BA STA

BOLETIM DE ANÁLISE ESTATÍSTICO

Resultados Finais de Analises Estatísticos utilizando Indicadores Sintéticos e Analíticos relacionados a Temas do GPS para Países da PUCSP, como parte do Projeto do Observatório da Rede Iberoamericana de Prospectiva - ORIBER

BASTA 2016 V2

ODS: PESSOAS
ANDRE NARDY

SAÚDE E MEIO AMBIENTE
CRISTINA ALCANTARA

MORADIA
RAFAEL CARDOSO

BEM ESTAR E QUALIDADE DE VIDA
TATIANA DUTRA

RESILIÊNCIA
VALERIA BERTONCELO

BIODIVERSIDADE - CAPITAL NATURAL E SOCIAL
MARCOS FERREIRA

INTEGRAÇÃO E ALIANÇAS ESTRATÉGICAS
FABIA SPADOTO

TRANSFORMAÇÃO PRODUTIVA E INOVAÇÃO SUSTENTÁVEL OVERNANÇA E EMPODERAMENTO DO CIDADÃO

LEDA DA SILVA REIS

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA

Eixos ODS: Pessoas

TRABALHO FINAL

Disciplina: **METODOS QUANTITATIVOS e QUALITATIVOS DA PESQUISA EMPIRICA**

Prof. Dr. Arnoldo José de Hoyos Guevara

DOUTORANDO –André Nardy (1º. Semestre de 2016)

Conteúdo

1. ANÁLISE EXPLORATÓRIA DE DADOS 7

a. INTRODUÇÃO 7

b. ENTENDENDO OS DADOS – BREVE APREENTAÇÃO DO PROJETO ORIBER 7

i. O tema de pesquisa: recursos básicos: agua, alimento, energia (basic features: water, food, energy) 7

A análise dos conceitos que compõem o tema foi extraída do material de aula Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP – Pontifícia Universidade Católica. Nesse estudo são apresentadas as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso. 7

Solução: *"GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."* 8

Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão "Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável." 8

c. O objeto de análise (indivíduos) 8

O objeto de análise deste trabalho constitui-se nos 132 países sobre os quais se dispõe de indicadores de referência internacional. Foram selecionadas dezessete variáveis, das bases de dados de 177 variáveis do GPS – Guia para o Desenvolvimento Sustentável. Descreve-se a seguir as variáveis analisadas. 8

d. As Variáveis do EIXO PESSOAS 8

e. ANÁLISE DAS VARIÁVEIS 11

i. Variáveis Quantitativas 11

ii. Estatística descritiva para as Variáveis selecionadas 12

f. ANÁLISE COMPARATIVA E CONSIDERAÇÕES FINAIS 12

i. IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO 12

ii. GOVERNANCE – Índice de Governança 13

iii. HPI – HAPPY PLANET INDEX 14

iv. GINNI Index - Índice de Progresso Social 14

v. SPI – Índice de Progresso Social 15

vi. ACCESS TO PIPED WATER – Acesso a agua encanada 16

vii. ACCESS TO IMPROVED SANITATION FACILITIES - Acesso a instalações sanitárias melhoradas 17

viii. SHELTER - Moradia 17

ix. STILLBIRTH RATE – Taxa de Não Nascidos 18

x. HOMICIDE RATE – Taxa de Homicídios 19

xi. INTERNET USERS – Usuários de Internet como % da população. 19

xii. INTERNATIONAL COOPERATION – Cooperação internatcional. 20

xiii. AGEWATCH INDEX – Idade e envelhecimento da população. 21

xiv. EPI – ENVIRONMENT PROTECTION INDEX – Proteção do meio ambiente 21
xv. DEMOCRACY INDEX – Democracia 22
xvi. NETWORK READINESS – Disponibilidade de rede 23
xvii. Efficiency of legal framework in settling disputes, 1-7 (best) – Eficiência do sistema legal em resolver disputas 23
g. Considerações Finais 24
2. ANÁLISE DE CORRELAÇÕES 24
a. CORRELAÇÕES 24
i. Análise da dispersão das principais correlações 27
ii. Análise da dispersão e Regressão das principais correlações 28
O Gráfico 3 apesenta as correlações com a linha de regressão resultante. 29
iii. Análise de similaridade entre as variáveis – emprego do gráfico de Dendograma 29
b. Considerações Finais 31
3. REGRESSÃO E STEPWISE 31
a. SELEÇÃO DAS VARIÁVEIS 32
b. REGRESSÃO – 1ª. RODADA SEM STEPWISE 33
i. INTERPRETAÇÃO DOS RESULTADOS INICIAIS 34
c. REGRESSÃO COM VARIÁVEIS SIGNIFICATIVAS 34
i. EMPREGO DA FUNÇÃO STEPWISE PARA IDENTIFICAR MELHORIAS NA REGRESSÃO 35
ii. EMPREGO DO AJUSTE DA AMOSTRA DA VARIÁVEL DEPENDENTE – "FIT" 36
d. NORMALIZAÇÃO DAS VARIÁVEIS 37
e. REGRESSÃO COM AS VARIÁVEIS NORMALIZADAS 37
f. REGRESSÃO EMPREGANDO A FUNÇÃO STEPWISE COM AS VARIÁVEIS NORMALIZADAS 39
4. ANALISE DE TENDENCIAS – EIXO PESSOAS E EIXO PLANETA - PESSOAS: Expectativa de Vida Ao Nascer e PIB per capita / PLANETA: Área Florestada (% total) e Emissões CO2 per capita – Comparativo entre Brasil, México, Portugal e Espanha. 41
a. EIXO PESSOAS 41
i. Breve descrição do EIXO PESSOAS. 41
Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão "Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável." 41
ii. Variáveis do EIXO PESSOAS 41
iii. A fonte e o tamanho da série de dados 42
b. ANÁLISE DAS VARIÁVEIS DO EIXO PESSOAS 42
i. Variável: Expectativa de Vida ao Nascer (em anos) 42
1. Análise de tendência e projeções - Brasil 42
2. Análise de tendência e projeções - México 45
3. Análise de tendência e projeções - Portugal 47
4. Análise de tendência e projeções - Portugal 50
ii. PIB per capita em USD constantes 52
1. Análise de Tendências e Projeções – Brasil 53
2. Análise de Tendências e Projeções – México 55
3. Análise de Tendências e Projeções – Portugal 57

4. Análise de Tendências e Projeções – Espanha ... 60

c. EIXO PLANETA ... 62

i. Breve descrição do EIXO PLANETA ... 62

Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão "Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável." ... 63

ii. Variáveis do EIXO PLANETA ... 63

iii. A fonte e o tamanho da série de dados ... 63

d. ANÁLISE DAS VARIÁVEIS DO EIXO PLANETA ... 64

i. Variável: Área Florestada como percentual da área total do país (%) ... 64

1. Análise de tendência e projeções - Brasil ... 64

2. Análise de tendência e projeções – México ... 66

3. Análise de tendência e projeções – Portugal ... 69

4. Análise de tendência e projeções – Espanha ... 72

e. Variáveis e fontes de dados ... 74

i. Variável Emissões de CO2 per capita. ... 74

1. Análise de tendência e projeções – Brasil ... 74

2. Análise de tendência e projeções – México ... 77

3. Análise de tendência e projeções – Portugal ... 79

4. Análise de tendência e projeções – Espanha ... 82

5. ANOVA E SIMILARIDADE ... 85

a. ANOVA DAS VARIÁVEIS DO EIXTO PESSOAS ... 88

b. ANÁLISE CONSOLIDADA DOS TESTES ANOVA: VARIÁVEIS DO EIXO PESSOAS VS REGIÕES ... 122

c. ANÁLISE ATRAVÉS DO GRÁFICO DE RADAR DAS VARIÁVEIS SINTÉTICAS E ANALÍTICAS DO EIXO PESSOAS VS REGIÕES ... 123

i. GRÁFICO DE RADAR DAS VARIÁVEIS SINTÉTICAS - MÉDIAS E F - EIXO PESSOAS VS REGIÕES ... 123

ii. GRÁFICO DE RADAR DAS VARIÁVEIS ANALÍTICAS - MÉDIAS E F - EIXO PESSOAS VS REGIÕES ... 124

6. AMOSTRAS ALEATÓRIAS DE VARIÁVEIS SELECIONADAS DO EIXO PESSOAS ... 125

a. ANÁLISE DAS AMOSTRAS ALEATÓRIAS PARA AS VARIÁVEIS ANALISADAS ... 126

i. VARIÁVEL IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO ... 126

1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES. ... 126

2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES ... 127

ii. VARIÁVEL SP I – ÍNDICE DE PROGRESSO SOCIAL ... 128

1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES. ... 128

2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES ... 130

iii. VARIÁVEL ÍNDICE DE GOVERNANÇA ... 131

1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES. ... 131

2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES ... 132

iv. COMPARAÇÃO DAS VARIÁVEIS POR REGIÃO – AIBER E AVECO ... 134

1. HDI – HUMAN DEVELOPMENT INDEX / IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO 134
2. SPI – SOCIAL PROGRESS INDEX / IPS – ÍNDICE DE PROGRESSO SOCIAL 135
3. GOVERNANCE INDEX – ÍNDICE DE GOVERNANÇA 136
b. CONCLUSÃO 137
7. ANÁLISE DE COMPONENTES PRINCIPAIS 137
a. ANÁLISE DE COMPONENTES PRINCIPAIS 138
i. ANÁLISE DAS VARIÁVEIS DA AMOSTRA – COMPONENTES PRINCIPAIS 138
ii. ANÁLISE DOS COMPONENTES PRINCIPAIS GERADOS 142
1. ANÁLISE COMPONENTE PRINCIPAL 1 142
2. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS 143
iii. ANÁLISE COMPONENTE PRINCIPAL 2 144
iv. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS 145
v. ANÁLISE COMPONENTE PRINCIPAL 3 146
vi. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS 147
vii. ANÁLISE COMPONENTE PRINCIPAL 4 148
viii. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS 149
b. CONSIDERAÇÕES FINAIS 150
8. AGRUPAMENTO DE OBSERVAÇÕES 151
a. NOVAS VARIÁVEIS EXPLICATIVAS DO EIXO PESSOAS - COMPONENTES PRINCIPAIS 151
b. CLUSTERS DE OBSERVAÇÕES COM AS VARIÁVEIS CHAVE DOS COMPONENTES PRINCIPAIS 152
i. EXERCICIO COM 2 CLUSTERS 152
ii. EXERCICIO COM 20 CLUSTERS 155
iii. EXERCICIO COM 3 CLUSTERS – COMPONENTES PRINCIPAIS 160
iv. EXERCICIO COM 20 CLUSTERS – COMPONENTES PRINCIPAIS 164
v. EXERCICIO COM 30 CLUSTERS – COMPONENTES PRINCIPAIS 168
c. GRÁFICO 3D DO EXERCICIO COM 20 CLUSTERS – COMPONENTES PRINCIPAIS 174
d. ANÁLISE DOS AGRUPAMENTOS POR REGIÃO – AIBER E AVECO 175
i. AIBER – VARIÁVEIS RELEVANTES E 20 CLUSTERS 175
ii. AIBER – VARIÁVEIS RELEVANTES E 8 CLUSTERS 178
iii. AIBER – COMPONENTES PRINCIPAIS 8 CLUSTERS 181
iv. AIBER – COMPONENTES PRINCIPAIS E 5 CLUSTERS 182
v. AIBER – ANÁLISE GRÁFICA COMPONENTES PRINCIPAIS E 5 CLUSTERS 184
e. AVECO – VARIÁVEIS RELEVANTES E 20 CLUSTERS 185
f. AVECO – VARIÁVEIS RELEVANTES E 8 CLUSTERS 189
g. AVECO – COMPONENTES PRINCIPAIS E 8 CLUSTERS 192
h. AVECO – COMPONENTES PRINCIPAIS E 5 CLUSTERS 194
i. AVECO – ANÁLISE GRÁFICA COMPONENTES PRINCIPAIS E 5 CLUSTERS 195
j. CONSIDERAÇÕES FINAIS 197

9. ANÁLISE DISCRIMINANTE ... 197

a. AGRUPAMENTO DAS OBSERVAÇÕES COM AS QUATRO COMPONENTES PRINCIPAIS EMPREGANDO EM CLUSTERS ATRAVÉS DO MÉTODO DE LIGAÇÃO COMPLETA E DISTANCIA POR MANHATTAN - AIBER, AVECO E OTHERS ... 197

i. OS 4 GRUPOS RESULTANTES ... 200

b. ANÁLISE DISCRIMINANTE ... 202

i. TESTE DA FUNÇÃO LINEAR ... 203

ii. TESTE DA FUNÇÃO QUADRÁTICA ... 205

c. CONCLUSÃO E CONSIDERAÇÕES FINAIS ... 208

10. REGRESSÃO LOGÍSTICA ... 208

a. AVALIAÇÃO E SELEÇÃO DO TIPO DE REGRESSÃO QUE MELHOR REPRESENTA OS GRUPOS DE PAÍSES ORGANIZADOS A PARTIR DAS QUATRO COMPONENTES PRINCIPAIS. ... 208

b. REGRESSÃO LOGÍSITCA ... 212

i. TESTE DA FUNÇÃO LOGÍSTICA ORDINAL PARA OS QUATRO AGRUPAMENTOS ... 212

ii. TESTE DA FUNÇÃO LOGÍSITCA NOMINAL PARA OS QUATRO AGRUPAMENTOS ... 214

iii. RESULTADOS DOS TESTE COM FUNÇÕES LOGÍSTICAS PARA OS QUATRO AGRUPAMENTOS DE PAÍSES. ... 215

c. TESTE DA FUNÇÕES LOGÍSTICA PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS ... 215

i. TESTE DA FUNÇÃO LOGÍSTICA ORDINAL PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS ... 216

ii. TESTE DA FUNÇÃO LOGÍSTICA NOMINAL PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS ... 217

iii. RESULTADOS DOS TESTE COM FUNÇÕES LOGÍSTICAS PARA O PERTENCIENTO A AIBER, AVECO E OTHERS. ... 218

d. CONCLUSÃO E CONSIDERAÇÕES FINAIS ... 218

11. ÁRVORE DE CLASSIFICAÇÃO ... 219

a. Variáveis chave dos Componentes Principais ... 219

b. APLICAÇÃO DA ANÁLISE DISCRIMINANTE COM SEPARAÇÃO DE CLUSTERS – MÉTODO MANHATAN ... 219

c. ANÁLISE DISCRIMINANTE COM OS GRUPOS AIBER, AVECO E OTHERS. ... 221

i. a. 84,8% of original grouped cases correctly classified. ... 222

ii. ... 222

d. EMPREGO DA ÁRVORE DE DECISÃO COMO FERRAMENTA DE CLASSIFICAÇÃO. 222

i. 4 CLUSTERS RESULTADOS DOS TESTES COM ÁRVORE DE DECISÃO ... 222

1. MÉTODO CHAID ... 222

2. MÉTODO EXHAUSTIVE CHAID ... 223

3. MÉTODO CRT ... 224

4. MÉTODO EXHAUSTIVE QUEST ... 227

ii. 3 GRUPOS DE PAÍSES: ABER, AVECO E OTHERS - RESULTADOS DOS TESTES COM ÁRVORE DE DECISÃO ... 229

1. MÉTODO CHAID ... 229

2. MÉTODO EXHAUSTIVE CHAID ... 230

3. MÉTODO CRT ... 231

4. MÉTODO QUEST ... 232

e. CONCLUSÕES E CONSIDERAÇÕES FINAIS 233
12. ANÁLISE DE CORRESPONDÊNCIA 233
a. A ANÁLISE DE CORRESPONDÊNCIA 233
i. CONSTRUÇÃO DA BASE DE ANÁLISE DOS GRUPOS 233
b. ANÁLISE DE CORRESPONDÊNCIA NO MINITAB ® 16 235
c. CONCLUSÃO 238

1. ANÁLISE EXPLORATÓRIA DE DADOS

a. INTRODUÇÃO

Este trabalho tem como objetivo realizar uma análise exploratória dos dados das variáveis do EIXO PESSOAS pra 132 países, comparados a partir de dezessete variáveis quantitativas, seis das quais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Empregou-se o Minitab ® para as análise estatísticas.

b. ENTENDENDO OS DADOS – BREVE APREENTAÇÃO DO PROJETO ORIBER

i. O tema de pesquisa: recursos básicos: agua, alimento, energia (basic features: water, food, energy)

A análise dos conceitos que compõem o tema foi extraída do material de aula Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica. Nesse estudo são apresentadas as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso.
Vale destacar o desafio citado no GPS e sua proposta de solução:
Desafio: *"A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?"*

Solução: *"GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."*

Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão "Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável."

c. O objeto de análise (indivíduos)

O objeto de análise deste trabalho constitui-se nos 132 países sobre os quais se dispõe de indicadores de referência internacional. Foram selecionadas dezessete variáveis, das bases de dados de 177 variáveis do GPS – Guia para o Desenvolvimento Sustentável. Descreve-se a seguir as variáveis analisadas.

d. As Variáveis do EIXO PESSOAS

Este tema da pesquisa apresenta dezessete variáveis, sendo oito índices sintéticos. Todas as variáveis são quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis componentes ou indicadores, a saber:

- **Índices sintéticos:** Índice de Progresso Social (SPI), o Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD – ONU, Índice de Governança, o Índice de Proteção Ambiental – EPI (*Environmental Protection Index*), o *Happy Planet Index* (HPI),Índice de Felicidade Planetária (HPI), Índice de Democracia, , Índice de observação etária (Age Watch Index), e o Índice de Ginni para os países selecionados.

- **Indicadores ou variáveis componentes:** Habitação (shelter), Taxa de não nascidos (stillbirth rate), acesso a água potável, acesso a saneamento básico melhorado, taxa de homicídios, percentual de usuários de internet sobre a população, cooperação internacional, disponibilidade de rede, eficiência das estruturaras legais em resolver disputas.

Quadro 1. As **17 Variáveis do EIXO PESSOAS**

Índice	Significado	Fonte	Tipo	Unidade de medida
IDH (atualizado para 2014)	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação	PNUD	Quantitativa	%

1. Governança	Governança eficácia de governo - consiste nas tradições e instituições pelas quais a autoridade de um país é exercido. Isso inclui o processo pelo qual os governos são selecionados, monitorados e substituídos; a capacidade do governo para formular e implementar com eficácia políticas sólidas; eo respeito dos cidadãos e do Estado pelas instituições que governam as interações econômicas e sociais	World Bank - World Governance Index	Quantitativa	%
Happy Planet Index	O Happy Planet Index (HPI) é a medida líder global de bem-estar sustentável.	happyplanetindex.org	Quantitativa	%
Índice Ginni	Indice de Gini mede a extensão em que a distribuição de renda (ou, em alguns casos, a despesa de consumo) entre indivíduos ou famílias dentro de uma economia desvia de uma distribuição perfeitamente igual. Um índice de Gini de 0 representa a igualdade perfeita, enquanto um índice de 100 indica desigualdade perfeita.	World Bank Development Research Group	Quantitativa	Indicador nominal de 0 a 100
SPI - Índice de Progresso Social	O SPI mede progresso social como "a capacidade de uma sociedade para satisfazer as necessidades humanas básicas dos seus cidadãos, estabelecer os blocos de construção que permitem que os cidadãos e as comunidades melhorem, manter a qualidade de suas vidas, e criar as condições para todos os indivíduos atingirem seu pleno potencial. O SPI mede uma ampla gama de componentes do desempenho social e ambiental e agrega-os num quadro geral." Fonte: Social progress imperative.	Social Progress Imperatives empregando dados da World Health Organization (WHO)-UNICEF	Quantitativa	%
AgeWatch Index	Índice de HelpAge International AgeWatch Mundial classifica os países por quão bem conduzem o processo de envelhecimento da população. Mensura quatro domínios que são elementos fundamentais do bem-estar das pessoas idosas: renda, saúde, capacidade e ambiente favorável.	HelpAge International	Quantitativa	%
EPI – Environmental Protective Index	O Índice de Desempenho Ambiental (EPI) classifica o desempenho dos países em questões ambientais de alta prioridade em duas grandes áreas políticas: proteção da saúde humana do mal e proteção dos ecossistemas ambiental.	Yale University	Quantitativa	%
Democracy Index	O Índice de Democracia é baseado em cinco categorias: processo eleitoral e pluralismo; liberdades civis; o funcionamento do governo; participação política; e cultura política. Com base nas suas pontuações em uma série de indicadores dentro dessas categorias, cada país é então classificado como um dos quatro tipos de regime: "democracias	The Economist Intelligence Unit	Quantitativo	%

	plenas"; "democracias imperfeitas"; "regimes híbridos"; e "regimes autoritários".			
Stillbirth rate (deaths/1,000 live births)	Mensura a taxa de não nascidos por milhar de habitantes de uma população nacional segundo a premissa de mortes no terceiro trimestre fetal (> 1000 gramas ou> 28 semanas), por 1.000 nascidos vivos .	World Health Organization	Quantitativo	Índice
Shelter	Composição de quatro indicadores de moradia, acesso a eletricidade e poluição.	Diversos	Quantitativo	%
Acesso a água canalizada	A percentagem da população com canalização de serviço de água conectada com canalização interna para um ou mais torneiras ou uma conexão de água canalizada a uma torneira colocada no quintal ou terreno fora da casa.	WHO-UNICEF	Quantitativa	%
Acesso a instalações sanitárias melhoradas (atualizado para 2015)	A percentagem da população com saneamento melhorado, incluindo autoclismos, sistemas de tubulações de esgoto, fossas sépticas, lavar / derramar rente à latrina, latrinas ventiladas melhoradas (VIP), latrina com laje, e banheiros de compostagem	WHO-UNICEF	Quantitativa	%
Taxa de Homicídios (1= <2/100,000; 5= >20/100,000	Número de homicídios, definida como a morte deliberadamente causados a uma pessoa por outra pessoa, por 100.000 pessoas. Pontuado em uma escala 1-5: 1 = 0-1,99 2 = 2-5,99 3 = 6-9,99 4 = 10-19,99 5 => 20	Institute for Economics and Peace Global Peace Index	Quantitativa	Escalar
Internet users	Percentual de usuários de internet sobre o total da população, baseado no acesso através de qualquer equipamento nos últimos 12 meses	International Telecommunications Union	Quantitativa	%

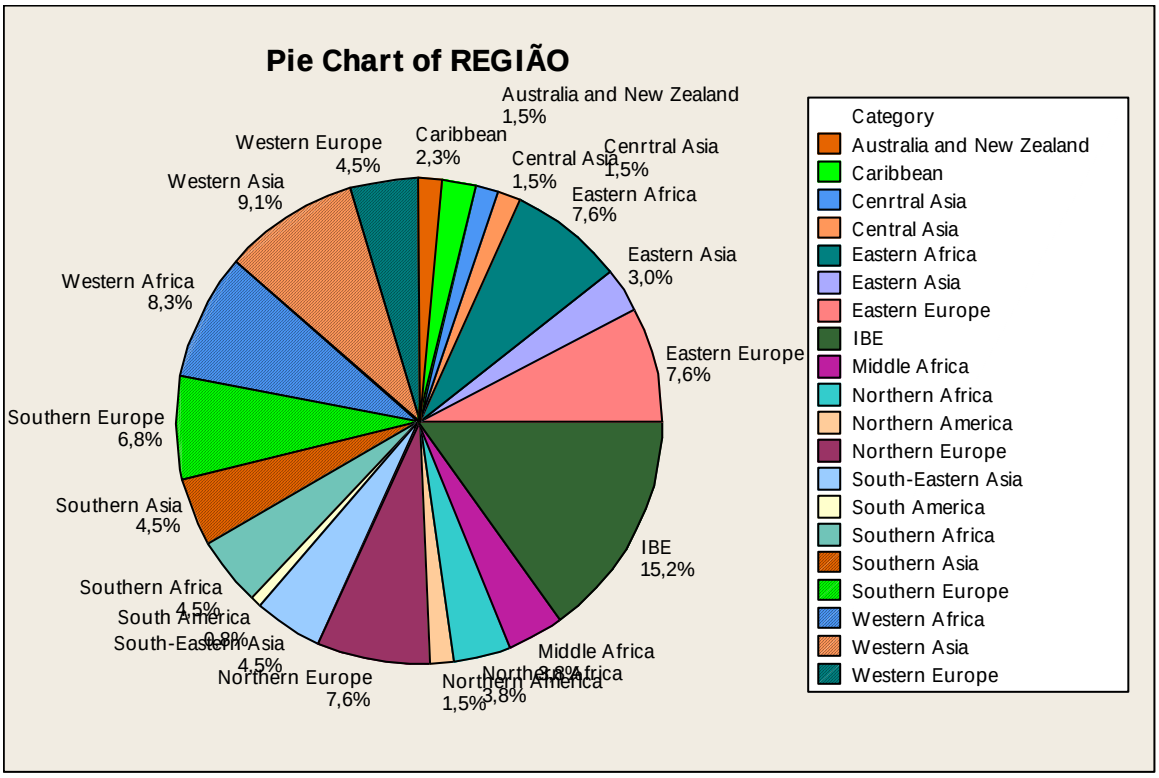
Network readiness	As medidas Networked Readiness Index, estabelecem uma escala de 1 (pior) a 7 (melhor) para o desempenho de 148 economias em alavancar tecnologias de informação e comunicação para aumentar a competitividade e bem-estar.	Global information technology report	Quantitativo	escalar
Efficiency of legal framework in settling disputes, 1-7 (best)	Mensura quão eficiente é o quadro jurídico para as empresas privadas na resolução de litígios: [1 = extremamente ineficiente; 7 = extremamente eficiente	World Economic Forum	Quantitativo	Escalar

Fonte: preparado pelo autor a partir dos dados da planilha de dados e das fontes consultadas

e. ANÁLISE DAS VARIÁVEIS

Apresentamos o gráfico circular ou gráfico de pizza ou gráfico de torta elaborado pelo MINITAB com as regiões dos países selecionados**.** Verificamos que pela classificação adotada 15,2% dos 132 países são definidos como mundo ibérico (IBE), subclassificação com maior número de países dentre as regiões.

Gráfico 1: Região dos Países Selecionados



i. Variáveis Quantitativas

As variáveis quantitativas são assim denominadas pois apresentam-se em escalas numéricas.

ii. Estatística descritiva para as Variáveis selecionadas

Empregou-se o Minitab®, função *statistic descriptives* para apresentação das variáveis. Os títulos das variáveis são apresentados no seu formato original de captura, em inglês.

Descriptive Statistics: 1. Governanç; EV - Water R; Happy Planet; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
1. Governança	132	0	53,38	1,88	21,59	15,22	37,11
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780
Social Progress Index	129	3	64,84	1,19	13,53	31,42	54,64
Access to piped water	129	3	64,68	3,02	34,34	1,58	25,38
Access to improved sanit	130	2	73,51	2,59	29,56	9,05	53,04
Q17 | International Coop	104	28	6,923	0,163	1,658	2,333	6,000
2016 EPI Score	132	0	69,87	1,26	14,42	37,10	61,10
Shelter	132	0	60,85	1,97	22,58	6,38	41,65
Stillbirth rate (deaths/	131	1	12,878	0,850	9,733	2,000	4,000
Homicide rate (1= <2/100	132	0	2,682	0,131	1,510	1,000	1,000
Internet users (% of pop	132	0	41,96	2,50	28,75	1,22	14,82
Human Development Index	131	1	0,7043	0,0137	0,1566	0,3483	0,5791
AgeWatch Index 2014	90	42	51,70	2,14	20,35	4,10	37,90
Demcracy Index 2014	132	0	5,873	0,180	2,068	1,490	4,100
The Networked Readiness	121	11	4,0189	0,0798	0,8783	2,2200	3,3500
Efficiency of legal fram	118	14	3,7298	0,0858	0,9316	1,5453	3,1724

Variable	Median	Q3	Maximum
1. Governança	48,46	70,36	100,00
Happy Planet Index	42,463	49,286	64,036
GINNI Index	38,160	44,550	63,140
Social Progress Index	64,98	74,08	88,36
Access to piped water	77,72	96,33	100,00
Access to improved sanit	88,04	98,22	100,00
Q17 | International Coop	7,000	8,000	10,000
2016 EPI Score	71,69	81,92	90,68
Shelter	65,95	78,81	91,81
Stillbirth rate (deaths/	10,000	21,000	47,000
Homicide rate (1= <2/100	2,000	4,000	5,000
Internet users (% of pop	41,87	64,75	96,21
Human Development Index	0,7330	0,8322	0,9439
AgeWatch Index 2014	51,25	65,08	93,40
Demcracy Index 2014	6,045	7,563	9,930
The Networked Readiness	3,9500	4,5800	6,0400
Efficiency of legal fram	3,6286	4,2840	6,0327

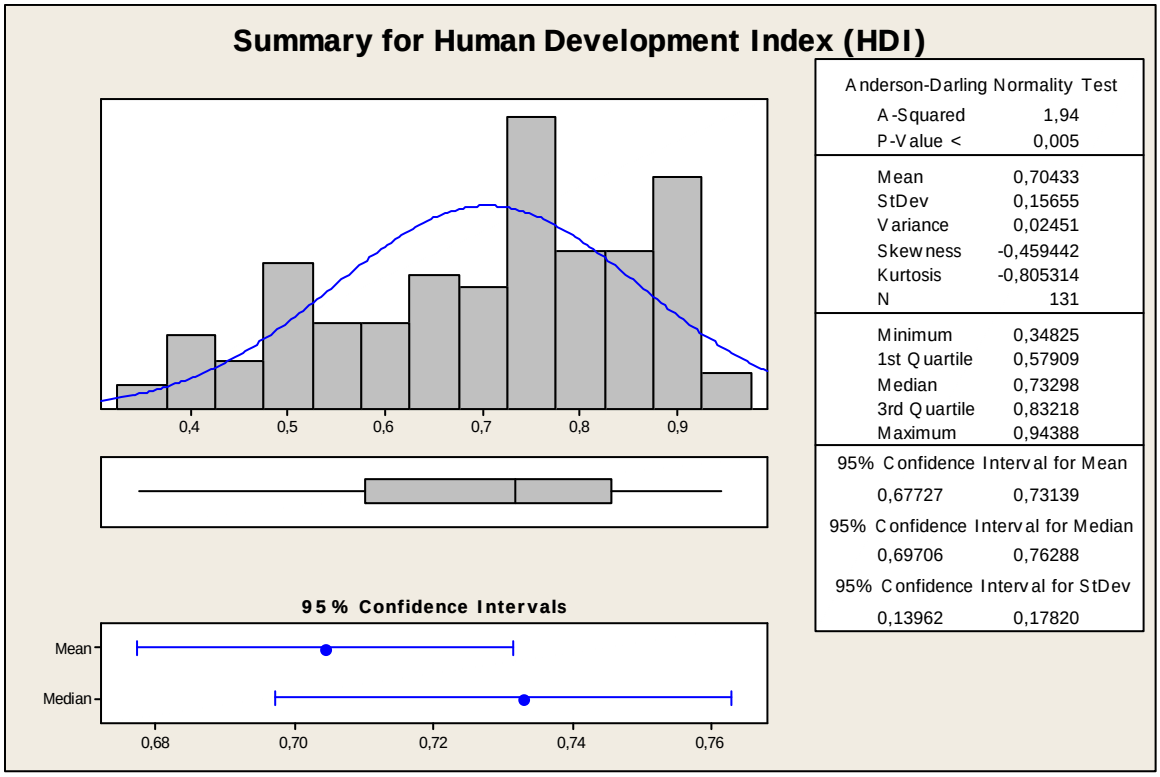
Legenda: N: observações válidas; N*: observações em branco; Mean: média; SE mean: erro padrão; STDev: desvio padrão da amostra; Minimum: valor mínimo observado; Q1: valor limite do primeiro quartil, ou 25% das observações; Median: mediana; Q3: valor limite do terceiro quartil ou 75% das observações; Maximun: valor máximo observado.

Fonte: preparado pelo autor.

f. ANÁLISE COMPARATIVA E CONSIDERAÇÕES FINAIS

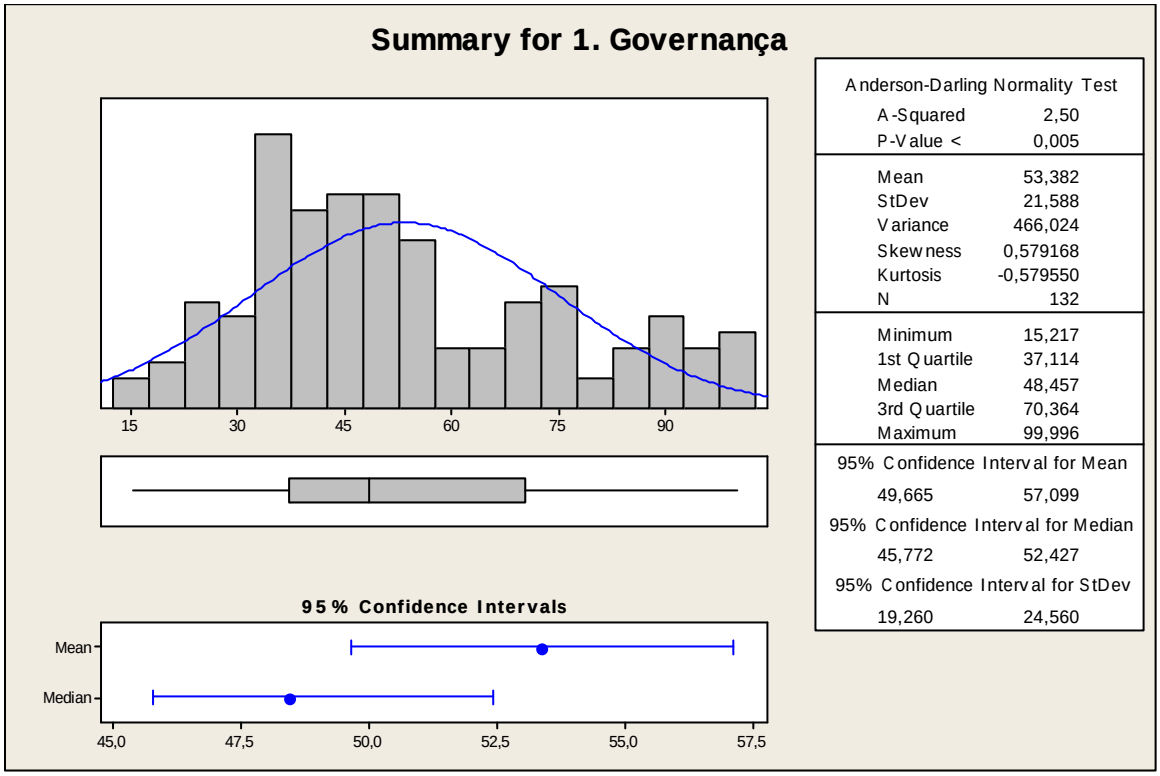
Analisa-se cada uma das 16 variáveis quantitativas com a função *basic statistics / graphical summary* do Minitab ®. Empregou-se grau de confiança de 95%.

i. IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO



Observa-se que o IDH possui distribuição deslocada para a direita, maiores observações, com uma mediana de 0,733, e média de 0,704, para 131 observações. Desta forma é possível afirmar que há mais países com melhores IDHs do que o contrário. No entanto não é possível afirmar que haja mais pessoas vivendo em nestes países. Não se encontram *outliers*.

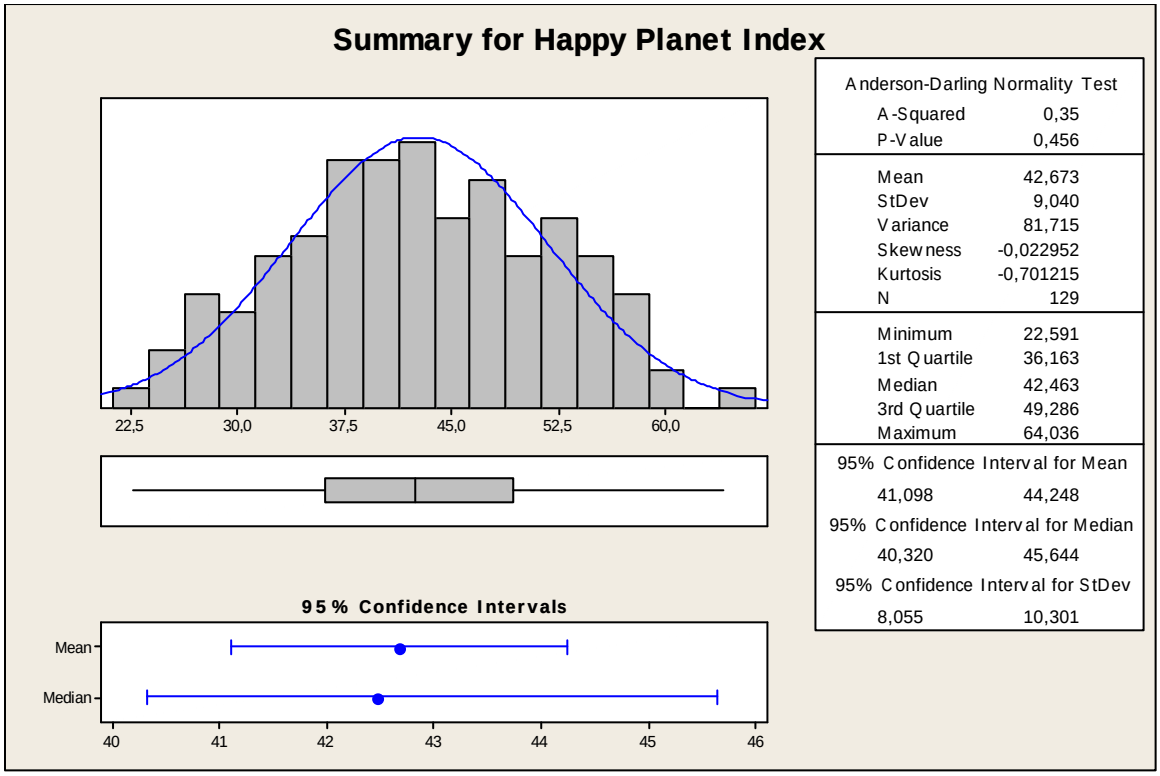
ii. GOVERNANCE – Índice de Governança



O índice de governança do banco mundial é uma composição de cinco vetores, Voz e Accountability, Estabilidade Política, Qualidade Regulatória, Força da Lei e Controle da Corrupção. Verifica-se que a

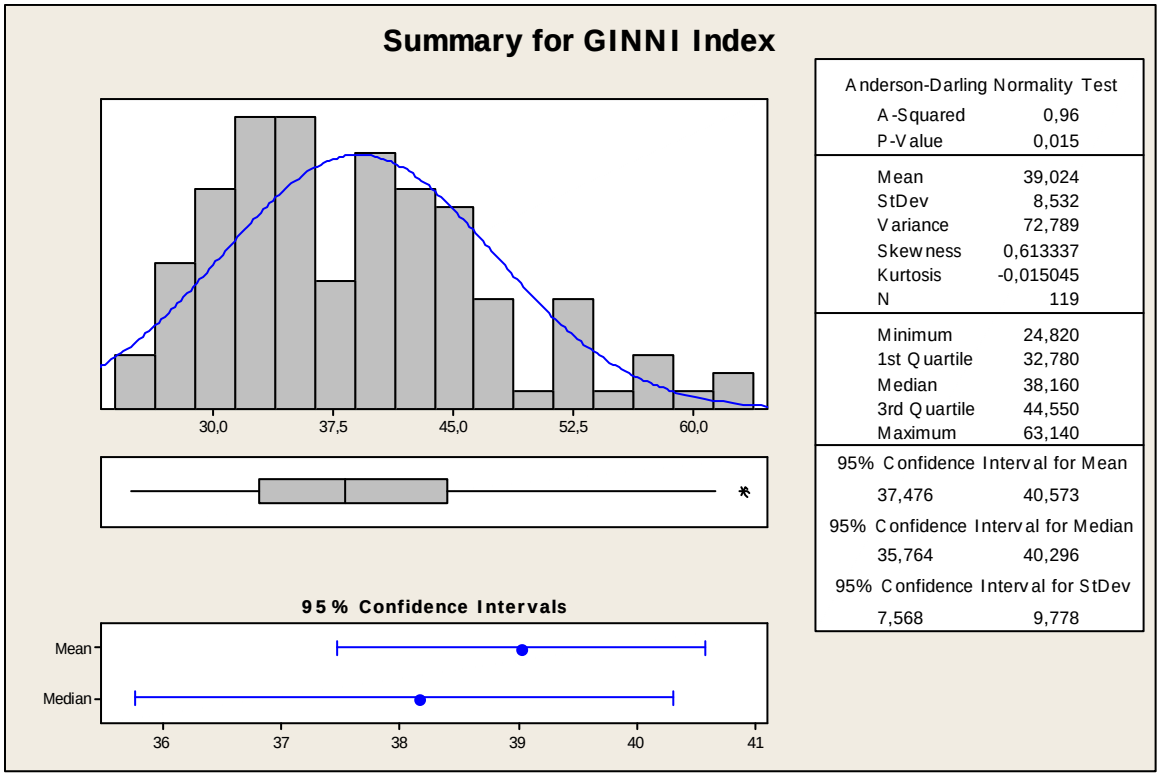
curva das observações é mais concentrada à esquerda, ou seja, com menores índices, há portanto mais observações (países) com piores governanças do que o contrario. A média é de 53,382 e a mediana da amostra é de 48,457. Todos os 132 países estudados possuem valores válidos.

iii. HPI – HAPPY PLANET INDEX



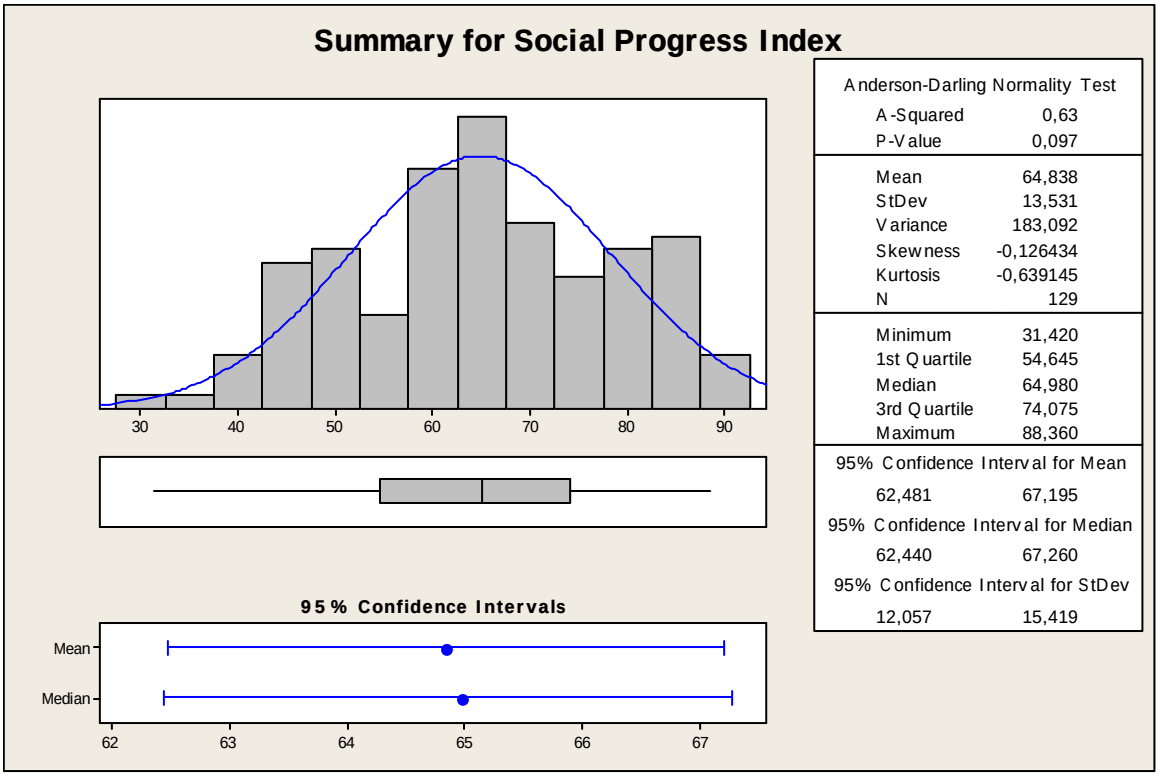
O Indicador de Felicidade Planetária ou HPI apresenta uma distribuição normal, com média de 42,67%, mediana de 42,46%, e desvio padrão de 9,04%. Não há *outliers*. A pior observação é encontrada no Lesotho com 23% e Chad com 24% a melhor na Costa Rica com 64% e Colômbia com 59%. Os oito melhores índices são encontrados na América Latina, o que não deixa de ser interessante.

iv. GINNI Index - Índice de Progresso Social



O índice de GINNI é um índice composto, como a maioria dos índices analisados até o memento, a exceção do índice de águas tratadas. Entretanto, diferentemente dos demais, o Indice de GINNI menor indica um maior desenvolvimento social, apresentando uma menor distância entre as populações de maior e menor renda. A pior distribuição de renda é também o único *outlier* da amostra, Africa do Sul com GINNI de 63,14.

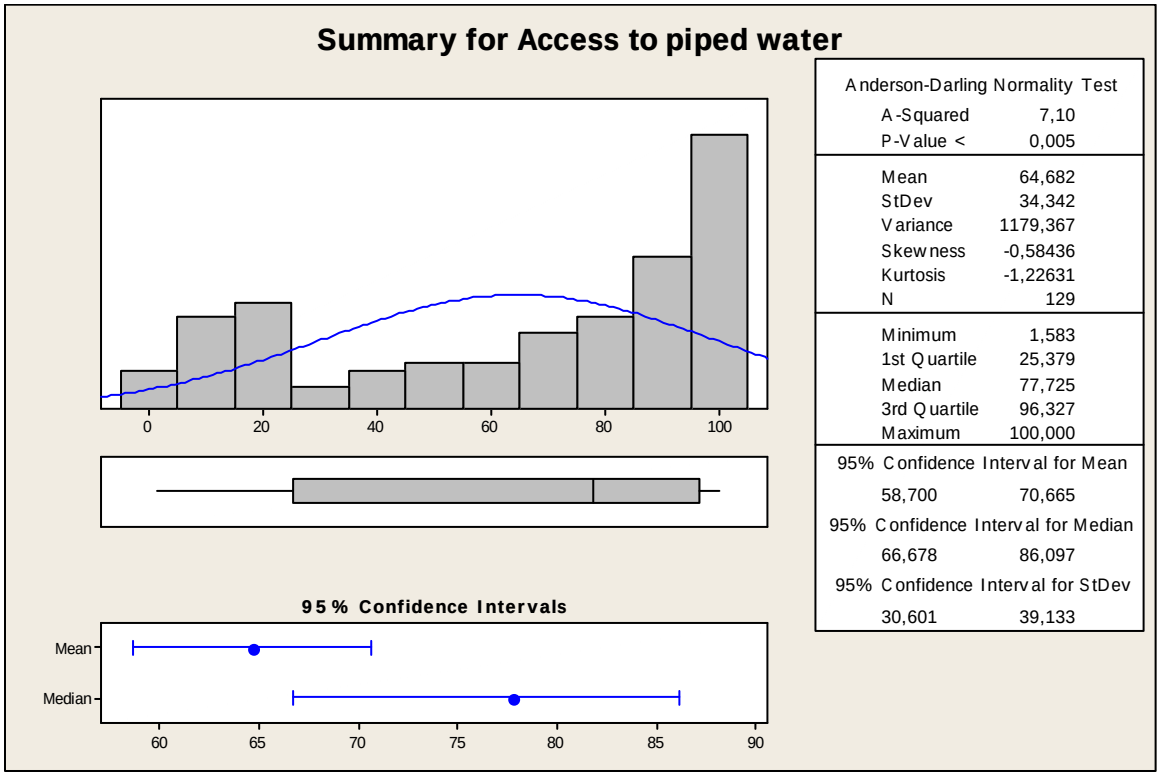
v. SPI – Índice de Progresso Social



Análise:
Variável: Social Progress Index – índice de Progresso Social
Origem e definição dado pelo portal: http://pt.knoema.com/jklbfre/the-social-progress-index
O Índice de Progresso Social é o resultado de um processo de dois anos de pesquisas envolvendo uma equipe renomada de especialistas que inclui, entre outros, os economistas Hernando de Soto e Michael Porter e do presidente da Fundação Rockefeller Dr. Judith Rodin. O índice sintetiza um total extenso de dados pesquisados para identificar as dimensões do desempenho das sociedades e medir o progresso social de forma abrangente e rigorosa. O Índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidadas em três dimensões do Progresso Social: necessidades humanas básicas, Fundações de Bem-estar e oportunidade. A primeira dimensão, necessidades humanas básicas, avalia o quão bem um país prevê necessidades essenciais de seu povo através da medição se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se tiverem acesso a água potável, se tiverem acesso a uma habitação condigna com serviços básicos, e se eles são seguros e protegidos.

Análise dos dados de 2015 indica que a distribuição aproxima-se da normal, para as 129 observações consideradas. A média das observações é 64,8% e a mediana 64,98%. Com grande concentração das observações dentro de um desvio padrão, de 13,53%. Mesmo com a normalidade da amostra observa-se uma maior concentração das observações à direita, com maiores índices de progresso social, acima de 85%.

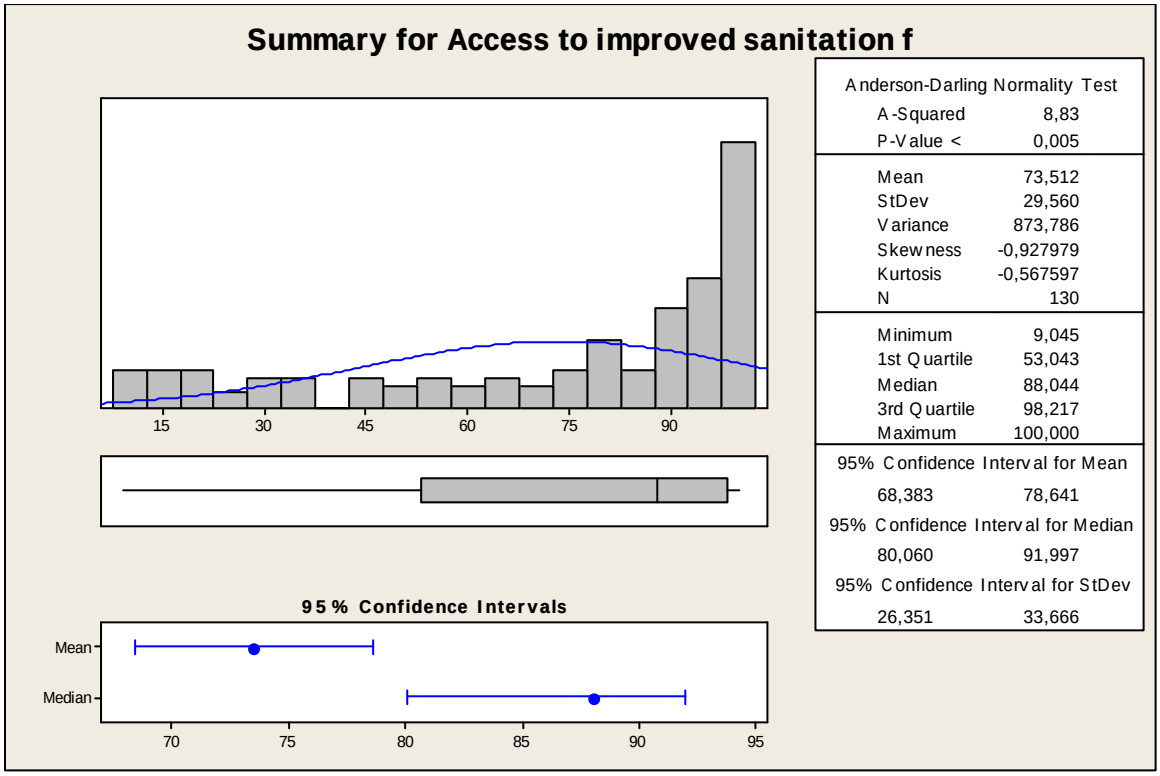
vi. ACCESS TO PIPED WATER – Acesso a agua encanada



Anderson-Darling Normality Test	
A-Squared	7,10
P-Value <	0,005
Mean	64,682
StDev	34,342
Variance	1179,367
Skewness	-0,58436
Kurtosis	-1,22631
N	129
Minimum	1,583
1st Quartile	25,379
Median	77,725
3rd Quartile	96,327
Maximum	100,000
95% Confidence Interval for Mean	
58,700	70,665
95% Confidence Interval for Median	
66,678	86,097
95% Confidence Interval for StDev	
30,601	39,133

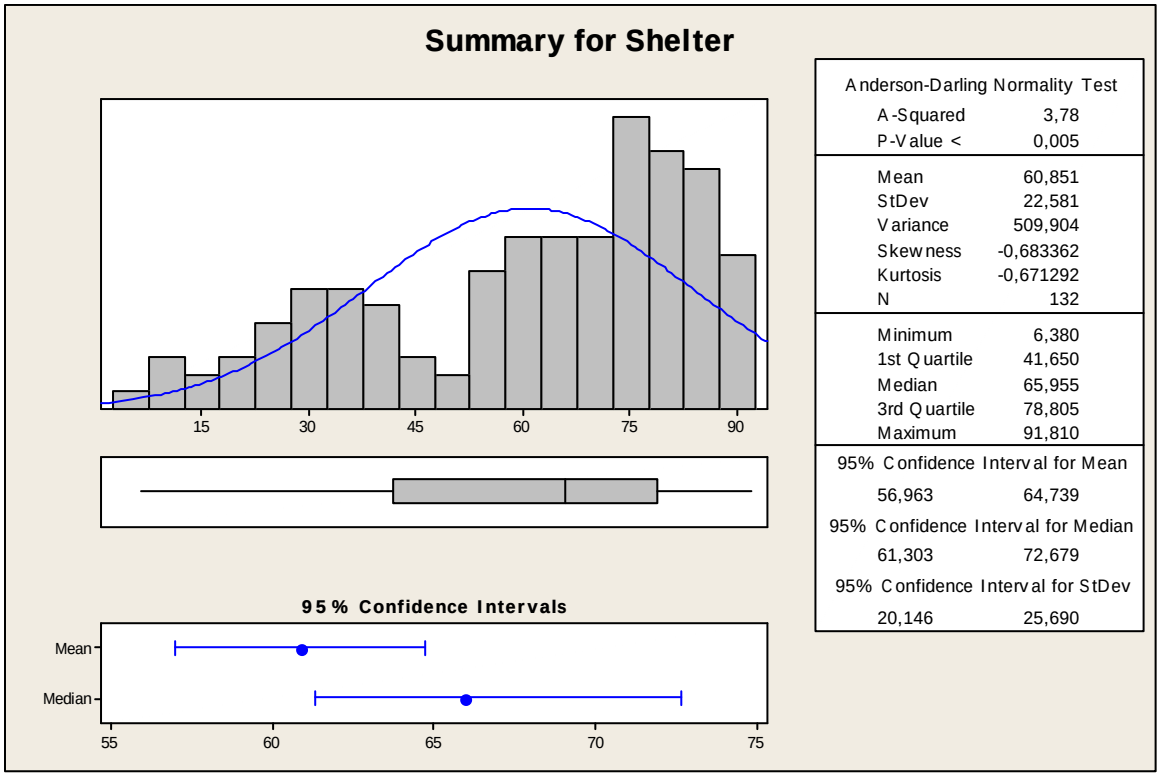
O índice de acesso a agua potável não apresenta distribuição normal nem tampouco *outliers*. Há 129 observações para este índice. A media é de 64,7%, mediana de 77,7% e desvio padrão de 34,3%. A amodtra-se concentra sua distribuição à direit ado gráfico, com grande quantidade de países com 90% ou mais de acesso a água potável.

vii. ACCESS TO IMPROVED SANITATION FACILITIES - Acesso a instalações sanitárias melhoradas



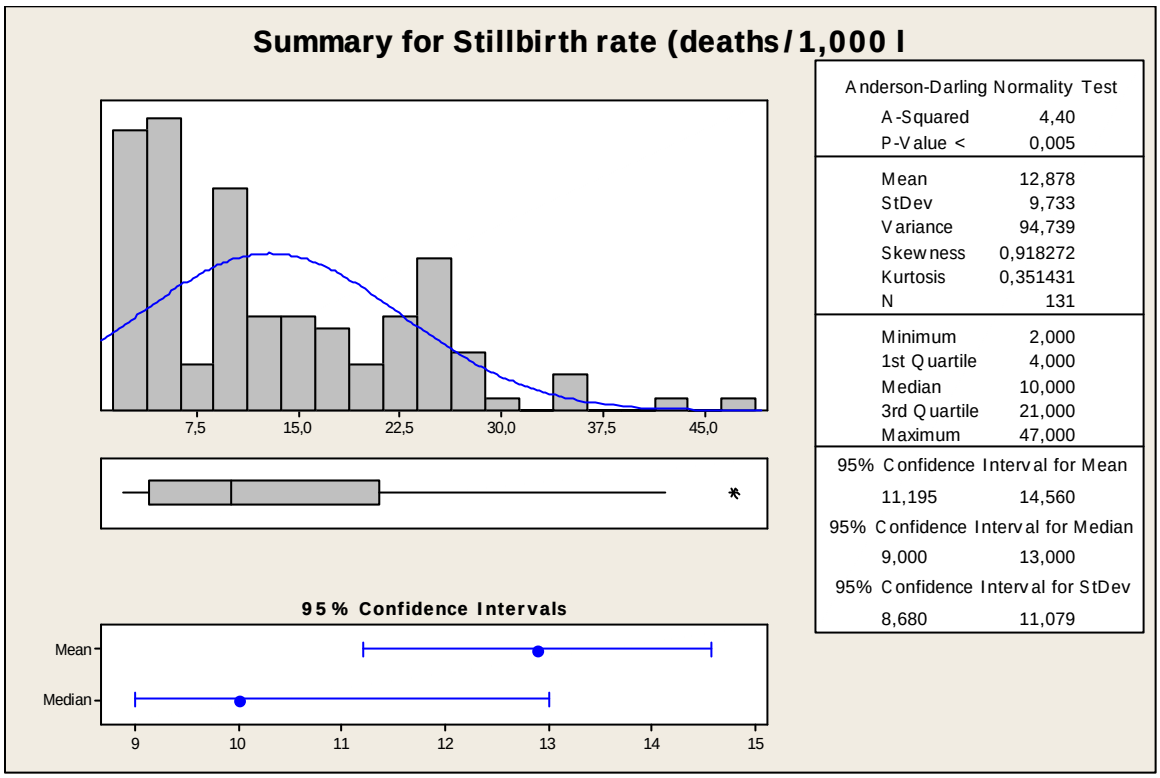
O índice de acesso a condições sanitárias melhoradas mede o percentual da população com a cesso a recursos sanitários. Quanto maior o índice, maior é o percentual da população com a cesso a estes recursos básicos. Verifica-se nesta amostra de 130 países que a amostra não é normalmente distribuída, com concentração À direita do gráfico, em especial nos valores acima de 90%, como indica o valor limite do terceiro quartile, 98,217%. A média da amostra é 73,5 e a mediana 88,0, com desvio padrão de 29,56.

viii. SHELTER - Moradia



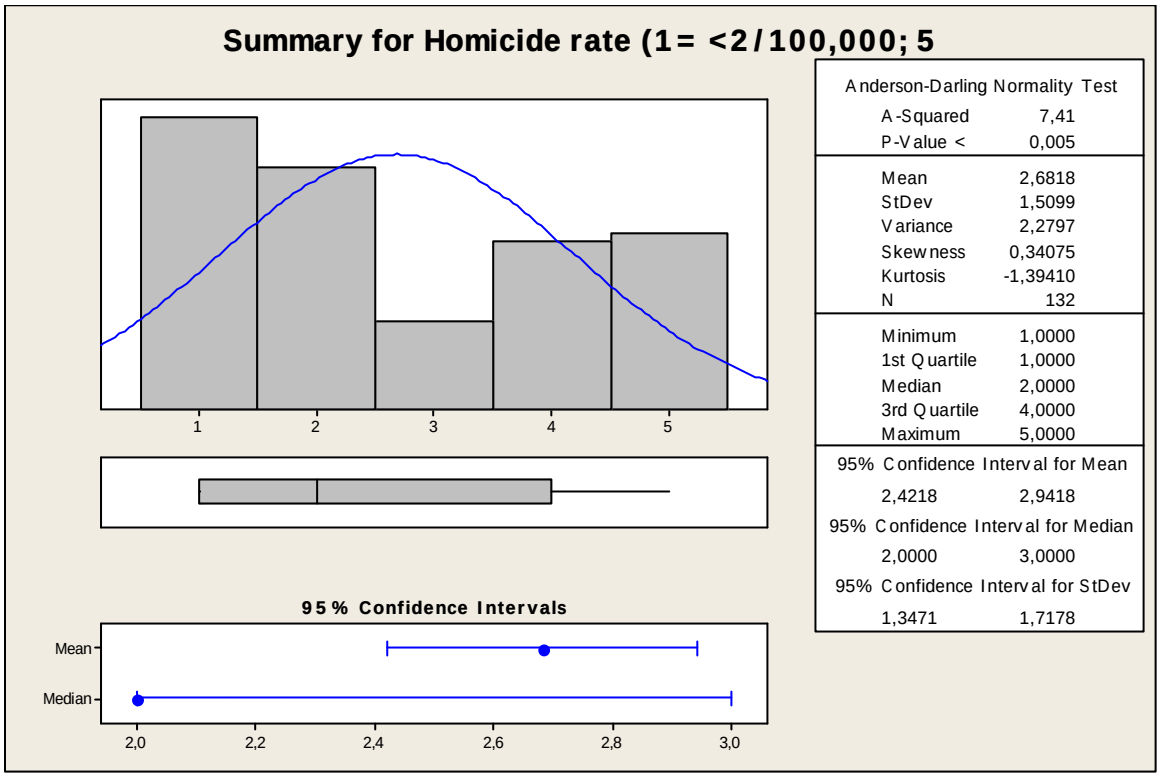
O índice de moradia indica melhores possibilidades de moradia em determinado país. Quanto maior o índice, melhores são as condições encontradas. Verifica-se nesta amostra de 132 países que a amostra não é normalmente distribuída, com concentração à direita do gráfico, em especial nos valores acima de 60%. A média da amostra é 60,85 e a mediana 65,96 com desvio padrão de 22,58.

ix. STILLBIRTH RATE – Taxa de Não Nascidos



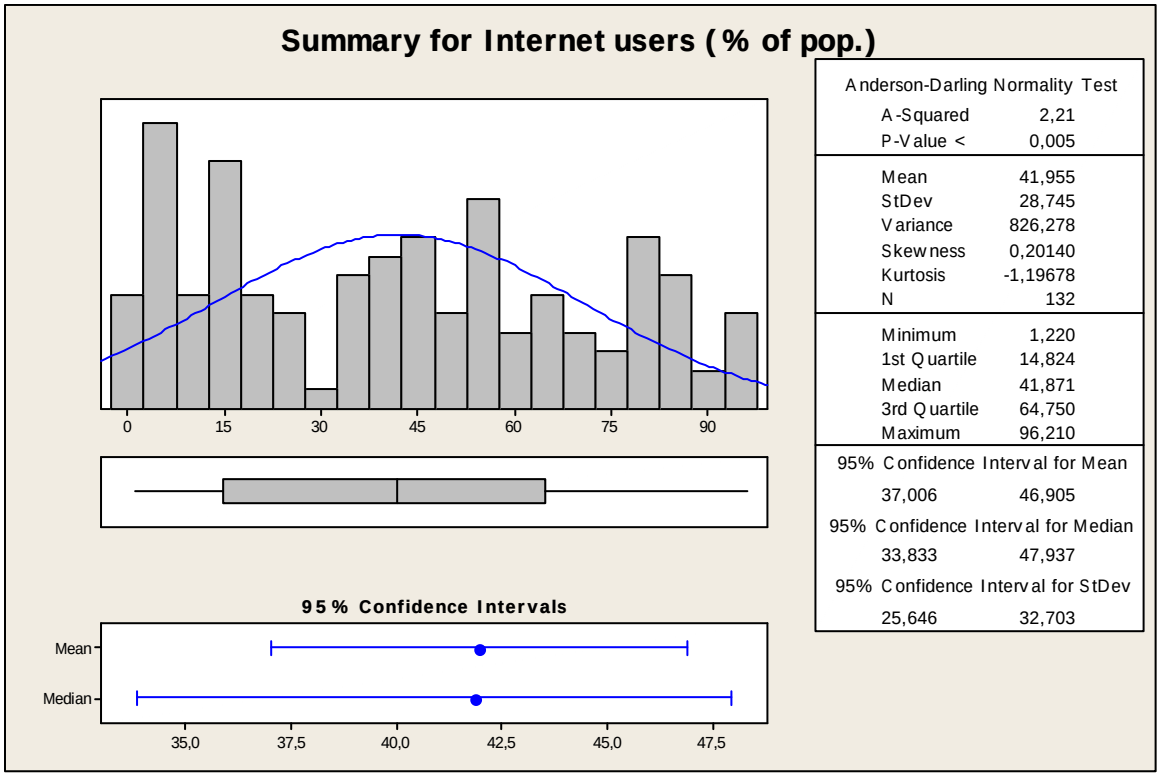
O índice de natimortos deve ser interpretado como uma escala inversa, quanto maior melhor o índice. Verifica-se nesta amostra de 131 países que a amostra não é normalmente distribuída, com concentração à esquerda (melhores medidas), do gráfico, em especial nos valores acima de 7,5 por mil. A média da amostra é 12,88, a mediana 10,00 e o desvio padrão de 9,73.Há um outlier, o Paquistão, com 47 natimortos por mil habitantes.

x. HOMICIDE RATE – Taxa de Homicídios



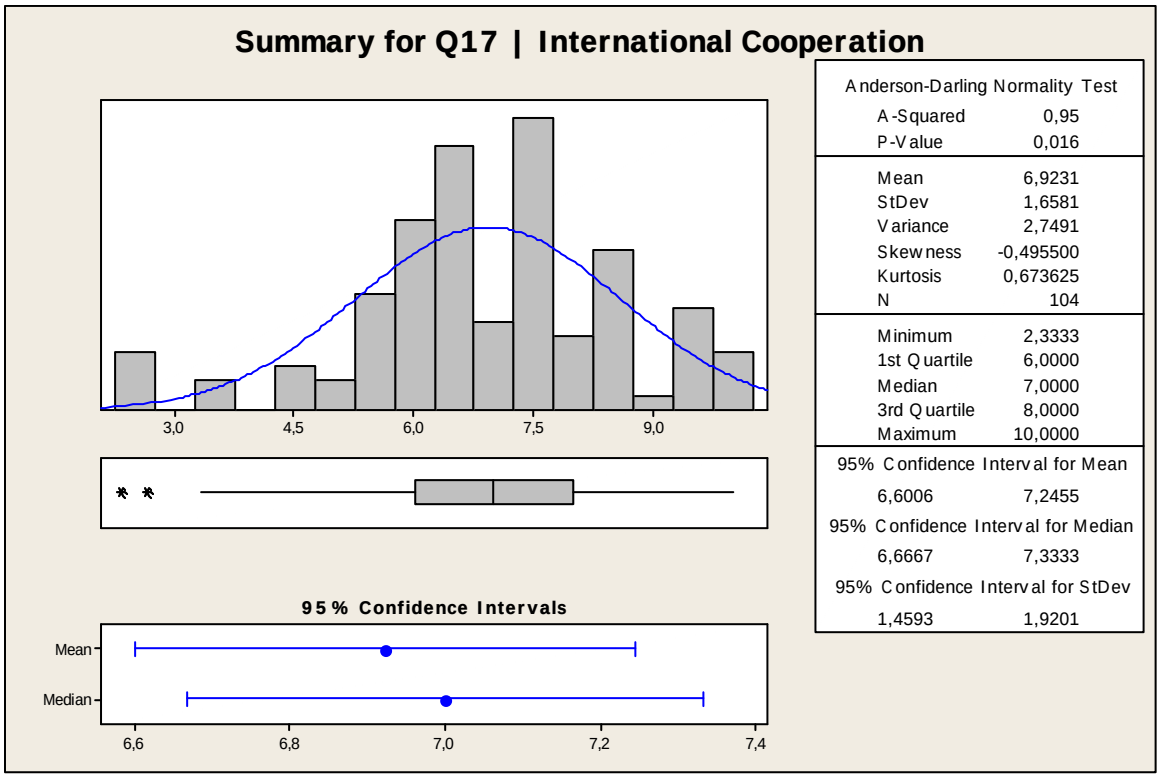
A taxa de homicídios compõe uma escala como apresentado no quadro da sessão anterior. Não é uma amostra com distribuição normal e o gráfico se concentra nos extremos, 1 (inferior a 2 homicídios por cem mil habitantes) e 5 (superior a 20 por cem mil habitantes). Não há outliers.

xi. INTERNET USERS – Usuários de Internet como % da população.



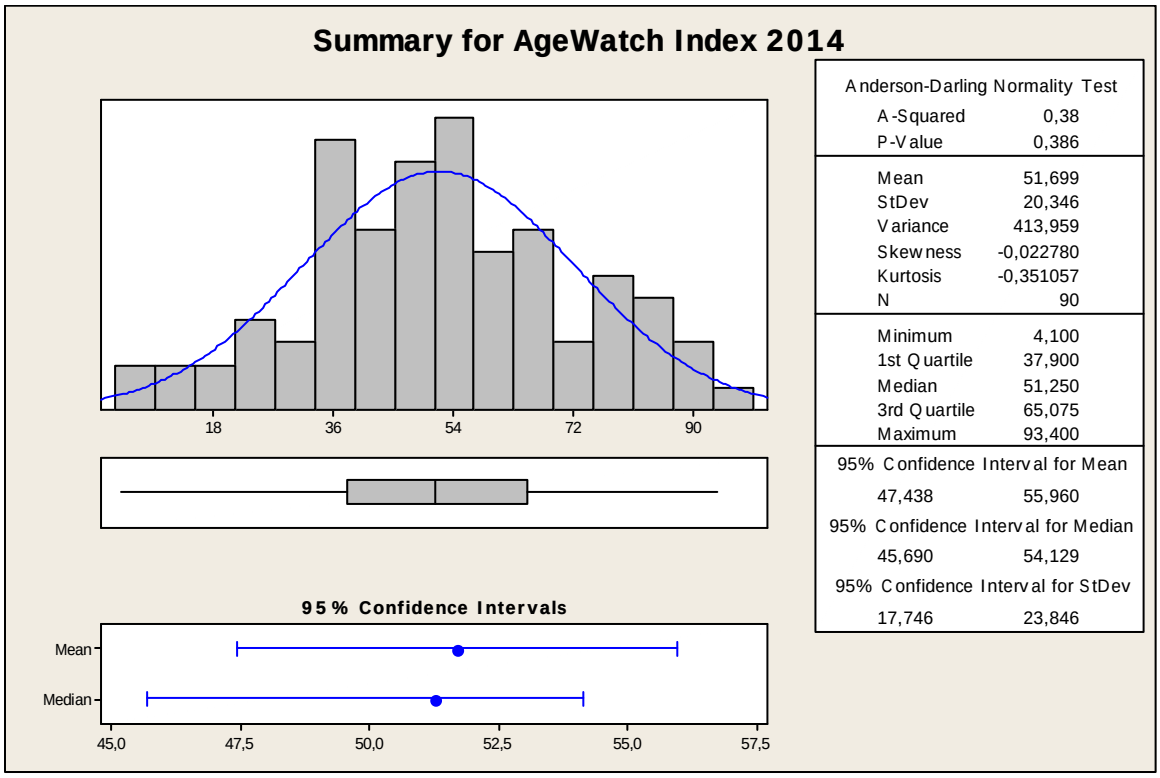
A taxa de usuários de internet não é uma amostra com distribuição normal para os 132 países observados. A média é de 41,955, mediana de 41,871, desvio padrão de 28,745. O r´fico novamente concentra-se nas extremidades, abaixo de 14,82% no 1º quartil e acima de 64,75% no 3º. Quartil.

xii. INTERNATIONAL COOPERATION – Cooperação internatcional.



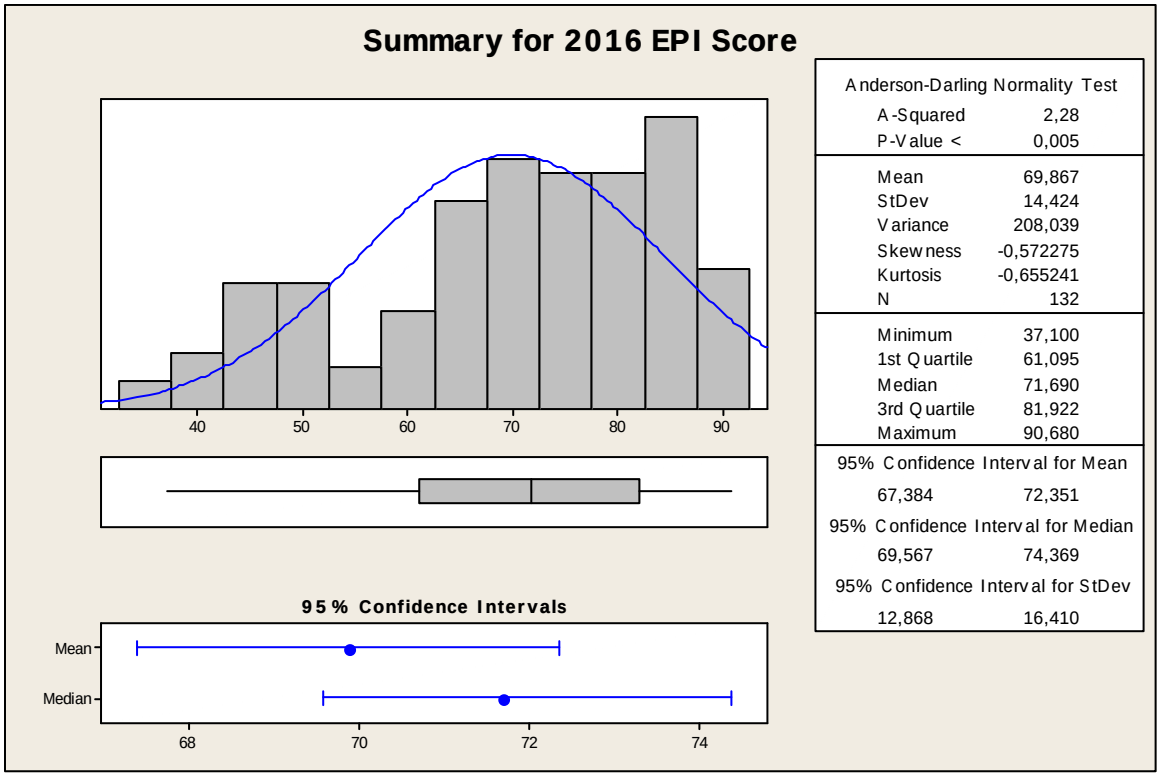
O indicador de Cooperação Internacional não apresenta distribuição normal. Há no extremo de menor cooperação dois outliers (Sudão e Rússia). Amédia é de 6,92, mediana de 7,000 e desvio padrão de 1,658. Cabe ressaltar que este índice não contempla medições para os países da Comunidade Européia, e países anglo saxônicos desenvolvidos, como EUA, Austrália, Nova Zelândia e Canadá, o que pode dar certo viés à amostra, mas também indica como possível intencionalidade de criação deste índice medir a cooperação para com os países omissos.

xiii. AGEWATCH INDEX – Idade e envelhecimento da população.



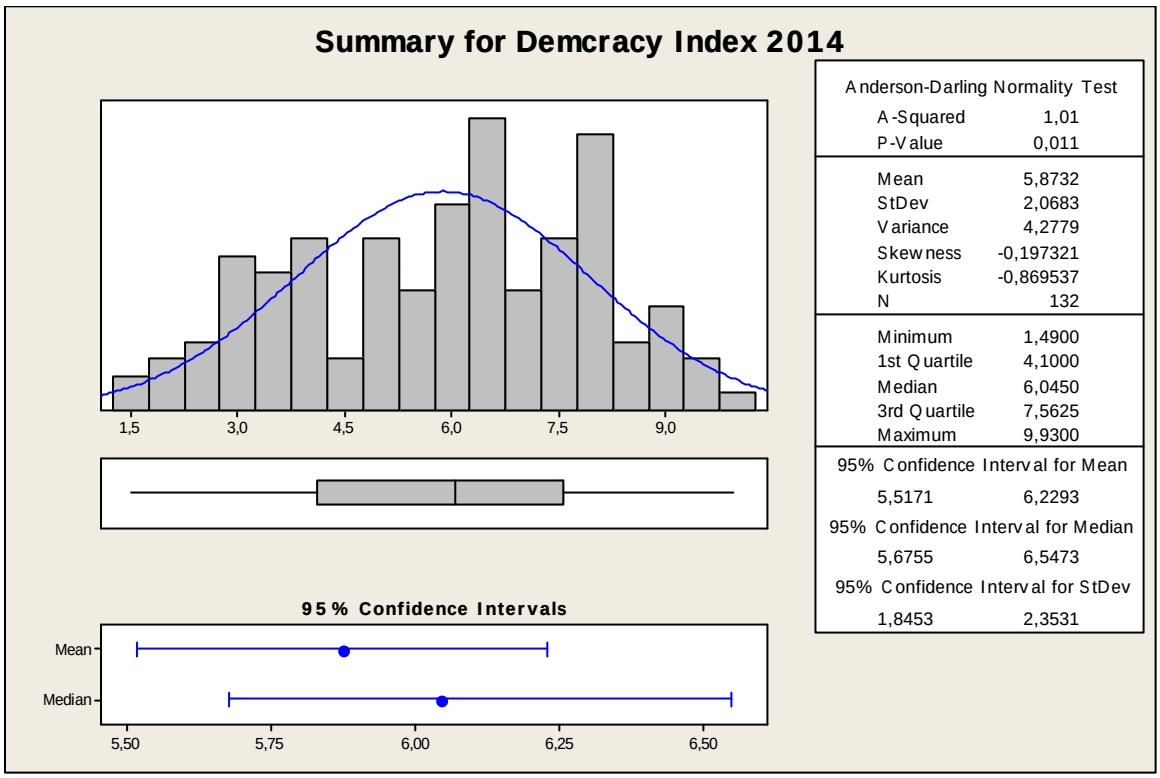
O Agewatch índex indica como uma sociedade lid com o envelhecimento de sua população. Verificamos nesta amostra de 90 países que ocorre uma distribuição normal, com média de 51,699, mediana de 51,250 e desvio padrão de 20,346. Não há outliers.

xiv. EPI – ENVIRONMENT PROTECTION INDEX – Proteção do meio ambiente



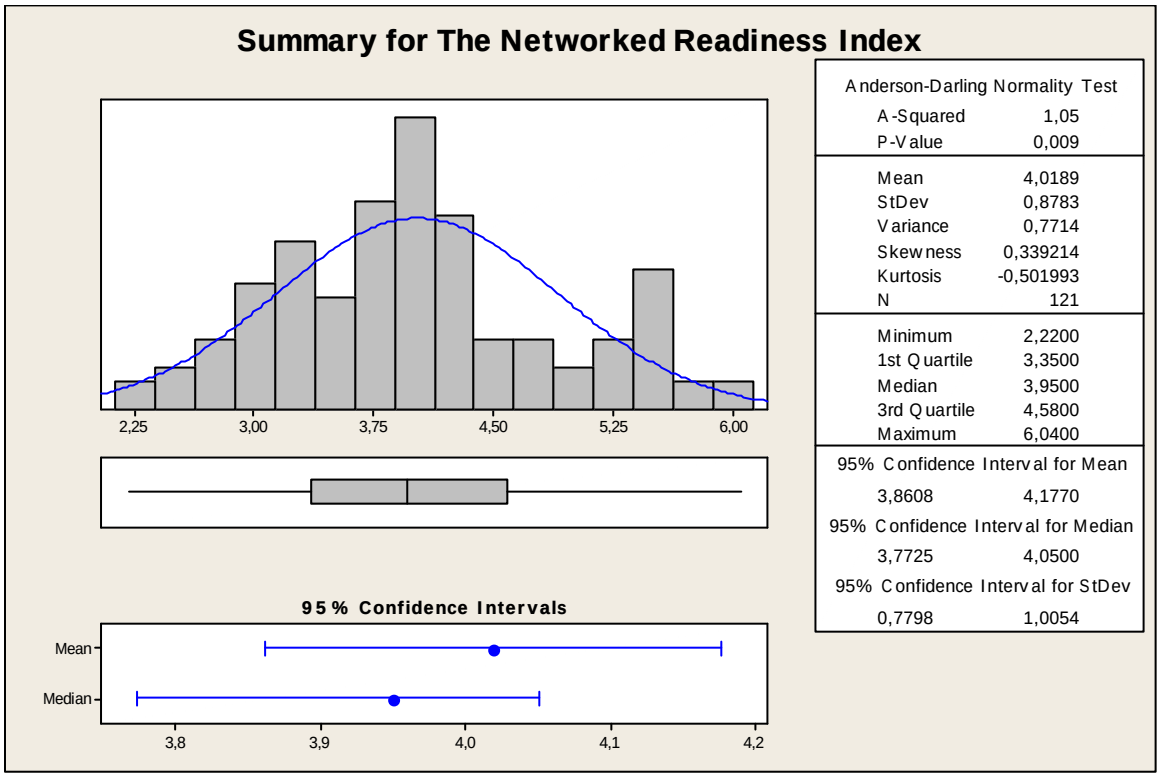
O EPI índex não possui distribuição normal, com concentração maior à direita do gráfico (melhores desempenhos), na combinação de indicadores de proteção à saúde humana e ao ecossistema. A mediana é 71,69, com o terceiro quartil iniciando-se em 81,922. Uma visão baseada no senso comum questionaria o índice, dado o cenário midiático de crises ambientais e humanas.

xv. DEMOCRACY INDEX – Democracia



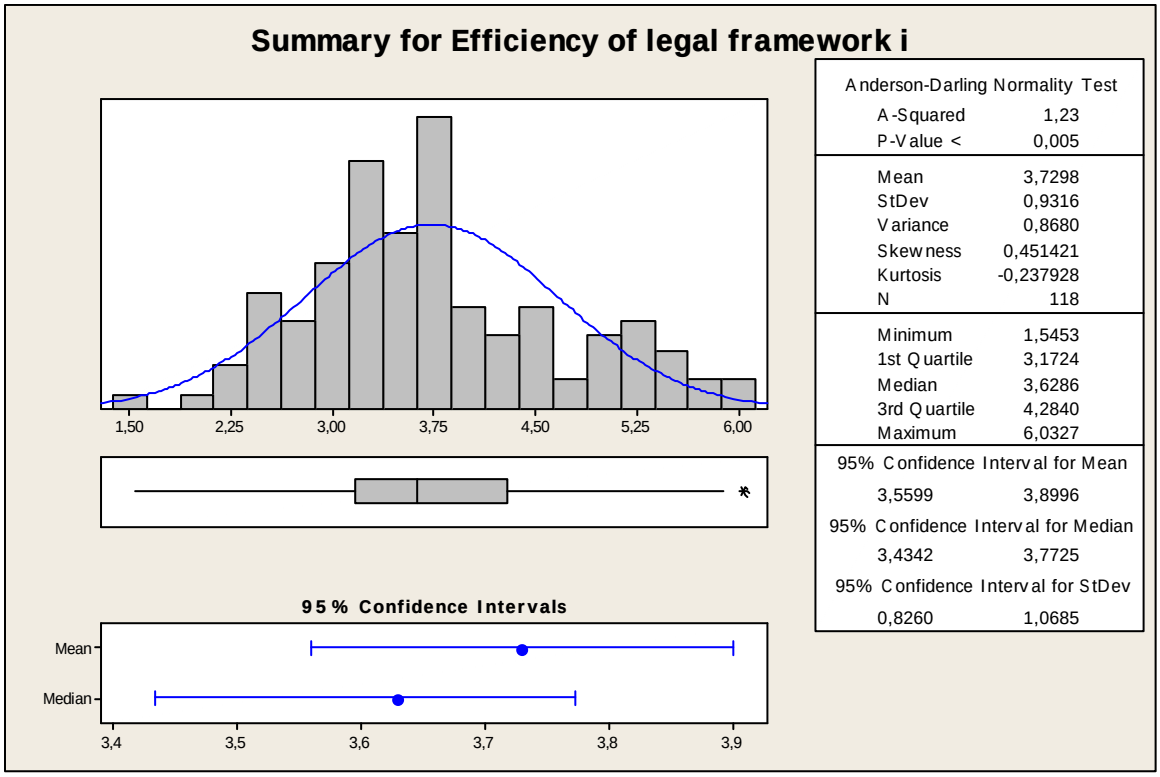
Neste índice, considerando o estudo original, cabe lembrar a classificação original: democracias plenas são valoradas até 8,05 (8 como corte), já dentro do quarto quartil, com apenas 20 países; democracias falhas até 6, 59 países, regimes híbridos até 4 (37 países) e ditaduras abaixo de 4 (20). Portnto países com traços democráticos mesmo que imperfeitos perfazem somente pouco mais da metade da amostra. A média ficaria como um regime híbrido e a mediana como um regime de democracia falha, como personificação da amosra de 132 países.

xvi. NETWORK READINESS – Disponibilidade de rede



A amostra de disponibilidade de rede apresenta em sua amostra de 121 países uma concentração emu ma das faixas à direita do gráfico, não possuindo distribuição normal. A media é de 4,0189, mediana de 3,9500 e desvio padrão de 0,8783.

xvii. Efficiency of legal framework in settling disputes, 1-7 (best) – Eficiência do sistema legal em resolver disputas



O índice de Eficiência do sistema legal em resolver disputas possui 118 observações que não estão normalmente distribu~idas. Há concentração à esquerda do gráfico (pior avaliação conforme a escala de medição). A média é 3,7298, A escala adotada é 1 extremamente ineficiente em resolver disputas e 7 como extremamente eficiente.

g. Considerações Finais

Os indicadores apresentados compõe um conjunto de índices compostos ou mensurados diretamente por diferentes organismos e entidades internacionais. Foram selecionados pela possibilidade de , em conjunto, darem uma medida da situação mundial por país da condição dos EIXO PESSOAS. São considerados 132 países no painel, porém nem todos os índices contém observações válidas para todos os países pesquisados. Optou-se por manter as informações assim como foram obtias em suas fontes originais, sem normalização ou ajustes. Ressalta-se a posição de destaque dos países latino-americanos no índice de felicidade pesquisado.

2. ANÁLISE DE CORRELAÇÕES

Este trabalho tem como objetivo realizar uma análise das correlações entre as variáveis do EIXO PESSOAS pra 132 países, comparados a partir de dezessete variáveis quantitativas, seis das quais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Empregou-se o Minitab ® para as análise estatísticas.

a. CORRELAÇÕES

Empregou-se o Minitab para construir uma análise de correlação entre as variáveis do EIXO PESSOAS, conforme a Tabela 1 a seguir. Foram encontradas (marcadas em amarelo) nesta tabela as trinta correlações entre variáveis com força explicativa superior a 0,80 ou 80% e p value significativo a 5% ou menos. Destas, seis possuem correlação superior a 90% e seis correçãoão entre 88,5% e inferiores a 90%. Selecionaram-se aquelas com força explicativa superior a 90% (chegando até 93,6%) e p value inferior a 1% para análise.

Tabela 1: Correlçaões entre as variáveis do Eixo Pessoas

Correlations: 1. Governanç; Happy Planet; GINNI Index; Social Progr; ...

	1. Governança	Happy Planet Ind	GINNI Index
Happy Planet Ind	0,089 0,317		
GINNI Index	-0,258 0,005	0,055 0,556	
Social Progress	0,887 0,000	0,295 0,001	-0,347 0,000
Access to piped	0,652 0,000	0,366 0,000	-0,240 0,009
Access to improv	0,561 0,000	0,381 0,000	-0,370 0,000
Q17 | Internatio	0,758 0,000	-0,010 0,919	0,103 0,329
2016 EPI Score	0,713 0,000	0,306 0,000	-0,314 0,001
Shelter	0,696 0,000	0,405 0,000	-0,316 0,000
Stillbirth rate	-0,621 0,000	-0,268 0,002	0,387 0,000
Homicide rate (1	-0,538 0,000	-0,075 0,401	0,648 0,000
Internet users (	**0,821** 0,000	0,146 0,099	-0,429 0,000
Human Developmen	0,767 0,000	0,328 0,000	-0,391 0,000
AgeWatch Index 2	0,805 0,000	0,152 0,154	-0,313 0,004
Demcracy Index 2	0,849 0,000	0,233 0,008	-0,126 0,171
The Networked Re	0,877 0,000	0,161 0,082	-0,409 0,000
Efficiency of le	0,657 0,000	0,017 0,853	0,009 0,930

	Social Progress	Access to piped	Access to improv
Access to piped	0,859 0,000		
Access to improv	0,802	0,875	

	0,000	0,000	
Q17 | Internatio	0,457	0,272	0,142
	0,000	0,005	0,151
2016 EPI Score	0,891	0,856	0,832
	0,000	0,000	0,000
Shelter	0,890	0,889	0,885
	0,000	0,000	0,000
Stillbirth rate	-0,804	-0,789	-0,796
	0,000	0,000	0,000
Homicide rate (1	-0,595	-0,531	-0,607
	0,000	0,000	0,000
Internet users (	0,905	0,824	0,787
	0,000	0,000	0,000
Human Developmen	0,936	0,886	0,879
	0,000	0,000	0,000
AgeWatch Index 2	0,887	0,688	0,676
	0,000	0,000	0,000
Demcracy Index 2	0,791	0,571	0,434
	0,000	0,000	0,000
The Networked Re	0,892	0,724	0,730
	0,000	0,000	0,000
Efficiency of le	0,473	0,232	0,262
	0,000	0,013	0,005

	Q17 | Internatio	2016 EPI Score	Shelter
2016 EPI Score	0,310		
	0,001		
Shelter	0,229	0,858	
	0,019	0,000	
Stillbirth rate	-0,174	-0,799	-0,800
	0,079	0,000	0,000
Homicide rate (1	-0,074	-0,515	-0,591
	0,458	0,000	0,000
Internet users (	0,315	0,841	0,846
	0,001	0,000	0,000
Human Developmen	0,249	0,911	0,919
	0,011	0,000	0,000
AgeWatch Index 2	0,370	0,720	0,770
	0,002	0,000	0,000
Demcracy Index 2	0,708	0,649	0,564
	0,000	0,000	0,000
The Networked Re	0,402	0,785	0,838
	0,000	0,000	0,000
Efficiency of le	0,214	0,300	0,425
	0,042	0,001	0,000

	Stillbirth rate	Homicide rate (1	Internet users (
Homicide rate (1	0,518		

	0,000		
Internet users (	-0,772	-0,643	
	0,000	0,000	
Human Developmen	-0,855	-0,640	0,904
	0,000	0,000	0,000
AgeWatch Index 2	-0,668	-0,520	0,813
	0,000	0,000	0,000
Demcracy Index 2	-0,492	-0,294	0,638
	0,000	0,001	0,000
The Networked Re	-0,755	-0,598	0,902
	0,000	0,000	0,000
Efficiency of le	-0,286	-0,265	0,441
	0,002	0,004	0,000

	Human Developmen	AgeWatch Index 2	Demcracy Index 2
AgeWatch Index 2	0,839		
	0,000		
Demcracy Index 2	0,646	0,739	
	0,000	0,000	
The Networked Re	0,877	0,829	0,669
	0,000	0,000	0,000
Efficiency of le	0,384	0,526	0,410
	0,000	0,000	0,000

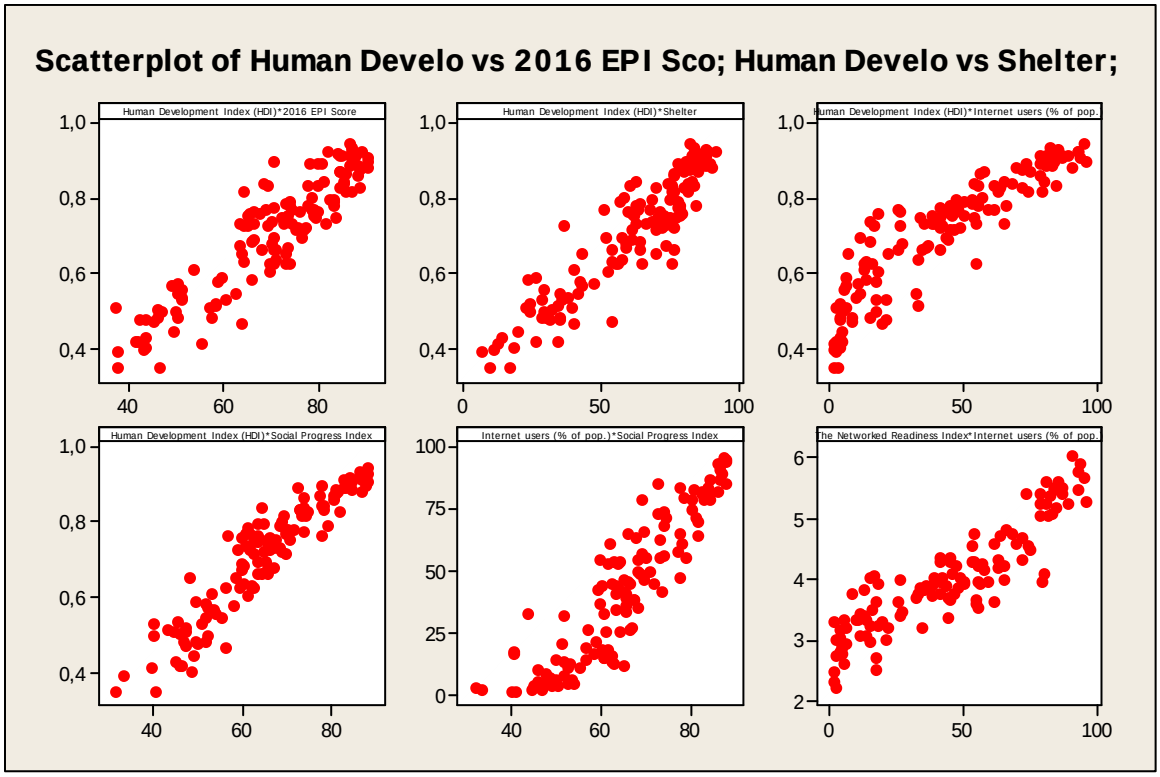
	The Networked Re
Efficiency of le	0,657
	0,000

Cell Contents: Pearson correlation
P-Value

Fonte: preparado pelo autor com o Minitab 16 ®

i. Análise da dispersão das principais correlações

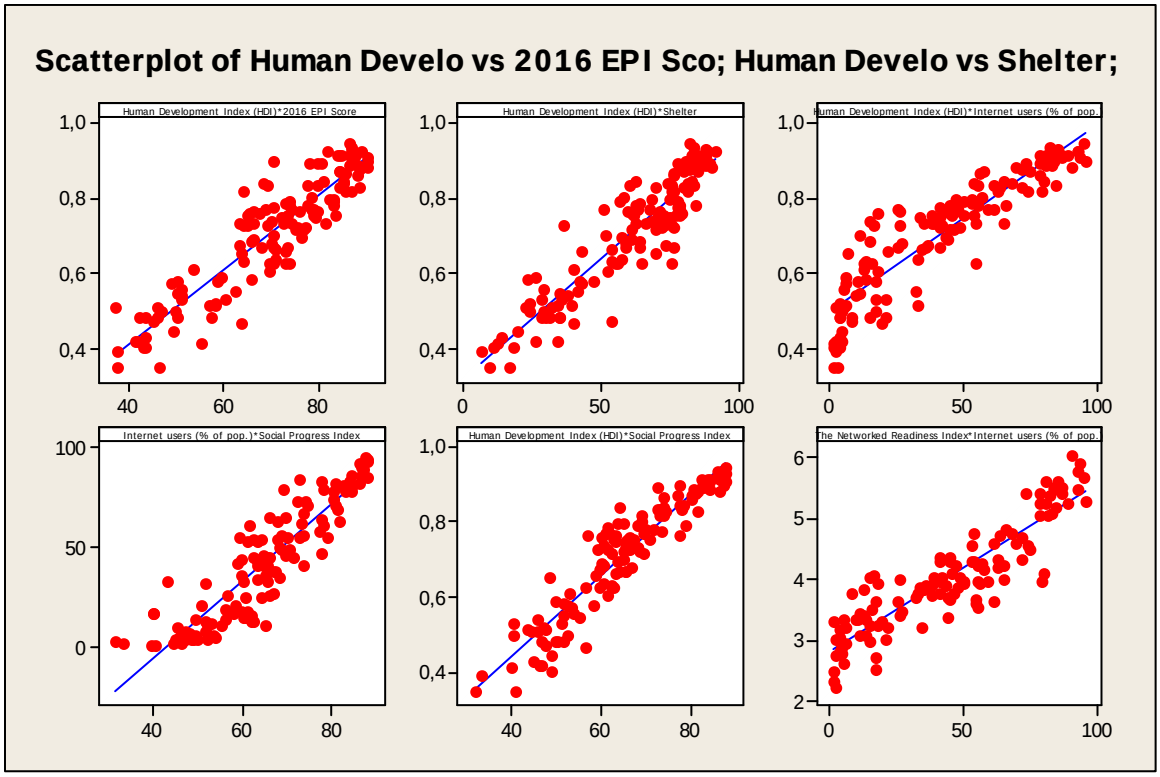
Gráfico 2: correlações Superiores a 90,0% e significativas a 1%.



Verificamos que as quatro primeiras das seis principais correlações entre variáveis ocorrem com a variável sintética IDH – Indice de Desenvolvimento Humano. Em todas há erta concentração na parte superior das observações dos dois índices. Para a correlação entre usuários de internet e SPI a concentração se dá nas observações de menor valor, assim como naquela entre internet users e Networked readiness.

ii. Análise da dispersão e Regressão das principais correlações

Gráfico 3: correlações Superiores a 90,0% e significativas a 1% com linha de Regressão.

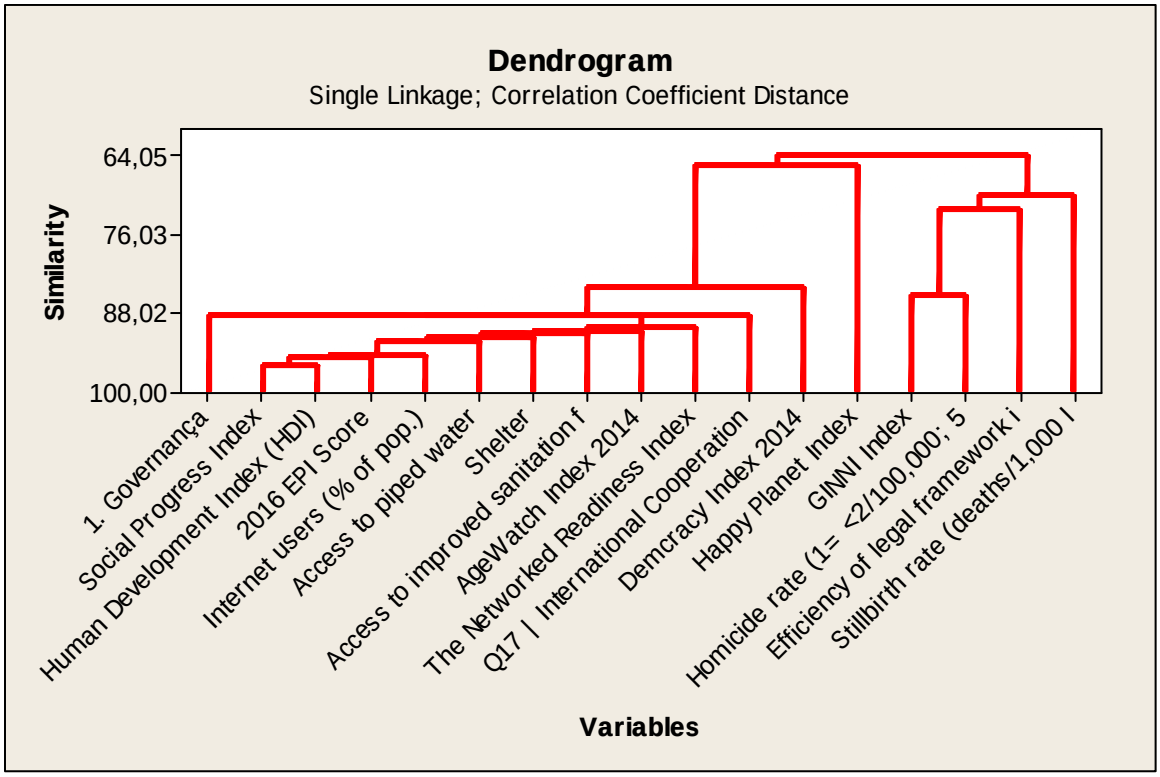


O Gráfico 3 apesenta as correlações com a linha de regressão resultante.

iii. Análise de similaridade entre as variáveis – emprego do gráfico de Dendograma

Um dos desafios de empregar-se análises com muitas variáveis é buscar identificar aquelas que são mais relevantes e em especial se o poder explicativo de acrescentar-se mais uma variável compensa em contrapartida À complexificação do modelo. A análise através do dendograma visa empregar o princípio da parcimônia na seleção das variáveis a serem utilizadas para caracterizar o EIXO PESSOAS. São realmente necessárias todas as dezessete variáveis para tanto?

Gráfico 3 – Dendograma 1



Fonte: Preparado pelo autor com o minitab 16 ®

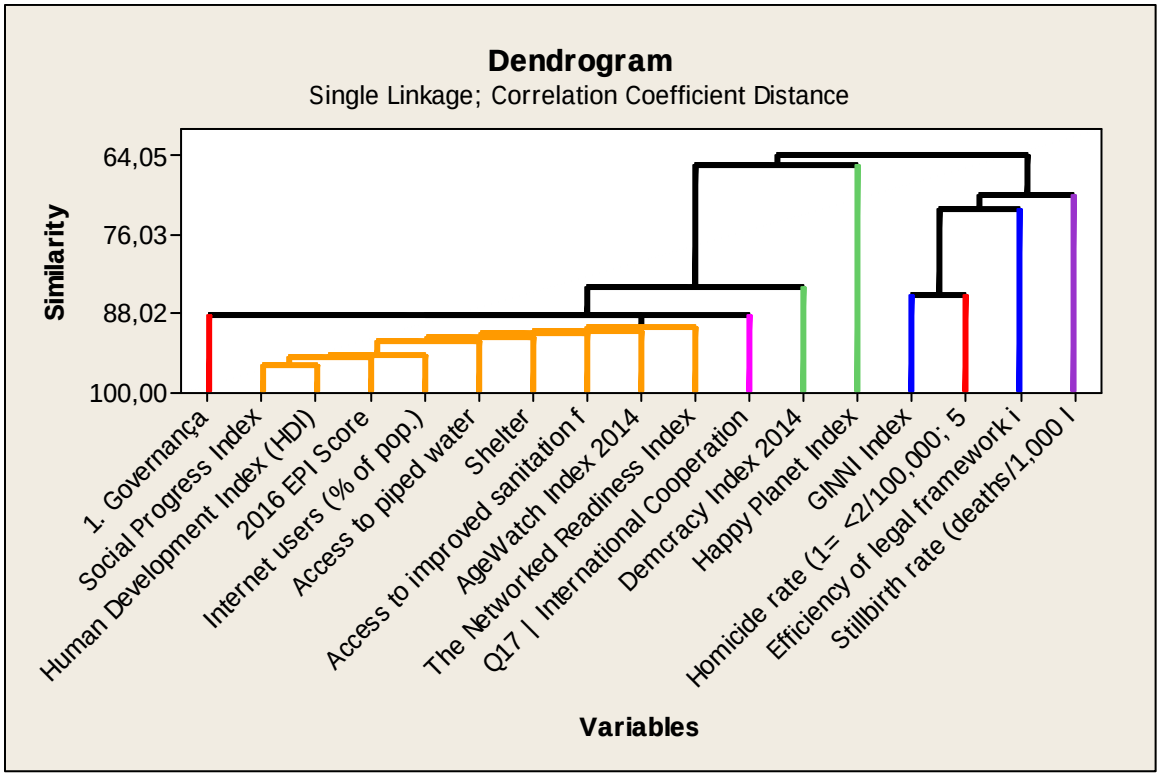
Quando se observa o dendograma do gráfico 3 observa-se 9 ou 10 variáveis poderiam ser selecionadas para explicar as observações, pois há similaridade de quase 90% aproximadamente entre algumas delas. A tabela 2 – análise de similaridade apresenta esta análise gráfica em números.

Nove variáveis seria necessárias para entender a amostra. AS oito demais poderiam ser descartadas por similaridade com as nove primeiras.

Poderiam ser selecionadas então do conjunto de variáveis aquelas com menor similaridade, e que dariam adequada significância ao conjunto:

1. Stillbirth Rate
2. Efficiency of legal framework to settle disputes
3. Ginni Index
4. Homicide Rate
5. Happy Planet Index Democracy Index
6. Democracy Index
7. International Cooperation
8. Governança
9. Uma única dentre todas as demais variáveis: SPI, HDI, EPI, Internet Users, Access to Piped Water, Shelter, Access to improved sanitation, Agewatch, Networked Readiness

Gráfico 3 – Dendograma 2



Quando se refaz a análise e se prepara um Dendograma com nove clusters, chega-se exatamente ao resultado proposto.

b. Considerações Finais

Como colcusão desta etapa verifica-se que seria possível desenvolver o EIXO PESSOAS que inicialmente possuía dezessete variáveis, com apenas nove delas, como resultado da análise de correlação e similaridade. São elas:

1. Stillbirth Rate
2. Efficiency of legal framework to settle disputes
3. Ginni Index
4. Homicide Rate
5. Happy Planet Index Democracy Index
6. Democracy Index
7. International Cooperation
8. Governança
9. Das demais variáveis, opta-se por empregas o Human Development Index, ou HDI, por sua maior correlação com as de mais variáveis.

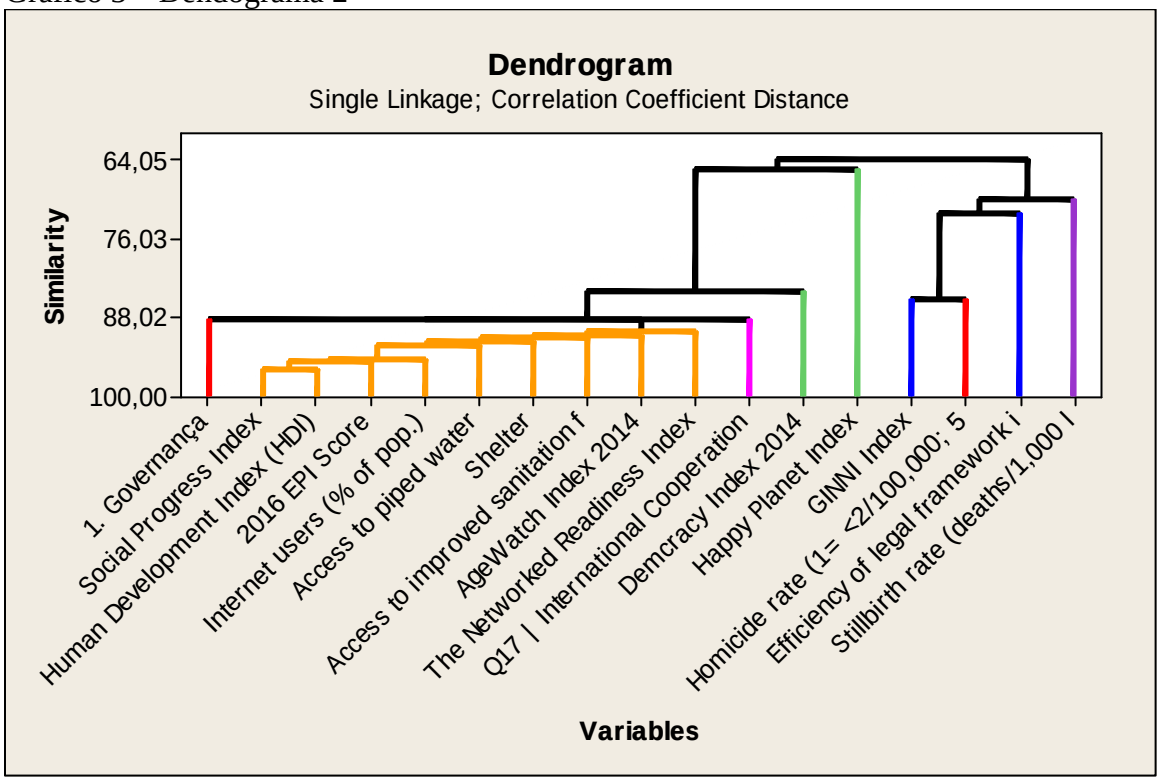
3. REGRESSÃO E STEPWISE

Este trabalho tem como objetivo construir um modelo de regressão múltipla entre as variáveis do EIXO PESSOAS pra 132 países, a partir da seleção de uma variável dependente e de variáveis dependentes explicativas. Empregou-se o Minitab ® para as análise estatísticas.

a. SELEÇÃO DAS VARIÁVEIS

Empregou-se o Minitab para construir um dendograma com as variáveis do EIXO PESSOAS, conforme a Tabela 1 a seguir.

Gráfico 3 – Dendograma 2



Verifica-se que seria possível desenvolver o EIXO PESSOAS com apenas nove variáveis, como resultado da análise de correlação e similaridade. São elas:

1. Stillbirth Rate
2. Efficiency of legal framework to settle disputes
3. Ginni Index
4. Homicide Rate
5. Happy Planet Index Democracy Index
6. Democracy Index
7. International Cooperation
8. Governança
9. Das demais variáveis, opta-se por empregas o Human Development Index, ou HDI, por sua maior correlação com as de mais variáveis.

Em análise anterior das correlações verificou-se que o IDH, Indice de Desenvolvimento Humano, um índice sintético amplo, possui forte correlação com outras quatro variáveis. Por estas duas características opta-se por estabelecer o IDH como variável dependente da regressão múltipla. As demais oito variáveis foram estabelecidas como independentes.

b. REGRESSÃO – 1ª. RODADA SEM STEPWISE

Uma primeira versão da regressão é desenvolvida no Minitab 16® coma s oito variáveis independentes e o IDH como dependente. O resultado é apresentado a seguir.

Regression Analysis: Human Develo versus Stillbirth r; Efficiency o; ...

```
The regression equation is
Human Development Index (HDI) = 0,732 - 0,00789 Stillbirth rate (deaths/1,000 l
                                - 0,0163 Efficiency of legal framework i
                                + 0,00009 GINNI Index
                                - 0,0146 Homicide rate (1= <2/100,000; 5
                                + 0,00137 Happy Planet Index
                                + 0,00959 Democracy index
                                - 0,0209 Q17 | International Cooperation
                                + 0,00397 1. Governança


78 cases used, 54 cases contain missing values


Predictor                             Coef    SE Coef      T      P
Constant                           0,73180    0,06291  11,63  0,000
Stillbirth rate (deaths/1,000 l  -0,007887   0,001187  -6,64  0,000
Efficiency of legal framework i   -0,01630    0,01283  -1,27  0,208
GINNI Index                        0,000092   0,001255   0,07  0,942
Homicide rate (1= <2/100,000; 5  -0,014560   0,007471  -1,95  0,055
Happy Planet Index                0,0013682  0,0008970   1,53  0,132
Democracy index                    0,009587   0,008327   1,15  0,254
Q17 | International Cooperation  -0,020911   0,008191  -2,55  0,013
1. Governança                      0,003967   0,001213   3,27  0,002


S = 0,0606505   R-Sq = 80,5%   R-Sq(adj) = 78,3%


Analysis of Variance

Source          DF       SS       MS      F      P
Regression       8  1,05066  0,13133  35,70  0,000
Residual Error  69  0,25382  0,00368
Total           77  1,30448


Source                           DF   Seq SS
Stillbirth rate (deaths/1,000 l   1  0,89683
Efficiency of legal framework i   1  0,00038
GINNI Index                       1  0,00403
Homicide rate (1= <2/100,000; 5   1  0,04147
Happy Planet Index                1  0,00032
Democracy index                   1  0,06384
Q17 | International Cooperation   1  0,00445
1. Governança                     1  0,03934


Unusual Observations

                          Human
     Stillbirth rate  Development
Obs  (deaths/1,000 l  Index (HDI)      Fit   SE Fit  Residual  St Resid
 32            27,0      0,87009  0,67996  0,03171   0,19013      3,68R
 92            42,0      0,51402  0,35447  0,02458   0,15955      2,88R

R denotes an observation with a large standardized residual.
```

i. INTERPRETAÇÃO DOS RESULTADOS INICIAIS

Verifica-se que somente três das variáveis possuem p value inferior a 10%, parâmetro de confiança que se estabeleceu como limite de corte de significância para esta regressão. São elas Stillbirth rate, Homicide rate e Governança. Adicionalmente, o inercepto, ou constante, também possui p value significativo e é mantido na regressão. Nota-se que o R-quadrado, ou força expliativa da regressão é relativamente alto, 79%, em especial para um estudo em ciências sociais aplicadas.

c. REGRESSÃO COM VARIÁVEIS SIGNIFICATIVAS

Roda-se novamente no Minitab 16® a regressão, desta vez com as trÊs variáveis significativas mencionadas acima. Os resultados são apresentados a seguir.

Regression Analysis: Human Develo versus 1. Governanç; Stillbirth r; ...

```
The regression equation is
Human Development Index (HDI) = 0,750 + 0,00228 1. Governança
                                - 0,00909 Stillbirth rate (deaths/1,000 l
                                - 0,0181 Homicide rate (1= <2/100,000; 5


130 cases used, 2 cases contain missing values


Predictor                            Coef    SE Coef       T      P
Constant                          0,74953    0,02992   25,05  0,000
1. Governança                   0,0022837  0,0003515    6,50  0,000
Stillbirth rate (deaths/1,000 l  -0,0090869  0,0007630  -11,91  0,000
Homicide rate (1= <2/100,000; 5   -0,018082   0,004571   -3,96  0,000


S = 0,0636032   R-Sq = 83,7%   R-Sq(adj) = 83,4%


Analysis of Variance

Source           DF       SS       MS       F      P
Regression        3  2,62536  0,87512  216,33  0,000
Residual Error  126  0,50972  0,00405
Total           129  3,13507


Source                           DF   Seq SS
1. Governança                     1  1,82271
Stillbirth rate (deaths/1,000 l   1  0,73934
Homicide rate (1= <2/100,000; 5   1  0,06330


Unusual Observations

                    Human
            1.  Development
Obs  Governança  Index (HDI)      Fit   SE Fit  Residual  St Resid
 18          42      0,40229  0,53782  0,00982  -0,13554     -2,16R
 23          16      0,35013  0,47754  0,01199  -0,12741     -2,04R
 32          75      0,87009  0,65807  0,01869   0,21202      3,49R
 77          35      0,41929  0,56640  0,00865  -0,14711     -2,33R
 91          37      0,34825  0,55293  0,00834  -0,20468     -3,25R
 92          27      0,51402  0,35626  0,01838   0,15776      2,59R
 94          28      0,53838  0,33109  0,02315   0,20729      3,50RX
102          37      0,79787  0,67086  0,01019   0,12701      2,02R
```

```
130          22      0,76225  0,60941  0,01476   0,15284      2,47R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

O R-quadrado da regressão somente com constante e as variáveis selecionadas aumenta para 83%, com todos os elementos significativos a 1%. A relação da variável dependente Governança é positiva com a variável dependente IDH. Stillbirth (natimortos) e Homicides (homicídios) possuem correlação negativa. O coeficiente é positivo.

i. EMPREGO DA FUNÇÃO STEPWISE PARA IDENTIFICAR MELHORIAS NA REGRESSÃO

Novamente roda-se a regressão no Minitab16 ® para verificar oportunidades de aprimoramento. O resultado apresenta-se a seguir.

Stepwise Regression: Human Develo versus 1. Governanç; Stillbirth r; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15


Response is Human Development Index (HDI) on 3 predictors, with N = 130
N(cases with missing observations) = 2 N(all cases) = 132


Step                                    1         2         3
Constant                           0,8827    0,6877    0,7495

Stillbirth rate (deaths/1,000 l  -0,01366  -0,00992  -0,00909
T-Value                           -18,62    -12,80    -11,91
P-Value                            0,000     0,000     0,000

1. Governança                                0,00273   0,00228
T-Value                                         7,77      6,50
P-Value                                        0,000     0,000

Homicide rate (1= <2/100,000; 5                         -0,0181
T-Value                                                   -3,96
P-Value                                                   0,000

S                                  0,0813    0,0672    0,0636
R-Sq                                73,04     81,72     83,74
R-Sq(adj)                           72,83     81,43     83,35
Mallows Cp                           82,9      17,6       4,0
```

Pelo emprego da funão Stepwise constata-se que a inclusão das três variáveis aumenta a força explicativa da regressão, que com uma única variável – Stillbirth – possuía r-quadrdo de 73,0% e passa a 81,7% com a segunda variável – Governança – e vai a 83,7% com a terceira variável – Homicide.
Conclui-se que até esta etapa o IDH – Índice de Desenvolvimento Humano – como variável a dependente selecionada como e representativa do EIXO PESSOAS é explicada por trÊs variáveis a 83,74%: Natimortos, Taxa de Homicídio e Governança, sendo que as duas primeiras possuem correlação negativa com o IDH.

ii. EMPREGO DO AJUSTE DA AMOSTRA DA VARIÁVEL DEPENDENTE – “FIT”

Aplica-se agora a adequação dos valores apontados na etapa de regressão para aquelas observações que menos se ajustavam à regressão resultante no item 4. São nove observações. Gera-se então nova regressão com a base ajustada assumindo-se o valor de “FIT, gerando-se uma nova variável IDH-FIT.

Como se observa abaixo na “Regression Analysis”, a força explictiva da nova regressão chega a 91,1% de R-quadrado.

Regression Analysis: IDH - FIT versus 1. Governanç; Stillbirth r; ...

```
The regression equation is
IDH - FIT = 0,771 + 0,00198 1. Governança
            - 0,0105 Stillbirth rate (deaths/1,000 l
            - 0,0133 Homicide rate (1= <2/100,000; 5


130 cases used, 2 cases contain missing values


Predictor                              Coef    SE Coef       T      P
Constant                            0,77094    0,02165   35,60  0,000
1. Governança                     0,0019751  0,0002544    7,76  0,000
Stillbirth rate (deaths/1,000 l  -0,0105046  0,0005521  -19,03  0,000
Homicide rate (1= <2/100,000; 5   -0,013316   0,003308   -4,03  0,000


S = 0,0460256   R-Sq = 91,1%   R-Sq(adj) = 90,9%


Analysis of Variance

Source           DF       SS       MS       F      P
Regression        3  2,72853  0,90951  429,35  0,000
Residual Error  126  0,26691  0,00212
Total           129  2,99544


Source                            DF   Seq SS
1. Governança                      1  1,76529
Stillbirth rate (deaths/1,000 l    1  0,92891
Homicide rate (1= <2/100,000; 5    1  0,03433


Unusual Observations

            1.
Obs  Governança  IDH - FIT      Fit   SE Fit  Residual  St Resid
  9          32    0,57010  0,42967  0,01307   0,14043      3,18R
 28          28    0,59062  0,48606  0,00788   0,10456      2,31R
 47          25    0,41128  0,50229  0,00780  -0,09101     -2,01R
 54          27    0,76559  0,66157  0,00873   0,10403      2,30R
 75          44    0,44542  0,53943  0,00814  -0,09401     -2,08R
 85          41    0,41643  0,51863  0,00804  -0,10220     -2,26R
 94          28    0,33109  0,29167  0,01675   0,03942      0,92 X
103          52    0,48324  0,57819  0,00700  -0,09495     -2,09R
111          46    0,75666  0,65649  0,00564   0,10017      2,19R
130          22    0,76225  0,63227  0,01068   0,12998      2,90R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

d. NORMALIZAÇÃO DAS VARIÁVEIS

Procede-se nesta etapa À normalização das variáveis, para que as mesmas estejam em escalas de medida comparáveis tanto numericamente quanto em sentido de interpretação, permitindo maior acurácia e objetividade na construção e análise das correlações e na construção de regressões.

Para tanto, a variável dependente IDH e as oito variáveis independente selecionadas anteriormente foram normalizadas empregando-se as funções de MINIMO e MÁXIMO no Minitab 16®, conforme a fórmula da Equação 1:

Equação 1:
Variável Normalizada = (Variável Original-MIN(Variável Original))/(MAX(Variável Original)-MIN(Variável Original))*100

Adicionalmente, transformou-se o sentido das variávels GINNI e Stillbirth para que se ajustassem à lógica de escala das demais variáveis, onde as maiores observações correspondem às melhores situações, empregando-se a equação 2:

Variável Ajustada = 100-Variável normalizada

e. REGRESSÃO COM AS VARIÁVEIS NORMALIZADAS

A regressão resultante possui R-quadrado de 78,3%, semelhante ao encontrado nas regressões anteriores. As variáveis Eficiência em solucionar disputas, Democracia, Happy Planet Index e Ginni são descartadas por possuírem p value superior a 10%.

Regression Analysis: N - (HDI)_1 versus N Efficiency; N - Democrac; ...

```
The regression equation is
N -  (HDI)_1 = - 3,78 - 0,123 N Efficiency of legal framework
               + 0,136 N - Democracy index_1 + 0,596 N - Stillbirth rate
               + 0,0978 N - Homicide rate - 0,269 N -  International Cooperati_1
               + 0,565 N - Governança_1 + 0,0952 N - Happy Planet Index_1
               - 0,0059 N - GINNI


78 cases used, 54 cases contain missing values


Predictor                            Coef  SE Coef      T      P
Constant                           -3,781    8,999  -0,42  0,676
N Efficiency of legal framework  -0,12281  0,09667  -1,27  0,208
N - Democracy index_1              0,1359   0,1180   1,15  0,254
N - Stillbirth rate               0,59584  0,08970   6,64  0,000
N - Homicide rate                 0,09778  0,05017   1,95  0,055
N -  International Cooperati_1    -0,2692   0,1054  -2,55  0,013
N - Governança_1                   0,5646   0,1726   3,27  0,002
N - Happy Planet Index_1          0,09520  0,06242   1,53  0,132
N - GINNI                        -0,00593  0,08072  -0,07  0,942
```

```
S = 10,1827   R-Sq = 80,5%   R-Sq(adj) = 78,3%


Analysis of Variance

Source          DF       SS      MS      F      P
Regression       8  29615,6  3702,0  35,70  0,000
Residual Error  69   7154,4   103,7
Total           77  36770,1


Source                            DF   Seq SS
N Efficiency of legal framework    1    117,8
N - Democracy index_1              1   9245,2
N - Stillbirth rate                1  17348,9
N - Homicide rate                  1   1597,6
N -  International Cooperati_1     1    144,8
N - Governança_1                   1    898,5
N - Happy Planet Index_1           1    262,4
N - GINNI                          1      0,6


Unusual Observations

     N Efficiency
         of legal      N -
Obs     framework  (HDI)_1    Fit  SE Fit  Residual  St Resid
 32            38    87,61  55,69    5,32     31,92      3,68R
 92            39    27,83   1,04    4,13     26,79      2,88R

R denotes an observation with a large standardized residual.
```

Rodando-se novament a regressão com as variáveis significativas, Stillbirth, Homicide, International Cooperation e Governança, temos uma nova solução, com R-quadrado de 77,1%, levemente superior à regressão anterior.
A correlação das variáveis é interesante, pois com as variáveis normalizadas, a cooperação internacional é inversamente proporcional ao IDH, o que parece contra-intuitivo.

Regression Analysis: N - (HDI)_1 versus N - Stillbir; N - Homicide; ...

```
The regression equation is
N -  (HDI)_1 = - 3,00 + 0,625 N - Stillbirth rate + 0,103 N - Homicide rate
               - 0,262 N -  International Cooperati_1 + 0,613 N - Governança_1


102 cases used, 30 cases contain missing values


Predictor                          Coef  SE Coef      T      P
Constant                         -3,003    5,309  -0,57  0,573
N - Stillbirth rate             0,62499  0,06548   9,54  0,000
N - Homicide rate               0,10283  0,03539   2,91  0,005
N -  International Cooperati_1  -0,26204  0,08713  -3,01  0,003
N - Governança_1                 0,6132   0,1153   5,32  0,000


S = 11,3551   R-Sq = 77,1%   R-Sq(adj) = 76,2%


Analysis of Variance

Source           DF     SS     MS      F      P
Regression        4  42172  10543  81,77  0,000
Residual Error   97  12507    129
```

```
Total           101  54679


Source                           DF  Seq SS
N - Stillbirth rate               1   35285
N - Homicide rate                 1    2817
N -  International Cooperati_1    1     423
N - Governança_1                  1    3647


Unusual Observations

            N -
     Stillbirth      N -
Obs        rate  (HDI)_1    Fit  SE Fit  Residual  St Resid
 32          44    87,61  53,45    4,09     34,16      3,22R
 91          53     0,00  28,22    2,56    -28,22     -2,55R
 92          11    27,83   2,20    3,48     25,63      2,37R
 94           0    31,92   4,22    4,81     27,70      2,69RX
102          82    75,49  65,58    4,41      9,90      0,95 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

f. REGRESSÃO EMPREGANDO A FUNÇÃO STEPWISE COM AS VARIÁVEIS NORMALIZADAS

Rodamos agora com a função STEPWISE, iniciando-se com as oito variáveis independentes.

Chegamos a uma regressão com cinco das oito variáveis independentes e R-quadrado de 78%. As cinco variáveis Stillbirth rate, Govenança, Homicide Rate, International Cooperation e Democracy índex possuem p value inferior a 5%.

Stepwise Regression: N - (HDI)_1 versus N Efficiency; N - Democrac; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15


Response is N -  (HDI)_1 on 8 predictors, with N = 78
N(cases with missing observations) = 54 N(all cases) = 132


Step                                1        2        3        4        5
Constant                      -14,219  -16,873  -16,392   -9,679   -9,755

N - Stillbirth rate             0,955    0,795    0,744    0,716    0,681
T-Value                         12,93    10,67     9,88     9,43     8,92
P-Value                         0,000    0,000    0,000    0,000    0,000

N - Governança_1                         0,368    0,340    0,529    0,392
T-Value                                   4,59     4,32     4,03     2,71
P-Value                                  0,000    0,000    0,000    0,008

N - Homicide rate                                 0,085    0,073    0,105
T-Value                                            2,37     2,04     2,74
P-Value                                           0,020    0,045    0,008

N -  International Cooperati_1                                -0,18    -0,25
T-Value                                                       -1,79    -2,40
P-Value                                                       0,078    0,019
```

N - Democracy index_1					0,22
T-Value					2,05
P-Value					0,044
S	12,3	10,9	10,6	10,5	10,2
R-Sq	68,75	75,60	77,32	78,27	79,47
R-Sq(adj)	68,34	74,95	76,40	77,08	78,04
Mallows Cp	36,8	14,5	10,4	9,1	6,8

Chega-se ao final a este modelo de cinco variáveis explicativas da variável dependente IDH. Como se observa da regressão Stepwise quatro variáveis normalizadas (quanto maior melhor o indicador) tem correlação positiva com o IDH, Stillbirth rate, Govenança, Homicide Rate, e Democracy índex, euma variável, International Cooperation possui correlação negativa. Uma possível explicação é que esta variável não contempla observações para os países da União Européia e para os países anglo saxônicos desenvolvidos, o que pode causar um viés na amostra e na regressão. Opta-se por rodar novamente a regressão somente comas quatro variáveis de correlação positiva, com a função Stepwise, para comparação dos resultados.
Nesta variante o R-quadrado obtido é superior, chegando a 84,5%, e com todas as variáveis com correlação positiva. Nota-se que a amostra válida passa de 130 em 132, ao passo que quando se inclui a variável Cooperação Internacional a amostra válida decai para 78. Nota-se que a constante é negativa. Deste modo opta-se por assumir a regressão com quatro variáveis e gerada coma função Stepwise como regressão que mais se adéqua a explicar o IDH dos países da amostra, considerando este como Proxy do EIXO PESSOAS:
‘

Equação 3 – Todas as Variáveis Normalizadas

IDH = -17,08 + 0,684 Stillbirth rate + 0,146 Governança + 0,151 Homicide Rate + 0,192 Democracy

Stepwise Regression: N - (HDI)_1 versus N - Governan; N - Stillbir; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is N - (HDI)_1 on 4 predictors, with N = 130
N(cases with missing observations) = 2 N(all cases) = 132

Step	1	2	3	4
Constant	-18,07	-14,28	-13,68	-17,08
N - Stillbirth rate	1,032	0,749	0,687	0,684
T-Value	18,62	12,80	11,91	12,10
P-Value	0,000	0,000	0,000	0,000
N - Governança_1		0,389	0,325	0,146
T-Value		7,77	6,50	1,70
P-Value		0,000	0,000	0,092
N - Homicide rate			0,121	0,151
T-Value			3,96	4,68
P-Value			0,000	0,000
N - Democracy index_1				0,192
T-Value				2,52
P-Value				0,013

S	13,6	11,3	10,7	10,5
R-Sq	73,04	81,72	83,74	84,53
R-Sq(adj)	72,83	81,43	83,35	84,03
Mallows Cp	91,8	23,7	9,4	5,0

4. ANALISE DE TENDENCIAS – EIXO PESSOAS E EIXO PLANETA - PESSOAS: Expectativa de Vida Ao Nascer e PIB per capita / PLANETA: Área Florestada (% total) e Emissões CO2 per capita – Comparativo entre Brasil, México, Portugal e Espanha.

O presente trabalho visa realizar uma análise de tendência e projeções das variáveis temporais e quantitativas relevantes dentro do EIXO PESSOAS, "expectativa de vida ao nascer" e "PIB per capita" e do EIXO PLANETA, . Os dados referem-se ao Brasil México, Portugal e Espanha, todos países do mundo ibero-americano. As bases de dados foram obtidas a partir das coleções de informação do Knoema. A primeira refere-se a análise do comportamento histórico através de gráficos. A segunda trata da análise de tendências, buscando encontrar a função que melhor se adapte à cada uma das séries de variáveis analisadas. A terceira busca a elaboração e análise de projeções através de extrapolações estatísticas. O software estatístico utilizado é o **MINITAB.**

a. EIXO PESSOAS

i. Breve descrição do EIXO PESSOAS.

Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão "Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável."

ii. Variáveis do EIXO PESSOAS

Este tema da pesquisa apresenta dezessete variáveis, sendo oito índices sintéticos. Todas as variáveis são quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis componentes ou indicadores, a saber:

- **Índices sintéticos:** Índice de Progresso Social (SPI), o Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD – ONU, Índice de Governança, o Índice de Proteção Ambiental – EPI (*Environmental Protection Index*), o *Happy Planet Index* (HPI),Índice de Felicidade Planetária (HPI), Índice de Democracia, , Índice de observação etária (Age Watch Index), e o Índice de Ginni para os países selecionados.

- **Indicadores ou variáveis componentes:** Habitação (shelter), Taxa de não nascidos (stillbirth rate), acesso a água potável, acesso a saneamento básico melhorado, taxa de homicídios, percentual de usuários de internet sobre a população, cooperação internacional, disponibilidade de rede, eficiência das estruturaras legais em resolver disputas.

iii. A fonte e o tamanho da série de dados

Fonte:
A fonte empregada é o site ww.knoema.org. Para as séries de expectativa de vida ao nascer, as informações da expectativa de vida masculina e feminina ao nascer advém de três fontes: (1) Divisão de População das Nações Unidas. Perspectivas da População Mundial, (2) Divisão de Estatística das Nações Unidas. População e Relatório de Estatísticas (vários anos), (3) Censo dos respectivos países".

Tamanho da Série de Dados:
As três séries de dados podem ser consideradas satisfatórias para a realização desta pesquisa, uma vez que todas possuem dados de 15 anos ou mais, sem interrupções.

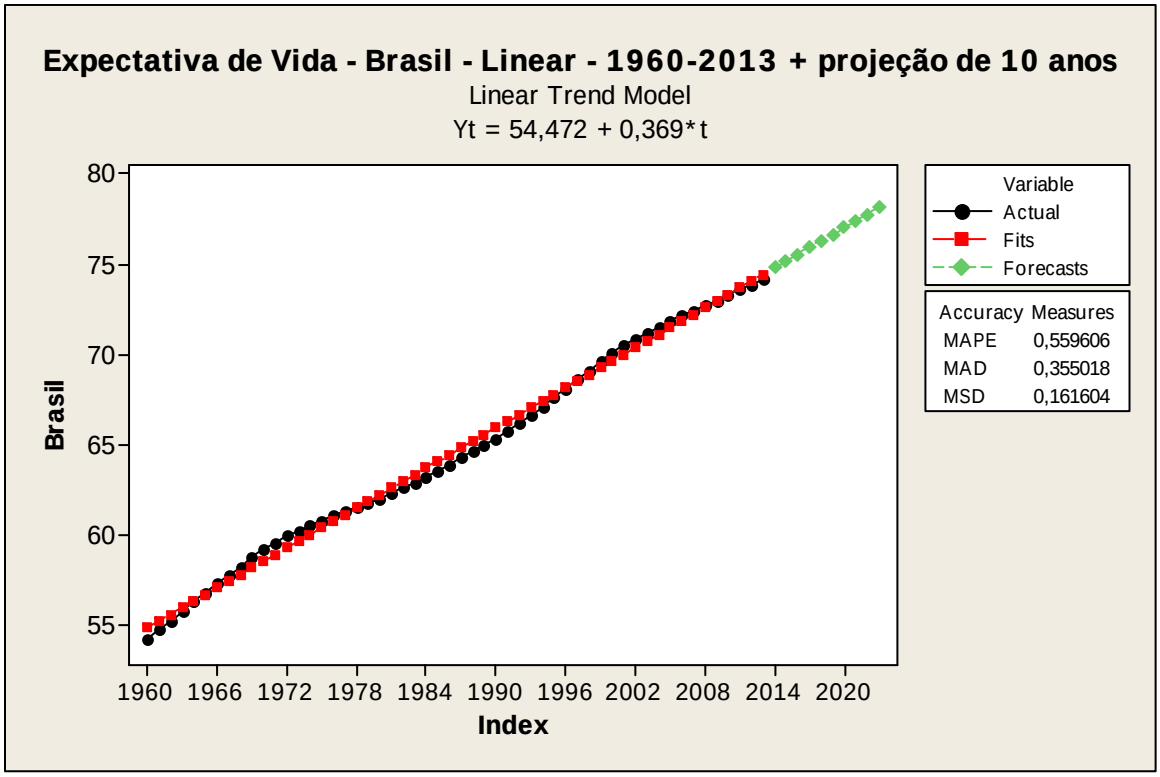
b. ANÁLISE DAS VARIÁVEIS DO EIXO PESSOAS

Cada variável é analisada a partir das séries dos quatro países mencionados. São desenvolvidos quatro modelos estatísticos: linear, quadrático, crescimento exponencial e curva s. Os resultados com os quatro modelos são analisados em relação ao fit, buscando-se interpretar qual dos modelos melhor explica o comportamento das observações. Uma linha de tendência de 10 anos (aunidade de tempo das observações) é traçada para estimação de comportamento futuro em cada modelo.

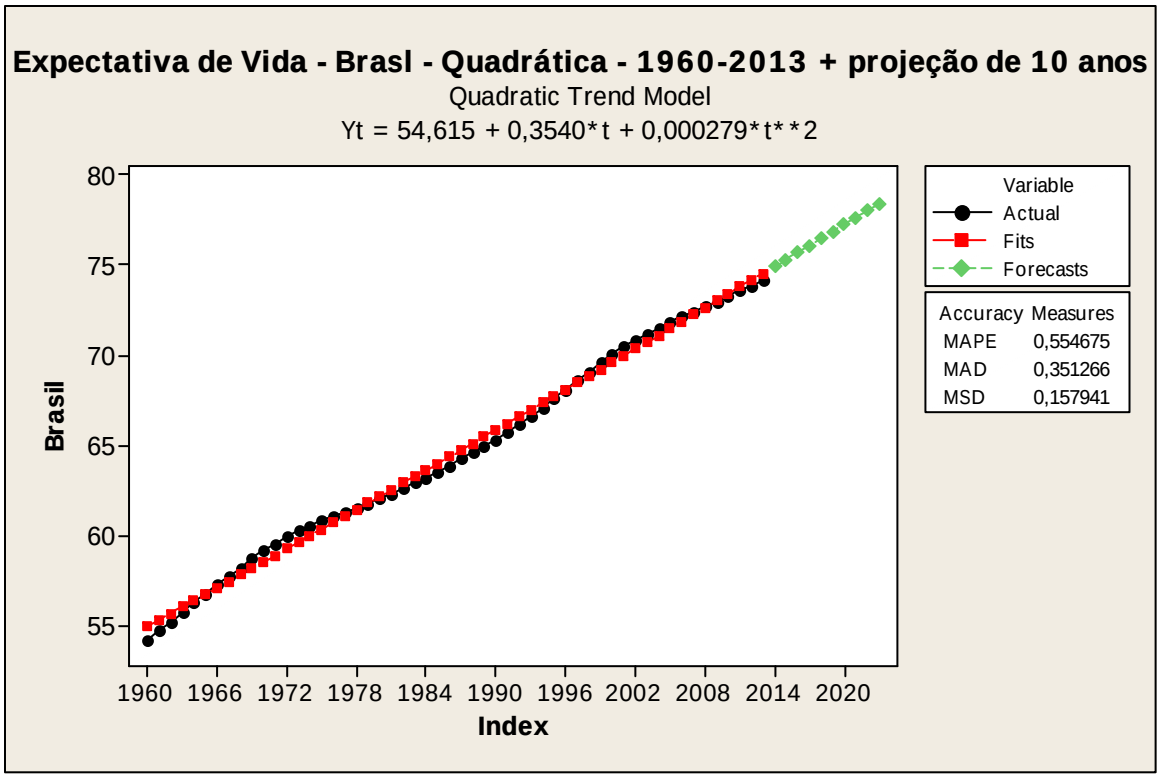
i. Variável: Expectativa de Vida ao Nascer (em anos)

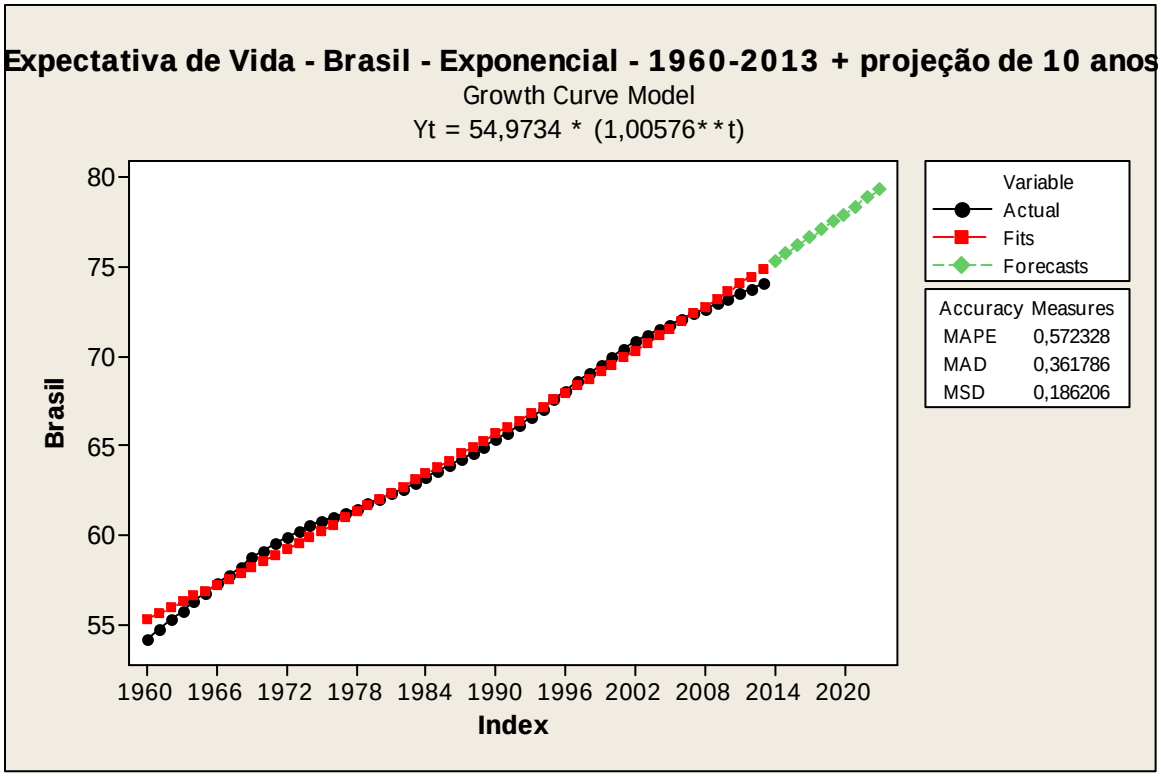
1. Análise de tendência e projeções - Brasil

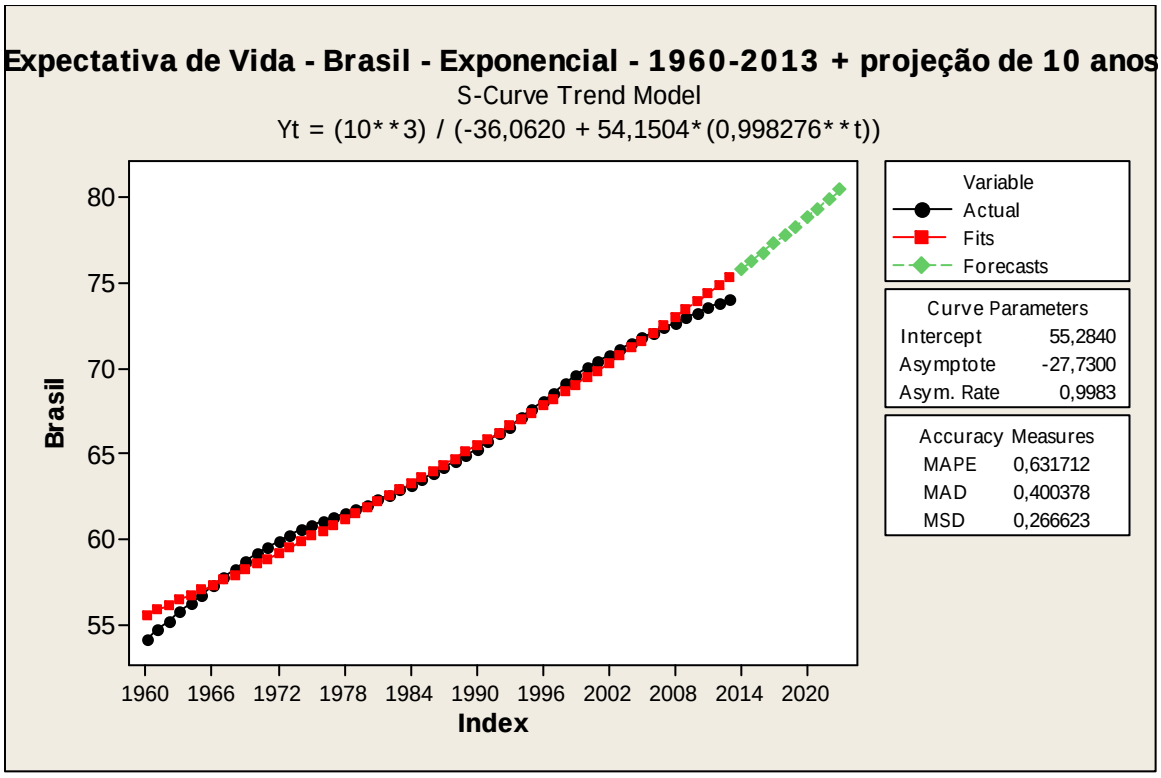
A expectativa de vida ao nascer no Brasil apresenta 54 observações ininterruptas entre 1960-2013, sempre crescentes. Apresenta-se a seguir os resultados de projeções e medidas de acurácia para os quatro modelos.



Para o Brasil verifica-se que a série tem dois momentos de maior crescimento, entre os anos 1960-1972 e entre 1996-2008.







Medidas de ajuste nos modelos - Brasil

	MAPE	MAD	MSD
LI - BR	0,559606	0,355018	0,161604
QD - BR	0,554675	**0,351266**	0,157941

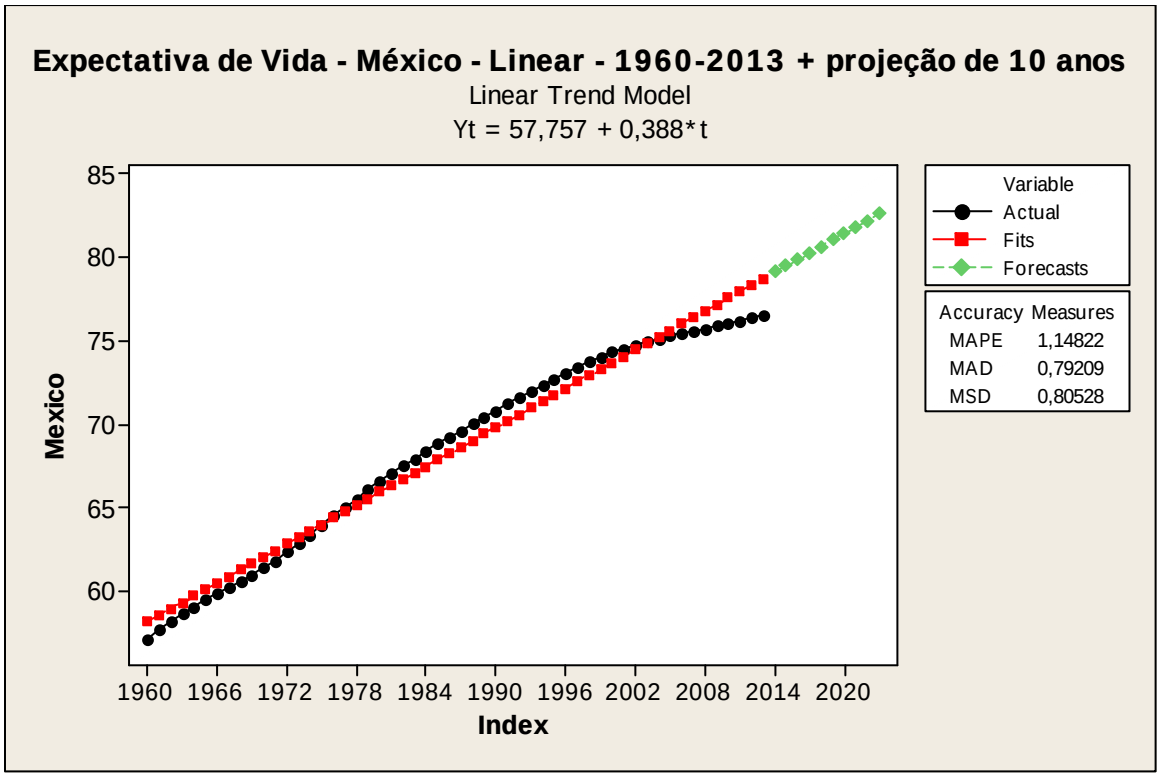
XP - BR	0,572328	0,361786	0,186206
S - BR	0,631712	0,400378	0,266623

Legenda: MAPE: Mean Absolute Percent Error; MAD: Median Average Deviation; MSD: Median Standard Deviation.

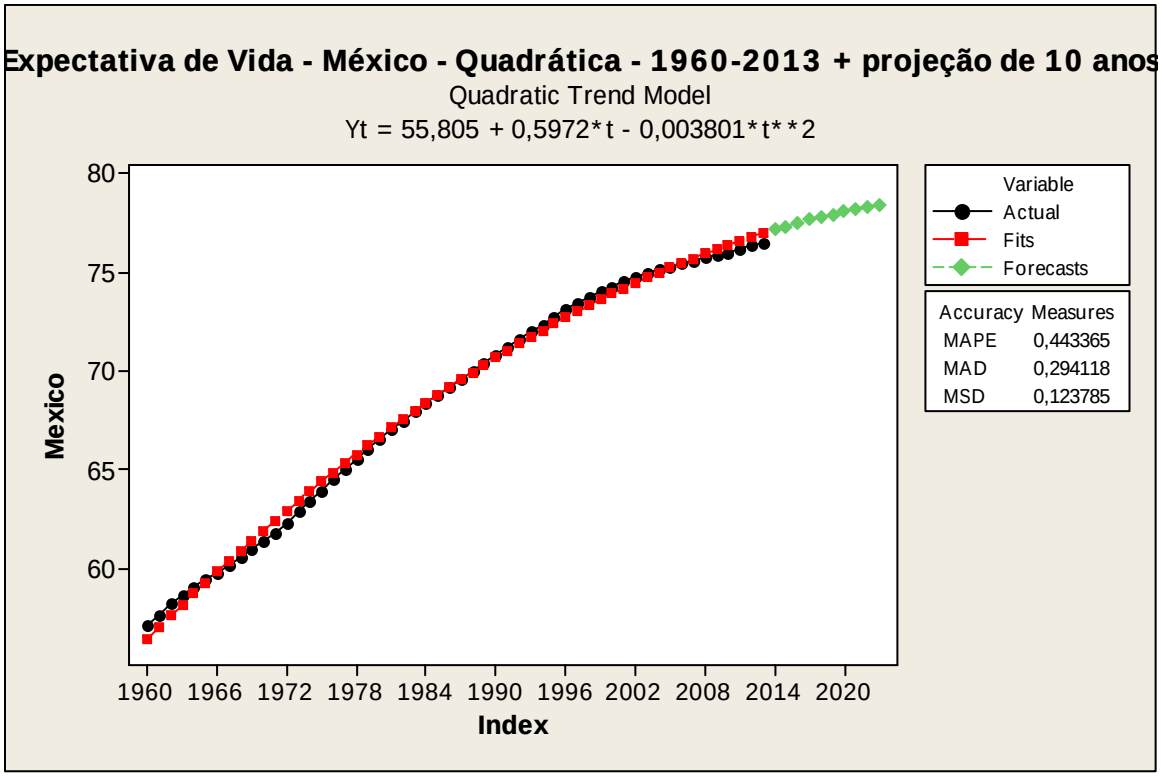
Verifica-se pelo gráfico e tabelas sobre as Medidas de ajuste nos modelos que para os dados do Brasil, aquele que possui a menor MAD – Mediana do desvio médio, é a função quadrátca, que então seria a melhor representação da curva. A projeção da mesma indica para os dez anos seguintes à série de dados as seguintes projeções:

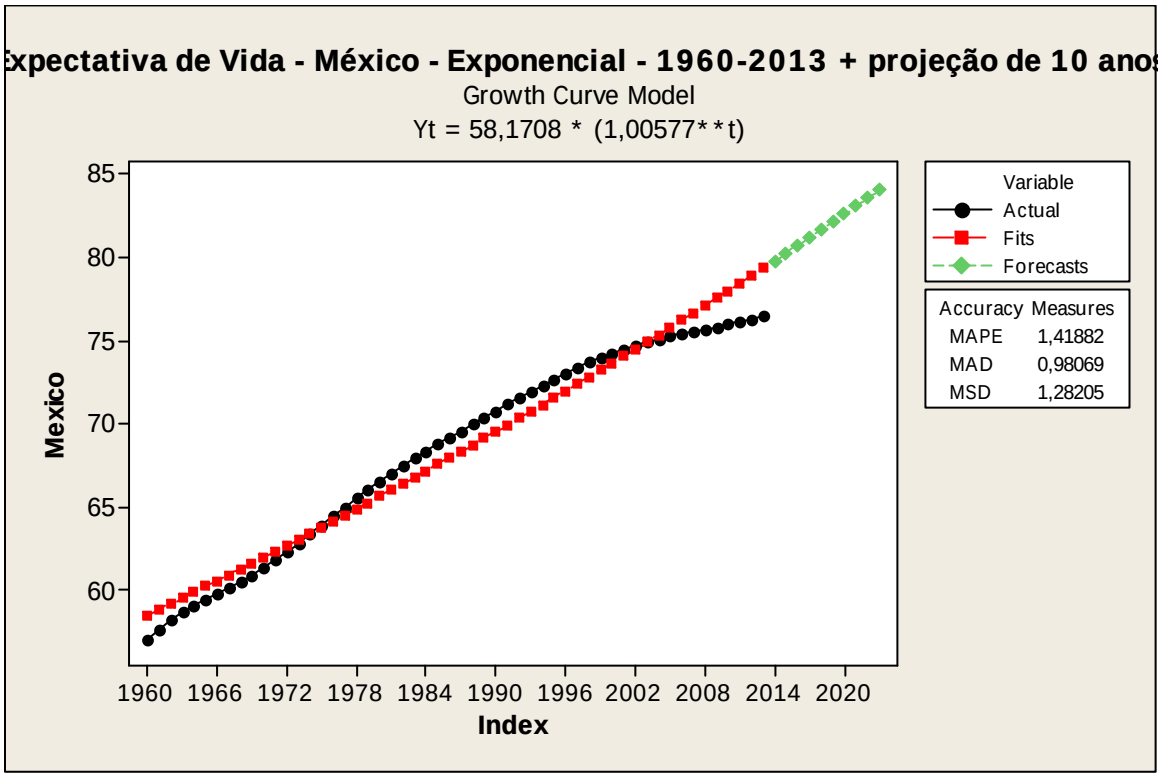
```
Period  Forecast
2014     74,9296
2015     75,3146
2016     75,7001
2017     76,0861
2018     76,4728
2019     76,8599
2020     77,2477
2021     77,6360
2022     78,0248
2023     78,4142
```

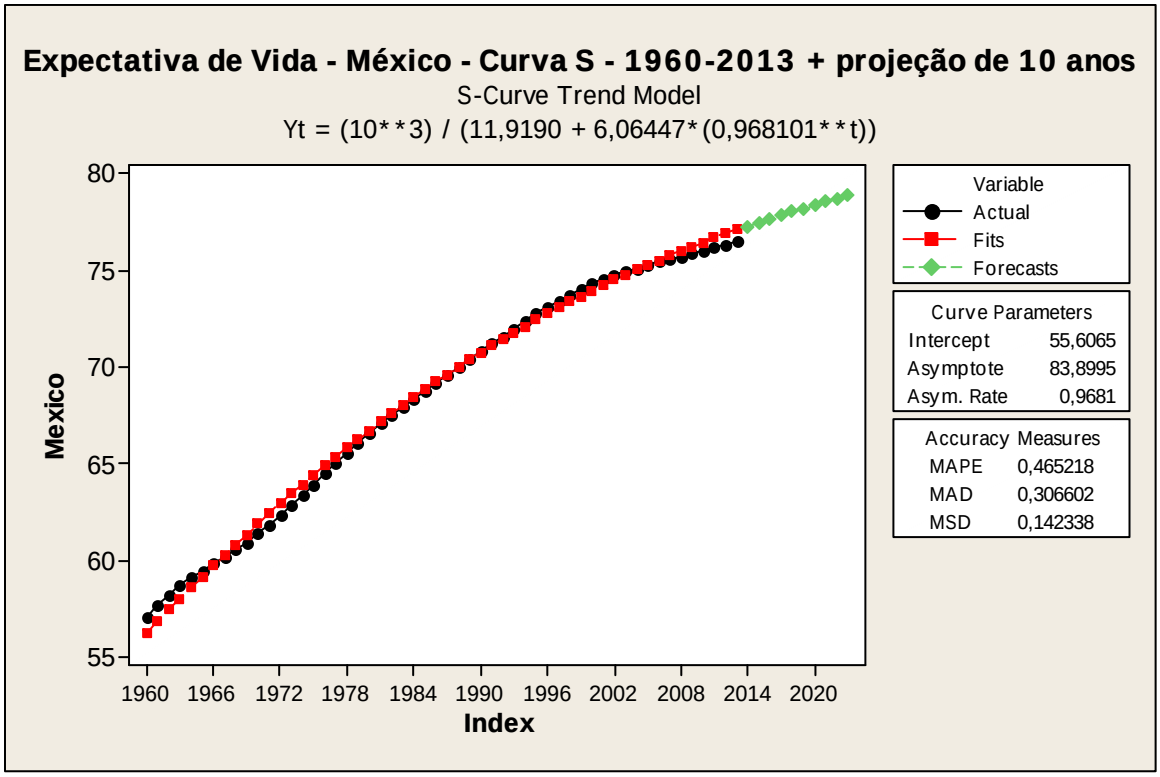
2. Análise de tendência e projeções - México



No México, ressalta-se a redução do crescimento a partir do ano 2000, após um crescimento significativo entre 1978 e 202.


Expectativa de Vida - México - Quadrática - 1960-2013 + projeção de 10 anos
Quadratic Trend Model
Yt = 55,805 + 0,5972*t - 0,003801*t**2
Mexico
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,443365
MAD 0,294118
MSD 0,123785


Expectativa de Vida - México - Exponencial - 1960-2013 + projeção de 10 anos
Growth Curve Model
Yt = 58,1708 * (1,00577**t)
Mexico
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 1,41882
MAD 0,98069
MSD 1,28205



Para a série de Expectativa de Vida ao Nascer para o México, a função quadrática apresenta o menor valor de desvio (MAD), sendo assim a função que melhor representa esta série de dados.

Medidas de ajuste nos modelos - México

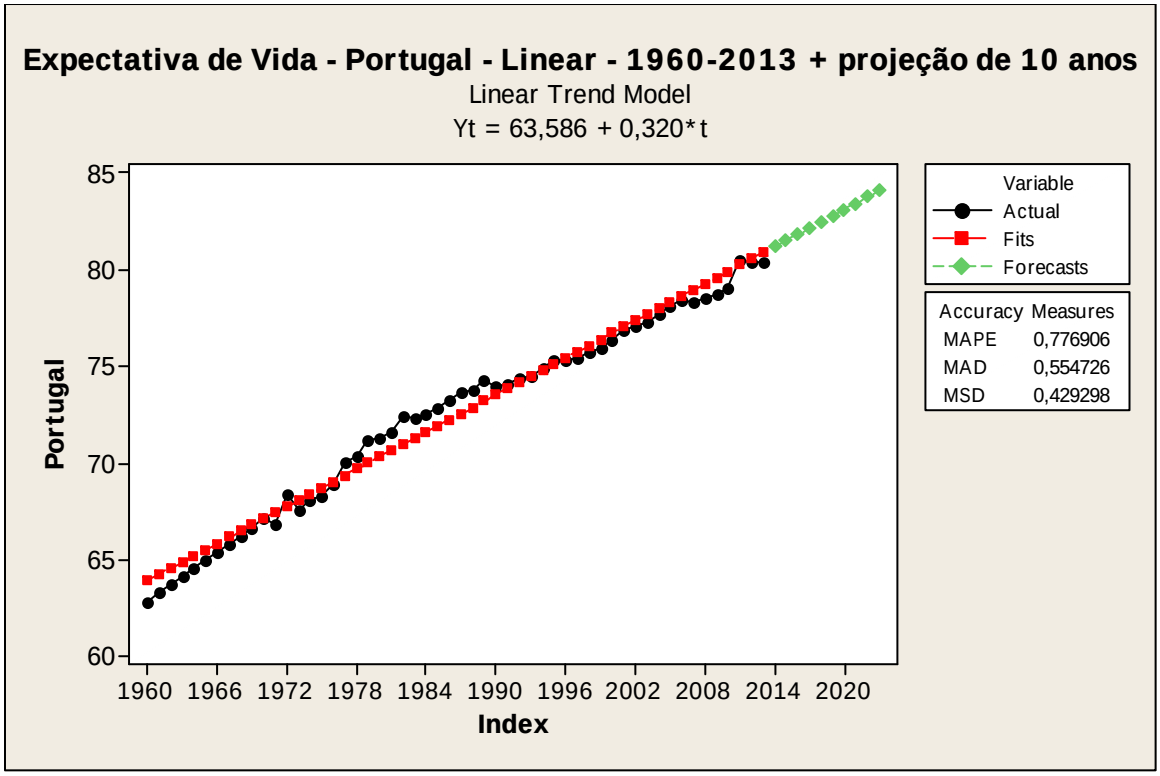
	LI - MX	QD - MX	XP - MX	S - MX
MAPE	1,14822	0,443365	1,41882	0,465218
MAD	0,79209	**0,294118**	0,98069	0,306602
MSD	0,80528	0,123785	1,28205	0,142338

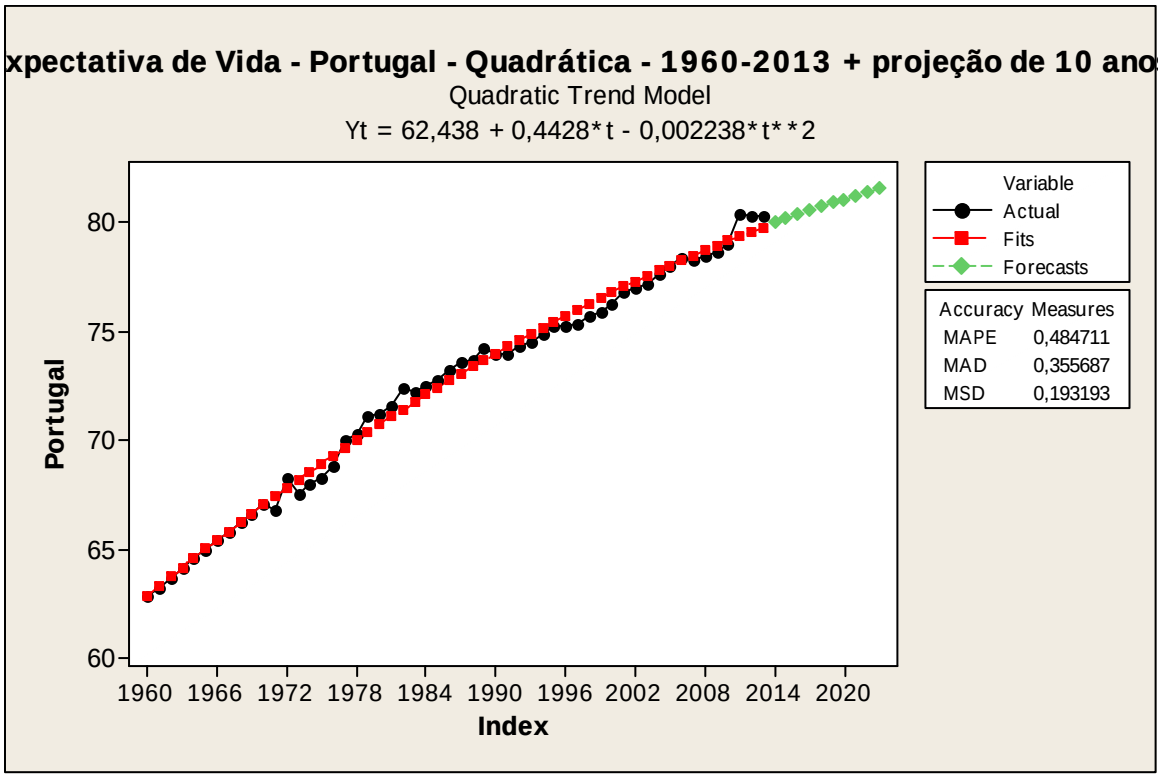
Abaixo apresenta-se a projeção de dados para o período de 10 anos após a série da dados, para a função quadrática mencionada acima para o México.

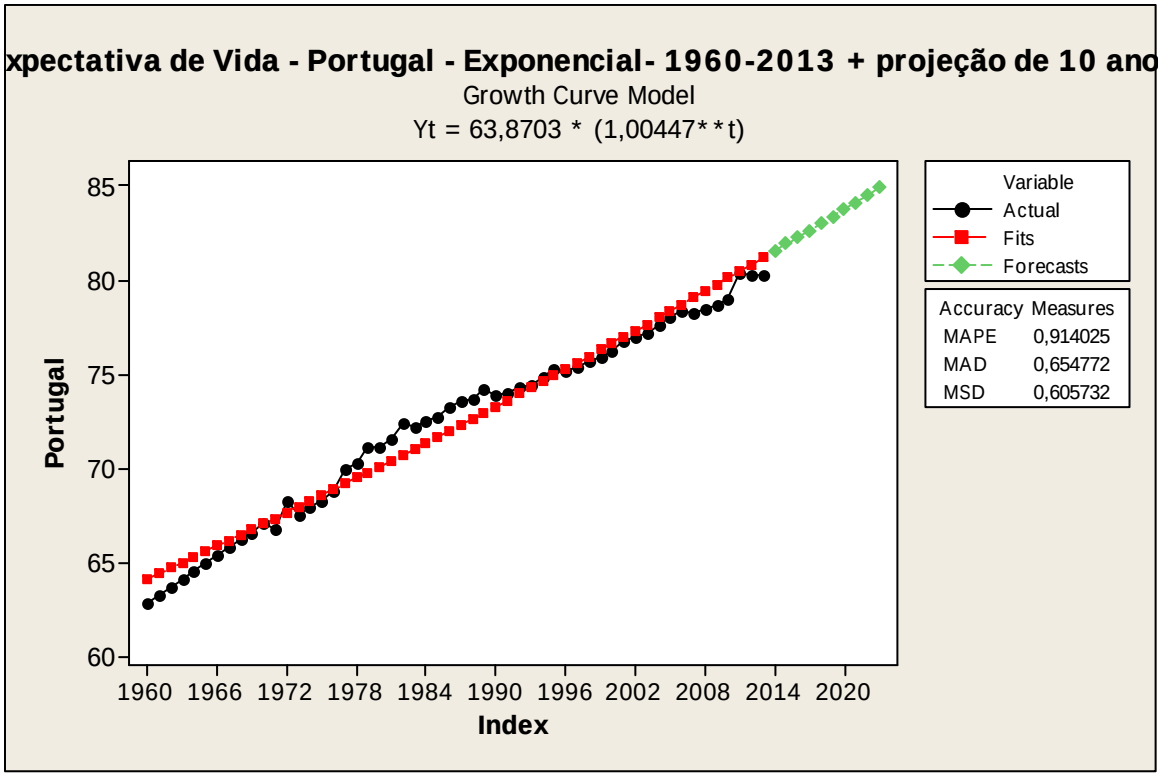
Period	Forecast
2014	77,1502
2015	77,3254
2016	77,4930
2017	77,6530
2018	77,8054
2019	77,9502
2020	78,0874
2021	78,2169
2022	78,3389
2023	78,4533

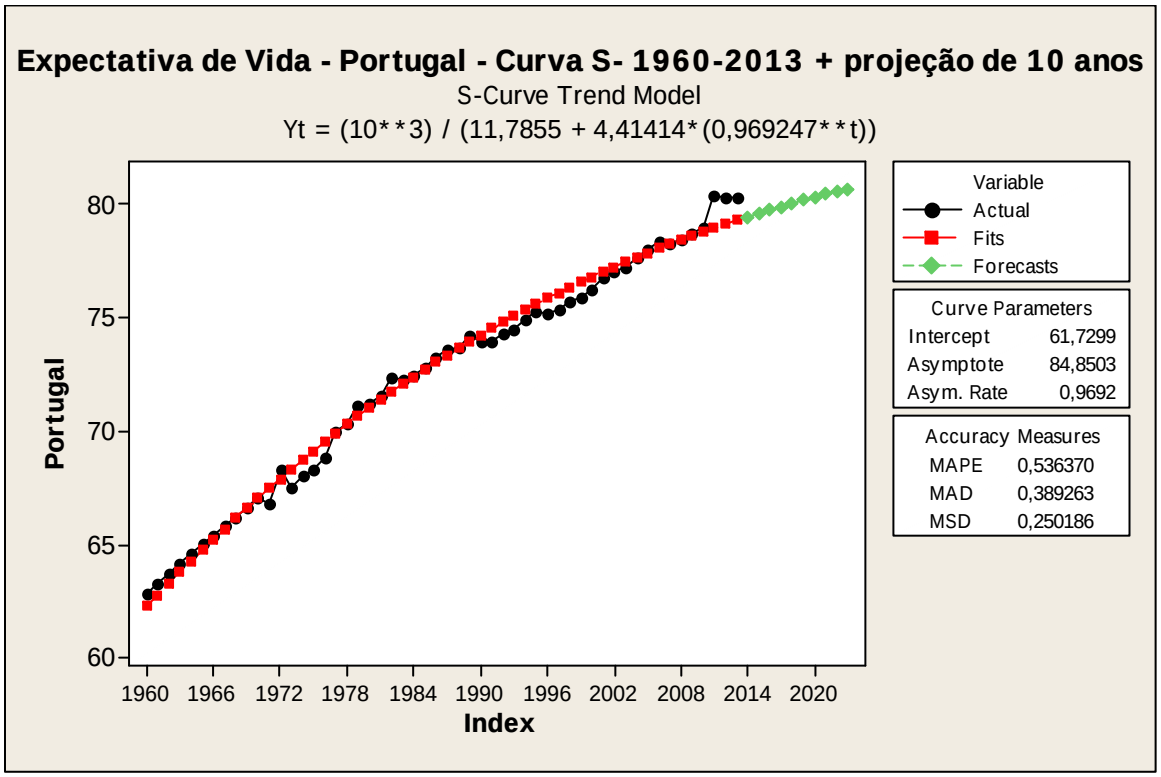
3. Análise de tendência e projeções - Portugal

Para Portugal a série apresenta um crescimento entre os anos 1970-1990.
Verificamos a seguir a análise dos modelos para esta série.









A análise dos modelos indica que a função quadrátiva é a que melhor representa os dados da série de dados temporais de Expectativa de vida ao nascer para Portugal, conforme dados a seguir.

Medida de ajuste nos modelos – Portugal

	LI - PT	QD - PT	XP - PT	S - PT
MAPE	0,776906	0,484711	0,914025	0,536370
MAD	0,554726	0,355687	0,654772	0,389263
MSD	0,429298	0,193193	0,605732	0,250186

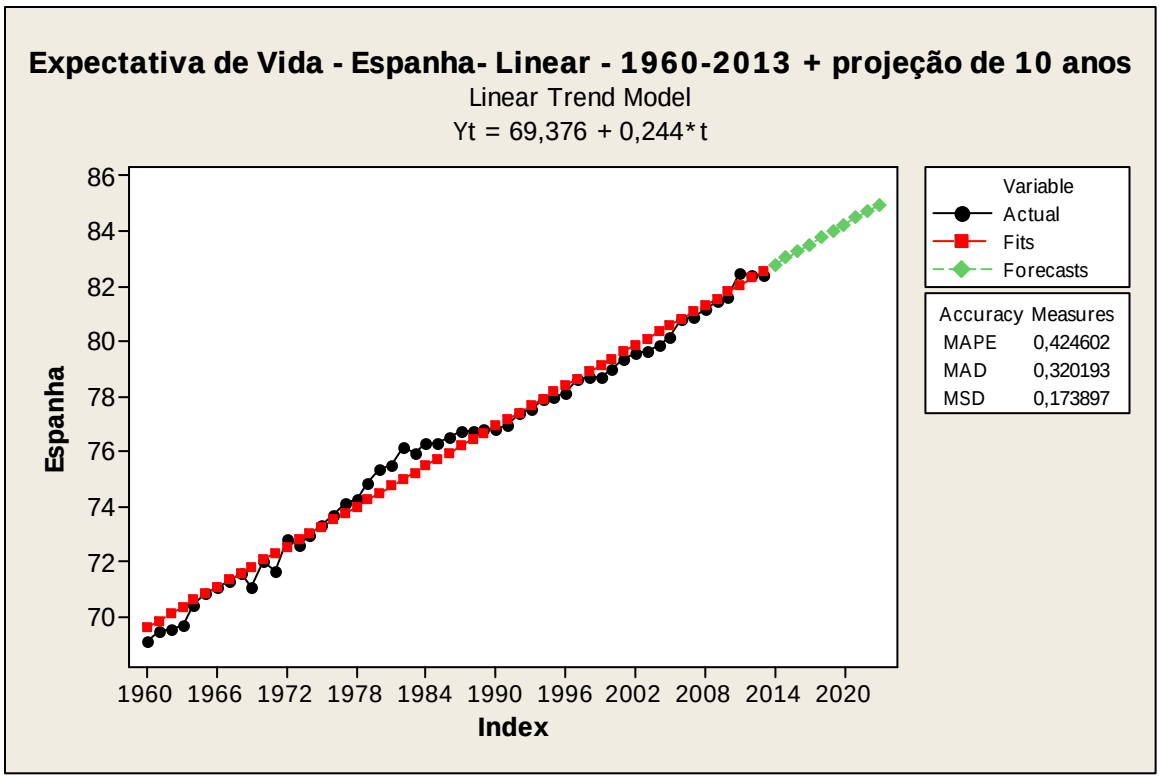
A série correspondente de projeções para 2014-2023 é apresentada:

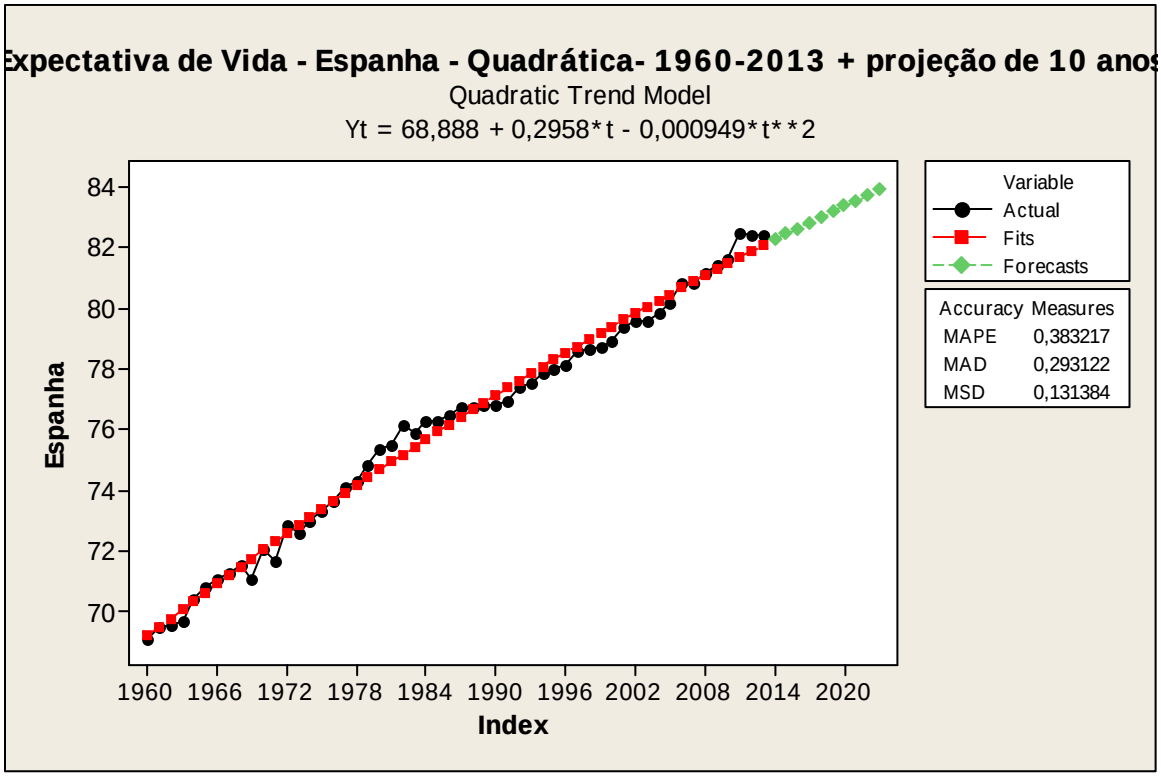
Forecasts

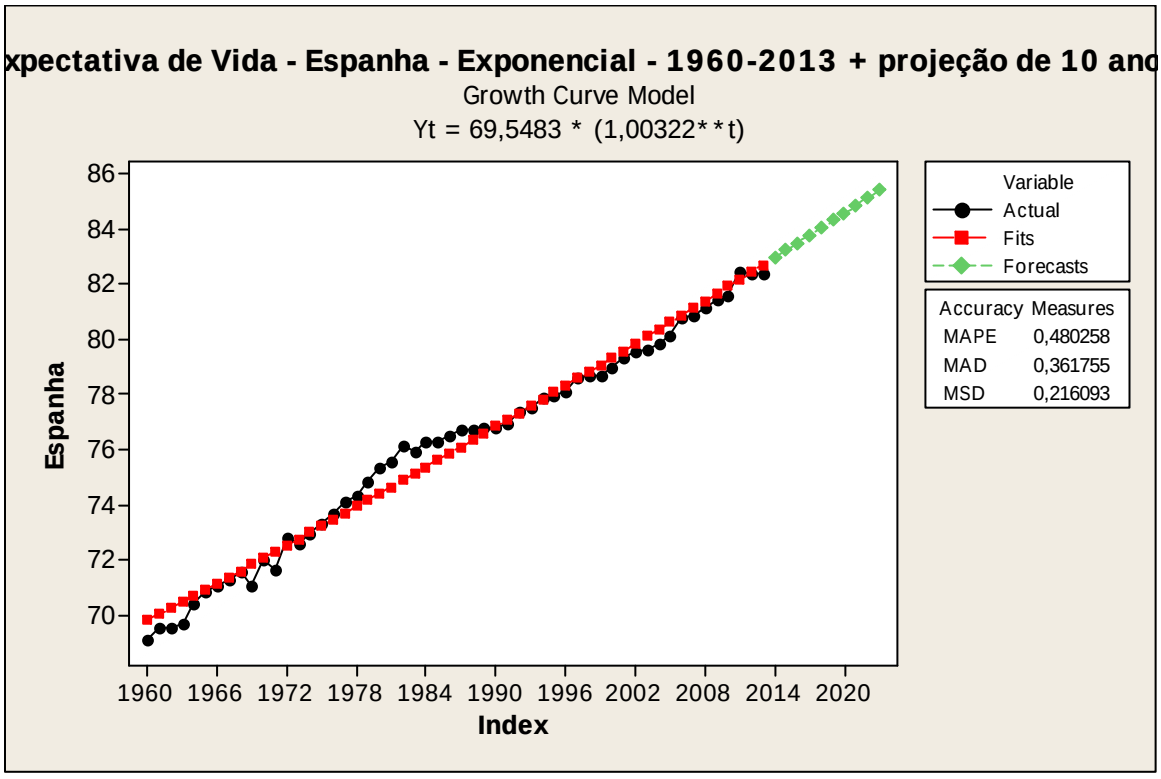
Period	Forecast
2014	80,0210
2015	80,2154
2016	80,4053
2017	80,5907
2018	80,7717
2019	80,9482
2020	81,1202
2021	81,2877
2022	81,4508
2023	81,6094

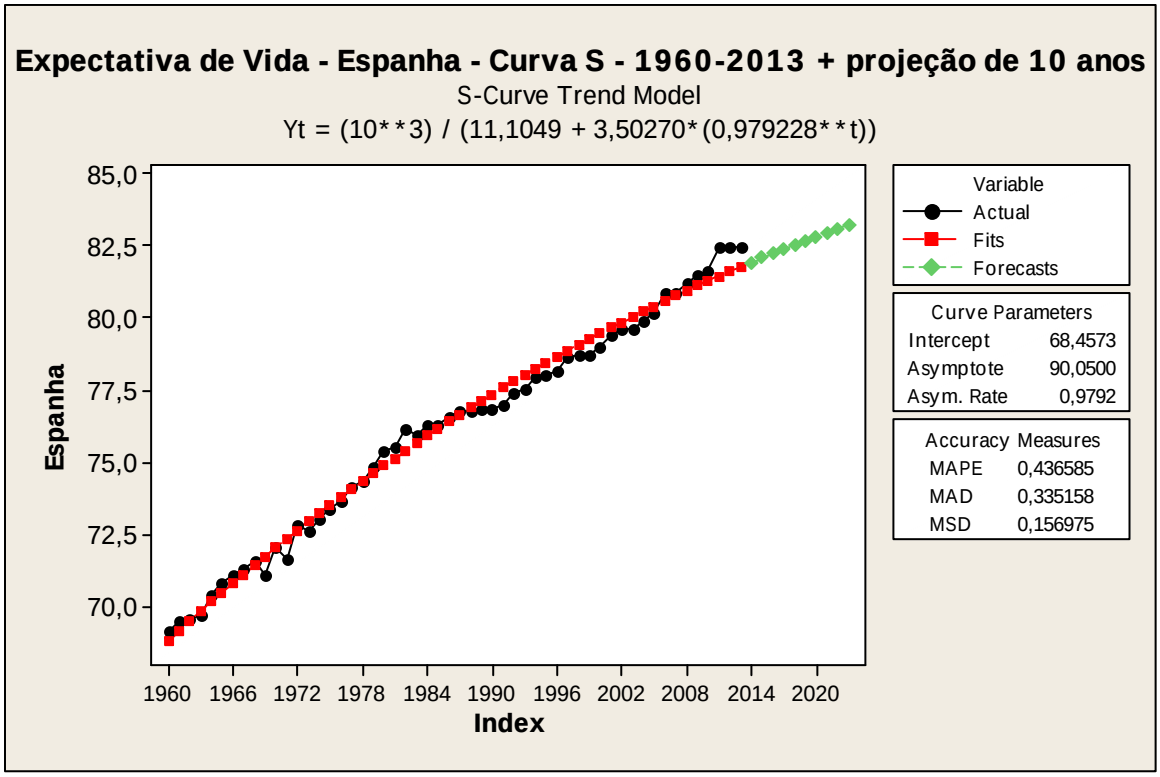
4. Análise de tendência e projeções - Portugal

Para Espenha, assim como notado em Portugal, a série apresenta um crescimento entre os anos 1970-1990 mais acentuado que nos demais períodos observados.
Verificamos a seguir a análise dos modelos para esta série.




Expectativa de Vida - Espanha - Quadrática- 1960-2013 + projeção de 10 anos
Quadratic Trend Model
Yt = 68,888 + 0,2958*t - 0,000949*t**2
Espanha
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,383217
MAD 0,293122
MSD 0,131384


xpectativa de Vida - Espanha - Exponencial - 1960-2013 + projeção de 10 ano
Growth Curve Model
Yt = 69,5483 * (1,00322**t)
Espanha
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,480258
MAD 0,361755
MSD 0,216093



A série tratada com uma função quadrática novamente é a que melhor representa a série de dados, neste caso a da Espanha.

Medida de ajuste nos modelos – Espanha

LI - ES		QD - ES		XP - ES		2 - ES
MAPE	0,424602		0,383217	MAPE	0,436585	
MAD	0,320193		0,293122	MAD	0,335158	
MSD	0,173897		0,131384	MSD	0,156975	

Seguem as projeções correspondentes à função quadrática.

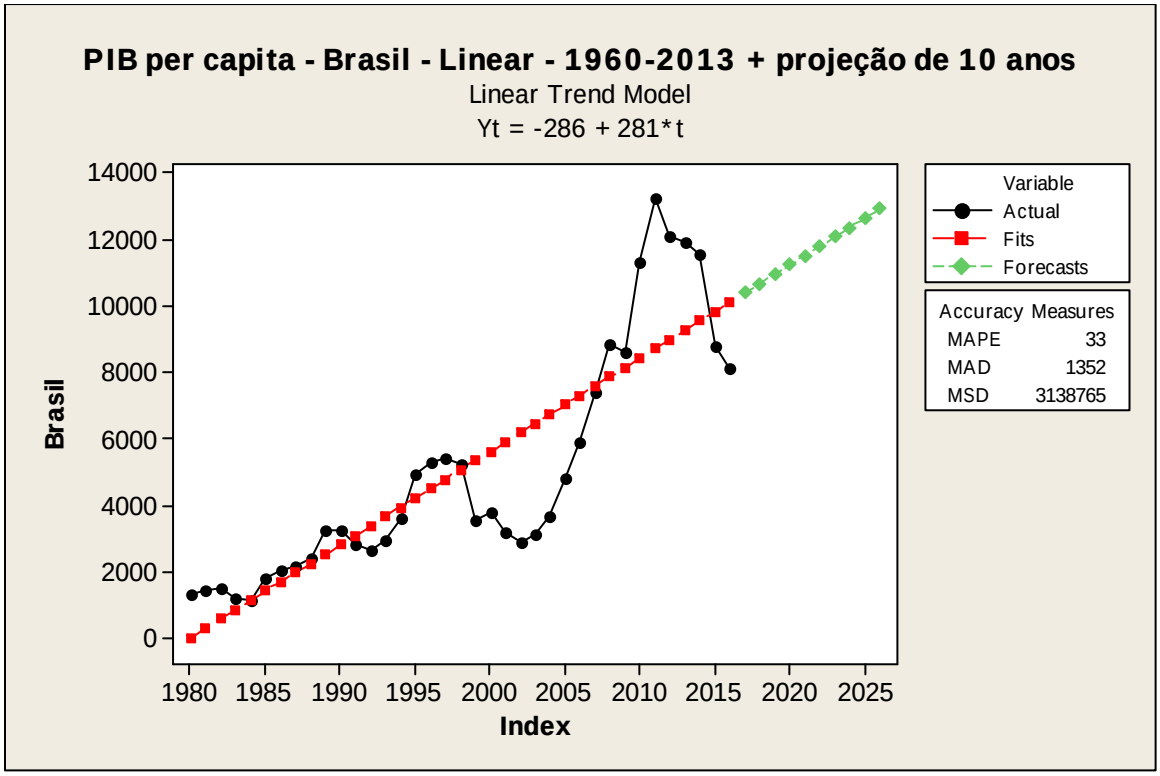
```
Forecasts

Period  Forecast
2014     82,2875
2015     82,4779
2016     82,6665
2017     82,8531
2018     83,0379
2019     83,2207
2020     83,4017
2021     83,5808
2022     83,7579
2023     83,9332
```

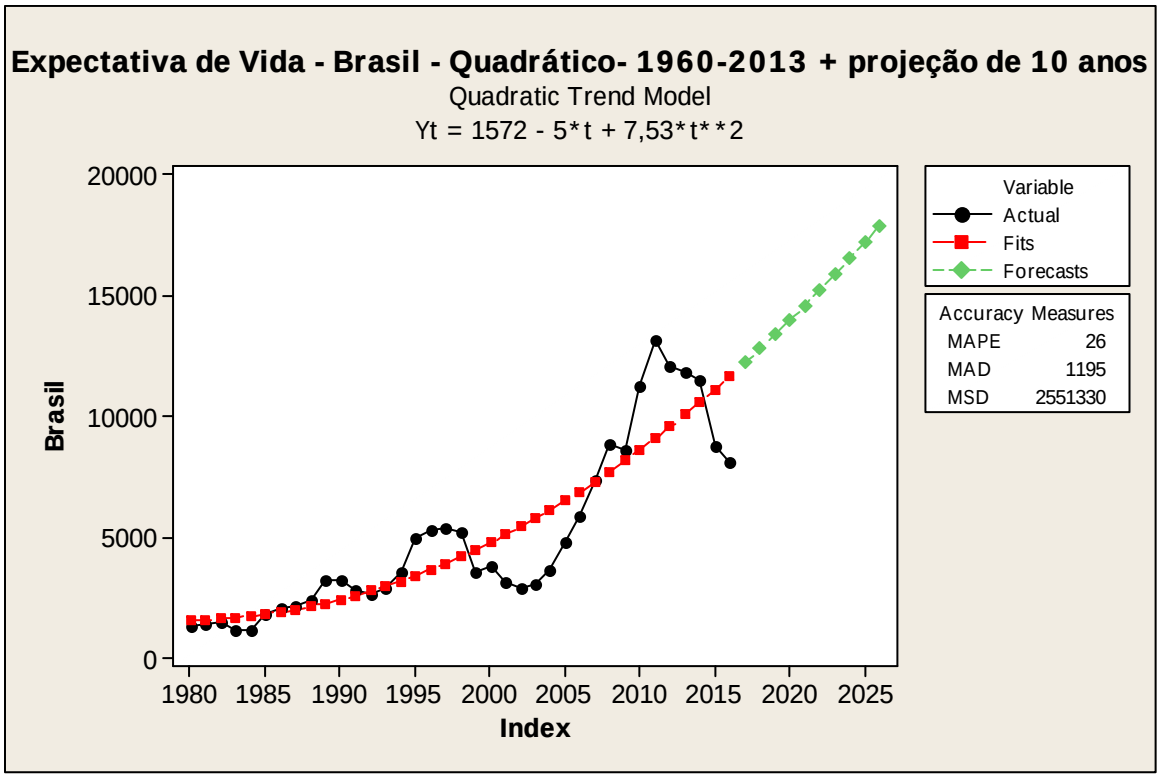
ii. PIB per capita em USD constantes

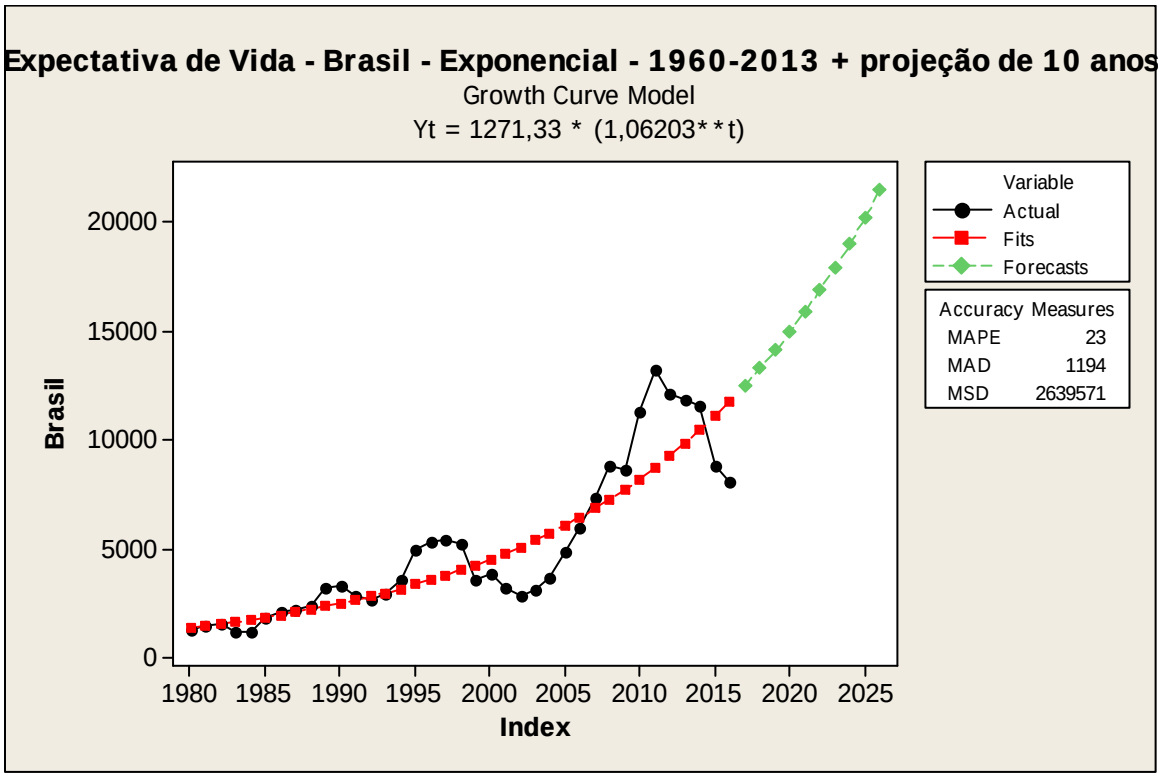
A variável PIB per capita em USD constantes foi obtida na base de dados do site www.knoema.org, e é originada das bases do Fundo Monetário Internacional (FMI – IMF), em seu relatório “World Economic Outlook (WEO), October 2015”.

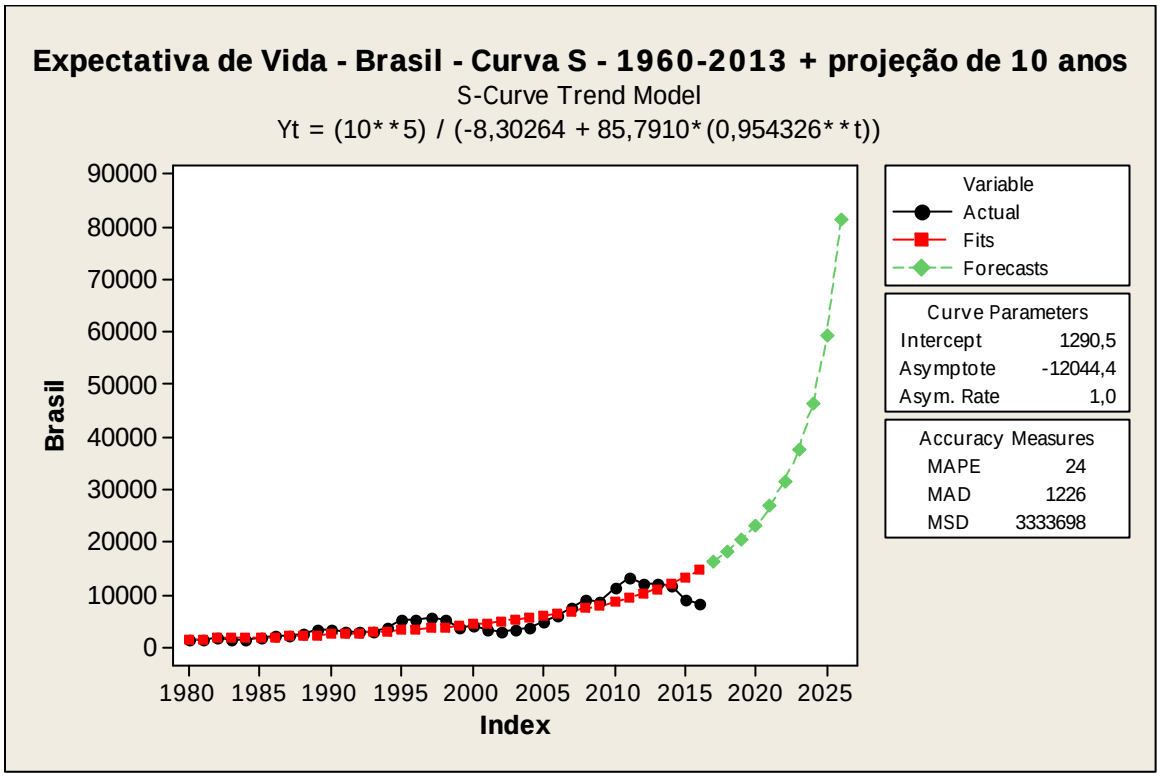
A série contempla dados de 1980 a 2016 (34 observações), para Brasil, México, Portugal e Espanha.

1. Análise de Tendências e Projeções – Brasil









Ao compararmos no quadro a seguir as medidas de ajuste nos quatro modelos para a variável Renda per capita em USD correntes observa-se que a menor mediana de desvio médio(MAD) é encontrada no modelo exponencial, que então é selecionado como o que melhor represneta esta variável para esta série de dados.

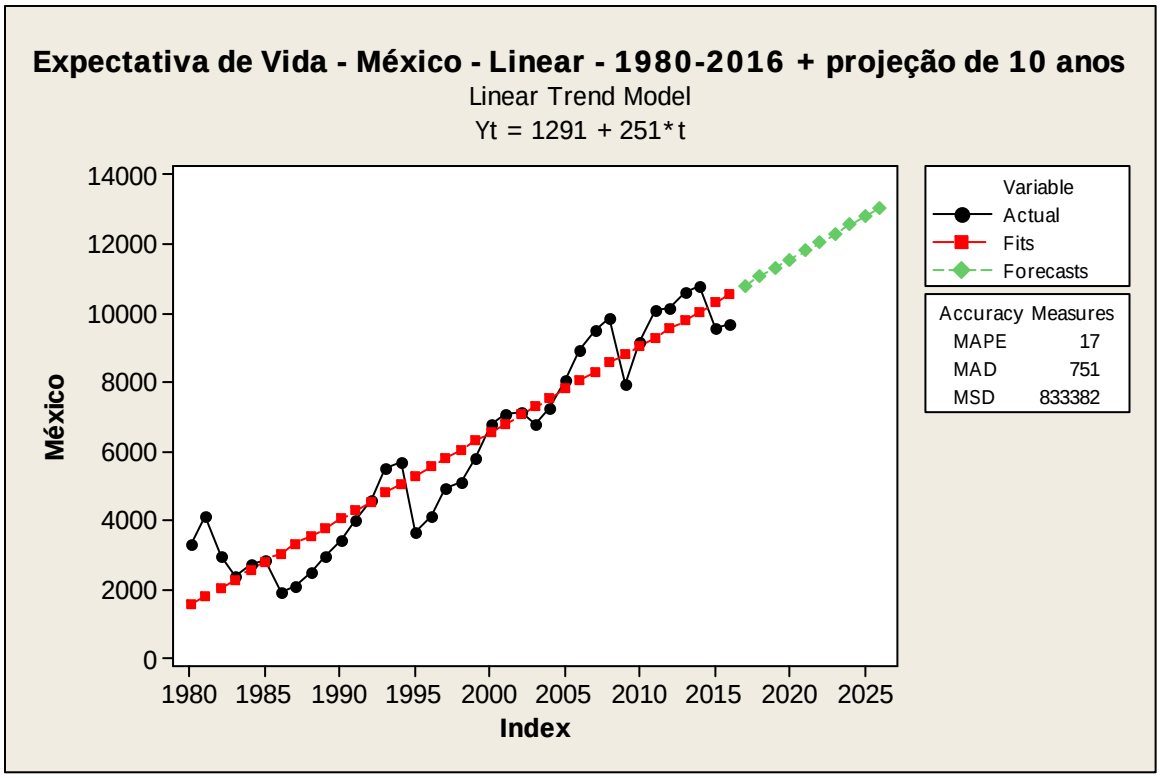
Medida de ajuste nos modelos – Brasil

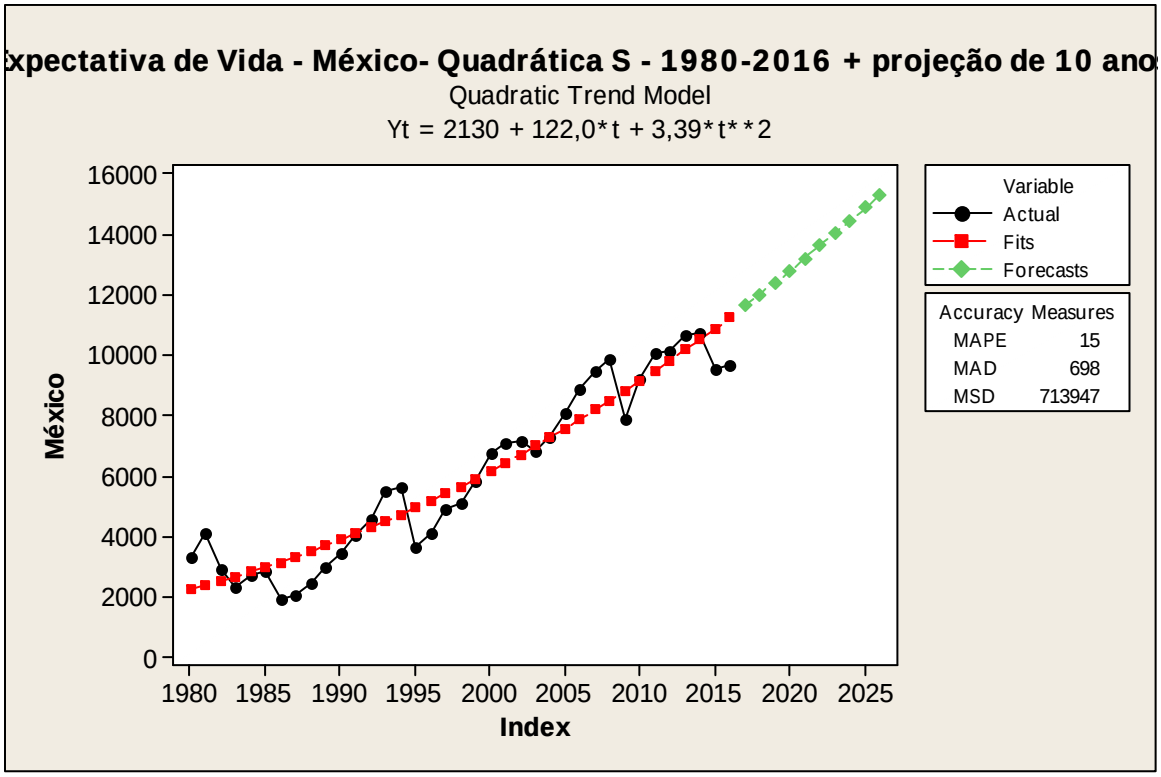
	BR - LI	BR - QD	BR - XP	BR - S
MAPE	33	26	23	24
MAD	1352	1195	**1194**	1226
MSD	3138765	2551330	2639571	3333698

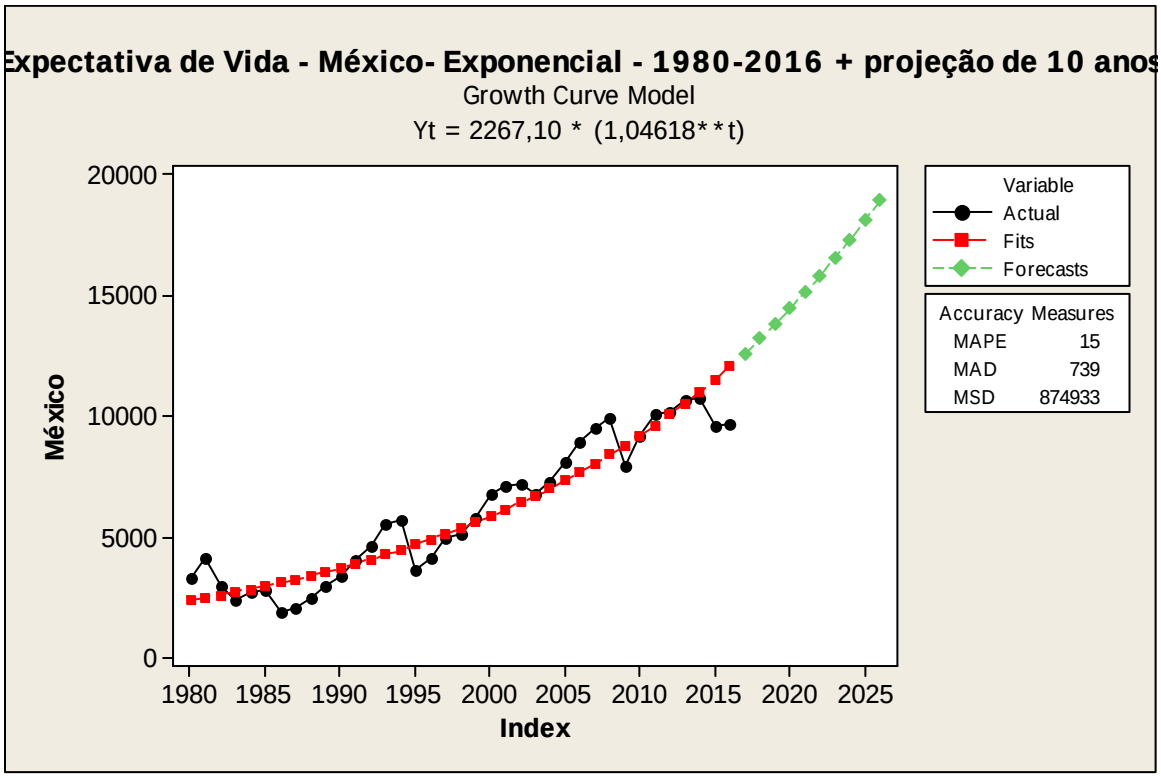
As projeções para a função Exponencial da Renda per capital em USD correntes para o Brasil, para os dez anos à frente da série, 2017 – 2026, são apresentadas a seguir.

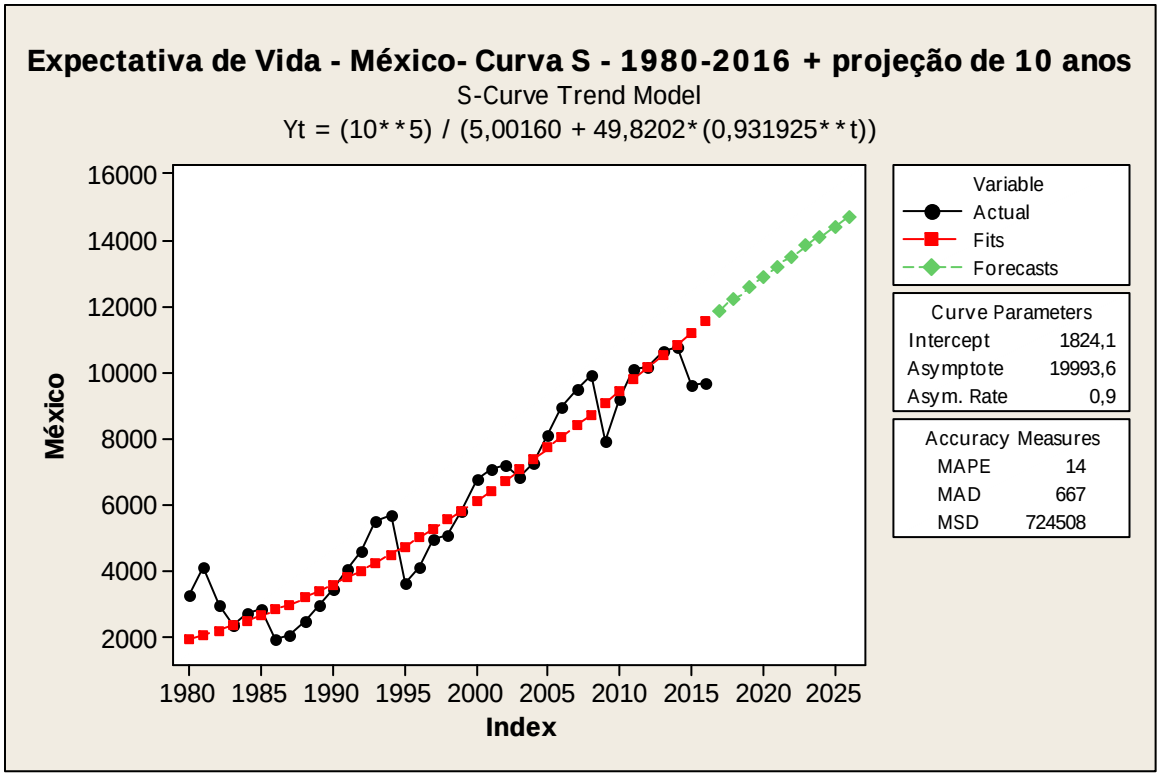
```
Forecasts

Period  Forecast
2017     12516,9
2018     13293,3
2019     14117,9
2020     14993,7
2021     15923,8
2022     16911,6
2023     17960,6
2024     19074,8
2025     20258,0
2026     21514,7
```

2. Análise de Tendências e Projeções – México




Expectativa de Vida - México- Quadrática S - 1980-2016 + projeção de 10 anos
Quadratic Trend Model
Yt = 2130 + 122,0*t + 3,39*t**2
México
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 15
MAD 698
MSD 713947


Expectativa de Vida - México- Exponencial - 1980-2016 + projeção de 10 anos
Growth Curve Model
Yt = 2267,10 * (1,04618**t)
México
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 15
MAD 739
MSD 874933



Para a variável analisada na série temporal do México, a função de Curva S é a que possui menor erro, e portanto assume-se como a que melhor representa as variáveis e permitiria melhor projeção. Os dados de erro comparativos são apresentados a seguir.

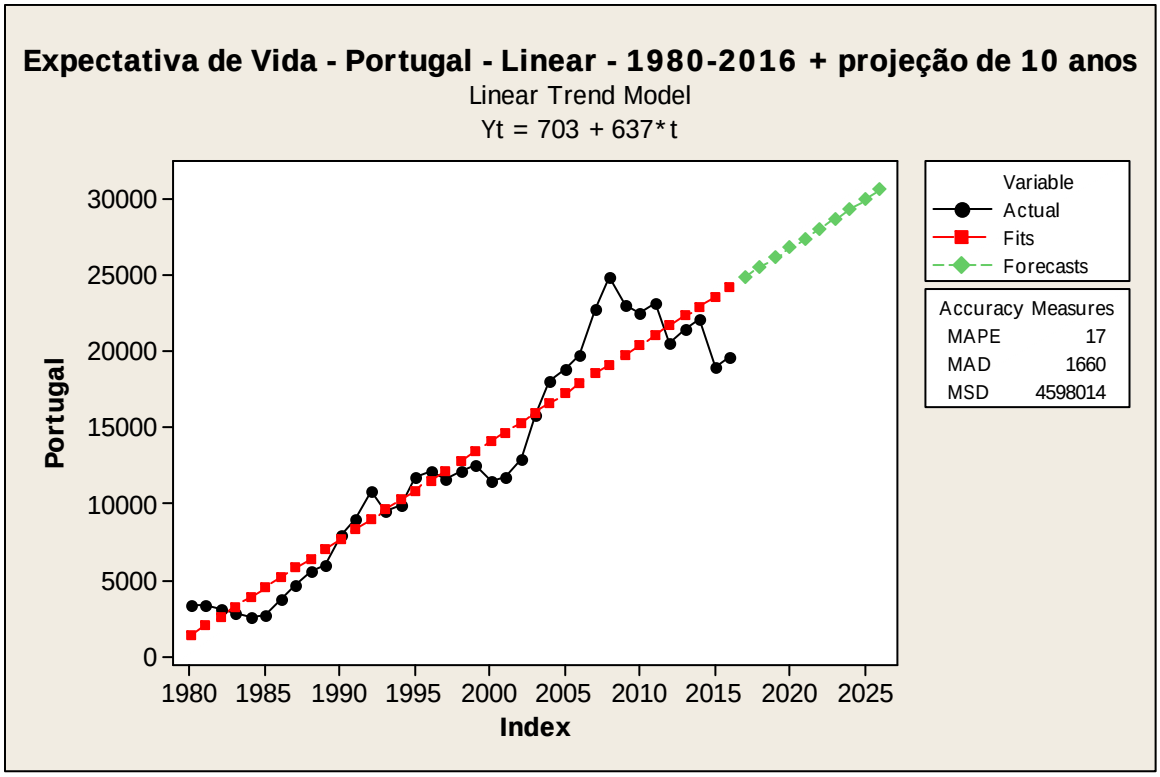
Medida de ajuste nos modelos – México

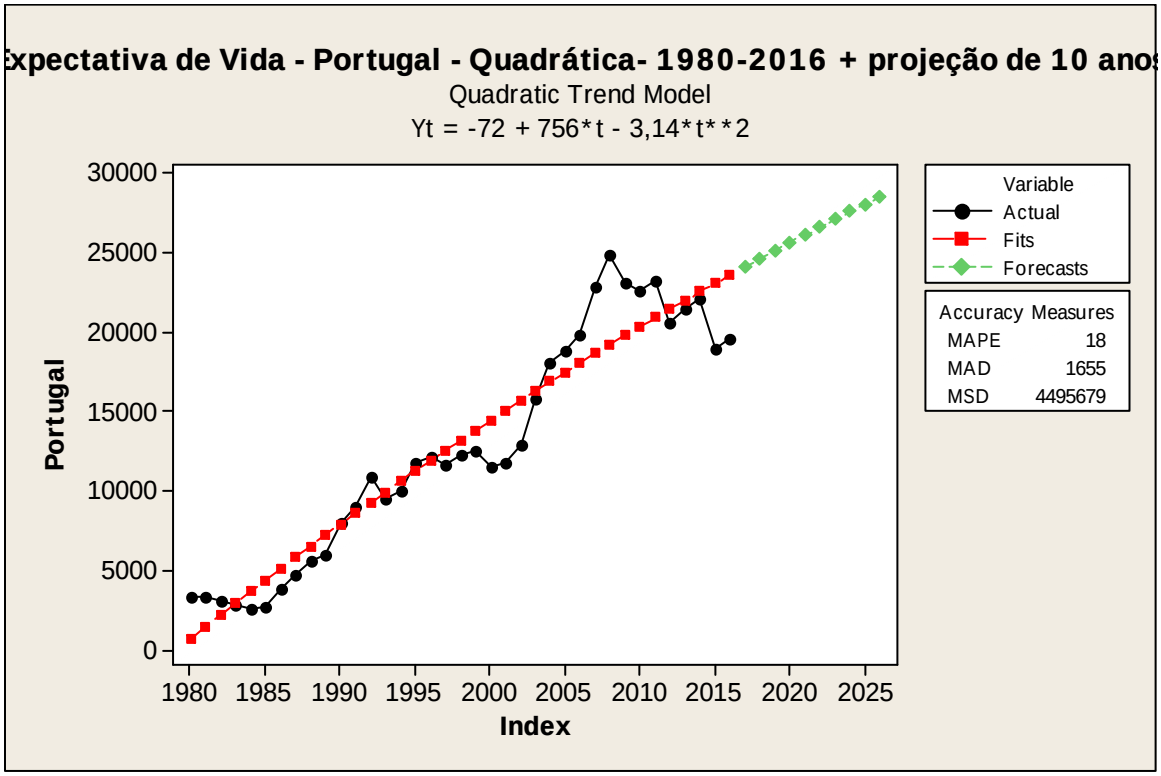
	MX - LI	MX - QD	MX - XP	MX - S
MAPE	17	15	15	14
MAD	751	698	739	**667**
MSD	833382	713947	874933	724508

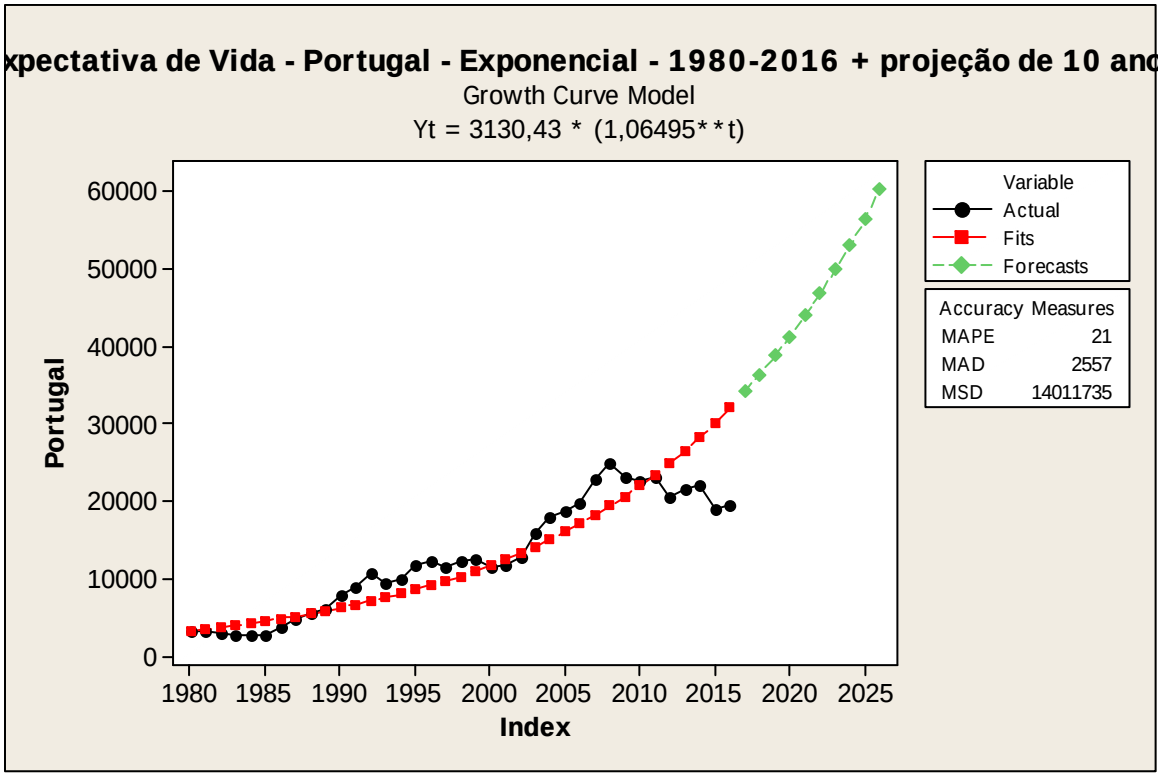
As projeções de dez anos geradas pela função de curva S para o PIB per capita do México a partir de 2017 são apresentadas a seguir.

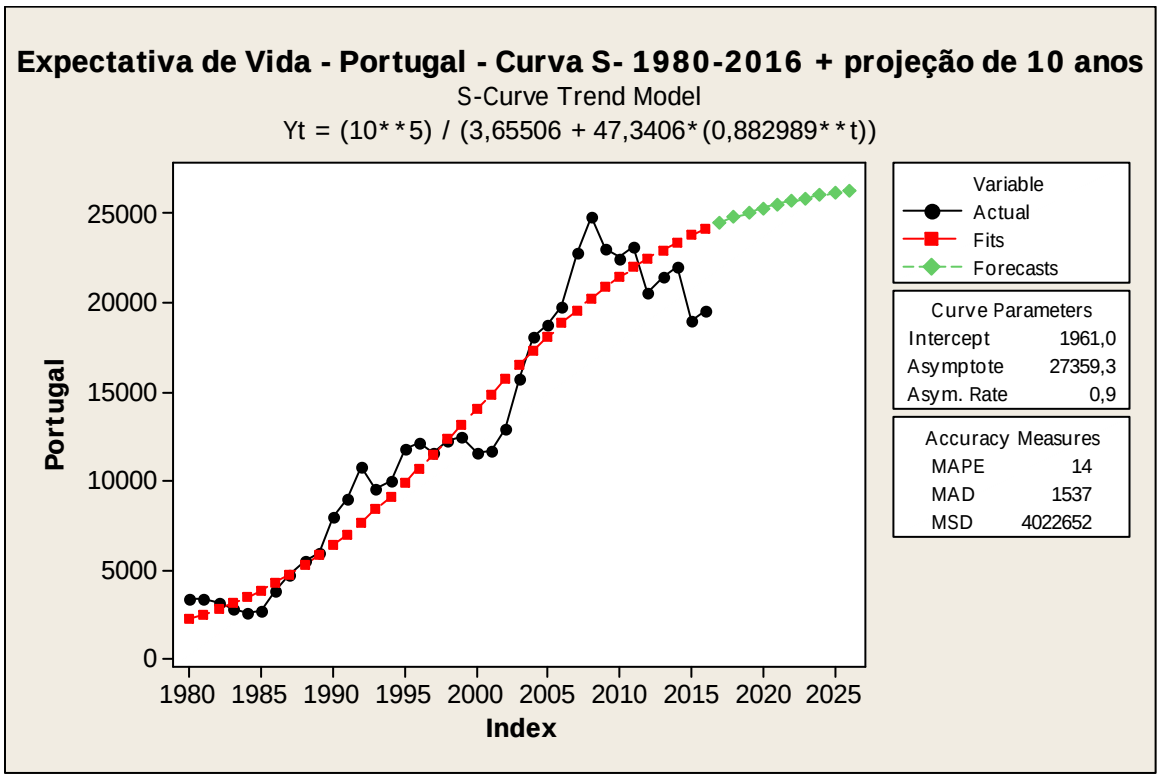
Period	Forecast
2017	11875,8
2018	12213,4
2019	12545,7
2020	12872,2
2021	13192,0
2022	13504,8
2023	13809,9
2024	14106,9
2025	14395,4
2026	14675,1

3. Análise de Tendências e Projeções – Portugal


Expectativa de Vida - Portugal - Linear - 1980-2016 + projeção de 10 anos
Linear Trend Model
Yt = 703 + 637*t
Portugal
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 17
MAD 1660
MSD 4598014


Expectativa de Vida - Portugal - Quadrática- 1980-2016 + projeção de 10 anos
Quadratic Trend Model
Yt = -72 + 756*t - 3,14*t**2
Portugal
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 18
MAD 1655
MSD 4495679





Para a variável analisada na série temporal de Portugal, a função de Curva S é a que possui menor erro, e portanto assume-se como a que melhor representa as variáveis e permitiria melhor projeção. Os dados de erro comparativos são apresentados a seguir.

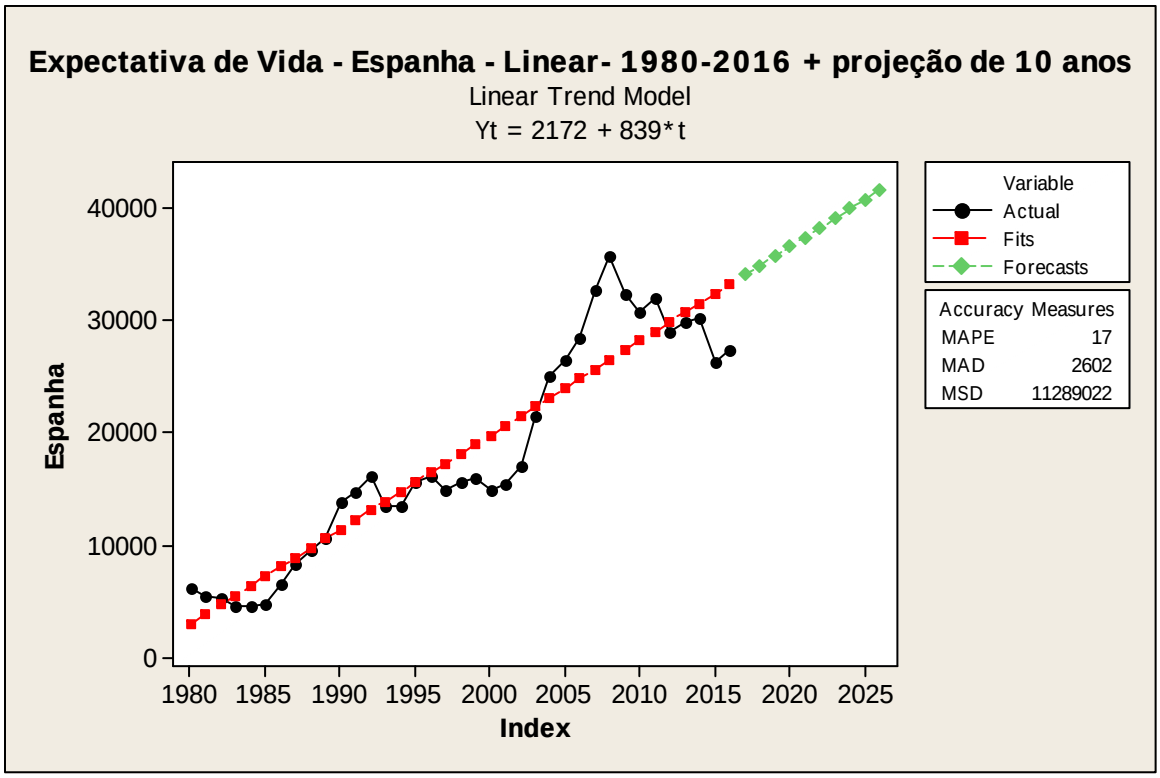
Medida de ajuste nos modelos – Portugal

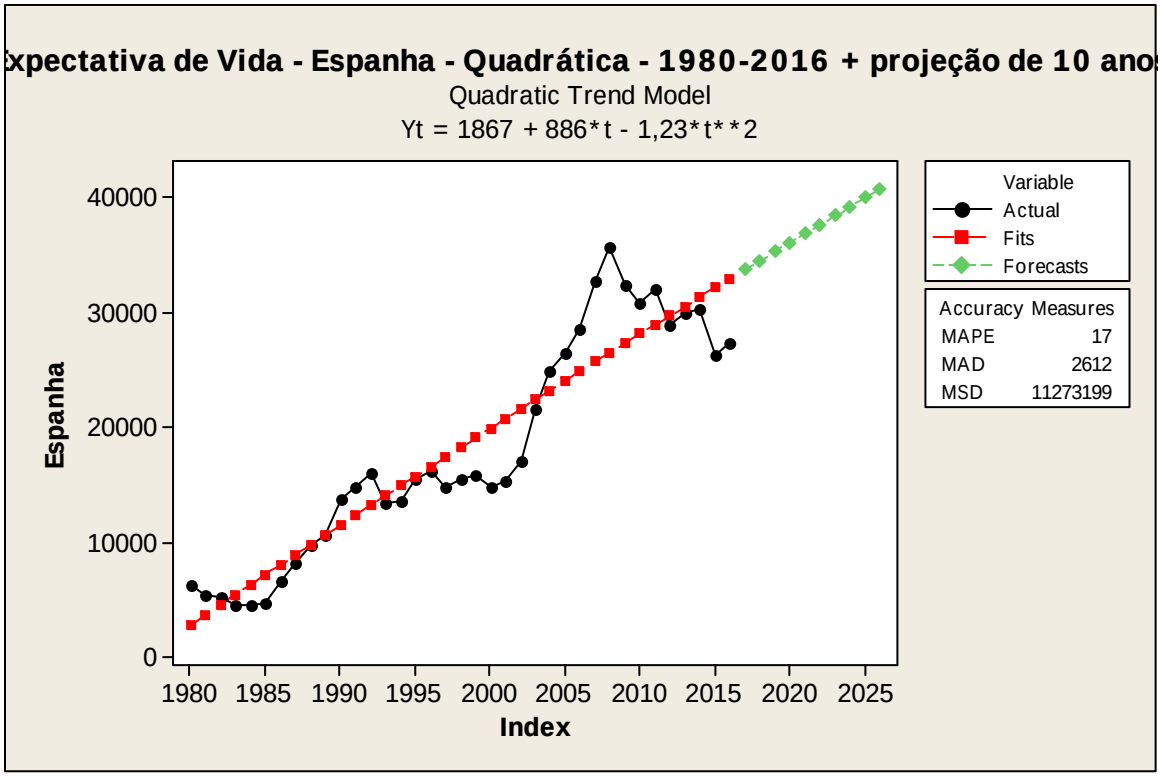
	PT - LI	PT - QD	PT - XP	PT - S
MAPE	17	18	21	14
MAD	1660	1655	2557	**1537**
MSD	4598014	4495679	14011735	4022652

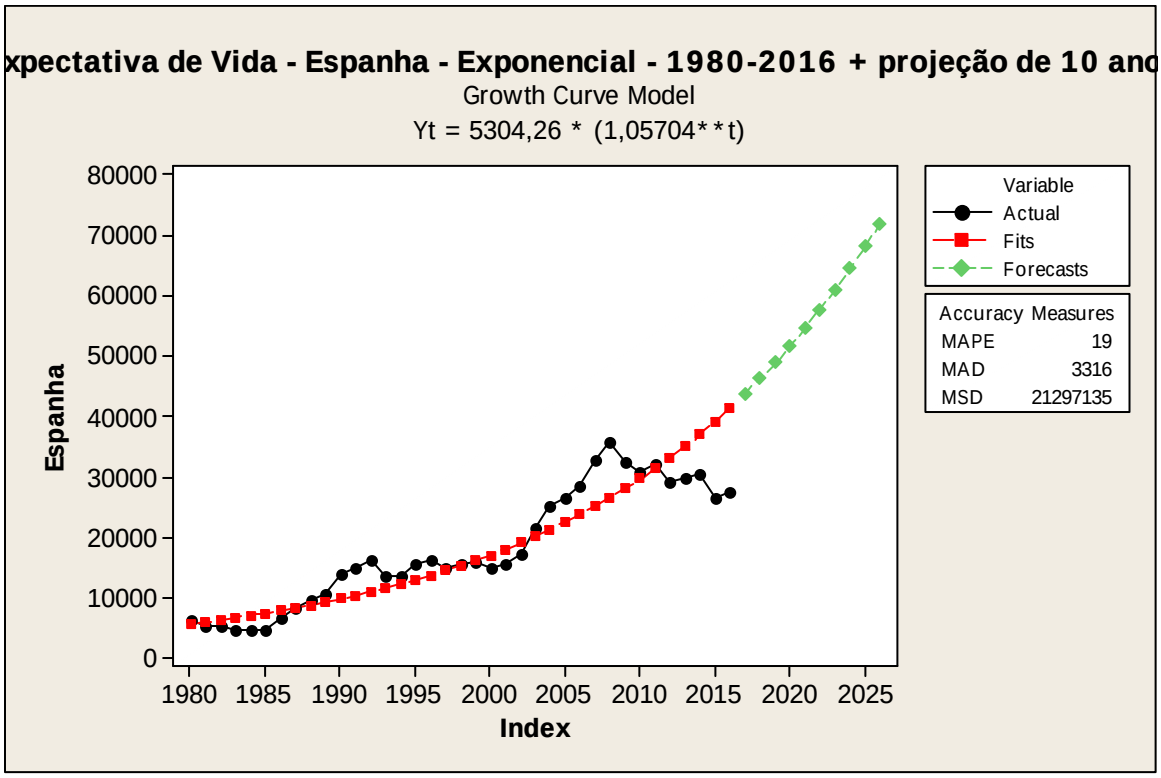
Abaixo as projeções de renda percapita geradas pelo modelo de Cruva S:

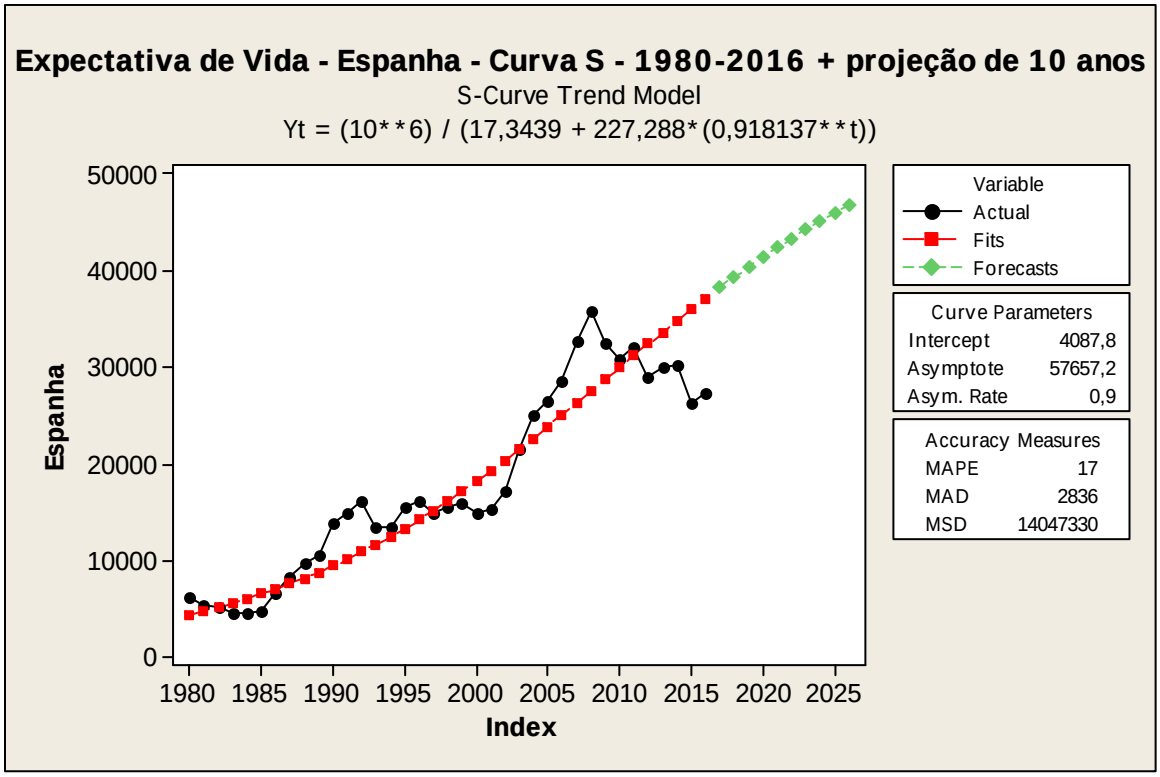
```
Forecasts

Period  Forecast
2017     24549,5
2018     24848,1
2019     25117,8
2020     25361,0
2021     25579,6
2022     25775,8
2023     25951,5
2024     26108,7
2025     26249,1
2026     26374,3
```

4. Análise de Tendências e Projeções – Espanha




xpectativa de Vida - Espanha - Quadrática - 1980-2016 + projeção de 10 anos
Quadratic Trend Model
Yt = 1867 + 886*t - 1,23*t**2
Espanha
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 17
MAD 2612
MSD 11273199


xpectativa de Vida - Espanha - Exponencial - 1980-2016 + projeção de 10 anos
Growth Curve Model
Yt = 5304,26 * (1,05704**t)
Espanha
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 19
MAD 3316
MSD 21297135



Para a variável analisada na série temporal com dados da Espanha, a função Linear é a que possui menor erro, e portanto assume-se como a que melhor representa as variáveis e permitiria melhor projeção. Os dados de erro comparativos são apresentados a seguir.

Medida de ajuste nos modelos – Espanha

	ES - LI	ES - QD	ES - XP	ES - S
MAPE	17	17	19	17
MAD	**2602**	2612	3316	2836
MSD	11289022	11273199	21297135	14047330

Os dados de projeção de dez anos desenvolvidos a pargir do modelo linear de menor erro são apresentados a seguir:

Period	Forecast
2017	34069,3
2018	34908,7
2019	35748,1
2020	36587,5
2021	37426,9
2022	38266,3
2023	39105,7
2024	39945,1
2025	40784,4
2026	41623,8

c. EIXO PLANETA

i. Breve descrição do EIXO PLANETA

Um dos cinco eixos centrais dos Objetivos do Desenvolvimento Sustentável – ODS, o EIXO PESSOAS tem como meta e visão “Acabar com a pobreza e a fome, em todas as usas formas e dimensões, e para assegurar que todos os seres humanos possam desenvolver todo seu potencial em dignidade e igualdade em um ambiente saudável.”

ii. Variáveis do EIXO PLANETA

Este tema da pesquisa apresenta dezessete variáveis, sendo oito índices sintéticos. Todas as variáveis são quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis componentes ou indicadores, a saber:

iii. A fonte e o tamanho da série de dados

Fonte:
A fonte empregada é o site www.knoema.org. Para as séries de Área Florestada como percentual da área total do país, a fonte de dados empregada utilizou-se de dados do banco mundial, World Development Indicators, de fevereiro de 2016.

Conforme o site knoema.org,

"FAO tem vindo a recolher e analisar dados sobre área florestal desde 1946. Isso é feito em intervalos de 5-10 anos como parte do Global Forest Resources Assessment (FRA). FAO relata dados para 229 países e territórios; para os restantes 56 pequenos Estados insulares e territórios onde nenhuma informação é fornecida, um relatório é preparado pela FAO utilizando a informação existente e uma pesquisa bibliográfica. Os dados são agregados a nível sub-regional, regional e global da equipe de FRA em FAO, e as estimativas são produzidas pelo somatório reta.

A defasagem entre o ano de referência e a produção efectiva de séries de dados, bem como a frequência da produção de dados varia entre os países. áreas desmatadas não incluem áreas exploradas, mas destina-se para a regeneração ou áreas degradadas pela recolha de lenha, precipitação ácida, ou incêndios florestais. Números negativos indicam um aumento na área florestal.

Os dados incluem superfícies com bambus e palmeiras; caminhos florestais, aceiros e outras zonas abertas de reduzida; floresta em parques nacionais, reservas naturais e outras áreas protegidas tais como as de interesse científico, histórico, cultural ou espiritual; quebra-ventos, abrigo e os corredores de árvores com uma área de mais de 0,5 hectares e largura superior a 20 metros; plantações usado principalmente para florestais ou de defesa fins, tais como plantações de árvore da borracha e de sobreiros. Dados exclui povoamentos em sistemas de produção agrícola, tais como plantações de frutas e sistemas agroflorestais. área de floresta também exclui árvores em parques urbanos e jardins. A proporção de área florestal para a área total do terreno é calculada e mudanças na proporção são computados para identificar tendências."

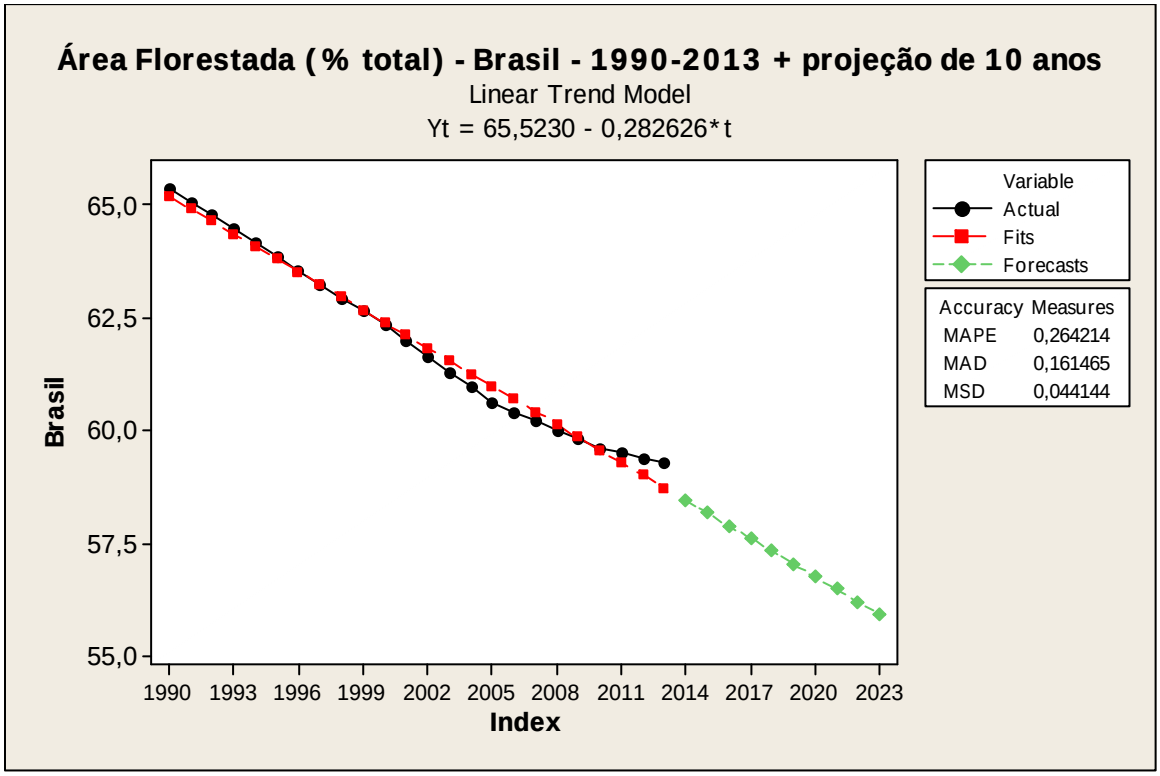
Tamanho da Série de Dados:
As três séries de dados podem ser consideradas satisfatórias para a realização desta pesquisa, uma vez que todas possuem dados de 15 anos ou mais, sem interrupções.

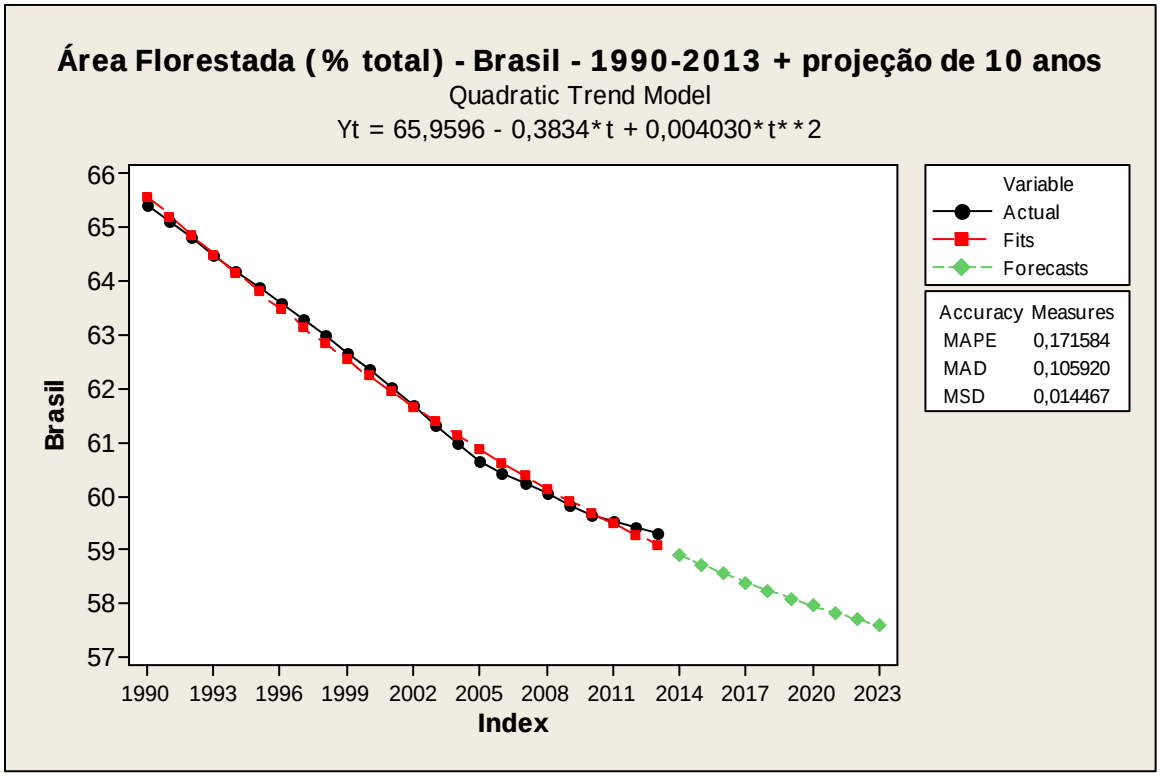
d. ANÁLISE DAS VARIÁVEIS DO EIXO PLANETA

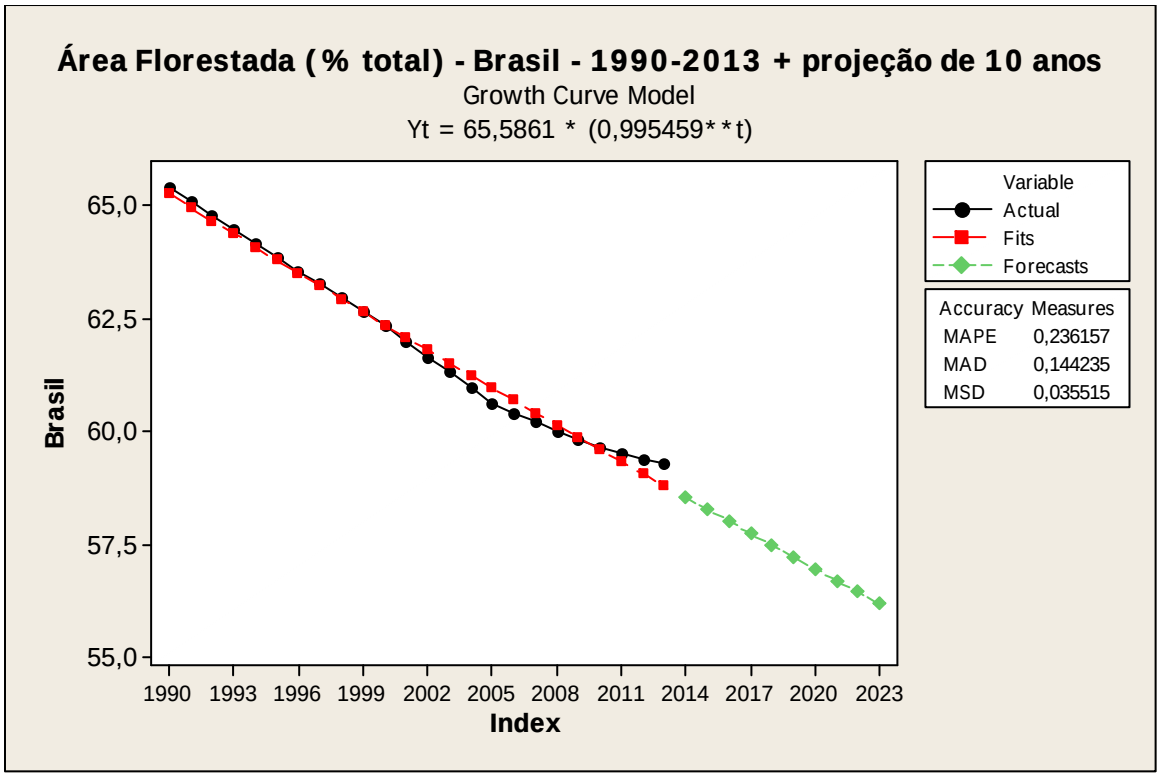
Cada variável é analisada a partir das séries dos quatro países mencionados. São desenvolvidos quatro modelos estatísticos: linear, quadrático, crescimento exponencial e curva s. Os resultados com os quatro modelos são analisados em relação ao fit, buscando-se interpretar qual dos modelos melhor explica o comportamento das observações. Uma linha de tendência de 10 anos (a unidade de tempo das observações) é traçada para estimação de comportamento futuro em cada modelo.

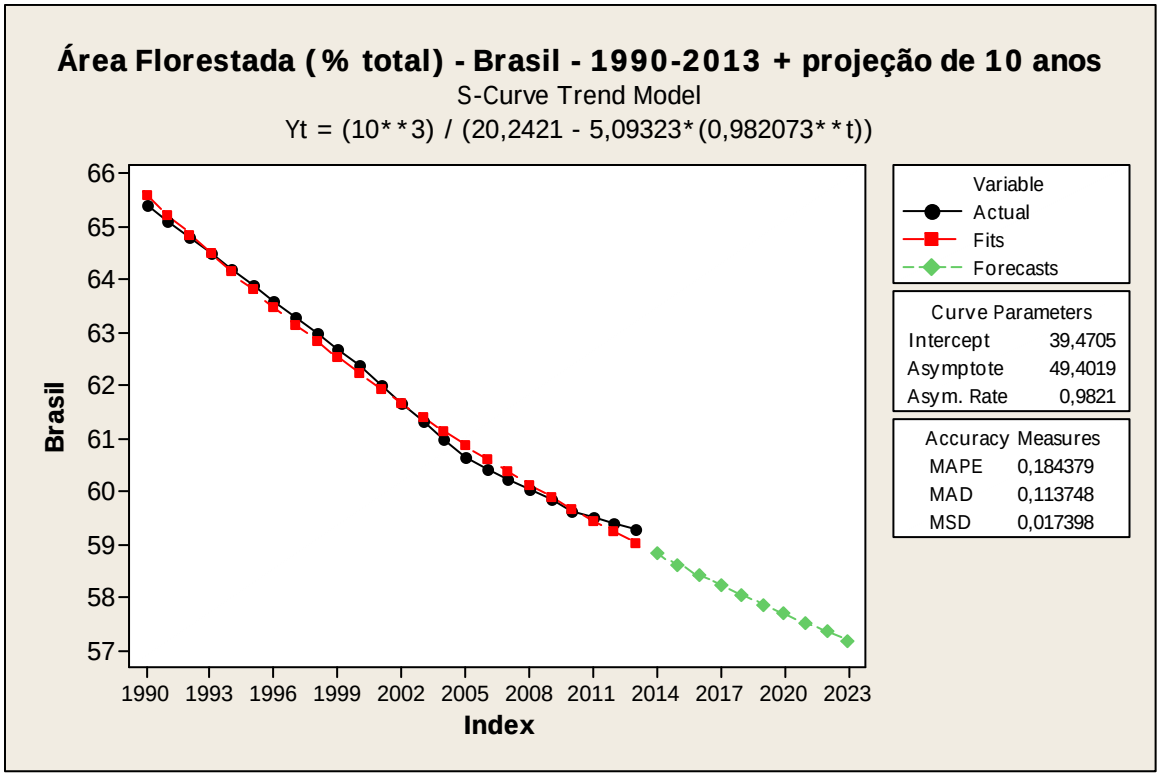
i. Variável: Área Florestada como percentual da área total do país (%)

1. Análise de tendência e projeções - Brasil




Área Florestada (% total) - Brasil - 1990-2013 + projeção de 10 anos
Quadratic Trend Model
Yt = 65,9596 - 0,3834*t + 0,004030*t**2
Brasil
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,171584
MAD 0,105920
MSD 0,014467


Área Florestada (% total) - Brasil - 1990-2013 + projeção de 10 anos
Growth Curve Model
Yt = 65,5861 * (0,995459**t)
Brasil
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,236157
MAD 0,144235
MSD 0,035515



Verifica-se que a área florestada Brasileira é decrescente para o período analisado.

Com relação aos modelos, a função quadrática é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

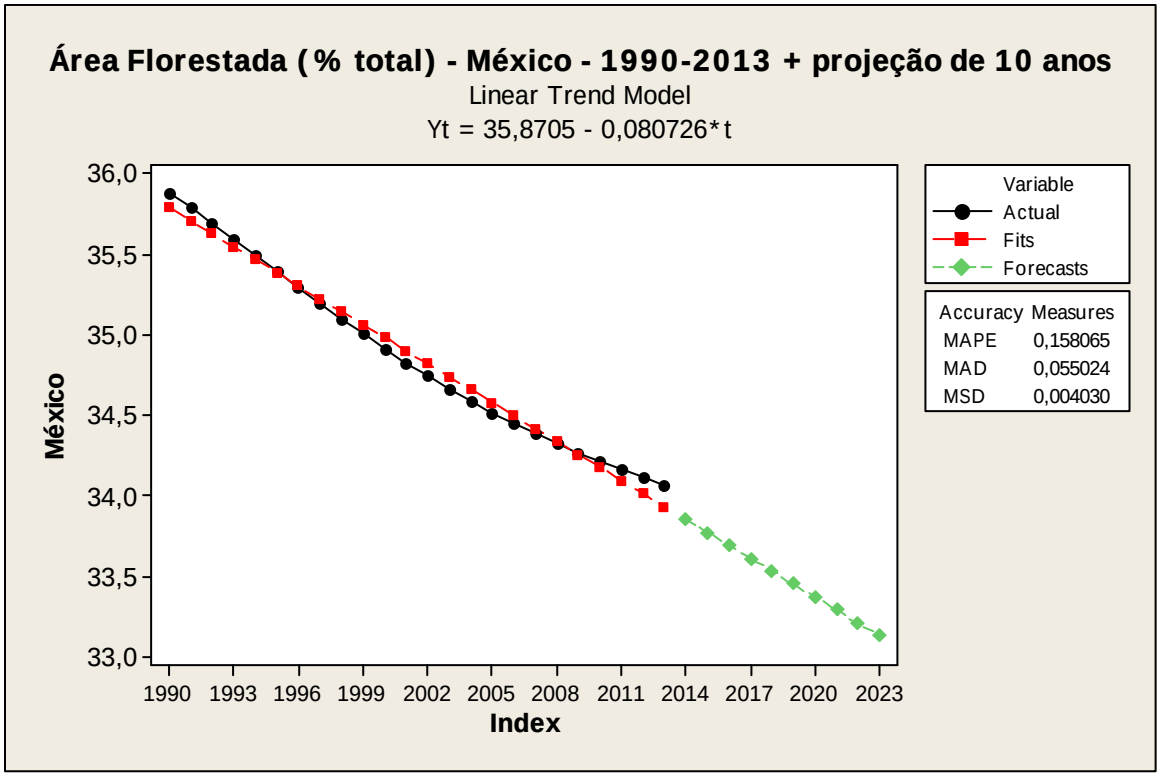
Medida de ajuste nos modelos – Brasil

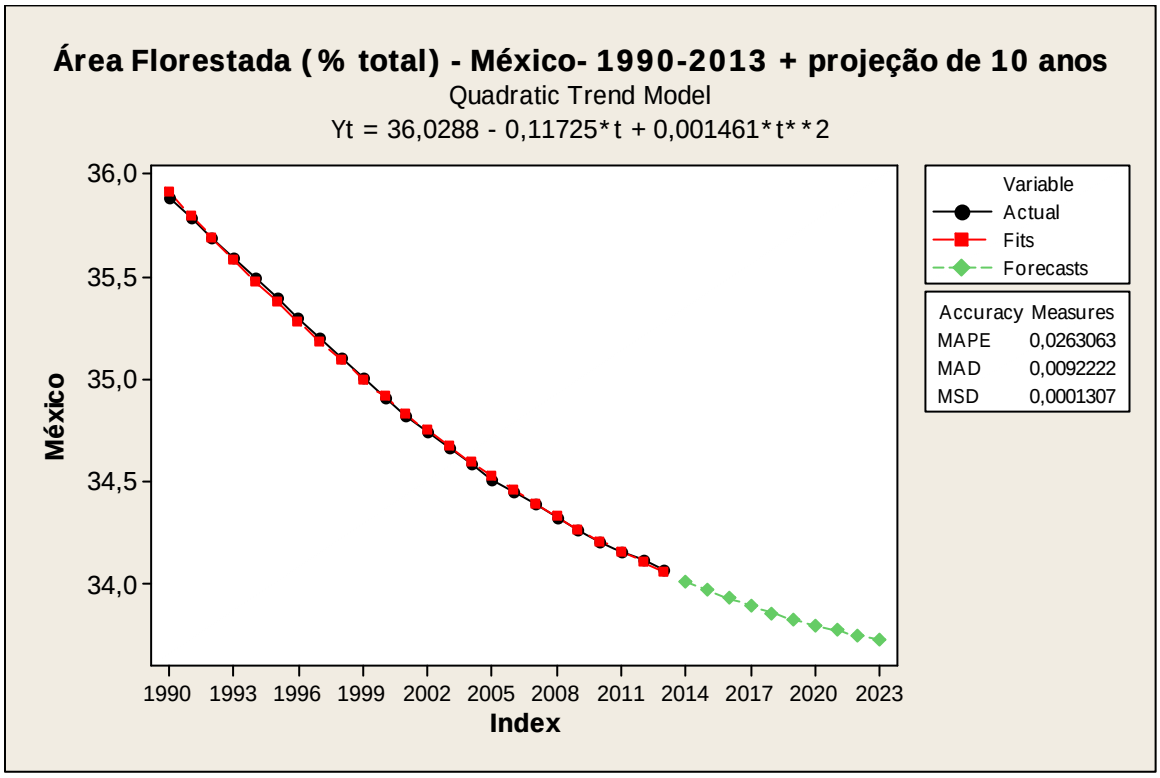
	BR - LI	BR - QD	BR - XP	BR - S
MAPE	0,264214	0,171584	0,236157	0,184379
MAD	0,161465	**0,105920**	0,144235	0,113748
MSD	0,044144	0,014467	0,035515	0,017398

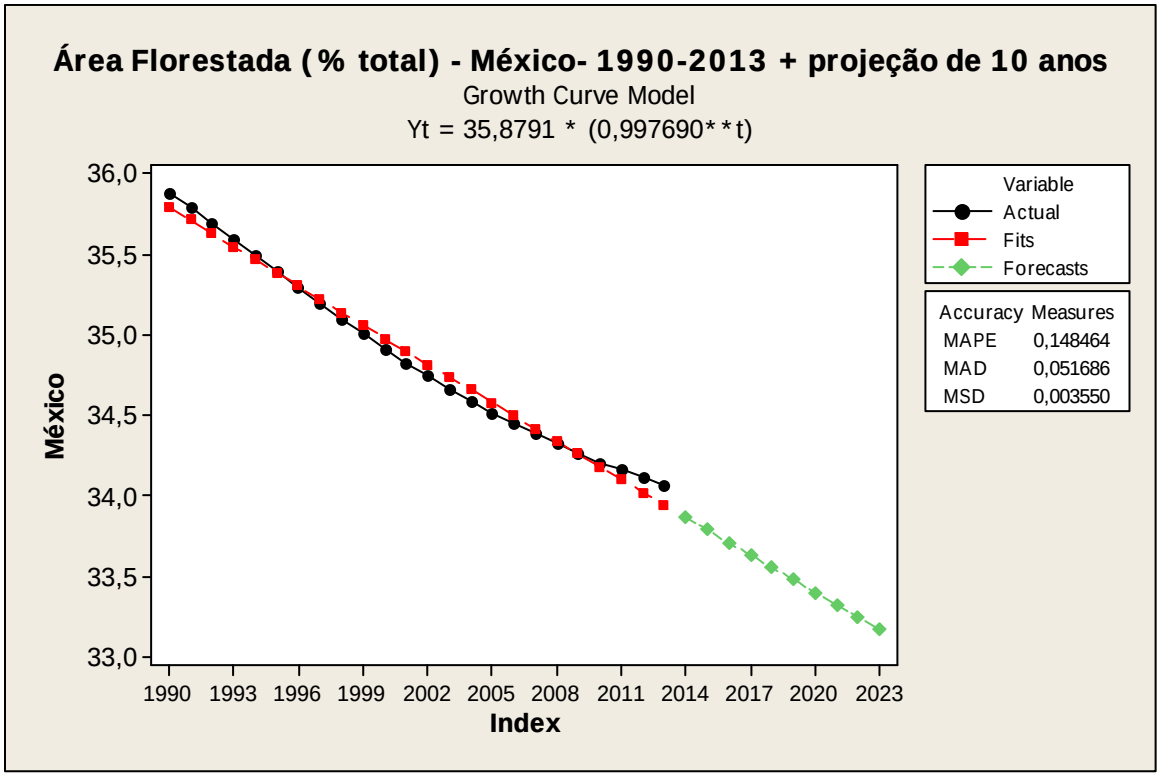
As projeções do modelo quadrático são apresentadas a seguir.

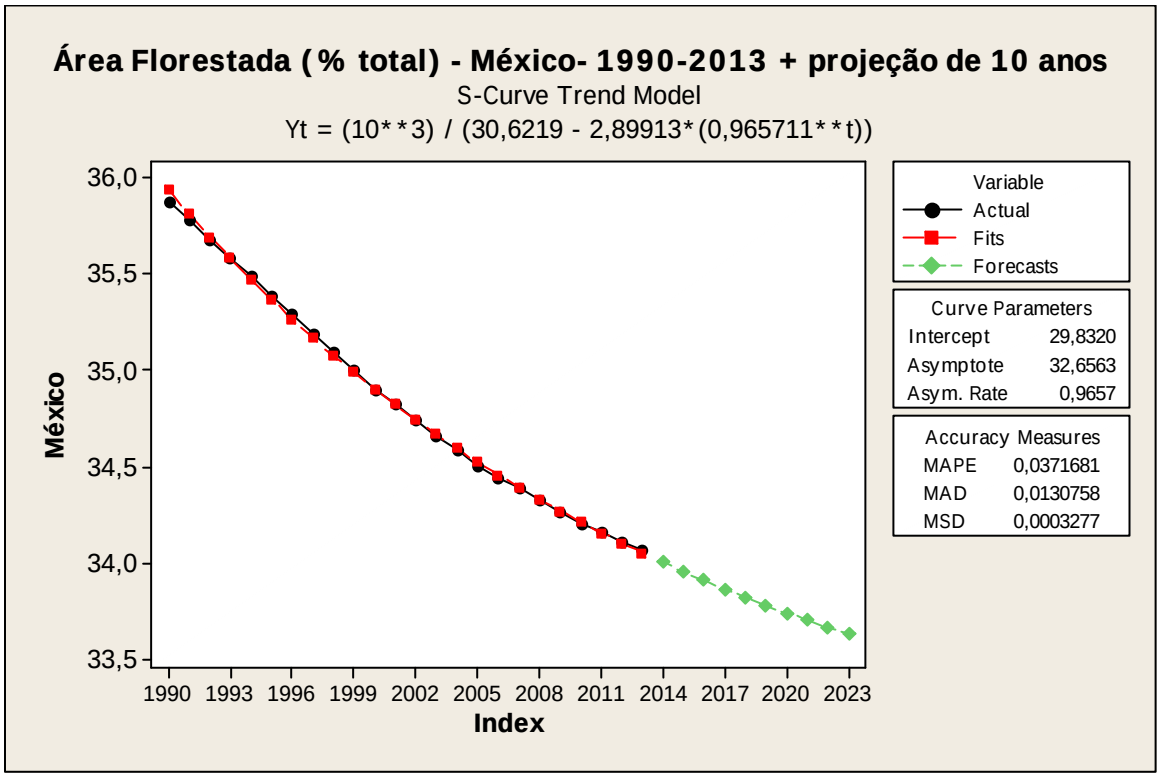
```
Forecasts (%)

Period  Forecast
2014     58,8940
2015     58,7161
2016     58,5463
2017     58,3846
2018     58,2310
2019     58,0854
2020     57,9478
2021     57,8183
2022     57,6969
2023     57,5836
```

2. Análise de tendência e projeções – México


Área Florestada (% total) - México - 1990-2013 + projeção de 10 anos
Linear Trend Model
Yt = 35,8705 - 0,080726*t
México
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,158065
MAD 0,055024
MSD 0,004030


Área Florestada (% total) - México- 1990-2013 + projeção de 10 anos
Quadratic Trend Model
Yt = 36,0288 - 0,11725*t + 0,001461*t**2
México
Index
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,0263063
MAD 0,0092222
MSD 0,0001307





Verifica-se que a área florestada Mexicana é decrescente para o período analisado.

Com relação aos modelos, a função de Curva S é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

Medida de ajuste nos modelos – México

MX - LI	MX - QD	MX - XP	MX - S

MAPE	0,158065	0,0263063	0,148464	0,0371681
MAD	0,055024	0,0092222	0,051686	0,0130758
MSD	0,004030	0,0001307	0,003550	0,0003277

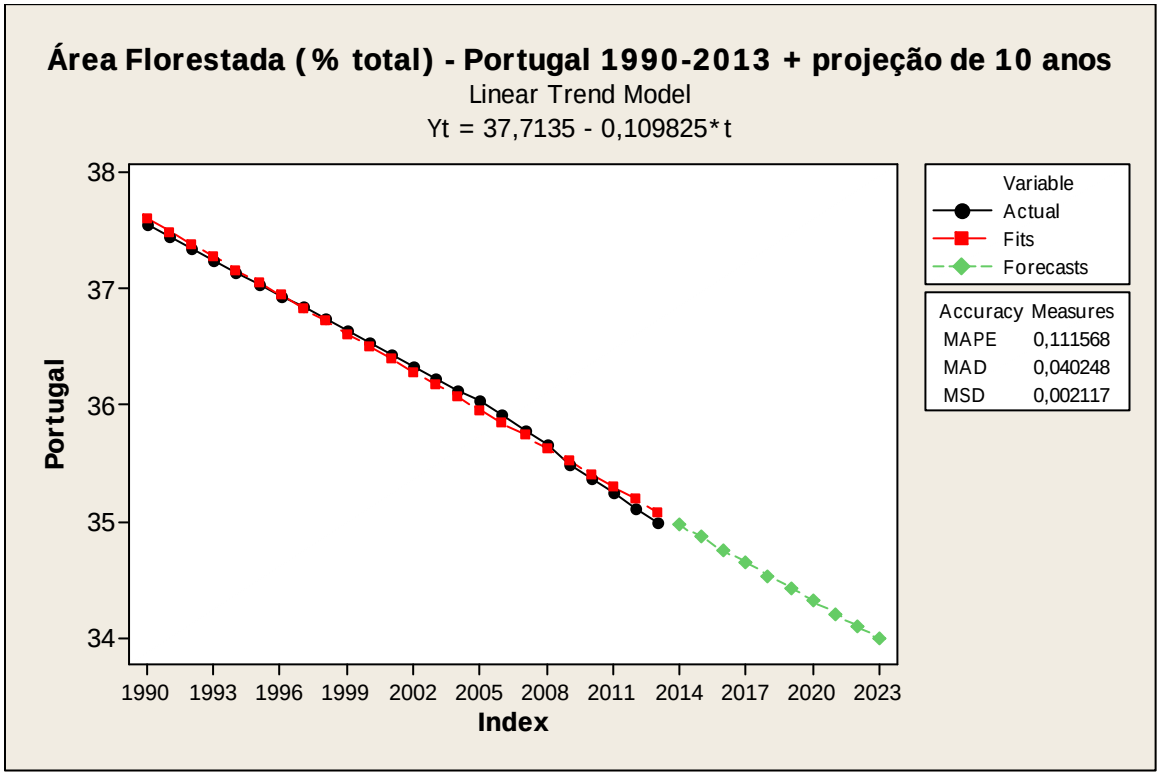
As projeções do modelo de Curva S são apresentadas a seguir.

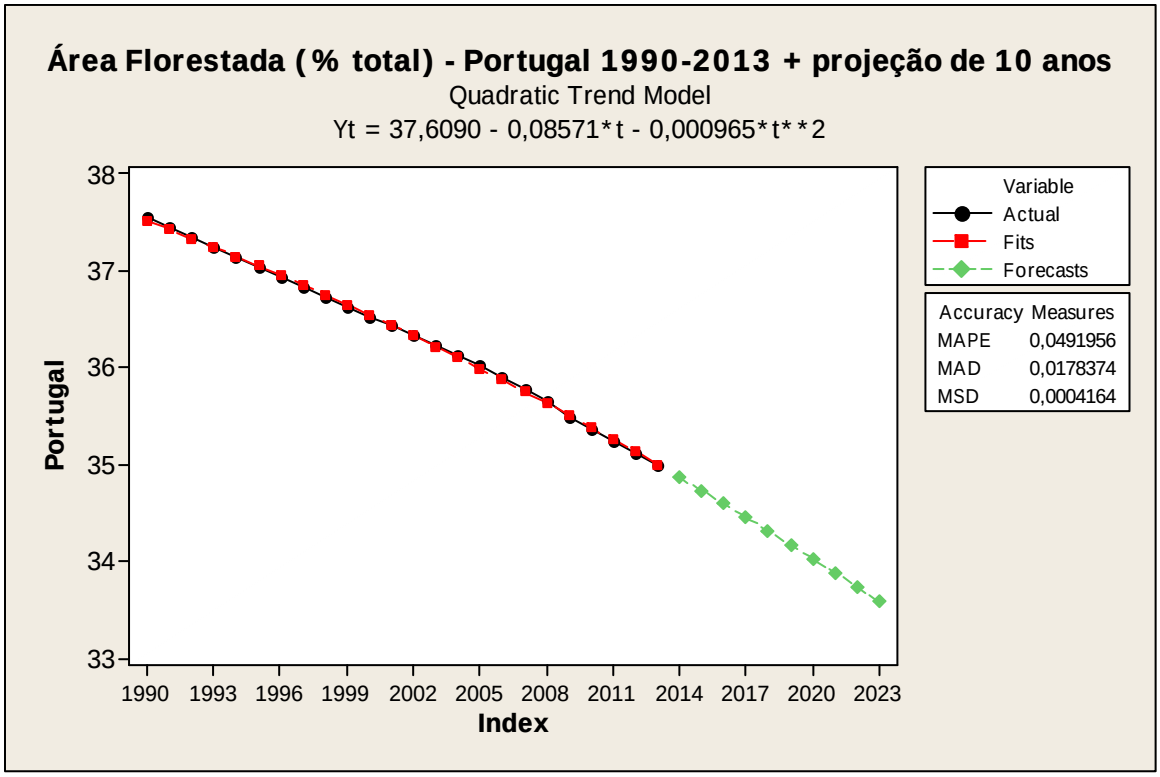
Forecasts (%)

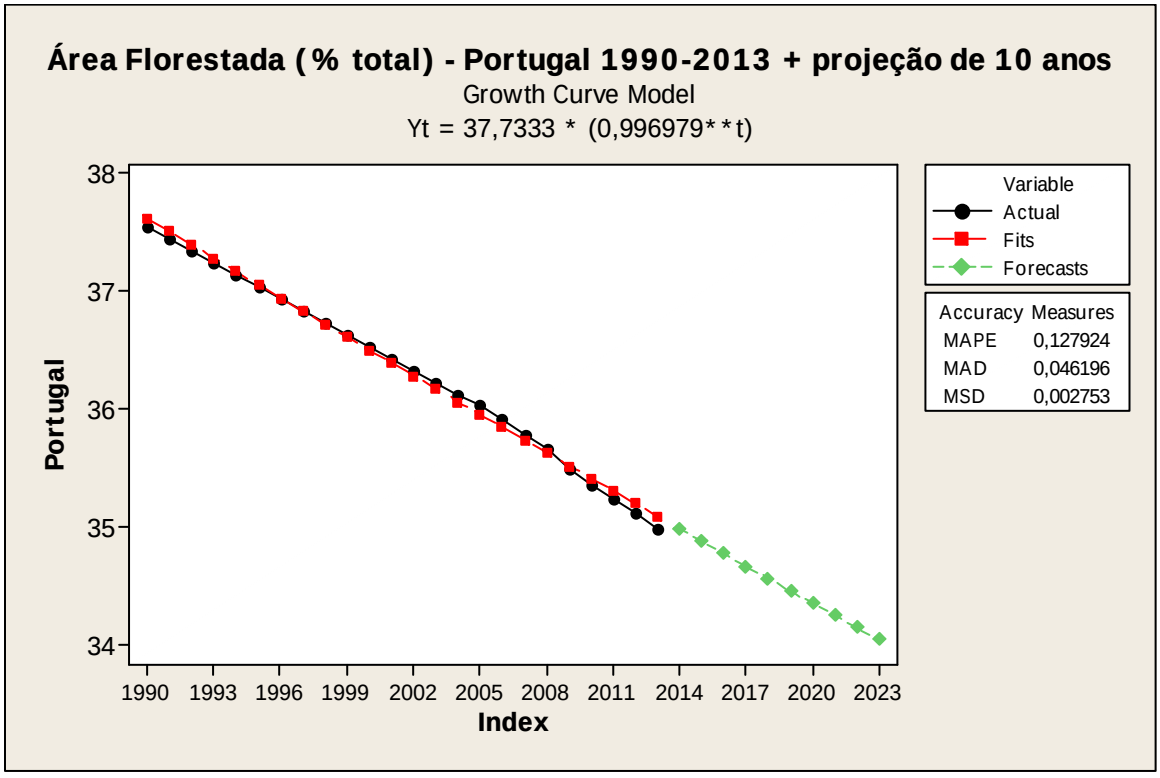
Period	Forecast
2014	34,0019
2015	33,9539
2016	33,9077
2017	33,8632
2018	33,8204
2019	33,7791
2020	33,7393
2021	33,7010
2022	33,6641
2023	33,6285

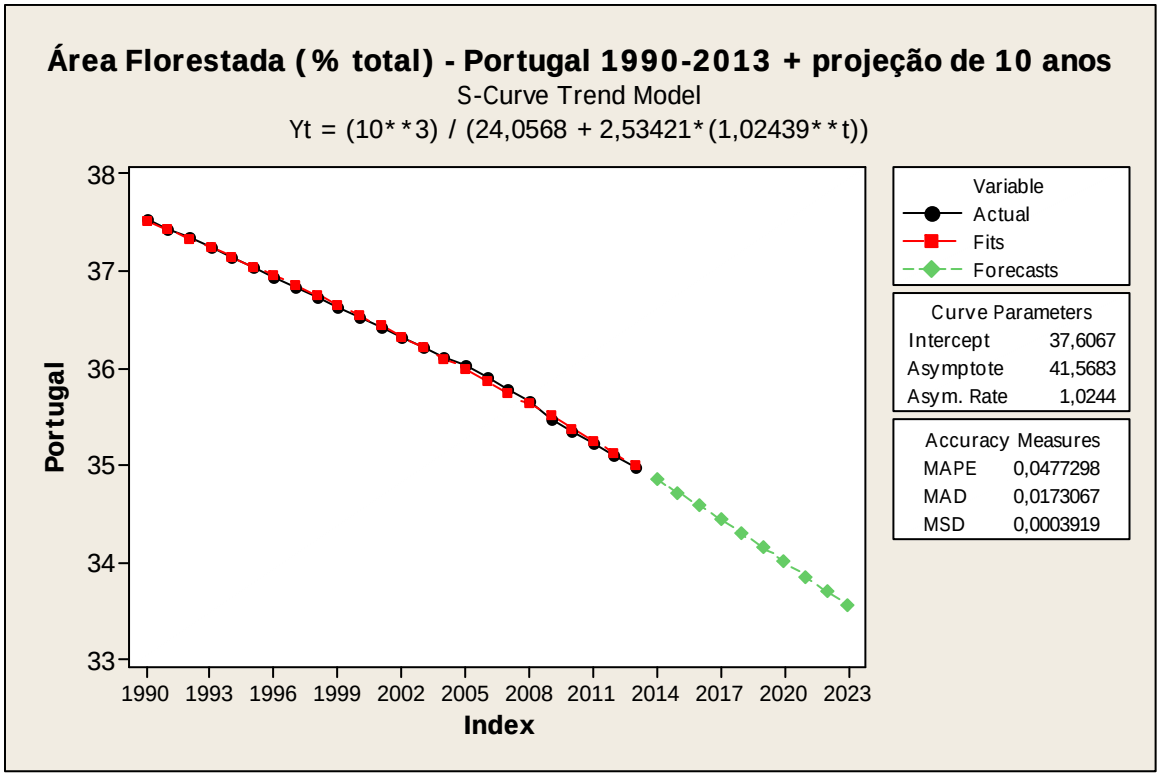
Nota-se que ao final do período projetado o México disporia conforme o model de 33,62% somente de sua área total coberta por florestas, quase metade do que o projetado para o mesmo ano para o Brasil, 57%.

3. Análise de tendência e projeções – Portugal




Área Florestada (% total) - Portugal 1990-2013 + projeção de 10 anos
Quadratic Trend Model
Yt = 37,6090 - 0,08571*t - 0,000965*t**2
Portugal
Index
38
37
36
35
34
33
1990 1993 1996 1999 2002 2005 2008 2011 2014 2017 2020 2023
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,0491956
MAD 0,0178374
MSD 0,0004164


Área Florestada (% total) - Portugal 1990-2013 + projeção de 10 anos
Growth Curve Model
Yt = 37,7333 * (0,996979**t)
Portugal
Index
38
37
36
35
34
1990 1993 1996 1999 2002 2005 2008 2011 2014 2017 2020 2023
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 0,127924
MAD 0,046196
MSD 0,002753



Verifica-se que a área florestada Portuguesa é decrescente para o período analisado. A redução de área florestada é significativamente menor em proporção ao analisado no início da série, do que para Brasil, e semelhante ao observado para o México.

Com relação aos modelos, a função de Curva S é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

Medida de ajuste nos modelos – Portugal

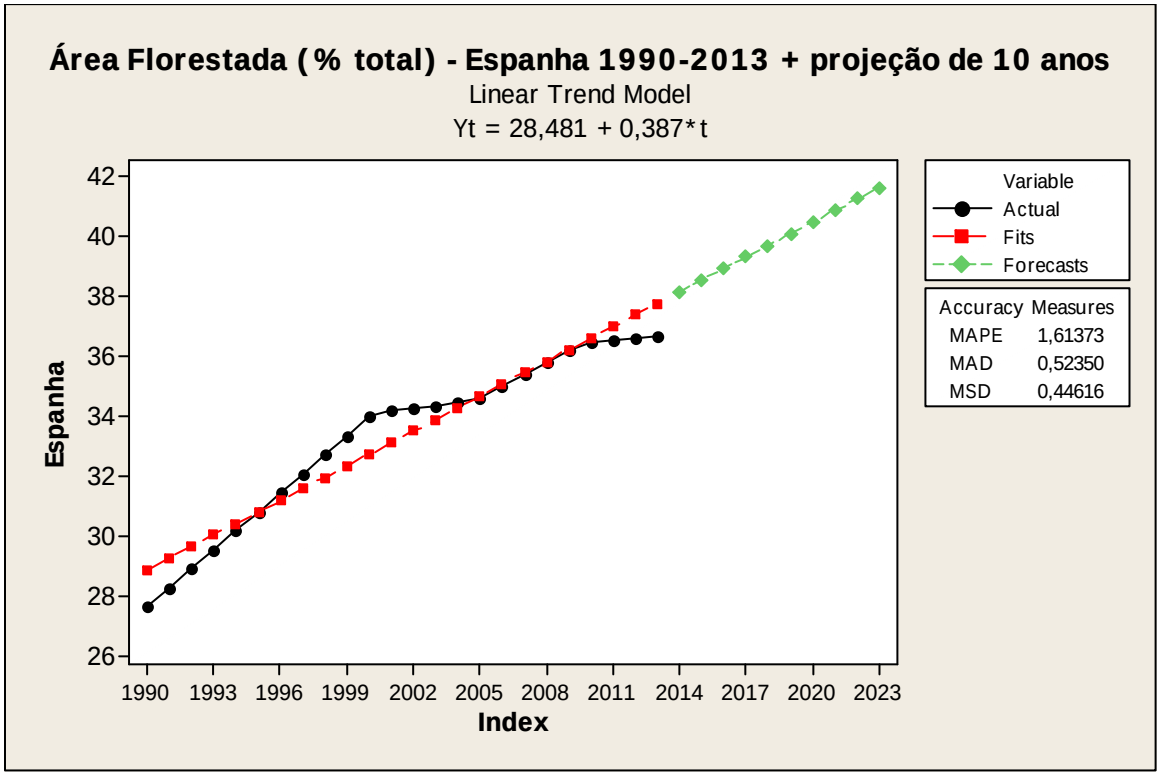
	PT - LI	PT - QD	PT - XP	PT - S
MAPE	0,111568	0,0491956	0,127924	0,0477298
MAD	0,040248	0,0178374	0,046196	0,0173067
MSD	0,002117	0,0004164	0,002753	0,0003919

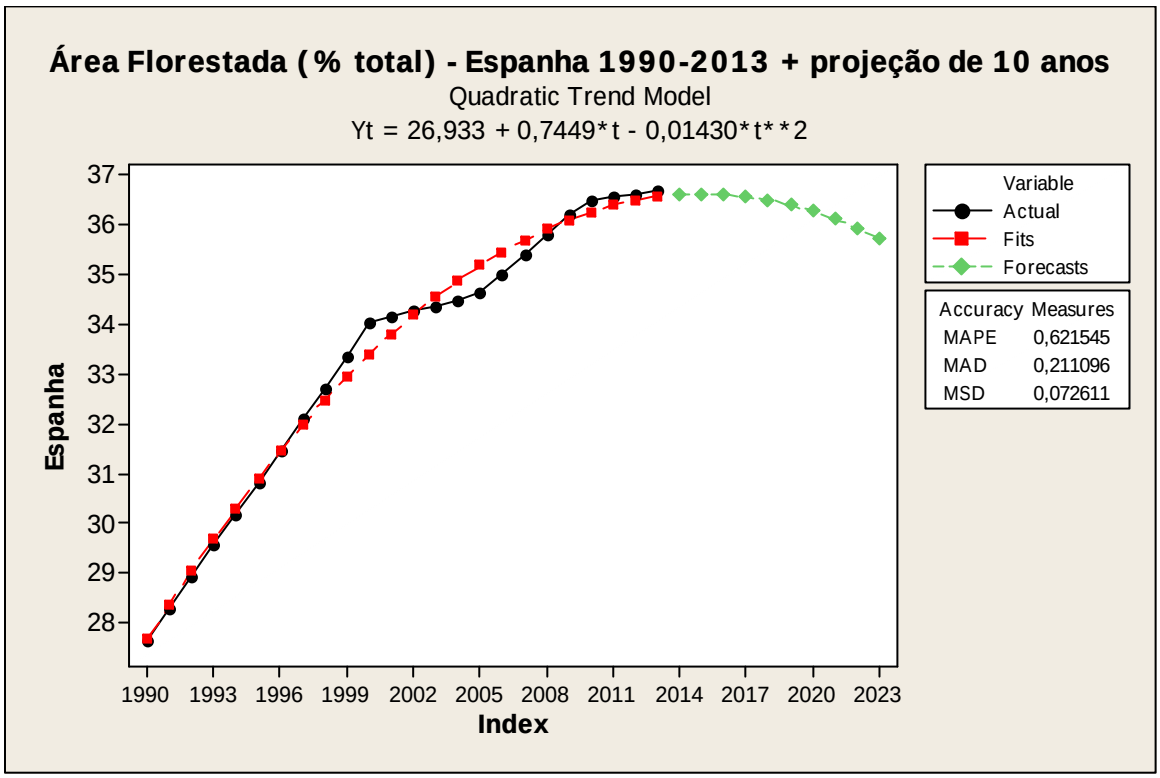
As projeções do modelo de Curva S são apresentadas a seguir.

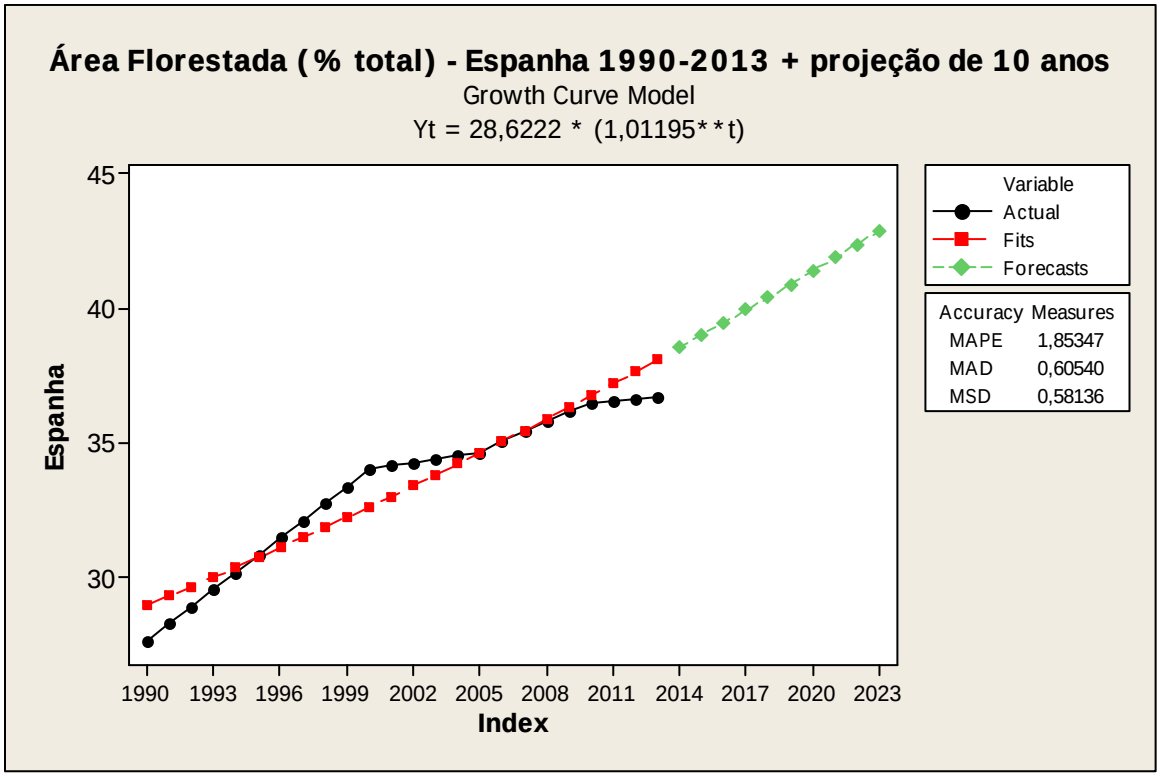
Forecasts (%)

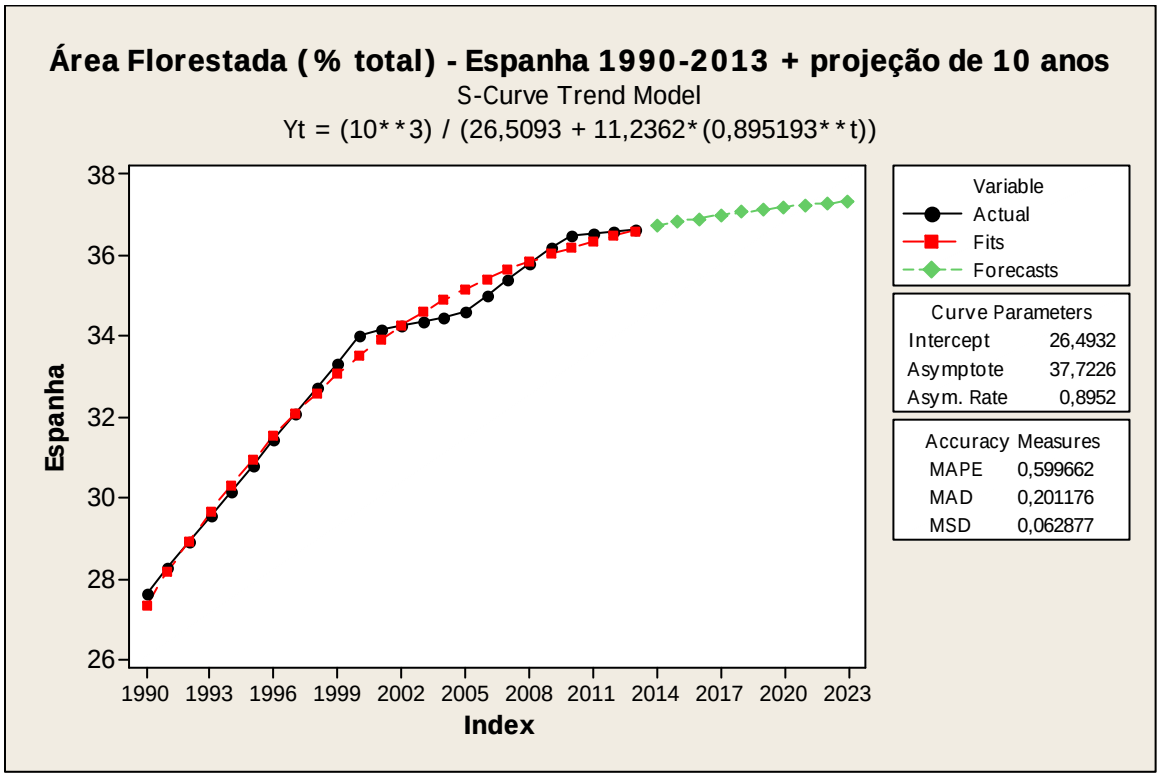
Period	Forecast
2014	34,8611
2015	34,7244
2016	34,5856
2017	34,4445
2018	34,3011
2019	34,1555
2020	34,0076
2021	33,8574
2022	33,7050
2023	33,5502

4. Análise de tendência e projeções – Espanha









Verifica-se que a área florestada Espanhola difere em tendência daquelas observadas nos outros três países, pois é crescente para o período analisado. O aumento de área florestada é relativamente constante no período analisado.

Com relação aos modelos, a função de Curva S é a que melhor representa a série de dados.

Os indicadores são apresentados a seguir.

Medida de ajuste nos modelos – Espanha

	ES - LI	ES - QD	ES - XP	ES - S
MAPE	1,61373	0,621545	1,85347	0,599662
MAD	0,52350	0,211096	0,60540	**0,201176**
MSD	0,44616	0,072611	0,58136	0,062877

As projeções do modelo de Curva S são apresentadas a seguir.

Forecasts (%)

Period	Forecast
2014	36,7446
2015	36,8447
2016	36,9348
2017	37,0158
2018	37,0886
2019	37,1541
2020	37,2128
2021	37,2656
2022	37,3130
2023	37,3555

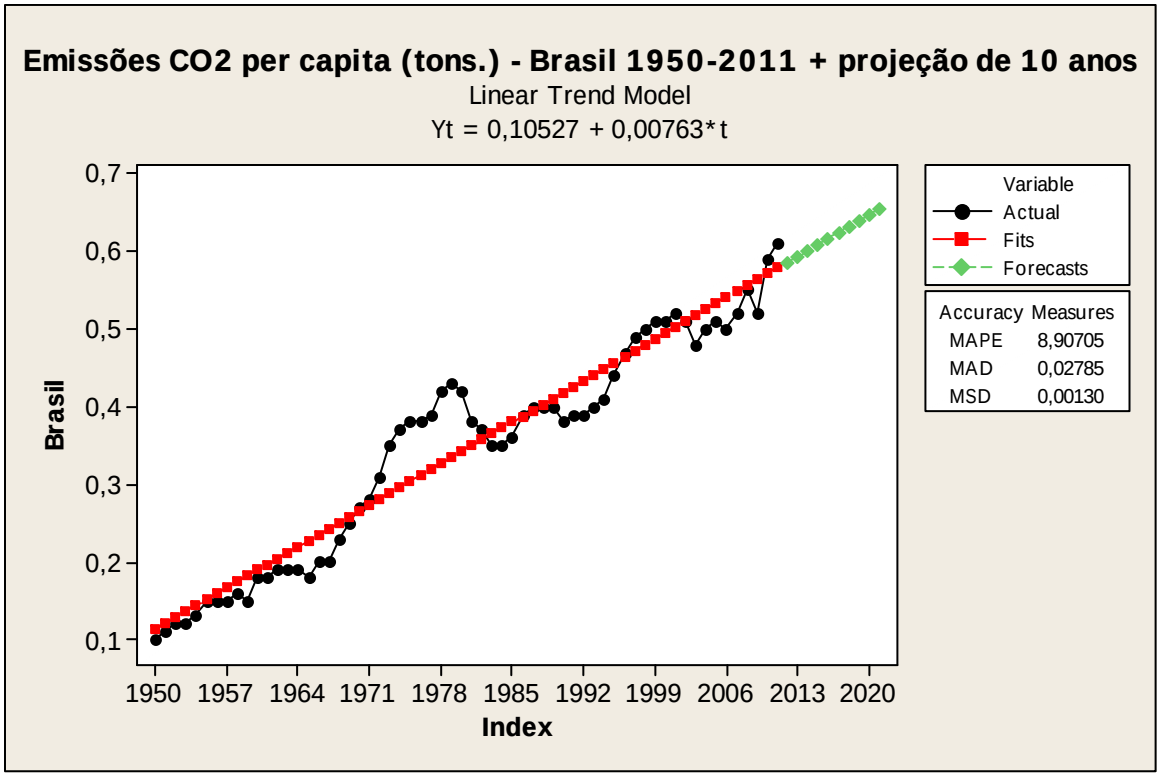
e. Variáveis e fontes de dados

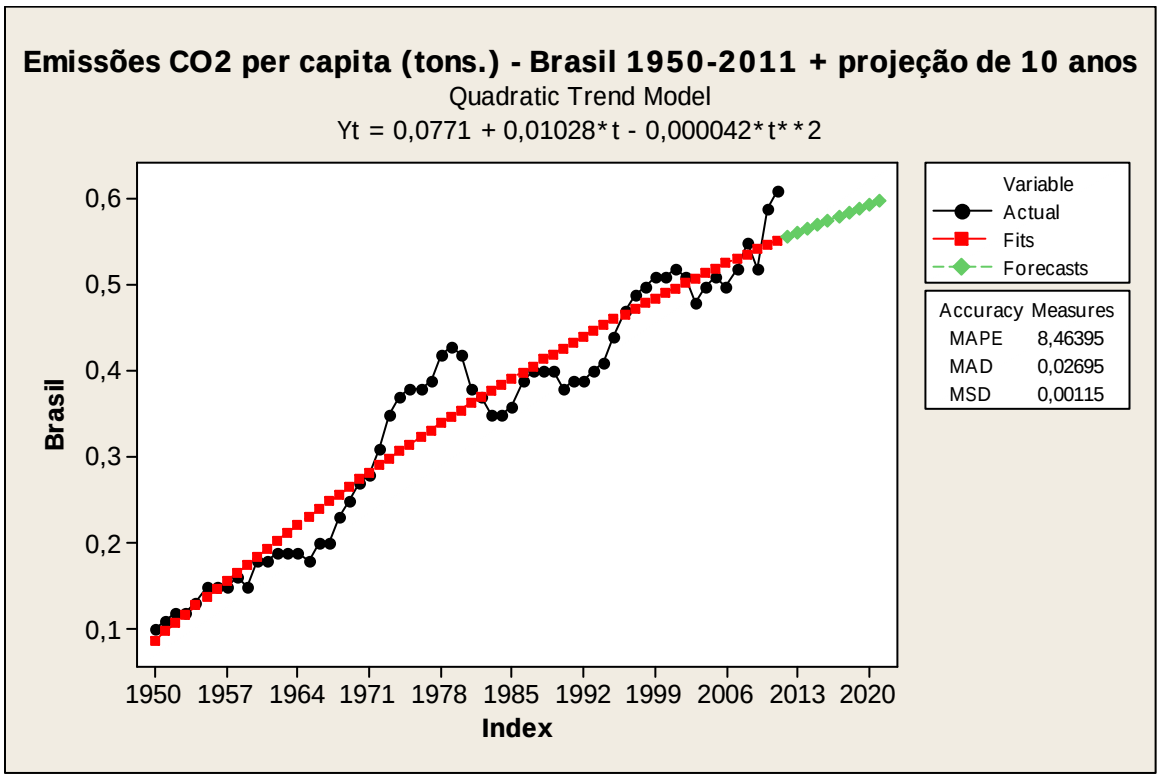
i. Variável Emissões de CO2 per capita.

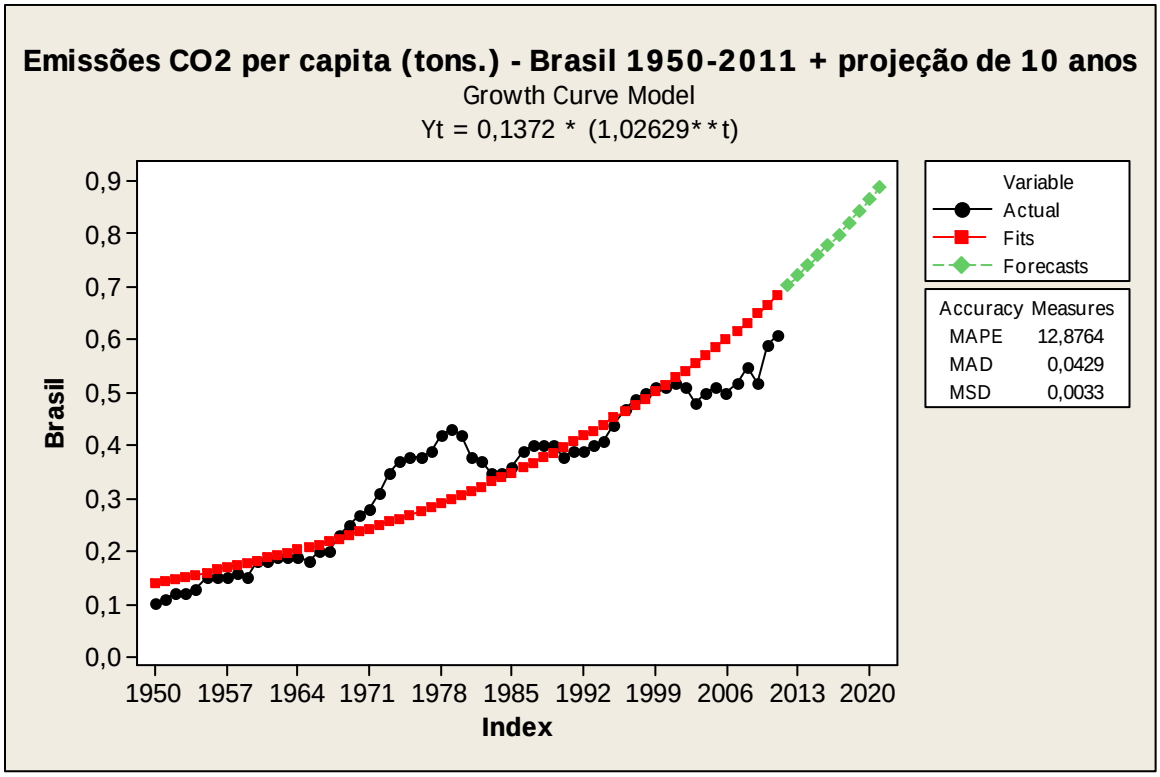
A variável em questão é definida como o total de emissões de CO2 provenientes de emissões per capita, medidas em tonelada per capita anual. Segudo o knoema, *"as emissões de dióxido de carbono são os decorrentes da queima de combustíveis fósseis e do fabrico de cimento. Eles incluem o dióxido de carbono produzido durante o consumo de combustíveis sólidos, líquidos e gases combustíveis e queima de gás"*.
A fonte de dados para esta variável é a *Carbon Dioxide Information Analysis Center*

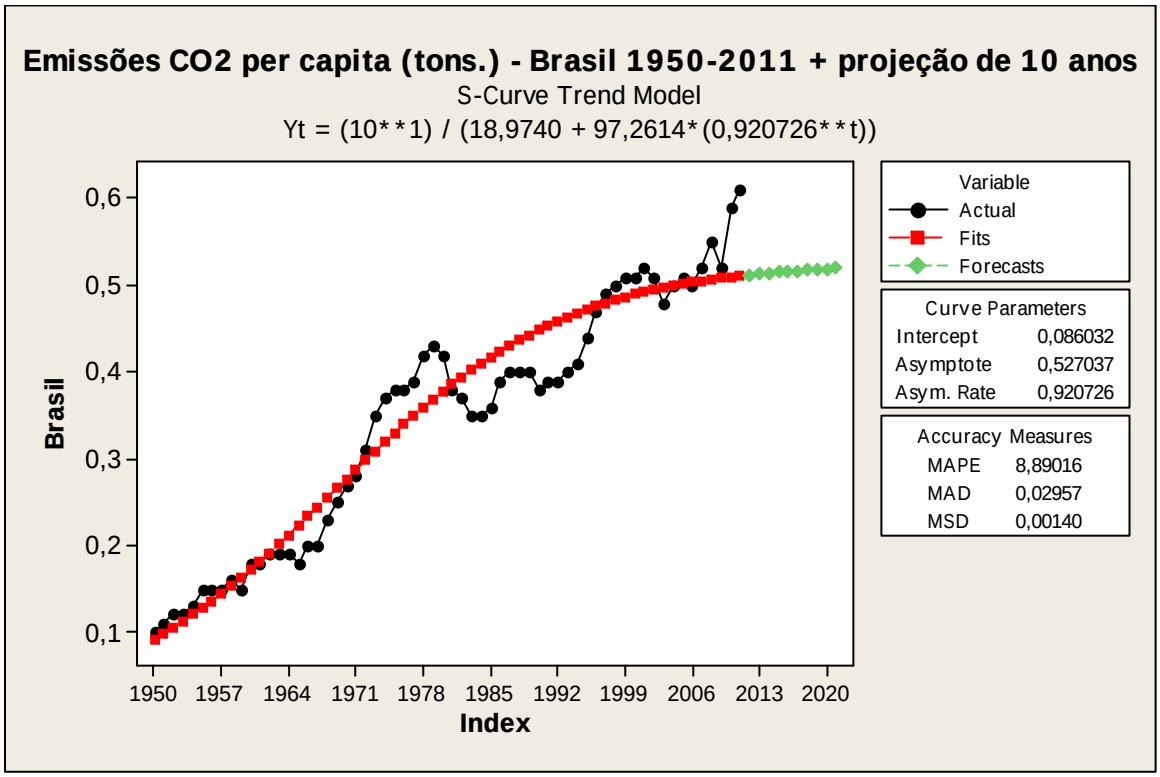
A série de dados analisada possui sessenta e uma observações, iniciando-se em 1950 e encerrando-se em 2011.

1. Análise de tendência e projeções – Brasil


Emissões CO2 per capita (tons.) - Brasil 1950-2011 + projeção de 10 anos
Linear Trend Model
Yt = 0,10527 + 0,00763*t
Brasil
Index
0,7
0,6
0,5
0,4
0,3
0,2
0,1
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 8,90705
MAD 0,02785
MSD 0,00130


Emissões CO2 per capita (tons.) - Brasil 1950-2011 + projeção de 10 anos
Quadratic Trend Model
Yt = 0,0771 + 0,01028*t - 0,000042*t**2
Brasil
Index
0,6
0,5
0,4
0,3
0,2
0,1
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 8,46395
MAD 0,02695
MSD 0,00115





Verifica-se que as emissões no Brasil decrescem ao final dos anos 70 e em parte dos anos 80,período correspondente à crise econômica provocada pelo *default* da divida externa.

Com relação aos modelos, a função quadrática é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

Medida de ajuste nos modelos – Brasil

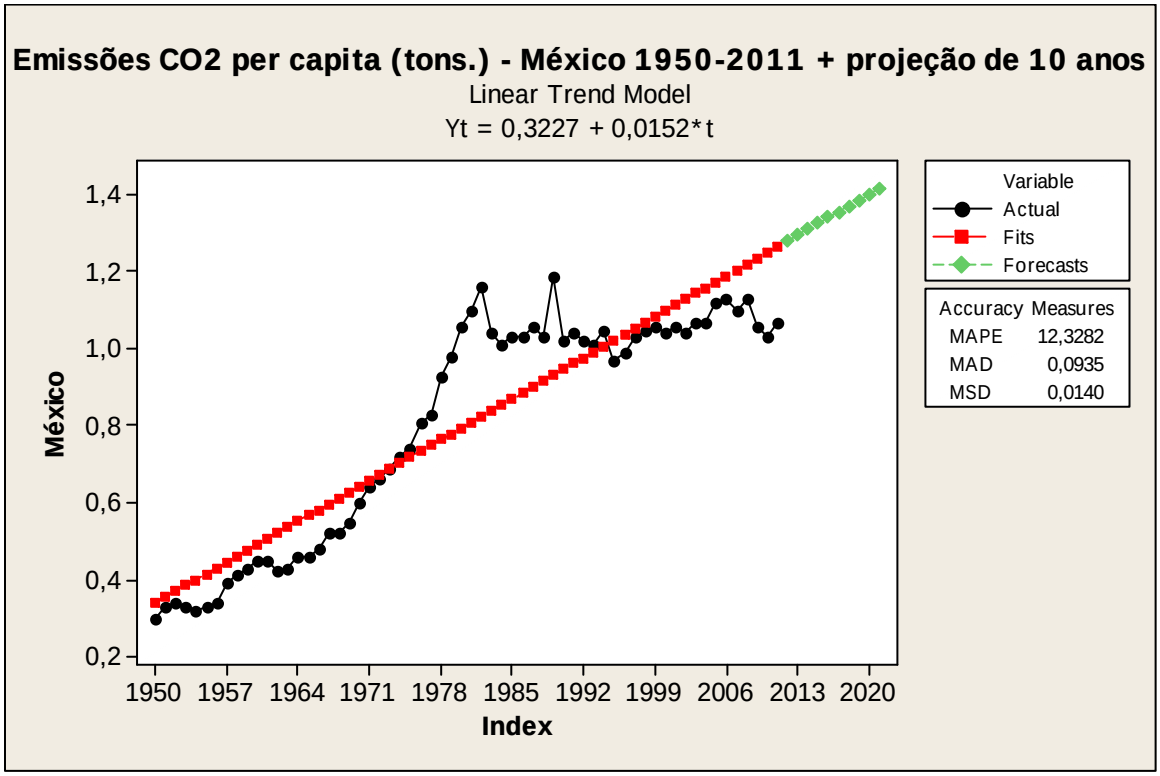
	BR - LI	BR - QD	BR - XP	BR - S
MAPE	8,90705	8,46395	12,8764	8,89016
MAD	0,02785	**0,02695**	0,0429	0,02957
MSD	0,00130	0,00115	0,0033	0,00140

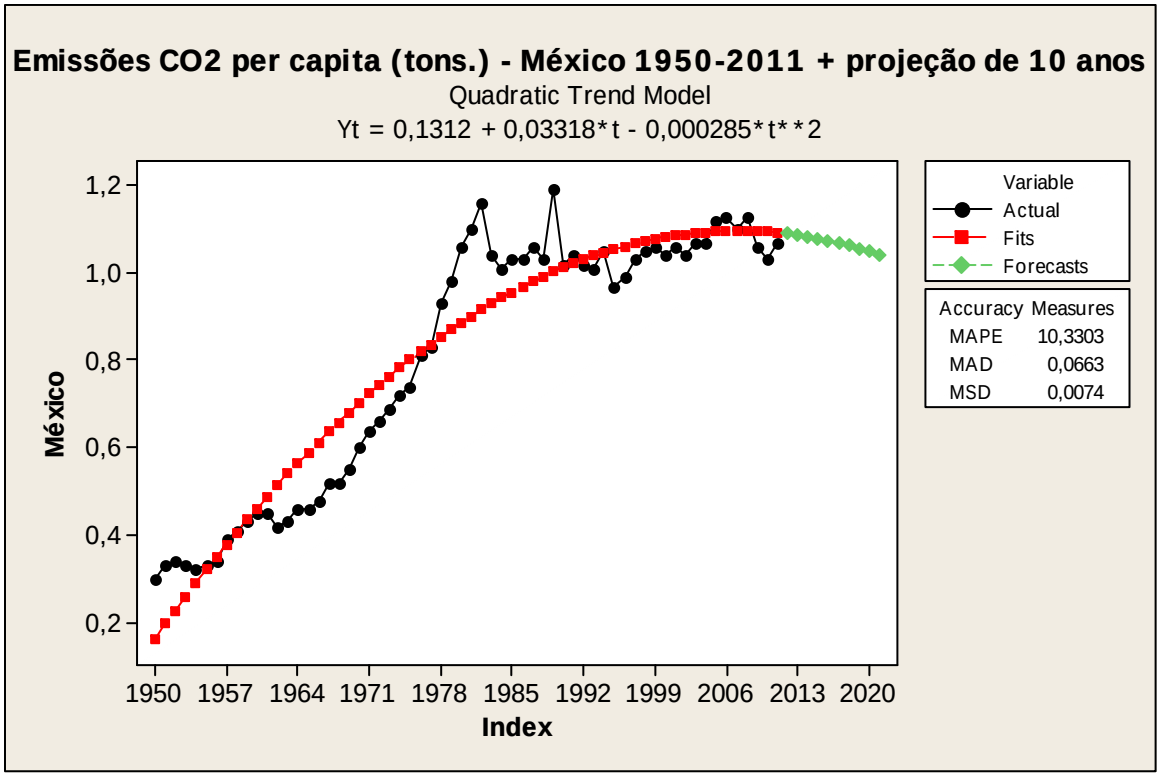
As projeções para dez anos do modelo com melhor adequação são apresentadas a seguir

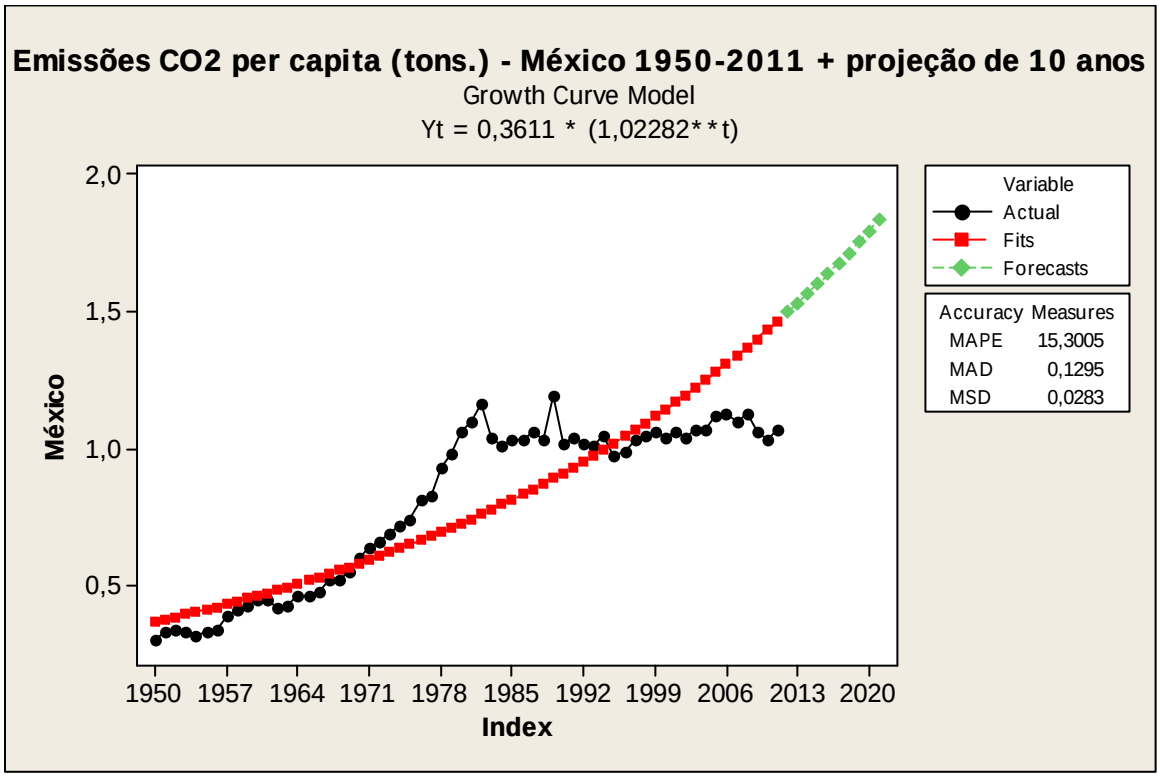
Forecasts

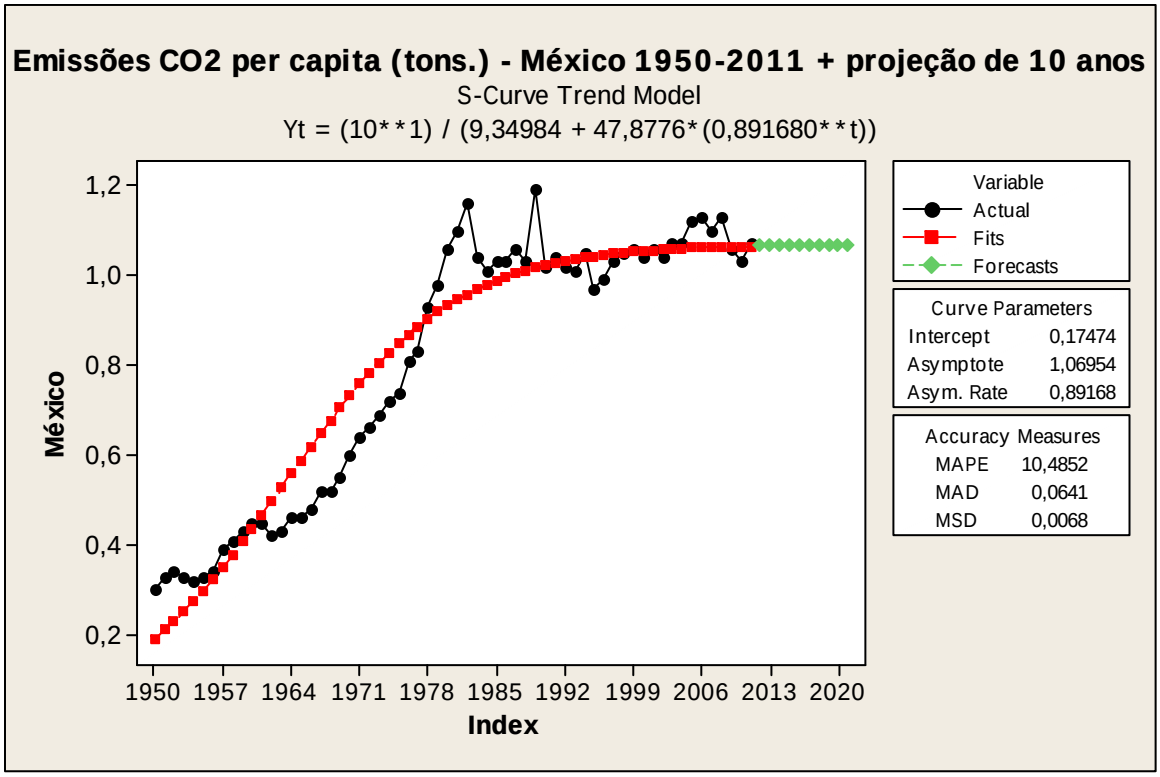
Period	Forecast
2012	0,557805
2013	0,562749
2014	0,567609
2015	0,572385
2016	0,577077
2017	0,581686
2018	0,586210
2019	0,590650
2020	0,595006
2021	0,599278

2. Análise de tendência e projeções – México




Emissões CO2 per capita (tons.) - México 1950-2011 + projeção de 10 anos
Quadratic Trend Model
Yt = 0,1312 + 0,03318*t - 0,000285*t**2
México
Index
1,2
1,0
0,8
0,6
0,4
0,2
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 10,3303
MAD 0,0663
MSD 0,0074


Emissões CO2 per capita (tons.) - México 1950-2011 + projeção de 10 anos
Growth Curve Model
Yt = 0,3611 * (1,02282**t)
México
Index
2,0
1,5
1,0
0,5
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 15,3005
MAD 0,1295
MSD 0,0283



Verifica-se que as emissões no México cresceram muito pouco desde os anos 80, exceto em 1987. Com relação aos modelos, a função quadrática é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

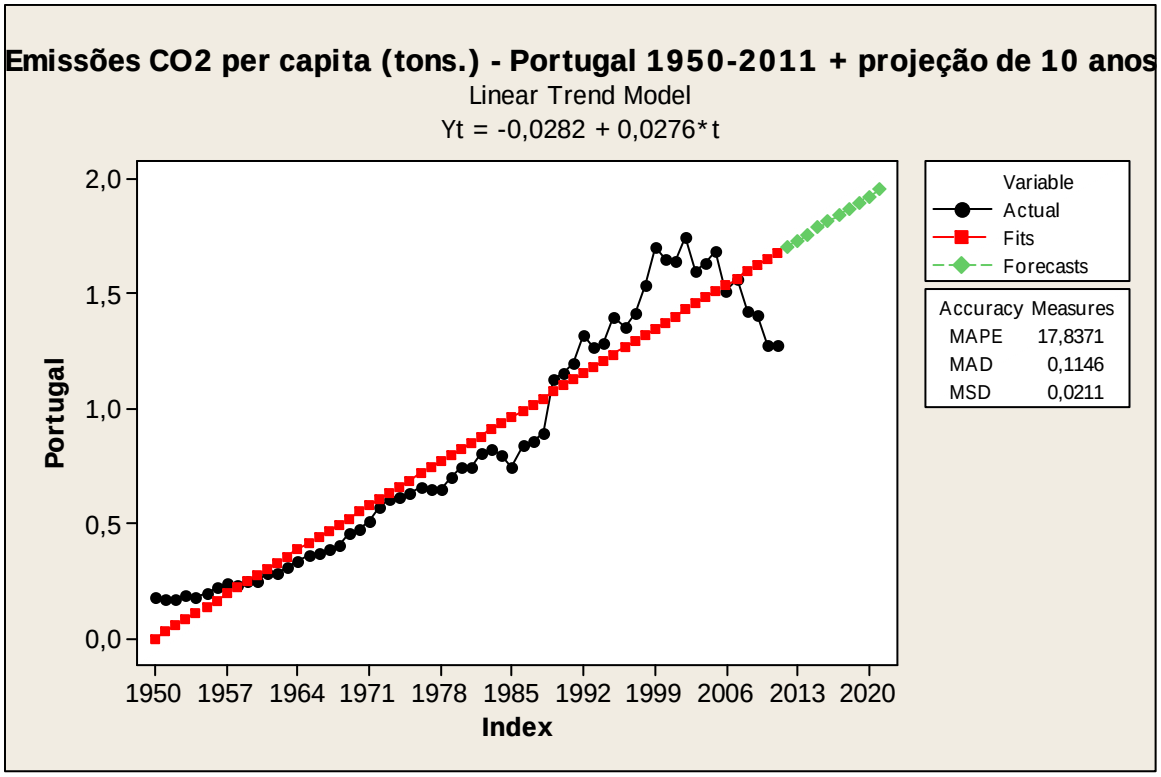
Medida de ajuste nos modelos – México

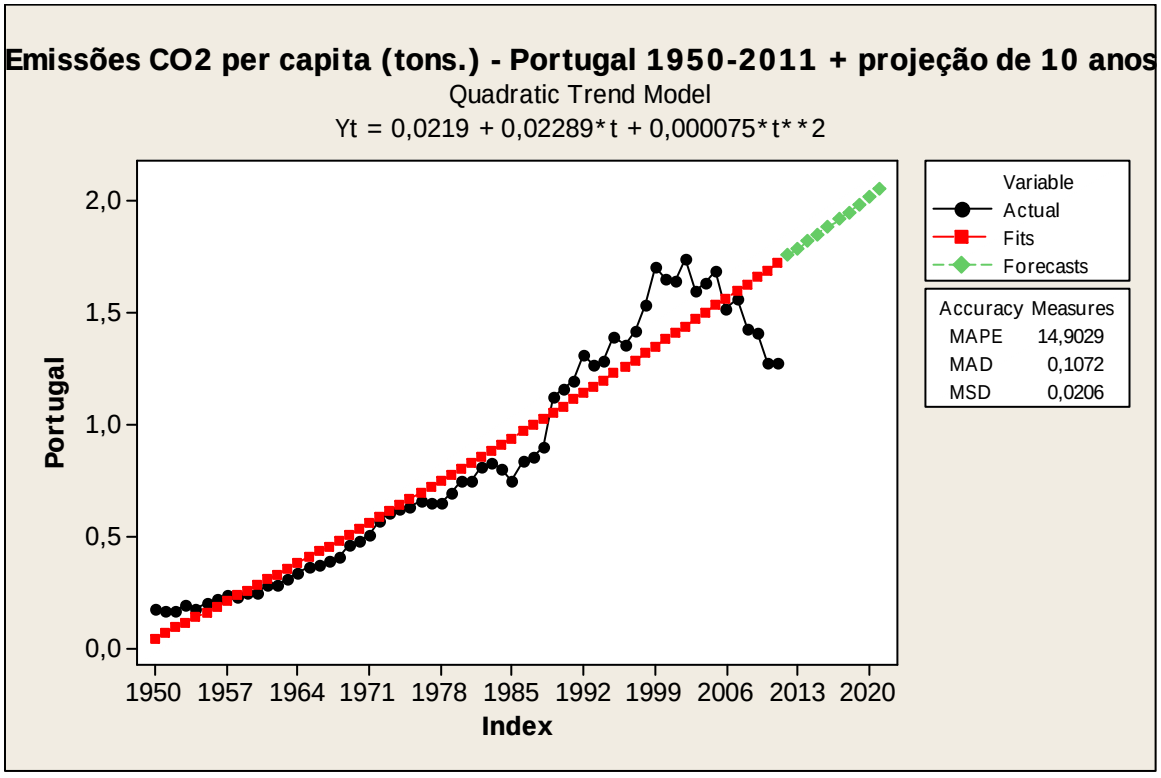
	MX - LI	MX - QD	MX - XP	MX - S
MAPE	12,3282	10,3303	15,3005	10,4852
MAD	0,0935	**0,0663**	0,1295	0,0641
MSD	0,0140	0,0074	0,0283	0,0068

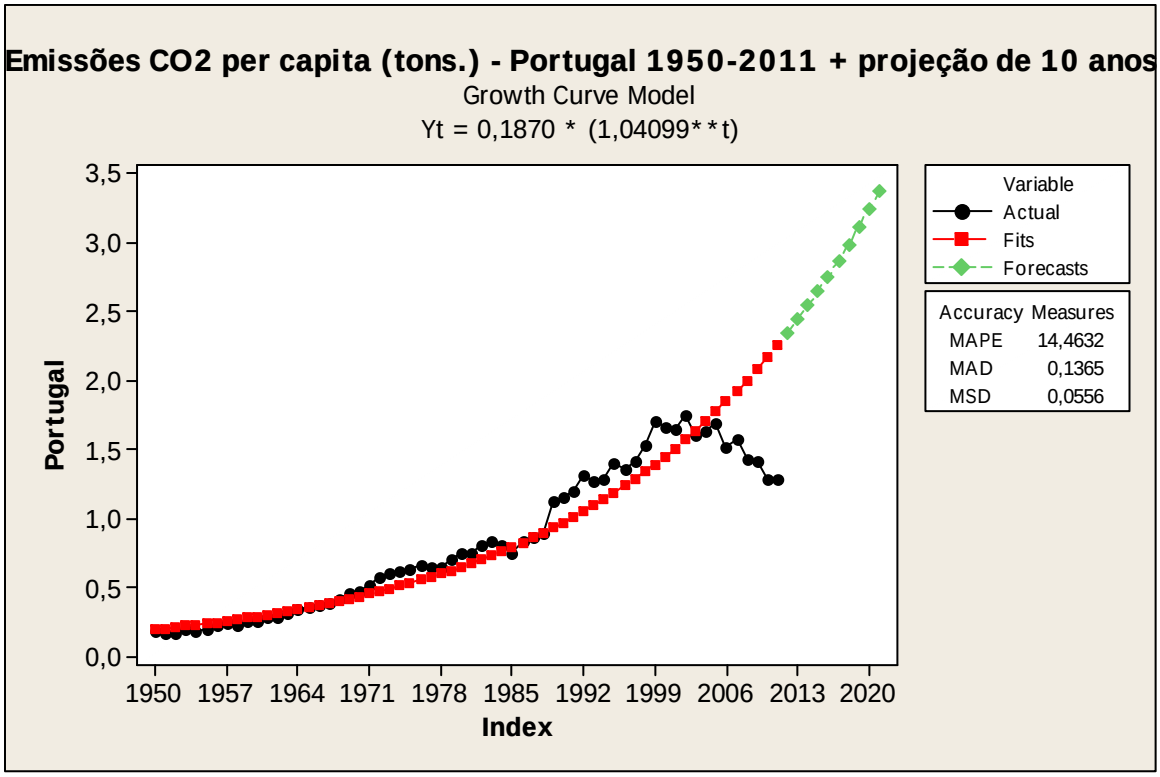
As projeções de dez anos para o modelo mais adequado são apresentadas a seguir.

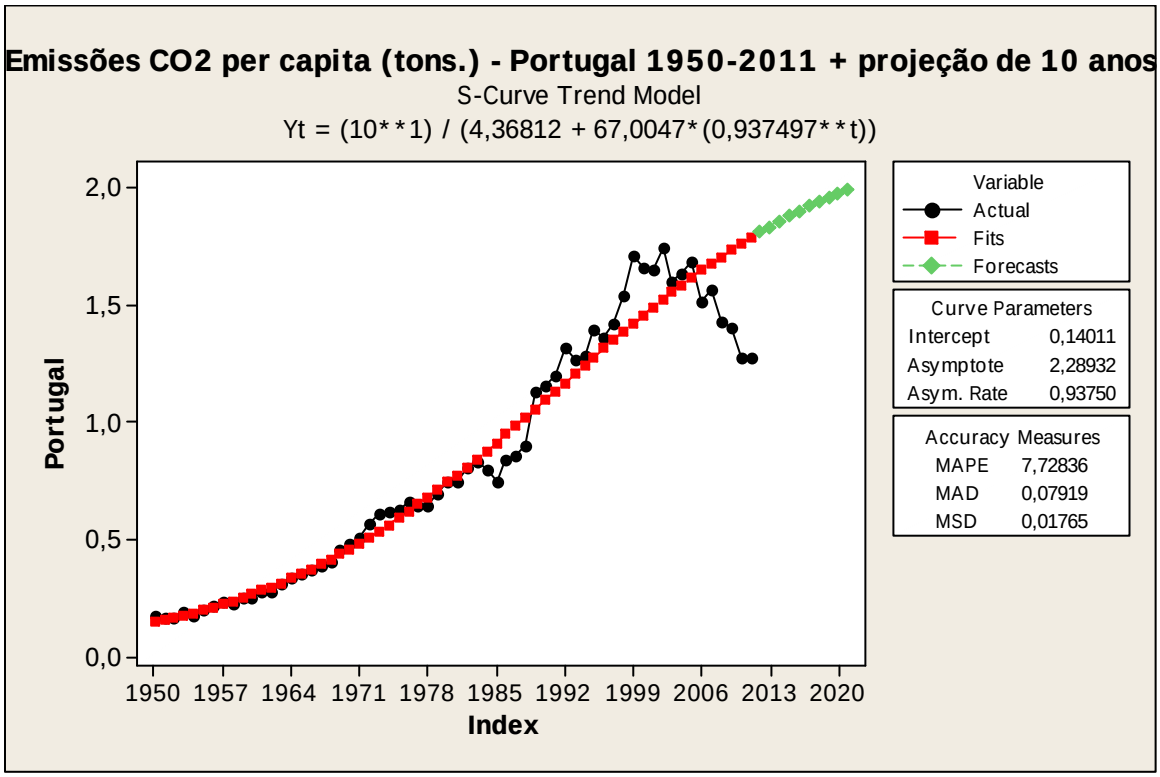
Period	Forecast
2012	1,08992
2013	1,08689
2014	1,08330
2015	1,07913
2016	1,07439
2017	1,06909
2018	1,06321
2019	1,05677
2020	1,04975
2021	1,04217

3. Análise de tendência e projeções – Portugal


Emissões CO2 per capita (tons.) - Portugal 1950-2011 + projeção de 10 anos
Linear Trend Model
Yt = -0,0282 + 0,0276*t
Portugal
Index
2,0
1,5
1,0
0,5
0,0
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 17,8371
MAD 0,1146
MSD 0,0211


Emissões CO2 per capita (tons.) - Portugal 1950-2011 + projeção de 10 anos
Quadratic Trend Model
Yt = 0,0219 + 0,02289*t + 0,000075*t**2
Portugal
Index
2,0
1,5
1,0
0,5
0,0
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 14,9029
MAD 0,1072
MSD 0,0206





Verifica-se que as emissões em Portugal decrescem desde o final dos anos noventa, ao contrário de todo o período anterior dentro das observações disponíveis. Os indicadores são apresentados a seguir. Como se observa, a função de Curva S é a que apresenta melhor adequação, por possuir o menor menor MAD (desvio médio das medianas)

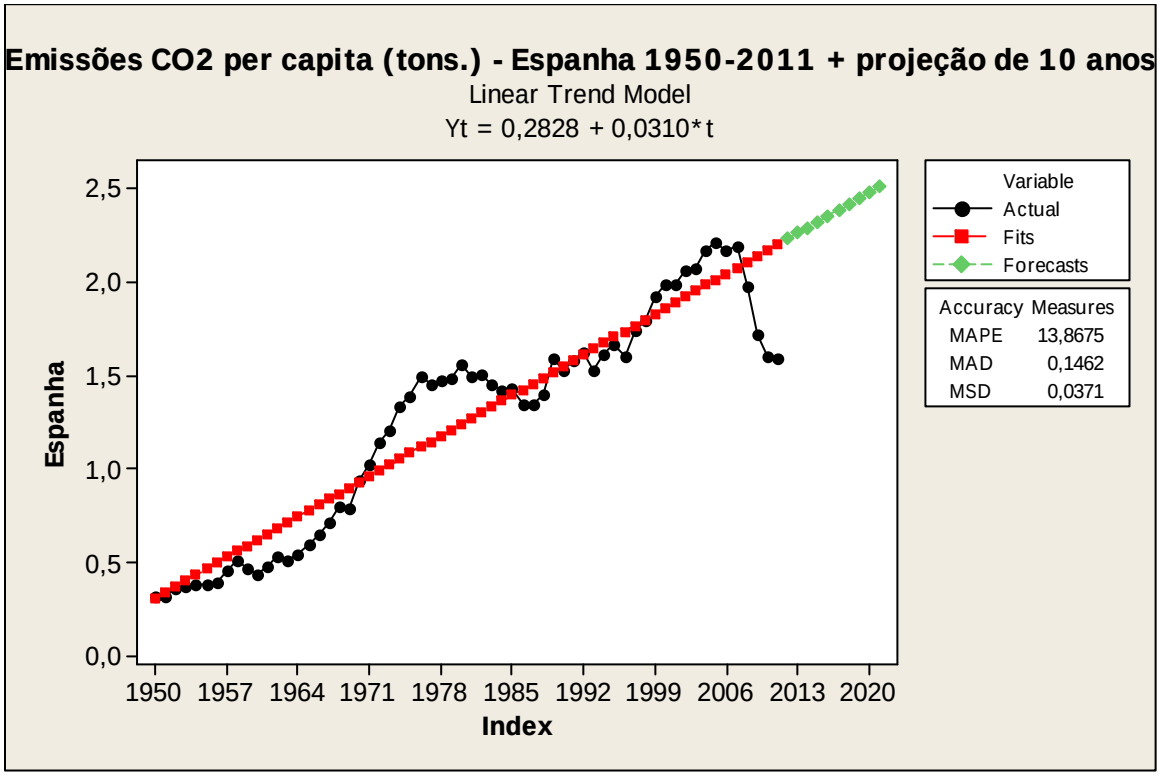
Medida de ajuste nos modelos – Portugal

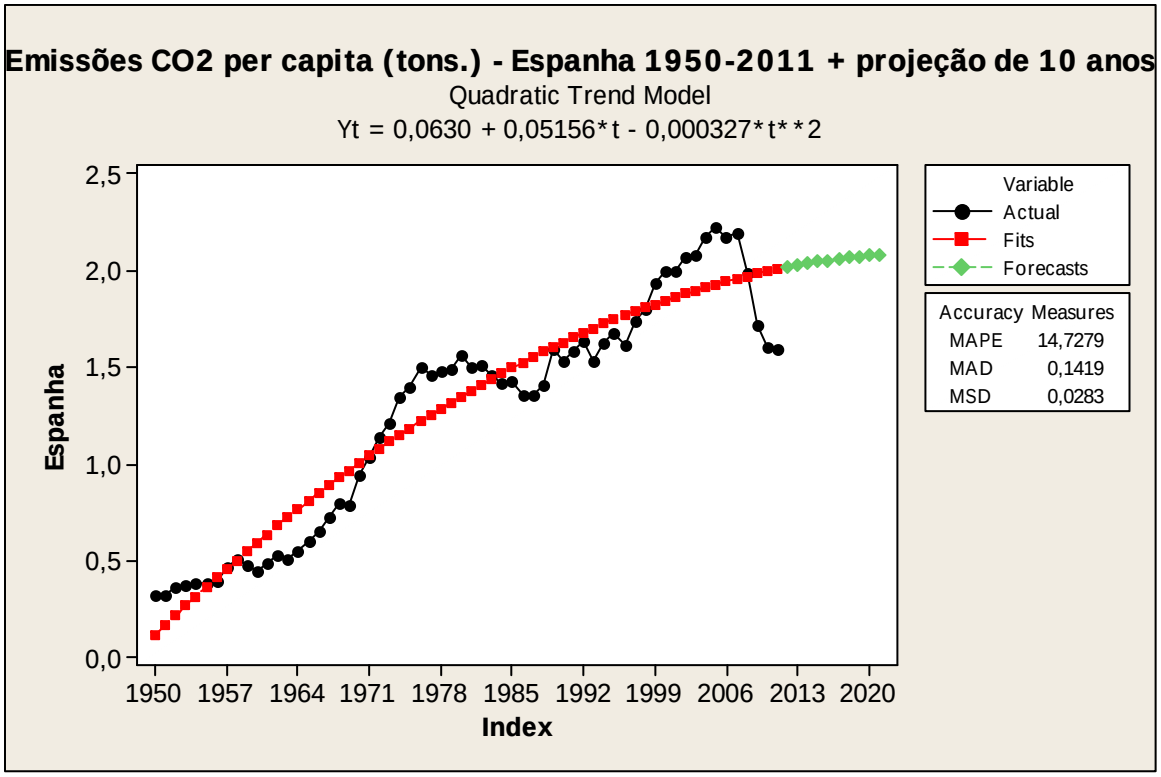
	PT - LI	PT - QD	PT - XP	PT - S
MAPE	17,8371	14,9029	14,4632	7,72836
MAD	0,1146	0,1072	0,1365	**0,07919**
MSD	0,0211	0,0206	0,0556	0,01765

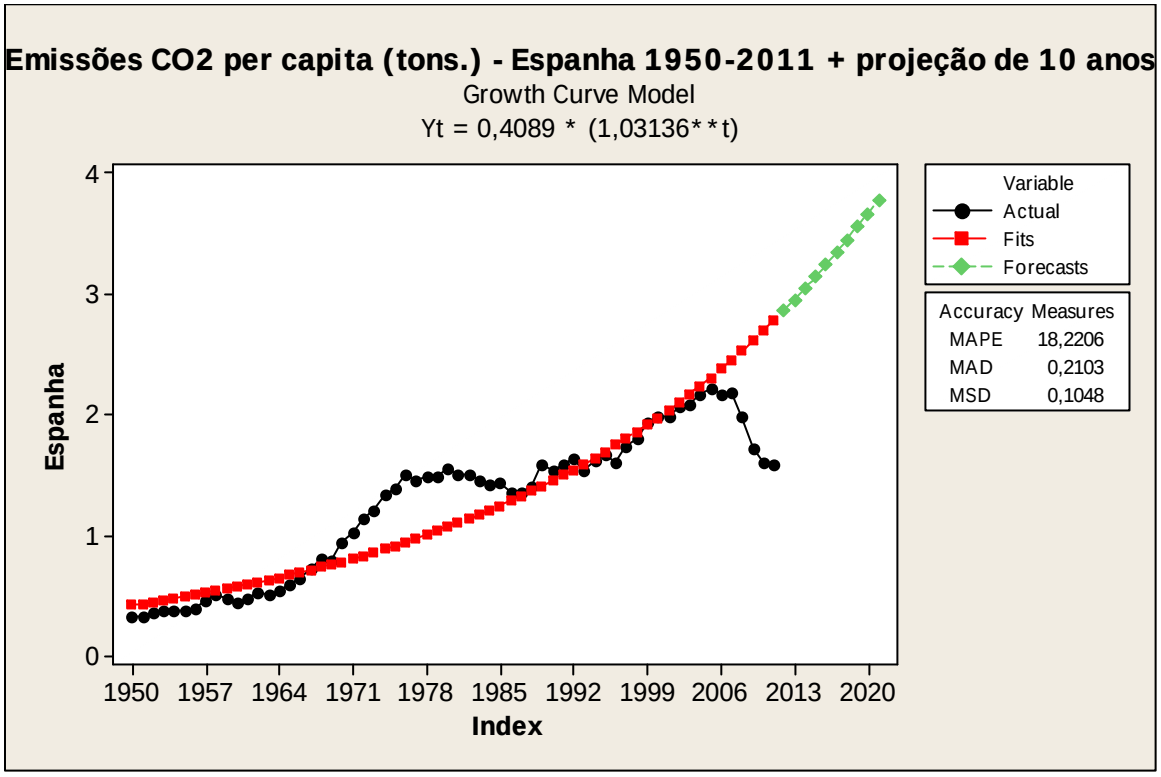
As projeções de dez anos do melhor modelo são apresentadas a seguir.

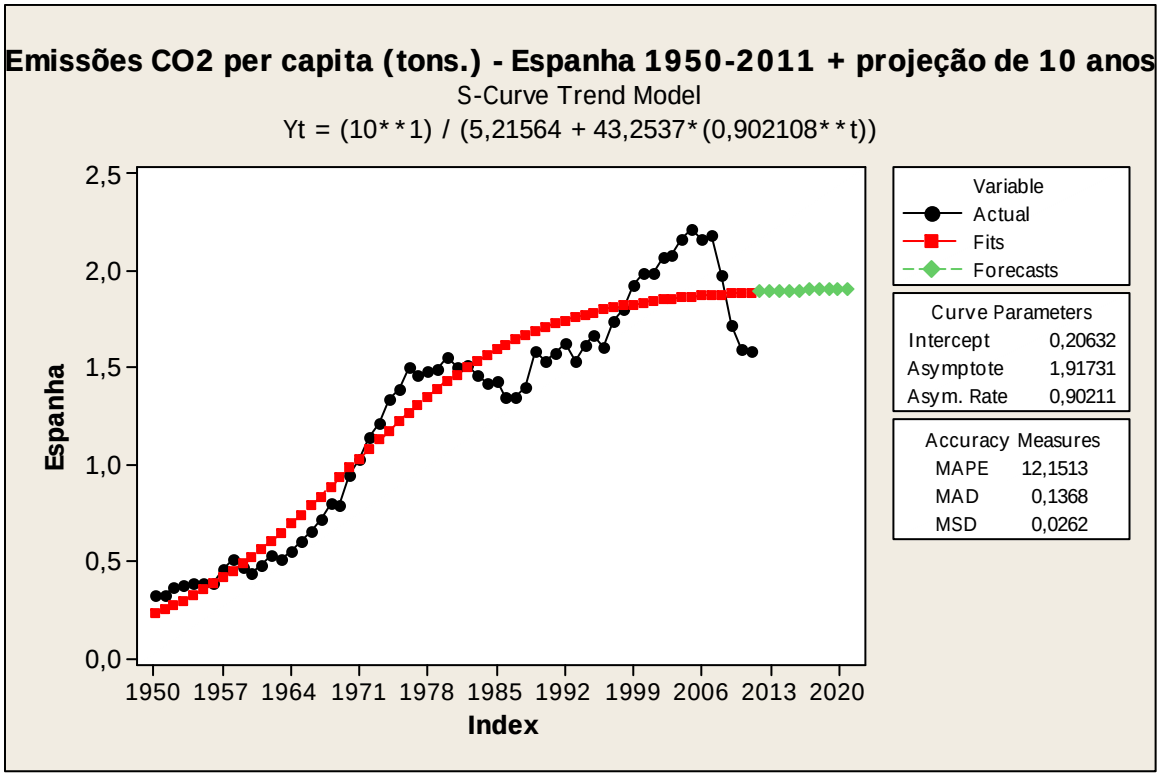
Period	Forecast
2012	1,81265
2013	1,83655
2014	1,85953
2015	1,88161
2016	1,90279
2017	1,92309
2018	1,94251
2019	1,96108
2020	1,97881
2021	1,99573

4. Análise de tendência e projeções – Espanha




Emissões CO2 per capita (tons.) - Espanha 1950-2011 + projeção de 10 anos
Quadratic Trend Model
Yt = 0,0630 + 0,05156*t - 0,000327*t**2
Espanha
Index
2,5
2,0
1,5
1,0
0,5
0,0
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 14,7279
MAD 0,1419
MSD 0,0283


Emissões CO2 per capita (tons.) - Espanha 1950-2011 + projeção de 10 anos
Growth Curve Model
Yt = 0,4089 * (1,03136**t)
Espanha
Index
4
3
2
1
0
1950 1957 1964 1971 1978 1985 1992 1999 2006 2013 2020
Variable
Actual
Fits
Forecasts
Accuracy Measures
MAPE 18,2206
MAD 0,2103
MSD 0,1048



Verifica-se que as emissões na Espanha reduzem-se a partir dos anos de 2005-6, em tendência, para o período analisado.

Com relação aos modelos, a função de Curva S é a que melhor representa a série de dados. Os indicadores são apresentados a seguir.

Medida de ajuste nos modelos – Espanha

	ES - LI	ES - QD	ES - XP	ES - S
MAPE	13,8675	14,7279	18,2206	12,1513
MAD	0,1462	0,1419	0,2103	**0,1368**
MSD	0,0371	0,0283	0,1048	0,0262

As projeções para dez anos do modelo com melhor adequação são apresentadas a seguir.

Forecasts

Period	Forecast
2012	1,89347
2013	1,89578
2014	1,89787
2015	1,89975
2016	1,90146
2017	1,90300
2018	1,90439
2019	1,90565
2020	1,90678
2021	1,90781

5. ANOVA E SIMILARIDADE

Este trabalho tem como objetivo analisar a similaridade ou dessemelhança entre agrupamentos de países através da aplicação de teste de análise de variância ANOVA *one way* para as variáveis do EIXO PESSOAS. As variáveis foram empregadas em seu formato normalizado com escala entre zero e um, e são apresentadas na Tabela 1. Os 132 países da amostra são analisados sob o ponto de vista de dois agrupamentos, um mais extenso, com oito regiões (6R) e outro agrupado em apenas três regiões (3R) , para destacar a comparação entre os países Ibero – americanos, países da conforme a tabela 2 e 3 a seguir. A análise baseia-se nos resultados numéricos e gráficos dos testes, apresentados no decorrer do trabalho. Uma tabela consolidada das variáveis conclui o estudo, indicando qual(is) variável(is) explica(m) melhor a(s) diferença(s) entre as regiões analisadas. Empregou-se o Minitab ® para as análises estatísticas.

Tabela 2 - Classificação Regional 8 e 3 países.

Classificação - 8 Regiões	Classificção 3 Regiões
AIBER – região Ibero-americana	AIBER – região Ibero-americana
AVECO – Economias avançadas: região da Europa Ocidental, países anglo-saxônicos, Coreia e Japão	AVECO – Economias avançadas: região da Europa Ocidental, países anglo-saxônicos, Coreia e Japão
CWEAS – Europa oriental	
CARLA – região do Caribe	
EMDEU – norte da Europa	
ESEAS – sudeste da Ásia	OUTROS
NSWUA – sul da Ásia	
SUSAF – África Subsaariana	

Preparado pelo autor com dados do Projeto Oriber e participantes das aulas de Métodos Quantitativos

Tabela 3 – Países por Região

No.	Country	Class 6R	N Class 3R
1	Argentina	AIBER	AIBER
2	Bolivia	AIBER	AIBER
3	Brazil	AIBER	AIBER
4	Chile	AIBER	AIBER
5	Colombia	AIBER	AIBER
6	Costa Rica	AIBER	AIBER
7	Dominican Republic	AIBER	AIBER
8	Ecuador	AIBER	AIBER
9	El Salvador	AIBER	AIBER
10	Guatemala	AIBER	AIBER
11	Honduras	AIBER	AIBER
12	Mexico	AIBER	AIBER
13	Nicaragua	AIBER	AIBER
14	Panama	AIBER	AIBER
15	Paraguay	AIBER	AIBER
16	Peru	AIBER	AIBER

17	Portugal	AIBER	AIBER
18	Spain	AIBER	AIBER
19	Uruguay	AIBER	AIBER
20	Venezuela	AIBER	AIBER
21	Australia	AVECO	AVECO
22	Austria	AVECO	AVECO
23	Belgium	AVECO	AVECO
24	Canada	AVECO	AVECO
25	Czech Republic	AVECO	AVECO
26	Denmark	AVECO	AVECO
27	Estonia	AVECO	AVECO
28	Finland	AVECO	AVECO
29	France	AVECO	AVECO
30	Germany	AVECO	AVECO
31	Greece	AVECO	AVECO
32	Iceland	AVECO	AVECO
33	Ireland	AVECO	AVECO
34	Israel	AVECO	AVECO
35	Italy	AVECO	AVECO
36	Japan	AVECO	AVECO
37	Korea, Republic of	AVECO	AVECO
38	Latvia	AVECO	AVECO
39	Netherlands	AVECO	AVECO
40	New Zealand	AVECO	AVECO
41	Norway	AVECO	AVECO
42	Slovakia	AVECO	AVECO
43	Slovenia	AVECO	AVECO
44	Sweden	AVECO	AVECO
45	Switzerland	AVECO	AVECO
46	United Kingdom	AVECO	AVECO
47	United States	AVECO	AVECO
48	Cuba	CARLA	AIBER
49	Guyana	CARLA	OTHERS
50	Jamaica	CARLA	OTHERS
51	Trinidad and Tobago	CARLA	OTHERS
52	Armenia	CWEAS	OTHERS
53	Azerbaijan	CWEAS	OTHERS
54	Belarus	CWEAS	OTHERS
55	Georgia	CWEAS	OTHERS
56	Kazakhstan	CWEAS	OTHERS
57	Kyrgyzstan	CWEAS	OTHERS
58	Moldova	CWEAS	OTHERS
59	Russia	CWEAS	OTHERS
60	Tajikistan	CWEAS	OTHERS
61	Ukraine	CWEAS	OTHERS
62	Uzbekistan	CWEAS	OTHERS

63	Albania	EMDEU	OTHERS
64	Bosnia and Herzegovina	EMDEU	OTHERS
65	Bulgaria	EMDEU	OTHERS
66	Croatia	EMDEU	OTHERS
67	Hungary	EMDEU	OTHERS
68	Lithuania	EMDEU	OTHERS
69	Macedonia	EMDEU	OTHERS
70	Montenegro	EMDEU	OTHERS
71	Poland	EMDEU	OTHERS
72	Romania	EMDEU	OTHERS
73	Serbia	EMDEU	OTHERS
74	Turkey	EMDEU	OTHERS
75	Bangladesh	ESEAS	OTHERS
76	Cambodia	ESEAS	OTHERS
77	China	ESEAS	OTHERS
78	India	ESEAS	OTHERS
79	Indonesia	ESEAS	OTHERS
80	Laos	ESEAS	OTHERS
81	Malaysia	ESEAS	OTHERS
82	Mongolia	ESEAS	OTHERS
83	Nepal	ESEAS	OTHERS
84	Philippines	ESEAS	OTHERS
85	Sri Lanka	ESEAS	OTHERS
86	Thailand	ESEAS	OTHERS
87	Algeria	NSWUA	OTHERS
88	Egypt	NSWUA	OTHERS
89	Iran	NSWUA	OTHERS
90	Iraq	NSWUA	OTHERS
91	Jordan	NSWUA	OTHERS
92	Kuwait	NSWUA	OTHERS
93	Lebanon	NSWUA	OTHERS
94	Mauritania	NSWUA	OTHERS
95	Morocco	NSWUA	OTHERS
96	Pakistan	NSWUA	OTHERS
97	Saudi Arabia	NSWUA	OTHERS
98	Tunisia	NSWUA	OTHERS
99	United Arab Emirates	NSWUA	OTHERS
100	Yemen	NSWUA	OTHERS
101	Angola	SUSAF	OTHERS
102	Benin	SUSAF	OTHERS
103	Botswana	SUSAF	OTHERS
104	Burkina Faso	SUSAF	OTHERS
105	Burundi	SUSAF	OTHERS
106	Cameroon	SUSAF	OTHERS
107	Central African Republic	SUSAF	OTHERS
108	Chad	SUSAF	OTHERS

109	Congo, Republic of	SUSAF	OTHERS
110	Djibouti	SUSAF	OTHERS
111	Ghana	SUSAF	OTHERS
112	Guinea	SUSAF	OTHERS
113	Kenya	SUSAF	OTHERS
114	Lesotho	SUSAF	OTHERS
115	Liberia	SUSAF	OTHERS
116	Madagascar	SUSAF	OTHERS
117	Malawi	SUSAF	OTHERS
118	Mali	SUSAF	OTHERS
119	Mauritius	SUSAF	OTHERS
120	Mozambique	SUSAF	OTHERS
121	Namibia	SUSAF	OTHERS
122	Niger	SUSAF	OTHERS
123	Nigeria	SUSAF	OTHERS
124	Rwanda	SUSAF	OTHERS
125	Senegal	SUSAF	OTHERS
126	South Africa	SUSAF	OTHERS
127	Sudan	SUSAF	OTHERS
128	Swaziland	SUSAF	OTHERS
129	Tanzania	SUSAF	OTHERS
130	Togo	SUSAF	OTHERS
131	Uganda	SUSAF	OTHERS
132	Zambia	SUSAF	OTHERS

Preparado pelo autor com dados do Projeto Oriber e participantes das aulas de Métodos Quantitativos

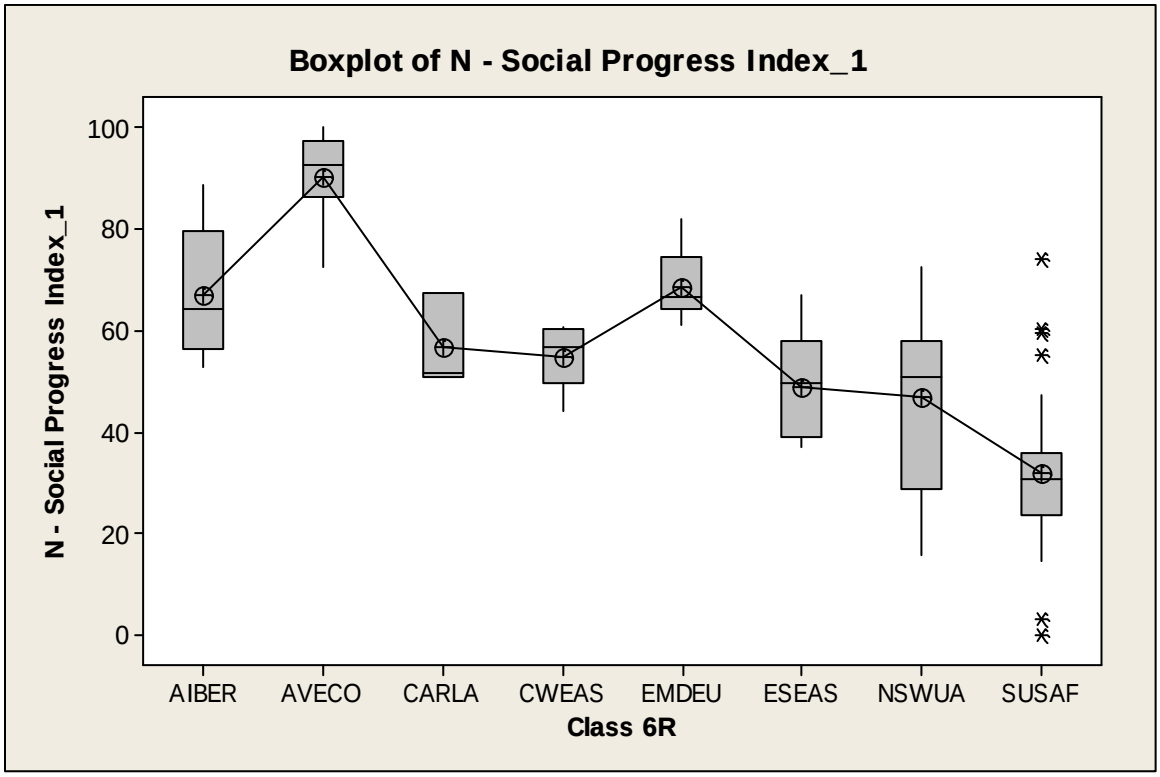
a. ANOVA DAS VARIÁVEIS DO EIXTO PESSOAS

———— 29/03/2016 13:39:15 ————————————————————

```
Welcome to Minitab, press F1 for help.
```

Results for: CODING REGIONS MINITAB.MTW

Boxplot of N - Social Progress Index_1



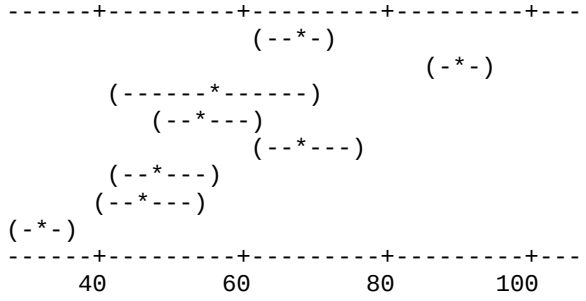
One-way ANOVA: N - Social Progress Index_1 versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	54628	7804	53,48	0,000
Error	121	17657	146		
Total	128	72285			

S = 12,08 R-Sq = 75,57% R-Sq(adj) = 74,16%

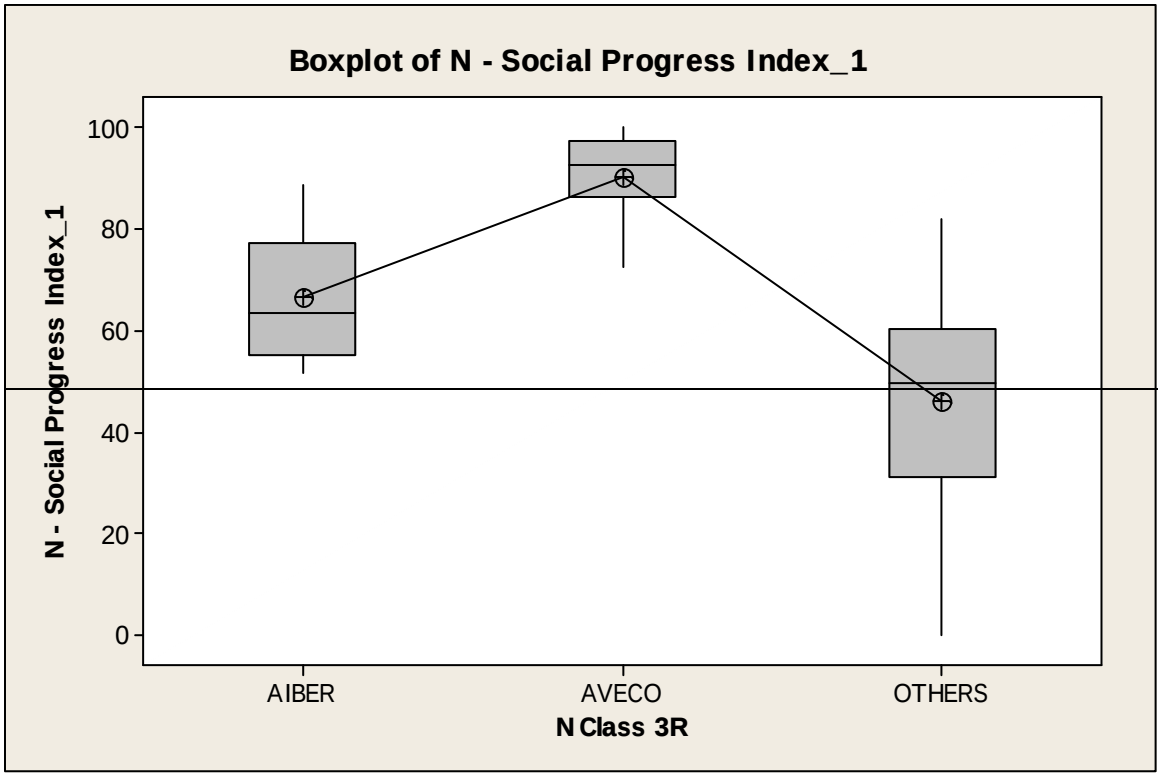
Level	N	Mean	StDev
AIBER	20	67,10	12,03
AVECO	27	90,31	8,14
CARLA	3	56,68	9,34
CWEAS	11	54,72	5,59
EMDEU	12	68,61	6,52
ESEAS	12	48,83	10,12
NSWUA	14	46,86	16,84
SUSAF	30	31,78	16,00

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev = 12,08

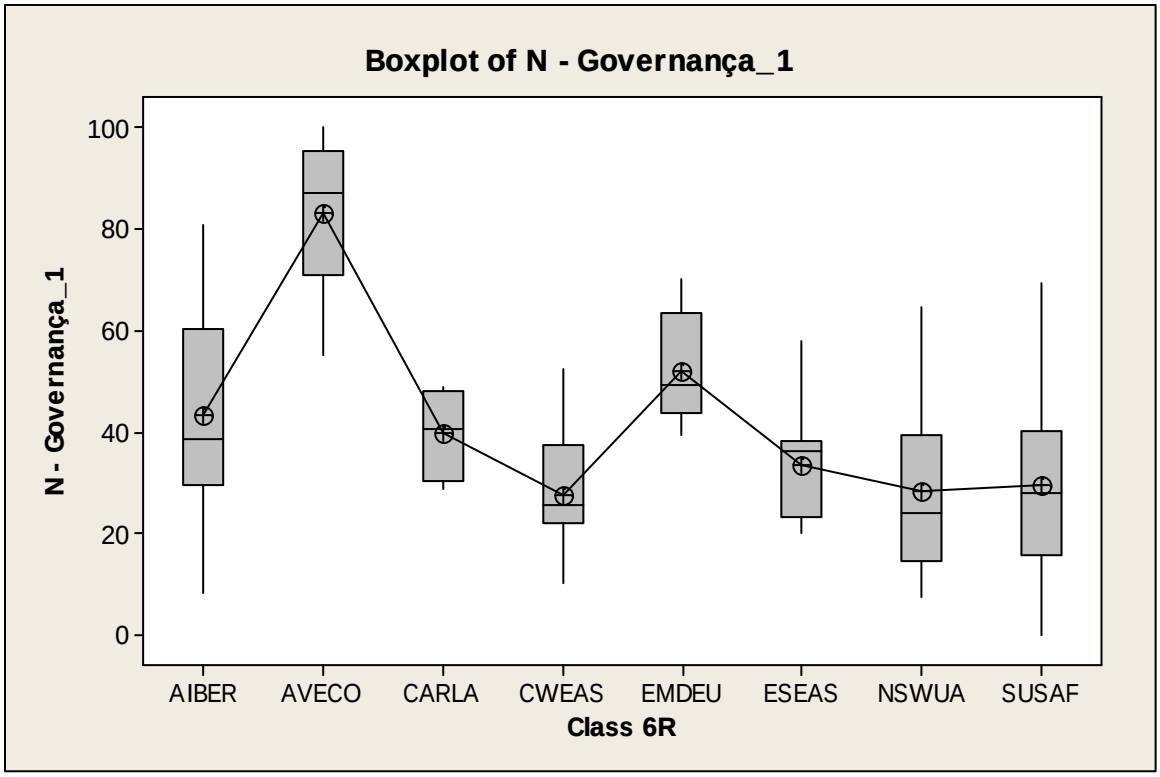
Boxplot of N - Social Progress Index_1



One-way ANOVA: N - Social Progress Index_1 versus N Class 3R

```
Source       DF     SS     MS      F      P

N Class 3R    2  40950  20475  82,33  0,000
Error       126  31335    249
Total       128  72285

S = 15,77   R-Sq = 56,65%   R-Sq(adj) = 55,96%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --+---------+---------+---------+-------
AIBER   21  66,37  12,20             (---*----)
AVECO   27  90,31   8,14                                (---*---)
OTHERS  81  46,16  18,25  (--*-)
                          --+---------+---------+---------+-------
                           45        60        75        90

Pooled StDev = 15,77
```

Boxplot of N - Governança_1 versus Class 6R

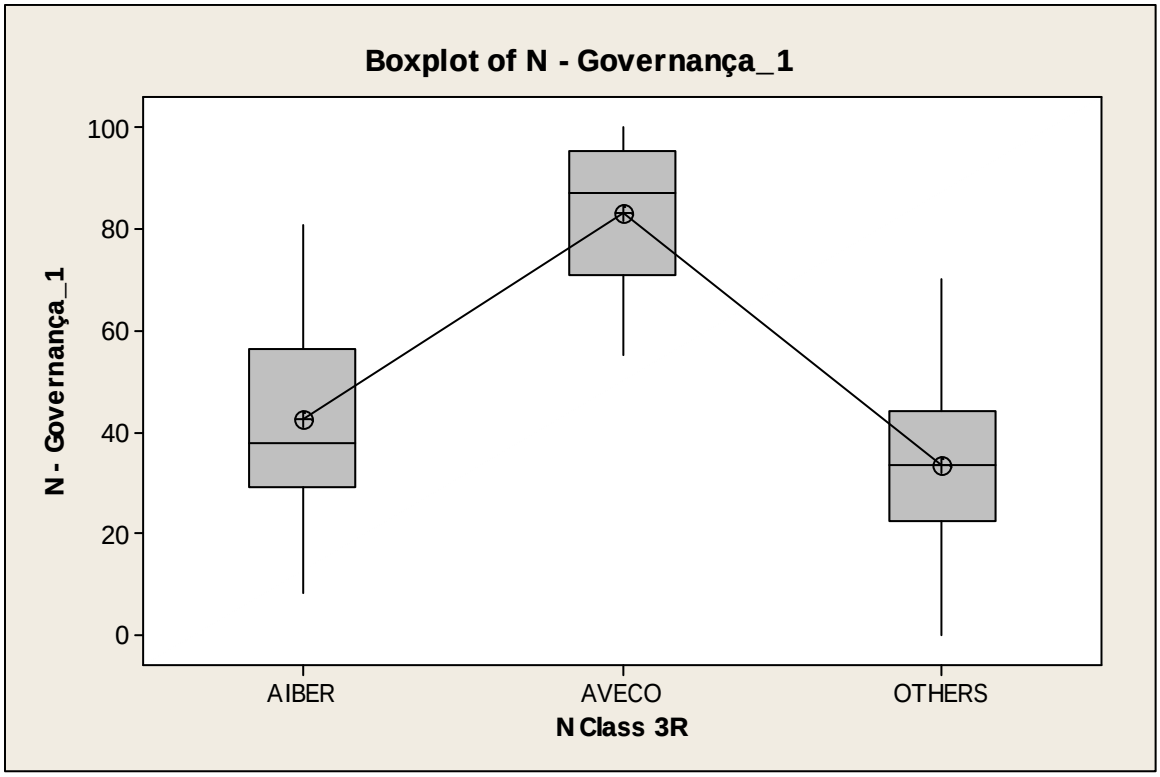


One-way ANOVA: N - Governança_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56677  8097  35,53  0,000
Error     124  28261   228
Total     131  84938

S = 15,10   R-Sq = 66,73%   R-Sq(adj) = 64,85%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  43,17  19,04            (---*--)
AVECO  27  83,25  13,38                                (--*--)
CARLA   4  39,62   9,54      (-------*------)
CWEAS  11  27,38  12,38   (----*---)
EMDEU  12  51,93  11,21                 (---*---)
ESEAS  12  33,61  10,47      (----*---)
NSWUA  14  28,46  16,21    (---*---)
SUSAF  32  29,58  16,95      (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,10
```

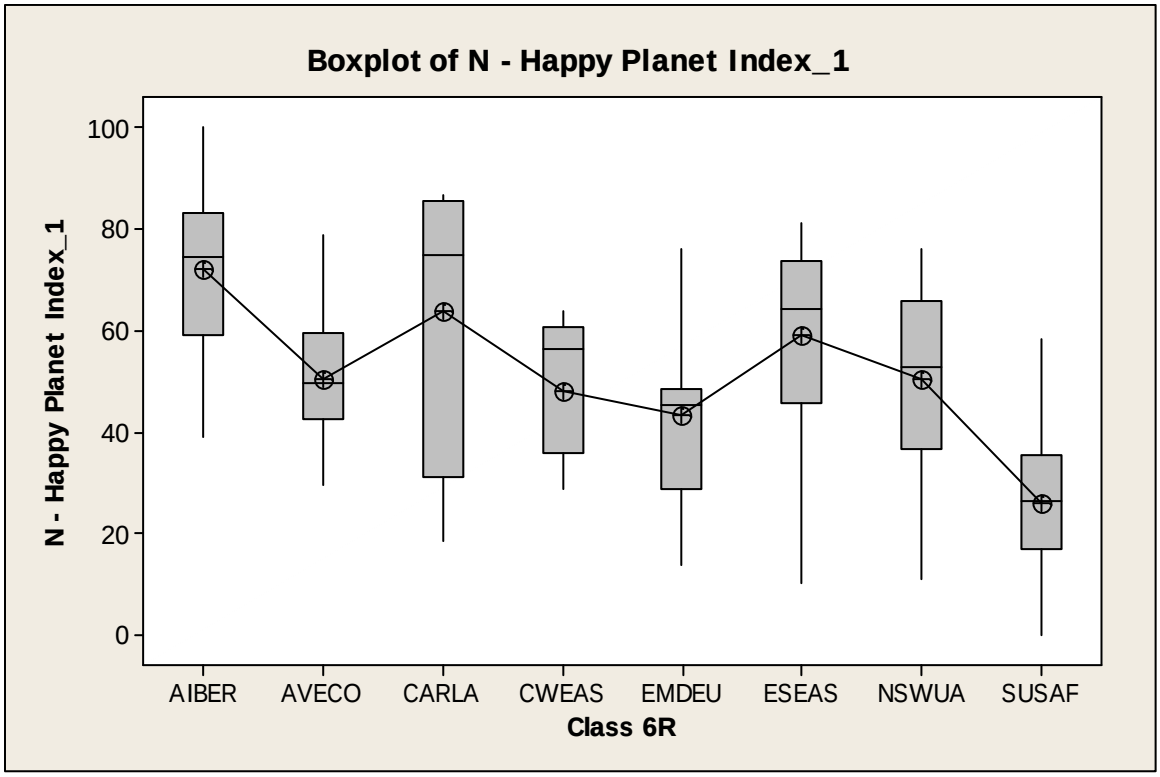
Boxplot of N - Governança_1



One-way ANOVA: N - Governança_1 versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  51019  25510  97,02  0,000
Error       129  33919    263
Total       131  84938

S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42,48  18,82       (---*----)
AVECO   27  83,25  13,38                                 (----*---)
OTHERS  84  33,36  16,35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75

Pooled StDev = 16,22
```

Boxplot of N - Happy Planet Index_1

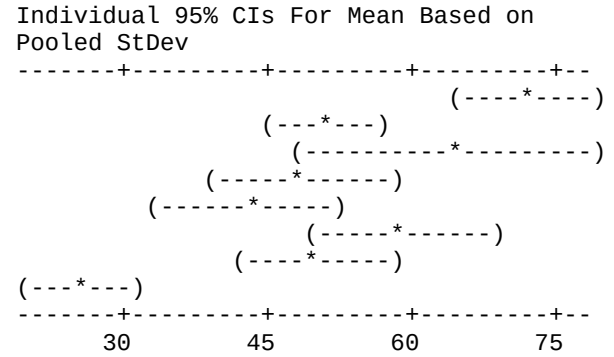


One-way ANOVA: N - Happy Planet Index_1 versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	29113	4159	15,83	0,000
Error	121	31781	263		
Total	128	60894			

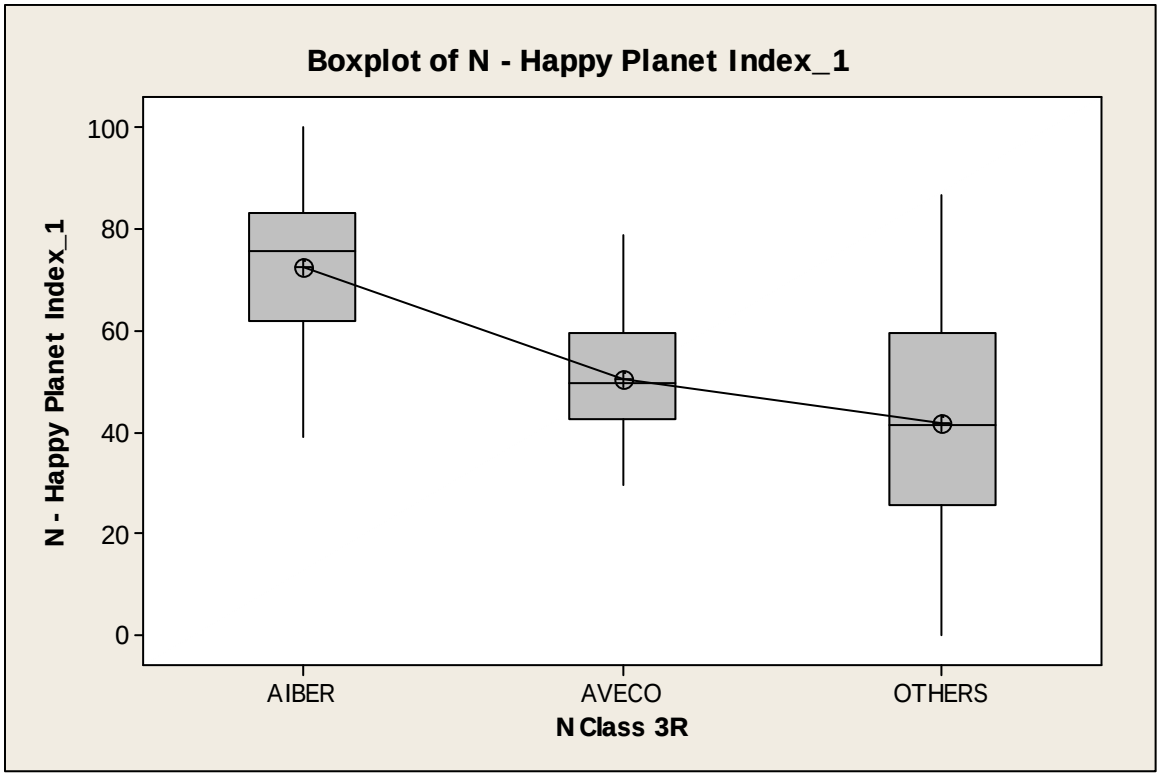
S = 16,21 R-Sq = 47,81% R-Sq(adj) = 44,79%

Level	N	Mean	StDev
AIBER	20	71,92	16,55
AVECO	27	50,58	12,84
CARLA	4	63,81	31,09
CWEAS	11	48,13	13,54
EMDEU	11	43,12	16,72
ESEAS	12	59,06	20,02
NSWUA	14	50,22	19,91
SUSAF	30	25,85	13,39



Pooled StDev = 16,21

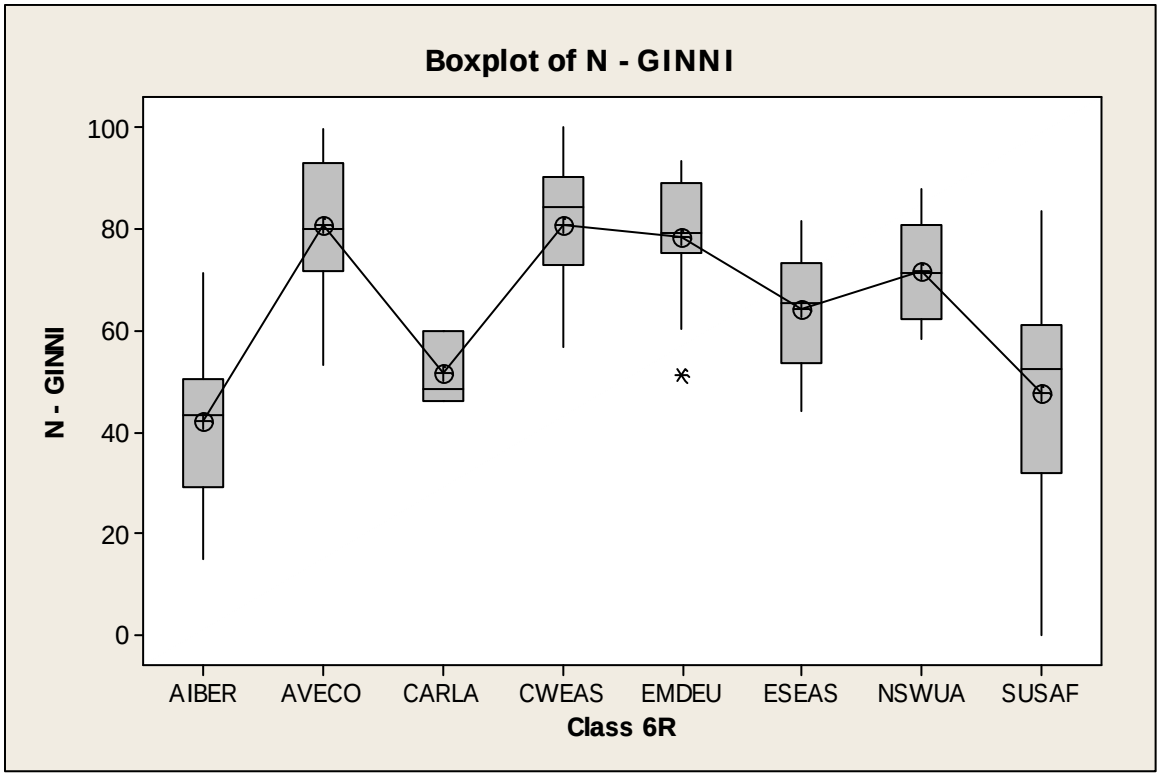
Boxplot of N - Happy Planet Index_1



One-way ANOVA: N - Happy Planet Index_1 versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  15984  7992  22,42  0,000
Error       126  44910   356
Total       128  60894

S = 18,88   R-Sq = 26,25%   R-Sq(adj) = 25,08%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72,36  16,25                          (-----*------)
AVECO   27  50,58  12,84       (-----*-----)
OTHERS  81  41,55  21,02  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18,88
```

Boxplot of N - GINNI versus Class 6R

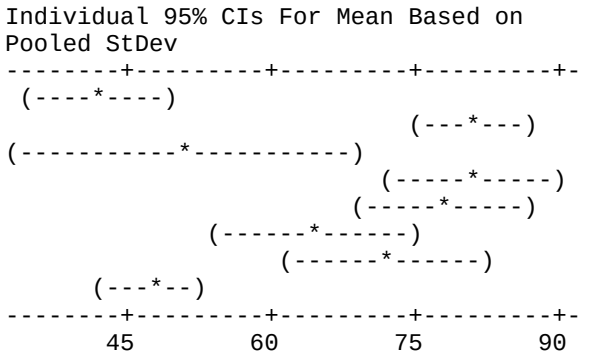


One-way ANOVA: N - GINNI versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	31043	4435	17,93	0,000
Error	111	27449	247		
Total	118	58492			

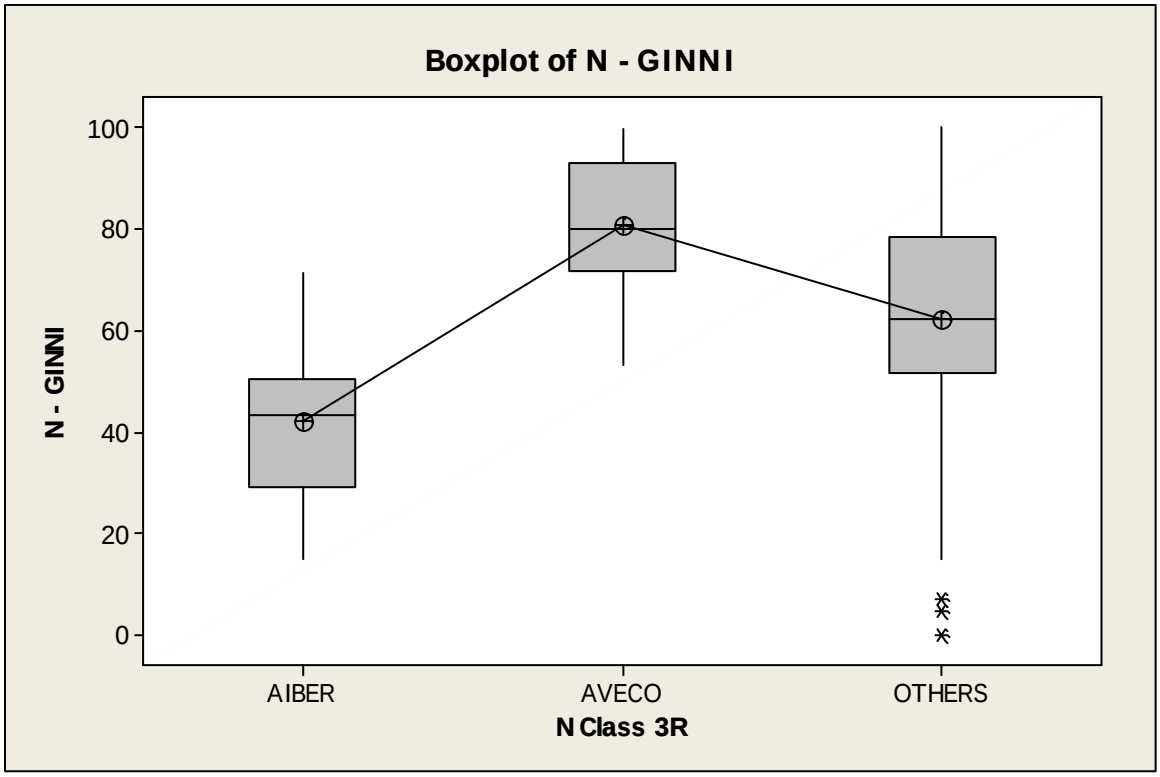
S = 15,73 R-Sq = 53,07% R-Sq(adj) = 50,11%

Level	N	Mean	StDev
AIBER	20	42,09	14,02
AVECO	26	80,57	12,19
CARLA	3	51,41	7,28
CWEAS	11	80,74	13,30
EMDEU	11	78,22	12,76
ESEAS	9	64,08	11,97
NSWUA	9	71,80	10,22
SUSAF	30	47,56	22,38



Pooled StDev = 15,73

Boxplot of N - GINNI



One-way ANOVA: N - GINNI versus N Class 3R

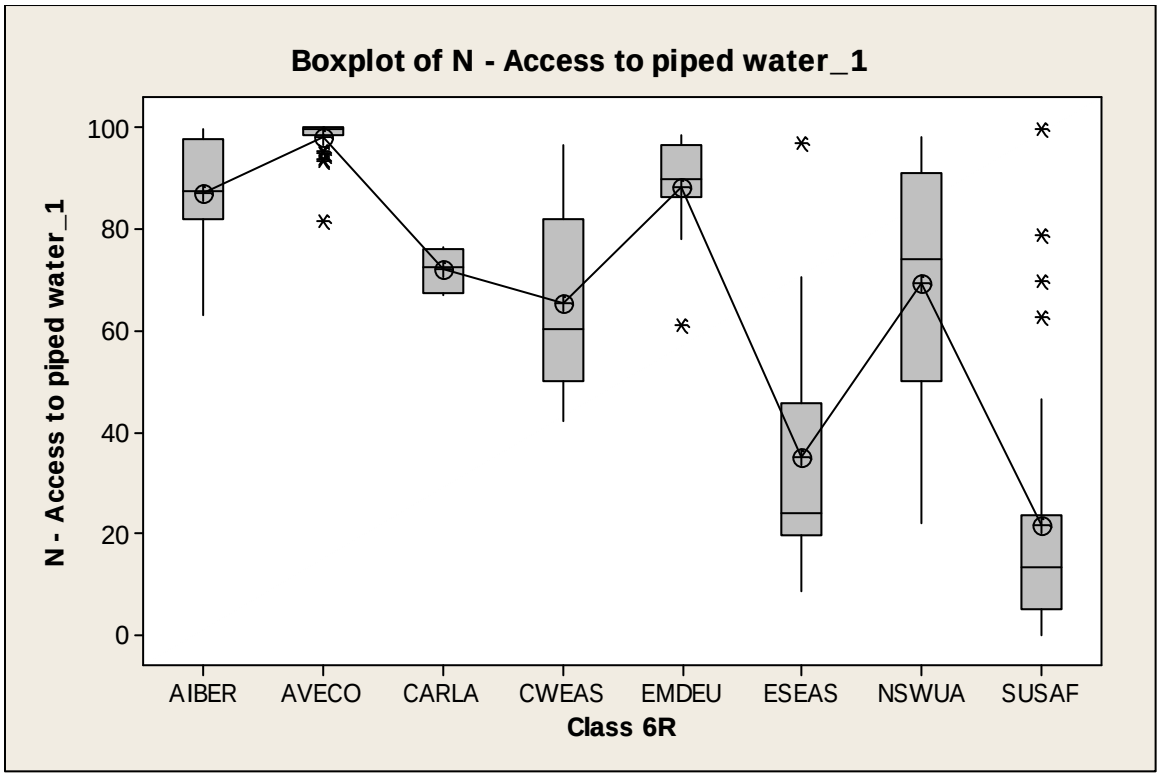
```
Source       DF     SS    MS      F      P
N Class 3R    2  16795  8397  23,36  0,000
Error       116  41698   359
Total       118  58492

S = 18,96   R-Sq = 28,71%   R-Sq(adj) = 27,48%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   20  42,09  14,02  (-----*-----)
AVECO   26  80,57  12,19                          (----*----)
OTHERS  73  62,36  21,81                  (--*--)
                          --------+---------+---------+---------+-
                                 45        60        75        90

Pooled StDev = 18,96
```

One-way ANOVA: N - Access to piped water_1 versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7  115010  16430  48,67  0,000
Error     121   40844    338
Total     128  155854

S = 18,37   R-Sq = 73,79%   R-Sq(adj) = 72,28%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
```

```
AIBER  20  87,07  10,91                          (--*--)
AVECO  25  98,12   4,04                             (--*--)
CARLA   4  72,09   4,54                  (------*------)
CWEAS  11  65,26  18,29                  (---*---)
EMDEU  12  88,23  10,36                           (---*---)
ESEAS  12  34,94  25,62       (---*---)
NSWUA  13  69,20  24,22                    (---*---)
SUSAF  32  21,63  24,98  (--*-)
                         ----+---------+---------+---------+-----
                            25        50        75       100

Pooled StDev = 18,37
```

Boxplot of N - Access to piped water_1

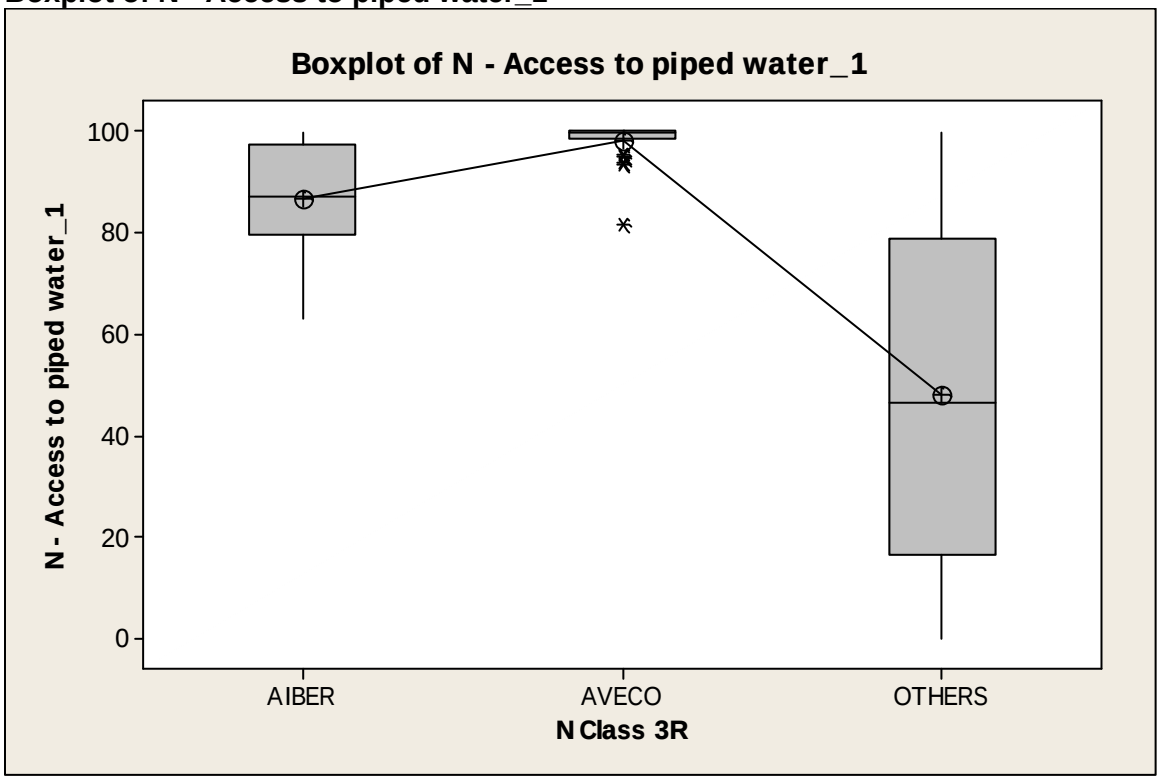


One-way ANOVA: N - Access to piped water_1 versus N Class 3R

```
Source       DF      SS     MS      F      P
N Class 3R    2   60546  30273  40,02  0,000
Error       126   95308    756
Total       128  155854

S = 27,50   R-Sq = 38,85%   R-Sq(adj) = 37,88%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  86,56  10,88                  (-----*-----)
AVECO   25  98,12   4,04                        (----*-----)
OTHERS  83  48,19  33,59  (--*--)
                          ---------+---------+---------+---------+
                                  60        80       100       120

Pooled StDev = 27,50
```

Boxplot of N - Access to piped water_1

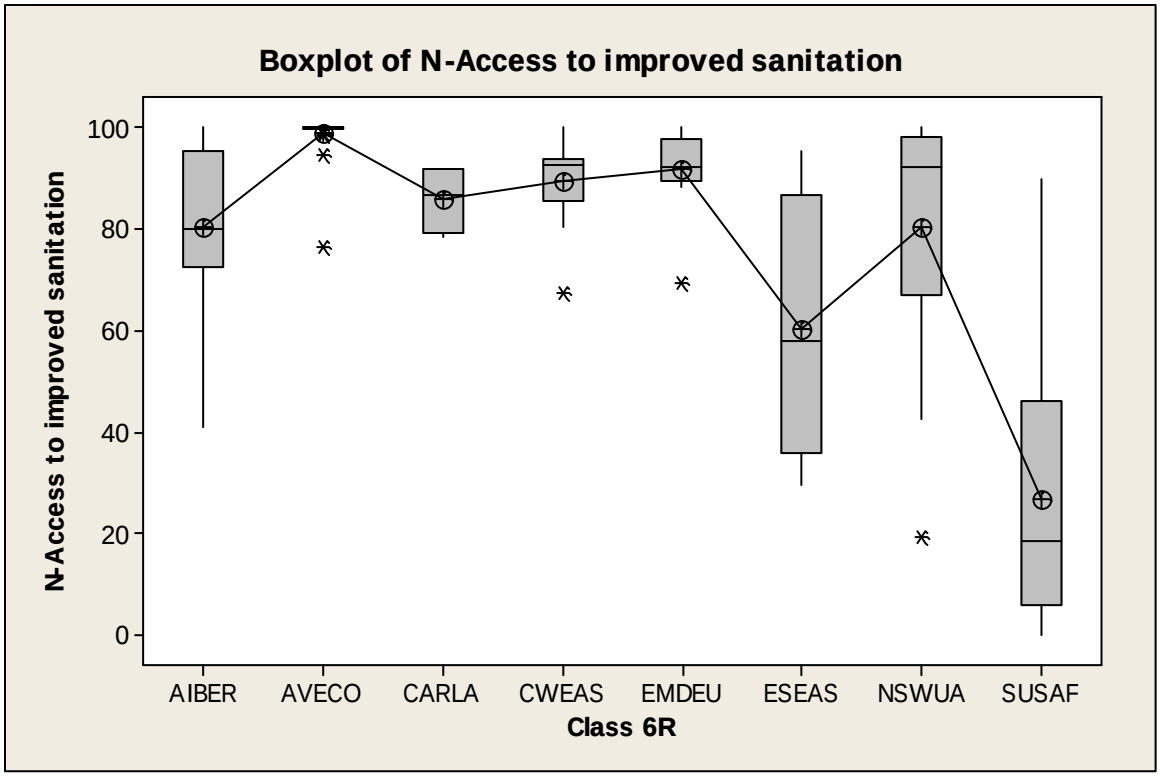


Outliers: 32, 34, 68, 64

One-way ANOVA: N-Access to improved sanitation versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7   96250  13750  41,94  0,000
Error     122   40002    328
Total     129  136252

S = 18,11   R-Sq = 70,64%   R-Sq(adj) = 68,96%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
AIBER  20  80,46  16,09                    (--*--)
AVECO  25  98,72   4,77                           (-*--)
CARLA   4  85,85   6,80                     (------*-------)
CWEAS  11  89,53   9,06                         (----*---)
EMDEU  12  91,82   8,23                           (---*---)
ESEAS  12  60,34  23,86               (---*---)
NSWUA  14  80,56  25,55                        (---*---)
SUSAF  32  26,72  24,15  (--*-)
                          --+---------+---------+---------+-------
                           25        50        75       100

Pooled StDev = 18,11
```

Boxplot of N-Access to improved sanitation

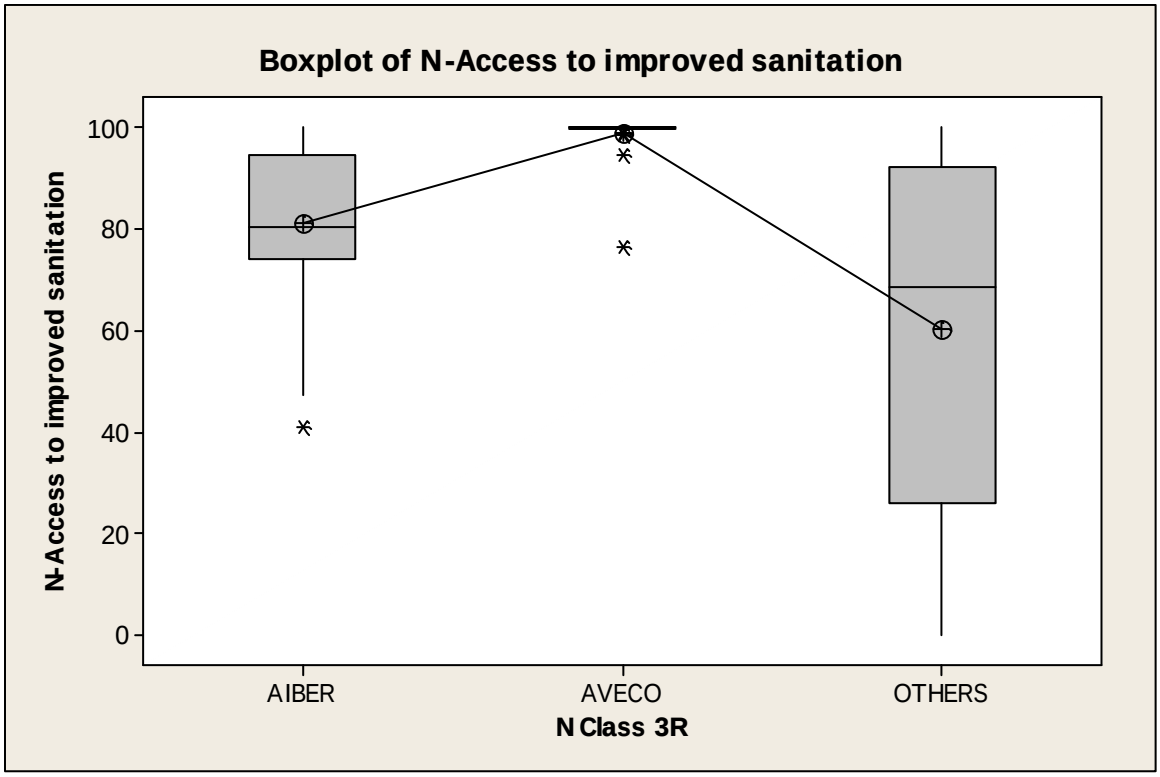


One-way ANOVA: N-Access to improved sanitation versus N Class 3R

```
Source       DF      SS     MS      F      P
N Class 3R    2   31359  15680  18,98  0,000
Error       127  104893    826
Total       129  136252

S = 28,74   R-Sq = 23,02%   R-Sq(adj) = 21,80%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  81,00  15,87           (-------*-------)
AVECO   25  98,72   4,77                     (-------*------)
OTHERS  84  60,06  34,59  (---*---)
                          ----+---------+---------+---------+-----
                             60        75        90       105

Pooled StDev = 28,74
```

Boxplot of N-Access to improved sanitation

Boxplot of N-Access to improved sanitation



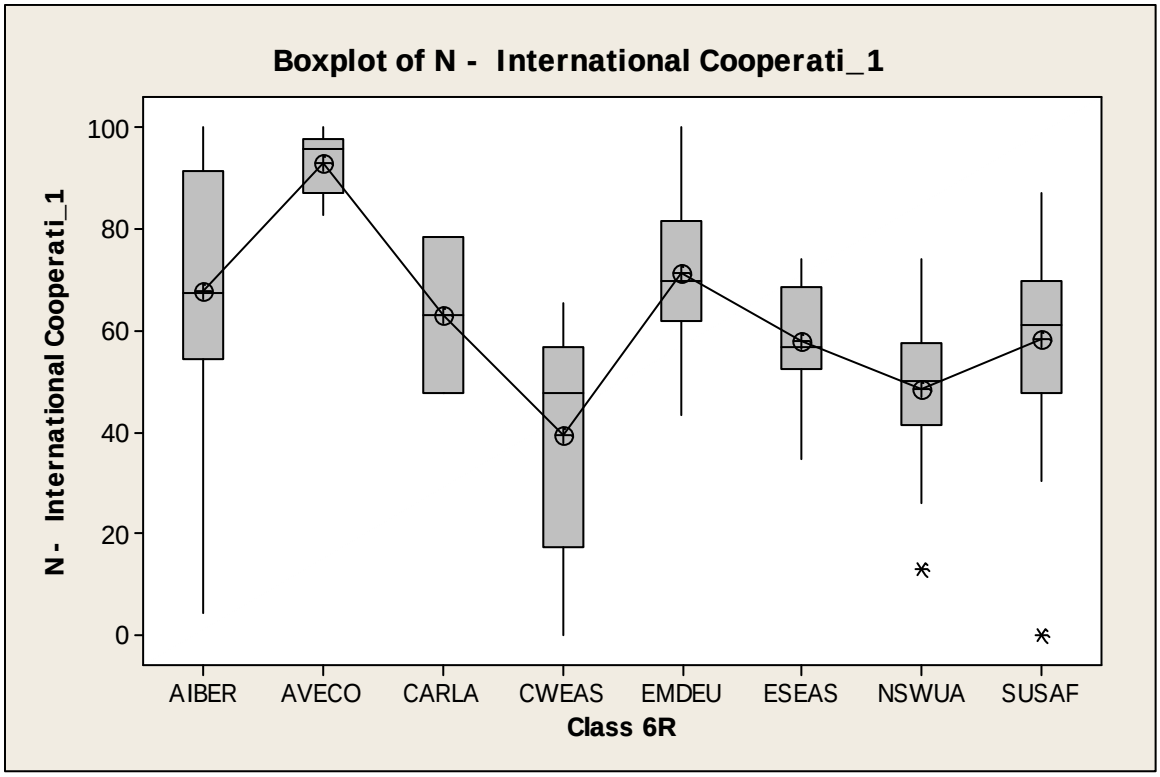
Outlier: Nicarágua

One-way ANOVA: N - International Cooperati_1 versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  14765  2109  6,06  0,000
Error      96  33410   348
Total     103  48175

S = 18,66   R-Sq = 30,65%   R-Sq(adj) = 25,59%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  18  67,87  24,15              (--*---)
AVECO   5  93,04   6,59                     (-----*------)
CARLA   2  63,04  21,52     (---------*----------)
CWEAS  11  39,53  23,37  (----*---)
EMDEU  12  71,38  16,94              (----*---)
ESEAS  12  57,97  12,06         (---*---)
NSWUA  14  48,45  15,71     (---*---)
SUSAF  30  58,26  18,04           (-*--)
                          ---------+---------+---------+---------+
                                  50        75       100       125

Pooled StDev = 18,66
```

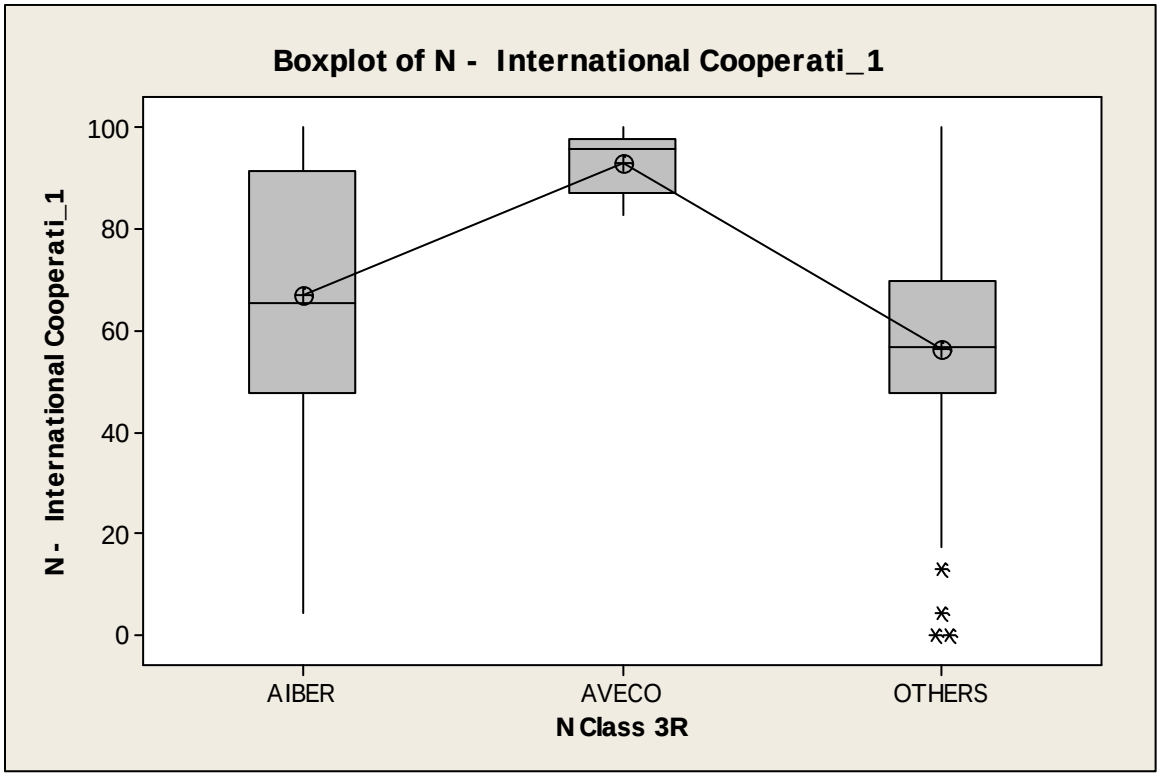
Boxplot of N - International Cooperati_1



One-way ANOVA: N - International Cooperati_1 versus N Class 3R

```
Source       DF     SS    MS     F      P
N Class 3R    2   7532  3766  9,36  0,000
Error       101  40643   402
Total       103  48175

S = 20,06   R-Sq = 15,64%   R-Sq(adj) = 13,96%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   19  66,82  23,91      (------*-----)
AVECO    5  93,04   6,59                 (-----------*-----------)
OTHERS  80  56,14  19,54  (--*--)
                          ------+---------+---------+---------+---
                               60        75        90       105

Pooled StDev = 20,06
```

Boxplot of N - International Cooperati_1

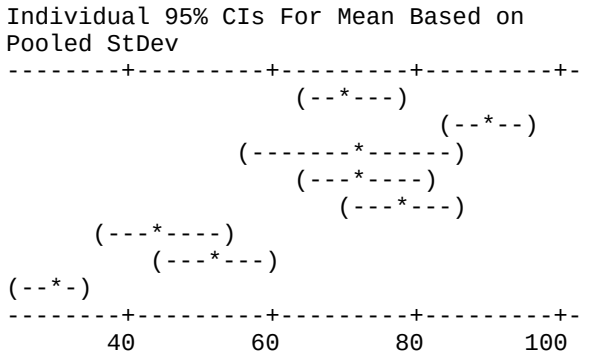


One-way ANOVA: N - 2016 EPI Score_1 versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	66838	9548	42,14	0,000
Error	124	28093	227		
Total	131	94932			

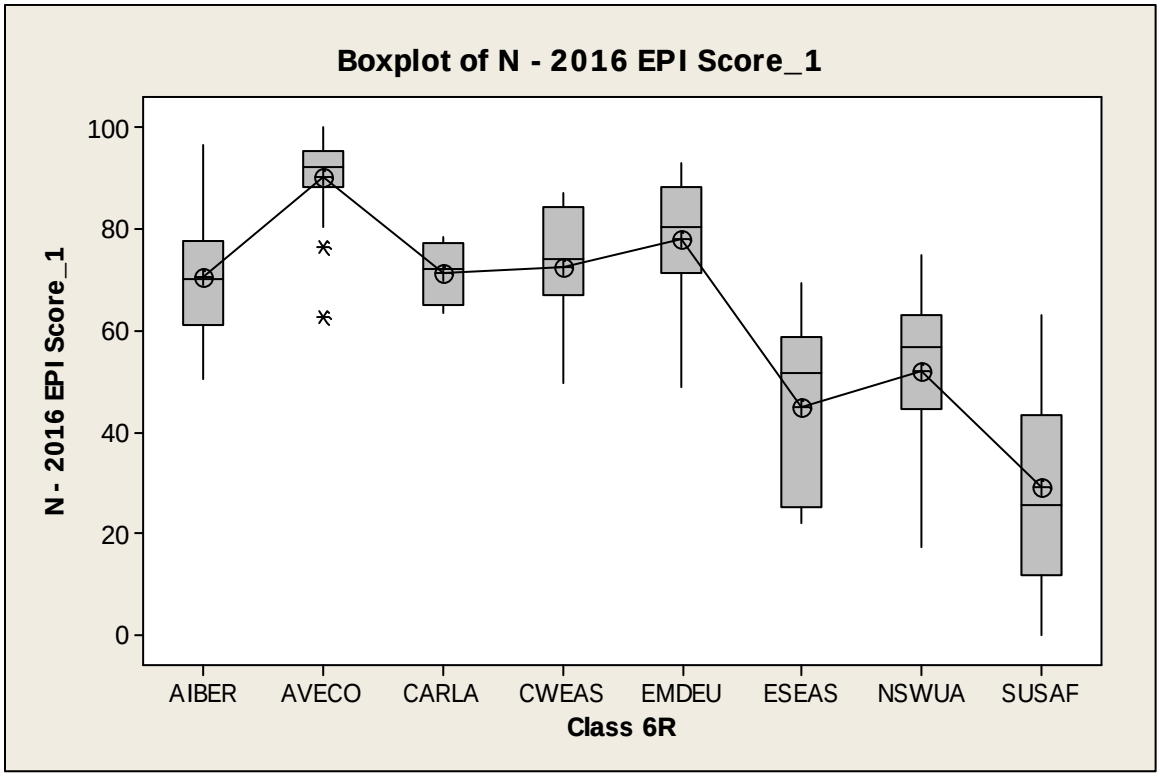
S = 15,05 R-Sq = 70,41% R-Sq(adj) = 68,74%

Level	N	Mean	StDev
AIBER	20	70,69	12,10
AVECO	27	90,29	8,04
CARLA	4	71,45	6,39
CWEAS	11	72,55	13,20
EMDEU	12	77,87	13,52
ESEAS	12	44,71	17,86
NSWUA	14	52,04	17,44
SUSAF	32	29,29	19,91



Pooled StDev = 15,05

Boxplot of N - 2016 EPI Score_1

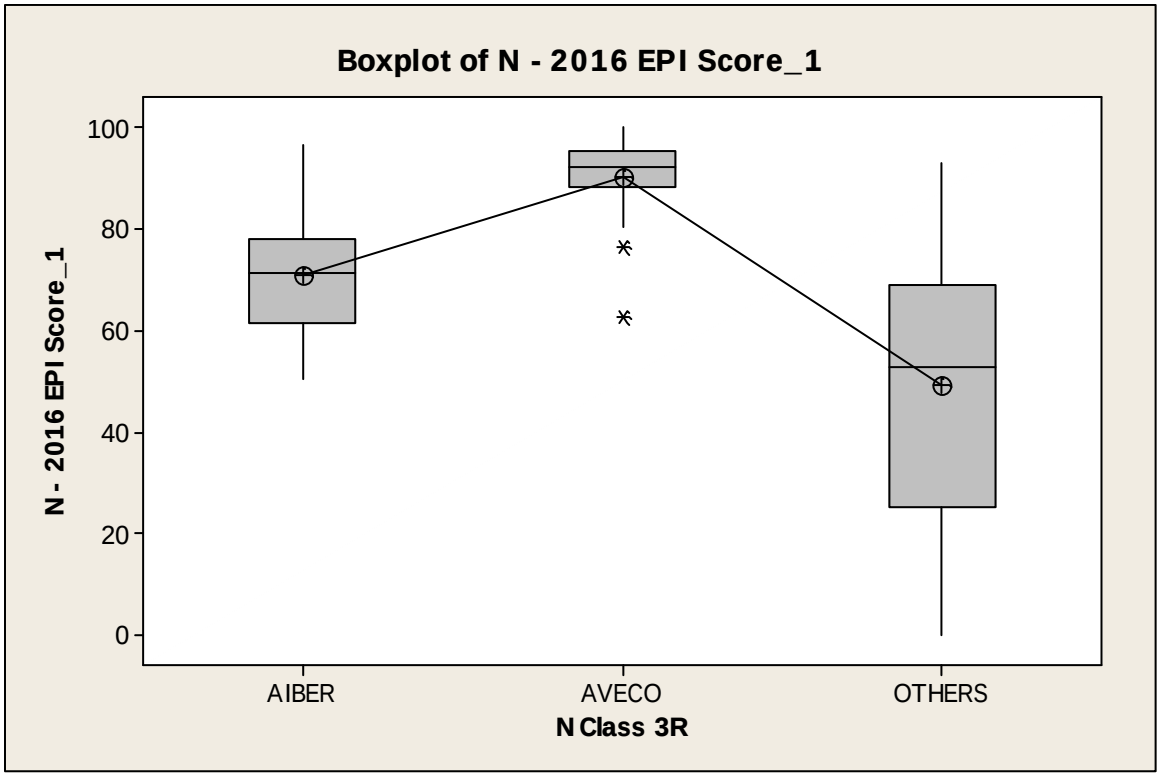


Outliers: 64, 57

One-way ANOVA: N - 2016 EPI Score_1 versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  36754  18377  40,75  0,000
Error       129  58178    451
Total       131  94932

S = 21,24   R-Sq = 38,72%   R-Sq(adj) = 37,77%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  71,05  11,91              (-----*-----)
AVECO   27  90,29   8,04                        (----*-----)
OTHERS  84  49,32  25,43   (--*--)
                           +---------+---------+---------+---------
                          45        60        75        90

Pooled StDev = 21,24
```

Boxplot of N - 2016 EPI Score_1

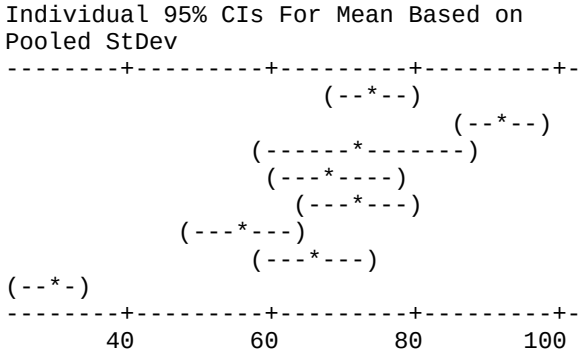


One-way ANOVA: N - Shelter_1 versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	63060	9009	39,24	0,000
Error	124	28465	230		
Total	131	91525			

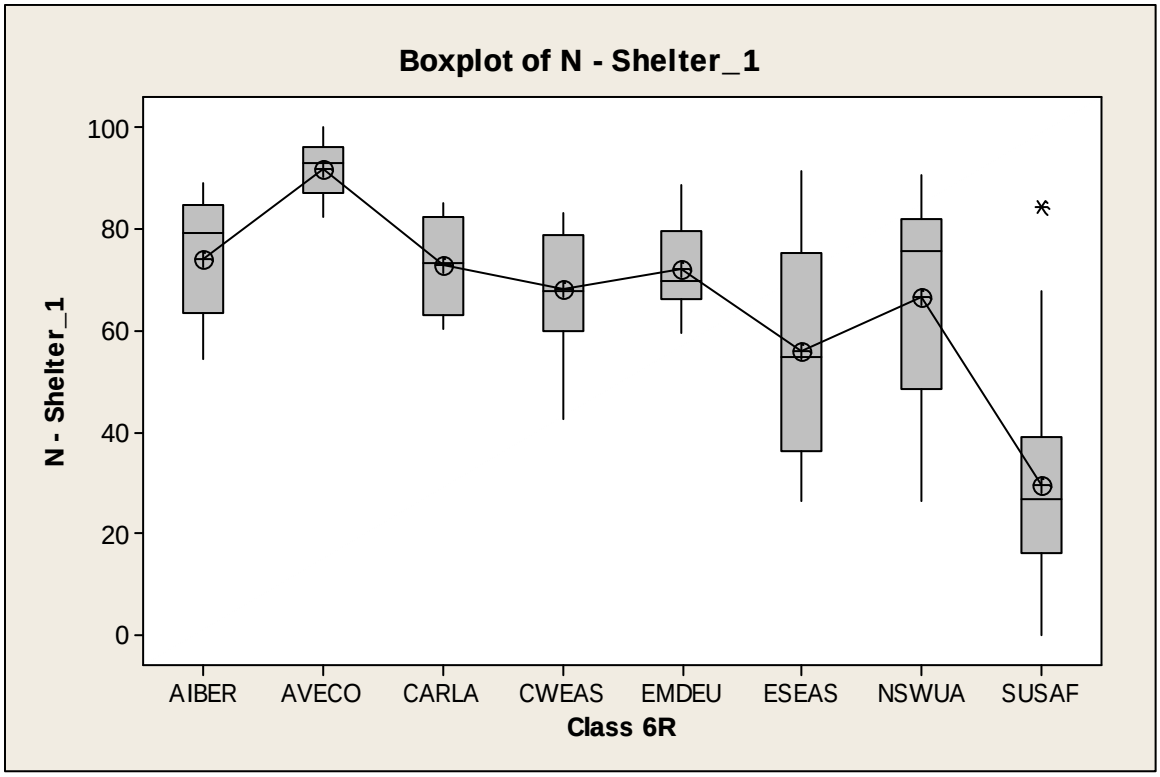
S = 15,15 R-Sq = 68,90% R-Sq(adj) = 67,14%

Level	N	Mean	StDev
AIBER	20	74,08	11,84
AVECO	27	91,72	5,24
CARLA	4	72,90	10,20
CWEAS	11	68,33	11,86
EMDEU	12	72,20	9,29
ESEAS	12	55,78	21,28
NSWUA	14	66,76	21,61
SUSAF	32	29,52	19,15



Pooled StDev = 15,15

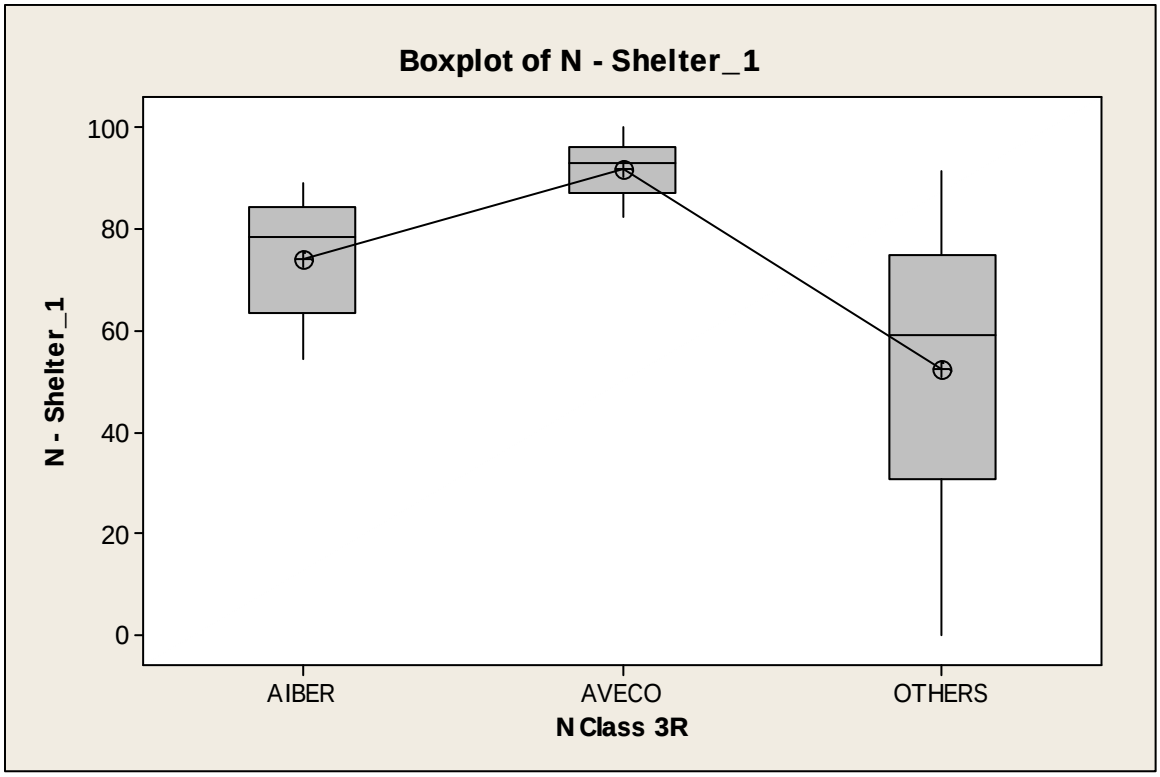
Boxplot of N - Shelter_1



One-way ANOVA: N - Shelter_1 versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  34482  17241  38,99  0,000
Error       129  57042    442
Total       131  91525

S = 21,03   R-Sq = 37,68%   R-Sq(adj) = 36,71%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  73,97  11,56            (-----*-----)
AVECO   27  91,72   5,24                      (----*----)
OTHERS  84  52,22  25,43  (--*--)
                          --------+---------+---------+---------+-
                                 60        75        90       105

Pooled StDev = 21,03
```

Boxplot of N - Shelter_1



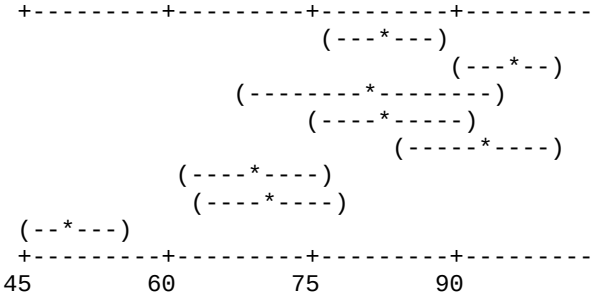
One-way ANOVA: N - Stillbirth rate versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	37020	5289	27,33	0,000
Error	123	23800	193		
Total	130	60820			

S = 13,91 R-Sq = 60,87% R-Sq(adj) = 58,64%

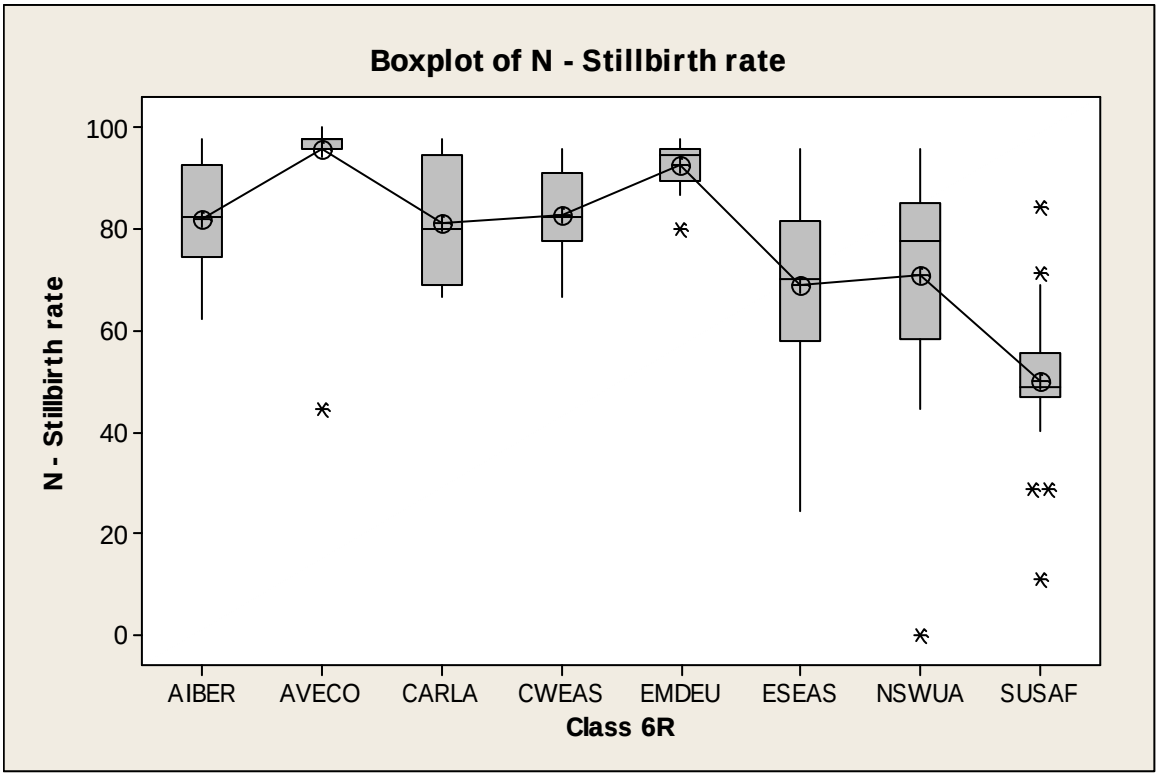
Level	N	Mean	StDev
AIBER	20	82,00	10,52
AVECO	27	95,64	10,34
CARLA	4	81,11	13,27
CWEAS	11	82,63	8,99
EMDEU	12	92,59	5,22
ESEAS	12	68,89	18,95
NSWUA	14	70,79	25,05
SUSAF	31	49,96	13,13

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev = 13,91

Boxplot of N - Stillbirth rate

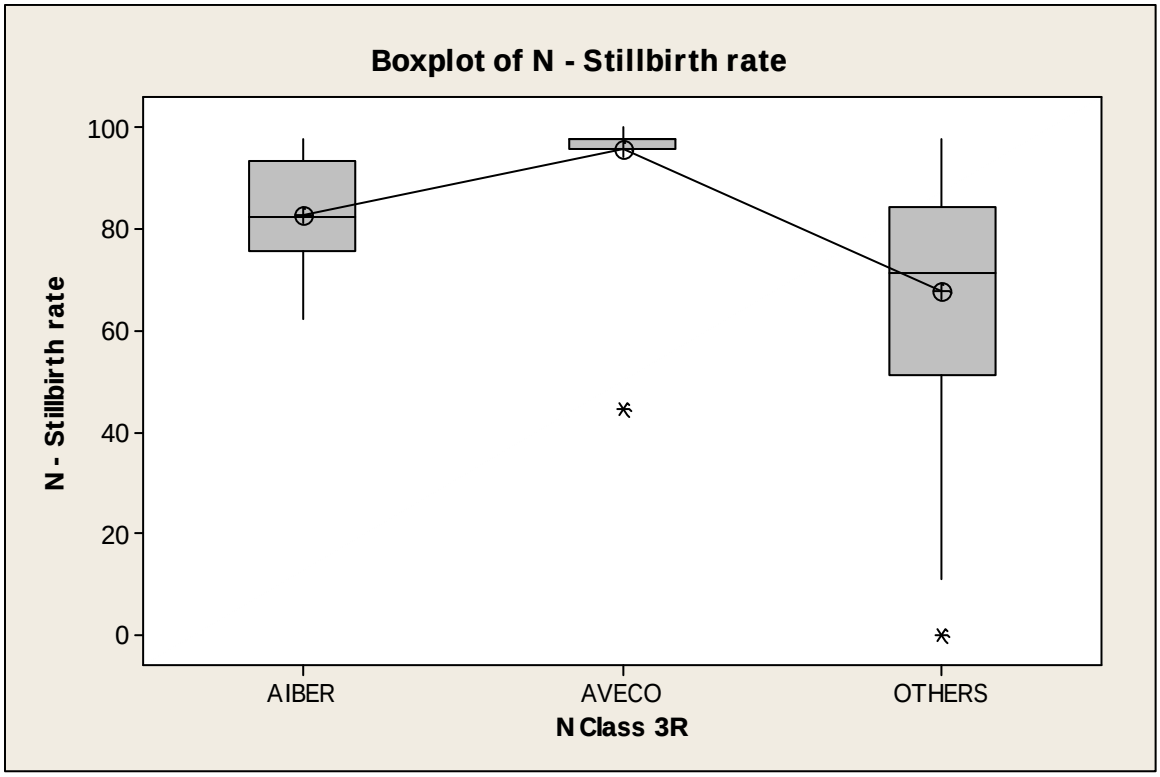


Outlier: 32

One-way ANOVA: N - Stillbirth rate versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  17180  8590  25,19  0,000
Error       128  43640   341
Total       130  60820

S = 18,46   R-Sq = 28,25%   R-Sq(adj) = 27,13%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  82,75  10,82            (-------*-------)
AVECO   27  95,64  10,34                      (------*------)
OTHERS  83  67,63  21,67  (---*---)
                          ------+---------+---------+---------+---
                               70        80        90       100

Pooled StDev = 18,46
```

Boxplot of N - Stillbirth rate

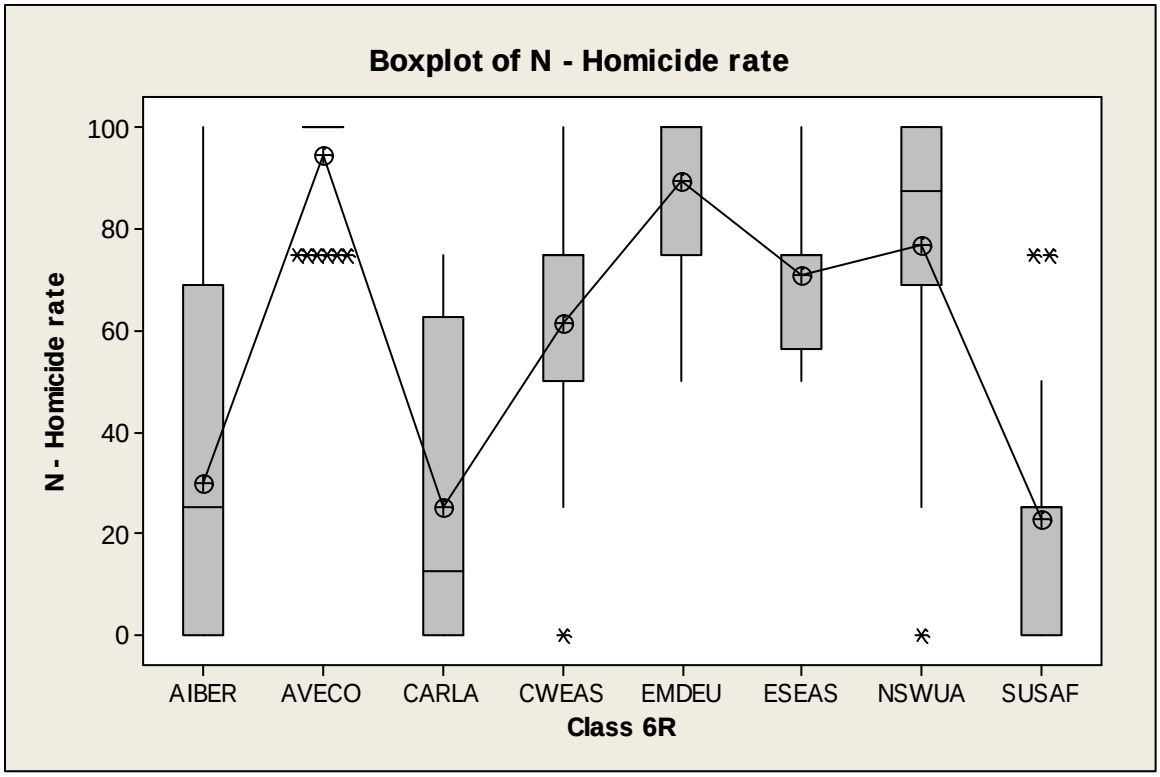


One-way ANOVA: N - Homicide rate versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7  114882  16412  28,36  0,000
Error     124   71765    579
Total     131  186648

S = 24,06   R-Sq = 61,55%   R-Sq(adj) = 59,38%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  20  30,00  35,91         (---*---)
AVECO  27  94,44  10,59                             (--*---)
CARLA   4  25,00  35,36  (-------*-------)
CWEAS  11  61,36  28,20                  (---*----)
EMDEU  12  89,58  16,71                           (----*---)
ESEAS  12  70,83  14,43                     (----*---)
NSWUA  14  76,79  31,72                       (----*---)
SUSAF  32  22,66  21,40       (--*-)
                          +---------+---------+---------+---------
                          0        30        60        90

Pooled StDev = 24,06
```

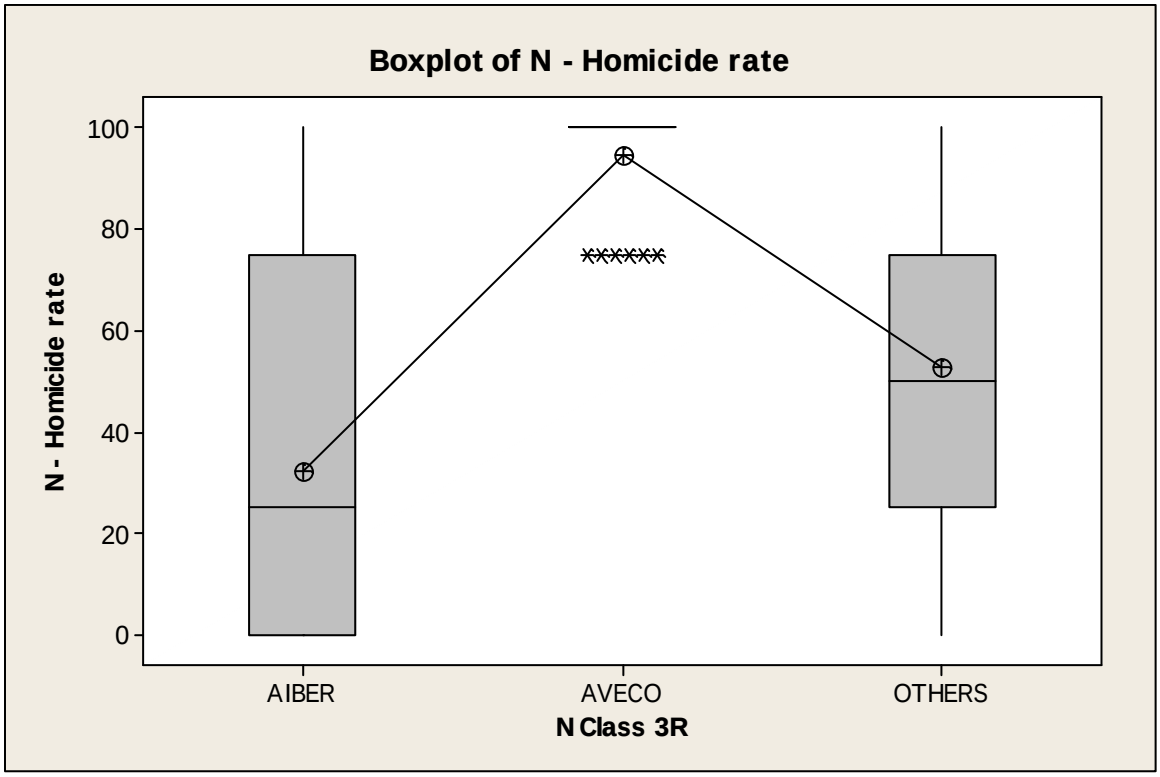
Boxplot of N - Homicide rate



One-way ANOVA: N - Homicide rate versus N Class 3R

```
Source       DF      SS     MS      F      P
N Class 3R    2   52280  26140  25,10  0,000
Error       129  134368   1042
Total       131  186648

S = 32,27   R-Sq = 28,01%   R-Sq(adj) = 26,89%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  32,14  36,35  (-----*----)
AVECO   27  94,44  10,59                              (----*----)
OTHERS  84  52,68  35,57            (--*--)
                          ---+---------+---------+---------+------
                            25        50        75       100

Pooled StDev = 32,27
```

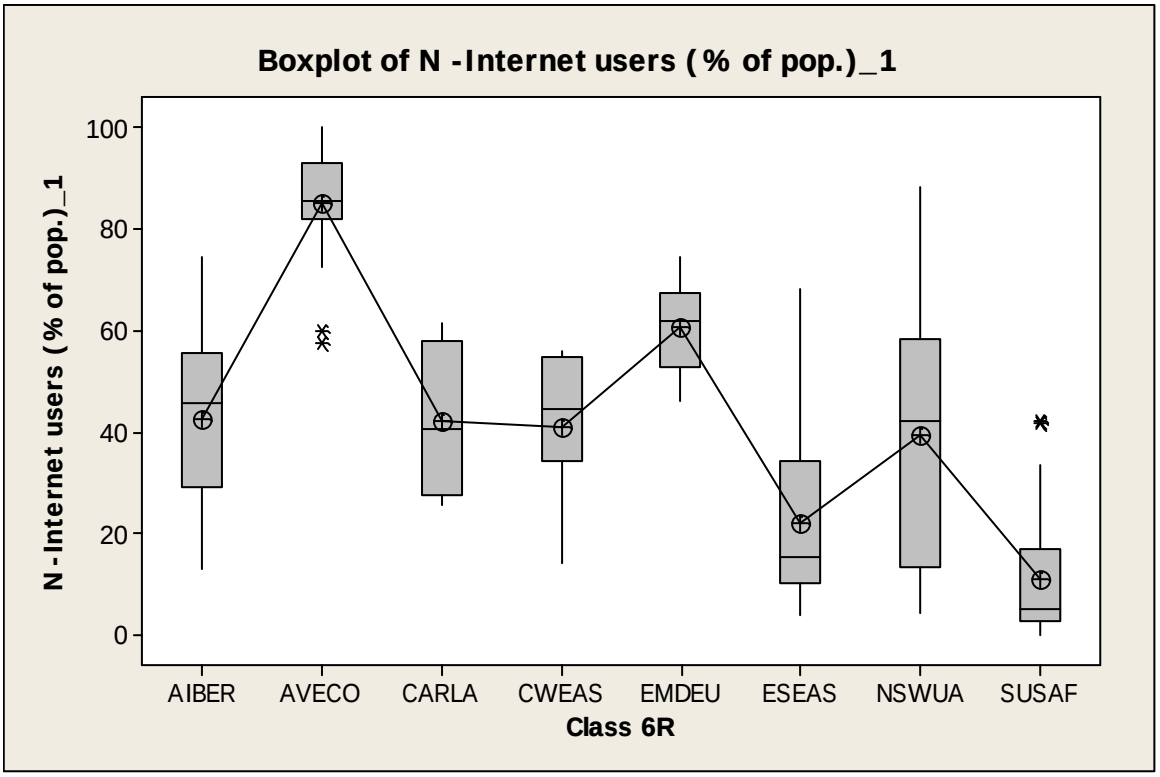
Boxplot of N - Homicide rate



One-way ANOVA: N -Internet users (% of pop.)_1 versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7   89746  12821  52,61  0,000
Error     124   30215    244
Total     131  119961

S = 15,61   R-Sq = 74,81%   R-Sq(adj) = 73,39%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   --------+---------+---------+---------+-
AIBER  20  42,67  17,04             (--*--)
AVECO  27  84,95  10,48                               (-*-)
CARLA   4  42,05  15,77           (-----*-----)
CWEAS  11  41,03  13,65             (--*---)
EMDEU  12  60,69   8,92                     (--*---)
ESEAS  12  22,10  18,81     (---*--)
NSWUA  14  39,48  27,59             (---*--)
SUSAF  32  10,87  12,14   (-*--)
                          --------+---------+---------+---------+-
                                 25        50        75       100

Pooled StDev = 15,61
```

Boxplot of N -Internet users (% of pop.)_1



Outliers: Grécia e Itália

One-way ANOVA: N -Internet users (% of pop.)_1 versus N Class 3R

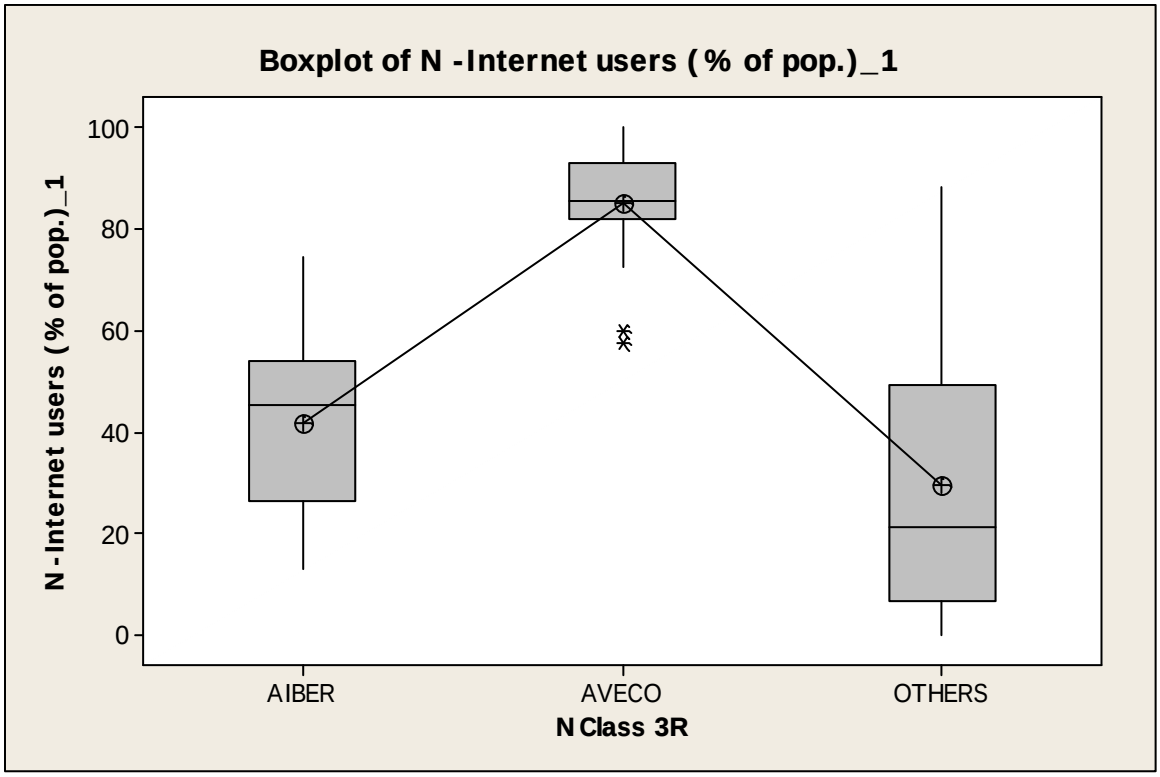
Source	DF	SS	MS	F	P
N Class 3R	2	62573	31287	70,33	0,000
Error	129	57388	445		
Total	131	119961			

S = 21,09 R-Sq = 52,16% R-Sq(adj) = 51,42%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	41,86	17,01	(----*---)
AVECO	27	84,95	10,48	(---*---)
OTHERS	84	29,62	24,23	(-*-)
				40 60 80 100

Pooled StDev = 21,09

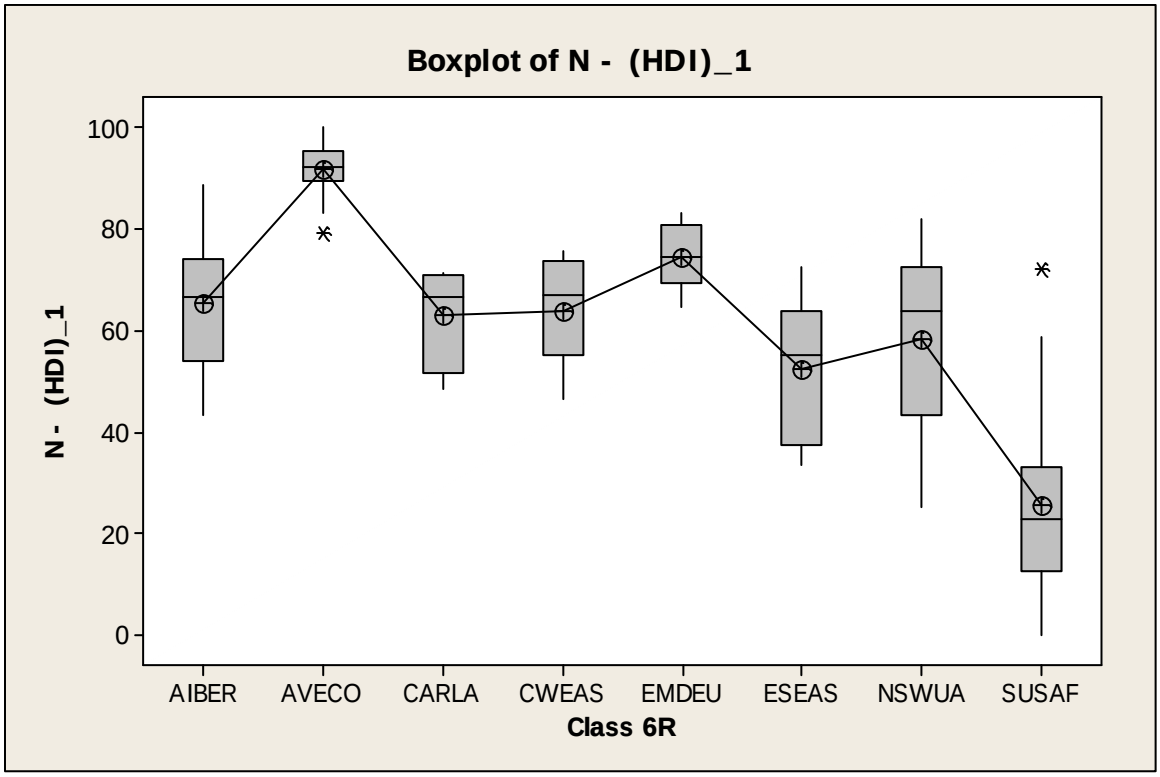
Boxplot of N -Internet users (% of pop.)_1



One-way ANOVA: N - (HDI)_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  68850  9836  57,72  0,000
Error     123  20961   170
Total     130  89811

S = 13,05   R-Sq = 76,66%   R-Sq(adj) = 75,33%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  20  65,42  12,72                    (--*--)
AVECO  27  91,89   4,78                                (--*-)
CARLA   4  63,05  10,68               (------*-----)
CWEAS  11  63,91   9,93                  (---*---)
EMDEU  11  74,35   6,59                       (---*---)
ESEAS  12  52,39  14,24            (---*---)
NSWUA  14  58,37  19,51                (--*---)
SUSAF  32  25,72  16,41  (-*-)
                          ---------+---------+---------+---------+
                                  40        60        80       100

Pooled StDev = 13,05
```

Boxplot of N - (HDI)_1

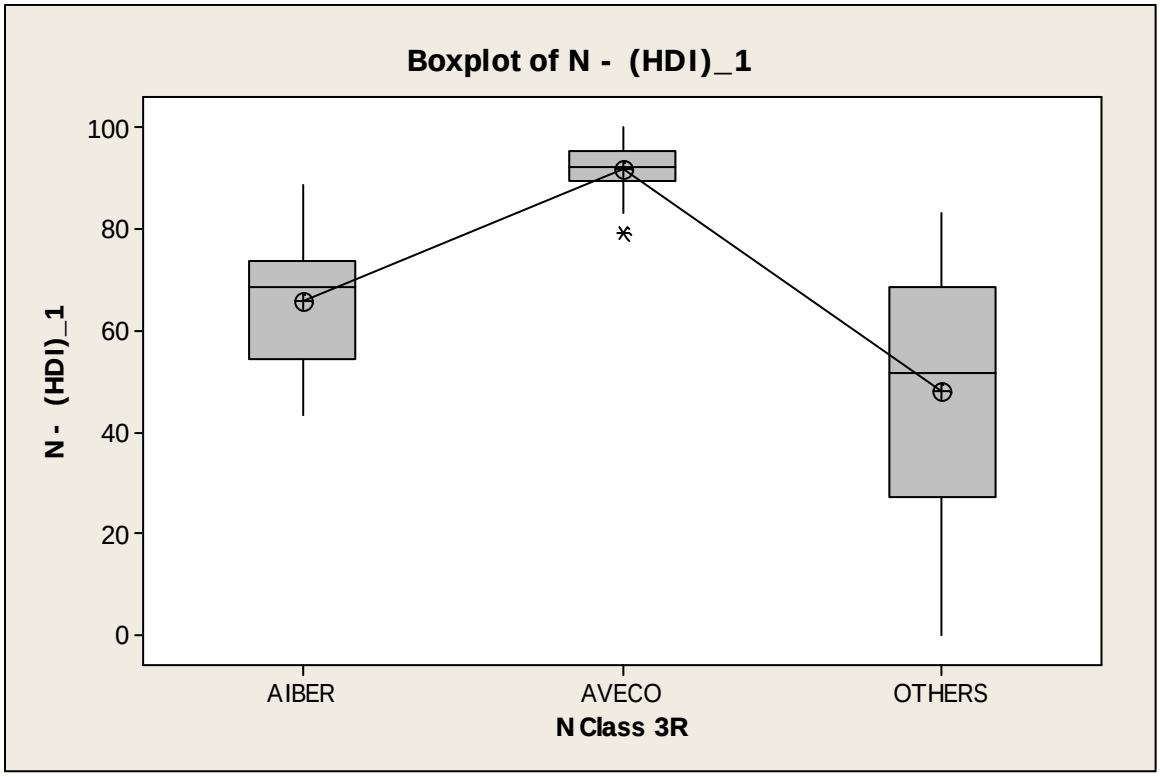


Outlier: Latvia

One-way ANOVA: N - (HDI)_1 versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  40389  20194  52,30  0,000
Error       128  49422    386
Total       130  89811

S = 19,65   R-Sq = 44,97%   R-Sq(adj) = 44,11%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65,67  12,45           (-----*----)
AVECO   27  91,89   4,78                               (----*----)
OTHERS  83  47,85  23,61  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 19,65
```

Boxplot of N - (HDI)_1



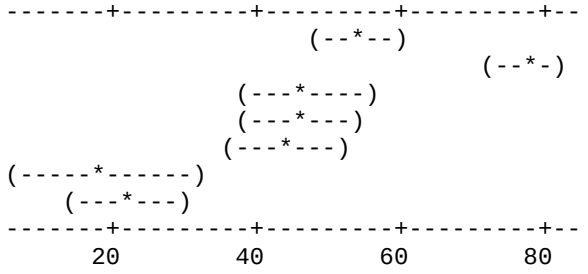
One-way ANOVA: N - AgeWatch Index 2014_1 versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	6	32220	5370	31,88	0,000
Error	83	13980	168		
Total	89	46200			

S = 12,98 R-Sq = 69,74% R-Sq(adj) = 67,55%

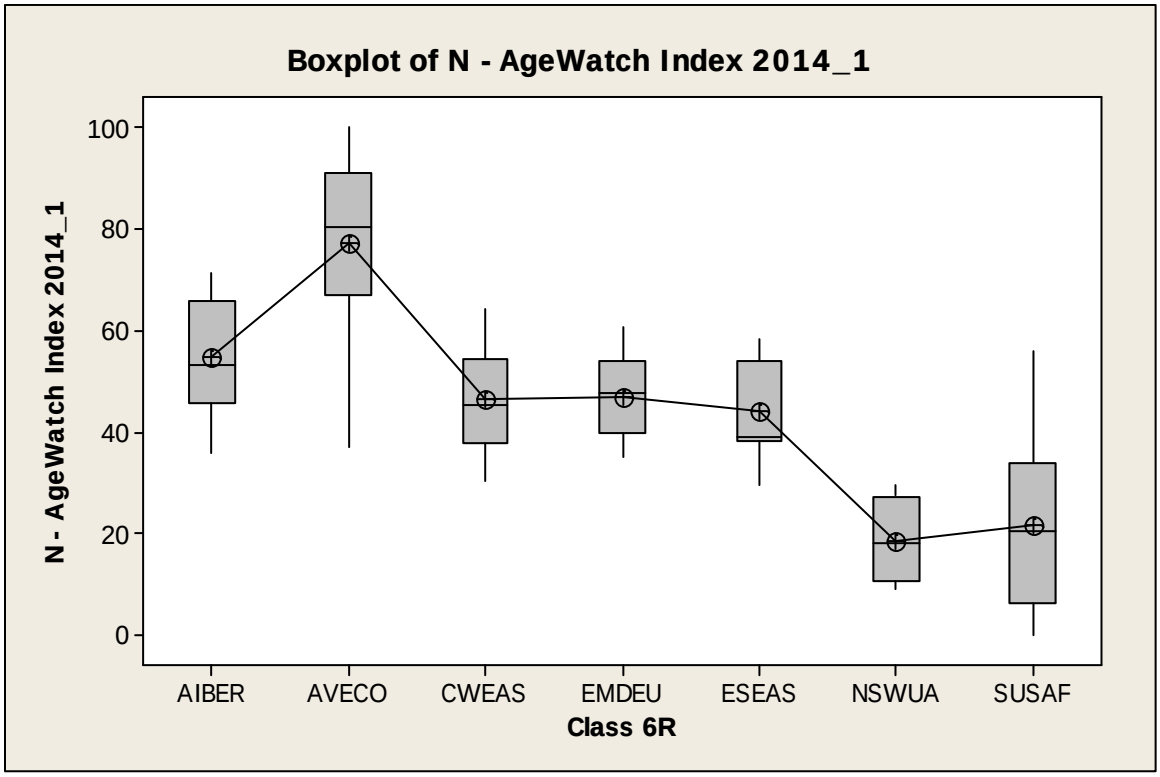
Level	N	Mean	StDev
AIBER	20	54,76	11,20
AVECO	27	77,33	15,76
CWEAS	8	46,32	10,75
EMDEU	10	46,79	8,29
ESEAS	11	44,10	9,57
NSWUA	4	18,56	8,73
SUSAF	10	21,65	16,90

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev = 12,98

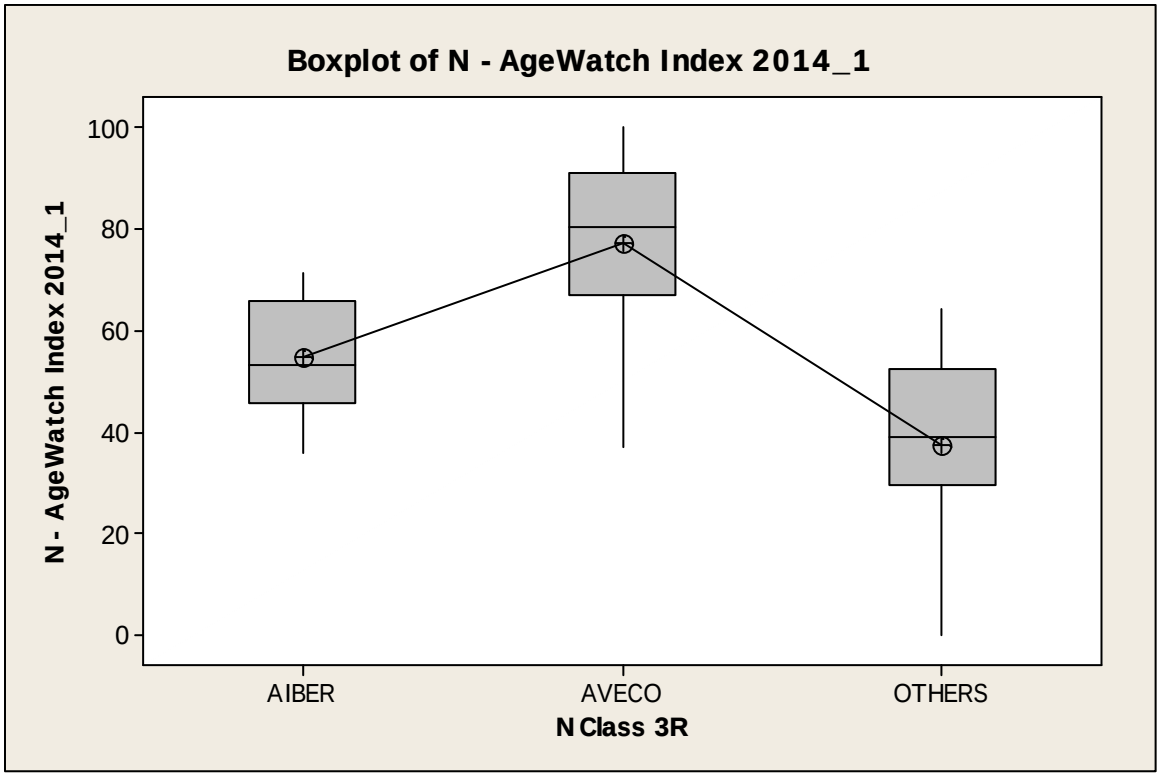
Boxplot of N - AgeWatch Index 2014_1



One-way ANOVA: N - AgeWatch Index 2014_1 versus N Class 3R

```
Source      DF     SS     MS      F      P
N Class 3R   2  26308  13154  57,53  0,000
Error       87  19892    229
Total       89  46200

S = 15,12   R-Sq = 56,94%   R-Sq(adj) = 55,95%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   20  54,76  11,20            (----*---)
AVECO   27  77,33  15,76                              (---*--)
OTHERS  43  37,54  16,22  (--*--)
                          --------+---------+---------+---------+-
                                 45        60        75        90

Pooled StDev = 15,12
```

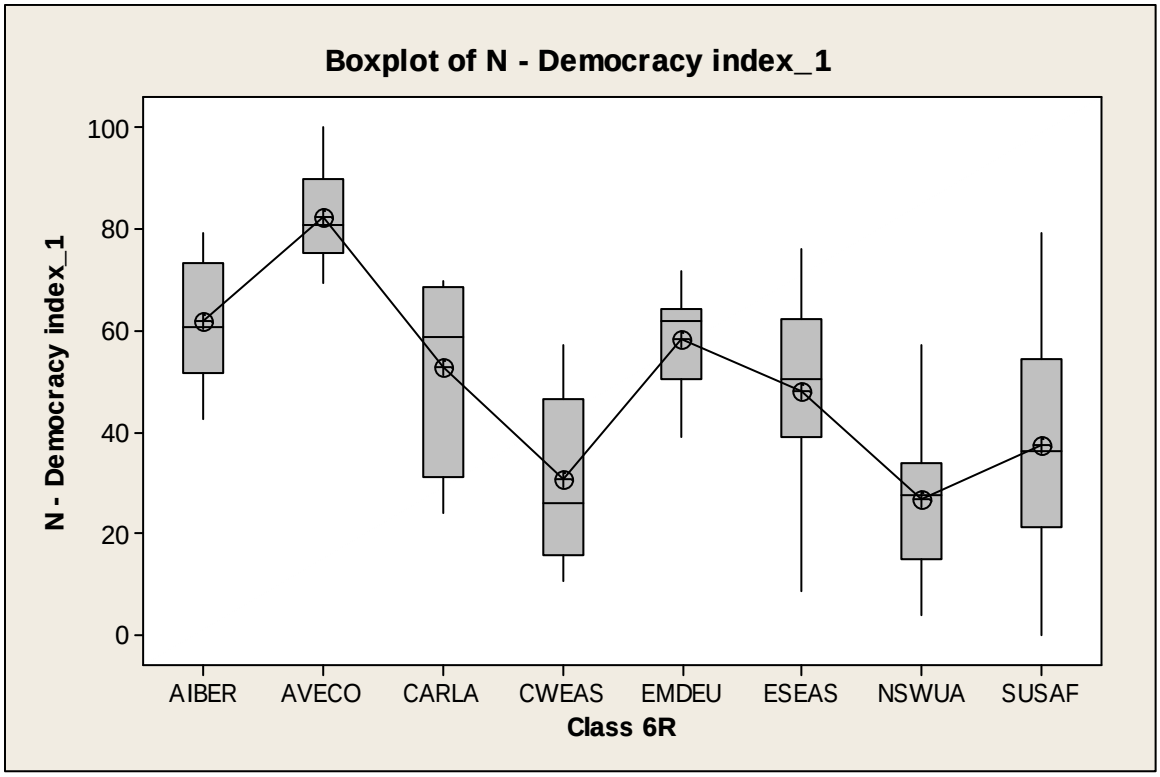
Boxplot of N - AgeWatch Index 2014_1



One-way ANOVA: N - Democracy index_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  48430  6919  28,37  0,000
Error     124  30243   244
Total     131  78672

S = 15,62   R-Sq = 61,56%   R-Sq(adj) = 59,39%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  61,77  11,09                    (---*--)
AVECO  27  82,48   8,87                              (--*--)
CARLA   4  52,87  20,59            (------*-------)
CWEAS  11  30,63  16,67    (---*----)
EMDEU  12  58,10  10,28                  (---*----)
ESEAS  12  47,88  19,58            (----*---)
NSWUA  14  26,55  14,34  (---*---)
SUSAF  32  37,53  21,06         (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,62
```

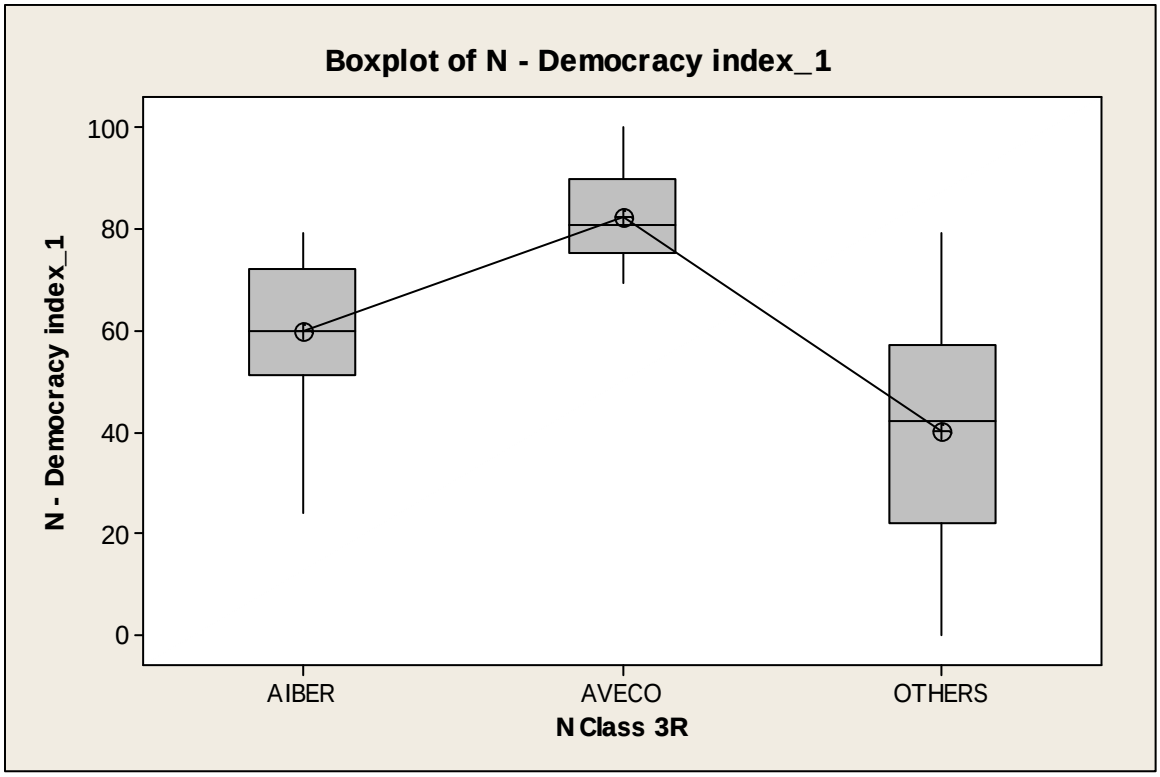
Boxplot of N - Democracy index_1



One-way ANOVA: N - Democracy index_1 versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  38300  19150  61,19  0,000
Error       129  40372    313
Total       131  78672

S = 17,69   R-Sq = 48,68%   R-Sq(adj) = 47,89%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  59,98  13,59            (----*----)
AVECO   27  82,48   8,87                       (----*---)
OTHERS  84  40,11  20,43  (--*-)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 17,69
```

Boxplot of N - Democracy index_1

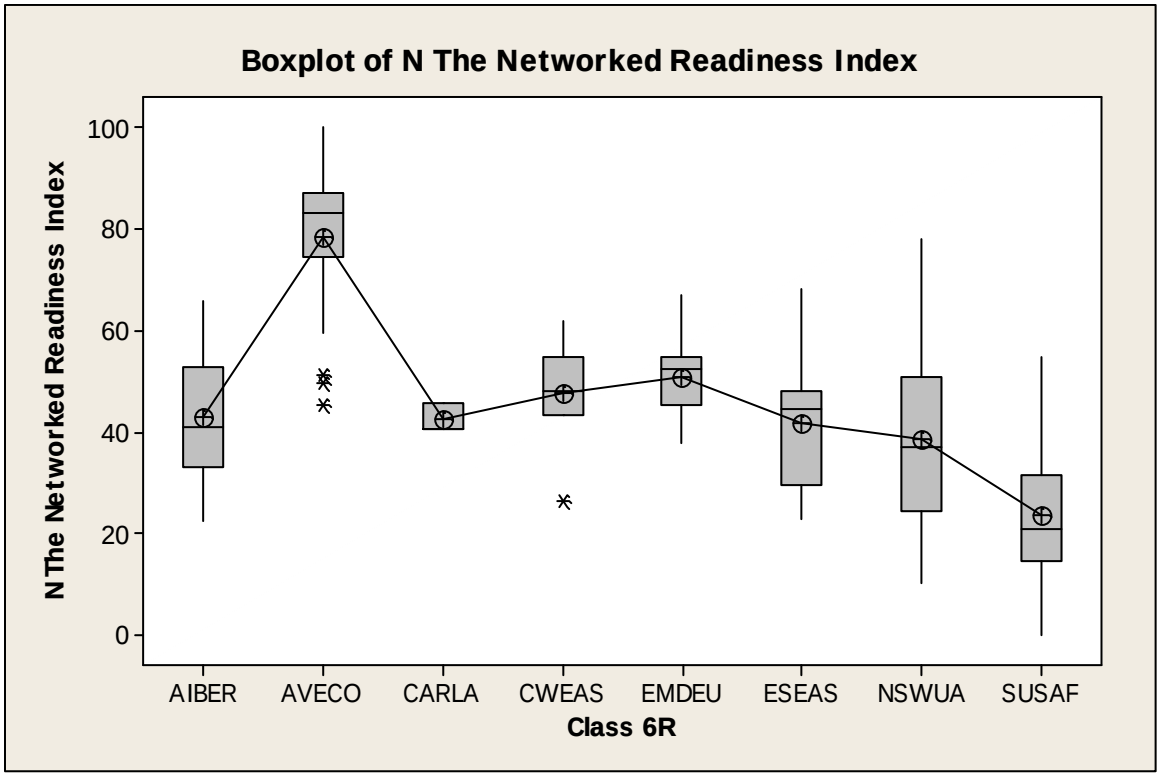


One-way ANOVA: N The Networked Readiness Index versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  42555  6079  32,90  0,000
Error     113  20882   185
Total     120  63437

S = 13,59   R-Sq = 67,08%   R-Sq(adj) = 65,04%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  42,83  13,02            (--*--)
AVECO  27  78,25  14,39                                (-*--)
CARLA   3  42,32   3,02       (-------*-------)
CWEAS   8  47,55  10,65             (----*----)
EMDEU  12  50,70   7,58               (---*---)
ESEAS  12  41,54  12,69           (---*---)
NSWUA  13  38,74  19,70          (--*---)
SUSAF  26  23,50  13,27   (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 13,59
```

Boxplot of N The Networked Readiness Index

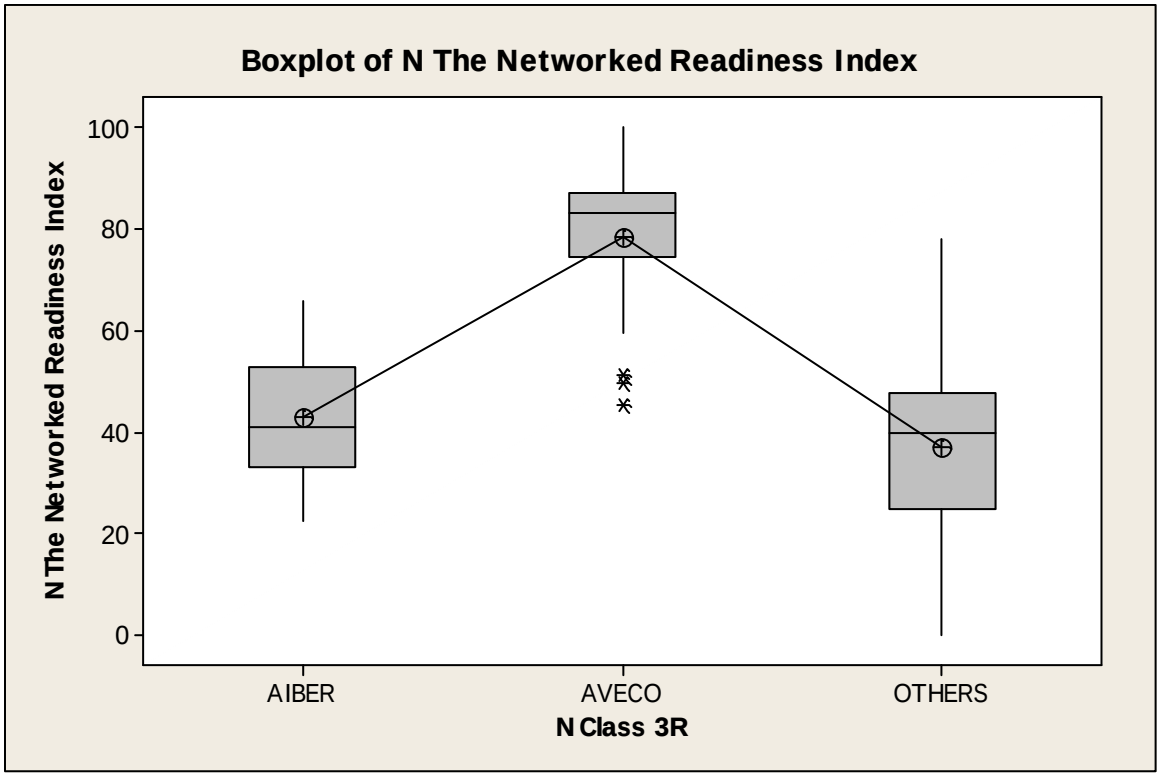


Outliers: 58, 107, 45

One-way ANOVA: N The Networked Readiness Index versus N Class 3R

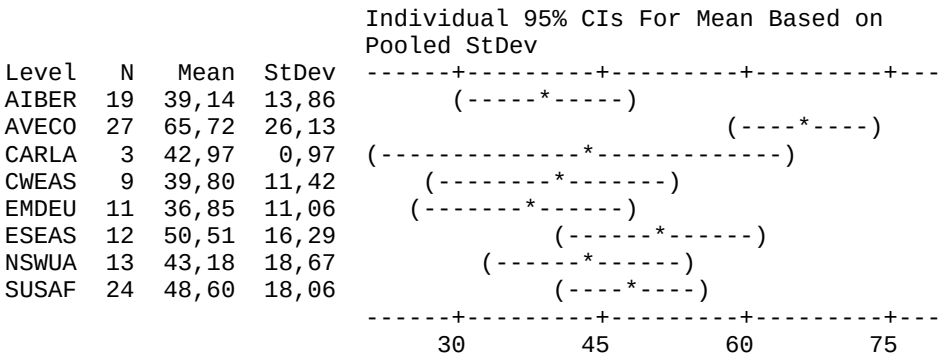
```
Source       DF     SS     MS      F      P
N Class 3R    2  34304  17152  69,48  0,000
Error       118  29132    247
Total       120  63437

S = 15,71   R-Sq = 54,08%   R-Sq(adj) = 53,30%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   20  42,83  13,02    (----*---)
AVECO   27  78,25  14,39                              (---*---)
OTHERS  74  36,88  16,77  (--*-)
                          --------+---------+---------+---------+-
                                 45        60        75        90

Pooled StDev = 15,71
```

Boxplot of N The Networked Readiness Index



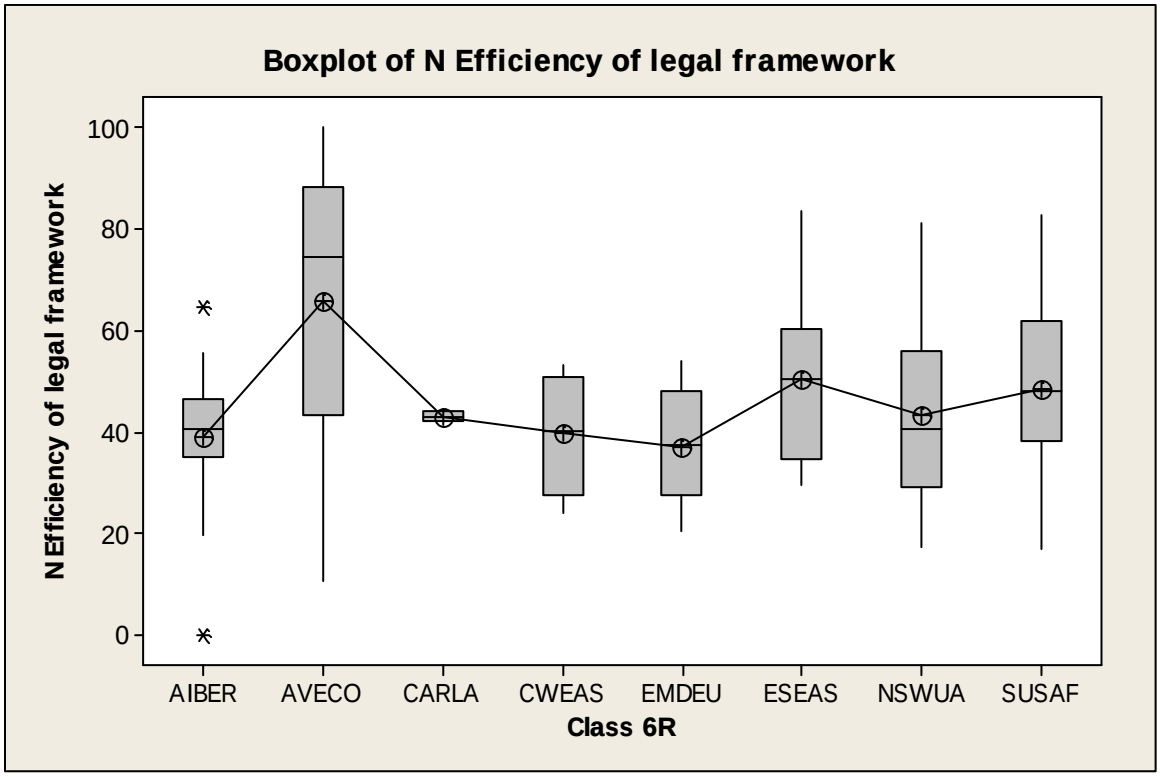
One-way ANOVA: N Efficiency of legal framework versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  12349  1764  5,10  0,000
Error     110  38082   346
Total     117  50431

S = 18,61   R-Sq = 24,49%   R-Sq(adj) = 19,68%
```




```
Pooled StDev = 18,61
```

Boxplot of N Efficiency of legal framework

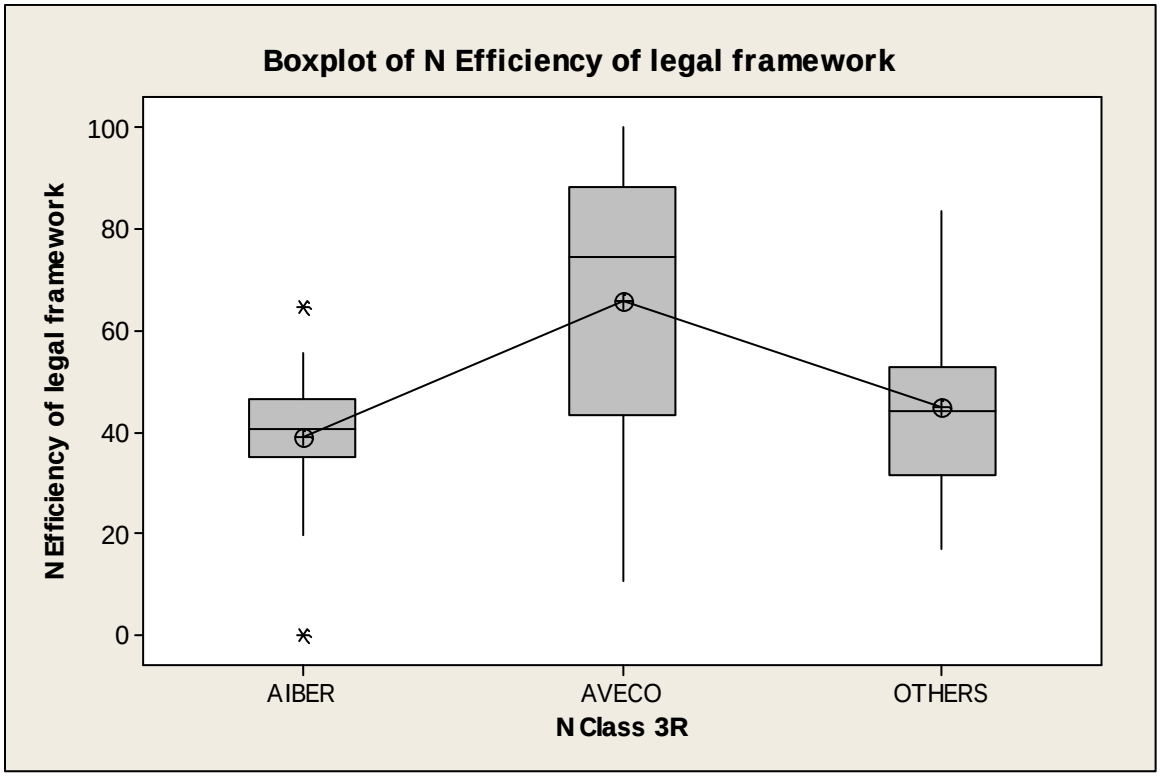


Outliers: Chile(+); Venezuela (-)

One-way ANOVA: N Efficiency of legal framework versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  10646  5323  15,39  0,000
Error       115  39785   346
Total       117  50431

S = 18,60   R-Sq = 21,11%   R-Sq(adj) = 19,74%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   19  39,14  13,86  (------*------)
AVECO   27  65,72  26,13                        (-----*-----)
OTHERS  72  44,81  16,17         (--*---)
                          ----+---------+---------+---------+-----
                             36        48        60        72

Pooled StDev = 18,60
```

Boxplot of N Efficiency of legal framework



b. ANÁLISE CONSOLIDADA DOS TESTES ANOVA: VARIÁVEIS DO EIXO PESSOAS VS REGIÕES

Analisa-se com o teste ANOVA a possibilidade de cada variável construir grupos com maior ou menor grau de similaridade entre os agrupamentos de países, para duas situações, com oito e com três grupos. Utiliza-se o resultado do F no teste ANOVA como valor comparativo. Cabe lembrar que todos os valores obtidos possuem p valor significativo a 1%. Na tabela 4 verifica-se o grau comparativo de proximidade obtido pra cada variável para as duas situações. Conclui-se que na situação de comparação entre os oito grupos de países (6R) a variável que mais discrimina, ou seja, que torna os grupos mais diversos entre si, ou com menor proximidade, é o Índice de Desenvolvimento Humano, (IDH ou HDI no original em inglês), com F de 57,62, seguido pelo *Social Progress Index*,SPI, com F de 53,48 e o % de usuários de internet como percentual da população, com F de 52,61. Ao realizar a mesma análise com três regiões – Países Ibero-Americanos, Países Desenvolvidos e Outros Países, as variáveis significativas de agrupamento encontradas não são todas as mesmas. E os valores de F discriminante são significativamente maiores, indicando que as diferenças encontradas são mais relevantes. A principal variável nesta segunda análise e Governança, com F igual a 97,02 (era 35,53 no teste com oito regiões), seguida pelo SPI com F igual a 82,33 e por fim o % de usuários de internet com F igual a 70,33. Portanto o modelo com três regiões parece indicar maiores diferenças entre os grupos do que o teste com oito grupos, a partir da análise do F no teste ANOVA.

Tabela 5 – Resultados Consolidados dos Testes ANOVA

Índice	Resultado com 8 Regiões (6R)		Resultado com 3 Regiões (3R)	
	F	P	F	P
SPI	53,48	0,00	82,33	0,00
Governança	35,53	0,00	97,92	0,00

HPI	15,83	0,00	22,42	0,00
GINNI	17,93	0,00	23,36	0,00
Access to Piped Water	48,67	0,00	40,02	0,00
Access to improved sanitation	41,94	0,00	18,98	0,00
International cooperation índex	06,06	0,00	09,36	0,00
EPI	42,14	0,00	40,75	0,00
Shelter	39,24	0,00	38,99	0,00
Stillbirth	27,33	0,00	25,19	0,00
Homicide Rate	28,36	0,00	25,10	0,00
% de usuários de Internet	52,61	0,00	70,33	0,00
IDH / HDI	57,62	0,00	52,30	0,00
Age Watch Index	31,88	0,00	57,53	0,00
Democracy Index	28,37	0,00	61,17	0,00
Network readiness	32,90	0,00	69,48	0,00
Efficiency of legal framework in settling disputes, 1-7 (best)	05,10	0,00	15,39	0,00

Fonte: preparado pelo autor com base nas análises executdas no Minitab 16®

c. ANÁLISE ATRAVÉS DO GRÁFICO DE RADAR DAS VARIÁVEIS SINTÉTICAS E ANALÍTICAS DO EIXO PESSOAS VS REGIÕES

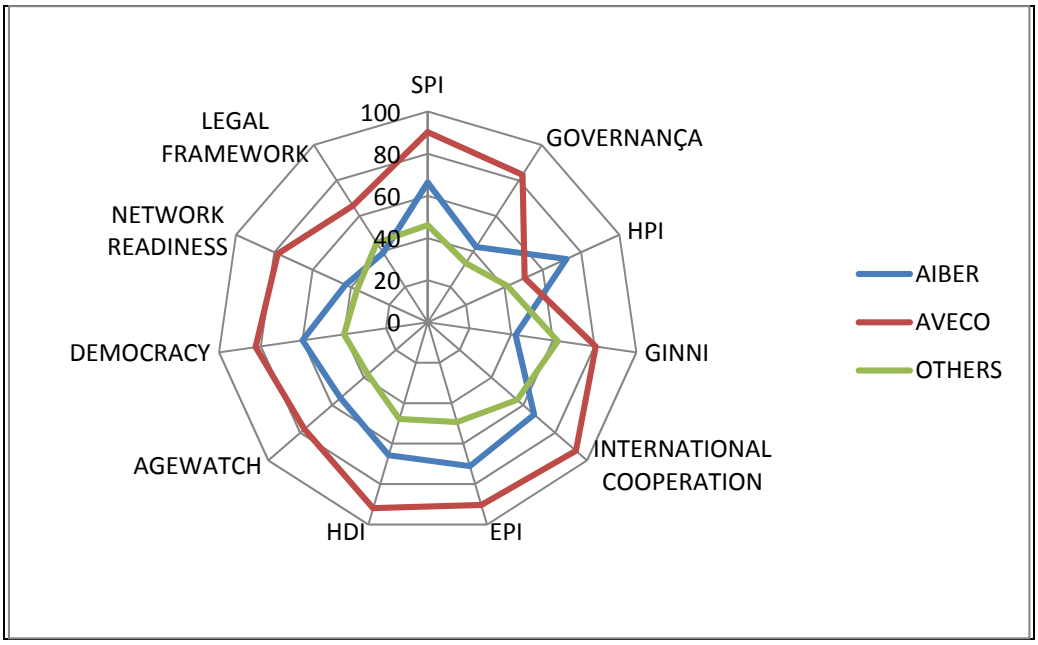
i. GRÁFICO DE RADAR DAS VARIÁVEIS SINTÉTICAS - MÉDIAS E F - EIXO PESSOAS VS REGIÕES

No gráfico de Radar fica claro o destaque para as variáveis Governança e SPI – Social Progresso Index como diferenciadora entre os países ibéricos (AIBER) e as economias avançadas (AVECO). Além das regiões opta-se por incluir também a significância de cada variável para auxiliar na comparação (F – ANOVA), de modo a incuir na avaliação o vetor de significância estatística das variáveis, o que dá mais peso ainda às diferenças percebidas nas duas variáveis mencionadas acima. Destaca-se também que na variável sintética HPI que mensura felicidade, os países ibéricos possuem média superior aos países avançados.

Tabela - Média das Variáveis Sintéticas e significância (F – ANOVA)

REGION	SPI	GOVERNANÇA	HPI	GINNI	INT. COOP.	EPI	HDI	AGEWATCH	DEMOCRACY	NETWORK READINESS	LEGAL FRAMEWORK
AIBER	66,4	42,48	72,4	42,09	66,82	71,1	65,7	54,76	59,98	42,83	39,14
AVECO	90,3	83,25	50,6	80,57	93,04	90,3	91,9	77,33	82,48	78,25	65,72
OTHERS	46,2	33,36	41,6	62,36	56,14	49,3	47,9	37,54	40,11	36,88	44,81
F (ANOVA)	82,3	97,02	22,4	23,36	9,36	40,8	52,3	57,53	61,19	69,48	15,39

Gráfico Radar - Média das Variáveis Sintéticas e significância (F – ANOVA)



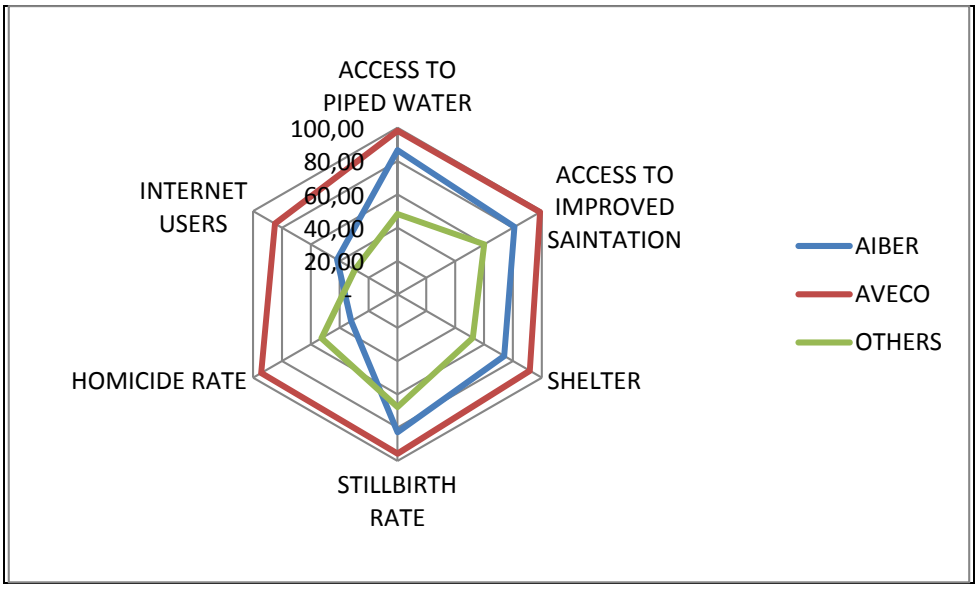
ii. GRÁFICO DE RADAR DAS VARIÁVEIS ANALÍTICAS - MÉDIAS E F - EIXO PESSOAS VS REGIÕES

Através do gráfico e da tabela comparativa sobre as variáveis analíticas do EIXO PESSOAS observa-se uma delas se destacam, taxa de homicídios. Uma segunda também surge em menor escala, usuários de internet. Cabe destacar que na análise de significância comparando os resultados do F para o teste ANOVA esta última se destaca como maior significância estatística.

Tabela - Média das Variáveis Analíticas e significância (F – ANOVA)

REGION	ACCESS TO PIPED WATER	ACCESS TO IMPROVED SAINTATION	SHELTER	STILLBIRTH RATE	HOMICIDE RATE	INTERNET USERS
AIBER	86,56	81,00	73,97	82,75	32,14	41,86
AVECO	98,12	98,72	91,72	95,64	94,44	84,95
OTHERS	48,19	60,06	52,22	67,63	52,68	29,62
F (ANOVA)	40,02	18,98	38,99	25,19	25,10	70,33

Gráfico Radar - Média das Variáveis Analíticas e significância (F – ANOVA)



6. AMOSTRAS ALEATÓRIAS DE VARIÁVEIS SELECIONADAS DO EIXO PESSOAS

Este trabalho tem como objetivo analisar a similaridade ou dessemelhança entre amostras aleatórias de três variáveis do eixo pessoas, geradas através de seleção randômica no apresentadas na Tabela 1.

Tabela1 1- Variáveis do EIXO PESSOAS aplicadas a esta etapa

Índice	Significado	Fonte	Tipo	Unidade de medida antes da normalização
IDH (atualizado para 2014)	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação	PNUD	Quantitativa	%
1. Governança	Governança eficácia de governo - consiste nas tradições e instituições pelas quais a autoridade de um país é exercido. Isso inclui o processo pelo qual os governos são selecionados, monitorados e substituídos; a capacidade do governo para formular e implementar com eficácia políticas sólidas; eo respeito dos cidadãos e do Estado pelas instituições que governam as interações econômicas e sociais	World Bank - World Governance Index	Quantitativa	%

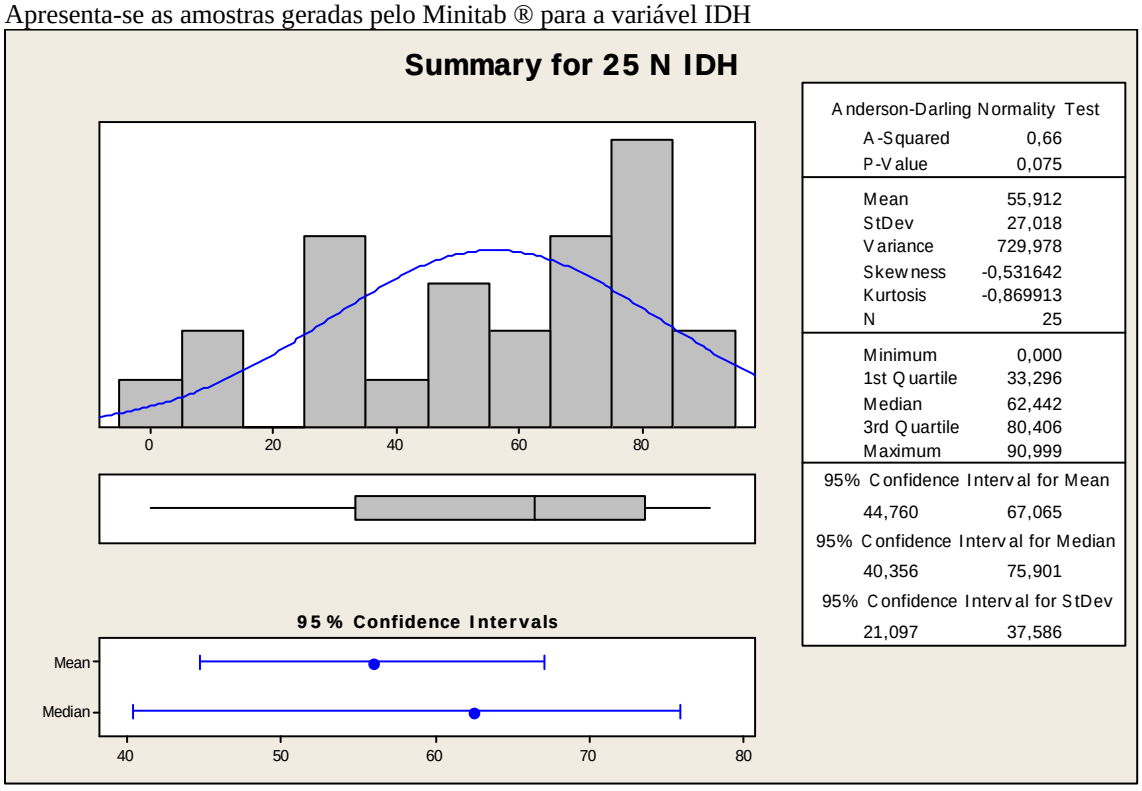
SPI - Índice de Progresso Social	O SPI mede progresso social como “a capacidade de uma sociedade para satisfazer as necessidades humanas básicas dos seus cidadãos, estabelecer os blocos de construção que permitem que os cidadãos e as comunidades melhorem, manter a qualidade de suas vidas, e criar as condições para todos os indivíduos atingirem seu pleno potencial. O SPI mede uma ampla gama de componentes do desempenho social e ambiental e agrega-os num quadro geral.” Fonte: Social progress imperative.	Social Progress Imperatives empregando dados da World Health Organization (WHO)-UNICEF	Quantitativa	%

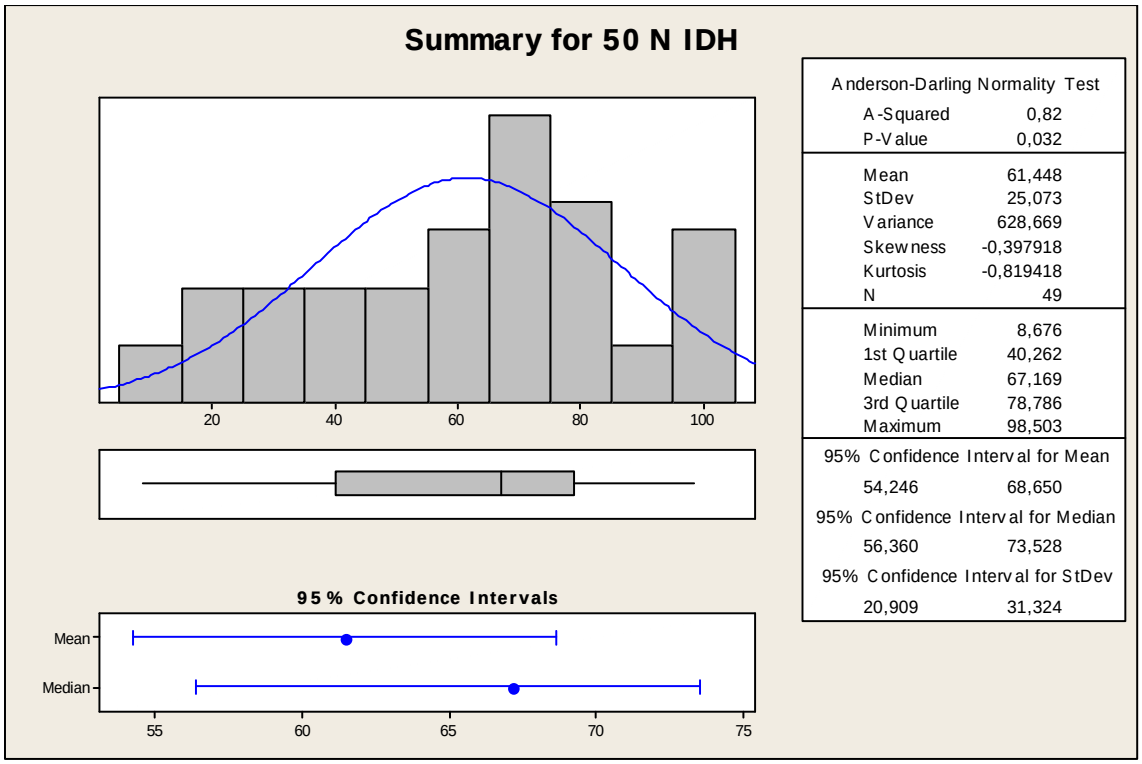
a. ANÁLISE DAS AMOSTRAS ALEATÓRIAS PARA AS VARIÁVEIS ANALISADAS

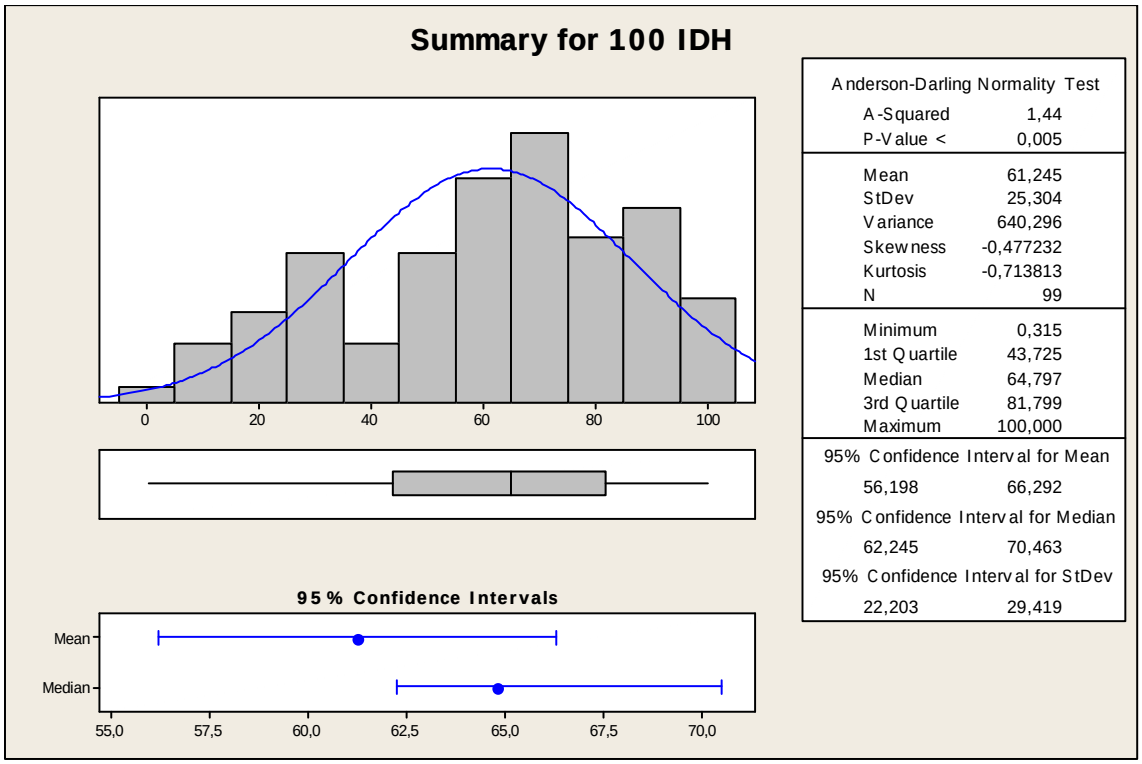
i. VARIÁVEL IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO

1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES.

Apresenta-se as amostras geradas pelo Minitab ® para a variável IDH

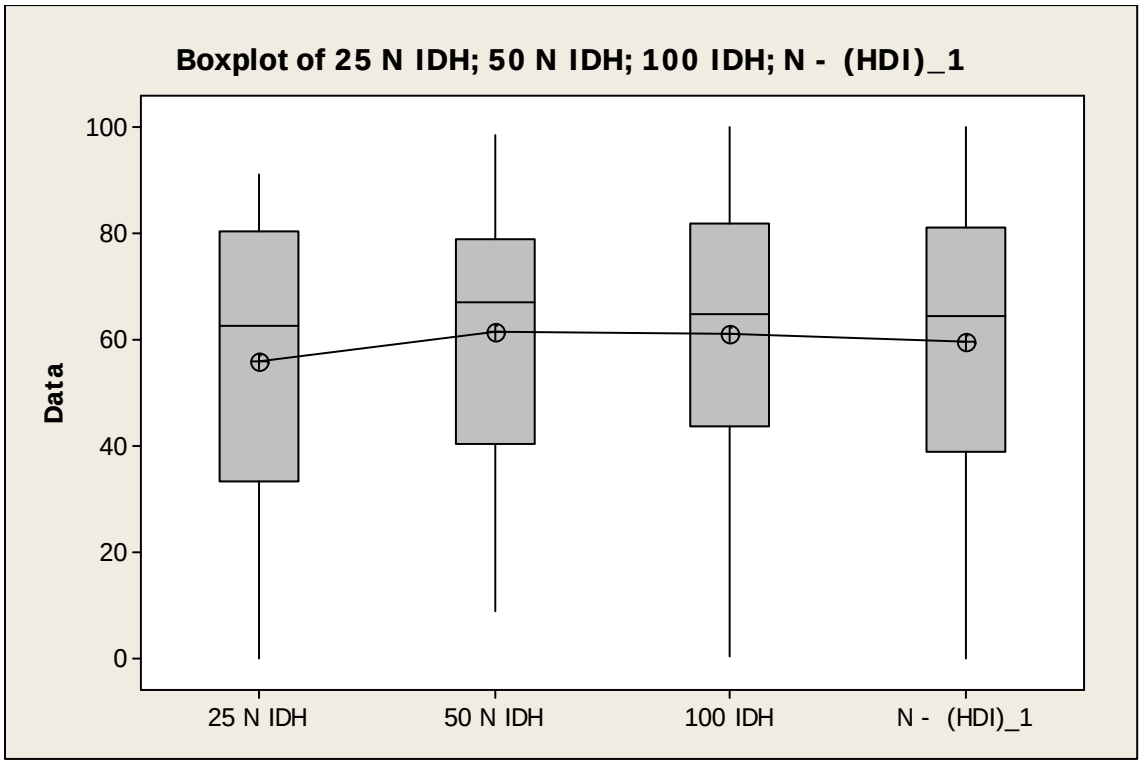






2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES

Boxplot – ANOVA – das Amostras aleatórias e total



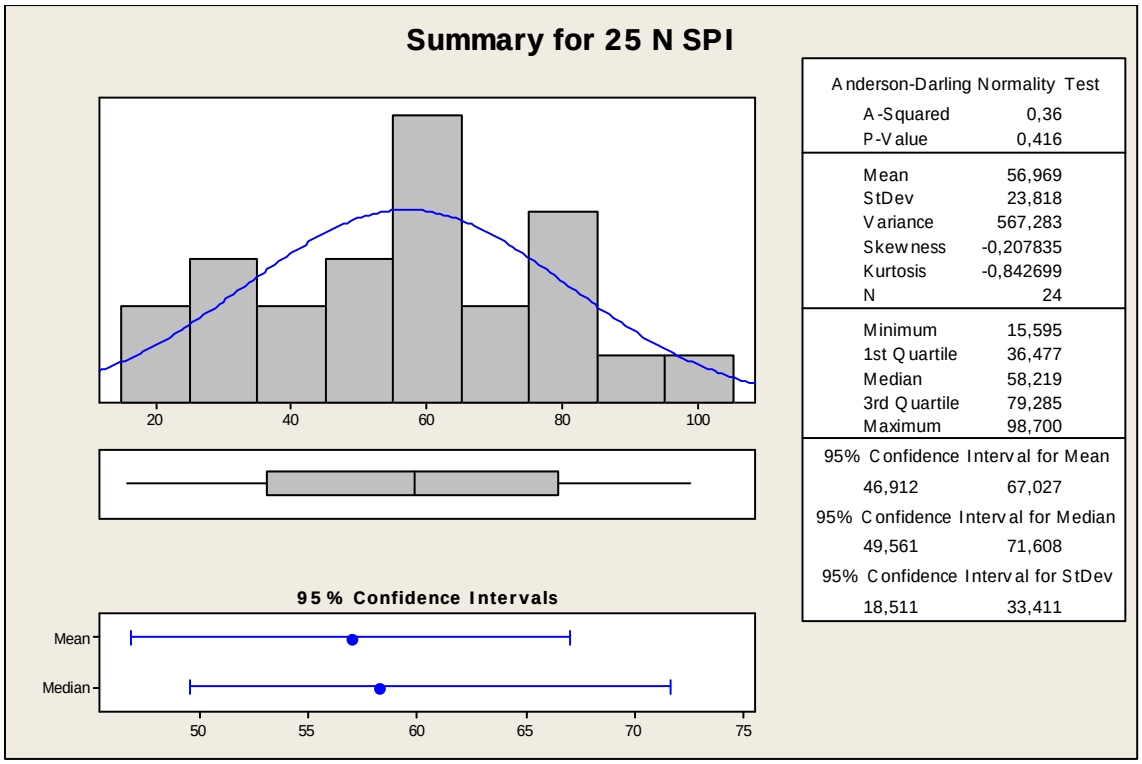
One-way ANOVA: 25 N IDH; 50 N IDH; 100 IDH; N - (HDI)_1

```
Source   DF      SS   MS     F      P
Factor    3     667  222  0,33  0,802
Error   300  200255  668
Total   303  200922

S = 25,84   R-Sq = 0,33%   R-Sq(adj) = 0,00%

                                    Individual 95% CIs For Mean Based on
                                    Pooled StDev
Level           N   Mean  StDev  ----+---------+---------+---------+-----
25 N IDH       25  55,91  27,02  (-----------------*----------------)
50 N IDH       49  61,45  25,07               (-----------*-----------)
100 IDH        99  61,24  25,30                   (-------*--------)
N -  (HDI)_1  131  59,78  26,28                 (-------*------)
                                    ----+---------+---------+---------+-----
                                     48,0      54,0      60,0      66,0

Pooled StDev = 25,84
```

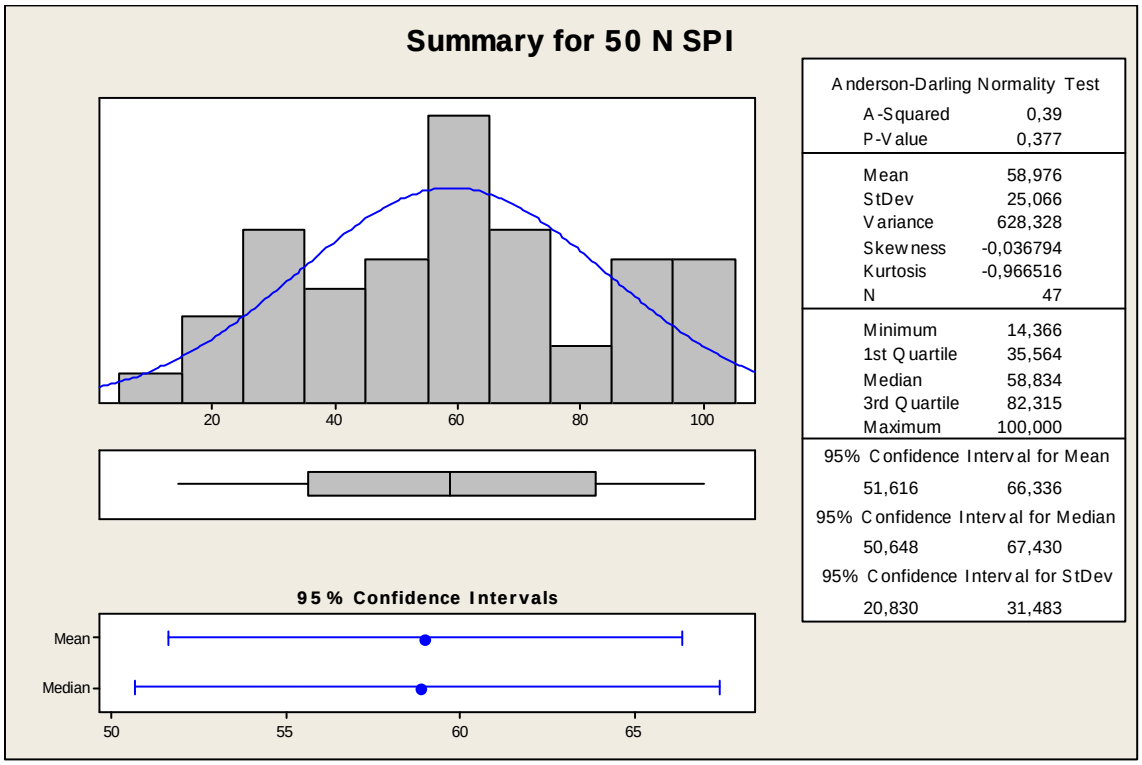
Através da análise Anova e do gráfico Boxplot verifica-se que a amostra com 100 observações é a que aproxima-se mais daquela do universo quanto ao desvio padrão. Ressalta-se que como há valores faltantes no universo, as amostras podem não possuir valores válidos exatos de 25, 50 ou 100 observações.

ii. VARIÁVEL SP I – ÍNDICE DE PROGRESSO SOCIAL

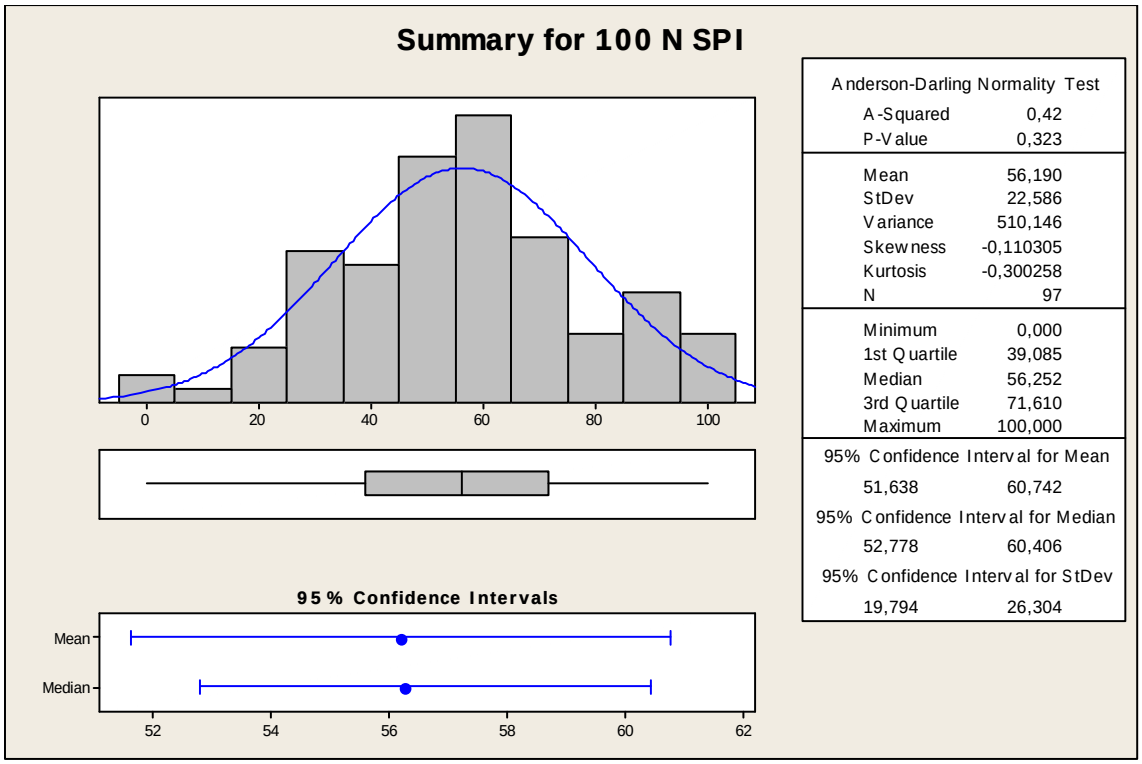
1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES.



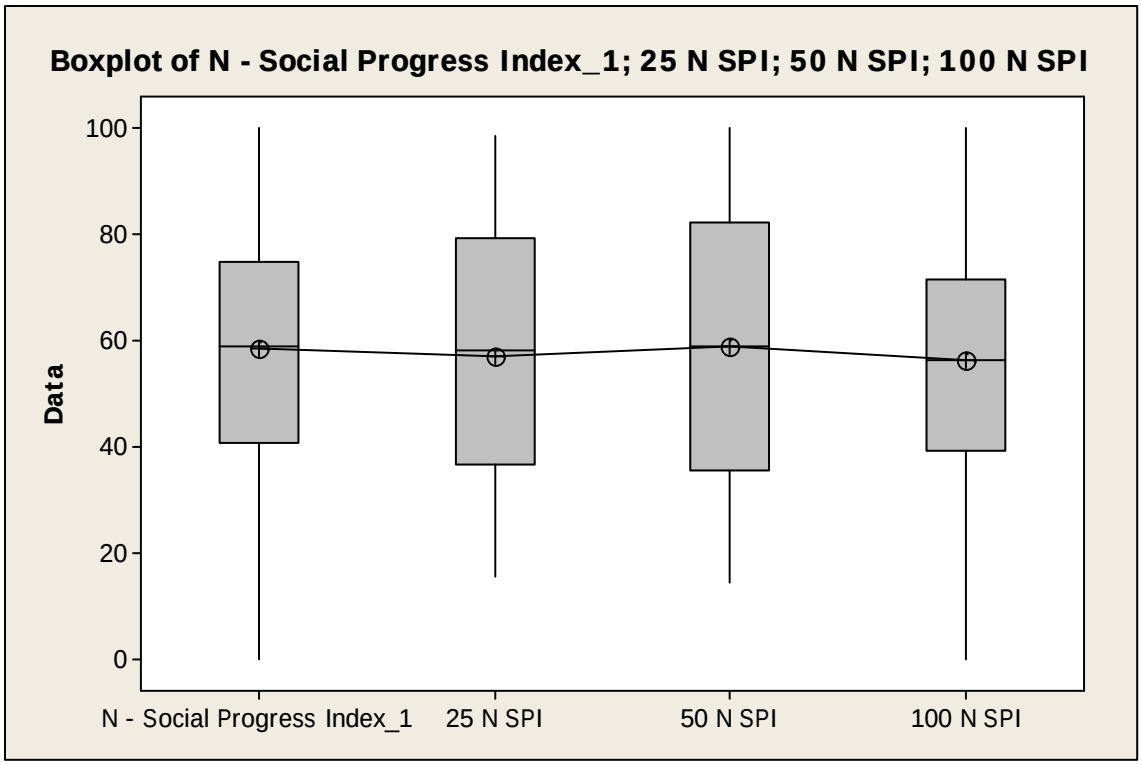
Anderson-Darling Normality Test	
A-Squared	0,36
P-Value	0,416
Mean	56,969
StDev	23,818
Variance	567,283
Skewness	-0,207835
Kurtosis	-0,842699
N	24
Minimum	15,595
1st Quartile	36,477
Median	58,219
3rd Quartile	79,285
Maximum	98,700
95% Confidence Interval for Mean	
46,912	67,027
95% Confidence Interval for Median	
49,561	71,608
95% Confidence Interval for StDev	
18,511	33,411



Anderson-Darling Normality Test	
A-Squared	0,39
P-Value	0,377
Mean	58,976
StDev	25,066
Variance	628,328
Skewness	-0,036794
Kurtosis	-0,966516
N	47
Minimum	14,366
1st Quartile	35,564
Median	58,834
3rd Quartile	82,315
Maximum	100,000
95% Confidence Interval for Mean	
51,616	66,336
95% Confidence Interval for Median	
50,648	67,430
95% Confidence Interval for StDev	
20,830	31,483



2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES



One-way ANOVA: N - Social Progress Index_1; 25 N SPI; 50 N SPI; 100 N SPI

Source	DF	SS	MS	F	P
Factor	3	435	145	0,26	0,854
Error	293	163209	557		

```
Total   296  163644

S = 23,60   R-Sq = 0,27%   R-Sq(adj) = 0,00%



Level                        N   Mean  StDev
N - Social Progress Inde   129  58,69  23,76
25 N SPI                    24  56,97  23,82
50 N SPI                    47  58,98  25,07
100 N SPI                   97  56,19  22,59

                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level                      -----+---------+---------+---------+----
N - Social Progress Inde                (-------*--------)
25 N SPI                   (------------------*------------------)
50 N SPI                            (-------------*-------------)
100 N SPI                         (--------*---------)
                           -----+---------+---------+---------+----
                             50,0      55,0      60,0      65,0

Pooled StDev = 23,60
```

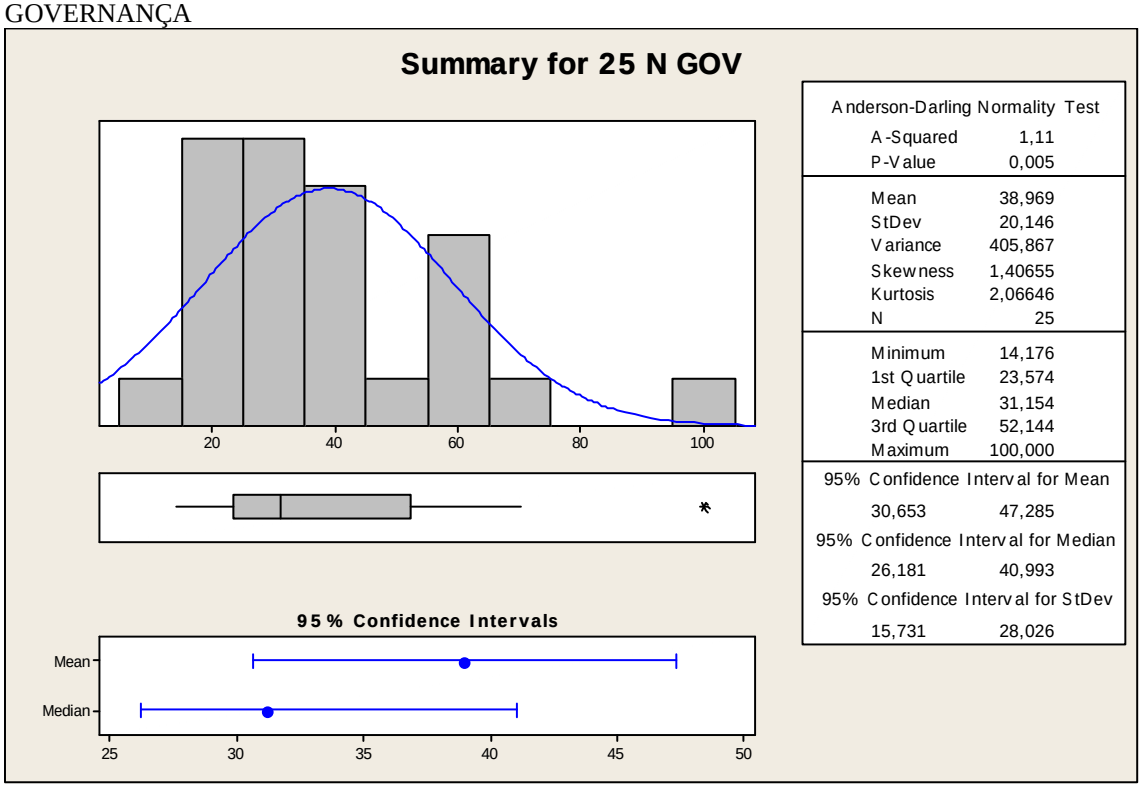
Boxplot of N - Social Progress Index_1; 25 N SPI; 50 N SPI; 100 N SPI

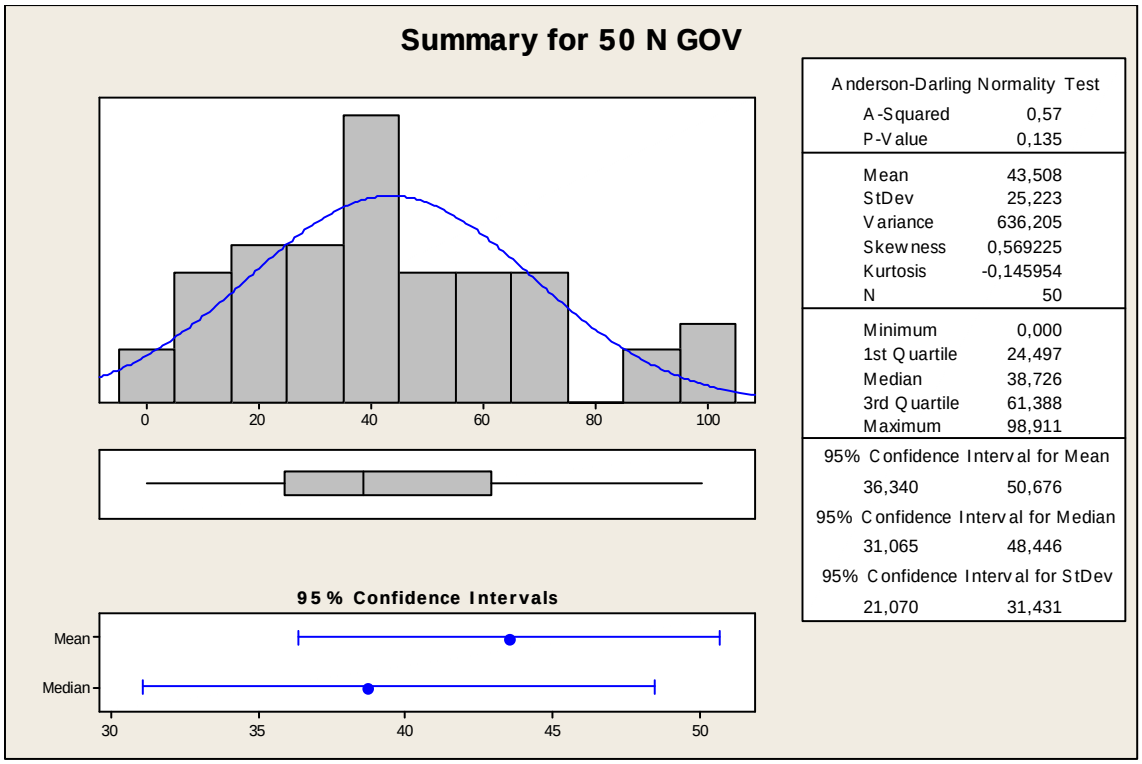
A variável SPI encontrou o desvio padrão e média mais próximo do universo da mesmo na amostra de 100 obsevações. Portanto esta última parece representar melhor o universo desta variável

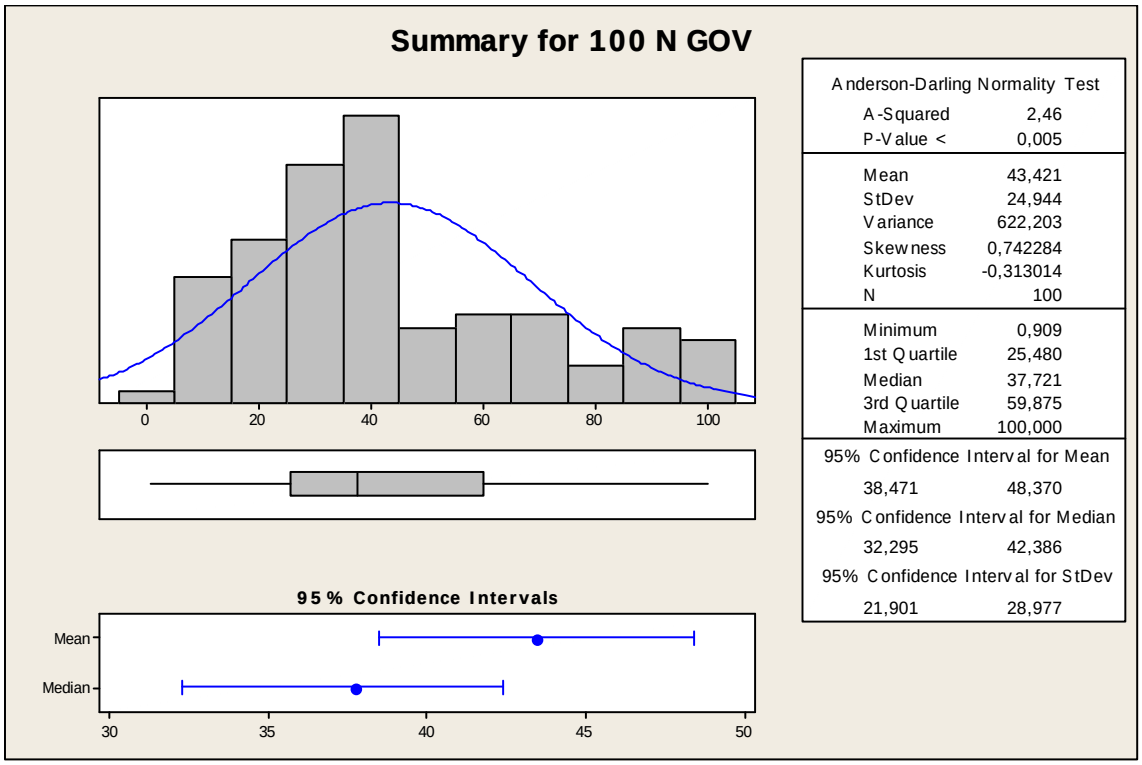
iii. VARIÁVEL ÍNDICE DE GOVERNANÇA

1. AMOSTRAS ALEATÓRIAS COM 25, 50 E 100 OBSERVAÇÕES.

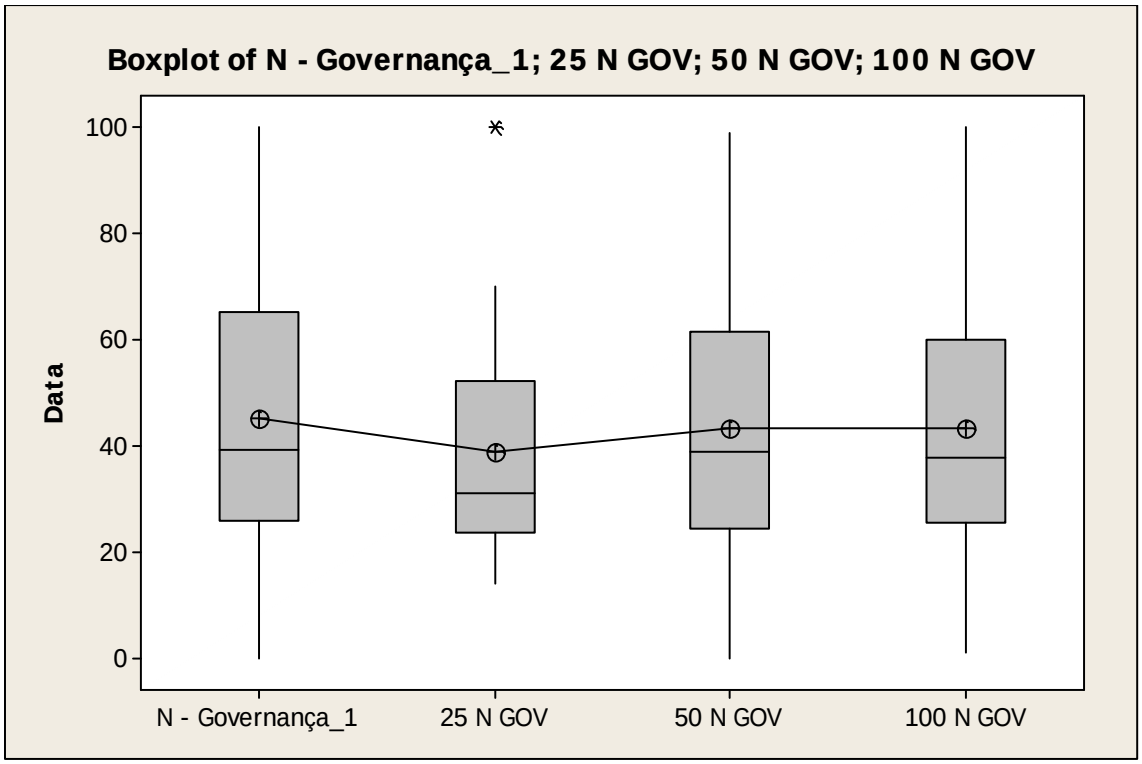
GOVERNANÇA







2. ANOVA COMPARATIVO ENTRE AMOSTRAS ALEATÓRIAS E UNIVERSO DAS OBSERVAÇÕES



One-way ANOVA: N - Governança_1; 25 N GOV; 50 N GOV; 100 N GOV

```
Source   DF      SS   MS     F      P
Factor    3     797  266  0,43  0,732
Error   303  187451  619
Total   306  188248

S = 24,87   R-Sq = 0,42%   R-Sq(adj) = 0,00%


Level               N   Mean  StDev
N - Governança_1  132  45,02  25,46
25 N GOV           25  38,97  20,15
50 N GOV           50  43,51  25,22
100 N GOV         100  43,42  24,94

                  Individual 95% CIs For Mean Based on Pooled StDev
Level               -+---------+---------+---------+--------
N - Governança_1                      (------*------)
25 N GOV            (---------------*---------------)
50 N GOV                        (-----------*----------)
100 N GOV                          (-------*--------)
                    -+---------+---------+---------+--------
                  30,0      36,0      42,0      48,0

Pooled StDev = 24,87
```

Boxplot of N - Governança_1; 25 N GOV; 50 N GOV; 100 N GOV

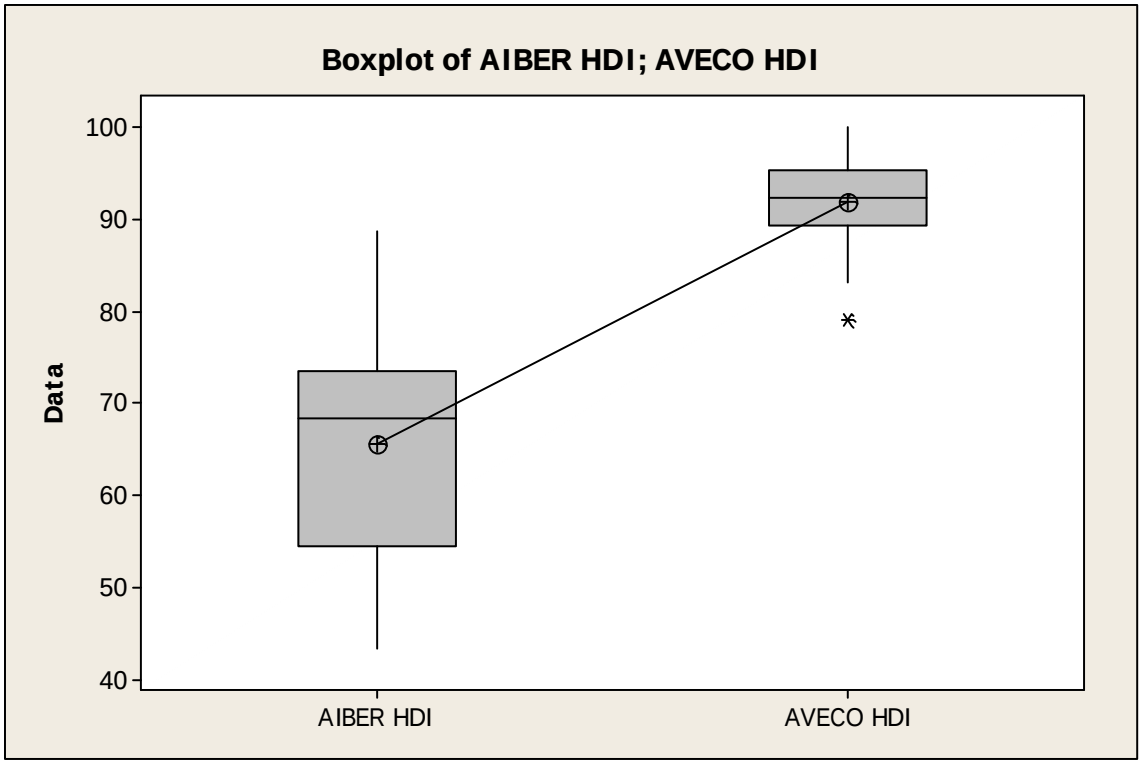
Novamente observamos que a amostra de 100 observações representa melhor o universo, em relação à média e variação medida pelo desvio padrão, em comparação às amostras de 25 e 50 observações.

iv. COMPARAÇÃO DAS VARIÁVEIS POR REGIÃO – AIBER E AVECO

Nesta seção compara-se as variáveis analisadas nas etapas anteriores tomando-se como base o total dos países das regiões AIBER e AVECO. Realiza-se comparação com teste ANOVA e análise gráfica Boxplot; e verifica-se se os resultados das observações dos dois grupos são estatisticamente distintos a partir dos parâmetros de um Teste T de Student com amostras não pareadas, a 95%. Emprega-se o Minitab ® como ferramental.

1. HDI – HUMAN DEVELOPMENT INDEX / IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO

Comparamos com o teste ANOVA o IDH das regiões AIBER E AVECO.



One-way ANOVA: AIBER HDI; AVECO HDI

```
Source  DF       SS      MS       F      P
Factor   1   8122,1  8122,1  101,08  0,000
Error   46   3696,1    80,4
Total   47  11818,2

S = 8,964   R-Sq = 68,73%   R-Sq(adj) = 68,05%
```

```
                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level       N    Mean   StDev  --------+---------+---------+---------+-
AIBER HDI  21  65,671  12,452  (---*---)
AVECO HDI  27  91,892   4,784                           (---*--)
                                --------+---------+---------+---------+-
                                       70        80        90       100

Pooled StDev = 8,964
```

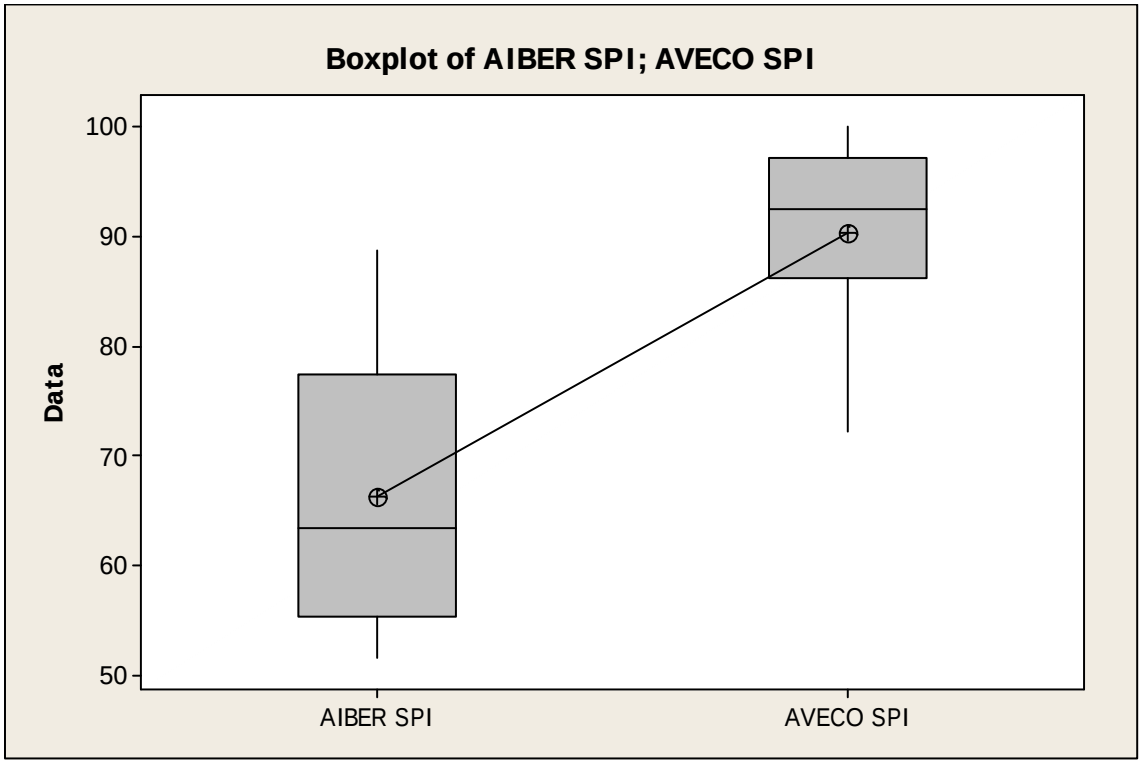
Boxplot of AIBER HDI; AVECO HDI

As médias parecem ser bastante diferentes, e o F resultante é de 101,08. Com o Teste T de Student para amostras não pareadas é possível verificar através do P value (inferior a 1%) que os valores de IDH de AVECO (média de 91,89) são significativamente superiores aos dos países ibéricos AIBER (média de 65,67), como apreentado nos resultados do teste a seguir.

Two-Sample T-Test and CI: AIBER HDI; AVECO HDI

```
Two-sample T for AIBER HDI vs AVECO HDI

            N   Mean  StDev  SE Mean
AIBER HDI  21   65,7   12,5      2,7
AVECO HDI  27  91,89   4,78     0,92


Difference = mu (AIBER HDI) - mu (AVECO HDI)
Estimate for difference:  -26,22
95% CI for difference:  (-32,14; -20,30)
T-Test of difference = 0 (vs not =): T-Value = -9,14  P-Value = 0,000  DF = 24
```

2. SPI – SOCIAL PROGRESS INDEX / IPS – ÍNDICE DE PROGRESSO SOCIAL



One-way ANOVA: AIBER SPI; AVECO SPI

```
Source  DF     SS    MS      F      P
Factor   1   6771  6771  66,32  0,000
Error   46   4697   102
Total   47  11468

S = 10,10   R-Sq = 59,05%   R-Sq(adj) = 58,16%
```

```
                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level       N   Mean  StDev  --------+---------+---------+---------+-
AIBER SPI  21  66,37  12,20  (---*----)
AVECO SPI  27  90,31   8,14                        (---*---)
                               --------+---------+---------+---------+-
                                      70        80        90       100

Pooled StDev = 10,10
```

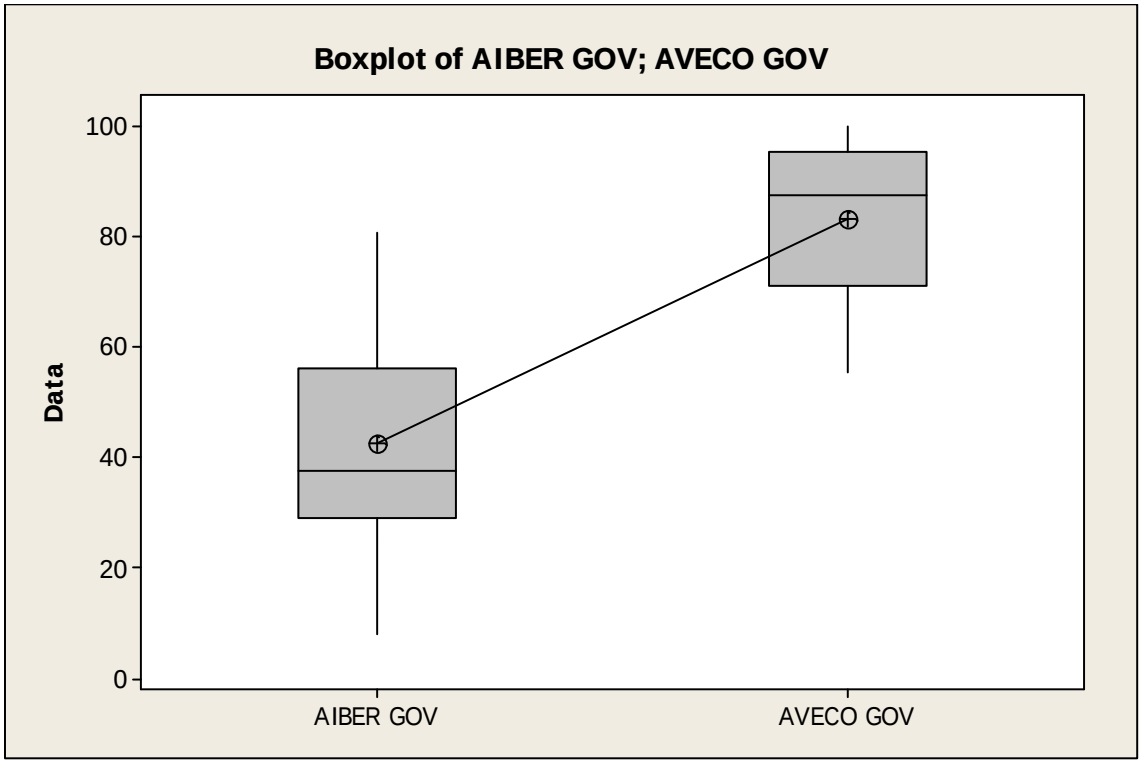
Boxplot of AIBER SPI; AVECO SPI

Ao compararmos os valores de SPI das duas regiões estes aparecem como novamente bastante diferentes, mas o F neste caso é menor, 66,32. Novamente estes valores são verificados como significativamente diferentes, sendo o valor de AVECO (90,31), superior ao 66,4 obtido nos países de AIBER, a 1%.

Two-Sample T-Test and CI: AIBER SPI; AVECO SPI

```
Two-sample T for AIBER SPI vs AVECO SPI

            N   Mean  StDev  SE Mean
AIBER SPI  21   66,4   12,2      2,7
AVECO SPI  27  90,31   8,14      1,6


Difference = mu (AIBER SPI) - mu (AVECO SPI)
Estimate for difference:  -23,94
95% CI for difference:  (-30,22; -17,66)
T-Test of difference = 0 (vs not =): T-Value = -7,75  P-Value = 0,000  DF = 33
```

3. GOVERNANCE INDEX – ÍNDICE DE GOVERNANÇA



One-way ANOVA: AIBER GOV; AVECO GOV

```
Source  DF     SS     MS      F      P
Factor   1  19633  19633  76,93  0,000
Error   46  11740    255
Total   47  31373

S = 15,98   R-Sq = 62,58%   R-Sq(adj) = 61,77%


                             Individual 95% CIs For Mean Based on
                             Pooled StDev
Level       N   Mean  StDev  ------+---------+---------+---------+---
AIBER GOV  21  42,48  18,82  (---*----)
AVECO GOV  27  83,25  13,38                            (----*---)
                             ------+---------+---------+---------+---
                                  45        60        75        90

Pooled StDev = 15,98
```

Boxplot of AIBER GOV; AVECO GOV

Novamente os valores de média e desvio padrão são bastante diferentes entre ACECO e AIBER, com F com valor intermediário entre SPI e HDI, 76,93. Novamente no teste T de Student observa-se diferença significativa entre os valores de AIBER (42,5) e AVECO (83,3), desta vez para Governança, conforem teste T de Student de amostra não pareadas, a 1%.

Two-Sample T-Test and CI: AIBER GOV; AVECO GOV

```
Two-sample T for AIBER GOV vs AVECO GOV

            N  Mean  StDev  SE Mean
AIBER GOV  21  42,5   18,8      4,1
AVECO GOV  27  83,3   13,4      2,6


Difference = mu (AIBER GOV) - mu (AVECO GOV)
Estimate for difference:  -40,77
95% CI for difference:  (-50,62; -30,92)
T-Test of difference = 0 (vs not =): T-Value = -8,41  P-Value = 0,000  DF = 34
```

b. CONCLUSÃO

Conclui-se dos testes realizados que para as variáveis Índice de Desenvolvimento Humano, Índice de Progresso Social e Índice de Governança, encontra-se diferenças significativas entre AIBER e AVECO, com a ultima região possuindo valores significativamente à primeira com significância a 1%.

7. ANÁLISE DE COMPONENTES PRINCIPAIS

Este trabalho tem como objetivo realizar uma análise das correlações e dos componentes principais dos dados das variáveis do EIXO PESSOAS para 132 países, comparados a partir de dezessete variáveis quantitativas, seis das quais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi

realizada através dos dados obtidos nas fontes de cada variável. Empregou-se o Minitab ® para as análises estatísticas.

a. ANÁLISE DE COMPONENTES PRINCIPAIS

A Análise de Componentes Principais tem como objetivo contornar uma potencial situação de correlação entre variáveis independentes de uma base de dados através da recombinação destas em variáveis novas sem correlação entre si, e que representam um conjunto de correlações entre as variáveis originais. Desta forma a análise prevê a construção dessas novas variáveis, o que é realizado através da aplicação dos procedimentos estatísticos através do Minitab® entre os quais a identificação do conteúdo e significado destas novas variáveis. É necessário também balancear com parcimônia a composição de quantidade de novas variáveis e sua força explicativa, de modo a reduzir a complexidade do novo modelo proposto dentro de uma força explicativa razoável.

i. ANÁLISE DAS VARIÁVEIS DA AMOSTRA – COMPONENTES PRINCIPAIS

Realiza-se uma análise de componentes principais para identificar a relação entre quantidade de novas variáveis e força explicativa, conforme apresentado na Tabela 2. Encontra-se a situação em que 4 variáveis explicam 83,3% das situações, com Eigenvalue 1,099. O Gráfico 1 apresenta a visão de distribuição e destas.variáveis e a força explicativa cumulativa das mesmas e o Gráfico 2 sua correlação..
Como forma de simplificar a análise e entendendo-se como adequado ao objetivo deste estudo tal força explicativa, procede-se como próximo passo à construção e entendimento destes quatro componentes.

Tabela 2 – Análise de Componentes Principais

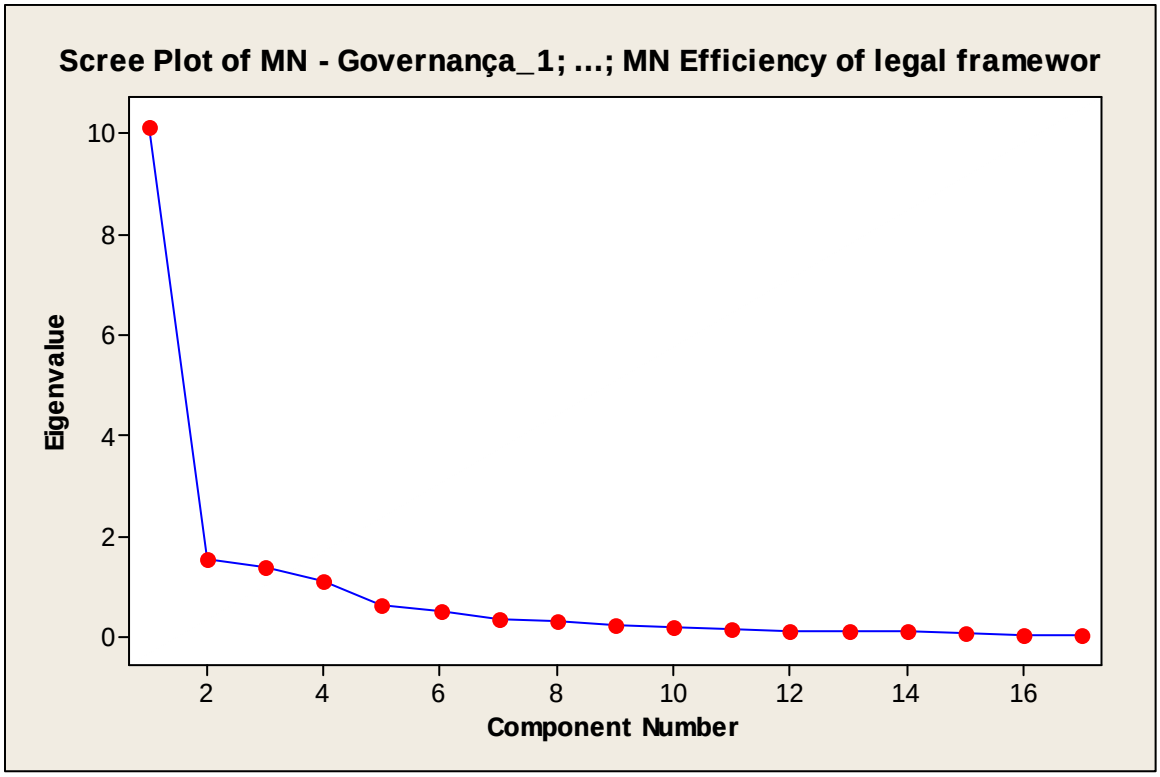
Principal Component Analysis: MN - Governa; MN - Happy P; MN - GINNI; MN - Soci

Eigenanalysis of the Correlation Matrix

Eigenvalue	10,108	1,543	1,403	1,099	0,628	0,506	0,352	0,323	0,216
Proportion	0,595	0,091	0,083	0,065	0,037	0,030	0,021	0,019	0,013
Cumulative	0,595	0,685	0,768	0,833	0,870	0,899	0,920	0,939	0,952

Eigenvalue	0,175	0,149	0,131	0,106	0,096	0,066	0,054	0,044
Proportion	0,010	0,009	0,008	0,006	0,006	0,004	0,003	0,003
Cumulative	0,962	0,971	0,978	0,985	0,990	0,994	0,997	1,000

Gráfico 1 – Análise de Componentes Principais



Como se verifica no Gráfico 2 é decrescente de forma acelerada a força explicativa de se adicionar variáveis novas a partir do quinto componente principal, pois as quatro primeiras variáveis tem força explicativa de 83% , com a quinta variável adicionada esta sobe para 87% e para a sexta, 90% (veja também Tabela 2).

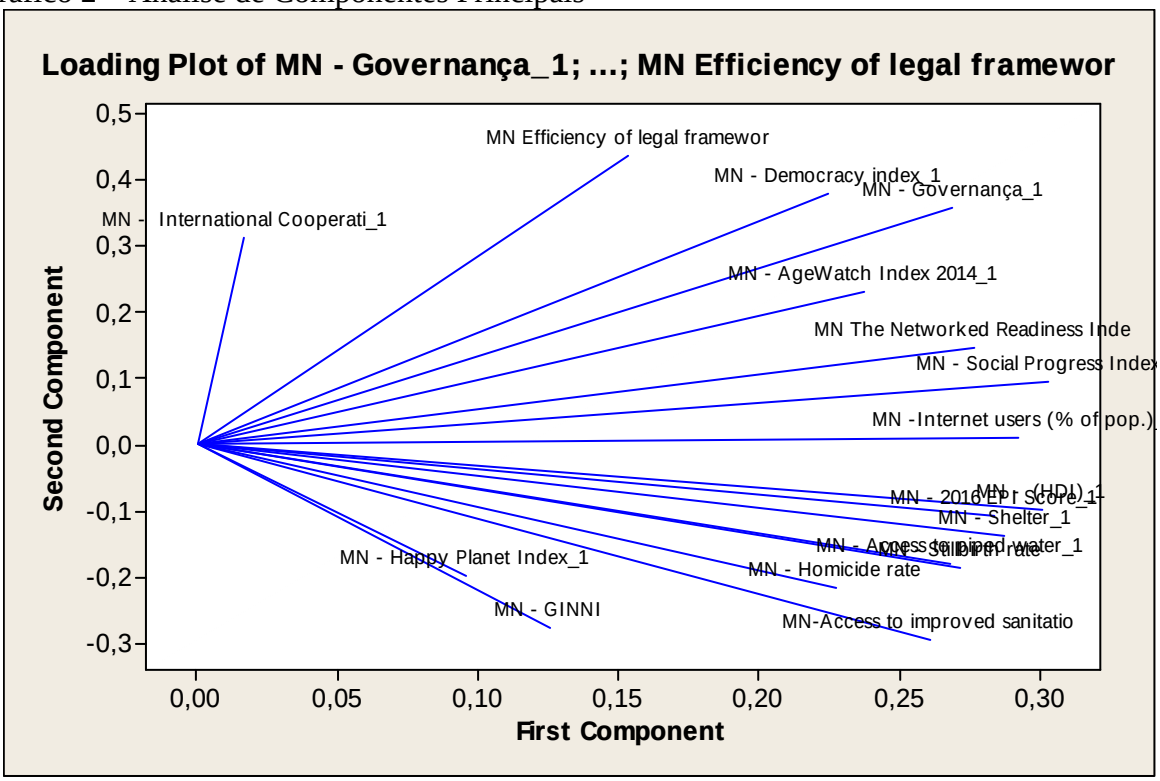
Através a aplicação da análise de componentes principais no Minitab ® gera-se quatro novas variáveis, que são exploradas quanto à sua composição e significado inicialmente através da análise da Tabela 3 e Gráfico 3, resultados desta operação. Na seção seguinte analisa-se as correlações internas entre as 17 variáveis para cada uma das variáveis novas geradas. Na tabela 3 apresenta-se os pesos das variáveis originais na composição de cada uma dos quatro novos indicadores ou Componentes Principais gerados, novas variáveis geradas a partir da recombinação das originais, privilegiando uma situação de não correlação ou minimização da mesma. Detaca-se a composição do componente principal 1 (PC1), com o peso concentrado no índice de progresso social, índice de desenvolvimento humano e percentual usuários de internet, os dois primeiros com peso superior a 0,3.

Tabela 3 – Composição dos Quatro componentes Principais Gerados na Análise

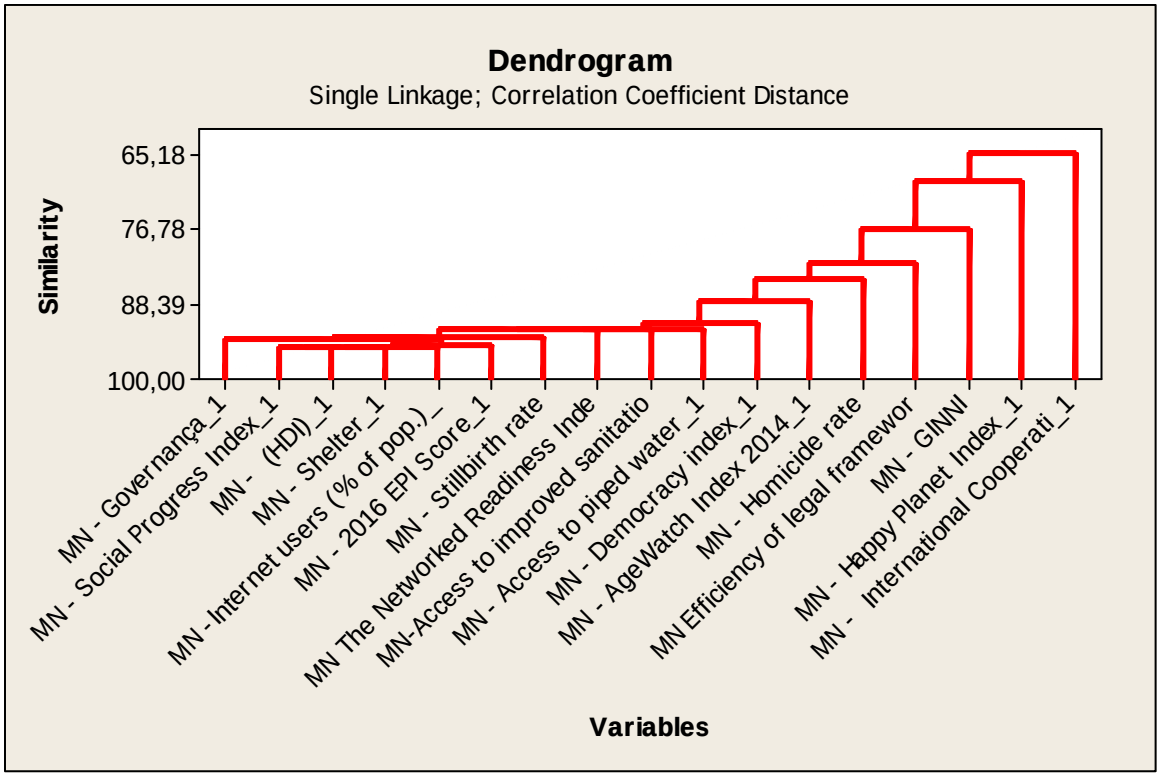
Variable	PC1	PC2	PC3	PC4
MN - Governança_1	0,269	0,357	-0,071	0,073
MN - Happy Planet Index_1	0,095	-0,198	0,527	-0,429
MN - GINNI	0,125	-0,277	-0,493	0,351
MN - Social Progress Index_1	0,303	0,093	0,053	0,033
MN - Access to piped water_1	0,268	-0,180	0,207	0,024
MN-Access to improved sanitatio	0,261	-0,297	0,107	0,003
MN - International Cooperati_1	0,017	0,313	0,444	0,639
MN - 2016 EPI Score_1	0,284	-0,109	0,121	0,063
MN - Shelter_1	0,287	-0,138	0,093	-0,084
MN - Stillbirth rate	0,271	-0,186	0,071	0,007
MN - Homicide rate	0,227	-0,216	-0,229	0,133
MN -Internet users (% of pop.)_	0,293	0,010	-0,118	0,072
MN - (HDI)_1	0,301	-0,100	0,033	0,002
MN - AgeWatch Index 2014_1	0,237	0,230	-0,127	-0,133
MN - Democracy index_1	0,224	0,378	0,161	0,138

MN The Networked Readiness Inde	0,276	0,146	-0,198	-0,107
MN Efficiency of legal framewor	0,154	0,436	-0,207	-0,442

Gráfico 2 – Análise de Componentes Principais



O Dendograma gerado com a mesma base de variáveis corrobora a avaliação dos componentes principais.



Como assinalado na tabela abaixo, quatro agrupamentos representariam 82% de similaridade entre as variáveis observadas.

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	16	95,0820	0,098359	4	13	4	2
2	15	95,0236	0,099529	4	9	4	3
3	14	94,9318	0,101364	4	12	4	4
4	13	94,6877	0,106245	4	8	4	5
5	12	93,8659	0,122682	1	4	1	6
6	11	93,5483	0,129034	1	10	1	7
7	10	92,3701	0,152598	1	16	1	8
8	9	92,3115	0,153771	1	6	1	9
9	8	92,2522	0,154956	1	5	1	10
10	7	91,3555	0,172890	1	15	1	11
11	6	87,8248	0,243505	1	14	1	12
12	5	84,5367	0,309265	1	11	1	13
13	4	82,1468	0,357064	1	17	1	14
14	3	76,6656	0,466689	1	3	1	15
15	2	69,2270	0,615459	1	2	1	16
16	1	65,1766	0,696467	1	7	1	17

Realiza-se uma análise de correlação entre os quatro componentes principais gerados e chega-se a uma correlação de zero entre eles, confirmando os resultados esperados para a aplicação desta metodologia (Tabela 4)

Tabela 4 – Correlação entre os componentes principais gerados

Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	0,000		
	1,000		

PC3	0,000	0,000	
	1,000	1,000	
PC4	-0,000	0,000	0,000
	1,000	1,000	1,000

Cell Contents: Pearson correlation
P-Value

ii. ANÁLISE DOS COMPONENTES PRINCIPAIS GERADOS

Aplica-se uma regressão multivariada com cada componente principal como variável dependente e as variáveis originais como independentes para identificar quais as variáveis que trazem maior força explicativa para cada componente principal, o que indicará por consequência o significado do componente principal analisado em função de suas variáveis componentes e seu respectivo peso. O método empregado é o *stepwise*, para otimizar o significado das regressões multivariadas resultantes. Atribui-se ao final de cada análise uma nomenclatura significativa para cada componente principal.

1. ANÁLISE COMPONENTE PRINCIPAL 1

A regressão *stepwise* gerada combina uma variável dependente que possui correlação com todas as variáveis dependentes, de forma otimizada, de forma que o *p value* destas últimas é sempre significativo. Desta forma, a análise do componente principal é realizada pelo crescimento do r-quadrado a cada etapa do procedimento *stepwise*. Nota-se nesta sequencia que a primeira variável adicionada é o SPI, ou índice de progresso social, com r-quadrado de 94%, seguido pelo HDI/IDH, ou índice de desenvolvimento humano, que elevado o r-quadrado para 97%.(Tabela 5). Ambos são índices sintéticos, porém privilegiam elementos diferentes. Pelo peso das duas variáveis opta-se por nomear esta primeira componente principal de INDICADORES DE PROGRESSO E DESENVLVIMENTO HUMANO.

Tabela 5 – Regressão *stepwise* - Componente Principal 1

Stepwise Regression: PC1 versus IDH - FIT; MN - Governança_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 18 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-7,811	-7,864	-7,894	-8,024	-8,179	-8,037
MN - Social Progress Index_1	0,1330	0,0759	0,0644	0,0639	0,0623	0,0480
T-Value	41,52	15,44	14,03	15,93	17,00	10,79
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
MN - (HDI)_1		0,0568	0,0494	0,0333	0,0314	0,0312
T-Value		12,90	12,43	7,75	8,01	8,65
P-Value		0,000	0,000	0,000	0,000	0,000
MN The Networked Readiness Inde			0,0243	0,0274	0,0259	0,0205
T-Value			6,62	8,46	8,76	7,00
P-Value			0,000	0,000	0,000	0,000

MN-Access to improved sanitatio				0,0138	0,0116	0,0158
T-Value				6,38	5,77	7,75
P-Value				0,000	0,000	0,000
MN - Homicide rate					0,0089	0,0085
T-Value					5,26	5,42
P-Value					0,000	0,000
MN - Governança_1						0,0153
T-Value						4,90
P-Value						0,000
S	0,845	0,561	0,486	0,424	0,386	0,355
R-Sq	92,99	96,94	97,72	98,27	98,58	98,81
R-Sq(adj)	92,93	96,89	97,67	98,22	98,53	98,76

2. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS

A regressão com as variáveis SPI, HDI e Internet Users indica que os betas das variáveis indica o peso das mesmas na explicação do Componente Principal 1, e o peso segue a sequencia apresentada SPI (0,0,639), HDI (0,033), Disponibilidade de rede representa 0,027 de peso e Acesso a recursos sanitários com peo 0,013. A força da regressão é de 98% (r-quadrado).

Regression Analysis: PC1 versus MN - (HDI)_1; MN - Social Prog; ...

The regression equation is
PC1 = - 8,02 + 0,0333 MN - (HDI)_1 + 0,0639 MN - Social Progress Index_1
+ 0,0138 MN-Access to improved sanitatio
+ 0,0274 MN The Networked Readiness Inde

Predictor	Coef	SE Coef	T	P
Constant	-8,0240	0,1036	-77,48	0,000
MN - (HDI)_1	0,033276	0,004296	7,75	0,000
MN - Social Progress Index_1	0,063944	0,004014	15,93	0,000
MN-Access to improved sanitatio	0,013809	0,002166	6,38	0,000
MN The Networked Readiness Inde	0,027384	0,003237	8,46	0,000

S = 0,424360 R-Sq = 98,3% R-Sq(adj) = 98,2%

Analysis of Variance

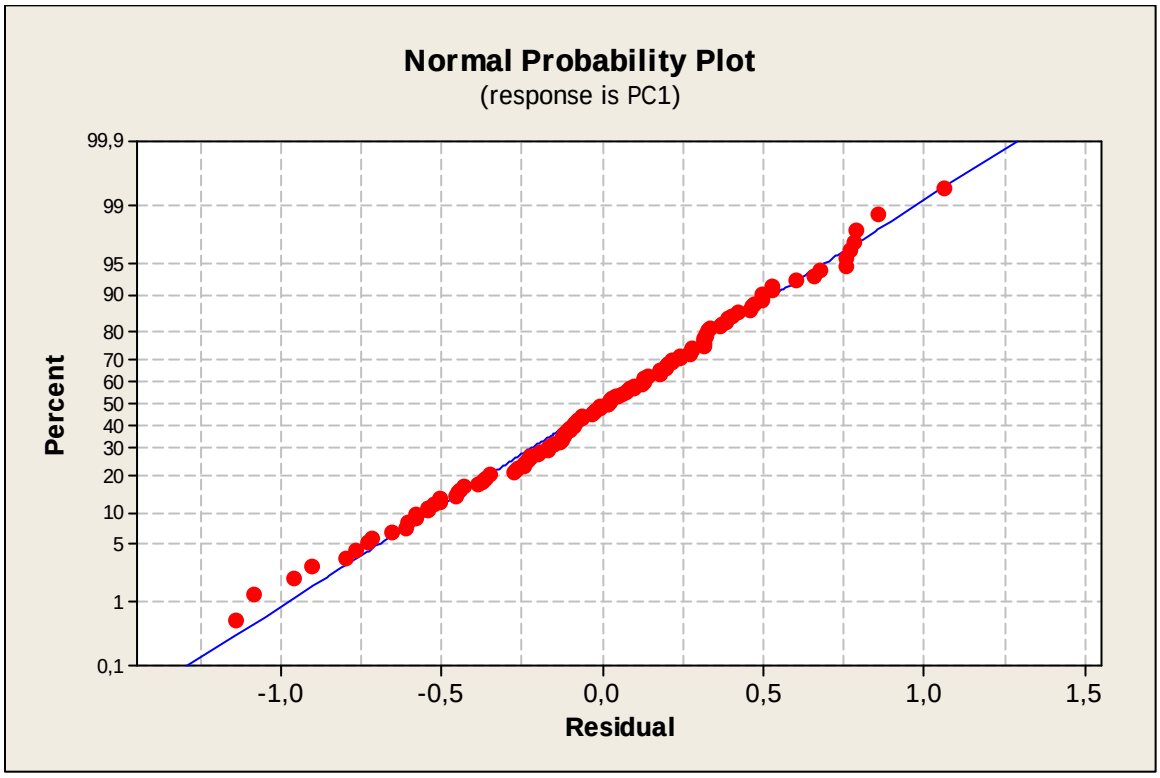
Source	DF	SS	MS	F	P
Regression	4	1301,27	325,32	1806,50	0,000
Residual Error	127	22,87	0,18		
Total	131	1324,14			

Source	DF	Seq SS
MN - (HDI)_1	1	1208,69
MN - Social Progress Index_1	1	74,91
MN-Access to improved sanitatio	1	4,78
MN The Networked Readiness Inde	1	12,89

Unusual Observations

Obs	MN - (HDI)_1	PC1	Fit	SE Fit	Residual	St Resid
19	9	-4,7147	-4,2086	0,1582	-0,5060	-1,29 X
23	0	-4,6732	-3,7127	0,1765	-0,9605	-2,49RX
24	7	-5,9158	-6,3070	0,1613	0,3912	1,00 X
58	88	2,2251	2,0958	0,1460	0,1293	0,32 X
65	79	0,5552	1,4605	0,0655	-0,9053	-2,16R
74	27	-4,3479	-5,2050	0,1024	0,8572	2,08R
82	64	-1,6287	-0,4881	0,0702	-1,1407	-2,73R
89	95	4,9757	4,1926	0,1511	0,7831	1,97 X
96	56	-1,2871	-0,1990	0,0697	-1,0882	-2,60R
131	25	-4,0922	-5,1514	0,0883	1,0592	2,55R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.



iii. ANÁLISE COMPONENTE PRINCIPAL 2

A construção da segunda Componente Principal e sua análise procede da mesma forma que na anterior. Nesta no entanto três variáveis são combinadas para obtenção de um r-quadrado maior do que 80% (87%), como se observa na Tabela 6: Eficiência do arcabouço legal em resolver disputas, Acesso a recursos sanitparios, Democracia e Cooperação internacional. Estes vetores são bem mais dispersos entre si do que o do primeiro componente. Opta-se por denominar esta variável RESOLUTIVIDADE LEGAL, RECURSOS SANITÁRIOS E DEMOCRACIA.

Stepwise Regression: PC2 versus IDH - FIT; MN - Governança_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 18 predictors, with N = 132

Step	1	2	3	4	5
Constant	-1,7052	-0,6437	-1,0902	-2,2955	-2,0516
MN Efficiency of legal framewor	0,0351	0,0437	0,0323	0,0375	0,0382
T-Value	7,35	11,24	11,07	16,61	18,46
P-Value	0,000	0,000	0,000	0,000	0,000
MN-Access to improved sanitatio		-0,0209	-0,0274	-0,0262	-0,0219
T-Value		-8,94	-15,72	-19,87	-14,87
P-Value		0,000	0,000	0,000	0,000
MN - Democracy index_1			0,0280	0,0205	0,0220
T-Value			11,46	10,29	11,94
P-Value			0,000	0,000	0,000
MN - International Cooperati_1				0,0219	0,0204
T-Value				9,92	10,02
P-Value				0,000	0,000
MN - Homicide rate					-0,0087
T-Value					-5,13
P-Value					0,000
MN - Happy Planet Index_1					
T-Value					
P-Value					
S	1,05	0,827	0,583	0,439	0,401
R-Sq	29,35	56,37	78,46	87,87	89,97

iv. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS

A análise de regressão do PC2 –com as três variáveis principais gerou uma regressão com r-quadrado de 78%. A variável Eficiência do sistema legal possui maior peso e as duas seguintes, Democracia e acesso a recursos sanitários possuem pesos muito próximos, porém esta última possui coeficiente negativo.

Regression Analysis: PC2 versus MN-Access to; MN - Democra; MN Efficienc

The regression equation is
PC2 = - 1,09 - 0,0274 MN-Access to improved sanitatio
+ 0,0280 MN - Democracy index_1 + 0,0323 MN Efficiency of legal framewor

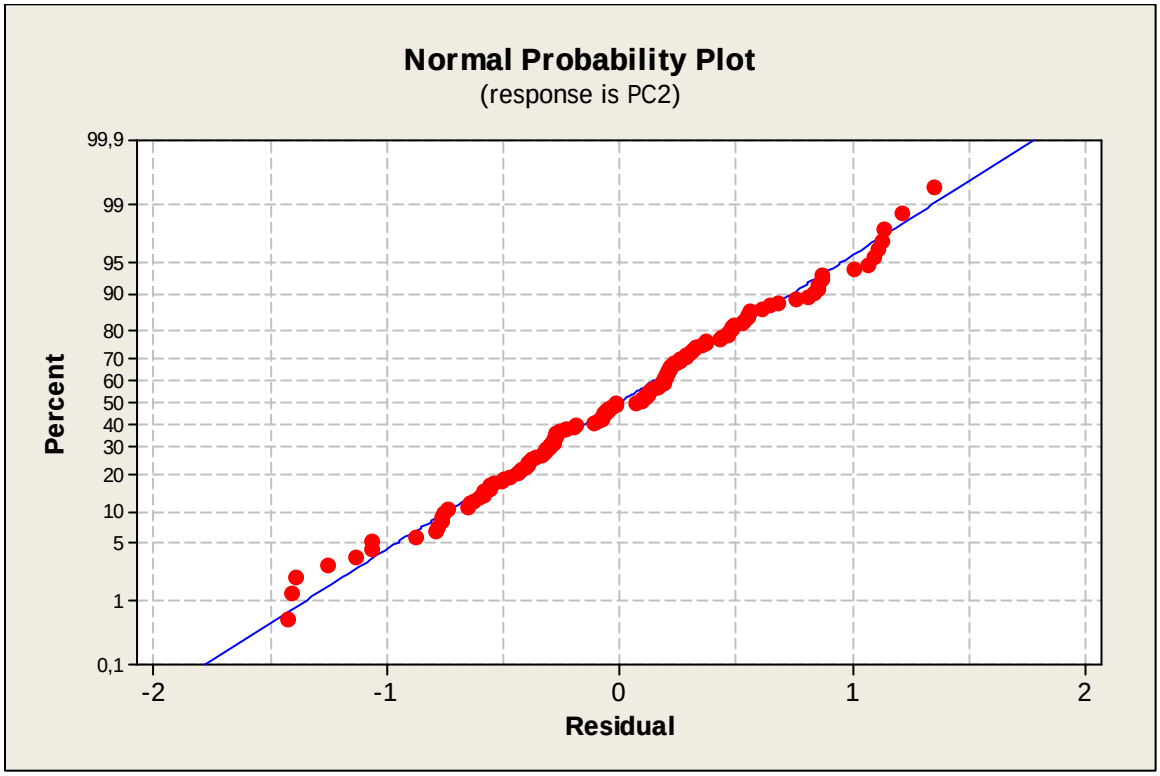
Predictor	Coef	SE Coef	T	P
Constant	-1,0902	0,1667	-6,54	0,000
MN-Access to improved sanitatio	-0,027428	0,001745	-15,72	0,000
MN - Democracy index_1	0,027997	0,002443	11,46	0,000
MN Efficiency of legal framewor	0,032270	0,002916	11,07	0,000

S = 0,583228 R-Sq = 78,5% R-Sq(adj) = 78,0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	158,632	52,877	155,45	0,000
Residual Error	128	43,540	0,340		
Total	131	202,172			

```
Source                            DF  Seq SS
MN-Access to improved sanitatio    1  27,554
MN - Democracy index_1             1  89,421
MN Efficiency of legal framewor    1  41,658


Unusual Observations

      MN-Access
     to improved
Obs    sanitatio      PC2      Fit  SE Fit  Residual  St Resid
 10           94  -2,7322  -1,4795  0,0985   -1,2527     -2,18R
 25           99   1,7284   0,3823  0,0785    1,3462      2,33R
 58           45  -1,2779   0,1294  0,1577   -1,4073     -2,51R
 89           45   1,9784   3,3702  0,1690   -1,3918     -2,49R
103           60   1,6774   0,4616  0,1491    1,2159      2,16R
125           97   0,3668  -0,7638  0,1747    1,1306      2,03R
130           90  -2,3428  -0,9176  0,0694   -1,4252     -2,46R

R denotes an observation with a large standardized residual.
```



v. ANÁLISE COMPONENTE PRINCIPAL 3

Seguindo a análise, observa-se na Tabela 7 que o componente principal 3 é explicado por três variáveis, Índice de felicdade planetária, Coperação internacional e índice de GINNI(indicador de distribuição de renda e igualdade social), com r-quadrado de 86%, o que leva a denominação desta componente para FELICIDADE, COOPERAÇÃO E EQUIDADE.

Stepwise Regression: PC3 versus IDH - FIT; MN - Governança_1; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15

Response is PC3 on 18 predictors, with N = 132
```

Step	1	2	3	4	5
Constant	-1,7032	-3,6418	-1,7176	-0,9543	-0,7591
MN - Happy Planet Index_1	0,03496	0,03534	0,03232	0,03323	0,02888
T-Value	9,11	12,56	17,87	24,45	25,41
P-Value	0,000	0,000	0,000	0,000	0,000
MN - International Cooperati_1		0,03346	0,02938	0,02799	0,02615
T-Value		10,63	14,47	18,32	22,04
P-Value		0,000	0,000	0,000	0,000
MN - GINNI			-0,0251	-0,0251	-0,0284
T-Value			-13,75	-18,34	-25,72
P-Value			0,000	0,000	0,000
MN Efficiency of legal framewor				-0,01497	-0,01758
T-Value				-10,03	-14,96
P-Value				0,000	0,000
MN - Access to piped water_1					0,00705
T-Value					9,52
P-Value					0,000
MN - Homicide rate					
T-Value					
P-Value					
S	0,929	0,681	0,434	0,326	0,249
R-Sq	38,95	67,45	86,86	92,67	95,74
R-Sq(adj)	38,48	66,95	86,55	92,44	95,57

vi. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS

O R-quadrado da regressão com as três principais variáveis chega a 86%, com o maior coeficiente para o índice de felicidade planetária(HPI), seguido pelo Índice da cooperação internacional e ao final o Índice de Ginni, com coeficiente negativo. Dado que na amostra países com maior HPI possuem também piores GINNIs, em especial nos países iberoamericanos, esta correlação parece ser coerente com o observado.

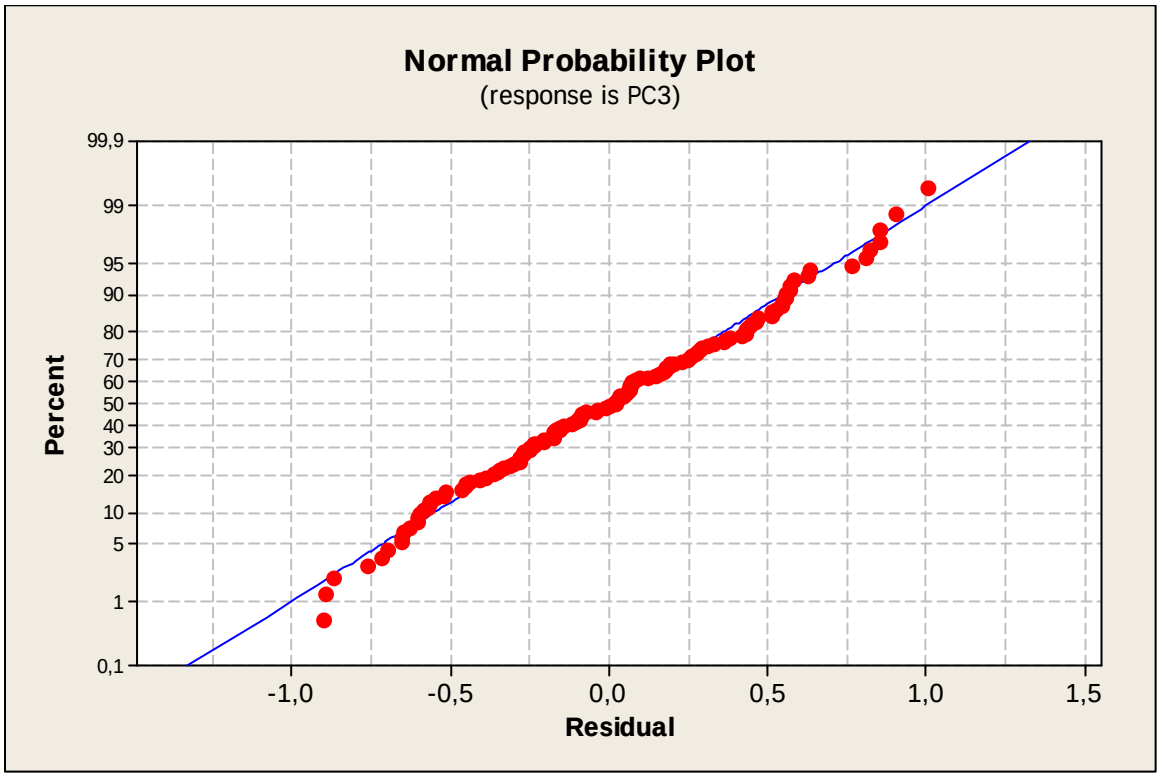
Regression Analysis: PC3 versus MN - GINNI; MN - Happy P; MN - Intern

```
The regression equation is
PC3 = - 1,72 - 0,0251 MN - GINNI + 0,0323 MN - Happy Planet Index_1
      + 0,0294 MN -  International Cooperati_1
```

Predictor	Coef	SE Coef	T	P
Constant	-1,7176	0,2054	-8,36	0,000
MN - GINNI	-0,025085	0,001825	-13,75	0,000
MN - Happy Planet Index_1	0,032319	0,001808	17,87	0,000
MN - International Cooperati_1	0,029375	0,002030	14,47	0,000

```
S = 0,434444   R-Sq = 86,9%   R-Sq(adj) = 86,6%

Analysis of Variance

Source           DF       SS      MS       F      P
```

```
Regression        3  159,681  53,227  282,01  0,000
Residual Error  128   24,159   0,189
Total           131  183,840


Source                            DF  Seq SS
MN - GINNI                         1  62,716
MN - Happy Planet Index_1          1  57,437
MN -  International Cooperati_1    1  39,529


Unusual Observations

      MN -
Obs  GINNI      PC3      Fit  SE Fit  Residual  St Resid
 22     77  -1,5570  -0,6886  0,0517   -0,8684     -2,01R
 45     74   0,0497  -0,8558  0,0498    0,9055      2,10R
 89     70  -0,7999   0,0981  0,0613   -0,8980     -2,09R
102     61  -1,1875  -2,1949  0,1208    1,0074      2,41R
125     45  -0,8547   0,0407  0,0757   -0,8955     -2,09R
130     48   0,6872  -0,1193  0,1320    0,8065      1,95 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```



vii. ANÁLISE COMPONENTE PRINCIPAL 4

A regressão *stepwise* com a componente principal 4 leva também a três variáveis relevantes que explicam 82% da variável dependente. Cooperação internacional, índice de Ginni e Eficiência do arcabouço legal em resolver conflitos, optando-se por denominar esta variável como COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL.

Stepwise Regression: PC4 versus IDH - FIT; MN - Governança_1; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15

Response is PC4 on 18 predictors, with N = 132
```

Step	1	2	3	4	5
Constant	-2,133	-3,792	-2,596	-1,617	-1,398
MN - International Cooperati_1	0,03716	0,04093	0,03886	0,03838	0,03479
T-Value	10,29	14,41	18,87	37,93	41,95
P-Value	0,000	0,000	0,000	0,000	0,000
MN - GINNI		0,02345	0,02328	0,02108	0,01918
T-Value		9,25	12,72	23,27	27,49
P-Value		0,000	0,000	0,000	0,000
MN Efficiency of legal framewor			-0,02193	-0,02060	-0,02452
T-Value			-10,92	-20,84	-29,60
P-Value			0,000	0,000	0,000
MN - Happy Planet Index_1				-0,01808	-0,02004
T-Value				-20,09	-28,83
P-Value				0,000	0,000
MN - Democracy index_1					0,00746
T-Value					10,21
P-Value					0,000
MN - Homicide rate					
T-Value					
P-Value					
S	0,781	0,608	0,439	0,216	0,160
R-Sq	44,89	66,87	82,85	95,90	97,76
R-Sq(adj)	44,46	66,36	82,45	95,77	97,67

viii. REGRESSÃO COM OS AS VARÍAVEIS RELVANTES DOS COMPONENTES PRINCIPAIS

A análise do componente principal 4 - COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL resultou um R-quadrado de 82%. Cooperação interncional é a varíavel de maior peso, seguida pelo Índice de Ginni, ambas com coeficiente positivo. Eficiência legal e moradia possui coeficiente negativo e é a variável de menor peso.

Regression Analysis: PC4 versus MN - GINNI; MN - Intern; MN Efficienc

```
The regression equation is
PC4 = - 2,60 + 0,0233 MN - GINNI + 0,0389 MN -  International Cooperati_1
      - 0,0219 MN Efficiency of legal framewor
```

Predictor	Coef	SE Coef	T	P
Constant	-2,5960	0,2093	-12,40	0,000
MN - GINNI	0,023284	0,001831	12,72	0,000
MN - International Cooperati_1	0,038860	0,002060	18,87	0,000
MN Efficiency of legal framewor	-0,021932	0,002008	-10,92	0,000

```
S = 0,439182   R-Sq = 82,9%   R-Sq(adj) = 82,5%
```

Analysis of Variance

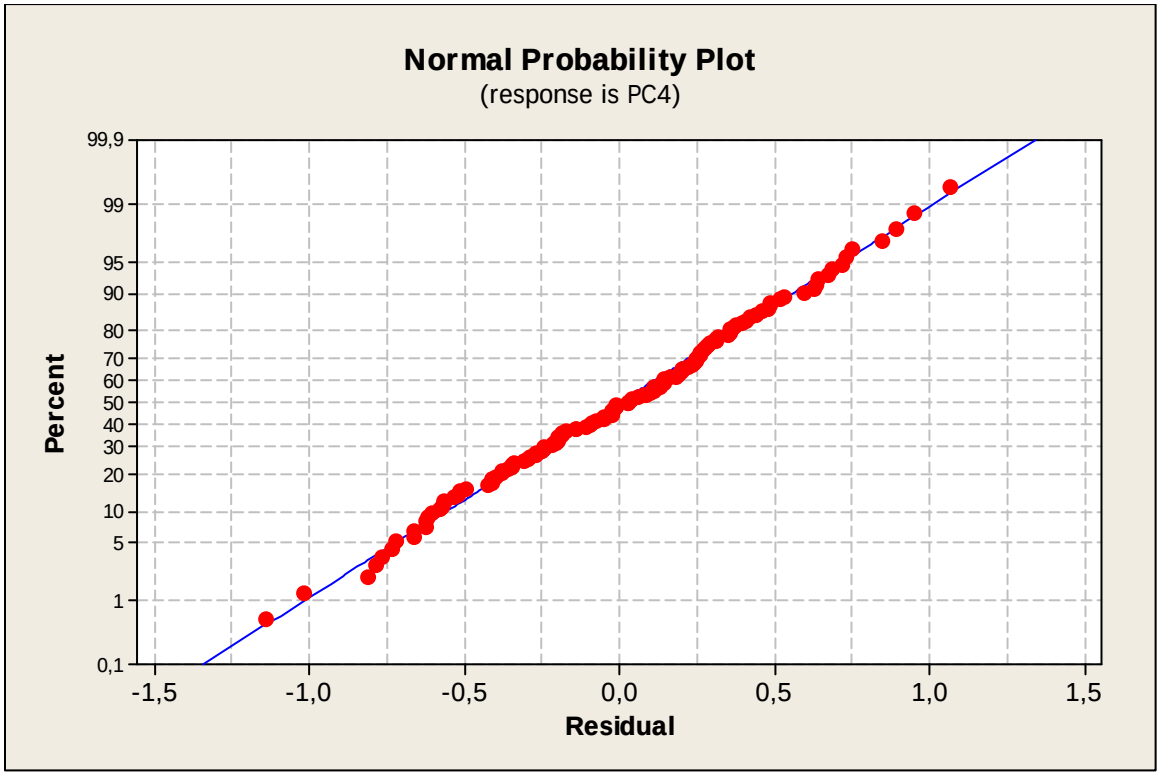
Source	DF	SS	MS	F	P
Regression	3	119,306	39,769	206,18	0,000
Residual Error	128	24,689	0,193		
Total	131	143,995			

Source	DF	Seq SS
MN - GINNI	1	19,468
MN - International Cooperati_1	1	76,820
MN Efficiency of legal framewor	1	23,018

Unusual Observations

Obs	MN - GINNI	PC4	Fit	SE Fit	Residual	St Resid
23	18	-0,5298	-1,4782	0,0971	0,9484	2,21R
29	38	-0,3939	0,7464	0,0858	-1,1404	-2,65R
38	56	0,3002	1,3177	0,0790	-1,0175	-2,36R
73	51	0,8438	-0,2221	0,0442	1,0659	2,44R
109	45	0,2373	-0,6548	0,0877	0,8922	2,07R

R denotes an observation with a large standardized residual.



b. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo aplicar a análise de componentes principais para construção de novas variáveis com maior força explicativa explicativa e menor correlação entre estas, e parcimônia quanto à quantidade final de variáveis.

O estudo foi iniciado com dezessete variáveis e chega-se à conclusão que é possível construir cerca de 84% das explicações a partir de quatro novas variáveis, com a aplicação da técnica de componentes principais (PC's):
PC1 - INDICADORES DE PROGRESSO E DESENVLVIMENTO HUMANO
PC2 - RESOLUTIVIDADE LEGAL, RECURSOS SANITÁRIOS E DEMOCRACIA.
PC3 - FELICIDADE, COOPERAÇÃO E EQUIDADE
PC4 - COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL

8. AGRUPAMENTO DE OBSERVAÇÕES

Este trabalho tem como objetivo realizar uma análise Dos agrupamentos de observações das principais variáveis que compõe as novas quarto Componentes Principais para o EIXO PESSOAS para 132 países, bem como das observações para os quartro componentes principais. Seis das variáveis originais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Todas as variáveis originais foram normalizadas e as observações faltantes foram substituídas pela média das observações não faltantes. Empregou-se o Minitab ® para as análises estatísticas.

a. NOVAS VARIÁVEIS EXPLICATIVAS DO EIXO PESSOAS - COMPONENTES PRINCIPAIS

Quatro componentes principais com correlação igual a zero foram gerados, com força explicativa de 83%.

Os novos componentes principais são formados por combinações de variáveis e são apresentados na lista a seguir, com suas equações explicativas baseadas nas principais variáveis constituintes, de modo a capturar também o peso relativo destas.

PC1 - INDICADORES DE PROGRESSO E DESENVLVIMENTO HUMANO.

```
PC1 = - 8,02 + 0,0333 MN -  (HDI)_1 + 0,0639 MN - Social Progress Index_1
      + 0,0138 MN-Access to improved sanitatio
      + 0,0274 MN The Networked Readiness Inde
```

PC2 - RESOLUTIVIDADE LEGAL, RECURSOS SANITÁRIOS E DEMOCRACIA.

```
The regression equation is
PC2 = - 1,09 - 0,0274 MN-Access to improved sanitatio
      + 0,0280 MN - Democracy index_1 + 0,0323 MN Efficiency of legal framewor
```

PC3 - FELICIDADE, COOPERAÇÃO E EQUIDADE.

```
The regression equation is
PC3 = - 1,72 - 0,0251 MN - GINNI + 0,0323 MN - Happy Planet Index_1
      + 0,0294 MN -  International Cooperati_1
```

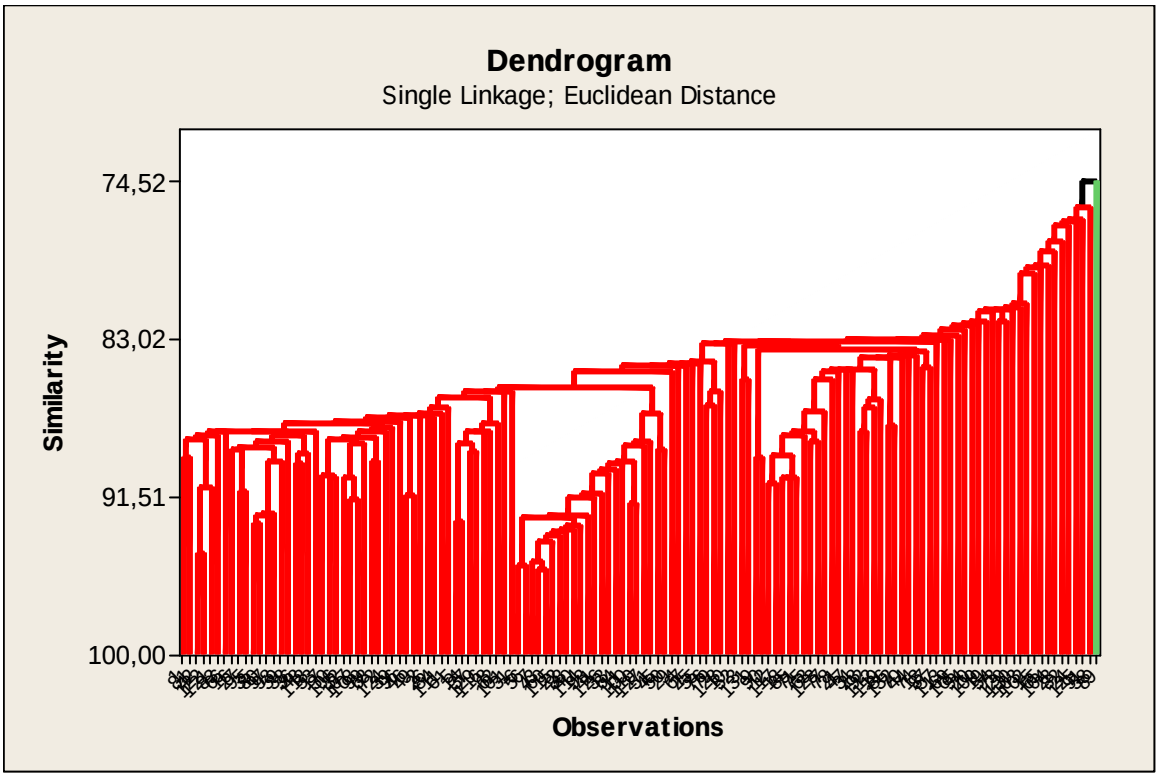
PC4 - COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL.

```
The regression equation is
PC4 = - 2,60 + 0,0233 MN - GINNI + 0,0389 MN -  International Cooperati_1
      - 0,0219 MN Efficiency of legal framewor
```

b. CLUSTERS DE OBSERVAÇÕES COM AS VARIÁVEIS CHAVE DOS COMPONENTES PRINCIPAIS

i. EXERCICIO COM 2 CLUSTERS

Um primeiro exercício com 2 clusters resultou em um com 131 observações e outro com uma única



———— 22/04/2016 17:30:15 ————————————————————

```
Welcome to Minitab, press F1 for help.
```

Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

```
Euclidean Distance, Single Linkage
Amalgamation Steps
```

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,4104	10,3210	43	60	43	2
2	130	95,2030	10,7872	6	56	6	2
3	129	94,9985	11,2472	7	43	7	3
4	128	94,5477	12,2609	42	122	42	2
5	127	93,9876	13,5205	7	115	7	4
6	126	93,6273	14,3306	7	22	7	5
7	125	93,4591	14,7091	7	88	7	6
8	124	93,2575	15,1624	7	40	7	7
9	123	93,1207	15,4699	7	114	7	8
10	122	93,0229	15,6899	35	80	35	2

11	121	92,8267	16,1311	5	14	5	2
12	120	92,6930	16,4317	6	7	6	10
13	119	92,5879	16,6680	35	97	35	3
14	118	92,5573	16,7370	6	93	6	11
15	117	92,4137	17,0598	35	98	35	4
16	116	91,8699	18,2827	100	110	100	2
17	115	91,7065	18,6500	107	108	107	2
18	114	91,5814	18,9314	6	41	6	12
19	113	91,4871	19,1435	46	49	46	2
20	112	91,3518	19,4478	6	126	6	13
21	111	91,2835	19,6015	27	95	27	2
22	110	91,0121	20,2117	42	121	42	3
23	109	90,8012	20,6860	12	117	12	2
24	108	90,5203	21,3176	18	85	18	2
25	107	90,5176	21,3237	32	107	32	3
26	106	90,5162	21,3269	18	71	18	3
27	105	90,5162	21,3269	30	106	30	2
28	104	90,3843	21,6235	17	30	17	3
29	103	90,2370	21,9548	6	33	6	14
30	102	90,1125	22,2347	6	64	6	15
31	101	89,7872	22,9663	36	48	36	2
32	100	89,7231	23,1104	6	11	6	16
33	99	89,6777	23,2124	6	100	6	18
34	98	89,6317	23,3158	68	72	68	2
35	97	89,5926	23,4038	35	59	35	5
36	96	89,4159	23,8011	1	81	1	2
37	95	89,3774	23,8877	9	90	9	2
38	94	89,2939	24,0755	12	18	12	5
39	93	89,1940	24,3003	36	118	36	3
40	92	89,0780	24,5610	37	129	37	2
41	91	88,9845	24,7713	16	27	16	3
42	90	88,9598	24,8269	45	50	45	2
43	89	88,8981	24,9657	16	35	16	8
44	88	88,7632	25,2689	6	127	6	19
45	87	88,6433	25,5386	32	99	32	4
46	86	88,6299	25,5686	5	37	5	4
47	85	88,5059	25,8476	63	123	63	2
48	84	88,5017	25,8570	16	38	16	9
49	83	88,5002	25,8604	6	51	6	20
50	82	88,4686	25,9313	17	32	17	7
51	81	88,3994	26,0871	1	42	1	5
52	80	88,2808	26,3536	17	68	17	9
53	79	88,2282	26,4720	16	36	16	12
54	78	88,1916	26,5544	1	83	1	6
55	77	88,1882	26,5620	12	75	12	6
56	76	88,0545	26,8626	1	66	1	7
57	75	88,0494	26,8741	17	124	17	10
58	74	88,0466	26,8804	5	116	5	5
59	73	88,0447	26,8846	12	63	12	8
60	72	88,0429	26,8887	1	16	1	19
61	71	88,0312	26,9151	28	92	28	2
62	70	88,0214	26,9371	1	53	1	20
63	69	88,0161	26,9489	5	8	5	6
64	68	87,7844	27,4701	17	39	17	11
65	67	87,7140	27,6283	86	132	86	2
66	66	87,6195	27,8408	5	62	5	7
67	65	87,5316	28,0385	1	17	1	31
68	64	87,4656	28,1870	1	96	1	32
69	63	87,2416	28,6906	26	84	26	2
70	62	87,2256	28,7266	1	46	1	34
71	61	87,1889	28,8092	1	2	1	35
72	60	87,1847	28,8187	1	26	1	37
73	59	86,9966	29,2416	1	111	1	38
74	58	86,9898	29,2569	6	45	6	22
75	57	86,9582	29,3279	12	77	12	9
76	56	86,7624	29,7683	28	112	28	3
77	55	86,7269	29,8481	1	61	1	39
78	54	86,6239	30,0796	76	79	76	2
79	53	86,3086	30,7887	28	119	28	4
80	52	86,2113	31,0077	1	5	1	46

81	51	85,8925	31,7244	76	128	76	3
82	50	85,8587	31,8004	1	101	1	47
83	49	85,8355	31,8526	1	31	1	48
84	48	85,5813	32,4243	1	6	1	70
85	47	85,2318	33,2103	3	131	3	2
86	46	85,1629	33,3653	12	78	12	10
87	45	84,7660	34,2577	12	21	12	11
88	44	84,7620	34,2667	1	29	1	71
89	43	84,6929	34,4221	12	47	12	12
90	42	84,6755	34,4613	12	34	12	13
91	41	84,6350	34,5523	12	28	12	17
92	40	84,5648	34,7100	13	87	13	2
93	39	84,4891	34,8804	1	4	1	72
94	38	84,3966	35,0883	1	67	1	73
95	37	84,3742	35,1388	1	25	1	74
96	36	84,3512	35,1904	1	55	1	75
97	35	84,2521	35,4133	1	76	1	78
98	34	84,0624	35,8399	12	86	12	19
99	33	83,9897	36,0034	12	70	12	20
100	32	83,8869	36,2346	12	44	12	21
101	31	83,8685	36,2759	12	74	12	22
102	30	83,6794	36,7011	12	13	12	24
103	29	83,6107	36,8558	9	12	9	26
104	28	83,3267	37,4943	3	9	3	28
105	27	83,2795	37,6004	1	10	1	79
106	26	83,2140	37,7478	1	52	1	80
107	25	83,1261	37,9455	1	3	1	108
108	24	83,1252	37,9476	1	57	1	109
109	23	83,0135	38,1987	1	113	1	110
110	22	82,9869	38,2584	1	69	1	111
111	21	82,8020	38,6742	1	105	1	112
112	20	82,5319	39,2816	1	94	1	113
113	19	82,3011	39,8006	1	20	1	114
114	18	82,2175	39,9886	1	109	1	115
115	17	82,1155	40,2180	73	120	73	2
116	16	82,0648	40,3321	1	19	1	116
117	15	81,4779	41,6519	1	54	1	117
118	14	81,4462	41,7231	1	73	1	119
119	13	81,3975	41,8326	1	130	1	120
120	12	81,3532	41,9323	1	103	1	121
121	11	81,0430	42,6299	1	102	1	122
122	10	79,5173	46,0608	1	15	1	123
123	9	79,1066	46,9843	1	65	1	124
124	8	78,9918	47,2424	1	104	1	125
125	7	78,2839	48,8344	1	23	1	126
126	6	77,7567	50,0199	1	82	1	127
127	5	76,8529	52,0523	1	24	1	128
128	4	76,6828	52,4350	1	125	1	129
129	3	76,5795	52,6671	1	91	1	130
130	2	75,9165	54,1582	1	58	1	131
131	1	74,5227	57,2924	1	89	1	132

Final Partition
Number of clusters: 2

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	131	630080	66,1739	128,717
Cluster2	1	0	0,0000	0,000

Cluster Centroids

Variable	Cluster1	Cluster2	Grand centroid
MN - Happy Planet Index_1	48,5555	69,8903	48,7172

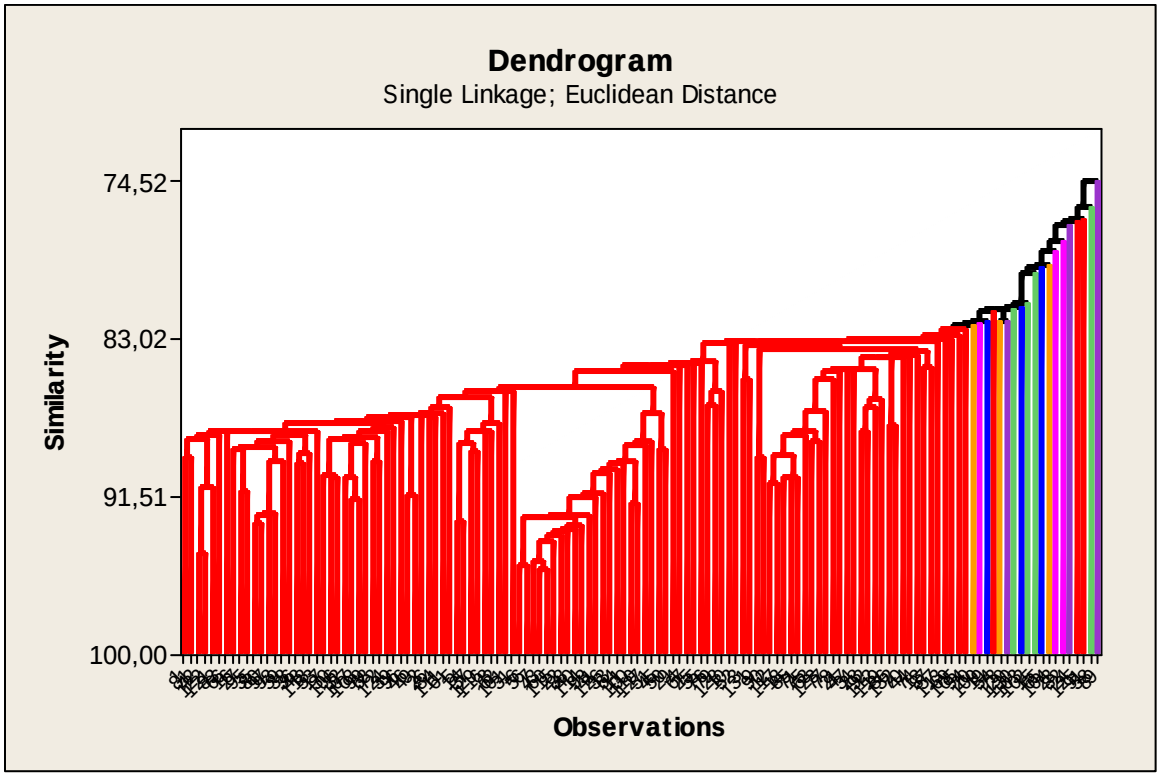
MN - GINNI	61,4415	70,3810	61,5092
MN - Social Progress Index_1	58,4223	97,7520	58,7203
MN-Access to improved sanitatio	71,0241	45,0180	70,8271
MN - International Cooperati_1	57,4932	45,0180	57,3987
MN - (HDI)_1	59,7456	94,8995	60,0120
MN - Democracy index_1	51,9708	92,0616	52,2745
MN The Networked Readiness Inde	47,0118	79,8429	47,2605
MN Efficiency of legal framewor	48,2669	96,6171	48,6332

Distances Between Cluster Centroids

	Cluster1	Cluster2
Cluster1	0,0000	95,7735
Cluster2	95,7735	0,0000

ii. EXERCICIO COM 20 CLUSTERS

Uma segunda tentativa é realizar o agrupamento em vintes clusters com as observações das variáveis relevantes dos componentes principais. Resulta em um cluster com 113 observações e outros dezenove com uma observação cada.



Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,4104	10,3210	43	60	43	2
2	130	95,2030	10,7872	6	56	6	2
3	129	94,9985	11,2472	7	43	7	3
4	128	94,5477	12,2609	42	122	42	2

5	127	93,9876	13,5205	7	115	7	4
6	126	93,6273	14,3306	7	22	7	5
7	125	93,4591	14,7091	7	88	7	6
8	124	93,2575	15,1624	7	40	7	7
9	123	93,1207	15,4699	7	114	7	8
10	122	93,0229	15,6899	35	80	35	2
11	121	92,8267	16,1311	5	14	5	2
12	120	92,6930	16,4317	6	7	6	10
13	119	92,5879	16,6680	35	97	35	3
14	118	92,5573	16,7370	6	93	6	11
15	117	92,4137	17,0598	35	98	35	4
16	116	91,8699	18,2827	100	110	100	2
17	115	91,7065	18,6500	107	108	107	2
18	114	91,5814	18,9314	6	41	6	12
19	113	91,4871	19,1435	46	49	46	2
20	112	91,3518	19,4478	6	126	6	13
21	111	91,2835	19,6015	27	95	27	2
22	110	91,0121	20,2117	42	121	42	3
23	109	90,8012	20,6860	12	117	12	2
24	108	90,5203	21,3176	18	85	18	2
25	107	90,5176	21,3237	32	107	32	3
26	106	90,5162	21,3269	18	71	18	3
27	105	90,5162	21,3269	30	106	30	2
28	104	90,3843	21,6235	17	30	17	3
29	103	90,2370	21,9548	6	33	6	14
30	102	90,1125	22,2347	6	64	6	15
31	101	89,7872	22,9663	36	48	36	2
32	100	89,7231	23,1104	6	11	6	16
33	99	89,6777	23,2124	6	100	6	18
34	98	89,6317	23,3158	68	72	68	2
35	97	89,5926	23,4038	35	59	35	5
36	96	89,4159	23,8011	1	81	1	2
37	95	89,3774	23,8877	9	90	9	2
38	94	89,2939	24,0755	12	18	12	5
39	93	89,1940	24,3003	36	118	36	3
40	92	89,0780	24,5610	37	129	37	2
41	91	88,9845	24,7713	16	27	16	3
42	90	88,9598	24,8269	45	50	45	2
43	89	88,8981	24,9657	16	35	16	8
44	88	88,7632	25,2689	6	127	6	19
45	87	88,6433	25,5386	32	99	32	4
46	86	88,6299	25,5686	5	37	5	4
47	85	88,5059	25,8476	63	123	63	2
48	84	88,5017	25,8570	16	38	16	9
49	83	88,5002	25,8604	6	51	6	20
50	82	88,4686	25,9313	17	32	17	7
51	81	88,3994	26,0871	1	42	1	5
52	80	88,2808	26,3536	17	68	17	9
53	79	88,2282	26,4720	16	36	16	12
54	78	88,1916	26,5544	1	83	1	6
55	77	88,1882	26,5620	12	75	12	6
56	76	88,0545	26,8626	1	66	1	7
57	75	88,0494	26,8741	17	124	17	10
58	74	88,0466	26,8804	5	116	5	5
59	73	88,0447	26,8846	12	63	12	8
60	72	88,0429	26,8887	1	16	1	19
61	71	88,0312	26,9151	28	92	28	2
62	70	88,0214	26,9371	1	53	1	20
63	69	88,0161	26,9489	5	8	5	6
64	68	87,7844	27,4701	17	39	17	11
65	67	87,7140	27,6283	86	132	86	2
66	66	87,6195	27,8408	5	62	5	7
67	65	87,5316	28,0385	1	17	1	31
68	64	87,4656	28,1870	1	96	1	32
69	63	87,2416	28,6906	26	84	26	2
70	62	87,2256	28,7266	1	46	1	34
71	61	87,1889	28,8092	1	2	1	35
72	60	87,1847	28,8187	1	26	1	37
73	59	86,9966	29,2416	1	111	1	38
74	58	86,9898	29,2569	6	45	6	22

75	57	86,9582	29,3279	12	77	12	9
76	56	86,7624	29,7683	28	112	28	3
77	55	86,7269	29,8481	1	61	1	39
78	54	86,6239	30,0796	76	79	76	2
79	53	86,3086	30,7887	28	119	28	4
80	52	86,2113	31,0077	1	5	1	46
81	51	85,8925	31,7244	76	128	76	3
82	50	85,8587	31,8004	1	101	1	47
83	49	85,8355	31,8526	1	31	1	48
84	48	85,5813	32,4243	1	6	1	70
85	47	85,2318	33,2103	3	131	3	2
86	46	85,1629	33,3653	12	78	12	10
87	45	84,7660	34,2577	12	21	12	11
88	44	84,7620	34,2667	1	29	1	71
89	43	84,6929	34,4221	12	47	12	12
90	42	84,6755	34,4613	12	34	12	13
91	41	84,6350	34,5523	12	28	12	17
92	40	84,5648	34,7100	13	87	13	2
93	39	84,4891	34,8804	1	4	1	72
94	38	84,3966	35,0883	1	67	1	73
95	37	84,3742	35,1388	1	25	1	74
96	36	84,3512	35,1904	1	55	1	75
97	35	84,2521	35,4133	1	76	1	78
98	34	84,0624	35,8399	12	86	12	19
99	33	83,9897	36,0034	12	70	12	20
100	32	83,8869	36,2346	12	44	12	21
101	31	83,8685	36,2759	12	74	12	22
102	30	83,6794	36,7011	12	13	12	24
103	29	83,6107	36,8558	9	12	9	26
104	28	83,3267	37,4943	3	9	3	28
105	27	83,2795	37,6004	1	10	1	79
106	26	83,2140	37,7478	1	52	1	80
107	25	83,1261	37,9455	1	3	1	108
108	24	83,1252	37,9476	1	57	1	109
109	23	83,0135	38,1987	1	113	1	110
110	22	82,9869	38,2584	1	69	1	111
111	21	82,8020	38,6742	1	105	1	112
112	20	82,5319	39,2816	1	94	1	113
113	19	82,3011	39,8006	1	20	1	114
114	18	82,2175	39,9886	1	109	1	115
115	17	82,1155	40,2180	73	120	73	2
116	16	82,0648	40,3321	1	19	1	116
117	15	81,4779	41,6519	1	54	1	117
118	14	81,4462	41,7231	1	73	1	119
119	13	81,3975	41,8326	1	130	1	120
120	12	81,3532	41,9323	1	103	1	121
121	11	81,0430	42,6299	1	102	1	122
122	10	79,5173	46,0608	1	15	1	123
123	9	79,1066	46,9843	1	65	1	124
124	8	78,9918	47,2424	1	104	1	125
125	7	78,2839	48,8344	1	23	1	126
126	6	77,7567	50,0199	1	82	1	127
127	5	76,8529	52,0523	1	24	1	128
128	4	76,6828	52,4350	1	125	1	129
129	3	76,5795	52,6671	1	91	1	130
130	2	75,9165	54,1582	1	58	1	131
131	1	74,5227	57,2924	1	89	1	132

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	113	528006	64,9277	117,491
Cluster2	1	0	0,0000	0,000
Cluster3	1	0	0,0000	0,000

Cluster4	1	0	0,0000	0,000
Cluster5	1	0	0,0000	0,000
Cluster6	1	0	0,0000	0,000
Cluster7	1	0	0,0000	0,000
Cluster8	1	0	0,0000	0,000
Cluster9	1	0	0,0000	0,000
Cluster10	1	0	0,0000	0,000
Cluster11	1	0	0,0000	0,000
Cluster12	1	0	0,0000	0,000
Cluster13	1	0	0,0000	0,000
Cluster14	1	0	0,0000	0,000
Cluster15	1	0	0,0000	0,000
Cluster16	1	0	0,0000	0,000
Cluster17	1	0	0,0000	0,000
Cluster18	1	0	0,0000	0,000
Cluster19	1	0	0,0000	0,000
Cluster20	1	0	0,0000	0,000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
MN - Happy Planet Index_1	51,6595	45,0180	19,1190	42,7841
MN - GINNI	62,7480	6,9937	77,9489	81,7328
MN - Social Progress Index_1	59,7618	59,3607	45,0180	39,5855
MN-Access to improved sanitatio	72,1677	60,7187	42,2584	30,5089
MN - International Cooperati_1	59,0058	86,9565	47,8261	34,7826
MN - (HDI)_1	60,8024	58,7057	8,6756	34,6702
MN - Democracy index_1	54,2277	75,5924	21,8009	38,9810
MN The Networked Readiness Inde	47,6548	31,6754	2,3560	29,8429
MN Efficiency of legal framewor	48,3266	64,2376	30,5073	33,7166

Variable	Cluster5	Cluster6	Cluster7	Cluster8
MN - Happy Planet Index_1	6,4309	5,0446	46,0906	57,3312
MN - GINNI	17,8497	51,7745	64,8747	72,0772
MN - Social Progress Index_1	45,0180	3,0734	44,6084	80,7165
MN-Access to improved sanitatio	13,7175	3,1706	88,3054	45,0180
MN - International Cooperati_1	43,4783	30,4348	13,0435	45,0180
MN - (HDI)_1	0,3151	7,3287	70,0672	88,1042
MN - Democracy index_1	45,0180	0,1185	5,8057	75,3555
MN The Networked Readiness Inde	45,0180	45,0180	31,4136	51,3089
MN Efficiency of legal framewor	45,0180	21,0348	40,3902	10,6805

Variable	Cluster9	Cluster10	Cluster11	Cluster12
MN - Happy Planet Index_1	10,907	13,7121	10,0743	69,8903
MN - GINNI	45,018	51,0960	69,4676	70,3810
MN - Social Progress Index_1	66,333	63,8743	52,8627	97,7520
MN-Access to improved sanitatio	100,000	90,4920	51,8255	45,0180
MN - International Cooperati_1	56,522	60,8696	73,9130	45,0180
MN - (HDI)_1	78,576	45,0180	63,5245	94,8995
MN - Democracy index_1	27,133	56,3981	60,7820	92,0616
MN The Networked Readiness Inde	45,550	51,5707	48,4293	79,8429
MN Efficiency of legal framewor	48,518	53,8565	36,7148	96,6171

Variable	Cluster13	Cluster14	Cluster15	Cluster16
MN - Happy Planet Index_1	10,2355	28,7783	34,4141	56,398
MN - GINNI	83,4551	61,1952	32,1503	45,018
MN - Social Progress Index_1	16,0520	56,5859	35,4408	57,692
MN-Access to improved sanitatio	45,0180	67,5199	60,1642	100,000
MN - International Cooperati_1	78,2609	4,3478	60,8696	34,783
MN - (HDI)_1	45,0180	75,4867	22,6630	82,102
MN - Democracy index_1	29,9763	22,5118	20,8531	3,910
MN The Networked Readiness Inde	28,5340	54,4503	40,8377	67,016
MN Efficiency of legal framewor	45,0180	36,6771	81,1346	63,164

Variable	Cluster17	Cluster18	Cluster19	Cluster20
MN - Happy Planet Index_1	13,5094	18,5213	22,1671	82,7132
MN - GINNI	45,0180	59,6816	45,0180	47,9384
MN - Social Progress Index_1	60,0983	45,0180	72,6554	56,2522
MN-Access to improved sanitatio	71,8916	91,3418	97,2920	89,9656
MN - International Cooperati_1	69,5652	45,0180	73,9130	4,3478
MN - (HDI)_1	53,3265	71,1254	81,7992	69,5067
MN - Democracy index_1	75,0000	65,1659	13,6256	42,4171
MN The Networked Readiness Inde	46,0733	45,8115	78,0105	30,6283
MN Efficiency of legal framewor	82,5392	42,8360	80,9886	45,0180

Variable	Grand centroid
MN - Happy Planet Index_1	48,7172
MN - GINNI	61,5092
MN - Social Progress Index_1	58,7203
MN-Access to improved sanitatio	70,8271
MN - International Cooperati_1	57,3987
MN - (HDI)_1	60,0120
MN - Democracy index_1	52,2745
MN The Networked Readiness Inde	47,2605
MN Efficiency of legal framewor	48,6332

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,000	70,954	93,086	67,931	108,156	132,751	73,453
Cluster2	70,954	0,000	123,195	111,668	103,941	153,813	123,940
Cluster3	93,086	123,195	0,000	51,466	85,123	82,792	95,893
Cluster4	67,931	111,668	51,466	0,000	85,816	84,084	80,833
Cluster5	108,156	103,941	85,123	85,816	0,000	76,380	129,937
Cluster6	132,751	153,813	82,792	84,084	76,380	0,000	125,140
Cluster7	73,453	123,940	95,893	80,833	129,937	125,140	0,000
Cluster8	64,136	104,724	121,454	86,496	132,793	153,201	104,015
Cluster9	62,077	91,160	110,093	104,475	123,631	143,874	70,174
Cluster10	47,241	74,644	95,328	89,109	99,798	134,282	87,614
Cluster11	51,419	81,831	87,616	71,025	97,171	119,040	98,947
Cluster12	93,428	114,305	167,861	132,151	159,950	196,282	140,999
Cluster13	79,856	109,093	64,542	63,519	99,214	91,632	99,070
Cluster14	70,346	120,238	101,254	75,859	115,125	117,548	42,630
Cluster15	75,260	81,629	84,584	83,353	77,099	103,752	92,699
Cluster16	72,685	113,045	130,759	111,440	141,803	151,411	56,707
Cluster17	59,513	58,635	110,165	99,200	100,988	145,361	110,542
Cluster18	46,346	85,447	102,820	85,221	115,540	136,994	74,544
Cluster19	79,256	102,241	138,754	127,580	140,243	162,299	97,762
Cluster20	70,826	111,240	119,886	93,964	138,553	153,384	56,610

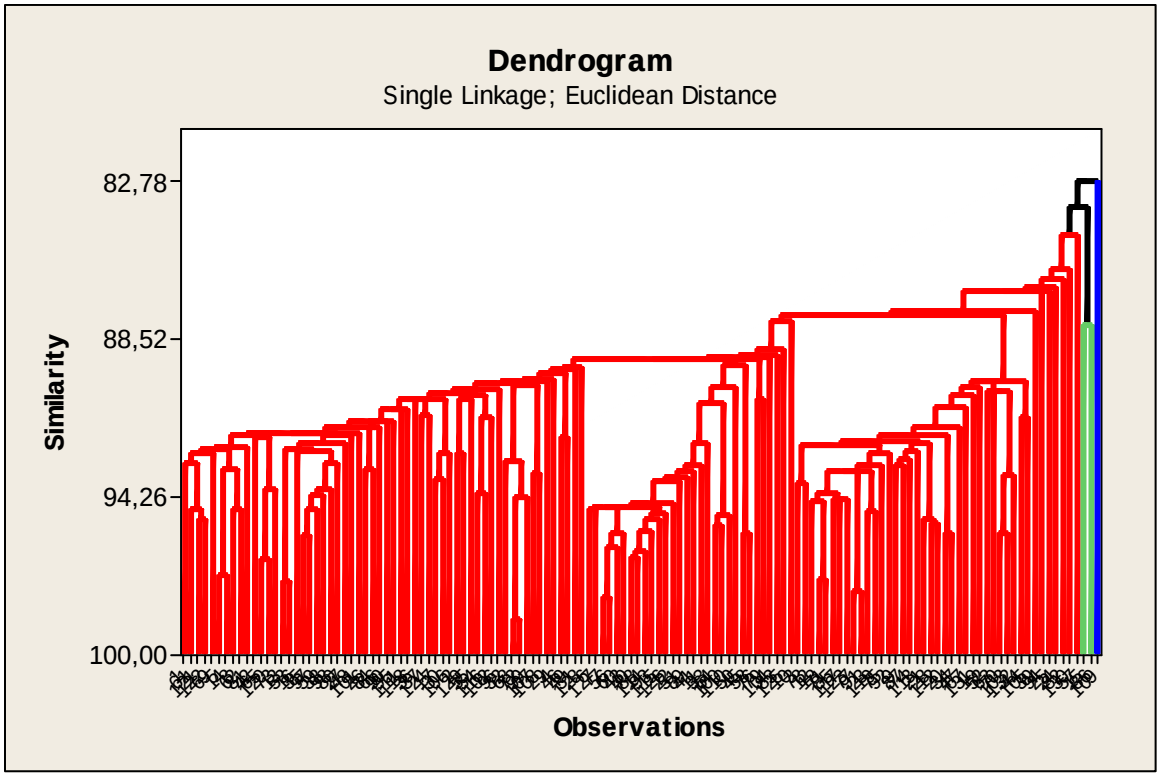
	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	64,136	62,077	47,241	51,419	93,428	79,856
Cluster2	104,724	91,160	74,644	81,831	114,305	109,093
Cluster3	121,454	110,093	95,328	87,616	167,861	64,542
Cluster4	86,496	104,475	89,109	71,025	132,151	63,519
Cluster5	132,793	123,631	99,798	97,171	159,950	99,214
Cluster6	153,201	143,874	134,282	119,040	196,282	91,632
Cluster7	104,015	70,174	87,614	98,947	140,999	99,070
Cluster8	0,000	100,651	95,004	73,483	94,738	115,098
Cluster9	100,651	0,000	46,984	70,069	127,423	94,539
Cluster10	95,004	46,984	0,000	52,940	115,683	83,806
Cluster11	73,483	70,069	52,940	0,000	114,075	58,071
Cluster12	94,738	127,423	115,683	114,075	0,000	152,143
Cluster13	115,098	94,539	83,806	58,071	152,143	0,000
Cluster14	85,520	68,582	80,527	84,627	122,806	100,882
Cluster15	130,293	86,776	71,681	88,447	138,324	78,507
Cluster16	111,950	61,998	84,402	106,036	121,557	118,830
Cluster17	103,938	71,361	41,723	59,119	99,612	89,653
Cluster18	80,848	49,689	40,218	52,406	114,101	83,349
Cluster19	122,913	52,701	70,848	92,714	117,002	113,742
Cluster20	92,398	92,951	96,623	113,156	119,409	127,956

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	70,346	75,260	72,685	59,513	46,346	79,256
Cluster2	120,238	81,629	113,045	58,635	85,447	102,241
Cluster3	101,254	84,584	130,759	110,165	102,820	138,754
Cluster4	75,859	83,353	111,440	99,200	85,221	127,580
Cluster5	115,125	77,099	141,803	100,988	115,540	140,243
Cluster6	117,548	103,752	151,411	145,361	136,994	162,299
Cluster7	42,630	92,699	56,707	110,542	74,544	97,762
Cluster8	85,520	130,293	111,950	103,938	80,848	122,913
Cluster9	68,582	86,776	61,998	71,361	49,689	52,701
Cluster10	80,527	71,681	84,402	41,723	40,218	70,848
Cluster11	84,627	88,447	106,036	59,119	52,406	92,714
Cluster12	122,806	138,324	121,557	99,612	114,101	117,002
Cluster13	100,882	78,507	118,830	89,653	83,349	113,742
Cluster14	0,000	97,602	65,240	100,991	66,436	94,481
Cluster15	97,602	0,000	90,773	72,974	90,163	90,445
Cluster16	65,240	90,773	0,000	102,669	82,010	58,884
Cluster17	100,991	72,974	102,669	0,000	58,638	80,602
Cluster18	66,436	90,163	82,010	58,638	0,000	84,411
Cluster19	94,481	90,445	58,884	80,602	84,411	0,000
Cluster20	68,247	105,669	70,886	111,186	82,438	115,566

	Cluster20
Cluster1	70,826
Cluster2	111,240
Cluster3	119,886
Cluster4	93,964
Cluster5	138,553
Cluster6	153,384
Cluster7	56,610
Cluster8	92,398
Cluster9	92,951
Cluster10	96,623
Cluster11	113,156
Cluster12	119,409
Cluster13	127,956
Cluster14	68,247
Cluster15	105,669
Cluster16	70,886
Cluster17	111,186
Cluster18	82,438
Cluster19	115,566
Cluster20	0,000

iii. EXERCICIO COM 3 CLUSTERS – COMPONENTES PRINCIPAIS

Uma terceira tentativa é realizar o agrupamento em vintes clusters com as observações das quatro componentes principais. Agrupa-se em três clusters. Resulta em um cluster com 129 observações e outros com dois e uma observações, respectivamente.



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98,7284	0,15442	68	72	68	2
2	130	97,9739	0,24606	7	56	7	2
3	129	97,7196	0,27694	77	91	77	2
4	128	97,3777	0,31846	13	53	13	2
5	127	97,2935	0,32869	18	71	18	2
6	126	97,1435	0,34690	5	14	5	2
7	125	96,5412	0,42006	65	120	65	2
8	124	96,4753	0,42806	40	88	40	2
9	123	96,3001	0,44934	40	114	40	3
10	122	96,1237	0,47077	7	60	7	3
11	121	95,6640	0,52658	35	97	35	2
12	120	95,6537	0,52784	7	43	7	4
13	119	95,6338	0,53026	45	58	45	2
14	118	95,5928	0,53523	20	34	20	2
15	117	95,5880	0,53582	70	103	70	2
16	116	95,5713	0,53785	40	93	40	4
17	115	95,3031	0,57042	64	100	64	2
18	114	95,2624	0,57536	28	112	28	2
19	113	95,1416	0,59003	121	122	121	2
20	112	95,1385	0,59041	21	28	21	3
21	111	95,0955	0,59563	40	115	40	5
22	110	94,9730	0,61051	64	110	64	3
23	109	94,8993	0,61946	40	126	40	6
24	108	94,8033	0,63112	23	119	23	2
25	107	94,7613	0,63622	61	121	61	3
26	106	94,7487	0,63775	35	59	35	3
27	105	94,7334	0,63960	8	62	8	2
28	104	94,7055	0,64300	6	127	6	2

29	103	94,6829	0,64574	6	7	6	6
30	102	94,6485	0,64992	6	40	6	12
31	101	94,5197	0,66556	6	22	6	13
32	100	94,4610	0,67269	12	18	12	3
33	99	94,3888	0,68146	117	123	117	2
34	98	94,3832	0,68214	85	117	85	3
35	97	94,3041	0,69174	68	99	68	3
36	96	94,2299	0,70075	35	38	35	4
37	95	94,1887	0,70576	12	85	12	6
38	94	94,1636	0,70881	54	116	54	2
39	93	93,9895	0,72995	65	73	65	3
40	92	93,9846	0,73055	35	98	35	5
41	91	93,7818	0,75518	3	78	3	2
42	90	93,7321	0,76121	6	51	6	14
43	89	93,6647	0,76939	17	30	17	2
44	88	93,5557	0,78263	6	89	6	15
45	87	93,4959	0,78989	70	132	70	3
46	86	93,3917	0,80255	107	108	107	2
47	85	93,3907	0,80267	6	41	6	16
48	84	93,3713	0,80503	12	77	12	8
49	83	93,3512	0,80747	12	23	12	10
50	82	93,3107	0,81239	5	8	5	4
51	81	93,2814	0,81594	26	84	26	2
52	80	93,1550	0,83130	6	11	6	17
53	79	93,1189	0,83568	12	75	12	11
54	78	93,1030	0,83761	9	87	9	2
55	77	93,0448	0,84468	35	80	35	6
56	76	93,0297	0,84651	1	61	1	4
57	75	92,9531	0,85581	32	68	32	4
58	74	92,8876	0,86377	9	74	9	3
59	73	92,7272	0,88325	12	92	12	12
60	72	92,7053	0,88592	17	106	17	3
61	71	92,6829	0,88863	2	129	2	2
62	70	92,6707	0,89011	1	69	1	5
63	69	92,6269	0,89543	9	113	9	4
64	68	92,6058	0,89799	16	35	16	7
65	67	92,5589	0,90369	1	5	1	9
66	66	92,5306	0,90712	13	16	13	9
67	65	92,4765	0,91370	1	81	1	10
68	64	92,4353	0,91870	3	12	3	14
69	63	92,3346	0,93093	13	27	13	10
70	62	92,3307	0,93141	6	33	6	18
71	61	92,3260	0,93198	9	21	9	7
72	60	92,2801	0,93754	3	9	3	21
73	59	92,1360	0,95505	46	49	46	2
74	58	92,1268	0,95616	42	65	42	4
75	57	92,0459	0,96599	3	20	3	23
76	56	92,0243	0,96861	13	48	13	11
77	55	92,0236	0,96870	1	42	1	14
78	54	92,0004	0,97152	1	13	1	25
79	53	91,9604	0,97637	1	101	1	26
80	52	91,7611	1,00058	26	66	26	3
81	51	91,7531	1,00155	3	47	3	24
82	50	91,7299	1,00437	1	26	1	29
83	49	91,5310	1,02853	1	90	1	30
84	48	91,5137	1,03062	1	95	1	31
85	47	91,4134	1,04281	44	105	44	2
86	46	91,3838	1,04640	54	102	54	3
87	45	91,3348	1,05235	37	124	37	2
88	44	91,0878	1,08235	1	111	1	32
89	43	91,0621	1,08547	1	118	1	33
90	42	91,0312	1,08923	3	67	3	25
91	41	90,8476	1,11152	6	64	6	21
92	40	90,7529	1,12302	1	36	1	34
93	39	90,7469	1,12375	76	79	76	2
94	38	90,7425	1,12429	1	37	1	36
95	37	90,7329	1,12545	1	17	1	39
96	36	90,6980	1,12969	2	55	2	3
97	35	90,5625	1,14614	3	131	3	26
98	34	90,5622	1,14618	2	54	2	6

99	33	90,5380	1,14912	1	2	1	45
100	32	90,4305	1,16217	63	70	63	4
101	31	90,4184	1,16364	52	63	52	5
102	30	90,3658	1,17003	1	96	1	46
103	29	90,2889	1,17937	3	19	3	27
104	28	90,2858	1,17975	6	50	6	22
105	27	90,2212	1,18760	32	107	32	6
106	26	90,1767	1,19299	1	32	1	52
107	25	90,0975	1,20262	1	29	1	53
108	24	90,0763	1,20520	3	52	3	32
109	23	90,0529	1,20804	3	44	3	34
110	22	89,9789	1,21702	1	4	1	54
111	21	89,7508	1,24473	1	46	1	56
112	20	89,6417	1,25797	1	31	1	57
113	19	89,5871	1,26460	1	128	1	58
114	18	89,4650	1,27943	6	45	6	24
115	17	89,2903	1,30064	1	6	1	82
116	16	89,2677	1,30339	1	76	1	84
117	15	89,1847	1,31348	1	104	1	85
118	14	89,0846	1,32563	1	83	1	86
119	13	88,9243	1,34509	1	125	1	87
120	12	88,0022	1,45708	15	86	15	2
121	11	87,9001	1,46948	1	10	1	88
122	10	87,6400	1,50108	1	3	1	122
123	9	87,4876	1,51958	1	39	1	123
124	8	86,7739	1,60625	1	94	1	124
125	7	86,7562	1,60841	1	25	1	125
126	6	86,6715	1,61869	1	24	1	126
127	5	86,3742	1,65480	1	82	1	127
128	4	86,0054	1,69958	1	130	1	128
129	3	84,7631	1,85045	1	57	1	129
130	2	83,7169	1,97752	1	15	1	131
131	1	82,7794	2,09136	1	109	1	132

Final Partition
Number of clusters: 3

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	129	1824,30	3,48085	6,44700
Cluster2	2	1,06	0,72854	0,72854
Cluster3	1	0,00	0,00000	0,00000

Cluster Centroids

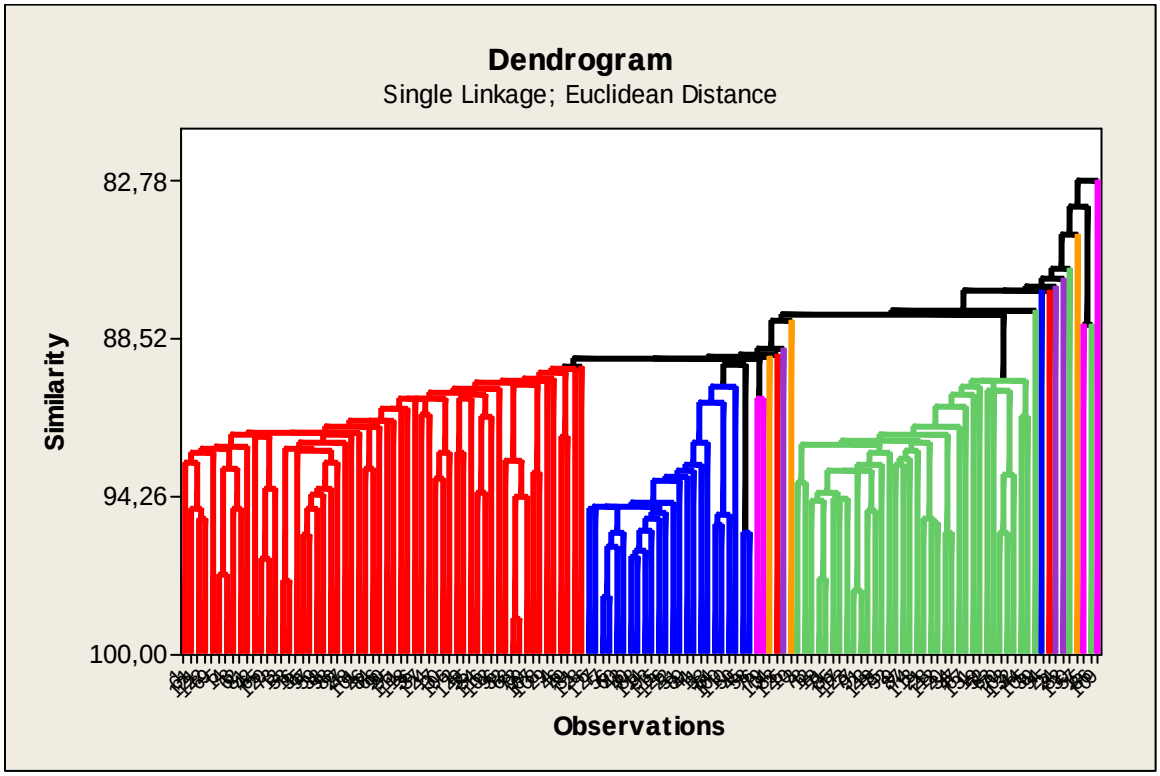
Variable	Cluster1	Cluster2	Cluster3	Grand centroid
PC1	0,0227843	-1,37831	-0,18255	0,0000000
PC2	-0,0538587	2,47471	1,99835	-0,0000000
PC3	-0,0295813	1,95357	-0,09115	-0,0000000
PC4	0,0055255	-0,47507	0,23735	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3
Cluster1	0,00000	3,53845	2,07636
Cluster2	3,53845	0,00000	2,51897
Cluster3	2,07636	2,51897	0,00000

iv. EXERCICIO COM 20 CLUSTERS – COMPONENTES PRINCIPAIS

Uma nova tentativa é realizar o agrupamento em vintes clusters com as observações das quatro componentes principais. Agrupa-se em três clusters. Resulta um cluster com 57 observações, outro com 34, um terceiro com 22 observações, e os demais com uma ou duas observações (países).



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98,7284	0,15442	68	72	68	2
2	130	97,9739	0,24606	7	56	7	2
3	129	97,7196	0,27694	77	91	77	2
4	128	97,3777	0,31846	13	53	13	2
5	127	97,2935	0,32869	18	71	18	2
6	126	97,1435	0,34690	5	14	5	2
7	125	96,5412	0,42006	65	120	65	2
8	124	96,4753	0,42806	40	88	40	2
9	123	96,3001	0,44934	40	114	40	3
10	122	96,1237	0,47077	7	60	7	3
11	121	95,6640	0,52658	35	97	35	2
12	120	95,6537	0,52784	7	43	7	4
13	119	95,6338	0,53026	45	58	45	2
14	118	95,5928	0,53523	20	34	20	2
15	117	95,5880	0,53582	70	103	70	2
16	116	95,5713	0,53785	40	93	40	4
17	115	95,3031	0,57042	64	100	64	2

18	114	95,2624	0,57536	28	112	28	2
19	113	95,1416	0,59003	121	122	121	2
20	112	95,1385	0,59041	21	28	21	3
21	111	95,0955	0,59563	40	115	40	5
22	110	94,9730	0,61051	64	110	64	3
23	109	94,8993	0,61946	40	126	40	6
24	108	94,8033	0,63112	23	119	23	2
25	107	94,7613	0,63622	61	121	61	3
26	106	94,7487	0,63775	35	59	35	3
27	105	94,7334	0,63960	8	62	8	2
28	104	94,7055	0,64300	6	127	6	2
29	103	94,6829	0,64574	6	7	6	6
30	102	94,6485	0,64992	6	40	6	12
31	101	94,5197	0,66556	6	22	6	13
32	100	94,4610	0,67269	12	18	12	3
33	99	94,3888	0,68146	117	123	117	2
34	98	94,3832	0,68214	85	117	85	3
35	97	94,3041	0,69174	68	99	68	3
36	96	94,2299	0,70075	35	38	35	4
37	95	94,1887	0,70576	12	85	12	6
38	94	94,1636	0,70881	54	116	54	2
39	93	93,9895	0,72995	65	73	65	3
40	92	93,9846	0,73055	35	98	35	5
41	91	93,7818	0,75518	3	78	3	2
42	90	93,7321	0,76121	6	51	6	14
43	89	93,6647	0,76939	17	30	17	2
44	88	93,5557	0,78263	6	89	6	15
45	87	93,4959	0,78989	70	132	70	3
46	86	93,3917	0,80255	107	108	107	2
47	85	93,3907	0,80267	6	41	6	16
48	84	93,3713	0,80503	12	77	12	8
49	83	93,3512	0,80747	12	23	12	10
50	82	93,3107	0,81239	5	8	5	4
51	81	93,2814	0,81594	26	84	26	2
52	80	93,1550	0,83130	6	11	6	17
53	79	93,1189	0,83568	12	75	12	11
54	78	93,1030	0,83761	9	87	9	2
55	77	93,0448	0,84468	35	80	35	6
56	76	93,0297	0,84651	1	61	1	4
57	75	92,9531	0,85581	32	68	32	4
58	74	92,8876	0,86377	9	74	9	3
59	73	92,7272	0,88325	12	92	12	12
60	72	92,7053	0,88592	17	106	17	3
61	71	92,6829	0,88863	2	129	2	2
62	70	92,6707	0,89011	1	69	1	5
63	69	92,6269	0,89543	9	113	9	4
64	68	92,6058	0,89799	16	35	16	7
65	67	92,5589	0,90369	1	5	1	9
66	66	92,5306	0,90712	13	16	13	9
67	65	92,4765	0,91370	1	81	1	10
68	64	92,4353	0,91870	3	12	3	14
69	63	92,3346	0,93093	13	27	13	10
70	62	92,3307	0,93141	6	33	6	18
71	61	92,3260	0,93198	9	21	9	7
72	60	92,2801	0,93754	3	9	3	21
73	59	92,1360	0,95505	46	49	46	2
74	58	92,1268	0,95616	42	65	42	4
75	57	92,0459	0,96599	3	20	3	23
76	56	92,0243	0,96861	13	48	13	11
77	55	92,0236	0,96870	1	42	1	14
78	54	92,0004	0,97152	1	13	1	25
79	53	91,9604	0,97637	1	101	1	26
80	52	91,7611	1,00058	26	66	26	3
81	51	91,7531	1,00155	3	47	3	24
82	50	91,7299	1,00437	1	26	1	29
83	49	91,5310	1,02853	1	90	1	30
84	48	91,5137	1,03062	1	95	1	31
85	47	91,4134	1,04281	44	105	44	2
86	46	91,3838	1,04640	54	102	54	3
87	45	91,3348	1,05235	37	124	37	2

88	44	91,0878	1,08235	1	111	1	32
89	43	91,0621	1,08547	1	118	1	33
90	42	91,0312	1,08923	3	67	3	25
91	41	90,8476	1,11152	6	64	6	21
92	40	90,7529	1,12302	1	36	1	34
93	39	90,7469	1,12375	76	79	76	2
94	38	90,7425	1,12429	1	37	1	36
95	37	90,7329	1,12545	1	17	1	39
96	36	90,6980	1,12969	2	55	2	3
97	35	90,5625	1,14614	3	131	3	26
98	34	90,5622	1,14618	2	54	2	6
99	33	90,5380	1,14912	1	2	1	45
100	32	90,4305	1,16217	63	70	63	4
101	31	90,4184	1,16364	52	63	52	5
102	30	90,3658	1,17003	1	96	1	46
103	29	90,2889	1,17937	3	19	3	27
104	28	90,2858	1,17975	6	50	6	22
105	27	90,2212	1,18760	32	107	32	6
106	26	90,1767	1,19299	1	32	1	52
107	25	90,0975	1,20262	1	29	1	53
108	24	90,0763	1,20520	3	52	3	32
109	23	90,0529	1,20804	3	44	3	34
110	22	89,9789	1,21702	1	4	1	54
111	21	89,7508	1,24473	1	46	1	56
112	20	89,6417	1,25797	1	31	1	57
113	19	89,5871	1,26460	1	128	1	58
114	18	89,4650	1,27943	6	45	6	24
115	17	89,2903	1,30064	1	6	1	82
116	16	89,2677	1,30339	1	76	1	84
117	15	89,1847	1,31348	1	104	1	85
118	14	89,0846	1,32563	1	83	1	86
119	13	88,9243	1,34509	1	125	1	87
120	12	88,0022	1,45708	15	86	15	2
121	11	87,9001	1,46948	1	10	1	88
122	10	87,6400	1,50108	1	3	1	122
123	9	87,4876	1,51958	1	39	1	123
124	8	86,7739	1,60625	1	94	1	124
125	7	86,7562	1,60841	1	25	1	125
126	6	86,6715	1,61869	1	24	1	126
127	5	86,3742	1,65480	1	82	1	127
128	4	86,0054	1,69958	1	130	1	128
129	3	84,7631	1,85045	1	57	1	129
130	2	83,7169	1,97752	1	15	1	131
131	1	82,7794	2,09136	1	109	1	132

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	57	296,940	2,12581	3,97914
Cluster2	34	82,343	1,49673	2,31974
Cluster3	22	45,475	1,25215	3,23120
Cluster4	1	0,000	0,00000	0,00000
Cluster5	1	0,000	0,00000	0,00000
Cluster6	1	0,000	0,00000	0,00000
Cluster7	1	0,000	0,00000	0,00000
Cluster8	1	0,000	0,00000	0,00000
Cluster9	2	0,141	0,26513	0,26513
Cluster10	1	0,000	0,00000	0,00000
Cluster11	2	0,631	0,56188	0,56188
Cluster12	1	0,000	0,00000	0,00000
Cluster13	1	0,000	0,00000	0,00000
Cluster14	1	0,000	0,00000	0,00000
Cluster15	1	0,000	0,00000	0,00000
Cluster16	1	0,000	0,00000	0,00000

Cluster17	1	0,000	0,00000	0,00000
Cluster18	1	0,000	0,00000	0,00000
Cluster19	1	0,000	0,00000	0,00000
Cluster20	1	0,000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	0,394549	-4,07530	4,76786	0,82789	-0,93497	-5,91577	2,92015
PC2	-0,772298	0,68179	0,74430	-2,73223	2,55667	-0,06017	1,72842
PC3	0,629921	-0,55700	-1,03977	-1,68497	2,39496	-2,12813	2,42128
PC4	0,176004	0,07759	-0,52776	-0,51760	-0,11082	-0,42358	0,13167

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
PC1	3,71464	2,20421	3,45616	2,04598	-1,62872	1,19502
PC2	1,46393	-1,51016	0,02768	1,27229	0,80819	-0,02087
PC3	0,23237	0,16604	0,53731	0,58583	-0,70718	1,15875
PC4	1,83555	0,88215	-1,39443	0,06656	1,75730	0,88940

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
PC1	-1,82165	-4,09166	0,95944	-0,18255	2,24004	2,41745
PC2	2,39276	-1,03798	-1,55029	1,99835	0,36682	1,34274
PC3	1,51218	0,39871	-0,63366	-0,09115	-0,85475	1,16062
PC4	-0,83932	-1,67095	-1,76850	0,23735	0,02357	1,54113

Variable	Cluster20	Grand centroid
PC1	-0,47767	0,0000000
PC2	-2,34275	-0,0000000
PC3	0,68724	-0,0000000
PC4	-2,64797	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,00000	4,84896	4,9708	3,14148	4,00591	6,9494	3,9804
Cluster2	4,84896	0,00000	8,8772	6,10929	4,70387	2,5802	7,6750
Cluster3	4,97082	8,87721	0,0000	5,29395	6,91217	10,7695	4,0983
Cluster4	3,14148	6,10929	5,2940	0,00000	6,92036	7,2679	6,4465
Cluster5	4,00591	4,70387	6,9122	6,92036	0,00000	7,2258	3,9506
Cluster6	6,94936	2,58021	10,7695	7,26788	7,22582	0,0000	10,1133
Cluster7	3,98035	7,67495	4,0983	6,44653	3,95063	10,1133	0,0000
Cluster8	4,35153	8,06278	2,9716	5,92912	5,59269	10,2832	2,8975
Cluster9	2,12912	6,73847	3,8854	2,96196	5,68750	8,6605	4,0805
Cluster10	3,53387	7,77915	2,3393	4,49802	5,54765	9,7922	3,0097
Cluster11	2,63087	6,25499	3,2685	4,79767	3,72023	8,5306	2,0846
Cluster12	3,29847	2,97417	6,8009	4,96998	4,08064	5,0900	5,8283
Cluster13	1,41209	5,64561	4,4936	4,18958	3,70260	7,9431	2,8643
Cluster14	4,09126	3,62334	7,2628	6,60385	1,45708	6,0169	4,9694
Cluster15	4,86429	2,63223	9,2219	5,72227	5,41343	3,4963	8,0099
Cluster16	2,51040	5,80902	4,6339	2,02096	5,69003	7,3165	5,2476
Cluster17	2,92117	4,13877	5,2502	5,14874	2,67952	6,4570	4,0029
Cluster18	2,63266	6,33043	2,6212	3,54688	5,04528	8,2777	3,6139
Cluster19	3,27273	6,90542	3,8736	5,60962	4,11885	9,2783	1,9943
Cluster20	3,34742	5,57384	6,6726	3,46728	5,79370	6,9034	6,2332

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	4,3515	2,12912	3,53387	2,63087	3,29847	1,41209
Cluster2	8,0628	6,73847	7,77915	6,25499	2,97417	5,64561
Cluster3	2,9716	3,88544	2,33931	3,26851	6,80091	4,49360
Cluster4	5,9291	2,96196	4,49802	4,79767	4,96998	4,18958
Cluster5	5,5927	5,68750	5,54765	3,72023	4,08064	3,70260
Cluster6	10,2832	8,66053	9,79224	8,53058	5,09001	7,94307
Cluster7	2,8975	4,08048	3,00968	2,08455	5,82829	2,86431
Cluster8	0,0000	3,46986	3,55744	2,46484	5,46538	3,21038
Cluster9	3,4699	0,00000	3,04187	2,93402	4,64699	2,05474
Cluster10	3,5574	3,04187	0,00000	2,38213	6,16014	3,27371
Cluster11	2,4648	2,93402	2,38213	0,00000	4,27191	1,84437
Cluster12	5,4654	4,64699	6,16014	4,27191	0,00000	3,59107

Cluster13	3,2104	2,05474	3,27371	1,84437	3,59107	0,00000
Cluster14	6,3487	6,01796	5,89130	4,22998	3,77041	4,24727
Cluster15	8,9175	6,81420	7,62895	6,78689	4,73817	6,00970
Cluster16	5,5151	3,03587	3,19917	3,74194	4,96967	3,55974
Cluster17	4,2582	4,29979	4,49237	2,44560	2,48950	2,82177
Cluster18	2,8006	2,30293	2,35425	1,71308	4,26494	2,45890
Cluster19	1,6266	3,09967	3,43721	1,62717	4,49362	1,94385
Cluster20	7,2371	4,54082	4,76317	5,17844	5,71000	4,57431

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	4,09126	4,86429	2,51040	2,92117	2,63266	3,27273
Cluster2	3,62334	2,63223	5,80902	4,13877	6,33043	6,90542
Cluster3	7,26282	9,22191	4,63395	5,25018	2,62117	3,87357
Cluster4	6,60385	5,72227	2,02096	5,14874	3,54688	5,60962
Cluster5	1,45708	5,41343	5,69003	2,67952	5,04528	4,11885
Cluster6	6,01692	3,49633	7,31650	6,45705	8,27774	9,27829
Cluster7	4,96944	8,00991	5,24763	4,00288	3,61394	1,99432
Cluster8	6,34871	8,91746	5,51505	4,25824	2,80058	1,62657
Cluster9	6,01796	6,81420	3,03587	4,29979	2,30293	3,09967
Cluster10	5,89130	7,62895	3,19917	4,49237	2,35425	3,43721
Cluster11	4,22998	6,78689	3,74194	2,44560	1,71308	1,62717
Cluster12	3,77041	4,73817	4,96967	2,48950	4,26494	4,49362
Cluster13	4,24727	6,00970	3,55974	2,82177	2,45890	1,94385
Cluster14	0,00000	4,34215	5,36191	2,56361	5,19121	4,98625
Cluster15	4,34215	0,00000	5,18183	5,32747	6,81956	7,67687
Cluster16	5,36191	5,18183	0,00000	4,26787	2,92841	4,96674
Cluster17	2,56361	5,32747	4,26787	0,00000	3,02648	3,23366
Cluster18	5,19121	6,81956	2,92841	3,02648	0,00000	2,71083
Cluster19	4,98625	7,67687	4,96674	3,23366	2,71083	0,00000
Cluster20	5,30877	3,97507	2,28288	5,27856	4,92367	6,30375

	Cluster20
Cluster1	3,34742
Cluster2	5,57384
Cluster3	6,67257
Cluster4	3,46728
Cluster5	5,79370
Cluster6	6,90343
Cluster7	6,23321
Cluster8	7,23707
Cluster9	4,54082
Cluster10	4,76317
Cluster11	5,17844
Cluster12	5,71000
Cluster13	4,57431
Cluster14	5,30877
Cluster15	3,97507
Cluster16	2,28288
Cluster17	5,27856
Cluster18	4,92367
Cluster19	6,30375
Cluster20	0,00000

v. EXERCICIO COM 30 CLUSTERS – COMPONENTES PRINCIPAIS

Uma nova tentativa é realizada com o agrupamento em trinta clusters com as observações das quatro componentes principais. Resulta um cluster com 46 observações, outro com 26, um terceiro com 21 observações, e os demais com uma a cinco observações (países). Não se verifica benefícios em identificar mais agrupamentos relevantes com 30 clusters, optando-se por desenvolver uma análise com os 20 clusters gerados na tentativa anterior.

Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98,7284	0,15442	68	72	68	2
2	130	97,9739	0,24606	7	56	7	2
3	129	97,7196	0,27694	77	91	77	2
4	128	97,3777	0,31846	13	53	13	2
5	127	97,2935	0,32869	18	71	18	2
6	126	97,1435	0,34690	5	14	5	2
7	125	96,5412	0,42006	65	120	65	2
8	124	96,4753	0,42806	40	88	40	2
9	123	96,3001	0,44934	40	114	40	3
10	122	96,1237	0,47077	7	60	7	3
11	121	95,6640	0,52658	35	97	35	2
12	120	95,6537	0,52784	7	43	7	4
13	119	95,6338	0,53026	45	58	45	2
14	118	95,5928	0,53523	20	34	20	2
15	117	95,5880	0,53582	70	103	70	2
16	116	95,5713	0,53785	40	93	40	4
17	115	95,3031	0,57042	64	100	64	2
18	114	95,2624	0,57536	28	112	28	2
19	113	95,1416	0,59003	121	122	121	2
20	112	95,1385	0,59041	21	28	21	3
21	111	95,0955	0,59563	40	115	40	5
22	110	94,9730	0,61051	64	110	64	3
23	109	94,8993	0,61946	40	126	40	6
24	108	94,8033	0,63112	23	119	23	2
25	107	94,7613	0,63622	61	121	61	3
26	106	94,7487	0,63775	35	59	35	3
27	105	94,7334	0,63960	8	62	8	2
28	104	94,7055	0,64300	6	127	6	2
29	103	94,6829	0,64574	6	7	6	6
30	102	94,6485	0,64992	6	40	6	12
31	101	94,5197	0,66556	6	22	6	13
32	100	94,4610	0,67269	12	18	12	3
33	99	94,3888	0,68146	117	123	117	2
34	98	94,3832	0,68214	85	117	85	3
35	97	94,3041	0,69174	68	99	68	3
36	96	94,2299	0,70075	35	38	35	4
37	95	94,1887	0,70576	12	85	12	6
38	94	94,1636	0,70881	54	116	54	2
39	93	93,9895	0,72995	65	73	65	3
40	92	93,9846	0,73055	35	98	35	5
41	91	93,7818	0,75518	3	78	3	2
42	90	93,7321	0,76121	6	51	6	14
43	89	93,6647	0,76939	17	30	17	2
44	88	93,5557	0,78263	6	89	6	15
45	87	93,4959	0,78989	70	132	70	3
46	86	93,3917	0,80255	107	108	107	2
47	85	93,3907	0,80267	6	41	6	16
48	84	93,3713	0,80503	12	77	12	8
49	83	93,3512	0,80747	12	23	12	10
50	82	93,3107	0,81239	5	8	5	4
51	81	93,2814	0,81594	26	84	26	2
52	80	93,1550	0,83130	6	11	6	17
53	79	93,1189	0,83568	12	75	12	11
54	78	93,1030	0,83761	9	87	9	2
55	77	93,0448	0,84468	35	80	35	6
56	76	93,0297	0,84651	1	61	1	4
57	75	92,9531	0,85581	32	68	32	4
58	74	92,8876	0,86377	9	74	9	3
59	73	92,7272	0,88325	12	92	12	12
60	72	92,7053	0,88592	17	106	17	3
61	71	92,6829	0,88863	2	129	2	2
62	70	92,6707	0,89011	1	69	1	5
63	69	92,6269	0,89543	9	113	9	4

64	68	92,6058	0,89799	16	35	16	7
65	67	92,5589	0,90369	1	5	1	9
66	66	92,5306	0,90712	13	16	13	9
67	65	92,4765	0,91370	1	81	1	10
68	64	92,4353	0,91870	3	12	3	14
69	63	92,3346	0,93093	13	27	13	10
70	62	92,3307	0,93141	6	33	6	18
71	61	92,3260	0,93198	9	21	9	7
72	60	92,2801	0,93754	3	9	3	21
73	59	92,1360	0,95505	46	49	46	2
74	58	92,1268	0,95616	42	65	42	4
75	57	92,0459	0,96599	3	20	3	23
76	56	92,0243	0,96861	13	48	13	11
77	55	92,0236	0,96870	1	42	1	14
78	54	92,0004	0,97152	1	13	1	25
79	53	91,9604	0,97637	1	101	1	26
80	52	91,7611	1,00058	26	66	26	3
81	51	91,7531	1,00155	3	47	3	24
82	50	91,7299	1,00437	1	26	1	29
83	49	91,5310	1,02853	1	90	1	30
84	48	91,5137	1,03062	1	95	1	31
85	47	91,4134	1,04281	44	105	44	2
86	46	91,3838	1,04640	54	102	54	3
87	45	91,3348	1,05235	37	124	37	2
88	44	91,0878	1,08235	1	111	1	32
89	43	91,0621	1,08547	1	118	1	33
90	42	91,0312	1,08923	3	67	3	25
91	41	90,8476	1,11152	6	64	6	21
92	40	90,7529	1,12302	1	36	1	34
93	39	90,7469	1,12375	76	79	76	2
94	38	90,7425	1,12429	1	37	1	36
95	37	90,7329	1,12545	1	17	1	39
96	36	90,6980	1,12969	2	55	2	3
97	35	90,5625	1,14614	3	131	3	26
98	34	90,5622	1,14618	2	54	2	6
99	33	90,5380	1,14912	1	2	1	45
100	32	90,4305	1,16217	63	70	63	4
101	31	90,4184	1,16364	52	63	52	5
102	30	90,3658	1,17003	1	96	1	46
103	29	90,2889	1,17937	3	19	3	27
104	28	90,2858	1,17975	6	50	6	22
105	27	90,2212	1,18760	32	107	32	6
106	26	90,1767	1,19299	1	32	1	52
107	25	90,0975	1,20262	1	29	1	53
108	24	90,0763	1,20520	3	52	3	32
109	23	90,0529	1,20804	3	44	3	34
110	22	89,9789	1,21702	1	4	1	54
111	21	89,7508	1,24473	1	46	1	56
112	20	89,6417	1,25797	1	31	1	57
113	19	89,5871	1,26460	1	128	1	58
114	18	89,4650	1,27943	6	45	6	24
115	17	89,2903	1,30064	1	6	1	82
116	16	89,2677	1,30339	1	76	1	84
117	15	89,1847	1,31348	1	104	1	85
118	14	89,0846	1,32563	1	83	1	86
119	13	88,9243	1,34509	1	125	1	87
120	12	88,0022	1,45708	15	86	15	2
121	11	87,9001	1,46948	1	10	1	88
122	10	87,6400	1,50108	1	3	1	122
123	9	87,4876	1,51958	1	39	1	123
124	8	86,7739	1,60625	1	94	1	124
125	7	86,7562	1,60841	1	25	1	125
126	6	86,6715	1,61869	1	24	1	126
127	5	86,3742	1,65480	1	82	1	127
128	4	86,0054	1,69958	1	130	1	128
129	3	84,7631	1,85045	1	57	1	129
130	2	83,7169	1,97752	1	15	1	131
131	1	82,7794	2,09136	1	109	1	132

Final Partition
Number of clusters: 30

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	46	175,786	1,84757	2,98715
Cluster2	26	52,655	1,35354	2,34734
Cluster3	1	0,000	0,00000	0,00000
Cluster4	21	34,537	1,13420	2,58417
Cluster5	1	0,000	0,00000	0,00000
Cluster6	1	0,000	0,00000	0,00000
Cluster7	1	0,000	0,00000	0,00000
Cluster8	1	0,000	0,00000	0,00000
Cluster9	1	0,000	0,00000	0,00000
Cluster10	1	0,000	0,00000	0,00000
Cluster11	1	0,000	0,00000	0,00000
Cluster12	4	0,912	0,45292	0,65026
Cluster13	1	0,000	0,00000	0,00000
Cluster14	2	0,544	0,52140	0,52140
Cluster15	2	0,141	0,26513	0,26513
Cluster16	2	0,456	0,47753	0,47753
Cluster17	1	0,000	0,00000	0,00000
Cluster18	5	4,677	0,90973	1,52862
Cluster19	1	0,000	0,00000	0,00000
Cluster20	2	0,631	0,56188	0,56188
Cluster21	1	0,000	0,00000	0,00000
Cluster22	1	0,000	0,00000	0,00000
Cluster23	1	0,000	0,00000	0,00000
Cluster24	1	0,000	0,00000	0,00000
Cluster25	1	0,000	0,00000	0,00000
Cluster26	2	0,322	0,40128	0,40128
Cluster27	1	0,000	0,00000	0,00000
Cluster28	1	0,000	0,00000	0,00000
Cluster29	1	0,000	0,00000	0,00000
Cluster30	1	0,000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	0,094681	-4,30661	1,37842	4,86853	0,82789	-0,93497	-4,71466
PC2	-0,882724	0,47305	-1,51248	0,83368	-2,73223	2,55667	-0,10182
PC3	0,485939	-0,67780	1,48507	-1,04324	-1,68497	2,39496	-1,75471
PC4	-0,024922	0,12176	-0,27224	-0,60218	-0,51760	-0,11082	0,86305

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
PC1	-5,91577	2,92015	2,07489	0,21528	2,58913	3,71464
PC2	-0,06017	1,72842	0,97443	-1,80402	0,35416	1,46393
PC3	-2,12813	2,42128	3,43507	1,24027	0,47235	0,23237
PC4	-0,42358	0,13167	-0,39393	-1,30174	2,47330	1,83555

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
PC1	-2,84422	2,20421	-1,19302	2,65370	-3,23709	3,45616
PC2	1,87067	-1,51016	-0,33109	-1,13275	1,44840	0,02768
PC3	0,12912	0,16604	2,55101	-0,96679	0,03626	0,53731
PC4	0,81801	0,88215	-1,13050	1,03505	-0,60530	-1,39443

Variable	Cluster20	Cluster21	Cluster22	Cluster23	Cluster24	Cluster25
PC1	2,04598	-1,62872	1,19502	-1,82165	-4,09166	0,95944
PC2	1,27229	0,80819	-0,02087	2,39276	-1,03798	-1,55029
PC3	0,58583	-0,70718	1,15875	1,51218	0,39871	-0,63366
PC4	0,06656	1,75730	0,88940	-0,83932	-1,67095	-1,76850

Variable	Cluster26	Cluster27	Cluster28	Cluster29	Cluster30	Grand centroid
PC1	3,24745	-0,18255	2,24004	2,41745	-0,47767	0,0000000
PC2	-0,91402	1,99835	0,36682	1,34274	-2,34275	-0,0000000
PC3	0,20024	-0,09115	-0,85475	1,16062	0,68724	-0,0000000

PC4	2,75716	0,23735	0,02357	1,54113	-2,64797	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,00000	4,75240	1,76182	5,3298	2,98560	4,06711	5,43585
Cluster2	4,75240	0,00000	6,41054	9,2180	6,16928	5,02049	1,48534
Cluster3	1,76182	6,41054	0,00000	4,9180	3,44967	4,77113	7,13446
Cluster4	5,32985	9,21796	4,91799	0,0000	5,42784	6,97940	9,76554
Cluster5	2,98560	6,16928	3,44967	5,4278	0,00000	6,92036	6,28887
Cluster6	4,06711	5,02049	4,77113	6,9794	6,92036	0,00000	6,28663
Cluster7	5,43585	1,48534	7,13446	9,7655	6,28887	6,28663	0,00000
Cluster8	6,61774	2,29665	8,26998	10,8770	7,26788	7,22582	1,79980
Cluster9	4,30946	7,96281	3,73095	4,1398	6,44653	3,95063	8,92266
Cluster10	4,02539	7,62604	3,23838	5,2842	6,44394	3,56717	8,70460
Cluster11	1,75003	5,59802	1,59928	5,8578	3,22626	4,80524	6,39202
Cluster12	3,74078	7,37685	3,67610	4,1450	5,12125	5,25780	7,81703
Cluster13	4,70495	8,31190	4,50877	3,0493	5,92912	5,59269	8,85431
Cluster14	4,12991	2,28644	5,68362	7,9971	6,30443	3,18003	3,30758
Cluster15	2,40185	6,90030	1,93763	4,0321	2,96196	5,68750	7,31736
Cluster16	2,72934	4,72593	3,14337	7,1622	5,30742	3,07731	5,91337
Cluster17	3,13769	7,20709	3,08078	3,3851	2,96958	6,25339	7,48371
Cluster18	4,13206	1,77022	5,68152	8,2003	6,08046	3,51226	3,15437
Cluster19	3,74253	8,01462	2,97434	2,4021	4,49802	5,54765	8,78228
Cluster20	2,91032	6,52640	3,02060	3,3556	4,79767	3,72023	7,32850
Cluster21	3,22942	3,15583	4,83253	6,9206	4,96998	4,08064	3,49973
Cluster22	1,80061	5,87146	1,92728	4,6150	4,18958	3,70260	6,58937
Cluster23	4,01466	3,94714	5,08071	7,3332	6,60385	1,45708	5,30688
Cluster24	4,50184	2,58883	5,76918	9,3279	5,72227	5,41343	3,51039
Cluster25	2,34244	5,94981	2,62770	4,7426	2,02096	5,69003	6,51730
Cluster26	4,21456	8,16726	3,83138	4,3028	4,84130	6,52538	8,45366
Cluster27	2,96298	4,43755	4,18410	5,3368	5,14874	2,67952	5,30185
Cluster28	2,82200	6,55063	3,13630	2,7485	3,54688	5,04528	7,07828
Cluster29	3,64083	7,16685	3,55323	3,9646	5,60962	4,11885	7,86845
Cluster30	3,06271	5,66782	3,22724	6,7714	3,46728	5,79370	6,42372

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	6,6177	4,3095	4,02539	1,75003	3,74078	4,7049
Cluster2	2,2966	7,9628	7,62604	5,59802	7,37685	8,3119
Cluster3	8,2700	3,7310	3,23838	1,59928	3,67610	4,5088
Cluster4	10,8770	4,1398	5,28421	5,85782	4,14502	3,0493
Cluster5	7,2679	6,4465	6,44394	3,22626	5,12125	5,9291
Cluster6	7,2258	3,9506	3,56717	4,80524	5,25780	5,5927
Cluster7	1,7998	8,9227	8,70460	6,39202	7,81703	8,8543
Cluster8	0,0000	10,1133	9,79138	7,26279	9,36266	10,2832
Cluster9	10,1133	0,0000	1,60841	4,82120	3,35853	2,8975
Cluster10	9,7914	1,6084	0,00000	4,10112	4,20094	4,2610
Cluster11	7,2628	4,8212	4,10112	0,00000	5,01334	5,8124
Cluster12	9,3627	3,3585	4,20094	5,01334	0,00000	1,7212
Cluster13	10,2832	2,8975	4,26101	5,81236	1,72124	0,0000
Cluster14	4,4496	6,2429	6,11547	5,34714	5,88888	6,6506
Cluster15	8,6605	4,0805	4,30172	3,15683	2,49989	3,4699
Cluster16	6,6912	4,7718	3,70239	2,42903	5,66413	6,4406
Cluster17	8,8353	4,5335	5,11797	4,09002	2,52087	3,1539
Cluster18	3,7641	6,6499	6,32758	4,94285	6,69404	7,3704
Cluster19	9,7922	3,0097	3,49317	3,78962	3,97768	3,5574
Cluster20	8,5306	2,0846	2,90168	3,88790	2,63500	2,4648
Cluster21	5,0900	5,8283	5,96073	4,83469	4,46098	5,4654
Cluster22	7,9431	2,8643	2,93144	2,99120	2,25035	3,2104
Cluster23	6,0169	4,9694	4,59245	4,69572	5,97207	6,3487
Cluster24	3,4963	8,0099	7,27505	4,47002	7,98444	8,9175
Cluster25	7,3165	5,2476	5,10514	2,08511	5,04965	5,5151
Cluster26	10,0116	4,3492	5,03335	5,24810	1,48199	2,5929
Cluster27	6,4570	4,0029	4,35628	4,33103	3,96265	4,2582
Cluster28	8,2777	3,6139	4,35583	3,86753	2,80792	2,8006
Cluster29	9,2783	1,9943	3,02830	4,77912	1,53278	1,6266
Cluster30	6,9034	6,2332	5,49095	1,69958	6,55379	7,2371

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19

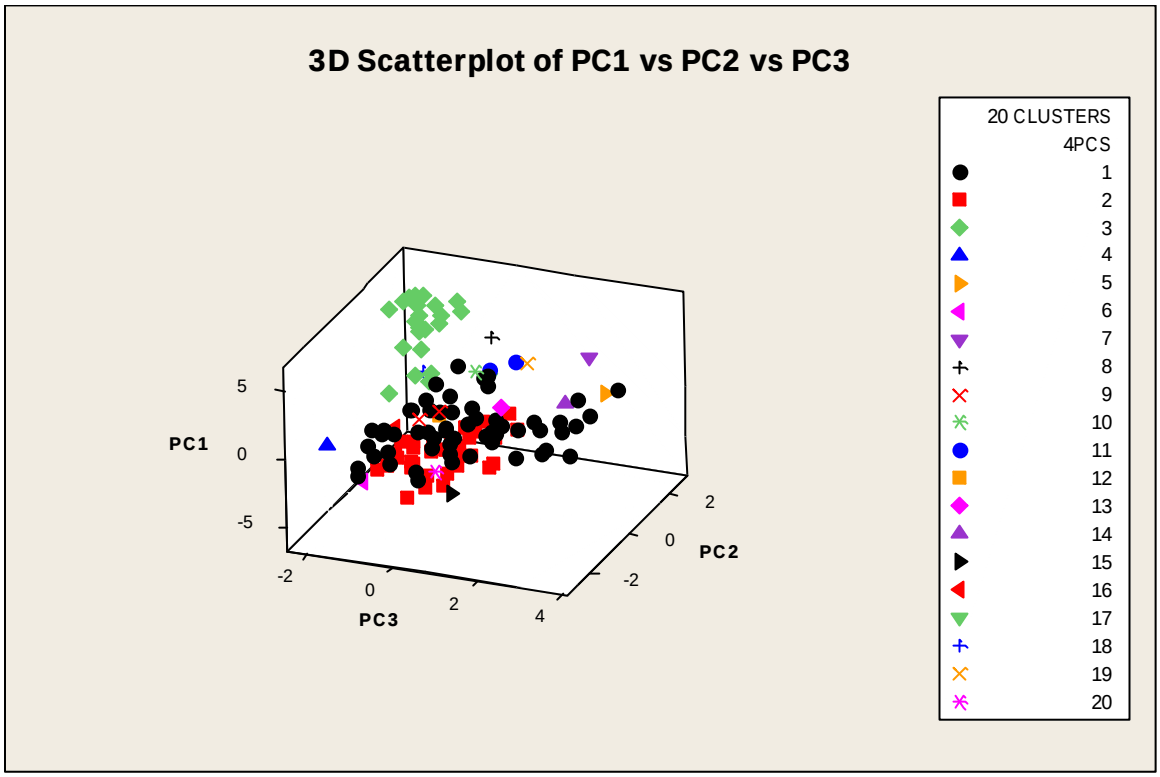
Cluster1	4,12991	2,40185	2,72934	3,13769	4,13206	3,74253
Cluster2	2,28644	6,90030	4,72593	7,20709	1,77022	8,01462
Cluster3	5,68362	1,93763	3,14337	3,08078	5,68152	2,97434
Cluster4	7,99707	4,03209	7,16218	3,38506	8,20026	2,40213
Cluster5	6,30443	2,96196	5,30742	2,96958	6,08046	4,49802
Cluster6	3,18003	5,68750	3,07731	6,25339	3,51226	5,54765
Cluster7	3,30758	7,31736	5,91337	7,48371	3,15437	8,78228
Cluster8	4,44964	8,66053	6,69118	8,83531	3,76414	9,79224
Cluster9	6,24286	4,08048	4,77175	4,53348	6,64992	3,00968
Cluster10	6,11547	4,30172	3,70239	5,11797	6,32758	3,49317
Cluster11	5,34714	3,15683	2,42903	4,09002	4,94285	3,78962
Cluster12	5,88888	2,49989	5,66413	2,52087	6,69404	3,97768
Cluster13	6,65057	3,46986	6,44056	3,15391	7,37041	3,55744
Cluster14	0,00000	6,07636	4,15168	6,36363	1,53854	6,93923
Cluster15	6,07636	0,00000	4,76132	1,28498	6,37104	3,04187
Cluster16	4,15168	4,76132	0,00000	5,70127	3,73425	5,08608
Cluster17	6,36363	1,28498	5,70127	0,00000	6,71272	3,18673
Cluster18	1,53854	6,37104	3,73425	6,71272	0,00000	6,90592
Cluster19	6,93923	3,04187	5,08608	3,18673	6,90592	0,00000
Cluster20	5,00453	2,93402	4,28449	3,08256	5,35679	2,38213
Cluster21	2,04645	4,64699	4,52138	4,76398	3,02180	6,16014
Cluster22	4,57806	2,05474	3,43763	2,81126	5,02956	3,27371
Cluster23	2,44496	6,01796	2,99641	6,48973	2,26459	5,89130
Cluster24	4,03534	6,81420	3,71837	7,39567	2,85996	7,62895
Cluster25	5,78296	3,03587	4,08273	3,31899	5,32962	3,19917
Cluster26	6,97340	2,22723	6,37947	2,17440	7,67876	4,27548
Cluster27	2,73614	4,29979	3,91144	4,38762	3,21853	4,49237
Cluster28	5,45072	2,30293	5,02036	1,85889	5,68841	2,35425
Cluster29	5,43606	3,09967	4,99079	3,31148	6,15275	3,43721
Cluster30	5,97311	4,54082	3,21479	5,25072	5,15592	4,76317

	Cluster20	Cluster21	Cluster22	Cluster23	Cluster24	Cluster25
Cluster1	2,91032	3,22942	1,80061	4,01466	4,50184	2,34244
Cluster2	6,52640	3,15583	5,87146	3,94714	2,58883	5,94981
Cluster3	3,02060	4,83253	1,92728	5,08071	5,76918	2,62770
Cluster4	3,35563	6,92062	4,61503	7,33319	9,32789	4,74261
Cluster5	4,79767	4,96998	4,18958	6,60385	5,72227	2,02096
Cluster6	3,72023	4,08064	3,70260	1,45708	5,41343	5,69003
Cluster7	7,32850	3,49973	6,58937	5,30688	3,51039	6,51730
Cluster8	8,53058	5,09001	7,94307	6,01692	3,49633	7,31650
Cluster9	2,08455	5,82829	2,86431	4,96944	8,00991	5,24763
Cluster10	2,90168	5,96073	2,93144	4,59245	7,27505	5,10514
Cluster11	3,88790	4,83469	2,99120	4,69572	4,47002	2,08511
Cluster12	2,63500	4,46098	2,25035	5,97207	7,98444	5,04965
Cluster13	2,46484	5,46538	3,21038	6,34871	8,91746	5,51505
Cluster14	5,00453	2,04645	4,57806	2,44496	4,03534	5,78296
Cluster15	2,93402	4,64699	2,05474	6,01796	6,81420	3,03587
Cluster16	4,28449	4,52138	3,43763	2,99641	3,71837	4,08273
Cluster17	3,08256	4,76398	2,81126	6,48973	7,39567	3,31899
Cluster18	5,35679	3,02180	5,02956	2,26459	2,85996	5,32962
Cluster19	2,38213	6,16014	3,27371	5,89130	7,62895	3,19917
Cluster20	0,00000	4,27191	1,84437	4,22998	6,78689	3,74194
Cluster21	4,27191	0,00000	3,59107	3,77041	4,73817	4,96967
Cluster22	1,84437	3,59107	0,00000	4,24727	6,00970	3,55974
Cluster23	4,22998	3,77041	4,24727	0,00000	4,34215	5,36191
Cluster24	6,78689	4,73817	6,00970	4,34215	0,00000	5,18183
Cluster25	3,74194	4,96967	3,55974	5,36191	5,18183	0,00000
Cluster26	3,68938	5,34474	3,06879	7,16146	8,57470	5,17849
Cluster27	2,44560	2,48950	2,82177	2,56361	5,32747	4,26787
Cluster28	1,71308	4,26494	2,45890	5,19121	6,81956	2,92841
Cluster29	1,62717	4,49362	1,94385	4,98625	7,67687	4,96674
Cluster30	5,17844	5,71000	4,57431	5,30877	3,97507	2,28288

	Cluster26	Cluster27	Cluster28	Cluster29	Cluster30
Cluster1	4,2146	2,96298	2,82200	3,64083	3,06271
Cluster2	8,1673	4,43755	6,55063	7,16685	5,66782
Cluster3	3,8314	4,18410	3,13630	3,55323	3,22724
Cluster4	4,3028	5,33677	2,74846	3,96456	6,77136
Cluster5	4,8413	5,14874	3,54688	5,60962	3,46728
Cluster6	6,5254	2,67952	5,04528	4,11885	5,79370

Cluster7	8,4537	5,30185	7,07828	7,86845	6,42372
Cluster8	10,0116	6,45705	8,27774	9,27829	6,90343
Cluster9	4,3492	4,00288	3,61394	1,99432	6,23321
Cluster10	5,0334	4,35628	4,35583	3,02830	5,49095
Cluster11	5,2481	4,33103	3,86753	4,77912	1,69958
Cluster12	1,4820	3,96265	2,80792	1,53278	6,55379
Cluster13	2,5929	4,25824	2,80058	1,62657	7,23707
Cluster14	6,9734	2,73614	5,45072	5,43606	5,97311
Cluster15	2,2272	4,29979	2,30293	3,09967	4,54082
Cluster16	6,3795	3,91144	5,02036	4,99079	3,21479
Cluster17	2,1744	4,38762	1,85889	3,31148	5,25072
Cluster18	7,6788	3,21853	5,68841	6,15275	5,15592
Cluster19	4,2755	4,49237	2,35425	3,43721	4,76317
Cluster20	3,6894	2,44560	1,71308	1,62717	5,17844
Cluster21	5,3447	2,48950	4,26494	4,49362	5,71000
Cluster22	3,0688	2,82177	2,45890	1,94385	4,57431
Cluster23	7,1615	2,56361	5,19121	4,98625	5,30877
Cluster24	8,5747	5,32747	6,81956	7,67687	3,97507
Cluster25	5,1785	4,26787	2,92841	4,96674	2,28288
Cluster26	0,0000	5,16538	3,35275	2,86059	6,73575
Cluster27	5,1654	0,00000	3,02648	3,23366	5,27856
Cluster28	3,3528	3,02648	0,00000	2,71083	4,92367
Cluster29	2,8606	3,23366	2,71083	0,00000	6,30375
Cluster30	6,7358	5,27856	4,92367	6,30375	0,00000

c. GRÁFICO 3D DO EXERCICIO COM 20 CLUSTERS – COMPONENTES PRINCIPAIS



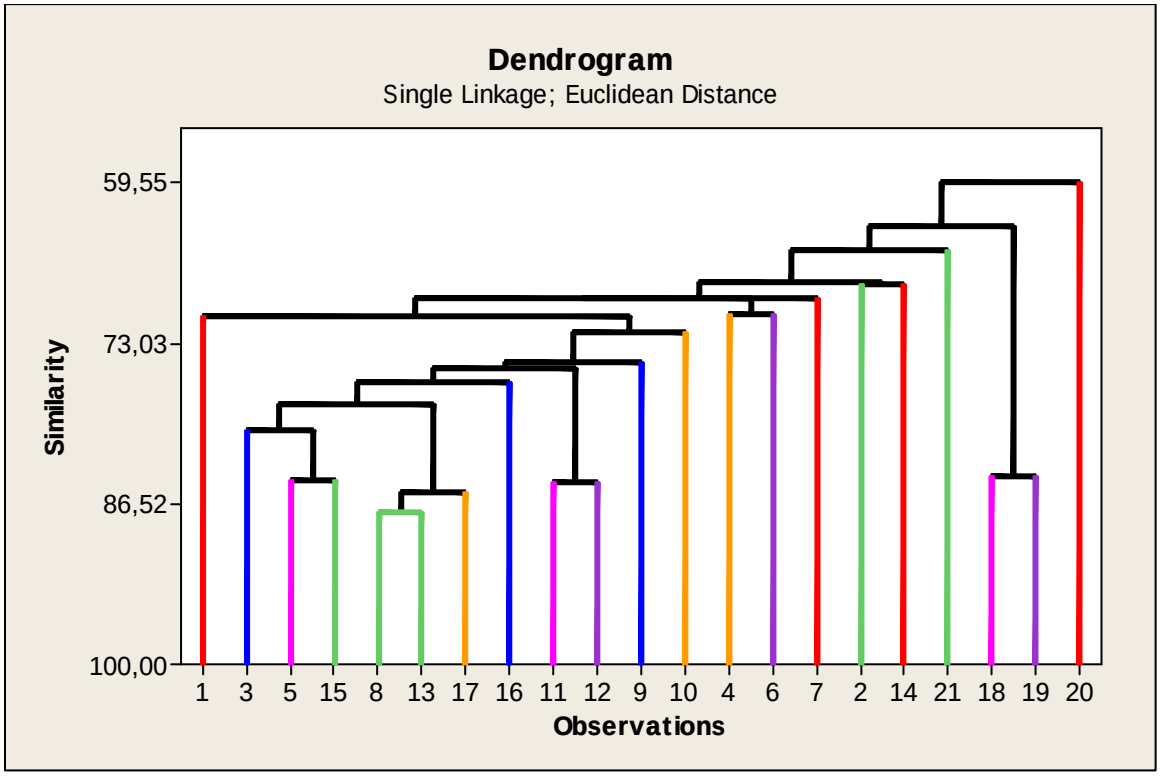
Com uma representação gráfica 3D com as três componentes principais mais relevantes, chegamos a uma visualização de três grupos principais, um mais concentrado em PC2 e PC3 e variando com PC1 mas com concentração em graus maiores desta variável (cluster 1), outro mais concentrado em PC 1 e PC3 e variando em PC2 (cluster 2), e um terceiro variando em PC2 e PC3 e menor variação em PC1 (cluster 3).

d. ANÁLISE DOS AGRUPAMENTOS POR REGIÃO – AIBER E AVECO

Repete-se o processo com as observações de AIBER – PAÍSES IBERICOS E IBEROAMERICA e AVECO – ECONOMIAS AVANÇDAS

i. AIBER – VARIÁVEIS RELEVANTES E 20 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 20 clusters com as oito variáveis principais. O resultado é bastante fragmentado.



Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	87,1805	14,7906	8	13	8	2
2	19	85,5817	16,6351	8	17	8	3
3	18	84,6721	17,6846	11	12	11	2
4	17	84,5351	17,8427	5	15	5	2
5	16	84,1797	18,2527	18	19	18	2
6	15	80,3371	22,6862	3	5	3	3
7	14	78,2053	25,1457	3	8	3	6
8	13	76,2922	27,3529	3	16	3	7
9	12	75,1942	28,6197	3	11	3	9
10	11	74,6732	29,2209	3	9	3	10
11	10	72,1068	32,1819	3	10	3	11
12	9	70,8922	33,5832	1	3	1	12
13	8	70,6597	33,8515	4	6	4	2
14	7	69,3875	35,3192	1	4	1	14

15	6	69,3311	35,3844	1	7	1	15
16	5	68,1922	36,6983	2	14	2	2
17	4	67,9230	37,0090	1	2	1	17
18	3	65,2768	40,0620	1	21	1	18
19	2	63,2640	42,3843	1	18	1	20
20	1	59,5477	46,6720	1	20	1	21

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	1	0,000	0,00000	0,00000
Cluster2	1	0,000	0,00000	0,00000
Cluster3	1	0,000	0,00000	0,00000
Cluster4	1	0,000	0,00000	0,00000
Cluster5	1	0,000	0,00000	0,00000
Cluster6	1	0,000	0,00000	0,00000
Cluster7	1	0,000	0,00000	0,00000
Cluster8	2	109,380	7,39528	7,39528
Cluster9	1	0,000	0,00000	0,00000
Cluster10	1	0,000	0,00000	0,00000
Cluster11	1	0,000	0,00000	0,00000
Cluster12	1	0,000	0,00000	0,00000
Cluster13	1	0,000	0,00000	0,00000
Cluster14	1	0,000	0,00000	0,00000
Cluster15	1	0,000	0,00000	0,00000
Cluster16	1	0,000	0,00000	0,00000
Cluster17	1	0,000	0,00000	0,00000
Cluster18	1	0,000	0,00000	0,00000
Cluster19	1	0,000	0,00000	0,00000
Cluster20	1	0,000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
MN - Happy Planet Index_1	75,9176	50,6379	73,2079	75,503
MN - GINNI	51,0699	43,0585	27,3225	28,914
MN - Social Progress Index_1	73,1647	56,0941	69,3186	82,315
MN-Access to improved sanitatio	96,9635	41,1035	79,4733	98,747
MN - International Cooperati_1	47,8261	56,5217	91,3043	100,000
MN - (HDI)_1	81,8165	52,6468	68,3381	81,247
MN - Democracy index_1	63,3886	50,9479	69,7867	74,763
MN Efficiency of legal framewor	23,5153	48,4903	36,9818	64,699

Variable	Cluster5	Cluster6	Cluster7	Cluster8
MN - Happy Planet Index_1	89,6612	100,000	81,0597	70,4092
MN - GINNI	25,0783	37,918	45,0180	42,4452
MN - Social Progress Index_1	65,7359	81,595	51,6509	58,9480
MN-Access to improved sanitatio	78,2305	93,298	91,8272	82,0365
MN - International Cooperati_1	73,9130	91,304	47,8261	65,2174
MN - (HDI)_1	62,4415	70,093	70,6415	65,0351
MN - Democracy index_1	59,9526	77,488	24,0521	61,4336
MN Efficiency of legal framewor	40,6974	49,628	45,0180	40,4014

Variable	Cluster9	Cluster10	Cluster11	Cluster12
MN - Happy Planet Index_1	72,1206	87,5764	82,6885	80,5517
MN - GINNI	43,2411	55,6889	28,1576	14,9791
MN - Social Progress Index_1	64,6821	57,7626	54,0393	52,7222
MN-Access to improved sanitatio	81,3867	67,5351	78,3498	77,9937
MN - International Cooperati_1	39,1304	91,3043	56,5217	65,2174
MN - (HDI)_1	64,3730	53,3106	46,8341	43,2825

MN - Democracy index_1	51,8957	59,7156	51,1848	51,5403
MN Efficiency of legal framewor	45,0180	42,4481	39,5835	46,2387

Variable	Cluster13	Cluster14	Cluster15	Cluster16
MN - Happy Planet Index_1	83,1763	84,9502	56,0612	71,8495
MN - GINNI	45,4332	29,3319	39,4833	46,4770
MN - Social Progress Index_1	54,0569	70,8992	62,6625	62,8908
MN-Access to improved sanitatio	47,3788	70,4832	77,7305	70,3758
MN - International Cooperati_1	47,8261	78,2609	78,2609	69,5652
MN - (HDI)_1	47,5431	72,4323	55,5570	64,7975
MN - Democracy index_1	45,3791	66,2322	56,5166	59,8341
MN Efficiency of legal framewor	38,1863	46,4629	19,4515	34,4131

Variable	Cluster17	Cluster18	Cluster19	Cluster20
MN - Happy Planet Index_1	38,814	51,8078	40,367	82,7132
MN - GINNI	64,431	71,4770	56,942	47,9384
MN - Social Progress Index_1	88,672	87,3727	83,930	56,2522
MN-Access to improved sanitatio	100,000	99,9880	96,091	89,9656
MN - International Cooperati_1	45,018	45,0180	100,000	4,3478
MN - (HDI)_1	80,897	88,6249	74,629	69,5067
MN - Democracy index_1	74,645	77,7251	79,147	42,4171
MN Efficiency of legal framewor	35,007	41,6235	55,454	45,0180

Variable	Grand centroid
MN - Happy Planet Index_1	72,3563
MN - GINNI	42,2309
MN - Social Progress Index_1	66,3673
MN-Access to improved sanitatio	80,9998
MN - International Cooperati_1	64,7429
MN - (HDI)_1	65,6706
MN - Democracy index_1	59,9752
MN Efficiency of legal framewor	41,8447

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,0000	76,2890	56,4472	71,595	51,3178	61,2040	51,8353
Cluster2	76,2890	0,0000	65,9360	91,471	62,1104	89,7637	68,2409
Cluster3	56,4472	65,9360	0,0000	39,793	27,1638	37,3394	69,9389
Cluster4	71,5954	91,4705	39,7927	0,000	52,3503	33,8515	84,0946
Cluster5	51,3178	62,1104	27,1638	52,350	0,0000	38,7110	53,9455
Cluster6	61,2040	89,7637	37,3394	33,851	38,7110	0,0000	77,9440
Cluster7	51,8353	68,2409	69,9389	84,095	53,9455	77,9440	0,0000
Cluster8	37,4526	49,7936	33,5195	57,228	28,5908	50,8411	44,9637
Cluster9	36,9411	51,0974	58,3288	75,691	44,1109	68,2978	35,3844
Cluster10	65,8438	59,5739	40,5500	63,382	38,8365	48,4867	65,2151
Cluster11	54,6770	52,4782	48,4804	74,358	28,6197	62,6202	43,3866
Cluster12	67,4228	56,7827	47,0198	71,115	29,9758	63,6899	53,8810
Cluster13	67,9887	36,6983	67,8993	95,882	51,6313	82,2632	54,9650
Cluster14	53,0046	59,5810	22,6862	44,432	17,8427	35,3192	61,6278
Cluster15	51,8489	52,7465	36,1404	69,502	43,3088	65,3127	61,9430
Cluster16	41,7581	44,2673	32,9181	61,976	30,1482	52,7712	50,7248
Cluster17	45,5234	82,4650	75,3892	80,959	80,9804	83,8016	79,8237
Cluster18	42,3843	86,2882	76,1113	77,582	79,7776	77,9541	78,2432
Cluster19	73,8727	85,5221	54,6766	46,672	74,2329	63,7510	93,4913
Cluster20	57,7499	80,8170	95,7740	109,046	77,5494	99,3258	47,5901

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	37,4526	36,9411	65,8438	54,6770	67,4228	67,989
Cluster2	49,7936	51,0974	59,5739	52,4782	56,7827	36,698
Cluster3	33,5195	58,3288	40,5500	48,4804	47,0198	67,899
Cluster4	57,2275	75,6907	63,3815	74,3578	71,1148	95,882
Cluster5	28,5908	44,1109	38,8365	28,6197	29,9758	51,631
Cluster6	50,8411	68,2978	48,4867	62,6202	63,6899	82,263
Cluster7	44,9637	35,3844	65,2151	43,3866	53,8810	54,965
Cluster8	0,0000	28,8117	38,8193	30,0866	38,9550	47,624

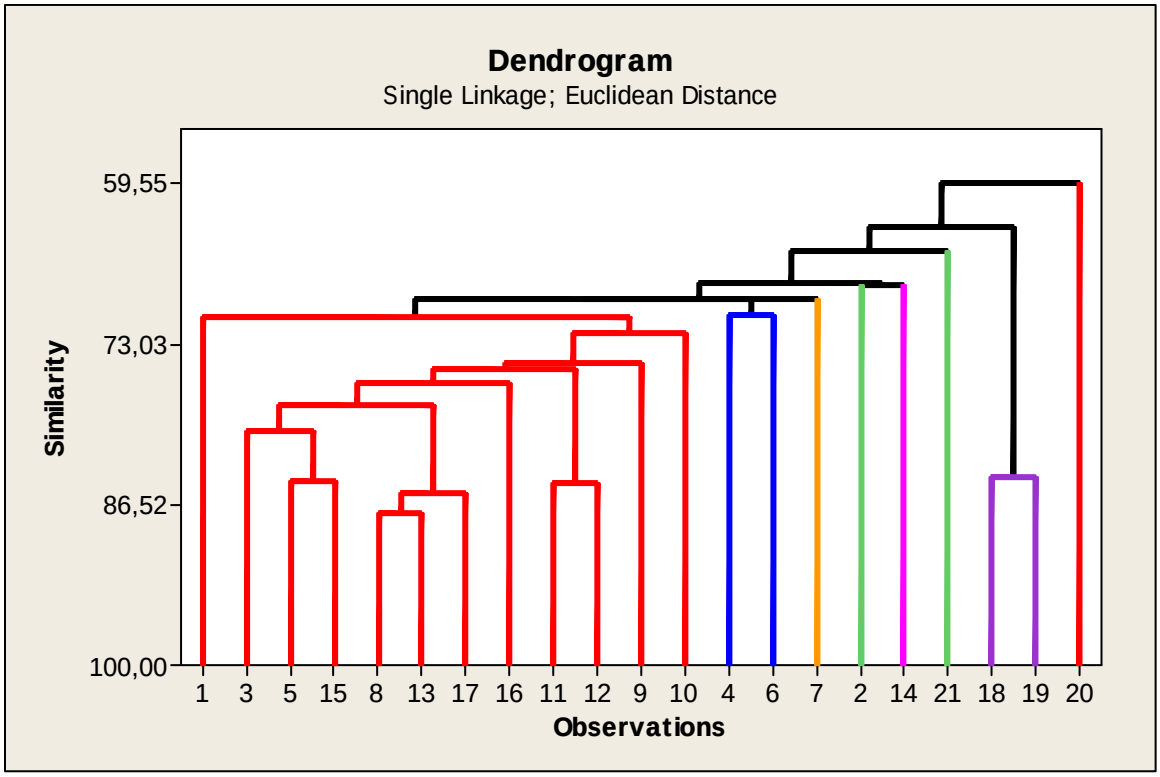
Cluster9	28,8117	0,0000	59,5468	33,1933	46,3824	42,947
Cluster10	38,8193	59,5468	0,0000	47,3861	51,9973	51,883
Cluster11	30,0866	33,1933	47,3861	0,0000	17,6846	37,009
Cluster12	38,9550	46,3824	51,9973	17,6846	0,0000	47,925
Cluster13	47,6245	42,9472	51,8832	37,0090	47,9250	0,000
Cluster14	30,7300	48,1598	38,4345	41,8837	42,9873	55,937
Cluster15	31,1452	50,7285	45,8043	43,5250	48,6028	56,846
Cluster16	15,0740	35,1700	32,1819	34,4873	43,8145	41,591
Cluster17	59,5910	58,3589	87,4371	82,2972	91,9078	91,011
Cluster18	59,5858	58,0710	84,6944	83,6472	94,1994	91,069
Cluster19	61,4538	80,2781	69,3353	86,7443	87,8445	100,621
Cluster20	66,0284	40,0620	93,3741	62,2389	75,6592	65,216

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	53,0046	51,8489	41,7581	45,5234	42,3843	73,873
Cluster2	59,5810	52,7465	44,2673	82,4650	86,2882	85,522
Cluster3	22,6862	36,1404	32,9181	75,3892	76,1113	54,677
Cluster4	44,4321	69,5016	61,9758	80,9588	77,5816	46,672
Cluster5	17,8427	43,3088	30,1482	80,9804	79,7776	74,233
Cluster6	35,3192	65,3127	52,7712	83,8016	77,9541	63,751
Cluster7	61,6278	61,9430	50,7248	79,8237	78,2432	93,491
Cluster8	30,7300	31,1452	15,0740	59,5910	59,5858	61,454
Cluster9	48,1598	50,7285	35,1700	58,3589	58,0710	80,278
Cluster10	38,4345	45,8043	32,1819	87,4371	84,6944	69,335
Cluster11	41,8837	43,5250	34,4873	82,2972	83,6472	86,744
Cluster12	42,9873	48,6028	43,8145	91,9078	94,1994	87,845
Cluster13	55,9372	56,8458	41,5906	91,0115	91,0692	100,621
Cluster14	0,0000	46,5487	29,1508	76,9829	74,4267	65,579
Cluster15	46,5487	0,0000	27,3529	66,4086	72,7163	63,128
Cluster16	29,1508	27,3529	0,0000	63,5408	63,6027	63,587
Cluster17	76,9829	66,4086	63,5408	0,0000	18,2527	59,974
Cluster18	74,4267	72,7163	63,6027	18,2527	0,0000	61,493
Cluster19	65,5792	63,1285	63,5867	59,9743	61,4932	0,000
Cluster20	83,5838	86,5031	72,3829	79,2056	76,5238	115,375

	Cluster20
Cluster1	57,750
Cluster2	80,817
Cluster3	95,774
Cluster4	109,046
Cluster5	77,549
Cluster6	99,326
Cluster7	47,590
Cluster8	66,028
Cluster9	40,062
Cluster10	93,374
Cluster11	62,239
Cluster12	75,659
Cluster13	65,216
Cluster14	83,584
Cluster15	86,503
Cluster16	72,383
Cluster17	79,206
Cluster18	76,524
Cluster19	115,375
Cluster20	0,000

ii. AIBER – VARIÁVEIS RELEVANTES E 8 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 8 clusters com as oito variáveis principais. O resultado é bastante fragmentado novamente, com um cluster com 12 observações e os demais com uma ou duas..



Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	87,1805	14,7906	8	13	8	2
2	19	85,5817	16,6351	8	17	8	3
3	18	84,6721	17,6846	11	12	11	2
4	17	84,5351	17,8427	5	15	5	2
5	16	84,1797	18,2527	18	19	18	2
6	15	80,3371	22,6862	3	5	3	3
7	14	78,2053	25,1457	3	8	3	6
8	13	76,2922	27,3529	3	16	3	7
9	12	75,1942	28,6197	3	11	3	9
10	11	74,6732	29,2209	3	9	3	10
11	10	72,1068	32,1819	3	10	3	11
12	9	70,8922	33,5832	1	3	1	12
13	8	70,6597	33,8515	4	6	4	2
14	7	69,3875	35,3192	1	4	1	14
15	6	69,3311	35,3844	1	7	1	15
16	5	68,1922	36,6983	2	14	2	2
17	4	67,9230	37,0090	1	2	1	17
18	3	65,2768	40,0620	1	21	1	18
19	2	63,2640	42,3843	1	18	1	20
20	1	59,5477	46,6720	1	20	1	21

Final Partition
Number of clusters: 8

Number of	Within cluster sum	Average distance from	Maximum distance from

	observations	of squares	centroid	centroid
Cluster1	12	8929,84	25,7440	41,1113
Cluster2	1	0,00	0,0000	0,0000
Cluster3	2	572,96	16,9257	16,9257
Cluster4	1	0,00	0,0000	0,0000
Cluster5	1	0,00	0,0000	0,0000
Cluster6	2	166,58	9,1264	9,1264
Cluster7	1	0,00	0,0000	0,0000
Cluster8	1	0,00	0,0000	0,0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
MN - Happy Planet Index_1	76,2836	50,6379	87,7517	81,0597
MN - GINNI	37,1434	43,0585	33,4160	45,0180
MN - Social Progress Index_1	62,6478	56,0941	81,9547	51,6509
MN-Access to improved sanitatio	78,5496	41,1035	96,0226	91,8272
MN - International Cooperati_1	68,4783	56,5217	95,6522	47,8261
MN - (HDI)_1	61,9378	52,6468	75,6701	70,6415
MN - Democracy index_1	59,4096	50,9479	76,1256	24,0521
MN Efficiency of legal framewor	37,9678	48,4903	57,1640	45,0180

Variable	Cluster5	Cluster6	Cluster7	Cluster8
MN - Happy Planet Index_1	83,1763	45,3109	40,367	82,7132
MN - GINNI	45,4332	67,9541	56,942	47,9384
MN - Social Progress Index_1	54,0569	88,0225	83,930	56,2522
MN-Access to improved sanitatio	47,3788	99,9940	96,091	89,9656
MN - International Cooperati_1	47,8261	45,0180	100,000	4,3478
MN - (HDI)_1	47,5431	84,7609	74,629	69,5067
MN - Democracy index_1	45,3791	76,1848	79,147	42,4171
MN Efficiency of legal framewor	38,1863	38,3152	55,454	45,0180

Variable	Grand centroid
MN - Happy Planet Index_1	72,3563
MN - GINNI	42,2309
MN - Social Progress Index_1	66,3673
MN-Access to improved sanitatio	80,9998
MN - International Cooperati_1	64,7429
MN - (HDI)_1	65,6706
MN - Democracy index_1	59,9752
MN Efficiency of legal framewor	41,8447

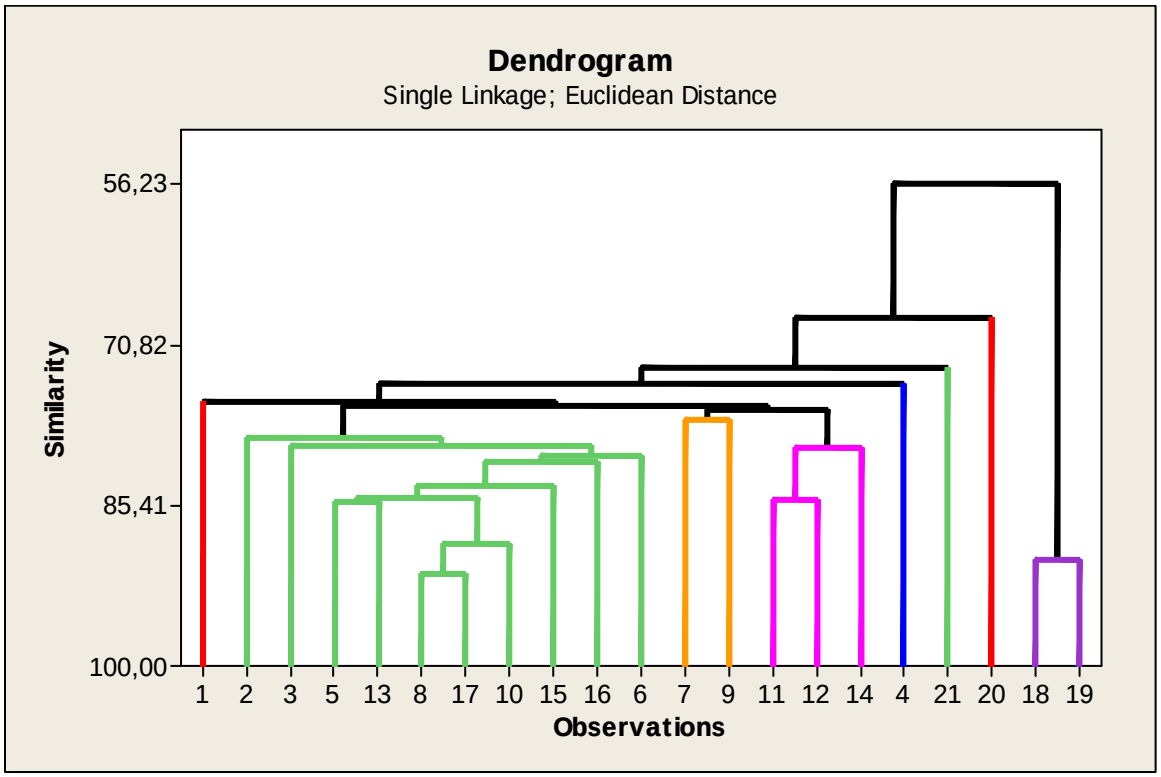
Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,0000	50,4922	48,973	46,7351	44,635	66,0693	65,519
Cluster2	50,4922	0,0000	89,026	68,2409	36,698	83,9033	85,522
Cluster3	48,9734	89,0264	0,000	79,2913	87,714	77,7716	53,242
Cluster4	46,7351	68,2409	79,291	0,0000	54,965	78,5087	93,491
Cluster5	44,6349	36,6983	87,714	54,9650	0,000	90,5818	100,621
Cluster6	66,0693	83,9033	77,772	78,5087	90,582	0,0000	60,049
Cluster7	65,5186	85,5221	53,242	93,4913	100,621	60,0490	0,000
Cluster8	69,5525	80,8170	102,917	47,5901	65,216	77,3396	115,375

	Cluster8
Cluster1	69,553
Cluster2	80,817
Cluster3	102,917
Cluster4	47,590
Cluster5	65,216
Cluster6	77,340
Cluster7	115,375
Cluster8	0,000

iii. AIBER – COMPONENTES PRINCIPAIS 8 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 8 clusters com as quatro componentes principais. O resultado continua bastante fragmentado, com um cluster com 10 observações, outro com três e os demais com uma ou duas.



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

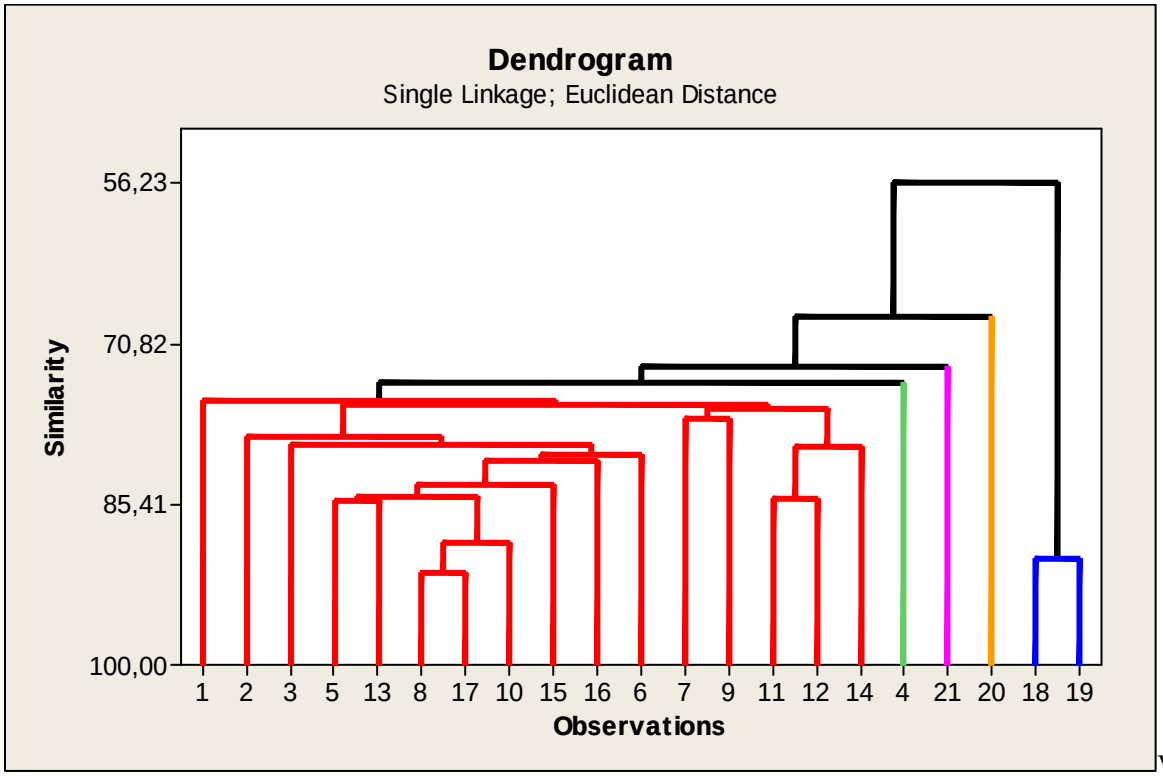
Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	91,6465	0,52658	8	17	8	2
2	19	90,3152	0,61051	18	19	18	2
3	18	88,8836	0,70075	8	10	8	3
4	17	85,1676	0,93499	5	13	5	2
5	16	84,8495	0,95505	11	12	11	2
6	15	84,7989	0,95824	5	8	5	5
7	14	83,6507	1,03062	5	15	5	6
8	13	81,4392	1,17003	5	16	5	7
9	12	80,9222	1,20262	5	6	5	8
10	11	80,2542	1,24473	11	14	11	3
11	10	80,0708	1,25629	3	5	3	9
12	9	79,2823	1,30599	2	3	2	10
13	8	77,7283	1,40395	7	9	7	2
14	7	76,8196	1,46123	7	11	7	5
15	6	76,4304	1,48577	2	7	2	15
16	5	76,0630	1,50893	1	2	1	16
17	4	74,4849	1,60841	1	4	1	17
18	3	73,0385	1,69958	1	21	1	18
19	2	68,3629	1,99432	1	20	1	19

```
  20          1     56,2337   2,75892   1    18        1       21


Final Partition
Number of clusters: 8


                                      Average   Maximum
                             Within  distance  distance
             Number of  cluster sum      from      from
          observations   of squares  centroid  centroid
Cluster1             1       0,0000   0,00000   0,00000
Cluster2            10      19,2037   1,25852   2,27245
Cluster3             1       0,0000   0,00000   0,00000
Cluster4             2       0,9855   0,70198   0,70198
Cluster5             3       1,3686   0,67092   0,77996
Cluster6             2       0,1864   0,30525   0,30525
Cluster7             1       0,0000   0,00000   0,00000
Cluster8             1       0,0000   0,00000   0,00000


Cluster Centroids


Variable  Cluster1  Cluster2  Cluster3  Cluster4  Cluster5  Cluster6  Cluster7
PC1        1,37842   0,27653   2,92015   0,07517  -1,41683   3,49230   2,41745
PC2       -1,51248   0,20413   1,72842  -1,14007  -0,33025  -0,60974   1,34274
PC3        1,48507   2,30313   2,42128   1,23203   2,23565  -0,23483   1,16062
PC4       -0,27224  -0,15885   0,13167  -1,48130  -1,18089   0,33727   1,54113

                       Grand
Variable  Cluster8  centroid
PC1       -0,47767   0,56610
PC2       -2,34275  -0,15396
PC3        0,68724   1,78509
PC4       -2,64797  -0,41270


Distances Between Cluster Centroids

          Cluster1  Cluster2  Cluster3  Cluster4  Cluster5  Cluster6  Cluster7
Cluster1   0,00000   2,20068   3,73095   1,83385   3,25577   2,93479   3,55323
Cluster2   2,20068   0,00000   3,06767   2,17798   2,04991   4,20606   3,17416
Cluster3   3,73095   3,06767   0,00000   4,50979   4,98044   3,59048   1,99432
Cluster4   1,83385   2,17798   4,50979   0,00000   1,99484   4,17335   4,55970
Cluster5   3,25577   2,04991   4,98044   1,99484   0,00000   5,70839   5,10546
Cluster6   2,93479   4,20606   3,59048   4,17335   5,70839   0,00000   2,89207
Cluster7   3,55323   3,17416   1,99432   4,55970   5,10546   2,89207   0,00000
Cluster8   3,22724   3,98275   6,23321   1,84661   3,07930   5,34096   6,30375

          Cluster8
Cluster1   3,22724
Cluster2   3,98275
Cluster3   6,23321
Cluster4   1,84661
Cluster5   3,07930
Cluster6   5,34096
Cluster7   6,30375
Cluster8   0,00000
```

iv. AIBER – COMPONENTES PRINCIPAIS E 5 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 5 clusters com as quatro componentes principais. O resultado continua bastante fragmentado, com um cluster com 14 observações, outros três com uma e outro com duas observações.



V

Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	91,6465	0,52658	8	17	8	2
2	19	90,3152	0,61051	18	19	18	2
3	18	88,8836	0,70075	8	10	8	3
4	17	85,1676	0,93499	5	13	5	2
5	16	84,8495	0,95505	11	12	11	2
6	15	84,7989	0,95824	5	8	5	5
7	14	83,6507	1,03062	5	15	5	6
8	13	81,4392	1,17003	5	16	5	7
9	12	80,9222	1,20262	5	6	5	8
10	11	80,2542	1,24473	11	14	11	3
11	10	80,0708	1,25629	3	5	3	9
12	9	79,2823	1,30599	2	3	2	10
13	8	77,7283	1,40395	7	9	7	2
14	7	76,8196	1,46123	7	11	7	5
15	6	76,4304	1,48577	2	7	2	15
16	5	76,0630	1,50893	1	2	1	16
17	4	74,4849	1,60841	1	4	1	17
18	3	73,0385	1,69958	1	21	1	18
19	2	68,3629	1,99432	1	20	1	19
20	1	56,2337	2,75892	1	18	1	21

Final Partition
Number of clusters: 5

	Within	Average distance	Maximum distance

	Number of observations	cluster sum of squares	from centroid	from centroid
Cluster1	16	42,2933	1,51657	2,71868
Cluster2	1	0,0000	0,00000	0,00000
Cluster3	2	0,1864	0,30525	0,30525
Cluster4	1	0,0000	0,00000	0,00000
Cluster5	1	0,0000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
PC1	0,00272	2,92015	3,49230	2,41745	-0,47767	0,56610
PC2	-0,17138	1,72842	-0,60974	1,34274	-2,34275	-0,15396
PC3	2,10546	2,42128	-0,23483	1,16062	0,68724	1,78509
PC4	-0,52287	0,13167	0,33727	1,54113	-2,64797	-0,41270

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,00000	3,55652	4,31116	3,64367	3,38719
Cluster2	3,55652	0,00000	3,59048	1,99432	6,23321
Cluster3	4,31116	3,59048	0,00000	2,89207	5,34096
Cluster4	3,64367	1,99432	2,89207	0,00000	6,30375
Cluster5	3,38719	6,23321	5,34096	6,30375	0,00000

v. AIBER – ANÁLISE GRÁFICA COMPONENTES PRINCIPAIS E 5 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 5 clusters com as quatro componentes principais. O resultado continua bastante fragmentado, com um cluster com 14 observações, outros três com uma e outro com duas observações.

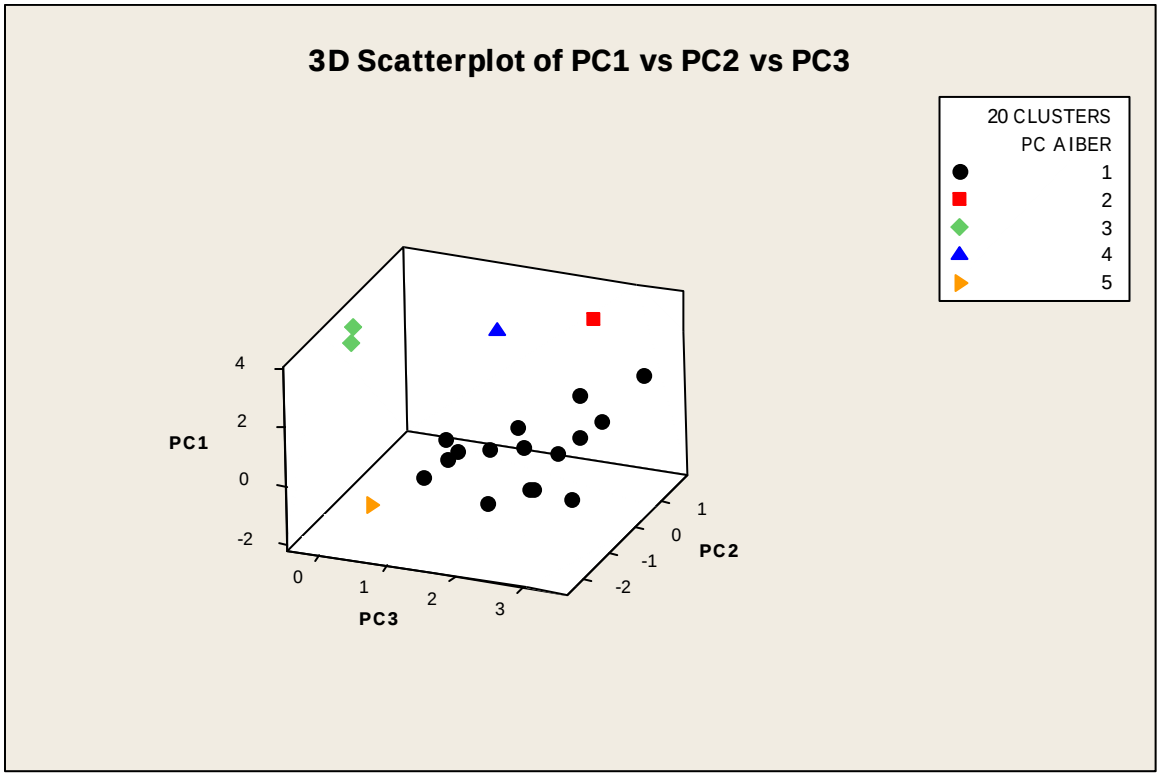


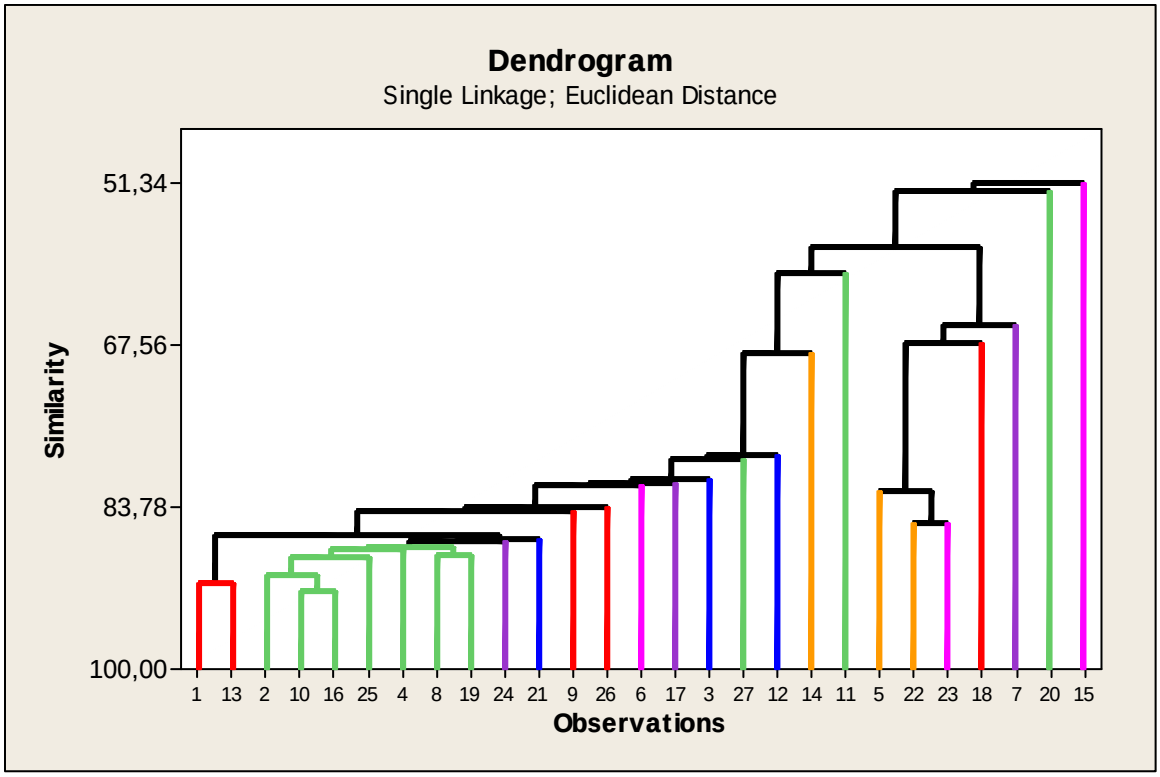
Tabela Países x Clusters x Região

PAISES	CLUSTER3R	20 CLUSTERS SEM REGIÃO	20 8 CLUSTERS AIBER
Argentina	**AIBER**	**1**	**1**
Bolivia	**AIBER**	**1**	**1**
Brazil	**AIBER**	**1**	**1**
Chile	**AIBER**	**7**	**2**
Colombia	**AIBER**	**1**	**1**
Costa Rica	**AIBER**	**1**	**1**
Cuba	**AIBER**	**1**	**1**
Dom. Rep.	**AIBER**	**1**	**1**
Ecuador	**AIBER**	**1**	**1**
El Salvador	**AIBER**	**1**	**1**
Guatemala	**AIBER**	**1**	**1**
Honduras	**AIBER**	**1**	**1**
Mexico	**AIBER**	**1**	**1**
Nicaragua	**AIBER**	**1**	**1**
Panama	**AIBER**	**1**	**1**
Paraguay	**AIBER**	**1**	**1**
Peru	**AIBER**	**1**	**1**
Portugal	**AIBER**	**3**	**3**
Spain	**AIBER**	**3**	**3**
Uruguay	**AIBER**	**19**	**4**
Venezuela	**AIBER**	**20**	**5**

Observamos pela tabela, que traduz os dados do gráfico em quais países compõe os clusters, que Chile compõe um cluster separado, Portugal e Espanha Outro, Uruguai outro, e Venezuela, outro. Todos os demais países são agrupados em um mesmo cluster sob o ponto de vista das três variáveis principais.

e. AVECO – VARIÁVEIS RELEVANTES E 20 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 20 clusters com as oito variáveis principais. O resultado é bastante fragmentado. Um clster com sete observações, um com duas e os demais com uma observação cada um.



Results for: Worksheet 6(3R CLUSTERS = AVECO)

Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	92,1522	9,3922	10	16	10	2
2	25	91,4080	10,2829	1	13	1	2
3	24	90,6347	11,2083	2	10	2	3
4	23	88,8634	13,3282	2	25	2	4
5	22	88,5775	13,6704	8	19	8	2
6	21	88,0328	14,3223	2	4	2	5
7	20	87,7358	14,6777	2	8	2	7
8	19	87,2458	15,2641	2	24	2	8
9	18	87,0175	15,5374	2	21	2	9
10	17	86,6145	16,0197	1	2	1	11
11	16	85,2922	17,6023	22	23	22	2
12	15	84,1875	18,9244	1	9	1	12
13	14	83,7755	19,4174	1	26	1	13
14	13	82,2401	21,2550	5	22	5	3
15	12	81,6566	21,9532	1	6	1	14
16	11	81,4770	22,1683	1	17	1	15
17	10	80,9079	22,8493	1	3	1	16
18	9	79,0735	25,0448	1	27	1	17
19	8	78,5024	25,7283	1	12	1	18
20	7	68,3041	37,9334	1	14	1	19
21	6	67,4242	38,9866	5	18	5	4
22	5	65,6262	41,1384	5	7	5	5
23	4	60,3092	47,5018	1	11	1	20
24	3	57,7545	50,5592	1	5	1	25
25	2	52,1819	57,2285	1	20	1	26
26	1	51,3426	58,2330	1	15	1	27

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	2	52,87	5,1414	5,1414
Cluster2	7	1305,82	12,8467	21,4639
Cluster3	1	0,00	0,0000	0,0000
Cluster4	1	0,00	0,0000	0,0000
Cluster5	1	0,00	0,0000	0,0000
Cluster6	1	0,00	0,0000	0,0000
Cluster7	1	0,00	0,0000	0,0000
Cluster8	1	0,00	0,0000	0,0000
Cluster9	1	0,00	0,0000	0,0000
Cluster10	1	0,00	0,0000	0,0000
Cluster11	1	0,00	0,0000	0,0000
Cluster12	1	0,00	0,0000	0,0000
Cluster13	1	0,00	0,0000	0,0000
Cluster14	1	0,00	0,0000	0,0000
Cluster15	1	0,00	0,0000	0,0000
Cluster16	1	0,00	0,0000	0,0000
Cluster17	1	0,00	0,0000	0,0000
Cluster18	1	0,00	0,0000	0,0000
Cluster19	1	0,00	0,0000	0,0000
Cluster20	1	0,00	0,0000	0,0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
MN - Happy Planet Index_1	47,2914	56,2985	34,985	40,4450
MN - GINNI	75,6654	84,4430	78,288	95,9029
MN - Social Progress Index_1	95,0474	95,2908	90,288	86,3541
MN-Access to improved sanitatio	99,4406	99,9698	100,000	99,9937
MN - International Cooperati_1	45,0180	45,0180	45,018	95,6522
MN - Democracy index_1	87,3815	86,2390	76,303	76,4218
MN The Networked Readiness Inde	78,9267	87,8459	74,346	59,4241
MN Efficiency of legal framewor	73,9614	86,6618	59,478	38,4783

Variable	Cluster5	Cluster6	Cluster7	Cluster8
MN - Happy Planet Index_1	33,831	29,809	57,745	43,2719
MN - GINNI	94,624	70,825	79,332	74,1127
MN - Social Progress Index_1	96,962	86,178	86,758	74,8332
MN-Access to improved sanitatio	100,000	94,699	100,000	98,5024
MN - International Cooperati_1	45,018	100,000	45,018	45,0180
MN - Democracy index_1	90,284	74,052	77,607	70,6161
MN The Networked Readiness Inde	85,864	79,843	75,131	45,2880
MN Efficiency of legal framewor	76,287	61,003	59,728	26,0220

Variable	Cluster9	Cluster10	Cluster11	Cluster12
MN - Happy Planet Index_1	42,380	78,689	57,3312	51,128
MN - GINNI	45,018	53,132	72,0772	82,333
MN - Social Progress Index_1	98,700	72,322	80,7165	81,279
MN-Access to improved sanitatio	100,000	100,000	45,0180	100,000
MN - International Cooperati_1	45,018	45,018	45,0180	45,018
MN - Democracy index_1	95,853	72,749	75,3555	77,844
MN The Networked Readiness Inde	80,628	83,770	51,3089	86,911
MN Efficiency of legal framewor	73,731	57,042	10,6805	43,303

Variable	Cluster13	Cluster14	Cluster15	Cluster16
MN - Happy Planet Index_1	29,6258	69,8903	69,580	42,3243
MN - GINNI	70,7463	70,3810	94,755	95,4071

MN - Social Progress Index_1	74,9912	97,7520	100,000	82,5957
MN-Access to improved sanitatio	76,4365	45,0180	100,000	99,7152
MN - International Cooperati_1	95,6522	45,0180	45,018	91,3043
MN - Democracy index_1	70,9716	92,0616	100,000	69,4313
MN The Networked Readiness Inde	61,7801	79,8429	91,099	49,7382
MN Efficiency of legal framewor	31,3553	96,6171	90,936	18,9669

Variable	Cluster17	Cluster18	Cluster19	Cluster20
MN - Happy Planet Index_1	42,426	56,898	61,128	35,5870
MN - GINNI	99,870	96,712	65,501	58,2724
MN - Social Progress Index_1	88,163	99,473	93,537	90,3231
MN-Access to improved sanitatio	100,000	100,000	100,000	99,9787
MN - International Cooperati_1	82,609	45,018	45,018	45,0180
MN - Democracy index_1	72,038	97,630	80,806	78,4360
MN The Networked Readiness Inde	62,304	97,120	86,911	88,7435
MN Efficiency of legal framewor	23,273	85,054	93,297	74,5857

Variable	Grand centroid
MN - Happy Planet Index_1	50,5832
MN - GINNI	79,2489
MN - Social Progress Index_1	90,3095
MN-Access to improved sanitatio	94,7419
MN - International Cooperati_1	53,9116
MN - Democracy index_1	82,4776
MN The Networked Readiness Inde	78,2529
MN Efficiency of legal framewor	65,7183

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,0000	20,017	23,1252	69,683	24,628	61,6414	22,4564
Cluster2	20,0168	0,000	39,2061	79,071	27,188	69,9423	32,5981
Cluster3	23,1252	39,206	0,0000	59,859	30,388	56,4635	23,1076
Cluster4	69,6825	79,071	59,8586	0,000	71,010	41,4602	61,9479
Cluster5	24,6283	27,188	30,3878	71,010	0,000	65,4472	38,2096
Cluster6	61,6414	69,942	56,4635	41,460	65,447	0,0000	62,7753
Cluster7	22,4564	32,598	23,1076	61,948	38,210	62,7753	0,0000
Cluster8	64,3342	80,184	48,1997	59,752	74,576	76,0778	49,5654
Cluster9	32,4309	45,513	43,0947	85,902	51,008	68,3361	45,9933
Cluster10	50,3405	55,545	54,5258	83,687	70,875	77,3191	37,9334
Cluster11	90,5136	103,018	81,0269	85,501	100,343	98,1242	78,0420
Cluster12	36,6568	46,678	27,9288	60,570	44,046	63,5587	22,1683
Cluster13	78,8042	91,652	66,4519	38,987	84,625	41,1384	71,7580
Cluster14	63,5908	60,044	77,5725	106,189	73,242	93,9390	70,4518
Cluster15	38,4912	22,872	58,2593	88,945	40,295	84,2172	47,7877
Cluster16	83,1296	94,533	69,5882	23,670	86,297	59,7434	70,9165
Cluster17	71,7730	82,259	58,4185	21,255	72,735	55,4308	59,9670
Cluster18	33,3803	19,666	50,3157	83,949	28,289	77,8764	44,6135
Cluster19	27,8963	21,439	46,6840	88,019	44,556	72,4771	39,0455
Cluster20	25,2765	36,689	29,0013	78,655	38,989	59,5142	36,7958

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	64,334	32,4309	50,3405	90,514	36,6568	78,804
Cluster2	80,184	45,5129	55,5449	103,018	46,6784	91,652
Cluster3	48,200	43,0947	54,5258	81,027	27,9288	66,452
Cluster4	59,752	85,9023	83,6871	85,501	60,5696	38,987
Cluster5	74,576	51,0082	70,8748	100,343	44,0457	84,625
Cluster6	76,078	68,3361	77,3191	98,124	63,5587	41,138
Cluster7	49,565	45,9933	37,9334	78,042	22,1683	71,758
Cluster8	0,000	74,7071	64,4264	58,233	47,5018	59,572
Cluster9	74,707	0,0000	53,8718	97,761	55,3363	85,298
Cluster10	64,426	53,8718	0,0000	84,371	43,7854	83,607
Cluster11	58,233	97,7614	84,3711	0,000	74,1952	70,067
Cluster12	47,502	55,3363	43,7854	74,195	0,0000	67,651
Cluster13	59,572	85,2984	83,6072	70,067	67,6513	0,000
Cluster14	103,663	70,4472	77,4548	94,494	82,9685	103,540
Cluster15	94,497	60,3171	67,2805	112,653	60,1883	105,685

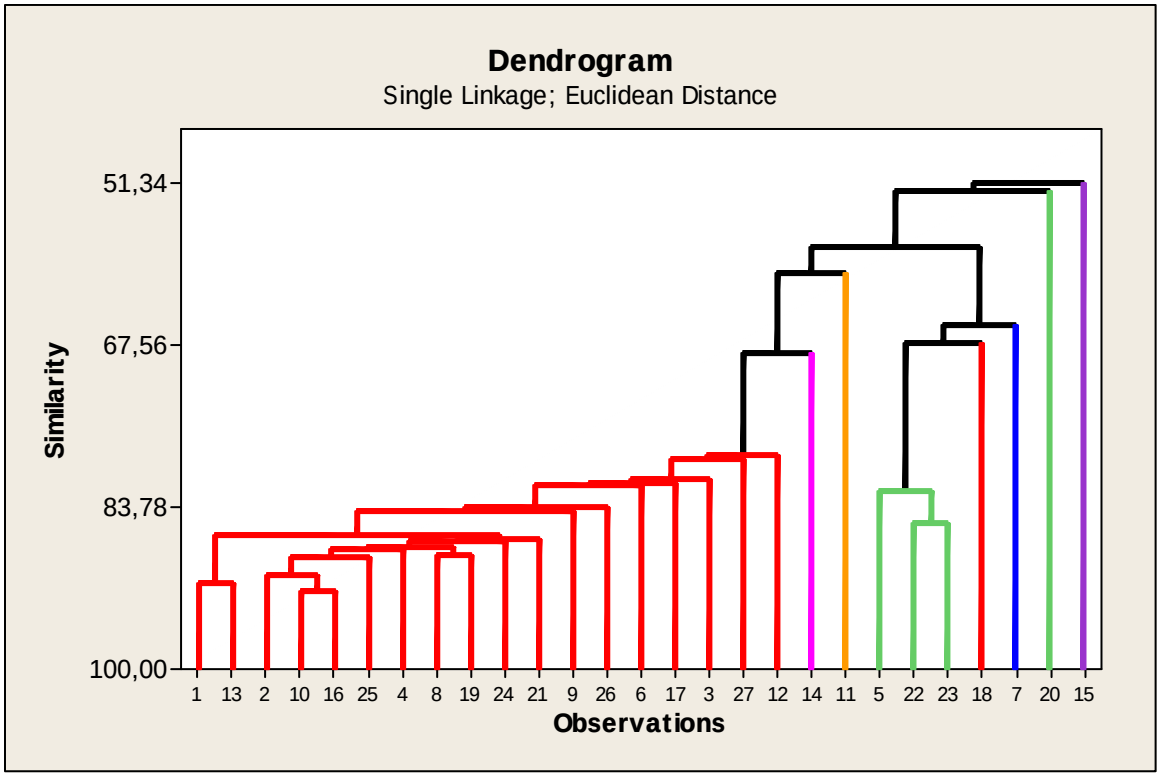
Cluster16	52,244	97,9404	89,3115	77,547	66,6146	41,096
Cluster17	50,559	89,3405	82,2504	76,004	53,7042	44,470
Cluster18	90,592	57,3324	68,5077	110,037	53,0096	99,898
Cluster19	84,305	38,0320	47,9651	106,619	55,1520	95,121
Cluster20	69,714	25,7283	50,7257	96,286	43,4113	78,714

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	63,591	38,491	83,130	71,773	33,380	27,896
Cluster2	60,044	22,872	94,533	82,259	19,666	21,439
Cluster3	77,573	58,259	69,588	58,419	50,316	46,684
Cluster4	106,189	88,945	23,670	21,255	83,949	88,019
Cluster5	73,242	40,295	86,297	72,735	28,289	44,556
Cluster6	93,939	84,217	59,743	55,431	77,876	72,477
Cluster7	70,452	47,788	70,916	59,967	44,614	39,046
Cluster8	103,663	94,497	52,244	50,559	90,592	84,305
Cluster9	70,447	60,317	97,940	89,341	57,332	38,032
Cluster10	77,455	67,281	89,311	82,250	68,508	47,965
Cluster11	94,494	112,653	77,547	76,004	110,037	106,619
Cluster12	82,969	60,188	66,615	53,704	53,010	55,152
Cluster13	103,540	105,685	41,096	44,470	99,898	95,121
Cluster14	0,000	62,002	119,156	110,634	65,965	57,698
Cluster15	62,002	0,000	104,945	92,228	15,537	36,885
Cluster16	119,156	104,945	0,000	17,602	100,246	102,695
Cluster17	110,634	92,228	17,602	0,000	86,282	92,513
Cluster18	65,965	15,537	100,246	86,282	0,000	38,503
Cluster19	57,698	36,885	102,695	92,513	38,503	0,000
Cluster20	71,760	57,602	91,228	81,042	50,633	32,772

	Cluster20
Cluster1	25,2765
Cluster2	36,6888
Cluster3	29,0013
Cluster4	78,6553
Cluster5	38,9894
Cluster6	59,5142
Cluster7	36,7958
Cluster8	69,7141
Cluster9	25,7283
Cluster10	50,7257
Cluster11	96,2863
Cluster12	43,4113
Cluster13	78,7138
Cluster14	71,7597
Cluster15	57,6024
Cluster16	91,2279
Cluster17	81,0420
Cluster18	50,6326
Cluster19	32,7719
Cluster20	0,0000

f. AVECO – VARIÁVEIS RELEVANTES E 8 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 8 clusters com as oito variáveis principais. O resultado é bastante fragmentado novamente, com um cluster com 18 observações, outro com três e os demais com uma.



Cluster Analysis of Observations: MN - Happy P; MN - GINNI; MN - Social ; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	92,1522	9,3922	10	16	10	2
2	25	91,4080	10,2829	1	13	1	2
3	24	90,6347	11,2083	2	10	2	3
4	23	88,8634	13,3282	2	25	2	4
5	22	88,5775	13,6704	8	19	8	2
6	21	88,0328	14,3223	2	4	2	5
7	20	87,7358	14,6777	2	8	2	7
8	19	87,2458	15,2641	2	24	2	8
9	18	87,0175	15,5374	2	21	2	9
10	17	86,6145	16,0197	1	2	1	11
11	16	85,2922	17,6023	22	23	22	2
12	15	84,1875	18,9244	1	9	1	12
13	14	83,7755	19,4174	1	26	1	13
14	13	82,2401	21,2550	5	22	5	3
15	12	81,6566	21,9532	1	6	1	14
16	11	81,4770	22,1683	1	17	1	15
17	10	80,9079	22,8493	1	3	1	16
18	9	79,0735	25,0448	1	27	1	17
19	8	78,5024	25,7283	1	12	1	18
20	7	68,3041	37,9334	1	14	1	19
21	6	67,4242	38,9866	5	18	5	4
22	5	65,6262	41,1384	5	7	5	5
23	4	60,3092	47,5018	1	11	1	20
24	3	57,7545	50,5592	1	5	1	25
25	2	52,1819	57,2285	1	20	1	26
26	1	51,3426	58,2330	1	15	1	27

Final Partition

Number of clusters: 8

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	18	10363,0	22,2466	38,4186
Cluster2	3	440,6	12,0366	13,8011
Cluster3	1	0,0	0,0000	0,0000
Cluster4	1	0,0	0,0000	0,0000
Cluster5	1	0,0	0,0000	0,0000
Cluster6	1	0,0	0,0000	0,0000
Cluster7	1	0,0	0,0000	0,0000
Cluster8	1	0,0	0,0000	0,0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
MN - Happy Planet Index_1	51,7741	41,7317	29,809	43,2719
MN - GINNI	79,8482	97,0598	70,825	74,1127
MN - Social Progress Index_1	94,1361	85,7043	86,178	74,8332
MN-Access to improved sanitatio	99,9249	99,9030	94,699	98,5024
MN - International Cooperati_1	45,0180	89,8551	100,000	45,0180
MN - Democracy index_1	86,2888	72,6303	74,052	70,6161
MN The Networked Readiness Inde	85,5294	57,1553	79,843	45,2880
MN Efficiency of legal framewor	78,3864	26,9060	61,003	26,0220

Variable	Cluster5	Cluster6	Cluster7	Cluster8
MN - Happy Planet Index_1	78,689	57,3312	29,6258	69,8903
MN - GINNI	53,132	72,0772	70,7463	70,3810
MN - Social Progress Index_1	72,322	80,7165	74,9912	97,7520
MN-Access to improved sanitatio	100,000	45,0180	76,4365	45,0180
MN - International Cooperati_1	45,018	45,0180	95,6522	45,0180
MN - Democracy index_1	72,749	75,3555	70,9716	92,0616
MN The Networked Readiness Inde	83,770	51,3089	61,7801	79,8429
MN Efficiency of legal framewor	57,042	10,6805	31,3553	96,6171

Variable	Grand centroid
MN - Happy Planet Index_1	50,5832
MN - GINNI	79,2489
MN - Social Progress Index_1	90,3095
MN-Access to improved sanitatio	94,7419
MN - International Cooperati_1	53,9116
MN - Democracy index_1	82,4776
MN The Networked Readiness Inde	78,2529
MN Efficiency of legal framewor	65,7183

Distances Between Cluster Centroids

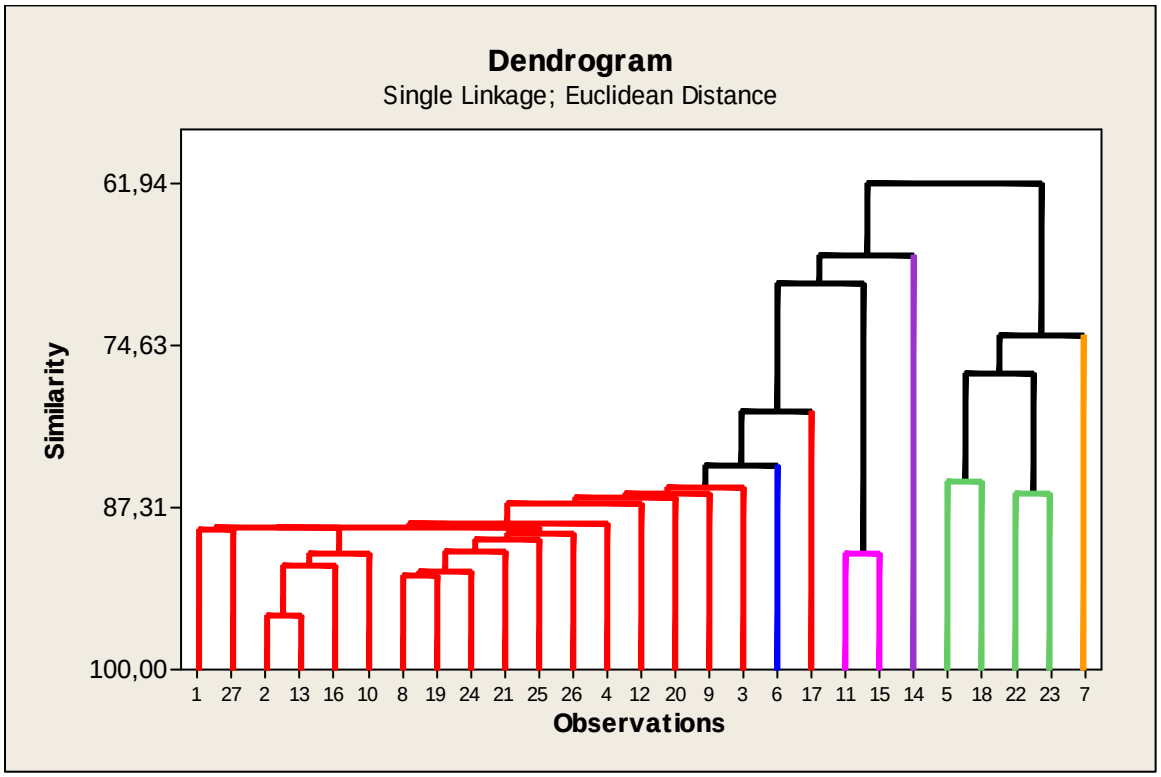
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,0000	78,233	64,5120	71,322	50,5574	95,7120	84,060
Cluster2	78,2332	0,000	51,3815	52,963	84,2705	78,8667	39,774
Cluster3	64,5120	51,381	0,0000	76,078	77,3191	98,1242	41,138
Cluster4	71,3221	52,963	76,0778	0,000	64,4264	58,2330	59,572
Cluster5	50,5574	84,271	77,3191	64,426	0,0000	84,3711	83,607
Cluster6	95,7120	78,867	98,1242	58,233	84,3711	0,0000	70,067
Cluster7	84,0596	39,774	41,1384	59,572	83,6072	70,0671	0,000
Cluster8	61,9975	111,465	93,9390	103,663	77,4548	94,4944	103,540

	Cluster8
Cluster1	61,997
Cluster2	111,465
Cluster3	93,939
Cluster4	103,663

```
Cluster5    77,455
Cluster6    94,494
Cluster7   103,540
Cluster8     0,000
```

g. AVECO – COMPONENTES PRINCIPAIS E 8 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 8 clusters com as quatro componentes principais. O resultado é semelhante, com um cluster com 17 observações, trÊs com duas e os demais com uma observação.



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

```
Euclidean Distance, Single Linkage
Amalgamation Steps
```

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	95,7760	0,24606	2	13	2	2
2	25	92,6516	0,42806	8	19	8	2
3	24	92,2863	0,44934	8	24	8	3
4	23	91,9185	0,47077	2	16	2	3
5	22	90,9387	0,52784	2	10	2	4
6	21	90,8972	0,53026	11	15	11	2
7	20	90,7669	0,53785	8	21	8	4
8	19	89,7751	0,59563	8	25	8	5
9	18	89,3659	0,61946	8	26	8	6
10	17	88,9619	0,64300	1	27	1	2
11	16	88,9147	0,64574	1	2	1	6
12	15	88,8430	0,64992	1	8	1	12

13	14	88,5746	0,66556	1	4	1	13
14	13	86,9326	0,76121	1	12	1	14
15	12	86,5649	0,78263	1	20	1	15
16	11	86,2228	0,80255	22	23	22	2
17	10	86,2209	0,80267	1	9	1	16
18	9	85,7293	0,83130	1	3	1	17
19	8	85,3085	0,85581	5	18	5	2
20	7	84,0108	0,93141	1	6	1	18
21	6	79,7680	1,17856	1	17	1	19
22	5	76,7572	1,35395	5	22	5	4
23	4	73,9138	1,51958	5	7	5	5
24	3	69,8125	1,75849	1	11	1	21
25	2	67,5705	1,88909	1	14	1	22
26	1	61,9427	2,21693	1	5	1	27

Final Partition
Number of clusters: 8

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	17	14,2941	0,855552	1,66834
Cluster2	2	0,3662	0,427906	0,42791
Cluster3	1	0,0000	0,000000	0,00000
Cluster4	1	0,0000	0,000000	0,00000
Cluster5	2	0,1406	0,265129	0,26513
Cluster6	1	0,0000	0,000000	0,00000
Cluster7	1	0,0000	0,000000	0,00000
Cluster8	2	0,3220	0,401277	0,40128

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	5,08734	2,65780	5,49338	3,71464	2,20421	3,45616	3,27638
PC2	1,07651	0,44279	0,87790	1,46393	-1,51016	0,02768	-0,45180
PC3	-1,11477	0,23787	-1,85318	0,23237	0,16604	0,53731	-0,63419
PC4	-0,79859	2,50692	0,10378	1,83555	0,88215	-1,39443	0,15194

Variable	Cluster8	Grand centroid
PC1	3,24745	4,39423
PC2	-0,91402	0,60206
PC3	0,20024	-0,72076
PC4	2,75716	-0,02173

Distances Between Cluster Centroids

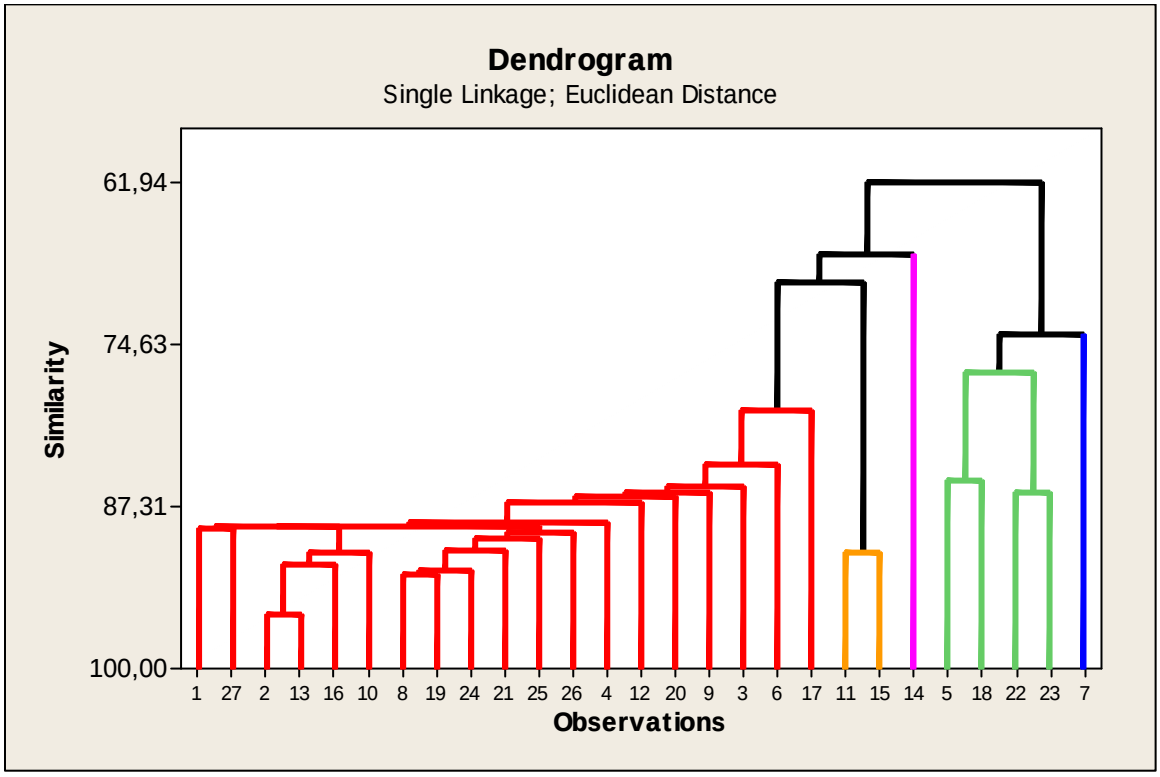
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,00000	4,36580	1,25053	3,28449	4,41232	2,61633	2,59803
Cluster2	4,36580	0,00000	4,28688	1,61567	2,58163	4,01496	2,73667
Cluster3	1,25053	4,28688	0,00000	3,29483	4,60484	3,58220	2,85858
Cluster4	3,28449	1,61567	3,29483	0,00000	3,46986	3,55744	2,72902
Cluster5	4,41232	2,58163	4,60484	3,46986	0,00000	3,04187	1,85559
Cluster6	2,61633	4,01496	3,58220	3,55744	3,04187	0,00000	2,00647
Cluster7	2,59803	2,73667	2,85858	2,72902	1,85559	2,00647	0,00000
Cluster8	4,66047	1,50089	4,41725	2,59293	2,22723	4,27548	2,77451

	Cluster8
Cluster1	4,66047
Cluster2	1,50089
Cluster3	4,41725
Cluster4	2,59293
Cluster5	2,22723
Cluster6	4,27548

```
Cluster7   2,77451
Cluster8   0,00000
```

h. AVECO – COMPONENTES PRINCIPAIS E 5 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 5 clusters com as quatro componentes principais. O resultado continua bastante fragmentado, com um cluster com 19 observações, outro com quatro, um com dois e dois com uma observação.



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	95,7760	0,24606	2	13	2	2
2	25	92,6516	0,42806	8	19	8	2
3	24	92,2863	0,44934	8	24	8	3
4	23	91,9185	0,47077	2	16	2	3
5	22	90,9387	0,52784	2	10	2	4
6	21	90,8972	0,53026	11	15	11	2
7	20	90,7669	0,53785	8	21	8	4
8	19	89,7751	0,59563	8	25	8	5
9	18	89,3659	0,61946	8	26	8	6
10	17	88,9619	0,64300	1	27	1	2
11	16	88,9147	0,64574	1	2	1	6
12	15	88,8430	0,64992	1	8	1	12
13	14	88,5746	0,66556	1	4	1	13
14	13	86,9326	0,76121	1	12	1	14
15	12	86,5649	0,78263	1	20	1	15

16	11	86,2228	0,80255	22	23	22	2
17	10	86,2209	0,80267	1	9	1	16
18	9	85,7293	0,83130	1	3	1	17
19	8	85,3085	0,85581	5	18	5	2
20	7	84,0108	0,93141	1	6	1	18
21	6	79,7680	1,17856	1	17	1	19
22	5	76,7572	1,35395	5	22	5	4
23	4	73,9138	1,51958	5	7	5	5
24	3	69,8125	1,75849	1	11	1	21
25	2	67,5705	1,88909	1	14	1	22
26	1	61,9427	2,21693	1	5	1	27

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	19	22,1626	0,970909	2,46074
Cluster2	4	2,9409	0,849794	1,01004
Cluster3	1	0,0000	0,000000	0,00000
Cluster4	2	0,1406	0,265129	0,26513
Cluster5	1	0,0000	0,000000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
PC1	5,01340	2,95263	3,71464	2,20421	3,45616	4,39423
PC2	0,98562	-0,23561	1,46393	-1,51016	0,02768	0,60206
PC3	-1,12834	0,21905	0,23237	0,16604	0,53731	-0,72076
PC4	-0,70107	2,63204	1,83555	0,88215	-1,39443	-0,02173

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,00000	4,32009	3,19398	4,27813	2,56861
Cluster2	4,32009	0,00000	2,02575	2,29118	4,07880
Cluster3	3,19398	2,02575	0,00000	3,46986	3,55744
Cluster4	4,27813	2,29118	3,46986	0,00000	3,04187
Cluster5	2,56861	4,07880	3,55744	3,04187	0,00000

i. AVECO – ANÁLISE GRÁFICA COMPONENTES PRINCIPAIS E 5 CLUSTERS

Primeiramente realiza-se agrupamento das observações em 5 clusters com as quatro componentes principais. O resultado continua bastante fragmentado, com um cluster com 14 observações, outros três com uma e outro com duas observações.

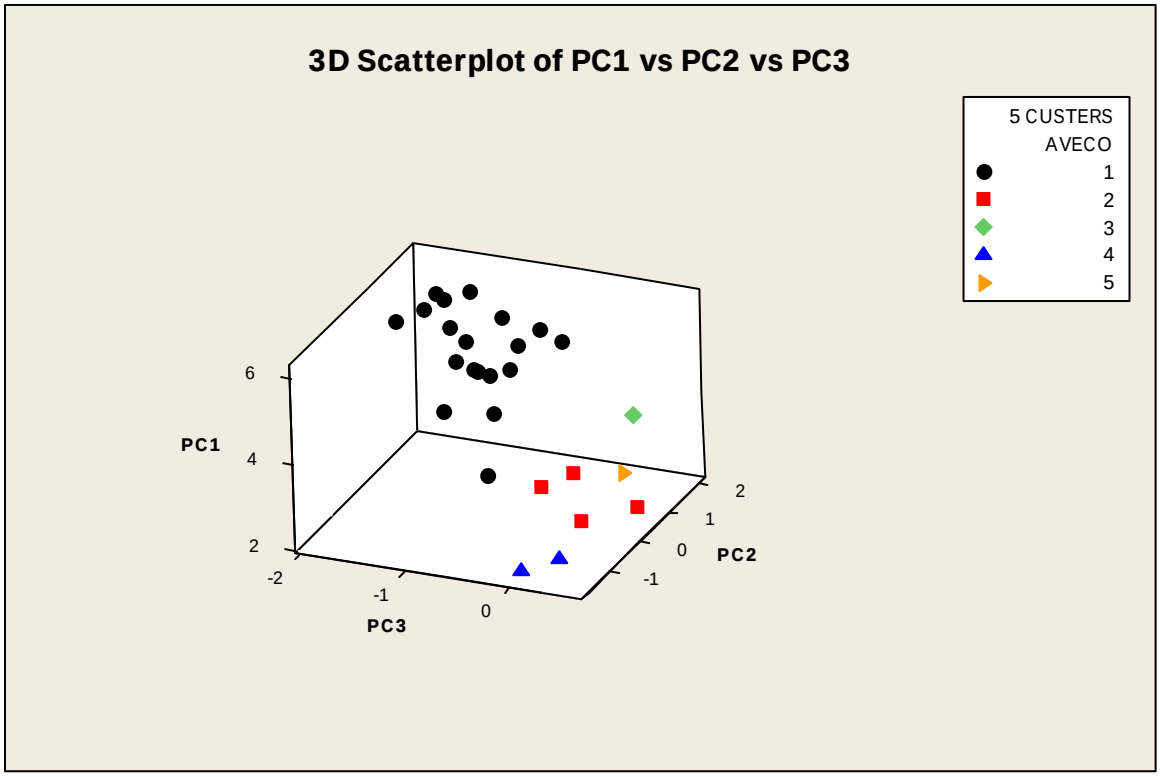


Tabela Países x Clusters x Região

PAISES	CLUSTER3R	20 CLUSTERS SEM REGIÃO	8 CLUSTERS AIBER
PAISES		**20 CLUSTER**	**5 CLUSTERS AVECO**
Australia	**AVECO**	**3**	**1**
Austria	**AVECO**	**3**	**1**
Belgium	**AVECO**	**3**	**1**
Canada	**AVECO**	**3**	**1**
Czech Rep.c	**AVECO**	**1**	**2**
Denmark	**AVECO**	**3**	**1**
Estonia	**AVECO**	**8**	**3**
Finland	**AVECO**	**3**	**1**
France	**AVECO**	**3**	**1**
Germany	**AVECO**	**3**	**1**
Greece	**AVECO**	**9**	**4**
Iceland	**AVECO**	**3**	**1**
Ireland	**AVECO**	**3**	**1**
Israel	**AVECO**	**10**	**5**
Italy	**AVECO**	**9**	**4**
Japan	**AVECO**	**3**	**1**
Korea, Rep.	**AVECO**	**3**	**1**
Latvia	**AVECO**	**1**	**2**
Netherlands	**AVECO**	**3**	**1**

New Zealand	**AVECO**	**3**	**1**
Norway	**AVECO**	**3**	**1**
Slovakia	**AVECO**	**1**	**2**
Slovenia	**AVECO**	**1**	**2**
Sweden	**AVECO**	**3**	**1**
Switzerland	**AVECO**	**3**	**1**
UK	**AVECO**	**3**	**1**
US	**AVECO**	**3**	**1**

Observamos pela tabela, que traduz os dados do gráfico em quais países compõe os clusters, que Israel e Estonia, cada um destes países, compõe um cluster separado, Grécia e Itália outro, países da Europa oriental e central, outro (Rep. Checa, Eslovenia, Eslováquia e Latvia). Todos os demais países são agrupados em um mesmo cluster sob o ponto de vista das três variáveis principais.

j. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo aplicar a análise de clusters de variáveis, buscando parcimônia no entendimento desta dimensão, associada ao agrupamento de variáveis do estudo de componentes principais para construção de novas variáveis com maior força explicativa.

Para os agrupamentos AIBER e AVECO, chega-se então a cinco grupos de países estruturados a partir dos três mais relevantes componentes principais.

9. ANÁLISE DISCRIMINANTE

Este trabalho tem como objetivo realizar uma análise dos agrupamentos de observações das principais variáveis que compõe as novas quatro Componentes Principais para o EIXO PESSOAS para 132 países, bem como estas últimas, **e propor modelos que auxiliem no seu entendimento**. Seis das variáveis originais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Todas as variáveis originais foram normalizadas e as observações faltantes foram substituídas pela média das observações não faltantes. Empregou-se o Minitab ® para as análises estatísticas.

a. AGRUPAMENTO DAS OBSERVAÇÕES COM AS QUATRO COMPONENTES PRINCIPAIS EMPREGANDO EM CLUSTERS ATRAVÉS DO MÉTODO DE LIGAÇÃO COMPLETA E DISTANCIA POR MANHATTAN - AIBER, AVECO E OTHERS

Aplica-se o método de construção de clusters por ligação completa e distancia Manhattan, definindo-se como resultado esperado quatro *clusters*. Estes quando gerados possuem respectivamente 32, 42, 17, 41 membros.

Number

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	of obs. in new cluster
1	131	98,3366	0,2661	39	90	39	2
2	130	97,5034	0,3993	95	105	95	2
3	129	97,1946	0,4487	23	34	23	2
4	128	96,3992	0,5759	2	77	2	2
5	127	96,2941	0,5927	60	89	60	2
6	126	96,0821	0,6267	52	57	52	2
7	125	95,7877	0,6737	40	45	40	2
8	124	95,6475	0,6962	84	124	84	2
9	123	95,4382	0,7296	25	29	25	2
10	122	95,3655	0,7413	23	37	23	3
11	121	95,0557	0,7908	8	17	8	2
12	120	94,8783	0,8192	88	112	88	2
13	119	94,7706	0,8364	32	36	32	2
14	118	94,2537	0,9191	18	19	18	2
15	117	94,2331	0,9224	67	118	67	2
16	116	94,0536	0,9511	41	47	41	2
17	115	93,7897	0,9933	62	69	62	2
18	114	93,7640	0,9974	18	38	18	3
19	113	93,6221	1,0201	93	102	93	2
20	112	93,5458	1,0323	53	82	53	2
21	111	93,5125	1,0377	125	126	125	2
22	110	93,2364	1,0818	22	48	22	2
23	109	93,1927	1,0888	42	46	42	2
24	108	93,1605	1,0940	64	123	64	2
25	107	92,9452	1,1284	10	80	10	2
26	106	92,5524	1,1912	121	127	121	2
27	105	92,2354	1,2419	51	96	51	2
28	104	92,1778	1,2511	92	104	92	2
29	103	92,0506	1,2715	78	120	78	2
30	102	91,5453	1,3523	2	108	2	3
31	101	91,5152	1,3571	59	68	59	2
32	100	91,3499	1,3835	39	109	39	3
33	99	91,1214	1,4201	5	13	5	2
34	98	91,1172	1,4208	81	125	81	3
35	97	91,1097	1,4220	31	42	31	3
36	96	91,1058	1,4226	84	91	84	3
37	95	90,8008	1,4714	56	60	56	3
38	94	90,6368	1,4976	50	130	50	2
39	93	90,6308	1,4986	43	44	43	2
40	92	90,4237	1,5317	22	23	22	5
41	91	90,2310	1,5625	24	30	24	2
42	90	90,0629	1,5894	87	98	87	2
43	89	89,9492	1,6076	66	101	66	2
44	88	89,9431	1,6086	14	74	14	2
45	87	89,8894	1,6172	54	119	54	2
46	86	89,8195	1,6283	70	128	70	2
47	85	89,7809	1,6345	26	39	26	4
48	84	89,3638	1,7012	33	41	33	3
49	83	89,2383	1,7213	11	12	11	2
50	82	88,9768	1,7631	31	40	31	5
51	81	88,8163	1,7888	72	114	72	2
52	80	88,7287	1,8028	63	67	63	3
53	79	88,6910	1,8088	50	79	50	3
54	78	88,6571	1,8142	49	81	49	4
55	77	88,6149	1,8210	85	87	85	3
56	76	88,4216	1,8519	52	53	52	4
57	75	88,2662	1,8768	106	121	106	3
58	74	87,9427	1,9285	8	10	8	4
59	73	87,7678	1,9565	88	132	88	3
60	72	87,5975	1,9837	7	9	7	2
61	71	87,5676	1,9885	94	97	94	2
62	70	87,5619	1,9894	78	111	78	3
63	69	87,5194	1,9962	25	31	25	7
64	68	87,3118	2,0294	56	64	56	5
65	67	87,0780	2,0668	2	117	2	4
66	66	87,0400	2,0729	59	115	59	3
67	65	86,8752	2,0993	62	63	62	5

68	64	86,6521	2,1349	5	15	5	3
69	63	86,5643	2,1490	76	83	76	2
70	62	86,2414	2,2006	32	75	32	3
71	61	85,3780	2,3387	51	61	51	3
72	60	85,1477	2,3756	71	100	71	2
73	59	84,6615	2,4533	66	122	66	3
74	58	84,5402	2,4727	11	14	11	4
75	57	84,0592	2,5497	84	110	84	4
76	56	83,6432	2,6162	56	93	56	7
77	55	83,4879	2,6410	20	28	20	2
78	54	83,3577	2,6619	58	103	58	2
79	53	83,2796	2,6744	95	106	95	5
80	52	83,1462	2,6957	3	5	3	4
81	51	83,0880	2,7050	1	49	1	5
82	50	82,9026	2,7347	72	76	72	4
83	49	82,8825	2,7379	54	92	54	4
84	48	82,5941	2,7840	25	27	25	8
85	47	82,4601	2,8054	22	24	22	7
86	46	81,9693	2,8839	51	73	51	4
87	45	80,7396	3,0806	8	16	8	5
88	44	80,5745	3,1070	52	70	52	6
89	43	80,3769	3,1386	4	6	4	2
90	42	80,1152	3,1805	86	107	86	2
91	41	79,9961	3,1995	18	129	18	4
92	40	79,2541	3,3182	26	43	26	6
93	39	78,1247	3,4989	54	86	54	6
94	38	77,3943	3,6157	55	113	55	2
95	37	76,8280	3,7063	66	85	66	6
96	36	76,6365	3,7369	56	95	56	12
97	35	76,5322	3,7536	51	131	51	5
98	34	76,2232	3,8030	7	21	7	3
99	33	76,1730	3,8110	71	94	71	4
100	32	76,1592	3,8132	55	78	55	5
101	31	75,3901	3,9363	32	59	32	6
102	30	75,0909	3,9841	52	84	52	10
103	29	74,9153	4,0122	62	65	62	6
104	28	73,9986	4,1588	50	66	50	9
105	27	73,9275	4,1702	72	88	72	7
106	26	73,6135	4,2204	2	116	2	5
107	25	73,4743	4,2427	25	33	25	11
108	24	72,7076	4,3653	18	35	18	5
109	23	70,3497	4,7424	22	25	22	18
110	22	69,8127	4,8283	26	32	26	12
111	21	69,6375	4,8564	8	11	8	9
112	20	68,5897	5,0240	1	50	1	14
113	19	67,7141	5,1640	51	56	51	17
114	18	67,0742	5,2663	18	71	18	9
115	17	66,2989	5,3903	3	4	3	6
116	16	64,5632	5,6680	2	99	2	6
117	15	63,5814	5,8250	20	26	20	14
118	14	63,3631	5,8599	1	7	1	17
119	13	62,1488	6,0541	51	62	51	23
120	12	61,7067	6,1249	8	58	8	11
121	11	55,8017	7,0693	2	72	2	13
122	10	54,9789	7,2009	1	55	1	22
123	9	52,5828	7,5842	18	20	18	23
124	8	51,4600	7,7638	51	54	51	29
125	7	46,0717	8,6256	1	52	1	32
126	6	43,2853	9,0713	3	8	3	17
127	5	34,4175	10,4897	2	51	2	42
128	4	23,3703	12,2566	18	22	18	41
129	3	16,7954	13,3082	1	3	1	49
130	2	9,4453	14,4839	1	18	1	90
131	1	0,0000	15,9946	1	2	1	132

Final Partition
Number of clusters: 4

	Average	Maximum

		Number of observations	Within cluster sum of squares	distance from centroid	distance from centroid
Cluster1	VERMELHO	32	83,511	1,53313	2,73551
Cluster2	LARANJA	42	168,697	1,87297	3,75273
Cluster3	AZUL	17	56,251	1,65681	3,27974
Cluster4	VERDE	41	233,455	2,29234	3,54600

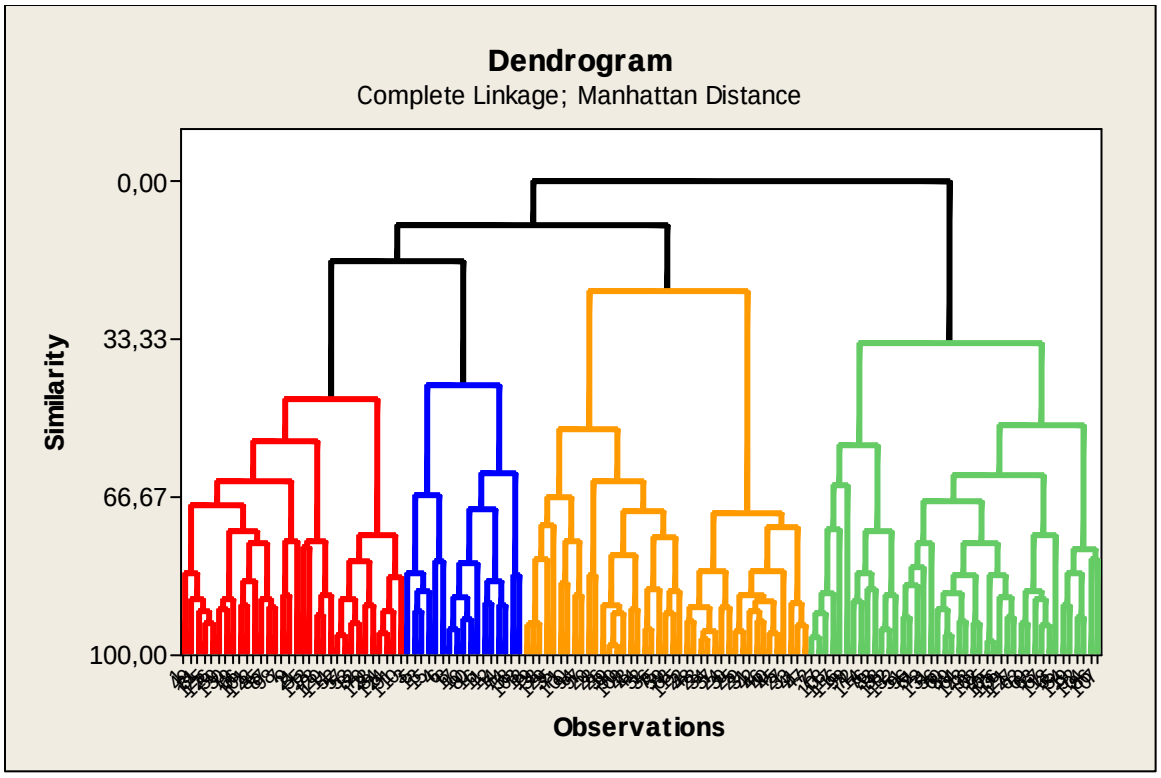
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
PC1	0,16468	-3,64920	-0,04299	3,62750	0,0000000
PC2	-1,52155	0,62389	0,43323	0,36882	-0,0000000
PC3	-0,14385	-0,44082	2,27974	-0,38141	-0,0000000
PC4	-0,25939	0,00520	-0,39646	0,36151	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,00000	4,39396	3,12360	4,00082
Cluster2	4,39396	0,00000	4,53915	7,29013
Cluster3	3,12360	4,53915	0,00000	4,59705
Cluster4	4,00082	7,29013	4,59705	0,00000

Dendrogram



i. OS 4 GRUPOS RESULTANTES

Os grupos resultantes são agrupados por maior similaridade. Chega-se a algumas caracterizações. Nos grupos 1 e 2 predominam países de OUTRAS Regiões (OTHERS),

sendo que no primeiro enquadram-se os países mais desenvolvidos de OTHERS, acrescido de alguns países de AIBER, e no grupo 2 os menos desenvolvidos. No grupo 3 encontra-se a maioria dos países de AIBER acrescidos de alguns países de OTHERS. E no grupo 4 encontram-se TODOS os países de AVECO, acrescidos de alguns de AIBER (Portugal, Espanha e Uruguai) e outros de OTHERS.

4 CLUSTERS MANHATTAN							
Grupo 1 - 32 países		**Grupo 2 - 42 países**		**Grupo 3 - 17 países**		**Grupo 4 - 41 países**	
País	**Região**	**País**	**Região**	**País**	**Região**	**País**	**Região**
Argentina	AIBER	Bolivia	AIBER	Brazil	AIBER	Portugal	AIBER
Cuba	AIBER	Angola	OTHERS	Chile	AIBER	Spain	AIBER
Ecuador	AIBER	Bangladesh	OTHERS	Colombia	AIBER	Uruguay	AIBER
Venezuela	AIBER	Benin	OTHERS	Costa Rica	AIBER	Australia	AVECO
Albania	OTHERS	Burkina Faso	OTHERS	Dominican Republic	AIBER	Austria	AVECO
Algeria	OTHERS	Burundi	OTHERS	El Salvador	AIBER	Belgium	AVECO
Armenia	OTHERS	Cambodia	OTHERS	Guatemala	AIBER	Canada	AVECO
Azerbaijan	OTHERS	Cameroon	OTHERS	Honduras	AIBER	Czech Republic	AVECO
Belarus	OTHERS	Central African Republic	OTHERS	Mexico	AIBER	Denmark	AVECO
Bosnia and Herzegovina	OTHERS	Chad	OTHERS	Nicaragua	AIBER	Estonia	AVECO
China	OTHERS	Congo, Republic of	OTHERS	Panama	AIBER	Finland	AVECO
Egypt	OTHERS	Djibouti	OTHERS	Paraguay	AIBER	France	AVECO
Iran	OTHERS	Ghana	OTHERS	Peru	AIBER	Germany	AVECO
Iraq	OTHERS	Guinea	OTHERS	Botswana	OTHERS	Greece	AVECO
Jordan	OTHERS	India	OTHERS	Guyana	OTHERS	Iceland	AVECO
Kazakhstan	OTHERS	Indonesia	OTHERS	Jamaica	OTHERS	Ireland	AVECO
Kuwait	OTHERS	Kenya	OTHERS	Namibia	OTHERS	Israel	AVECO
Kyrgyzstan	OTHERS	Laos	OTHERS			Italy	AVECO
Lebanon	OTHERS	Lesotho	OTHERS			Japan	AVECO
Macedonia	OTHERS	Liberia	OTHERS			Korea, Republic of	AVECO
Moldova	OTHERS	Madagascar	OTHERS			Latvia	AVECO
Morocco	OTHERS	Malawi	OTHERS			Netherlands	AVECO
Romania	OTHERS	Mali	OTHERS			New Zealand	AVECO
Russia	OTHERS	Mauritania	OTHERS			Norway	AVECO
Saudi Arabia	OTHERS	Mongolia	OTHERS			Slovakia	AVECO
Tajikistan	OTHERS	Mozambique	OTHERS			Slovenia	AVECO
Thailand	OTHERS	Nepal	OTHERS			Sweden	AVECO
Trinidad and Tobago	OTHERS	Niger	OTHERS			Switzerland	AVECO

Tunisia	OTHERS	Nigeria	OTHERS		United Kingdom	AVECO
Turkey	OTHERS	Pakistan	OTHERS		United States	AVECO
Ukraine	OTHERS	Philippines	OTHERS		Bulgaria	OTHERS
Uzbekistan	OTHERS	Rwanda	OTHERS		Croatia	OTHERS
		Senegal	OTHERS		Georgia	OTHERS
		South Africa	OTHERS		Hungary	OTHERS
		Sri Lanka	OTHERS		Lithuania	OTHERS
		Sudan	OTHERS		Malaysia	OTHERS
		Swaziland	OTHERS		Mauritius	OTHERS
		Tanzania	OTHERS		Montenegro	OTHERS
		Togo	OTHERS		Poland	OTHERS
		Uganda	OTHERS		Serbia	OTHERS
		Yemen	OTHERS		United Arab Emirates	OTHERS
		Zambia	OTHERS			

b. ANÁLISE DISCRIMINANTE

Procede-se nesta etapa ao teste de melhor adequação das funções ao dado disponível, de modo a obter-se melhor previsibilidade sobre o . A variável dependente é o pertencimento a cada um dos 4 *clusters* gerados na etapa anterior. E as variáveis independentes utilizadas são as oito variáveis relevantes constituintes dos componentes principais:
Social Progress Index
Access to improved sanitation
Network Readiness
Democracy índex
Efficiency of the legal framework to settle disputes
GINNI
Happy Planet Index
International Cooperation

Estas foram extraídas das equações obtidas anteriormente:

PC1 - INDICADORES DE PROGRESSO E DESENVLVIMENTO HUMANO.

```
PC1 = - 8,02 + 0,0333 MN -  (HDI)_1 + 0,0639 MN - Social Progress Index_1
      + 0,0138 MN-Access to improved sanitatio
      + 0,0274 MN The Networked Readiness Inde
```

PC2 - RESOLUTIVIDADE LEGAL, RECURSOS SANITÁRIOS E DEMOCRACIA.

```
The regression equation is
PC2 = - 1,09 - 0,0274 MN-Access to improved sanitatio
      + 0,0280 MN - Democracy index_1 + 0,0323 MN Efficiency of legal framewor
```

PC3 - FELICIDADE, COOPERAÇÃO E EQUIDADE.

```
The regression equation is
PC3 = - 1,72 - 0,0251 MN - GINNI + 0,0323 MN - Happy Planet Index_1
      + 0,0294 MN -  International Cooperati_1
```

PC4 - COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL.

```
The regression equation is
PC4 = - 2,60 + 0,0233 MN - GINNI + 0,0389 MN -  International Cooperati_1
      - 0,0219 MN Efficiency of legal framewor
```

São testadas funções linear e quadrática.

i. TESTE DA FUNÇÃO LINEAR

A maior distância entre grupos, medida como distancia quadrática, é encontrada entre os grupos 2 e 4 (33,25). As demais distâncias ficam entre 11,57 (1 e 4), 13,43 (1 e 3), 16,66(1 e 2), 18,08 (2 e 3), e 18,16(3 e 4).

Squared Distance Between Groups

	1	2	3	4
1	0,0000	16,6602	13,4267	11,5721
2	16,6602	0,0000	18,0818	33,2466
3	13,4267	18,0818	0,0000	18,1596
4	11,5721	33,2466	18,1596	0,0000

O teste da função linear obteve 92,4% de acertos em predizer a classificação dos países, ou 122 acertos e dez erros de classificação dentre as 132 observações (países).

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
ARGENTINA 1**	1	3	1	11,42	0,374
			2	34,72	0,000
			3	10,59	0,568
			4	15,17	0,058
EQUADOR 9**	1	3	1	5,808	0,475
			2	19,376	0,001
			3	5,636	0,517
			4	14,071	0,008
BANGLADESH54	2	3	1	17,749	0,007
			2	9,785	0,360
			3	8,658	0,633
			4	35,303	0,000
GEORGIA 71**	4	1	1	3,041	0,777
			2	17,344	0,001
			3	6,181	0,162
			4	8,158	0,060
GUYANA 74**	3	1	1	5,689	0,638
			2	12,783	0,018
			3	6,939	0,342
			4	17,944	0,001
MACEDONIA91**	1	4	1	10,242	0,380
			2	23,232	0,001
			3	19,696	0,003

			4	9,278	0,616
MONTENEGRO **	4	3	1	9,569	0,133
			2	23,234	0,000
			3	6,425	0,640
			4	8,499	0,227
NAMIBIA 103**	3	2	1	35,79	0,000
			2	15,87	0,567
			3	16,41	0,432
			4	43,22	0,000
FILIPINAS08**	2	3	1	8,062	0,051
			2	11,652	0,009
			3	2,255	0,937
			4	13,707	0,003
SRI LANKA17**	2	1	1	8,589	0,717
			2	14,313	0,041
			3	10,771	0,241
			4	22,376	0,001

Conforme os quadros abaixo, do grupo 1, com 32 países, 29 foram corretamente classificados e os três restantes foram classificados dois no grupo 3 um no grupo 4. Para o grupo 2, 39 acertos e três erros de classificação, 1 no grupo 1 e dois no grupo 3. Para o grupo 3, 15 acertos e dois erros, um para o grupo 1 e outro no grupo 2. E grupo 4, 39 acertos e dois erros, 1 no grupo 1 e outro no grupo 3.

Discriminant Analysis: 4 CLUSTERS M versus MN - Social ; MN-Access to; ...

Linear Method for Response: 4 CLUSTERS MANHATTAN

Predictors: MN - Social Progress Index_1; MN-Access to improved sanitatio;
MN The Networked Readiness Inde; MN - Democracy index_1;
MN Efficiency of legal framewor; MN - GINNI;
MN - Happy Planet Index_1; MN - International Cooperati_1

Group	1	2	3	4
Count	32	42	17	41

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	29	1	1	1
2	0	39	1	0
3	2	2	15	1
4	1	0	0	39
Total N	32	42	17	41
N correct	29	39	15	39
Proportion	0,906	0,929	0,882	0,951

N = 132 N Correct = 122 Proportion Correct = 0,924

Squared Distance Between Groups

	1	2	3	4
1	0,0000	16,6602	13,4267	11,5721
2	16,6602	0,0000	18,0818	33,2466
3	13,4267	18,0818	0,0000	18,1596
4	11,5721	33,2466	18,1596	0,0000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-34,404	-20,489	-36,334	-51,704
MN - Social Progress Index_1	0,359	0,126	0,374	0,516
MN-Access to improved sanitatio	0,180	0,006	0,095	0,157
MN The Networked Readiness Inde	0,016	-0,019	0,048	0,157
MN - Democracy index_1	-0,139	-0,031	-0,065	-0,055
MN Efficiency of legal framewor	0,097	0,184	0,059	0,032
MN - GINNI	0,250	0,229	0,103	0,242
MN - Happy Planet Index_1	0,114	0,113	0,208	0,062
MN - International Cooperati_1	0,204	0,242	0,310	0,245

ii. TESTE DA FUNÇÃO QUADRÁTICA

A maior distância entre grupos, medida como distancia quadrática, é encontrada entre os grupos 2 e 4 (127,20), ou seja, entre os menos e os mais desenvolvidos dos grupos.

From Generalized Squared Distance to Group

Group	1	2	3	4
1	39,42	53,66	96,69	60,64
2	89,01	42,27	72,45	127,20
3	52,26	59,25	34,71	59,97
4	66,06	70,92	106,93	39,48

O teste da função linear retorna um resultado ainda melhor do que o obtido pela Linear, com 96,2% de acertos contra 92,4%, para predizer a classificação dos países, ou 127 acertos e 5 erros de classificação dentre as 132 observações (países).

Conforme os quadros abaixo, do grupo 1, com 32 países, 31 foram corretamente classificados e um classificado no grupo 2. Para o grupo 2, 41 acertos e um erro de classificação, no grupo 3. Para o grupo 3, nenhum erro, e para o grupo 4, três erros, incluídos no grupo 1.

Os países que foram classificados de forma equivocada no modelo linear foram:

Discriminant Analysis: 4 CLUSTERS M versus MN - Social ; MN-Access to; ...

Quadratic Method for Response: 4 CLUSTERS MANHATTAN

Predictors: MN - Social Progress Index_1; MN-Access to improved sanitatio;
MN The Networked Readiness Inde; MN - Democracy index_1;
MN Efficiency of legal framewor; MN - GINNI;
MN - Happy Planet Index_1; MN - International Cooperati_1

Group	1	2	3	4
Count	32	42	17	41

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	31	0	0	3
2	1	41	0	0
3	0	1	17	0

4	0	0	0	38
Total N	32	42	17	41
N correct	31	41	17	38
Proportion	0,969	0,976	1,000	0,927

N = 132 N Correct = 127 Proportion Correct = 0,962

From Generalized Squared Distance to Group

Group	1	2	3	4
1	39,42	53,66	96,69	60,64
2	89,01	42,27	72,45	127,20
3	52,26	59,25	34,71	59,97
4	66,06	70,92	106,93	39,48

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
CHINA 66**	1	2	1	51,77	0,361
			2	50,63	0,639
			3	154,22	0,000
			4	77,91	0,000
GEORGIA 71**	4	1	1	42,38	0,948
			2	54,83	0,002
			3	61,82	0,000
			4	48,25	0,050
FILIPINAS08**	2	3	1	48,15	0,016
			2	50,56	0,005
			3	39,96	0,979
			4	56,51	0,000
SERVIA 115**	4	1	1	46,27	0,536
			2	71,66	0,000
			3	79,76	0,000
			4	46,56	0,464
UAE 129**	4	1	1	64,72	0,967
			2	85,39	0,000
			3	431,85	0,000
			4	71,46	0,033

Discriminant Analysis: 3R CLUSTERS versus MN - Social ; MN-Access to; ...

Linear Method for Response: 3R CLUSTERS

Predictors: MN - Social Progress Index_1; MN-Access to improved sanitatio;
MN The Networked Readiness Inde; MN - Democracy index_1;
MN Efficiency of legal framewor; MN - GINNI;
MN - Happy Planet Index_1; MN - International Cooperati_1

Group	AIBER	AVECO	OTHERS
Count	21	27	84

Summary of classification

	True Group		
Put into Group	AIBER	AVECO	OTHERS
AIBER	19	0	12
AVECO	2	26	1
OTHERS	0	1	71
Total N	21	27	84
N correct	19	26	71

Proportion 0,905 0,963 0,845

N = 132 N Correct = 116 Proportion Correct = 0,879

Squared Distance Between Groups

	AIBER	AVECO	OTHERS
AIBER	0,0000	15,1869	7,1121
AVECO	15,1869	0,0000	14,5139
OTHERS	7,1121	14,5139	0,0000

Linear Discriminant Function for Groups

	AIBER	AVECO	OTHERS
Constant	-24,343	-37,927	-21,030
MN - Social Progress Index_1	0,194	0,210	0,053
MN-Access to improved sanitatio	-0,054	-0,106	-0,035
MN The Networked Readiness Inde	-0,031	0,104	-0,033
MN - Democracy index_1	-0,005	0,086	-0,015
MN Efficiency of legal framewor	0,152	0,203	0,235
MN - GINNI	0,174	0,311	0,268
MN - Happy Planet Index_1	0,215	0,139	0,128
MN - International Cooperati_1	0,194	0,124	0,200

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
9**	OTHERS	AIBER	AIBER	5,110	0,841
			AVECO	25,204	0,000
			OTHERS	8,440	0,159
15**	OTHERS	AIBER	AIBER	10,74	0,883
			AVECO	30,55	0,000
			OTHERS	14,79	0,117
30**	OTHERS	AIBER	AIBER	8,871	0,479
			AVECO	11,504	0,128
			OTHERS	9,270	0,393
48**	OTHERS	AIBER	AIBER	4,264	0,587
			AVECO	16,884	0,001
			OTHERS	4,973	0,412
50**	OTHERS	AVECO	AIBER	13,577	0,017
			AVECO	6,012	0,763
			OTHERS	8,499	0,220
59**	OTHERS	AIBER	AIBER	1,662	0,990
			AVECO	18,351	0,000
			OTHERS	10,938	0,010
69**	OTHERS	AIBER	AIBER	7,590	0,679
			AVECO	25,050	0,000
			OTHERS	9,093	0,320
83**	OTHERS	AIBER	AIBER	5,047	0,513
			AVECO	16,485	0,002
			OTHERS	5,160	0,485
86**	OTHERS	AIBER	AIBER	11,38	0,825
			AVECO	28,39	0,000
			OTHERS	14,48	0,174
87**	OTHERS	AIBER	AIBER	4,893	0,538
			AVECO	20,927	0,000
			OTHERS	5,195	0,462
98**	OTHERS	AIBER	AIBER	1,742	0,782
			AVECO	12,027	0,005
			OTHERS	4,341	0,213
99**	OTHERS	AIBER	AIBER	10,54	0,601
			AVECO	13,55	0,133
			OTHERS	12,17	0,266
100**	AIBER	AVECO	AIBER	13,176	0,056
			AVECO	7,550	0,937
			OTHERS	17,575	0,006

107**	AVECO	OTHERS	AIBER	14,30	0,295
			AVECO	14,23	0,305
			OTHERS	13,69	0,400
110**	AIBER	AVECO	AIBER	10,385	0,037
			AVECO	3,897	0,957
			OTHERS	14,314	0,005
118**	OTHERS	AIBER	AIBER	3,591	0,576
			AVECO	12,784	0,006
			OTHERS	4,228	0,419

c. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo aplicar a análise de clusters de variáveis e identificar a distância entre os grupos, bem como a similaridade internamente. Para os quatro grupos estabelecidos, um deles tem maior pertencimento de países de AVECO (4), outro de AIBER (3) e os demais, um combina os países de OTHERS e alguns de AIBER mais avançados (1) e outro (2) os menos avançados de OTHERS acrescido de Bolívia (AIBER).
Observa-se que a função quadrática tem maior poder explicativo sobre o pertencimento aos grupos.

10.REGRESSÃO LOGÍSTICA

Este trabalho tem como objetivo realizar uma análise da capacidade explicativa de modelos baseados em regressões logísticas para caracterizar os dos agrupamentos de observações das principais variáveis que compõe as novas quatro Componentes Principais para o EIXO PESSOAS para 132 países, bem como estas últimas. É complementar ao emprego da análise discriminante, de modo que os principais resultados dos modelos com funções linear e quadrática são também trazidos neste trabalho para comparação com aqueles a serem obtidos com as regressões logísticas ordinais e nominais.Seis das variáveis originais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Todas as variáveis originais foram normalizadas e as observações faltantes foram substituídas pela média das observações não faltantes. Empregou-se o Minitab ® para as análises estatísticas.

a. AVALIAÇÃO E SELEÇÃO DO TIPO DE REGRESSÃO QUE MELHOR REPRESENTA OS GRUPOS DE PAÍSES ORGANIZADOS A PARTIR DAS QUATRO COMPONENTES PRINCIPAIS.

Ao aplicar-se o método de construção de clusters por ligação completa e distancia Manhattan, definindo-se como resultado esperado quatro *clusters*, obteve-se os seguintes resultaods que são o ponto de partida para definir a variável dependente da regressão:

	Average	Maximum

			Within cluster sum of squares	distance from centroid	distance from centroid
	Number of observations				
Cluster1	VERMELHO	32	83,511	1,53313	2,73551
Cluster2	LARANJA	42	168,697	1,87297	3,75273
Cluster3	AZUL	17	56,251	1,65681	3,27974
Cluster4	VERDE	41	233,455	2,29234	3,54600

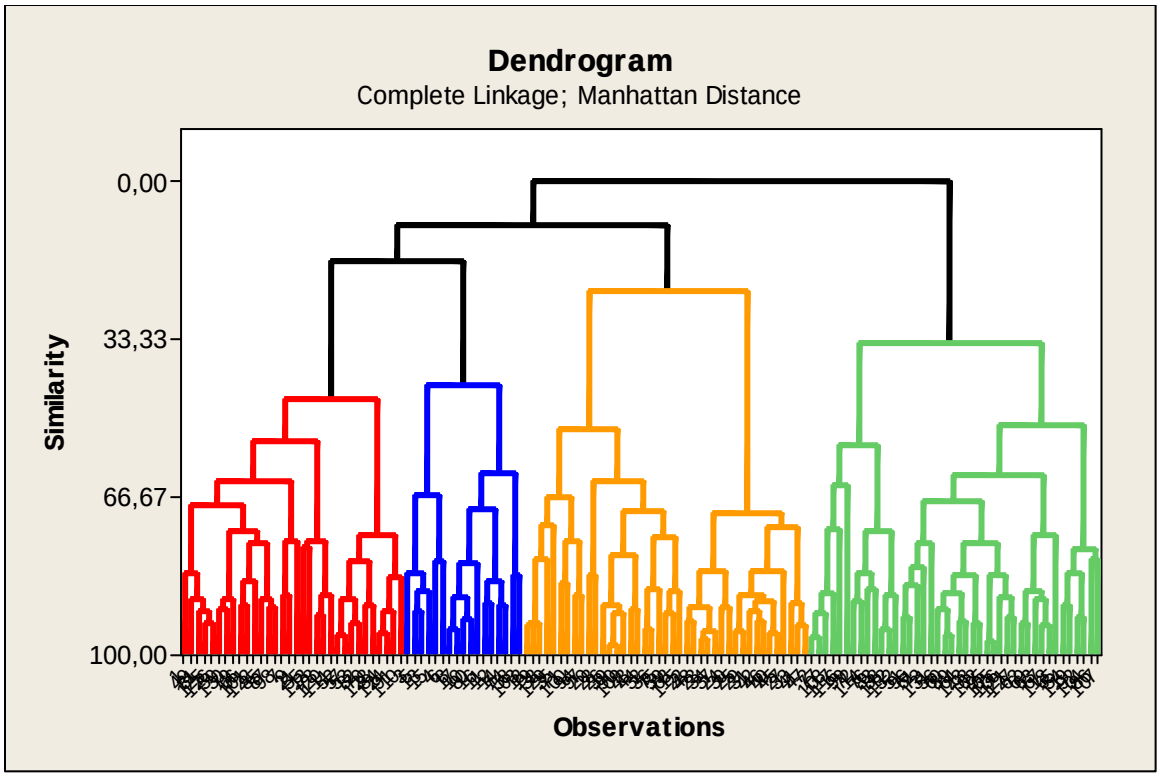
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
PC1	0,16468	-3,64920	-0,04299	3,62750	0,0000000
PC2	-1,52155	0,62389	0,43323	0,36882	-0,0000000
PC3	-0,14385	-0,44082	2,27974	-0,38141	-0,0000000
PC4	-0,25939	0,00520	-0,39646	0,36151	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,00000	4,39396	3,12360	4,00082
Cluster2	4,39396	0,00000	4,53915	7,29013
Cluster3	3,12360	4,53915	0,00000	4,59705
Cluster4	4,00082	7,29013	4,59705	0,00000

Dendrogram



Procedeu-se a um teste Anova para verificar e caracterizar os graus de diferença entre os quatro clusters a que chegamos no dendograma. A variável escolhida para análise é o IDH em sua forma normalizada (HDI no original em inglês), que altamente relacionada ao SPI, um das variáveis encontradas no estudo das componentes principais e empregadas na construção dos clusters. Uma das vantagens do emprego do IDH em substituição ao SPI é que o primeiro é

mais facilmente encontrado para diversos níveis, países, cidades, etc, enquanto o último é calculado somente para países. Piores IDHs concentram-se no grupo 2, com média de 29,91. Os grupos 1 e 3 (onde concentram-se os países do AIBER) tem IDHs mais próximos, com o primeiro ligeiramente superior e com menor dispersão. Pode-se esperar que alguns países do grupo 3 possuam IDH superiores a elementos do grupo 1. Já o grupo 4, onde se classificaram os países do AVECO, possuem IDH médio de 86,9, muito superior também ao grupo 1, de 65,3 e ao 3, 59,5, e são mais homogêneos entre si.

One-way ANOVA: MN - (HDI)_1 versus 4 CLUSTERS MANHATTAN

```
Source                 DF     SS     MS       F      P
4 CLUSTERS MANHATTAN    3  68679  22893  162,92  0,000
Error                 128  17987    141
Total                 131  86666

S = 11,85   R-Sq = 79,25%   R-Sq(adj) = 78,76%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -------+---------+---------+---------+--
1      32  65,29   9,82                    (-*-)
2      42  29,91  15,74   (-*-)
3      17  59,53  10,80                (--*--)
4      41  86,93   8,63                              (*-)
                          -------+---------+---------+---------+--
                                40        60        80       100

Pooled StDev = 11,85
```

Os grupos resultantes são agrupados por maior similaridade. Chega-se a algumas caracterizações. Nos grupos 1 e 2 predominam países de OUTRAS Regiões (OTHERS), sendo que no primeiro enquadram-se os países mais desenvolvidos de OTHERS, acrescido de alguns países de AIBER, e no grupo 2 os menos desenvolvidos. No grupo 3 encontra-se a maioria dos países de AIBER acrescidos de alguns países de OTHERS. E no grupo 4 encontram-se TODOS os países de AVECO, acrescidos de alguns de AIBER (Portugal, Espanha e Uruguai) e outros de OTHERS.

4 CLUSTERS MANHATTAN							
Grupo 1 - 32 países		**Grupo 2 - 42 países**		**Grupo 3 - 17 países**		**Grupo 4 - 41 países**	
País	**Região**	**País**	**Região**	**País**	**Região**	**País**	**Região**
Argentina	AIBER	Bolivia	AIBER	Brazil	AIBER	Portugal	AIBER
Cuba	AIBER	Angola	OTHERS	Chile	AIBER	Spain	AIBER
Ecuador	AIBER	Bangladesh	OTHERS	Colombia	AIBER	Uruguay	AIBER
Venezuela	AIBER	Benin	OTHERS	Costa Rica	AIBER	Australia	AVECO
Albania	OTHERS	Burkina Faso	OTHERS	Dominican Republic	AIBER	Austria	AVECO
Algeria	OTHERS	Burundi	OTHERS	El Salvador	AIBER	Belgium	AVECO
Armenia	OTHERS	Cambodia	OTHERS	Guatemala	AIBER	Canada	AVECO
Azerbaijan	OTHERS	Cameroon	OTHERS	Honduras	AIBER	Czech Republic	AVECO
Belarus	OTHERS	Central African Republic	OTHERS	Mexico	AIBER	Denmark	AVECO
Bosnia and Herzegovina	OTHERS	Chad	OTHERS	Nicaragua	AIBER	Estonia	AVECO

China	OTHERS	Congo, Republic of	OTHERS	Panama	AIBER	Finland	AVECO
Egypt	OTHERS	Djibouti	OTHERS	Paraguay	AIBER	France	AVECO
Iran	OTHERS	Ghana	OTHERS	Peru	AIBER	Germany	AVECO
Iraq	OTHERS	Guinea	OTHERS	Botswana	OTHERS	Greece	AVECO
Jordan	OTHERS	India	OTHERS	Guyana	OTHERS	Iceland	AVECO
Kazakhstan	OTHERS	Indonesia	OTHERS	Jamaica	OTHERS	Ireland	AVECO
Kuwait	OTHERS	Kenya	OTHERS	Namibia	OTHERS	Israel	AVECO
Kyrgyzstan	OTHERS	Laos	OTHERS			Italy	AVECO
Lebanon	OTHERS	Lesotho	OTHERS			Japan	AVECO
Macedonia	OTHERS	Liberia	OTHERS			Korea, Republic of	AVECO
Moldova	OTHERS	Madagascar	OTHERS			Latvia	AVECO
Morocco	OTHERS	Malawi	OTHERS			Netherlands	AVECO
Romania	OTHERS	Mali	OTHERS			New Zealand	AVECO
Russia	OTHERS	Mauritania	OTHERS			Norway	AVECO
Saudi Arabia	OTHERS	Mongolia	OTHERS			Slovakia	AVECO
Tajikistan	OTHERS	Mozambique	OTHERS			Slovenia	AVECO
Thailand	OTHERS	Nepal	OTHERS			Sweden	AVECO
Trinidad and Tobago	OTHERS	Niger	OTHERS			Switzerland	AVECO
Tunisia	OTHERS	Nigeria	OTHERS			United Kingdom	AVECO
Turkey	OTHERS	Pakistan	OTHERS			United States	AVECO
Ukraine	OTHERS	Philippines	OTHERS			Bulgaria	OTHERS
Uzbekistan	OTHERS	Rwanda	OTHERS			Croatia	OTHERS
		Senegal	OTHERS			Georgia	OTHERS
		South Africa	OTHERS			Hungary	OTHERS
		Sri Lanka	OTHERS			Lithuania	OTHERS
		Sudan	OTHERS			Malaysia	OTHERS
		Swaziland	OTHERS			Mauritius	OTHERS
		Tanzania	OTHERS			Montenegro	OTHERS
		Togo	OTHERS			Poland	OTHERS
		Uganda	OTHERS			Serbia	OTHERS
		Yemen	OTHERS			United Arab Emirates	OTHERS
		Zambia	OTHERS				

b. REGRESSÃO LOGÍSITCA

A variável dependente é definida como o pertencimento a cada um dos 4 *clusters* gerados na etapa anterior. E as variáveis independentes utilizadas são as oito variáveis relevantes constituintes dos componentes principais:
Social Progress Index
Access to improved sanitation
Network Readiness
Democracy índex
Efficiency of the legal framework to settle disputes
GINNI
Happy Planet Index
International Cooperation

Estas foram extraídas das equações obtidas anteriormente:

PC1 - INDICADORES DE PROGRESSO E DESENVLVIMENTO HUMANO.

```
PC1 = - 8,02 + 0,0333 MN -  (HDI)_1 + 0,0639 MN - Social Progress Index_1
      + 0,0138 MN-Access to improved sanitatio
      + 0,0274 MN The Networked Readiness Inde
```

PC2 - RESOLUTIVIDADE LEGAL, RECURSOS SANITÁRIOS E DEMOCRACIA.

```
The regression equation is
PC2 = - 1,09 - 0,0274 MN-Access to improved sanitatio
      + 0,0280 MN - Democracy index_1 + 0,0323 MN Efficiency of legal framewor
```

PC3 - FELICIDADE, COOPERAÇÃO E EQUIDADE.

```
The regression equation is
PC3 = - 1,72 - 0,0251 MN - GINNI + 0,0323 MN - Happy Planet Index_1
      + 0,0294 MN -  International Cooperati_1
```

PC4 - COOPERAÇÃO, EQUIDADE E RESOLUTIVIDADE LEGAL.

```
The regression equation is
PC4 = - 2,60 + 0,0233 MN - GINNI + 0,0389 MN -  International Cooperati_1
      - 0,0219 MN Efficiency of legal framewor
```

São testadas funções logísticas ordinal (onde os valores atribuídos ao pertencimento a cada cluster são organizados de forma crescente de 1 a 4 de modo a auxiliar com esta caracterização significativa da variável à execução da regressão logística; e nominal, onde não se estabelece este ordenamento.

Ao final do trabalho os resultados são adicionados e comparados com os obtidos com as funções de regressão multivariada linear e quadrática.

i. TESTE DA FUNÇÃO LOGÍSTICA ORDINAL PARA OS QUATRO AGRUPAMENTOS

Os grupos são reordenados do pior para o melhor em termos dos indicadores. Como apresentado em seções anteriores, empregou-se para a classificação um teste Anova com o IDH como variável *Proxy* para maior clarificação da posição dos grupos frente aos quatro componentes principais, para classificar os grupos.

O pior grupo por esta classificação é aquele composto pelo originalmente grupo 2, agora classificado como grupo 1 (pior IDH médio), o segundo o grupo 2, o terceiro e quatro mantendo sua denominação. O resultado obtido é uma regressão logística que encontra 95,3% de concordância, ou capacidade de explicar os resultados das oito variáveis das componentes principais dos 132 países. A descrição do teste é apresentada a seguir.

Ordinal Logistic Regression: 4R ORD versus MN - Governa; MN - Happy P; ...

```
Link Function: Logit


Response Information

Variable  Value  Count
4R ORD    1         42
          2         32
          3         17
          4         41
          Total    132


Logistic Regression Table

                                                                        Odds
Predictor                              Coef    SE Coef      Z      P  Ratio
Const(1)                            6,96395    2,02639   3,44  0,001
Const(2)                            10,5452    2,19983   4,79  0,000
Const(3)                            12,5733    2,30025   5,47  0,000
MN - Governança_1                -0,0909715  0,0352507  -2,58  0,010   0,91
MN - Happy Planet Index_1        -0,0121190  0,0129825  -0,93  0,351   0,99
MN - GINNI                        0,0219323  0,0133629   1,64  0,101   1,02
MN - Social Progress Index_1      -0,104868  0,0391102  -2,68  0,007   0,90
MN-Access to improved sanitatio  -0,0518674  0,0155273  -3,34  0,001   0,95
MN -  International Cooperati_1   0,0053595  0,0153540   0,35  0,727   1,01
MN - Democracy index_1            0,0154648  0,0188227   0,82  0,411   1,02
MN Efficiency of legal framewor   0,0523020  0,0207111   2,53  0,012   1,05

                                    95% CI
Predictor                         Lower  Upper
Const(1)
Const(2)
Const(3)
MN - Governança_1                 0,85   0,98
MN - Happy Planet Index_1         0,96   1,01
MN - GINNI                        1,00   1,05
MN - Social Progress Index_1      0,83   0,97
MN-Access to improved sanitatio   0,92   0,98
MN -  International Cooperati_1   0,98   1,04
MN - Democracy index_1            0,98   1,05
MN Efficiency of legal framewor   1,01   1,10


Log-Likelihood = -71,876
Test that all slopes are zero: G = 208,695, DF = 8, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      239,620  385  1,000
```

```
Deviance     143,751  385  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    6034     95,3  Somers' D              0,91
Discordant     297      4,7  Goodman-Kruskal Gamma  0,91
Ties             2      0,0  Kendall's Tau-a        0,66
Total         6333    100,0
```

ii. TESTE DA FUNÇÃO LOGÍSITCA NOMINAL PARA OS QUATRO AGRUPAMENTOS

Novamente com os grupos reordenados do pior para o melhor em termos dos indicadores, Procede-se à execução de uma função logística. Desta vez não se informa a classificação dos grupos. o mantendo sua denominação. O resultado obtido é uma regressão logística que encontra 95,3% de concordância, ou capacidade de explicar os resultados das oito variáveis das componentes principais dos 132 países. A descrição do teste é apresentada a seguir.

Nominal Logistic Regression: 4 CLUSTERS M versus MN - Governa; MN - GINNI; ...

```
Response Information

Variable              Value  Count
4 CLUSTERS MANHATTAN  4         41  (Reference Event)
                      3         17
                      2         42
                      1         32
                      Total    132


Logistic Regression Table

                                                                      Odds
Predictor                              Coef    SE Coef      Z      P  Ratio
Logit 1: (3/4)
Constant                            69,0271    28,9237   2,39  0,017
MN - Governança_1                 -0,0804289  0,0998257  -0,81  0,420   0,92
MN - GINNI                         -0,488690   0,181314  -2,70  0,007   0,61
MN -  (HDI)_1                      -0,199826   0,164465  -1,22  0,224   0,82
MN-Access to improved sanitatio    -0,275494   0,155328  -1,77  0,076   0,76
MN Efficiency of legal framewor   -0,0600779  0,0904864  -0,66  0,507   0,94
Logit 2: (2/4)
Constant                            65,5670    28,2969   2,32  0,020
MN - Governança_1                  -0,356978   0,149104  -2,39  0,017   0,70
MN - GINNI                         -0,250709   0,163648  -1,53  0,126   0,78
MN -  (HDI)_1                      -0,284582   0,180460  -1,58  0,115   0,75
MN-Access to improved sanitatio    -0,401340   0,166657  -2,41  0,016   0,67
MN Efficiency of legal framewor     0,273611   0,122043   2,24  0,025   1,31
Logit 3: (1/4)
Constant                            26,1947    20,1078   1,30  0,193
MN - Governança_1                  -0,415096   0,145944  -2,84  0,004   0,66
MN - GINNI                         0,0147673  0,0656891   0,22  0,822   1,01
MN -  (HDI)_1                     -0,0635876   0,150098  -0,42  0,672   0,94
MN-Access to improved sanitatio   -0,0726449  0,0886260  -0,82  0,412   0,93
MN Efficiency of legal framewor    0,0896438  0,0747905   1,20  0,231   1,09

                                      95% CI
Predictor                          Lower  Upper
Logit 1: (3/4)
Constant
```

```
MN - Governança_1                0,76   1,12
MN - GINNI                       0,43   0,88
MN -  (HDI)_1                    0,59   1,13
MN-Access to improved sanitatio  0,56   1,03
MN Efficiency of legal framewor  0,79   1,12
Logit 2: (2/4)
Constant
MN - Governança_1                0,52   0,94
MN - GINNI                       0,56   1,07
MN -  (HDI)_1                    0,53   1,07
MN-Access to improved sanitatio  0,48   0,93
MN Efficiency of legal framewor  1,04   1,67
Logit 3: (1/4)
Constant
MN - Governança_1                0,50   0,88
MN - GINNI                       0,89   1,15
MN -  (HDI)_1                    0,70   1,26
MN-Access to improved sanitatio  0,78   1,11
MN Efficiency of legal framewor  0,94   1,27


Log-Likelihood = -19,759
Test that all slopes are zero: G = 312,928, DF = 15, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      57,6056  378  1,000
Deviance     39,5187  378  1,000
```

Não se encontra nesta tentativa de produzir uma regressão logística nominal, um resultado válido, o que eventualmente pode ocorrer.

iii. RESULTADOS DOS TESTE COM FUNÇÕES LOGÍSTICAS PARA OS QUATRO AGRUPAMENTOS DE PAÍSES.

Para entendimento dos quatro grupos de países observa-se que a regressão logística nominal não é capaz de encontrar uma resposta válida enquanto a função ordinal provou-se melhor, trazendo uma regressão capaz de explicar 95,3% das observações.

c. TESTE DA FUNÇÕES LOGÍSTICA PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS

Procede-se ao desenvolvimento de funções logísticas ordinal e nominal para os grupos AIBER, AVECO e OTHERS. O teste Anova como o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3), o que será empregado na função Ordinal.

One-way ANOVA: MN - (HDI)_1 versus 3R ORD

```
Source   DF     SS     MS      F      P
3R ORD    2  39538  19769  54,11  0,000
Error   129  47128    365
Total   131  86666

S = 19,11   R-Sq = 45,62%   R-Sq(adj) = 44,78%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
1      84  48,35  22,88   (--*--)
2      21  65,67  12,45            (-----*----)
3      27  91,89   4,78                              (----*----)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 19,11
```

i. TESTE DA FUNÇÃO LOGÍSTICA ORDINAL PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS

A função logística ordinal com a variável dependente 3R ORD, ou classificação dos países entre AIBER, AVECO e OTHERS, ordenada como indicado no item 2.7. e as 8 variáveis das quatro Componentes Principais possui uma consistência de 91,2%. A descrição do procedimento no Minitab ® encontra-se a seguir.

Ordinal Logistic Regression: 3R ORD versus MN - Governa; MN - GINNI; ...

```
Link Function: Logit


Response Information

Variable  Value  Count
3R ORD    1         84
          2         21
          3         27
          Total    132


Logistic Regression Table

                                                                    Odds
Predictor                            Coef    SE Coef      Z      P  Ratio
Const(1)                          6,20979    1,41829   4,38  0,000
Const(2)                          8,09137    1,53560   5,27  0,000
MN - Governança_1               -0,0431922  0,0196911  -2,19  0,028   0,96
MN - GINNI                       0,0255702  0,0128843   1,98  0,047   1,03
MN -  (HDI)_1                    -0,128171  0,0367094  -3,49  0,000   0,88
MN-Access to improved sanitatio  0,0316443  0,0195090   1,62  0,105   1,03
MN Efficiency of legal framewor  0,0179547  0,0152671   1,18  0,240   1,02

                                   95% CI
Predictor                        Lower  Upper
Const(1)
Const(2)
MN - Governança_1                 0,92   1,00
MN - GINNI                        1,00   1,05
MN -  (HDI)_1                     0,82   0,95
MN-Access to improved sanitatio   0,99   1,07
MN Efficiency of legal framewor   0,99   1,05


Log-Likelihood = -66,521
Test that all slopes are zero: G = 105,796, DF = 5, P-Value = 0,000


Goodness-of-Fit Tests
```

Method	Chi-Square	DF	P
Pearson	179,515	257	1,000
Deviance	133,041	257	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	4192	91,2	Somers' D	0,82
Discordant	401	8,7	Goodman-Kruskal Gamma	0,83
Ties	6	0,1	Kendall's Tau-a	0,44
Total	4599	100,0		

ii. TESTE DA FUNÇÃO LOGÍSTICA NOMINAL PARA AS TRÊS REGIÕES DE PAÍSES: AIBER, AVECO E OTHERS

A função logística nominal com a variável dependente 3R ORD, ou classificação dos países entre AIBER, AVECO e OTHERS, ordenada como indicado no item 2.7. e as 8 variáveis das quatro Componentes Principais não gerou resultados no Minitab ® apresentado a seguir.

Nominal Logistic Regression: 3R ORD versus MN - Governa; MN - Happy P; ...

Response Information

Variable	Value	Count	
3R ORD	3	27	(Reference Event)
	2	21	
	1	84	
	Total	132	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Logit 1: (2/3)					
Constant	9,23985	12,0321	0,77	0,443	
MN - Governança_1	-0,284489	0,186833	-1,52	0,128	0,75
MN - Happy Planet Index_1	0,0447189	0,0575921	0,78	0,437	1,05
MN - GINNI	-0,175736	0,0647833	-2,71	0,007	0,84
MN - Social Progress Index_1	0,400284	0,213099	1,88	0,060	1,49
MN-Access to improved sanitatio	0,0339986	0,0495005	0,69	0,492	1,03
MN - International Cooperati_1	0,0293645	0,0324757	0,90	0,366	1,03
MN - Democracy index_1	-0,242788	0,161892	-1,50	0,134	0,78
MN Efficiency of legal framewor	-0,0199959	0,0664123	-0,30	0,763	0,98
Logit 2: (1/3)					
Constant	26,4388	11,5533	2,29	0,022	
MN - Governança_1	-0,131537	0,178691	-0,74	0,462	0,88
MN - Happy Planet Index_1	-0,0660609	0,0524605	-1,26	0,208	0,94
MN - GINNI	-0,0299675	0,0486243	-0,62	0,538	0,97
MN - Social Progress Index_1	-0,125474	0,180858	-0,69	0,488	0,88
MN-Access to improved sanitatio	0,0693506	0,0583216	1,19	0,234	1,07
MN - International Cooperati_1	0,0401640	0,0339882	1,18	0,237	1,04
MN - Democracy index_1	-0,208848	0,155092	-1,35	0,178	0,81
MN Efficiency of legal framewor	0,0576287	0,0627812	0,92	0,359	1,06

Predictor	95% CI Lower	Upper
Logit 1: (2/3)		

```
Constant
MN - Governança_1                 0,52   1,09
MN - Happy Planet Index_1         0,93   1,17
MN - GINNI                        0,74   0,95
MN - Social Progress Index_1      0,98   2,27
MN-Access to improved sanitatio   0,94   1,14
MN -  International Cooperati_1   0,97   1,10
MN - Democracy index_1            0,57   1,08
MN Efficiency of legal framewor   0,86   1,12
Logit 2: (1/3)
Constant
MN - Governança_1                 0,62   1,24
MN - Happy Planet Index_1         0,84   1,04
MN - GINNI                        0,88   1,07
MN - Social Progress Index_1      0,62   1,26
MN-Access to improved sanitatio   0,96   1,20
MN -  International Cooperati_1   0,97   1,11
MN - Democracy index_1            0,60   1,10
MN Efficiency of legal framewor   0,94   1,20


Log-Likelihood = -22,446
Test that all slopes are zero: G = 193,946, DF = 16, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      157,988  246  1,000
Deviance      44,892  246  1,000
```

iii. RESULTADOS DOS TESTE COM FUNÇÕES LOGÍSTICAS PARA O PERTENCIENTO A AIBER, AVECO E OTHERS.

Para entendimento do pertencimento de um país a um dos trÊs grupos, AIBER, AVECO ou OTHERS, verifica-se que a regressão logística nominal não é capaz de encontrar uma resposta válida enquanto a função ordinal provou-se melhor, trazendo uma regressão capaz de explicar 91,2% das observações.

d. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo combinar três técnicas. Primeiro e segundo, respectivamente, aplicar os resultados da analise discriminante com funções linear e quadrática como forma de explicação dos resultados análise de clusters de variáveis para identificar a distância entre os grupos, bem como a similaridade internamente. Terceiro, comparar os resultados com a aplicação de função logística ordinal e nominal, e verificar qual mais se adequava para explicar o pertencimento aos quatro clustrrs de países organizaç~do pelos quatro componentes principais, e segundo, pelo pertencimetno a um dos três grupos, AIBER, AVECO e OTHERS. As regressões Quadráticas, como desenvolvidas e apresentadas no trabalho 9, retronaram melhor significância para explicação do que as lineares e seus resultados são apresentados no quadro “Resultados das Regressões”. Da mesma forma, as regressões ordinais foram mais eficazes para explicar o petrtencimento a cada um dos dois conjuntos de grupos do que as regressões log´sitcas nominais e também são apresentadas no mesmo quadro, que é apresentado na sequência.

Resultados das Regressões

Var. Dependente / Método	4 Agrupamentos (PCs)	AIBER / AVECO / OTHERS
Análise Discriminante	96,2% (Q)	91,2% (Q)
Regressão Logística	94,6 %(O)	90,7% (O)

Legenda: Q = Função Quadrátca; O = Regressão Logística Ordinal.

11. ÁRVORE DE CLASSIFICAÇÃO

Este trabalho tem como objetivo realizar uma análise da capacidade explicativa de modelos baseados em Árvores de Classificação para caracterizar dos agrupamentos de observações das oito principais variáveis que compõe as novas quatro Componentes Principais para o EIXO PESSOAS para 132 países, considerando os agrupamentos gerados pela análise discriminante e os grupos de países AIBER, AVECO e OTHERS. Seis das variáveis originais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Todas as variáveis originais foram normalizadas e as observações faltantes foram substituídas pela média das observações não faltantes. Empregou-se neste trabalho o IBM SPSS19® para as análises estatísticas.

a. Variáveis chave dos Componentes Principais

Considerando o princípio da parcimônia selecionamos as oito principais variáveis independentes a partir do entendimento dos componentes principais:

```
Índice de Governança
HappyPlanetIndex
GINNI
Access to improved sanitation
International Cooperation
HDI / IDH
Democracy índex
Efficiency of legal framework to settle disputes index
```

b. APLICAÇÃO DA ANÁLISE DISCRIMINANTE COM SEPARAÇÃO DE CLUSTERS – MÉTODO MANHATAN

Ao aplicar-se no IBM SPSS19® o método de construção de clusters por ligação completa e distancia Manhattan, definindo-se como resultado esperado quatro *clusters*, obteve-se os seguintes resultados que são o ponto de partida para definir a variável dependente da regressão:

Aplicando-se o método *stepwise* para identificar a função que melhor discrimina os agrupamentos, chega-se a seis variáveis relevantes, conforme a Tabela 2, todas com significância estatística..

Tabela 2 – Stepwise – Variáveis inseridas nas funções

Variables Entered/Removed[a,b,c,d]

Ste	Entered	Wilks' Lambda

p		Statistic	df1	df2	df3	Exact F				Approximate F			
						Statistic	df1	df2	Sig.	Statistic	df1	df2	Sig.
1	MN - (HDI)_1	,208	1	3	128,000	162,916	3	128,000	,000				
2	MN - Governança_1	,106	2	3	128,000	87,541	6	254,000	,000				
3	MN - GINNI	,072	3	3	128,000					66,352	9	306,801	,000
4	MN - Happy Planet Index_1	,057	4	3	128,000					54,092	12	331,010	,000
5	MN-Access to improved sanitatio	,045	5	3	128,000					47,308	15	342,711	,000
6	MN Efficiency of legal framewor	,037	6	3	128,000					43,005	18	348,382	,000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 16.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

As funções resultantes seguem os pesos das variáveis descritos na Tabela 3.

Tabela 3 - Standardized Canonical Discriminant Function Coefficients

	Function		
	1	2	3
MN - Governança_1	,543	-1,197	-,263
MN - Happy Planet Index_1	-,172	-,006	-,683
MN - GINNI	,061	,219	,780
MN-Access to improved sanitatio	,460	,390	-,196
MN - (HDI)_1	,522	,408	,127
MN Efficiency of legal framewor	-,410	,427	,495

Todos os 132 países foram classificados. A capacidade preditiva da técnica é relevante como apresentado na Tabela 4. O grupo 1 consegue pelo modelo 87,5% de predição de resultados

(28 de 32 elementos), o grupo 2 88,1% (37 em 42), o grupo 3 100% dos 17 elementos pertencentes e o grupo 4, 90,2%, ou 37 em 41.

Tabela 4 - Classification Results (a)

		4 CLUSTERS MANHATTAN	Predicted Group Membership				Total
			1	2	3	4	
Original	Count	1	28	0	2	2	32
		2	3	37	2	0	42
		3	0	0	17	0	17
		4	3	0	1	37	41
	%	1	87,5	,0	6,3	6,3	100,0
		2	7,1	88,1	4,8	,0	100,0
		3	,0	,0	100,0	,0	100,0
		4	7,3	,0	2,4	90,2	100,0

a. 90,2% of original grouped cases correctly classified.

c. ANÁLISE DISCRIMINANTE COM OS GRUPOS AIBER, AVECO E OTHERS.

Emprega-se agora o método Manhattan em outr classificação.A variável dependente é definida como o pertencimento a cada um dos 3 *clusters* AIBER, AVECO e OTHERS. Na Tabela 4, o estudo seguem com cinco variáveis mostrando-se significativas, gerado a partir do desenvolvimento de funções logísticas ordinal e nominal para os grupos. O teste Anova realizado anteriormente com o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3), o que será empregado na função Ordinal.

Tabela 4 - Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda							
						Exact F			
		Statistic	df1	df2	df3	Statistic	df1	df2	Sig.
1	MN - Governança_1	,397	1	2	129,000	97,930	2	129,000	,000
2	MN - Happy Planet Index_1	,296	2	2	129,000	53,720	4	256,000	,000
3	MN - GINNI	,237	3	2	129,000	44,696	6	254,000	,000
4	MN - International Cooperati_1	,204	4	2	129,000	38,166	8	252,000	,000
5	MN Efficiency of legal framewor	,185	5	2	129,000	33,075	10	250,000	,000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 16.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

A classificação dos três clusters gerou resultados também satisfatórios. Para o pertencimento no OTHERS (1), chega-se a 83% de acertos com o método empregado. Para AIBER (2), 81%, e AVECO (3), 93%, conforme a Tabela 5.

Tabela 5 - Classification Results[a]

			Predicted Group Membership			
		3R ORD	1	2	3	Total
Original	Count	1	70	12	2	84
		2	2	17	2	21
		3	1	1	25	27
	%	1	83,3	14,3	2,4	100,0
		2	9,5	81,0	9,5	100,0
		3	3,7	3,7	92,6	100,0

i. **a. 84,8% of original grouped cases correctly classified.**

ii.

d. EMPREGO DA ÁRVORE DE DECISÃO COMO FERRAMENTA DE CLASSIFICAÇÃO.

Utilizam-se quatro metodos de árvore de decisão para classificação dos 4 *clusters* e os 3 grupos: CHAID, CHAID EXAUSTIVO, CRT e QUEST.

i. 4 CLUSTERS RESULTADOS DOS TESTES COM ÁRVORE DE DECISÃO

Emprega-se neste agrupamento os 4 métodos.

1. MÉTODO CHAID

Ao se empregar o método Chaid para os 4 clusters (nomeados pelo resultado de análise Anova pelo IDH), verificamos que somente um grupo, o médio, possui percentual de acerto inferior a 80% de predições corretas, 71%.

Classification

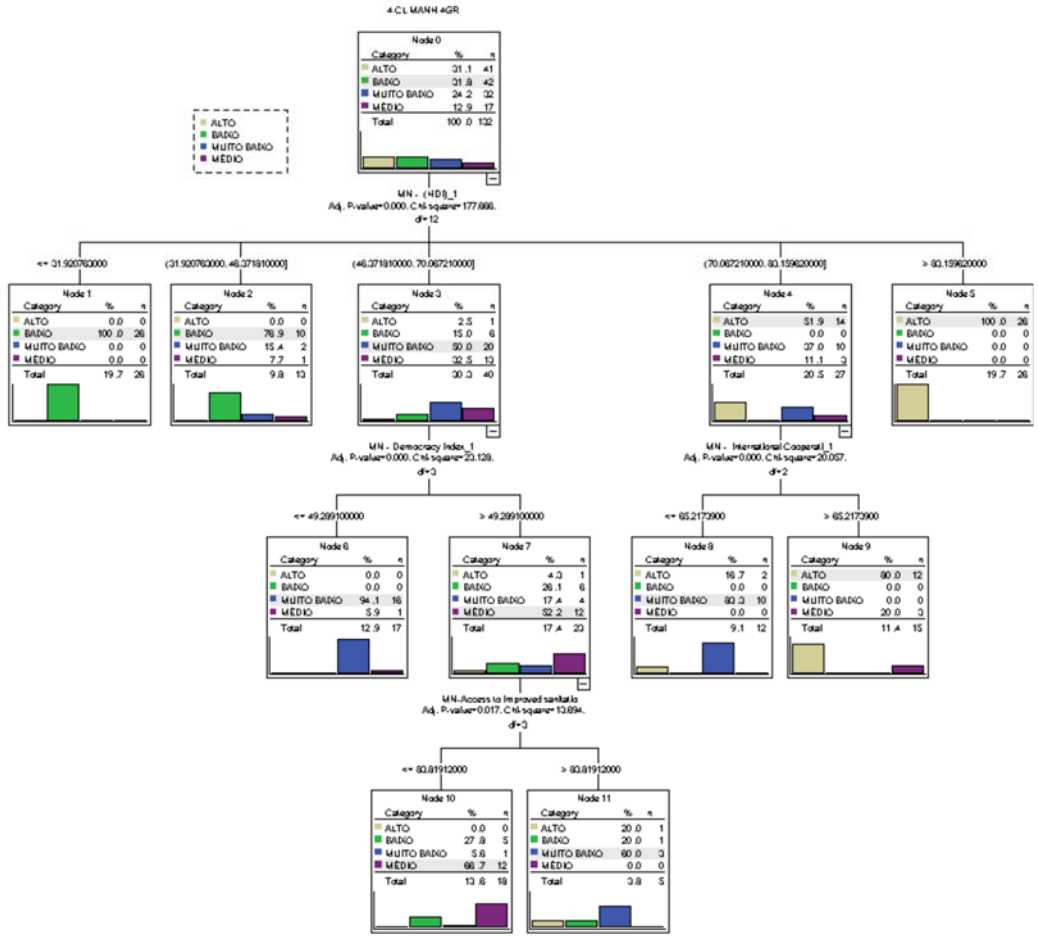
	Predicted				
Observed	ALTO	BAIXO	MUITO BAIXO	MÉDIO	Percent Correct
ALTO	38	0	3	0	92,7%
BAIXO	0	36	1	5	85,7%
MUITO BAIXO	0	2	29	1	90,6%
MÉDIO	3	1	1	12	70,6%
Overall Percentage	31,1%	29,5%	25,8%	13,6%	87,1%

Classification

Observed	Predicted				
	ALTO	BAIXO	MUITO BAIXO	MÉDIO	Percent Correct
ALTO	38	0	3	0	92,7%
BAIXO	0	36	1	5	85,7%
MUITO BAIXO	0	2	29	1	90,6%
MÉDIO	3	1	1	12	70,6%
Overall Percentage	31,1%	29,5%	25,8%	13,6%	87,1%

Growing Method: CHAID

Dependent Variable: 4 CL MANH 4GR



2. MÉTODO EXHAUSTIVE CHAID

Ao se empregar o método Exhaustive Chaid para os 4 clusters (nomeados pelo resultado de análise Anova pelo IDH), verificamos que o grupo de médio IDH possui somente 47% de predições corretas.

Tabela 6 – Exhaustive Chaid

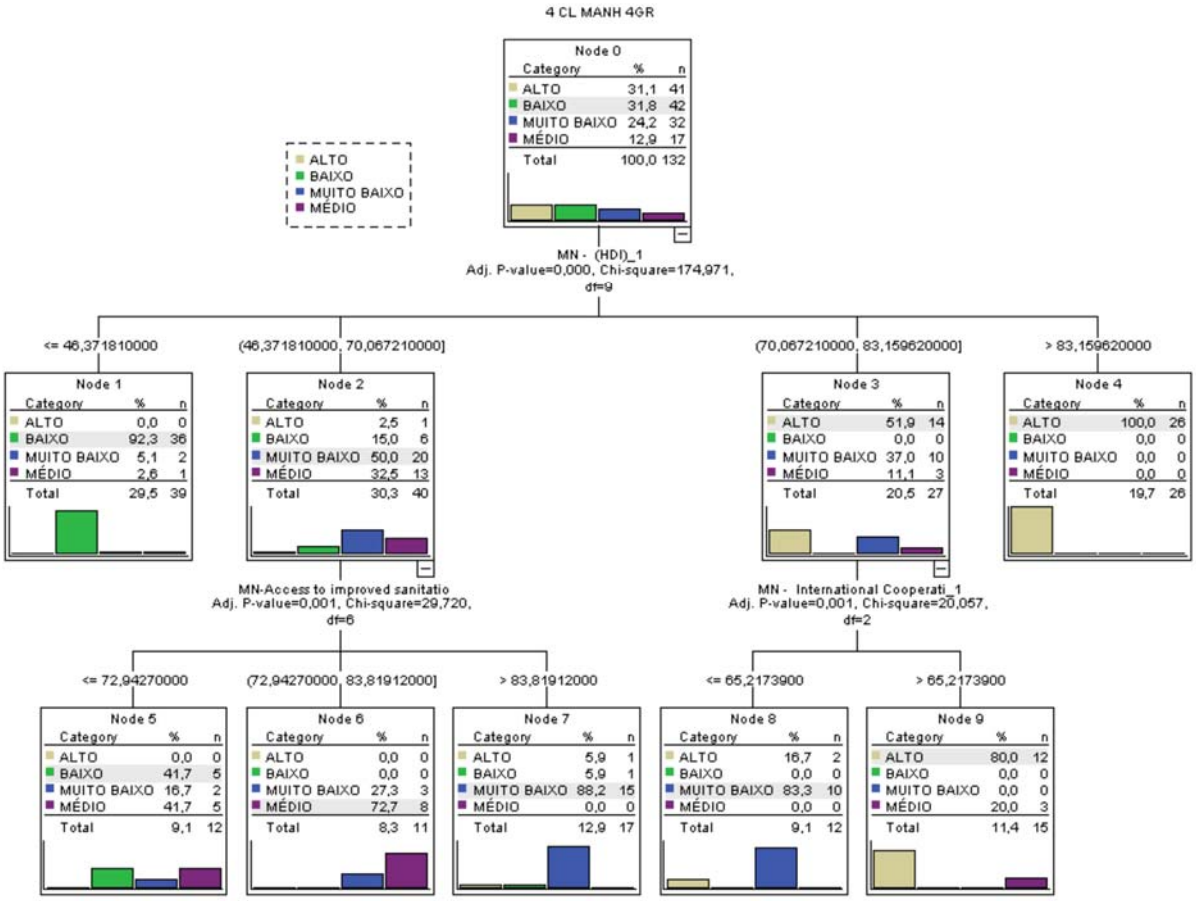
Classification

Observed	Predicted				
	ALTO	BAIXO	MUITO BAIXO	MÉDIO	Percent Correct
ALTO	38	0	3	0	92,7%
BAIXO	0	41	1	0	97,6%
MUITO BAIXO	0	4	25	3	78,1%
MÉDIO	3	6	0	8	47,1%
Overall Percentage	31,1%	38,6%	22,0%	8,3%	84,8%

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: 4 CL MANH 4GR

Gráfico 2 - Exhaustive Chaid



3. MÉTODO CRT

Ao se empregar o método CRT para os 4 clusters (nomeados pelo resultado de análise Anova pelo IDH), verificamos que o grupo de médio IDH possui somente 47% de predições corretas.

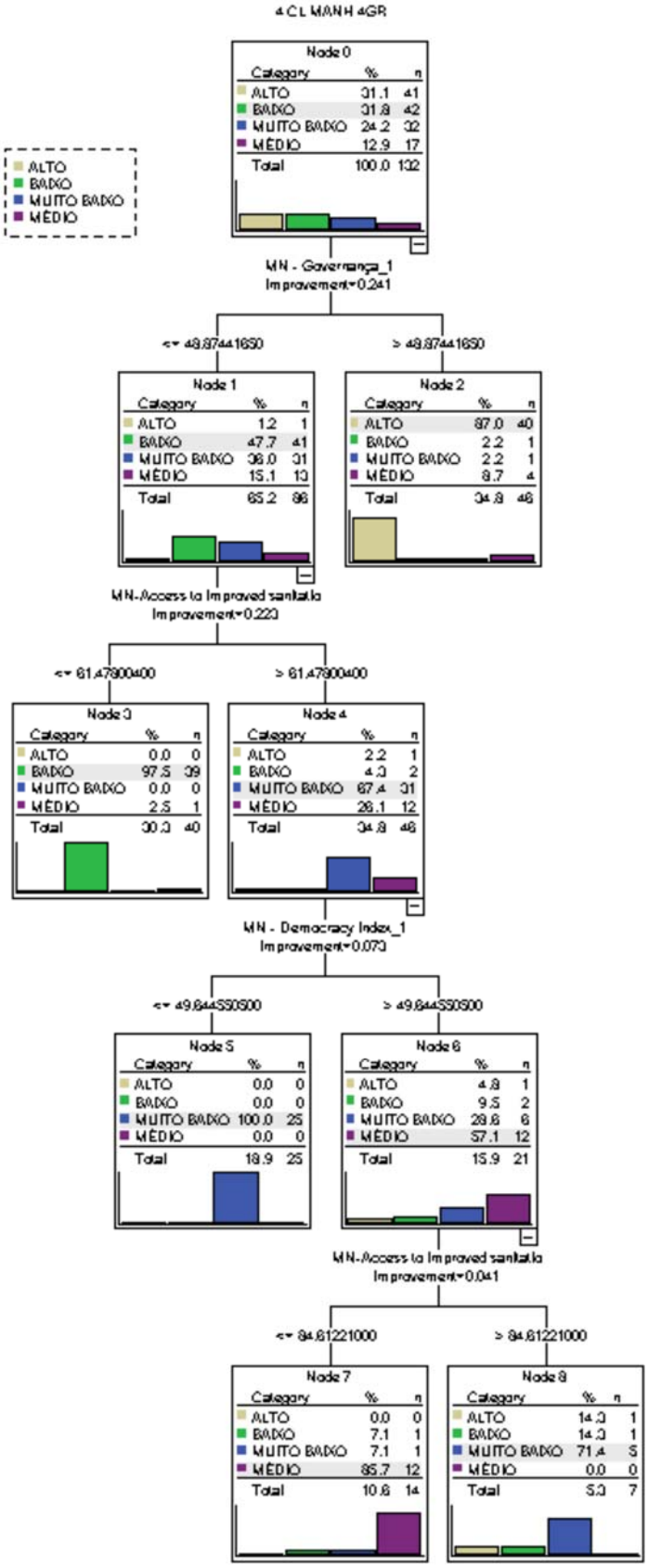
Classification

	Predicted				
Observed	ALTO	BAIXO	MUITO BAIXO	MÉDIO	Percent Correct
ALTO	40	0	1	0	97,6%
BAIXO	1	39	1	1	92,9%
MUITO BAIXO	1	0	30	1	93,8%
MÉDIO	4	1	0	12	70,6%
Overall Percentage	34,8%	30,3%	24,2%	10,6%	91,7%

Growing Method: CRT

Dependent Variable: 4 CL MANH 4GR

Gráfico 3 - CRT


4 CLMANH 4GR
ALTO
BAIXO
MUITO BAIXO
MÉDIO
Node 0
Category % n
ALTO 31.1 41
BAIXO 31.8 42
MUITO BAIXO 24.2 32
MÉDIO 12.9 17
Total 100.0 132
MN - Governança_1
Improvement=0.241
<= 48.87441650
> 48.87441650
Node 1
Category % n
ALTO 1.2 1
BAIXO 47.7 41
MUITO BAIXO 36.0 31
MÉDIO 15.1 13
Total 65.2 86
Node 2
Category % n
ALTO 87.0 40
BAIXO 2.2 1
MUITO BAIXO 2.2 1
MÉDIO 8.7 4
Total 34.8 46
MN-Access to Improved sanitatio
Improvement=0.223
<= 61.47800400
> 61.47800400
Node 3
Category % n
ALTO 0.0 0
BAIXO 97.5 39
MUITO BAIXO 0.0 0
MÉDIO 2.5 1
Total 30.3 40
Node 4
Category % n
ALTO 2.2 1
BAIXO 4.3 2
MUITO BAIXO 67.4 31
MÉDIO 26.1 12
Total 34.8 46
MN - Democracy Index_1
Improvement=0.073
<= 49.64450900
> 49.64450900
Node 5
Category % n
ALTO 0.0 0
BAIXO 0.0 0
MUITO BAIXO 100.0 25
MÉDIO 0.0 0
Total 18.9 25
Node 6
Category % n
ALTO 4.8 1
BAIXO 9.5 2
MUITO BAIXO 28.6 6
MÉDIO 57.1 12
Total 15.9 21
MN-Access to Improved sanitatio
Improvement=0.041
<= 84.61221000
> 84.61221000
Node 7
Category % n
ALTO 0.0 0
BAIXO 7.1 1
MUITO BAIXO 7.1 1
MÉDIO 85.7 12
Total 10.6 14
Node 8
Category % n
ALTO 14.3 1
BAIXO 14.3 1
MUITO BAIXO 71.4 5
MÉDIO 0.0 0
Total 5.3 7

4. MÉTODO EXHAUSTIVE QUEST

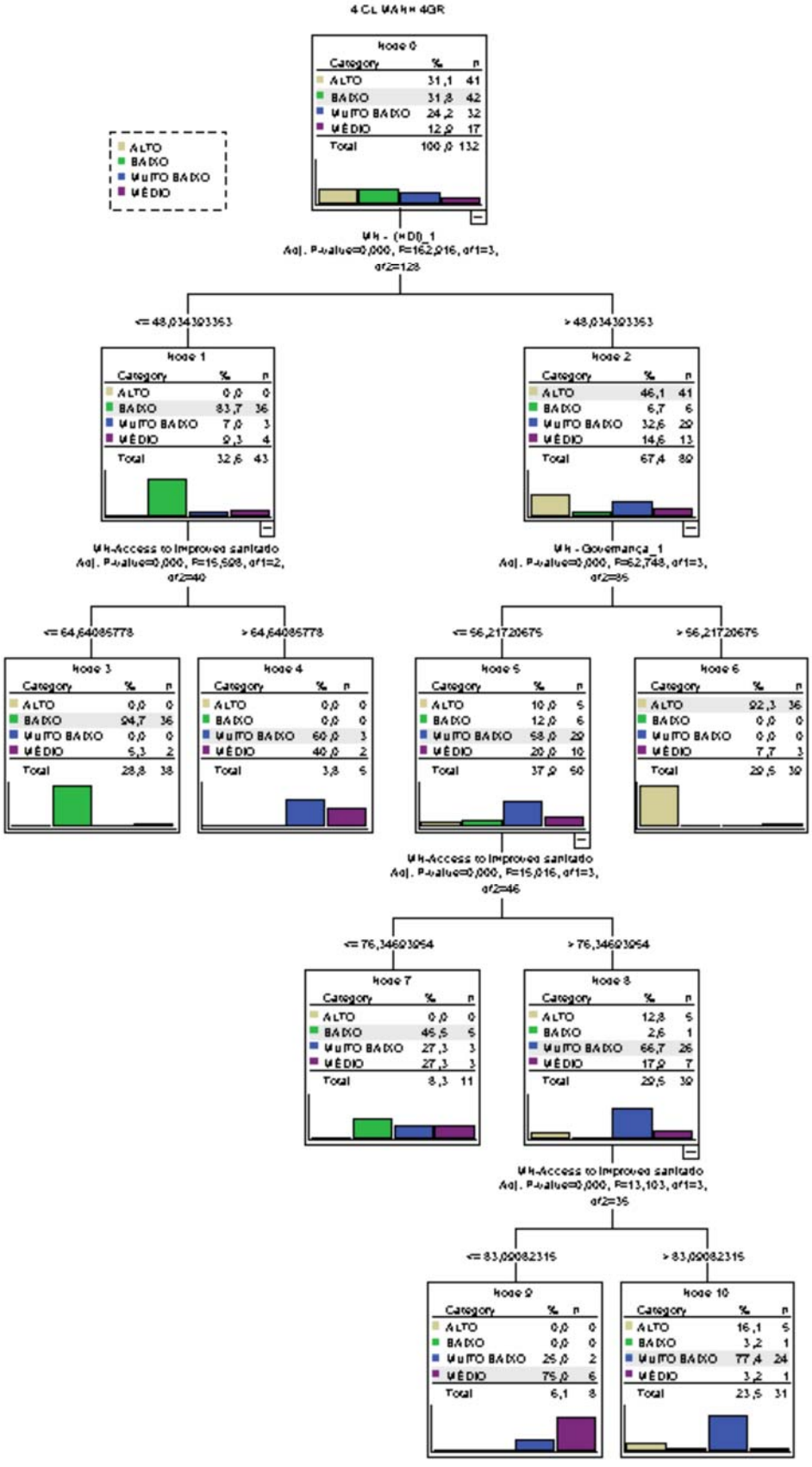
Ao se empregar o método QUEST para os 4 clusters (nomeados pelo resultado de análise Anova pelo IDH), verificamos que o grupo de médio IDH possui somente 35% de predições corretas.

Classification

Observed	Predicted				
	ALTO	BAIXO	MUITO BAIXO	MÉDIO	Percent Correct
ALTO	36	0	5	0	87,8%
BAIXO	0	41	1	0	97,6%
MUITO BAIXO	0	3	27	2	84,4%
MÉDIO	3	5	3	6	35,3%
Overall Percentage	29,5%	37,1%	27,3%	6,1%	83,3%

Growing Method: QUEST

Dependent Variable: 4 CL MANH 4GR


4 CL MAIN 4GR
ALTO
BAIXO
MUITO BAIXO
MÉDIO
Node 0
Category % n
ALTO 31,1 41
BAIXO 31,8 42
MUITO BAIXO 24,2 32
MÉDIO 12,9 17
Total 100,0 132
MH - (HDI)_1
Adj. P-value=0,000, F=162,016, df1=3, df2=128
<= 48,034303363
> 48,034303363
Node 1
Category % n
ALTO 0,0 0
BAIXO 83,7 36
MUITO BAIXO 7,0 3
MÉDIO 9,3 4
Total 32,6 43
Node 2
Category % n
ALTO 46,1 41
BAIXO 6,7 6
MUITO BAIXO 32,6 29
MÉDIO 14,6 13
Total 67,4 89
MH-Access to improved sanitado
Adj. P-value=0,000, F=15,598, df1=2, df2=40
MH - Governança_1
Adj. P-value=0,000, F=62,748, df1=3, df2=85
<= 64,64086778
> 64,64086778
<= 56,21720675
> 56,21720675
Node 3
Category % n
ALTO 0,0 0
BAIXO 94,7 36
MUITO BAIXO 0,0 0
MÉDIO 5,3 2
Total 28,8 38
Node 4
Category % n
ALTO 0,0 0
BAIXO 0,0 0
MUITO BAIXO 60,0 3
MÉDIO 40,0 2
Total 3,8 5
Node 5
Category % n
ALTO 10,0 5
BAIXO 12,0 6
MUITO BAIXO 58,0 29
MÉDIO 20,0 10
Total 37,9 50
Node 6
Category % n
ALTO 92,3 36
BAIXO 0,0 0
MUITO BAIXO 0,0 0
MÉDIO 7,7 3
Total 29,5 39
MH-Access to improved sanitado
Adj. P-value=0,000, F=15,016, df1=3, df2=46
<= 76,34603064
> 76,34603064
Node 7
Category % n
ALTO 0,0 0
BAIXO 45,5 5
MUITO BAIXO 27,3 3
MÉDIO 27,3 3
Total 8,3 11
Node 8
Category % n
ALTO 12,8 5
BAIXO 2,6 1
MUITO BAIXO 66,7 26
MÉDIO 17,9 7
Total 29,5 39
MH-Access to improved sanitado
Adj. P-value=0,000, F=13,103, df1=3, df2=35
<= 83,09082315
> 83,09082315
Node 9
Category % n
ALTO 0,0 0
BAIXO 0,0 0
MUITO BAIXO 25,0 2
MÉDIO 75,0 6
Total 6,1 8
Node 10
Category % n
ALTO 16,1 5
BAIXO 3,2 1
MUITO BAIXO 77,4 24
MÉDIO 3,2 1
Total 23,5 31

ii. 3 GRUPOS DE PAÍSES: ABER, AVECO E OTHERS - RESULTADOS DOS TESTES COM ÁRVORE DE DECISÃO

Emprega-se neste agrupamento os 4 métodos.

1. MÉTODO CHAID

Procede-se ao desenvolvimento de funções logísticas ordinal e nominal para os grupos AIBER, AVECO e OTHERS. O teste Anova como o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3), o que será empregado na função Ordinal. Não se encontrab bons resultados na identificação do pertencimento ao AIBER (43% de aertos somente). E o gráfico 1 apresenta visualmente as conclusões.

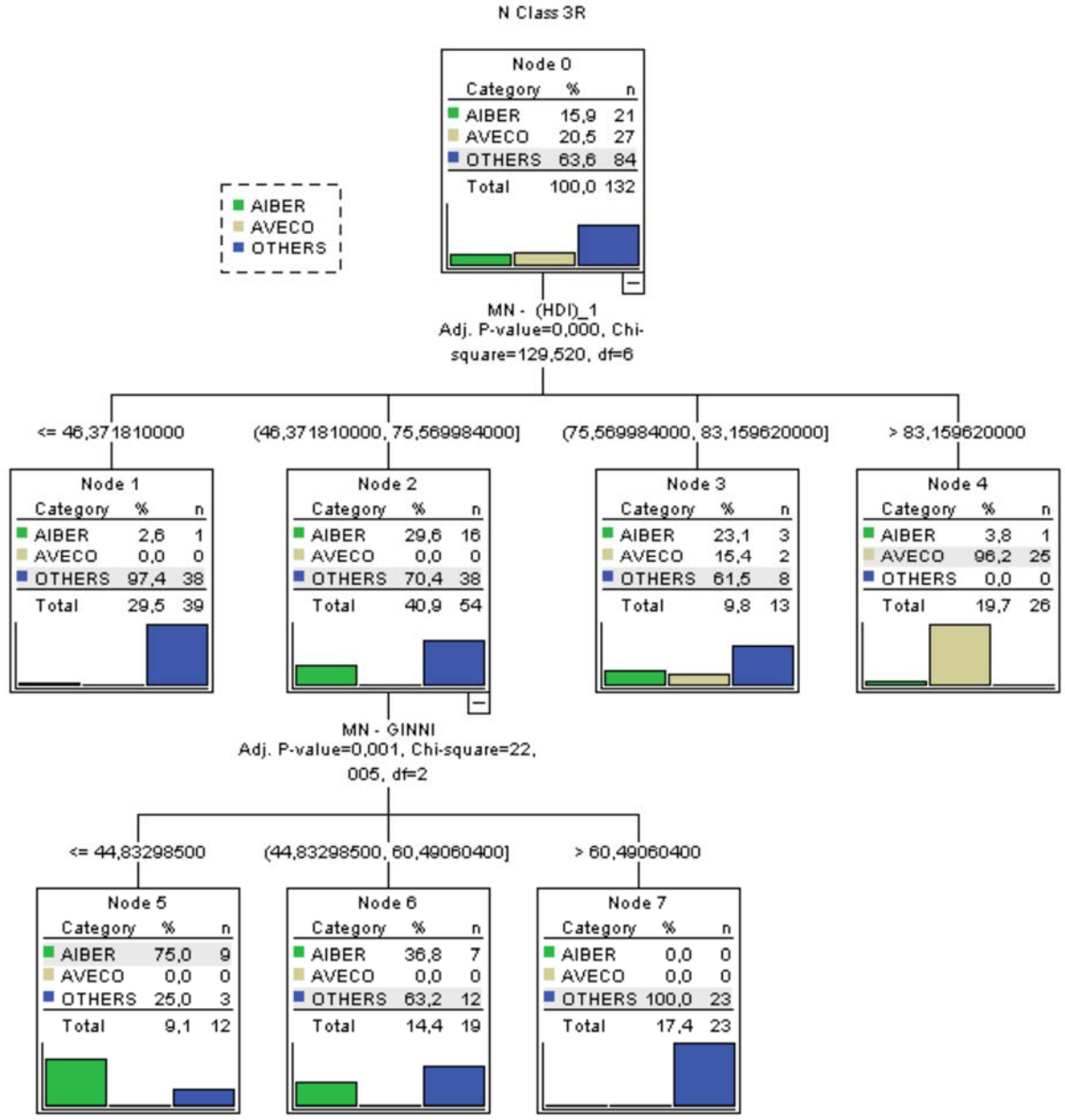
Tabela 10 – Classification

	Predicted			
Observed	AIBER	AVECO	OTHERS	Percent Correct
AIBER	9	1	11	42,9%
AVECO	0	25	2	92,6%
OTHERS	3	0	81	96,4%
Overall Percentage	9,1%	19,7%	71,2%	87,1%

Growing Method: CHAID

Dependent Variable: N Class 3R

Gráfico 5 – Arvore de Decisão – 3R - CHAID



2. MÉTODO EXHAUSTIVE CHAID

Procede-se ao desenvolvimento da árvore de decisão com os grupos AIBER, AVECO e OTHERS. O teste Anova como o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3).

Classification

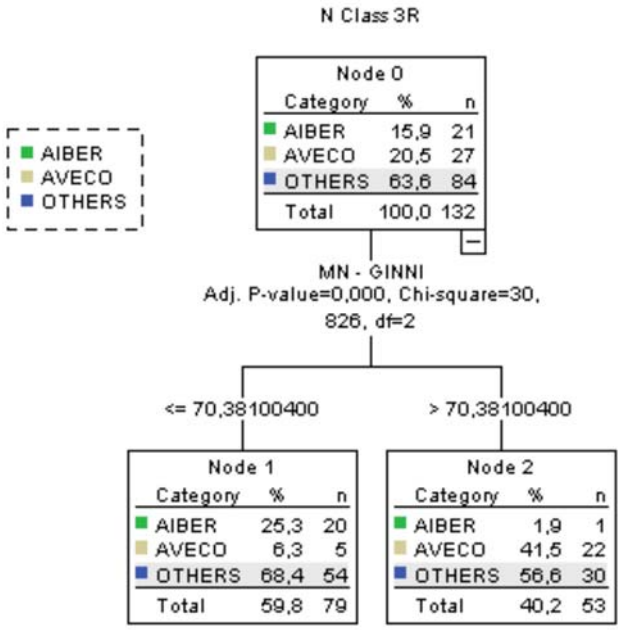
Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	0	0	21	,0%

AVECO	0	0	27	,0%
OTHERS	0	0	84	100,0%
Overall Percentage	,0%	,0%	100,0%	63,6%

Growing Method: CHAID

Dependent Variable: N Class 3R

Somente para OTHERS é que se encontra alguma relevância. O modelo não foi capaz de prever o pertencimento aos outros grupos.



3. MÉTODO CRT

Procede-se ao desenvolvimento da árvore de decisão com os grupos AIBER, AVECO e OTHERS. O teste Anova como o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3). Somente o OTHERS encontrou significância.

Classification

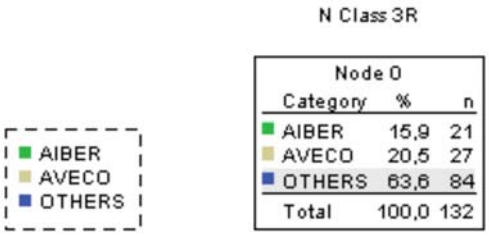
	Predicted			
Observed	AIBER	AVECO	OTHERS	Percent Correct
AIBER	0	0	21	,0%
AVECO	0	0	27	,0%
OTHERS	0	0	84	100,0%
Overall Percentage	,0%	,0%	100,0%	63,6%

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	0	0	21	,0%
AVECO	0	0	27	,0%
OTHERS	0	0	84	100,0%
Overall Percentage	,0%	,0%	100,0%	63,6%

Growing Method: CRT

Dependent Variable: N Class 3R



4. MÉTODO QUEST

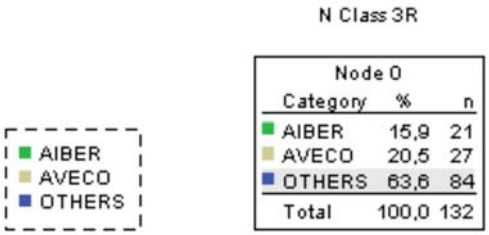
Procede-se ao desenvolvimento da árvore de decisão com os grupos AIBER, AVECO e OTHERS. O teste Anova como o IDH indica que o ordenamento do pior para o melhor IDH deve ser OTHERS (1), AIBER (2) e AVECO (3). Somente o OTHERS encontrou significância novamente.

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	0	0	21	,0%
AVECO	0	0	27	,0%
OTHERS	0	0	84	100,0%
Overall Percentage	,0%	,0%	100,0%	63,6%

Growing Method: QUEST

Dependent Variable: N Class 3R



e. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Os grupos de 4 clusters caracterizado pelo IDH aplicando o método CRT chega de 70% a 90% de capacidade explicativa. Ao contrário, neste estudo somente CHAID consegue desenvolver um modelo para os grupos AIBER, AVECO e OTHERS.

12. ANÁLISE DE CORRESPONDÊNCIA

Este trabalho tem como objetivo realizar uma análise da correspondência entre agrupamentos de países e as variáveis mais significativas, baseados em Árvores de Classificação para caracterizar dos agrupamentos de observações das oito principais variáveis que compõe as novas quatro Componentes Principais para o EIXO PESSOAS para 132 países, considerando os agrupamentos gerados pela análise discriminante e os grupos de países AIBER, AVECO e OTHERS. Seis das variáveis originais foram atualizadas. O critério de atualização foi possuírem estas variáveis bases de dados de 2015. A análise exploratória dos dados foi realizada através dos dados obtidos nas fontes de cada variável. Todas as variáveis originais foram normalizadas e as observações faltantes foram substituídas pela média das observações não faltantes. Empregou-se neste trabalho o Minitab16® para as análises estatísticas.

a. A ANÁLISE DE CORRESPONDÊNCIA

i. CONSTRUÇÃO DA BASE DE ANÁLISE DOS GRUPOS

ANOVA Grupos e Variáveis Chave

Procedeu-se a um teste Anova para verificar e caracterizar os graus de diferença entre os quatro clusters a que chegamos no dendograma. A variável escolhida para análise é o IDH em sua forma normalizada (HDI no original em inglês), que altamente relacionada ao SPI, um das variáveis encontradas no estudo das componentes principais e empregadas na construção dos clusters. Uma das vantagens do emprego do IDH em substituição ao SPI é que o primeiro é mais facilmente encontrado para diversos níveis, países, cidades, etc, enquanto o último é calculado somente para países. Piores IDHs concentram-se no grupo 2, com média de 29,91. Os grupos 1 e 3 (onde concentram-se os países do AIBER) tem IDHs mais próximos, com o primeiro ligeiramente superior e com menor dispersão. Pode-se esperar que alguns países do

grupo 3 possuam IDH superiores a elementos do grupo 1. Já o grupo 4, onde se classificaram os países do AVECO, possuem IDH médio de 86,9, muito superior também ao grupo 1, de 65,3 e ao 3, 59,5, e são mais homogêneos entre si.

One-way ANOVA: MN - (HDI)_1 versus 4 CLUSTERS MANHATTAN

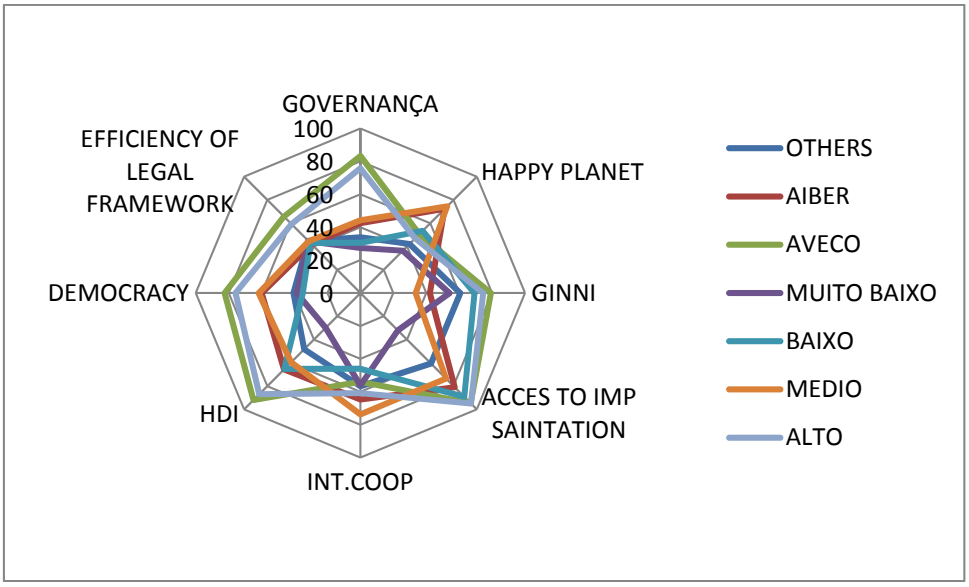
```
Source                 DF     SS     MS       F      P
4 CLUSTERS MANHATTAN    3  68679  22893  162,92  0,000
Error                 128  17987    141
Total                 131  86666

S = 11,85   R-Sq = 79,25%   R-Sq(adj) = 78,76%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
1      32  65,29   9,82                    (-*-)
2      42  29,91  15,74  (-*-)
3      17  59,53  10,80               (--*--)
4      41  86,93   8,63                                 (*-)
                          -------+---------+---------+---------+--
                                40        60        80       100

Pooled StDev = 11,85
```

Os grupos resultantes são agrupados por maior similaridade. Chega-se a algumas caracterizações. Nos grupos 1 e 2 predominam países de OUTRAS Regiões (OTHERS), sendo que no primeiro enquadram-se os países mais desenvolvidos de OTHERS, acrescido de alguns países de AIBER, e no grupo 2 os menos desenvolvidos. No grupo 3 encontra-se a maioria dos países de AIBER acrescidos de alguns países de OTHERS. E no grupo 4 encontram-se TODOS os países de AVECO, acrescidos de alguns de AIBER (Portugal, Espanha e Uruguai) e outros de OTHERS.

Portanto temos quatro grupos baseados em sua caracterização quanto ao IDH: ALTO, MÉDIO, BAIXO e MUITO BAIXO.
Outros três grupos originais são utilizados para a análise, caracterizados por sua definição original no presente trabalho: AIBER, com os países de origem ibérica na Eurpoa e América, AVECO, congregando as economias avançadas, e OTHERS, com os demis países dentre os 132 da amostra estudada.

O Gráfico de Radar apresenta uma visão comparativa dos grupos a partir das oito principais variáveis, permitindo uma comparação visual dos mesmos. AIBER, em espcial, caracteriza-se pela maior média no HAPPY PLANET INDEX, e AVECO por uma maior média de DEMOCRACIA, GOVERNANÇA e IDH, similar ao grupo ALTO, do qual muito dos países deAVECO fazem parte.

Gráfico de Radar – Todos os Grupos



b. ANÁLISE DE CORRESPONDÊNCIA NO MINITAB ® 16

Realiza-se análise de correspondência no MINITAB® com os sete grupos mencionados anteriormente e as oito Variáveis Chave, para trabalhar sua caracterização e distinção. Os dois primeiros eixos de análise explicam 76,5% das diferenças, e são representados no gráfico ROW PLOT.

O Eixo 1 tem como principal contribuição as variáveis HAPPY PLANET INDEX e INTERNATIONAL COOPERATION. O Eixo 2 possui como principais variáveis o índice de GINI e o índice de eficiência das estruturas legais para resolver disputas.

Verifica-se proximidade de posição entre os grupos MÉDIO e AIBER, entre ALTO e AVECO e em menor escala (maior distância) entr OTHER e BAIXO e MUITO BAIXO.

Tabela

Simple Correspondence Analysis: GOVERNANÇA; HAPPY PLANET; GINNI; ACCES TO IMP;

Contingency Table

	GOV	HAP	GIN	ACS	INC	HDI	DEM
OTHERS	33,900	42,210	60,630	60,600	56,680	48,350	40,640
AIBER	42,480	72,360	42,230	81,000	64,740	65,670	59,980
AVECO	83,250	50,580	79,250	94,740	53,910	91,890	82,480
MUITO BAIXO	27,420	36,420	54,140	32,140	56,110	29,910	38,610
BAIXO	30,480	53,380	69,340	89,180	46,020	65,290	34,990
MEDIO	44,150	74,470	33,530	73,530	73,750	59,530	61,730
ALTO	75,850	46,990	74,550	95,010	60,820	86,930	75,840
Total	337,530	376,410	413,670	526,200	412,030	447,570	394,270

	EFF	Total
OTHERS	44,840	387,850
AIBER	41,840	470,300
AVECO	65,720	601,820
MUITO BAIXO	44,310	319,060
BAIXO	43,040	431,720
MEDIO	44,480	465,170

ALTO	59,140	575,130
Total	343,370	3251,050

Chi-Square Distances

	GOV	HAP	GIN	ACS	INC	HDI	DEM	EFF
OTHERS	1,007	0,162	2,578	0,075	1,152	0,477	0,870	0,367
AIBER	0,825	5,890	5,183	0,313	0,442	0,013	0,152	1,235
AVECO	6,903	5,235	0,093	0,073	6,557	0,986	1,235	0,073
MUITO BAIXO	0,983	0,007	4,517	7,364	6,075	4,472	0,000	3,341
BAIXO	4,589	0,231	3,779	5,333	1,382	0,577	5,761	0,143
MEDIO	0,356	7,889	11,124	0,041	3,713	0,318	0,501	0,440
ALTO	4,362	5,769	0,026	0,040	1,999	0,759	0,532	0,042
Total	19,024	25,182	27,299	13,239	21,320	7,601	9,051	5,642

	Total
OTHERS	6,687
AIBER	14,054
AVECO	21,156
MUITO BAIXO	26,760
BAIXO	21,794
MEDIO	24,381
ALTO	13,528
Total	128,359

Relative Inertias

	GOV	HAP	GIN	ACS	INC	HDI	DEM	EFF	Total
OTHERS	0,008	0,001	0,020	0,001	0,009	0,004	0,007	0,003	0,052
AIBER	0,006	0,046	0,040	0,002	0,003	0,000	0,001	0,010	0,109
AVECO	0,054	0,041	0,001	0,001	0,051	0,008	0,010	0,001	0,165
MUITO BAIXO	0,008	0,000	0,035	0,057	0,047	0,035	0,000	0,026	0,208
BAIXO	0,036	0,002	0,029	0,042	0,011	0,004	0,045	0,001	0,170
MEDIO	0,003	0,061	0,087	0,000	0,029	0,002	0,004	0,003	0,190
ALTO	0,034	0,045	0,000	0,000	0,016	0,006	0,004	0,000	0,105
Total	0,148	0,196	0,213	0,103	0,166	0,059	0,071	0,044	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0180	0,4560	0,4560	******************************
2	0,0122	0,3091	0,7651	********************
3	0,0089	0,2255	0,9906	**************
4	0,0003	0,0081	0,9987	
5	0,0001	0,0013	1,0000	
Total	0,0395			

Row Contributions

					Component 1		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr
1	OTHERS	0,947	0,119	0,052	-0,017	0,017	0,002
2	AIBER	0,977	0,145	0,109	-0,139	0,648	0,156
3	AVECO	0,928	0,185	0,165	0,172	0,842	0,304
4	MUITO BAIXO	0,791	0,098	0,208	-0,076	0,069	0,032
5	BAIXO	0,075	0,133	0,170	0,009	0,002	0,001
6	MEDIO	0,949	0,143	0,190	-0,198	0,746	0,311
7	ALTO	0,931	0,177	0,105	0,141	0,844	0,195

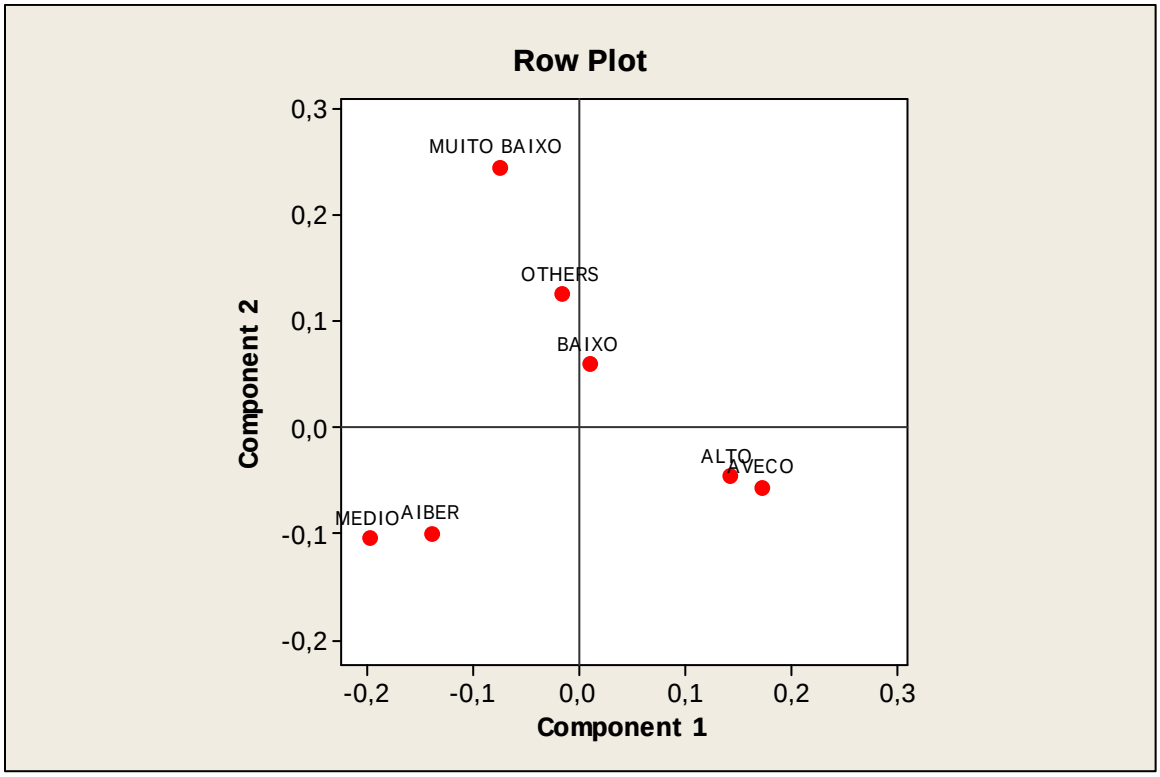
		Component 2		
ID	Name	Coord	Corr	Contr
1	OTHERS	0,127	0,929	0,157
2	AIBER	-0,099	0,328	0,116
3	AVECO	-0,055	0,086	0,046
4	MUITO BAIXO	0,246	0,721	0,486
5	BAIXO	0,061	0,073	0,040

6	MEDIO	-0,103	0,203	0,125
7	ALTO	-0,045	0,087	0,030

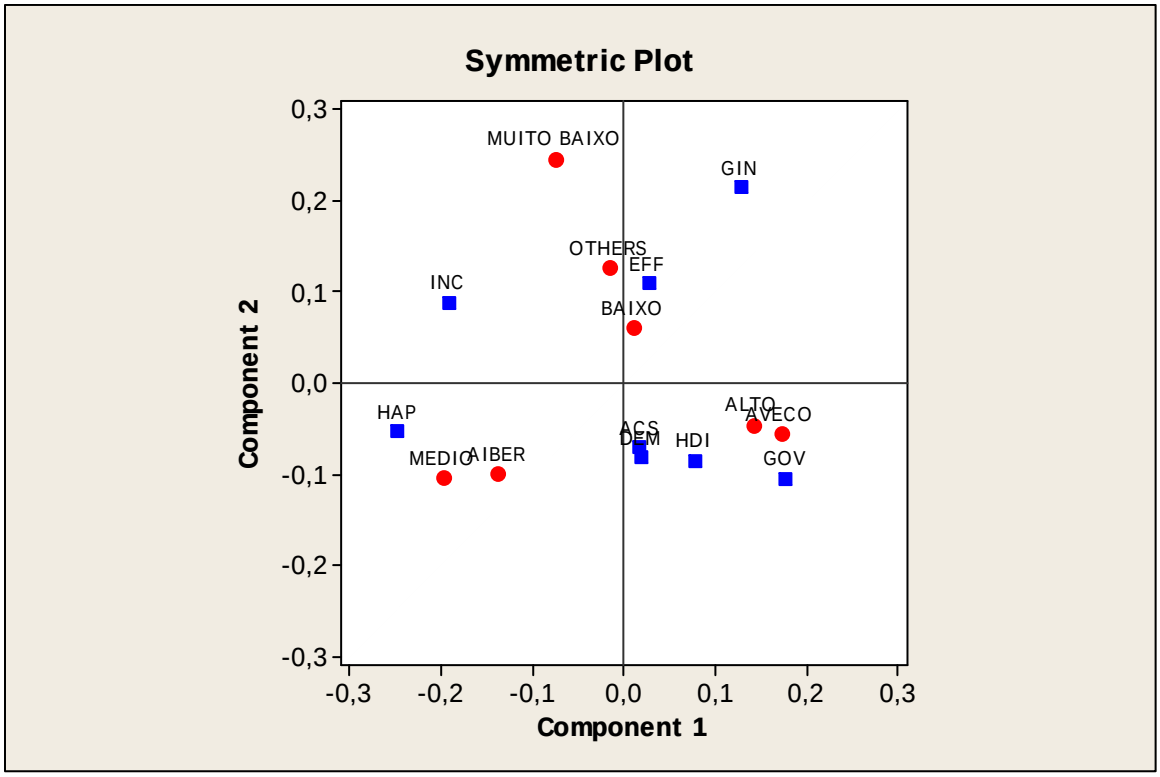
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	GOV	0,743	0,104	0,148	0,176	0,549	0,178	-0,105	0,194	0,093
2	HAP	0,958	0,116	0,196	-0,247	0,915	0,394	-0,053	0,043	0,027
3	GIN	0,946	0,127	0,213	0,128	0,247	0,115	0,215	0,699	0,481
4	ACS	0,213	0,162	0,103	0,017	0,011	0,003	-0,071	0,202	0,067
5	INC	0,860	0,127	0,166	-0,192	0,709	0,258	0,088	0,151	0,081
6	HDI	0,774	0,138	0,059	0,077	0,348	0,045	-0,085	0,426	0,082
7	DEM	0,301	0,121	0,071	0,019	0,015	0,002	-0,081	0,285	0,065
8	EFF	0,775	0,106	0,044	0,027	0,045	0,004	0,110	0,730	0,104

Row Plot



Symmetric Plot



O gráfico SYMMETRIC PLOT , que inclui grupos e variáveis reforça a caracterização dos primeiros a partir dos segundos. AIBER e MÉDIO distinguem-se pelo HAPPY PLANET INDEX, ALTO e AVECO por GOVERNANCE, HDI, DEMOCRACY e ACCESS TO IMPROVED SANITATION (todos altos), e OTHERS próximo a BAIXO e MUITO BAIXO, e próximo à variáveis GINI, INTERNATIONAL COOPERATION, EFFICIENCY.

c. CONCLUSÃO

Esta etapa de análise de correspondência verificou a possibilidade de caracterizar os grupos de países selecionados em função das principais variáveis e concluiu que é possível verificar-se graficamente algumas caracterizações: AIBER e MÉDIO possuem maior HAPPY PLANET INDEX, ALTO e AVECO maior GOVERNANCE, HDI, DEMOCRACY e ACCESS TO IMPROVED SANITATION (todos altos), e OTHERS, BAIXO e MUITO BAIXO, valores semelhantes para as variáveis INTERNATIONAL COOPERATION, EFFICIENCY e GINI.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER
SAÚDE E MEIO AMBIENTE

METODOS QUANTITATIVOS E QUALITATIVOS DA PESQUISA EMPIRICA – TRABALHO FINAL

Prof. Dr. Arnoldo José de Hoyos Guevara
Monitor Prof. Luciano Ferreira
Doutoranda. Cristina Mª Alcântara de Brito Vieitez

São Paulo, 22 de Junho de 2016.

SUMÁRIO

1.INTRODUÇÃO 5

2. ENTENDENDO OS DADOS 5

2.1 OS INDIVÍDUOS 7

2.1.1 A Tabela de Dados 9

3. ANÁLISE EXPLORATÓRIA DOS DADOS 10

3.1 ANÁLISE DESCRITIVA 10

3.2 SEGREGAÇÃO POR REGIÕES 11

3.3 ANÁLISE EXPLORATÓRIA DAS VARIÁVEIS 12

3.3.1 ÍNDICE DE DESEMPENHO AMBIENTE (EPI) (The Environmental Performance Index) 12

3.3.2 GOVERNANÇA Effectiveness Governance 13

3.3.3 DESPESAS DE SAÚDE, PÚBLICO (% do total das despesas de saúde) 2012 Health expenditure, public (% of total health expenditure) 2012 14

3.3.4 MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR EXTERIOR (mortes / 100.000) Outdoor air pollution attributable deaths (deaths/100,000) 16

3.3.5 MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR INTERIOR (mortes / 100.000) Indoor air pollution attributable deaths (deaths/100,000) 16

3.3.6 TAXA DE OBESIDADE (% DA POP.) Obesity rate (% of pop.) 17

3.3.7 MORTES POR DOENÇAS NÃO TRANSMISSÍVEIS ENTRE AS IDADES DE 30 E 70 (PROBABILIDADE DE MORRER) Non-communicable disease deaths between the ages of 30 and 70 (probability of dying) 18

3.3.8 ACESSO A INSTALAÇÕES SANITÁRIAS MELHORADAS (% DA POP.) Access to improved sanitation facilities (% of pop.) 19

3.3.9 ACESSO A ÁGUA CANALIZADA (% DA POP.) Access to piped water (% of pop.) 20

3.3.10 MORTES POR DOENÇAS INFECCIOSAS TRANSMISSÍVEIS (IDADE PADRONIZADA DE MORTES / 100.000) Deaths from infectious diseases (deaths/100,000) 21

3.3.11 TAXA DE MORTALIDADE INFANTIL (MORTES / 1.000 NASCIMENTOS) Child mortality rate (deaths/1,000 live births) 22

3.3.12 TAXA DE MORTALIDADE MATERNA (ÓBITOS / 100.000 NASCIDOS VIVOS) Maternal mortality rate (deaths/100,000 live births) 23

3.3.13 TAXA DE NATIMORTOS (MORTES / 1.000 NASCIMENTOS) Stillbirth rate (deaths/1,000 live births) 24

3.3.14 ÍNDICE DE PROGRESSO SOCIAL Social Progress Index 25

3.3.15 ÍNDICE DE PLANETA FELIZ (HPI) (Happy Planet Index 2012) 26

3.3.16 ÍNDICE GINNI (World Bank estimate) 27

3.3.17 ÍNDICE DE SAÚDE DO OCEANO (OHI) Ocean Health Index 28

3.3.18 IDH - 2013 30

3.3.19 RECURSOS EXTERNOS PARA A SAÚDE (% da despesa total em saúde) 2012 - External resources for health (% of total expenditure on health) 31

3.3.20 A EXPECTATIVA DE VIDA (ANOS) 31

3.4. NORMALIZAÇÃO DOS DADOS 32

3.5. MATRIZ DE CORRELAÇÃO 44

3.6. GRÁFICOS DE DISPERSÃO 47

3.7. ANÁLISE DE CLUSTERS E DENDOGRAMAS 50

3.8. REGRESSÃO LINEAR 52

3.9. ANOVA 61

3.10. ANÁLISE AMOSTRAL 103

3.10.1 Deaths for infectious desease para Amostra de 25 países. Ver Quadro 4 103

3.10.2 Deaths for infectious desease para Amostra de 50 países. Ver Quadro 5 104

3.10.3 Deaths for infectious desease para Amostra de 100 países. Ver Quadro 6 105

3.10.4 Deaths for Infectious Deseases: ANOVA 107

3.10.5 Outdoor Air Pollution: MESMOS PROCEDIMENTOS PARA Indicador de Bem-Estar Social... 110
3.11 CRIAÇÃO DE NOVOS COMPONENTES 113
3.12 ANÁLISE DE COMPONENTES PRINCIPAIS 116
3.13 CLASSIFICAÇÃO NÃO SUPERVISIONADA, ANÁLISE DISCRIMINANTE..... 131
3.13.1 CLASSIFICAÇÃO NÃO SUPERVISONADA 132
3.13.2 ANÁLISE DISCRIMINANTE: 4 GRUP OR VERSUS PC1; PC2; PC3; PC4..... 139
3.14 ANÁLISE MULTIVARIADA MÉTODOS LINEAR E QUADRÁTICO..... 142
3.15 CLASSIFICAÇÃO SUPERVISIONADA E REGRESSÃO LOGÍSTICA..... 151
3.16 CLASSIFICAÇÃO SUPERVISIONADA E ÁRVORES DE DECISÃO 157
3.16.1 O SPSS 157
3.16.2 ANÁLISE DOS NOVOS COMPONENTES..... 158
3.16.3 ÁRVORE DE DECISÕES..... 164
3.17 ANÁLISE DE CORRESPONDÊNCIAS 167
3.18. ANÁLISE PAÍSES IBERO AMERICANOS 181
4. CONSIDERAÇÕES FINAIS..... 196

1.INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados relativos ao tema SAÚDE E MEIO AMBIENTE, analisando-se um conjunto de vinte e três variáveis previamente selecionadas que refletem os dados sobre os principais indicadores que influenciam em tal tema, considerando 132 países. Este estudo, portanto, apresenta conclusões preliminares, identificando correlações entre as variáveis que compõem o tema, em especial, sobre as mais relevantes após as regressões. Foram focalizados os países Ibero-Americanos (RIBER), denominados assim pelo material ORIBER do GUIA GPS: Gestão Pública Sustentável, uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica de São Paulo. O objetivo final é um *ranking* entre eles, compreendendo os fatores que mais influenciam na Saúde e Meio Ambiente da sua população.

A Organização Mundial de Saúde (OMS) estabeleceu um marco normativo para indicadores de saúde ambiental, indicando as directrizes essencias para: consideração ao contexto sociodemográfico (população, densidade, taxa de crescimiento demográfico, expectativa de vida e mortalidade infantil, nível de poluição do ar, (quantidade de ar interno e externo, mortalidade infantil devido a enfermedades respiratórias agudas e capacidade de moniàtorar a qualidade do ar, saneamento básico em função do tratamento dos residuos organicos e de doenças como a diarréia, associadas a baixos padrões de higiene e qualidade da agua, acesso a água potable, enfermedades transmitidas por insetos (por picadura de mosquito), vivenda em função da porcentagem de famiílias morando em áreas invadidas, não regulamentadas e inseguras, acidentes domésticos e planificação urbana.

2. ENTENDENDO OS DADOS

Antes da apresentação dos indivíduos desta análise que envolve 132 países e os indicadores de Saúde e Meio Ambiente selecionados, apresenta-se o tema abordado.

Entende-se por Meio Ambiente, segundo o Guia GPS:

- ambiente natural: água, terra, ar;
- ambiente ocupacional: das indústrias químicas e outros locais de trabalho;
- ambiente de consume: compreendido por alimentos e medicamentos;
- ambiente social e cultural: compreende o estilo e hábitos de vida.

Vale destacar o desafio citado no GPS e sua proposta de solução:
Desafio: "A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?"

> **Solução:** "O Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do

> desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."

O Guia sugere uma abordagem sistêmica e um planejamento estratégico. As diretrizes estabelecidas devem constar no Plano Director e no Plano Plurianual. Para que ocorra, deve haver uma adequação à realidade de cada país ou região.

Os países europeus Portugal e Espanha, por exemplo, devem ter semelhanças com os demais, porém também muitas idiossincrasias. O mesmo deve ser dito sobre os países do Cone Sul como Argentina, Chile e Uruguai e assim por diante. O ORIBER relembra que países grandes e de destaque, com alto PIB per cápita como Brasil, México, Colombia e Venezuela, guardam muitas características internas distintas quanta à sua urbanização. É histórico e sabido que os demais países da América do Sul e Central apresentam um nível de desenvolvimento socioeconómico ainda inferior.

Também é importante dizer que ainda se pode encontra na região sérios problemas de mando politico e governança, bem como de estabilidade institucional e democrática. A Gestão Pública ainda tem muito a evoluir.

Os dados devem ser entendidos neste contexto. De acordo com a Organização Mundial de Saúde (OMS), o tema que será debatido aqui, apresenta uma dificuldade ímpar de especificação devido à variedade de interconexões e das diferentes prioridades e intereses dos *stakeholders*.

Um dos fatores de fundamental importância diz respeito à falta de infraestructura da saúde pública dos países Ibero-Americanos. Não é de se estranhar que a saúde ambiental esteja correlacionada com o desenvolvimento e riqueza dos mesmos, e mesmo de gerir bem tal riqueza.

De qualquer forma, baseado nos dados da OMS, o Projeto ORIBER acredita que apesar das diferenças socioculturais e econômicas, parecer ser que as semelhanças permitem a criação de uma pauta comum de respuestas para a superação das dificultades dos países da Região.

2.1 OS INDIVÍDUOS

Os indivíduos desta análise são os 132 países analisados por uma série de indicadores de referência internacional. Os dados analisados de cada país são as variáveis que descrevemos a seguir.

- Índices sintéticos: São sete: Índice De Progresso Social, Índice De Desenvolvimento Humano – IDH, Índice De Governança, Índice De Desempenho Ambiental – EPI, Índice De Planeta Feliz (HPI), Índice De Saúde Dos Oceanos – OHI e o Índice De Desigualdade Social ou Ginni.

- Indicadores ou Variáveis Componentes: Inclui 13 Variáveis Como: Taxa De Mortalidade Materna, Taxa De Natimortos, Taxa De Mortalidade Infantil, Mortes Por Doenças Infecciosas, Acesso a Água Canalizada, Acesso A Instalações Sanitárias Melhoradas, Mortes Atribuíveis À Poluição Do Ar Interior, Expectativa De Vida (Anos), Mortes Por Doenças Não Transmissíveis Entre As Idades De 30 E 70 (Probabilidade de Morrer), Taxa De Obesidade, Mortes Atribuíveis À Poluição Do Ar Exterior, Despesas De Saúde Pública, Recursos Externos Para A Saúde.

Quadro 1. As **20** Variáveis referentes ao Eixo temático Saúde e Meio Ambiente são:

1.	**TAXA DE MORTALIDADE MATERNA** (ÓBITOS / 100.000 NASCIDOS VIVOS) *Maternal mortality rate (deaths/100,000 live births)*
2.	**TAXA DE NATIMORTOS** (MORTES / 1.000 NASCIMENTOS) *Stillbirth rate (deaths/1,000 live births)*
3.	**TAXA DE MORTALIDADE INFANTIL** (MORTES / 1.000 NASCIMENTOS) *Child mortality rate (deaths/1,000 live births)*
4.	**MORTES POR DOENÇAS INFECCIOSAS** (MORTES / 100.000) *Deaths from infectious diseases (deaths/100,000)*
5.	**ACESSO A ÁGUA CANALIZADA** (% DA POP.) *Access to piped water (% of pop.)*
6.	**ACESSO A INSTALAÇÕES SANITÁRIAS MELHORADAS** (% DA POP.) *Access to improved sanitation facilities (% of pop.)*
7.	**MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR INTERIOR** (mortes / 100.000) *Indoor air pollution attributable deaths (deaths/100,000)*
8.	**EXPECTATIVA DE VIDA** (ANOS) *Life Expectancy*
9.	**MORTES POR DOENÇAS NÃO TRANSMISSÍVEIS ENTRE AS IDADES DE 30 E 70** (Probabilidade de morrer) *(% of pop.)* *Non-communicable disease deaths between the ages of 30 and 70*
10.	**TAXA DE OBESIDADE** (% Da pop.) *Obesity rate (% of pop.)*
11.	**MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR EXTERIOR** (Mortes / 100.000) *Outdoor air pollution attributable deaths* (deaths/100,000)
12.	**DESPESAS DE SAÚDE PÚBLICA** (% do total das despesas de saúde) *Health expenditure, public (% of total health expenditure) 2012*
13.	**RECURSOS EXTERNOS PARA A SAÚDE** (% da despesa total em saúde) *External resources for health (% of total expenditure on health) 2012*
14.	**IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO** De 0 a 1 (Mede renda, saúde e educação). *Human development index 2013*
15.	**GOVERNANÇA** Indicador Nominal, mede a capacidade de um país de eleger, monitorar e substituir seus governantes. *Effectiveness Governance*
16.	**IPS – ÍNDICE DE PROGRESSO SOCIAL** De 0 a 100 *Social Progress Index (SPI)*
17.	**ÍNDICE DE DESEMPENHO AMBIENTE** Indicador nominal, medido pela mortalidade infantil (de 1 a 5 anos), qualidade do ar, nível de água tratada e sanitização adequada). *The Environmental Performance Index (EPI)*
18.	**ÍNDICE DE PLANETA FELIZ** Indicador nominal que mede longevidade, felicidade e vida sustentável percebida. *Happy Planet Index (HPI) 2012*
19.	**ÍNDICE DE SAÚDE DO OCEANO** *Ocean Health Index (OHI)*
20.	**ÍNDICE ÍNDICE DE DESIGUALDADE SOCIAL** Indicador nominal que mede de desigualdade social: distribuição de renda ou riqueza. *GINNI (Indicador Nominal)*

Fonte: GPS, a partir de dados do Banco Mundial.

2.1.1 A Tabela de Dados

Em função da extensão da tabela 1, vamos representar aqui **apenas alguns países** e **algumas** variáveis selecionadas para corroborar o Quadro 1.

Tabela 1: Países e variáveis selecionadas

País	Social Progress Index(índice de progresso social)	Access to piped water (% of pop.)(água encanada)	Rural vs. urban access to improved water source (absolute difference between % of pop.)(agua tratada zona rural e urbana)	Access to improved sanitation facilities (% of pop.)(instalações sanitárias)
Albania	69,13	81,8	1,8	93,9
Algeria	59,13	73,7	6,0	95,1
Angola	39,93	20,0	31,6	58,7
Brazil	69,97	91,7	15,1	80,8
Bulgaria	70,24	96,8	0,7	100,0
Burkina Faso	47,33	6,9	22,3	18,0
Canada	86,95	87,7	1,0	99,8
El Salvador	64,70	71,8	12,8	70,0
Estonia	81,28	90,3	2,3	97,9
Finland	86,91	99,4	0,0	100,0
France	81,11	100,0	0,0	100,0
Hungary	73,87	94,3	0,0	100,0
Iceland	88,07	100,0	0,0	100,0
India	50,24	25,2	6,8	35,1
Kuwait	70,66		0,0	100,0
Paraguay	62,65	65,6	33,4	70,8
Peru	66,29	75,3	24,7	71,6
Philippines	65,86	42,7	0,7	74,2
Poland	77,44	98,0		89,3
Portugal	80,49	99,7	0,0	100,0
Russia	60,79	82,0	6,6	70,4
Rwanda	49,46	3,6	13,1	61,3
Saudi Arabia	64,38	66,7	0,0	100,0
South Africa	62,96	68,8	19,7	74,0
Spain	80,77	99,3	0,1	100,0
Sri Lanka	59,71	29,5	7,3	91,1
Venezuela	63,78	86,8	19,1	90,9
Yemen	40,23	40,2	25,4	53,0
Zambia	49,88	15,0	35,9	42,1

Fonte: Dados Guia GPS.

3. ANÁLISE EXPLORATÓRIA DOS DADOS

Depois de haver apresentado o conjunto de 20 (vinte) – Quadro 1 – variáveis selecionadas previamente, com o objetivo de explicar as relações entre saúde e meio ambiente num universo de 132 países (Tabela 1), se analisará cada uma das variáveis separadamente quanto a sua forma de distribuição, os valores atípicos, medidas de centro e dispersão, através do *Display Descriptive Statistics* que está no *link* principal *Basic Statistics* do *software* Minitab.

Para tal se contará com o auxílio de gráficos (*pie chart*, barras, histogramas, dendogramas, curvas de densidade) e tabelas de medidas numéricas (média, mediana, quartis, desvio-padrão, variância, intervalo de confiança e teste de normalidade de Anderson-Darling).

Segue-se com a aplicação do *link* principal *Data*, acessando o *Code – Numeric to Numeric*, sendo que este recurso permite que os dados ausentes que aparecem na tabela com o símbolo asterisco (*) seja substituído pelo valor determinado que é a média encontrada.

Em seguida, se apresentará os gráficos de dispersão e se procederá na sequência à normalização dos mesmos que é um ajuste com vista a afastar os *outliers*, ou seja, os pontos fora da curva – é como se dos dados analisados não possuíssem uma distribuição normal então se faz um ajuste para que todos fiquem dentro de uma certa normalidade.

Apresentar-se-ão os valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas. Para essa etapa será usado o comando do MINITAB denominado *Calculator* e será aplicada a fórmula (a seguir) indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.

$$(c9-MIN(c9))/(MAX(c9)-MIN(c9))*100$$

Buscar-se-á comparar as análises efetuadas através da correlação e posteriormente apresentar-se-á dendogramas.

Nesta primeira parte, se apresentará o conjunto dos gráficos gerados pelo recurso *Graphical Summary* que está no item *Basic Statistic*, que por sua vez está no item *Static* na barra de ferramentas do MINITAB.

3.1 ANÁLISE DESCRITIVA

A estatística descritiva é a etapa inicial da análise utilizada para descrever e resumir os dados. Nela é possível ter a ideia de como os dados se apresentam. Também é possível identificar se o banco de dados possui valores ausentes (*missings values*). A seguir, será descrito o passo-a-passo.

A Tabela 2 apresentou a estatística descritiva de todas as variáveis do estudo. Tem-se, por exemplo, que a média da expectativa de vida entre os países é de 70 anos, com mínimo de 46 e máximo de 82 anos.

Pode-se observar também que a análise acusou valores ausentes em algumas variáveis, indicado pelo N*. Para solucionar este fato, no lugar desses valores, serão colocadas as respectivas médias de cada variável e será apresentada uma nova tabela de análises descritivas (Tabela 3).

Tabela 2: Análise Descritiva

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
Maternal mortality rate	132	0	161,6	19,2	221,0	2,0	12,0	61,5	240,0	1100,0
Stillbirth rate (deaths/	131	1	12,878	0,850	9,733	2,000	4,000	10,000	21,000	47,000
Child mortality rate (de	132	0	33,83	3,18	36,49	2,30	6,30	17,60	54,75	163,50
Deaths from infectious d	132	0	260,3	28,4	326,0	11,0	32,3	90,5	384,5	1255,0
Access to piped water (%	129	3	63,98	3,03	34,36	2,53	24,72	75,33	95,34	100,00
Access to improved sanit	130	2	73,75	2,55	29,09	9,58	52,97	88,00	98,66	100,00
Indoor air pollution att	131	1	46,00	4,19	48,00	0,00	3,69	24,87	81,62	222,32
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319	73,398	76,905	82,695
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000	23,000	28,000	38,000
Obesity rate (% of pop.)	132	0	17,373	0,822	9,444	1,100	7,225	19,000	24,275	42,800
Outdoor air pollution at	130	2	15,75	1,12	12,76	0,00	7,00	12,00	20,25	67,00
Health expenditure, publ	132	0	57,47	1,49	17,09	18,01	45,02	58,98	71,06	94,15
External resources for h	93	39	10,09	1,41	13,61	0,00	0,56	3,65	15,01	55,69
IDH - 2013	125	7	0,6894	0,0144	0,1611	0,3370	0,5595	0,7220	0,8200	0,9440
1. GovernanÁa	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163	42,463	49,287	64,036
OHI	98	34	64,975	0,881	8,722	45,050	59,645	65,140	71,497	82,140
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780	38,160	44,550	63,140

Fonte: Elaborado pela autora

Na Tabela 3, com os valores ausentes substituídos pela média, apenas a contagem de valores ausentes é para sofrer alteração, as outras permanecem inalteradas.

Tabela 3: Análise Descritiva após substituição dos valores ausentes

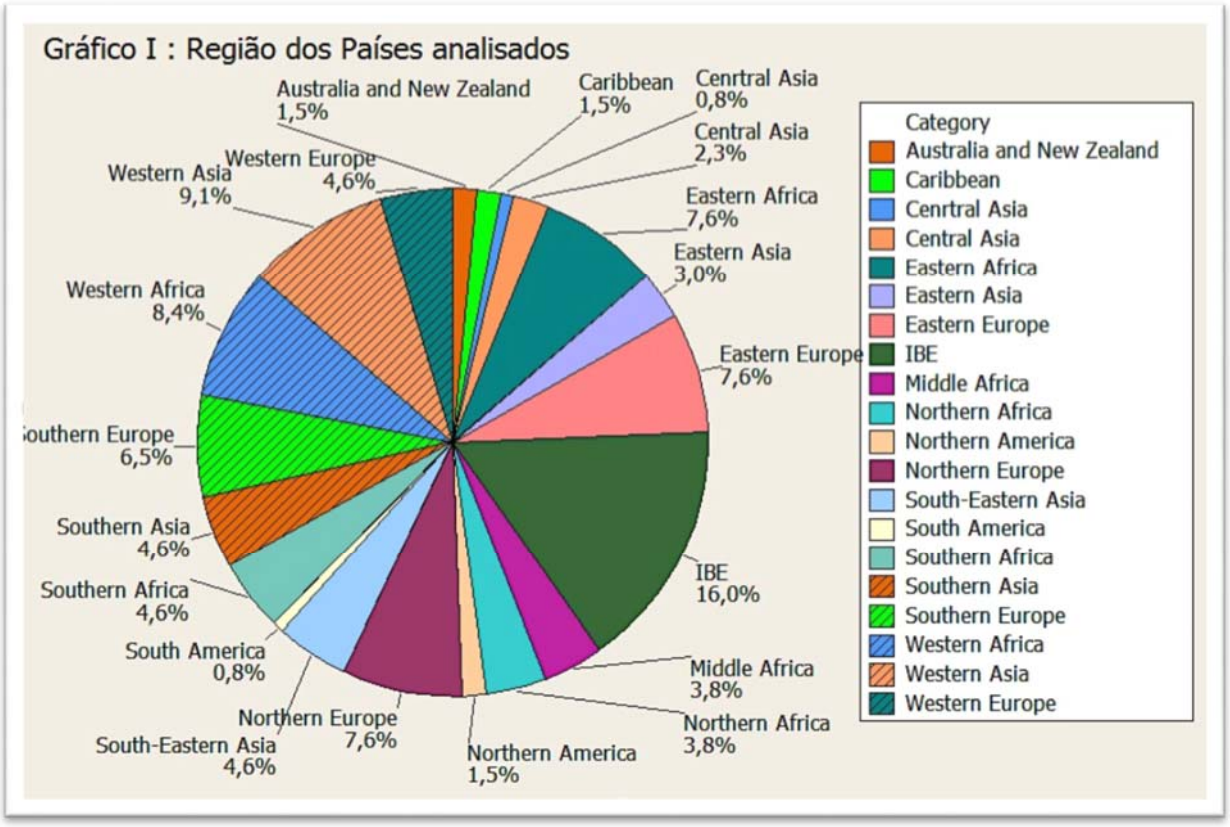
Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
Maternal mortality rate	132	0	161,6	19,2	221,0	2,0	12,0	61,5	240,0	1100,0
Stillbirth rate (deaths/	132	0	12,878	0,844	9,696	2,000	4,000	10,000	20,750	47,000
Child mortality rate (de	132	0	33,83	3,18	36,49	2,30	6,30	17,60	54,75	163,50
Deaths from infectious d	132	0	260,3	28,4	326,0	11,0	32,3	90,5	384,5	1255,0
Access to piped water (%	132	0	63,98	2,96	33,97	2,53	26,30	74,80	95,03	100,00
Access to improved sanit	132	0	73,75	2,51	28,87	9,58	53,00	86,37	98,54	100,00
Indoor air pollution att	132	0	45,73	4,17	47,92	0,00	3,70	24,42	81,59	222,32
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319	73,398	76,905	82,695
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000	23,000	28,000	38,000
Obesity rate (% of pop.)	132	0	17,373	0,822	9,444	1,100	7,225	19,000	24,275	42,800
Outdoor air pollution at	132	0	15,75	1,10	12,66	0,00	7,00	12,50	19,75	67,00
Health expenditure, publ	132	0	57,47	1,49	17,09	18,01	45,02	58,98	71,06	94,15
External resources for h	132	0	10,088	0,993	11,404	0,000	1,305	10,090	10,090	55,685
IDH - 2013	132	0	0,6894	0,0136	0,1567	0,3370	0,5700	0,7155	0,8147	0,9440
1. GovernanÁa	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
Happy Planet Index	132	0	42,673	0,778	8,936	22,591	36,587	42,627	49,175	64,036
OHI	132	0	64,975	0,653	7,505	45,050	60,833	64,975	69,545	82,140
GINNI Index	132	0	39,024	0,705	8,097	24,820	33,065	39,024	43,553	63,140

Fonte: Elaborado pela autora

3.2 SEGREGAÇÃO POR REGIÕES

A seguir, será feito um agrupamento dos 132 países em 19 regiões dentro dos continentes: Europeu, Africano, Oceania, Asiático e Americano. O Gráfico 1 apresenta a distribuição dos países nas regiões mencionadas.

Gráfico 1: Agrupamentos dos países em regiões



Fonte: Elaborado pela autora

A análise primordial de este estúdio se centrará na região IBE 16%, dos países Ibero Americanos.

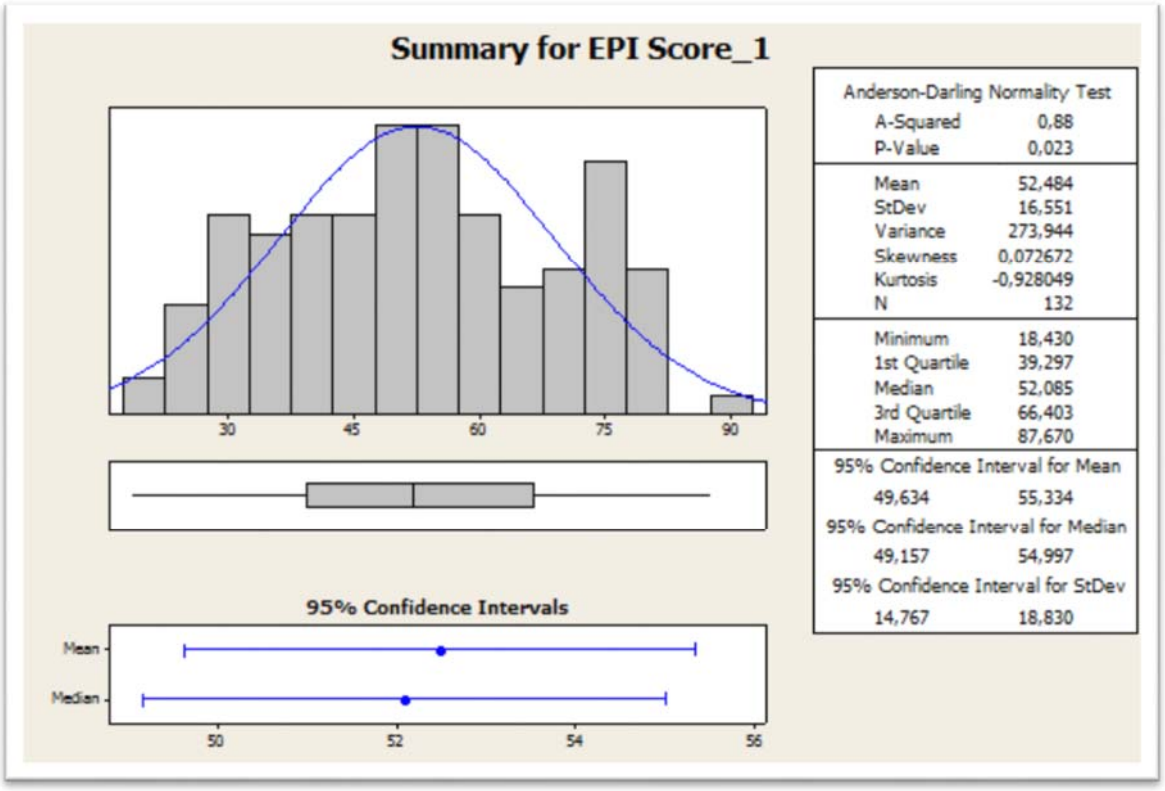
3.3 ANÁLISE EXPLORATÓRIA DAS VARIÁVEIS

A análise exploratória das variáveis é realizada para se conhecer um pouco mais das variáveis em termos de distribuição.

3.3.1 ÍNDICE DE DESEMPENHO AMBIENTE (EPI) (*The Environmental Performance Index*)

O Índice de Desempenho Ambiental (EPI) classifica como os países bem executam as questões ambientais de alta prioridade em duas grandes áreas políticas: proteção da saúde humana e proteção do ecossistema ambiental. Indicador EPI: %

Gráfico 2: Descritivo - EPI



Fonte: Autor – Dados Guia GPS

Análise:
O gráfico 2 mostra que índice EPI tem média de 52,484 (intervalo de confiança de 95% entre 49,634 e 55,334) e desvio padrão de 16,551(intervalo de confiança de 95% entre 14,767 e 18,830). Neste caso se nota uma distribuição próxima a uma distribuição normal, com média e mediana muito próximas. Isso significa que a maioria dos países encontra-se em torno da média, sendo ela um pouco superior a 50%. Nota-se um pico de países com a nota próxima a 75%, representado maioritariamente pelos países europeus, destacando a Suíça com melhor desempenho no grupo (máxima).

3.3.2 GOVERNANÇA *Effectiveness Governance*

Consiste nas tradições e instituições pelas quais a autoridade de um país é exercido. Isso inclui o processo pelo qual os governos são selecionados, monitorados e substituídos; a capacidade do governo para formular e implementar com eficácia políticas sólidas; e o respeito dos cidadãos e do Estado pelas instituições que governam as interações econômicas e sociais entre eles.

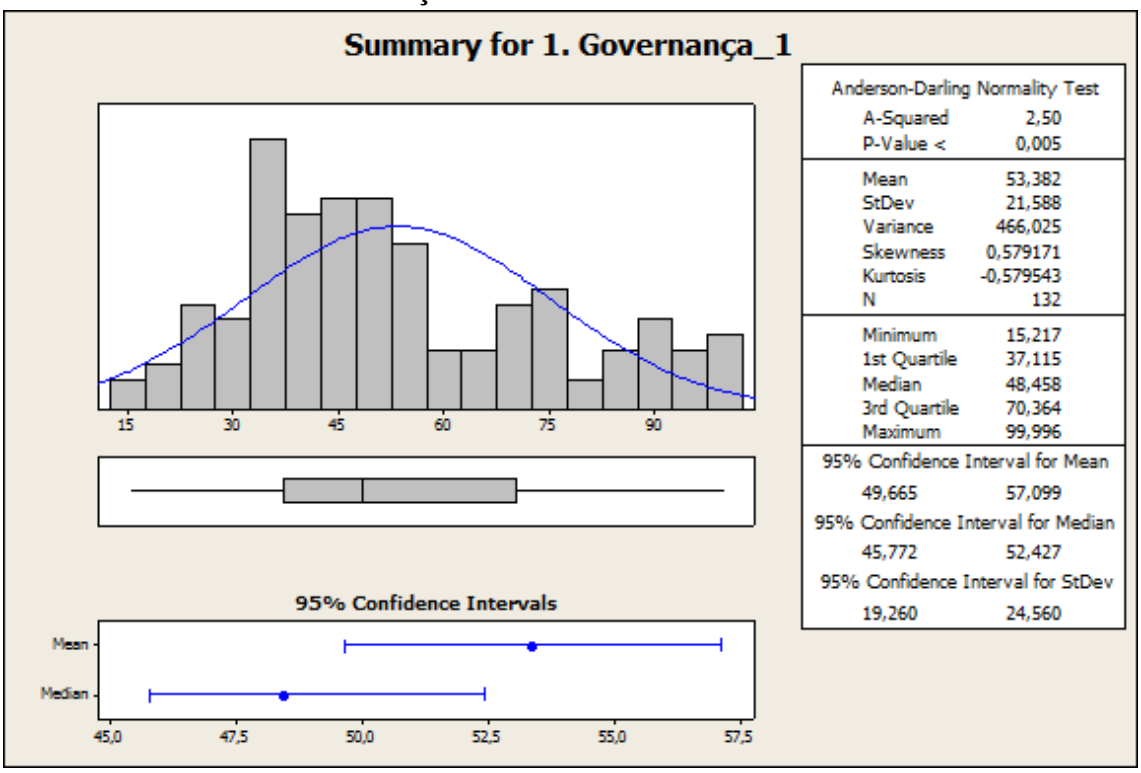
Os Indicadores de Governança Mundial segundo o relatório do Banco Mundial sobre seis grandes dimensões de governança para 215 países durante o período de 1996-2013 refere-se a:

- Voz e Responsabilização

- Estabilidade política
- Qualidade regulatória
- Estado de Direito
- Controle da Corrupção

Em resumo é um % que indica como os governos são indicados; como a autoridade do país é exercida, como as politicas são implementadas e qual a capacidade de formulá-las.

Gráfico 3: Descritivo - Governança



Fonte: Autor – Dados Guia GPS.

Análise:

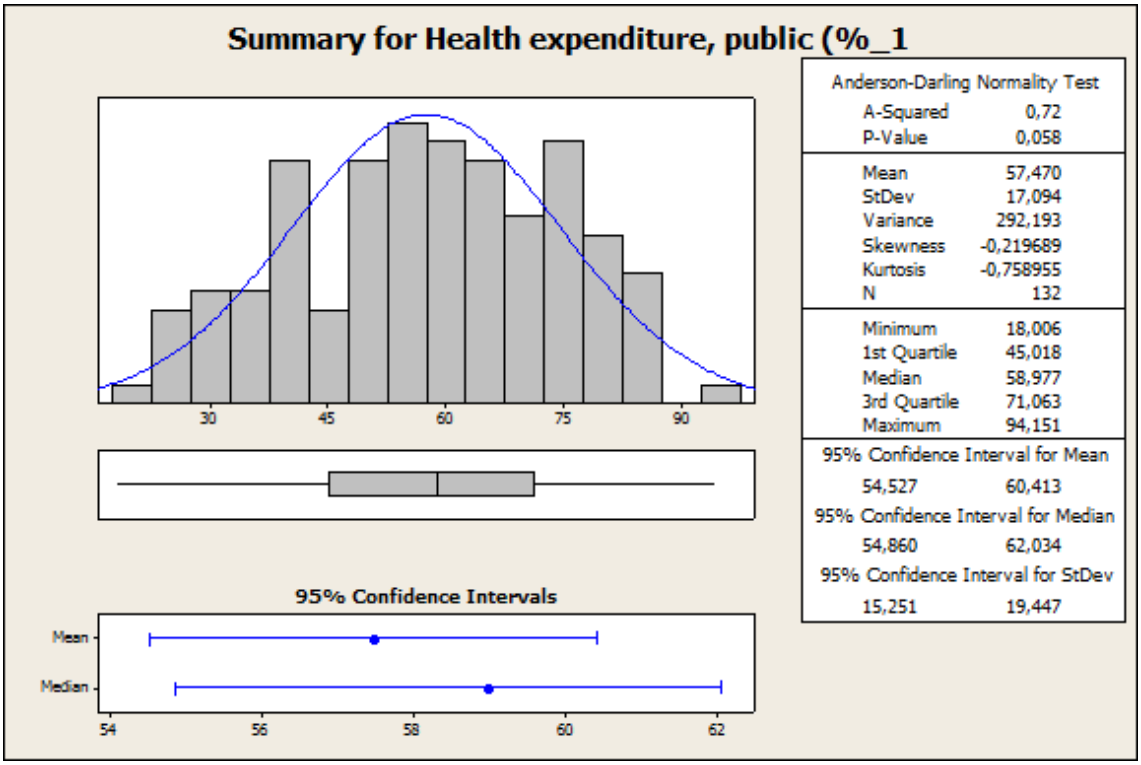
O Gráfico 3, mostra que índice em questão tem média de 53,382 (intervalo de confiança de 95%, entre 49,665 e 57,099) e desvio padrão de 21,588 (intervalo de confiança de 95%, entre 19,260 e 24,560). Pode-se observar pelo histograma que a maior parte dos países tem o índice de 0 a 50. Neste caso nota-se a mediana abaixo da média, indicando que a maioria dos países estão com resultado abaixo da média (distribuição não uniforme, com mais países com notas mais baixas e menos países com notas muito mais altas) gerando um deslocamento para o lado esquerdo. Do lado superior, mais à direita, mais uma vez os países Europeus, Austrália, Nova Zelândia, Japão e Chile. Do lado com menores resultados os países da África e Ásia Central (em sua maioria).

3.3.3 DESPESAS DE SAÚDE, PÚBLICO (% do total das despesas de saúde) 2012
Health expenditure, public (% of total health expenditure) 2012

Consistem em despesas correntes e de capital dos orçamentos do governo (central e local), empréstimos e subsídios (incluindo doações de agências internacionais e organizações não

governamentais) externos e fundos de seguros de saúde sociais (ou obrigatórios). Despesa total em saúde é a soma das despesas de saúde público e privado. Ela abrange a prestação de serviços de saúde (preventivos e curativos), atividades de planejamento familiar, atividades de nutrição e ajuda de emergência designada para a saúde, mas não inclui o fornecimento de água e saneamento do Banco Mundial.

Gráfico 4: Descritivo – *Health Expenditure*



Fonte: Autor – Dados Guia GPS.

Análise:

A média da Despesa com Saúde é de 57,470 (intervalo de confiança de 95%, entre 54,527 e 60,413) nos países pesquisados, e desvio padrão de 17,094 (intervalo de confiança de 95%,15,251 e 19,447), conforme apresentado no Gráfico 4.

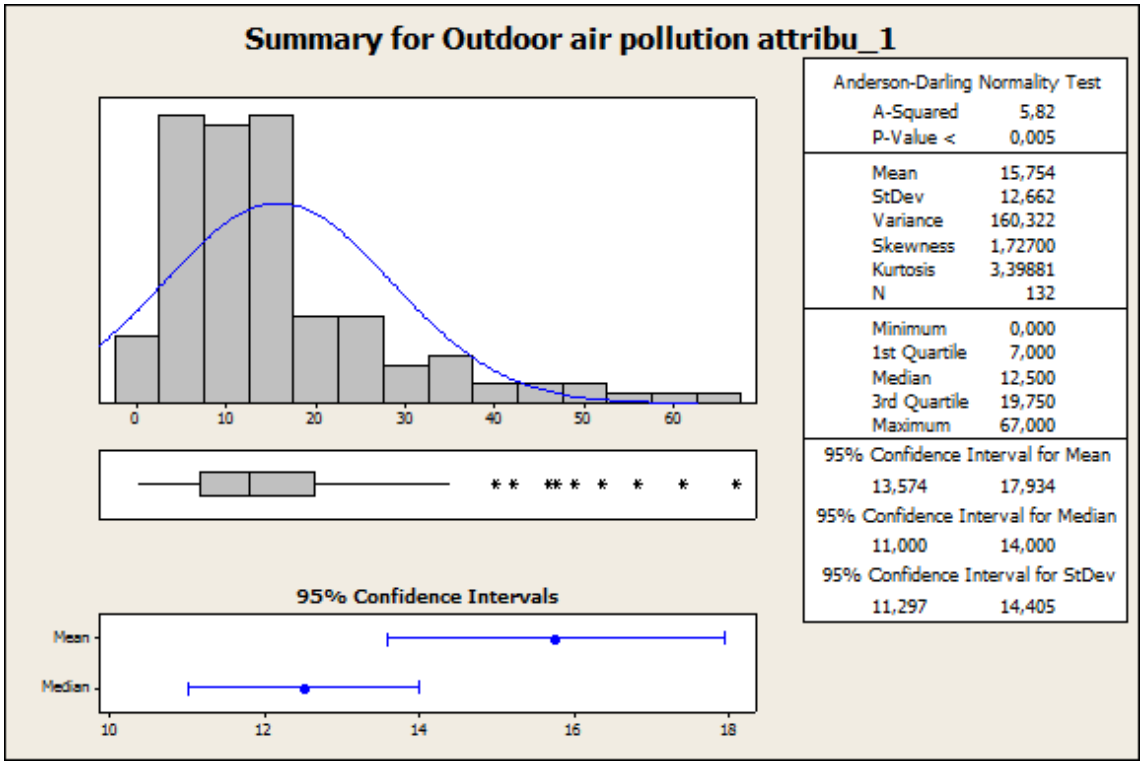
Pelo teste de Anderson-Darling para normalidade, os dados podem ser considerados próximos a normais (p-valor = 0,058), para o nível de significância de 0,05. Pela curva se percebe, porém com baixa frequência no substrato de 45 e um pouco mais alta no substrato de 40. Média e mediana muito próximas.

Não há um padrão nos países que gastam mais com saúde. Tantos países ricos, quanto podres, podem ter altas despesas com saúde.

3.3.4 MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR EXTERIOR (mortes / 100.000) Outdoor air pollution attributable deaths (deaths/100,000)

Este indicador refere-se ao número de mortes resultantes de emissões da atividade industrial, das residências, carros e caminhões, expressa como a taxa por 100.000 pessoas, segundo a Organização Mundial de Saúde.

Gráfico 5: Descritivo – *Outdoor Air Pollution*



Fonte: Autor – Dados Guia GPS.

Análise:

O índice de mortes de devido à poluição do ar exterior, tem média 15,754 (intervalo de confiança de 95%, entre 13,574 e 17,934) e desvio padrão de 12,662 (intervalo de confiança de 95%, entre 11,297 e 14,405), conforme Gráfico 5.

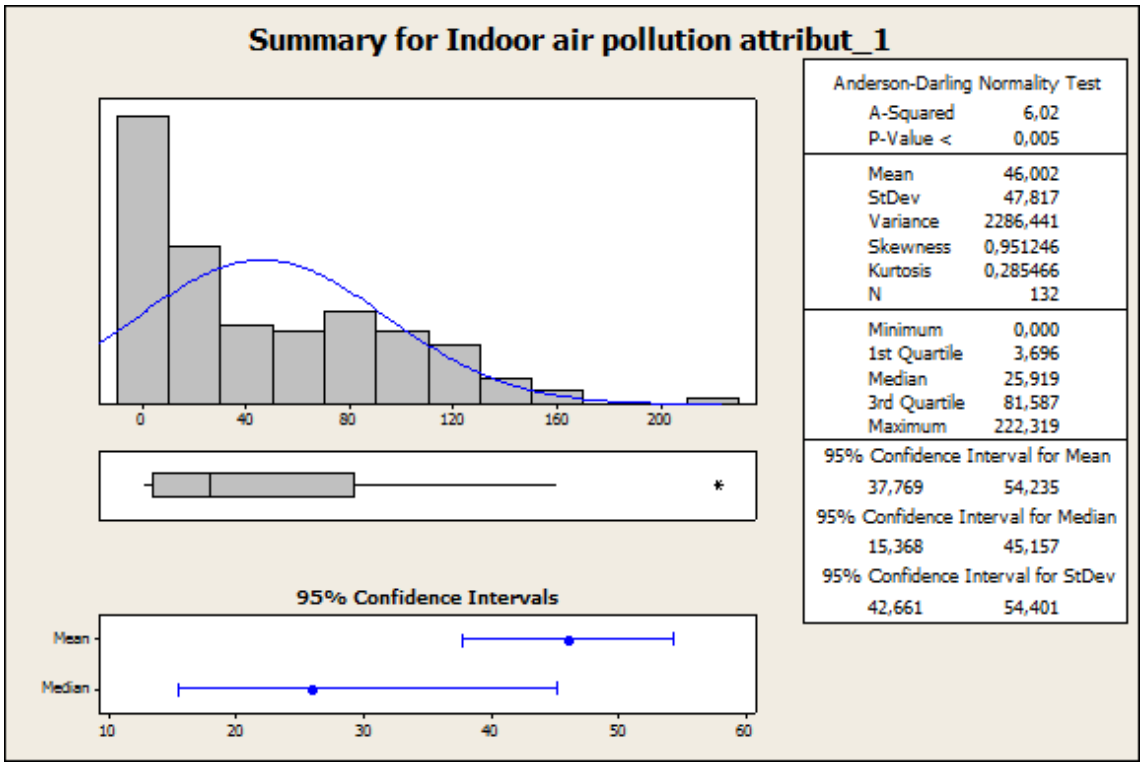
No caso, a mediana menor que a média, deslocando o gráfico para o lado esquerdo. A grande maioria dos países classificadas no substrato até 25, e países do Leste Europeu aparecendo com máximas.

3.3.5 MORTES ATRIBUÍVEIS À POLUIÇÃO DO AR INTERIOR (mortes / 100.000) Indoor air pollution attributable deaths (deaths/100,000)

Segundo o Instituto para a Avaliação e Métricas Globais de Saúde, refere-se a mortes causadas em idade padrão pela poluição do ar interior, incluindo casos derivados da gripe Influenza, pneumonia pneumocócica, pneumonia tipo B influenza H, doenças respiratórias, pneumonia provocada pelo vírus sincicial, outras infecções do trato respiratório inferior,

traquéia, brônquios e cancer do pulmão, doença isquémica cardíaca, acidente vascular encefálico isquêmico e não-isquêmicas, doença pulmonar obstrutiva crónica, e catarata por 100.000 pessoas. No modelo SPI, os dados são dimensionados de 3 (<30 mortes por 100.000 pessoas) para 1 (> 100 mortes por 100.000 pessoas), segundo o *Institute for Global Health Metrics and Evaluation*.

Gráfico 6: Descritivo – *Indoor Air Pollution*



Fonte: Autor – Dados Guia GPS.

Análise:

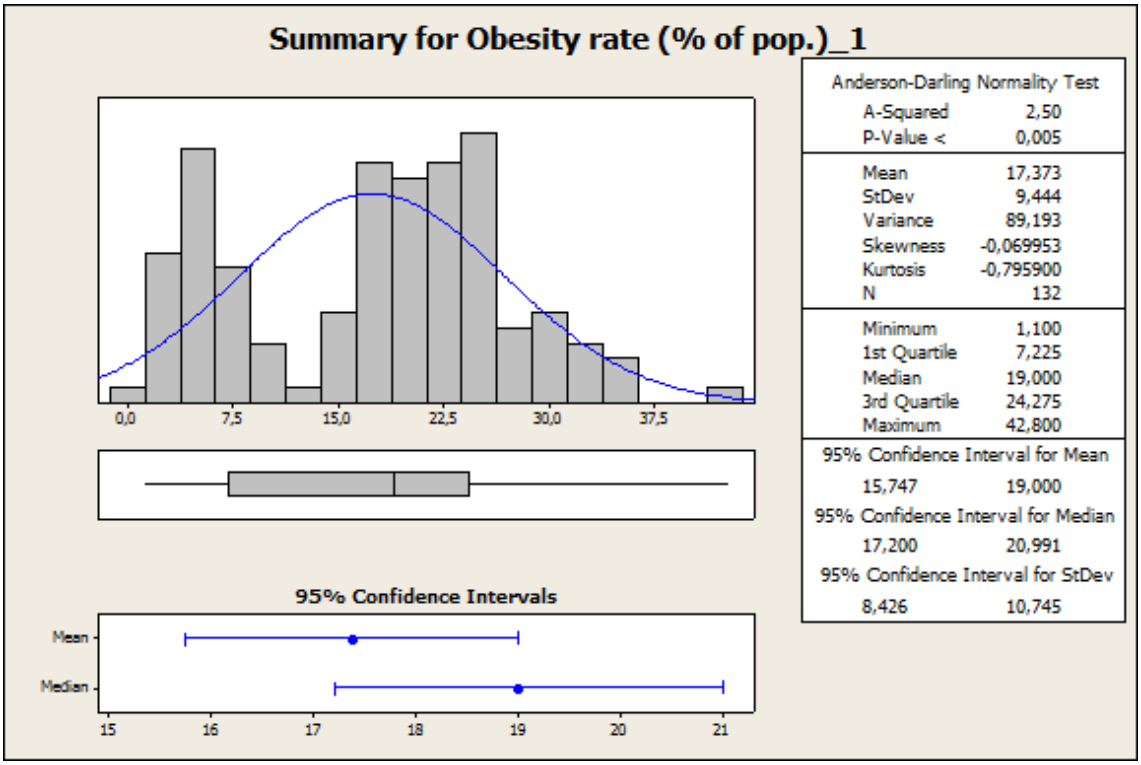
O valor do primeiro quartil de 3,696 indica que 25% da amostra possui valores entre 0 e 3,696. O índice Poluição do Ar Interior tem média 46,002 (intervalo de confiança de 95%, entre 37,769 e 54,235) e desvio padrão de 47,817 (intervalo de confiança de 95%, entre 42,661 e 54,401). Estes dados indicam que há uma alta variabiliade na amostra, ou seja, uma desproporção muito grande entre os países.

3.3.6 TAXA DE OBESIDADE (% DA POP.) Obesity rate (% of pop.)

É considerada obesa a percentagem da população com um índice de massa corporal (IMC) de 30 kg/m^2 ou superior (estimativa idade-padronizados), de ambos os sexos, segundo a Organização Mundial de Saúde.

Sendo assim, a distribuição de obesidade da população nos países pesquisados se apresenta conforme o Gráfico7, a seguir.

Gráfico 7: Descritivo – *Obesity Rate*



Fonte: Autor – Dados Guia GPS.

Análise:

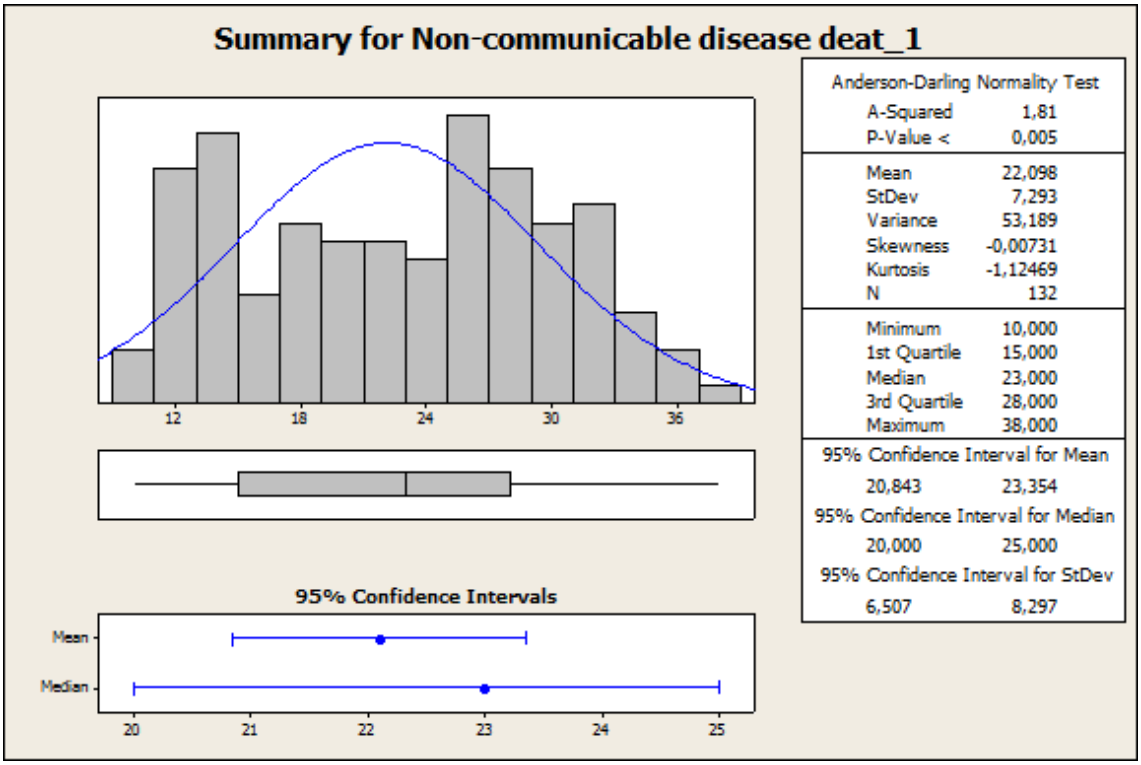
A média de Obesidade nos países pesquisados é de 17,373 (intervalo de confiança de 95%, entre 15,747 e 19) com desvio padrão de 9,444 (intervalo de confiança de 95%, entre 8,426 e 10,745), conforme apresentado no Gráfico 7.

Aqui se apresenta uma análise interessante, onde praticamente podemos ver quase “duas distribuições” de dados... Um dos substratos de 0 a 14, representados basicamente por África e Ásia, e uma segunda com países de Europa, Austrália e Nova Zelândia, Ibéricos e Estados Unidos.

3.3.7 MORTES POR DOENÇAS NÃO TRANSMISSÍVEIS ENTRE AS IDADES DE 30 E 70 (PROBABILIDADE DE MORRER) Non-communicable disease deaths between the ages of 30 and 70 (probability of dying)

Segundo a Organização Mundial de Saúde, “Mortes por doenças não transmissíveis entre as idades de 30 e 70 anos” é um índice que se refere à probabilidade de morrer entre as idades de 30 e 70 anos de doenças cardiovasculares, câncer, diabetes ou doença respiratória crônica.

Gráfico 8: Descritivo – *Non-communicable desease deaths*



Fonte: Autor – Dados Guia GPS.

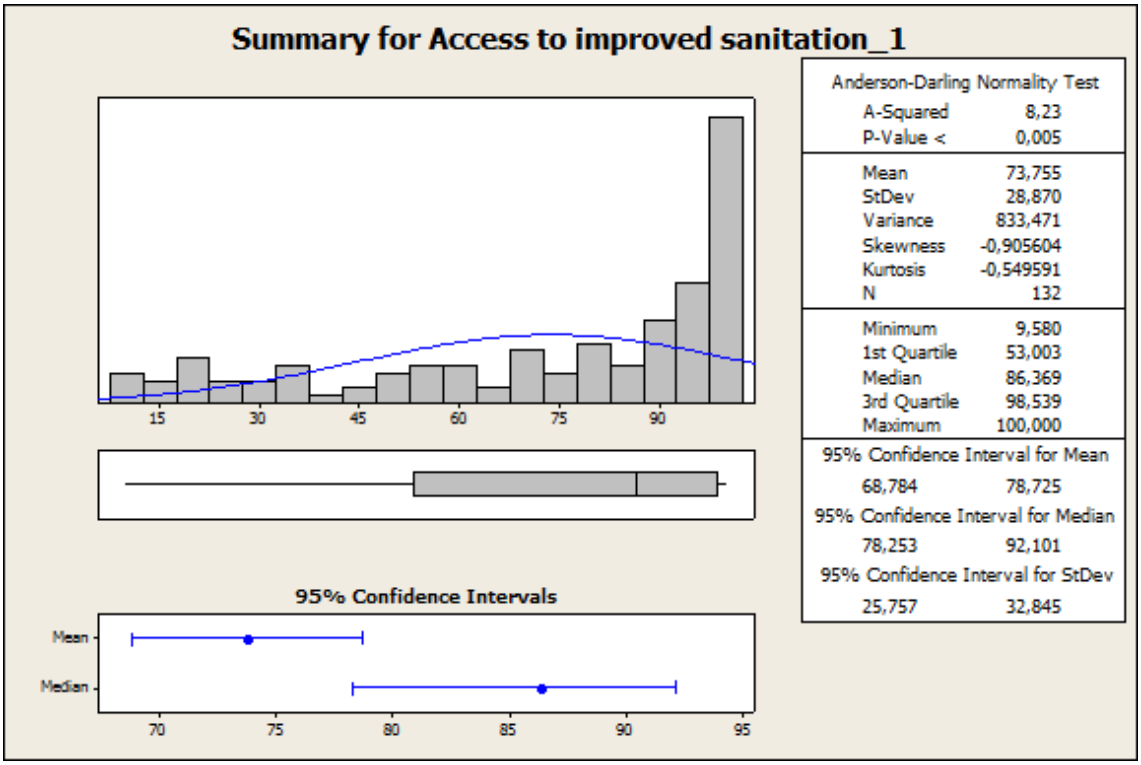
Análise:

Conforme mostrado no Gráfico 8, o índice de Doenças Não Transmissíveis tem média de 22,098 (intervalo de confiança de 95%, entre 20,843 e 23,354) e desvio padrão de 7,293 (intervalo de confiança de 95%, entre 6,507 e 8,297).

3.3.8 ACESSO A INSTALAÇÕES SANITÁRIAS MELHORADAS (% DA POP.)
Access to improved sanitation facilities (% of pop.)

Acesso a instalações sanitárias melhoradas (% da pop.) refere-se à percentagem da população com saneamento melhorado, incluindo autoclismos, sistemas de tubulações de esgoto, fossas sépticas, latrinas melhoradas e ventiladas (VIP), latrina com laje, casa de banhos, segundo o Programa de Monitorização da OMS e o Abastecimento de Água e Saneamento da UNICEF.

Gráfico 9: Descritivo – *Access to Improved Sanitation*



Fonte: Autor – Dados Guia GPS.

Análise:

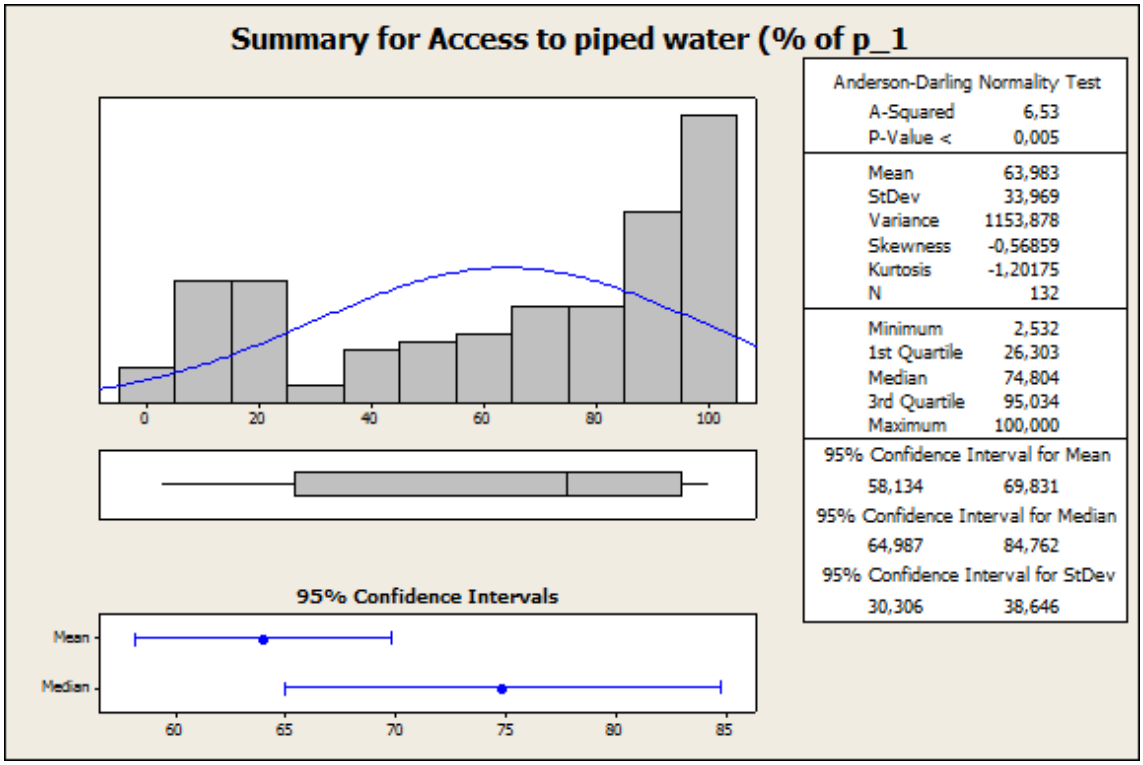
O índice em questão tem média de 73,755 (intervalo de confiança de 95%, entre 68,784 e 78,725) e desvio padrão de 28,870 (intervalo de confiança de 95%, entre 25,757 e 32,845). Esse é mais um índice que mostra a discrepância entre países de primeiro e de terceiro mundo, conforme mostra o Gráfico 8.

Aqui vemos a Mediana à direita da média, estando a mediana em 86,369. O mínimo em 9,58 e máximo em 100 (22 países com score 100 e mais 41 países entre 90 e 100 – maioria representada pela Europa e países mais desenvolvidos das outras áreas). A parte inferior mais uma vez para o grupo de países da África e Ásia (em sua maioria).

3.3.9 ACESSO A ÁGUA CANALIZADA (% DA POP.) Access to piped water (% of pop.)

Acesso a água canalizada refere-se à percentagem da população com um tubo de serviço de água conectada com canalização interna para um ou mais torneiras ou uma conexão de água canalizada a uma torneira colocada no quintal ou terreno fora da casa, segundo a OMS - Programa de Monitorização / UNICEF para o Abastecimento de Água e Saneamento.

Gráfico 10: Descritivo – *Access to Piped Water*



Fonte: Autor – Dados Guia GPS.

Análise:

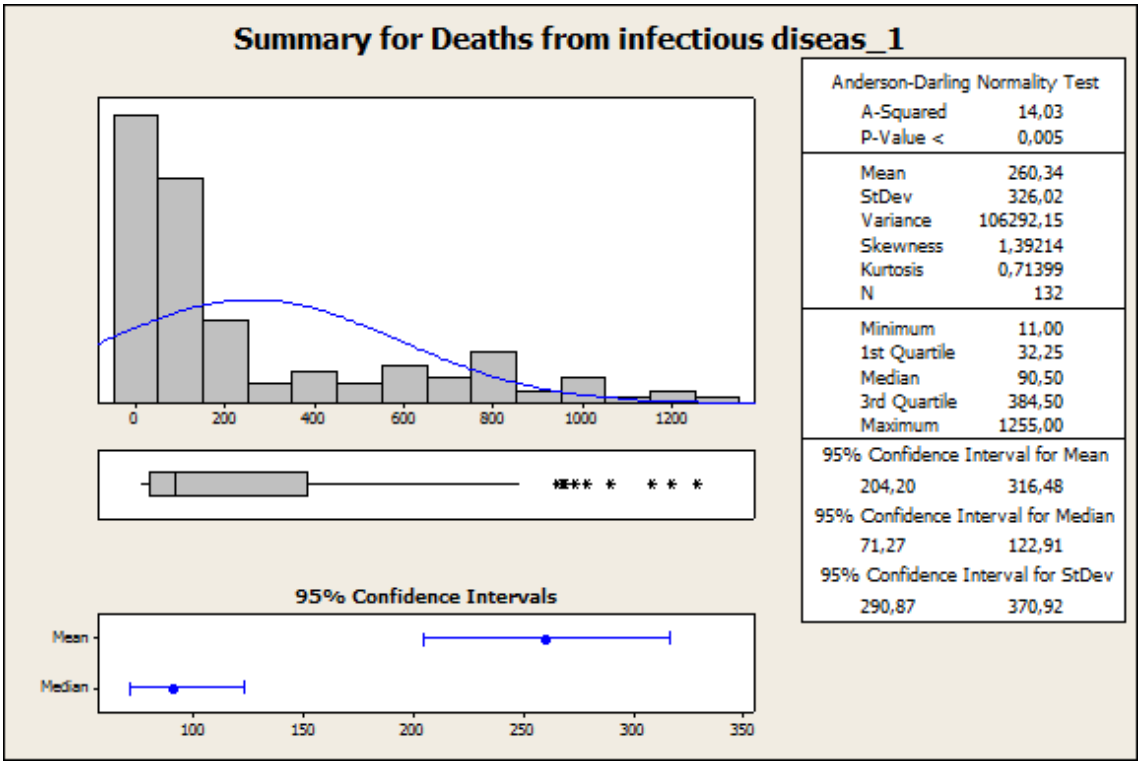
O índice de Acesso à Água Encanada médio é de 63,983 (intervalo de confiança de 95%, entre 58,134 e 69,831) e desvio padrão de 33,969 (intervalo de confiança de 95%, entre 30,306 e 38,646), conforme o Gráfico 10. Pode-se observar que o desvio-padrão desse índice é bem elevado, isso mostra que há uma alta variabilidade nessa variável.

E assim como os índices anteriores, os valores mais baixos desse índice estão na África e os mais baixos estão na Europa e Austrália. Com poucas exceções, segue bastante o mesmo perfil do grupo de Instalações de Saneamento melhoradas.

3.3.10 MORTES POR DOENÇAS INFECCIOSAS TRANSMISSÍVEIS (IDADE PADRONIZADA DE MORTES / 100.000) Deaths from infectious diseases (deaths/100,000)

Segundo a Organização Mundial da Saúde, refere-se à taxa de mortalidade de mortes causadas por tuberculose, doenças sexualmente transmissíveis, HIV / AIDS, diarréia, tosse convulsiva, poliomielite, sarampo, tétano, meningite, hepatite B, hepatite C, malária, tripanossomíase, doença de Chagas, esquistossomose, leishmaniose, filariose linfática, a oncocercose, hanseníase, dengue, encephaltitis japonesa, tracoma, infecções intestinais e outras doenças infecciosas por 100 mil pessoas.

Gráfico 11: Descritivo – *Deaths from Infectious Disease*



Fonte: Autor – Dados Guia GPS.

Análise:

O Gráfico 11 mostra que o índice de Morte por Doença Infecciosa médio é de 260,34% (intervalo de confiança de 95%, entre 204,20 e 316,48) e desvio padrão de 326,02 (intervalo de confiança de 95%, entre 290,87 E 370,92).

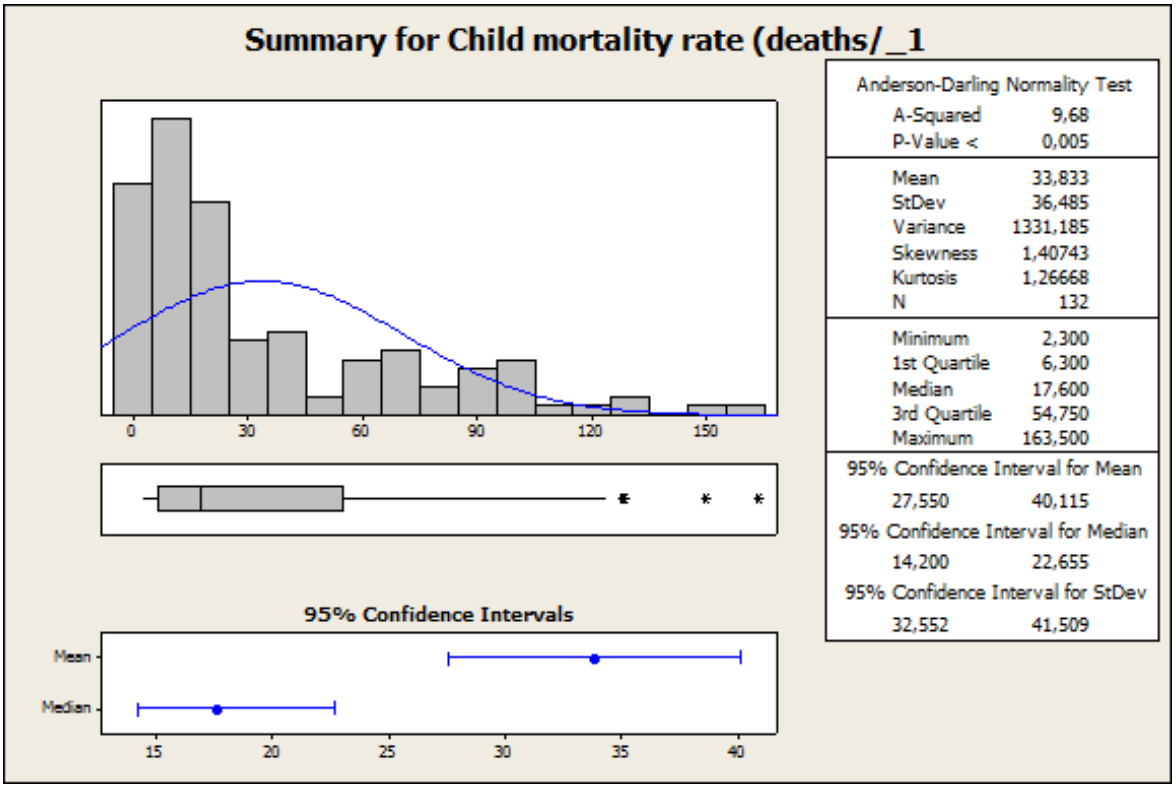
A Mediana (90,50) à Esquerda da Média (260,34 – que é quase 3 vezes maior) determinando intervalos de confiança bastante diferenciados. **A dispersão é muito alta**, com mínimo em 11 e máximo em 1255. A análise dos países nos substratos apresentados, com poucas **exceções são inversamente proporcionais a acesso a água encanada e instalações de saneamento melhoradas.**

Obs: O Mesmo Vale Para Os Próximos 3 Gráficos – Mortalidade Infantil, Mortalidade Maternal e Natimortos.

3.3.11 TAXA DE MORTALIDADE INFANTIL (MORTES / 1.000 NASCIMENTOS) Child mortality rate (deaths/1,000 live births)

Segundo o Grupo Interagências das Nações Unidas para Estimativas sobre Mortalidade Infantil, esta taxa refere-se à probabilidade de uma criança nascida em um ano específico morrer antes de completar cinco anos de idade por 1.000 nascidos vivos.

Gráfico 12: Descritivo – *Child Mortality Rate*



Fonte: Autor – Dados Guia GPS.

Análise:

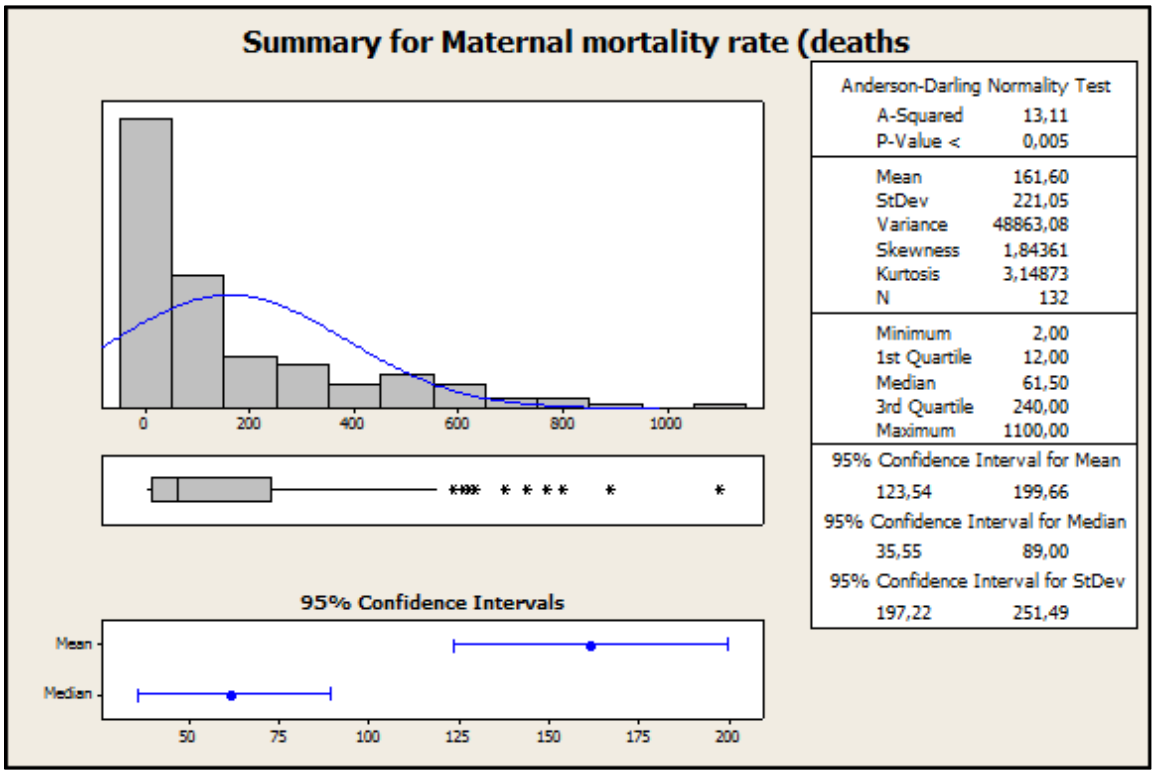
Conforme apresentado no Gráfico 12, a média para Mortalidade Infantil é de 33,833 (intervalo de confiança de 95%, entre 27,550 e 40,115) e desvio padrão de 36,485 (intervalo de confiança de 95%, entre 32,552 e 41,509).

A distribuição do gráfico da Mortalidade Infantil é parecido com o da Mortalidade Maternal e com o anterior.

3.3.12 TAXA DE MORTALIDADE MATERNA (ÓBITOS / 100.000 NASCIDOS VIVOS) Maternal mortality rate (deaths/100,000 live births)

Segundo a Organização Mundial de Saúde, refere-se ao número anual de mortes de mulheres por qualquer causa relacionada ou agravada pela gravidez ou sua gestão (excluindo causas acidentais ou incidentais) durante a gravidez e no parto ou no prazo de 42 dias após o término da gravidez, independentemente da duração e local da gravidez, por 100.000 nascidos vivos.

Gráfico 13: Descritivo – *Maternal Mortality Rate*



Fonte: Autor – Dados Guia GPS.

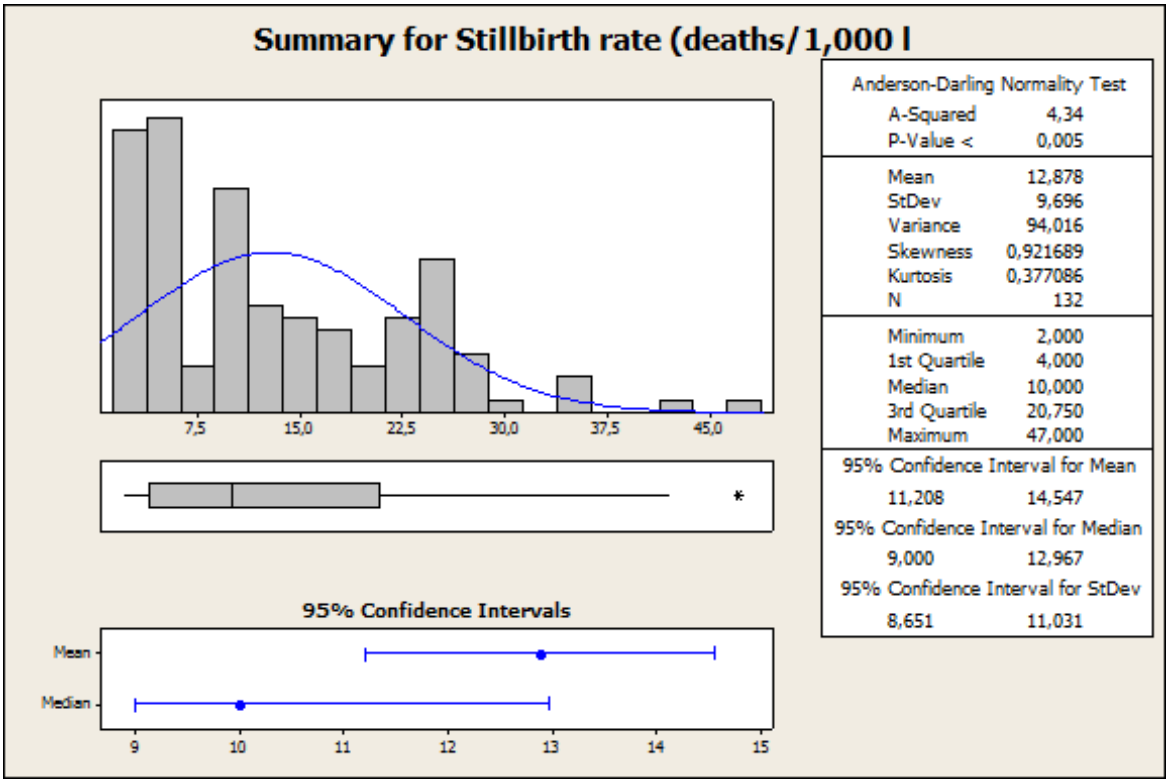
Análise:

O Gráfico 13 mostra que a Mortalidade Maternal tem média de 161,60 (intervalo de confiança de 95%, entre 123,54 e 199,66) e desvio padrão de 221,05 (intervalo de confiança de 95%, entre 197,22 e 251,49). Para os extremos, será desconsiderado o valor mínimo 2 e o Máximo de 1100, demonstrando dispersão muito alta.

3.3.13 TAXA DE NATIMORTOS (MORTES / 1.000 NASCIMENTOS) Stillbirth rate (deaths/1,000 live births)

Segundo a Organização Mundial de Saúde, refere-se aos óbitos fetais até o terceiro trimestre (> 1000 gramas ou> 28 semanas), por 1.000 nascidos vivos ou natimortalidade.

Gráfico 14: Descritivo – *Stilbirth Rate*



Fonte: Autor – Dados Guia GPS.

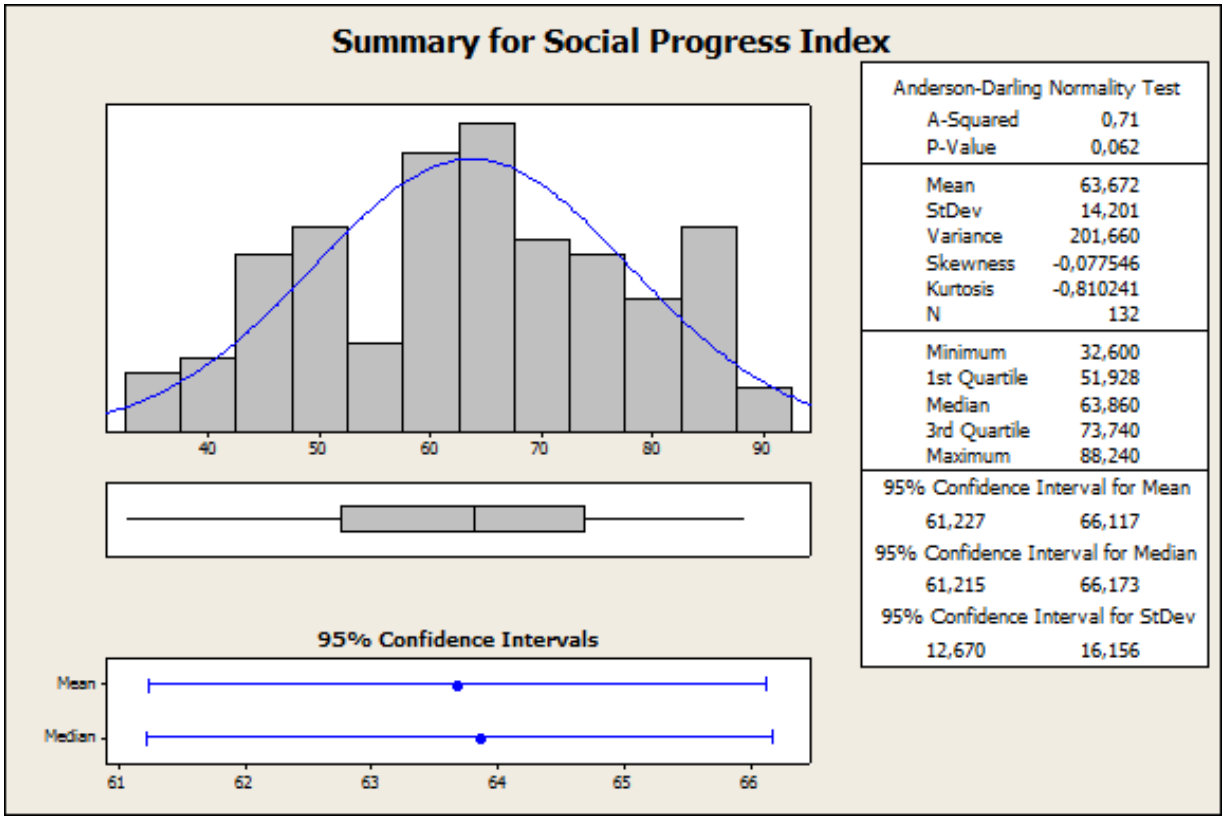
Análise:

O Gráfico 14 mostra que o índice de Natimorto médio é de 12,878 (intervalo de confiança de 95%, entre 11,208 e 14,547) e desvio padrão de 9,696 (intervalo de confiança de 95%, entre 8,651 e 11,031). Para os extremos, tem-se o mínimo 2 e o máximo 47.

3.3.14 ÍNDICE DE PROGRESSO SOCIAL Social Progress Index

Compreende necessidades humanas básicas atendidas, índices de bem-estar social e de oportunidades, ou seja, vários indicadores individuais que aqui estão listados e contidos no índice de progresso social (IPS).

Gráfico 15: Descritivo – *Social Progress Index*



Fonte: Autor – Dados Guia GPS.

Análise

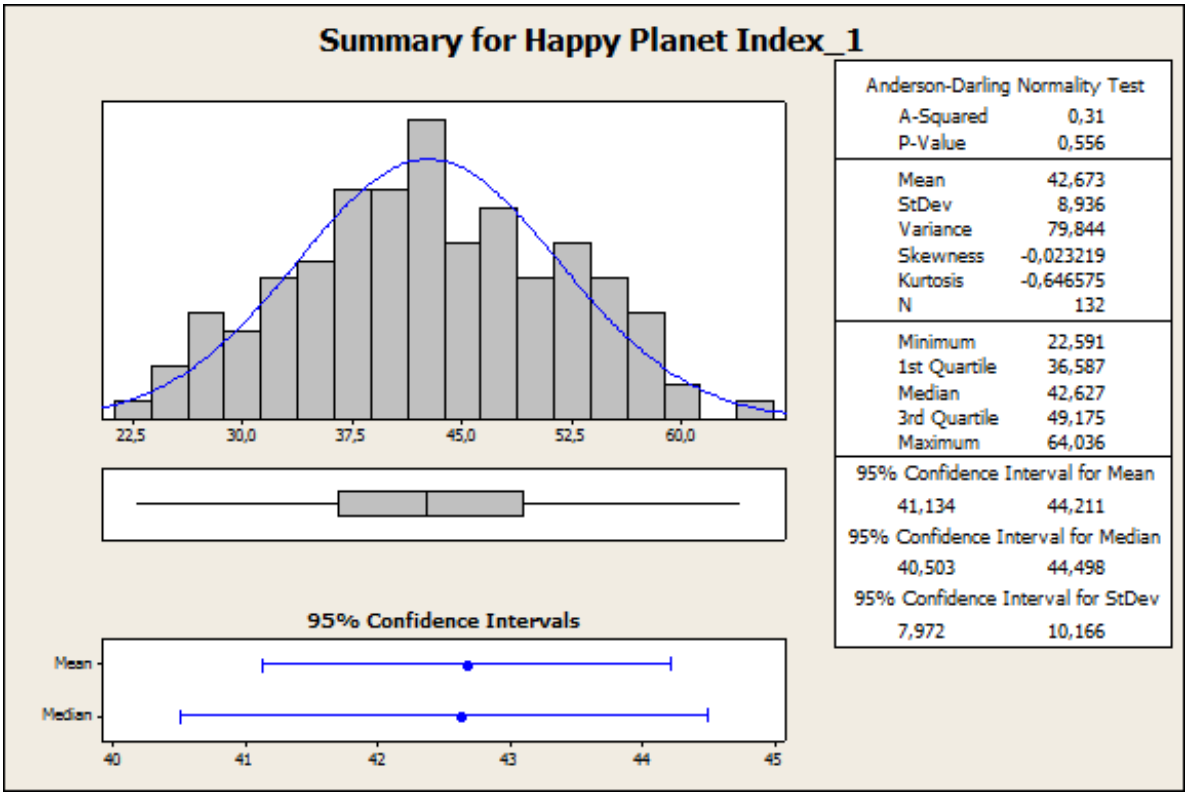
O Gráfico 15 demonstra o IPS médio de 63,672 (intervalo de confiança de 95%, entre 61,227 e 66,117) e desvio padrão de 14,201 (intervalo de confiança de 95%, entre 12,670 e 16,156).

Índice de Progresso Social segue uma distribuição normal, com média e mediana muito próximas. O Progresso social apresenta alguns substratos diferenciados (55 com poucas mostras e 85 com muitas). Muito similar ao que acontece no índice HPI.

3.3.15 ÍNDICE DE PLANETA FELIZ (HPI) (Happy Planet Index 2012)

Segundo Abdallah S, Michaelson J, Shah S, Stoll L, Marcas N (2012), o Índice de Happy Planet (HPI) é a medida líder global de bem-estar sustentável. Ele utiliza dados globais sobre a expectativa de vida, bem-estar experimentado e medidas ecológicas para calculá-lo. O índice é uma medida de eficiência, que classifica os países de quantas vidas longas e felizes eles produzem por unidade de insumo ambiental. O relatório de 2012 HPI classifica 151 países e é a terceira vez que o índice foi publicado. - Veja mais em:

Gráfico 16: Descritivo – *Happy Planet Index*



Fonte: Autor – Dados Guia GPS.

Análise:

O Índice de Planeta feliz segue bastante a distribuição normal. Observa-se que a Média de 42,673 e a Mediana de 42,627 são bem próximos. O Gráfico 16 demonstra o HPI médio com intervalo de confiança de 95%, entre 41,134 e 44,211 e desvio padrão de 8,936 (intervalo de confiança de 95%, entre 7,972 e 10,166).

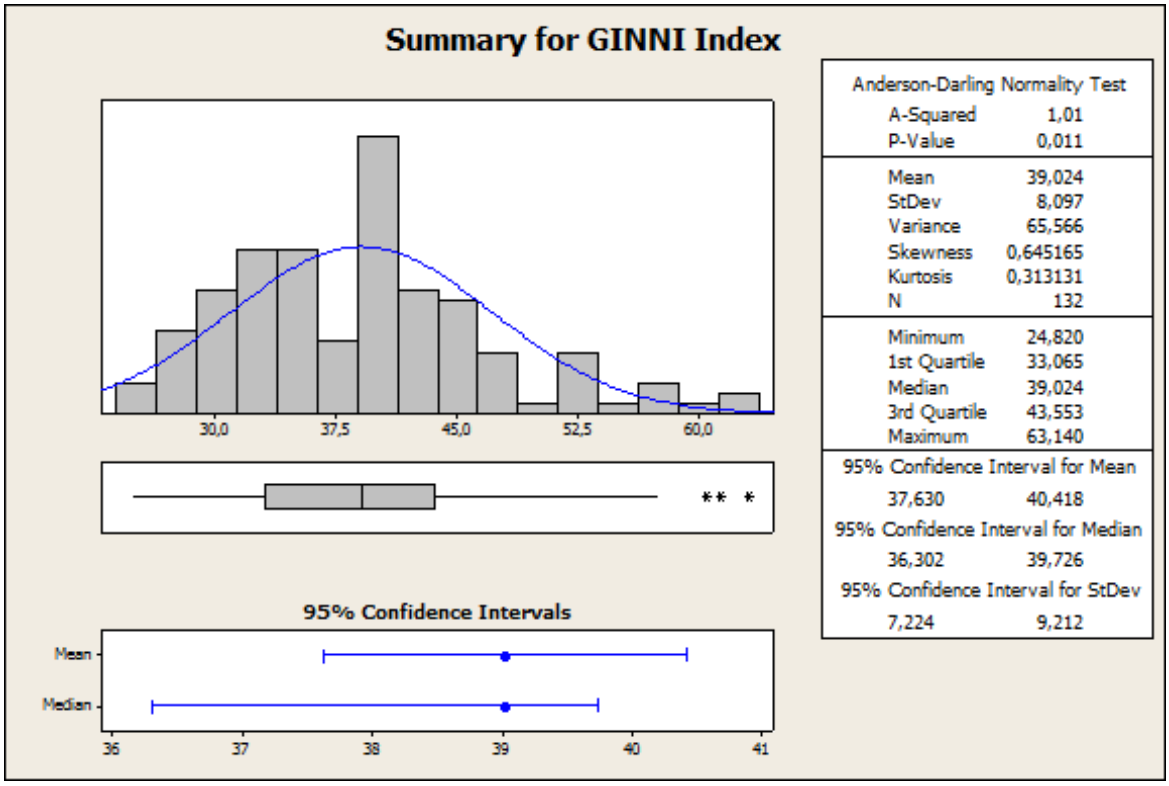
Pelo teste de Anderson-Darling para normalidade, os dados podem ser considerados normais (p-valor = 0,556), para o nível de significância de 0,05. Pode-se observar que a maioria dos dados se encontra em torno da média.

3.3.16 ÍNDICE GINNI (World Bank estimate)

É uma estimativa do Banco Mundial que mede a extensão em que a distribuição de renda (ou, em alguns casos, a despesa de consumo) entre indivíduos ou famílias dentro de uma economia desvia de uma distribuição perfeitamente igual, através de uma curva de Lorenz com as percentagens cumulativas de rendimento total recebido com o número acumulado de beneficiários, partindo dos indivíduos ou famílias mais pobres. O índice de Ginni mede a área entre a curva de Lorenz e uma linha hipotética de igualdade absoluta, expressa em percentagem da área máxima abaixo da linha. Assim, um índice de Ginni de 0 representa a igualdade perfeita, enquanto um índice de 100 indica desigualdade perfeita, Segundo o Banco Mundial, *Development Research Group*.

Os dados são baseados em dados de pesquisas domiciliares primárias obtidos a partir de órgãos de estatística e departamentos governamentais nacionais do Banco Mundial. O índice de Ginni fornece uma medida resumo conveniente do grau de desigualdade. Os dados sobre a distribuição de renda ou de consumo vêm de pesquisas domiciliares nacionalmente representativas. Onde os dados originais do inquérito aos agregados familiares estavam disponíveis, eles têm sido usados para calcular as quotas de renda ou de consumo por quintil. Caso contrário, as ações foram calculadas a partir dos melhores dados agrupados disponíveis. Os dados de distribuição foram ajustados para o tamanho do agregado familiar, fornecendo uma medida mais consistente de renda ou consumo per capita. Nenhum ajuste foi feito para diferenças espaciais em custo de vida nos países, porque os dados necessários para tais cálculos são geralmente indisponíveis.

Gráfico 17: Descritivo - GiNNI



Fonte: Autor – Dados Guia GPS.

Análise:
Segundo o Gráfico 17, para o Índice de Ginni, houve um pico cuja mediana foi de 39,024, equivalente à média. O desvio padrão foi de 8,097. Tem uma distribuição não uniforme, com uma maior quantidade de países à esquerda.

3.3.17 ÍNDICE DE SAÚDE DO OCEANO (OHI) Ocean Health Index

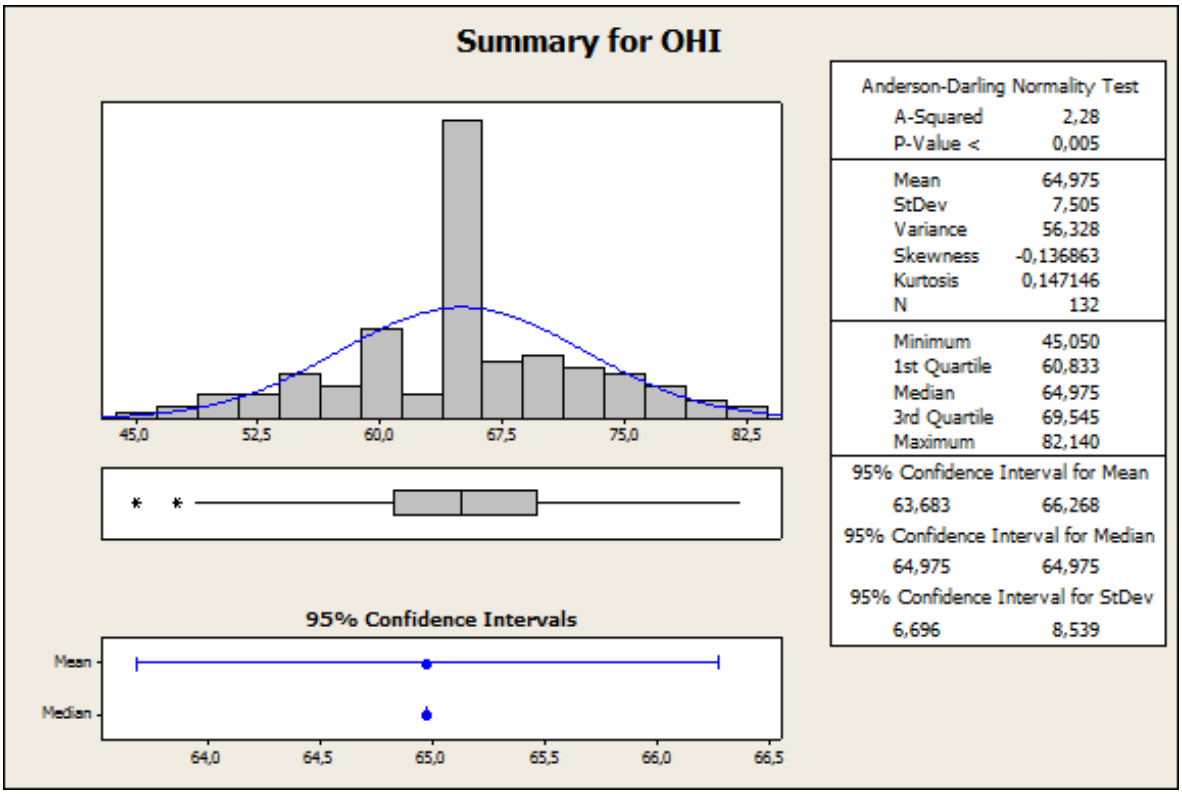
O Índice de Saúde do Oceano estabelece pontos de referência para atingir 10 objetivos sócio-ecológicos amplamente aceitos, e contagens dos oceanos adjacentes para 171 países e territórios. Avaliadas globalmente e por país, estas 10 metas públicas representam a ampla

gama de benefícios que um oceano saudável pode proporcionar; cuja pontuação geral de cada país é a média de suas respectivas pontuações objetivo.

1. fornecimento de alimentos (colheita de frutos do mar de forma sustentável),
2. as possibilidades de pesca artesanal (garantindo alimento para as comunidades locais), (habitat de preservação que absorvem carbono)
3. produtos naturais (colheita recursos oceânicos não alimentares de forma sustentável),
4. armazenamento de carbono,
5. proteção costeira (habitats de preservação que protegem shores) ,
6. coastal modos de vida & economias (empregos sustentáveis e prósperas economias costeiras),
7. turismo e lazer (mantendo a atração de destinos costeiros),
8. sense of place (proteção de espécies emblemáticas e lugares especiais),
9. águas limpas (poluição minimizando),
10. biodiversidade (apoiando saudável) .

O cálculo do Índice de Saúde do Oceano para 2014 representa o primeiro ano em que o Índice de Saúde do Oceano calculou uma pontuação para todo o oceano global. Todas essas pontuações em conjunto produzem a primeira pontuação total mundial Índice de Saúde do Oceano, que é de 67. As maiores foram Habitats (91), a biodiversidade 83 e economias & Modos de Vida (82). As menores foram Turismo e Recreação (44), produção de alimentos (51) e os produtos naturais (53).
No caso dos High Seas, demonstram necessidades de cuidado:
Chile + Ilha de Páscoa, marcam 68
Brasil + Trindade, marcam 66
Equador + Galapagos, marcam 64

Gráfico 18: Descritivo - OHI



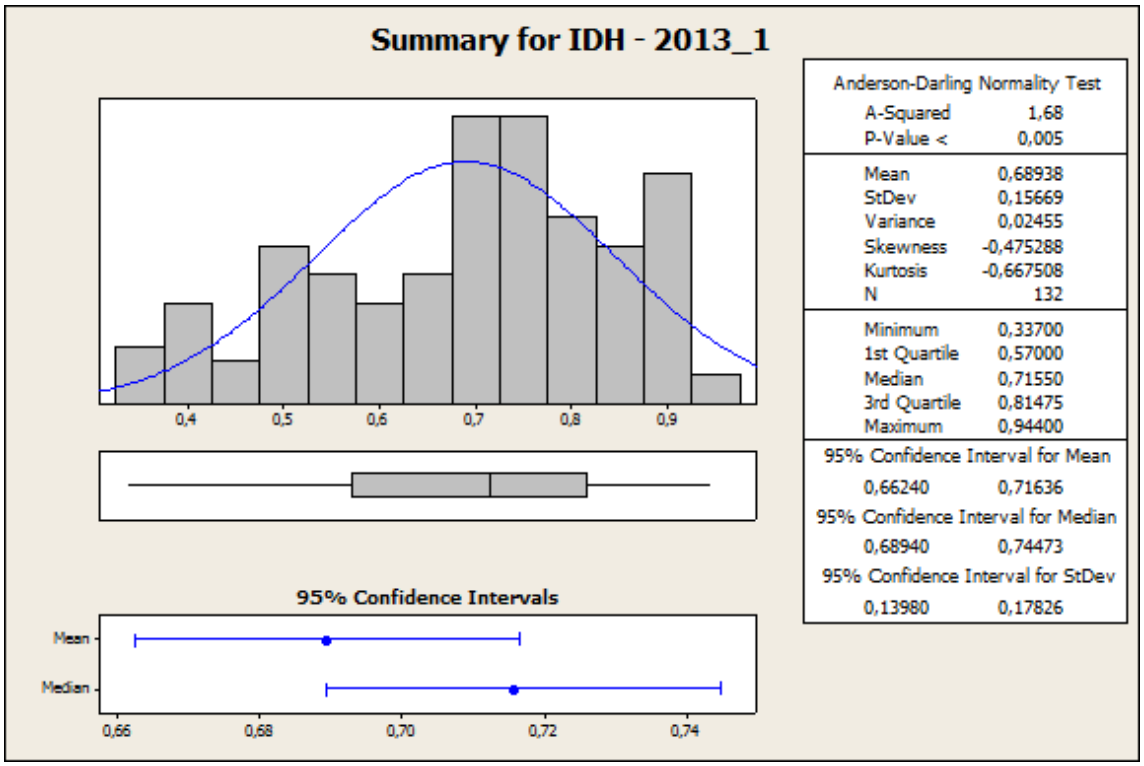
Fonte: Autor – Dados Guia GPS.

Análise:
Conforme apresentado no Gráfico 18, a média para OHI é de 64,975 (intervalo de confiança de 95%, entre 63,683 e 66,268) e desvio padrão de 7,505 (intervalo de confiança de 95%, entre 6,696 e 8,539).

3.3.18 IDH - 2013

O Índice de Desenvolvimento Humano (IDH) mede o Progresso de uma Nação a partir de três dimensões: renda, saúde e educação, segundo o Programa das Nações Unidas para o Desenvolvimento. O indicador nominal varia de 0 a 1 e que, a partir dos indicadores renda, saúde e educação, mede o progresso de uma nação (PNUD - ONU, 2013).

Gráfico 19: Descritivo - IDH



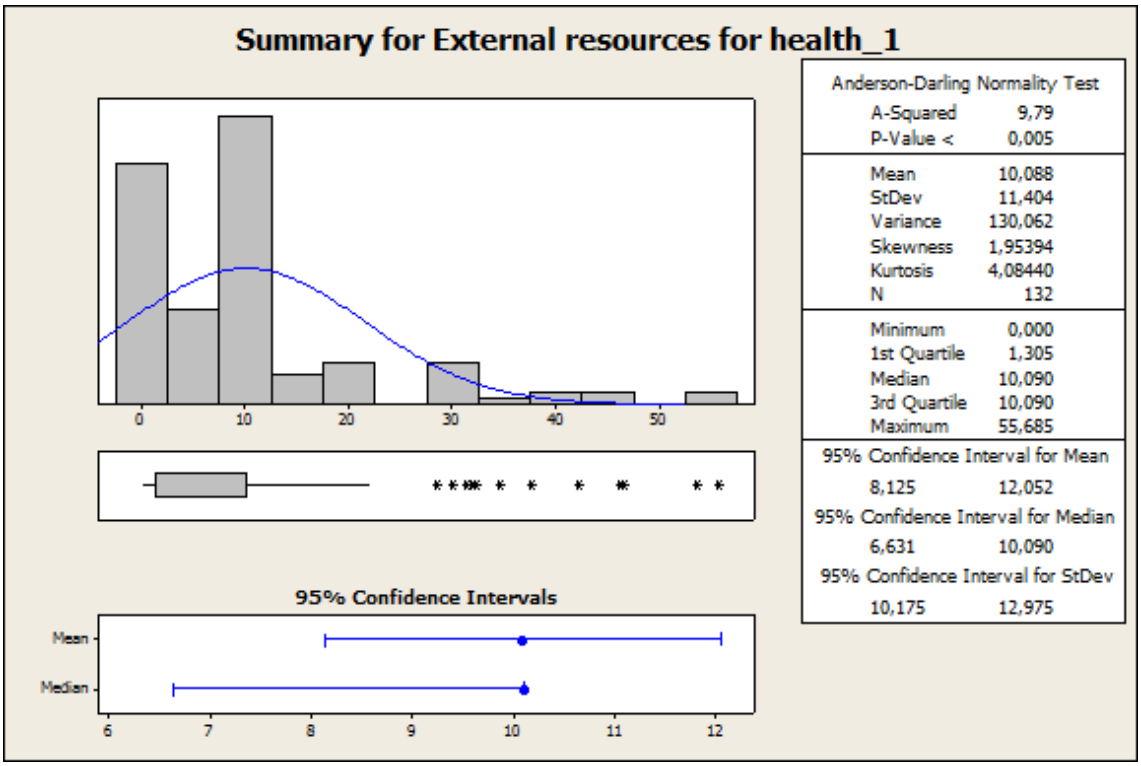
Fonte: Autor – Dados Guia GPS.

Análise:
A mediana de 0,71 está bem próxima à média de 0,69 aproximadamente. Do lado esquerdo do gráfico, se apresentam os países com baixo índice de IDH, os países chamados subdesenvolvidos. Já do lado direito, os países desenvolvidos. O desvio padrão foi de 0,15669, com 95% de confiança de 0,13980 e 0,17826.

3.3.19 RECURSOS EXTERNOS PARA A SAÚDE (% da despesa total em saúde) 2012 - External resources for health (% of total expenditure on health)

Segundo o Banco Mundial e Indicadores de Desenvolvimento Mundial (WDI) 2012, os recursos externos para a saúde são fundos ou serviços em espécie que são fornecidos por entidades que não façam parte do país em questão. Os recursos podem ser provenientes de organizações internacionais, em outros países através de acordos bilaterais ou organizações não-governamentais estrangeiras. Estes recursos são parte do total das despesas de saúde.

Gráfico 20: Descritivo – *External Resources for Health*



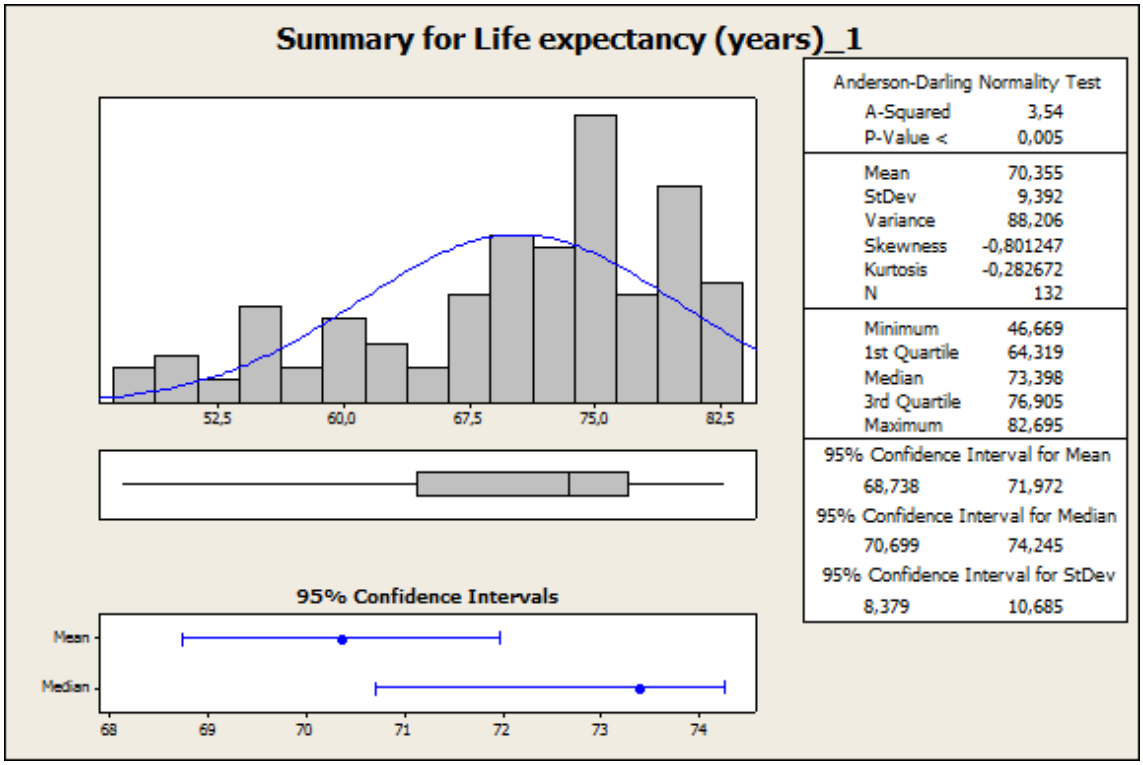
Fonte: Autor – Dados Guia GPS.

Análise:
Segundo o Gráfico 20, a média é de 10 (Com intervalo de confiança de 95% entre 8,125 e 12,052) e desvio padrão de 11,404, (com intervalo de confiança de 95% entre 10,175 e 12,975). Percebe-se uma concentração dos países à esquerda, com média e mediana próximas, e um desvio-padrão alto, mostrando alta variabilidade.

3.3.20 A EXPECTATIVA DE VIDA (ANOS)

Segundo os Indicadores de Desenvolvimento Mundial, a expectativa de vida (anos) refere-se ao número de anos que um recém-nascido viveria se os padrões de mortalidade prevalecentes na altura do seu nascimento fossem permanecer o mesmo durante toda a sua vida. Este indicador revela o tempo médio de vida de um indivíduo, com base no ano de seu nascimento, sua idade atual e outros fatores demográficos, incluindo sexo.

Gráfico 21: Descritivo – *Life Expectancy*



Fonte: Autor – Dados Guia GPS.

Análise:

Conforme mostrado no Gráfico 21, a Expectativa de Vida média 70,355 anos (intervalo de confiança de 95%, entre 68,738 e 71,972) e desvio padrão de 9,392 (intervalo de confiança de 95%, entre 8,379 e 10,685).

Observa-se ainda que a média e a mediana estão próximas dos 70 anos. A maioria dos países estão do lado direito do gráfico, o que pode ser considerado bom, pois indica uma expectativa de vida mais alta. As maiores expectativas de vidas estão nos países de primeiro mundo.

3.4. NORMALIZAÇÃO DOS DADOS

O primeiro tratamento realizado foi a substituição de células vazias (marcadas com asterisco) pelo valor médio das variáveis em cada variável. Na análise abaixo foi possível identificar o número de valores faltantes na coluna “N*” e a média de cada variável, que pôde ser observado pela Tabela 2, reproduzido aqui novamente.

Tabela 2. Dados Descritivos

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
Maternal mortality rate	132	0	161,6	19,2	221,0	2,0	12,0	61,5	240,0	1100,0
Stillbirth rate (deaths/	131	1	12,878	0,850	9,733	2,000	4,000	10,000	21,000	47,000
Child mortality rate (de	132	0	33,83	3,18	36,49	2,30	6,30	17,60	54,75	163,50
Deaths from infectious d	132	0	260,3	28,4	326,0	11,0	32,3	90,5	384,5	1255,0
Access to piped water (%	129	3	63,98	3,03	34,36	2,53	24,72	75,33	95,34	100,00
Access to improved sanit	130	2	73,75	2,55	29,09	9,58	52,97	88,00	98,66	100,00
Indoor air pollution att	131	1	46,00	4,19	48,00	0,00	3,69	24,87	81,62	222,32
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319	73,398	76,905	82,695
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000	23,000	28,000	38,000

Obesity rate (% of pop.)	132	0	17,373	0,822	9,444	1,100	7,225	19,000	24,275	42,800
Outdoor air pollution at	130	2	15,75	1,12	12,76	0,00	7,00	12,00	20,25	67,00
Health expenditure, publ	132	0	57,47	1,49	17,09	18,01	45,02	58,98	71,06	94,15
External resources for h	93	39	10,09	1,41	13,61	0,00	0,56	3,65	15,01	55,69
IDH - 2013	125	7	0,6894	0,0144	0,1611	0,3370	0,5595	0,7220	0,8200	0,9440
1. GovernanÁa	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163	42,463	49,287	64,036
OHI	98	34	64,975	0,881	8,722	45,050	59,645	65,140	71,497	82,140
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780	38,160	44,550	63,140

Fonte: Elaborado pela autora

Análise:

Apesar de ainda não haverem sido normalizados através de outro procedimento do Minitab, pode-se perceber que, os índices podem se aproximar a uma escala normal e quando distribuídos, a área sob sua curva pode se aproximar a uma soma 100%. Será exemplificado com a variável Progresso Social, que tem o maior valor de 88, 24, a média de 63,67, o mínimo 32,60 e a mediana muito próxima à media 63,86, em uma escala até 100, indicando que em geral os países estão dentro da escala normal.

Países com lacunas em branco, foram substituídas pela média. São eles:

- Sudão: substituí pela média 12,878 na variável **Stillbirth rate;**
- Australia, E.U.A., Kuwait: Substituir pela média 63,98 Na variável **Access to piped Water**;
- Italy, New Zealand, Substituir pela média 73,75 na variável **Access to improved sanitation faciities**;
- Spain Substituir pela média 45,73 na variável **Indoor Air Pollution Attributable Deaths**;
- Montenegro, Serbia: media 15,75 na variável **Outdoor Air Pollution;**
- Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Kuwait, Latvia, Mauritius, Morocco, Netherlands, New Zeland, Norway, Poland, Portugal, Romania, Russia, Slovaquia, Spain, Sweden, Switzerland,Trinidad and Tobago, United Arabes Emirates, United Kingdom, United States substituí pela média 10,09 na variável **External resources for health;**
- Algeria, Kazakhstan, Korea, Kyrgyzstan, Latvia, Mauritius, Netherlands substituí "*"pela média 0,6894 na variável **IDH;**
- Lesotho, Montenegro, Swaziland: substituí "*" pela média 42,673 na variável **Happy Planet;**
- Armenia, Austria, Azerbaijan, Belarus, Bolivia, Botswana, Burkina Faso, Burundi, Central African Republic, Chad, Czech Republic, Hungary, Kazakhstan, Kyrgyzstan, Laos, Lesotho; Macedonia, Malawi, Mali, Moldova, Mongolia, Nepal, Nigeria, Paraguay, Rwanda, Serbia, Slovenia, Sri Lanka, Swaziland, Switzerland, Tajikistan, Uganda, Uzbekistan, Zambia: substituí "*" pela média 64,975 na variável **OHI;**
- Bangladesh, Cuba, Guinea, Iceland, Kuwait, Lebanon, , Mauritius, Montenegro, Nepal, Pakistan, Saudi Arabia, Sri Lanka, United Arab Emirates: substituí "*" pela média 39,024 **GInni INDEX.**

A Tabela 4 mostra a confirmação da eliminação dos dados ausentes depois que foram substituídos pela Média de cada uma das variáveis.

Tabela 4. Observação da Eliminação dos Dados Ausentes (N* = 0)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
Maternal mortality rate	132	0	161,6	19,2	221,0	2,0	12,0	61,5	240,0	1100,0
Stillbirth rate (deaths/	132	0	12,878	0,844	9,696	2,000	4,000	10,000	20,750	47,000
Child mortality rate (de	132	0	33,83	3,18	36,49	2,30	6,30	17,60	54,75	163,50
Deaths from infectious d	132	0	260,3	28,4	326,0	11,0	32,3	90,5	384,5	1255,0
Access to piped water (%	132	0	63,98	2,96	33,97	2,53	26,30	74,80	95,03	100,00
Access to improved sanit	132	0	73,75	2,51	28,87	9,58	53,00	86,37	98,54	100,00
Indoor air pollution att	132	0	45,73	4,17	47,92	0,00	3,70	24,42	81,59	222,32
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319	73,398	76,905	82,695
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000	23,000	28,000	38,000
Obesity rate (% of pop.)	132	0	17,373	0,822	9,444	1,100	7,225	19,000	24,275	42,800
Outdoor air pollution at	132	0	15,75	1,10	12,66	0,00	7,00	12,50	19,75	67,00
Health expenditure, publ	132	0	57,47	1,49	17,09	18,01	45,02	58,98	71,06	94,15
External resources for h	132	0	10,088	0,993	11,404	0,000	1,305	10,090	10,090	55,685
IDH - 2013	132	0	0,6894	0,0136	0,1567	0,3370	0,5700	0,7155	0,8147	0,9440
1. GovernanÁa	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
Happy Planet Index	132	0	42,673	0,778	8,936	22,591	36,587	42,627	49,175	64,036
OHI	132	0	64,975	0,653	7,505	45,050	60,833	64,975	69,545	82,140
GINNI Index	132	0	39,024	0,705	8,097	24,820	33,065	39,024	43,553	63,140

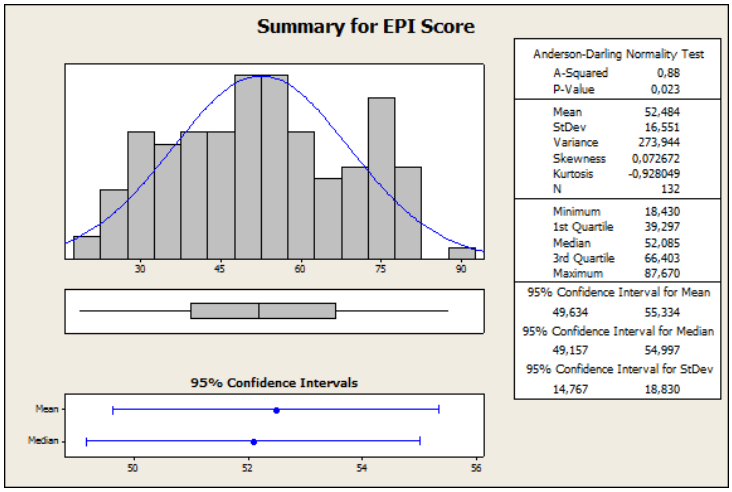
Fonte: Elaborado pela autora

Procedimento:

Após a substituição dos dados ausentes identificados por asterisco (*) pela média de cada variável, foram rodados novamente os indicadores de número de elementos (N), dados ausentes (N*), média, erro da média (SE mean), desvio padrão (Stdev), valor mínimo e máximo; apenas para confirmar se foram lançados corretamente. Portanto, sem alteração nos respectivos valores. Conforme imaginado, se se substitui os dados em branco pela média, os resultados não devem se alterar, as médias seguem as mesmas, mínimos e máximos também, assim como a mediana.

A seguir, serão apresentados os gráficos dos indicadores e analisados alguns deles.

Gráfico 22: Descritivo – EPI (normalizado)



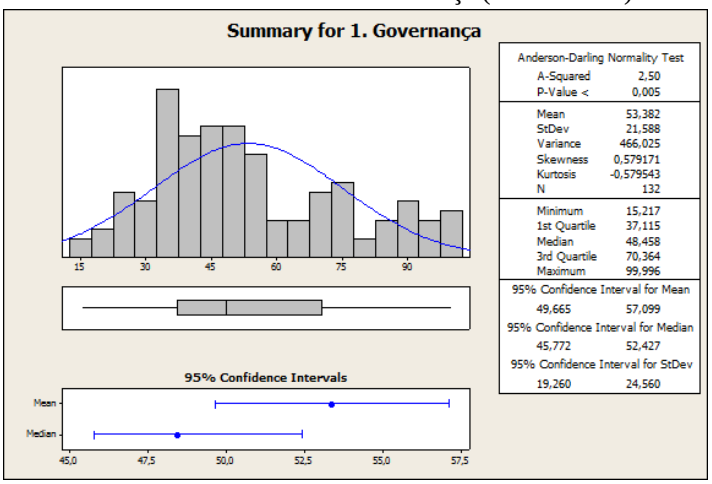
Fonte: Autor – Dados Guia GPS.

Análise:

Neste caso se nota uma distribuição próxima a uma distribuição normal, com média e mediana muito próximas. Isso significa que quase metade dos países estão na parte superior, e

a outra metade na inferior, sendo a média um pouco superior aos 50%. Nota-se um pico de países com a nota próxima a 75%, representado maioritariamente pelos paises europeus, destacando a Suíça com melhor desempenho no grupo (máxima).

Gráfico 23: Descritivo – Governança (normalizado)

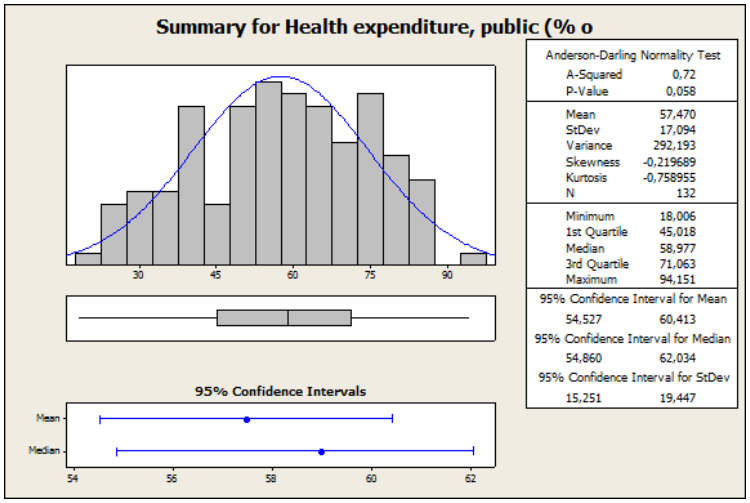


Fonte: Autor – Dados Guia GPS.

Análise:

Neste caso nota-se a mediana abaixo da média, indicando que a maioria dos países estão com resultado abaixo da média (distribuição não uniforme, com mais países com notas mais baixas e menos paises com notas muito mais altas) gerando um deslocamento para o lado esquerdo. Do lado superior, mais à direita, mais uma vez os países Europeus, Austrália, Nova Zelândia, Japão e Chile. Do lado com menores resultados os países da África e Ásia Central (em sua maioria).

Gráfico 24: Descritivo – *Health Expenditure, public* (normalizado)

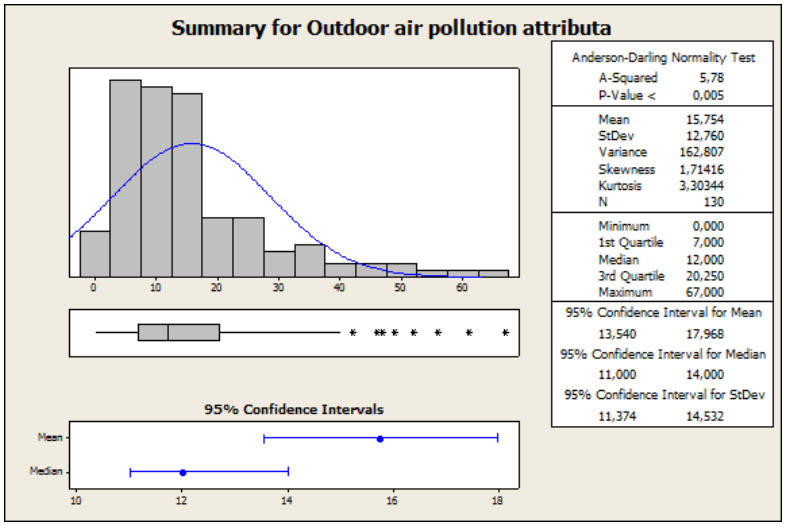


Fonte: Autor – Dados Guia GPS.

Análise:

Curva próxima a uma curva normal, porém com baixa frequência no substrato de 45 e um pouco mais alta no substrato de 40. Média e mediana muito próximas.

Gráfico 25: Descritivo – *Outdoor Air Pollution* (normalizado)

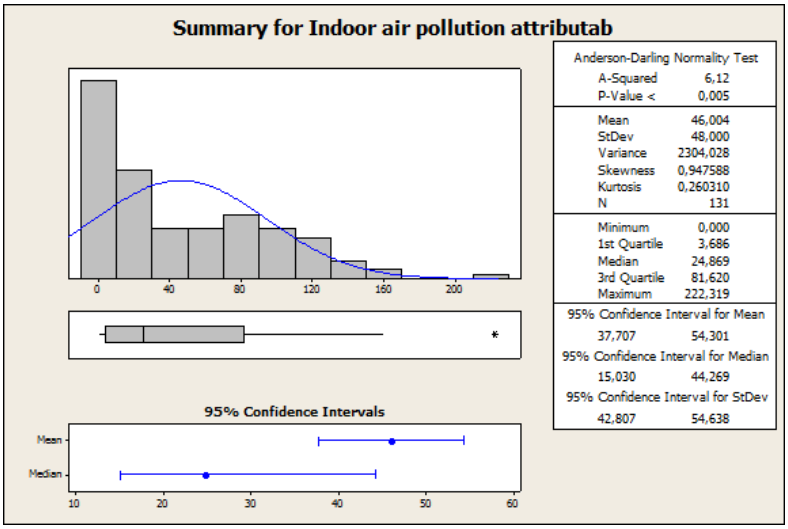


Fonte: Autor – Dados Guia GPS.

Análise:

No caso, a Mediana menor que a média, deslocando o gráfico para o lado esquerdo. A grande maioria dos países classificadas no substrato até 25, porém países do Easter Europe aparecendo com máximas.

Gráfico 26: Descritivo – *Indoor Air Pollution* (normalizado)

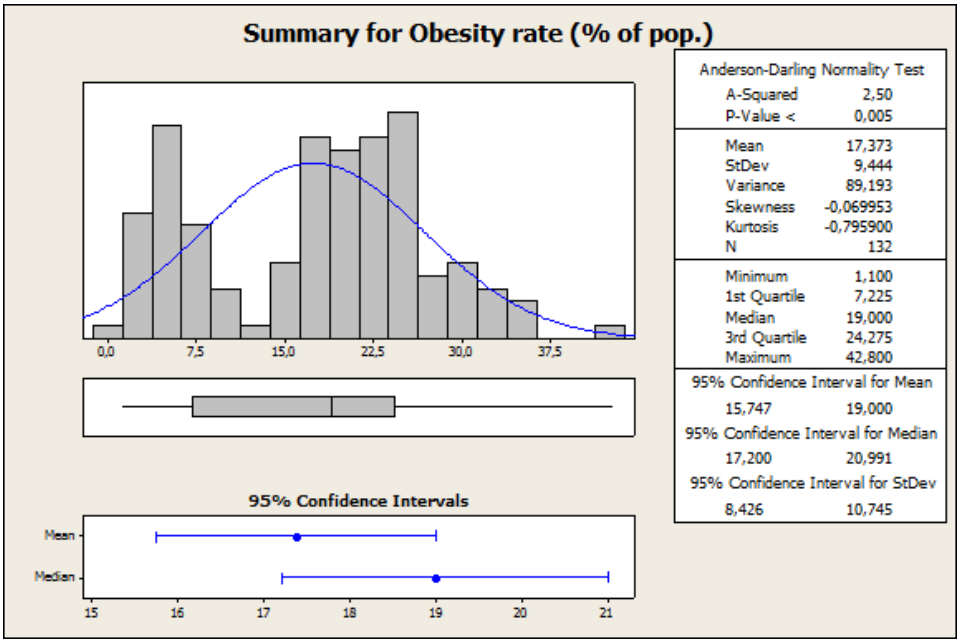


Fonte: Autor – Dados Guia GPS.

Análise:

Aqui o primeiro quartil em 3,686 demonstra que mais de 25% da amostra com valor inferior a 5, e mediana em 24,8 (50% da amostra), porém o 3º quartil com um valor de 81 e o desvio padrão em 48 (bem acima da média de 46) e o valor máximo chegando a 222, mostrando um desproporção muito grande entre a mediana, média e valor máximo.

Gráfico 27: Descritivo – *Obesity* (normalizado)

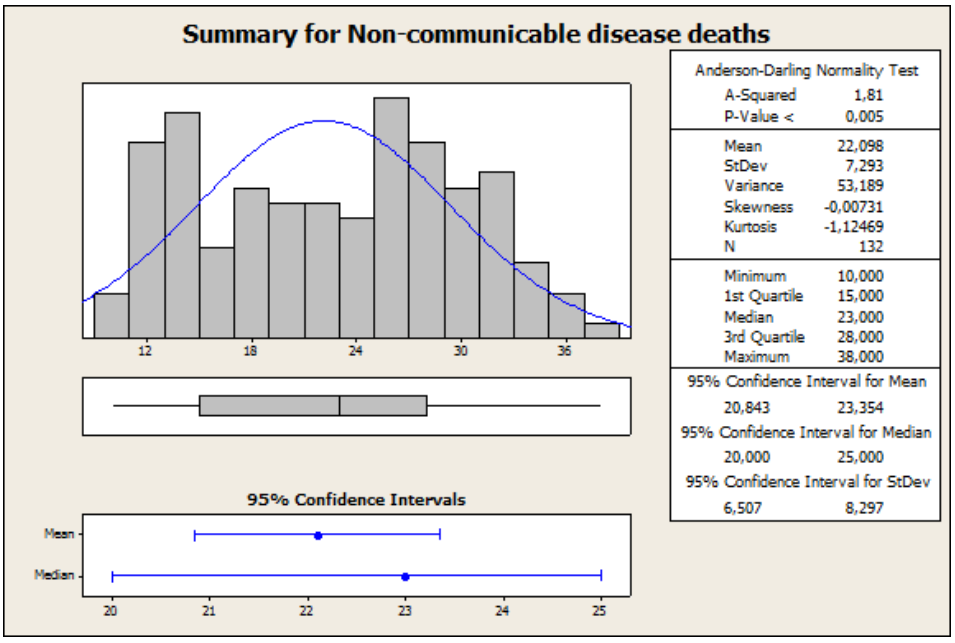


Fonte: Autor – Dados Guia GPS.

Análise:

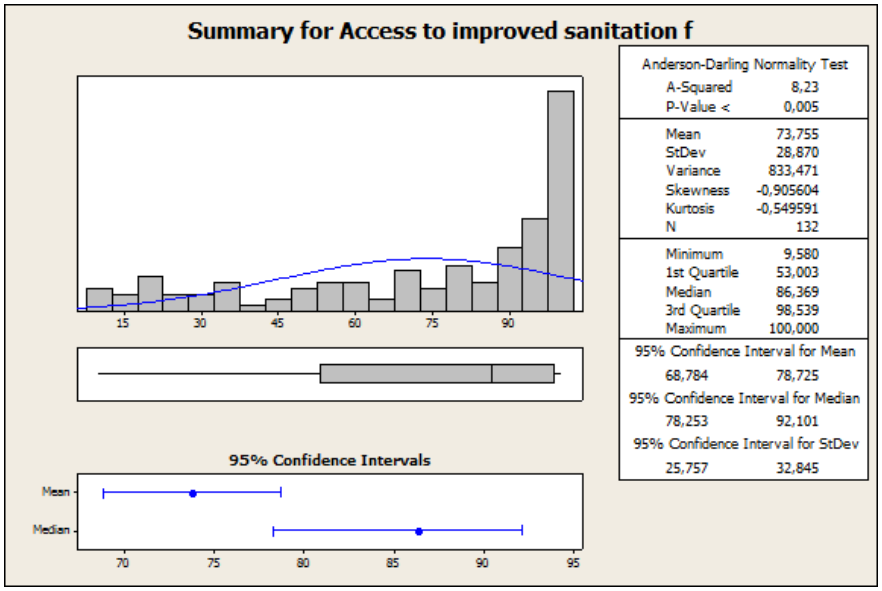
Aqui se apresenta uma análise interessante, onde praticamente podemos ver quase "duas distribuições" de dados... Um dos substratos de 0 a 14, representados basicamente por África e Ásia, e uma segunda com países de Europa, Austrália e Nova Zelandia, Ibéricos e Estados Unidos. Na parte superior voltam alguns países da Western Asia como Emirados Árabes, Kwait, Egito, Arábia Saudita.

Gráfico 28: Descritivo – *Obesity* (normalizado)



Fonte: Autor – Dados Guia GPS.

Gráfico 28: Descritivo – *Access to improved sanitation* (normalizado)

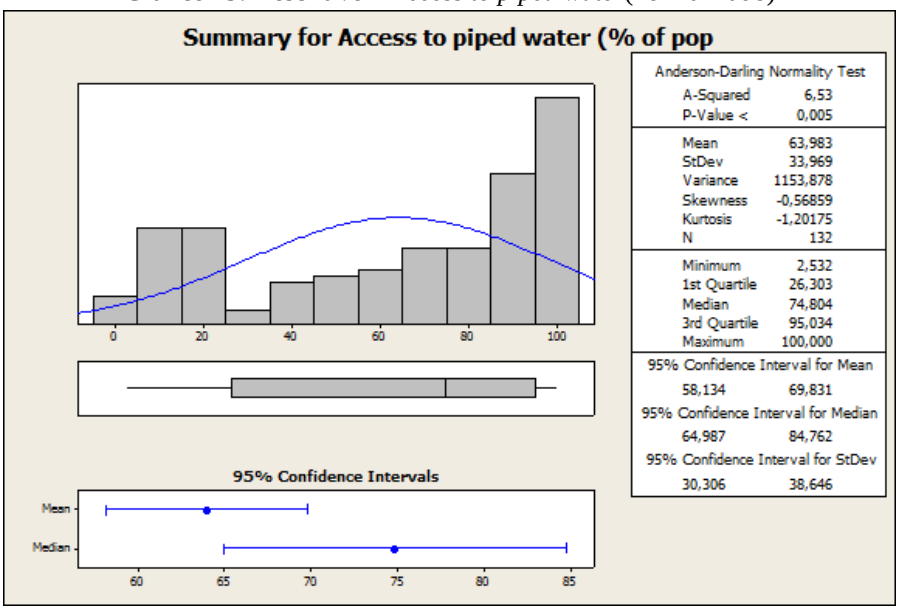


Fonte: Autor – Dados Guia GPS.

Análise:

Aqui vemos a Mediana à direita da média, estando a média em 73,755 e mediana em 86,369. O mínimo em 9,58 e máximo em 100 (22 países com score 100 e mais 41 países entre 90 e 100 – maioria representada pela Europa e países mais desenvolvidos das outras áreas). A parte inferior mais uma vez para o grupo de países da África e Ásia (em sua maioria).

Gráfico 29: Descritivo – *Access to piped water*(normalizado)

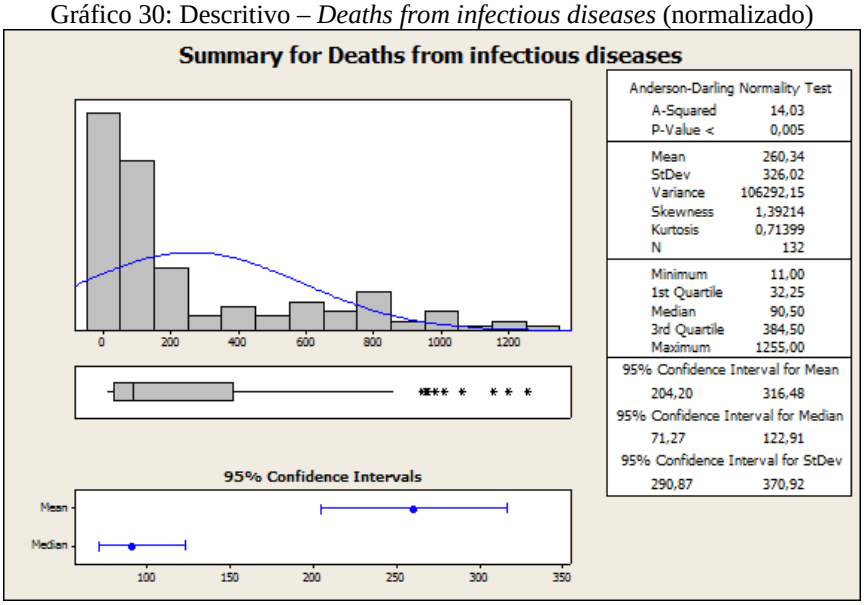


Fonte: Autor – Dados Guia GPS.

Análise:

Com poucas exceções, segue bastante o mesmo perfil do grupo de Instalações de Saneamento melhoradas. Análise bastante similar.

Gráfico 30: Descritivo – *Deaths from infectious diseases* (normalizado)



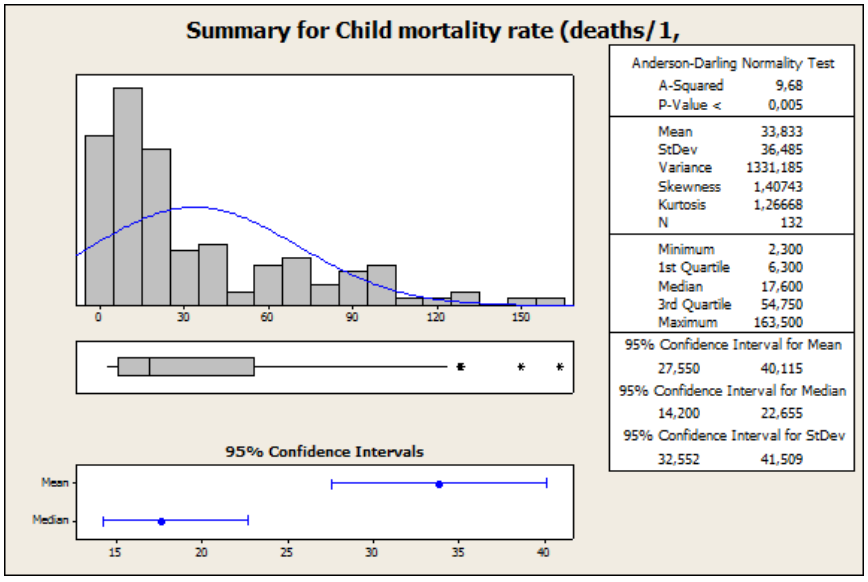
Fonte: Autor – Dados Guia GPS.

Análise:

Mediana (90,50) à Esquerda da Média (260,34 – quase 3 vezes maior) determinando intervalos de confiança bastante diferenciados. A dispersão é muito alta, com mínimo em 11 e máximo em 1255. A análise dos países nos substratos apresentados, com poucas **exceções são inversamente proporcionais a acesso a água encanada e instalações de saneamento melhoradas.**

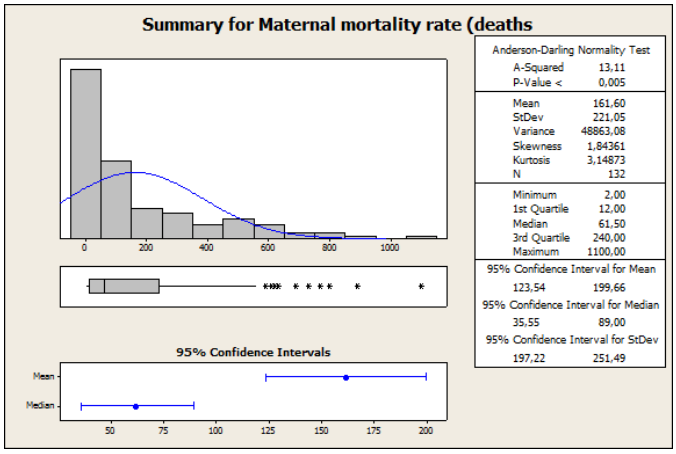
O MESMO VALE PARA OS PRÓXIMOS 3 GRÁFICOS – MORTALIDADE INFANTIL, MATERNAL MORTALITY RATE E STILLBIRTH RATE

Gráfico 31: Descritivo – *Child Mortality* (normalizado)



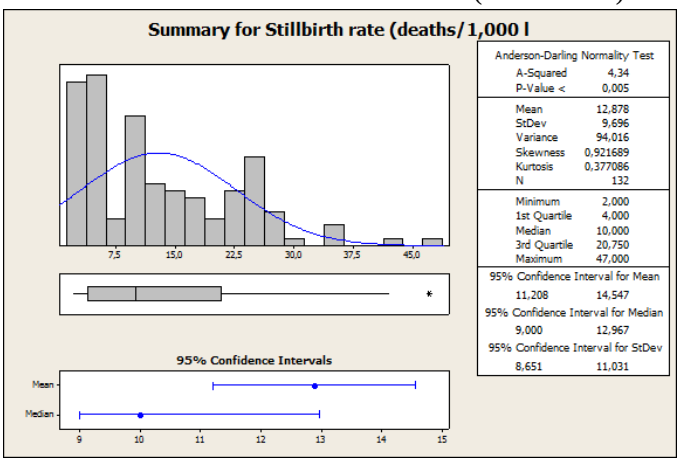
Fonte: Autor – Dados Guia GPS.

Gráfico 32: Descritivo – *Maternal Mortality* (normalizado)



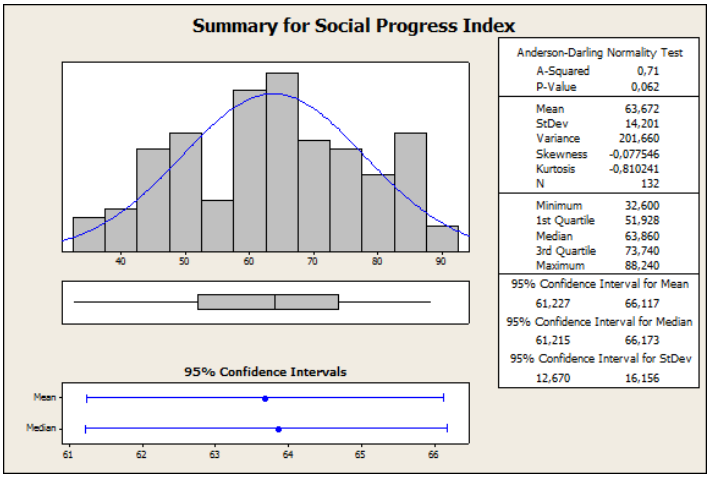
Fonte: Autor – Dados Guia GPS.

Gráfico 33: Descritivo – *Stillbirth rate* (normalizado)



Fonte: Autor – Dados Guia GPS.

Gráfico 34: Descritivo – *Social Progress Index* (normalizado)



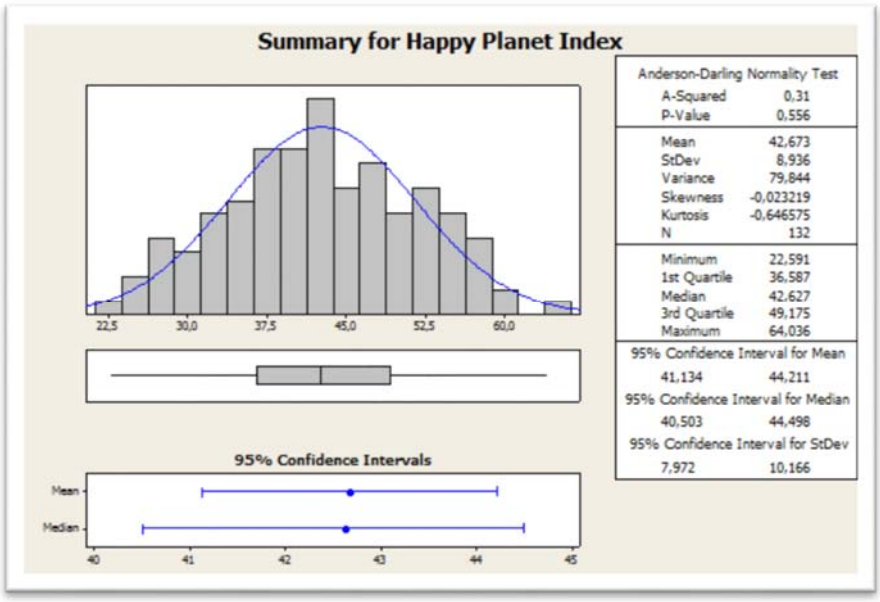
Fonte: Autor – Dados Guia GPS.

Análise Conjunta:

Índice de Progresso Social (acima) e Planeta Feliz (abaixo) seguem uma distribuição normal, com média e mediana muito próximas. O Progresso social apresenta alguns substratos

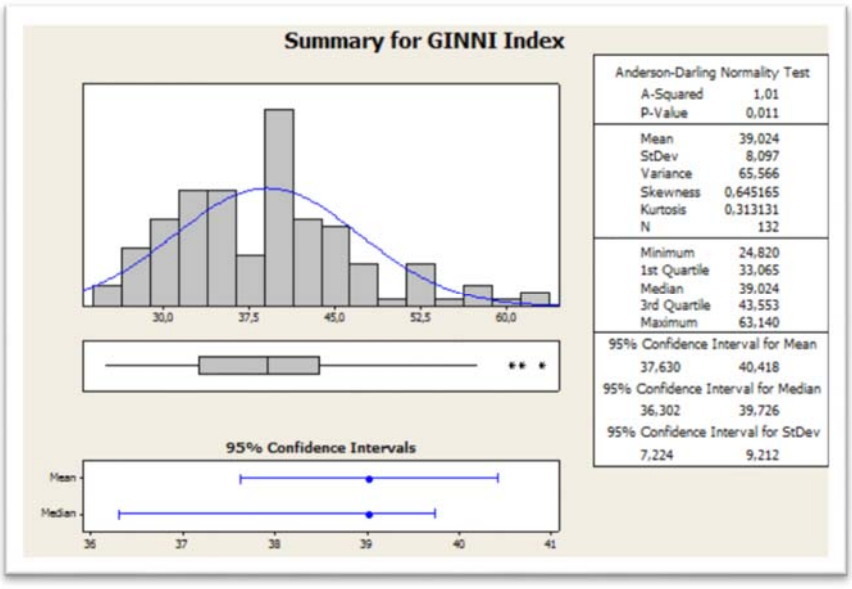
diferenciados (55 com poucas mostras e 85 com muitas), porém o Planeta feliz segue bastante a distribuição normal incluindo os substratos.

Gráfico 35: Descritivo – *Happy Planet Index* (normalizado)



Fonte: Autor – Dados Guia GPS.

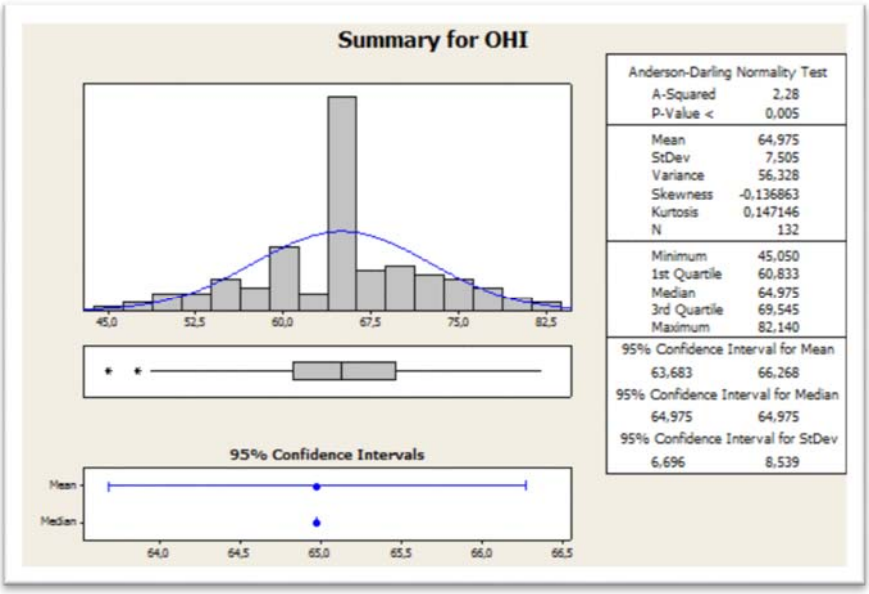
Gráfico 36: Descritivo – *Ginni* (normalizado)



Fonte: Autor – Dados Guia GPS.

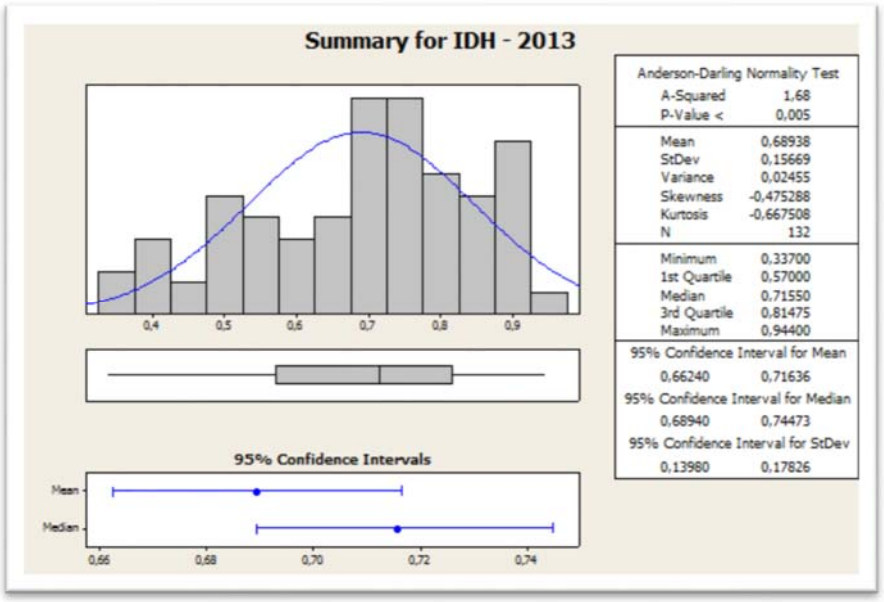
Análise: Para o Índice de Ginni, houve um pico cuja mediana foi de 39,024, equivalente à média. Tem uma distribuição bem normalizada, com uma maior quantidade de países à direita.

Gráfico 37: Descritivo – *OHI* (normalizado)



Fonte: Autor – Dados Guia GPS.

Gráfico 38: Descritivo – *IDH* (normalizado)

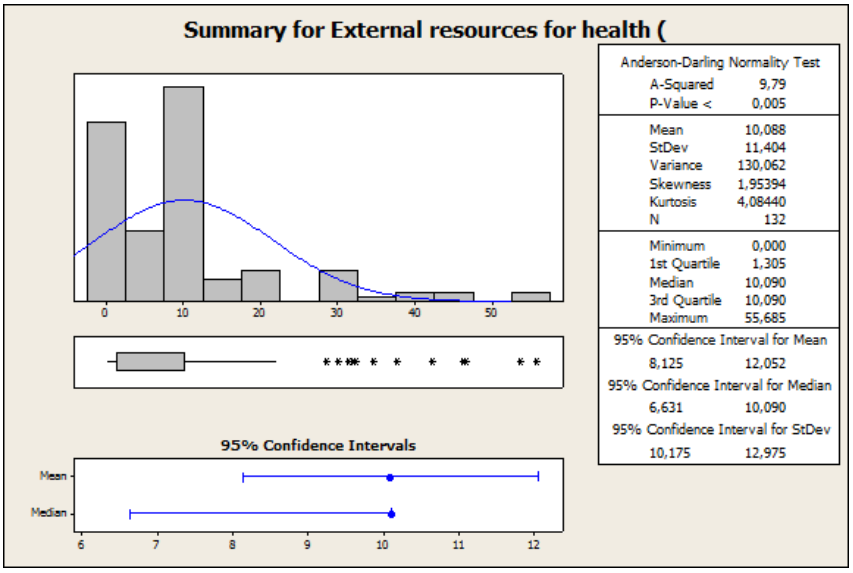


Fonte: Autor – Dados Guia GPS.

Análise:

A mediana de 71% está bem próxima à média de 69% aproxmadamente. Do lado esquerdo do gráfico, se apresentam os países com baixo índice de IDH, os países chamados subdesenvolvidos, ou aproximando-se mais à média, em desenvolvimento.

Gráfico 39: Descritivo – *External Resources for Health* (normalizado)

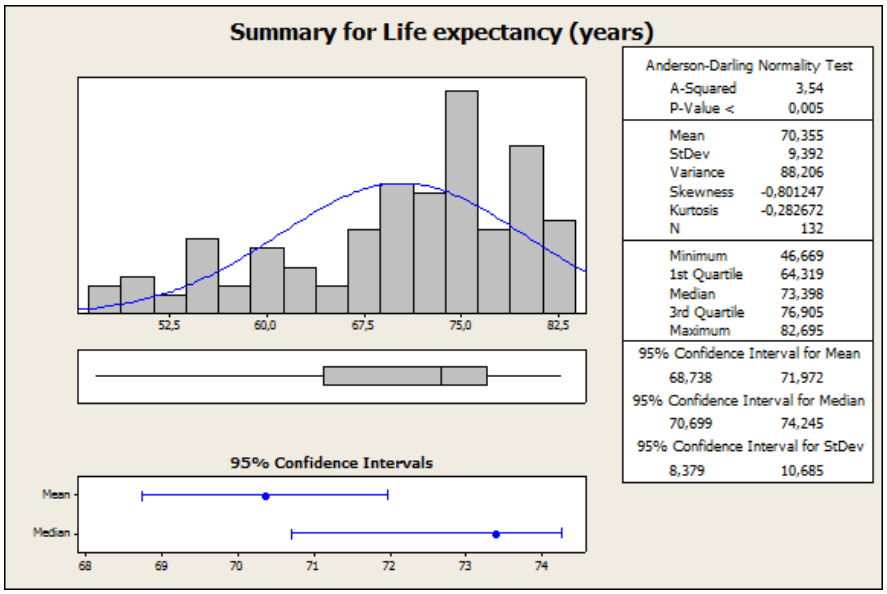


Fonte: Autor – Dados Guia GPS.

Análise:

Percebe-se uma distribuição com uma concentração pico à esquerda do gráfico Média de 10%, igual à Mediana, e uma maior quantidade de países à direita**.**

Gráfico 40: Descritivo – *Life Expectancy* (normalizado)



Fonte: Autor – Dados Guia GPS.

Análise:

Observa-se uma distribuição próxima à normal, com média e mediana bem próximas, respectivamente 70,35% e 73,39%. Uma concentração de países maior à direita do gráfico.

Considerações Preliminares:

Após as análises iniciais, parece que existe uma correlação forte entre o nível de desenvolvimento dos fatores ambientais (água tratada, esgotos, poluição interna e externa) com o nível de mortes por infecções, mortalidade infantil e maternal. Averigüaremos a seguir.

3.5. MATRIZ DE CORRELAÇÃO

A matriz de correlação com todas as 20 variáveis (Tabela 5) foi feita no decorrer do semestre e, devido a este fato e também ao seu tamanho, ela não será colocada no presente trabalho. Algumas correlações mais relevantes serão apresentadas (Quadro 2).

A matriz de correlação linear apresenta o teste de significância *P-Value*. Para a correlação foi utilizado o índice de *Pearson*. A correlação é sempre um número entre 0 e 1 e mede a intensidade de relações entre as variáveis, se positiva ou negativa. Abaixo está a matriz de correlação feita com as variáveis estudadas.

Tabela 5. Matriz de Correlação

	Social Progress	Maternal mortali	Stillbirth rate
Maternal mortali	-0,782 0,000		
Stillbirth rate	-0,805 0,000	0,739 0,000	
Child mortality	-0,843 0,000	0,901 0,000	0,813 0,000
Deaths from infe	-0,747 0,000	0,871 0,000	0,759 0,000
Access to piped	0,846 0,000	-0,799 0,000	-0,779 0,000
Access to improv	0,791 0,000	-0,814 0,000	-0,779 0,000
Indoor air pollu	-0,710 0,000	0,621 0,000	0,605 0,000
Life expectancy	0,855 0,000	-0,832 0,000	-0,789 0,000
Non-communicable	-0,748 0,000	0,547 0,000	0,580 0,000
Obesity rate (%	0,574 0,000	-0,621 0,000	-0,601 0,000
Outdoor air poll	0,096 0,273	-0,260 0,003	-0,191 0,028
Health expenditu	0,593 0,000	-0,411 0,000	-0,532 0,000
External resourc	-0,358 0,000	0,549 0,000	0,430 0,000
IDH - 2013_1	0,868 0,000	-0,817 0,000	-0,803 0,000
1. GovernanÁa_1	0,890 0,000	-0,545 0,000	-0,613 0,000
EPI Score_1	0,917 0,000	-0,702 0,000	-0,760 0,000
Happy Planet Ind	0,296 0,001	-0,431 0,000	-0,265 0,002
OHI_1	0,561 0,000	-0,363 0,000	-0,446 0,000
GINNI Index_1	-0,308 0,000	0,305 0,000	0,346 0,000

	Child mortality	Deaths from infe	Access to piped
Deaths from infe	0,888 0,000		
Access to piped	-0,837 0,000	-0,805 0,000	
Access to improv	-0,833 0,000	-0,804 0,000	0,856 0,000
Indoor air pollu	0,670	0,636	-0,803

	0,000	0,000	0,000
Life expectancy	-0,906	-0,924	0,831
	0,000	0,000	0,000
Non-communicable	0,616	0,601	-0,684
	0,000	0,000	0,000
Obesity rate (%	-0,646	-0,594	0,731
	0,000	0,000	0,000
Outdoor air poll	-0,253	-0,295	0,282
	0,003	0,001	0,001
Health expenditu	-0,442	-0,316	0,503
	0,000	0,000	0,000
External resourc	0,508	0,654	-0,561
	0,000	0,000	0,000
IDH - 2013_1	-0,848	-0,763	0,826
	0,000	0,000	0,000
1. GovernanÁa_1	-0,609	-0,469	0,633
	0,000	0,000	0,000
EPI Score_1	-0,776	-0,697	0,813
	0,000	0,000	0,000
Happy Planet Ind	-0,448	-0,465	0,370
	0,000	0,000	0,000
OHI_1	-0,383	-0,296	0,390
	0,000	0,001	0,000
GINNI Index_1	0,311	0,457	-0,247
	0,000	0,000	0,004

	Access to improv	Indoor air pollu	Life expectancy
Indoor air pollu	-0,706		
	0,000		
Life expectancy	0,808	-0,666	
	0,000	0,000	
Non-communicable	-0,546	0,657	-0,756
	0,000	0,000	0,000
Obesity rate (%	0,704	-0,663	0,554
	0,000	0,000	0,000
Outdoor air poll	0,278	-0,166	0,196
	0,001	0,057	0,024
Health expenditu	0,449	-0,410	0,425
	0,000	0,000	0,000
External resourc	-0,534	0,439	-0,525
	0,000	0,000	0,000
IDH - 2013_1	0,834	-0,689	0,819
	0,000	0,000	0,000
1. GovernanÁa_1	0,547	-0,553	0,645
	0,000	0,000	0,000
EPI Score_1	0,782	-0,717	0,807
	0,000	0,000	0,000
Happy Planet Ind	0,373	-0,258	0,508
	0,000	0,003	0,000
OHI_1	0,413	-0,360	0,378
	0,000	0,000	0,000
GINNI Index_1	-0,346	0,215	-0,424
	0,000	0,013	0,000

	Non-communicable	Obesity rate (%	Outdoor air poll
Obesity rate (%	-0,404		
	0,000		
Outdoor air poll	0,137	0,252	
	0,117	0,004	
Health expenditu	-0,453	0,350	0,038
	0,000	0,000	0,669
External resourc	0,314	-0,509	-0,275
	0,000	0,000	0,001
IDH - 2013_1	-0,628	0,692	0,195
	0,000	0,000	0,025
1. GovernanÁa_1	-0,663	0,348	-0,011

	0,000	0,000	0,903
EPI Score_1	-0,695	0,594	0,147
	0,000	0,000	0,093
Happy Planet Ind	-0,416	0,222	-0,070
	0,000	0,011	0,427
OHI_1	-0,304	0,195	0,100
	0,000	0,025	0,252
GINNI Index_1	0,132	-0,097	-0,285
	0,131	0,268	0,001

	Health expenditu	External resourc	IDH - 2013_1
External resourc	-0,047		
	0,591		
IDH - 2013_1	0,562	-0,494	
	0,000	0,000	
1. GovernanÁa_1	0,595	-0,071	0,697
	0,000	0,421	0,000
EPI Score_1	0,563	-0,329	0,824
	0,000	0,000	0,000
Happy Planet Ind	0,109	-0,419	0,333
	0,215	0,000	0,000
OHI_1	0,353	0,037	0,481
	0,000	0,673	0,000
GINNI Index_1	-0,181	0,129	-0,297
	0,038	0,139	0,001

	1. GovernanÁa_1	EPI Score_1	Happy Planet Ind
EPI Score_1	0,811		
	0,000		
Happy Planet Ind	0,089	0,197	
	0,312	0,024	
OHI_1	0,625	0,545	-0,143
	0,000	0,000	0,103
GINNI Index_1	-0,247	-0,365	0,051
	0,004	0,000	0,562

	OHI_1
GINNI Index_1	-0,293
	0,001

Cell Contents: Pearson correlation
P-Value

Fonte: Preparado pela autora

Quadro 2. Correlações >0,8

Child Mortality	e	Matern. Mortality	0,901	
Child Mortality	e	Stillbirth	0,813	
Deaths from In.	e	Matern. Mortalit.	0,871	
Access to Piped	e	Social Progress	0,846	
Life Expectancy	e	Social Progress	0,855	
IDH	e	Social Progress	0,868	
Governança	e	Social Progress	0,890	
EPI	e	Social Progress	0,917	
Deaths from Infe.	e	Child Mortality	0,888	
Access to Improv.	e	Access to Piped	0,856	
Life Expectancy	e	Access to Piped	0,831	
IDH	e	Access to Piped	0,826	
EPI	e	Access to Piped	0,813	
Life Expectancy	e	Access to Improv.	0,808	
IDH	e	Access to Improv.		0,834
EPI Score	e	Life Expectancy	0,807	
EPI	e	IDH		0,824
EPI	e	Governança		0,811

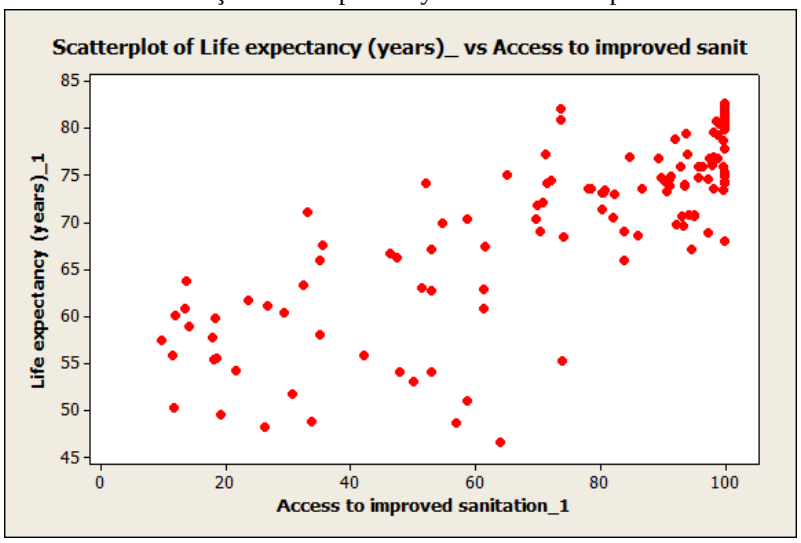
Fonte: Preparado pela autora

Serão apresentados neste trabalho os gráficos de dispersão das correlações acima de 0,86, item a seguir.

3.6. GRÁFICOS DE DISPERSÃO

Os gráficos com correlações destacadas no Quadro 2. serão mostrados e analisados a continuação.

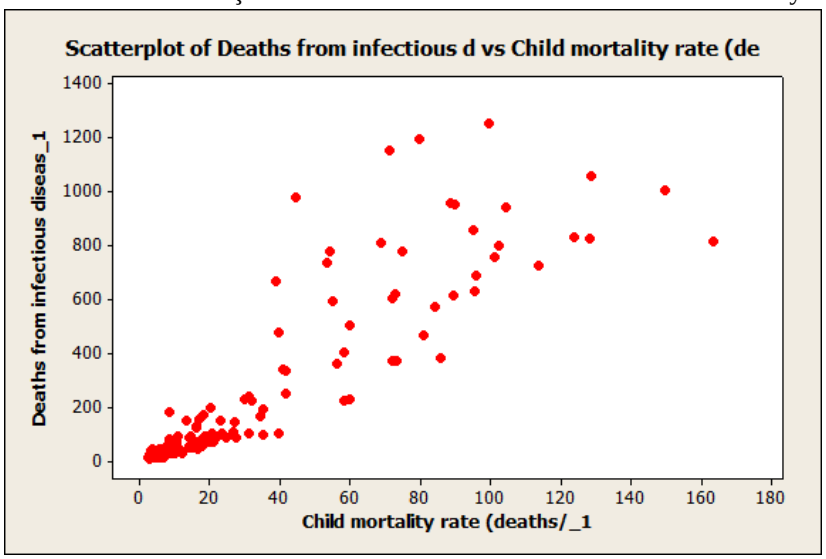
Gráfico 41: Correlação Life Expectancy vs Access to Improved Sanitation



Fonte: Autor – Dados Guia GPS.

Análise:
Correlação entre *access to improved sanitation* e *life expectancy* na ordem de 0,808. Ou seja, quanto maior o acesso a melhorias de esgoto e saneamento, maior a expectativa de vida da população.

Gráfico 42: Correlação Deaths from infectious desease vs child mortality

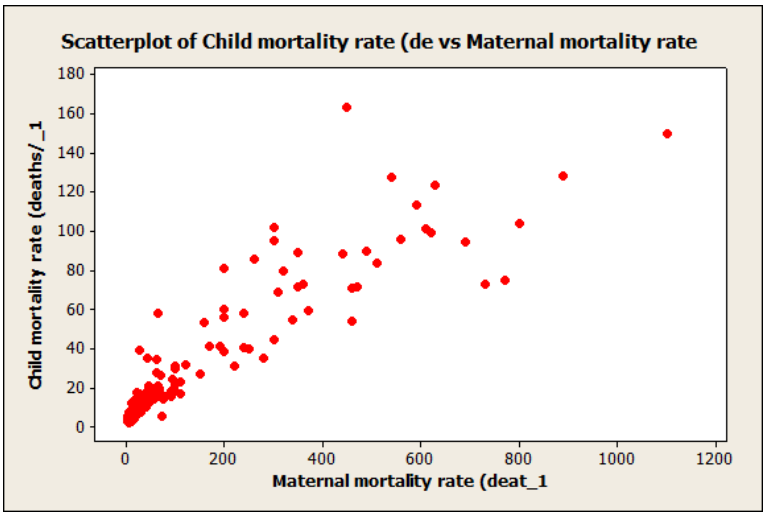


Fonte: Autor – Dados Guia GPS.

Análise:
A correlação entre ambos ficou na ordem de 0,888. Neste gráfico vemos uma concentração de dados quando temos mortalidade infantil abaixo de 40, e também vemos que a mortalidade por doenças infecciosas também é baixa (abaixo de 200, no quartil inferior). Com mortalidade infantil acima de 40, vemos grande dispersão de dados, porém com claro crescente, sendo uma variação de 200 a 1300 no tocante a mortes por doenças infecciosas.

A mesma análise vale para os próximos 2 gráficos: Gráfico 43. *Maternal mortality x child mortality*, Gráfico 44. *Maternal mortality x Deaths from infectious desease.*
A correlação entre maternal mortality e child mortality foi de 0,901, apresentando uma correlação altíssima, ver gráfico 43.

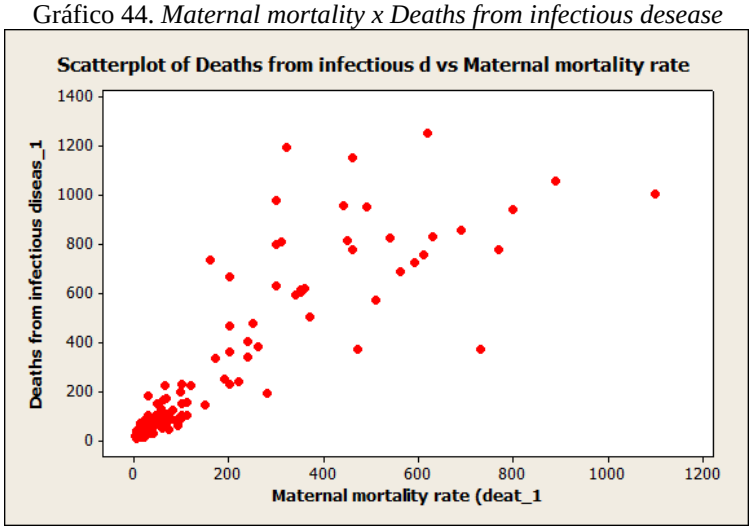
Gráfico 43. *Maternal mortality x child mortality*



Fonte: Autor – Dados Guia GPS.

A correlação entre Mortalidade Maternal e Morte por Doença Infecciosa foi de 0,871, o que indica uma forte correlação positiva. Tal fato pode ser visualizado no Gráfico 44 a seguir.

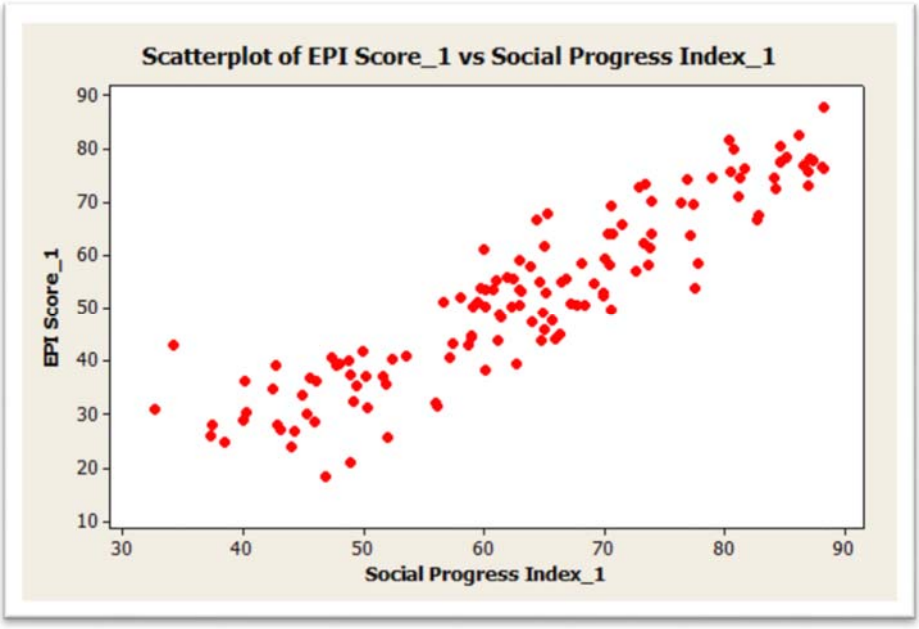
Gráfico 44. *Maternal mortality x Deaths from infectious desease*



Fonte: Autor – Dados Guia GPS.

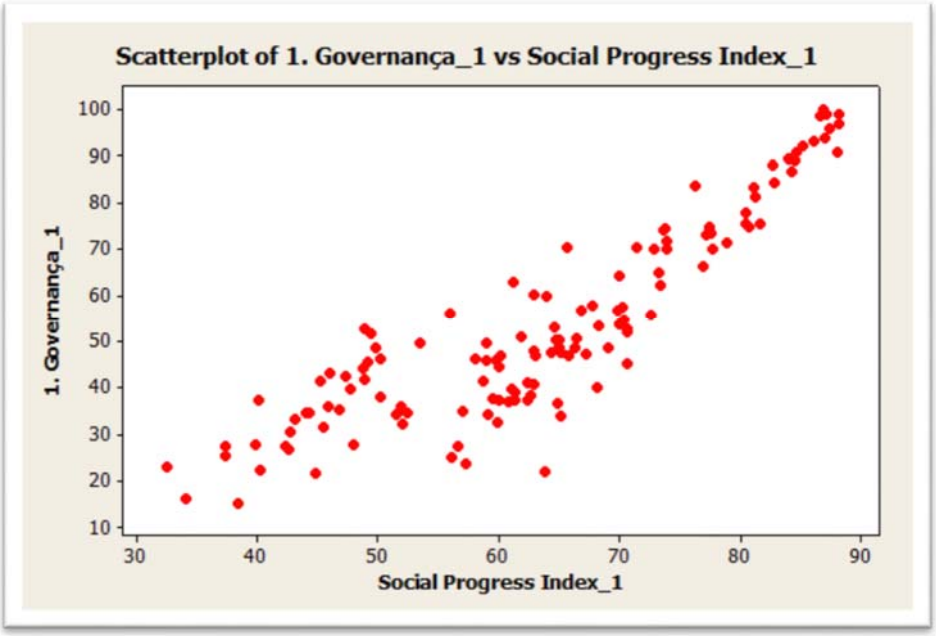
Para os próximos 4 gráficos (45, 46 e 47)(Social progress x EPI, Governança x Social Progress e IDH vs. Social PRogress) vemos uma forte correlação linear entre o Índice de progressão social e as variáveis IDH, governança e EPI score. Algumas com mais dispersão ou outliers que outras, porém todos demonstram claramente uma relação linear crescente.

Gráfico 45. *EPI x Social Progress*



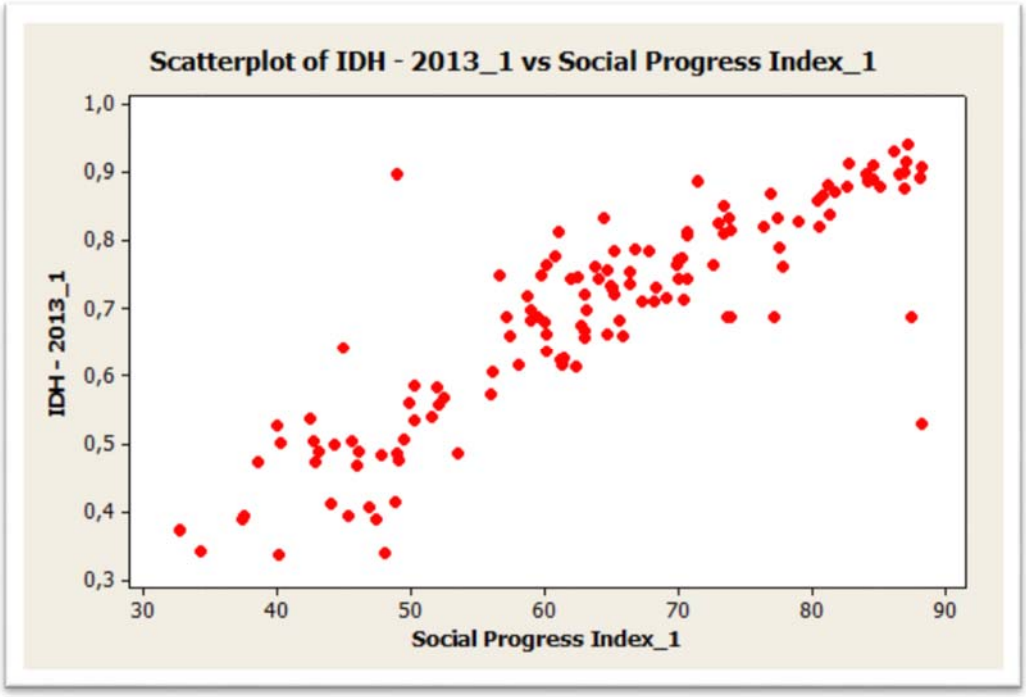
Fonte: Autor – Dados Guia GPS.

Gráfico 46. Governança *x Social Progress*



Fonte: Autor – Dados Guia GPS.

Gráfico 47. IDH *x Social Progress*



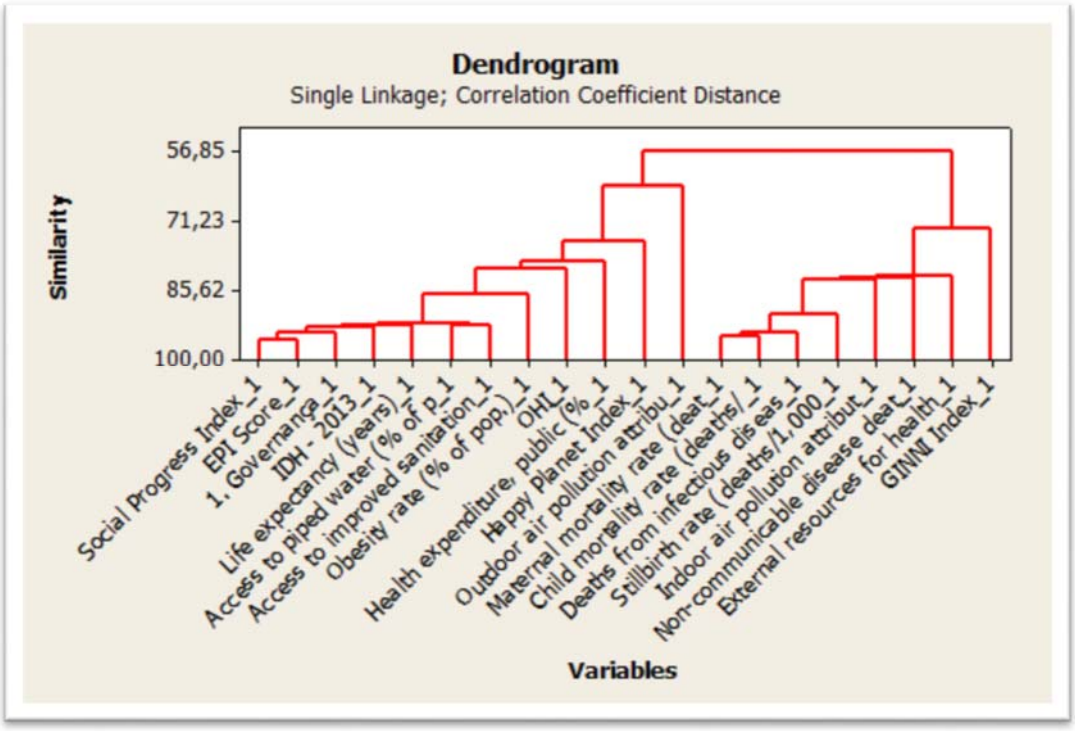
Fonte: Autor – Dados Guia GPS.

Uma vez que agora já se tem uma idéia de como se relacionam as variáveis, será importante identificar como tais variáveis podem ser agrupadas. Seguindo com a análise, será apresentado, portanto, o dendograma, o qual indica as variáveis mais próximas em relação ao coeficiente de correlação.

3.7. ANÁLISE DE CLUSTERS E DENDOGRAMAS

Um Dendograma é um tipo específico de diagrama ou representação icônica que organiza determinados fatores e variáveis. É um diagrama de similaridade. O objetivo é unificar objetivos em *clusters* ou grupos sucessivamente maiores através da utilização de medida de similaridade ou de distância. Abaixo se apresenta o Dendograma que apresenta a similaridade entre as variáveis, com a escolha de 1 cluster, por isso, uma única cor.

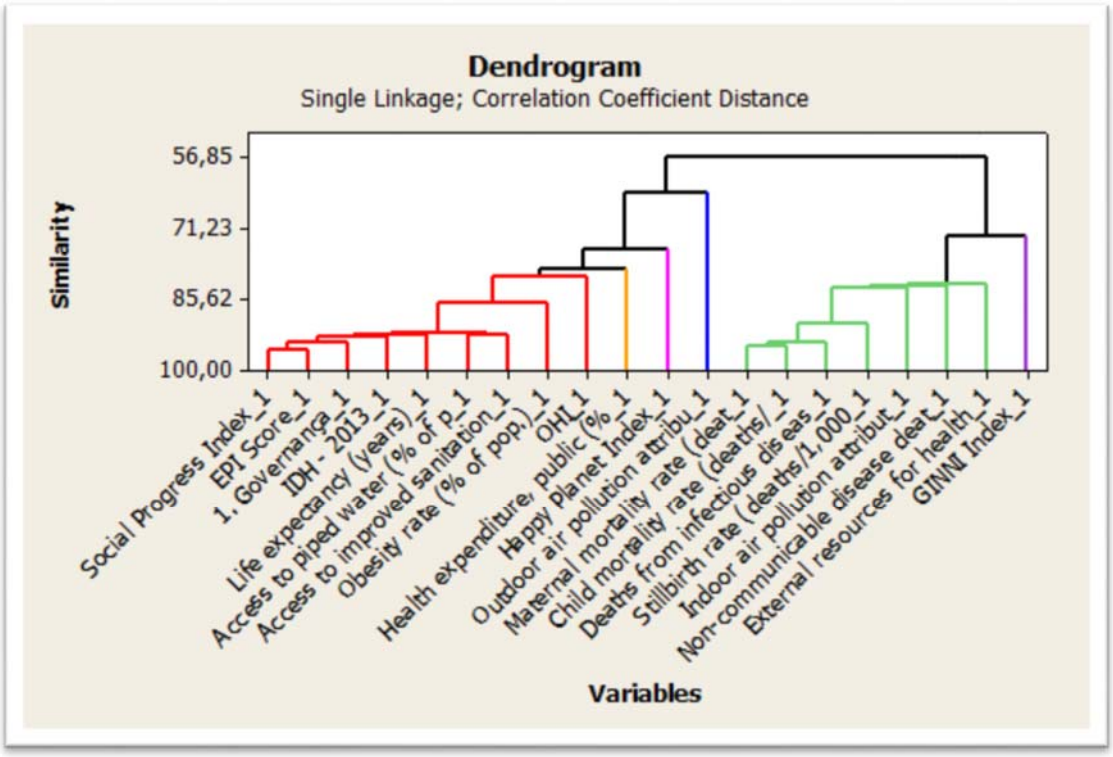
Gráfico 48 – Dendograma das Variáveis, Cluster Único



Fonte: Preparado pela autora

No gráfico seguinte, se confeccionou um dendograma com 6 clusters, com base no que foi visto acima. Pelos níveis, do mais alto ao mais baixo, se pôde observar que há 6 grupos ou classes de variáveis similares, agrupadas por cores abaixo.

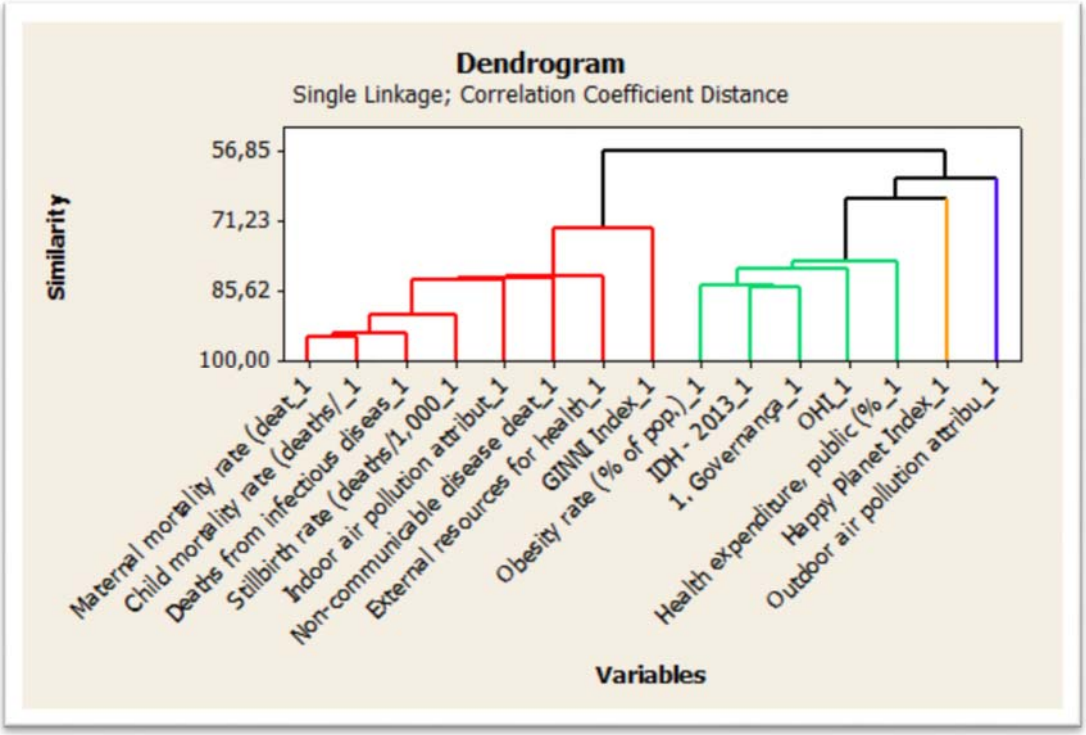
Gráfico 49 – Dendograma das Variáveis, 6 Clusters



Fonte: Preparado pela autora

Pelo critério da parcimônia, foi escolhida Governança como representante do cluster lateralizado à esquerda e como substituto de: Social Progress Index, Epi, IDH, Life Expectancy, Access to piped water, Access to improved sanitation. Dessa forma, ficou assim:

Gráfico 50 – Dendograma das Variáveis, 4 Clusters



Fonte: Preparado pela autora

Dando continuação a análise, será proposto um modelo de regressão para relacionar as variáveis.

3.8. REGRESSÃO LINEAR

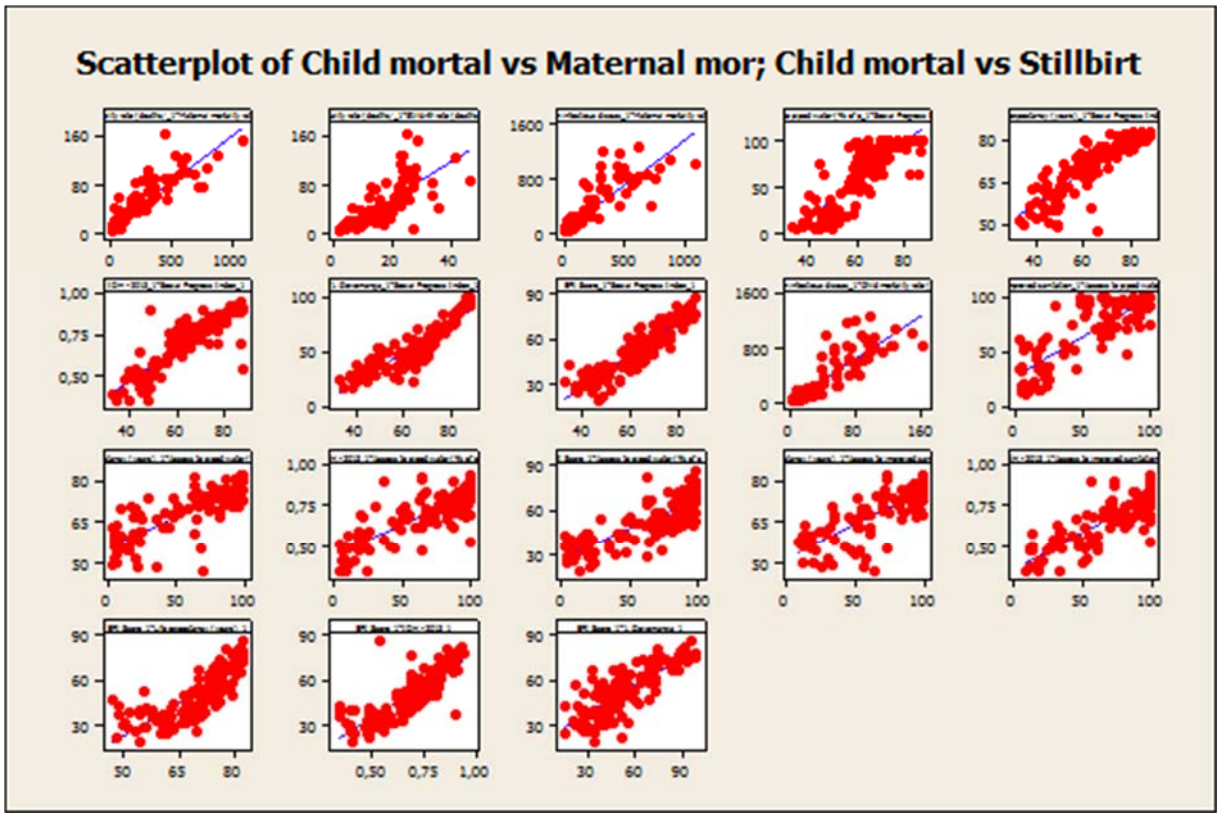
A correlação mede a direção e a intensidade da relação linear (linha reta) entre duas variáveis quantitativas. Se um diagrama de dispersão mostra uma relação linear, é interessante resumir esse padrão geral traçando uma reta no diagrama de dispersão. Uma reta de regressão resume a relação entre duas variáveis, mas somente em um contexto específico: quando uma das variáveis ajuda a explicar ou predizer a outra, ou seja, a regressão descreve uma relação entre uma variável explanatória e uma variável resposta.

A regressão linear assume sempre a forma de uma equação linear:

Y = a + bx, sendo:
Y = Variável dependente;
a = uma constante, o intercepto;
b = a inclinação na reta;
x = variável independente ou explicativa.

O “b”, ou seja, a declividade é dada pela multiplicação do índice de correlação pela divisão dos desvios-padrão entre as variáveis x e y. E “a” é dada pela média de “Y” menos a multiplicação de “b” pela média de “x”. Assim, percebe-se muito claramente que a regressão depende da correlação entre as variáveis, além de medidas de centro de cada uma das variáveis.

Gráficos Conjuntos de Regressão Linear Múltipla n. 51



Fonte: Preparado pela autora

Análise:

Aqui, apresenta-se de forma resumida a relação linear múltipla entre as variáveis. A análise é similar a que foi feita nos gráficos de dispersão, e por isso mesmo, não será repetida aqui caso a caso. A regressão nada mais é do que a expressão numérica da reta de tendência que foram vistos nos gráficos de dispersão.

A seguir, através da técnica StepWise, se entenderá melhor o peso de cada uma delas na Saúde e Bem-Estar de cada país analisado.

Em primeiro lugar, devido ao Dendograma, Gráfico 50, foi escolhida a variável Social Progress para isolá-la e compará-la com as outras, pelo critério da Parcimônia.

Neste caso, não foram usadas as variáveis Governança, EPI, IDH, Life Expec., Access to piped water, nem Access to Improved Sanitation, pelo Dendograma, pois, a principio, daria “no mesmo” que usar a Social Progress, pois esta última seria suficiente para explicar o fenômeno.

Regression Analysis: Social Progr versus Maternal mor; Stillbirth r; ...

```
The regression equation is
Social Progress Index_1 = 63,1
                      - 0,00860 Maternal mortality rate (deat_1
                           - 0,191 Stillbirth rate (deaths/1,000_1
                           - 0,143 Child mortality rate (deaths/_1
                           + 0,00294 Deaths from infectious diseas_1
                           - 0,0193 Indoor air pollution attribut_1
                           - 0,533 Non-communicable disease deat_1
                           + 0,0115 Obesity rate (% of pop.)_1
                           - 0,0499 Outdoor air pollution attribu_1
                           + 0,0956 Health expenditure, public (%_1
                           + 0,0020 External resources for health_1
                           - 0,0797 Happy Planet Index_1 + 0,330 OHI_1
                           - 0,0464 GINNI Index_1
```

Tabela 6. Regressão com Social Progress como Variável Preditora

Predictor	Coef	SE Coef	T	P
Constant	63,136	8,851	7,13	0,000
Maternal mortality rate (deat_1	-0,008598	0,005283	-1,63	0,106
Stillbirth rate (deaths/1,000_1	-0,19084	0,09272	-2,06	0,042
Child mortality rate (deaths/_1	-0,14318	0,03918	-3,65	0,000
Deaths from infectious diseas_1	0,002940	0,004649	0,63	0,528
Indoor air pollution attribut_1	-0,01925	0,01633	-1,18	0,241
Non-communicable disease deat_1	-0,5331	0,1068	-4,99	0,000
Obesity rate (% of pop.)_1	0,01146	0,07670	0,15	0,882
Outdoor air pollution attribu_1	-0,04990	0,04335	-1,15	0,252
Health expenditure, public (%_1	0,09562	0,03503	2,73	0,007
External resources for health_1	0,00195	0,06227	0,03	0,975
Happy Planet Index_1	-0,07970	0,06900	-1,16	0,250
OHI_1	0,32954	0,07817	4,22	0,000
GINNI Index_1	-0,04637	0,07591	-0,61	0,542

Fonte: Elaborado pela autora

S = 5,21214 R-Sq = 87,9% (R quadrado) R-Sq(adj) = 86,5% (R quadrado ajustado)

Análise:

A regressão nada mais é o valor da reta de tendência que foram vistos nos gráficos de dispersão. Esta equação tem um poder explicativo de mais de 87,9% (ver Tabela 6), como se pode ver acima, que é o R-Quadrado, considerando o modelo em que o Social Progress é a variável dependente. Decisão: P > 5% => retirar, grau de imprecisão de explicação.

Tabela 7. Análise de Variância com SPI como Variável Dependente

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	13	23211,8	1785,5	65,73	0,000
Residual Error	118	3205,6	27,2		
Total	131	26417,5			

Source	DF	Seq SS
Maternal mortality rate (deat_1	1	16158,3
Stillbirth rate (deaths/1,000_1	1	2999,9
Child mortality rate (deaths/_1	1	785,3
Deaths from infectious diseas_1	1	67,9
Indoor air pollution attribut_1	1	777,5
Non-communicable disease deat_1	1	1281,9
Obesity rate (% of pop.)_1	1	25,7
Outdoor air pollution attribu_1	1	3,9
Health expenditure, public (%_1	1	313,7
External resources for health_1	1	84,8
Happy Planet Index_1	1	199,9
OHI_1	1	502,9
GINNI Index_1	1	10,1

Fonte: Elaborado pela autora

Análise:
Como pode ser visto pela tabela 7, a confiabilidade é garantida com o p-value=0. A Tabela 8 apresenta os valores pouco usuais.

Tabela 8. Observações Pouco Usuais para SPI como Variável Dependente

Obs	Maternal mortality rate (deat_1	Social Progress Index_1	Fit	SE Fit	Residual	St Resid
2	97	59,130	70,858	1,467	-11,728	-2,34R
3	450	39,930	32,247	3,244	7,683	1,88 X
31	73	61,070	73,457	1,588	-12,387	-2,50R
55	63	44,840	62,261	1,502	-17,421	-3,49R
82	63	58,970	58,750	3,053	0,220	0,05 X
109	300	62,960	58,873	2,958	4,087	0,95 X
115	8	88,190	78,036	1,269	10,154	2,01R
131	200	40,230	50,594	1,327	-10,364	-2,06R

Fonte: Elaborado pela autora

Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Análise:
Recomenda-se retirar P> 5%, pelo grau de imprecisão de explicação: Retirar ainda então: Death…., Obesity…., External…., Ginni…. E gerar nova Regressão, conform epode ser visto na Tabela 9.

Regression Analysis: Social Progr versus Maternal mor; Stillbirth r; ...

The regression equation is
Social Progress Index_1 = 60,6 - 0,00725 Maternal mortality rate (deat_1
- 0,183 Stillbirth rate (deaths/1,000_1
- 0,135 Child mortality rate (deaths/_1
- 0,0194 Indoor air pollution attribut_1
- 0,512 Non-communicable disease deat_1
- 0,0526 Outdoor air pollution attribu_1
+ 0,104 Health expenditure, public (%_1
- 0,0927 Happy Planet Index_1 + 0,340 OHI_1

Tabela 9. Regressão Pós "Limpeza da Tabela 8"

Predictor	Coef	SE Coef	T	P
Constant	60,649	7,542	8,04	0,000
Maternal mortality rate (deat_1	-0,007255	0,004743	-1,53	0,129
Stillbirth rate (deaths/1,000_1	-0,18257	0,08696	-2,10	0,038
Child mortality rate (deaths/_1	-0,13532	0,03509	-3,86	0,000
Indoor air pollution attribut_1	-0,01942	0,01438	-1,35	0,179
Non-communicable disease deat_1	-0,51158	0,09977	-5,13	0,000
Outdoor air pollution attribu_1	-0,05261	0,04086	-1,29	0,200
Health expenditure, public (%_1	0,10417	0,03220	3,24	0,002
Happy Planet Index_1	-0,09269	0,06519	-1,42	0,158
OHI_1	0,34045	0,07303	4,66	0,000

Fonte: Elaborado pela autora

S = 5,13924 R-Sq = 87,8% R-Sq(adj) = 86,9%

Tabela 10. Nova Análise de Variância

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	9	23195,2	2577,2	97,58	0,000
Residual Error	122	3222,2	26,4		
Total	131	26417,5			

Source	DF	Seq SS

Maternal mortality rate (deat_1	1	16158,3
Stillbirth rate (deaths/1,000_1	1	2999,9
Child mortality rate (deaths/_1	1	785,3
Indoor air pollution attribut_1	1	729,9
Non-communicable disease deat_1	1	1170,1
Outdoor air pollution attribu_1	1	30,7
Health expenditure, public (%_1	1	453,4
Happy Planet Index_1	1	293,7
OHI_1	1	574,0

Fonte: Elaborado pela autora

Análise:

O *P-Value* deve ser o mais próximo a zero que demostra a confiabilidade, e apresenta R-Sq considerável para haver um grau de explicação e distinguir uma média entre a variável analisada.

Tabela 11. Observações Pouco Usuais para SPI como Variável Dependente

Obs	Maternal mortality rate (deat_1	Social Progress Index_1	Fit	SE Fit	Residual	St Resid
2	97	59,130	70,682	1,188	-11,552	-2,31R
3	450	39,930	32,596	3,119	7,334	1,80 X
31	73	61,070	73,598	1,512	-12,528	-2,55R
55	63	44,840	61,668	1,136	-16,828	-3,36R
82	63	58,970	59,364	2,714	-0,394	-0,09 X
115	8	88,190	77,646	0,968	10,544	2,09R
131	200	40,230	50,705	1,120	-10,475	-2,09R

Fonte: Elaborado pela autora

Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Tabela12. StepWise Social Progress como Variável Preditora

Stepwise Regression: Social Progr versus Maternal mor; Stillbirth r; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is Social Progress Index_1 on 19 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	22,36	22,97	39,93	37,42	38,50	39,18
EPI Score_1	0,787	0,491	0,242	0,176	0,177	0,165
T-Value	26,28	12,22	6,81	4,87	5,02	4,80
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
1. GovernanÁa_1		0,280	0,294	0,299	0,288	0,304
T-Value		9,09	13,56	14,81	14,26	14,93
P-Value		0,000	0,000	0,000	0,000	0,000
Child mortality rate (deaths/_1			-0,137	-0,103	-0,106	-0,072
T-Value			-11,54	-7,73	-8,12	-4,24
P-Value			0,000	0,000	0,000	0,000
Access to piped water (% of p_1				0,071	0,079	0,068
T-Value				4,56	5,10	4,38
P-Value				0,000	0,000	0,000
Outdoor air pollution attribu_1					-0,058	-0,061
T-Value					-2,73	-2,98
P-Value					0,007	0,003
Deaths from infectious diseas_1						-0,0051
T-Value						-2,94
P-Value						0,004
S	5,67	4,45	3,12	2,91	2,84	2,75
R-Sq	84,16	90,35	95,27	95,94	96,16	96,41
R-Sq(adj)	84,04	90,20	95,16	95,81	96,01	96,24
Mallows Cp	430,4	214,3	42,8	21,3	15,3	8,6

Fonte: Elaborado pela autora

Análise:
A primeira variável EPI explicou 84,16%, a segunda, Governança, apresentou uma explicação acumulada de 90,35%, e a Terceira, Child Mortality, 95,27%. Então, não há necessidade de gerar uma nova regressão, pois o R2 são bem altos, de maneira que o modelo 3 já é suficiente para explicar o Social Progress Index.

Tabela 12. Continuação

Step	7	8	9
Constant	35,15	37,35	34,68
EPI Score_1	0,150	0,143	0,152
T-Value	4,32	4,13	4,35
P-Value	0,000	0,000	0,000
1. GovernanÁa_1	0,297	0,296	0,300
T-Value	14,63	14,69	14,84
P-Value	0,000	0,000	0,000
Child mortality rate (deaths/_1	-0,059	-0,054	-0,048
T-Value	-3,33	-2,99	-2,61
P-Value	0,001	0,003	0,010
Access to piped water (% of p_1	0,060	0,058	0,054
T-Value	3,87	3,70	3,42
P-Value	0,000	0,000	0,001
Outdoor air pollution attribu_1	-0,061	-0,060	-0,050
T-Value	-3,01	-2,98	-2,34
P-Value	0,003	0,004	0,021
Deaths from infectious diseas_1	-0,0050	-0,0047	-0,0042
T-Value	-2,93	-2,74	-2,46
P-Value	0,004	0,007	0,015
IDH - 2013_1	7,6	6,4	6,0
T-Value	2,23	1,84	1,75
P-Value	0,027	0,068	0,082
Stillbirth rate (deaths/1,000_1		-0,081	-0,094
T-Value		-1,75	-1,99
P-Value		0,082	0,048
Happy Planet Index_1			0,050
T-Value			1,49
P-Value			0,138
S	2,71	2,69	2,68
R-Sq	96,55	96,63	96,69
R-Sq(adj)	96,35	96,41	96,45
Mallows Cp	5,7	4,7	4,6

Fonte: Elaborado pela autora

Tabela 13. Observações Pouco Usuais para SPI como V.D.

Obs	Maternal mortality rate (deat_1	Social Progress Index_1	Fit	SE Fit	Residual	St Resid
2	97	59,130	70,682	1,188	-11,552	-2,31R
3	450	39,930	32,596	3,119	7,334	1,80 X
31	73	61,070	73,598	1,512	-12,528	-2,55R
55	63	44,840	61,668	1,136	-16,828	-3,36R
82	63	58,970	59,364	2,714	-0,394	-0,09 X
115	8	88,190	77,646	0,968	10,544	2,09R
131	200	40,230	50,705	1,120	-10,475	-2,09R

Fonte: Elaborado pela autora

Análise:
Os países das linhas 2, 3, 31, 55, 82, 115 e 131 da Tabela 13, precisam de um ajuste correspondente aos valores da coluna residual para dar o valor da coluna FIT. ENTAO, RECOMENDA-SE SUBSTITUIR, NA PLANILHA COM AS MÉDIAS, CADA LINHA DE OBSERVAÇÃO (NÚMEROS ACIMA) DA COLUNA DE SOCIAL PROGRESS, PELO FIT.

Depois, segue-sem à normalização dos dados, através da equação (c9-MIN(c9))/(MAX(c9)-MIN(c9))*100 e se seleciona na caixa de opções as funções MINIMO e MAXIMO. Assim se cria uma planilha com os dados normalizados.

Como já explicado anteriormente, o Social Progress Index é uma variável bastante complexa para se analisar. Portanto, se tomará o IDH, Índice de Desenvolvimento Humano como Substituto e se analisará a Regressão do mesmo, uma vez que também pelo Dendograma, pelo critério da Parcimônia, seria uma variável possível de ser trabalhada como substituta.

Regression Analysis: IDH - 2013_1 versus EPI Score_1; Happy Planet; ...

```
The regression equation is
IDH - 2013_1 = - 0,000 + 0,000359 EPI Score_1 - 0,000013 Happy Planet Index_1
               + 0,00172 OHI_1 - 0,000268 GINNI Index_1
               + 0,00295 Social Progress Index_1
               - 0,000134 Maternal mortality rate (deat_1
               - 0,00129 Stillbirth rate (deaths/1,000_1
               - 0,000520 Child mortality rate (deaths/_1
               + 0,000142 Deaths from infectious diseas_1
               - 0,000565 Access to piped water (% of p_1
               + 0,000710 Access to improved sanitation_1
               - 0,000005 Indoor air pollution attribut_1
               + 0,00415 Life expectancy (years)_1
               + 0,00039 Non-communicable disease deat_1
               + 0,00299 Obesity rate (% of pop.)_1
               - 0,000030 Outdoor air pollution attribu_1
               + 0,000953 Health expenditure, public (%_1
               - 0,00217 External resources for health_1
```

Tabela 14. Regressão com IDH como variável Preditora

Predictor	Coef	SE Coef	T	P
Constant	-0,0000	0,2367	-0,00	1,000
EPI Score_1	0,0003594	0,0009538	0,38	0,707
Happy Planet Index_1	-0,0000129	0,0009154	-0,01	0,989
OHI_1	0,001718	0,001006	1,71	0,090
GINNI Index_1	-0,0002682	0,0009669	-0,28	0,782
Social Progress Index_1	0,002948	0,001408	2,09	0,038
Maternal mortality rate (deat_1	-0,00013423	0,00006602	-2,03	0,044
Stillbirth rate (deaths/1,000_1	-0,001295	0,001144	-1,13	0,260
Child mortality rate (deaths/_1	-0,0005202	0,0005266	-0,99	0,325
Deaths from infectious diseas_1	0,00014235	0,00006726	2,12	0,037
Access to piped water (% of p_1	-0,0005649	0,0004845	-1,17	0,246
Access to improved sanitation_1	0,0007099	0,0004595	1,54	0,125
Indoor air pollution attribut_1	-0,0000050	0,0002141	-0,02	0,981
Life expectancy (years)_1	0,004154	0,002828	1,47	0,145
Non-communicable disease deat_1	0,000386	0,001662	0,23	0,817
Obesity rate (% of pop.)_1	0,0029873	0,0009822	3,04	0,003
Outdoor air pollution attribu_1	-0,0000301	0,0005485	-0,05	0,956
Health expenditure, public (%_1	0,0009531	0,0004344	2,19	0,030
External resources for health_1	-0,0021719	0,0007533	-2,88	0,005

Fonte: Elaborado pela autora

S = 0,0620504 R-Sq = 86,5% R-Sq(adj) = 84,3%

Tabela 15. Análise de Variância para IDH como V.Dependente

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	18	2,78119	0,15451	40,13	0,000
Residual Error	113	0,43508	0,00385		
Total	131	3,21627			

Source	DF	Seq SS
EPI Score_1	1	2,18226
Happy Planet Index_1	1	0,09833
OHI_1	1	0,02986
GINNI Index_1	1	0,00042
Social Progress Index_1	1	0,15887
Maternal mortality rate (deat_1	1	0,13692
Stillbirth rate (deaths/1,000_1	1	0,03920
Child mortality rate (deaths/_1	1	0,00577
Deaths from infectious diseas_1	1	0,00199
Access to piped water (% of p_1	1	0,00948
Access to improved sanitation_1	1	0,02504
Indoor air pollution attribut_1	1	0,00150
Life expectancy (years)_1	1	0,00126
Non-communicable disease deat_1	1	0,00060
Obesity rate (% of pop.)_1	1	0,04489
Outdoor air pollution attribu_1	1	0,00027
Health expenditure, public (%_1	1	0,01252
External resources for health_1	1	0,03201

Fonte: Elaborado pela autora

Tabela 16. Observações Pouco Usuais para IDH como V.Dependente

Obs	EPI Score_1	IDH - 2013_1	Fit	SE Fit	Residual	St Resid
3	28,7	0,52600	0,46916	0,04205	0,05684	1,25 X
15	47,6	0,68300	0,60902	0,04337	0,07398	1,67 X
28	39,4	0,33800	0,47597	0,02168	-0,13797	-2,37R
75	40,1	0,41400	0,53339	0,03148	-0,11939	-2,23R
88	77,8	0,68940	0,86816	0,01478	-0,17876	-2,97R
113	37,4	0,89800	0,64574	0,03360	0,25226	4,84R
115	87,7	0,53000	0,83215	0,02145	-0,30215	-5,19R

Fonte: Elaborado pela autora

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Tabela 17 . Stepwise para IDH

Stepwise Regression: IDH - 2013_1 versus Social Progr; Maternal mor; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is IDH - 2013_1 on 19 predictors, with N = 132

Step	1	2	3	4	5
Constant	0,07990	0,14152	0,14350	0,25308	0,25204
Social Progress Index_1	0,00957	0,00612	0,00602	0,00527	0,00456
T-Value	19,89	8,86	9,13	7,56	5,96
P-Value	0,000	0,000	0,000	0,000	0,000
Access to improved sanitation_1		0,00215	0,00146	0,00099	0,00099
T-Value		6,32	3,90	2,45	2,50
P-Value		0,000	0,000	0,016	0,014
Obesity rate (% of pop.)_1			0,00313	0,00287	0,00278
T-Value			3,67	3,43	3,36
P-Value			0,000	0,001	0,001
Maternal mortality rate (deat_1				-0,00013	-0,00014
T-Value				-2,81	-3,05
P-Value				0,006	0,003
Health expenditure, public (%_1					0,00085
T-Value					2,10
P-Value					0,038
External resources for health_1					

T-Value					
P-Value					
S	0,0782	0,0686	0,0655	0,0638	0,0630
R-Sq	75,26	81,11	82,91	83,91	84,46
R-Sq(adj)	75,07	80,82	82,51	83,41	83,84
Mallows Cp	77,7	31,0	18,1	11,7	9,2

Fonte: Elaborado pela autora

Análise:

A primeira variável Social Progress explicou 75,26%, e a segunda Access to improved sanitation 81,11%, então, não há necessidade de gerar uma nova regressão, pois o R2 de ambos são bem altos. O grau de explicação do modelo 2 é satisfatório.

Tabela 17 . Continuação

Step	**6**
Constant	**0,25366**
Social Progress Index_1	**0,00474**
T-Value	**6,19**
P-Value	**0,000**
Access to improved sanitation_1	**0,00087**
T-Value	**2,18**
P-Value	**0,031**
Obesity rate (% of pop.)_1	**0,00245**
T-Value	**2,91**
P-Value	**0,004**
Maternal mortality rate (deat_1	**-0,00012**
T-Value	**-2,47**
P-Value	**0,015**
Health expenditure, public (%_1	**0,00101**
T-Value	**2,46**
P-Value	**0,015**
External resources for health_1	**-0,00112**
T-Value	**-1,78**
P-Value	**0,077**
S	**0,0625**
R-Sq	**84,84**
R-Sq(adj)	**84,11**
Mallows Cp	**8,0**

Fonte: Elaborado pela autora

Tabela 17 . Continuação

Step	**7**	**8**	**9**
Constant	**0,1866**	**0,2298**	**0,2178**
Social Progress Index_1	**0,00418**	**0,00363**	**0,00366**
T-Value	**5,11**	**4,02**	**4,06**
P-Value	**0,000**	**0,000**	**0,000**
Access to improved sanitation_1	**0,00077**	**0,00067**	**0,00075**
T-Value	**1,94**	**1,66**	**1,89**
P-Value	**0,055**	**0,100**	**0,061**
Obesity rate (% of pop.)_1	**0,00262**	**0,00251**	**0,00249**
T-Value	**3,12**	**2,99**	**2,96**
P-Value	**0,002**	**0,003**	**0,004**
Maternal mortality rate (deat_1	**-0,00012**	**-0,00007**	
T-Value	**-2,56**	**-1,21**	
P-Value	**0,012**	**0,230**	
Health expenditure, public (%_1	**0,00103**	**0,00109**	**0,00110**
T-Value	**2,55**	**2,68**	**2,71**
P-Value	**0,012**	**0,008**	**0,008**
External resources for health_1	**-0,00143**	**-0,00143**	**-0,00160**

T-Value	-2,21	-2,23	-2,54
P-Value	0,029	0,027	0,012
OHI_1	0,00167	0,00184	0,00189
T-Value	1,80	1,98	2,03
P-Value	0,074	0,050	0,044
Child mortality rate (deaths/_1		-0,00060	-0,00090
T-Value		-1,46	-2,69
P-Value		0,148	0,008
S	0,0619	0,0616	0,0617
R-Sq	85,23	85,48	85,31
R-Sq(adj)	84,40	84,54	84,48
Mallows Cp	6,8	6,7	6,1

Fonte: Elaborado pela autora

3.9. ANOVA

Uma vez que se utilizou a Regressão Múltipla e o Stepwise para identificar quais variáveis com melhor poder explicativo, se buscará agrupar os países por regiões. Da classificação de 8 regiões (Quadro 3), nas quais os 132 países estavam distribuídos: AIBER, AVECO, CARLA, CWEAS, EMDEU, ESEAS, NSWUA E SUSAF. Será observado o que acontecerá caso se proceda a uma classificação para 6 regiões e 3 regiões e se proceda o teste da diferença entre as médias das regiões classificadas em: 6R e depois 3R. Para isto, o método estatístico ANOVA (*Analysis of Variance*) é o mais recomendado. Será analisada cada ANOVA referente a cada variável comparando os grupos. Relembra-se neste momento que a Região Ibero-Americana é a região de maior interesse para este estudo. A ANOVA pretende identificar o quanto as médias dos grupos são iguais.

Quadro 3. Oito Regiões

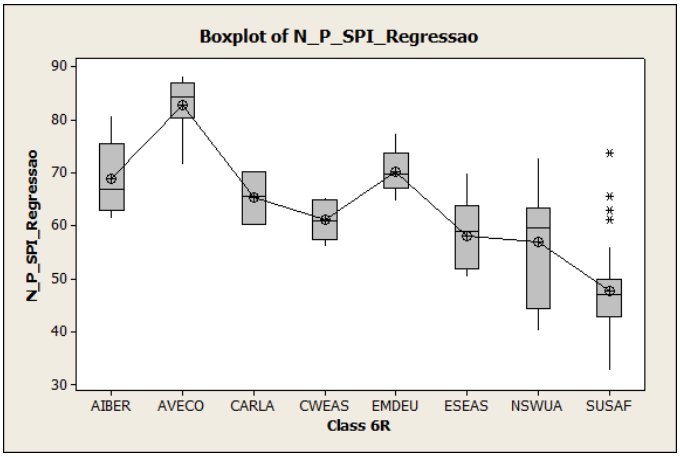
Nova Sigla	Nomenclatura da Região
AIBER	Ibero-Americanos
AVECO	Economias Avançadas
CARLA	América Latina e o Caribe
CWEAS	Comunidade dos Estados Independentes
ESEAS	Emergentes e em Desenvolvimento da Ásia
EMDEU	Emergentes e em Desenvolvimento da Europa
NSWUA	Oriente Médio, Norte da África e Paquistão
SUSAF	África Subsariana

Fonte: Elaborado pela autora

Regiões 6R

O gráfico abaixo apresenta uma interpretação visual das médias dos grupos. Ele sugere que há diferença entre os grupos.

Gráfico 52.



Fonte: Elaborado pela autora

A seguir, será realizada a análise para confirmar tal fato.

One-way ANOVA: N_P_SPI_Regressao versus Class 6R

```
Source     DF       SS      MS      F      P
Class 6R    7  20270,6  2895,8  58,42  0,000
Error     124   6146,9    49,6
Total     131  26417,5

S = 7,041   R-Sq = 76,73%   R-Sq(adj) = 75,42%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean   StDev  --+---------+---------+---------+-------
AIBER  20  68,804   6,599                  (-*--)
AVECO  27  82,843   4,838                            (-*-)
CARLA   4  65,350   5,545            (----*-----)
CWEAS  11  61,125   3,408          (---*--)
EMDEU  12  70,068   3,918                   (--*---)
ESEAS  12  57,958   6,520         (--*---)
NSWUA  14  56,945  10,426       (--*---)
SUSAF  32  47,643   8,895  (-*-)
                          --+---------+---------+---------+-------
                           48        60        72        84

Pooled StDev = 7,041
```

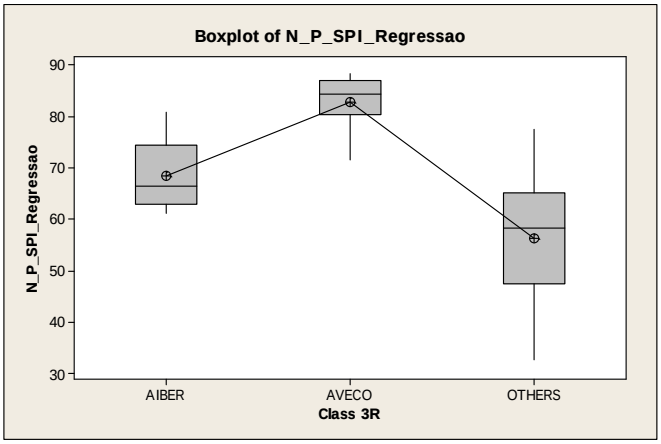
Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 58,42 (p-valor = 0,000). A análise gráfica e a análise por intervalo de confiança sugerem que a região que possui o maior índice é a AVECO e a menor é a SUSAF. Porém, é necessária uma análise mais específica para comparar tal fato. Percebe-se que há 4 outliers no grupo SUSAF.

Regiões 3R
AIBER, AVECO e OTHERS

O gráfico abaixo novamente apresenta uma interpretação visual das médias dos grupos. Ele sugere que há diferença entre os grupos.

Gráfico 53.



Fonte: Elaborado pela autora

A seguir, será realizada a análise para confirmar tal fato.

One-way ANOVA: N_P_SPI_Regressao versus Class 3R

```
Source     DF       SS      MS      F      P
Class 3R    2  14940,9  7470,4  83,97  0,000
Error     129  11476,6    89,0
Total     131  26417,5

S = 9,432   R-Sq = 56,56%   R-Sq(adj) = 55,88%

                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level    N    Mean   StDev  ------+---------+---------+---------+---
AIBER   21  68,436   6,650           (---*----)
AVECO   27  82,843   4,838                             (---*--)
OTHERS  84  56,319  10,967  (-*-)
                            ------+---------+---------+---------+---
                                 60        70        80        90

Pooled StDev = 9,432
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 83,97 (p-valor = 0,000).

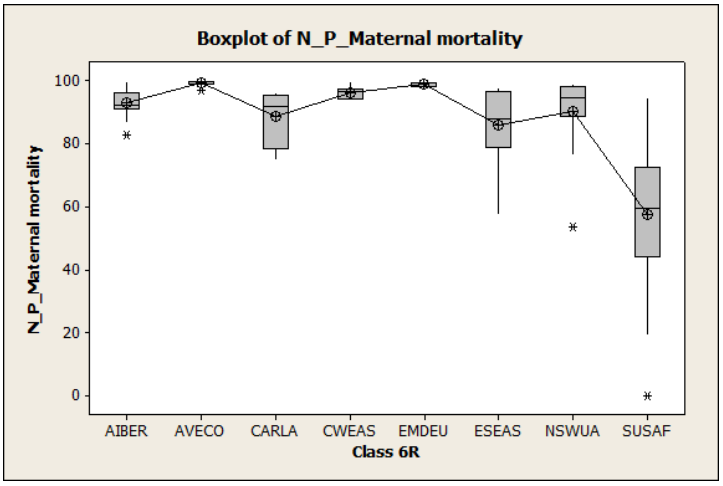
A análise gráfica e a análise por intervalo de confiança sugerem que a região que possui o maior índice é a AVECO e a menor é a OTHERS. Porém, é necessária uma análise mais específica para comparar tal fato.

Variável N_P_Maternal mortality

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 54.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Maternal mortality versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  35172  5025  34,76  0,000
Error     124  17922   145
Total     131  53094

S = 12,02   R-Sq = 66,24%   R-Sq(adj) = 64,34%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
AIBER  20  92,99   4,20                       (---*---)
AVECO  27  99,37   0,61                           (--*--)
CARLA   4  88,59   9,58                 (-------*-------)
CWEAS  11  96,32   1,83                         (----*----)
EMDEU  12  98,86   0,70                           (----*---)
ESEAS  12  86,06  11,94                    (---*----)
NSWUA  14  90,35  12,41                       (---*---)
SUSAF  32  57,52  21,02  (-*--)
                          ----+---------+---------+---------+-----
                             60        75        90       105

Pooled StDev = 12,02
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 34,76 (p-valor = 0,000).

Maior índice: Não definido.

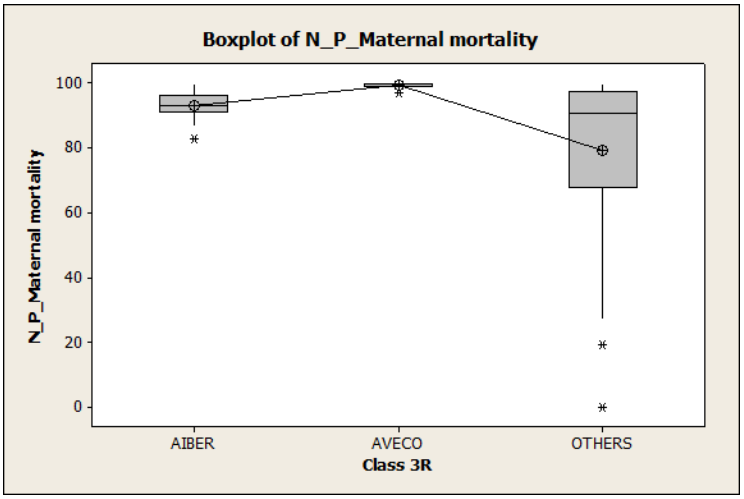
Menor índice: SUSAF.

Há outliers no AIBER, AVECO, NSWUA e SUSAF.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 55.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Maternal mortality versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2   9813  4907  14,62  0,000
Error     129  43281   336
Total     131  53094

S = 18,32   R-Sq = 18,48%   R-Sq(adj) = 17,22%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  93,02   4,10              (---------*---------)
AVECO   27  99,37   0,61                        (--------*--------)
OTHERS  84  79,11  22,74  (----*----)
                          ------+---------+---------+---------+---
                             80,0      88,0      96,0     104,0

Pooled StDev = 18,32
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 14,62 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

A análise gráfica e a análise por intervalo de confiança sugerem que a região que possui o melhor índice é a AVECO e a pior é a OTHERS.

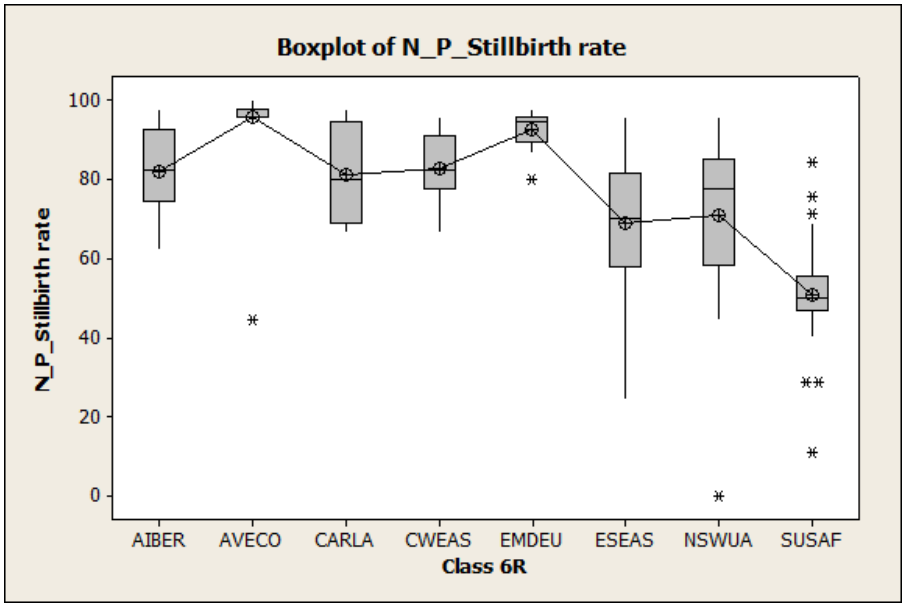
Há outliers nas 3 regiões.

Variável N_P_Stillbirth rate

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 56.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Stillbirth rate versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  36372  5196  26,35  0,000
Error     124  24448   197
Total     131  60820

S = 14,04   R-Sq = 59,80%   R-Sq(adj) = 57,53%
```

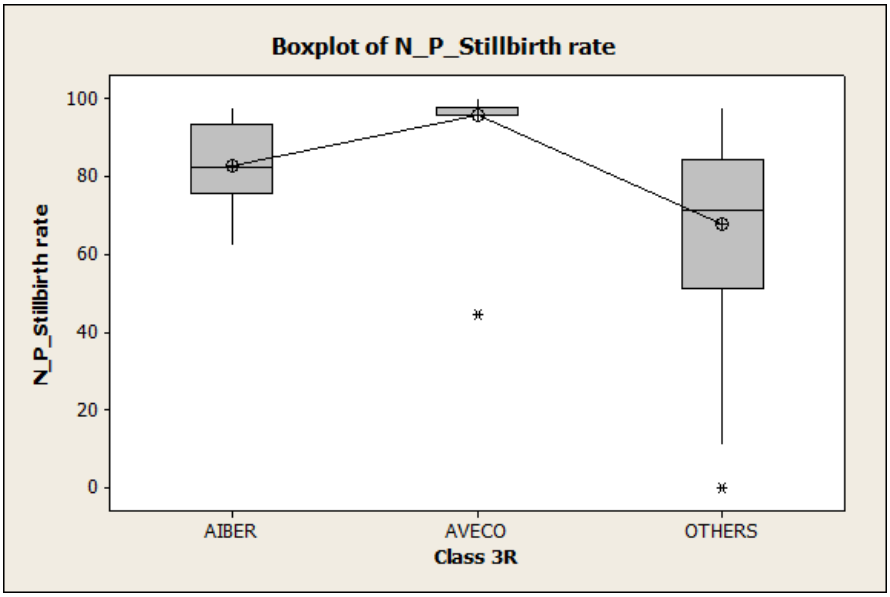
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---------+---------+---------+---------+
AIBER  20  82,00  10,52                     (---*---)
AVECO  27  95,64  10,34                               (---*--)
CARLA   4  81,11  13,27               (--------*--------)
CWEAS  11  82,63   8,99                   (-----*-----)
EMDEU  12  92,59   5,22                         (-----*----)
ESEAS  12  68,89  18,95           (----*----)
NSWUA  14  70,79  25,05            (----*----)
SUSAF  32  50,77  13,70  (--*--)
                          ---------+---------+---------+---------+
                                  60        75        90       105

Pooled StDev = 14,04
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 26,35 (p-valor = 0,000).
Maior índice: AVECO e EMDEU.
Menor índice: SUSAF.
Não há outliers na AIBER, CARLA e CWEAS.

Regiões 3R
O gráfico sugere que há diferença entre os grupos.

Gráfico 57.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Stillbirth rate versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  17113  8557  25,25  0,000
Error     129  43707   339
Total     131  60820

S = 18,41   R-Sq = 28,14%   R-Sq(adj) = 27,02%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  82,75  10,82            (-------*-------)
AVECO   27  95,64  10,34                      (------*------)
OTHERS  84  67,73  21,56  (---*---)
                          ------+---------+---------+---------+---
                               70        80        90       100

Pooled StDev = 18,41
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 25,25 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

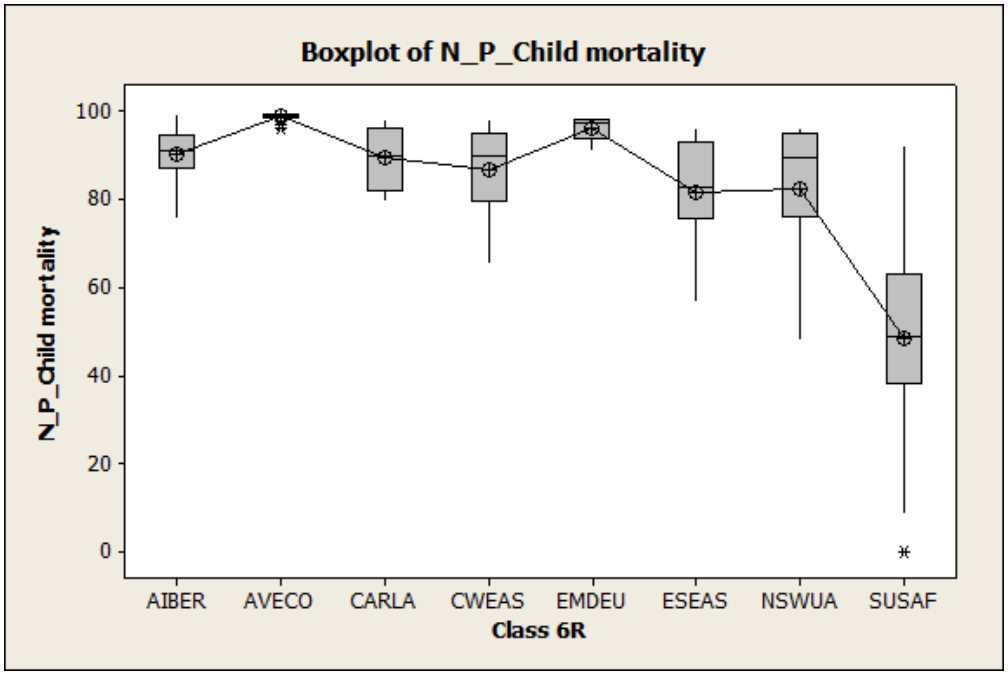
Não há outlier na AIBER.

Variável N_P_Child mortality

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 58.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Child mortality versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  47826  6832  43,94  0,000
Error     124  19283   156
Total     131  67109

S = 12,47   R-Sq = 71,27%   R-Sq(adj) = 69,64%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  90,33   5,74                              (--*---)
AVECO  27  98,73   0,91                                    (--*--)
CARLA   4  89,30   7,61                          (--------*-------)
CWEAS  11  86,85   9,63                            (----*----)
EMDEU  12  96,10   2,53                                  (----*----)
ESEAS  12  81,58  12,07                          (---*----)
NSWUA  14  82,53  16,61                           (---*---)
SUSAF  32  48,30  19,90  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 12,47
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 43,94 (p-valor = 0,000).

Maior índice: Não definido.

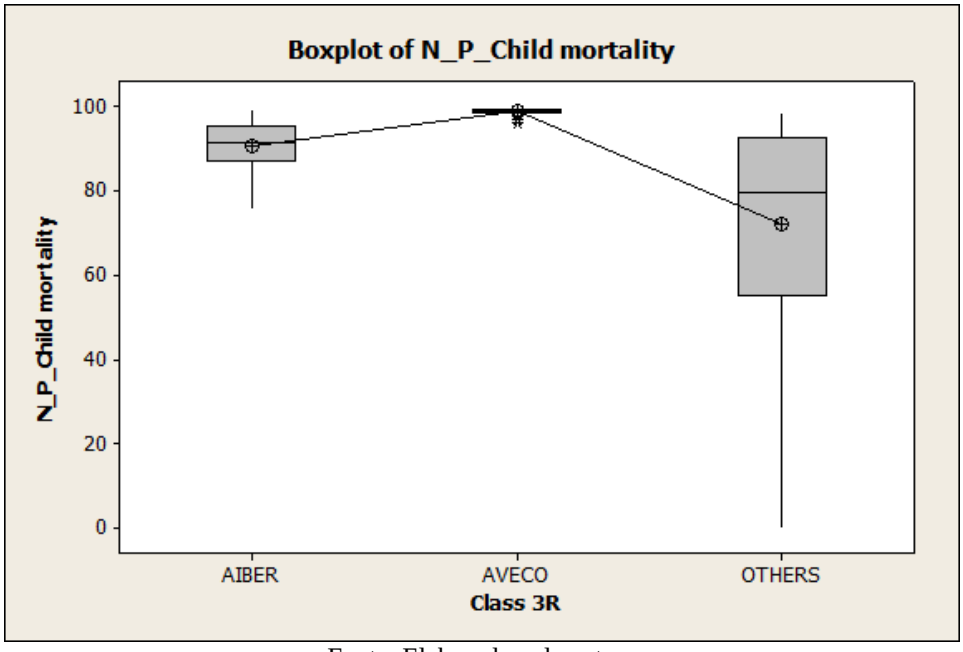
Menor índice: SUSAF.

Há outliers na AVECO e na SUSAF.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 59.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Child mortality versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  17229  8614  22,28  0,000
Error     129  49880   387
Total     131  67109

S = 19,66   R-Sq = 25,67%   R-Sq(adj) = 24,52%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --+---------+---------+---------+-------
AIBER   21  90,70   5,84               (--------*-------)
AVECO   27  98,73   0,91                         (-------*------)
OTHERS  84  72,00  24,34  (---*---)
                          --+---------+---------+---------+-------
                           70        80        90       100

Pooled StDev = 19,66
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 22,28 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

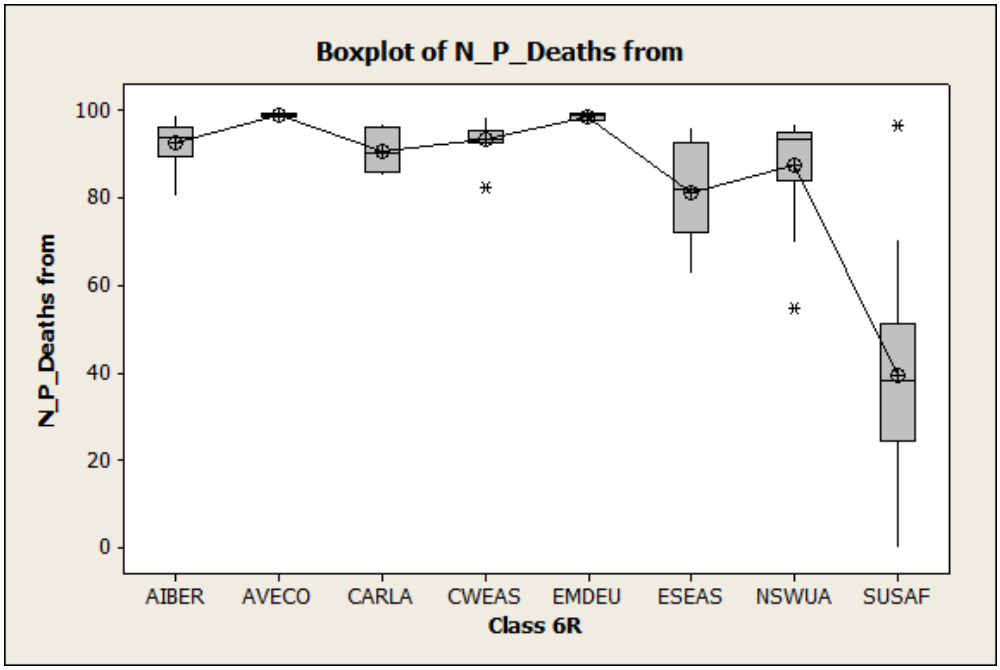
Há outlier na AVECO.

Variável N_P_Deaths from

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 60.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Deaths from versus Class 6R

```
Source     DF     SS     MS      F      P
Class 6R    7  73298  10471  77,85  0,000
Error     124  16679    135
Total     131  89977

S = 11,60   R-Sq = 81,46%   R-Sq(adj) = 80,42%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
AIBER  20  92,68   4,95                          (-*--)
AVECO  27  98,90   0,60                             (-*--)
CARLA   4  90,76   5,18                        (----*-----)
CWEAS  11  93,27   4,05                           (---*--)
EMDEU  12  98,57   0,97                              (--*---)
ESEAS  12  81,20  10,92                      (---*--)
NSWUA  14  87,66  12,02                          (--*--)
SUSAF  32  39,28  20,29  (-*-)
                          --+---------+---------+---------+-------
                           40        60        80       100

Pooled StDev = 11,60
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 77,85 (p-valor = 0,000).

Maior índice: Não definido.

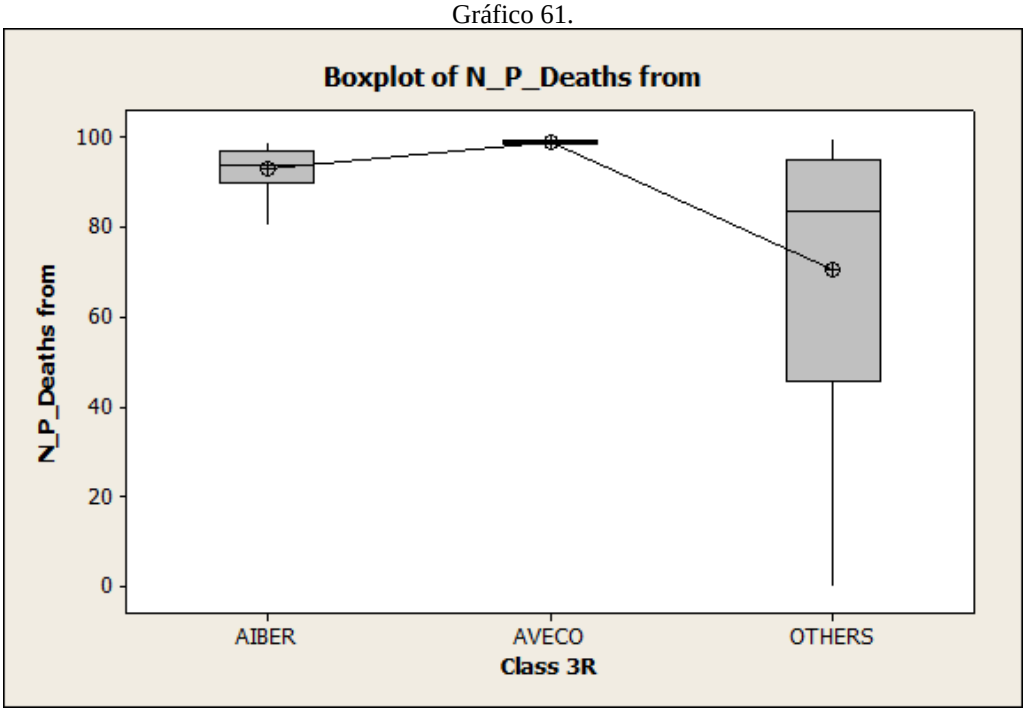
Menor índice: SUSAF.

Há outliers na CWEAS, NSWUA e na SUSAF

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 61.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Deaths from versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  20501  10251  19,03  0,000
Error     129  69476    539
Total     131  89977

S = 23,21   R-Sq = 22,78%   R-Sq(adj) = 21,59%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  92,89   4,92               (-------*--------)
AVECO   27  98,90   0,60                     (------*-------)
OTHERS  84  70,63  28,83  (---*---)
                          -----+---------+---------+---------+----
                              72        84        96       108

Pooled StDev = 23,21
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 19,03 (p-valor = 0,000).

Maior índice: AVECO.

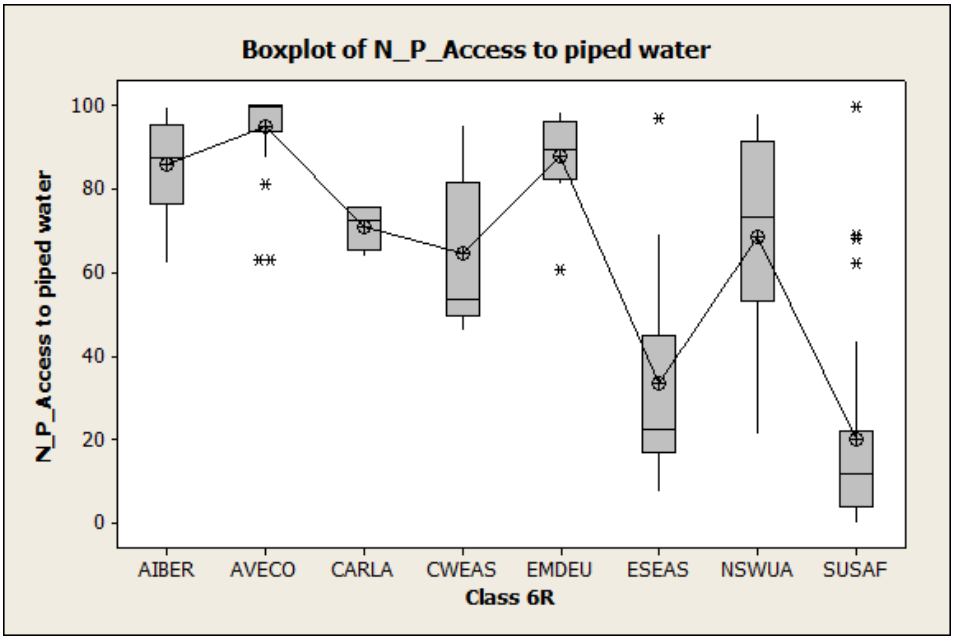
Menor índice: OTHERS.

Variável N_P_Access to piped water

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 62.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Access to piped water versus Class 6R

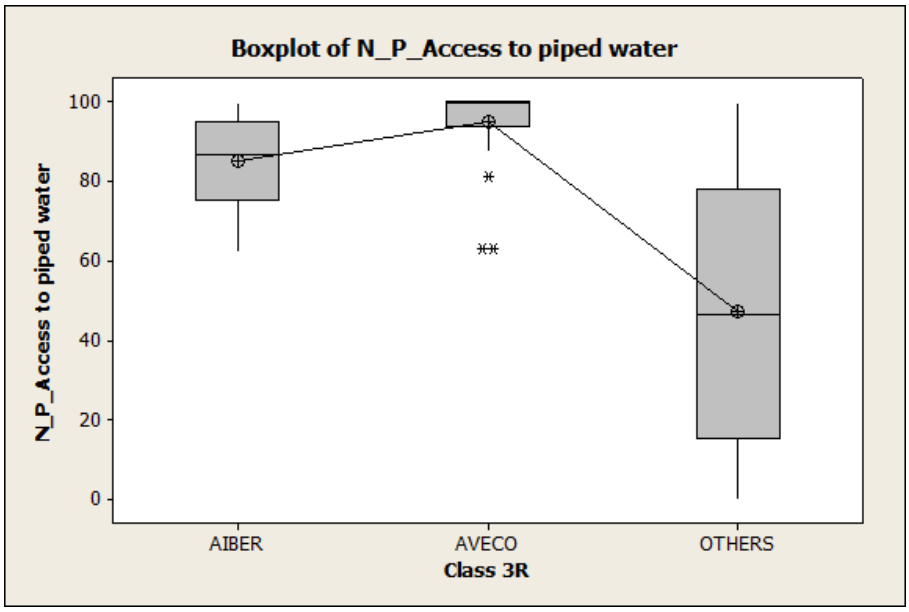
```
Source     DF      SS     MS      F      P
Class 6R    7  115589  16513  47,04  0,000
Error     124   43525    351
Total     131  159113

S = 18,74   R-Sq = 72,65%   R-Sq(adj) = 71,10%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -----+---------+---------+---------+----
AIBER  20  85,70  11,90                          (--*---)
AVECO  27  94,89  10,29                             (--*--)
CARLA   4  71,05   5,54                 (------*-------)
CWEAS  11  64,69  18,52                 (----*---)
EMDEU  12  87,72  10,31                          (---*---)
ESEAS  12  33,41  26,28      (---*----)
NSWUA  14  68,43  23,81                   (---*---)
SUSAF  32  19,96  24,38  (--*--)
                          -----+---------+---------+---------+----
                              25        50        75       100

Pooled StDev = 18,74
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 47,04 (p-valor = 0,000).
Maior índice: AVECO.
Menor índice: SUSAF.
Não há outliers na AIBER e na CWEAS.

Regiões 3R
O gráfico sugere que há diferença entre os grupos.

Gráfico 63.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Access to piped water versus Class 3R

```
Source     DF      SS     MS      F      P
Class 3R    2   58626  29313  37,63  0,000
Error     129  100488    779
Total     131  159113

S = 27,91   R-Sq = 36,85%   R-Sq(adj) = 35,87%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  85,22  11,81                  (-----*-----)
AVECO   27  94,89  10,29                       (----*-----)
OTHERS  84  47,27  33,82  (--*--)
                          ---------+---------+---------+---------+
                                  60        80       100       120

Pooled StDev = 27,91
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 37,63 (p-valor = 0,000).

Maior índice: Não definido.

Menor índice: OTHERS.

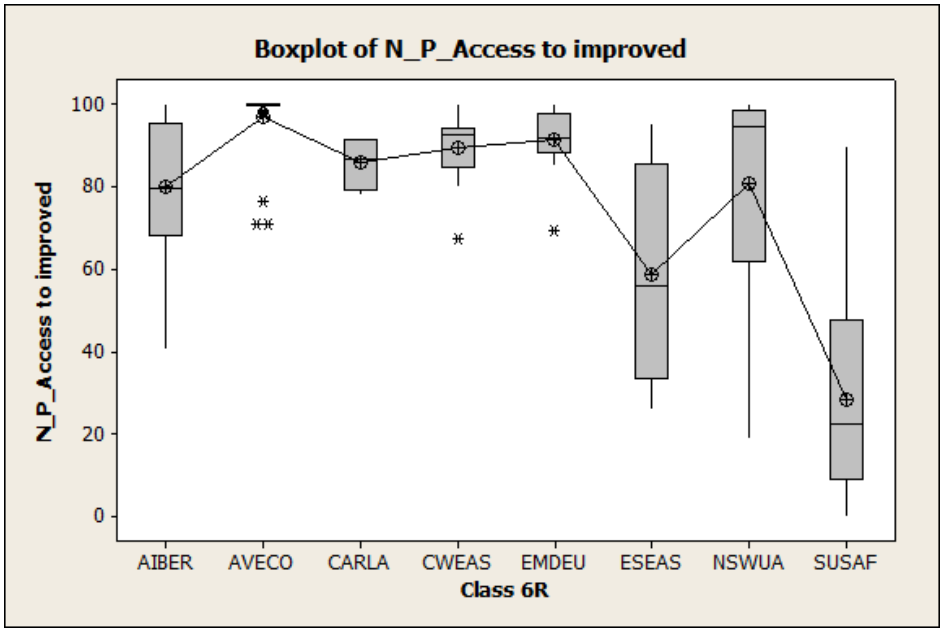
Há outliers na AVECO.

Variável N_P_Access to improved

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 64.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Access to improved versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7   90985  12998  37,87  0,000
Error     124   42562    343
Total     131  133547

S = 18,53   R-Sq = 68,13%   R-Sq(adj) = 66,33%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  80,10  16,92                    (--*--)
AVECO  27  96,75   8,71                           (--*--)
CARLA   4  85,69   6,64                  (------*-------)
CWEAS  11  89,43   9,16                      (----*---)
EMDEU  12  91,30   8,49                       (----*---)
ESEAS  12  58,62  24,81           (---*----)
NSWUA  14  80,57  26,35                     (---*---)
SUSAF  32  28,15  23,82  (-*--)
                          -+---------+---------+---------+--------
                          25        50        75       100

Pooled StDev = 18,53
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 37,87 (p-valor = 0,000).

Maior índice: Não definido.

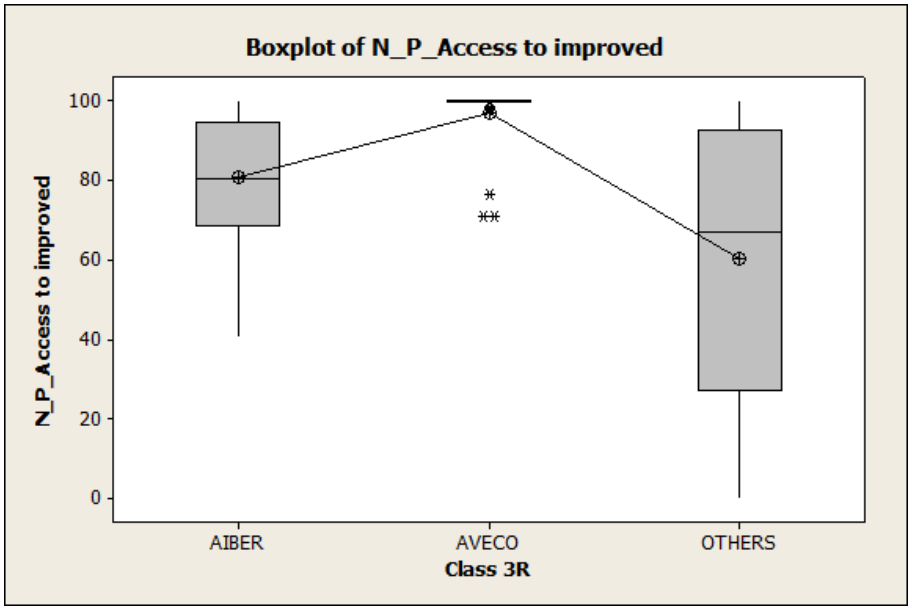
Menor índice: SUSAF.

Há outliers na AVECO, CWEAS e na EMDEU.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 65.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Access to improved versus Class 3R

```
Source     DF      SS     MS      F      P
Class 3R    2   29519  14759  18,30  0,000
Error     129  104028    806
Total     131  133547

S = 28,40   R-Sq = 22,10%   R-Sq(adj) = 20,90%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  80,63  16,67           (-------*-------)
AVECO   27  96,75   8,71                     (-------*------)
OTHERS  84  60,27  34,10  (---*---)
                          ----+---------+---------+---------+-----
                             60        75        90       105

Pooled StDev = 28,40
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 18,30 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

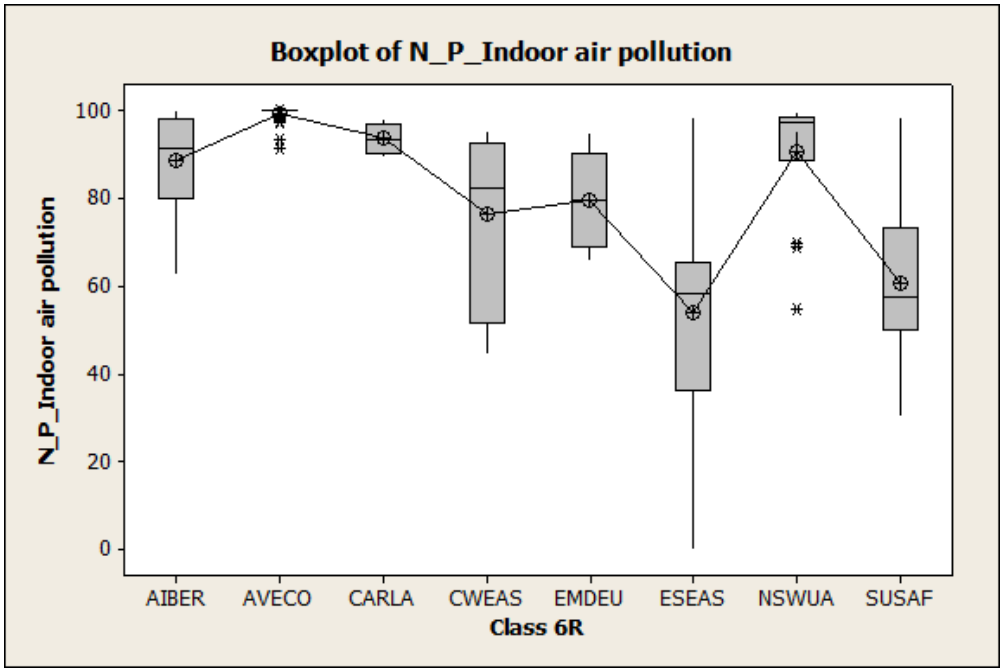
Há outliers na AVECO.

Variável N_P_Indoor air pollution

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 66.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Indoor air pollution versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  34113  4873  22,81  0,000
Error     124  26488   214
Total     131  60601

S = 14,62   R-Sq = 56,29%   R-Sq(adj) = 53,82%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
AIBER  20  88,47  11,09                      (---*---)
AVECO  27  99,21   2,12                              (--*--)
CARLA   4  93,60   3,58                     (---------*---------)
CWEAS  11  76,58  18,96               (-----*----)
EMDEU  12  79,49  10,35                 (-----*----)
ESEAS  12  53,92  25,93  (-----*----)
NSWUA  14  90,52  14,51                        (----*---)
SUSAF  32  60,48  17,12        (--*--)
                          --+---------+---------+---------+-------
                           48        64        80        96

Pooled StDev = 14,62
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 22,81 (p-valor = 0,000).

Maior índice: Não definido.

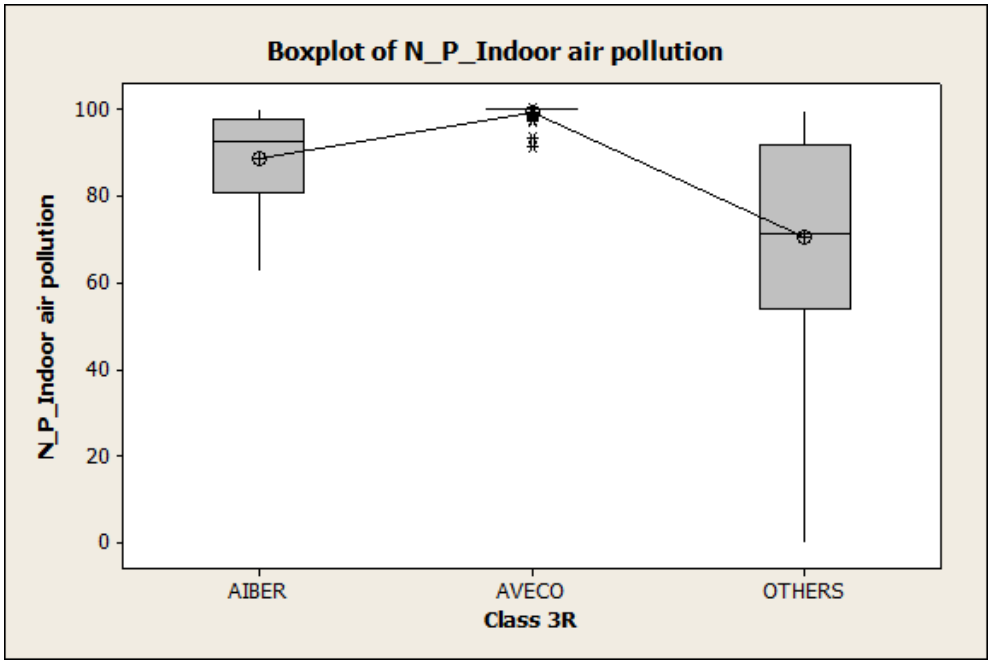
Menor índice: Não definido.

Há outliers na AVECO e na NSWUA.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 67.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Indoor air pollution versus Class 3R

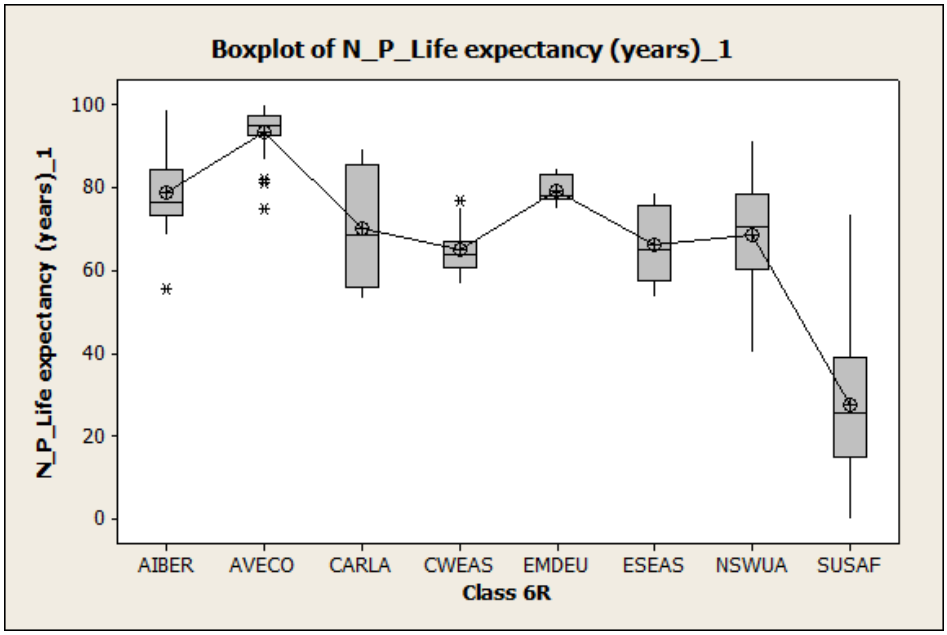
```
Source     DF     SS    MS      F      P
Class 3R    2  18982  9491  29,42  0,000
Error     129  41619   323
Total     131  60601

S = 17,96   R-Sq = 31,32%   R-Sq(adj) = 30,26%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  88,71  10,87               (-------*------)
AVECO   27  99,21   2,12                         (------*------)
OTHERS  84  70,56  21,72  (---*--)
                          ---+---------+---------+---------+------
                            70        80        90       100

Pooled StDev = 17,96
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 29,42 (p-valor = 0,000).
Maior índice: AVECO.
Menor índice: OTHERS.
Há outliers na AVECO.

Variável N_P_Life expectancy (years)_1
Regiões 6R
O gráfico sugere que há diferença entre os grupos.

Gráfico 68.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Life expectancy (years)_1 versus Class 6R

```
Source     DF     SS     MS      F      P
Class 6R    7  73311  10473  82,61  0,000
Error     124  15720    127
Total     131  89031

S = 11,26   R-Sq = 82,34%   R-Sq(adj) = 81,35%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  20  78,94  10,14                        (-*--)
AVECO  27  93,43   6,01                                  (-*-)
CARLA   4  70,01  15,32                 (-----*-----)
CWEAS  11  65,10   6,28                 (---*--)
EMDEU  12  79,19   3,28                        (---*--)
ESEAS  12  65,99   9,08                  (--*--)
NSWUA  14  68,47  14,45                    (--*--)
SUSAF  32  27,51  16,05  (-*-)
                          --------+---------+---------+---------+-
                                 40        60        80       100

Pooled StDev = 11,26
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 82,61 (p-valor = 0,000).

Maior índice: AVECO.

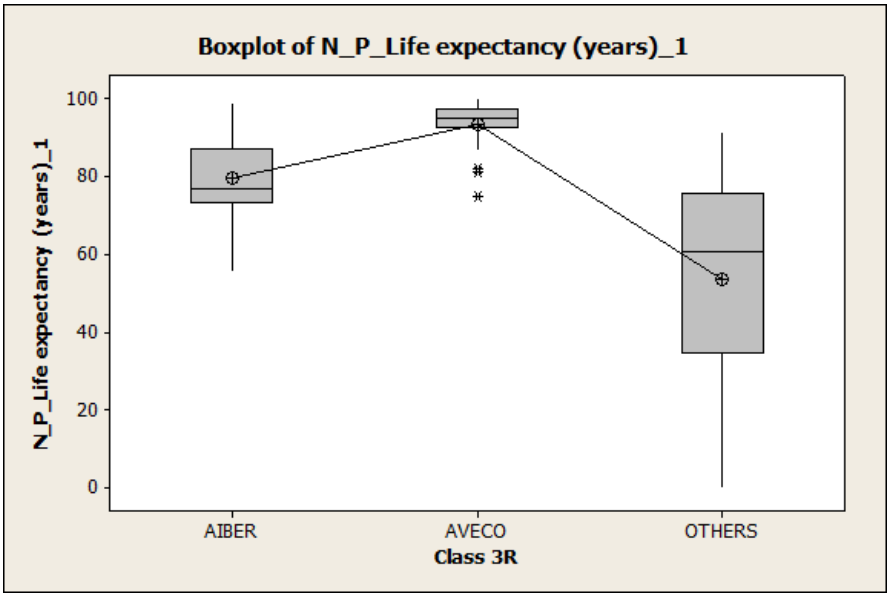
Menor índice: SUSAF.

Há outliers na AVECO, AIBER e CWEAS.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 69.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P_Life expectancy (years)_1 versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  37374  18687  46,67  0,000
Error     129  51656    400
Total     131  89031

S = 20,01   R-Sq = 41,98%   R-Sq(adj) = 41,08%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  79,44  10,14                (-----*-----)
AVECO   27  93,43   6,01                          (----*----)
OTHERS  84  53,43  24,21  (--*-)
                          -------+---------+---------+---------+--
                                60        75        90       105

Pooled StDev = 20,01
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 46,67 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

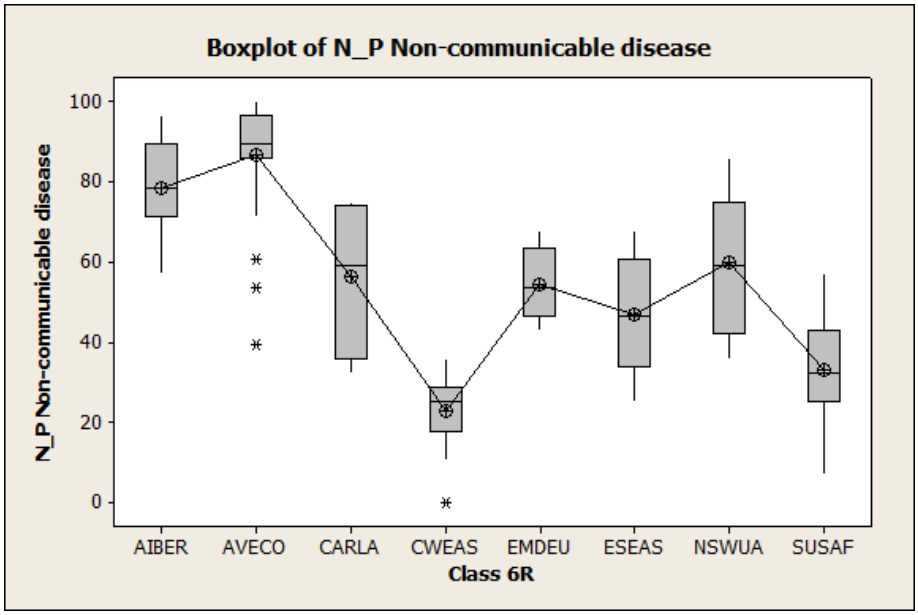
Há outliers na AVECO.

Variável N_P_Non-communicable disease

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 70.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Non-communicable disease versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  65979  9426  51,05  0,000
Error     124  22895   185
Total     131  88874

S = 13,59   R-Sq = 74,24%   R-Sq(adj) = 72,78%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  20  78,57  11,23                                (--*--)
AVECO  27  86,77  14,52                                     (-*--)
CARLA   4  56,25  20,49                 (------*------)
CWEAS  11  22,73  10,51  (---*---)
EMDEU  12  54,46   8,36                    (---*---)
ESEAS  12  46,73  14,96                 (---*---)
NSWUA  14  59,69  18,18                       (---*--)
SUSAF  32  33,04  12,86          (--*-)
                          ---+---------+---------+---------+------
                            20        40        60        80

Pooled StDev = 13,59
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 51,05 (p-valor = 0,000).

Maior índice: AVECO.

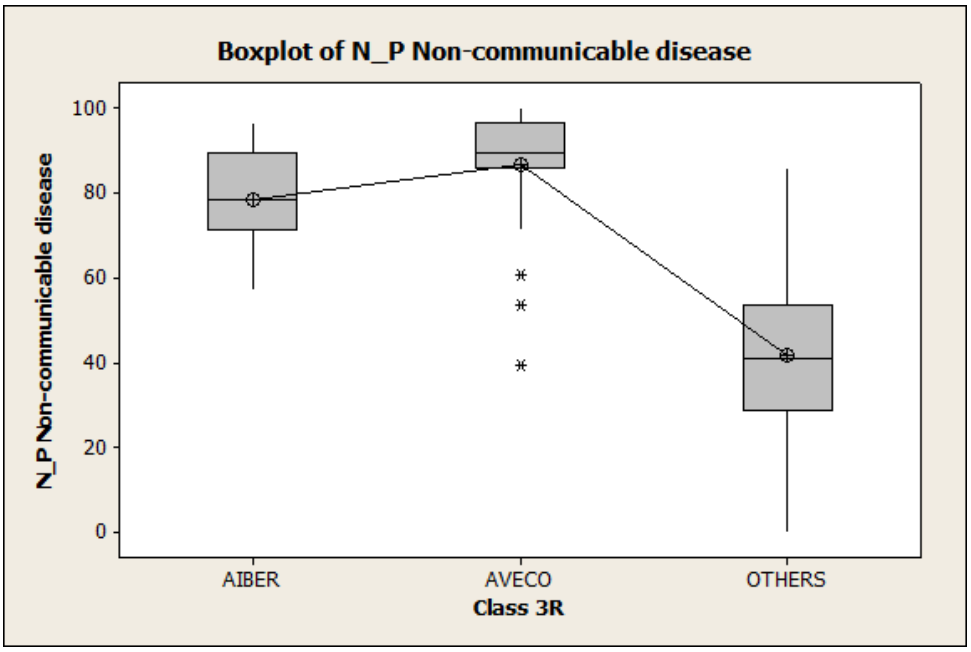
Menor índice: CWEAS.

Há outliers na AVECO e na CWEAS.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 71.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Non-communicable disease versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  53076  26538  95,63  0,000
Error     129  35798    278
Total     131  88874

S = 16,66   R-Sq = 59,72%   R-Sq(adj) = 59,10%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  78,40  10,98                       (----*----)
AVECO   27  86,77  14,52                              (---*---)
OTHERS  84  41,75  18,34  (--*-)
                          -----+---------+---------+---------+----
                              45        60        75        90

Pooled StDev = 16,66
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 95,63 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

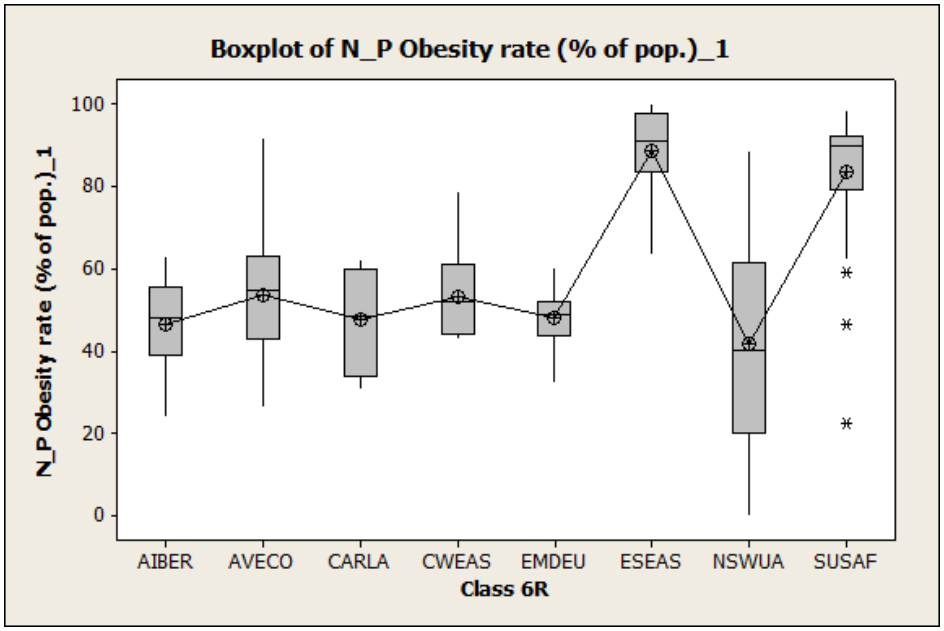
Há outliers na AVECO.

Variável N_P Obesity rate (% of pop.)_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 72.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Obesity rate (% of pop.)_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  39502  5643  25,27  0,000
Error     124  27692   223
Total     131  67194

S = 14,94   R-Sq = 58,79%   R-Sq(adj) = 56,46%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ----+---------+---------+---------+-----
AIBER  20  46,34  10,85       (--*--)
AVECO  27  53,73  14,61           (--*--)
CARLA   4  47,48  13,50   (-------*------)
CWEAS  11  53,17  10,80         (----*---)
EMDEU  12  48,12   6,95       (---*---)
ESEAS  12  88,55  11,81                            (---*----)
NSWUA  14  41,83  24,76    (---*---)
SUSAF  32  83,45  16,22                           (--*-)
                          ----+---------+---------+---------+-----
                             40        60        80       100

Pooled StDev = 14,94
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 25,27 (p-valor = 0,000).
Maior índice: ESEAS e a SUSAF.
Menor índice: Não definido.
Há outliers na SUSAF.

Regiões 3R
O gráfico sugere que há diferença entre os grupos.

Gráfico 73.

Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Obesity rate (% of pop.)_1 versus Class 3R

```
Source     DF     SS    MS     F      P
Class 3R    2   8629  4315  9,50  0,000
Error     129  58565   454
Total     131  67194

S = 21,31   R-Sq = 12,84%   R-Sq(adj) = 11,49%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  46,68  10,68  (---------*---------)
AVECO   27  53,73  14,61           (-------*-------)
OTHERS  84  66,88  24,72                        (----*---)
                          ---+---------+---------+---------+------
                            40        50        60        70

Pooled StDev = 21,31
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 9,5 (p-valor = 0,000).

Maior índice: OTHERS.

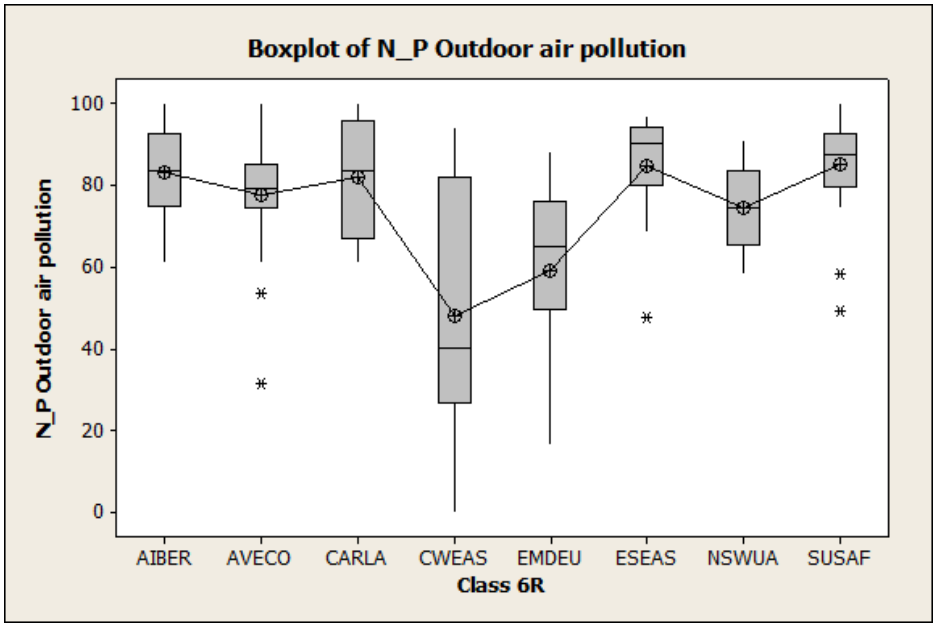
Menor índice: Não definido.

Variável N_P Outdoor air pollution

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 74.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Outdoor air pollution versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  16808  2401  9,93  0,000
Error     124  29978   242
Total     131  46786

S = 15,55   R-Sq = 35,93%   R-Sq(adj) = 32,31%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
AIBER  20  83,13  10,58                          (---*----)
AVECO  27  77,50  13,38                      (---*---)
CARLA   4  82,09  15,94                  (----------*---------)
CWEAS  11  47,90  31,64  (-----*-----)
EMDEU  12  59,02  21,79        (-----*-----)
ESEAS  12  84,58  14,28                           (-----*-----)
NSWUA  14  74,63  10,51                   (-----*----)
SUSAF  32  84,93  10,66                             (---*--)
                          ----+---------+---------+---------+-----
                             45        60        75        90

Pooled StDev = 15,55
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 9,93 (p-valor = 0,000).

Maior índice: CWEAS.

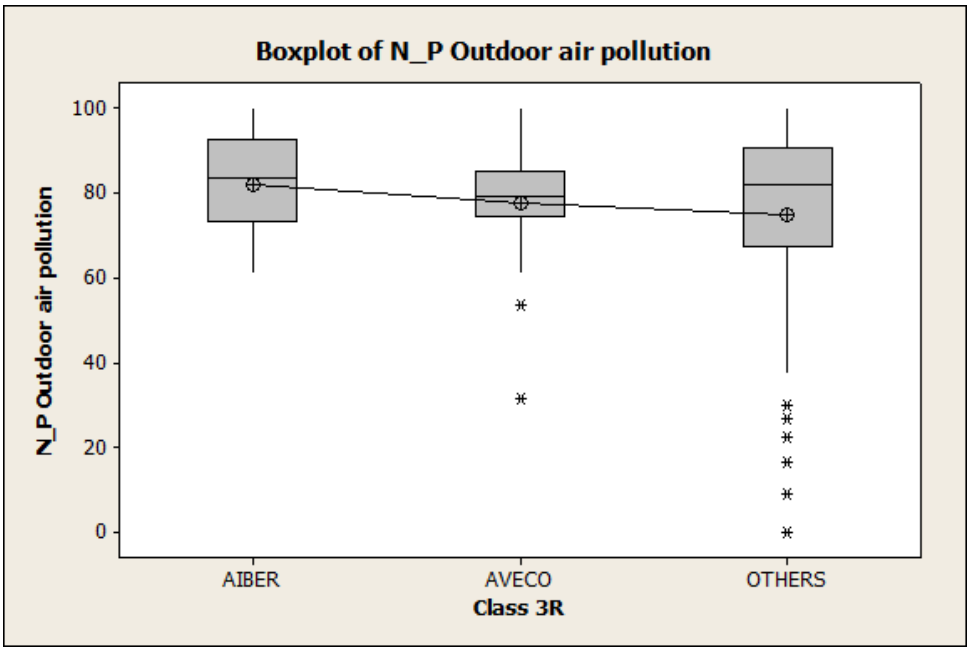
Menor índice: Não definido.

Há outliers na AVECO, ESEAS e SUSAF.

Regiões 3R

O gráfico sugere que não há diferença entre os grupos.

Gráfico 75.



Fonte: Elaborado pela autora

One-way ANOVA: N_P Outdoor air pollution versus Class 3R

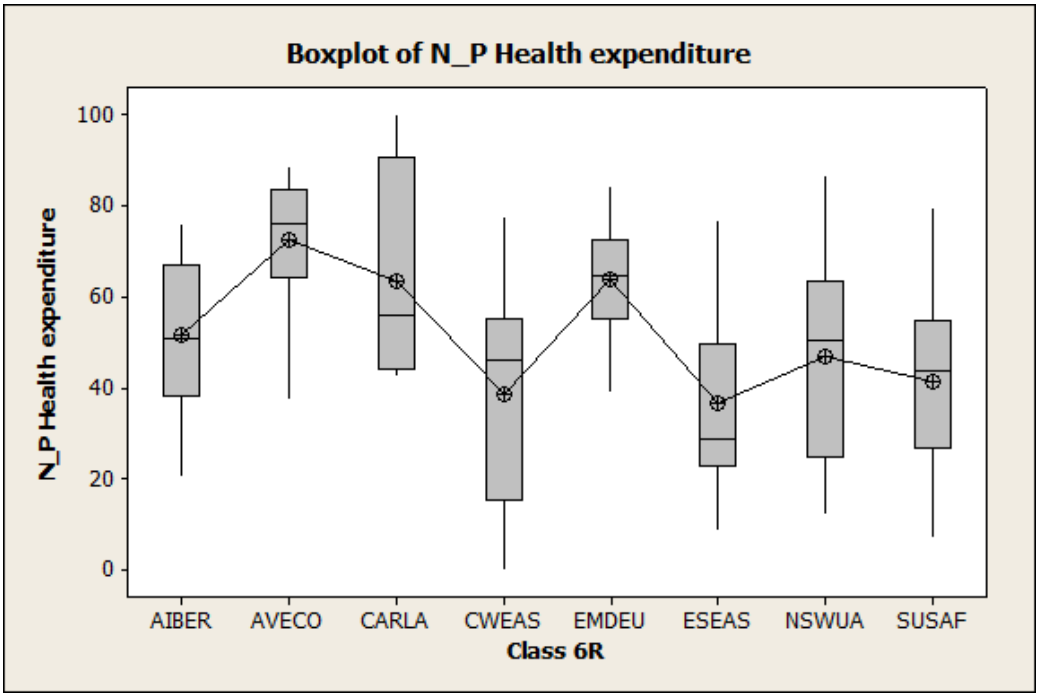
```
Source     DF     SS   MS     F      P
Class 3R    2    937  469  1,32  0,271
Error     129  45848  355
Total     131  46786

S = 18,85   R-Sq = 2,00%   R-Sq(adj) = 0,48%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  82,09  11,37         (---------------*---------------)
AVECO   27  77,50  13,38  (-------------*-------------)
OTHERS  84  74,76  21,57  (--------*-------)
                          ---------+---------+---------+---------+
                                75,0      80,0      85,0      90,0

Pooled StDev = 18,85
```

Análise:

A ANOVA realizada mostrou que não há diferença significativa entre os grupos, F = 1,32 (p-valor = 0,271). Ou seja, as medias foram consideradas estatisticamente iguais.
Não há outliers na AIBER

Variável N_P Health expenditure
Regiões 6R
O gráfico sugere que há diferença entre os grupos.

Gráfico 76.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Health expenditure versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  22492  3213  9,15  0,000
Error     124  43525   351
Total     131  66017

S = 18,74   R-Sq = 34,07%   R-Sq(adj) = 30,35%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  20  51,41  17,23             (----*-----)
AVECO  27  72,52  13,50                       (---*----)
CARLA   4  63,51  25,82               (-----------*------------)
CWEAS  11  38,65  23,68   (-------*------)
EMDEU  12  63,74  12,73                    (------*-------)
ESEAS  12  36,57  19,23  (------*-------)
NSWUA  14  46,65  24,61         (------*------)
SUSAF  32  41,22  19,52       (---*----)
                          ---+---------+---------+---------+------
                            30        45        60        75

Pooled StDev = 18,74
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 9,15 (p-valor = 0,000).

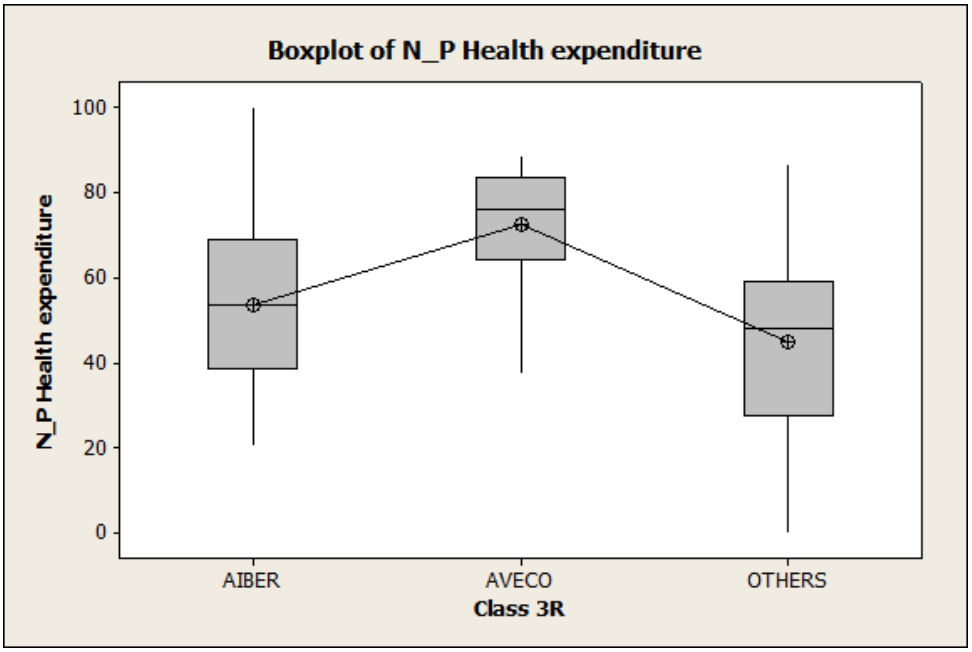
Maior índice: AVECO mas parece não definido.

Menor índice: ESEAS, mas parece não definido.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 77.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Health expenditure versus Class 3R

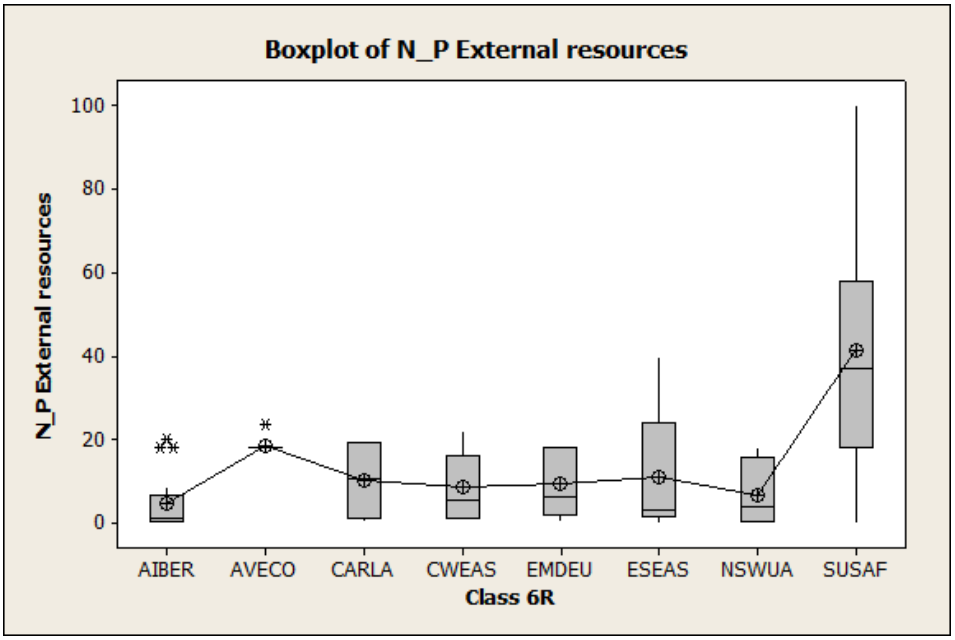
```
Source     DF     SS    MS      F      P
Class 3R    2  15901  7951  20,47  0,000
Error     129  50116   388
Total     131  66017

S = 19,71   R-Sq = 24,09%   R-Sq(adj) = 22,91%

                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  53,72  19,86        (--------*-------)
AVECO   27  72,52  13,50                          (-------*------)
OTHERS  84  44,70  21,25   (----*---)
                           +---------+---------+---------+---------
                          40        50        60        70

Pooled StDev = 19,71
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 20,47 (p-valor = 0,000).
Maior índice: AVECO.
Menor índice: OTHERS.

Variável N_P External resources
Regiões 6R
O gráfico sugere que há diferença entre os grupos.

Gráfico 78.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P External resources versus Class 6R

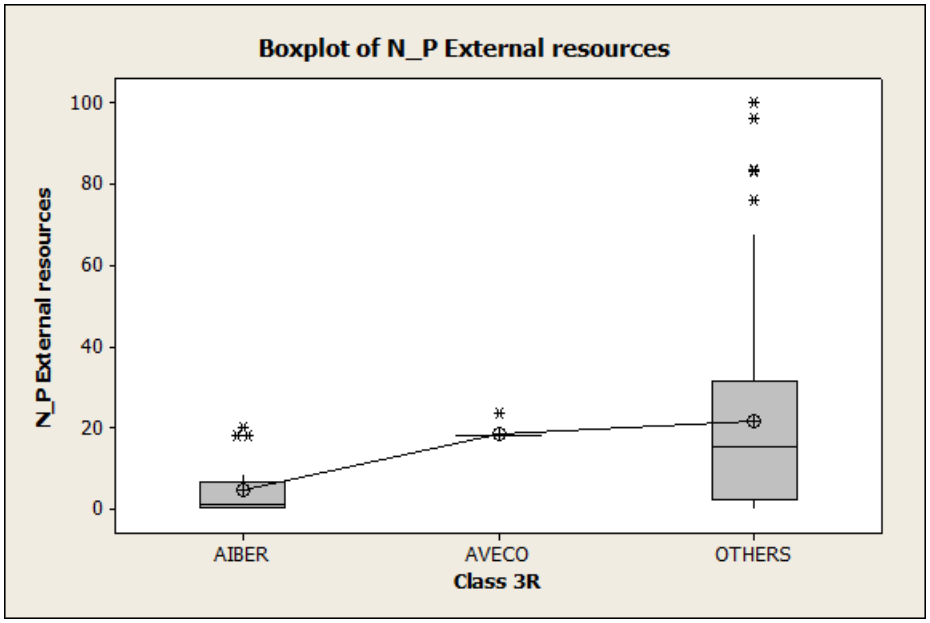
```
Source     DF     SS    MS      F      P
Class 6R    7  25703  3672  15,57  0,000
Error     124  29244   236
Total     131  54947

S = 15,36   R-Sq = 46,78%   R-Sq(adj) = 43,77%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  20   4,81   6,56    (---*----)
AVECO  27  18,32   1,06              (---*---)
CARLA   4  10,25  10,00  (----------*----------)
CWEAS  11   8,48   7,90     (-----*-----)
EMDEU  12   9,28   8,01     (-----*-----)
ESEAS  12  11,02  14,19      (-----*-----)
NSWUA  14   6,71   7,47    (----*-----)
SUSAF  32  41,52  27,72                              (---*--)
                          ---+---------+---------+---------+------
                             0        15        30        45

Pooled StDev = 15,36
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 15,57 (p-valor = 0,000).
Maior índice: SUSAF.
Menor índice: AIBER(Parece ser, mas não muito definido).
Não há outliers na AIBER, nem na AVECO.

Regiões 3R
O gráfico sugere que há diferença entre os grupos.

Gráfico 79.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P External resources versus Class 3R

```
Source     DF     SS    MS     F      P
Class 3R    2   4762  2381  6,12  0,003
Error     129  50185   389
Total     131  54947

S = 19,72   R-Sq = 8,67%   R-Sq(adj) = 7,25%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21   4,60   6,47  (----------*---------)
AVECO   27  18,32   1,06                    (--------*--------)
OTHERS  84  21,43  24,38                           (-----*----)
                          -----+---------+---------+---------+----
                             0,0       8,0      16,0      24,0

Pooled StDev = 19,72
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 6,12 (p-valor = 0,003).

Maior índice: Others (Parece não definido).

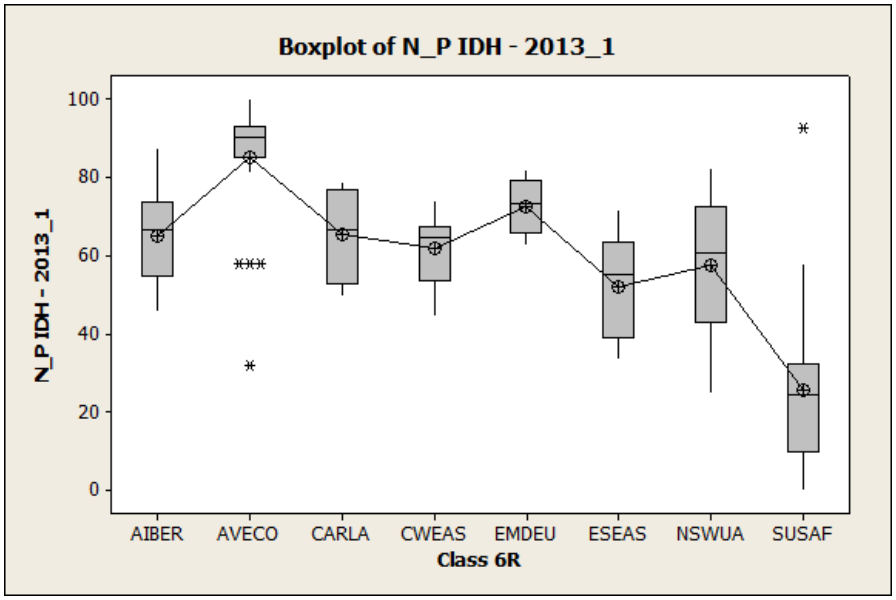
Menor índice: AIBER.

Variável N_P IDH – 2013_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 80.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P IDH - 2013_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  57555  8222  34,29  0,000
Error     124  29737   240
Total     131  87292

S = 15,49   R-Sq = 65,93%   R-Sq(adj) = 64,01%
```

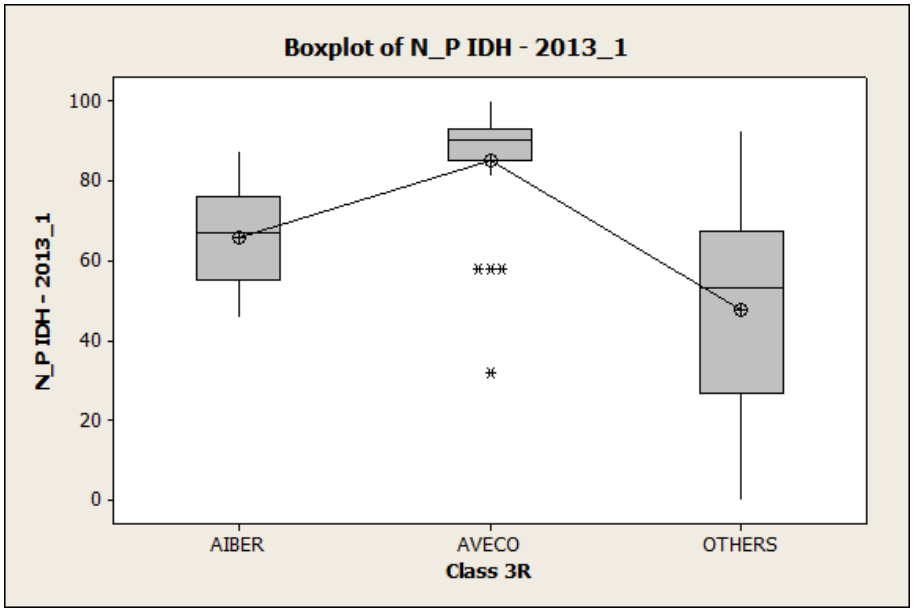
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  20  64,96  11,95                    (--*---)
AVECO  27  85,11  15,53                              (--*--)
CARLA   4  65,32  12,46                 (-------*------)
CWEAS  11  61,73   9,01                   (----*---)
EMDEU  12  72,60   6,81                         (---*----)
ESEAS  12  52,16  13,52               (---*----)
NSWUA  14  57,58  19,21                  (---*---)
SUSAF  32  25,69  19,80   (--*--)
                          +---------+---------+---------+---------
                          20        40        60        80

Pooled StDev = 15,49
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 34,29 (p-valor = 0,000).
Maior índice: AVECO.
Menor índice: SUSAF.

Regiões 3R
O gráfico sugere que há diferença entre os grupos.

Gráfico 81.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P IDH - 2013_1 versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  30392  15196  34,45  0,000
Error     129  56900    441
Total     131  87292

S = 21,00   R-Sq = 34,82%   R-Sq(adj) = 33,81%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65,62  12,03           (-----*-----)
AVECO   27  85,11  15,53                      (-----*----)
OTHERS  84  47,46  23,98  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 21,00
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 34,45 (p-valor = 0,000).

Maior índice: AVECO.

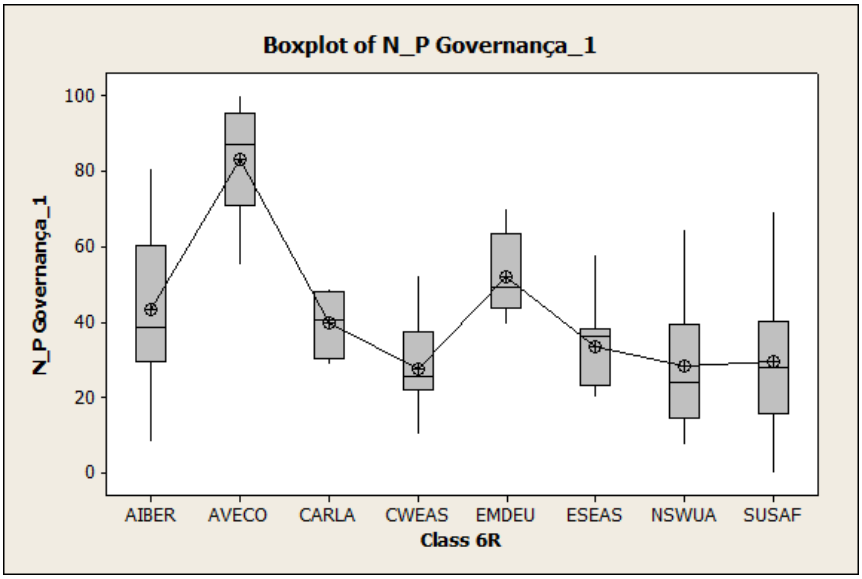
Menor índice: OTHERS.

Variável N_P IDH – 2013_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 82.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P GovernanÁa_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56677  8097  35,53  0,000
Error     124  28261   228
Total     131  84938

S = 15,10   R-Sq = 66,73%   R-Sq(adj) = 64,85%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  43,17  19,04           (---*--)
AVECO  27  83,25  13,38                              (--*--)
CARLA   4  39,62   9,54     (-------*------)
CWEAS  11  27,38  12,38  (----*---)
EMDEU  12  51,93  11,21               (---*---)
ESEAS  12  33,61  10,47     (----*---)
NSWUA  14  28,46  16,21   (---*---)
SUSAF  32  29,58  16,95     (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,10
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 35,53 (p-valor = 0,000).

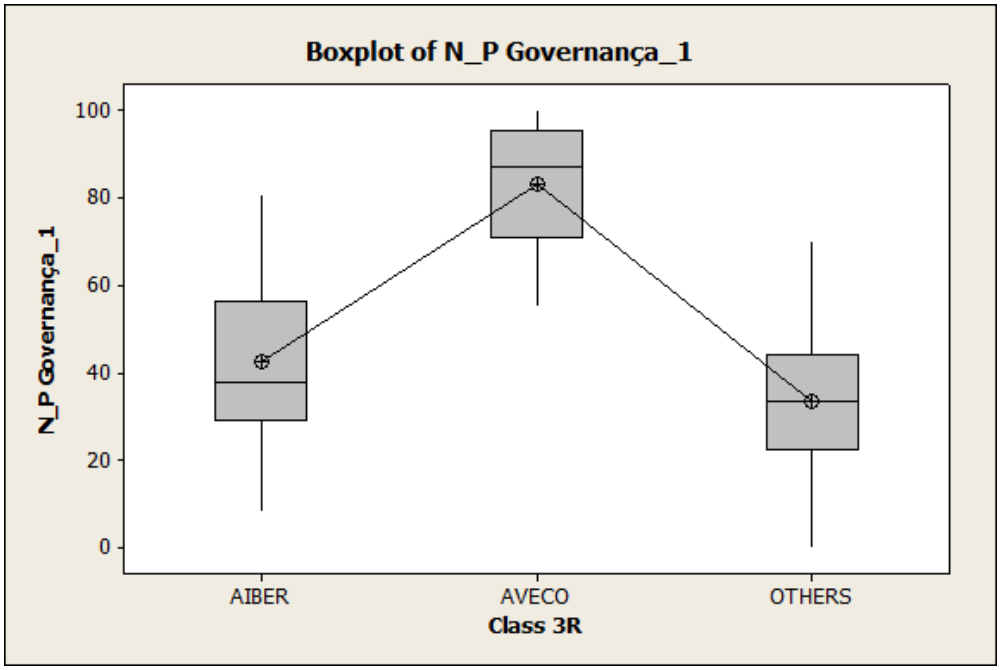
Maior índice: AVECO.

Menor índice: CWEAS, NSWUA ou SUSAF (Parece não muito definido).

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 83.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P GovernanÁa_1 versus Class 3R

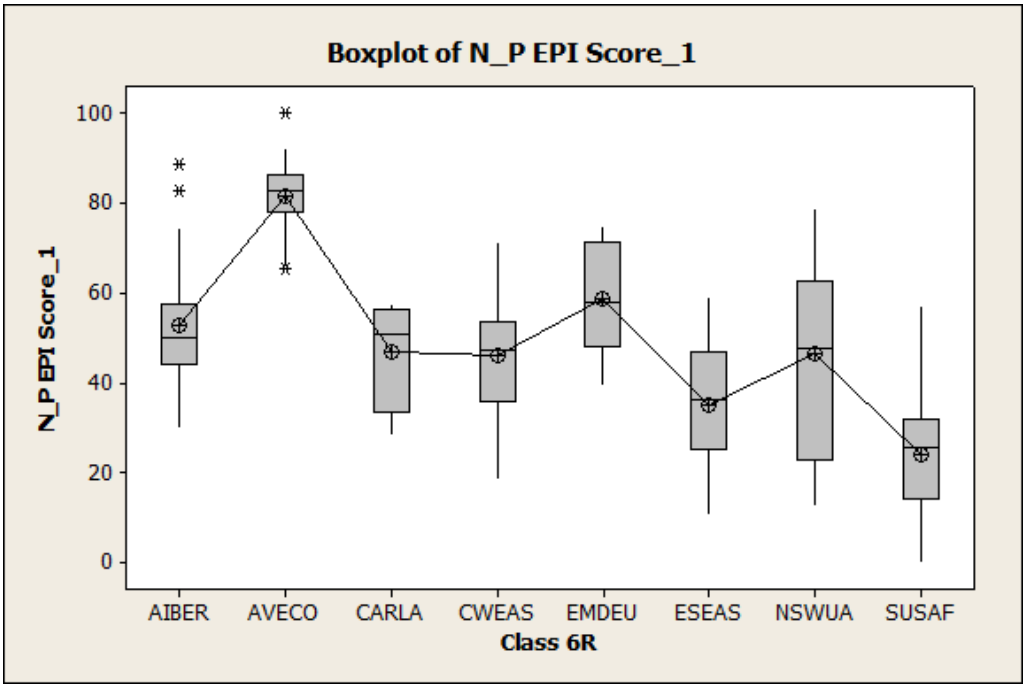
```
Source     DF     SS     MS      F      P
Class 3R    2  51019  25510  97,02  0,000
Error     129  33919    263
Total     131  84938

S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42,48  18,82       (---*----)
AVECO   27  83,25  13,38                                  (----*---)
OTHERS  84  33,36  16,35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75

Pooled StDev = 16,22
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 92,02 (p-valor = 0,000).
Maior índice: AVECO.
Menor índice: OTHERS.

Variável N_P EPI Score_1
Regiões 6R
O gráfico sugere que há diferença entre os grupos.

Gráfico 84.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P EPI Score_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  52052  7436  40,44  0,000
Error     124  22803   184
Total     131  74855

S = 13,56   R-Sq = 69,54%   R-Sq(adj) = 67,82%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  20  52,69  14,68               (--*--)
AVECO  27  81,39   8,14                                 (--*-)
CARLA   4  46,92  12,87          (-----*------)
CWEAS  11  46,14  14,39            (---*---)
EMDEU  12  58,63  11,91                 (---*---)
ESEAS  12  34,97  13,95       (--*---)
NSWUA  14  46,41  20,67             (--*---)
SUSAF  32  24,14  12,90   (-*-)
                          +---------+---------+---------+---------
                          20        40        60        80
Pooled StDev = 13,56
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 40,44 (p-valor = 0,000).

Maior índice: AVECO.

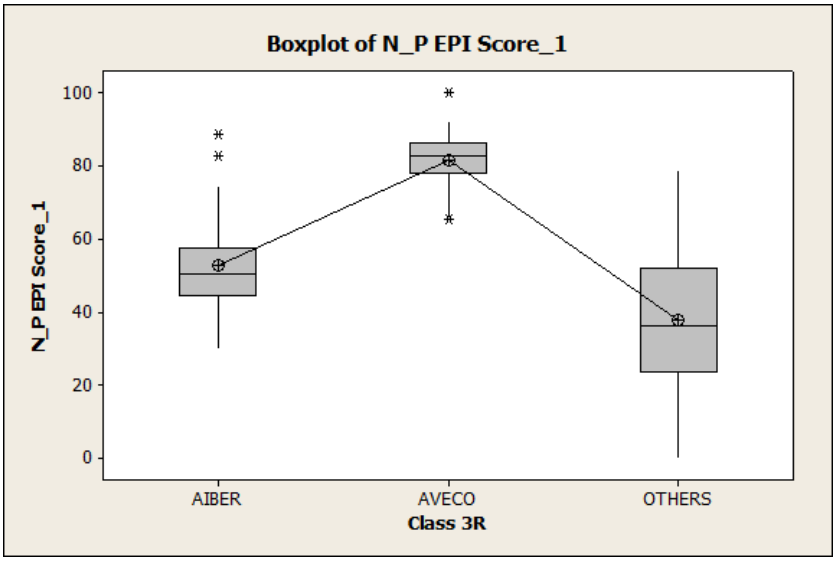
Menor índice: SUSAF.

Aiber e AVECO apresentam outliers.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 85.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P EPI Score_1 versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  38872  19436  69,68  0,000
Error     129  35982    279
Total     131  74855

S = 16,70   R-Sq = 51,93%   R-Sq(adj) = 51,18%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  52,70  14,31         (----*----)
AVECO   27  81,39   8,14                             (---*----)
OTHERS  84  37,95  19,06  (-*--)
                          -------+---------+---------+---------+--
                                45        60        75        90

Pooled StDev = 16,70
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 69,68 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: OTHERS.

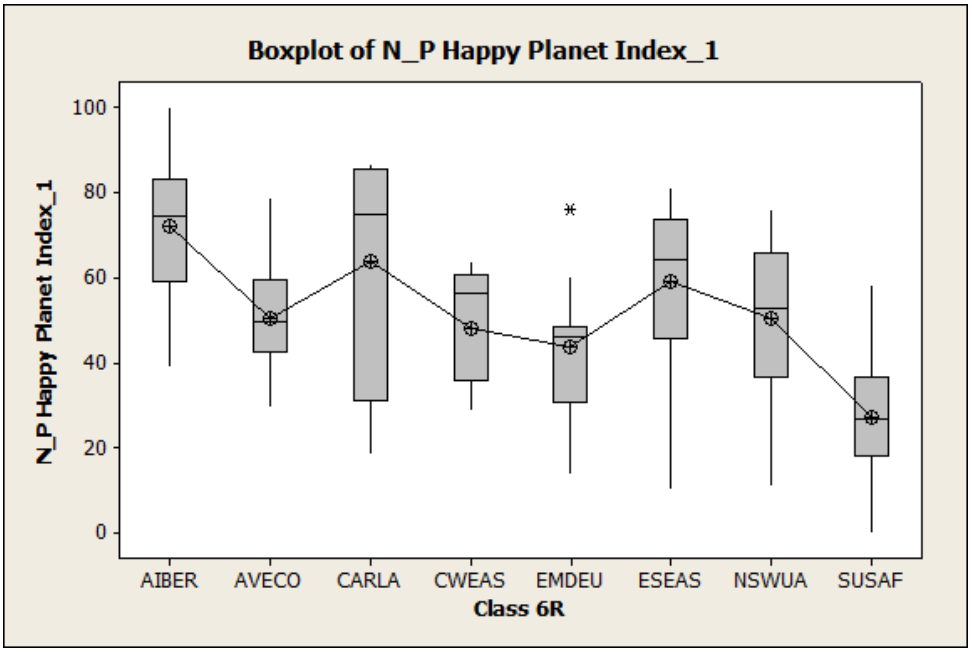
Há outliers em AIBER e AVECO.

Variável N_P Happy Planet Index_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 86.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Happy Planet Index_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  28129  4018  15,21  0,000
Error     124  32765   264
Total     131  60893

S = 16,26   R-Sq = 46,19%   R-Sq(adj) = 43,16%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  20  71,92  16,55                              (----*----)
AVECO  27  50,58  12,84                 (---*---)
CARLA   4  63,81  31,09                   (----------*---------)
CWEAS  11  48,13  13,54              (-----*------)
EMDEU  12  43,56  16,02           (-----*-----)
ESEAS  12  59,06  20,02                    (-----*------)
NSWUA  14  50,22  19,91               (----*-----)
SUSAF  32  27,27  14,09  (---*---)
                          ------+---------+---------+---------+---
                               30        45        60        75

Pooled StDev = 16,26
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 15,21 (p-valor = 0,000).

Maior índice: Não definido.

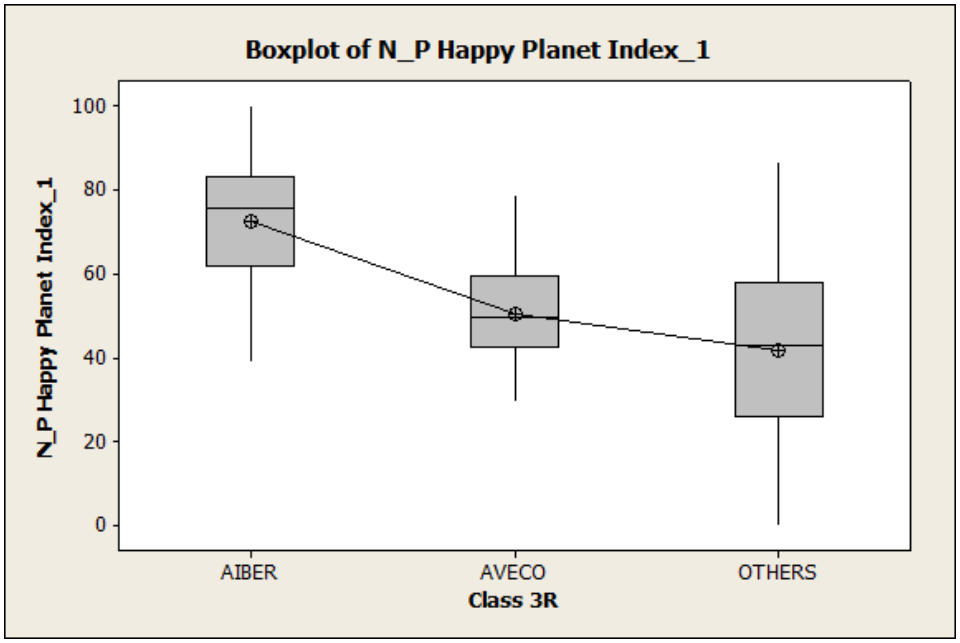
Menor índice: SUSAF.

Há um outlier em EMDEU.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 87.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P Happy Planet Index_1 versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  15845  7923  22,69  0,000
Error     129  45048   349
Total     131  60893

S = 18,69   R-Sq = 26,02%   R-Sq(adj) = 24,87%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72,36  16,25                        (-----*------)
AVECO   27  50,58  12,84       (-----*-----)
OTHERS  84  41,79  20,68  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18,69
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 22,69 (p-valor = 0,000).

Maior índice: AIBER.

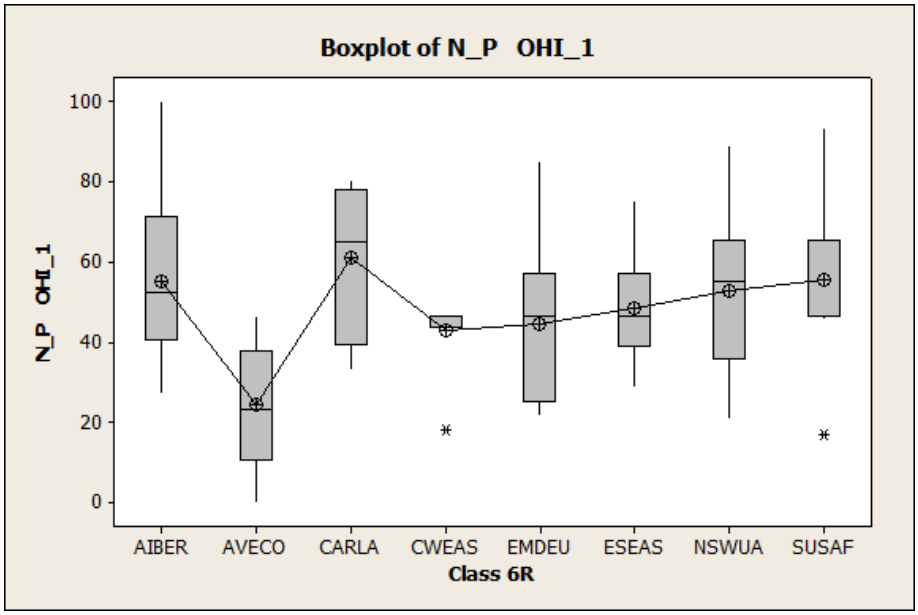
Menor índice: OTHERS.

Variável N_P OHI_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 88.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P OHI_1 versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  19147  2735  9,83  0,000
Error     124  34492   278
Total     131  53639

S = 16,68   R-Sq = 35,70%   R-Sq(adj) = 32,07%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   --------+---------+---------+---------+-
AIBER  20  55,28  19,51                     (----*----)
AVECO  27  24,24  15,48  (---*---)
CARLA   4  60,89  20,61                  (----------*----------)
CWEAS  11  43,04   8,44            (------*-----)
EMDEU  12  44,50  18,58             (------*-----)
ESEAS  12  48,38  13,03                (-----*------)
NSWUA  14  52,90  18,68                   (-----*-----)
SUSAF  32  55,52  16,89                     (---*---)
                          --------+---------+---------+---------+-
                                 30        45        60        75

Pooled StDev = 16,68
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 9,83 (p-valor = 0,000).

Maior índice: CARLA.

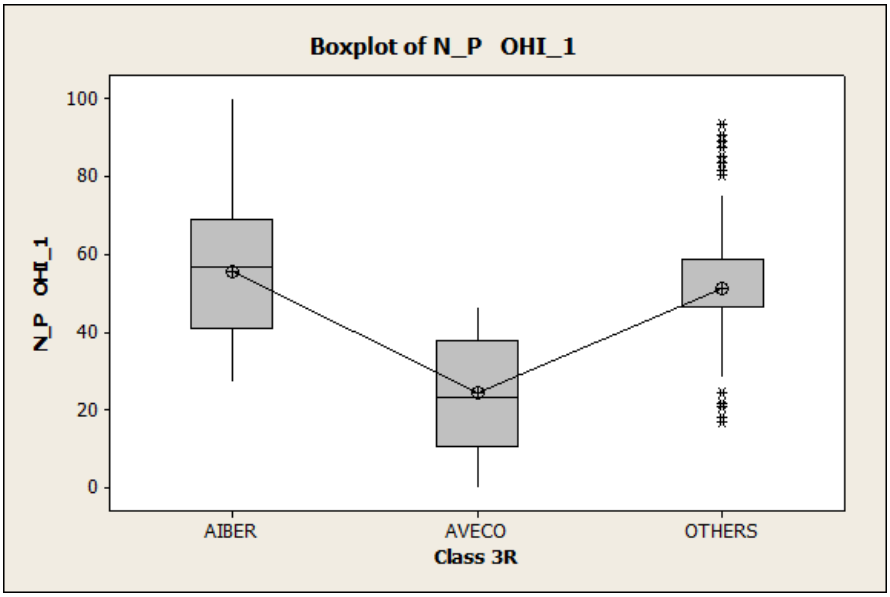
Menor índice: AVECO.

Há um outlier em CWEAS e outro em SUSAF.

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 89.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P OHI_1 versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  16816  8408  29,45  0,000
Error     129  36823   285
Total     131  53639

S = 16,90   R-Sq = 31,35%   R-Sq(adj) = 30,29%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  55,45  19,03                          (-----*-----)
AVECO   27  24,24  15,48  (----*-----)
OTHERS  84  51,07  16,77                          (--*--)
                          -----+---------+---------+---------+----
                              24        36        48        60

Pooled StDev = 16,90
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 29,45 (p-valor = 0,000).

Maior índice: Não definido.

Menor índice: AVECO.

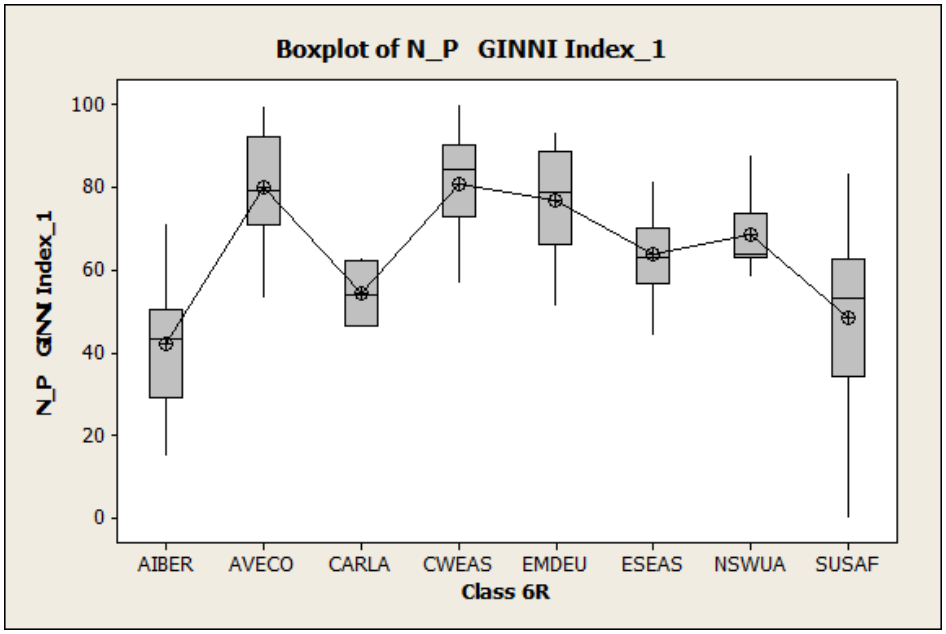
Há vários outliers em OTHERS.

Variável N_P GINNI Index_1

Regiões 6R

O gráfico sugere que há diferença entre os grupos.

Gráfico 90.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P GINNI Index_1 versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  29731  4247  18,31  0,000
Error     124  28761   232
Total     131  58492

S = 15,23   R-Sq = 50,83%   R-Sq(adj) = 48,05%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  20  42,09  14,02  (---*----)
AVECO  27  79,91  12,42                          (---*---)
CARLA   4  54,29   8,28    (---------*---------)
CWEAS  11  80,74  13,30                           (-----*-----)
EMDEU  12  76,95  12,94                        (-----*-----)
ESEAS  12  63,79  10,22               (-----*----)
NSWUA  14  68,63   9,15                  (-----*----)
SUSAF  32  48,52  21,97       (--*---)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 15,23
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 18,31 (p-valor = 0,000).

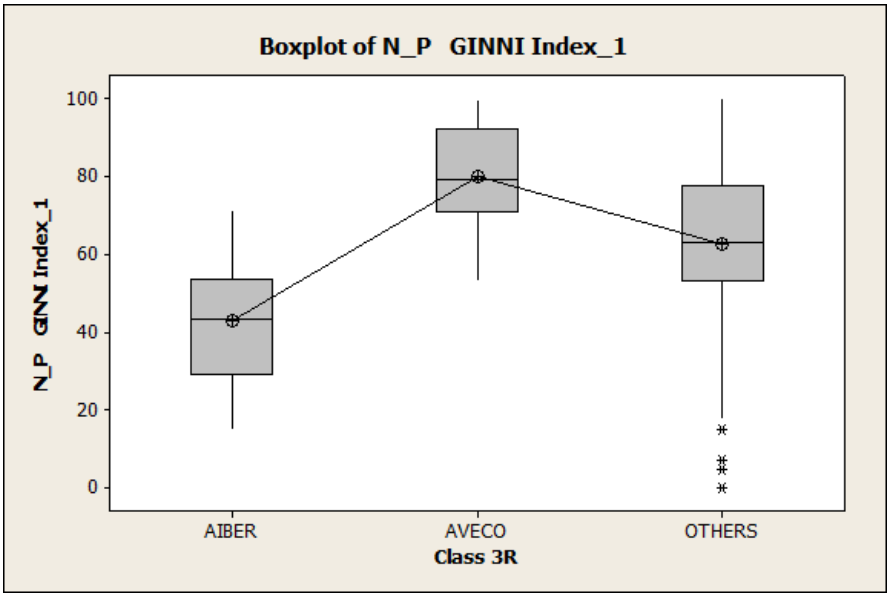
Maior índice: CWEAS (Mas parece não definido).

Menor índice: AIBER (Mas parece não definido).

Regiões 3R

O gráfico sugere que há diferença entre os grupos.

Gráfico 91.



Fonte: Elaborado pela autora

ANOVA

One-way ANOVA: N_P GINNI Index_1 versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  16078  8039  24,45  0,000
Error     129  42414   329
Total     131  58492

S = 18,13   R-Sq = 27,49%   R-Sq(adj) = 26,36%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  43,08  14,40  (----*----)
AVECO   27  79,91  12,42                          (---*----)
OTHERS  84  62,44  20,32                (--*-)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 18,13
```

Análise:

A ANOVA realizada mostrou que realmente há diferença significativa entre os grupos, F = 24,45 (p-valor = 0,000).

Maior índice: AVECO.

Menor índice: AIBER.

Há vários outliers em OTHERS.

Resumindo todas as análises, podemos montar uma tabela com todos os valores da estatística F, conforme mostrado abaixo.

Tabela 18.

Variável	6R
N_P_Life expectancy (years)_1	82,61
N_P_Deaths from	77,85
N_P_SPI_Regressao	58,42
N_P_Access to piped water	47,04
N_P Non-communicable disease	46,67
N_P_Child mortality	43,94
N_P EPI Score_1	40,44
N_P_Access to improved	37,87
N_P GovernanÁa_1	35,53
N_P_Maternal mortality	34,76
N_P IDH - 2013_1	34,29
N_P_Stillbirth rate	26,35
N_P Obesity rate (% of pop.)_1	25,27
N_P_Indoor air pollution	22,81
N_P GINNI Index_1	18,31
N_P External resources	15,57
N_P Happy Planet Index_1	15,21
N_P Outdoor air pollution	9,93
N_P OHI_1	9,83
N_P Health expenditure	9,15

Fonte: Elaborado pela autora

Tabela 19.

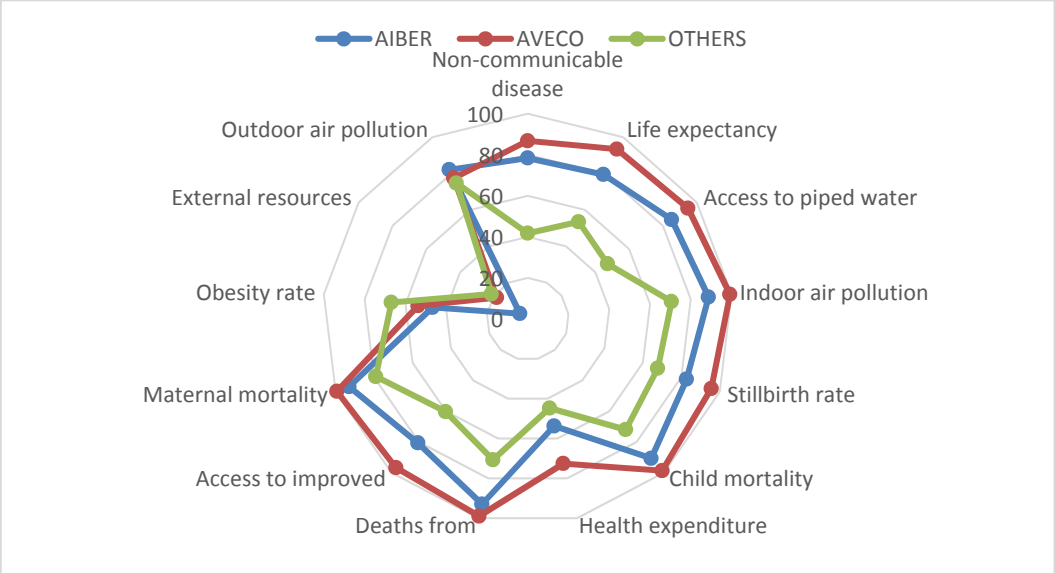
Variável	3R
N_P GovernanÁa_1	97,02
N_P Non-communicable disease	95,63
N_P_SPI_Regressao	83,97
N_P EPI Score_1	69,68
N_P_Life expectancy (years)_1	46,67
N_P_Access to piped water	37,63
N_P IDH - 2013_1	34,45
N_P OHI_1	29,45
N_P_Indoor air pollution	29,42
N_P_Stillbirth rate	25,25
N_P GINNI Index_1	24,45
N_P Happy Planet Index_1	22,69
N_P_Child mortality	22,28
N_P Health expenditure	20,47
N_P_Deaths from	19,03
N_P_Access to improved	18,30
N_P_Maternal mortality	14,62
N_P Obesity rate (% of pop.)_1	9,50
N_P External resources	6,12
N_P Outdoor air pollution	1,32

Fonte: Elaborado pela autora

Análise:

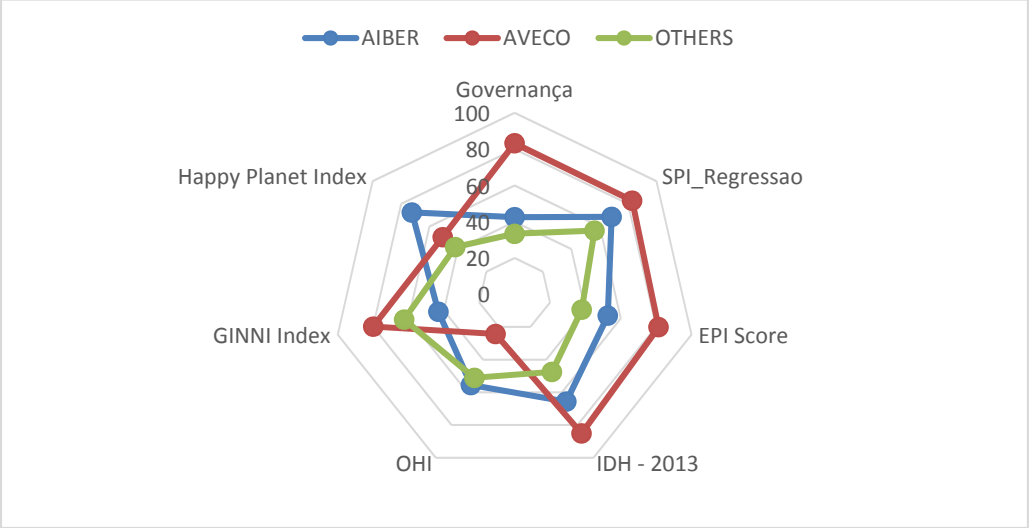
Exceto para a variável N_P Outdoor air pollution da tabela com 3 regiões, todos os outros valores de F foram significativos. Ou seja, para cada variável, as médias das regiões foram consideradas estatisticamente diferentes.

Gráfico 92.



Fonte: Elaborado pela autora

Gráfico 93.



Fonte: Elaborado pela autora

3.10. ANÁLISE AMOSTRAL

Escolheu-se para a análise amostral duas variáveis: a variável *Deaths for infectious desease*, referente à saúde; e outra referente ao Bem-Estar Social: *Outdoor Air Pollution* Percebe-se que essa variável sempre se encontra entre as principais variáveis correlacionadas com o tema Saúde e Bem-estar para a região AIBER (em destaque nos Quadros a seguir). Trabalhou-se, portanto, com três amostras aleatórias, sendo uma com 25 países, uma com 50 países e outra com 100 países. Como os dados foram normalizados e positivados, coloca-se "N_P" antes da variável.

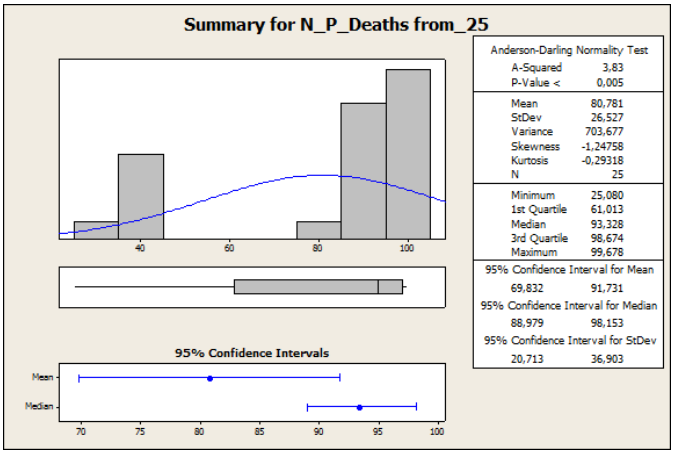
3.10.1 *Deaths for infectious desease* para Amostra de 25 países. Ver Quadro 4

Quadro 4.

N_P_Deaths from_25	
Macedonia	98,9550
Kazakhstan	93,2476
Angola	35,0482
Denmark	98,7138
Netherlands	98,6334
Azerbaijan	92,6849
Botswana	41,4791
Bolivia	80,5466
Bulgaria	98,3923
New Zealand	99,6785
Nicaragua	93,8907
Tanzania	38,0225
Uruguay	96,4630
Burkina Faso	36,4952
Uganda	35,7717
Brazil	93,0868
Albania	97,1865
Serbia	99,5177
Iran	94,2926
Ukraine	93,3280
Burundi	25,0804
Sweden	99,2765
Jamaica	88,1029
Trinidad and Tobago	92,5241
Bosnia and Herzegovina	99,1158

Fonte: Elaborado pela autora

Gráfico 92.



Fonte: Elaborado pela autora

Análise:

Significa que os 25 países amostrais têm comportamentos distintos quanto à mortes por doenças infecciosas. Segundo o Gráfico 92, a média é de 80,78 (Com intervalo de confiança de 95% entre 69,83 e 91,73) e desvio padrão de 26,52, (com intervalo de confiança de 95% entre 20,71 e 36,90). Percebe-se uma concentração dos países à direita; e a formação de dois grupos distintos. Os países Ibero Americanos parecem estar mais à direito do histograma, e os africanos, mais à esquerda.

3.10.2 *Deaths for infectious desease* para Amostra de 50 países. Ver Quadro 5

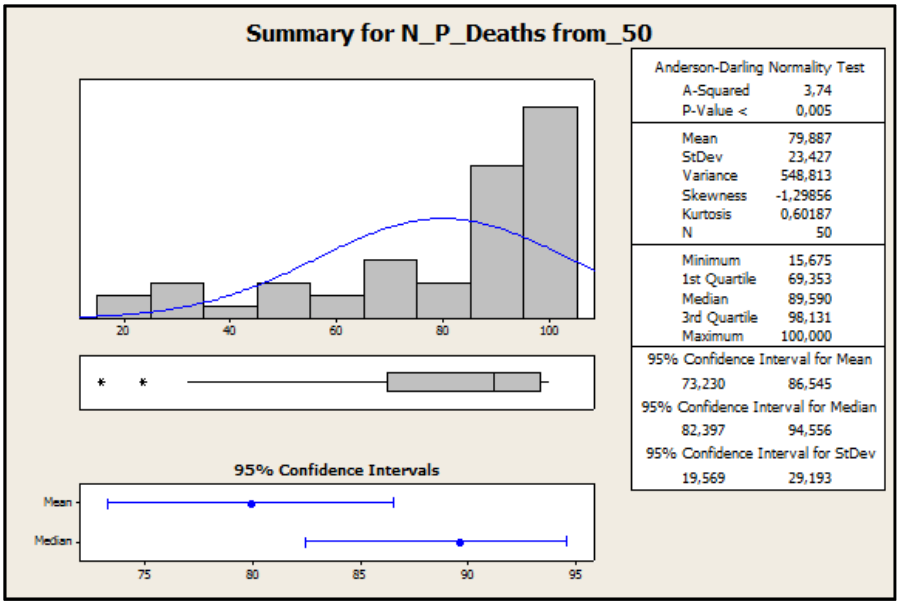
Quadro 5.

N_P_Deaths from_50	
Madagascar	68,087
Tunisia	90,113
Djibouti	63,103
Netherlands	98,633
Yemen	82,235
Togo	49,839
Honduras	88,666
El Salvador	90,675
Central African Republic	15,675
Italy	99,598
Bangladesh	73,232
Dominican Republic	89,068
Sudan	70,579
Latvia	98,473
Australia	99,437
Uganda	35,772
Malaysia	86,013
Georgia	95,096
Venezuela	95,177
Mali	34,405
Cameroon	31,672
Jamaica	88,103
Guyana	85,289
Ukraine	93,328
Switzerland	99,518
Mauritius	96,704
Cambodia	62,460
Slovakia	98,071
Portugal	97,186
Nigeria	34,003
Costa Rica	98,312
Panama	93,489
Peru	86,977
Iceland	99,437
Czech Republic	98,875
Rwanda	53,055

Iran	94,293
Laos	70,659
Croatia	99,277
Tajikistan	82,476
Poland	98,633
Argentina	93,891
Zambia	23,633
Indonesia	81,270
Trinidad and Tobago	92,524
United Kingdom	97,990
Finland	100,00
Pakistan	69,775
Kenya	50,723
Greece	98,875

Fonte: Elaborado pela autora

Gráfico 93.



Fonte: Elaborado pela autora

Análise:

Significa que os 50 países amostrais têm comportamentos distintos quanto à mortes por doenças infecciosas, mais similares até o terceiro quartil. Segundo o Gráfico 92, a média é de 79,887 (Com intervalo de confiança de 95% entre 73,23 e 86,54) e desvio padrão de 23,42, (com intervalo de confiança de 95% entre 19,569 e 29,193). Os países Ibero Americanos parecem estar mais à direita do histograma, no último quartil. Finlândia, pertencente a AVECO, é o país que apresenta melhor indicador de todos os países.

3.10.3 *Deaths for infectious desease* para Amostra de 100 países. Ver Quadro 6

Quadro 6.

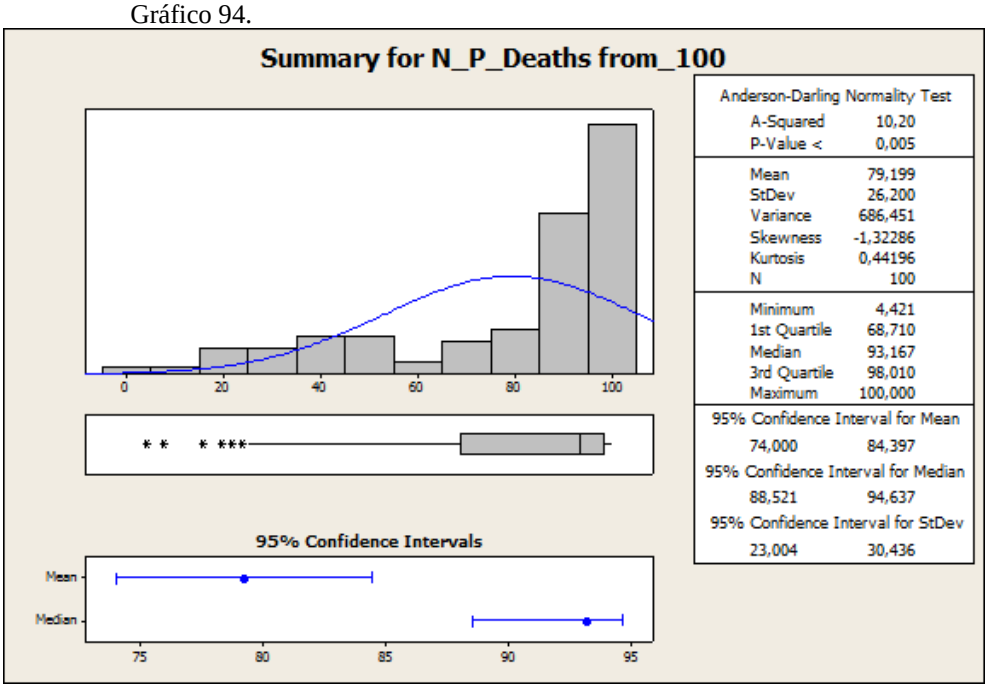
N_P_Deaths from_100	
Malaysia	86,013
Iran	94,293
United States	98,151
Sweden	99,277
Indonesia	81,270
Honduras	88,666
Central African Republic	15,675
Laos	70,659

Mauritania	54,662
Ukraine	93,328
Australia	99,437
China	96,222
Romania	97,830
Peru	86,977
Cuba	97,106
Germany	99,196
France	99,035
Botswana	41,479
Rwanda	53,055
Nicaragua	93,891
Nepal	73,714
Kenya	50,723
Canada	99,035
Austria	99,759
Togo	49,839
Uzbekistan	92,524
Paraguay	93,408
Brazil	93,087
Nigeria	34,003
Sri Lanka	94,534
United Arab Emirates	95,016
Mongolia	93,730
Denmark	98,714
Namibia	47,026
Malawi	7,958
Chad	19,775
Tajikistan	82,476
Finland	100,00
Croatia	99,277
Panama	93,489
Jordan	94,132
Cameroon	31,672
Kazakhstan	93,248
Trinidad and Tobago	92,524
Greece	98,875
Madagascar	68,087
Algeria	84,646
Spain	98,955
Philippines	82,315
Macedonia	98,955
El Salvador	90,675
Moldova	96,141
Lithuania	98,151
Colombia	95,740
Bosnia and Herzegovina	99,116
Turkey	96,624
Armenia	94,936
Mali	34,405
Korea, Republic of	98,553
Bolivia	80,547
Ecuador	92,444
Dominican Republic	89,068
Belgium	98,232
Burundi	25,080
Liberia	38,023
Russia	95,177
Yemen	82,235
Venezuela	95,177
Montenegro	99,518
Ireland	98,553
Kyrgyzstan	91,881
Estonia	99,196
Georgia	95,096
Jamaica	88,103
Swaziland	4,421
Slovakia	98,071
Sudan	70,579
Niger	42,203
Italy	99,598
Lebanon	97,106
Switzerland	99,518
India	71,704
Mozambique	23,955
Angola	35,048
Congo, Republic of	45,257

Mexico	95,418
Iraq	87,379
Poland	98,633
South Africa	21,865
Azerbaijan	92,685
Senegal	59,968
Burkina Faso	36,495
Cambodia	62,460
Tanzania	38,023
Guatemala	82,797
Argentina	93,891
Portugal	97,186
Belarus	98,473
Tunisia	90,113
Kuwait	96,624

Fonte: Elaborado pela autora

Gráfico 94.



Fonte: Elaborado pela autora

Análise:

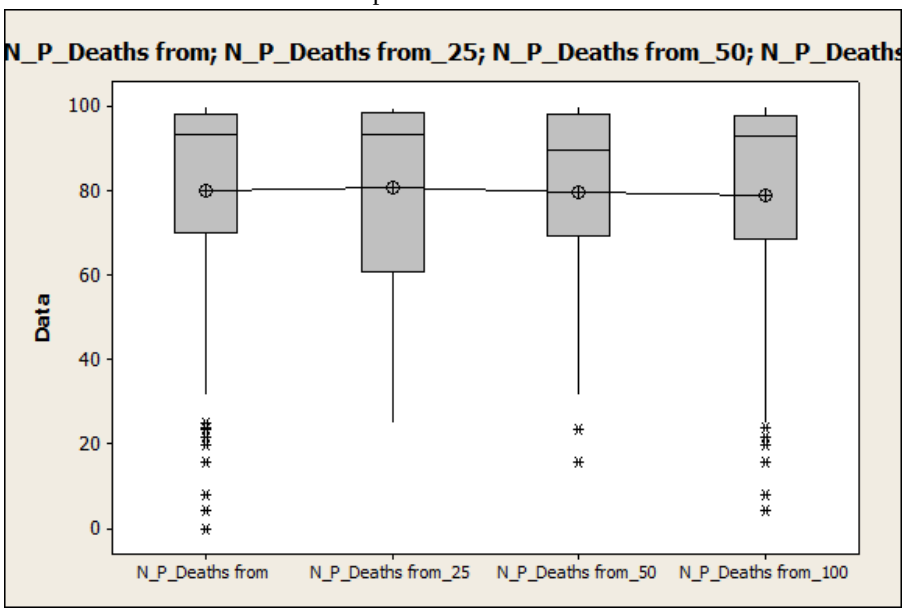
Numa amostra de 100, observa-se que a taxa de sobrevivência por doenças infecciosas diminuiu, e a taxa de sobrevida em alguns países permaneceu mais ou menos a mesma, o que pode ser observado pelo último quartil, na parte mais lateralizada à direita do gráfico 94.

3.10.4 *Deaths for Infectious Deseases*: ANOVA

Após calcular as amostras, foi calculada uma ANOVA comparando o indicador para os diferentes cálculos amostrais de 25, 50 e 100 variáveis. Sabe-se que a ANOVA é uma análise que testa a importância de um ou mais fatores comparando as médias das variáveis de resposta em diferentes níveis dos fatores. A hipótese nula afirma que todas as médias das populações (médias dos níveis dos fatores) são iguais, enquanto a hipótese alternativa afirma que pelo menos uma é diferente.

Boxplot of N_P_Deaths from; N_P_Deaths from_25; N_P_Deaths from_50; N_P_Deaths

Gráfico 95. ANOVA para os diferentes cálculos amostrais



Fonte: Elaborado pela autora

One-way ANOVA: N_P_Deaths f; N_P_Deaths f; N_P_Deaths f; N_P_Deaths f

Source	DF	SS	MS	F	P
Factor	3	64	21	0,03	0,992
Error	303	201716	666		
Total	306	201779			

S = 25,80 R-Sq = 0,03% R-Sq(adj) = 0,00%

Level	N	Mean	StDev	
N_P_Deaths from		132	79,96	26,21
N_P_Deaths from_25	25	80,78	26,53	
N_P_Deaths from_50	50	79,89	23,43	
N_P_Deaths from_100	100	79,20	26,20	

Individual 95% CIs For Mean Based on Pooled StDev

```
Level                --+---------+---------+---------+-------
N_P_Deaths from             (------*-------)
N_P_Deaths from_25   (----------------*----------------)
N_P_Deaths from_50       (-----------*-----------)
N_P_Deaths from_100         (-------*-------)
                     --+---------+---------+---------+-------
                     72,0      78,0      84,0      90,0
```

Pooled StDev = 25,80

O menor intervalo pela ANOVA, seja pelo gráfico ou pelos números é o N_P_Deaths from_100 100 79,20 26,20
Com 95% de confiança.

Análise:

A média está entre 79,20 pontos para uma amostra de 100 e 80,78 para uma amostra de 25, e se confirma o que se viu nos gráficos. A menor taxa de sobrevida por doença infecciosa se observou para uma amostra maior, a de 100 países.

Pode-se observar essa diferença por região. O Minitab oferece diferentes tipos de ANOVA para fatores adicionais, tipos de fatores e experimentos diferentes para atender a diferentes necessidades. Parte-se, portanto, para a classificação agora apenas para as seguintes regiões:

- AIBER – REGIÃO DE PAÍSES IBERO AMERICANOS
- AVECO – PAISES DESENVOLVIDOS
- OTHERS – OUTROS PAÍSES

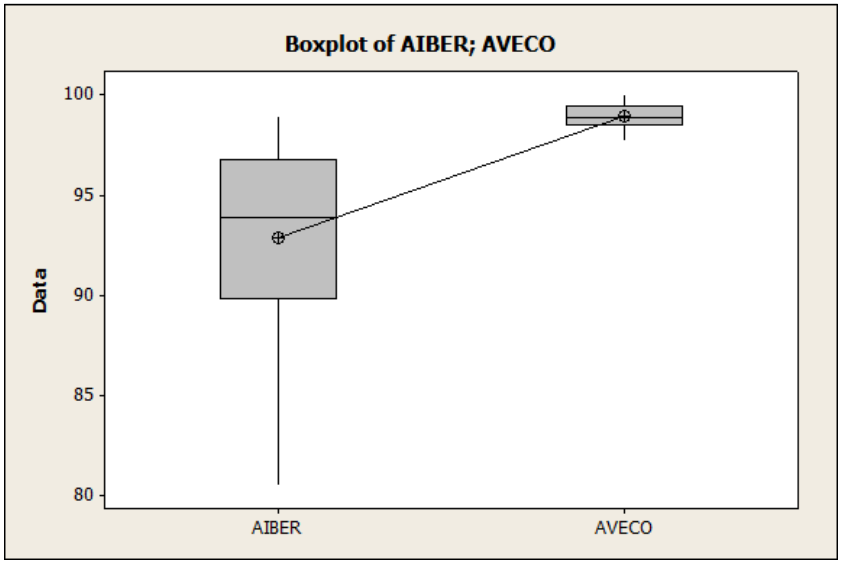
Foram retirados, portanto, os outros países e comparados apenas Aiber e Aveco para o mesmo indicador.

One-way ANOVA: AIBER; AVECO

```
Source  DF     SS     MS      F      P
Factor   1  426,2  426,2  39,67  0,000
Error   46  494,2   10,7
Total   47  920,4

S = 3,278   R-Sq = 46,31%   R-Sq(adj) = 45,14%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
AIBER  21  92,89   4,92  (-----*----)
AVECO  27  98,90   0,60                          (----*----)
                          ----+---------+---------+---------+-----
                           92,5      95,0      97,5     100,0

Pooled StDev = 3,28
```

Boxplot of AIBER; AVECO

Gráfico 96. ANOVA para AIBER e AVECO



Fonte: Elaborado pela autora

Análise:

Aqui também se observa que nos países AVECO (média 98,8) há uma maior sobrevida por doenças infecciosas, o que tem sentido, uma vez que o meio ambiente e as melhores condições sanitárias e de tratamento geral com a saúde nos países desenvolvidos devem evitar algumas mortes.

Two-Sample T-Test and CI: AIBER; AVECO

```
Two-sample T for AIBER vs AVECO

        N    Mean  StDev  SE Mean
AIBER  21   92,89   4,92      1,1
AVECO  27  98,898  0,598     0,12


Difference = mu (AIBER) - mu (AVECO)
Estimate for difference:  -6,01
95% upper bound for difference:  -4,14
T-Test of difference = 0 (vs <): T-Value = -5,56  P-Value = 0,000  DF = 20
```

Análise:

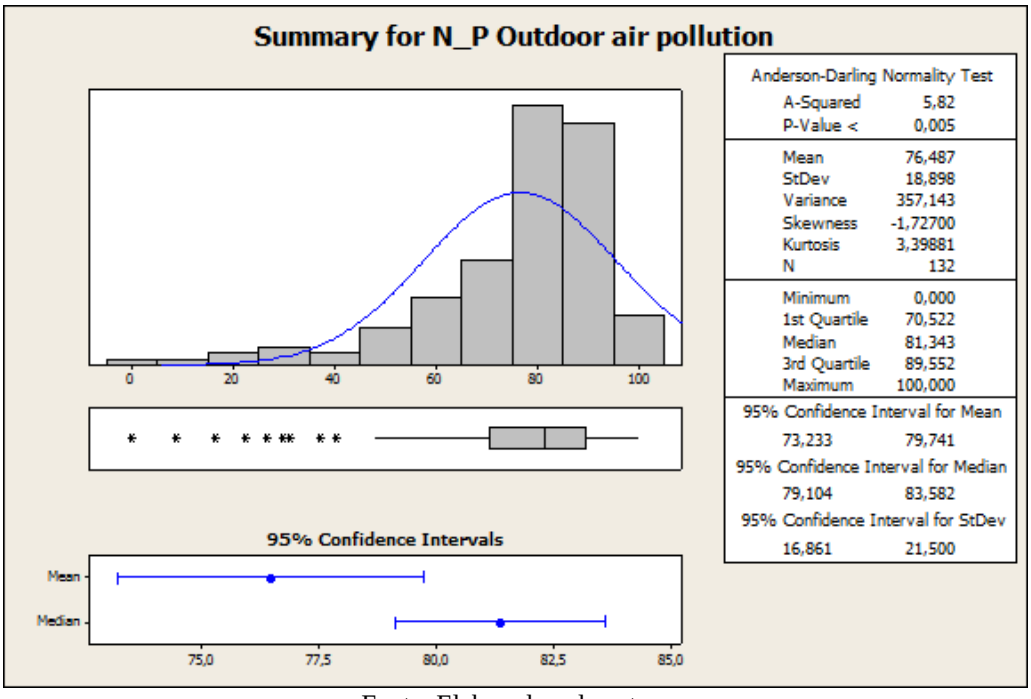
Aqui, o p-value = 0, já no gráfico boxplot gerado pela ANOVA da comparação das amostras, o p-value era de 0,992, o que significa que há uma maior confiança, menor p-value.

3.10.5 *Outdoor Air Pollution*: MESMOS PROCEDIMENTOS PARA Indicador de Bem-Estar Social

Agora, se vai analisar um indicador de Bem-estar Social: Outdoor Air Pollution, para 25 países, 50 e outra amostra de 100 países.

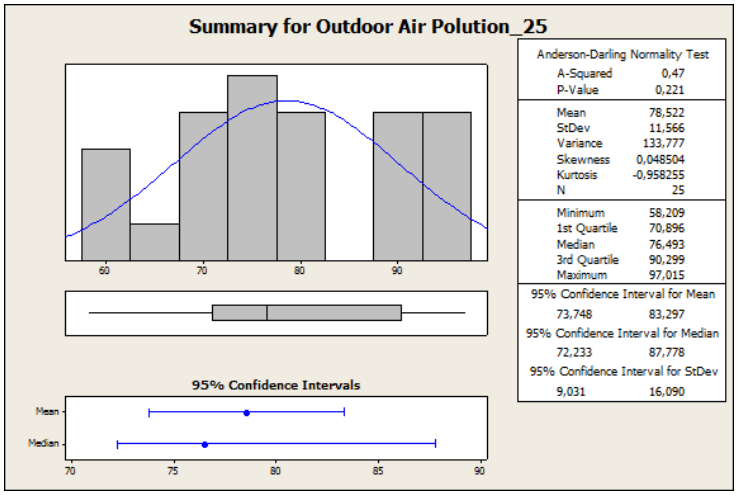
Depois, se seguirá o mesmo procedimento acima, comparando as amostras e depois analisando e comparando pelas regiões dadas.

Gráfico 97.



Fonte: Elaborado pela autora

Gráfico 98.

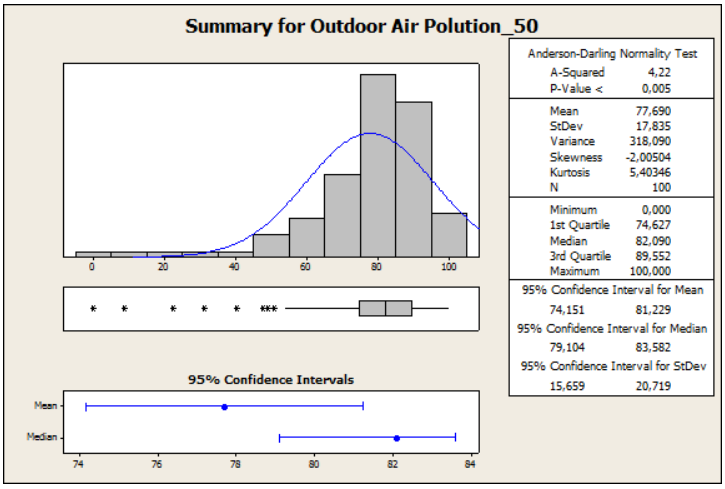


Fonte: Elaborado pela autora

Análise:

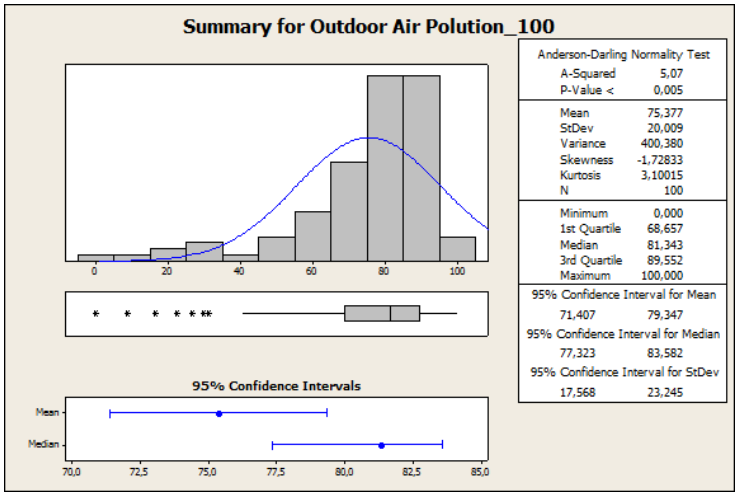
Interessante que para os 25 países, a curva normal é também evidenciada como nas amostras seguintes para de 50 e 100 países.

Gráfico 99.



Fonte: Elaborado pela autora

Gráfico 100.



Fonte: Elaborado pela autora

Calculando One-way ANOVA: Outdoor Air Polu; Outdoor Air Polu; Outdoor Air Polution, tem-se:

Source	DF	SS	MS	F	P
Factor	2	355	178	0,53	0,589
Error	222	74339	335		
Total	224	74695			

S = 18,30 R-Sq = 0,48% R-Sq(adj) = 0,00%

Level	N	Mean	StDev
Outdoor Air Polution_25	25	78,52	11,57
Outdoor Air Polution_50	100	77,69	17,84
Outdoor Air Polution_100	100	75,38	20,01

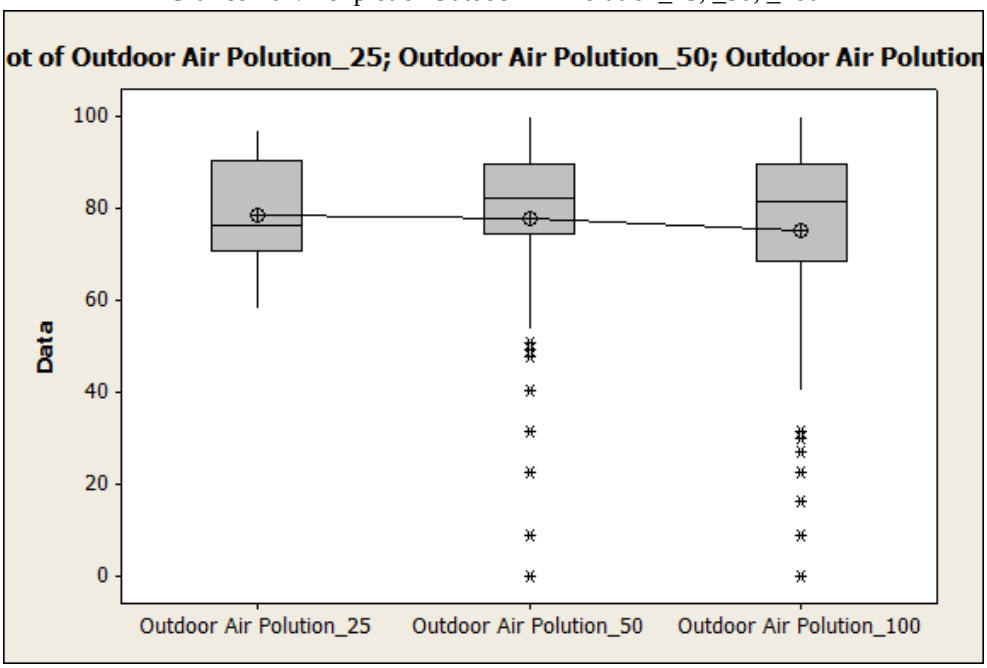
Análise:
A melhor amostra é a de 100.

Individual 95% CIs For Mean Based on Pooled StDev

```
Level                    --+---------+---------+---------+-------
Outdoor Air Polution_25  (-----------------*-----------------)
Outdoor Air Polution_50           (--------*--------)
Outdoor Air Polution_100  (--------*--------)
                         --+---------+---------+---------+-------
                          72,0      76,0      80,0      84,0
```

Pooled StDev = 18,30

Gráfico 101. Boxplot of Outdoor Air Polution_25; _50; _100



Fonte: Elaborado pela autora

Análise:
Pode-se observar outliers na amostra de 100: Ukraine 0,00; Armenia 8,955; Bulgaria 16,418; Bosnia and Herzegovina 22,388; Georgia 26,866; Russia 29,851; Latvia 31,343.

Outliers na amostra de 50: Ukraine 0,00; Armenia 8,955; Bosnia and Herzegovina 22,388; Latvia 31,343; Kazakhstan 40,299; China 47,761; Djibouti 49,254; Turkey 50,746.

Fazendo a ANOVA entre AVECO e AIBER do indicador Outdoor Air Polution
One-way ANOVA: aiber_OAP; AVECO_OAP

Source	DF	SS	MS	F	P
Factor	1	249	249	1,58	0,215
Error	46	7240	157		
Total	47	7488			

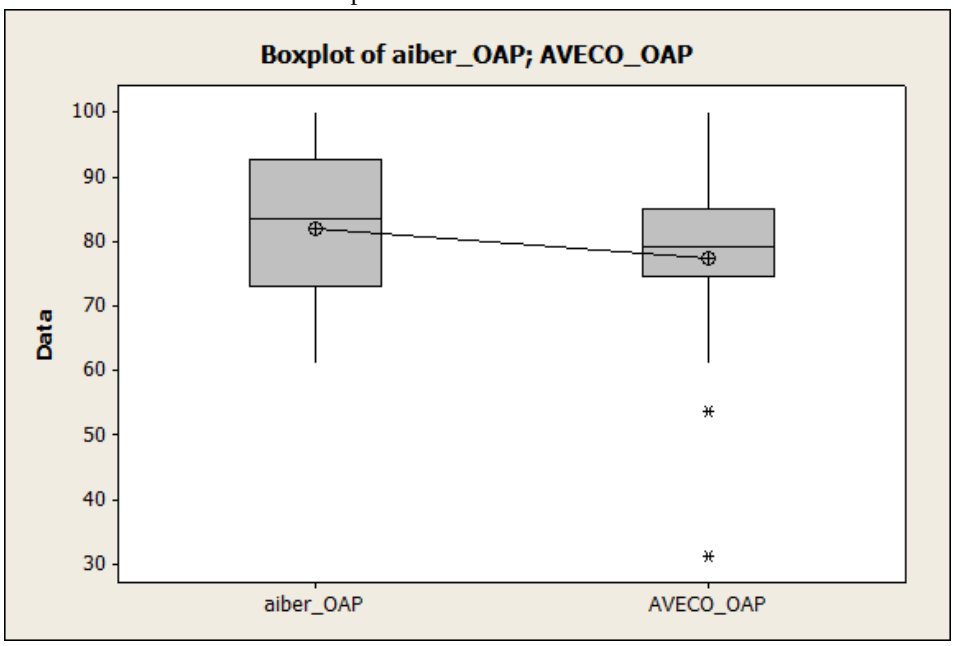
S = 12,55 R-Sq = 3,32% R-Sq(adj) = 1,22%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
aiber_OAP	21	82,09	11,37	(----------*----------)
AVECO_OAP	27	77,50	13,38	(----------*----------)

76,0 80,0 84,0 88,0

Pooled StDev = 12,55

Gráfico 102. Boxplot de AIBER e AVECO referente OAP



Fonte: Elaborado pela autora

Análise:

Melhorou o nível de confiança, mas ainda o p-value ainda não é 0. Há dois outliers na AVECO: linha 11 Greece 53,732, e linha 18 Latvia 31,343. A melhor região é a AVECO.

Quando há uma maior amostra, parece ser que há um menor intervalo e uma maior amplitude.

3.11 CRIAÇÃO DE NOVOS COMPONENTES

Após o uso do recurso do Minitab Stat< Multivariate< Principal Component Analysis, solicitou-se o agrupamento das variáveis em 4 grupos não correlacionados entre si. O Resultado pode ser visto no Tabela 20. abaixo, e para maiores detalhes, vide o próprio arquivo.

Tabela 20. PCs por país e região

País_1_1	Country Code_1_1	REGIÃO_1_1	PC1	PC2	PC3	PC4
Albania	ALB	Southern Europe	1,48393	1,30652	0,31069	-1,51779
Algeria	DZA	Northern Africa	1,03567	0,86331	0,94299	-0,21206
Angola	AGO	Middle Africa	-5,49677	0,65184	0,02368	1,02992
Argentina	ARG	IBE	2,44912	1,27289	0,44639	0,96058
Armenia	ARM	Western Asia	1,32243	1,76578	-3,12608	0,16597
Australia	AUS	Australia and New Zealand	4,17977	-1,94043	0,48571	-0,36877
Austria	AUT	Western Europe	4,09943	-1,13667	0,48816	-0,79600
Azerbaijan	AZE	Western Asia	0,01518	1,56683	-1,33242	-0,08090
Bangladesh	BGD	Southern Asia	-3,27916	1,17868	0,67583	-2,05453
Belarus	BLR	Eastern Europe	2,01811	0,80352	-2,50084	0,44341
Belgium	BEL	Western Europe	3,90512	-2,06783	-0,27637	-0,18985
Benin	BEN	Western Africa	-4,82270	-1,36098	-0,01075	-0,00136
Bolivia	BOL	IBE	-0,34998	-0,02524	1,59180	0,64677
Bosnia and Herzegovina	BIH	Southern Europe	1,75342	1,55297	-2,07762	0,20028
Botswana	BWA	Southern Africa	-1,88650	-1,84571	0,86190	2,97366
Brazil	BRA	IBE	1,29637	1,31408	1,50887	0,70323
Bulgaria	BGR	Eastern Europe	2,01885	0,11781	-3,21214	0,63312
Burkina Faso	BFA	Western Africa	-5,38294	-1,92116	-0,12963	-0,51724
Burundi	BDI	Eastern Africa	-6,66439	-2,48343	-0,52557	-0,65985
Cambodia	KHM	South-Eastern Asia	-3,59211	0,69025	0,00336	-2,37238
Cameroon	CMR	Middle Africa	-5,14722	0,11572	-0,69857	0,63042
Canada	CAN	Northern America	4,11741	-1,68265	0,59889	-0,32577
Central African Republic	CAF	Eastern Africa	-7,56276	-1,66070	-0,08847	1,49735
Chad	TCD	Middle Africa	-7,85996	-1,11574	-0,68132	0,69785
Chile	CHL	IBE	3,23952	0,58341	1,46836	1,01362
China	CHN	Eastern Asia	0,17167	1,25745	-0,73146	-0,64146
Colombia	COL	IBE	1,32628	1,68880	2,54797	0,57018
Congo, Republic of	COG	Middle Africa	-5,30548	0,42763	-0,86882	-0,02069
Costa Rica	CRI	IBE	2,99378	1,10838	2,86121	0,19386
Croatia	HRV	Southern Europe	2,80068	-0,81430	-0,91709	0,14269
Cuba	CUB	Caribbean	2,33218	1,32332	0,48608	0,08433
Czech Republic	CZE	Eastern Europe	3,19040	-1,00073	-0,37957	0,03429
Denmark	DNK	Northern Europe	4,40779	-2,83228	-0,84664	-0,61297
Djibouti	DJI	Eastern Africa	-2,99172	0,46089	-1,21879	0,72419
Dominican Republic	DOM	IBE	0,58054	1,24660	0,86960	0,44863
Ecuador	ECU	IBE	1,47315	1,13021	1,53740	0,22221
Egypt	EGY	Northern Africa	1,33424	1,30590	-1,14169	0,58701
El Salvador	SLV	IBE	0,80913	1,76136	1,95832	-0,03556
Estonia	EST	Northern Europe	3,11876	-2,22210	-0,06039	0,04092
Finland	FIN	Northern Europe	4,51422	-2,41651	-0,14636	-0,82594
France	FRA	Western Europe	3,84871	-1,30066	0,51296	-0,76133
Georgia	GEO	Western Asia	0,44840	1,50677	-2,21501	-0,00638
Germany	DEU	Western Europe	4,43342	-1,70653	0,19397	-0,58483
Ghana	GHA	Western Africa	-3,69408	-0,38801	0,63428	-0,22166

Greece	GRC	Southern Europe	3,31271	-0,37891	-0,72227	-0,05678
Guatemala	GTM	IBE	0,26590	1,20948	1,89511	0,21365
Guinea	GIN	Western Africa	-6,45545	0,82211	-0,71541	-0,14092
Guyana	GUY	South America	-0,90985	1,08994	0,87132	0,45900
Honduras	HND	IBE	-0,41912	2,33597	2,28171	0,62295
Hungary	HUN	Eastern Europe	2,52066	-0,35664	-1,31690	-0,07493
Iceland	ISL	Northern Europe	4,31187	-1,66664	0,65368	0,24429
India	IND	Southern Asia	-3,05998	1,27634	-0,01185	-2,08915
Indonesia	IDN	South-Eastern Asia	-1,46041	1,29891	0,66310	-1,81598
Iran	IRN	Southern Asia	0,92060	2,17133	-0,53497	0,43228
Iraq	IRQ	Western Asia	-0,00040	2,46223	-0,90594	0,13842
Ireland	IRL	Northern Europe	3,92086	-1,30595	0,73379	-0,21394
Israel	ISR	Western Asia	3,55544	0,42498	0,78934	0,27003
Italy	ITA	Southern Europe	3,47367	-0,89505	0,53512	-0,46941
Jamaica	JAM	Caribbean	1,10538	1,79218	1,89187	0,26366
Japan	JPN	Eastern Asia	3,81219	-1,28044	0,65729	-1,35008
Jordan	JOR	Western Asia	1,62827	1,79361	-0,27168	0,66370
Kazakhstan	KAZ	Cenrtral Asia	0,43698	1,29828	-3,26303	0,49278
Kenya	KEN	Eastern Africa	-4,33675	-1,44173	0,82968	-0,16760
Korea, Republic of	KOR	Eastern Asia	2,77153	-0,56094	-0,34408	-1,33543
Kuwait	KWT	Western Asia	2,51055	-0,44825	-0,17003	1,93194
Kyrgyzstan	KGZ	Cenrtral Asia	-0,69882	0,92829	-0,77944	-0,98434
Laos	LAO	South-Eastern Asia	-3,40285	-0,35041	0,28126	-1,91887
Latvia	LVA	Northern Europe	2,07250	-0,73935	-2,65345	0,66205
Lebanon	LBN	Western Asia	1,73915	1,92329	-0,84224	0,63389
Lesotho	LSO	Southern Africa	-5,60602	-1,97454	1,42811	1,82880
Liberia	LBR	Western Africa	-6,33718	-0,26432	0,37092	-0,82656
Lithuania	LTU	Northern Europe	2,63625	-0,84182	-1,39084	0,48380
Macedonia	MKD	Southern Europe	1,28995	0,39092	-1,30388	0,96442
Madagascar	MDG	Eastern Africa	-4,31944	0,07073	0,93259	-1,68795
Malawi	MWI	Southern Africa	-5,51708	-2,77985	0,46810	0,12977
Malaysia	MYS	South-Eastern Asia	1,88926	-0,12534	0,68957	0,49280
Mali	MLI	Western Africa	-6,11144	-1,79968	-0,67415	-0,65844
Mauritania	MRT	Western Africa	-4,11429	-0,54816	-0,60508	0,55341
Mauritius	MUS	Eastern Africa	1,60700	-1,08993	-0,00702	0,09305
Mexico	MEX	IBE	2,02824	1,38572	1,19329	0,96792
Moldova	MDA	Eastern Europe	0,44681	1,44644	-2,14935	-0,36019
Mongolia	MNG	Eastern Asia	-1,68782	-0,10610	-1,83626	-0,64698
Montenegro	MNE	Southern Europe	1,57108	1,46682	0,19300	0,19638
Morocco	MAR	Northern Africa	-0,09713	0,69086	0,06445	-0,28077
Mozambique	MOZ	Eastern Africa	-6,35905	-2,27938	-0,15028	0,37474
Namibia	NAM	Southern Africa	-2,40990	-0,51565	1,39162	1,51251
Nepal	NPL	Southern Asia	-3,17985	0,06654	0,63743	-1,87413
Netherlands	NLD	Western Europe	4,05741	-2,19565	-0,23706	-0,92118
New Zealand	NZL	Australia and New Zealand	4,43838	-1,84129	0,75432	-0,05390
Nicaragua	NIC	IBE	-0,35621	2,06076	2,27532	-0,28970
Niger	NER	Western Africa	-5,88683	-1,06174	-0,32341	-1,10274
Nigeria	NGA	Western Africa	-5,85071	0,00962	-0,68382	0,43873
Norway	NOR	Northern	4,86118	-2,21757	0,34335	-1,08614

		Europe				
Pakistan	PAK	Southern Asia	-4,22648	2,71578	0,25954	-1,14431
Panama	PAN	IBE	1,88766	1,10679	2,63627	0,67272
Paraguay	PRY	IBE	-0,35178	1,21308	1,23149	-0,00265
Peru	PER	IBE	0,96655	1,20705	1,58925	-0,10347
Philippines	PHL	South-Eastern Asia	-0,99654	1,40684	0,87577	-1,12645
Poland	POL	Eastern Europe	2,95832	-0,87048	-0,57647	-0,14867
Portugal	PRT	IBE	3,60743	-0,96603	-0,39263	0,33612
Romania	ROU	Eastern Europe	1,52950	-0,13419	-1,73060	-0,64751
Russia	RUS	Eastern Europe	0,71808	0,92303	-2,52519	1,52909
Rwanda	RWA	Eastern Africa	-4,07184	-2,08644	0,95638	-0,12690
Saudi Arabia	SAU	Western Asia	2,00983	1,59192	0,02310	1,09349
Senegal	SEN	Western Africa	-3,32670	-0,20282	0,34459	0,13219
Serbia	SRB	Southern Europe	2,04340	0,50302	-0,95134	-0,40177
Slovakia	SVK	Eastern Europe	3,34902	-1,47713	-0,79073	-0,37106
Slovenia	SVN	Southern Europe	3,80527	-1,07918	-0,19100	-0,50979
South Africa	ZAF	Southern Africa	-1,53049	-0,23859	0,59883	4,24607
Spain	ESP	IBE	3,73870	-0,78625	0,36091	-0,21720
Sri Lanka	LKA	Southern Asia	-0,05860	0,89668	0,82986	-1,79202
Sudan	SDN	Northern Africa	-4,76396	1,82397	-0,48020	-0,68989
Swaziland	SWZ	Southern Africa	-3,49837	-1,12218	0,76771	2,35417
Sweden	SWE	Northern Europe	4,63984	-2,24704	0,07955	-1,17643
Switzerland	CHE	Western Europe	3,66058	-1,04751	1,01581	-1,36596
Tajikistan	TJK	Central Asia	-1,75738	1,29744	-0,36601	-1,81522
Tanzania	TZA	Eastern Africa	-4,79334	-1,81966	-0,19741	-0,50584
Thailand	THA	South-Eastern Asia	1,22002	0,08129	0,84119	-0,94297
Togo	TGO	Western Africa	-5,19416	-0,92992	0,08922	0,16302
Trinidad and Tobago	TTO	Caribbean	1,13890	-0,52261	-0,16880	1,12822
Tunisia	TUN	Northern Africa	1,72105	0,89998	0,41810	-0,07282
Turkey	TUR	Western Asia	2,08737	1,36822	-0,45113	1,08832
Uganda	UGA	Eastern Africa	-5,02274	-1,03456	-0,03281	-0,15733
Ukraine	UKR	Eastern Europe	1,16057	1,64340	-4,26626	0,24094
United Arab Emirates	ARE	Western Asia	3,08020	-1,37624	-0,03986	1,09657
United Kingdom	GBR	Northern Europe	4,34919	-1,46615	0,30024	0,28660
United States	USA	Northern America	3,26910	-0,78381	-0,11021	0,76649
Uruguay	URY	IBE	2,47177	0,37593	0,09116	0,83740
Uzbekistan	UZB	Central Asia	-0,30992	1,49586	-0,74711	-0,35671
Venezuela	VEM	IBE	1,21492	3,05134	1,53601	0,52593
Yemen	YEM	Western Asia	-3,09022	1,89436	-0,16888	-0,63340
Zambia	ZMB	Middle Africa	-4,92392	-1,73444	0,58594	1,31498

Fonte: Elaborado pela autora

3.12 ANÁLISE DE COMPONENTES PRINCIPAIS

Na análise dos componentes principais, com base nas 20 variáveis pré-selecionadas para os países em questão, uma vez criadas as 4 variáveis PC1, PC2, PC3 e PC4, pôde-se perceber quais dos agrupamentos é melhor representativo para explicar as demais variáveis (Tabela 1).

Tabela 21: Indica o peso de cada variável nas componentes PC1, PC2, PC3 e PC4.

Variable	PC1	PC2	PC3	PC4
N_P_SPI_Regressao	0,274	-0,176	0,073	-0,024
N_P_Maternal mortality	0,258	0,133	-0,034	-0,062
N_P_Stillbirth rate	0,253	-0,036	-0,054	-0,004
N_P_Child mortality	0,271	0,092	-0,006	-0,069
N_P_Deaths from	0,259	0,212	-0,078	-0,229
N_P_Access to piped water	0,272	0,062	0,008	0,169
N_P_Access to improved	0,263	0,087	-0,081	0,069
N_P_Indoor air pollution	0,232	0,005	0,049	0,251
N_P_Life expectancy (years)_1	0,273	0,073	0,044	-0,262
N_P Non-communicable disease	0,216	-0,100	0,363	-0,119
N_P Obesity rate (% of pop.)_1	-0,208	-0,164	0,042	-0,501
N_P Outdoor air pollution	-0,068	-0,211	0,613	-0,180
N_P Health expenditure	0,165	-0,299	0,075	0,162
N_P External resources	-0,158	-0,458	0,023	-0,069
N_P IDH - 2013_1	0,270	-0,028	-0,003	0,088
N_P GovernanÁa_1	0,216	-0,395	0,089	-0,028
N_P EPI Score_1	0,263	-0,185	-0,012	0,027
N_P Happy Planet Index_1	0,117	0,357	0,425	-0,306
N_P OHI_1	-0,144	0,419	0,210	0,022
N_P GINNI Index_1	0,109	-0,071	-0,475	-0,584

Fonte: Elaborado pela autora

Análise:

- Analisando PC1. Há 4 variáveis com contribuição negativa: Obesity rate (% of pop.), *Outdoor air pollution*, *External resources* e OHI_1. Uma delas próximo a zero: *Outdoor air pollution*. E o resto das variáveis que variam de 0,109 a 0,274.
- Analisando PC2. Há 11 variáveis com contribuição negativa: *Social Progress*, *Stillbirth rate*, *Non-communicable disease*, *Obesity rate*, *Outdoor air pollution*, *Health expenditure*, *External resources*, IDH, Governança, EPI *Score*, GINNI *Index*; 8 variáveis se aproximam de zero: *Stillbirth rate*, *Child mortality*, *Access to piped water*, *Access to improved*, *Indoor air pollution*, *Life expectancy*, *Obesity rate*, IDH, GINNI; as com contribuição positiva acima de 0,1 variam de 0,133 a 0,419.
- Analisando PC3. Há 8 variáveis com contribuição negativa,15 variáveis com contribuição próxima a zero, e postivias acima de 0,1, tem-se 5 com contribuição entre 0,210 e 0,613.
- Analisando PC4. Há 13 variáveis com contribuição negativa, 10 com contribuição próxima a zero, e as demais variam entre 0,162 e 0,251.

Tabela 22. Matriz de Correlações *Eigenanalysis*

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Eigenvalue	11,658	1,983	1,479	0,972	0,647	0,585	0,508	0,389	0,375
Proportion	0,583	0,099	0,074	0,049	0,032	0,029	0,025	0,019	0,019
Cumulative	0,583	0,682	0,756	0,805	0,837	0,866	0,892	0,911	0,930

	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18
Eigenvalue	0,264	0,240	0,208	0,164	0,135	0,114	0,089	0,083	0,058
Proportion	0,013	0,012	0,010	0,008	0,007	0,006	0,004	0,004	0,003
Cumulative	0,943	0,955	0,965	0,974	0,980	0,986	0,990	0,995	0,997

	PC19	PC20
Eigenvalue	0,029	0,023
Proportion	0,001	0,001
Cumulative	0,999	1,000

Fonte: Elaborado pela autora

Análise:

Segundo Las Casas e Guevara (2010)[2], por convenção só se deve considerar componentes com contribuição >= 1, pois só acima ou igual a l é que ajuda a explicar o fenômeno. Neste caso, só PC1, PC2 e PC3 têm Eigenvalue que atendem a esse critério. Se se inclui o PC4 por uma questão de ampliar o espectro de análise, apesar de se perceber que não se altera significativamente (todos os cálculos foram feitos dessa forma anteriormente), também tais agrupamentos explicam, cumulativamente 80,5% das variáveis. Se se observa o peso de explicação de PC1 é de 58,3%. Quando se passa para o grupo de variáveis de PC2, há um acréscimo de apenas 9,9% de poder de explicação; de PC2 para PC3, há um acúmulo acrescido de apenas 7,4% de poder de explicação, com o qual, essas primeiras variáveis são suficientes para explicar os indicadores em questão de Saúde e Meio Ambiente dos Países pesquisados para os 20 indicadores pré-selecionados. Para comprovar que os PC's sào complementares, calculamos a correlação entre eles. Ver tabela 3.

Tabela 23. Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	0,000		
	1,000		
PC3	-0,000	0,000	
	1,000	1,000	
PC4	0,000	-0,000	-0,000
	1,000	1,000	1,000

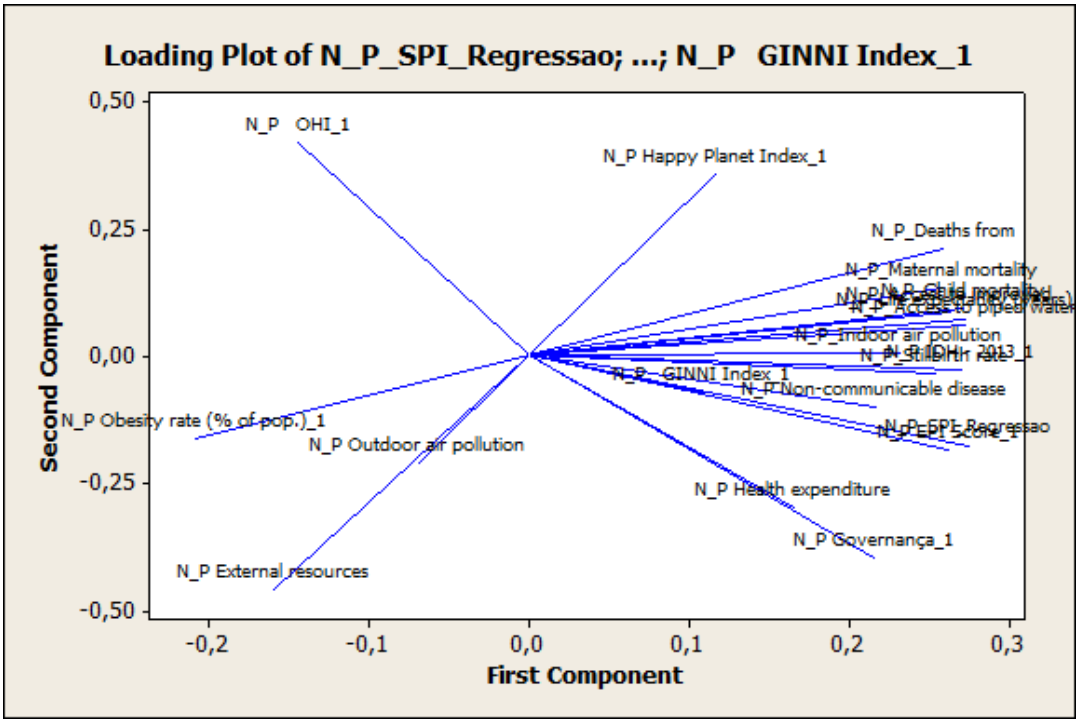
Fonte: Elaborado pela autora

```
Cell Contents: Pearson correlation
               P-Value
```

Análise:

Correlação = 0 (nula), o que significa que cada coisa está coisa diferente explicando. São explicações complementares, conforme esperado.

Gráfico 103.



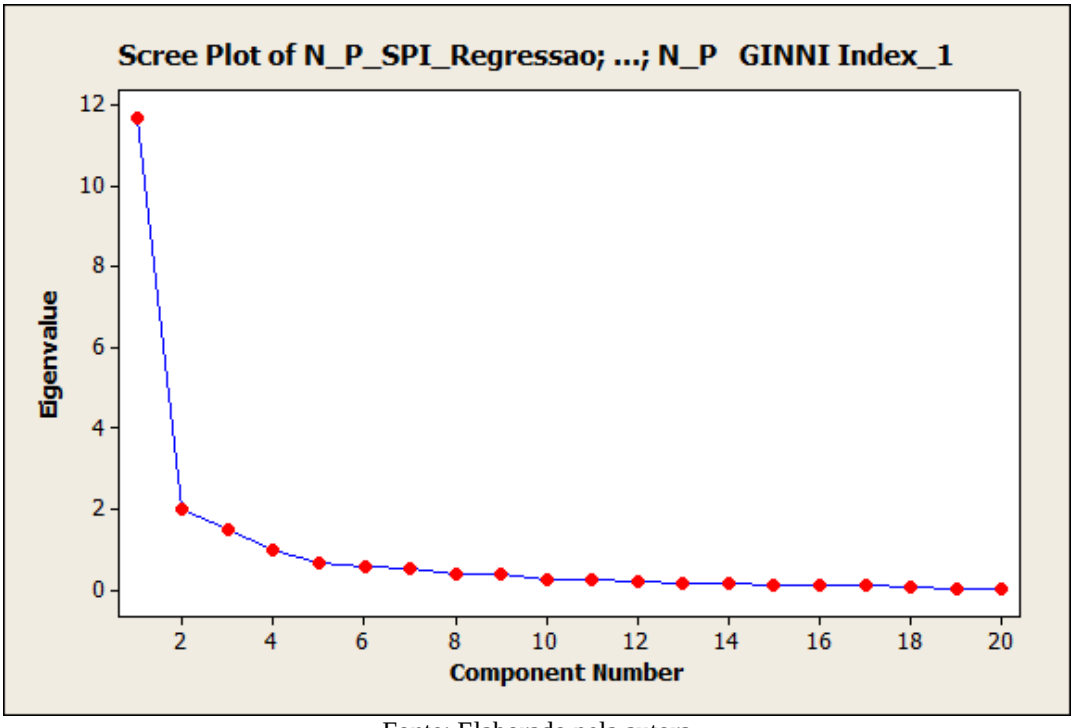
Fonte: Elaborado pela autora

Análise:

A primeira componente complementa a segunda componente e vice-versa, mas não se correlacionam.

Percebe-se um grupo de variáveis que mesmo depois de positivada continua negativa, devido ao contraste. São elas as referentes ao meio ambiente e a obesidade. E do outro lado, direito, mais indicadores e variáveis de natureza mais voltada para à saúde de forma mais direta, além de indicadores financeiros.

Gráfico 104.



Fonte: Elaborado pela autora

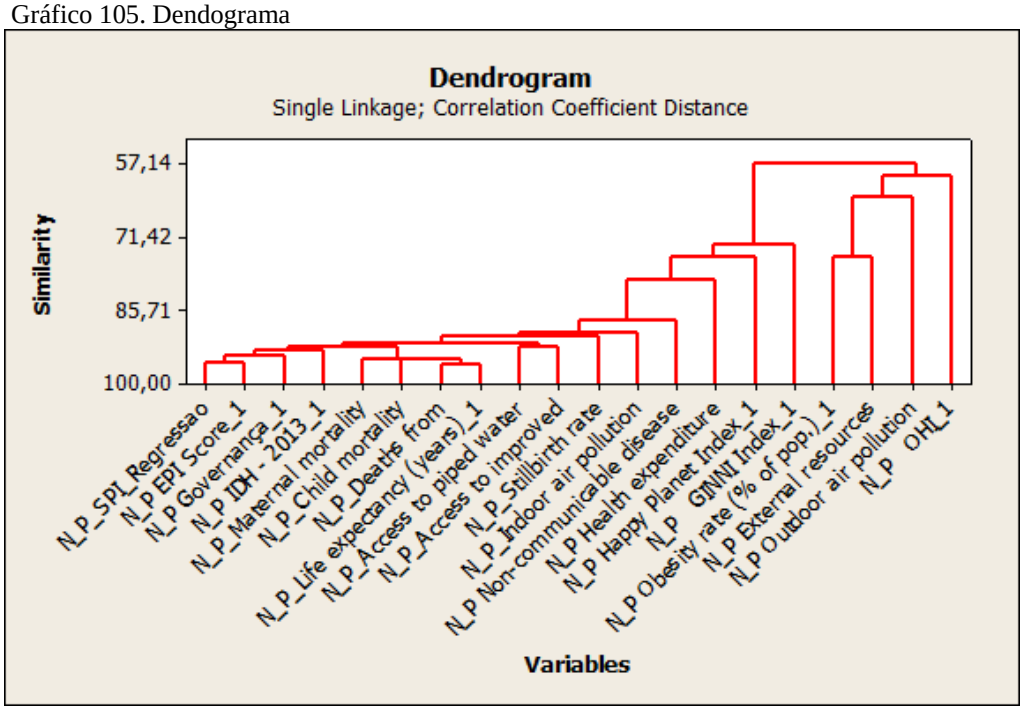
Análise:
PELO Stepwise feito a seguir serão descobertos quais itens correspondem às 4 primeiras bolas vermelhas, da esquerda para a direita, que respondem fundamentalmente pelo Modelo.

Cluster Analysis of Variables: N_P_SPI_Regr; N_P_Maternal; N_P_Stillbir; ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	19	96,2149	0,075703	5	9	5	2
2	18	95,8701	0,082598	1	17	1	2
3	17	95,2932	0,094137	4	5	4	3
4	16	95,0342	0,099316	2	4	2	4
5	15	94,4907	0,110187	1	16	1	3
6	14	93,3759	0,132483	1	15	1	4
7	13	92,8215	0,143569	6	7	6	2
8	12	92,7700	0,144600	1	2	1	8
9	11	92,2870	0,154260	1	6	1	10
10	10	90,6406	0,187187	1	3	1	11
11	9	90,1264	0,197472	1	8	1	12
12	8	87,8186	0,243628	1	10	1	13
13	7	79,7595	0,404811	1	13	1	14
14	6	75,4511	0,490979	11	14	11	2
15	5	75,4142	0,491717	1	18	1	15
16	4	72,8474	0,543051	1	20	1	16
17	3	63,7578	0,724844	11	12	11	3
18	2	59,7293	0,805415	11	19	11	4
19	1	57,1374	0,857251	1	11	1	20

Gráfico 105. Dendograma



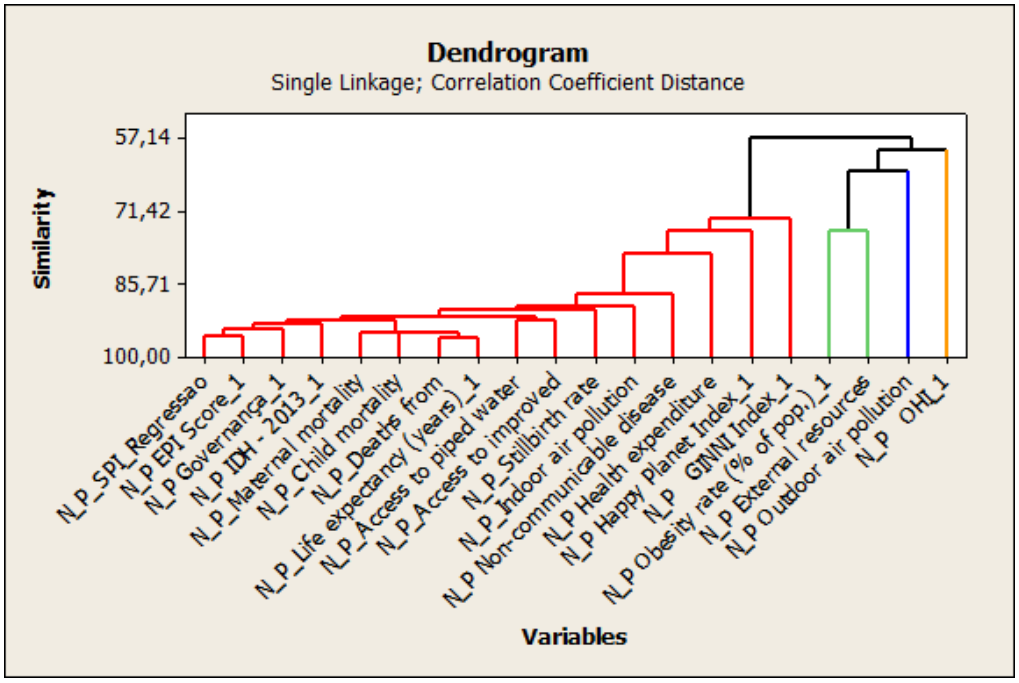
Fonte: Elaborado pela autora

Reagrupando em 4 clusters:

Cluster Analysis of Variables: N_P_SPI_Regr; N_P_Maternal; N_P_Stillbir; ...

```
Correlation Coefficient Distance, Single Linkage
Amalgamation Steps
                                                                 Number
                                                                of obs.
      Number of  Similarity  Distance  Clusters      New   in new
Step   clusters       level     level   joined   cluster  cluster
   1         19     96,2149  0,075703   5     9        5        2
   2         18     95,8701  0,082598   1    17        1        2
   3         17     95,2932  0,094137   4     5        4        3
   4         16     95,0342  0,099316   2     4        2        4
   5         15     94,4907  0,110187   1    16        1        3
   6         14     93,3759  0,132483   1    15        1        4
   7         13     92,8215  0,143569   6     7        6        2
   8         12     92,7700  0,144600   1     2        1        8
   9         11     92,2870  0,154260   1     6        1       10
  10         10     90,6406  0,187187   1     3        1       11
  11          9     90,1264  0,197472   1     8        1       12
  12          8     87,8186  0,243628   1    10        1       13
  13          7     79,7595  0,404811   1    13        1       14
  14          6     75,4511  0,490979  11    14       11        2
  15          5     75,4142  0,491717   1    18        1       15
  16          4     72,8474  0,543051   1    20        1       16
  17          3     63,7578  0,724844  11    12       11        3
  18          2     59,7293  0,805415  11    19       11        4
  19          1     57,1374  0,857251   1    11        1       20


Final Partition

Cluster 1
  N_P_SPI_Regressao  N_P_Maternal mortality  N_P_Stillbirth rate  N_P_Child
     mortality  N_P_Deaths from  N_P_Access to piped water  N_P_Access to
     improved  N_P_Indoor air pollution  N_P_Life expectancy (years)_1  N_P
     Non-communicable disease  N_P Health expenditure  N_P IDH - 2013_1  N_P
     GovernanÁa_1  N_P EPI Score_1  N_P Happy Planet Index_1  N_P   GINNI
     Index_1
Cluster 2
  N_P Obesity rate (% of pop.)_1  N_P External resources
Cluster 3
  N_P Outdoor air pollution
Cluster 4
  N_P   OHI_1
```

Gráfico 106.

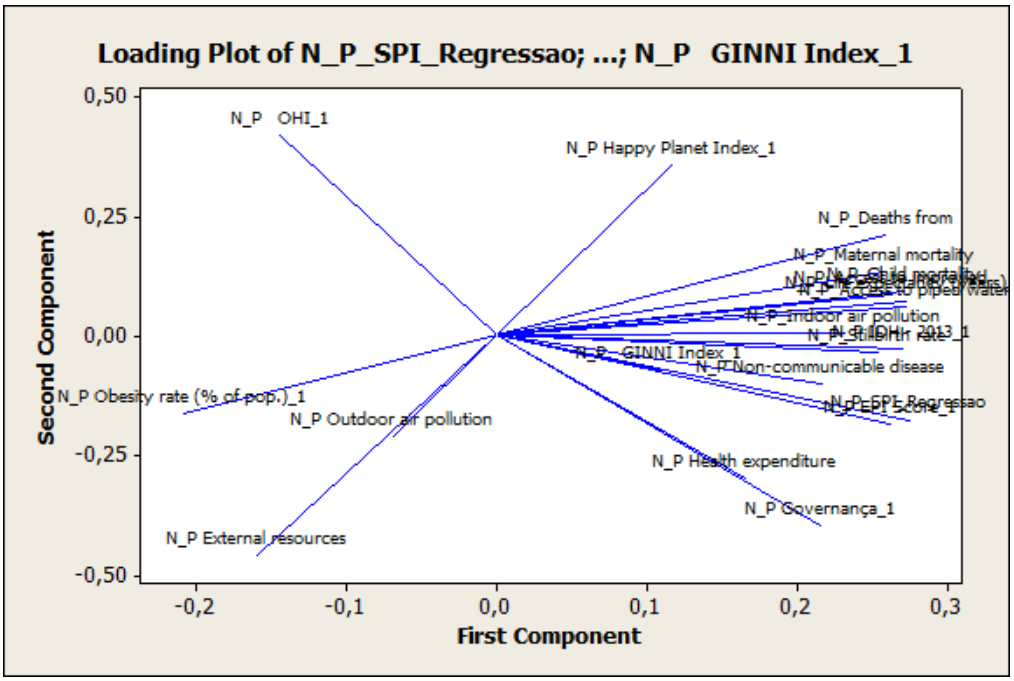


Fonte: Elaborado pela autora

Análise Gráficos 106 e 107:

Obesidade (no verde) está interferindo no quadro geral, indicando que vai puxar o indicador pra baixo. Explicação: pelos contrastes. Tal indicador puxa para o lado esquerdo, no sentido inverso. Gerando os 4 PC's, relembrando o que já foi exposto neste trabalho que do lado esquerdo do gráfico abaixo havia justamente os mesmos indicadores acima: *Obesity, External Resources, Outdoor Pollution* e OHL.

Gráfico 107.



Fonte: Elaborado pela autora

Próximos passos: se vai proceder *Stepwise Regression* em cada componente PC1, PC2, PC3 e PC4. E – criar um novo nome.

Stepwise PC1

```
Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15
Response is PC1 on 20 predictors, with N = 132
```

Tabela 24. Stepwise Regression PC1 versus N_P_SPI_Regressa; N_P_Maternal mor; ...

Step	1	2	3	4	5	6
Constant	-14,33	-13,91	-12,00	-11,09	-11,71	-11,98
N_P_SPI_Regressao	0,2251	0,1501	0,1103	0,0796	0,0776	0,0708
T-Value	30,41	21,89	16,51	13,08	14,00	13,62
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
N_P_Deaths from		0,0544	0,0391	0,0339	0,0346	0,0317
T-Value		14,64	12,03	13,22	14,82	14,49
P-Value		0,000	0,000	0,000	0,000	0,000
N_P_Access to piped water			0,0293	0,0244	0,0179	0,0162
T-Value			9,62	10,13	7,15	7,07
P-Value			0,000	0,000	0,000	0,000
S	1,20	0,741	0,566	0,436	0,396	0,360
R-Sq	87,68	95,37	97,31	98,42	98,70	98,94
R-Sq(adj)	87,58	95,30	97,25	98,37	98,65	98,89

Fonte: Elaborado pela autora

Análise:

Bom, sabe-se portanto que a primeira bola vermelha do gráfico 104 corresponde a *Social Progress*. As variáveis seguintes que aparecem incrementam bem menos à componente PC1: *Deaths from infectious desease*, *Access to piped water* e IDH. A *Social progress* explica 87,68%, depois há um acréscimo de 7,69%. Depois, o acréscimo é de apenas 1,94%. Assim os dois primeiros são suficientes como variáveis explicativas do PC1. Observa-se que PC1 explica sozinho (Social Progress e *Deaths from infectious desease*) mais de 95,37% da variabilidade, sendo as demais variáveis componentes de baixa contribuição. Depois de fazer a Regressão Múltipla, recomenda-se batizar o nome da componente.

Stepwise Regression: PC2 versus N_P_SPI_Regressa; N_P_Maternal mor; ...

```
Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15
Response is PC2 on 20 predictors, with N = 132
```

Tabela 25. Stepwise PC2

Step	19	20
Constant	1,084	1,069
N_P External resources	-0,02239	-0,02238
T-Value	-1215,05	*
P-Value	0,000	*
N_P GovernanÁa_1	-0,01552	-0,01553
T-Value	-498,39	*
P-Value	0,000	*
N_P Happy Planet Index_1	0,01658	0,01658
T-Value	1027,00	*
P-Value	0,000	*
S	1,08	0,664	0,515	0,452	0,356	0,279
R-Sq	41,66	78,12	86,91	90,00	93,85	96,25
R-Sq(adj)	41,21	77,78	86,60	89,68	93,61	96,07

Fonte: Elaborado pela autora

Análise:
As seguintes bolas vermelhas do Gráfico 104 – depois do Social Progress são, portanto: External Resources, Governança e Happy Planet, que explicam 86,91% – ver a linha amarela - são os números no “Cumulative” na tabela.

Há de se chamar atenção que *Happy Planet* influencia aqui. Esse novo indicador PC2, com dois indicadores negativos e todos próximos a zero, têm um poder de explicação baixo de variabilidade. Poderia ser batizado com o nome de PC2: **Índice de Planeta Viável**, incluindo recursos externos, governança e *happy planet*, depois da Regressão Múltipla.

Tabela 26. *Stepwise* PC3

Step	19	20			
Constant	-3,585	-3,584			
N_P Outdoor air pollution	0,03242	0,03242			
T-Value	5499,33	*			
P-Value	0,000	*			
N_P Happy Planet Index_1	0,01971	0,01971			
T-Value	3200,57	*			
P-Value	0,000	*			
N_P GINNI Index_1	-0,02248	-0,02248			
T-Value	-3786,79	*			
P-Value	0,000	*			
S	0,814	0,585	0,377	0,220	0,158
R-Sq	55,50	77,22	90,61	96,83	98,38
R-Sq(adj)	55,16	76,87	90,39	96,73	98,31

Fonte: Elaborado pela autora

Análise:
Happy Planet está influenciando bastante no PC3 também. Porém com grau de explicação baixo, próximos a zero, com um valor negativo de Ginni.

Tabela 27. *Stepwise PC4*

Step	17	18			
Constant	4,756	4,750			
N_P GINNI Index_1	-0,02754	-0,02760			
T-Value	-448,68	-850,64			
P-Value	0,000	0,000			
N_P Obesity rate (% of pop.)_1	-0,02223	-0,02215			
T-Value	-313,61	-590,01			
P-Value	0,000	0,000			
N_P Happy Planet Index_1	-0,01432	-0,01420			
T-Value	-232,31	-427,23			
P-Value	0,000	0,000			
S	0,810	0,599	0,376	0,332	0,276
R-Sq	33,11	63,70	85,81	89,02	92,49
R-Sq(adj)	32,60	63,13	85,47	88,67	92,19

Fonte: Elaborado pela autora

Análise:
No PC4, aparece também obesidade, de forma negativa, próximo a zero.

Antes de batizar as Componentes Principais é recomendado, logo após realizadas as Regressões Stepwise, rodar uma Regressão Multiple para cada componente utilizando só as variáveis selecionadas no Stepwise. Dessa forma se pode avaliar o peso com o qual cada variável entra na Componente Principal.

Regressão PC1
Social Progress, Deaths from infectious desease, Access to piped water

```
The regression equation is
PC1 = - 12,0 + 0,110 N_P_SPI_Regressao + 0,0391 N_P_Deaths from
      + 0,0293 N_P_Access to piped water
```

Tabela 28.

Predictor	Coef	SE Coef	T	P
Constant	-12,0011	0,3023	-39,70	0,000
N_P_SPI_Regressao	0,110311	0,006680	16,51	0,000
N_P_Deaths from	0,039140	0,003254	12,03	0,000
N_P_Access to piped water	0,029310	0,003046	9,62	0,000

Fonte: Elaborado pela autora

S = 0,566290 R-Sq = 97,3% R-Sq(adj) = 97,2%

Tabela 29.

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	3	1486,15	495,38	1544,76	0,000
Residual Error	128	41,05	0,32		
Total	131	1527,19			

Source	DF	Seq SS
N_P_SPI_Regressao	1	1339,00
N_P_Deaths from	1	117,46
N_P_Access to piped water	1	29,69

Fonte: Elaborado pela autora

Tabela 30.

Unusual Observations						
Obs	N_P_SPI_Regressao	PC1	Fit	SE Fit	Residual	St Resid
31	61,1	2,3322	0,7531	0,0804	1,5791	2,82R
49	61,3	-0,4191	0,7753	0,0915	-1,1944	-2,14R
55	44,8	-0,0004	-1,4748	0,1596	1,4744	2,71R
70	48,9	-5,6060	-6,0500	0,1913	0,4440	0,83 X
79	73,7	1,6070	2,8338	0,0750	-1,2268	-2,19R
104	64,4	2,0098	0,7083	0,0639	1,3015	2,31R
109	63,0	-1,5305	-2,2066	0,2022	0,6761	1,28 X
113	48,9	-3,4984	-5,4249	0,1961	1,9265	3,63RX
118	65,1	1,2200	0,0059	0,0886	1,2141	2,17R

Fonte: Elaborado pela autora

```
Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

General Regression Analysis: PC1 versus N_P_SPI_Regr; N_P_Deaths f; ...

```
Regression Equation
PC1  =  -12,0011 + 0,110311 N_P_SPI_Regressao + 0,0391396 N_P_Deaths from +
        0,0293102 N_P_Access to piped water
```

Tabela 31.

Coefficients				
Term	Coef	SE Coef	T	P
Constant	-12,0011	0,302273	-39,7030	0,000
N_P_SPI_Regressao	0,1103	0,006680	16,5142	0,000
N_P_Deaths from	0,0391	0,003254	12,0300	0,000
N_P_Access to piped water	0,0293	0,003046	9,6220	0,000

Fonte: Elaborado pela autora

Summary of Model

S = 0,566290 R-Sq = 97,31% R-Sq(adj) = 97,25%
PRESS = 44,5924 R-Sq(pred) = 97,08%

Tabela 32.

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	1486,15	1486,15	495,382	1544,76	0,0000000
N_P_SPI_Regressao	1	1339,00	87,46	87,457	272,72	0,0000000
N_P_Deaths from	1	117,46	46,41	46,410	144,72	0,0000000
N_P_Access to piped water	1	29,69	29,69	29,690	92,58	0,0000000
Error	128	41,05	41,05	0,321		
Total	131	1527,19				

Fonte: Elaborado pela autora

Tabela 33.

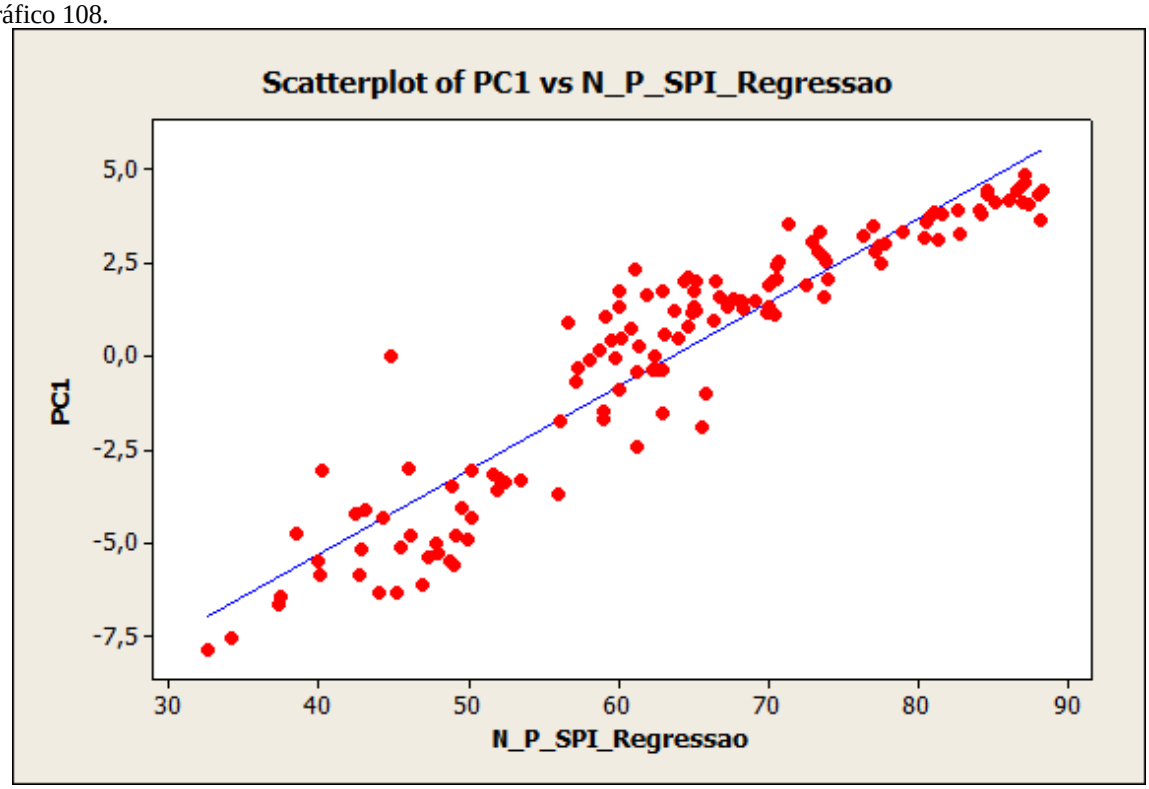
Fits and Diagnostics for Unusual Observations

Obs	PC1	Fit	SE Fit	Residual	St Resid		
31	2,33218	0,75312	0,080411	1,57906	2,81698	R	
49	-0,41912	0,77527	0,091536	-1,19438	-2,13724	R	
55	-0,00040	-1,47477	0,159567	1,47436	2,71349	R	
70	-5,60602	-6,05000	0,191288	0,44398	0,83297		X
79	1,60700	2,83380	0,075007	-1,22680	-2,18563	R	
104	2,00983	0,70833	0,063915	1,30150	2,31307	R	
109	-1,53049	-2,20656	0,202237	0,67607	1,27813		X
113	-3,49837	-5,42488	0,196138	1,92651	3,62645	R	X
118	1,22002	0,00591	0,088639	1,21410	2,17071	R	

Fonte: Elaborado pela autora

Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Gráfico 108.



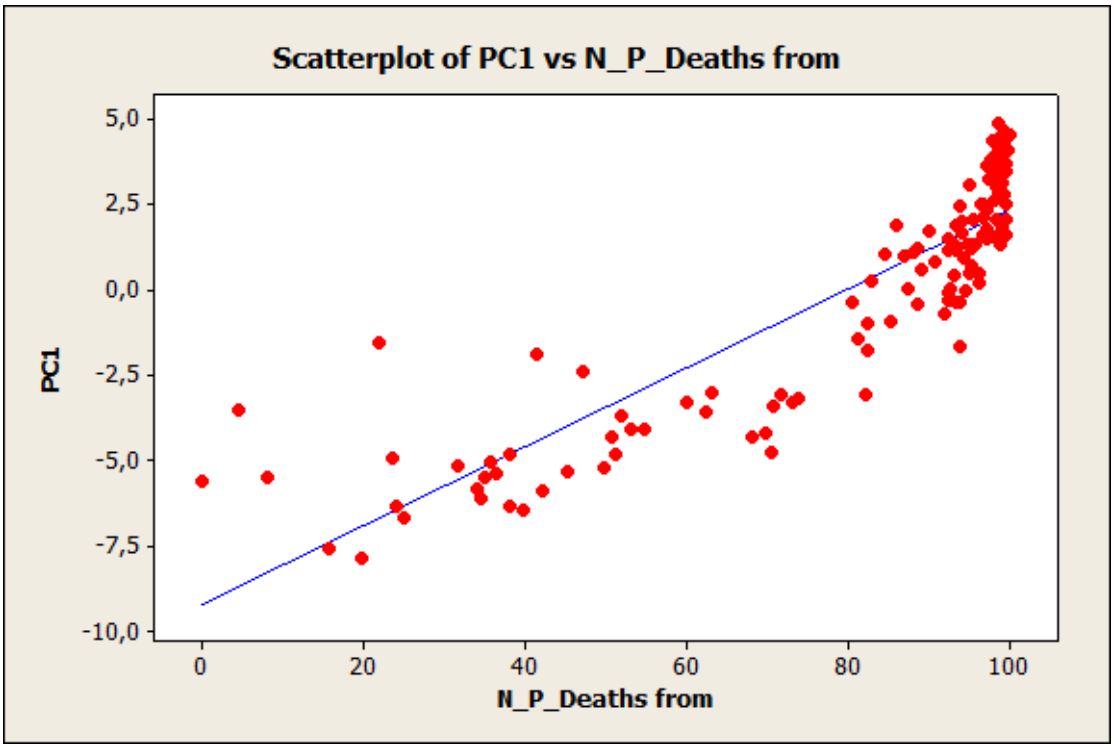
Fonte: Elaborado pela autora

Análise:

O peso de importância de SPI é muito alto, com r^2 de 87,46%, inclusive mais alto que os outros abaixo, e é um indicador muito complexo, difícil de se trabalhar. Portanto, recomenda-

se que se use em substituição o IDH para equilibrar melhor os pesos dos indicadores e facilitar análises futuras.

Gráfico 109.

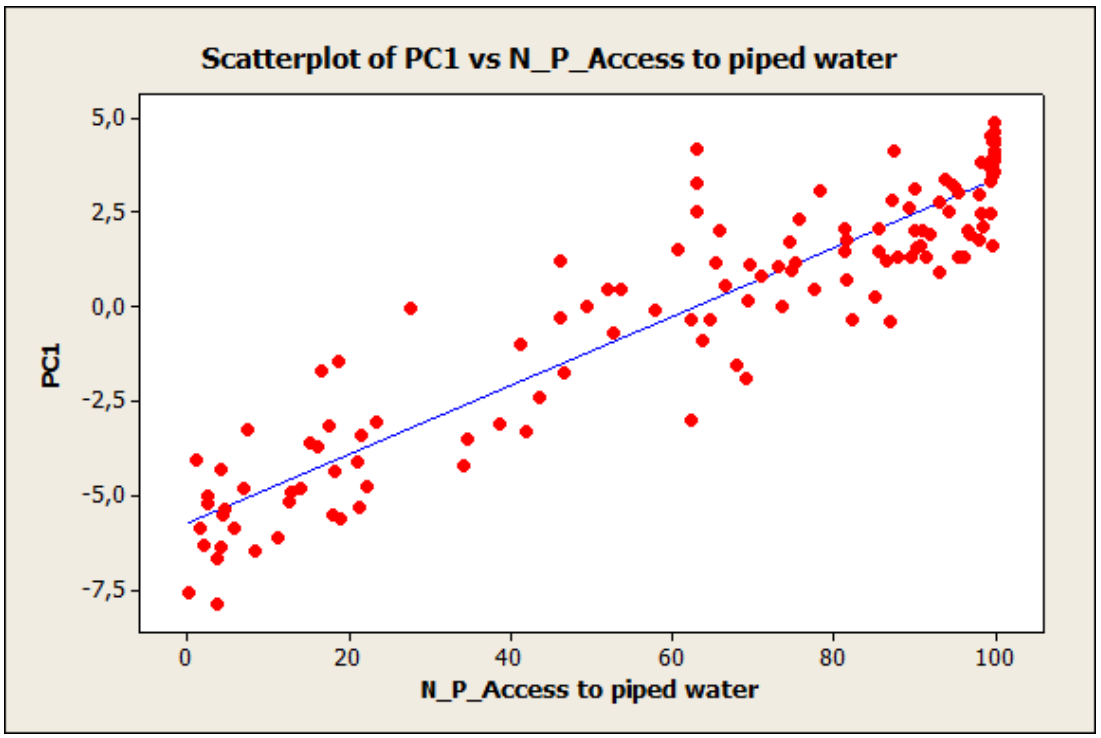


Fonte: Elaborado pela autora

Análise:

Deaths from Infectious Deasease tem um R^2 de 46,41, baixo, com relação positiva. Essa variável sozinha tem um peso baixo de explicação de PC1.

Gráfico 110.



Fonte: Elaborado pela autora

Análise:

- Access to piped water tem um R^2 de 29,69, baixo, com relação positiva. Essa variável sozinha tem um peso muito baixo de explicação da variabilidade de PC1.
- Todos os P = 0, o que era necessário acontecer.
- O peso de Deaths from Infectious Desease é de 3 x, e o de Access to piped water é de 4 x.

Regressão PC2

PC2 nas variáveis External Resources, Governança e Happy Planet.

Regression Equation

PC2 = 1,20802 - 0,0379871 N_P External resources - 0,0345507 N_P GovernanÁa_1 + 0,0213721 N_P Happy Planet Index_1

Tabela 34.

Coefficients

Term	Coef	SE Coef	T	P
Constant	1,20802	0,161202	7,4938	0,000
N_P External resources	-0,03799	0,002424	-15,6742	0,000
N_P GovernanÁa_1	-0,03455	0,001777	-19,4438	0,000
N_P Happy Planet Index_1	0,02137	0,002306	9,2700	0,000

Summary of Model

S = 0,515483 R-Sq = 86,91% R-Sq(adj) = 86,60%
PRESS = 36,2523 R-Sq(pred) = 86,04%

Fonte: Elaborado pela autora

Tabela 35.

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	225,760	225,760	75,253	283,203	0,0000000
N_P External resources	1	108,223	65,283	65,283	245,682	0,0000000
N_P GovernanÁa_1	1	94,703	100,459	100,459	378,060	0,0000000
N_P Happy Planet Index_1	1	22,834	22,834	22,834	85,932	0,0000000
Error	128	34,012	34,012	0,266		
Total	131	259,773				

Fonte: Elaborado pela autora

Tabela 36.

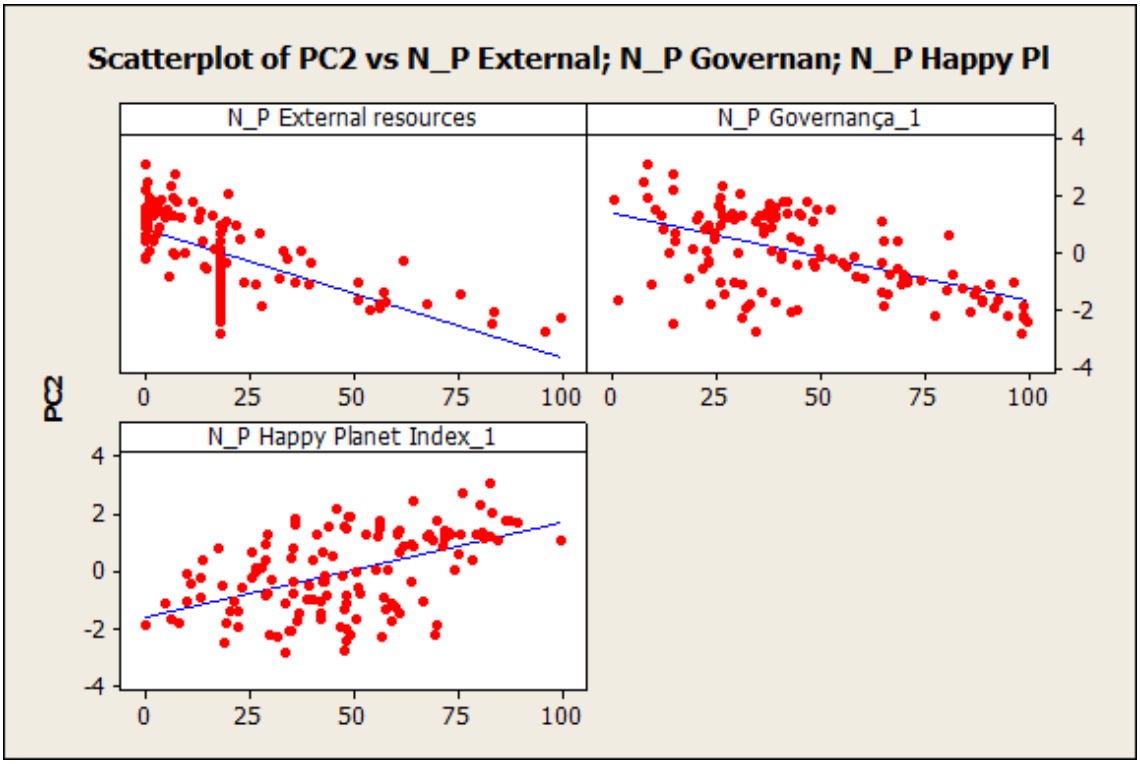
Fits and Diagnostics for Unusual Observations

Obs	PC2	Fit	SE Fit	Residual	St Resid		
2	0,86331	1,95662	0,082590	-1,09331	-2,14870	R	
5	1,76578	0,55227	0,049338	1,21351	2,36497	R	
19	-2,48343	-2,03341	0,156463	-0,45003	-0,91624		X
23	-1,66070	-0,63758	0,126863	-1,02312	-2,04776	R	
24	-1,11574	-0,00720	0,117814	-1,10854	-2,20896	R	
42	1,50677	0,39418	0,054283	1,11259	2,17041	R	
71	-0,26432	-1,29719	0,111615	1,03288	2,05239	R	
75	-2,77985	-2,60137	0,194344	-0,17847	-0,37381		X
85	-2,27938	-2,97924	0,191478	0,69986	1,46231		X
118	0,08129	1,43467	0,074363	-1,35338	-2,65322	R	

Fonte: Elaborado pela autora

Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Gráficos (Conjunto) 111.



Fonte: Elaborado pela autora

Análise: Curiosamente, entrou Ginni – desigualdade nos Estados Unidos (AVECO), laranja, entre os melhores.

Regressão PC3 nas variáveis:
Outdoor Pollution, Happy Planet e Ginni, segundo o Stepwise.

Tabela 37.

Regression Equation				
PC3 = -2,83456 + 0,0388839 N_P Outdoor air pollution + 0,0256764 N_P Happy Planet Index_1 - 0,0219867 N_P GINNI Index_1				
Coefficients				
Term	**Coef**	**SE Coef**	**T**	**P**
Constant	-2,83456	0,209380	-13,5379	0,000
N_P Outdoor air pollution	0,03888	0,001821	21,3511	0,000
N_P Happy Planet Index_1	0,02568	0,001532	16,7604	0,000
N_P GINNI Index_1	-0,02199	0,001627	-13,5144	0,000
Summary of Model				
S = 0,376920 R-Sq = 90,61% R-Sq(adj) = 90,39%				
PRESS = 19,2081 R-Sq(pred) = 90,08%				

Fonte: Elaborado pela autora

Tabela 38.

Analysis of Variance						
Source	**DF**	**Seq SS**	**Adj SS**	**Adj MS**	**F**	**P**
Regression	3	175,533	175,533	58,5111	411,852	0
N_P Outdoor air pollution	1	107,520	64,764	64,7643	455,867	0
N_P Happy Planet Index_1	1	42,066	39,909	39,9087	280,912	0
N_P GINNI Index_1	1	25,947	25,947	25,9470	182,638	0
Error	128	18,185	18,185	0,1421		
Total	131	193,718				

Fonte: Elaborado pela autora

Tabela 39

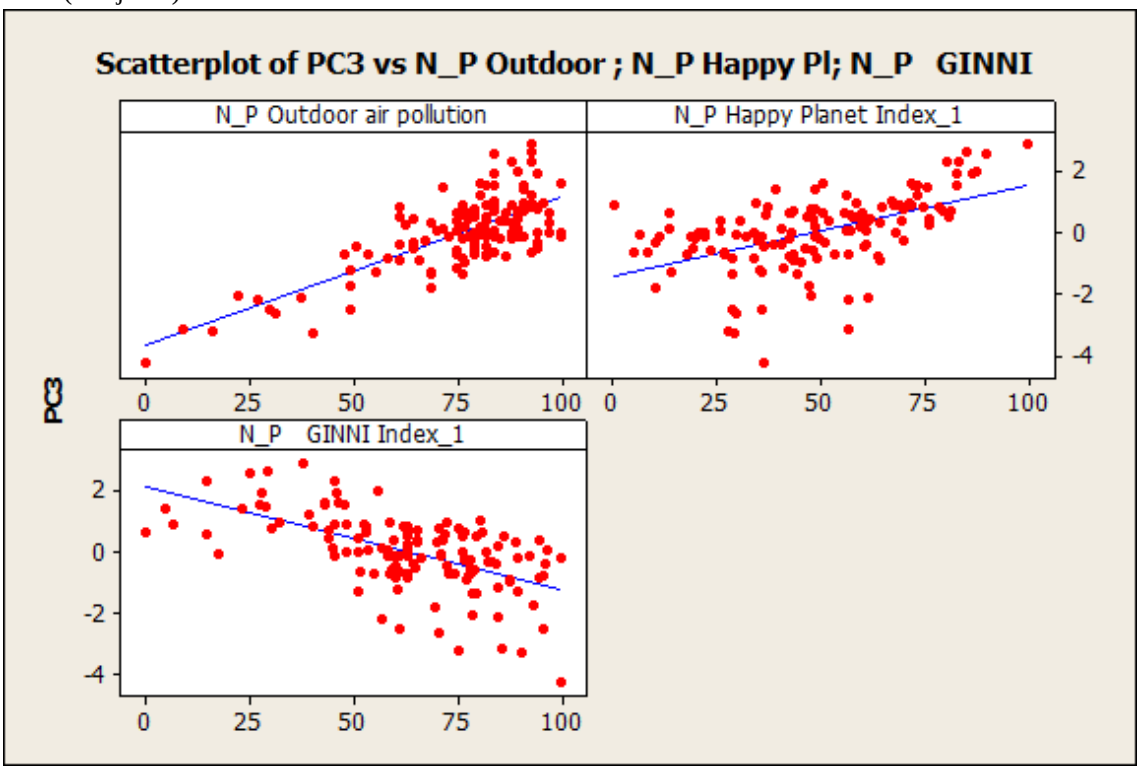
Fits and Diagnostics for Unusual Observations

Obs	PC3	Fit	SE Fit	Residual	St Resid		
5	-3,12608	-2,92015	0,123907	-0,20593	-0,57850		X
8	-1,33242	-0,46898	0,041797	-0,86343	-2,30498	R	
15	0,86190	0,43577	0,121845	0,42613	1,19469		X
60	0,65729	-0,16949	0,047473	0,82678	2,21113	R	
62	-3,26303	-2,50124	0,079350	-0,76179	-2,06743	R	
66	-0,77944	0,29164	0,049606	-1,07108	-2,86661	R	
67	0,28126	1,03735	0,057902	-0,75609	-2,03007	R	
109	0,59883	0,53017	0,121155	0,06867	0,19239		X
116	-0,36601	0,52550	0,065493	-0,89151	-2,40179	R	
124	-4,26626	-4,10443	0,139080	-0,16183	-0,46195		X
129	-0,74711	0,20417	0,041102	-0,95128	-2,53897	R	

Fonte: Elaborado pela autora

```
Legenda:
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Gráficos (Conjunto) 112.



Fonte: Elaborado pela autora

Os Novos Componentes foram batizados em: PC1 é o componente de Desenvolvimento Humano, o PC2 é de Planeta Viável e o PC3 é o de Ambiente Sustentável.

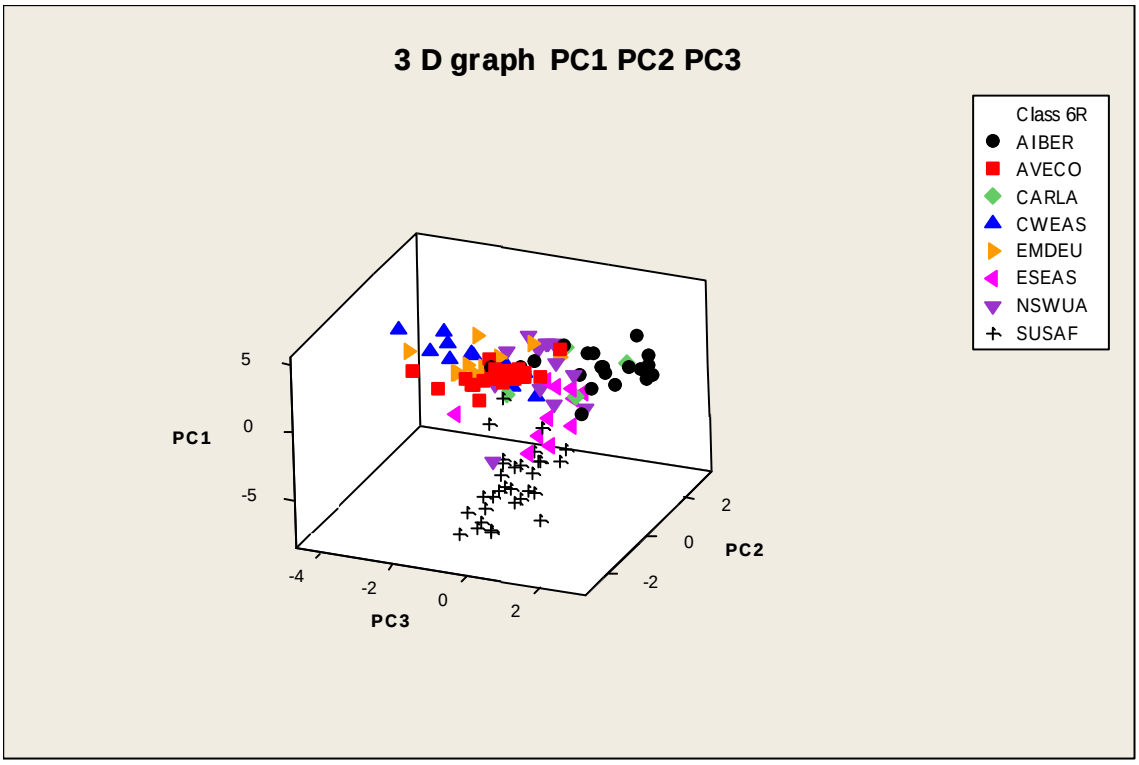
3.13 CLASSIFICAÇÃO NÃO SUPERVISIONADA, ANÁLISE DISCRIMINANTE

Após haver feito uma análise multivariada dos dados compilados pelo Banco Mundial, com a utilização do *software* MINITAB, apresentado análises das correlações dos componentes principais dos dados dimensionadores da Saúde e Meio Ambiente dos países analisados, apresentando também análises de estatística descritiva, prosseguiu-se a apresentação de dendogramas e regressões com o uso do método *stepwise*. As regiões foram classificadas em 6 e 3, e foram criados novos componentes PC1 (Desenvolvimento Humano), PC2 (Planeta Viável) e PC3 (Ambiente Sustentável).

A seguir será apresentada uma classificação não supervisionada (pressuposição de classificação dos países de acordo com as variáveis encontradas no *Stepwise*) e análise discriminante (AD). A AD é uma técnica multivariada que entre outras coisas ajuda a identificar se os dados são bons, válidos. Seu objetivo é identificar quais são as características distintitvas dos indivíduos em cada grupo e poder utilizá-las para estimar o grupo ao que pertencem outros indivíduos. A continuação, analisa-se os resultados pelo Método Linear e Quadrático.

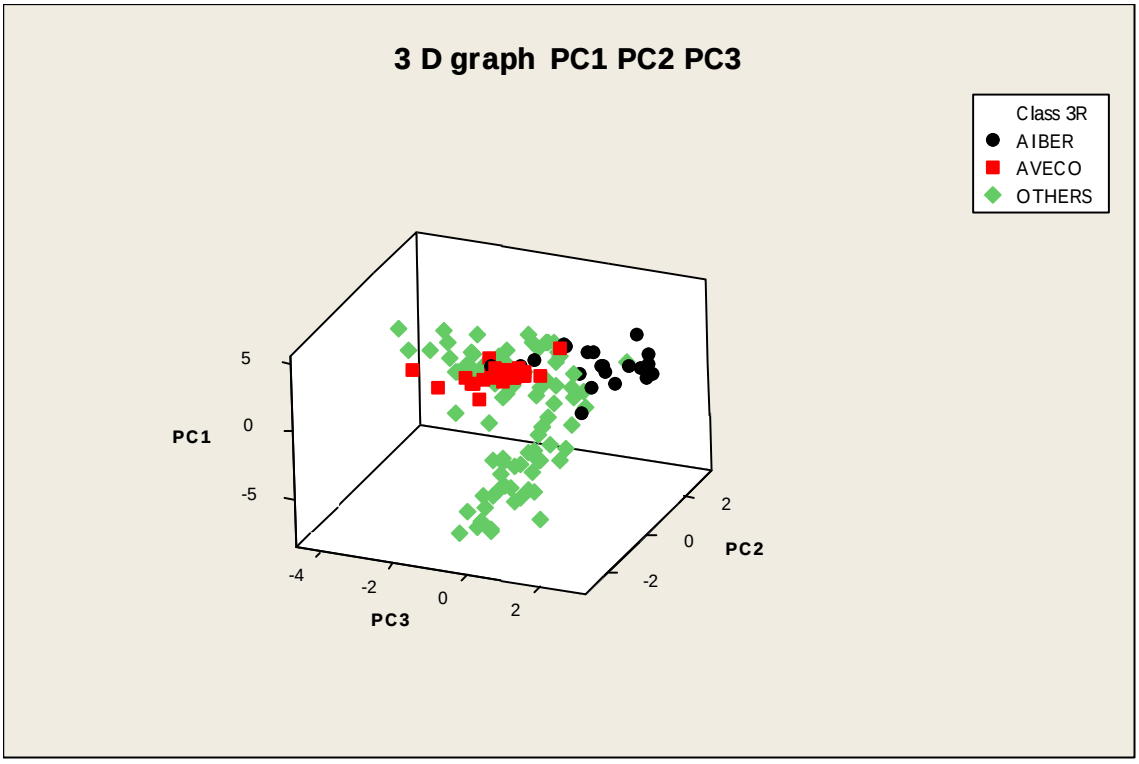
Os gráficos 113 (6 regiões) e 114 (3 regiões) mostram como estão distribuídos os países por região, de acordo com as componentes.

Gráfico 113.



Fonte: Elaborado pela autora

Gráfico 114.



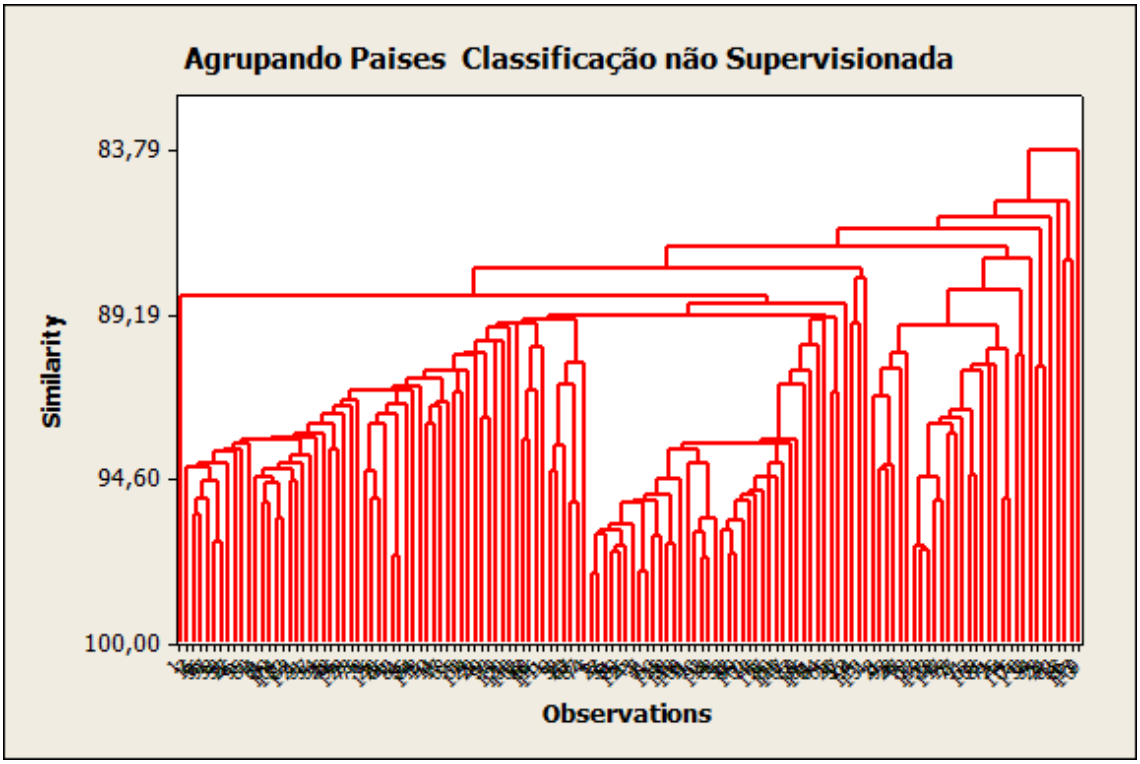
Fonte: Elaborado pela autora

3.13.1 CLASSIFICAÇÃO NÃO SUPERVISONADA

Chama-se de classificação não supervisionada ao método estatístico de agrupar alvos por similaridade, sem o conhecimento prévio do classificador sobre os atributos dos componentes envolvidos. O processo de classificar significa agrupar alvos iguais e separar os diferentes, e no caso em questão será utilizado o recurso do MINITAB: *Stat* > *Multivariate* > *Clusters Observations.*

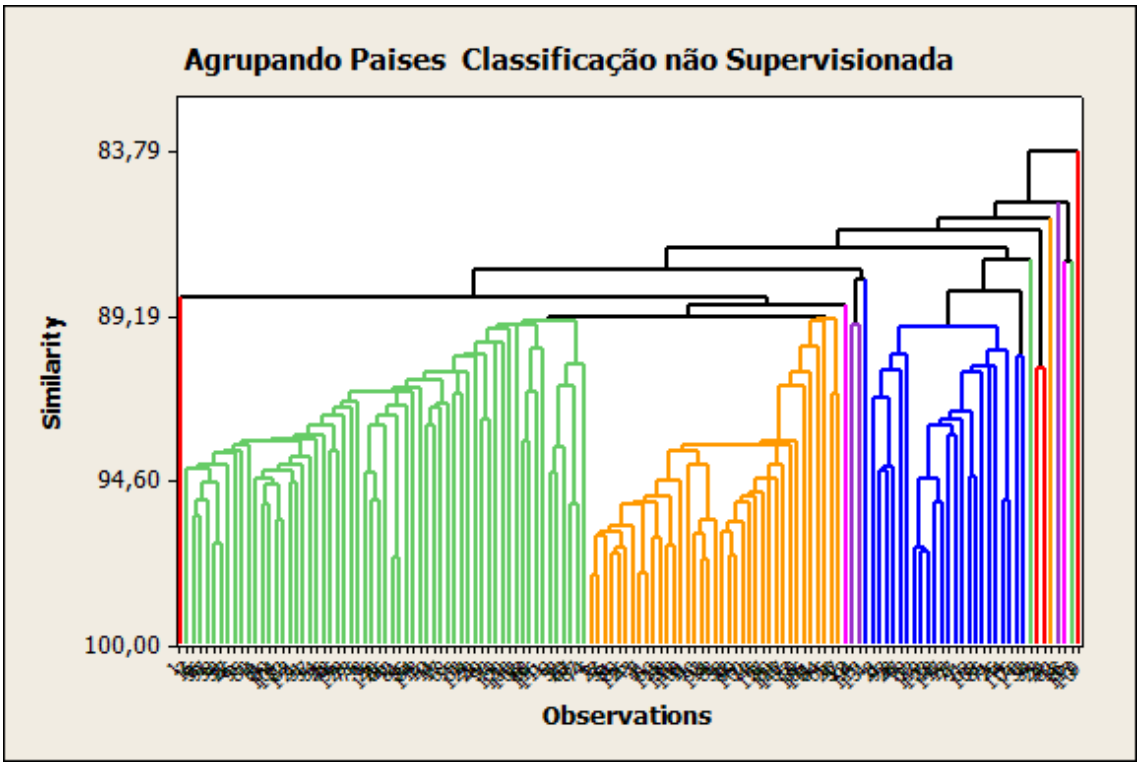
Será feita a observação dos *Clusters* PC1, PC2 e PC3 através de dendogramas. O primeiro deles foi apresentado com um único cluster, conforme pode ser visto no Dendograma monocromático do Gráfico 115. Posteriormente em vários mais (10), Gráfico 116.

Gráfico 115.



Fonte: Elaborado pela autora

Gráfico 116.



Fonte: Elaborado pela autora

One-way ANOVA: N_P_SPI_Regressao versus 4 grupos

Tabela 40

Source	DF	SS	MS	F	P
4 grupos	3	19816,1	6605,4	128,08	0,000
Error	128	6601,4	51,6		
Total	131	26417,5			

Fonte: Elaborado pela autora

S = 7,181 R-Sq = 75,01% R-Sq(adj) = 74,43%

Tabela 41.

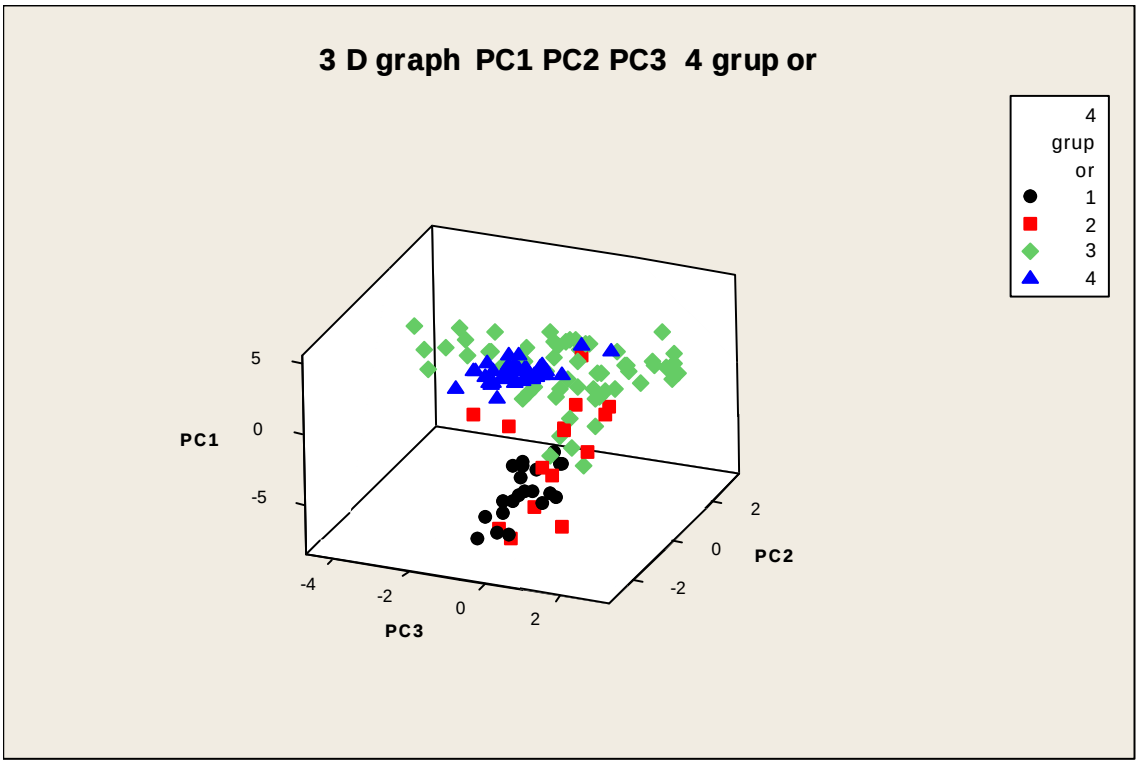
				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	----+---------+---------+---------+-----
azul	21	45,767	4,829	(-*--)
laranja	37	80,559	6,002	(-*-)
o	15	50,816	11,947	(--*--)
verde	59	62,723	6,998	(*-)
				----+---------+---------+---------+-----
				48 60 72 84

Fonte: Elaborado pela autora

Pooled StDev = 7,181

preto -> 1 vermelho - > 2 verde - > 3 azul -> 4

Gráfico 117



Fonte: Elaborado pela autora

Tabela 42. Tabulated statistics: Class 3R; 4 grup or

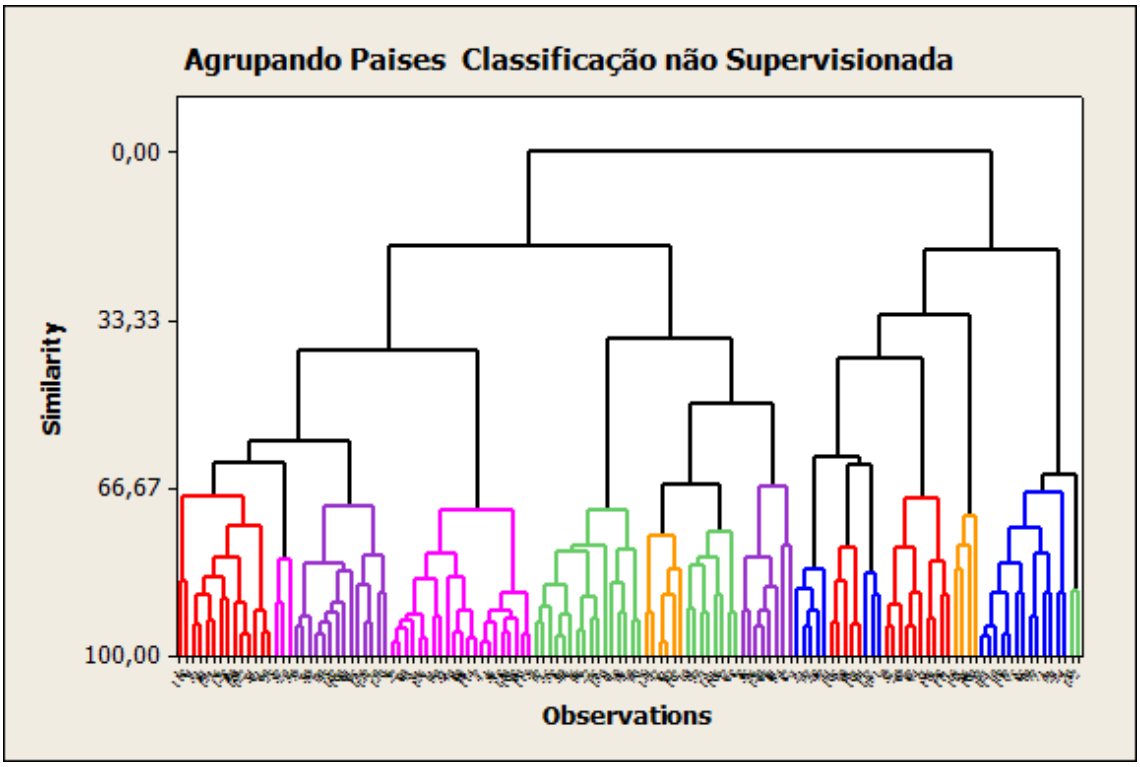
Rows: Class 3R	Columns: 4 grup or				
	1	2	3	4	All
AIBER	0	1	17	3	21
AVECO	0	0	1	26	27
OTHERS	21	14	41	8	84
All	21	15	59	37	132

Fonte: Elaborado pela autora

Para o Gráfico 118:
MINITAB: Stat > Multivariate > Clusters Observations

- Complete > Manhattan – Outra maneira de ver os dados, primeiro com 10 clusters, depois com os PC's

Gráfico 118.



Fonte: Elaborado pela autora

Diminuindo o dendograma de manhattan para 4 clusters, utilizando os principais indicadores das variáveis pc1, pc2, pc3 e pc4, tem-se:

Tabela 43. *Cluster Analysis of Observations*: **N_P_SPI_Regr; N_P_Deaths f; N_P_Access t; ...**

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	97,2008	18,187	7	43	7	2
2	130	96,2670	24,254	99	110	99	2
3	129	95,4121	29,809	32	108	32	2
4	128	95,3147	30,442	93	114	93	2
5	127	95,2492	30,867	22	56	22	2
6	126	94,3442	36,748	46	49	46	2
7	125	94,0161	38,880	67	87	67	2
8	124	93,9952	39,015	32	107	32	3
9	123	93,9645	39,215	7	41	7	3
10	122	93,9610	39,237	40	93	40	3
11	121	93,6327	41,371	11	51	11	2
12	120	93,6116	41,508	21	92	21	2
13	119	93,4480	42,570	30	45	30	2
14	118	93,3689	43,085	50	72	50	2
15	117	92,9214	45,992	38	59	38	2
16	116	92,8532	46,435	16	27	16	2
17	115	92,5972	48,098	71	117	71	2
18	114	92,4020	49,367	89	126	89	2
19	113	92,2082	50,626	69	122	69	2
20	112	92,0991	51,334	35	97	35	2
21	111	92,0950	51,361	98	118	98	2
22	110	92,0747	51,493	18	71	18	3
23	109	92,0735	51,501	58	99	58	3
24	108	91,7078	53,877	48	84	48	2
25	107	91,6864	54,016	83	106	83	2
26	106	91,6433	54,296	13	96	13	2
27	105	91,2121	57,098	66	129	66	2
28	104	91,2080	57,125	6	22	6	3
29	103	91,1991	57,182	100	128	100	2
30	102	91,1604	57,434	29	95	29	2
31	101	90,8442	59,488	33	88	33	2
32	100	90,7407	60,161	3	21	3	3
33	99	90,7177	60,310	5	14	5	2
34	98	90,4686	61,929	75	85	75	2
35	97	90,4005	62,371	120	125	120	2
36	96	90,3662	62,594	7	60	7	4
37	95	90,1971	63,692	63	103	63	2
38	94	90,1736	63,845	53	111	53	2
39	93	90,1022	64,309	12	123	12	2
40	92	89,9260	65,454	35	121	35	3
41	91	89,7835	66,380	39	79	39	2
42	90	89,7317	66,717	67	74	67	3
43	89	89,5584	67,842	77	91	77	2
44	88	89,5359	67,988	16	36	16	3
45	87	89,4901	68,286	4	80	4	2
46	86	88,7143	73,327	17	68	17	2
47	85	88,4698	74,915	30	58	30	5
48	84	88,3277	75,839	24	47	24	2
49	83	87,5463	80,916	53	98	53	4
50	82	87,5248	81,055	50	100	50	4
51	81	87,5225	81,070	2	66	2	3
52	80	87,3582	82,138	8	62	8	2
53	79	87,2727	82,693	3	78	3	4
54	78	87,1126	83,734	9	67	9	4
55	77	87,0793	83,950	35	38	35	5
56	76	86,7396	86,157	12	119	12	3
57	75	86,7232	86,264	11	89	11	4
58	74	86,6943	86,451	4	61	4	3
59	73	86,3219	88,871	65	120	65	3
60	72	86,3000	89,013	33	40	33	5
61	71	86,1765	89,815	10	54	10	2
62	70	85,9918	91,016	17	102	17	3
63	69	85,9851	91,060	5	42	5	3
64	68	85,9496	91,290	16	46	16	5
65	67	85,8750	91,774	112	131	112	2
66	66	85,7780	92,405	48	90	48	3
67	65	85,4898	94,277	25	29	25	3
68	64	85,0963	96,834	81	101	81	2
69	63	85,0602	97,069	6	11	6	7

70	62	84,8384	98,510	18	19	18	4
71	61	84,8130	98,675	20	44	20	2
72	60	84,6819	99,527	7	33	7	9
73	59	84,4534	101,011	1	35	1	6
74	58	84,4014	101,349	10	37	10	3
75	57	84,2089	102,600	31	104	31	2
76	56	84,1479	102,996	52	94	52	2
77	55	84,1203	103,176	3	28	3	5
78	54	84,0204	103,824	30	64	30	6
79	53	83,5384	106,957	31	69	31	4
80	52	83,4768	107,357	73	83	73	3
81	51	82,9291	110,915	34	105	34	2
82	50	82,5330	113,489	32	50	32	7
83	49	82,3956	114,381	18	77	18	6
84	48	82,1620	115,899	2	116	2	4
85	47	82,0938	116,343	70	132	70	2
86	46	81,7983	118,262	4	130	4	4
87	45	81,7182	118,783	15	86	15	2
88	44	81,3317	121,294	13	16	13	7
89	43	81,0039	123,424	7	115	7	10
90	42	80,8136	124,660	63	75	63	4
91	41	80,6842	125,501	4	31	4	8
92	40	80,6755	125,557	6	127	6	8
93	39	80,1058	129,259	73	76	73	4
94	38	80,0506	129,618	32	39	32	9
95	37	79,9586	130,216	8	81	8	4
96	36	79,7763	131,400	25	57	25	4
97	35	79,5493	132,875	5	17	5	6
98	34	78,5918	139,096	8	124	8	5
99	33	78,4341	140,121	9	52	9	6
100	32	78,0169	142,831	1	48	1	9
101	31	77,8307	144,041	2	53	2	8
102	30	77,7656	144,464	30	32	30	15
103	29	77,1587	148,407	23	24	23	3
104	28	76,7645	150,968	12	20	12	5
105	27	76,2879	154,065	1	13	1	16
106	26	76,2247	154,476	10	55	10	4
107	25	75,5543	158,832	9	112	9	8
108	24	74,3613	166,583	8	26	8	6
109	23	73,8676	169,791	3	12	3	10
110	22	73,4070	172,784	15	109	15	3
111	21	72,2703	180,169	4	10	4	12
112	20	71,2888	186,546	18	23	18	9
113	19	70,7908	189,782	25	73	25	8
114	18	70,3868	192,406	7	30	7	25
115	17	69,6810	196,992	63	70	63	6
116	16	68,8844	202,168	5	8	5	12
117	15	68,6673	203,579	2	82	2	9
118	14	67,3442	212,175	3	34	3	12
119	13	65,2176	225,992	6	7	6	33
120	12	64,2565	232,237	1	25	1	24
121	11	62,3553	244,590	3	9	3	20
122	10	60,7415	255,075	4	65	4	15
123	9	59,8706	260,734	18	63	18	15
124	8	59,0480	266,078	4	5	4	27
125	7	58,1911	271,646	15	113	15	4
126	6	50,6986	320,327	2	4	2	36
127	5	49,4630	328,355	1	6	1	57
128	4	46,6691	346,508	3	18	3	35
129	3	46,0554	350,495	1	2	1	93
130	2	33,5572	431,700	3	15	3	39
131	1	0,0000	649,732	1	3	1	132

Fonte: Elaborado pela autora

Tabela 44.

Final Partition

Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	57	123904	45,0921	76,1307
Cluster2	36	96448	50,2263	77,2127
Cluster3	35	86900	47,8116	73,3511
Cluster4	4	9003	46,5853	61,3661

Fonte: Elaborado pela autora

Análise:

- Menor distância do centróide: 45 no Cluster 1.
- Maior distância: 77 no Cluster 2.

Tabela 45. *Cluster Centroids*

Variable	Cluster1	Cluster2	Cluster3	Cluster4
N_P_SPI_Regressao	75,5526	62,8731	45,6049	59,6550
N_P_Deaths from	95,9737	93,4753	45,8245	28,6977
N_P_Access to piped water	89,4350	69,6358	14,3543	53,7779
N_P IDH - 2013_1	74,8033	65,5757	22,5371	62,3970
N_P External resources	12,3110	6,5356	39,2134	20,4799
N_P GovernanÁa_1	64,5979	33,6798	24,0761	51,2754
N_P Happy Planet Index_1	56,8065	51,6254	34,2344	25,3185
N_P GINNI Index_1	65,5664	70,3575	56,9940	10,5363
N_P Outdoor air pollution	80,1650	61,6501	84,4776	87,6866
N_P Obesity rate (% of pop.)_1	52,1435	48,7543	88,7359	53,8969

Variable	Grand centroid
N_P_SPI_Regressao	63,6721
N_P_Deaths from	79,9565
N_P_Access to piped water	63,0469
N_P IDH - 2013_1	58,0523
N_P External resources	18,1166
N_P GovernanÁa_1	45,0176
N_P Happy Planet Index_1	48,4542
N_P GINNI Index_1	62,9325
N_P Outdoor air pollution	76,4869
N_P Obesity rate (% of pop.)_1	60,9749

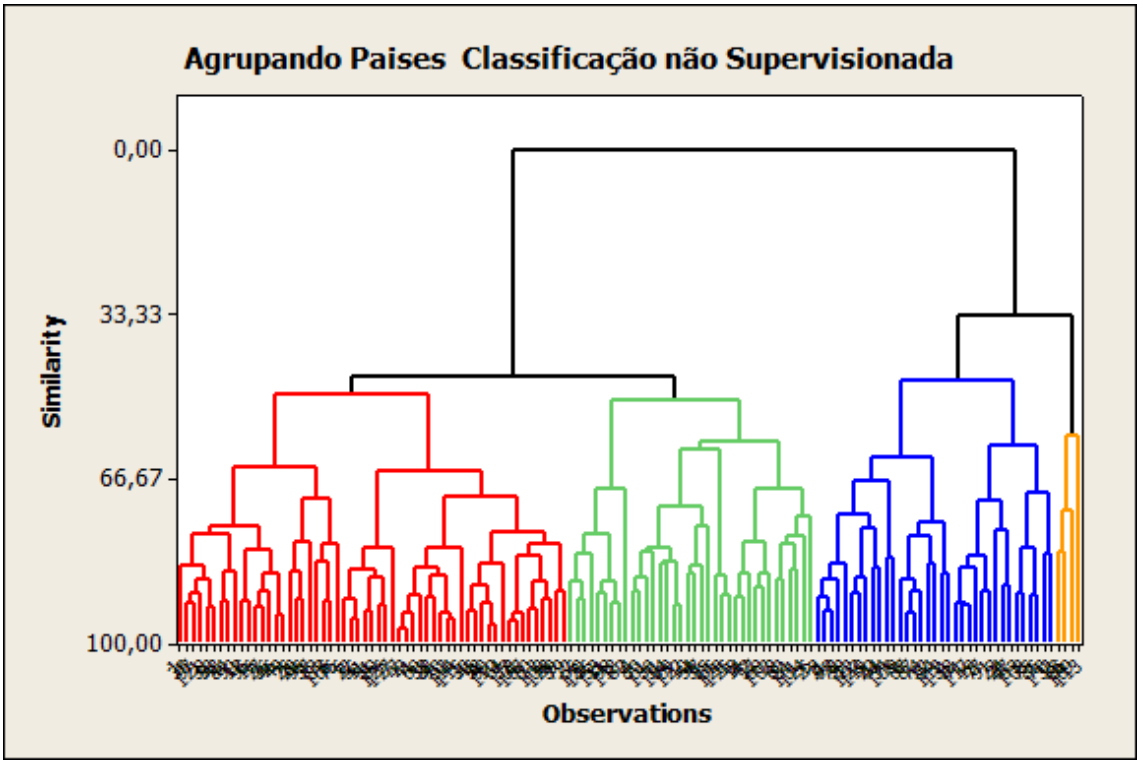
Fonte: Elaborado pela autora

Tabela 46. ***Distances Between Cluster Centroids***

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,000	45,139	126,835	102,605
Cluster2	45,139	0,000	106,009	99,735
Cluster3	126,835	106,009	0,000	90,479
Cluster4	102,605	99,735	90,479	0,000

Fonte: Elaborado pela autora

Gráfico 119.



Fonte: Elaborado pela autora

Legenda: Vermelho Cluster 1 ; Verde: Cluster 2; Azul: Cluster 3; Laranja: Cluster 4.

Análise:

- Menor distância entre os centroides dos Clusters 1 e 2: 45.
- Maior distância entre os centroids dos clusters 1 e 3: 126.

3.13.2 ANÁLISE DISCRIMINANTE: 4 GRUP OR VERSUS PC1; PC2; PC3; PC4

Na Coluna do Minitab C34, foi botado o resultado da análise discriminante dos 4 clusters.

a) Pelo método Linear:

Predictors: PC1; PC2; PC3; PC4

Tabela 47

Group	1	2	3	4
Count	21	15	59	37

Summary of classification				
	True Group			
Put into Group	1	2	3	4
1	16	4	2	0
2	5	7	0	0
3	0	4	55	3
4	0	0	2	34
Total N	21	15	59	37
N correct	16	7	55	34
Proportion	0,762	0,467	0,932	0,919

Fonte: Elaborado pela autora

N = 132 N Correct = 112 **Proportion Correct = 0,848**

Análise:
O grau de acerto foi de 84,8%, de maneira que o melhor grupo foi o 3, com grau de acerto de 93,2%. O grupo 1 teve um grau de acerto de 76,2%; o 2, de 46,7% e o 4 de 91,9%. Dos 20 erros: 5 foram no grupo 1, 8 no gruo 2, e 3 no grupo 4.

Tabela 48. *Squared Distance Between Groups*

	1	2	3	4
1	0,0000	2,6637	20,0279	40,3475
2	2,6637	0,0000	12,0380	33,3340
3	20,0279	12,0380	0,0000	12,2290
4	40,3475	33,3340	12,2290	0,0000

Fonte: Elaborado pela autora

Análise:

- A tabela 48 indica a distância entre os grupos;
- O melhor grupo é o 3;
- Menor distância está entre os grupos 1 e 2;
- Maior distância está entre os grupos 1 e 4.

b) Pelo Método Quadrático

```
Predictors: PC1; PC2; PC3; PC4
```

Tabela 49.

Group	1	2	3	4
Count	21	15	59	37

Summary of classification				
	True Group			
Put into Group	1	2	3	4
1	21	0	0	0
2	0	14	3	0
3	0	1	55	1
4	0	0	1	36
Total N	21	15	59	37
N correct	21	14	55	36
Proportion	1,000	0,933	0,932	0,973

Fonte: Elaborado pela autora

```
N = 132          N Correct = 126   Proportion Correct = 0,955
```

Análise:
Pode-se observer que houve um grau de acerto de 95,5%. Analisando cada grupo o grupo 1 foi o melhor, pois teve 100% de acerto, o 2, teve um grau de acerto de 93,3%, o 3, de 93,2% e o 4, grau de acerto de 97,3%. Dos 6 erros, 1 foi do grupo 2, 4 do grupo 3 e 1 do grupo 4.

Na tabela 50 a seguir, se observará qual país mudou de grupo e quais erros podem ser considerados na classificação não supervisionada.

Tabela 50. *Summary of Misclassified Observations*

Observation	True Group	Pred Group	Group	Squared Distance	Probability	
25**	4	3	1	82,935	0,000	Chile
			2	9,994	0,059	
			3	5,040	0,702	
			4	7,197	0,239	
48**	3	2	1	20,970	0,000	Guyana
			2	4,294	0,721	
			3	6,195	0,279	
			4	116,987	0,000	
74**	3	2	1	18,196	0,011	Madagascar
			2	9,800	0,701	
			3	11,577	0,288	
			4	538,893	0,000	
82**	2	3	1	62,384	0,000	Mongolia
			2	9,865	0,127	
			3	6,018	0,873	

Fonte: Elaborado pela autora

Análise:

- GRAU de acerto total melhorou com o método quadrático, com 95,5% (subiu de 84,8, ou seja, 10,7%).
- O melhor grupo é o 1, com grau de acerto de 100%.
- Na Classificação não Supervisionada Chile pertencia ao Grupo 4, mas na realidade, pela similaridade ele pertence ao Grupo 3, Mongolia também presumia-se que pertencia ao grupo 2, mas pertence ao 3. Guyana e Madagascar pertencem ao Grupo 2 e não ao 3 como se presumia.

One-way ANOVA

Tabela 51.

One_way ANOVA N_P_SPI_Regressao versus 4 Grup

Source	DF	SS	MS	F	P
4 Grup	3	18969,6	6323,2	108,67	0,000
Error	128	7447,9	58,2		
Total	131	26417,5			

S = 7,628 R-Sq = 71,81% R-Sq(adj) = 71,15%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	52	76,556	8,736	(--*-)
2	38	63,010	6,071	(-*-)
3	27	50,684	8,222	(--*--)
4	15	44,064	5,572	(---*---)
				40 50 60 70

Fonte: Elaborado pela autora

Pooled StDev = 7,628

Se se muda o SPI pelo IDH devido a sua complexidade, como já foi feito anteriormente, tem-se:

Tabela 52. One-way ANOVA: N_P IDH - 2013_1 versus 4 Grup

```
Source   DF     SS     MS       F      P
4 Grup    3  64229  21410  118,82  0,000
Error   128  23063    180
Total   131  87292

S = 13,42   R-Sq = 73,58%   R-Sq(adj) = 72,96%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
1      52  79,99  13,18                                    (-*-)
2      38  59,56  10,27                           (-*-)
3      27  37,59  18,36                 (--*-)
4      15  15,00  10,45  (---*--)
                          ------+---------+---------+---------+---
                               20        40        60        80
```

Fonte: Elaborado pela autora

```
Pooled StDev = 13,42
```

Análise:

Como já foi visto, no ANOVA One-Way se analisa grupos independentes. Nos dois que foram rodados se buscou analisar as variáveis Social Progress e IDH nos 4 grupos em comparação, rodando-as em separado.

No primeiro ANOVA One-Way computado com o Social Progress, o melhor é o grupo 1, com a média 76,55. No Segundo ANOVA, rodado com o IDH, a média subiu para 79,99, e também o melhor grupo é o 1. Como o SPI é uma variável muito complexa e de pouca maleabilidade, o IDH é o mais recomendável.

Como visto acima, o melhor Grupo foi o 1, e o melhor método foi o Quadrático, com 95,5% de acerto.

Aqui, foi visto a principal variável explicativa (SPI => IDH) comparando a sua incidência em cada um dos componentes/ grupos. Abaixo, serão comparadas todas as variáveis mais relevantes resultantes do Stepwise e sua incidência comparativa nos grupos, a fim de distinguir qual o melhor grupo e melhor modelo, linear ou quadrático.

3.14 ANÁLISE MULTIVARIADA MÉTODOS LINEAR E QUADRÁTICO

Fazendo uma Análise discriminante das variáveis mais representativas dos 4 P's, ou seja, dos 4 clusters – resultado do Stepwise feito previamente, calculamos a seguir qual o método se encaixa melhor no Modelo.

c) Linear

```
Linear Method for Response: 4 Grup:


Predictors: N_P_Deaths from; N_P_Access to piped water;
            N_P Obesity rate (% of pop.)_1; N_P Outdoor air pollution;
            N_P IDH - 2013_1; N_P GovernanÁa_1; N_P EPI Score_1;
```

N_P Happy Planet Index_1; N_P OHI_1; N_P GINNI Index_1

Tabela 53.

Group	1	2	3	4
Count	52	38	27	15

Summary of classification				
	True Group			
Put into Group	1	2	3	4
1	44	1	0	0
2	8	35	1	0
3	0	2	22	2
4	0	0	4	13
Total N	52	38	27	15
N correct	44	35	22	13
Proportion	0,846	0,921	0,815	0,867

Fonte: Elaborado pela autora

N = 132 N Correct = 114 Proportion Correct = 0,864

Análise

- Grau de acerto total 86,4%
- Melhor grupo é o 2, com 92,1% de acerto, seguido pelo 4, 1 e 3, nesta ordem.
- Os erros foram de 8 erros no grupo 1; 3 no grupo 2; 5 no grupo 3 e 2 no grupo 4, totalizando 18.

Tabela 54.

Squared Distance Between Groups				
	1	2	3	4
1	0,0000	6,4388	39,8579	81,4904
2	6,4388	0,0000	25,5412	62,5587
3	39,8579	25,5412	0,0000	10,5485
4	81,4904	62,5587	10,5485	0,0000

Fonte: Elaborado pela autora

Análise:

- Menor distância foi entre os grupos 1 e 2;
- Maior distância entre os grupos 1 e 4.

Tabela 55. Qual a distância linear para o grupo

Linear Discriminant Function for Groups				
	1	2	3	4
Constant	-100,02	-79,69	-66,26	-56,27
N_P_Deaths from	0,48	0,48	0,17	-0,05
N_P_Access to piped water	0,47	0,42	0,24	0,17
N_P Obesity rate (% of pop.)_1	0,33	0,35	0,50	0,53
N_P Outdoor air pollution	0,38	0,32	0,37	0,38
N_P IDH - 2013_1	0,49	0,44	0,45	0,33
N_P GovernanÁa_1	-0,01	-0,08	-0,02	0,00
N_P EPI Score_1	0,10	0,07	0,03	0,07
N_P Happy Planet Index_1	-0,14	-0,12	-0,14	-0,16
N_P OHI_1	0,37	0,35	0,40	0,38
N_P GINNI Index_1	0,19	0,14	0,16	0,18

Fonte: Elaborado pela autora

Tabela 56. Erros na Classificação Não Supervisionada

Observation	True Group	Pred Group	Group	Squared Distance	Probability	
4**	1	2	1	11,063	0,151	Argentina
			2	7,611	0,849	
			3	51,216	0,000	
			4	98,108	0,000	
12**	4	3	1	57,292	0,000	Benin
			2	41,251	0,000	
			3	4,579	0,734	
			4	6,611	0,266	
21**	3	4	1	68,507	0,000	Cameron
			2	50,452	0,000	
			3	7,975	0,082	
			4	3,148	0,918	
28**	3	4	1	70,90	0,000	Congo
			2	48,70	0,000	
			3	14,87	0,145	
			4	11,32	0,855	
31**	1	2	1	10,934	0,059	Cuba
			2	5,401	0,941	
			3	35,653	0,000	
			4	80,656	0,000	
34**	2	3	1	38,48	0,000	Djibouti
			2	21,12	0,168	
			3	17,92	0,832	
			4	32,81	0,000	
47**	3	4	1	78,797	0,000	Guinea
			2	58,559	0,000	
			3	11,064	0,092	
			4	6,489	0,908	
54**	1	2	1	13,357	0,073	Iran
			2	8,283	0,927	
			3	42,776	0,000	
			4	86,532	0,000	
69**	1	2	1	10,023	0,124	Lebanon
			2	6,117	0,876	
			3	50,144	0,000	
			4	96,017	0,000	
80**	1	2	1	9,507	0,253	Mexico
			2	7,341	0,747	
			3	49,665	0,000	
			4	94,021	0,000	
82**	2	3	1	34,74	0,003	Mongolia
			2	25,43	0,339	
			3	24,10	0,658	
			4	52,00	0,000	
92**	3	4	1	77,961	0,000	Nigeria
			2	57,683	0,000	
			3	11,640	0,038	
			4	5,155	0,962	
106**	2	1	1	4,654	0,844	Serbia
			2	8,031	0,156	
			3	41,977	0,000	
			4	81,526	0,000	
118**	1	2	1	19,46	0,025	Thailand
			2	12,35	0,877	
			3	16,73	0,098	
			4	46,45	0,000	
119**	4	3	1	69,070	0,000	Togo
			2	48,615	0,000	
			3	7,754	0,521	
			4	7,923	0,479	
121**	1	2	1	4,543	0,283	Tunisia
			2	2,686	0,717	
			3	28,515	0,000	
			4	64,235	0,000	
122**	1	2	1	7,612	0,268	Turkey

			2	5,605	0,732	
			3	49,062	0,000	
			4	94,565	0,000	
131**	3	2	1	27,37	0,000	Yemen
			2	12,91	0,665	
			3	14,28	0,335	
			4	36,76	0,000	

Fonte: Elaborado pela autora

Análise:

- Grau de acerto total foi de 86,4%.
- Menor distância foi entre os grupos 1 e 2.
- Maior distância entre 1 e 4.
- Grau de acerto total melhorou em relação a anterior Linear (0,848) em 1,6%.
- Grau de acerto total piorou em relação à quadrática (0,955) em 0,91%
- O melhor grupo é o 2, com grau de acerto de 92,1%
- Destaques: Pelo Método Linear, os países que eram pretensamente do Grupo 1 eram Argentina, Cuba, México, Líbano, Irã, Tailandia, Tunisia, Turquia, mas na verdade, eram do Grupo 2, curiosamente, o que aparentemente signigica que pioraram a sua classificação. Os do Grupo 2 Djibouti e Mongolia pertencem ao 3 e Serbia ao 1, melhorando a sua classificação. Mongolia já havia sido detectado anteriormente. E do Grupo 3 Cameron, Congo, Guinea, Nigeria pertencem ao 4, e Yemen, ao 2. E do Grupo 4, Benin e Togo pertencem ao 3.

c) Quadrática

```
Predictors: N_P_Deaths from; N_P_Access to piped water;
            N_P Obesity rate (% of pop.)_1; N_P Outdoor air pollution;
            N_P IDH - 2013_1; N_P GovernanÁa_1; N_P EPI Score_1;
            N_P Happy Planet Index_1; N_P   OHI_1; N_P   GINNI Index_1
```

Tabela 57.

Group	1	2	3	4
Count	52	38	27	15
Summary of classification				
		True Group		
Put into Group	1	2	3	4
1	48	1	0	0
2	3	37	0	0
3	1	0	27	0
4	0	0	0	15
Total N	52	38	27	15
N correct	48	37	27	15
Proportion	0,923	0,974	1,000	1,000

Fonte: Elaborado pela autora

```
N = 132          N Correct = 127   Proportion Correct = 0,962
```

Tabela 58. *From Generalized Squared Distance to Group*

Group	1	2	3	4
1	47,48	68,09	100,69	2723,12
2	55,16	49,53	74,57	2687,29
3	369,87	99,19	49,39	346,64
4	1044,02	204,67	58,60	39,19

Fonte: Elaborado pela autora

Análise:

- Menor Distâancia: Do Grupo 1 Para o 2
- Maior Distância Do Grupo 1 Para o 4

Tabela 59. *Summary of Misclassified Observations*- Qual País Que Mudou De Grupo Ou Erros

Observation	True Group	Pred Group	Group	Squared Distance	Probability	
29**	1	2	1	65,35	0,380	Costa Rica
			2	64,37	0,620	
			3	107,79	0,000	
			4	3529,66	0,000	
76**	1	2	1	75,42	0,092	Malaysia
			2	70,85	0,899	
			3	79,95	0,009	
			4	2282,58	0,000	
106**	2	1	1	56,71	0,908	Serbia
			2	61,28	0,092	
			3	102,06	0,000	
			4	3389,85	0,000	
118**	1	3	1	74,33	0,001	Thailand
			2	68,80	0,015	
			3	60,38	0,984	
			4	1527,19	0,000	
121**	1	2	1	57,08	0,132	Tunisia
			2	53,31	0,868	
			3	77,30	0,000	
			4	2976,65	0,000	

Fonte: Elaborado pela autora

Análise

- O grau de acerto total é de 96,20%, o que indica que o Método quadrático é superior ao Linear, apresentando um Grau de Acerto Maior.
- Os grupos 3 e 4 são os melhores com 100% de acerto.
- A maior Distância é do grupo 1 e 4;
- Os grupos 1 e 2 são os mais próximos;
- Destaques: A Thailand melhorou sua classificação, da 1 para o 3. Argentina piorou sua classificação. Costa Rica, Malaysia e Tunisia eram pretensamente do Grupo 1, mas foram reclassificados para o grupo 2, também melhorando um pouco sua classificação, Serbia, como já dito, era do 2 e foi para o 1, piorando sua classificação.

O método quadrático teve o maior poder explicativo, e os grupos 3 e 4 apresentaram-se como melhores, ambos apresentando os componentes importantes que identificam a Saúde e Bem-estar social dos países investigados.

Adicionalmente, agora recomenda-se classificar os países segundo a classificação: AIBER/ AVECO/ OTHERS.

Discriminant Analysis: Class 3R versus N_P_Deaths f; N_P_Access t; ...

e) Linear: Class 3R

```
Predictors: N_P_Deaths from; N_P_Access to piped water;
            N_P Obesity rate (% of pop.)_1; N_P Outdoor air pollution;
            N_P IDH - 2013_1; N_P GovernanÁa_1; N_P EPI Score_1;
            N_P Happy Planet Index_1; N_P   OHI_1; N_P   GINNI Index_1


Group     AIBER     AVECO     OTHERS
Count        21        27         84
```

Tabela 60.

Summary of classification			
	True Group		
Put into Group	AIBER	AVECO	OTHERS
AIBER	18	1	8
AVECO	2	25	4
OTHERS	1	1	72
Total N	21	27	84
N correct	18	25	72
Proportion	0,857	0,926	0,857

Fonte: Elaborado pela autora

```
N = 132           N Correct = 115           Proportion Correct = 0,871
```

Análise:

- Grau de acerto 87,1%.
- Melhor grupo AVECO.
- Aiber e Others têm o mesmo grau de acerto proporcional de 85,7%, embora nominalmente AIBER tenha 3 erros pela classificação não supervisionada e Others, 12 erros.

Tabela 61. Squared Distance Between Groups

	AIBER	AVECO	OTHERS
AIBER	0,0000	17,8675	7,3150
AVECO	17,8675	0,0000	13,3945
OTHERS	7,3150	13,3945	0,0000

Fonte: Elaborado pela autora

Tabela 62. Linear Discriminant Function for Groups

	AIBER	AVECO	OTHERS
Constant	-66,741	-85,698	-60,252
N_P_Deaths from	0,076	-0,023	0,033
N_P_Access to piped water	0,276	0,250	0,220
N_P Obesity rate (% of pop.)_1	0,535	0,618	0,551
N_P Outdoor air pollution	0,375	0,369	0,333
N_P IDH - 2013_1	0,261	0,319	0,311
N_P GovernanÁa_1	0,023	0,171	0,035
N_P EPI Score_1	0,171	0,262	0,145
N_P Happy Planet Index_1	-0,033	-0,071	-0,105
N_P OHI_1	0,332	0,300	0,324
N_P GINNI Index_1	0,096	0,273	0,200

Fonte: Elaborado pela autora

Tabela 63. Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability	
2**	OTHERS	AIBER	AIBER	7,273	0,514	Algeria
			AVECO	18,122	0,002	
			OTHERS	7,398	0,483	
48**	OTHERS	AIBER	AIBER	4,047	0,823	Guyana
			AVECO	24,185	0,000	
			OTHERS	7,122	0,177	
50**	OTHERS	AVECO	AIBER	16,519	0,002	Hungary
			AVECO	4,693	0,872	
			OTHERS	8,568	0,126	
57**	AVECO	AIBER	AIBER	6,062	0,785	Israel
			AVECO	9,877	0,117	
			OTHERS	10,220	0,098	
59**	OTHERS	AIBER	AIBER	5,916	0,979	Jamaica
			AVECO	22,529	0,000	
			OTHERS	13,658	0,020	
61**	OTHERS	AIBER	AIBER	10,50	0,549	Jordan
			AVECO	21,64	0,002	
			OTHERS	10,90	0,449	
68**	AVECO	OTHERS	AIBER	25,09	0,008	Latvia
			AVECO	17,97	0,289	
			OTHERS	16,19	0,702	
72**	OTHERS	AVECO	AIBER	17,157	0,003	Lithuania
			AVECO	6,629	0,674	
			OTHERS	8,107	0,322	
76**	OTHERS	AIBER	AIBER	10,12	0,664	Malaysia
			AVECO	13,18	0,144	
			OTHERS	12,60	0,192	
79**	OTHERS	AVECO	AIBER	17,87	0,174	Mauritius
			AVECO	15,14	0,683	
			OTHERS	18,27	0,143	
86**	OTHERS	AIBER	AIBER	10,06	0,770	Namibia
			AVECO	23,81	0,001	
			OTHERS	12,49	0,229	
99**	OTHERS	AVECO	AIBER	12,160	0,004	Poland
			AVECO	1,420	0,961	
			OTHERS	8,085	0,034	
100**	AIBER	AVECO	AIBER	13,454	0,007	Portugal
			AVECO	3,649	0,961	
			OTHERS	10,478	0,032	
109**	OTHERS	AIBER	AIBER	24,34	0,629	SulAfric
			AVECO	40,40	0,000	
			OTHERS	25,40	0,370	
110**	AIBER	AVECO	AIBER	13,412	0,007	Spain
			AVECO	3,670	0,978	
			OTHERS	12,117	0,014	
122**	OTHERS	AIBER	AIBER	5,050	0,641	Turkey
			AVECO	18,280	0,001	
			OTHERS	6,214	0,358	
128**	AIBER	OTHERS	AIBER	12,04	0,381	Uruguay
			AVECO	15,74	0,060	
			OTHERS	11,27	0,559	

Fonte: Elaborado pela autora

Análise:

- ✓ Menor distância entre grupos: AIBER para AVECO;
- ✓ A proporção total de acerto pelo método linear foi de 87,10%;
- ✓ O melhor grupo é AVECO, com 92,60% de acerto;
- ✓ Destaques: Portugal, Hungary, Lithuania, Mauritius, Poland e Spain melhoraram a sua classificação, migrando para o Grupo AVECO. Latvia e Israel pioraram nos componentes analisados. Os demais não sofreram alterações de grande peso.

f) Quadrática: Class 3R

```
Predictors: N_P_Deaths from; N_P_Access to piped water;
            N_P Obesity rate (% of pop.)_1; N_P Outdoor air pollution;
            N_P IDH - 2013_1; N_P GovernanÁa_1; N_P EPI Score_1;
            N_P Happy Planet Index_1; N_P   OHI_1; N_P   GINNI Index_1

Group     AIBER     AVECO     OTHERS
Count        21        27         84
```

Tabela 64.

Summary of classification			
	True Group		
Put into Group	AIBER	AVECO	OTHERS
AIBER	20	1	7
AVECO	1	26	4
OTHERS	0	0	73
Total N	21	27	84
N correct	20	26	73
Proportion	0,952	0,963	0,869

Fonte: Elaborado pela autora

N = 132 N Correct = 119 Proportion Correct = 0,902

Análise:

- Grau de acerto total de 90,2%, com total de 13 erros nominais devido à classificação não supervisionada.
- Melhor grupo é o AVECO com 96,3% de acertos e 1 erro nominal.
- Segundo melhor grupo é AIBER, com 95,2% de acertos e 1 erro nominal, seguido pelo Others com 86,9% de acerts e 11 erros nominais.

Tabela 65.

From	Generalized Squared Distance to Group		
Group	AIBER	AVECO	OTHERS
AIBER	44,70	188,62	61,91
AVECO	61,71	41,80	70,49
OTHERS	99,90	2803,35	55,08

Fonte: Elaborado pela autora

Tabela 66. Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability	
2**	OTHERS	AIBER	AIBER	61,13	0,586	Algeria
			AVECO	776,20	0,000	
			OTHERS	61,82	0,414	
30**	OTHERS	AVECO	AIBER	59,12	0,075	Croatia
			AVECO	54,21	0,871	
			OTHERS	59,78	0,054	
48**	OTHERS	AIBER	AIBER	53,64	0,985	Guyana
			AVECO	691,30	0,000	
			OTHERS	62,02	0,015	
50**	OTHERS	AVECO	AIBER	60,59	0,041	Hungary
			AVECO	54,31	0,953	
			OTHERS	64,49	0,006	
54**	OTHERS	AIBER	AIBER	64,72	0,636	Iran
			AVECO	128,73	0,000	
			OTHERS	65,84	0,364	

57**	AVECO	AIBER	AIBER	51,84	0,812	Israel
			AVECO	54,79	0,186	
			OTHERS	63,46	0,002	
59**	OTHERS	AIBER	AIBER	60,18	0,991	Jamaica
			AVECO	483,22	0,000	
			OTHERS	69,49	0,009	
72**	OTHERS	AVECO	AIBER	58,70	0,048	Lithuania
			AVECO	52,71	0,949	
			OTHERS	63,86	0,004	
76**	OTHERS	AIBER	AIBER	58,76	0,978	Malaysia
			AVECO	642,95	0,000	
			OTHERS	66,32	0,022	
83**	OTHERS	AIBER	AIBER	58,12	0,982	Montenegro
			AVECO	91,66	0,000	
			OTHERS	66,15	0,018	
99**	OTHERS	AVECO	AIBER	57,44	0,002	Poland
			AVECO	44,61	0,998	
			OTHERS	63,77	0,000	
110**	AIBER	AVECO	AIBER	54,14	0,037	Spain
			AVECO	47,62	0,963	
			OTHERS	67,03	0,000	
121**	OTHERS	AIBER	AIBER	52,74	0,939	Uruguay
			AVECO	321,54	0,000	
			OTHERS	58,21	0,061	

Fonte: Elaborado pela autora

Análise:

- ✓ A proporção total de acerto é de 90,20%, tendo aumentado com relação ao método linear em 3,1%.
- ✓ O melhor grupo é o AVECO, com 96,3% de acertos.
- ✓ O pior grupo é others, com 86,9%.
- ✓ A menor distância entre os grupos é entre AIBER e OTHERS.
- ✓ Destaques: Croatia, Hungary, Lithuania, Poland e Spain melhoraram sua classificação consideravelmente, para AVECO. Os demais melhoraram de OTHERS para AIBER, então se vê que todos foram melhor reclassificados excetuando-se Israel, que sofreu uma piora, provavelmente pela instabilidade da região com ameaças constantes de guerra.

A função quadrática ainda tem melhor poder de explicação e a melhor região é a AVECO. O melhor modelo é o que apresentou o maior número de acertos, como o apresentado a seguir:

Tabela 67. Comparação entre os Métodos

Análise Discriminante	Linear	Quadrática
AD P1xP2xP3xP4	84,80%	95,50%
AD 4R	86,40%	96,20%
AD 3R AIBERxAVECOxO	87,10%	90,20%

Fonte: Elaborado pela autora

3.15 CLASSIFICAÇÃO SUPERVISIONADA E REGRESSÃO LOGÍSTICA

Seguindo com as comparações... *One-way* ANOVA: N_P_SPI_Regressao versus 4 grup or

Tabela 68.

Source	DF	SS	MS	F	P
4 grup or	3	19816,1	6605,4	128,08	0,000
Error	128	6601,4	51,6		
Total	131	26417,5			

S = 7,181 R-Sq = 75,01% R-Sq(adj) = 74,43%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	21	45,767	4,829	(-*--)(pior)
2	15	50,816	11,947	(--*--)
3	59	62,723	6,998	(*-)
4	37	80,559	6,002	(-*-)(melhor)
				----+---------+---------+---------+-----
				48 60 72 84

Fonte: Elaborado pela autora Pooled StDev = 7,181

Análise:

- Melhor: maior média: 4
- Pior: 1

Fazer o reordenamento pela média, do pior para o melhor (na planilha, já estava ordenado, do pior para o melhor).

g) Método Regressão Logística: 4 *group or versus* N_P_SPI_Regr; N_P_Deaths f; ...

Variable	Value	Count
4 grup or	1	21
	2	15
	3	59
	4	37
	Total	132

Tabela 68. Regressão Logística

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	11,3040	3,15254	3,59	0,000	
Const(2)	14,0839	3,35166	4,20	0,000	
Const(3)	20,3121	3,67679	5,52	0,000	
N_P_SPI_Regressao	-0,0594911	0,0770085	-0,77	0,440	0,94
N_P_Deaths from	-0,0315683	0,0210674	-1,50	0,134	0,97
N_P_Access to piped water	-0,0104500	0,0183303	-0,57	0,569	0,99
N_P IDH - 2013_1	-0,0646445	0,0213548	-3,03	0,002	0,94
N_P External resources	0,0009475	0,0162551	0,06	0,954	1,00
N_P GovernanÁa_1	-0,0416527	0,0303652	-1,37	0,170	0,96
N_P Happy Planet Index_1	-0,0234952	0,0140848	-1,67	0,095	0,98
N_P Outdoor air pollution	-0,0075887	0,0149461	-0,51	0,612	0,99
N_P GINNI Index_1	-0,0425037	0,0154789	-2,75	0,006	0,96
N_P Obesity rate (% of pop.)_1	-0,0035499	0,0181599	-0,20	0,845	1,00

Predictor	95% CI Lower	Upper
Const(1)		
Const(2)		

Const(3)		
N_P_SPI_Regressao	**0,81**	**1,10**
N_P_Deaths from	**0,93**	**1,01**
N_P_Access to piped water	**0,95**	**1,03**
N_P IDH - 2013_1	**0,90**	**0,98**
N_P External resources	**0,97**	**1,03**
N_P GovernanÁa_1	**0,90**	**1,02**
N_P Happy Planet Index_1	**0,95**	**1,00**
N_P Outdoor air pollution	**0,96**	**1,02**
N_P GINNI Index_1	**0,93**	**0,99**
N_P Obesity rate (% of pop.)_1	**0,96**	**1,03**

Fonte: Elaborado pela autora

Log-Likelihood = -64,908
Test that all slopes are zero: G = 201,774, DF = 10, P-Value = 0,000

Tabela 69.

Goodness-of-Fit Tests			
Method	**Chi-Square**	**DF**	**P**
Pearson	**1705,59**	**383**	**0,000**
Deviance	**129,82**	**383**	**1,000**

Fonte: Elaborado pela autora

Tabela 79.

Measures of Association:				
(Between the Response Variable and Predicted Probabilities)				
Pairs	**Number**	**Percent**	**Summary Measures**	
Concordant	**5753**	**96,6**	**Somers' D**	**0,93**
Discordant	**194**	**3,3**	**Goodman-Kruskal Gamma**	**0,93**
Ties	**7**	**0,1**	**Kendall's Tau-a**	**0,64**
Total	**5954**	**100,0**		

Fonte: Elaborado pela autora

Como há muitos indicadores com p-value acima de 10%, denotando que o modelo não é adequado, buscar-se-á um novo modelo rodando só com aquelas variáveis cujos p-values são inferiores a 10%: IDH, Happy-Planet, Ginni.

Ordinal Logistic Regression: 4 grup or versus N_P IDH - 20; N_P GINNI ; ...

Variable	Value	Count
4 grup or	1	21
	2	15
	3	59
	4	37
	Total	132

Tabela 80.

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	Upper
Const(1)	6,18662	1,01234	6,11	0,000			
Const(2)	8,52403	1,23968	6,88	0,000			
Const(3)	13,5787	1,66666	8,15	0,000			
N_P IDH - 2013_1	-0,125730	0,0158852	-7,91	0,000	0,88	0,85	0,91
N_P GINNI Index_1	-0,0397723	0,0107800	-3,69	0,000	0,96	0,94	0,98
N_P Happy Planet Index_1	-0,0236295	0,0105731	-2,23	0,025	0,98	0,96	1,00

Fonte: Elaborado pela autora

Log-Likelihood = -80,991
Test that all slopes are zero: G = 169,609, DF = 3, P-Value = 0,000

Tabela 81.

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	835,927	390	0,000
Deviance	161,982	390	1,000

Fonte: Elaborado pela autora

Tabela 82.

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	5567	93,5	Somers' D	0,87
Discordant	376	6,3	Goodman-Kruskal Gamma	0,87
Ties	11	0,2	Kendall's Tau-a	0,60
Total	5954	100,0		

Fonte: Elaborado pela autora

Agora, rodando só IDH e Ginni, que deram o p-value zero, ou seja, garantindo a confiança do modelo, tem-se: (*Ordinal Logistic Regression: 4 grup or versus* N_P IDH - 20; N_P GINNI)

Variable	Value	Count
4 grup or	1	21
	2	15
	3	59
	4	37
	Total	132

Tabela 83.

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	Upper
Const(1)	5,25625	0,896773	5,86	0,000			
Const(2)	7,22291	1,01777	7,10	0,000			
Const(3)	12,2748	1,50529	8,15	0,000			
N_P IDH - 2013_1	-0,126952	0,0154312	-8,23	0,000	0,88	0,85	0,91
N_P GINNI Index_1	-0,0371605	0,0107360	-3,46	0,001	0,96	0,94	0,98

Fonte: Elaborado pela autora

```
Log-Likelihood = -83,510
Test that all slopes are zero: G = 164,571, DF = 2, P-Value = 0,000
```

Tabela 84.

Goodness-of-Fit Tests			
Method	Chi-Square	DF	P
Pearson	825,257	391	0,000
Deviance	167,020	391	1,000

Fonte: Elaborado pela autora

Tabela 85.

Measures of Association:			
(Between the Response Variable and Predicted Probabilities)			
Pairs	Number	Percent	Summary Measures
Concordant	5552	93,2	Somers' D 0,87
Discordant	384	6,4	Goodman-Kruskal Gamma 0,87
Ties	18	0,3	Kendall's Tau-a 0,60
Total	5954	100,0	

Fonte: Elaborado pela autora

Agora, rodando-se só com o IDH.

Ordinal Logistic Regression: 4 grup or versus N_P IDH - 2013_1

Variable	Value	Count
4 grup or	1	21
	2	15
	3	59
	4	37
	Total	132

Tabela 86. *Logistic Regression Table*

					Odds	95% CI	
Predictor	Coef	SE Coef	Z	P	Ratio	Lower	Upper
Const(1)	3,25203	0,597406	5,44	0,000			
Const(2)	5,16639	0,728314	7,09	0,000			
Const(3)	9,62895	1,10834	8,69	0,000			
N_P IDH - 2013_1	-0,125273	0,0146514	-8,55	0,000	0,88	0,86	0,91

Fonte: Elaborado pela autora

```
Log-Likelihood = -90,211
Test that all slopes are zero: G = 151,169, DF = 1, P-Value = 0,000
```

Tabela 87.

Goodness-of-Fit Tests			
Method	Chi-Square	DF	P
Pearson	1175,00	344	0,000
Deviance	166,50	344	1,000

Fonte: Elaborado pela autora

Tabela 88. *Measures of Association:*
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	5466	91,8	Somers' D 0,84
Discordant	458	7,7	Goodman-Kruskal Gamma 0,85
Ties	30	0,5	Kendall's Tau-a 0,58
Total	5954	100,0	

Fonte: Elaborado pela autora

Análise:
O melhor modelo encontrado para o Método de Regressão Logística foi rodando-se IDH e Ginni, com um percentual de "*Concordant*" de 93,2% e p-value=0.

Uma vez que se encontrou o melhor modelo através do Método de Regressão Logística com o IDH e GINNI, parte-se-á para o reordenamento das 3 regiões. Recomenda-se que a COLUNA do Minitab 3R_OR seja reclassificada só com 3 grupos de clusters, do pior para o melhor: 1-Others, 2-Aiber, 3- Aveco.

Utilizando-se esta coluna para rodar a Regressão Logística, considerando IDH e Ginni como variáveis dependentes, tem-se:

Ordinal Logistic Regression: 3RORD versus N_P IDH - 20; N_P GINNI

```
Response Information
Variable  Value  Count
3RORD     1         84
          2         21
          3         27
          Total    132
```

Tabela 89. *Logistic Regression Table*

					Odds	95% CI	
Predictor	**Coef**	**SE Coef**	**Z**	**P**	**Ratio**	**Lower**	**Upper**
Const(1)	**6,00725**	**1,09913**	**5,47**	**0,000**			
Const(2)	**7,28647**	**1,17726**	**6,19**	**0,000**			
N_P IDH - 2013_1	**-0,0948732**	**0,0161924**	**-5,86**	**0,000**	**0,91**	**0,88**	**0,94**
N_P GINNI Index_1	**0,0109634**	**0,0106544**	**1,03**	**0,303**	**1,01**	**0,99**	**1,03**

Fonte: Elaborado pela autora

```
Log-Likelihood = -85,976
Test that all slopes are zero: G = 66,885, DF = 2, P-Value = 0,000
```

Tabela 90. *Goodness-of-Fit Tests*

Method	**Chi-Square**	**DF**	**P**
Pearson	**372,333**	**260**	**0,000**
Deviance	**171,952**	**260**	**1,000**

Fonte: Elaborado pela autora

Tabela 91. *Measures of Association*
(Between the Response Variable and Predicted Probabilities)

Pairs	**Number**	**Percent**	**Summary Measures**	
Concordant	**3888**	**84,5**	**Somers' D**	**0,69**
Discordant	**703**	**15,3**	**Goodman-Kruskal Gamma**	**0,69**
Ties	**8**	**0,2**	**Kendall's Tau-a**	**0,37**
Total	**4599**	**100,0**		

Fonte: Elaborado pela autora

Análise:
Como o p-value de Ginni (0,303) foi acima de 0,10, ele foi retirado, pois o modelo não é adequado, não é confiável. Rodando novamente só com IDH, teve-se:

Ordinal Logistic Regression: 3RORD versus N_P IDH - 2013_1

```
Link Function: Logit
Response Information

Variable  Value  Count
3RORD     1         84
          2         21
          3         27
          Total    132
```

Tabela 92. *Logistic Regression Table*

					Odds	95% CI	
Predictor	Coef	SE Coef	Z	P	Ratio	Lower	Upper
Const(1)	6,23975	1,04512	5,97	0,000			
Const(2)	7,51060	1,12699	6,66	0,000			
N_P IDH - 2013_1	-0,0884634	0,0145867	-6,06	0,000	0,92	0,89	0,94

Fonte: Elaborado pela autora

```
Log-Likelihood = -86,523
Test that all slopes are zero: G = 65,791, DF = 1, P-Value = 0,000
```

Tabela 93. *Goodness-of-Fit Tests*

Method	Chi-Square	DF	P
Pearson	270,346	229	0,032
Deviance	155,168	229	1,000

Fonte: Elaborado pela autora

Tabela 94. *Measures of Association*
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	3839	83,5	Somers' D	0,67
Discordant	744	16,2	Goodman-Kruskal Gamma	0,68
Ties	16	0,3	Kendall's Tau-a	0,36
Total	4599	100,0		

Fonte: Elaborado pela autora

Análise:
Aqui, crê-se que se pode satisfazer com essa variável (IDH), não havendo necessidade de gerar outras regressões logísticas, como por exemplo, a regressão logística nominal. O p-value=0, e o percentual de *Concordant equivale* a 83,5%, o que é um bom percentual.

Observa-se que a análise multivariada aplicada aos dados que melhor explica a saúde e bem-estar nos países pesquisados refere-se ao Modelo apresentado pela Análise Discriminante Quadrática. O modelo de 4 clusters também melhor explica os dados. No entanto, uma vez que se trata de um modelo de complexa aplicação e a Linear (Ver Tabela 67.) também foi bastante acorde com um bom grau de acertos, talvez se pudesse aplicá-la.

Tabela 95. Tabela Final Comparativa AD e RL

	4R		3R
AD (Quadrática)	**Grau de Acertos**	96,20%	90,20%
RL (Regressão Logística)	**Concordance**	93,20%	83,50%

Fonte: Elaborado pela autora

3.16 CLASSIFICAÇÃO SUPERVISIONADA E ÁRVORES DE DECISÃO

Segue-se com a análise multivariada dos dados compilados pelo Banco Mundial, porém, muda-se para a utilização do *software* SPSS. A finalidade é apresentar análise discriminante (AD), regressão logística e árvores de decisão para os dados componentes mais importantes na explicação da Saúde e Meio Ambiente dos 132 países analisados. A princípio, a análise se fará relativa aos 4 grupos de variáveis para todos os países, e depois, classificando-se os países pelo seu desenvolvimento: AIBER – Grupos de países IBERO AMERICANOS, AVECO – Grupo de Países Desenvolvidos, OTHERS – Outros países. Objetiva-se comparar qual das análises apresenta melhor modelo explicativo.

Como já foi dito, as técnicas multivariadas, entre outras coisas, ajudam a identificar se os dados são válidos e verificar as características distintas dos indivíduos em cada grupo, estimando o grupo ao que pertencem outros indivíduos. Neste etapa do estudo, será feita análise discriminante e agrega-se como novidade a árvore de decisão que permite estabelecer pretensos relacionamentos entre os grupos e predizer eventos futuros, de forma inclusive bastante visual, com o uso da nova ferramenta estatística, o SPSS.

3.16.1 O SPSS

Segundo a IBM®, o SPSS® *Decision Trees* auxilia a identificar facilmente grupos e resultados previstos. Uma vez identificados, é possível apresentar resultados categóricos de maneira intuitiva. Ele fornece técnicas de crescimento em árvore especializadas para classificação com quatro algoritmos:

- CHAID: estatístico e de várias vias que explora dados com rapidez e eficiência e constrói segmentos e perfis com relação ao resultado desejado.
- CHAID Exaustivo: uma modificação do CHAID, que examina todas as divisões possíveis para cada variável preditora.
- Classification and regression trees (C&RT): um algoritmo de árvore binária completo que particiona dados e produz subgrupos homogêneos precisos.
- QUEST: estatístico, que seleciona variáveis sem propensão e constrói árvores binárias precisas com rapidez e eficiência.

O método escolhido devido à sua funcionalidade foi o Método CHAID.

3.16.2 ANÁLISE DOS NOVOS COMPONENTES

Na análise dos componentes principais, com base nas 20 variáveis pré-selecionadas para os países em questão, uma vez criados PC1, PC2, PC3 e PC4, e classificando-os pelas regiões AIBER, AVECO e Others, pôde-se perceber:

Tabela 96: Resumo Tabela 67 e Tabela 95

ANÁLISE DISCRIMINANTE	LINEAR	QUADRÁTICA	RL*
AD 4R	86,40%	96,20%	**93,20% (Concord.)**
AD 3R	87,10%	90,20%	83,50% (Concord.)
AIBERxAVECOxOt			

Fonte: Elaborado pela autora * *RL – Regressão Logística; AD – Análise Discriminante*

Depois do Stepwise, as principais variáveis dos componentes PC1, PC2, PC3 e PC4 foram:

1. *Deaths from Infectious Desease,*
2. *Access to Piped Water,*
3. *Obesity Rate,*
4. *Outdoor Air Pollution,*
5. IDH**,
6. Governança,
7. EPI Score,
8. *Happy Planet,*
9. OHI,
10. Ginni.

**Relembra-se que o *Social Progress* é um indicador com o maior poder explicativo em relação às demais variáveis, porém, com uma grande dificuldade operacional e de grande complexidade. Por essa razão, decidiu-se substitui-lo pelo IDH, que é uma variável que o representa bastante bem, e de mais fácil aplicabilidade.

Dando continuidade à análise, fazendo a Análise de Discriminante, através do uso do software SPSS, tem-se: a variável 3R_ORD (1; 3) como dependente e as outras 10 mostradas acima como variáveis independentes, com a opção do spss “enter” escolhido, gerando os seguintes resultados:

Tabela 96.

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		132	100,0
Excluded	Missing or out-of-range group codes	0	,0
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	0	,0
Total		132	100,0

Fonte: Elaborado pela autora

Análise: A tabela acima mostra que a base de dados possui um total de 132 países pesquisados e não possui dados faltantes (*missings*), uma vez que tem 100% de dados válidos.

Tabela 97.

Group Statistics

@3RORD		Mean	Std. Deviation	Valid N (listwise) Unweighted	Valid N (listwise) Weighted
1,00	N_P_Deaths_from	70,6343	28,82890	84	84,000
	N_P_Access_to_piped_water	47,2670	33,82187	84	84,000
	N_P_Obesity_rate__of_pop#_1	66,8751	24,72358	84	84,000
	N_P_Outdoor_air_pollution	74,7601	21,56769	84	84,000
	N_P_IDH__2013_1	47,4633	23,98084	84	84,000
	N_P_Governança_1	33,3613	16,34673	84	84,000
	N_P_EPI_Score_1	37,9500	19,06338	84	84,000
	N_P_Happy_Planet_Index_1	41,7945	20,67539	84	84,000
	N_P___OHI_1	51,0698	16,77312	84	84,000
	N_P___GINNI_Index_1	62,4368	20,31560	84	84,000
2,00	N_P_Deaths_from	92,8916	4,92402	21	21,000
	N_P_Access_to_piped_water	85,2194	11,80871	21	21,000
	N_P_Obesity_rate__of_pop#_1	46,6827	10,68488	21	21,000
	N_P_Outdoor_air_pollution	82,0896	11,36683	21	21,000
	N_P_IDH__2013_1	65,6154	12,03438	21	21,000
	N_P_Governança_1	42,4836	18,82010	21	21,000
	N_P_EPI_Score_1	52,6987	14,31093	21	21,000
	N_P_Happy_Planet_Index_1	72,3562	16,25029	21	21,000
	N_P___OHI_1	55,4546	19,03183	21	21,000
	N_P___GINNI_Index_1	43,0841	14,39843	21	21,000
3,00	N_P_Deaths_from	98,8984	,59816	27	27,000
	N_P_Access_to_piped_water	94,8947	10,29073	27	27,000
	N_P_Obesity_rate__of_pop#_1	53,7348	14,60695	27	27,000
	N_P_Outdoor_air_pollution	77,5014	13,38124	27	27,000
	N_P_IDH__2013_1	85,1132	15,53183	27	27,000
	N_P_Governança_1	83,2524	13,38237	27	27,000
	N_P_EPI_Score_1	81,3938	8,14104	27	27,000
	N_P_Happy_Planet_Index_1	50,5831	12,83975	27	27,000
	N_P___OHI_1	24,2353	15,47696	27	27,000
	N_P___GINNI_Index_1	79,9124	12,42135	27	27,000
Total	N_P_Deaths_from	79,9565	26,20778	132	132,000
	N_P_Access_to_piped_water	63,0469	34,85121	132	132,000
	N_P_Obesity_rate__of_pop#_1	60,9749	22,64801	132	132,000
	N_P_Outdoor_air_pollution	76,4869	18,89823	132	132,000
	N_P_IDH__2013_1	58,0523	25,81380	132	132,000
	N_P_Governança_1	45,0176	25,46339	132	132,000
	N_P_EPI_Score_1	49,1826	23,90418	132	132,000
	N_P_Happy_Planet_Index_1	48,4542	21,56002	132	132,000
	N_P___OHI_1	46,2785	20,23508	132	132,000
	N_P___GINNI_Index_1	62,9325	21,13069	132	132,000

Fonte: Elaborado pela autora

Análise:
De acordo com a tabela acima, vê-se um resumo para cada variável, com médias e desvios-padrão para cada grupo e para o total, sendo 1 = OTHERS (pior grupo, de acordo com trabalhos anteriores), 2 = AIBER e 3 = AVECO (melhor região).

Tabela 98.

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
N_P_Deaths_from	,772	19,033	2	129	,000
N_P_Access_to_piped_water	,632	37,630	2	129	,000
N_P_Obesity_rate__of_pop#_1	,872	9,504	2	129	,000
N_P_Outdoor_air_pollution	,980	1,319	2	129	,271
N_P_IDH__2013_1	,652	34,451	2	129	,000
N_P_Governança_1	,399	97,017	2	129	,000
N_P_EPI_Score_1	,481	69,680	2	129	,000
N_P_Happy_Planet_Index_1	,740	22,688	2	129	,000
N_P___OHI_1	,686	29,455	2	129	,000
N_P___GINNI_Index_1	,725	24,451	2	129	,000

Fonte: Elaborado pela autora

Análise:
Como o nome da tabela diz, aqui foi feito um teste de igualdade de média entre os grupos. Nele, foram identificadas as variáveis mais importantes para discriminar os grupos dos países (Others, Aiber e Aveco). A primeira coluna: a Wilk's Lambda varia de 0 a 1, e quanto mais próximo a 0, melhor, pois ela mede a diferença das médias entre os grupos. Assim, o melhor foi: Governança, com 0,399 , com o: > F. A pior foi: *Outdoor air pollution*, pois tem o valor maior = 0,980.
Seria possível, numa análise futura, retirar a variável *Outdoor air pollution*, pois o p-value> 0,05, sendo = 0,271.

Tabela 99.

Pooled Within-Groups Matrices

		N_P_Deaths_from	N_P_Access_to_piped_water	N_P_Obesity_rate__of_pop#_1	N_P_Outdoor_air_pollution	N_P_IDH__2013_1	N_P_Governança_1	N_P_EPI_Score_1	N_P_Happy_Planet_Index_1	N_P___OHI_1	N_P___GINNI_Index_1
Correlation	N_P_Deaths_from	1,000	,738	-,533	-,400	,691	,267	,620	,387	-,159	,543
	N_P_Access_to_piped_water	,738	1,000	-,715	-,443	,748	,453	,741	,217	-,246	,273
	N_P_Obesity_rate__of_pop#_1	-,533	-,715	1,000	,326	-,696	-,309	-,637	-,064	,167	-,170
	N_P_Outdoor_air_pollution	-,400	-,443	,326	1,000	-,306	-,045	-,286	-,003	,128	-,298
	N_P_IDH__2013_1	,691	,748	-,696	-,306	1,000	,503	,718	,257	-,307	,234
	N_P_Governança_1	,267	,453	-,309	-,045	,503	1,000	,591	-,027	-,403	-,043
	N_P_EPI_Score_1	,620	,741	-,637	-,286	,718	,591	1,000	,084	-,317	,265
	N_P_Happy_Planet_Index_1	,387	,217	-,064	-,003	,257	-,027	,084	1,000	,184	,130
	N_P___OHI_1	-,159	-,246	,167	,128	-,307	-,403	-,317	,184	1,000	-,057
	N_P___GINNI_Index_1	,543	,273	-,170	-,298	,234	-,043	,265	,130	-,057	1,000

Fonte: Elaborado pela autora

Análise:
Aqui, continua-se a análise da relação entre as variáveis. Pelo critério da parcimônia se pode excluir do modelo aquelas variáveis com correlação muito grande entre si. Correlações positivas altas (próximas de +1) indicam variáveis diretamente proporcionais e correlações negativas altas (próximas de -1) indicam variáveis inversamente proporcionais.

Quanto maior *Access to piped water* maior a Sobrevivência (Oposto de *Deaths*) from Infectious Deseases, com uma correlação entre as duas variáveis de 0,738. Talvez não fosse necessário usar ambas, e só uma.

Quanto maior o IDH, maior o *Access to piped wate*r, com uma correlação de 0,748. Talvez não fosse necessário usar ambas, e só uma. Então, provavelmente seria recomendável retirar a variável *Access to piped water*.

Tabela 100.

Test Results

Box's M		428,919
F	Approx.	3,231
	df1	110
	df2	10483,230
	Sig.	,000

Tests null hypothesis of equal population covariance matrices.

Fonte: Elaborado pela autora

Análise:
Teste de hipótese nula, com significância igual a zero (p-value).

Tabela 101.

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2,184[a]	69,6	69,6	,828
2	,953[a]	30,4	100,0	,698

a. First 2 canonical discriminant functions were used in the analysis.

Fonte: Elaborado pela autora

Tabela 102.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	,161	227,507	20	,000
2	,512	83,312	9	,000

Fonte: Elaborado pela autora

Análise:
A tabela discriminante acima: a coluna dos eigenvalues refere-se à capacidade de discriminar os dados – quanto maiores os seus valores, melhores discriminados os dados. Sendo assim, a primeira – função 1 – é melhor para dizer o quanto os grupos são similares ou não entre si, com uma variação de 69,6%. Já a segunda – função 2 – , contribui menos para dizer o quanto o grupo é similar, com um grau de variação de 30,4%. Cada função foi significativa, com o p-value = 0.

Tabela 103.

Structure Matrix

	Function	
	1	2
N_P_Governança_1	,818*	,214
N_P_EPI_Score_1	,668*	,332
N_P___OHI_1	-,453*	,093
N_P_IDH__2013_1	,446*	,323
N_P_Happy_Planet_Index_1	,038	,605*
N_P_Access_to_piped_water	,394	,506*
N_P___GINNI_Index_1	,326	-,392*
N_P_Deaths_from	,282	,357*
N_P_Obesity_rate__of_pop#_1	-,117	-,351*
N_P_Outdoor_air_pollution	,018	,144*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Fonte: Elaborado pela autora

Análise:
Pode-se observar que para a primeira função as 4 primeiras variáveis foram as mais importantes. Já para a segunda função, as 6 últimas.

Outra tabela que o SPSS nos fornece é a dos coeficientes das funções discriminantes, dada a seguir:

Tabela 104.

Canonical Discriminant Function Coefficients

	Function	
	1	2
N_P_Deaths_from	-,018	,016
N_P_Access_to_piped_water	,005	,021
N_P_Obesity_rate__of_pop#_1	,019	-,006
N_P_Outdoor_air_pollution	,008	,016
N_P_IDH__2013_1	,005	-,018
N_P_Governança_1	,038	-,004
N_P_EPI_Score_1	,031	,010
N_P_Happy_Planet_Index_1	,005	,027
N_P___OHI_1	-,007	,003
N_P___GINNI_Index_1	,026	-,038
(Constant)	-5,753	-1,629

Unstandardized coefficients

Fonte: Elaborado pela autora

Análise:
Cada variável e seu valor respectivo para cada coeficiente padrão das funções 1 e 2.

Tabela 105.

Classification Function Coefficients

	@3RORD		
	1,00	2,00	3,00
N_P_Deaths_from	,033	,076	-,023
N_P_Access_to_piped_water	,220	,276	,250
N_P_Obesity_rate__of_pop#_1	,551	,535	,618
N_P_Outdoor_air_pollution	,333	,375	,369
N_P_IDH__2013_1	,311	,261	,319
N_P_Governança_1	,035	,023	,171
N_P_EPI_Score_1	,145	,171	,262
N_P_Happy_Planet_Index_1	-,105	-,033	-,071
N_P___OHI_1	,324	,332	,300
N_P___GINNI_Index_1	,200	,096	,273
(Constant)	-60,704	-68,580	-87,285

Fisher's linear discriminant functions

Fonte: Elaborado pela autora

Análise:
Os dados acima apresentam as variáveis do *Stepwise* (as 10 citadas), utilizadas para a Análise Discriminante, e os valores referentes a cada função divididas para cada grupo de países: 1- Others, 2- Aiber; 3- AVECO.

A última tabela apresentada pelo programa é a tabela de classificação dos resultados, conforme abaixo.

Tabela 106.

Classification Results[a]

			Predicted Group Membership			
		@3RORD	1,00	2,00	3,00	Total
Original	Count	1,00	79	2	3	84
		2,00	2	17	2	21
		3,00	2	1	24	27
	%	1,00	94,0	2,4	3,6	100,0
		2,00	9,5	81,0	9,5	100,0
		3,00	7,4	3,7	88,9	100,0

a. 90,9% of original grouped cases correctly classified.

Fonte: Elaborado pela autora

Análise:
Esta tabela apresenta os erros. Sendo assim, 2 (9,5%) de Aiber e 2 (7,4%) de AVECO deveriam ter sido classificados como Others. 2 (2,4%) e 1 (3,7%) que foram classificados como Others e AVECO, respectivamente, deveriam ter sido classificados como AIBER. E 3 (3,6%) de Others e 2 (9,5%) de Aiber deveriam ter sido classificados como AVECO.

3.16.3 ÁRVORE DE DECISÕES

A partir daqui, através do recurso estatística do SPSS, buscou-se gerar Árvore de Decisão, mas o SPSS não deu resultados satisfatórios, proporcionando as seguintes análises:

Quadro 7.

```
GET
  FILE='C:\Users\aofu117407\Documents\Dados Minitab Saude Bem Estar_1.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
* Decision Tree.
TREE @3RORD [n] BY Deaths_from_Infectious_Deseases [s] Access_to_piped_water [s] Ob
esity_Rate [s]
    Outdoor_Air_Pollution [s] IDH [s] Governanca [s] EPI_Score [s] Happy_Planet [s]
 OHI [s] Ginni [s]
  /TREE DISPLAY=LEFTTORIGHT NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCAL
E=AUTO
  /DEPCATEGORIES USEVALUES=[VALID]
  /PRINT MODELSUMMARY CLASSIFICATION RISK
  /METHOD TYPE=CHAID
  /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO CHISQUARE=PEARSON CONVERGE=
0.001
    MAXITERATIONS=100 ADJUST=BONFERRONI INTERVALS=10.
```

Fonte: Elaborado pela autora

Classification Tree

[DataSet1] C:\Users\aofu117407\Documents\Dados Minitab Saude Bem Estar_1.sav

Warnings

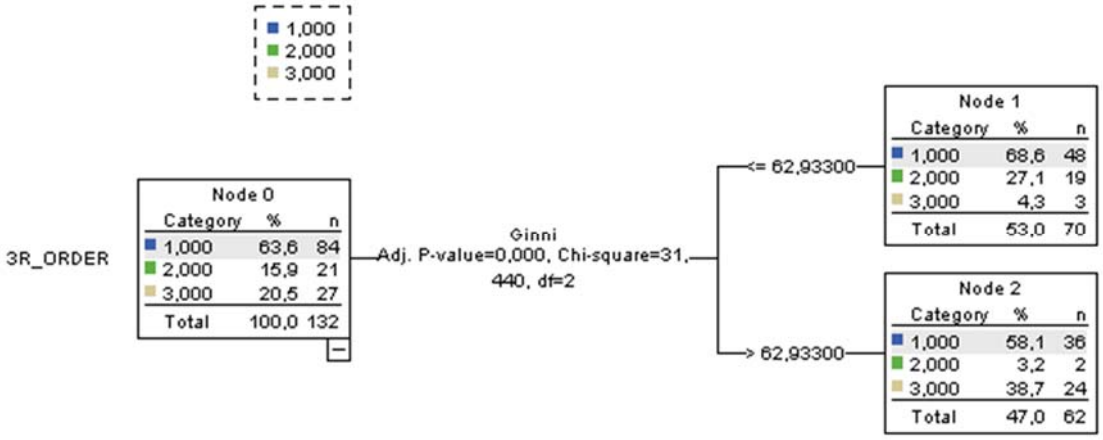
Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	3R_ORDER
	Independent Variables	Deaths_from_Infectious_Deseases, Access_to_piped_water, Obesity_Rate, Outdoor_Air_Pollution, IDH, Governanca, EPI_Score, Happy_Planet, OHI, Ginni
	Validation	None
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	100
	Minimum Cases in Child Node	50
Results	Independent Variables Included	Ginni
	Number of Nodes	3
	Number of Terminal Nodes	2
	Depth	1

Fonte: Elaborado pela autora

Árvore de Decisão 1.



Fonte: Elaborado pela autora

Tabela 107.

Risk

Estimate	Std. Error
,364	,042

Growing Method: CHAID
Dependent Variable: 3R_ORDER

Fonte: Elaborado pela autora

Tabela 108.

Classification					
Observed		Predicted			
1,00	2,00	3,00	Percent Correct		
1,00	84	0	0	100,0%	
2,00	21	0	0	0,0%	
3,00	27	0	0	0,0%	
Overall Percentage			100,0%	0,0%	0,0% 63,6%
Growing Method: CHAID					
Dependent Variable: 3R_ORDER					

Fonte: Elaborado pela autora

Análise:

Como resultado da Árvore de Decisões, infelizmente não foi possível encontrar resultados satisfatórios. A única variável encontrada nesta etapa foi o Ginni, o que dificulta qualquer análise por meio deste método CHAID, com os dados disponíveis.

3.17 ANÁLISE DE CORRESPONDÊNCIAS

Cálculo do ANOVA para cada uma, e obtenção da média por região, para cada uma das variáveis do *Stepwise*:

Tabela 109. *One-way* ANOVA (N_P IDH - 2013_1 versus 3RORD)

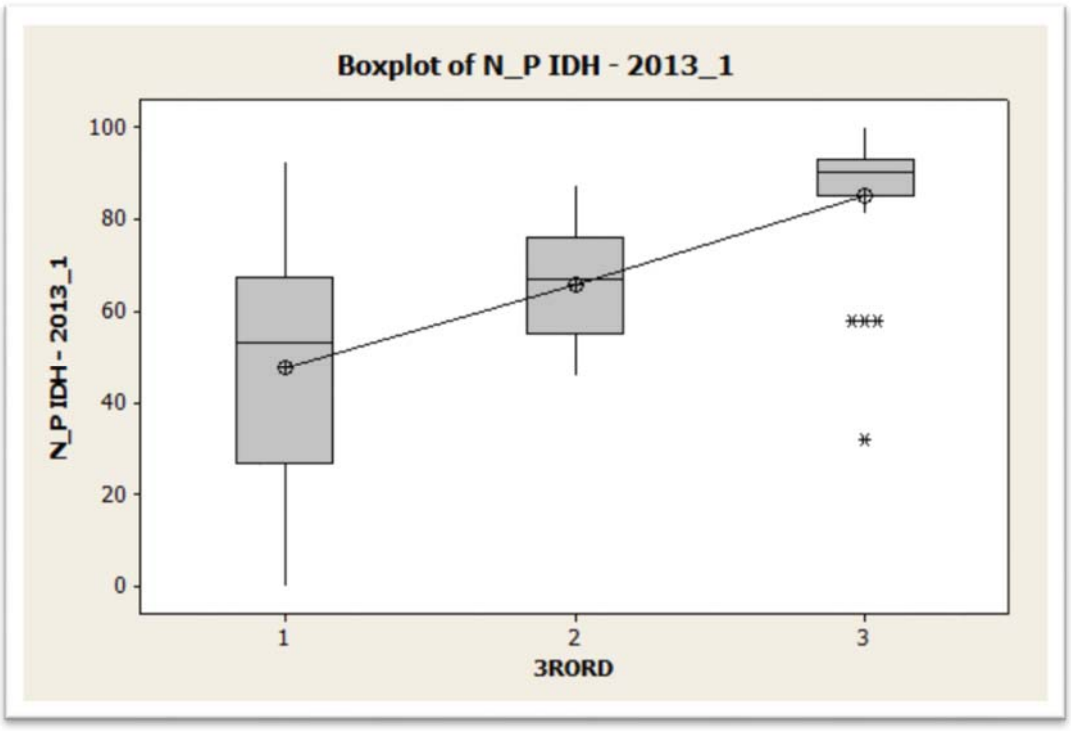
Source	DF	SS	MS	F	P
3RORD	2	30392	15196	34,45	0,000
Error	129	56900	441		
Total	131	87292			

S = 21,00 R-Sq = 34,82% R-Sq(adj) = 33,81%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	-+---------+---------+---------+--------
1	84	47,46	23,98	(--*--)
2	21	65,62	12,03	(-----*-----)
3	27	85,11	15,53	(-----*----)
				-+---------+---------+---------+--------
				45 60 75 90

Fonte: Elaborado pela autora Pooled StDev = 21,00

Gráfico 120. Boxplot of N_P IDH - 2013_1



Fonte: Elaborado pela autora

Tabela 110. *One-way* ANOVA: N_P Obesity rate (% of pop.)_1 versus 3RORD

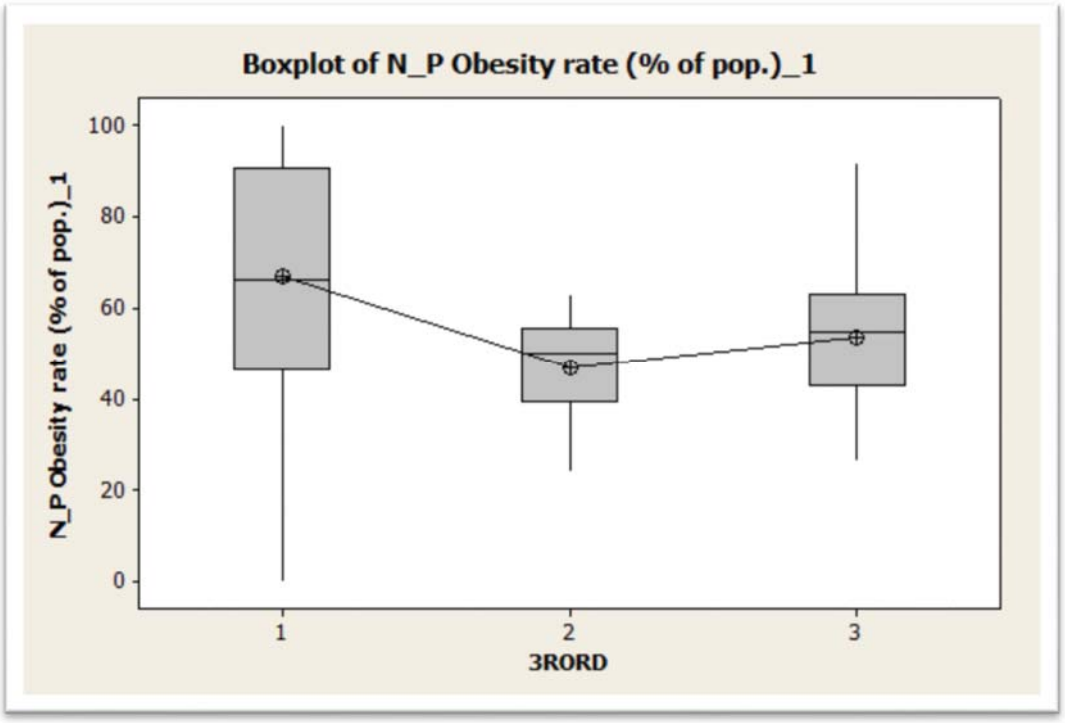
Source	DF	SS	MS	F	P
3RORD	2	8629	4315	9,50	0,000
Error	129	58565	454		
Total	131	67194			

S = 21,31 R-Sq = 12,84% R-Sq(adj) = 11,49%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	84	66,88	24,72	(----*---)
2	21	46,68	10,68	(---------*--------)
3	27	53,73	14,61	(-------*-------)
				40 50 60 70

Fonte: Elaborado pela autora Pooled StDev = 21,31

Gráfico 121. Boxplot of N_P Obesity rate (% of pop.)_1



Fonte: Elaborado pela autora

Tabela 111. *One-way ANOVA: N_P_Deaths from versus 3RORD*

Source	DF	SS	MS	F	P
3RORD	2	20501	10251	19,03	0,000
Error	129	69476	539		
Total	131	89977			

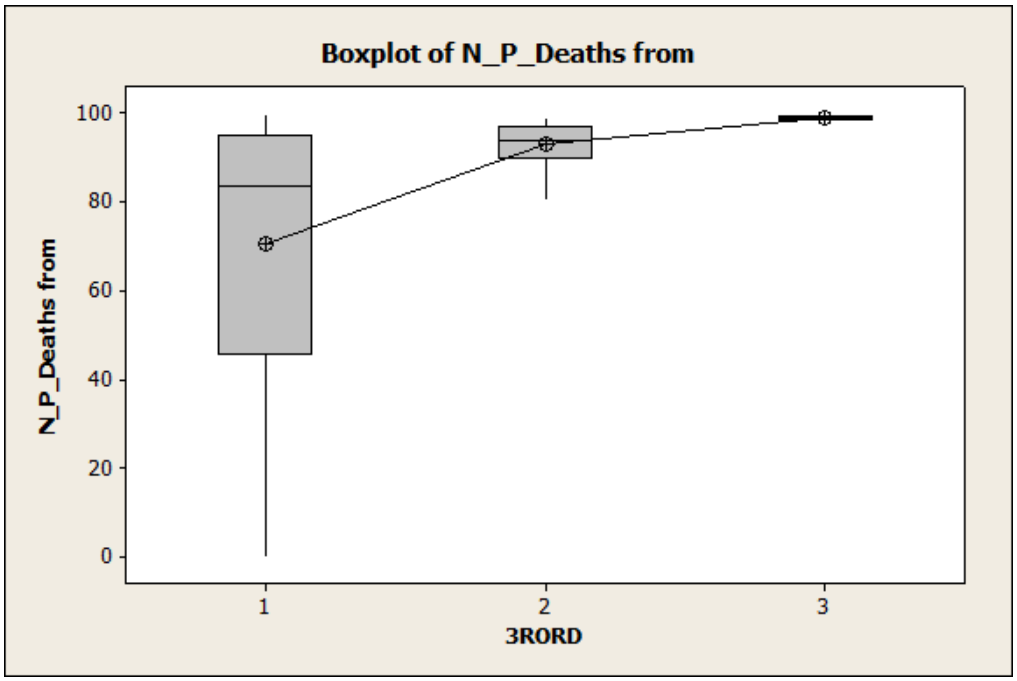
S = 23,21 R-Sq = 22,78% R-Sq(adj) = 21,59%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	-----+---------+---------+---------+----
1	84	70,63	28,83	(---*---)
2	21	92,89	4,92	(-------*--------)
3	27	98,90	0,60	(------*-------)
				-----+---------+---------+---------+----
				72 84 96 108

Fonte: Elaborado pela autora `Pooled StDev = 23,21`

Gráfico 122. Boxplot of N_P_Deaths from



Fonte: Elaborado pela autora

Tabela 112. *One-way ANOVA: N_P_Access to piped water versus 3RORD*

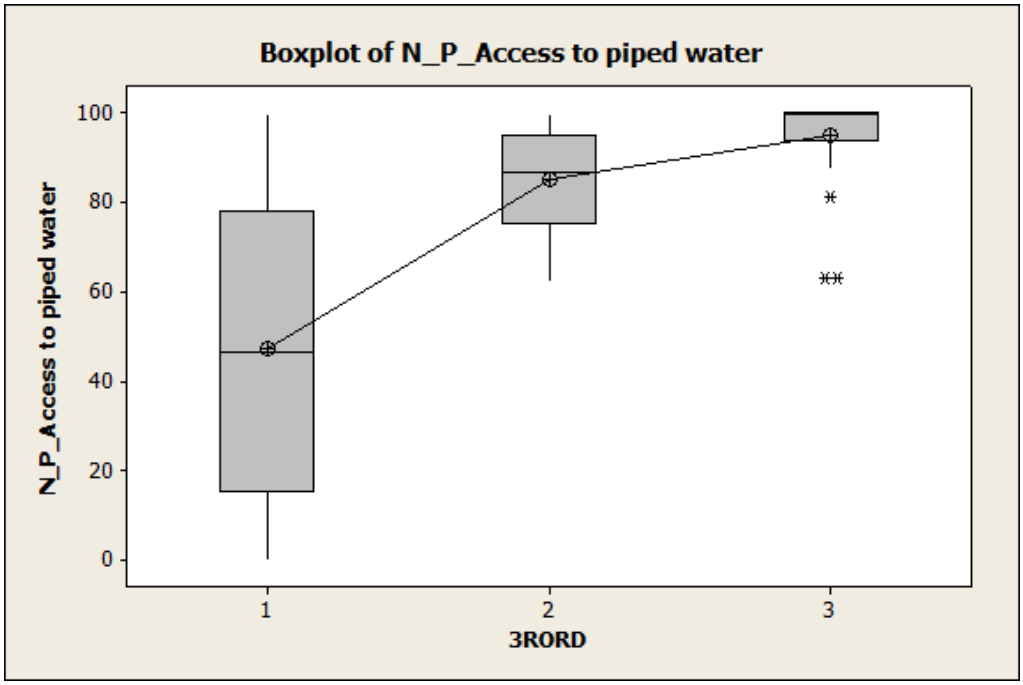
Source	DF	SS	MS	F	P
3RORD	2	58626	29313	37,63	0,000
Error	129	100488	779		
Total	131	159113			

S = 27,91 R-Sq = 36,85% R-Sq(adj) = 35,87%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	---------+---------+---------+---------+
1	84	47,27	33,82	(--*--)
2	21	85,22	11,81	(-----*-----)
3	27	94,89	10,29	(----*-----)
				---------+---------+---------+---------+
				60 80 100 120

Fonte: Elaborado pela autora `Pooled StDev = 27,91`

Gráfico 123. Boxplot of N_P_Access to piped water



Fonte: Elaborado pela autora

Tabela 113. *One-way ANOVA: N_P Outdoor air pollution versus 3RORD*

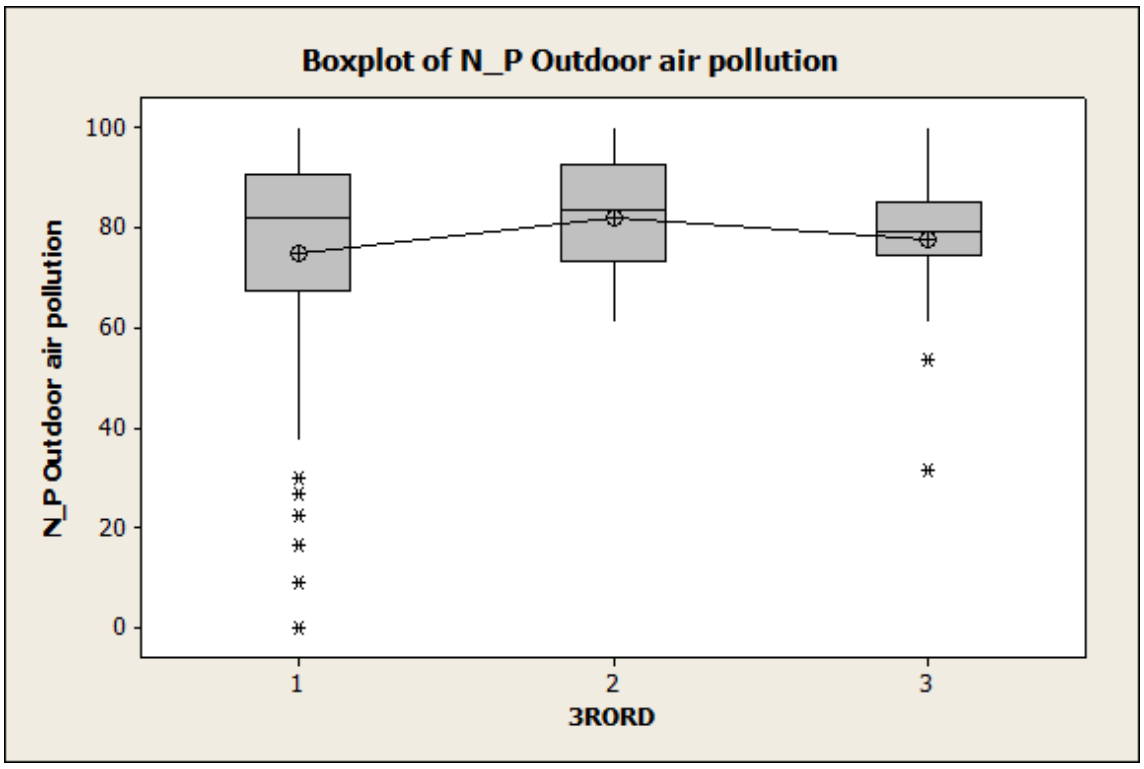
Source	DF	SS	MS	F	P
3RORD	2	937	469	1,32	0,271
Error	129	45848	355		
Total	131	46786			

S = 18,85 R-Sq = 2,00% R-Sq(adj) = 0,48%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	---------+---------+---------+---------+
1	84	74,76	21,57	(--------*-------)
2	21	82,09	11,37	(--------------*--------------)
3	27	77,50	13,38	(-------------*-------------)
				---------+---------+---------+---------+
				75,0 80,0 85,0 90,0

Fonte: Elaborado pela autora Pooled StDev = 18,85

Gráfico 124. Boxplot of N_P Outdoor air pollution



Fonte: Elaborado pela autora

Tabela 114. *One-way ANOVA:* N_P GovernanÁa_1 versus 3RORD

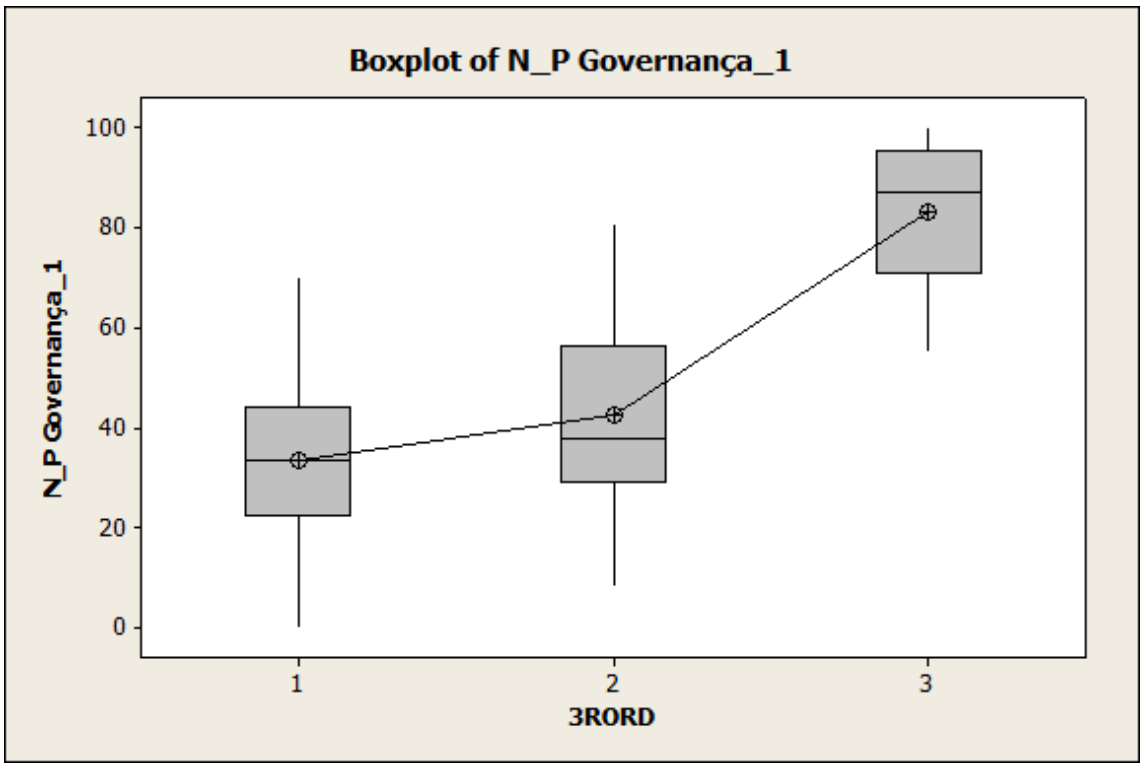
Source	DF	SS	MS	F	P
3RORD	2	51019	25510	97,02	0,000
Error	129	33919	263		
Total	131	84938			

S = 16,22 R-Sq = 60,07% R-Sq(adj) = 59,45%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	+---------+---------+---------+---------
1	84	33,36	16,35	(-*--)
2	21	42,48	18,82	(---*----)
3	27	83,25	13,38	(----*---)
				+---------+---------+---------+---------
				30 45 60 75

Fonte: Elaborado pela autora Pooled StDev = 16,22

Gráfico 125. Boxplot of N_P Governança



Fonte: Elaborado pela autora

Tabela 115. *One-way* ANOVA: N_P EPI Score_1 versus 3RORD

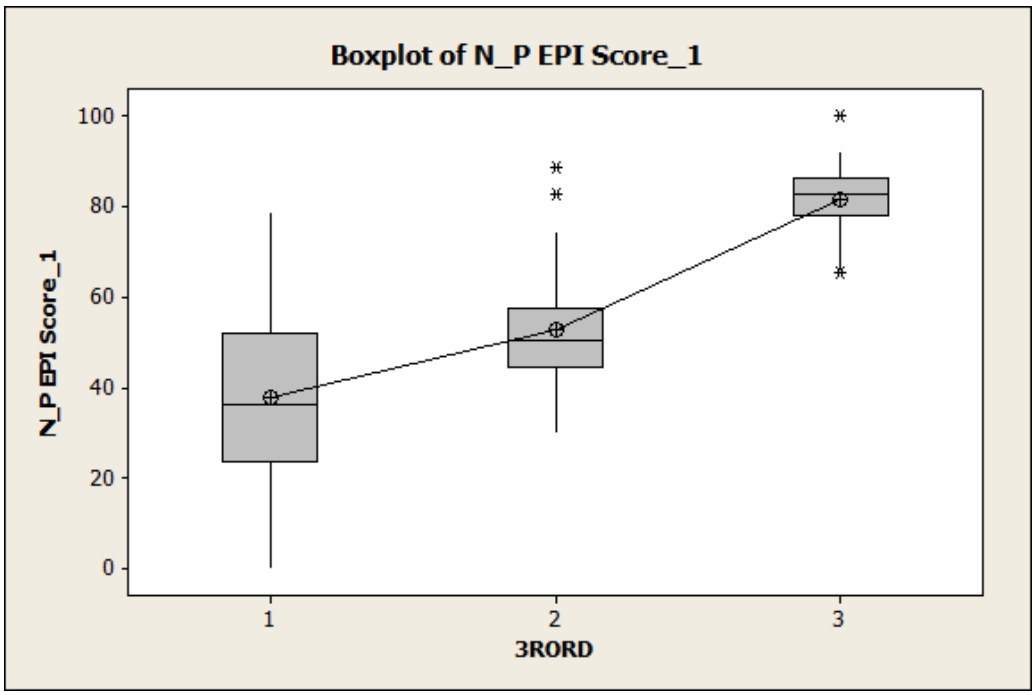
Source	DF	SS	MS	F	P
3RORD	2	38872	19436	69,68	0,000
Error	129	35982	279		
Total	131	74855			

S = 16,70 R-Sq = 51,93% R-Sq(adj) = 51,18%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	-------+---------+---------+---------+--
1	84	37,95	19,06	(-*--)
2	21	52,70	14,31	(----*----)
3	27	81,39	8,14	(---*----)
				-------+---------+---------+---------+--
				45 60 75 90

Fonte: Elaborado pela autora Pooled StDev = 16,70

Gráfico 126. Boxplot of N_P EPI Score_1



Fonte: Elaborado pela autora

Tabela 116. *One-way ANOVA: N_P Happy Planet Index_1 versus 3RORD*

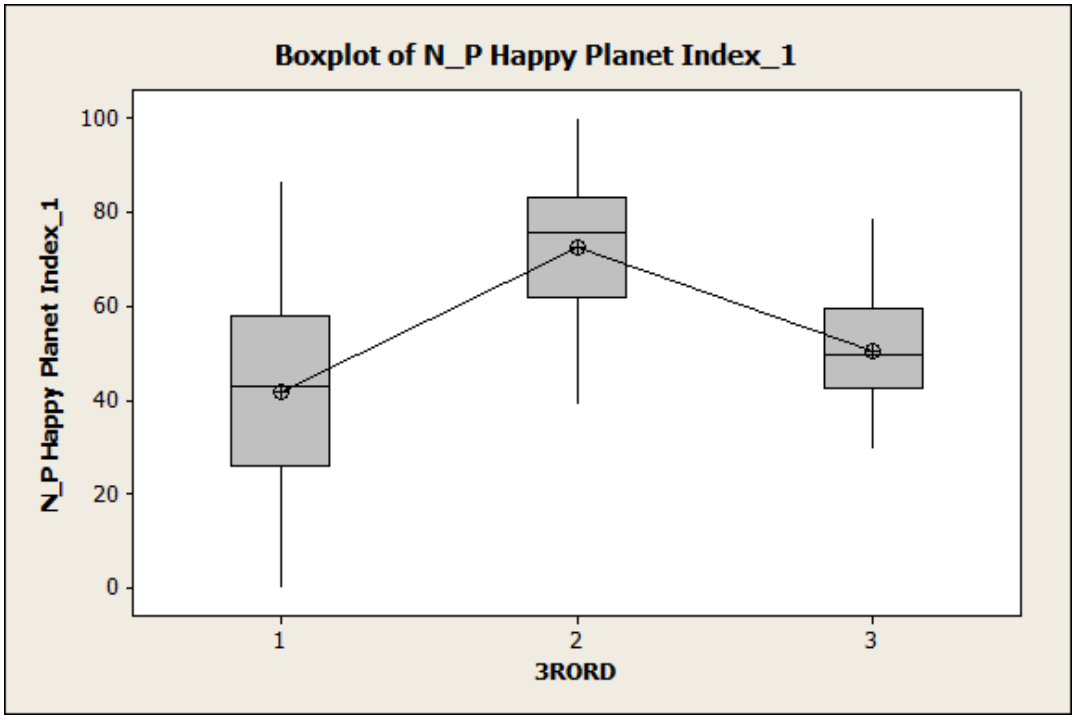
Source	DF	SS	MS	F	P
3RORD	2	15845	7923	22,69	0,000
Error	129	45048	349		
Total	131	60893			

S = 18,69 R-Sq = 26,02% R-Sq(adj) = 24,87%

				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	---------+---------+---------+---------+
1	84	41,79	20,68	(---*--)
2	21	72,36	16,25	(-----*------)
3	27	50,58	12,84	(-----*-----)
				---------+---------+---------+---------+
				48 60 72 84

Fonte: Elaborado pela autora `Pooled StDev = 18,69`

Gráfico 127. Boxplot of N_P Happy Planet Index_1



Fonte: Elaborado pela autora

Tabela 117. *One-way ANOVA: N_P OHI_1 versus 3RORD*

Source	DF	SS	MS	F	P
3RORD	2	16816	8408	29,45	0,000
Error	129	36823	285		
Total	131	53639			

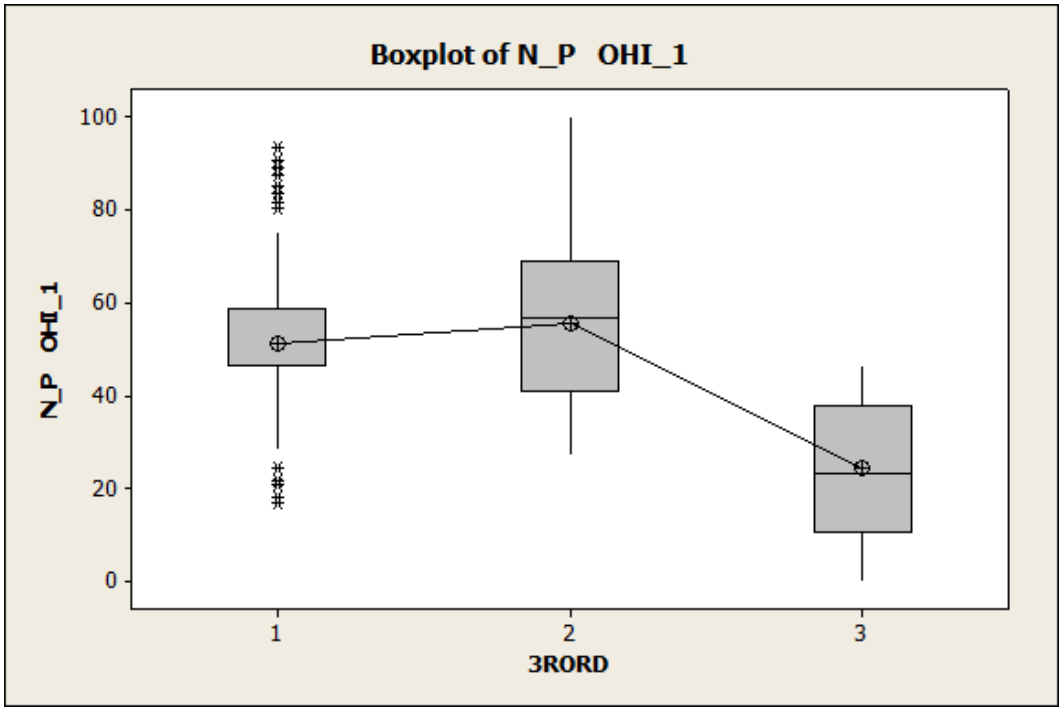
S = 16,90 R-Sq = 31,35% R-Sq(adj) = 30,29%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	-----+---------+---------+---------+----
1	84	51,07	16,77	(--*--)
2	21	55,45	19,03	(-----*-----)
3	27	24,24	15,48	(----*-----)
				-----+---------+---------+---------+----
				24 36 48 60

Fonte: Elaborado pela autora Pooled StDev = 16,90

Gráfico 128. Boxplot of N_P OHI_1



Fonte: Elaborado pela autora

Tabela 118. *One-way ANOVA: N_P GINNI Index_1 versus 3RORD*

Source	DF	SS	MS	F	P
3RORD	2	16078	8039	24,45	0,000
Error	129	42414	329		
Total	131	58492			

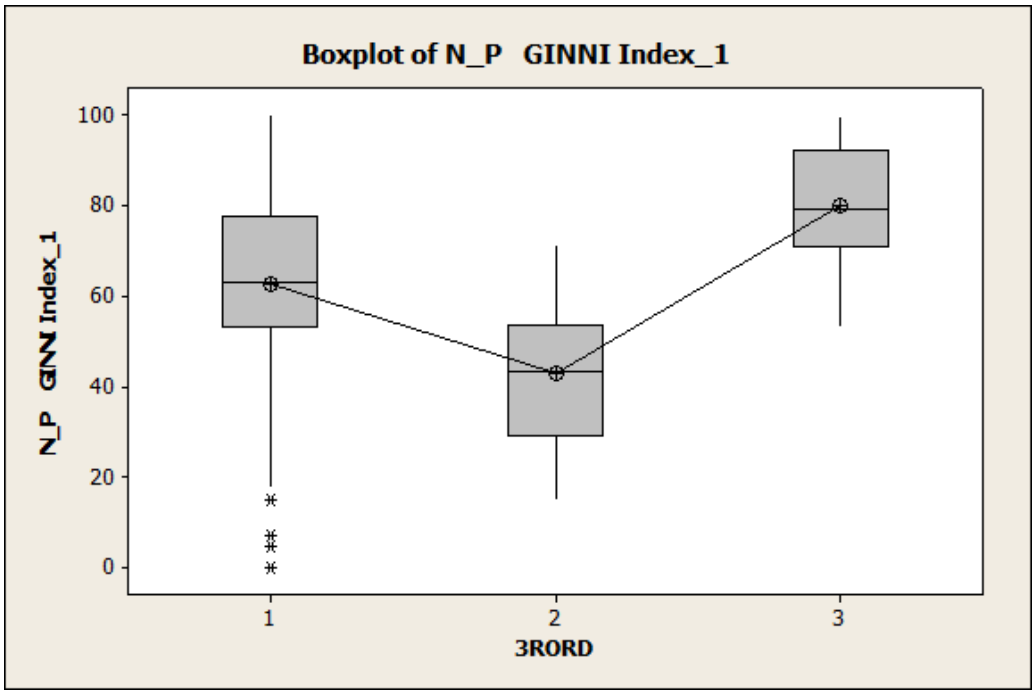
S = 18,13 R-Sq = 27,49% R-Sq(adj) = 26,36%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	------+---------+---------+---------+---
1	84	62,44	20,32	(--*-)
2	21	43,08	14,40	(----*----)
3	27	79,91	12,42	(---*----)
				------+---------+---------+---------+---
				45 60 75 90

Fonte: Elaborado pela autora Pooled StDev = 18,13

Gráfico 129. Boxplot of N_P GINNI Index_1



Fonte: Elaborado pela autora

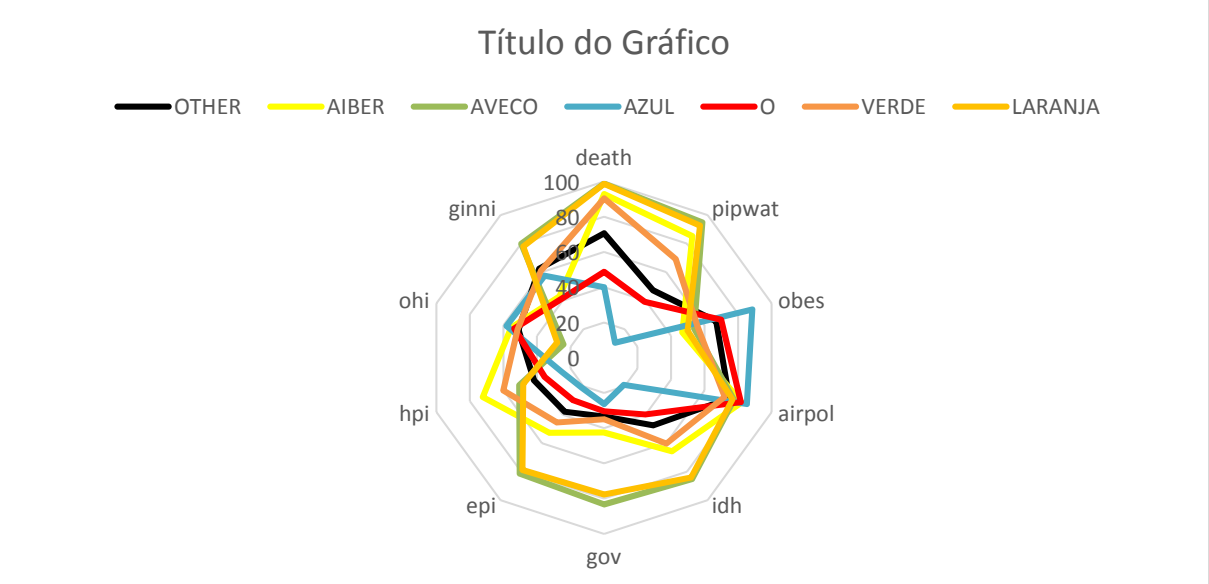
Após o cálculo das médias por região referente a cada variável, através do ANOVA, foi criada uma tabela com todas as médias, e foi possível fazer as relações entre as mesmas. Ver a seguir (Tabela 119).

Tabela 119. Médias

	death	pipwat	obes	airpol	idh	gov	epi	hpi	ohi	ginni
OTHER	70,63428	47,26704	66,87507	74,76013	47,46332	33,36126	37,94995	41,79449	51,06979	62,43681
AIBER	92,89159	85,21942	46,68265	82,08955	65,61544	42,48362	52,69869	72,35625	55,45456	43,084054
AVECO	98,89842	94,89466	53,73479	77,50138	85,11319	83,2524	81,39376	50,58306	24,23534	79,912433
AZUL	39,96325	10,52248	88,48921	85,21677	19,09469	26,33463	21,19543	25,45519	58,21297	57,474153
O	48,73526	39,19803	69,83213	81,59204	39,82427	30,36538	29,84691	35,40934	53,5688	41,253653
VERDE	90,40683	69,19961	55,91188	71,95168	60,26303	34,95099	45,49002	60,10612	51,65927	60,978822
LARANJA	98,64865	92,71556	49,84121	76,69423	84,02778	77,61353	78,79409	48,21623	27,96922	77,934605

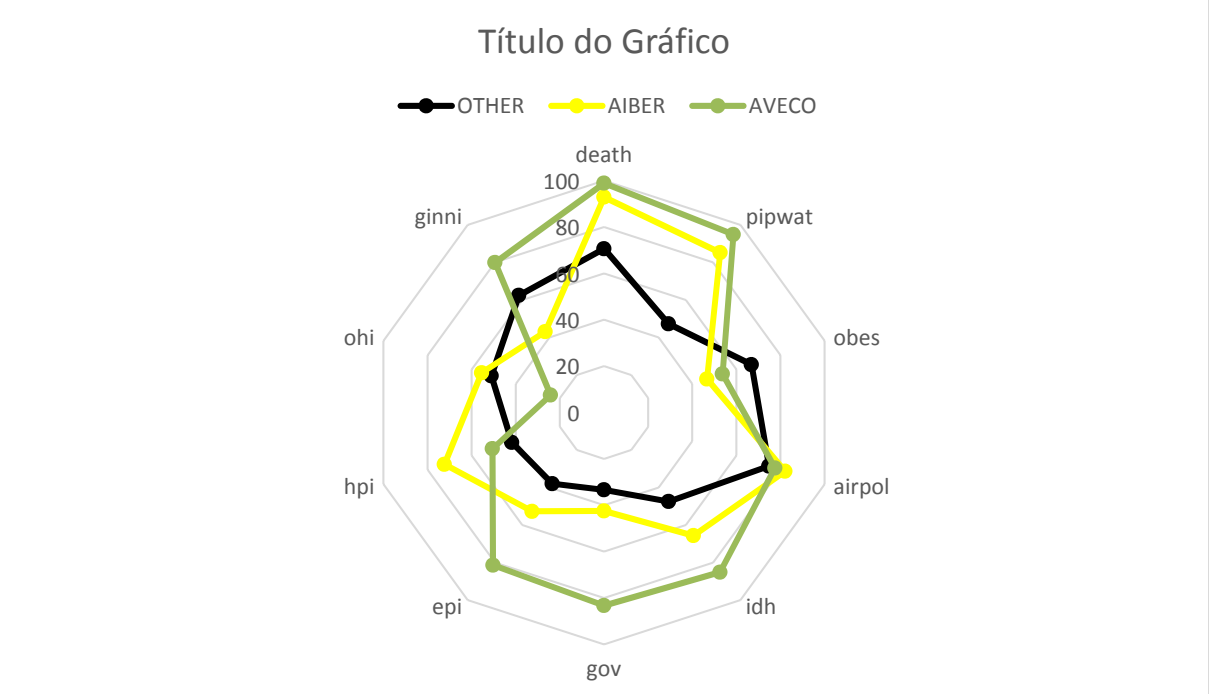
Fonte: Elaborado pela autora (Excel)

Gráfico 130



Fonte: Elaborado pela autora (Excel)

Gráfico 131



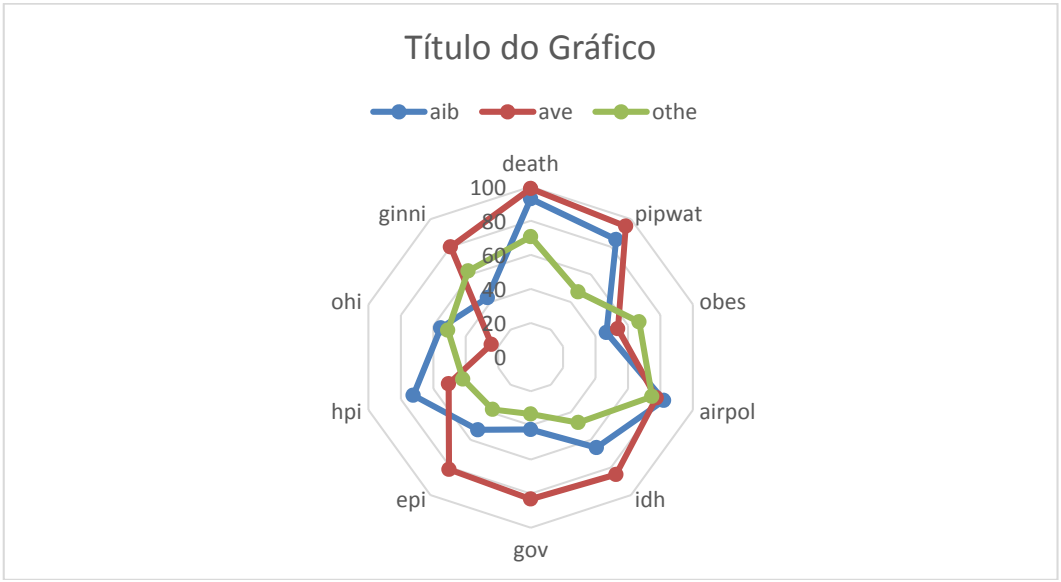
Fonte: Elaborado pela autora (Excel)

Tabela 120.

	death	pipwat	obes	airpol	idh	gov	epi	hpi	ohi	ginni
aib	92,89	85,22	46,68	82,09	65,62	42,48	52,7	72,36	55,45	43,08
ave	98,9	94,89	53,73	77,5	85,11	83,25	81,39	50,58	24,24	79,91
othe	70,63	47,27	66,88	74,76	47,46	33,36	37,95	41,79	51,07	62,44
g1	39,96	10,52	88,49	85,22	19,09	26,33	21,2	25,46	58,21	57,47
g2	48,74	39,2	69,83	81,59	39,82	30,37	29,85	35,41	53,57	41,25
g3	90,41	69,2	55,91	71,95	60,26	34,95	45,49	60,11	51,66	60,98
g4	98,65	92,72	49,84	76,69	84,03	77,61	78,79	48,22	27,97	77,93

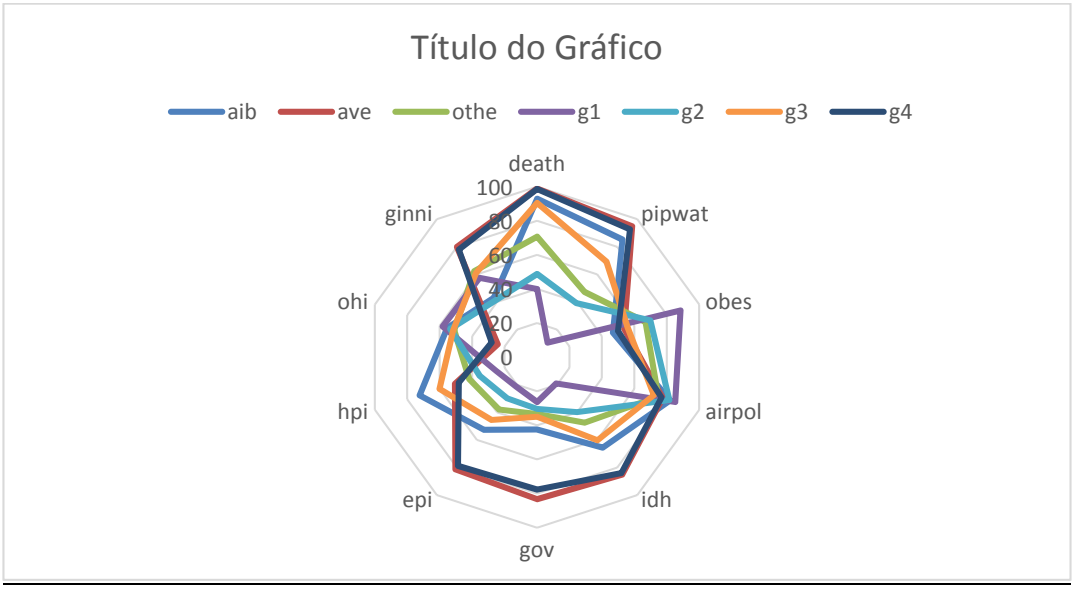
Fonte: Elaborado pela autora (Excel)

Gráfico 132



Fonte: Elaborado pela autora (Excel)

Gráfico 133



Fonte: Elaborado pela autora (Excel)

Através do *Minitab> Stat> Multivariate> Simple Correspondence Analysis*, utilizando-se 3ORD com as 3 regiões principais e as 10 variáveis Stepwiser, obteve-se:

Simple Correspondence Analysis: Mean_Deaths; Mean_Access; Mean_Obesity; Mean_Ou

Tabela 119. *Contingency Table*

	Deaths	Access	Obesity	Outdoor	IDH	Governança	EPI
OTHER	70,634	47,267	66,875	74,760	47,463	33,361	37,950
AIBER	92,892	85,219	46,683	82,090	65,615	42,484	52,699
AVECO	98,898	94,895	53,735	77,501	85,113	83,252	81,394
Total	262,424	227,381	167,293	234,351	198,192	159,097	172,042

	Happy	OHI	GINNI	Total
OTHER	41,794	51,070	62,437	533,612
AIBER	72,356	55,455	43,084	638,576
AVECO	50,583	24,235	79,912	729,519
Total	164,734	130,760	185,433	1901,707

Fonte: Elaborado pela autora

Tabela 120. *Chi-Square Distances*

	Deaths	Access	Obesity	Outdoor	IDH	Governança	EPI	Happy
OTHER	0,122	4,285	8,465	1,232	1,194	2,851	2,208	0,424
AIBER	0,258	1,030	1,604	0,147	0,013	2,240	0,445	5,249
AVECO	0,031	0,674	1,699	1,710	1,085	8,090	3,592	2,517
Total	0,412	5,989	11,767	3,089	2,293	13,181	6,245	8,190

	OHI	GINNI	Total
OTHER	5,635	2,081	28,498
AIBER	3,036	5,910	19,933
AVECO	13,400	1,083	33,881
Total	22,071	9,074	82,311

Fonte: Elaborado pela autora

Tabela 121. *Relative Inertias*

	Deaths	Access	Obesity	Outdoor	IDH	Governança	EPI	Happy	OHI
OTHER	0,001	0,052	0,103	0,015	0,015	0,035	0,027	0,005	0,068
AIBER	0,003	0,013	0,019	0,002	0,000	0,027	0,005	0,064	0,037
AVECO	0,000	0,008	0,021	0,021	0,013	0,098	0,044	0,031	0,163
Total	0,005	0,073	0,143	0,038	0,028	0,160	0,076	0,100	0,268

	GINNI	Total
OTHER	0,025	0,346
AIBER	0,072	0,242
AVECO	0,013	0,412
Total	0,110	1,000

Fonte: Elaborado pela autora

Tabela 122. *Analysis of Contingency Table*

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0297	0,6854	0,6854	******************************
2	0,0136	0,3146	1,0000	*************
Total	0,0433			

Fonte: Elaborado pela autora

Tabela 123. *Row Contributions*

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	OTHER	1,000	0,281	0,346	-0,185	0,640	0,323	0,139	0,360	0,396
2	AIBER	1,000	0,336	0,242	-0,089	0,253	0,090	-0,153	0,747	0,575
3	AVECO	1,000	0,384	0,412	0,213	0,978	0,587	0,032	0,022	0,029

Fonte: Elaborado pela autora

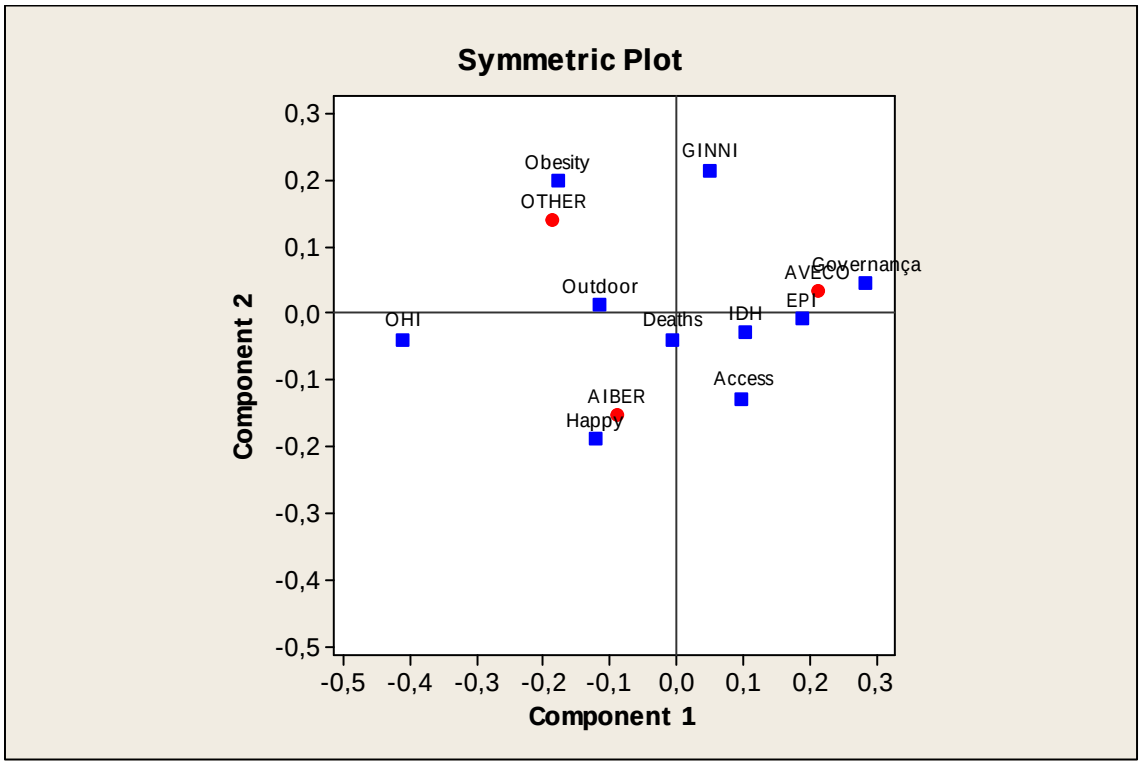
Tabela 124. *Column Contributions*

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr

1	Deaths	1,000	0,138	0,005	-0,005	0,019	0,000	-0,039	0,981	0,016
2	Access	1,000	0,120	0,073	0,100	0,377	0,040	-0,128	0,623	0,144
3	Obesity	1,000	0,088	0,143	-0,176	0,439	0,092	0,199	0,561	0,255
4	Outdoor	1,000	0,123	0,038	-0,114	0,989	0,054	0,012	0,011	0,001
5	IDH	1,000	0,104	0,028	0,103	0,922	0,037	-0,030	0,078	0,007
6	Governança	1,000	0,084	0,160	0,284	0,976	0,228	0,044	0,024	0,012
7	EPI	1,000	0,090	0,076	0,190	0,998	0,111	-0,008	0,002	0,000
8	Happy	1,000	0,087	0,100	-0,119	0,286	0,042	-0,188	0,714	0,226
9	OHI	1,000	0,069	0,268	-0,409	0,991	0,388	-0,040	0,009	0,008
10	GINNI	1,000	0,098	0,110	0,052	0,055	0,009	0,215	0,945	0,331

Fonte: Elaborado pela autora

Gráfico 134.



Fonte: Elaborado pela autora

Análise:

O grupo AVECO está próximo de Governança e EPI e distante de OHI. O grupo AIBER está próximo do *Happy Planet* e distante de GINNI. E o grupo OTHERS está próximo de *Obesity* e longe de *Access to Piped Water*.

3.18. ANÁLISE PAÍSES IBERO AMERICANOS

Por tudo que foi apresentado até agora, percebe-se que os países Ibero Americanos se encontram quase sempre numa posição intermediária. Às vezes se apresenta em uma melhor posição devido a países como Portugal, Espanha, Uruguai e Chile. Os países desenvolvidos, AVECO, estão na dianteira com os melhores índices até agora, e os países africanos estão com um pior posicionamento

Primeiramente, recomenda-se observar os clusters novamente, dessa vez só para os 21 países do AIBER investigados:

Quadro 8. SAÚDE E BEM-ESTAR PAÍSES IBERO-AMERICANOS

PAÍSES IBERO-AMERICANOS	VARIÁVEIS DE MAIOR RELEVÂNCIA
1. Argentina	***1. Deaths from Infectious Deseases***
2. Bolivia	***2. Access to piped water***
3. Brazil	***3. Obesity rate***
4. Chile	***4. Outdoor air pollutioni***
5. Colombia	***5. IDH***
6. Costa Rica	***6. Governance***
7. Cuba	***7. EPI Score***
8. Dominican Republic	***8. Happy Planet***
9. Ecuador	***9. Ocean Health Index***
10. El Salvador	**10. *Ginni***
11. Guatemala	
12. Honduras	
13. Mexico	
14. Nicaragua	
15. Panama	
16. Paraguay	
17. Peru	
18. Portugal	
19. Spain	
20. Uruguay	
21. Venezuela	

Fonte: Elaborado pela Autora

Voltando-se mais uma vez para o foco das 10 variáveis mais correlacionadas, e que norteiam esse trabalho desde as Regressões mais “finas”, conforme mostra novamente o Quadro 8., segue-se os Clusters e Dendogramas só referentes a tais países.

Cluster Analysis of Observations: N_P_Deaths f; N_P_Access t; N_P Obesity ; ...

Tabela 124. Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	83,1279	55,427	8	17	8	2
2	19	82,2079	58,449	6	15	6	2
3	18	81,1046	62,074	3	13	3	2
4	17	79,2737	68,089	2	16	2	2
5	16	78,4072	70,935	5	12	5	2
6	15	77,7907	72,960	18	19	18	2
7	14	75,8352	79,384	3	9	3	3
8	13	72,3956	90,684	10	14	10	2
9	12	69,3570	100,666	18	20	18	3
10	11	69,1172	101,454	1	7	1	2
11	10	67,7011	106,106	2	8	2	4
12	9	63,8169	118,866	5	11	5	3
13	8	61,4607	126,606	3	6	3	5
14	7	60,2542	130,570	10	21	10	3
15	6	56,0740	144,302	4	18	4	4
16	5	54,8872	148,201	2	5	2	7
17	4	51,0488	160,810	1	3	1	7
18	3	43,0570	187,065	1	10	1	10
19	2	41,7147	191,474	1	2	1	17
20	1	0,0000	328,512	1	4	1	21

Fonte: Elaborado pela Autora

Tabela 125. Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	7	5663,29	27,6653	38,1961
Cluster2	7	6230,60	29,3335	36,3854
Cluster3	4	3669,35	29,6870	38,5162
Cluster4	3	2026,26	25,7468	30,6750

Fonte: Elaborado pela Autora

Tabela 126. Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
N_P_Deaths from	94,8209	88,1718	97,5080	93,2476
N_P_Access to piped water	89,7880	78,3177	98,2302	73,3157
N_P Obesity rate (% of pop.)_1	42,8229	56,3549	43,6451	37,1703
N_P Outdoor air pollution	80,5970	87,8465	69,7761	88,5572
N_P IDH - 2013_1	70,6755	55,9426	80,5189	56,5074
N_P GovernanÁa_1	41,8395	32,2939	73,3156	26,6532
N_P EPI Score_1	53,1320	42,6488	74,1659	46,5145
N_P Happy Planet Index_1	80,0531	71,3076	51,6232	84,4879
N_P OHI_1	50,5643	56,4149	40,3950	84,7039
N_P GINNI Index_1	41,5919	34,6854	55,4410	49,6868

Fonte: Elaborado pela Autora

Tabela 127. Grande Centroide

Variable	Grand centroid
N_P_Deaths from	92,8916
N_P_Access to piped water	85,2194
N_P Obesity rate (% of pop.)_1	46,6827
N_P Outdoor air pollution	82,0896
N_P IDH - 2013_1	65,6154
N_P GovernanÁa_1	42,4836
N_P EPI Score_1	52,6987
N_P Happy Planet Index_1	72,3562
N_P OHI_1	55,4546
N_P GINNI Index_1	43,0841

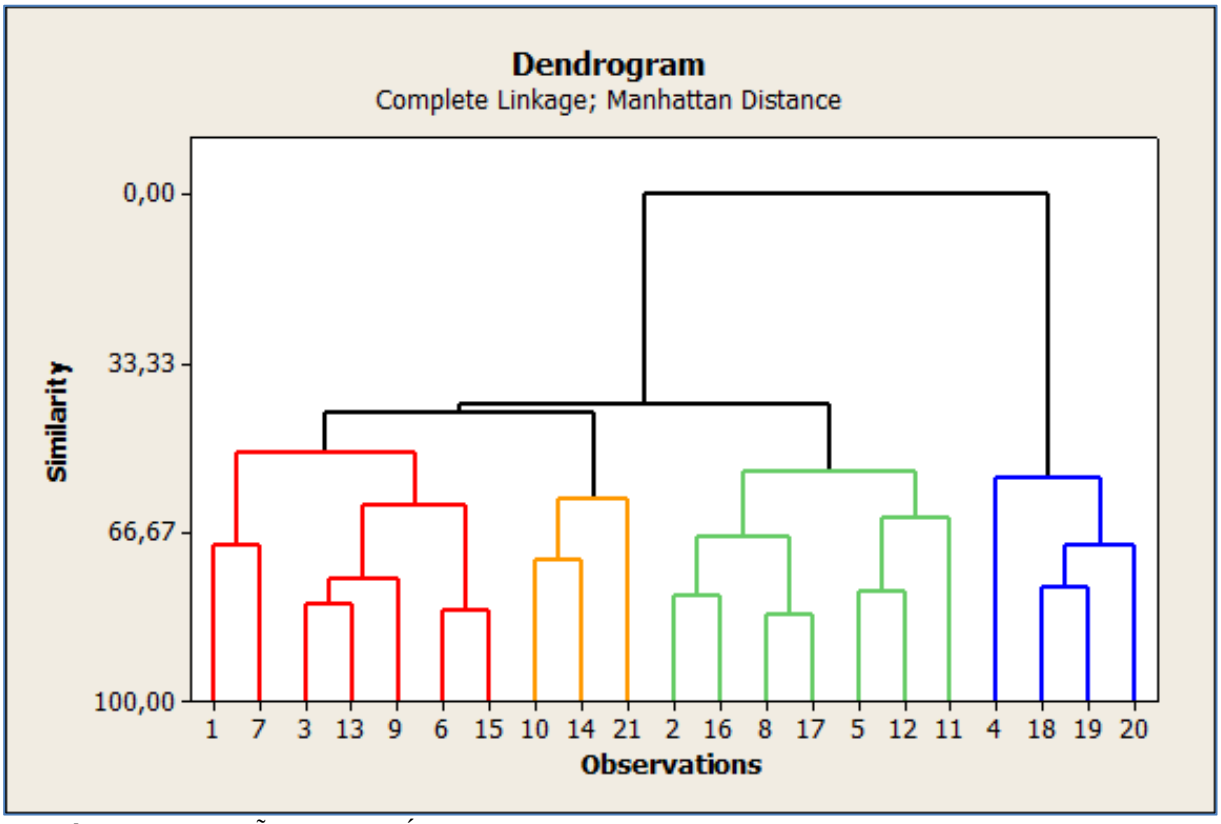
Fonte: Elaborado pela Autora

Tabela 128. Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,0000	31,4315	53,1958	45,7710
Cluster2	31,4315	0,0000	72,9847	40,8099
Cluster3	53,1958	72,9847	0,0000	87,3415
Cluster4	45,7710	40,8099	87,3415	0,0000

Fonte: Elaborado pela Autora

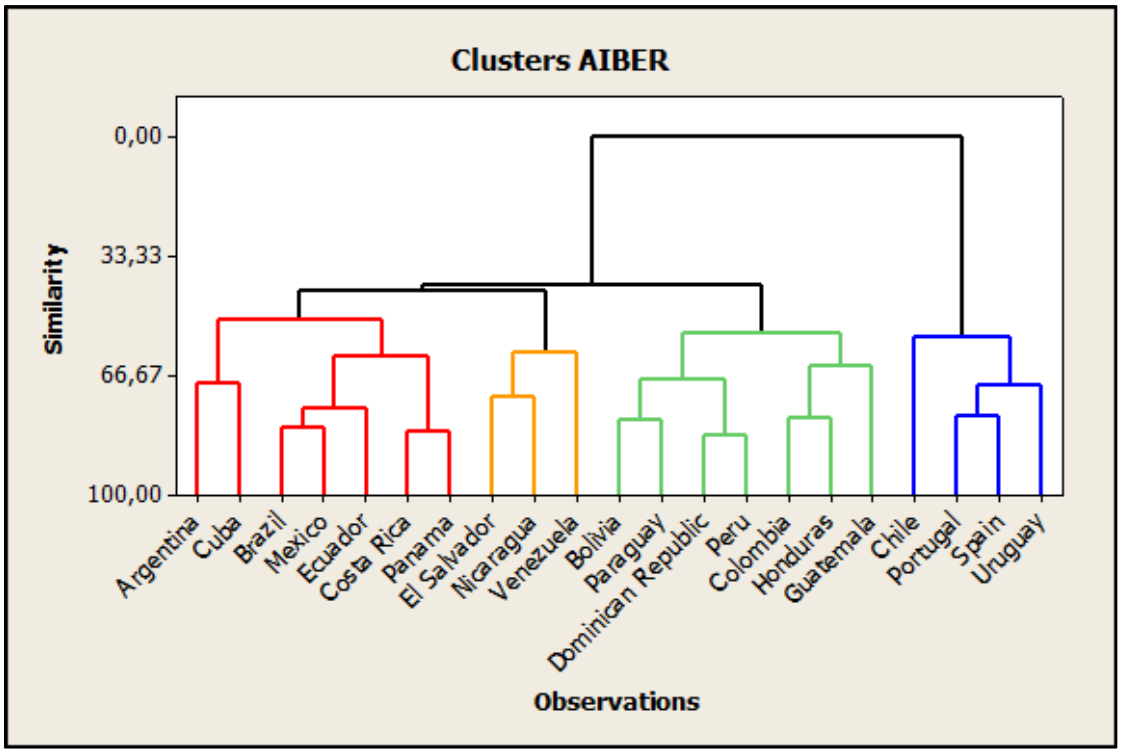
Gráfico 135. Dendrograma Manhattan



Legenda: EMBAIXO SÃO OS 21 PAÍSES: De 1 a 20.
Fonte: Elaborado pela Autora

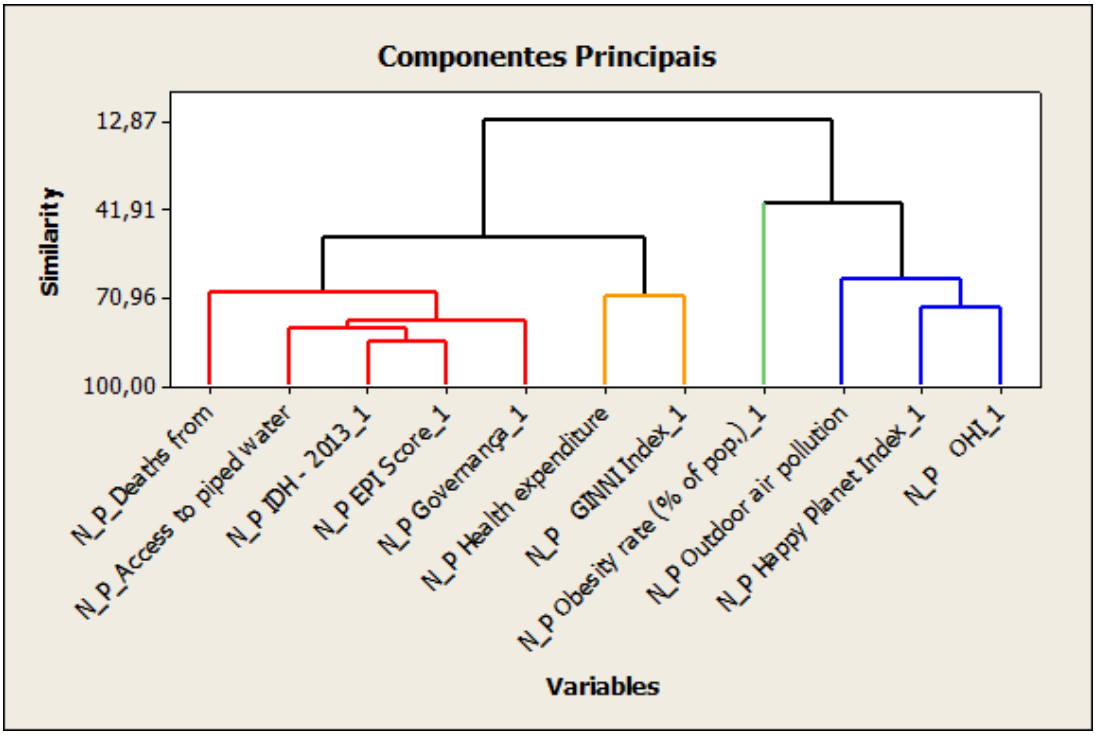
Agora, buscando analisar através do Comando STAT> MULTIVARIATE> CLUSTER VARIABLES SÓ COM OS PAÍSES AIBER, tem-se:

Gráfico 136.



Fonte: Elaborado pela Autora

Gráfico 137. Dendograma Comp. Principais



Fonte: Elaborado pela Autora

Tabela 138. Cluster Analysis of Observations: N_P_Deaths f; N_P_Access t; N_P Obesity ; ...

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	83,1279	55,427	8	17	8	2
2	19	82,2079	58,449	6	15	6	2
3	18	81,1046	62,074	3	13	3	2
4	17	79,2737	68,089	2	16	2	2
5	16	78,4072	70,935	5	12	5	2
6	15	77,7907	72,960	18	19	18	2
7	14	75,8352	79,384	3	9	3	3
8	13	72,3956	90,684	10	14	10	2
9	12	69,3570	100,666	18	20	18	3
10	11	69,1172	101,454	1	7	1	2
11	10	67,7011	106,106	2	8	2	4
12	9	63,8169	118,866	5	11	5	3
13	8	61,4607	126,606	3	6	3	5
14	7	60,2542	130,570	10	21	10	3
15	6	56,0740	144,302	4	18	4	4
16	5	54,8872	148,201	2	5	2	7
17	4	51,0488	160,810	1	3	1	7
18	3	43,0570	187,065	1	10	1	10
19	2	41,7147	191,474	1	2	1	17
20	1	0,0000	328,512	1	4	1	21

Tabela 139. Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	7	5663,29	27,6653	38,1961
Cluster2	7	6230,60	29,3335	36,3854
Cluster3	4	3669,35	29,6870	38,5162
Cluster4	3	2026,26	25,7468	30,6750

Tabela 140. Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
N_P_Deaths from	94,8209	88,1718	97,5080	93,2476
N_P_Access to piped water	89,7880	78,3177	98,2302	73,3157
N_P Obesity rate (% of pop.)_1	42,8229	56,3549	43,6451	37,1703
N_P Outdoor air pollution	80,5970	87,8465	69,7761	88,5572
N_P IDH - 2013_1	70,6755	55,9426	80,5189	56,5074
N_P GovernanÁa_1	41,8395	32,2939	73,3156	26,6532
N_P EPI Score_1	53,1320	42,6488	74,1659	46,5145
N_P Happy Planet Index_1	80,0531	71,3076	51,6232	84,4879
N_P OHI_1	50,5643	56,4149	40,3950	84,7039
N_P GINNI Index_1	41,5919	34,6854	55,4410	49,6868

Fonte: Elaborado pela Autora

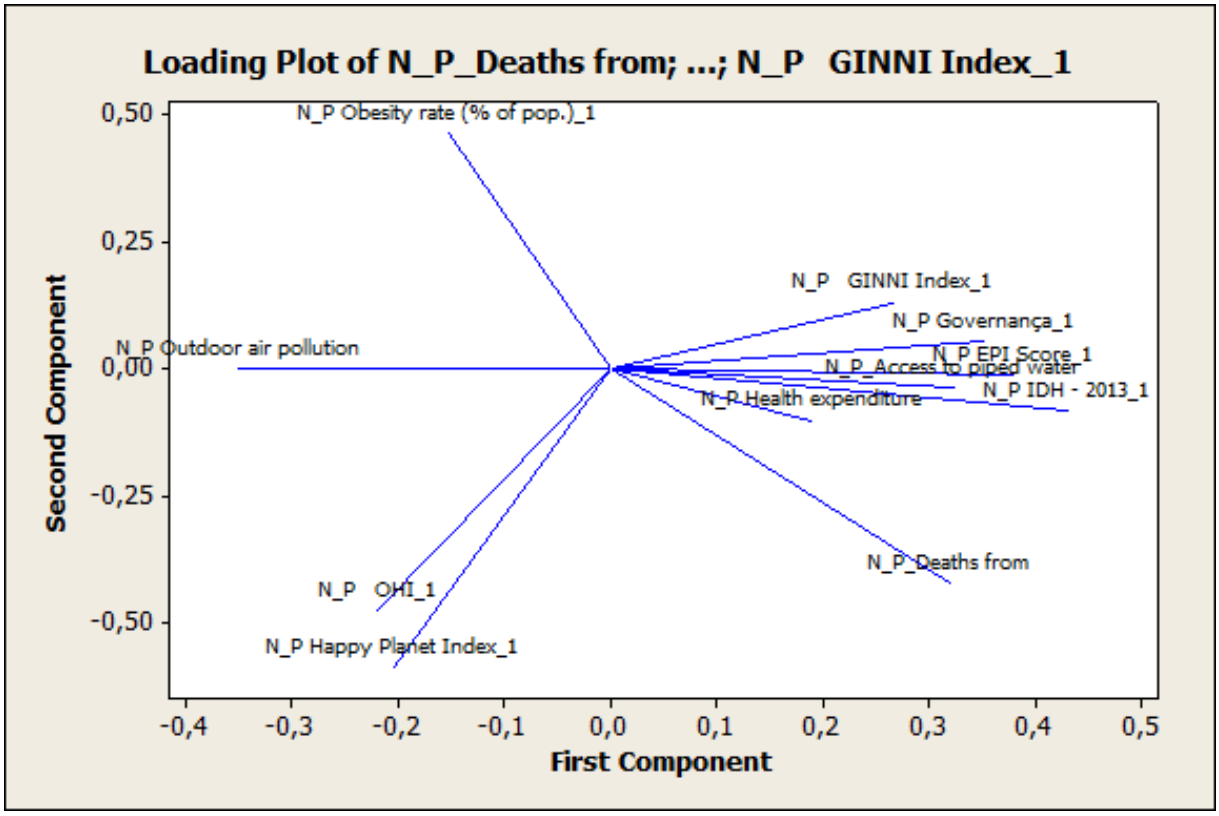
Tabela 141. Grand Centroid

Variable	Grande centroid
N_P_Deaths from	92,8916
N_P_Access to piped water	85,2194
N_P Obesity rate (% of pop.)_1	46,6827
N_P Outdoor air pollution	82,0896
N_P IDH - 2013_1	65,6154
N_P GovernanÁa_1	42,4836
N_P EPI Score_1	52,6987
N_P Happy Planet Index_1	72,3562
N_P OHI_1	55,4546
N_P GINNI Index_1	43,0841

Tabela 142. Distances Between Cluster Centroids

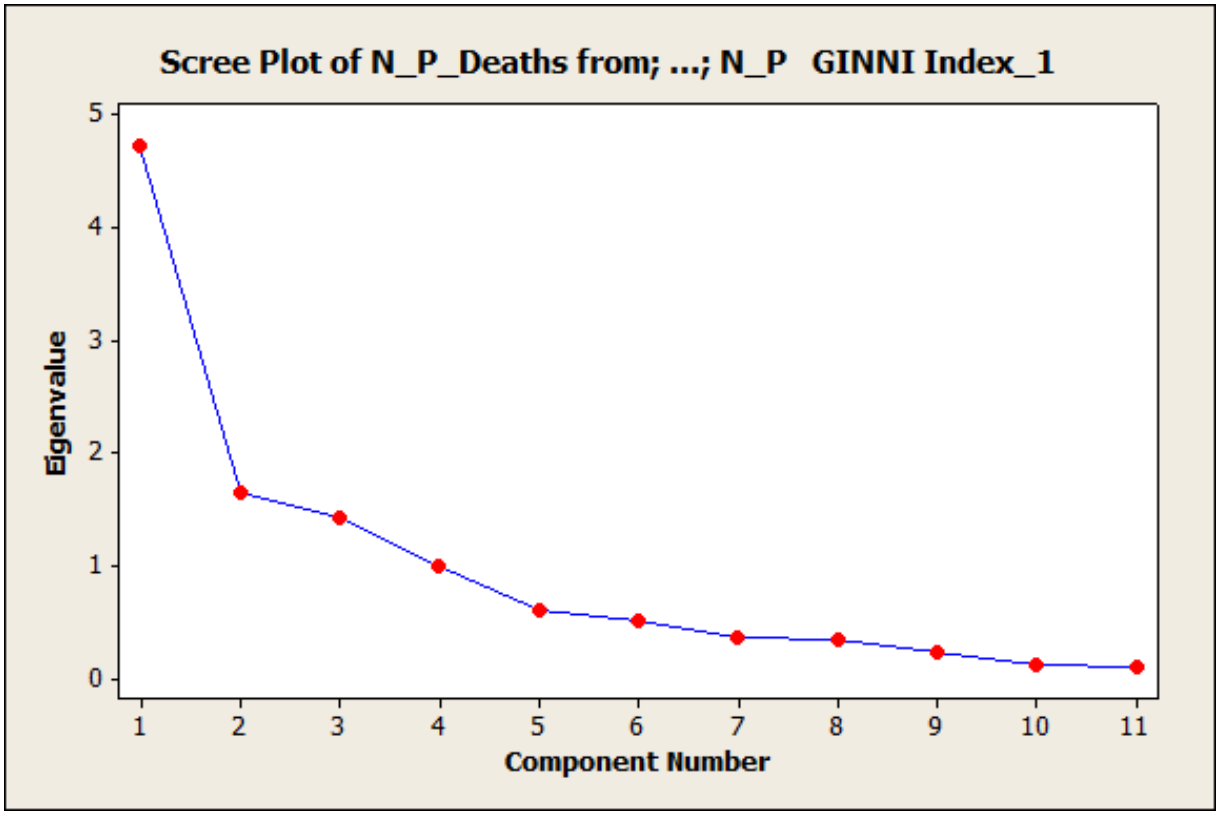
	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,0000	31,4315	53,1958	45,7710
Cluster2	31,4315	0,0000	72,9847	40,8099
Cluster3	53,1958	72,9847	0,0000	87,3415
Cluster4	45,7710	40,8099	87,3415	0,0000

Gráfico 138.



Fonte: Elaborado pela Autora

Gráfico 139.



Fonte: Elaborado pela Autora

Tabela 143. One-way ANOVA: N_P IDH - 2013_1 versus 4 Cluster AIBER

Source	DF	SS	MS	F	P
4 Cluster AIBER	3	1971,5	657,2	12,08	0,000
Error	17	925,0	54,4		
Total	20	2896,5			

```
S = 7,377   R-Sq = 68,06%   R-Sq(adj) = 62,43%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N    Mean   StDev   +---------+---------+---------+---------
1      7  70,675   5,934                (----*----)
2      7  55,943   7,177     (----*----)
3      4  80,519   5,361                        (-----*------)
4      3  56,507  12,620   (------*-------)
                           +---------+---------+---------+---------
                          48        60        72        84
Pooled StDev = 7,377
```

Fonte: Elaborado pela Autora

Tabela 144. Cluster Analysis of Variables: N_P_Deaths f; N_P_Access t; N_P Obesity ; ...

Correlation Coefficient Distance, Complete Linkage							
Amalgamation Steps							
							Number of obs.
Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	in new cluster
1	10	85,6615	0,28677	5	7	5	2
2	9	81,1083	0,37783	2	5	2	3
3	8	78,8196	0,42361	2	6	2	4
4	7	74,4905	0,51019	9	10	9	2
5	6	70,4225	0,59155	8	11	8	2
6	5	69,6450	0,60710	1	2	1	5
7	4	64,9480	0,70104	4	9	4	3
8	3	51,4055	0,97189	1	8	1	7
9	2	40,0803	1,19839	3	4	3	4
10	1	12,8678	1,74264	1	3	1	11

Final Partition

Cluster 1
  N_P_Deaths from N_P_Access to piped water N_P IDH - 2013_1 N_P
     GovernanÁa_1 N_P EPI Score_1
Cluster 2
  N_P Obesity rate (% of pop.)_1
Cluster 3
  N_P Outdoor air pollution N_P Happy Planet Index_1 N_P   OHI_1
Cluster 4
  N_P Health expenditure N_P   GINNI Index_1

NOS CÁLCULOS SEGUINTES, FOI INCLUÍDA A VARIÁVEL *HEALTH EXPENDITURE*, PARA ANALISAR O COMPORTAMENTO DOS DADOS.

Tabela 145. Principal Component Analysis: N_P_Deaths f; N_P_Access t; N_P Obesity ; N_P Out

Eigenanalysis of the Correlation Matrix

Eigenvalue	4,7108	1,6408	1,4263	0,9887	0,5987	0,4999	0,3529	0,3318
Proportion	0,428	0,149	0,130	0,090	0,054	0,045	0,032	0,030
Cumulative	0,428	0,577	0,707	0,797	0,851	0,897	0,929	0,959
Eigenvalue	0,2337	0,1227	0,0937					
Proportion	0,021	0,011	0,009					
Cumulative	0,980	0,991	1,000					

Variable	PC1	PC2	PC3	PC4
N_P_Deaths from	0,320	-0,422	-0,074	-0,037
N_P_Access to piped water	0,323	-0,038	0,389	0,313
N_P Obesity rate (% of pop.)_1	-0,153	0,461	-0,298	0,515
N_P Outdoor air pollution	-0,349	-0,001	0,216	0,171
N_P IDH - 2013_1	0,430	-0,082	-0,028	-0,015
N_P GovernanÁa_1	0,351	0,054	0,134	0,335
N_P Health expenditure	0,189	-0,102	-0,547	0,499
N_P EPI Score_1	0,379	-0,014	0,169	0,039
N_P Happy Planet Index_1	-0,204	-0,587	0,078	0,226
N_P OHI_1	-0,219	-0,476	-0,349	0,041
N_P GINNI Index_1	0,266	0,129	-0,482	-0,436

Fonte: Elaborado pela Autora

OBSERVOU-SE QUE *HEALTH EXPENDITURE* NÃO É TÃO IMPORTANTE, SEGUIU-SE, POR TANTO, COM AS 10 PREVISTAS.

Tabela 146. One-way ANOVA: Pc1_AIBER versus 4 Cluster AIBER

Source	DF	SS	MS	F	P
4 Cluster AIBER	3	80,947	26,982	34,57	0,000
Error	17	13,269	0,781		
Total	20	94,216			

S = 0,8835 R-Sq = 85,92% R-Sq(adj) = 83,43%

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N     Mean   StDev  ----+---------+---------+---------+-----
1      7   0,5306  0,8534              (---*--)
2      7  -1,8140  0,7670   (---*--)
3      4   3,4648  1,0564                           (---*----)
4      3  -1,6253  1,0051  (-----*----)
                          ----+---------+---------+---------+-----
                           -2,0       0,0       2,0       4,0
Pooled StDev = 0,8835
```

Tabela 147. One-way ANOVA: N_P IDH - 2013_1 versus 4 Cluster AIBER

Source	DF	SS	MS	F	P
4 Cluster AIBER	3	1971,5	657,2	12,08	0,000
Error	17	925,0	54,4		
Total	20	2896,5			

S = 7,377 R-Sq = 68,06% R-Sq(adj) = 62,43%

```
                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level  N    Mean   StDev  +---------+---------+---------+---------
1      7  70,675   5,934               (----*----)
2      7  55,943   7,177    (----*----)
3      4  80,519   5,361                       (-----*------)
4      3  56,507  12,620  (------*-------)
                         +---------+---------+---------+---------
                        48        60        72        84
Pooled StDev = 7,377
```

Tabela 148. One-way ANOVA: PC1 versus 4 Cluster AIBER

Source	DF	SS	MS	F	P
4 Cluster AIBER	3	27,639	9,213	21,46	0,000
Error	17	7,298	0,429		
Total	20	34,937			

S = 0,6552 R-Sq = 79,11% R-Sq(adj) = 75,42%

```
                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level  N    Mean   StDev  --+---------+---------+---------+-------
1      7  2,0658  0,5853               (---*----)
2      7  0,2883  0,7001  (---*----)
3      4  3,2644  0,5691                       (-----*-----)
4      3  0,5559  0,8156  (------*-----)
                         --+---------+---------+---------+-------
                         0,0       1,2       2,4       3,6
Pooled StDev = 0,6552
```

Fonte: Elaborado pela Autora

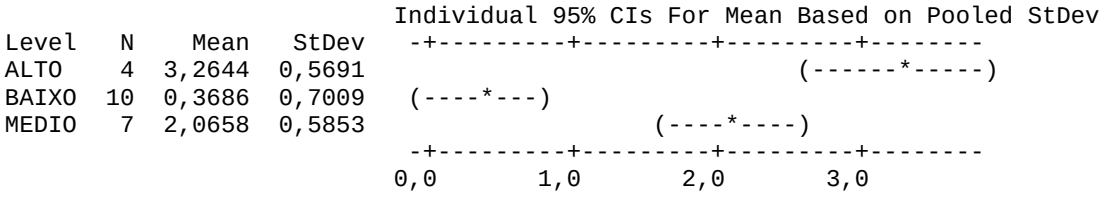
Através da Codificação, ordenar a planilha de países no MINITAB.
De 2 para 1 ALTO
De 4 para 2 MEDIO
De 1 para 3 BAIXO
De 3 para 4 MUITO BAIXO

VOLTANDO PARA 3 CLUSTERS

Tabela 149. One-way ANOVA: PC1 versus C47

Source	DF	SS	MS	F	P
C47	2	27,489	13,744	33,21	0,000
Error	18	7,449	0,414		
Total	20	34,937			

S = 0,6433 R-Sq = 78,68% R-Sq(adj) = 76,31%

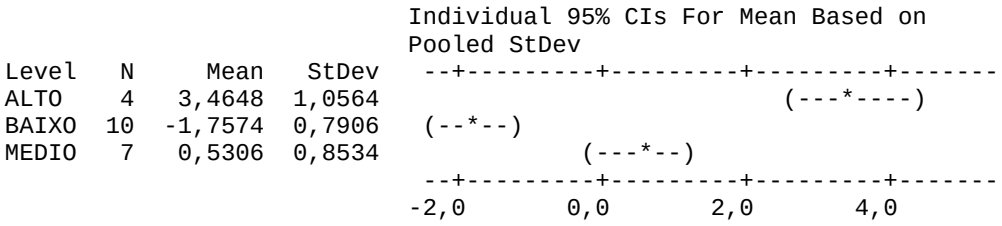


Pooled StDev = 0,6433

Tabela 150. One-way ANOVA: Pc1_AIBER versus C47

Source	DF	SS	MS	F	P
C47	2	80,872	40,436	54,55	0,000
Error	18	13,344	0,741		
Total	20	94,216			

S = 0,8610 R-Sq = 85,84% R-Sq(adj) = 84,26%



Pooled StDev = 0,8610

Fonte: Elaborado pela Autora

Análise
Em virtude de todos os procedimentos e métodos estatísticos utilizados, gerou-se a seguinte situação dos dados: tudo indica que realmente o PC1 é o mais adequado componente. Ainda assim será feito um novo Stepwise para identificar o peso das variáveis, e com isso, se rodarão os outros PCs.

Tabela 152. PC1/ IBERO-AMERICANO
E LATINO AMERICANO

País_1_2 Clusters	PC1_AIBER
Argentina	1,67553
Bolivia	-1,84819
Brazil	-0,28918
Chile	2,66526
Colombia	-0,93789
Costa Rica	0,72040
Cuba	1,43389
Dominican Republic	-1,00245
Ecuador	-0,56408
El Salvador	-1,63223
Guatemala	-2,64956
Honduras	-2,79936
Mexico	0,74627
Nicaragua	-2,62687
Panama	-0,00838
Paraguay	-2,18820
Peru	-1,27202
Portugal	4,33744
Spain	4,41429
Uruguay	2,44209
Venezuela	-0,61676

Fonte: Elaborado pela Autora

Análise:
Reordenando tais países, rankeando-os pelo índice de melhor situação de saúde e bem-estar social, do mais alto bem-estar social e saúde, para o mais baixo, após a positivação dos dados normalizados, encontrou-se (Ver Quadro 9).

Como já informado, recomenda-se ainda proceder com uma analise *stepwise* dos componentes para selecionar as variáveis-chave. Ver tabelas 153 e 154.

Tabela 153. Stepwise Regression: Pc1_Aiber Versus N_P_Deaths F; N_P_Access T; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1_AIBER on 10 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-11,052	-10,347	-10,428	-5,578	-6,439	-7,594
N_P IDH - 2013_1	0,168	0,135	0,116	0,086	0,080	0,059
T-Value	11,39	9,66	8,05	5,43	5,64	5,10
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
N_P GovernanÁa_1		0,0355	0,0267	0,0272	0,0292	0,0263
T-Value		3,99	3,09	3,76	4,56	5,69
P-Value		0,001	0,007	0,002	0,000	0,000
N_P EPI Score_1			0,0318	0,0353	0,0330	0,0259
T-Value			2,47	3,26	3,44	3,67
P-Value			0,025	0,005	0,004	0,003
N_P Outdoor air pollution				-0,0377	-0,0310	-0,0341
T-Value				-2,89	-2,63	-4,04
P-Value				0,011	0,019	0,001
N_P GINNI Index_1					0,0175	0,0284
T-Value					2,39	4,81
P-Value					0,030	0,000
N_P_Access to piped water						0,0334
T-Value						3,95
P-Value						0,001
S	0,796	0,596	0,526	0,440	0,386	0,275
R-Sq	87,22	93,21	95,00	96,72	97,62	98,87
R-Sq(adj)	86,55	92,46	94,12	95,90	96,83	98,39
Mallows Cp	388,7	200,4	145,7	93,2	66,5	28,7

Step	7	8	9
Constant	-10,08	-11,33	-11,38
N_P IDH - 2013_1	0,0501	0,0403	0,0435
T-Value	4,29	4,25	4,88
P-Value	0,001	0,001	0,000
N_P GovernanÁa_1	0,0253	0,0229	0,0216
T-Value	5,83	6,68	6,67
P-Value	0,000	0,000	0,000
N_P EPI Score_1	0,0243	0,0229	0,0225
T-Value	3,67	4,49	4,78
P-Value	0,003	0,001	0,001
N_P Outdoor air pollution	-0,0333	-0,0319	-0,0303
T-Value	-4,24	-5,26	-5,36
P-Value	0,001	0,000	0,000
N_P GINNI Index_1	0,0300	0,0308	0,0271
T-Value	5,39	7,20	6,09
P-Value	0,000	0,000	0,000
N_P_Access to piped water	0,0365	0,0347	0,0337
T-Value	4,53	5,59	5,87
P-Value	0,001	0,000	0,000
N_P_Deaths from	0,030	0,058	0,062
T-Value	1,79	3,72	4,27
P-Value	0,097	0,003	0,001
N_P OHI_1		-0,0102	-0,0085
T-Value		-3,17	-2,70
P-Value		0,008	0,021
N_P Happy Planet Index_1			-0,0062
T-Value			-1,78
P-Value			0,103
S	0,256	0,197	0,181
R-Sq	99,10	99,51	99,62
R-Sq(adj)	98,61	99,18	99,30
Mallows Cp	23,7	12,6	11,1

Análise:
As principais variáveis que compõem Saúde e Bem-Estar social para os 21 países analisados para o component PC1 são IDH, que responde por 87% de explicação, depois, de forma acumulada, junto com o IDH, a Governança, que explica 93,21% do fenômeno estudado (agregando 6,21%), em seguida, EPI, que cumulativamente responde por 95% (1,79% a mais), Outdoor air pollution (agregando-se 1,72%) por 96,72%; Ginni, 97,62% (0,90% a mais) e 98,87% Access to piped water (1,25%). O modelo 6 do Stepwise que inclui todas essas variáveis garante 98,87% do fenômeno estudado para os países pesquisados.

Tabela 154. Stepwise Regression: PC2_Aiber Versus N_P_Deaths F; N_P_Access T; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15
Response is PC2_AIBER on 10 predictors, with N = 21

Step	1	2	3	4	5
Constant	4,285	17,047	12,337	11,352	11,633
N_P Happy Planet Index_1	-0,0592	-0,0586	-0,0544	-0,0406	-0,0406
T-Value	-4,96	-8,02	-9,09	-17,74	-23,97
P-Value	0,000	0,000	0,000	0,000	0,000
N_P_Deaths from		-0,1378	-0,1073	-0,0970	-0,0890
T-Value		-5,71	-5,02	-13,69	-15,68
P-Value		0,000	0,000	0,000	0,000
N_P Obesity rate (% of pop.)_1			0,0335	0,0398	0,0382
T-Value			3,34	11,88	15,19
P-Value			0,004	0,000	0,000
N_P OHI_1				-0,0227	-0,0254
T-Value				-11,87	-16,00
P-Value				0,000	0,000
N_P_Access to piped water					-0,0094
T-Value					-3,76
P-Value					0,002
N_P IDH - 2013_1					
T-Value					
P-Value					
S	0,867	0,531	0,425	0,140	0,104
R-Sq	56,45	84,52	90,64	99,05	99,51
R-Sq(adj)	54,15	82,80	88,99	98,81	99,34
Mallows Cp	1640,3	574,1	343,1	25,3	9,7

Step	6
Constant	11,218
N_P Happy Planet Index_1	-0,0412
T-Value	-24,98
P-Value	0,000
N_P_Deaths from	-0,0806
T-Value	-10,70
P-Value	0,000
N_P Obesity rate (% of pop.)_1	0,0380
T-Value	15,79
P-Value	0,000
N_P OHI_1	-0,0264
T-Value	-16,06
P-Value	0,000
N_P_Access to piped water	-0,0080
T-Value	-3,10
P-Value	0,008
N_P IDH - 2013_1	-0,0058
T-Value	-1,59
P-Value	0,133

S	0,0987
R-Sq	99,58
R-Sq(adj)	99,41
Mallows Cp	8,8
Step	7
Constant	10,74
N_P Happy Planet Index_1	-0,0390
T-Value	-23,32
P-Value	0,000
N_P_Deaths from	-0,0790
T-Value	-12,14
P-Value	0,000
N_P Obesity rate (% of pop.)_1	0,0391
T-Value	18,46
P-Value	0,000
N_P OHI_1	-0,0272
T-Value	-18,76
P-Value	0,000
N_P_Access to piped water	-0,0048
T-Value	-1,85
P-Value	0,087
N_P IDH - 2013_1	-0,0107
T-Value	-2,88
P-Value	0,013
N_P GINNI Index_1	0,0051
T-Value	2,43
P-Value	0,030
S	0,0849
R-Sq	99,71
R-Sq(adj)	99,56
Mallows Cp	5,9

Análise:

As principais variáveis que compõem o tema Saúde e Bem-Estar social para os 21 países analisados de acordo com o componente PC2 são Happy Planet, que responde por 56,45% de explicação, depois, de forma acumulada, junto com o Deaths from Infectious Deseases dá um pulo significativo para 84,52% (agregando 28,07%), em seguida, Obesity, que cumulativamente responde por 90,64% (6,12% a mais), seguido por OHI que responde por 99,05% (8,41% a mais) de explicação para os países pesquisados.

Tabela 155. Principal Component Analysis: N_P_Deaths F; N_P_Access T; N_P Obesity ; N_P Out

Eigenanalysis of the Correlation Matrix

	PC1	PC2	PC3	PC4				
Eigenvalue	4,5750	1,6359	1,2263	0,7360	0,5627	0,3549	0,3390	0,2403
Proportion	0,457	0,164	0,123	0,074	0,056	0,035	0,034	0,024
Cumulative	0,457	0,621	0,744	0,817	0,874	0,909	0,943	0,967

Eigenvalue	0,2330	0,0969
Proportion	0,023	0,010
Cumulative	0,990	1,000

Variable	PC1	PC2	PC3	PC4
N_P_Deaths from	0,322	-0,418	0,179	0,310
N_P_Access to piped water	0,332	-0,075	-0,494	-0,037
N_P Obesity rate (% of pop.)_1	-0,168	0,501	0,009	0,712
N_P Outdoor air pollution	-0,350	-0,017	-0,309	-0,010
N_P IDH - 2013_1	0,433	-0,082	0,060	-0,022

N_P	GovernanÁa_1	0,355	0,042	-0,245	0,398
N_P	EPI Score_1	0,390	-0,035	-0,129	0,139
N_P	Happy Planet Index_1	-0,214	-0,584	-0,147	0,046
N_P	OHI_1	-0,237	-0,435	0,355	0,429
N_P	GINNI Index_1	0,259	0,169	0,632	-0,177

Análise

O Componente PC1 é realmente o principal componente para explicar o tema estudado para os países RIBER em questão. Observa-se que a proporção de 45,70% da Tabela 155. é bem superior às proporções dos outros PCs. Também pode-se dizer que quanto maior o valor indicado na Tabela 55 relativo às variáveis, maior o peso da variável no PC. Neste sentido, a coluna PC1 corrobora as variáveis IDH, EPI, Governança, Outdoor Air Pollution como mais relevantes para o componente. Os países do RIBER, com seus respectivos PC's já normalizados e positivados estão ilustrados no Quadro 9 e Mapa (Figura 1).

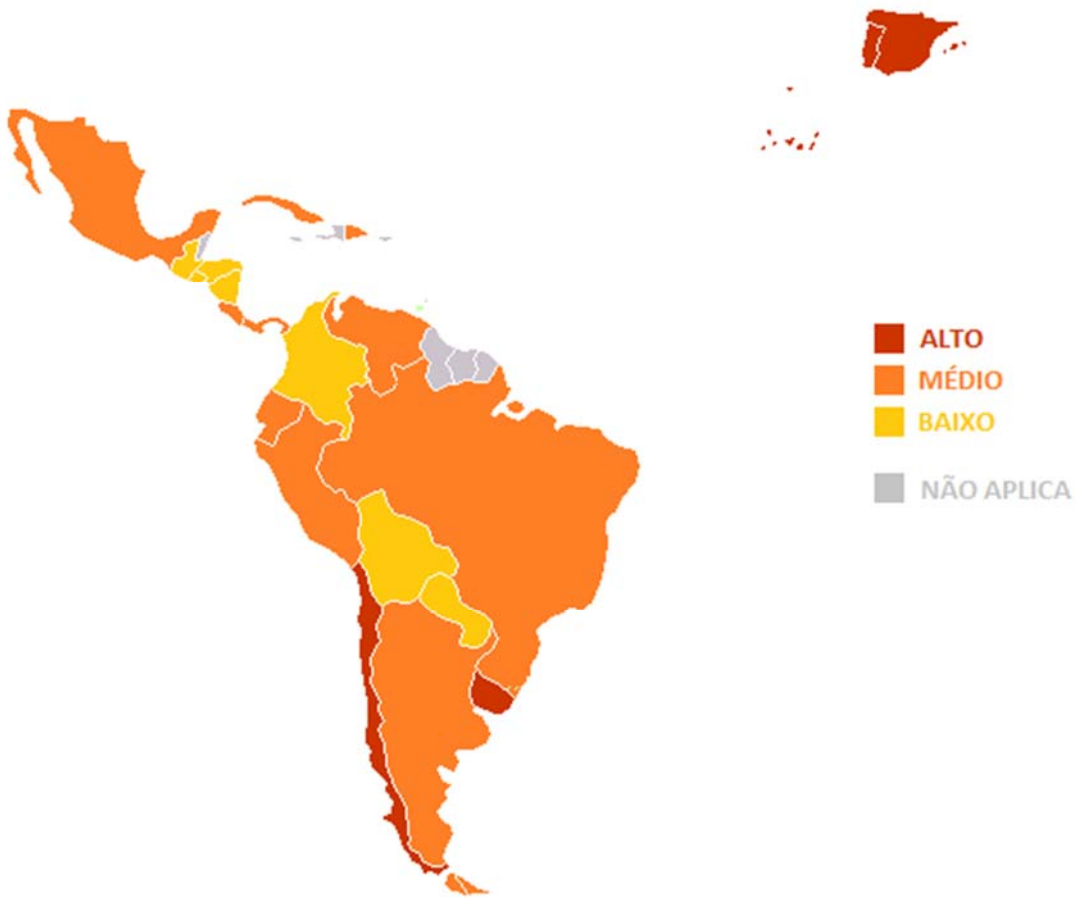
O Componente P1 será batizado de Indicador CRISMA (Cr_I_Saúde_Meio_Ambiente) – aceitando a sugestão do Prof. Dr. Hoyos.

QUADRO 9. PAÍSES IBERO E LATINO AMERICANOS SEGUNDO O RANKING DE SAÚDE E BEM-ESTAR SOCIAL

PAISES PC1_ORD AIBER	RANKING	PC1_AIBER POSITIVADOS
1. Spain	ALTO	100,000
2. Portugal	ALTO	98,935
3. Chile	ALTO	75,754
4. Uruguay	ALTO	72,660
5. Argentina	MEDIO	62,034
6. Cuba	MEDIO	58,684
7. Mexico	MEDIO	49,152
8. Costa Rica	MEDIO	48,793
9. Panama	MEDIO	38,690
10. Brazil	MEDIO	34,798
11. Ecuador	MEDIO	30,987
12. Venezuela	BAIXO	30,257
13. Colombia	BAIXO	25,805
14. Dominican Republic	BAIXO	24,910
15. Peru	BAIXO	21,173
16. El Salvador	BAIXO	16,179
17. Bolivia	BAIXO	13,186
18. Paraguay	BAIXO	8,472
19. Nicaragua	BAIXO	2,391
20. Guatemala	BAIXO	2,077
21. Honduras	BAIXO	0,000

Fonte: Elaborado pela autora e pelo Professor Dr. Hoyos

FIGURA 1. MAPA ILUSTRATIVO SAÚDE E BEM-ESTAR SOCIAL – PAÍSES DO RIBER



Fonte: Elaborado pela autora

4. CONSIDERAÇÕES FINAIS

O cuidado com o meio ambiente e a preservação ambiental da região Ibero-Americana tem um impacto não só em termos da saúde da sua população, como também do planeta. Como previsto pelo Projeto ORIBER, Guia de Gestão Pública Sustentável (GPS), pela análise dos dados referentes à Saúde e Meio Ambiente, o resumo das conclusões se encontra em sua maioria no subitem anterior.

As principais variáveis que compõem Saúde e Bem-Estar social para os 21 países analisados compondo o novo indicador CRI_S_MA Index são: IDH, Governança, EPI, Outdoor air pollution e Ginni, cobrindo 97,62% de explicação do tema para os países pesquisados.

Segundo o GPS – Guia de Países Sustentáveis, os países latinoamericanos e caribenhos sofrem de uma grande lacuna entre o conhecimento teórico acerca da saúde ambiental e a sua práctica em políticas públicas. O acesso e benefícios à população ficam prejudicados. Este fato contrasta com outro, o de representarem regiões ricas em biodiversidade, apesar de extremamente ameaçadas. A maior parte da população está concentrada nos centros urbanos, vive em condições precárias e seu modelo de desenvolvimento está historicamente ligado à extração predadora de recursos naturais e exploração da terra devido a grandes monoculturas

destinadas à exportação, inicialmente para as colônias. Este processso hoje em dia é chamado de neo-colonialismo extrativista.

Refletindo sobre tal fato e o resultado dos métodos estatísticos aplicados sobre os dados, percebe-se que a maioria dos países Ibero-Americanos está em um baixo nível de desenvolvimento e de proteção do meio ambiente e da saúde, corroborando com os dados históricos. Isto acontece, a depeito de que no mapa (figura 1) parece que a maioria tem um médio patamar de avanço quanto ao tema. Essa aparência se deve ao fato dos países Argentina e Brasil terem um grande território e se encontrarem nesta faixa intermediária de avanços.

Alguns pontos merecem destaque. Na América Latina, o GPS previa que Argentina, Chile e Uruguai tinham um desempenho um pouco melhor em termos de *Environmental Performance Index* (EPI) por exemplo, o que continua sendo verdade, apesar da Argentina ter descido um pouco nesta escala. O Brasil melhora um pouco nos últimos anos, de maneira que o IDH sobe um pouco o seu patamar, mas ainda muito baixo, estando apenas acima de Equador entre os de *score* Mediano (1º Argentina, 2º Cuba, 3º Mexico, 4º Costa Rica, 5º Panama, depois Brasil e Equador). Chile e Uruguai continuam sendo os mais avançados na América Latina, e presume-se que além dos avanços sociais, como mortalidade infantil e acesso a saneamento básico já apontados no passado, a combinação de IDH, Governança e EPI tenham também contribuído de alguma forma. Eles estão logo depois de Espanha e Portugal, que têm os mais altos *scores* e descendem de uma outra história e cultura, apresentando altos padrões de saúde e bem-estar social, assim como de preservação ambiental.

Como esperado, fatores ambientais podem ser melhorados em geral na América Latina. A infraestrutura da saúde pública disponível em cada país também deve ser foco de atenção para dirimir as predisposições às enfermedades. Por exemplo, parece existir uma relação do cancer (80% está relacionado ao meio ambiente), que apresenta grandes fatores de risco.

Um grande problema dos países Ibero-Americanos em geral, e em especial na Latino América, refere-se à utilização massiva de agrotóxicos (aditivos, antibióticos, pesticidas). O Brasil, por exemplo, é o maior Mercado de agrotóxicos do mundo e o maior país da região (Segundo o Guia), representando 16% da venda mundial. As maiores empresas do setor de agrotóxicos do mundo concentram 80% das suas vendas no país, incluindo o Acefato e o Endosulfam, já proibidos nos EUA, União Européia, Argentina, Nigeria, Senegal, Mauritania, entre outros. Tais substâncias contaminam os alimentos, são nocivas também aos trabalhadores rurais, são predadoras ao meio ambiente, envenenam consumidores e destroem a vida animal e vegetal, além de impactos de longo prazo na população em geral.

Portanto, todos estes dados ajudam a informar as instituições responsáveis que se implicam no controle de riscos ambientais para a saúde, no intuito de controlar e desenvolver medidas preventivas e campanhas para evitar impactos, e promover o bem-estar da população dos países Ibero-Americanos.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

FEA - Faculdade de Economia e Administração

Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA: MORADIA

TRABALHO FINAL

Disciplina: MÉTODOS QUANTITATIVOS E QUALITATIVOS DA PESQUISA EMPÍRICA

Prof. Dr. Arnoldo José de Hoyos Guevara

MESTRANDO – Rafael Rodrigues Cardoso

Trabalho Final

Sumário

CAPITULO 1 – ANÁLISE EXPLORATÓRIA DOS DADOS 7
1. INTRODUÇÃO 7
2. ENTENDENDO OS DADOS 8
2.1 O tema de pesquisa: Moradia 8
2.2 Os indivíduos 10
2.3 As Variáveis do tema: MORADIA 10
2.3.1 A Tabela de Dados 13
3. ANÁLISE DAS VARIÁVEIS 14
3.1 – Variáveis Quantitativas 15
3.2 – Média, Mínimo e Máximo das variáveis com uso do MINITAB 15
3.2.1 – Usando o link principal Data, acessando o Code – Numeric to Numeric para substituir os dados ausentes que aparecem na tabela com o símbolo asterisco (*) pela média encontrada. 17
3.2.2 – Normalização dos dados 18
4. ANÁLISE COMPARATIVA 20
4.1 SPI – Índice de Progresso Social 21
4.2 – Shelter (Abrigo) 23
4.3 – Access to piped water (% of pop.)(água encanada) 26
4.4 - Access to improved sanitation facilities (% of pop)(acesso a instalações sanitárias melhoradas) 27
4.5 – Availability of affordable housing (% satisfied) 29
4.6 – Access to electricity (% of pop.)(eletricidade) 30
4.7 – Quality of electricity supply (1=low; 7=high)(qualidade fornecimento eletricidade) 31
4.8 - Private property rights (0=none; 100=full) 32
4.9 IDH (2013) - Índice de Desenvolvimento Humano (IDH/PNUD) 33
4.10 Índice de Governança 34
4.11 EPI - Pontuação no Índice de Proteção Ambiental 35
4.12 HPI - Pontuação no Happy Planet Index 36
4.13 Índice de GINNI 37
5. Considerações Finais 39
CAPITULO 2 - MODELO DE REGRESSÃO MULTIPLA LINEAR 40
1 RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS 40
1.1 Correlação linear 40
1.2 Gráficos de dispersão 42

1.3 Linhas de tendência 43
1.3.1 Scatterplot de Social Progress Index 43
1.4 Dendograma ou Dendrograma 47
2. Considerações finais 49
CAPITULO 3 – REGRESSÃO LINEAR 50
1. REGRESSÃO 50
1.1 – Normalização dos dados 53
2. STEPWISE 55
CAPITULO 4 – ANÁLISE DE TENDÊNCIAS 59
1. INTRODUÇÃO 59
1.1 Uma breve descrição do SOFI – *State of the Future Index* 59
2. ENTENDENDO OS DADOS 59
2.1 Os Indivíduos 59
2.2 As Variáveis 59
2.3 A fonte e o tamanho da série de dados 60
2.4 A Tabela de Dados – Crescimento do PIB de 1970 a 2014 países selecionados 60
2.6 A Tabela de Dados – Saldo Balança Comercial de 1992 a 2014 países selecionados 61
3. ANÁLISE DA VARIÁVEL PIB 62
3.1 Variável: Crescimento do PIB 62
3.1.1. O comportamento da variável 62
3.1.2. A análise de tendência - Brasil 63
3.1.3. As previsões 64
3.1.4 Análise de tendência – Espanha, México e Portugal 65
4. ANÁLISE DA VARIÁVEL BALANÇA COMERCIAL 67
4.1 Variável: Crescimento da Balança Comercial 67
4.1.1 O comportamento da variável 67
4.2 Análise de tendência Brasil 69
CAPITULO 5 – COMPARAÇÕES: Um Estudo dos Indicadores de Moradia da América Latina Comparado às Demais Regiões do Mundo 72
1. INTRODUÇÃO 72
2. OS INDIVIDUOS 72
2. ENTENDENDO OS DADOS 73
2.2 As Variáveis 73
3. ANÁLISE DAS VARIÁVEIS 75
3.1 One-way ANOVA: Social Progress Index_1N versus Class 6R 76
3.2 One-way ANOVA: Shelter_1N versus N Class 3R 78
3.3 One-way ANOVA: Access to piped water_1N versus N Class 3R 80

3.4 One-way ANOVA: Access to improved sanitation_N versus N Class 3R.....81
3.5 One-way ANOVA: Availability of affordable ho_N versus N Class 3R.....83
3.6 One-way ANOVA: Access to electricity (% of p_N versus N Class 3R.....84
3.9 One-way ANOVA: Access to electricity (% of p_N versus N Class 3R.....85
3.7 One-way ANOVA: Quality of electricity supply_N versus N Class 3R.....86
3.8 One-way ANOVA: Private property rights (0=no_N versus N Class 3R 87
3.9 One-way ANOVA: IDH - 2013_1_N versus N Class 3R.....89
3.10 One-way ANOVA: 1. Governança_1_N versus N Class 3R.....91
3.11 One-way ANOVA: EPI Score_1_N versus N Class 3R.....92
3.12 One-way ANOVA: Happy Planet Index_1_N versus N Class 3R 94
3.13 One-way ANOVA: GINNI P versus N Class 3R 95
CAPITULO 6 – AMOSTRAGEM.....99
1 – INTRODUÇÃO 99
1.1 – Países analisados.....99
2. AS VARIÁVEIS 102
3. PESQUISA POR AMOSTRAGEM 103
3.1 Análise Exploratória para Amostra de 25 Indivíduos.....103
3.2 Social Progress Index.....103
3.3 GOVERNANÇA 106
3.4 PRIVATE PROPERTY RIGHTS.....109
4. CONSIDERAÇÕES FINAIS.....111
CAPITULO 7 - ANÁLISE MULTIVARIADA: Correlações, Dendogramas e Análise dos Componentes Principais 112
1. INTRODUÇÃO 112
2. ENTENDENDO OS DADOS.....112
2.2 As Variáveis.....112
2.3 Fonte de Dados 113
3. CORRELAÇÕES E DENDOGRAMAS 114
3.1 Stepwise PC1 117
3.2 Stepwise PC2 119
3.3 Stepwise PC3 120
3.4 Stepwise PC4 122
4. RENOMEANDO AS NOVAS VARIÁVEIS.....124
4.1 General Regression Analysis: PC1 versus Social Progr; Shelter_1N; Access to im.124

4.1.1 REGRESSÃO PC1 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH 125
4.2 General Regression Analysis: PC2 versus Happy Planet; Private prop; 126
4.3 General Regression Analysis: PC3 versus Availability; GINNI P; Private prop 126
4.4 General Regression Analysis: PC4 versus GINNI P; Availability; Social Progr 127
4.4.1 REGRESSÃO PC4 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH 127
5. CONSIDERAÇÕES FINAIS 128
CAPITULO 8 – ANÁLISE DE CONGLOMERADOS 129
1. INTRODUÇÃO 129
2. FORMAÇÃO DOS CONGLOMERADOS 130
2.1 One-way ANOVA: MeDS versus N Class 3R 132
2.2 One-way ANOVA: McFA versus N Class 3R 133
2.3 One-way ANOVA: UDF versus N Class 3R 134
2.4 One-way ANOVA: IDM versus N Class 3R 134
3. Dendogramas com as principais variáveis de todos os Componentes Principais 136
3.1 Cluster Analysis of Observations: MeDS; DMeQV; UDF; IDM 138
4. ANÁLISE PARA AS REGIÕES AVECO, AIBER E OTHERS 141
4.1 Cluster Analysis of Observations: Social Progr; Shelter_1N; Access to pi; 141
CAPITULO 9 – ANÁLISE DISCRIMINANTE 143
1. ANÁLISE DE COMPONENTES PRINCIPAIS 143
1.1 Regressão PC1 145
1.1.1 REGRESSÃO PC1 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH 146
1.2 Regressão PC2 148
1.3 Regressão PC3 149
1.4 Regressão PC4 151
1.4.1 REGRESSÃO PC4 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH 152
2. FUNÇÃO LINEAR 155
2.1 Discriminant Analysis: 4 clusters versus Social Progr; Shelter_1N; 155
3. FUNÇÃO QUADRÁTICA 156
3.1 Discriminant Analysis: 4 clusters versus Social Progr; Shelter_1N; 156
4. FUNÇÃO LINEAR 158
4.1 Discriminant Analysis: N Class 3R versus Social Progr; Shelter_1N; 158

5. FUNÇÃO QUADRÁTICA 159
5.1 Discriminant Analysis: N Class 3R versus Social Progr; Shelter_1N; 159
6. CLASSIFICAÇÃO NÃO SUPERVISONADA 161
CAPITULO 10 – REGRESSÃO LOGÍSTICA 169
1. Regressão Logística com as principais variáveis 169
2. Regressão Logística retirando variáveis com elevado p-value 170
3. Regressão Logística com regiões em ordem de força 171
4. Teste para Regressão Logística 172
5. Comparação: Análise Discriminante X Regressão Logística 173
CAPITULO 11 - ÁRVORE DE REGRESSÃO 174
1. INTRODUÇÃO 174
2. TÉCNICAS DE CRESCIMENTO EM ÁRVORE 174
2.1 Discriminant 175
2.2 Stepwise Statistics 177
2.3 Summary of Canonical Discriminant Functions 181
2.4 Classification Statistics 182
3. ANÁLISE CHAID EXAUSTIVE 184
4. ANÁLISE CHAID 186
5. ANÁLISE CRT 188
6. ANÁLISE QUEST 190
7. ANÁLISE LOGÍSTICA 193
CAPITULO 12 – ANÁLISE POR CORRESPONDÊNCIA 195
1. INTRODUÇÃO 195
2. Análise por Correspondência das Médias 195
CAPITULO 13 – ANÁLISE POR CORRESPONDÊNCIA SOMENTE AIBER 203
1. INTRODUÇÃO 203
2. Results for: Somente Aiber 203
3. MAPA DE DESENVOLVIMENTO PARA MORADIA AIBER 205
3.1 One-way ANOVA: IDH - 2013_1_N versus 3cluster 206
3.2 One-way ANOVA: PC1 versus 3cluster 206
4. CONSIDERAÇÕES FINAIS 210

CAPITULO 1 – ANÁLISE EXPLORATÓRIA DOS DADOS

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados relativos ao tema da Moradia que possui um conjunto de treze variáveis previamente selecionadas que refletem o desenvolvimento humano nesta área considerando 132 países, conforme mostrado no Quadro 1 que apresenta as Variáveis e suas respectivas categorizações, bem como o tipo da variável, unidade de medida e fontes, enquanto a Tabela 1 nos apresenta algumas variáveis relacionadas aos 132 países.

Iniciamos o trabalho apresentando o objetivo do Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável, uma publicação do Núcleo de Estudos do Futuro da PUC/SP – Pontifícia Universidade Católica. Depois partimos para o entendimento dos dados apresentando: média, mínimo, máximo, padronização e normalização dos dados e, por fim, a análise exploratória dos dados sobre o tema: **Moradia** empregando o software estatístico MINITAB, em especial seus recursos: Display Descriptive Statistics que está no link principal Basic Statistics, que nos permite obter a média, o valor mínimo e o valor máximo do objeto estudado.

Seguimos com a aplicação do link principal Data, acessando o Code – Numeric to Numeric, sendo que este recurso permite que os dados ausentes que aparecem na tabela com o símbolo asterisco (*) seja substituído pelo valor determinado que é a média encontrada.

Na sequência temos que fazer a normalização dos dados que é um ajuste nos dados com vista a afastar os *outliers*, ou seja, é como se os dados analisados não possuíssem uma distribuição normal, então fazemos o ajuste para que todos fiquem dentro de uma certa normalidade. No nosso caso, usamos como parâmetro média entre o mínimo e o máximo obtido na coleta de dados. Temos que apresentar os valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas.

Para essa etapa usamos o link da barra de ferramentas do MINITAB denominada Calculator e aplicamos a fórmula desejada indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.

Por fim, apresentamos os gráficos gerados pelo recurso Graphical Summary que está no item Basic Statistic encontrado no item Static na barra de ferramentas.

O Objetivo é explorar os dados sobre **Moradia** de tal modo que possamos com os recursos citados anteriormente apresentar conclusões de correlações entre as variáveis que compõem o tema.

2. ENTENDENDO OS DADOS

Antes da apresentação dos indivíduos desta análise que envolve 132 países e os indicadores de desenvolvimento humano selecionados a partir de referência internacional, citamos a origem do trabalho que é o Projeto ORIBER – Observatório da Rede Ibero-Americana de Prospectiva[1].

2.1 O tema de pesquisa: Moradia

A análise dos conceitos que compõem o tema foi extraída do material de aula Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica. Nesse estudo são apresentadas as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso.

Vale destacar o desafio citado no GPS e sua proposta de solução:

Desafio: *"A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?"*

Solução: "GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de

[1] Para conhecer melhor o RIBER e o ORIBER é possível acessar o sítio do Núcleo de Estudos do Futuro no link da Cátedra Ignacy Sachs. Disponível em: <nef.org.br/catedra>

informações e experiências com outros países da região ibero-americana e suas redes de contatos.”

Para chegar a uma conclusão do estágio em que se encontra o país e como é possível adotar uma solução, o GPS propõe um levantamento dos dados a partir da análise SWOT, conforme se apresenta na figura 1.

Figura 1: Análise SWOT de cada país

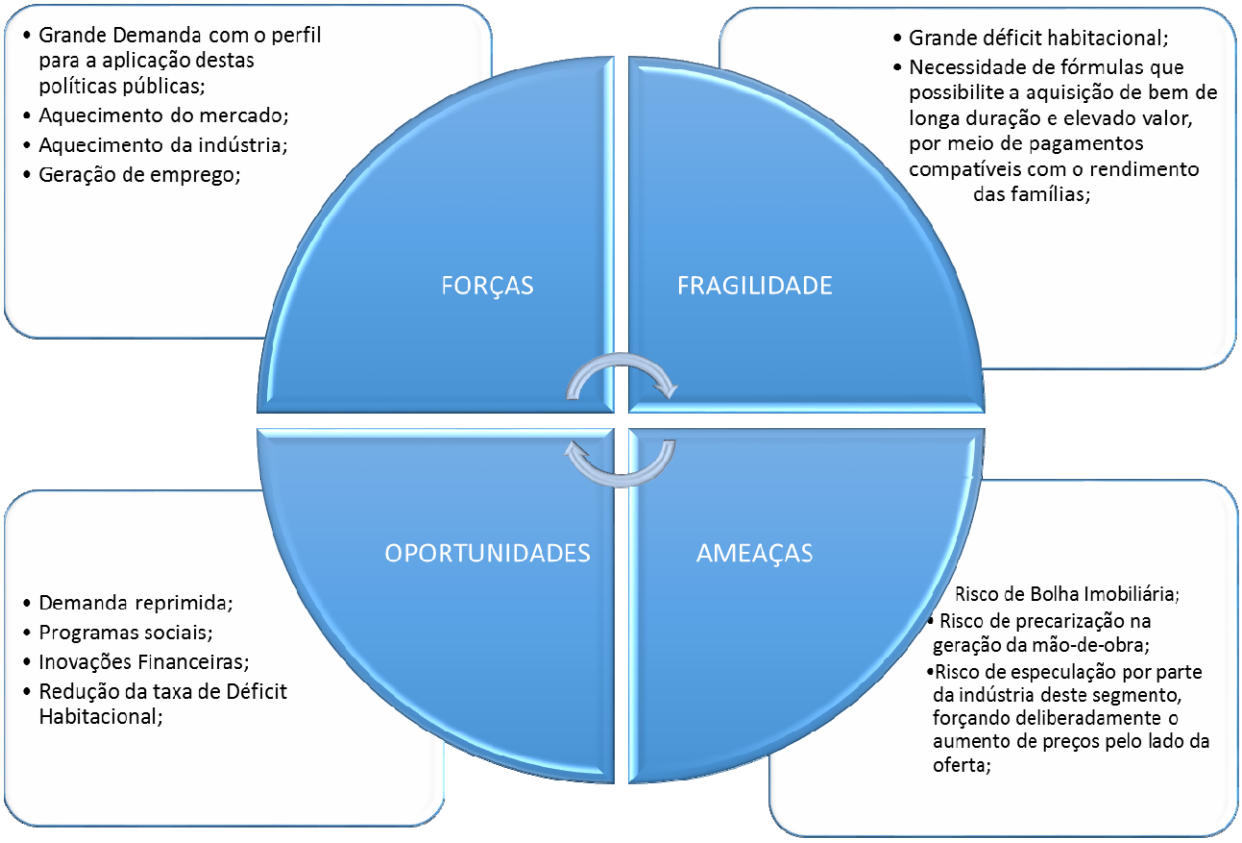
fonte: autor, adaptado de GPS, p.8

Na figura 1 é possível obter informações sobre o país analisado considerando sua Forças (*strenghts)* – quais são os seus pontos fortes em cada um dos eixos temáticos. As fragilidades *(weaknesses),* ameaças *(threats)* e oportunidades *(opportunities)* que cada eixo temático apresenta para o país em questão.

Como exemplo inicial desse trabalho que tem o tema: **Moradia** podemos destacar o cenário em que um país possui grande déficit habitacional (fragilidade) e que está densamente povoado em sua área urbana, forçando a expulsão das famílias para áreas de risco (ameaças), mas que começam a conceder incentivos à sua população mais carente, por meio de programas sociais e inovações financeiras que permita a aquisição do imóvel com pagamento em maior prazo e políticas de juros socialmente justa (oportunidades), motivando a indústria da construção civil e gerando empregos a partir do aumento da demanda que aqueceu tanto o mercado de trás - aquele que alimenta a indústria - fornecendo aço, metal, vidro e concreto, quanto o mercado da frente - após a sua construção - como cerâmica, eletrodomésticos e mobiliários para a nova moradia (Forças).

Para ajudar o leitor a entender a contribuição da matriz SWOT, apresentamos um exemplo com o tema: Moradia, citado na figura 2.

Figura 2: Análise SWOT - Moradia



Fonte: Elaboração Própria.

Se tomarmos como ponto de partida o quadrante do Ponto Forte e avançarmos no sentido horário da matriz, até alcançarmos as Oportunidades, perceberemos a alta complexidade na resolução do desafio do déficit habitacional, tornando o estudo da moradia um desafio perene, levando em consideração que este está intrinsicamente ligado a interesses por todos os lados, desde os beneficiários, passando pelo Governo e principalmente pelas incorporadoras e construtoras.

2.2 Os indivíduos

Os indivíduos desta análise são os 132 países analisados por uma série de indicadores de referência internacional. Os dados analisados de cada país são as variáveis que descrevemos a seguir.

2.3 As Variáveis do tema: MORADIA

Este tema da pesquisa apresenta treze variáveis, sendo todas elas quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis-componentes ou indicadores, a saber:

- **Índices sintéticos:** São sete: Índice de Progresso Social, Shelter, Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD/ONU, Índice de Governança, EPI (*Environmental Protection Index*), o *Happy Planet Index* (HPI), e o Índice de Gini para os países selecionados.

- **Indicadores ou variáveis componentes:** Inclui seis variáveis como: Access to piped water (% of pop.)(água encanada), Access to improved sanitation facilities (% of pop.)(acesso a tratamento sanitário melhorado), Availability of affordable housing (% satisfied)(disponibilidade e acessibilidade do imóvel), Access to electricity (% of pop.)(eletricidade), Quality of electricity supply (1=low;7=high)(qualidade fornecimento eletricidade) e Private Property Rights (0=none; 100=full)(direitos de propriedade privada), conforme apresentado no Quadro 1 que classifica o tipo de variável e a unidade de medida.

Quadro 1. As **13** Variáveis do Eixo temático: **Moradia**

Variável	Significado	Tipo	Unidade de Medida	Fonte
SOCIAL PROGRESS INDEX	O índice identifica os pontos fortes e fracos dos países a partir de um conjunto de informações e análises conduzidas para cada um deles.	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
SHELTER (Abrigo)	Este índice indica se as pessoas possuem abrigos (lar) com suas necessidades básicas atendidas?	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
Acesso a Água e saneamento Básico	Porcentagem da população que tem uma ou mais torneiras de água encanada tratada.	Variável Quantitativa	Percentual	WHO UNICEF
Acesso a tratamento sanitário Melhorado	Inclui a população com melhores condições de saneamento, canalização de esgoto, fossas sépticas, latrinas com laje ou melhoradas para as condições sanitárias adequadas	Variável Quantitativa	Percentual	WHO UNICEF
Disponibilidade e Acessibilidade do Imóvel	Percentual da população que responderam em pesquisa se “em sua cidade ou região em que vive, está satisfeito com a disponibilidade e acessibilidade do imóvel”	Variável Quantitativa	Percentual	Galoop World Poll
Acesso a Eletricidade	Percentual de pessoas com acesso a energia elétrica.	Variável Quantitativa	Percentual	UN Sustain Able energy for All Project

Variável	Significado	Tipo	Unidade de Medida	Fonte
Qualidade no fornecimento de energia elétrica	Média das respostas para a pergunta: "Em seu país, como você avalia a confiança no fornecimento de energia?". Usando Escala onde 1 não é confiável e 7 extremamente confiável para mostrar a qualidade do fornecimento de energia elétrica.	Variável Quantitativa	Indicador nominal que varia de 1 a 7	World Economc Forum Global Competitiveness Report
Direitos de Propriedade Privada	O grau com que um país protege com leis os Direitos com Propriedade Privada. Medido em escala onde 0 corresponde a nenhuma proteção pelo estado ao direito de propriedade privada e elevado grau de corrupção e 100 corresponde ao total suporte do Governo e nenhum grau de corrupção.	Variável Quantitativa	Indicador nominal que varia de 0 a 100	Heritage Foundation
IDH Índice de Desenvolvimento Humano	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável Quantitativa	Indicador nominal que varia de 0 a 1	PNUD
Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as políticas são implementadas e qual a capacidade de formulá-las.	Variável Quantitativa	Percentual	WGI 2014
EPI- Indice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	Variável Quantitativa	Percentual	EPI 2014
HPI – Bem estar sustentável do país	Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	Variável Quantitativa	Percentual	HPI 2014
GINNI index	Mede a distribuição de renda e de despesas das famílias	Variável Quantitativa	Percentual	Banco Mundial

Fonte: Elaboração própria; IPS, Unicef, GWP, WEF, Heritage Foundation, PNUD, WGI, EPI, Banco Mundial

2.3.1 A Tabela de Dados

Em função da extensão da tabela, vamos representar aqui **apenas os 30 primeiros países quando filtrados em ordem alfabética** e **algumas das principais** variáveis para corroborar o Quadro 1.

Tabela 1: Países e variáveis selecionadas

País	1 Social Progress Index	7 Shelter	55 (0=none; 100=full)	94 IDH - 2013	106 1. Governança	138 EPI Score	143 Happy Planet Index	145 GINNI Index
Albania	69,13	66,72	30	0,716	48,51	54,73	54,05118	28,96
Algeria	59,13	68,57	30		34,23	50,08	52,181298	35,33
Angola	39,93	28,09	15	0,526	27,83	28,69	33,201432	42,66
Argentina	70,59	60,75	15	0,808	45,15	49,55	54,055042	43,57
Armenia	65,03	62,92	30	0,73	50,16	61,67	46,003186	30,3
Australia	86,10	84,12	90	0,933	93,03	82,4	41,979812	34,01
Austria	85,11	88,38	90	0,881	92,06	78,32	47,085135	30,04
Azerbaijan	62,44	77,50	20	0,747	37,16	55,47	40,88457	33,03
Bangladesh	52,04	42,91	20	0,558	32,16	25,61	56,292001	
Belarus	65,20	74,99	20	0,786	33,73	67,69	37,414602	26,46
Belgium	82,63	86,11	80	0,881	88,16	66,61	37,090528	33,14
Benin	49,11	31,47	30	0,476	45,53	32,42	31,083209	43,53
Bolivia	62,90	53,78	10	0,667	40,67	50,48	43,577908	46,64
Bosnia and Herz.	64,99	72,76	20	0,731	48,49	45,79	42,35456	33,04
Botswana	65,60	51,54	70	0,683	70,44	47,6	22,591174	60,46
Brazil	69,97	73,20	50	0,744	53,73	52,97	52,931996	52,67
Bulgaria	70,24	62,83	30	0,777	57,21	64,01	34,145379	34,28
Burkina Faso	47,33	18,43	30	0,388	42,42	40,52	31,79385	39,78
Burundi	37,33	10,96	20	0,389	27,23	25,78	30,515012	33,27
Cambodia	51,89	28,96	30	0,584	36,02	35,44	40,322921	31,82

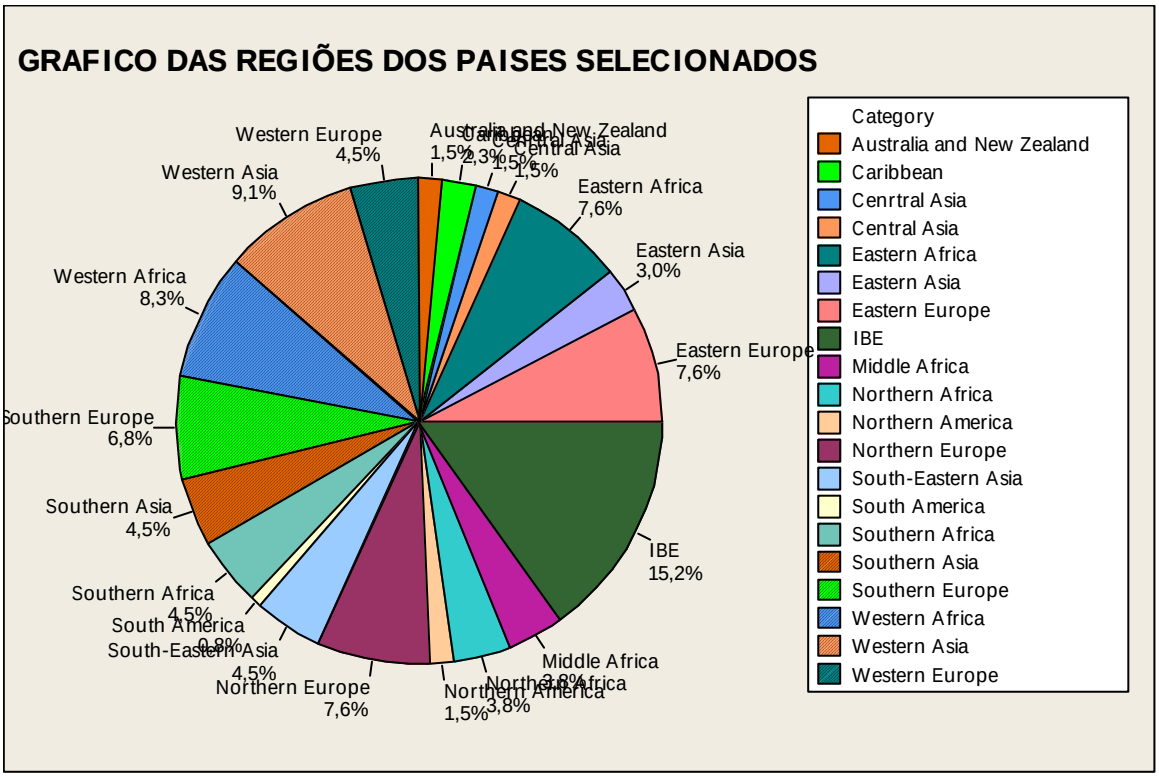
Cameroon	45,51	39,15	30	0,504	31,34	36,68	33,686686	40,72
Canada	86,95	88,64	90	0,902	93,91	73,14	43,559839	33,68
Central African Rep.	34,17	9,61	10	0,341	15,99	42,94	25,256461	56,3
Chad	32,60	6,38	20	0,372	23,02	31,02	24,681898	43,3
Chile	76,30	76,37	90	0,822	83,61	69,93	53,883338	52,06
China	58,67	72,75	20	0,719	41,36	43	44,660738	42,06
Colombia	67,24	76,88	50	0,711	47,25	50,77	59,751023	53,53
Congo, Republic of	47,99	26,20	10	0,338	27,75	39,44	34,547037	40,17
Costa Rica	77,75	78,83	50	0,763	69,96	58,53	64,035926	48,61
Croatia	73,31	81,95	40	0,812	64,83	62,23	40,623723	33,61

Fonte: Elaboração Própria com a Base de Dados da aula Métodos Quantitativos

3. ANÁLISE DAS VARIÁVEIS

Apresentamos o gráfico circular ou gráfico de pizza ou gráfico de torta elaborado pelo MINITAB com as regiões dos países selecionados**.**

Gráfico 1: Região dos Países Selecionados



3.1 – Variáveis Quantitativas

Como apresentado na terceira coluna e corroborado pelo indicador de medida na quarta coluna do Quadro 1, as variáveis são denominadas quantitativas. Essa informação aparece na Tabela 1 com uma seleção de países e indicadores com os números em percentual.

Com isso, a variável é denominada quantitativa pois mostra as que podem ser medidas em escala quantitativa, isto é, com números (SANTOS; PARRA FILHO, 2011).

Elas são Variáveis Contínuas, porque assumem valores fracionados já que estão em percentual **(IDH e GINNI).**

3.2 – Média, Mínimo e Máximo das variáveis com uso do MINITAB

Para facilitar o acompanhamento das análises vamos expor os dados do MINITAB citando com títulos que representem a situação de estudo.

- Valor de N é o total de elementos da amostra que neste caso é de 132 países;
- **N*** indica o total de dados ausentes na tabela em cada variável. Por exemplo, a variável access do piped water (%of population) possui três países sem dados, a saber: Austrália, Kuait e Estados Unidos.
- Mean (média) mostra o valor médio da variável para os 132 países.
- SE Mean mostra o erro padrão da amostra, isto é, ele mostra o quanto a média varia em relação as outras da amostra.
- StDev é o desvio padrão e mostra a distância estimada entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação à média.
- Minimum mostra o valor mínimo dos dados dos 132 países
- Maximum mostra o valor máximo dos dados dos 132 países

Tabela com dados originais da base de dados do tema: **Moradia**

Descriptive Statistics: Social Progr; Shelter; Access to pi; Access to im; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93
Shelter	132	0	60,85	1,97	22,58	6,38	41,65
Access to piped water (%	129	3	63,98	3,03	34,36	2,53	24,71
Access to improved sanit	130	2	73,75	2,55	29,09	9,58	52,97
Availability of affordab	132	0	44,57	1,24	14,22	12,00	34,25
Access to electricity (%	132	0	78,33	2,80	32,23	4,00	58,00
Quality of electricity s	121	11	4,515	0,142	1,567	1,273	3,223
Private property rights	130	2	43,88	2,11	24,03	5,00	30,00
IDH - 2013	125	7	0,6894	0,0144	0,1611	0,3370	0,5595
1. Governança	132	0	53,38	1,88	21,59	15,22	37,11
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780

Variable	Median	Q3	Maximum
Social Progress Index	63,86	73,74	88,24
Shelter	65,95	78,81	91,81
Access to piped water (%	75,33	95,34	100,00
Access to improved sanit	88,00	98,65	100,00
Availability of affordab	44,00	54,00	85,00
Access to electricity (%	99,00	100,00	100,00
Quality of electricity s	4,848	5,660	6,754
Private property rights	40,00	60,00	95,00
IDH - 2013	0,7220	0,8200	0,9440
1. Governança	48,46	70,36	100,00
EPI Score	52,09	66,40	87,67
Happy Planet Index	42,463	49,286	64,036
GINNI Index	38,160	44,550	63,140

Análise: Selecionamos o Índice de Progresso Social para esse comentário que mostra a Média de 63,67 em uma escala até 100, indicando que de uma maneira geral os países estão dentro da escala normal. Embora o valor mínimo observado é de 32,60 referente ao Chad no centro da África com 12,300 milhões de pessoas e tem seus vizinhos: Líbia, Egito, Nigeria e Sudão. E o valor máximo observado foi 88,24 para a Nova Zelândia que é um país formado por um conjunto de ilhas sendo as maiores denominadas: Ilha do Norte e Ilha do Sul. Por ter uma característica de economia desenvolvida esse indicador confirma essa qualidade.

3.2.1 – Usando o link principal Data, acessando o Code – Numeric to Numeric para substituir os dados ausentes que aparecem na tabela com o símbolo asterisco (*) pela média encontrada.

Foram substituídos:

Access to piped water (% of pop.)(água encanada) incluímos em três países a média 63,98: Austrália, Kwait, Estados unidos

Access to improved sanitation facilities (% of pop)(acesso a instalações sanitárias melhoradas) Aqui tivemos dois países que são Itália e Nova Zelândia com média de 73,75.

Quality of electricity supply (1=low; 7=high)(qualidade fornecimento eletricidade) esse dado está com 11 países sem a informação, logo o uso do recurso Data, acessando o Code – Numeric to Numeric para substituir os dados ausentes que aparecem na tabela com o símbolo asterisco (*) pela média encontrada foi extremamente útil.

Os países são: Belarus (Bielorussia), República Centro Africana, República do Congo, Cuba, Djibout ou Jibout (Republica do Djibuti – Africa), Iraque, Nigéria, Sudão, Tajiquistão, Togo, Uzebesquitão. Foi substituído o dado ausente pela média de 4,515.

Private Property Rights apresentava a ausência de dados para dois países: Iraque e Sudão.

IDH – Índice de Desenvolvimento Humano contava com sete países sem a informação: Argélia, Cazaquistão, República da Coreia, Quirguistão, Letônia, República da Mauricia, Holanda.

Happy Planet Index(bem estar sustentável - o país oferece vida feliz sustentável) com três países: Lesoto, Montenegro, Reino da Suazilândia (África austral) que receberam o valor de 42,673.

GINNI Index estava com 13 países sem informação: Bangladesh, Cuba, Nova Guiné, Islandia, Kwite, Libano, República da Mauricia, Montenegro, Nepal, Paquistão, Arabia Saudita, Sri Lanka, Emirados Arabes. Esses países receberam o valor da média: 39,095.

Tabela Confirmando a eliminação dos dados ausentes depois que foram substituídos pela Média de cada uma das variáveis.

Descriptive Statistics: Social Progr; Shelter_1; Access to pi; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index_1	132	0	63,67	1,24	14,20	32,60	51,93
Shelter_1	132	0	60,85	1,97	22,58	6,38	41,65
Access to piped water (%	132	0	63,98	2,96	33,97	2,53	26,30
Access to improved sanit	132	0	73,75	2,51	28,87	9,58	53,00

Availability of affordab	132	0	44,57	1,24	14,22	12,00	34,25
Access to electricity (%	132	0	78,33	2,80	32,23	4,00	58,00
Quality of electricity s	132	0	4,515	0,131	1,500	1,273	3,305
Private property rights	132	0	43,88	2,08	23,84	5,00	30,00
IDH - 2013_1	132	0	0,6894	0,0136	0,1567	0,3370	0,5700
1. Governança_1	132	0	53,38	1,88	21,59	15,22	37,11
EPI Score_1	132	0	52,48	1,44	16,55	18,43	39,30
Happy Planet Index_1	132	0	42,673	0,778	8,936	22,591	36,587
GINNI Index_1	132	0	39,024	0,705	8,097	24,820	33,065

Variable	Median	Q3	Maximum
Social Progress Index_1	63,86	73,74	88,24
Shelter_1	65,95	78,81	91,81
Access to piped water (%	74,80	95,03	100,00
Access to improved sanit	86,37	98,54	100,00
Availability of affordab	44,00	54,00	85,00
Access to electricity (%	99,00	100,00	100,00
Quality of electricity s	4,707	5,628	6,754
Private property rights	40,00	58,75	95,00
IDH - 2013_1	0,7155	0,8147	0,9440
1. Governança_1	48,46	70,36	100,00
EPI Score_1	52,09	66,40	87,67
Happy Planet Index_1	42,627	49,175	64,036
GINNI Index_1	39,024	43,553	63,140

Análise: Após a substituição dos dados ausentes identificados por asterisco (*) pela média de cada variável foi rodado novamente os indicadores de número de elementos (N), dados ausentes (N*), média, erro da média (SE mean), desvio padrão (Stdev), valor mínimo e máximo; apenas para confirmar se foram lançados corretamente. Portanto, sem alteração nos respectivos valores.

3.2.2 – Normalização dos dados

Para a normalização foram empregadas as seguintes rotinas:

1 – Estabeleceu como parâmetro para normalização o valor máximo e o mínimo.

Primeiro: Clique em Calc no menu e abrirá uma janela, então, siga a rotina: Clique em calc (calculadora no MINITAB) em seguida aparece várias opções, então, clique em Calculator e aparece uma janela pedindo: Store result in variable (indique a coluna para onde vai o dado) logo abaixo está: Expression insira a fórmula: **(c9-MIN(c9))/(MAX(c9)-MIN(c9))*100**.

Importante lembrar que o C9 aqui é o número da coluna do exemplo e você deve substituir pela coluna de sua pesquisa que quer alterar);
Depois clique OK e vai aparecer os dados alterados na coluna que você indicou.

2 – Deixar os resultados com duas casas depois da virgula, foi empregado a fórmula: 0,01*ROUND(100*número da coluna desejada).
Exemplo: se o interesse era alterar os dados da Coluna 22, então no lugar do número da coluna inserimos C22 e o MINITAB gerou os valores com duas casas depois da virgula.
Rotina: Clique em calc (calculadora no MINITAB) em seguida aparece várias opções, então, clique em Calculator e aparece uma janela pedindo: Store result in variable (indique a coluna que quer alterar) logo abaixo está: Expression (insira a expressão 0,01*ROUND(100*número da coluna desejada).

Após a normalização e deixar os resultados com duas casas depois da vírgula rodamos novamente os valores de estatística descritiva: N, N*, Mean, SE mean, StDev, Minimo e máximo; conforme se apresentam a seguir:

Descriptive Statistics: Social Progr; Shelter_1_1; Access to pi; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index_1_	132	0	55,84	2,22	25,52	0,00	34,74
Shelter_1_1	132	0	63,76	2,30	26,43	0,00	41,28
Access to piped water (%	132	0	63,05	3,03	34,85	0,00	24,39
Access to improved sanit	132	0	70,97	2,78	31,93	0,00	48,02
Availability of affordab	132	0	44,61	1,69	19,47	0,00	30,48
Access to electricity (%	132	0	77,43	2,92	33,57	0,00	56,25
Quality of electricity s	132	0	59,15	2,38	27,36	0,00	37,08
Private property rights	132	0	43,21	2,31	26,49	0,00	27,78
IDH - 2013_1_1	132	0	58,05	2,25	25,81	0,00	38,38
1. Governança_1_1	132	0	45,02	2,22	25,46	0,00	25,83
EPI Score_1_1	132	0	49,18	2,08	23,90	0,00	30,14
Happy Planet Index_1_1	132	0	48,45	1,88	21,56	0,00	33,77
GINNI Index_1_1	132	0	37,07	1,84	21,13	0,00	21,52

Variable	Median	Q3	Maximum
Social Progress Index_1_	56,19	73,94	100,00
Shelter_1_1	69,73	84,78	100,00
Access to piped water (%	74,15	94,91	100,00

Access to improved sanit	84,92	98,39	100,00
Availability of affordab	43,84	57,53	100,00
Access to electricity (%	98,96	100,00	100,00
Quality of electricity s	62,66	79,45	100,00
Private property rights	38,89	59,72	100,00
IDH - 2013_1_1	62,36	78,71	100,00
1. Governança_1_1	39,21	65,05	100,00
EPI Score_1_1	48,61	69,28	100,00
Happy Planet Index_1_1	48,34	64,14	100,00
GINNI Index_1_1	37,07	48,88	100,00

Os dados foram normalizados, isto é, foram eliminados os outliers pelo valor mínimo de cada variável e foram lançados nas colunas do MINITAB de número C36 até C48. A seguir apresentamos os dados da média, mínimo e máximo considerando as colunas com os dados normalizados, isto é, de número C36 a C48.

Análise: Após a normalização dos dados observou-se que as variáveis tiveram alteração porque o critério estabelecido foi a diferença entre o valor máximo e o valor mínimo, dado pela expressão: (c9-MIN(c9))/(MAX(c9)-MIN(c9))*100. Novamente o registro de que o C9 é apenas referência da coluna no exemplo.

Então, a coluna C4 que representa Social Progress Index originalmente apresentava uma média de: 63,67 e com a normalização passou para 55,84 com duas casas depois da vírgula.

Isto foi observado em outras variáveis como a coluna C5 que mostra os dados de Shelter (lar) com valor médio original de 60,85 e passou com a normalização para 63,76. Isto não elimina o poder de interpretação e análise mesmo com a diferença para maior.

- Considerando o desvio padrão (StDev) – que mostra a distância estimativa entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação à média. Nesse caso podemos citar como exemplo IDH com média de 58,13 e desvio de 25,82. Outros indicadores serão explorados ao longo do trabalho.

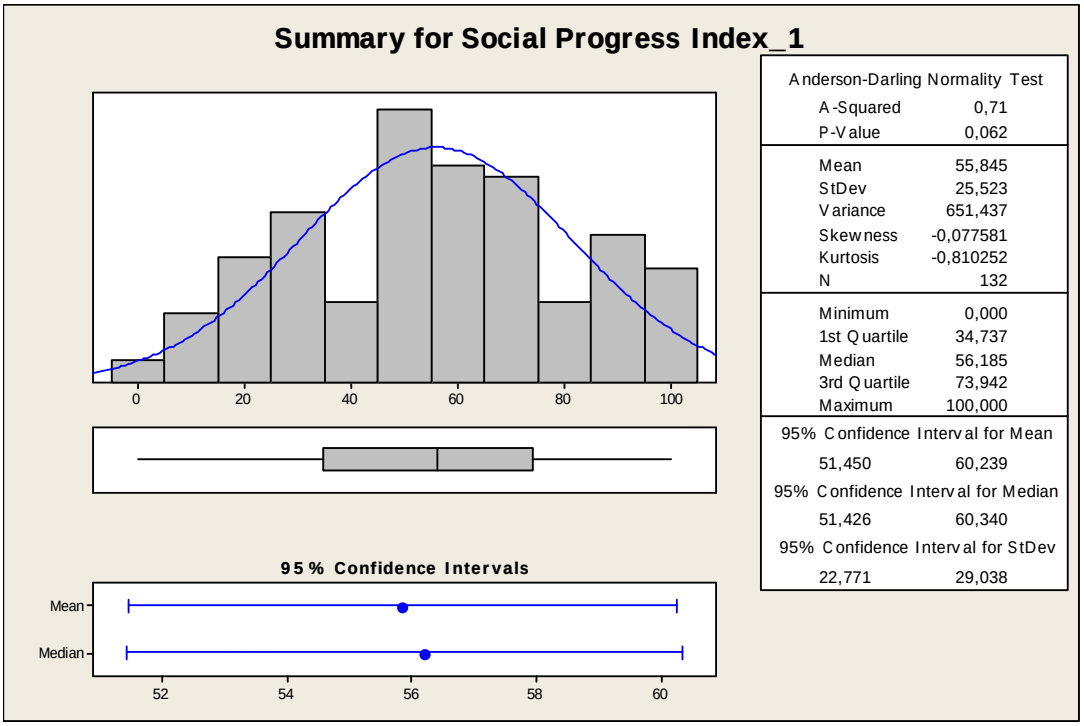
4. ANÁLISE COMPARATIVA

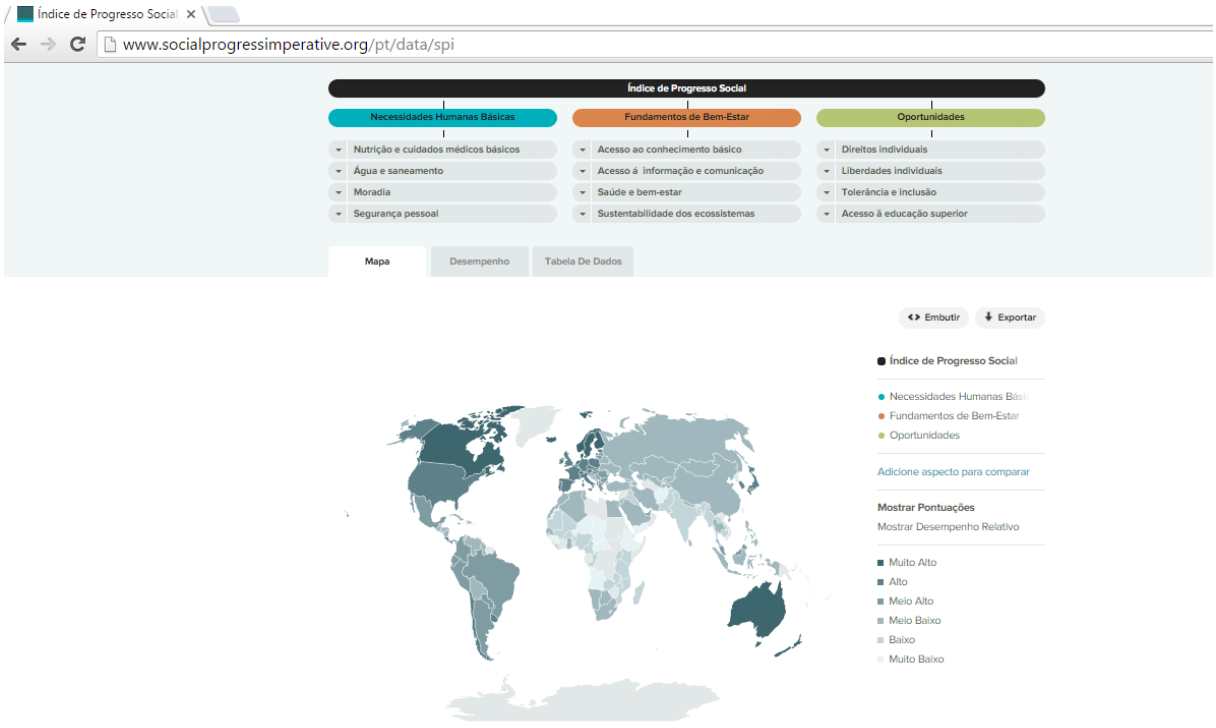
Aqui são apresentados os gráficos gerados pelo Graphical Summary de acordo com a rotina:

Primeiro: clique em Stat em seguida passe o mouse sobre Basic Statistics e aparece uma janela onde o Graphical Summary é o terceiro de cima para baixo. Clique nele.

Segundo: Após clicar, abriu uma janela e você deve escolher a coluna que quer o gráfico na janela da Variables logo a primeira. Está escrito o grau de confiança de 95,0. Clique OK e o gráfico será gerado em uma janela. Copie e cole no seu trabalho.

4.1 SPI – Índice de Progresso Social





Fonte: Sítio do Social Progress Imperative. Disponível em: >http://www.socialprogressimperative.org/pt/data/spi

Análise:

Variável: Social Progress Index – índice de Progresso Social

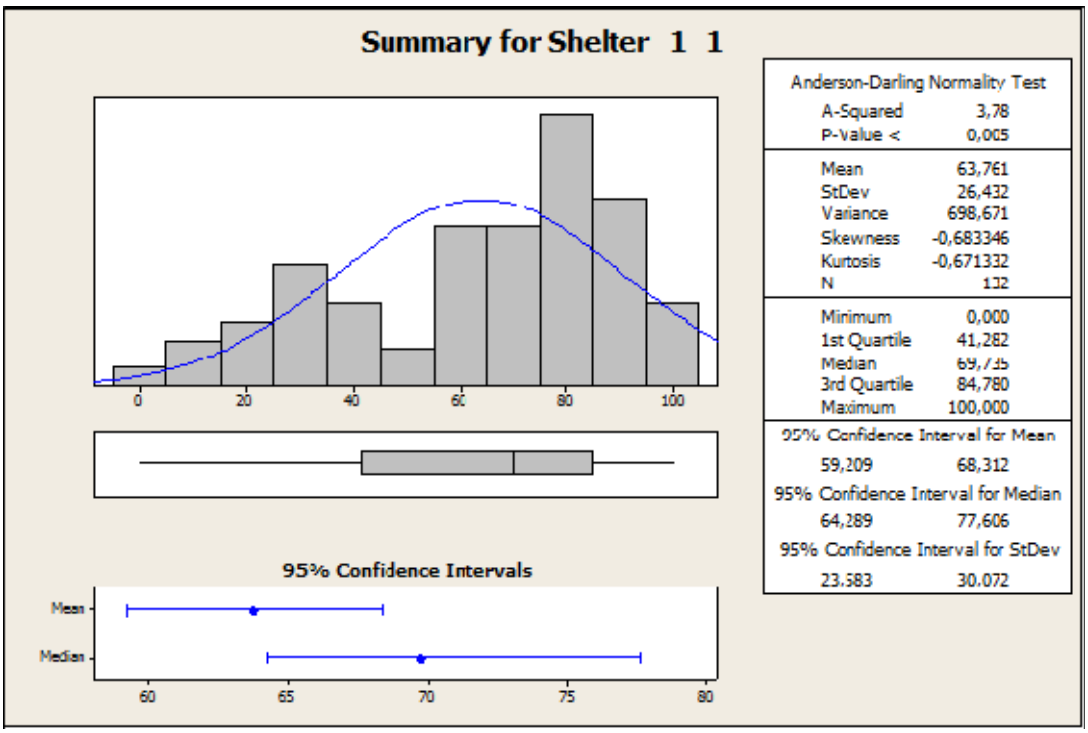
Origem e definição dado pelo portal: http://pt.knoema.com/jklbfre/the-social-progress-index

O Índice de Progresso Social é o resultado de um processo de dois anos de pesquisas envolvendo uma equipe renomada de especialistas que inclui, entre outros, os economistas Hernando de Soto e Michael Porter e do presidente da Fundação Rockefeller Dr. Judith Rodin. O índice sintetiza um total extenso de dados pesquisados para identificar as dimensões do desempenho das sociedades e medir o progresso social de forma abrangente e rigorosa. O Índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidadas em três dimensões do Progresso Social: necessidades humanas básicas, Fundações de Bem-estar e oportunidade. A primeira dimensão, necessidades humanas básicas, avalia o quão bem um país prevê necessidades essenciais de seu povo através da medição se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se tiverem acesso a água potável, se tiverem acesso a uma habitação condigna com serviços básicos, e se eles são seguros e protegidos.

Análise dos dados da nossa pesquisa mostra que a distribuição aproxima-se da normal, considerando os 132 países.

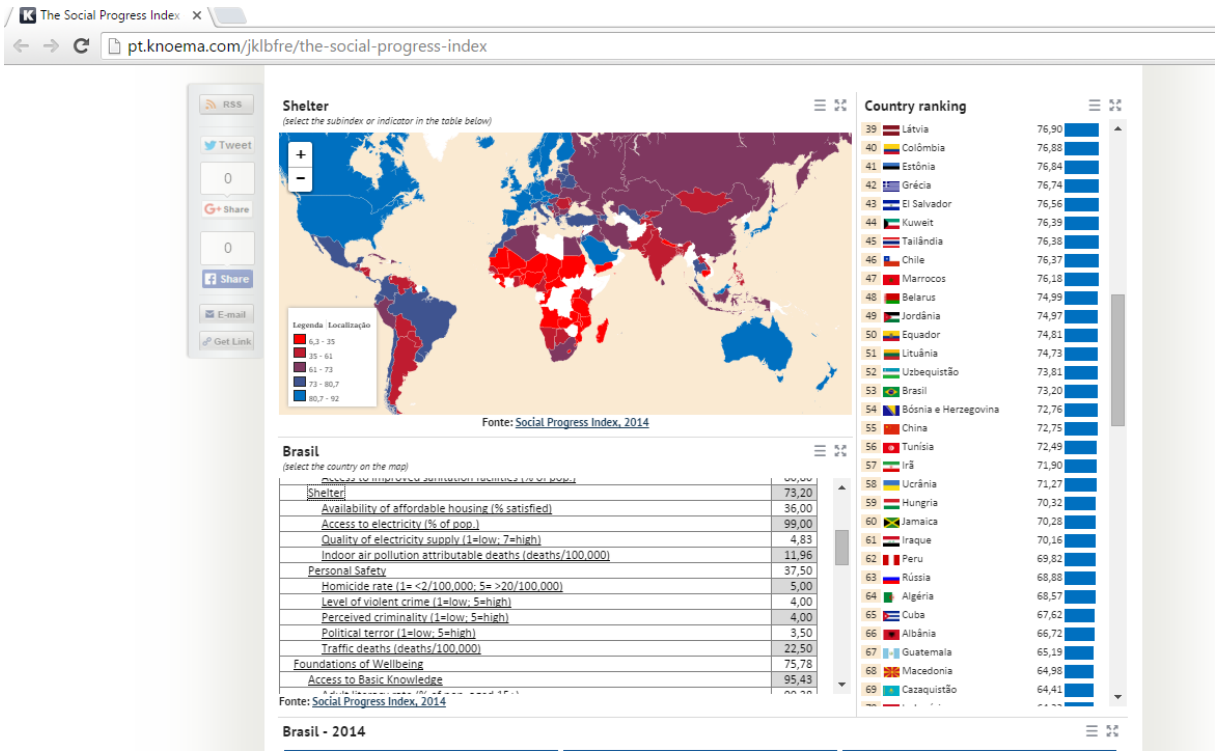
Percebe-se um número menor de países do que o esperado classificados entre 35 e 45 com apenas 7 observações, e o mesmo fenômeno se repete para a faixa de pontuação 75 à 85, que também apresenta apenas 7 observações. O maior número de observações encontra-se na faixa de 45 a 55, com 24 observações. Esta faixa corresponde a última faixa imediatamente inferior à mediana e corresponde a um nível de desenvolvimento considerado “médio” pelo índice. A mediana de 56,185 e o terceiro quartil de 73,942 indicam o poder discricionário do índice, que consegue separar o grupo de países em 11 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados.

4.2 – Shelter (Abrigo)



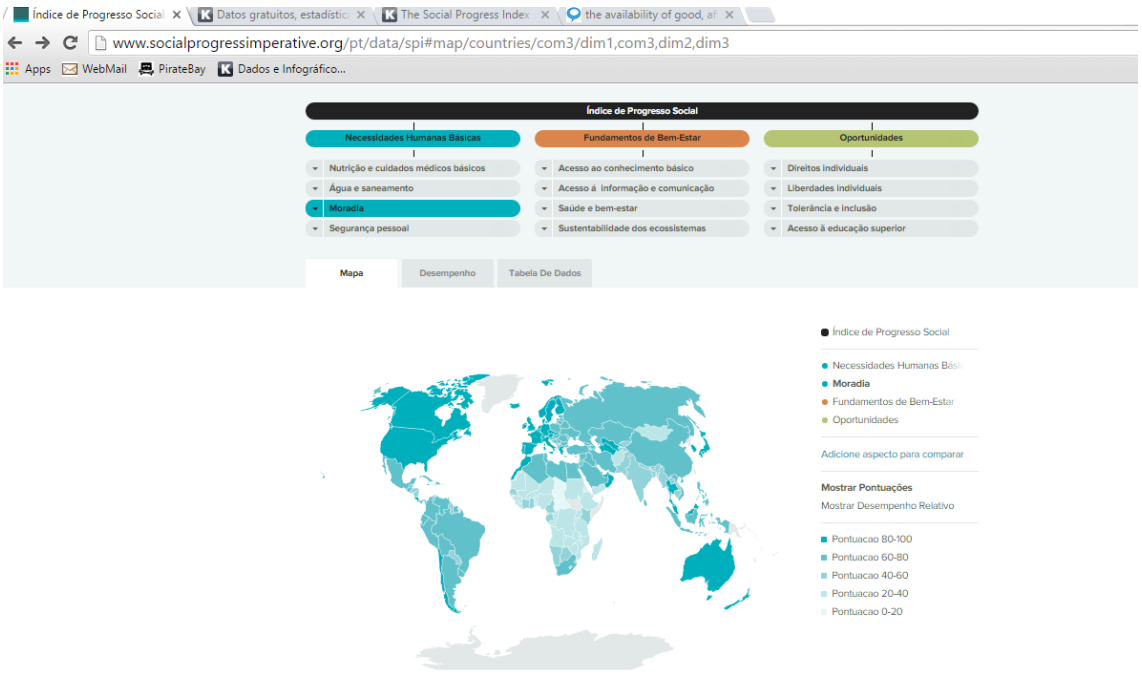
Análise:

Variável Shelter: De acordo com o sítio de fornecimento de dados e estatísticas *Knoema* a variável *Shelter* (Abrigo) está classificada na dimensão das Necessidades Básicas Humanas e possui como seus indicadores quais as condições de dignidade e qualidade de vida apresentadas.

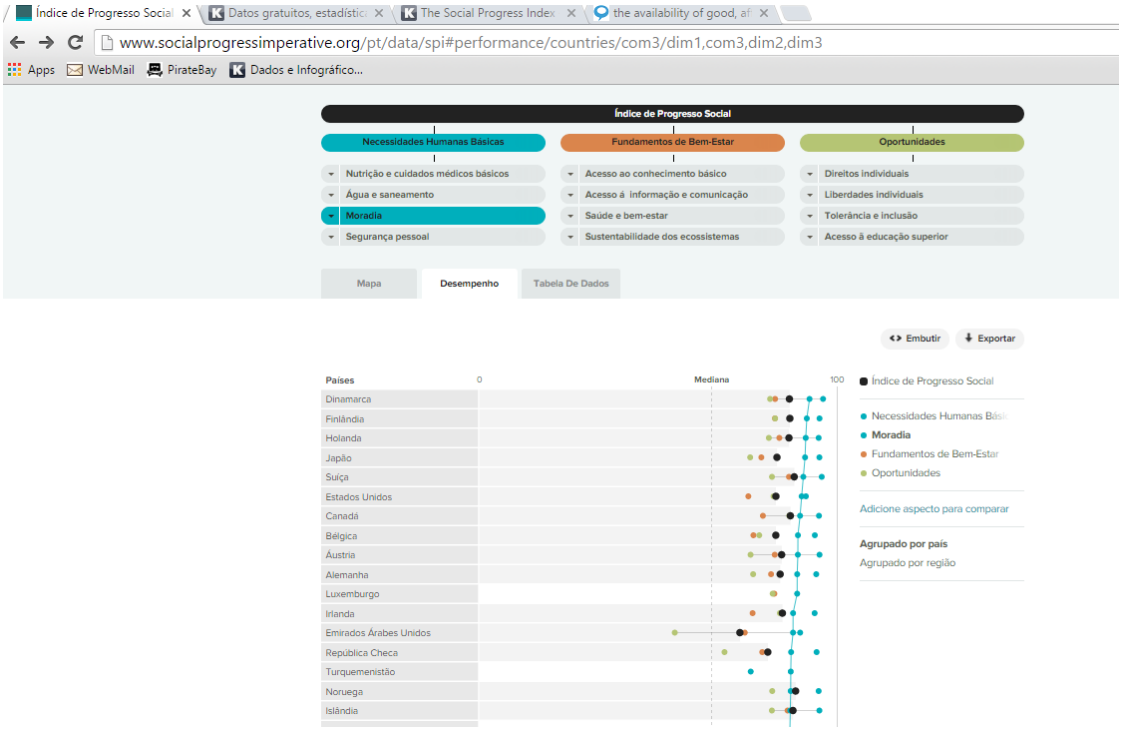


Fonte: Sítio do knoema.com. Disponível em: > http://pt.knoema.com/jklbfre/the-social-progress-index

Análise dos dados da nossa pesquisa: Temos nesta variável um gráfico de distribuição normal, onde ocorre assimetria negativa, ou à esquerda. Podemos observar que a média se encontra em 63,761 com mediana de 69,735. Observamos maior concentração de países no intervalo 75 a 85, mesma faixa do 3º quartil, com 29 países. Na “elite” desta distribuição temos 09 países, à saber: Dinamarca com 91,81% de satisfação, seguida de Finlândia com 90,66%, Japão com 89,74%, Holanda 88,92, Canadá 88,64, Irlanda 88,61, Suíça 88,48, Áustria 88,38 e Estados Unidos com 87,99% de satisfação. Na rabeira temos Central African Republic e Chad com 9,61% e 6,38% respectivamente. O Brasil está com 73,20%, área em que a cauda da curtose inicia sua descida.

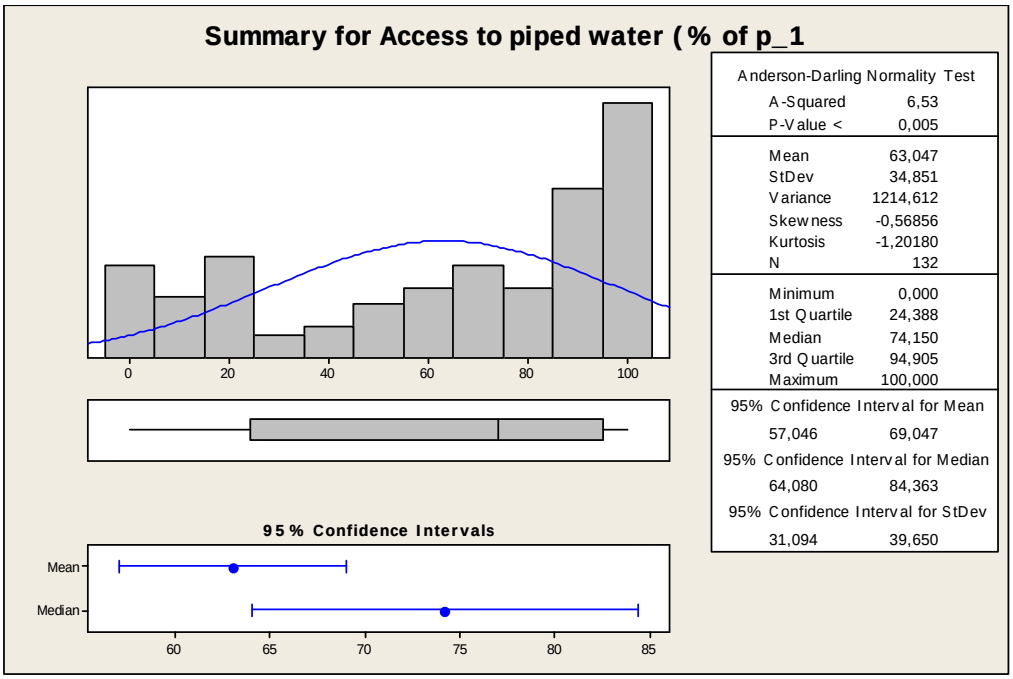


A ordem abaixo apresentada difere um pouco dos dados que estamos trabalhando, mas elucida didaticamente o tema ora estudado.



Fonte: Sítio do Social Progress Imperative. Disponível em: > http://www.socialprogressimperative.org/pt/data/spi/components/com3#map/countries/com3/

4.3 – Access to piped water (% of pop.)(água encanada)



Análise: Access to piped water (% of pop.)(água encanada) mostra quantas residências possuem agua encanada ou outro tipo de processo que leva a água tratada para o consumo.

A análise da pesquisa mostrada no gráfico com uma tendência quase normal porque existem dois extremos quase distintos. Do lado esquerdo do gráfico estão 3 estratos distintos, sendo o mais próximo da origem com 12 países dentro de um intervalo de -5 e 5%, do lado deste estrato está o estrato com 8 países entre 5 e 15% e colado está o estrato com 13 países entre 15 e 25.

A mediana é 74,150 com o terceiro quartil com o valor de 94,905 com 11 estratos distintos. Vale o registro do lado direito do gráfico com os estratos bem aparentes entre o 80 e o 100. Sendo que na faixa dos 80 temos 9 países entre 75 e 85; 22 países entre 85 e 95 e 33 países entre 95 e 105.

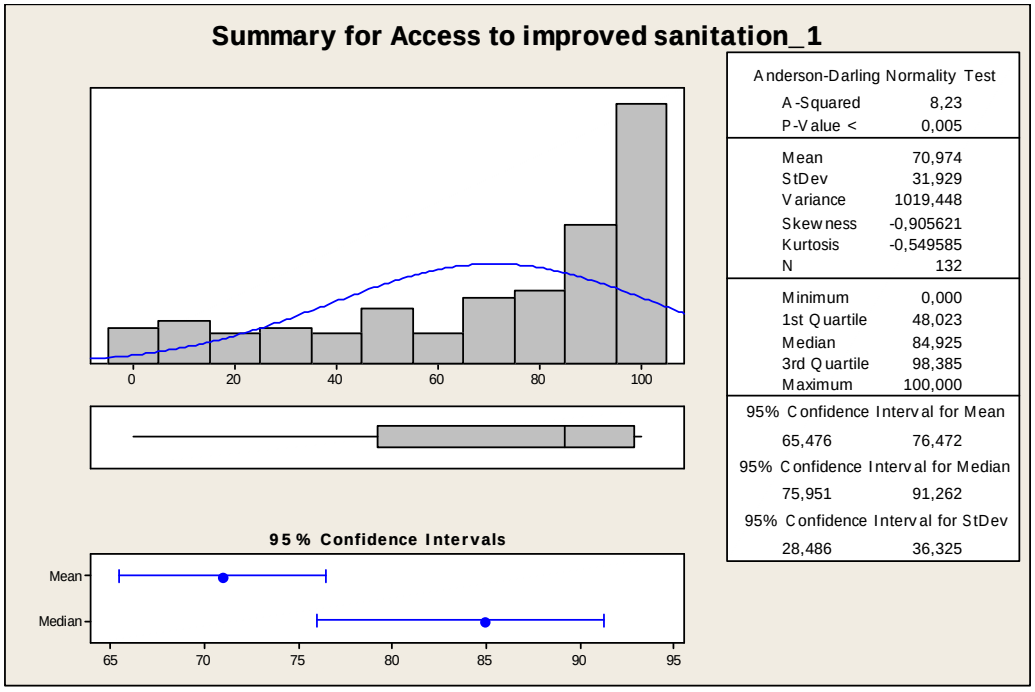
De acordo com os dados da UNICEF (2010) a evolução da utilização de diferentes tipos de fontes de água a partir de 1990-2010, por regiões do MDG - Millenium Development Goals (objetivos de desenvolvimento do milênio) apresentou dois grupos distintos de evolução. O primeiro é um conjunto de regiões em que a utilização de água canalizada para uma habitação, terreno ou quintal é baixa (30 por cento ou menos). Ele inclui a África subsaariana, Oceania, Sul da Ásia e do Sudeste Asiático. Embora os ganhos na utilização de água canalizada nas instalações foram feitas nessas regiões, o progresso foi principalmente na

qualidade da origem da água. Destacamos que 65% da população do sul da Ásia estão usando outras fontes naturais ao invés de água encanada no local.

O segundo conjunto de regiões é composto por Ásia oriental, África do Norte, Ásia Ocidental e na América Latina e no Caribe, onde pelo menos 70% da população está usando água encanada. Ásia Oriental (destaque para a China) teve um aumento significativo no abastecimento de água canalizada desde 1990, ganhando 35 pontos percentuais na cobertura nesta categoria em 20 anos. Isto representa 562 milhões de novos usuários que foram adicionados durante um período em que o mundo como um todo aumentou apenas 9%. Ásia Oriental é também a região com o aumento mais expressivo no uso de água potável de diversas fontes com qualidade, a partir de 68% em 1990 e passou para 91% de cobertura em 2010. Isso representa um aumento de 23 ponto percentual, muito maior que qualquer outra região.

http://www.unicef.org/media/files/JMPreport2012.pdf

4.4 - Access to improved sanitation facilities (% of pop)(acesso a instalações sanitárias melhoradas)



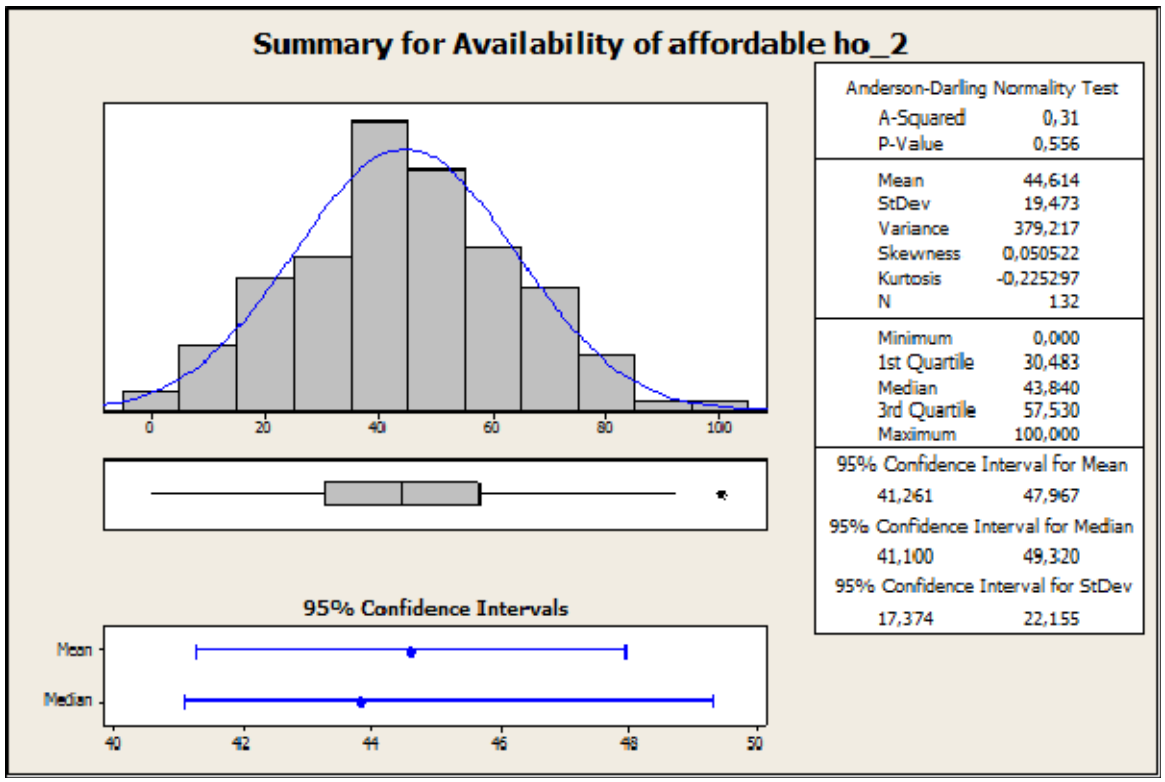
Análise: Access to improved sanitation facilities (% of pop)(acesso a instalações sanitárias melhoradas)

Os dados apresentados estão com a concentração do lado direito com uma média de 70,974 e mediana de 84,925, sendo que o terceiro quartil está com 98,385 indicando que esse estrato possui 43 países com melhores condições sanitárias.
Do lado oposto estão 6 países com indicadores -5 e 5 positivos. Do lado a este estrato estão 6 países com valores entre 5 e 15. Os países estão na tabela a seguir:

Tabela 2 – Países com valores entre -5 e 15 no item Acesso a Instalações Sanitárias Melhoradas

País	Índice
Benin	5,08
Burkina Faso	9,33
República do Chade	2,38
República do Congo	9,06
Ghana	4,29
Republica da Guine	9,84
Libéria	9,53
Madagascar	4,51
Mali	13,31
Moçambique	10,48
Republica do Niger	0,00
Tanazania	2,58
Republica do Togo	2,01

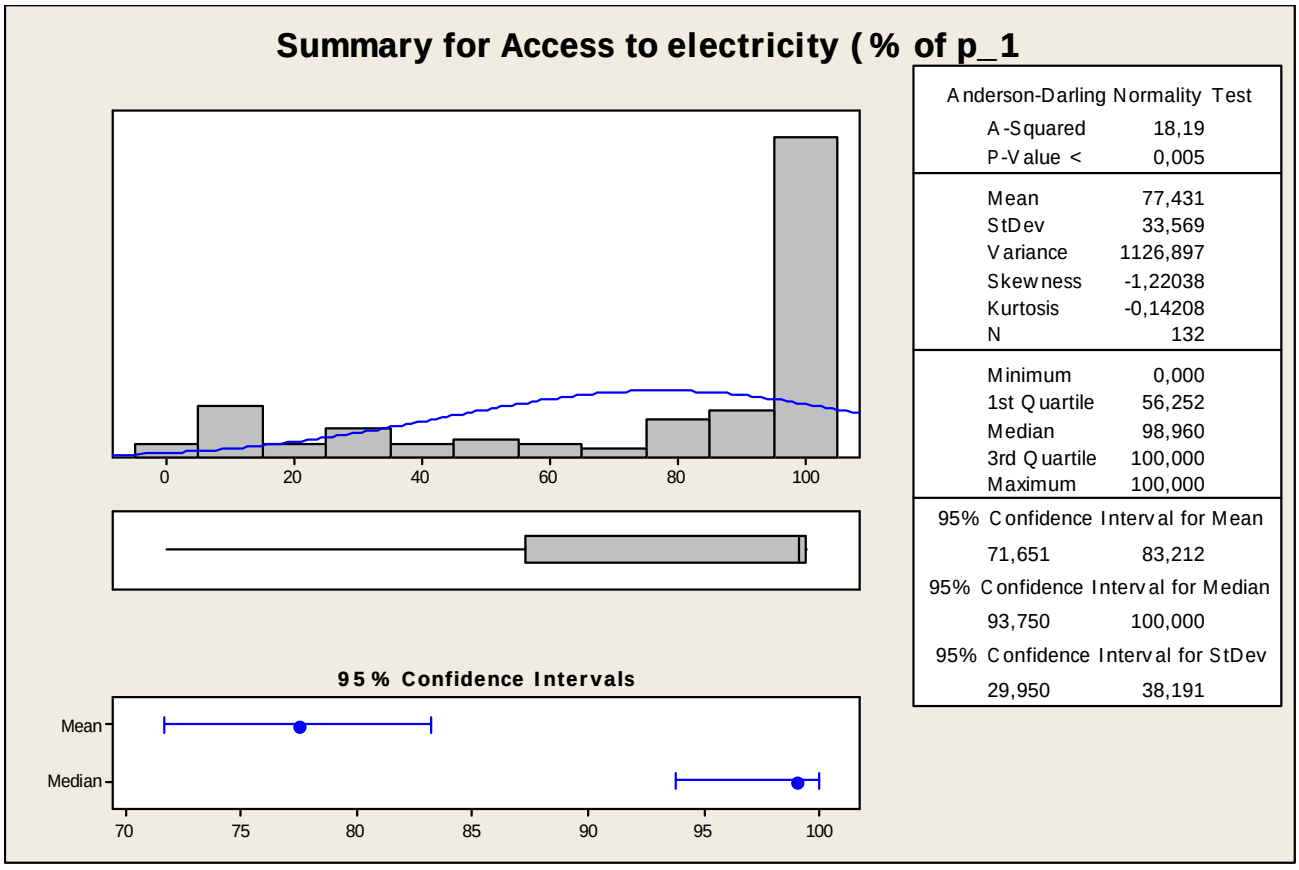
4.5 – Availability of affordable housing (% satisfied)



Análise: *Availability of affordable housing* diz respeito ao percentual da população de determinado país que indicam se estão satisfeitos ou insatisfeitos com a disponibilidade das habitações e se os preços estão acessíveis em suas regiões.

Neste caso podemos perceber que média e mediana estão bem próximas, estando a primeira estimada em 44,61 e a segunda ligeiramente abaixo em 43,84. Observa-se a maior concentração de países, totalizando 30, no intervalo entre 35 e 45, evidenciando que há ainda uma necessidade mundial em se ajustar e melhorar a sensação de satisfação da população e principalmente a praticar preços mais "justos" em suas regiões. Na cauda à esquerda da curtose estão países como Mongólia e Cuba com 12% e 14% de satisfação respectivamente. Na outra ponta temos Japão e Tailândia com 79% e 85% de satisfação respectivamente. O Brasil encontra-se justamente no intervalo com a maior concentração de países acima citada com 36% de satisfação (exatamente no ponto de máximo da curtose).

4.6 – Access to electricity (% of pop.)(eletricidade)



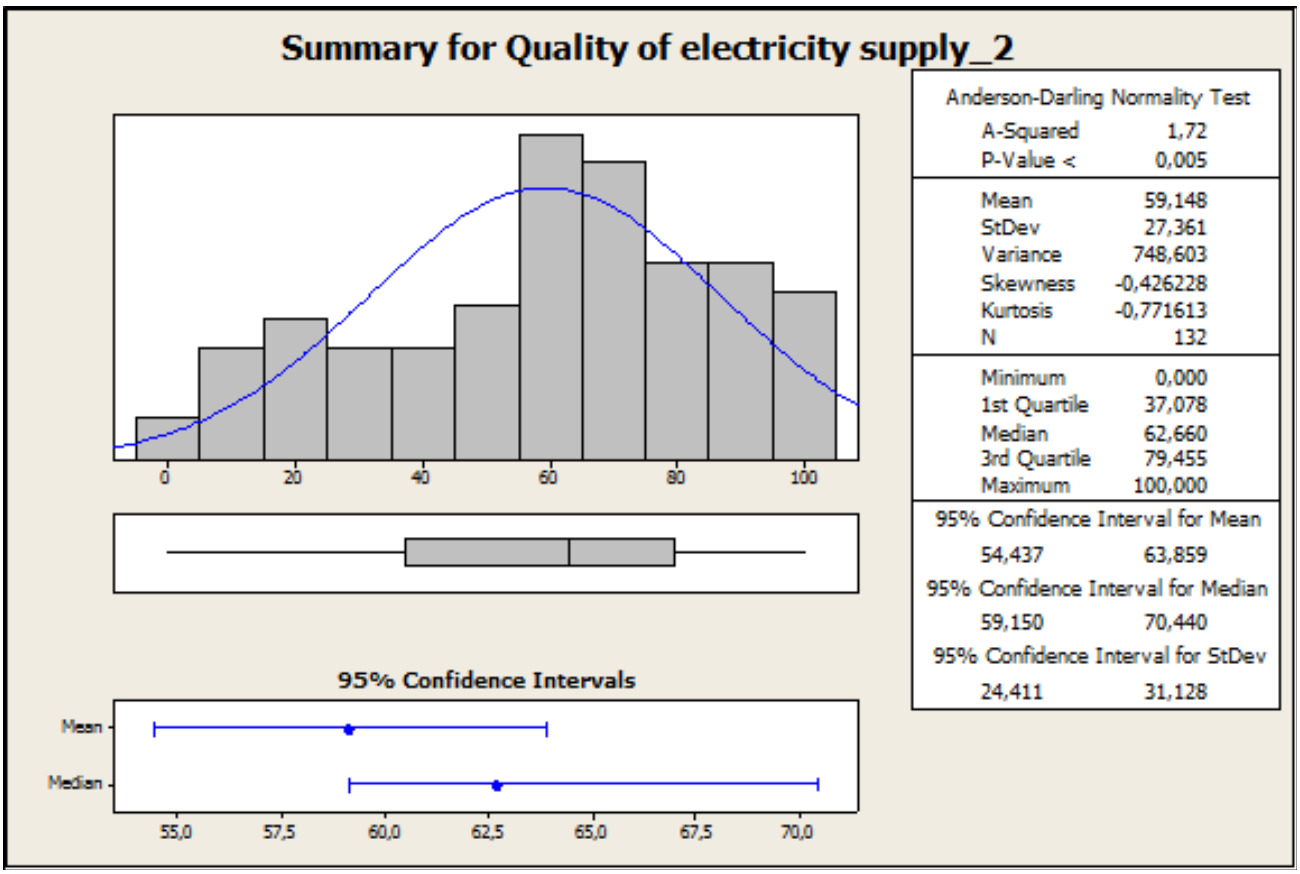
Análise:

Access to electricity (% of pop.)(eletricidade)

Este indicador mostra o percentual da população de cada país que tem acesso à energia elétrica. De uma maneira geral existe um grande número de pessoas com acesso à energia elétrica com o estrato entre 95 e 105 para 75 países, com média de 77,431 e mediana de 98,96.

Isto quer dizer que metade da população mais significativa tem energia elétrica e outra metade bem menor não tem. Isto aparece no estrato entre 5 e 15 com 12 países.

4.7 – Quality of electricity supply (1=low; 7=high)(qualidade fornecimento eletricidade)



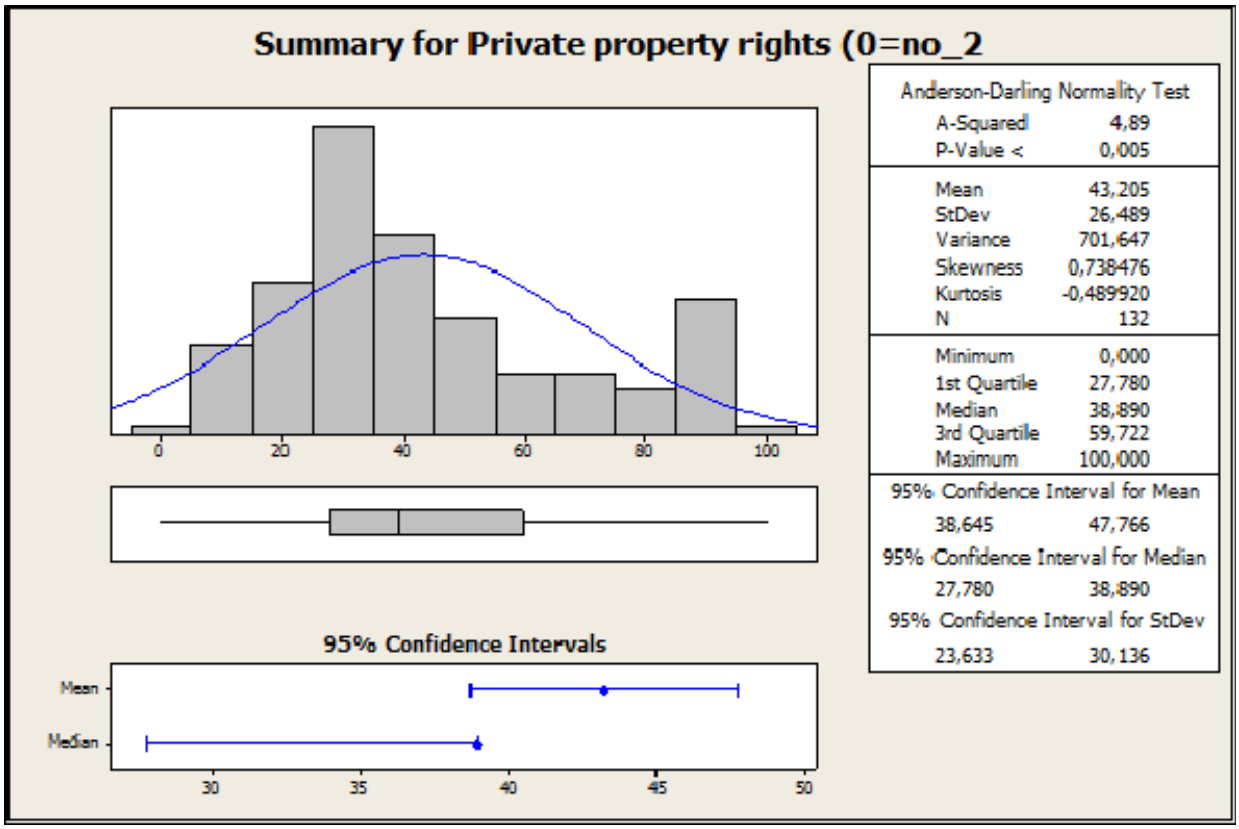
Análise:

Quality of electricity supply (1=low; 7=high)(qualidade fornecimento eletricidade)

Ter acesso a energia não significa que ela esteja disponível 24 horas por dia ou sempre que o consumidor necessitar.

Esse item procura identificar a qualidade do fornecimento da energia e o gráfico mostra uma distribuição normal entre os dados. A média está com o valor de 59,148 e mediana 62,66, sendo que o terceiro quartil com 79,455 representado pelo estrato com 21 países.

4.8 - Private property rights (0=none; 100=full)



Análise: *Private property rights* diz respeito ao rigor das leis de cada país para proteger e assegurar o direito à propriedade privada, seja ela física como um imóvel ou automóvel seja ela intelectual como fórmulas ou descobertas cientificas, por exemplo.

Nesta amostra temos claro que à esquerda estão países que não protegem com rigor este direito fundamental, encontrando o caso extremo da Venezuela isolada na primeira barra com índice 5, onde subentende-se que praticamente tudo é propriedade do Governo. Lembramos que a Coréia do Norte não está relacionada nesta lista de 132 países, caso estivesse certamente estaria no mesmo patamar da Venezuela ou abaixo.

Por outro lado, temos 15 países que se distanciam do miolo da mediana (que está no intervalo entre 27 e 38). Estes países oferecem maior segurança para seus proprietários individuais. Alguns destes países são: Alemanha, Canadá, Chile, Holanda e Reino Unido, todos com índice 90. Acima deles isolada com a maior pontuação está e Nova Zelândia com 95.

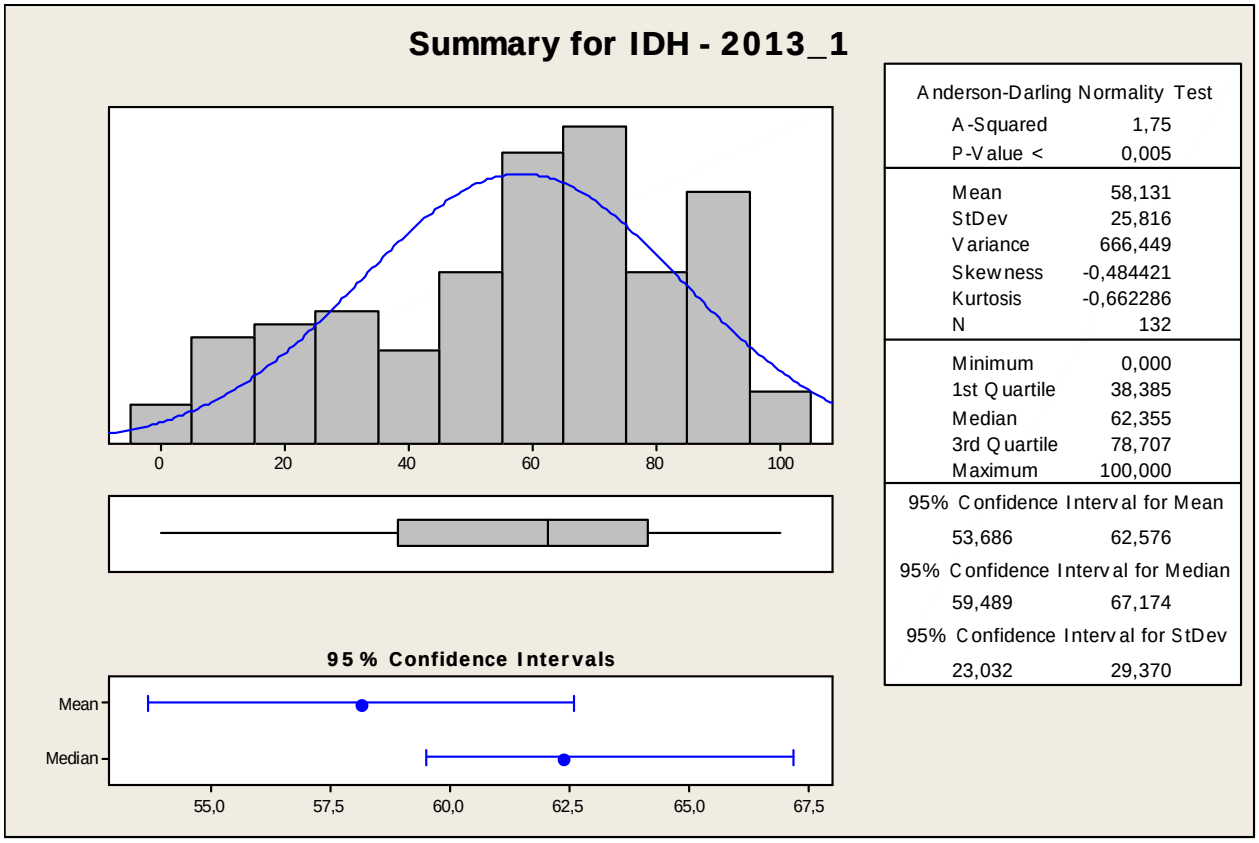
O Brasil encontra-se na barra da faixa entre 45 e 55, onde ocorre o ponto de máxima da curtose, nesta mesma faixa com índice 50 estão outros 12 países como: Colômbia, México, Itália, Costa Rica, Ghana e Índia.

4.9 IDH (2013) - Índice de Desenvolvimento Humano (IDH/PNUD)

O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.

O trabalho publicado preliminar publicado por PNUD - Programa de las Naciones Unidas para el Desarrollo sobre o desenvolvimento humano em 2014 aponta que o IDH é progresso real da população em relação a expectativa de vida, educação, saúde, habitação, segurança e condições para um desenvolvimento sustentável.

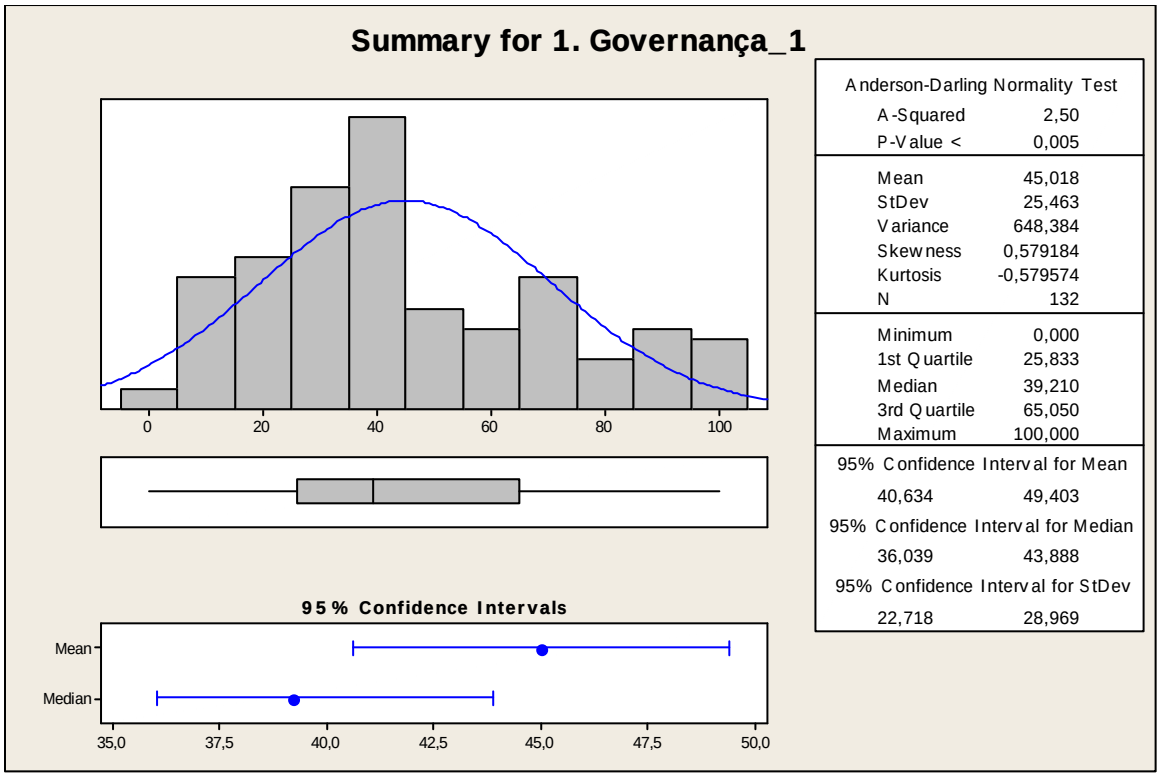
http://hdr.undp.org/sites/default/files/hdr14-summary-es.pdf



Anderson-Darling Normality Test	
A-Squared	1,75
P-Value <	0,005
Mean	58,131
StDev	25,816
Variance	666,449
Skewness	-0,484421
Kurtosis	-0,662286
N	132
Minimum	0,000
1st Quartile	38,385
Median	62,355
3rd Quartile	78,707
Maximum	100,000
95% Confidence Interval for Mean	
53,686	62,576
95% Confidence Interval for Median	
59,489	67,174
95% Confidence Interval for StDev	
23,032	29,370

O gráfico para o IDH aproxima-se de uma distribuição normal, mas deslocada para a direita, com uma concentração maior de países dos 65 aos 75 pontos, totalizando 25 países. Isso pode ser constatado também pela mediana, em 62,355. No estrato do 55 a 65 pontos, encontra-se o Paraguai, com IDH de 55,85 pontos. Do lado extremo esquerdo do gráfico estão 3 países no estrato de -5 e 5 pontos. Estes são considerados países de baixo desenvolvimento ou subdesenvolvidos.

4.10 Índice de Governança



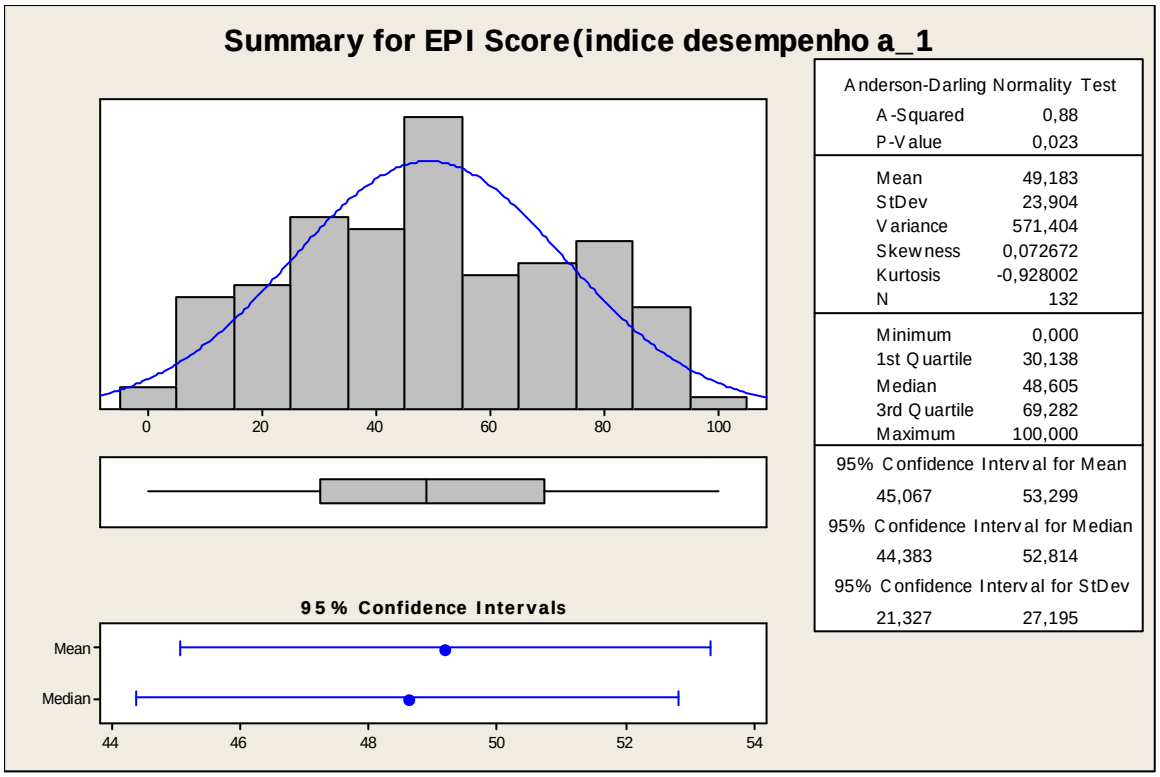
Análise:

Governança: World Economic Fórum ressalta que as estratégias de atuação: Simplificar os processos administrativos e fazer uso de alianças regionais e de tecnologia para coibir a corrupção e os abusos; aprimorar os instrumentos e mecanismos regionais que possibilitem uma ação coletiva capaz de prevenir violações da democracia e garantir a preservação da ordem democrática (http://www3.weforum.org/docs/LA11/WEF_LA11_Report_PT.pdf).

O gráfico para a variável GOV aproxima-se de uma distribuição normal deslocada para a esquerda, o primeiro quartil de 25,833 e a mediana de 39,21 confirmam esse deslocamento, indicando uma concentração crescente de países entre as faixas de 5 a 15 (13 países), de 15 a 25 (15 países), de 25 a 35 (22 países) até chegar ao pico – a faixa 35 a 45 (29 países). Essas faixas agrupam o equivalente a 60% dos países analisados.

Em relação ao pico, os países mais bem colocados são a Macedônia (44,74), a Turquia (44,69) e o Kuwait (44,24) e na faixa inferior ficam Argentina (35,30), Benin (35,75) e Indonesia (35,96). Aparentemente trata-se de grupo bastante heterogêneo, não sendo possível estabelecer relações imediatas.

4.11 EPI - Pontuação no Índice de Proteção Ambiental



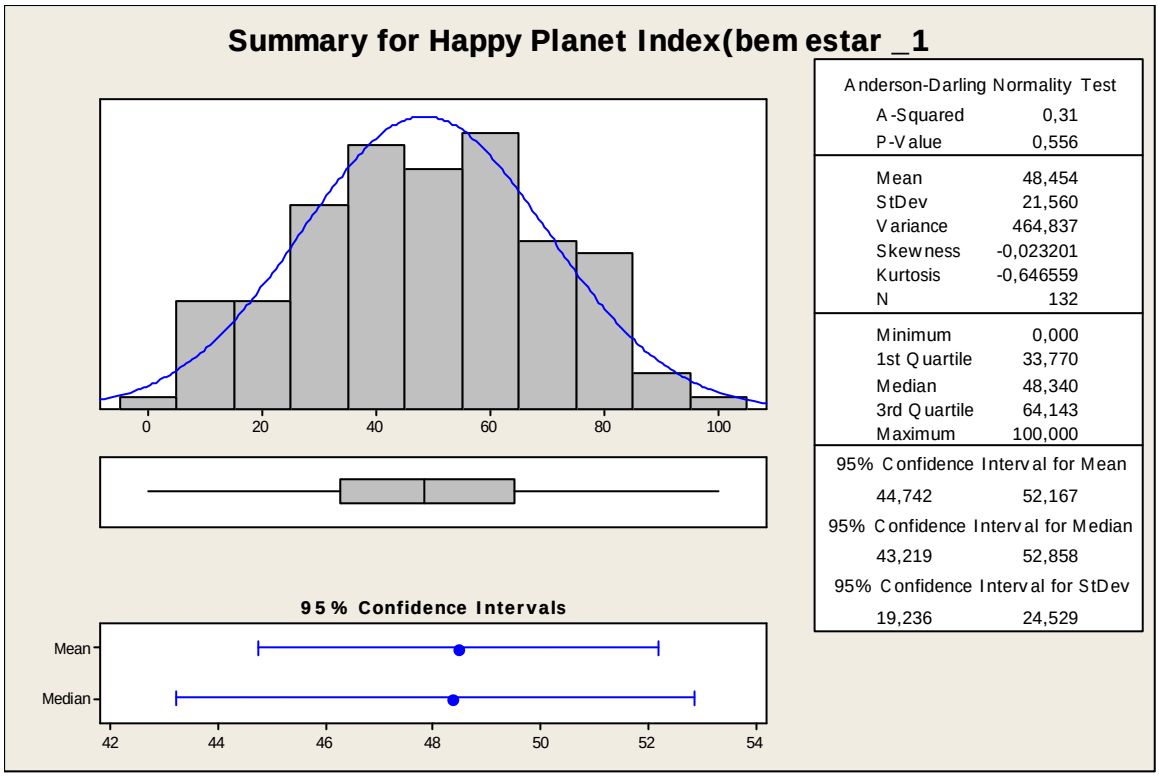
Análise:

A análise dos dados do Índice de Proteção Ambiental apresenta uma distribuição normal, com média de 49,18 e mediana de 48,60, ou seja, um ligeiro deslocamento para a esquerda. Há apenas 1 país no estrato superior (nota normalizada acima de 95), que é a Suiça, com EPI de 87,67 seguida por 9 países no segundo estrato (de 85 a 95), sendo todos países de alto IDH e IPS (Austrália, República Tcheca, Alemanha, etc) sendo o último país do estrato a Dinamarca com EPI de 76,92.

Já nos dois estratos inferiores temos a seguinte composição: no primeiro estrato (nota normalizada até 5) há dois países – Mali (EPI = 18,43) e Lesoto (EPI = 20,81) , seguidos por um grupo de 10 países no segundo estrato que vai de 5 a15. Destes 10 países, 9 encontram-se na África, sendo Bangladesh a única exceção do grupo. O país melhor colocado é Angola, com EPI de 39,93.

Já a faixa de pico, que vai de 45 a 55, é composta da seguinte maneira: na parte inferior, Algéria (EPI = 50,08) e Libano (EPI = 50,15) seguindo por um grupo heterogêneo de 26 países nos quais Jordânia (EPI =55,78) e Montenegro (EPI= 55,52). O Brasil está neste mesmo intervalo, abrigado pela média (45 a 53) com pontuação de 52,97, justamente onde a curtose começa a decrescer.

4.12 HPI - Pontuação no Happy Planet Index.



Análise:

A distribuição para o Índice de Felicidade Bruta (FIB) aproxima-se de uma normal, com maior concentração de países entre a marca do 35 até 65, que formam quase um plateau, com os três maiores picos da distribuição. Nas extremidades encontramos um pequeno número de países. Costa Rica ficou em primeiro lugar, com um FIB de 64,03, seguida por um estrato de notas normalizadas de 85 a 95, que é composto por três países: Colômbia (FIB=59,75), El Salvador (58,88) e Jamaica (58,53).

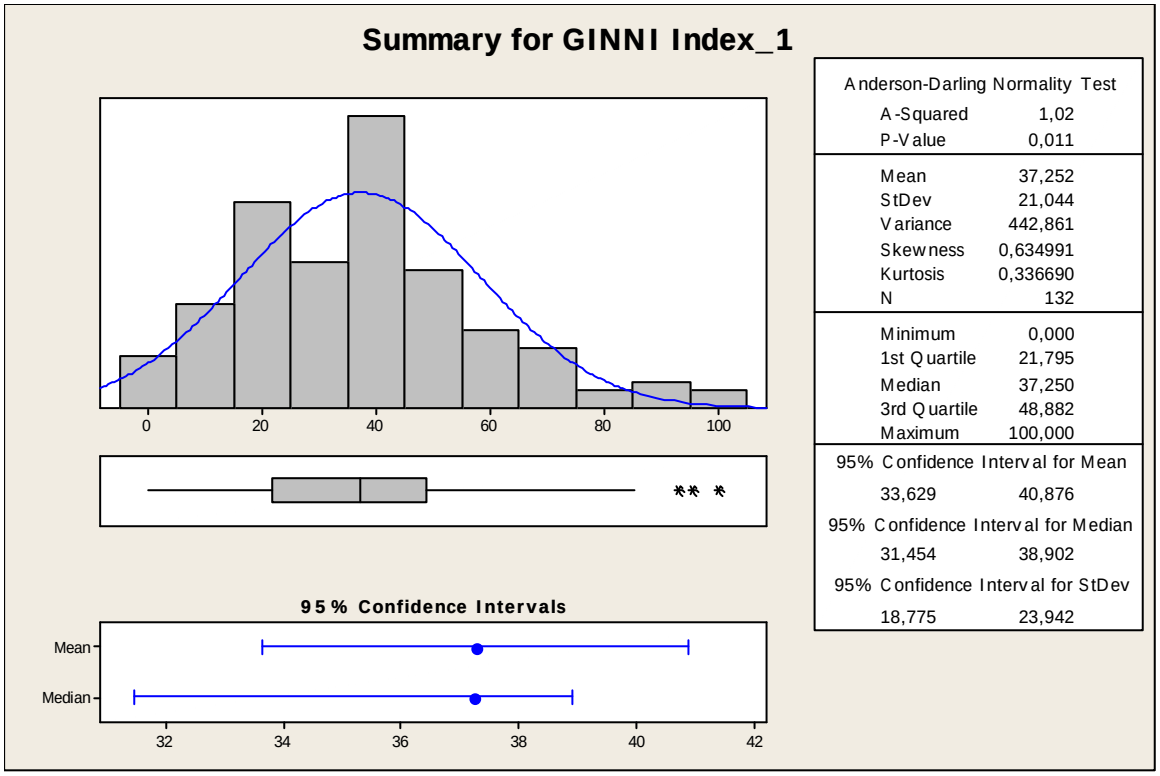
No estrato inferior, que vai até 5 pontos, figura também apenas um país: Botswana, com FIB = 22,59, seguida pelo segundo estrato, que vai de 5 a 15 pontos e compreende 9 países. A menor pontuação é do Chade (FIB=24,68) e a maior é da Macedônia (FIB=28,27). Chama a atenção neste grupo de países a presença da África do Sul, que figura com um FIB de 28,19, o oitavo pior colocado na classificação geral do FIB.

A maior concentração em número de países está alocada entre 35 e 65 pontos, sendo Djibouti (Lesta da África) com FIB=37,23, Estados Unidos (FIB=37,34) e Hungria (FIB=37,40) os piores colocados do grupo, e Sri Lanka (FIB=49,38), Iraque (FIB=49,19) e Laos (FIB=49,14) os melhores colocados do grupo.

O Índice de Felicidade Bruta mede aspectos subjetivos e objetivos da vida humana, evidenciando pesquisas que apontam que, a partir de certo nível de renda, o nível de felicidade médio reportado não aumenta na mesma proporção do aumento da renda, até decaindo com o tempo. Outro tópico relacionado é a questão da resiliência construída por pessoas em situação de privação em países de baixo IDH e IPS, que acabam adaptando-se às situações adversas.

Uma análise mais aprofundada poderá revelar o baixo nível de correlação deste indicador com o PIB e inclusive com o IDH e o IPS, pois sua metodologia é bastante distinta.

4.13 Índice de GINNI



Análise:

Para o Índice de Gini, que mede a desigualdade de renda, faltaram dados relativos a 13 países, que foram substituídos pela média da distribuição original, que foi de 39,02, que normalizado de 0 a 100 tornou-se 62,93, o que explica o pico registrado na coluna entre 55 e 65 – dos 33 países, na realidade apenas 20 tem sua pontuação original registrada neste intervalo.

Assim percebemos uma distribuição basicamente uniforme, com tendência para concentração de países mais a direita, como já demonstra o primeiro quartil, em 51,11.

São 6 países no estrato superior, acima de 95 pontos na escala normalizada, conforme a Tabela 13, abaixo, que como podemos ver é formada pelos antigos países socialistas

Tabela 13 – Países com menor índice de Gini

PAÍS	CÓD	REGIÃO	GINI
Ukraine	UKR	Eastern Europe	24,820
Slovenia	SVN	Southern Europe	24,870
Sweden	SWE	Northern Europe	26,080
Czech Republic	CZE	Eastern Europe	26,390
Belarus	BLR	Eastern Europe	26,460
Slovakia	SVK	Eastern Europe	26,580

Os países com maior desigualdade de renda estão divididos nos três primeiros estratos: de 5 a 15 (2 países – África do Sul e Namíbia), de 15 a 25 (3 países – Botswana, Zâmbia e Honduras) e de 25 a 35 (2 países – República Central Africana e Lesoto).

Tabela 13 – Países de maior índice de Gini (mais desiguais)

PAÍS	CÓD	REGIÃO	GINI
Lesotho	LSO	Southern Africa	54,170
Central African Republic	CAF	Eastern Africa	56,300
Honduras	HND	IBE	57,400
Zambia	ZMB	Middle Africa	57,490
Botswana	BWA	Southern Africa	60,460
Namibia	NAM	Southern Africa	61,320
South Africa	ZAF	Southern Africa	63,140

Botswana, Namíbia e África do Sul aparecem como *outliers* no gráfico.

5. Considerações Finais

Está sendo uma grande experiência a oportunidade de aprendizado com a utilização do Software Minitab, principalmente pela possibilidade de acesso ao universo de dados e estatísticas que englobam os principais países de nossa Casa Comum, o Planeta Terra.

Após termos feito a apresentação dos dados pela estatística descritiva como média, desvio padrão e variância; por exemplo, é possível constatar que os países de uma forma geral corroboram o que as análises macroeconômicas de diversos organismos tais como: UNICEF, Banco Mundial, WEF – World Economica Forum, FAO e outros órgãos preocupados em identificar a evolução dos povos do globo nos apresentam.

Indicadores como o IDH – Índice de Desenvolvimento Humano apresentado neste trabalho com a normalização não muda a situação real da economia analisada. Por exemplo: quando se pegam os dados de países da Africa Sub-saariana, Republica centro-africana e Lesotho apresentam indicadores que mostram que as variáveis que fazem parte deste indicador como saúde, educação, uso de água potável, instalações sanitária, que podem refletir a qualidade de vida, estão baixos quando comparados com outros países.

Especificamente a Zambia difere dos demais principalmente quando analisamos seus dados de IDH de 1998 para 2013. Esse indicador salta de 0,41 em 1998 para 0,56 em 2013, ou seja, o indicador nos diz que quanto mais próximo de 1 melhor a qualidade de vida da população.

http://pt.knoema.com/atlas/Z%C3%A2mbia/%C3%8Dndice-de-Desenvolvimento-Humano

Para melhor entendimento dos dados estatísticos sugerimos que o pesquisador examine a definição do conceito na estatística para aplicar ao fato analisado. Dessa maneira, temos que a mediana que divide a metade inferior da metade superior das amostras. Como temos 132 países, a mediana é calculada pela média dos dois centrais depois de coladas em ordem crescente os valores. Por exemplo: se pegarmos a variável IDH (poderia ser qualquer uma) e colocarmos em ordem crescente os valores dos 132 países, a mediana estará entre os números da posição 66 e 67 na linha do excel. Logo, se somarmos os valores e dividirmos por 2 temos a mediana.

Especificamente para o tema **Moradia,** efetuamos o estudo das 13 variáveis apresentadas nos itens 4.1 a 4.13, explorando as principais informações obtidas graficamente como: médias, medianas, curtose e outros dados obtidos por meio de uma estatística de Anderson-Darling.

CAPITULO 2 - MODELO DE REGRESSÃO MULTIPLA LINEAR

1 RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS

1.1 Correlação linear

A matriz de correlação linear apresenta o teste de significância *P-Value*. Para a correlação foi utilizado o índice de *Pearson*. Conforme Las Casas e Guevara (2010)[2], a correlação (valores entre -1 e 1) é expressada pelo grau de relação linear que existe entre elas. Valores próximos a zero significam nenhuma relação, e próximos a 1 e -1 mostram alta relação positiva ou negativa, respectivamente. Vale ressaltar que o índice de correlação entre as variáveis não requer, necessariamente, que exista uma relação de causa-efeito entre as variáveis para dar sentido às relações. A seguir, apresentamos a matriz de correlação das variáveis estudadas:

Correlations: Social Progr; Shelter_1; Access to pi; Access to im; ...

	Social Progress	Shelter_1	Access to piped
Shelter_1	**0,894** 0,000		
Access to piped	**0,846** 0,000	**0,878** 0,000	
Access to improv	0,791 0,000	**0,869** 0,000	**0,856** 0,000
Availability of	0,307 0,000	0,308 0,000	0,109 0,213
Access to electr	0,769 0,000	**0,887** 0,000	**0,846** 0,000
Quality of elect	0,785 0,000	0,818 0,000	0,680 0,000
Private property	0,764 0,000	0,617 0,000	0,529 0,000
IDH - 2013_1	**0,868** 0,000	**0,867** 0,000	0,826 0,000
1. Governança_1	**0,890** 0,000	0,696 0,000	0,633 0,000
EPI Score_1	**0,917**	0,837	0,813

[2] LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisas de Marketing. São Paulo: Atlas, 2010.

	0,000	0,000	0,000
Happy Planet Ind	0,296	0,399	0,370
	0,001	0,000	0,000
GINNI Index_1	-0,308	-0,301	-0,247
	0,000	0,000	0,004

	Access to improv	Availability of	Access to electr
Availability of	0,163		
	0,063		
Access to electr	**0,878**	0,169	
	0,000	0,053	
Quality of elect	0,656	0,214	0,608
	0,000	0,014	0,000
Private property	0,466	0,270	0,392
	0,000	0,002	0,000
IDH - 2013_1	0,834	0,182	0,823
	0,000	0,037	0,000
1. Governança_1	0,547	0,290	0,480
	0,000	0,001	0,000
EPI Score_1	0,782	0,175	0,724
	0,000	0,045	0,000
Happy Planet Ind	0,373	0,251	0,476
	0,000	0,004	0,000
GINNI Index_1	-0,346	0,040	-0,353
	0,000	0,647	0,000

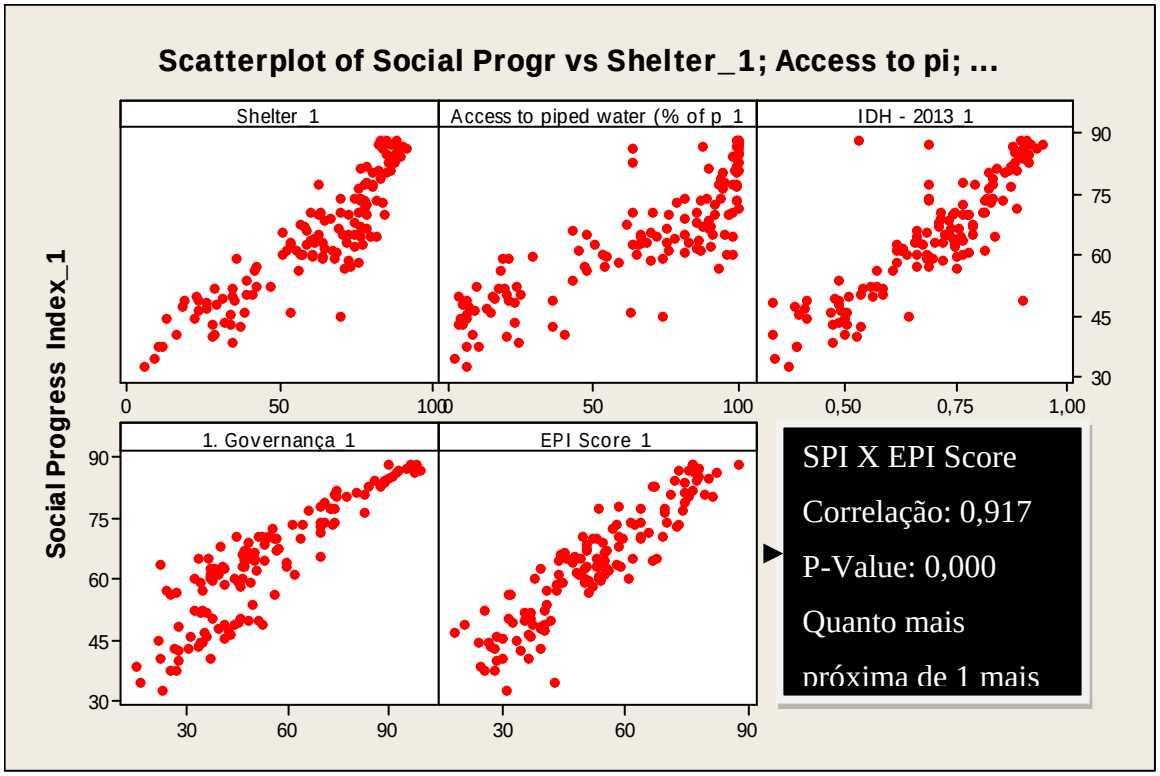
	Quality of elect	Private property	IDH - 2013_1
Private property	0,661		
	0,000		
IDH - 2013_1	0,700	0,602	
	0,000	0,000	
1. Governança_1	0,727	**0,914**	0,697
	0,000	0,000	0,000
EPI Score_1	0,769	0,712	0,824
	0,000	0,000	0,000
Happy Planet Ind	0,179	0,067	0,333
	0,040	0,444	0,000
GINNI Index_1	-0,322	-0,262	-0,297
	0,000	0,002	0,001

	1. Governança_1	EPI Score_1	Happy Planet Ind
EPI Score_1	0,811		
	0,000		
Happy Planet Ind	0,089	0,197	
	0,312	0,024	
GINNI Index_1	-0,247	-0,365	0,051
	0,004	0,000	0,562

Cell Contents: Pearson correlation
P-Value

1.2 Gráficos de dispersão

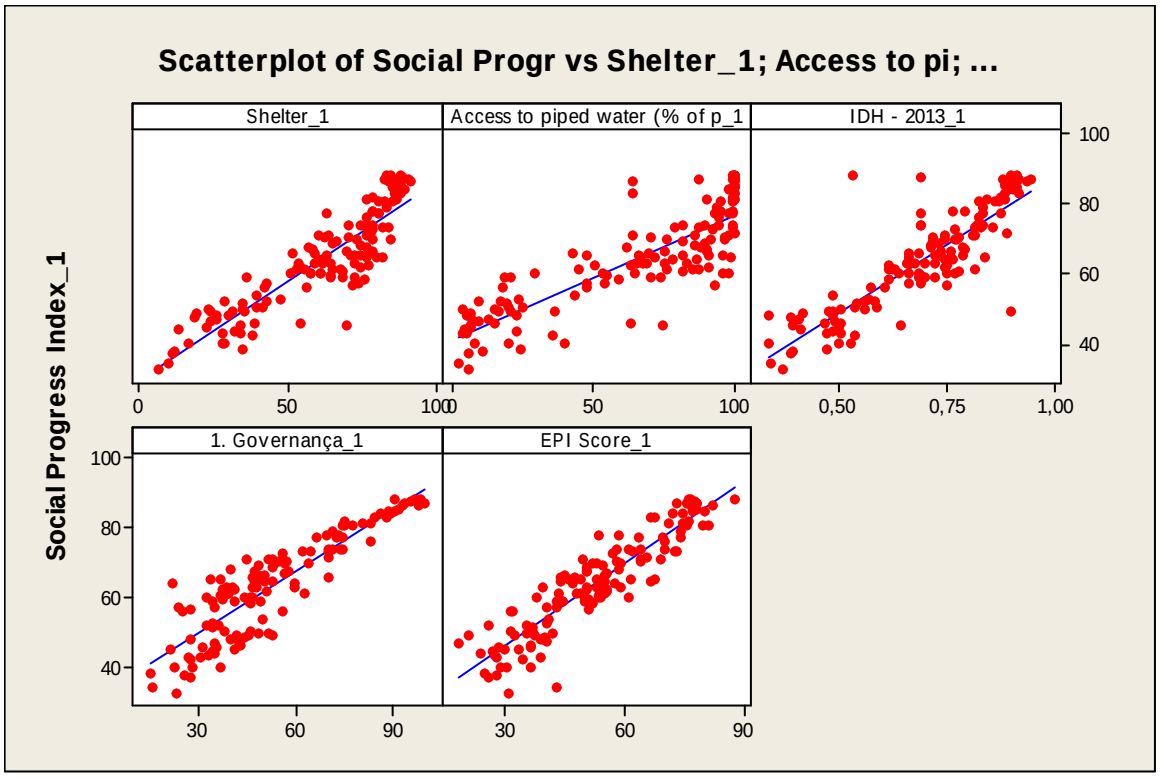
Para efeito de análise foi realizada a relação entre as variáveis com maior grau de correlação. Para não tornarmos a análise desnecessariamente extensa, avaliamos que o grau aceito como de grande correlação seriam todos os dados com banda entre 0,84 até 1 ou -0,84 até -1. Lembramos também que todas as variáveis consideradas, possuem P-Value = 0,000 ou seja, apresentando alto grau de confiabilidade. Alertamos que estamos cruzando unicamente as variáveis presentes no tema **MORADIA.**



Podemos perceber que ao correlacionarmos a variável **SPI – Social Progress Index** com as variáveis Shelter, Access to piped water, IDH-2013, Governança e EPI Score todos os gráficos apresentaram correlação linear positiva. Sendo que a correlação SPI X EPI Score demonstra-se visulamente mais agrupada (com correlação 0,917) e a correlação SPI X Access to piped water encontra-se mais desagrupada (com correlação 0,846) revelando-nos gráficamente que quanto mais próximas de 1 estiverem as comparações, mais as amostras estarão próximas e quanto mais as amostras estiverem se aproximando de 0 mais elas irão se afastando.

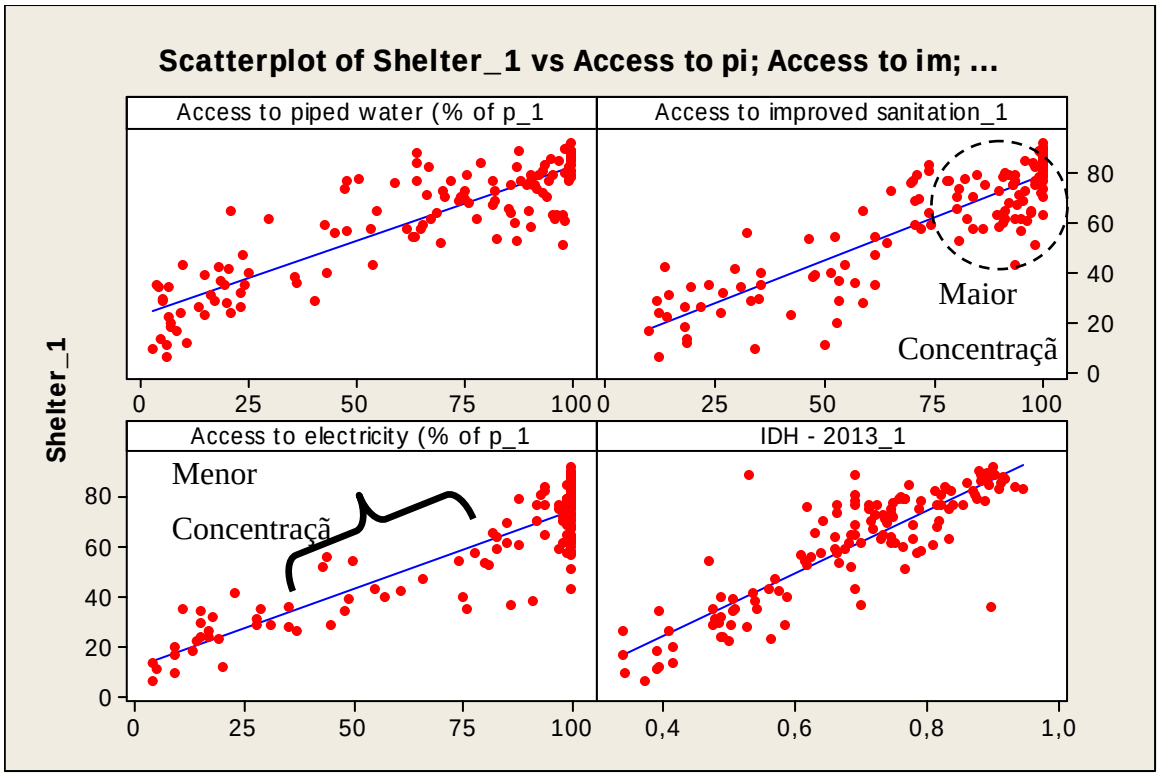
1.3 Linhas de tendência

1.3.1 Scatterplot de Social Progress Index



Agora temos a mesma representação gráfica do Item 1.2, porém adicionamos as linhas de tendência. Neste caso podemos verificar ainda mais claramente para onde as amostras estão se afastando ou se concentrando. Esta ferramenta nos auxilia a compreender o tipo de movimento que nossas representações gráficas estão formando, levando-se em consideração os pontos máximos e mínimos relativos, constatamos se há algum padrão para seguirmos adiante em nossa análise. Quanto mais próximas de um padrão (agrupamento) mais as variáveis estarão correlacionadas.

1.3.2 Scatterplot de Shelter

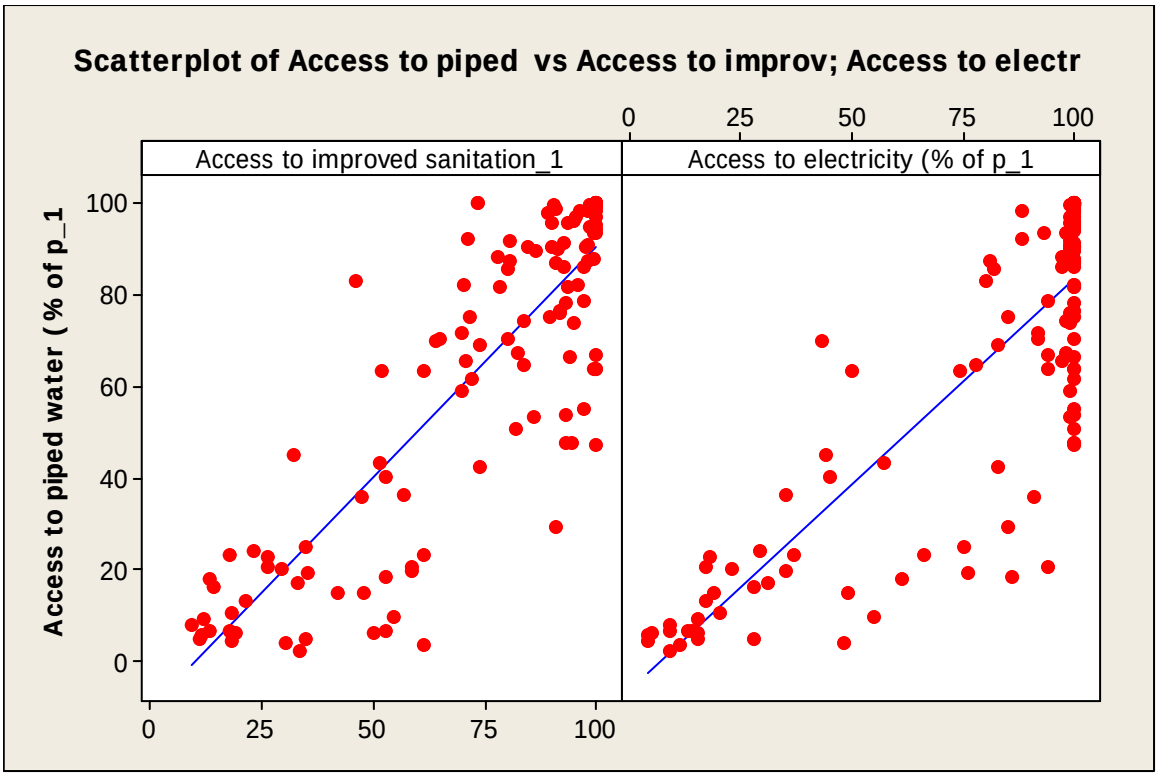


Passaremos a seguir, a analisar as correlações calculadas pelo MINITAB com valores acima de 0,84 para a variável **Shelter.** Neste caso obtivemos grau de comparação com Access to piped water, access to improved sanitation, access to eletricity e IDH-2013. Vemos que não há nenhuma representação gráfica linear negativa (linha de procura/demanda), ou seja todas as representações gráficas seguem uma tendência de resistência (linha de oferta).

Comparando estas representações gráficas com as do item 1.3.1 percebemos que agora há um menor condensamento de amostras ao longo da linha, ou seja, a correlação entre Shelter X Access to eletricity e Shelter X Access to improved sanitation, por exemplo, nos apontam que as amostras estão ocupando com maior densidade o extremo superior do gráfico.

Mesmo assim, as variáveis Shelter X Access to eletricity são as que possuem maior correlação com índice 0,887, pois nela há um menor afastamento das amostras em relação à linha de tendência, seguidas de Shelter X Access to improved sanitation com índice 0,869, Shelter X Access to piped water índice 0,878 e por fim Shelter X IDH-2013 com índice 0,867.

1.3.3 Scatterplot de Access to piped water



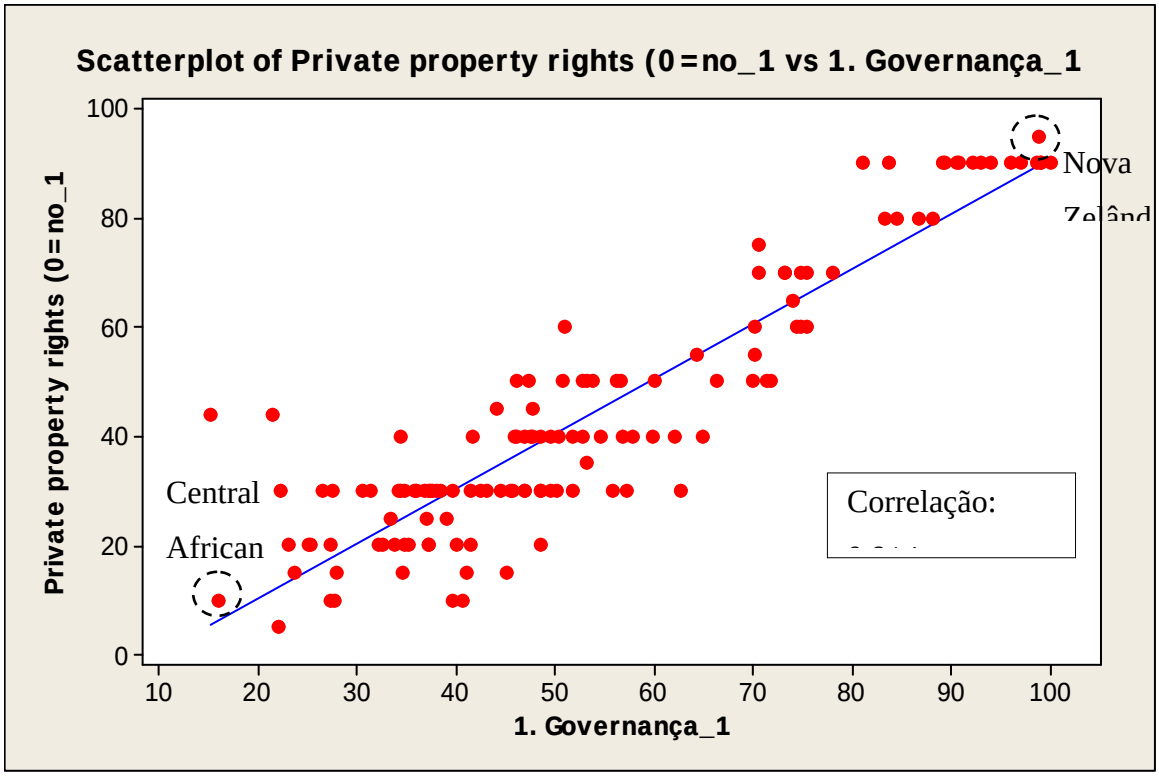
No Scatterplot acima temos as variáveis que apresentaram correlação superior a 0,84 com Access to piped water (Acesso à água potável). Estas variáveis são Access to improved sanitation (Accesso a instalação sanitária melhorada) com correlação 0,856 e Access to eletricity (Acesso a eletricidade) com correlação 0,0846.

5.3.4 Demais Scatterplots

O scatterplot entre Private Property Rights e Governança foi o que nos apresentou a maior correlação com índice de 0,914. Esta informação nos leva a crer que há forte relação entre estas variáveis.

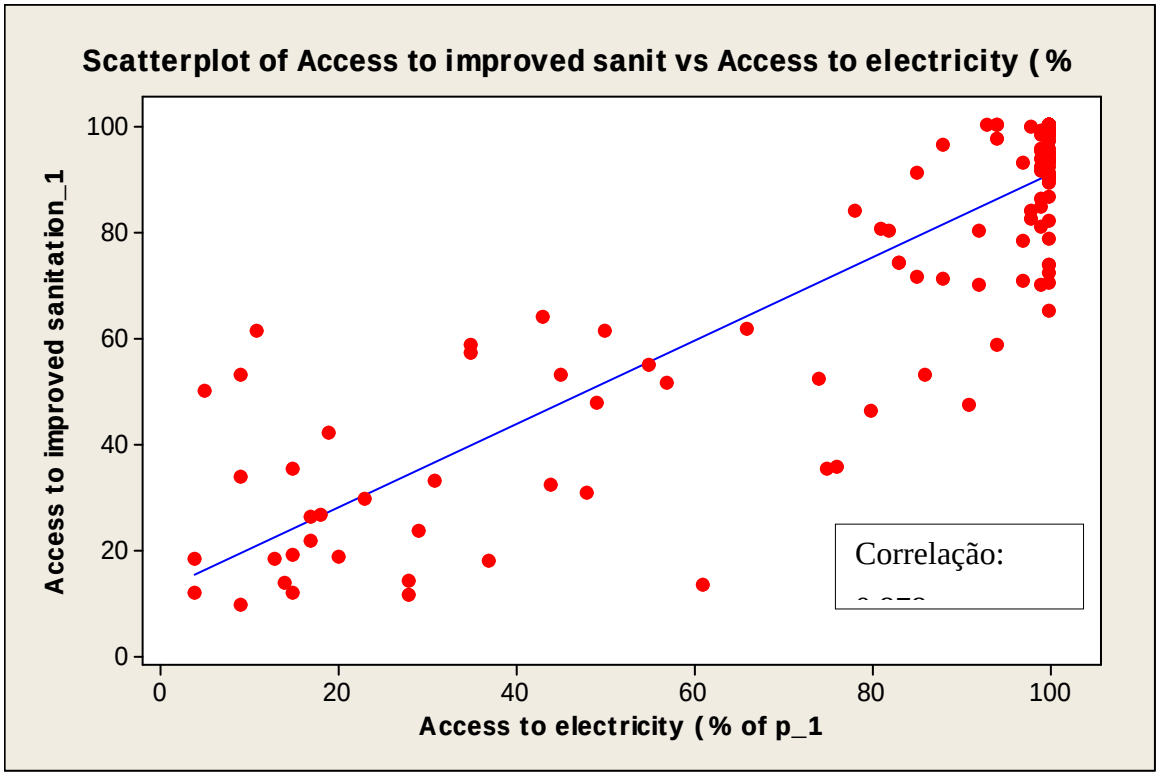
Interpretamos esta representação gráfica afirmando que quanto maior o grau de governança de determinado País, maior será o rigor e a aplicabilidade de leis para Proteção aos Direitos de Propriedade Privada.

Temos portanto, isolada na parte mais elevada do gráfico a Nova Zelândia que possui índice 98,908 de Governança e 95,000 de Private Property Rights. Por outro lado isolada no extremo inferior do gráfico temos a República Central da África com índice 15,987 de Governança e 10,000 de Private Property Rights, consubstanciando a metodologia aplicada.



Por fim, no gráfico abaixo temos o scatterplot entre Access to Improved Sanitation e Access to Eletricity. Temos nestas variáveis correlação de 0,878, nos possibilitando avaliar que existe relação entre a qualidade de sistemas avançados de instalações sanitárias e ao acesso à eletricidade, convém ressaltar que esta relação não se explica unicamente por estas variáveis, mas existe de fato relevância e necessidade de condições adequadas nos países de aterramento de córregos e eficientes sistemas de políticas sanitárias para maximizar a chegada e acesso da infraestrutura de instalações elétricas.

Dos 132 países analisados apenas 20 possuem índice 100 em ambas variáveis, nenhum deles está na América Latina. Dentre os países Íbero-Americanos apenas portugal possuí índice 100 em Access to improved sanitation e 100 em Access to eletricity.



1.4 Dendograma ou Dendrograma

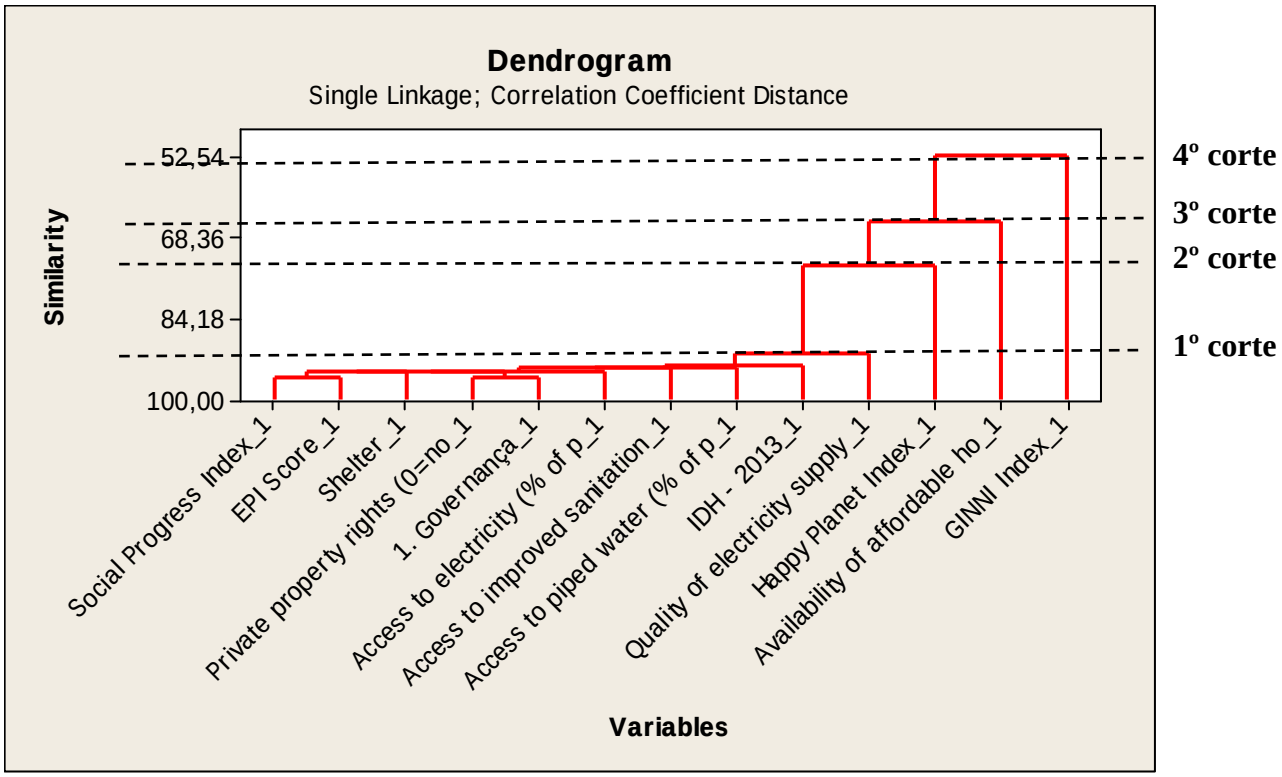
Um dendograma ilustra as informações na tabela de amalgamação na forma de um diagrama de árvore. Dendro = Árvore. Sua representação gráfica se dá por meio de um tipo específico de diagrama que organiza determinados fatores e variáveis.

O Dendograma é o resultado de uma análise estatística de uma série de dados, em que se emprega a técnica de um método quantitativo que leva a agrupamentos e a uma ordenação hierárquica ascendente, onde os termos gráficos se assemelham aos ramos de uma árvore que vão se conectando a outros sucessivamente.

Por padrão, o nível de similaridade é medido no eixo vertical (também é possível exibir o nível de distância) e as diferentes observações são listadas ao longo do eixo horizontal. O gráfico mostra como os agrupamentos são formados: unindo duas observações individuais ou pareando uma observação individual com um agrupamento existente. É possível ver em que nível de similaridade os agrupamentos são formados, e a composição dos agrupamentos da partição final.

Para alguns conjuntos de dados, os métodos média, centróide, mediana e Ward podem não gerar um dendograma hierárquico. Ou seja, as distâncias de amalgamação nem sempre

aumentam com cada etapa. No dendograma, essa etapa geraria uma junção para baixo no lugar de para cima.



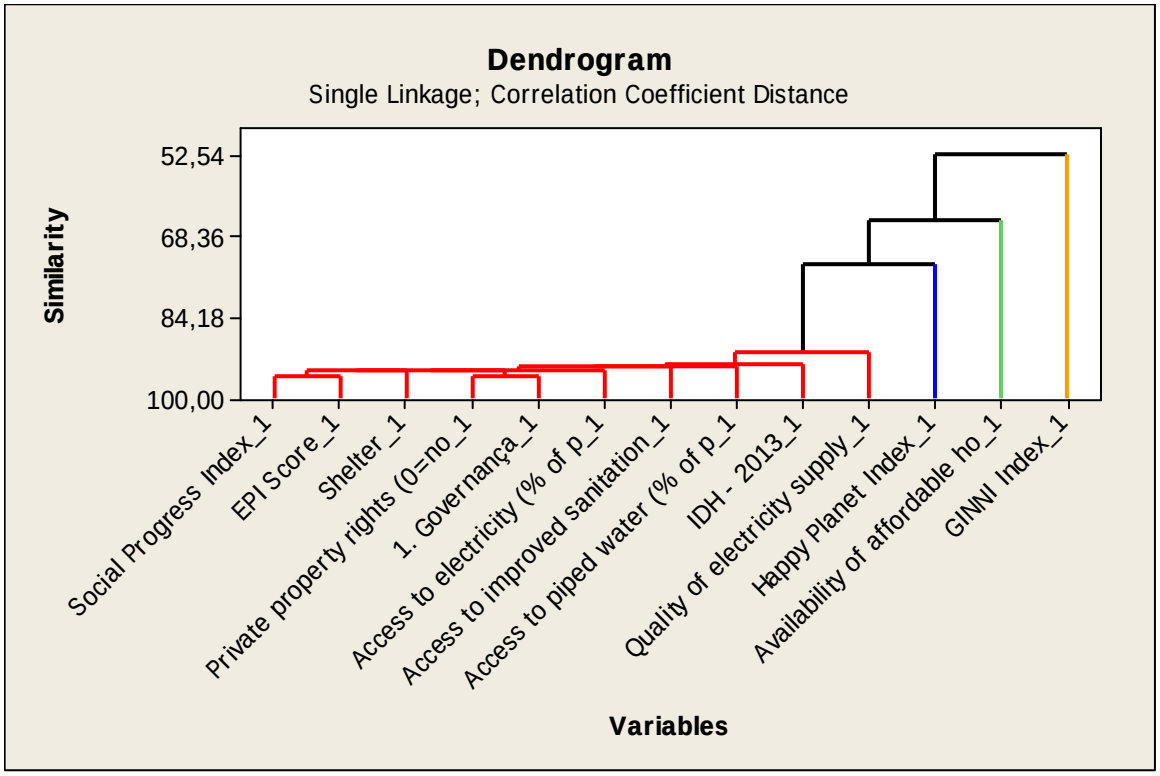
O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação. Assim, verifica-se que as variáves Social Progress até Quality of eletricity supply apresentam alto grau de similaridade. Deste modo podemos inferir que se fizéssemos um corte neste trecho teríamos um primeiro agrupamento. Pertenceriam a este grupo as variáveis: (SPI, EPI Score, Shelter, PPR, Governança, Access to eletricity, Acces to improved sanitation, Access to piped water, IDH e Quality of eletricity supply).

Seguindo este sistema subjetivo de dedução, poderíamos visualmente acrescentar um segundo corte a partir do trecho em que ocorre um afastamento desproporcional de similaridade em relação à variável anterior. Este segundo corte compreenderia todas as variáveis do primeiro corte mais Happy Planet Index.

Analogamente ao passo anterior, poderíamos acrescentar um terceiro corte que compreenderia todas as variáveis do primeiro e segundo corte mais a variável Availability of affordable house.

Finalmente, podemos acrescentar um quarto corte, onde compreenderia todas as variáveis anteriores mais Ginni Index.

Este exercício resulta graficamente no dendograma abaixo, onde o próprio software Minitab calculou os clusters e nos indicou as distâncias de similaridade dos coeficientes de correlação. Estes cortes ficam destacados pela alteração nas cores das barras verticais.



2. Considerações finais

Separando as variáveis em *clusters*, como visualiza-se no dendograma acima, é possível classificar as variáveis em 4 grupos de acordo com as suas similaridades.

Lembramos que enquanto um coeficiente de correlação é uma medida de similaridade, a distância euclidiana é uma medida de dissimilaridade.

A distância euclidiana cresce à medida que cresce o número de variáveis e uma das formas que se apresenta graficamente é pelo Método de Agrupamento Hierárquico, conforme utilizado no presente dendograma, deste modo os indivíduos são colocados nos grupos em diferentes etapas, de modo hierárquico, de acordo com seu maior grau de correlação com a variável anterior, tendo como produto uma Árvore de Classificação apontando um encadeamento crescente entre estas distâncias.

CAPITULO 3 – REGRESSÃO LINEAR

1. REGRESSÃO

No capítulo I e II do presente trabalho alguns países tiveram que ter variáveis preenchidas com a média geral para podermos dar continuidade nas análises, devido à falta de informação na base de dados. A partir do momento em que foi possível analisar a dispersão das variáveis e gerar o dendograma procedeu-se em seguida com a regressão, e agora é possível verificar a similaridade entre as variáveis para obter as médias mais próximas da realidade.

Utilizando ainda o programa MINITAB foi feita a Análise de Regressão, onde o critério foi isolar a Variável *Shelter*, ou seja, escolher a principal variável que melhor explica o grupo em que se estabeleceu no 1º corte do Dendograma (vide Item 5.4). Descartando, portanto desta análise as variáveis do corte 2, 3 e 4 por não apresentarem grau de explicabilidade.

Obtivemos os resultados descriminados abaixo:

Regression Analysis: Shelter_1 versus Social Progr; Access to pi; ...

```
The regression equation is
Shelter_1 = - 18,2 + 0,675 Social Progress Index_1
            + 0,0619 Access to piped water (% of p_1
            + 0,0520 Access to improved sanitation_1
            + 0,239 Access to electricity (% of p_1
            + 4,40 Quality of electricity supply_1
            + 0,135 Private property rights (0=no_1 + 4,68 IDH - 2013_1
            - 0,243 1. Governança_1 - 0,125 EPI Score_1
```

Predictor	Coef	SE Coef	T	P
Constant	-18,179	4,871	-3,73	0,000
Social Progress Index_1	0,6753	0,1858	3,63	0,000
Access to piped water (% of p_1	0,06189	0,03621	1,71	0,090
Access to improved sanitation_1	0,05202	0,04407	1,18	0,240
Access to electricity (% of p_1	0,23883	0,04190	5,70	0,000
Quality of electricity supply_1	4,4022	0,5662	7,77	0,000
Private property rights (0=no_1	0,13535	0,05570	2,43	0,017
IDH - 2013_1	4,684	7,698	0,61	0,544
1. Governança_1	-0,2432	0,1081	-2,25	0,026
EPI Score_1	-0,12455	0,08142	-1,53	0,129

Desconsiderare mos as variáveis com p.value

S = 5,78845 R-Sq = 93,9% R-Sq(adj) = 93,4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	9	62709,7	6967,7	207,95	0,000
Residual Error	122	4087,7	33,5		
Total	131	66797,4			

Source	DF	Seq SS
Social Progress Index_1	1	53409,7
Access to piped water (% of p_1	1	3466,1
Access to improved sanitation_1	1	1759,4
Access to electricity (% of p_1	1	1706,5
Quality of electricity supply_1	1	2072,6
Private property rights (0=no_1	1	38,8
IDH - 2013_1	1	8,3
1. Governança_1	1	169,9
EPI Score_1	1	78,4

Unusual Observations

Obs	Social Progress Index_1	Shelter_1	Fit	SE Fit	Residual	St Resid
23	34,2	9,610	22,552	2,476	-12,942	-2,47R
66	57,1	42,740	56,670	1,541	-13,930	-2,50R
82	59,0	36,390	51,517	1,900	-15,127	-2,77R
99	77,4	62,930	77,986	0,907	-15,056	-2,63R
112	38,5	34,770	38,696	2,784	-3,926	-0,77 X
113	48,9	35,680	41,126	3,010	-5,446	-1,10 X
115	88,2	88,480	86,516	3,111	1,964	0,40 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Obs. referência da linha de Países que possuem diferença e que deverão ser ajustadas pelo novo índice apontado na coluna **FIT**

Após rodarmos a primeira Regressão (acima visualizada) procederemos com a checagem das variáveis que apontaram P-Value superior a 10% ou 0,10 e ajustaremos os índices das linhas apontadas na coluna “Obs”, substituindo na Coluna “*shelter”* os novos índices informados na coluna “*Fit”*. Após esta etapa rodaremos novamente a Variável *shelter*

como Preditora, retirando da análise as variáveis: *access to improved sanitation*, IDH-2013 e *EPI-Score*, por apresentarem P-Value superior a 0,10.

Importante ressaltar que nesta etapa ao aplicarmos os ajustes apontados na coluna "*Fit*" os índices da variável "*shelter*" foram ajustados após a Normalização, nos dando uma percepção mais próxima da realidade. Para melhor compreensão do que ocorreu demonstramos abaixo como ficaria se aplicássemos a normalização sem antes ajustarmos os índices pelo "*Fit*" e como ficou depois do "*Fit*", conforme tabela abaixo:

Tabela: Comparativo antes e depois de aplicar *Fit*

				NORMALIZAÇÃO	
Linha no Minitab	País no Minitab	Shelter	Fit	Antes do Fit	Depois do Fit
23	Central African Republic	9,610	22,552	3,78	18,93
66	Kyrgystan	42,740	56,670	42,56	58,87
82	Mongólia	36,390	51,517	35,13	52,84
99	Phillipines	62,930	77,986	66,19	83,82
112	Sudan	34,770	38,696	33,23	37,83
113	Swaziland (South Africa)	35,680	41,126	34,30	40,67
115	Switzerland	88,480	86,516	96,10	93,80

Todos os países foram apontados para *Fit* por meio do MINITAB.

Após rodarmos pela segunda vez a Regressão, obtivemos os resultados descriminados abaixo:

Regression Analysis: Shelter_1N versus Social Progr; Access to pi; ...

```
The regression equation is
Shelter_1N = 5,58 + 0,506 Social Progress Index_1N
             + 0,0804 Access to piped water_1N
             + 0,344 Access to electricity (% of p_N
             + 0,177 Private property rights (0=no_N - 0,197 1. Governança_1_N
```

Predictor	Coef	SE Coef	T	P
Constant	5,578	1,846	3,02	0,003
Social Progress Index_1N	0,5059	0,1246	4,06	0,000
Access to piped water_1N	0,08038	0,04490	1,79	0,076
Access to electricity (% of p_N	0,34432	0,04777	7,21	0,000
Private property rights (0=no_N	0,17714	0,06602	2,68	0,008
1. Governança_1_N	-0,1973	0,1202	-1,64	0,103

```
S = 7,71741   R-Sq = 91,5%   R-Sq(adj) = 91,2%


Analysis of Variance

Source           DF     SS     MS       F      P
Regression        5  80923  16185  271,74  0,000
Residual Error  126   7504     60
Total           131  88427

Source                           DF  Seq SS
Social Progress Index_1N          1   70996
Access to piped water_1N          1    4486
Access to electricity (% of p_N   1    4999
Private property rights (0=no_N   1     282
1. Governança_1_N                 1     161


Unusual Observations

       Social
     Progress
Obs  Index_1N  Shelter_1N     Fit  SE Fit  Residual  St Resid
 34        24      55,840  39,339   2,108    16,501      2,22R
 55        22      74,660  62,556   3,243    12,104      1,73 X
 69        49      52,270  70,693   1,764   -18,423     -2,45R
 87        34      33,380  50,525   1,748   -17,145     -2,28R
104        57      88,530  71,442   0,935    17,088      2,23R
112        11      37,830  29,300   3,237     8,530      1,22 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

1.1 – Normalização dos dados

Para a normalização foram empregadas as seguintes rotinas:

1 – Estabeleceu como parâmetro para normalização o valor máximo e o mínimo.

Primeiro: Clique em Calc no menu e abrirá uma janela, então, siga a rotina: Clique em calc (calculadora no MINITAB) em seguida aparece várias opções, então, clique em Calculator e aparece uma janela pedindo: Store result in variable (indique a coluna para onde vai o dado) logo abaixo está: Expression insira a fórmula: **(c9-MIN(c9))/(MAX(c9)-MIN(c9))*100**.

Importante lembrar que o C9 aqui é o número da coluna do exemplo e você deve substituir pela coluna de sua pesquisa que quer alterar);
Depois clique OK e vai aparecer os dados alterados na coluna que você indicou.

2 – Deixar os resultados com duas casas depois da virgula, foi empregado a fórmula: 0,01*ROUND(100*número da coluna desejada).
Exemplo: se o interesse era alterar os dados da Coluna 22, então no lugar do número da coluna inserimos C22 e o MINITAB gerou os valores com duas casas depois da virgula.
Rotina: Clique em calc (calculadora no MINITAB) em seguida aparece várias opções, então, clique em Calculator e aparece uma janela pedindo: Store result in variable (indique a coluna que quer alterar) logo abaixo está: Expression (insira a expressão 0,01*ROUND(100*número da coluna desejada).

Após a normalização e deixar os resultados com duas casas depois da vírgula rodamos novamente os valores de Estatística Descritiva: N, N*, Mean, SE mean, StDev, Minimo e máximo; conforme se apresentam a seguir:

Descriptive Statistics: Social Progr; Shelter_1N; Access to pi; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index_1N	132	0	55,84	2,22	25,52	0,00	34,74
Shelter_1N	132	0	64,33	2,26	25,98	0,00	42,13
Access to piped water_1N	132	0	63,05	3,03	34,85	0,00	24,39
Access to improved sanit	132	0	70,97	2,78	31,93	0,00	48,02
Availability of affordab	132	0	44,61	1,69	19,47	0,00	30,48
Access to electricity (%	132	0	77,43	2,92	33,57	0,00	56,25
Quality of electricity s	132	0	59,15	2,38	27,36	0,00	37,08
Private property rights	132	0	43,21	2,31	26,49	0,00	27,78
IDH - 2013_1_N	132	0	58,05	2,25	25,81	0,00	38,38
1. Governança_1_N	132	0	45,02	2,22	25,46	0,00	25,83
EPI Score_1_N	132	0	49,18	2,08	23,90	0,00	30,14
Happy Planet Index_1_N	132	0	48,45	1,88	21,56	0,00	33,77
GINNI Index_1_N	132	0	37,07	1,84	21,13	0,00	21,52

Variable	Median	Q3	Maximum
Social Progress Index_1N	56,19	73,94	100,00
Shelter_1N	71,16	84,78	100,00
Access to piped water_1N	74,15	94,91	100,00
Access to improved sanit	84,92	98,39	100,00

Availability of affordab	43,84	57,53	100,00
Access to electricity (%	98,96	100,00	100,00
Quality of electricity s	62,66	79,45	100,00
Private property rights	38,89	59,72	100,00
IDH - 2013_1_N	62,36	78,71	100,00
1. Governança_1_N	39,21	65,05	100,00
EPI Score_1_N	48,61	69,28	100,00
Happy Planet Index_1_N	48,34	64,14	100,00
GINNI Index_1_N	37,07	48,88	100,00

Os dados foram normalizados, isto é, foram eliminados os outliers pelo valor mínimo de cada variável e foram lançados em nova planilha do MINITAB. A seguir apresentamos os dados da média, mínimo e máximo considerando as colunas com os dados normalizados.

Análise: Após a normalização dos dados observou-se que as variáveis tiveram alteração porque o critério estabelecido foi a diferença entre o valor máximo e o valor mínimo, dado pela expressão: (c9-MIN(c9))/(MAX(c9)-MIN(c9))*100. Novamente o registro de que o C9 é apenas referência da coluna no exemplo.

Então, a coluna C4 que representa Social Progress Index originalmente apresentava uma média de: 63,67 e com a normalização passou para 55,84 com duas casas depois da vírgula.

Isto foi observado em outras variáveis como a coluna C5 que mostra os dados de Shelter (lar) com valor médio original de 60,85 e passou com a normalização para 64,33. Isto não elimina o poder de interpretação e análise mesmo com a diferença para maior.

- Considerando o desvio padrão (StDev) – que mostra a distância estimada entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação à média. Nesse caso podemos citar como exemplo IDH com média de 58,13 e desvio de 25,82.

2. STEPWISE

Stepwise Regression: Shelter_1N versus Social Progr; Access to pi; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is Shelter_1N on 5 predictors, with N = 132

Step	1	2	3	4	5
Constant	13,396	5,170	4,559	5,571	5,578

Social Progress Index_1N	0,912	0,516	0,411	0,343	0,506
T-Value	23,01	12,14	6,04	4,54	4,06
P-Value	0,000	0,000	0,000	0,000	0,000
Access to electricity (% of p_N		0,392	0,426	0,384	0,344
T-Value		12,13	11,73	9,27	7,21
P-Value		0,000	0,000	0,000	0,000
Private property rights (0=no_N			0,089	0,098	0,177
T-Value			1,96	2,16	2,68
P-Value			0,052	0,033	0,008
Access to piped water_1N				0,090	0,080
T-Value				2,00	1,79
P-Value				0,047	0,076
1. Governança_1_N					-0,20
T-Value					-1,64
P-Value					0,103
S	11,6	7,95	7,86	7,77	7,72
R-Sq	80,29	90,79	91,06	91,33	91,51
R-Sq(adj)	80,14	90,65	90,85	91,06	91,18
Mallows Cp	164,7	10,7	8,8	6,7	6,0

Com a regressão Stepwise acima, onde a variável *Shelter* é apontada como *Response* podemos verificar que obtivemos grande grau de explicação com R-Quadrado em 80,29%. Todos com P-Value abaixo de 0,10 demonstrando grande confiabilidade.

Stepwise Regression: Social Progr versus Shelter_1N; Access to pi; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is Social Progress Index_1N on 5 predictors, with N = 132

Step	1	2	3	4	5
Constant	-0,7835	-1,3753	0,5839	0,6624	-0,2175
Shelter_1N	0,880	0,527	0,353	0,367	0,229
T-Value	23,01	17,93	8,08	8,92	4,06

P-Value	0,000	0,000	0,000	0,000	0,000
1. Governança_1_N		0,518	0,509	0,694	0,727
T-Value		17,25	18,47	13,61	14,58
P-Value		0,000	0,000	0,000	0,000
Access to piped water_1N			0,153	0,134	0,100
T-Value			5,07	4,65	3,43
P-Value			0,000	0,000	0,001
Private property rights (0=no_N				-0,188	-0,172
T-Value				-4,22	-4,01
P-Value				0,000	0,000
Access to electricity (% of p_N					0,126
T-Value					3,45
P-Value					0,001
S	11,4	6,28	5,75	5,41	5,19
R-Sq	80,29	94,04	95,04	95,65	96,02
R-Sq(adj)	80,14	93,94	94,92	95,51	95,86
Mallows Cp	496,4	62,9	33,2	15,9	6,0

Agora apontamos a variável *Social Progress Index* como *Response* e percebemos que se manteve elevado o R-Quadrado em 80,29%, e P-Value sempre abaixo de 0,10.

Stepwise Regression: 1. Governanç versus Social Progr; Shelter_1N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is 1. Governança_1_N on 5 predictors, with N = 132

Step	1	2	3	4
Constant	7,0751	-3,9777	0,5733	1,1170
Private property rights (0=no_N	0,878	0,540	0,407	0,417
T-Value	25,61	14,83	12,24	12,50
P-Value	0,000	0,000	0,000	0,000
Social Progress Index_1N		0,460	0,791	0,840
T-Value		12,17	15,96	15,10

P-Value		0,000	0,000	0,000
Access to electricity (% of p_N			-0,224	-0,173
T-Value			-8,46	-4,59
P-Value			0,000	0,000
Shelter_1N				-0,119
T-Value				-1,87
P-Value				0,064
S	10,4	7,12	5,72	5,67
R-Sq	83,46	92,30	95,06	95,20
R-Sq(adj)	83,33	92,18	94,95	95,04
Mallows Cp	310,4	78,1	6,8	5,3

Temos agora a variável Governança como *Response* e percebemos que houve aumento no R-Quadrado para 83,46% e todos os P-Value com alto grau de confiabilidade.

CAPITULO 4 – ANÁLISE DE TENDÊNCIAS

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise de tendência e projeções das variáveis temporais e quantitativas “PIB”, “CRESCIMENTO DO PIB” e “BALANÇA COMERCIAL”. Todos os dados utilizados são referentes ao Brasil, Espanha, México e Portugal. Estas variáveis integram a pesquisa SOFI (*State of the Future Index*). A análise de cada variável está dividida em três partes. A primeira refere-se a análise do comportamento histórico através de gráficos. A segunda trata da análise de tendências, buscando encontrar a função que melhor se adapte à cada uma das séries de variáveis analisadas. A terceira busca a elaboração e análise de projeções através de extrapolações estatísticas. O software estatístico utilizado é o **MINITAB.**

1.1 Uma breve descrição do SOFI – *State of the Future Index*

A pesquisa SOFI é constituída por variáveis chaves utilizadas para previsões que, quando agregadas, possibilitam inferências sobre as tendências dos próximos 10 anos. Trata-se de um estudo do futuro. O objetivo é demonstrar as direções e a intensidade das mudanças para os próximos anos, buscando identificar os fatores responsáveis por estas alterações. São utilizadas 20 variáveis, que foram definidas e classificadas previamente levando em consideração a relevância para as previsões e a disponibilidade de dados nos últimos 20 anos nos diversos países que a compõe. Trata-se de um trabalho contínuo que vem sendo atualizado ano após ano.

2. ENTENDENDO OS DADOS

2.1 Os Indivíduos

Os dados são séries históricas referentes aos países já mencionados na introdução, portanto, tratam-se de séries temporais. As séries vão de 1970 a 2014 para as variáveis “PIB” e “CRESCIMENTO DO PIB”, e de 1992 a 2014 para a variável “BALANÇA COMERCIAL”.

2.2 As Variáveis

São 3 as variáveis desta pesquisa, mais o ano a que se referem os dados. As mesmas são melhor explicadas na Tabela 1.

Tabela 1. As Variáveis

Variável	Significado	Tipo	Unidade de Medida
Ano	É o ano a que se refere o dado de cada variável.	Variável Categórica	Formato: AAAA
PIB	Produto Interno Bruto é a conta nacional onde se calcula todo o valor produzido em bens e serviços, dentro do território nacional, em um determinado período de tempo.	Variável Quantitativa	(US$ - 1Bi)
CRESCIMENTO DO PIB	É a variação percentual crescente ou decrescente do PIB imediatamente anterior ou posterior ao PIB analisado.	Variável Quantitativa	Percentual (%)
BALANÇA COMERCIAL	É a conta do Balanço de Pagamentos onde estão registradas as Exportações e Importações. O Saldo desta conta poderá ser superavitário, quando a Balança Comercial apresenta maior valor em exportações, deficitário, quando a Balança Comercial apresenta menor valor em exportações ou nula quando o valor de exportações é exatamente igual ao das importações.	Variável Quantitativa	(US$ - 1Bi)

2.3 A fonte e o tamanho da série de dados

Fonte:

Todas as informações desta pesquisa foram obtidas diretamente da base de dados do sítio *knoema*, especializado em análises estatísticas e infográficos sócio-econômicos.

Tamanho da Série de Dados:

As três séries de dados podem ser consideradas satisfatórias para a realização desta pesquisa, uma vez que todas possuem dados de 10 anos ou mais.

2.4 A Tabela de Dados – Crescimento do PIB de 1970 a 2014 países selecionados

Valores em U$Bi	Brasil	Espanha	México	Portugal
1970	**42.327**	40.881	**35.541**	**8.109**
1990	**461.951**	535.101	**262.709**	**78.721**
2000	**657.216**	595.402	**683.647**	**118.358**
2005	**892.103**	1.157.276	**866.346**	**197.304**
2008	**1.694.585**	1.634.989	**1.101.275**	**262.007**
2009	**1.664.586**	1.499.074	**893.368**	**243.745**
2010	**2.209.433**	1.431.672	**1.049.925**	**238.317**
2011	**2.615.234**	1.487.924	**1.169.362**	**244.879**
2012	**2.413.135**	1.339.946	**1.184.499**	**216.368**
2014	**2.346.076**	1.381.342	**1.294.689**	**230.116**

2.6 A Tabela de Dados – Saldo Balança Comercial de 1992 a 2014 países selecionados

Ano	Brasil	Espanha	México	Portugal
1992	$15.239	-$35.424	-$18.017	-$11.231
1993	$13.299	-$12.922	-$15.661	-$8.784
1994	$10.466	-$12.167	-$21.104	-$9.072
1995	-$3.466	-$15.688	$5.115	-$9.827
1996	-$5.599	-$13.978	$4.021	-$10.584
1997	-$6.753	-$14.914	-$2.690	-$11.100
1998	-$6.623	-$24.689	-$11.612	-$13.642
1999	-$1.289	-$30.913	-$9.693	-$15.433
2000	-$732	-$40.892	-$13.097	-$15.589
2001	$2.685	-$37.990	-$14.492	-$15.382
2002	$13.196	-$39.418	-$12.405	-$14.248
2003	$24.877	-$52.455	-$9.643	-$15.443
2004	$33.842	-$75.708	-$14.280	-$19.161
2005	$44.929	-$96.142	-$14.033	-$25.183
2006	$46.456	-$114.979	-$13.515	-$25.934
2007	$40.032	-$136.004	-$18.425	-$29.647
2008	$24.957	-$139.310	-$27.039	-$37.279
2009	$25.273	-$65.880	-$11.803	-$27.452
2010	$20.147	-$72.598	-$11.900	-$28.343
2011	$29.793	-$70.055	-$11.499	-$22.830
2012	$19.395	-$42.088	-$9.834	-$14.013
2013	$2.286	-$22.765	-$11.004	-$12.802
2014	-$4.036	-$33.199	-$14.046	-$14.004

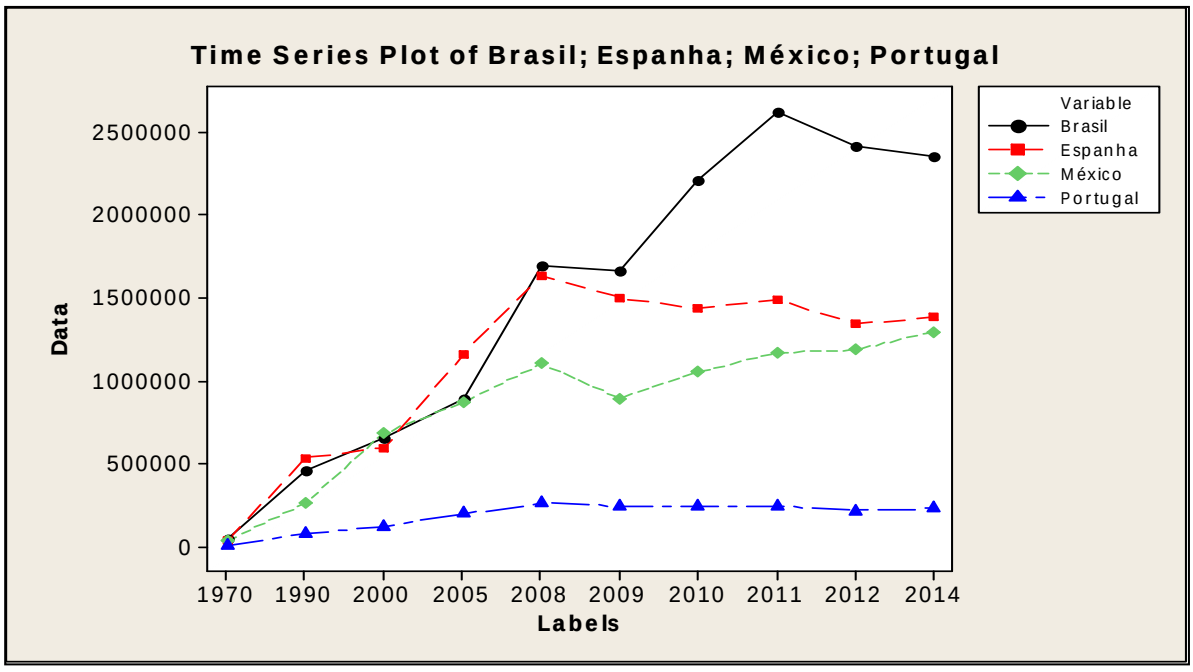
3. ANÁLISE DA VARIÁVEL PIB

Conforme mencionado acima, cada variável será analisada utilizando gráficos temporais para demonstrar o comportamento histórico da série, linhas de tendência, funções, erros das funções, além de extrapolações estatísticas.

3.1 Variável: Crescimento do PIB

3.1.1. O comportamento da variável

Para análise do comportamento histórico da variável: **Crescimento do PIB** utilizaremos, inicialmente, gráfico de série histórica, sobrepondo os resultados obtidos por todos os países analisados na mesma imagem, nos permitindo observar seus respectivos desempenhos em uma verificação visual preliminar.



De acordo com o gráfico acima, percebemos tratar-se de uma série temporal com tendência ascendente, porém com muitas oscilações durante o período, não demonstrando necessariamente ser uma série linear. Inicialmente percebemos que dentre todos os países selecionados é o Brasil quem mais se destaca sendo o único a alcançar o patamar de mais de US$2Tri em seu Produto Interno Bruto. Em seguida temos Espanha e México que também

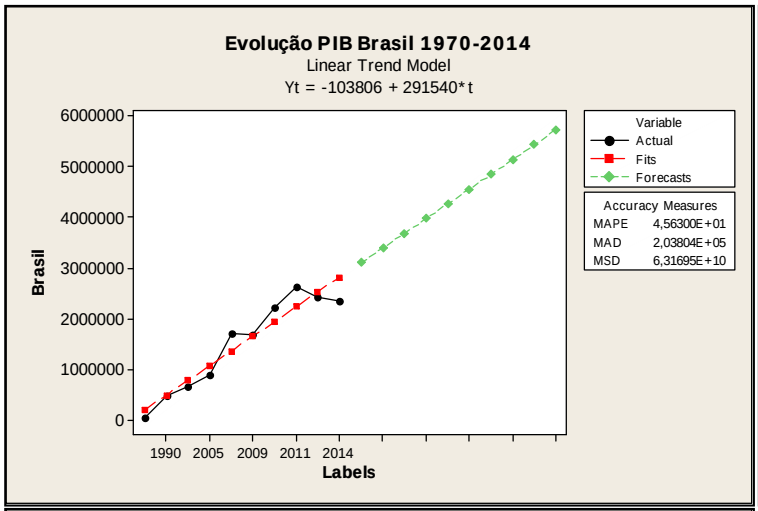
demonstram ascensão em seus respectivos PIBs, sobretudo no período entre 2000 a 2008, passando a se desestabilizar a partir da crise mundial de 2008, sendo que a Espanha segue em relativa baixa até 2014 e México demonstra ter se recuperado mais rapidamente alcançando novamente crescimento do PIB, porém em menor ritmo em relação ao período entre 2000 a 2008.

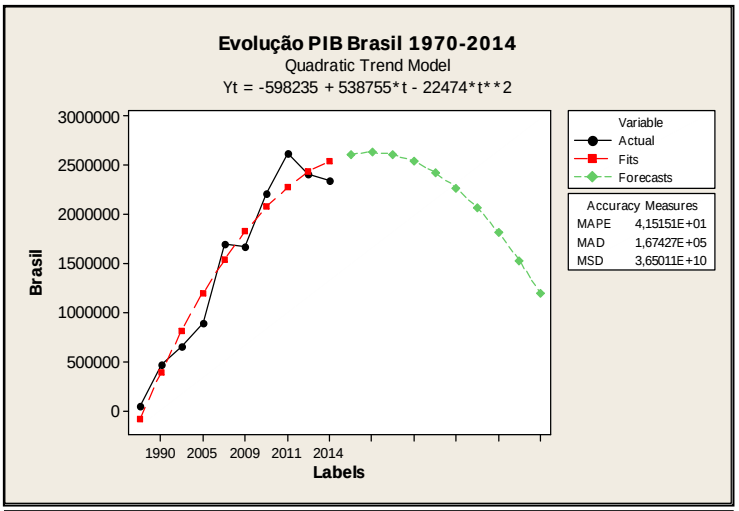
Já Portugal durante todo o decorrer da série histórica, que compreende um período de mais de quatro décadas nunca superou a marca dos US$300Bi. Percebe-se que eles também foram beneficiados pelo período de grande crescimento mundial entre 2000 a 2008, mas também sofrem queda de desempenho a partir de 2009.

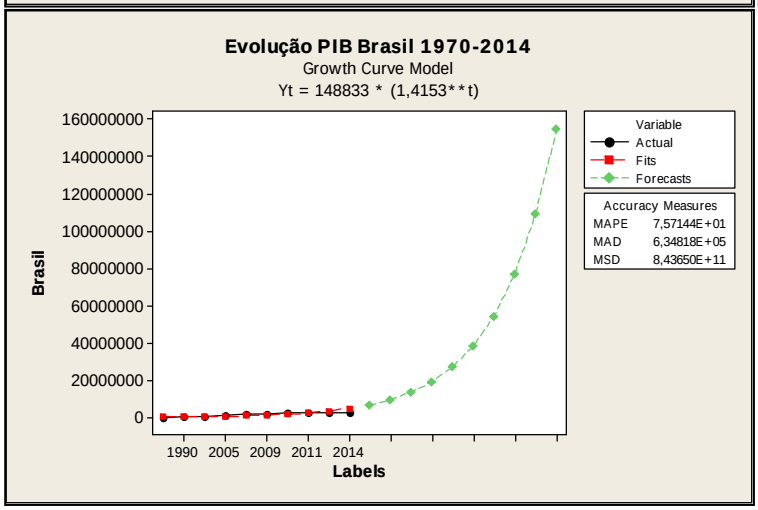
3.1.2. A análise de tendência - Brasil

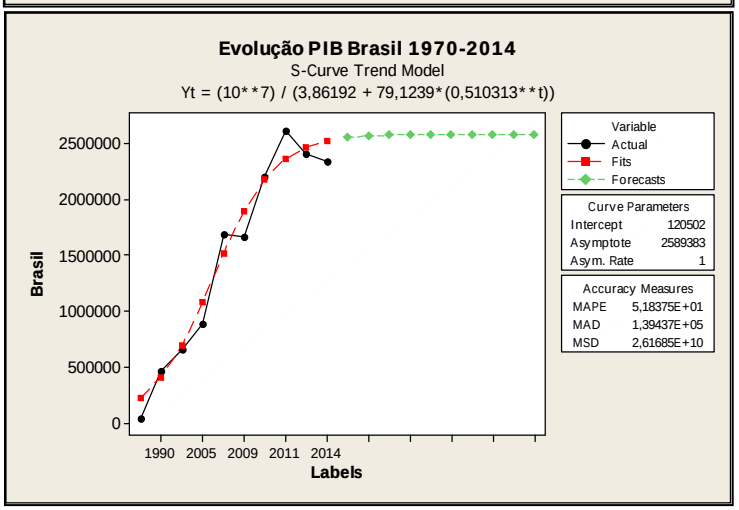
Nesta seção analisaremos a tendência da série de dados, buscando encontrar a função (linear, quadrática, exponencial ou s-curve) que melhor a represente. Para medir a eficiência da função ou da curva de tendência, analisaremos os erros: MAPE (*Mean Average Percentual Error)*, MAD (*Mean Absolute Deaviation)* e MSD (*Mean Standard Deviation).*

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

	BRASIL			
	linear	Quadratic	Exponencial	S-Curve
MAPE	4,5630	4,1515	7,5714	5,1837
MAD	2,0380	1,6742	6,3481	1,3943
MSD	6,3169	3,6501	8,4365	2,6168

Ao analisarmos os gráficos estritamente pelos índices de erros temos a função S-Curve como a que apresenta menor índice de erros para as variáveis *MAD* E *MSD*, sendo esta, portanto a que melhor explica as tendências para os próximos dez anos. Percebemos que nesta previsão o PIB brasileiro ainda permaneceria sempre acima da casa dos US$2,5Tri, conforme tabela abaixo

3.1.3. As previsões

Utilizando a equação da função S-Curve, apresentada no gráfico acima, temos os seguintes valores previstos para os próximos 10 anos em relação ao PIB no Brasil.

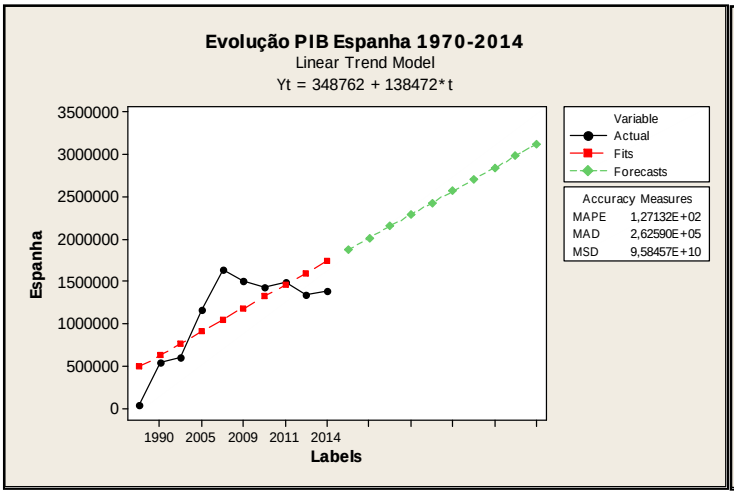
Period	Forecast US$bi
2015	2557357
2016	2572940
2017	2580966
2018	2585081
2019	2587186
2020	2588261
2021	2588810
2022	2589091
2023	2589234
2024	2589307

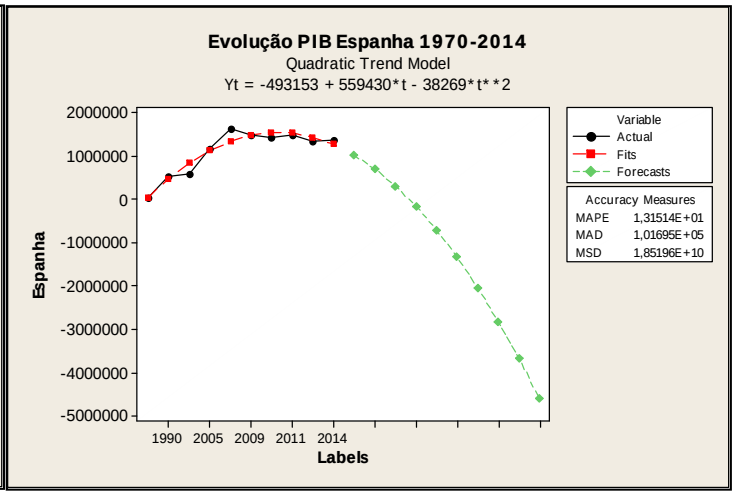
De fato a S-Curve foi dentre as demais curvas de tendência a que mais se aproximou da realidade do PIB Brasileiro em 2015. Ocorre que o País obteve recessão técnica e percebeu uma retração de 3,8% de seu PIB em relação a 2014. Além deste fato com a desvalorização do Câmbio o país fechou seu PIB em Dólar na casa dos US$1,8Tri. Os demais gráficos de tendência apontaram PIB superior a US$2,5Tri o que não se concretizou na realidade.
Fonte: IBGE

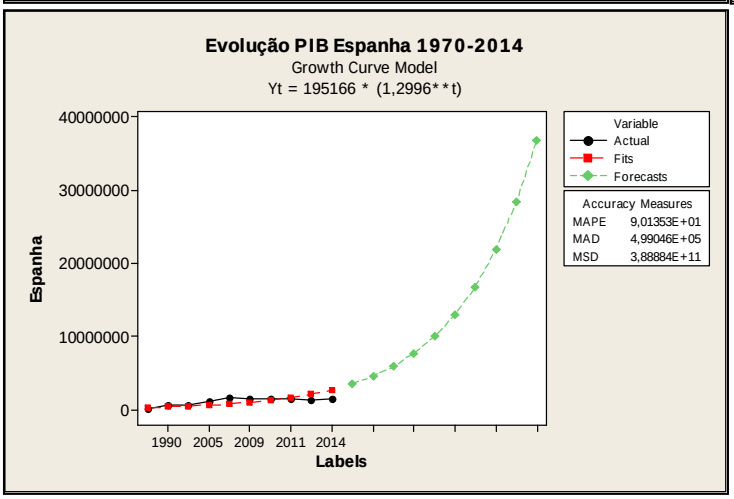
Vale ressaltar que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o

comportamento passado se repetirá. Estudos econômicos vêm demonstrando que a previsão do PIB tem relação com uma série de variáveis como nível de investimentos, taxas de desemprego etc. Assim, talvez a utilização de regressões múltiplas poderia ser mais eficiente neste trabalho de previsão.

3.1.4 Análise de tendência – Espanha, México e Portugal

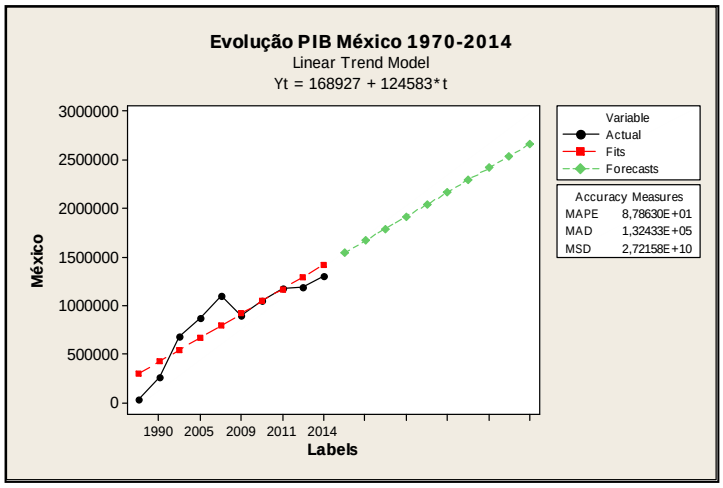


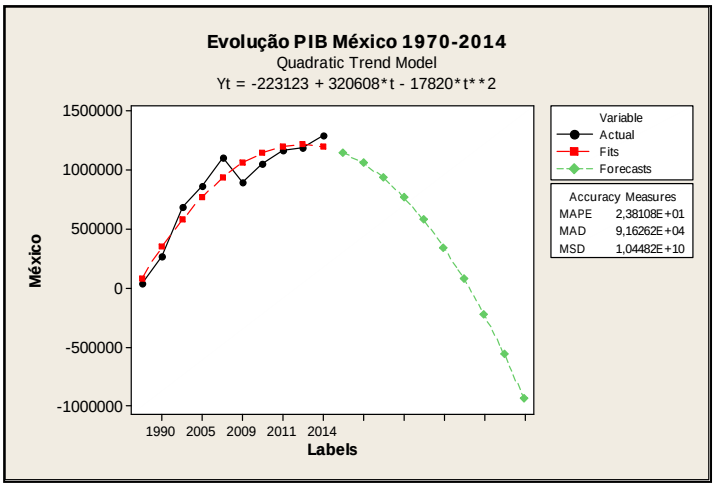


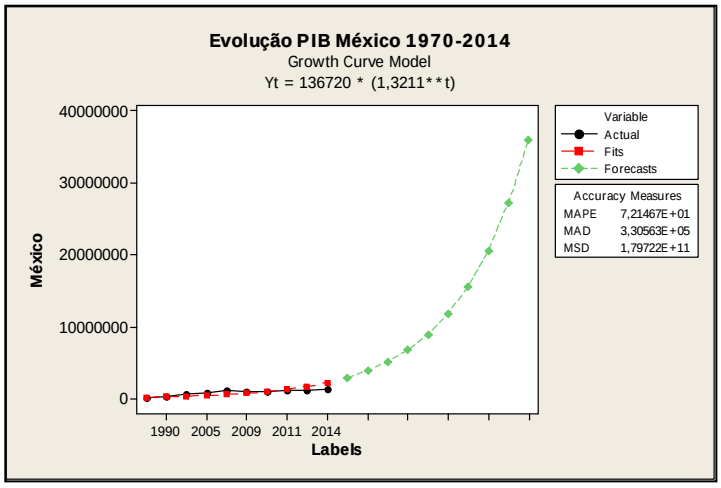


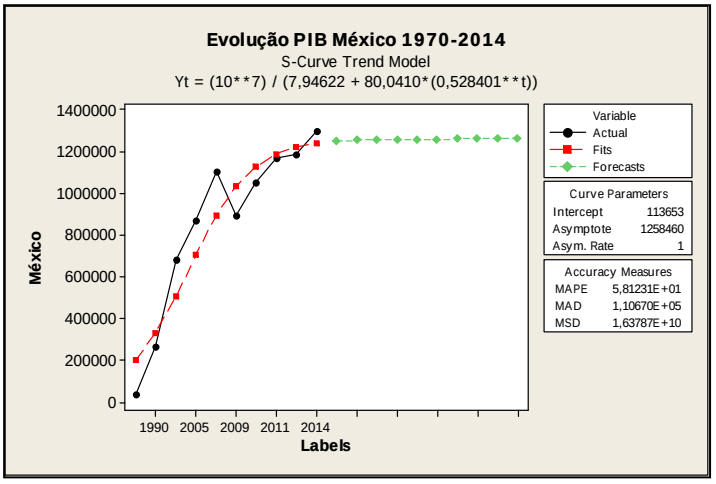
	ESPANHA			
	linear	Quadratic	Exponencial	scurve
MAPE	1,2713	1,3151	9,0135	error
MAD	2,6259	1,0169	4,9904	error
MSD	9,5845	1,8519	3,8888	error

Para a Espanha percebemos que o gráfico que melhor explica a tendência seria sem dúvidas a Quadrática onde as três variáveis de erro apresentam os menores índices e que aponta claramente queda acentuada no Produto Interno Bruto do País. No entanto foi a Espanha que segundo a própria União Europeia liderou a retomada do crescimento europeu no ano de 2015 (http://brasil.elpais.com/brasil/2015/05/05/economia/1430814645_238655.html), alcançando crescimento entre 2,8% a 3,5% não tendo sido formalizado ainda pela *eurostat* (departamento de estatística da UE). Portanto, na realidade observada em 2015 foi a regressão Linear que de fato mais se aproximou do PIB alcançado pela Espanha no ano de 2015.



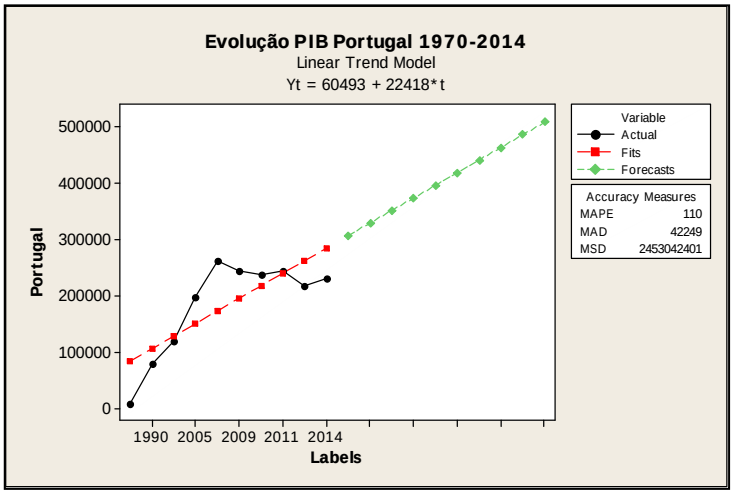


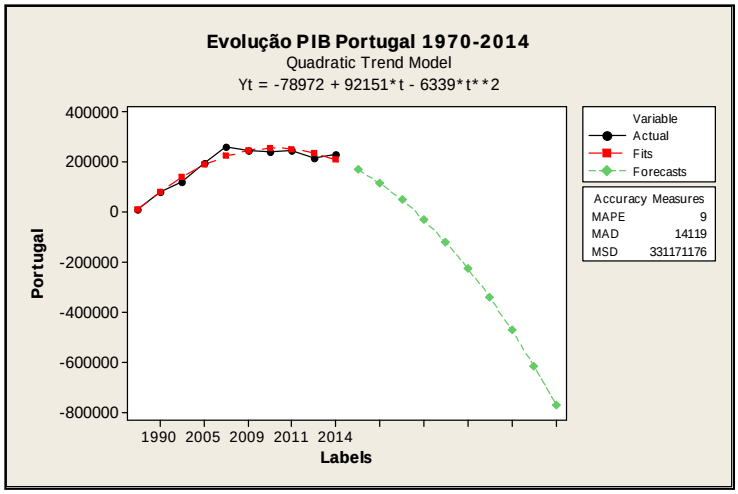


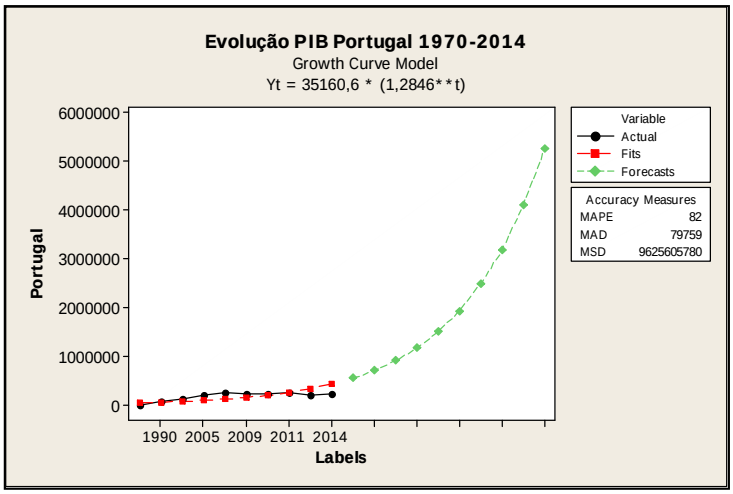


	MÉXICO			
	linear	Quadratic	Exponencial	S-Curve
MAPE	8,7863	2,3810	7,2146	5,8123
MAD	1,3243	9,1626	3,3056	1,1067
MSD	2,7215	1,0448	1,7972	1,6378

Para a análise mexicana percebemos que o gráfico que melhor explica a tendência para os próximos dez anos seria a S-Curve, pois aponta menor índice de erro para as variáveis *MAD* E *MSD*.







	PORTUGAL			
	linear	Quadratic	Exponencial	scurve
MAPE	110	9	82	error
MAD	42249	14119	79759	error
MSD	245304	331171	962560	error

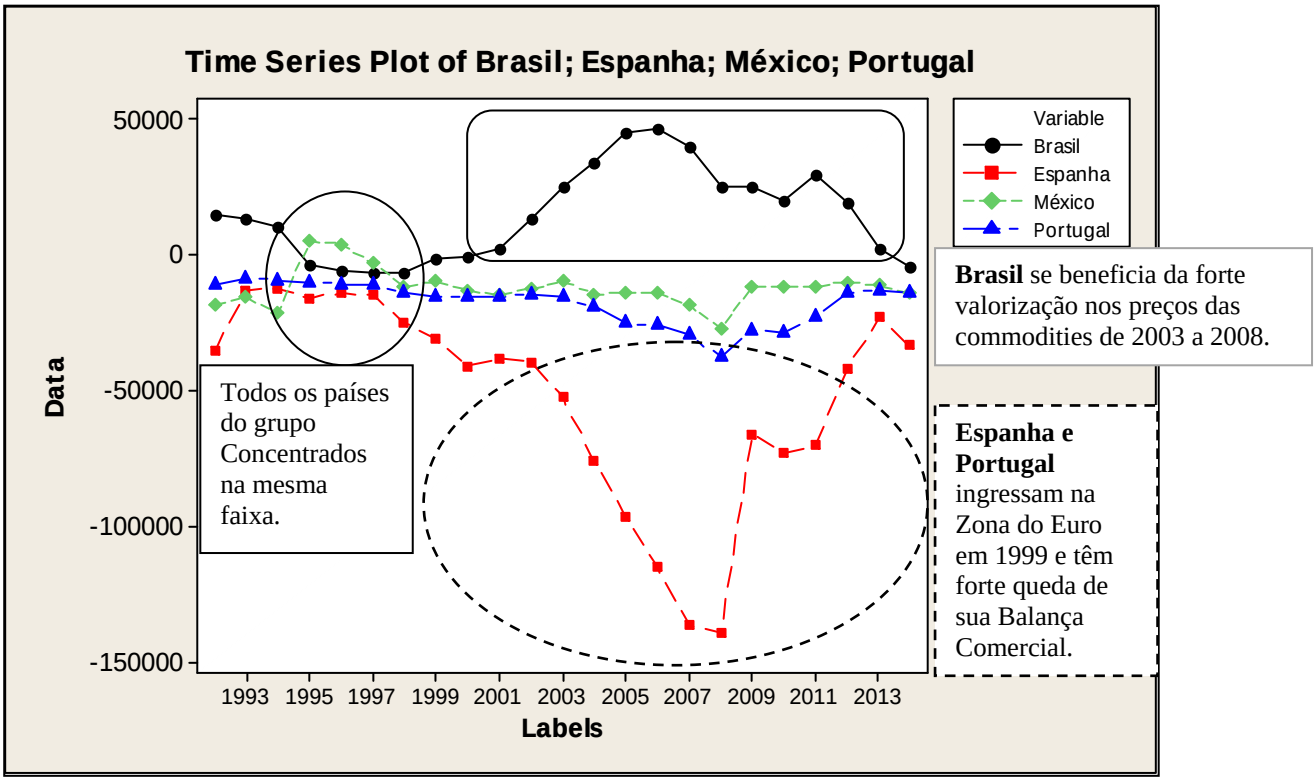
Para Portugal podemos verificar que a Análise Quadrática é que apresenta menores índices de erros para *MAPE* e *MAD*.

4. ANÁLISE DA VARIÁVEL BALANÇA COMERCIAL

4.1 Variável: Crescimento da Balança Comercial

4.1.1 O comportamento da variável

Para análise do comportamento histórico da variável: **Balança Comercial** utilizaremos, inicialmente, gráfico de série histórica, sobrepondo os resultados obtidos por todos os países analisados na mesma imagem, nos permitindo observar seus respectivos desempenhos em uma verificação visual preliminar.



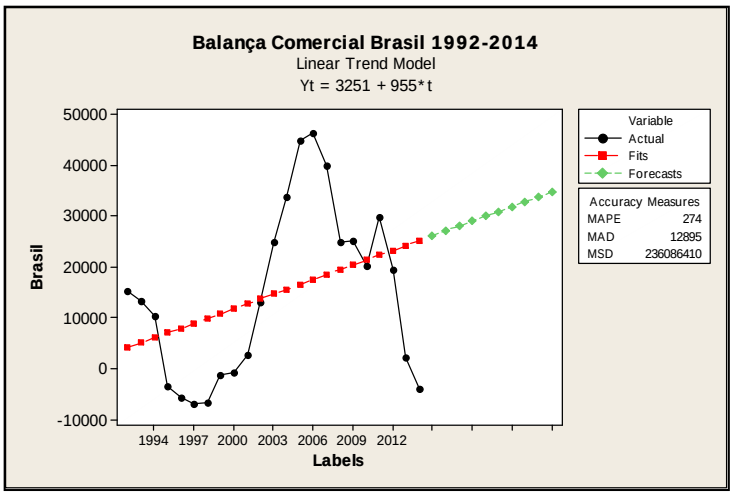
O gráfico acima demonstra a evolução da Balança Comercial desde 1992 a 2014 nos seguintes países: Brasil, Espanha, México e Portugal.

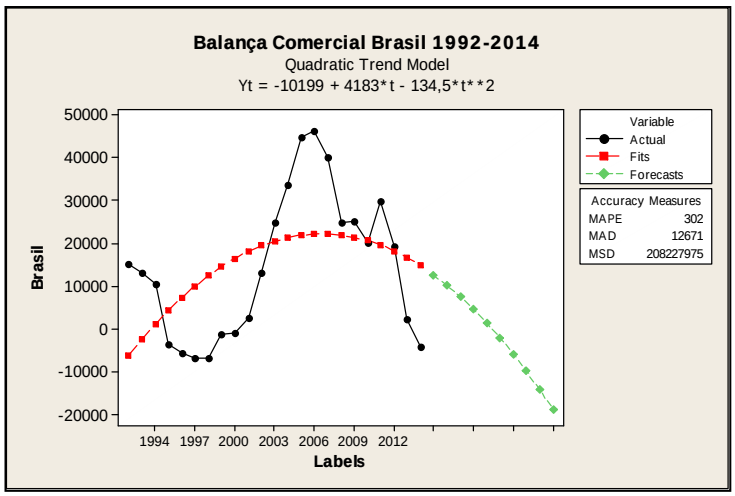
Podemos verificar que a exemplo do crescimento do PIB analisados no item 3.1.1 foi o Brasil quem obteve maior destaque também no crescimento da Balança Comercial.

Percebemos que até 1997 todos os países do grupo estavam concentrados numa mesma faixa do gráfico, oscilando ora em uma pequena balança superavitária ora deficitária.

No entanto, a partir de 1999 Espanha e Portugal ingressam na Zona do Euro e para a Espanha o resultado desfavorável em sua Balança Comercial é imediato, uma vez que com acesso ao crédito, sua população passava a consumir os produtos produzidos principalmente na Alemanha e França. Portugal enfrentou problema semelhante, porém, embora sua balança comercial também permanecesse sempre desfavorável, o impacto foi menor do que os apresentados pela Espanha.

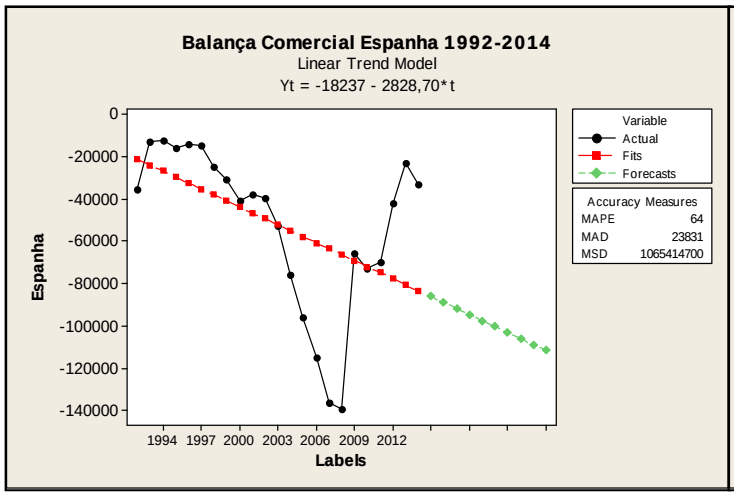
4.2 Análise de tendência Brasil

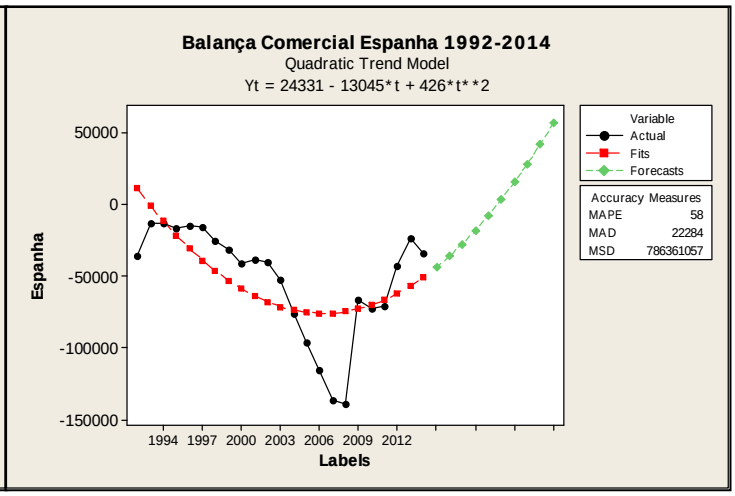




	BRASIL			
	linear	Quadratic	Exponencial	S-Curve
MAPE	274	302	error	error
MAD	12895	12671	error	error
MSD	236086	208227	error	error

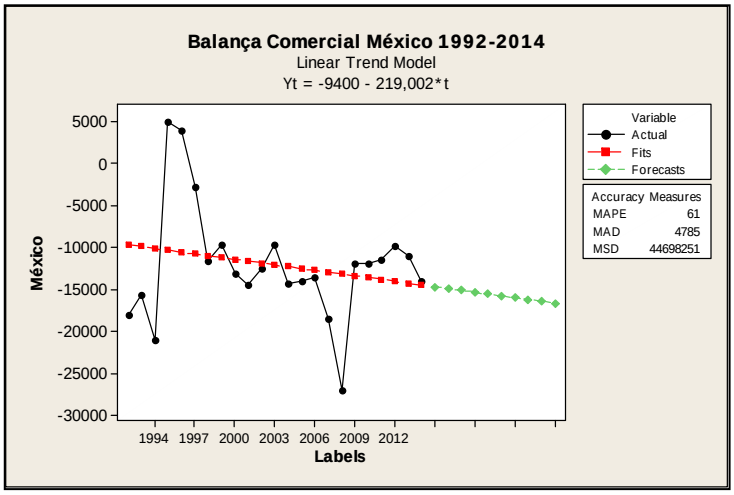
De acordo com a tabela de erros acima, podemos verificar que para o Brasil a Curva de Tendência que melhor se aplicaria seria a Quadrática, com índices *MAD* e *MSD* inferiores. No entanto na realidade, já a partir de 2015 nossa balança comercial volta a se recuperar apresentando saldo positivo de US$19,69Bi, ante o saldo negativo apresentado em 2014 de (-)US$4,03Bi.

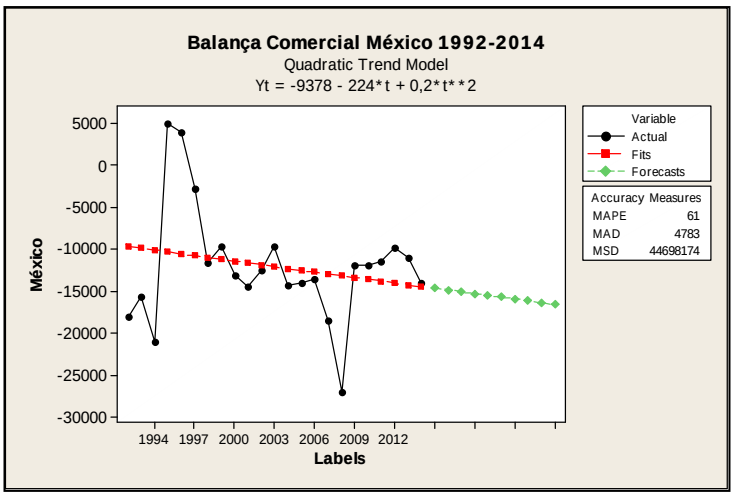


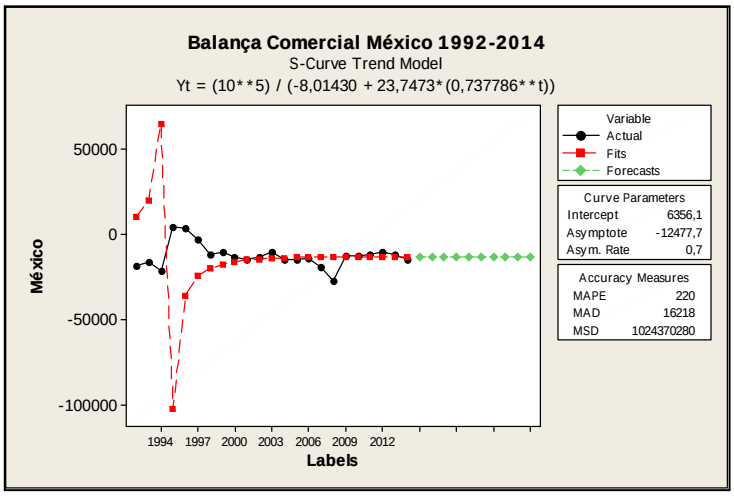


	ESPANHA			
	linear	Quadratic	Exponencial	scurve
MAPE	64	58	error	error
MAD	23831	22284	error	error
MSD	106514	786361	error	error

Para análise da Balança Comercial da Espanha, podemos verificar que de acordo com a tabela de erros poderíamos escolher a função quadrática para prever a tendência de seu saldo com índices *MAPE* e *MAD* inferiores. No entanto as tendências geradas com o gráfico quadrático apontam recuperação já para o ano de 2015, fato que não se concretizou, uma vez que, segundo a *eurostat* o país obteve queda em sua balança comercial no ano de 2015 com relação a 2014.



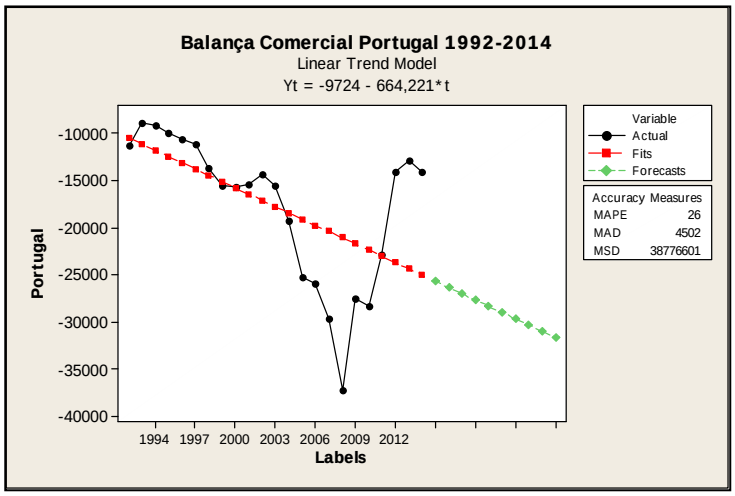


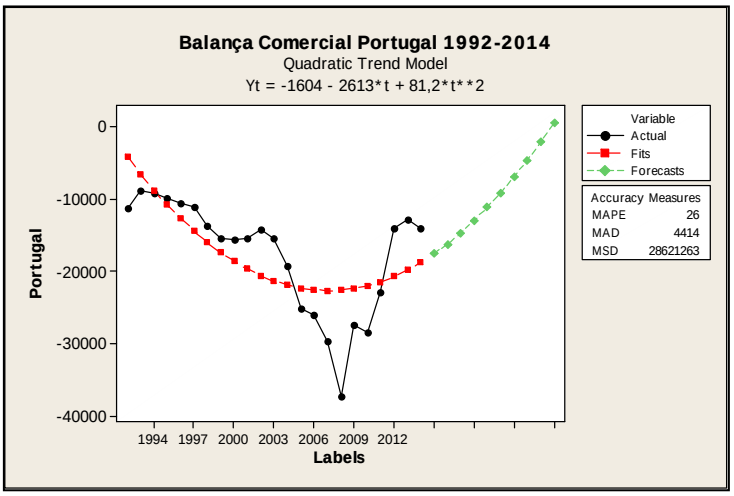


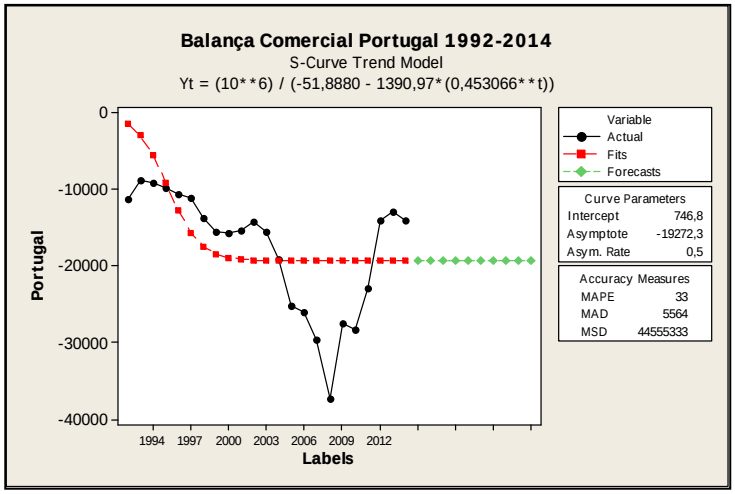
	MÉXICO			
	linear	Quadratic	Exponencial	scurve
MAPE	61	61	error	220
MAD	4785	4783	error	16218
MSD	446982	446981	error	102437

Passando a analisar os gráficos da Balança Comercial do México, podemos verificar que as tendências lineares e quadráticas estão muito semelhantes, ambas apontando tendências decrescente para os dez próximos períodos, sendo que a quadrática ficou com uma diferença mínima inferior nos índices de erros *MAPE*, *MAD* e *MSD*.

O gráfico com a tendência S-Curve aponta para uma estagnação para o próximo período de dez anos, porém apresenta índices de erro muito superiores em relação às análises anteriores.







	PORTUGAL			
	linear	Quadratic	Exponencial	scurve
MAPE	26	26	error	33
MAD	4502	4414	error	5564
MSD	387766	286212	error	445553

Finalmente, temos acima os gráficos de tendência da Balança Comercial de Portugal. Podemos verificar que a função quadrática é que apresenta os menores índices de erro. Portugal, a exemplo de México e Espanha estão com saldo negativo em suas respectivas balanças comerciais no decorrer de toda série histórica que estamos analisando (1992 a 2014). Exceção ao México que teve saldo positivo apenas nos anos de 1995 e 1996.

CAPITULO 5 – COMPARAÇÕES: Um Estudo dos Indicadores de Moradia da América Latina Comparado às Demais Regiões do Mundo

1. INTRODUÇÃO

O presente capitulo tem por objetivo efetuar uma análise comparativa de médias e intervalos de confiança de dados dimensionadores do nível de Moradia de diversos países do mundo. O principal propósito é comparar a América Latina com as demais regiões do mundo. Os dados são compilados pelo Banco Mundial. O software estatístico utilizado é o **MINITAB.**

2. OS INDIVIDUOS

Inicialmente nosso trabalho estava divido em OITO principais regiões globais, conforme tabela abaixo. No entanto, para não tornar nossa análise desnecessariamente longa, centralizamos nossa análise para apenas três regiões, onde mantivemos inalteradas as regiões AIBER e AVECO e juntamos as demais regiões passando a nomeá-las como: OUTROS.

Tabela 1. Oito Regiões Globais

AIBER	AVECO	CARLA	CWEAS	EMDEU	ESEAS	NSWUA	SUSAF
Argentina	Australia	Guyana	Armenia	Albania	Bangladesh	Algeria	Angola
Bolivia	Austria	Jamaica	Azerbaijan	Bosnia and Herzegovina	Cambodia	Egypt	Benin
Brazil	Belgium	Trinidad and Tobago	Belarus	Bulgaria	China	Iran	Botswana
Chile	Canada		Georgia	Croatia	India	Iraq	Burkina Faso
Colombia	Czech Republic		Kazakhstan	Hungary	Indonesia	Jordan	Burundi
Costa Rica	Denmark		Kyrgyzstan	Lithuania	Laos	Kuwait	Cameroon
Cuba	Estonia		Moldova	Macedonia	Malaysia	Lebanon	Central African Republic
Dominican Republic	Finland		Russia	Montenegro	Mongolia	Mauritania	Chad
Ecuador	France		Tajikistan	Poland	Nepal	Morocco	Congo, Republic of
El Salvador	Germany		Ukraine	Romania	Philippines	Pakistan	Djibouti
Guatemala	Greece		Uzbekistan	Serbia	Sri Lanka	Saudi Arabia	Ghana
Honduras	Iceland			Turkey	Thailand	Tunisia	Guinea
Mexico	Ireland					United Arab Emirates	Kenya
Nicaragua	Israel					Yemen	Lesotho
Panama	Italy						Liberia
Paraguay	Japan						Madagascar
Peru	Korea, Republic of						Malawi
Portugal	Latvia						Mali
Spain	Netherlands						Mauritius
Uruguay	New Zealand						Mozambique
Venezuela	Norway						Namibia
	Slovakia						Niger
	Slovenia						Nigeria
	Sweden						Rwanda
	Switzerland						Senegal
	United Kingdom						South Africa
	United States						Sudan
							Swaziland
							Tanzania
							Togo
							Uganda
							Zambia

2. ENTENDENDO OS DADOS

Os dados são referentes ao ano de 2014. As variáveis são indicadores agregados da qualidade de moradia de cada país. Os indivíduos desta análise que envolve 132 países e os indicadores de desenvolvimento humano, foram selecionados a partir de referência internacional, citamos a origem do trabalho que é o Projeto ORIBER – Observatório da Rede Ibero-Americana de Prospectiva.

2.2 As Variáveis

Este tema da pesquisa apresenta treze variáveis, sendo todas elas quantitativas.

Tabela 2. As **13** Variáveis do Eixo temático: **Moradia**

Variável	Significado	Tipo	Unidade de Medida	Fonte
SOCIAL PROGRESS INDEX	O índice identifica os pontos fortes e fracos dos países a partir de um conjunto de informações e análises conduzidas para cada um deles.	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
SHELTER (Abrigo)	Este índice indica se as pessoas possuem abrigos (lar) com suas necessidades básicas atendidas?	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
Acesso a Água e saneamento Básico	Porcentagem da população que tem uma ou mais torneiras de água encanada tratada.	Variável Quantitativa	Percentual	WHO UNICEF
Acesso a tratamento sanitário Melhorado	Inclui a população com melhores condições de saneamento, canalização de esgoto, fossas sépticas, latrinas com laje ou melhoradas para as condições sanitárias adequadas	Variável Quantitativa	Percentual	WHO UNICEF
Disponibilidade e Acessibilidade do Imóvel	Percentual da população que responderam em pesquisa se “em sua cidade ou região em que vive, está satisfeito com a disponibilidade e acessibilidade do imóvel”	Variável Quantitativa	Percentual	Galoop World Poll
Acesso a Eletricidade	Percentual de pessoas com acesso a energia elétrica.	Variável Quantitativa	Percentual	UN Sustanable energy for All Project

Variável	Significado	Tipo	Unidade de Medida	Fonte
Qualidade no fornecimento de energia elétrica	Média das respostas para a pergunta: "Em seu país, como você avalia a confiança no fornecimento de energia?". Usando Escala onde 1 não é confiável e 7 extremamente confiável para mostrar a qualidade do fornecimento de energia elétrica.	Variável Quantitativa	Indicador nominal que varia de 1 a 7	World Economc Forum Global Competitiveness Report
Direitos de Propriedade Privada	O grau com que um país protege com leis os Direitos com Propriedade Privada. Medido em escala onde 0 corresponde a nenhuma proteção pelo estado ao direito de propriedade privada e elevado grau de corrupção e 100 corresponde ao total suporte do Governo e nenhum grau de corrupção.	Variável Quantitativa	Indicador nominal que varia de 0 a 100	Heritage Foundation
IDH Índice de Desenvolvimento Humano	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável Quantitativa	Indicador nominal que varia de 0 a 1	PNUD
Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as politicas são implementadas e qual a capacidade de formulá-las.	Variável Quantitativa	Percentual	WGI 2014
EPI- Indice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	Variável Quantitativa	Percentual	EPI 2014
HPI – Bem estar sustentável do país	Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	Variável Quantitativa	Percentual	HPI 2014
GINNI index	Mede a distribuição de renda e de despesas das famílias	Variável Quantitativa	Percentual	Banco Mundial

Fonte: Elaboração própria; IPS, Unicef, GWP, WEF, Heritage Foundation, PNUD, WGI, EPI, Banco Mundial

3. ANÁLISE DAS VARIÁVEIS

Lembramos que o objetivo de nosso trabalho é comparar a América Latina com as demais regiões do mundo, para cada variável. Assim, cada variável será separada por região. Em outras palavras, os dados de cada variável para a América Latina serão comparados com cada uma das demais regiões: AVECO (que compõe o grupo dos 27 países desenvolvidos) e com: Ásia Leste, Ásia Sul, Europa Oriental e Bálticos, Antiga União Soviética, Caribe, Oriente Médio e África Norte, África Subsaariana (que compõem o grupo Outros). O procedimento se repete para cada variável.

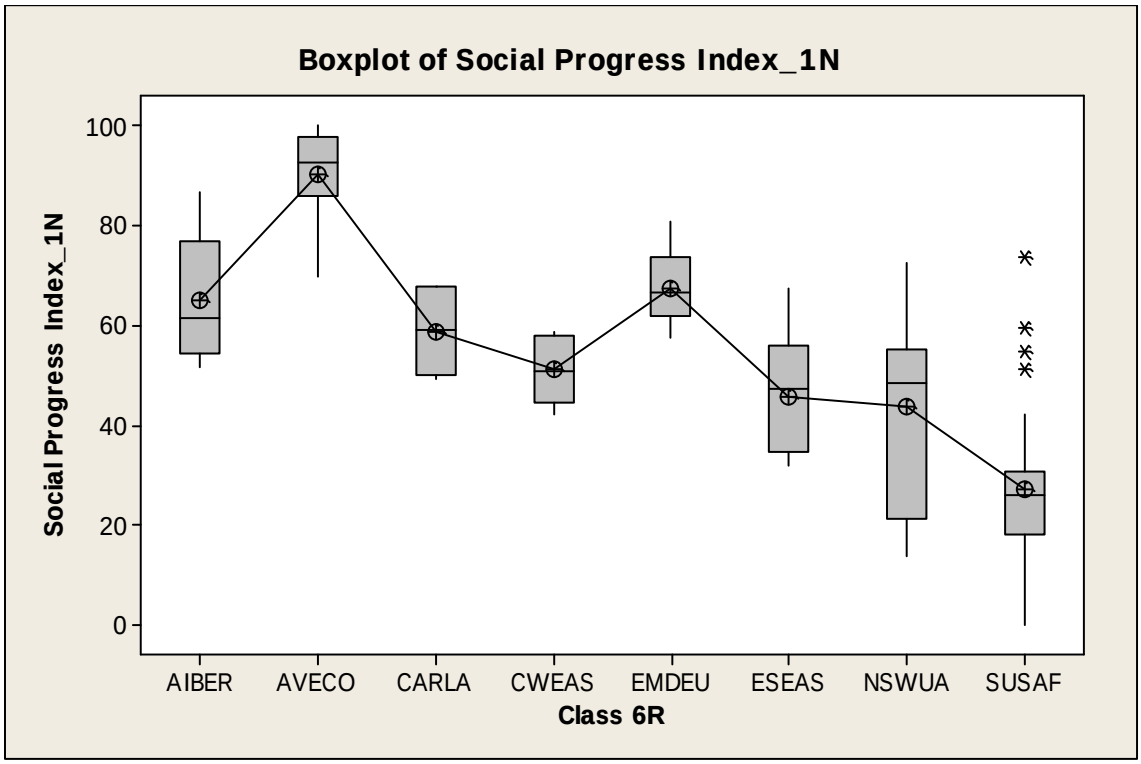
Para as comparações serão analisados: os resumos descritivos de cada série de dados, a média, o intervalo de confiança para a média e a variância.

Geramos no software MINITAB um único gráfico boxplot com as oito regiões citadas na tabela 1, utilizamos a variável *Social Progress index*. Podemos verificar que desta forma obtivemos índice F = 58,42 e P-Value = 0,000 demonstrando que há grande confiabilidade nos resultados.

Em seguida geramos novo gráfico boxplot apenas com as três regiões *AIBER*, *AVECO* e Outros. Podemos observar que tivemos aumento do índice F que foi de 58,42 com oito regiões para 83,97 e manteve-se o P-Value = 0,000 mantendo o grau de confiabilidade.

Por esta razão passaremos a gerar para as demais variáveis apenas boxplot com três regiões.

3.1 One-way ANOVA: Social Progress Index_1N versus Class 6R



One-way ANOVA: Social Progress Index_1N versus Class 6R

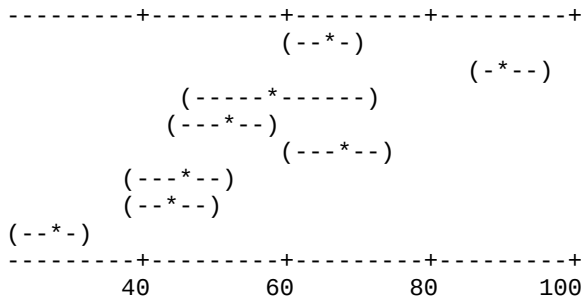
Source	DF	SS	MS	F	P
Class 6R	7	65481	9354	58,42	0,000
Error	124	19857	160		
Total	131	85338			

F = O GRAU QUE ESTA DISCRIMANDO AS REGIÕES.

S = 12,65 R-Sq = 76,73% R-Sq(adj) = 75,42%

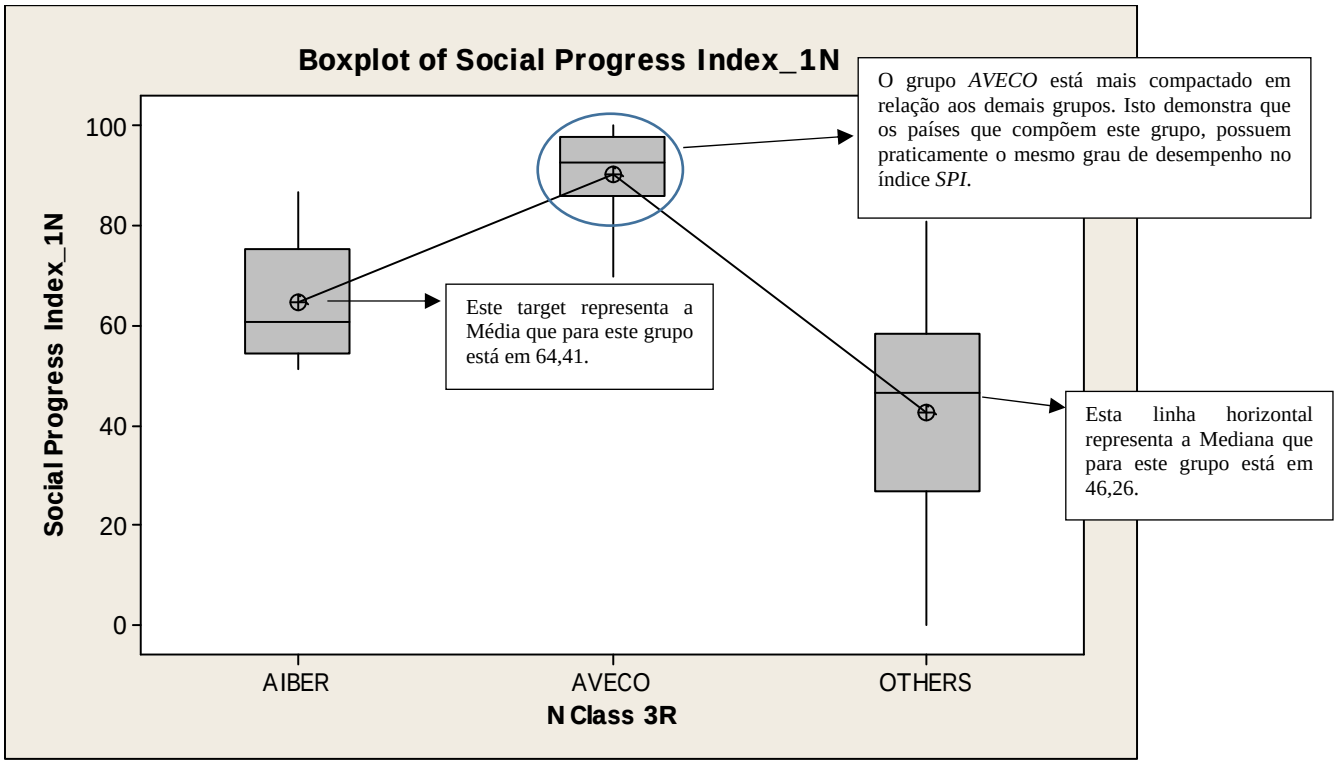
Level	N	Mean	StDev
AIBER	20	65,07	11,86
AVECO	27	90,30	8,70
CARLA	4	58,86	9,97
CWEAS	11	51,27	6,13
EMDEU	12	67,34	7,04
ESEAS	12	45,57	11,72
NSWUA	14	43,76	18,74
SUSAF	32	27,03	15,99

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev = 12,65

Boxplot of Social Progress Index_1N



3.1.1 One-way ANOVA: Social Progress Index_1N versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  48264  24132  83,97  0,000
Error       129  37074    287
Total       131  85338

S = 16,95   R-Sq = 56,56%   R-Sq(adj) = 55,88%
```

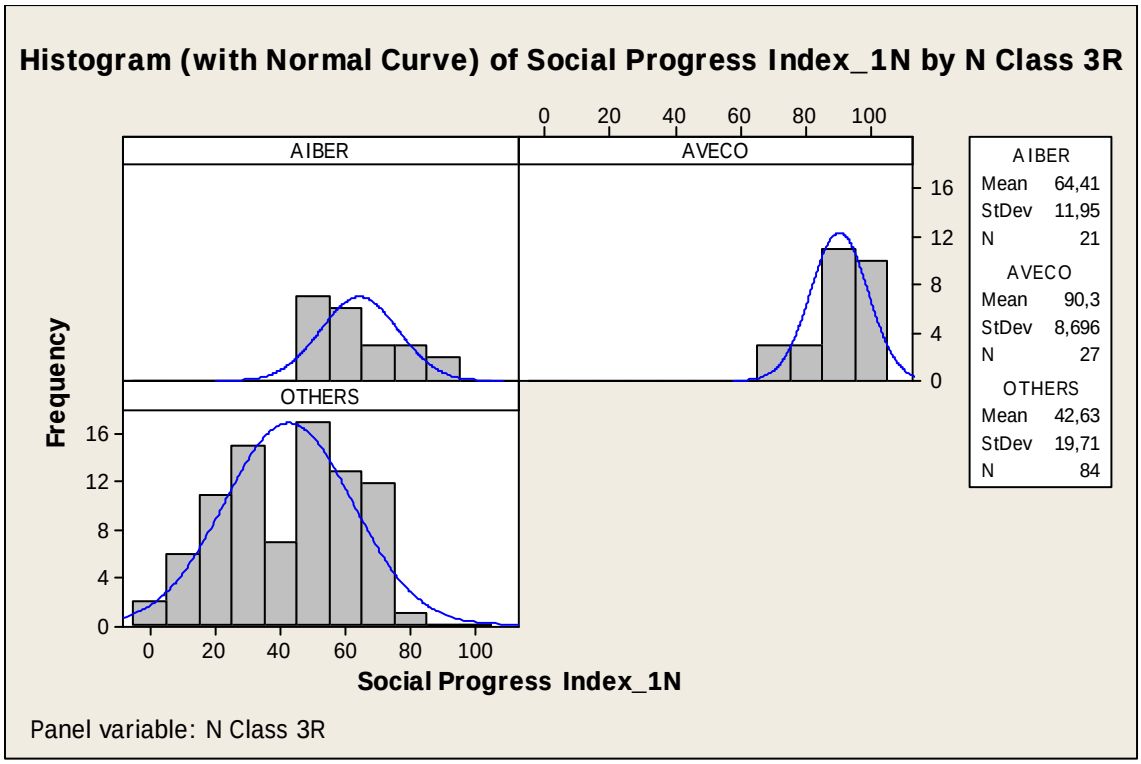
UTILIZAR F MAIS ALTO

F = Variação entre grupos / Variação dentro dos

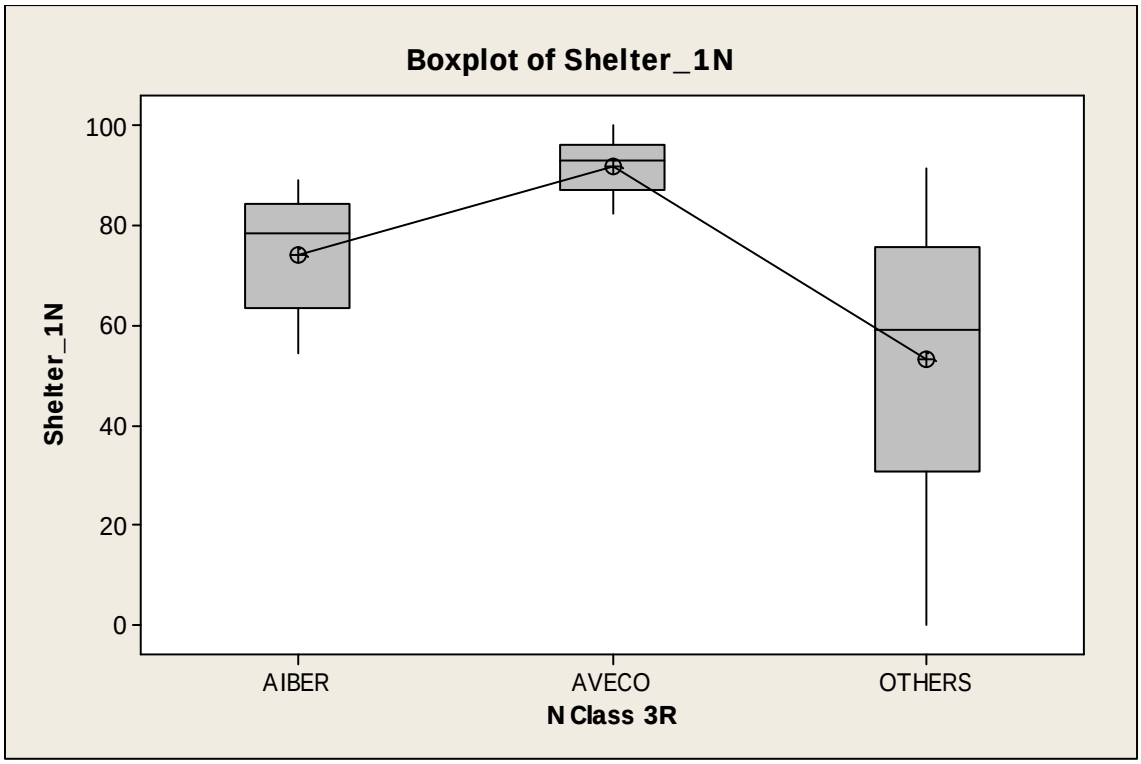
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  64,41  11,95             (----*----)
AVECO   27  90,30   8,70                                (---*----)
OTHERS  84  42,63  19,71  (-*--)
                          ----+---------+---------+---------+-----
                             45        60        75        90

Pooled StDev = 16,95
```

Análise: Para a Variável Social Progress Index, que é o índice que identifica os pontos fortes e fracos dos países a partir de um conjunto de informações e análises conduzidas para cada um deles. Podemos verificar claramente pelo boxplot acima que o grupo *AVECO* encontra-se com melhor desempenho em relação ao grupo *AIBER* e Outros.

Verificamos também que *AVECO* mantem seus países num grupo mais compactado do gráfico, demonstrando que há praticamente o mesmo grau de desempenho entre seus integrantes. Nos grupos *AIBER* e Outros há uma maior dispersão entre os países.



3.2 One-way ANOVA: Shelter_1N versus N Class 3R



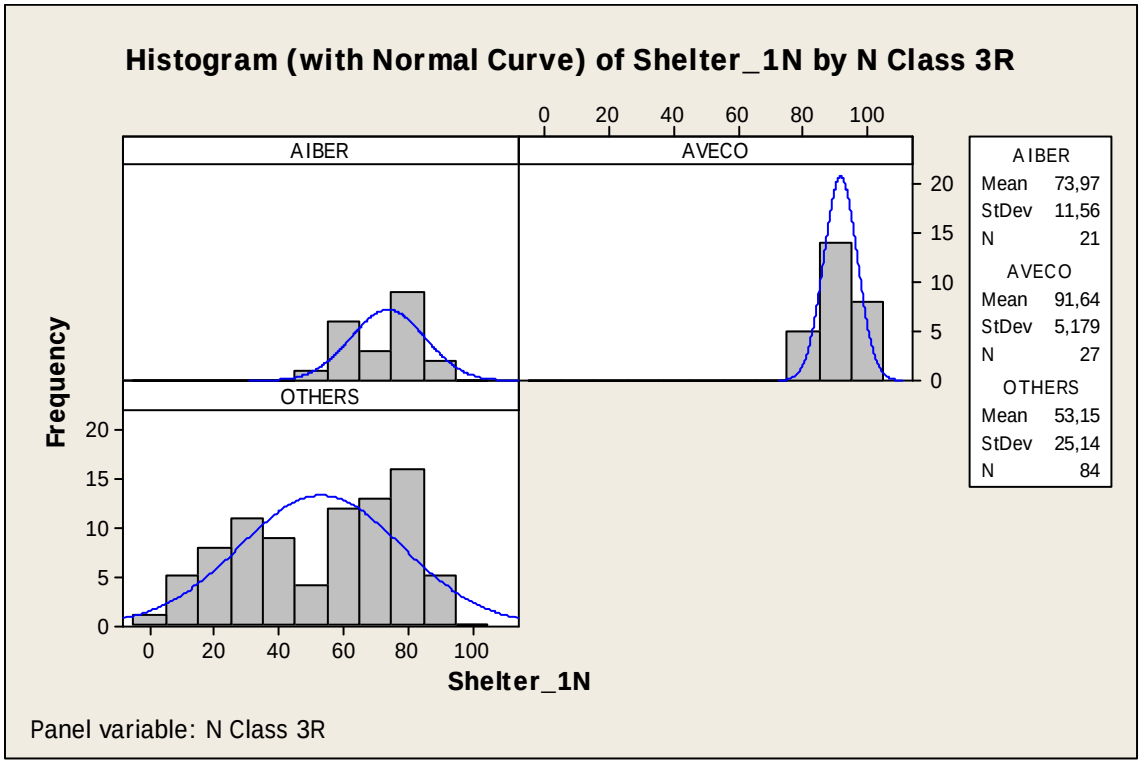
Source	DF	SS	MS	F	P
N Class 3R	2	32589	16295	37,64	0,000
Error	129	55838	433		
Total	131	88427			

S = 20,81 R-Sq = 36,85% R-Sq(adj) = 35,88%

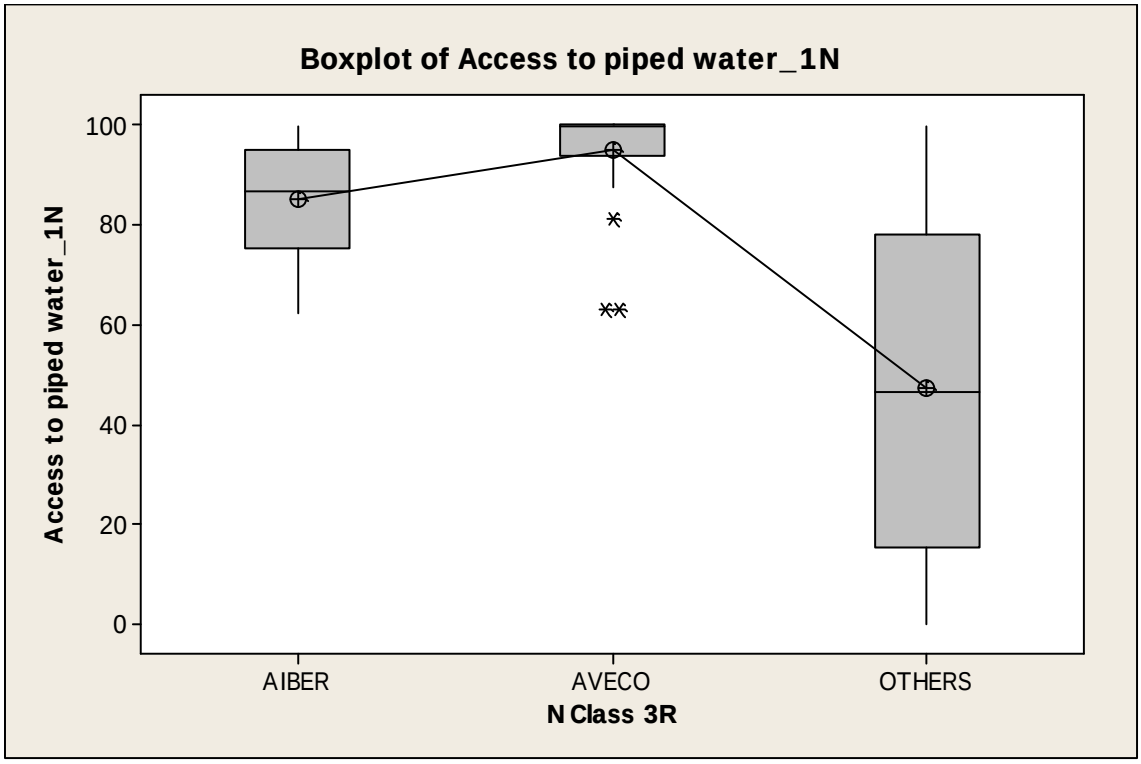
Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	73,97	11,56	(-----*-----)
AVECO	27	91,64	5,18	(----*----)
OTHERS	84	53,15	25,14	(--*--)

60 75 90 105

Pooled StDev = 20,81



3.3 One-way ANOVA: Access to piped water_1N versus N Class 3R



Source	DF	SS	MS	F	P
N Class 3R	2	58625	29313	37,63	0,000
Error	129	100489	779		
Total	131	159114			

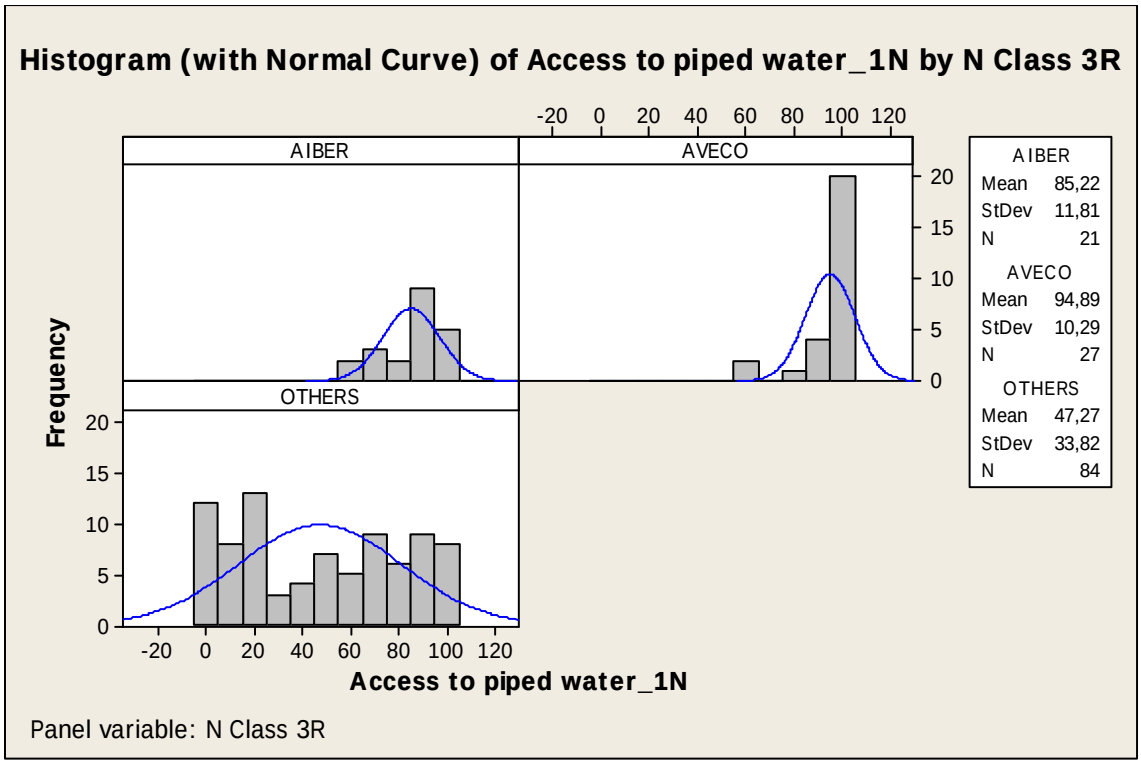
S = 27,91 R-Sq = 36,84% R-Sq(adj) = 35,87%

Level	N	Mean	StDev
AIBER	21	85,22	11,81
AVECO	27	94,89	10,29
OTHERS	84	47,27	33,82

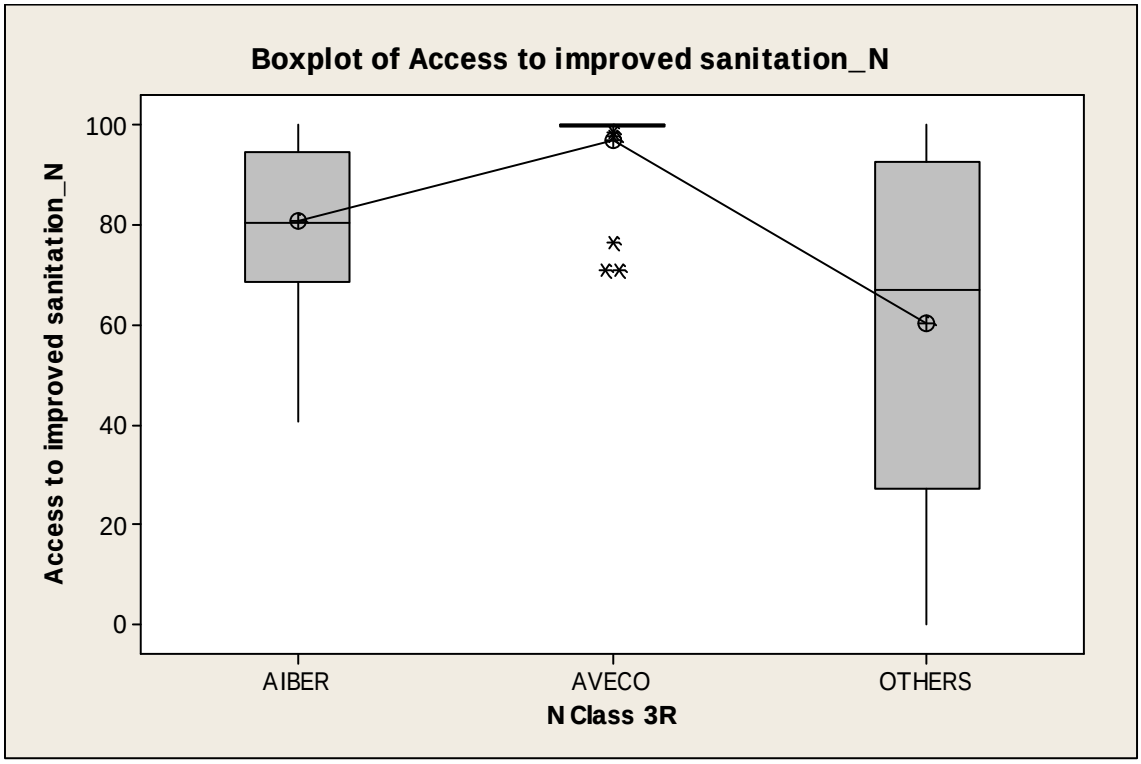
Individual 95% CIs For Mean Based on Pooled StDev

```
        (-----*-----)
             (----*-----)
(--*--)
---------+---------+---------+---------+
        60        80       100       120
```

Pooled StDev = 27,91



3.4 One-way ANOVA: Access to improved sanitation_N versus N Class 3R



One-way ANOVA: Access to improved sanitation_N versus N Class 3R

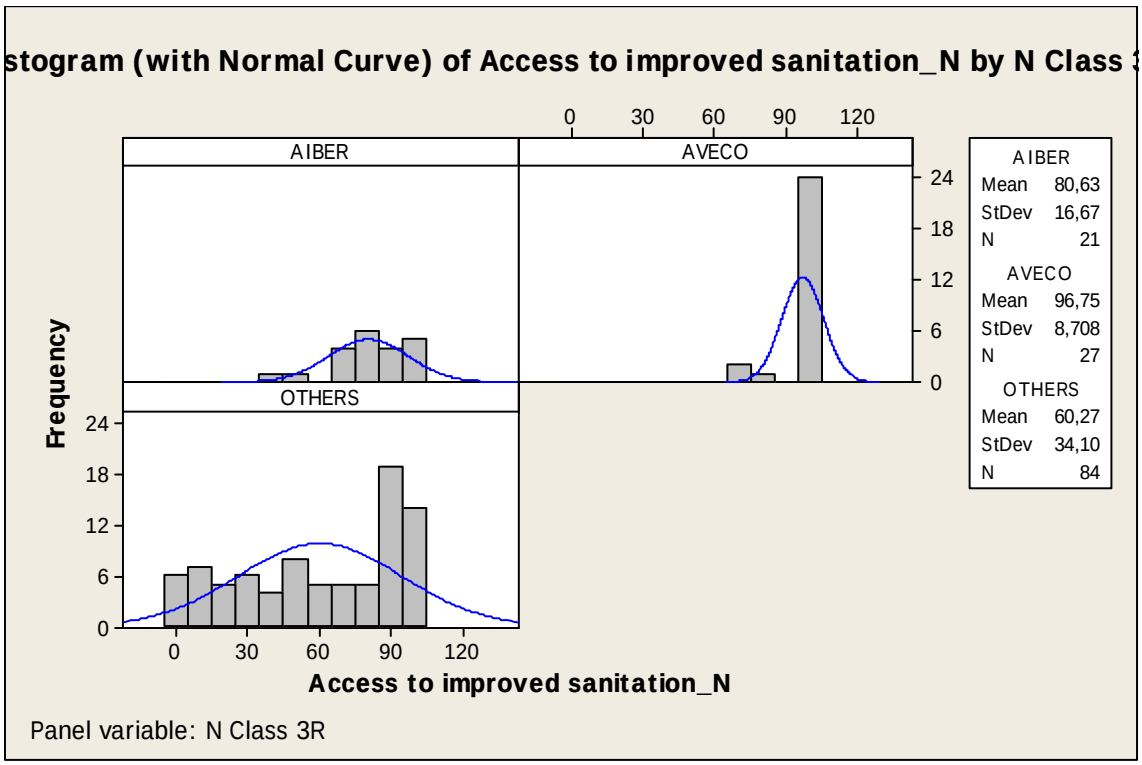
Source	DF	SS	MS	F	P
N Class 3R	2	29518	14759	18,30	0,000
Error	129	104030	806		
Total	131	133548			

S = 28,40 R-Sq = 22,10% R-Sq(adj) = 20,90%

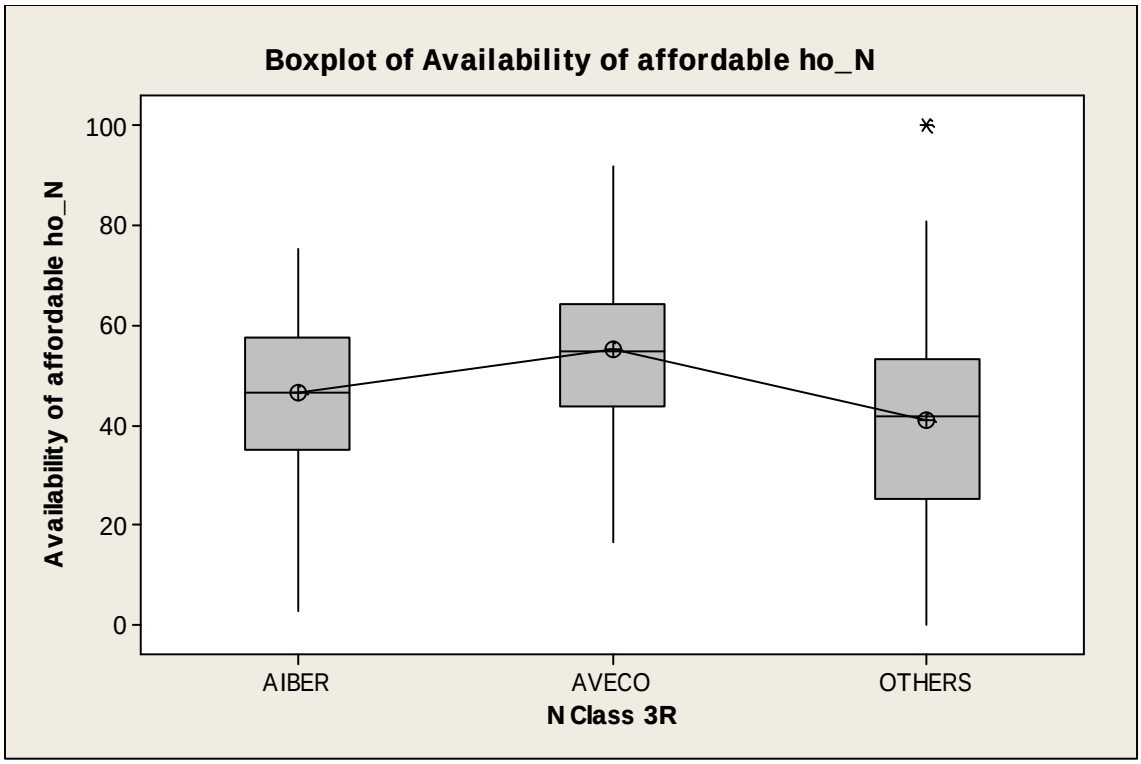
Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	80,63	16,67	(-------*-------)
AVECO	27	96,75	8,71	(-------*------)
OTHERS	84	60,27	34,10	(---*---)

60 75 90 105

Pooled StDev = 28,40



3.5 One-way ANOVA: Availability of affordable ho_N versus N Class 3R



Source	DF	SS	MS	F	P
N Class 3R	2	4253	2126	6,04	0,003
Error	129	45425	352		
Total	131	49677			

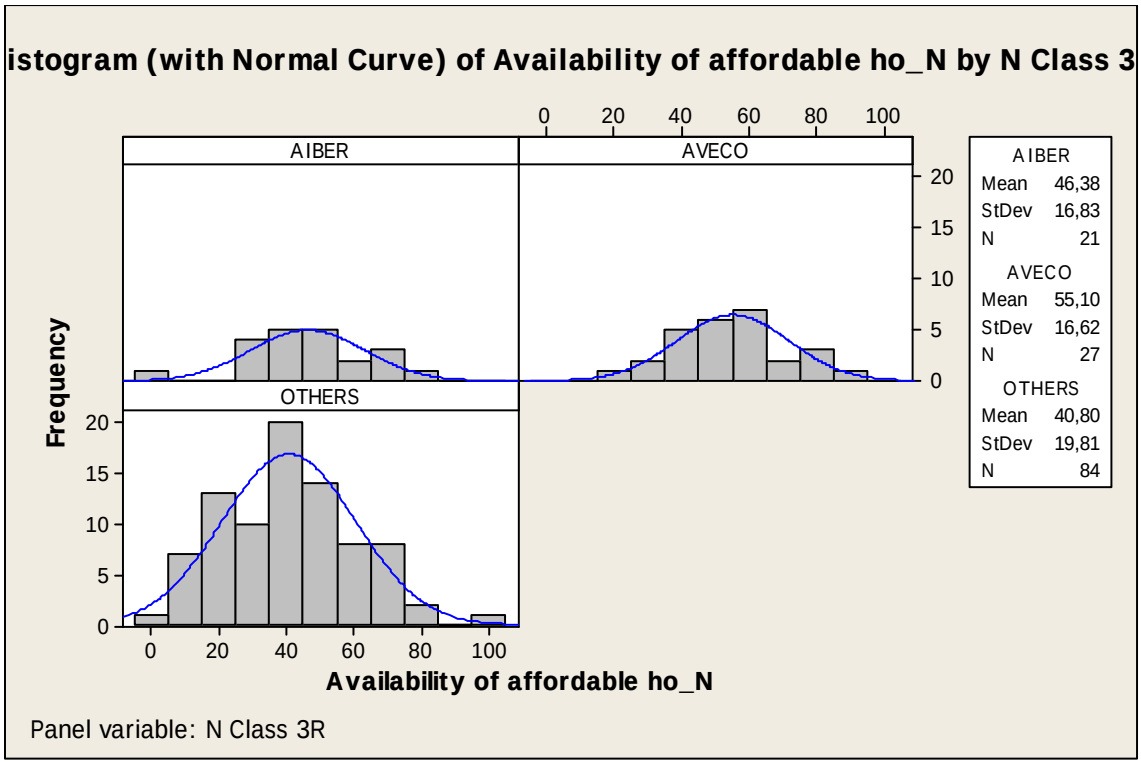
S = 18,77 R-Sq = 8,56% R-Sq(adj) = 7,14%

Level	N	Mean	StDev
AIBER	21	46,38	16,83
AVECO	27	55,10	16,62
OTHERS	84	40,80	19,81

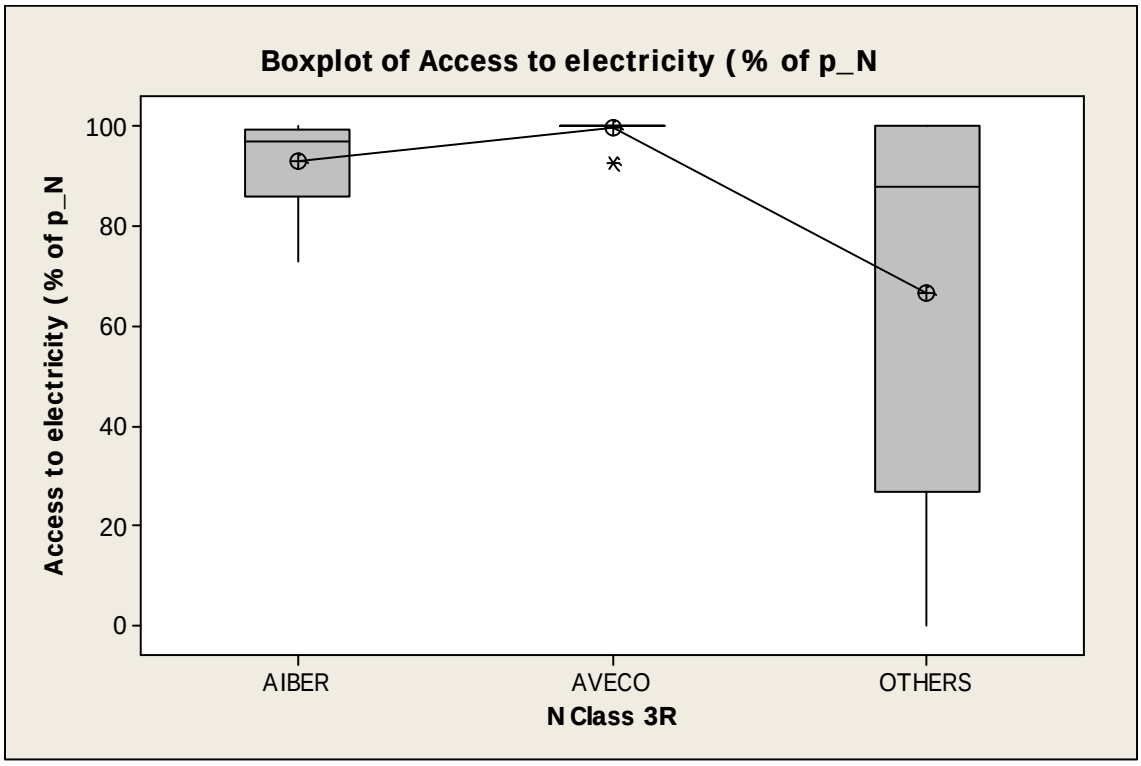
Individual 95% CIs For Mean Based on Pooled StDev

```
-------+---------+---------+---------+--
  (----------*----------)
                (---------*---------)
(----*-----)
-------+---------+---------+---------+--
    42,0      49,0      56,0      63,0
```

Pooled StDev = 18,77



3.6 One-way ANOVA: Access to electricity (% of p_N versus N Class 3R



3.9 One-way ANOVA: Access to electricity (% of p_N versus N Class 3R

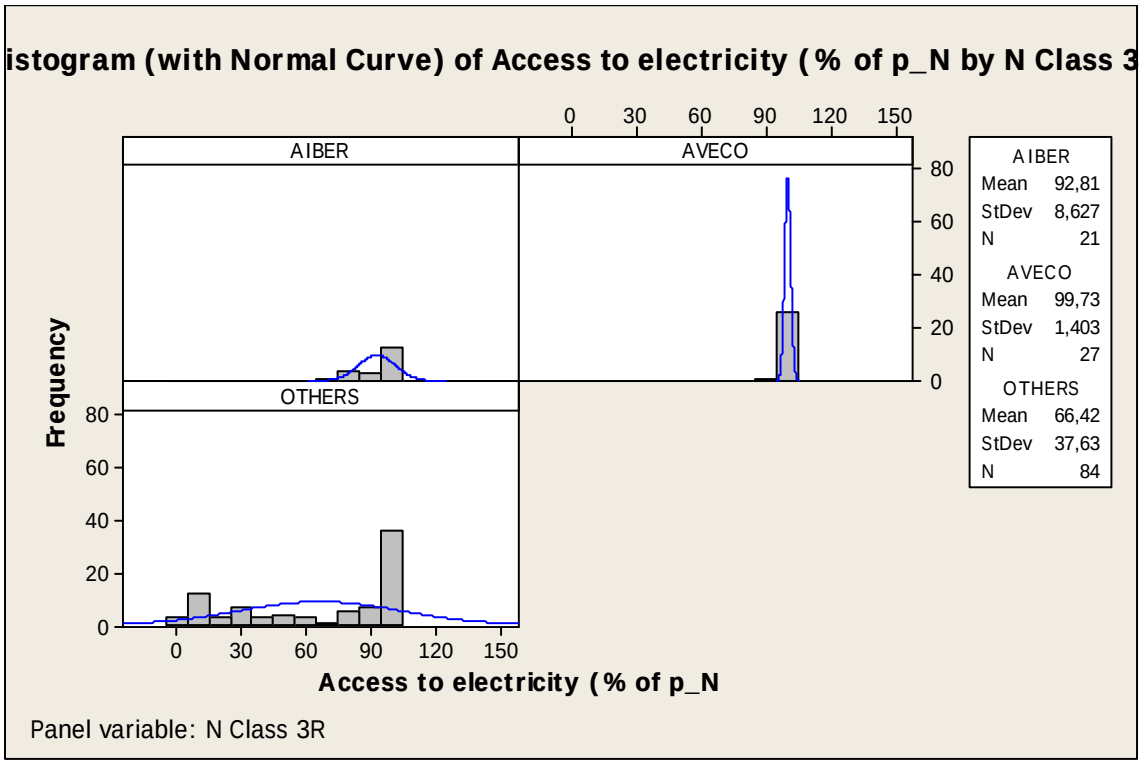
Source	DF	SS	MS	F	P
N Class 3R	2	28577	14289	15,48	0,000
Error	129	119046	923		
Total	131	147624			

S = 30,38 R-Sq = 19,36% R-Sq(adj) = 18,11%

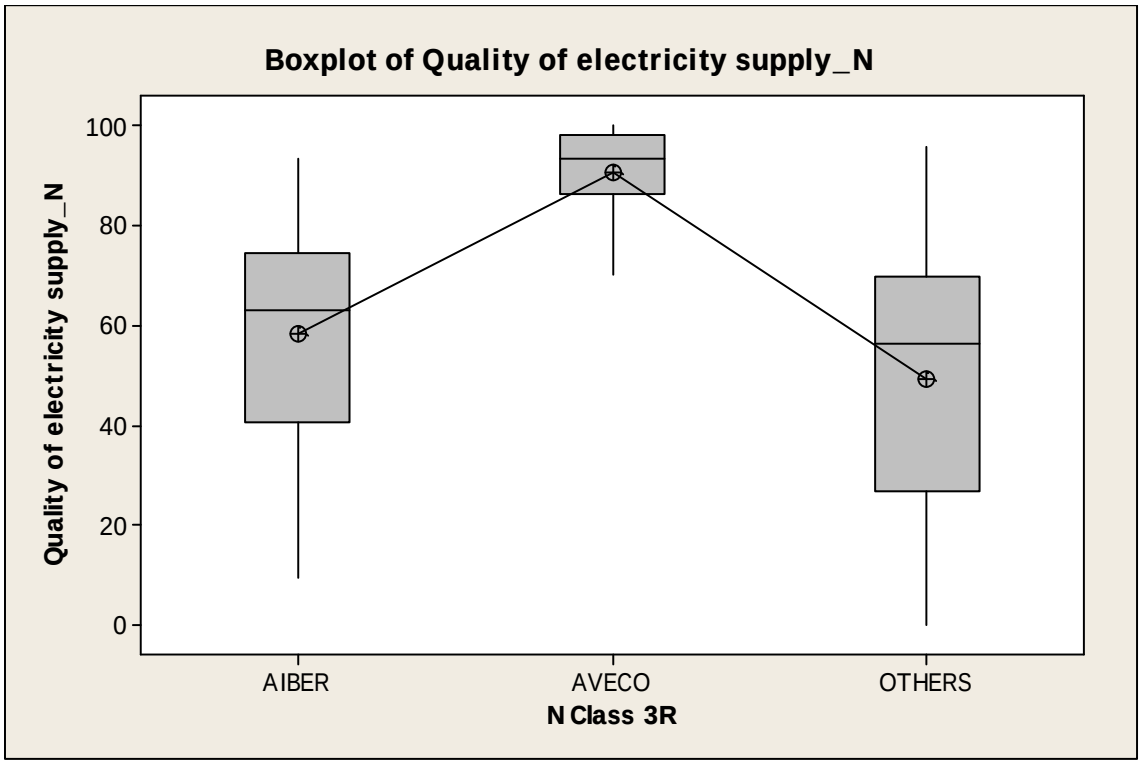
Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	92,81	8,63	(--------*--------)
AVECO	27	99,73	1,40	(------*-------)
OTHERS	84	66,42	37,63	(---*----)

60 75 90 105

Pooled StDev = 30,38

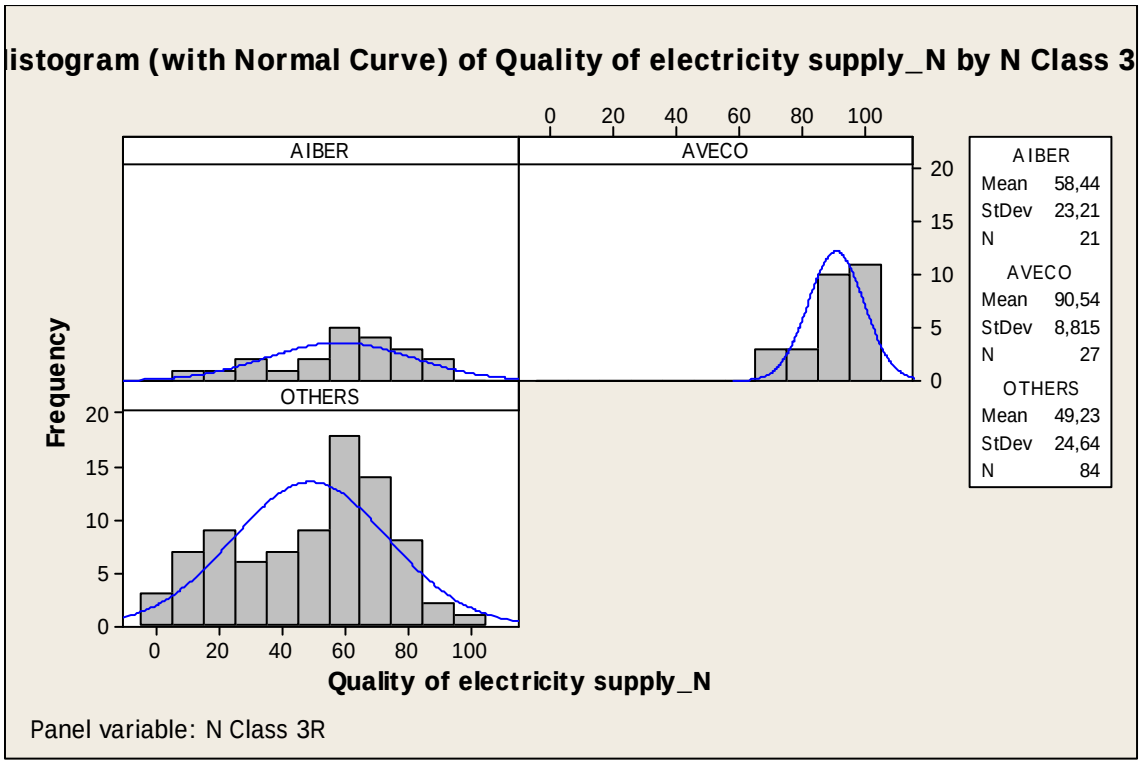


3.7 One-way ANOVA: Quality of electricity supply_N versus N Class 3R

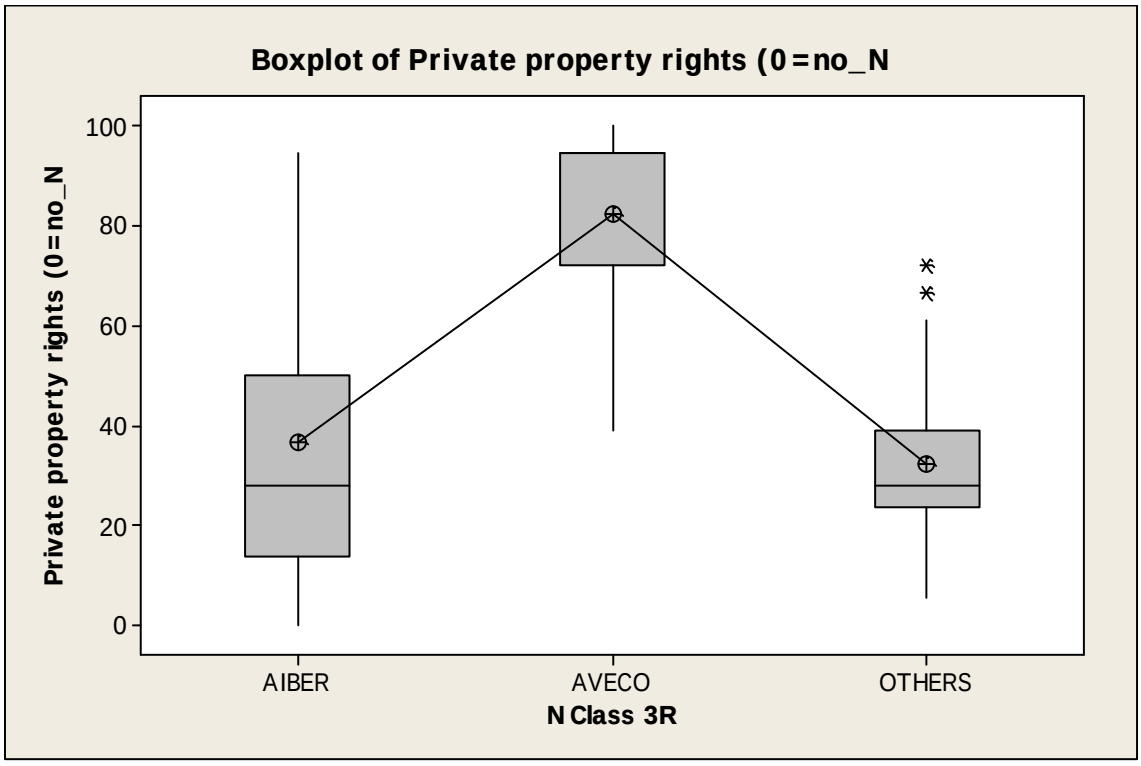


One-way ANOVA: Quality of electricity supply_N versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  34876  17438  35,60  0,000
Error       129  63191    490
Total       131  98067

S = 22,13   R-Sq = 35,56%   R-Sq(adj) = 34,56%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  58,44  23,21      (-----*-----)
AVECO   27  90,54   8,82                           (----*-----)
OTHERS  84  49,23  24,64   (--*--)
                           +---------+---------+---------+---------
                          45        60        75        90

Pooled StDev = 22,13
```



3.8 One-way ANOVA: Private property rights (0=no_N versus N Class 3R



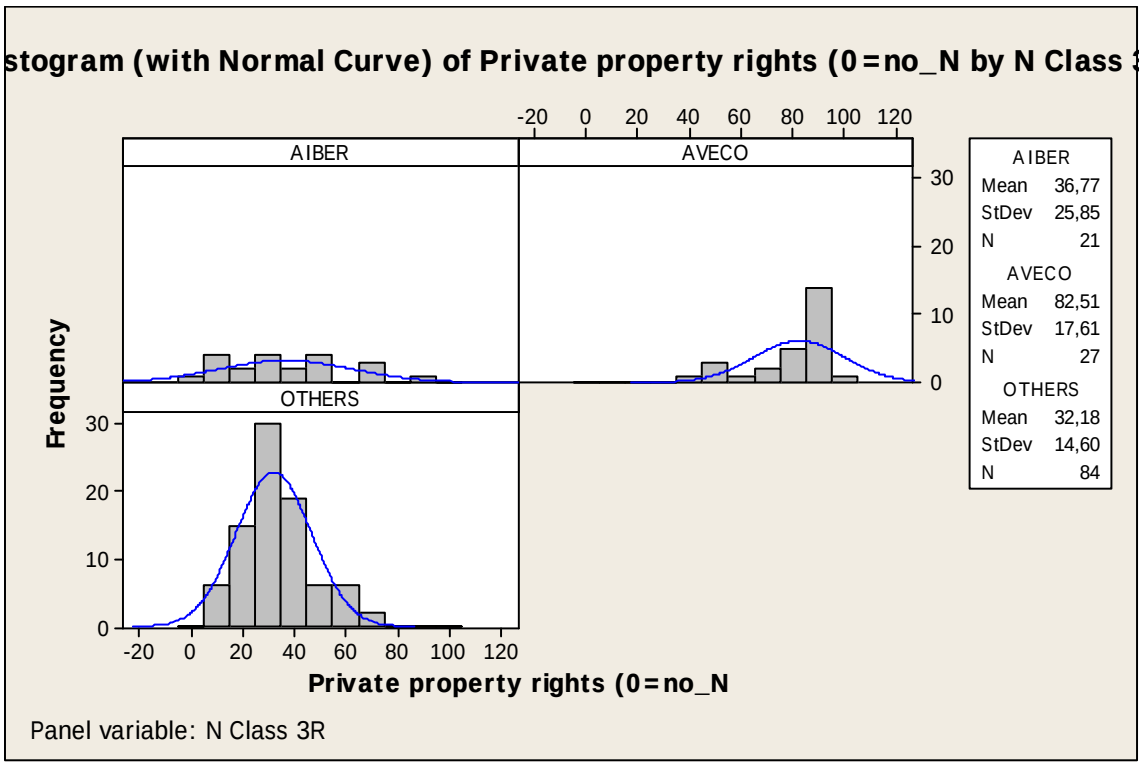
One-way ANOVA: Private property rights (0=no_N versus N Class 3R

Source	DF	SS	MS	F	P
N Class 3R	2	52784	26392	87,00	0,000
Error	129	39132	303		
Total	131	91916			

S = 17,42 R-Sq = 57,43% R-Sq(adj) = 56,77%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	36,77	25,85	(----*----)
AVECO	27	82,51	17,61	(----*---)
OTHERS	84	32,18	14,60	(-*-)
				32 48 64 80

Pooled StDev = 17,42



3.9 One-way ANOVA: IDH - 2013_1_N versus N Class 3R

Source	DF	SS	MS	F	P
N Class 3R	2	35137	17568	43,46	0,000
Error	129	52153	404		
Total	131	87290			

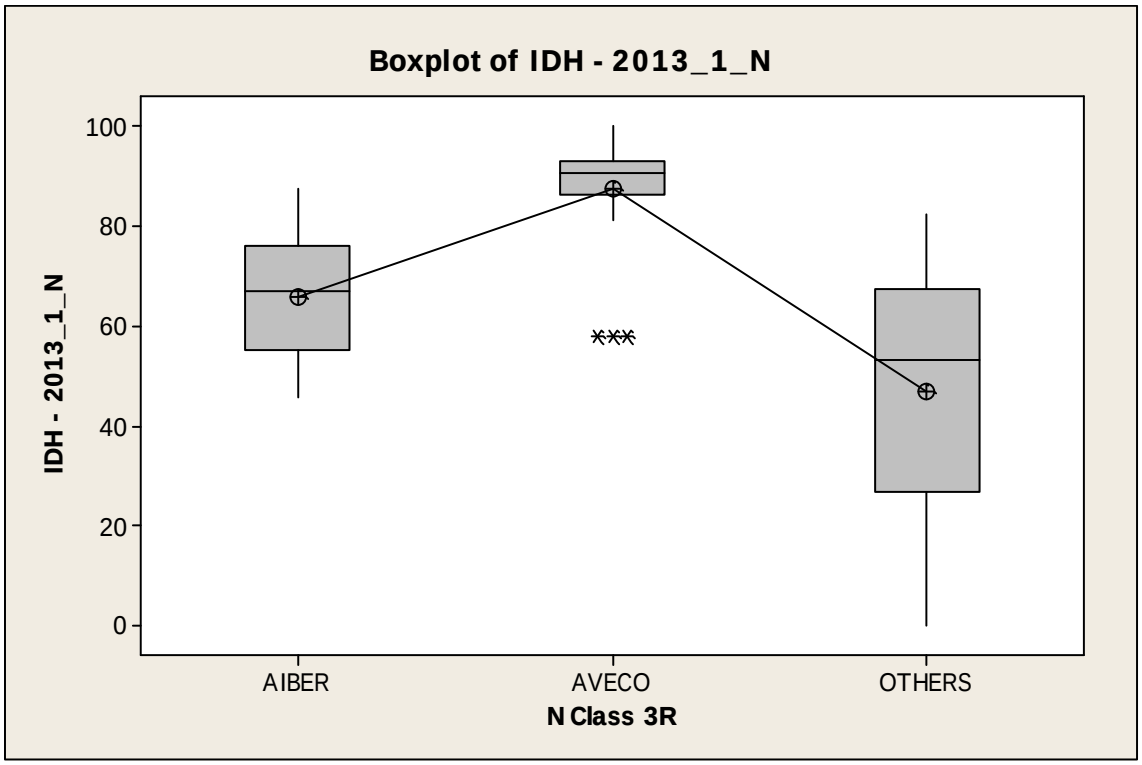
S = 20,11 R-Sq = 40,25% R-Sq(adj) = 39,33%

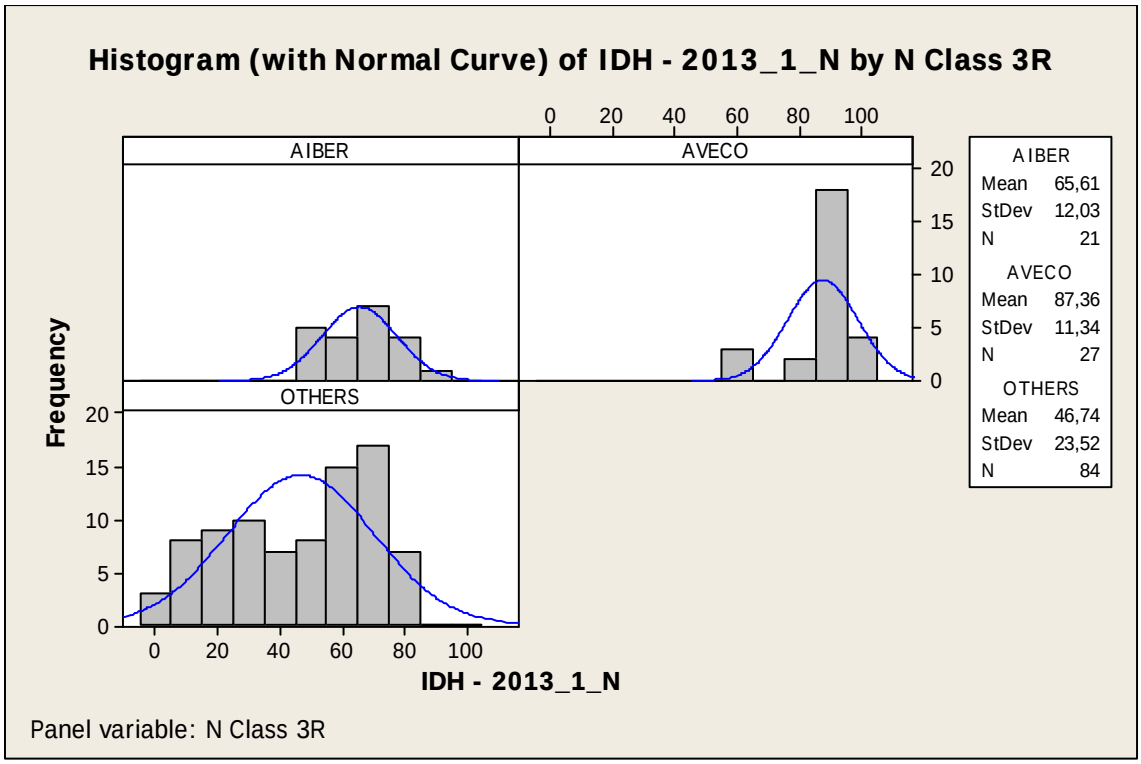
Level	N	Mean	StDev
AIBER	21	65,61	12,03
AVECO	27	87,36	11,34
OTHERS	84	46,74	23,52

Individual 95% CIs For Mean Based on Pooled StDev

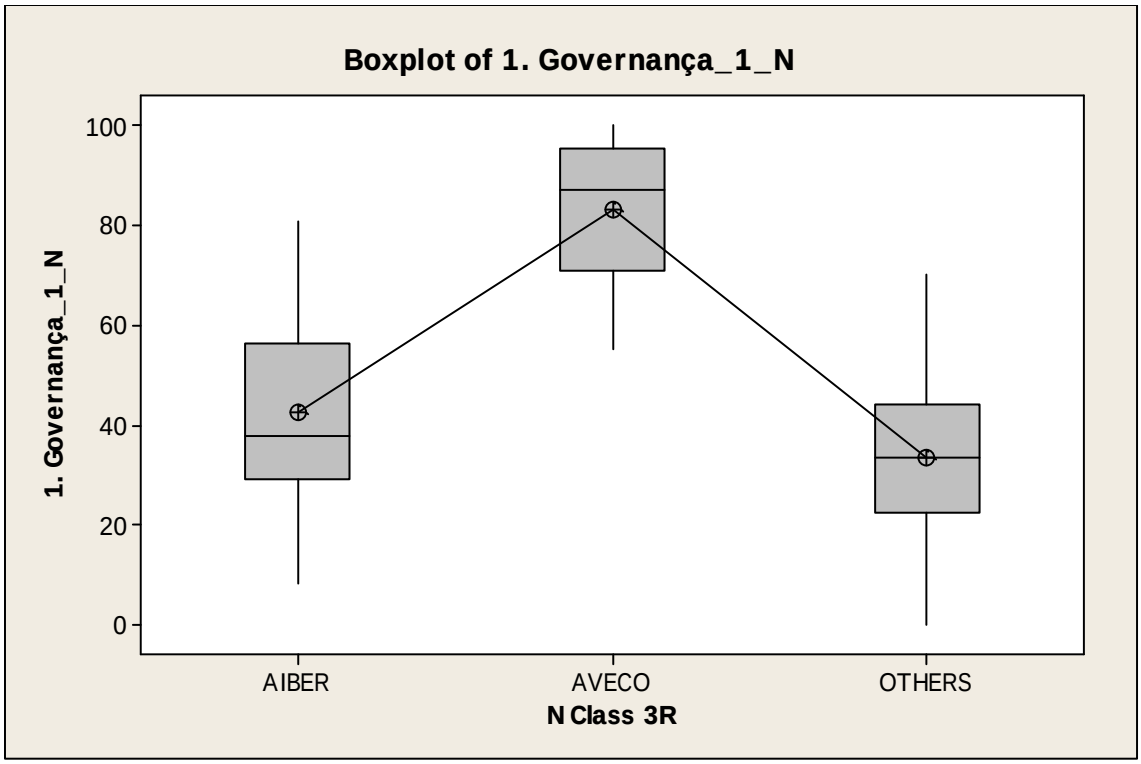
```
--+---------+---------+---------+-------
          (-----*-----)
                              (----*----)
(--*--)
--+---------+---------+---------+-------
 45        60        75        90
```

Pooled StDev = 20,11




Histogram (with Normal Curve) of IDH - 2013_1_N by N Class 3R
AIBER
AVECO
OTHERS
Frequency
IDH - 2013_1_N
AIBER
Mean 65,61
StDev 12,03
N 21
AVECO
Mean 87,36
StDev 11,34
N 27
OTHERS
Mean 46,74
StDev 23,52
N 84
Panel variable: N Class 3R

3.10 One-way ANOVA: 1. Governança_1_N versus N Class 3R



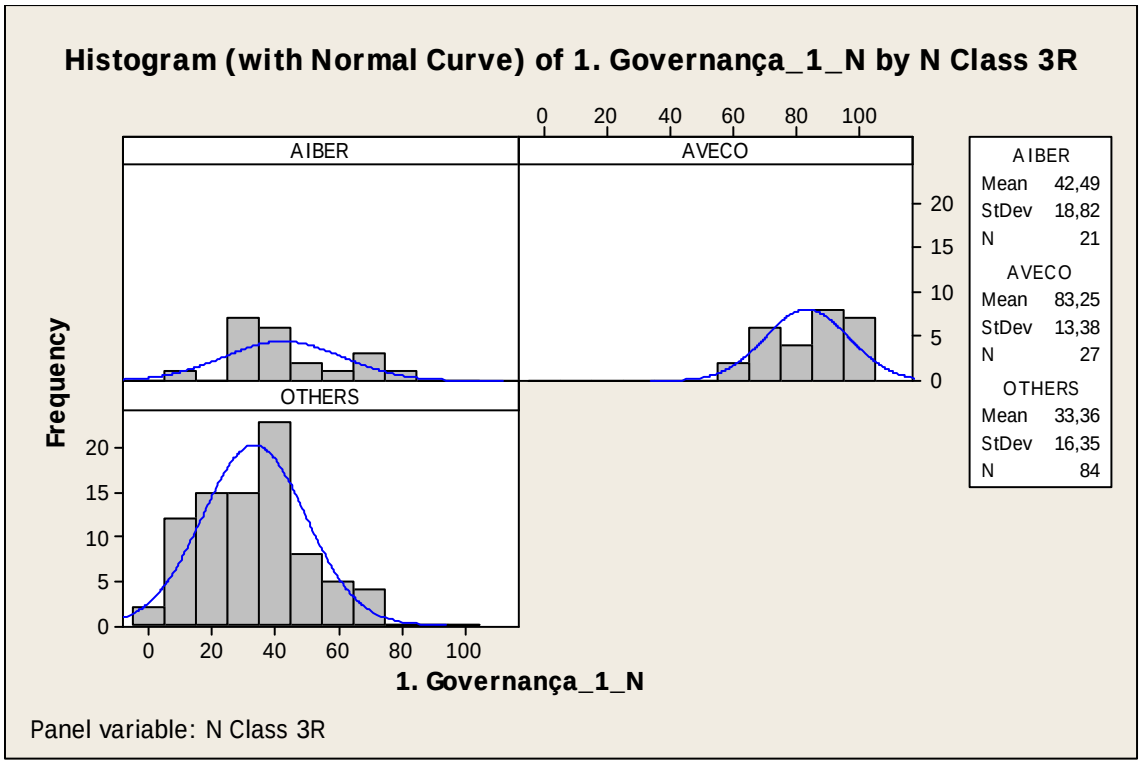
One-way ANOVA: 1. Governança_1_N versus N Class 3R

Source	DF	SS	MS	F	P
N Class 3R	2	51020	25510	97,02	0,000
Error	129	33918	263		
Total	131	84938			

S = 16,22 R-Sq = 60,07% R-Sq(adj) = 59,45%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	42,49	18,82	(---*----)
AVECO	27	83,25	13,38	(----*---)
OTHERS	84	33,36	16,35	(-*--)
				30 45 60 75

Pooled StDev = 16,22



3.11 One-way ANOVA: EPI Score_1_N versus N Class 3R

Source	DF	SS	MS	F	P
N Class 3R	2	38873	19436	69,68	0,000
Error	129	35981	279		
Total	131	74854			

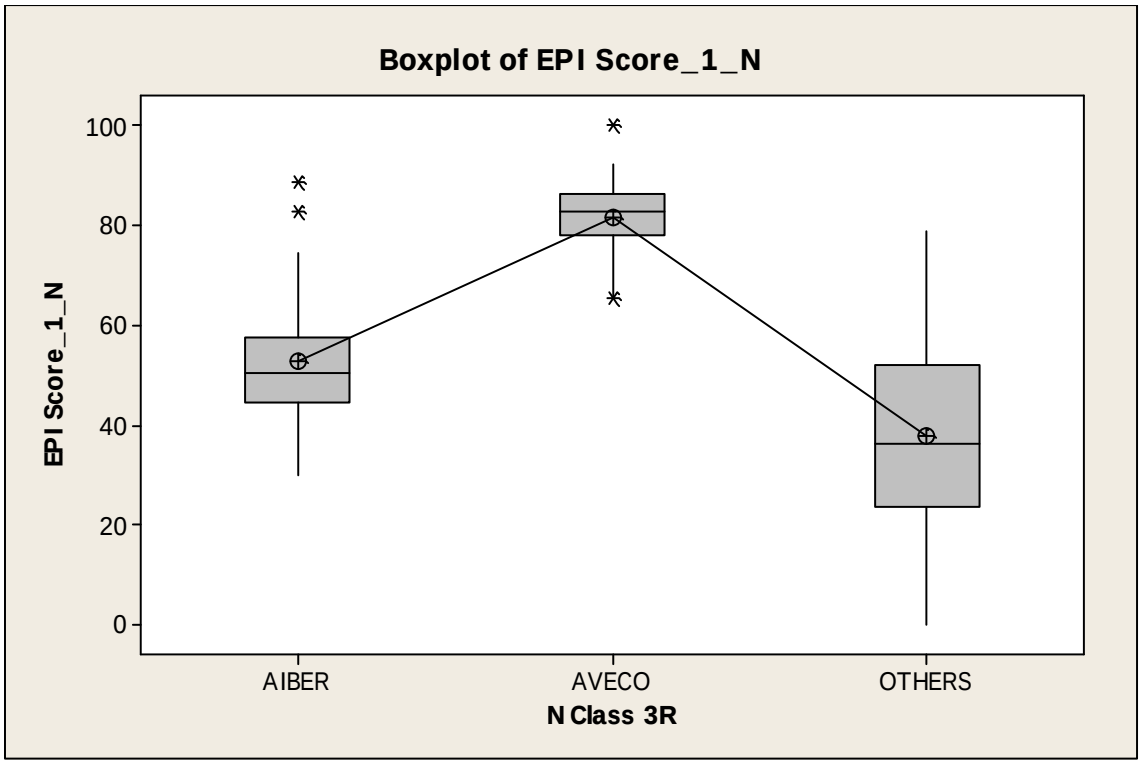
S = 16,70 R-Sq = 51,93% R-Sq(adj) = 51,19%

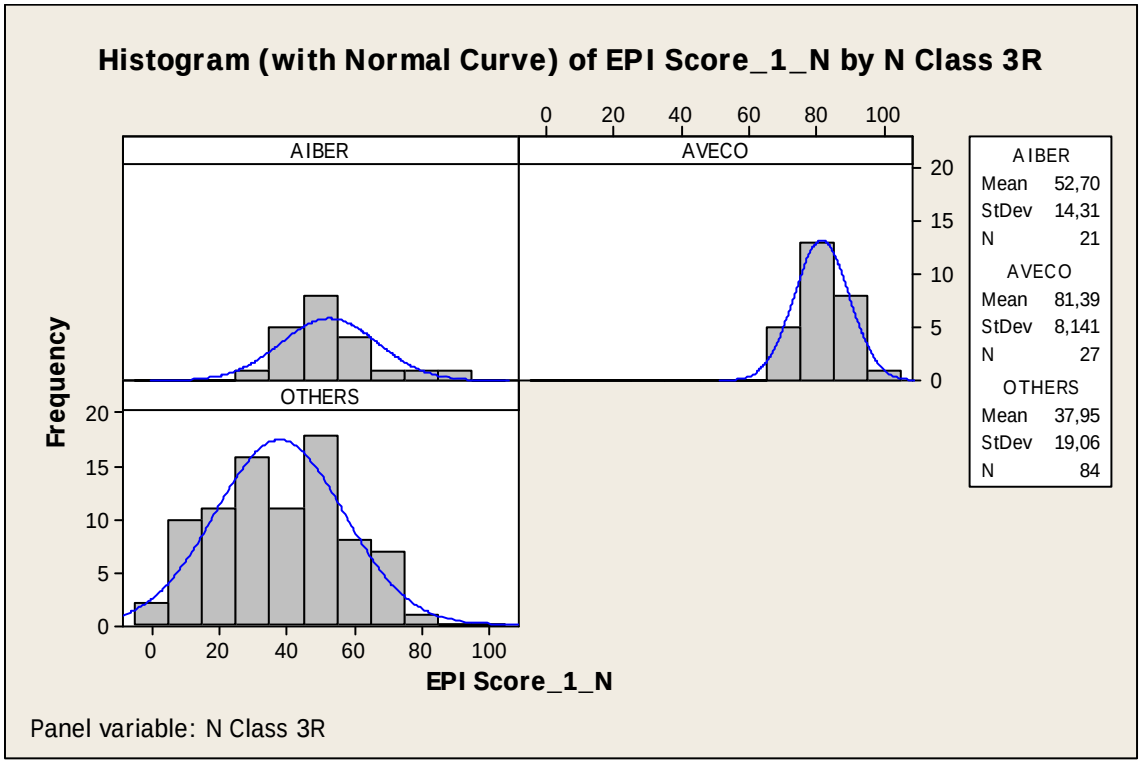
Level	N	Mean	StDev
AIBER	21	52,70	14,31
AVECO	27	81,39	8,14
OTHERS	84	37,95	19,06

Individual 95% CIs For Mean Based on Pooled StDev

```
-------+---------+---------+---------+--
       (----*----)
                              (---*----)
(-*--)
-------+---------+---------+---------+--
      45        60        75        90
```

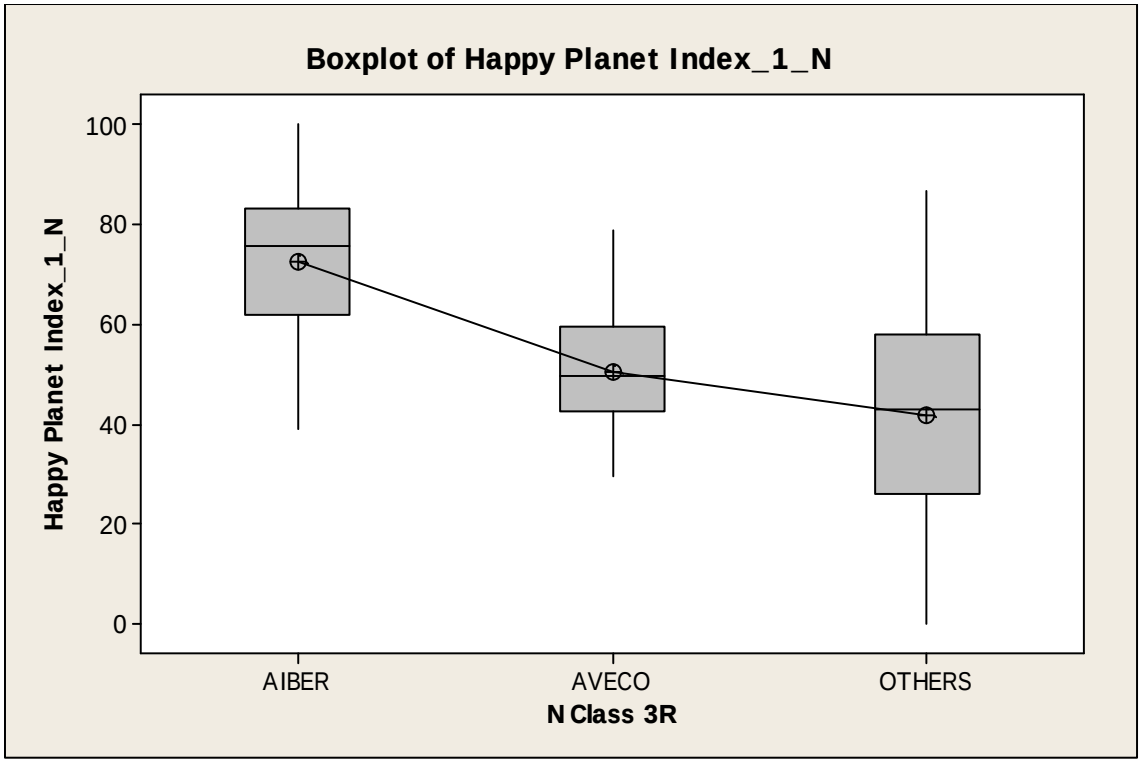
Pooled StDev = 16,70

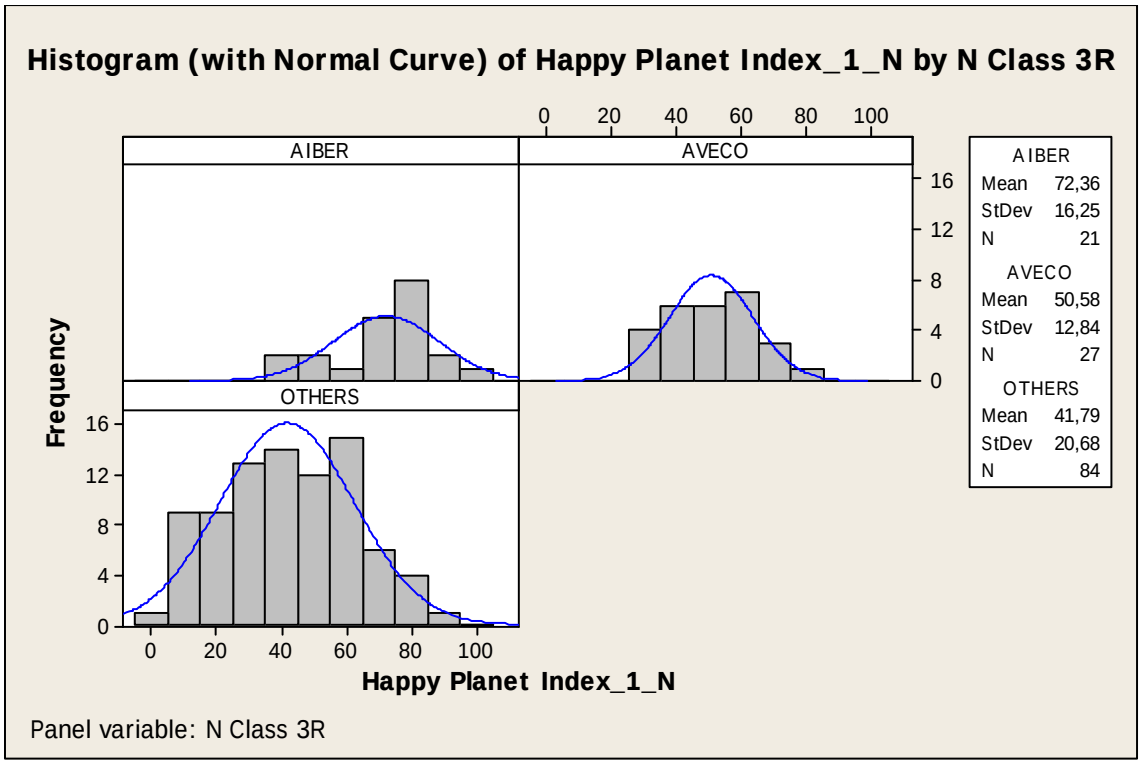

Boxplot of EPI Score_1_N
EPI Score_1_N
100
80
60
40
20
0
AIBER
AVECO
OTHERS
N Class 3R


Histogram (with Normal Curve) of EPI Score_1_N by N Class 3R
AIBER
AVECO
OTHERS
Frequency
EPI Score_1_N
AIBER
Mean 52,70
StDev 14,31
N 21
AVECO
Mean 81,39
StDev 8,141
N 27
OTHERS
Mean 37,95
StDev 19,06
N 84
Panel variable: N Class 3R

3.12 One-way ANOVA: Happy Planet Index_1_N versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  15846  7923  22,69  0,000
Error       129  45048   349
Total       131  60894

S = 18,69   R-Sq = 26,02%   R-Sq(adj) = 24,88%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72,36  16,25                        (-----*------)
AVECO   27  50,58  12,84       (-----*-----)
OTHERS  84  41,79  20,68  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18,69
```





3.13 One-way ANOVA: GINNI P versus N Class 3R

Source	DF	SS	MS	F	P
N Class 3R	2	16078	8039	24,45	0,000
Error	129	42416	329		
Total	131	58494			

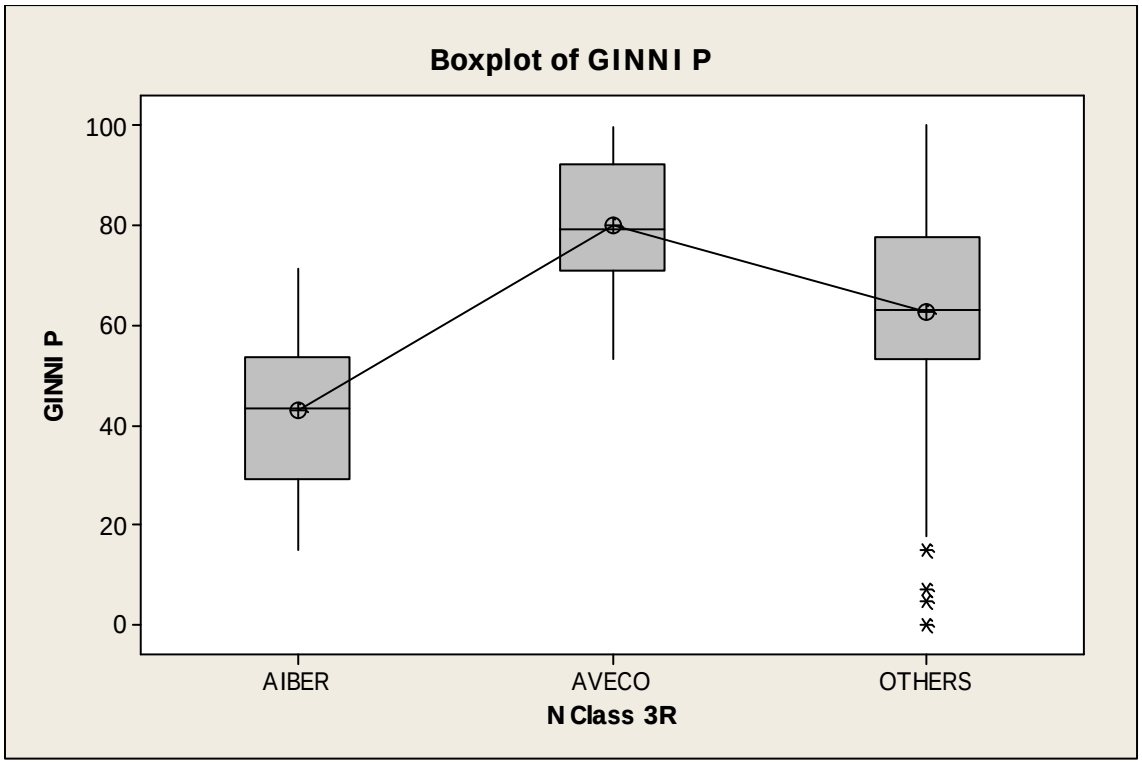
S = 18,13 R-Sq = 27,49% R-Sq(adj) = 26,36%

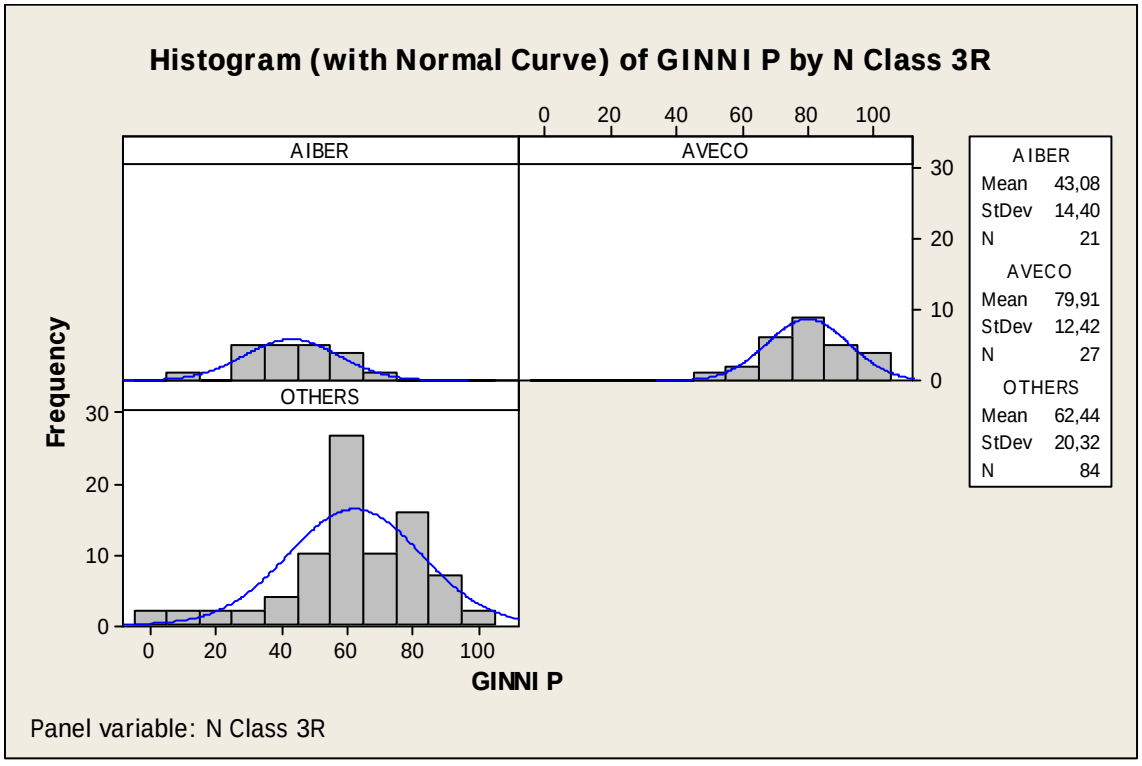
Level	N	Mean	StDev
AIBER	21	43,08	14,40
AVECO	27	79,91	12,42
OTHERS	84	62,44	20,32

Individual 95% CIs For Mean Based on Pooled StDev

```
------+---------+---------+---------+---
(----*----)
                          (---*----)
               (--*-)
------+---------+---------+---------+---
     45        60        75        90
```

Pooled StDev = 18,13


Boxplot of GINNI P
GINNI P
100
80
60
40
20
0
AIBER
AVECO
OTHERS
N Class 3R

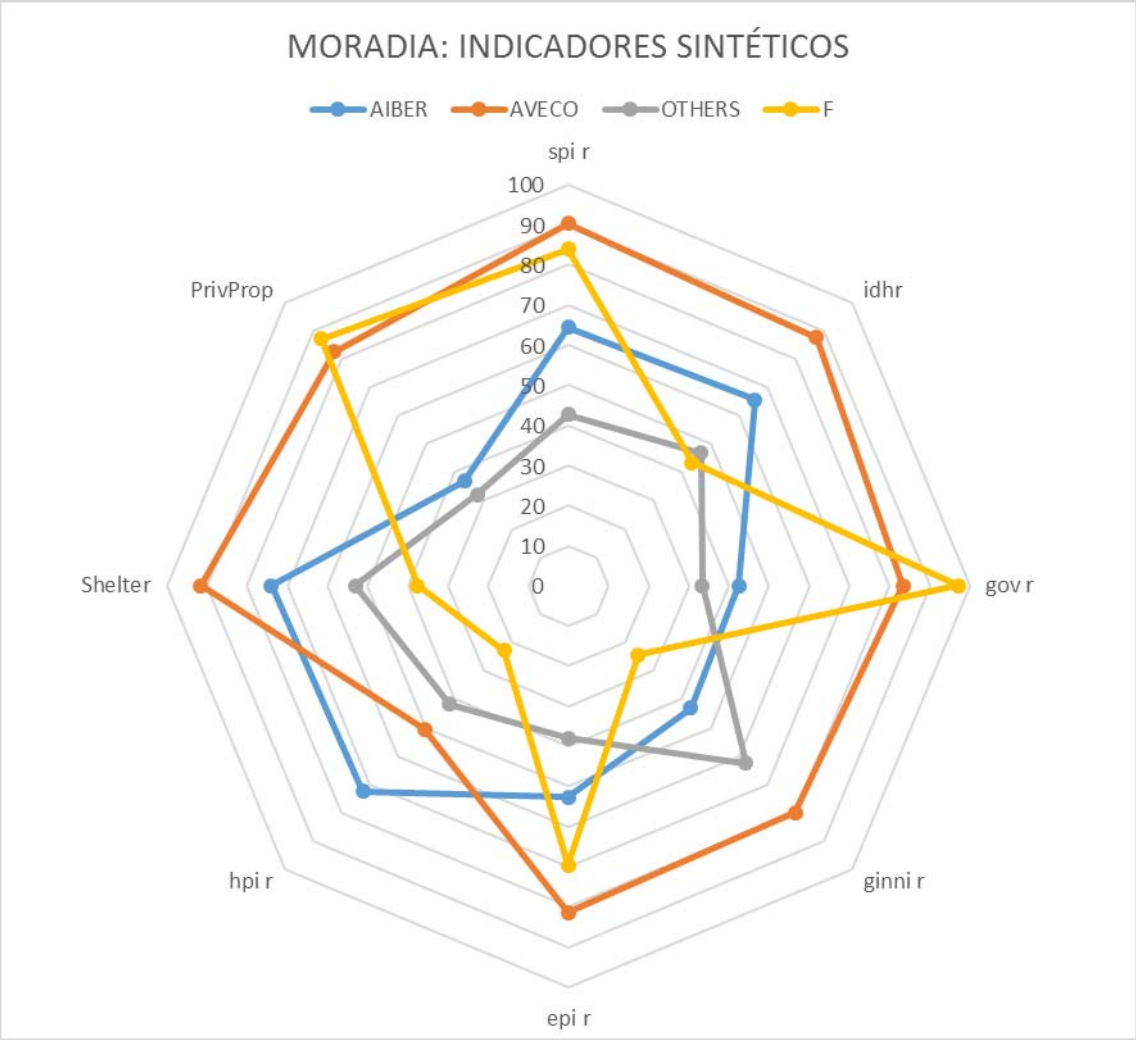

Histogram (with Normal Curve) of GINNI P by N Class 3R
AIBER
AVECO
OTHERS
Frequency
GINNI P
0
20
40
60
80
100
30
20
10
0
AIBER
Mean 43,08
StDev 14,40
N 21
AVECO
Mean 79,91
StDev 12,42
N 27
OTHERS
Mean 62,44
StDev 20,32
N 84
Panel variable: N Class 3R

GRÁFICOS DE RADAR: Comparando 3 Regiões

TEMA: MORADIA

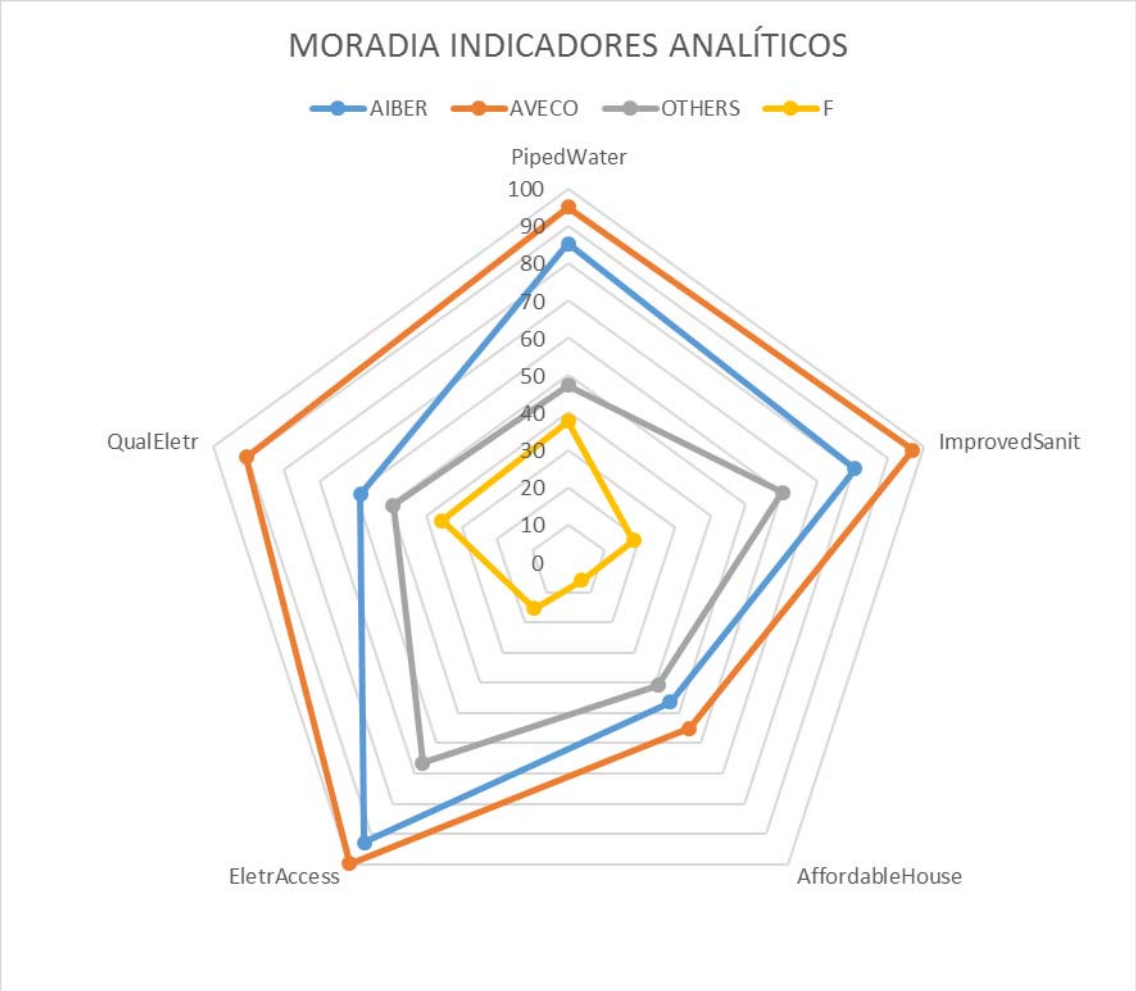
MÉDIAS e F dos INDICADORES SINTÉTICOS

Regiões	spi r	idhr	gov r	ginni r	epi r	hpi r	Shelter	PrivProp
AIBER	64,41	65,61	42,49	43,08	52,70	72,36	73,97	36,77
AVECO	90,3	87,36	83,25	79,91	81,39	50,58	91,64	82,51
OTHERS	42,63	43,46	33,36	62,44	37,95	41,79	53,15	32,18
F	83,97	43,46	97,02	24,45	69,68	22,69	37,64	87,00



MÉDIAS e F dos INDICADORS ANALÍTICOS

Regiões	PipedWater	ImprovedSanit	AffordableHouse	EletrAccess	QualEletr
AIBER	85,22	80,63	46,38	92,81	58,44
AVECO	94,89	96,75	55,1	99,73	90,54
OTHERS	47,27	60,27	40,8	66,42	49,23
F	37,63	18,3	6,04	15,48	35,6



CAPITULO 6 – AMOSTRAGEM

1 – INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma pesquisa por amostragem de dados dimensionadores do nível de Moradia, representado pelos índices SPI, Governança e Private Property Index de diversos países do mundo. O software estatístico utilizado é o **MINITAB.**

1.1 – Países analisados

Este trabalho consiste em fazer análises com amostragens dos indicadores SPI, Governança e Private Property Index na região em que 132 países estão inseridos, conforme o Quadro 1.

Quadro 1 – Países analisados

QUANT.	PAÍS	COD. PAÍS	REGIÃO
1	Albania	ALB	EMDEU
2	Algeria	DZA	NSWUA
3	Angola	AGO	SUSAF
4	Argentina	ARG	AIBER
5	Armenia	ARM	CWEAS
6	Australia	AUS	AVECO
7	Austria	AUT	AVECO
8	Azerbaijan	AZE	CWEAS
9	Bangladesh	BGD	ESEAS
10	Belarus	BLR	CWEAS
11	Belgium	BEL	AVECO
12	Benin	BEN	SUSAF
13	Bolivia	BOL	AIBER
14	Bosnia and Herzegovina	BIH	EMDEU
15	Botswana	BWA	SUSAF
16	Brazil	BRA	AIBER
17	Bulgaria	BGR	EMDEU
18	Burkina Faso	BFA	SUSAF
19	Burundi	BDI	SUSAF
20	Cambodia	KHM	ESEAS
21	Cameroon	CMR	SUSAF
22	Canada	CAN	AVECO
23	Central African Republic	CAF	SUSAF
24	Chad	TCD	SUSAF
25	Chile	CHL	AIBER
26	China	CHN	ESEAS
27	Colombia	COL	AIBER
28	Congo, Republic of	COG	SUSAF
29	Costa Rica	CRI	AIBER
30	Croatia	HRV	EMDEU
31	Cuba	CUB	CARLA

32	Czech Republic	CZE	AVECO
33	Denmark	DNK	AVECO
34	Djibouti	DJI	SUSAF
35	Dominican Republic	DOM	AIBER
36	Ecuador	ECU	AIBER
37	Egypt	EGY	NSWUA
38	El Salvador	SLV	AIBER
39	Estonia	EST	AVECO
40	Finland	FIN	AVECO
41	France	FRA	AVECO
42	Georgia	GEO	CWEAS
43	Germany	DEU	AVECO
44	Ghana	GHA	SUSAF
45	Greece	GRC	AVECO
46	Guatemala	GTM	AIBER
47	Guinea	GIN	SUSAF
48	Guyana	GUY	CARLA
49	Honduras	HND	AIBER
50	Hungary	HUN	EMDEU
51	Iceland	ISL	AVECO
52	India	IND	ESEAS
53	Indonesia	IDN	ESEAS
54	Iran	IRN	NSWUA
55	Iraq	IRQ	NSWUA
56	Ireland	IRL	AVECO
57	Israel	ISR	AVECO
58	Italy	ITA	AVECO
59	Jamaica	JAM	CARLA
60	Japan	JPN	AVECO
61	Jordan	JOR	NSWUA
62	Kazakhstan	KAZ	CWEAS
63	Kenya	KEN	SUSAF
64	Korea, Republic of	KOR	AVECO
65	Kuwait	KWT	NSWUA
66	Kyrgyzstan	KGZ	CWEAS
67	Laos	LAO	ESEAS
68	Latvia	LVA	AVECO
69	Lebanon	LBN	NSWUA
70	Lesotho	LSO	SUSAF
71	Liberia	LBR	SUSAF
72	Lithuania	LTU	EMDEU
73	Macedonia	MKD	EMDEU
74	Madagascar	MDG	SUSAF
75	Malawi	MWI	SUSAF
76	Malaysia	MYS	ESEAS
77	Mali	MLI	SUSAF
78	Mauritania	MRT	NSWUA
79	Mauritius	MUS	SUSAF
80	Mexico	MEX	AIBER

81	Moldova	MDA	CWEAS
82	Mongolia	MNG	ESEAS
83	Montenegro	MNE	EMDEU
84	Morocco	MAR	NSWUA
85	Mozambique	MOZ	SUSAF
86	Namibia	NAM	SUSAF
87	Nepal	NPL	ESEAS
88	Netherlands	NLD	AVECO
89	New Zealand	NZL	AVECO
90	Nicaragua	NIC	AIBER
91	Niger	NER	SUSAF
92	Nigeria	NGA	SUSAF
93	Norway	NOR	AVECO
94	Pakistan	PAK	NSWUA
95	Panama	PAN	AIBER
96	Paraguay	PRY	AIBER
97	Peru	PER	AIBER
98	Philippines	PHL	ESEAS
99	Poland	POL	EMDEU
100	Portugal	PRT	AIBER
101	Romania	ROU	EMDEU
102	Russia	RUS	CWEAS
103	Rwanda	RWA	SUSAF
104	Saudi Arabia	SAU	NSWUA
105	Senegal	SEN	SUSAF
106	Serbia	SRB	EMDEU
107	Slovakia	SVK	AVECO
108	Slovenia	SVN	AVECO
109	South Africa	ZAF	SUSAF
110	Spain	ESP	AIBER
111	Sri Lanka	LKA	ESEAS
112	Sudan	SDN	SUSAF
113	Swaziland	SWZ	SUSAF
114	Sweden	SWE	AVECO
115	Switzerland	CHE	AVECO
116	Tajikistan	TJK	CWEAS
117	Tanzania	TZA	SUSAF
118	Thailand	THA	ESEAS
119	Togo	TGO	SUSAF
120	Trinidad and Tobago	TTO	CARLA
121	Tunisia	TUN	NSWUA
122	Turkey	TUR	EMDEU
123	Uganda	UGA	SUSAF
124	Ukraine	UKR	CWEAS
125	United Arab Emirates	ARE	NSWUA
126	United Kingdom	GBR	AVECO
127	United States	USA	AVECO
128	Uruguay	URY	AIBER
129	Uzbekistan	UZB	CWEAS

130	Venezuela	VEN	AIBER
131	Yemen	YEM	NSWUA
132	Zambia	ZMB	SUSAF

2. AS VARIÁVEIS

Esta pesquisa é composta por 3 variáveis quantitativas e 3 categóricas – o nome dos países, seu código e região. As mesmas são mais bem explicadas no Quadro 2. As variáveis podem ainda ser divididas em índices sintéticos e variáveis-componentes ou indicadores

Quadro 2 - As variáveis de pesquisa

Variável	Significado	Tipo	Unidade de Medida
País	Nome dos países que serão analisados.	Variável Categórica	N/A
Código do país	Abreviação que representa o país.	Variável Categórica	N/A
Região	Região em que o país está localizado.	Variável Categórica	N/A
SOCIAL PROGRESS INDEX	O índice identifica os pontos fortes e fracos dos países a partir de um conjunto de informações e análises conduzidas para cada um deles.	Variável Quantitativa	Indicador nominal de 0 a 100
Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as políticas são implementadas e qual a capacidade de formulá-las.	Variável Quantitativa	Percentual
Direitos de Propriedade Privada	O grau com que um país protege com leis os Direitos com Propriedade Privada. Medido em escala onde 0 corresponde a nenhuma proteção pelo estado ao direito de propriedade privada e elevado grau de corrupção e 100 corresponde ao total suporte do Governo e nenhum grau de corrupção.	Variável Quantitativa	Indicador nominal que varia de 0 a 100

3. PESQUISA POR AMOSTRAGEM

A pesquisa por amostragem foi feita em três amostras, uma de 25 indivíduos, ou países, a outra composta por 50 indivíduos e por fim, com 100 países. Para cada amostra foram efetuadas uma análise exploratória de dados, amostragens e Boxplots.

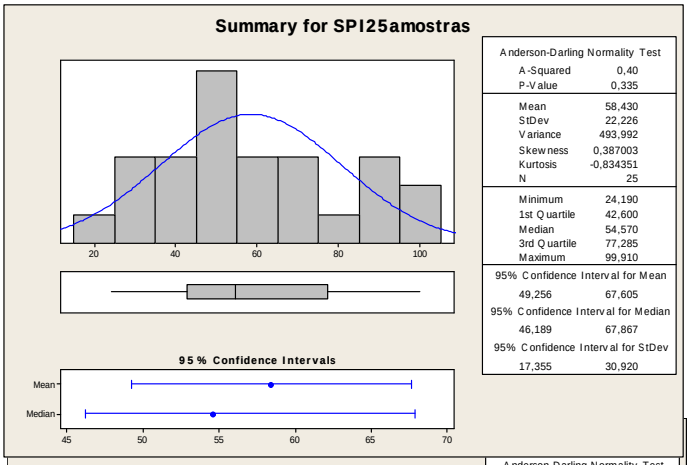
3.1 Análise Exploratória para Amostra de 25 Indivíduos

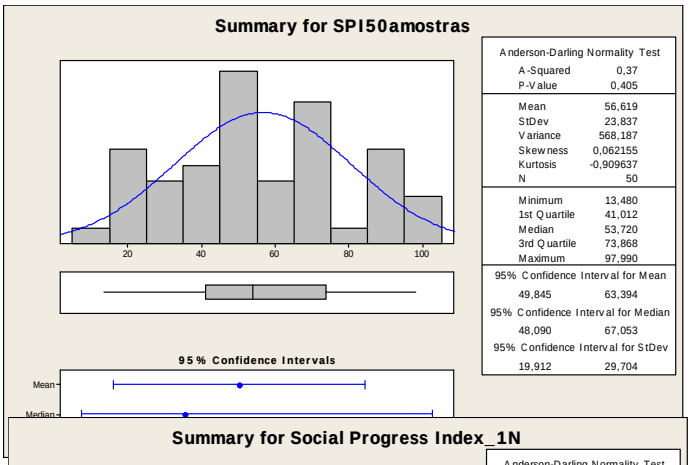
A amostra de 25, 50 e 100 indivíduos foram feitas de seguinte maneira:

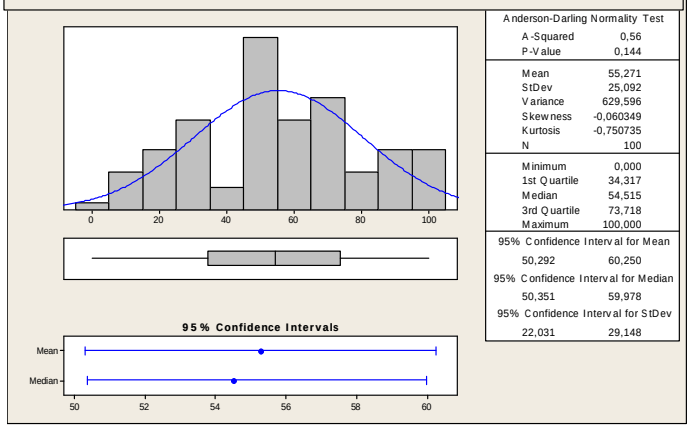
- **Amostra Aleatória**: foi feita através a funcionalidade “Random Data” do MINITAB.

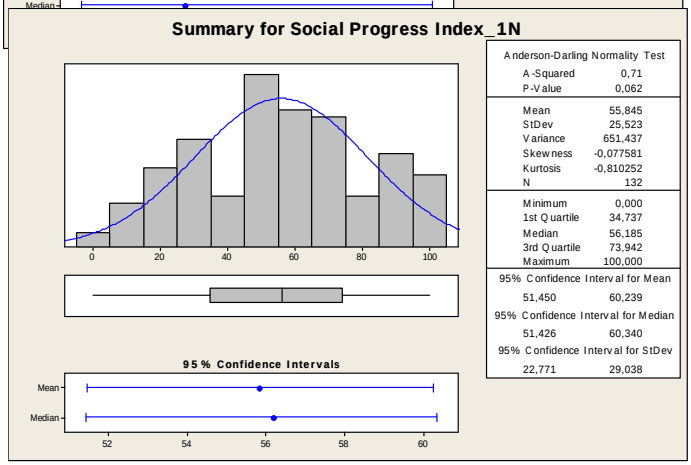
3.2 Social Progress Index

Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países. Em seguida os Boxplots que fazem o comparativo das regiões.









Conforme os gráficos acima expostos para as análises com 25, 50, 100 e com todos os 132 países podemos verificar inicialmente que com o “sorteio” na função “Random Data” do Minitab o bloco com 25 países culminou por ficar com a média mais elevada em 58,43. Seguido pelo bloco de 50 países, depois pelo bloco original com 132 países e por fim pelo bloco de 100 países com médias de 56,619, 55,845 e 55,271 respectivamente.

País25amostras	SPI25amostras
Belgium	89,92>
Ghana	41,98<
Nepal	34,11<
Turkey	57,55<
Poland	80,59>
Iran	43,22<
Canada	97,68>
Burkina Faso	26,47<
Slovenia	88,16>
Moldova	49,46<
Laos	35,60<
Egypt	49,19<
Bulgaria	67,65>
Lithuania	73,98>
Tanzania	24,19<
Jamaica	67,92>
Czech Republic	85,93>
Morocco	45,67<
El Salvador	57,69<
Zambia	31,06<
Tunisia	54,57<
Switzerland	99,91>
Kazakhstan	48,29<
Namibia	51,38<
Belarus	58,59>

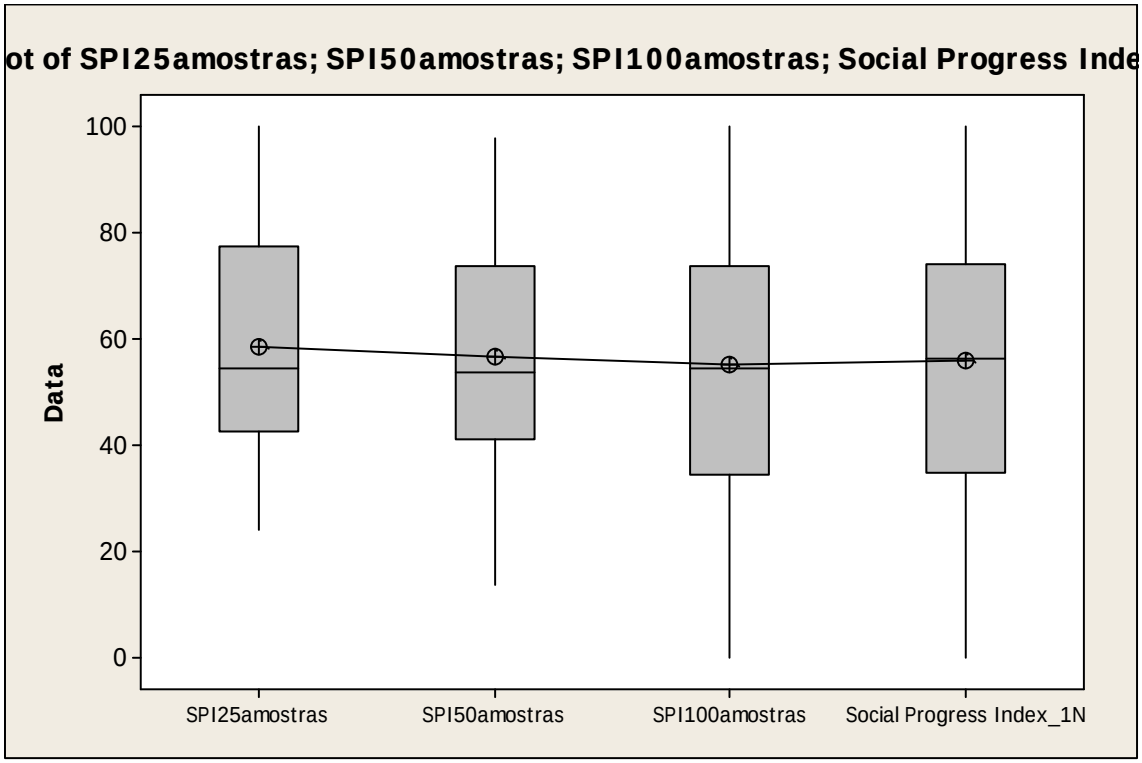
Dentre os países sorteados pelo Random Data nas amostras para 25 países, 10 possuem índice superior à média de 58,43 (estes países estão marcados na tabela ao lado com o símbolo >) e outros 25 estão abaixo da média (marcados na tabela ao lado com o símbolo <)

País 50 amostras	SPI50amostras	País 50 amostras	SPI50amostras
Colombia	62,26 >	United States	90,17>
South Africa	54,57<	Uzbekistan	44,46<
Brazil	67,16>	Japan	92,76>
Djibouti	23,99<	United Arab Emirates	72,47>
Kazakhstan	48,29<	Kenya	31,63<
Egypt	49,19<	Hungary	74,17>
Russia	50,66<	Finland	97,61>
Mauritius	73,83>	Mozambique	22,70<
Nicaragua	53,43<	Germany	93,48>
Estonia	87,49>	Thailand	58,48<
Niger	13,48<	Greece	73,38>
Spain	86,57>	Nigeria	18,06<
Togo	18,33<	Jamaica	67,92>
Namibia	51,38<	Lithuania	73,98>
Kyrgyzstan	44,00<	Iran	43,22<
Cambodia	34,67<	Algeria	47,68<
Madagascar	20,99<	Jordan	52,70<
Tajikistan	42,15<	Tanzania	24,19<
Morocco	45,67<	Congo, Republic of	27,66<
Trinidad and Tobago	67,00>	Rwanda	30,30<
Moldova	49,46<	Portugal	86,07>
Canada	97,68>	Senegal	37,60<
Slovakia	83,27>	Serbia	68,31<
Peru	60,55>	Norway	97,99>
Paraguay	54,01<	Ecuador	63,89>

Dentre os países sorteados pelo Random Data nas amostras para 50 países, 21 possuem índice superior à média de 56,619 (estes países estão marcados na tabela ao lado com o símbolo >) e outros 29 estão abaixo da média (marcados na tabela ao lado com o símbolo <) fazendo com que a mediana tenha ficado em 53,72.

País 100 amostras	SPI100amostras	País 100 amostras	SPI100amostras	País 100 amostras	SPI100amostras	País 100 amostras	SPI100amostras
France	87,19	Poland	80,59	Albania	65,65	Portugal	86,07
Uruguay	80,72	Israel	69,73	Turkey	57,55	Burundi	8,50
Finland	97,61	Guatemala	51,71	Jamaica	67,92	Dominican Republic	54,69
Cuba	51,17	Latvia	74,25	Kuwait	68,40	El Salvador	57,69
Egypt	49,19	Pakistan	17,61	Panama	71,85	Lebanon	49,34
Argentina	68,28	Chad	0,00	Djibouti	23,99	United Arab Emirates	72,47
Iraq	22,00	Jordan	52,70	Spain	86,57	Togo	18,33
Indonesia	47,41	Kazakhstan	48,29	Belarus	58,59	Mauritius	73,83
Mongolia	47,39	Russia	50,66	Benin	29,67	Czech Republic	85,93
Philippines	59,78	China	46,85	Malaysia	67,22	Bosnia and Herzego	58,21
Saudi Arabia	57,12	New Zealand	100,00	Yemen	13,71	Algeria	47,68
India	31,70	Zambia	31,06	Paraguay	54,01	Laos	35,60
Netherlands	98,44	Mali	25,61	Switzerland	99,91	Peru	60,55
Kenya	31,63	Bangladesh	34,94	Italy	79,67	Nicaragua	53,43
Sudan	10,51	Tanzania	24,19	Denmark	96,96	Lesotho	29,37
Angola	13,17	Sweden	97,92	Bulgaria	67,65	Malawi	29,10
Thailand	58,48	Honduras	51,55	Swaziland	29,24	Moldova	49,46
Guyana	49,35	Burkina Faso	26,47	Nepal	34,11	Australia	96,15
Lithuania	73,98	Belgium	89,92	Bolivia	54,46	Costa Rica	81,15
Congo, Republic of	27,66	Hungary	74,17	Ecuador	63,89	Azerbaijan	53,63
Morocco	45,67	Mexico	60,77	Namibia	51,38	Ireland	92,47
Ukraine	58,07	United Kingdom	93,39	Nigeria	18,06	Mozambique	22,70
Senegal	37,60	Mauritania	18,89	Sri Lanka	48,72	Greece	73,38
Estonia	87,49	Chile	78,54	South Africa	54,57	Canada	97,68
Uzbekistan	44,46	Serbia	68,31	Romania	63,12	Guinea	8,64

Por fim, temos acima o Random Data efetuado pelo Minitab para 100 países. Podemos verificar que mesmo aumentando em 100% a quantidade de amostras em relação ao quadro anterior, ou seja, aumentado de 50 para 100 países, verificamos que nossa média, embora tenha recuado de 56,62 (50 amostras) para 55,27 (100 amostras) permaneceram extremamente próximas uma da outra.



Diante do boxplot acima apresentado, podemos corroborar a análise que efetuamos por uso dos gráficos anteriores, uma vez que, temos claro que para todas as amostragens obtivemos muita proximidade tanto nas médias, quanto nas medianas. Sendo nas amostragens com 25 e

50 países há mais amostras com índices acima da média e da mediana e a amostragem com 100 países foi a que mais se aproximou da realidade da variável original com 132 países.

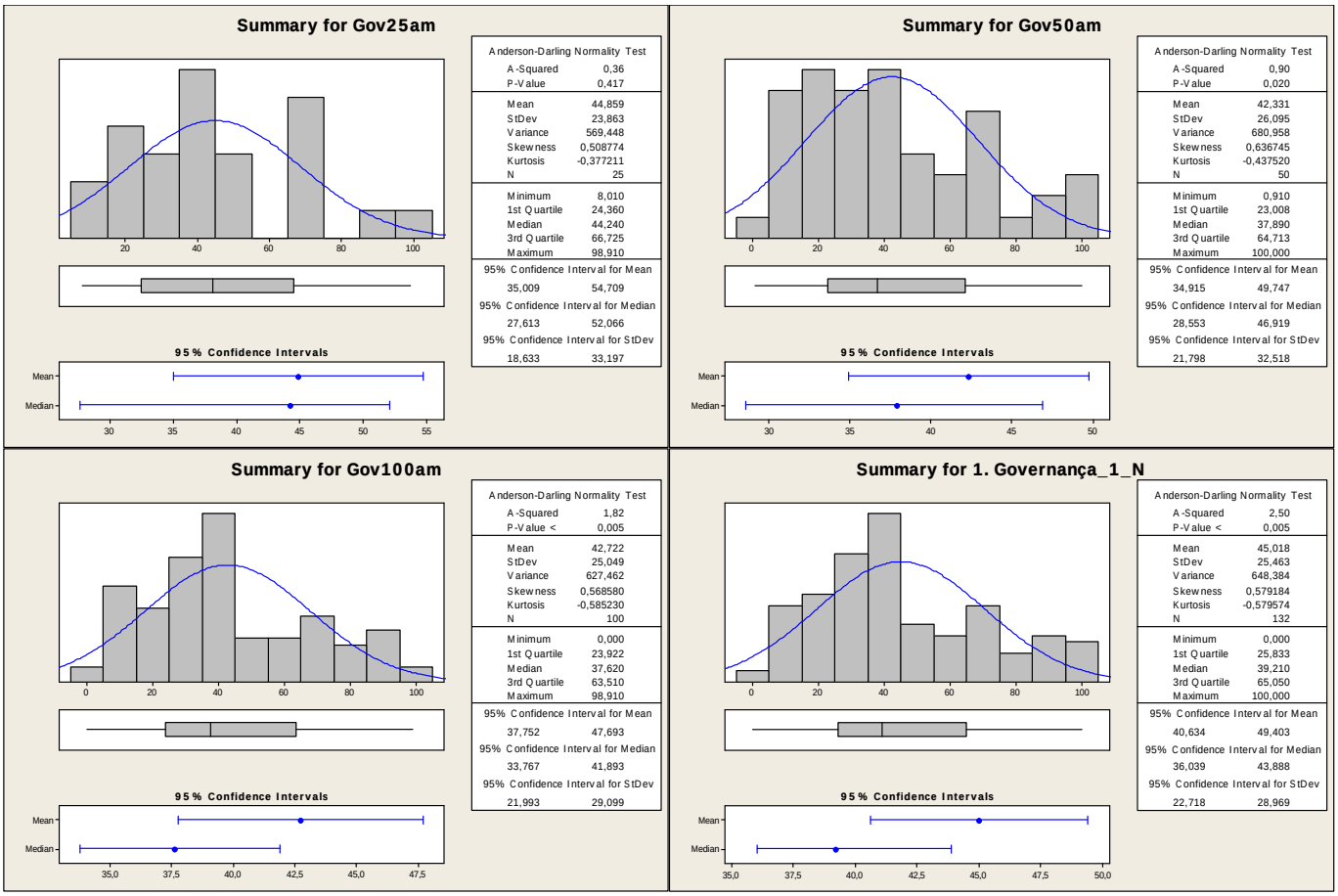
One-way ANOVA: SPI25amostras; SPI50amostras; SPI100amostras; Social Progress

```
Source   DF      SS   MS     F      P
Factor    3     223   74  0,12  0,948
Error   303  187365  618
Total   306  187588

S = 24,87   R-Sq = 0,12%   R-Sq(adj) = 0,00%



Level                        N   Mean  StDev
SPI25amostras               25  58,43  22,23
SPI50amostras               50  56,62  23,84
SPI100amostras             100  55,27  25,09
Social Progress Index_1N   132  55,84  25,52



                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level                      ---+---------+---------+---------+------
SPI25amostras              (------------------*------------------)
SPI50amostras                (-------------*-------------)
SPI100amostras                 (---------*--------)
Social Progress Index_1N         (--------*-------)
                           ---+---------+---------+---------+------
                           50,0      55,0      60,0      65,0

Pooled StDev = 24,87
```

Diante da analise ANOVA que foi apresentada acima observa-se que a maioria das regiões estão sobrepostas, tendo as médias parecidas. Na amostragem com 100 países foi onde a média se apresentou com a menor numeração. Na análise do MINITAB F apresenta o número 0,12 que significa que as regiões não são todas iguais e o gráfico mostra que os intervalos não estão totalmente separados uns dos outros.

3.3 GOVERNANÇA

Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países. Em seguida o Boxplot que faz o comparativo das regiões.



Analisando os gráficos acima gerados pelo MINITAB para a Variável Governança, podemos verificar que há um grande distanciamento entre as médias e as medianas (com exceção para o sorteio de 25 países). Isto ocorre pois após filtrarmos os países que foram sorteados por ordem crescente de seus índices de Governança a amostra que ficou no centro da lista ainda está bem abaixo da média de seu grupo. Por exemplo, no gráfico *summary for gov100am* temos uma média de 42,72, porém conforme podemos verificar a grande concentração de países estão localizadas próximas a mediana em 37,62.

One-way ANOVA: Gov25am; Gov50am; Gov100am; 1. Governança_1_N

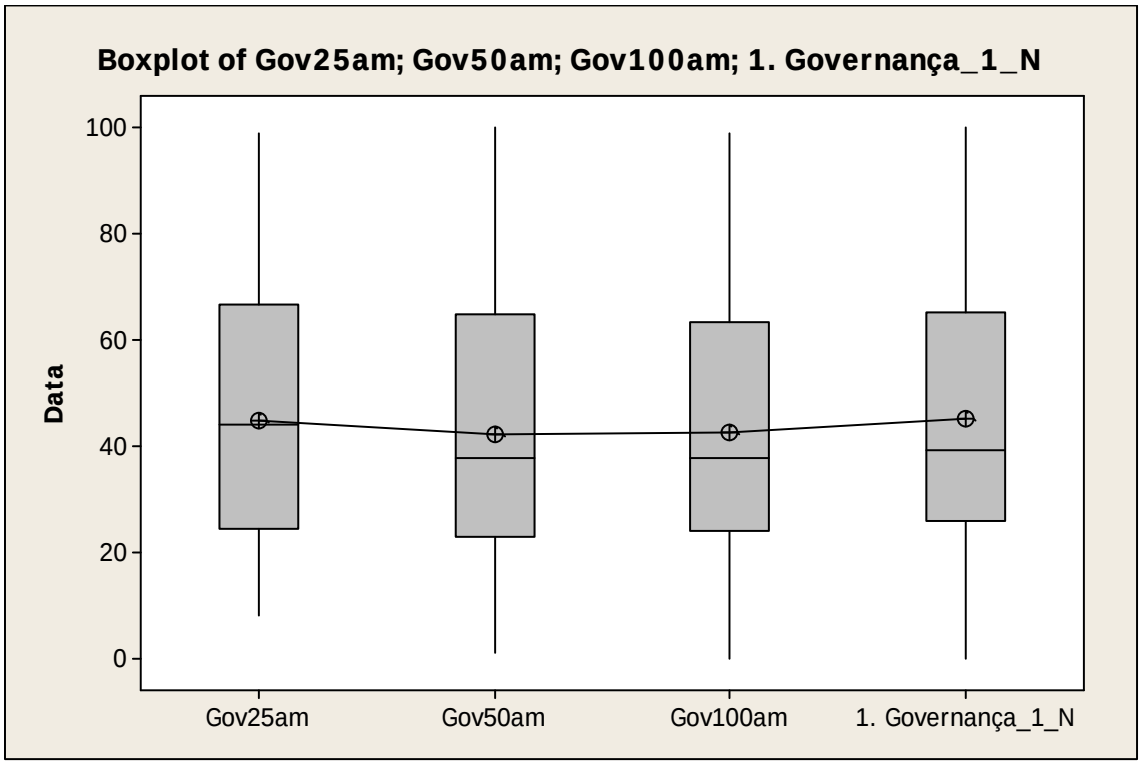
Source	DF	SS	MS	F	P
Factor	3	448	149	0,23	0,873
Error	303	194091	641		
Total	306	194539			

S = 25,31 R-Sq = 0,23% R-Sq(adj) = 0,00%

Level	N	Mean	StDev
Gov25am	25	44,86	23,86

```
Gov50am             50  42,33  26,10
Gov100am           100  42,72  25,05
1. Governança_1_N  132  45,02  25,46


                   Individual 95% CIs For Mean Based on Pooled StDev
Level                 +---------+---------+---------+---------
Gov25am               (-------------------*-------------------)
Gov50am                (-------------*-------------)
Gov100am                     (---------*---------)
1. Governança_1_N                 (--------*--------)
                      +---------+---------+---------+---------
                     35,0      40,0      45,0      50,0

Pooled StDev = 25,31
```

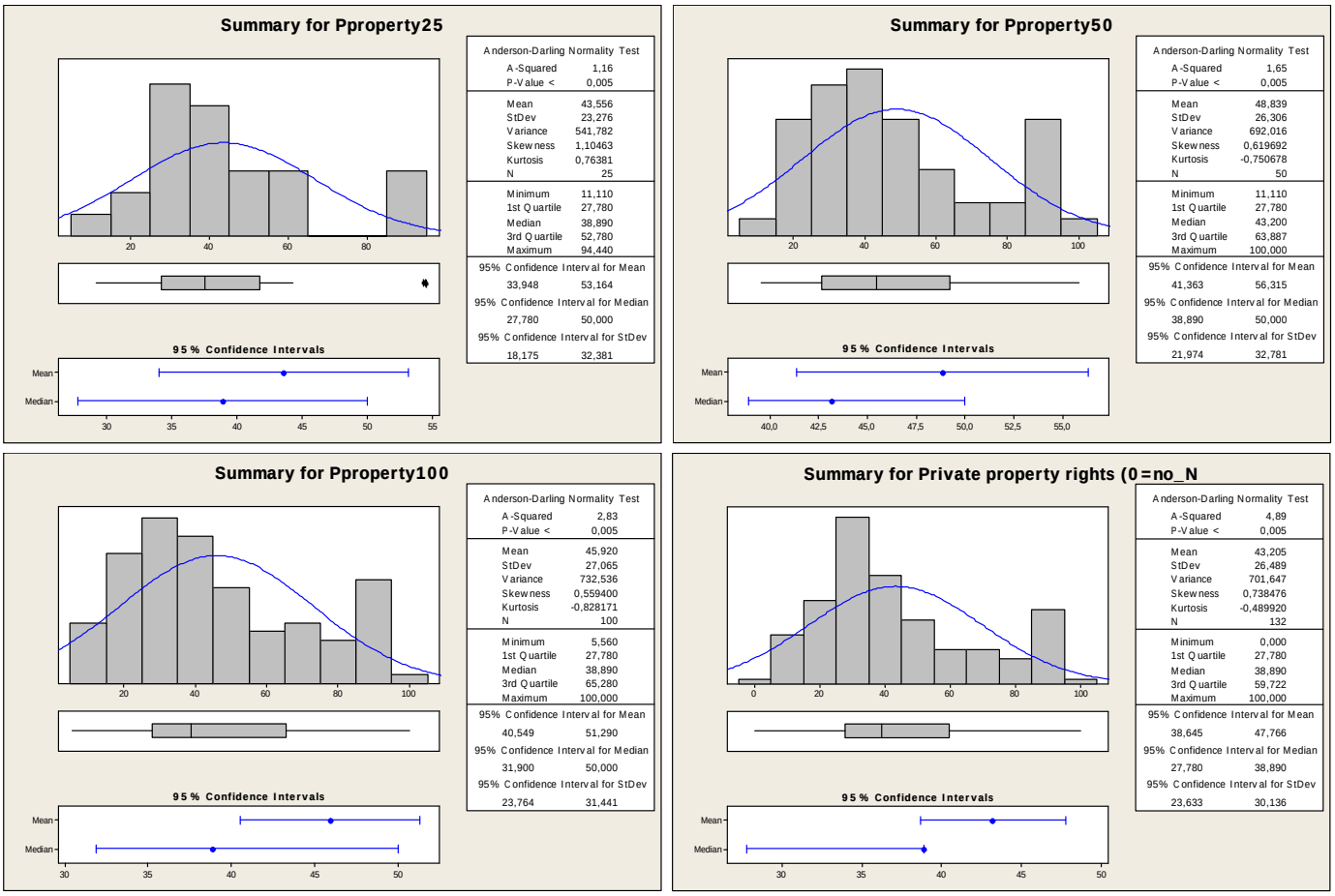
Diante da analise ANOVA que foi apresentada acima observa-se que a maioria das regiões estão sobrepostas, tendo as médias parecidas. Na amostragem com 50 e 100 países as médias foram praticamente as mesmas 42,33 e 42,72 respectivamente. Na análise do MINITAB F apresenta o número 0,23 que significa que as regiões não são todas iguais e o gráfico mostra que os intervalos não estão totalmente separados uns dos outros.



Na análise apresentada no Boxplot acima não houve nenhum outlier e as médias estão bem próximas uma da outra, com exceção da amostra com 25 países que apresenta média um pouco superior às demais.

3.4 PRIVATE PROPERTY RIGHTS

Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países. Em seguida o Boxplot que faz o comparativo das regiões.



Por fim temos acima as análises gráficas da variável Private Property Rights.

Podemos verificar que em todos os gráficos as médias ficam muito próximas entre 43,20 a 48,84 à exemplo das medianas que também variam pouco, 38,89 para as amostras com 25, 100 e 132 países e 43,20 para a amostra com 50 países.

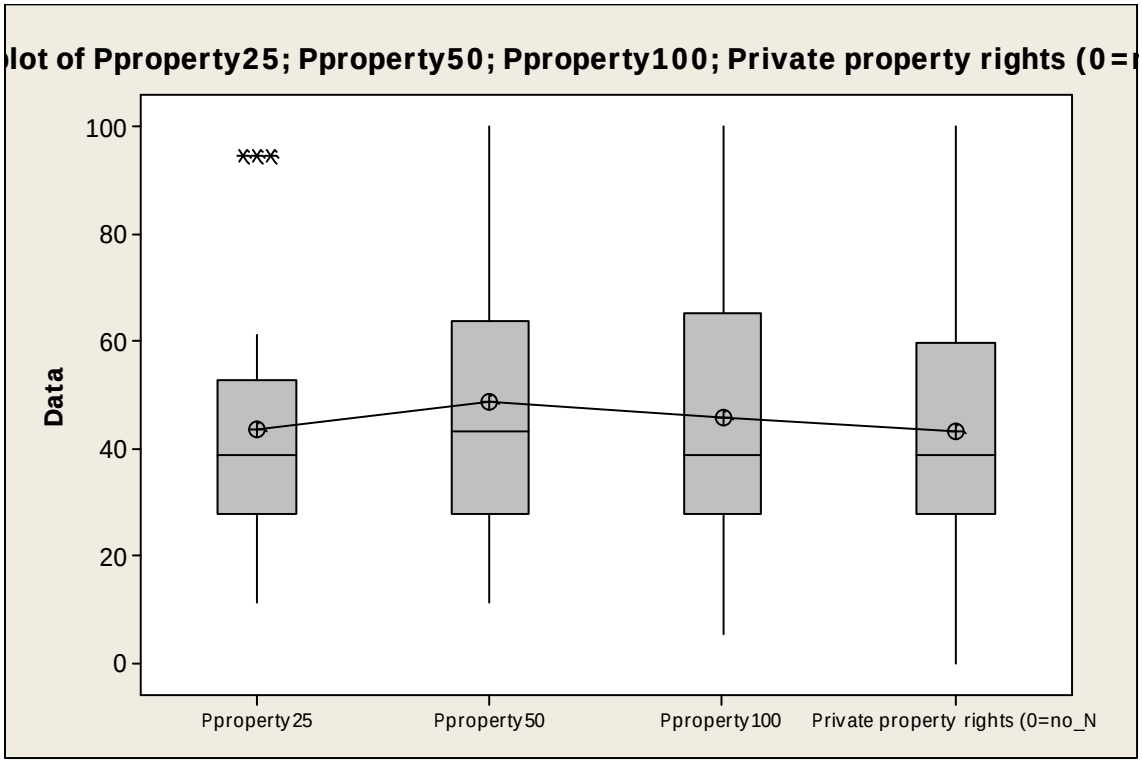
One-way ANOVA: Pproperty25; Pproperty50; Pproperty100; Private property rights

Source	DF	SS	MS	F	P
Factor	3	1298	433	0,62	0,602
Error	303	211348	698		
Total	306	212647			

S = 26,41 R-Sq = 0,61% R-Sq(adj) = 0,00%

Level	N	Mean	StDev
Pproperty25	25	43,56	23,28

```
Pproperty50                50  48,84  26,31
Pproperty100              100  45,92  27,07
Private property rights   132  43,21  26,49

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level                     -----+---------+---------+---------+----
Pproperty25               (-----------------*-----------------)
Pproperty50                              (-----------*-----------)
Pproperty100                           (--------*-------)
Private property rights             (-------*-------)
                          -----+---------+---------+---------+----
                            36,0      42,0      48,0      54,0

Pooled StDev = 26,41
```



Diante da analise ANOVA que foi apresentada acima observa-se que as regiões não estão tão sobrepostas umas às outras, conforme observamos nas análises de Social Progress Index e Governança.

Na amostragem com 25, 50 e 132 países as médias variaram pouco, entre 43,20 a 48,84. Havendo similaridade na média quando a análise foi feita com 25 países e 132 países. Na análise do MINITAB F apresenta o número 0,62 onde as médias apresentam grau de menor similaridade em comparação com as outras duas variáveis que analisamos no presente trabalho.

4. CONSIDERAÇÕES FINAIS

Diante das análises realizadas observou-se que a média das amostragens foram bem parecidas e até mesmo tiveram resultados iguais. É possível verificar que fazendo a análise com amostras diferentes esses grupos de países que foram escolhidos aleatoriamente, demonstram nos resultados que podem ser obtidos utilizando todos os países ou com uma quantidade menor. Sendo assim, em uma pesquisa não se faz necessário utilizar uma análise com todo universo, pois uma amostra mais reduzida já pode apresentar os resultados, desde que essa análise seja executada mediante uma busca aleatória das variáveis.

CAPITULO 7 - ANÁLISE MULTIVARIADA: Correlações, Dendogramas e Análise dos Componentes Principais

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise das correlações e dos componentes principais (análise multivariada) de dados dimensionadores do nível de MORADIA de diversos países do mundo separados por AIBER, AVECO e Outros. Para tal, iniciamos com análise da estatística descritiva. Em seguida passamos para a análise das correlações e dendogramas. Na terceira parte utilizamos a análise dos componentes principais. Por fim, fazemos as considerações finais. Os dados são compilados pelo Banco Mundial. O software estatístico utilizado é o **MINITAB.**

2. ENTENDENDO OS DADOS

Os dados são referentes ao ano de 2013. As variáveis são indicadores agregados da qualidade de moradia de cada país.

2.2 As Variáveis

São 13 as variáveis desta pesquisa, além do nome dos países. As mesmas são melhor explicadas na Tabela 1.

Tabela 1. As Variáveis

Variável	Significado	Tipo	Unidade de Medida	Fonte
SOCIAL PROGRESS INDEX	O índice identifica os pontos fortes e fracos dos países a partir de um conjunto de informações e análises conduzidas para cada um deles.	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
SHELTER (Abrigo)	Este índice indica se as pessoas possuem abrigos (lar) com suas necessidades básicas atendidas?	Variável Quantitativa	Indicador nominal de 0 a 100	IPS
Acesso a Água e saneamento Básico	Porcentagem da população que tem uma ou mais torneiras de água encanada tratada.	Variável Quantitativa	Percentual	WHO UNICEF
Acesso a tratamento sanitário Melhorado	Inclui a população com melhores condições de saneamento, canalização de esgoto, fossas sépticas, latrinas com laje ou melhoradas para as condições sanitárias adequadas	Variável Quantitativa	Percentual	WHO UNICEF
Disponibilidade e Acessibilidade do Imóvel	Percentual da população que responderam em pesquisa se “em sua cidade ou região em que vive, está satisfeito com a disponibilidade e acessibilidade do imóvel”	Variável Quantitativa	Percentual	Galoop World Poll

Variável	Significado	Tipo	Unidade de Medida	Fonte
Acesso a Eletricidade	Percentual de pessoas com acesso a energia elétrica.	Variável Quantitativa	Percentual	UN Sustain Able energy for All Project
Qualidade no fornecimento de energia elétrica	Média das respostas para a pergunta: "Em seu país, como você avalia a confiança no fornecimento de energia?". Usando Escala onde 1 não é confiável e 7 extremamente confiável para mostrar a qualidade do fornecimento de energia elétrica.	Variável Quantitativa	Indicador nominal que varia de 1 a 7	World Economc Forum Global Competiti veness Report
Direitos de Propriedade Privada	O grau com que um país protege com leis os Direitos com Propriedade Privada. Medido em escala onde 0 corresponde a nenhuma proteção pelo estado ao direito de propriedade privada e elevado grau de corrupção e 100 corresponde ao total suporte do Governo e nenhum grau de corrupção.	Variável Quantitativa	Indicador nominal que varia de 0 a 100	Heritage Foundation
IDH Índice de Desenvolvimento Humano	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável Quantitativa	Indicador nominal que varia de 0 a 1	PNUD
Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as políticas são implementadas e qual a capacidade de formulá-las.	Variável Quantitativa	Percentual	WGI 2014
EPI- Indice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	Variável Quantitativa	Percentual	EPI 2014
HPI – Bem estar sustentável do país	Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	Variável Quantitativa	Percentual	HPI 2014
GINNI index	Mede a distribuição de renda e de despesas das famílias	Variável Quantitativa	Percentual	Banco Mundial

Fonte: Elaboração própria; IPS, Unicef, GWP, WEF, Heritage Foundation, PNUD, WGI, EPI, Banco Mundial

2.3 Fonte de Dados

Todos os dados desta pesquisa foram obtidos no sítio de fornecimento de dados e estatísticas *Knoema:* *http://pt.knoema.com/atlas*

3. CORRELAÇÕES E DENDOGRAMAS

O objetivo do presente trabalho é, por meio da análise dos componentes principais, tentarmos reduzir o número de variáveis, ou seja, percebermos as relações entre as variáveis e a possibilidade de agruparmos as mesmas. Certamente a análise de correlações e dendogramas nos darão uma ideia das possibilidades de agrupamento nos apontando onde há elevados índices de correlação entre todas as variáveis:

Segue abaixo o resultado das análises dos componentes principais juntamente com o gráfico de Loadind Plot, Scree Plot, Dendograma e gráficos 3d.

Principal Component Analysis: Social Progr; Shelter_1N; Access to pi; Access to

Eigenanalysis of the Correlation Matrix

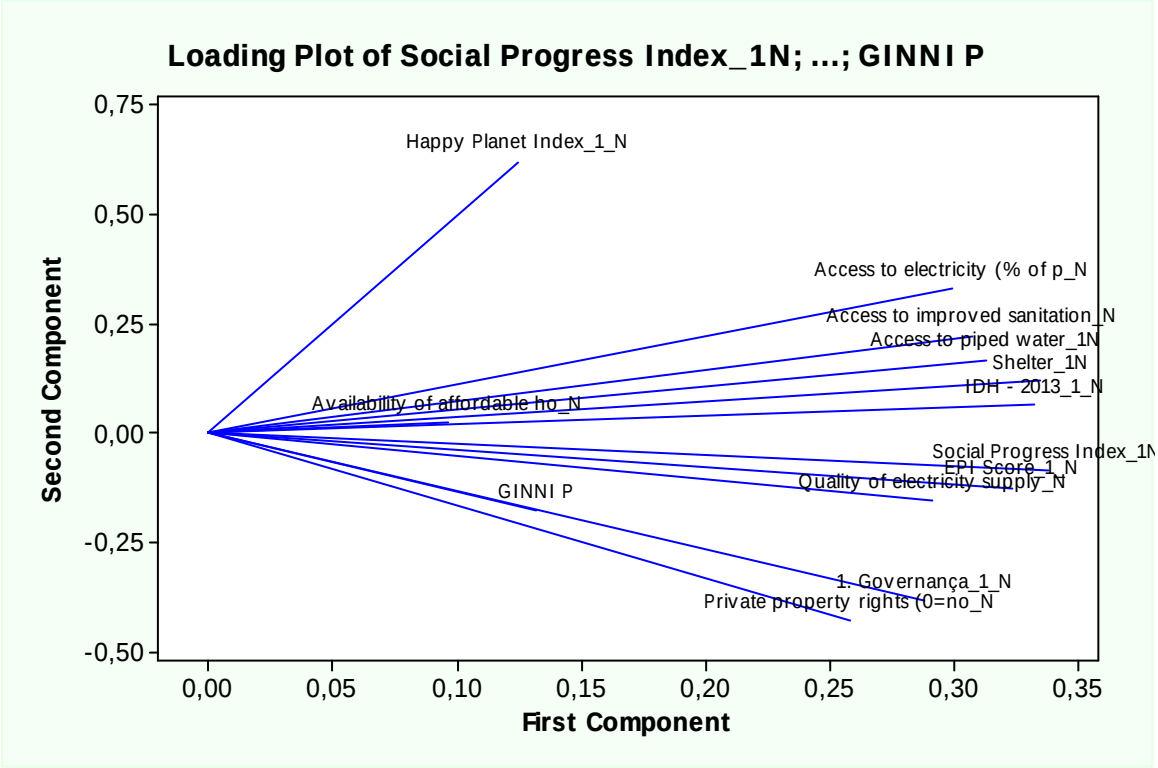
Eigenvalue	8,1912	1,3952	1,1919	0,7780	0,4933	0,3318	0,1617	0,1352
Proportion	0,630	0,107	0,092	0,060	0,038	0,026	0,012	0,010
Cumulative	0,630	0,737	0,829	0,889	0,927	0,952	0,965	0,975

Eigenvalue	0,1284	0,0831	0,0609	0,0315	0,0178
Proportion	0,010	0,006	0,005	0,002	0,001
Cumulative	0,985	0,992	0,996	0,999	1,000

Variable	PC1	PC2	PC3	PC4
Social Progress Index_1N	0,338	-0,086	0,081	-0,058
Shelter_1N	0,335	0,120	0,003	-0,010
Access to piped water_1N	0,313	0,167	-0,113	-0,196
Access to improved sanitation_N	0,307	0,222	-0,159	0,009
Availability of affordable ho_N	0,096	0,024	0,717	0,552
Access to electricity (% of p_N	0,299	0,330	-0,156	0,067
Quality of electricity supply_N	0,291	-0,157	0,010	-0,019
Private property rights (0=no_N	0,258	-0,431	0,204	-0,062
IDH - 2013_1_N	0,332	0,064	-0,071	-0,068
1. Governança_1_N	0,288	-0,382	0,192	-0,117
EPI Score_1_N	0,324	-0,125	-0,064	-0,083
Happy Planet Index_1_N	0,124	0,619	0,266	0,080
GINNI P	0,132	-0,177	-0,507	0,783

Conforme podemos verificar na Tabela 1 (pg. 110), temos à disposição 13 Variáveis para trabalharmos o assunto de Moradia nos 132 países que compõem o campo de estudos do World Economic Forum. No entanto, grande parcela destas variáveis possui grau de similaridade entre si, sendo, portanto, pertinente reduzirmos estas informações, a fim de deixarmos nossas análises mais condensadas.

Uma forma de condessarmos estas variáveis, ou seja, tentar reduzi-las sem ao mesmo tempo perdermos índices e dados desejáveis ao projeto é aplicarmos a Análise de Componentes Principais, que nos auxiliará de forma estatística a gerar quatro novas variáveis que nos permitirão substituir as 13 trabalhadas até o presente momento.



O Loading Plot acima é a representação gráfica do posicionamento das 13 variáveis nos primeiros e segundos componentes.

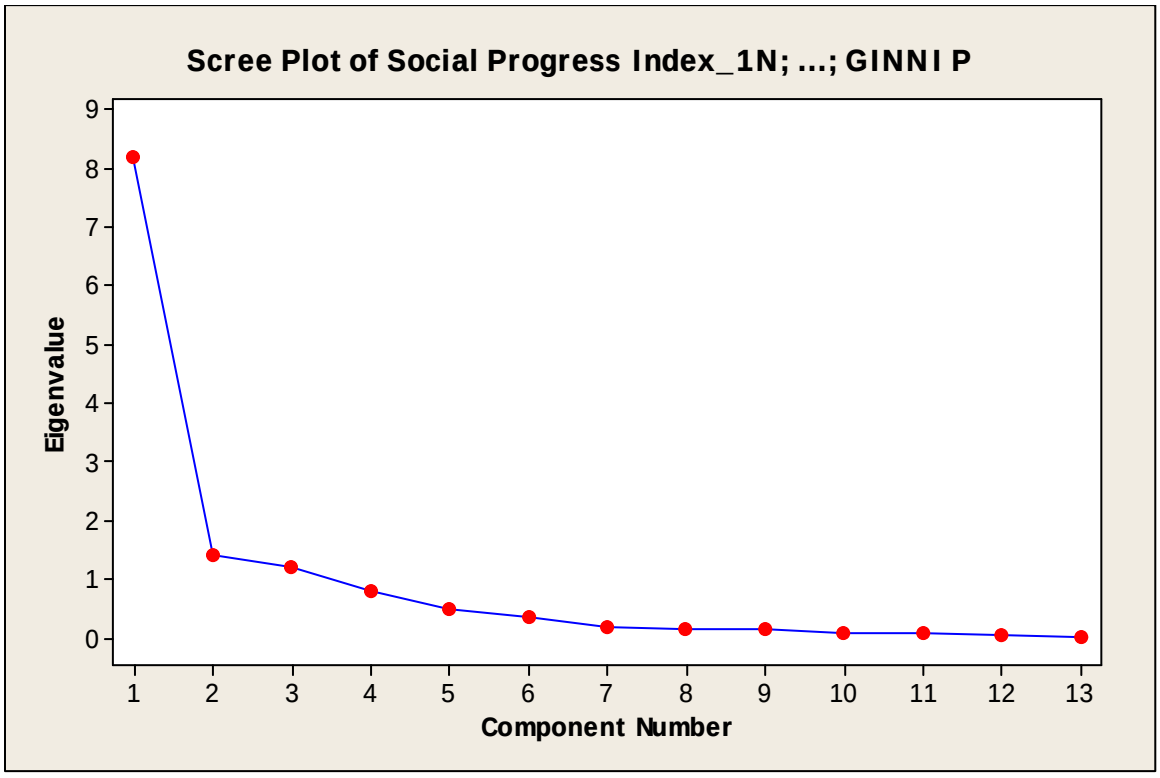
A primeira componente (eixo das abscissas) complementa a segunda componente (eixo das ordenadas) e vice-versa, mas não se correlacionam.

Percebe-se que em nosso gráfico não encontramos nenhuma variável negativada, ou seja, que tenha ficado direcionada à esquerda dos eixos, embora este fenômeno seja possível de acontecer nos demais temas.

Ao verificarmos nosso gráfico, percebemos que a variável HPI é a que mais se distanciou das demais variáveis, pois ficou com sua segunda componente em 0,619.

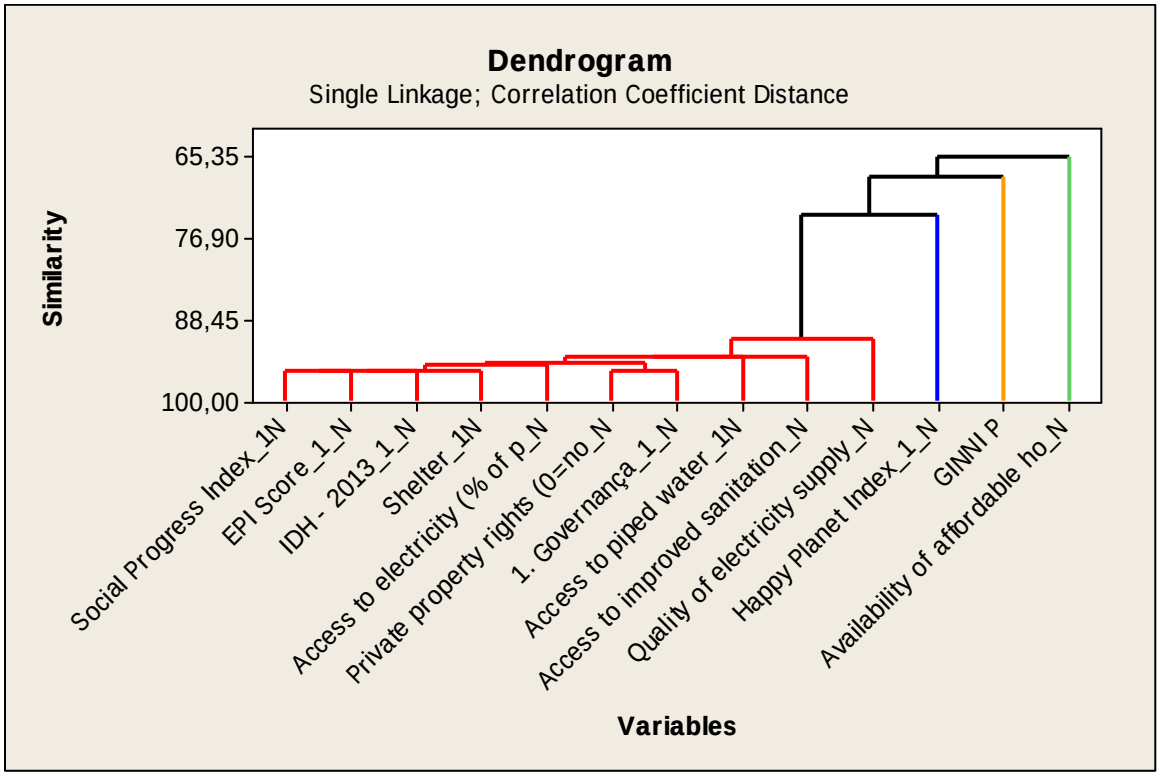
No outro extremo as variáveis Governança e Private Property Rights foram as que mais se distanciaram das demais variáveis tendo ficado com a segunda componente em -0,382 -0,431 respectivamente.

Na primeira componente a variável que mais se distanciou foi Social Progress Index com 0,338.



O Scree Plot acima é a representação gráfica do posicionamento das 13 variáveis para suas respectivas Eigenvalue.

As 4 primeiras bolas vermelhas, da esquerda para a direita, são referente as variáveis mais importantes para darmos seguimento ao nosso estudo: Social Progress é a principal, seguida de Happy Planet Index, Affordable House e Ginni.



3.1 Stepwise PC1

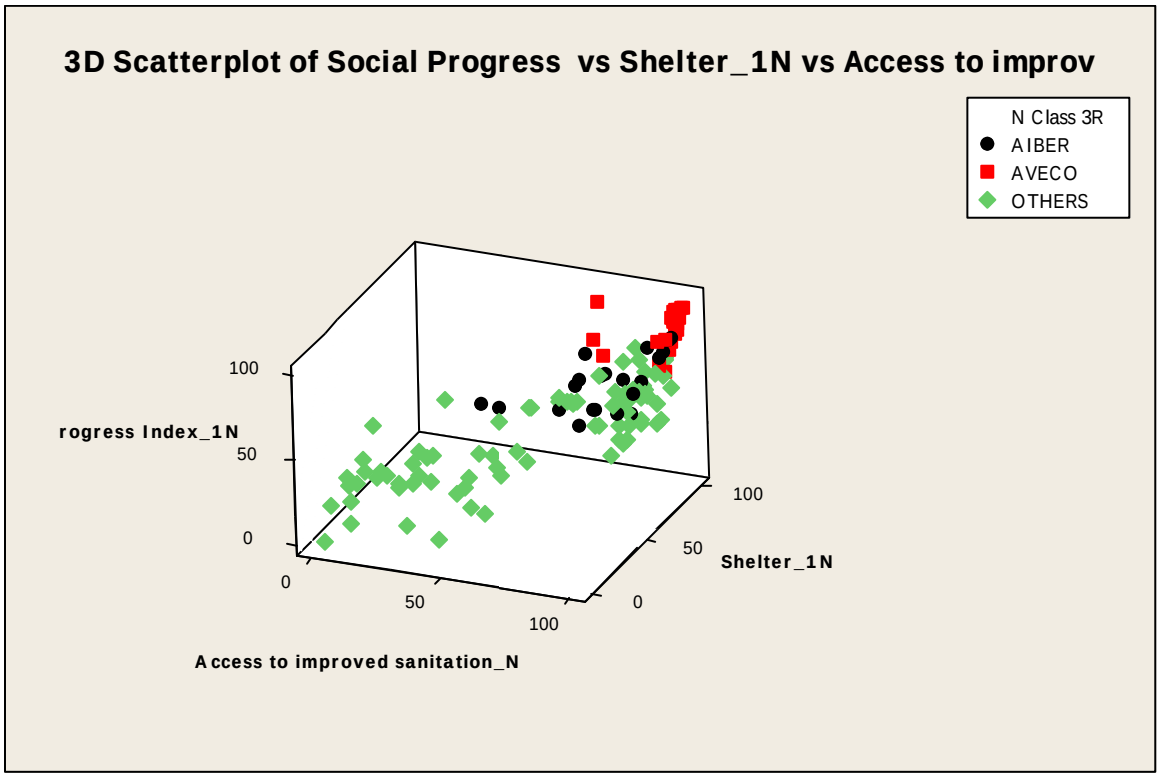
Stepwise Regression: PC1 versus Social Progress ; Shelter_1N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 13 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-6,054	-6,742	-6,767	-6,824	-6,813	-6,846
Social Progress Index_1N	0,1084	0,0615	0,0609	0,0456	0,0350	0,0332
T-Value	43,06	18,42	20,87	14,88	10,83	11,41
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
Shelter_1N		0,0514	0,0371	0,0380	0,0383	0,0293
T-Value		15,68	10,22	12,75	14,54	10,33
P-Value		0,000	0,000	0,000	0,000	0,000
Access to improved sanitation_N			0,0137	0,0176	0,0153	0,0118
T-Value			6,40	9,63	9,26	7,33
P-Value			0,000	0,000	0,000	0,000
Private property rights (0=no_N				0,0135	0,0123	0,0153

T-Value				7,97	8,11	10,51
P-Value				0,000	0,000	0,000
EPI Score_1_N					0,0158	0,0168
T-Value					6,05	7,17
P-Value					0,000	0,000
Access to electricity (% of p_N						0,0100
T-Value						5,66
P-Value						0,000
S	0,735	0,433	0,378	0,310	0,274	0,246
R-Sq	93,45	97,75	98,29	98,86	99,12	99,30
R-Sq(adj)	93,40	97,71	98,25	98,83	99,08	99,26



Análise:

Social Progress Index, Shelter e Access to Improved Sanitation são as variáveis principais. A Social Progress explica 93,45%, depois há um acréscimo de 4,3%. Na terceira variável, o acréscimo é de apenas 0,54%. Assim os dois primeiros são suficientes como variáveis explicativas do PC1. Observa-se que no PC1 as variáveis Social Progress e Shelter sozinhas explicam mais de 97% da variabilidade, sendo as demais variáveis componentes de baixa contribuição. Sugerimos para este Componente Principal a nomenclatura: Moradia e Desenvolvimento Social (MeDS).

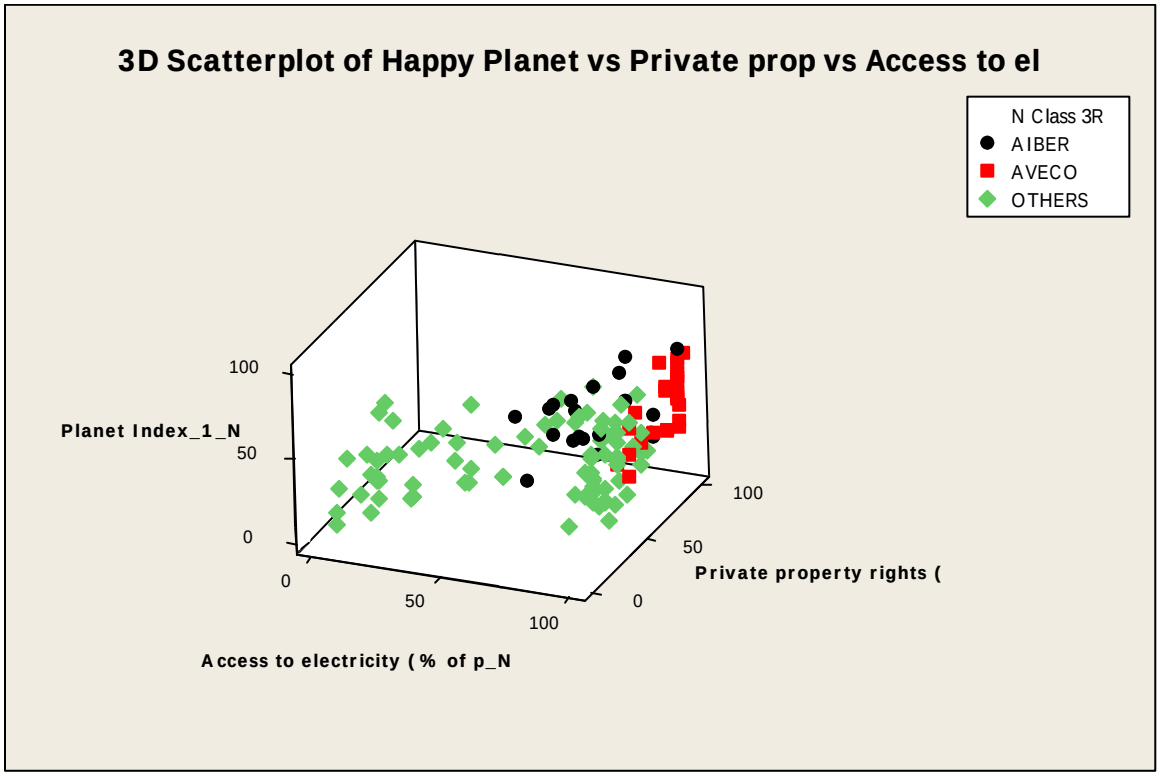
3.2 Stepwise PC2

Stepwise Regression: PC2 versus Social Progress ; Shelter_1N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 13 predictors, with N = 132

Step	1	2	3	4
Constant	-1,9422	-0,9618	-1,2755	-0,8210
Happy Planet Index_1_N	0,04008	0,04214	0,03278	0,03043
T-Value	12,24	22,38	24,12	26,78
P-Value	0,000	0,000	0,000	0,000
Private property rights (0=no_N		-0,02500	-0,03110	-0,03032
T-Value		-16,31	-29,42	-35,20
P-Value		0,000	0,000	0,000
Access to electricity (% of p_N			0,01331	0,01579
T-Value			14,07	19,19
P-Value			0,000	0,000
GINNI P				-0,00900
T-Value				-8,26
P-Value				0,000
1. Governança_1_N				
T-Value				
P-Value				
Access to improved sanitation_N				
T-Value				
P-Value				
S	0,808	0,464	0,292	0,236
R-Sq	53,53	84,83	94,04	96,12
R-Sq(adj)	53,17	84,59	93,90	96,00



Análise:

Happy Planet Index influencia bastante aqui. Esse novo indicador possui indicadores muito próximos a zero, têm um poder de explicação baixo de variabilidade. Poderia ser batizado com o nome: Moradia como fator de alegria (McFA).

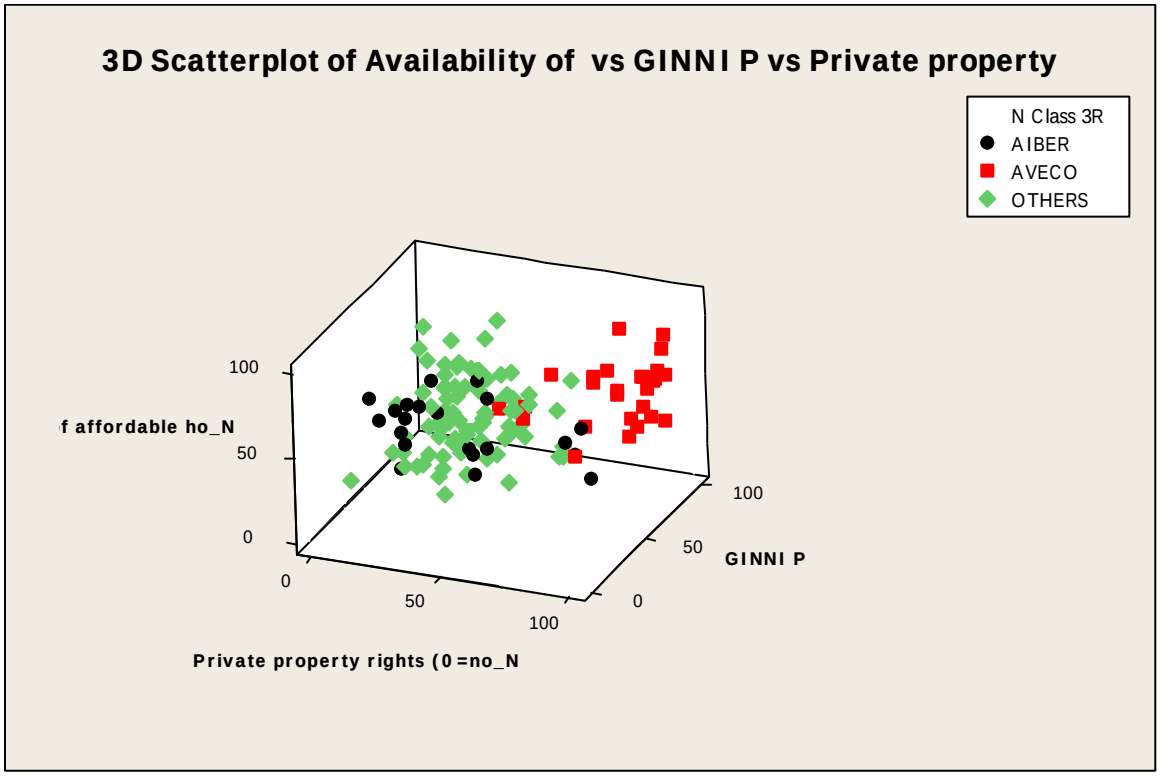
3.3 Stepwise PC3

Stepwise Regression: PC3 versus Social Progress ; Shelter_1N; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15


Response is PC3 on 13 predictors, with N = 132
```

Step	1	2	3	4
Constant	-1,95861	-0,20423	-0,23822	-0,05344
Availability of affordable ho_N	0,04390	0,04272	0,03986	0,04080
T-Value	14,36	25,69	26,32	37,72
P-Value	0,000	0,000	0,000	0,000
GINNI P		-0,02704	-0,02960	-0,02649
T-Value		-17,65	-21,26	-25,73
P-Value		0,000	0,000	0,000

Private property rights (0=no_N			0,00747	0,01128
T-Value			6,48	12,69
P-Value			0,000	0,000
Access to improved sanitation_N				-0,00827
T-Value				-11,20
P-Value				0,000
Happy Planet Index_1_N				
T-Value				
P-Value				
Access to electricity (% of p_N				
T-Value				
P-Value				
S	0,682	0,370	0,323	0,230
R-Sq	61,32	88,67	91,47	95,71
R-Sq(adj)	61,02	88,49	91,27	95,57



Análise:

A variável que mais influencia no PC3 é Availability of Affordabel House. Porém com grau de explicação baixo, próximo a zero, com valor negativo de Ginni e valor extremamente baixo de Private Property Rights. PC3: Urbanização e Direitos Fundamentais.

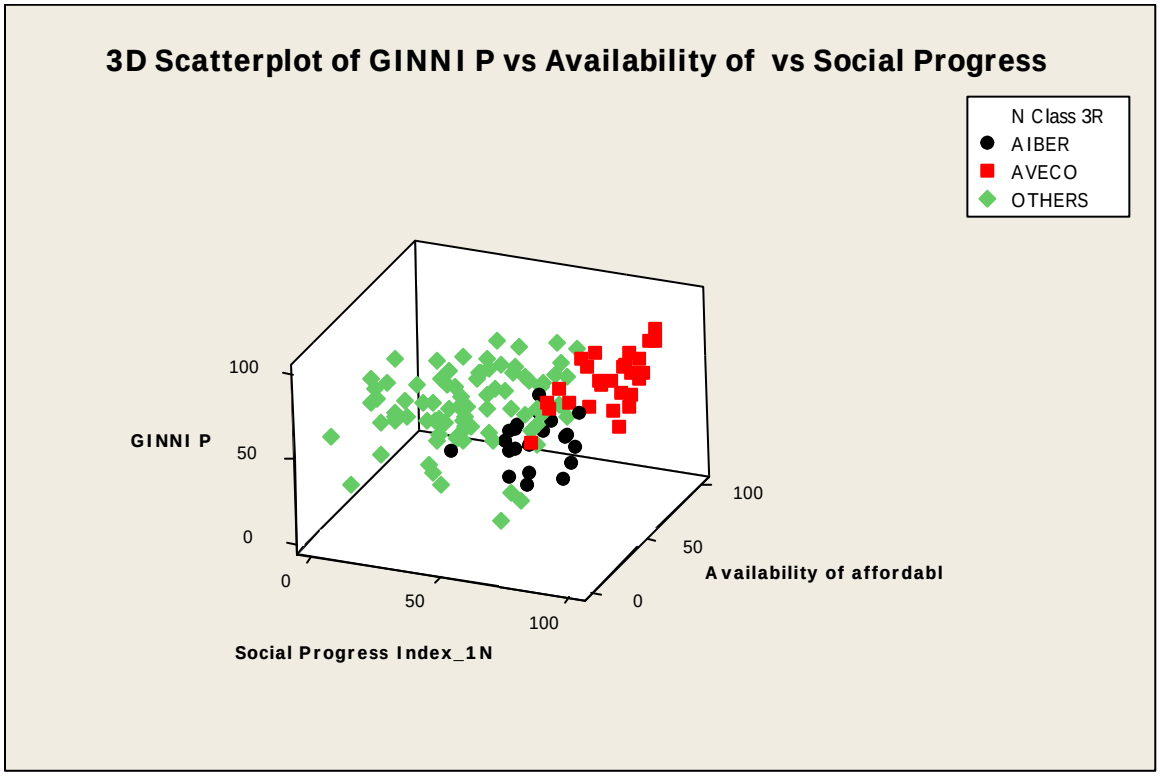
3.4 Stepwise PC4

Stepwise Regression: PC4 versus Social Progress ; Shelter_1N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4 on 13 predictors, with N = 132

Step	1	2	3	4	5
Constant	-1,815	-2,910	-2,675	-2,846	-2,835
GINNI P	0,02884	0,02970	0,03686	0,03751	0,03757
T-Value	10,89	15,91	55,36	72,66	87,57
P-Value	0,000	0,000	0,000	0,000	0,000
Availability of affordable ho_N		0,02334	0,03107	0,03024	0,03064
T-Value		11,52	43,03	53,83	65,23
P-Value		0,000	0,000	0,000	0,000
Social Progress Index_1N			-0,01845	-0,01963	-0,01640
T-Value			-31,88	-42,52	-28,69
P-Value			0,000	0,000	0,000
Happy Planet Index_1_N				0,00480	0,00390
T-Value				9,48	8,93
P-Value				0,000	0,000
Private property rights (0=no_N					-0,00392
T-Value					-7,61
P-Value					0,000
Access to piped water_1N					
T-Value					
P-Value					
S	0,640	0,451	0,151	0,116	0,0967
R-Sq	47,72	74,23	97,12	98,31	98,84
R-Sq(adj)	47,32	73,83	97,05	98,26	98,80



Análise:

No PC4, temos Ginni próximo a zero. Observa-se que Ginni explica apenas 47% da variabilidade, A variável Affordable House acrescenta muito a esta variável, acrescendo a explicabilidade a 74%, por fim Social Progress Index finaliza a explicabilidade com 97%.

Nomeamos esta nova variável como: Incentivos ao Direito de Moradia.

Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	-0,000		
	1,000		
PC3	0,000	-0,000	
	1,000	1,000	
PC4	-0,000	0,000	0,000
	1,000	1,000	1,000

Cell Contents: Pearson correlation
P-Value

Análise:

Correlação = 0 nula o que significa que

Cada variável está explicando dados variados, portanto trata-se de explicações complementares.

4. RENOMEANDO AS NOVAS VARIÁVEIS

Quadro 1: Relação de Principal Component analysis

	Variável 1	Variável 2	Variável 3	Nova Variável
PC1	Social Progress Index ou IDH	Shelter	Access to Improved Sanitation	Moradia e Desenvolvimento Social
PC2	Happy Planet Index	Private Property Rights	Access to Eletricity	Direitos de Moradia e Qualidade de Vida
PC3	Availability of Affordable House	Ginni	Private Property Rights	Urbanização e Direitos Fundamentais
PC4	Ginni	Availability of Affordable House	Social Progress Index ou IDH	Incentivos ao Direito de Moradia

4.1 General Regression Analysis: PC1 versus Social Progr; Shelter_1N; Access to im

```
Regression Equation

PC1  =  -6,76656 + 0,0609229 Social Progress Index_1N + 0,0371388 Shelter_1N +
        0,0137389 Access to improved sanitation_N


Coefficients

Term                                Coef    SE Coef         T      P
Constant                        -6,76656  0,0883427  -76,5944  0,000
Social Progress Index_1N         0,06092  0,0029186   20,8740  0,000
Shelter_1N                       0,03714  0,0036336   10,2210  0,000
Access to improved sanitation_N  0,01374  0,0021474    6,3979  0,000


Summary of Model

S = 0,378370     R-Sq = 98,29%        R-Sq(adj) = 98,25%
PRESS = 19,5169  R-Sq(pred) = 98,18%
```

4.1.1 REGRESSÃO PC1 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH

Há um consenso de que há muita complexidade na composição da Variável Social Progress Index, por este motivo estamos rodando nova simulação com a variável IDH para verificarmos seu potencial e possibilidade de seguirmos adiante em nossa análise com o IDH no lugar de Social Progress Index, uma vez que estas variáveis possuem em sua gênese grande grau de correlação.

Era esperado obtermos R^2 inferior de IDH em relação a SPI. No nosso caso na primeira regressão tínhamos grau de explicabilidade de 98,3%, enquanto na segunda regressão a explicabilidade ficou em 95,2%, portanto menos 3,1%. Sendo assim com a pequena queda de explicabilidade optaremos por darmos continuidade em nossas análises à partir das variáveis que compõem a segunda regressão de MeDS (PC1) com IDH.

General Regression Analysis: PC1 versus IDH - 2013_1; Shelter_1N; Access to im

```
Regression Equation

PC1  =  -6,77697 + 0,0481585 IDH - 2013_1_N + 0,0561479 Shelter_1N + 0,00520075
        Access to improved sanitation_N


Coefficients

Term                                  Coef   SE Coef         T      P
Constant                          -6,77697  0,147418  -45,9710  0,000
IDH - 2013_1_N                     0,04816  0,005580    8,6308  0,000
Shelter_1N                         0,05615  0,005842    9,6113  0,000
Access to improved sanitation_N    0,00520  0,003760    1,3830  0,169


Summary of Model

S = 0,631315     R-Sq = 95,25%        R-Sq(adj) = 95,13%
PRESS = 54,9448  R-Sq(pred) = 94,88%


Analysis of Variance
```

4.2 General Regression Analysis: PC2 versus Happy Planet; Private prop; ...

Regression Equation

PC2 = -1,2755 + 0,032779 Happy Planet Index_1_N - 0,031102 Private property rights (0=no_N + 0,0133149 Access to electricity (% of p_N

Coefficients

Term	Coef	SE Coef	T	P
Constant	-1,27550	0,0764985	-16,6735	0,000
Happy Planet Index_1_N	0,03278	0,0013590	24,1201	0,000
Private property rights (0=no_N	-0,03110	0,0010572	-29,4178	0,000
Access to electricity (% of p_N	0,01331	0,0009465	14,0678	0,000

Summary of Model

S = 0,291695 R-Sq = 94,04% R-Sq(adj) = 93,90%
PRESS = 11,6405 R-Sq(pred) = 93,63%

4.3 General Regression Analysis: PC3 versus Availability; GINNI P; Private prop

Regression Equation

PC3 = -0,238216 + 0,0398624 Availability of affordable ho_N - 0,0296013 GINNI P + 0,00746772 Private property rights (0=no_N

Coefficients

Term	Coef	SE Coef	T	P
Constant	-0,238216	0,111747	-2,1317	0,035
Availability of affordable ho_N	0,039862	0,001514	26,3240	0,000
GINNI P	-0,029601	0,001392	-21,2618	0,000
Private property rights (0=no_N	0,007468	0,001153	6,4786	0,000

Summary of Model

S = 0,322607 R-Sq = 91,47% R-Sq(adj) = 91,27%
PRESS = 14,1361 R-Sq(pred) = 90,95%

4.4 General Regression Analysis: PC4 versus GINNI P; Availability; Social Progr

```
Regression Equation

PC4  =  -2,67532 + 0,0368571 GINNI P + 0,0310721 Availability of affordable
        ho_N - 0,0184519 Social Progress Index_1N


Coefficients

Term                                 Coef    SE Coef         T      P
Constant                          -2,67532  0,0529286  -50,5458  0,000
GINNI P                            0,03686  0,0006658   55,3569  0,000
Availability of affordable ho_N    0,03107  0,0007222   43,0269  0,000
Social Progress Index_1N          -0,01845  0,0005787  -31,8841  0,000


Summary of Model

S = 0,151480     R-Sq = 97,12%        R-Sq(adj) = 97,05%
PRESS = 3,14325  R-Sq(pred) = 96,92%
```

4.4.1 REGRESSÃO PC4 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH

A exemplo do procedimento que realizamos para o PC1 estamos realizando nova regressão para o PC4 substituindo a variável SPI por IDH.
Como a queda de explicabilidade ficou em apenas 4,4% optamos por darmos continuidade em nossas análises com a variável IDH.

General Regression Analysis: PC4 versus GINNI P; Availability; IDH - 2013_1

```
Regression Equation

PC4  =  -2,59502 + 0,0364707 GINNI P + 0,0275097 Availability of affordable
        ho_N - 0,0159766 IDH - 2013_1_N


Coefficients

Term                                 Coef    SE Coef         T      P
Constant                          -2,59502  0,0850451  -30,5135  0,000
GINNI P                            0,03647  0,0010637   34,2858  0,000
Availability of affordable ho_N    0,02751  0,0011046   24,9046  0,000
```

```
IDH - 2013_1_N                  -0,01598  0,0008850  -18,0522  0,000


Summary of Model

S = 0,240553     R-Sq = 92,73%        R-Sq(adj) = 92,56%
PRESS = 7,91258  R-Sq(pred) = 92,24%
```

5. CONSIDERAÇÕES FINAIS

De acordo com todas as análises acima, percebemos claramente que o agrupamento de variáveis é bastante pertinente no caso das variáveis de moradia, independentemente da região analisada.

Isto pôde ser observado inicialmente pelas matrizes de correlação e dendogramas e depois comprovados pela análise dos componentes principais.

Assim, ao invés de trabalharmos com um grupo grande de variáveis (13) poderíamos utilizar apenas três ou quatro índices (PC1, PC2, PC3 e PC4) que as represente satisfatoriamente.

CAPITULO 8 – ANÁLISE DE CONGLOMERADOS

1. INTRODUÇÃO

Para a análise de conglomerados usaremos o software Minitab para criar uma nova classificação e agrupar os países da base nesses novos grupos.

No Minitab usamos a análise do componente principal para criar um novo índice, esse índice batizado como **MeDS (Moradia e Desenvolvimento Social)** é uma síntese das variáveis a seguir:

1. SPI Social Progress Index;
2. Shelter;
3. Access to improved Sanitation;

Stepwise Regression: PC1 versus Social Progress ; Shelter_1N; ...

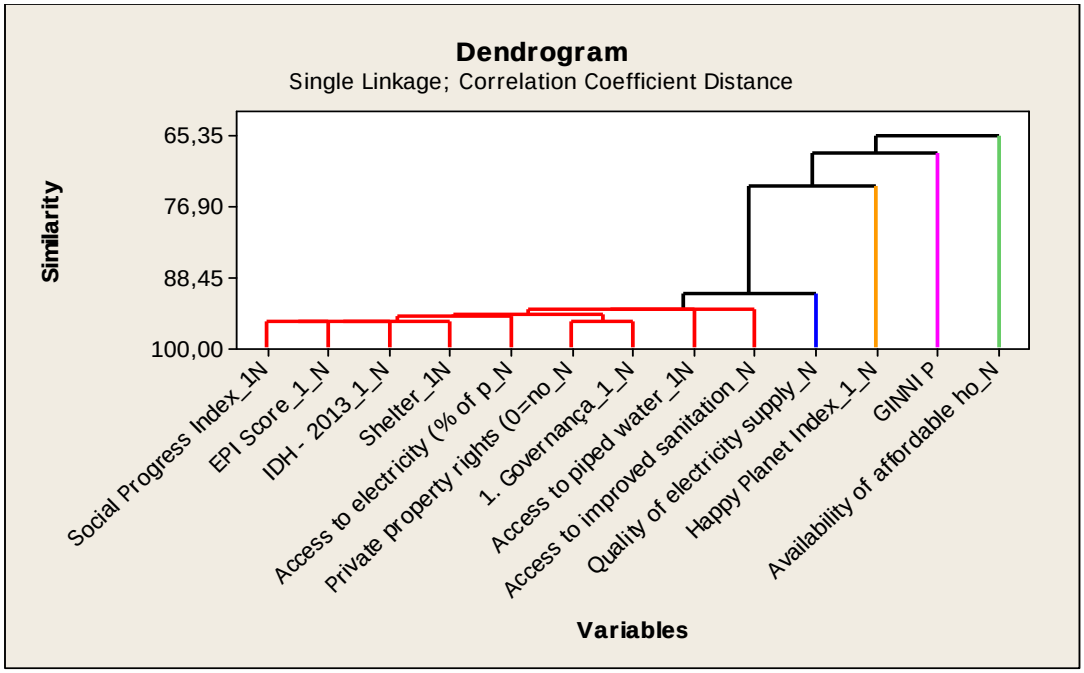
Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 13 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-6,054	-6,742	-6,767	-6,824	-6,813	-6,846
Social Progress Index_1N	0,1084	0,0615	0,0609	0,0456	0,0350	0,0332
T-Value	43,06	18,42	20,87	14,88	10,83	11,41
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
Shelter_1N		0,0514	0,0371	0,0380	0,0383	0,0293
T-Value		15,68	10,22	12,75	14,54	10,33
P-Value		0,000	0,000	0,000	0,000	0,000
Access to improved sanitation_N			0,0137	0,0176	0,0153	0,0118
T-Value			6,40	9,63	9,26	7,33
P-Value			0,000	0,000	0,000	0,000
S	0,735	0,433	0,378	0,310	0,274	0,246
R-Sq	93,45	97,75	98,29	98,86	99,12	99,30
R-Sq(adj)	93,40	97,71	98,25	98,83	99,08	99,26

O índice MeDS explica 98,29% do comportamento dessas variáveis, abaixo o Dendograma gerado com 5 clusters:

2. FORMAÇÃO DOS CONGLOMERADOS



Results for: CODING REGIONS MINITAB (1).MTW

Cluster Analysis of Variables: Social Progr; Shelter_1N; Access to pi; ...

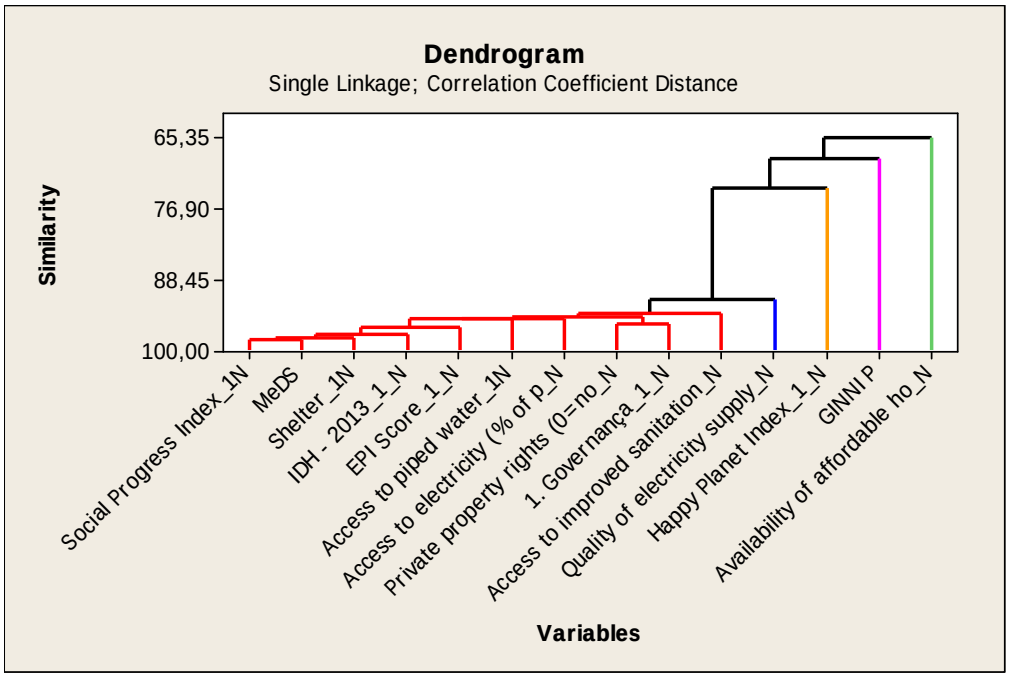
Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	12	95,8702	0,082596	1	11	1	2
2	11	95,8578	0,082843	1	9	1	3
3	10	95,7798	0,084405	1	2	1	4
4	9	95,6780	0,086441	8	10	8	2
5	8	94,7981	0,104038	1	6	1	5
6	7	94,4910	0,110180	1	8	1	7
7	6	93,9000	0,122000	1	3	1	8
8	5	93,8971	0,122058	1	4	1	9
9	4	91,2356	0,175288	1	7	1	10
10	3	73,8013	0,523973	1	12	1	11
11	2	68,2390	0,635220	1	13	1	12
12	1	65,3462	0,693077	1	5	1	13

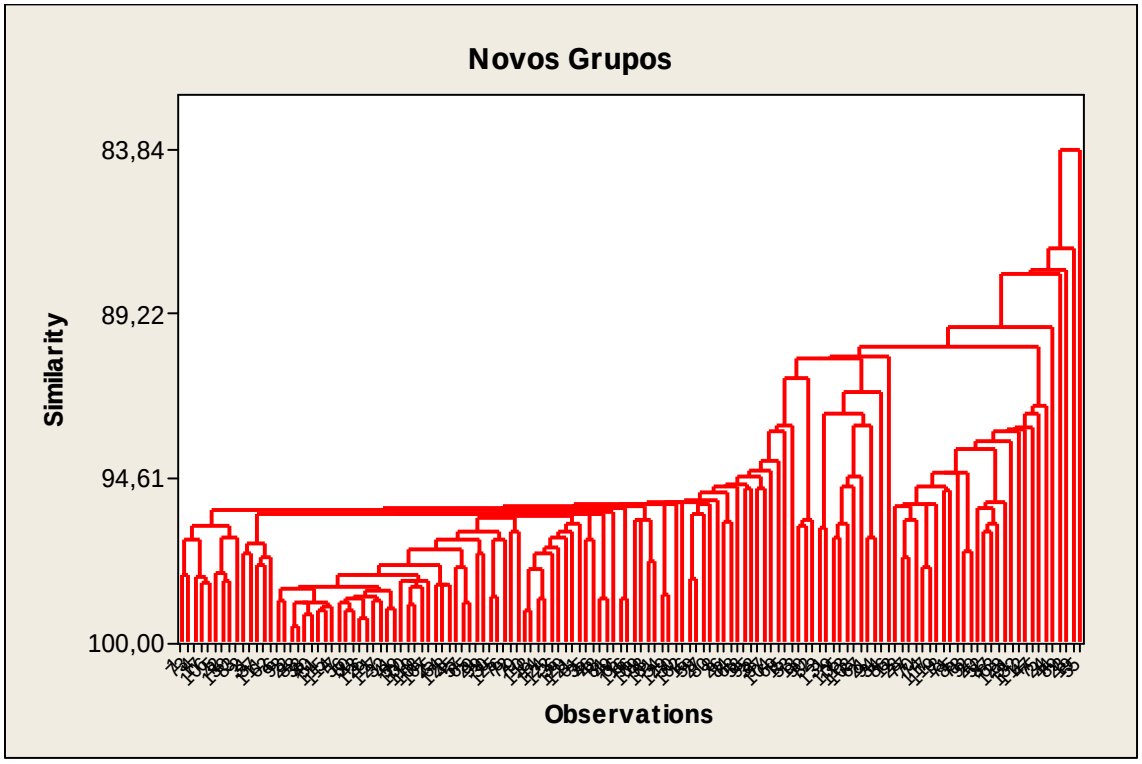
Final Partition

Cluster 1
  Social Progress Index_1N Shelter_1N Access to piped water_1N Access to improved sanitation_N Access to electricity (% of p_N Private property rights (0=no_N IDH - 2013_1_N 1. Governança_1_N EPI Score_1_N
Cluster 2
  Availability of affordable ho_N
Cluster 3
  Quality of electricity supply_N
Cluster 4
  Happy Planet Index_1_N
Cluster 5
  GINNI P

Abaixo o Dendograma com o novo índice MeDS e 5 clusters, onde podemos observar claramente o alto grau de similaridade do novo índice no primeiro cluster com 98,34% de explicabilidade.



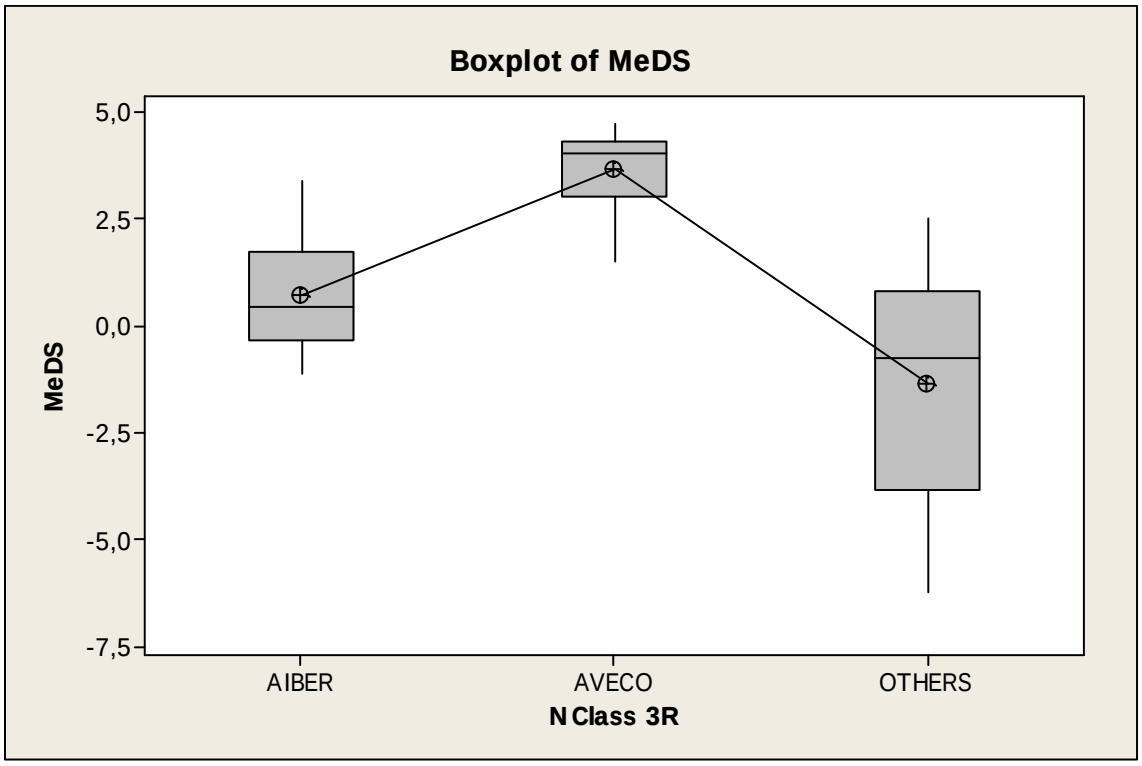
O Dendograma abaixo mostra a similaridade entre os países com relação as variáveis: Social Progress Index, Shelter e Access to Improved Sanitaion.



Temos Abaixo Anova com as novas Variáveis MeDS, McFA, UDF e IDM para as três regiões (AIBER, AVECO e Outros),

2.1 One-way ANOVA: MeDS versus N Class 3R

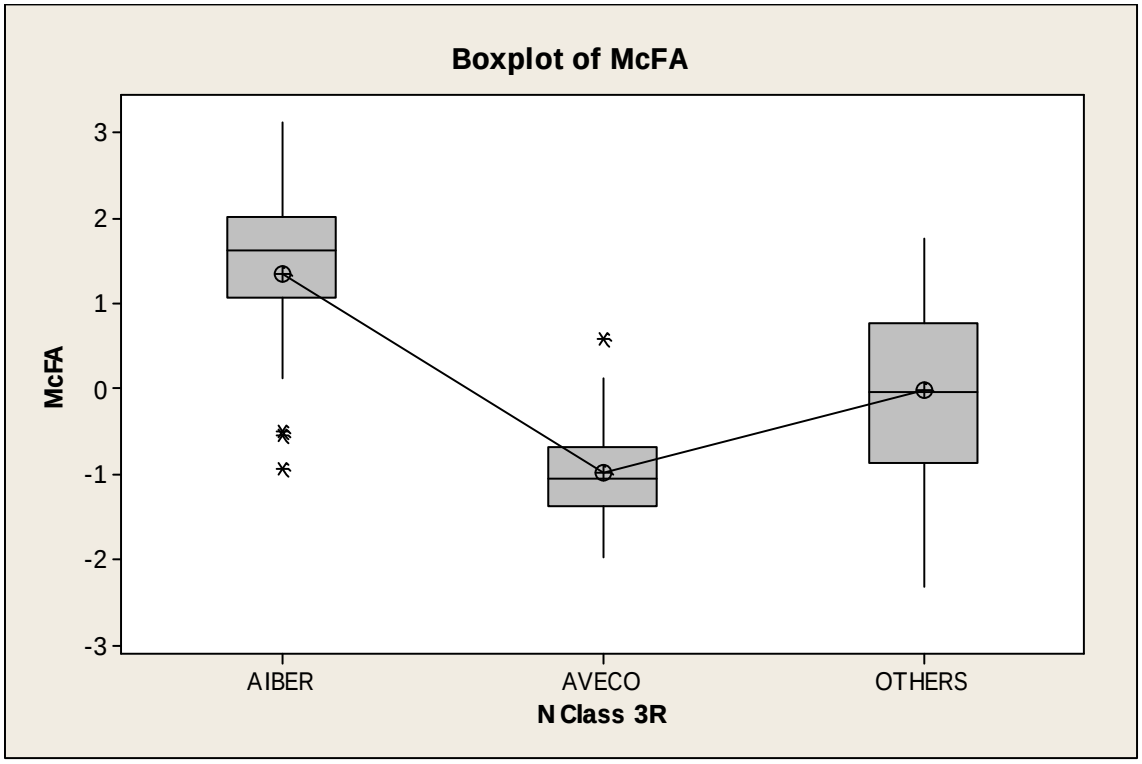
```
Source       DF       SS      MS      F      P
N Class 3R    2   532,38  266,19  63,51  0,000
Error       129   540,67    4,19
Total       131  1073,05

S = 2,047   R-Sq = 49,61%   R-Sq(adj) = 48,83%


                           Individual 95% CIs For Mean Based on Pooled StDev
Level    N    Mean  StDev    -+---------+---------+---------+--------
AIBER   21   0,736  1,367              (-----*----)
AVECO   27   3,674  0,826                                  (----*----)
OTHERS  84  -1,365  2,419    (-*--)
                             -+---------+---------+---------+--------
                           -1,6       0,0       1,6       3,2

Pooled StDev = 2,047
```



Podemos verificar que a Variável Moradia e Desenvolvimento Social (MeDS) está melhor aplicada nos países que compõem a região AVECO com os países mais desenvolvidos, seguido dos países do grupo AIBER que estão melhores posicionados em relação ao restante do mundo.

2.2 One-way ANOVA: McFA versus N Class 3R

```
Source       DF       SS      MS      F      P
N Class 3R    2   64,701  32,351  35,35  0,000
Error       129  118,068   0,915
Total       131  182,769

S = 0,9567   R-Sq = 35,40%   R-Sq(adj) = 34,40%

                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level    N     Mean   StDev  -------+---------+---------+---------+--
AIBER   21   1,3426  1,0220                                (----*----)
AVECO   27  -0,9968  0,5695  (----*---)
OTHERS  84  -0,0153  1,0340               (--*-)
                                -------+---------+---------+---------+--
                                    -0,80      0,00      0,80      1,60

Pooled StDev = 0,9567
```



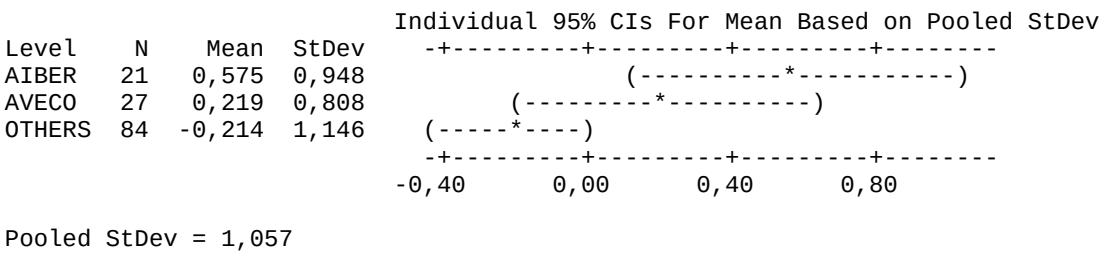
No boxplot acima Moradia como Fator de Alegria (McFA) podemos verificar que os grupos AIBER e Outros estão melhor posicionados que o grupo AVECO. Acreditamos que este fato ocorre, pois para este Componente Principal o MINITAB considerou a variável Happy Planet Index, que como sabemos puxa o grupo de países classificados como Desenvolvidos para baixo em comparação com os demais grupos.

Outliers AIBER: Portugal, Espanha e Uruguai

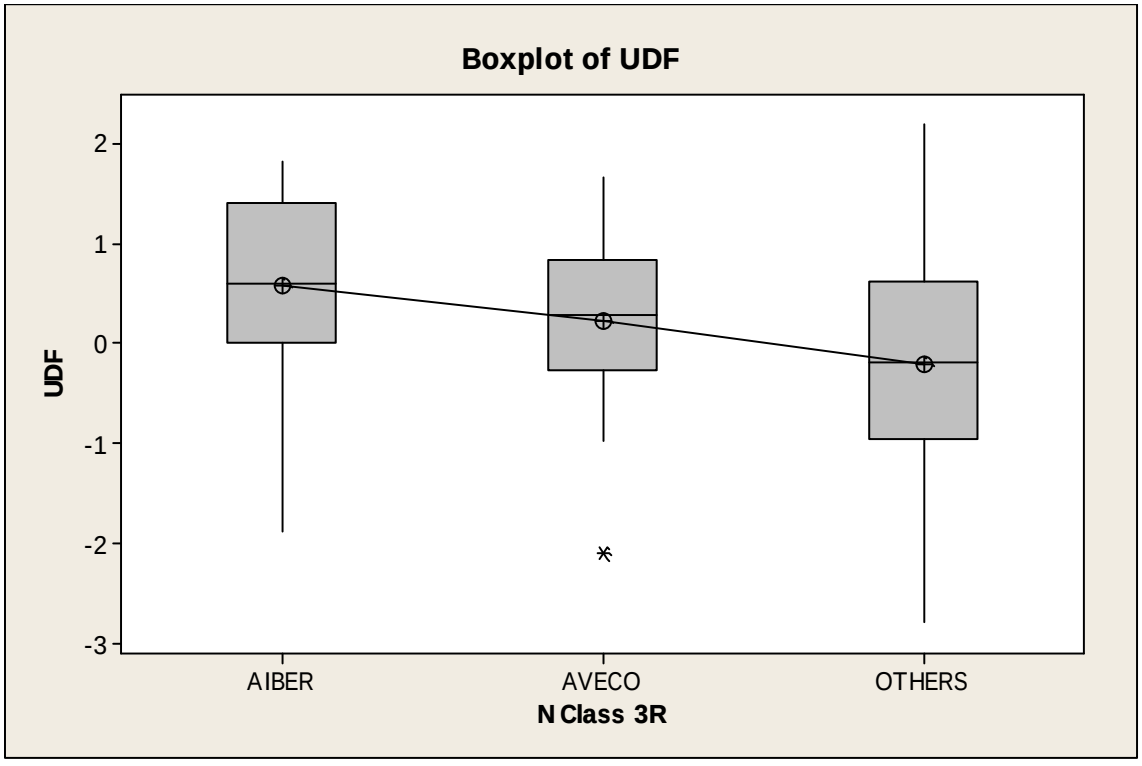
Outliers AVECO: Israel

2.3 One-way ANOVA: UDF versus N Class 3R

```
Source       DF      SS    MS     F      P
N Class 3R    2   12,09  6,05  5,41  0,006
Error       129  144,04  1,12
Total       131  156,13

S = 1,057   R-Sq = 7,74%   R-Sq(adj) = 6,31%
```




```
Pooled StDev = 1,057
```

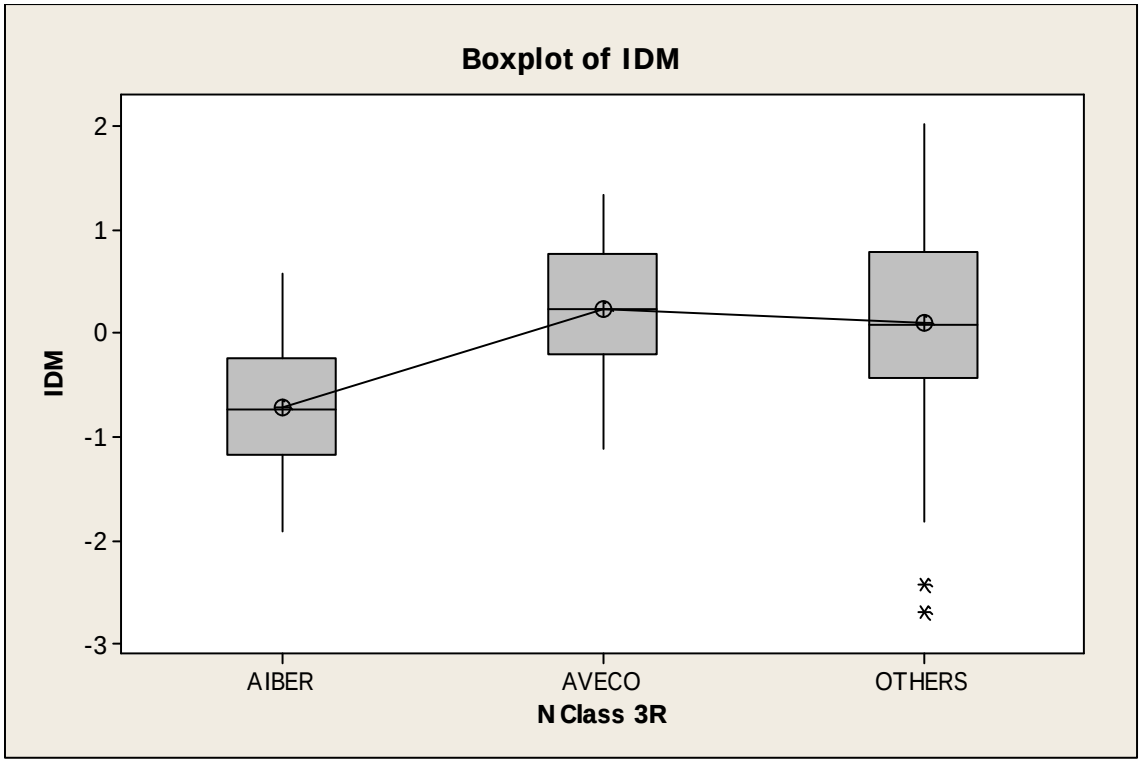


No boxplot Urbanização e Direitos Fundamentais (UDF) AIBER está um pouco acima de AVECO.
Outiler AVECO: Slovenia

2.4 One-way ANOVA: IDM versus N Class 3R

```
Source       DF       SS     MS     F      P
N Class 3R    2   13,181  6,590  9,58  0,000
Error       129   88,733  0,688
Total       131  101,914

S = 0,8294   R-Sq = 12,93%   R-Sq(adj) = 11,58%
```

```
                                Individual 95% CIs For Mean Based on Pooled StDev
Level    N     Mean   StDev     -+---------+---------+---------+--------
AIBER   21  -0,7160  0,6238     (------*------)
AVECO   27   0,2386  0,6219                     (------*-----)
OTHERS  84   0,1023  0,9242                     (---*---)
                                -+---------+---------+---------+--------
                               -1,00     -0,50      0,00      0,50

Pooled StDev = 0,8294
```

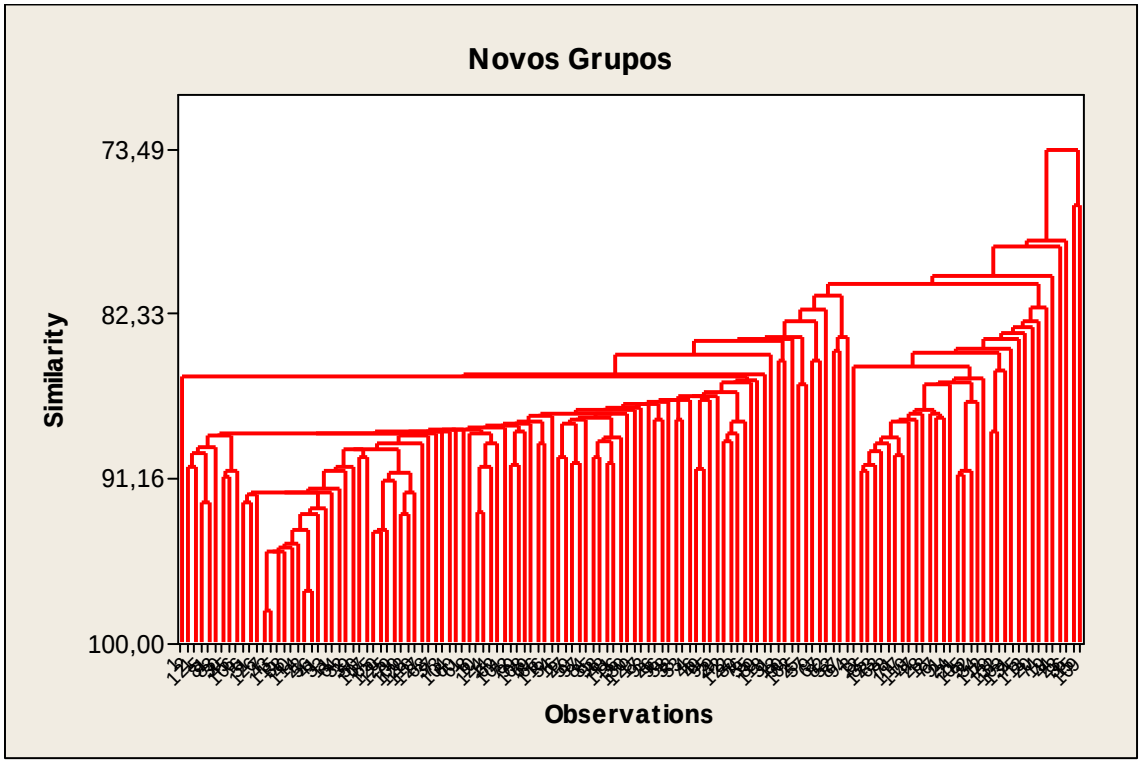


No boxplot acima Incentivos ao Direito de Moradia (IDM) podemos verificar que AIBER encontra-se atrás de AVECO e de Outros. Demonstrando que ainda há diversas resistências, inclusive governamentais para oferecer Incentivos aos Direitos de Moradia.

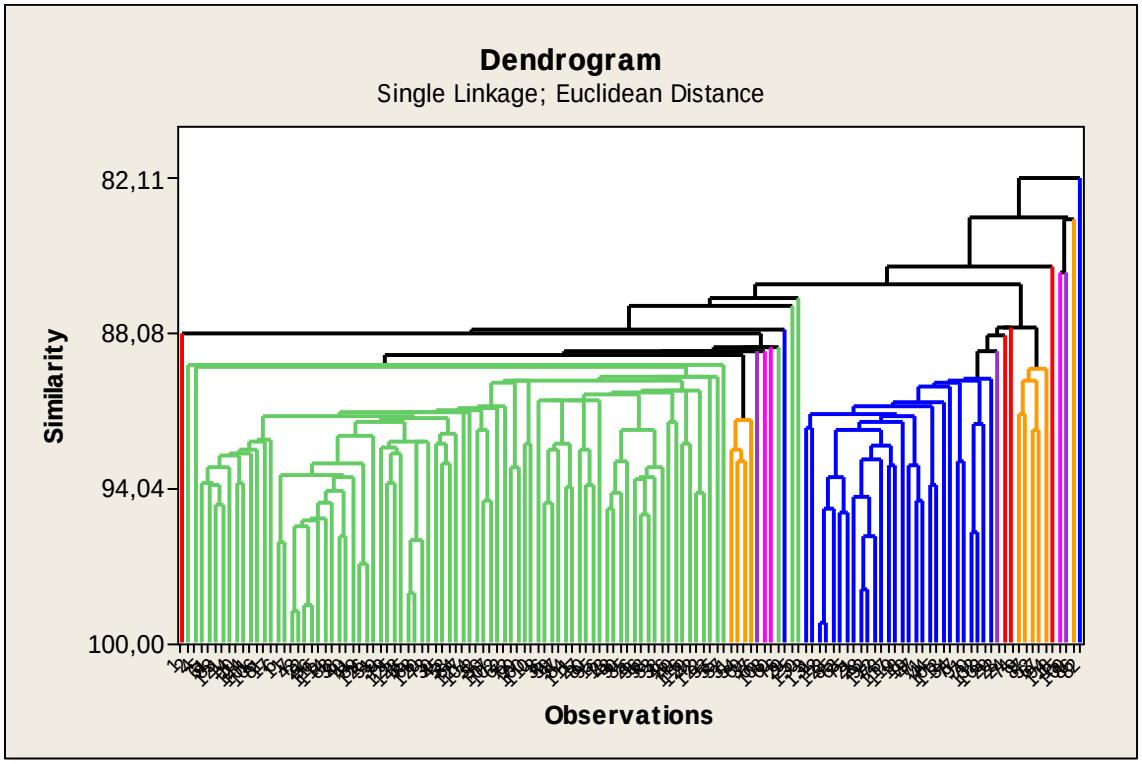
Outlier Outros: Botswana e South Africa.

3. Dendogramas com as principais variáveis de todos os Componentes Principais

PC1: SPI, Shelter e Access to Improved Sanitation;
PC2: HPI, Private Property Rights e Access to Eletricity;
PC3: Avaialability of Affordable House, Ginni e Private Property Rights;
PC4: Ginni, Avaialability of Affordable House e SPI

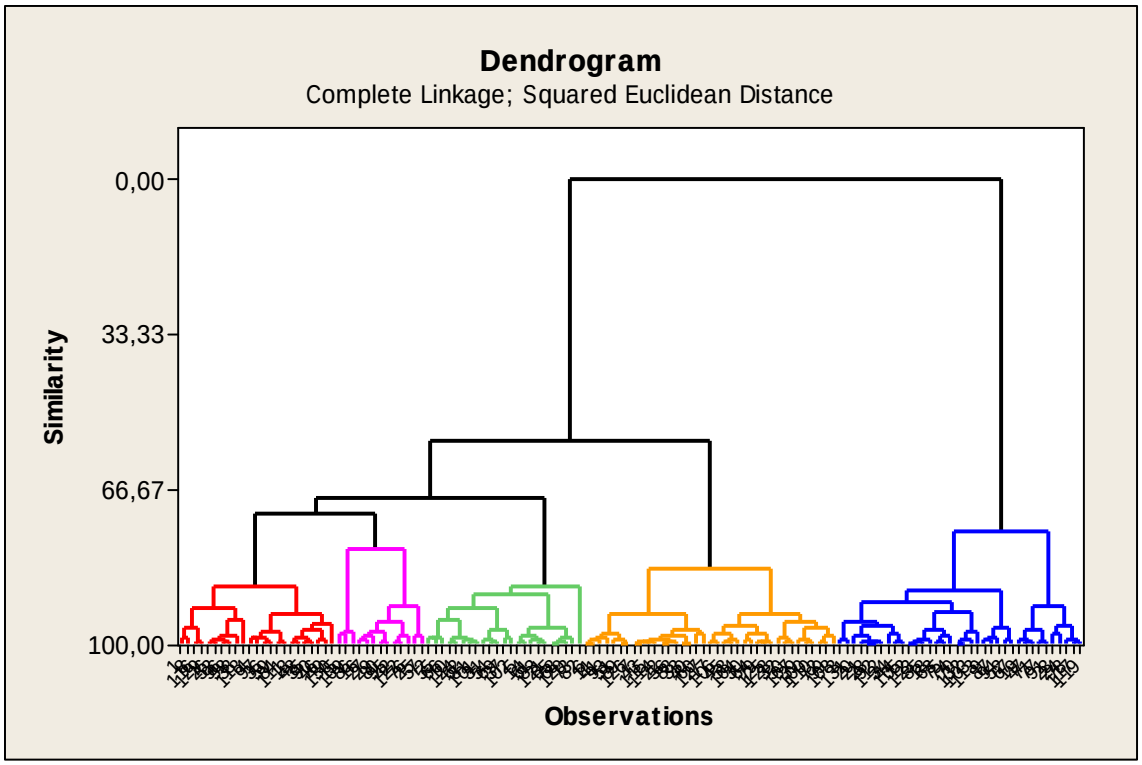


Abaixo o Dendograma com 20 clusters dos 4 Componentes Principais: PC1, PC2, PC3 e PC4



Podemos notar três grandes agrupamentos dentre os 20 selecionados.

O Dendograma abaixo com cinco clusters em Complete Linkage.



3.1 Cluster Analysis of Observations: MeDS; DMeQV; UDF; IDM

Squared Euclidean Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	99,9942	0,007	12	123	12	2
2	130	99,9853	0,018	7	43	7	2
3	129	99,9787	0,026	22	56	22	2
4	128	99,9628	0,046	65	68	65	2
5	127	99,9585	0,052	28	92	28	2
6	126	99,9097	0,112	89	126	89	2
7	125	99,8486	0,188	6	11	6	2
8	124	99,8351	0,205	33	40	33	2
9	123	99,8195	0,224	70	132	70	2
10	122	99,7705	0,285	65	120	65	3
11	121	99,7633	0,294	7	115	7	3
12	120	99,7536	0,306	96	98	96	2
13	119	99,7472	0,314	63	75	63	2
14	118	99,7305	0,335	13	48	13	2
15	117	99,7231	0,344	12	85	12	3
16	116	99,7152	0,354	83	121	83	2
17	115	99,7124	0,358	36	59	36	2
18	114	99,7027	0,369	28	112	28	3
19	113	99,7011	0,372	50	107	50	2
20	112	99,6987	0,375	47	71	47	2
21	111	99,6698	0,410	22	88	22	3
22	110	99,6682	0,412	42	122	42	2
23	109	99,6381	0,450	84	111	84	2
24	108	99,6346	0,454	80	95	80	2
25	107	99,6345	0,454	44	105	44	2
26	106	99,6198	0,473	10	124	10	2
27	105	99,6180	0,475	79	125	79	2
28	104	99,5974	0,500	26	53	26	2
29	103	99,5478	0,562	5	14	5	2
30	102	99,5406	0,571	72	99	72	2
31	101	99,5366	0,576	21	28	21	4
32	100	99,5348	0,578	117	119	117	2
33	99	99,5272	0,588	45	58	45	2
34	98	99,5168	0,601	77	91	77	2
35	97	99,5149	0,603	54	69	54	2
36	96	99,4605	0,671	35	90	35	2
37	95	99,4488	0,685	79	128	79	3
38	94	99,4169	0,725	100	110	100	2
39	93	99,4123	0,731	116	129	116	2
40	92	99,4031	0,742	81	101	81	2
41	91	99,3871	0,762	7	114	7	4
42	90	99,3845	0,765	17	106	17	2
43	89	99,3509	0,807	30	64	30	2
44	88	99,3381	0,823	61	83	61	3
45	87	99,3333	0,829	52	67	52	2
46	86	99,3143	0,852	3	131	3	2
47	85	99,3083	0,860	6	41	6	3
48	84	99,2846	0,889	46	49	46	2
49	83	99,2466	0,937	19	47	19	3
50	82	99,2449	0,939	51	89	51	3
51	81	99,2255	0,963	9	87	9	2
52	80	99,2238	0,965	78	117	78	3
53	79	99,1568	1,048	7	22	7	7
54	78	99,1224	1,091	16	27	16	2
55	77	99,1116	1,104	62	81	62	3
56	76	99,0744	1,151	70	103	70	3
57	75	99,0484	1,183	12	18	12	4
58	74	99,0378	1,196	26	96	26	4

59	73	99,0312	1,204	45	104	45	3
60	72	99,0234	1,214	31	54	31	3
61	71	98,9495	1,306	39	100	39	3
62	70	98,9282	1,332	13	35	13	4
63	69	98,8691	1,406	32	50	32	3
64	68	98,8629	1,413	4	97	4	2
65	67	98,8335	1,450	65	73	65	4
66	66	98,7616	1,539	2	55	2	2
67	65	98,7446	1,560	6	93	6	4
68	64	98,6205	1,715	37	102	37	2
69	63	98,6108	1,727	4	36	4	4
70	62	98,5883	1,755	1	8	1	2
71	61	98,5456	1,808	10	62	10	5
72	60	98,5323	1,824	9	94	9	3
73	59	98,5247	1,834	2	66	2	3
74	58	98,4950	1,871	26	38	26	5
75	57	98,4035	1,984	5	61	5	5
76	56	98,3939	1,996	20	21	20	5
77	55	98,3304	2,075	7	33	7	9
78	54	98,3261	2,081	25	57	25	2
79	53	98,3042	2,108	72	108	72	3
80	52	98,2719	2,148	16	29	16	3
81	51	98,1428	2,309	63	74	63	3
82	50	98,1048	2,356	42	76	42	3
83	49	98,0156	2,467	60	127	60	2
84	48	97,9793	2,512	34	44	34	3
85	47	97,9675	2,526	15	109	15	2
86	46	97,9046	2,605	6	51	6	7
87	45	97,7260	2,827	30	79	30	5
88	44	97,6703	2,896	17	45	17	5
89	43	97,6691	2,897	23	24	23	2
90	42	97,5156	3,088	70	113	70	4
91	41	97,3197	3,332	4	84	4	6
92	40	97,2847	3,375	16	80	16	5
93	39	97,2708	3,392	15	86	15	3
94	38	97,2060	3,473	12	63	12	7
95	37	97,1671	3,521	31	37	31	5
96	36	97,0940	3,612	19	77	19	5
97	35	97,0319	3,689	13	46	13	6
98	34	96,9214	3,827	7	60	7	11
99	33	96,6818	4,124	1	116	1	4
100	32	96,6183	4,203	5	65	5	9
101	31	96,5500	4,288	39	72	39	6
102	30	96,5066	4,342	9	52	9	5
103	29	96,3869	4,491	17	30	17	10
104	28	96,3798	4,500	20	34	20	8
105	27	96,1439	4,793	23	78	23	5
106	26	95,7109	5,331	13	130	13	7
107	25	95,3193	5,818	16	42	16	8
108	24	95,2767	5,871	3	20	3	10
109	23	95,2365	5,921	32	39	32	9
110	22	94,9605	6,264	26	118	26	6
111	21	94,8618	6,387	2	10	2	8
112	20	93,7651	7,750	4	13	4	13
113	19	93,6930	7,840	6	7	6	18
114	18	93,5565	8,009	17	32	17	19
115	17	93,4262	8,171	12	70	12	11
116	16	92,5689	9,237	1	26	1	10
117	15	92,4644	9,367	2	31	2	13
118	14	92,0846	9,839	19	23	19	10
119	13	91,8482	10,133	16	25	16	10
120	12	91,2446	10,883	3	12	3	21
121	11	89,2578	13,352	2	5	2	22
122	10	88,4086	14,408	3	9	3	26
123	9	87,8255	15,133	1	4	1	23
124	8	87,7382	15,241	2	82	2	23
125	7	83,9039	20,007	6	17	6	37
126	6	79,8227	25,080	15	16	15	13
127	5	75,9428	29,903	3	19	3	36
128	4	72,0295	34,767	1	15	1	36

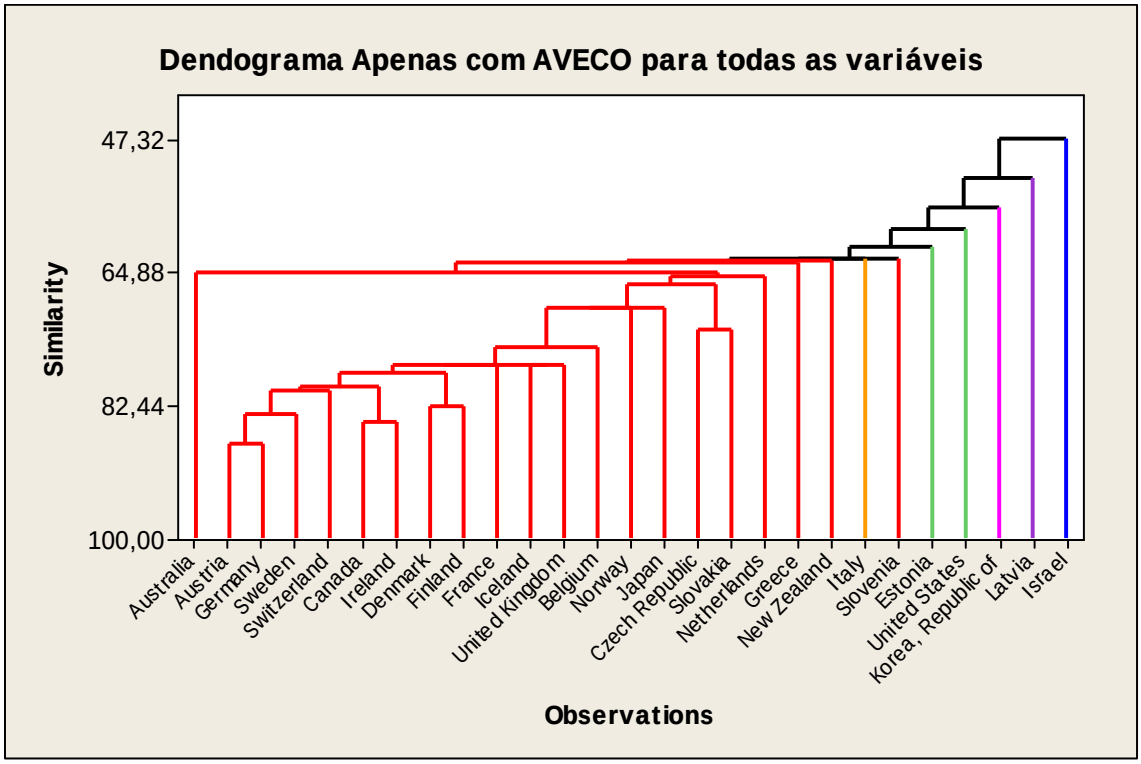
```
 129          3     68,7094    38,894    1    2        1       59
 130          2     56,3299    54,282    1    6        1       96
 131          1      0,0000   124,300    1    3        1      132


Final Partition
Number of clusters: 5


                                      Average   Maximum
                              Within  distance  distance
             Number of  cluster sum      from      from
          observations   of squares  centroid  centroid
Cluster1            23       49,989   1,40347   2,52997
Cluster2            23       51,720   1,40234   2,79859
Cluster3            36      108,444   1,61193   2,91441
Cluster4            37       90,389   1,48800   2,52497
Cluster5            13       52,967   1,89347   3,50307


Cluster Centroids

Variable  Cluster1  Cluster2  Cluster3  Cluster4  Cluster5  Grand centroid
MeDS      -0,13509   0,32812  -3,86814   3,27902   1,03768       0,0000000
McFA       1,50364   0,56054  -0,63457  -0,86181   0,55812      -0,0000000
UDF        0,60113  -1,53450   0,20759  -0,03395   1,17311       0,0000000
IDM        0,24538  -0,10884   0,16926   0,14413  -1,12049       0,0000000


Distances Between Cluster Centroids

          Cluster1  Cluster2  Cluster3  Cluster4  Cluster5
Cluster1   0,00000   2,40632   4,32068   4,20299   2,11238
Cluster2   2,40632   0,00000   4,70628   3,61200   2,97624
Cluster3   4,32068   4,70628   0,00000   7,15489   5,29955
Cluster4   4,20299   3,61200   7,15489   0,00000   3,17743
Cluster5   2,11238   2,97624   5,29955   3,17743   0,00000
```

Quadro 1: Relação de Principal Component analysis

	Variável 1	**Variável 2**	**Variável 3**	**Nova Variável**
PC1	**Social Progress Index ou IDH**	**Shelter**	**Access to Improved Sanitation**	**Moradia e Desenvolvimento Social (MeDS)**
PC2	**Happy Planet Index**	**Private Property Rights**	**Access to Eletricity**	**Moradia como Fator de Alegria (McFA)**
PC3	**Availability of Affordable House**	**Ginni**	**Private Property Rights**	**Urbanização e Direitos Fundamentais (UDF)**
PC4	**Ginni**	**Availability of Affordable House**	**Social Progress Index ou IDH**	**Incentivos ao Direito de Moradia (IDM)**

Após esta etapa iremos efetuar análises para as regiões AVECO e AIBER com todas as variáveis que compõem o estudo de MORADIA

4. ANÁLISE PARA AS REGIÕES AVECO, AIBER E OTHERS



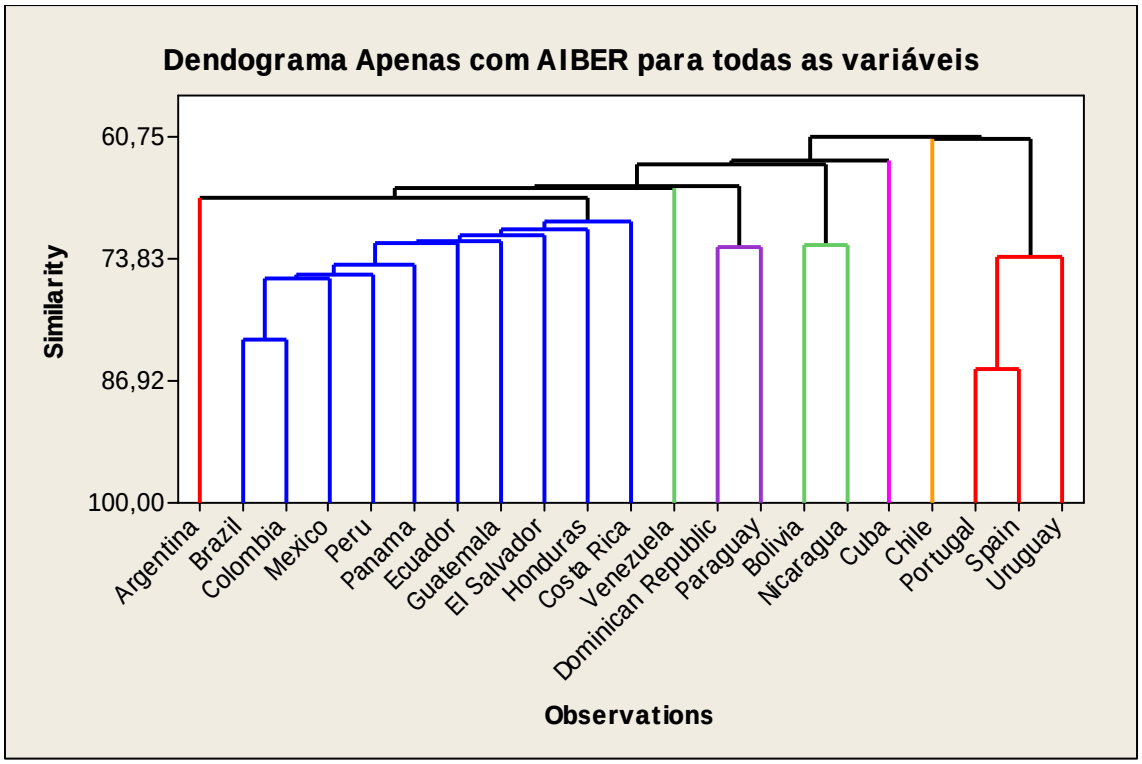
Podemos verificar que há um grande conglomerado de países no cluster vermelho.

4.1 Cluster Analysis of Observations: Social Progr; Shelter_1N; Access to pi; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	87,4668	12,2721	2	10	2	2
2	25	84,7037	14,9775	4	13	4	2
3	24	83,6135	16,0451	2	24	2	3
4	23	82,6622	16,9765	6	8	6	2
5	22	80,5829	19,0125	2	25	2	4
6	21	79,9576	19,6247	2	4	2	6
7	20	78,2113	21,3346	2	6	2	8
8	19	77,1929	22,3318	12	26	12	2
9	18	77,1909	22,3338	2	9	2	9
10	17	76,9847	22,5357	2	12	2	11
11	16	74,6853	24,7871	2	3	2	12
12	15	72,3760	27,0483	5	22	5	2
13	14	69,5899	29,7763	2	21	2	13
14	13	69,4568	29,9067	2	16	2	14
15	12	66,5694	32,7339	2	5	2	16
16	11	65,4932	33,7878	2	19	2	17
17	10	64,8119	34,4549	1	2	1	18
18	9	63,6174	35,6244	1	11	1	19
19	8	63,3645	35,8720	1	20	1	20
20	7	62,9885	36,2402	1	15	1	21
21	6	62,9341	36,2935	1	23	1	22
22	5	61,4425	37,7540	1	7	1	23
23	4	59,1867	39,9628	1	27	1	24
24	3	56,3385	42,7516	1	17	1	25
25	2	52,3235	46,6830	1	18	1	26
26	1	47,3184	51,5837	1	14	1	27

Final Partition - Number of clusters: 8



Podemos verificar um maior conglomerado de países no cluster azul.

Results for: Somente Aiber

Cluster Analysis of Observations: Social Progr; Shelter_1N; Access to pi; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	85,6916	21,2725	18	19	18	2
2	19	82,5704	25,9129	3	5	3	2
3	18	75,9762	35,7166	3	13	3	3
4	17	75,5203	36,3943	3	17	3	4
5	16	74,5005	37,9105	3	15	3	5
6	15	73,6098	39,2348	18	20	18	3
7	14	72,6820	40,6141	8	16	8	2
8	13	72,3423	41,1192	2	14	2	2
9	12	72,1987	41,3326	3	9	3	6
10	11	72,0235	41,5931	3	11	3	7
11	10	71,4135	42,5000	3	10	3	8
12	9	70,6982	43,5634	3	12	3	9
13	8	69,8948	44,7579	3	6	3	10
14	7	67,3795	48,4975	1	3	1	11
15	6	66,2735	50,1418	1	21	1	12
16	5	65,9843	50,5717	1	8	1	14
17	4	63,7343	53,9169	1	2	1	16
18	3	63,3269	54,5225	1	7	1	17
19	2	61,0195	57,9530	4	18	4	4
20	1	60,7486	58,3558	1	4	1	21

Final Partition
Number of clusters: 8

CAPITULO 9 – ANÁLISE DISCRIMINANTE

1. ANÁLISE DE COMPONENTES PRINCIPAIS

Na análise dos componentes principais, com base nas 13 variáveis pré-selecionadas para os países em questão, uma vez criadas as 4 variáveis PC1, PC2, PC3 e PC4, pôde-se perceber quais dos agrupamentos é melhor representativo para explicar as demais variáveis.
Ver Tabela 1.

Tabela 1: Indica o peso de cada variável nas componentes PC1, PC2, PC3 e PC4.

Variable	PC1	PC2	PC3	PC4
Social Progress Index_1N	0,338	-0,086	0,081	-0,058
Shelter_1N	0,335	0,120	0,003	-0,010
Access to piped water_1N	0,313	0,167	-0,113	-0,196
Access to improved sanitation_N	0,307	0,222	-0,159	0,009
Availability of affordable ho_N	0,096	0,024	0,717	0,552
Access to electricity (% of p_N	0,299	0,330	-0,156	0,067
Quality of electricity supply_N	0,291	-0,157	0,010	-0,019
Private property rights (0=no_N	0,258	-0,431	0,204	-0,062
IDH - 2013_1_N	0,332	0,064	-0,071	-0,068
1. Governança_1_N	0,288	-0,382	0,192	-0,117
EPI Score_1_N	0,324	-0,125	-0,064	-0,083
Happy Planet Index_1_N	0,124	0,619	0,266	0,080
GINNI P	0,132	-0,177	-0,507	0,783

Análise:

- Analisando PC1. Podemos verificar que não há nenhuma variável negativa. A variável com maior grau de contribuição é SPI com 0,338 seguida de Shelter 0,335 e IDH 0,332.
- Analisando PC2. Há 06 variáveis com contribuição negativa: Social Progress, Quality of eletricity, Private Property Rights, Governança, EPI Score e HPI; a variável com maior grau de contribuição é Happy Planet Index com 0,619.
- Analisando PC3. Há 6 variáveis com contribuição negativa; a variável com maior grau de contribuição é Availability of Affordable House.
- Analisando PC4. Há 08 variáveis com contribuição negativa; a variável com maior grau de contribuição é Ginni

Tabela 2. Matriz de Correlações

Principal Component Analysis: Social Progr; Shelter_1N; Access to pi; Access to

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,1912	1,3952	1,1919	0,7780	0,4933	0,3318	0,1617	0,1352
Proportion	0,630	0,107	0,092	0,060	0,038	0,026	0,012	0,010
Cumulative	0,630	0,737	0,829	0,889	0,927	0,952	0,965	0,975

Eigenvalue	0,1284	0,0831	0,0609	0,0315	0,0178
Proportion	0,010	0,006	0,005	0,002	0,001
Cumulative	0,985	0,992	0,996	0,999	1,000

Análise:

Segundo Las Casas e Guevara (2010)[2], por convenção só se deve considerar componentes com contribuição >ou= 1, pois só acima ou igual a 1 é que ajuda a explicar o fenômeno. Neste caso, só PC1, PC2 e PC3 têm Eigenvalue que atendem a esse critério. Porém, incluiremos o PC4 por uma questão de ampliar o espectro da análise, apesar de se perceber que não se altera significativamente (todos os cálculos foram feitos dessa forma anteriormente), também tais agrupamentos explicam, cumulativamente 88,9% das variáveis. Observa-se o peso de explicação de PC1 em 63,0%. Quando se passa para o grupo de variáveis de PC2, há um acréscimo razoável de 10,7% de poder de explicação; de PC2 para PC3, há um acúmulo inferior de apenas 9,2% de poder de explicação, com o qual, essas primeiras variáveis são suficientes para explicar os indicadores em questão de Moradia dos Países pesquisados para os 20 indicadores pré-selecionados. Para comprovar que os PC's são complementares, calculamos a correlação entre eles. Ver tabela 3.

Tabela 3. Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	-0,000		
	1,000		
PC3	0,000	-0,000	
	1,000	1,000	
PC4	-0,000	0,000	0,000
	1,000	1,000	1,000

Cell Contents: Pearson correlation
P-Value

Análise:

Correlação = 0 nula o que significa que

Cada variável está explicando dados variados, portanto trata-se de explicações complementares.

1.1 Regressão PC1

Regression Analysis: MeDS (PC1) versus Social Progr; Shelter_1N; Access to im

The regression equation is
MeDS = - 6,77 + 0,0609 Social Progress Index_1N + 0,0371 Shelter_1N
+ 0,0137 Access to improved sanitation_N

Predictor	Coef	SE Coef	T	P
Constant	-6,76656	0,08834	-76,59	0,000
Social Progress Index_1N	0,060923	0,002919	20,87	0,000
Shelter_1N	0,037139	0,003634	10,22	0,000
Access to improved sanitation_N	0,013739	0,002147	6,40	0,000

S = 0,378370 R-Sq = 98,3% R-Sq(adj) = 98,3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	1054,72	351,57	2455,75	0,000
Residual Error	128	18,32	0,14		
Total	131	1073,05			

Source	DF	Seq SS
Social Progress Index_1N	1	1002,75
Shelter_1N	1	46,11
Access to improved sanitation_N	1	5,86

Unusual Observations

Obs	Social Progress Index_1N	MeDS	Fit	SE Fit	Residual	St Resid
4	68	0,3164	1,0742	0,0738	-0,7578	-2,04R
19	9	-5,2858	-5,4347	0,1150	0,1489	0,41 X
52	32	-2,1468	-2,9962	0,0571	0,8494	2,27R
55	22	-0,7301	-1,5244	0,1302	0,7943	2,24RX
57	70	2,7674	1,9636	0,0458	0,8038	2,14R
91	13	-4,6209	-5,5086	0,0816	0,8877	2,40R
94	18	-2,7466	-3,7490	0,0632	1,0024	2,69R

```
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

1.1.1 REGRESSÃO PC1 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH

Há um consenso de que há muita complexidade na composição da Variável Social Progress Index, por este motivo estamos rodando nova simulação com a variável IDH para verificarmos seu potencial e possibilidade de seguirmos adiante em nossa análise com o IDH no lugar de Social Progress Index, uma vez que estas variáveis possuem em sua gênese grande grau de correlação.

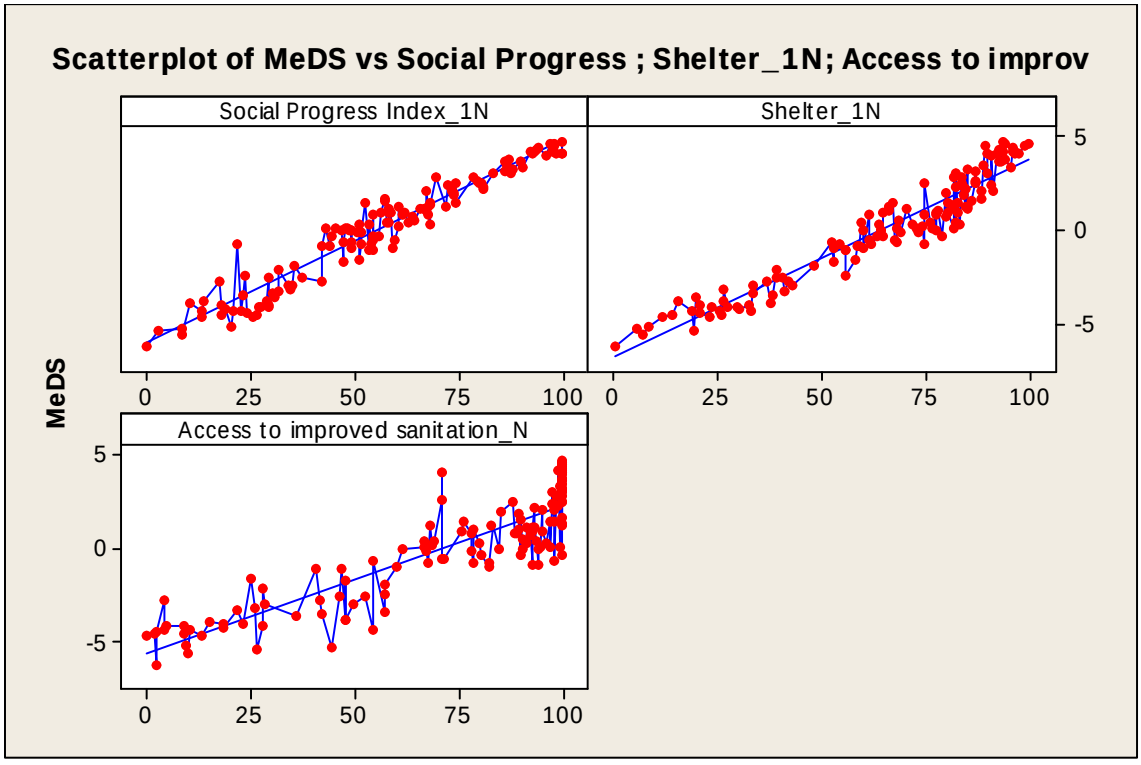
Era esperado obtermos R^2 inferior de IDH em relação a SPI. No nosso caso na primeira regressão tínhamos grau de explicabilidade de 98,3%, enquanto na segunda regressão a explicabilidade ficou em 95,2%, portanto menos 3,1%. Sendo assim com a pequena queda de explicabilidade optaremos por darmos continuidade em nossas análises à partir das variáveis que compõem a segunda regressão de MeDS (PC1) com IDH.

Regression Analysis: MeDS (PC1) versus IDH - 2013_1; Shelter_1N; Access to im

```
The regression equation is
MeDS = - 6,78 + 0,0482 IDH - 2013_1_N + 0,0561 Shelter_1N
       + 0,00520 Access to improved sanitation_N

Predictor                          Coef   SE Coef       T      P
Constant                        -6,7770    0,1474  -45,97  0,000
IDH - 2013_1_N                 0,048158  0,005580    8,63  0,000
Shelter_1N                     0,056148  0,005842    9,61  0,000
Access to improved sanitation_N 0,005201  0,003760    1,38  0,169


S = 0,631315   R-Sq = 95,2%   R-Sq(adj) = 95,1%
Analysis of Variance

Source           DF       SS      MS       F      P
Regression        3  1022,03  340,68  854,77  0,000
Residual Error  128    51,02    0,40
Total           131  1073,05


Source                          DF  Seq SS
IDH - 2013_1_N                   1  970,47
Shelter_1N                       1   50,80
```

```
Access to improved sanitation_N   1    0,76


Unusual Observations

        IDH -
Obs  2013_1_N     MeDS      Fit   SE Fit  Residual  St Resid
  8        68   0,2812   1,5671  0,0846   -1,2860     -2,06R
 19         9  -5,2858  -5,8305  0,1921    0,5448      0,91 X
 28         0  -4,1102  -5,4195  0,1532    1,3093      2,14R
 75        13  -3,7793  -5,0539  0,1629    1,2746      2,09R
 88        58   4,0487   1,9642  0,1813    2,0845      3,45R
 91         0  -4,6209  -6,1167  0,1406    1,4958      2,43R
102        73  -0,1053   1,1795  0,0910   -1,2849     -2,06R
115        92   4,7134   3,4606  0,0953    1,2528      2,01R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Normplot of Residuals for MeDS



1.2 Regressão PC2

Regression Analysis: McFA (PC2) versus Happy Planet Ind; Private property; ...

The regression equation is
McFA = - 1,28 + 0,0328 Happy Planet Index_1_N
- 0,0311 Private property rights (0=no_N
+ 0,0133 Access to electricity (% of p_N

Predictor	Coef	SE Coef	T	P
Constant	-1,27550	0,07650	-16,67	0,000
Happy Planet Index_1_N	0,032779	0,001359	24,12	0,000
Private property rights (0=no_N	-0,031102	0,001057	-29,42	0,000
Access to electricity (% of p_N	0,0133149	0,0009465	14,07	0,000

S = 0,291695 R-Sq = 94,0% R-Sq(adj) = 93,9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	171,878	57,293	673,35	0,000
Residual Error	128	10,891	0,085		
Total	131	182,769			

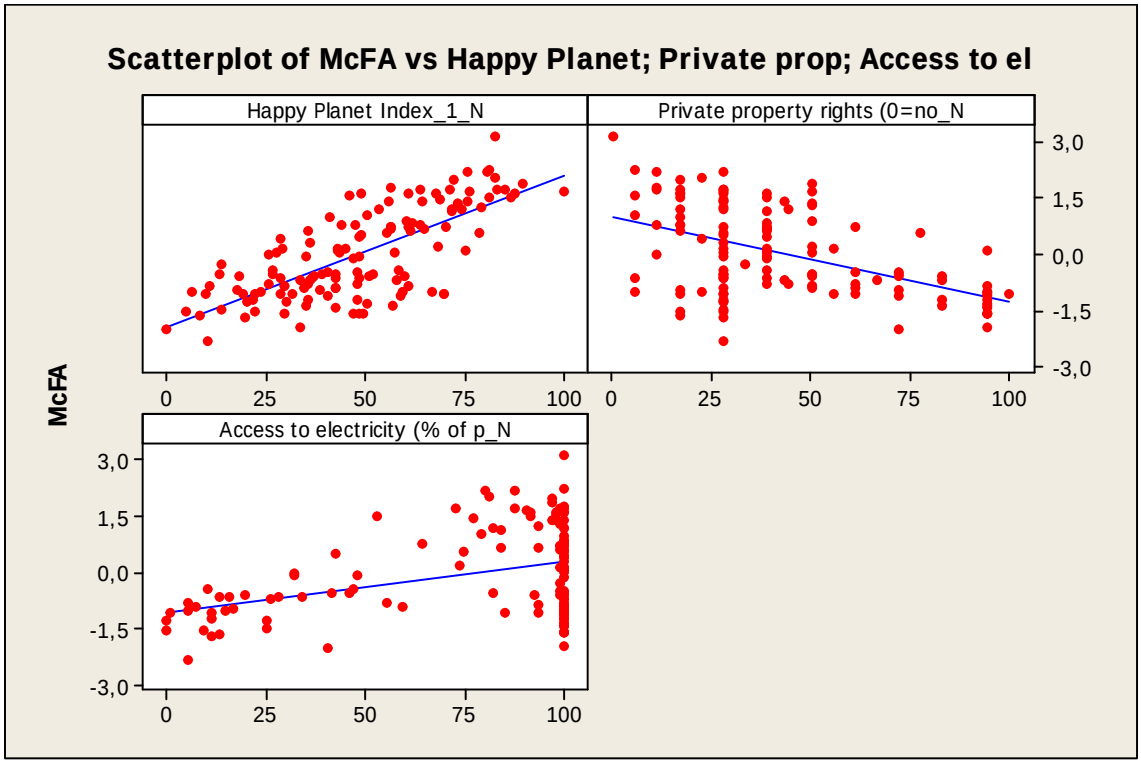
Source	DF	Seq SS
Happy Planet Index_1_N	1	97,832
Private property rights (0=no_N	1	57,207
Access to electricity (% of p_N	1	16,839

Unusual Observations

Obs	Happy Planet Index_1_N	McFA	Fit	SE Fit	Residual	St Resid
15	0	-1,9863	-2,9807	0,0719	0,9944	3,52R
28	29	-0,6686	-0,0450	0,0502	-0,6236	-2,17R
49	81	2,1650	1,5688	0,0503	0,5962	2,07R
55	64	1,3916	0,7884	0,0327	0,6031	2,08R
91	10	-2,3346	-1,7345	0,0643	-0,6001	-2,11R
109	14	-0,5617	-1,2921	0,0553	0,7303	2,55R

R denotes an observation with a large standardized residual.

Normplot of Residuals for McFA



1.3 Regressão PC3

Regression Analysis: UDF (PC3) versus Availability; GINNI P; Private prop

The regression equation is

UDF = - 0,238 + 0,0399 Availability of affordable ho_N - 0,0296 GINNI P
+ 0,00747 Private property rights (0=no_N

Predictor	Coef	SE Coef	T	P
Constant	-0,2382	0,1117	-2,13	0,035
Availability of affordable ho_N	0,039862	0,001514	26,32	0,000
GINNI P	-0,029601	0,001392	-21,26	0,000
Private property rights (0=no_N	0,007468	0,001153	6,48	0,000

S = 0,322607 R-Sq = 91,5% R-Sq(adj) = 91,3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	142,813	47,604	457,40	0,000
Residual Error	128	13,322	0,104		
Total	131	156,135			

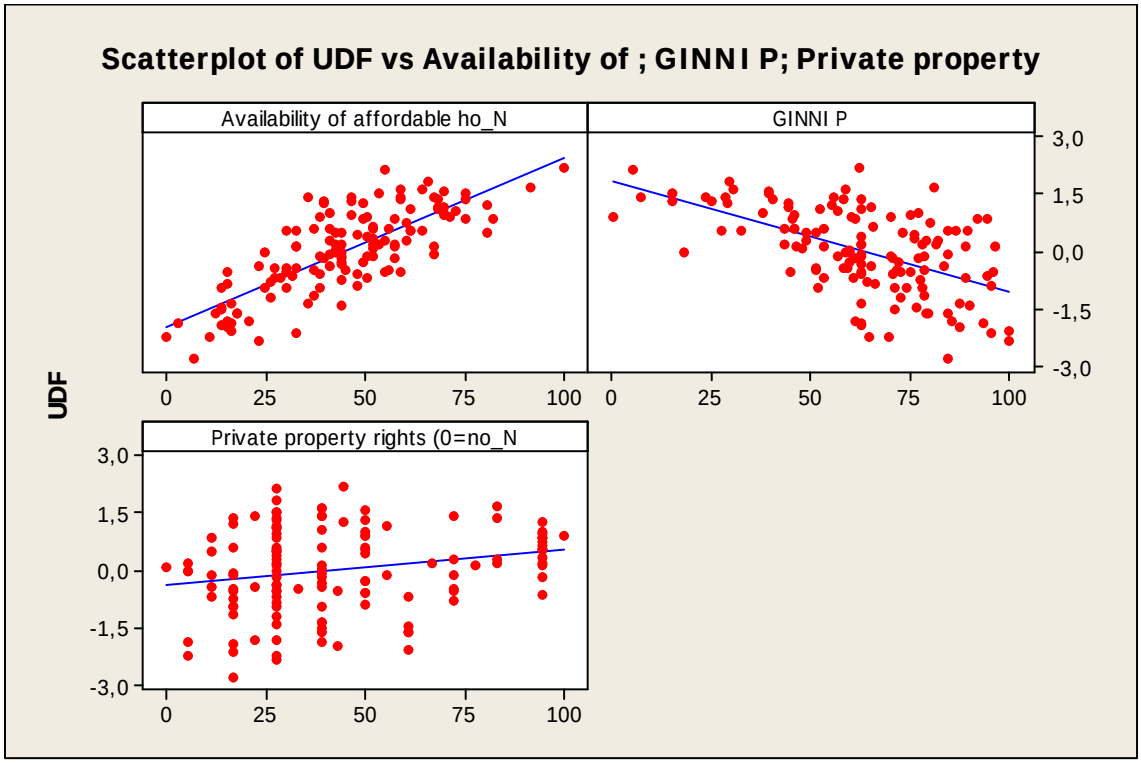
Source	DF	Seq SS
Availability of affordable ho_N	1	95,743
GINNI P	1	42,703
Private property rights (0=no_N	1	4,368

Unusual Observations

Obs	Availability of affordable ho_N	UDF	Fit	SE Fit	Residual	St Resid
52	48	0,4321	-0,2120	0,0337	0,6442	2,01R
65	48	-0,5688	0,1838	0,0292	-0,7526	-2,34R
71	23	-0,3818	-1,0321	0,0428	0,6503	2,03R
73	45	-0,4662	0,3002	0,0330	-0,7665	-2,39R
74	64	1,6392	0,8798	0,0421	0,7595	2,37R
109	38	0,9222	1,6643	0,0962	-0,7421	-2,41R

R denotes an observation with a large standardized residual.

Normplot of Residuals for UDF



1.4 Regressão PC4

Regression Analysis: IDM versus GINNI P; Availability; Social Progr

The regression equation is

IDM = - 2,68 + 0,0369 GINNI P + 0,0311 Availability of affordable ho_N
- 0,0185 Social Progress Index_1N

Predictor	Coef	SE Coef	T	P
Constant	-2,67532	0,05293	-50,55	0,000
GINNI P	0,0368571	0,0006658	55,36	0,000
Availability of affordable ho_N	0,0310721	0,0007222	43,03	0,000
Social Progress Index_1N	-0,0184519	0,0005787	-31,88	0,000

S = 0,151480 R-Sq = 97,1% R-Sq(adj) = 97,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	98,977	32,992	1437,80	0,000
Residual Error	128	2,937	0,023		
Total	131	101,914			

Source	DF	Seq SS
GINNI P	1	48,638
Availability of affordable ho_N	1	27,012
Social Progress Index_1N	1	23,327

Unusual Observations

Obs	GINNI P	IDM	Fit	SE Fit	Residual	St Resid
9	63	1,5130	1,1275	0,0275	0,3855	2,59R
15	7	-2,4264	-2,0647	0,0400	-0,3617	-2,48R
53	65	1,2968	0,9826	0,0238	0,3142	2,10R
98	52	0,4399	0,0712	0,0186	0,3688	2,45R

R denotes an observation with a large standardized residual.

1.4.1 REGRESSÃO PC4 - SUBSTITUINDO SOCIAL PROGRESS INDEX POR IDH

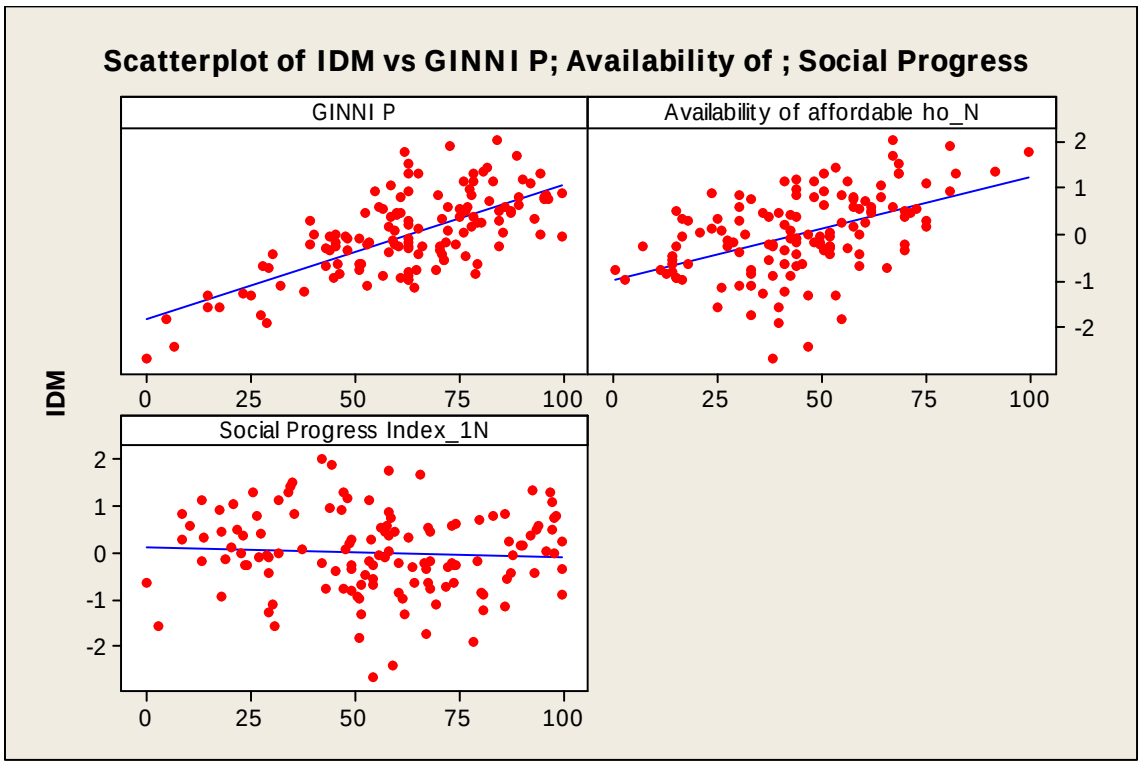
A exemplo do procedimento que realizamos para o PC1 estamos realizando nova regressão para o PC4 substituindo a variável SPI por IDH.
Como a queda de explicabilidade ficou em apenas 4,4% optamos por darmos continuidade em nossas análises com a variável IDH.

Regression Analysis: IDM versus GINNI P; Availability; IDH - 2013_1

```
The regression equation is
IDM = - 2,60 + 0,0365 GINNI P + 0,0275 Availability of affordable ho_N
      - 0,0160 IDH - 2013_1_N


Predictor                             Coef    SE Coef       T      P
Constant                          -2,59502    0,08505  -30,51  0,000
GINNI P                           0,036471   0,001064   34,29  0,000
Availability of affordable ho_N   0,027510   0,001105   24,90  0,000
IDH - 2013_1_N                  -0,0159766  0,0008850  -18,05  0,000


S = 0,240553   R-Sq = 92,7%   R-Sq(adj) = 92,6%


Analysis of Variance

Source           DF       SS      MS       F      P
Regression        3   94,507  31,502  544,40  0,000
Residual Error  128    7,407   0,058
Total           131  101,914


Source                           DF  Seq SS
GINNI P                           1  48,638
Availability of affordable ho_N   1  27,012
IDH - 2013_1_N                    1  18,857


Unusual Observations

Obs  GINNI P      IDM      Fit  SE Fit  Residual  St Resid
  9       63   1,5130   1,0025  0,0414    0,5105      2,15R
 31       63  -0,9820  -1,4827  0,0572    0,5007      2,14R
 53       65   1,2968   0,7551  0,0341    0,5417      2,27R
```

```
 88        89    0,7761    1,5100   0,0430    -0,7339      -3,10R
130        48   -0,0512   -0,5387   0,0306     0,4875       2,04R

R denotes an observation with a large standardized residual.
```

Normplot of Residuals for IDM



54 Países	Cluster 1	23 Países	Cluster 2	36 Países	Cluster 3	19 Países	Cluster 4
Albania	1	Angola	2	Australia	3	Botswana	4
Algeria	1	Benin	2	Austria	3	Brazil	4
Argentina	1	Burkina Faso	2	Belgium	3	Chile	4
Armenia	1	Burundi	2	Canada	3	Colombia	4
Azerbaijan	1	Cambodia	2	Croatia	3	Costa Rica	4
Bangladesh	1	Cameroon	2	Czech Republic	3	Djibouti	4
Belarus	1	Central African Rep	2	Denmark	3	Ghana	4
Bolivia	1	Chad	2	Estonia	3	Israel	4
Bosnia and Herz	1	Congo, Republic of	2	Finland	3	Kenya	4
Bulgaria	1	Guinea	2	France	3	Lesotho	4
China	1	Liberia	2	Germany	3	Malawi	4
Cuba	1	Madagascar	2	Greece	3	Mexico	4
Dominican Republic	1	Mali	2	Hungary	3	Namibia	4
Ecuador	1	Mauritania	2	Iceland	3	Panama	4
Egypt	1	Mongolia	2	Ireland	3	Rwanda	4
El Salvador	1	Mozambique	2	Italy	3	Senegal	4
Georgia	1	Niger	2	Japan	3	South Afr	4
Guatemala	1	Nigeria	2	Korea, Republic of	3	Swaziland	4
Guyana	1	Sudan	2	Lithuania	3	Zambia	4
Honduras	1	Tanzania	2	Malaysia	3		
India	1	Togo	2	Mauritius	3		
Indonesia	1	Uganda	2	Netherlands	3		
Iran	1	Yemen	2	New Zealand	3		
Iraq	1			Norway	3		
Jamaica	1			Poland	3		
Jordan	1			Portugal	3		
Kazakhstan	1			Saudi Arabia	3		
Kuwait	1			Slovakia	3		
Kyrgyzstan	1			Slovenia	3		
Laos	1			Spain	3		
Latvia	1			Sweden	3		
Lebanon	1			Switzerland	3		
Macedonia	1			United Arab Emir.	3		
Moldova	1			United Kingdom	3		
Montenegro	1			United States	3		
Morocco	1			Uruguay	3		
Nepal	1						
Nicaragua	1						
Pakistan	1						
Paraguay	1						
Peru	1						
Philippines	1						
Romania	1						
Russia	1						
Serbia	1						
Sri Lanka	1						

Tajikistan	1
Thailand	1
Trinidad and Tobago	1
Tunisia	1
Turkey	1
Ukraine	1
Uzbekistan	1
Venezuela	1

2. FUNÇÃO LINEAR

2.1 Discriminant Analysis: 4 clusters versus Social Progr; Shelter_1N; ...

Linear Method for Response: 4 clusters

Predictors: Social Progress Index_1N; Shelter_1N; Access to piped water_1N;
Access to improved sanitation_N; Availability of affordable ho_N;
Access to electricity (% of p_N; Quality of electricity supply_N;
Private property rights (0=no_N; IDH - 2013_1_N; 1. Governança_1_N;
EPI Score_1_N; Happy Planet Index_1_N; GINNI P

Group	1	2	3	4
Count	54	23	36	19

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	51	1	1	1
2	0	22	0	3
3	2	0	35	1
4	1	0	0	14
Total N	54	23	36	19
N correct	51	22	35	14
Proportion	0,944	0,957	0,972	0,737

N = 132 N Correct = 122 Proportion Correct = 0,924

Squared Distance Between Groups

	1	2	3	4
1	0,0000	23,4169	17,5543	12,4762
2	23,4169	0,0000	38,5339	15,3331
3	17,5543	38,5339	0,0000	18,8219
4	12,4762	15,3331	18,8219	0,0000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-31,106	-12,804	-47,937	-18,996
Social Progress Index_1N	-0,171	-0,173	-0,235	-0,280
Shelter_1N	-0,026	0,099	0,109	0,112
Access to piped water_1N	-0,001	0,019	0,042	0,031
Access to improved sanitation_N	0,119	-0,014	0,107	0,103
Availability of affordable ho_N	0,184	0,137	0,206	0,141
Access to electricity (% of p_N	0,207	-0,058	0,027	0,013

Quality of electricity supply_N	0,052	0,006	0,054	0,033
Private property rights (0=no_N	-0,185	-0,046	-0,096	-0,096
IDH - 2013_1_N	0,056	0,065	0,104	0,055
1. Governança_1_N	0,279	0,132	0,414	0,382
EPI Score_1_N	0,037	0,109	0,152	0,042
Happy Planet Index_1_N	0,141	0,070	0,103	0,138
GINNI P	0,248	0,271	0,298	0,131

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
GHANA 44**	4	2	1	28,42	0,010
			2	19,87	0,725
			3	42,14	0,000
			4	21,88	0,265
HONDURAS 49**	1	4	1	18,78	0,420
			2	40,54	0,000
			3	43,19	0,000
			4	18,13	0,580
ISRAEL 57**	4	3	1	22,88	0,018
			2	46,38	0,000
			3	15,59	0,697
			4	17,39	0,284
KENYA 63**	4	2	1	30,148	0,000
			2	7,892	0,964
			3	37,250	0,000
			4	14,459	0,036
KUWAIT 65**	1	3	1	19,49	0,428
			2	34,30	0,000
			3	18,96	0,557
			4	26,23	0,015
LATVIA 68**	1	3	1	17,45	0,194
			2	36,96	0,000
			3	14,65	0,789
			4	22,35	0,017
MALAWI 75**	4	2	1	40,20	0,000
			2	19,88	0,743
			3	47,49	0,000
			4	22,00	0,257
MEXICO 80**	4	1	1	10,06	0,621
			2	33,86	0,000
			3	17,70	0,014
			4	11,12	0,365
MONGOLIA 82**		2	1	27,24	0,968
			2	34,67	0,024
			3	46,45	0,000
			4	36,80	0,008
SAUDI ARABIA 104**		3	1	16,64	0,705
			2	38,31	0,000
			3	18,60	0,264
			4	22,88	0,031

3. FUNÇÃO QUADRÁTICA

3.1 Discriminant Analysis: 4 clusters versus Social Progr; Shelter_1N; ...

Social Progress Index_1N is highly correlated with other predictors in group 4.

Shelter_1N is highly correlated with other predictors in group 4.

Access to electricity (% of p_N is highly correlated with other predictors in

group 4.

Quadratic Method for Response: 4 clusters

Predictors: Social Progress Index_1N; Shelter_1N; Access to piped water_1N; Access to improved sanitation_N; Availability of affordable ho_N; Access to electricity (% of p_N; Quality of electricity supply_N; Private property rights (0=no_N; IDH - 2013_1_N; 1. Governança_1_N; EPI Score_1_N; Happy Planet Index_1_N; GINNI P

Group	1	2	3	4
Count	54	23	36	19

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	52	0	0	0
2	0	23	0	0
3	2	0	36	0
4	0	0	0	19
Total N	54	23	36	19
N correct	52	23	36	19
Proportion	0,963	1,000	1,000	1,000

N = 132 N Correct = 130 Proportion Correct = 0,985

From Generalized Squared Distance to Group

Group	1	2	3	4
1	62,49	130,02	106,47	159,28
2	148,01	56,87	3491,03	82,44
3	86,70	512,39	48,22	373,52
4	86,47	198,40	1038,27	58,30

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
LATVIA 68**	1	3	1	81,62	0,063
			2	286,41	0,000
			3	76,23	0,937
			4	256,19	0,000
TRINIDAD120**	1	3	1	72,63	0,137
			2	314,59	0,000
			3	68,95	0,863
			4	593,17	0,000

4. FUNÇÃO LINEAR

4.1 Discriminant Analysis: N Class 3R versus Social Progr; Shelter_1N; ...

Linear Method for Response: N Class 3R

Predictors: Social Progress Index_1N; Shelter_1N; Access to piped water_1N;
Access to improved sanitation_N; Availability of affordable ho_N;
Access to electricity (% of p_N; Quality of electricity supply_N;
Private property rights (0=no_N; IDH - 2013_1_N; 1. Governança_1_N;
EPI Score_1_N; Happy Planet Index_1_N; GINNI P

Group	AIBER	AVECO	OTHERS
Count	21	27	84

Summary of classification

	True Group		
Put into Group	AIBER	AVECO	OTHERS
AIBER	18	0	4
AVECO	2	25	2
OTHERS	1	2	78
Total N	21	27	84
N correct	18	25	78
Proportion	0,857	0,926	0,929

N = 132 N Correct = 121 Proportion Correct = 0,917

Squared Distance Between Groups

	AIBER	AVECO	OTHERS
AIBER	0,0000	18,3297	10,1886
AVECO	18,3297	0,0000	15,2228
OTHERS	10,1886	15,2228	0,0000

Linear Discriminant Function for Groups

	AIBER	AVECO	OTHERS
Constant	-17,424	-35,137	-13,819
Social Progress Index_1N	0,144	-0,112	-0,259
Shelter_1N	0,134	0,186	0,128
Access to piped water_1N	0,082	0,041	0,008
Access to improved sanitation_N	-0,075	-0,065	-0,007
Availability of affordable ho_N	0,102	0,167	0,135
Access to electricity (% of p_N	-0,167	-0,222	-0,095
Quality of electricity supply_N	-0,038	-0,037	-0,008
Private property rights (0=no_N	0,030	0,087	-0,013
IDH - 2013_1_N	0,011	0,090	0,068
1. Governança_1_N	-0,032	0,186	0,218
EPI Score_1_N	0,118	0,227	0,105
Happy Planet Index_1_N	0,208	0,152	0,100
GINNI P	0,141	0,317	0,240

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
Cuba	31**	AIBER	OTHERS	AIBER	20,34	0,383
				AVECO	39,03	0,000
				OTHERS	19,39	0,617
Greece	45**	AVECO	OTHERS	AIBER	14,192	0,056
				AVECO	9,999	0,460
				OTHERS	9,896	0,484
Guyana	48**	OTHERS	AIBER	AIBER	10,65	0,690
				AVECO	30,00	0,000
				OTHERS	12,25	0,310
Hungary	50**	OTHERS	AVECO	AIBER	25,102	0,000
				AVECO	9,050	0,832
				OTHERS	12,259	0,167
Jamaica	59**	OTHERS	AIBER	AIBER	4,808	0,998
				AVECO	22,276	0,000
				OTHERS	17,719	0,002
Latvia	68**	AVECO	OTHERS	AIBER	21,61	0,019
				AVECO	17,46	0,153
				OTHERS	14,08	0,828
Namibia	86**	OTHERS	AIBER	AIBER	22,46	0,639
				AVECO	38,15	0,000
				OTHERS	23,60	0,361
Filipinas	98**	OTHERS	AIBER	AIBER	8,907	0,812
				AVECO	24,969	0,000
				OTHERS	11,835	0,188
Poland	99**	OTHERS	AVECO	AIBER	13,837	0,048
				AVECO	8,058	0,870
				OTHERS	12,782	0,082
Portugal	100**	AIBER	AVECO	AIBER	12,858	0,052
				AVECO	7,130	0,908
				OTHERS	13,365	0,040
Spain	110**	AIBER	AVECO	AIBER	13,683	0,011
				AVECO	4,744	0,986
				OTHERS	16,494	0,003

5. FUNÇÃO QUADRÁTICA

5.1 Discriminant Analysis: N Class 3R versus Social Progr; Shelter_1N; ...

Shelter_1N is highly correlated with other predictors in group AVECO.

Availability of affordable ho_N is highly correlated with other predictors in group AVECO.

Quadratic Method for Response: N Class 3R

Predictors: Social Progress Index_1N; Shelter_1N; Access to piped water_1N; Access to improved sanitation_N; Availability of affordable ho_N; Access to electricity (% of p_N; Quality of electricity supply_N; Private property rights (0=no_N; IDH - 2013_1_N; 1. Governança_1_N; EPI Score_1_N; Happy Planet Index_1_N; GINNI P

Group	AIBER	AVECO	OTHERS
Count	21	27	84

Summary of classification

	True Group

Put into Group	AIBER	AVECO	OTHERS
AIBER	19	0	3
AVECO	2	27	3
OTHERS	0	0	78
Total N	21	27	84
N correct	19	27	78
Proportion	0,905	1,000	0,929

N = 132 N Correct = 124 Proportion Correct = 0,939

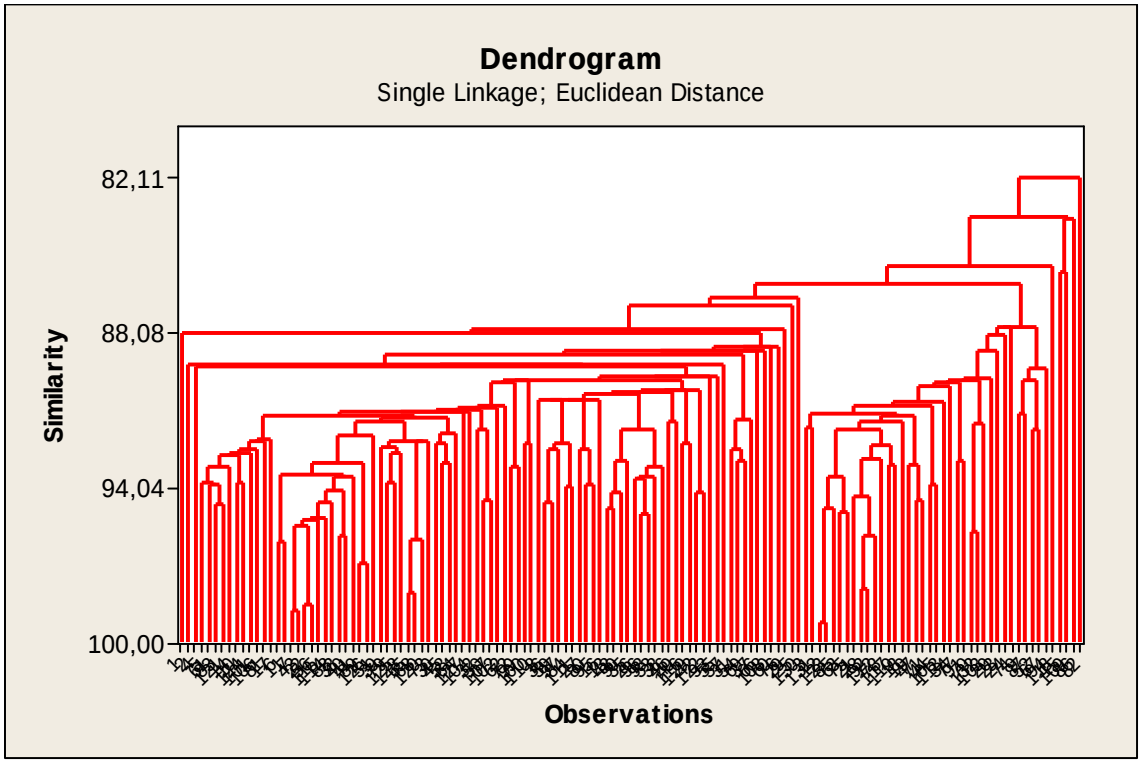
From Generalized Squared Distance to Group

Group	AIBER	AVECO	OTHERS
AIBER	55,47	163,47	76,28
AVECO	73,76	43,19	84,14
OTHERS	108,73	2049,38	66,21

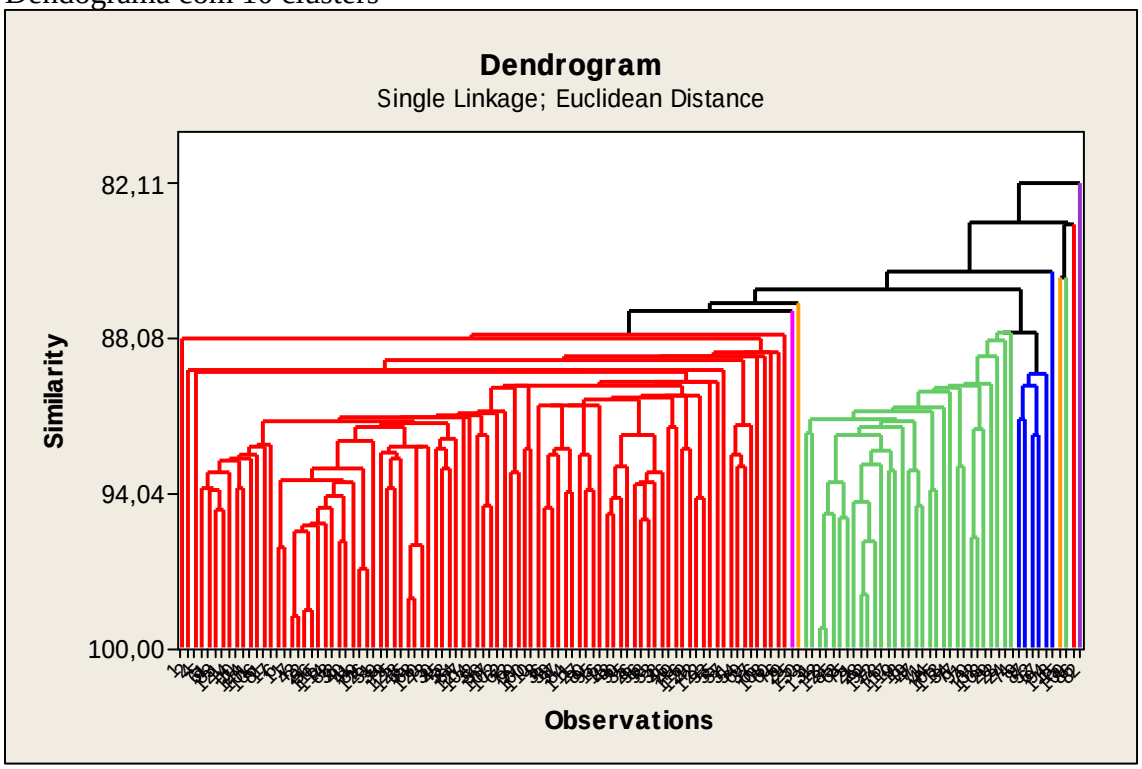
Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
Croatia 30**	OTHERS	AVECO	AIBER	75,56	0,000
			AVECO	59,66	0,999
			OTHERS	73,85	0,001
Jamaica 59**	OTHERS	AIBER	AIBER	68,48	0,999
			AVECO	122,72	0,000
			OTHERS	82,53	0,001
Lithuania72**	OTHERS	AIBER	AIBER	72,08	0,550
			AVECO	72,87	0,371
			OTHERS	75,97	0,079
Malaysia 76**	OTHERS	AVECO	AIBER	83,97	0,002
			AVECO	71,74	0,803
			OTHERS	74,57	0,195
Mauritius79**	OTHERS	AVECO	AIBER	102,15	0,000
			AVECO	66,88	0,999
			OTHERS	81,36	0,001
Polanda 99**	OTHERS	AIBER	AIBER	75,38	0,806
			AVECO	298,23	0,000
			OTHERS	78,23	0,194
Portugal100**	AIBER	AVECO	AIBER	64,01	0,256
			AVECO	61,87	0,744
			OTHERS	80,59	0,000
Spain 110**	AIBER	AVECO	AIBER	66,96	0,001
			AVECO	52,79	0,999
			OTHERS	83,22	0,000

6. CLASSIFICAÇÃO NÃO SUPERVISONADA



Dendograma com 10 clusters



Cluster Analysis of Observations: MeDS; McFA; UDF; IDM

```
Euclidean Distance, Single Linkage
Amalgamation Steps
```

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	99,2416	0,08455	12	123	12	2
2	130	98,7895	0,13496	7	43	7	2
3	129	98,5406	0,16271	22	56	22	2
4	128	98,0717	0,21498	65	68	65	2
5	127	97,9636	0,22704	28	92	28	2
6	126	96,9958	0,33494	89	126	89	2
7	125	96,1087	0,43384	6	11	6	2
8	124	96,0340	0,44216	65	120	65	3
9	123	95,9395	0,45270	33	40	33	2
10	122	95,9388	0,45279	28	112	28	3
11	121	95,7518	0,47363	70	132	70	2
12	120	95,5062	0,50101	7	22	7	4
13	119	95,3072	0,52320	7	115	7	5
14	118	95,2455	0,53008	7	114	7	6
15	117	95,0362	0,55341	96	98	96	2
16	116	94,9725	0,56051	63	75	63	2
17	115	94,8241	0,57706	12	85	12	3
18	114	94,8088	0,57877	13	48	13	2
19	113	94,6631	0,59501	83	121	83	2
20	112	94,6370	0,59791	36	59	36	2
21	111	94,6368	0,59794	7	88	7	7
22	110	94,5327	0,60955	50	107	50	2
23	109	94,5107	0,61201	47	71	47	2
24	108	94,3562	0,62922	21	28	21	4
25	107	94,2496	0,64111	13	90	13	3
26	106	94,2398	0,64221	42	122	42	2
27	105	94,1772	0,64918	7	33	7	9
28	104	93,9843	0,67069	84	111	84	2
29	103	93,9548	0,67398	80	95	80	2
30	102	93,9544	0,67402	44	105	44	2
31	101	93,9001	0,68008	61	83	61	3
32	100	93,8667	0,68380	5	61	5	4
33	99	93,8340	0,68745	10	124	10	2
34	98	93,8193	0,68908	79	125	79	2
35	97	93,6775	0,70490	26	96	26	3
36	96	93,6609	0,70675	12	63	12	5
37	95	93,6549	0,70742	26	53	26	4
38	94	93,6363	0,70949	7	41	7	10
39	93	93,5841	0,71530	6	7	6	12
40	92	93,2754	0,74973	5	14	5	5
41	91	93,2342	0,75432	26	38	26	5
42	90	93,2223	0,75564	72	99	72	2
43	89	93,1796	0,76040	117	119	117	2
44	88	93,1424	0,76456	19	47	19	3
45	87	93,1239	0,76662	45	58	45	2
46	86	93,1140	0,76771	6	89	6	14
47	85	93,0485	0,77502	77	91	77	2
48	84	93,0349	0,77654	54	69	54	2
49	83	93,0342	0,77662	13	35	13	4
50	82	92,9555	0,78538	21	78	21	5
51	81	92,7562	0,80761	5	10	5	7
52	80	92,7415	0,80925	79	128	79	3
53	79	92,7093	0,81284	5	101	5	8
54	78	92,5813	0,82711	36	97	36	3
55	77	92,5778	0,82750	27	80	27	3
56	76	92,5436	0,83131	31	54	31	3
57	75	92,5345	0,83233	5	106	5	9
58	74	92,5092	0,83515	39	79	39	4
59	73	92,3796	0,84960	21	117	21	7
60	72	92,3636	0,85138	100	110	100	2
61	71	92,3336	0,85473	116	129	116	2
62	70	92,3118	0,85715	30	45	30	3
63	69	92,3065	0,85775	36	84	36	5
64	68	92,2973	0,85878	65	73	65	4
65	67	92,2893	0,85966	39	65	39	8

66	66	92,2741	0,86136	5	81	5	10
67	65	92,1547	0,87467	5	17	5	11
68	64	92,0240	0,88925	6	51	6	15
69	63	91,9435	0,89822	30	64	30	4
70	62	91,8345	0,91037	52	67	52	2
71	61	91,8163	0,91240	12	21	12	12
72	60	91,7940	0,91489	13	26	13	9
73	59	91,7841	0,91598	32	50	32	3
74	58	91,7193	0,92321	3	131	3	2
75	57	91,6014	0,93636	70	103	70	3
76	56	91,5419	0,94299	46	49	46	2
77	55	91,5123	0,94629	6	39	6	23
78	54	91,5009	0,94757	12	18	12	13
79	53	91,4058	0,95817	31	37	31	4
80	52	91,3631	0,96293	6	30	6	27
81	51	91,2706	0,97324	5	6	5	38
82	50	91,2567	0,97479	12	19	12	16
83	49	91,2178	0,97913	3	12	3	18
84	48	91,1995	0,98116	9	87	9	2
85	47	91,1086	0,99130	5	127	5	39
86	46	91,0969	0,99261	5	104	5	40
87	45	91,0789	0,99461	5	32	5	43
88	44	91,0070	1,00263	5	102	5	44
89	43	90,9115	1,01327	3	44	3	20
90	42	90,8848	1,01625	5	62	5	45
91	41	90,7407	1,03232	3	113	3	21
92	40	90,6957	1,03733	8	36	8	6
93	39	90,6947	1,03745	8	27	8	9
94	38	90,6320	1,04444	8	16	8	10
95	37	90,4708	1,06241	8	13	8	19
96	36	90,4007	1,07022	8	46	8	21
97	35	90,2690	1,08491	8	116	8	23
98	34	90,2583	1,08610	8	42	8	25
99	33	90,1476	1,09844	3	34	3	22
100	32	90,0074	1,11408	3	77	3	24
101	31	90,0053	1,11430	5	72	5	47
102	30	89,9311	1,12258	5	100	5	49
103	29	89,8858	1,12763	5	8	5	74
104	28	89,8849	1,12773	9	52	9	4
105	27	89,8806	1,12821	3	70	3	27
106	26	89,8134	1,13570	3	20	3	28
107	25	89,7949	1,13777	5	93	5	75
108	24	89,7801	1,13942	5	55	5	76
109	23	89,4514	1,17606	9	94	9	5
110	22	89,4023	1,18153	4	5	4	77
111	21	89,3018	1,19274	4	57	4	78
112	20	89,2859	1,19451	2	4	2	79
113	19	88,9447	1,23255	2	31	2	83
114	18	88,7981	1,24890	3	23	3	29
115	17	88,7939	1,24937	2	66	2	84
116	16	88,7627	1,25285	2	108	2	85
117	15	88,6398	1,26655	2	60	2	86
118	14	88,5996	1,27102	2	76	2	87
119	13	88,1943	1,31622	3	24	3	30
120	12	88,1184	1,32468	1	2	1	88
121	11	87,9689	1,34135	1	29	1	89
122	10	87,8858	1,35061	3	74	3	31
123	9	87,8344	1,35634	3	9	3	36
124	8	87,0619	1,44246	1	25	1	90
125	7	86,7289	1,47960	1	130	1	91
126	6	86,2212	1,53619	1	3	1	127
127	5	85,7435	1,58945	15	109	15	2
128	4	85,4823	1,61857	1	118	1	128
129	3	83,7087	1,81632	15	86	15	3
130	2	83,6362	1,82439	1	15	1	131
131	1	82,1132	1,99419	1	82	1	132

Final Partition
Number of clusters: 10

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	89	544,498	2,34080	4,06132
Cluster2	31	73,135	1,43515	2,59852
Cluster3	5	4,240	0,91063	1,08467
Cluster4	1	0,000	0,00000	0,00000
Cluster5	1	0,000	0,00000	0,00000
Cluster6	1	0,000	0,00000	0,00000
Cluster7	1	0,000	0,00000	0,00000
Cluster8	1	0,000	0,00000	0,00000
Cluster9	1	0,000	0,00000	0,00000
Cluster10	1	0,000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
MeDS	1,57924	-4,08242	-2,53963	-0,97014	2,84666	-1,75147	-1,57746
McFA	0,25327	-0,88737	0,93273	-1,98626	0,11250	-1,05883	-0,52318
UDF	-0,14903	0,13949	0,62978	1,41974	1,24261	-2,20881	2,13608
IDM	0,02032	0,01195	1,14455	-2,42645	-1,91967	-0,76483	-1,81835

Variable	Cluster8	Cluster9	Cluster10	Grand centroid
MeDS	-0,54575	1,08708	-0,38832	0,0000000
McFA	-0,56172	1,20794	3,11319	-0,0000000
UDF	0,92223	2,18221	0,09650	0,0000000
IDM	-2,68995	1,76845	-0,05121	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,00000	5,78262	4,39286	4,46795	2,70673	4,20409	4,37836
Cluster2	5,78262	0,00000	2,68628	4,29868	7,34573	3,40303	3,70726
Cluster3	4,39286	2,68628	0,00000	4,93558	6,28092	4,03619	3,75409
Cluster4	4,46795	4,29868	4,93558	0,00000	4,38873	4,17109	1,84183
Cluster5	2,70673	7,34573	6,28092	4,38873	0,00000	5,98003	4,55911
Cluster6	4,20409	3,40303	4,03619	4,17109	5,98003	0,00000	4,50612
Cluster7	4,37836	3,70726	3,75409	1,84183	4,55911	4,50612	0,00000
Cluster8	3,69770	4,53067	4,58235	1,58945	3,55795	3,90005	1,81632
Cluster9	3,10552	6,19451	4,00350	5,71083	4,33373	6,23652	4,79206
Cluster10	3,48078	5,44579	3,33118	5,80824	4,92686	5,00875	4,68187

	Cluster8	Cluster9	Cluster10
Cluster1	3,69770	3,10552	3,48078
Cluster2	4,53067	6,19451	5,44579
Cluster3	4,58235	4,00350	3,33118
Cluster4	1,58945	5,71083	5,80824
Cluster5	3,55795	4,33373	4,92686
Cluster6	3,90005	6,23652	5,00875
Cluster7	1,81632	4,79206	4,68187
Cluster8	0,00000	5,22138	4,60158
Cluster9	5,22138	0,00000	3,66990
Cluster10	4,60158	3,66990	0,00000

Dendrogram

Cluster Analysis of Observations: MeDS; McFA; UDF; IDM

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	99,0364	0,1523	12	123	12	2
2	130	98,5290	0,2325	7	43	7	2
3	129	98,0784	0,3038	22	56	22	2
4	128	97,8861	0,3342	65	68	65	2
5	127	97,5246	0,3913	28	92	28	2
6	126	96,1699	0,6055	89	126	89	2
7	125	95,8460	0,6567	33	40	33	2
8	124	95,1925	0,7600	70	132	70	2
9	123	94,6952	0,8386	6	11	6	2
10	122	94,4930	0,8706	63	75	63	2
11	121	94,1987	0,9171	61	121	61	2
12	120	94,1905	0,9184	36	59	36	2
13	119	94,0678	0,9378	65	120	65	3
14	118	94,0367	0,9427	21	28	21	3
15	117	93,9100	0,9627	7	115	7	3
16	116	93,7909	0,9816	84	111	84	2
17	115	93,7779	0,9836	13	48	13	2
18	114	93,7545	0,9873	42	122	42	2
19	113	93,6917	0,9972	12	85	12	3
20	112	93,3870	1,0454	26	53	26	2
21	111	93,3649	1,0489	96	98	96	2
22	110	93,0289	1,1020	50	107	50	2
23	109	92,9520	1,1142	44	105	44	2
24	108	92,9137	1,1202	45	58	45	2
25	107	92,8950	1,1232	88	114	88	2
26	106	92,8754	1,1263	47	71	47	2
27	105	92,7538	1,1455	54	69	54	2
28	104	92,6514	1,1617	10	124	10	2
29	103	92,6461	1,1625	77	91	77	2
30	102	92,6454	1,1626	35	90	35	2
31	101	92,6268	1,1656	5	14	5	2
32	100	92,5771	1,1734	80	95	80	2
33	99	92,3797	1,2046	79	125	79	2
34	98	92,3549	1,2086	17	106	17	2
35	97	92,2894	1,2189	117	119	117	2
36	96	92,1901	1,2346	72	99	72	2
37	95	91,7979	1,2966	3	78	3	2
38	94	91,6465	1,3206	6	41	6	3
39	93	91,3631	1,3654	7	22	7	5
40	92	91,0998	1,4070	9	87	9	2
41	91	91,0909	1,4084	83	102	83	2
42	90	90,9841	1,4253	21	112	21	4
43	89	90,8196	1,4513	81	101	81	2
44	88	90,7601	1,4607	116	129	116	2
45	87	90,6241	1,4822	79	128	79	3
46	86	90,2846	1,5359	46	49	46	2
47	85	89,6125	1,6421	51	89	51	3
48	84	89,5950	1,6449	100	110	100	2
49	83	89,5018	1,6596	30	64	30	2
50	82	89,3862	1,6779	19	47	19	3
51	81	89,2644	1,6971	31	54	31	3
52	80	89,2029	1,7069	52	67	52	2
53	79	88,7874	1,7725	32	50	32	3
54	78	88,7653	1,7760	5	61	5	4
55	77	88,6192	1,7991	4	97	4	2
56	76	88,4475	1,8263	8	62	8	2
57	75	88,4379	1,8278	33	88	33	4
58	74	88,3804	1,8369	16	27	16	2
59	73	88,3199	1,8464	38	96	38	3
60	72	88,3160	1,8471	70	103	70	3
61	71	88,2338	1,8601	45	104	45	3
62	70	87,3788	1,9952	12	18	12	4
63	69	87,1978	2,0238	39	100	39	3
64	68	87,0230	2,0515	6	7	6	8
65	67	86,8995	2,0710	10	81	10	4

66	66	86,7600	2,0930	65	73	65	4
67	65	86,6668	2,1078	55	66	55	2
68	64	85,9184	2,2261	13	35	13	4
69	63	85,4136	2,3059	2	55	2	3
70	62	85,3810	2,3110	4	36	4	4
71	61	85,3683	2,3130	52	94	52	3
72	60	85,3013	2,3236	30	79	30	5
73	59	85,2292	2,3350	63	70	63	5
74	58	85,0786	2,3588	1	118	1	2
75	57	84,9011	2,3869	20	21	20	5
76	56	84,8771	2,3907	26	38	26	5
77	55	84,5724	2,4389	72	108	72	3
78	54	84,5156	2,4478	25	57	25	2
79	53	84,2534	2,4893	34	44	34	3
80	52	84,2107	2,4960	6	127	6	9
81	51	84,1557	2,5047	24	117	24	3
82	50	83,9600	2,5357	51	93	51	4
83	49	83,4902	2,6099	15	109	15	2
84	48	83,4879	2,6103	16	29	16	3
85	47	83,1493	2,6638	45	76	45	4
86	46	82,0330	2,8403	34	113	34	4
87	45	81,7383	2,8869	3	131	3	3
88	44	81,4120	2,9385	16	80	16	5
89	43	80,9285	3,0149	32	39	32	6
90	42	80,7331	3,0458	10	37	10	5
91	41	80,0571	3,1527	8	84	8	4
92	40	79,6430	3,2181	5	17	5	6
93	39	79,1986	3,2884	31	83	31	5
94	38	79,0496	3,3119	1	42	1	4
95	37	78,9249	3,3316	33	60	33	5
96	36	78,8260	3,3473	12	74	12	5
97	35	78,7937	3,3524	19	24	19	6
98	34	78,5252	3,3948	15	86	15	3
99	33	77,8792	3,4970	13	46	13	6
100	32	77,6614	3,5314	30	45	30	9
101	31	77,5957	3,5418	5	65	5	10
102	30	77,3227	3,5849	6	51	6	13
103	29	76,8283	3,6631	19	23	19	7
104	28	75,5334	3,8678	2	8	2	7
105	27	75,5191	3,8701	4	13	4	10
106	26	75,4978	3,8734	12	20	12	10
107	25	75,0251	3,9481	9	52	9	5
108	24	74,4899	4,0327	4	130	4	11
109	23	72,9502	4,2762	26	116	26	7
110	22	72,2808	4,3820	12	77	12	12
111	21	72,1252	4,4066	10	31	10	10
112	20	71,6680	4,4788	32	72	32	9
113	19	71,1360	4,5630	6	33	6	18
114	18	68,5569	4,9707	3	12	3	15
115	17	68,2854	5,0136	16	25	16	7
116	16	66,7532	5,2558	34	63	34	9
117	15	65,5397	5,4476	2	10	2	17
118	14	64,9680	5,5380	9	26	9	12
119	13	63,4120	5,7840	30	32	30	18
120	12	58,5085	6,5592	2	5	2	27
121	11	57,9682	6,6446	19	82	19	8
122	10	56,3017	6,9080	1	9	1	16
123	9	51,2819	7,7016	15	34	15	12
124	8	48,5391	8,1352	6	30	6	36
125	7	46,4914	8,4589	1	4	1	27
126	6	43,9014	8,8683	3	19	3	23
127	5	35,9407	10,1268	15	16	15	19
128	4	33,3867	10,5305	1	2	1	54
129	3	25,5296	11,7726	1	15	1	73
130	2	13,6275	13,6542	1	6	1	109
131	1	0,0000	15,8085	1	3	1	132

Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	54	217,201	1,91364	3,64898
Cluster2	23	51,802	1,34629	3,40534
Cluster3	36	82,134	1,43967	2,33866
Cluster4	19	144,184	2,67733	3,95962

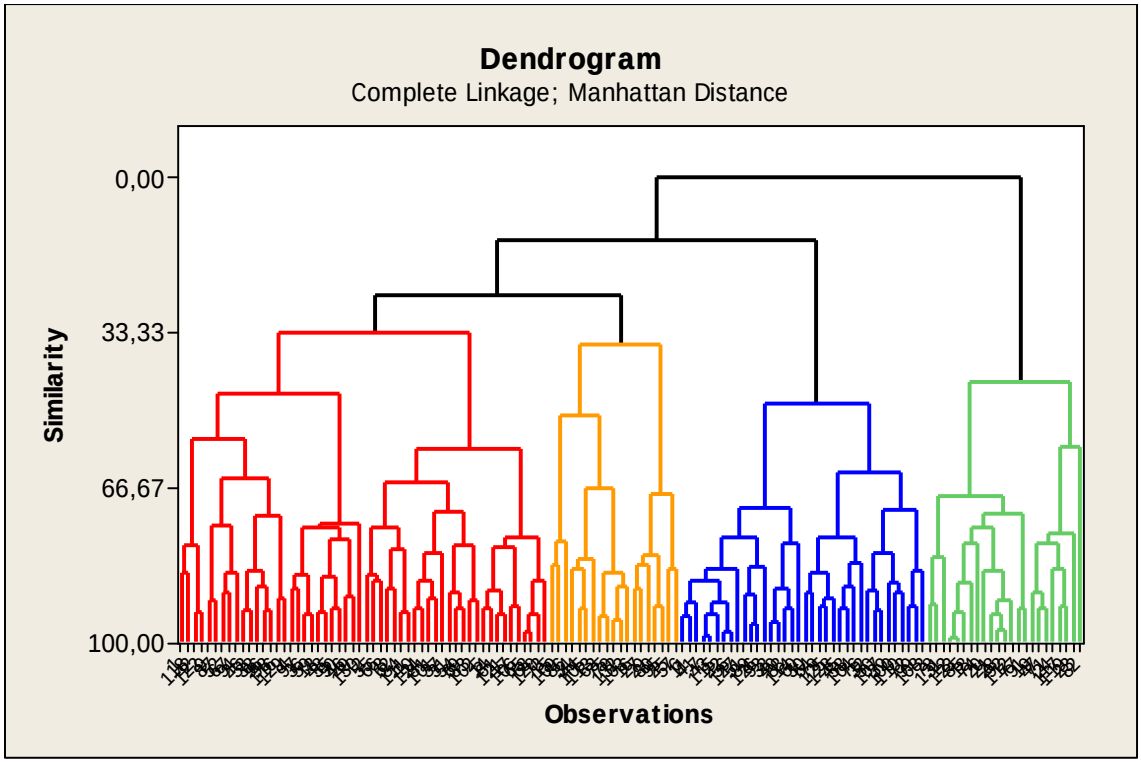
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
MeDS	-0,02766	-4,35068	3,35427	-1,01022	0,0000000
McFA	1,01797	-1,00535	-0,88754	0,00546	-0,0000000
UDF	-0,29730	-0,22901	0,04962	1,02816	0,0000000
IDM	0,21539	0,19328	0,11086	-1,05618	0,0000000

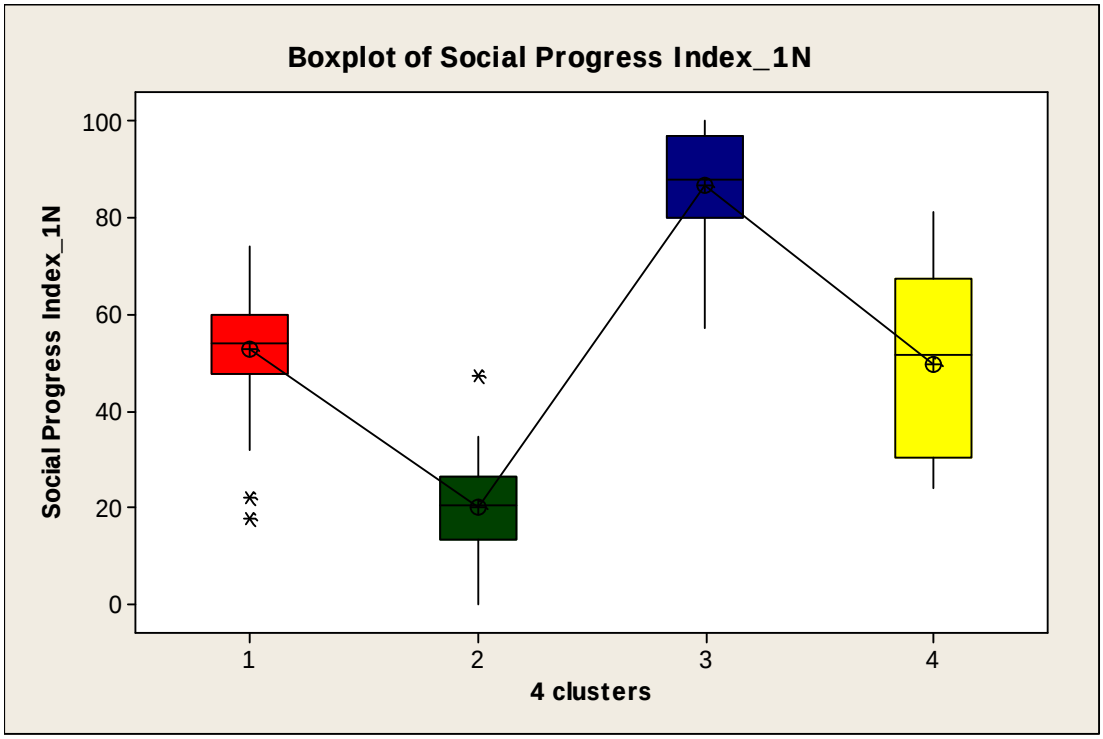
Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,00000	4,77362	3,89868	2,31611
Cluster2	4,77362	0,00000	7,71132	3,91433
Cluster3	3,89868	7,71132	0,00000	4,70805
Cluster4	2,31611	3,91433	4,70805	0,00000

MAIOR DISTÂNCIA
MENOR DISTÂNCIA



Cluster 1 Vermelho: 54 países está como o segundo cluster mais fraco;
Cluster 2 Verde: 23 Países é o mais fraco;
Cluster 3 Azul: 36 Países Mais forte;
Cluster 4 Amarelo: 19 Países Segundo Mais forte;



One-way ANOVA: Social Progress Index_1N versus 4 clusters

```
Source       DF     SS     MS       F      P
4 clusters    3  65253  21751  138,61  0,000
Error       128  20085    157
Total       131  85338

S = 12,53   R-Sq = 76,46%   R-Sq(adj) = 75,91%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---+---------+---------+---------+------
1Verme 54  52,85  11,46                  (*-)
2Verde 23  19,84  10,58  (--*--)
3Azul  36  86,67  10,78                                    (-*-)
4Amare 19  49,53  19,13                (--*--)
                          ---+---------+---------+---------+------
                            20        40        60        80

Pooled StDev = 12,53
```

Ordenado os clusters do mais fraco para o mais forte

2 → 1
4 → 2
1 → 3
3 → 4

Conforme podemos verificar acima reordenamos a posição dos clusters para darmos sequência no trabalho 10 (capitulo 10) com as regressões logísticas.

CAPITULO 10 – REGRESSÃO LOGÍSTICA

1. Regressão Logística com as principais variáveis

Ordinal Logistic Regression: 3R ORD versus IDH - 2013_1; Shelter_1N; ...

Link Function: Logit

Response Information

Variable	Value	Count
3R ORD	1	23
	2	73
	3	36
	Total	132

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	10,9621	2,57704	4,25	0,000	
Const(2)	23,8358	4,70798	5,06	0,000	
IDH - 2013_1_N	-0,0883431	0,0399441	-2,21	0,027	0,92
Shelter_1N	-0,0999264	0,0477330	-2,09	0,036	0,90
Access to improved sanitation_N	-0,0578996	0,0249647	-2,32	0,020	0,94
Happy Planet Index_1_N	0,0021636	0,0190191	0,11	0,909	1,00
Private property rights (0=no_N	-0,0692108	0,0253119	-2,73	0,006	0,93
Access to electricity (% of p_N	0,0278383	0,0255301	1,09	0,276	1,03
Availability of affordable ho_N	-0,0368878	0,0205854	-1,79	0,073	0,96
GINNI P	-0,0134707	0,0210644	-0,64	0,522	0,99

Predictor	95% CI Lower	Upper
Const(1)		
Const(2)		
IDH - 2013_1_N	0,85	0,99
Shelter_1N	0,82	0,99
Access to improved sanitation_N	0,90	0,99
Happy Planet Index_1_N	0,97	1,04
Private property rights (0=no_N	0,89	0,98
Access to electricity (% of p_N	0,98	1,08
Availability of affordable ho_N	0,93	1,00
GINNI P	0,95	1,03

Log-Likelihood = -32,288
Test that all slopes are zero: G = 195,831, DF = 8, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	82,2723	254	1,000
Deviance	64,5759	254	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	5051	98,4	Somers' D	0,97
Discordant	83	1,6	Goodman-Kruskal Gamma	0,97
Ties	1	0,0	Kendall's Tau-a	0,57
Total	5135	100,0		

Ao rodarmos nossa primeira Regressão logística “Ordinal Logistic Regression” com as variáveis indicadas no “Principal Componentes PC1, PC2, PC3 e PC4” que compõem o tema moradia percebemos que existem variáveis com baixo grau de confiabilidade, ou seja, P-Value superior a 10% (marcadas em vermelho na regressão acima).

Desta forma iremos rodar a seguir nova Regressão Logística retirando estas variáveis:

2. Regressão Logística retirando variáveis com elevado p-value

Ordinal Logistic Regression: 3R ORD versus IDH - 2013_1; Shelter_1N; ...

```
Link Function: Logit


Response Information

Variable  Value  Count
3R ORD    1         23
          2         73
          3         36
          Total    132


Logistic Regression Table

                                                                        Odds
Predictor                              Coef    SE Coef      Z      P  Ratio
Const(1)                            10,0340    2,04187   4,91  0,000
Const(2)                            22,6718    4,28081   5,30  0,000
IDH - 2013_1_N                   -0,0795674  0,0375335  -2,12  0,034   0,92
Shelter_1N                       -0,0745848  0,0407722  -1,83  0,067   0,93
Access to improved sanitation_N  -0,0553169  0,0249841  -2,21  0,027   0,95
Private property rights (0=no_N  -0,0722383  0,0252755  -2,86  0,004   0,93
Availability of affordable ho_N  -0,0312542  0,0188962  -1,65  0,098   0,97

                                    95% CI
Predictor                         Lower  Upper
Const(1)
Const(2)
IDH - 2013_1_N                     0,86   0,99
Shelter_1N                         0,86   1,01
Access to improved sanitation_N    0,90   0,99
Private property rights (0=no_N    0,89   0,98
Availability of affordable ho_N    0,93   1,01


Log-Likelihood = -33,100
Test that all slopes are zero: G = 194,207, DF = 5, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      109,183  257  1,000
Deviance      66,200  257  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    5055     98,4  Somers' D              0,97
Discordant      80      1,6  Goodman-Kruskal Gamma  0,97
Ties             0      0,0  Kendall's Tau-a        0,58
Total         5135    100,0
```

Podemos verificar que em ambos os casos a eficiência do modelo ficou em 98,4%.

Agora iremos rodar Regressão Logística colocando como “response” as 3 regiões por ordem de força, onde a região 1 representa a AVECO, região 2 representa AIBER e região 3 representa OTHERS:

3. Regressão Logística com regiões em ordem de força

Ordinal Logistic Regression: NClass ORD versus IDH - 2013_1; Shelter_1N; ...

```
Link Function: Logit


Response Information

Variable    Value  Count
NClass ORD  1         27
            2         21
            3         84
            Total    132


Logistic Regression Table

                                                                    Odds
Predictor                              Coef    SE Coef      Z      P  Ratio
Const(1)                          -10,2334    1,64829  -6,21  0,000
Const(2)                          -8,51162    1,53183  -5,56  0,000
IDH - 2013_1_N                   0,0802787  0,0269904   2,97  0,003   1,08
Shelter_1N                       0,0480104  0,0298229   1,61  0,107   1,05
Access to improved sanitation_N -0,0322263  0,0187752  -1,72  0,086   0,97
Private property rights (0=no_N  0,0225717  0,0119326   1,89  0,059   1,02
Availability of affordable ho_N  0,0213044  0,0129036   1,65  0,099   1,02

                                    95% CI
Predictor                         Lower  Upper
Const(1)
Const(2)
IDH - 2013_1_N                     1,03   1,14
Shelter_1N                         0,99   1,11
Access to improved sanitation_N    0,93   1,00
Private property rights (0=no_N    1,00   1,05
Availability of affordable ho_N    1,00   1,05


Log-Likelihood = -69,879
Test that all slopes are zero: G = 99,079, DF = 5, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      175,517  257  1,000
Deviance     139,758  257  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    4100     89,1  Somers' D              0,78
Discordant     490     10,7  Goodman-Kruskal Gamma  0,79
Ties             9      0,2  Kendall's Tau-a        0,42
Total         4599    100,0
```

4. Teste para Regressão Logística

Concluída esta etapa iremos efetuar teste para Regressão Logística Nominal, onde, verifica-se que o MINITAB não conseguiu resolver a Regressão Logística Nominal.

Nominal Logistic Regression: 3clusters versus IDH - 2013_1; Shelter_1N; ...

```
* WARNING * Algorithm has not converged after 20 iterations.
* WARNING * Convergence has not been reached for the parameter estimates
            criterion.
* WARNING * The results may not be reliable.
* WARNING * Try increasing the maximum number of iterations.
```

Response Information

Variable	Value	Count	
3clusters	3	36	(Reference Event)
	2	23	
	1	73	
	Total	132	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower
Logit 1: (2/3)						
Constant	4072,07	39439,0	0,10	0,918		
IDH - 2013_1_N	-2,13135	169,255	-0,01	0,990	0,12	0,00
Shelter_1N	-25,3256	273,676	-0,09	0,926	0,00	0,00
Access to improved sanitation_N	-5,15408	268,703	-0,02	0,985	0,01	0,00
Happy Planet Index_1_N	0,205951	108,276	0,00	0,998	1,23	0,00
Private property rights (0=no_N	-6,36606	128,527	-0,05	0,960	0,00	0,00
Access to electricity (% of p_N	-6,42426	201,642	-0,03	0,975	0,00	0,00
Availability of affordable ho_N	-2,06469	233,899	-0,01	0,993	0,13	0,00
GINNI P	-6,27843	156,143	-0,04	0,968	0,00	0,00
Logit 2: (1/3)						
Constant	3952,52	36094,6	0,11	0,913		
IDH - 2013_1_N	-2,75274	41,5427	-0,07	0,947	0,06	0,00
Shelter_1N	-23,2479	220,226	-0,11	0,916	0,00	0,00
Access to improved sanitation_N	-4,61534	71,9552	-0,06	0,949	0,01	0,00
Happy Planet Index_1_N	1,76540	18,5982	0,09	0,924	5,84	0,00
Private property rights (0=no_N	-4,28302	38,8761	-0,11	0,912	0,01	0,00
Access to electricity (% of p_N	-5,93067	106,907	-0,06	0,956	0,00	0,00
Availability of affordable ho_N	-1,92397	49,4233	-0,04	0,969	0,15	0,00
GINNI P	-8,14614	73,0430	-0,11	0,911	0,00	0,00

Predictor	Upper
Logit 1: (2/3)	
Constant	
IDH - 2013_1_N	1,39492E+143
Shelter_1N	9,01210E+221
Access to improved sanitation_N	3,03475E+226
Happy Planet Index_1_N	1,79368E+92
Private property rights (0=no_N	4,34106E+106
Access to electricity (% of p_N	7,04665E+168
Availability of affordable ho_N	1,57916E+198
GINNI P	1,52320E+130
Logit 2: (1/3)	
Constant	
IDH - 2013_1_N	1,46446E+34
Shelter_1N	2,29129E+177
Access to improved sanitation_N	1,75355E+59

```
Happy Planet Index_1_N            3,95852E+16
Private property rights (0=no_N   1,70346E+31
Access to electricity (% of p_N   2,65303E+88
Availability of affordable ho_N   1,71257E+41
GINNI P                           4,33009E+58


Log-Likelihood = -0,000
Test that all slopes are zero: G = 260,406, DF = 16, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson    0,0000360  246  1,000
Deviance   0,0000720  246  1,000
```

5. Comparação: Análise Discriminante X Regressão Logística

Quadro comparative: Análise Discriminante X Regressão Logística

	3 Clusters	3 Regiões
AD- Análise Discriminante	93,2% Quadrática	85,6% Quadrática
RL- Regressão Logística	98,4%	89,1%

De acordo com a comparação acima, podemos verificar que a Análise que apresentou maior percentual de confiança foi a Regressão Logística para 3 Clusters com 98,4%, seguida da Análise Discriminante Quadrática para 3 Clusters com 93,2% de confiança.

Nos resultados da Análise Discriminante, demos preferência para as Quadráticas por apresentarem maior porcentagem de acertos, além de maior número de n-correct.

CAPITULO 11 - ÁRVORE DE REGRESSÃO

1. INTRODUÇÃO

Para esta etapa do trabalho, passaremos a trabalhar com o software estatístico IBM® SPSS® Decision Trees, que nos ajudará a melhor identificar grupos e descobrir relacionamentos entre estes grupos para predizer eventos futuros.

Pretendemos, portanto apresentar de maneira visual resultados categóricos de maneira intuitiva, no entanto, iremos experimentar diferentes tipos de “crescimento em árvore” para localizarmos qual que melhor se ajusta aos nossos dados.

2. TÉCNICAS DE CRESCIMENTO EM ÁRVORE

Temos neste software técnicas de crescimento em árvore especializadas para classificação dentro do ambiente do IBM SPSS Statistics. Os quatro algoritmos de crescimento em árvore incluem:

- **CHAID**—um algoritmo de árvore rápido, estatístico e de várias vias que explora dados com rapidez e eficiência e constrói segmentos e perfis com relação ao resultado desejado.
- **CHAID Exaustivo**—uma modificação do CHAID, que examina todas as divisões possíveis para cada preditor.
- **Classification and regression trees (C&RT)**—um algoritmo de árvore binária completo que particiona dados e produz subgrupos homogêneos precisos.
- **QUEST**—um algoritmo estatístico que seleciona variáveis sem propensão e constrói árvores binárias precisas com rapidez e eficiência.

2.1 Discriminant

Notes

Output Created		16-Mai-2016 22h8min55s
Comments		
Input	Data	C:\Users\Rafael\Desktop\Árvores Aula 10_05_16.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing in the analysis phase.
	Cases Used	In the analysis phase, cases with no user- or system-missing values for any predictor variable are used. Cases with user-, system-missing, or out-of-range values for the grouping variable are always excluded.
Syntax		DISCRIMINANT /GROUPS=VAR00012(1 3) /VARIABLES=VAR00001 VAR00002 VAR00003 VAR00004 VAR00005 VAR00006 VAR00007 VAR00009 /ANALYSIS ALL /METHOD=WILKS /FIN=3.84 /FOUT=2.71 /PRIORS EQUAL /HISTORY /STATISTICS=TABLE /CLASSIFY=NONMISSING POOLED.
Resources	Processor Time	00:00:00,032
	Elapsed Time	00:00:00,069

[DataSet1] C:\Users\Rafael\Desktop\Árvores Aula 10_05_16.sav

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		132	100,0
Excluded	Missing or out-of-range group codes	0	,0
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	0	,0
Total		132	100,0

Group Statistics

3R ORD		Valid N (listwise)	
		Unweighted	Weighted
1,0000	IDH_N	27	27,000
	SHELTER	27	27,000
	IMPROVED SANITATION	27	27,000
	HPI	27	27,000
	PRIVATE PROPERTY HOUSE	27	27,000
	ACCESS TO ELETRICITY	27	27,000
	AFFORDABLE HOUSE	27	27,000
	GINNI	27	27,000
2,0000	IDH_N	21	21,000
	SHELTER	21	21,000
	IMPROVED SANITATION	21	21,000
	HPI	21	21,000
	PRIVATE PROPERTY HOUSE	21	21,000
	ACCESS TO ELETRICITY	21	21,000
	AFFORDABLE HOUSE	21	21,000
	GINNI	21	21,000

3,0000	IDH_N	84	84,000
	SHELTER	84	84,000
	IMPROVED SANITATION	84	84,000
	HPI	84	84,000
	PRIVATE PROPERTY HOUSE	84	84,000
	ACCESS TO ELETRICITY	84	84,000
	AFFORDABLE HOUSE	84	84,000
	GINNI	84	84,000
Total	IDH_N	132	132,000
	SHELTER	132	132,000
	IMPROVED SANITATION	132	132,000
	HPI	132	132,000
	PRIVATE PROPERTY HOUSE	132	132,000
	ACCESS TO ELETRICITY	132	132,000
	AFFORDABLE HOUSE	132	132,000
	GINNI	132	132,000

Analysis 1

2.2 Stepwise Statistics

Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda			
		Statistic	df1	df2	df3
1	PRIVATE PROPERTY HOUSE	,426	1	2	129,000
2	HPI	,316	2	2	129,000
3	GINNI	,243	3	2	129,000
4	IDH_N	,225	4	2	129,000
5	ACCESS TO ELETRICITY	,211	5	2	129,000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 16.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables Entered/Removed[a,b,c,d]

Step	Wilks' Lambda			
	Exact F			
	Statistic	df1	df2	Sig.
1	87,002	2	129,000	,000
2	49,784	4	256,000	,000
3	43,566	6	254,000	,000
4	34,844	8	252,000	,000
5	29,476	10	250,000	,000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 16.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	PRIVATE PROPERTY HOUSE	1,000	87,002	
2	PRIVATE PROPERTY HOUSE	1,000	85,653	,740
	HPI	1,000	22,124	,426
3	PRIVATE PROPERTY HOUSE	,998	74,745	,529
	HPI	,983	22,976	,331
	GINNI	,981	19,215	,316
4	PRIVATE PROPERTY HOUSE	,826	37,841	,361
	HPI	,906	13,029	,272
	GINNI	,890	20,214	,298
	IDH_N	,710	4,875	,243

5	PRIVATE PROPERTY HOUSE	,805	26,226	,299
	HPI	,811	11,125	,248
	GINNI	,815	22,383	,286
	IDH_N	,232	6,253	,232
	ACCESS TO ELETRICITY	,216	4,402	,225

Variables Not in the Analysis

Step		Tolerance	Min. Tolerance	F to Enter	Wilks' Lambda
0	IDH_N	1,000	1,000	43,455	,597
	SHELTER	1,000	1,000	37,645	,631
	IMPROVED SANITATION	1,000	1,000	18,302	,779
	HPI	1,000	1,000	22,688	,740
	PRIVATE PROPERTY HOUSE	1,000	1,000	87,002	,426
	ACCESS TO ELETRICITY	1,000	1,000	15,483	,806
	AFFORDABLE HOUSE	1,000	1,000	6,039	,914
	GINNI	1,000	1,000	24,449	,725
1	IDH_N	,863	,863	9,372	,371
	SHELTER	,861	,861	9,174	,372
	IMPROVED SANITATION	,940	,940	4,445	,398
	HPI	1,000	1,000	22,124	,316
	ACCESS TO ELETRICITY	,959	,959	6,070	,389
	AFFORDABLE HOUSE	,992	,992	1,421	,416
	GINNI	,998	,998	18,379	,331
2	IDH_N	,783	,783	3,967	,298
	SHELTER	,759	,759	2,694	,303
	IMPROVED SANITATION	,842	,842	,932	,312
	ACCESS TO ELETRICITY	,791	,791	,578	,314
	AFFORDABLE HOUSE	,941	,941	,899	,312
	GINNI	,981	,981	19,215	,243
3	IDH_N	,710	,710	4,875	,225
	SHELTER	,701	,701	3,939	,229
	IMPROVED SANITATION	,746	,746	2,038	,235
	ACCESS TO ELETRICITY	,660	,660	3,047	,232
	AFFORDABLE HOUSE	,911	,911	2,234	,235
4	SHELTER	,241	,241	,330	,224
	IMPROVED SANITATION	,298	,284	1,334	,221
	ACCESS TO ELETRICITY	,216	,216	4,402	,211
	AFFORDABLE HOUSE	,907	,707	2,224	,218
5	SHELTER	,146	,131	1,738	,205
	IMPROVED SANITATION	,237	,171	,102	,210

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda	
1	PRIVATE PROPERTY HOUSE	1,000	87,002		
2	PRIVATE PROPERTY HOUSE	1,000	85,653	,740	
	HPI	1,000	22,124	,426	
3	PRIVATE PROPERTY HOUSE	,998	74,745	,529	
	HPI	,983	22,976	,331	
	GINNI	,981	19,215	,316	
4	PRIVATE PROPERTY HOUSE	,826	37,841	,361	
	HPI	,906	13,029	,272	
	GINNI	,890	20,214	,298	
	IDH_N	,710	4,875	,243	
5	PRIVATE PROPERTY HOUSE	,805	26,226	,299	
	HPI	,811	11,125	,248	
	GINNI	,815	22,383	,286	
	IDH_N	,232	6,253	,232	
	AFFORDABLE HOUSE	,896	,213	2,768	,202

Wilks' Lambda

Step	Number of Variables	Lambda	df1	df2	df3
1	1	,426	1	2	129
2	2	,316	2	2	129
3	3	,243	3	2	129
4	4	,225	4	2	129
5	5	,211	5	2	129

Wilks' Lambda

Step	Exact F			
	Statistic	df1	df2	Sig.
1	87,002	2	129,000	,000
2	49,784	4	256,000	,000
3	43,566	6	254,000	,000
4	34,844	8	252,000	,000
5	29,476	10	250,000	,000

2.3 Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1,851[a]	73,6	73,6	,806
2	,665[a]	26,4	100,0	,632

a. First 2 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	,211	197,839	10	,000
2	,600	64,784	4	,000

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
IDH_N	,693	,448
HPI	,052	,680
PRIVATE PROPERTY	,752	-,020
HOUSE		
ACCESS TO ELETRICITY	-,685	,018
GINNI	,470	-,671

Structure Matrix

	Function	
	1	2
PRIVATE PROPERTY	,847*	,178
HOUSE		
IDH_N	,538*	,455
HPI	,024	,726*
GINNI	,365	-,446*
SHELTER[a]	,411	,436*
ACCESS TO ELETRICITY	,266	,404*
IMPROVED SANITATION[a]	,349	,366*
AFFORDABLE HOUSE[a]	-,028	,249*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

a. This variable not used in the analysis.

Functions at Group Centroids

3R ORD	Function	
	1	2
1,0000	2,650	,069
2,0000	-,812	1,789
3,0000	-,649	-,469

Unstandardized canonical discriminant functions evaluated at group means

2.4 Classification Statistics

Classification Processing Summary

Processed		132
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		132

Prior Probabilities for Groups

3R ORD	Prior	Cases Used in Analysis	
		Unweighted	Weighted
1,0000	,333	27	27,000
2,0000	,333	21	21,000
3,0000	,333	84	84,000
Total	1,000	132	132,000

Classification Results[a]

		3R ORD	Predicted Group Membership			
			1,0000	2,0000	3,0000	Total
Original	Count	1,0000	25	0	2	27
		2,0000	3	18	0	21
		3,0000	4	16	64	84
	%	1,0000	92,6	,0	7,4	100,0
		2,0000	14,3	85,7	,0	100,0
		3,0000	4,8	19,0	76,2	100,0

a. 81,1% of original grouped cases correctly classified.

Classification Tree

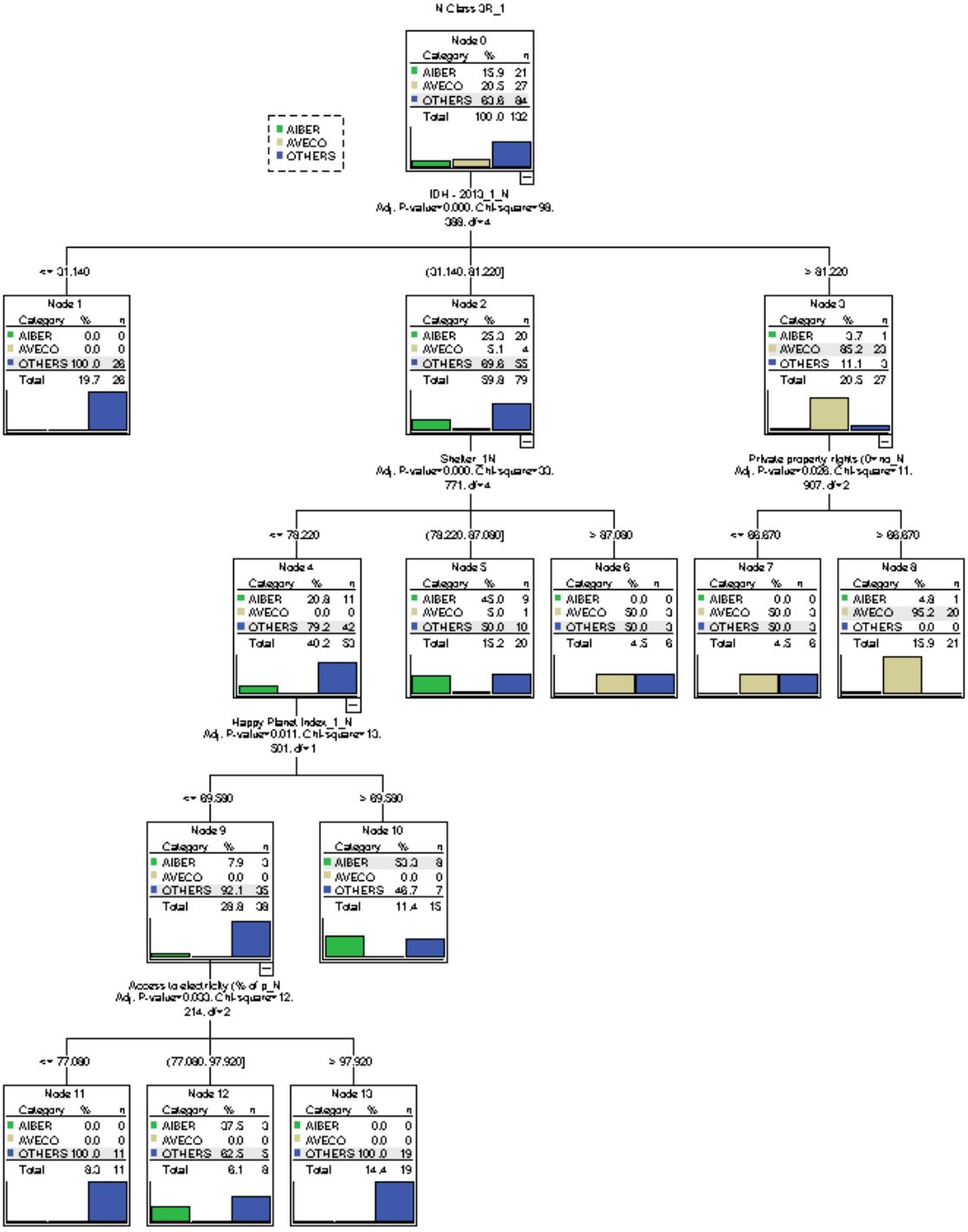
3. ANÁLISE CHAID EXAUSTIVE

Warnings

Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	N Class 3R_1
	Independent Variables	IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Private property rights (0=no_N, Access to electricity (% of p_N, Availability of affordable ho_N, GINNI P
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH - 2013_1_N, Shelter_1N, Happy Planet Index_1_N, Access to electricity (% of p_N, Private property rights (0=no_N
	Number of Nodes	14
	Number of Terminal Nodes	9
	Depth	4



Risk

Estimate	Std. Error
,205	,035

Growing Method:

EXHAUSTIVE CHAID

Dependent Variable: N

Class 3R_1

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	8	1	12	38,1%
AVECO	0	20	7	74,1%
OTHERS	7	0	77	91,7%
Overall Percentage	11,4%	15,9%	72,7%	79,5%

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: N Class 3R_1

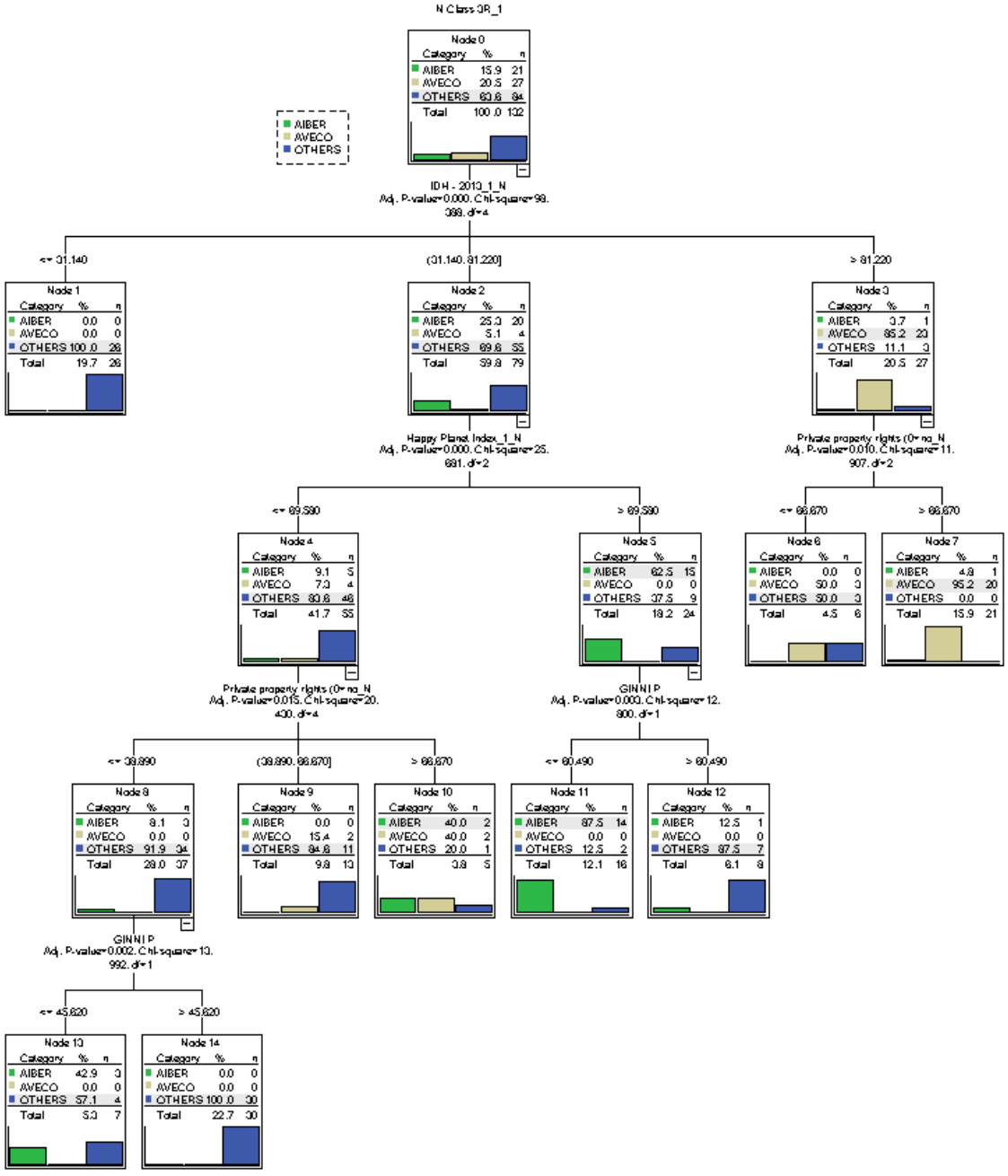
4. ANÁLISE CHAID

Warnings

Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	N Class 3R_1
	Independent Variables	IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Private property rights (0=no_N, Access to electricity (% of p_N, Availability of affordable ho_N, GINNI P
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH - 2013_1_N, Happy Planet Index_1_N, Private property rights (0=no_N, GINNI P
	Number of Nodes	15
	Number of Terminal Nodes	9
	Depth	4



Risk

Estimate	Std. Error
,114	,028

Growing Method: CHAID

Dependent Variable: N

Class 3R_1

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	16	1	4	76,2%
AVECO	2	20	5	74,1%
OTHERS	3	0	81	96,4%
Overall Percentage	15,9%	15,9%	68,2%	88,6%

Growing Method: CHAID

Dependent Variable: N Class 3R_1

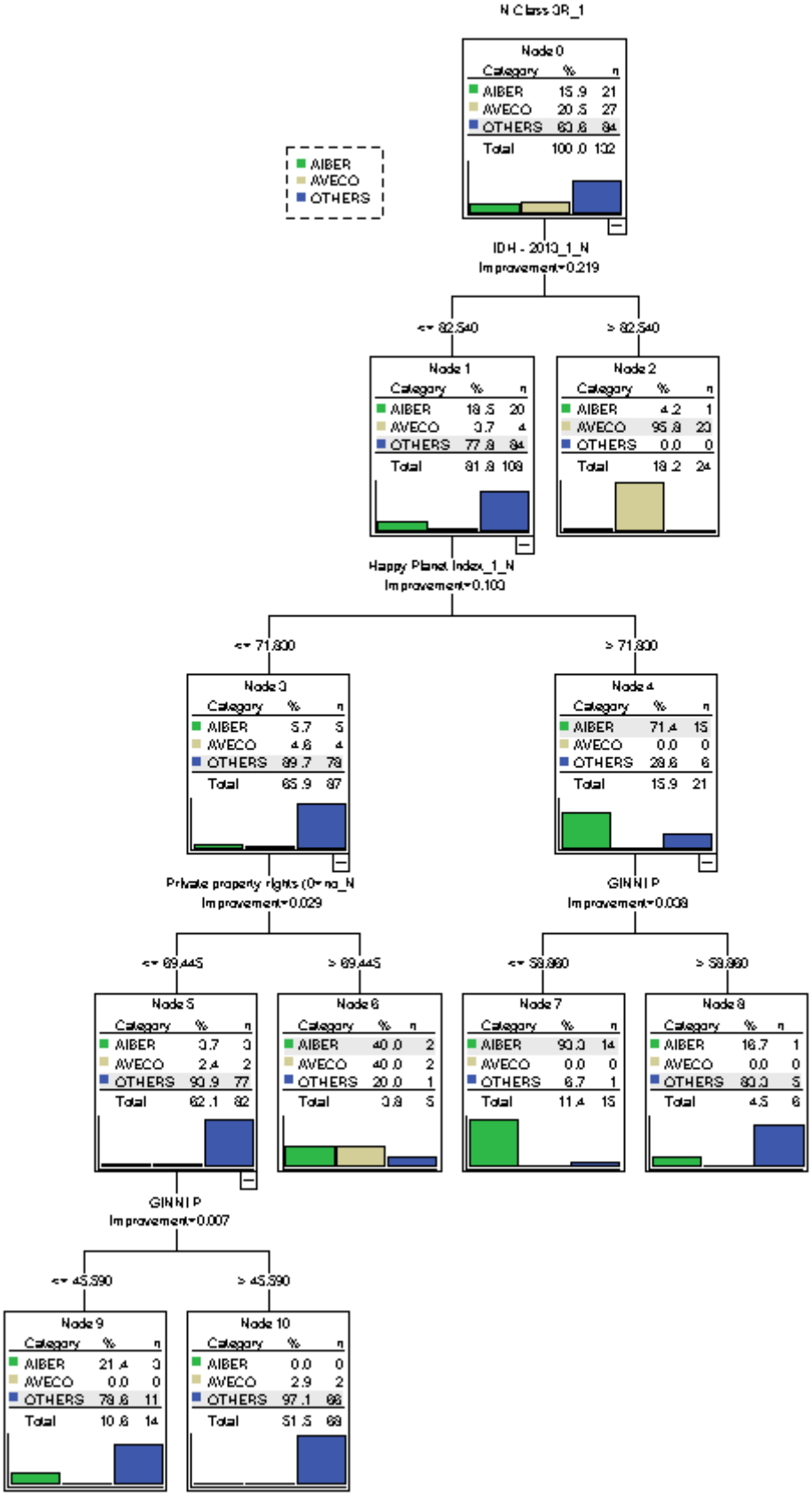
5. ANÁLISE CRT

Warnings

Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	CRT
	Dependent Variable	N Class 3R_1
	Independent Variables	IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Private property rights (0=no_N, Access to electricity (% of p_N, Availability of affordable ho_N, GINNI P
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH - 2013_1_N, Private property rights (0=no_N, Shelter_1N, Access to improved sanitation_N, GINNI P, Availability of affordable ho_N, Happy Planet Index_1_N, Access to electricity (% of p_N
	Number of Nodes	11
	Number of Terminal Nodes	6
	Depth	4


N Class OR_1
AIBER
AVECO
OTHERS
Node 0
Category % n
AIBER 15.9 21
AVECO 20.5 27
OTHERS 63.6 84
Total 100.0 132
IDH - 2013_1_N
Improvement=0.219
<= 82.540
> 82.540
Node 1
Category % n
AIBER 18.5 20
AVECO 3.7 4
OTHERS 77.8 84
Total 81.8 108
Node 2
Category % n
AIBER 4.2 1
AVECO 95.8 23
OTHERS 0.0 0
Total 18.2 24
Happy Planet Index_1_N
Improvement=0.100
<= 71.800
> 71.800
Node 3
Category % n
AIBER 5.7 5
AVECO 4.6 4
OTHERS 89.7 78
Total 65.9 87
Node 4
Category % n
AIBER 71.4 15
AVECO 0.0 0
OTHERS 28.6 6
Total 15.9 21
Private property rights (0= no_N
Improvement=0.029
GINNI P
Improvement=0.038
<= 69.445
> 69.445
<= 59.880
> 59.880
Node 5
Category % n
AIBER 3.7 3
AVECO 2.4 2
OTHERS 93.9 77
Total 62.1 82
Node 6
Category % n
AIBER 40.0 2
AVECO 40.0 2
OTHERS 20.0 1
Total 3.8 5
Node 7
Category % n
AIBER 93.3 14
AVECO 0.0 0
OTHERS 6.7 1
Total 11.4 15
Node 8
Category % n
AIBER 16.7 1
AVECO 0.0 0
OTHERS 83.3 5
Total 4.5 6
GINNI P
Improvement=0.007
<= 45.390
> 45.390
Node 9
Category % n
AIBER 21.4 3
AVECO 0.0 0
OTHERS 78.6 11
Total 10.6 14
Node 10
Category % n
AIBER 0.0 0
AVECO 2.9 2
OTHERS 97.1 68
Total 51.5 68

Risk

Estimate	Std. Error
,083	,024

Growing Method: CRT

Dependent Variable: N

Class 3R_1

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	16	1	4	76,2%
AVECO	2	23	2	85,2%
OTHERS	2	0	82	97,6%
Overall Percentage	15,2%	18,2%	66,7%	91,7%

Growing Method: CRT

Dependent Variable: N Class 3R_1

6. ANÁLISE QUEST

Warnings

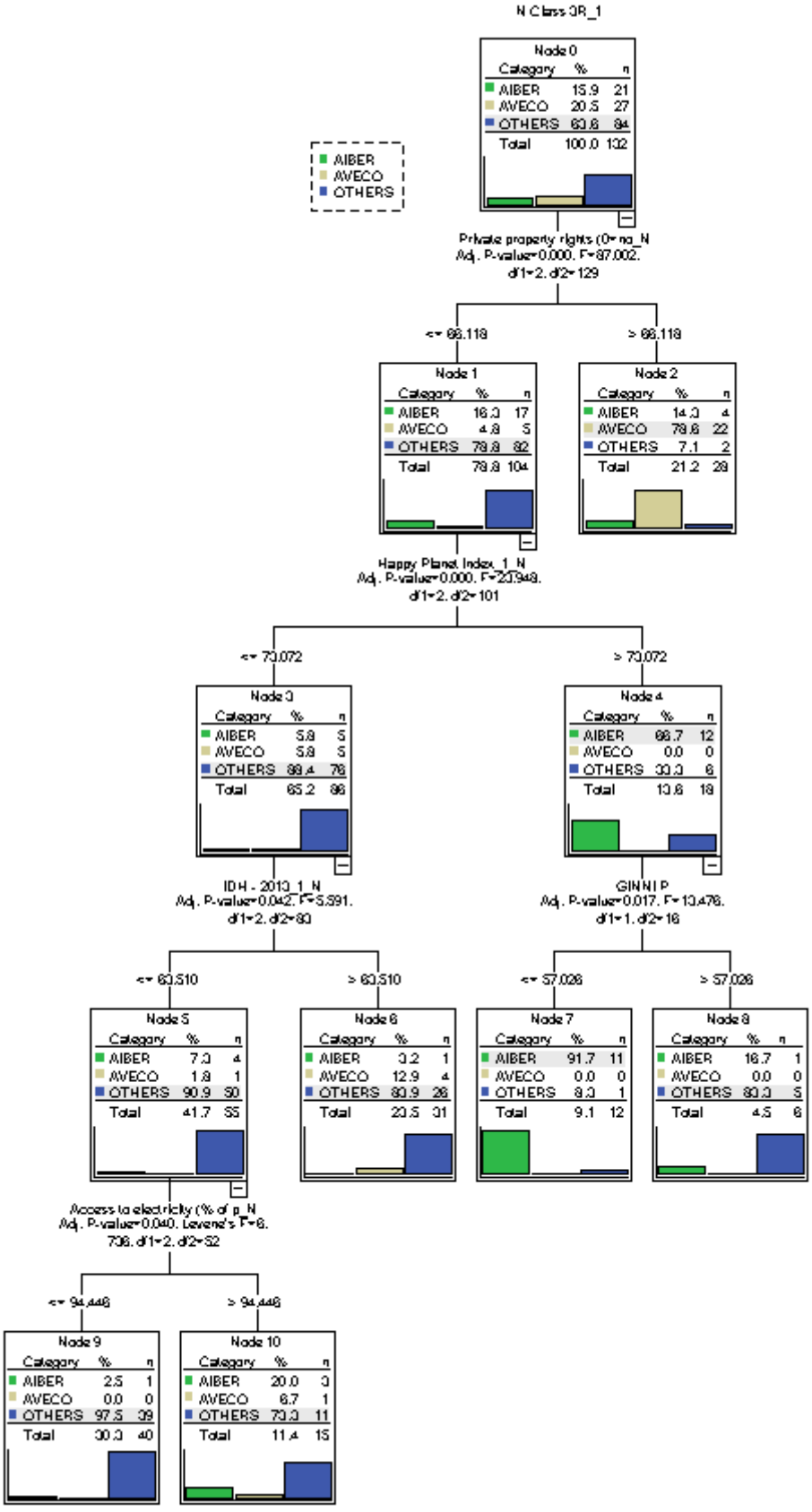
Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	QUEST
	Dependent Variable	N Class 3R_1
	Independent Variables	IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Private property rights (0=no_N, Access to electricity (% of p_N, Availability of affordable ho_N, GINNI P
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	Private property rights (0=no_N, IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Availability of affordable ho_N, Access to electricity (% of p_N, GINNI P
	Number of Nodes	11
	Number of Terminal Nodes	6

Model Summary

Specifications	Growing Method	QUEST
	Dependent Variable	N Class 3R_1
	Independent Variables	IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Private property rights (0=no_N, Access to electricity (% of p_N, Availability of affordable ho_N, GINNI P
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	Private property rights (0=no_N, IDH - 2013_1_N, Shelter_1N, Access to improved sanitation_N, Happy Planet Index_1_N, Availability of affordable ho_N, Access to electricity (% of p_N, GINNI P
	Number of Nodes	11
	Number of Terminal Nodes	6
	Depth	4


N Class OR_1
AIBER
AVECO
OTHERS
Node 0
Category % n
AIBER 15.9 21
AVECO 20.5 27
OTHERS 63.6 84
Total 100.0 132
Private property rights (0=no_N
Adj. P-value=0.000, F=87.002,
df1=2, df2=129
<= 68.118
> 68.118
Node 1
Category % n
AIBER 16.3 17
AVECO 4.8 5
OTHERS 78.8 82
Total 78.8 104
Node 2
Category % n
AIBER 14.3 4
AVECO 78.6 22
OTHERS 7.1 2
Total 21.2 28
Happy Planet Index_1_N
Adj. P-value=0.000, F=23.948,
df1=2, df2=101
<= 70.072
> 70.072
Node 3
Category % n
AIBER 5.8 5
AVECO 5.8 5
OTHERS 88.4 76
Total 65.2 86
Node 4
Category % n
AIBER 66.7 12
AVECO 0.0 0
OTHERS 33.3 6
Total 13.6 18
IDH - 2013_1_N
Adj. P-value=0.042, F=5.591,
df1=2, df2=83
GINNI P
Adj. P-value=0.017, F=13.476,
df1=1, df2=16
<= 60.510
> 60.510
<= 57.026
> 57.026
Node 5
Category % n
AIBER 7.3 4
AVECO 1.8 1
OTHERS 90.9 50
Total 41.7 55
Node 6
Category % n
AIBER 3.2 1
AVECO 12.9 4
OTHERS 83.9 26
Total 23.5 31
Node 7
Category % n
AIBER 91.7 11
AVECO 0.0 0
OTHERS 8.3 1
Total 9.1 12
Node 8
Category % n
AIBER 16.7 1
AVECO 0.0 0
OTHERS 83.3 5
Total 4.5 6
Access to electricity (% of p_N
Adj. P-value=0.040, Levene's F=6.
736, df1=2, df2=52
<= 94.446
> 94.446
Node 9
Category % n
AIBER 2.5 1
AVECO 0.0 0
OTHERS 97.5 39
Total 30.3 40
Node 10
Category % n
AIBER 20.0 3
AVECO 6.7 1
OTHERS 73.3 11
Total 11.4 15

Risk

Estimate	Std. Error
,136	,030

Growing Method: QUEST

Dependent Variable: N

Class 3R_1

Classification

Observed	Predicted			
	AIBER	AVECO	OTHERS	Percent Correct
AIBER	11	4	6	52,4%
AVECO	0	22	5	81,5%
OTHERS	1	2	81	96,4%
Overall Percentage	9,1%	21,2%	69,7%	86,4%

Growing Method: QUEST

Dependent Variable: N Class 3R_1

7. ANÁLISE LOGÍSTICA

PLUM - Ordinal Regression

Warnings

There are 264 (66,7%) cells (i.e., dependent variable levels by combinations of predictor variable values) with zero frequencies.

The log-likelihood value is practically zero. There may be a complete separation in the data. The maximum likelihood estimates do not exist.

The PLUM procedure continues despite the above warning(s). Subsequent results shown are based on the last iteration. Validity of the model fit is uncertain.

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	238,837			
Final	,000	238,837	131	,000

Link function: Logit.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	,000	131	1,000
Deviance	,001	131	1,000

Link function: Logit.

Pseudo R-Square

Cox and Snell	,836
Nagelkerke	1,000
McFadden	1,000

Link function: Logit.

CAPITULO 12 – ANÁLISE POR CORRESPONDÊNCIA

1. INTRODUÇÃO

A análise por correspondência é um método de análise fatorial para variáveis categóricas. Basicamente, converte uma tabela de dados não negativos de duas ou múltiplas entradas em um tipo de representação gráfica em que as linhas e as colunas são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico. Este método permite mostrar como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe.

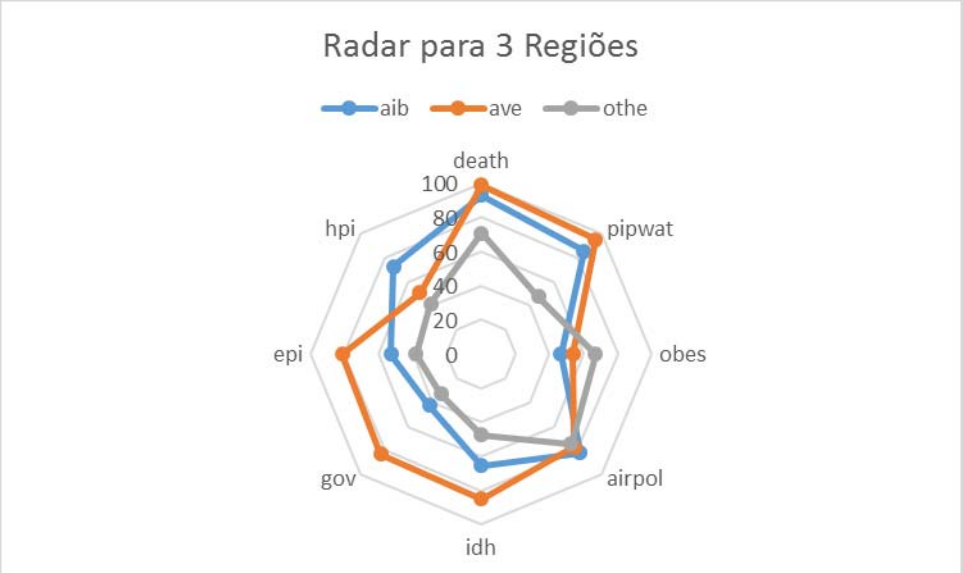
As análises foram efetuadas para cruzamento e conhecer a correspondência, de média e desvio padrão, entre 8 variáveis selecionadas (IDH, Shelter, Access to Improved Sanitation, Happy Planet Index, Private Property Rights, Access to Eletricity, Availability of Affordable House e Ginni) e as 3 clusters ordenadas.

2. Análise por Correspondência das Médias

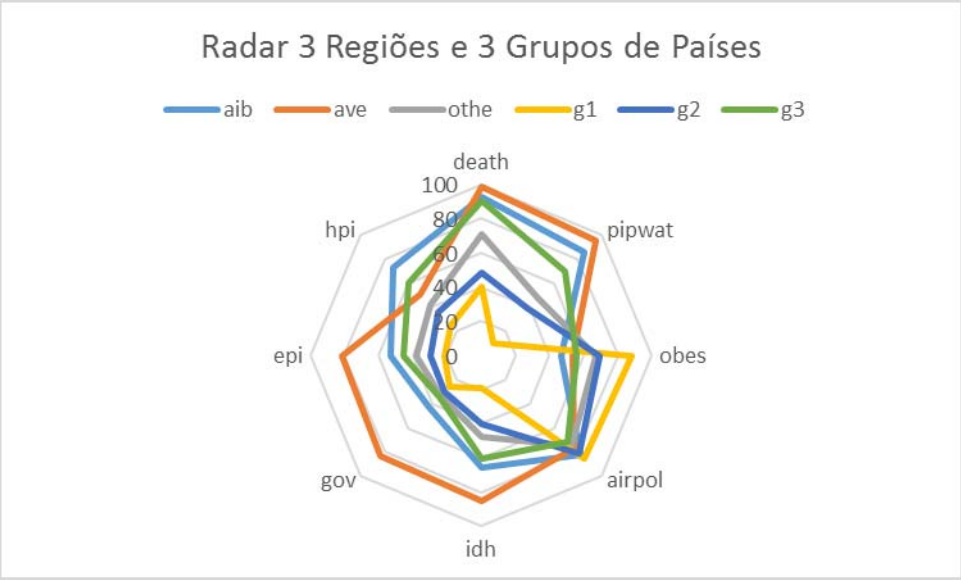
Abaixo a tabela com dados para análise de correspondência e respectivos gráficos radares.

Tabela de Dados – Variáveis, Clusters e Médias

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni
aib	73,97	65,61	80,63	72,36	36,77	92,81	46,38	43,08
ave	91,64	87,36	96,75	50,58	82,51	99,73	55,1	79,91
othe	53,15	46,74	60,27	41,79	32,18	66,42	40,8	62,44
g1	23,31	19,43	19,73	24,11	24,1	21,83	34,31	60,9
g2	64,45	56,85	74,33	56,73	32,78	84,12	44,19	56,46
g3	90,29	85,17	96,9	47,23	76,54	99,39	52,05	77,35



Podemos verificar no radar acima que para todas as variáveis analisadas (exceto HPI – Happy Planet Index), o grupo AVECO é que se mostra melhor classificado. Em seguida temos AIBER melhor classificado, exceto na variável Ginni, onde se encontra abaixo do grupo Others. Na Variável Private Property Rights os grupor AIBER e Others estão muito próximos, enquanto AVECO está bem distante, ou seja, muito melhor classificada.



Podemos verificar no Radar acima, que quando juntamos na mesma representação gráfica as 3 regiões e os 3 grupos de países, percebemos que o g1 está bastante defasado em relação às médias das demais variáveis. Percebemos que há grande relação entre AVECO e g3, onde no radar estão representados quase sempre sobrepostos. Há também forte relação entre o g2 e AIBER, mas neste caso não visualizamos estes índices se sobrepondo em nenhuma das variáveis.

Simple Correspondence Analysis: shelter; IDH; ImpSan; HPI; PPR; AccElet; AffHou

Contingency Table

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni
aib	73,970	65,610	80,630	72,360	36,770	92,810	46,380	43,080
ave	91,640	87,360	96,750	50,580	82,510	99,730	55,100	79,910
othe	53,150	46,740	60,270	41,790	32,180	66,420	40,800	62,440
g1	23,310	19,430	19,730	24,110	24,100	21,830	34,310	60,900
g2	64,450	56,850	74,330	56,730	32,780	84,120	44,190	56,460
g3	90,290	85,170	96,900	47,230	76,540	99,390	52,050	77,350
Total	396,810	361,160	428,610	292,800	284,880	464,300	272,830	380,140

	Total
aib	511,610
ave	643,580
othe	403,790
g1	227,720
g2	469,910
g3	624,920
Total	2881,530

Chi-Square Distances

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
aib	0,176	0,034	0,270	7,985	3,771	1,306	0,088	8,831	22,459
ave	0,102	0,556	0,011	3,357	5,604	0,152	0,559	0,294	10,634
othe	0,108	0,296	0,001	0,014	1,501	0,028	0,173	1,579	3,700
g1	2,066	2,909	5,904	0,041	0,112	6,020	7,538	31,698	56,288
g2	0,001	0,071	0,281	1,689	4,027	0,933	0,002	0,494	7,498
g3	0,208	0,598	0,168	4,169	3,525	0,017	0,857	0,314	9,856
Total	2,662	4,464	6,635	17,254	18,539	8,456	9,216	43,209	110,434

Relative Inertias

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
aib	0,002	0,000	0,002	0,072	0,034	0,012	0,001	0,080	0,203
ave	0,001	0,005	0,000	0,030	0,051	0,001	0,005	0,003	0,096
othe	0,001	0,003	0,000	0,000	0,014	0,000	0,002	0,014	0,034
g1	0,019	0,026	0,053	0,000	0,001	0,055	0,068	0,287	0,510
g2	0,000	0,001	0,003	0,015	0,036	0,008	0,000	0,004	0,068
g3	0,002	0,005	0,002	0,038	0,032	0,000	0,008	0,003	0,089
Total	0,024	0,040	0,060	0,156	0,168	0,077	0,083	0,391	1,000

Analysis of Contingency Table

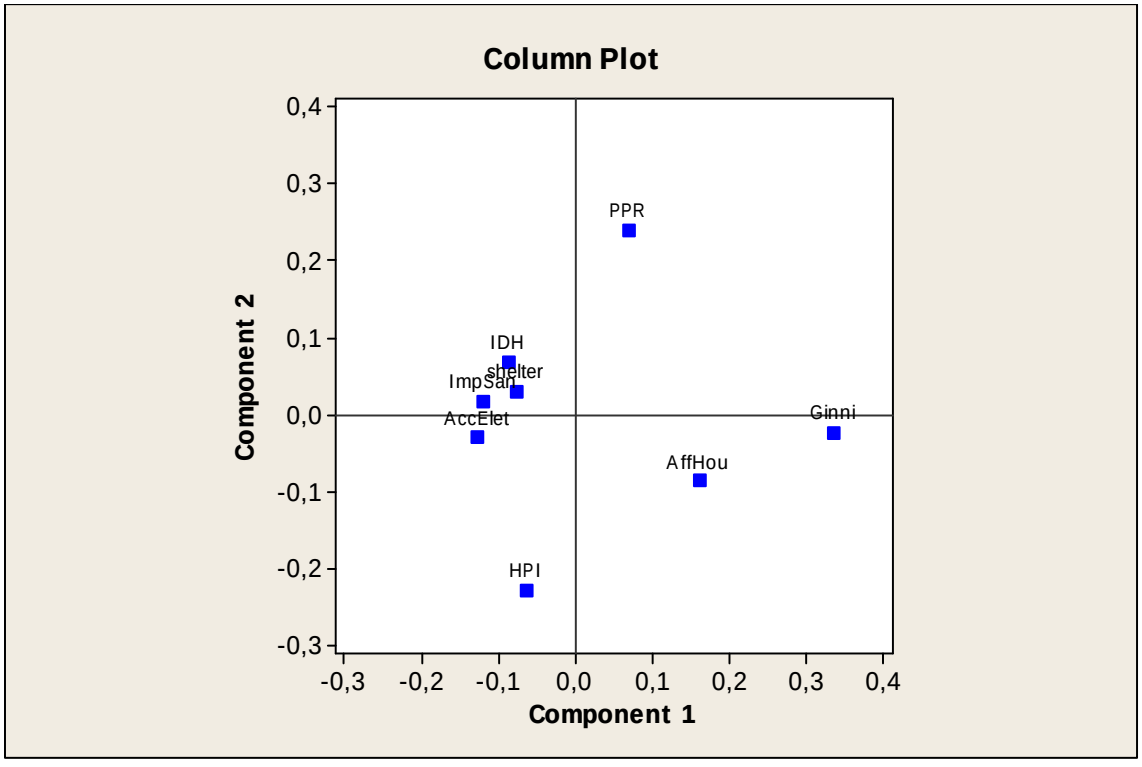
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0246	0,6426	0,6426	******************************
2	0,0127	0,3315	0,9741	***************
3	0,0010	0,0258	0,9999	*
4	0,0000	0,0001	1,0000	
Total	0,0383			

Row Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	aib	0,959	0,178	0,203	-0,162	0,599	0,189	-0,126	0,360	0,221
2	ave	0,984	0,223	0,096	-0,008	0,004	0,001	0,127	0,981	0,285
3	othe	0,712	0,140	0,034	0,051	0,284	0,015	-0,063	0,428	0,043
4	g1	0,996	0,079	0,510	0,488	0,962	0,763	-0,091	0,033	0,051
5	g2	0,945	0,163	0,068	-0,065	0,267	0,028	-0,104	0,678	0,139
6	g3	0,996	0,217	0,089	-0,021	0,027	0,004	0,124	0,968	0,261

Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	shelter	0,996	0,138	0,024	-0,076	0,855	0,032	0,031	0,141	0,010
2	IDH	0,996	0,125	0,040	-0,087	0,615	0,039	0,069	0,381	0,046
3	ImpSan	0,934	0,149	0,060	-0,119	0,914	0,085	0,018	0,020	0,004
4	HPI	0,952	0,102	0,156	-0,064	0,069	0,017	-0,228	0,883	0,416
5	PPR	0,962	0,099	0,168	0,068	0,072	0,019	0,241	0,890	0,451
6	AccElet	0,954	0,161	0,077	-0,128	0,904	0,108	-0,030	0,050	0,012
7	AffHou	0,988	0,095	0,083	0,161	0,768	0,100	-0,086	0,220	0,055
8	Ginni	0,992	0,132	0,391	0,335	0,987	0,601	-0,023	0,005	0,006



Simple Correspondence Analysis: shelter; IDH; ImpSan; HPI; PPR; AccElet; AffHou

Contingency Table

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni
aib	73,970	65,610	80,630	72,360	36,770	92,810	46,380	43,080
ave	91,640	87,360	96,750	50,580	82,510	99,730	55,100	79,910
othe	53,150	46,740	60,270	41,790	32,180	66,420	40,800	62,440
g1	23,310	19,430	19,730	24,110	24,100	21,830	34,310	60,900
g2	64,450	56,850	74,330	56,730	32,780	84,120	44,190	56,460
g3	90,290	85,170	96,900	47,230	76,540	99,390	52,050	77,350
Total	396,810	361,160	428,610	292,800	284,880	464,300	272,830	380,140

	Total
aib	511,610
ave	643,580
othe	403,790
g1	227,720
g2	469,910
g3	624,920
Total	2881,530

Chi-Square Distances

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
aib	0,176	0,034	0,270	7,985	3,771	1,306	0,088	8,831	22,459
ave	0,102	0,556	0,011	3,357	5,604	0,152	0,559	0,294	10,634
othe	0,108	0,296	0,001	0,014	1,501	0,028	0,173	1,579	3,700
g1	2,066	2,909	5,904	0,041	0,112	6,020	7,538	31,698	56,288
g2	0,001	0,071	0,281	1,689	4,027	0,933	0,002	0,494	7,498
g3	0,208	0,598	0,168	4,169	3,525	0,017	0,857	0,314	9,856
Total	2,662	4,464	6,635	17,254	18,539	8,456	9,216	43,209	110,434

Relative Inertias

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
aib	0,002	0,000	0,002	0,072	0,034	0,012	0,001	0,080	0,203
ave	0,001	0,005	0,000	0,030	0,051	0,001	0,005	0,003	0,096
othe	0,001	0,003	0,000	0,000	0,014	0,000	0,002	0,014	0,034
g1	0,019	0,026	0,053	0,000	0,001	0,055	0,068	0,287	0,510
g2	0,000	0,001	0,003	0,015	0,036	0,008	0,000	0,004	0,068
g3	0,002	0,005	0,002	0,038	0,032	0,000	0,008	0,003	0,089
Total	0,024	0,040	0,060	0,156	0,168	0,077	0,083	0,391	1,000

Analysis of Contingency Table

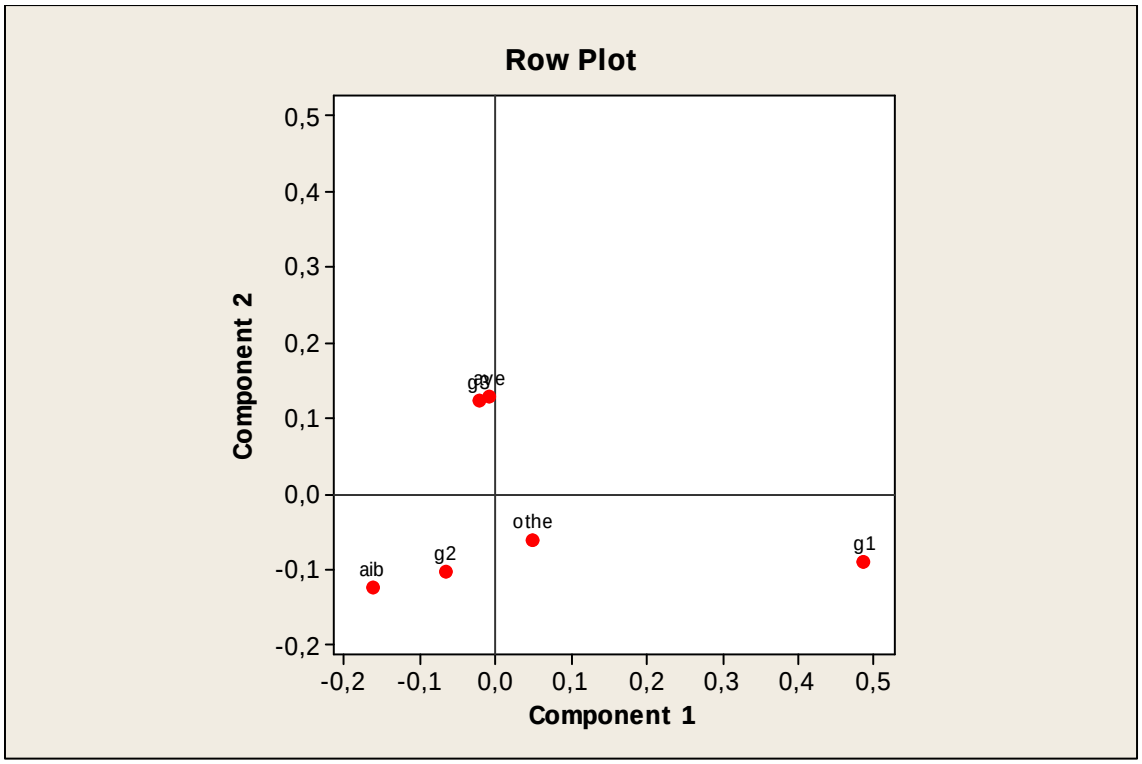
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0246	0,6426	0,6426	******************************
2	0,0127	0,3315	0,9741	***************
3	0,0010	0,0258	0,9999	*
4	0,0000	0,0001	1,0000	
Total	0,0383			

Row Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	aib	0,959	0,178	0,203	-0,162	0,599	0,189	-0,126	0,360	0,221
2	ave	0,984	0,223	0,096	-0,008	0,004	0,001	0,127	0,981	0,285
3	othe	0,712	0,140	0,034	0,051	0,284	0,015	-0,063	0,428	0,043
4	g1	0,996	0,079	0,510	0,488	0,962	0,763	-0,091	0,033	0,051
5	g2	0,945	0,163	0,068	-0,065	0,267	0,028	-0,104	0,678	0,139
6	g3	0,996	0,217	0,089	-0,021	0,027	0,004	0,124	0,968	0,261

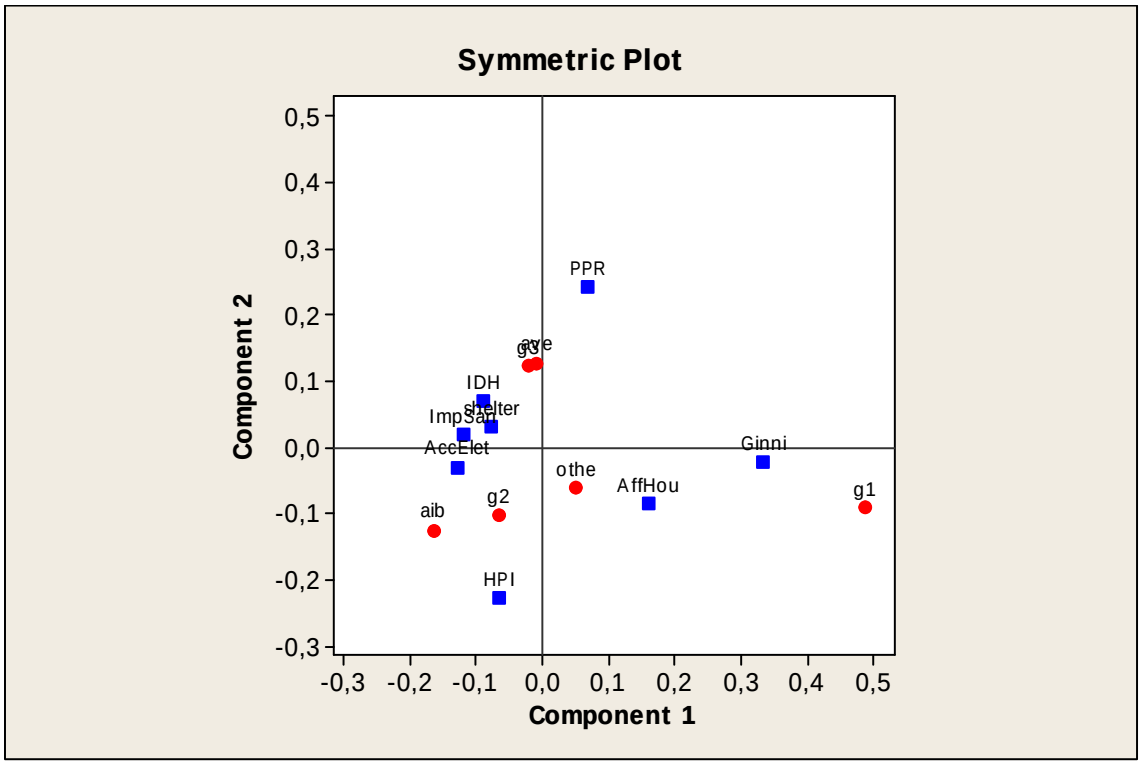
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	shelter	0,996	0,138	0,024	-0,076	0,855	0,032	0,031	0,141	0,010
2	IDH	0,996	0,125	0,040	-0,087	0,615	0,039	0,069	0,381	0,046
3	ImpSan	0,934	0,149	0,060	-0,119	0,914	0,085	0,018	0,020	0,004
4	HPI	0,952	0,102	0,156	-0,064	0,069	0,017	-0,228	0,883	0,416
5	PPR	0,962	0,099	0,168	0,068	0,072	0,019	0,241	0,890	0,451
6	AccElet	0,954	0,161	0,077	-0,128	0,904	0,108	-0,030	0,050	0,012
7	AffHou	0,988	0,095	0,083	0,161	0,768	0,100	-0,086	0,220	0,055
8	Ginni	0,992	0,132	0,391	0,335	0,987	0,601	-0,023	0,005	0,006



Principais Considerações: Temos na representação gráfica acima a posição em coordenadas das regiões AIBER, AVECO e Others e dos três grupos de países por ordem crescente de g1 a g3. Onde g1 estão os países piores classificados e g3 os países melhores classificados.

Verificamos portanto, que há forte correlação entre os países AVECO e o grupo g3, ao passo que também há correlação entre AIBER e g2, porém com uma aproximação um pouco maior entre estes grupos.



Principais Considerações: No gráfico acima, podemos verificar que a variável Ginni está distante da maioria dos grupos de países. As variáveis que mais se aproximam do g2/AIBER são Happy Planet Index e Access to Eletricity e as variáveis que mais se aproximam do g2/AVECO são IDH, Shelter e Improved Sanitation.

CAPITULO 13 – ANÁLISE POR CORRESPONDÊNCIA SOMENTE AIBER

1. INTRODUÇÃO

O trabalho a seguir corresponde a uma análise multivariada dos dados compilados pelo Banco Mundial, com a utilização do *software* MINITAB. A finalidade é apresentar uma análise das correlações dos componentes principais de dados dimensionadores de MORADIA apenas dos países que compõem a região AIBER. A princípio, faz-se necessário uma análise de estatística descritiva, prosseguindo-se a continuação, às correlações, dendogramas, regressões com o uso do método stepwise.

2. Results for: Somente Aiber

Cluster Analysis of Observations: Shelter_1N; IDH - 2013_1; Access to im; ...

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	88,6334	37,07	18	20	18	2
2	19	86,4226	44,28	11	12	11	2
3	18	84,0370	52,06	3	5	3	2
4	17	83,6599	53,29	18	19	18	3
5	16	81,2099	61,28	8	9	8	2
6	15	78,9685	68,59	2	14	2	2
7	14	77,1594	74,49	3	13	3	3
8	13	76,9724	75,10	8	21	8	3
9	12	75,1633	81,00	10	15	10	2
10	11	72,8145	88,66	3	6	3	4
11	10	70,5179	96,15	1	7	1	2
12	9	68,0128	104,32	10	17	10	3
13	8	63,7813	118,12	11	16	11	3
14	7	62,5211	122,23	4	18	4	4
15	6	58,7772	134,44	8	11	8	6
16	5	57,9493	137,14	8	10	8	9
17	4	41,4007	191,11	2	8	2	11
18	3	41,0235	192,34	1	3	1	6
19	2	35,1302	211,56	1	4	1	10
20	1	0,0000	326,13	1	2	1	21

Final Partition
Number of clusters: 3

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	6	7844,4	34,7578	54,6075
Cluster2	11	11586,1	31,6335	46,2882
Cluster3	4	2558,8	23,2821	40,0709

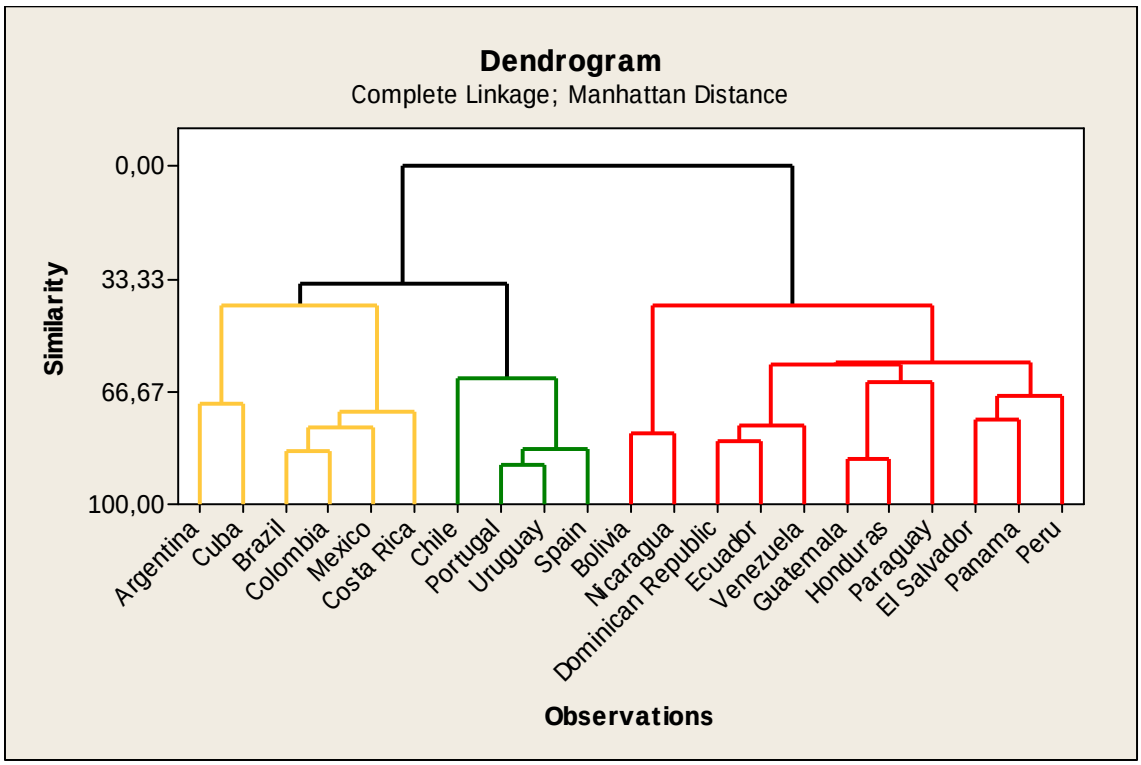
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Grand centroid
Shelter_1N	77,5933	67,7636	85,6000	73,9695
IDH - 2013_1_N	70,7017	57,4209	80,5175	65,6148
Access to improved sanitation_N	86,3067	70,7336	99,3325	80,6305
Happy Planet Index_1_N	82,1617	74,5482	51,6225	72,3567

Private property rights (0=no_N	36,1117	22,2236	77,7750	36,7729
Access to electricity (% of p_N	96,8767	88,0709	99,7400	92,8095
Availability of affordable ho_N	37,6733	55,1664	35,2775	46,3800
GINNI P	40,6083	39,9409	55,4400	43,0838
Social Progress Index_1N	65,1317	57,2609	82,9750	64,4076

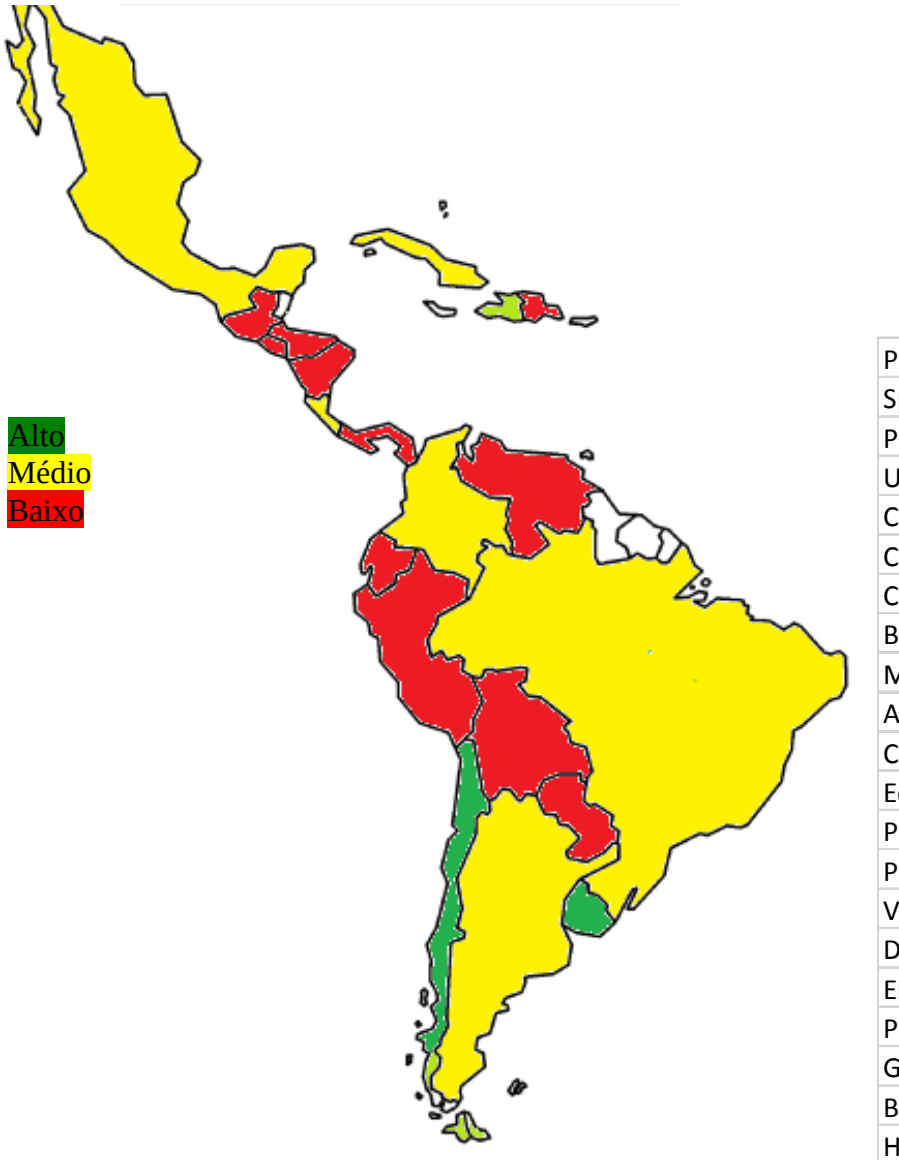
Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3
Cluster1	0,0000	34,8182	59,5895
Cluster2	34,8182	0,0000	81,9402
Cluster3	59,5895	81,9402	0,0000



Acima temos o dendograma gerado para o Ranking dos países, conforme mapa no item 3.

3. MAPA DE DESENVOLVIMENTO PARA MORADIA AIBER



Países Ranking	Ranking	Ranking0a100
Spain	1	100
Portugal	1	97,68
Uruguay	1	87,04
Chile	1	80,40
Costa Rica	2	66,60
Cuba	2	58,83
Brazil	2	55,26
Mexico	2	53,80
Argentina	2	51,30
Colombia	2	45,29
Ecuador	3	48,97
Panama	3	41,85
Peru	3	41,75
Venezuela	3	40,74
Dominican Republic	3	38,61
El Salvador	3	35,61
Paraguay	3	27,74
Guatemala	3	16,02
Bolivia	3	10,07
Honduras	3	8,07
Nicaragua	3	0,00

ANÁLISE: Conforme podemos observar no mapa acima, avaliamos que os Países que compõem a AIBER ainda possuem diversos desafios para alcançar grau de desenvolvimento na questão da moradia. Estes desafios se concentram em diversos setores, desde Governança a Infra-estrutura, até condições sanitárias e de instalações de redes elétricas e de esgoto.

Dos 4 países que possuem elevado nível de condições em moradia 2 estão na Península Ibérica - Espanha e Portugal primeira e segunda colocadas, respectivamente – Brasil e Argentina se juntam a Costa Rica, Cuba, México e Colômbia compondo o grupo dos países em posição média no ranking. Fechando o Ranking temos Paraguai, Guatemala, Bolívia, Honduras e Nicarágua nas últimas colocações.

3.1 One-way ANOVA: IDH - 2013_1_N versus 3cluster

```
Source    DF      SS     MS      F      P
3cluster   2  1782,2  891,1  14,39  0,000
Error     18  1114,3   61,9
Total     20  2896,4

S = 7,868   R-Sq = 61,53%   R-Sq(adj) = 57,26%
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  --------+---------+---------+---------+-
1       6  70,702  6,501            (------*-----)
2      11  57,421  9,038  (----*----)
3       4  80,517  5,359                      (--------*-------)
                          --------+---------+---------+---------+-
                                  60        70        80        90

Pooled StDev = 7,868
```

3.2 One-way ANOVA: PC1 versus 3cluster

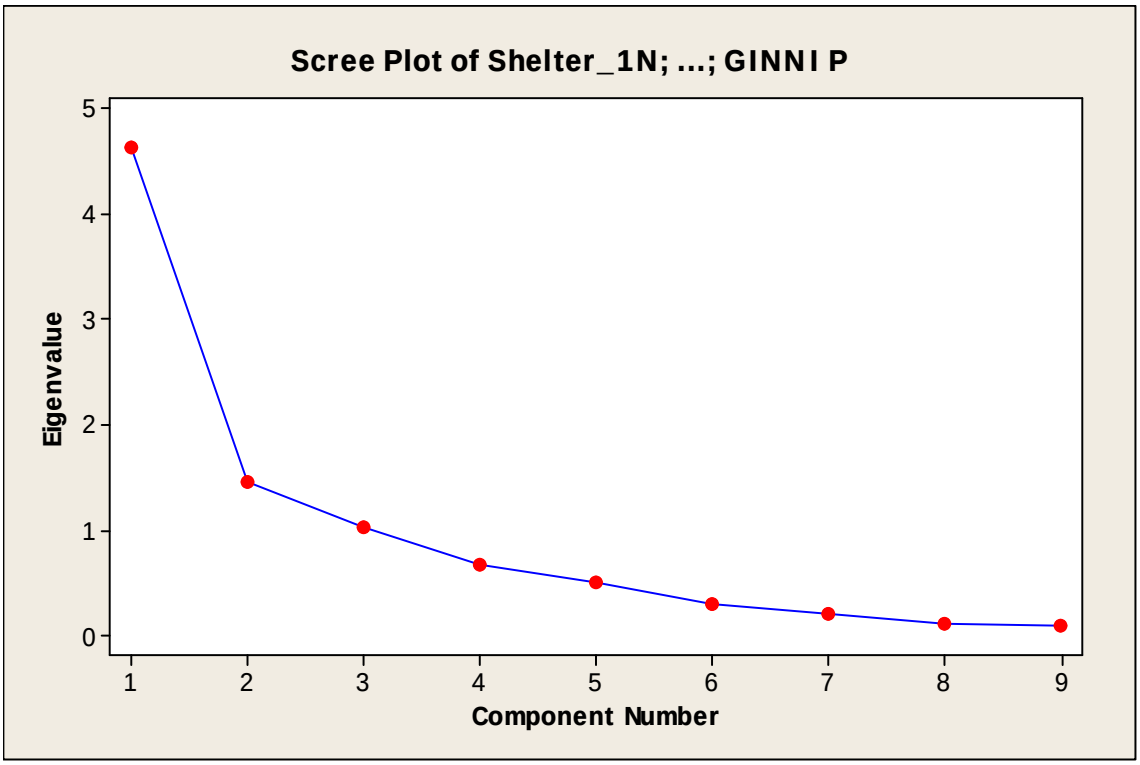
```
Source    DF     SS     MS      F      P
3cluster   2  72,58  36,29  33,01  0,000
Error     18  19,79   1,10
Total     20  92,37

S = 1,049   R-Sq = 78,58%   R-Sq(adj) = 76,19%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean  StDev    -+---------+---------+---------+--------
1       6   0,564  0,556              (----*---)
2      11  -1,527  1,294    (--*---)
3       4   3,355  0,710                       (-----*----)
                            -+---------+---------+---------+--------
                          -2,0       0,0       2,0       4,0

Pooled StDev = 1,049
```

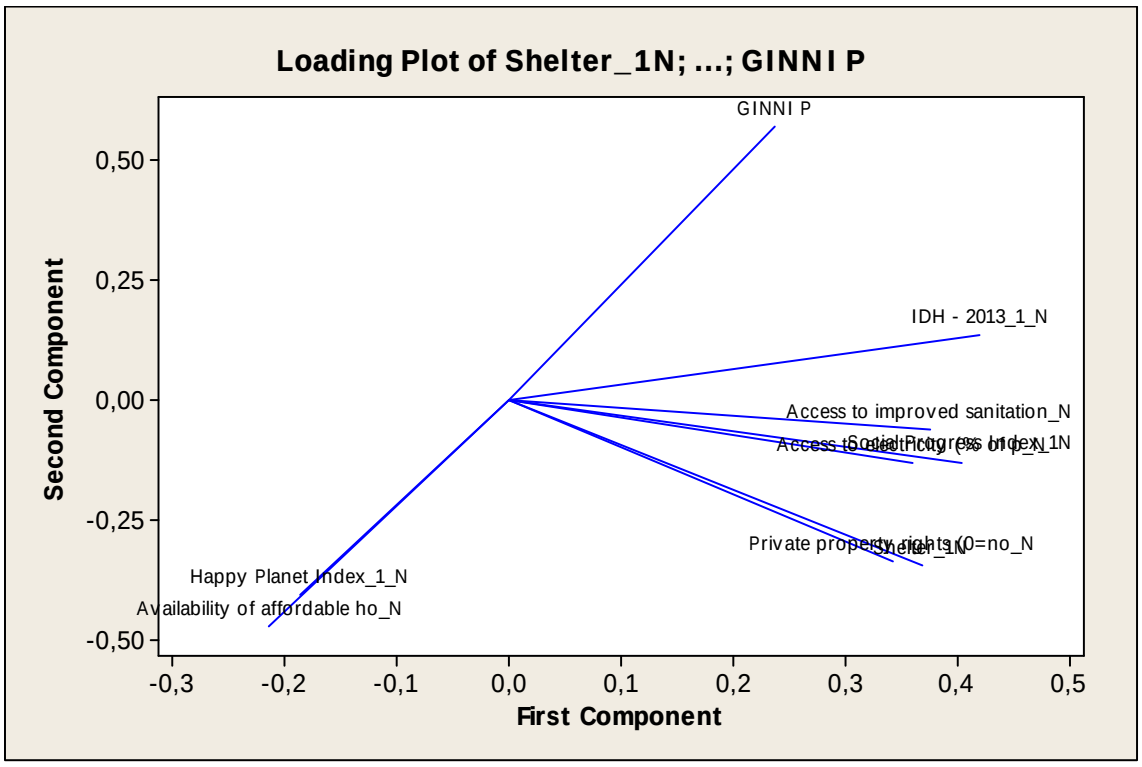
Rodamos duas ANOVAS para conferirmos qual nos daria melhores resultados de confiança para explicabilidade do ranking e do mapa. Conforme podemos observar a ANOVA com o PC1 para 3 Clusters nos apresentou R^2=78,58% e F=33,01, ficando portanto acima da ANOVA IDH 2013 para 3 Clusters que ficou com R^2=61,53 e F=14,39.

Principal Component Analysis: Shelter_1N; Social Progr; IDH - 2013_1; Access to

```
Eigenanalysis of the Correlation Matrix

Eigenvalue  4,6186  1,4625  1,0255  0,6803  0,5077  0,2952  0,2032  0,1206
Proportion   0,513   0,162   0,114   0,076   0,056   0,033   0,023   0,013
Cumulative   0,513   0,676   0,790   0,865   0,922   0,954   0,977   0,990

Eigenvalue  0,0863
Proportion   0,010
Cumulative   1,000

Variable                          PC1     PC2     PC3     PC4
Shelter_1N                      0,368  -0,346   0,046  -0,033
Social Progress Index_1N        0,403  -0,130   0,201   0,307
IDH - 2013_1_N                  0,420   0,134  -0,200   0,016
Access to improved sanitation_N 0,376  -0,064  -0,325  -0,159
Happy Planet Index_1_N         -0,186  -0,406  -0,689   0,090
```

Private property rights (0=no_N	0,343	-0,338	0,349	0,322
Access to electricity (% of p_N	0,360	-0,132	-0,204	-0,552
Availability of affordable ho_N	-0,214	-0,471	0,407	-0,592
GINNI P	0,237	0,568	0,089	-0,335





Simple Correspondence Analysis: shelter; IDH; ImpSan; HPI; PPR; AccElet; AffHou

Contingency Table

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni
g1	85,600	80,510	99,330	51,620	77,780	99,740	35,280	55,440
g2	77,590	70,700	86,310	82,160	36,110	96,880	37,670	40,610
g3	67,760	57,420	70,730	74,550	22,220	88,070	55,170	39,940
Total	230,950	208,630	256,370	208,330	136,110	284,690	128,120	135,990

	Total
g1	585,300
g2	528,030
g3	475,860
Total	1589,190

Chi-Square Distances

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
g1	0,003	0,175	0,255	8,216	15,252	0,249	3,004	0,572	27,728
g2	0,009	0,027	0,015	2,419	1,837	0,055	0,564	0,463	5,390
g3	0,028	0,408	0,475	2,374	8,430	0,094	7,362	0,015	19,186
Total	0,041	0,611	0,745	13,009	25,519	0,398	10,931	1,051	52,304

Relative Inertias

	shelter	IDH	ImpSan	HPI	PPR	AccElet	AffHou	Ginni	Total
g1	0,000	0,003	0,005	0,157	0,292	0,005	0,057	0,011	0,530
g2	0,000	0,001	0,000	0,046	0,035	0,001	0,011	0,009	0,103
g3	0,001	0,008	0,009	0,045	0,161	0,002	0,141	0,000	0,367
Total	0,001	0,012	0,014	0,249	0,488	0,008	0,209	0,020	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0298	0,9065	0,9065	******************************
2	0,0031	0,0935	1,0000	***
Total	0,0329			

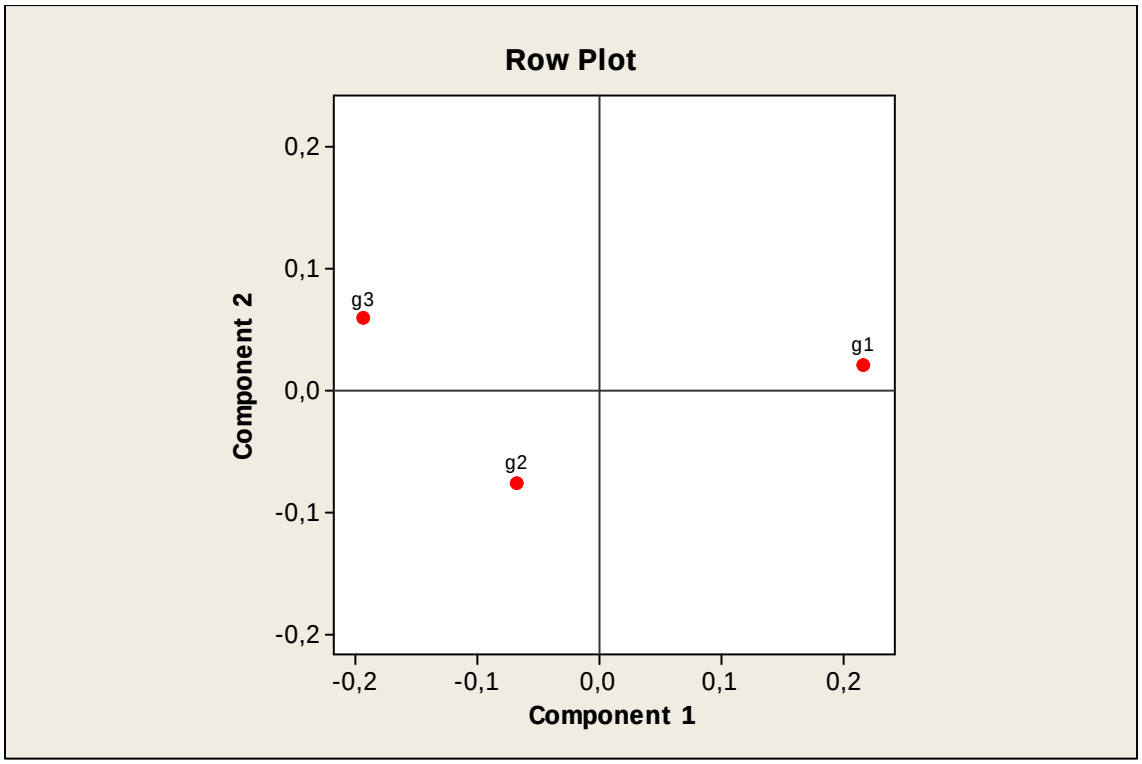
Row Contributions

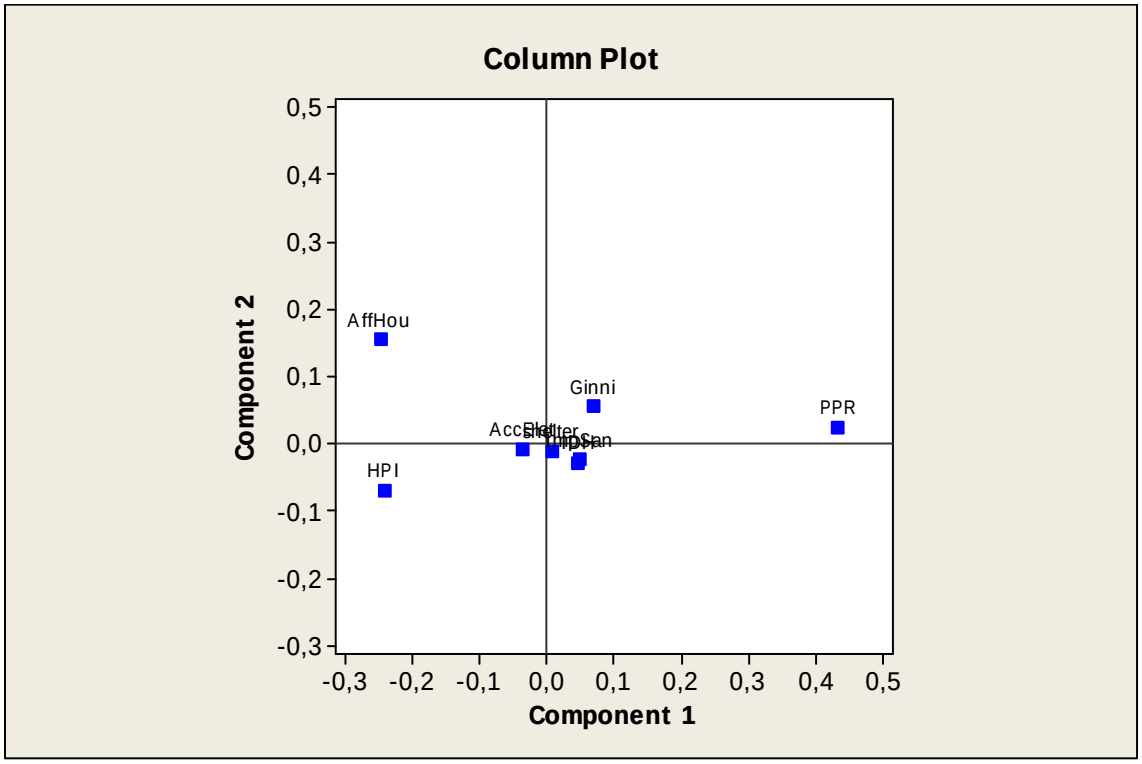
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	g1	1,000	0,368	0,530	0,217	0,991	0,579	0,021	0,009	0,052
2	g2	1,000	0,332	0,103	-0,067	0,439	0,050	-0,076	0,561	0,618
3	g3	1,000	0,299	0,367	-0,192	0,916	0,371	0,058	0,084	0,330

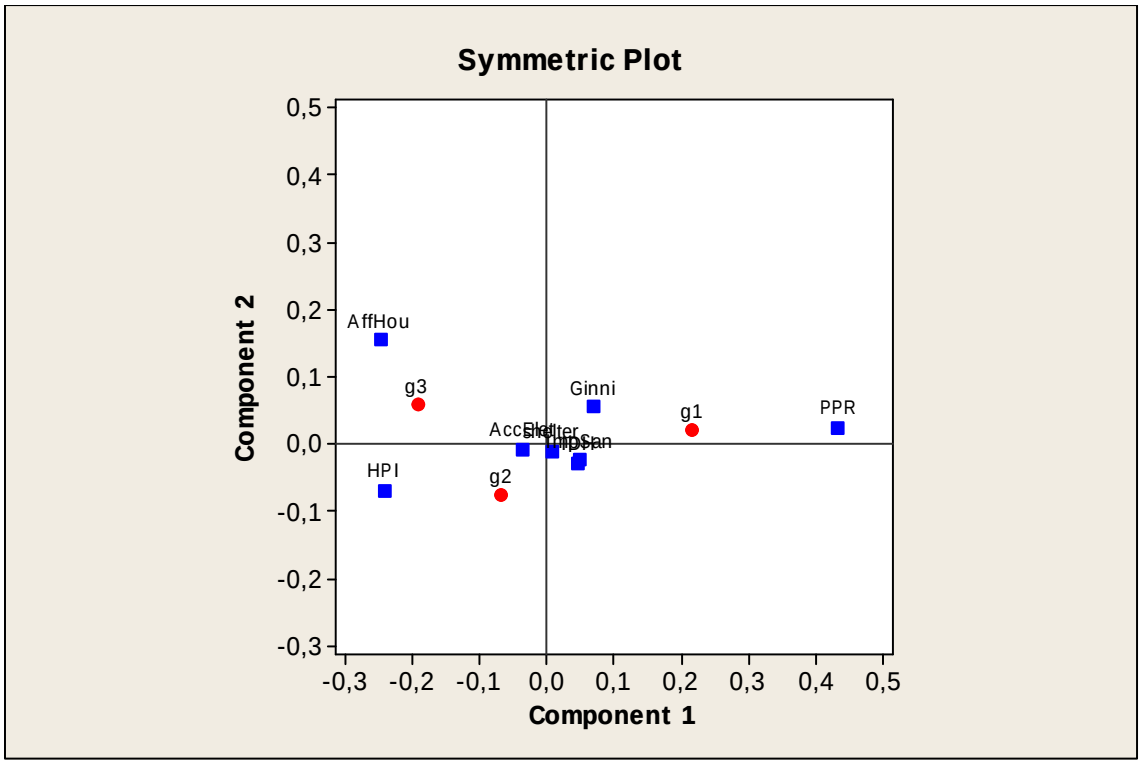
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	shelter	1,000	0,145	0,001	0,008	0,380	0,000	-0,010	0,620	0,005
2	IDH	1,000	0,131	0,012	0,046	0,736	0,009	-0,028	0,264	0,033
3	ImpSan	1,000	0,161	0,014	0,049	0,810	0,013	-0,023	0,190	0,029
4	HPI	1,000	0,131	0,249	-0,240	0,924	0,254	-0,069	0,076	0,201
5	PPR	1,000	0,086	0,488	0,432	0,997	0,536	0,025	0,003	0,017
6	AccElet	1,000	0,179	0,008	-0,037	0,962	0,008	-0,007	0,038	0,003
7	AffHou	1,000	0,081	0,209	-0,248	0,719	0,166	0,155	0,281	0,628
8	Ginni	1,000	0,086	0,020	0,069	0,613	0,014	0,055	0,387	0,083

* NOTE * There are no supplementary points to plot


Row Plot
Component 2
Component 1
g1
g2
g3
0,2
0,1
0,0
-0,1
-0,2
-0,2
-0,1
0,0
0,1
0,2


Column Plot
Component 2
Component 1
AffHou
Ginni
PPR
HPI
0,5
0,4
0,3
0,2
0,1
0,0
-0,1
-0,2
-0,3
-0,3
-0,2
-0,1
0,0
0,1
0,2
0,3
0,4
0,5



g1 representa os países melhores desenvolvidos, g2 os países em desenvolvimento médio e g3 os países piores avaliados.

4. CONSIDERAÇÕES FINAIS

Ao sobrepormos os grupos de países (g1, g2 e g3) com as 8 principais variáveis para o tema Moradia, podemos verificar que o g1 (melhores desenvolvidos) é o grupo que mais se aproxima de *Private Property Rights* e *Ginni*. O grupo g2 se aproxima de *Happy Planet Index* e *Access to Eletricity*. Curiosamente o grupo g3 ficou próximo de *Availability of affordable house*, dando-nos a sensação de avaliarem positivamente as condições de moradia e abrigo a que estão inseridas.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA: BEM-ESTAR E QUALIDADE DE VIDA

TRABALHO FINAL

Disciplina: Métodos Quantitativos e Qualitativos da Pesquisa Empírica

Prof. Dr. Arnoldo José de Hoyos Guevara

Aluna: Tatiana Almendra Dutra

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados relativos ao tema que possui um conjunto de vinte variáveis previamente selecionadas que refletem o desenvolvimento humano considerando 132 países, conforme mostrado no Quadro

1 que apresenta as Variáveis e sua categorização em três e dezessete quantitativas, enquanto a Tabela 1 que apresenta algumas variáveis relacionadas com os 132 países.
Iniciamos o trabalho apresentando o objetivo do Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP – Pontifícia Universidade Católica. Depois partimos para o entendimento dos dados apresentando: média, mínimo, máximo, padronização e normalização dos dados e, por fim, a análise exploratória dos dados sobre o tema: bem-estar e qualidade de vida empregando software estatístico MINITAB, em especial seus recursos: Display Descriptive Statistics que está no link principal Basic Statistics, ele permite obter a média, o valor mínimo e o valor máximo do objeto estudado.
Seguimos com a aplicação do link principal Data, acessando o Code – Numeric to Numeric, sendo que este recurso permite que os dados ausentes que aparecem na tabela com o símbolo asterisco (*) seja substituído pelo valor determinado que é a média encontrada.
Na sequencia temos que fazer a normalização dos dados que é um ajuste nos dados com vista a afastar os *outliers*, ou seja, é como se dos dados analisados não possuíssem uma distribuição normal então fazemos o ajuste para que todos fiquem dentro de uma certa normalidade. No nosso caso, usamos como parâmetro médio entre o mínimo e o máximo obtido na coleta de dados. Temos que apresentar os valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas.
Para essa etapa usamos o link da barra de ferramentas do MINITAB denominada Calculator e aplicamos a fórmula desejada indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.
Por fim, esse trabalho apresentamos os gráficos gerados pelo recurso Graphical Summary que está no item Basic Statistic que por sua vez está no item Static na barra de ferramentas.
O Objetivo é explorar os dados sobre bem-estar e qualidade de vida de tal modo que possamos efetuar uma análise das correlações e dos componentes principais (análise multivariada) de dados dimensionadores em diversos países do mundo separados por AVECo, AIBER e OTHERS. Para tal, iniciamos com análise da estatística descritiva. Em seguida passamos para a análise das correlações e dendogramas. Na terceira parte utilizamos a análise dos componentes principais.

2. ENTENDENDO OS DADOS

Antes da apresentação dos indivíduos desta análise que envolve 132 países e os indicadores de desenvolvimento humano selecionados a partir de referência internacional, citamos a origem do trabalho que é o Projeto ORIBER.

2.1 O tema de pesquisa: bem-estar e qualidade de vida

A análise dos conceitos que compõem o tema foi extraída do material de aula Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica. Nesse estudo são apresentadas

as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso.
Vale destacar o desafio citado no GPS e sua proposta de solução:
Desafio: *"A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?"*

Solução: *"GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."*

Para chegar a uma conclusão do estágio em que se encontra o país e como é possível adotar uma solução, o GPS propõe um levantamento dos dados a partir da análise SWOT, conforme se apresenta na figura 1.

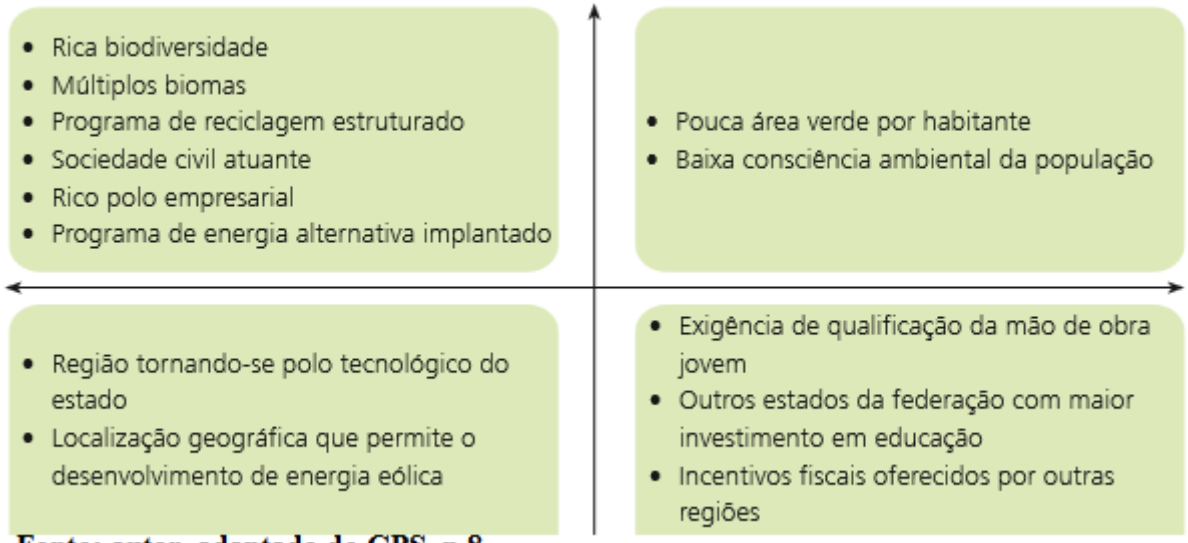
Figura 1: Análise SWOT de cada país

fonte: autor, adaptado de GPS, p.8

Na figura 1 é possível obter informações sobre o país analisado considerando sua Forças – quais são os seus pontos fortes em cada um dos eixos temáticos. As fragilidades, ameaças e oportunidades que cada eixo temático apresenta para o país em questão.
Como exemplo inicial desse trabalho que tem o tema: bem-estar e qualidade de vida, podemos destacar o alimento. Supondo que o país analisado seja um grande produtor de alimento (ponto forte) e que tenha uma estiagem em seu território (fragilidade) e que perceba uma ameaça (outro país produtor) mas percebe uma oportunidade a de planejar a produção daquele alimento usando sistema de irrigação por gotejamento a partir do armazenamento e do uso racional desse recurso.

Para ajudar o leitor a entender a contribuição da matriz SWOT, o GPS apresenta um exemplo com o tema: Bens naturais Comuns, citado na figura 2.

Figura 2: Análise SWOT do eixo temático Bens Naturais Comuns

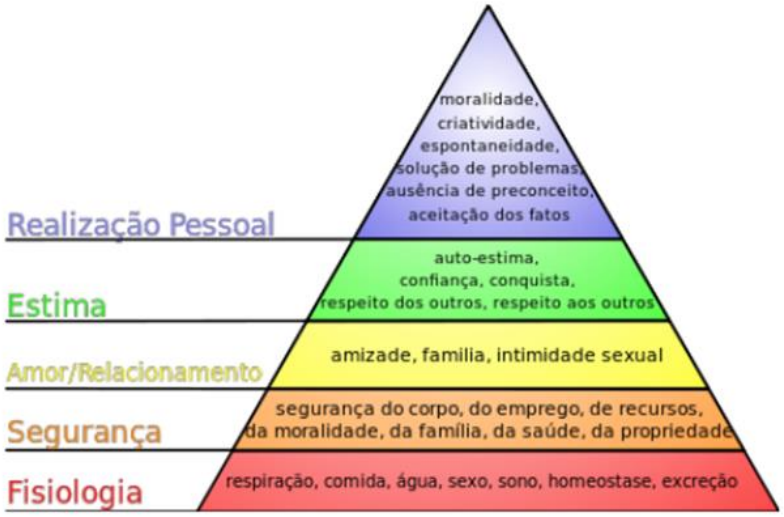


Fonte: autor, adaptado de GPS, p.8

Se tomarmos como ponto de partida o quadrante do Ponto Forte Rica diversidade e avançarmos no sentido horário da matriz encontraremos a relação com pouca área verde por habitante que indica a fragilidade. Aparece como fragilidade a Exigência de qualificação da mão de obra jovem e em seguida como oportunidade o fato de a região se tornar um Polo tecnológico. Note que as variáveis vão sendo encontradas e analisadas como proposta de solução para o problema do tema desse país.

De forma geral os temas ‘bem-estar’ e ‘qualidade de vida’ estão relacionados ao atendimento das necessidades de um indivíduo ou grupo social, cuja percepção se relaciona aos aspectos objetivos e subjetivos e dependem de uma estrutura que proporcione condições favoráveis para o seu desenvolvimento. Logo, desenvolver planos que promovam o bem-estar implica no reconhecimento das condições de cada localidade e na percepção das demandas da população e da formulação de políticas públicas, visando a disponibilização de serviços essenciais á manutenção da vida e ao bem-estar nas suas varias dimensões.

Na verdade um modelo integral do nível de ‘qualidade de vida’ precisa contar com parâmetros subjetivos, como os de bem-estar, felicidade, amor, prazer e realização pessoal; bem como parâmetros objetivos, cujas referências são a satisfação das necessidades básicas e das necessidades criadas pelo grau de desenvolvimento econômico e social de determinada sociedade. Isto aparece mais explicitamente na tradicional Pirâmide da Hierarquia das Necessidades de Maslow:



2.2 Os indivíduos

Os indivíduos desta análise são os 132 países analisados por uma série de indicadores de referência internacional. Os dados analisados de cada país são as variáveis que descrevemos a seguir.

	País	Código do País	Região	Todas as áreas
1	Albania	ALB	Southern Europe	EMDEU
2	Algeria	DZA	Northern Africa	NSWUA
3	Angola	AGO	Middle Africa	SUSAF
4	Argentina	ARG	IBE	AIBER
5	Armenia	ARM	Western Asia	CWEAS
6	Australia	AUS	Australia and New Zealand	AVECO
7	Austria	AUT	Western Europe	AVECO
8	Azerbaijan	AZE	Western Asia	CWEAS
9	Bangladesh	BGD	Southern Asia	ESEAS
10	Belarus	BLR	Eastern Europe	CWEAS
11	Belgium	BEL	Western Europe	AVECO
12	Benin	BEN	Western Africa	SUSAF
13	Bolivia	BOL	IBE	AIBER
14	Bosnia and Herzegovina	BIH	Southern Europe	EMDEU
15	Botswana	BWA	Southern Africa	SUSAF
16	Brazil	BRA	IBE	AIBER
17	Bulgaria	BGR	Eastern Europe	EMDEU
18	Burkina Faso	BFA	Western Africa	SUSAF
19	Burundi	BDI	Eastern Africa	SUSAF
20	Cambodia	KHM	South-Eastern Asia	ESEAS
21	Cameroon	CMR	Middle Africa	SUSAF
22	Canada	CAN	Northern America	AVECO
23	Central African Republic	CAF	Eastern Africa	SUSAF

24	Chad	TCD	Middle Africa	SUSAF
25	Chile	CHL	IBE	AIBER
26	China	CHN	Eastern Asia	ESEAS
27	Colombia	COL	IBE	AIBER
28	Congo, Republic of	COG	Middle Africa	SUSAF
29	Costa Rica	CRI	IBE	AIBER
30	Croatia	HRV	Southern Europe	EMDEU
31	Cuba	CUB	Caribbean	CARLA
32	Czech Republic	CZE	Eastern Europe	AVECO
33	Denmark	DNK	Northern Europe	AVECO
34	Djibouti	DJI	Eastern Africa	SUSAF
35	Dominican Republic	DOM	IBE	AIBER
36	Ecuador	ECU	IBE	AIBER
37	Egypt	EGY	Northern Africa	NSWUA
38	El Salvador	SLV	IBE	AIBER
39	Estonia	EST	Northern Europe	AVECO
40	Finland	FIN	Northern Europe	AVECO
41	France	FRA	Western Europe	AVECO
42	Georgia	GEO	Western Asia	CWEAS
43	Germany	DEU	Western Europe	AVECO
44	Ghana	GHA	Western Africa	SUSAF
45	Greece	GRC	Southern Europe	AVECO
46	Guatemala	GTM	IBE	AIBER
47	Guinea	GIN	Western Africa	SUSAF
48	Guyana	GUY	South America	CARLA
49	Honduras	HND	IBE	AIBER
50	Hungary	HUN	Eastern Europe	EMDEU
51	Iceland	ISL	Northern Europe	AVECO
52	India	IND	Southern Asia	ESEAS
53	Indonesia	IDN	South-Eastern Asia	ESEAS
54	Iran	IRN	Southern Asia	NSWUA
55	Iraq	IRQ	Western Asia	NSWUA
56	Ireland	IRL	Northern Europe	AVECO
57	Israel	ISR	Western Asia	AVECO
58	Italy	ITA	Southern Europe	AVECO
59	Jamaica	JAM	Caribbean	CARLA
60	Japan	JPN	Eastern Asia	AVECO
61	Jordan	JOR	Western Asia	NSWUA
62	Kazakhstan	KAZ	Cenrtral Asia	CWEAS
63	Kenya	KEN	Eastern Africa	SUSAF
64	Korea, Republic of	KOR	Eastern Asia	AVECO
65	Kuwait	KWT	Western Asia	NSWUA
66	Kyrgyzstan	KGZ	Cenrtral Asia	CWEAS
67	Laos	LAO	South-Eastern Asia	ESEAS
68	Latvia	LVA	Northern Europe	AVECO
69	Lebanon	LBN	Western Asia	NSWUA

70	Lesotho	LSO	Southern Africa	SUSAF
71	Liberia	LBR	Western Africa	SUSAF
72	Lithuania	LTU	Northern Europe	EMDEU
73	Macedonia	MKD	Southern Europe	EMDEU
74	Madagascar	MDG	Eastern Africa	SUSAF
75	Malawi	MWI	Southern Africa	SUSAF
76	Malaysia	MYS	South-Eastern Asia	ESEAS
77	Mali	MLI	Western Africa	SUSAF
78	Mauritania	MRT	Western Africa	NSWUA
79	Mauritius	MUS	Eastern Africa	SUSAF
80	Mexico	MEX	IBE	AIBER
81	Moldova	MDA	Eastern Europe	CWEAS
82	Mongolia	MNG	Eastern Asia	ESEAS
83	Montenegro	MNE	Southern Europe	EMDEU
84	Morocco	MAR	Northern Africa	NSWUA
85	Mozambique	MOZ	Eastern Africa	SUSAF
86	Namibia	NAM	Southern Africa	SUSAF
87	Nepal	NPL	Southern Asia	ESEAS
88	Netherlands	NLD	Western Europe	AVECO
89	New Zealand	NZL	Australia and New Zealand	AVECO
90	Nicaragua	NIC	IBE	AIBER
91	Niger	NER	Western Africa	SUSAF
92	Nigeria	NGA	Western Africa	SUSAF
93	Norway	NOR	Northern Europe	AVECO
94	Pakistan	PAK	Southern Asia	NSWUA
95	Panama	PAN	IBE	AIBER
96	Paraguay	PRY	IBE	AIBER
97	Peru	PER	IBE	AIBER
98	Philippines	PHL	South-Eastern Asia	ESEAS
99	Poland	POL	Eastern Europe	EMDEU
100	Portugal	PRT	IBE	AIBER
101	Romania	ROU	Eastern Europe	EMDEU
102	Russia	RUS	Eastern Europe	CWEAS
103	Rwanda	RWA	Eastern Africa	SUSAF
104	Saudi Arabia	SAU	Western Asia	NSWUA
105	Senegal	SEN	Western Africa	SUSAF
106	Serbia	SRB	Southern Europe	EMDEU
107	Slovakia	SVK	Eastern Europe	AVECO
108	Slovenia	SVN	Southern Europe	AVECO
109	South Africa	ZAF	Southern Africa	SUSAF
110	Spain	ESP	IBE	AIBER
111	Sri Lanka	LKA	Southern Asia	ESEAS
112	Sudan	SDN	Northern Africa	SUSAF
113	Swaziland	SWZ	Southern Africa	SUSAF
114	Sweden	SWE	Northern Europe	AVECO
115	Switzerland	CHE	Western Europe	AVECO

116	Tajikistan	TJK	Central Asia	CWEAS
117	Tanzania	TZA	Eastern Africa	SUSAF
118	Thailand	THA	South-Eastern Asia	ESEAS
119	Togo	TGO	Western Africa	SUSAF
120	Trinidad and Tobago	TTO	Caribbean	CARLA
121	Tunisia	TUN	Northern Africa	NSWUA
122	Turkey	TUR	Western Asia	EMDEU
123	Uganda	UGA	Eastern Africa	SUSAF
124	Ukraine	UKR	Eastern Europe	CWEAS
125	United Arab Emirates	ARE	Western Asia	NSWUA
126	United Kingdom	GBR	Northern Europe	AVECO
127	United States	USA	Northern America	AVECO
128	Uruguay	URY	IBE	AIBER
129	Uzbekistan	UZB	Central Asia	CWEAS
130	Venezuela	VEN	IBE	AIBER
131	Yemen	YEM	Western Asia	NSWUA
132	Zambia	ZMB	Middle Africa	SUSAF

2.3. As Variáveis do tema: bem-estar e qualidade de vida

Este tema da pesquisa apresenta trinta e três variáveis, sendo três categóricas e dezessete quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis-componentes ou indicadores, a saber:

- **Índices sintéticos:** São oito: (1) Índice de Progresso Social; (2) Índice de liberdade de imprensa; (3) Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD – ONU; (4) Índice de Governança; (5) Índice de Proteção Ambiental – EPI (*Environmental Protection Index*); (6) *Happy Planet Index* (HPI); (7) Índice de Gini para os países selecionados; e (8) Índice de Democracia 2014
- **Indicadores ou variáveis componentes:** São vinte e cinco: (1) Nutrição e cuidados médicos básicos; (2) Água e saneamento; (3) Segurança social; (4) Acesso à Educação Avançada; (5) Desnutrição (% da pop.); (6) Acesso a água canalizada (% da pop.); (7) Acesso a boas instalações sanitárias (% da pop.); (8) Acesso a eletricidade (% da pop.); (9) Qualidade do fornecimento de eletricidade (1 = baixa; 7 = Alto); (10) A expectativa de vida (anos); (11) Mortes por doenças não transmissíveis entre as idades de 30 e 70 (probabilidade de morrer); (12) Taxa de obesidade (% da pop.); (13) Taxa de suicídio (mortes / 100.000); (14) Escravidão moderna, tráfico de seres humanos e casamento infantil (1 = baixa; 100 = alto); (15) Demanda de satisfação com métodos contraceptivos (% das mulheres); (16) As mulheres tratadas com respeito (0 = baixo; 100 = alta); (17) Tolerância aos imigrantes (0 = baixa; 100 = alto); (18) Tolerância para homossexuais (0 = baixa; 100 = alto); (19) Discriminação e violência contra as minorias (0 = baixo, 10 = alto); (20) A tolerância religiosa (1 = baixo; 4 = alto); (21) Rede de segurança comunitária (0 = baixa; 100 = alta); (22) Anos de escolaridade terciária; (23) Média de anos das mulheres na escola; (24) A desigualdade na obtenção de educação (0 = baixo, 1 = alto); e (25) bem-estar.

Tabela 1. As 33 Variáveis do Eixo temático: bem-estar e qualidade de vida

Variável	Significado	Tipo	Unidade de Medida
Índice de Progresso Social	Índice que oferece um modelo de medição do Progresso Social. A performance das nações é baseada em 52 indicadores nas áreas de necessidades humanas básicas, bem-estar e oportunidades ao crescimento. É calculado pela associação sem fins lucrativos Social Progress Imperative	Variável Quantitativa	Percentual
Nutrição e cuidados médicos básicos	Refere-se a como está em percentual a nutrição e o acesso a cuidados médicos básicos no país observado.	Variável Quantitativa	Percentual
Água e Saneamento	Inclui a população com acesso a água e instalações sanitárias	Variável Quantitativa	Percentual
Segurança social	Refere-se a como se encontra a segurança dos indivíduos na sociedade em que vivem	Variável Quantitativa	Percentual
Acesso à Educação Avançada	Refere-se ao acesso a educação mais avançada.	Variável Quantitativa	Percentual
Desnutrição (% da pop.)	A percentagem da população que possuem uma dieta com consumo insuficiente de alimentos para manter uma energia constante. Os dados que mostram que 5% significa uma prevalência de desnutrição abaixo de 5%.	Variável Quantitativa	Percentual
Acesso a água canalizada (% da pop.)	A percentagem da população com um cano de água conectada com canalização interna para um ou mais torneiras ou uma conexão de água canalizada a uma torneira colocada no quintal ou terreno fora da casa.	Variável Quantitativa	Percentual
Acesso a boas instalações sanitárias (% da pop.)	A percentagem da população com saneamento melhorado, incluindo autoclismos, sistemas de tubulações de esgoto, fossas sépticas, lavar / derramar rente à latrina, ventilado melhoradas latrinas (VIP), latrina com laje, e banheiros de compostagem.	Variável Quantitativa	Percentual
Acesso a eletricidade (% da pop.)	Porcentagem da população com acesso a eletricidade	Variável Quantitativa	Percentual
Qualidade do fornecimento de eletricidade (1 = baixa; 7 = Alto)	Avaliação de resposta para a pergunta: "No seu país, como você avaliaria a confiabilidade do fornecimento de eletricidade (falta de interrupções e da falta de flutuações de tensão)?" [1 = não confiável a todos; 7 = extremamente confiável]	Variável Quantitativa	Indicador nominal que varia de 1 a 7
Índice de liberdade de imprensa (0 = mais livre; 100 = menos livre)	O grau de liberdade que os jornalistas, organizações de notícias, e internautas desfrutar em cada país, e os esforços envidados pelas autoridades para respeitar e assegurar o respeito por esta liberdade. No modelo SPI, as pontuações são agrupados em 7 bandas, de 1 (menos livre) a 7 (mais livre).	Variável Quantitativa	Indicador nominal que varia de 1 a 7
A expectativa de	O número de anos que um recém-nascido viveria se	Variável	Percentual

vida (anos)	os padrões de mortalidade permanecessem os mesmos do momento em que ele nasceu durante toda a sua vida.	Quantitativa	
Mortes por doenças não transmissíveis entre as idades de 30 e 70 (probabilidade de morrer)	A probabilidade de morrer entre as idades de 30 e 70 anos por doenças cardiovasculares, câncer, diabetes ou doença respiratória crônica.	Variável Quantitativa	Percentual
Taxa de obesidade (% da pop.)	A percentagem da população com um índice de massa corporal (IMC) de 30 kg / m2 ou superior (estimativa idade-padronizados), ambos os sexos.	Variável Quantitativa	Percentual
Taxa de suicídio (mortes / 100.000)	A mortalidade devido a lesão auto infligida, por 100.000 pessoas, a idade ajustada.	Variável Quantitativa	Percentual
Escravidão moderna, tráfico de seres humanos e casamento infantil (1 = baixa; 100 = alto)	A medida combinada de três variáveis: prevalência estimada de escravidão moderna em cada país (que representam 95% do total); uma medida do nível de tráfico de seres humanos para e de cada país (contas para 2,5%); e uma medida do nível da criança e do casamento precoce em cada país (contas para 2,5%), escala de 1 (baixo escravidão) a 100 (alta escravidão).	Variável Quantitativa	Percentual
Demanda de satisfação com métodos contraceptivos (% das mulheres)	A taxa de prevalência contraceptiva para mulheres com idade entre 15-49 anos que são casados ou em união através da soma da taxa de prevalência de contraceptivos e necessidade não atendida de contracepção.	Variável Quantitativa	Percentual
As mulheres tratadas com respeito (0 = baixo; 100 = alta)	A percentagem de respondentes do sexo feminino que respondem sim à pergunta, "Você acredita que as mulheres neste país são tratadas com respeito e dignidade?"	Variável Quantitativa	Percentual
Tolerância aos imigrantes (0 = baixa; 100 = alto)	A percentagem de inquiridos que respondem sim à pergunta: "Será que a cidade ou área onde você vive é um bom lugar para um imigrantes de outros países viver?"	Variável Quantitativa	Percentual
Tolerância para homossexuais (0 = baixa; 100 = alto)	O percentual de entrevistados que respondem sim à pergunta: "Será que a cidade ou área onde você vive é um bom lugar para pessoas gay ou lésbica viverem?"	Variável Quantitativa	Percentual
Discriminação e violência contra as minorias (0 = baixo, 10 = alto)	A discriminação, a impotência, a violência étnica, violência comunal, a violência sectária e violência religiosa, medido em uma escala de 0 (baixa pressão) a 10 (pressões muito altas) .	Variável Quantitativa	Percentual
A tolerância religiosa (1 =	Uma medida de 13 tipos de hostilidade religiosa por particulares, organizações ou grupos da sociedade,	Variável Quantitativa	Indicador nominal

baixo; 4 = alto)	incluindo conflitos relacionados com religião armados ou terrorismo, a multidão ou a violência sectária, o assédio sobre o traje por razões religiosas ou outras intimação ou abuso relacionados com a religião. No modelo SPI, as pontuações variam de 1 (baixo) a 4 (muito alto).		que varia de 0 a 4
Rede de segurança comunitária (0 = baixa; 100 = alta)	O percentual de entrevistados que respondem sim à pergunta: "Você tem parentes ou amigos com quem contar para ajudá-lo sempre que se encontrar em uma situação de apuros?	Variável Quantitativa	Percentual
Anos de escolaridade terciária	A média de anos de educação superior concluído entre as pessoas com idade acima de 25.	Variável Quantitativa	Percentual
Média de anos das mulheres na escola	O número médio de anos de escola frequentada por mulheres entre 25 e 34 anos de idade, incluindo primário, secundário e ensino superior.	Variável Quantitativa	Percentual
A desigualdade na obtenção de educação (0 = baixo, 1 = alto)	A perda na educação potencial devido à desigualdade, calculada como a diferença percentual entre o Índice de Desenvolvimento Humano Índice de Educação, que compreende média de anos de escolaridade e anos de escolaridade esperados, e o Índice de Educação Ajustado à Desigualdade.	Variável Quantitativa	Indicador nominal que varia de 0 a 1
IDH - Índice de Desenvolvimento Humano – 2013	Mede o progresso de um país por indicadores de qualidade de vida: renda, saúde e educação são os principais.	Variável Quantitativa	Indicador nominal que varia de 0 a 1
Bem-estar	Medido por três componentes: (1) expectativa de vida; (2) experiência de bem-estar; e (3) pegada ecológica.	Variável Quantitativa	Percentual
Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as politicas são implementadas e qual a capacidade de formulá-las.	Variável Quantitativa	Percentual
EPI -Índice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	Variável Quantitativa	Percentual
HPI – Índice de Felicidade do Planeta	Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	Variável Quantitativa	Percentual
Índice de Ginni	Mede a distribuição de renda e de despesas das famílias (estimativa do banco mundial)	Variável Quantitativa	Percentual
Índice de Democracia 2014	Índice de Democracia da Economist Intelligence Unit do fornece um instantâneo do estado da democracia em todo o mundo para 165 estados independentes e dois territórios-isto cobre quase toda a população do mundo e a grande maioria dos Estados do mundo (micro estados são excluídos). O Índice de Democracia é baseado em cinco categorias: processo eleitoral e pluralismo; liberdades civis; o funcionamento do governo; participação política; e cultura política. Com base	Variável Quantitativa	Percentual

	nas suas pontuações em uma série de indicadores dentro dessas categorias, cada país é então classificado como um dos quatro tipos de regime: "democracias plenas"; "democracias imperfeitas"; "regimes híbridos"; e "regimes autoritários".		

Fonte: autor a partir dos dados da planilha estatística e do GPS p.16

3. ANÁLISE EEXPLORATÓRIA DOS DADOS

Depois de apresentar o conjunto de 33 (trinta e três) variáveis relacionadas ao tema bem-estar e qualidade de vidas, com o objetivo de exlicar a sua relação com o universo dos 132 países também já apresentados, se analisará cada uma das variáveis separadamente quanto a sua forma de distribuição, os valores atípicos, medidas de centro e dispersão, através do *Display Descriptive Statistics* que está no *link* principal *Basic Statistics* do *software* Minitab, contando-se com o auxílio de gráficos (*pie chart*, barras, histogramas, dendogramas, curvas de densidade) e tabelas de medidas numéricas (média, mediana, quartis, desvio-padrão, variância, intervalo de confiança e teste de normalidade de Anderson-Darling).

Segue-se com a aplicação do *link* principal *Data*, acessando o *Code – Numeric to Numeric*, sendo que este recurso permite que os dados ausentes que aparecem na tabela com o símbolo asterisco (*) seja substituído pelo valor determinado que é a média encontrada.

Em seguida, se apresentará os gráficos de dispersão e se procederá na sequência à normalização dos mesmos que é um ajuste com vista a afastar os *outliers*, ou seja, os pontos fora da curva – é como se dos dados analisados não possuíssem uma distribuição normal então se faz um ajuste para que todos fiquem dentro de uma certa normalidade.

Apresentar-se-ão os valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas. Para essa etapa será usado o comando do MINITAB denominado *Calculator* e será aplicada a fórmula (a seguir) indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.

Buscar-se-á comparar as análises efetuadas através da correlação e posteriormente apresentar-se-á dendogramas.

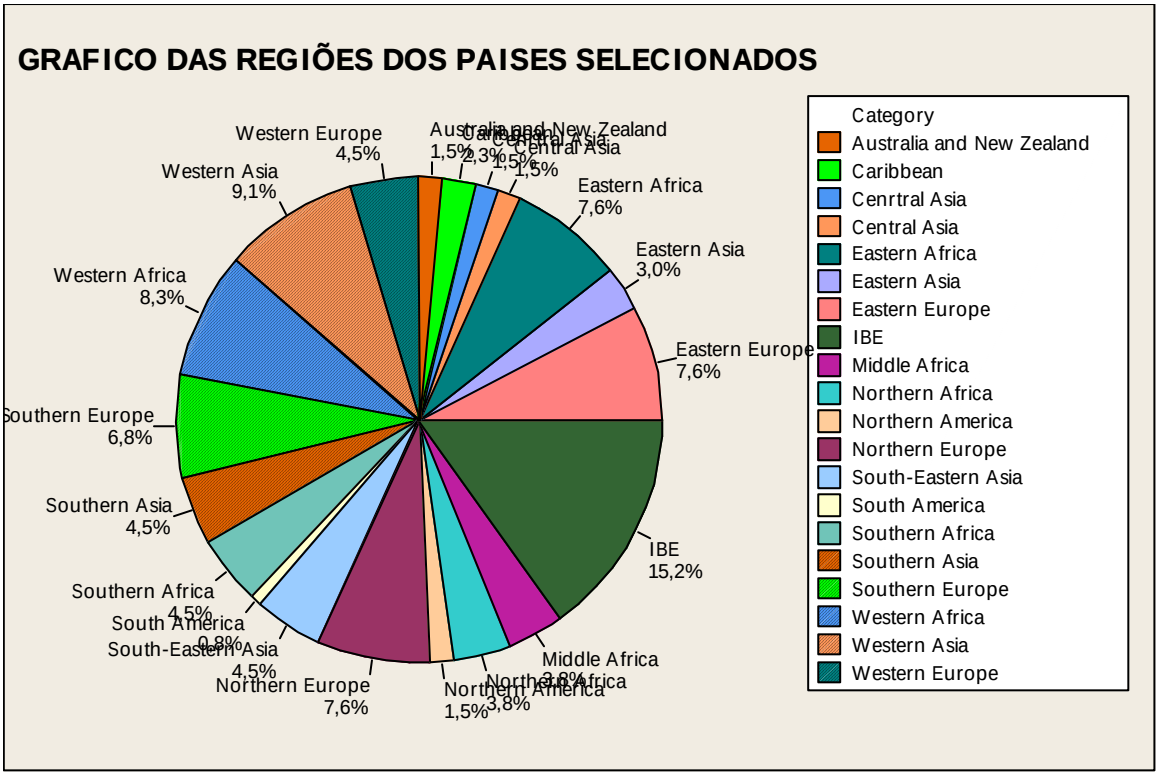
Nesta primeira parte, se apresentará o conjunto dos gráficos gerados pelo recurso *Graphical Summary* que está no item *Basic Statistic*, que por sua vez está no item *Static* na barra de ferramentas do MINITAB.

3.1 Seguegação por região

A seguir será apresntado o gráfico circular ou gráfico de pizza ou gráfico de torta elaborado pelo MINITAB a partir de um agrupamento dos 132 países em 19 regiões dentro dos

continentes: Europeu, Africano, Oceania, Asiático e Americano. O Gráfico 1 apresenta a distribuição dos países nas regiões mencionadas

Gráfico 1: Região dos Países Selecionados



3.2 Análise Descritivados dados

Média, Mínimo e Máximo das variáveis com uso do MINITAB

Para facilitar o acompanhamento das análises vamos expor os dados do MINITAB citando com títulos que representem a situação de estudo.

- Valor de N é o total de elementos da amostra que nota caso ela é total: 132 países
- **N*** indica o total de dados ausentes na tabela em cada variável. Por exemplo, a variável Depth of food deficit (calories/undernourished person)(déficit alimentar ou calorias consumidas) mostra que não tem uma informação que o Iraque.
- Mean (média) mostra o valor médio da variável para os 132 países.
- SE Mean mostra o erro padrão da amostra, isto é, ele mostra o quanto a média varia em relação as outra da amostra.
- StDev é o desvio padrão mostra a distância estimativa entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação a média.
- Minimum mostra o valor mínimo dos dados dos 132 países
- Maximum mostra o valor máximo dos dados dos 132 países

Tabela com dados originais da base de dados do tema: **recursos básicos: água, alimento, energia (basic features: water, food, energy)**

Descriptive Statistics: Social Progr; Nutrition an; Water and Sa; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93
Nutrition and Basic Medi	132	0	84,12	1,38	15,90	27,58	72,20
Water and Sanitation	132	0	73,11	2,30	26,46	4,75	51,94
Personal Safety	132	0	61,94	1,55	17,86	21,52	48,42
Access to Advanced Educa	132	0	39,78	1,75	20,05	4,72	22,56
Undernourishment (% of p	131	1	10,82	1,19	13,57	0,50	0,50
Access to piped water (%	129	3	63,98	3,03	34,36	2,53	24,71
Access to improved sanit	130	2	73,75	2,55	29,09	9,58	52,97
Access to electricity (%	132	0	78,33	2,80	32,23	4,00	58,00
Quality of electricity s	121	11	4,515	0,142	1,567	1,273	3,223
Press Freedom Index (0=m	132	0	30,91	1,31	15,04	6,38	22,95
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000
Obesity rate (% of pop,)	132	0	17,373	0,822	9,444	1,100	7,225
Suicide rate (deaths/100	132	0	10,225	0,610	7,009	0,475	4,391
Modern slavery, human tr	132	0	10,441	0,931	10,691	1,000	4,010
Satisfied demand for con	131	1	74,58	1,68	19,18	20,30	66,90
Women treated with respe	129	3	61,57	1,67	19,01	17,00	47,50
Tolerance for immigrants	132	0	57,82	1,42	16,32	20,00	46,00
Tolerance for homosexual	125	7	28,62	2,13	23,85	1,00	8,50
Discrimination and viole	132	0	5,990	0,176	2,019	1,000	4,450
Religious tolerance (1=l	132	0	2,7803	0,0856	0,9832	1,0000	2,0000
Community safety net (0=	132	0	79,42	1,10	12,63	30,00	73,00
Years of tertiary school	117	15	0,4475	0,0344	0,3717	0,0100	0,1400
Womens aerage years in s	131	1	9,471	0,327	3,743	1,300	6,600
Inequality in the attain	117	15	0,1960	0,0136	0,1474	0,0130	0,0555
IDH - 2013	125	7	0,6894	0,0144	0,1611	0,3370	0,5595
Well-being (0-10)	128	4	5,439	0,105	1,186	2,808	4,492
1, Governança	132	0	53,38	1,88	21,59	15,22	37,11
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780
Democracy Index 2014	132	0	5,873	0,180	2,068	1,490	4,100

Variable	Median	Q3	Maximum
Social Progress Index	63,86	73,74	88,24
Nutrition and Basic Medi	91,25	97,00	98,78
Water and Sanitation	81,71	95,87	100,00
Personal Safety	61,41	75,67	93,45
Access to Advanced Educa	41,74	52,48	89,37
Undernourishment (% of p	0,50	19,80	73,40
Access to piped water (%	75,33	95,34	100,00
Access to improved sanit	88,00	98,65	100,00
Access to electricity (%	99,00	100,00	100,00
Quality of electricity s	4,848	5,660	6,754
Press Freedom Index (0=m	28,45	37,72	73,40
Life expectancy (years)	73,398	76,905	82,695
Non-communicable disease	23,000	28,000	38,000
Obesity rate (% of pop,)	19,000	24,275	42,800
Suicide rate (deaths/100	9,760	13,356	32,700
Modern slavery, human tr	8,466	15,249	97,897
Satisfied demand for con	82,20	88,30	96,20
Women treated with respe	62,00	77,50	99,00
Tolerance for immigrants	56,00	71,75	93,00
Tolerance for homosexual	20,00	47,00	85,00
Discrimination and viole	6,000	7,683	10,025
Religious tolerance (1=l	3,0000	4,0000	4,0000
Community safety net (0=	82,00	89,00	98,00
Years of tertiary school	0,3500	0,6250	1,7100
Womens aerage years in s	10,400	12,500	15,000
Inequality in the attain	0,1780	0,3095	0,4980
IDH - 2013	0,7220	0,8200	0,9440
Well-being (0-10)	5,264	6,338	7,771
1, Governança	48,46	70,36	100,00

EPI Score	52,09	66,40	87,67
Happy Planet Index	42,463	49,286	64,036
GINNI Index	38,160	44,550	63,140
Democracy Index 2014	6,045	7,563	9,930

Análise: Selecionamos o Indice de Progresso Social para esse comentário que mostra a Média de 63,67 em uma escala até 100, indicando que de uma maneira geral os países estão dentro da escala normal. Embora o valor mínimo observado é de 32,60 cujo país desse número é o Chad no centro da África com 12,300 milhões de pessoas e tem seus vizinhos: Líbia, Egito, Nigeria e Sudão. E o valor máximo observado foi 88,24 para o país Nova Zelândia que é um país formado por um conjunto de ilhas sendo as maiores denominadas: Ilha do Norte e Ilha do Sul. Por ter uma característica de economia desenvolvida esse indicador confirma essa qualidade.

Usando o link principal Data, acessando o Code – Numeric to Numeric para substituir os dados ausentes que aparecem na tabela com o símbolo asterisco (*) pela média encontrada.
Foram substituídos:
Acesso a água canalizada (Access to piped water) foi incluída a média 63,98 em dois países: Australia e estados Unidos incluímos a média 82,85 no Iraque
Acesso a boas instalações sanitárias (Access to improved sanitation facilities) foi incluída a média 73,75 em dois países: Nova Zelândia e Itália
Qualidade do fornecimento de eletricidade (Quality of electricity supply) foi incluída a média 4,515 em dez países: República do Congo, República da África Central, Cuba, Iraque, Djibouti, Nigéria, Sudão, Tajiquistão, Togo e Uzbequistão.
Demanda de satisfação com métodos contraceptivos (Satisfied demand for contraception) foi incluída a média 74,58 em um país que é a Islândia.
As mulheres tratadas com respeito (Women treated with respect) foi incluída a média 61,57 em três países: Cuba, Namibia e Guiana.
Tolerância para homossexuais (Tolerance for homosexuals) foi incluída a média 28,62 em sete países: Egito, Iran, Kuwait, Jordânia, Arábia Saudita, Emirados Arabes e Iémen. Nota-se que a média para tolerância a homossexualidade é bastante baixa e que os países que não demostraram os índices são situados na África e na Ásia, muitos deles com uma cultura muçulmana, bastante machista.
Anos de escolaridade terciária (Years of tertiary schooling) foi incluída a média 0,445 em quinze países, sendo na África: Angola, Burkina Faso, Chad, Djibouti, Guiné, Lesoto, Madagascar e Nigéria; na Ásia: Azerbaijão, Georgia e Uzbekistan; na Europa: Belarus, Bósnia, Macedônia e Montenegro. Nota-se que a média de pessoas com educação universitária concluída acima de 25 anos de idade é muito baixa.
Média de anos das mulheres na escola (Women's average years in school) foi incluída a média 9,471 em um país que é a Islândia. A média dos países é muito baixa.
A desigualdade na obtenção de educação (Inequality in the attainment of education) foi incluída a média 0,196 em catorze países: Argélia, Bósnia, Burundi, Cuba, Iran, Kuwait, Malásia, Nova Zelândia, Emirados Árabes Unidos, Arábia Saudita, Tunísia, Sudão, Rússia e Paraguai.

IDH - Índice de Desenvolvimento Humano foi incluída a média 0,689 em sete países: Argélia, Cazaquistão, República Coreana, Letônia, Quirguistão, Mauritius e Países Baixos.
Bem-estar (Well-being) foi incluída a média 5, 439 em quatro países: Guiné, Lesoto, Montenegro.
HPI – Índice de Felicidade do Planeta (Happy Planet Index) foi incluída a média 42,673 em dois países: Lesoto, Montenegro e Swaziland.
Índice de Ginni foi incluída a média 39,024 em treze países: Azerbaijão, Cuba, Guiana, Islândia, Kuwait, Mauritius, Montenegro, Nepal, Paquistão, Arábia Saudita, Sri Lanka e Emirados Árabes Unidos.
Tabela Confirmando a eliminação dos dados ausentes depois que foram substituídos pela Média de cada uma das variáveis.

Descriptive Statistics: Democracy In; GINNI Index; Happy Planet; EPI Score; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Democracy Index 2014	132	0	5,873	0,180	2,068	1,490	4,100
GINNI Index	132	0	39,024	0,705	8,097	24,820	33,065
Happy Planet Index	132	0	42,673	0,778	8,936	22,591	36,587
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30
1, Governança	132	0	53,38	1,88	21,59	15,22	37,11
Well-being (0-10)	132	0	5,438	0,102	1,168	2,808	4,591
IDH - 2013	132	0	0,6889	0,0136	0,1567	0,3370	0,5700
Inequality in the attain	132	0	0,1953	0,0121	0,1387	0,0130	0,0615
Womens aerage years in s	132	0	9,471	0,325	3,728	1,300	6,625
Years of tertiary school	132	0	0,4467	0,0304	0,3498	0,0100	0,1525
Community safety net (0=	132	0	79,42	1,10	12,63	30,00	73,00
Religious tolerance (1=l	132	0	2,7803	0,0856	0,9832	1,0000	2,0000
Discrimination and viole	132	0	5,990	0,176	2,019	1,000	4,450
Tolerance for homosexual	132	0	28,62	2,02	23,20	1,00	9,00
Tolerance for immigrants	132	0	57,82	1,42	16,32	20,00	46,00
Women treated with respe	132	0	61,57	1,64	18,79	17,00	48,00
Satisfied demand for con	132	0	74,58	1,66	19,11	20,30	66,95
Modern slavery, human tr	132	0	10,441	0,931	10,691	1,000	4,010
Suicide rate (deaths/100	132	0	10,225	0,610	7,009	0,475	4,391
Obesity rate (% of pop,)	132	0	17,373	0,822	9,444	1,100	7,225
Non-communicable disease	132	0	22,098	0,635	7,293	10,000	15,000
Life expectancy (years)	132	0	70,355	0,817	9,392	46,669	64,319
Press Freedom Index (0=m	132	0	30,91	1,31	15,04	6,38	22,95
Quality of electricity s	132	0	4,513	0,131	1,500	1,273	3,305
Access to electricity (%	132	0	78,33	2,80	32,23	4,00	58,00
Access to improved sanit	132	0	73,75	2,51	28,87	9,58	53,00
Access to piped water (%	132	0	63,98	2,96	33,97	2,53	26,30
Undernourishment (% of p	132	0	10,82	1,18	13,51	0,50	0,50
Access to Advanced Educa	132	0	39,78	1,75	20,05	4,72	22,56
Personal Safety	132	0	61,94	1,55	17,86	21,52	48,42
Water and Sanitation	132	0	73,11	2,30	26,46	4,75	51,94
Nutrition and Basic Medi	132	0	84,12	1,38	15,90	27,58	72,20
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93

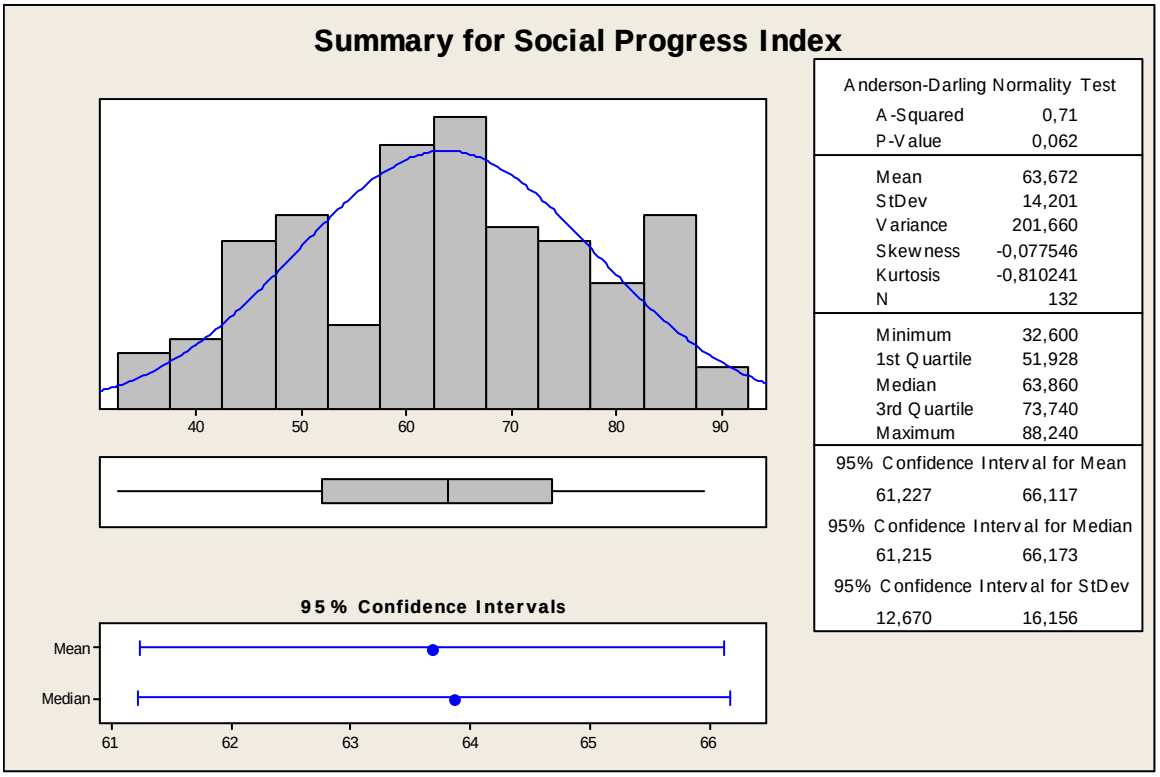
Variable	Median	Q3	Maximum
Democracy Index 2014	6,045	7,563	9,930
GINNI Index	39,020	43,553	63,140
Happy Planet Index	42,627	49,175	64,036
EPI Score	52,09	66,40	87,67
1, Governança	48,46	70,36	100,00
Well-being (0-10)	5,343	6,271	7,771
IDH - 2013	0,7155	0,8147	0,9440
Inequality in the attain	0,1900	0,2970	0,4980
Womens aerage years in s	10,400	12,500	15,000
Years of tertiary school	0,4250	0,6075	1,7100
Community safety net (0=	82,00	89,00	98,00
Religious tolerance (1=l	3,0000	4,0000	4,0000

Discrimination and viole	6,000	7,683	10,025
Tolerance for homosexual	21,50	45,75	85,00
Tolerance for immigrants	56,00	71,75	93,00
Women treated with respe	62,00	77,00	99,00
Satisfied demand for con	82,15	88,30	96,20
Modern slavery, human tr	8,466	15,249	97,897
Suicide rate (deaths/100	9,760	13,356	32,700
Obesity rate (% of pop,)	19,000	24,275	42,800
Non-communicable disease	23,000	28,000	38,000
Life expectancy (years)	73,398	76,905	82,695
Press Freedom Index (0=m	28,45	37,72	73,40
Quality of electricity s	4,707	5,628	6,754
Access to electricity (%	99,00	100,00	100,00
Access to improved sanit	86,37	98,54	100,00
Access to piped water (%	74,80	95,03	100,00
Undernourishment (% of p	2,80	19,43	73,40
Access to Advanced Educa	41,74	52,48	89,37
Personal Safety	61,41	75,67	93,45
Water and Sanitation	81,71	95,87	100,00
Nutrition and Basic Medi	91,25	97,00	98,78
Social Progress Index	63,86	73,74	88,24

Análise: Após a substituição dos dados ausentes identificados por asterisco (*) pela média de cada variável foi rodado novamente os indicadores de número de elementos (N), dados ausentes (N*), média, erro da média (SE mean), desvio padrão (Stdev), valor mínimo e máximo; apenas para confirmar se foram lançados corretamente. Portanto, sem alteração nos respectivos valores.

3.2 Análise exploratória das variáveis

4.1 SPI – Social Progress Index (Índice de Progresso Social)



Variável: Social Progress Index – índice de Progresso Social

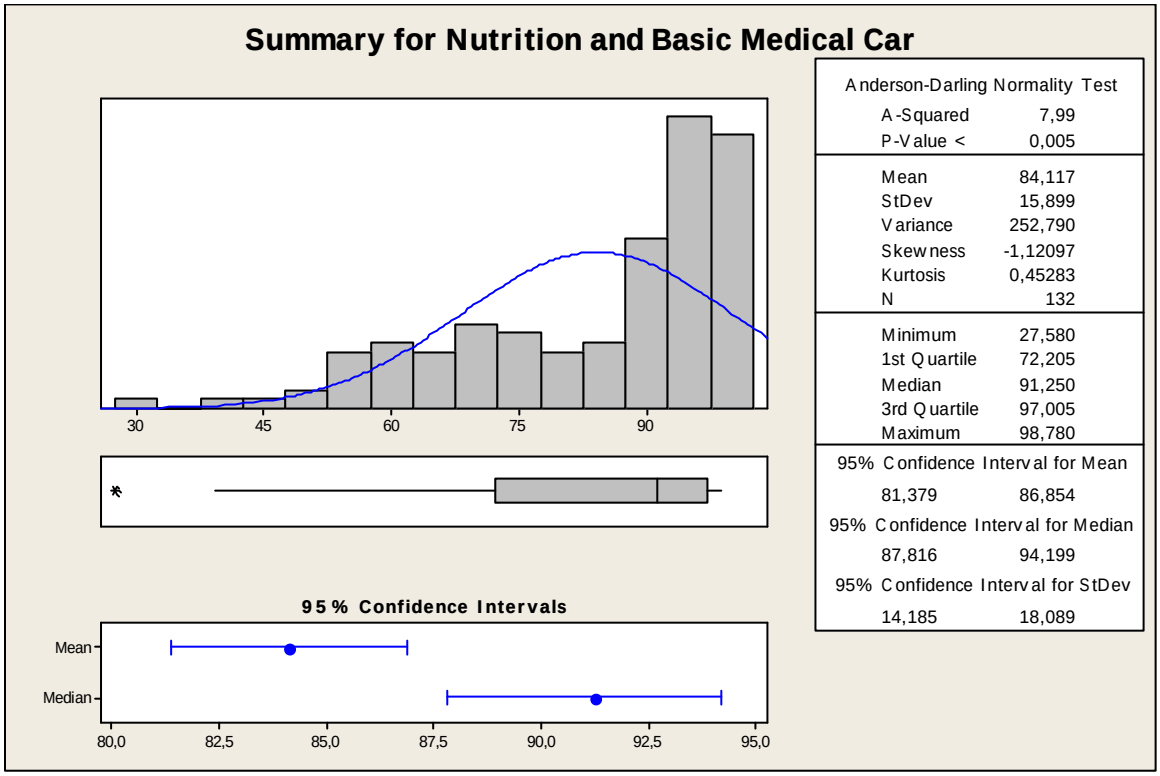
Origem e definição dado pelo portal: http://pt.knoema.com/jklbfre/the-social-progress-index

O Índice de Progresso Social é o resultado de um processo de dois anos de pesquisas envolvendo uma equipe renomada de especialistas que inclui, entre outros, os economistas Hernando de Soto e Michael Porter e do presidente da Fundação Rockefeller Dr. Judith Rodin. O índice sintetiza um total extenso de dados pesquisados para identificar as dimensões do desempenho das sociedades e medir o progresso social de forma abrangente e rigorosa. O Índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidadas em três dimensões do Progresso Social: necessidades humanas básicas, Fundações de Bem-estar e oportunidade. A primeira dimensão, necessidades humanas básicas, avalia o quão bem um país prevê necessidades essenciais de seu povo através da medição se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se tiverem acesso a água potável, se tiverem acesso a uma habitação condigna com serviços básicos, e se eles são seguros e protegidos.

Análise dos dados da nossa pesquisa:

O gráfico mostra que a distribuição aproxima-de da normal, considerando os 132 países. Percebe-se que no intervalo de 62,5 a 67,5 que compreende a média de 63,6 e a mediana de 63,8 possui 21 observaçõe. Os dois intervalo anteriores 57,5 a 62,5 e 52,5 a 57,5 temos 19 e 6 observações, nos dois posterior de 67,5 a 72,5 e 72,5 a 77,5 temos 13 e 12 observações, sendo assim se somarmos os 5 intervalos teremos uma concentração de 71 países próximos da mediana. Observa-se uma concentração de 84 países no segundo e terceiro quartil, 21 no primeiro e 24 no quarto quartil. A mediana de 63,8 e o terceiro quartil de 73,7 indicam o poder discricionário do índice, que consegue separar o grupo de países em 12 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados. O brasil encomtra-se com 69,97 acima da média e da mediana.

4.2 – Nutrition and Basic Medical Care (Nutrição e cuidados médicos básicos)



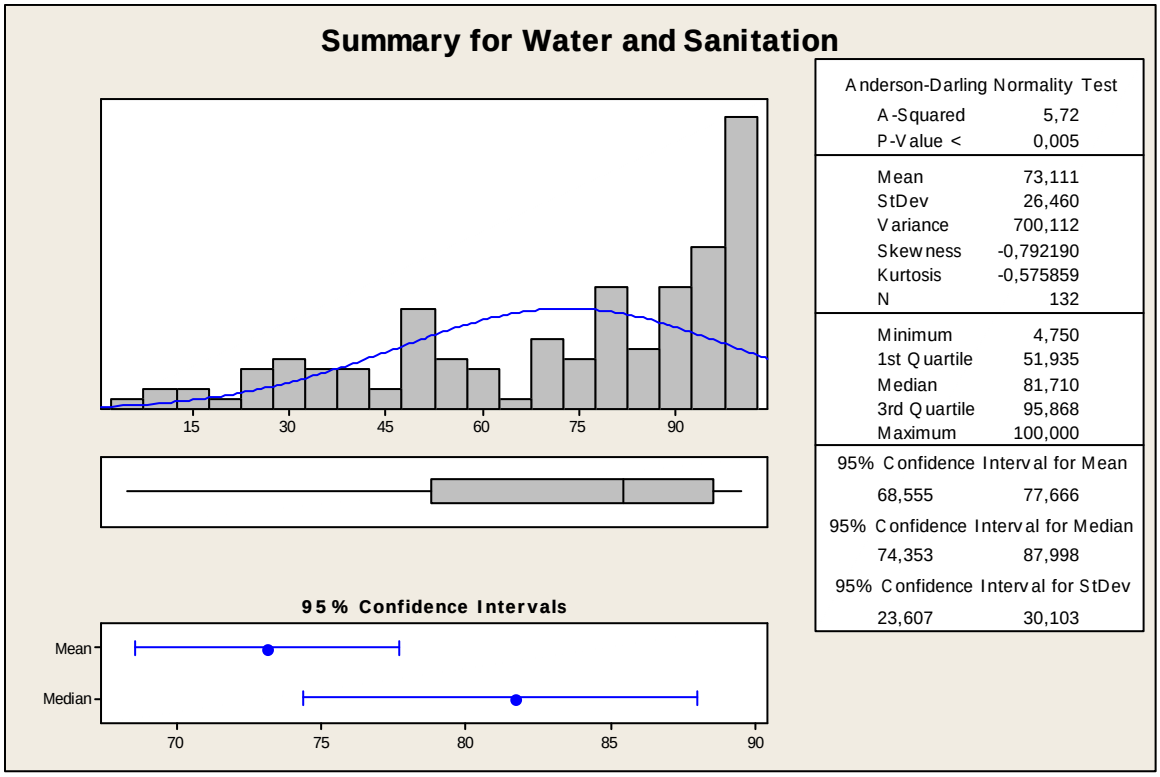
Variável: Nutrition and Basic Medical Care

.

<u>Análise dos dados da nossa pesquisa</u>

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a direita o que indica que os dados estão concentrados no quarto quartil, nos três últimos intervalos entre 87,5 e 102,5 com 78 observações, que se encontram acima da média de 84,1 e próximo ou acima da mediada de 91,2. Já a menor concentração encontra-se no primeiro quartil nos quatro primeiros intervalos que vão de 27,5 a 47,5 com apenas 3 observações, as três encotram-se no continente Africano, são elas Burundi, Chad e Central African Republic. O Brasil encontra-se perto da mediana com 92.

4.3 – Water and Sanitation (Água e Saneamento)

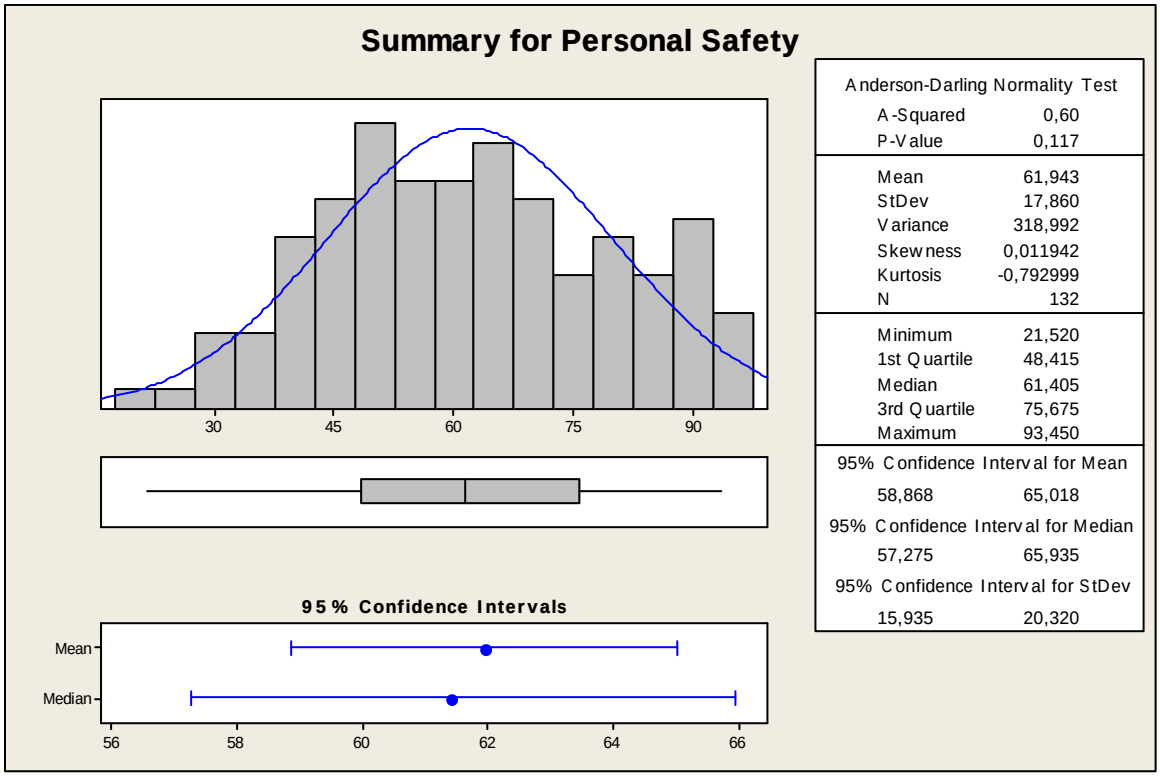


Variável: Water and Sanitation

<u>Análise dos dados da nossa pesquisa:</u>

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a direita o que indica que os dados estão concentrados no quarto quartil, nos cinco últimos intervalos entre 77,5 e 102,5 com 75 observações, que se encontram acima da média de 73,1 e próximo ou acima da mediada de 81,7. Já a menor concentração encontra-se no primeiro quartil nos quatro primeiros intervalos que vão de 2,5 a 27,5 com apenas 10 observações, as três encotram-se no continente Africano, são eles Nigéria (4,7), Congo (10), Togo (11,3), Madagascar (16), Mozambique (17,3), Tanzania (22) Central African Republic (23,6), Mali (25,2), Chad (26,1) e Liberia (26,2). O Brasil se encontra próximo da mediana com 81,6.

4.4 –Personal Safety (Segurança pessoal)

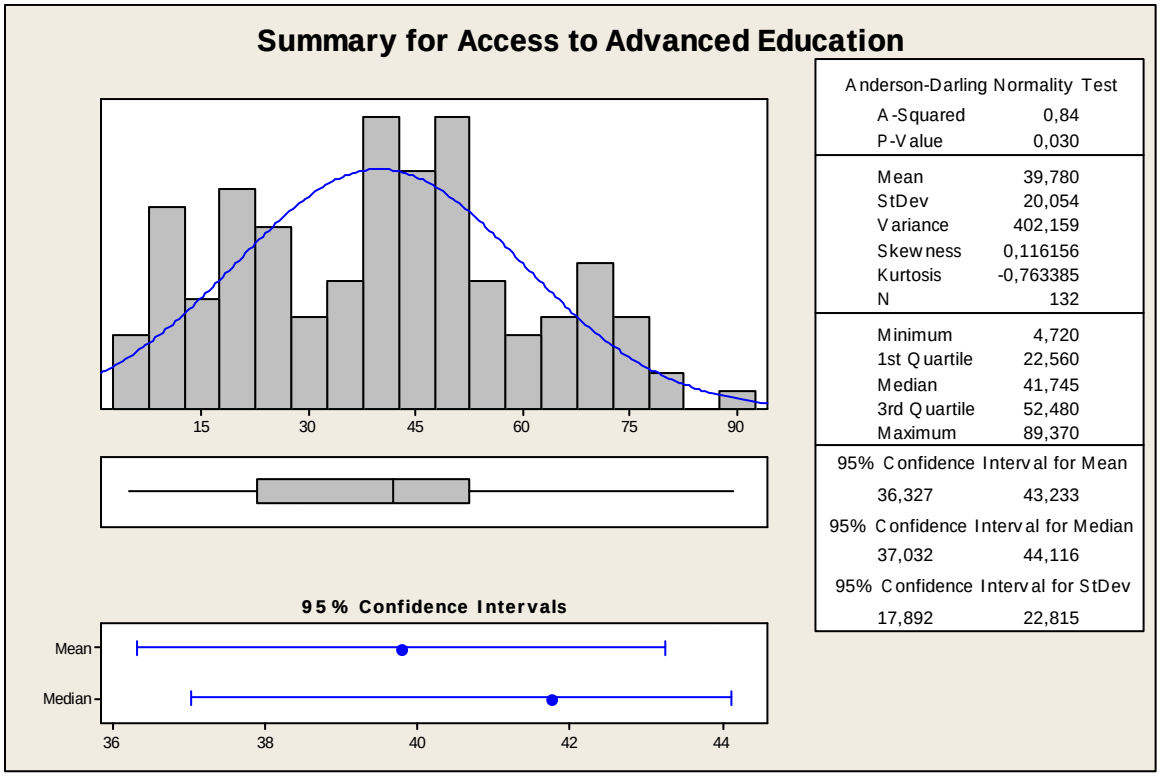


Variável: Personal Safety

<u>Análise dos dados da nossa pesquisa</u>:

O gráfico mostra que a distribuição aproxima-de da normal, considerando os 132 países. Percebe-se que a maior observação de 47 encontra-se no segundo quartil com intervalo de 37,5 a 57,5. No entanto tanto a média de 61,9, quanto a mediana de 61,4 encontram-se no terceiro quartil com intervalo de 57,5 a 77,5 com 44 observações. No quarto quartil nos intervalos de 77,5 para 97,5 encontram-se 31 observações e no primeiro quartil de 17,5 a 37,5 nota-se 10 observações nas quais encontram-se 5 países do continente Africano: Chad, Central African Republic, Sudan, South African e Nigéria; 4 do continente Ibero Americano: Mexico, Velezuela, Honduras, Dominican Republic, e 1 do continente Asiático:Iraq. O Brasil se encontra no segundo quartil com 37,5.

4.5 –Access to Advanced Education (Acesso a Educação Avançada)



Variável: Access to Advanced Education

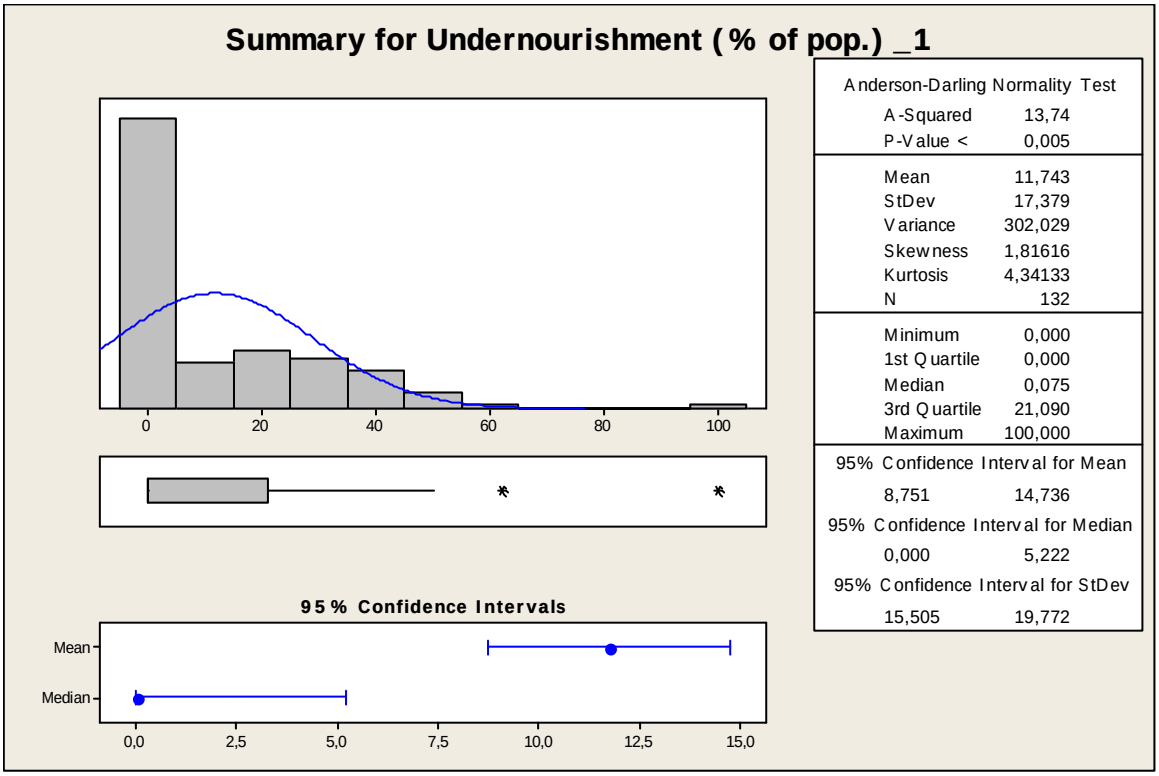
<u>Análise dos dados da nossa pesquisa</u>:

O gráfico mostra que a maior concentração de observações com 46 está no segundo quartil no intervalo de 25 a 47,5, que é onde se encontra a média de 39,7 e a mediana de 41,7. A segunda maior concentração está no terceiro quatil com 36 observações no intervalo de 47,5 a 60 e a terceira maior com 34 observações no primeiro quartil no intervalo de 2,5 a 25. Nota-se que tem-se apenas 12 observações no quarto quartil no intervalo de 70 a 92,5, nos quais se encontram os países destacados na tabela 2 a seguir. O Brasil encontra-se no segundo quartil com 38.

Tabela 2. Países com melhor índice de acesso a educação avançada.

País	Cód.	Região	Acesso à Educação Avançada
Australia	AUS	Australia and New Zealand	75,32
Canada	CAN	Northern America	82,21
Finland	FIN	Northern Europe	67,89
Ireland	IRL	Northern Europe	70,44
Israel	ISR	Western Asia	72,82
Japan	JPN	Eastern Asia	77,74
Netherlands	NLD	Western Europe	69,67
New Zealand	NZL	Australia and New Zealand	76,84
Russia	RUS	Eastern Europe	75,35
Sweden	SWE	Northern Europe	68,63
United Kingdom	GBR	Northern Europe	73,35
United States	USA	Northern America	89,37

4.6 –Undernourishment (% of pop.) (Desnutrição)



Undernourishment (% of pop.) (DESNUTRICAO)

Definição dado por:
http://pt.knoema.com/search?query=undernourishment%20o%20que%20%C3%A9

Proporção da população em estado de subnutrição. A desnutrição refere-se à condição de pessoas cujo consumo de energia alimentar é continuamente abaixo de um requisito de energia da ração mínima para a manutenção de uma vida saudável e da realização de uma atividade física leve, com um corpo de peso mínimo aceitável para atingido altura.

População abaixo do nível mínimo de consumo de energia da dieta (também referida como prevalência de desnutrição) mostra a percentagem da população com a ingestão de alimentos é insuficiente para atender às necessidades de energia na dieta de forma contínua. Mostrando os dados como 5 significa uma prevalência de desnutrição abaixo de 5%.

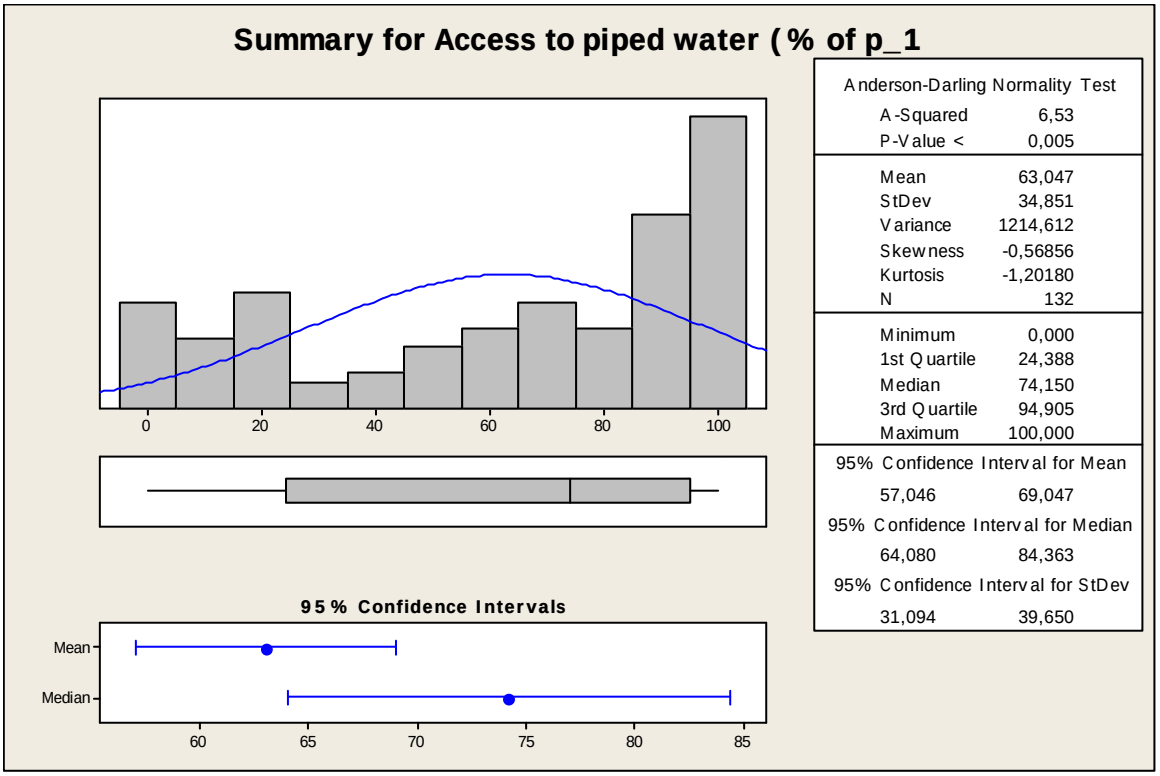
A análise na nossa pesquisa:

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a esquerda o que indica que os dados estão entre -5 e 5,0 com 76 observações. Nessa situação indicada pelo MINITAB está a Malásia com valor zero.

Na análise apresentada pelo Knoema em 2009 mostrava que a Malásia juntamente com Egito, Republica Dominicana, Jordania e Kwait apresentavam fatores igual a 5,0 %.

Percebe-se um número menor de países do que o esperado classificados entre 45 e 55 com apenas 4 observações, e 10 observações para a faixa de pontuação 35 à 45. O maior número de observações encontra-se na faixa de -5 a 5, com 76 observações. A mediana de 0,075 e o terceiro quartil de 21,090 indicam o poder discricionário do índice, que consegue separar o grupo de países em 6 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados. **O Brasil apresenta 6,9 apresentando-se acima da mediana, mas ainda no primeiro quartil, assim como a maior parte dos países com 82 observalções no intervalo de -5 a 10.**

4.7 – Access to piped water (% of pop.)(Água Encanada)



Access to piped water (% of pop.)(água encanada)
Definição dado por: http://www.wssinfo.org/data-estimates/table/
Mostra quantas residências possuem agua encanada ou outro tipo de processo que leva a água tratada para o consumo.

A análise da pesquisa:

O gráfico mostra uma tendência quase normal porque existem dois extremos quase distintos. Do lado esquerdo do gráfico estão 3 estratos distintos, sendo o mais próximo da origem com 12 países dentro de um intervalo de -5 e 5%, do lado deste estrato está o estrato com 8 países entre 5 e 15% e colado está o estrato com 13 países entre 15 e 25.

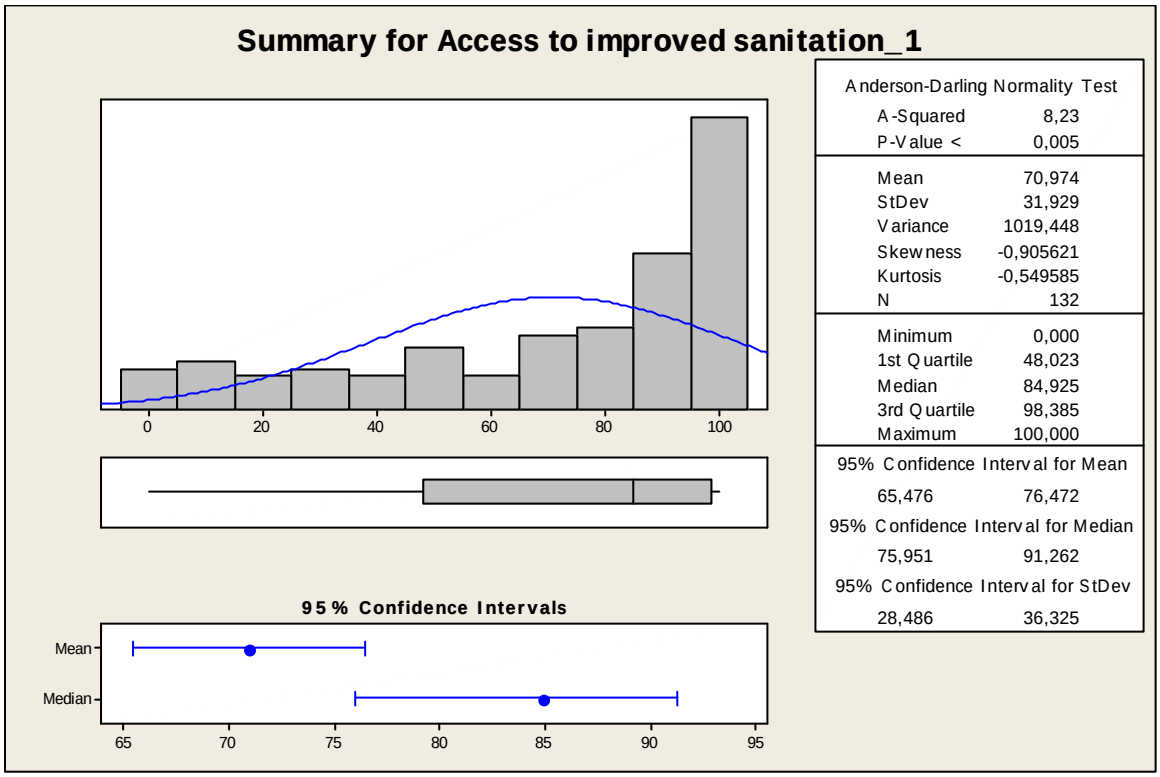
A mediana é 74,150 com o terceiro quartil com o valor de 94,905 com 11 estratos distintos. Vale o registro do lado esquerdo do gráfico com os estratos bem aparentes entre o 80 e 0 100. Sendo na faixa dos 80 temos 9 países entre 75 e 85; 22 países entre 85 e 95 e 33 países entre 95 e 105.

De acordo com os dados da UNICEF (2010) a evolução da utilização de diferentes tipos de fontes de água a partir de 1990-2010, por regiões do MDG - Millenium Development Goals (objetivos de desenvolvimento do milênio) apresentou dois grupos distintos de evolução. O primeiro é um conjunto de regiões em que a utilização de água canalizada para uma habitação, terreno ou quintal é baixa (30 por cento ou menos). Ele inclui a Afica sub-saariana, Oceania, Sul da Ásia e do Sudeste Asiático. Embora os ganhos na utilização de água canalizada nas instalações foram feitas nessas regiões, o progresso foi principalmente na qualidade da origem da água. Destacamos que 65 por cento da população do sul da Ásia estão usando outras fontes naturais ao invés de água encanada no local.

O segundo conjunto de regiões é composto por Ásia oriental, África do Norte, Ásia Ocidental e na América Latina e no Caribe, onde pelo menos 70 por cento da população está usando água encanada. Ásia Oriental (destaque para a China) teve um aumento significativo no abastecimento de água canalizada desde 1990, ganhando 35 pontos percentuais na cobertura nesta categoria em 20 anos. Isto representa 562 milhões de novos usuários que foram adicionados durante um período em que o mundo como um todo aumentou apenas 9 %. Ásia Oriental é também a região com o aumento mais expressivo no uso de água potável de diversas fontes com qualidade, a partir de 68 por cento em 1990 e passou para 91 por cento de cobertura em 2010. Isso representa um aumento de 23 ponto percentual, muito maior que qualquer outra região. **O Brasil encontra-se com 91.7 no quarto quartil, bem acima da mediana e da média.**

Site da Unicef: http://www.unicef.org/media/files/JMPreport2012.pdf

4.8 - Access to improved sanitation facilities (% of pop) (Acesso a instalações sanitárias melhoradas)



Anderson-Darling Normality Test	
A-Squared	8,23
P-Value <	0,005
Mean	70,974
StDev	31,929
Variance	1019,448
Skewness	-0,905621
Kurtosis	-0,549585
N	132
Minimum	0,000
1st Quartile	48,023
Median	84,925
3rd Quartile	98,385
Maximum	100,000
95% Confidence Interval for Mean	
65,476	76,472
95% Confidence Interval for Median	
75,951	91,262
95% Confidence Interval for StDev	
28,486	36,325

Análise: Access to improved sanitation facilities (% of pop)(acesso a instalações sanitárias melhoradas)

Os dados apresentados estão com a concentração do lado direito com uma média de 70,974 e mediana de 84,925, sendo que o terceiro quartil está com 98,385 indicando que esse estrato possui 43 países com melhores condições sanitárias.

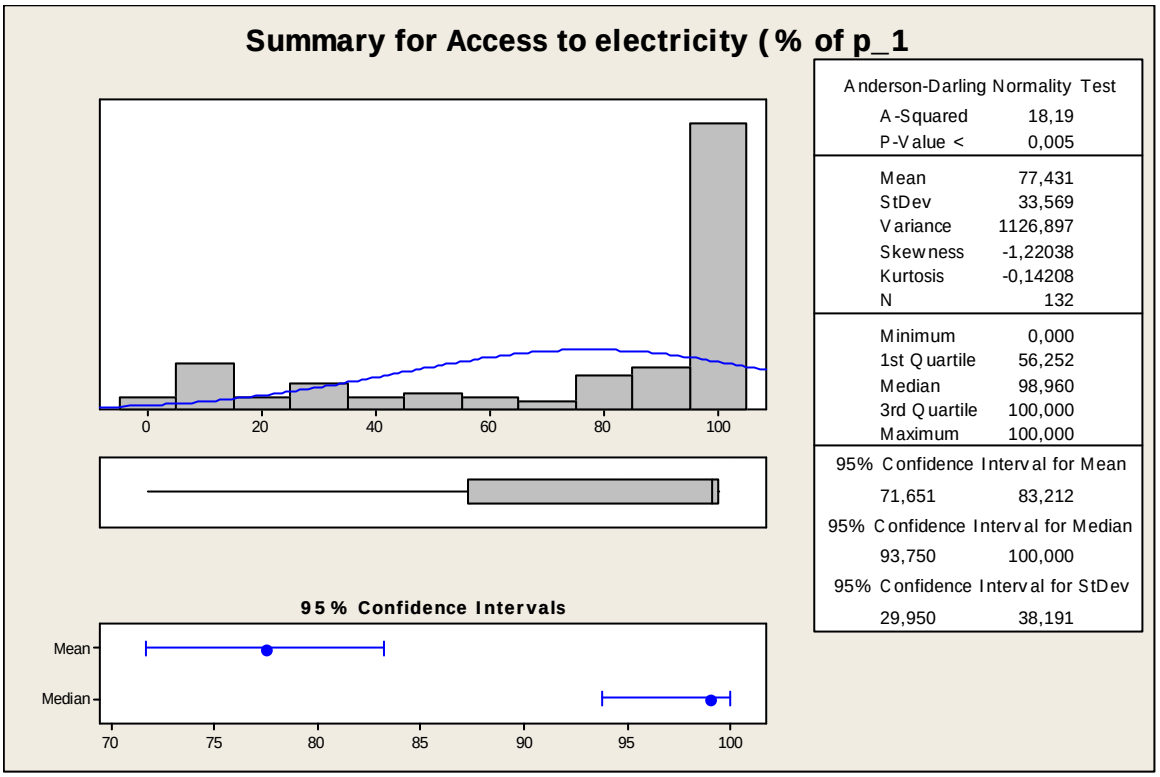
Do lado oposto estão 6 países com indicadores -5 e 5 positivos. Do lado a este estrato estão 6 países com valores entre 5 e 15. Os países estão na tabela a seguir.

O Brasil encontra-se com 80,8 acima da média e um pouco abaico da mediana.

Tabela 3 – Países com valores entre -5 e 15 no item Acesso a instalações sanitárias melhoradas

País	Índice
Benin	5,08
Burkina Faso	9,33
República do Chade	2,38
República do Congo	9,06
Ghana	4,29
Republica da Guine	9,84
Libéria	9,53
Madagascar	4,51
Mali	13,31
Moçambique	10,48
Republica do Niger	0,00
Tanazania	2,58
Republica do Togo	2,01

4.9 – Access to electricity (% of pop.) (Eletricidade)

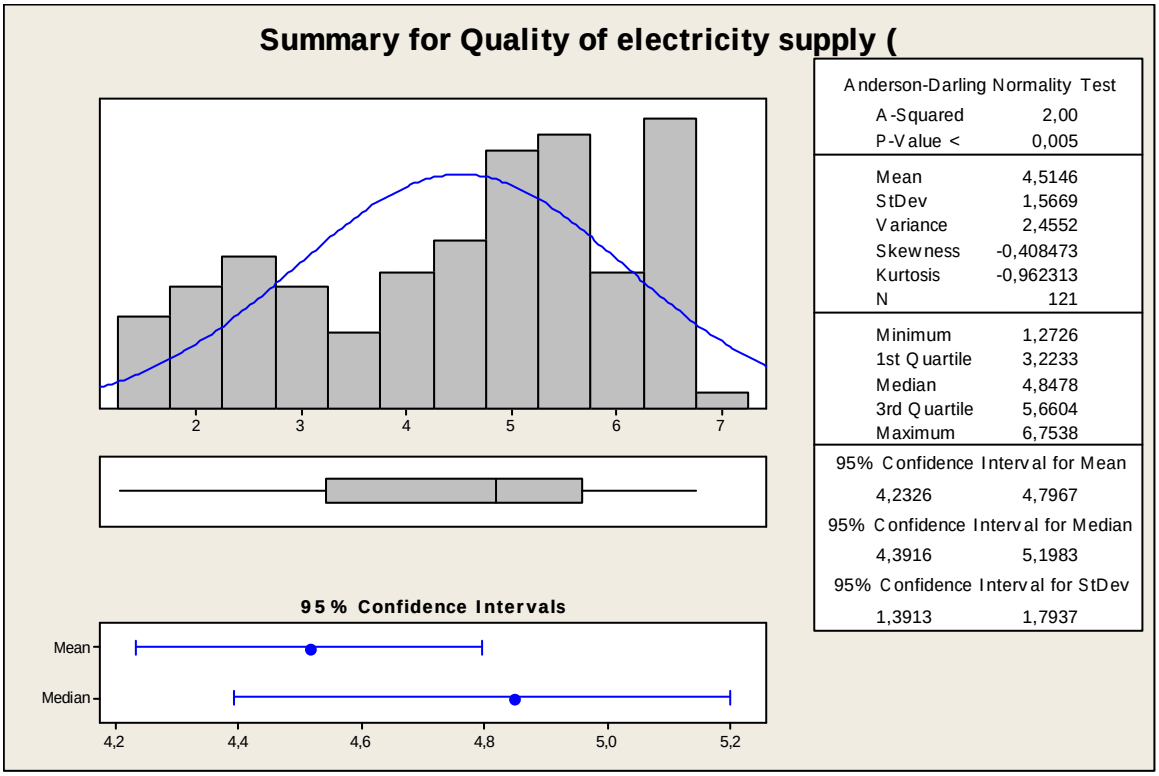


Análise:Access to electricity (% of pop.)(eletricidade)

Este indicador mostra o percentual da população de cada país que tem acesso a energia elétrica. De uma maneira geral existe um grande número de pessoas com acesso a energia elétrica com o estrato entre 95 e 105 para 75 países, com média de 77,431 e mediana de 98,96, o **Brasil encontra-se entre eles com 99**.

Isto quer dizer que metade da população mais significativa tem energia elétrica e outra metade bem menor não tem. Isto aparece no estrato entre 5 e 15 com 12 países.

4.10 – Quality of electricity supply (1=low; 7=high) (Qualidade fornecimento eletricidade)

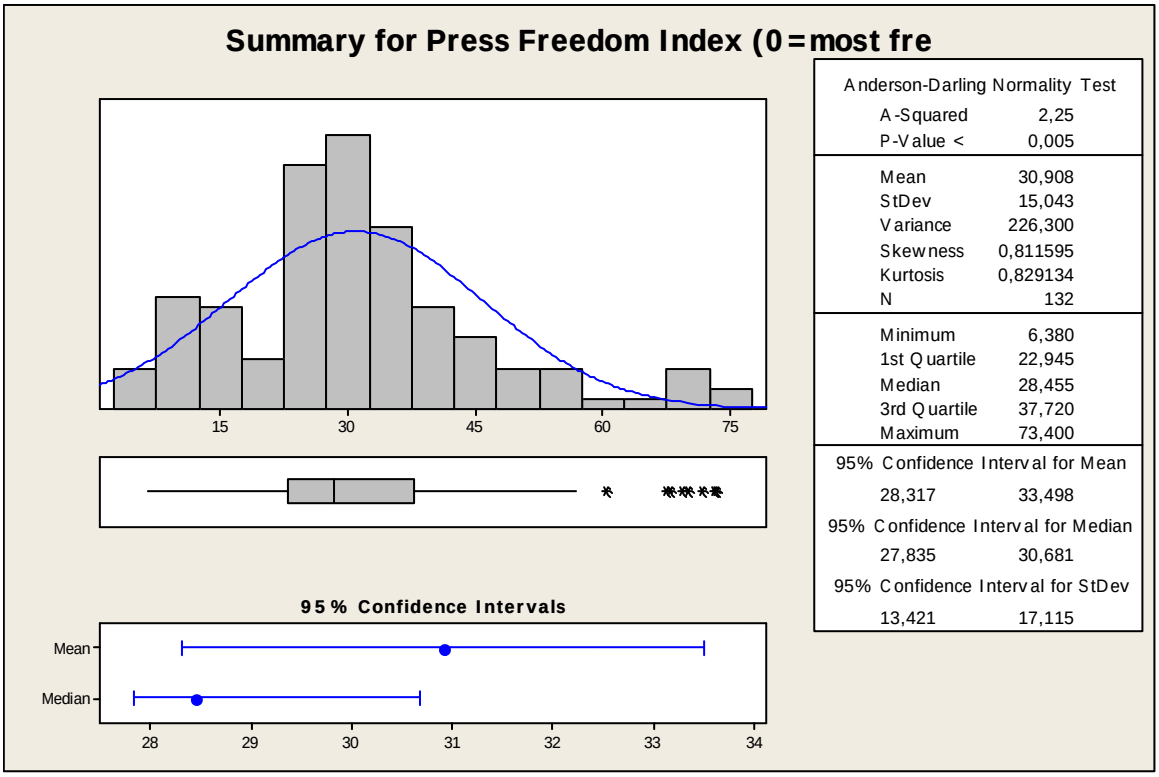


Análise: Quality of electricity supply (1=low; 7=high)(qualidade fornecimento eletricidade)

Ter acesso a energia não significa que ela esteja disponível 24 horas por dia ou sempre que o consumidor necessitar.

Esse item procura identificar a qualidade do fornecimento da energia e o gráfico mostra uma distribuição normal entre os dados. A média está com o valor de 4,5 e mediana 4,84, sendo que o terceiro quartil com 5,66 representado pelo estrato com 36 países. **O Brasil encontra-se entre eles com 4,8.**

4.11 - Press Freedom Index (0=most free; 100=least free) (Índica de Liberdade de Imprensa)



Variável: Press Freedom Index (0=most free; 100=least free)

Origem e definição dado pelo portal: http://en.rsf.org/

O grau de liberdade que os jornalistas, organizações de notícias, e internautas desfrutar em cada país, e os esforços envidados pelas autoridades para respeitar e assegurar o respeito por esta liberdade. No modelo SPI, as pontuações são agrupados em 7 bandas, de 1 (menos livre) a 7 (mais livre).

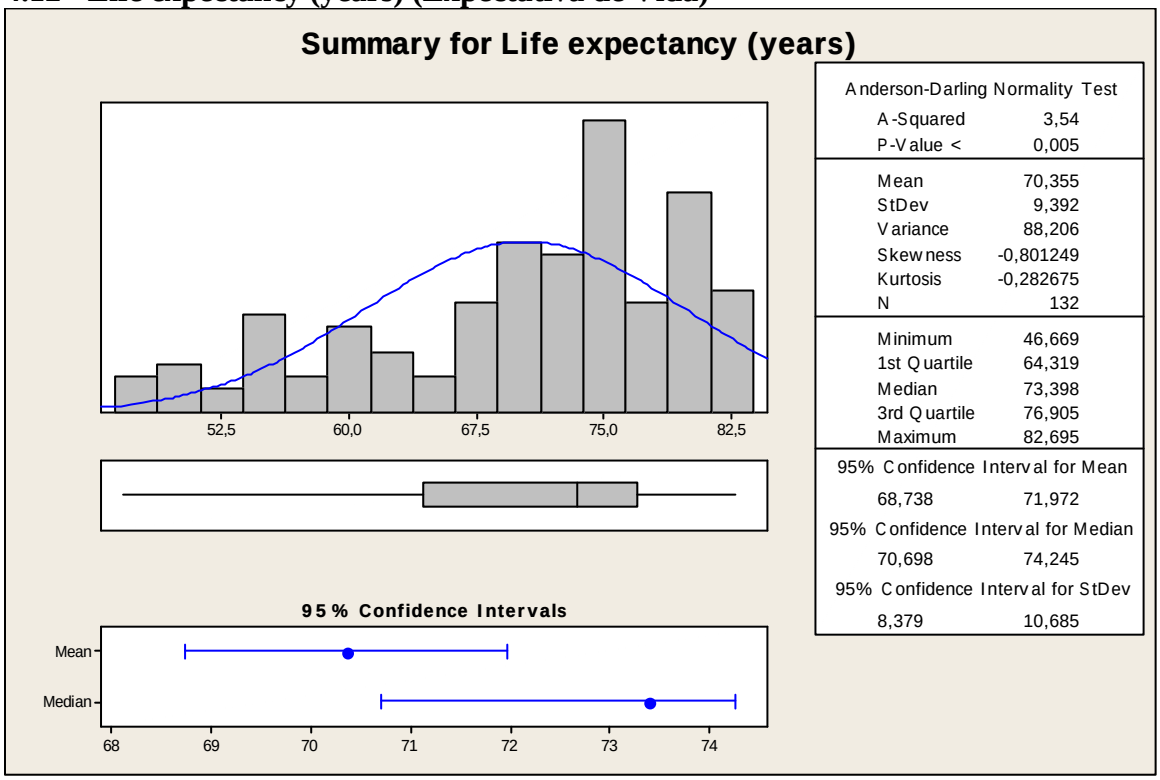
Análise dos dados da nossa pesquisa

O gráfico mostra uma curva pendendo levemente para a esquerda, com média de 30,90 e mediana de 28,45. A grande concentração de observações 78 encontra-se no segundo quartil no intervalo de 18,5 a 40, o Brasil esncontra-se neste quartil com 32,8, um pouco acima da mádia e da mediana. No primeiro quadrante encontram-se 26 observações no intervalo de 2,5 a 18,5, sendo assim 104 dos 132 países se encontram nos dois primeiros quadrantes com um índice menos do que 40% de liberdade de imprensa. Os dez países mais críticos estão em sua maioria no norte europeu conforme observa-se na tabela 4.

Tabela 4. Países com pior índice de liberdade de imprensa.

País	Cód.	Região	Liberdade de Imprensa
Denmark	DNK	Northern Europe	7,1
Estonia	EST	Northern Europe	9,3
Finland	FIN	Northern Europe	6,4
Iceland	ISL	Northern Europe	8,5
Jamaica	JAM	Caribbean	9,9
Netherlands	NLD	Western Europe	6,5
New Zealand	NZL	Australia and New Zealand	8,4
Norway	NOR	Northern Europe	6,5
Sweden	SWE	Northern Europe	9,2
Switzerland	CHE	Western Europe	9,9

4.12 - Life expectancy (years) (Expectativa de Vida)



Variável: Life expectancy (years)

Origem e definição dado pelo portal: http://data.worldbank.org/indicator/SP.DYN.LE00.IN

O número de anos que um recém-nascido viveria se os padrões de mortalidade permanecessem os mesmos do momento em que ele nasceu durante toda a sua vida.

Análise dos dados da nossa pesquisa:

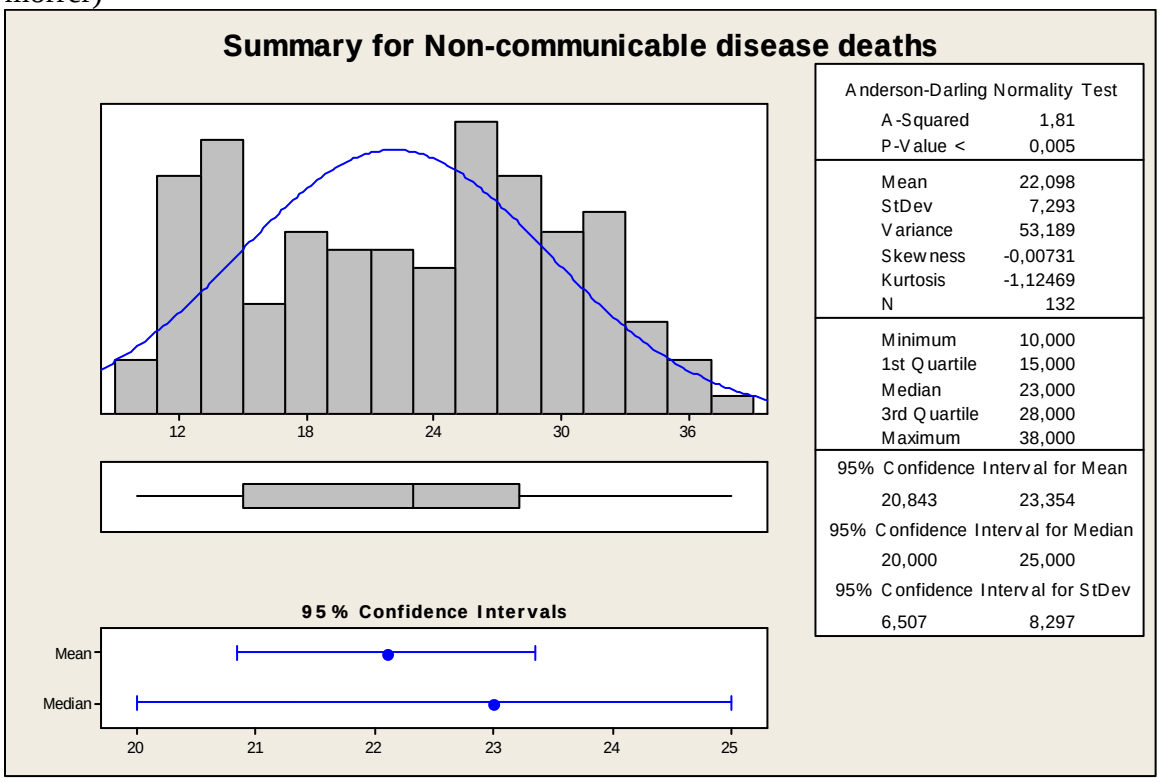
O gráfico mostra que a grande concentração de observações estão no terceiro e quarto quartil com 50 e 49 respectivamente, com uma média de 70,3 e uma mediana de 73,3 em um intervalo de 65 a 83,75, o Brasil encontra-se na mediana com 73,3. Observam-se 13 países no primeiro quartil com menor expectativa de vida, todos estão no continente africano conforme visualizado na tabela 5.

Tabela 5. Países com pior expectativa de vida em anos.

País	Cód.	Região	Expectativa de Vida
Angola	AGO	Middle Africa	51,1
Botswana	BWA	Southern Africa	46,7
Burkina Faso	BFA	Western Africa	55,4
Burundi	BDI	Eastern Africa	53,1
Cameroon	CMR	Middle Africa	54,1
Central African Republic	CAF	Eastern Africa	48,8
Chad	TCD	Middle Africa	50,2
Lesotho	LSO	Southern Africa	48,2
Malawi	MWI	Southern Africa	54,1
Mali	MLI	Western Africa	54,2
Mozambique	MOZ	Eastern Africa	49,5
Nigeria	NGA	Western Africa	51,7

South Africa	ZAF	Southern Africa	55,3
Swaziland	SWZ	Southern Africa	48,7

4.13 - Non-communicable disease deaths between the ages of 30 and 70 (probability of dying) Mortes por doenças não transmissíveis entre as idades de 30 e 70 (probabilidade de morrer)



Variável: Non-communicable disease deaths between the ages of 30 and 70 (probability of dying) Mortes por doenças não transmissíveis entre as idades de 30 e 70 (probabilidade de morrer)

Origem e definição dado pelo portal: http://apps.who.int/gho/data/node.main.A857?lang=en

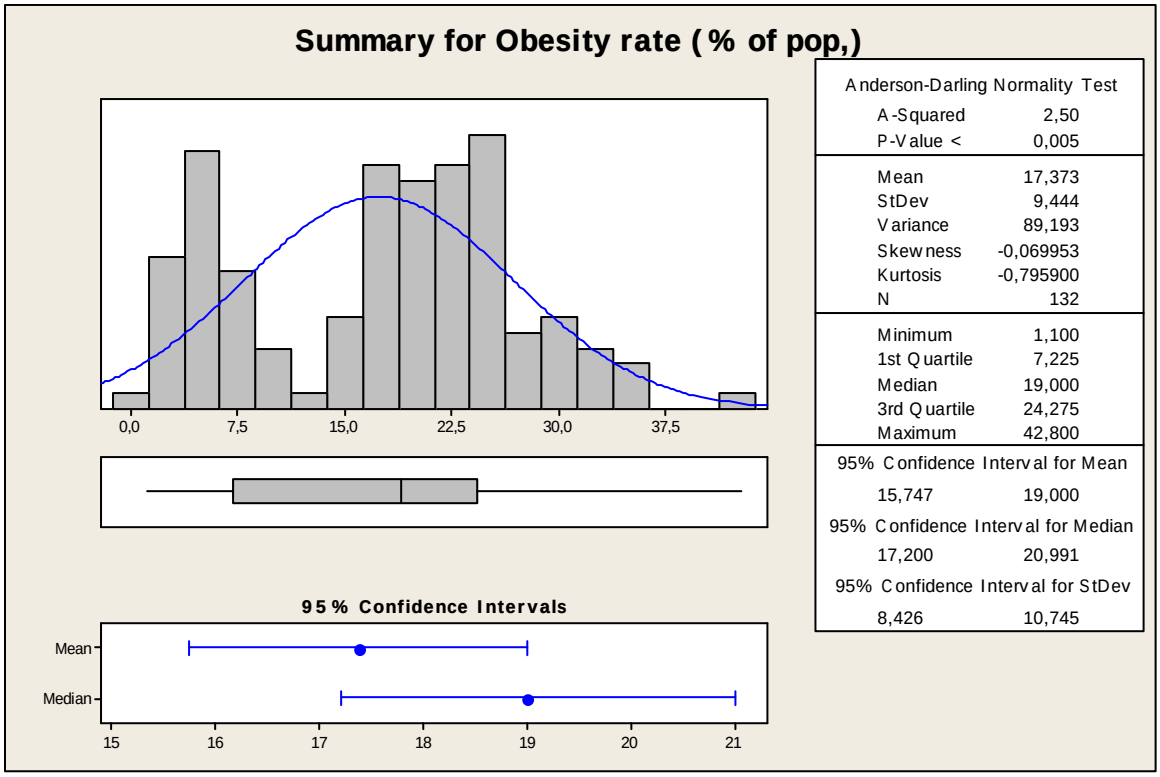
A probabilidade de morrer entre as idades de 30 e 70 anos por doenças cardiovasculares, câncer, diabetes ou doença respiratória crônica.

.

Análise dos dados da nossa pesquisa

O gráfico mostra ua curva quase perfeita e uma concentração maior de observações no primeiro e no terceiro quartil. Os países com maior probabilidade de contrair doenças cardiovasculares, câncer, diabetes e doenças respiratórias então no continente africanos, os cinco mais críticos são: Kazakhstan com 38, Malawi com 36, Swaziland com 35, Zambia e Guinea com 34. A média de 22 e a mediana de 23 estão no segundo quartil, no qual se insere o Brasil com 20.

4.14 - Obesity rate (% of pop.) Taxa de Obesidade



Variável: Obesity rate (% of pop.) Taxa de Obesidade

Origem e definição dado pelo portal: http://apps.who.int/gho/data/node.main.A900?lang=en

A percentagem da população com um índice de massa corporal (IMC) de 30 kg / m2 ou superior (estimativa idade-padronizados), ambos os sexos.

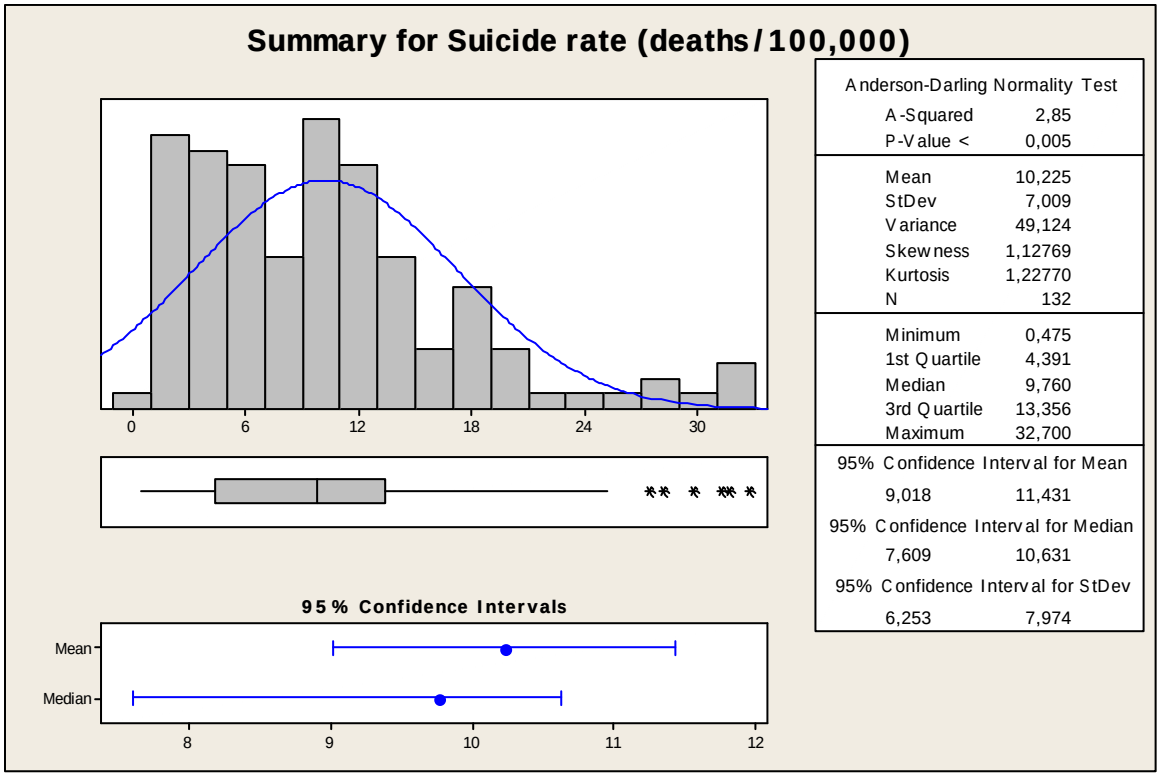
Análise dos dados da nossa pesquisa:

O gráfico mostra que o intervalo de 1,25 a 10 possui 39 observações enquanto que o intervalo de 10 a 20 possui 40 observações, nele estão contidos a média 17,3 e a mediana 19 e o Brasil 19,5 está entre eles, sendo assim a maior concentração está no primeiro e segundo quartis com os menores índices de obesidade. No entanto o maior numero de observações está no terceiro quartil no intervalo de 20 a 32,75 com 47. Os países com maior taxa de obesidade estão contidos na tabela 6.

Tabela 6. Países com maior taxa de obesidade.

País	Cód.	Região	Taxa de Obesidade
Egypt	EGY	Northern Africa	34,6
Jordan	JOR	Western Asia	34,3
Kuwait	KWT	Western Asia	42,8
Mexico	MEX	IBE	32,8
Saudi Arabia	SAL	Western Asia	35,2
South Africa	ZAF	Southern Africa	33,5
Trinidad and Tobago	TTO	Caribbean	30
United Arab Emirates	ARE	Western Asia	33,7
United States	USA	Northern America	31,8
Venezuela	VEM	IBE	30,8

4.15 - Suicide rate (deaths/100,000) (Taxa de Suicídio)



Variável: Suicide rate (deaths/100,000) (Taxa de Suicídio)
Origem e definição dado pelo portal:
https://cloud.ihme.washington.edu/public.php?service=files&t=d559026958b38c3f4d12029b36d783da&download&path=/2010/IHME_GBD_2010_RESULTS_2010_SELF-HARM_Y2013M11D15.zip

A mortalidade devido a lesão auto infligida, por 100.000 pessoas, a idade ajustada.

Análise dos dados da nossa pesquisa:

O gráfico mostra que o maior número de observações está no segundo quartil entre 7,5 e 16 com 55 observações aonde encontram-se a média 10,2 e a mediana de 9,7.O Brasil com 6,5 encontra-se no primeiro quartil entre -1 e 7,5 com 54 observações. Sendo assim observa-se que a maior parte dos países encontra-se 81 encontram-se no intervalo de -1 a 11 com baixas taxas de suicídios. A seguir são apenas 7 os países que encontram-se no quarto quartil de 25,5 a 33 e encontram-se listados na tabela 7, assim como os sete países com menor taxa de suicídio.

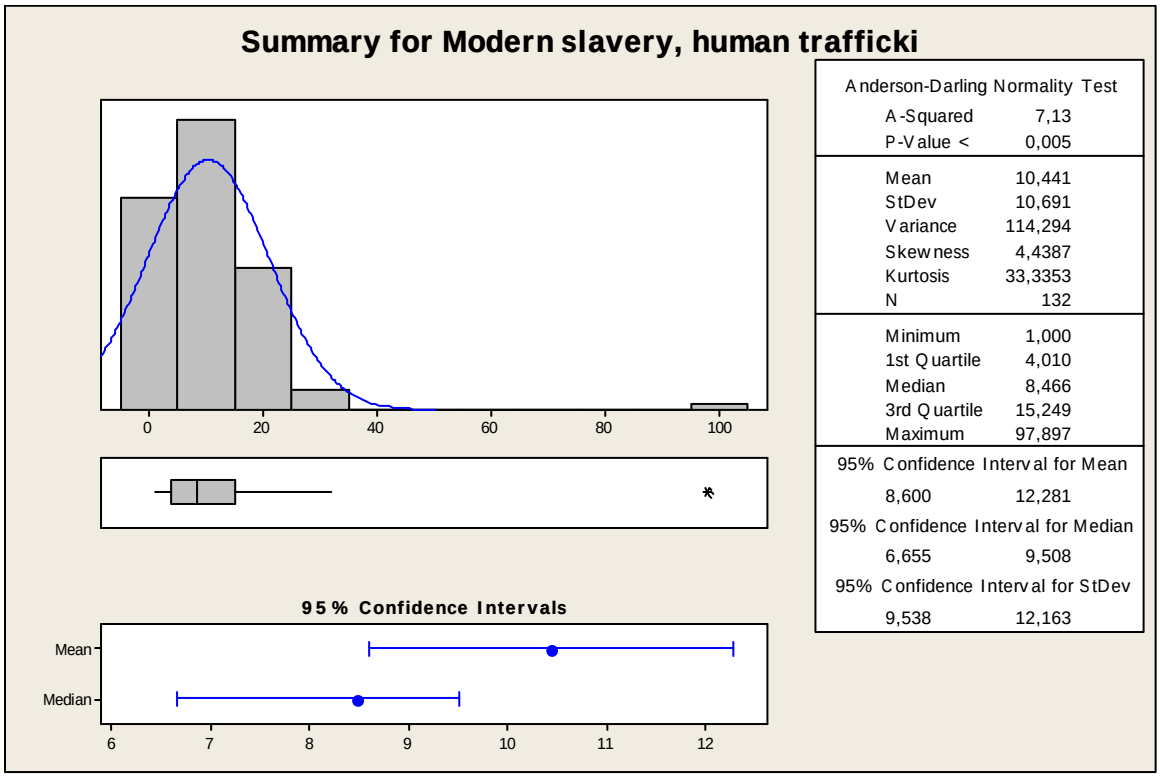
Tabela 7. Países com maior taxa de suicídios.

País	Cód.	Região	Taxa de Suicídio
Belarus	BLR	Eastern Europe	28,2
Guyana	GUY	South America	29,7
India	IND	Southern Asia	25,2
Kazakhstan	KAZ	Cenrtral Asia	31,6
Lithuania	LTU	Northern Europe	31,2
Russia	RUS	Eastern Europe	27,5
Sri Lanka	LKA	Southern Asia	32,7

Tabela 7. Países com menor taxa de suicídios.

País	Cód.	Região	Taxa de Suicídio
Egypt	EGY	Northern Africa	2,3
Ghana	GHA	Western Africa	2,1
Jamaica	JAM	Caribbean	1,2
Kuwait	KWT	Western Asia	2,2
Lebanon	LBN	Western Asia	2,1
Saudi Arabia	SAU	Western Asia	0,5
Tunisia	TUN	Northern Africa	2,3

4.16 -Modern slavery, human trafficking and child marriage (1=low; 100=high) Escravidão moderna, tráfico de seres humanos e casamento infantil



Variável: Modern slavery, human trafficking and child marriage (1=low; 100=high) Escravidão moderna, tráfico de seres humanos e casamento infantil

Origem e definição dado pelo portal: http://www.globalslaveryindex.org/report/

A medida combinada de três variáveis: prevalência estimada de escravidão moderna em cada país (que representam 95% do total); uma medida do nível de tráfico de seres humanos para e de cada país (contas para 2,5%); e uma medida do nível da criança e do casamento precoce em cada país (contas para 2,5%), escala de 1 (baixo escravidão) a 100 (alta escravidão).

Análise dos dados da nossa pesquisa:

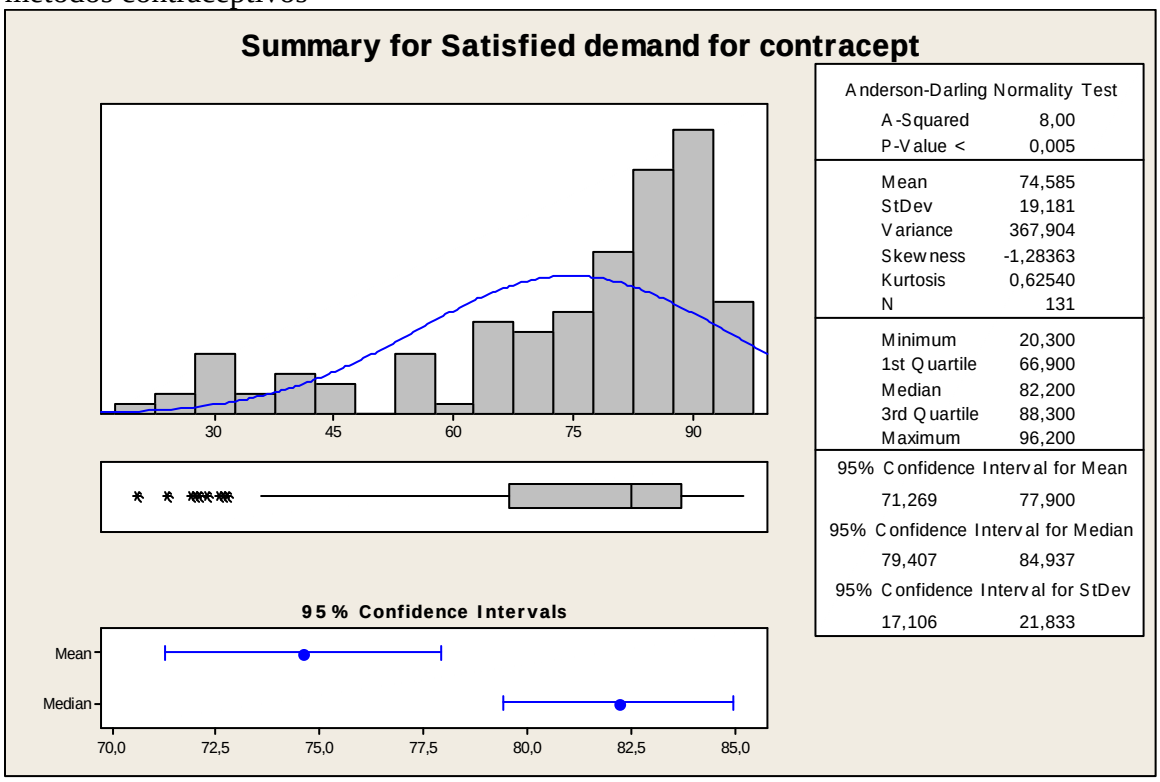
O gráfico mostra que temos uma curva tendendo para a esquerda com um outlier a direita que é a Mauritania com 97,9 conforme descrito na tabela 10, na qual estão destacados também os 4 países com maior índice depois deste, que se encontram no intervalo de 25 a 35. A maior parte dos países 57, encontram-se no intervalo de 5 a 15, no qual se encontram a média 10,4 e a mediana de 8,4. 42 países encontram-se no intervalo de -5 a 5. Nota-se com isso que ainda existem essas questões a serem trabalhadas na maior parte dos países estudados

em maior ou menor grau. Uma sugestão é que em futuras pesquisas este indicador seja desmembrado para que se tenha maior precisão dos dados.

Tabela 8. Países com maior taxa de escravidão, trafico humano e casamento infantil.

País	Cód.	Região	Escravidão e Trafico
India	IND	Southern Asia	30,8
Mauritania	MRT	Western Africa	97,9
Moldova	MDA	Eastern Europe	25,7
Nepal	NPL	Southern Asia	26,6
Pakistan	PAK	Southern Asia	32,1

4.17 - Satisfied demand for contraception (% of women) Demanda de satisfação com métodos contraceptivos



Variável: Satisfied demand for contraception (% of women) Demanda de satisfação com métodos contraceptivos
Origem e definição dado pelo portal:
http://www.un.org/esa/population/publications/WCU2012/Alkema-etal_2013_Natl-regl&global-rates&trends-family-planning_WEB-APPENDIX.pdf

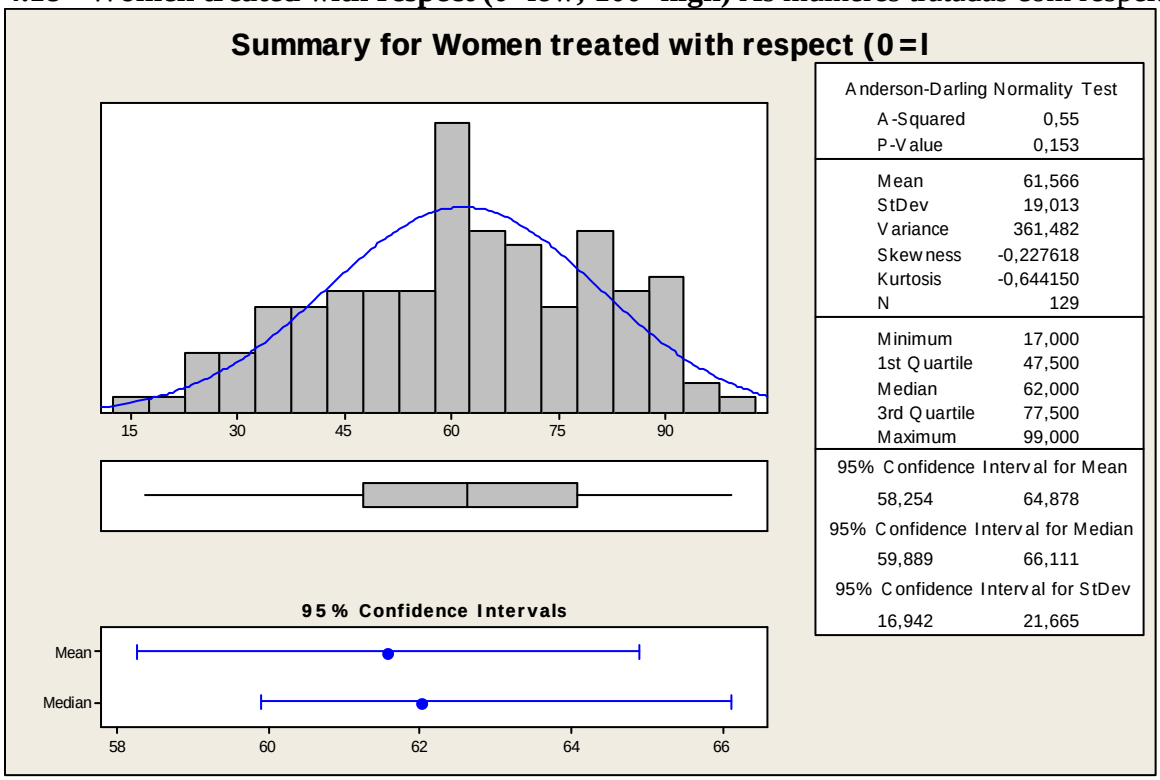
A taxa de prevalência contraceptiva para mulheres com idade entre 15-49 anos que são casados ou em união através da soma da taxa de prevalência de contraceptivos e necessidade não atendida de contracepção.

Análise dos dados da nossa pesquisa:

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a direita o que indica que os dados estão concentrados no quarto quartil com intervalo de 75,5 a 97,5 com 81 observações ou seja a maior parte dos países, neste quartil encontra-se também a mediana com 82,2. Jã a média de 74,5 encontra-se no final do terceiro quartil com intervalo de 60 a 75.5 com 26 observações. Pode-se dizer que grande parte dos países possuem um bom índice de satisfação com métodos contraceptivos o que corrobora com a diminuição do índice

mundial de natalidade, segundo a DSW Fundação Alemã População Mundial em 2014 o número de filhos por mulher caiu pela metade desde 1970 se for considerada a população mundial, a 40 anos a média por mulher era de 4,7 filhos, hoje é de 2,5.

4.18 - Women treated with respect (0=low; 100=high) As mulheres tratadas com respeito



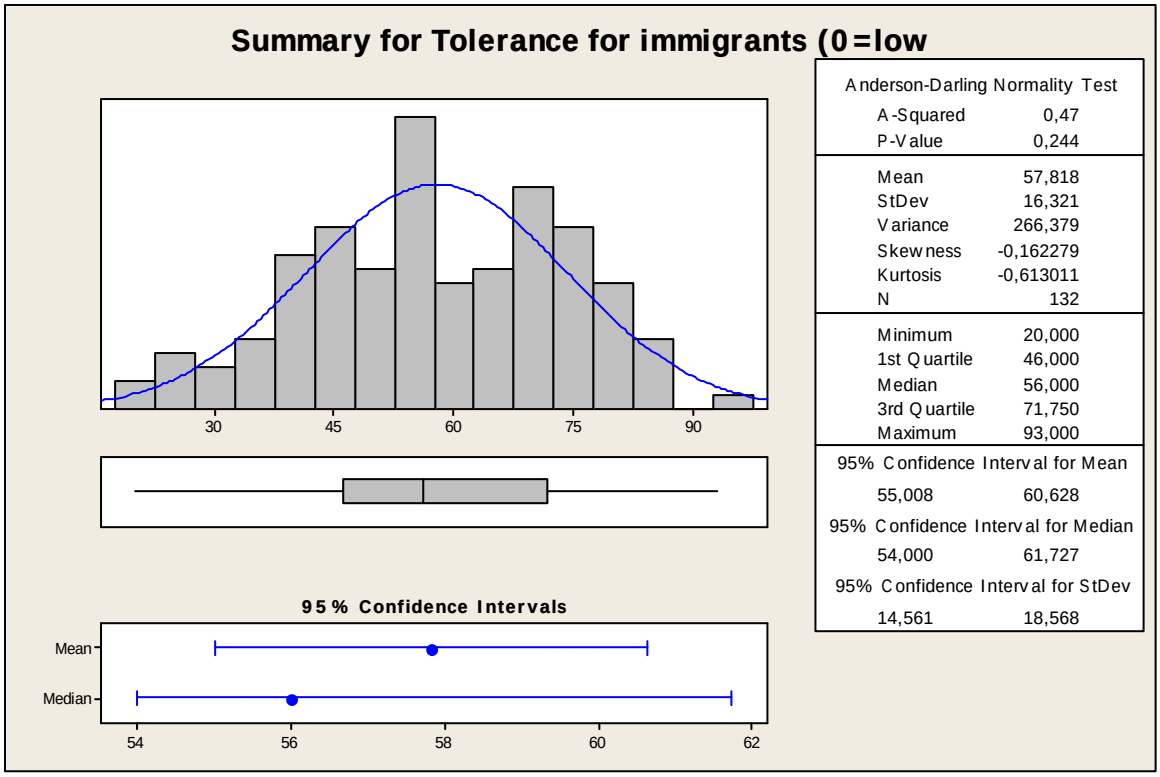
Variável: Women treated with respect (0=low; 100=high) As mulheres tratadas com respeito

Origem e definição dado pelo portal: A percentagem de respondentes do sexo feminino que respondem sim à pergunta, "Você acredita que as mulheres neste país são tratadas com respeito e dignidade?"

<u>Análise dos dados da nossa pesquisa:</u>

O gráfico mostra uma curva normal, a maior parte das observações 54 estão no terceiro quartil com um intervalo de 57,5 a 80 na qual está inserida a média de 61,5 e a mediana de 62.

4.19 Tolerance for immigrants (0=low; 100=high)



Variável: Tolerance for immigrants (0=low; 100=high) Tolerância aos Imigrantes

Origem e definição dado pelo portal:

A percentagem de inquiridos que respondem sim à pergunta: "Será que a cidade ou área onde você vive é um bom lugar para um imigrantes de outros países viver?"

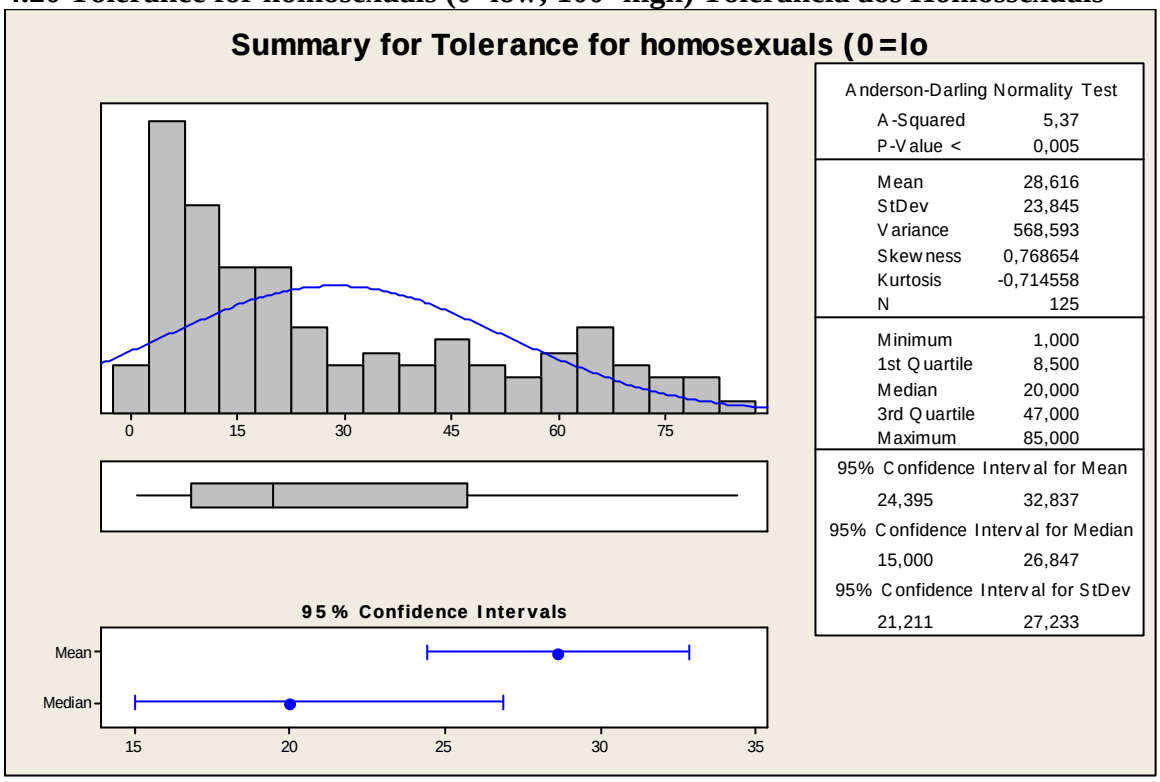
<u>Análise dos dados da nossa pesquisa</u>:

O gráfico mostra uma curva normal, a maior parte das observações 55 estão no segundo quartil com um intervalo de 35,5 a 57,5 na qual está inserida a média de 57,8 e a mediana de 56. 48 observações são encontradas no terceiro quartil com intervalo de 57,5 a 77,5. Nota-se 14 observações no primeiro quartil e 15 no quarto quatil, sendo assim ainda é visivel um percentual acima de 50% de discriminação a imigrantes em 62 países. O Brasil está entre eles com um percentual de 69, acima da média.

Tabela 9. Países com maior taxa de discriminação ao imigrante.

País	Cód.	Região	Discriinação ao imigrante
Australia	AUS	Australia and New Zealand	87
Burkina Faso	BFA	Western Africa	86
Canada	CAN	Northern America	93
Ireland	IRL	Northern Europe	84
Mali	MLI	Western Africa	80
New Zealand	NZL	Australia and New Zealand	87
Paraguay	PRY	IBE	87
Spain	ESP	IBE	81
Sweden	SWE	Northern Europe	80
United Arab Emirates	ARE	Western Asia	80
United States	USA	Northern America	82
Uruguay	URY	IBE	82

4.20 Tolerance for homosexuals (0=low; 100=high) Tolerância aos Homossexuais



Variável: Tolerance for homosexuals (0=low; 100=high) Tolerância aos Homossexuais

Origem e definição: O percentual de entrevistados que respondem sim à pergunta: "Será que a cidade ou área onde você vive é um bom lugar para pessoas gay ou lésbica viverem?"

<u>Análise dos dados da nossa pesquisa</u>:

Ainda é observada uma enorme intolerância ao homossexual na maior parte dos países, -2,5 a 42,5 com 89 observações, sendo que a grande concentração está no intervalo de 2,5 a 25 com 66 observações.

Tabela 10. Países menos tolerantes ao homossexual.

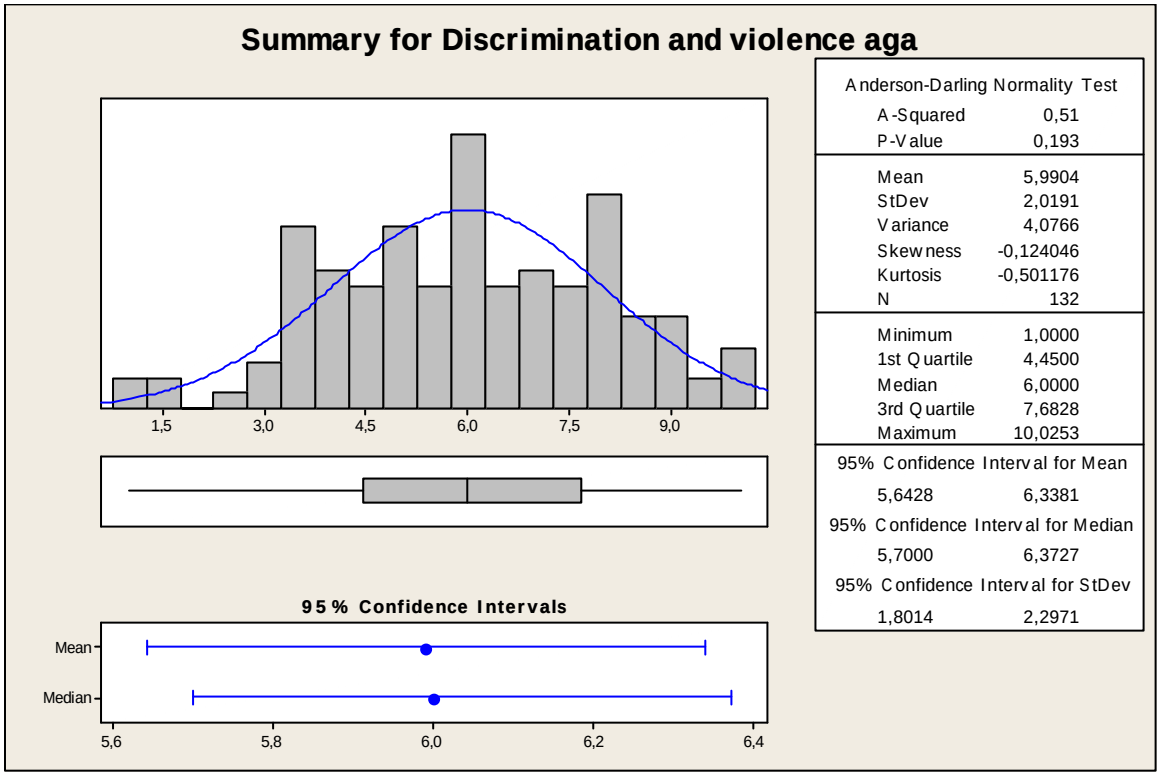
País	Cód.	Região	Discriinação ao homosexual
Burundi	BDI	Eastern Africa	3
Chad	TCD	Middle Africa	4
Ghana	GHA	Western Africa	4
Indonesia	IDN	South-Eastern Asia	2
Mali	MLI	Western Africa	2
Niger	NER	Western Africa	3
Rwanda	RWA	Eastern Africa	2
Senegal	SEN	Western Africa	3
Sri Lanka	LKA	Southern Asia	4
Tajikistan	TJK	Central Asia	1

Tabela 10. Países mais tolerantes ao homossexual.

País	Cód.	Região	Discriinação ao homosexual
Canada	CAN	Northern America	81

Iceland	ISL	Northern Europe	79
Ireland	IRL	Northern Europe	75
Netherlands	NLD	Western Europe	85
Spain	ESP	IBE	80
United Kingdom	GBR	Northern Europe	75
United States	USA	Northern America	70
Uruguay	URY	IBE	76

4.21 - Discrimination and violence against minorities (0=low; 10=high) Discriminação e violência contra as minorias



Variável: Discrimination and violence against minorities (0=low; 10=high) Discriminação e violência contra as minorias

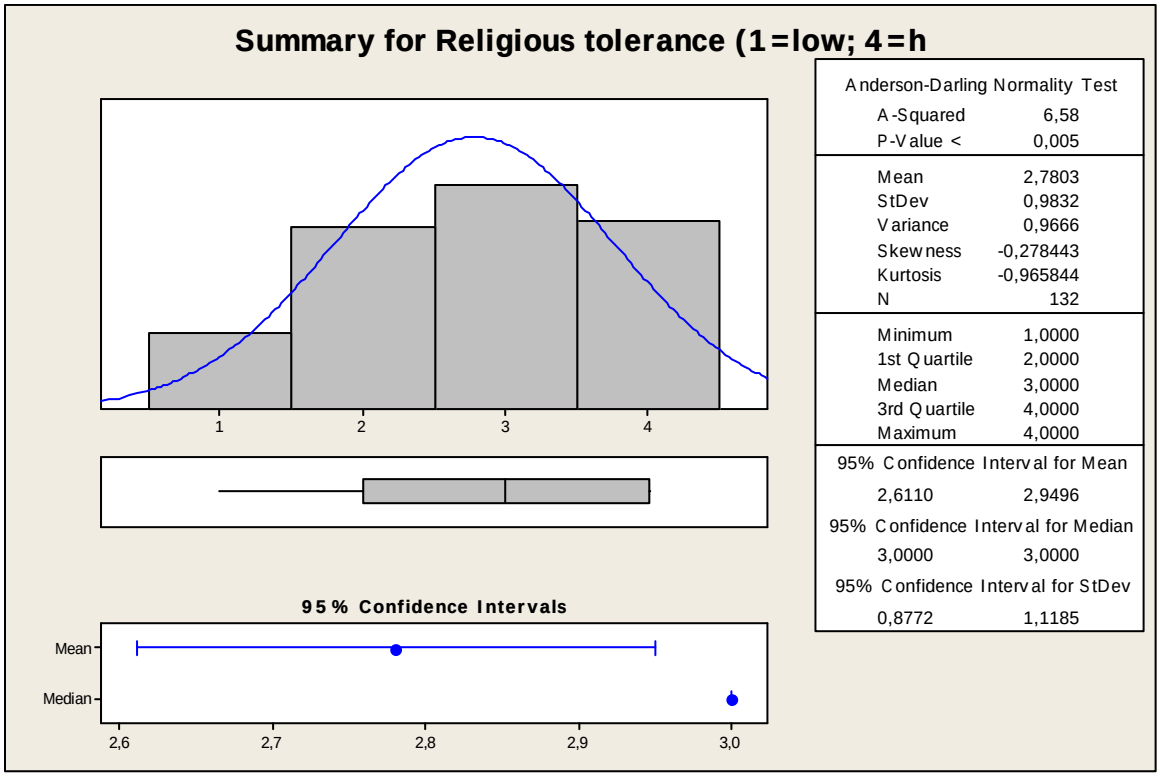
<u>Origem e definição dado pelo portal:</u> http://ffp.statesindex.org/rankings-2013-sortable

A discriminação, a impotência, a violência étnica, violência comunal, a violência sectária e violência religiosa, medido em uma escala de 0 (baixa pressão) a 10 (pressões muito altas)

<u>Análise dos dados da nossa pesquisa</u>:

Nota-se no gráfico uma curva normal no qual tanto a média de 5,99 quanto a mediana de 6 encontra-se no terceiro quartil onde está a maior concentração de observações com 17 delas no interbalod e 5,75 a 6,25.

4.22 - Religious tolerance (1=low; 4=high) Tolerância Religiosa



Variável: Religious tolerance (1=low; 4=high) Tolerância Religiosa

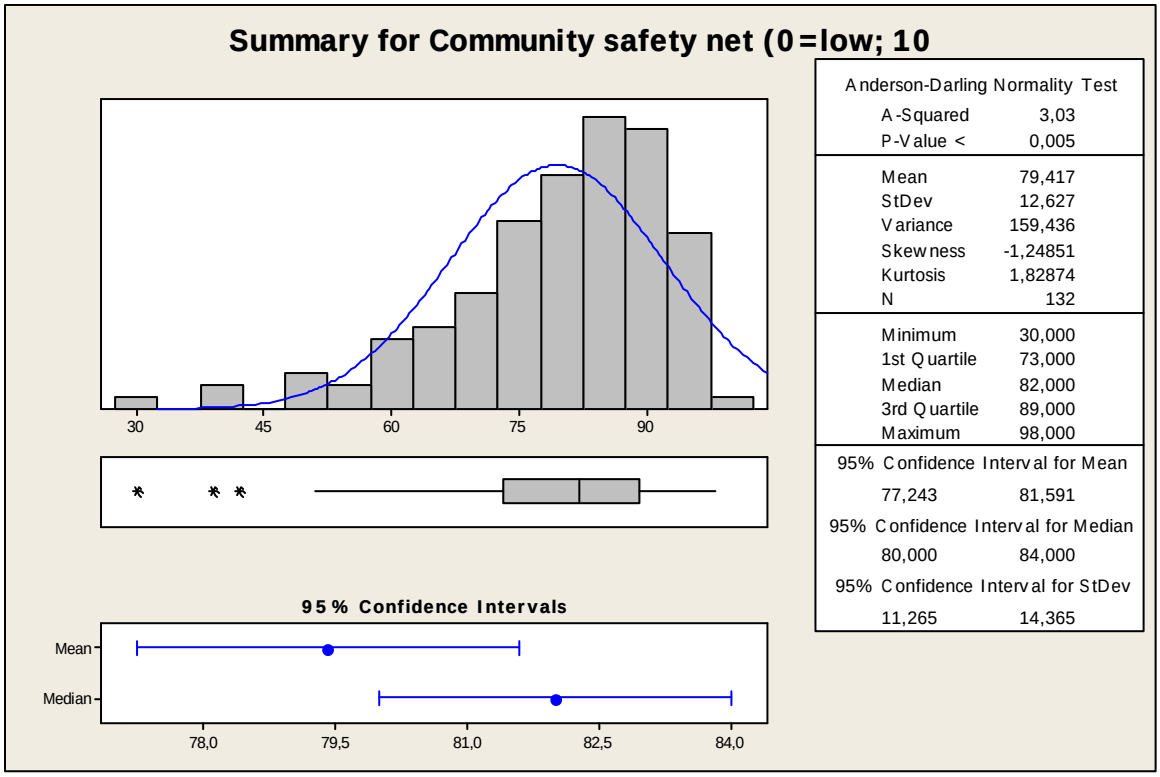
Origem e definição dado pelo portal: http://www.pewforum.org/2014/01/14/appendix-3-social-hostilities-index/

Uma medida de 13 tipos de hostilidade religiosa por particulares, organizações ou grupos da sociedade, incluindo conflitos relacionados com religião armados ou terrorismo, a multidão ou a violência sectária, o assédio sobre o traje por razões religiosas ou outras intimação ou abuso relacionados com a religião. No modelo SPI, as pontuações variam de 1 (baixo) a 4 (muito alto).

Análise dos dados da nossa pesquisa:

Nota-se que ainda é fote a intolerência religiosa, com média 2,9 e mediana 3, sendo que 44 observações estão no intevalo de 2,5 a 3,5 e 37 observações estão no interbalo 3,5 a 4,5, sendo que 4 e considerado como alta intolerância.

4.23 - Community safety net (0=low; 100=high) Rede de segurança comunitária



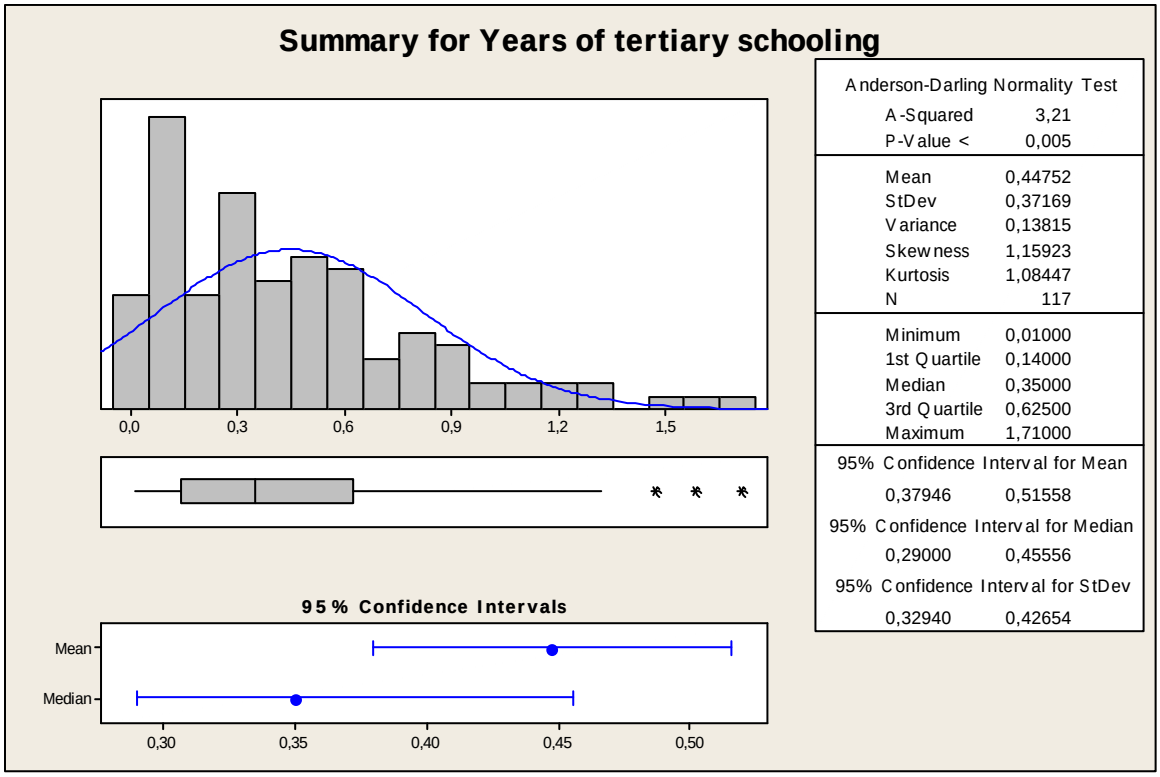
Variável: Community safety net (0=low; 100=high) Rede de segurança comunitária

Origem e definição dado: O percentual de entrevistados que respondem sim à pergunta: "Você tem parentes ou amigos com quem contar para ajudá-lo sempre que se encontrar em uma situação de apuros?

<u>Análise dos dados da nossa pesquisa</u>:

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a direita o que indica que os dados estão concentrados no quarto quartil com intervalo de 81,5 a 102,5 com 70 observações, isso indica que a maior parte dos países possui uma boa rede de segurança comunitária, com uma mediana de 82 e uma média de 79,4.

4.24 - Years of tertiary schooling Anos de escolaridade terciária

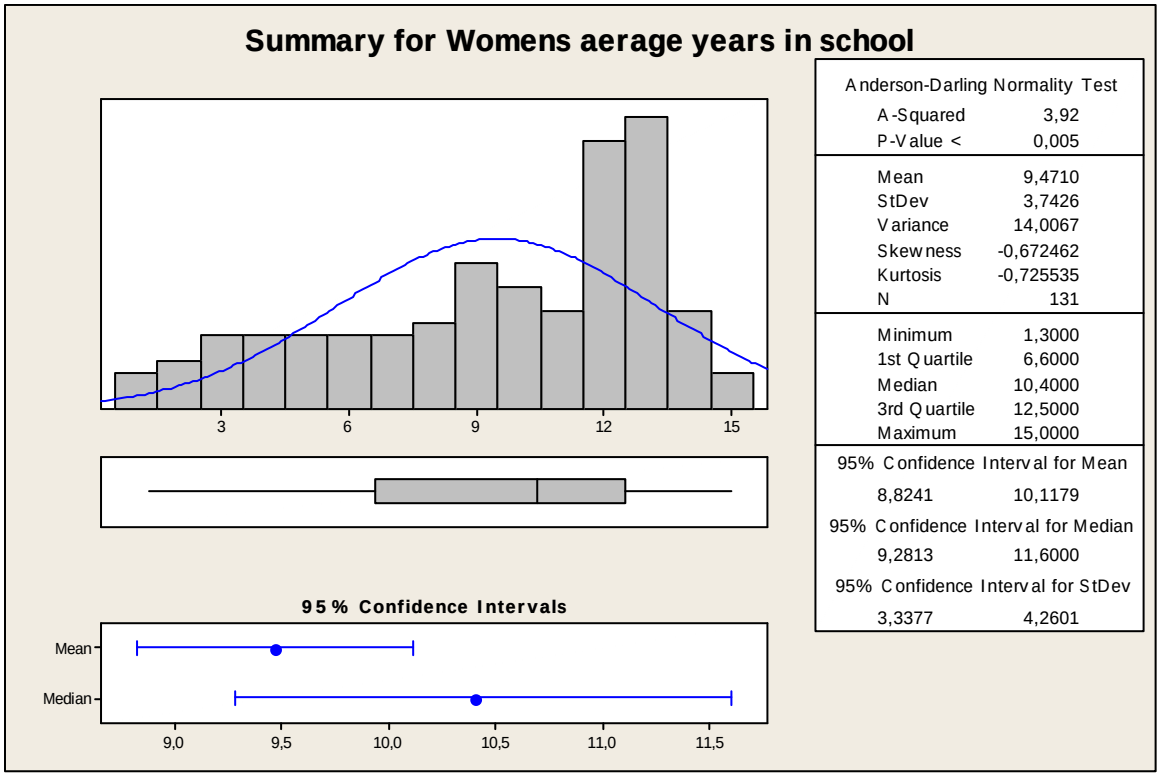


Variável: Years of tertiary schooling Anos de escolaridade terciária
Origem e definição dado pelo portal:
http://databank.worldbank.org/Data/Views/VariableSelection/SelectVariables.aspx?source=Education%20Statistics
A média de anos de educação superior concluído entre as pessoas com idade acima de 25.

Análise dos dados da nossa pesquisa:

O gráfico indica que a distribuição com a curtose (curva do gráfico) para a esquerda o que indica que os dados estão concentrados no priemeiro quartil com intervalo de -0,05 a 0,40 com 63 observações e 0,40 a 0,8 com 38 observações, sendo que nota-se apenas 5 observações no intervalo de 1,30 a 1,75, demostrando quão baixa é a média de pessoas que concluem a educação superior no mundo acima de 25 anos.

4.25 - Women's average years in school Média de anos das mulheres na escola



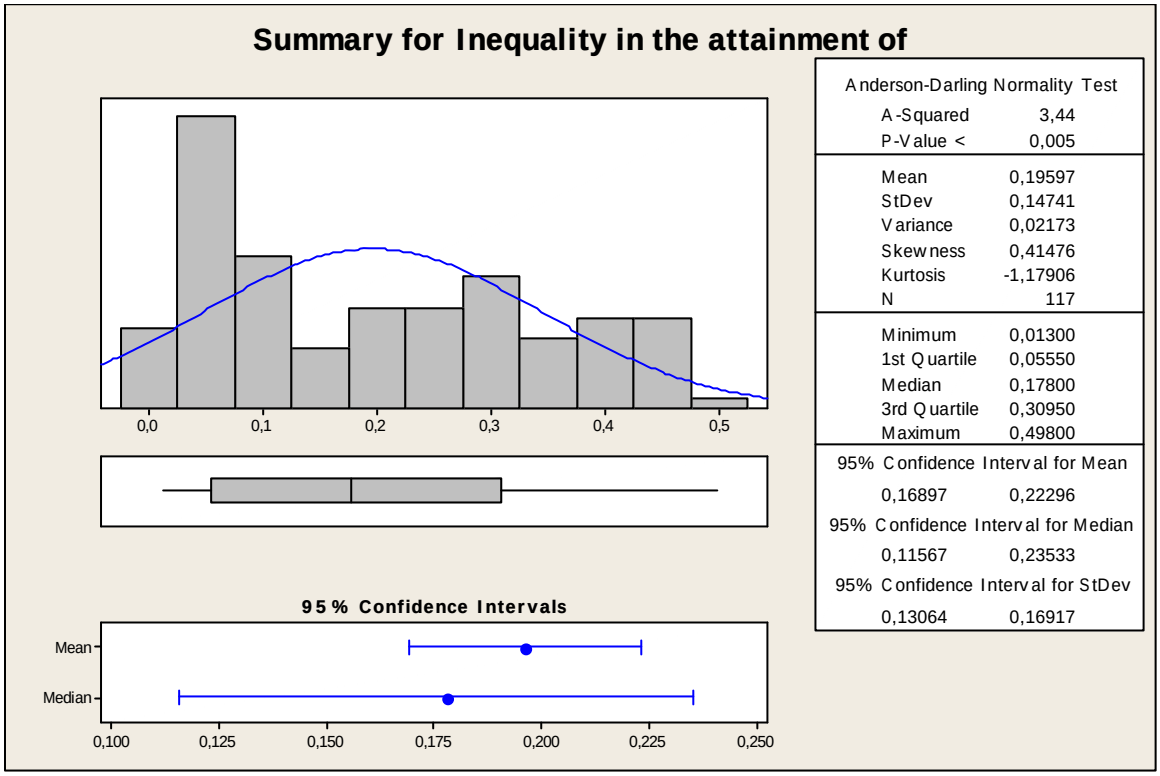
Variável: Women's average years in school Média de anos das mulheres na escola
Origem e definição dado pelo portal: http://www.gapminder.org/data/

O número médio de anos de escola frequentada por mulheres entre 25 e 34 anos de idade, incluindo primário, secundário e ensino superior.

Análise dos dados da nossa pesquisa:

O gráfico

4.26 - Inequality in the attainment of education (A desigualdade na obtenção de educação)



Variável: Inequality in the attainment of education (0=low; 1=high)

Origem e definição dado pelo portal:

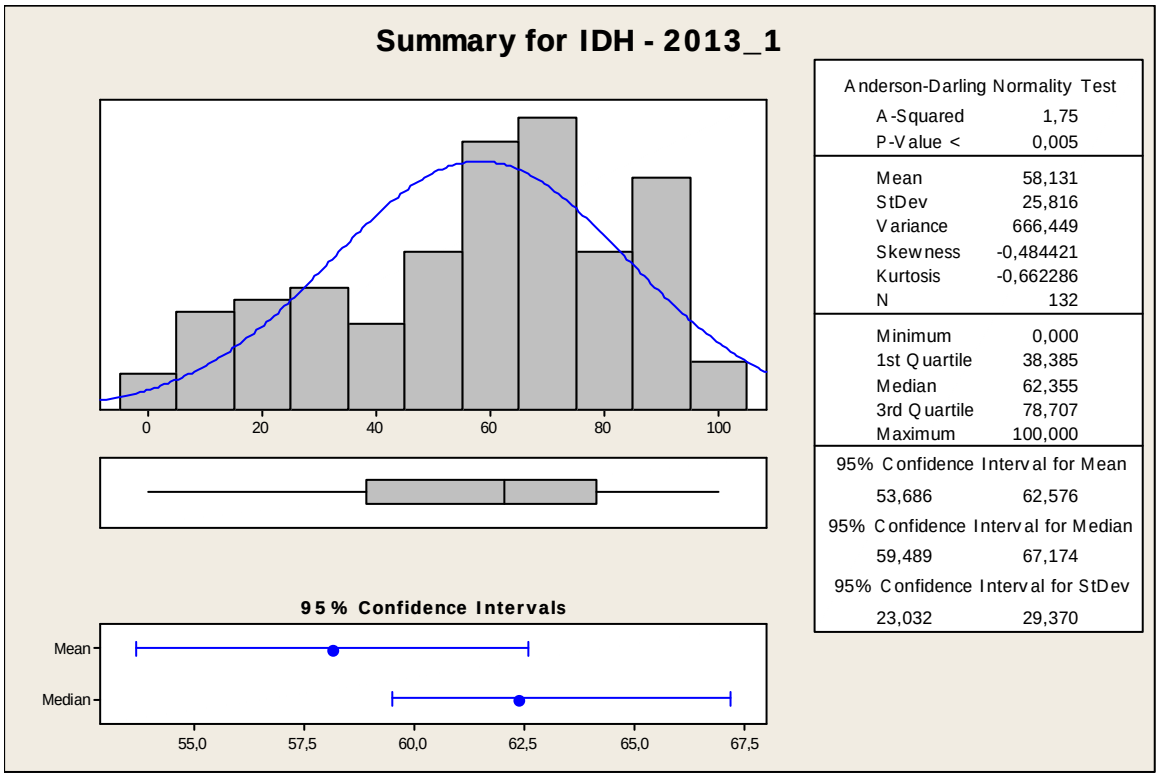
A perda na educação potencial devido à desigualdade, calculada como a diferença percentual entre o Índice de Desenvolvimento Humano Índice de Educação, que compreende média de anos de escolaridade e anos de escolaridade esperados, e o Índice de Educação Ajustado à Desigualdade.

Análise dos dados da nossa pesquisa:

4.27 IDH (2013) - Índice de Desenvolvimento Humano (IDH/PNUD)

O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.

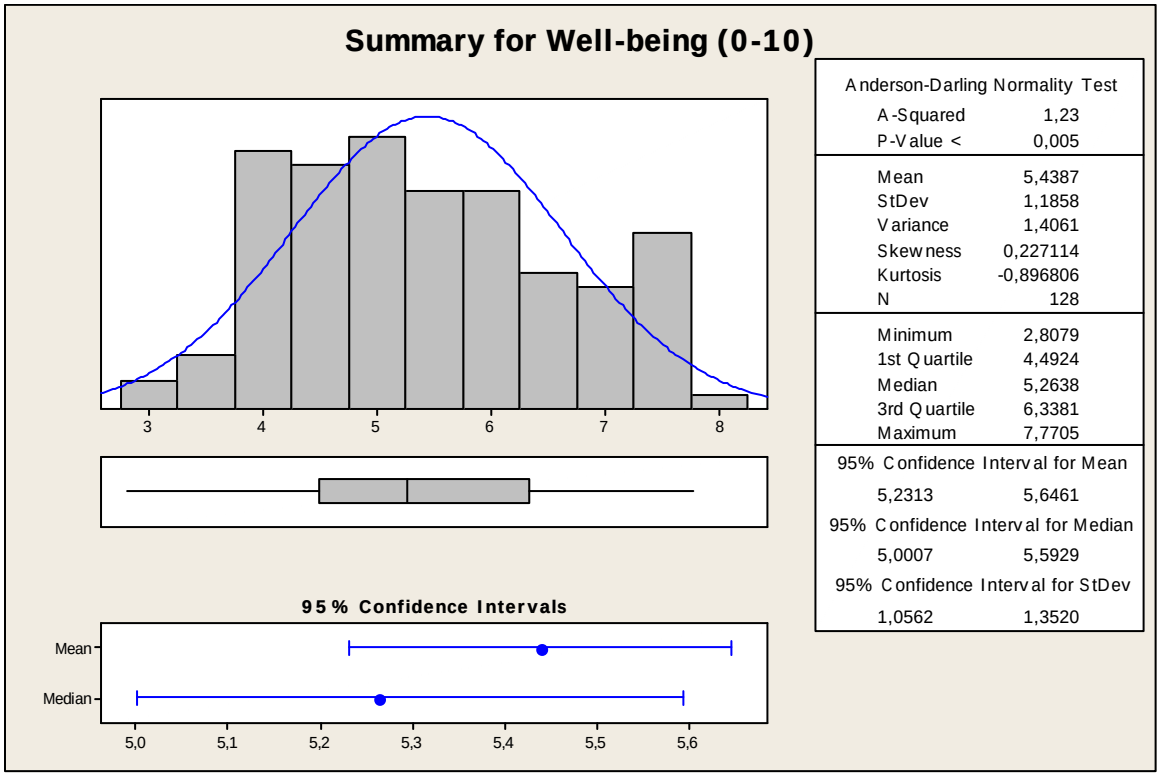
O trabalho publicado preliminar publicado por PNUD - Programa de las Naciones Unidas para el Desarrollo sobre o desenvolvimento humano em 2014 aponta que o IDH é progresso real da população em relação a expectativa de vida, educação, saúde, habitação, segurança e condições para um desenvolvimento sustentável.
http://hdr.undp.org/sites/default/files/hdr14-summary-es.pdf



O gráfico para o IDH aproxima-se de uma distribuição normal, mas deslocada para a direita, com uma concentração maior de países dos 65 aos 75 pontos, totalizando 25 países. Isso pode ser constatado também pela mediana, em 62,355. No estrato do 55 a 65 pontos, encontra-se o Paraguai, com IDH de 55,85 pontos. Do lado extremo esquerdo do gráfico estão

3 paíss no estrato de -5 e 5 pontos. Estes são considerados países de baixo desenvolvimento ou subdesenvolvidos.

4.28 - Well-being (Bem-estar)



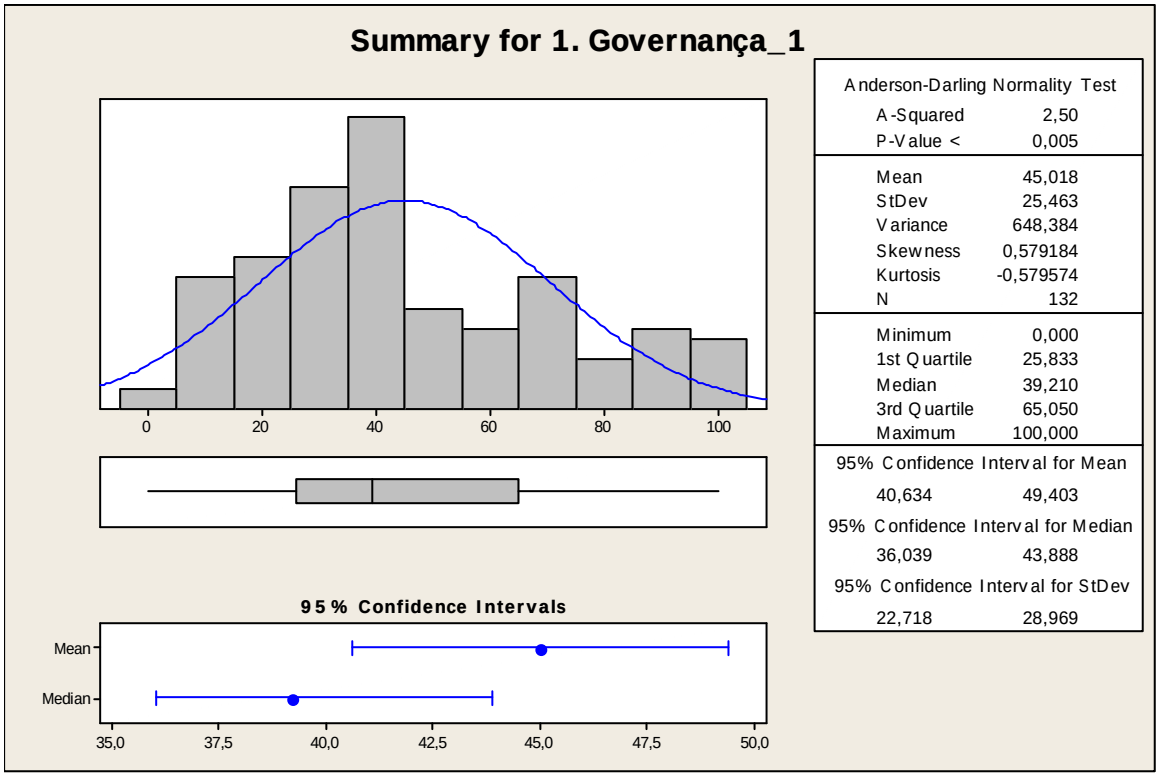
Variável: Well-being (0-10)

<u>Origem e definição dado pelo portal</u>: http://www.happyplanetindex.org/data/#table-view

Cada uma das três medidas componentes - a expectativa de vida, experimentou bem-estar e Pegada Ecológica - é atribuída uma pontuação de semáforo com base em limites para o bem (verde), médio (âmbar) e mau desempenho (vermelho). Estas pontuações são combinados a um semáforo seis cores expandida para a pontuação HPI em geral, onde, para alcançar verde brilhante - a melhor das seis cores, um país teria para um bom desempenho em todos os três componentes individuais.

<u>Análise dos dados da nossa pesquisa</u>:
O gráfico mostra uma concentração de observações no segundo e no terceiro quartil.

4.29 Índice de Governança



Origem e definição dado pelo portal:

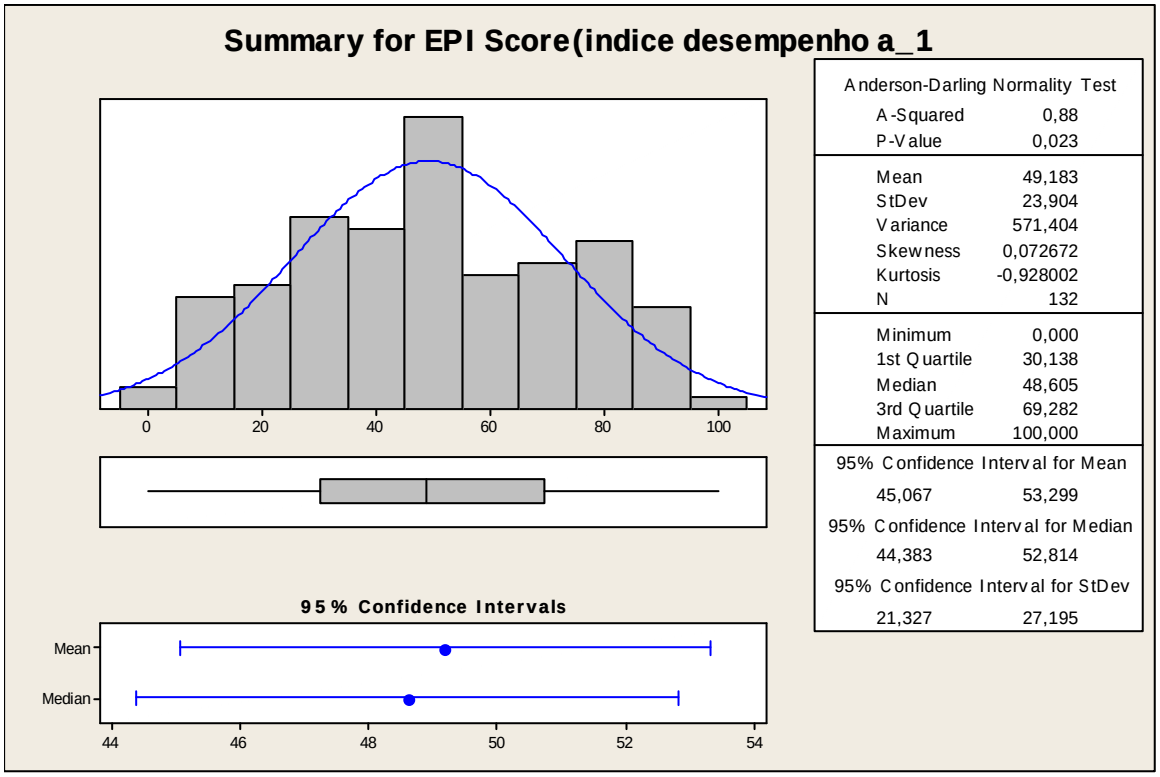
Governança: World Economic Fórum ressalta que as estratégias de atuação: Simplificar os processos administrativos e fazer uso de alianças regionais e de tecnologia para coibir a corrupção e os abusos; aprimorar os instrumentos e mecanismos regionais que possibilitem uma ação coletiva capaz de prevenir violações da democracia e garantir a preservação da ordem democrática (http://www3.weforum.org/docs/LA11/WEF_LA11_Report_PT.pdf).

Análise dos dados da nossa pesquisa:

O gráfico para a variável GOV aproxima-se de uma distribuição normal deslocada para a esquerda, o primeiro quartil de 25,833 e a mediana de 39,21 confirmam esse deslocamento, indicando uma concentração crescente de países entre as faixas de 5 a 15 (13 países), de 15 a 25 (15 países), de 25 a 35 (22 países) até chegar ao pico – a faixa 35 a 45 (29 países). Essas faixas agrupam o equivalente a 60% dos países analisados.

Em relação ao pico, os países mais bem colocados são a Macedônia (44,74), a Turquia (44,69) e o Kuwait (44,24) e na faixa inferior ficam Argentina (35,30), Benin (35,75) e Indonesia (35,96). Aparentemente trata-se de grupo bastante heterogêneo, não sendo possível estabelecer relações imediatas.

4.30 EPI - Pontuação no Índice de Proteção Ambiental



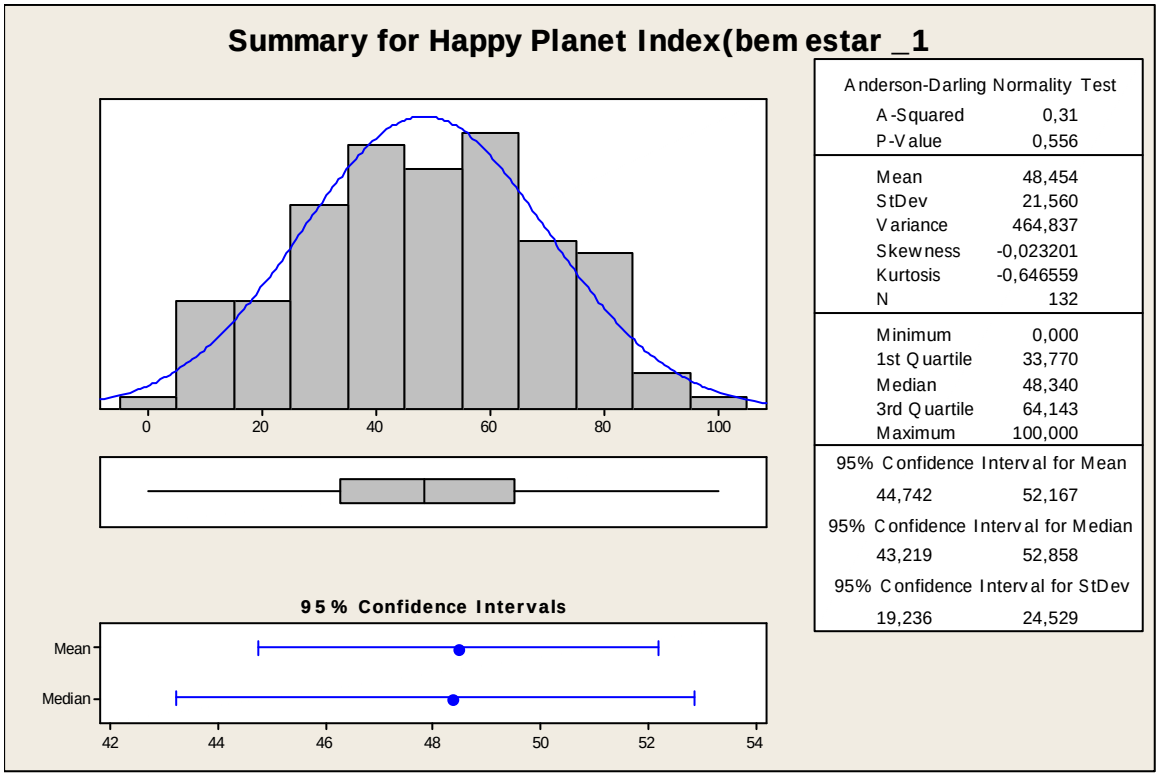
<u>Análise dos dados da nossa pesquisa</u>:

A análise dos dados do Índice de Proteção Ambiental apresenta uma distribuição normal, com média de 49,18 e mediana de 48,60, ou seja, um ligeiro deslocamento para a esquerda. Há apenas 1 país no estrato superior (nota normalizada acima de 95), que é a Suiça, com EPI de 87,67 , seguida por 9 países no segundo estrato (de 85 a 95), sendo todos países de alto IDH e IPS (Austrália, República Tcheca, Alemanha, etc) sendo o último país do estrato a Dinamarca com EPI de 76,92.

Já nos dois estratos inferiores temos a seguinte composição: no primeiro estrato (nota normalizada até 5) há dois países – Mali (EPI = 18,43) e Lesoto (EPI = 20,81) , seguidos por um grupo de 10 países no segundo estrato que vai de 5 a15. Destes 10 países, 9 encontram-se na África, sendo Bangladesh a única exceção do grupo. O país melhor colocado é Angola, com EPI de 39,93.

Já a faixa de pico, que vai de 45 a 55, é composta da seguinte maneira: na parte inferior, Algéria (EPI = 50,08) e Libano (EPI = 50,15) seguindo por um grupo heterogêneo de 26 países nos quais Jordânia (EPI =55,78) e Montenegro (EPI= 55,52).

4.31 HPI - Pontuação no Happy Planet Index.



<u>Análise dos dados da nossa pesquisa</u>:

A distribuição para o Índice de Felicidade Bruta (FIB) aproxima-se de uma normal, com maior concentração de países entre a marca do 35 até 65, que formam quase um plateau, com os três maiores picos da distribuição. Nas extremidades encontramos um pequeno número de países : Costa Rica ficou um primeiro lugar, com um FIB de 64,03, seguida por um estrato de notas normalizadas de 85 a 95, que é composto por três países : Colômbia (FIB=59,75), El Salvador (58,88) e Jamaica (58,53).

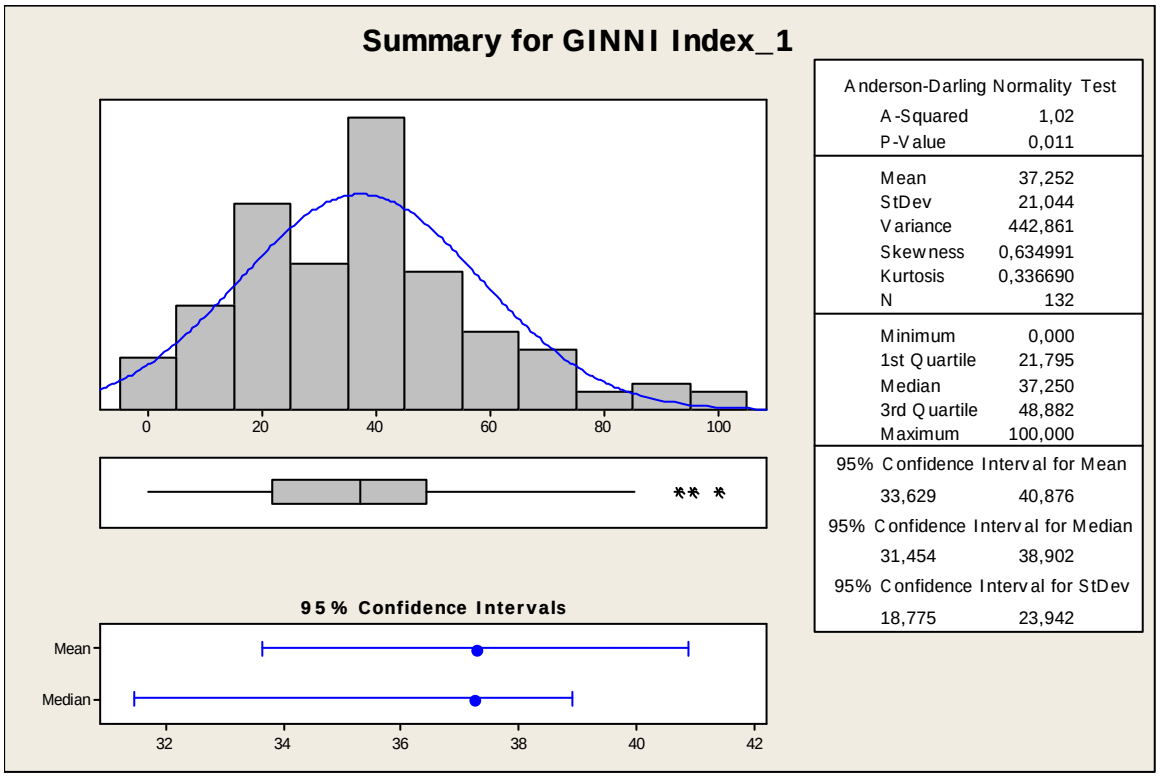
No estrato inferior, que vai até 5 pontos, figura também apenas um país :Botswana, com FIB = 22,59, seguida pelo segundo estrato, que vai de 5 a 15 pontos e compreende 9 paíse. A menor pontuação é do Chade (FIB=24,68) e a maior é da Macedônia (FIB=28,27). Chama a atenção neste grupo de países a presença da África do Sul, que figura com um FIB de 28,19, o oitavo pior colocado na classificação geral do FIB.

A maior concentração em número de países está alocada entre 35 e 65 pontos, sendo Djibouti (FIB=37,23), Estados Unidos (FIB=37,34) e Hungria (FIB=37,40) os piores colocados do grupo, e Sri Lanka (FIB=49,38), Iraque (FIB=49,19) e Laos (FIB=49,14) os melhores colocados do grupo.

O Índice de Felicidade Bruta mede aspectos subjetivos e objetivos da vida humana, evidenciando pesquisas que apontam que, a partir de certo nível de renda, o nível de felicidade médio reportado não aumenta na mesma proporção do aumento da renda, até decaindo com o tempo. Outro tópico relaciona é a questão da resiliência construída por pessoas em situação de privação em países de baixo IDH e IPS, que acabam adaptando-se às situações adversas.

Uma análise mais aprofundada poderá revelar o baixo nível de correlação deste indicados com o PIB e inclusive com o IDH e o IPS, pois sua metodologia é bastante distinta.

4.32 Índice de GINI



Análise dos dados da nossa pesquisa:

Para o Índice de Gini, que mede a desigualdade de renda, faltaram dados relativos a 13 países, que foram substituídos pela média da distribuição original, que foi de 39,02, que normalizado de 0 a 100 tornou-se 62,93, o que explica o pico registrado na coluna entre 55 e 65 – dos 33 países, na realidade apenas 20 tem sua pontuação original registrada neste intervalo.

Assim percebemos uma distribuição basicamente uniforme, com tendência para concentração de países mais a direita, como já demonstra o primeiro quartil, em 51,11.

São 6 países no estrato superior, acima de 95 pontos na escala normalizada, conforme a Tabela 13, abaixo, que como podemos ver é formada pelos antigos países socialistas

Tabela 11 – Países com menor índice de Gini

PAÍS	CÓD	REGIÃO	GINI
Ukraine	UKR	Eastern Europe	24,820
Slovenia	SVN	Southern Europe	24,870
Sweden	SWE	Northern Europe	26,080
Czech Republic	CZE	Eastern Europe	26,390
Belarus	BLR	Eastern Europe	26,460
Slovakia	SVK	Eastern Europe	26,580

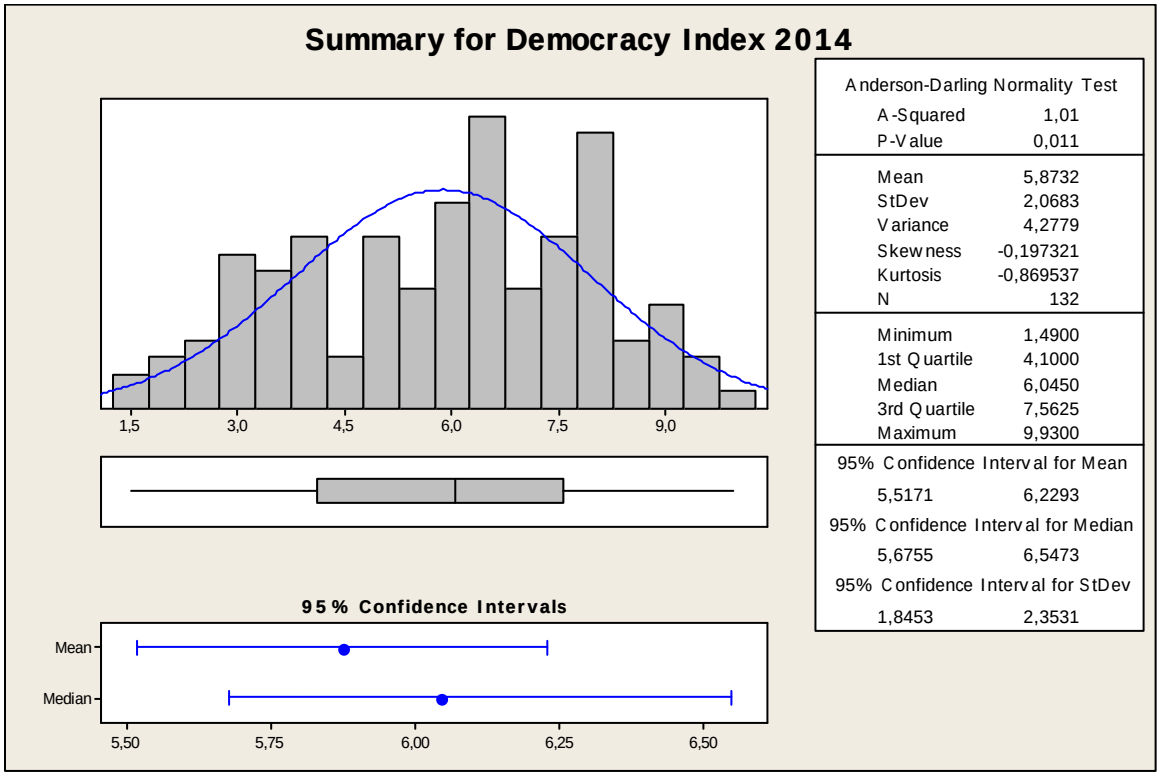
Os países com maior desigualdade de renda estão divididos nos três primeiros estratos : de 5 a 15 (2 países – África do Sul e Namíbia), de 15 a 25 (3 países – Botswana, Zâmbia e Honduras) e de 25 a 35 (2 países – República Central Africana e Lesoto).

Tabela 11 – Países de maior índice de Gini (mais desiguais)

PAÍS	CÓD	REGIÃO	GINI
Lesotho	LSO	Southern Africa	54,170
Central African Republic	CAF	Eastern Africa	56,300
Honduras	HND	IBE	57,400
Zambia	ZMB	Middle Africa	57,490
Botswana	BWA	Southern Africa	60,460
Namibia	NAM	Southern Africa	61,320
South Africa	ZAF	Southern Africa	63,140

Botswana, Namíbia e África do Sul aparecem como *outliners* no gráfico.

4.33 - Democracy Index 2014 (Índice de Democracia 2014)



Variável: Democracy Index 2014

<u>Origem e definição dado pelo portal:</u>

http://www.sudestada.com.uy/Content/Articles/421a313a-d58f-462e-9b24-2504a37f6b56/Democracy-index-2014.pdf

Índice de Democracia da Economist Intelligence Unit do fornece um apanhado instantâneo do estado da democracia em todo o mundo para 165 estados independentes e dois territórios-isto cobre quase toda a população do mundo e a grande maioria dos Estados do mundo (micro estados são excluídos). O Índice de Democracia é baseado em cinco categorias: processo eleitoral e pluralismo; liberdades civis; o funcionamento do governo; participação política; e cultura política. Com base nas suas pontuações em uma série de indicadores dentro

dessas categorias, cada país é então classificado como um dos quatro tipos de regime: "democracias plenas"; "democracias imperfeitas"; "regimes híbridos"; e "regimes autoritários".

Esta é a sétima edição do Índice de Democracia. Ele reflete a situação no final de 2014, um ano em que descontentamentos da democracia estavam em ascensão. Como tem sido o padrão nos últimos anos, houve pouca mudança na pontuação agregada global. O mesmo número de países (48) registrou uma melhoria na sua pontuação como registrado uma deterioração; o restante (71) manteve a mesma pontuação que em 2013. Três regiões experimentaram uma regressão (América Latina, Oriente Médio e África do Norte e Sub-sahara Africano) como indicado por um declínio na sua pontuação média regional. Duas regiões Ásia e Europa Oriental registraram uma ligeira melhoria em sua pontuação média. Não houve alteração na pontuação média para a América do Norte ou Europa Ocidental. Nessas duas regiões, no entanto, o descontentamento popular com a democracia foi expressa no crescimento de populistas e de protestos, que na Europa, têm vindo a representar um desafio cada vez maior para a ordem política estabelecida, na medida em que vários distúrbios políticos estão na agenda em 2015.

Análise dos dados da nossa pesquisa:

A pesquisa demonstra que o terceiro quartil possui 54 observações, ou seja, a maior parte dos países no intervalo de 5,75 a 8 nos quais estão contidas a média de 5,8 e a mediana de 6. As maiores concentrações de observações estão nos intervalos de 6,25 a 6, 75 e 7,75 a 8,25 com 17 e 16 observações respectivamente. Nota-se que as observações tanto do primeiro quartil, quanto do quarto quartil contem 22 observações. Na tabela 14 abaixo estão destacados os dez países com menor índice de democracia e os dez países com maior índice, nota-se que os países de menor índice estão na região da África e Ásia e os de maior índice na região da Europa e Austrália e Nova Zelândia.

Tabela 12 – Países com menor índice de Democracia

País	Cód.	Região	Índice de Democracia
Central African Republic	CAF	Eastern Africa	1,49
Chad	TCD	Middle Africa	1,5
Iran	IRN	Southern Asia	1,98
Laos	LAO	South-Eastern Asia	2,21
Saudi Arabia	SAU	Western Asia	1,82
Sudan	SDN	Northern Africa	2,54
Tajikistan	TJK	Central Asia	2,37
United Arab Emirates	ARE	Western Asia	2,64
Uzbekistan	UZB	Central Asia	2,45
Yemen	YEM	Western Asia	2,79

Tabela 12 – Países com maior índice de Democracia

País	Cód.	Região	Índice de Democracia
Australia	AUS	Australia and New Zealand	9,01
Canada	CAN	Northern America	9,08
Denmark	DNK	Northern Europe	9,11
Finland	FIN	Northern Europe	9,03

Iceland	ISL	Northern Europe	9,58
Netherlands	NLD	Western Europe	8,92
New Zealand	NZL	Australia and New Zealand	9,26
Norway	NOR	Northern Europe	9,93
Sweden	SWE	Northern Europe	9,73
Switzerland	CHE	Western Europe	9,09

Considerações Finais

Os índices apresentados a partir de dados de pesquisa de 132 paises que refletem o grau de evolução no quesito humano, principalmente o tema deste trabalho que é: bem-estar e qualidade de vida**.**
Depois de ter feito a apresentação dos dados pela estatística descritiva como média, desvio padrão e variância; por exemplo, é possível deduzir que os países de uma forma geral corroboram o que as análises macroeconômica de diversos organismos tais como: UNICEF, Banco Mundial, WEF – World Economica Forum, FAO e outros órgãos preocupados em identificar a evolução dos povos do globo.
Indicadores como o IDH – índice de Desenvolvimento econômico apresentado nesse trabalho com a normalização não muda a situação real da economia analisada. Por exemplo: quando se pegam os dados de países da Africa Sub-saariana, Republica centro-africana, Lesotho apresentam indicadores que mostram que as variáveis que fazem parte deste indicador como saúde, educação, uso de água potável, instalações sanitária que podem refletir a qualidade de vida estão baixos quando comparados com outros países.
Especificamente a Zambia difere dos demais principalmente quando analisamos seus dados de IDH de 1998 para 2013. Esse indicador salta de 0,41 em 1998 para 0,56 em 2013, ou seja, o indicador nos diz que quanto mais próximo de um melhor a qualidade de vida da população.

5. RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS

5.1 Correlação linear

A matriz de correlação linear apresenta o teste de significância *P-Value*. Para a correlação foi utilizado o índice de *Pearson*. Conforme Las Casas e Guevara (2010)[1], a correlação (valores entre -1 e 1) é expressada pelo grau de relação linear que existe entre elas. Valores próximos a zero significam nenhuma relação, e próximos a 1 e -1 mostram alta relação positiva ou negativa, respectivamente. Vale ressaltar que o índice de correlação entre as variáveis não requer, necessariamente, que exista uma relação de causa-efeito entre as variáveis para dar sentido às relações. A seguir, apresentamos a matriz de correlação das variáveis estudadas:

[1] LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisas de Marketing. São Paulo: Atlas, 2010.

Correlations: Social Progr; Nutrition an; Water and Sa; Personal Saf; ...

	Social Progress	Nutrition and Ba	Water and Sanita
Nutrition and Ba	0,853		
	0,000		
Water and Sanita	0,856	0,878	
	0,000	0,000	
Personal Safety	0,766	0,607	0,593
	0,000	0,000	0,000
Access to Advanc	0,879	0,792	0,791
	0,000	0,000	0,000
Undernourishment	-0,695	-0,869	-0,724
	0,000	0,000	0,000
Access to piped	0,846	0,862	0,923
	0,000	0,000	0,000
Access to improv	0,791	0,849	0,941
	0,000	0,000	0,000
Access to electr	0,769	0,908	0,865
	0,000	0,000	0,000
Quality of elect	0,786	0,686	0,670
	0,000	0,000	0,000
Press Freedom In	-0,566	-0,247	-0,252
	0,000	0,004	0,004
Life expectancy	0,855	0,904	0,850
	0,000	0,000	0,000
Non-communicable	-0,748	-0,623	-0,618
	0,000	0,000	0,000
Obesity rate (%	0,574	0,681	0,716
	0,000	0,000	0,000
Suicide rate (de	0,161	0,141	0,145
	0,065	0,106	0,097
Modern slavery,	-0,578	-0,539	-0,530
	0,000	0,000	0,000
Satisfied demand	0,743	0,789	0,746
	0,000	0,000	0,000
Women treated wi	0,162	0,084	0,113
	0,064	0,341	0,199
Tolerance for im	0,243	-0,031	0,001
	0,005	0,727	0,987
Tolerance for ho	0,721	0,499	0,546
	0,000	0,000	0,000
Discrimination a	-0,656	-0,395	-0,368
	0,000	0,000	0,000
Religious tolera	0,196	-0,001	0,005
	0,024	0,993	0,950
Community safety	0,681	0,558	0,543

	0,000	0,000	0,000
Years of tertiar	0,626	0,566	0,555
	0,000	0,000	0,000
Womens aerage ye	0,829	0,791	0,795
	0,000	0,000	0,000
Inequality in th	-0,772	-0,687	-0,718
	0,000	0,000	0,000
IDH - 2013	0,866	0,859	0,869
	0,000	0,000	0,000
Well-being (0-10	0,712	0,590	0,630
	0,000	0,000	0,000
1, Governança	0,890	0,612	0,646
	0,000	0,000	0,000
EPI Score	0,917	0,797	0,831
	0,000	0,000	0,000
Happy Planet Ind	0,296	0,395	0,368
	0,001	0,000	0,000
GINNI Index	-0,308	-0,390	-0,325
	0,000	0,000	0,000
Democracy Index	0,789	0,528	0,537
	0,000	0,000	0,000

	Personal Safety	Access to Advanc	Undernourishment
Access to Advanc	0,670		
	0,000		
Undernourishment	-0,524	-0,648	
	0,000	0,000	
Access to piped	0,557	0,746	-0,702
	0,000	0,000	0,000
Access to improv	0,551	0,764	-0,673
	0,000	0,000	0,000
Access to electr	0,479	0,736	-0,741
	0,000	0,000	0,000
Quality of elect	0,703	0,733	-0,543
	0,000	0,000	0,000
Press Freedom In	-0,446	-0,404	0,260
	0,000	0,000	0,003
Life expectancy	0,634	0,777	-0,691
	0,000	0,000	0,000
Non-communicable	-0,526	-0,567	0,481
	0,000	0,000	0,000
Obesity rate (%	0,273	0,527	-0,609
	0,002	0,000	0,000
Suicide rate (de	0,227	0,287	-0,036
	0,009	0,001	0,679
Modern slavery,	-0,364	-0,554	0,354
	0,000	0,000	0,000

Satisfied demand	0,430	0,717	-0,553
	0,000	0,000	0,000
Women treated wi	0,444	0,188	-0,149
	0,000	0,031	0,088
Tolerance for im	0,219	0,122	-0,007
	0,012	0,163	0,934
Tolerance for ho	0,477	0,607	-0,385
	0,000	0,000	0,000
Discrimination a	-0,635	-0,497	0,353
	0,000	0,000	0,000
Religious tolera	0,146	0,034	0,066
	0,095	0,700	0,456
Community safety	0,446	0,655	-0,444
	0,000	0,000	0,000
Years of tertiar	0,470	0,802	-0,491
	0,000	0,000	0,000
Womens aerage ye	0,587	0,919	-0,639
	0,000	0,000	0,000
Inequality in th	-0,631	-0,836	0,508
	0,000	0,000	0,000
IDH - 2013	0,632	0,849	-0,696
	0,000	0,000	0,000
Well-being (0-10	0,414	0,654	-0,488
	0,000	0,000	0,000
1, Governança	0,799	0,758	-0,524
	0,000	0,000	0,000
EPI Score	0,730	0,831	-0,683
	0,000	0,000	0,000
Happy Planet Ind	-0,008	0,240	-0,215
	0,926	0,005	0,013
GINNI Index	-0,534	-0,368	0,378
	0,000	0,000	0,000
Democracy Index	0,559	0,634	-0,427
	0,000	0,000	0,000

	Access to piped	Access to improv	Access to electr
Access to improv	0,856		
	0,000		
Access to electr	0,846	0,878	
	0,000	0,000	
Quality of elect	0,681	0,657	0,609
	0,000	0,000	0,000
Press Freedom In	-0,303	-0,157	-0,134
	0,000	0,071	0,125
Life expectancy	0,831	0,808	0,848
	0,000	0,000	0,000
Non-communicable	-0,684	-0,546	-0,546
	0,000	0,000	0,000

Obesity rate (%	0,731	0,704	0,685
	0,000	0,000	0,000
Suicide rate (de	0,060	0,165	0,132
	0,494	0,058	0,133
Modern slavery,	-0,574	-0,584	-0,554
	0,000	0,000	0,000
Satisfied demand	0,734	0,763	0,799
	0,000	0,000	0,000
Women treated wi	-0,038	0,119	-0,016
	0,663	0,173	0,858
Tolerance for im	0,044	-0,044	-0,108
	0,619	0,618	0,219
Tolerance for ho	0,577	0,472	0,448
	0,000	0,000	0,000
Discrimination a	-0,415	-0,314	-0,252
	0,000	0,000	0,004
Religious tolera	0,098	0,007	-0,069
	0,262	0,941	0,435
Community safety	0,541	0,517	0,487
	0,000	0,000	0,000
Years of tertiar	0,523	0,514	0,510
	0,000	0,000	0,000
Womens aerage ye	0,764	0,791	0,778
	0,000	0,000	0,000
Inequality in th	-0,648	-0,742	-0,652
	0,000	0,000	0,000
IDH - 2013	0,825	0,832	0,821
	0,000	0,000	0,000
Well-being (0-10	0,621	0,582	0,541
	0,000	0,000	0,000
1, Governança	0,633	0,547	0,480
	0,000	0,000	0,000
EPI Score	0,813	0,782	0,724
	0,000	0,000	0,000
Happy Planet Ind	0,370	0,373	0,476
	0,000	0,000	0,000
GINNI Index	-0,247	-0,346	-0,353
	0,004	0,000	0,000
Democracy Index	0,554	0,422	0,438
	0,000	0,000	0,000

	Quality of elect	Press Freedom In	Life expectancy
Press Freedom In	-0,327		
	0,000		
Life expectancy	0,701	-0,269	
	0,000	0,002	
Non-communicable	-0,610	0,410	-0,756

	0,000	0,000	0,000
Obesity rate (%	0,462	-0,168	0,554
	0,000	0,054	0,000
Suicide rate (de	0,126	-0,036	0,124
	0,151	0,679	0,157
Modern slavery,	-0,451	0,156	-0,536
	0,000	0,074	0,000
Satisfied demand	0,586	-0,177	0,751
	0,000	0,043	0,000
Women treated wi	0,287	0,049	0,109
	0,001	0,580	0,213
Tolerance for im	0,152	-0,441	0,048
	0,082	0,000	0,586
Tolerance for ho	0,531	-0,446	0,591
	0,000	0,000	0,000
Discrimination a	-0,492	0,640	-0,375
	0,000	0,000	0,000
Religious tolera	0,127	-0,301	0,000
	0,146	0,000	0,997
Community safety	0,465	-0,362	0,512
	0,000	0,000	0,000
Years of tertiar	0,501	-0,293	0,581
	0,000	0,001	0,000
Womens aerage ye	0,670	-0,356	0,729
	0,000	0,000	0,000
Inequality in th	-0,660	0,390	-0,646
	0,000	0,000	0,000
IDH - 2013	0,700	-0,296	0,818
	0,000	0,001	0,000
Well-being (0-10	0,578	-0,378	0,632
	0,000	0,000	0,000
1, Governança	0,728	-0,694	0,645
	0,000	0,000	0,000
EPI Score	0,769	-0,487	0,807
	0,000	0,000	0,000
Happy Planet Ind	0,179	0,083	0,508
	0,040	0,346	0,000
GINNI Index	-0,322	0,113	-0,424
	0,000	0,199	0,000
Democracy Index	0,482	-0,737	0,569
	0,000	0,000	0,000

	Non-communicable	Obesity rate (%	Suicide rate (de
Obesity rate (%	-0,404		
	0,000		
Suicide rate (de	0,129	-0,105	
	0,139	0,231	

Modern slavery,	0,478	-0,428	-0,056
	0,000	0,000	0,522
Satisfied demand	-0,527	0,557	0,235
	0,000	0,000	0,007
Women treated wi	-0,041	-0,106	-0,039
	0,643	0,224	0,653
Tolerance for im	-0,310	0,028	-0,168
	0,000	0,751	0,055
Tolerance for ho	-0,711	0,420	0,009
	0,000	0,000	0,921
Discrimination a	0,430	-0,279	-0,129
	0,000	0,001	0,142
Religious tolera	-0,128	0,082	0,032
	0,144	0,350	0,714
Community safety	-0,464	0,462	0,178
	0,000	0,000	0,041
Years of tertiar	-0,482	0,367	0,185
	0,000	0,000	0,033
Womens aerage ye	-0,462	0,608	0,290
	0,000	0,000	0,001
Inequality in th	0,363	-0,500	-0,325
	0,000	0,000	0,000
IDH - 2013	-0,629	0,692	0,145
	0,000	0,000	0,096
Well-being (0-10	-0,636	0,540	0,067
	0,000	0,000	0,444
1, Governança	-0,663	0,348	0,162
	0,000	0,000	0,063
EPI Score	-0,695	0,594	0,148
	0,000	0,000	0,091
Happy Planet Ind	-0,416	0,222	-0,013
	0,000	0,011	0,881
GINNI Index	0,132	-0,097	-0,275
	0,131	0,268	0,001
Democracy Index	-0,628	0,243	0,171
	0,000	0,005	0,050

	Modern slavery,	Satisfied demand	Women treated wi
Satisfied demand	-0,580		
	0,000		
Women treated wi	0,037	-0,063	
	0,672	0,472	
Tolerance for im	-0,080	-0,099	0,091
	0,360	0,256	0,301
Tolerance for ho	-0,431	0,487	0,091
	0,000	0,000	0,300
Discrimination a	0,325	-0,313	-0,137
	0,000	0,000	0,117

Religious tolera	0,003	0,045	-0,129
	0,970	0,609	0,140
Community safety	-0,409	0,501	0,054
	0,000	0,000	0,542
Years of tertiar	-0,437	0,443	0,115
	0,000	0,000	0,190
Womens aerage ye	-0,554	0,773	0,067
	0,000	0,000	0,448
Inequality in th	0,484	-0,634	-0,142
	0,000	0,000	0,103
IDH - 2013	-0,579	0,764	0,110
	0,000	0,000	0,209
Well-being (0-10	-0,441	0,542	0,074
	0,000	0,000	0,398
1, Governança	-0,437	0,514	0,283
	0,000	0,000	0,001
EPI Score	-0,553	0,683	0,205
	0,000	0,000	0,018
Happy Planet Ind	-0,291	0,502	-0,242
	0,001	0,000	0,005
GINNI Index	0,141	-0,159	-0,323
	0,106	0,069	0,000
Democracy Index	-0,372	0,511	-0,007
	0,000	0,000	0,938

	Tolerance for im	Tolerance for ho	Discrimination a
Tolerance for ho	0,500		
	0,000		
Discrimination a	-0,428	-0,559	
	0,000	0,000	
Religious tolera	0,324	0,211	-0,506
	0,000	0,015	0,000
Community safety	0,212	0,606	-0,454
	0,015	0,000	0,000
Years of tertiar	0,151	0,484	-0,332
	0,083	0,000	0,000
Womens average ye	-0,018	0,469	-0,464
	0,837	0,000	0,000
Inequality in th	-0,065	-0,443	0,481
	0,458	0,000	0,000
IDH - 2013	0,035	0,591	-0,469
	0,691	0,000	0,000
Well-being (0-10	0,314	0,723	-0,504
	0,000	0,000	0,000
1, Governança	0,330	0,678	-0,737
	0,000	0,000	0,000
EPI Score	0,202	0,644	-0,560

	0,020	0,000	0,000
Happy Planet Ind	-0,114	0,323	0,031
	0,195	0,000	0,720
GINNI Index	0,120	-0,061	0,139
	0,169	0,488	0,111
Democracy Index	0,264	0,637	-0,629
	0,002	0,000	0,000

	Religious tolera	Community safety	Years of tertiar
Community safety	0,135		
	0,124		
Years of tertiar	0,003	0,444	
	0,977	0,000	
Womens aerage ye	0,085	0,632	0,622
	0,332	0,000	0,000
Inequality in th	-0,121	-0,595	-0,485
	0,168	0,000	0,000
IDH - 2013	0,013	0,625	0,608
	0,886	0,000	0,000
Well-being (0-10	0,126	0,631	0,522
	0,151	0,000	0,000
1, Governança	0,208	0,574	0,555
	0,017	0,000	0,000
EPI Score	0,096	0,611	0,607
	0,273	0,000	0,000
Happy Planet Ind	-0,053	0,210	0,166
	0,543	0,016	0,058
GINNI Index	0,206	-0,135	-0,309
	0,018	0,122	0,000
Democracy Index	0,192	0,546	0,440
	0,028	0,000	0,000

	Womens aerage ye	Inequality in th	IDH - 2013
Inequality in th	-0,852		
	0,000		
IDH - 2013	0,842	-0,736	
	0,000	0,000	
Well-being (0-10	0,594	-0,487	0,712
	0,000	0,000	0,000
1, Governança	0,650	-0,636	0,696
	0,000	0,000	0,000
EPI Score	0,772	-0,723	0,823
	0,000	0,000	0,000
Happy Planet Ind	0,274	-0,134	0,334
	0,001	0,127	0,000
GINNI Index	-0,310	0,406	-0,294
	0,000	0,000	0,001
Democracy Index	0,559	-0,539	0,576

	0,000	0,000	0,000

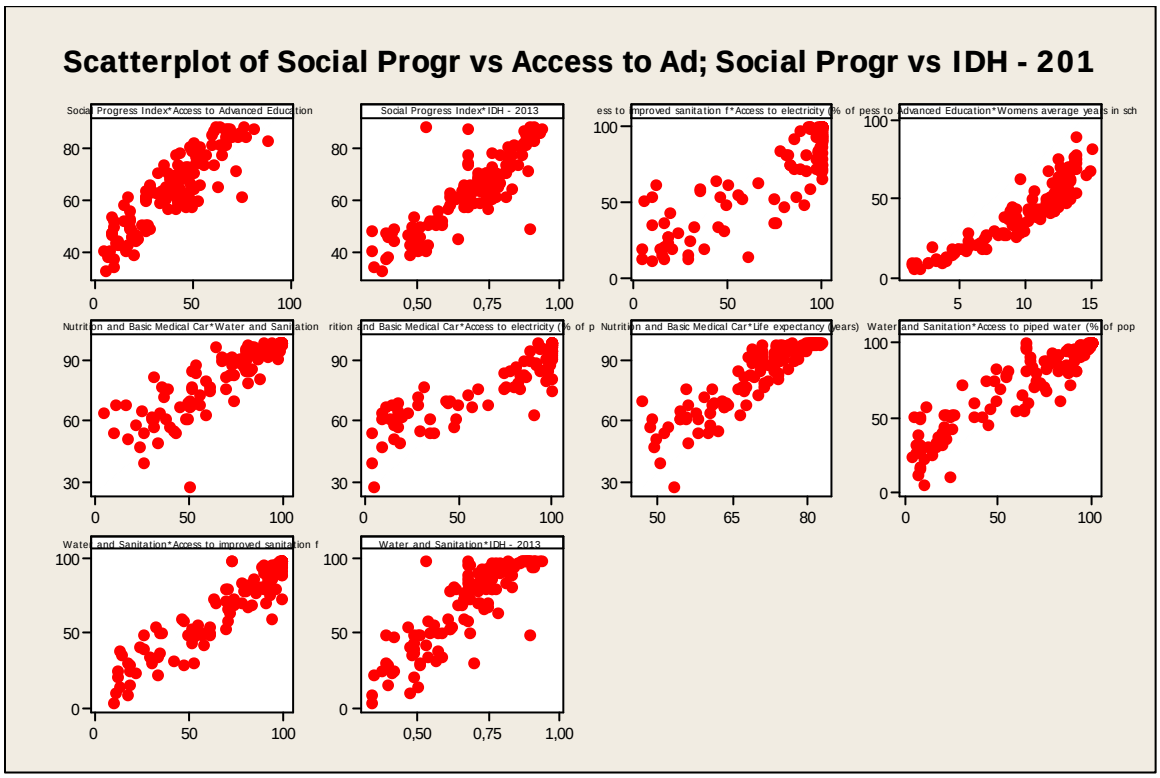
	Well-being (0-10	1, Governança	EPI Score
1, Governança	0,625		
	0,000		
EPI Score	0,639	0,811	
	0,000	0,000	
Happy Planet Ind	0,473	0,089	0,197
	0,000	0,312	0,024
GINNI Index	-0,096	-0,247	-0,365
	0,272	0,004	0,000
Democracy Index	0,554	0,849	0,633
	0,000	0,000	0,000

	Happy Planet Ind	GINNI Index
GINNI Index	0,051	
	0,562	
Democracy Index	0,231	-0,118
	0,008	0,177

Cell Contents: Pearson correlation
P-Value

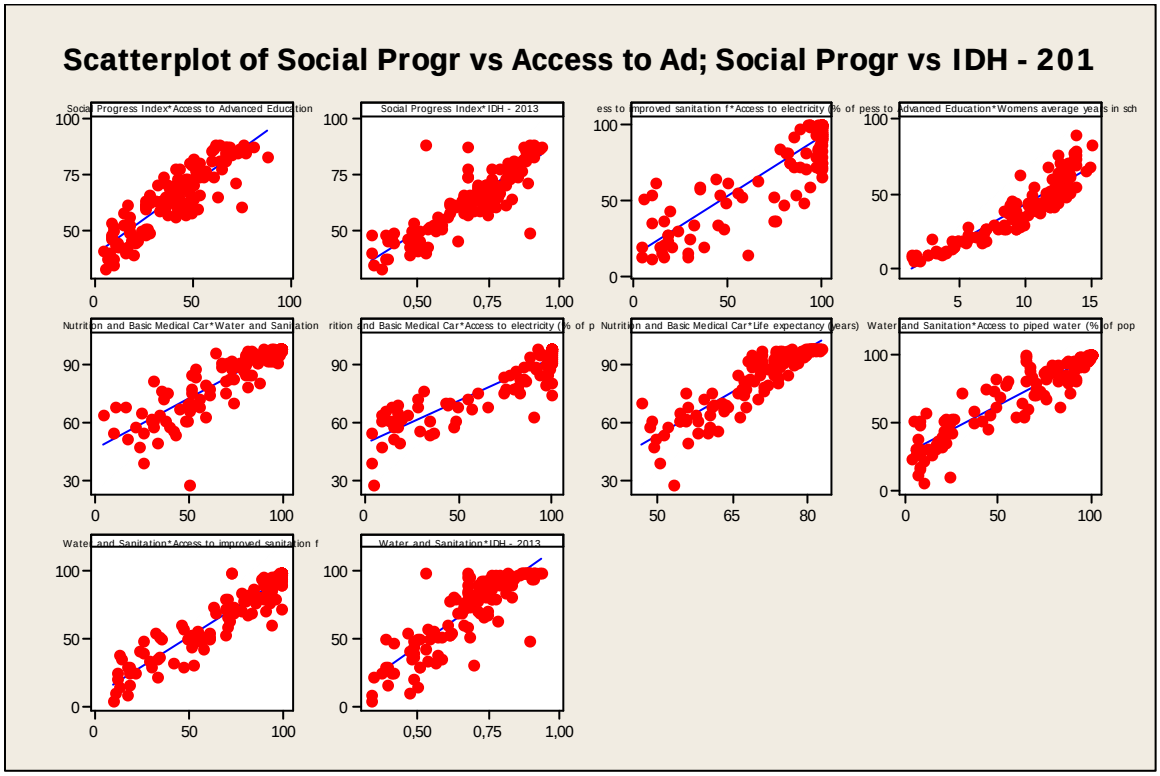
5.2 Gráficos de dispersão

Para efeito de nálises relação entre as vairáveis foram selecionadas aquelas com grau de correlação acima de 0,860.

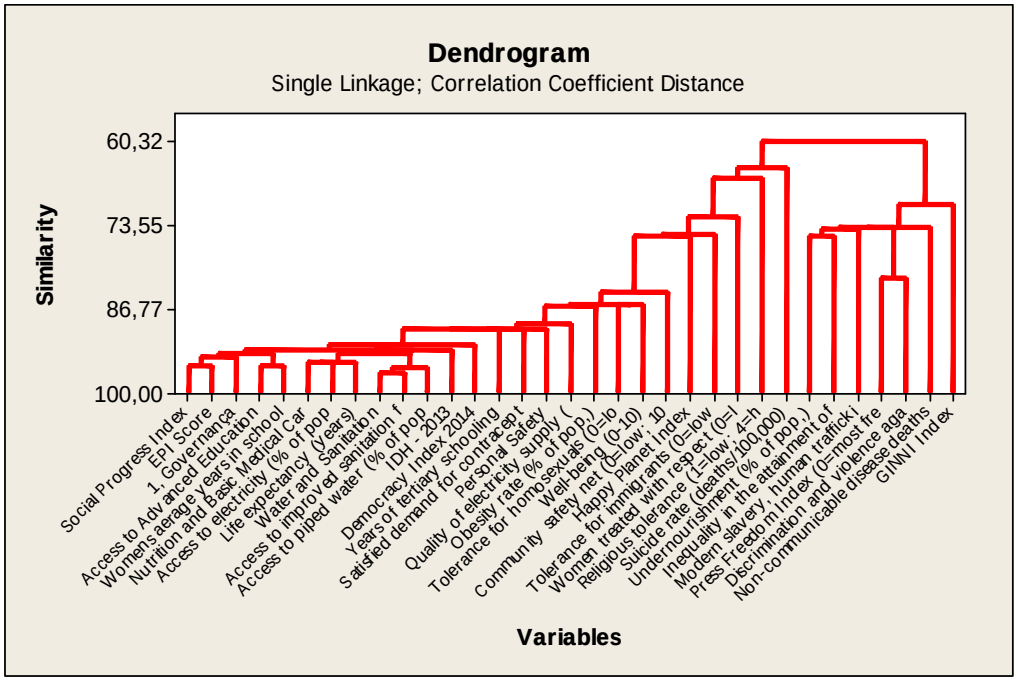


Pode-se perceber a relação entre as correlação linear positiva entre as variáveis com pouca dispersão.

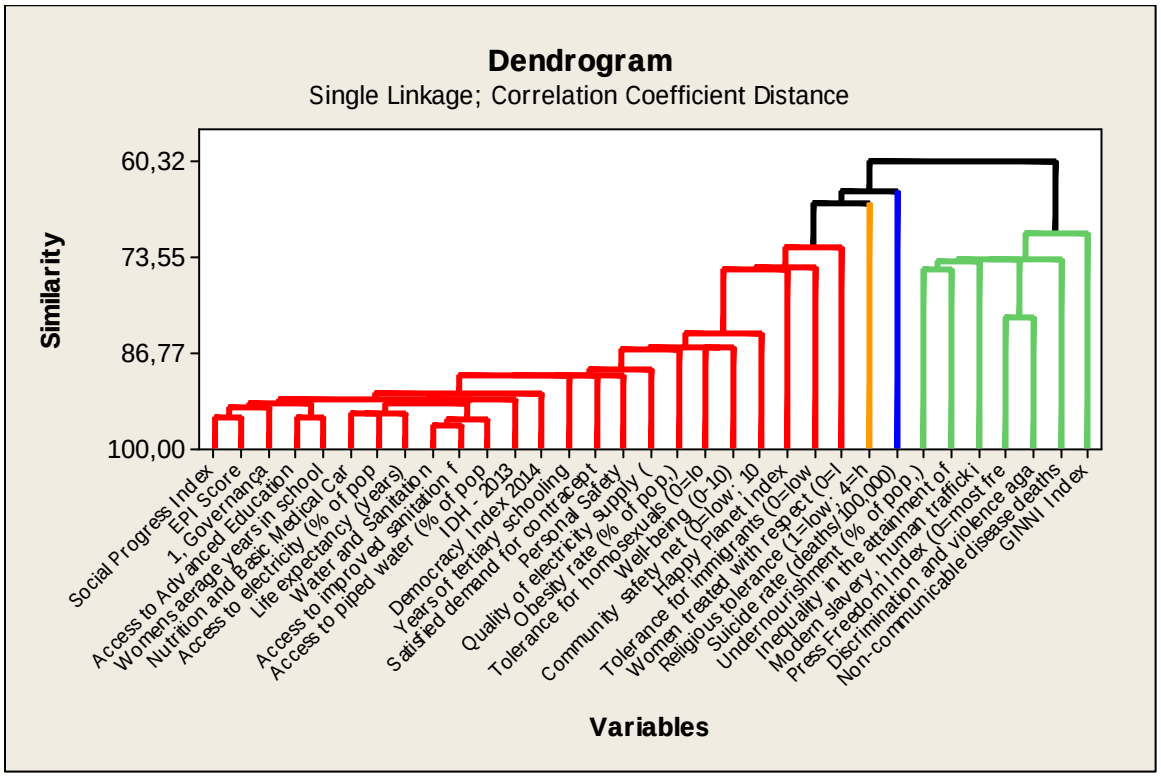
5.3 Linhas de tendência



5.3 Dendograma



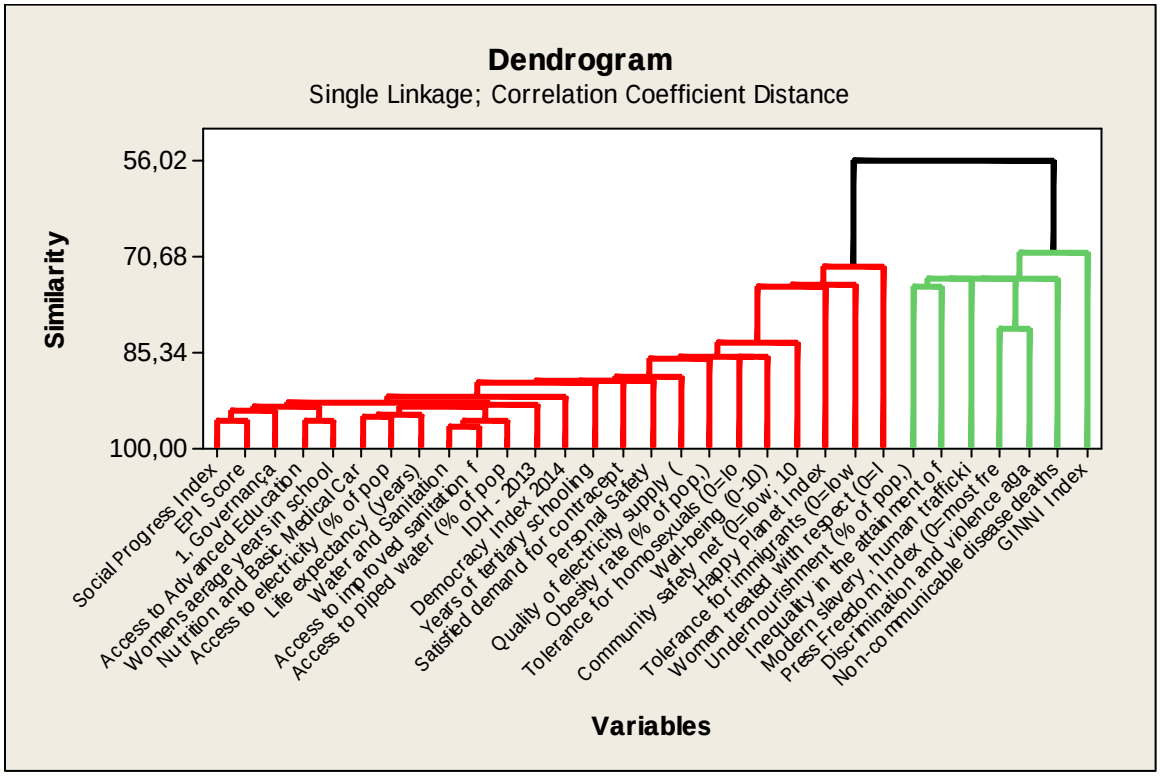
O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação. Assim, verifica-se que As variáves Social Progress até Democracy Index 2014 apresentam alto grau de similaridade.



Separando as variáveis em *clusters*, como visualiza-se no dendograma acima, e possível classificar as variáveis em 4 grupos de acordo com as suas similaridades, no entanto observa-se um grande grupo com 24 variáveis que vai de Social Progress Index até Women treated with respect, um segundo grupo com 7 variáveis que vai do Undernourishment a Gini Index, e

dois grupos com apenas uma variável um com a variável Religious Tolerance e outro com a variável Suicide Rate, observado no gráfico a baixo.

Fica claro que as variáveis Religiouns Tolerance e Suicide rate, não apresentam um nível desejado de correlação com as demais variáveis, não oferendo assim, um poder explicativo.



Ao solicitar apenas dois agrupamentos no dendograma, verifica-se que as duas variáveis que tinham ficado isoladas passam a fazer parte do grande grupo composto agora por 26 variáveis que vai do Social Progress Index até o Women Treated With Respect, o segundo grupo continua com 7 variáveis.

6. REGRESSÃO

No início deste trabalho alguns países tiveram que ter variáveis preenchidas com a média geral por causa da falta de informação de tais dados, para poder dar continuidade nas analises. A partir do momento em que foi possível analisar a dispersão das variáveis e gerar o dendograma procedeu-se em seguida com a regressão, e agora é possível verificar a similaridade entre as variáveis para obter as médias mais próximas da realidade.

Utilizando ainda o programa MINITAB foi feita a regressão com critérios e obtiveram-se os resultados descriminados abaixo.

Regression Analysis: IDH - 2013 versus Well-being (; Quality of e; ...

```
The regression equation is
IDH - 2013 = 0,0350 + 0,0282 Well-being (0-10)
             + 0,00344 Quality of electricity supply (
             + 0,00468 Obesity rate (% of pop,)
             + 0,00246 Satisfied demand for contracept
             - 0,000613 Tolerance for homosexuals (0=lo
             + 0,000680 Community safety net (0=low; 10
             + 0,0558 Years of tertiary schooling + 0,00214 Personal Safety
             + 0,00448 Democracy Index 2014


Predictor                              Coef    SE Coef      T      P
Constant                            0,03498    0,04854   0,72  0,473
Well-being (0-10)                  0,028243   0,008298   3,40  0,001
Quality of electricity supply (    0,003441   0,006331   0,54  0,588
Obesity rate (% of pop,)          0,0046778  0,0008040   5,82  0,000
Satisfied demand for contracept   0,0024605  0,0004276   5,75  0,000
Tolerance for homosexuals (0=lo  -0,0006133  0,0003970  -1,54  0,125
Community safety net (0=low; 10   0,0006804  0,0006379   1,07  0,288
Years of tertiary schooling         0,05580    0,02047   2,73  0,007
Personal Safety                   0,0021351  0,0004926   4,33  0,000
Democracy Index 2014               0,004481   0,004128   1,09  0,280


S = 0,0647227   R-Sq = 84,1%   R-Sq(adj) = 82,9%


Analysis of Variance

Source           DF       SS       MS      F      P
Regression        9  2,70579  0,30064  71,77  0,000
Residual Error  122  0,51106  0,00419
Total           131  3,21685


Source                            DF   Seq SS
Well-being (0-10)                  1  1,62923
Quality of electricity supply (    1  0,40171
Obesity rate (% of pop,)           1  0,28074
Satisfied demand for contracept    1  0,21010
Tolerance for homosexuals (0=lo    1  0,00056
Community safety net (0=low; 10    1  0,02459
Years of tertiary schooling        1  0,05531
Personal Safety                    1  0,09860
Democracy Index 2014               1  0,00494


Unusual Observations

     Well-being
Obs      (0-10)  IDH - 2013      Fit   SE Fit  Residual  St Resid
 28        3,98     0,33800  0,51174  0,01586  -0,17374     -2,77R
 64        6,12     0,68000  0,81585  0,02221  -0,13585     -2,23R
 70        5,43     0,48600  0,63002  0,01249  -0,14402     -2,27R
 88        7,50     0,68000  0,84519  0,01731  -0,16519     -2,65R
 90        5,69     0,61400  0,74545  0,01515  -0,13145     -2,09R
113        7,52     0,89800  0,74861  0,02390   0,14939      2,48R
115        5,43     0,53000  0,80535  0,01783  -0,27535     -4,43R

R denotes an observation with a large standardized residual.
```

Na analise realizada no IDH observa-se que as variáveis Personal Safety, Obesity Rate e Satisfied Demand for Contracept tem o *P-Value* 0,000 que demostra a confiabilidade, e apresentam R-Sq considerável para haver um grau de explicação e distinguir uma média entre a variável analisada. A variável Well- being deu 0,001.

Se eliminarmos as variáveis com P-Value maior que 10%, serão eliminadas Democrecy Index 2014; Quality of eletricity; Tolerance for Homossexual; e Community Safety. A seguir estão os dados da nova regressão rodada sem essas quatro variáveis. Sendo assim temos na analise realizada no IDH as variáveis: Personal Safety; Obesity Rate; Satisfied Demand for Contracept; e Well-being com o P-Value 0,000. Avariável Years of tertiary schooling ficu com P-Value 0,006, para a qual foram sugeridos alguns ajustes.

Regression Analysis: IDH - 2013 versus Years of ter; Satisfied de; ...

The regression equation is
IDH - 2013 = 0,0864 + 0,0567 Years of tertiary schooling
+ 0,00265 Satisfied demand for contracept + 0,00239 Personal Safety
+ 0,00463 Obesity rate (% of pop,) + 0,0279 Well-being (0-10)

Predictor	Coef	SE Coef	T	P
Constant	0,08635	0,03290	2,62	0,010
Years of tertiary schooling	0,05665	0,02030	2,79	0,006
Satisfied demand for contracept	0,0026475	0,0003963	6,68	0,000
Personal Safety	0,0023864	0,0003770	6,33	0,000
Obesity rate (% of pop,)	0,0046302	0,0007683	6,03	0,000
Well-being (0-10)	0,027856	0,006587	4,23	0,000

S = 0,0647046 R-Sq = 83,6% R-Sq(adj) = 83,0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	2,68933	0,53787	128,47	0,000
Residual Error	126	0,52752	0,00419		
Total	131	3,21685			

Source	DF	Seq SS
Years of tertiary schooling	1	1,18867
Satisfied demand for contracept	1	0,97796
Personal Safety	1	0,19833
Obesity rate (% of pop,)	1	0,24950
Well-being (0-10)	1	0,07486

Unusual Observations

Obs	Years of tertiary schooling	IDH - 2013	Fit	SE Fit	Residual	St Resid
28	0,09	0,33800	0,52090	0,01150	-0,18290	-2,87R
60	1,21	0,89000	0,76574	0,01986	0,12426	2,02R

64	1,13	0,68000	0,80488	0,01790	-0,12488	-2,01R
70	0,04	0,48600	0,61713	0,00969	-0,13113	-2,05R
88	0,82	0,68000	0,85714	0,01455	-0,17714	-2,81R
90	0,52	0,61400	0,76259	0,00773	-0,14859	-2,31R
102	1,58	0,77800	0,79526	0,02655	-0,01726	-0,29 X
113	0,14	0,89800	0,75049	0,01712	0,14751	2,36R
115	0,62	0,53000	0,80690	0,01323	-0,27690	-4,37R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Stepwise Regression: IDH - 2013 versus Personal Saf; Obesity rate; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is IDH - 2013 on 5 predictors, with N = 132

Step	1	2	3	4
Constant	0,2218564	0,0536968	0,0008744	0,0524223
Satisfied demand for contracept	0,00626	0,00439	0,00368	0,00278
T-Value	13,48	9,46	8,69	6,88
P-Value	0,000	0,000	0,000	0,000
Well-being (0-10)		0,0566	0,0464	0,0331
T-Value		7,46	6,75	5,12
P-Value		0,000	0,000	0,000
Personal Safety			0,00260	0,00269
T-Value			6,21	7,25
P-Value			0,000	0,000
Obesity rate (% of pop,)				0,00474
T-Value				6,02
P-Value				0,000
Years of tertiary schooling				
T-Value				
P-Value				
S	0,102	0,0852	0,0750	0,0664
R-Sq	58,29	70,87	77,62	82,59
R-Sq(adj)	57,97	70,42	77,09	82,04
Mallows Cp	192,5	97,8	48,0	11,8

Step	5
Constant	0,0863501
Satisfied demand for contracept	0,00265
T-Value	6,68
P-Value	0,000
Well-being (0-10)	0,0279
T-Value	4,23
P-Value	0,000
Personal Safety	0,00239
T-Value	6,33
P-Value	0,000
Obesity rate (% of pop,)	0,00463
T-Value	6,03
P-Value	0,000

Years of tertiary schooling	0,057
T-Value	2,79
P-Value	0,006
S	0,0647
R-Sq	83,60
R-Sq(adj)	82,95
Mallows Cp	6,0

Após a regressão obteve-se para a variável Years of tertiary schooling a média de nove países, sendo eles: República do Congo, Japão, República da Corrêa, Lesoto, Países Baixos, Nicarágua, Rússia, e Suíça. Os países mencionados tiveram a média alterada, pois a análise de regressão apresentaram as médias coerentes.

CONCLUSÃO

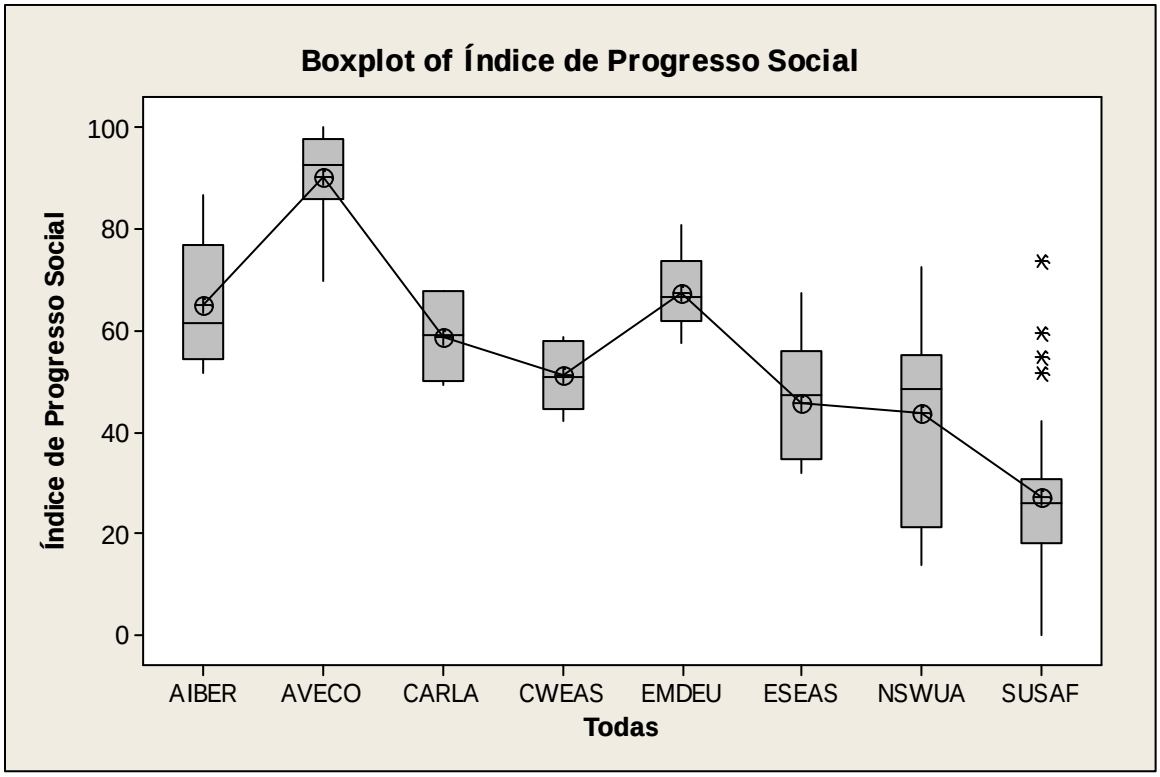
Conclui-se que os indicadores das variáveis quantitativas analisadas vem agregar valor para a transformação de como se calcula os resultados econômicos e de desenvolvimento dos países. O índice mostra uma correlação positiva entre o Índice de Desenvolvimento Humano com a variável Bem Estar, o que é bastante coerente para este trabalho que visa estudar o tema Bem Estar e Qualidade de Vida. Houve também uma correlação positiva com as variáveis relacionadas a segurança pessoal, taxa de obesidade e satisfação de demandas contraceptivas, questões bastante centrais para a qualidade de vida e para o bem estar das pessoas.

7. BOX-PLOT

Segue abaixo os gráficos Box-Plot considerando todas as regiões para cada uma das variáveis. Eles apenas corroboram visualmente o que o P-value e as demais análises acima já nos indicavam.

1.1. ANÁLISES

1.1.1. Índice de Progresso Social



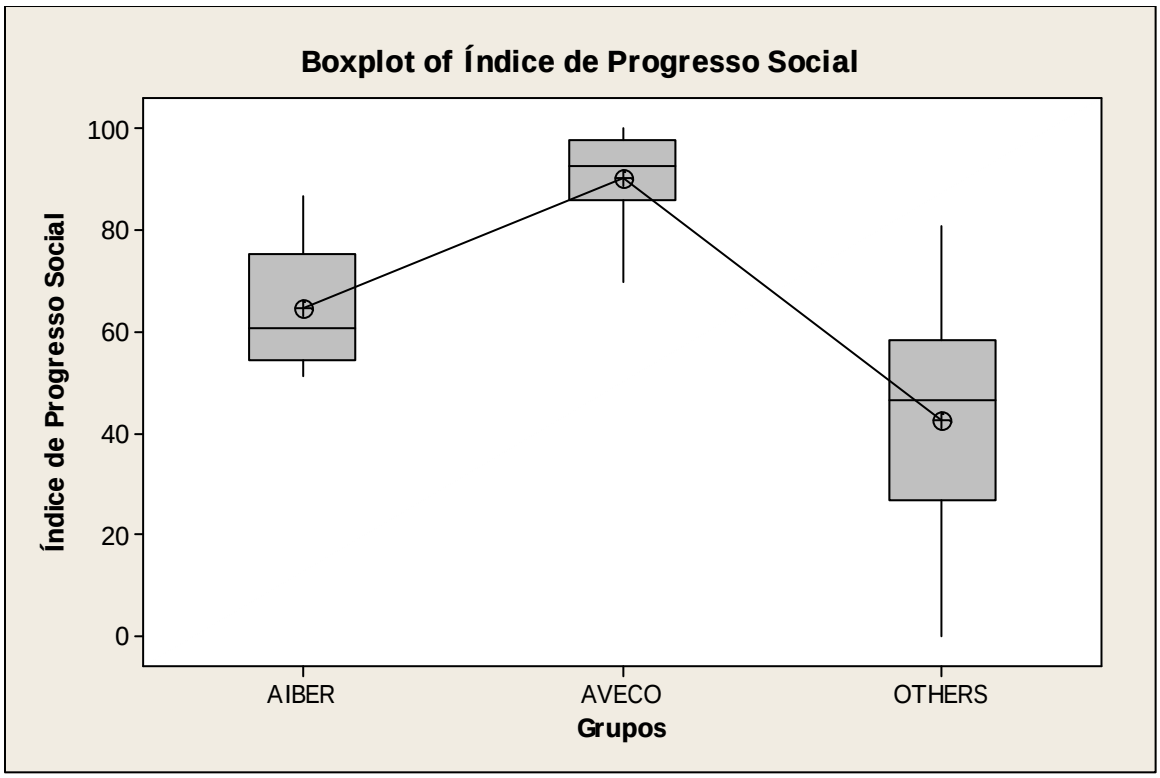
One-way ANOVA: Índice de Progresso Social versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  65477  9354  58,42  0,000
Error   124  19856   160
Total   131  85333

S = 12,65   R-Sq = 76,73%   R-Sq(adj) = 75,42%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  20  65,07  11,86                    (--*-)
AVECO  27  90,30   8,70                              (-*--)
CARLA   4  58,86   9,96             (------*------)
CWEAS  11  51,27   6,12            (---*--)
EMDEU  12  67,34   7,04                     (---*--)
ESEAS  12  45,57  11,72          (---*--)
NSWUA  14  43,75  18,74          (--*--)
SUSAF  32  27,04  15,99  (--*-)
                          ---------+---------+---------+---------+
                                  40        60        80       100

Pooled StDev = 12,65
```

Boxplot of Índice de Progresso Social

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica na maior parte dos agrupamentos, com exceção das regiões AVECO, CARLA E CWEAS. A AVECO apresenta o maior percentual do índice de progresso social. A SUSAF é a região com menor índice de progresso social, é a única região que apresenta outliers poisi quatro países possuem o índice de projeto social a cima de 50, bastante distantes da média de 27 e da mediana de 26,

que estão bastante próximas, esses países são: Tanzania (51,3) Togo (54,5), Uganda (59,2) e Zambia (73,8). Nesre gráfico o agrupamento CARLA mostra um maior intervalo deconfiança, os menores intervalos de confiança estão nos agrupamentos AIBER, AVECO e SUSAF.



One-way ANOVA: Índice de Progresso Social versus Grupos

Source	DF	SS	MS	F	P
Grupos	2	48262	24131	83,97	0,000
Error	129	37071	287		
Total	131	85333			

S = 16,95 R-Sq = 56,56% R-Sq(adj) = 55,88%

Level	N	Mean	StDev
AIBER	21	64,41	11,95
AVECO	27	90,30	8,70
OTHERS	84	42,63	19,71

Individual 95% CIs For Mean Based on Pooled StDev

```
----+---------+---------+---------+-----
        (----*----)
                                (---*----)
(-*--)
----+---------+---------+---------+-----
   45        60        75        90
```

Pooled StDev = 16,95

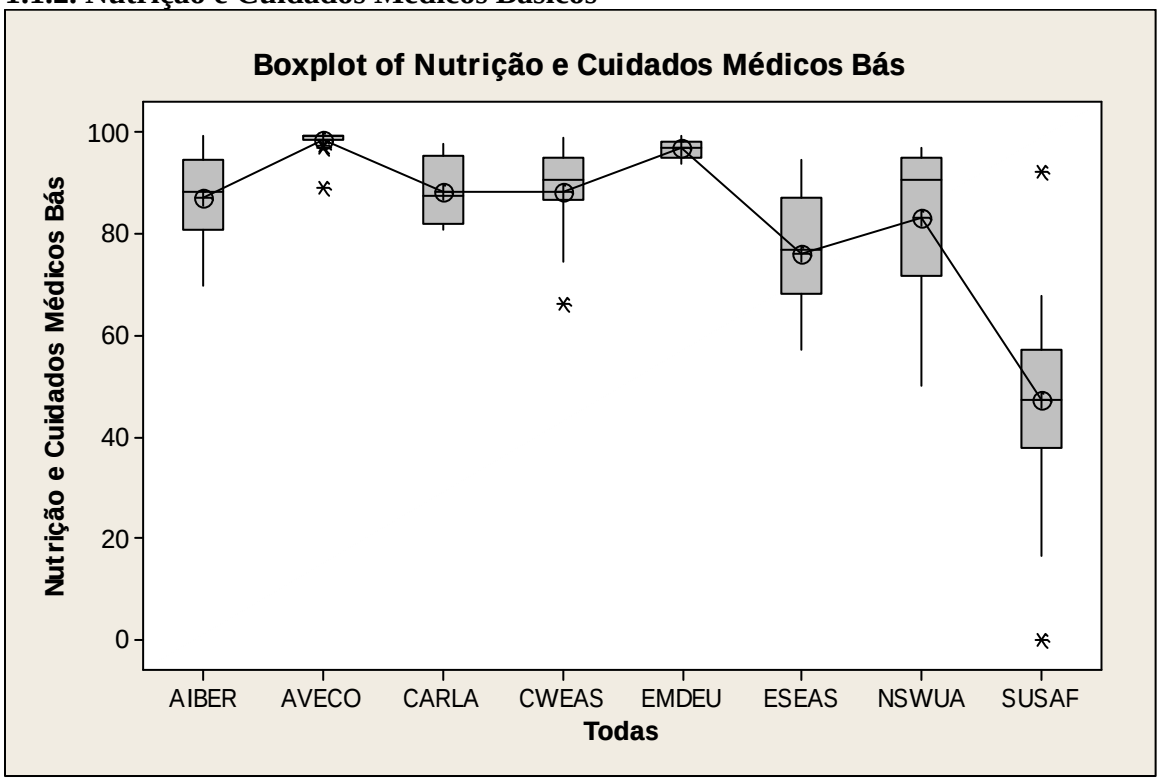
Boxplot of Índice de Progresso Social

Análise do gráfico: O agrupamento AVECO possui um maior índice de progresso social com média 90 estando a abaixo da mediana 92,7 apresentando dados assimétricos, assim

como os demais agrupamentos. O agrupamento AIBER está melhor posicionado que os demais, com uma média de 64,4 acima da mediana de 60,7. Os o agrupamento OTHERS possui uma média de 42,6, estando abaixo da mediana de 46.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto por os três grupos AVECO, AIBER e OTHERS.

1.1.2. Nutrição e Cuidados Médicos Básicos



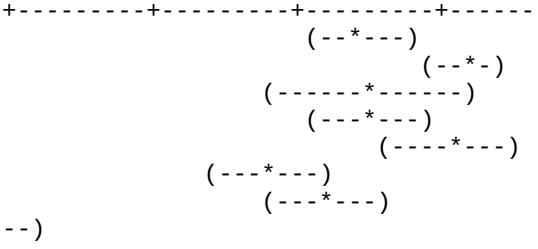
One-way ANOVA: Nutrição e Cuidados Médicos Bás versus Todas

Source	DF	SS	MS	F	P
Todas	7	49517	7074	55,49	0,000
Error	124	15806	127		
Total	131	65324			

S = 11,29 R-Sq = 75,80% R-Sq(adj) = 74,44%

Level	N	Mean	StDev
AIBER	20	87,09	8,83
AVECO	27	98,65	2,10
CARLA	4	88,36	7,05
CWEAS	11	88,19	9,65
EMDEU	12	96,83	1,70
ESEAS	12	76,21	11,63
NSWUA	14	83,25	15,83
SUSAF	32	47,20	16,42

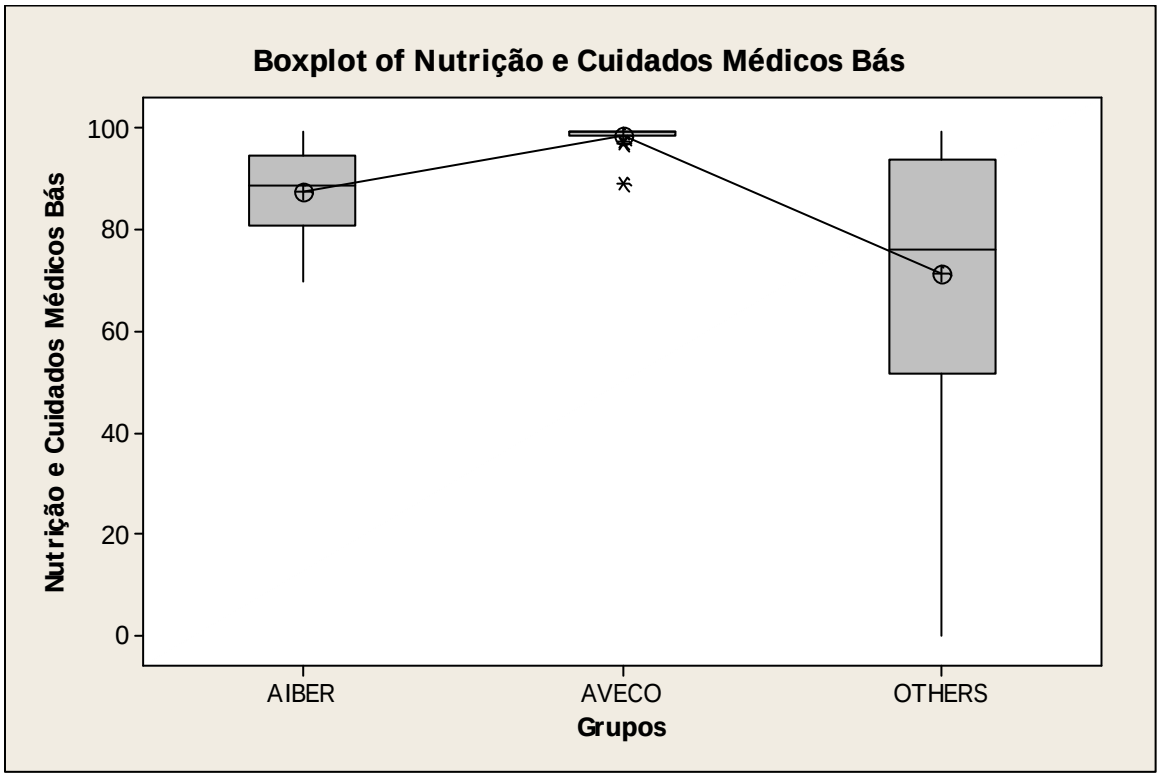
Individual 95% CIs For Mean Based on Pooled StDev



```
                    ---+---------+---------+---------+------
                      48        64        80        96

Pooled StDev = 11,29
```

Boxplot of Nutrição e Cuidados Médicos Bás

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica na maior parte dos agrupamentos. Os agrupamentos que apresentam o melhor percentual de nutrição e cuidados médicos básicos são o AVECO com média 98 e o EMDEU com média 96. A SUSAF é a região com menor índice de progresso social, apresentando dois outliers um para cima o Mauritius com 92,3 e outro para baixo o Burundi com 0, bastante diatantes da média de 47,3. O AVECO também possui um outlier para baixo, no entanto, ainda sim possui um alto percentual de 88,9 apenas dez ponto distante da média do agrupamento. O CWEAS também possui um outlier, com uma distância de 66 da média 88,1 do agrupamento. Nesre gráfico o agrupamento CARLA mostra um maior intervalo deconfiança, os menores intervalos de confiança estão nos agrupamentos AVECO e SUSAF.



One-way ANOVA: Nutrição e Cuidados Médicos Bás versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  17102  8551  22,88  0,000
Error   129  48221   374
Total   131  65324

S = 19,33   R-Sq = 26,18%   R-Sq(adj) = 25,04%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
```

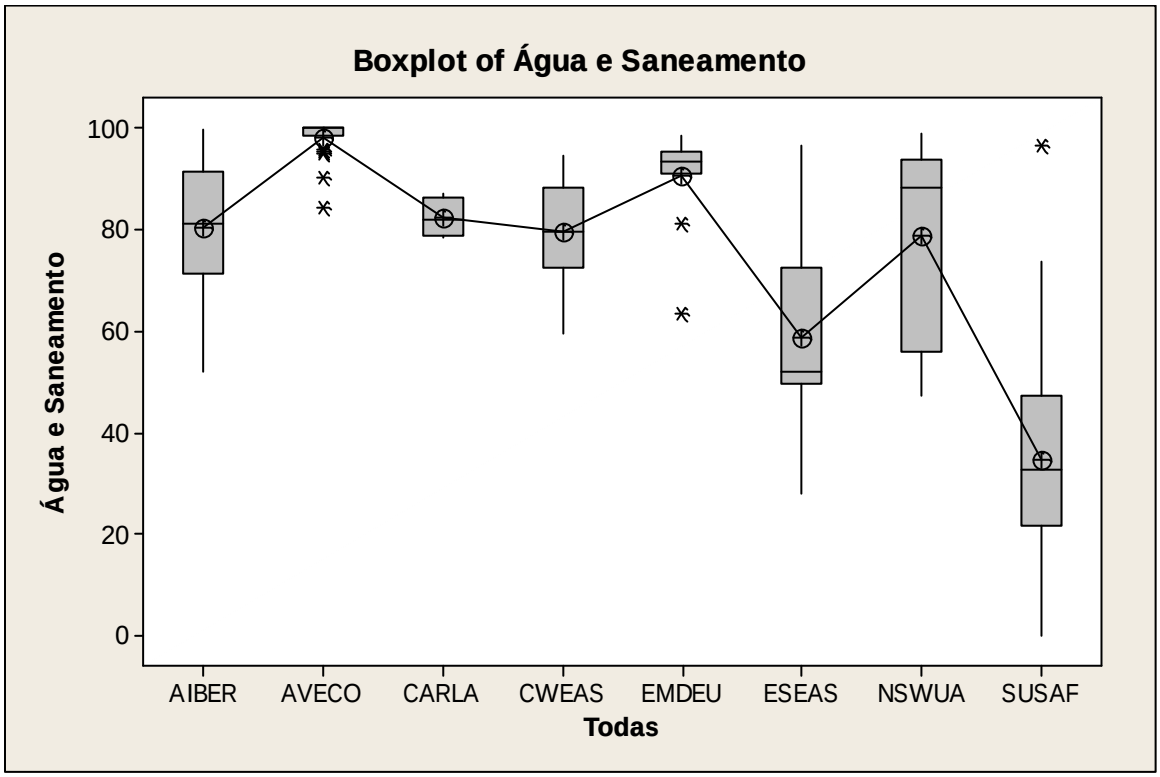
```
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  87,59   8,91              (--------*-------)
AVECO   27  98,65   2,10                        (-------*------)
OTHERS  84  71,17  23,67  (---*---)
                          ---+---------+---------+---------+------
                            70        80        90       100

Pooled StDev = 19,33
```

Boxplot of Nutrição e Cuidados Médicos Bás

Análise do gráfico: O agrupamento AVECO possui um maior índice de nutrição e cuidados médicos básicos com média 98,6 estando próximo da mediana 99,2 apresentando dados assimétricos, assim como os demais agrupamentos. O agrupamento AIBER está melhor posicionado que os demais, com uma média de 87,5. Os o agrupamento OTHERS possui uma média de 71,1.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto por por todos os agrupamentos.

1.1.3. Água tratada e Saneamento Básico



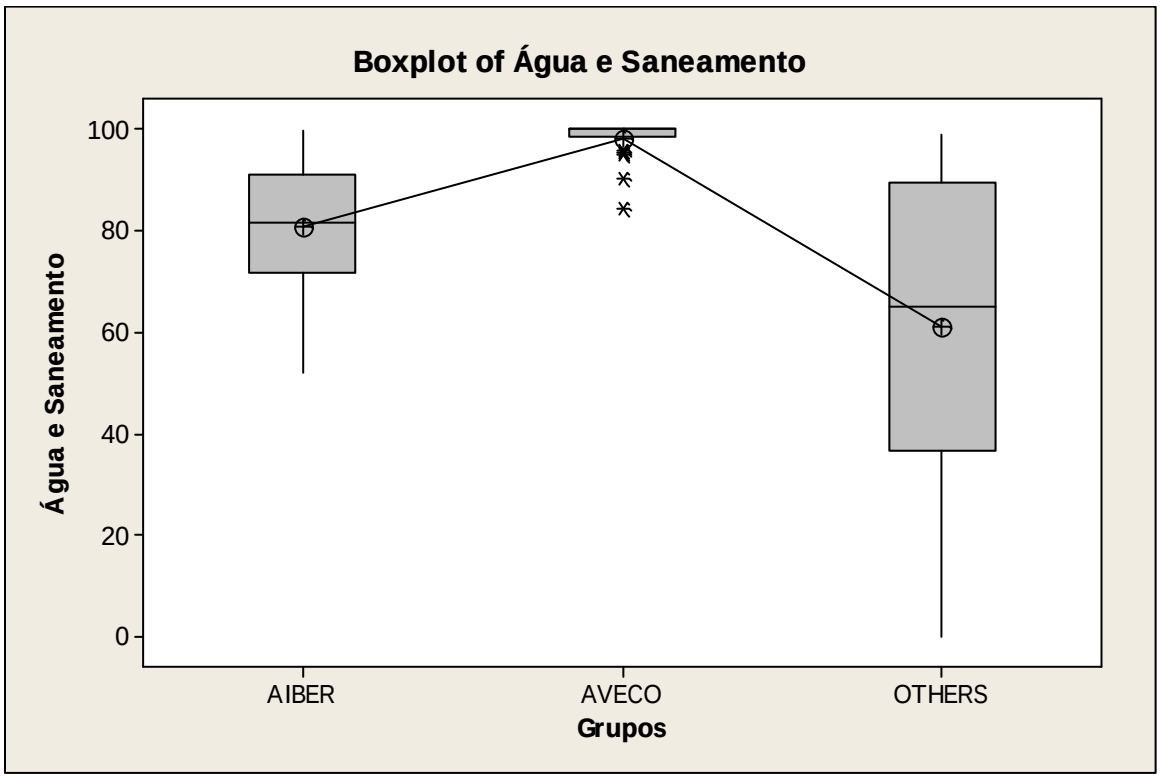
One-way ANOVA: Água e Saneamento versus Todas

```
Source   DF      SS     MS      F      P
Todas     7   72377  10340  44,65  0,000
Error   124   28713    232
Total   131  101090
```

```
S = 15,22   R-Sq = 71,60%   R-Sq(adj) = 69,99%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
AIBER  20  80,44  13,98                      (--*---)
AVECO  27  98,29   3,61                                (--*--)
CARLA   4  82,37   3,78                    (------*-------)
CWEAS  11  79,67  11,04                     (----*---)
EMDEU  12  90,56   9,69                           (---*----)
ESEAS  12  58,63  19,95           (---*----)
NSWUA  14  78,71  19,65                     (---*---)
SUSAF  32  34,78  20,45  (-*--)
                          -----+---------+---------+---------+----
                              40        60        80       100

Pooled StDev = 15,22
```

Boxplot of Água e Saneamento

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica na maior parte dos agrupamentos, com exceção das regiões AIBER E CWEAS. A AVECO apresenta o maior percentual do índice de água tratada e saneamento básico com média 98,2 e mediana 99,9, apresentando três outliers para baixo com valores de 95,5, 90,2 e 84,2, que ainda sim apresentam um nom percentual com relação a água tratada e saneamento básico. O agrupamento EMDEUA é o segundo melhor com relação a água tratada e saneamento básico com uma média de 90,5 e uma mediana de 93,3 e possui também dois outliers para baixo com 81 e 63,3, o primeiro demosntra ainda um bom percentual no entanto o segundo jã se encontra bastante abaixo da maioria dos países. O agrupamento SUSAF é a região com menor água tratada e saneamento básico, apresentando uma média de 34,5 e uma mediana de 32,5, o agrupamento apresenta apenas um outlier que é a Zambia com 96,3.



One-way ANOVA: Água e Saneamento versus Grupos

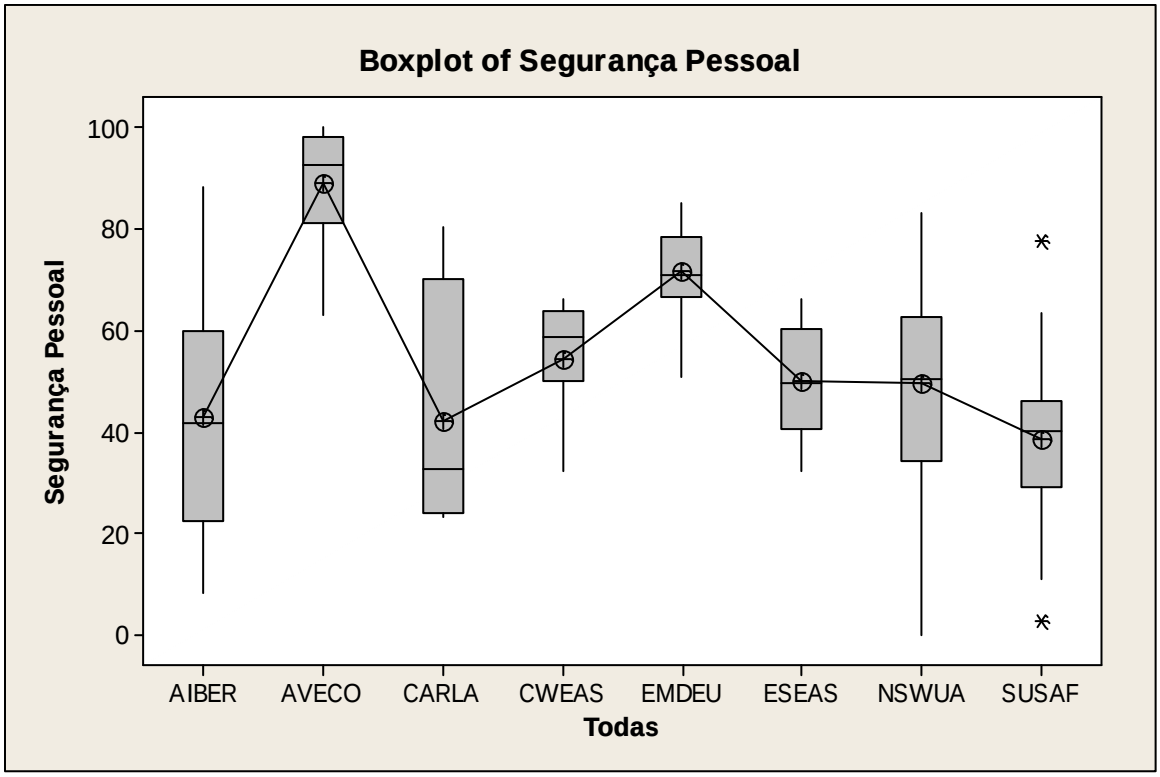
```
Source   DF      SS     MS      F      P
Grupos    2   30286  15143  27,59  0,000
Error   129   70804    549
Total   131  101090

S = 23,43   R-Sq = 29,96%   R-Sq(adj) = 28,87%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  80,59  13,65            (------*-----)
AVECO   27  98,29   3,61                       (-----*----)
OTHERS  84  61,04  28,36  (---*--)
                          ---+---------+---------+---------+------
                            60        75        90       105

Pooled StDev = 23,43
```

Boxplot of Água e Saneamento

Análise do gráfico: O agrupamento AVECO possui um maior índice de água tratada e saneamento básico com média 98,2 estando próximo da mediana 99,9 apresentando dados assimétricos, apresentando três outliers para baixo com valores de 95,5, 90,2 e 84,2, que ainda sim apresentam um nom percentual com relação a água tratada e saneamento básico. outliersssim como os demais agrupamentos. O agrupamento AIBER está melhor posicionado que os demais, com uma média de 80,5 e ediana de 81,5. Os o agrupamento OTHERS possui uma média de 61 e ua mediana de 65,1.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto por todos os agrupamentos.

1.1.4. Segurança Pessoal



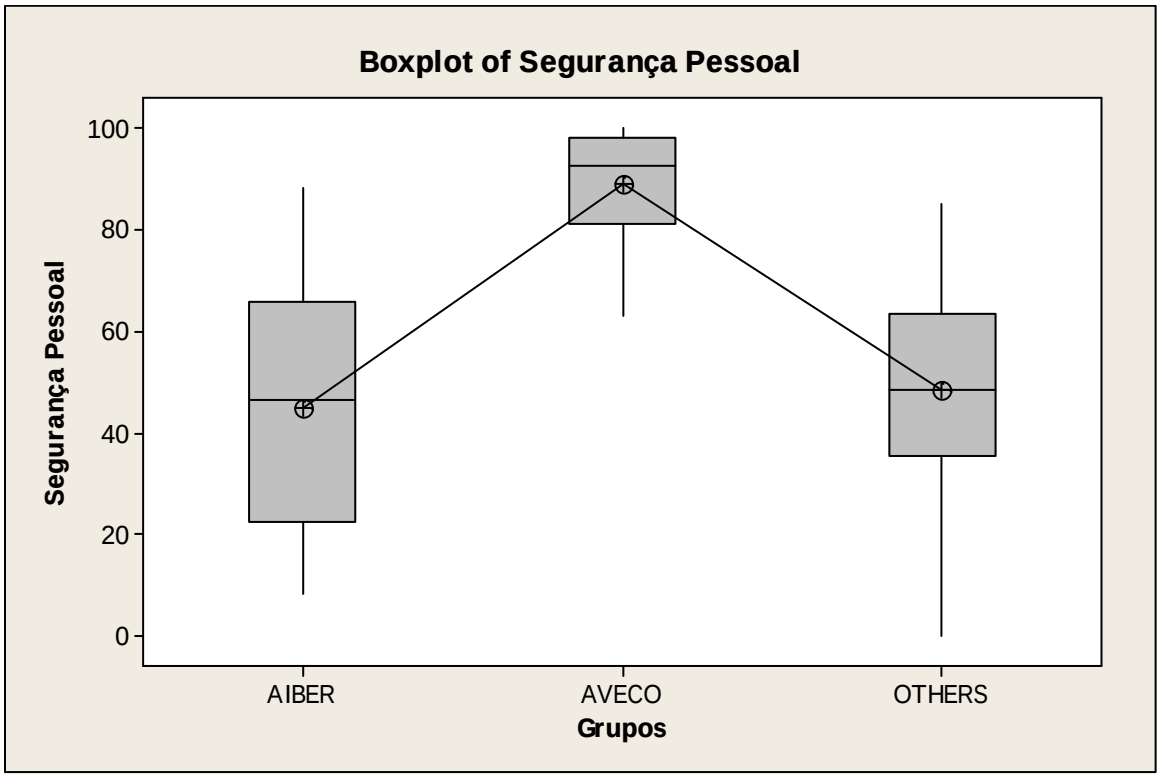
One-way ANOVA: Segurança Pessoal versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  47001  6714  24,66  0,000
Error   124  33765   272
Total   131  80766

S = 16,50   R-Sq = 58,19%   R-Sq(adj) = 55,83%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -------+---------+---------+---------+--
AIBER  20  43,02  23,76        (---*--)
AVECO  27  88,96  10,98                             (--*---)
CARLA   4  42,24  26,43   (-------*-------)
CWEAS  11  54,21  11,67           (----*----)
EMDEU  12  71,53   9,08                     (----*---)
ESEAS  12  49,99  11,28         (----*----)
NSWUA  14  49,67  21,60         (----*---)
SUSAF  32  38,65  16,14     (--*--)
                          -------+---------+---------+---------+--
                                40        60        80       100

Pooled StDev = 16,50
```

Boxplot of Segurança Pessoal

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica na maior parte dos agrupamentos, com exceção do agrupamento ESEAS. A AVECO apresenta o maior percentual em segurança pessoal, com uma média de 88,6. O agrupamento SUSAF é a região com menor segurança social, apresentando uma média de 38,6, o agrupamento apresenta dois outliers um para cima que é a Zambia com 77,5 e outro para baixo que é Angola com 2,8.



One-way ANOVA: Segurança Pessoal versus Grupos

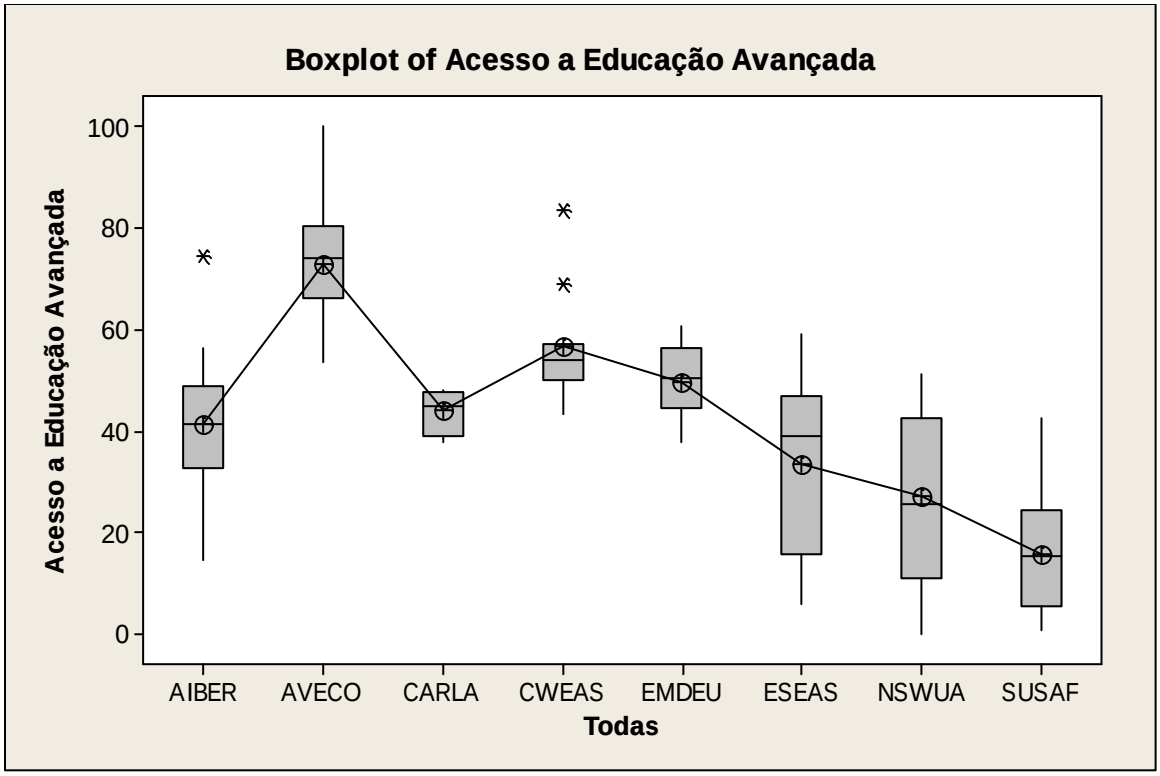
```
Source   DF     SS     MS      F      P
Grupos    2  36657  18329  53,60  0,000
Error   129  44109    342
Total   131  80766

S = 18,49   R-Sq = 45,39%   R-Sq(adj) = 44,54%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  44,80  24,55  (----*----)
AVECO   27  88,96  10,98                              (---*----)
OTHERS  84  48,52  18,66       (-*--)
                          -----+---------+---------+---------+----
                              45        60        75        90

Pooled StDev = 18,49
```

Boxplot of Segurança Pessoal

Análise do gráfico: O agrupamento AVECO possui um maior percentual de segurança pessoal com média 88,9 estando acima da mediana. O agrupamento AIBER está próximo dos demais com média de 44,8, sendo menor do que a média de 48,5 do agrupamento OTHERS.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto apenas pelos três agrupamentos: AIBER. AVECO e OTHER.

1.1.5. Acesso à Educação Avançada



One-way ANOVA: Acesso a Educação Avançada versus Todas

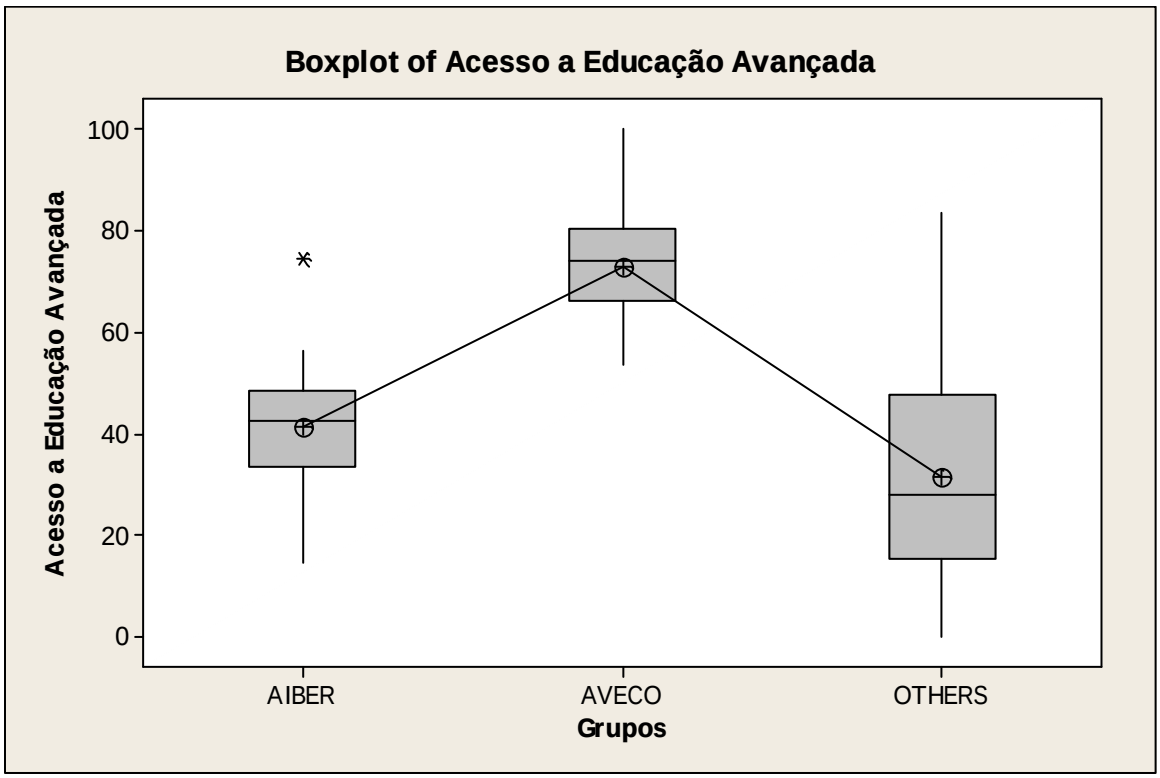
```
Source   DF     SS    MS      F      P
Todas     7  54967  7852  52,48  0,000
Error   124  18555   150
Total   131  73522

S = 12,23   R-Sq = 74,76%   R-Sq(adj) = 73,34%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ----+---------+---------+---------+-----
AIBER  20  41,20  12,96             (--*-)
AVECO  27  72,81  11,14                               (-*--)
CARLA   4  43,92   4,54           (-----*-----)
CWEAS  11  56,57  10,85                     (--*---)
EMDEU  12  49,79   7,09                 (---*--)
ESEAS  12  33,60  17,94         (---*--)
NSWUA  14  27,17  16,10      (---*--)
SUSAF  32  15,57  10,53  (-*-)
                          ----+---------+---------+---------+-----
                             20        40        60        80

Pooled StDev = 12,23
```

Boxplot of Acesso a Educação Avançada

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica na maior parte dos agrupamentos, com exceção dos AIBER, EMDEU e SUSAF. A AVECO apresenta o maior percentual de acesso à educação avançada com média de 72,8. O agrupamento CWEAS é o segundo melhor com uma média de 56,5, próximo da mediana de 54, apresentando dois outliers para cima um com 83,4 e outro com 68,7. O agrupamento AIBER com média 41,2 e

mediana 41,3, apresenta um outlier bem acima com 74,2 muito próximo da média do agrupamento AVECO. O agrupamento SUSAF é a região com menor acesso a educação avançada, com média de 15,5 bem próxima da mediana 15,2, e com o menor grau de confiança apresentando uma amplitude de 0,5 a 42,4.O maior grau de confiança está no agrupamento CARLA que apresenta uma amplitude de aproximadamente 10 pontos percentuais que vai de 37,3 a 47,8, sua mediana é de 45 e a média de 43,9.



One-way ANOVA: Acesso a Educação Avançada versus Grupos

Source	DF	SS	MS	F	P
Grupos	2	35187	17594	59,21	0,000
Error	129	38334	297		
Total	131	73522			

S = 17,24 R-Sq = 47,86% R-Sq(adj) = 47,05%

Level	N	Mean	StDev
AIBER	21	41,46	12,69
AVECO	27	72,81	11,14
OTHERS	84	31,32	19,60

Individual 95% CIs For Mean Based on Pooled StDev

```
--+---------+---------+---------+-------
      (----*----)
                              (----*---)
(--*-)
--+---------+---------+---------+-------
 30        45        60        75
```

Pooled StDev = 17,24

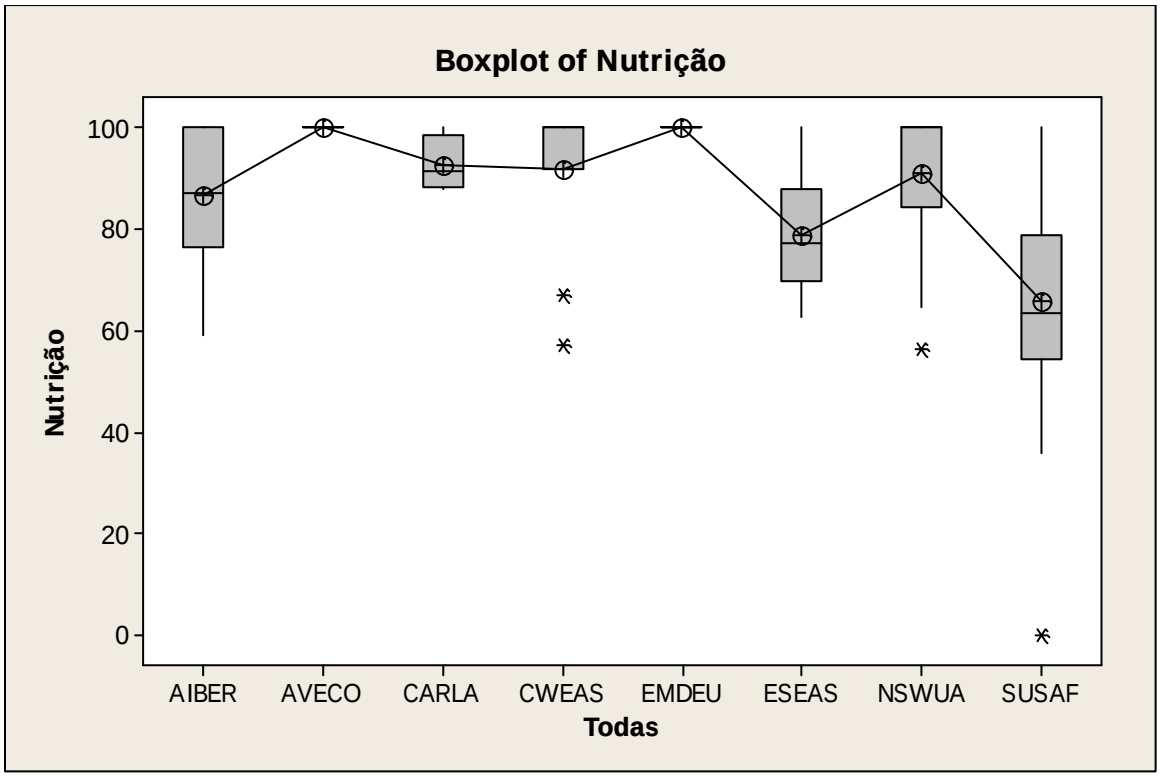
Boxplot of Acesso a Educação Avançada

Análise do gráfico: O agrupamento AVECO possui um maior percentual acesso à educação avançada com média de 72,8 estando próximo da mediana. O agrupamento AIBER está próximo dos demais com média de 41,4, sendo apenas 10 pontos percentuais maior que o

agrupamento OTHERS que possui uma média de 31,3. O menor grau de confiança apresentado é do agrupamento OTHERS com uma amplitude de 0 a 83,4.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto apenas pelos três agrupamentos: AIBER. AVECO e OTHER, no entanto, os valores foram muito próximos.

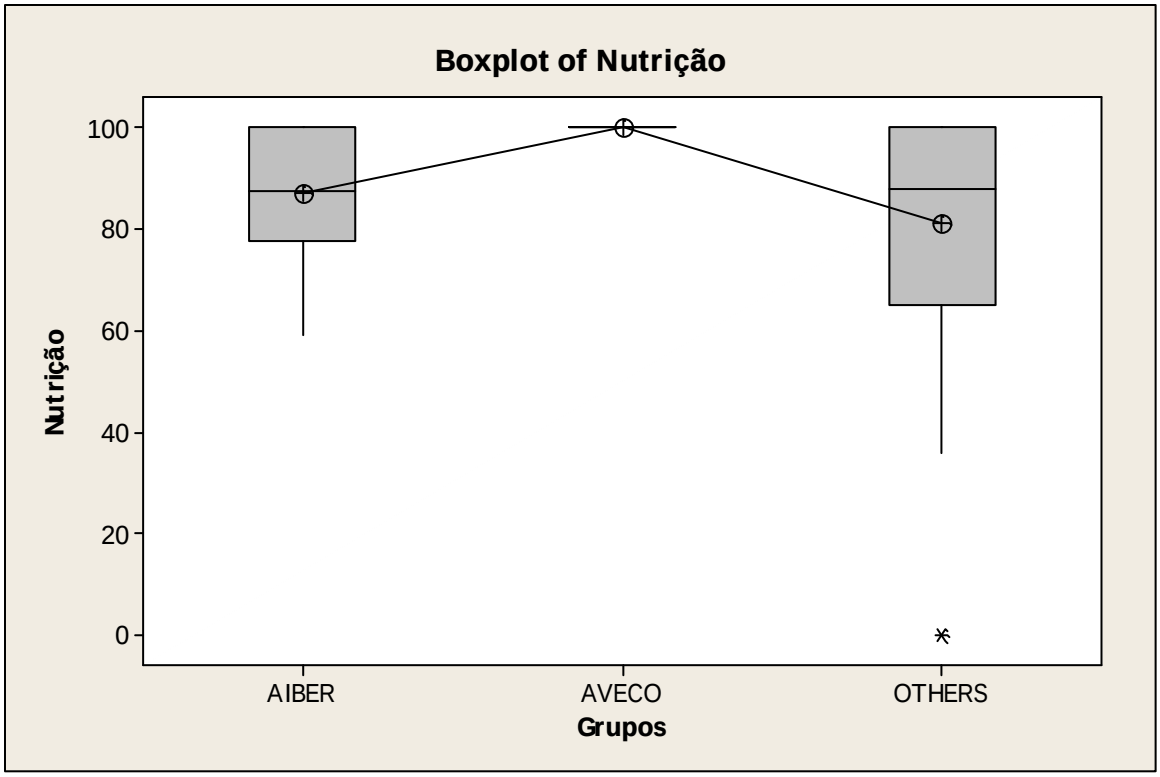
1.1.6. Nutrição



One-way ANOVA: Nutrição versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  22438  3205  17,60  0,000
Error   124  22581   182
Total   131  45019

S = 13,49   R-Sq = 49,84%   R-Sq(adj) = 47,01%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  ---------+---------+---------+---------+
AIBER  20   86,52  13,05                (----*----)
AVECO  27  100,00   0,00                          (---*----)
CARLA   4   92,59   5,56              (----------*----------)
CWEAS  11   91,66  15,14                  (-----*------)
EMDEU  12  100,00   0,00                         (-----*------)
ESEAS  12   78,93  10,71        (------*-----)
NSWUA  14   91,10  15,07                  (-----*-----)
SUSAF  32   65,61  20,28  (---*---)
                          ---------+---------+---------+---------+
                                  72        84        96       108

Pooled StDev = 13,49
```

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica em todos os agrupamentos. Os agrupamentos AVECO e EMDEU possuem os maiores percentuais de nutrição com 100%. O ior agrupamento com relação a nutrição é a SUSAF com uma média de 65,6, um bom índice com relação aos demais já analisados para este agrupamento, que possui o menor intervalo de confiança, apresentando uma amplitude de 35,6 a 100, este agrupamento apresenta um outlier em 0 que é Angola. O agrupamento NSWUA, tamb~em apsenta um outlier muito próximo da média do SUSAF, com 56,2 , próximo também aos dois outliers do CEWAS com 57,2 e 68,2.



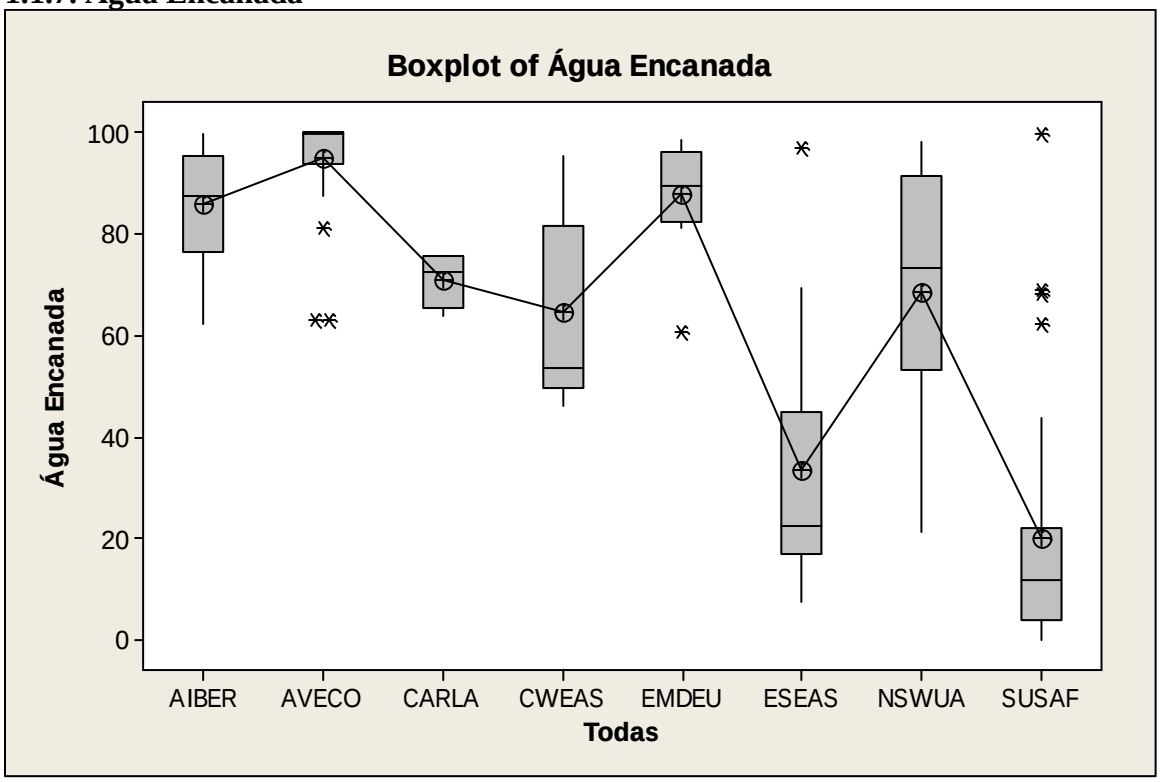
One-way ANOVA: Nutrição versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2   7449  3725  12,79  0,000
Error   129  37570   291
Total   131  45019

S = 17,07   R-Sq = 16,55%   R-Sq(adj) = 15,25%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N    Mean  StDev  ---+---------+---------+---------+------
AIBER   21   87,16  13,06     (--------*--------)
AVECO   27  100,00   0,00                       (-------*-------)
OTHERS  84   80,96  20,29  (---*----)
                          ---+---------+---------+---------+------
                          80,0      88,0      96,0     104,0

Pooled StDev = 17,07
```

Análise do gráfico: O agrupamento AVECO possui um maior percentual de nutrição com 100%. O agrupamento AIBER está próximo dos demais com média de 87,1, sendo aproximadamente 6 pontos percentuais maior que o agrupamento OTHERS que possui uma média de 80,9. O menor grau de confiança apresentado é do agrupamento OTHERS com uma amplitude de 35,6 a 100, este agrupamento possui também um outlier com 0 que é Angola.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto por todos os agrupamentos, no entanto os números estão muito próximos.

1.1.7. Água Encanada



One-way ANOVA: Água Encanada versus Todas

Source	DF	SS	MS	F	P
Todas	7	115588	16513	47,04	0,000
Error	124	43524	351		
Total	131	159113			

S = 18,74 R-Sq = 72,65% R-Sq(adj) = 71,10%

Level	N	Mean	StDev
AIBER	20	85,70	11,90
AVECO	27	94,89	10,29
CARLA	4	71,05	5,54

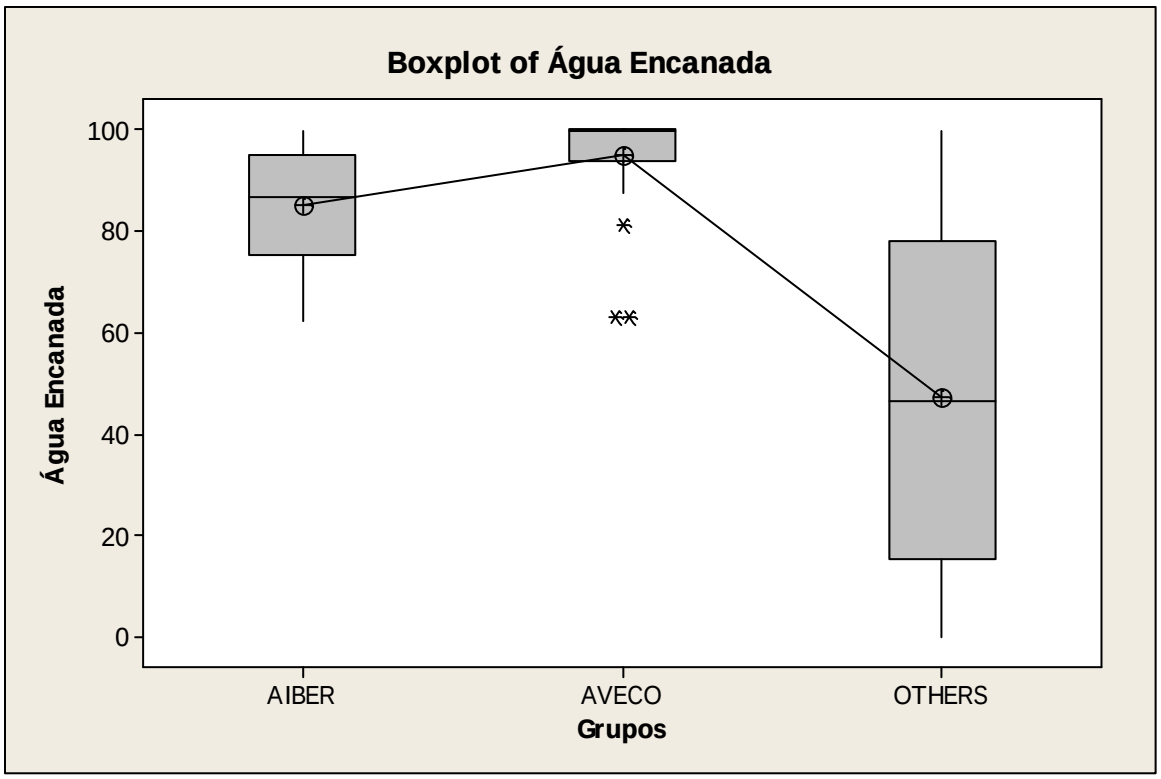
Individual 95% CIs For Mean Based on Pooled StDev

```
-----+---------+---------+---------+----
                          (--*---)
                             (--*--)
               (------*-------)
```

```
CWEAS  11  64,69  18,52                  (----*---)
EMDEU  12  87,72  10,31                           (---*---)
ESEAS  12  33,41  26,28      (---*----)
NSWUA  14  68,43  23,81                    (---*---)
SUSAF  32  19,96  24,38  (--*--)
                         -----+---------+---------+---------+----
                             25        50        75       100

Pooled StDev = 18,74
```

Análise do gráfico: No gráfico a cima a distribuição dos dados é assimétrica em todos os agrupamentos. O agrupamento AVECO possue os maior percentuais de água encanada com média 94,8, com três outliers para baixo. Um com 81,2 e dois com 63. Aparecem outliers também nos agrupamentos EMDEU que possui um bem abaixo da média que é 87,7 com 60,6, no ESEAS que possui um bem acima da média que é 33,4 para cima 86,8 com e no SUSAF com três bem acima da média que é 19,9 acima 99,7, 69 e 62,3.

Boxplot of Água Encanada



One-way ANOVA: Água Encanada versus Grupos

```
Source   DF      SS     MS      F      P
Grupos    2   58626  29313  37,63  0,000
Error   129  100487    779
Total   131  159113

S = 27,91   R-Sq = 36,85%   R-Sq(adj) = 35,87%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  85,22  11,81                (-----*-----)
AVECO   27  94,89  10,29                     (----*-----)
```

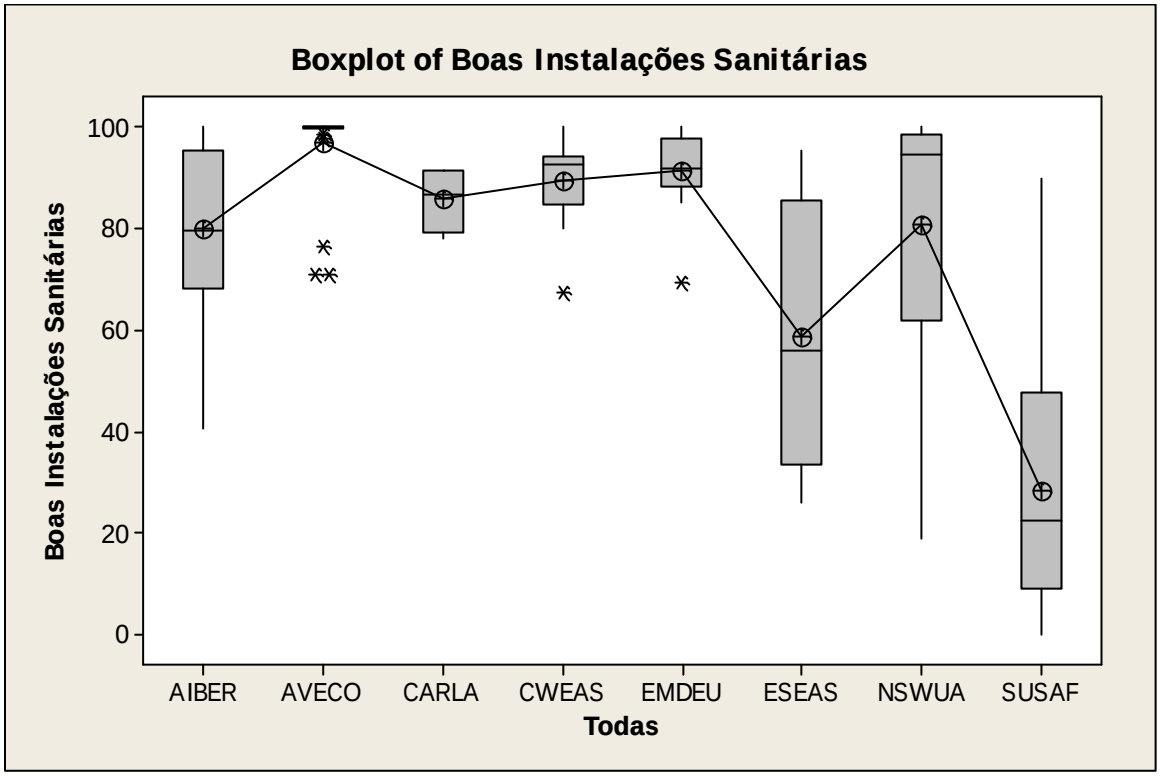
```
OTHERS  84  47,27  33,82  (--*--)
                          ---------+---------+---------+---------+
                                   60        80       100       120

Pooled StDev = 27,91
```

Boxplot of Água Encanada

Análise do gráfico: O agrupamento AVECOEU possue os maior percentuais de água encanada com média 94,8, com três outliers para baixo. O agrupamento AIBER está próximo do AVECO com média de 85,2, bastante acima do agrupamento OTHERS com média 47,2. O menor grau de confiança apresentado é do agrupamento OTHERS com uma amplitude de 0 a 97,7.

Análise da variável a partir dos gráficos: Os F foi melhor no gráfico composto por todos os agrupamentos, no entanto os números estão muito próximos.

1.1.8. Boas Instalações Sanitárias

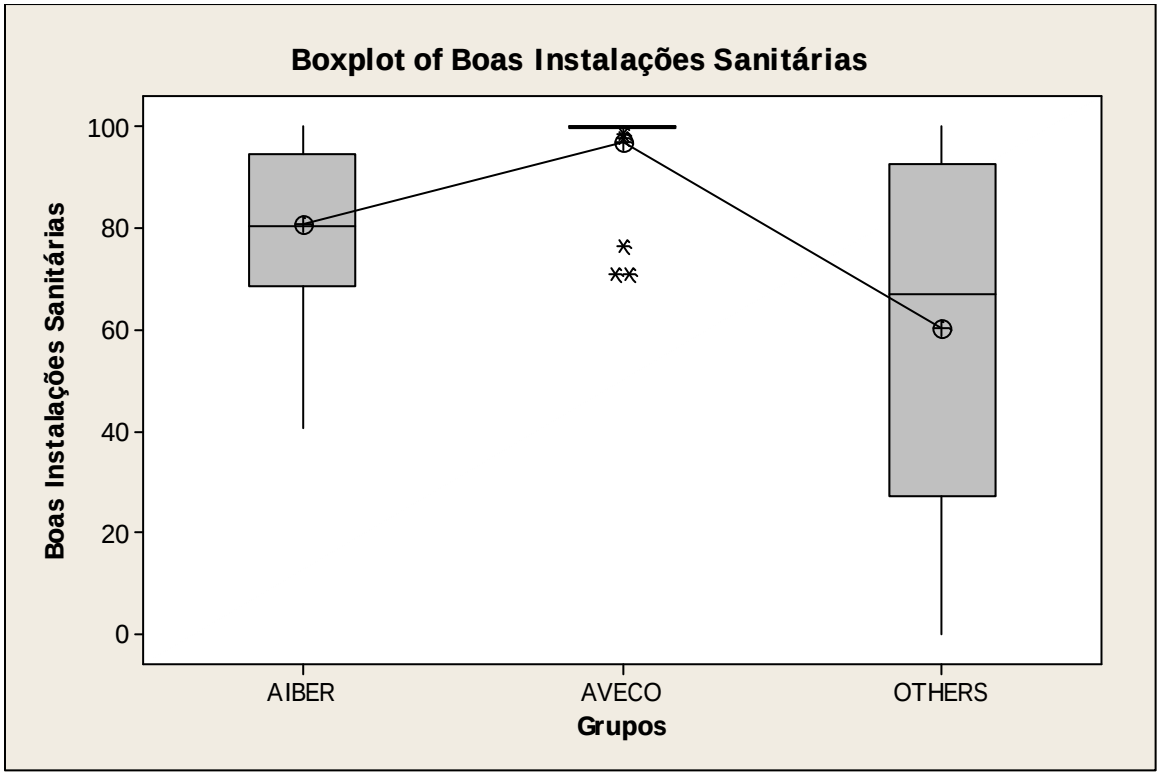


One-way ANOVA: Boas Instalações Sanitárias versus Todas

```
Source   DF      SS     MS      F      P
Todas     7   90984  12998  37,87  0,000
Error   124   42562    343
Total   131  133545

S = 18,53   R-Sq = 68,13%   R-Sq(adj) = 66,33%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
```

```
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  80,10  16,92                    (--*--)
AVECO  27  96,75   8,71                           (--*--)
CARLA   4  85,69   6,64                   (------*-------)
CWEAS  11  89,43   9,16                       (----*---)
EMDEU  12  91,30   8,49                        (----*---)
ESEAS  12  58,62  24,81           (---*----)
NSWUA  14  80,57  26,35                     (---*---)
SUSAF  32  28,15  23,82  (-*--)
                         -+---------+---------+---------+--------
                         25        50        75       100

Pooled StDev = 18,53
```

Boxplot of Boas Instalações Sanitárias



One-way ANOVA: Boas Instalações Sanitárias versus Grupos

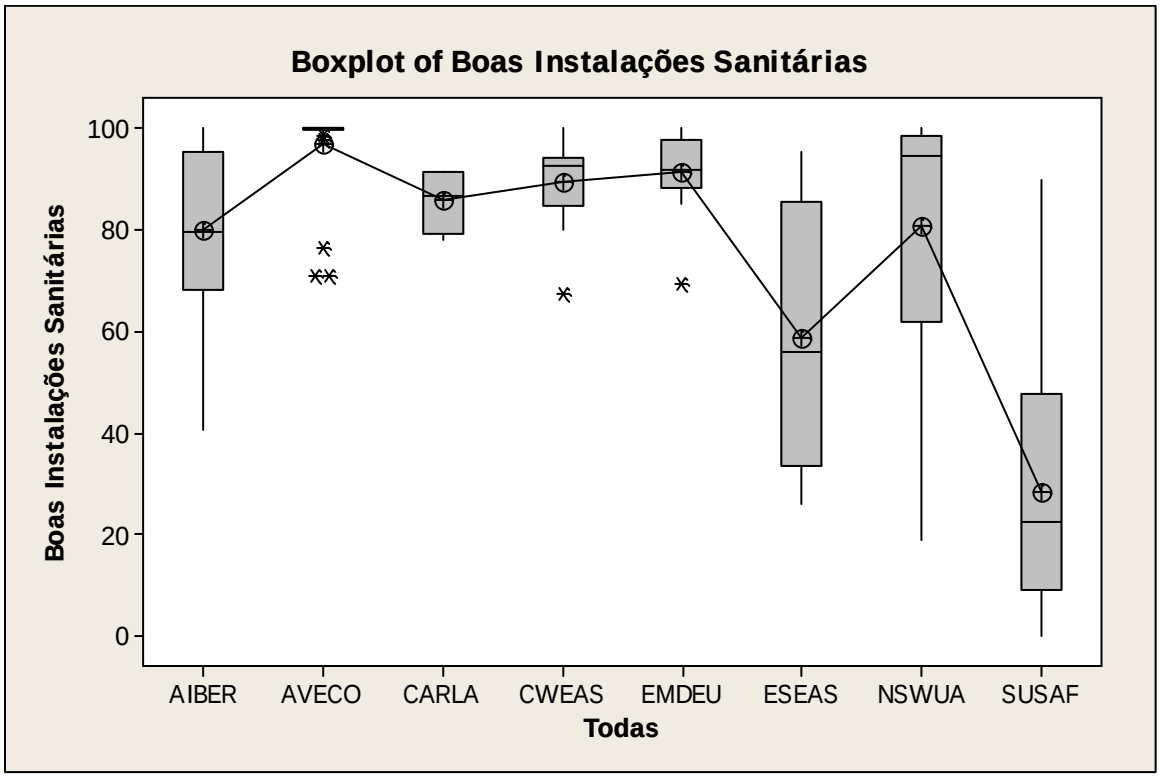
```
Source   DF      SS     MS      F      P
Grupos    2   29518  14759  18,30  0,000
Error   129  104027    806
Total   131  133545

S = 28,40   R-Sq = 22,10%   R-Sq(adj) = 20,90%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  80,63  16,67            (-------*-------)
AVECO   27  96,75   8,71                      (-------*------)
OTHERS  84  60,27  34,10  (---*---)
                          ----+---------+---------+---------+-----
                             60        75        90       105
```

```
Pooled StDev = 28,40
```

Boxplot of Boas Instalações Sanitárias

1.1.9. Acesso a eletricidade

1.1.10. Qualidade no Fornecimento de Eletricidade

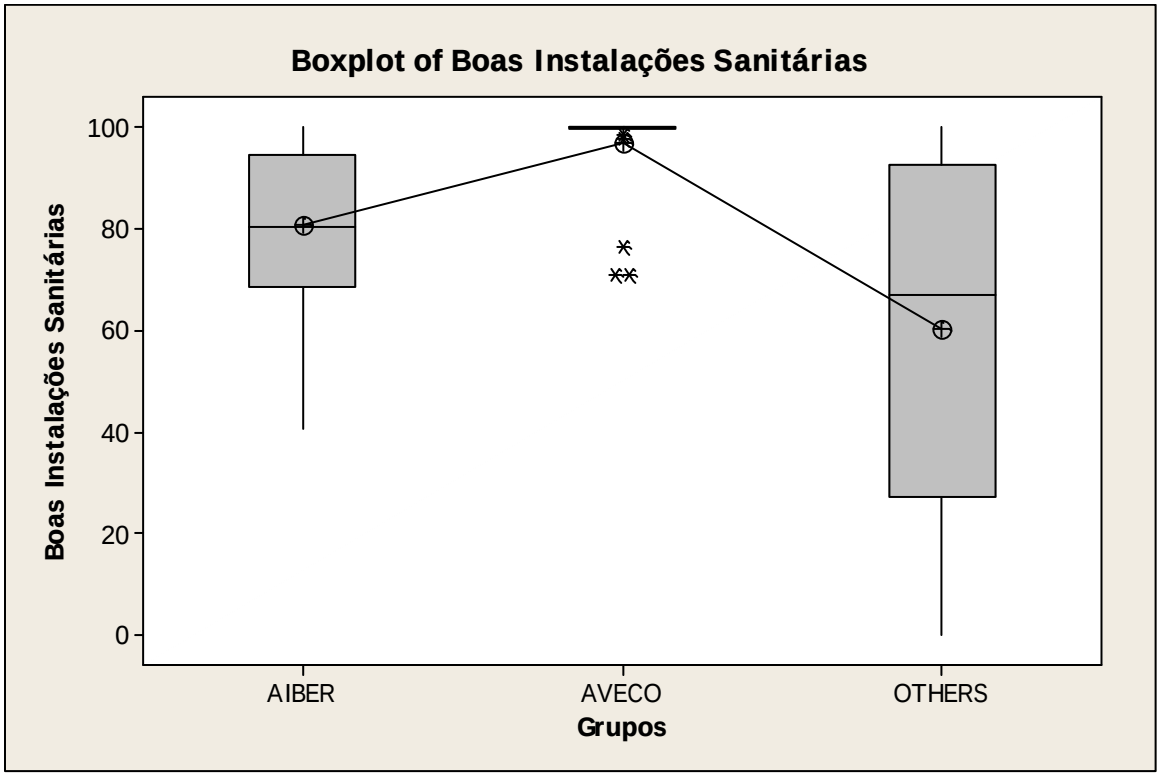


One-way ANOVA: Boas Instalações Sanitárias versus Todas

```
Source   DF      SS     MS      F      P
Todas     7   90984  12998  37,87  0,000
Error   124   42562    343
Total   131  133545

S = 18,53   R-Sq = 68,13%   R-Sq(adj) = 66,33%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  80,10  16,92                    (--*--)
AVECO  27  96,75   8,71                           (--*--)
CARLA   4  85,69   6,64                   (------*-------)
CWEAS  11  89,43   9,16                       (----*---)
EMDEU  12  91,30   8,49                        (----*---)
ESEAS  12  58,62  24,81           (---*----)
NSWUA  14  80,57  26,35                     (---*---)
SUSAF  32  28,15  23,82  (-*--)
                          -+---------+---------+---------+--------
```

```
                          25        50        75       100

Pooled StDev = 18,53
```

Boxplot of Boas Instalações Sanitárias

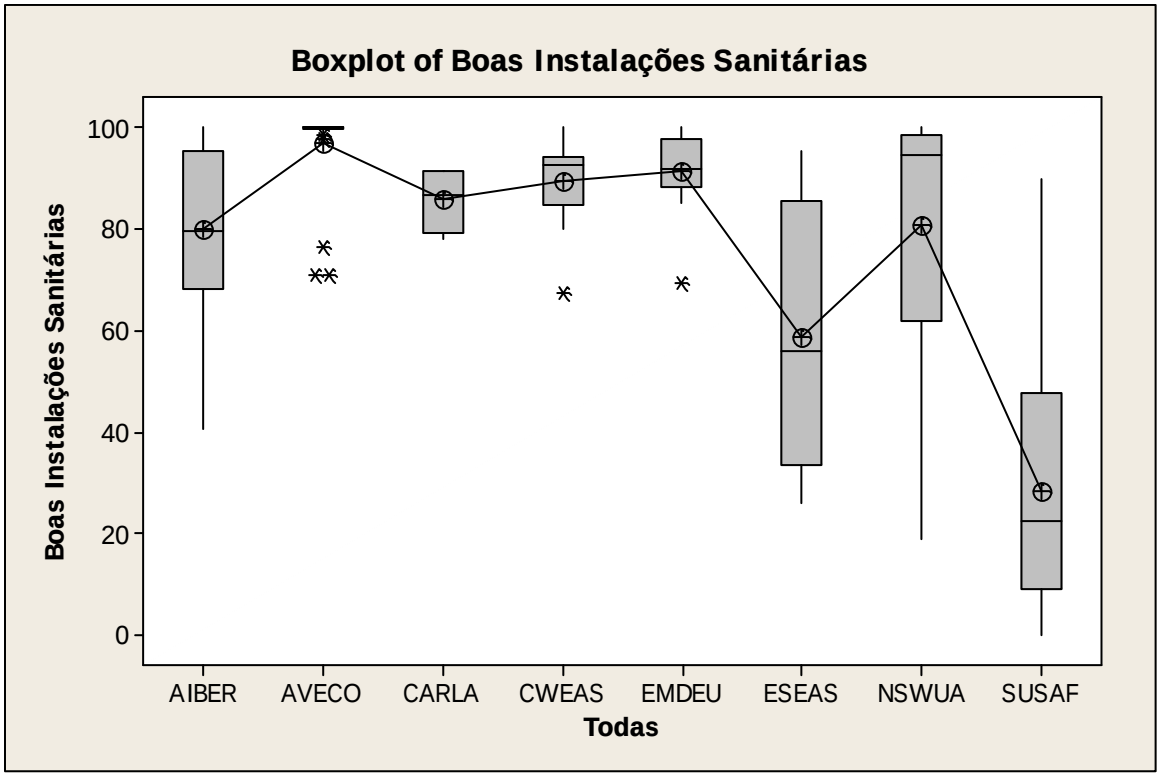


One-way ANOVA: Boas Instalações Sanitárias versus Grupos

```
Source   DF      SS     MS      F      P
Grupos    2   29518  14759  18,30  0,000
Error   129  104027    806
Total   131  133545

S = 28,40   R-Sq = 22,10%   R-Sq(adj) = 20,90%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  80,63  16,67            (-------*-------)
AVECO   27  96,75   8,71                      (-------*------)
OTHERS  84  60,27  34,10  (---*---)
                          ----+---------+---------+---------+-----
                             60        75        90       105

Pooled StDev = 28,40
```

Boxplot of Boas Instalações Sanitárias

1.1.11. Índice de Liberdade de Imprensa

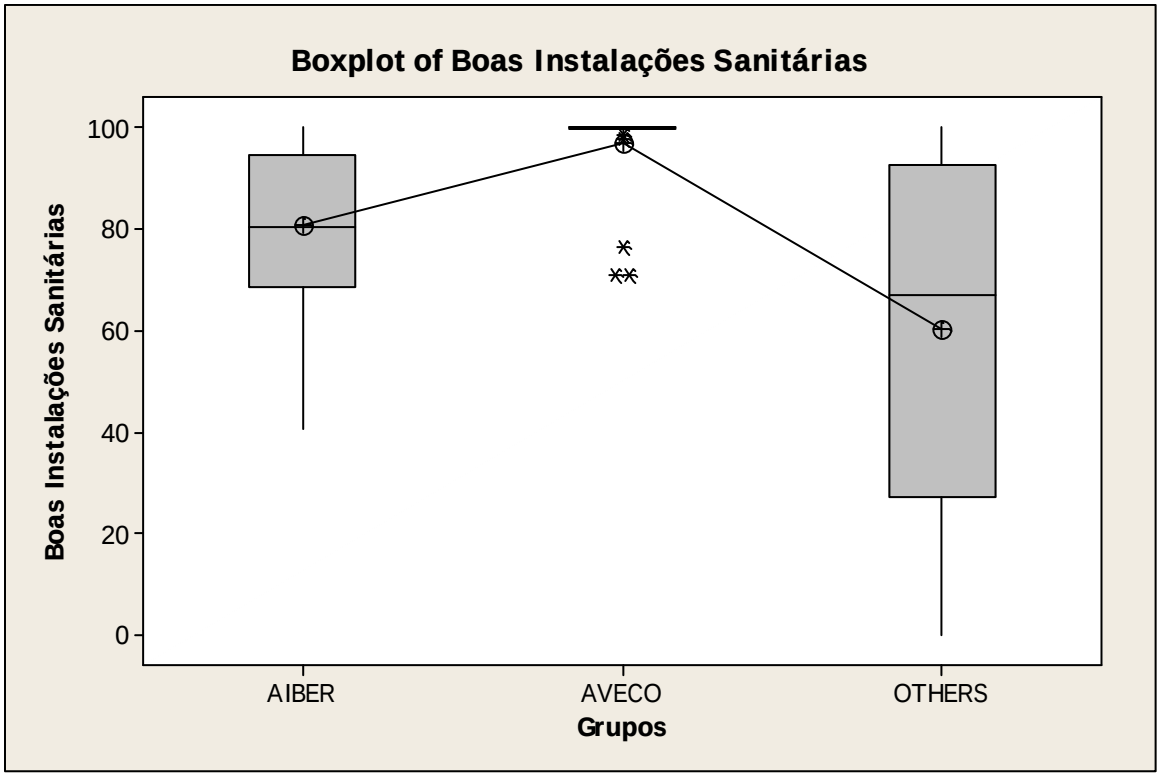


One-way ANOVA: Boas Instalações Sanitárias versus Todas

```
Source   DF      SS     MS      F      P
Todas     7   90984  12998  37,87  0,000
Error   124   42562    343
Total   131  133545

S = 18,53   R-Sq = 68,13%   R-Sq(adj) = 66,33%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  80,10  16,92                    (--*--)
AVECO  27  96,75   8,71                            (--*--)
CARLA   4  85,69   6,64                   (------*-------)
CWEAS  11  89,43   9,16                       (----*---)
EMDEU  12  91,30   8,49                        (----*---)
ESEAS  12  58,62  24,81           (---*----)
NSWUA  14  80,57  26,35                    (---*---)
SUSAF  32  28,15  23,82  (-*--)
                          -+---------+---------+---------+--------
                          25        50        75       100

Pooled StDev = 18,53
```

Boxplot of Boas Instalações Sanitárias

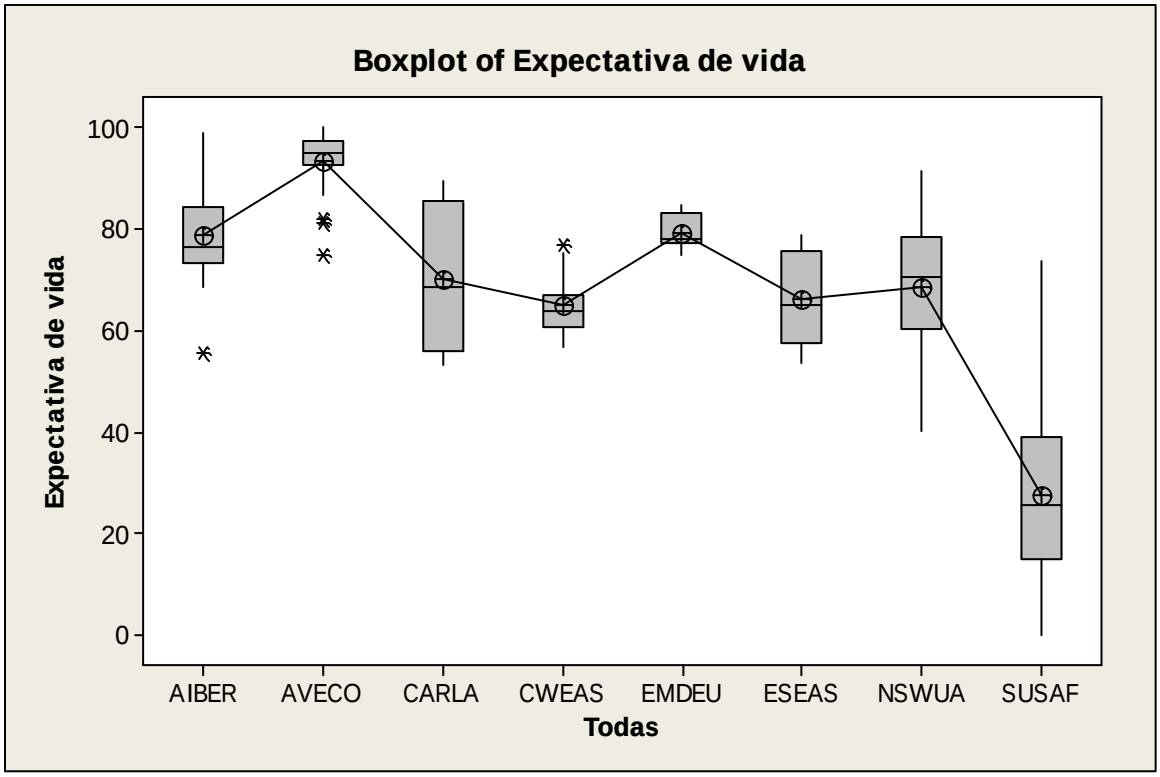


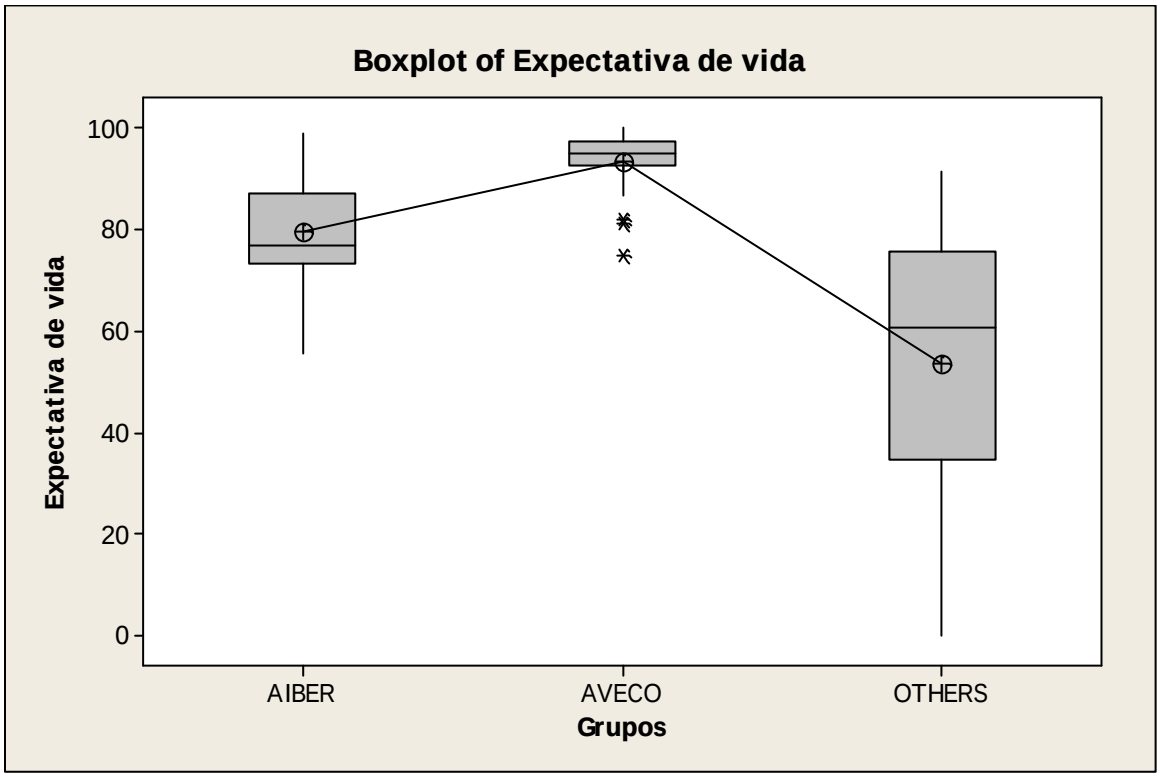
One-way ANOVA: Boas Instalações Sanitárias versus Grupos

```
Source   DF      SS     MS      F      P
Grupos    2   29518  14759  18,30  0,000
Error   129  104027    806
Total   131  133545

S = 28,40   R-Sq = 22,10%   R-Sq(adj) = 20,90%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  80,63  16,67            (-------*-------)
AVECO   27  96,75   8,71                      (-------*------)
OTHERS  84  60,27  34,10  (---*---)
                          ----+---------+---------+---------+-----
                             60        75        90       105

Pooled StDev = 28,40
```

Boxplot of Boas Instalações Sanitárias

1.1.12. Expectativa de Vida





One-way ANOVA: Expectativa de vida versus Grupos

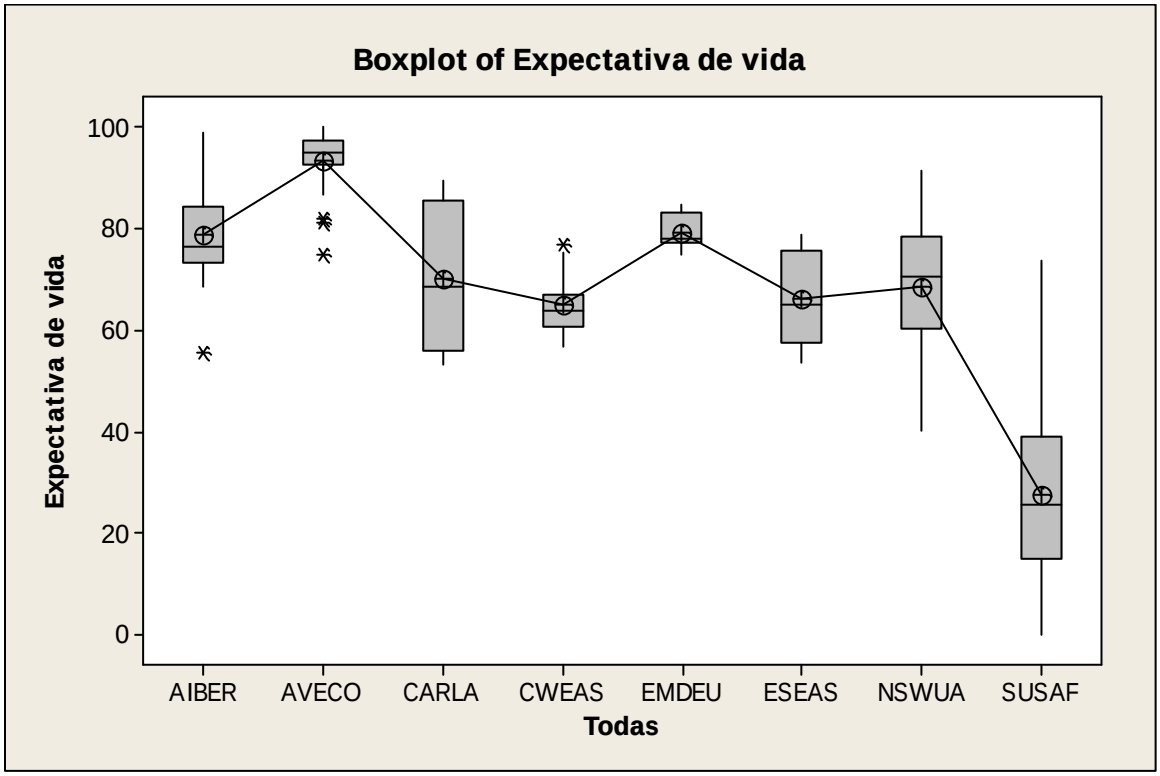
Source	DF	SS	MS	F	P
Grupos	2	37373	18687	46,67	0,000
Error	129	51656	400		
Total	131	89030			

S = 20,01 R-Sq = 41,98% R-Sq(adj) = 41,08%

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  79,44  10,14               (-----*-----)
AVECO   27  93,43   6,01                         (----*----)
OTHERS  84  53,43  24,21  (--*-)
                          -------+---------+---------+---------+--
                                60        75        90       105

Pooled StDev = 20,01
```

Boxplot of Expectativa de vida

1.1.13. Expectativa de Vida entre 30 e 70 anos

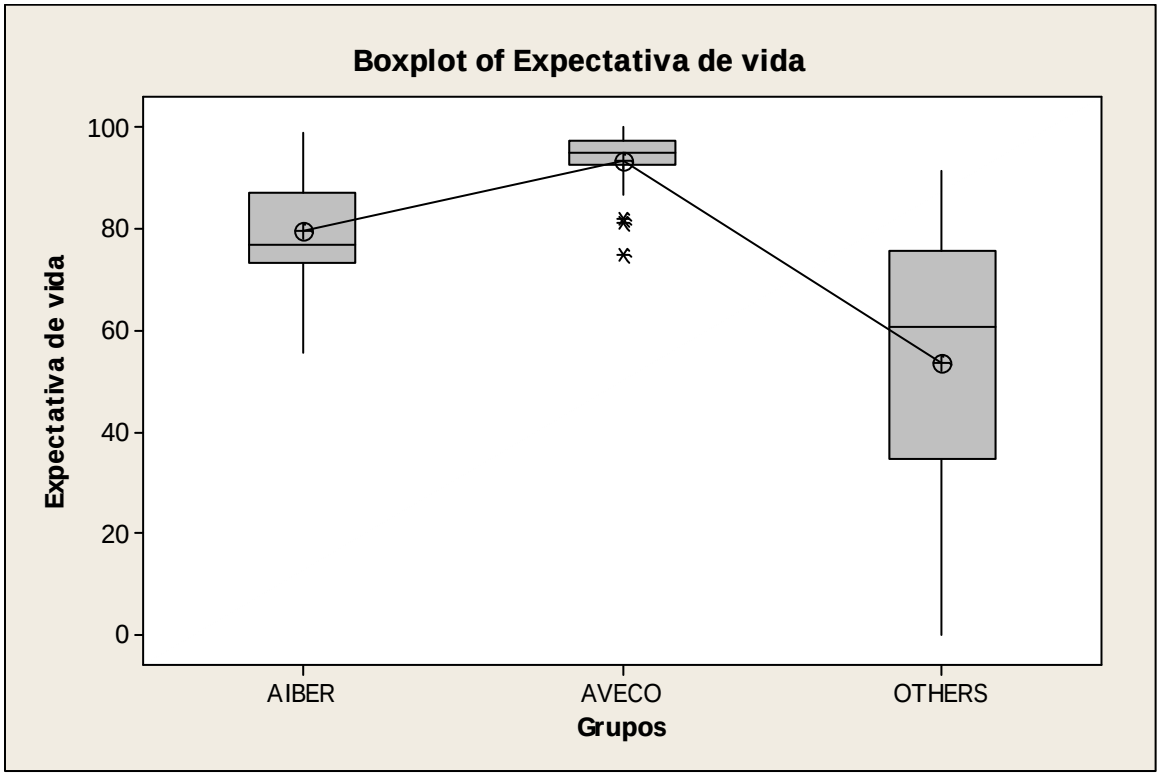


One-way ANOVA: Expectativa de vida versus Todas

```
Source   DF     SS     MS      F      P
Todas     7  73310  10473  82,61  0,000
Error   124  15719    127
Total   131  89030

S = 11,26   R-Sq = 82,34%   R-Sq(adj) = 81,35%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  20  78,94  10,14                           (-*--)
```

```
AVECO  27  93,43   6,01                                  (-*-)
CARLA   4  70,01  15,32                    (-----*-----)
CWEAS  11  65,10   6,28                    (---*--)
EMDEU  12  79,19   3,28                           (---*--)
ESEAS  12  65,99   9,08                     (--*--)
NSWUA  14  68,47  14,45                      (--*--)
SUSAF  32  27,51  16,05  (-*-)
                         --------+---------+---------+---------+-
                                 40        60        80       100

Pooled StDev = 11,26
```

Boxplot of Expectativa de vida

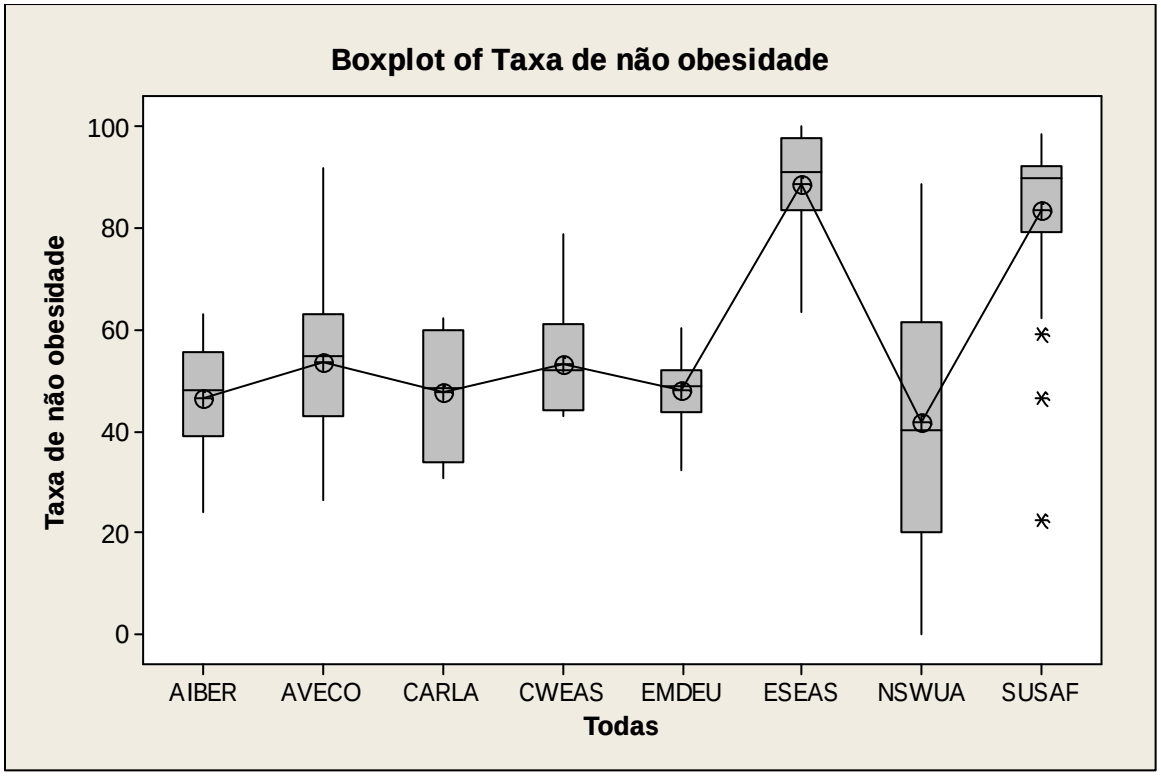


One-way ANOVA: Expectativa de vida versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  37373  18687  46,67  0,000
Error   129  51656    400
Total   131  89030

S = 20,01   R-Sq = 41,98%   R-Sq(adj) = 41,08%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  79,44  10,14                (-----*-----)
AVECO   27  93,43   6,01                          (----*----)
OTHERS  84  53,43  24,21  (--*-)
                          -------+---------+---------+---------+--
                                60        75        90       105

Pooled StDev = 20,01
```

Boxplot of Expectativa de vida

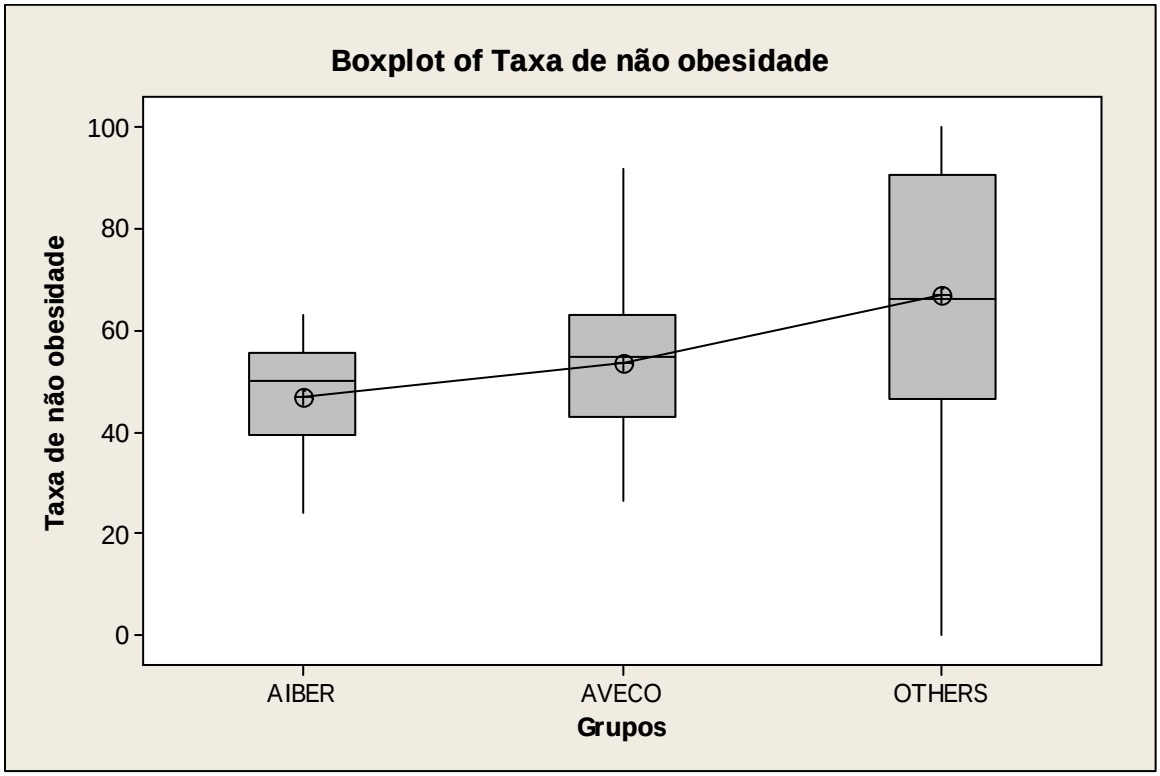
1.1.14. Taxa de não obesidade



One-way ANOVA: Taxa de não obesidade versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  39502  5643  25,27  0,000
Error   124  27692   223
Total   131  67194

S = 14,94   R-Sq = 58,79%   R-Sq(adj) = 56,46%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ----+---------+---------+---------+-----
AIBER  20  46,34  10,85       (--*--)
AVECO  27  53,73  14,61           (--*--)
CARLA   4  47,48  13,50  (-------*------)
CWEAS  11  53,17  10,80         (----*---)
EMDEU  12  48,12   6,95       (---*---)
ESEAS  12  88,55  11,81                          (---*----)
NSWUA  14  41,83  24,76    (---*---)
SUSAF  32  83,45  16,22                        (--*-)
                          ----+---------+---------+---------+-----
                             40        60        80       100

Pooled StDev = 14,94
```

Boxplot of Taxa de não obesidade

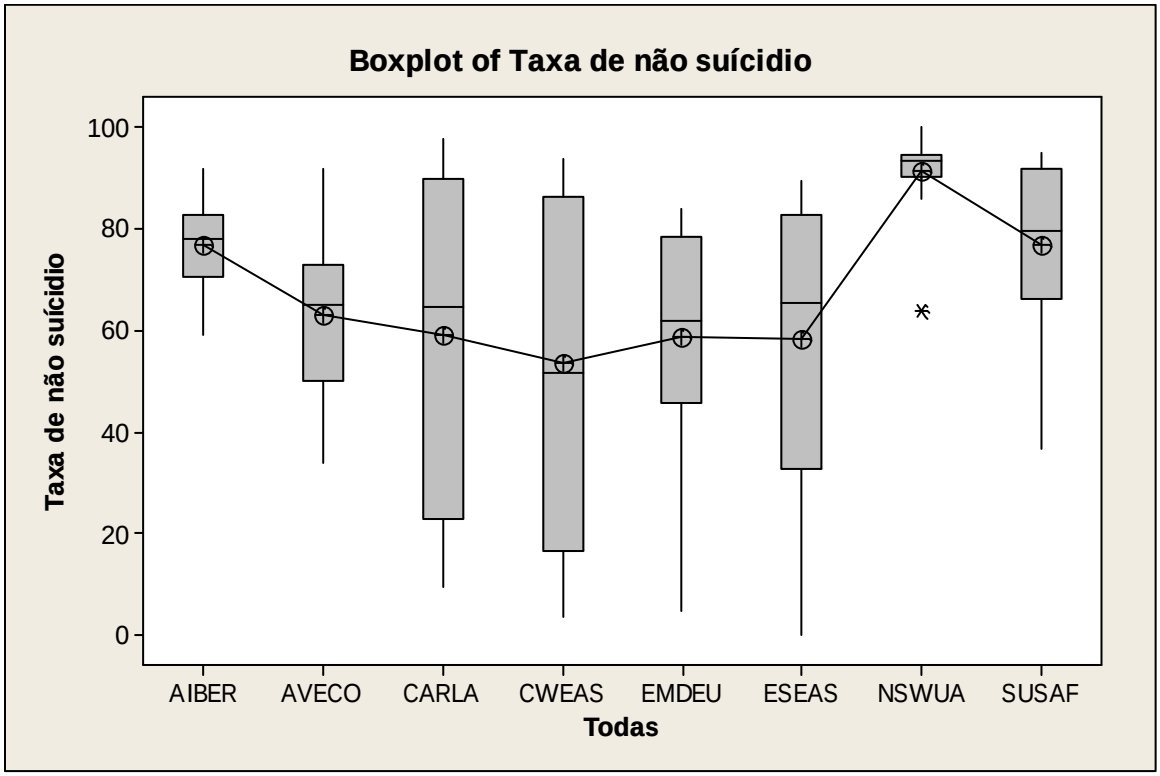


One-way ANOVA: Taxa de não obesidade versus Grupos

```
Source   DF     SS    MS     F      P
Grupos    2   8629  4315  9,50  0,000
Error   129  58565   454
Total   131  67194

S = 21,31   R-Sq = 12,84%   R-Sq(adj) = 11,49%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  46,68  10,68  (---------*--------)
AVECO   27  53,73  14,61           (-------*-------)
OTHERS  84  66,88  24,72                        (----*---)
                          ---+---------+---------+---------+------
                            40        50        60        70

Pooled StDev = 21,31
```

Boxplot of Taxa de não obesidade

1.1.15. Taxa de não suicídio



One-way ANOVA: Taxa de não suícidio versus Todas

Source	DF	SS	MS	F	P
Todas	7	16527	2361	6,44	0,000
Error	124	45441	366		
Total	131	61968			

S = 19,14 R-Sq = 26,67% R-Sq(adj) = 22,53%

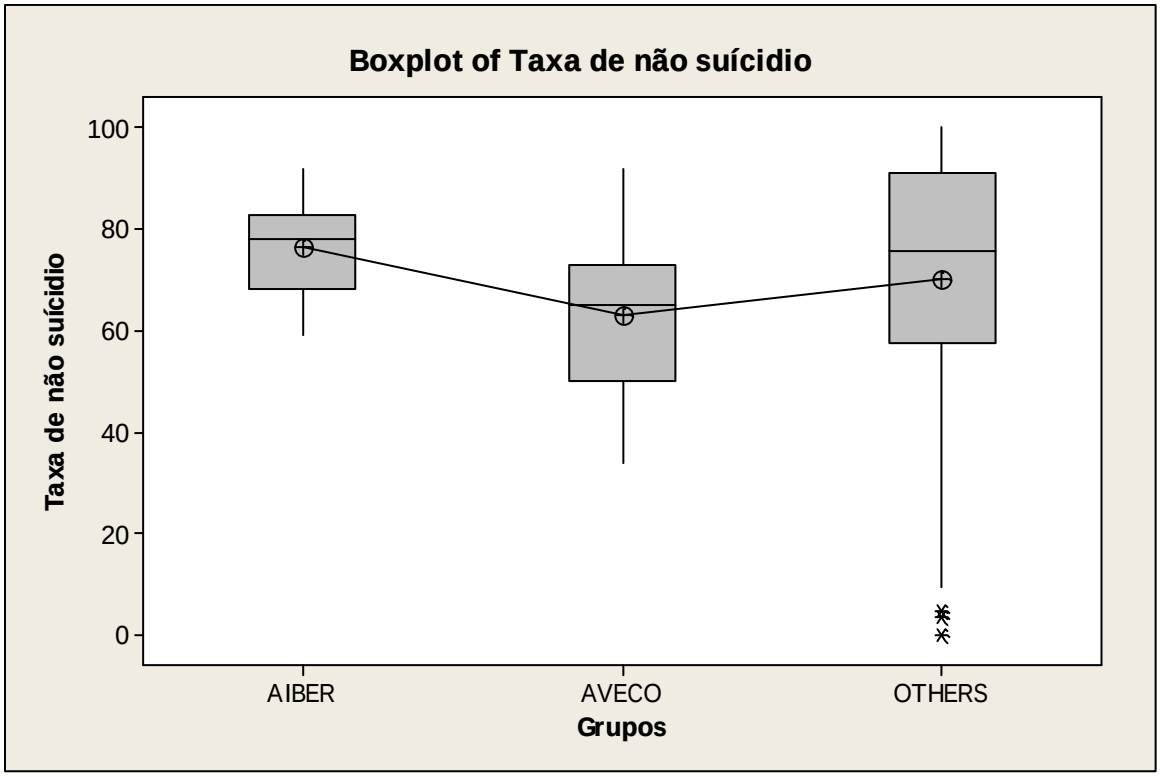
Level	N	Mean	StDev
AIBER	20	76,97	8,54
AVECO	27	63,17	13,86
CARLA	4	59,03	36,72
CWEAS	11	53,72	33,13
EMDEU	12	58,61	22,51
ESEAS	12	58,40	28,30
NSWUA	14	91,25	8,60
SUSAF	32	76,65	16,75

Individual 95% CIs For Mean Based on Pooled StDev

```
-----+---------+---------+---------+----
                   (----*----)
          (---*----)
(-----------*-----------)
 (-------*------)
     (------*-----)
     (------*-----)
                          (-----*-----)
                     (---*---)
-----+---------+---------+---------+----
    48        64        80        96
```

Pooled StDev = 19,14

Boxplot of Taxa de não suícidio

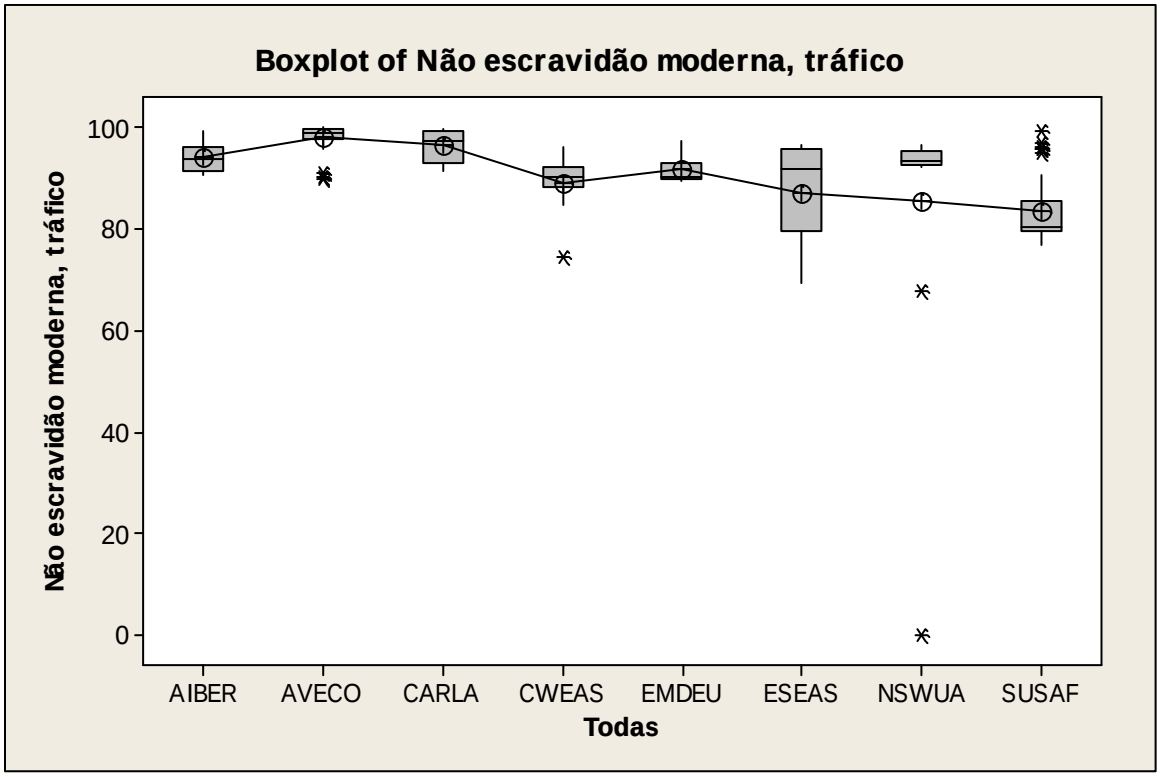


One-way ANOVA: Taxa de não suícidio versus Grupos

```
Source   DF     SS    MS     F      P
Grupos    2   2117  1059  2,28  0,106
Error   129  59850   464
Total   131  61968

S = 21,54   R-Sq = 3,42%   R-Sq(adj) = 1,92%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev    -+---------+---------+---------+--------
AIBER   21  76,41   8,72                  (-----------*----------)
AVECO   27  63,17  13,86    (---------*---------)
OTHERS  84  70,19  25,35                (-----*-----)
                            -+---------+---------+---------+--------
                          56,0      64,0      72,0      80,0

Pooled StDev = 21,54
```

Boxplot of Taxa de não suícidio

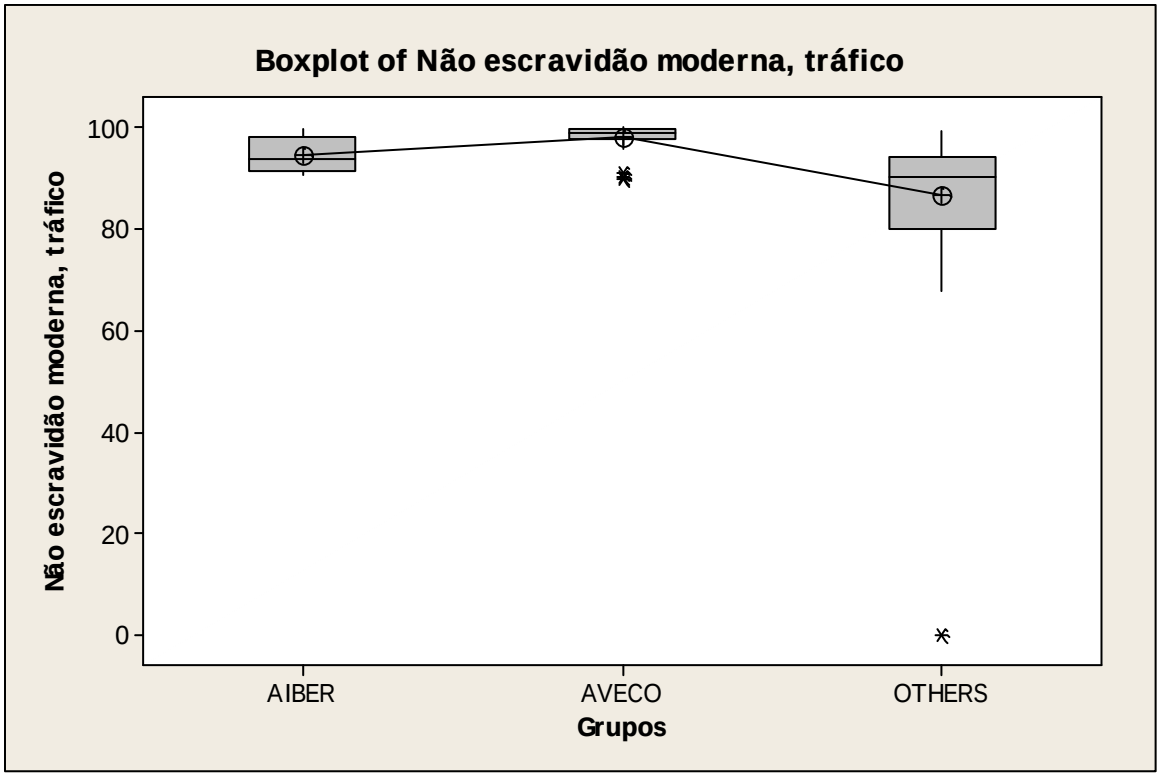
1.1.16. Não escravidão moderna, tráfico de seres humanos e casamento infantil.



One-way ANOVA: Não escravidão moderna, tráfico versus Todas

```
Source   DF       SS     MS     F      P
Todas     7   4036,7  576,7  6,00  0,000
Error   124  11910,3   96,1
Total   131  15947,0

S = 9,801   R-Sq = 25,31%   R-Sq(adj) = 21,10%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  20  94,28   3,15               (------*-----)
AVECO  27  98,03   2,96                     (----*----)
CARLA   4  96,54   3,59           (-------------*-------------)
CWEAS  11  89,11   5,60       (-------*--------)
EMDEU  12  91,65   2,76           (-------*-------)
ESEAS  12  86,99   9,80    (-------*-------)
NSWUA  14  85,68  25,68   (------*-------)
SUSAF  32  83,49   6,79  (----*----)
                          ------+---------+---------+---------+---
                             84,0      91,0      98,0     105,0

Pooled StDev = 9,80
```

Boxplot of Não escravidão moderna, tráfico

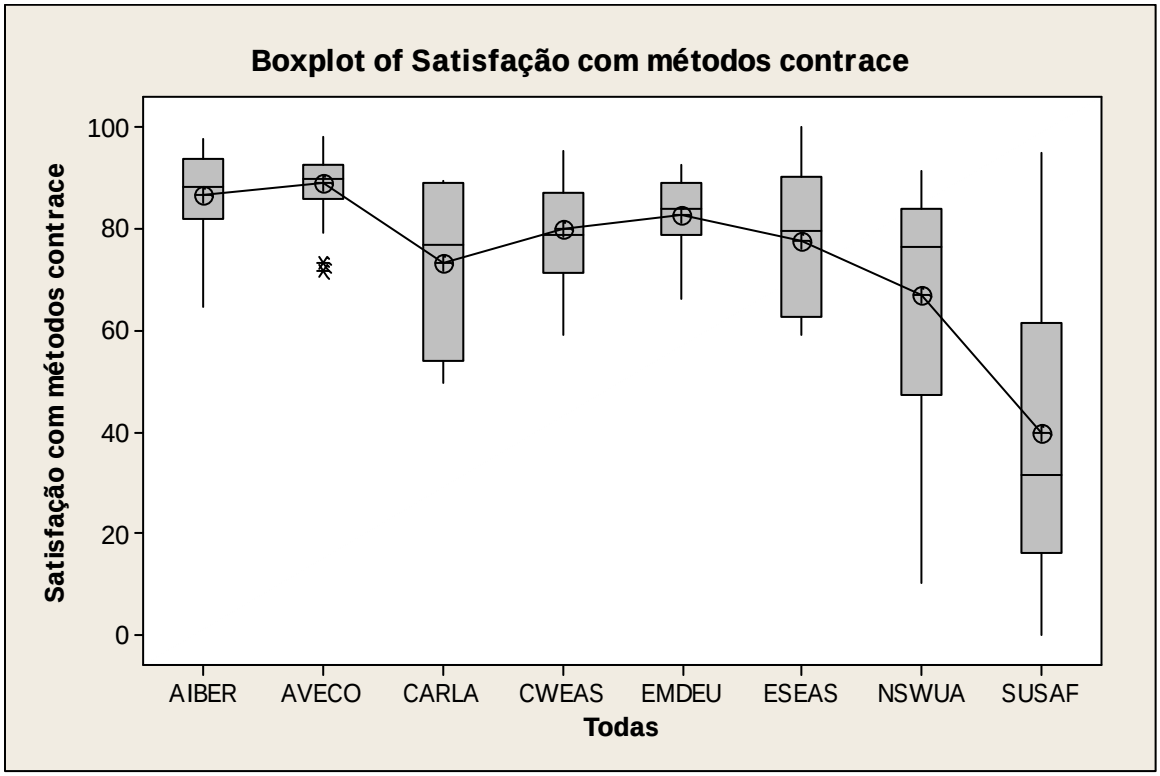


One-way ANOVA: Não escravidão moderna, tráfico versus Grupos

```
Source   DF       SS      MS      F      P
Grupos    2   3088,6  1544,3  15,49  0,000
Error   129  12858,4    99,7
Total   131  15947,0

S = 9,984   R-Sq = 19,37%   R-Sq(adj) = 18,12%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev    -+---------+---------+---------+--------
AIBER   21  94,54   3,29               (--------*--------)
AVECO   27  98,03   2,96                        (-------*-------)
OTHERS  84  86,69  12,23    (---*----)
                             -+---------+---------+---------+--------
                           85,0      90,0      95,0     100,0

Pooled StDev = 9,98
```

Boxplot of Não escravidão moderna, tráfico

1.1.17. Satisfação com métodos contraceptivos



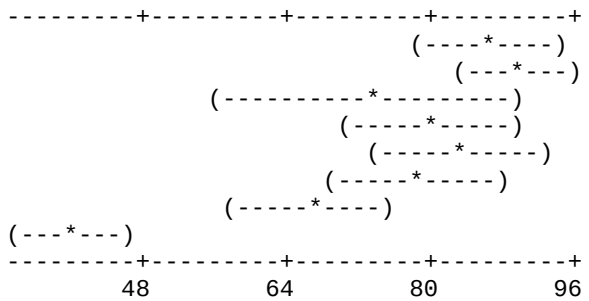
One-way ANOVA: Satisfação com métodos contrace versus Todas

Source	DF	SS	MS	F	P
Todas	7	47885	6841	24,14	0,000
Error	124	35137	283		
Total	131	83022			

S = 16,83 R-Sq = 57,68% R-Sq(adj) = 55,29%

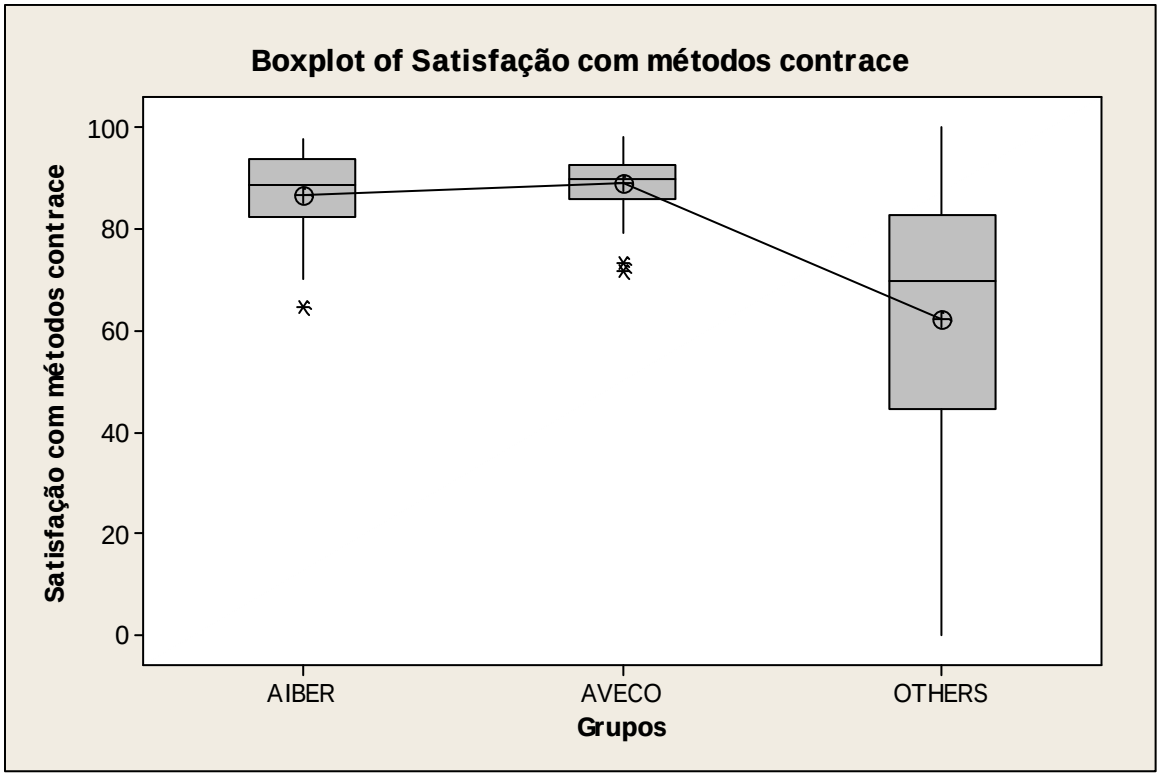
Level	N	Mean	StDev
AIBER	20	86,55	9,03
AVECO	27	88,99	6,47
CARLA	4	73,19	18,82
CWEAS	11	79,91	10,33
EMDEU	12	82,55	8,25
ESEAS	12	77,71	14,17
NSWUA	14	66,98	22,77
SUSAF	32	39,83	25,83

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev = 16,83

Boxplot of Satisfação com métodos contrace

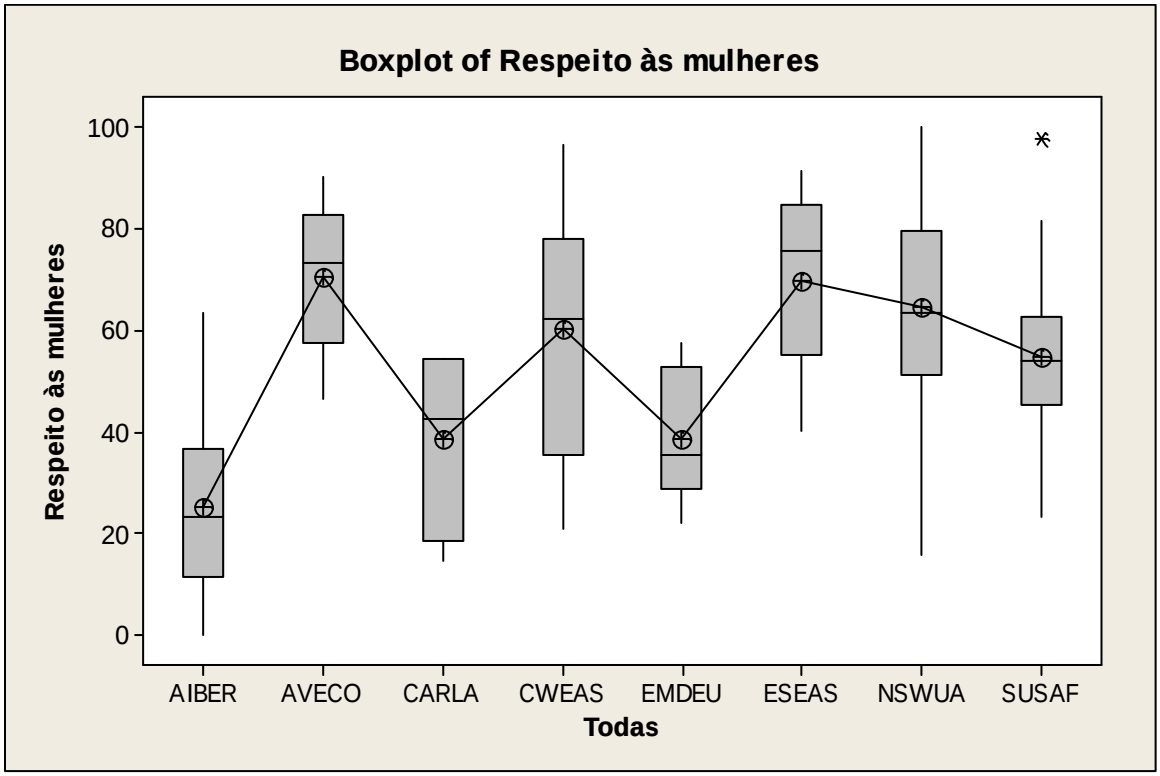


One-way ANOVA: Satisfação com métodos contrace versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  20502  10251  21,15  0,000
Error   129  62520    485
Total   131  83022

S = 22,01   R-Sq = 24,69%   R-Sq(adj) = 23,53%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --+---------+---------+---------+-------
AIBER   21  86,69   8,82                  (-------*-------)
AVECO   27  88,99   6,47                     (------*------)
OTHERS  84  62,11  26,86  (---*---)
                          --+---------+---------+---------+-------
                           60        72        84        96

Pooled StDev = 22,01
```

Boxplot of Satisfação com métodos contrace

1.1.18. Respeito às mulheres

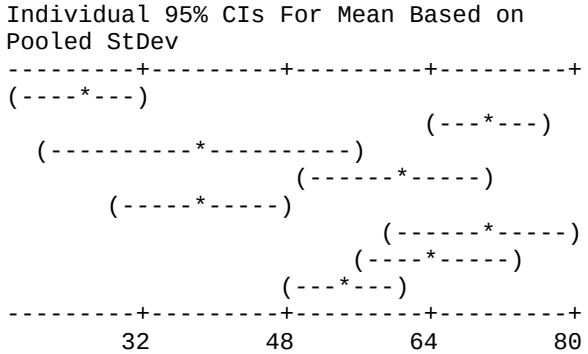


One-way ANOVA: Respeito às mulheres versus Todas

Source	DF	SS	MS	F	P
Todas	7	32505	4644	15,86	0,000
Error	124	36308	293		
Total	131	68813			

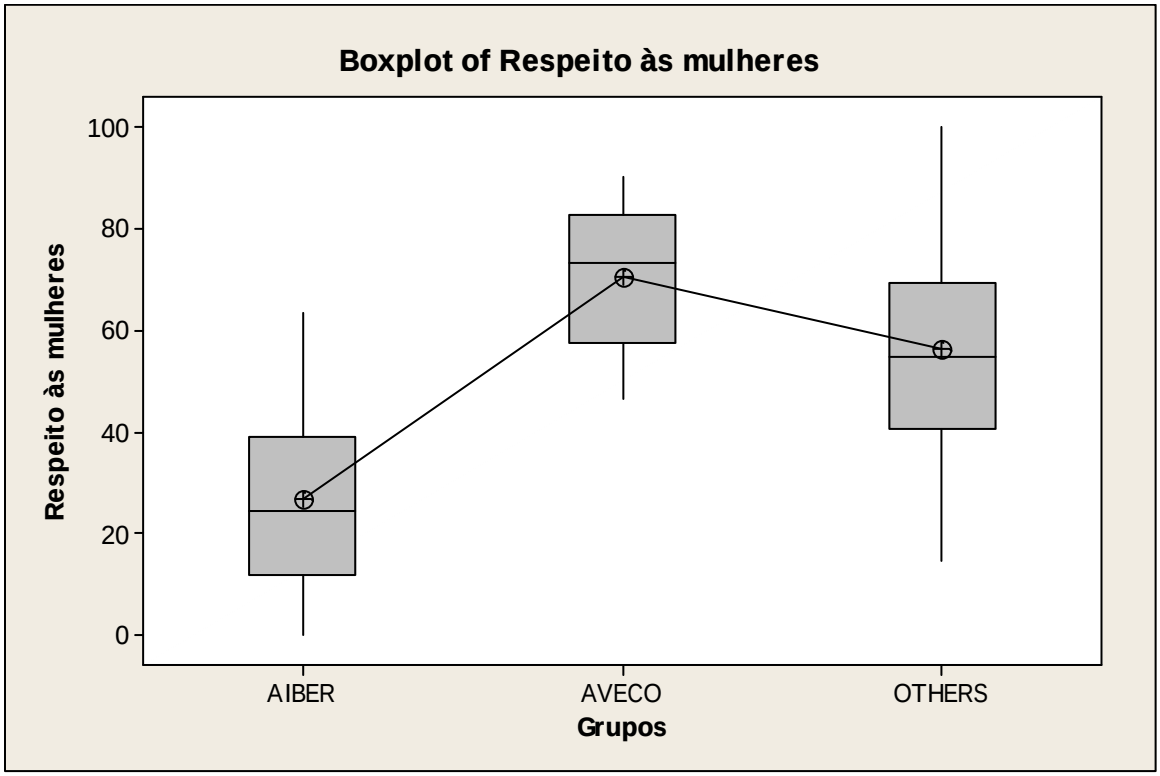
S = 17,11 R-Sq = 47,24% R-Sq(adj) = 44,26%

Level	N	Mean	StDev
AIBER	20	25,18	17,17
AVECO	27	70,33	13,59
CARLA	4	38,48	19,48
CWEAS	11	60,42	23,44
EMDEU	12	38,72	12,41
ESEAS	12	69,61	17,63
NSWUA	14	64,55	22,05
SUSAF	32	54,67	15,90



Pooled StDev = 17,11

Boxplot of Respeito às mulheres

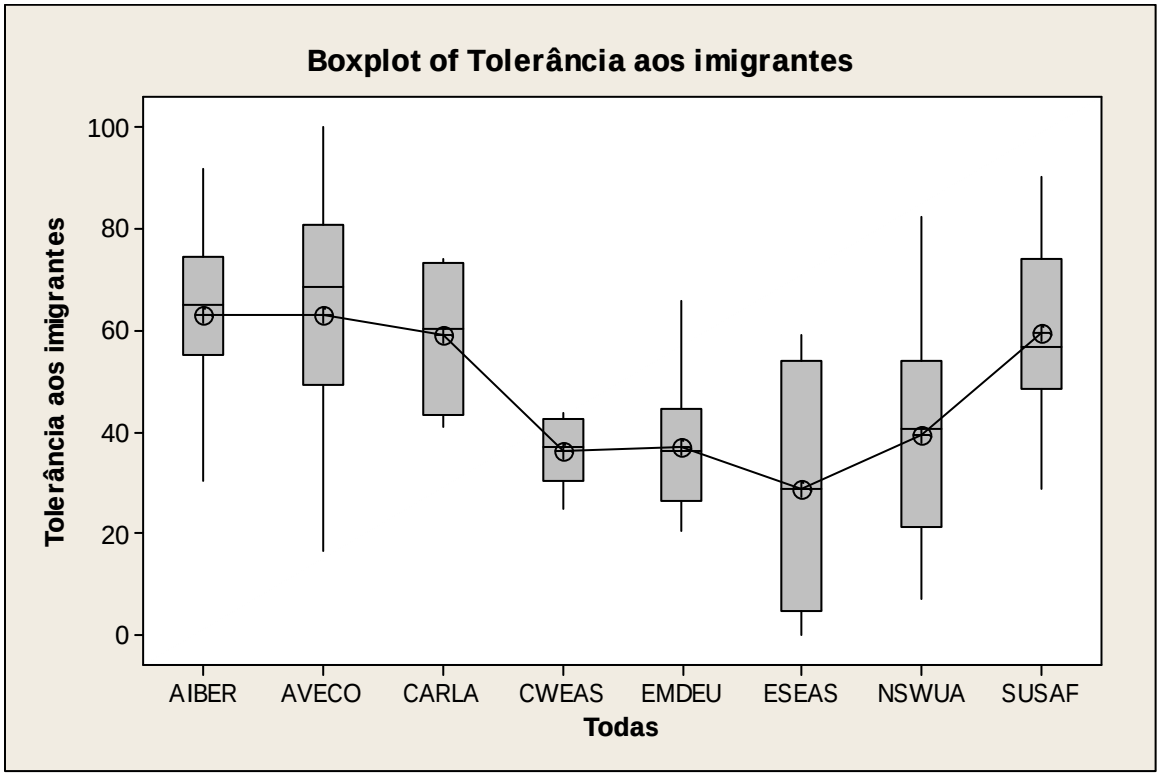


One-way ANOVA: Respeito às mulheres versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  23367  11684  33,16  0,000
Error   129  45445    352
Total   131  68813

S = 18,77   R-Sq = 33,96%   R-Sq(adj) = 32,93%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  26,57  17,90  (-----*----)
AVECO   27  70,33  13,59                              (----*----)
OTHERS  84  56,16  20,31                        (-*--)
                          --------+---------+---------+---------+-
                                 30        45        60        75

Pooled StDev = 18,77
```

Boxplot of Respeito às mulheres

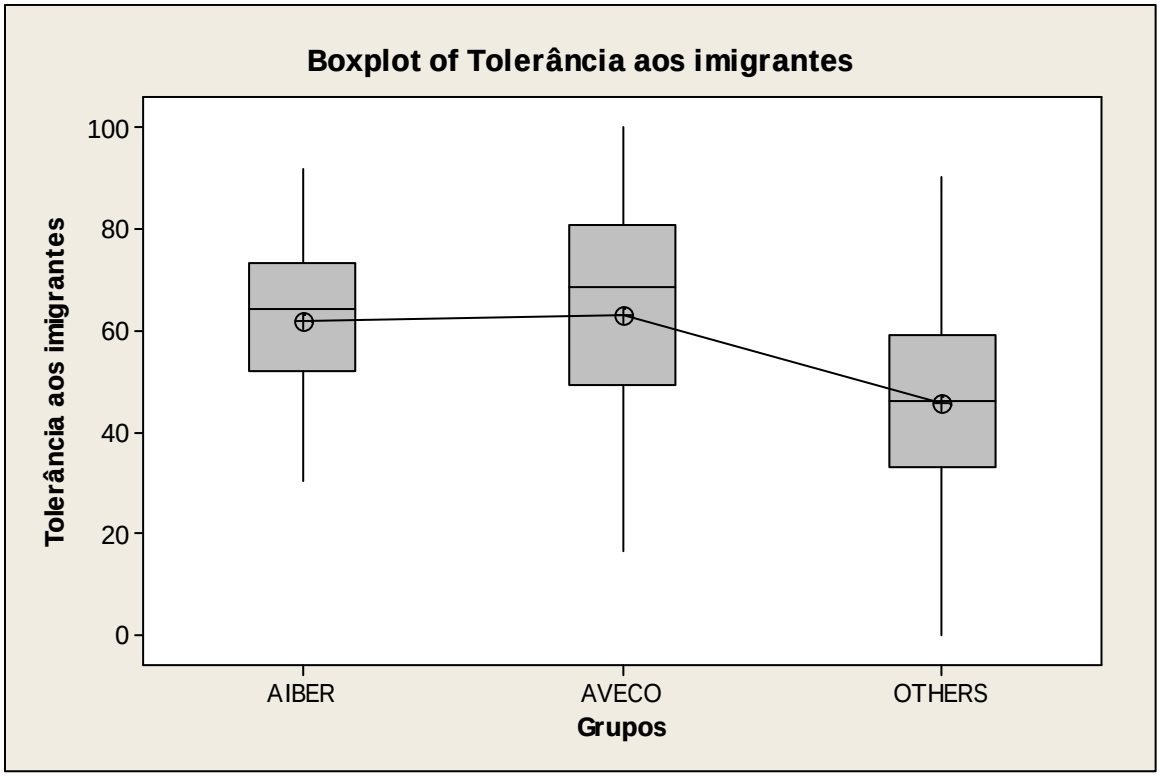
1.1.19. Tolerância ao Imigrante



One-way ANOVA: Tolerância aos imigrantes versus Todas

```
Source   DF     SS    MS     F      P
Todas     7  21917  3131  8,91  0,000
Error   124  43566   351
Total   131  65483

S = 18,74   R-Sq = 33,47%   R-Sq(adj) = 29,71%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  20  63,01  15,94                         (-----*-----)
AVECO  27  62,86  24,27                          (----*----)
CARLA   4  58,90  16,21                (-----------*------------)
CWEAS  11  36,11   6,46      (------*-------)
EMDEU  12  37,10  12,67       (------*------)
ESEAS  12  28,54  23,08  (------*------)
NSWUA  14  39,43  22,78         (-----*------)
SUSAF  32  59,63  15,91                        (----*---)
                          --------+---------+---------+---------+-
                                 30        45        60        75

Pooled StDev = 18,74
```

Boxplot of Tolerância aos imigrantes

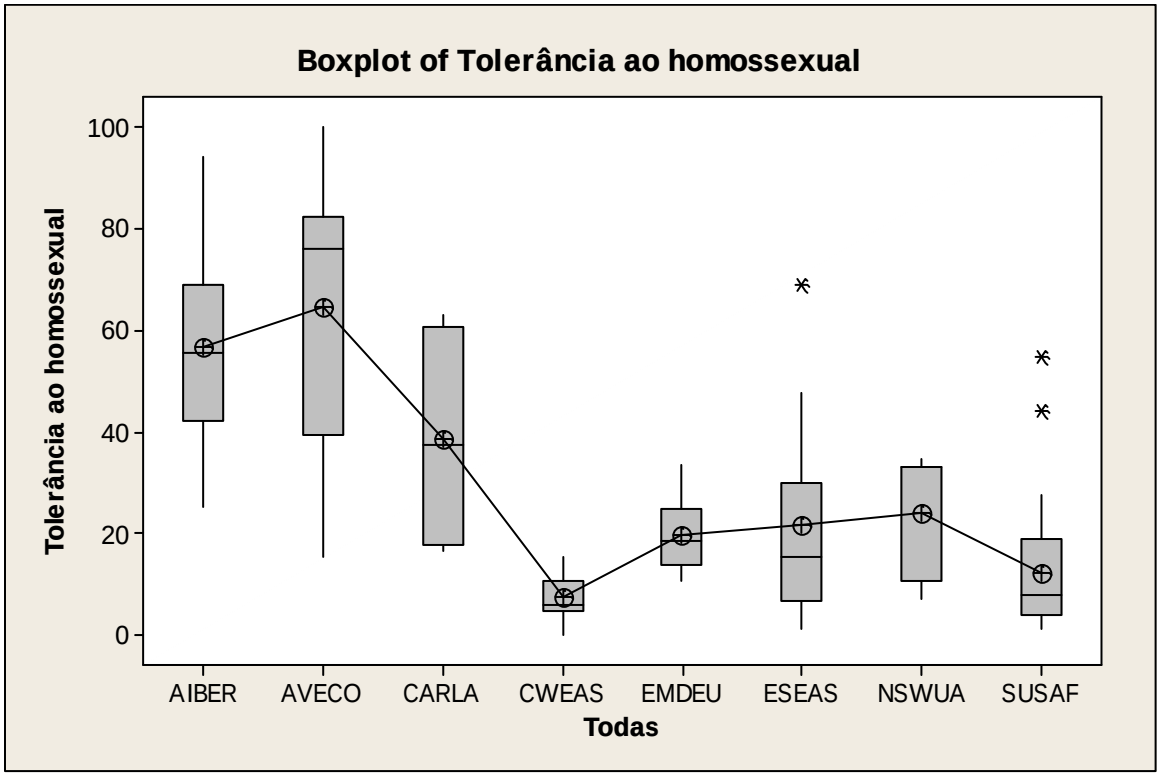


One-way ANOVA: Tolerância aos imigrantes versus Grupos

```
Source   DF     SS    MS     F      P
Grupos    2   8590  4295  9,74  0,000
Error   129  56893   441
Total   131  65483

S = 21,00   R-Sq = 13,12%   R-Sq(adj) = 11,77%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  61,97  16,25                (----------*----------)
AVECO   27  62,86  24,27                   (---------*---------)
OTHERS  84  45,71  20,91  (-----*-----)
                          ---------+---------+---------+---------+
                                48,0      56,0      64,0      72,0

Pooled StDev = 21,00
```

Boxplot of Tolerância aos imigrantes

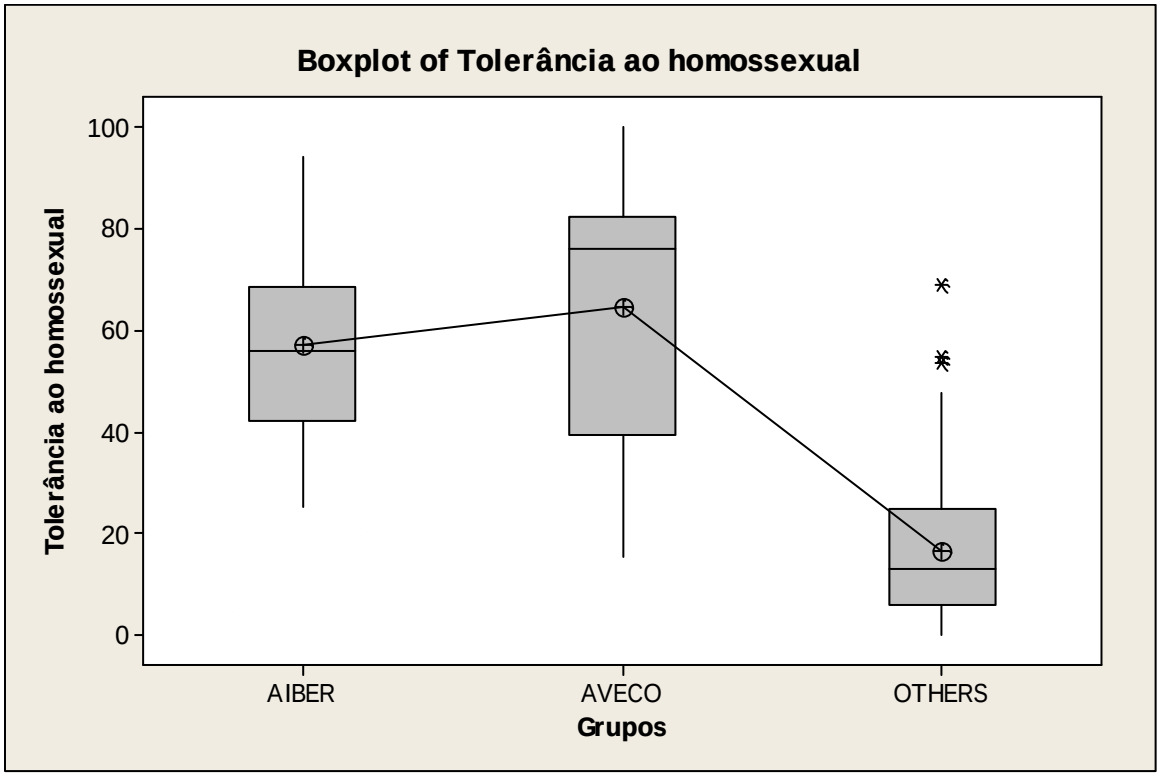
1.1.20. Tolerância ao Homosexual



One-way ANOVA: Tolerância ao homossexual versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  64105  9158  31,70  0,000
Error   124  35817   289
Total   131  99923

S = 17,00   R-Sq = 64,15%   R-Sq(adj) = 62,13%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  56,85  17,83                            (--*---)
AVECO  27  64,51  25,29                                (--*--)
CARLA   4  38,69  23,09                      (-------*--------)
CWEAS  11   7,25   4,59  (----*----)
EMDEU  12  19,74   6,69        (----*----)
ESEAS  12  21,73  19,95         (----*----)
NSWUA  14  24,00  11,53           (---*---)
SUSAF  32  12,28  12,37      (--*--)
                          -+---------+---------+---------+--------
                           0        20        40        60

Pooled StDev = 17,00
```

Boxplot of Tolerância ao homossexual

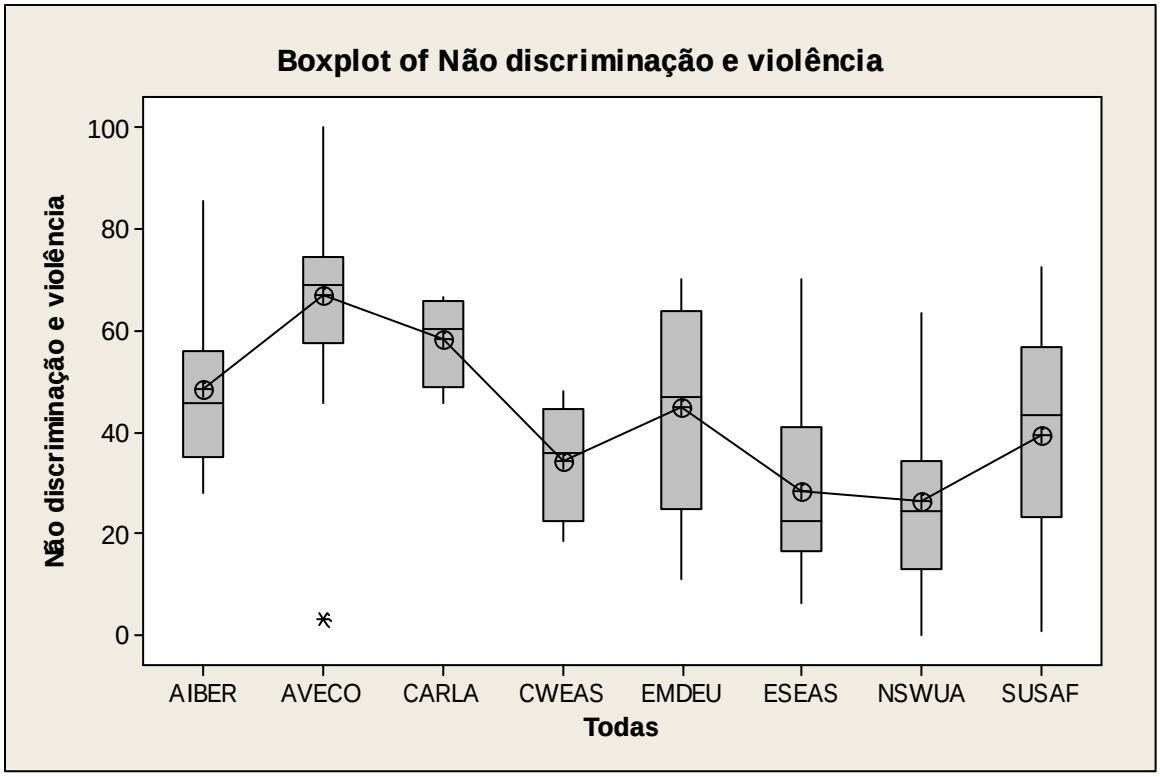


One-way ANOVA: Tolerância ao homossexual versus Grupos

```
Source   DF     SS     MS       F      P
Grupos    2  61519  30760  103,32  0,000
Error   129  38404    298
Total   131  99923

S = 17,25   R-Sq = 61,57%   R-Sq(adj) = 60,97%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  57,14  17,43                        (----*----)
AVECO   27  64,51  25,29                              (---*---)
OTHERS  84  16,64  13,76  (-*--)
                          -+---------+---------+---------+--------
                          15        30        45        60

Pooled StDev = 17,25
```

Boxplot of Tolerância ao homossexual

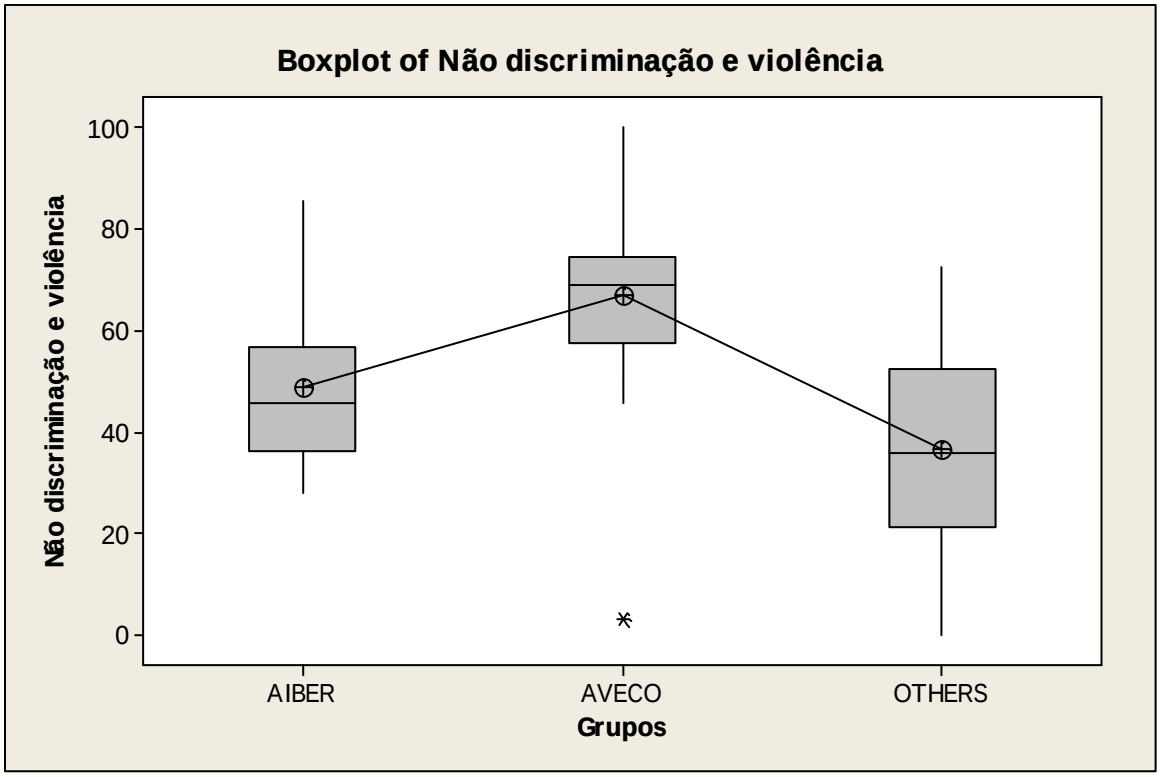
1.1.21. Não discriminação e violência



One-way ANOVA: Não discriminação e violência versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  24646  3521  10,67  0,000
Error   124  40915   330
Total   131  65561

S = 18,16   R-Sq = 37,59%   R-Sq(adj) = 34,07%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---------+---------+---------+---------+
AIBER  20  48,51  16,36                 (----*-----)
AVECO  27  67,10  19,37                           (----*---)
CARLA   4  58,17   9,06                 (-----------*-----------)
CWEAS  11  34,26  10,61       (------*------)
EMDEU  12  44,86  20,39              (------*------)
ESEAS  12  28,24  17,60    (------*------)
NSWUA  14  26,21  18,64   (-----*------)
SUSAF  32  39,55  19,84             (---*----)
                          ---------+---------+---------+---------+
                                  30        45        60        75

Pooled StDev = 18,16
```

Boxplot of Não discriminação e violência

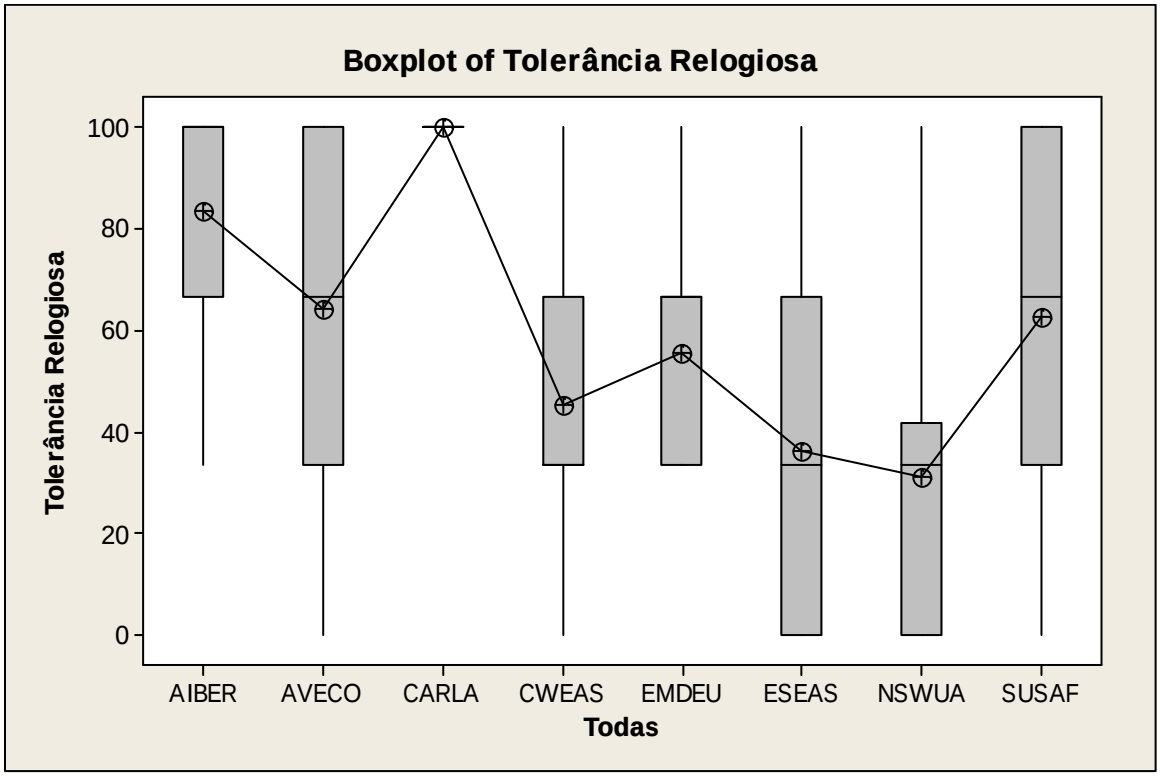


One-way ANOVA: Não discriminação e violência versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  19650  9825  27,61  0,000
Error   129  45911   356
Total   131  65561

S = 18,87   R-Sq = 29,97%   R-Sq(adj) = 28,89%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  48,96  16,08        (------*------)
AVECO   27  67,10  19,37                      (-----*-----)
OTHERS  84  36,45  19,32  (--*---)
                          ---+---------+---------+---------+------
                            36        48        60        72

Pooled StDev = 18,87
```

Boxplot of Não discriminação e violência

1.1.22. Tolerância Religiosa

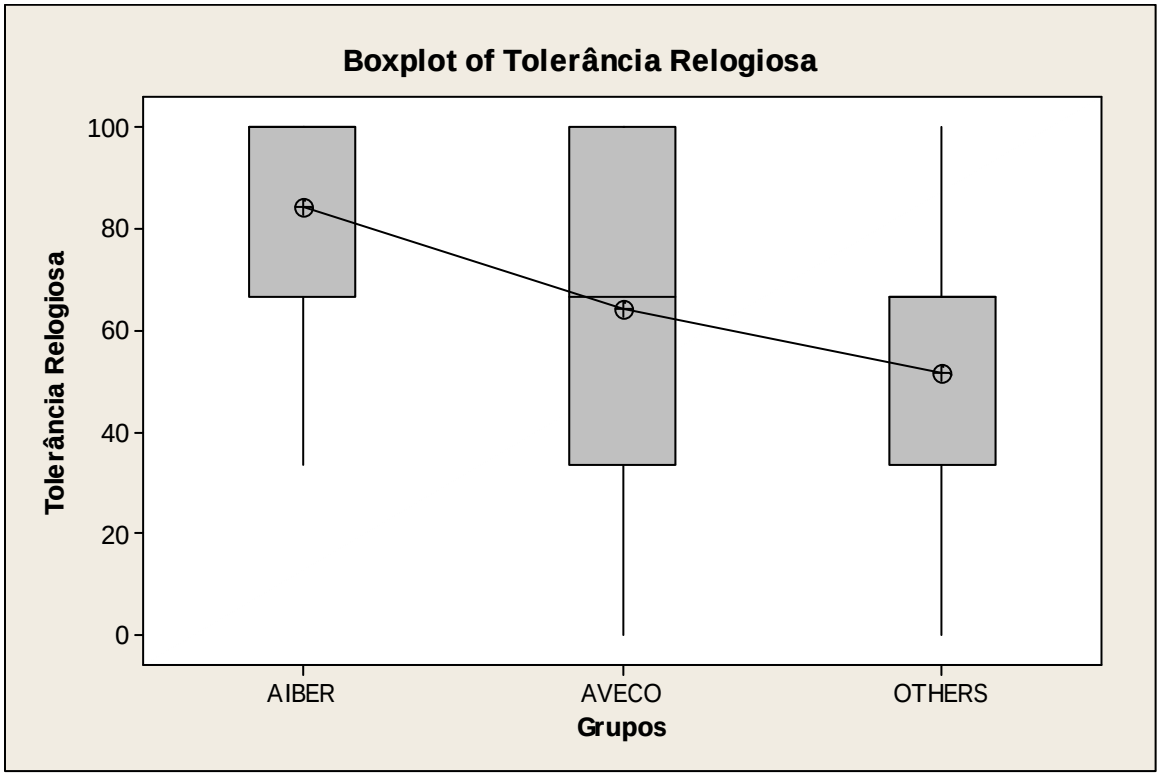


One-way ANOVA: Tolerância Relogiosa versus Todas

```
Source   DF      SS    MS     F      P
Todas     7   39133  5590  6,83  0,000
Error   124  101566   819
Total   131  140699

S = 28,62   R-Sq = 27,81%   R-Sq(adj) = 23,74%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  -----+---------+---------+---------+----
AIBER  20   83,33  22,94                    (---*---)
AVECO  27   64,20  27,62               (--*---)
CARLA   4  100,00   0,00                    (--------*---------)
CWEAS  11   45,45  26,97      (-----*-----)
EMDEU  12   55,56  21,71           (-----*----)
ESEAS  12   36,11  38,82    (----*----)
NSWUA  14   30,95  30,56  (----*----)
SUSAF  32   62,50  31,40              (---*--)
                          -----+---------+---------+---------+----
                              30        60        90       120

Pooled StDev = 28,62
```

Boxplot of Tolerância Relogiosa

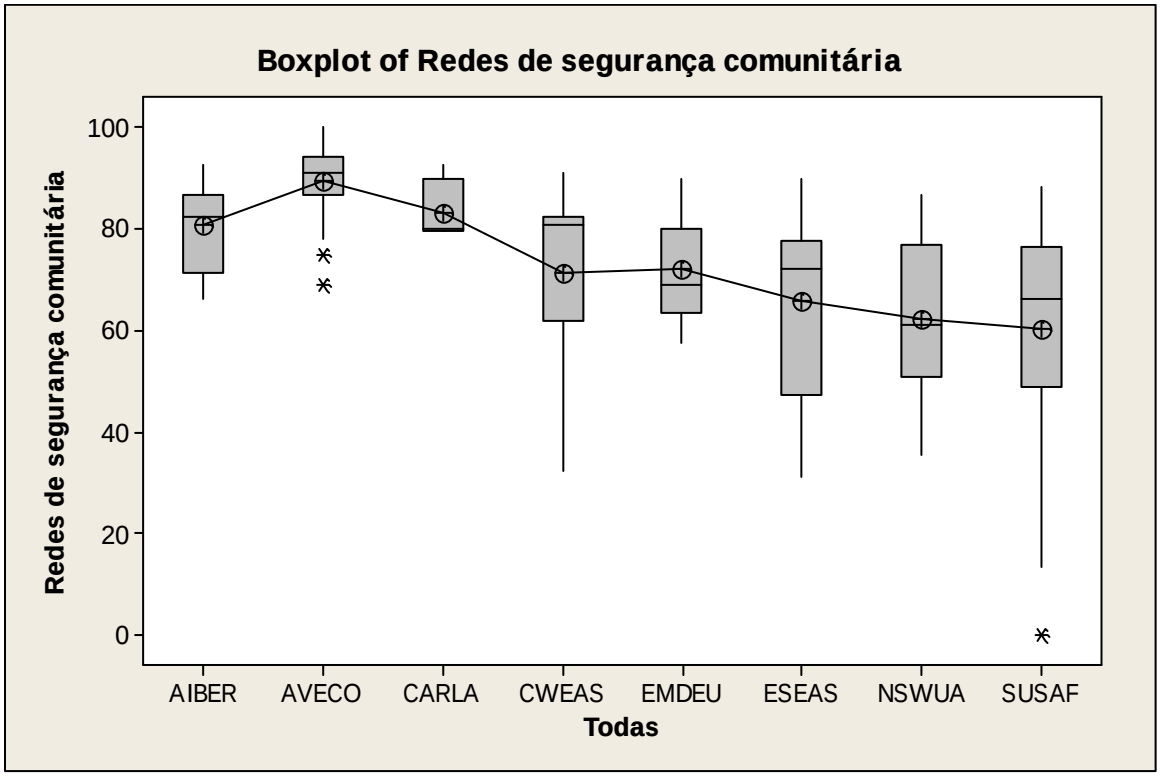


One-way ANOVA: Tolerância Relogiosa versus Grupos

```
Source   DF      SS    MS     F      P
Grupos    2   18588  9294  9,82  0,000
Error   129  122111   947
Total   131  140699

S = 30,77   R-Sq = 13,21%   R-Sq(adj) = 11,87%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  84,13  22,65                   (--------*--------)
AVECO   27  64,20  27,62        (-------*-------)
OTHERS  84  51,59  33,30   (---*----)
                           +---------+---------+---------+---------
                          45        60        75        90

Pooled StDev = 30,77
```

Boxplot of Tolerância Relogiosa

1.1.23. Redes de segurança comunitária

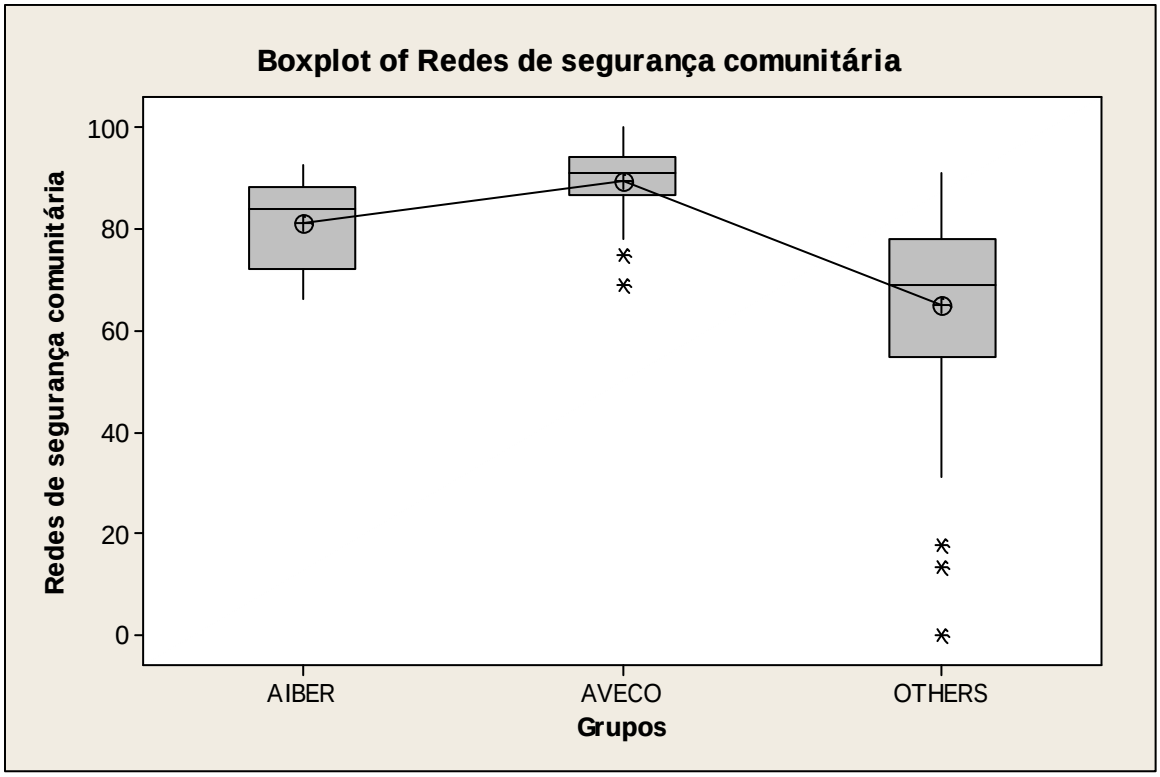


One-way ANOVA: Redes de segurança comunitária versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  16530  2361  10,22  0,000
Error   124  28639   231
Total   131  45169

S = 15,20   R-Sq = 36,60%   R-Sq(adj) = 33,02%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -----+---------+---------+---------+----
AIBER  20  80,74   8,99                  (----*-----)
AVECO  27  89,43   6,99                        (----*---)
CARLA   4  83,09   6,41             (-----------*-----------)
CWEAS  11  71,39  17,23         (------*-------)
EMDEU  12  71,94  10,48          (-------*------)
ESEAS  12  65,81  18,88     (------*------)
NSWUA  14  62,08  15,28  (------*-----)
SUSAF  32  60,11  21,68   (---*----)
                          -----+---------+---------+---------+----
                              60        72        84        96

Pooled StDev = 15,20
```

Boxplot of Redes de segurança comunitária

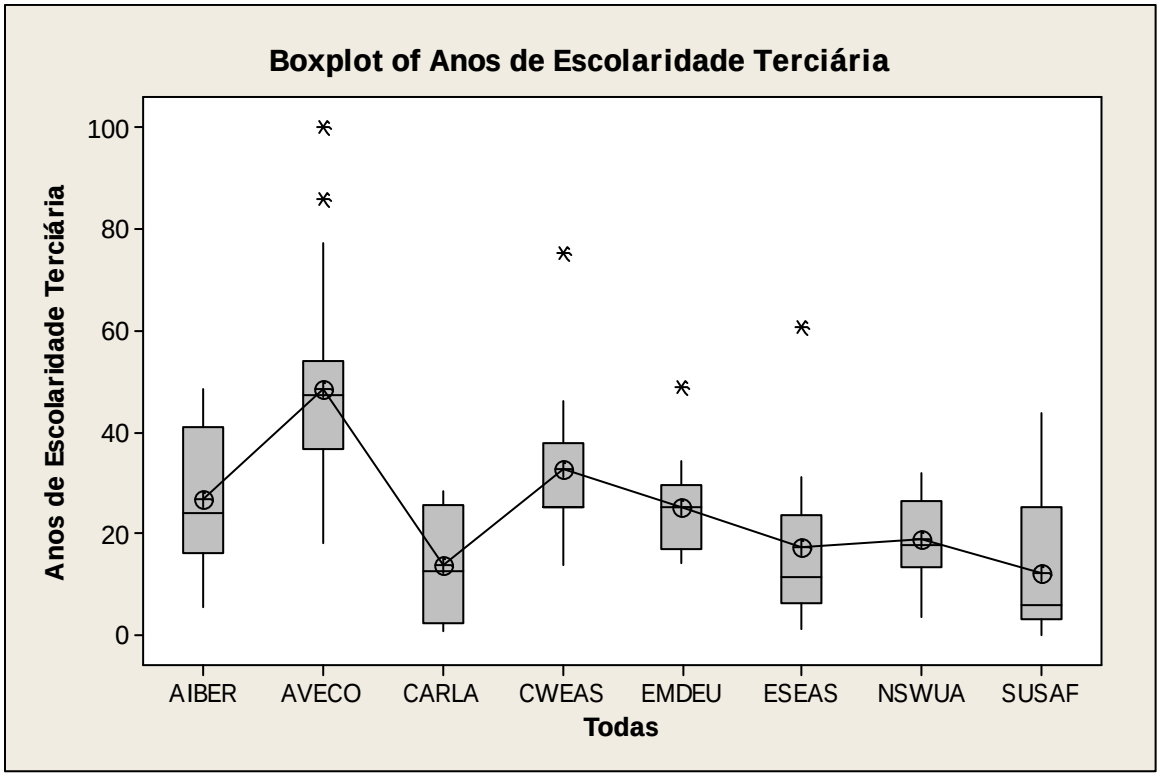


One-way ANOVA: Redes de segurança comunitária versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  13933  6966  28,77  0,000
Error   129  31236   242
Total   131  45169

S = 15,56   R-Sq = 30,85%   R-Sq(adj) = 29,77%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  81,30   9,14               (-----*------)
AVECO   27  89,43   6,99                         (----*-----)
OTHERS  84  65,13  18,46  (--*--)
                          --------+---------+---------+---------+-
                                 70        80        90       100

Pooled StDev = 15,56
```

Boxplot of Redes de segurança comunitária

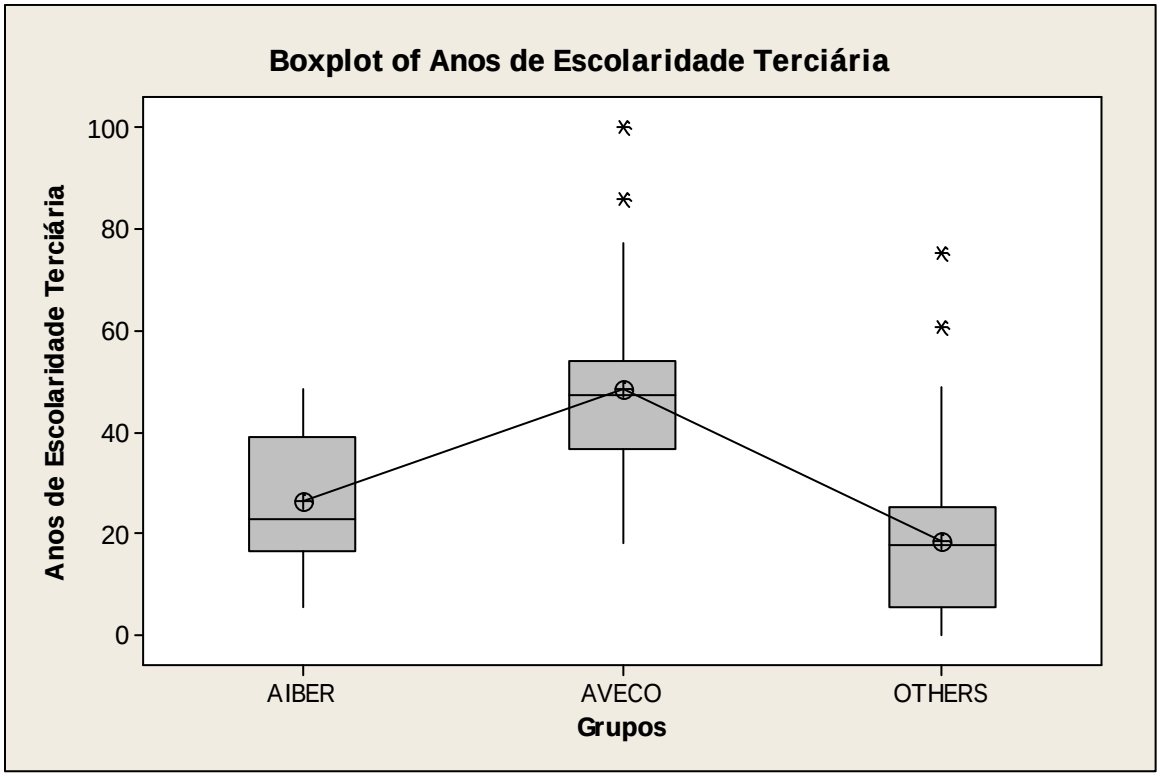
1.1.24. Anos de Escolaridade Terciária



One-way ANOVA: Anos de Escolaridade Terciária versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  22460  3209  15,35  0,000
Error   124  25914   209
Total   131  48373

S = 14,46   R-Sq = 46,43%   R-Sq(adj) = 43,41%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  26,85  13,40                (---*---)
AVECO  27  48,39  19,50                          (--*---)
CARLA   4  13,53  12,05   (---------*---------)
CWEAS  11  32,67  16,45                   (-----*-----)
EMDEU  12  25,25   9,60              (-----*----)
ESEAS  12  17,40  16,31         (-----*----)
NSWUA  14  18,82   8,62          (-----*----)
SUSAF  32  12,10  12,14        (--*--)
                          -+---------+---------+---------+--------
                           0        15        30        45

Pooled StDev = 14,46
```

Boxplot of Anos de Escolaridade Terciária



One-way ANOVA: Anos de Escolaridade Terciária versus Grupos

Source	DF	SS	MS	F	P
Grupos	2	18197	9098	38,89	0,000
Error	129	30176	234		
Total	131	48373			

S = 15,29 R-Sq = 37,62% R-Sq(adj) = 36,65%

Level	N	Mean	StDev
AIBER	21	26,41	13,21
AVECO	27	48,39	19,50
OTHERS	84	18,55	14,23

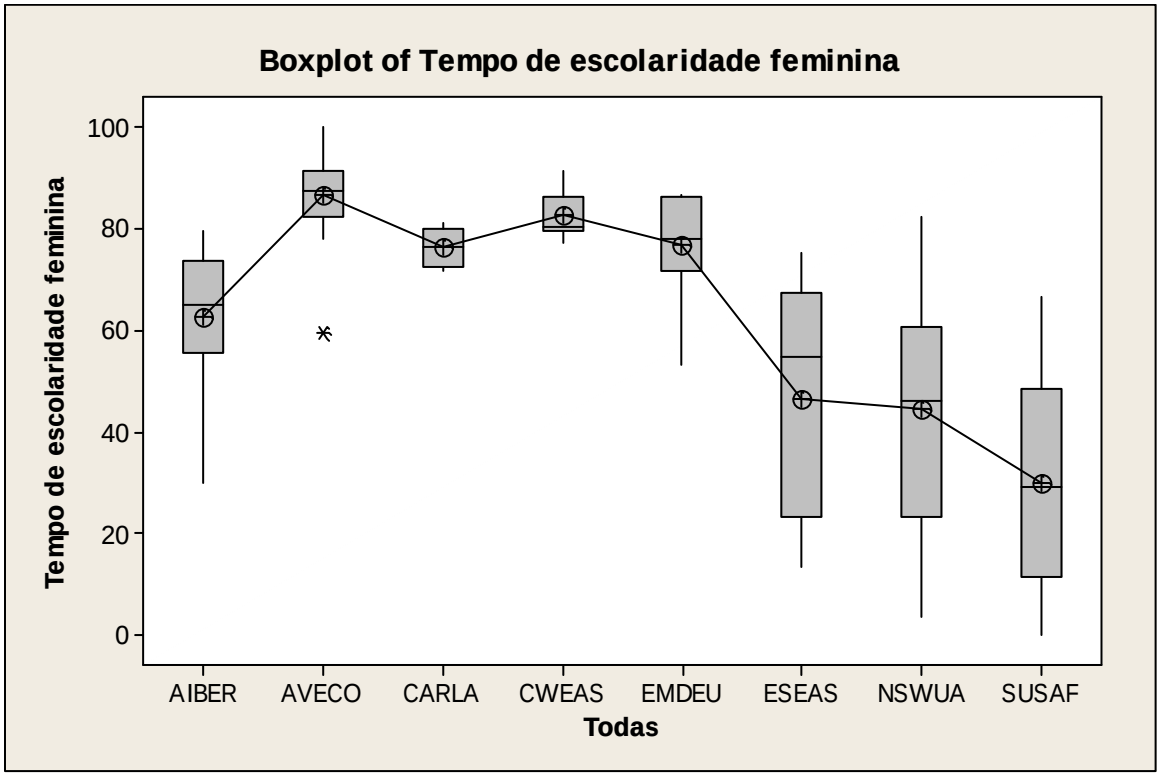
Individual 95% CIs For Mean Based on Pooled StDev

```
-----+---------+---------+---------+----
     (-----*------)
                               (----*-----)
(---*--)
-----+---------+---------+---------+----
    20        30        40        50
```

Pooled StDev = 15,29

Boxplot of Anos de Escolaridade Terciária

1.1.25. Tempo de escolaridade feminina

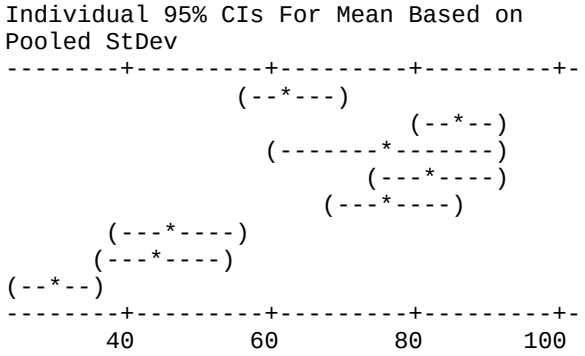


One-way ANOVA: Tempo de escolaridade feminina versus Todas

Source	DF	SS	MS	F	P
Todas	7	64364	9195	34,92	0,000
Error	124	32651	263		
Total	131	97015			

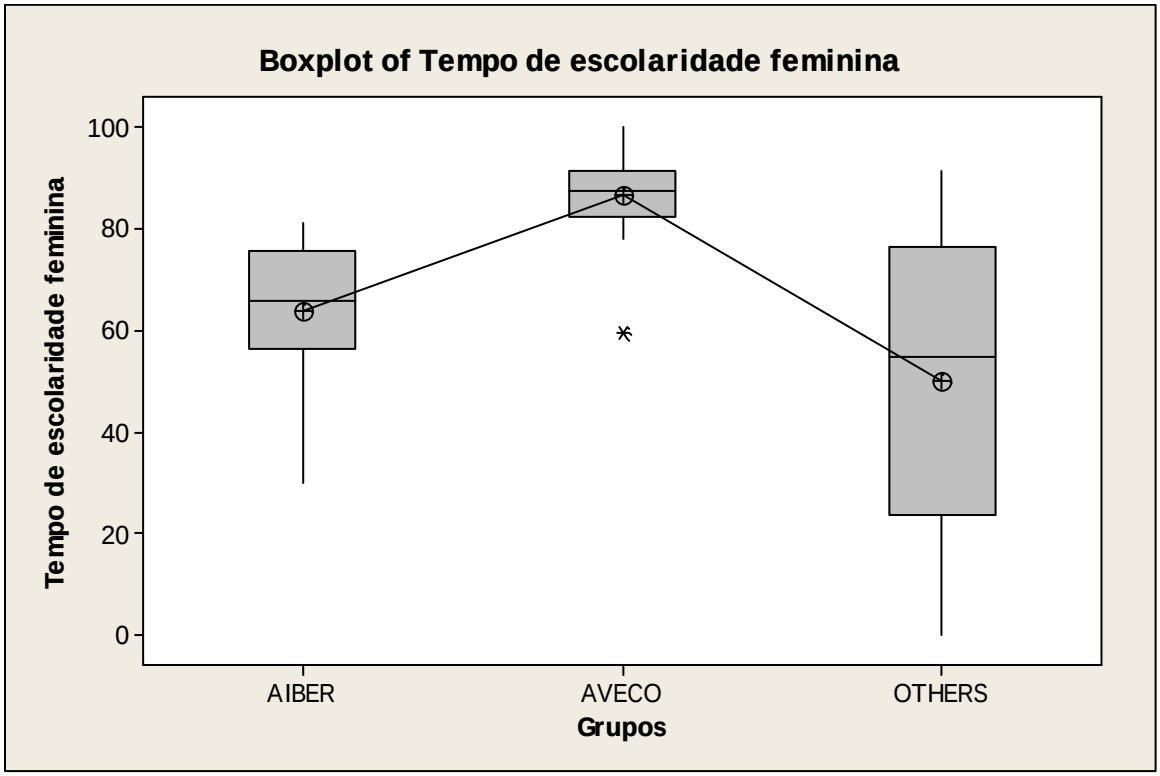
S = 16,23 R-Sq = 66,34% R-Sq(adj) = 64,44%

Level	N	Mean	StDev
AIBER	20	62,74	12,71
AVECO	27	86,80	7,72
CARLA	4	76,28	3,89
CWEAS	11	82,75	4,38
EMDEU	12	76,89	10,31
ESEAS	12	46,53	22,24
NSWUA	14	44,58	23,65
SUSAF	32	29,81	21,19



Pooled StDev = 16,23

Boxplot of Tempo de escolaridade feminina



One-way ANOVA: Tempo de escolaridade feminina versus Grupos

Source	DF	SS	MS	F	P
Grupos	2	28180	14090	26,40	0,000
Error	129	68835	534		
Total	131	97015			

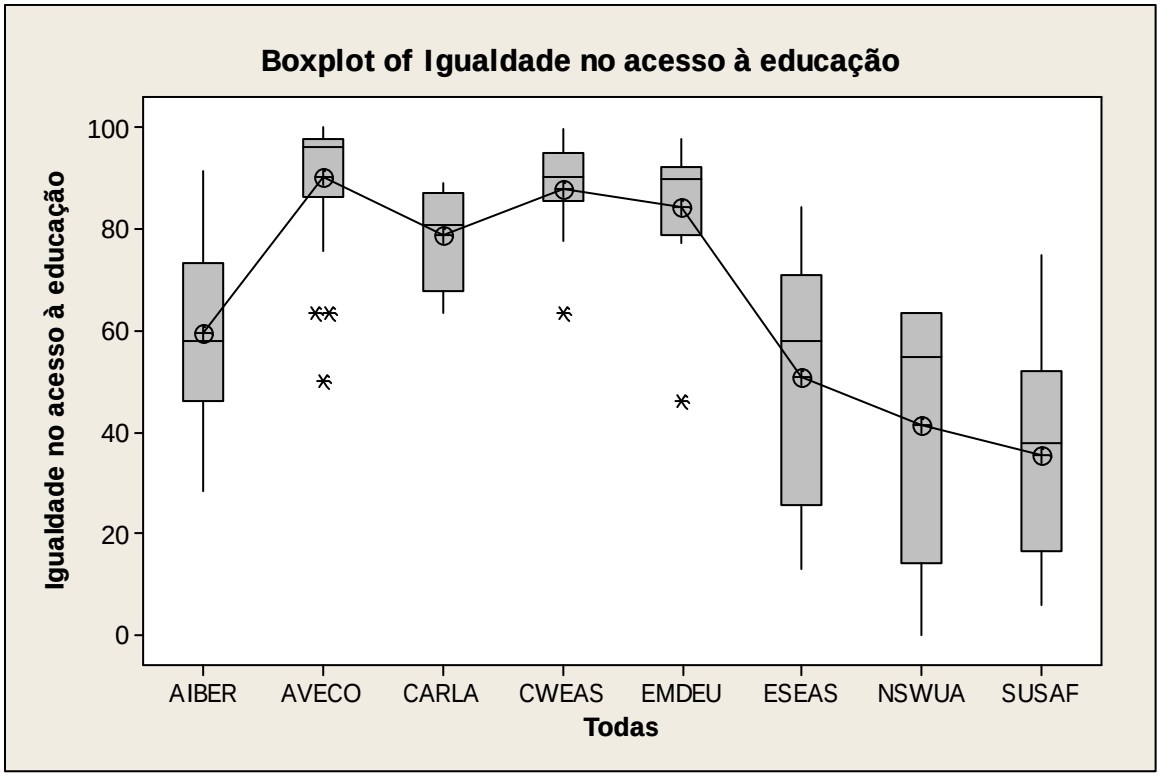
S = 23,10 R-Sq = 29,05% R-Sq(adj) = 27,95%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	63,61	13,02	(-----*------)
AVECO	27	86,80	7,72	(-----*-----)
OTHERS	84	49,92	27,75	(--*---)
				45 60 75 90

Pooled StDev = 23,10

Boxplot of Tempo de escolaridade feminina

1.1.26. Igualdade no acesso à educação

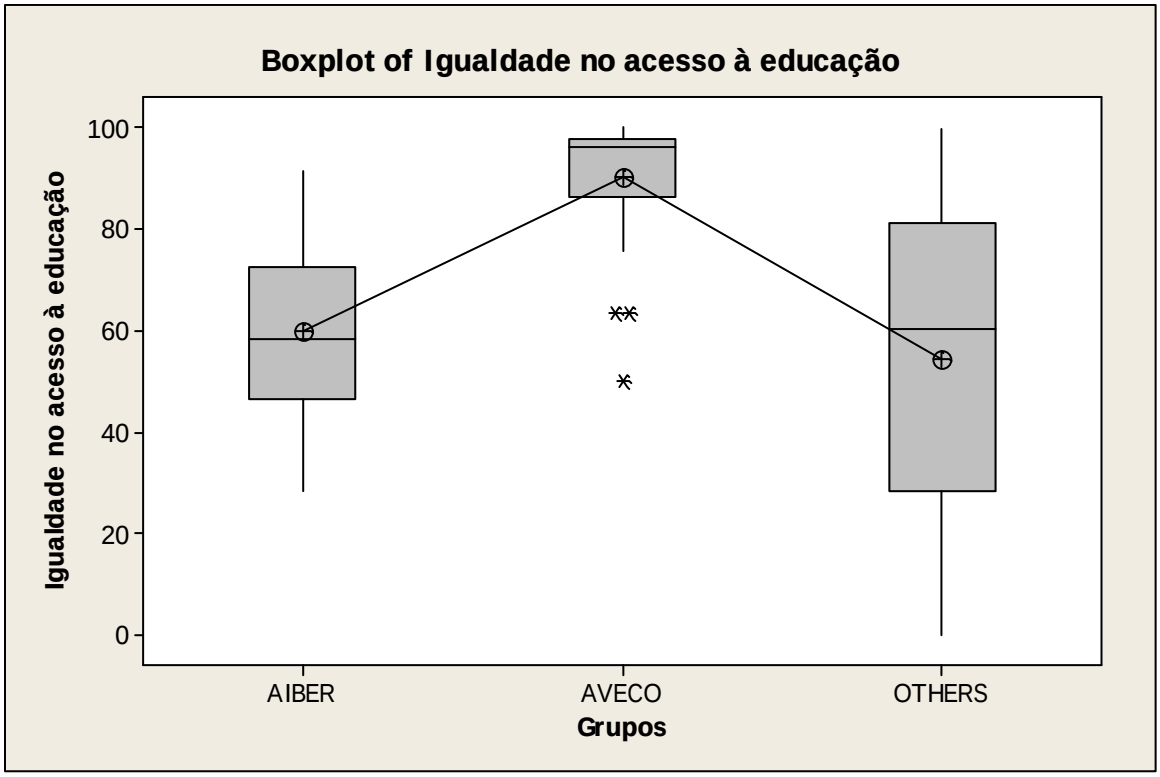


One-way ANOVA: Igualdade no acesso à educação versus Todas

```
Source   DF      SS    MS      F      P
Todas     7   65950  9421  28,33  0,000
Error   124   41230   333
Total   131  107181

S = 18,23   R-Sq = 61,53%   R-Sq(adj) = 59,36%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
AIBER  20  59,58  17,93            (---*---)
AVECO  27  90,06  12,98                       (--*---)
CARLA   4  78,61  10,78                  (--------*--------)
CWEAS  11  87,89   9,95                           (----*----)
EMDEU  12  84,45  13,74                        (----*----)
ESEAS  12  50,70  24,19       (----*-----)
NSWUA  14  41,22  25,35    (----*---)
SUSAF  32  35,48  20,11   (--*--)
                          -----+---------+---------+---------+----
                              40        60        80       100

Pooled StDev = 18,23
```

Boxplot of Igualdade no acesso à educação

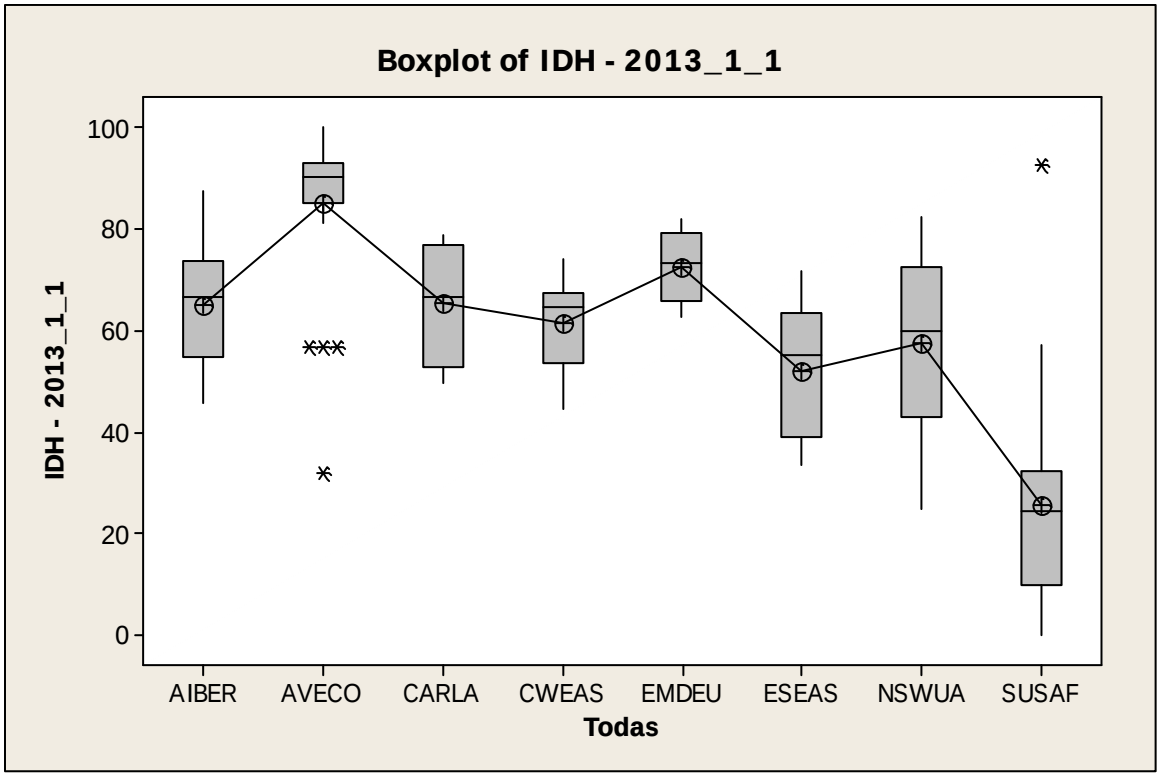


One-way ANOVA: Igualdade no acesso à educação versus Grupos

```
Source   DF      SS     MS      F      P
Grupos    2   26457  13229  21,14  0,000
Error   129   80723    626
Total   131  107181

S = 25,02   R-Sq = 24,68%   R-Sq(adj) = 23,52%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  59,76  17,50  (------*------)
AVECO   27  90,06  12,98                       (-----*-----)
OTHERS  84  54,19  29,09  (--*---)
                          -------+---------+---------+---------+--
                                60        75        90       105

Pooled StDev = 25,02
```

Boxplot of Igualdade no acesso à educação

1.1.27. IDH

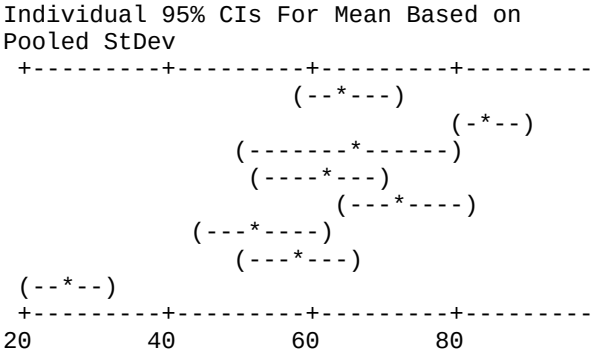


One-way ANOVA: IDH - 2013_1_1 versus Todas

Source	DF	SS	MS	F	P
Todas	7	57384	8198	33,97	0,000
Error	124	29924	241		
Total	131	87308			

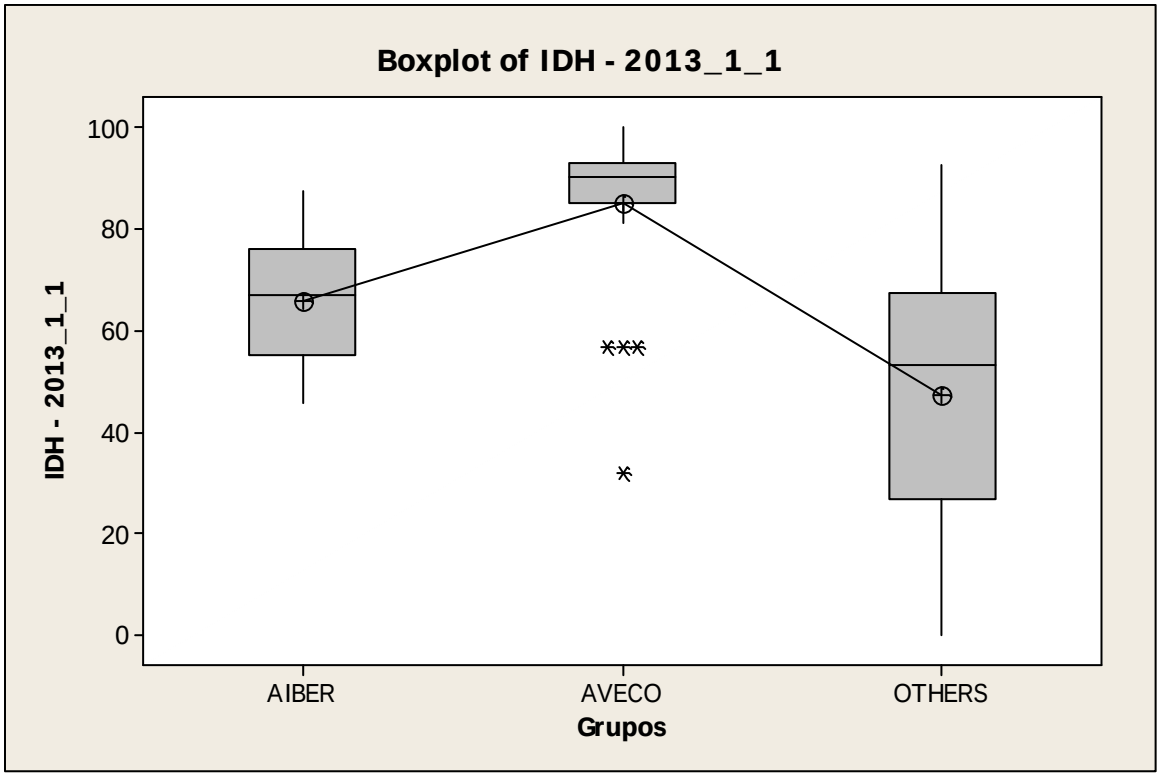
S = 15,53 R-Sq = 65,73% R-Sq(adj) = 63,79%

Level	N	Mean	StDev
AIBER	20	64,96	11,95
AVECO	27	84,94	15,85
CARLA	4	65,32	12,46
CWEAS	11	61,45	9,16
EMDEU	12	72,60	6,81
ESEAS	12	52,16	13,52
NSWUA	14	57,47	19,21
SUSAF	32	25,64	19,72



Pooled StDev = 15,53

Boxplot of IDH - 2013_1_1

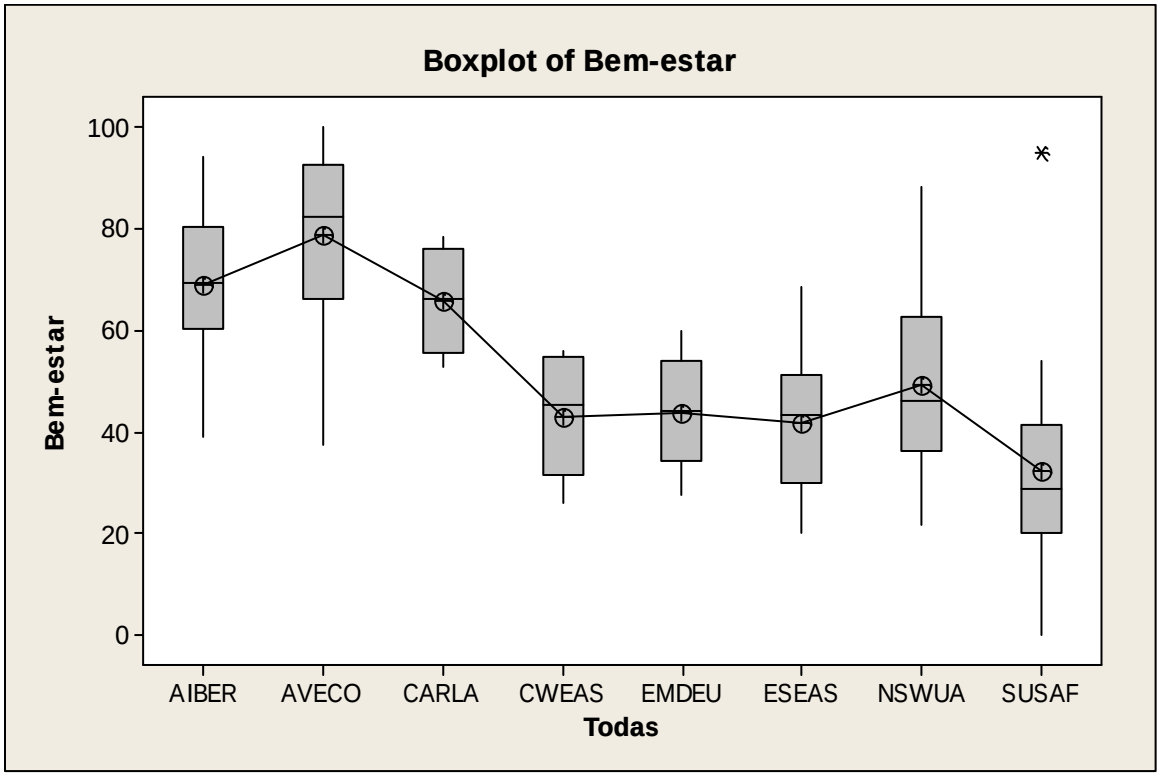


One-way ANOVA: IDH - 2013_1_1 versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  30272  15136  34,23  0,000
Error   129  57036    442
Total   131  87308

S = 21,03   R-Sq = 34,67%   R-Sq(adj) = 33,66%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65,62  12,03           (-----*-----)
AVECO   27  84,94  15,85                      (-----*----)
OTHERS  84  47,39  23,95  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 21,03
```

Boxplot of IDH - 2013_1_1

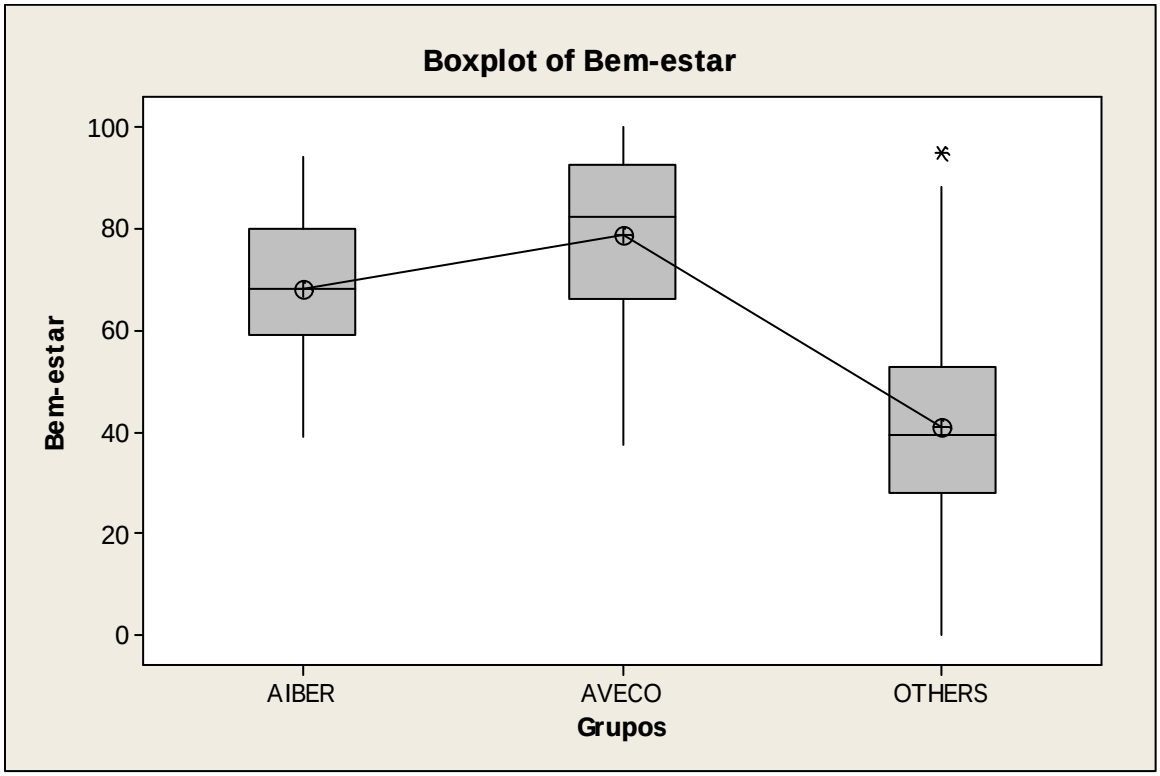
1.1.28. Índice de Bem-estar



One-way ANOVA: Bem-estar versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  41238  5891  23,36  0,000
Error   124  31272   252
Total   131  72510

S = 15,88   R-Sq = 56,87%   R-Sq(adj) = 54,44%


                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
AIBER  20  69,07  14,95                        (----*----)
AVECO  27  78,79  16,71                               (----*---)
CARLA   4  65,91  10,68                (----------*---------)
CWEAS  11  42,91  11,53     (------*-----)
EMDEU  12  43,83  10,91      (-----*-----)
ESEAS  12  41,83  13,52     (-----*-----)
NSWUA  14  49,23  20,17          (-----*----)
SUSAF  32  32,36  17,49  (---*--)
                         --+---------+---------+---------+-------
                          30        45        60        75

Pooled StDev = 15,88
```

Boxplot of Bem-estar

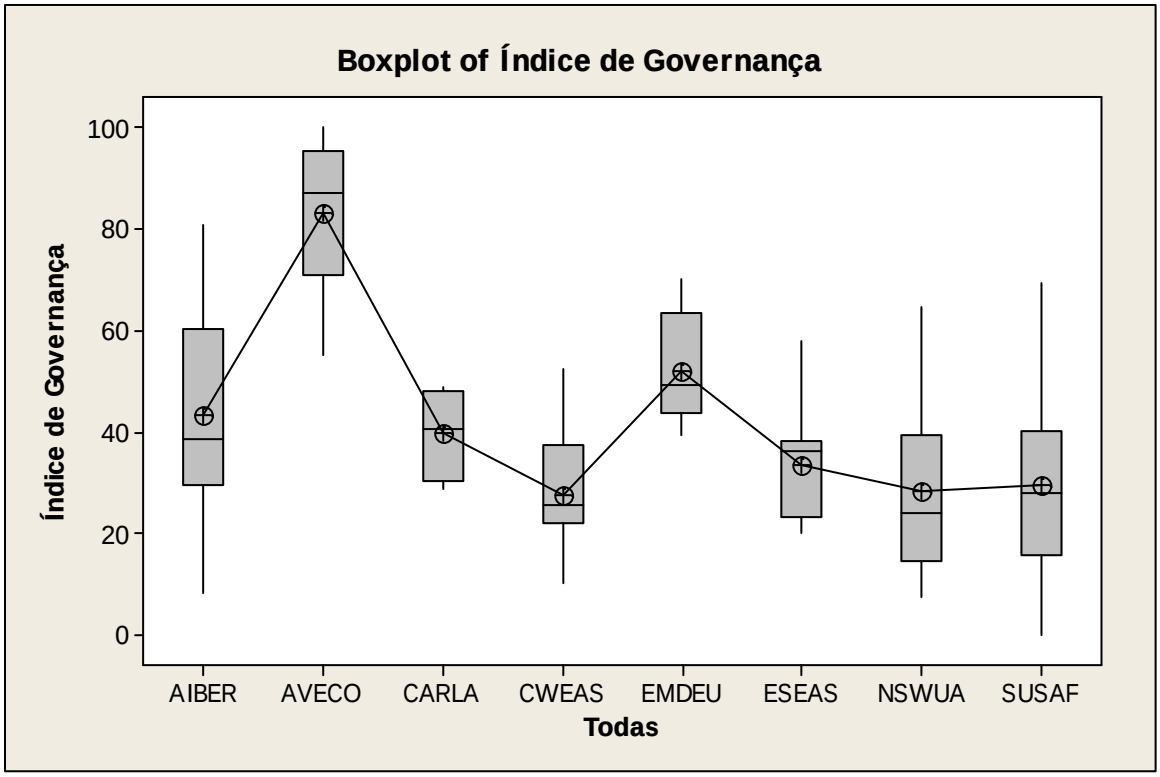


One-way ANOVA: Bem-estar versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  35150  17575  60,69  0,000
Error   129  37360    290
Total   131  72510

S = 17,02   R-Sq = 48,48%   R-Sq(adj) = 47,68%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  68,28  15,01                (----*---)
AVECO   27  78,79  16,71                       (----*---)
OTHERS  84  40,90  17,56  (-*--)
                          -----+---------+---------+---------+----
                              45        60        75        90

Pooled StDev = 17,02
```

Boxplot of Bem-estar

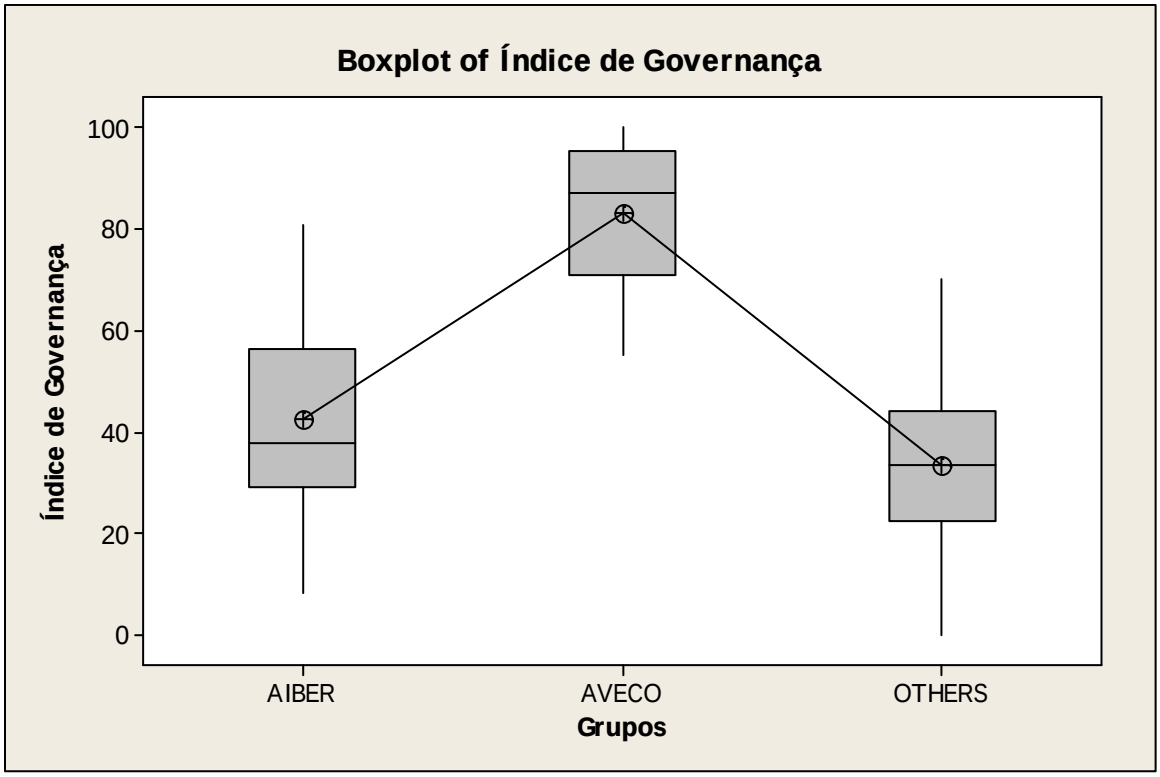
1.1.29. Índice de Governança



One-way ANOVA: Índice de Governança versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  56677  8097  35,53  0,000
Error   124  28261   228
Total   131  84938

S = 15,10   R-Sq = 66,73%   R-Sq(adj) = 64,85%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  20  43,17  19,04           (---*--)
AVECO  27  83,25  13,38                                (--*--)
CARLA   4  39,62   9,54     (-------*------)
CWEAS  11  27,38  12,38  (----*---)
EMDEU  12  51,93  11,21               (---*---)
ESEAS  12  33,61  10,47     (----*---)
NSWUA  14  28,46  16,21   (---*---)
SUSAF  32  29,58  16,95     (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,10
```

Boxplot of Índice de Governança

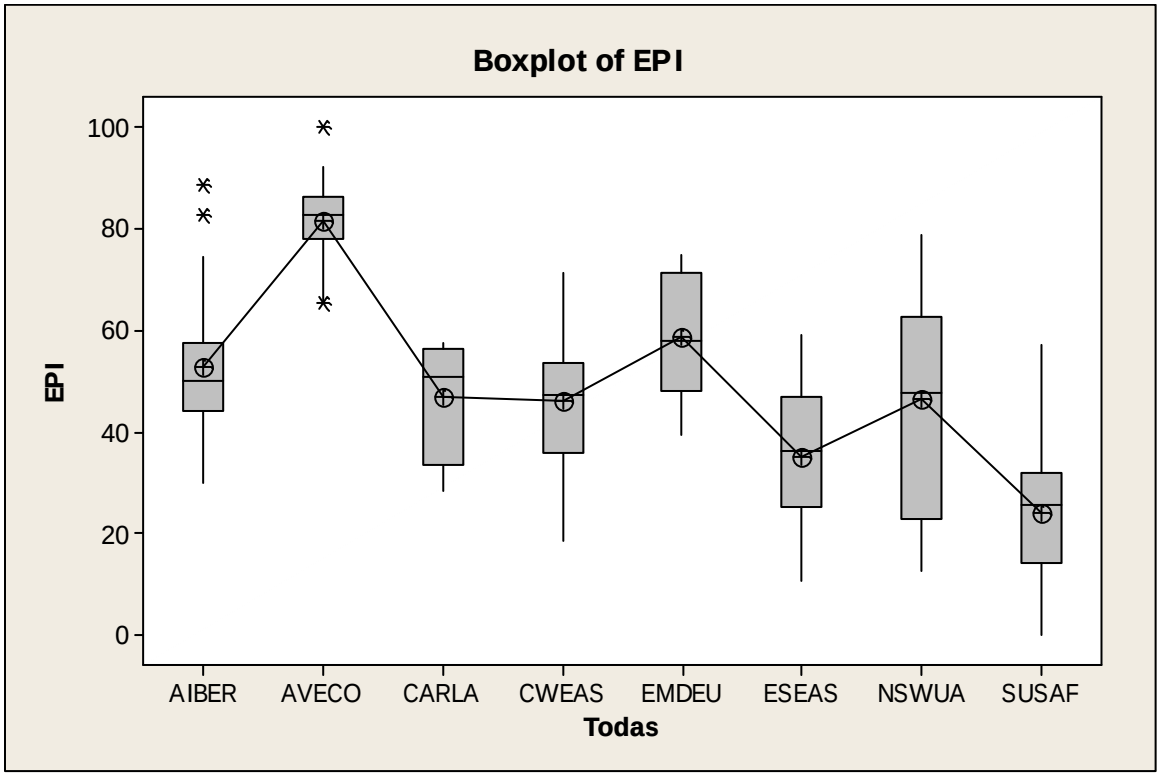


One-way ANOVA: Índice de Governança versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  51019  25510  97,02  0,000
Error   129  33919    263
Total   131  84938

S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42,48  18,82       (---*----)
AVECO   27  83,25  13,38                                 (----*---)
OTHERS  84  33,36  16,35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75

Pooled StDev = 16,22
```

Boxplot of Índice de Governança

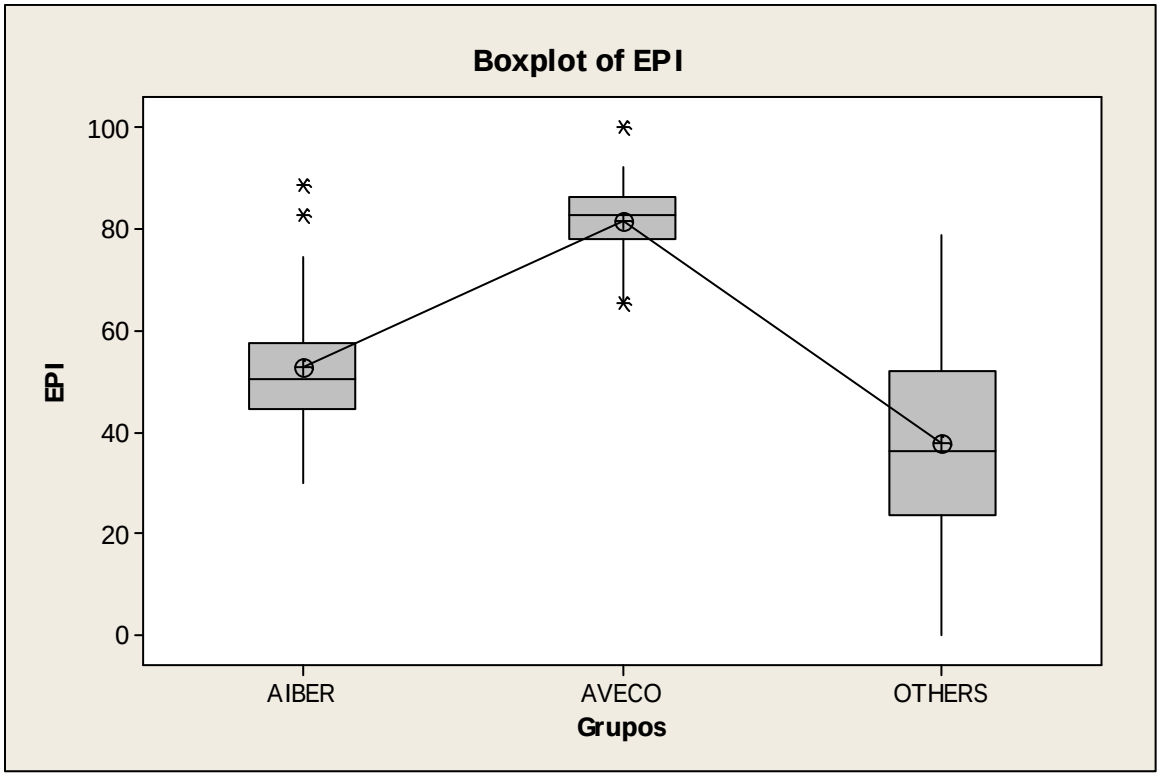
1.1.30. EPI



One-way ANOVA: EPI versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  52052  7436  40,44  0,000
Error   124  22803   184
Total   131  74855

S = 13,56   R-Sq = 69,54%   R-Sq(adj) = 67,82%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  20  52,69  14,68               (--*--)
AVECO  27  81,39   8,14                                 (--*-)
CARLA   4  46,92  12,87         (-----*------)
CWEAS  11  46,14  14,39           (---*---)
EMDEU  12  58,63  11,91                 (---*---)
ESEAS  12  34,97  13,95      (--*---)
NSWUA  14  46,41  20,67            (--*---)
SUSAF  32  24,14  12,90   (-*-)
                          +---------+---------+---------+---------
                          20        40        60        80

Pooled StDev = 13,56
```

Boxplot of EPI

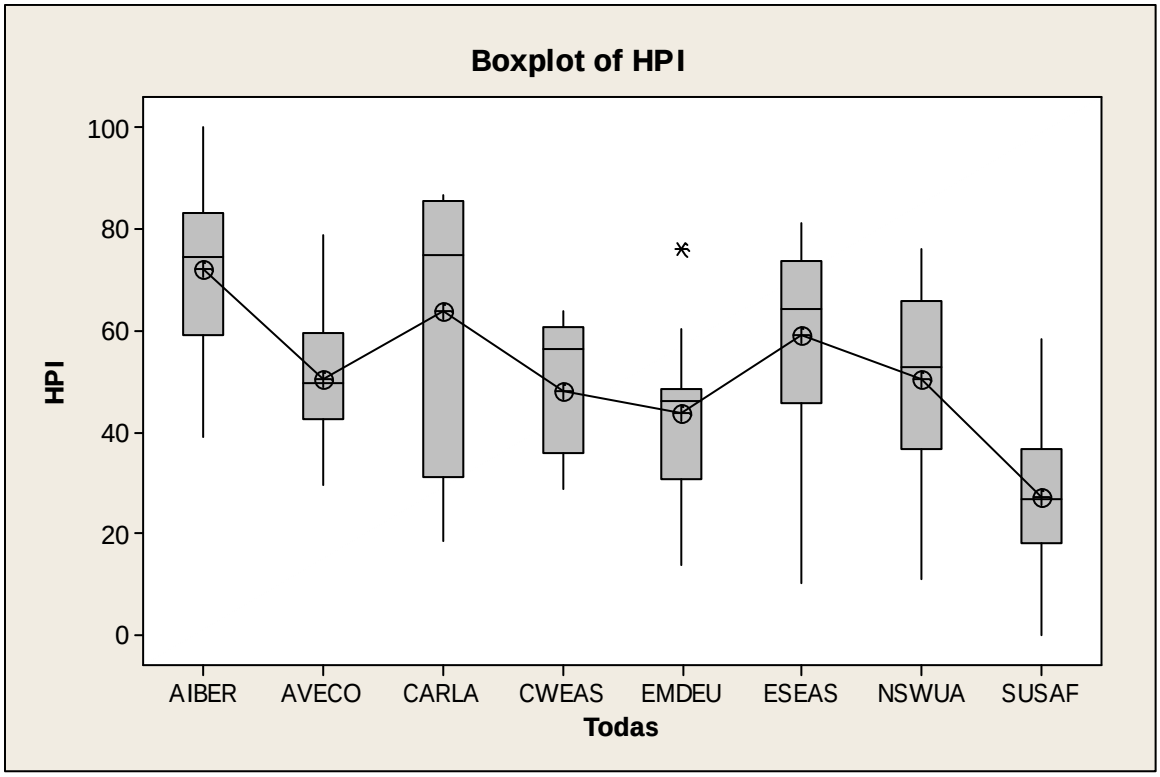


One-way ANOVA: EPI versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  38872  19436  69,68  0,000
Error   129  35982    279
Total   131  74855

S = 16,70   R-Sq = 51,93%   R-Sq(adj) = 51,18%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  52,70  14,31         (----*----)
AVECO   27  81,39   8,14                             (---*----)
OTHERS  84  37,95  19,06  (-*--)
                          -------+---------+---------+---------+--
                                45        60        75        90

Pooled StDev = 16,70
```

Boxplot of EPI

1.1.31. HPI



One-way ANOVA: HPI versus Todas

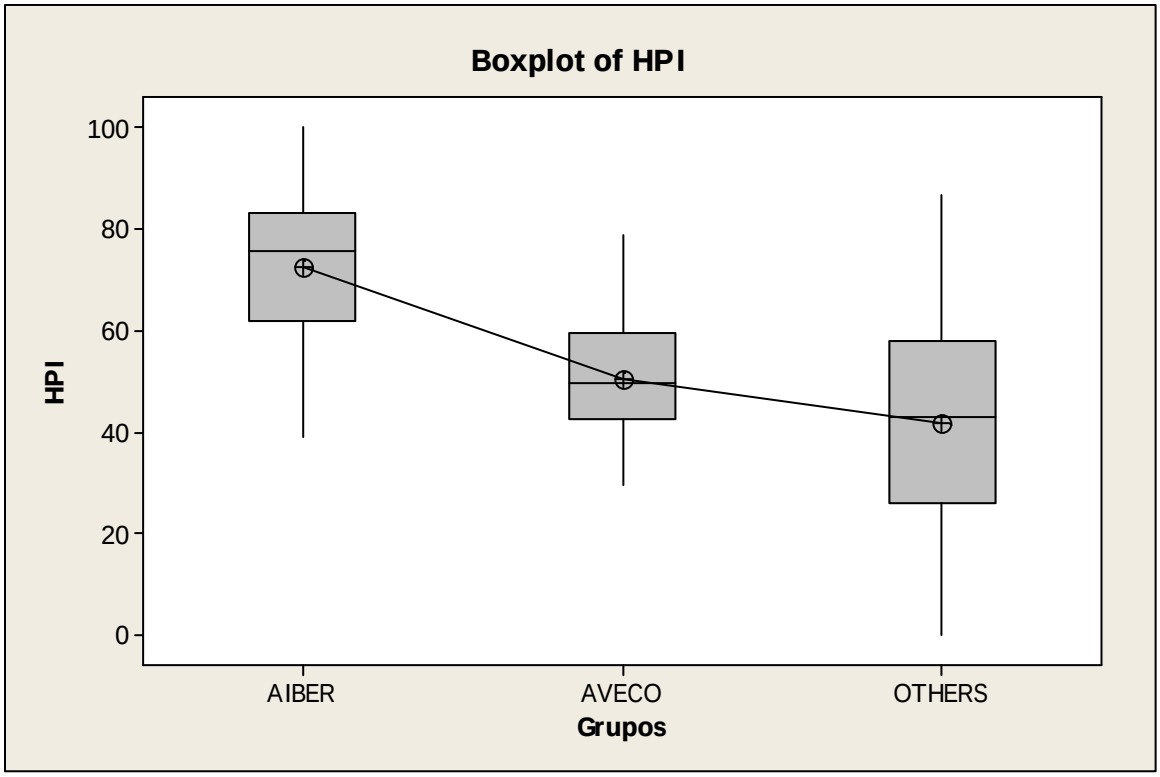
Source	DF	SS	MS	F	P
Todas	7	28129	4018	15,21	0,000
Error	124	32765	264		
Total	131	60894			

S = 16,26 R-Sq = 46,19% R-Sq(adj) = 43,16%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	20	71,92	16,55	(----*----)
AVECO	27	50,58	12,84	(---*---)
CARLA	4	63,81	31,09	(----------*---------)
CWEAS	11	48,13	13,54	(-----*------)
EMDEU	12	43,56	16,02	(-----*-----)
ESEAS	12	59,06	20,02	(-----*------)
NSWUA	14	50,22	19,91	(----*-----)
SUSAF	32	27,27	14,09	(---*---)
				30 45 60 75

Pooled StDev = 16,26

Boxplot of HPI

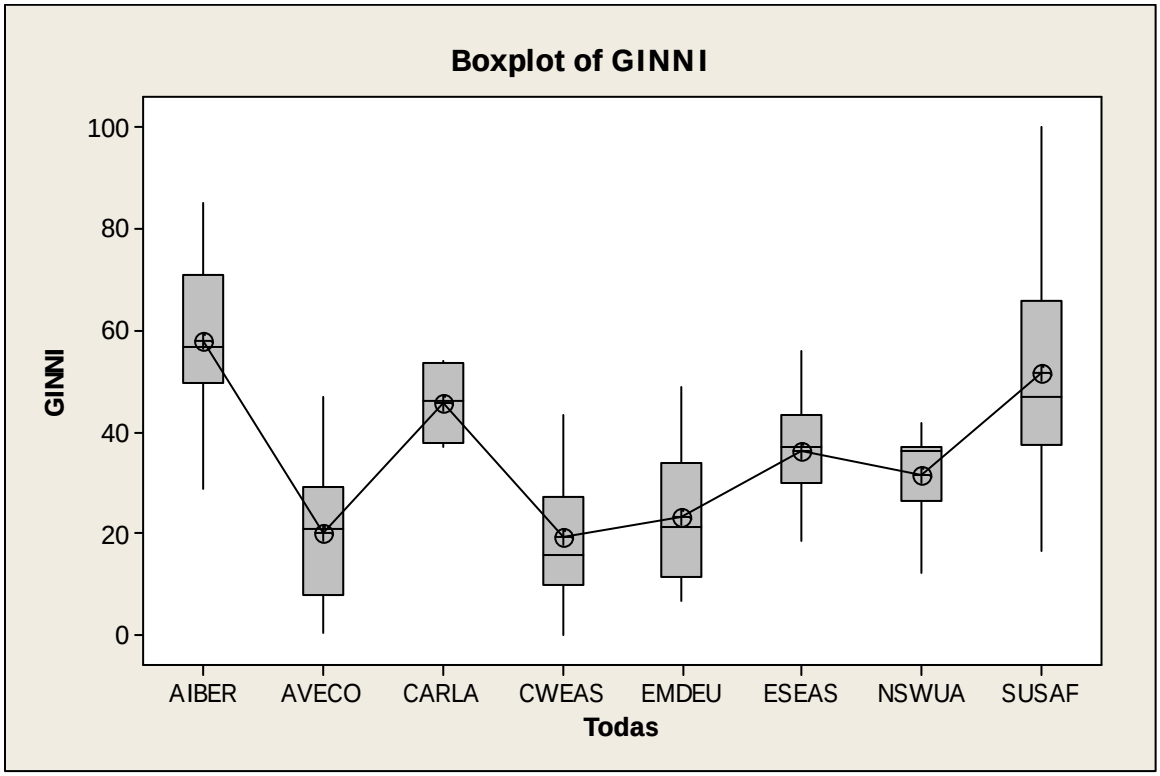


One-way ANOVA: HPI versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  15846  7923  22,69  0,000
Error   129  45048   349
Total   131  60894

S = 18,69   R-Sq = 26,02%   R-Sq(adj) = 24,87%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72,36  16,25                       (-----*------)
AVECO   27  50,58  12,84      (-----*-----)
OTHERS  84  41,79  20,68  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18,69
```

Boxplot of HPI

1.1.32. Índice de Ginni



One-way ANOVA: GINNI versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  29731  4247  18,31  0,000
Error   124  28761   232
Total   131  58492

S = 15,23   R-Sq = 50,83%   R-Sq(adj) = 48,05%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  20  57,91  14,02                            (----*---)
AVECO  27  20,09  12,42     (--*---)
CARLA   4  45,71   8,28               (---------*----------)
CWEAS  11  19,26  13,30  (-----*-----)
EMDEU  12  23,05  12,94     (----*-----)
ESEAS  12  36,21  10,22              (-----*-----)
NSWUA  14  31,36   9,15           (----*----)
SUSAF  32  51,48  21,97                         (--*---)
                          ---+---------+---------+---------+------
                            15        30        45        60

Pooled StDev = 15,23
```

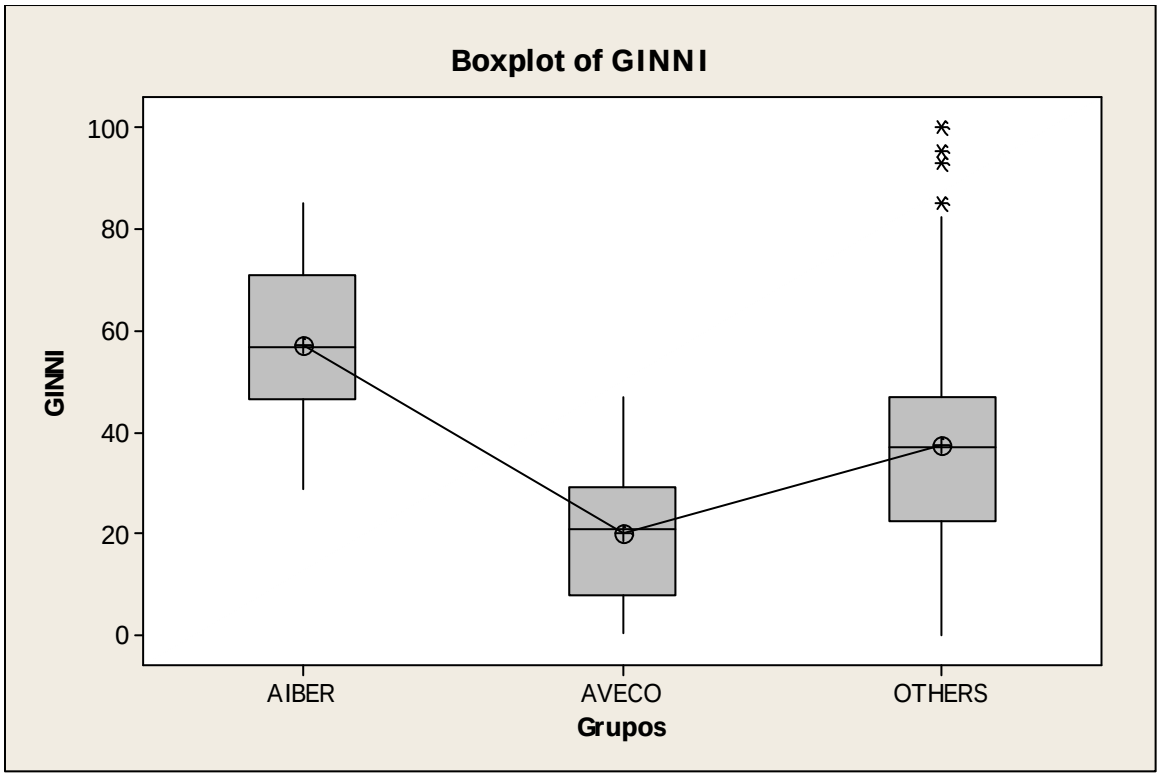
Boxplot of GINNI

One-way ANOVA: GINNI versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  16078  8039  24,45  0,000
Error   129  42414   329
Total   131  58492

S = 18,13   R-Sq = 27,49%   R-Sq(adj) = 26,36%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  56,92  14,40                        (----*----)
AVECO   27  20,09  12,42  (---*----)
OTHERS  84  37,56  20,32              (--*--)
                          -+---------+---------+---------+--------
                          15        30        45        60

Pooled StDev = 18,13
```

Boxplot of GINNI



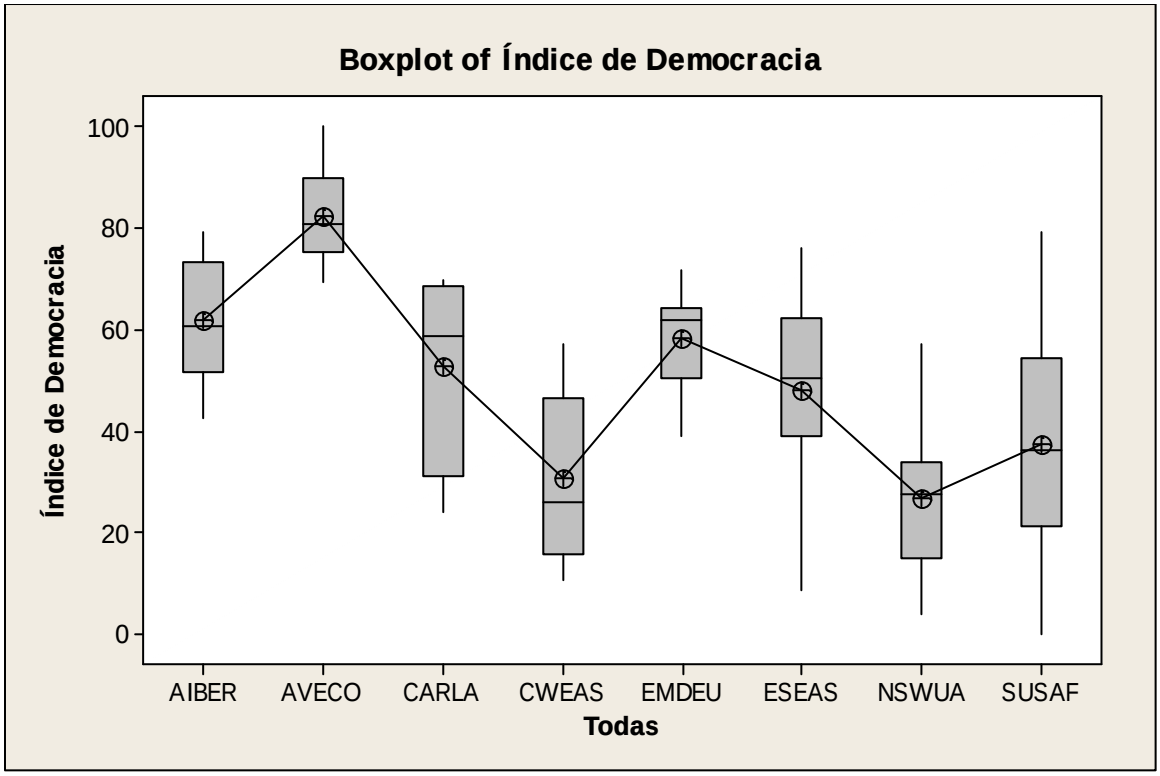
One-way ANOVA: GINNI versus Grupos

```
Source   DF     SS    MS      F      P
Grupos    2  16078  8039  24,45  0,000
Error   129  42414   329
Total   131  58492

S = 18,13   R-Sq = 27,49%   R-Sq(adj) = 26,36%
```


```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  56,92  14,40                        (----*----)
AVECO   27  20,09  12,42  (---*----)
OTHERS  84  37,56  20,32              (--*--)
                          -+---------+---------+---------+--------
                          15        30        45        60

Pooled StDev = 18,13
```


Boxplot of GINNI

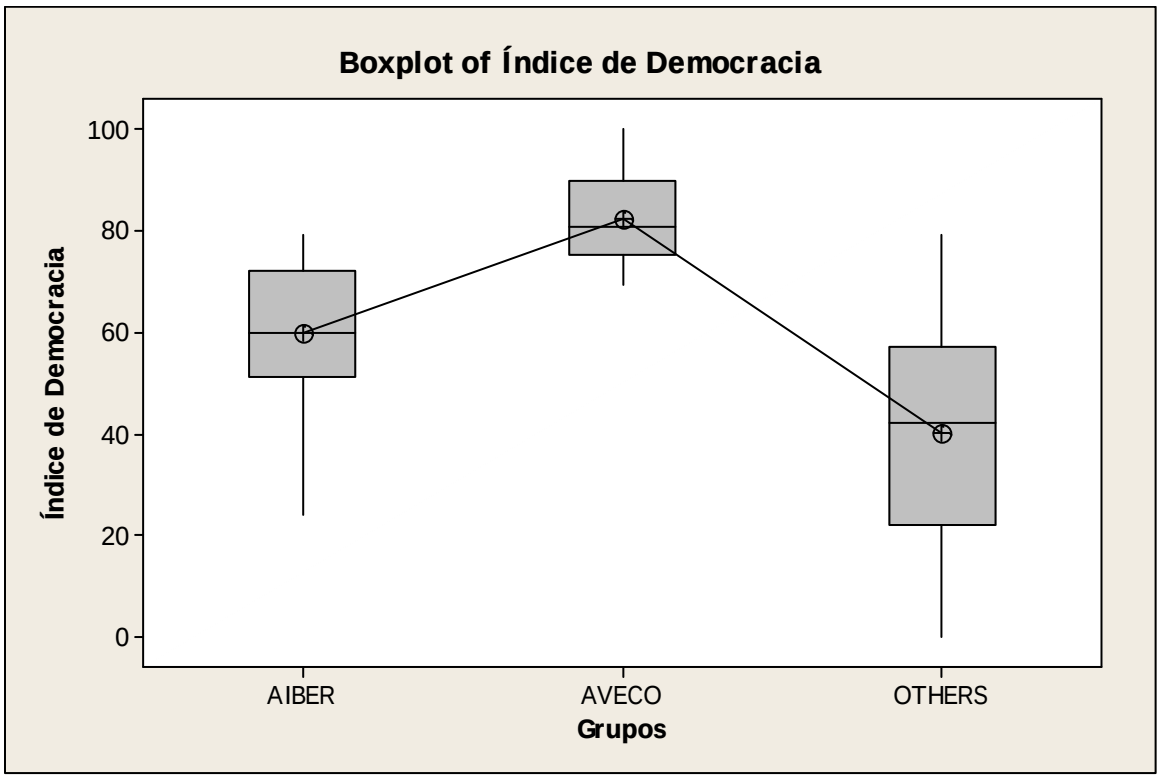
1.1.33. Índice de Democracia



One-way ANOVA: Índice de Democracia versus Todas

```
Source   DF     SS    MS      F      P
Todas     7  48430  6919  28,37  0,000
Error   124  30243   244
Total   131  78672

S = 15,62   R-Sq = 61,56%   R-Sq(adj) = 59,39%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  20  61,77  11,09                   (---*--)
AVECO  27  82,48   8,87                             (--*--)
CARLA   4  52,87  20,59            (------*-------)
CWEAS  11  30,63  16,67    (---*----)
EMDEU  12  58,10  10,28                  (---*----)
ESEAS  12  47,88  19,58            (----*---)
NSWUA  14  26,55  14,34  (---*---)
SUSAF  32  37,53  21,06         (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,62
```

Boxplot of Índice de Democracia



One-way ANOVA: Índice de Democracia versus Grupos

```
Source   DF     SS     MS      F      P
Grupos    2  38300  19150  61,19  0,000
Error   129  40372    313
Total   131  78672

S = 17,69   R-Sq = 48,68%   R-Sq(adj) = 47,89%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  59,98  13,59            (----*----)
AVECO   27  82,48   8,87                       (----*---)
OTHERS  84  40,11  20,43  (--*-)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 17,69
```

Boxplot of Índice de Democracia

CONSIDERAÇÕES FINAIS

O presente trabalho efetuou uma análise das médias amostrais, intervalos de confiança para a média populacional e análise de variância referente a 33 variáveis de diversos países do mundo separador por região. O objetivo foi comparar a América Latina com as demais regiões, na comparação com a AVECO agrupamento formado por países com alto grau de

desenvolvimento e o agrupamento Others, composto pelos demais países, os países da América Latina ficaram no meio termo.

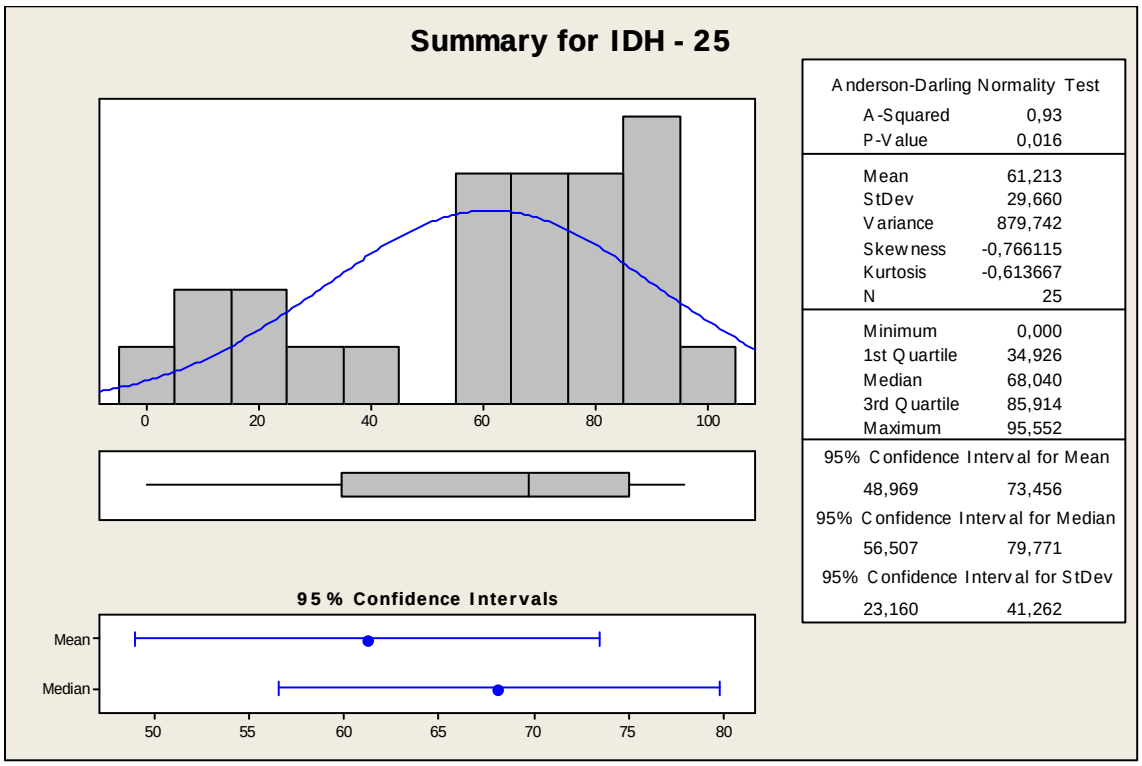
Com relação a todos os agrupamentos os países da América Latina aparecem na média, a AVECO lidera em todos os quesitos e a SUSAF fica sempre abaixo, liderando em amplitude,e invariavelmente com o menor intervalo de confiança e quase sempre apresentando outliers, Angola aparece como outlier próximo de 0 na maior parte dos itens analisados.

Ao analisar os Box-plots pode-se dizer que grande parte dos gráficos apresentam assimetria.

Entrega 6 – Amostragem

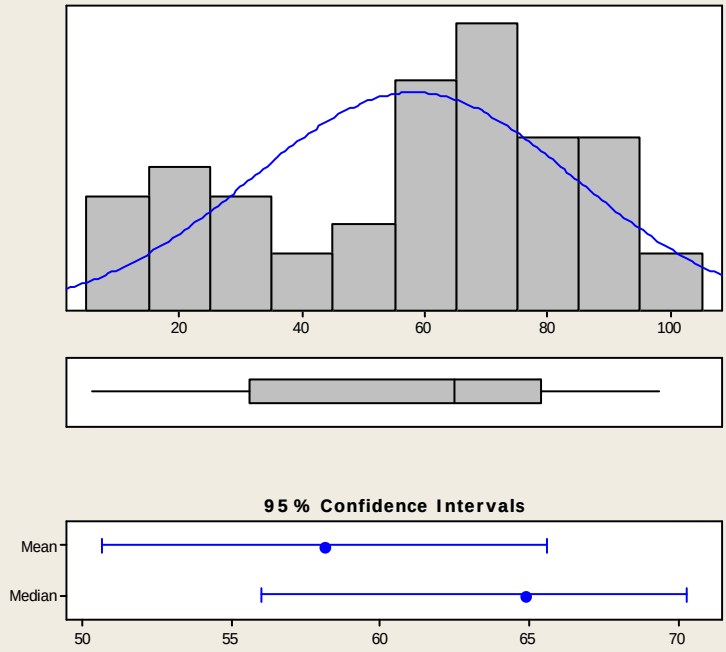
Para este trabalho cujo tema é “Bem-estar e Qualidade de Vida” escolheu-se para a análise por amostragem três variáveis analisadas a seguir: IDH, Bem-estar e Índice de Progresso Social.
Trabalhou-se com três amostras aleatórias sendo, uma com 25 países, uma com 50 países e outra com 100 países.

1.IDH



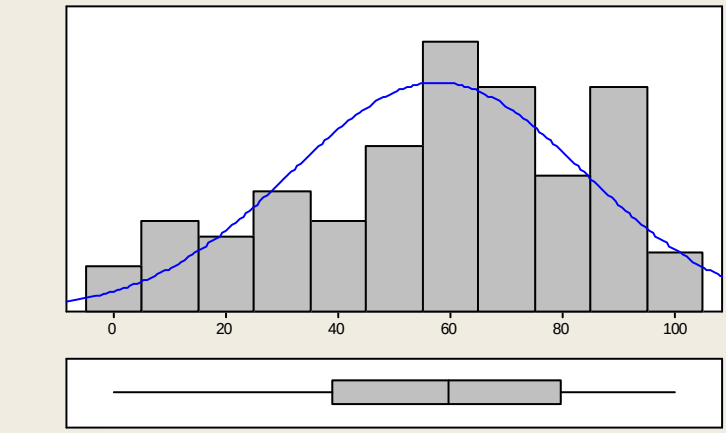
Anderson-Darling Normality Test	
A-Squared	0,93
P-Value	0,016
Mean	61,213
StDev	29,660
Variance	879,742
Skewness	-0,766115
Kurtosis	-0,613667
N	25
Minimum	0,000
1st Quartile	34,926
Median	68,040
3rd Quartile	85,914
Maximum	95,552
95% Confidence Interval for Mean	
48,969	73,456
95% Confidence Interval for Median	
56,507	79,771
95% Confidence Interval for StDev	
23,160	41,262

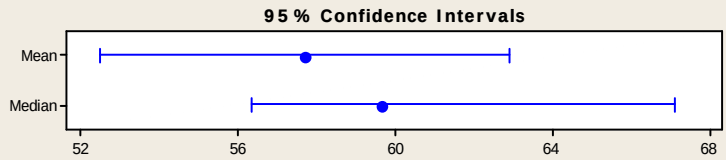
Summary for IDH - 50



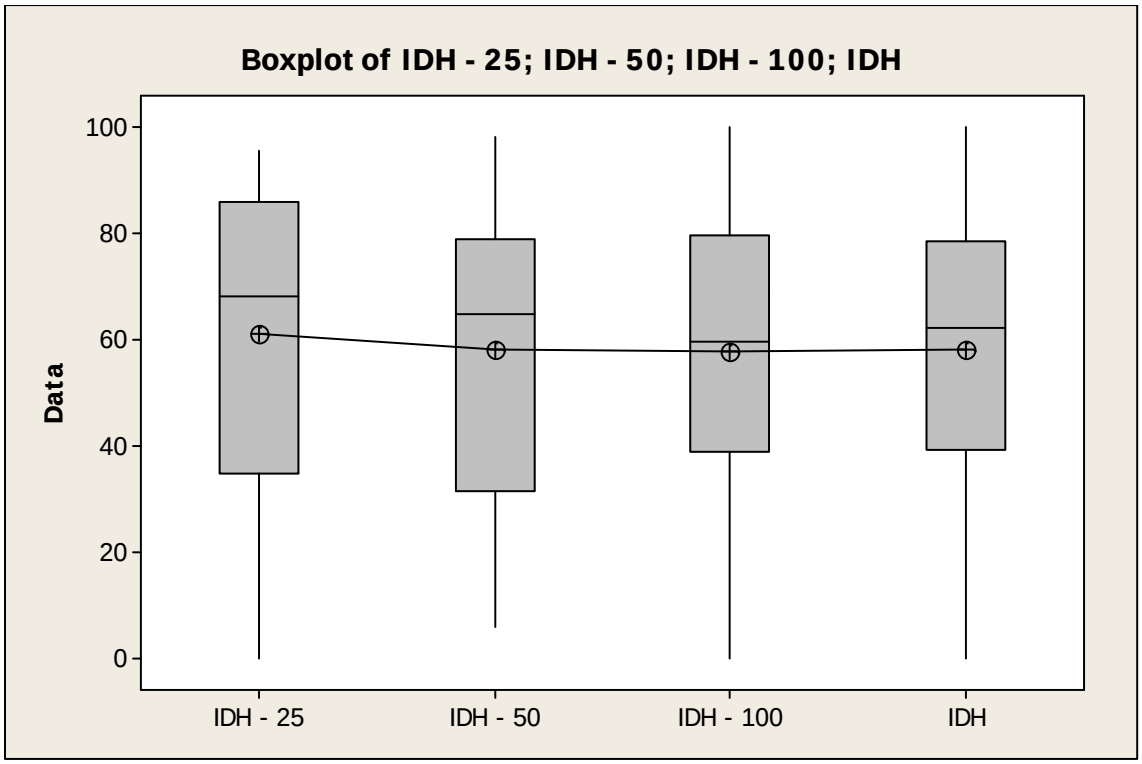
Anderson-Darling Normality Test	
A-Squared	1,12
P-Value	0,006
Mean	58,102
StDev	26,297
Variance	691,538
Skewness	-0,448637
Kurtosis	-0,902054
N	50
Minimum	5,766
1st Quartile	31,425
Median	64,827
3rd Quartile	78,871
Maximum	98,188
95% Confidence Interval for Mean	
50,629	65,576
95% Confidence Interval for Median	
56,026	70,289
95% Confidence Interval for StDev	
21,967	32,770

Summary for IDH - 100





Anderson-Darling Normality Test	
A-Squared	1,03
P-Value	0,010
Mean	57,695
StDev	26,107
Variance	681,575
Skewness	-0,441723
Kurtosis	-0,638617
N	100
Minimum	0,000
1st Quartile	38,839
Median	59,638
3rd Quartile	79,613
Maximum	100,000
95% Confidence Interval for Mean	
52,515	62,875
95% Confidence Interval for Median	
56,338	67,094
95% Confidence Interval for StDev	
22,922	30,328



One-way ANOVA: IDH - 25; IDH - 50; IDH - 100; IDH

Source	DF	SS	MS	F	P
Factor	3	254	85	0,12	0,946
Error	303	208706	689		
Total	306	208960			

S = 26,24 R-Sq = 0,12% R-Sq(adj) = 0,00%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
IDH - 25	25	61,21	29,66	(----------*----------)
IDH - 50	50	58,10	26,30	(-------*-------)
IDH - 100	100	57,70	26,11	(----*----)
IDH	132	58,09	25,66	(----*---)

54,0 60,0 66,0 72,0

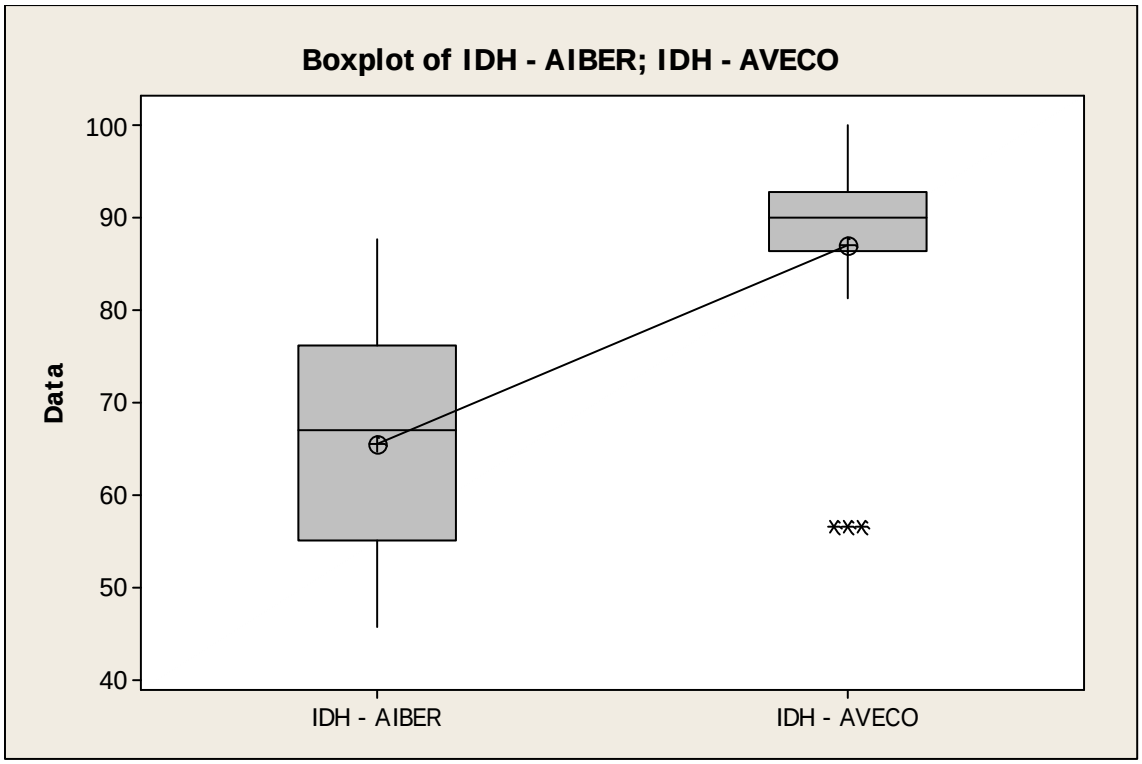
Pooled StDev = 26,24

N é o numero da amostra

Boxplot of IDH - 25; IDH - 50; IDH - 100; IDH

Análises

Nota-se que com a amostra de 100 a curva normal ficou mais bem definida, com a média em 57,70 além disso, observa-se que o intervalo ficou menos, a amostra de 100 está entre 52,5 e 62,8, um pouco menos de dez pontos, na amostra de 50 esse intervalo está entre 50,6 e 65,5, praticamente 15 pontos e na amostra de 25 está entre 48,9 e 73,4, aproximadamente 25 pontos.



Two-Sample T-Test and CI: IDH - AIBER; IDH - AVECO

```
Two-sample T for IDH - AIBER vs IDH - AVECO

              N  Mean  StDev  SE Mean
IDH - AIBER  21  65,6   12,0      2,6
IDH - AVECO  27  87,0   11,7      2,3


Difference = mu (IDH - AIBER) - mu (IDH - AVECO)
Estimate for difference:  -21,35
95% upper bound for difference:  -15,53
T-Test of difference = 0 (vs <): T-Value = -6,17  P-Value = 0,000  DF = 42
```

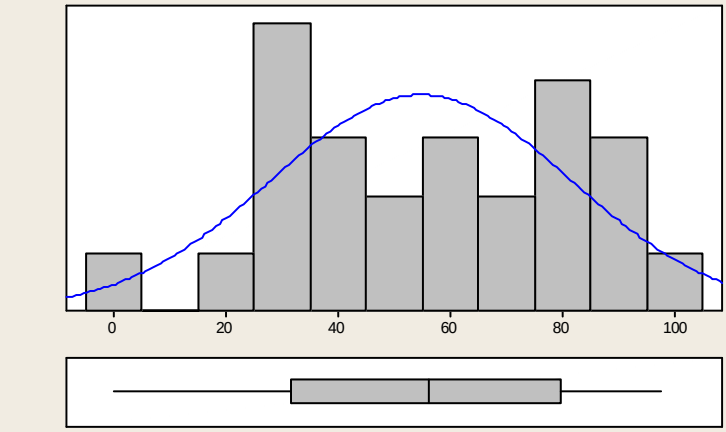
Boxplot of IDH - AIBER; IDH - AVECO

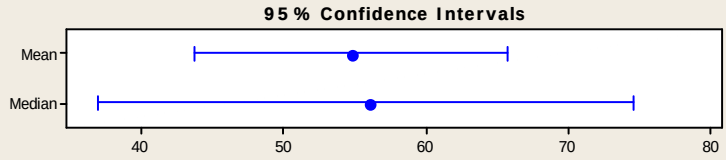
Análises

Observa-se neste gráfico um P-Value de 0, enquanto no gráfico acima ele é de 0,946, mostrando uma maior confiança. No gráfico acima observam-se três outliers no AVECO, são eles Bulgária, Burkina Faso e Burundi 56,7, a média dos países da AVECO para o IDH é 87.

2. Bem-estar

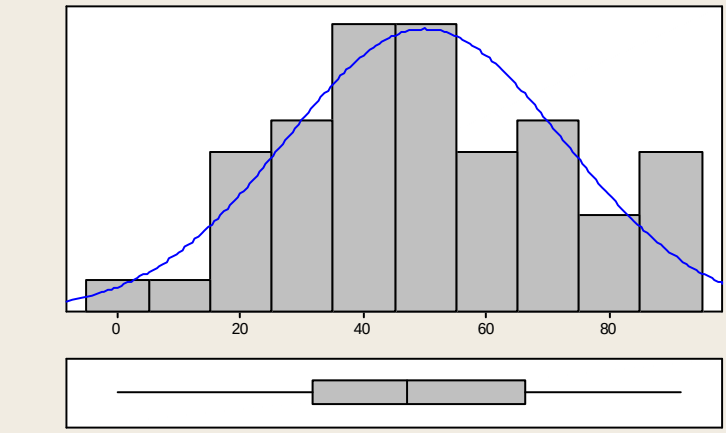
Summary for Bem-estar 25

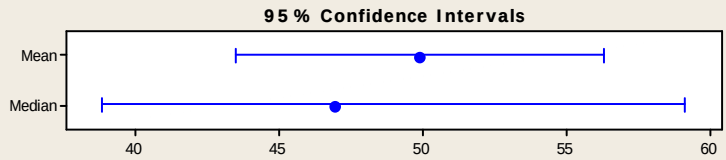




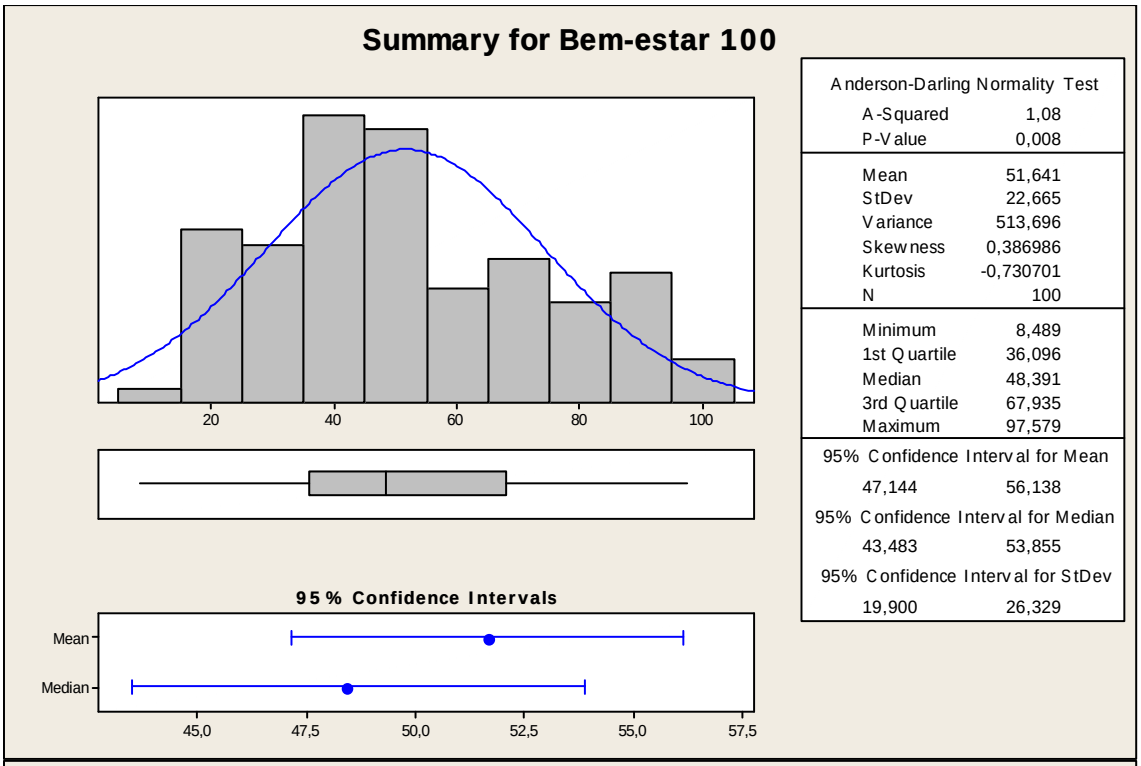
Anderson-Darling Normality Test	
A-Squared	0,30
P-Value	0,557
Mean	54,733
StDev	26,604
Variance	707,768
Skewness	-0,120805
Kurtosis	-0,881317
N	25
Minimum	0,000
1st Quartile	31,719
Median	56,056
3rd Quartile	79,707
Maximum	97,579
95% Confidence Interval for Mean	
43,751	65,714
95% Confidence Interval for Median	
37,035	74,544
95% Confidence Interval for StDev	
20,773	37,010

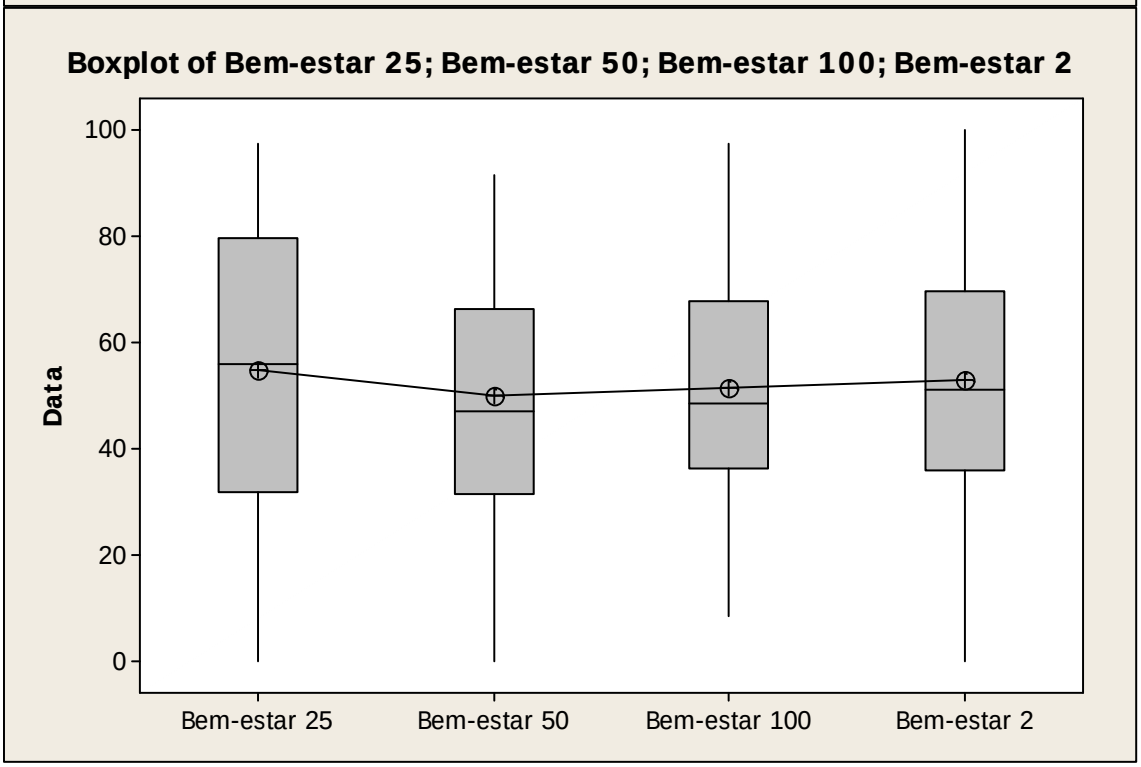
Summary for Bem-estar 50





Anderson-Darling Normality Test	
A-Squared	0,31
P-Value	0,550
Mean	49,902
StDev	22,582
Variance	509,947
Skewness	0,116143
Kurtosis	-0,621497
N	50
Minimum	0,000
1st Quartile	31,628
Median	46,964
3rd Quartile	66,338
Maximum	91,706
95% Confidence Interval for Mean	
43,485	56,320
95% Confidence Interval for Median	
38,851	59,140
95% Confidence Interval for StDev	
18,864	28,140





One-way ANOVA: Bem-estar 25; Bem-estar 50; Bem-estar 100; Bem-estar 2

Source	DF	SS	MS	F	P
Factor	3	542	181	0,33	0,803
Error	303	165340	546		
Total	306	165881			

S = 23,36 R-Sq = 0,33% R-Sq(adj) = 0,00%

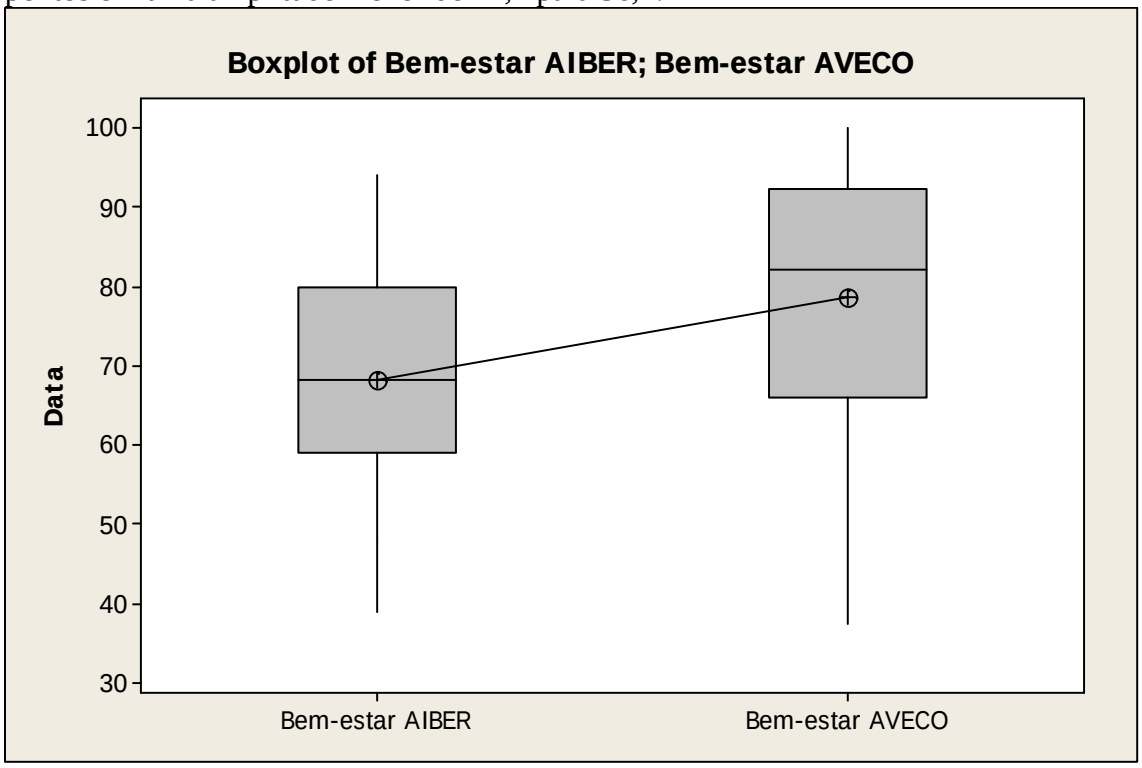
Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
Bem-estar 25	25	54,73	26,60	(--------------*--------------)

```
Bem-estar 50    50  49,90  22,58  (----------*----------)
Bem-estar 100  100  51,64  22,66        (-------*-------)
Bem-estar 2    132  53,01  23,53            (-----*------)
                                 --------+---------+---------+---------+-
                                      48,0      54,0      60,0      66,0

Pooled StDev = 23,36
```

Boxplot of Bem-estar 25; Bem-estar 50; Bem-estar 100; Bem-estar 2

Análises

Observa-se que na amostra com 25 países ao intervalo de confiança é muito maior que os demais com uma amplitude de 43,7 a 65,7, com aproximadamente 22 pontos, na amostra com 50 países o intervalo de confiança diminui para 13 pontos, com um a amplitude de 43,4 a 56,3, e na amostra com 100 países este intervalo de confian;ca diminui ainda mais para 9 pontos em uma amplitude menor de 47,1 para 56,1.



Two-Sample T-Test and CI: Bem-estar AIBER; Bem-estar AVECO

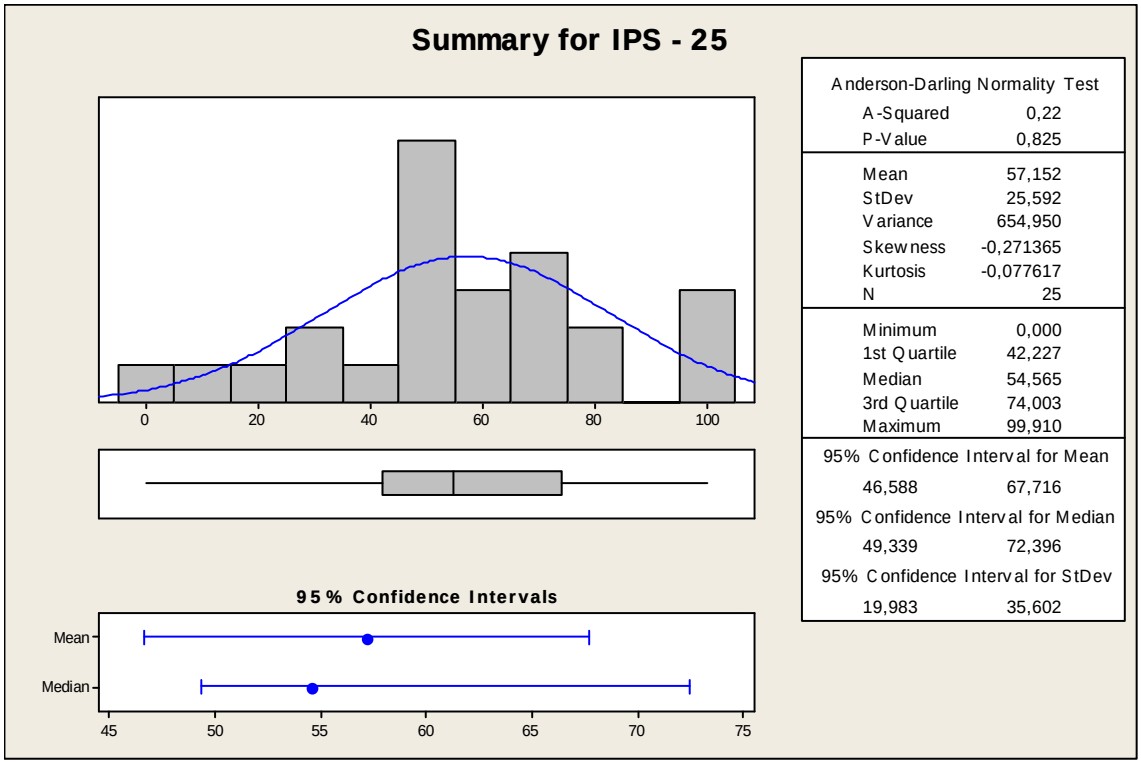
```
Two-sample T for Bem-estar AIBER vs Bem-estar AVECO

                  N  Mean  StDev  SE Mean
Bem-estar AIBER  21  68,3   15,0      3,3
Bem-estar AVECO  27  78,8   16,7      3,2


Difference = mu (Bem-estar AIBER) - mu (Bem-estar AVECO)
Estimate for difference:  -10,50
95% upper bound for difference:  -2,79
T-Test of difference = 0 (vs <): T-Value = -2,29  P-Value = 0,013  DF = 44
```

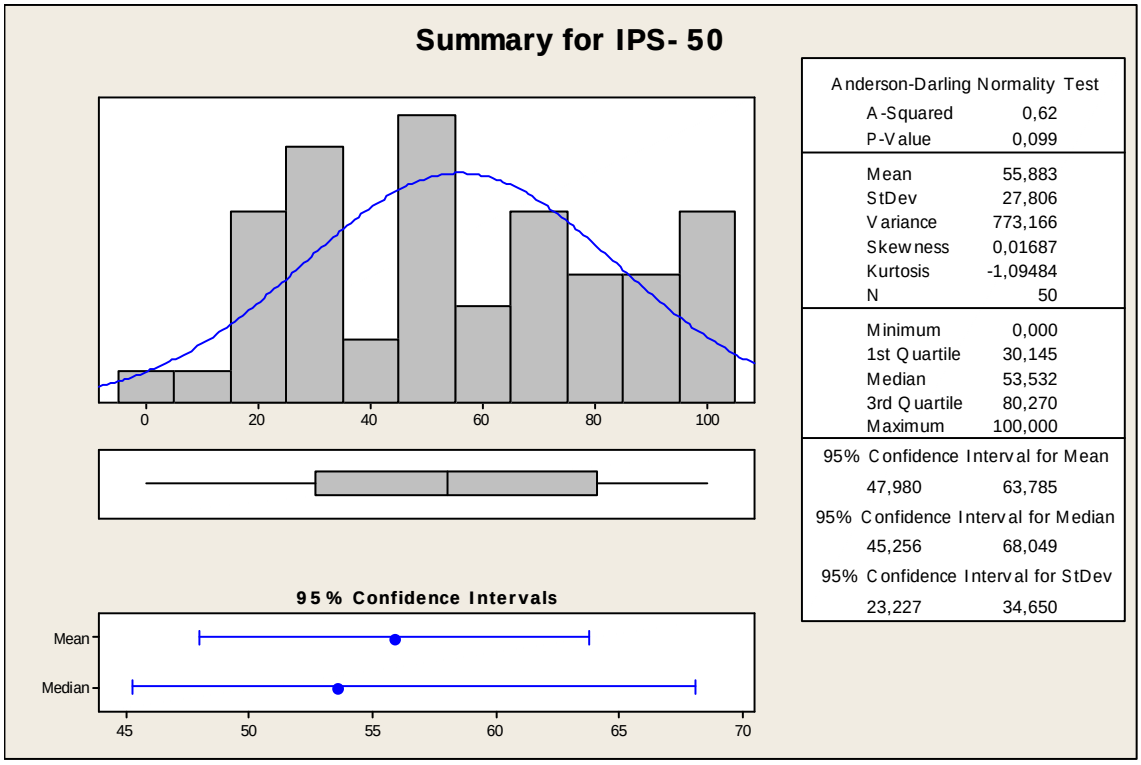
Boxplot of Bem-estar AIBER; Bem-estar AVECO

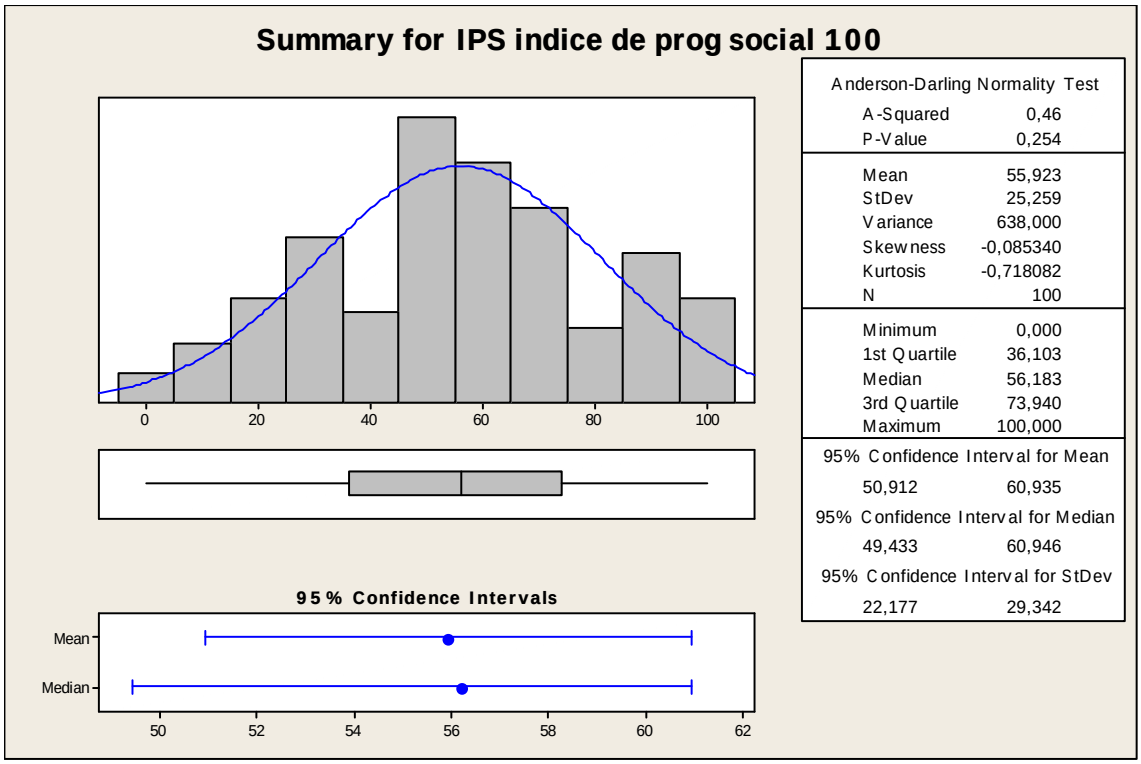
Análises

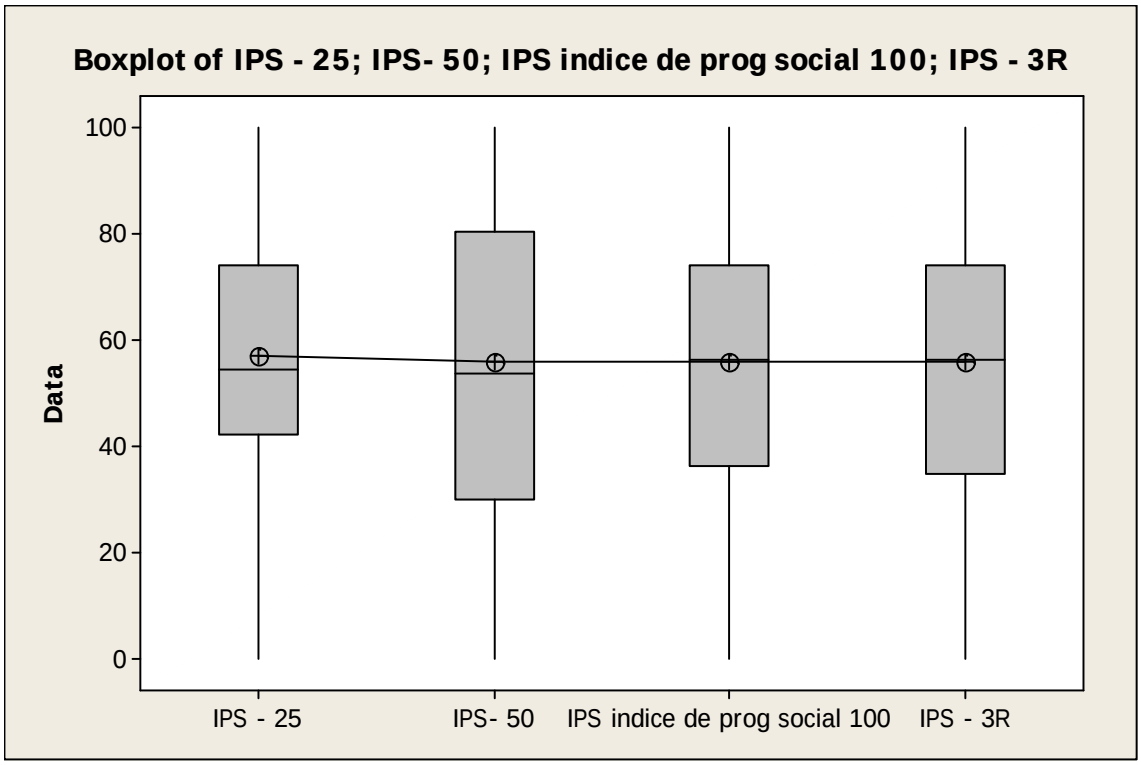
Observa-se neste gráfico um P-Value de 0,013, enquanto no gráfico acima ele é de 0,803, mostrando uma maior confiança.

3.IPS – Índice de Progresso Social









One-way ANOVA: IPS - 25; IPS- 50; IPS indice de prog social 100; IPS - 3R

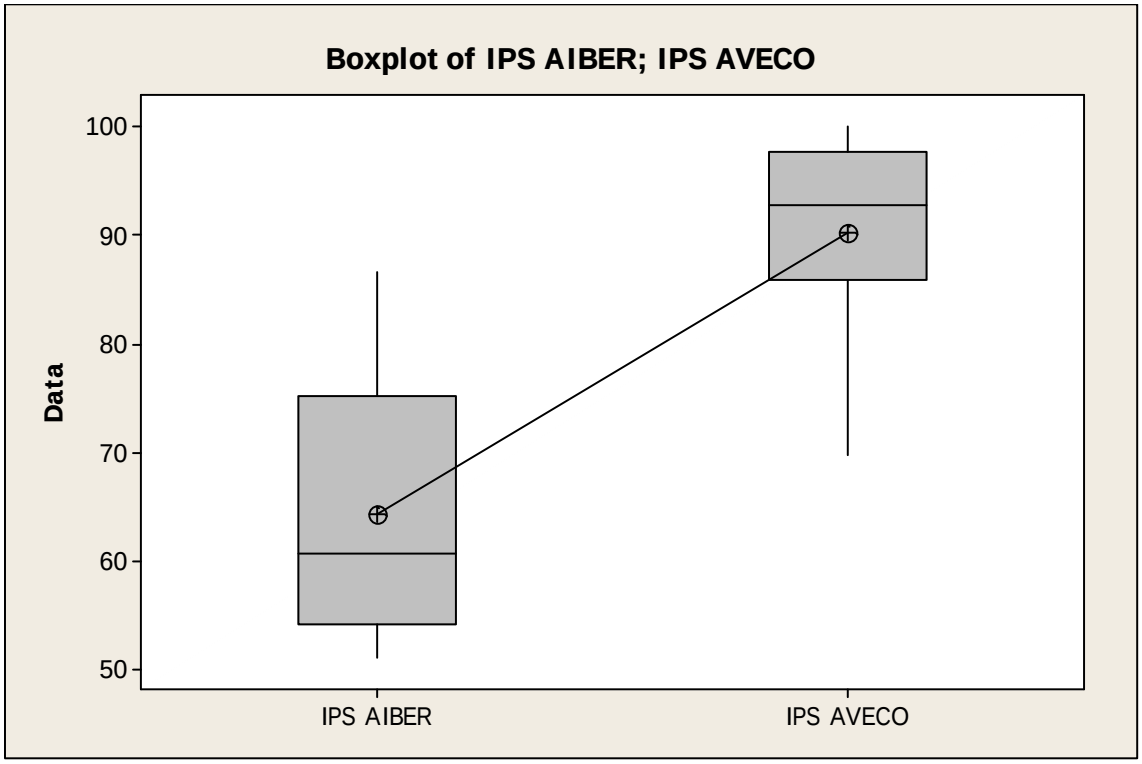
Source	DF	SS	MS	F	P
Factor	3	38	13	0,02	0,997
Error	303	202099	667		
Total	306	202136			

```
S = 25,83   R-Sq = 0,02%   R-Sq(adj) = 0,00%



Level                       N   Mean  StDev
IPS - 25                   25  57,15  25,59
IPS- 50                    50  55,88  27,81
IPS indice de prog socia  100  55,92  25,26
IPS - 3R                  132  55,84  25,52

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level                      --+---------+---------+---------+-------
IPS - 25                   (----------------*----------------)
IPS- 50                       (-----------*-----------)
IPS indice de prog socia          (-------*--------)
IPS - 3R                           (------*------)
                           --+---------+---------+---------+-------
                          48,0      54,0      60,0      66,0

Pooled StDev = 25,83
```

Boxplot of IPS - 25; IPS- 50; IPS indice de prog social 100; IPS - 3R

Análises

Observa-se que na amostra com 25 países ao intervalo de confiança é muito maior que os demais com uma amplitude de 46,5 a 67,7, com aproximadamente 21 pontos, na amostra com 50 países o intervalo de confiança diminui para 16 pontos, com um a amplitude de 47,9 a 63,7, e na amostra com 100 países este intervalo de confianca diminui ainda mais para 10 pontos em uma amplitude menor de 46,5 para 67,7.



Two-Sample T-Test and CI: IPS AIBER; IPS AVECO

```
Two-sample T for IPS AIBER vs IPS AVECO

            N   Mean  StDev  SE Mean
IPS AIBER  21   64,4   12,0      2,6
IPS AVECO  27  90,30   8,70      1,7


Difference = mu (IPS AIBER) - mu (IPS AVECO)
Estimate for difference:  -25,89
95% upper bound for difference:  -20,66
T-Test of difference = 0 (vs <): T-Value = -8,36  P-Value = 0,000  DF = 35
```

Boxplot of IPS AIBER; IPS AVECO

Análises

Observa-se neste gráfico um P-Value de 0, enquanto no gráfico acima ele é de 0,997, mostrando uma maior confiança. A diferença entre les é 25 isso significa que o AIBER e o AVECO estão muito longe um do outro.

Considerações finais

Pode-se observar nas análises que quanto maior a amostra menor o intervalo de confiança e menor a amplitude da amostra.

8. ANÁLISE EXPLORATÓRIA DOS DADOS

Iniciamos o trabalho apresentando o objetivo do Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP – Pontifícia Universidade Católica. Depois partimos para o entendimento dos dados apresentando: média, mínimo, máximo, padronização e normalização dos dados e, por fim, a análise exploratória dos dados sobre o tema: bem-estar e qualidade de vida empregando software estatístico MINITAB, em especial seus recursos: Display Descriptive Statistics que está no link principal Basic Statistics, ele permite obter a média, o valor mínimo e o valor máximo do objeto estudado.

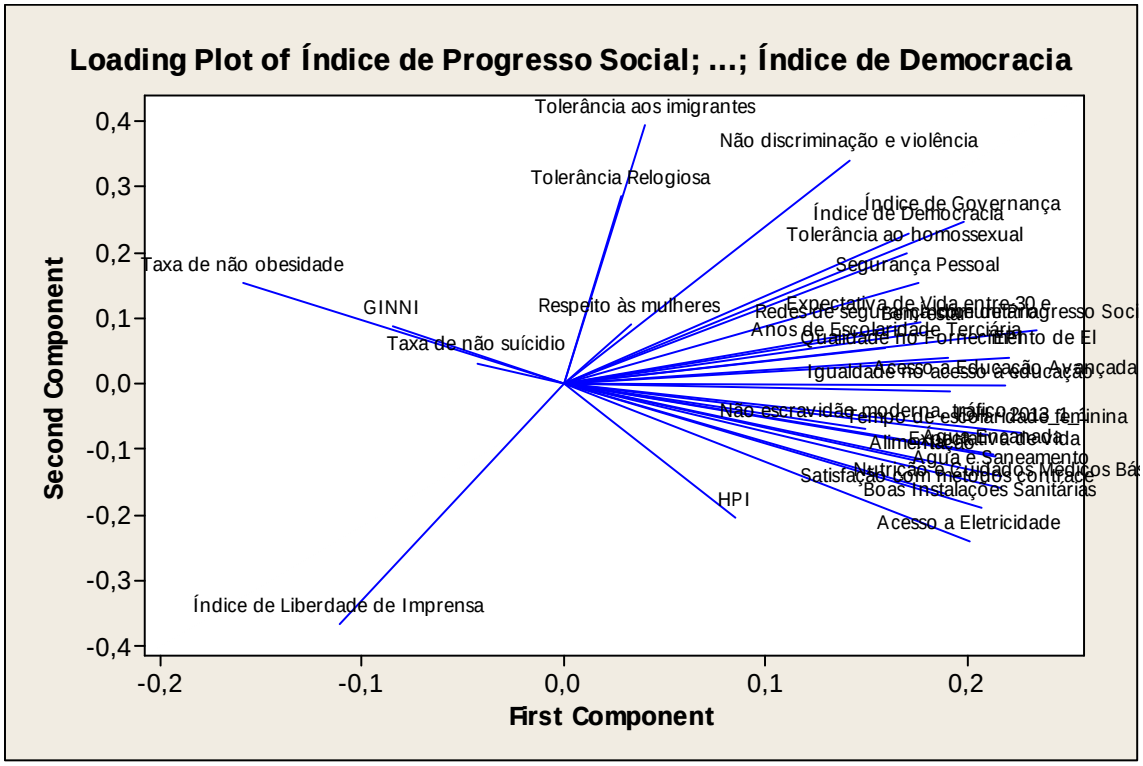
Seguimos com a aplicação do link principal Data, acessando o Code – Numeric to Numeric, sendo que este recurso permite que os dados ausentes que aparecem na tabela com o símbolo asterisco (*) seja substituído pelo valor determinado que é a média encontrada.

Na sequencia temos que fazer a normalização dos dados que é um ajuste nos dados com vista a afastar os *outliers*, ou seja, é como se dos dados analisados não possuíssem uma distribuição normal então fazemos o ajuste para que todos fiquem dentro de uma certa normalidade. No nosso caso, usamos como parâmetro médio entre o mínimo e o máximo obtido na coleta de dados. Temos que apresentar os valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas.

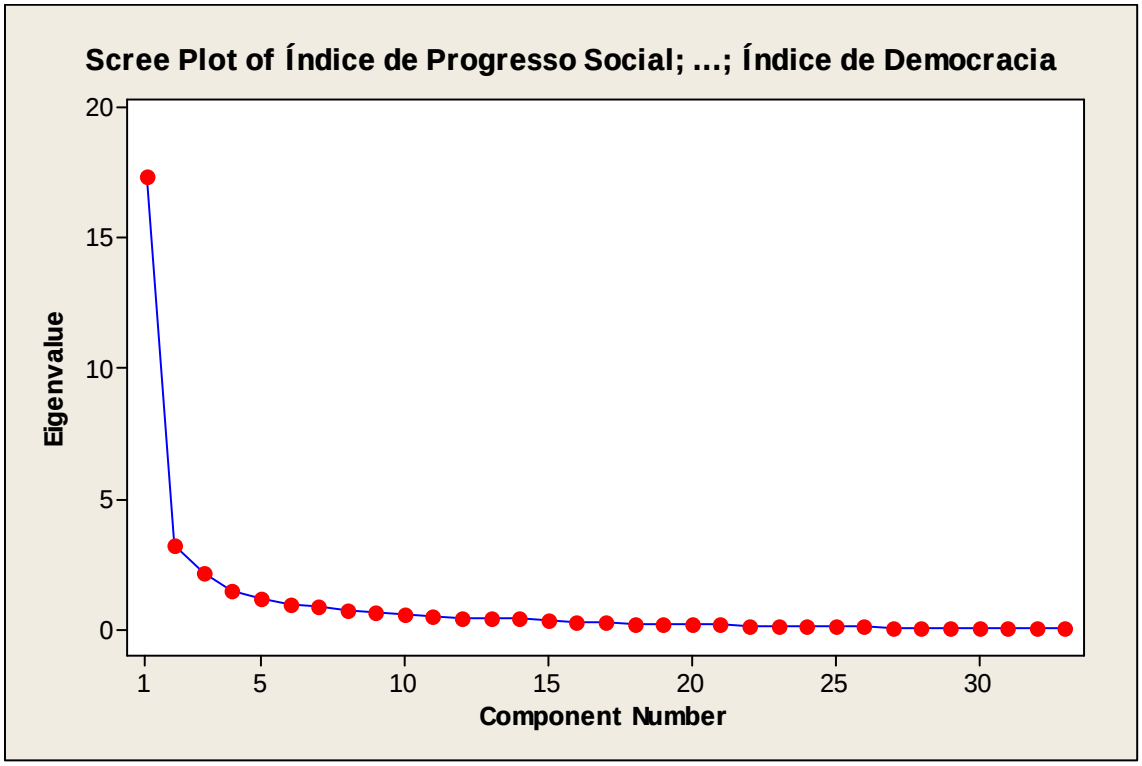
Para essa etapa usamos o link da barra de ferramentas do MINITAB denominada Calculator e aplicamos a fórmula desejada indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.

Por fim, esse trabalho apresentamos os gráficos gerados pelo recurso Graphical Summary que está no item Basic Statistic que por sua vez está no item Static na barra de ferramentas.
O Objetivo é explorar os dados sobre bem-estar e qualidade de vida de tal modo que possamos com os recursos citados anteriormente apresentar conclusões de correlações entre as variáveis que compõem o tema.

8.2. ANÁLISE DOS COMPONETES PRINCIPAIS



Análise:
A primeira componente é a que mais está explicando, a segunda componente vai para outra direção. A primeira complementa a segunda componente e vice-versa, mas não se correlacionam.



Análises:

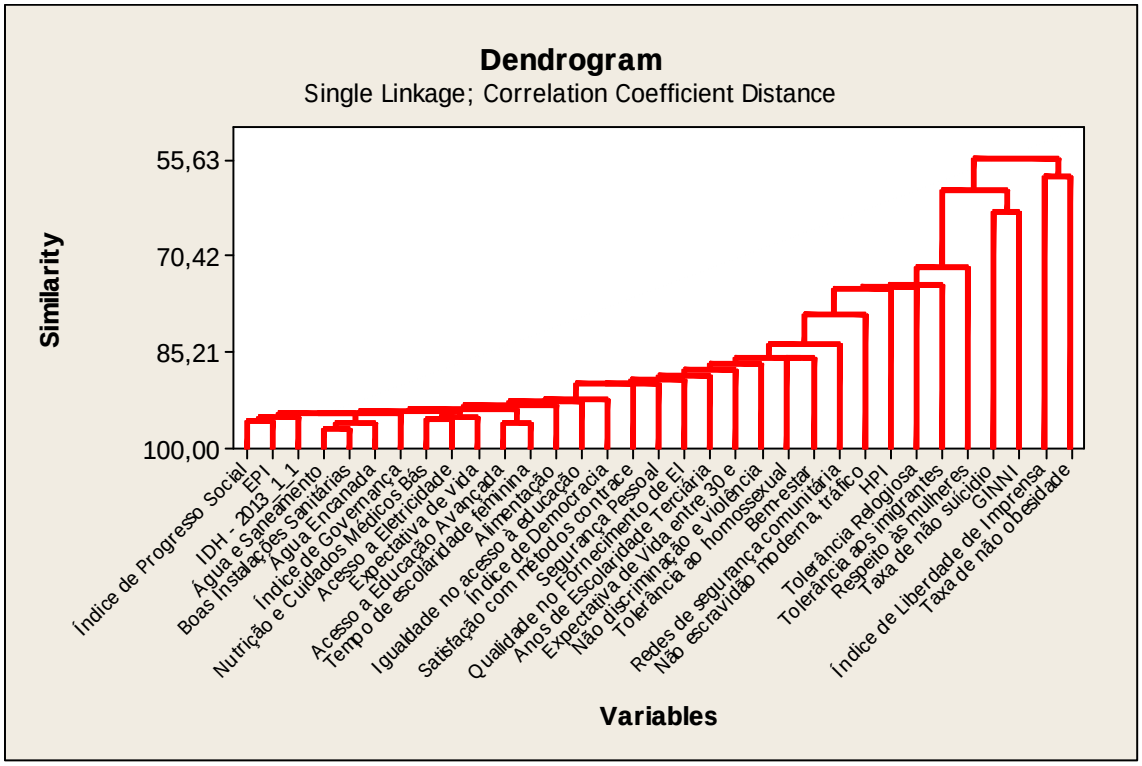
As 4 primeiras variáveis representadas pelas bolas vermelhas, da esquerda para a direita, são suficientes pois já na linha 4 tenho 73% de confiabilidade, na quinta 76%, na sexta 79%, começa a aumentar pouco.

Matriz de Correlação com 33 variáveis (PC1 a PC33)

Principal Component Analysis: Índice de Pr; Nutrição e C; Água e Sanea; Seguran

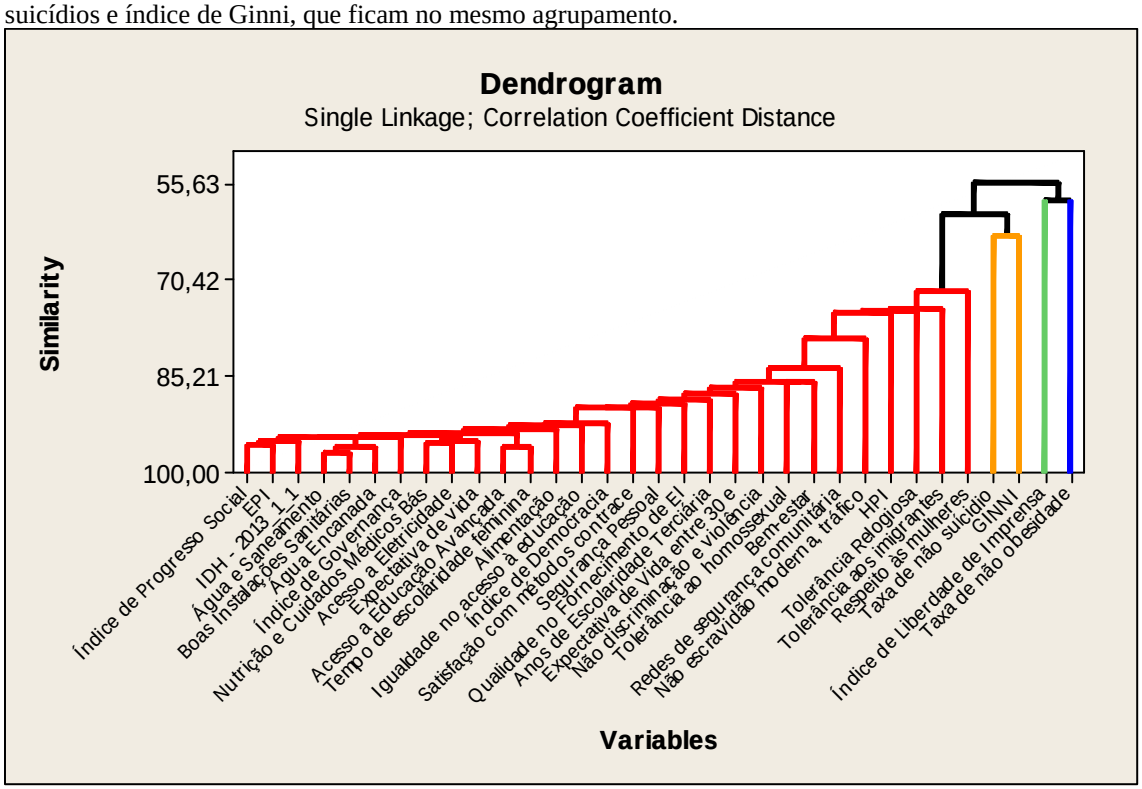
Eigenanalysis of the Correlation Matrix

Eigenvalue	17,400	3,203	2,177	1,447	1,133	0,958	0,849	0,704	0,668
Proportion	0,527	0,097	0,066	0,044	0,034	0,029	0,026	0,021	0,020
Cumulative	0,527	0,624	0,690	0,734	0,769	0,798	0,823	0,845	0,865
Eigenvalue	0,576	0,473	0,441	0,399	0,379	0,331	0,253	0,220	0,191
Proportion	0,017	0,014	0,013	0,012	0,011	0,010	0,008	0,007	0,006
Cumulative	0,882	0,897	0,910	0,922	0,934	0,944	0,951	0,958	0,964
Eigenvalue	0,169	0,151	0,142	0,117	0,112	0,104	0,097	0,089	0,061
Proportion	0,005	0,005	0,004	0,004	0,003	0,003	0,003	0,003	0,002
Cumulative	0,969	0,973	0,978	0,981	0,985	0,988	0,991	0,993	0,995
Eigenvalue	0,048	0,033	0,026	0,023	0,018	0,008			
Proportion	0,001	0,001	0,001	0,001	0,001	0,000			
Cumulative	0,997	0,998	0,999	0,999	1,000	1,000			



Análise:

O dendograma indica as variáveis mais dispersas estão no final do gráfico da esquerda para a direita são elas: Taxa de obesidade, Índice de liberdade de imprensa, Ginni e Taxa de não Suicídio. Esse aspecto fica mais claro no gráfico a baixo no qual se pede 4 agrupamentos, as variáveis não se agrupam, com exceção da indicie de suicídios e índice de Ginni, que ficam no mesmo agrupamento.



8.3. ANÁLISE DE CADA UM DOS COMPONETES PRINCIPAIS

Realizou-se a análise a partir das 32 variáveis pré-selecionadas para os 132 em questão, foram criadas as 4 componentes PC1, PC2, PC3 e PC4, com as quais pode-se verificar os agrupamentos mais e menos representativos para explicar as demais variáveis, conforme a tabela abaixo

Variable	PC1	PC2	PC3	PC4
Índice de Progresso Social	0,234	0,081	-0,005	-0,006
Nutrição e Cuidados Médicos Bás	0,216	-0,160	-0,013	0,042
Água e Saneamento	0,217	-0,143	0,005	0,039
Segurança Pessoal	0,176	0,152	-0,305	0,095
Acesso a Educação Avançada	0,219	-0,004	-0,107	-0,091
Alimentação	0,179	-0,119	-0,068	0,152
Água Encanada	0,213	-0,108	0,101	0,026
Boas Instalações Sanitárias	0,207	-0,191	-0,015	0,007
Acesso a Eletricidade	0,201	-0,242	0,041	-0,000
Qualidade no Fornecimento de El	0,190	0,040	-0,113	0,109
Índice de Liberdade de Imprensa	-0,111	-0,367	-0,015	0,129
Expectativa de vida	0,214	-0,112	0,020	0,108
Expectativa de Vida entre 30 e	0,176	0,091	0,207	0,260
Taxa de não obesidade	-0,159	0,152	-0,170	-0,051
Taxa de não suicídio	-0,042	0,031	0,274	0,572
Não escravidão moderna, tráfico	0,149	-0,070	0,107	-0,028
Satisfação com métodos contracep.	0,190	-0,168	0,103	-0,171
Respeito às mulheres	0,033	0,088	-0,419	0,437
Tolerância aos imigrantes	0,040	0,393	0,149	0,208
Tolerância ao homossexual	0,170	0,198	0,204	0,129
Não discriminação e violência	0,142	0,340	-0,027	-0,115
Tolerância Religiosa	0,029	0,285	0,158	-0,268
Redes de segurança comunitária	0,165	0,081	0,064	-0,126
Anos de Escolaridade Terciária	0,160	0,054	-0,031	0,018
Tempo de escolaridade feminina	0,210	-0,078	-0,070	-0,213
Igualdade no acesso à educação	0,192	-0,011	-0,174	-0,230
IDH - 2013_1_1	0,227	-0,075	0,002	0,011
Bem-estar	0,179	0,077	0,200	0,073
Índice de Governança	0,198	0,246	-0,089	0,018
EPI	0,221	0,040	-0,064	0,063
HPI	0,085	-0,205	0,364	0,004
GINNI	-0,085	0,087	0,450	-0,090
Índice de Democracia	0,171	0,229	0,066	-0,143

Análises:

- Observaram-se quatro variáveis com contribuição negativa, são elas: Índice de Liberdade de Imprensa (0,111), Taxa de não suicídio (0,159), Taxa de não obesidade (0,042) e GINNI (0,085).
- Na PC2, destacam-se 15 variáveis com contribuição negativa, são elas: Índice de Progresso social (0,005), Nutrição e Cuidados Médicos Básicos de Saúde (0,013), Segurança Pessoal (0,305),Acesso a Educação Avançada (0,107), Alimentação (0,068), Boas Instalações Sanitárias (0,015), Acesso a Eletricidade (0,242), Índice de Liberdade de Imprensa (0,367), Expectativa de Vida (0,112), Não escravidão moderna, tráfico de pessoas (0,70), Satisfação com métodos contraceptivos (0,168), Tempo de Escolaridade Feminina (0,078), Igualdade no acesso à educação (0,011), IDH (0,075) e HPI (0,205). Duas delas muito próximas de 0, Acesso a Educação Avançada com 0,004 e Igualdade no acesso à educação com 0,011.
- Na PC3, destacam-se 16 variáveis com contribuição negativa, são elas: Nutrição e Cuidados Médicos Básicos de Saúde (0,160), Água e Saneamento (0,143), Acesso a Educação Avançada (0,004), alimentação (0,119), Água Encanada (0,108), Boas Instalações Sanitárias (0,191), Qualidade no Fornecimento de Eletricidade (0,116), Índice de Liberdade de Imprensa (0,015), Taxa de não obesidade (0,170), Respeito a Mulher (0,419), Não discriminação e violência (0,027), Anos de escolaridade terciária (0,031), Tempo de Escolaridade Feminina (0,070), Igualdade no acesso à educação (0,174), Índice de Governaça (0,089) e EPI (0,064).
- Na PC4, destacam-se 13 variáveis com contribuição negativa, são elas: Nutrição e Cuidados Médicos Básicos de Saúde (0,006), Acesso a Educação Avançada (0,091), Acesso a Eletricidade (0,000), Taxa de não obesidade (0,051), Não escravidão moderna, tráfico de pessoas (0,028), Satisfação com métodos contraceptivos (0,171), Não discriminação e violência (0,115), Tolerância Religiosa (0, 268), Redes de segurança comunitária (0,126), Tempo de Escolaridade Feminina (0,313), Igualdade no acesso à educação (0,330), GINNI (0.090) e Índice de Democracia (0,143).

A seguir serão analisados separadamente cada um dos quatro componentes destacando os principais agrupamentos de variáveis para cada componente.

8.3.1 Stepwise Regression: PC1 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 33 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-8,915	-9,403	-9,203	-7,497	-7,635	-8,087
Índice de Progresso Social	0,1596	0,1288	0,1071	0,1052	0,0934	0,0883
T-Value	51,93	34,77	24,41	27,59	24,86	24,93
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
Boas Instalações Sanitárias		0,0311	0,0268	0,0183	0,0121	0,0136
T-Value		10,51	10,37	7,05	4,97	6,10
P-Value		0,000	0,000	0,000	0,000	0,000
Acesso a Educação Avançada			0,0318	0,0337	0,0337	0,0321
T-Value			7,11	8,65	10,00	10,46
P-Value			0,000	0,000	0,000	0,000
Taxa de não obesidade				-0,0175	-0,0186	-0,0157
T-Value				-6,59	-8,06	-7,28
P-Value				0,000	0,000	0,000
Expectativa de vida					0,0197	0,0187
T-Value					6,55	6,83
P-Value					0,000	0,000
Bem-estar						0,0112
T-Value						5,31
P-Value						0,000
S	0,898	0,662	0,562	0,487	0,422	0,383
R-Sq	95,40	97,52	98,23	98,68	99,01	99,20
R-Sq(adj)	95,37	97,48	98,18	98,64	98,97	99,16

Análise:
No componente PC1, temos seis variáveis de destaque, são elas: Índice de Progresso Social, Boas Instalações Sanitárias, Acesso a Educação Avançada, Taxa de Obesidade, Expectativa de Vida e Bem-estar. O Índice de progresso social explica 95,40% unindo-se a Boas Instalações Sanitárias o percentual sobre para 97,48%, sequencialmente nota-se que o aumento é muito pequeno, sendo assim o Índice de Progresso Social explica sozinho o PC1, sendo assim será nomeado de **Fator de Bem-estar Social (FBS).**

8.3.2. Stepwise Regression: PC2 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 33 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2,91570	-0,87554	1,71882	0,43668	0,06879
Tolerância aos imigrantes	0,0563	0,0411	0,0335	0,0265	0,0259
T-Value	11,27	8,76	10,53	11,39	12,79
P-Value	0,000	0,000	0,000	0,000	0,000
Índice de Liberdade de Imprensa		-0,0343	-0,0425	-0,0266	-0,0184
T-Value		-7,33	-13,38	-10,06	-7,03
P-Value		0,000	0,000	0,000	0,000
Acesso a Eletricidade			-0,0245	-0,0288	-0,0335

T-Value			-12,77	-20,51	-23,54
P-Value			0,000	0,000	0,000
Não discriminação e violência				0,0312	0,0228
T-Value				11,41	8,40
P-Value				0,000	0,000
Índice de Governança					0,0187
T-Value					6,45
P-Value					0,000
Tolerância Relogiosa					
T-Value					
P-Value					
S	1,28	1,08	0,717	0,506	0,440
R-Sq	49,43	64,30	84,30	92,25	94,18
R-Sq(adj)	49,04	63,75	83,94	92,01	93,94

Análise:

No componente PC2, temos seis variáveis de destaque, são elas: Tolerância aos Imigrantes, Índice de Liberdade de Imprensa, Acesso a Eletricidade, Não Discriminação e Violência, Índice de Governança e Tolerância Religiosa. A variável Tolerância aos Imigrantes explica apenas 49,43%, somada a variável Índice de Liberdade de Imprensa esse percentual sobre para 64,75%, somados a variável Acesso a Eletricidade o percentual aumenta para 83,94%, que somados a variável Não Discriminação e Violência passa a explicar 92,01%, a soma da quinta variável, Índice de Governança, aumenta muito pouco o percentual para 93,94%. Sendo assim conclui-se que para uma boa explicação do componente PC2 é necessária pelo menos à soma das três primeiras variáveis. Nomeia-se o PC2 como **Fator de Integridade e Igualdade Social 1 (FIIS1)**

8.3.3. Stepwise Regression: PC3 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3 on 33 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-1,720	-3,325	-1,592	-2,548	-2,563	-1,942
GINNI	0,0464	0,0446	0,0371	0,0305	0,0309	0,0261
T-Value	10,14	13,12	12,55	12,43	17,39	14,61
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
HPI		0,0345	0,0292	0,0285	0,0219	0,0210
T-Value		10,36	10,32	12,75	12,61	13,58
P-Value		0,000	0,000	0,000	0,000	0,000
Respeito às mulheres			-0,0221	-0,0249	-0,0279	-0,0248
T-Value			-7,86	-11,11	-16,96	-15,89
P-Value			0,000	0,000	0,000	0,000
Taxa de não suícidio				0,0199	0,0202	0,0186
T-Value				8,81	12,35	12,57
P-Value				0,000	0,000	0,000
Tolerância ao homossexual					0,0142	0,0186
T-Value					10,81	13,38
P-Value					0,000	0,000
Segurança Pessoal						-0,0108
T-Value						-5,84
P-Value						0,000
S	1,11	0,821	0,677	0,535	0,387	0,345

R-Sq	44,14	69,50	79,44	87,24	93,38	94,80
R-Sq(adj)	43,71	69,03	78,95	86,83	93,11	94,55

Análise:

No componente PC3, temos cinco variáveis de destaque, são elas: Índice de GINNI, HPI, Respeito às Mulheres, Taxa de não Suicídio, Tolerância ao Homossexual e Segurança pessoal. A variável Índice de GINNI explica apenas 43,14%, somada a variável HPI esse percentual sobre para 69,50%, somados a variável Respeito às Mulheres o percentual aumenta para 79,44%, que somados a variável Taxa de não Suicídio passa a explicar 87,24%. Sendo assim conclui-se que para uma boa explicação do componente PC3 é necessária, pelo menos, a soma das três primeiras variáveis. Nomeia-se o PC3 como **Fator de Integridade e Igualdade Social 2 (FIIS2)**

8.3.4. Stepwise Regression: PC4 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4 on 33 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2,655	-4,003	-3,367	-3,779	-3,043
Taxa de não suícidio	0,0381	0,0370	0,0366	0,0348	0,0289
T-Value	10,81	14,40	15,91	17,46	16,56
P-Value	0,000	0,000	0,000	0,000	0,000
Respeito às mulheres		0,0262	0,0246	0,0239	0,0249
T-Value		10,76	11,17	12,63	16,30
P-Value		0,000	0,000	0,000	0,000
Tolerância Relogiosa			-0,00881	-0,01008	-0,00989
T-Value			-5,73	-7,55	-9,22
P-Value			0,000	0,000	0,000
Expectativa de Vida entre 30 e				0,0115	0,0184
T-Value				6,82	11,62
P-Value				0,000	0,000
Tempo de escolaridade feminina					-0,0131
T-Value					-8,41
P-Value					0,000
Índice de Democracia					
T-Value					
P-Value					
S	0,876	0,639	0,572	0,491	0,395
R-Sq	47,35	72,24	77,90	83,83	89,64
R-Sq(adj)	46,94	71,81	77,38	83,32	89,23

Análise:

No componente PC4, temos cinco variáveis de destaque, são elas: Taxa de não Suicídio, Respeito às Mulheres, Tolerância Religiosa, Expectativa de Vida entre 30 70 anos, Tempo de Escolaridade Feminina e Índice de Democracia. A variável Taxa de não Suicídio explica apenas 47,35%, somada a variável Tolerância Religiosa esse percentual sobre para 72,24%, somados a variável Respeito às Mulheres o percentual aumenta para 77,90%, que somados a variável Expectativa de Vida entre 30 70 anos o passa a explicar 83,83%. Sendo assim conclui-se que para uma boa explicação do componente PC4 é necessária, pelo menos, a soma das três primeiras variáveis. Nomeia-se o PC4 como **Fator de Qualidade de Vida (FQV)**

CONSIDERAÇÕES FINAIS

Nota-se que as mesmas variáveis são destacadas com pouca correlação, tanto no dendograma quanto no gráfico de análise dos componentes principais.

Nota-se também que ao destacar os componentes com exceção do componente PC1 os que é explicado já com a primeira variável os demais componentes necessitam da soma das três primeiras variáveis.

A seguir serão analisados separadamente cada um dos quatro componentes destacando os principais agrupamentos de variáveis para cada componente.

3.1 Stepwise Regression: PC1 versus Índice de Progre; Nutrição e Cuida; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15


Response is PC1 on 33 predictors, with N = 132
```

Step	1	2	3	4	5	6
Constant	-8,915	-9,403	-9,203	-7,497	-7,635	-8,087
Índice de Progresso Social	0,1596	0,1288	0,1071	0,1052	0,0934	0,0883
T-Value	51,93	34,77	24,41	27,59	24,86	24,93
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
Boas Instalações Sanitárias		0,0311	0,0268	0,0183	0,0121	0,0136
T-Value		10,51	10,37	7,05	4,97	6,10
P-Value		0,000	0,000	0,000	0,000	0,000
Acesso a Educação Avançada			0,0318	0,0337	0,0337	0,0321
T-Value			7,11	8,65	10,00	10,46
P-Value			0,000	0,000	0,000	0,000
Taxa de não obesidade				-0,0175	-0,0186	-0,0157
T-Value				-6,59	-8,06	-7,28
P-Value				0,000	0,000	0,000
Expectativa de vida					0,0197	0,0187
T-Value					6,55	6,83
P-Value					0,000	0,000
Bem-estar						0,0112
T-Value						5,31
P-Value						0,000
S	0,898	0,662	0,562	0,487	0,422	0,383
R-Sq	95,40	97,52	98,23	98,68	99,01	99,20
R-Sq(adj)	95,37	97,48	98,18	98,64	98,97	99,16

Análise:
No componente PC1, temos seis variáveis de destaque, são elas: Índice de Progresso Social, Boas Instalações Sanitárias, Acesso a Educação Avançada, Taxa de Obesidade, Expectativa de Vida e Bem-estar. O Índice de progresso social explica 95,40% unindo-se a Boas Instalações Sanitárias o percentual sobre para 97,48%, sequencialmente nota-se que o aumento é muito pequeno, sendo assim o Índice de Progresso Social explica sozinho o PC1, sendo assim será nomeado de **Fator de Bem-estar Social (FBS).**

General Regression Analysis: PC1 versus Índice de Pr; Acesso a Edu; ...

```
Regression Equation

PC1  =  -9,20324 + 0,107078 Índice de Progresso Social + 0,0318469 Acesso a
        Educação Avançada + 0,0268328 Boas Instalações Sanitárias


Coefficients

Term                          Coef    SE Coef         T      P
Constant                  -9,20324  0,127646  -72,0999  0,000
```

```
Índice de Progresso Social    0,10708  0,004386   24,4123  0,000
Acesso a Educação Avançada    0,03185  0,004476    7,1148  0,000
Boas Instalações Sanitárias   0,02683  0,002588   10,3673  0,000


Summary of Model

S = 0,562193     R-Sq = 98,23%        R-Sq(adj) = 98,18%
PRESS = 43,0358  R-Sq(pred) = 98,11%


Analysis of Variance

Source                          DF   Seq SS   Adj SS   Adj MS        F
Regression                       3  2238,92  2238,92  746,307  2361,27
  Índice de Progresso Social     1  2174,54   188,36  188,361   595,96
  Acesso a Educação Avançada     1    30,41    16,00   15,999    50,62
  Boas Instalações Sanitárias    1    33,97    33,97   33,970   107,48
Error                          128    40,46    40,46    0,316
Total                          131  2279,38

Source                                P
Regression                    0,0000000
  Índice de Progresso Social  0,0000000
  Acesso a Educação Avançada  0,0000000
  Boas Instalações Sanitárias 0,0000000
Error
Total


Fits and Diagnostics for Unusual Observations

Obs       PC1       Fit    SE Fit  Residual  St Resid
 15  -2,11825  -0,35459  0,089476  -1,76366  -3,17761  R
 19  -8,23408  -6,90348  0,115762  -1,33060  -2,41863  R
 31   1,85052   0,20893  0,083922   1,64159   2,95306  R
 39   3,77696   4,89399  0,079010  -1,11703  -2,00683  R
 55  -2,75910  -3,98609  0,138963   1,22699   2,25240  R
 60   4,87170   6,15944  0,112957  -1,28774  -2,33825  R
 90  -0,18134  -1,31743  0,081878   1,13608   2,04258  R
 98  -0,12625   0,99193  0,083203  -1,11818  -2,01111  R
102   0,56872   0,68495  0,212124  -0,11623  -0,22325     X
103  -5,26243  -3,93715  0,081123  -1,32529  -2,38229  R
116  -2,10773  -0,78574  0,113766  -1,32199  -2,40117  R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Como o IPS – Índice de Progresso Social tem muito peso na equação, sendo um indicador muito complexo com muitas variáveis é mais interessante trocar para melhorar esse peso. Conforme se vê na regressão abaixo.

General Regression Analysis: PC1 versus IDH - 2013_1; Boas Instala; ...

```
Regression Equation

PC1  =  -8,80033 + 0,0780681 IDH - 2013_1_1 + 0,0234493 Boas Instalações
        Sanitárias + 0,0622501 Acesso a Educação Avançada
```

Os 93,34% de explicação dessa componente se dá pela soma de -8,80 + 0,78 de IDH + 0,02 de Boas Instalações + 0,06 de Acesso a Educação

```
Coefficients

Term                          Coef   SE Coef         T      P
Constant                  -8,80033  0,243165  -36,1908  0,000
IDH - 2013_1_1             0,07807  0,009359    8,3413  0,000
Boas Instalações Sanitárias  0,02345  0,005664    4,1397  0,000
Acesso a Educação Avançada   0,06225  0,008235    7,5591  0,000
```

Summary of Model

S = 1,07616 R-Sq = 93,50% R-Sq(adj) = 93,34%
PRESS = 161,860 R-Sq(pred) = 92,90%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F
Regression	3	2131,14	2131,14	710,380	613,395
IDH - 2013_1_1	1	2039,98	80,58	80,578	69,577
Boas Instalações Sanitárias	1	24,98	19,85	19,847	17,137
Acesso a Educação Avançada	1	66,17	66,17	66,174	57,139
Error	128	148,24	148,24	1,158	
Total	131	2279,38			

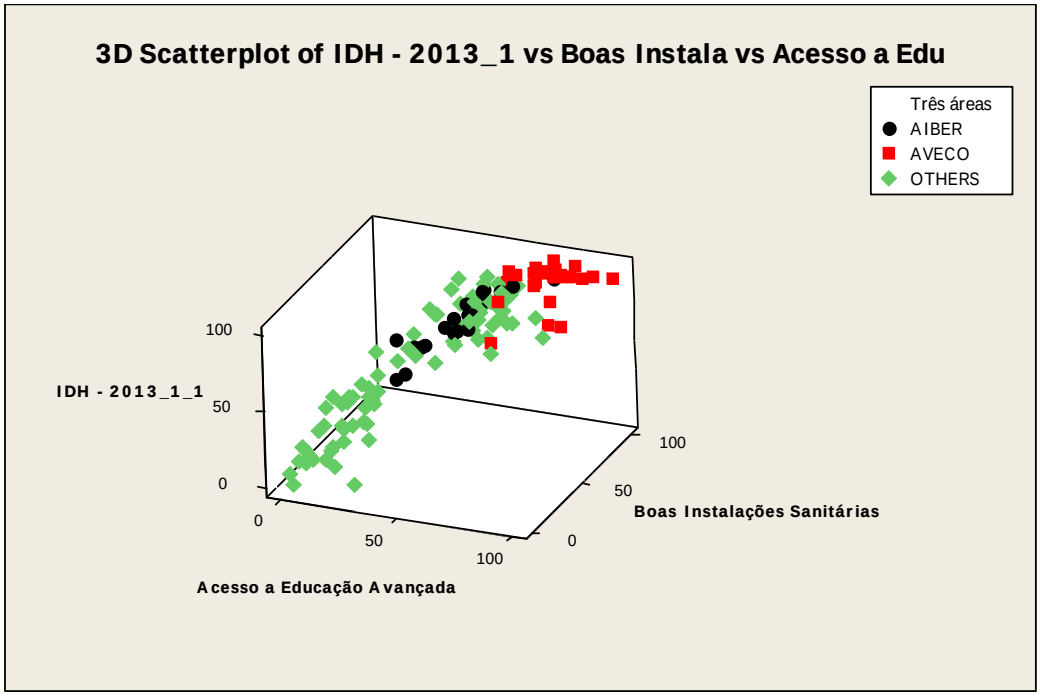
Source	P
Regression	0,0000000
IDH - 2013_1_1	0,0000000
Boas Instalações Sanitárias	0,0000626
Acesso a Educação Avançada	0,0000000
Error	
Total	

Fits and Diagnostics for Unusual Observations

Obs	PC1	Fit	SE Fit	Residual	St Resid		
28	-5,50215	-6,97766	0,345001	1,47551	1,44749		X
29	3,91352	1,70554	0,125372	2,20798	2,06579	R	
88	5,99208	2,73233	0,353006	3,25975	3,20648	R	X
90	-0,18134	-2,36967	0,121833	2,18832	2,04662	R	
102	0,56872	3,64364	0,299247	-3,07492	-2,97464	R	
109	-0,16541	2,69456	0,322149	-2,85997	-2,78530	R	

Aqui eu tenho os países que destoam

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.



127 MELHOR VERMELHO, 68 PIOR VERMELHO. Dois pretos finais 90 e 13, melhor pretos 50 e 110 Melhor verde 50, pior 24 e 91
Neste gráfico pode-se observar que os países da AVECO se destacam nessas três variáveis, sendo que o primeiro deles é os EUA com 100% de educação avançada, 99,54 % de Boas Instalações Sanitárias e 95% de IDH, e o último é a com

3.2. Stepwise Regression: PC2 versus Índice de Progre; Nutrição e Cuida; ...

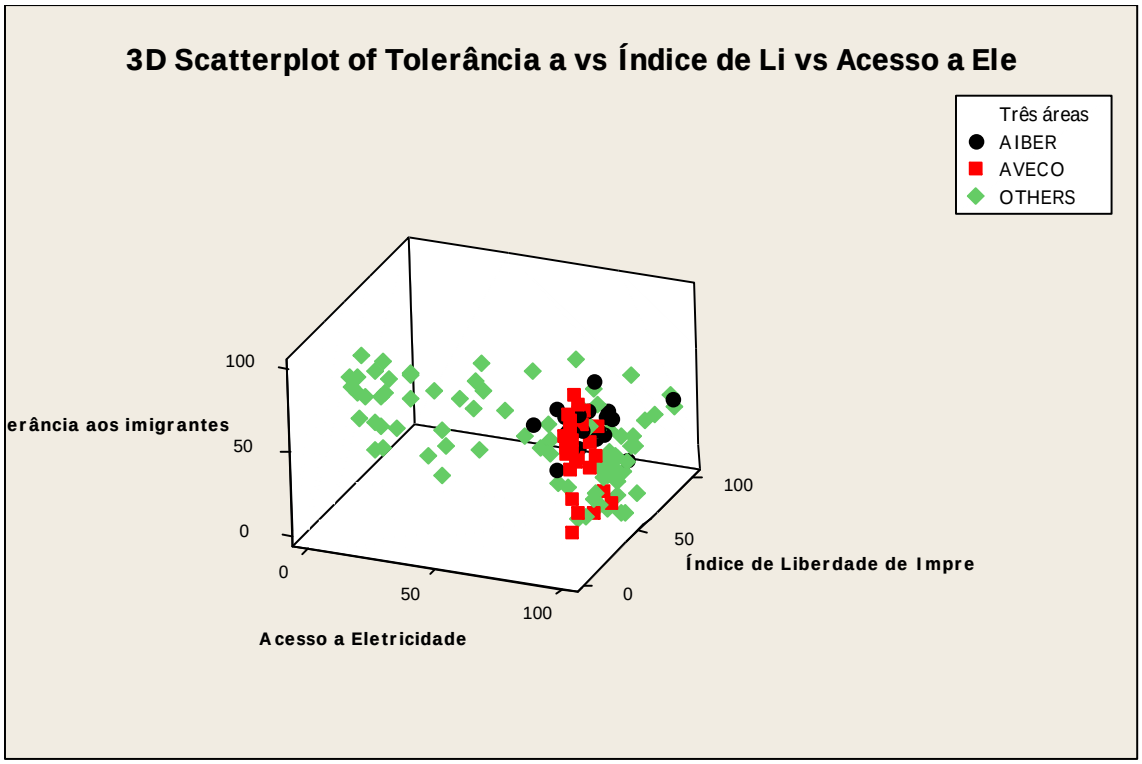
Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 33 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2,91570	-0,87554	1,71882	0,43668	0,06879
Tolerância aos imigrantes	0,0563	0,0411	0,0335	0,0265	0,0259
T-Value	11,27	8,76	10,53	11,39	12,79
P-Value	0,000	0,000	0,000	0,000	0,000
Índice de Liberdade de Imprensa		-0,0343	-0,0425	-0,0266	-0,0184
T-Value		-7,33	-13,38	-10,06	-7,03
P-Value		0,000	0,000	0,000	0,000
Acesso a Eletricidade			-0,0245	-0,0288	-0,0335
T-Value			-12,77	-20,51	-23,54
P-Value			0,000	0,000	0,000
Não discriminação e violência				0,0312	0,0228
T-Value				11,41	8,40
P-Value				0,000	0,000
Índice de Governança					0,0187
T-Value					6,45
P-Value					0,000
Tolerância Relogiosa					
T-Value					
P-Value					
S	1,28	1,08	0,717	0,506	0,440
R-Sq	49,43	64,30	84,30	92,25	94,18
R-Sq(adj)	49,04	63,75	83,94	92,01	93,94

Análise:

No componente PC2, temos seis variáveis de destaque, são elas: Tolerância aos Imigrantes, Índice de Liberdade de Imprensa, Acesso a Eletricidade, Não Discriminação e Violência, Índice de Governança e Tolerância Religiosa. A variável Tolerância aos Imigrantes explica apenas 49,43%, somada a variável Índice de Liberdade de Imprensa esse percentual sobre para 64,75%, somados a variável Acesso a Eletricidade o percentual aumenta para 83,94%, que somados a variável Não Discriminação e Violência passa a explicar 92,01%, a soma da quinta variável, Índice de Governança, aumenta muito pouco o percentual para 93,94%. Sendo assim conclui-se que para uma boa explicação do componente PC2 é necessária pelo menos à soma das três primeiras variáveis. Nomeia-se o PC2 como **Fator de Integridade e Igualdade Social 1 (FIIS1)**



Melhr preto 31 e 96 , pior preto 46 melhor vermelho 22 pior 39. Vrrde 67 104

3.3. Stepwise Regression: PC3 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

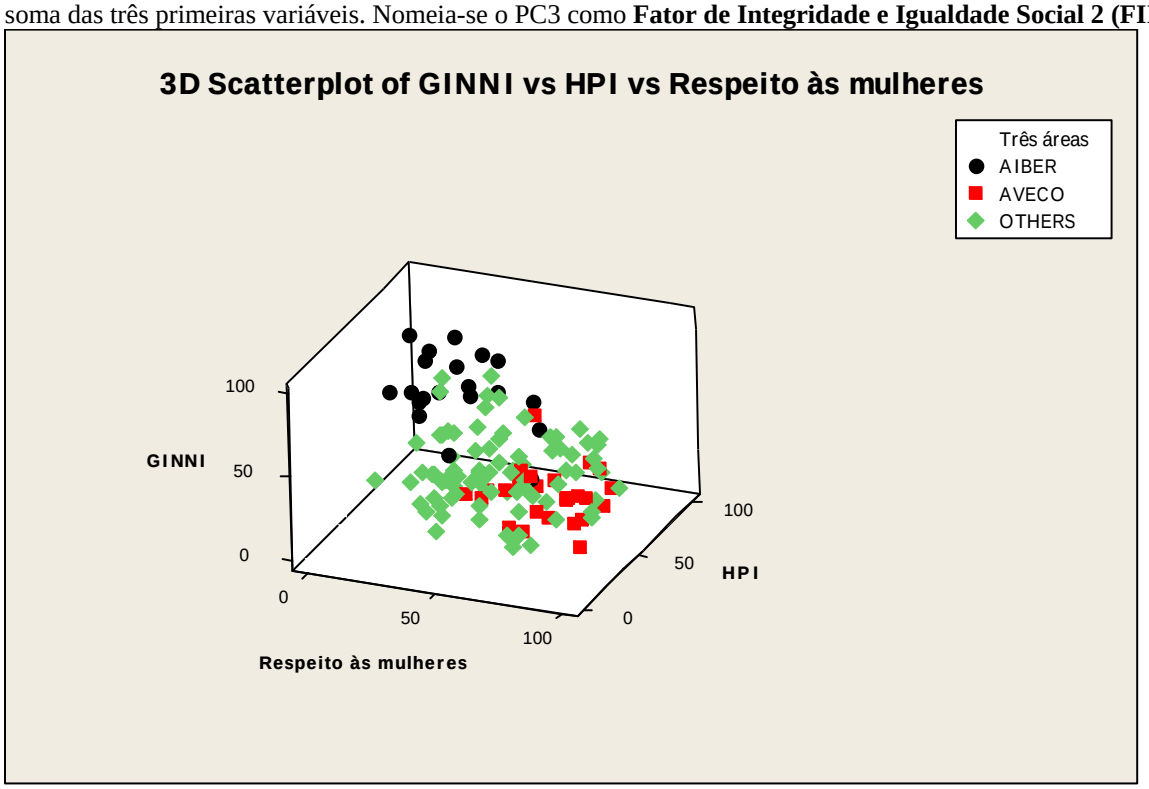
Response is PC3 on 33 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-1,720	-3,325	-1,592	-2,548	-2,563	-1,942
GINNI	0,0464	0,0446	0,0371	0,0305	0,0309	0,0261
T-Value	10,14	13,12	12,55	12,43	17,39	14,61
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
HPI		0,0345	0,0292	0,0285	0,0219	0,0210
T-Value		10,36	10,32	12,75	12,61	13,58
P-Value		0,000	0,000	0,000	0,000	0,000
Respeito às mulheres			-0,0221	-0,0249	-0,0279	-0,0248
T-Value			-7,86	-11,11	-16,96	-15,89
P-Value			0,000	0,000	0,000	0,000
Taxa de não suícidio				0,0199	0,0202	0,0186
T-Value				8,81	12,35	12,57
P-Value				0,000	0,000	0,000
Tolerância ao homossexual					0,0142	0,0186
T-Value					10,81	13,38
P-Value					0,000	0,000
Segurança Pessoal						-0,0108
T-Value						-5,84
P-Value						0,000
S	1,11	0,821	0,677	0,535	0,387	0,345

R-Sq	44,14	69,50	79,44	87,24	93,38	94,80
R-Sq(adj)	43,71	69,03	78,95	86,83	93,11	94,55

Análise:

No componente PC3, temos cinco variáveis de destaque, são elas: Índice de GINNI, HPI, Respeito às Mulheres, Taxa de não Suicídio, Tolerância ao Homossexual e Segurança pessoal. A variável Índice de GINNI explica apenas 43,14%, somada a variável HPI esse percentual sobre para 69,50%, somados a variável Respeito às Mulheres o percentual aumenta para 79,44%, que somados a variável Taxa de não Suicídio passa a explicar 87,24%. Sendo assim conclui-se que para uma boa explicação do componente PC3 é necessária, pelo menos, a soma das três primeiras variáveis. Nomeia-se o PC3 como **Fator de Integridade e Igualdade Social 2 (FIIS2)**



49 e 27 preto melhor , 128 pior. Vermelho 57 pior 33. 84 m4lgor verde pior 124

3.4. Stepwise Regression: PC4 versus Índice de Progre; Nutrição e Cuida; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

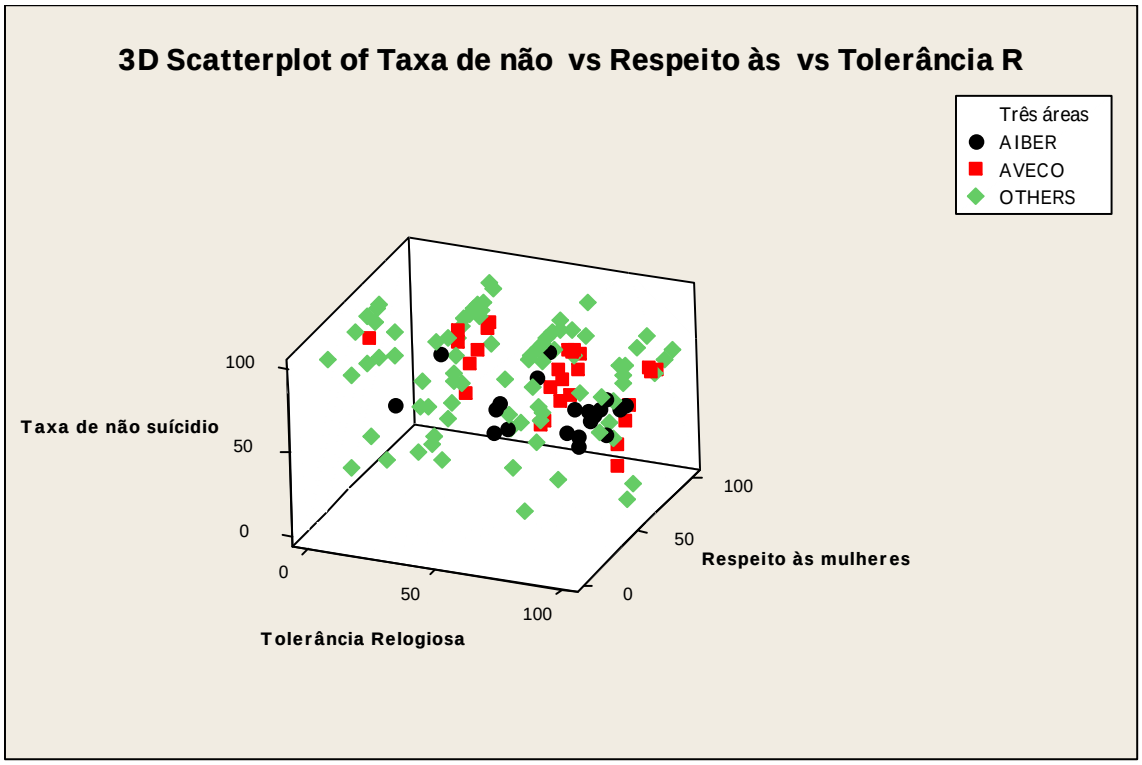
Response is PC4 on 33 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2,655	-4,003	-3,367	-3,779	-3,043
Taxa de não suícidio	0,0381	0,0370	0,0366	0,0348	0,0289
T-Value	10,81	14,40	15,91	17,46	16,56
P-Value	0,000	0,000	0,000	0,000	0,000
Respeito às mulheres		0,0262	0,0246	0,0239	0,0249
T-Value		10,76	11,17	12,63	16,30
P-Value		0,000	0,000	0,000	0,000
Tolerância Relogiosa			-0,00881	-0,01008	-0,00989
T-Value			-5,73	-7,55	-9,22
P-Value			0,000	0,000	0,000
Expectativa de Vida entre 30 e				0,0115	0,0184
T-Value				6,82	11,62
P-Value				0,000	0,000

Tempo de escolaridade feminina					-0,0131
T-Value					-8,41
P-Value					0,000
Índice de Democracia					
T-Value					
P-Value					
S	0,876	0,639	0,572	0,491	0,395
R-Sq	47,35	72,24	77,90	83,83	89,64
R-Sq(adj)	46,94	71,81	77,38	83,32	89,23

Análise:

No componente PC4, temos cinco variáveis de destaque, são elas: Taxa de não Suicídio, Respeito às Mulheres, Tolerância Religiosa, Expectativa de Vida entre 30 70 anos, Tempo de Escolaridade Feminina e Índice de Democracia. A variável Taxa de não Suicídio explica apenas 47,35%, somada a variável Tolerância Religiosa esse percentual sobre para 72,24%, somados a variável Respeito às Mulheres o percentual aumenta para 77,90%, que somados a variável Expectativa de Vida entre 30 70 anos o passa a explicar 83,83%. Sendo assim conclui-se que para uma boa explicação do componente PC4 é necessária, pelo menos, a soma das três primeiras variáveis. Nomeia-se o PC4 como **Fator de Qualidade de Vida (FQV)**



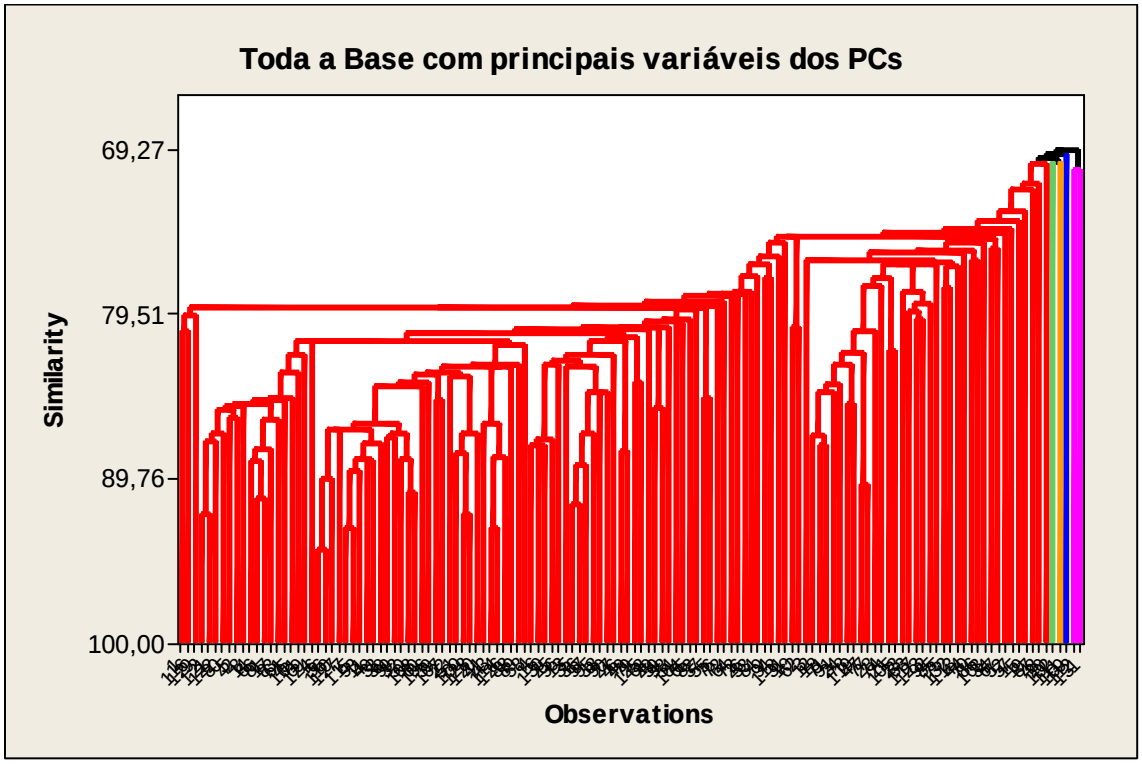
Vermelho 64 preto 27 verde 62 e 102

Análises Finais:

Nota-se que as mesmas variáveis são destacadas com pouca correlação, tanto no dendograma quanto no gráfico de análise dos componentes principais.
Nota-se também que ao destacar os componentes com exceção do componente PC1 os que é explicado já com a primeira variável os demais componentes necessitam da soma das três primeiras variáveis.
.

5. DEFINIÇÃO DAS VARIÁVEIS MAIS RELEVANTES

Foram identificados três agrupamentos.



General Regression Analysis: PC1 versus Índice de Pr; Acesso a Edu; ...

Regression Equation

PC1 = -9,20324 + 0,107078 Índice de Progresso Social + 0,0318469 Acesso a Educação Avançada + 0,0268328 Boas Instalações Sanitárias

Coefficients

Term	Coef	SE Coef	T	P
Constant	-9,20324	0,127646	-72,0999	0,000
Índice de Progresso Social	0,10708	0,004386	24,4123	0,000
Acesso a Educação Avançada	0,03185	0,004476	7,1148	0,000
Boas Instalações Sanitárias	0,02683	0,002588	10,3673	0,000

Summary of Model

S = 0,562193 R-Sq = 98,23% R-Sq(adj) = 98,18%
PRESS = 43,0358 R-Sq(pred) = 98,11%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	2238,92	2238,92	746,307	2361,27	0,0000000
Índice de Progresso Social	1	2174,54	188,36	188,361	595,96	0,0000000
Acesso a Educação Avançada	1	30,41	16,00	15,999	50,62	0,0000000
Boas Instalações Sanitárias	1	33,97	33,97	33,970	107,48	0,0000000
Error	128	40,46	40,46	0,316		
Total	131	2279,38				

```
Error
Total


Fits and Diagnostics for Unusual Observations

Obs       PC1       Fit    SE Fit  Residual  St Resid
 15  -2,11825  -0,35459  0,089476  -1,76366  -3,17761  R
 19  -8,23408  -6,90348  0,115762  -1,33060  -2,41863  R
 31   1,85052   0,20893  0,083922   1,64159   2,95306  R
 39   3,77696   4,89399  0,079010  -1,11703  -2,00683  R
 55  -2,75910  -3,98609  0,138963   1,22699   2,25240  R
 60   4,87170   6,15944  0,112957  -1,28774  -2,33825  R
 90  -0,18134  -1,31743  0,081878   1,13608   2,04258  R
 98  -0,12625   0,99193  0,083203  -1,11818  -2,01111  R
102   0,56872   0,68495  0,212124  -0,11623  -0,22325     X
103  -5,26243  -3,93715  0,081123  -1,32529  -2,38229  R
116  -2,10773  -0,78574  0,113766  -1,32199  -2,40117  R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Cluster Analysis of Observations: Acesso a Edu; Boas Instala; IDH - 2013_1; ...

```
Euclidean Distance, Single Linkage
Amalgamation Steps
```

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	94,2551	12,4069	6	56	6	2
2	130	92,8432	15,4561	43	114	43	2
3	129	92,8183	15,5099	7	115	7	2
4	128	92,0541	17,1604	2	121	2	2
5	127	92,0190	17,2362	100	128	100	2
6	126	91,4598	18,4438	35	36	35	2
7	125	91,0254	19,3820	66	81	66	2
8	124	90,6091	20,2811	50	108	50	2
9	123	90,2270	21,1063	47	77	47	2
10	122	89,8026	22,0229	6	110	6	3
11	121	89,2634	23,1873	7	93	7	3
12	120	88,9209	23,9269	35	97	35	3
13	119	88,7274	24,3449	14	66	14	3
14	118	88,6843	24,4380	11	40	11	2
15	117	88,6126	24,5928	7	11	7	5
16	116	88,5936	24,6339	32	50	32	3
17	115	88,4950	24,8467	43	126	43	3
18	114	88,2533	25,3688	51	100	51	3
19	113	88,0639	25,7778	46	49	46	2
20	112	87,9657	25,9899	14	17	14	4
21	111	87,7302	26,4984	16	130	16	2
22	110	87,6898	26,5857	4	16	4	3
23	109	87,6841	26,5980	18	71	18	2
24	108	87,6448	26,6829	7	33	7	6
25	107	87,5136	26,9663	2	122	2	3
26	106	87,3366	27,3485	4	25	4	4
27	105	87,1827	27,6809	30	99	30	2
28	104	87,1192	27,8179	12	18	12	3
29	103	86,9679	28,1447	30	32	30	5
30	102	86,9655	28,1499	51	120	51	4
31	101	86,9524	28,1782	2	80	2	4
32	100	86,9318	28,2227	35	95	35	4
33	99	86,7908	28,5273	6	127	6	4
34	98	86,6825	28,7611	6	7	6	10
35	97	86,3590	29,4598	41	43	41	4
36	96	86,3088	29,5682	6	30	6	15

37	95	86,1517	29,9074	14	83	14	5
38	94	85,9384	30,3682	5	42	5	2
39	93	85,5218	31,2679	2	5	2	6
40	92	85,3663	31,6037	29	59	29	2
41	91	85,2582	31,8370	2	8	2	7
42	90	85,1408	32,0908	78	119	78	2
43	89	85,0827	32,2160	2	14	2	12
44	88	84,9414	32,5213	39	107	39	2
45	87	84,9053	32,5993	2	61	2	13
46	86	84,8261	32,7703	37	55	37	2
47	85	84,7077	33,0261	2	45	2	14
48	84	84,4421	33,5995	38	90	38	2
49	83	84,3815	33,7304	35	38	35	6
50	82	84,3542	33,7894	12	91	12	4
51	81	84,0394	34,4694	6	60	6	16
52	80	83,8766	34,8210	12	24	12	5
53	79	83,8278	34,9263	6	106	6	17
54	78	83,7780	35,0338	79	125	79	2
55	77	83,4560	35,7293	22	51	22	5
56	76	83,2927	36,0820	6	39	6	19
57	75	83,1663	36,3549	2	101	2	15
58	74	83,1333	36,4262	6	22	6	24
59	73	82,8469	37,0446	41	88	41	5
60	72	82,8230	37,0962	6	41	6	29
61	71	82,7909	37,1656	13	35	13	7
62	70	82,7230	37,3123	4	96	4	5
63	69	82,6936	37,3757	6	89	6	30
64	68	82,6392	37,4932	12	78	12	7
65	67	82,4897	37,8160	4	13	4	12
66	66	82,0976	38,6630	4	27	4	13
67	65	82,0516	38,7622	2	69	2	16
68	64	81,9612	38,9574	12	47	12	9
69	63	81,7800	39,3489	21	105	21	2
70	62	81,4971	39,9597	6	58	6	31
71	61	81,2038	40,5933	2	124	2	17
72	60	81,1899	40,6233	4	46	4	15
73	59	81,1518	40,7055	2	6	2	48
74	58	80,9718	41,0942	4	79	4	17
75	57	80,7036	41,6734	2	4	2	65
76	56	80,6515	41,7859	1	116	1	2
77	55	80,5696	41,9629	12	28	12	10
78	54	80,4556	42,2091	2	68	2	66
79	53	80,4058	42,3167	2	29	2	68
80	52	80,3297	42,4809	10	62	10	2
81	51	80,3029	42,5388	2	98	2	69
82	50	80,0466	43,0924	54	104	54	2
83	49	80,0152	43,1602	2	54	2	71
84	48	79,8272	43,5661	2	65	2	72
85	47	79,8174	43,5873	117	123	117	2
86	46	79,5518	44,1610	1	129	1	3
87	45	79,3925	44,5050	63	117	63	3
88	44	79,1909	44,9404	2	82	2	73
89	43	79,1713	44,9828	1	2	1	76
90	42	79,1144	45,1056	1	37	1	78
91	41	78,9355	45,4920	1	76	1	79
92	40	78,8213	45,7385	63	92	63	4
93	39	78,7405	45,9132	1	72	1	80
94	38	78,4260	46,5923	1	64	1	81
95	37	78,3072	46,8488	1	73	1	82
96	36	78,2545	46,9626	1	26	1	83
97	35	78,1545	47,1787	23	63	23	5
98	34	78,0905	47,3169	1	53	1	84
99	33	77,9203	47,6845	75	132	75	2
100	32	77,7693	48,0105	12	74	12	11
101	31	77,2402	49,1531	12	21	12	13
102	30	77,2365	49,1612	9	94	9	2
103	29	77,1288	49,3938	1	84	1	85
104	28	76,5168	50,7155	75	113	75	3
105	27	76,4732	50,8097	12	23	12	18
106	26	76,4513	50,8570	12	85	12	19

107	25	76,4060	50,9548	1	9	1	87
108	24	76,2536	51,2838	12	75	12	22
109	23	76,1704	51,4636	70	86	70	2
110	22	76,1608	51,4842	3	12	3	23
111	21	75,9707	51,8949	1	118	1	88
112	20	75,6727	52,5384	3	44	3	24
113	19	75,5166	52,8756	3	70	3	26
114	18	75,4317	53,0590	34	67	34	2
115	17	75,0740	53,8314	1	31	1	89
116	16	75,0719	53,8360	3	103	3	27
117	15	74,8624	54,2883	3	34	3	29
118	14	74,7303	54,5738	1	10	1	91
119	13	74,7136	54,6098	1	3	1	120
120	12	74,4857	55,1019	1	52	1	121
121	11	74,2321	55,6497	1	57	1	122
122	10	73,7452	56,7012	1	15	1	123
123	9	73,0558	58,1901	1	19	1	124
124	8	71,7276	61,0585	1	87	1	125
125	7	71,3819	61,8051	1	48	1	126
126	6	70,5177	63,6714	112	131	112	2
127	5	70,1366	64,4944	1	20	1	127
128	4	70,0310	64,7225	102	111	102	2
129	3	69,8168	65,1851	1	102	1	129
130	2	69,4968	65,8763	1	109	1	130
131	1	69,2676	66,3712	1	112	1	132

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	127	838089	78,3901	134,989
Cluster2	1	0	0,0000	0,000
Cluster3	1	0	0,0000	0,000
Cluster4	1	0	0,0000	0,000
Cluster5	2	2027	31,8357	31,836

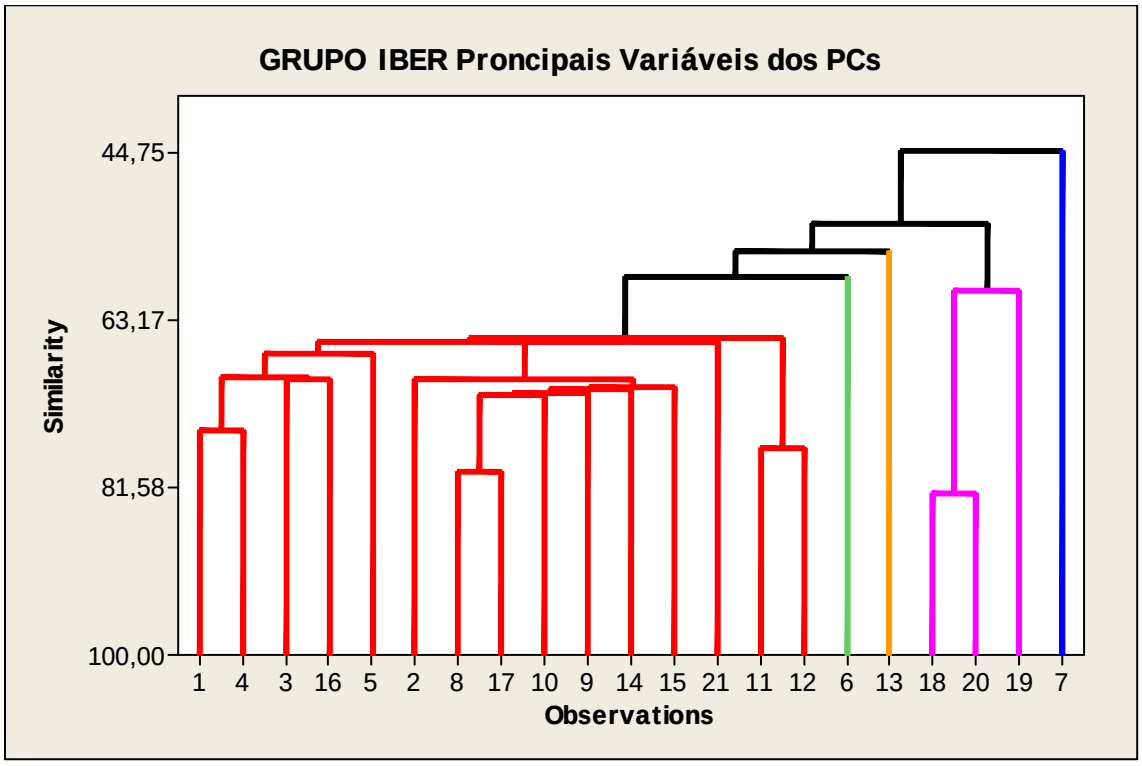
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Acesso a Educação Avançada	41,5536	83,438	42,457	45,3396
Boas Instalações Sanitárias	71,4680	67,300	71,203	90,1885
IDH - 2013_1_1	58,4603	72,652	92,000	68,0395
GINNI	36,6980	38,805	100,000	37,0564
Tolerância Relogiosa	61,1549	0,000	66,667	0,0000
Acesso a Eletricidade	77,8379	100,000	82,292	84,3750
Tolerância aos imigrantes	52,5510	30,137	49,315	32,8767
Índice de Liberdade de Imprensa	35,3131	55,267	27,126	74,9179
HPI	48,8481	28,778	13,509	64,6437
Taxa de não suícidio	70,3922	16,284	90,531	0,0000

Variable	Cluster5	Grand centroid
Acesso a Educação Avançada	9,3207	41,4180
Boas Instalações Sanitárias	31,7184	70,9740
IDH - 2013_1_1	24,6293	58,3819
GINNI	28,1315	37,0664
Tolerância Relogiosa	0,0000	59,3434
Acesso a Eletricidade	34,3750	77,4306
Tolerância aos imigrantes	26,0274	51,8057
Índice de Liberdade de Imprensa	94,3897	36,5974
HPI	42,6569	48,4542
Taxa de não suícidio	79,8454	69,7448

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,000	102,174	83,157	106,768	117,418
Cluster2	102,174	0,000	132,202	64,722	138,803
Cluster3	83,157	132,202	0,000	150,726	158,530
Cluster4	106,768	64,722	150,726	0,000	128,316
Cluster5	117,418	138,803	158,530	128,316	0,000



Cluster Analysis of Observations: Acesso a Edu; Boas Instala; IDH - 2013_1; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	82,3808	21,8412	18	20	18	2
2	19	79,8159	25,0207	8	17	8	2
3	18	77,3987	28,0172	11	12	11	2
4	17	75,4270	30,4614	1	4	1	2
5	16	71,3856	35,4712	8	10	8	3
6	15	71,2554	35,6326	8	9	8	4
7	14	70,8151	36,1785	8	14	8	5
8	13	70,5364	36,5239	8	15	8	6
9	12	69,7616	37,4844	2	8	2	7
10	11	69,6443	37,6298	3	16	3	2
11	10	69,5462	37,7514	1	3	1	4
12	9	66,9868	40,9241	1	5	1	5
13	8	65,6726	42,5532	2	21	2	8
14	7	65,6586	42,5706	1	2	1	13
15	6	65,2646	43,0590	1	11	1	15
16	5	59,9043	49,7038	18	19	18	3

17	4	58,4730	51,4780	1	6	1	16
18	3	55,6074	55,0304	1	13	1	17
19	2	52,5579	58,8106	1	18	1	20
20	1	44,7488	68,4909	1	7	1	21

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	15	24448,7	39,3928	59,8206
Cluster2	1	0,0	0,0000	0,0000
Cluster3	1	0,0	0,0000	0,0000
Cluster4	1	0,0	0,0000	0,0000
Cluster5	3	2286,9	26,0904	36,9536

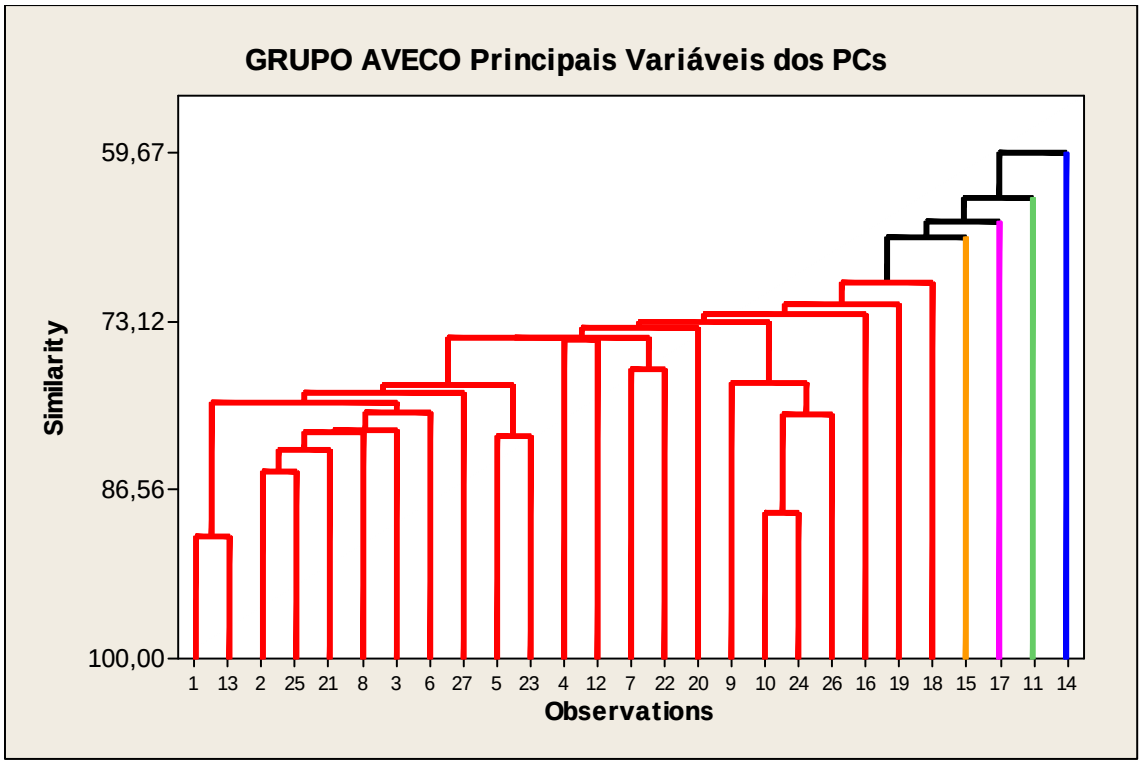
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
Acesso a Educação Avançada	37,4672	45,706	46,627	45,3869	56,9955	41,4626
Boas Instalações Sanitárias	75,1423	93,042	91,240	83,0808	99,5833	80,6308
IDH - 2013_1_1	61,1862	70,181	78,748	69,0280	80,7249	65,6154
GINNI	61,8841	62,082	37,056	60,6733	35,7168	56,9154
HPI	75,6212	100,000	81,060	73,1163	43,6628	72,3563
Tolerância Relogiosa	84,4444	100,000	100,000	33,3333	88,8889	84,1270
Respeito às mulheres	19,1870	24,390	54,390	41,4634	50,0000	26,5738
Taxa de não suícidio	77,7434	78,196	65,102	87,5696	69,1624	76,4050
Tolerância aos imigrantes	60,5479	78,082	41,096	30,1370	81,2785	61,9700
Índice de Liberdade de Imprensa	36,5851	8,505	97,374	58,0722	16,9253	36,3573
Acesso a Eletricidade	90,1389	98,958	100,000	98,9583	99,6528	92,8075

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,0000	50,065	84,317	69,745	70,8248
Cluster2	50,0652	0,000	106,774	102,004	71,3650
Cluster3	84,3174	106,774	0,000	86,981	99,1327
Cluster4	69,7454	102,004	86,981	0,000	99,2037
Cluster5	70,8248	71,365	99,133	99,204	0,0000

Dendrogram



Welcome to Minitab, press F1 for help.

Cluster Analysis of Observations: Acesso a Edu; Boas Instala; IDH - 2013_1; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	90,3248	13,8243	1	13	1	2
2	25	88,3714	16,6154	10	24	10	2
3	24	85,0760	21,3240	2	25	2	2
4	23	83,4168	23,6948	2	21	2	3
5	22	82,2392	25,3774	5	23	5	2
6	21	82,0426	25,6583	2	8	2	4
7	20	81,8832	25,8861	2	3	2	5
8	19	80,6290	27,6781	10	26	10	3
9	18	80,3931	28,0152	2	6	2	6
10	17	79,7089	28,9928	1	2	1	8
11	16	78,8994	30,1494	1	27	1	9
12	15	78,2088	31,1362	1	5	1	11
13	14	77,9976	31,4380	9	10	9	4
14	13	76,9848	32,8851	7	22	7	2
15	12	74,6327	36,2459	4	12	4	2
16	11	74,4494	36,5078	1	4	1	13
17	10	74,3894	36,5936	1	7	1	15
18	9	73,7169	37,5544	1	20	1	16
19	8	73,1547	38,3577	1	9	1	20
20	7	72,6575	39,0681	1	16	1	21
21	6	71,7992	40,2944	1	19	1	22
22	5	70,1526	42,6473	1	18	1	23
23	4	66,4783	47,8972	1	15	1	24
24	3	65,1460	49,8009	1	17	1	25
25	2	63,3271	52,3998	1	11	1	26
26	1	59,6739	57,6196	1	14	1	27

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	23	44299,6	41,3236	81,6822
Cluster2	1	0,0	0,0000	0,0000
Cluster3	1	0,0	0,0000	0,0000
Cluster4	1	0,0	0,0000	0,0000
Cluster5	1	0,0	0,0000	0,0000

Cluster Centroids

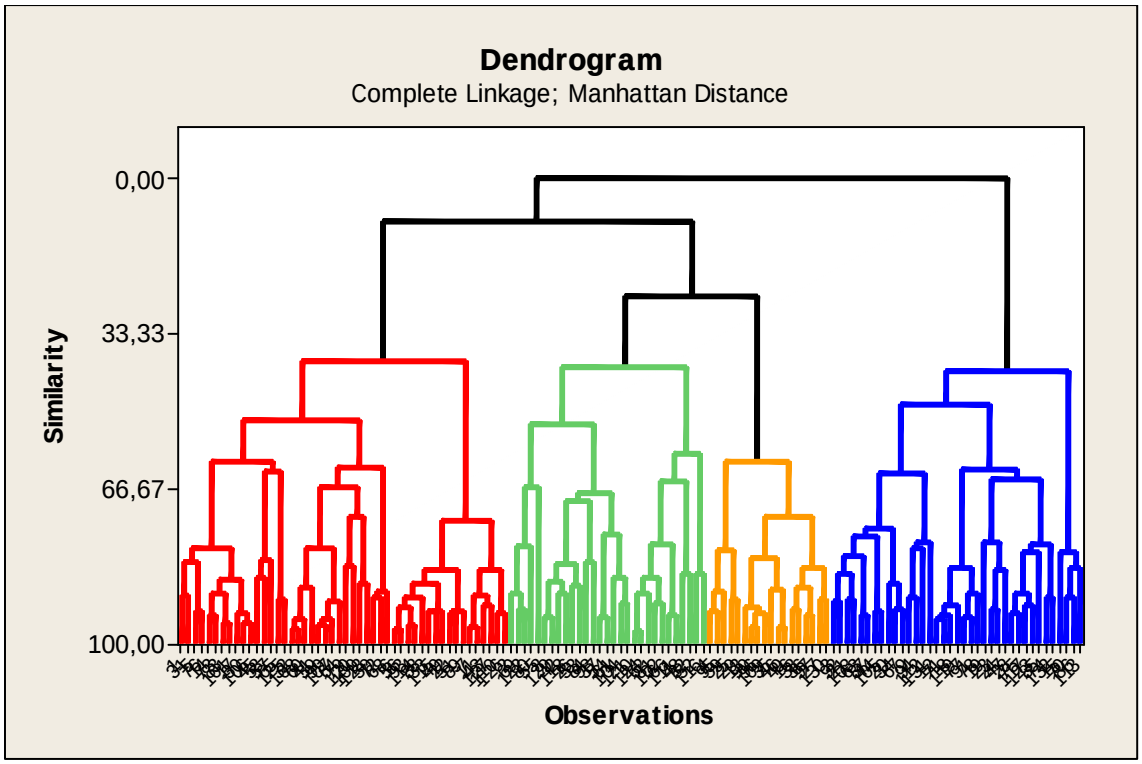
Variable	Cluster1	Cluster2	Cluster3	Cluster4
Acesso a Educação Avançada	73,263	67,005	80,449	62,197
Boas Instalações Sanitárias	97,516	98,489	100,000	70,969
IDH - 2013_1_1	88,158	85,008	90,774	88,138
Tolerância Relogiosa	68,116	33,333	0,000	33,333
Respeito às mulheres	73,224	57,317	53,659	57,317
Taxa de não suícidio	61,471	91,964	81,040	84,717
GINNI	18,435	25,887	46,868	27,923
HPI	49,362	43,272	78,689	57,331
Tolerância aos imigrantes	66,766	27,397	16,438	68,493
Índice de Liberdade de Imprensa	9,359	32,945	39,675	29,439
Acesso a Eletricidade	100,000	100,000	100,000	100,000

Variable	Cluster5	Grand centroid
Acesso a Educação Avançada	71,270	72,8140
Boas Instalações Sanitárias	100,000	96,7526
IDH - 2013_1_1	56,507	86,9653
Tolerância Relogiosa	100,000	64,1975
Respeito às mulheres	46,341	70,3252
Taxa de não suícidio	33,903	63,1652
GINNI	17,667	20,0872
HPI	51,128	50,5832
Tolerância aos imigrantes	49,315	62,8615
Índice de Liberdade de Imprensa	27,007	12,7530
Acesso a Eletricidade	92,708	99,7299

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,000	68,1285	102,911	58,2246	64,684
Cluster2	68,129	0,0000	57,620	52,3998	97,297
Cluster3	102,911	57,6196	0,000	77,1391	128,222
Cluster4	58,225	52,3998	77,139	0,0000	98,199
Cluster5	64,684	97,2975	128,222	98,1994	0,000

9. DEFINIÇÃO DAS VARIÁVEIS MAIS RELEVANTES



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	97,4619	0,5342	10	124	10	2
2	130	97,1991	0,5895	6	56	6	2
3	129	97,0265	0,6258	32	108	32	2
4	128	96,7476	0,6846	90	96	90	2
5	127	96,5720	0,7215	7	11	7	2
6	126	96,5323	0,7299	46	49	46	2
7	125	96,1052	0,8197	39	99	39	2
8	124	95,9745	0,8473	93	114	93	2
9	123	95,6777	0,9097	30	106	30	2
10	122	95,6132	0,9233	83	101	83	2
11	121	95,1167	1,0278	32	60	32	3
12	120	95,0208	1,0480	39	107	39	3
13	119	94,8157	1,0912	12	71	12	2
14	118	94,5657	1,1438	77	91	77	2
15	117	94,5225	1,1529	37	54	37	2
16	116	94,3525	1,1886	8	129	8	2
17	115	94,2432	1,2117	66	111	66	2
18	114	94,1490	1,2315	36	97	36	2
19	113	94,0269	1,2572	12	18	12	3
20	112	93,9166	1,2804	14	73	14	2
21	111	93,8529	1,2938	63	87	63	2
22	110	93,7031	1,3253	110	126	110	2
23	109	93,6595	1,3345	90	109	90	3
24	108	93,6547	1,3355	17	30	17	3
25	107	93,6368	1,3393	28	75	28	2
26	106	93,4758	1,3732	22	51	22	2
27	105	93,4065	1,3877	4	95	4	2
28	104	93,3170	1,4066	5	76	5	2
29	103	93,2896	1,4124	33	115	33	2
30	102	93,2532	1,4200	22	89	22	3

31	101	93,1714	1,4372	44	105	44	2
32	100	93,1059	1,4510	33	40	33	3
33	99	92,8983	1,4947	69	122	69	2
34	98	92,8517	1,5045	34	67	34	2
35	97	92,8370	1,5076	16	38	16	2
36	96	92,6700	1,5428	23	24	23	2
37	95	92,4514	1,5888	13	90	13	4
38	94	92,2360	1,6341	6	93	6	4
39	93	92,0409	1,6752	12	119	12	4
40	92	91,9856	1,6868	117	123	117	2
41	91	91,7718	1,7318	43	127	43	2
42	90	91,6137	1,7651	104	121	104	2
43	89	91,3314	1,8245	81	102	81	2
44	88	91,2764	1,8361	39	64	39	4
45	87	90,9018	1,9149	3	92	3	2
46	86	90,8064	1,9350	21	103	21	2
47	85	90,7067	1,9560	25	29	25	2
48	84	90,6398	1,9701	27	130	27	2
49	83	90,6270	1,9728	65	125	65	2
50	82	90,3588	2,0292	50	72	50	2
51	81	90,3337	2,0345	74	132	74	2
52	80	90,3186	2,0377	70	86	70	2
53	79	90,3035	2,0409	6	88	6	5
54	78	90,1687	2,0692	9	94	9	2
55	77	89,8897	2,1280	1	31	1	2
56	76	89,6815	2,1718	7	43	7	4
57	75	89,5234	2,2051	77	78	77	3
58	74	89,5218	2,2054	42	118	42	2
59	73	89,4452	2,2215	28	117	28	4
60	72	89,4000	2,2310	10	62	10	3
61	71	89,2536	2,2618	2	69	2	3
62	70	89,2441	2,2638	14	83	14	4
63	69	89,0502	2,3046	50	68	50	3
64	68	88,9383	2,3282	4	59	4	3
65	67	88,4768	2,4253	35	46	35	3
66	66	88,0838	2,5080	32	41	32	4
67	65	88,0080	2,5240	16	36	16	4
68	64	87,7128	2,5861	26	53	26	2
69	63	87,4836	2,6344	20	34	20	3
70	62	87,4102	2,6498	6	33	6	8
71	61	87,3250	2,6677	100	128	100	2
72	60	86,8818	2,7610	8	42	8	4
73	59	86,6309	2,8138	14	17	14	7
74	58	85,9978	2,9471	61	104	61	3
75	57	85,9718	2,9526	10	81	10	5
76	56	85,8186	2,9848	45	58	45	2
77	55	85,1282	3,1301	52	116	52	2
78	54	85,0528	3,1460	48	82	48	2
79	53	85,0456	3,1475	3	21	3	4
80	52	84,4163	3,2799	6	22	6	11
81	51	84,1362	3,3389	7	110	7	6
82	50	84,1139	3,3436	23	47	23	3
83	49	83,9363	3,3810	16	27	16	6
84	48	83,9141	3,3856	70	113	70	3
85	47	83,7441	3,4214	12	77	12	7
86	46	83,5713	3,4578	79	120	79	2
87	45	82,8758	3,6042	8	26	8	6
88	44	82,5056	3,6821	84	98	84	2
89	43	82,4307	3,6979	1	5	1	4
90	42	82,0345	3,7812	45	57	45	3
91	41	81,8310	3,8241	13	35	13	7
92	40	81,3722	3,9207	3	63	3	6
93	39	80,4224	4,1206	15	70	15	4
94	38	80,2833	4,1498	28	85	28	5
95	37	80,1142	4,1854	4	25	4	5
96	36	79,8481	4,2414	32	39	32	8
97	35	79,7574	4,2605	1	14	1	11
98	34	79,6219	4,2890	9	112	9	3
99	33	79,1242	4,3938	2	80	2	4
100	32	78,8659	4,4482	10	66	10	7

101	31	78,6902	4,4851	28	74	28	7
102	30	78,3744	4,5516	9	131	9	4
103	29	78,1675	4,5951	19	23	19	4
104	28	77,2487	4,7885	3	44	3	8
105	27	76,8581	4,8707	37	61	37	5
106	26	75,2919	5,2004	3	20	3	11
107	25	73,8288	5,5083	6	7	6	17
108	24	73,0627	5,6696	79	100	79	4
109	23	72,9487	5,6936	13	16	13	13
110	22	69,4155	6,4372	8	84	8	8
111	21	67,8252	6,7719	8	37	8	13
112	20	66,5651	7,0372	2	55	2	5
113	19	66,4108	7,0696	32	79	32	12
114	18	65,3013	7,3031	10	48	10	9
115	17	64,9791	7,3710	19	28	19	11
116	16	63,6342	7,6540	3	9	3	15
117	15	63,1856	7,7484	45	65	45	5
118	14	62,7437	7,8415	12	19	12	18
119	13	62,2723	7,9407	32	50	32	15
120	12	60,8344	8,2433	1	45	1	16
121	11	60,7814	8,2545	4	13	4	18
122	10	59,3993	8,5454	10	52	10	11
123	9	53,1233	9,8663	2	8	2	18
124	8	52,0412	10,0940	1	32	1	31
125	7	48,7155	10,7940	3	12	3	33
126	6	41,4153	12,3305	3	15	3	37
127	5	40,5140	12,5202	2	10	2	29
128	4	39,3007	12,7756	1	6	1	48
129	3	25,5301	15,6739	2	4	2	47
130	2	9,4418	19,0601	1	2	1	95
131	1	0,0000	21,0473	1	3	1	132

Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	48	369,367	2,64285	3,85943
Cluster2	29	192,351	2,42454	4,05557
Cluster3	37	226,791	2,31754	4,17899
Cluster4	18	56,897	1,60313	3,36716

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
PC1	4,06234	-0,23845	-5,52190	0,90183	-0,0000000
PC2	0,50340	-2,17106	1,08410	-0,07303	0,0000000
PC3	-0,62827	-0,48614	-0,27564	3,02522	-0,0000000
PC4	0,07455	-0,00179	0,15788	-0,52046	-0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,00000	5,06711	9,60865	4,90134
Cluster2	5,06711	0,00000	6,21134	4,27792
Cluster3	9,60865	6,21134	0,00000	7,34568
Cluster4	4,90134	4,27792	7,34568	0,00000

Cluster 1 Vermelho
Cluster 2 Verde
Cluster 3 Azul
Cluster 4 Laraja

Análise:

Os mais parecidos são os

TRABALHO 9 – Análise Discriminante

O que separa um grupo do outro?

Discriminant Analysis: C47 versus Acesso a Edu; Boas Instala; ...

Linear Method for Response: C47

Predictors: Acesso a Educação Avançada; Boas Instalações Sanitárias;
IDH - 2013_1_1; HPI; GINNI; Tolerância Relogiosa;
Respeito às mulheres; Taxa de não suícidio;
Índice de Liberdade de Imprensa; Acesso a Eletricidade;
Tolerância aos imigrantes

Group	1	2	3	4
Count	48	29	37	18

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	44	2	0	0
2	4	27	1	0
3	0	0	36	0
4	0	0	0	18
Total N	48	29	37	18
N correct	44	27	36	18
Proportion	0,917	0,931	0,973	1,000

N = 132 N Correct = 125 Proportion Correct = 0,947 (ACERTOU 94%)

44 ELE ACEROU NO GRUPO 1 QUE ERA O MELHOR

O GRUPO 4 FOI O MELHHOR POIS TEVE 100% DE ACERTO

Squared Distance Between Groups

	1	2	3	4
1	0,0000	8,0061	56,6824	22,1656
2	8,0061	0,0000	37,3201	19,5743
3	56,6824	37,3201	0,0000	46,3329
4	22,1656	19,5743	46,3329	0,0000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-60,546	-51,210	-25,206	-66,250
Acesso a Educação Avançada	0,324	0,297	0,093	0,205
Boas Instalações Sanitárias	0,142	0,098	-0,034	0,083
IDH - 2013_1_1	0,105	-0,026	0,005	0,092
HPI	0,031	0,065	0,110	0,185
GINNI	0,046	0,099	0,156	0,215
Tolerância Relogiosa	0,137	0,110	0,097	0,169
Respeito às mulheres	-0,039	-0,004	0,054	-0,157
Taxa de não suícidio	0,191	0,176	0,147	0,207
Índice de Liberdade de Imprensa	0,101	0,177	0,181	0,149
Acesso a Eletricidade	0,503	0,508	0,153	0,470
Tolerância aos imigrantes	0,092	0,069	0,132	0,164

Summary of Misclassified Observations
A BAIXO ESTÁO OS PAISE PIORES CLASSIFICADOS – ELE DÁ O NUMERO DOS PAISES AO LADO DOS ASTERISCOS

Observation	True Group	Pred Group	Group	Squared Distance	Probability
10**	2	1	1	15,51	0,662
			2	16,85	0,338
			3	65,59	0,000
			4	43,06	0,000
31**	1	2	1	29,52	0,015
			2	22,12	0,603
			3	62,15	0,000
			4	23,03	0,382
73**	1	2	1	10,633	0,411
			2	9,912	0,589
			3	54,531	0,000
			4	27,788	0,000
76**	1	2	1	20,16	0,021
			2	12,44	0,979
			3	59,76	0,000
			4	40,41	0,000
79**	1	2	1	11,54	0,385
			2	10,61	0,613
			3	45,60	0,000
			4	22,06	0,002
94**	3	2	1	50,06	0,000
			2	22,76	0,973
			3	29,90	0,027
			4	42,70	0,000
124**	2	1	1	13,63	0,711
			2	15,43	0,289
			3	76,00	0,000
	4	46,11		0,000	

A seguir será analisado o quadrático, para se verificar se o modo quadrático propicia uma análise melhor do que o linear.

Discriminant Analysis: 4 GRUPOS versus Acesso a Edu; Boas Instala; ...

Quadratic Method for Response: 4 GRUPOS

Predictors: Acesso a Educação Avançada; Boas Instalações Sanitárias; IDH - 2013_1_1; HPI; GINNI; Tolerância Relogiosa; Respeito às mulheres; Taxa de não suícidio; Índice de Liberdade de Imprensa; Acesso a Eletricidade; Tolerância aos imigrantes

Group	1	2	3	4
Count	48	29	37	18

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	47	2	0	0
2	1	27	0	0
3	0	0	37	0
4	0	0	0	18
Total N	48	29	37	18
N correct	47	27	37	18
Proportion	0,979	0,931	1,000	1,000

NESTE CASO O QUADRÁTICO É O MELHOR COM 97% DE ACERTO E DOIS GRUPOS COM 100%

N = 132 N Correct = 129 Proportion Correct = 0,977

From Generalized Squared Distance to Group

Group	1	2	3	4
1	53,13	75,72	183,51	195,94
2	67,39	58,05	109,45	238,58
3	2880,85	212,48	59,89	278,76
4	152,45	75,68	105,59	49,72

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
64**	1	2	1	88,42	0,000
			2	70,98	1,000
			3	219,77	0,000
			4	435,58	0,000
69**	2	1	1	64,65	0,811
			2	67,56	0,189
			3	138,31	0,000
			4	302,93	0,000
122**	2	1	1	63,94	0,515
			2	64,06	0,485
			3	113,29	0,000
			4	116,44	0,000

CONSIDERAÇÕES FINAIS

Nota-se que o melhor resultado está no quadrático com dois grupos com acerto de 100%.

10. REGRESSÃO LOGÍSTICA

É uma técnica estatística que tem como objetivo produzir, a partir de um conjunto de observações, um modelo que permita a predição de valores tomados por uma variável categórica, frequentemente binária, a partir de uma série de variáveis explicativas contínuas e/ou binárias.

A regressão logística é amplamente usada em ciências médicas e sociais, e tem outras denominações, como modelo logístico, modelo logit, e classificador de máxima entropia. A regressão logística é utilizada em áreas como as seguintes:

- Em medicina, permite por exemplo determinar os fatores que caracterizam um grupo de indivíduos doentes em relação a indivíduos sãos.
- No domínio dos seguros, permite encontrar fracções da clientela que sejam sensíveis a determinada política securitária em relação a um dado risco particular.
- Em instituições financeiras, pode detectar os grupos de risco para a subscrição de um crédito.

- Em econometria, permite explicar uma variável discreta, como por exemplo as intenções de voto em atos eleitorais.

O êxito da regressão logística assenta-se sobretudo nas numerosas ferramentas que permitem interpretar de modo aprofundado os resultados obtidos, que distingue-se essencialmente pelo fato de a variável resposta ser categórica. Enquanto método de predição para variáveis categóricas, a regressão logística é comparável às técnicas supervisionadas propostas em aprendizagem automática (árvores de decisão, redes neuronais, etc.), ou ainda a análise discriminante preditiva em estatística exploratória. É útil para modelar a probabilidade de um evento ocorrer como função de outros fatores.

Esta etapa do trabalho tem por objetivo efetuar a regressão logística e comparações utilizando o software Minitab para analisar se o modelo usado para a criação dos conglomerados são satisfatórios, e se os 132 países da base de dados estão bem distribuídos e agrupados corretamente de acordo com as variáveis usadas para a definição desse modelo de conglomerados que são: AVECO, AIBER e OTHERS.

Foram destacadas as 33 variáveis relacionadas à bem estar e qualidade de vida e essas foram agrupadas em quatro grupos sendo destacados os componentes principais de cada um dos grupos, são eles:

1. PC1: índice de progresso social (sendo substituído pelo IDH); boas instalações sanitárias e acesso a educação avançada.
2. PC2: tolerância aos imigrantes; índice de liberdade de imprensa; e acesso a eletricidade.
3. PC3: GINNI, HPI e respeito às mulheres.
4. PC4: taxa de não suicídio; respeito às mulheres e tolerância religiosa.

ANÁLISE DISCRIMINANTE

10.1.Quatro agrupamentos:

One-way ANOVA: IDH - 2013_1_1 versus 4 GRUPOS

Source	DF	SS	MS	F	P
4 GRUPOS	3	68031	22677	150,32	0,000
Error	128	19311	151		
Total	131	87342			

S = 12,28 R-Sq = 77,89% R-Sq(adj) = 77,37%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	48	81,12	12,11	(-*)
2	29	60,88	9,70	(-*--)
3	37	24,46	14,16	(-*-)

```
4      18  63,46  12,28                      (--*--)
                          +---------+---------+---------+---------
                         20        40        60        80

Pooled StDev = 12,28
```

<u>Análise</u>: O PValue no cálculo da ANOVA é zero, o que significa que temos uma grande confiança no modelo. Desta forma, fez-se o reordenamento dos quatro clusters pela média, ordenando do melhor para o pior, obtendo-se a seguinte ordem: (1) level 1 obteve a melhor média e o menor intervalo, level 4 na sequencia, level 2 e level 3.

REORDENAMENTO PELA MÉDIA:

1 VIRA 4

2 VIRA 2

3 VIRA 1

4 VIRA 3

A Seguir será realizada a análise dos quatro agrupamentos com as principais variáveis identificadas para bem estar e qualidade de vida.

10.2. 4R ORD

Ordinal Logistic Regression: 4R ORD versus Boas Instala; Índice de Pr; ...

Link Function: Logit

Response Information

Variable	Value	Count
4R ORD	1	1133
	2	1753
	3	887
	4	3871
	Total	7644

Frequency: IDH - 2013_1_1

* NOTE * 129 cases were used
* NOTE * 3 cases contained missing values or was a case with zero frequency.

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower
Const(1)	1,03778	0,134413	7,72	0,000		
Const(2)	2,79855	0,136863	20,45	0,000		
Const(3)	3,61309	0,139710	25,86	0,000		
Boas Instalações Sanitárias	-0,0017292	0,0019566	-0,88	0,377	1,00	0,99
Índice de Progresso Social	-0,0948180	0,0028870	-32,84	0,000	0,91	0,90
Acesso a Eletricidade	0,0384259	0,0021100	18,21	0,000	1,04	1,03
Respeito às mulheres	-0,0395152	0,0011923	-33,14	0,000	0,96	0,96

Tolerância aos imigrantes	0,0222371	0,0014033	15,85	0,000	1,02	1,02
IDH - 2013_1_1	0,0036265	0,0028699	1,26	0,206	1,00	1,00

Predictor	Upper
Const(1)	
Const(2)	
Const(3)	
Boas Instalações Sanitárias	1,00
Índice de Progresso Social	0,91
Acesso a Eletricidade	1,04
Respeito às mulheres	0,96
Tolerância aos imigrantes	1,03
IDH - 2013_1_1	1,01

Log-Likelihood = -6959,768
Test that all slopes are zero: G = 4657,902, DF = 6, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	19698,3	378	0,000
Deviance	13919,5	378	0,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	16057454	83,8	Somers' D	0,68
Discordant	3061796	16,0	Goodman-Kruskal Gamma	0,68
Ties	32064	0,2	Kendall's Tau-a	0,44
Total	19151314	100,0		

Análises: As variáveis cujos PValue foram superiores a 10% foram desprezadas, IDH e Boas Instalações Sanitárias, sendo rodado novamente a seguir sem essas duas variáveis, conforme demonstrado a seguir.

Ordinal Logistic Regression: 4R ORD versus Taxa de não ; Respeito às ; ...

Link Function: Logit

Response Information

Variable	Value	Count
4R ORD	1	37
	2	29
	3	18
	4	48
	Total	132

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	13,7656	3,65542	3,77	0,000	
Const(2)	19,5538	4,41162	4,43	0,000	
Const(3)	21,7576	4,51958	4,81	0,000	
Taxa de não suicídio	-0,0041664	0,0137569	-0,30	0,762	1,00
Respeito às mulheres	0,0216802	0,0150434	1,44	0,150	1,02
IDH - 2013_1_1	-0,0610369	0,0375760	-1,62	0,104	0,94
Tolerância Relogiosa	-0,0199903	0,0106108	-1,88	0,060	0,98

HPI	0,0167154	0,0145412	1,15	0,250	1,02
GINNI	0,0183854	0,0178403	1,03	0,303	1,02
Tolerância aos imigrantes	0,0109439	0,0156556	0,70	0,485	1,01
Acesso a Eletricidade	-0,121155	0,0442266	-2,74	0,006	0,89
Índice de Liberdade de Imprensa	0,0092310	0,0213094	0,43	0,665	1,01
Acesso a Educação Avançada	0,0296080	0,0294662	1,00	0,315	1,03
Boas Instalações Sanitárias	0,0127015	0,0223672	0,57	0,570	1,01
Índice de Progresso Social	-0,155980	0,0468395	-3,33	0,001	0,86

	95% CI	
Predictor	Lower	Upper
Const(1)		
Const(2)		
Const(3)		
Taxa de não suícidio	0,97	1,02
Respeito às mulheres	0,99	1,05
IDH - 2013_1_1	0,87	1,01
Tolerância Relogiosa	0,96	1,00
HPI	0,99	1,05
GINNI	0,98	1,05
Tolerância aos imigrantes	0,98	1,04
Acesso a Eletricidade	0,81	0,97
Índice de Liberdade de Imprensa	0,97	1,05
Acesso a Educação Avançada	0,97	1,09
Boas Instalações Sanitárias	0,97	1,06
Índice de Progresso Social	0,78	0,94

Log-Likelihood = -55,182
Test that all slopes are zero: G = 240,497, DF = 12, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	147,464	381	1,000
Deviance	110,363	381	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	6093	96,8	Somers' D	0,94
Discordant	198	3,1	Goodman-Kruskal Gamma	0,94
Ties	2	0,0	Kendall's Tau-a	0,68
Total	6293	100,0		

Análises: Os itens que demonstraram um p-value acima de 0,10 serão descartados, pois isso significa que o modelo não se mostra confiável, tornando-se necessário roda-lo novamente a segui apenas com os itens: Tolerância religiosa, Índice de Progresso social e tolerância religiosa.

Ordinal Logistic Regression: 4R ORD versus Índice de Pr; Acesso a Ele; ...

Link Function: Logit

Response Information

Variable	Value	Count
4R ORD	1	37
	2	29
	3	18
	4	48
	Total	132

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower
Const(1)	15,6189	2,75203	5,68	0,000		
Const(2)	21,0919	3,57657	5,90	0,000		
Const(3)	23,0946	3,71612	6,21	0,000		
Índice de Progresso Social	-0,161023	0,0273977	-5,88	0,000	0,85	0,81
Acesso a Eletricidade	-0,117911	0,0300753	-3,92	0,000	0,89	0,84
Tolerância Relogiosa	-0,0172680	0,0089837	-1,92	0,055	0,98	0,97

Continuaram boas

Predictor	Upper
Const(1)	
Const(2)	
Const(3)	
Índice de Progresso Social	0,90
Acesso a Eletricidade	0,94
Tolerância Relogiosa	1,00

Log-Likelihood = -59,570
Test that all slopes are zero: G = 231,721, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	157,078	390	1,000
Deviance	119,139	390	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	6052	96,2	Somers' D	0,92
Discordant	239	3,8	Goodman-Kruskal Gamma	0,92
Ties	2	0,0	Kendall's Tau-a	0,67
Total	6293	100,0		

Análises: Nota-se que este é o melhor modelo encontrado mostrando-se absolutamente confiável, pois o p-value dos três itens: Tolerância religiosa, Índice de Progresso social e tolerância religiosa foram iguais a 0.

A seguir serão ordenadas as 3 regiões. A COLUNA do Minitab 3R OR classificando só 3 grupos de clusters, do pior para o melhor: 1-Others, 2-Aiber, 3- Aveco. Nas principais variáveis identificadas para bem-estar e qualidade de vida.

10.3. 3R ORD

Ordinal Logistic Regression: 3R ORD versus Acesso a Edu; Boas Instala; ...

Link Function: Logit

Response Information

Variable	Value	Count
3R ORD	1	3540
	2	1343

```
          3       2424
          Total   7307

Frequency: Índice de Progresso Social

* NOTE * 131 cases were used
* NOTE * 1 cases contained missing values or was a case with zero frequency.


Logistic Regression Table

                                                                          Odds
Predictor                              Coef    SE Coef       Z       P  Ratio
Const(1)                           13,1244   0,570955   22,99   0,000
Const(2)                           15,2864   0,582458   26,24   0,000
Acesso a Educação Avançada       -0,0972481  0,0037312  -26,06   0,000   0,91
Boas Instalações Sanitárias       0,0327957  0,0035166    9,33   0,000   1,03
Acesso a Eletricidade            -0,0334943  0,0061730   -5,43   0,000   0,97
Índice de Liberdade de Imprensa   0,0475433  0,0022910   20,75   0,000   1,05
Tolerância aos imigrantes        -0,0058632  0,0017191   -3,41   0,001   0,99
IDH - 2013_1_1                   -0,0480804  0,0033290  -14,44   0,000   0,95
GINNI                             0,0005078  0,0021688    0,23   0,815   1,00
HPI                              -0,0535723  0,0021273  -25,18   0,000   0,95
Respeito às mulheres             -0,0094531  0,0019818   -4,77   0,000   0,99
Tolerância Relogiosa             -0,0310230  0,0014069  -22,05   0,000   0,97
Taxa de não suícidio             -0,0093441  0,0019629   -4,76   0,000   0,99
Tirar o GINNI que tem mais de 10% de P Value
                                       95% CI
Predictor                           Lower  Upper
Const(1)
Const(2)
Acesso a Educação Avançada           0,90   0,91
Boas Instalações Sanitárias          1,03   1,04
Acesso a Eletricidade                0,96   0,98
Índice de Liberdade de Imprensa      1,04   1,05
Tolerância aos imigrantes            0,99   1,00
IDH - 2013_1_1                       0,95   0,96
GINNI                                1,00   1,00
HPI                                  0,94   0,95
Respeito às mulheres                 0,99   0,99
Tolerância Relogiosa                 0,97   0,97
Taxa de não suícidio                 0,99   0,99


Log-Likelihood = -3712,326
Test that all slopes are zero: G = 7605,505, DF = 11, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      11455,0  249  0,000
Deviance      7424,7  249  0,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs         Number  Percent  Summary Measures
Concordant  15522168     93,6  Somers' D              0,87
Discordant   1049333      6,3  Goodman-Kruskal Gamma  0,87
Ties           19111      0,1  Kendall's Tau-a        0,54
Total       16590612    100,0
```

Análises: O único item que demonstraram um p-value acima de 0,10 é o Índice de Ginni que será descartado, pois isso significa que o modelo não se mostra confiável, tornando-se necessário roda-lo novamente a segui apenas com os demais itens.

Ordinal Logistic Regression: 3R ORD versus Acesso a Edu; Boas Instala; ...

Link Function: Logit

Response Information

Variable	Value	Count
3R ORD	1	3540
	2	1343
	3	2424
	Total	7307

Frequency: Índice de Progresso Social

* NOTE * 131 cases were used
* NOTE * 1 cases contained missing values or was a case with zero frequency.

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	13,1706	0,531600	24,78	0,000	
Const(2)	15,3325	0,544003	28,18	0,000	
Acesso a Educação Avançada	-0,0973268	0,0037268	-26,12	0,000	0,91
Boas Instalações Sanitárias	0,0327993	0,0035144	9,33	0,000	1,03
Acesso a Eletricidade	-0,0339980	0,0057916	-5,87	0,000	0,97
Índice de Liberdade de Imprensa	0,0476923	0,0022008	21,67	0,000	1,05
Tolerância aos imigrantes	-0,0057692	0,0016802	-3,43	0,001	0,99
IDH - 2013_1_1	-0,0479292	0,0032804	-14,61	0,000	0,95
HPI	-0,0535190	0,0021188	-25,26	0,000	0,95
Respeito às mulheres	-0,0096170	0,0018604	-5,17	0,000	0,99
Tolerância Relogiosa	-0,0310276	0,0014074	-22,05	0,000	0,97
Taxa de não suícidio	-0,0092130	0,0018834	-4,89	0,000	0,99

Predictor	95% CI Lower	Upper
Const(1)		
Const(2)		
Acesso a Educação Avançada	0,90	0,91
Boas Instalações Sanitárias	1,03	1,04
Acesso a Eletricidade	0,96	0,98
Índice de Liberdade de Imprensa	1,04	1,05
Tolerância aos imigrantes	0,99	1,00
IDH - 2013_1_1	0,95	0,96
HPI	0,94	0,95
Respeito às mulheres	0,99	0,99
Tolerância Relogiosa	0,97	0,97
Taxa de não suícidio	0,99	0,99

Log-Likelihood = -3712,353
Test that all slopes are zero: G = 7605,450, DF = 10, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	11449,5	250	0,000
Deviance	7424,7	250	0,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	15525585	93,6	Somers' D	0,87
Discordant	1049333	6,3	Goodman-Kruskal Gamma	0,87

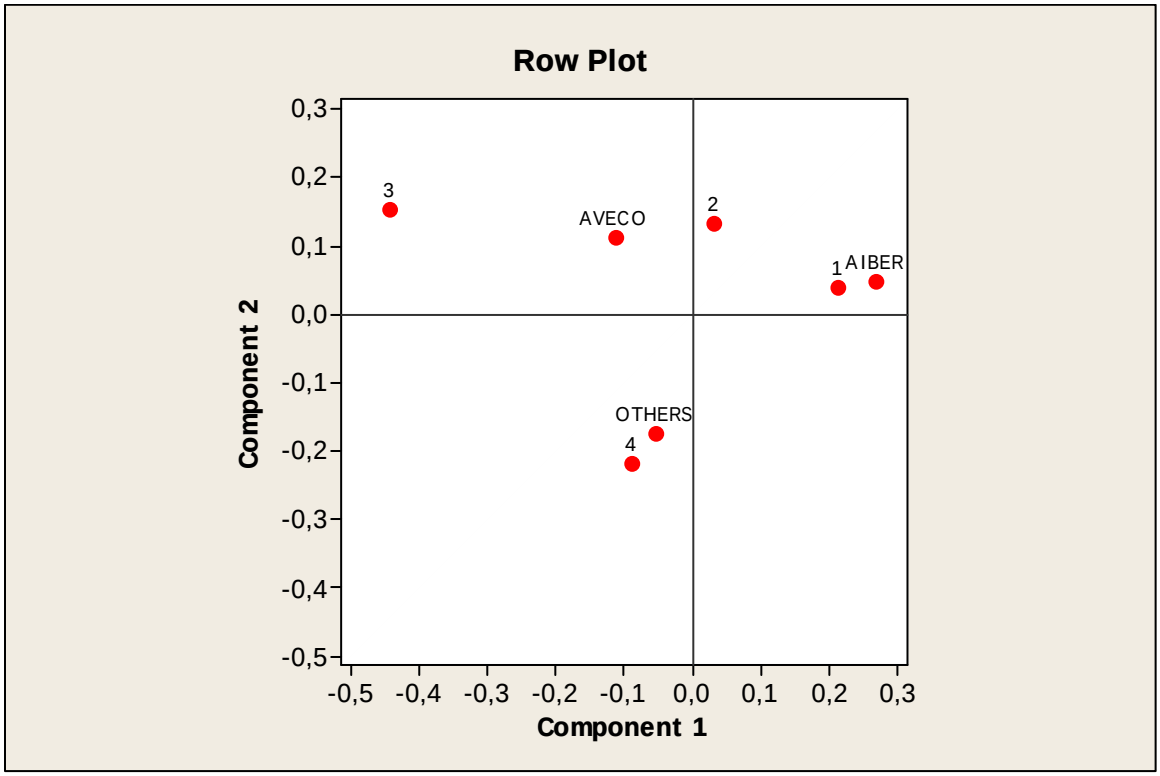
Ties	15694	0,1	Kendall's Tau-a	0,54
Total	16590612	100,0		

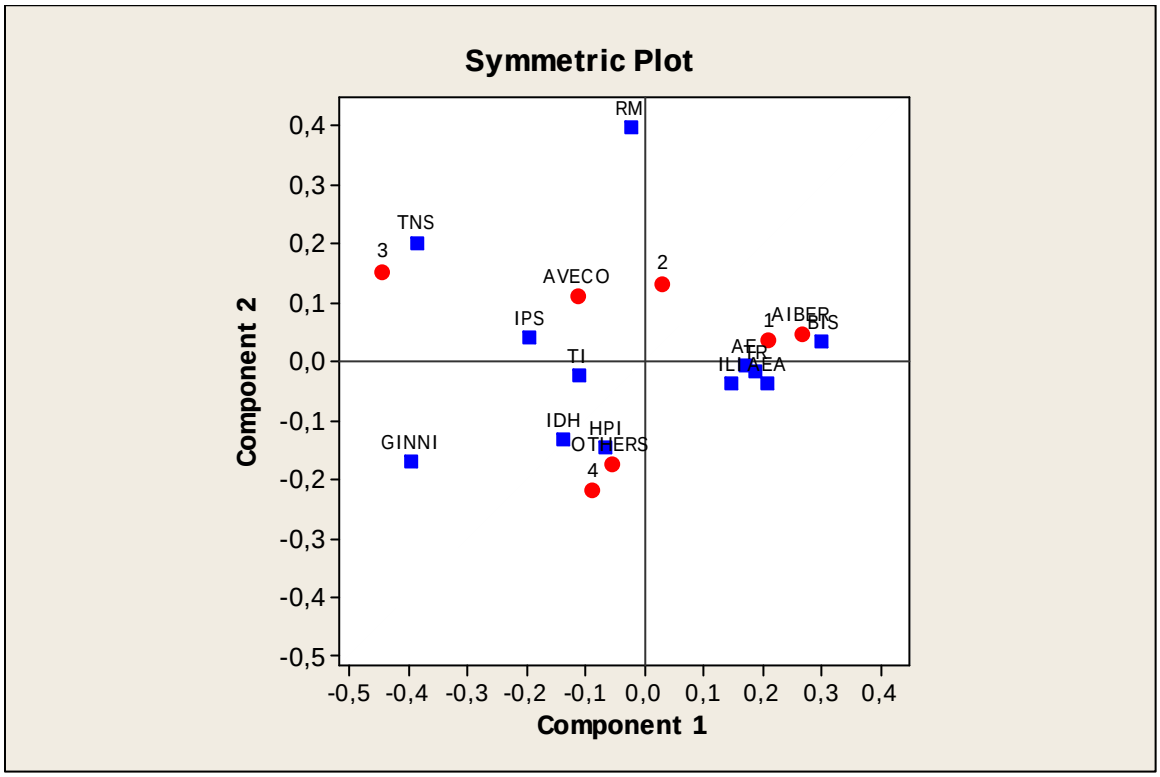
<u>Análises</u>: Nota-se que este é o melhor modelo encontrado mostrando-se absolutamente confiável, pois o p-value dos dez itens foram iguais a 0.

<u>Conclusões:</u>

Nota-se que a regressão das três regiões é mais confiável e necessita a utilização de mais itens.

12. ANÁLISE DE CORRESPONDENCIA





Simple Correspondence Analysis: TNS; IPS; AEA M; BIS; AE; ILI; RM; TI; IDH; TR;

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI
OTHERS	76,410	64,410	41,460	80,630	92,810	36,360	26,570	61,970
AIBER	63,170	90,300	72,810	96,750	99,730	12,750	70,330	62,860
AVECO	70,190	42,630	31,320	60,270	66,420	44,320	56,160	45,710
1	64,750	81,260	62,820	95,220	99,500	21,480	61,960	56,990
2	65,710	49,490	43,740	83,380	95,910	55,620	58,440	33,300

3	74,590	25,030	13,110	27,310	28,070	43,170	57,670	54,790
4	79,600	61,650	38,780	76,080	90,280	32,760	20,660	61,640
Total	494,420	414,770	304,040	519,640	572,720	246,460	351,790	377,260

	TR	IDH	HPI	GINNI	Total
OTHERS	65,620	84,130	72,360	56,920	759,650
AIBER	86,970	64,200	50,580	20,090	790,540
AVECO	47,390	51,590	41,790	37,560	595,350
1	81,120	63,890	46,360	23,860	759,210
2	60,880	36,780	55,810	29,790	668,850
3	24,460	58,560	32,910	47,010	486,680
4	63,460	85,190	74,140	63,570	747,810
Total	429,900	444,340	373,950	278,800	4808,090

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR
OTHERS	0,037	0,019	0,900	0,026	0,060	0,171	15,142	0,094	0,078
AIBER	4,040	7,164	10,417	1,498	0,329	19,034	2,697	0,011	3,753
AVECO	1,314	1,483	1,063	0,258	0,285	6,243	3,645	0,022	0,641
1	2,273	3,796	4,569	2,113	0,909	7,813	0,740	0,112	2,581
2	0,137	1,168	0,049	1,702	3,310	13,277	1,845	7,010	0,019
3	12,037	6,846	10,140	12,158	15,423	13,311	13,668	7,219	8,344
4	0,095	0,127	1,531	0,278	0,016	0,810	21,196	0,150	0,173
Total	19,933	20,603	28,671	18,034	20,332	60,659	58,933	14,617	15,590

	IDH	HPI	GINNI	Total
OTHERS	2,763	2,984	3,761	26,036
AIBER	1,074	1,934	14,465	66,415
AVECO	0,214	0,440	0,267	15,875
1	0,561	2,726	9,235	37,428
2	10,137	0,276	2,086	41,017
3	4,102	0,645	12,510	116,405
4	3,742	4,390	9,417	41,925
Total	22,593	13,395	51,741	345,100

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
OTHERS	0,000	0,000	0,003	0,000	0,000	0,000	0,044	0,000	0,000	0,008
AIBER	0,012	0,021	0,030	0,004	0,001	0,055	0,008	0,000	0,011	0,003
AVECO	0,004	0,004	0,003	0,001	0,001	0,018	0,011	0,000	0,002	0,001
1	0,007	0,011	0,013	0,006	0,003	0,023	0,002	0,000	0,007	0,002
2	0,000	0,003	0,000	0,005	0,010	0,038	0,005	0,020	0,000	0,029
3	0,035	0,020	0,029	0,035	0,045	0,039	0,040	0,021	0,024	0,012
4	0,000	0,000	0,004	0,001	0,000	0,002	0,061	0,000	0,001	0,011
Total	0,058	0,060	0,083	0,052	0,059	0,176	0,171	0,042	0,045	0,065

	HPI	GINNI	Total
OTHERS	0,009	0,011	0,075
AIBER	0,006	0,042	0,192
AVECO	0,001	0,001	0,046
1	0,008	0,027	0,108
2	0,001	0,006	0,119
3	0,002	0,036	0,337
4	0,013	0,027	0,121
Total	0,039	0,150	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0422	0,5874	0,5874	******************************
2	0,0192	0,2682	0,8556	*************
3	0,0101	0,1408	0,9964	*******
4	0,0002	0,0027	0,9992	
5	0,0001	0,0008	1,0000	
Total	0,0718			

Row Contributions

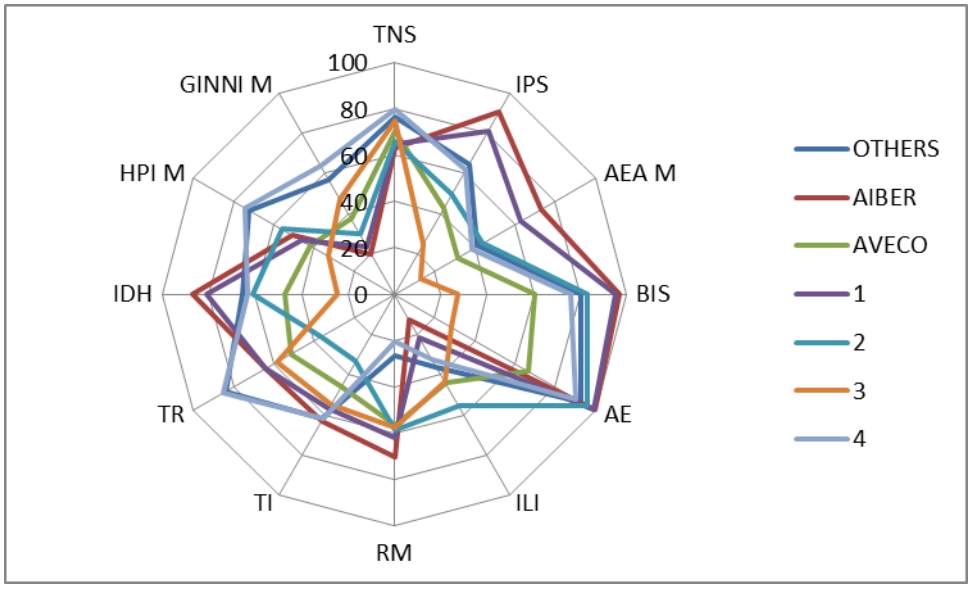
ID	Name	Qual	Mass	Inert	Component 1 Coord	Component 1 Corr	Component 1 Contr	Component 2 Coord	Component 2 Corr	Component 2 Contr
1	OTHERS	0,979	0,158	0,075	-0,056	0,090	0,012	-0,175	0,889	0,250
2	AIBER	0,874	0,164	0,192	0,267	0,847	0,278	0,047	0,027	0,019
3	AVECO	0,947	0,124	0,046	-0,114	0,485	0,038	0,111	0,462	0,079
4	1	0,937	0,158	0,108	0,212	0,908	0,168	0,038	0,029	0,012
5	2	0,306	0,139	0,119	0,031	0,015	0,003	0,134	0,291	0,129
6	3	0,922	0,101	0,337	-0,444	0,824	0,473	0,153	0,098	0,124
7	4	0,994	0,156	0,121	-0,088	0,139	0,029	-0,219	0,854	0,387

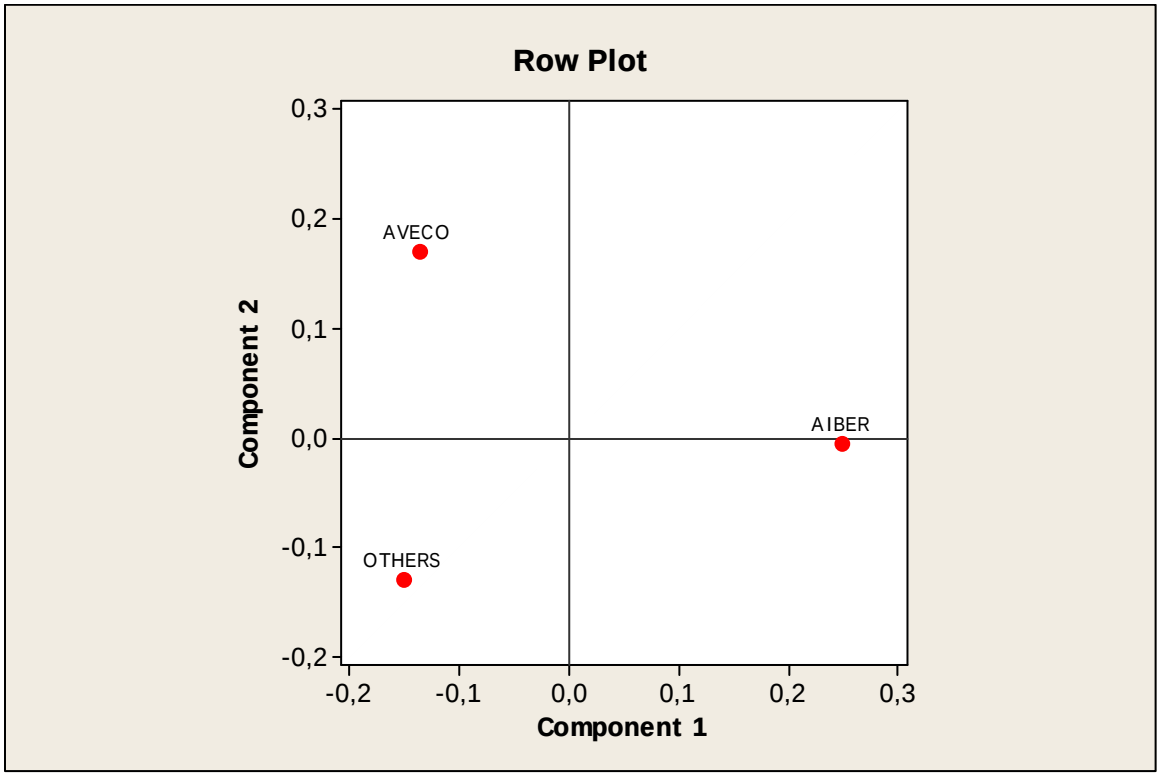
Column Contributions

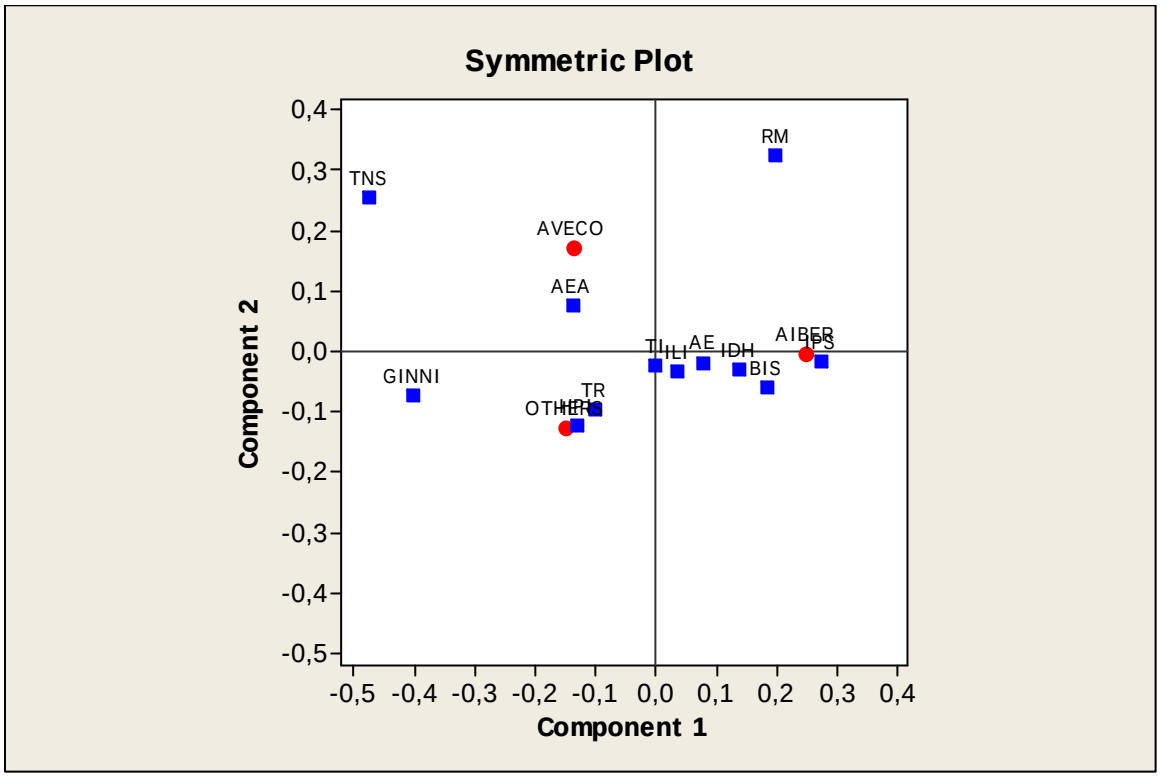
ID	Name	Qual	Mass	Inert	Component 1 Coord	Component 1 Corr	Component 1 Contr	Component 2 Coord	Component 2 Corr	Component 2 Contr
1	IPS	0,983	0,103	0,058	-0,195	0,944	0,093	0,039	0,038	0,008
2	AEA	0,910	0,086	0,060	0,209	0,881	0,090	-0,038	0,029	0,007
3	BIS	0,989	0,063	0,083	0,303	0,977	0,138	0,034	0,012	0,004
4	AE	0,852	0,108	0,052	0,172	0,851	0,076	-0,007	0,002	0,000
5	ILI	0,660	0,119	0,059	0,148	0,620	0,062	-0,037	0,040	0,009
6	TNS	0,765	0,051	0,176	-0,385	0,601	0,180	0,201	0,164	0,108
7	RM	0,946	0,073	0,171	-0,022	0,003	0,001	0,398	0,943	0,601
8	TI	0,330	0,078	0,042	-0,111	0,315	0,023	-0,024	0,015	0,002
9	TR	0,999	0,089	0,045	0,190	0,991	0,076	-0,017	0,008	0,001
10	IDH	0,709	0,092	0,065	-0,135	0,357	0,040	-0,134	0,352	0,086
11	HPI	0,708	0,078	0,039	-0,064	0,115	0,008	-0,146	0,593	0,086
12	GINNI	0,998	0,058	0,150	-0,395	0,840	0,214	-0,172	0,158	0,089

Row Plot

Symmetric Plot







Simple Correspondence Analysis: AEA_1; TNS_1; IPS_1; BIS_1; AE_1; ILI_1; RM_1;

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI
OTHERS	41,460	76,410	64,410	80,630	92,810	36,360	26,570	61,970

AIBER	72,810	63,170	90,300	96,750	99,730	12,750	70,330	62,860
AVECO	31,320	70,190	42,630	60,270	66,420	44,320	56,160	45,710
Total	145,590	209,770	197,340	237,650	258,960	93,430	153,060	170,540

	TR	IDH	HPI	GINNI	Total
OTHERS	84,130	65,620	72,360	56,920	759,650
AIBER	64,200	86,970	50,580	20,090	790,540
AVECO	51,590	47,390	41,790	37,560	595,350
Total	199,920	199,980	164,730	114,570	2145,540

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
OTHERS	1,974	0,062	0,427	0,147	0,014	0,325	14,079	0,042	2,516	0,380
AIBER	6,848	2,580	4,255	0,964	0,195	13,647	3,443	0,000	1,215	2,396
AVECO	2,040	2,467	2,686	0,488	0,411	13,052	4,412	0,055	0,272	1,183
Total	10,862	5,108	7,368	1,598	0,620	27,024	21,934	0,097	4,004	3,958

	HPI	GINNI	Total
OTHERS	3,378	6,594	29,937
AIBER	1,686	11,595	48,823
AVECO	0,336	1,047	28,449
Total	5,400	19,236	107,209

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
OTHERS	0,018	0,001	0,004	0,001	0,000	0,003	0,131	0,000	0,023	0,004
AIBER	0,064	0,024	0,040	0,009	0,002	0,127	0,032	0,000	0,011	0,022
AVECO	0,019	0,023	0,025	0,005	0,004	0,122	0,041	0,001	0,003	0,011
Total	0,101	0,048	0,069	0,015	0,006	0,252	0,205	0,001	0,037	0,037

	HPI	GINNI	Total
OTHERS	0,032	0,062	0,279
AIBER	0,016	0,108	0,455
AVECO	0,003	0,010	0,265
Total	0,050	0,179	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0361	0,7215	0,7215	******************************
2	0,0139	0,2785	1,0000	***********
Total	0,0500			

Row Contributions

						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	OTHERS	1,000	0,354	0,279	-0,151	0,579	0,224	-0,129	0,421	0,422
2	AIBER	1,000	0,368	0,455	0,248	1,000	0,631	-0,004	0,000	0,001
3	AVECO	1,000	0,277	0,265	-0,137	0,394	0,145	0,170	0,606	0,578

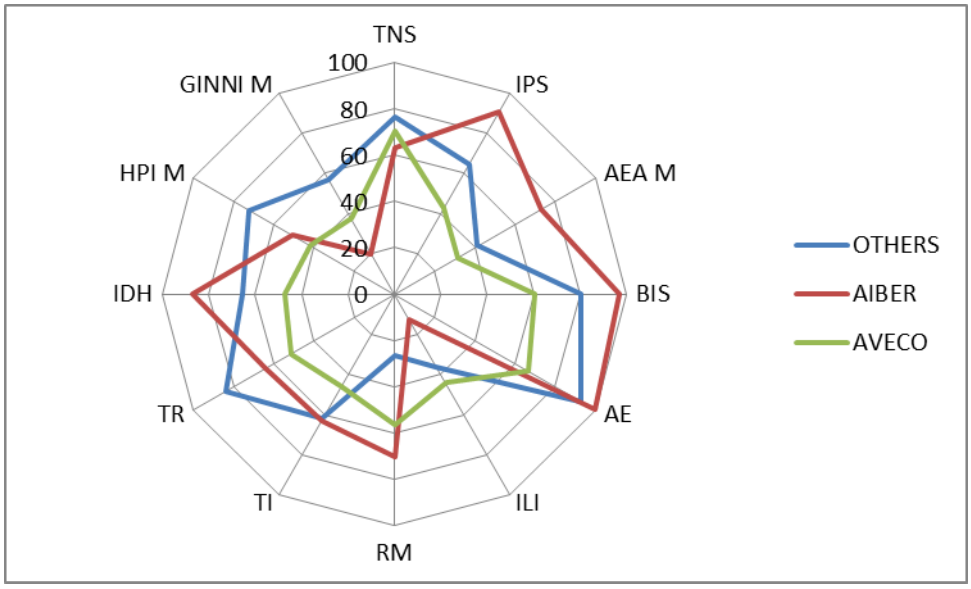
Column Contributions

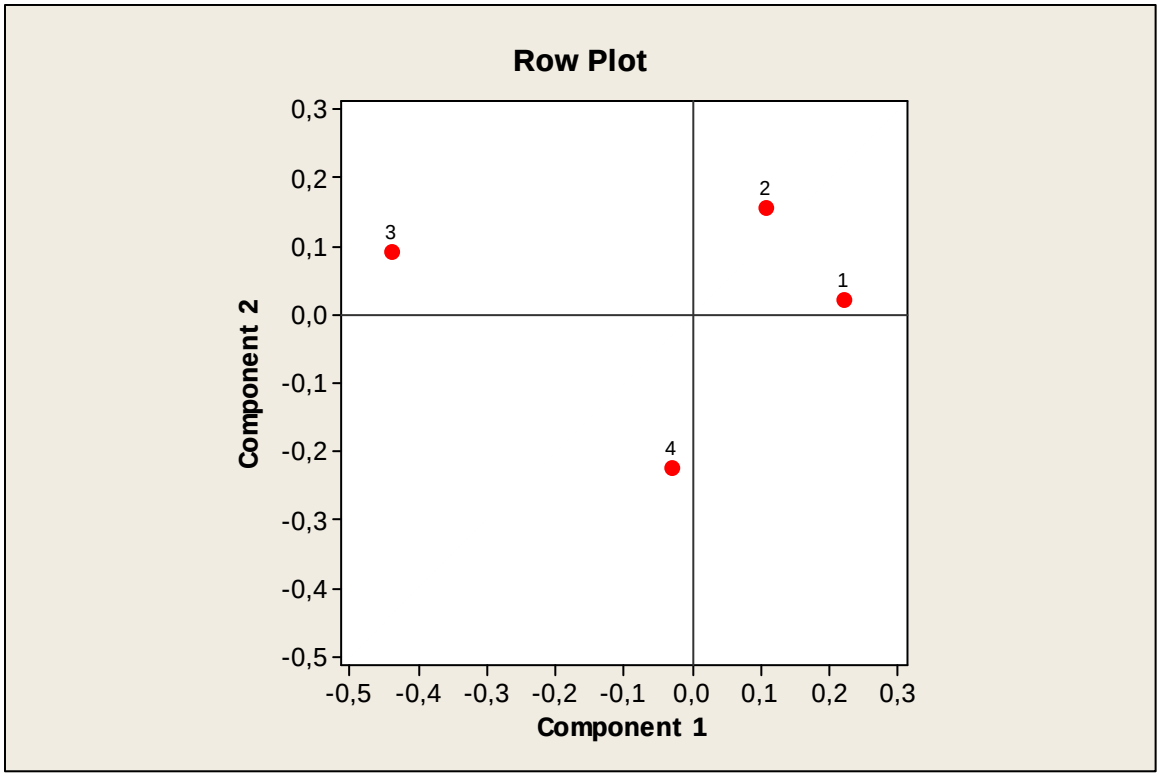
						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IPS	1,000	0,068	0,101	0,272	0,995	0,140	-0,019	0,005	0,002
2	AEA	1,000	0,098	0,048	-0,137	0,776	0,051	0,074	0,224	0,038
3	BIS	1,000	0,092	0,069	0,183	0,897	0,085	-0,062	0,103	0,025
4	AE	1,000	0,111	0,015	0,080	0,942	0,019	-0,020	0,058	0,003
5	ILI	1,000	0,121	0,006	0,034	0,469	0,004	-0,036	0,531	0,011
6	TNS	1,000	0,044	0,252	-0,474	0,776	0,271	0,255	0,224	0,203
7	RM	1,000	0,071	0,205	0,198	0,274	0,078	0,323	0,726	0,533
8	TI	1,000	0,079	0,001	-0,000	0,000	0,000	-0,024	1,000	0,003

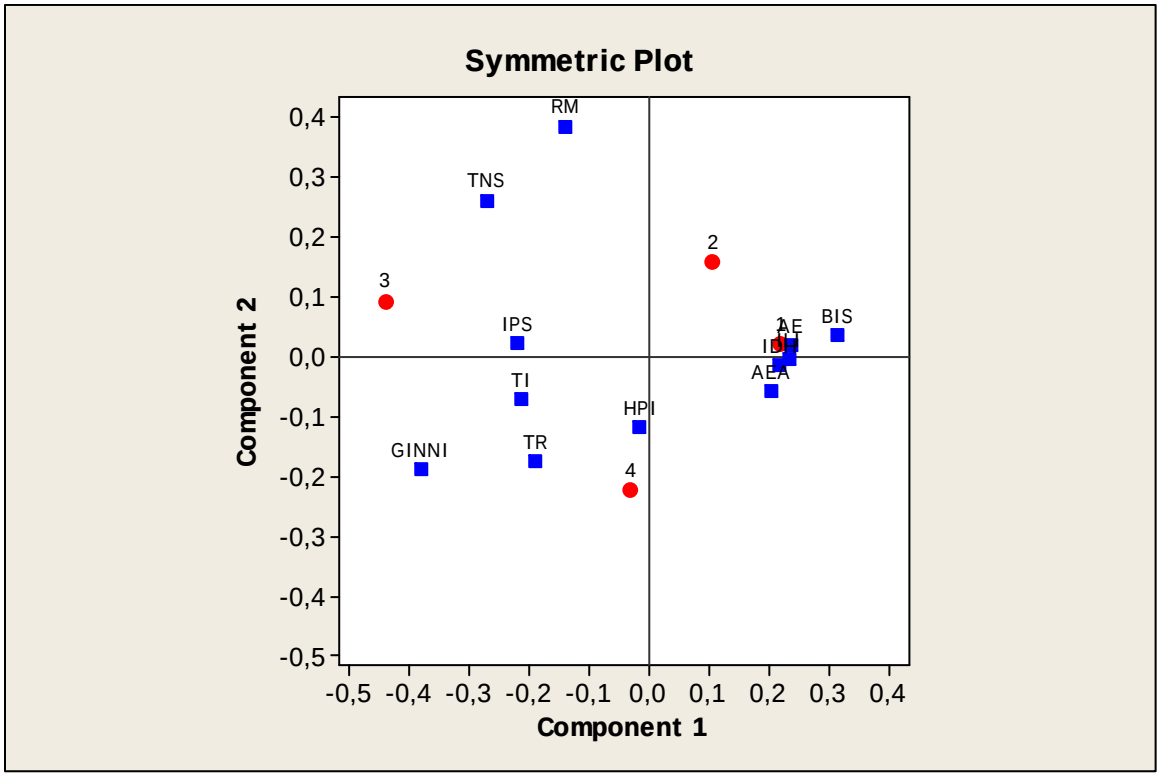
9	TR	1,000	0,093	0,037	-0,101	0,510	0,026	-0,099	0,490	0,066
10	IDH	1,000	0,093	0,037	0,137	0,946	0,048	-0,033	0,054	0,007
11	HPI	1,000	0,077	0,050	-0,131	0,523	0,037	-0,125	0,477	0,086
12	GINNI	1,000	0,053	0,179	-0,403	0,966	0,240	-0,076	0,034	0,022

Row Plot

Symmetric Plot







Simple Correspondence Analysis: TNS_2; IPS_2; AEA M_2; BIS_2; AE_2; ILI_2; RM_2

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI
1	64,750	81,260	62,820	95,220	99,500	21,480	61,960	56,990
2	65,710	49,490	43,740	83,380	95,910	55,620	58,440	33,300
3	74,590	25,030	13,110	27,310	28,070	43,170	57,670	54,790
4	79,600	61,650	38,780	76,080	90,280	32,760	20,660	61,640
Total	284,650	217,430	158,450	281,990	313,760	153,030	198,730	206,720

	TR	IDH	HPI	GINNI	Total
1	63,890	81,120	46,360	23,860	759,210
2	36,780	60,880	55,810	29,790	668,850
3	58,560	24,460	32,910	47,010	486,680
4	85,190	63,460	74,140	63,570	747,810
Total	244,420	229,920	209,220	164,230	2662,550

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR
1	3,320	5,984	6,886	2,729	1,125	11,249	0,494	0,065	0,483
2	0,470	0,482	0,389	2,221	3,706	7,676	1,453	6,683	9,872
3	9,782	5,447	8,677	11,394	14,950	8,258	12,542	7,652	4,314
4	0,002	0,006	0,736	0,123	0,053	2,430	22,143	0,221	3,986
Total	13,573	11,918	16,688	16,466	19,834	29,613	36,633	14,621	18,656

	IDH	HPI	GINNI	Total
1	3,693	2,964	11,266	50,259
2	0,169	0,201	3,187	36,509
3	7,343	0,744	9,617	100,718
4	0,019	4,024	6,597	40,339
Total	11,223	7,933	30,666	227,826

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
1	0,015	0,026	0,030	0,012	0,005	0,049	0,002	0,000	0,002	0,016
2	0,002	0,002	0,002	0,010	0,016	0,034	0,006	0,029	0,043	0,001
3	0,043	0,024	0,038	0,050	0,066	0,036	0,055	0,034	0,019	0,032
4	0,000	0,000	0,003	0,001	0,000	0,011	0,097	0,001	0,017	0,000
Total	0,060	0,052	0,073	0,072	0,087	0,130	0,161	0,064	0,082	0,049

	HPI	GINNI	Total
1	0,013	0,049	0,221
2	0,001	0,014	0,160
3	0,003	0,042	0,442
4	0,018	0,029	0,177
Total	0,035	0,135	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0523	0,6108	0,6108	******************************
2	0,0220	0,2568	0,8677	************
3	0,0113	0,1323	1,0000	******
Total	0,0856			

Row Contributions

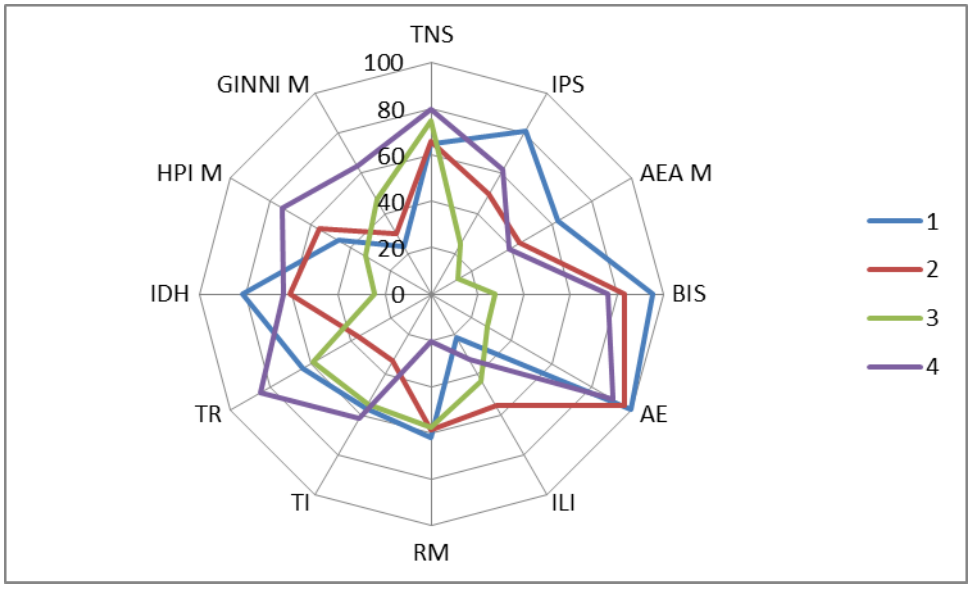
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	1	0,732	0,285	0,221	0,219	0,725	0,262	0,022	0,007	0,006
2	2	0,660	0,251	0,160	0,106	0,206	0,054	0,157	0,454	0,283
3	3	0,980	0,183	0,442	-0,441	0,938	0,679	0,093	0,042	0,072
4	4	0,944	0,281	0,177	-0,030	0,017	0,005	-0,224	0,926	0,639

Column Contributions

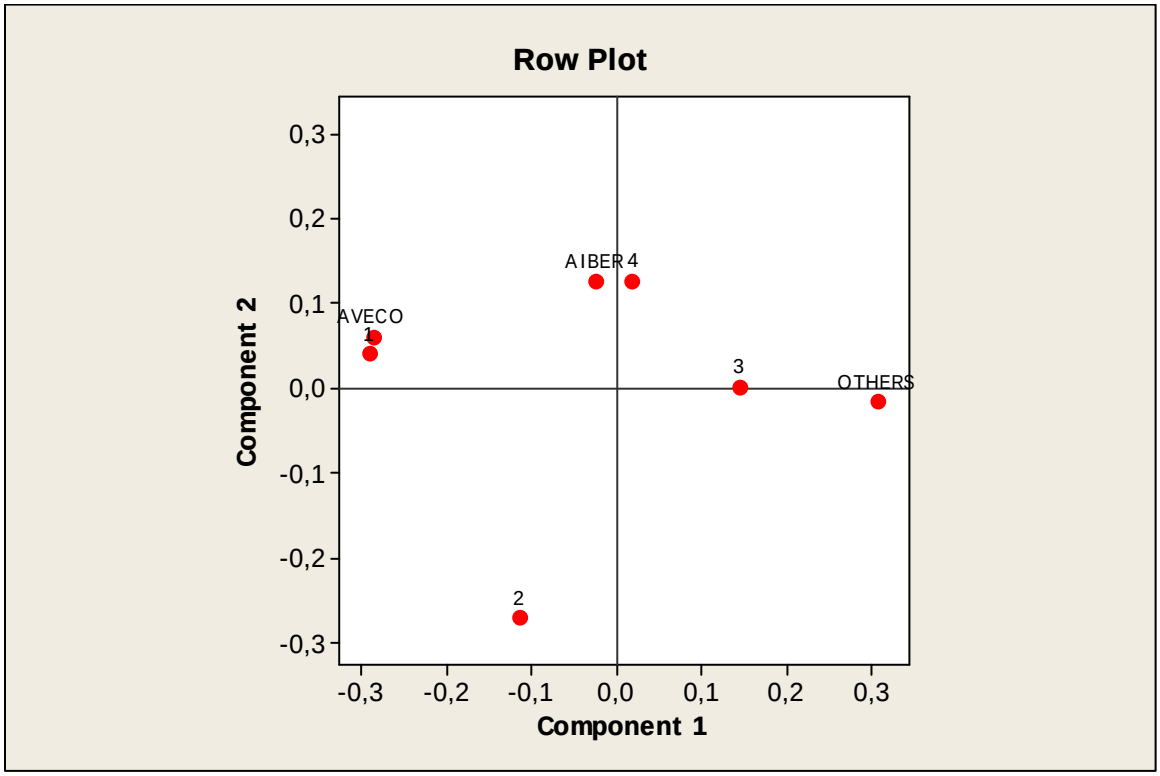
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IPS	1,000	0,107	0,060	-0,217	0,990	0,097	0,021	0,010	0,002
2	AEA	0,821	0,082	0,052	0,204	0,759	0,065	-0,058	0,062	0,013
3	BIS	0,959	0,060	0,073	0,316	0,947	0,114	0,035	0,011	0,003
4	AE	0,976	0,106	0,072	0,238	0,971	0,115	0,018	0,005	0,002
5	ILI	0,873	0,118	0,087	0,235	0,873	0,124	-0,006	0,001	0,000
6	TNS	0,725	0,057	0,130	-0,269	0,374	0,080	0,261	0,351	0,178
7	RM	0,901	0,075	0,161	-0,138	0,103	0,027	0,384	0,798	0,500
8	TI	0,707	0,078	0,064	-0,212	0,634	0,067	-0,072	0,073	0,018
9	TR	0,873	0,092	0,082	-0,188	0,463	0,062	-0,177	0,410	0,131
10	IDH	0,988	0,086	0,049	0,219	0,983	0,079	-0,016	0,005	0,001
11	HPI	0,383	0,079	0,035	-0,014	0,005	0,000	-0,120	0,378	0,051
12	GINNI	0,967	0,062	0,135	-0,380	0,773	0,170	-0,190	0,194	0,102

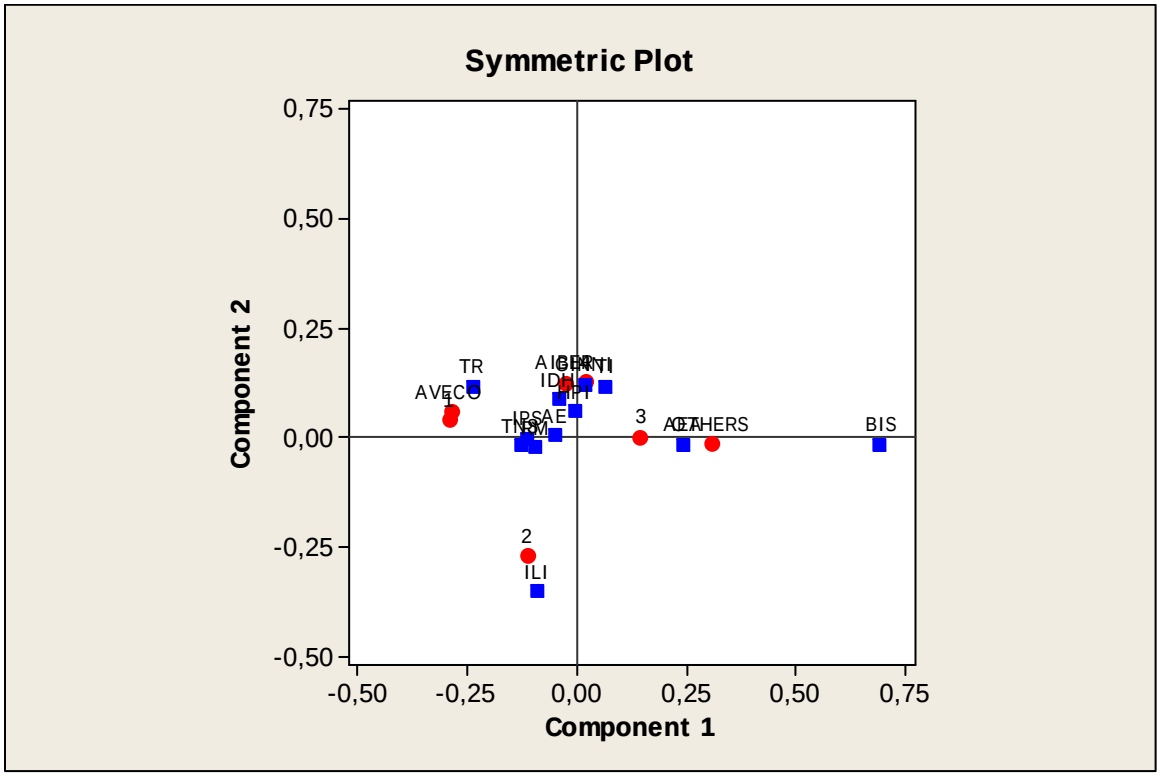
Row Plot

Symmetric Plot



DADOS GERADOS A PARTIR DO DESVIO PADRÃO DO ANOVA





Simple Correspondence Analysis: AEA DP; BIS DP; AE DP; ILI DP; TNS DP; RM DP; T

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI
AIBER	12,690	16,670	8,630	18,290	8,720	17,900	22,650	12,030
AVECO	11,140	8,710	1,400	11,500	13,860	13,590	27,620	11,700
OTHERS	19,600	34,100	37,630	20,680	25,350	20,310	33,300	23,940
1	15,300	8,210	1,620	17,660	17,390	19,590	26,480	12,110
2	14,870	17,600	7,510	18,080	32,880	21,120	31,300	9,700
3	8,390	19,680	22,520	21,540	18,210	17,420	35,490	14,160
4	10,520	15,500	8,590	11,270	8,610	13,860	20,520	12,280
Total	92,510	120,470	87,900	119,020	125,020	123,790	197,360	95,920

	TR	IDH	HPI	GINNI	Total
AIBER	16,250	16,250	14,400	11,950	176,430
AVECO	24,270	12,840	12,420	8,700	157,750
OTHERS	20,910	20,680	20,320	19,710	296,530
1	25,000	15,780	13,530	13,580	186,250
2	15,080	16,260	14,440	8,350	207,190
3	19,360	18,960	20,060	12,250	228,040
4	13,210	19,170	14,040	9,120	156,690
Total	134,080	119,940	109,210	83,660	1408,880

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
AIBER	0,105	0,166	0,514	0,769	3,073	0,371	0,173	0,000	0,017	0,101
AVECO	0,059	1,693	7,241	0,250	0,001	0,005	1,380	0,086	5,708	0,026
OTHERS	0,001	3,016	19,780	0,762	0,035	1,266	1,634	0,697	1,894	0,825
1	0,771	3,738	8,606	0,236	0,045	0,636	0,006	0,026	2,986	0,000
2	0,118	0,001	2,270	0,019	11,427	0,467	0,179	1,376	1,091	0,108
3	2,895	0,002	4,833	0,269	0,203	0,342	0,394	0,120	0,253	0,011
4	0,005	0,330	0,144	0,292	2,016	0,001	0,093	0,244	0,194	2,549
Total	3,954	8,945	43,387	2,598	16,800	3,088	3,857	2,549	12,143	3,619

	HPI	GINNI	Total
AIBER	0,038	0,207	5,534

AVECO	0,003	0,048	16,501
OTHERS	0,309	0,251	30,471
1	0,057	0,574	17,681
2	0,163	1,270	18,488
3	0,321	0,123	9,764
4	0,295	0,004	6,166
Total	1,188	2,477	104,605

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
AIBER	0,001	0,002	0,005	0,007	0,029	0,004	0,002	0,000	0,000	0,001
AVECO	0,001	0,016	0,069	0,002	0,000	0,000	0,013	0,001	0,055	0,000
OTHERS	0,000	0,029	0,189	0,007	0,000	0,012	0,016	0,007	0,018	0,008
1	0,007	0,036	0,082	0,002	0,000	0,006	0,000	0,000	0,029	0,000
2	0,001	0,000	0,022	0,000	0,109	0,004	0,002	0,013	0,010	0,001
3	0,028	0,000	0,046	0,003	0,002	0,003	0,004	0,001	0,002	0,000
4	0,000	0,003	0,001	0,003	0,019	0,000	0,001	0,002	0,002	0,024
Total	0,038	0,086	0,415	0,025	0,161	0,030	0,037	0,024	0,116	0,035

	HPI	GINNI	Total
AIBER	0,000	0,002	0,053
AVECO	0,000	0,000	0,158
OTHERS	0,003	0,002	0,291
1	0,001	0,005	0,169
2	0,002	0,012	0,177
3	0,003	0,001	0,093
4	0,003	0,000	0,059
Total	0,011	0,024	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0456	0,6146	0,6146	******************************
2	0,0152	0,2043	0,8189	*********
3	0,0056	0,0749	0,8938	***
4	0,0047	0,0639	0,9577	***
5	0,0025	0,0342	0,9920	*
6	0,0006	0,0080	1,0000	
Total	0,0742			

Row Contributions

						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	0,526	0,125	0,053	-0,025	0,020	0,002	0,126	0,506	0,131
2	AVECO	0,820	0,112	0,158	-0,286	0,784	0,201	0,061	0,036	0,028
3	OTHERS	0,922	0,210	0,291	0,307	0,920	0,436	-0,014	0,002	0,003
4	1	0,903	0,132	0,169	-0,290	0,886	0,244	0,041	0,017	0,014
5	2	0,960	0,147	0,177	-0,114	0,145	0,042	-0,270	0,814	0,704
6	3	0,492	0,162	0,093	0,145	0,492	0,075	0,003	0,000	0,000
7	4	0,423	0,111	0,059	0,019	0,009	0,001	0,128	0,414	0,119

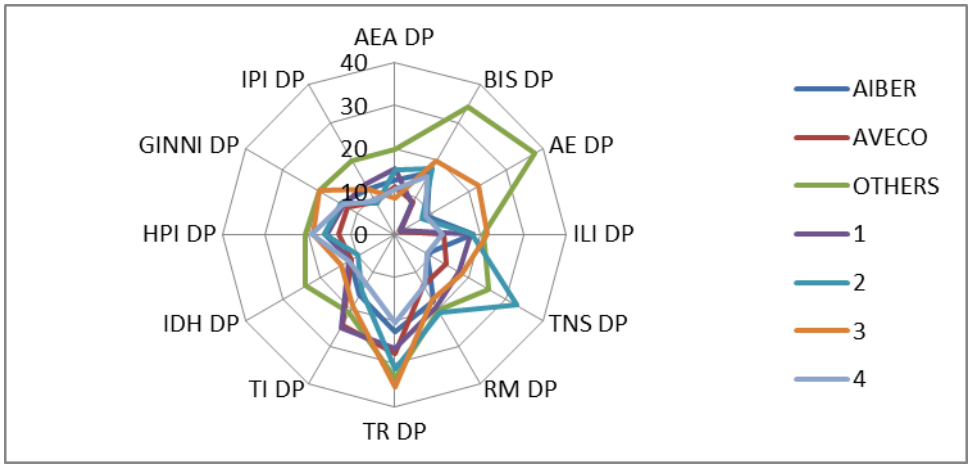
Column Contributions

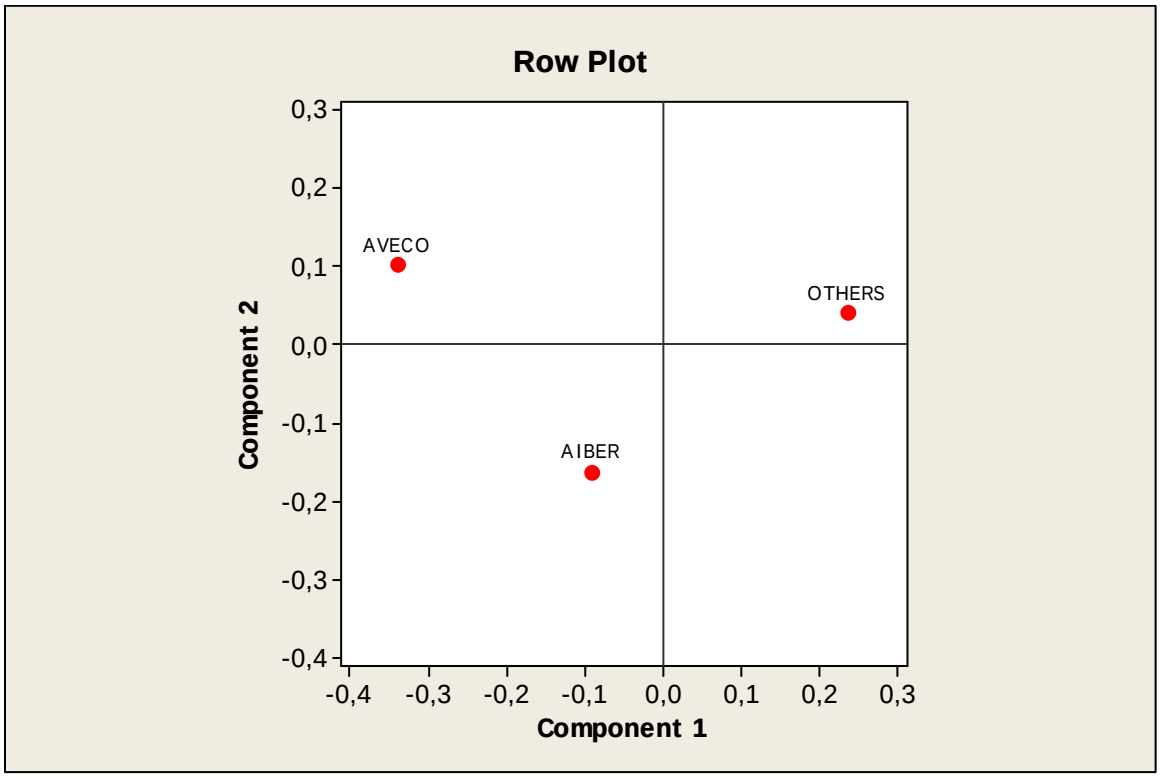
						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IPS	0,290	0,066	0,038	-0,111	0,289	0,018	-0,002	0,000	0,000
2	AEA	0,819	0,086	0,086	0,246	0,816	0,114	-0,016	0,003	0,001
3	BIS	0,980	0,062	0,415	0,695	0,980	0,661	-0,015	0,000	0,001
4	AE	0,109	0,084	0,025	-0,049	0,108	0,004	0,004	0,001	0,000
5	ILI	0,982	0,089	0,161	-0,089	0,059	0,015	-0,352	0,923	0,725
6	TNS	0,674	0,088	0,030	-0,128	0,661	0,032	-0,018	0,014	0,002
7	RM	0,477	0,140	0,037	-0,094	0,448	0,027	-0,024	0,029	0,005
8	TI	0,679	0,068	0,024	0,067	0,170	0,007	0,116	0,509	0,061
9	TR	0,779	0,095	0,116	-0,239	0,629	0,119	0,117	0,150	0,085

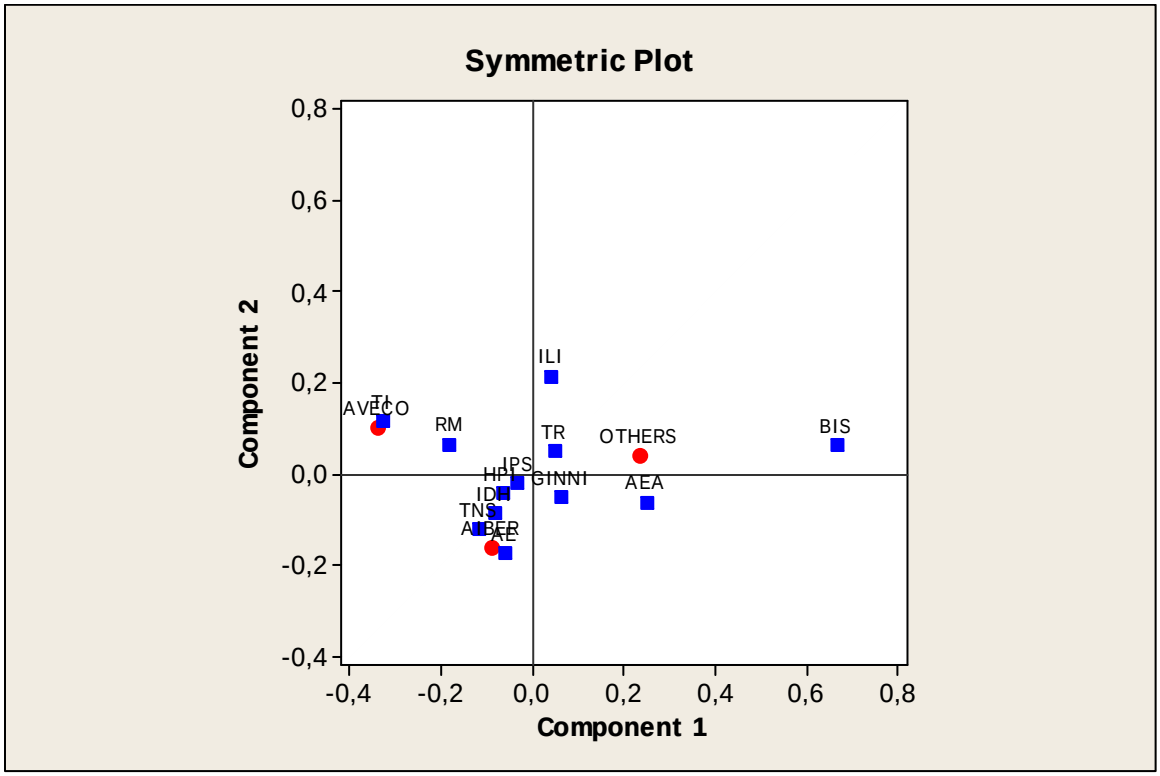
10	IDH	0,310	0,085	0,035	-0,041	0,055	0,003	0,088	0,255	0,043
11	HPI	0,318	0,078	0,011	-0,003	0,001	0,000	0,059	0,317	0,018
12	GINNI	0,513	0,059	0,024	0,018	0,011	0,000	0,122	0,502	0,058

Row Plot

Symmetric Plot







Simple Correspondence Analysis: AEA DP_1; BIS DP_1; AE DP_1; ILI DP_1; TNS DP_1

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR
AIBER	12,690	16,670	8,630	18,290	8,720	17,900	22,650	16,250	12,030
AVECO	11,140	8,710	1,400	11,500	13,860	13,590	27,620	24,270	11,700
OTHERS	19,600	34,100	37,630	20,680	25,350	20,310	33,300	20,910	23,940
Total	43,430	59,480	47,660	50,470	47,930	51,800	83,570	61,430	47,670

	IDH	HPI	GINNI	Total
AIBER	16,250	14,400	11,950	176,430
AVECO	12,840	12,420	8,700	157,750
OTHERS	20,680	20,320	19,710	296,530
Total	49,770	47,140	40,360	630,710

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
AIBER	0,024	0,000	1,658	1,233	1,639	0,802	0,023	0,051	0,128	0,389
AVECO	0,007	2,556	9,285	0,100	0,292	0,031	2,159	5,162	0,004	0,012
OTHERS	0,033	1,346	10,341	0,392	0,352	0,671	0,913	2,200	0,104	0,316
Total	0,064	3,902	21,285	1,724	2,283	1,505	3,095	7,413	0,236	0,718

	HPI	GINNI	Total
AIBER	0,112	0,039	6,097
AVECO	0,034	0,193	19,835
OTHERS	0,153	0,028	16,851
Total	0,299	0,260	42,783

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
AIBER	0,001	0,000	0,039	0,029	0,038	0,019	0,001	0,001	0,003	0,009
AVECO	0,000	0,060	0,217	0,002	0,007	0,001	0,050	0,121	0,000	0,000
OTHERS	0,001	0,031	0,242	0,009	0,008	0,016	0,021	0,051	0,002	0,007

Total	0,001	0,091	0,498	0,040	0,053	0,035	0,072	0,173	0,006	0,017

	HPI	GINNI	Total
AIBER	0,003	0,001	0,143
AVECO	0,001	0,005	0,464
OTHERS	0,004	0,001	0,394
Total	0,007	0,006	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0570	0,8406	0,8406	******************************
2	0,0108	0,1594	1,0000	*****
Total	0,0678			

Row Contributions

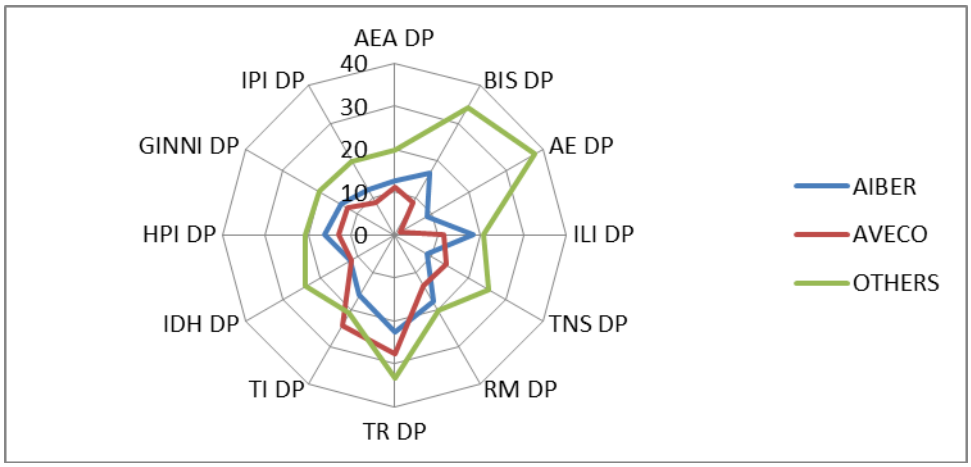
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	1,000	0,280	0,143	-0,091	0,240	0,041	-0,162	0,760	0,680
2	AVECO	1,000	0,250	0,464	-0,339	0,916	0,505	0,103	0,084	0,245
3	OTHERS	1,000	0,470	0,394	0,235	0,969	0,454	0,042	0,031	0,076

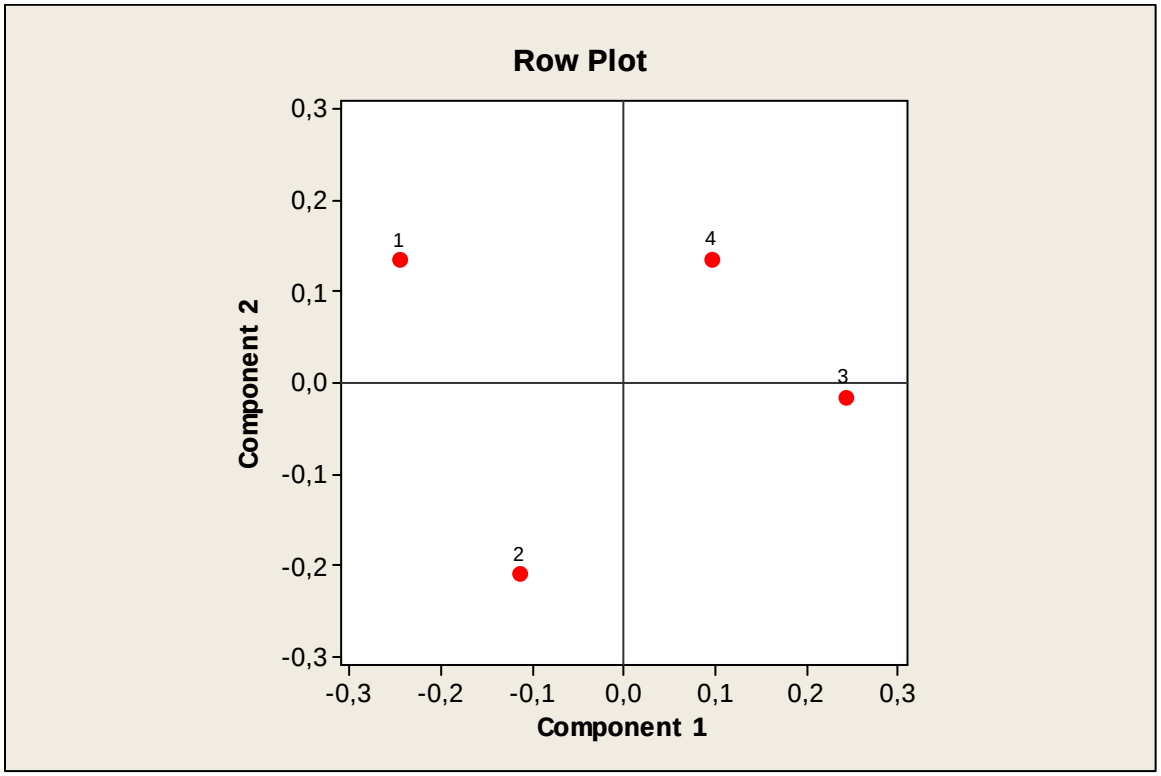
Column Contributions

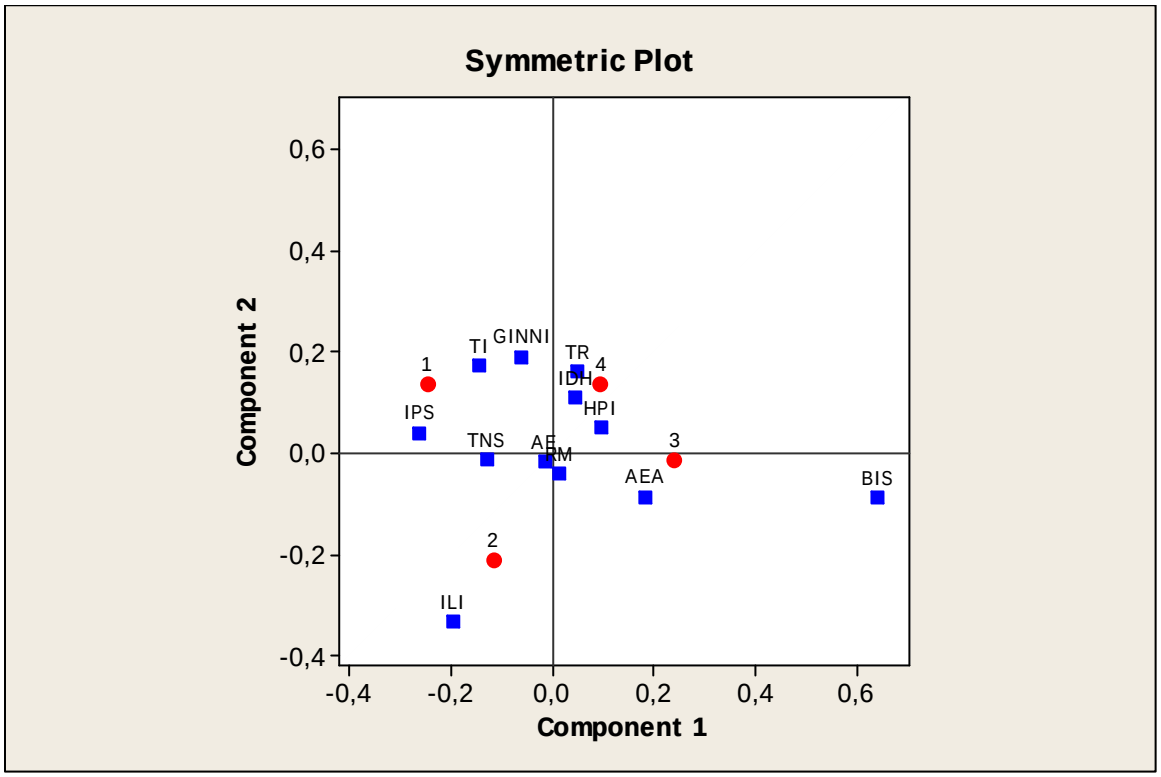
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IPS	1,000	0,069	0,001	-0,032	0,710	0,001	-0,021	0,290	0,003
2	AEA	1,000	0,094	0,091	0,249	0,941	0,102	-0,062	0,059	0,034
3	BIS	1,000	0,076	0,498	0,665	0,991	0,586	0,064	0,009	0,028
4	AE	1,000	0,080	0,040	-0,059	0,103	0,005	-0,175	0,897	0,227
5	ILI	1,000	0,076	0,053	0,040	0,033	0,002	0,215	0,967	0,324
6	TNS	1,000	0,082	0,035	-0,119	0,489	0,020	-0,122	0,511	0,113
7	RM	1,000	0,133	0,072	-0,181	0,888	0,076	0,064	0,112	0,051
8	TI	1,000	0,097	0,173	-0,328	0,890	0,184	0,115	0,110	0,119
9	TR	1,000	0,076	0,006	0,049	0,477	0,003	0,051	0,523	0,018
10	IDH	1,000	0,079	0,017	-0,083	0,475	0,009	-0,087	0,525	0,055
11	HPI	1,000	0,075	0,007	-0,067	0,714	0,006	-0,043	0,286	0,013
12	GINNI	1,000	0,064	0,006	0,061	0,574	0,004	-0,052	0,426	0,016

Row Plot

Symmetric Plot







Simple Correspondence Analysis: AEA DP_2; BIS DP_2; AE DP_2; ILI DP_2; TNS DP_2

Contingency Table

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR
1	15,300	8,210	1,620	17,660	17,390	19,590	26,480	25,000	12,110
2	14,870	17,600	7,510	18,080	32,880	21,120	31,300	15,080	9,700
3	8,390	19,680	22,520	21,540	18,210	17,420	35,490	19,360	14,160

4	10,520	15,500	8,590	11,270	8,610	13,860	20,520	13,210	12,280
Total	49,080	60,990	40,240	68,550	77,090	71,990	113,790	72,650	48,250

	IDH	HPI	GINNI	Total
1	15,780	13,530	13,580	186,250
2	16,260	14,440	8,350	207,190
3	18,960	20,060	12,250	228,040
4	19,170	14,040	9,120	156,690
Total	70,170	62,070	43,300	778,170

Chi-Square Distances

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
1	1,075	2,795	6,664	0,096	0,061	0,323	0,021	3,332	0,027	0,061
2	0,249	0,114	0,958	0,002	7,436	0,199	0,033	0,940	0,771	0,314
3	2,497	0,183	9,759	0,105	0,850	0,641	0,138	0,175	0,000	0,125
4	0,041	0,844	0,029	0,465	3,078	0,028	0,250	0,138	0,677	1,798
Total	3,861	3,936	17,411	0,667	11,425	1,191	0,442	4,584	1,475	2,299

	HPI	GINNI	Total
1	0,118	0,998	15,571
2	0,263	0,876	12,155
3	0,192	0,015	14,680
4	0,190	0,018	7,557
Total	0,764	1,908	49,963

Relative Inertias

	IPS	AEA	BIS	AE	ILI	TNS	RM	TI	TR	IDH
1	0,022	0,056	0,133	0,002	0,001	0,006	0,000	0,067	0,001	0,001
2	0,005	0,002	0,019	0,000	0,149	0,004	0,001	0,019	0,015	0,006
3	0,050	0,004	0,195	0,002	0,017	0,013	0,003	0,004	0,000	0,003
4	0,001	0,017	0,001	0,009	0,062	0,001	0,005	0,003	0,014	0,036
Total	0,077	0,079	0,348	0,013	0,229	0,024	0,009	0,092	0,030	0,046

	HPI	GINNI	Total
1	0,002	0,020	0,312
2	0,005	0,018	0,243
3	0,004	0,000	0,294
4	0,004	0,000	0,151
Total	0,015	0,038	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0369	0,5742	0,5742	******************************
2	0,0199	0,3105	0,8846	***************
3	0,0074	0,1154	1,0000	******
Total	0,0642			

Row Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	1	0,946	0,239	0,312	-0,246	0,725	0,393	0,136	0,222	0,223
2	2	0,976	0,266	0,243	-0,116	0,227	0,096	-0,210	0,748	0,587
3	3	0,905	0,293	0,294	0,241	0,901	0,461	-0,014	0,003	0,003
4	4	0,571	0,201	0,151	0,095	0,186	0,049	0,136	0,385	0,188

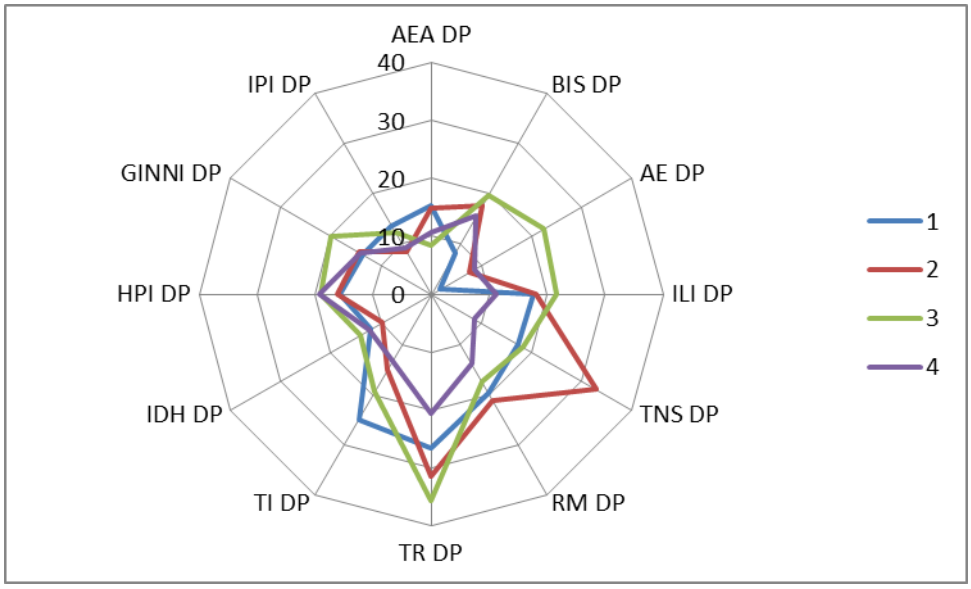
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IPS	0,891	0,063	0,077	-0,262	0,870	0,117	0,040	0,021	0,005
2	AEA	0,640	0,078	0,079	0,184	0,525	0,072	-0,086	0,115	0,029

3	BIS	0,976	0,052	0,348	0,644	0,957	0,581	-0,089	0,019	0,021
4	AE	0,047	0,088	0,013	-0,014	0,019	0,000	-0,017	0,028	0,001
5	ILI	0,998	0,099	0,229	-0,194	0,255	0,101	-0,332	0,743	0,547
6	TNS	0,980	0,093	0,024	-0,127	0,972	0,040	-0,012	0,009	0,001
7	RM	0,518	0,146	0,009	0,016	0,070	0,001	-0,042	0,448	0,013
8	TI	0,788	0,093	0,092	-0,142	0,319	0,051	0,172	0,469	0,139
9	TR	0,915	0,062	0,030	0,051	0,085	0,004	0,159	0,830	0,079
10	IDH	0,428	0,090	0,046	0,046	0,065	0,005	0,109	0,363	0,054
11	HPI	0,979	0,080	0,015	0,098	0,775	0,021	0,050	0,204	0,010
12	GINNI	0,904	0,056	0,038	-0,059	0,080	0,005	0,191	0,824	0,101

Row Plot

Symmetric Plot



CONSIDERAÇÕES FINAIS

Vemos que os países desenvolvidos, sobretudo da Europa Ocidental são os mais bem posicionados, enquanto os países africanos estão bastante atrás. Os países ibero-americanos se encontram quase sempre em uma posição intermediária, muitas vezes puxada por Portugal, Espanha, Chile, Costa Rica Uruguai – excluídos tais países, muito provavelmente a região estaria em um patamar bem pior. É o que veremos em breve, ao explorar em detalhe e profundidade apenas essa região na análise da base de dados.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA: RESILIÊNCIA

TRABALHO FINAL

Disciplina: **METODOS QUANTITATIVOS e QUALITATIVOS DA PESQUISA EMPIRICA**

Prof. Dr. Arnoldo José de Hoyos Guevara

Valeria Regina Bertoncelo

SUMÁRIO

1. INTRODUÇÃO 1
2. ENTENDENDO OS DADOS 1
2.1 HISTÓRICO, OBJETIVO DO GUIA GPS 1
2.2 A GESTÃO PUBLICA E SUSTENTÁVEL DOS PAÍSES 2
2.4 MAPEAMENTO ESTRATÉGIO DO PAÍS 3
2.5 DEFINIÇÃO DO EIXO TEMATÍCO: RESILIÊNCIA 3
2.6 AS VARIÁVEIS DO EIXO TEMÁTICO RESILIÊNCIA 4
2.7 OS INDIVIDUOS 5
3. ANÁLISE EXPLORATÓRIA DOS DADOS 10
3.1 SEGREGAÇÃO POR REGIÕES 10
3.2 ANÁLISE EXPLORATÓRIA DAS VARIÁVEIS 10
4 RELAÇÃO ENTRE AS VARIÁVEIS QUANTITAVAS 20
4.1 CORRELAÇÃO LINEAR 20
4.2 GRÁFICOS DE DISPERSÃO E LINHAS DE TENDÊNCIA 21
4.3 DENDOGRAMA 23
5 REGRESSÃO LINEAR MULTIPLA 25
5.1 CORRELAÇÃO ENTRE AS VARIÁVEIS DO EIXO TEMÁTICO RESILIÊNCIA COM DADOS NORMALIZADOS 28
5.2 GRÁFICOS DE DISPERSÃO E LINHAS DE TENDÊNCIA 29
5.3 DENDOGRAMA 31
5.4 REGRESSÃO LINEAR MULTIPLA 32
6 AMOSTRA 34
7. ANÁLISE DOS COMPONENTES PRINCIPAIS 45
8 ANALISE DE CLUSTERS 48
9 ANÁLISE DE DISCRIMINANTE 51
10 REGRESSÃO LOGÍSTICA 55
12 ANÁLISE DE CORRESPONDÊNCIA 67
13. REGIÃO IBE(AIBER) 73
14. OS DADOS DA REGIÃO IBE(AIBER) 73
15 ANOVA PARA A REGIÃO IBE 73
16. ANÁLISE DOS COMPONENTES PRINCIPAIS 76
17 ANALISE DE CLUSTERS PARA A REGIÃO IBE (AIBER) 84
18 ANÁLISE DE DISCRIMINANTE 90
19 REGRESSÃO LOGÍSTICA 91
20 ÁRVORE CLASSIFICATÓRIA 93
21 ANÁLISE DE CORRESPONDÊNCIA 96
22. MAPA DA REGIÃO IBE(AIBER) 99
23. CONCLUSÃO 103

1. INTRODUÇÃO

O presente trabalho tem por objetivo realizar análise exploratória dos dados sobre fatores de **resiliência** para 132 países, de distintos continentes, constantes como um dos 15 eixos temáticos do Programa Países Sustentáveis descritos no Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP.

Em primeiro lugar, buscamos compreender a base de dados, incluindo a definição dos indivíduos e das variáveis, suas classificações em categorias ou quantitativas, os significados e unidades de medida, além da apresentação da tabela de dados.

As etapas seguintes deste trabalho de pesquisa foram: (1) apresentação do objetivo do Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica; (2) aplicação da Estatística Descritiva, através do software MINITAB e SPSS , para a geração dados básicos como por exemplo: média, mediana, quartis, desvio-padrão, variância, intervalo de confiança e teste de normalidade de Anderson-Darling; (3) análise exploratória dos resultados de cada variável relativa ao tema: **resiliência.**

2. ENTENDENDO OS DADOS

2.1 HISTÓRICO, OBJETIVO DO GUIA GPS

O mundo atravessa um período de rápidas transformações, como resultado do acelerado desenvolvimento da era industrial e pós-industrial, em que se manifestam crises sistêmicas de governanças nas esferas local e global, envolvendo aspectos, socioeconômicos e ambientais para os países em desenvolvimento.

A dinâmica dessas mudanças tem impactos diferenciados, conforme seu próprio contexto político e estágio de desenvolvimento; porém na recente pesquisa global: Meu Mundo (myworld2015) realizada pela ONU, as três principais prioridades encontradas para Um Mundo Melhor foram: Educação de Qualidade, Melhor Cuidado com Saúde e Melhores Oportunidade de Emprego.

Pouco tempo atrás o Projeto Milênio publicou o informe Latino américa 2030 trabalhando com quatro cenários, que contrastam os Avanços Sócio-Políticos com os Tecno-Econômicos da região. Estes cenários consideram múltiplas possibilidades futuras que dependem das ações presentes. Globalmente existem grandes preocupações em relação com o futuro do trabalho e aumento na desigualdade, o que é fundamental. Junto com outras variáveis estas precisam ser levadas em conta para evitar as piores situações e alcançar as melhores possibilidades ao nível Ibero-americano. Num futuro com múltiplas alternativas, as intervenções presentes podem determinar em grande parte os acontecimentos futuros. As decisões tomadas hoje serão fundamentais para toda essa região no amanhã. Isso em particular representa um grande desafio para a região devido as deficiências em termos de Governabilidade e Democracia.

Neste contexto, a grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?

O GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o

sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos.

O guia GPS de gestão pública sustentável é um exemplo de isso, e representa uma nova versão para países do que foi desenvolvido pouco tempo atrás como um Guia para o Gestão Pública Sustentável de Cidades do Programa de Cidades Sustentáveis da Rede Nossa São Paulo.

2.2 A GESTÃO PUBLICA E SUSTENTÁVEL DOS PAÍSES

Cerca de metade da humanidade vive hoje nas cidades, percentual que deverá chegar a 60% em 2030, de acordo com as estimativas. Na América Latina, o Brasil é o país mais urbanizado, resultado de um intenso processo de estruturação das cidades iniciado na década de 1950, que provocou a concentração de 85% de sua população nas áreas urbanas. As estimativas dão conta de que esse percentual possa chegar a 90% até 2020. Na América Latina como a urbanização está perto de 80%, o crescimento da população que vive nas cidades acarreta novos e complexos desafios para os gestores públicos locais, pressionando a infraestrutura e o consumo dos recursos naturais. Com efeito, as cidades passaram a ser protagonistas nos processos de decisão, ao mesmo tempo em que tiveram de enfrentar problemas relacionados à desigualdade social, à poluição, às dificuldades de mobilidade, ao excesso de resíduos, à falta de saneamento básico, habitações precárias, violência e mudanças climáticas. Tais problemas demandam em particular a criação de um novo modelo de gestão pública municipal, que possa incluir planos estratégicos eficientes e equipes bem preparadas para desenvolvê-los.

A elaboração de um planejamento estratégico se mostra necessário uma abordagem sistêmica e sustentável deve ser capaz de interagir com o campo o econômico, cultural, social, ecológico, tecnológico, tributário e demográfico de cada país.

O planejamento estratégico baseado em uma abordagem sistêmica deve considerar, assim, a execução dos projetos sob uma visão de curto, médio e longo prazo, a fim de assegurar a continuidade dos programas, especialmente dos investimentos em infraestrutura, dada a característica de longo prazo. Também, é importante o estabelecimento de metas passíveis de monitoramento público em longo prazo e que estejam contempladas no Plano Diretor e no Plano Plurianual.

Países bem-sucedidos na concepção e execução de seu Plano Diretor e do Plano de Metas utilizaram como subsídio o mapeamento dos projetos já existentes e o seu alinhamento com o contexto externo.

2.3 EIXOS TEMÁTICOS DO PROGRAMA PAÍSES SUSTENTÁVEIS

O Guia GPS adotou 15 (quinze) eixos temáticos do Programa de Países sustentáveis para fins de análise exploratória de dados e da avaliação das políticas públicas contidas no planejamento estratégico e sustentável de cada país, como seguem:

1. Recursos básicos: água, alimento e energia
2. Moradia
3. Segurança e paz
4. Saúde e meio ambiente
5. Desigualdade e inclusão social de gênero
6. O futuro da educação e do trabalho
7. Transformação produtiva e inovação sustentável
8. Integração e alianças estratégicas
9. Tendências de longevidade
10. Mudança climática e energias renováveis
11. Biodiversidade: capital natural e social
12. **Resiliência**
13. Democracia e redes sociais
14. Governança e empoderamento cidadão
15. Bem-estar e qualidade de vida

2.4 MAPEAMENTO ESTRATÉGIO DO PAÍS

O planejamento das metas necessita de dados referentes aos respectivos indicadores, os quais precisam ser coletados, organizados, e analisados previamente. Após a organização dos dados, o passo seguinte é a descrição de um diagnóstico do país. Este diagnóstico deve exprimir as perspectivas: (i) interna - uma visão sobre o funcionamento do país isoladamente, considerando seus pontos fortes (quais os indicadores já atingiram as metas sugeridas ou estão próximos de atingir) e fracos (quais indicadores possuem valores distantes das metas sugeridas) e (ii) externa sobre o país - deve-se visualizar o país inserido na sua região mais próxima ou Ibero-americana, levando em conta as oportunidades oferecidas no âmbito dessas três esferas. Simultaneamente, é necessário projetar possíveis ônus que as dinâmicas nos contextos regionais possam trazer ao país.

2.5 DEFINIÇÃO DO EIXO TEMATÍCO: RESILIÊNCIA

A resiliência é, pois, um conceito que tem sido utilizado para explicar fenômenos que superam ou transcendem situações desfavoráveis. Trata-se de um tema de extrema importância, pois implica em maior conhecimento de como podemos fortalecer a capacidade de resistência da sociedade e da natureza, aprendendo a lidar com as tensões causadas pelas alterações climáticas e outros impactos ambientais.

Cidades ou países resilientes são os que conseguem se recuperar rapidamente dos impactos de catástrofes naturais ou de efeitos de condições econômicas e/ou sociais desfavoráveis ou de situações extremas como revoltas ou guerras.

Uma série de evidências e de Relatórios Climáticos aponta para uma situação em que os períodos de mudança abrupta devem se tornar mais presentes, aumentando tanto em frequência como em magnitude. Isto desafia a capacidade de adaptação das sociedades modernas.

Vivemos um tempo caracterizado pela velocidade e amplitude de transformações tecnológicas, sociais e econômicas. Isso demanda uma grande capacidade de adaptação ao novo, tanto de indivíduos como de instituições. Busca-se a flexibilidade de estrutura e até mesmo de vida pessoal, como forma de ajuste às novas condições e necessidades. Por vezes, mudanças podem ocorrer de forma gradual, onde os impactos se movem de forma mais ou menos contínua e previsível. Em outros casos, a mudança é súbita, desorganizada e turbulenta, refletindo em profundos impactos climáticos, impondo imensos desafios para as regiões do globo mais vulneráveis ou propensas a desastres naturais. O estudo da resiliência enfoca a interação e a dinâmica entre os períodos de mudança gradual e súbita e como devemos nos adaptar a estas circunstâncias.

De acordo com os pesquisadores do ***Stockholm Resilience Centre***, a resiliência pode ser definida como a capacidade de longo prazo para um sistema lidar com variações ou mudanças e continuar a se desenvolver. Para um ecossistema como uma floresta, isso pode envolver lidar com tempestades, incêndios e poluição, enquanto para uma sociedade envolve a habilidade para lidar com as incertezas políticas ou desastres naturais de uma forma que seja sustentável em longo prazo.

2.6 AS VARIÁVEIS DO EIXO TEMÁTICO RESILIÊNCIA

Os dados analisados para cada região sob o eixo temático resiliência ssão as variáveis que descritas no **Quadro 1** abaixo.

Quadro 1: As variáveis do eixo temático: Resiliência

	Variável Sintética	Significado	Tipo	Unidade de Medida
Região	Região	Área geográfica/ politica/ linguística comum	Categórica	Unidade
IPS	Índice de Progresso Social	Desempenho social por (1) Necessidades Básicas, (2) Bem-Estar, (3) Oportunidades	Quantitativa	Percentual
RISCN	Índice Mundial de risco a desastres naturais	O Índice de Risco Mundial consiste em quatro componentes: Exposição para os perigos naturais, a suscetibilidade, capacidades de enfrentamento e capacidades adaptativas. O Índice é calculado a partir de 28 indicadores que utilizam dados disponibilizados pelos países e acessíveis a todos.	Quantitativa	Percentual
EXP	Índice Mundial de risco a desastres naturais - **Exposição**	A exposição refere-se à população, condições de áreas construídas, componentes de infraestrutura e ambiental expostos aos impactos de um ou mais riscos naturais, tais como: terremotos, ciclones, secas, inundações e elevação do nível do mar etc.	Quantitativa	Percentual
SUSC	Índice Mundial de risco a desastres naturais – **Suscetibilidade**	Susceptibilidade refere-se à probabilidade de sofrer perda, dano ou interrupção face a eventos extremos ou de desastres naturais. Assim, a susceptibilidade descreve as características estruturais disponíveis, tais como: moradia, distribuição de renda etc. e o respectivo enquadramento social.	Quantitativa	Percentual
VULN	Índice Mundial de risco a desastres naturais – **Vulnerabilidade**	Compreende à suscetibilidade, associadas a falta de capacidade de resposta e de adaptação da sociedade, país à fatores sociais, físicos, econômicos e ambientais que os tornam a suscetíveis aos impactos dos riscos de desastres naturais, aos efeitos adversos das alterações climáticas ou de processos de	Quantitativa	Percentual

	Variável Sintética	Significado	Tipo	Unidade de Medida
		transformação. É, também, demonstram a capacidade da sociedade, país em enfrentar e adaptar-se aos impactos negativos dos desastres naturais.		
SDIF	Índice Mundial de risco a desastres naturais – **Capacidade de enfrentar situações difíceis.**	É a habilidade da sociedade em minimizar os impactos negativos oriundos de desastres naturais, às alterações climáticas via ação direta e da disponibilidade de recursos para redução de danos e prejuízos na ocorrência destes eventos.	Quantitativa	Percentual
CADAPT	Índice Mundial de risco a desastres naturais – **Capacidade de adaptação**	Adaptação, ao contrário da capacidade de enfrentar e sobreviver, é entendida como um processo de longo prazo que inclui mudanças estruturais. Concentra-se em medidas e estratégias que abordam os impactos negativos dos desastres naturais e das mudanças climáticas em longo prazo.	Quantitativa	Percentual
IDH	Índice Desenvolvimento humano	Mede o progresso de um país por indicadores de qualidade de vida: renda, saúde e educação são os principais.	Quantitativa	Indicador nominal que varia de 0 a 1
GO	Governança	Capacidade de conduzir os processos de formulação, execução e avaliação de políticas públicas integrando instrumentos e mecanismos de gestão.	Quantitativa	Percentual
EPI	Índice de desempenho ambiental	Mede (1) proteção da saúde frente a riscos ambientais e (2) proteção dos ecossistemas.	Quantitativa	Percentual
HPI	Bem-estar sustentável do país	Medida de bem-estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	Quantitativa	Percentual
GINI	GINI index	Mede a extensão em que a distribuição de renda (ou, em alguns casos, a despesa de consumo) entre indivíduos ou famílias dentro de uma economia desvia de uma distribuição perfeitamente igual.	Quantitativa	Percentual

Fonte: Adaptado pelo autor – GPS - Guia para Gestão Pública Sustentável – Acesso em 04/03/2016.

2.7 OS INDIVIDUOS

Os indivíduos desta análise exploratória de dados são 132 países de todos os continentes, divididos em 8 regiões: AIBER, AVECO, CARLA, CWEAS, ESEAS, NSWUA, SUSAF, EMDEU e 3 regiões: AIBER, AVECO e OTHERS, conforme demonstrado em quadro 2 a seguir.

Quadro 2 -

	Pais	REGIÃO	Class 6R	Class 3R	Clusters 3R
1	Australia	Australia and New Zealand	AVECO	AVECO	1
2	New Zealand	Australia and New Zealand	AVECO	AVECO	1
3	Jamaica	Caribbean	CARLA	OTHERS	2
4	Trinidad and Tobago	Caribbean	CARLA	OTHERS	2
5	Kazakhstan	Central Asia	CWEAS	OTHERS	2
6	Kyrgyzstan	Central Asia	CWEAS	OTHERS	3
7	Tajikistan	Central Asia	CWEAS	OTHERS	3
8	Uzbekistan	Central Asia	CWEAS	OTHERS	3
9	Burundi	Eastern Africa	SUSAF	OTHERS	3

	Pais	REGIÃO	Class 6R	Class 3R	Clusters 3R
10	Central African Republic	Eastern Africa	SUSAF	OTHERS	3
11	Djibouti	Eastern Africa	SUSAF	OTHERS	3
12	Kenya	Eastern Africa	SUSAF	OTHERS	3
13	Madagascar	Eastern Africa	SUSAF	OTHERS	3
14	Mauritius	Eastern Africa	SUSAF	OTHERS	1
15	Mozambique	Eastern Africa	SUSAF	OTHERS	3
16	Rwanda	Eastern Africa	SUSAF	OTHERS	3
17	Tanzania	Eastern Africa	SUSAF	OTHERS	3
18	Uganda	Eastern Africa	SUSAF	OTHERS	3
19	China	Eastern Asia	ESEAS	OTHERS	2
20	Japan	Eastern Asia	AVECO	AVECO	1
21	Korea, Republic of	Eastern Asia	AVECO	AVECO	1
22	Mongolia	Eastern Asia	ESEAS	OTHERS	2
23	Belarus	Eastern Europe	CWEAS	OTHERS	1
24	Bulgaria	Eastern Europe	EMDEU	OTHERS	1
25	Czech Republic	Eastern Europe	AVECO	AVECO	1
26	Hungary	Eastern Europe	EMDEU	OTHERS	1
27	Moldova	Eastern Europe	CWEAS	OTHERS	2
28	Poland	Eastern Europe	EMDEU	OTHERS	1
29	Romania	Eastern Europe	EMDEU	OTHERS	2
30	Russia	Eastern Europe	CWEAS	OTHERS	2
31	Slovakia	Eastern Europe	AVECO	AVECO	1
32	Ukraine	Eastern Europe	CWEAS	OTHERS	2
33	Argentina	IBE	AIBER	AIBER	2
34	Bolivia	IBE	AIBER	AIBER	3
35	Brazil	IBE	AIBER	AIBER	2
36	Chile	IBE	AIBER	AIBER	1
37	Colombia	IBE	AIBER	AIBER	2
38	Costa Rica	IBE	AIBER	AIBER	1
39	Cuba	IBE	CARLA	AIBER	2
40	Dominican Republic	IBE	AIBER	AIBER	2
41	Ecuador	IBE	AIBER	AIBER	2
42	El Salvador	IBE	AIBER	AIBER	2
43	Guatemala	IBE	AIBER	AIBER	2
44	Honduras	IBE	AIBER	AIBER	2
45	Mexico	IBE	AIBER	AIBER	2
46	Nicaragua	IBE	AIBER	AIBER	2
47	Panama	IBE	AIBER	AIBER	2
48	Paraguay	IBE	AIBER	AIBER	3
49	Peru	IBE	AIBER	AIBER	2
50	Portugal	IBE	AIBER	AIBER	1
51	Spain	IBE	AIBER	AIBER	1
52	Uruguay	IBE	AIBER	AIBER	1
53	Venezuela	IBE	AIBER	AIBER	2
54	Angola	Middle Africa	SUSAF	OTHERS	3
55	Cameroon	Middle Africa	SUSAF	OTHERS	3
56	Chad	Middle Africa	SUSAF	OTHERS	3
57	Congo, Republic of	Middle Africa	SUSAF	OTHERS	3
58	Zambia	Middle Africa	SUSAF	OTHERS	3
59	Algeria	Northern Africa	NSWUA	OTHERS	3
60	Egypt	Northern Africa	NSWUA	OTHERS	2
61	Morocco	Northern Africa	NSWUA	OTHERS	2
62	Sudan	Northern Africa	SUSAF	OTHERS	3

	Pais	REGIÃO	Class 6R	Class 3R	Clusters 3R
63	Tunisia	Northern Africa	NSWUA	OTHERS	2
64	Canada	Northern America	AVECO	AVECO	1
65	United States	Northern America	AVECO	AVECO	1
66	Denmark	Northern Europe	AVECO	AVECO	1
67	Estonia	Northern Europe	AVECO	AVECO	1
68	Finland	Northern Europe	AVECO	AVECO	1
69	Iceland	Northern Europe	AVECO	AVECO	1
70	Ireland	Northern Europe	AVECO	AVECO	1
71	Latvia	Northern Europe	AVECO	AVECO	1
72	Lithuania	Northern Europe	EMDEU	OTHERS	1
73	Norway	Northern Europe	AVECO	AVECO	1
74	Sweden	Northern Europe	AVECO	AVECO	1
75	United Kingdom	Northern Europe	AVECO	AVECO	1
76	Cambodia	South-Eastern Asia	ESEAS	OTHERS	3
77	Indonesia	South-Eastern Asia	ESEAS	OTHERS	2
78	Laos	South-Eastern Asia	ESEAS	OTHERS	3
79	Malaysia	South-Eastern Asia	ESEAS	OTHERS	2
80	Philippines	South-Eastern Asia	ESEAS	OTHERS	2
81	Thailand	South-Eastern Asia	ESEAS	OTHERS	2
82	Guyana	South America	CARLA	OTHERS	2
83	Botswana	Southern Africa	SUSAF	OTHERS	2
84	Lesotho	Southern Africa	SUSAF	OTHERS	3
85	Malawi	Southern Africa	SUSAF	OTHERS	3
86	Namibia	Southern Africa	SUSAF	OTHERS	2
87	South Africa	Southern Africa	SUSAF	OTHERS	2
88	Swaziland	Southern Africa	SUSAF	OTHERS	3
89	Bangladesh	Southern Asia	ESEAS	OTHERS	3
90	India	Southern Asia	ESEAS	OTHERS	3
91	Iran	Southern Asia	NSWUA	OTHERS	2
92	Nepal	Southern Asia	ESEAS	OTHERS	3
93	Pakistan	Southern Asia	NSWUA	OTHERS	3
94	Sri Lanka	Southern Asia	ESEAS	OTHERS	2
95	Albania	Southern Europe	EMDEU	OTHERS	2
96	Bosnia and Herzegovina	Southern Europe	EMDEU	OTHERS	2
97	Croatia	Southern Europe	EMDEU	OTHERS	1
98	Greece	Southern Europe	AVECO	AVECO	1
99	Italy	Southern Europe	AVECO	AVECO	1
100	Macedonia	Southern Europe	EMDEU	OTHERS	2
101	Montenegro	Southern Europe	EMDEU	OTHERS	2
102	Serbia	Southern Europe	EMDEU	OTHERS	2
103	Slovenia	Southern Europe	AVECO	AVECO	1
104	Benin	Western Africa	SUSAF	OTHERS	3
105	Burkina Faso	Western Africa	SUSAF	OTHERS	3
106	Ghana	Western Africa	SUSAF	OTHERS	3
107	Guinea	Western Africa	SUSAF	OTHERS	3
108	Liberia	Western Africa	SUSAF	OTHERS	3
109	Mali	Western Africa	SUSAF	OTHERS	3
110	Mauritania	Western Africa	NSWUA	OTHERS	3
111	Niger	Western Africa	SUSAF	OTHERS	3
112	Nigeria	Western Africa	SUSAF	OTHERS	3
113	Senegal	Western Africa	SUSAF	OTHERS	3
114	Togo	Western Africa	SUSAF	OTHERS	3
115	Armenia	Western Asia	CWEAS	OTHERS	2

	Pais	REGIÃO	Class 6R	Class 3R	Clusters 3R
116	Azerbaijan	Western Asia	CWEAS	OTHERS	2
117	Georgia	Western Asia	CWEAS	OTHERS	2
118	Iraq	Western Asia	NSWUA	OTHERS	3
119	Israel	Western Asia	AVECO	AVECO	1
120	Jordan	Western Asia	NSWUA	OTHERS	2
121	Kuwait	Western Asia	NSWUA	OTHERS	1
122	Lebanon	Western Asia	NSWUA	OTHERS	2
123	Saudi Arabia	Western Asia	NSWUA	OTHERS	1
124	Turkey	Western Asia	EMDEU	OTHERS	2
125	United Arab Emirates	Western Asia	NSWUA	OTHERS	1
126	Yemen	Western Asia	NSWUA	OTHERS	3
127	Austria	Western Europe	AVECO	AVECO	1
128	Belgium	Western Europe	AVECO	AVECO	1
129	France	Western Europe	AVECO	AVECO	1
130	Germany	Western Europe	AVECO	AVECO	1
131	Netherlands	Western Europe	AVECO	AVECO	1
132	Switzerland	Western Europe	AVECO	AVECO	1

Fonte: Adaptado pelo autor – GPS - Guia para Gestão Pública Sustentável – Acesso em 04/03/2016.

2.8 OS DADOS

Os dados utilizados neste trabalho foram obtidos pelo Núcleo de Estudos do Futuro – Guia GPS, onde foi aplicado o software Minitab para a obtenção da seguinte estatística descritiva para N=132 países: Média (Mean), Erro padrão da Amostra (SE Mean), desvio padrão (STDEV), mínimo (Minimum), máximo (Máximum), conforme demonstrado na **Tabela 1**, abaixo. Os resultados aplicados para a amostra de 132 países indicam a ausência de alguns dados de países para: **(i)** índice desenvolvimento humano **(IDH)** – Algéria, Kasakhstan, Korea, Kyrgystan, Latvia, Mauritius e Netherlands; **(ii)** índice mundial de riscos a desastres naturais **(RISCN)**, exposição **(EXP)**, vulnerabilidade **(VUL)**, suscetibilidade **(SUS)**, capacidade de sobrevivência à situações difíceis **(SIF)** e da adaptação **(CADAPT)** – Montenegro; **(iii)** do bem-estar **(HPI)** - Lesotho, Montenegro e Swaziland; e do **(iv)** GINI **(GINI)** – Bangladesh, Cuba, Guinea, Iceland, Kuwait, Lebanon, Mauritius, Montenegro, Nepal, Pakistan, Saudi Arabia, Sri lanka e United Arab Emirates.

Tabela 1 – Resultado Estatística Descritiva para eixo temático: Resiliência –

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
IPS	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
IDH	125	7	0,69	0,01	0,16	0,34	0,56	0,72	0,82	0,94
RISCN	131	1	6,96	0,36	4,16	1,17	3,85	6,42	8,25	28,25
EXP	131	1	14,51	0,68	7,80	2,93	10,21	12,34	15,77	52,46
VULN	131	1	47,51	1,21	13,89	22,27	36,08	45,90	60,03	77,12
SUSC	131	1	30,31	1,36	15,60	10,47	17,76	23,64	40,91	69,89
SDIF	130	2	68,36	1,35	15,37	37,61	55,89	70,69	81,13	93,05
CADAPT	131	1	43,79	1,09	12,47	20,96	33,56	44,02	53,98	71,21
GO	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
HPI	129	3	42,67	0,80	9,04	22,59	36,16	42,46	49,29	64,04
GINNI	119	13	39,02	0,78	8,53	24,82	32,78	38,16	44,55	63,14

Face a ausência de dados para algumas variáveis do eixo resiliência para alguns países, adotou-se pela inclusão dos valores médios demonstrados na coluna Mean, na **Tabela 1.** A seguir, na Tabela 2, demonstra os resultados estatísticos após os ajustes de inclusão.

Tabela 2 – Resultado Estatística Descritiva para eixo temático: Resiliência -

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
IPS	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
IDH	132	0	0,69	0,01	0,16	0,34	0,57	0,72	0,81	0,94
RISCN	132	0	6,96	0,36	4,14	1,17	3,86	6,47	8,25	28,25
EXP	132	0	14,51	0,68	7,77	2,93	10,22	12,40	15,73	52,46
VULN	132	0	47,51	1,20	13,84	22,27	36,19	45,98	59,98	77,12
SUSC	132	0	30,31	1,35	15,54	10,47	17,80	23,81	40,75	69,89
SDIF	132	0	68,36	1,33	15,25	37,61	56,10	70,20	80,94	93,05
CADAPT	132	0	43,79	1,08	12,42	20,96	33,59	43,91	53,97	71,21
GO	132	0	53,38	1,88	21,59	15,22	37,11	48,46	70,36	100,00
EPI	132	0	52,48	1,44	16,55	18,43	39,30	52,09	66,40	87,67
HPI	132	0	42,67	0,78	8,94	22,59	36,59	42,63	49,18	64,04
GINNI	132	0	39,02	0,71	8,10	24,82	33,07	39,02	43,55	63,14

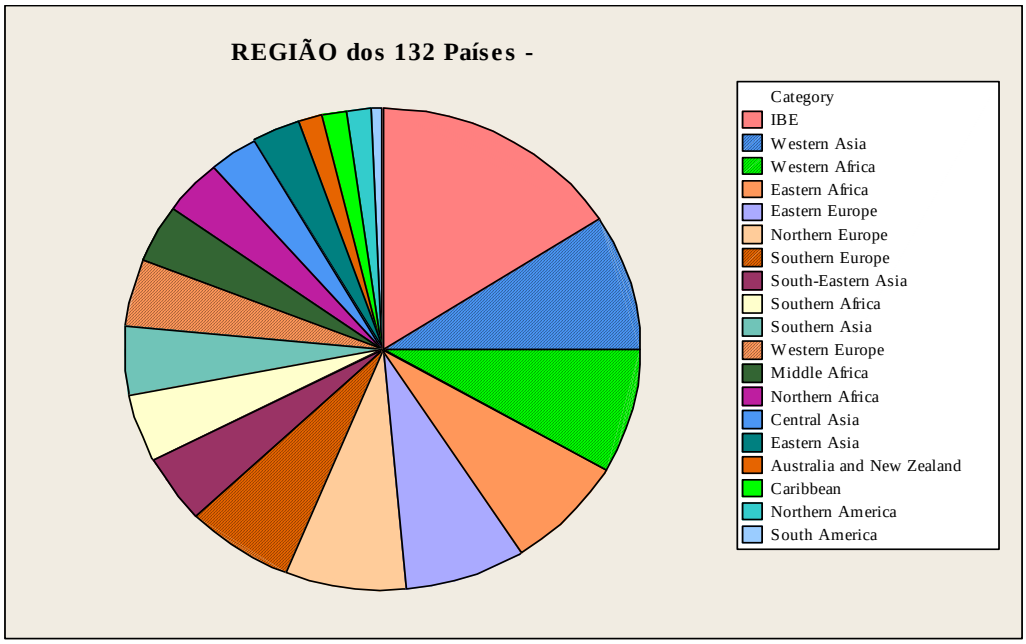
Fonte: Adaptado pelo autor – GPS - Guia para Gestão Pública Sustentável – Acesso em 04/03/2016.

3. ANÁLISE EXPLORATÓRIA DOS DADOS

3.1 SEGREGAÇÃO POR REGIÕES

A amostra do trabalho são 132 países categorizados em 19 regiões dos continentes Europeu, Africano, Oceania, Asiático e Americano – Quadro 2

Gráfico 1 – Região dos 132 países para análise do eixo: Resiliência



Fonte: Autor – Dados Guia GPS –

3.2 ANÁLISE EXPLORATÓRIA DAS VARIÁVEIS

O eixo temático: resiliência do Programa de Países sustentáveis representadas pelas variáveis: IPS, IDH, Risco Mundial à Desastres Naturais e seus desdobramentos, Governança, EPI, HPI e GINI, os quais foram utilizadas para a análise exploratória de dados, demonstram que ainda são necessárias ações da iniciativa pública e privada para que haja certa uniformidade entre os 132 países categorizados por regiões. Observa-se que os países categorizados por regiões com IDH e IPS superiores à média, também, tendem a ter melhor aderência a resiliência, ou seja, boa capacidade de superação aos efeitos adversos.

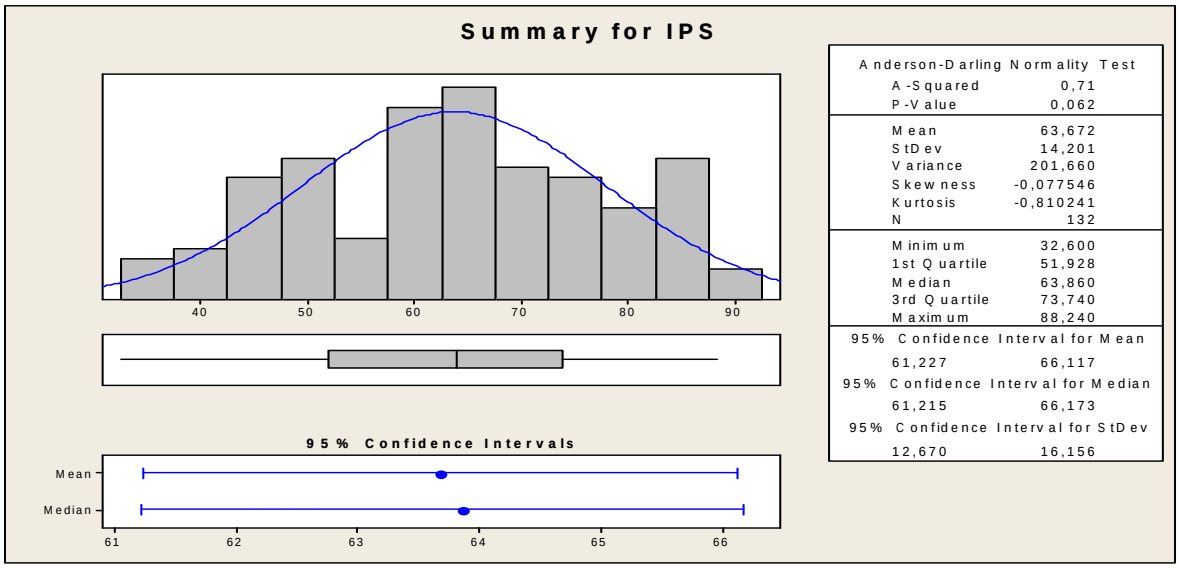
IPS – ÍNDICE DE PROGRESSO SOCIAL –

O índice de progresso social leva em conta três atributos para a composição do índice: (i) capacidade dos países de possibilitar as mínimas condições fisiológicas de vida; (ii) de bem ou de ter acesso a conhecimentos básicos, à informação e comunicação, saúde e bem-estar e eco sustentabilidade; (iii) condições de oportunidade, ou de direitos pessoais, liberdade pessoal e de eleição, tolerância e inclusão, acesso à educação superior.

Gráfico 2 - Variável: Índice de Progresso Social das Países, categorizados em Regiões, demonstra o IPS médio de 63,672%, para um intervalo de confiança de 95% e desvio padrão de 14,201%. O histograma demonstra que 60 países das regiões categorizadas estão abaixo do IPS médio e inseridas entre 32,5-62,5%: **Central Asia**, **Eastern Africa**, **Middle Africa**, **Northern Africa**, **South America**, **South-Eastern Asia**, **Southern Africa**, **Western Africa**, **Western Asia**.

O remanescente dos 70 países se encontram acima de 62,5%, dos quais 38 atingem índices superiores a 72,5% de IPS e 34% estão entre 62,5% a 72,5%.
Entre os extremos: mínimo (32,60%) e máximo (88,24%), há: Chad (**Middle Africa**) e Nova Zelândia (**Austrália e New Zealand)**, respectivamente.
De forma geral, observamos que 45% da amostra ainda não promove à sociedade de forma igualitária de bem-estar, oportunidade e de condições mínimas de vida a sua população.
É importante lembrar que o Índice IPS exclui de sua base de formação do índice a variável: produto interno bruto (PIB), uma vez que há o entendimento de que o progresso social nem sempre é explicado pela elevada renda per capita, como é, por exemplo, o caso dos Emirados Árabes e da Costa Rica.

Gráfico 2: Variável: Índice de Progresso Social dos Países, categorizados em Regiões

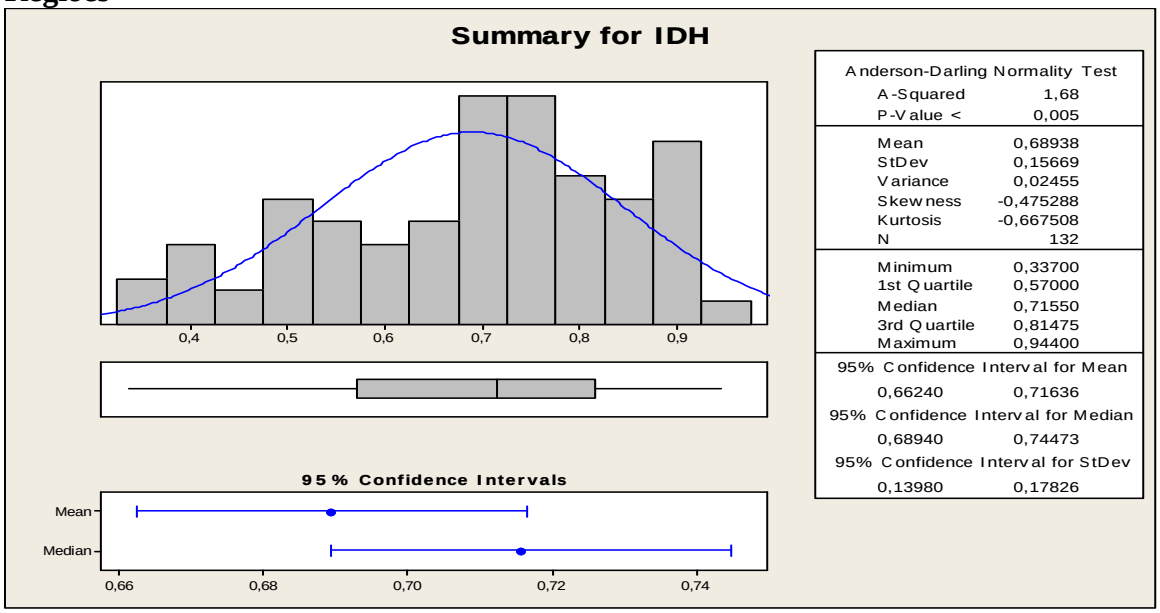


Fonte: Autor – Dados Guia GPS –

IDH – ÍNDICE DE DESENVOLVIMENTO HUMANO

O IDH mede o progresso de um país por indicadores de qualidade de vida: renda, saúde e educação, como os principais atributos. Para o PNUD – Programa de las Naciones Unidas para el Desarrollo, o IDH indica o progresso real da população em relação a expectativa de vida, educação, saúde, habitação segurança e condições para desenvolvimento sustentável.
Gráfico 3 - Variável: Índice de Desenvolvimento Humano dos Países, categorizados em Regiões, demonstra o IDH médio de 0,68938, para um intervalo de confiança de 95% e desvio padrão de 0,15669. O histograma abaixo, demonstra que 82 países das regiões categorizadas se encontram entre a escala de 0,675 a 0,975, portanto, igual ou superior à média para a amostra que é de 0,68938. Os 50 demais países estão na escala entre 0,325-0,675. Também, demonstra que 60 países das regiões categorizadas estão abaixo do IPS médio e inseridas entre 0,325-0,625: **Central Asia, Eastern Africa, Middle Africa, Northern Africa, South America, South-Eastern Asia, Southern Africa, Southern Asia, Western Africa.**
Entre os extremos: mínimo (0,337) e máximo (88,24%), há: Niger (**Western Africa**) e Noruega (**Northern Europe)**, respectivamente.

Gráfico 3: Variável: Índice de Desenvolvimento Humano dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

RISCN - INDICE MUNDIAL DE RISCOS A DESASTRES NATURAIS

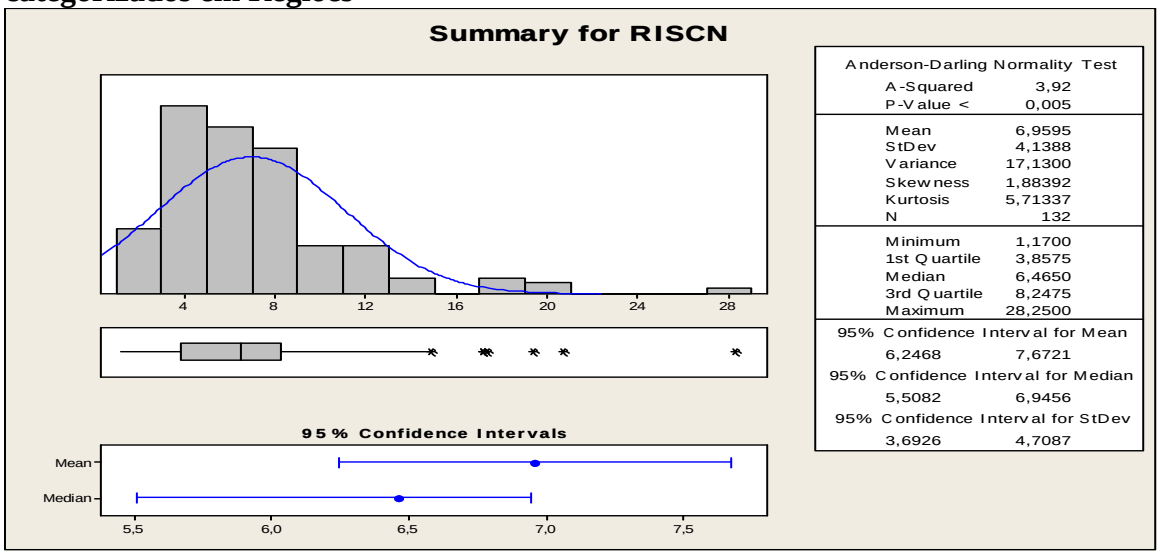
O Índice de Risco Mundial a Desastres Naturais (RISCN) é uma ferramenta utilizada para avaliar e estimar o risco destes em um país. Considera fatores externos e internos, ou seja, ameaças, perigos naturais como terremotos, ciclones, inundações, secas e elevação do nível do mar e condições sociais. A combinação da exposição dos países aos riscos naturais e as condições sociais, econômicas e ecológicas destes, é possível calcular o risco potencial de cada país. O objetivo do **RISCN** é demonstrar que não só a magnitude ou intensidade de um evento natural influencia o risco de desastre, mas uma multiplicidade de fatores diferentes, tais como as estruturas políticas e institucionais, o estado da infraestrutura ou a situação nutricional, condições económicas e ambientais de um país determinar se um perigo natural se transformará em um desastre (BIRKMANN et al 2011; WELLE et al 2012, 2013; IPCC 2012).

O RISCN é o resultado da ponderação dos indicadores principais de suscetibilidade, exposição, capacidade de adaptação e de enfrentar situações difíceis. Estes por sua vez, totalizam cerca de 28 indicadores. Neste caso quanto menor for o índice ou próximo a 1, indica que o país possa não ser totalmente suscetível a desastres naturais ou de que tem políticas públicas e de infraestrutura para minimizar eventual risco à desastres naturais.

Gráfico 4 - Variável: Índice Mundial de Riscos a Desastres Naturais dos Países, categorizados em Regiões, demonstra o **RISCN** médio de 6,9595, para um intervalo de confiança de 95% e desvio padrão de 4,1388. O histograma abaixo, demonstra que 85 países das regiões categorizadas se encontram entre a escala de 5 a 29, portanto, igual ou superior à média para a amostra que é de 6,9595, sendo: **Caribbean**, **Central Asia**, **Eastern Asia, Eastern Africa**, **IBE**, **Middle Africa**, **Northern Africa**, **South America**, **South-Eastern Asia**, **Southern Africa**, **Southern Asia, Southern Europe, Western Africa**.

Entre os extremos: mínimo (1,17) e máximo (28,25), há: Filipinas (**South-Eastern Asia**) e Arábia Saudita (**Western Asia)**, respectivamente.

Gráfico 4: Variável: Índice Mundial de Riscos a Desastres Naturais dos Países, categorizados em Regiões



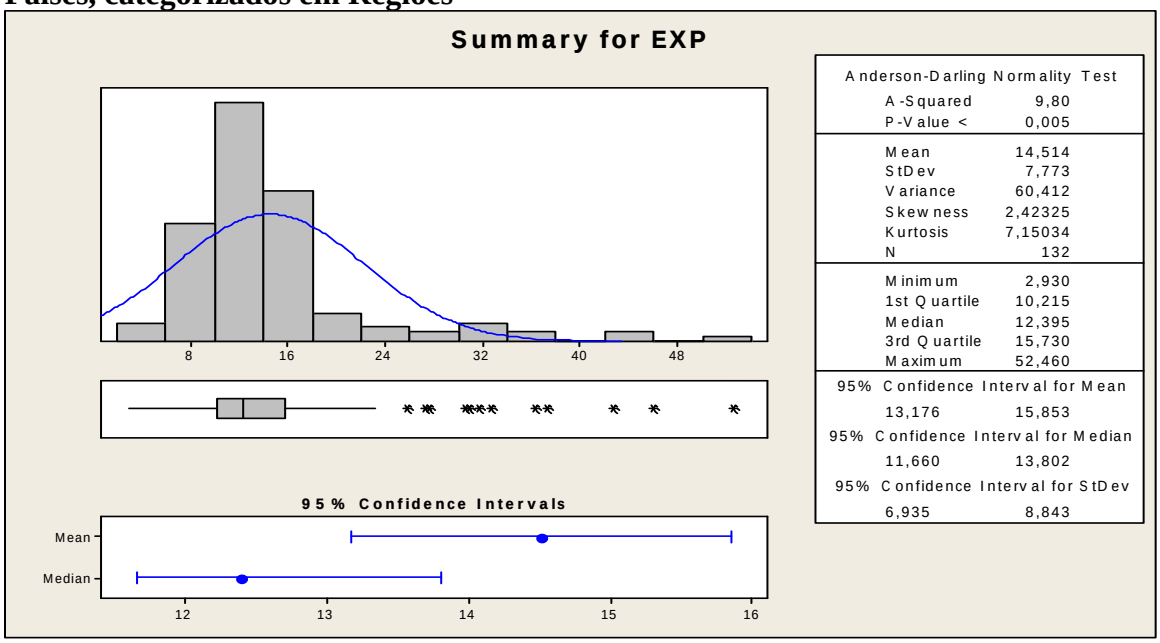
Fonte: Autor – Dados Guia GPS –

EXP - ÍNDICE MUNDIAL DE RISCO A DESASTRES NATURAIS - EXPOSIÇÃO

A exposição (EXP) refere-se à população, condições de áreas construídas, componentes de infraestrutura, da área ambiental, sendo expostos aos impactos de um ou mais riscos naturais: terremotos, ciclones, secas, inundações e elevação do nível do mar.

Gráfico 5 - Variável: Índice Mundial de Riscos a Desastres Naturais - Exposição dos Países, categorizados em Regiões, demonstra o **EXP** médio de 14,514, para um intervalo de confiança de 95% e desvio padrão de 7,773. É importante observar que os países com menor exposição, estão abaixo da média, ou no quartil inferior, portanto, contrário aos índices de suscetibilidade, vulnerabilidade e de adaptação. Entre os extremos: mínimo (2,93) e máximo (52,46), há: Arábia Saudita (**Western Asia)** e Filipinas (**South-Eastern Asia**), respectivamente.

Gráfico 5: Variável: Índice Mundial de Riscos a Desastres Naturais - Exposição dos Países, categorizados em Regiões



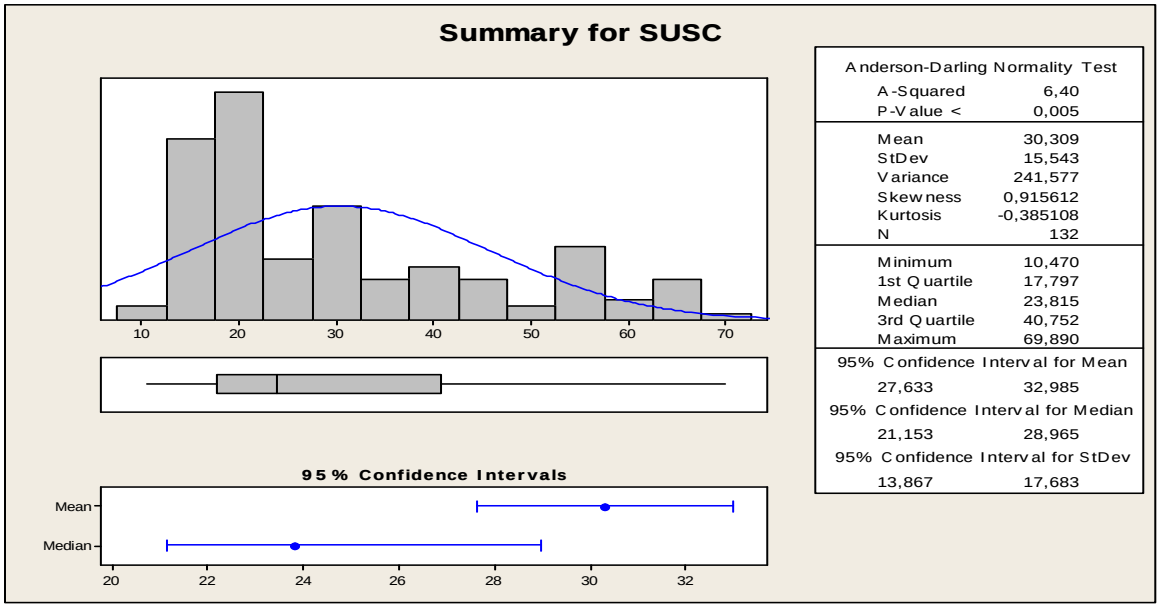
Fonte: Autor – Dados Guia GPS –

SUSC - ÍNDICE MUNDIAL DE RISCO A DESASTRES NATURAIS – SUSCETIBILIDADE

Suscetibilidade refere-se à probabilidade que o país tem a sofrer danos, perdas face a desastres naturais e ou eventos extremos. Portanto, avalia as características estruturais, tais como: infraestrutura pública, condições de moradia, capacidade econômica, distribuição de renda, consequentemente, a índices de pobreza e de nutrição da população.

Gráfico 6 - Variável: Índice Mundial de Riscos a Desastres Naturais - Vulnerabilidade dos Países, categorizados em Regiões, demonstra o **SUSC** médio de 30,309, para um intervalo de confiança de 95% e desvio padrão de 15,543. Entre os extremos: mínimo (10,47) e máximo (69,89), há: Emirados Árabes (**Western Asia)** e Moçambique (**Eastern Africa**), respectivamente. De forma geral, todo o continente Africano demonstra índices superiores à média, exceto pelos países categorizados como **Southern Africa.**

Gráfico 6: Variável: Índice Mundial de Riscos a Desastres Naturais - Suscetibilidade dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

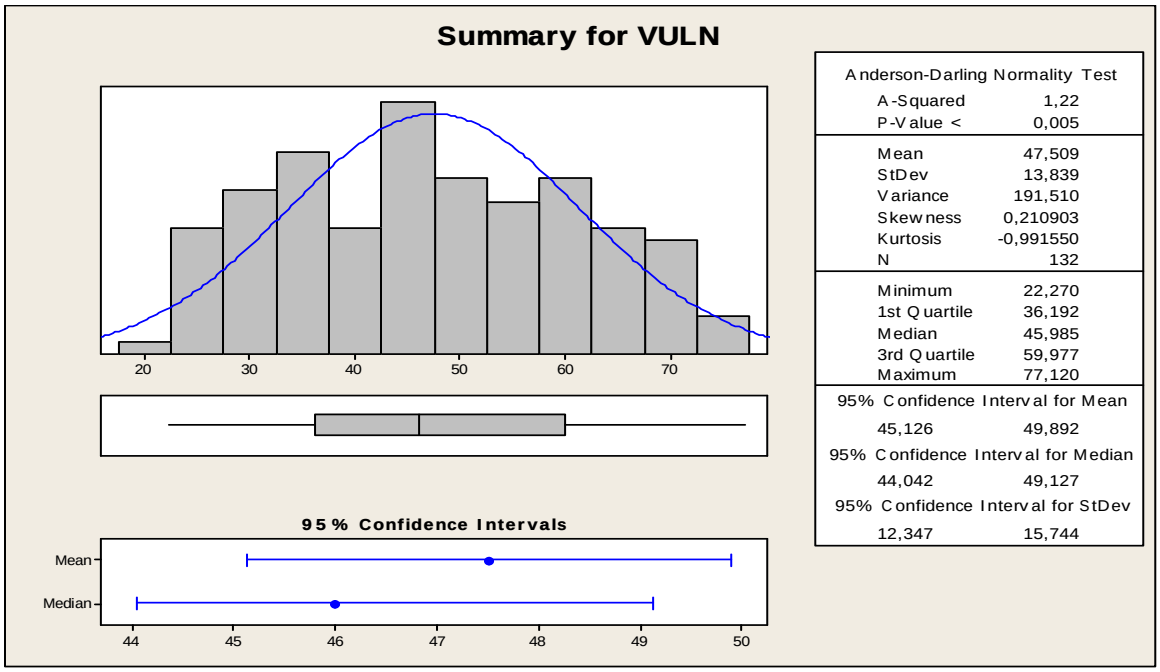
VULN - ÍNDICE MUNDIAL DE RISCO A DESASTRES NATURAIS - VULNERABILIDADE

Vulnerabilidade é compreendida como uma variável da análise da suscetibilidade dos países. Contudo, avalia a falta de capacidade de adaptação e da falta de capacidade de recuperação a situações difíceis, devido a fatores sociais, físicos, econômicos e ambientais que tornam as pessoas ou sistemas suscetíveis à impactos dos riscos naturais, os efeitos adversos das alterações climáticas ou outros processos de transformação.

Gráfico 7 - Variável: Índice Mundial de Riscos a Desastres Naturais - Vulnerabilidade dos Países, categorizados em Regiões, demonstra o **VULN** médio de 47,509, para um intervalo de confiança de 95% e desvio padrão de 13,839. A vulnerabilidade medida entre os países é cerca de 17 pontos superiores aos índices de suscetibilidade, indicando que há muitos países suscetíveis aos desastres naturais, mas não tem capacidade suficiente de recuperação. Com isto, além de todas as regiões da África, exceto por **Southern Africa**, acrescenta as

regiões: **Central Asia**, **Western Asia**. Entre os extremos: mínimo (22,27) e máximo (77,12), há: Nova Zelândia (**Australia e New Zealand)** e Niger (**Western Africa**), respectivamente.

Gráfico 7: Variável: Índice Mundial de Riscos a Desastres Naturais - Vulnerabilidade dos Países, categorizados em Regiões



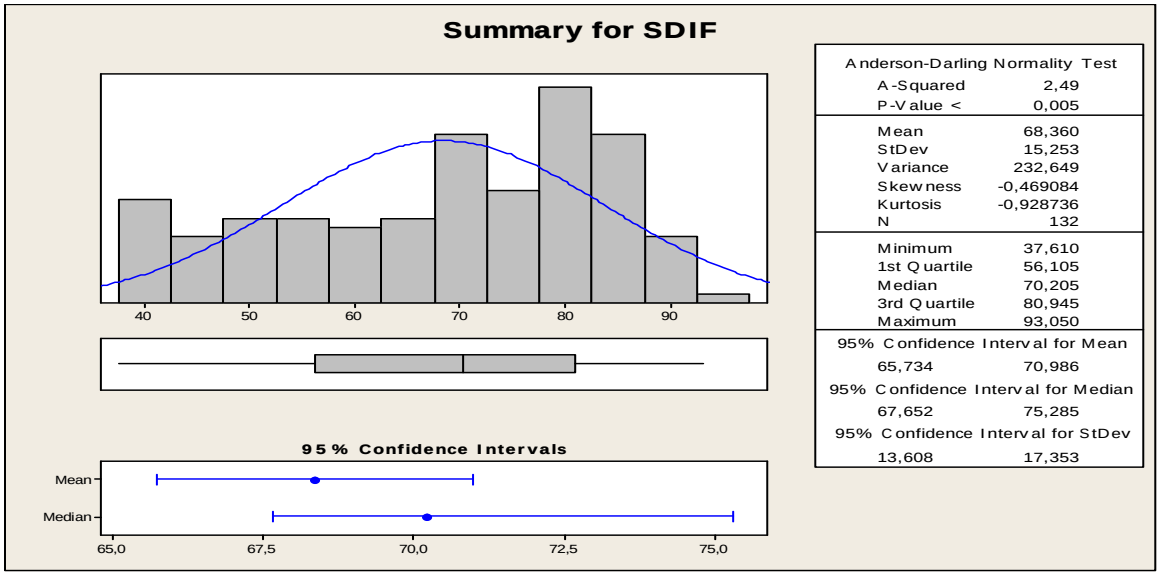
Fonte: Autor – Dados Guia GPS –

SDIF - ÍNDICE MUNDIAL DE RISCO A DESASTRES NATURAIS – ADAPTAÇÃO A SITUAÇÕES DIFICEIS

A capacidade de adaptação às situações difíceis é a habilidade da sociedade em minimizar os impactos negativos oriundos de desastres naturais, às alterações climáticas via ação direta e da disponibilidade de recursos para redução destes danos, prejuízos. De acordo com World Risk Report (2014), os atributos analisados nos países são a boa governança, corrupção dos governos, número adequado de hospitais e médicos por habitante, mecanismos de antecipação e de preparação para desastres naturais e da cobertura de seguros para tal.

Gráfico 8 - Variável: Índice Mundial de Riscos a Desastres Naturais - Vulnerabilidade dos Países, categorizados em Regiões, demonstra o **SDIF** médio de 68,360, para um intervalo de confiança de 95% e desvio padrão de 15,253. Grande parte dos países categorizados em regiões estão próximos a média, exceto pelas regiões: **Australia e New Zealand**, **Northern America**, **Northern Europe** e **Western Europe**, com índices entre 40 a 47. Entre os extremos: mínimo (37,61) e máximo (93,05), há: Áustria **(Western Europe)** e Sudão (**Northern Africa**), respectivamente.

Gráfico 8: Variável: Índice Mundial de Riscos a Desastres Naturais – Adaptação a Situações Difíceis dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

CADAPT - ÍNDICE MUNDIAL DE RISCO A DESASTRES NATURAIS – ADAPTAÇÃO

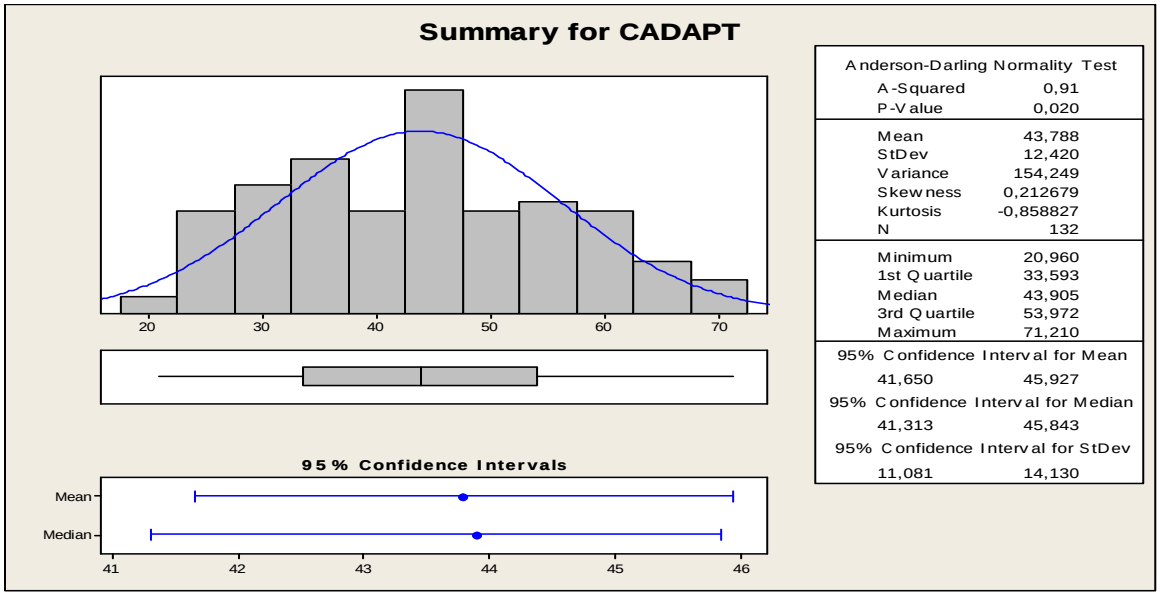
Adaptação, ao contrário da capacidade de enfrentar e sobreviver, é entendida como um processo de longo prazo que inclui mudanças estruturais. Concentra-se em medidas e estratégias que abordam os impactos negativos dos desastres naturais e das mudanças climáticas em longo prazo. Entre estas medidas e estratégias públicas estão os investimentos para educação e pesquisa, com o objetivo de aumentar a alfabetização e a frequência escolar; políticas de proteção ao meio ambiente e da promoção da igualdade de gênero.

Gráfico 9 - Variável: Índice Mundial de Riscos a Desastres Naturais - Adaptação dos Países, categorizados em Regiões, demonstra o **CADAPT** médio de 43,78 para um intervalo de confiança de 95% e desvio padrão de 12,42.

Entre os extremos: mínimo (20,96) e máximo (71,21), há: Austrália (**Australia e New Zealand)** e Mali (**Western Africa**), respectivamente. As regiões: **Central Asia, Eastern Africa, Middle Africa, Northern Africa, Western Africa, South America, Southern Africa, South-Eastern Asia, Southern Asia** tem índices superiores a media, onde a busca pelo ideal é de índices inferiores a 43,78.

Destaca-se que a região IBE tem o índice 40,22, cerca de 4 pontos abaixo da média, mas tem países como Costa Rica, Argentina, Cuba, Espanha, Portugal com índices 25-35, demonstrando que os países com índices bons, alguns comparáveis ao continente Europeu.

Gráfico 9: Variável: Índice Mundial de Riscos a Desastres Naturais – Adaptação dos Países, categorizados em Regiões



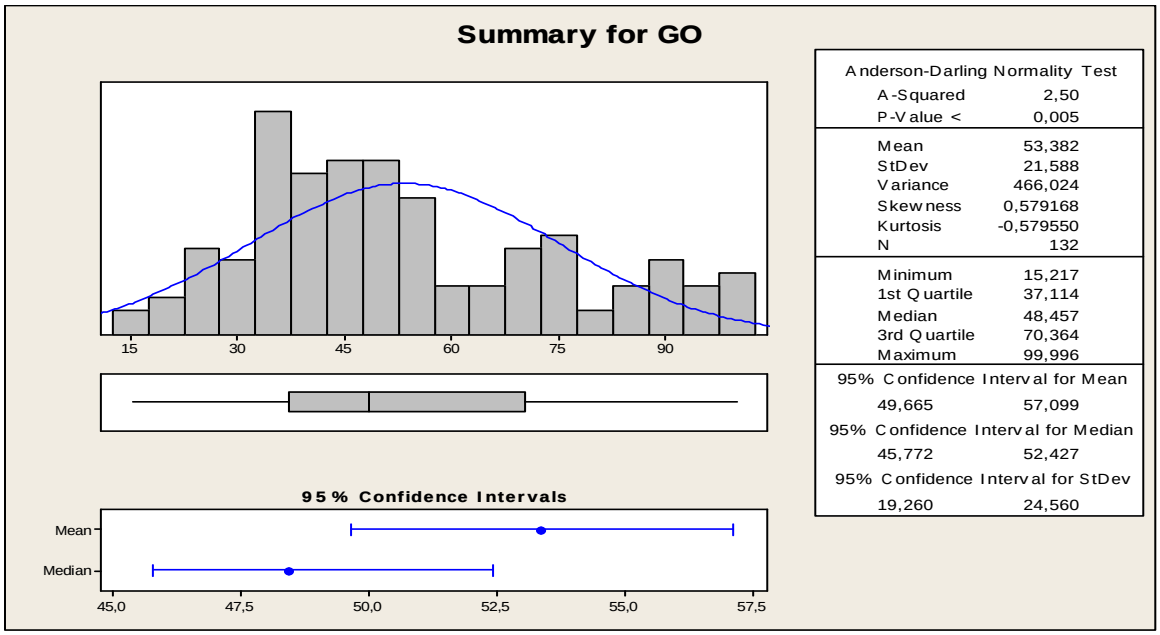
Fonte: Autor – Dados Guia GPS –

GO – GOVERNANÇA

Capacidade de conduzir os processos de formulação, execução e avaliação de políticas públicas integrando instrumentos e mecanismos de gestão, capaz de coibir a corrupção, abusos e prevenção das violações da democracia.

Gráfico 10 - Variável: Governança dos Países, categorizados em Regiões, demonstra o **GO** médio de 53,382, para um intervalo de confiança de 95% e desvio padrão de 21,58. Entre os extremos: mínimo (15,217) e máximo (99,996), há: Sudão (**Northern Africa)** e Finlândia (**Northern Europe**), respectivamente. Comparativamente ao índice de adaptação a situações difíceis, observa-se que é inversamente proporcional a estes resultados, muito embora, haja a avaliação de outros atributos, inclui a governança.

Gráfico 10: Variável: Governança dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

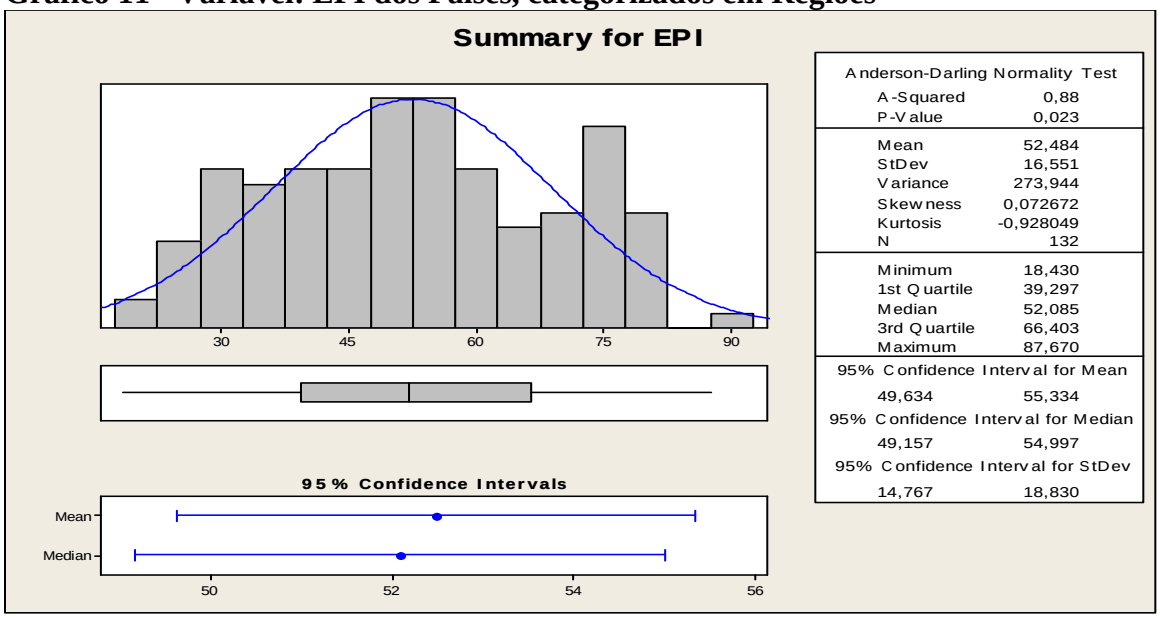
EPI – INDICE DE DESEMPENHO AMBIENTAL

O Índice de Desempenho Ambiental, ou Índice de Performance Ambiental -Environmental Performance Index (EPI) é um método para quantificar e classificar numericamente o desempenho ambiental das políticas de um país. O EPI foi antecedido pelo Índice de Sustentabilidade Ambiental, publicado entre 1999 e 2005. Estes indicadores foram desenvolvidos pela Universidade de Yale em conjunto Universidade de Columbia.

Gráfico 11 - Variável: EPI dos Países, categorizados em Regiões, demonstra o **EPI** médio de 52,484, para um intervalo de confiança de 95% e desvio padrão de 16,551.

Entre os extremos: mínimo (18,43) e máximo (87,67), há: Mali (**Western Africa)** e Suíça (**Western Europe**), respectivamente. Estes extremos estão correlacionados com os resultados de IDH e IPS.

Gráfico 11 - Variável: EPI dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

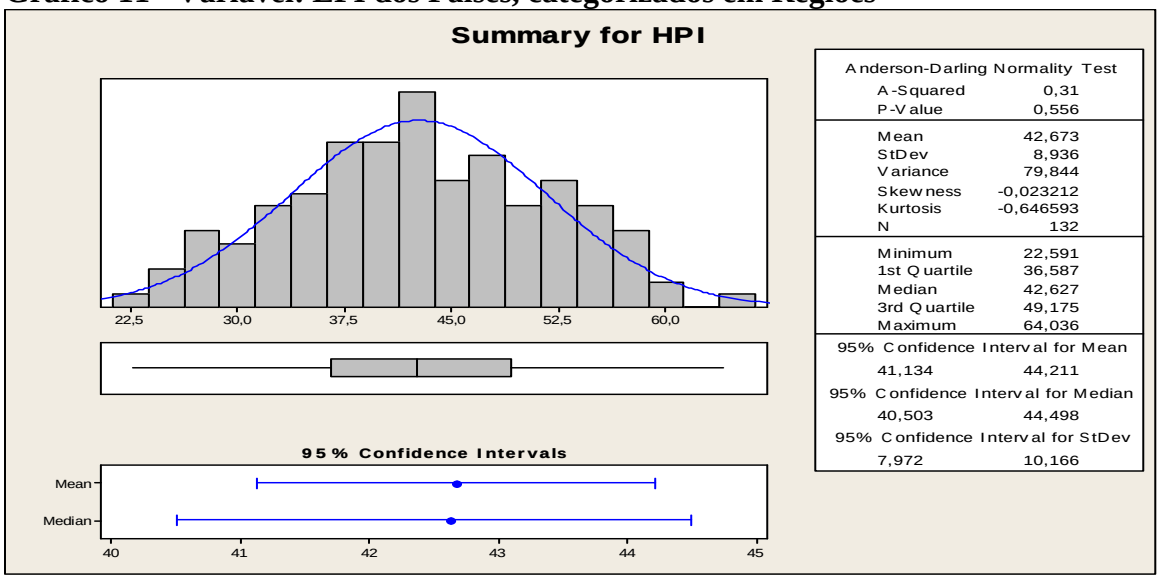
HPI – HAPPY PLANET INDEX

O HPI ou Índice do Planeta Feliz mede a capacidade que cada país tem para proporcionar um bem-estar sustentável aos seus cidadãos. Foi introduzido pela New Economics Foundation, uma organização não governamental ecológica britânica em Julho de 2006. O índice não é uma forma de quantificar quais são os países mais felizes, mas uma forma de medir a eficiência com que uma nação converte os seus recursos naturais em longevidade e felicidade para seus cidadãos. É a associação do nível subjetivo de satisfação, esperança média de via e eco políticas aplicadas per capita.

Gráfico 12 - Variável: HPI dos Países, categorizados em Regiões, demonstra o **HPI** médio de 42,673, para um intervalo de confiança de 95% e desvio padrão de 8,936.

Entre os extremos: mínimo (22,591) e máximo (64,036), há: Botswana (**Southern Africa)** e Costa Rica (**IBE**), respectivamente. Observa-se que países com bons **HPI**, também tem boa pontuação no **IPS**, haja visto que ambos buscam quantificar a sustentabilidade das políticas públicas.

Gráfico 11 - Variável: EPI dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

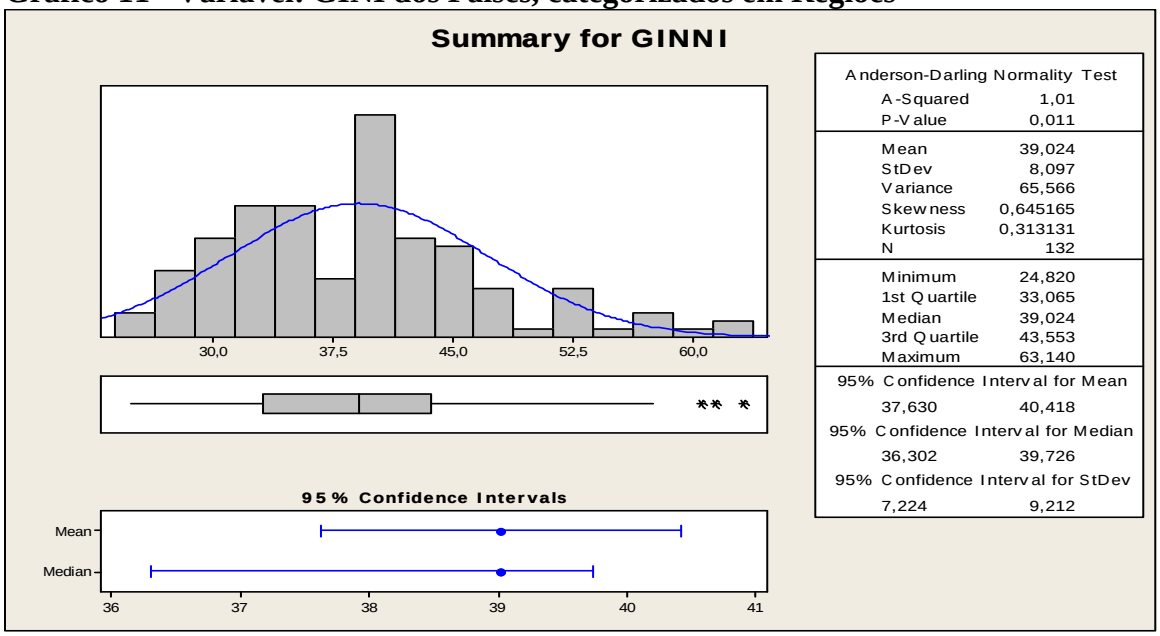
GINI -

O índice de GINI, de acordo com o The World Bank, mede a extensão em que a distribuição de renda, ou, em alguns casos, a despesa de consumo - entre indivíduos ou famílias dentro de uma economia, se desvia de uma distribuição perfeitamente igual. Assim, um índice de GINI igual zero (0) representa a igualdade perfeita, por outro lado, um índice igual a cem (100) indica desigualdade perfeita.

Gráfico 13 - Variável: GINI dos Países, categorizados em Regiões, demonstra o **GINI** médio de 39,024, para um intervalo de confiança de 95% e desvio padrão de 8,097.

Entre os extremos: mínimo (24,82) e máximo (63,14), há: Ucrânia **(Eastern Europe)** e Africa do Sul (**Southern Africa**), respectivamente.

Gráfico 11 - Variável: GINI dos Países, categorizados em Regiões



Fonte: Autor – Dados Guia GPS –

4 RELAÇÃO ENTRE AS VARIÁVEIS QUANTITAVAS

4.1 CORRELAÇÃO LINEAR

Garson (2009) afirma que correlação é uma medida de associação bivariada (força) do grau de relacionamento entre duas variáveis. Para Moore (2007), “A correlação mensura a direção e o grau da relação linear entre duas variáveis quantitativas” (Moore, 2007). Para Las Casas e Guevara (2010) expressa o grau de relação linear existente entre as variáveis, mas não requer que exista uma relação causa e efeito (LAS CASAS e GUEVARA, 2010).

Este coeficiente varia entre os valores -1 e 1. O valor 0 (zero) significa que não há relação linear, o valor 1 indica uma relação linear perfeita e o valor -1 também indica uma relação linear perfeita, mas inversa, ou seja, quando uma das variáveis aumenta a outra diminui. Quanto mais próximo estiver de 1 ou -1, mais forte é a associação linear entre as duas variáveis.

O Quadro 3 abaixo demonstra a Matriz de Correlação e sua significância estatística entre as variáveis do eixo do temático: Resiliência, como segue:

Quadro 3 – Correlação das variáveis do eixo temático: Resiliência

	IPS	IDH	RISCN	EXP	VULN	SUSC	SDIF	CADAPT	GO	EPI	HPI	GINNI
IDH	0,8680											
	0,0000											
RISCN	-0,3730	-0,4410										
	0,0000	0,0000										
EXP	0,0590	-0,0500	0,8670									
	0,5010	0,5680	0,0000									
VULN	-0,9540	-0,9060	0,4940	0,0510								
	0,0000	0,0000	0,0000	0,5630								
SUSC	-0,8380	-0,8720	0,4250	0,0220	0,9150							
	0,0000	0,0000	0,0000	0,8000	0,0000							
SDIF	-0,9260	-0,8120	0,5020	0,0780	0,9380	0,7750						
	0,0000	0,0000	0,0000	0,3720	0,0000	0,0000						
CADAPT	-0,9320	-0,8790	0,4740	0,0450	0,9650	0,8480	0,8970					
	0,0000	0,0000	0,0000	0,6080	0,0000	0,0000	0,0000					
GO	0,8900	0,6970	-0,3560	0,0200	-0,8120	-0,6130	-0,9040	-0,7950				
	0,0000	0,0000	0,0000	0,8220	0,0000	0,0000	0,0000	0,0000				
EPI	0,9170	0,8240	-0,4720	-0,0610	-0,9260	-0,8110	-0,8890	-0,9290	0,8110			
	0,0000	0,0000	0,0000	0,4890	0,0000	0,0000	0,0000	0,0000	0,0000			
HPI	0,2960	0,3330	0,2130	0,3560	-0,2550	-0,3180	-0,0980	-0,2660	0,0890	0,1970		
	0,0010	0,0000	0,0140	0,0000	0,0030	0,0000	0,2630	0,0020	0,3120	0,0240		
GINI	-0,3080	-0,2970	0,2920	0,1370	0,3930	0,4180	0,4110	0,3060	-0,2470	-0,3650	0,0510	
	0,0000	0,0010	0,0010	0,1180	0,0000	0,0000	0,0000	0,0000	0,0040	0,0000	0,5620	

Fonte: Preparado pela autora.

O resultado da matriz de correlação entre as variáveis demonstra que há três índices: **GINI, HPI e Risco Mundial a Riscos**, que não têm forte correlação as demais variáveis do eixo temático: Resiliência.

4.2 GRÁFICOS DE DISPERSÃO E LINHAS DE TENDÊNCIA

As variáveis com coeficiente de correlação superior a (+/-) 0,8 e pvalue < 0,0005, portanto, com alta correlação entre as variáveis, foram selecionadas para a demonstração dos gráficos de dispersão e de linha de tendência, conforme demonstrados abaixo:

Gráfico 12

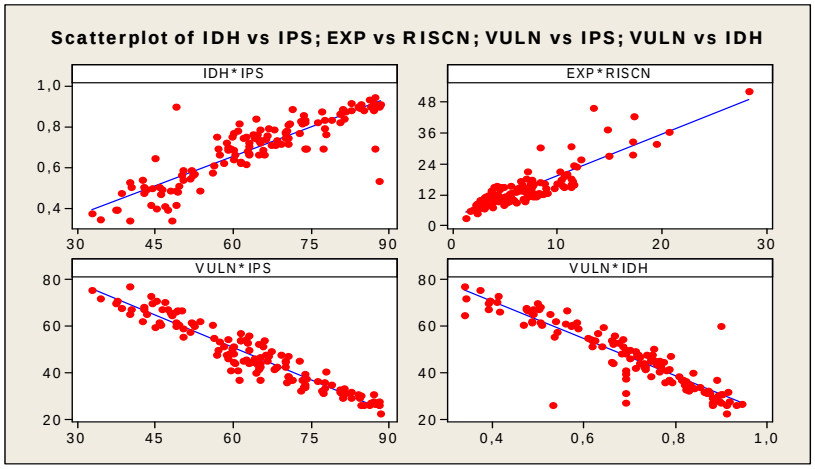


Gráfico 13

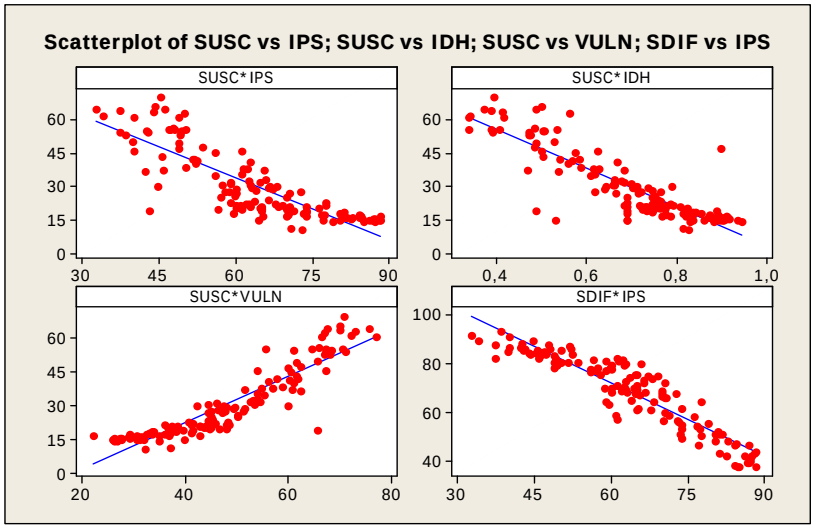


Gráfico 14

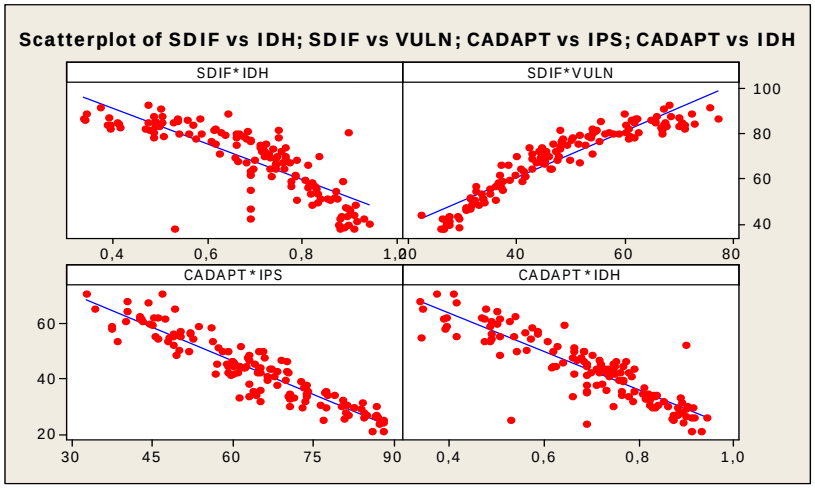


Gráfico 15

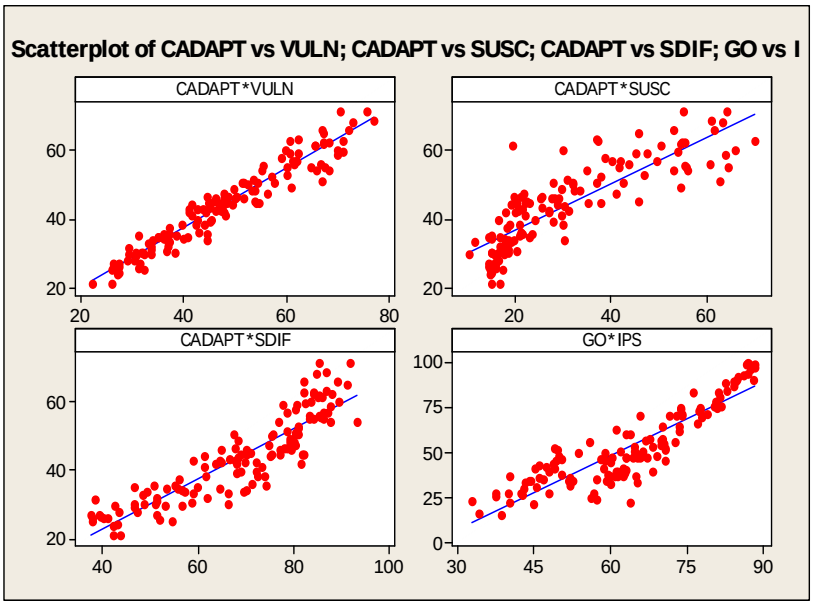


Gráfico 16

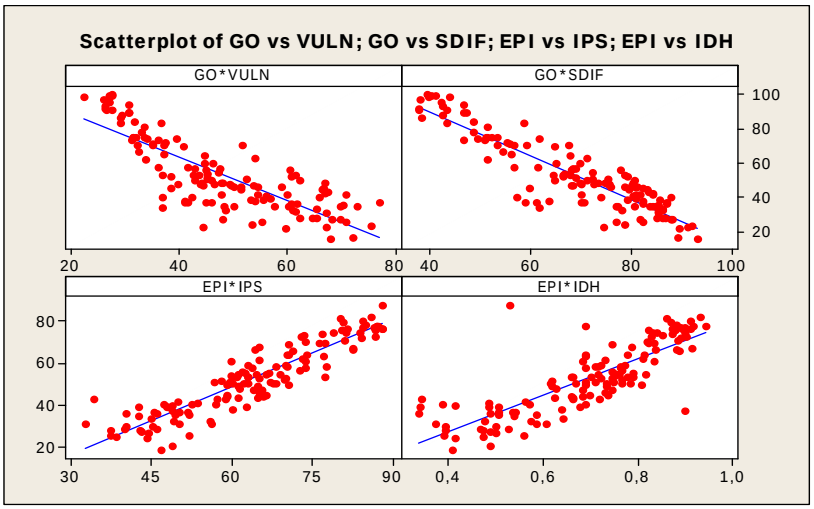


Gráfico 17

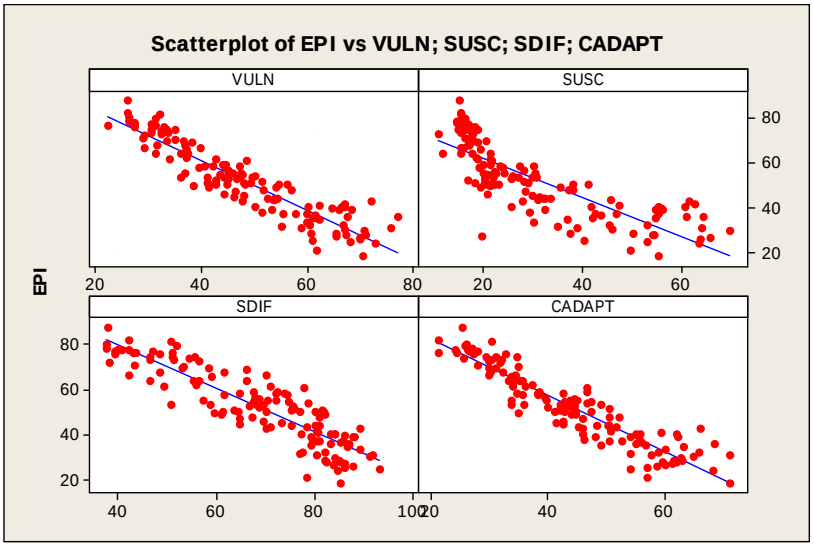
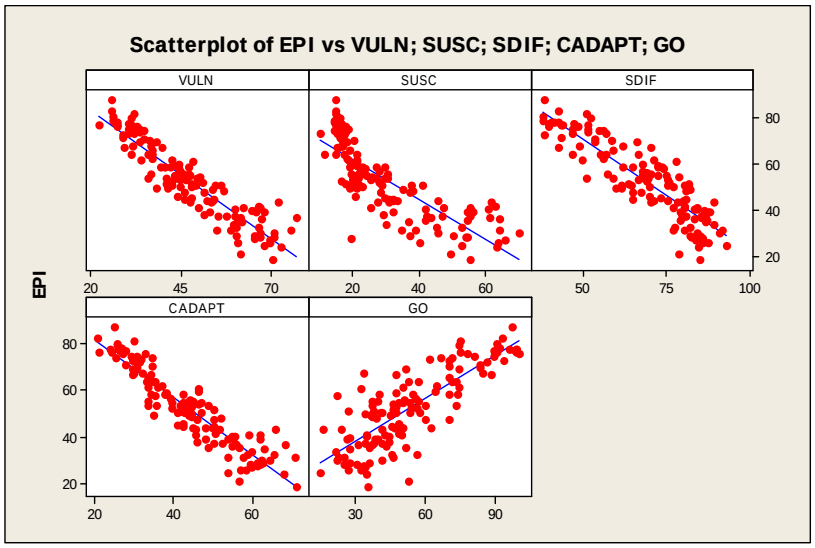


Gráfico 18

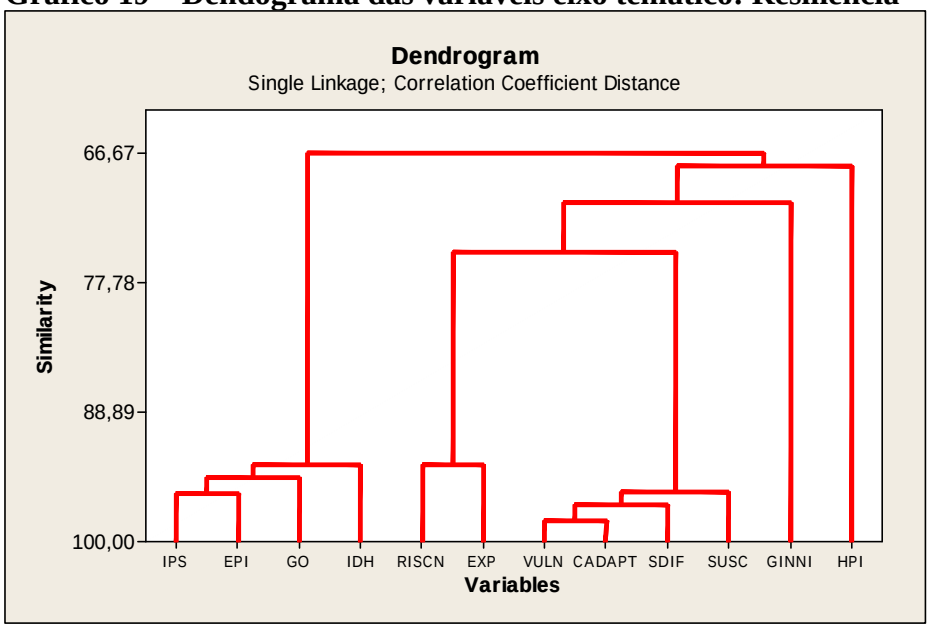


Os gráficos de dispersão com as respectivas linhas de tendência demonstram que as variáveis que tem índices superiores > 70%, tais como: Governança (GO), EPI, IPS e IDH tem correlação positiva, entretanto, quando combinados com os desdobros do índice de risco mundial a desastres naturais: SUSC, VULN, SDIF, CADAPT, há uma correlação negativa. Isto faz sentido, uma vez que estes índices devem ser menores a uma escala de 10% para avaliação de uma boa condição dos países aos riscos a desastres naturais.

4.3 DENDOGRAMA

O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação. O Gráfico 19 demonstra o agrupamento em um único cluster, resultando um grau de similaridade de 66,67%.

Gráfico 19 – Dendograma das variáveis eixo temático: Resiliência



Fonte: Preparado pela autora.

De acordo o quadro 4, abaixo, observa-se que há pelo menos um agrupamento por 5 clusters que resultariam em 93,3553% de similaridades e que 10 clusters resultariam em 96,9039%.

Estas variáveis são: IPS, EPI, GO IDH, RISCN e EXP seguidos por VULN, CADAPT, SDIF e SUSC. A utilização de análise de cluster é para a classificação de variáveis em grupos que em um primeiro momento não são conhecidas. O agrupamento reduz o número de variáveis.

Quadro 4 – Análise de Cluster

Step	No. clusters	Similarity Level	Distance	Clusters Joined		New Cluster	In new Cluster
1	11	98,2739	0,034521	5	8	5	2
2	10	96,9039	0,061922	5	7	5	3
3	9	95,8701	0,082598	1	10	1	2
4	8	95,7454	0,085092	5	6	5	4
5	7	94,4907	0,110186	1	9	1	3
6	6	93,3759	0,132483	1	2	1	4
7	5	93,3553	0,132894	3	4	3	2
8	4	75,1065	0,497871	3	5	3	6
9	3	70,8806	0,582388	3	12	3	7
10	2	67,8002	0,643996	3	11	3	8
11	1	66,6672	0,666657	1	3	1	12

Fonte: Preparado pela autora.

5 REGRESSÃO LINEAR MULTIPLA

O resultado das correlações entre as variáveis e do cluster para o eixo temático resiliência, conclui-se que as variáveis GINI e HPI não apresentam similaridade, portanto, excluídas do procedimento de regressão linear múltipla, ou para construção de modelo capaz de explicar o comportamento do eixo temático resiliência que será representado pela variável: vulnerabilidade (VULN). A regressão das variáveis preditoras: IPS, IDH, RISCN, EXP, SUSC, SDIF, CADAPT, GO e EPI para a vulnerabilidade (VULN) resultou na seguinte equação:

Equação 1: **VULN** = 15,9 - 0,160**IPS** - 2,90**IDH** + 0,289**RISCN** - 0,109**EXP** + 0,229**SUSC** + 0,286**SDIF** + 0,332**CADAPT** + 0,0449**GO** - 0,0013**EPI** =>>para um R2 de 99% e R2 ajustado de 98,9%.

O quadro 5 abaixo, demonstra que há variáveis, muito embora, com alta correlação, apresentam baixa significância estatística como IDH, RISCN, EXP.

Quadro 5 – Regressão Linear Multipla para eixo temático Resiliência

Predictor	Coef	SE Coef	T-Value	P-Value	Observações
Constant	15,936000	4,732000	3,370000	0,001000	
IPS	-0,160440	0,043340	-3,700000	0,000000	
IDH	-2,904000	1,992000	-1,460000	0,147000	Excluída p>0,5
RISCN	0,289400	0,146200	1,980000	0,050000	Excluída p>0,5
EXP	-0,109480	0,066510	-1,650000	0,102000	Excluída p>0,5
SUSC	0,229330	0,020960	10,940000	0,000000	
SDIF	0,286130	0,030730	9,310000	0,000000	
CADAPT	0,331820	0,037300	8,900000	0,000000	
GO	0,044860	0,018410	2,440000	0,016000	
EPI	-0,001330	0,022840	-0,060000	0,954000	Excluída p>0,5

Fonte: preparada pela autora.

A exclusão das variáveis com *p-value* igual ou superior a 0,05 resultou na equação 2 de regressão múltipla linear e os respectivos t e p-value para cada variável:

Equação 2: **VULN** = 8,79-0,145**IPS**+0,252**SUSC**+0,317**SDIF**+0,368**CADAPT**+0,0466**GO** =>> para um **R2** de 98,9% e **R2 ajustado** de 98,9%.

Quadro 5 – Regressão Linear Multipla para eixo temático Resiliência Ajustado ao P-value

Predictor	Coef	SE Coef	*T-value*	*P-value*
Constant	8,79400	3,65300	2,41000	0,01800
IPS	-0,14507	0,03938	-3,68000	0,00000
SUSC	0,25247	0,01894	13,33000	0,00000
SDIF	0,31736	0,02784	11,40000	0,00000
CADAPT	0,36810	0,03181	11,57000	0,00000
GO	0,04659	0,01860	2,50000	0,01400

Fonte: preparada pela autora.

O quadro 6 de observações não usuais demonstra a casos com resíduos padronizados superiores a 2, consequentemente, sugere que estes casos não seguem a equação de regressão linear múltipla proposta na equação 2. Para estas observações não usuais é necessária a investigação do registro e coleta dos dados ou a substituição dos dados por aqueles constantes na coluna FIT, que tem por objetivo ajustar estes dados a equação da regressão linear múltipla.

Quadro 6 – Observações não usuais

Obs	IPS	VULN	Fit	SE Fit	Residual	St Resid
36 – Equador	68,2	42,23	45,728	0,374	-3,498	-2,50R
47- Guinea	37,4	70,94	66,212	0,391	4,728	3,39R
63 – Kenya	50,2	55,51	64,915	0,279	-9,405	-6,62R
78 – Mauritania	43,1	65,51	58,795	0,673	6,715	5,24RX
89 – Nova Zelândia	88,2	22,27	26,502	0,35	-4,232	-3,01R
91 – Nigeria	40,1	77,12	72,886	0,345	4,234	3,01R
132 – Zambia	49,9	66,95	63,834	0,42	3,116	2,25R

Fonte: preparada pela autora.

A regressão com a utilização do *stepwise* demonstra a contribuição a cada inserção de cada uma das 5 variáveis para a demonstração final da equação 2 de regressão linear múltipla. Abaixo, demonstra-se a evolução da equação de regressão linear múltipla, onde observamos que a partir da inclusão de 3 variáveis no modelo, há um aumento significativo do poder de explicação com um R2=98,83%. As demais variáveis contribuem com 0,12% de aumento no R2.

1 variável equação de regressão:
VULN = 0,402+1,076**CADAPT** com R2=93,21%
2 variáveis na equação de regressão:
VULN = 5,3306+0,752**CADAPT**+0,305SUSC com R2=96,53%
3 variáveis na equação de regressão:
VULN = 1,2073+0,425**CADAPT**+0,2890**SUSC**+0,312S**DIF** com R2=98,83%
4 variáveis na equação de regressão:
VULN=8,54770+0386**CADAPT**+0,2780**SUSC**+0,277S**DIF**-0,0840**IPS** com R2=98,89%
5 variáveis na equação de regressão:
VULN = 8,79-0,145**IPS**+0,252**SUSC**+0,317**SDIF**+0,368**CADAPT**+0,0466**GO** com R2=98,95%

O quadro 7 sumariza as equações descritas anteriormente, entretanto é importante lembrar que o stepwise para as variáveis CADAPT, SUSC, SDIF, IPS, GO demonstra a contribuição adicional de cada variável inserida no modelo.

Quadro 7 – Regressão Setwise das variáveis:CADAPT, SUSC, SDIF, IPS, GO

Step	1 Variável	2 Variáveis	3 Variáveis	4 Variáveis	5 Variáveis
Constant	**0,40220**	**5,33060**	**-1,20730**	**8,54770**	**8,79410**
CADAPT	**1,07600**	**0,75200**	**0,42500**	**0,38600**	**0,36800**
T-Value	42,26000	21,84000	14,79000	12,21000	11,57000
P-Value	0,00000	0,00000	0,00000	0,00000	0,00000
SUSC		**0,30500**	**0,28900**	**0,27800**	**0,25200**
T-Value		11,10000	18,03000	17,11000	13,33000
P-Value		0,00000	0,00000	0,00000	0,00000
SDIF			**0,31200**	**0,27700**	**0,31700**
T-Value			15,89000	11,93000	11,40000
P-Value			0,00000	0,00000	0,00000
IPS				**-0,08400**	**-0,14500**
T-Value				-2,66000	-3,68000
P-Value				0,00900	0,00000
GO					**0,04700**
T-Value					2,50000
P-Value					0,01400
S	**3,62**	**2,6**	**1,51**	**1,48**	**1,45**
R-Sq	**93,21**	**96,53**	**98,83**	**98,89**	**98,95**
R-Sq(adj)	**93,16**	**96,48**	**98,81**	**98,86**	**98,9**

Mallows Cp	**683,7**	**289**	**15,6**	**10,3**	**6**

Fonte: preparada pela autora.

O *stepwise* também é utilizado para determinar qual modelo e variáveis tem melhor poder de explicação, utilizando a técnica de entrada e saída de varíaveis segundo um alpha < 0,15. Isto significa que todas as variáveis do modelo poderiam ser agregadas e o stepwise resultaria com a sugestão, conforme demonstrado no quadro 8 a seguir.

Quadro 8 – Regressão Stepwise das 9 variáveis do eixo temático resiliência

Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Constant	**0,4022**	**5,3306**	**-1,2073**	**8,5477**	**8,7941**	**11,6910**	**13,6400**	**15,8400**
CADAPT	**1,0760**	**0,7520**	**0,4250**	**0,3860**	**0,3680**	**0,3550**	**0,3440**	**0,3330**
T-Value	42,2600	21,8400	14,7900	12,2100	11,5700	10,9300	10,3400	9,8200
P-Value	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
SUSC		**0,3050**	**0,2890**	**0,2780**	**0,2520**	**0,2480**	**0,2410**	**0,2290**
T-Value		11,1000	18,0300	17,1100	13,3300	13,0500	12,3400	11,0200
P-Value		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
SDIF			**0,3120**	**0,2770**	**0,3170**	**0,3010**	**0,2860**	**0,2860**
T-Value			15,8900	11,9300	11,4000	10,2200	9,3200	9,3600
P-Value			0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
IPS				**-0,0840**	**-0,1450**	**-0,1690**	**-0,1720**	**-0,1610**
T-Value				-2,6600	-3,6800	-4,0500	-4,1400	-3,8200
P-Value				0,0090	0,0000	0,0000	0,0000	0,0000
GO					**0,0470**	**0,0470**	**0,0460**	**0,0450**
T-Value					2,5000	2,5300	2,5300	2,4500
P-Value					0,0140	0,0130	0,0130	0,0160
RISCN						**0,0640**	**0,2900**	**0,2900**
T-Value						1,6300	1,9600	1,9900
P-Value						0,1050	0,0520	0,0480
EXP							**-0,1050**	**-0,1100**
T-Value							-1,5800	-1,6500
P-Value							0,1170	0,1010
IDH								**-2,9000**
T-Value								-1,4700
P-Value								0,1450
S	**3,6200**	**2,6000**	**1,5100**	**1,4800**	**1,4500**	**1,4400**	**1,4300**	**1,4200**
R-Sq	**93,2100**	**96,5300**	**98,8300**	**98,8900**	**98,9500**	**98,9700**	**98,9900**	**99,0100**
R-Sq(adj)	**93,1600**	**96,4800**	**98,8100**	**98,8600**	**98,9000**	**98,9200**	**98,9300**	**98,9400**
Mallows C	**705,2000**	**300,0000**	**19,3000**	**13,8000**	**9,3000**	**8,6000**	**8,1000**	**8,0000**

Fonte: preparada pela autora.

5.1 CORRELAÇÃO ENTRE AS VARIÁVEIS DO EIXO TEMÁTICO RESILIÊNCIA COM DADOS NORMALIZADOS

Quadro 9 – Correlação das variáveis do eixo temático: Resiliência – Dados normalizados

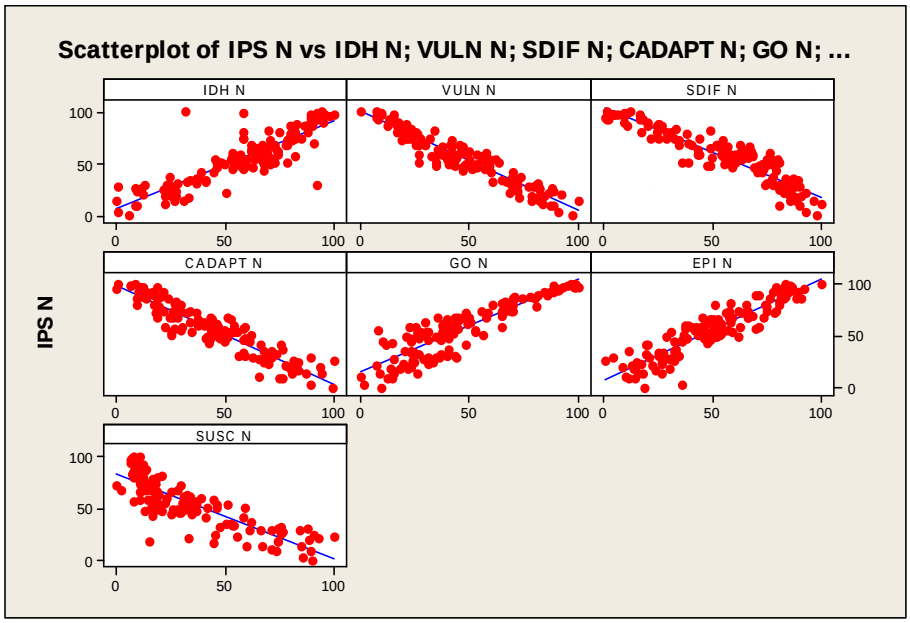
	IPSN	IDHN	RISCNN	EXPN	VULNN	SUSCN	SDIFN	CADAPTN	GON	EPIN	HPIN
IDHN	0,868										
	0,000										
RISCNN	-0,373	-0,441									
	0,000	0,000									
EXPN	0,059	-0,050	0,867								
	0,501	0,568	0,000								
VULNN	-0,954	-0,906	0,494	0,051							
	0,000	0,000	0,000	0,563							
SUSCN	-0,838	-0,872	0,425	0,022	0,915						
	0,000	0,000	0,000	0,800	0,000						
SDIFN	-0,926	-0,812	0,502	0,078	0,938	0,775					
	0,000	0,000	0,000	0,372	0,000	0,000					
CADAPTN	-0,932	-0,879	0,474	0,045	0,965	0,848	0,897				
	0,000	0,000	0,000	0,608	0,000	0,000	0,000				
GON	0,890	0,697	-0,356	0,020	-0,812	-0,613	-0,904	-0,795			
	0,000	0,000	0,000	0,822	0,000	0,000	0,000	0,000			
EPIN	0,917	0,824	-0,472	-0,061	-0,926	-0,811	-0,889	-0,929	0,811		
	0,000	0,000	0,000	0,489	0,000	0,000	0,000	0,000	0,000		
HPIN	0,296	0,333	0,213	0,356	-0,255	-0,318	-0,098	-0,266	0,089	0,197	
	0,001	0,000	0,014	0,000	0,003	0,000	0,263	0,002	0,312	0,024	
GINNIN	-0,308	-0,297	0,292	0,137	0,393	0,418	0,411	0,306	-0,247	-0,365	0,051
	0,000	0,001	0,001	0,118	0,000	0,000	0,000	0,000	0,004	0,000	0,562

Fonte: preparada pela autora.

5.2 GRÁFICOS DE DISPERSÃO E LINHAS DE TENDÊNCIA

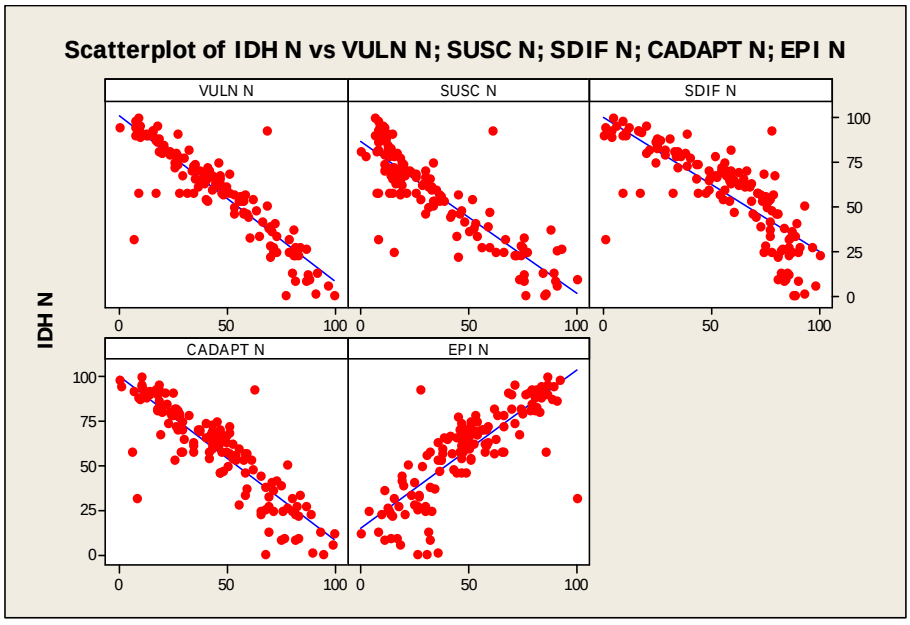
As variáveis com coeficiente de correlação superior a (+/-) 0,8 e pvalue < 0,0005, portanto, com alta correlação entre as variáveis, foram selecionadas para a demonstração dos gráficos de dispersão e de linha de tendência, conforme demonstrados abaixo:

Grafico 20



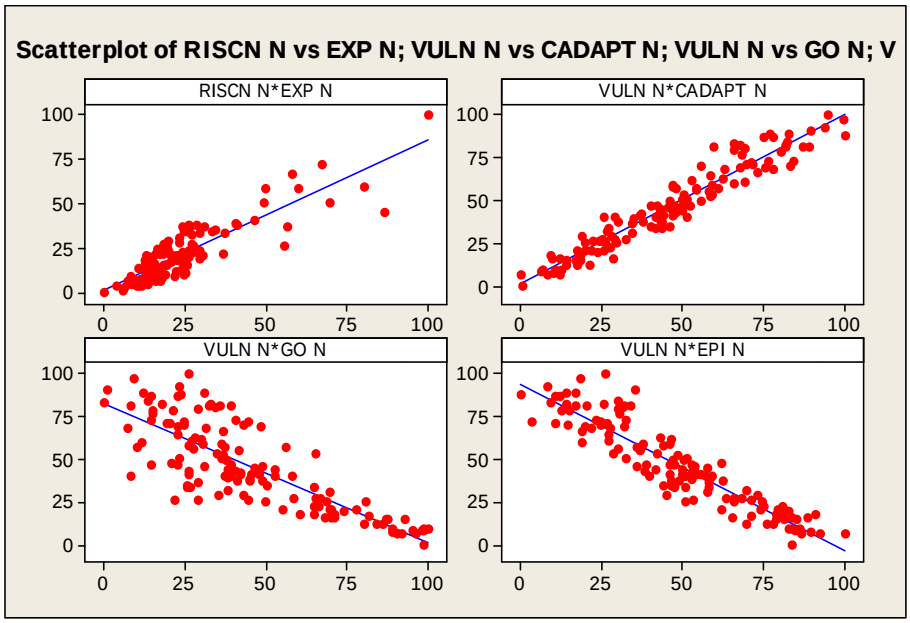
Fonte: preparada pela autora.

Grafico 21



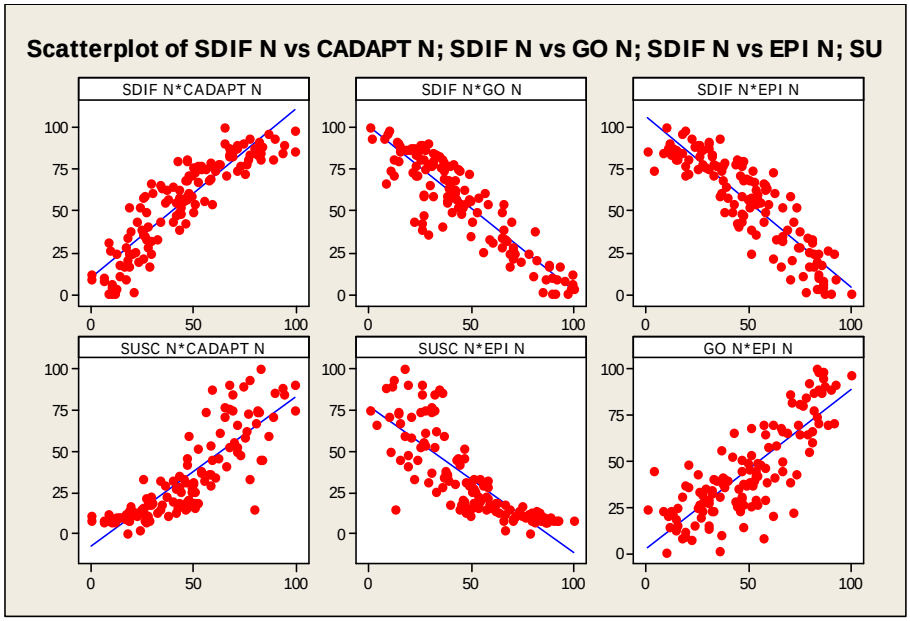
Fonte: preparada pela autora.

Grafico 22



Fonte: preparada pela autora.

Grafico 23 -

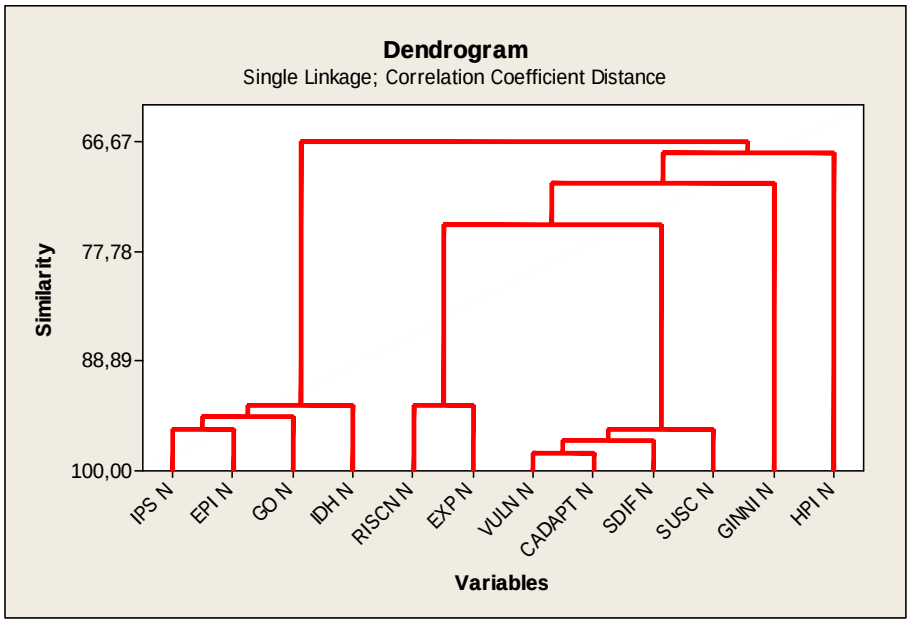


Fonte: preparada pela autora.

5.3 DENDOGRAMA

O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação. O Gráfico 24 demonstra o agrupamento em um único cluster, resultando um grau de similaridade de 66,67%.

Grafico 24 -



Fonte: preparada pela autora.

De acordo o quadro 10, observa-se que há pelo menos um agrupamento por 5 clusters que resultariam em 93,3553% de similaridades e que 11 clusters resultariam em 98,27%. A utilização de análise de cluster é para a classificação de variáveis em grupos que em um primeiro momento não são conhecidas. O agrupamento reduz o número de variáveis.

Quadro 10 – Análise de Cluster

Step	No. clusters	Similarity level	Distance level		Clusters joined	New cluster	No. of obs in new cluster
1	11	98,2739	0,034521	5	8	5	2
2	10	96,9039	0,061922	5	7	5	3
3	9	95,8701	0,082598	1	10	1	2
4	8	95,7454	0,085092	5	6	5	4
5	7	94,4907	0,110186	1	9	1	3
6	6	93,3759	0,132483	1	2	1	4
7	5	93,3553	0,132894	3	4	3	2
8	4	75,1065	0,497871	3	5	3	6
9	3	70,8806	0,582388	3	12	3	7
10	2	67,8002	0,643996	3	11	3	8
11	1	66,6672	0,666657	1	3	1	12

Fonte: preparada pela autora.

5.4 REGRESSÃO LINEAR MULTIPLA

O resultado das correlações entre as variáveis e do cluster para o eixo temático resiliência, conclui-se que as variáveis GINI e HPI não apresentam similaridade, portanto, excluídas do procedimento de stepwise para construção de modelo de regressão linear múltipla capaz de explicar o comportamento do eixo temático resiliência que será representado pela variável: vulnerabilidade (VULN) e as variáveis preditoras: IPS, IDH, RISCN, EXP, SUSC, SDIF, CADAPT, GO e EPI.

O *stepwise* também é utilizado para determinar qual modelo e variáveis tem melhor poder de explicação, utilizando a técnica de entrada e saída de varíaveis segundo um alpha < 0,15. Isto significa que todas as variáveis do modelo poderiam ser agregadas e o stepwise resultaria com a sugestão, conforme demonstrado no quadro 11 a seguir.

Quadro 11 – Setpwise 9 variáveis do eixo temático resiliência --

Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Constant	**1,241**	**3,677**	**0,362**	**9,083**	**8,748**	**11,039**	**12,710**	**14,940**
CADAPT N	**0,986**	**0,689**	**0,389**	**0,354**	**0,337**	**0,326**	**0,315**	**0,305**
T-Value	42,260	21,840	14,790	12,210	11,570	10,930	10,340	9,820
P-Value	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
SUSC N		**0,331**	**0,314**	**0,301**	**0,274**	**0,269**	**0,261**	**0,249**
T-Value		11,100	18,030	17,110	13,330	13,050	12,340	11,020
P-Value		0,000	0,000	0,000	0,000	0,000	0,000	0,000
SDIF N			**0,316**	**0,280**	**0,321**	**0,304**	**0,289**	**0,289**
T-Value			15,890	11,930	11,400	10,220	9,320	9,360
P-Value			0,000	0,000	0,000	0,000	0,000	0,000
IPS N				**-0,085**	**-0,147**	**-0,172**	**-0,175**	**-0,163**
T-Value				-2,660	-3,680	-4,050	-4,140	-3,820
P-Value				0,009	0,000	0,000	0,000	0,000
GO N					**0,072**	**0,072**	**0,072**	**0,069**
T-Value					2,500	2,530	2,530	2,450
P-Value					0,014	0,013	0,013	0,016
RISCN N						**0,031**	**0,141**	**0,143**
T-Value						1,630	1,960	1,990
P-Value						0,105	0,052	0,048
EXP N							**-0,095**	**-0,099**
T-Value							-1,580	-1,650
P-Value							0,117	0,101
IDH N								**-0,032**
T-Value								-1,470
P-Value								0,145
S	**6,600**	**4,740**	**2,760**	**2,690**	**2,640**	**2,620**	**2,610**	**2,600**
R-Sq	**93,210**	**96,530**	**98,830**	**98,890**	**98,950**	**98,970**	**98,990**	**99,010**
R-Sq(adj)	**93,160**	**96,480**	**98,810**	**98,860**	**98,900**	**98,920**	**98,930**	**98,940**
Mallows C	**705,200**	**300,000**	**19,300**	**13,800**	**9,300**	**8,600**	**8,100**	**8,000**

Fonte: preparada pela autora.

As variáveis RISCN, EXP N IDH N para o modelo proposto de 8 variáveis resulta um p-value acima de 0,05%, inferindo que tais variáveis não responderão de forma adequada ao modelo de regressão proposto. Além disto, próprio modelo excluiu a variável EPIN.

O Quadro 12 demonstra que o modelo de regressão linear ideal é para 5 variáveis preditoras para a correspondente: vulnerabilidade (VULNN), com a seguinte equação 3:

VULNN= 8,748+0,337**CADAPTN**+0,274**SUSCN**+ 0,321**SDIFN**-0,147**IPSN**+0,072**GON** =>> com R2=98,95%

Quadro 12 – Setpwise 6 variáveis do eixo temático resiliência --

Step	**1**	**2**	**3**	**4**	**5**
Constant	**1,241**	**3,677**	**0,362**	**9,083**	**8,748**
CADAPT N	**0,986**	**0,689**	**0,389**	**0,354**	**0,337**
T-Value	42,260	21,840	14,790	12,210	11,570
P-Value	0,000	0,000	0,000	0,000	0,000
SUSC N		**0,331**	**0,314**	**0,301**	**0,274**
T-Value		11,100	18,030	17,110	13,330
P-Value		0,000	0,000	0,000	0,000
SDIF N			**0,316**	**0,280**	**0,321**
T-Value			15,890	11,930	11,400
P-Value			0,000	0,000	0,000
IPS N				**-0,085**	**-0,147**
T-Value				-2,660	-3,680
P-Value				0,009	0,000
GO N					**0,072**
T-Value					2,500
P-Value					0,014
S	**6,600**	**4,740**	**2,760**	**2,690**	**2,640**
R-Sq	**93,210**	**96,530**	**98,830**	**98,890**	**98,950**
R-Sq(adj)	**93,160**	**96,480**	**98,810**	**98,860**	**98,900**
Mallows Cp	**683,700**	**289,000**	**15,600**	**10,300**	**6,000**

Fonte: preparada pela autora.

6 AMOSTRA

A definição da amostra aleatória simples é o número (n) de indíviduos de uma população, escolhidos de tal forma que cada conjunto de (n) indivíduos tem a mesma chance de constituir a a amostra efetivamente selecionada. Utilizou o Mini tab em sua função **Calc>Random Data> Sample** from columns para obter as amostras de 25, 50 e 100 países demonstrados nos quadros 3, 4 e 5 a seguir cuja população são de 132 países descritos no quadro 13.

Quadro 13 – Amostra para 25 países

PAÍS	REGIÃO	SDIF N 25	GOV N 25	IPS N 25
Burundi	**OTHERS**	90,20563	14,17571	8,501078
Malaysia	**OTHERS**	54,02237	57,91025	67,21783
Niger	**OTHERS**	88,70851	25,81576	13,47951
Ecuador	**AIBER**	65,20563	29,14948	63,89288
Greece	**AVECO**	24,53102	55,29758	73,38246
Belgium	**AVECO**	8,603896	86,04408	89,91733
Sri Lanka	**OTHERS**	73,79149	36,25195	48,72394
Estonia	**AVECO**	24,4228	77,71434	87,49101
Guyana	**OTHERS**	75,50505	34,56365	49,35298
Slovakia	**AVECO**	32,55772	66,15245	83,26743
Tunisia	**OTHERS**	62,95094	38,29809	54,56506
South Africa	**OTHERS**	57,66595	52,83468	54,56506
Canada	**AVECO**	15,94517	92,81882	97,68152
Morocco	**OTHERS**	68,72294	36,4718	45,66858
Togo	**OTHERS**	85,93074	17,98043	18,33214
Tajikistan	**OTHERS**	70,72511	11,62803	42,14594
Bulgaria	**OTHERS**	34,1811	49,52957	67,64917
Austria	**AVECO**	0,00000	90,6356	94,37455
Guinea	**OTHERS**	80,59163	11,93563	8,64486
Venezuela	**AIBER**	66,07143	8,006227	56,03882
Mauritius	**OTHERS**	43,41631	69,40726	73,83178
Mali	**OTHERS**	85,75036	23,43103	25,61107
Jamaica	**OTHERS**	62,33766	46,45301	67,91876
Kuwait	**OTHERS**	51,64141	44,23912	68,40403
Montenegro	**OTHERS**	55,46537	49,02399	61,46657

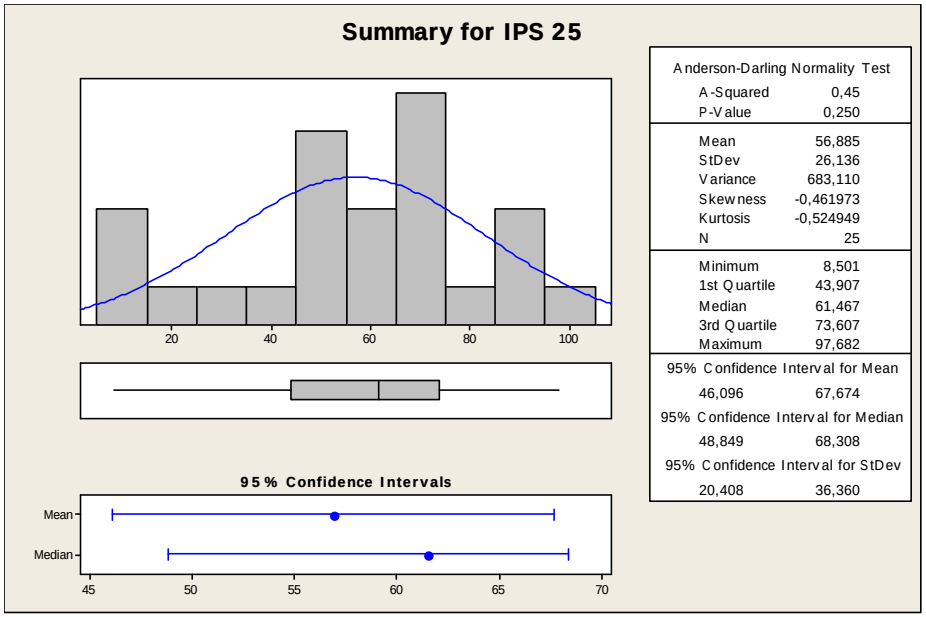
Quadro 14 – Amostra para 50 países

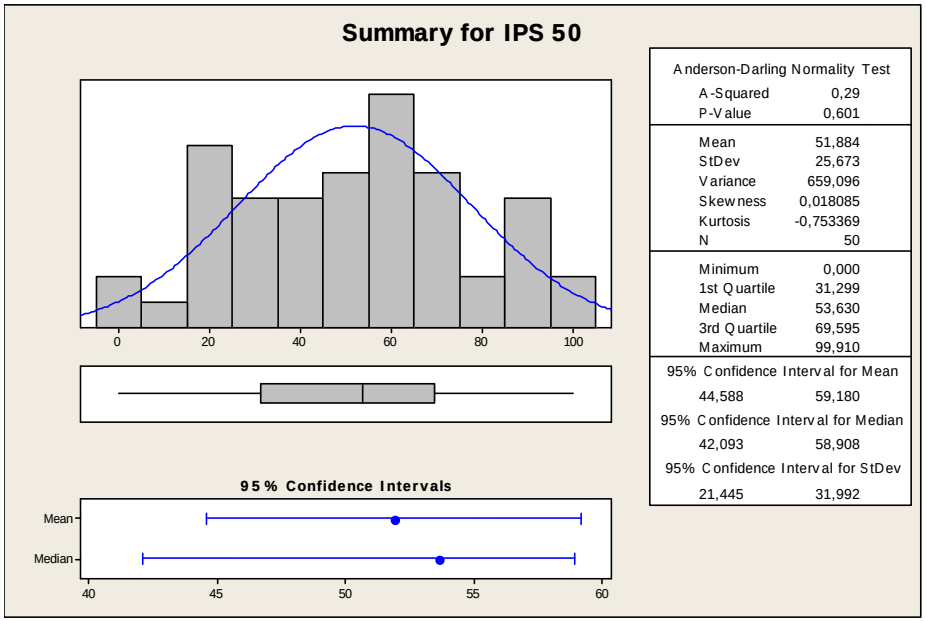
PAÍS	REGIÃO	SDIF N 50	GOV N 50	IPS N 50
Lebanon	**OTHERS**	58,42352	25,86971	49,33501
Hungary	**OTHERS**	28,24675	64,77849	74,17326
Netherlands	**AVECO**	8,189033	95,21018	98,43638
Iran	**OTHERS**	79,31097	14,1909	43,2243
Bosnia and Herzegovina	**OTHERS**	57,77417	39,24162	58,21352
Kenya	**OTHERS**	86,59812	26,85225	31,63192
India	**OTHERS**	77,12843	36,43542	31,70381
Laos	**OTHERS**	83,67605	22,80752	35,60388
Saudi Arabia	**OTHERS**	58,51371	37,99324	57,11718
Slovenia	**AVECO**	24,4228	70,88677	88,156
Kuwait	**OTHERS**	51,64141	44,23912	68,40403
Thailand	**OTHERS**	68,27201	38,25448	58,48311
Greece	**AVECO**	24,53102	55,29758	73,38246
Madagascar	**OTHERS**	83,00866	22,60583	20,99209
Ireland	**AVECO**	16,16162	87,50518	92,46945
Chad	**OTHERS**	97,74531	9,201681	0
Switzerland	**AVECO**	0,559163	96,50587	99,91014
Georgia	**OTHERS**	49,06205	52,5262	56,32638
Korea, Republic of	**AVECO**	16,21573	68,29349	80,12221
Mauritania	**OTHERS**	87,19336	21,38472	18,88929
South Africa	**OTHERS**	57,66595	52,83468	54,56506
Ukraine	**OTHERS**	42,46032	25,40268	58,06973
Namibia	**OTHERS**	60,26335	55,97836	51,3839
Central African Republic	**OTHERS**	92,94733	0,908823	2,821711
Angola	**OTHERS**	85,28139	14,87265	13,17398
Togo	**OTHERS**	85,93074	17,98043	18,33214
Mali	**OTHERS**	85,75036	23,43103	25,61107
Djibouti	**OTHERS**	80,23088	24,3634	23,99353
Mozambique	**OTHERS**	83,94661	30,88328	22,6995
Rwanda	**OTHERS**	74,92785	43,03677	30,30194
Peru	**AIBER**	64,33983	39,19666	60,54996
Ghana	**OTHERS**	72,18615	48,30986	41,98418
Philippines	**OTHERS**	76,46104	37,3383	59,77714
Iraq	**OTHERS**	93,23593	7,322459	21,99856
Germany	**AVECO**	0,21645	88,95797	93,47592
Czech Republic	**AVECO**	23,91775	70,87671	85,92739
Serbia	**OTHERS**	51,51515	43,11438	68,31416
Croatia	**OTHERS**	33,11688	58,51761	73,16679
Senegal	**OTHERS**	77,41703	40,54297	37,59885
Sri Lanka	**OTHERS**	73,79149	36,25195	48,72394
Macedonia	**OTHERS**	48,28644	44,73852	64,21639
Cameroon	**OTHERS**	85,96681	19,01826	23,20273
United States	**AVECO**	19,76912	81,66411	90,16894
Costa Rica	**AIBER**	48,7013	64,56897	81,14666
Tajikistan	**OTHERS**	70,72511	11,62803	42,14594
El Salvador	**AIBER**	68,07359	41,42528	57,69231
Argentina	**AIBER**	39,88095	35,30174	68,27822
Bangladesh	**OTHERS**	87,37374	19,98928	34,93889
Russia	**OTHERS**	38,2215	25,71017	50,66499
Jordan	**OTHERS**	56,24098	42,16	52,6959

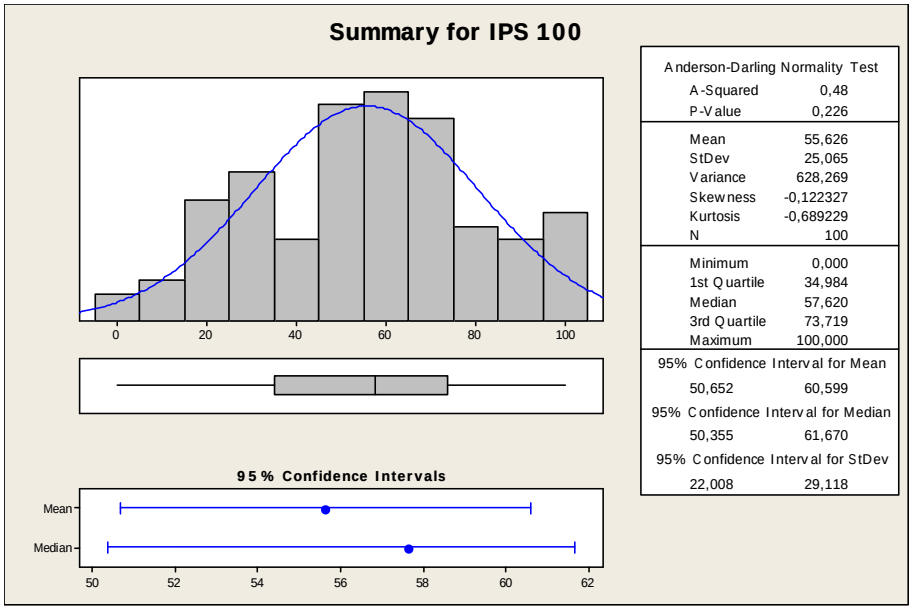
Quadro 15 – Amostra para 100 países

		SDIF N 100	GOV N 100	IPS N 100
El Salvador	**AIBER**	68,07359	41,42528	57,69231
Finland	**AVECO**	3,210678	100	97,60963
Bolivia	**AIBER**	76,80375	30,01834	54,45723
Moldova	**OTHERS**	54,92424	37,34581	49,46082
Norway	**AVECO**	4,401154	98,91135	97,98706
Djibouti	**OTHERS**	80,23088	24,3634	23,99353
Nigeria	**OTHERS**	90,99928	13,40851	18,06254
Morocco	**OTHERS**	68,72294	36,4718	45,66858
Saudi Arabia	**OTHERS**	58,51371	37,99324	57,11718
Uganda	**OTHERS**	90,31385	28,89067	27,22861
Indonesia	**OTHERS**	78,22872	35,96207	47,41193
Ireland	**AVECO**	16,16162	87,50518	92,46945
Trinidad and Tobago	**OTHERS**	56,18687	48,72485	67,00216
Lesotho	**OTHERS**	73,75541	44,18279	29,36736
Yemen	**OTHERS**	96,35642	8,278221	13,71316
Hungary	**OTHERS**	28,24675	64,77849	74,17326
Cameroon	**OTHERS**	85,96681	19,01826	23,20273
Denmark	**AVECO**	3,391053	98,28356	96,96262
Thailand	**OTHERS**	68,27201	38,25448	58,48311
Canada	**AVECO**	15,94517	92,81882	97,68152
Iceland	**AVECO**	9,992785	89,00054	99,69446
Uruguay	**AIBER**	23,79149	68,4302	80,71531
Russia	**OTHERS**	38,2215	25,71017	50,66499
Swaziland	**OTHERS**	77,86797	31,15361	29,24155
Rwanda	**OTHERS**	74,92785	43,03677	30,30194
Greece	**AVECO**	24,53102	55,29758	73,38246
New Zealand	**AVECO**	11,14719	98,71693	100
Benin	**OTHERS**	80,19481	35,75805	29,6729
Mauritius	**OTHERS**	43,41631	69,40726	73,83178
Lebanon	**OTHERS**	58,42352	25,86971	49,33501
Slovenia	**AVECO**	24,4228	70,88677	88,156
Jordan	**OTHERS**	56,24098	42,16	52,6959
Spain	**AIBER**	25,95599	70,133	86,57441
Congo, Republic of	**OTHERS**	87,57215	14,78931	27,65996
Chad	**OTHERS**	97,74531	9,201681	0
Liberia	**OTHERS**	84,7583	22,95955	20,5248
Macedonia	**OTHERS**	48,28644	44,73852	64,21639
United States	**AVECO**	19,76912	81,66411	90,16894
Argentina	**AIBER**	39,88095	35,30174	68,27822
Azerbaijan	**OTHERS**	59,07287	25,87723	53,63048
Burundi	**OTHERS**	90,20563	14,17571	8,501078
Peru	**AIBER**	64,33983	39,19666	60,54996
Brazil	**AIBER**	52,29076	45,42651	67,16391
Netherlands	**AVECO**	8,189033	95,21018	98,43638
Romania	**OTHERS**	42,83911	50,19397	63,12006
Mauritania	**OTHERS**	87,19336	21,38472	18,88929
Guatemala	**AIBER**	77,97619	28,1691	51,7074
Slovakia	**AVECO**	32,55772	66,15245	83,26743
Malaysia	**OTHERS**	54,02237	57,91025	67,21783
Israel	**AVECO**	38,45599	65,15645	69,734
Algeria	**OTHERS**	71,08586	22,42798	47,68152
Montenegro	**OTHERS**	55,46537	49,02399	61,46657
Estonia	**AVECO**	24,4228	77,71434	87,49101
Korea, Republic of	**AVECO**	16,21573	68,29349	80,12221
Mexico	**AIBER**	62,31962	41,79955	60,76564
Costa Rica	**AIBER**	48,7013	64,56897	81,14666
Iraq	**OTHERS**	93,23593	7,322459	21,99856

		SDIF N 100	GOV N 100	IPS N 100
Turkey	**OTHERS**	54,0404	44,69895	57,54853
Malawi	**OTHERS**	82,12482	34,11461	29,09777
Kuwait	**OTHERS**	51,64141	44,23912	68,40403
Zambia	**OTHERS**	77,00216	39,22069	31,05679
Uzbekistan	**OTHERS**	73,61111	10,00424	44,46441
Central African Republic	**OTHERS**	92,94733	0,908823	2,821711
Kyrgyzstan	**OTHERS**	71,21212	23,14217	43,99712
Ghana	**OTHERS**	72,18615	48,30986	41,98418
Togo	**OTHERS**	85,93074	17,98043	18,33214
Lithuania	**OTHERS**	21,19408	69,87021	73,97556
Belarus	**OTHERS**	43,43434	21,84122	58,59094
Japan	**AVECO**	1,208514	84,28548	92,75701
Armenia	**OTHERS**	60,38961	41,21751	58,28541
Iran	**OTHERS**	79,31097	14,1909	43,2243
Paraguay	**AIBER**	74,87374	27,32517	54,00791
Poland	**OTHERS**	28,04834	70,14508	80,5895
Botswana	**OTHERS**	53,62554	65,13659	59,30985
Guyana	**OTHERS**	75,50505	34,56365	49,35298
Honduras	**AIBER**	80,32107	26,11486	51,54565
Guinea	**OTHERS**	80,59163	11,93563	8,64486
Sri Lanka	**OTHERS**	73,79149	36,25195	48,72394
Chile	**AIBER**	37,75253	80,67439	78,54062
Madagascar	**OTHERS**	83,00866	22,60583	20,99209
Bulgaria	**OTHERS**	34,1811	49,52957	67,64917
Egypt	**OTHERS**	72,60101	20,48642	49,19123
Ukraine	**OTHERS**	42,46032	25,40268	58,06973
Senegal	**OTHERS**	77,41703	40,54297	37,59885
Ecuador	**AIBER**	65,20563	29,14948	63,89288
Burkina Faso	**OTHERS**	83,78427	32,0929	26,47376
Pakistan	**OTHERS**	88,56421	14,53862	17,61323
Jamaica	**OTHERS**	62,33766	46,45301	67,91876
Sweden	**AVECO**	5,934343	98,82058	97,91517
Colombia	**AIBER**	67,64069	37,78721	62,25737
Kenya	**OTHERS**	86,59812	26,85225	31,63192
Tajikistan	**OTHERS**	70,72511	11,62803	42,14594
Latvia	**AVECO**	31,70996	66,58802	74,24515
Nepal	**OTHERS**	77,14646	22,56615	34,11215
United Arab Emirates	**OTHERS**	34,09091	64,78364	72,46585
Italy	**AVECO**	30,30303	60,32703	79,6729
Venezuela	**AIBER**	66,07143	8,006227	56,03882
Bosnia and Herzegovina	**OTHERS**	57,77417	39,24162	58,21352
Kazakhstan	**OTHERS**	46,8254	26,44716	48,2926
Croatia	**OTHERS**	33,11688	58,51761	73,16679



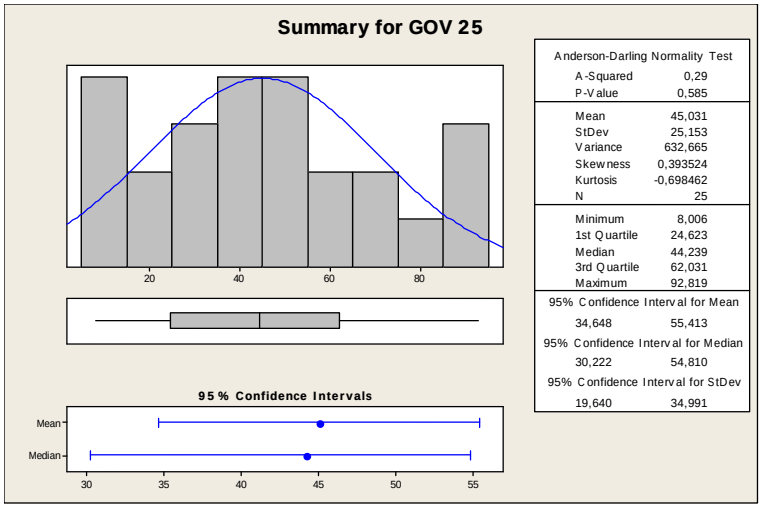


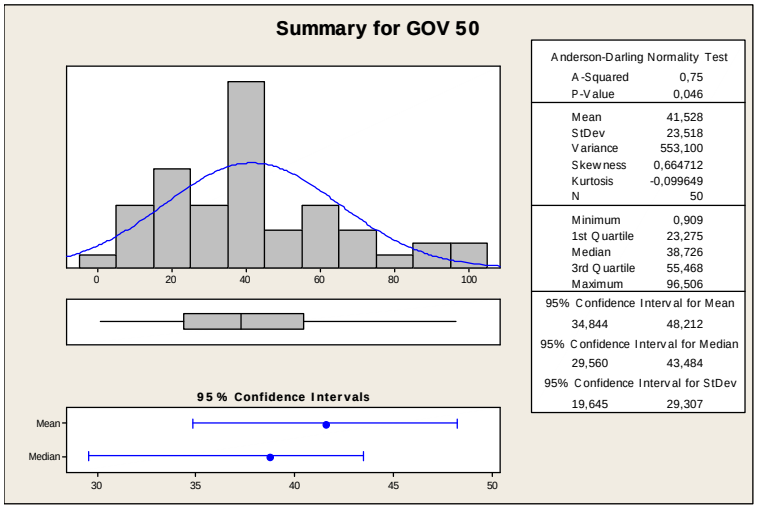


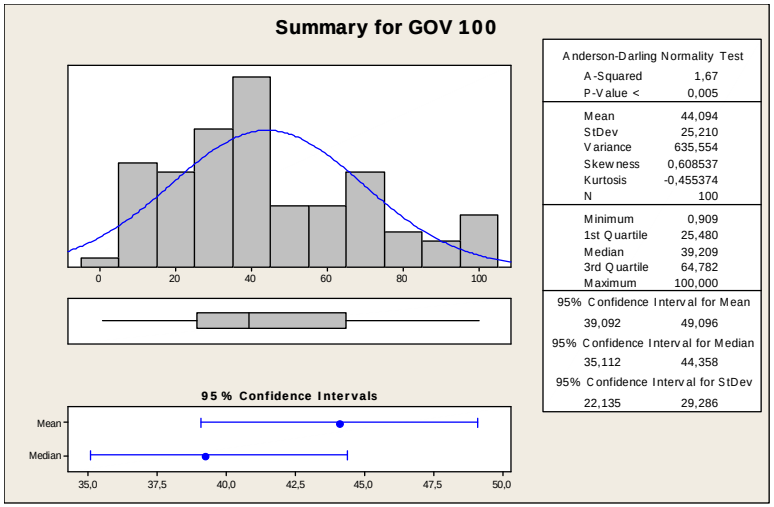
Intervalo de confiança 95% para média IPS com amostra de 25, 50 e 100 países

IPS 25	46,096	67,674
IPS 50	44,588	59,908
IPS 100	50,652	60,599

De acordo com os resultados da média do índice de progresso social – IPS, a amostra adequada (< distância do intervalo de confiança para as amostras), para análise do eixo temático resiliência, seria de 100 países – vide quadro 5 com a composição.



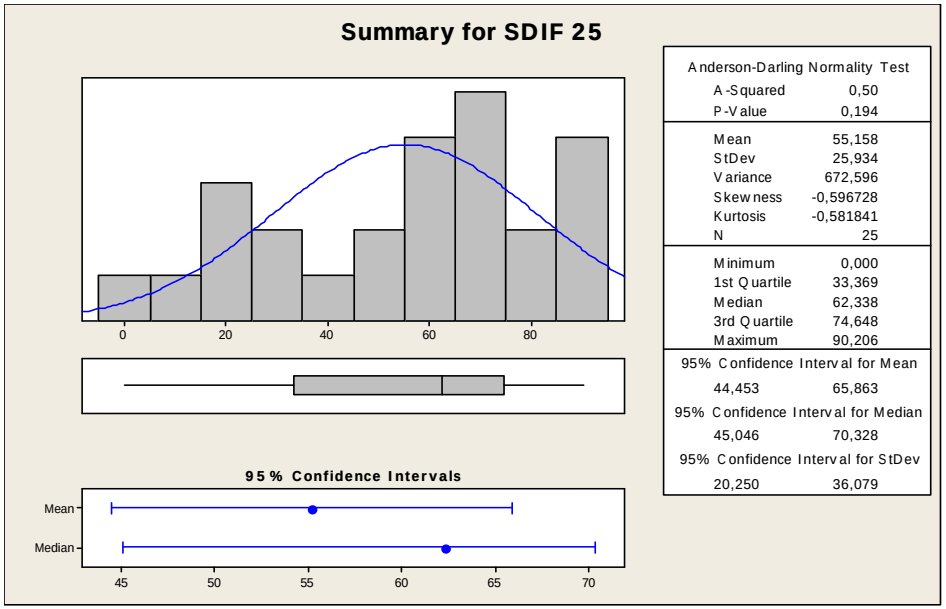


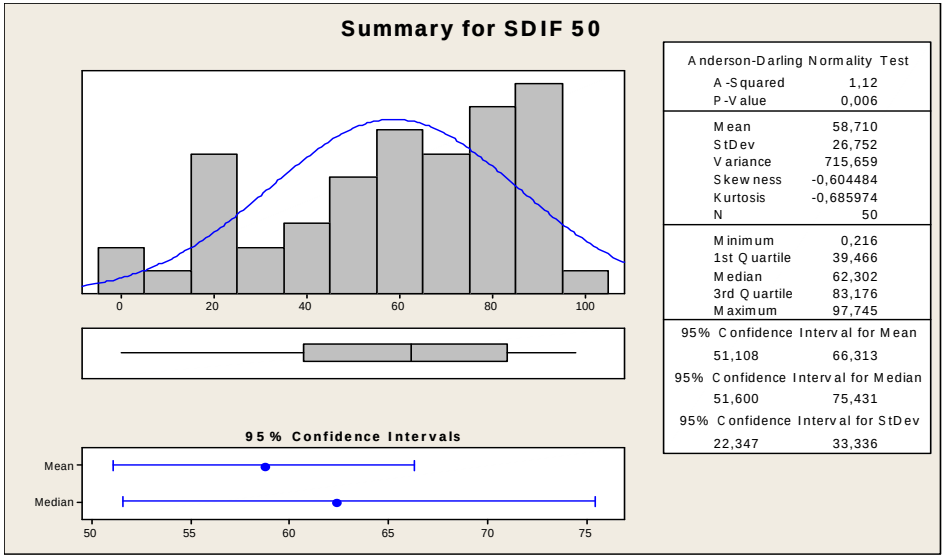


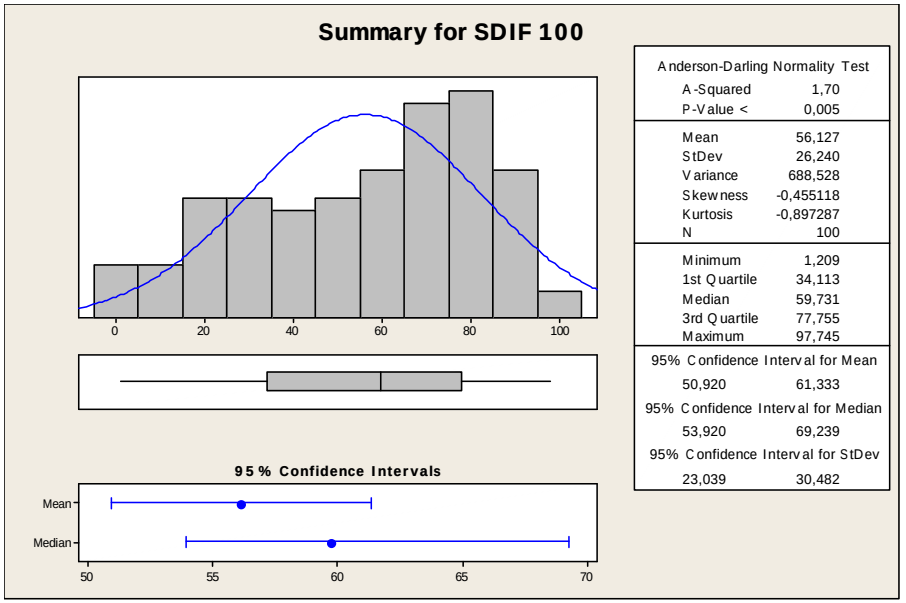
Intervalo de confiança 95% para média GOV com amostra de 25, 50 e 100 países

GOV 25	34,648	55,413
GOV 50	34,844	48,212
GOV 100	39,092	49,096

De acordo com os resultados da média do índice de governança – GO, a amostra adequada (< distância do intervalo de confiança para as amostras), para análise do eixo temático resiliência, seria de 100 países – vide quadro 5 com a composição.







Intervalo de confiança 95% para média SDIF com amostra de 25, 50 e 100 países

SDIF 25	44,453	65,863
SDIF 50	51,108	66,313
SDIF 100	50,920	61,333

De acordo com os resultados da média do índice SDIF, a amostra adequada (< distância do intervalo de confiança para as amostras), para análise do eixo temático resiliência, seria de 100 países – vide quadro 5 com a composição.

6 ANOVA

ANOVA - Análise de variância é semelhante à regressão, na medida em que é usada para investigar e modelar a relação entre uma variável de resposta e uma ou mais variáveis de previsão. Entretanto, a análise de variância difere da regressão no seguinte: as variáveis de previsão são consideradas categóricas e ou qualitativos e não há suposição sobre a natureza da relação, portanto, não incluem coeficientes para as variáveis. ANOVA-análise de variância executa o F-teste para igualdade da amostra ou então, homogeneidade das variâncias (Bartlet e Levene test). Isto porque, se baseiam no pressuposto de que as amostras de diferentes populações têm a mesma variância. Em sequência, apresenta-se os resultados dos testes com a utilização do modelo ANOVA para o eixo temático resiliência, demonstram que para nenhuma das variáveis analisadas o F teste é elevado e tão pouco p-value tem significância estatística. Neste caso, para a amostra aleatória em análise, rejeita-se todas.

One-way ANOVA: SDIF 25; SDIF 50; SDIF 100

```
Source   DF      SS   MS     F      P
Factor    2     294  147  0,21  0,809
Error   172  119374  694
Total   174  119668

S = 26,34   R-Sq = 0,25%   R-Sq(adj) = 0,00%


Individual 95% CIs For Mean Based on Pooled StDev
Level       N   Mean  StDev  -----+---------+---------+---------+----
SDIF 25    25  55,16  25,93  (----------------*----------------)
SDIF 50    50  58,71  26,75            (-----------*-----------)
SDIF 100  100  56,13  26,24          (--------*-------)
                             -----+---------+---------+---------+----
                               48,0      54,0      60,0      66,0
Pooled StDev = 26,34
```

One-way ANOVA: GOV 25; GOV 50; GOV 100

```
Source   DF      SS   MS     F      P
Factor    2     288  144  0,24  0,790
Error   172  105206  612
Total   174  105494

S = 24,73   R-Sq = 0,27%   R-Sq(adj) = 0,00%


Individual 95% CIs For Mean Based on Pooled StDev
Level      N   Mean  StDev   --+---------+---------+---------+-------
GOV 25    25  45,03  25,15    (---------------*---------------)
GOV 50    50  41,53  23,52   (----------*----------)
GOV 100  100  44,09  25,21         (-------*--------)
                             --+---------+---------+---------+-------
                            36,0      42,0      48,0      54,0

Pooled StDev = 24,73
```

One-way ANOVA: IPS 25; IPS 50; IPS 100

Source	DF	SS	MS	F	P
Factor	**2**	**601**	**301**	**0,47**	**0,628**
Error	172	110889	645		
Total	174	111490			

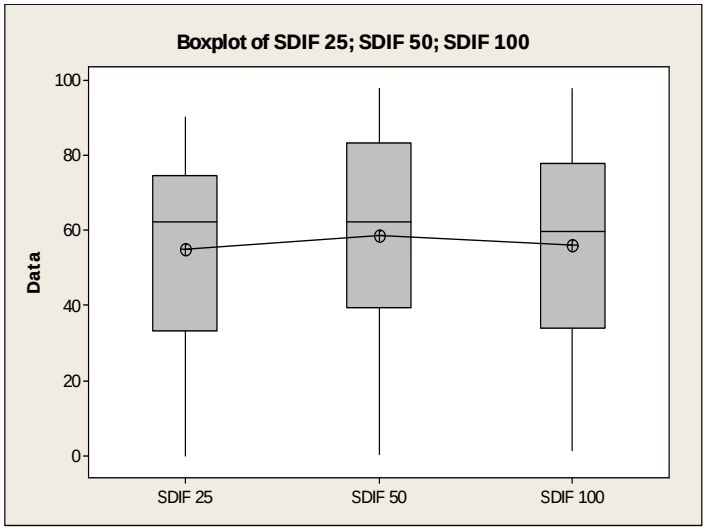
S = 25,39 R-Sq = 0,54% R-Sq(adj) = 0,00%

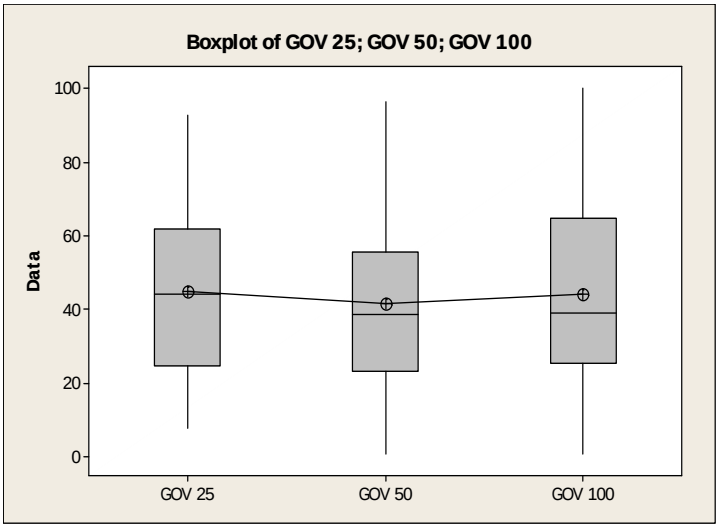
Individual 95% CIs For Mean Based on Pooled StDev

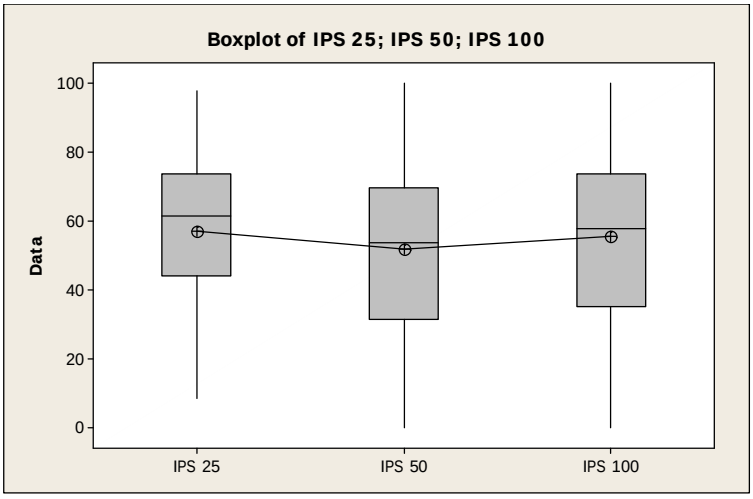
Level	N	Mean	StDev	95% CIs
IPS 25	25	56,88	26,14	(----------*----------)
IPS 50	50	51,88	25,67	(----------*----------)
IPS 100	100	55,63	25,07	(--------*-------)

48,0 54,0 60,0 66,0

Pooled StDev = 25,39



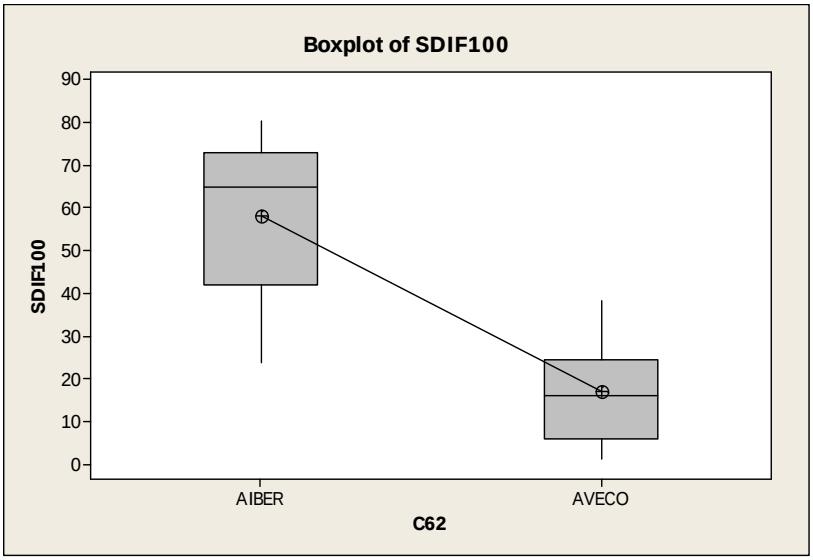




A estratificação da amostra aleatória simples foi a alocação dos países em regiões AIBER, AVECO e OTHERS, contudo, neste estudo foi excluído a categorização Others, isto porquê há objetivo de análise dos grupos com certa homogeneidade.
Nos resultados da ANOVA – observamos que há homogeneidade entre os grupos tanto para a amostra de 50 quanto para 100 países, entretanto o F test é maior para 100 países e o p-value é próximo a zero. Por outro lado, observa-se que há dois grupos bem distintos e que a aplicação de do two sample test para as amostras não seria aplicável.

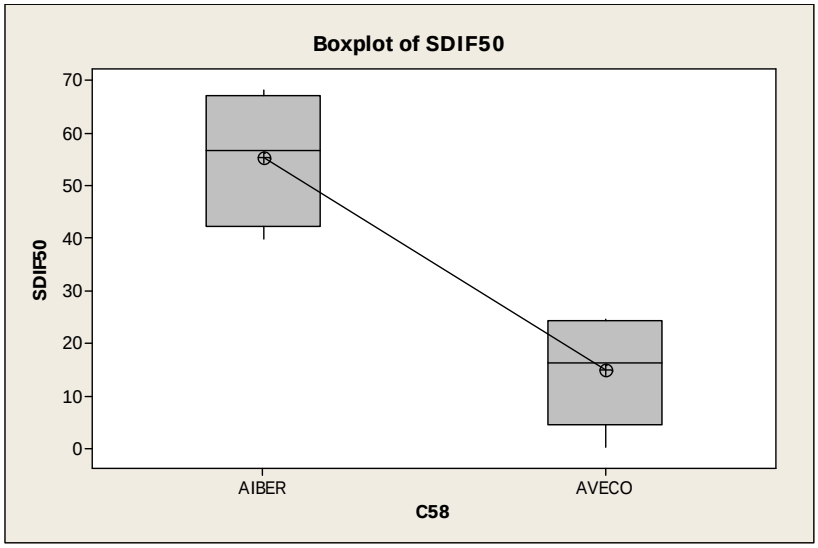
One-way ANOVA: SDIF100 versus Aiber e Aveco

```
Source  DF     SS     MS      F      P
C62      1  14818   14818   67,00  0,000
Error   33   7299    221
Total   34  22117

S = 14,87   R-Sq = 67,00%   R-Sq(adj) = 66,00%


Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  16  58,25  18,18                            (----*----)
AVECO  19  16,95  11,40  (---*----)
                         ---+---------+---------+---------+------
                           15        30        45        60
Pooled StDev = 14,87
```



One-way ANOVA: SDIF50 versus C58

```
Source  DF    SS    MS      F      P
C58      1  4511  4511   38,64  0,000
Error   11  1284   117
Total   12  5796

S = 10,81   R-Sq = 77,84%   R-Sq(adj) = 75,83%

Individual 95% CIs For Mean Based on Pooled StDev
Level  N   Mean  StDev  ------+---------+---------+---------+---
AIBER  4  55,25  13,24                        (-------*------)
AVECO  9  14,89   9,73  (----*----)
                        ------+---------+---------+---------+---
                             16        32        48        64
Pooled StDev = 10,81
```



Aplicação do Two sample – t

```
Two-sample T for SDIF100 vs SDIF50

          N  Mean  StDev  SE Mean
SDIF100  35  35,8   25,5      4,3
SDIF50   13  27,3   22,0      6,1

Difference = mu (SDIF100) - mu (SDIF50)
Estimate for difference:  8,52
95% upper bound for difference:  21,29
T-Test of difference = 0 (vs <): T-Value = 1,14  P-Value = 0,868  DF = 24
```

7. ANÁLISE DOS COMPONENTES PRINCIPAIS

A análise de componentes principais é comumente usada como um passo de uma série de análises que o estudo será objeto. Este primeiro passo é para a tentativa de redução do número de variáveis e, assim, evita a multicolinearidade, ou quando há muitos indicadores relativos ao número de observações. A melhor utilização da análise dos componentes principais é formar um número menor de variáveis não correlacionadas, a partir de um grande conjunto de dados. O objetivo da análise de componentes principais é explicar a quantidade máxima de variância com o menor número de componentes principais. Análise de componentes principais é comumente usada nas ciências sociais, pesquisa de mercado, e outras indústrias que utilizam grandes conjuntos de dados. Para a presente pesquisa do eixo de temático: Resiliência há 12 variáveis para uma amostra de 132 países, onde se observa com aplicação da análise dos componentes que com 4 variáveis há cerca de 93,5% de correlação entre as variáveis.

Principal Component Analysis: IPS N; IDH N; RISCN N; EXP N; VULN N; SUSC N; SDI

Eigenanalysis of the Correlation Matrix

Eigenvalue	7,5173	2,0192	0,8675	0,8166	0,3344	0,1530	0,1097	0,0766
Proportion	0,626	0,168	0,072	0,068	0,028	0,013	0,009	0,006
Cumulative	0,626	0,795	0,867	0,935	0,963	0,976	0,985	0,991

Eigenvalue	0,0458	0,0310	0,0213	0,0075
Proportion	0,004	0,003	0,002	0,001
Cumulative	0,995	0,998	0,999	1,000

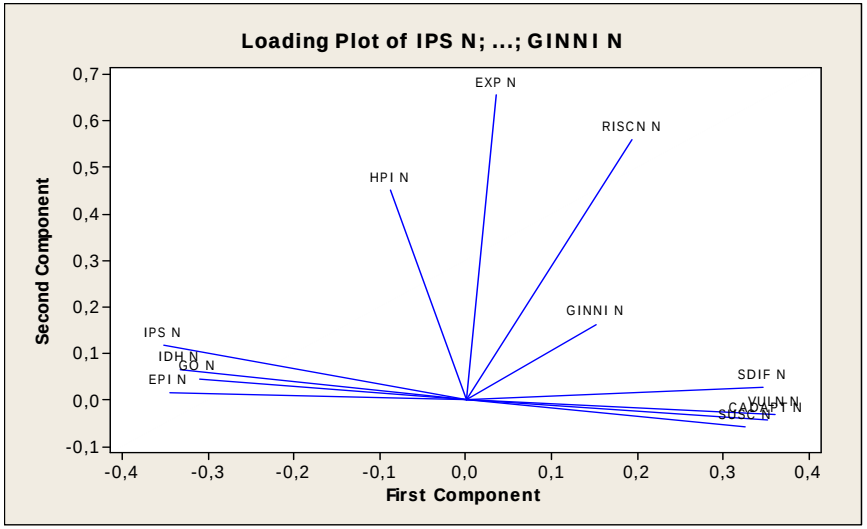
Variable	PC1	PC2	PC3	PC4
IPS N	-0,352	0,116	-0,020	-0,124
IDH N	-0,333	0,063	-0,148	0,122
RISCN N	0,194	0,559	0,246	-0,131
EXP N	0,036	0,656	0,292	-0,176
VULN N	0,361	-0,031	0,007	-0,011
SUSC N	0,326	-0,059	0,000	-0,257
SDIF N	0,347	0,027	-0,110	0,199
CADAPT N	0,351	-0,044	0,089	0,043
GO N	-0,311	0,045	0,064	-0,428
EPI N	-0,345	0,015	0,011	-0,085
HPI N	-0,088	0,451	-0,408	0,615
GINNI N	0,153	0,162	-0,801	-0,496

O primeiro componente principal tem variância (eigenvalue) 7,5173, sendo responsável por 62,6% da variância total. Os coeficientes listados em PC1 demonstram como seria a equação que representaria o eixo temático: resiliência, como segue: PC1 = -0,352IPSN-0,333IDHN+0,194RISCN+0,036EXPN+0,361VULNN+0,326SUSCN+0,347SDIFN+0,351CADAPT N-0,311GO N-0,345EPI N-0,088HPI N+0,153GINNI N

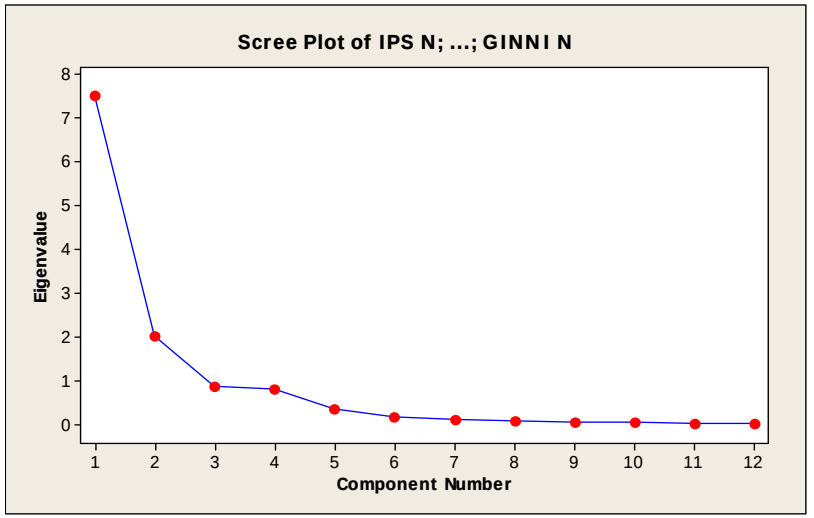
Note que para PC2 tem variância (eigenvalue) 2,0192, sendo responsável por 16,8% da variabilidade total e que PC1 e PC2 teriam 79,5%. Sucessivamente, a equação para PC2 seria: PC2 = 0,116IPS N+0,063IDH N+0,559RISC N+0,656EXP N+0,031VULN N-0,059SUSCN+0,027SDIF N+0,044CADAPT N+0,045GO N+0,015EPI N+0,451HPI N+0,162GINNIN

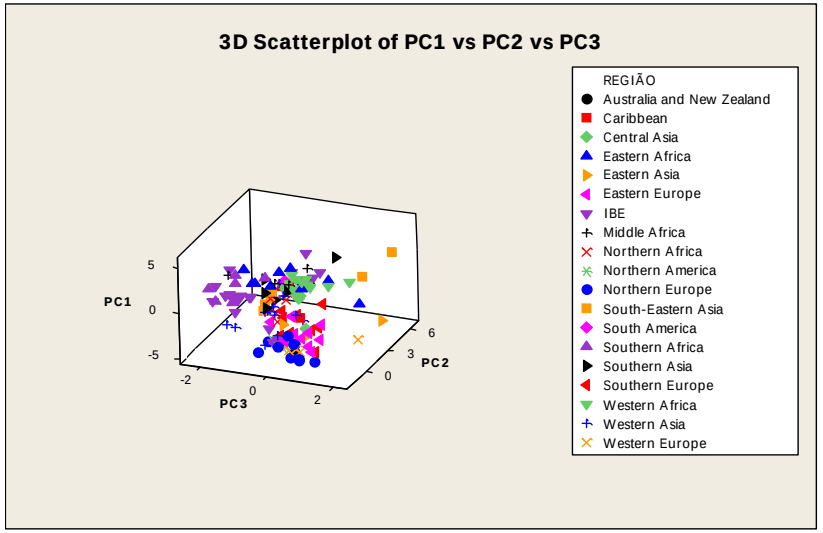
De forma geral, na análise dos 4 componentes principais, nota-se que não há uma relevância significativa entre as variáveis, ou seja, não há uma em específico que tenha uma maior ponderação que a outra, exceto pelos componentes PC3 e PC4. Finalmente, observa-se que os componentes principais em PC4 teriam variância de 0,8166 e corresponderia a 93,5% do total.

Os gráficos abaixam demonstram dois eixos opostos, um para os índices de desenvolvimento humano e econômico e outro com as políticas públicas para tornar o país resiliente - SDIF, CADAPT, VULV.

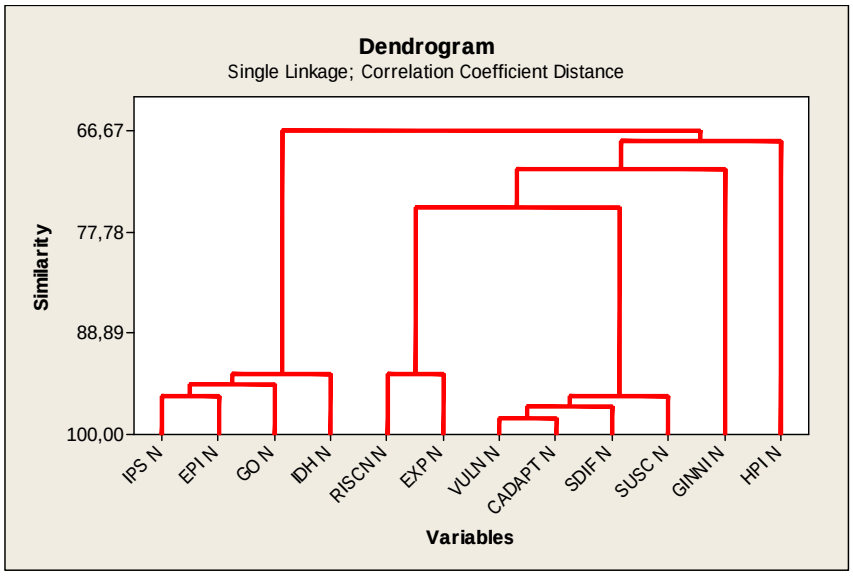


O gráfico SCREE PLOT, abaixo, demonstra que 4 componentes são suficientes para a construção de um modelo de análise para o eixo temático resiliência.





O dendograma abaixo, demonstra que há um grupo de variáveis VULV, CADAPT, SDIF, SUSC tem homogeneidade de categorias de grupo e capazes de explicar em 90%, corroborando com as análises de principais componentes.



A próxima sequência de dados teve por objetivo identificar entre os PC1, as sequências de variáveis que melhor representariam a análise sobre resiliência dos países, como segue abaixo:

Stepwise regression PC1 e outras variáveis demonstra que somente variável VULV N tem um R sqd em 97,83% e demais contribuem em cerca de 1% a cada adição.

Stepwise Regression: PC1 versus IPS N; IDH N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15
Response is PC1 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-4,9463	-3,6893	-2,2602	-0,6220	-0,6867	-0,5345
VULN N	0,1075	0,0950	0,0811	0,0659	0,0622	0,0575
T-Value	76,85	47,68	28,91	20,41	24,13	28,69
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
GO N		-0,01519	-0,01212	-0,01361	-0,01479	-0,01619
T-Value		-7,69	-6,75	-8,81	-12,08	-17,30
P-Value		0,000	0,000	0,000	0,000	0,000
EPI N			-0,0188	-0,0197	-0,0185	-0,0181
T-Value			-6,36	-7,83	-9,27	-12,03
P-Value			0,000	0,000	0,000	0,000
IDH N				-0,0142	-0,0146	-0,0167
T-Value				-7,02	-9,14	-13,61
P-Value				0,000	0,000	0,000
RISCN N					0,01157	0,01053
T-Value					8,85	10,61
P-Value					0,000	0,000
GINNI N						0,00677
T-Value						9,79
P-Value						0,000
S	0,404	0,336	0,294	0,250	0,197	0,149
R-Sq	97,85	98,52	98,88	99,19	99,50	99,72
R-Sq(adj)	97,83	98,50	98,85	99,17	99,48	99,70

8 ANALISE DE CLUSTERS

A análise de conglomerados é não uma técnica de inferência estatística em que parâmetros para uma amostra são avaliados como sendo possivelmente representativos de uma população. Ao invés disto, a análise de conglomerados é uma metodologia objetiva para quantificar características estruturais de um conjunto de observações. Os requisitos de normalidade, linearidade, homoscedasticidade são importantes para outras técnicas, mas não para Clusters.

A análise dos componentes principais demonstrou que 5 variáveis constantes no PC1: VULN, GO, EPI IDH e RISCN têm a capacidade de explicação do modelo em cerca de 96%. Para fins de análise de clusters estas variáveis foram utilizadas. A determinação ou partição de grupos foi dada para 4 grupos e o método *linkage* ou da distância mínima entre dois objetos e assim sucessivamente.

Cluster Analysis of Observations: IDH N; RISCN N; VULN N; GO N; EPI N

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98,3417	6,203	73	74	73	2
2	130	98,1779	6,815	25	51	25	2
3	129	97,5535	9,151	35	120	35	2
4	128	97,5426	9,192	70	75	70	2
5	127	97,3095	10,064	45	124	45	2
6	126	97,2865	10,150	66	68	66	2
7	125	97,0086	11,189	25	103	25	3
8	124	96,9142	11,543	49	96	49	2
9	123	96,8204	11,894	37	81	37	2
10	122	96,8081	11,939	127	130	127	2
11	121	96,7730	12,071	4	47	4	2
12	120	96,4827	13,157	63	115	63	2
13	119	96,4329	13,343	85	105	85	2
14	118	96,3432	13,678	50	67	50	2
15	117	96,2719	13,945	28	72	28	2
16	116	96,1783	14,295	5	32	5	2
17	115	96,0313	14,845	40	95	40	2
18	114	95,9864	15,013	64	70	64	3
19	113	95,9769	15,049	31	125	31	2
20	112	95,9084	15,305	9	107	9	2
21	111	95,8599	15,486	12	92	12	2
22	110	95,7962	15,725	128	129	128	2
23	109	95,6889	16,126	69	127	69	3
24	108	95,6739	16,182	35	45	35	4
25	107	95,6664	16,210	116	122	116	2
26	106	95,6315	16,341	24	121	24	2
27	105	95,6279	16,354	66	73	66	4
28	104	95,4418	17,050	17	18	17	2
29	103	95,4099	17,169	108	109	108	2
30	102	95,2316	17,836	13	114	13	2
31	101	95,1849	18,011	29	100	29	2
32	100	95,1670	18,078	21	71	21	2
33	99	95,1634	18,092	54	93	54	2
34	98	95,0104	18,664	37	94	37	3
35	97	95,0100	18,665	27	87	27	2
36	96	94,8538	19,250	5	30	5	3
37	95	94,8069	19,425	4	101	4	3
38	94	94,5627	20,338	6	59	6	2
39	93	94,4761	20,662	65	128	65	3
40	92	94,4334	20,822	24	97	24	3
41	91	94,2010	21,692	98	99	98	2
42	90	94,1479	21,890	31	50	31	4
43	89	93,8884	22,861	13	110	13	3
44	88	93,7978	23,200	44	46	44	2
45	87	93,7343	23,437	19	49	19	3
46	86	93,7239	23,476	29	117	29	3
47	85	93,5997	23,941	2	66	2	5
48	84	93,4165	24,626	90	106	90	2

49	83	93,2718	25,167	64	69	64	6
50	82	93,1233	25,723	3	40	3	3
51	81	93,0585	25,965	41	63	41	3
52	80	92,9630	26,322	91	116	91	3
53	79	92,6624	27,447	16	113	16	2
54	78	92,6477	27,502	4	79	4	4
55	77	92,5914	27,712	12	78	12	3
56	76	92,5774	27,765	15	108	15	3
57	75	92,5045	28,038	77	82	77	2
58	74	92,4934	28,079	26	98	26	3
59	73	92,2591	28,955	112	126	112	2
60	72	92,2481	28,996	33	39	33	2
61	71	92,2318	29,057	14	38	14	2
62	70	92,1892	29,217	9	56	9	3
63	69	92,1722	29,280	83	86	83	2
64	68	92,1531	29,352	28	119	28	3
65	67	92,1276	29,447	11	55	11	2
66	66	91,9737	30,023	7	118	7	2
67	65	91,9165	30,237	25	31	25	7
68	64	91,8216	30,592	34	48	34	2
69	63	91,8159	30,613	37	61	37	4
70	62	91,5952	31,439	54	112	54	4
71	61	91,5789	31,500	1	64	1	7
72	60	91,4463	31,995	17	58	17	3
73	59	91,2313	32,800	76	89	76	2
74	58	91,1855	32,971	19	22	19	4
75	57	91,1534	33,091	53	91	53	4
76	56	90,9172	33,975	27	35	27	6
77	55	90,7544	34,584	16	17	16	5
78	54	90,3805	35,982	10	57	10	2
79	53	90,1035	37,019	85	104	85	3
80	52	90,0526	37,209	11	13	11	5
81	51	89,9452	37,610	4	29	4	7
82	50	89,7289	38,420	23	123	23	2
83	49	89,6434	38,739	1	2	1	12
84	48	89,4500	39,463	20	36	20	2
85	47	89,1900	40,435	42	77	42	3
86	46	89,1877	40,444	6	8	6	3
87	45	88,5973	42,652	28	65	28	6
88	44	88,4778	43,100	21	52	21	3
89	43	88,1560	44,303	84	90	84	3
90	42	88,0540	44,685	41	102	41	4
91	41	87,8201	45,560	25	26	25	10
92	40	87,5645	46,516	5	33	5	5
93	39	87,4425	46,972	43	44	43	3
94	38	86,7204	49,673	6	34	6	5
95	37	86,2764	51,334	19	37	19	8
96	36	86,2239	51,530	53	60	53	5
97	35	85,6795	53,567	23	24	23	5
98	34	85,6492	53,680	4	27	4	13
99	33	85,5790	53,943	54	62	54	5
100	32	84,8599	56,632	9	15	9	6
101	31	84,5704	57,715	85	111	85	4
102	30	84,4432	58,191	7	12	7	5
103	29	83,4873	61,767	16	84	16	8
104	28	83,1588	62,996	3	42	3	6
105	27	82,9820	63,657	4	41	4	17
106	26	82,6241	64,995	11	54	11	10
107	25	82,6193	65,013	131	132	131	2
108	24	82,1892	66,622	25	28	25	16
109	23	81,6696	68,566	19	83	19	10
110	22	80,4896	72,980	9	85	9	10
111	21	80,3419	73,532	21	25	21	19
112	20	80,2798	73,765	4	5	4	22
113	19	79,0315	78,434	6	7	6	10
114	18	78,0807	81,990	3	19	3	16
115	17	76,8149	86,725	6	88	6	11
116	16	75,9378	90,006	43	80	43	4
117	15	75,7958	90,537	4	53	4	27
118	14	75,4312	91,901	11	16	11	18
119	13	74,2754	96,224	14	20	14	4
120	12	74,2152	96,450	21	23	21	24
121	11	73,4972	99,135	9	10	9	12
122	10	72,4096	103,203	11	76	11	20
123	9	72,0374	104,596	1	131	1	14
124	8	66,5776	125,018	3	43	3	20
125	7	63,6735	135,881	6	11	6	31
126	6	63,0202	138,325	14	21	14	28

127	5	57,2997	159,723	6	9	6	43
128	4	55,5964	166,094	3	4	3	47
129	3	48,1238	194,046	1	14	1	42
130	2	37,2386	234,763	3	6	3	90
131	1	0,0000	374,056	1	3	1	132

Final Partition
Number of clusters: 3

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	42	34705,0	25,6620	60,9498
Cluster2	47	29646,9	21,7750	78,6680
Cluster3	43	35535,9	26,4447	64,2954

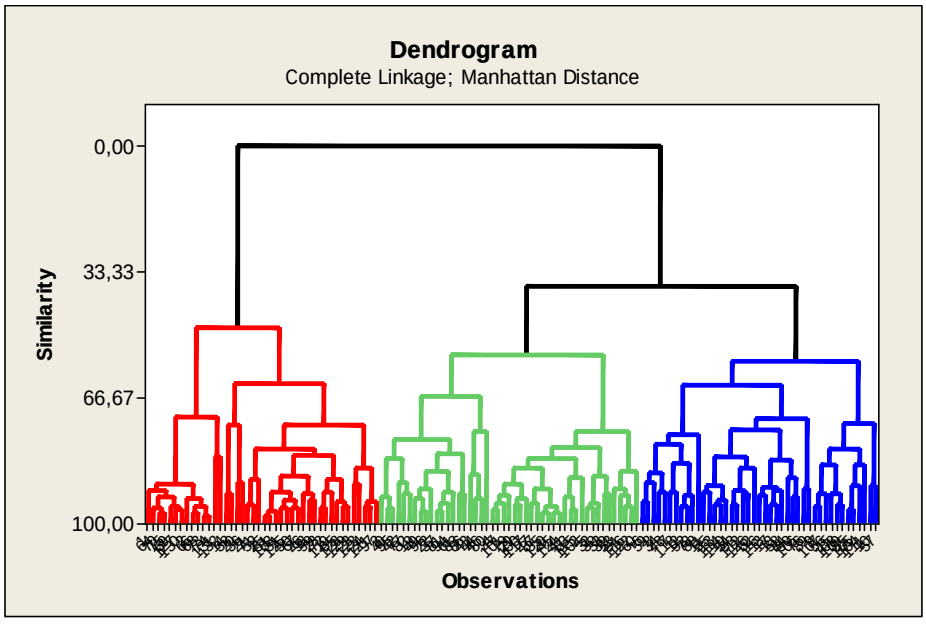
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Grand centroid
IDH N	82,3237	62,7586	29,2012	58,0523
RISCN N	12,4332	23,9164	27,3436	21,3791
VULN N	17,7124	44,2844	75,5498	46,0146
GO N	75,1610	37,8705	23,3884	45,0180
EPI N	77,0079	48,4427	22,8131	49,1826

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3
Cluster1	0,000	58,5430	109,576
Cluster2	58,543	0,0000	54,607
Cluster3	109,576	54,6075	0,000

Dendrogram



9 ANÁLISE DE DISCRIMINANTE

No contexto de análise multivariada, uma das principais questões se refere a utilização de suas técnicas para fins de classificação e posterior previsão dos elementos que estão sendo observados. Seja essa classificação com intuito apenas de organização ou separação simples entre grupos, o que se observou foi que essa procura conduziu ao desenvolvimento de diversas técnicas para auxiliar nesse propósito. Uma das técnicas mais famosas para resolver problemas de classificação e previsão de elementos é a análise de discriminante.

A análise discriminante é utilizada para a classificação de observações em dois ou mais grupos de uma amostra com grupos conhecidos. Também, é utilizada para investigar como variáveis contribuem para a separação do grupo. É possível a utilização da função linear e quadrática. Onde, se linear: se assemelha a uma equação de regressão múltipla, portanto, composta por variáveis independentes que representam as características do elemento, que são ponderadas pelo nível de sua importância e ou impacto que causam na variável dependente. Na função linear, todos os grupos são considerados ter a mesma matriz covariância, no entanto, se quadrática, não faz essa suposição, bem como de suas propriedades.

Sua característica básica é a utilização de um conjunto de informações obtidas acerca de variáveis consideradas independentes para conseguir um valor de uma variável dependente que possibilite a classificação desejada. Na análise discriminante é de natureza qualitativa, ou seja, categórica, discreta, já que seu valor representa uma classificação estabelecida. A classificação se realiza mediante o confronto do valor obtido com os valores de outros elementos. Exemplificando: modelo de análise de risco de crédito pode ser dividido em categorias de clientes rotulados como alto, médio e baixo risco, contribuindo com informação de maior relevância para o gestor.

No presente trabalho, na etapa de análise de clusters foi realizado por alocação de similaridades de clusters, segundo a cinco principais variáveis constantes em 4 componentes principais: IDH, RISCN, VULVN, GO e EPI – variáveis independentes. Cabe ressaltar que a determinação das variáveis foi em de acordo com a ANOVA dos 4 componentes principais.

Análise Discriminante – SPSS para o Cluster 3R

Resumo de processamento de caso de análise

Casos não ponderados		N	Porcentagem
Válido		132	100,0
Excluídos	Códigos de grupo ausentes ou fora do intervalo	0	,0
	Pelo menos uma variável discriminante ausente	0	,0
	Códigos de grupo ausentes ou fora do intervalo e pelo menos uma variável discriminadora ausente	0	,0
	Total	0	,0
Total		132	100,0

Estatísticas de grupo

Clusters 3R		N válido (de lista)	
		Não ponderado	Ponderado
1,0	IDH N	42	42,000
	RISCN N	42	42,000
	VULN N	42	42,000
	GO N	42	42,000
	EPI N	42	42,000
2,0	IDH N	47	47,000
	RISCN N	47	47,000
	VULN N	47	47,000
	GO N	47	47,000
	EPI N	47	47,000
3,0	IDH N	43	43,000
	RISCN N	43	43,000
	VULN N	43	43,000
	GO N	43	43,000
	EPI N	43	43,000
Total	IDH N	132	132,000
	RISCN N	132	132,000
	VULN N	132	132,000
	GO N	132	132,000
	EPI N	132	132,000

Análise 1 - Estatísticas em etapas

Variáveis Inseridas/Removidas[a,b,c,d]

Etapa	Inseridas	Lambda de Wilks							
						F exato			
		Estatística	df1	df2	df3	Estatística	df1	df2	Sig.
1	VULN N	,145	1	2	129,000	380,174	2	129,000	,000
2	GO N	,109	2	2	129,000	130,243	4	256,000	,000
3	EPI N	,095	3	2	129,000	95,227	6	254,000	,000
4	RISCN N	,086	4	2	129,000	75,955	8	252,000	,000

Em cada etapa, a variável que minimiza o Lambda de Wilks geral é inserida.
a. O número máximo de etapas é 10.
b. O F parcial mínimo a ser inserido é 3.84.
c. O F parcial máximo a ser removido é 2.71.
d. Nível f, tolerância ou VIN insuficiente para cálculos adicionais.

Variáveis na análise

Etapa		Tolerância	F a ser removido	Lambda de Wilks
1	VULN N	1,000	380,174	
2	VULN N	,892	104,372	,286
	GO N	,892	21,512	,145
3	VULN N	,669	22,637	,128
	GO N	,879	16,270	,119
	EPI N	,695	9,288	,109
4	VULN N	,602	28,491	,125
	GO N	,861	16,958	,109
	EPI N	,684	9,978	,100
	RISCN N	,815	6,431	,095

Lambda de Wilks

Etapa	Número de variáveis	Lambda	df1	df2	df3	F exato			
						Estatística	df1	df2	Sig.
1	1	,145	1	2	129	380,174	2	129,000	,000
2	2	,109	2	2	129	130,243	4	256,000	,000
3	3	,095	3	2	129	95,227	6	254,000	,000
4	4	,086	4	2	129	75,955	8	252,000	,000

Resumo de funções discriminantes canônicas

Valores próprios

Função	Autovalor	% de variância	% cumulativa	Correlação canônica
1	7,891[a]	96,2	96,2	,942
2	,309[a]	3,8	100,0	,486

a. As primeiras 2 funções discriminantes canônicas foram usadas na análise.

Lambda de Wilks

Teste de funções	Lambda de Wilks	Qui-quadrado	df	Sig.
1 até 2	,086	312,907	8	,000
2	,764	34,315	3	,000

Coeficientes de funções discriminantes canônicas padronizados

	Função	
	1	2
RISCN N	,241	-,512
VULN N	-,634	,825
GO N	,213	,934
EPI N	,468	,155

Matriz de estruturas

	Função	
	1	2
VULN N	-,863*	,230
EPI N	,798*	,113
IDH N[b]	,527*	-,196
GO N	,545	,713*
RISCN N	-,153	-,247*

Correlações entre grupos no conjunto entre variáveis discriminantes e funções discriminantes canônicas padronizadas

Variáveis ordenadas por tamanho absoluto de correlação na função.

*. Maior correlação absoluta entre cada variável e qualquer função discriminante

b. Essa variável não é usada na análise.

Funções em centroides de grupo

Clusters 3R	Função	
	1	2
1,0	3,496	,410
2,0	,011	-,739
3,0	-3,426	,407

Funções discriminantes canônicas não padronizadas avaliadas em médias de grupo

Estatísticas de classificação

Resumo de processamento de classificação

Processado		132
Excluídos	Códigos de grupo ausentes ou fora do intervalo	0
	Pelo menos uma variável discriminante ausente	0
Usado em saída		132

Probabilidades a priori para grupos

Clusters 3R	A priori	Casos utilizados na análise	
		Não ponderado	Ponderado
1,0	,333	42	42,000
2,0	,333	47	47,000
3,0	,333	43	43,000
Total	1,000	132	132,000

Resultados da classificação[a]

			Associação ao grupo prevista			
		Clusters 3R	1,0	2,0	3,0	Total
Original	Contagem	1,0	39	3	0	42
		2,0	1	45	1	47
		3,0	0	3	40	43
	%	1,0	92,9	7,1	,0	100,0
		2,0	2,1	95,7	2,1	100,0
		3,0	,0	7,0	93,0	100,0

a. 93,9% de casos agrupados originais classificados corretamente.

A análise de discriminante para a classificação feita por Clusters 3R tem chance de correta classificação dos países em 93,9%.

10 REGRESSÃO LOGÍSTICA

A técnica de regressão logística foi desenvolvida por volta de 1960 em resposta ao desafio de realizar predições ou explicar a ocorrência de determinados fenômenos quando a variável dependente fosse de natureza binária ou dicotômica, assume entre dois resultados e pode ser de natureza qualitativa. Isto exige que o resultado da análise possibilite associações a certas categorias, como por exemplo: positivo ou negativo, aceitar ou rejeitar, morrer ou sobreviver etc.

REGIÃO_6R versus IDH N; RISCN; VULV; GO, EPI

A análise ANOVA abaixo tem como objetivo a classificação das 6 regiões com base no índice de governança. Este procedimento se faz necessário, pois que a regressão logística é uma análise dicotômica, ou seja, necessita de um parâmetro para que os testes possam ser realizados. A coluna classificação demonstrada nos resultados abaixo, será utilizada nos testes de regressão logística para a região 6R, lembrando que foi uma inferência e não produzido pelos testes ANOVA.

One-way ANOVA: GO N versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56677  8097  35,53  0,000
Error     124  28261   228
Total     131  84938
S = 15,10   R-Sq = 66,73%   R-Sq(adj) = 64,85%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------   Classificação
AIBER  20  43,17  19,04           (---*--)                           (3)
AVECO  27  83,25  13,38                                (--*--)       (1)
CARLA   4  39,62   9,54     (-------*------)                         (4)
CWEAS  11  27,38  12,38  (----*---)                                  (8)
EMDEU  12  51,93  11,21                (---*---)                     (2)
ESEAS  12  33,61  10,47     (----*---)                               (5)
NSWUA  14  28,46  16,21   (---*---)                                  (7)
SUSAF  32  29,58  16,95     (--*-)                                   (6)
                         -+---------+---------+---------+--------
                         20        40        60        80
Pooled StDev = 15,10
```

Ordinal Logistic Regression: N Class 6R versus IDH N; RISCN N; ...

De acordo com o a tabela de Logistic Regression, somente o índice de governança (GO) tem um p-value inferior a 0,5%, com isto tendem não contribuir com o modelo de regressão logística, portanto, serão retirados do modelo para novos testes com regressão logística, conforme demonstrado a seguir.

Goodness-of-Fit Tests demonstra a adequacidade do modelo Logit em relação as variáveis IDH N 1; RISCN N 1; VULVN1; GO N 1, EPI N 1. Neste estudo, valor-p para o teste de Pearson é 0,95, e o valor p para o teste de desvio é 1,000, o que indica que não há provas suficientes para afirmar que o modelo não se encaixa nos dados de forma adequada. Se o p-value é menor do que o a-level (0,05), o teste rejeita a hipótese nula de que o modelo se ajusta aos dados de forma adequada. Os testes de hipótese é um procedimento que avalia duas afirmações mutuamente exclusivas sobre uma população. Um teste de hipótese usa dados para determinar qual afirmação é melhor suportada pelos dados. Estas afirmações são chamadas de hipótese nula e a hipótese alternativa.

```
Link Function: Logit
Response Information
Variable    Value  Count
N Class 6R  1         27
            2         12
            3         20
            4          4
            5         12
            6         32
            7         14
            8         11
            Total    132

Logistic Regression Table
                                                    Odds     95% CI
Predictor       Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)    -7,74211    2,57483  -3,01  0,003
Const(2)    -6,67815    2,55248  -2,62  0,009
Const(3)    -5,42288    2,53000  -2,14  0,032
Const(4)    -5,20356    2,52657  -2,06  0,039
Const(5)    -4,58130    2,51713  -1,82  0,069
Const(6)    -2,91001    2,49729  -1,17  0,244
Const(7)    -1,84088    2,49996  -0,74  0,462
IDH N      0,0023050  0,0165670   0,14  0,889   1,00   0,97   1,04
RISCN N    0,0178456  0,0122997   1,45  0,147   1,02   0,99   1,04
VULN N     0,0093961  0,0247054   0,38  0,704   1,01   0,96   1,06
GO N       0,0746124  0,0131349   5,68  0,000   1,08   1,05   1,11
EPI N      0,0239517  0,0185216   1,29  0,196   1,02   0,99   1,06

Log-Likelihood = -202,264
Test that all slopes are zero: G = 107,907, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests
Method    Chi-Square   DF      P
Pearson      842,912  912  0,950
Deviance     404,528  912  1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs       Number  Percent  Summary Measures
Concordant    5914     80,7  Somers' D              0,62
Discordant    1403     19,2  Goodman-Kruskal Gamma  0,62
Ties             8      0,1  Kendall's Tau-a        0,52
Total         7325    100,0
```

Ordinal Logistic Regression: N Class 6R versus GO N

```
Link Function: Logit
Response Information
Variable    Value  Count
N Class 6R  1         27
            2         12
            3         20
            4          4
            5         12
            6         32
            7         14
            8         11
            Total    132
Logistic Regression Table
                                                    Odds     95% CI
Predictor       Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)    -5,92207   0,630396  -9,39  0,000
Const(2)    -4,85482   0,543020  -8,94  0,000
Const(3)    -3,61959   0,453391  -7,98  0,000
Const(4)    -3,41019   0,441678  -7,72  0,000
Const(5)    -2,82267   0,413561  -6,83  0,000
Const(6)    -1,19790   0,377656  -3,17  0,002
Const(7)   -0,133538   0,417041  -0,32  0,749
GO N       0,0822319  0,0095571   8,60  0,000   1,09   1,07   1,11

Log-Likelihood = -204,223
Test that all slopes are zero: G = 103,988, DF = 1, P-Value = 0,000

Goodness-of-Fit Tests
Method    Chi-Square   DF      P
Pearson      836,994  916  0,970
Deviance     408,447  916  1,000
```

```
Measures of Association: (Between the Response Variable and Predicted Probabilities)
Pairs       Number  Percent  Summary Measures
Concordant    5867     80,1  Somers' D              0,60
Discordant    1436     19,6  Goodman-Kruskal Gamma  0,61
Ties            22      0,3  Kendall's Tau-a        0,51
Total         7325    100,0
```

A regressão logística para a os países que compõem a 6R após testes ANOVA indicam que tem 80,1% de probabilidade de acerto nas alocações propostas.

REGIÃO_3R versus IDH N; RISCN; VULV; GO, EPI

One-way ANOVA: GO N versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  51019  25510  97,02  0,000
Error     129  33919    263
Total     131  84938

S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%

Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+--------- Classificação
AIBER   21  42,48  18,82       (---*----)                            (2)
AVECO   27  83,25  13,38                                 (----*---)  (1)
OTHERS  84  33,36  16,35   (-*--)                                    (3)
                           +---------+---------+---------+---------
                          30        45        60        75
Pooled StDev = 16,22
```

De acordo com o a tabela de Logistic Regression, somente o índice de governança (GO) tem um p-value inferior a 0,5%, com isto tendem não contribuir com o modelo de regressão logística, portanto, serão retirados do modelo para novos testes com regressão logística, conforme demonstrado a seguir.

Goodness-of-Fit Tests demonstra a adequacidade do modelo Logit em relação as variáveis IDH N 1; RISCN N 1; VULVN1; GO N 1, EPI N 1. Neste estudo, valor-p para o teste de Pearson é 1,000 e o valor p para o teste de desvio é 1,000, o que indica que não há provas suficientes para afirmar que o modelo não se encaixa nos dados de forma adequada. Se o p-value é menor do que o a-level (0,05), o teste rejeita a hipótese nula de que o modelo se ajusta aos dados de forma adequada.

Ordinal Logistic Regression: N Class 3R versus IDH N; RISCN; VULV; GO, EPI

```
Link Function: Logit
Response Information
Variable    Value  Count
N Class 3R  1         27
            2         21
            3         84
            Total    132

Logistic Regression Table
                                                    Odds     95% CI
Predictor         Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)     -7,39275    3,69148  -2,00  0,045
Const(2)     -5,61818    3,64238  -1,54  0,123
IDH N        0,0199319  0,0228124   0,87  0,382   1,02   0,98   1,07
RISCN N      0,0239635  0,0164209   1,46  0,144   1,02   0,99   1,06
VULN N      -0,0193560  0,0381811  -0,51  0,612   0,98   0,91   1,06
GO N         0,0351707  0,0174744   2,01  0,044   1,04   1,00   1,07
EPI N        0,0468898  0,0298642   1,57  0,116   1,05   0,99   1,11

Log-Likelihood = -70,623
Test that all slopes are zero: G = 97,592, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests
Method    Chi-Square   DF      P
Pearson      170,995  257  1,000
Deviance     141,245  257  1,000

Measures of Association: (Between the Response Variable and Predicted Probabilities)
```

```
Pairs       Number  Percent  Summary Measures
Concordant    4071     88,5  Somers' D              0,77
Discordant     520     11,3  Goodman-Kruskal Gamma  0,77
Ties             8      0,2  Kendall's Tau-a        0,41
Total         4599    100,0
```

Ordinal Logistic Regression: N Class 3R versus GO N

```
Link Function: Logit
Response Information
Variable    Value  Count
N Class 3R  1         27
            2         21
            3         84
            Total    132

Logistic Regression Table
                                                    Odds     95% CI
Predictor       Coef    SE Coef       Z      P  Ratio  Lower  Upper
Const(1)    -5,96448   0,757718  -7,87  0,000
Const(2)    -4,37296   0,613992  -7,12  0,000
GO N        0,0828339  0,0116209   7,13  0,000   1,09   1,06   1,11

Log-Likelihood = -78,262
Test that all slopes are zero: G = 82,314, DF = 1, P-Value = 0,000

Goodness-of-Fit Tests
Method    Chi-Square   DF      P
Pearson      228,576  261  0,927
Deviance     156,524  261  1,000

Measures of Association:(Between the Response Variable and Predicted Probabilities)
Pairs       Number  Percent  Summary Measures
Concordant    3893     84,6  Somers' D              0,70
Discordant     696     15,1  Goodman-Kruskal Gamma  0,70
Ties            10      0,2  Kendall's Tau-a        0,37
Total         4599    100,0
```

A regressão logística para a os países que compõem a 3R após testes ANOVA indicam que tem 84,6% de probabilidade de acerto nas alocações propostas.

11 ÁRVORES DE DECISÃO

É um modelo de classificação baseado em árvore. Classifica os dados em grupos ou prevê valores de uma variável dependente (destino) com base em valores de variáveis (Predictor) independentes. Este procedimento fornece ferramentas de validação para análise de classificação exploratória e confirmatória.

O procedimento pode ser utilizado para: (a)segmentação- identificação pessoas que possam ser membros de um grupo particular; (b)Estratificação - atribui casos em uma das várias categorias, tais como alta, média, e os grupos de baixo risco; (c) preditores - cria regras e os usa para eventos futuros, previsão, como por exemplo: risco de crédito; (d) redução de dados e seleção de variáveis; (e)identificação e as interações - identificar relações que dizem respeito apenas aos subgrupos específicos para a construção de um modelo estatístico;

CLUSTERS 3R – é a classificação proposta pela análise de conglomerados dada a similiaridade pelas variáveis escolhidas: IDH N; RISCN N; VULN N; GO N; EPI N.

MODELO CHAID

Resumo do modelo

Especificações	Método crescente	CHAID
	Variável dependente	3 R ordem
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	3
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	EPI N, IDH N, GO N
	Número de nós	11
	Número de nós de terminal	7
	Espessura	2

Risco

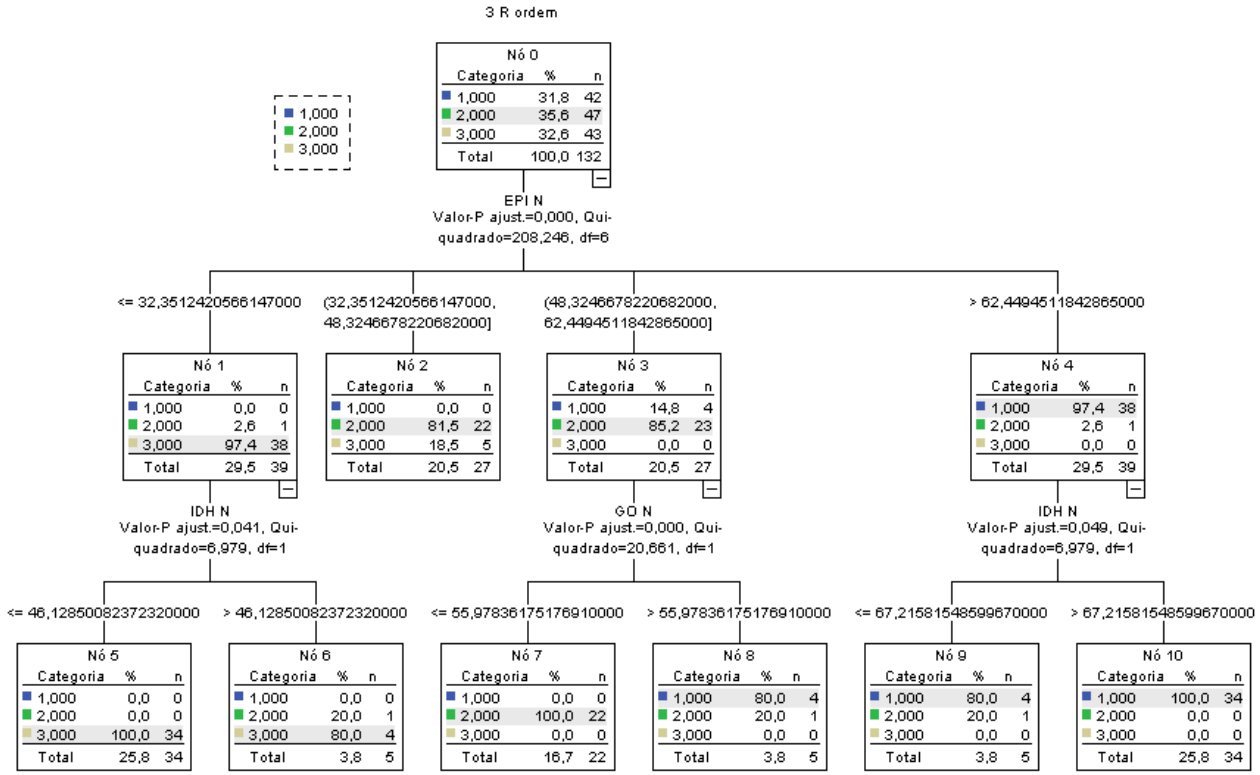
Estimativa	Erro Padrão
,061	,021

Método Crescente: CHAID
Variável Dependente: 3 R ordem

Classificação

	Previsto			
Observado	1,0	2,0	3,0	Porcentagem Correta
1,0	42	0	0	100,0%
2,0	2	44	1	93,6%
3,0	0	5	38	88,4%
Porcentagem global	33,3%	37,1%	29,5%	93,9%

Método Crescente: CHAID
Variável Dependente: 3 R ordem



MODELO MODELO CHAID EXAUSTIVO

Resumo do modelo

Especificações	Método crescente	CHAID EXAUSTIVO
	Variável dependente	3 R ordem
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	3
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	VULN N, GO N
	Número de nós	7
	Número de nós de terminal	5
	Espessura	2

Risco

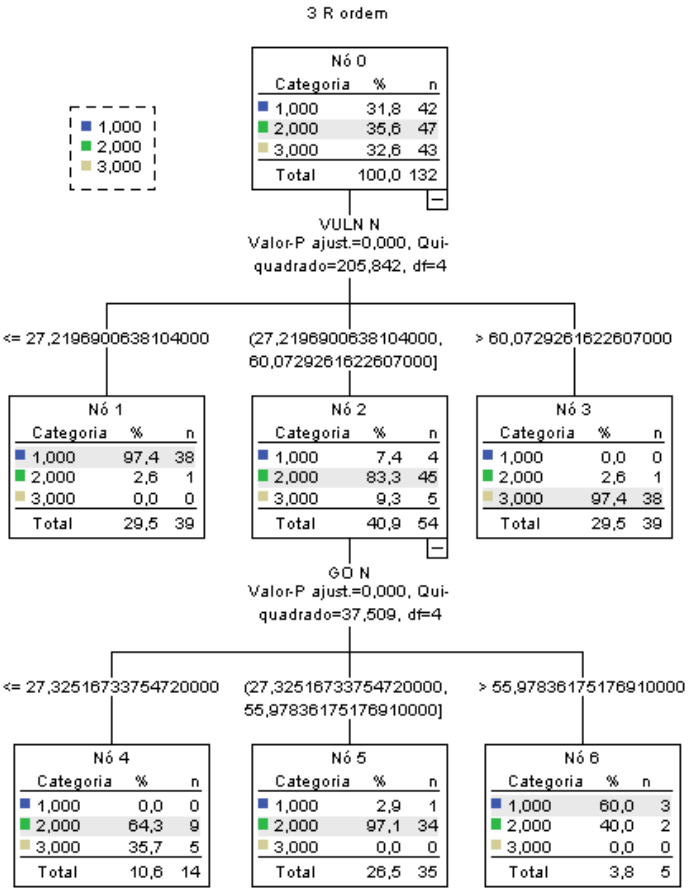
Estimativa	Erro Padrão
,076	,023

Método Crescente: CHAID EXAUSTIVO
Variável Dependente: 3 R ordem

Classificação

	Previsto			
Observado	1,0	2,0	3,0	Porcentagem Correta
1,0	41	1	0	97,6%
2,0	3	43	1	91,5%
3,0	0	5	38	88,4%
Porcentagem global	33,3%	37,1%	29,5%	92,4%

Método Crescente: CHAID EXAUSTIVO
Variável Dependente: 3 R ordem



MODELO MODELO CRT

Resumo do modelo

Especificações	Método crescente	CRT
	Variável dependente	3 R ordem
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	5
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	VULN N, EPI N, GO N, IDH N, RISCN N
	Número de nós	9
	Número de nós de terminal	5
	Espessura	3

Risco

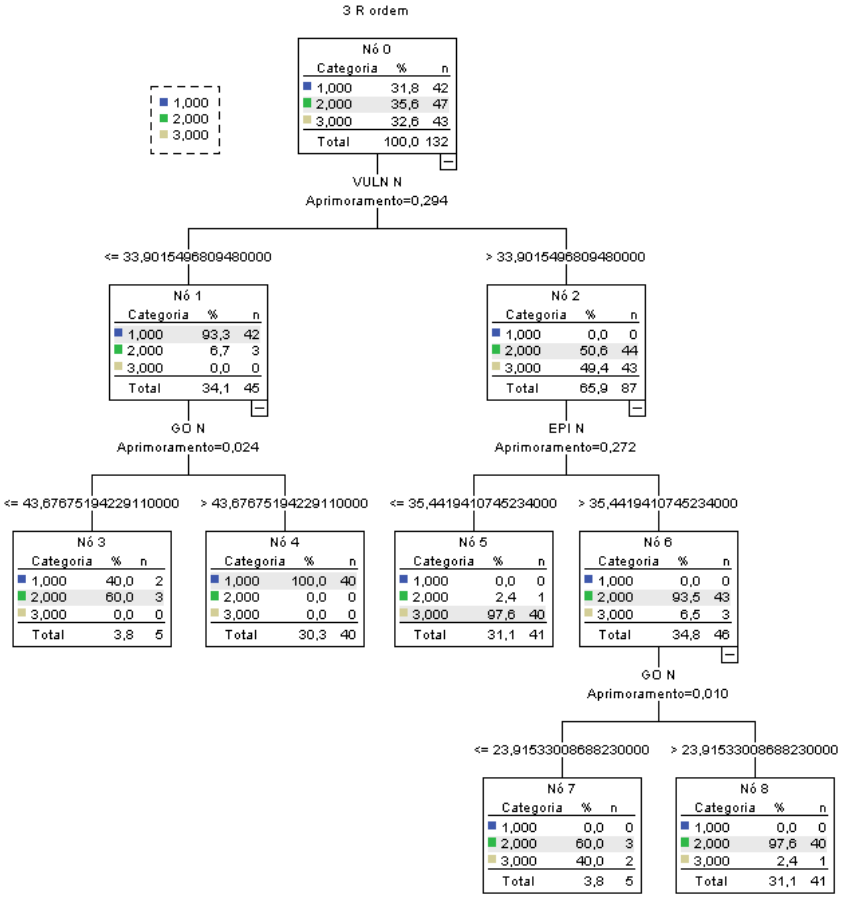
Estimativa	Erro Padrão
,045	,018

Método Crescente: CRT
Variável Dependente: 3 R ordem

Classificação

	Previsto			
Observado	1,0	2,0	3,0	Porcentagem Correta
1,0	40	2	0	95,2%
2,0	0	46	1	97,9%
3,0	0	3	40	93,0%
Porcentagem global	30,3%	38,6%	31,1%	95,5%

Método Crescente: CRT
Variável Dependente: 3 R ordem



MODELO MODELO QUEST

Resumo do modelo

Especificações	Método crescente	QUEST
	Variável dependente	3 R ordem
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	5
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	VULN N, IDH N, EPI N, GO N, RISCN N
	Número de nós	5
	Número de nós de terminal	3
	Espessura	2

Risco

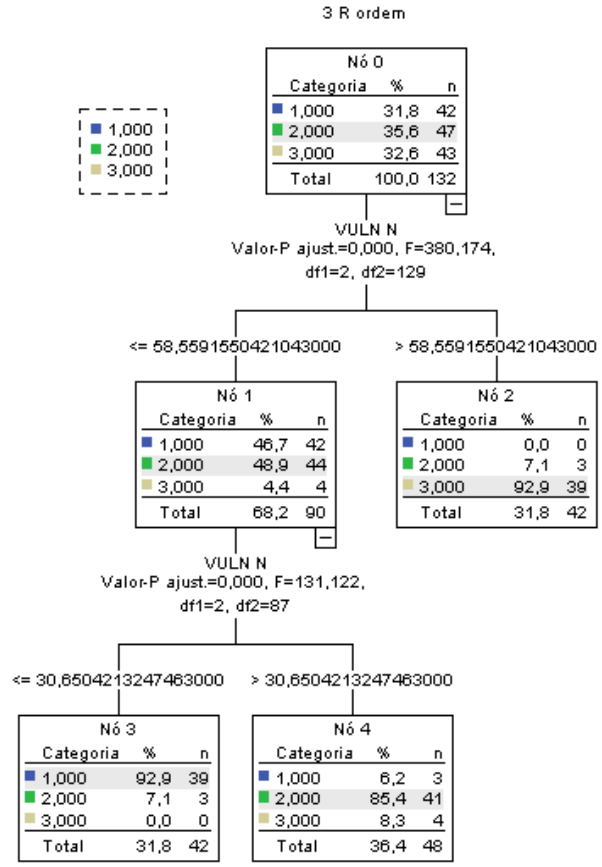
Estimativa	Erro Padrão
,098	,026

Método Crescente: QUEST
Variável Dependente: 3 R ordem

Classificação

	Previsto			
Observado	1,0	2,0	3,0	Porcentagem Correta
1,0	39	3	0	92,9%
2,0	3	41	3	87,2%
3,0	0	4	39	90,7%
Porcentagem global	31,8%	36,4%	31,8%	90,2%

Método Crescente: QUEST
Variável Dependente: 3 R ordem



AIBER, AVECO E OTHERS – classificação dos 132 países em 3 regiões pré-definidas.

Resumo do modelo

Especificações	Método crescente	CHAID
	Variável dependente	Class 3R
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	3
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	EPI N, IDH N
	Número de nós	7
	Número de nós de terminal	5
	Espessura	2

Risco

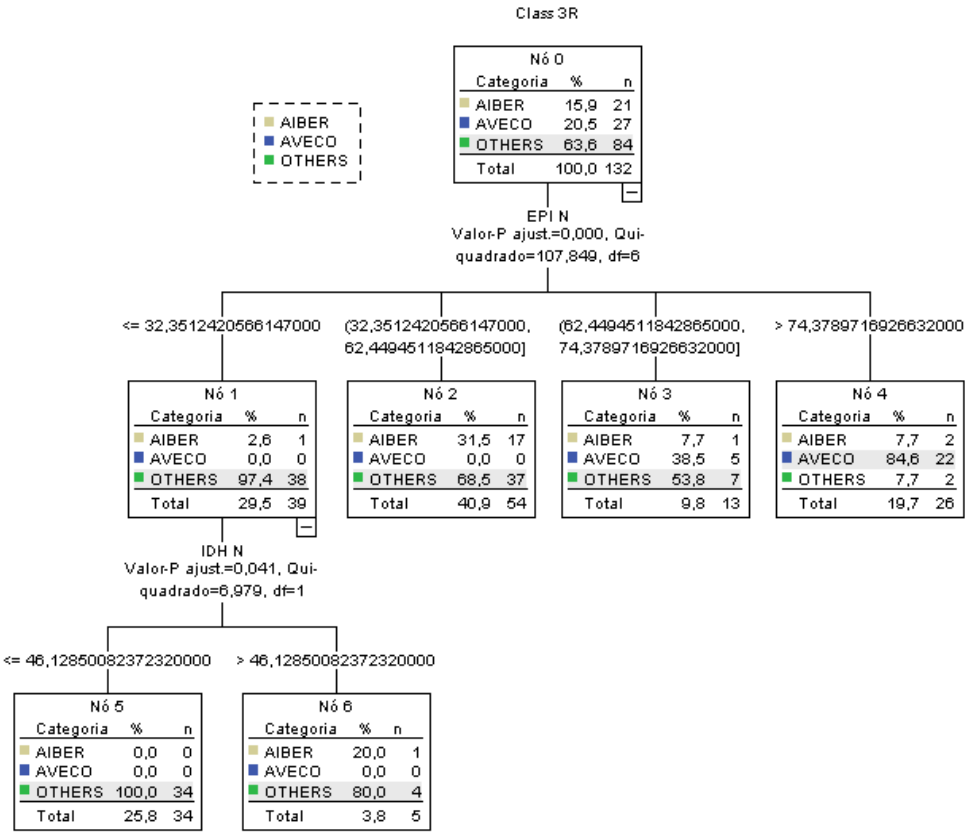
Estimativa	Erro Padrão
,212	,036

Método Crescente: CHAID
Variável Dependente: Class 3R

Classificação

	Previsto			
Observado	AIBER	AVECO	OTHERS	Porcentagem Correta
AIBER	0	2	19	0,0%
AVECO	0	22	5	81,5%
OTHERS	0	2	82	97,6%
Porcentagem global	0,0%	19,7%	80,3%	78,8%

Método Crescente: CHAID
Variável Dependente: Class 3R



MODELO CHAID EXAUSTIVO

Resumo do modelo

Especificações	Método crescente	CHAID EXAUSTIVO
	Variável dependente	Class 3R
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	3
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	EPI N, GO N
	Número de nós	8
	Número de nós de terminal	6
	Espessura	2

Risco

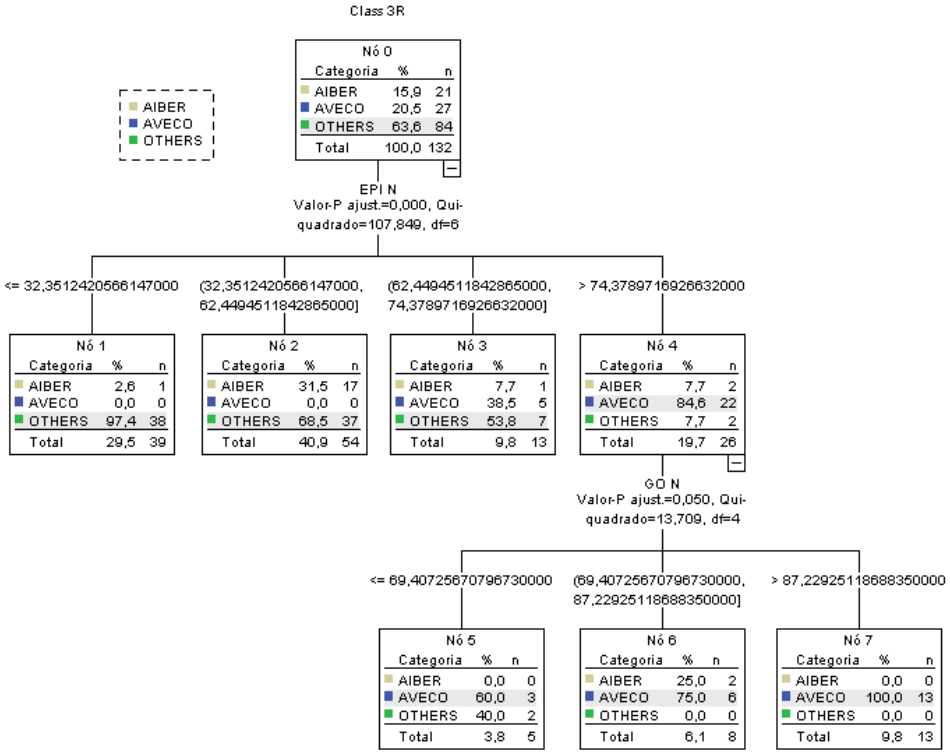
Estimativa	Erro Padrão
,212	,036

Método Crescente: CHAID EXAUSTIVO
Variável Dependente: Class 3R

Classificação

	Previsto			
Observado	AIBER	AVECO	OTHERS	Porcentagem Correta
AIBER	0	2	19	0,0%
AVECO	0	22	5	81,5%
OTHERS	0	2	82	97,6%
Porcentagem global	0,0%	19,7%	80,3%	78,8%

Método Crescente: CHAID EXAUSTIVO
Variável Dependente: Class 3R



MODELO CRT

Resumo do modelo

Especificações	Método crescente	CRT
	Variável dependente	Class 3R
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	5
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	VULN N, EPI N, GO N, IDH N, RISCN N
	Número de nós	9
	Número de nós de terminal	5
	Espessura	3

Risco

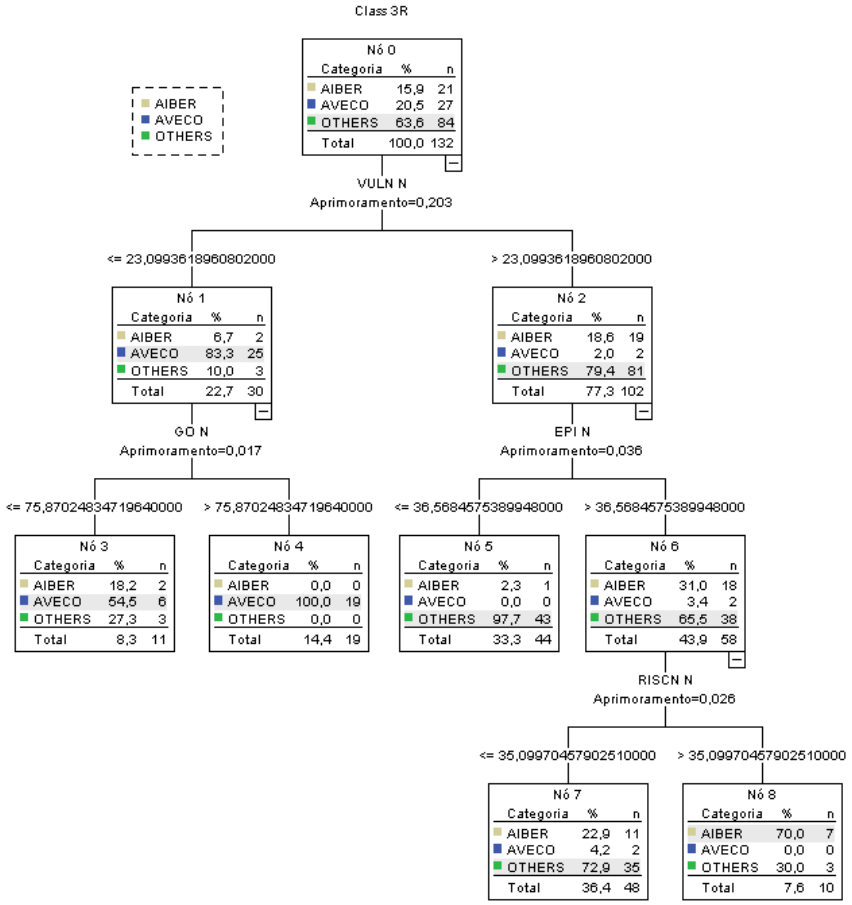
Estimativa	Erro Padrão
,167	,032

Método Crescente: CRT
Variável Dependente: Class 3R

Classificação

	Previsto			
Observado	AIBER	AVECO	OTHERS	Porcentagem Correta
AIBER	7	2	12	33,3%
AVECO	0	25	2	92,6%
OTHERS	3	3	78	92,9%
Porcentagem global	7,6%	22,7%	69,7%	83,3%

Método Crescente: CRT
Variável Dependente: Class 3R



MODELO QUEST

Resumo do modelo

Especificações	Método crescente	QUEST
	Variável dependente	Class 3R
	Variáveis independentes	IDH N, RISCN N, VULN N, EPI N, GO N
	Validação	Nenhum
	Profundidade máxima de árvore	5
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	GO N, VULN N, EPI N, IDH N, RISCN N
	Número de nós	5
	Número de nós de terminal	3
	Espessura	2

Risco

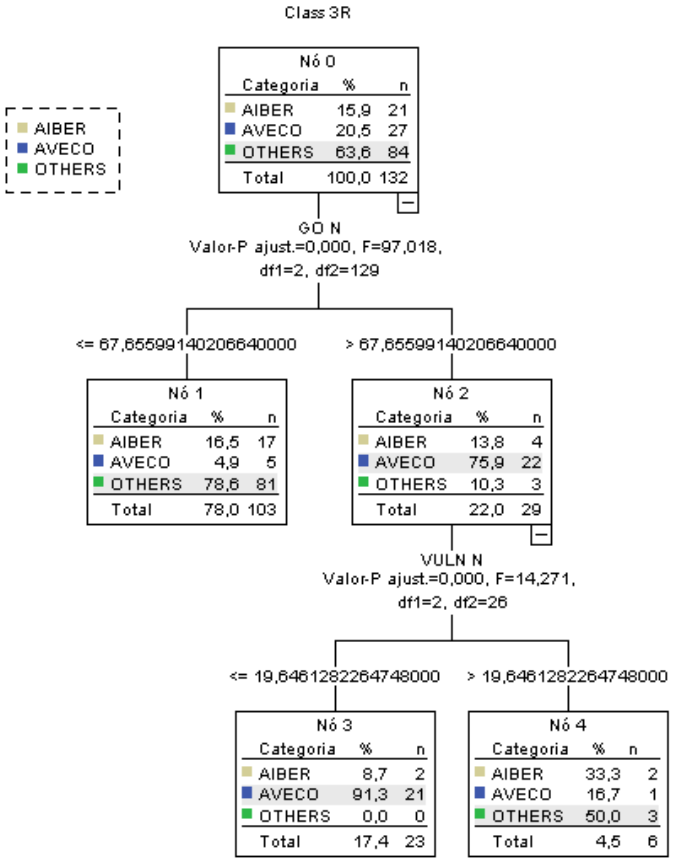
Estimativa	Erro Padrão
,205	,035

Método Crescente: QUEST
Variável Dependente: Class 3R

Classificação

	Previsto			
Observado	AIBER	AVECO	OTHERS	Porcentagem Correta
AIBER	0	2	19	0,0%
AVECO	0	21	6	77,8%
OTHERS	0	0	84	100,0%
Porcentagem global	0,0%	17,4%	82,6%	79,5%

Método Crescente: QUEST
Variável Dependente: Class 3R



As técnicas estatísticas de analise discriminante e de arvore de decisão para o eixo temático resiliência com as variáveis: IDH, RISCN, VULVN, GO e EPI, demonstrou que a arvore de decisão tem probabilidade de êxito na segregação dos grupos com base nas variáveis independentes em 95,5% através do modelo CRT. Para as regiões pré-classificadas em AIBER, AVECO e OTHERS tem o melhor modelo CRT em 83,3%, o que já foi observado nas técnicas de analise discriminante e regressão logística realizados pela ferramenta Mini-tab.

O resultado demonstrou que para o eixo temático resiliência e suas regiões que a técnica de arvore de decisão tem sua melhor alocação com o método CRT, com as variáveis: VULV, EPI E GOV.

12 ANÁLISE DE CORRESPONDÊNCIA

A análise de correspondência foi desenvolvida na década de 1960 por um grupo de estatísticos franceses . Faz parte de um conjunto de métodos utilizados para a análise descritiva exploratória de grandes tabelas. A concepção geral é semelhante à análise de componentes principais e à análise fatorial, diferenciando-se destas, entre outros aspectos, por permitir a inclusão de variáveis categóricas.

É uma técnica de análise exploratória de dados adequada para analisar tabelas de duas entradas ou tabelas de múltiplas entradas, levando em conta algumas medidas de correspondência entre linhas e colunas. Analise de correspondenca converte uma matriz de dados não negativos em um tipo particular de representação gráfica em que as linhas e colunas da matriz são simultaneamente representadas em dimensão reduzida, isto é, por

pontos no gráfico. Este método permite estudar as relações e semelhanças existentes entre (a) as categorias de linhas e entre as categorias de colunas de uma tabela de contingência, (b) o conjunto de categorias de linhas e o conjunto categorias de colunas.

O processo de execução da análise de correspondência requer a execução da ANOVA para obtenção das médias para as variáveis: IDH, RISCN, VULV, GO e EPI, os quais foram escolhidas através da análise de componentes para o eixo temático resiliência.

One-way ANOVA: IDH N versus Clusters 3R

```
Source        DF     SS     MS       F      P
Clusters 3R    2  61576  30788  154,44  0,000
Error        129  25716    199
Total        131  87292
S = 14,12   R-Sq = 70,54%   R-Sq(adj) = 70,08%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
1      42  82,32  13,55                                   (-*--)
2      47  62,76   8,60                        (-*--)
3      43  29,20  18,76  (-*--)
                         ----+---------+---------+---------+-----
                            32        48        64        80
Pooled StDev = 14,12
```

One-way ANOVA: VULN N versus Clusters 3R

```
Source        DF       SS       MS       F      P
Clusters 3R    2  71293,5  35646,7  380,17  0,000
Error        129  12095,6     93,8
Total        131  83389,1
S = 9,683   R-Sq = 85,50%   R-Sq(adj) = 85,27%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  -+---------+---------+---------+--------
1      42  17,712   8,091  (-*-)
2      47  44,284   8,518                 (-*)
3      43  75,550  12,026                                    (-*-)
                          -+---------+---------+---------+--------
                          16        32        48        64
Pooled StDev = 9,683
```

One-way ANOVA: RISCN N versus Clusters 3R

```
Source        DF     SS    MS      F      P
Clusters 3R    2   5194  2597  13,18  0,000
Error        129  25407   197
Total        131  30601
S = 14,03   R-Sq = 16,97%   R-Sq(adj) = 15,68%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
1      42  12,43  12,88  (------*------)
2      47  23,92  17,39                (------*------)
3      43  27,34  10,57                      (------*------)
                         ------+---------+---------+---------+---
                            12,0      18,0      24,0      30,0
Pooled StDev = 14,03
```

One-way ANOVA: GO N versus Clusters 3R

```
Source        DF     SS     MS       F      P
Clusters 3R    2  60680  30340  161,34  0,000
Error        129  24258    188
Total        131  84938
S = 13,71   R-Sq = 71,44%   R-Sq(adj) = 71,00%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
1      42  75,16  17,84                                   (--*--)
2      47  37,87  11,16           (--*-)
3      43  23,39  11,42  (--*-)
                         --------+---------+---------+---------+-
                                32        48        64        80
Pooled StDev = 13,71
```

One-way ANOVA: EPI N versus Clusters 3R

```
Source        DF       SS       MS       F      P
Clusters 3R    2  62444,0  31222,0  324,53  0,000
```

```
Error        129  12410,7     96,2
Total        131  74854,7
S = 9,808   R-Sq = 83,42%   R-Sq(adj) = 83,16%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  --------+---------+---------+---------+-
1      42  77,008  10,633                                    (-*-)
2      47  48,443   8,432                  (*-)
3      43  22,813  10,356  (-*-)
                           --------+---------+---------+---------+-
                                  32        48        64        80
Pooled StDev = 9,808
```

One-way ANOVA: IDH N versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  30392  15196  34,45  0,000
Error     129  56900    441
Total     131  87292
S = 21,00   R-Sq = 34,82%   R-Sq(adj) = 33,81%
Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65,62  12,03           (-----*-----)
AVECO   27  85,11  15,53                      (-----*----)
OTHERS  84  47,46  23,98  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90
Pooled StDev = 21,00
```

One-way ANOVA: RISCN N versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2   4413  2206  10,87  0,000
Error     129  26188   203
Total     131  30601
S = 14,25   R-Sq = 14,42%   R-Sq(adj) = 13,09%
Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  27,07  19,09                      (-------*-------)
AVECO   27  10,25   8,74  (------*------)
OTHERS  84  23,53  14,27                      (--*---)
                          ----+---------+---------+---------+-----
                            8,0      16,0      24,0      32,0
Pooled StDev = 14,25
```

One-way ANOVA: VULN N versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  39457  19728  57,93  0,000
Error     129  43933    341
Total     131  83389
S = 18,45   R-Sq = 47,32%   R-Sq(adj) = 46,50%
Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  41,83  14,32                    (----*----)
AVECO   27  13,74   6,58  (----*----)
OTHERS  84  57,44  21,60                               (-*--)
                          ------+---------+---------+---------+---
                               15        30        45        60
Pooled StDev = 18,45
```

One-way ANOVA: GO N versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  51019  25510  97,02  0,000
Error     129  33919    263
Total     131  84938
S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%
Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42,48  18,82       (---*----)
AVECO   27  83,25  13,38                                 (----*---)
OTHERS  84  33,36  16,35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75
Pooled StDev = 16,22
```

One-way ANOVA: EPI N versus Class 3R

```
Source     DF     SS     MS      F      P
```

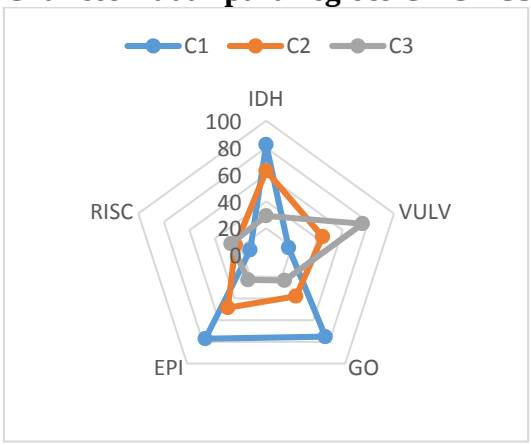
```
Class 3R    2  38872  19436  69,68  0,000
Error     129  35982    279
Total     131  74855
S = 16,70   R-Sq = 51,93%   R-Sq(adj) = 51,18%
Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  52,70  14,31         (----*----)
AVECO   27  81,39   8,14                             (---*----)
OTHERS  84  37,95  19,06  (-*--)
                          -------+---------+---------+---------+--
                                45        60        75        90
Pooled StDev = 16,70
```

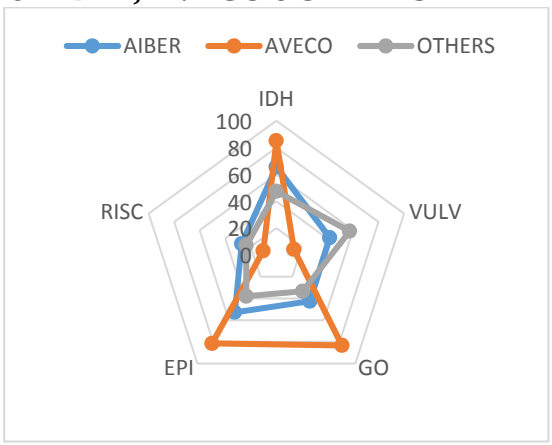
A tabela abaixo é o resumo das médias apresentadas para cada variável: IDH, RISCN, VULV, GO e EPI, cuja entrada dupla de dados será a base para a tabela de contingência a ser executada na análise de correspondência. Lembrando que C1, C2 e C3 são os países classificados anteriormente pela análise de cluster.

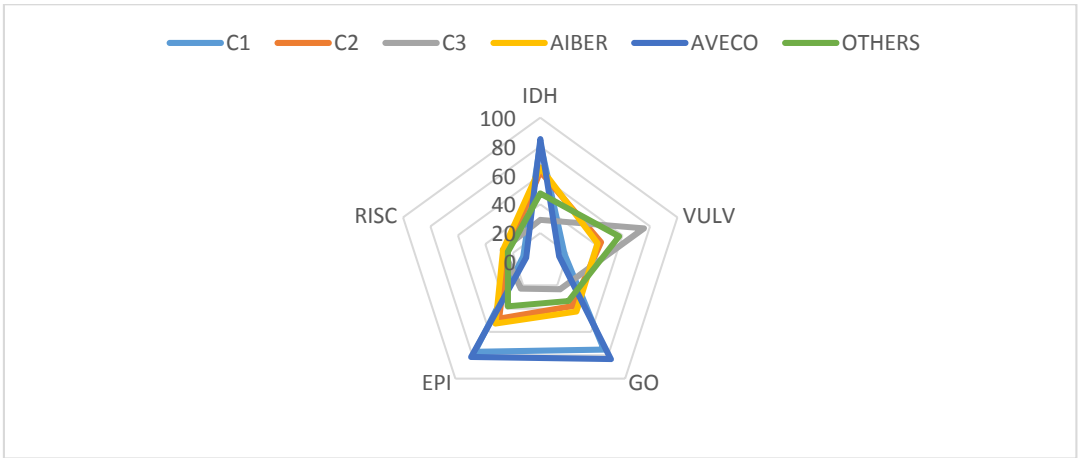
Tabela de Contingência -

XXXXX	**Mean**					XXXXX
Rows	**IDH**	**VULV**	**GO**	**EPI**	**RISC**	**Columns**
C1	82,32	17,712	75,16	77,008	12,43	IDH
C2	62,76	44,284	37,87	48,443	23,92	VUL
C3	29,2	75,55	23,39	22,813	27,34	GO
AIBER	65,62	41,83	42,48	52,7	27,07	EPI
AVECO	85,11	13,74	83,25	81,39	10,25	RISC
OTHERS	47,46	57,44	33,36	37,95	23,53	

Gráficos Radar para regiões C1 C2 C3 e AIBER, AVECO e OTHERS







Simple Correspondence Analysis: IDH_1; VULV; GO; EPI_1; RISC

Contingency Table

	IDH	VUL	GO	EPI	RIS	Total
C1	82,320	17,712	75,160	77,008	12,430	264,630
C2	62,760	44,284	37,870	48,443	23,920	217,277
C3	29,200	75,550	23,390	22,813	27,340	178,293
AIBER	65,620	41,830	42,480	52,700	27,070	229,700
AVECO	85,110	13,740	83,250	81,390	10,250	273,740
OTHERS	47,460	57,440	33,360	37,950	23,530	199,740
Total	372,470	250,556	295,510	320,304	124,540	1363,380

Chi-Square Distances

	IDH	VUL	GO	EPI	RIS	Total
C1	1,390	19,659	5,525	3,541	5,705	35,820
C2	0,195	0,475	1,807	0,133	0,836	3,445
C3	7,814	55,865	6,022	8,686	7,502	85,888
AIBER	0,131	0,003	1,072	0,030	1,766	3,003
AVECO	1,426	26,579	9,641	4,536	8,707	50,889
OTHERS	0,926	11,710	2,279	1,717	1,531	18,162
Total	11,881	114,292	26,346	18,642	26,046	197,207

Relative Inertias

	IDH	VUL	GO	EPI	RIS	Total
C1	0,007	0,100	0,028	0,018	0,029	0,182
C2	0,001	0,002	0,009	0,001	0,004	0,017
C3	0,040	0,283	0,031	0,044	0,038	0,436
AIBER	0,001	0,000	0,005	0,000	0,009	0,015
AVECO	0,007	0,135	0,049	0,023	0,044	0,258
OTHERS	0,005	0,059	0,012	0,009	0,008	0,092
Total	0,060	0,580	0,134	0,095	0,132	1,000

A inércia total é o percentual da variância que é explicada pela aplicação da Análise de Correspondência e corresponde à soma ponderada das distâncias dos pontos do conjunto a seu centróide. Pode-se entender a Análise de Correspondência como um método de decomposição da inércia total em dimensões que melhor expliquem a variabilidade dos dados. Neste caso, observamos que a componente 1 no axis 1 está com 96,66%, já com a componente 2, sobe para 99,83%.

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,1397	0,9660	0,9660	******************************
2	0,0047	0,0323	0,9983	*
3	0,0002	0,0017	1,0000	
4	0,0000	0,0000	1,0000	
Total	0,1446			

Row Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	C1	1,000	0,194	0,182	-0,367	0,994	0,187	-0,027	0,005	0,031
2	C2	0,973	0,159	0,017	0,091	0,525	0,009	0,084	0,448	0,242
3	C3	1,000	0,131	0,436	0,688	0,982	0,443	-0,091	0,017	0,233
4	AIBER	0,954	0,168	0,015	0,050	0,194	0,003	0,100	0,759	0,357
5	AVECO	1,000	0,201	0,258	-0,428	0,983	0,263	-0,055	0,017	0,132
6	OTHERS	0,996	0,147	0,092	0,301	0,995	0,095	-0,013	0,002	0,005

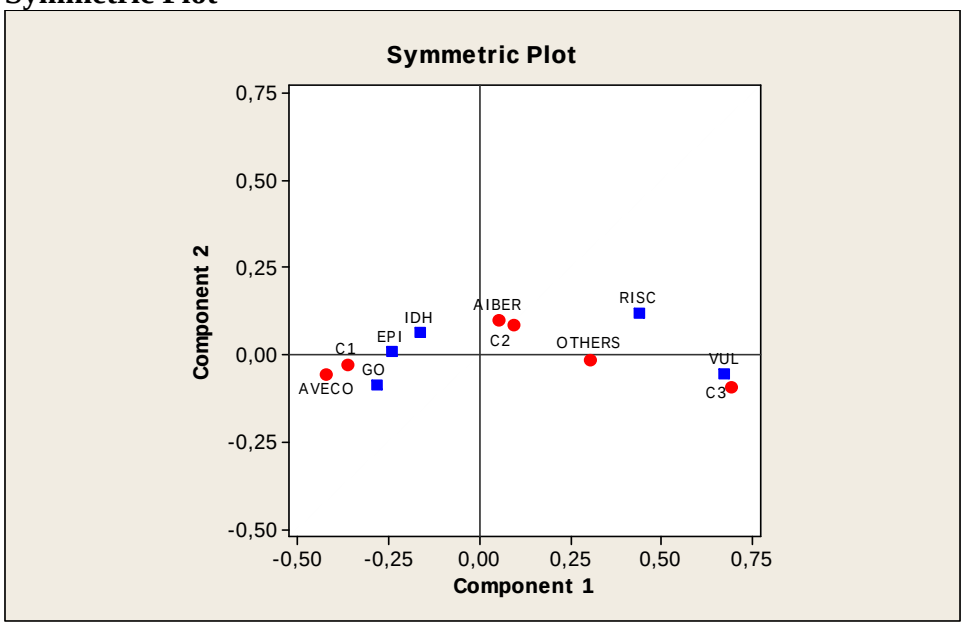
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	IDH	0,993	0,273	0,060	-0,167	0,870	0,054	0,063	0,123	0,230
2	VUL	1,000	0,184	0,580	0,673	0,993	0,596	-0,057	0,007	0,126
3	GO	0,998	0,217	0,134	-0,285	0,909	0,126	-0,089	0,089	0,367
4	EPI	1,000	0,235	0,095	-0,241	0,999	0,098	0,007	0,001	0,003
5	RIS	0,994	0,091	0,132	0,440	0,926	0,127	0,119	0,067	0,275

A aplicação da técnica de análise de correspondência demonstra que as duas componentes correspondem por 99,83% do modelo e demonstra que as variáveis: VULV tem maior correspondência para com os países do cluster C3 e que os índices de governança (GO) estão para os países da AVECO e C1.

O gráfico Symmetric Plot demonstra que a variável VULV e países que compõem o cluster C3 tem maior correspondência, enquanto que os países do grupo AIBER, OTHERS e C2 tem correspondência com a variável RISC e finalmente, os países do AIBER e C1 tem sua maior correspondência com o índice de Governança (GO).

Symmetric Plot



13. REGIÃO IBE(AIBER)

A presente seção deste trabalho tem por objetivo realizar análise exploratória dos dados sobre fatores de **resiliência** para 21 países, da região IBER(AIBER) para o eixo temático: Resiliência.

A aplicação de diversos técnicas estatísticas tiveram por objetivo análise da base de dados em suas variáveis e indivíduos, para posterior categorização em modelos para a diferenciação ou explicação do desenvolvimento do conceito de resiliência.

14. OS DADOS DA REGIÃO IBE(AIBER)

Pais	IPS	IDH	RISC	EXP	VULN	SUSC	SDIF	CADP	GO	EPI	HPI	GINI
Argentina	68,278	77,595	9,269	13,366	29,681	17,789	39,881	27,741	35,302	44,945	75,918	48,930
Bolivia	54,457	54,366	14,291	12,215	61,750	51,229	76,804	52,478	30,018	46,288	50,638	56,942
Brazil	67,164	67,051	11,558	13,325	41,604	25,345	52,291	44,159	45,427	49,884	73,208	72,677
Chile	78,541	79,901	37,408	56,572	25,998	16,409	37,753	19,622	80,674	74,379	75,503	71,086
Colombia	62,257	61,614	20,901	22,027	49,353	30,882	67,641	45,990	37,787	46,707	89,661	74,922
Costa Rica	81,147	70,181	59,675	80,113	33,564	21,054	48,701	26,866	64,569	57,915	100,00	62,082
Cuba	51,168	78,748	19,387	29,316	26,472	15,399	35,335	25,075	28,758	52,917	81,060	37,067
Rep. Domin.	54,691	59,802	38,146	40,804	49,991	32,447	55,465	55,582	37,449	50,274	67,702	54,436
Ecuador	63,893	61,614	23,855	26,691	36,390	32,582	65,206	34,090	29,149	57,929	72,121	56,759
El Salvador	57,692	53,542	58,900	59,903	55,150	36,402	68,074	58,050	41,425	36,626	87,576	44,311
Guatemala	51,707	47,941	72,046	67,373	63,282	46,163	77,976	62,149	28,169	42,793	82,688	71,842
Honduras	51,546	46,129	35,561	34,484	57,830	41,669	80,321	47,065	26,115	43,963	80,552	85,021
Mexico	60,766	69,028	18,833	22,027	41,933	22,753	62,320	37,194	41,800	52,860	73,116	60,673
Nicaragua	53,433	45,634	50,591	49,061	58,997	45,978	79,527	46,667	30,530	46,057	83,176	54,567
Panama	71,855	70,511	23,043	27,297	41,495	29,367	54,582	36,498	47,875	55,474	84,950	70,668
Paraguay	54,008	55,848	9,490	8,278	56,354	36,772	74,874	54,010	27,325	30,069	56,061	60,517
Peru	60,550	65,898	21,196	23,158	46,910	32,144	64,340	40,199	39,197	38,446	71,849	53,523
Portugal	86,071	79,901	9,010	16,152	19,581	12,521	19,426	23,423	74,026	82,857	38,814	35,569
Spain	86,574	87,644	7,496	14,739	16,408	9,441	25,956	9,512	70,133	88,619	51,808	28,523
Uruguay	80,715	74,629	10,451	16,495	25,123	18,664	23,791	29,493	68,430	50,809	40,367	43,058
Venezuela	56,039	70,346	17,356	20,634	40,492	22,164	66,071	29,055	8,006	56,860	82,713	52,062

Fonte: preparado pelo autor.

15 ANOVA PARA A REGIÃO IBE

ANOVA - Análise de variância é semelhante à regressão, na medida em que é usada para investigar e modelar a relação entre uma variável de resposta e uma ou mais variáveis de previsão. Entretanto, a análise de variância difere da regressão no seguinte: as variáveis de previsão são consideradas categóricas e ou qualitativos e não há suposição sobre a natureza da relação, portanto, não incluem coeficientes para as variáveis. ANOVA-análise de variância executa o F-teste para igualdade da amostra ou então, homogeneidade das variâncias (Bartlet e Levene test). Isto porque, se baseiam no pressuposto de que as amostras de diferentes populações têm a mesma variância.

O resultado dos testes com a utilização do modelo ANOVA para o eixo temático resiliência, demonstram um F relativamente baixo quando comparado com os resultados dos 21 países da região IBE (AIBER), se deve ao tamanho da amostra. As variáveis de maior representatividade são HPI e EPI.

One-way ANOVA: VULN N_1 versus Cluster 1

```
Source     DF    SS   MS     F      P
Cluster 1   4  2398  600  5,64  0,005
Error      16  1701  106
Total      20  4099
S = 10,31   R-Sq = 58,51%   R-Sq(adj) = 48,13%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
1      15  44,88  10,96                    (-*--)
2       2  59,05   3,82                      (-------*------)
3       1  26,00      *   (----------*----------)
```

```
4       2  17,99   2,24  (-------*-------)
5       1  25,12      *   (----------*---------)
                          ---------+---------+---------+---------+
                                  20        40        60        80
Pooled StDev = 10,31
```

One-way ANOVA: HPI N_1 versus Cluster 1

```
Source     DF      SS      MS      F      P
Cluster 1   4  4193,9  1048,5  15,42  0,000
Error      16  1087,8    68,0
Total      20  5281,6
S = 8,245   R-Sq = 79,40%   R-Sq(adj) = 74,26%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean  StDev  ---------+---------+---------+---------+
1      15  80,419  8,403                              (-*-)
2       2  53,350  3,835           (------*-----)
3       1  75,503      *                    (--------*-------)
4       2  45,311  9,188       (------*-----)
5       1  40,367      *  (--------*--------)
                          ---------+---------+---------+---------+
                                  40        60        80       100
Pooled StDev = 8,245
```

One-way ANOVA: IPS N_1 versus Cluster 1

```
Source     DF      SS     MS     F      P
Cluster 1   4  1827,2  456,8  7,10  0,002
Error      16  1029,6   64,4
Total      20  2856,8
S = 8,022   R-Sq = 63,96%   R-Sq(adj) = 54,95%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean  StDev  --+---------+---------+---------+-------
1      15  60,812  8,575           (--*-)
2       2  54,233  0,318  (-------*-------)
3       1  78,541      *              (----------*-----------)
4       2  86,323  0,356                       (-------*-------)
5       1  80,715      *                (-----------*----------)
                          --+---------+---------+---------+-------
                            45        60        75        90
Pooled StDev = 8,022
```

One-way ANOVA: IDH N_1 versus Cluster 1

```
Source     DF    SS   MS     F      P
Cluster 1   4  1265  316  3,10  0,046
Error      16  1632  102
Total      20  2897
S = 10,10   R-Sq = 43,67%   R-Sq(adj) = 29,59%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
1      15  63,04  10,69            (--*---)
2       2  55,11   1,05  (--------*---------)
3       1  79,90      *             (------------*------------)
4       2  83,77   5,48                   (--------*---------)
5       1  74,63      *          (-------------*------------)
                          -----+---------+---------+---------+----
                              48        64        80        96
Pooled StDev = 10,10
```

One-way ANOVA: EPI N_1 versus Cluster 1

```
Source     DF      SS     MS      F      P
Cluster 1   4  3293,8  823,4  16,42  0,000
Error      16   802,3   50,1
Total      20  4096,1
S = 7,081   R-Sq = 80,41%   R-Sq(adj) = 75,52%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  ------+---------+---------+---------+---
1      15  48,910   6,836          (*-)
2       2  38,179  11,469  (----*----)
3       1  74,379       *                   (------*-------)
4       2  85,738   4,075                            (----*----)
5       1  50,809       *      (------*-------)
                          ------+---------+---------+---------+---
                                40        60        80       100
Pooled StDev = 7,081
```

16. ANÁLISE DOS COMPONENTES PRINCIPAIS

A melhor utilização da análise dos componentes principais é formar um número menor de variáveis não correlacionadas, a partir de um grande conjunto de dados, portanto, é explicar a quantidade máxima de variância com o menor número de componentes principais. Análise de componentes principais é comumente usada nas ciências sociais, pesquisa de mercado, e outras indústrias que utilizam grandes conjuntos de dados. Para o eixo de temático: Resiliência há 12 variáveis para uma amostra de 21 países, alocados na região IBE(AIBER) onde se observa com aplicação da análise de pelo menos 4 componentes.

REGIÃO IBE(AIBER) - Principal Component Analysis: IPS N_1; IDH N_1; RISCN N_1; EXP N_1; VULN N_1; S

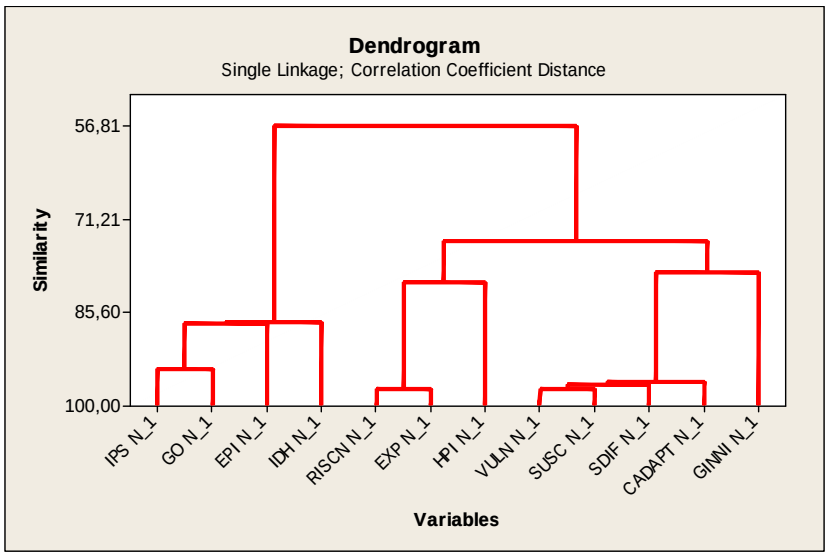
O primeiro componente principal tem variância (eigenvalue) 7,2766, sendo responsável por 60,6% da variância total. As variáveis com os respectivos coeficientes listados em PC1 demonstram como seria a equação que representaria o eixo temático: resiliência, como segue: PC1=-0,315IPS-0,351IDH+0,204RISC+0,118EXP+0,359SUSC+0,351SDIF+0,333CADAPT-0,270GO-0,300EPI+0,178HPI. Assim, sucessivamente, tem-se para os demais componentes de PC2, PC3 e PC4. Note que para PC2 tem variância (eigenvalue) 2,2113, sendo responsável por 18,4% da variabilidade total e que PC1 e PC2 teriam 79,1% e assim, sucessivamente, PC3 e PC4 acumulariam 93,3%.

Eigenanalysis of the Correlation Matrix

Eigenvalue	7,2766	2,2113	0,9365	0,7736	0,3849	0,1585	0,0964	0,0635
Proportion	0,606	0,184	0,078	0,064	0,032	0,013	0,008	0,005
Cumulative	0,606	0,791	0,869	0,933	0,965	0,978	0,986	0,992

Eigenvalue	0,0429	0,0390	0,0103	0,0065
Proportion	0,004	0,003	0,001	0,001
Cumulative	0,995	0,999	0,999	1,000

Variable	PC1	PC2	PC3	PC4
IPS N_1	-0,315	0,188	-0,259	0,265
IDH N_1	-0,351	0,013	0,250	0,008
RISCN N_1	0,204	0,527	-0,151	-0,258
EXP N_1	0,118	0,617	-0,087	-0,219
VULN N_1	0,359	-0,068	-0,168	0,043
SUSC N_1	0,340	-0,075	-0,306	0,025
SDIF N_1	0,351	-0,057	0,054	0,107
CADAPT N_1	0,333	-0,085	-0,292	-0,032
GO N_1	-0,270	0,273	-0,468	0,236
EPI N_1	-0,300	0,163	-0,100	0,072
HPI N_1	0,178	0,403	0,625	0,065
GINNI N_1	0,219	0,148	0,099	0,858



Stepwise Regression: PC1_1 versus IPS N_1; IDH N_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15
Response is PC1_1 on 12 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-7,639	-9,439	-5,778	1,322	2,308	2,636
VULN N_1	0,1826	0,1712	0,1469	0,0898	0,0804	0,0736
T-Value	17,21	21,83	14,36	6,12	6,24	5,86
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
HPI N_1		0,0315	0,0286	0,0289	0,0283	0,0253
T-Value		4,56	4,95	7,25	8,34	7,15
P-Value		0,000	0,000	0,000	0,000	0,000
IPS N_1			-0,0379	-0,0403	-0,0334	-0,0376
T-Value			-3,06	-4,71	-4,34	-5,01
P-Value			0,007	0,000	0,001	0,000
IDH N_1				-0,070	-0,072	-0,072
T-Value				-4,45	-5,40	-5,83
P-Value				0,000	0,000	0,000
EPI N_1					-0,0163	-0,0166
T-Value					-2,70	-2,96
P-Value					0,017	0,010
GINNI N_1						0,0084
T-Value						1,84
P-Value						0,088
S	0,680	0,476	0,393	0,271	0,230	0,213
R-Sq	93,97	97,20	98,20	99,19	99,46	99,56
R-Sq(adj)	93,65	96,89	97,88	98,99	99,28	99,37

Step	7	8	9	10	11	12
Constant	2,42872	1,20546	0,52575	-0,02761	-0,60985	-0,60728
VULN N_1	0,07294	0,07445	0,06188	0,05040	0,02400	0,02510
T-Value	6,45	9,35	7,46	5,56	8,67	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
HPI N_1	0,01814	0,00927	0,01110	0,01385	0,01117	0,01098
T-Value	3,85	2,29	3,22	4,22	14,78	*
P-Value	0,002	0,041	0,008	0,002	0,000	*
IPS N_1	-0,04343	-0,02263	-0,02504	-0,02434	-0,02751	-0,02640
T-Value	-5,93	-3,00	-3,95	-4,39	-21,93	*
P-Value	0,000	0,011	0,002	0,001	0,000	*
IDH N_1	-0,05908	-0,04665	-0,03732	-0,03564	-0,02774	-0,02915
T-Value	-4,62	-4,87	-4,25	-4,62	-15,36	*
P-Value	0,000	0,000	0,001	0,001	0,000	*
EPI N_1	-0,01964	-0,02178	-0,02312	-0,02018	-0,02159	-0,02096
T-Value	-3,73	-5,82	-7,32	-6,53	-31,08	*
P-Value	0,003	0,000	0,000	0,000	0,000	*
GINNI N_1	0,01205	0,01686	0,01703	0,01677	0,01557	0,01521
T-Value	2,68	4,95	6,01	6,77	27,94	*
P-Value	0,019	0,000	0,000	0,000	0,000	*
RISCN N_1	0,00855	0,01760	0,01733	0,01573	0,01723	0,01069
T-Value	2,06	4,67	5,52	5,53	26,87	*
P-Value	0,059	0,001	0,000	0,000	0,000	*
GO N_1		-0,01790	-0,01687	-0,01752	-0,01316	-0,01436
T-Value		-3,78	-4,26	-5,05	-15,81	*

P-Value		0,003	0,001	0,000	0,000	*
SUSC N_1			0,02255	0,02815	0,02898	0,02883
T-Value			2,52	3,41	15,78	*
P-Value			0,029	0,007	0,000	*
CADAPT N_1				0,01175	0,02236	0,02372
T-Value				2,10	15,34	*
P-Value				0,062	0,000	*
SDIF N_1					0,01880	0,01831
T-Value					13,90	*
P-Value					0,000	*
EXP N_1						0,00582
T-Value						*
P-Value						*
S	0,192	0,135	0,112	0,0981	0,0218	0,000000
R-Sq	99,67	99,85	99,90	99,93	100,00	100,00
R-Sq(adj)	99,49	99,75	99,83	99,87	99,99	100,00

Stepwise Regression: PC2_1 versus IPS N_1; IDH N_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2_1 on 12 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-2,096	-4,770	-6,774	-6,817	-6,395	-6,657
EXP N_1	0,0673	0,0686	0,0571	0,0566	0,0511	0,0520
T-Value	10,02	17,49	18,66	23,53	24,16	28,87
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
IPS N_1		0,0409	0,0510	0,0395	0,0179	0,0205
T-Value		6,15	11,79	8,26	2,94	3,95
P-Value		0,000	0,000	0,000	0,010	0,001
HPI N_1			0,0236	0,0249	0,0304	0,0277
T-Value			5,81	7,75	11,73	11,64
P-Value			0,000	0,000	0,000	0,000
EPI N_1				0,0135	0,0144	0,0156
T-Value				3,41	5,22	6,63
P-Value				0,004	0,000	0,000
GO N_1					0,0163	0,0142
T-Value					4,23	4,28
P-Value					0,001	0,001
GINNI N_1						0,0050
T-Value						2,71
P-Value						0,017
S	0,608	0,355	0,211	0,166	0,116	0,0969
R-Sq	84,10	94,87	98,28	99,01	99,55	99,70
R-Sq(adj)	83,26	94,31	97,98	98,76	99,40	99,58

Step	7	8	9	10	11	12
Constant	-6,152	-6,244	-5,703	-5,650	-5,679	-5,495
EXP N_1	0,0537	0,0404	0,0335	0,0328	0,0331	0,0305
T-Value	33,54	9,84	12,06	12,13	15,25	20,02
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
IPS N_1	0,0194	0,0197	0,0155	0,0149	0,0158	0,0151
T-Value	4,53	6,16	7,58	7,58	9,83	14,51
P-Value	0,001	0,000	0,000	0,000	0,000	0,000

HPI N_1	0,02609	0,02815	0,02661	0,02639	0,02543	0,02465
T-Value	12,81	17,20	26,56	26,85	29,55	42,35
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
EPI N_1	0,01449	0,01532	0,01269	0,01230	0,01251	0,01129
T-Value	7,36	10,29	12,46	12,95	16,34	19,74
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
GO N_1	0,01062	0,01167	0,01472	0,01549	0,01478	0,01547
T-Value	3,54	5,16	10,13	12,35	14,26	22,61
P-Value	0,004	0,000	0,000	0,000	0,000	0,000
GINNI N_1	0,00826	0,00866	0,00958	0,00944	0,00962	0,00988
T-Value	4,33	6,07	11,26	11,20	14,17	22,51
P-Value	0,001	0,000	0,000	0,000	0,000	0,000
SDIF N_1	-0,0061	-0,0088	-0,0018			
T-Value	-2,80	-4,87	-1,04			
P-Value	0,015	0,000	0,321			
RISCN N_1		0,0140	0,0229	0,0237	0,0239	0,0274
T-Value		3,37	7,60	8,12	10,24	15,98
P-Value		0,006	0,000	0,000	0,000	0,000
VULN N_1			-0,0148	-0,0173	-0,0128	-0,0092
T-Value			-4,94	-9,64	-5,88	-5,65
P-Value			0,000	0,000	0,000	0,000
SUSC N_1					-0,0056	-0,0073
T-Value					-2,77	-5,40
P-Value					0,018	0,000
CADAPT N_1						-0,0045
T-Value						-4,11
P-Value						0,002
S	0,0794	0,0592	0,0345	0,0346	0,0277	0,0177
R-Sq	99,81	99,90	99,97	99,97	99,98	99,99
R-Sq(adj)	99,71	99,84	99,95	99,95	99,97	99,99
Step	13	14				
Constant	-5,514	-5,599				
EXP N_1	0,03072	0,03044				
T-Value	84,69	*				
P-Value	0,000	*				
IPS N_1	0,01575	0,01570				
T-Value	62,41	*				
P-Value	0,000	*				
HPI N_1	0,02495	0,02482				
T-Value	177,95	*				
P-Value	0,000	*				
EPI N_1	0,01149	0,01139				
T-Value	83,97	*				
P-Value	0,000	*				
GO N_1	0,01452	0,01452				
T-Value	81,35	*				
P-Value	0,000	*				
GINNI N_1	0,01019	0,01027				
T-Value	95,23	*				
P-Value	0,000	*				
SDIF N_1	-0,00322	-0,00296				
T-Value	-12,96	*				
P-Value	0,000	*				

RISCN N_1	0,02712	0,02761
T-Value	66,62	*
P-Value	0,000	*
VULN N_1	-0,00448	-0,00474
T-Value	-8,41	*
P-Value	0,000	*
SUSC N_1	-0,00681	-0,00637
T-Value	-21,04	*
P-Value	0,000	*
CADAPT N_1	-0,00618	-0,00605
T-Value	-21,22	*
P-Value	0,000	*
IDH N_1		0,00108
T-Value		*
P-Value		*
S	0,00421	0,000000
R-Sq	100,00	100,00
R-Sq(adj)	100,00	100,00

Stepwise Regression: PC3_1 versus IPS N_1; IDH N_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3_1 on 12 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-2,60608	-3,71500	-3,28459	-2,57710	0,03524	-0,02376
HPI N_1	0,0360	0,0675	0,0745	0,0700	0,0444	0,0457
T-Value	3,31	8,62	9,77	9,24	8,68	11,80
P-Value	0,004	0,000	0,000	0,000	0,000	0,000
RISCN N_1		-0,0431	-0,0427	-0,0408	-0,0188	-0,0195
T-Value		-6,47	-7,16	-7,15	-4,56	-5,43
P-Value		0,000	0,000	0,000	0,000	0,000
GINNI N_1			-0,0166	-0,0177	0,0017	
T-Value			-2,33	-2,64	0,39	
P-Value			0,032	0,018	0,701	
GO N_1				-0,0088	-0,0310	-0,0303
T-Value				-1,81	-7,99	-9,04
P-Value				0,089	0,000	0,000
SUSC N_1					-0,0534	-0,0515
T-Value					-7,23	-9,27
P-Value					0,000	0,000
S	0,791	0,446	0,399	0,375	0,183	0,178
R-Sq	36,58	80,92	85,55	88,01	97,32	97,30
R-Sq(adj)	33,24	78,80	83,00	85,02	96,43	96,62

Step	7	8	9	10	11	12
Constant	0,31208	1,45409	0,27203	0,07576	0,45434	0,57323
HPI N_1	0,04432	0,04493	0,04476	0,04037	0,03785	0,03867
T-Value	14,60	20,93	22,82	20,71	38,50	58,27
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
RISCN N_1	-0,01760	-0,01920	-0,01780	-0,01470	-0,01250	-0,01293
T-Value	-6,18	-9,37	-8,88	-8,41	-14,24	-22,52
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
GINNI N_1				0,00551	0,00651	0,00679
T-Value				3,39	8,45	13,57
P-Value				0,005	0,000	0,000

GO N_1	-0,03208	-0,02193	-0,02299	-0,02587	-0,02641	-0,02634
T-Value	-12,07	-6,97	-7,87	-11,08	-24,25	-37,58
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
SUSC N_1	-0,0381	-0,0377	-0,0295	-0,0312	-0,0306	-0,0262
T-Value	-6,50	-9,10	-5,21	-7,37	-15,52	-15,72
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
CADAPT N_1	-0,0156	-0,0209	-0,0183	-0,0190	-0,0239	-0,0211
T-Value	-3,39	-5,96	-5,33	-7,41	-17,05	-18,73
P-Value	0,004	0,000	0,000	0,000	0,000	0,000
IPS N_1		-0,0215	-0,0211	-0,0210	-0,0195	-0,0205
T-Value		-4,02	-4,33	-5,79	-11,44	-18,25
P-Value		0,001	0,001	0,000	0,000	0,000
IDH N_1			0,0130	0,0176	0,0195	0,0187
T-Value			1,95	3,44	8,12	12,02
P-Value			0,073	0,005	0,000	0,000
EPI N_1					-0,00646	-0,00631
T-Value					-6,66	-10,10
P-Value					0,000	0,000
VULN N_1						-0,0075
T-Value						-4,07
P-Value						0,002
S	0,138	0,0975	0,0890	0,0662	0,0308	0,0198
R-Sq	98,47	99,29	99,45	99,72	99,94	99,98
R-Sq(adj)	97,96	98,98	99,15	99,53	99,90	99,96

Step	13	14
Constant	0,5816	0,4948
HPI N_1	0,03884	0,03846
T-Value	76,08	*
P-Value	0,000	*
RISCN N_1	-0,00868	-0,00790
T-Value	-5,58	*
P-Value	0,000	*
GINNI N_1	0,00705	0,00691
T-Value	17,94	*
P-Value	0,000	*
GO N_1	-0,02563	-0,02488
T-Value	-43,39	*
P-Value	0,000	*
SUSC N_1	-0,02614	-0,02600
T-Value	-20,50	*
P-Value	0,000	*
CADAPT N_1	-0,02221	-0,02074
T-Value	-23,60	*
P-Value	0,000	*
IPS N_1	-0,02113	-0,02170
T-Value	-23,78	*
P-Value	0,000	*
IDH N_1	0,01951	0,02081
T-Value	15,97	*
P-Value	0,000	*
EPI N_1	-0,00670	-0,00696
T-Value	-13,49	*
P-Value	0,000	*

VULN N_1	-0,00772	-0,01174
T-Value	-5,51	*
P-Value	0,000	*
EXP N_1	-0,00381	-0,00430
T-Value	-2,85	*
P-Value	0,019	*
SDIF N_1		0,00284
T-Value		*
P-Value		*
S	0,0151	0,000000
R-Sq	99,99	100,00
R-Sq(adj)	99,98	100,00

Stepwise Regression: PC4_1 versus IPS N_1; IDH N_1; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15
Response is PC4_1 on 12 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-2,622	-5,725	-5,182	-4,821	-5,355	-5,453
GINNI N_1	0,04607	0,05647	0,06349	0,06343	0,05994	0,06025
T-Value	5,01	8,77	35,42	39,63	44,88	63,37
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
IPS N_1		0,0390	0,0329	0,0237	0,0269	0,0237
T-Value		5,02	15,47	5,37	8,76	10,17
P-Value		0,000	0,000	0,000	0,000	0,000
RISCN N_1			-0,02026	-0,02134	-0,02295	-0,02312
T-Value			-15,23	-16,71	-24,52	-34,74
P-Value			0,000	0,000	0,000	0,000
GO N_1				0,0062	0,0094	0,0097
T-Value				2,31	4,83	7,02
P-Value				0,035	0,000	0,000
SDIF N_1					0,0077	0,0085
T-Value					4,47	6,89
P-Value					0,000	0,000
EPI N_1						0,0044
T-Value						3,98
P-Value						0,001
S	0,593	0,393	0,106	0,0943	0,0638	0,0453
R-Sq	56,89	82,05	98,77	99,08	99,61	99,81
R-Sq(adj)	54,62	80,06	98,56	98,85	99,47	99,74

Step	7	8	9	10	11	12
Constant	-5,452	-5,612	-5,666	-5,584	-5,604	-5,656
GINNI N_1	0,06081	0,05947	0,05933	0,05950	0,05951	0,05956
T-Value	70,18	164,05	228,93	283,35	907,79	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
IPS N_1	0,02409	0,02202	0,02217	0,02172	0,02216	0,02213
T-Value	11,79	27,16	38,58	45,85	143,32	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
RISCN N_1	-0,01732	-0,01384	-0,01504	-0,01367	-0,01383	-0,01353
T-Value	-6,71	-13,09	-18,39	-17,15	-55,47	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
GO N_1	0,00973	0,01275	0,01258	0,01287	0,01254	0,01254
T-Value	8,00	22,20	30,79	38,59	114,68	*
P-Value	0,000	0,000	0,000	0,000	0,000	*

SDIF N_1	0,00729	0,00757	0,00642	0,00643	0,00542	0,00558
T-Value	6,03	16,44	14,08	18,11	35,61	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
EPI N_1	0,00474	0,00526	0,00552	0,00506	0,00511	0,00505
T-Value	4,82	13,89	19,91	18,89	60,98	*
P-Value	0,000	0,000	0,000	0,000	0,000	*
EXP N_1	-0,00523	-0,01043	-0,00971	-0,01082	-0,01061	-0,01079
T-Value	-2,30	-9,99	-12,72	-15,26	-47,74	*
P-Value	0,038	0,000	0,000	0,000	0,000	*
HPI N_1		0,00367	0,00432	0,00417	0,00408	0,00401
T-Value		8,84	12,54	15,22	47,54	*
P-Value		0,000	0,000	0,000	0,000	*
SUSC N_1			0,00257	0,00276	0,00184	0,00211
T-Value			3,61	4,94	9,30	*
P-Value			0,004	0,001	0,000	*
CADAPT N_1				-0,00125	-0,00236	-0,00228
T-Value				-2,86	-13,24	*
P-Value				0,017	0,000	*
VULN N_1					0,00316	0,00300
T-Value					9,68	*
P-Value					0,000	*
IDH N_1						0,00066
T-Value						*
P-Value						*
S	0,0396	0,0150	0,0106	0,00827	0,00258	0,000000
R-Sq	99,87	99,98	99,99	100,00	100,00	100,00
R-Sq(adj)	99,80	99,97	99,99	99,99	100,00	100,00

17 ANALISE DE CLUSTERS PARA A REGIÃO IBE (AIBER)

A análise de conglomerados é não uma técnica de inferência estatística em que parâmetros para uma amostra são avaliados como sendo possivelmente representativos de uma população. Ao invés disto, a análise de conglomerados é uma metodologia objetiva para quantificar características estruturais de um conjunto de observações. Para fins de análise de clusters, os 4 componentes principais foram utilizados para avaliação do grau de similaridade e da utilização destes conglomerados para a análise de discriminante, logística e árvores. A determinação ou partição de grupos foi dada para 5 grupos e o método *linkage* ou da distância mínima entre dois objetos e assim sucessivamente.

Cluster Analysis of Observations: VULN N_1; HPI N_1; IPS N_1; IDH N_1; EPI N_1

Manhattan Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	95,8279	8,2670	12	14	12	2
2	19	94,5840	10,7318	11	12	11	3
3	18	94,0600	11,7701	3	13	3	2
4	17	89,6003	20,6070	15	21	15	2
5	16	89,3603	21,0825	3	15	3	4
6	15	88,8226	22,1482	3	9	3	5
7	14	87,8863	24,0034	3	17	3	6
8	13	85,3794	28,9709	2	16	2	2
9	12	84,9382	29,8451	10	11	10	4
10	11	84,7718	30,1748	18	19	18	2
11	10	84,5568	30,6009	5	10	5	5
12	9	84,3494	31,0116	3	8	3	7
13	8	84,2390	31,2305	1	3	1	8
14	7	83,4020	32,8889	1	7	1	9
15	6	82,5857	34,5066	1	5	1	14
16	5	82,3151	35,0426	1	6	1	15
17	4	80,5805	38,4799	1	2	1	17
18	3	76,7350	46,0997	1	4	1	18
19	2	74,8825	49,7705	18	20	18	3
20	1	70,1666	59,1151	1	18	1	21

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	15	5954,25	18,7905	32,5099
Cluster2	2	162,00	8,9999	8,9999
Cluster3	1	0,00	0,0000	0,0000
Cluster4	2	136,16	8,2510	8,2510
Cluster5	1	0,00	0,0000	0,0000

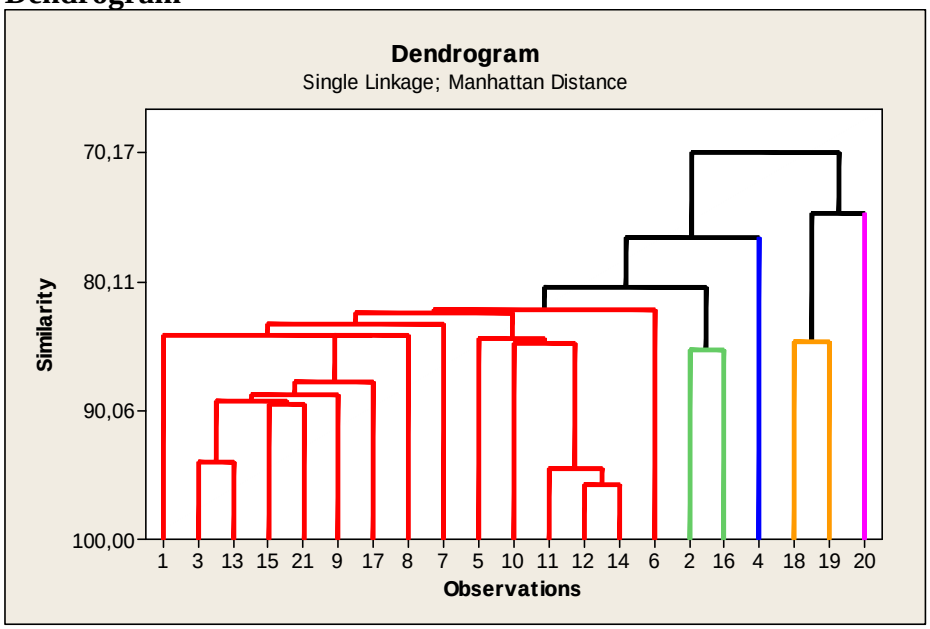
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
VULN N_1	44,8763	59,0520	25,9982	17,9945	25,1231	41,8266
HPI N_1	80,4194	53,3495	75,5033	45,3109	40,3665	72,3563
IPS N_1	60,8124	54,2326	78,5406	86,3228	80,7153	64,4072
IDH N_1	63,0423	55,1071	79,9012	83,7727	74,6293	65,6154
EPI N_1	48,9101	38,1788	74,3790	85,7380	50,8088	52,6987

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,0000	33,9874	40,3451	66,2726	50,2834
Cluster2	33,9874	0,0000	64,0267	76,5749	50,6132
Cluster3	40,3451	64,0267	0,0000	34,3543	42,7017
Cluster4	66,2726	76,5749	34,3543	0,0000	37,5547
Cluster5	50,2834	50,6132	42,7017	37,5547	0,0000

Dendrogram



Cluster Analysis of Observations: EXP N_1; IPS N_1; HPI N_1; EPI N_1; GO N_1

Manhattan Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	91,8046	17,6514	7	12	7	2
2	19	91,2116	18,9285	1	3	1	2
3	18	90,8861	19,6296	13	17	13	2
4	17	89,8812	21,7938	1	13	1	4
5	16	88,5941	24,5660	18	19	18	2
6	15	88,1151	25,5978	7	14	7	3
7	14	87,8587	26,1499	7	11	7	4
8	13	87,6933	26,5062	1	9	1	5
9	12	86,9062	28,2015	1	5	1	6
10	11	86,6646	28,7218	2	16	2	2
11	10	86,2143	29,6916	1	7	1	10
12	9	83,2268	36,1262	1	8	1	11
13	8	82,8758	36,8822	1	15	1	12
14	7	82,4654	37,7661	1	10	1	13
15	6	81,4746	39,9001	1	21	1	14
16	5	79,1553	44,8954	18	20	18	3
17	4	78,2347	46,8781	1	2	1	16
18	3	65,8787	73,4906	1	18	1	19
19	2	61,3642	83,2139	4	6	4	2
20	1	55,2917	96,2928	1	4	1	21

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	14	6624,59	20,3751	37,1575
Cluster2	2	157,71	8,8801	8,8801
Cluster3	1	0,00	0,0000	0,0000
Cluster4	1	0,00	0,0000	0,0000
Cluster5	3	969,97	17,1383	23,9477

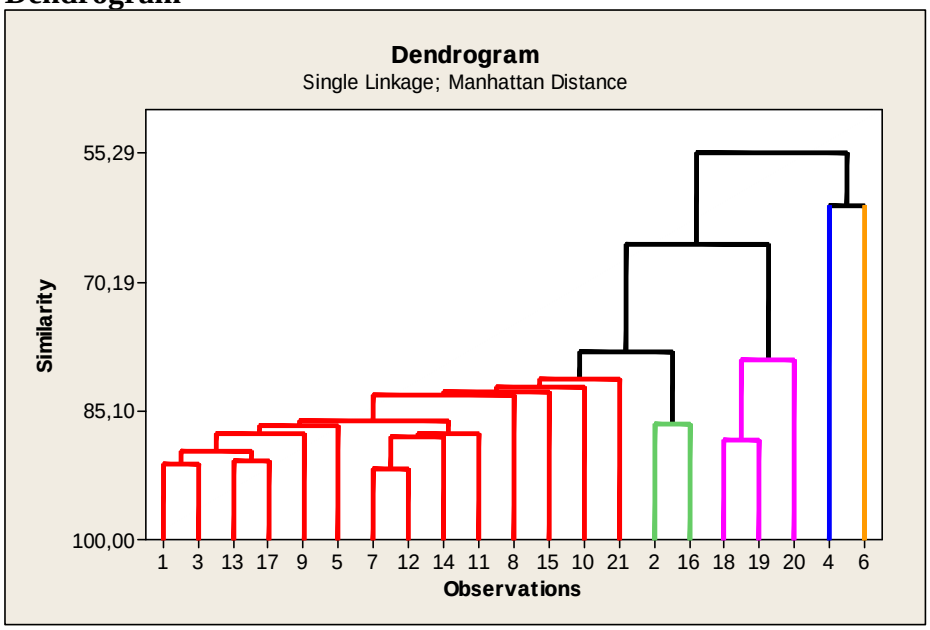
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
EXP N_1	32,1046	10,2463	56,5718	80,113	15,7951	31,1442
IPS N_1	59,3599	54,2326	78,5406	81,147	84,4536	64,4072
HPI N_1	79,0208	53,3495	75,5033	100,000	43,6628	72,3563
EPI N_1	48,2669	38,1788	74,3790	57,915	74,0949	52,6987
GO N_1	34,0706	28,6718	80,6744	64,569	70,8631	42,4840

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,0000	35,9723	61,9086	65,1370	64,5497
Cluster2	35,9723	0,0000	85,1026	97,2631	64,0938
Cluster3	61,9086	85,1026	0,0000	41,1283	52,9892
Cluster4	65,1370	97,2631	41,1283	0,0000	87,3100
Cluster5	64,5497	64,0938	52,9892	87,3100	0,0000

Dendrogram



Cluster Analysis of Observations: HPI N_1; RISCN N_1; GINNI N_1; GO N_1; ...

Manhattan Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	92,5227	16,6505	9	17	9	2
2	19	90,0213	22,2209	18	20	18	2
3	18	89,8021	22,7089	5	15	5	2
4	17	89,7728	22,7741	9	13	9	3
5	16	89,1947	24,0616	8	9	8	4
6	15	88,5086	25,5892	3	8	3	5
7	14	87,1895	28,5266	18	19	18	3
8	13	86,1019	30,9487	2	16	2	2
9	12	85,7613	31,7070	3	5	3	7
10	11	84,0260	35,5713	1	3	1	8
11	10	83,8078	36,0571	1	7	1	9
12	9	81,9488	40,1968	1	2	1	11
13	8	81,2450	41,7640	11	14	11	2
14	7	80,4942	43,4360	10	11	10	3
15	6	79,2546	46,1963	1	21	1	12
16	5	78,2205	48,4992	1	10	1	15
17	4	74,7044	56,3289	1	12	1	16
18	3	68,8064	69,4626	1	6	1	17
19	2	68,5708	69,9872	1	4	1	18
20	1	65,5976	76,6080	1	18	1	21

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	15	11288,2	25,2652	49,8764
Cluster2	1	0,0	0,0000	0,0000
Cluster3	1	0,0	0,0000	0,0000
Cluster4	1	0,0	0,0000	0,0000
Cluster5	3	271,3	9,1888	11,7328

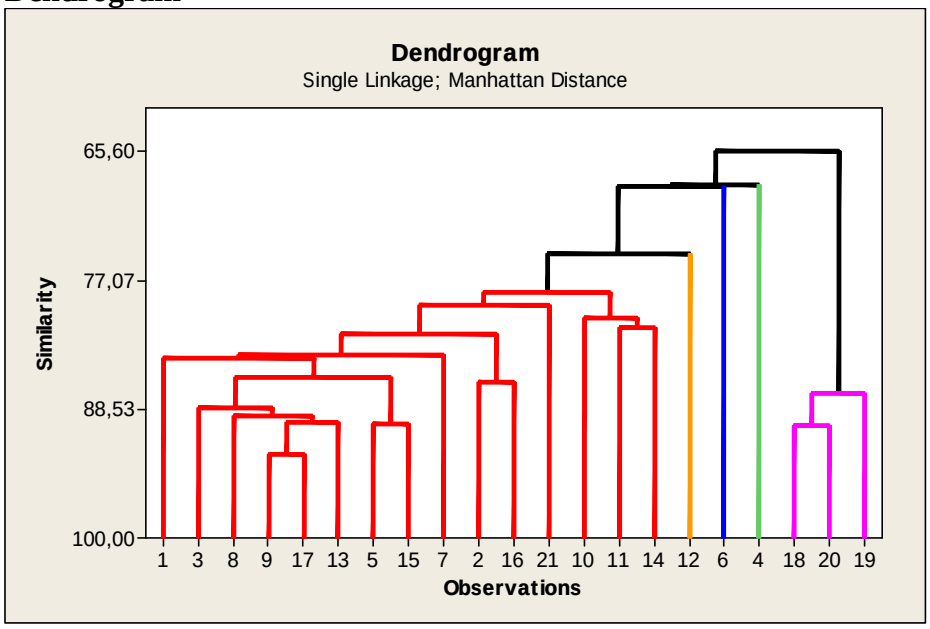
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
HPI N_1	75,4959	75,5033	100,000	80,5517	43,6628	72,3563
RISCN N_1	27,2575	37,4077	59,675	35,5613	8,9857	27,0697
GINNI N_1	57,9930	71,0856	62,082	85,0209	35,7168	56,9159
GO N_1	33,8811	80,6744	64,569	26,1149	70,8631	42,4840
SUSC N_1	31,8277	16,4086	21,054	41,6695	13,5420	28,4368

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,0000	51,9789	52,2101	31,3400	59,5439
Cluster2	51,9789	0,0000	38,1830	61,9511	56,3654
Cluster3	52,2101	38,1830	0,0000	58,2204	80,8361
Cluster4	31,3400	61,9511	58,2204	0,0000	85,3903
Cluster5	59,5439	56,3654	80,8361	85,3903	0,0000

Dendrogram



Cluster Analysis of Observations: GINNI N_1; IPS N_1; RISCN N_1; GO N_1; ...

Manhattan Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	94,3514	13,4482	2	16	2	2
2	19	93,9716	14,3524	13	17	13	2
3	18	91,8154	19,4859	18	19	18	2
4	17	91,5362	20,1506	9	13	9	3
5	16	90,3713	22,9241	3	15	3	2
6	15	89,8157	24,2466	18	20	18	3
7	14	88,5998	27,1415	5	9	5	4
8	13	86,7063	31,6496	2	5	2	6

9	12	85,5743	34,3447	2	8	2	7
10	11	83,5739	39,1073	2	3	2	9
11	10	82,7538	41,0597	2	21	2	10
12	9	81,9665	42,9340	2	14	2	11
13	8	81,3644	44,3676	2	11	2	12
14	7	81,0259	45,1735	2	10	2	13
15	6	79,1306	49,6858	1	2	1	14
16	5	78,9227	50,1808	1	7	1	15
17	4	77,9153	52,5792	1	12	1	16
18	3	74,4074	60,9307	4	6	4	2
19	2	71,1406	68,7083	1	18	1	19
20	1	70,1375	71,0966	1	4	1	21

Final Partition
Number of clusters: 5

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	15	11057,2	24,7914	50,2007
Cluster2	1	0,0	0,0000	0,0000
Cluster3	1	0,0	0,0000	0,0000
Cluster4	1	0,0	0,0000	0,0000
Cluster5	3	169,7	7,3458	8,7452

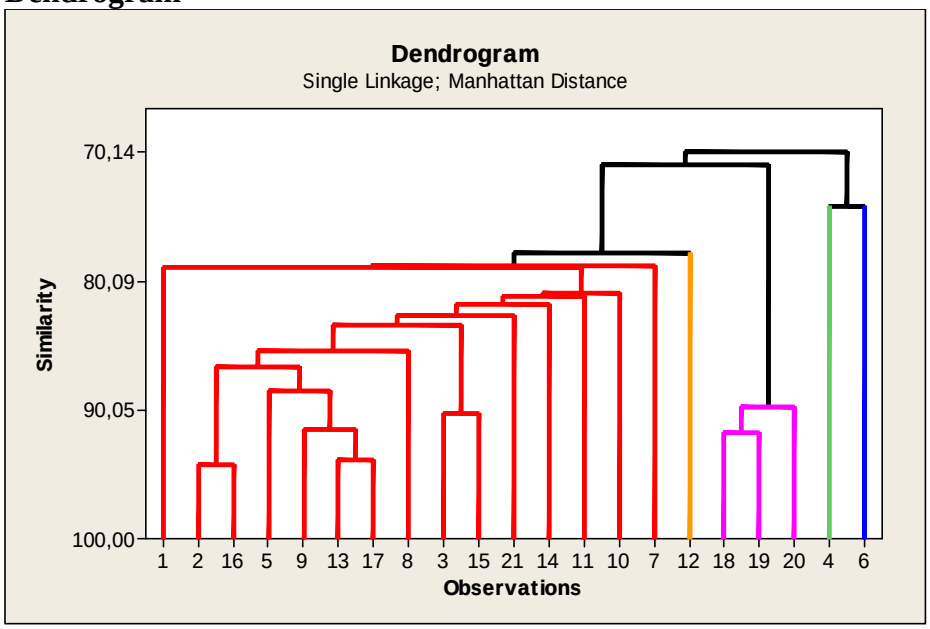
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Grand centroid
GINNI N_1	57,9930	71,0856	62,0825	85,0209	35,7168	56,9159
IPS N_1	59,1972	78,5406	81,1467	51,5457	84,4536	64,4072
RISCN N_1	27,2575	37,4077	59,6750	35,5613	8,9857	27,0697
GO N_1	33,8811	80,6744	64,5690	26,1149	70,8631	42,4840
SDIF N_1	62,6924	37,7525	48,7013	80,3211	23,0580	56,0159

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0,0000	58,8236	51,8353	35,0583	66,3817
Cluster2	58,8236	0,0000	31,0313	75,5986	49,0501
Cluster3	51,8353	31,0313	0,0000	66,8011	63,0293
Cluster4	35,0583	75,5986	66,8011	0,0000	97,4760
Cluster5	66,3817	49,0501	63,0293	97,4760	0,0000

Dendrogram



O resultado da análise de clusters para os componentes principais resultou em maior grau de similaridade em cerca de 82,3151% para o cluster 1 para as seguintes variáveis: Cluster Analysis of Observations: **VULN N_1; HPI N_1; IPS N_1; IDH N_1; EPI N_1**, conforme quadro abaixo.

País	Região		Cluster 1
Argentina	IBE	AIBER	1
Brazil	IBE	AIBER	1
Colombia	IBE	AIBER	1
Costa Rica	IBE	AIBER	1
Cuba	IBE	CARLA	1
Dominican Republic	IBE	AIBER	1
Ecuador	IBE	AIBER	1
El Salvador	IBE	AIBER	1
Guatemala	IBE	AIBER	1
Honduras	IBE	AIBER	1
Mexico	IBE	AIBER	1
Nicaragua	IBE	AIBER	1
Panama	IBE	AIBER	1
Peru	IBE	AIBER	1
Venezuela	IBE	AIBER	1
Bolivia	IBE	AIBER	2
Paraguay	IBE	AIBER	2
Chile	IBE	AIBER	3
Portugal	IBE	AIBER	4
Spain	IBE	AIBER	4
Uruguay	IBE	AIBER	5

18 ANÁLISE DE DISCRIMINANTE

Uma das técnicas mais famosas para resolver problemas de classificação e previsão de elementos é a análise de discriminante. Para análise do cluster 1, da região IBER (AIBER), foi realizada com base nos resultados da técnica de componentes principais PC1, PC2, PC3, PC4. A escolha do cluster foi com base na partição e das variáveis que apresentou maior similaridade para a região IBE(AIBER), resultando na utilização do PC1 para as variáveis: VULN N_1; HPI N_1; IPS N_1; IDH N_1; EPI N_1. Para fins de análise de discriminante, foi utilizado os dados da ANOVA para a variável EPI aplicado a região IBE para o reagrupamento em variáveis qualitativas: Ótimo e regular. É importante ressaltar que utilizou pelo menos de 3 classificações usuais: ótimo, bom e regular, entretanto, dado a amostra e a similaridade entre os países, tal classificação para fins de validação de análise de discriminante não foi possível, por isto, houve o reagrupamento para ótimo e regular.

One-way ANOVA: EPI N_1 versus Cluster 1

```
Source     DF      SS     MS      F      P
Cluster 1   4  3293,8  823,4  16,42  0,000
Error      16   802,3   50,1
Total      20  4096,1
S = 7,081   R-Sq = 80,41%   R-Sq(adj) = 75,52%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  ------+---------+---------+---------+--- Classificação
1      15  48,910   6,836           (*-)                             Regular
2       2  38,179  11,469  (----*----)                              Regular
3       1  74,379       *                  (------*-------)         Otimo
4       2  85,738   4,075                          (----*----)      Ótimo
5       1  50,809       *      (------*-------)                     Regular
                           ------+---------+---------+---------+---
                                40        60        80       100
Pooled StDev = 7,081
```

O resultado da análise discriminante resultou em 100% de acerto nas alocações, uma vez que a amostra é reduzida e há uma evidência da existência de 2 grupos distintos, conforme resultados a seguir.

Discriminant Analysis: Regroup versus IPS N_1; IDH N_1; ...

```
Linear Method for Response: Regroup
Predictors: IPS N_1; IDH N_1; VULN N_1; EPI N_1; HPI N_1
Group   Otimo   Regular
Count       3        18
Summary of classification
                  True Group
Put into Group  Otimo  Regular
Otimo              3        0
Regular            0       18
Total N            3       18
N correct          3       18
Proportion     1,000    1,000
N = 21           N Correct = 21           Proportion Correct = 1,000

Squared Distance Between Groups
           Otimo  Regular
Otimo     0,0000  25,7475
Regular  25,7475   0,0000

Linear Discriminant Function for Groups
            Otimo  Regular
Constant  -572,00  -501,28
IPS N_1      2,10     1,95
IDH N_1      7,04     6,87
VULN N_1     7,90     7,58
EPI N_1      2,73     2,01
HPI N_1      0,01     0,15
```

19 REGRESSÃO LOGÍSTICA

A técnica de regressão logística foi desenvolvida por volta de 1960 em resposta ao desafio de realizar predições ou explicar a ocorrência de determinados fenômenos quando a variável dependente fosse de natureza binária ou dicotômica, assume entre dois resultados e pode ser de natureza qualitativa. Isto exige que o resultado da análise possibilite associações a certas categorias, como por exemplo: positivo ou negativo, aceitar ou rejeitar, morrer ou sobreviver etc.

Para a aplicação da regressão logística dos países que compõem a região IBE, os clusters conforme o Principal Components 1 (PC1) obtidos em sessão anterior, foram realocados para o seguinte:

One-way ANOVA: EPI N_1 versus Cluster 1

```
Source     DF      SS     MS      F      P
Cluster 1   4  3293,8  823,4  16,42  0,000
Error      16   802,3   50,1
Total      20  4096,1
S = 7,081   R-Sq = 80,41%   R-Sq(adj) = 75,52%

Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  ------+---------+---------+---------+--- Classificação
1      15  48,910   6,836           (*-)                             Regular
2       2  38,179  11,469  (----*----)                              Pessimo
3       1  74,379       *                  (------*-------)         Bom
4       2  85,738   4,075                          (----*----)      Ótimo
5       1  50,809       *      (------*-------)                     Regular
                           ------+---------+---------+---------+---
                                40        60        80       100
Pooled StDev = 7,081
```

De acordo com o resultado da aplicação do Logistic Regression, os índices: IDH, VULV, EPI HPI tem tem p-value inferior a 5% e as demais, tendem não contribuir com o modelo de regressão logística, conforme demonstrado a seguir.

Ordinal Logistic Regression: Regroup Logi versus IDH N_1; VULN N_1; ...

```
Link Function: Logit
Response Information
Variable         Value    Count
Regroup Logit_1  Bom          1
                 Otimo        2
                 Pessimo      2
                 Regular     16
                 Total       21

Logistic Regression Table
                                                    Odds     95% CI
Predictor         Coef    SE Coef       Z      P  Ratio  Lower  Upper
Const(1)     -40,4720    31,1478  -1,30  0,194
Const(2)     -38,5266    31,0226  -1,24  0,214
Const(3)     -36,7195    30,8452  -1,19  0,234
IDH N_1       0,279585   0,299512   0,93  0,351   1,32   0,74   2,38
VULN N_1      0,303930   0,236370   1,29  0,199   1,36   0,85   2,15
EPI N_1     0,0981665  0,0814674   1,20  0,228   1,10   0,94   1,29
HPI N_1    -0,0863167  0,0516931  -1,67  0,095   0,92   0,83   1,02
IPS N_1     0,0730627   0,164473   0,44  0,657   1,08   0,78   1,49
Log-Likelihood = -10,182
Test that all slopes are zero: G = 13,238, DF = 5, P-Value = 0,021
```

Goodness-of-Fit Tests demonstra a adequacidade do modelo Logit em relação as variáveis: VULV, HPI, IPS, IDH, EPI. Neste estudo, p-value para o teste de Pearson é 0,465 e o p-value para o teste de desvio é 1,000, o que indica que não há provas suficientes para afirmar que o

modelo não se encaixa nos dados de forma adequada. Se o p-value é menor do que o a-level 5%, o teste rejeita a hipótese nula de que o modelo se ajusta aos dados de forma adequada. Os testes de hipótese é um procedimento que avalia duas afirmações mutuamente exclusivas sobre uma população. Um teste de hipótese usa dados para determinar qual afirmação é melhor suportada pelos dados. Estas afirmações são chamadas de hipótese nula e a hipótese alternativa.

```
Goodness-of-Fit Tests
Method    Chi-Square  DF      P
Pearson      55,2458  55  0,465
Deviance     20,3638  55  1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs       Number  Percent  Summary Measures
Concordant      80     90,9  Somers' D              0,82
Discordant       8      9,1  Goodman-Kruskal Gamma  0,82
Ties             0      0,0  Kendall's Tau-a        0,34
Total           88    100,0
```

Reprocedendo os testes de regresssão logistica para com as variáveis com p-value inferior a 0,05 temos 90,8% da adequada categorização dos países da região e que as variáveis que contribuem para esta alocação são IDH, VULV, EPI e HPI.

Ordinal Logistic Regression: Regroup Logi versus IDH N_1; VULN N_1; ...

```
Link Function: Logit
Response Information

Variable         Value    Count
Regroup Logit_1  Bom          1
                 Otimo        2
                 Pessimo      2
                 Regular     16
                 Total       21

Logistic Regression Table
                                                    Odds     95% CI
Predictor         Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)     -33,2260    26,4685  -1,26  0,209
Const(2)     -31,2862    26,2980  -1,19  0,234
Const(3)     -29,4975    26,1153  -1,13  0,259
IDH N_1       0,275409   0,300574   0,92  0,360   1,32   0,73   2,37
VULN N_1      0,255571   0,205994   1,24  0,215   1,29   0,86   1,93
EPI N_1       0,106084  0,0782455   1,36  0,175   1,11   0,95   1,30
HPI N_1     -0,0950595  0,0525021  -1,81  0,070   0,91   0,82   1,01
Log-Likelihood = -10,289
Test that all slopes are zero: G = 13,023, DF = 4, P-Value = 0,011

Goodness-of-Fit Tests
Method    Chi-Square  DF      P
Pearson      61,5864  56  0,283
Deviance     20,5788  56  1,000

Measures of Association: (Between the Response Variable and Predicted
Probabilities)
Pairs       Number  Percent  Summary Measures
Concordant      80     90,9  Somers' D              0,82
Discordant       8      9,1  Goodman-Kruskal Gamma  0,82
Ties             0      0,0  Kendall's Tau-a        0,34
Total           88    100,0
```

20 ÁRVORE CLASSIFICATÓRIA

É um modelo de classificação baseado em árvore. Classifica os dados em grupos ou prevê valores de uma variável dependente (destino) com base em valores de variáveis (Predictor) independentes. Este procedimento fornece ferramentas de validação para análise de classificação exploratória e confirmatória. Para os países classificados na região IBE (AIBER) estão classificados em de acordo com o **Cluster 1**, com as variáveis independentes: **IDH, HPI, VULV, IPS, EPI.**

MODELO CHAID

Resumo do modelo

Especificações	Método crescente	CHAID
	Variável dependente	Cluster 1
	Variáveis independentes	EPI N_1, HPI N_1, VULN N_1, IDH N_1, IPS N_1
	Validação	Nenhum
	Profundidade máxima de árvore	3
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	HPI N_1
	Número de nós	3
	Número de nós de terminal	2
	Espessura	1

Risco

Estimativa	Erro Padrão
,238	,093

Método Crescente: CHAID

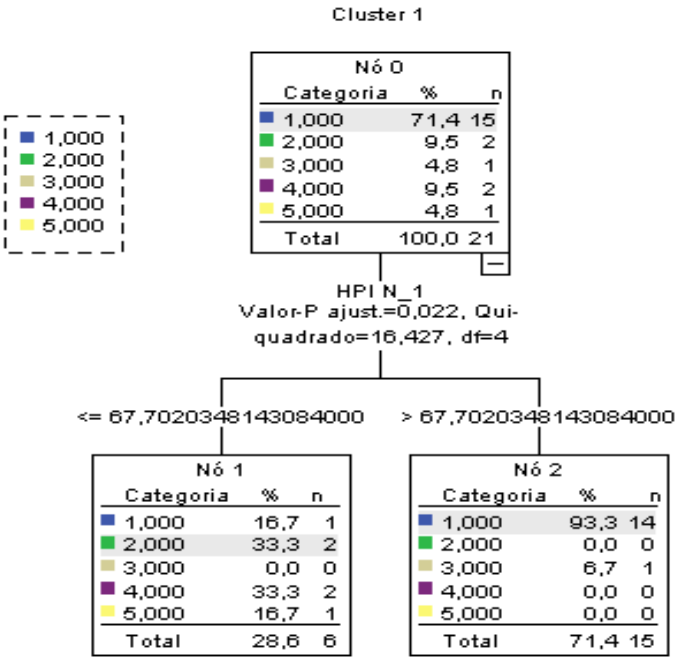
Variável Dependente: Cluster 1

Classificação

	Previsto					
Observado	1,0	2,0	3,0	4,0	5,0	Porcentagem Correta
1,0	14	1	0	0	0	93,3%
2,0	0	2	0	0	0	100,0%
3,0	1	0	0	0	0	0,0%
4,0	0	2	0	0	0	0,0%
5,0	0	1	0	0	0	0,0%
Porcentagem global	71,4%	28,6%	0,0%	0,0%	0,0%	76,2%

Método Crescente: CHAID

Variável Dependente: Cluster 1



MODELO CRT-

Resumo do modelo

Especificações	Método crescente	CRT
	Variável dependente	Cluster 1
	Variáveis independentes	EPI N_1, HPI N_1, VULN N_1, IDH N_1, IPS N_1
	Validação	Nenhum
	Profundidade máxima de árvore	5
	Casos mínimos em nó pai	10
	Casos mínimos em nó filho	5
Resultados	Variáveis independentes incluídas	HPI N_1, VULN N_1, EPI N_1, IPS N_1, IDH N_1
	Número de nós	3
	Número de nós de terminal	2
	Espessura	1

Risco

Estimativa	Erro Padrão
,190	,086

Método Crescente: CRT

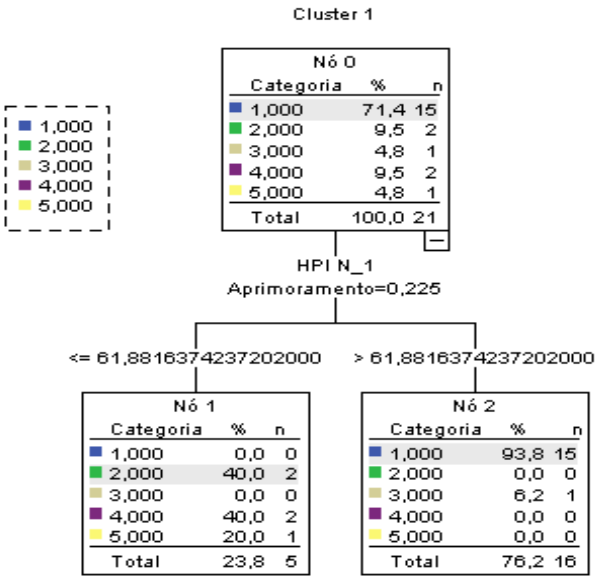
Variável Dependente: Cluster 1

Classificação

Observado	Previsto					
	1,0	2,0	3,0	4,0	5,0	Porcentagem Correta
1,0	15	0	0	0	0	100,0%
2,0	0	2	0	0	0	100,0%
3,0	1	0	0	0	0	0,0%
4,0	0	2	0	0	0	0,0%
5,0	0	1	0	0	0	0,0%
Porcentagem global	76,2%	23,8%	0,0%	0,0%	0,0%	81,0%

Método Crescente: CRT

Variável Dependente: Cluster 1



21 ANÁLISE DE CORRESPONDÊNCIA

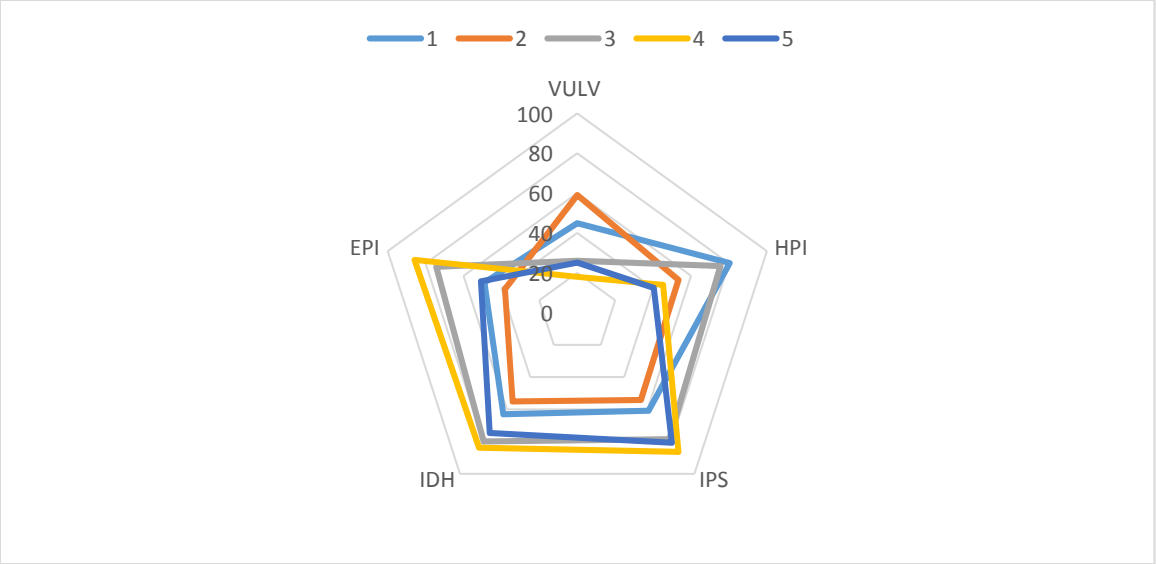
A análise de correspondência é uma técnica de análise exploratória de dados adequada para analisar tabelas de duas entradas ou tabelas de múltiplas entradas, levando em conta algumas medidas de correspondência entre linhas e colunas. Analise de correspondenca converte uma matriz de dados não negativos em um tipo particular de representação gráfica em que as linhas e colunas da matriz são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico.

A execução da ANOVA, na seção 15, foi para obtenção das médias para as variáveis: VULV, HPI, IPS, IDH e EPI, cujo resultado será utilizado para análise de correspondência. A tabela abaixo é o resumo das médias apresentadas para cada variável: **VULV, HPI, IPS, IDH e EPI**, cuja entrada dupla de dados será a base para a tabela de contingência a ser executada na análise de correspondência. Lembrando que 1, 2, 3, 4 e 5 são os países classificados anteriormente pela análise de cluster.

Tabela de Contingência –

XXXXX	**Mean**					XXXXX
Rows	**VULV**	**HPI**	**IPS**	**IDH**	**EPI**	**Columns**
1	44,88	80,419	60,812	63,04	48,910	V
2	59,05	53,350	54,233	55,11	38,179	H
3	26	75,503	78,541	79,9	74,379	PS
4	17,99	45,311	86,323	83,77	85,738	DH
5	25,12	40,367	80,715	74,63	50,809	EP

Gráficos Radar para regiões do Cluster 1,2, 3, 4 e 5



Simple Correspondence Analysis: VULV; HPI; IPS; IDH; EPI

Contingency Table

	V	H	PS	DH	EP	Total
1	44,880	80,419	60,812	63,040	48,910	298,061
2	59,050	53,350	54,233	55,110	38,179	259,922
3	26,000	75,503	78,541	79,900	74,379	334,323
4	17,990	45,311	86,323	83,770	85,738	319,132
5	25,120	40,367	80,715	74,630	50,809	271,641
Total	173,040	294,950	360,624	356,450	298,015	1483,079

Chi-Square Distances

	V	H	PS	DH	EP	Total
1	2,935	7,540	1,877	1,032	2,014	15,399
2	27,205	0,053	1,273	0,867	3,780	33,178
3	4,338	1,222	0,093	0,003	0,771	6,427
4	9,947	5,194	0,981	0,651	7,283	24,056
5	1,364	3,452	3,255	1,337	0,261	9,669
Total	45,788	17,462	7,479	3,890	14,109	88,728

Relative Inertias

	V	H	PS	DH	EP	Total
1	0,033	0,085	0,021	0,012	0,023	0,174
2	0,307	0,001	0,014	0,010	0,043	0,374
3	0,049	0,014	0,001	0,000	0,009	0,072
4	0,112	0,059	0,011	0,007	0,082	0,271
5	0,015	0,039	0,037	0,015	0,003	0,109
Total	0,516	0,197	0,084	0,044	0,159	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0473	0,7910	0,7910	******************************
2	0,0100	0,1673	0,9583	******
3	0,0025	0,0416	1,0000	*
4	0,0000	0,0000	1,0000	
Total	0,0598			

Row Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	1	0,990	0,201	0,174	0,198	0,756	0,166	-0,110	0,234	0,242
2	2	0,987	0,175	0,374	0,337	0,888	0,420	0,112	0,099	0,221
3	3	1,000	0,225	0,072	-0,083	0,358	0,033	-0,111	0,642	0,278
4	4	0,951	0,215	0,271	-0,266	0,939	0,322	0,030	0,012	0,020
5	5	0,802	0,183	0,109	-0,124	0,435	0,060	0,114	0,368	0,239

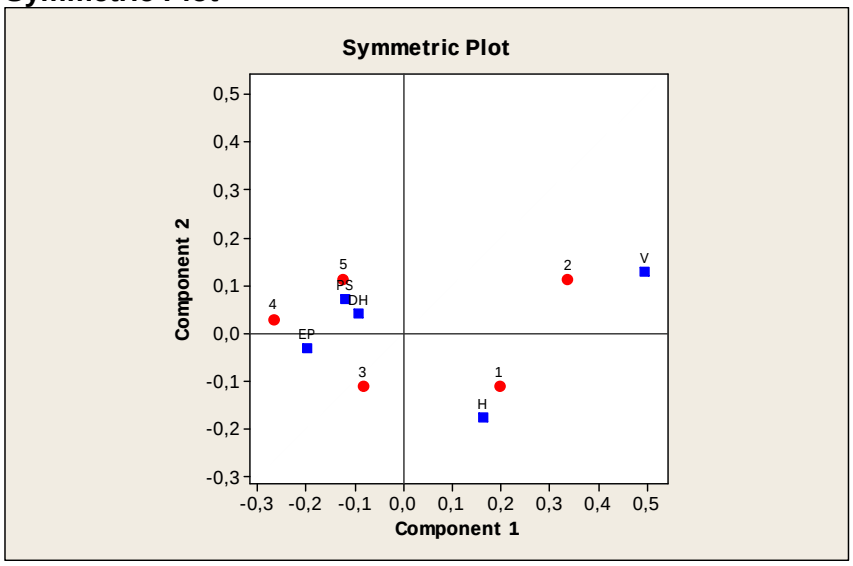
Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	V	0,993	0,117	0,516	0,496	0,931	0,607	0,129	0,062	0,193
2	H	0,988	0,199	0,197	0,164	0,454	0,113	-0,178	0,534	0,629
3	PS	0,905	0,243	0,084	-0,118	0,669	0,071	0,070	0,235	0,119
4	DH	0,938	0,240	0,044	-0,093	0,786	0,044	0,041	0,152	0,040
5	EP	0,842	0,201	0,159	-0,197	0,820	0,165	-0,032	0,021	0,020

A inércia total é o percentual da variância que é explicada pela aplicação da Análise de Correspondência e corresponde à soma ponderada das distâncias dos pontos do conjunto a seu centróide. Pode-se entender a Análise de Correspondência como um método de decomposição da inércia total em dimensões que melhor expliquem a variabilidade dos dados. Neste caso, observamos que a componente 1 no axis 1 está com 79,10%, já com a componente 2, sobe para 95,83%.

O gráfico Symmetric Plot demonstra que a variável VULV e países que compõem o cluster 2 tem maior correspondência, enquanto que o cluster 4 e 5 estão para IDH e IPS e finalmente, o cluster 1 tem maior correspondência com a variável H- Happy Index.

Symmetric Plot



A anállise de correspondência permite estudar as relações e semelhanças existentes entre (a) as categorias de linhas e entre as categorias de colunas de uma tabela de contingência, (b) o conjunto de categorias de linhas e o conjunto categorias de colunas, cujos os resultados demonstram que a região é categorizada em 3 grandes grupos com caracteristicas distintas, conforme demonstrados no quadro abaixo, como segue:

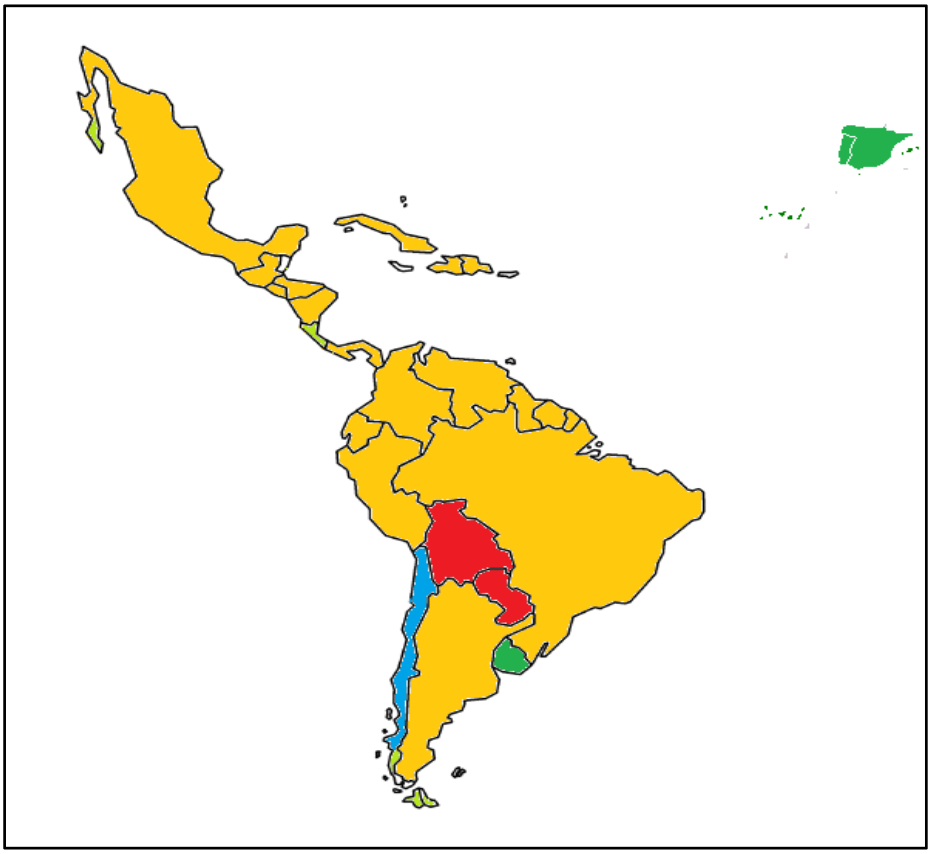
País	Região		Cluster 1	Variáveis	Mapa Alocação
Argentina	IBE	AIBER	1	(H) Happy Index	Amarelo
Brazil	IBE	AIBER	1	(H) Happy Index	Amarelo
Colombia	IBE	AIBER	1	(H) Happy Index	Amarelo
Costa Rica	IBE	AIBER	1	(H) Happy Index	Amarelo
Cuba	IBE	CARLA	1	(H) Happy Index	Amarelo
Dominican Republic	IBE	AIBER	1	(H) Happy Index	Amarelo
Ecuador	IBE	AIBER	1	(H) Happy Index	Amarelo
El Salvador	IBE	AIBER	1	(H) Happy Index	Amarelo
Guatemala	IBE	AIBER	1	(H) Happy Index	Amarelo
Honduras	IBE	AIBER	1	(H) Happy Index	Amarelo
Mexico	IBE	AIBER	1	(H) Happy Index	Amarelo
Nicaragua	IBE	AIBER	1	(H) Happy Index	Amarelo
Panama	IBE	AIBER	1	(H) Happy Index	Amarelo
Peru	IBE	AIBER	1	(H) Happy Index	Amarelo
Venezuela	IBE	AIBER	1	(H) Happy Index	Amarelo
Bolivia	IBE	AIBER	2	(V) Vunerabilidade	Vermelho
Paraguay	IBE	AIBER	2	(V) Vunerabilidade	Vermelho
Chile	IBE	AIBER	3	(EP) EPI	Azul
Portugal	IBE	AIBER	4	(DH), (PS) - IDH, IPS	Verde
Spain	IBE	AIBER	4	(DH), (PS) - IDH, IPS	Verde
Uruguay	IBE	AIBER	5	(DH), (PS) - IDH, IPS	Verde

Fonte: preparado pelo autor.

22. MAPA DA REGIÃO IBE(AIBER)

Os países da região IBE(AIBER) para o eixo temático resiliência , são categorizados em 4 clusters distintos, onde a maioria, 15 deles, tem maior correspondência para o índice: *Happy Index*, outros, tais como: Bolívia e Paraguai tem sua correspondência na vulnerabilidade e o Uruguai, praticamente comparado aos países da Europa: Portugal, Espanha. Finalmente, o Chile teve correspondência com o índice *Enviromental Progress Index*.
Em contrapartida, quando estes países da região IBE (AIBER) são analisados na amostra de 132 países, os índices que categorizam a maior correspondência entre os países são as variáveis VULV - vulnerabilidade e GO – governança que praticamente na utilização da técnica de stepwise não tem F-test significativo.

Mapa categorizado por correspondência -



Fonte: adaptado pelo autor.

Em seção 15 deste trabalho, a análise estatística das variáveis: IPS, IDH, HPI, VULV e EPI resultou em maior R-Sq(adj) igual a 74,26% o índice Happy index (HPI).
O HPI ou Índice do Planeta Feliz mede a capacidade que cada país tem para proporcionar um bem-estar sustentável aos seus cidadãos. Foi introduzido pela *New Economics Foundation*, uma organização não governamental ecológica britânica em Julho de 2006. O índice não é uma forma de quantificar quais são os países mais felizes, mas uma forma de medir a eficiência com que uma nação converte os seus recursos naturais em longevidade e felicidade para seus cidadãos. É a associação do nível subjetivo de satisfação, esperança média de via e eco políticas aplicadas per capita.

O índice é uma medida de eficiência, que classifica os países sob o aspecto de quanto há longevidade e felicidade produzida por cada unidade de insumo ambiental.
O índice HPI tem escala de 0 a 100, cuja obtenção é do índice é dado pela seguinte divisão do seguinte denominador:

$$\text{Happy Planet Index} \approx \frac{\text{Experienced well-being} \times \text{Life expectancy}}{\text{Ecological footprint}}$$

O HPI utiliza a medida de *Ecological footprint* com o objetivo de medir o consumo de recursos. É uma medida per capita da quantidade de terra necessária para sustentar os padrões de consumo de um país, medido em termos de hectares globais (g ha) que representam um hectare de terra com a biocapacidade produtiva média. Este cálculo é realizado por uma entidade não governamental, a WWF.
Os resultados confirmam que ainda não estamos vivendo em um planeta feliz. Nenhum país é capaz de combinar o sucesso ao longo dos três objetivos da expectativa de vida elevado, elevado bem-estar e viver dentro dos limites ambientais experientes.
De acordo com as últimas divulgações realizadas pela *New Economics Foundation*, os países de alta renda têm baixa pontuação face a extensa utilização dos recursos naturais ou do seu *Ecological Foot print*, em contrapartida, os países de baixa renda, i.e. Continente Africano, tem a tendência de baixa expectativa de vida e bem-estar, mas baixo índice de *Ecological Footprint*. E, para os países em desenvolvimento, i.e. Latino Americanos, tem pontuação alta para o índice, face os resultados razoáveis de expectativa de vida e bem-estar e baixo índice de *Ecological Foot print*.

O resultado da ANOVA para o índice HPI demonstra a aderência dos resultados obtidos com a divulgação das pesquisas pelo instituto, como seguem:

One-way ANOVA: HPI N_1 versus Cluster 1

```
Source     DF      SS      MS      F      P
Cluster 1   4  4193,9  1048,5  15,42  0,000
Error      16  1087,8    68,0
Total      20  5281,6
S = 8,245   R-Sq = 79,40%   R-Sq(adj) = 74,26%
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  ---------+---------+---------+---------+ CLASSFICAÇÃO
1      15  80,419  8,403                              (-*-)              1
2       2  53,350  3,835          (------*-----)                      3
3       1  75,503      *                   (--------*-------)          2
4       2  45,311  9,188       (------*-----)                         4
5       1  40,367      *  (--------*--------)                         5
                          ---------+---------+---------+---------+
                                  40        60        80       100
Pooled StDev = 8,245
```

A categorização dos clusters em uma escala de 1-5 é corresponde do melhor para o pior resultado observado através das médias de cada país em seus respectivos clusters.

Abaixo, faz a representação dos resultados em Mapa da Região em cores e uma tabela classificatória dos países da região IBER(AIBER).

MAPA DA REGIÃO IBER(AIBER) POR HAPPY PLANET INDEX –

País	Cluster 1	ANOVA Class	MAPA
Argentina	1	1	
Brazil	1	1	
Colombia	1	1	
Costa Rica	1	1	
Cuba	1	1	
Dominican Rep	1	1	
Ecuador	1	1	
El Salvador	1	1	
Guatemala	1	1	
Honduras	1	1	
Mexico	1	1	
Nicaragua	1	1	
Panama	1	1	
Peru	1	1	
Venezuela	1	1	
Bolivia	2	3	
Paraguay	2	3	
Chile	3	2	
Portugal	4	4	
Spain	4	4	
Uruguay	5	5	

O quadro abaixo demonstra a classificação dos países face ao Happy Planet Index que estão alinhados com os resultados, em termos gerais, sobre a elevada pontuação dos países da América Latina em detrimento dos países com elevado IDH e IPS, como Espanha e Portugal.

Quadro – Classificação dos países por Happy Planet Index

	Pais	HPI	IPS	IDH	EPI	VULN	GO	RISC	EXP	SUSC	SDIF	CADP	GINI
1	Costa Rica	100	81,147	70,181	57,915	33,564	64,569	59,675	80,113	21,054	48,701	26,866	62,082
2	Colombia	89,661	62,257	61,614	46,707	49,353	37,787	20,901	22,027	30,882	67,641	45,99	74,922
3	El Salvador	87,576	57,692	53,542	36,626	55,15	41,425	58,9	59,903	36,402	68,074	58,05	44,311
4	Panama	84,95	71,855	70,511	55,474	41,495	47,875	23,043	27,297	29,367	54,582	36,498	70,668
5	Nicaragua	83,176	53,433	45,634	46,057	58,997	30,53	50,591	49,061	45,978	79,527	46,667	54,567
6	Venezuela	82,713	56,039	70,346	56,86	40,492	8,006	17,356	20,634	22,164	66,071	29,055	52,062
7	Guatemala	82,688	51,707	47,941	42,793	63,282	28,169	72,046	67,373	46,163	77,976	62,149	71,842
8	Cuba	81,06	51,168	78,748	52,917	26,472	28,758	19,387	29,316	15,399	35,335	25,075	37,067
9	Honduras	80,552	51,546	46,129	43,963	57,83	26,115	35,561	34,484	41,669	80,321	47,065	85,021
10	Argentina	75,918	68,278	77,595	44,945	29,681	35,302	9,269	13,366	17,789	39,881	27,741	48,93
11	Chile	75,503	78,541	79,901	74,379	25,998	80,674	37,408	56,572	16,409	37,753	19,622	71,086
12	Brazil	73,208	67,164	67,051	49,884	41,604	45,427	11,558	13,325	25,345	52,291	44,159	72,677
13	Mexico	73,116	60,766	69,028	52,86	41,933	41,8	18,833	22,027	22,753	62,32	37,194	60,673
14	Ecuador	72,121	63,893	61,614	57,929	36,39	29,149	23,855	26,691	32,582	65,206	34,09	56,759
15	Peru	71,849	60,55	65,898	38,446	46,91	39,197	21,196	23,158	32,144	64,34	40,199	53,523
16	Rep. Domin.	67,702	54,691	59,802	50,274	49,991	37,449	38,146	40,804	32,447	55,465	55,582	54,436
17	Paraguay	56,061	54,008	55,848	30,069	56,354	27,325	9,49	8,278	36,772	74,874	54,01	60,517
18	Spain	51,808	86,574	87,644	88,619	16,408	70,133	7,496	14,739	9,441	25,956	9,512	28,523
19	Bolivia	50,638	54,457	54,366	46,288	61,75	30,018	14,291	12,215	51,229	76,804	52,478	56,942
20	Uruguay	40,367	80,715	74,629	50,809	25,123	68,43	10,451	16,495	18,664	23,791	29,493	43,058
21	Portugal	38,814	86,071	79,901	82,857	19,581	74,026	9,01	16,152	12,521	19,426	23,423	35,569

Com o objetivo de corroborar com o entendimento *New Economics Foundation*, reproduziu a análise com a variável Enviromental Performance Index (EPI). O Índice de Desempenho Ambiental (EPI) classifica os países desempenho em questões ambientais de alta prioridade em duas áreas: proteção da saúde e proteção dos direitos humanos ecossistemas. O resultado

da ANOVA para o índice EPI demonstra a aderência dos resultados obtidos com a divulgação das pesquisas pelo *New Economics Foundation*, como seguem:

One-way ANOVA: EPI N_1 versus Cluster 1

```
Source      DF      SS     MS      F      P
Cluster 1    4  3293,8  823,4  16,42  0,000
Error       16   802,3   50,1
Total       20  4096,1
S = 7,081   R-Sq = 80,41%   R-Sq(adj) = 75,52%
Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev  ------+---------+---------+---------+--- CLASSFICAÇÃO
1      15  48,910   6,836           (*-)                              5
2       2  38,179  11,469  (----*----)                               4
3       1  74,379       *                  (------*-------)          2
4       2  85,738   4,075                           (----*----)      1
5       1  50,809       *       (------*-------)                     3
                            ------+---------+---------+---------+---
                                  40        60        80       100
Pooled StDev = 7,081
```

MAPA DA REGIÃO IBER(AIBER) POR ENVIROMENTAL PERFORMANCE INDEX –

País	Cluster 1	ANOVA Class	MAPA
Argentina	1	5	
Brazil	1	5	
Colombia	1	5	
Costa Rica	1	5	
Cuba	1	5	
Dominican Rep	1	5	
Ecuador	1	5	
El Salvador	1	5	
Guatemala	1	5	
Honduras	1	5	
Mexico	1	5	
Nicaragua	1	5	
Panama	1	5	
Peru	1	5	
Venezuela	1	5	
Bolivia	2	4	
Paraguay	2	4	
Chile	3	2	
Portugal	4	1	
Spain	4	1	
Uruguay	5	3	

Quadro – Classificação dos países por Enviromental Performance Index

	Pais	EPI	HPI	IPS	IDH	GO	RISC	EXP	VULN	SUSC	SDIF	CADP	GINI
1	Spain	88,619	51,808	86,574	87,644	70,133	7,496	14,739	16,408	9,441	25,956	9,512	28,523
2	Portugal	82,857	38,814	86,071	79,901	74,026	9,01	16,152	19,581	12,521	19,426	23,423	35,569
3	Chile	74,379	75,503	78,541	79,901	80,674	37,408	56,572	25,998	16,409	37,753	19,622	71,086
4	Ecuador	57,929	72,121	63,893	61,614	29,149	23,855	26,691	36,39	32,582	65,206	34,09	56,759
5	Costa Rica	57,915	100	81,147	70,181	64,569	59,675	80,113	33,564	21,054	48,701	26,866	62,082
6	Venezuela	56,86	82,713	56,039	70,346	8,006	17,356	20,634	40,492	22,164	66,071	29,055	52,062
7	Panama	55,474	84,95	71,855	70,511	47,875	23,043	27,297	41,495	29,367	54,582	36,498	70,668
8	Cuba	52,917	81,06	51,168	78,748	28,758	19,387	29,316	26,472	15,399	35,335	25,075	37,067
9	Mexico	52,86	73,116	60,766	69,028	41,8	18,833	22,027	41,933	22,753	62,32	37,194	60,673
10	Uruguay	50,809	40,367	80,715	74,629	68,43	10,451	16,495	25,123	18,664	23,791	29,493	43,058
11	Rep. Domin.	50,274	67,702	54,691	59,802	37,449	38,146	40,804	49,991	32,447	55,465	55,582	54,436
12	Brazil	49,884	73,208	67,164	67,051	45,427	11,558	13,325	41,604	25,345	52,291	44,159	72,677
13	Colombia	46,707	89,661	62,257	61,614	37,787	20,901	22,027	49,353	30,882	67,641	45,99	74,922
14	Bolivia	46,288	50,638	54,457	54,366	30,018	14,291	12,215	61,75	51,229	76,804	52,478	56,942
15	Nicaragua	46,057	83,176	53,433	45,634	30,53	50,591	49,061	58,997	45,978	79,527	46,667	54,567
16	Argentina	44,945	75,918	68,278	77,595	35,302	9,269	13,366	29,681	17,789	39,881	27,741	48,93
17	Honduras	43,963	80,552	51,546	46,129	26,115	35,561	34,484	57,83	41,669	80,321	47,065	85,021
18	Guatemala	42,793	82,688	51,707	47,941	28,169	72,046	67,373	63,282	46,163	77,976	62,149	71,842
19	Peru	38,446	71,849	60,55	65,898	39,197	21,196	23,158	46,91	32,144	64,34	40,199	53,523
20	El Salvador	36,626	87,576	57,692	53,542	41,425	58,9	59,903	55,15	36,402	68,074	58,05	44,311
21	Paraguay	30,069	56,061	54,008	55,848	27,325	9,49	8,278	56,354	36,772	74,874	54,01	60,517

23. CONCLUSÃO

A palavra resiliência, de acordo com Oliveira et al. (2008), é originária do latim resilo e se refere à capacidade de retornar ao estado anterior. Esse conceito foi aplicado primeiramente às ciências exatas, com ênfase na Engenharia e na Física, para se referir à capacidade elástica de um corpo físico voltar ao seu estado normal após ser submetido a algum tipo de pressão (Barlach, Limongi-França & Malvezzi, 2008). Ao ser adaptado para as ciências humanas, em especial para a Psicologia, tal conceito foi descrito em oposição ao de vulnerabilidade, sendo definido, a princípio, como um conjunto de traços de personalidade, capacidades ou habilidades que tornam as pessoas resistentes à doenças psíquicas quando passam por experiências traumáticas (Souza & Cerveny, 2006). Souza e Cerveny (2006), ao se referirem à importância do contexto e das redes de relacionamento para o desenvolvimento saudável em condições adversas, destacam a relação complexa entre indivíduo e ambiente na produção de fatores de risco, no provimento de fatores protetores e no próprio desenvolvimento da resiliência.

Para o World Risk Report (2014) o crescimento urbano expõe as pessoas aos desastres naturais, como por exemplo: enchentes, utilização desordenada dos recursos naturais. Também, promove um fluxo substancial de pessoas, levando ao desenvolvimento de assentamentos informais e ou, favelas - muitas vezes em áreas que são particularmente expostas aos perigos naturais, como em margens de rios ou encostas íngremes. Centenas de milhões de pessoas em todo o mundo vivem atualmente em assentamentos informais ou favelas como os moradores da cidade sem quaisquer direitos cívicos, e têm de competir uns contra os outros por empregos mal pagos e suprimentos de alimentos escassos.

Com isto, notamos no presente trabalho que os países desenvolvidos, constantes nos continentes América do Norte, Europeu, Oceania e o Asiático: Japão, representados em sua maior parte pela AVECO, têm políticas públicas privadas que os tornam mais resilientes. A variável que melhor representa o grupo é o índice de governança (GO), elevado IDH e IPS. Por outro lado, os países da AIBER, OTHERS ou aqueles situados América Latina e Central, África, Ásia tem índices muito aquém do desejado, sendo que a variável que tem maior correspondência entre os mesmos é a VULV- vulnerabilidade e a altos índices de percepção sobre a corrupção, ausência de democracia em cada região. O desenvolvimento do conceito de resiliência dos países é uma tarefa árdua para os governos, mas traz equilíbrio e equidade a sociedade em longo prazo.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

Tema: BIODIVERSIDADE - CAPITAL NATURAL E SOCIAL

TRABALHO FINAL DE SEMESTRE

Disciplina: Métodos Quantitativos e Qualitativos em Administração

Professor: Dr. Arnoldo José de Hoyos Guevara

Marcos Antonio Ferreira

05/Jul/2016

1. INTRODUÇÃO

O presente trabalho tem por objetivo analisar um conjunto de dados disponibilizados pelo Banco Mundial relativos ao tema **Biodiversidade - Capital Natural e Social**, que possui um conjunto de 33 (trinta e três) variáveis previamente selecionadas e que refletem o grau de engajamento sobre esta temática, considerando os 132 países objetos deste estudo.

Iniciamos o trabalho apresentando os objetivos do programa Países Sustentáveis – Gestão Pública Sustentável (GPS), uma publicação do Núcleo de Estudos do Futuro da PUC-SP. Em seguida algumas definições e contextualização sobre o que se entende por Biodiversidade, Capital Natural e Social.

Depois partimos para o entendimento dos dados apresentando resultados de análise exploratória dos dados sobre o tema, empregando os *softwares* estatísticos MINITAB e SPSS,

Listar principais análises estatísticas utilizadas nas 12 sessões...

Por fim, apresentamos os resultados e considerações....

O objetivo é explorar os dados sobre **Biodiversidade: Capital Natural e Social** de tal modo que possamos com os recursos citados anteriormente apresentar conclusões de correlações entre as variáveis que compõem o tema.

1.1. ENTENDENDO OS DADOS

Antes da apresentação dos indivíduos desta análise que envolve 132 países e os indicadores de desenvolvimento humano selecionados a partir de referência internacional, citamos a origem do trabalho, que é o **Projeto ORIBER – Obervatório RIBER**

1.1.1 O tema de pesquisa: Biodiversidade - Capital Natural e Social

A análise dos conceitos que compõem o tema foi extraída do material de aula e do Guia GPS. Nesse estudo são apresentadas as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso.

Vale destacar o desafio citado no GPS e sua proposta de solução:

Desafio: *"A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa ?"*

Solução: *"O GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."*

Para chegar a uma conclusão do estágio em que se encontra o país e como é possível adotar uma solução, o GPS propõe um levantamento dos dados a partir da análise SWOT, conforme se apresenta na figura 1.

Figura 1: Análise SWOT de cada país

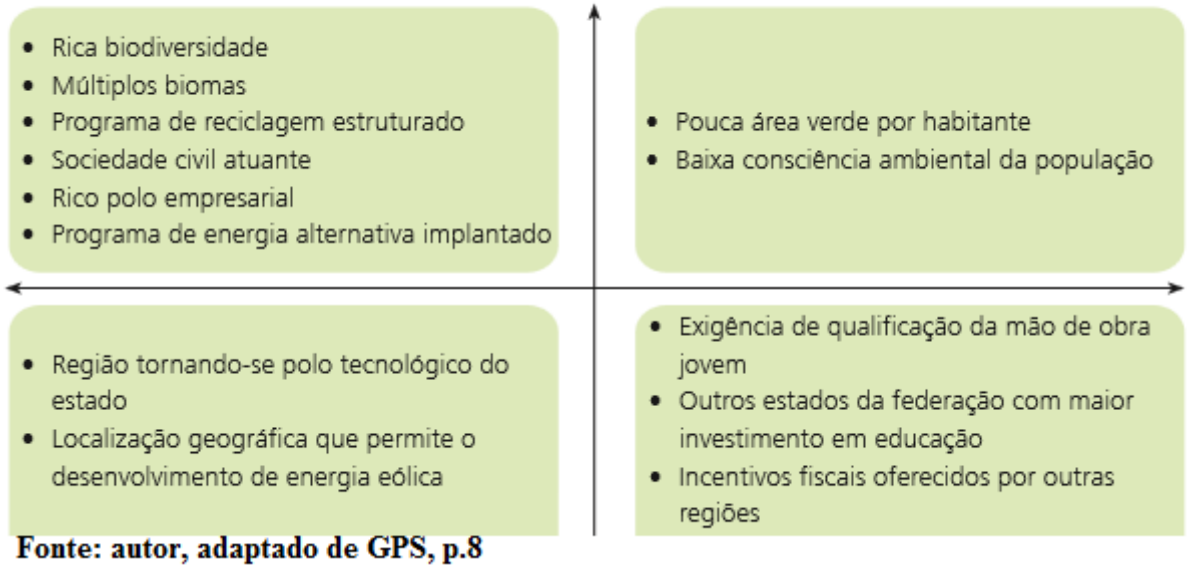
fonte: autor, adaptado de GPS, p.8

Na figura 1 é possível obter informações sobre o país analisado considerando suas Forças – quais são os seus pontos fortes em cada um dos eixos temáticos e suas Fragilidades, Ameaças e Oportunidades que cada eixo temático apresenta para o país em questão.

Como exemplo inicial deste trabalho referente ao tema: **Biodiversidade - Capital Natural e Social,** podemos destacar o “Alimento”. Supondo que o país analisado seja um grande produtor de alimento (ponto forte), tenha uma estiagem em seu território (fragilidade) e perceba uma ameaça (outro país produtor) mas identifica uma oportunidade, a de planejar a produção de certo alimento usando sistema de irrigação por gotejamento a partir do armazenamento e do uso racional desse recurso.

Para ajudar o leitor a entender a contribuição da matriz SWOT, o GPS apresenta um exemplo com o tema “Bens naturais comuns”, conforme mostra a figura 2.

Figura 2: Análise SWOT do eixo temático Bens Naturais Comuns



Fonte: autor, adaptado de GPS, p.8

Se tomarmos como ponto de partida o item do quadrante do ponto forte “Rica Biodiversidade” e avançarmos no sentido horário da matriz encontraremos a relação com “Pouca área verde por habitante”, que indica a fragilidade. Aparece como ameaça a “Exigência de qualificação da mão de obra jovem” e em seguida como oportunidade o fato de a “Região se tornar um polo tecnológico do estado”. Note que as variáveis vão sendo encontradas e analisadas como proposta de solução para o problema do tema daquele país.

Projeto Milênio – Observatório RIBER - ORIBER

A Rede Ibero-americana do Projeto Milênio – RIBER, tem como objetivo promover o estudo da prospectiva na Iberoamérica através de uma comunidade com competências, conceitos, métodos e práticas prospectivas, que permita analisar e intercambiar informações sobre os diferentes países da região e contribuir na aprendizagem coletiva dos atores regionais para o desenvolvimento de competências em prospectiva, fornecendo referências das melhores práticas existentes na região e no mundo como um todo.
Fonte: Guia GPS – Gestão Pública Sustentável

1.1.2 Os indivíduos

Os indivíduos desta análise são os 132 países analisados por uma série de indicadores de referência internacional. Os dados analisados de cada país são as variáveis que descrevemos a seguir.

1.1.2.1 As variáveis do tema: Biodiversidade - Capital Natural e Social

Este tema da pesquisa apresenta 33 (trinta e três) variáveis, sendo três categóricas – País, Country code e REGIÃO e e trinta quantitativas, descritas abaixo no Quadro 1. Estas ainda podem ser divididas em: índices sintéticos e variáveis-componentes ou indicadores, a saber:

Índices sintéticos: São 16 (dezesseis): Índice de Progresso Social - Social Progress Index, Índice de Desenvolvimento Humano - IDH – 2013, Índice de Governança, Índice de Proteção Ambiental - EPI Score, EV - Water Resources – tratamento de águas residuais, EV Agriculture, EV Forests, EV Fisheries - Ecosystem Vitality (índice de vitalidade do ecossistema), Basic Human Needs, Foundations of Wellbeing, Opportunity, Health and Wellness, Personal Freedom and Choice, Índice de Felicidade - Happy Planet Index, Índice de GINI e Índice de Democracia - Democracy Index – 2014.

Indicadores ou variáveis componentes: Inclui 14 variáveis como: Satisfied demand for contraception (% of women), Greenhouse gas emissions (CO2 equivalents per GDP) - (4=low; 0=high), Water withdrawals as a percent of resources, Biodiversity and habitat (0=no protection; 100=high protection), Modern slavery, human trafficking and child marriage (1=low; 100=high), Women treated with respect (0=low; 100=high), Tolerance for immigrants (0=low; 100=high), Tolerance for homosexuals (0=low; 100=high), Discrimination and violence against minorities (0=low; 10=high), Religious tolerance (1=low; 4=high), Community safety net (0=low; 100=high), Years of tertiary schooling, Women's average years in school e Inequality in the attainment of education (0=low; 1=high)

Quadro 1. As 30 variáveis do eixo temático: **Biodiversidade - Capital Natural e Social**

Variável	Significado	Tipo	Unidade de Medida
Social Progress Index	O Índice de Progresso Social (Social Progress Index), é composto por três dimensões do progresso social: Necessidades Humanas Básicas, Fundamentos de Bem-Estar e Oportunidades. fonte: www.socialprogressimperative.org	Variável Quantitativa	Indicador nominal
Basic Human Needs	Necessidades básicas do ser humano (alimentação básica e cuidados médicos). É medido por uma escala de 1 a 100, sendo: 1 para baixa atenção e 100 para alta atenção às necessidades básicas do ser humano.	Variável Quantitativa	Indicador nominal
Foundations of Wellbeing	Fundamentos do Bem Estar e melhoria de vida através da educação	Variável Quantitativa	Indicador nominal
Opportunity	Indicador composto por: Direitos individuais, Liberdades individuais, Tolerância e inclusão e Acesso à educação superior	Variável Quantitativa	Indicador nominal
Health and Wellness	Qualidade de vida, através de melhores condições de saúde e de bem estar	Variável Quantitativa	Indicador nominal

Personal Freedom and Choice	Liberdade na tomada de decisão e escolhas de vida do indivíduo e da população	Variável Quantitativa	Percentual
Greenhouse gas emissions (CO2 equivalents per GDP)	Emissão de poluentes na atmosfera. Escala de 0 a 4, sendo: 4: < 100, 3: 100 – 200, 2: 200 – 1000, 1: 1000 – 2000 e 0: > 2000	Variável Quantitativa	Indicador nominal que varia de 0 a 4
Water withdrawals as a percent of resources	Uso de água como porcentagem dos recursos. Obedece uma escala de 0 a 5, sendo: 4–5: Extremely high stress (>80%), 3–4: High stress (40–80%), 2–3: Medium-high stress (20–40%), 1–2: Low-medium stress (10–20%) e 0–1: Low stress (<10%)	Variável Quantitativa	Indicador nominal que varia de 0 a 5
Biodiversity and habitat (0=no protection; 100=high protection)	Variável composta por componenentes que retratam o nível de proteção das áreas terrestres e marinhas, bem como as espécies ameaçadas ou em vias de extinção e seus biomas.	Variável Quantitativa	Indicador nominal que varia de 1 a 100
Modern slavery, human trafficking and child marriage (1=low; 100=high)	Esta variável é a combinação de 3 variáveis que medem: a escravidão nos tempos atuais (privação da liberdade e trabalho forçado), o tráfico de pessoas e exploração, e o casamento forçado de indivíduos menores de 18 anos.	Variável Quantitativa	Indicador nominal que varia de 1 a 100
Satisfied demand for contraception (% of women)	Variável que mede o grau de satisfação de mulheres com idade entre 15 e 49 anos, satisfeitas com os métodos contraceptivos disponíveis.	Variável Quantitativa	Percentual
Women treated with respect (0=low; 100=high)	Variável que contabiliza respostas SIM (positivas) de mulheres respondentes sobre sua percepção sobre serem tratadas com respeito e dignidade em seus países.	Variável Quantitativa	Percentual
Tolerance for immigrants (0=low; 100=high)	Variável que contabiliza respostas SIM (positivas) da população sobre a percepção de onde vivem ser um bom lugar para imigrantes de outros países.	Variável Quantitativa	Percentual
Tolerance for homosexuals (0=low; 100=high)	Variável que contabiliza respostas SIM (positivas) da população sobre a percepção de onde vivem ser um bom lugar para a comunidade LGBT (gays, lésbicas,...).	Variável Quantitativa	Percentual
Discrimination and violence against minorities (0=low; 10=high)	Variável que mede o grau de intolerância, discriminação e violëncia contra as minorias e desfavorecidos (éticas, religiosas,...). É medida por uma escala de 0 a 10, sendo: 0 (low pressures) to 10 (very high pressures).	Variável Quantitativa	Indicador nominal que varia de 0 a 10
Religious tolerance (1=low; 4=very high)	Esta variável é a medida de 13 tipos de hostilidades causadas por indivíduos, organizações ou sociedade (assédio, abuso, intimidação,..)	Variável Quantitativa	Indicador nominal que varia de 1 a 4

Community safety net (0=low; 100=high)	Variável que contabiliza respostas SIM (positivas) da população sobre uma situação problemática, onde possam contar com seus familiares ou amigos onde estiverem.	Variável Quantitativa	Percentual
Years of tertiary schooling	Média de anos para completer o terceiro grau (ensino superior) para pessoas com idade superior a 25 anos de idade. Quanto maior o indicador, melhor o resultado do país.	Variável Quantitativa	Indicador nominal (em anos)
Women's average years in school	Média de anos para mulheres que frequentaram a escola, com idade entre 25 e 34 anos. Incluindo primeiro, segundo e terceiro grau	Variável Quantitativa	Indicador nominal (em anos)
Inequality in the attainment of education (0=low; 1=high)	Variável mede a perda no aprendizado devido a falta de qualidade no ensino. Escala de 0, para baixa perda e 1 para alta perda.	Variável Quantitativa	Indicador nominal que varia de 0 a 1
IDH - 2013	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável Quantitativa	Indicador nominal
1. Governança	Indicador que mede a capacidade de um país de eleger, monitorar e substituir seus governantes eleitos. Mede também nível de estabilidade política, medidas contra corrupção, entre outros.	Variável Quantitativa	Indicador nominal
EPI Score	Índice de Performance Ambiental, composto por: Saúde Ambiental (saúde humana e dos ecossistemas), Impactos na Saúde (medido pela mortalidade infantil – idade entre 1 e 5 anos), Qualidade do Ar (poluição do ar, partículas sólidas e seu impacto na saúde) e Nível de Água tratada e Sanitização adequada.	Variável Quantitativa	Indicador nominal
EV - Water Resources	Wastewater treatment level weighted by connection to wastewater treatment rate.	Variável Quantitativa	Indicador nominal
EV - Agriculture	Variável que mede os subsídios destinados à Agricultura e a Regulamentação no uso de pesticidas conforme tratado assinado na Convenção de Estocolmo.	Variável Quantitativa	Indicador nominal
EV - Forests	Variável que demonstra a variação das áreas cobertas por floresta (desmatamento, queima e exploração).	Variável Quantitativa	Indicador nominal
EV - Fisheries	Pesca costeira intensiva com a utilização de equipamentos de arrasto e dragagem (medido em ton métrica), e nível de estocagem do excedente produzido. Quanto maior o índice, mais intensiva é a prática. (Coastal Shelf Fishing Pressure and Fish Stocks).	Variável Quantitativa	Indicador nominal
Happy Planet Index	HPI, Índice de Felicidade do Planeta, medida do Bem Estar Sustentável (Longevidade, Felicidade e Vida sustentável percebida pela população do país).	Variável Quantitativa	Indicador nominal
GINI Index	Medida de desigualdade social utilizada para medir a desigualdade na distribuição de renda, ou desigualdade da riqueza. Valores mais baixos	Variável Quantitativa	Indicador nominal

	representam menor desigualdade.		
Democracy Index - 2014	A variável Índice de Demogracia é formada por 5 categorias: Processo eleitoral e pluralismo, Liberdade civil, Funcionamento do governo, Participação política e Cultura política.	Variável Quantitativa	Indicador nominal

Fonte: autor a partir dos dados da planilha estatística e do GPS p.16

1.1.2.2 A Tabela de Dados

Em função da extensão da tabela, vamos representar aqui **apenas alguns países** e **algumas** variáveis selecionadas para ilustrar as definições do Quadro 1.

Tabela 1: Países e variáveis selecionadas

PAÍS (132)	Country Code	1	48	49	50	58	59	61	62	63
		Social Progress Index	Greenhouse gas emissions (CO2 equivalents per GDP)	Water withdrawals as a percent of resources	**Biodiversity and habitat (0=no protection; 100=high protection)**	Modern slavery, human trafficking and child marriage (1=low; 100=high)	Satisfied demand for contraception (% of women)	Women treated with respect (0=low; 100=high)	Tolerance for immigrants (0=low; 100=high)	Tolerance for homosexuals (0=low; 100=high)
Albania	ALB	69.13	267.57	1.90	63.19	11.07	82.40	44.00	45.00	12.00
Algeria	DZA	59.13	629.65	3.44	34.86	7.29	82.50	49.00	49.00	21.00
Angola	AGO	39.93	2227.70	1.54	41.70	4.84	31.20	62.00	41.00	15.00
Argentina	ARG	70.59	618.53	2.51	44.88	4.35	86.90	26.00	72.00	64.00
Armenia	ARM	65.03	886.10	3.07	80.16	10.57	74.50	68.00	52.00	3.00
Australia	AUS	86.10	769.09	3.51	83.08	2.15	87.00	83.00	87.00	72.00
Austria	AUT	85.11	284.46	0.32	86.59	1.14	80.40	78.00	63.00	48.00
Azerbaijan	AZE	62.44	795.71	3.39	58.66	11.07	78.60	77.00	50.00	6.00
Bangladesh	BGD	52.04	588.57	0.65	39.68	9.54	82.40	81.00	57.00	8.00
Belarus	BLR	65.20	816.25	0.59	41.25	4.82	86.50	67.00	46.00	10.00
Belgium	BEL	82.63	439.90	3.16	57.56	1.14	89.90	77.00	64.00	67.00
Benin	BEN	49.11	1370.12	0.00	65.88	23.57	40.60	65.00	77.00	4.00
Bolivia	BOL	62.90	1387.13	0.68	81.04	9.30	75.10	34.00	69.00	37.00
Bosnia and Herzegovina	BIH	64.99	1021.09	0.02	2.50	10.45	70.60	42.00	51.00	10.00
Botswana	BWA	65.60	458.01	1.36	100.00	19.92	74.00	71.00	60.00	24.00
Brazil	BRA	69.97	590.89	0.91	66.74	7.16	91.50	28.00	69.00	62.00
Bulgaria	BGR	70.24	713.96	1.27	68.85	10.80	85.50	61.00	46.00	29.00
Burkina Faso	BFA	47.33	1088.93	0.00	83.75	20.65	35.80	67.00	86.00	5.00
Burundi	BDI	37.33	8442.23	0.00	30.29	19.64	43.00	52.00	57.00	3.00
Cambodia	KHM	51.89	934.71	0.44	78.93	19.99	67.60	92.00	25.00	22.00
Cameroon	CMR	45.51	2241.41	0.11	45.67	20.87	53.50	62.00	74.00	8.00
Canada	CAN	86.95	604.51	1.21	58.40	1.73	90.20	85.00	93.00	81.00

Central African Republic	CAF	34.17	28763.64	0.00	97.00	20.36	53.40	61.00	74.00	11.00
Chad	TCD	32.60	2523.22	0.52	55.34	20.19	20.30	36.00	61.00	4.00
Chile	CHL	76.30	418.10	3.21	60.16	7.44	82.20	37.00	68.00	58.00
China	CHN	58.67	1138.49	2.94	66.63	8.59	96.20	89.00	45.00	14.00
Colombia	COL	67.24	443.28	0.33	80.40	9.50	90.40	17.00	68.00	47.00
Congo, Republic of	COG	47.99	976.94	0.43	66.29	21.01	69.10	54.00	72.00	9.00
Costa Rica	CRI	77.75	308.71	1.94	89.55	1.66	94.50	37.00	77.00	59.00
Croatia	HRV	73.31	380.19	0.02	70.87	10.32	85.10	47.00	35.00	22.00
Cuba	CUB	61.07	471.29	3.19	56.06	1.28	88.30		50.00	54.00
Czech Republic	CZE	80.41	469.03	1.13	98.35	10.82	87.60	70.00	49.00	43.00
Denmark	DNK	86.55	322.54	0.95	67.74	1.14	88.90	91.00	77.00	67.00
Djibouti	DJI	45.95	634.45	3.39	15.45	10.00	44.30	84.00	56.00	5.00
Dominican Republic	DOM	63.03	365.06	3.49	92.94	9.01	84.90	18.00	63.00	34.00
Ecuador	ECU	68.15	461.80	1.86	88.09	9.78	88.90	43.00	60.00	36.00
Egypt	EGY	59.97	639.20	1.33	65.26	5.09	84.30	69.00	25.00	
El Salvador	SLV	64.70	331.03	1.49	24.53	7.12	89.30	27.00	61.00	50.00
Estonia	EST	81.28	992.42	2.75	100.00	3.47	84.30	61.00	35.00	27.00
Finland	FIN	86.91	413.48	0.98	61.86	1.14	90.70	84.00	72.00	66.00
France	FRA	81.11	283.98	1.75	54.45	2.13	93.00	69.00	66.00	62.00
Georgia	GEO	63.94	691.05	1.51	29.59	11.56	73.90	81.00	47.00	6.00
Germany	DEU	84.61	337.60	1.90	100.00	2.63	89.00	78.00	75.00	63.00
Ghana	GHA	55.96	724.48	0.11	34.26	21.04	40.30	69.00	59.00	4.00
Greece	GRC	73.43	434.13	3.27	66.49	1.14	86.50	64.00	40.00	21.00
Guatemala	GTM	61.37	566.93	1.01	57.55	5.75	69.30	26.00	45.00	22.00
Guinea	GIN	37.41	2030.19	0.06	41.00	21.26	30.50	57.00	74.00	6.00
Guyana	GUY	60.06	1833.51	1.78	18.66	9.29	58.10		56.00	19.00
Honduras	HND	61.28	705.68	0.07	76.81	5.42	81.70	35.00	56.00	32.00
Hungary	HUN	73.87	409.82	0.49	29.57	10.82	90.70	58.00	53.00	24.00
Iceland	ISL	88.07	407.88	0.14	81.90	1.00		80.00	78.00	79.00
India	IND	50.24	618.09	3.58	39.18	30.84	80.00	61.00	30.00	11.00
Indonesia	IDN	58.98	883.43	3.26	78.08	5.08	82.10	66.00	22.00	2.00
Iran	IRN	56.65	934.71	4.78	37.06	5.70	89.60	65.00	50.00	
Iraq	IRQ	44.84	2189.29	3.48	0.00	5.61	71.40	30.00	37.00	10.00
Ireland	IRL	84.05	374.49	2.92	18.74	1.00	85.30	88.00	84.00	75.00
Israel	ISR	71.40	447.89	4.83	39.34	5.23	88.60	61.00	32.00	31.00
Italy	ITA	76.93	314.44	3.35	79.77	3.12	84.40	64.00	70.00	56.00
Jamaica	JAM	70.39	680.86	5.00	88.04	4.03	86.50	29.00	74.00	15.00
Japan	JPN	84.21	328.44	3.05	73.53	3.31	75.90	62.00	58.00	34.00
Jordan	JOR	61.92	810.02	4.59	8.52	7.87	80.80	74.00	31.00	
Kazakhstan	KAZ	59.47	1376.80	4.02	14.11	8.51	79.80	73.00	52.00	7.00
Kenya	KEN	50.20	771.43	0.68	67.96	5.72	65.60	48.00	54.00	12.00
Korea, Republic of	KOR	77.18	512.75	3.54	50.40	2.32	93.40	55.00	56.00	14.00
Kuwait	KWT	70.66	1666.98	4.96	92.88	6.63	76.00	92.00	69.00	
Kyrgyzstan	KGZ	57.08	994.05	4.82	15.77	8.73	80.20	59.00	47.00	5.00
Laos	LAO	52.41	1490.45	0.01	93.85	20.38	66.70	88.00	63.00	41.00

Latvia	LVA	73.91	422.72	0.35	95.17	3.21	85.50	56.00	34.00	20.00
Lebanon	LBN	60.05	424.72	4.54	1.71	4.72	81.40	60.00	53.00	7.00
Lesotho	LSO	48.94	850.05	3.97	1.30	19.34	67.10	61.00	44.00	18.00
Liberia	LBR	44.02	833.81	0.27	4.13	20.32	27.00	72.00	71.00	17.00
Lithuania	LTU	73.76	447.46	1.19	91.83	3.63	79.50	48.00	36.00	16.00
Macedonia	MKD	68.33	593.75	2.65	27.54	8.94	82.60	64.00	51.00	12.00
Madagascar	MDG	44.28	2699.67	1.25	37.08	5.22	66.30	38.00	54.00	9.00
Malawi	MWI	48.79	1491.23	0.11	87.81	20.20	62.70	62.00	55.00	5.00
Malaysia	MYS	70.00	738.12	2.09	93.37	4.48	76.50	82.00	20.00	6.00
Mali	MLI	46.85	1684.87	0.20	14.30	20.07	24.20	84.00	80.00	2.00
Mauritania	MRT	43.11	1191.34	0.60	32.63	97.90	28.10	79.00	54.00	10.00
Mauritius	MUS	73.68	397.30		43.70	1.81	92.30	73.00	79.00	24.00
Mexico	MEX	66.41	487.36	3.52	62.32	5.57	86.90	51.00	42.00	39.00
Moldova	MDA	60.12	1083.18	1.46	7.51	25.68	85.30	34.00	41.00	9.00
Mongolia	MNG	58.97	2354.61	4.05	73.25	5.45	81.60	50.00	38.00	14.00
Montenegro	MNE	66.80	6220.00	0.31	51.60	10.94	72.60	45.00	54.00	15.00
Morocco	MAR	58.01	322.89	4.24	31.20	7.23	85.60	56.00	47.00	7.00
Mozambique	MOZ	45.23	1356.28	0.82	63.17	20.07	35.90	40.00	77.00	38.00
Namibia	NAM	61.19	1185.09	1.88	94.39	18.89	74.50		56.00	18.00
Nepal	NPL	51.58	1135.44	2.40	61.79	26.56	65.00	50.00	60.00	23.00
Netherlands	NLD	87.37	465.89	1.73	94.69	2.13	86.50	86.00	79.00	85.00
New Zealand	NZL	88.24	703.28	1.35	76.05	1.10	89.10	82.00	87.00	69.00
Nicaragua	NIC	62.33	920.17	1.01	91.67	5.47	87.50	38.00	64.00	57.00
Niger	NER	40.10	2242.32	0.11	77.21	20.50	39.00	78.00	72.00	3.00
Nigeria	NGA	42.65	915.58	0.29	57.54	13.99	40.40	51.00	62.00	7.00
Norway	NOR	87.12	243.66	0.40	71.68	1.14	93.70	86.00	74.00	66.00
Pakistan	PAK	42.40	753.36	4.31	41.43	32.11	55.50	69.00	38.00	30.00
Panama	PAN	72.58	525.57	0.42	86.96	1.67	73.40	47.00	53.00	44.00
Paraguay	PRY	62.65	1167.57	0.01	31.20	9.49	92.70	32.00	87.00	45.00
Peru	PER	66.29	320.51	3.20	70.36	10.04	91.30	24.00	61.00	34.00
Philippines	PHL	65.86	448.50	3.33	64.67	6.73	68.70	78.00	63.00	59.00
Poland	POL	77.44	568.80	1.31	91.76	10.32	87.50	63.00	39.00	19.00
Portugal	PRT	80.49	345.44	3.34	64.36	1.64	92.80	58.00	75.00	48.00
Romania	ROU	67.72	559.77	0.84	63.51	4.00	88.00	35.00	40.00	17.00
Russia	RUS	60.79	1151.93	1.23	53.39	12.43	92.70	46.00	42.00	12.00
Rwanda	RWA	49.46	419.68	0.00	56.32	19.62	70.60	97.00	56.00	2.00
Saudi Arabia	SAU	64.38	799.83	4.99	93.70	8.73	56.50	89.00	69.00	
Senegal	SEN	53.52	986.77	0.21	85.36	22.31	27.70	62.00	77.00	3.00
Serbia	SRB	70.61	975.08	0.58	99.78	10.45	81.70	40.00	68.00	21.00
Slovakia	SVK	78.93	399.86	0.20	84.33	10.32	88.60	65.00	41.00	27.00
Slovenia	SVN	81.65	385.03	0.03	100.00	9.83	90.60	65.00	56.00	38.00
South Africa	ZAF	62.96	1176.39	3.04	63.96	4.99	83.40	54.00	56.00	47.00
Spain	ESP	80.77	329.19	3.73	56.56	1.64	83.50	69.00	81.00	80.00
Sri Lanka	LKA	59.71	411.41	3.01	67.70	4.81	90.80	72.00	44.00	4.00
Sudan	SDN	38.45	2084.31	0.91	16.29	20.86	29.00	60.00	53.00	7.00
Swaziland	SWZ	48.87	489.39	3.11	17.69	3.95	79.00	56.00	55.00	18.00

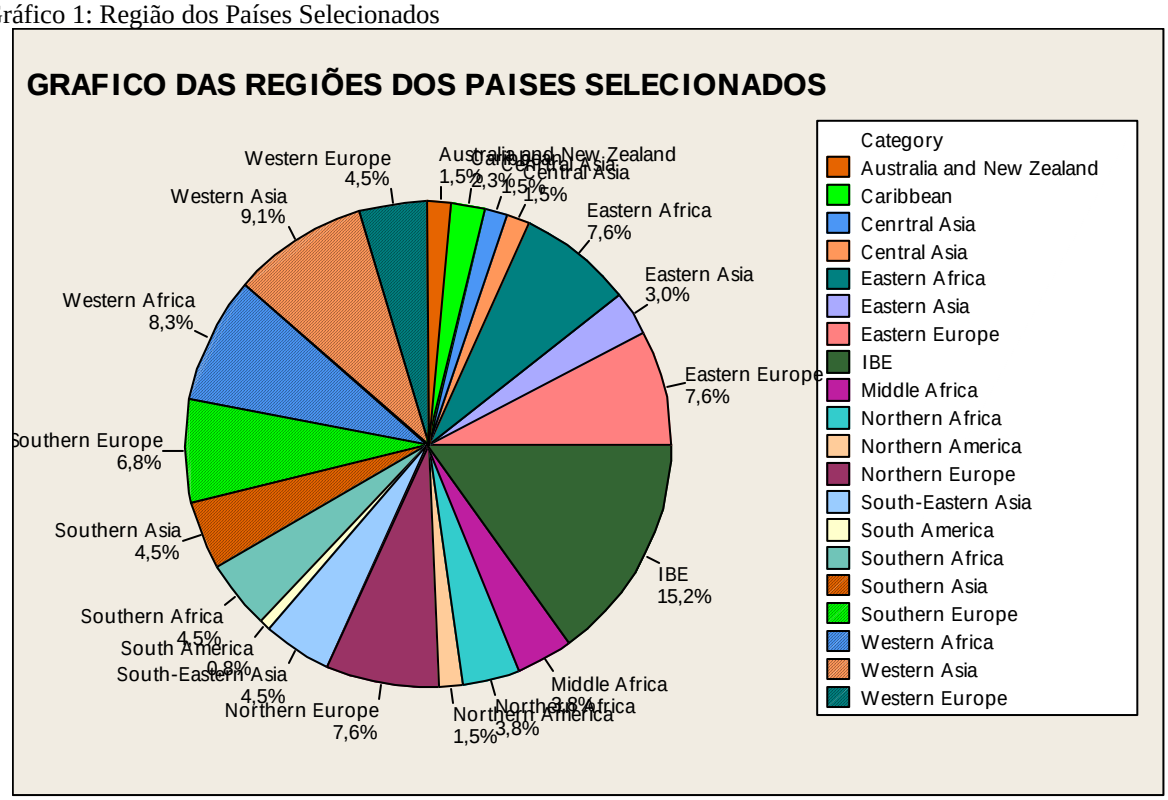
Sweden	SWE	87.08	231.19	1.30	62.38	1.14	88.30	73.00	80.00	69.00
Switzerland	CHE	88.19	170.50	1.06	100.00	1.14	92.10	90.00	70.00	65.00
Tajikistan	TJK	56.05	782.66	3.53	24.28	9.59	65.00	87.00	44.00	1.00
Tanzania	TZA	46.06	1292.02	1.50	79.12	20.47	57.30	55.00	43.00	16.00
Thailand	THA	65.14	720.14	1.70	70.19	20.78	93.80	80.00	23.00	27.00
Togo	TGO	42.80	1493.51	0.12	44.23	21.53	31.60	58.00	67.00	8.00
Trinidad and Tobago	TTO	69.88	1734.18	5.00	83.48	2.81	70.50	42.00	72.00	46.00
Tunisia	TUN	62.96	310.47	3.44	31.08	4.35	83.90	80.00	56.00	11.00
Turkey	TUR	64.62	422.04	3.02	32.62	7.40	89.30	38.00	47.00	14.00
Uganda	UGA	47.75	716.27	0.00	91.31	20.73	44.30	67.00	63.00	5.00
Ukraine	UKR	64.91	1411.36	2.10	41.46	8.42	86.30	44.00	38.00	14.00
United Arab Emirates	ARE	72.92	832.63	5.00	96.47	7.74	66.90	99.00	80.00	
United Kingdom	GBR	84.56	307.12	2.63	70.11	1.00	94.70	88.00	77.00	75.00
United States	USA	82.77	528.55	2.89	63.35	2.77	92.90	75.00	82.00	70.00
Uruguay	URY	77.51	817.95	0.86	24.62	9.51	91.50	47.00	82.00	76.00
Uzbekistan	UZB	57.34	2796.96	4.32	12.87	15.67	87.70	96.00	51.00	5.00
Venezuela	VEN	63.78	829.55	2.30	92.92	9.34	85.50	59.00	67.00	51.00
Yemen	YEM	40.23	507.26	4.67	23.59	7.28	53.40	68.00	25.00	
Zambia	ZMB	49.88	3734.91	0.08	100.00	19.83	63.10	61.00	68.00	5.00

Fonte: autor com base na tabela de dados da aula de Métodos Quantitativos

1.3. ANÁLISE DAS VARIÁVEIS

Apresentamos o gráfico circular ou gráfico de pizza (ou gráfico de torta) elaborado pelo *software* MINITAB com as regiões dos países selecionados**.**

Gráfico 1: Região dos Países Selecionados



1.3.1 – Variáveis Quantitativas

Vemos no Quadro 1, as variáveis e suas descrições. Essa informação aparece na Tabela 1 com uma seleção de países e indicadores com os números em percentual.
Com isso, a variável é denominada quantitativa pois mostra que ela pode ser medida em uma escala quantitativa, isto é, com números (SANTOS; PARRA FILHO, 2011).
Elas são Variáveis Contínuas porque assumem valores fracionados já que estão em percentual **(IDH e GINI).**

1.3.2 – Média, Mínimo e Máximo das variáveis com uso do MINITAB

Para facilitar o acompanhamento das análises, algumas definições se fazem necessárias para entender a apresentação dos dados do *software* MINITAB.

- Valor de N é o total de elementos da amostra, que neste caso é de: 132 países
- **N*** indica o total de dados ausentes na tabela em cada variável. Por exemplo, a variável EV – Fisheries, mostra um total de 37 países não divulgaram seus resultados.
- Mean (média) mostra o valor médio da variável para os 132 países.
- SE Mean mostra o erro padrão da amostra, isto é, ele mostra o quanto a média varia em relação as outras da amostra.
- StDev é o desvio-padrão mostra a distância estimada entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação a média.
- Minimum mostra o valor mínimo dos dados dos 132 países
- Maximum mostra o valor máximo dos dados dos 132 países

Abaixo a tabela de estatística descritiva para os dados da base referente ao tema **Biodiversidade - Capital Natural e Social**.

Descriptive Statistics: Social Progr, Basic Human , Foundations , ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63.67	1.24	14.20	32.60	51.93
Basic Human Needs	132	0	70.00	1.63	18.77	25.94	55.14
Foundations of Wellbeing	132	0	67.529	0.994	11.424	41.340	59.465
Opportunity	132	0	53.48	1.35	15.56	24.31	42.38
Health and Wellness	132	0	72.458	0.606	6.958	49.930	68.225
Personal Freedom and Cho	132	0	67.52	1.12	12.87	17.62	59.93
Greenhouse gas emissions	132	0	1177	228	2623	170	427
Water withdrawals as a p	131	1	1.933	0.136	1.559	0.000	0.442
Biodiversity and habitat	132	0	59.51	2.43	27.90	0.00	37.60
Modern slavery, human tr	132	0	10.441	0.931	10.691	1.000	4.010
Satisfied demand for con	131	1	74.58	1.68	19.18	20.30	66.90
Women treated with respe	129	3	61.57	1.67	19.01	17.00	47.50
Tolerance for immigrants	132	0	57.82	1.42	16.32	20.00	46.00
Tolerance for homosexual	125	7	28.62	2.13	23.85	1.00	8.50
Discrimination and viole	132	0	5.990	0.176	2.019	1.000	4.450
Religious tolerance (1=l	132	0	2.7803	0.0856	0.9832	1.0000	2.0000
Community safety net (0=	132	0	79.42	1.10	12.63	30.00	73.00
Years of tertiary school	117	15	0.4475	0.0344	0.3717	0.0100	0.1400
Women`s average years in	131	1	9.471	0.327	3.743	1.300	6.600
Inequality in the attain	117	15	0.1960	0.0136	0.1474	0.0130	0.0555
IDH - 2013	125	7	0.6894	0.0144	0.1611	0.3370	0.5595
1. Governança	132	0	53.38	1.88	21.59	15.22	37.11
EPI Score	132	0	52.48	1.44	16.55	18.43	39.30
EV - Water Resources	132	0	27.77	2.85	32.71	0.00	0.54

EV - Agriculture	132	0	71.37	1.65	18.98	14.66	58.52
EV - Forests	110	22	39.72	3.02	31.65	0.00	14.76
EV - Fisheries	95	37	18.09	1.34	13.08	0.00	2.52
Happy Planet Index	129	3	42.673	0.796	9.040	22.591	36.163
GINNI Index	119	13	39.024	0.782	8.532	24.820	32.780
Democracy Index - 2014	132	0	5.873	0.180	2.068	1.490	4.100

Conforme observamos na tabela acima, identificamos dados faltantes (*missing values*) para algumas variáveis, que estão evidenciadas pelo valor diferente de 0 (zero) na coluna N*. Para fins deste estudo, as células que estiverem em branco serão preenchidas com o valor correspondente a média da variável em questão.

Após verificação na base de dados, as seguintes células foram substituídas para os respectivos países que apresentaram dados faltantes (*missing values*):

Water withdrawals as a percent of resources (1 caso) - Incluímos a média 1.933 no país: Mauritius

Satisfied demand for contraception (% of women) (1 caso) - Incluímos a média 74.58 no país: Iceland

Women treated with respect (3 casos) - Incluímos a média 61.57 nos países: Cuba, Guyana e Namibia

Tolerance for homosexuals (7 casos) - Incluímos a média 28.62 nos países: Egypt, Iran, Jordan, Kuwait, Saudi Arabia, United Arab Emirates e Yemen

Years of tertiary schooling (15 casos) - Incluímos a média 0.4475 nos países: Angola, Azerbaijan, Belarus, Bosnia and Herzegovina, Burkina Faso, Chad, Djibouti, Georgia, Guinea, Lebanon, Macedonia, Madagascar, Montenegro, Nigeria e Uzbekistan

Women's average years in school (1 caso) - Incluímos a média 9.471 no país: Iceland

Inequality in the attainment of education (15 casos) - Incluímos a média 0.1960 nos países: Algeria, Botswana, Burundi, Cuba, Iran, Japan, Kuwait, Malaysia, New Zealand, Paraguay, Russia, Saudi Arabia, Sudan, Tunisia e United Arab Emirates

IDH-2013 (7 casos) - Incluímos a média 0.6894 nos países: Algeria, Kazakhstan, Korea, Republic of Kyrgyzstan, Latvia, Mauritius e Netherlands

EV-Forests (22 casos) - Incluímos a média 39.72 nos países: Benin, Botswana, Burkina Faso, Chad, Djibouti, Egypt, Iceland, Iraq, Israel, Jordan, Kuwait, Lesotho, Mali, Mauritania, Namibia, Niger, Saudi Arabia, Senegal, Sudan, Tajikistan, United Arab Emirates e Yemen

EV - Fisheries (37 casos) - Incluímos a média 18.09 nos países: Armenia, Austria, Azerbaijan, Belarus, Bolivia, Bosnia and Herzegovina, Botswana, Burkina Faso, Burundi, Central African Republic, Chad, Czech Republic, Hungary, Iraq, Jordan, Kazakhstan, Kyrgyztan, Laos, Lesotho, Macedonia, Malawi, Mali, Moldova, Mongolia, Nepal, Niger, Paraguay, Rwanda, Serbia, Slovakia, Slovenia, Swaziland, Switzerland, Tajikistan, Uganda, Uzbekistan e Zambia

Happy Planet Index (3 casos) - Incluímos a média 42.673 nos países: Lesotho, Montenegro e Swaziland

GINI Index (13 casos) - Incluímos a média 39.024 nos países: Bangladesh, Cuba, Guinea, Iceland, Kuwait, Lebanon, Mauritius, Montenegro, Nepal, Pakistan, Saudi Arabia, Sri Lanka e United Arab Emirates

Após a inclusão dos dados ausentes pela média de cada variável, foi executado novamente os indicadores de número de elementos (N), dados ausentes (N*), média, erro da média (SE mean), desvio padrão (Stdev), valor mínimo e máximo; apenas para confirmar se foram substituídos corretamente. Portanto, sem alteração nos demais valores.

Descriptive Statistics: Social Progr, Basic Human , Foundations , ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63.67	1.24	14.20	32.60	51.93
Basic Human Needs	132	0	70.00	1.63	18.77	25.94	55.14
Foundations of Wellbeing	132	0	67.529	0.994	11.424	41.340	59.465
Opportunity	132	0	53.48	1.35	15.56	24.31	42.38
Health and Wellness	132	0	72.458	0.606	6.958	49.930	68.225
Personal Freedom and Cho	132	0	67.52	1.12	12.87	17.62	59.93
Greenhouse gas emissions	132	0	1177	228	2623	170	427
Water withdrawals as a p	132	0	1.933	0.135	1.553	0.000	0.455
Biodiversity and habitat	132	0	59.51	2.43	27.90	0.00	37.60
Modern slavery, human tr	132	0	10.441	0.931	10.691	1.000	4.010
Satisfied demand for con	132	0	74.58	1.66	19.11	20.30	66.95
Women treated with respe	132	0	61.57	1.64	18.79	17.00	48.00
Tolerance for immigrants	132	0	57.82	1.42	16.32	20.00	46.00
Tolerance for homosexual	132	0	28.62	2.02	23.20	1.00	9.00
Discrimination and viole	132	0	5.990	0.176	2.019	1.000	4.450
Religious tolerance (1=l	132	0	2.7803	0.0856	0.9832	1.0000	2.0000
Community safety net (0=	132	0	79.42	1.10	12.63	30.00	73.00
Years of tertiary school	132	0	0.4475	0.0304	0.3498	0.0100	0.1525
Women`s average years in	132	0	9.471	0.325	3.728	1.300	6.625
Inequality in the attain	132	0	0.1960	0.0121	0.1387	0.0130	0.0615
IDH - 2013	132	0	0.6894	0.0136	0.1567	0.3370	0.5700
1. Governança	132	0	53.38	1.88	21.59	15.22	37.11
EPI Score	132	0	52.48	1.44	16.55	18.43	39.30
EV - Water Resources	132	0	27.77	2.85	32.71	0.00	0.54
EV - Agriculture	132	0	71.37	1.65	18.98	14.66	58.52
EV - Forests	132	0	39.72	2.51	28.87	0.00	18.04
EV - Fisheries	132	0	18.089	0.964	11.080	0.000	15.057
Happy Planet Index	132	0	42.673	0.778	8.936	22.591	36.587
GINNI Index	132	0	39.024	0.705	8.097	24.820	33.065
Democracy Index - 2014	132	0	5.873	0.180	2.068	1.490	4.100

1.4. ANÁLISE COMPARATIVA E CONSIDERAÇÕES

Aqui são apresentados os gráficos gerados pela funcionalidade *Graphical Summary* de acordo com os passos:
Primeiro: clique em *Stat* em seguida clique em *Basic Statistics*, e em *Graphical Summary*.
Segundo: Aparecerá uma janela e deverá ser escolhida a coluna (variável) com a qual deseja construir o gráfico. Não é necessário alterar o grau de confiança de 95,0. Clique OK e o gráfico será gerado em uma janela. Copie e cole no seu trabalho.

4.1 Definições e Análise das Variáveis deste estudo

4.1.1 Biodiversity and Habitat

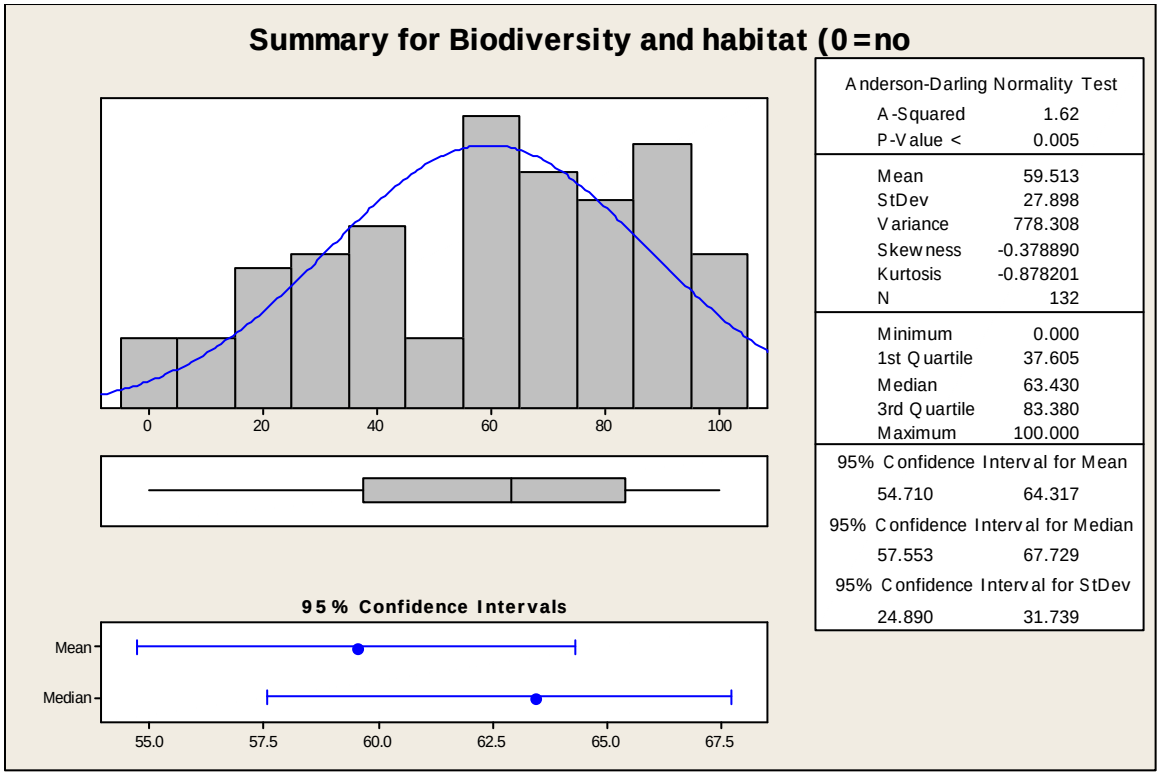


Figure 03: graph - Biodiversity and habitat (0=no protection; 100=high protection)

Definição:
A variável *Biodiversity and Habitat* (Biodiversidade: Capital Natural e Social), é formada por componentes que retratam o nível de proteção das áreas terrestres e marinhas, bem como as espécies ameaçadas ou em via de extinção e seus biomas.
source: http://epi.yale.edu/files/biodiversity_and_habitat.xls (tradução livre do autor).

Análise da variável:
A variável Biodiversity and Habitat (Biodiversidade: Capital Natural e Social) se apresenta numa escala de 0 a 100, onde o valor 0 representa nenhuma proteção à biodiversidade, e 100, alta proteção.
A curva de densidade no histograma acima indica uma distribuição próxima da normal entre os dados analisados, porém levemente deslocada para a direita. O que é evidenciada pelos valores da média (59.513) e a mediana (63.430), que resultaram relativamente próximas em seus valores. Pelo fato da média e mediana figurarem a direita do gráfico, isso nos mostra que temos mais países ativamente empenhados em cuidar de seu bioma, que o contrário. Por isso vemos 49 países abaixo da média (59.513) em contraste com outros 73 acima da média (59.513). Somente 11 países estão na faixa do valor máximo na escala, o que representa menos de 10% do total. Entre os quais somente Alemanha, Estônia, Eslovênia, Suiça, Zâmbia, Botswana, apresentaram valor máximo (100=*high protection*). O resultado do Brasil foi de 66.74 pontos, o que o coloca acima da média, e no estrato onde figuram 17 países, tais como Bulgária, Dinamarca e Egito, para citar alguns.
Identificamos 5 países no estrato com os valores mais baixos do gráfico (<5), são eles, Bósnia e Herzegovínia, Iraque, Líbano, Lesotho e Libéria. O Iraque é o único com valor 0 para esta variável (0=no protection). Este país também figura na lista de países com os menores índices

de Governança e de Bem Estar para sua população e aparece no estrato dos maiores emissores de poluentes na atmosfera (Greenhouse gas emissions > 2000).
Com 95% de confiança, podemos afirmar que a média encontra-se no intervalo entre 54.710 e 64.317 pontos.

4.1.2 SPI – Social Progress Index – Índice de Progresso Social

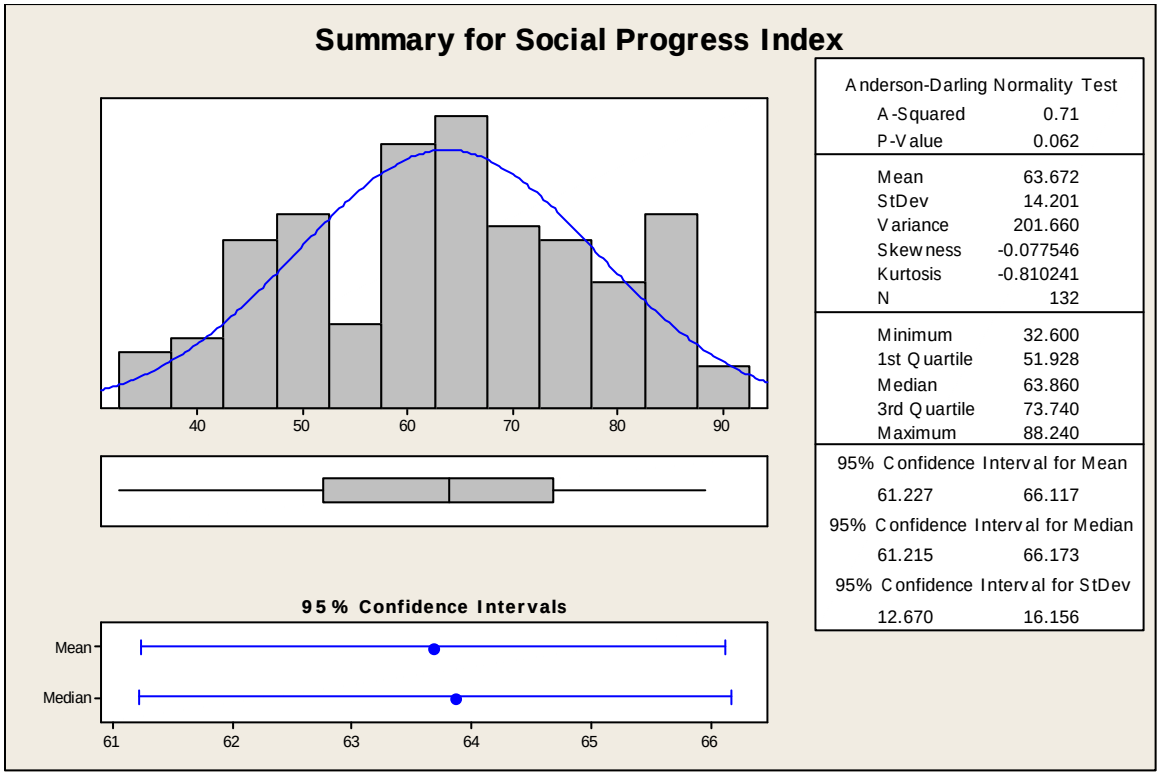


Figure 04: graph - Social Progress Index

Definição:
O Índice de Progresso Social (*Social Progress Index*) é o resultado de um processo de dois anos de pesquisas envolvendo uma equipe renomada de especialistas que inclui, entre outros, os economistas Hernando de Soto e Michael Porter e da presidente da Fundação Rockefeller, Dr. Judith Rodin. O índice sintetiza uma extensa quantidade de dados pesquisados para identificar as dimensões do desempenho das sociedades e medir o progresso social de forma abrangente e rigorosa. O índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidados em três dimensões do Progresso Social: Necessidades Humanas Básicas, Fundações de Bem-estar e Oportunidade. A primeira dimensão, necessidades humanas básicas, por exemplo, avalia o quão bem um país prevê as necessidades essenciais de seu povo através da medição que considera se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se têm acesso à água potável, acesso a uma habitação condigna com serviços básicos, e se eles estão seguros e protegidos.
Fonte: origem e definição dados pelo portal: http://pt.knoema.com/jklbfre/the-social-progress-index (tradução livre do autor).

Análise da variável:
Considerando os 132 países analisados, o histograma acima apresenta uma distribuição próxima da normal.
Há uma concentração maior de países no estrato entre 62.5 e 67.5, totalizando 21 países. A média (63.672) e a mediana (63.860) são praticamente idênticas, diferenciando-se por 0.2 (dois décimos) apenas. Somente 3 países ocupam a posição mais elevada do gráfico, compreendida entre 87.5 e 92.5, que são os países Suiça, Nova Zelândia e Islândia. Ao passo que no estrato de 32.5 e 37.5, que representa os menores valores obtidos nesta análise, há 4 observações, representadas pelos países Burundi, República da África Central, Chade e Guinea. Cerca de 70 países estão acima da média (63.672) do índice (Upper Middle Social Progress Countries).

4.1.3 Basic Human Needs – Necessidades Básicas do Ser Humano

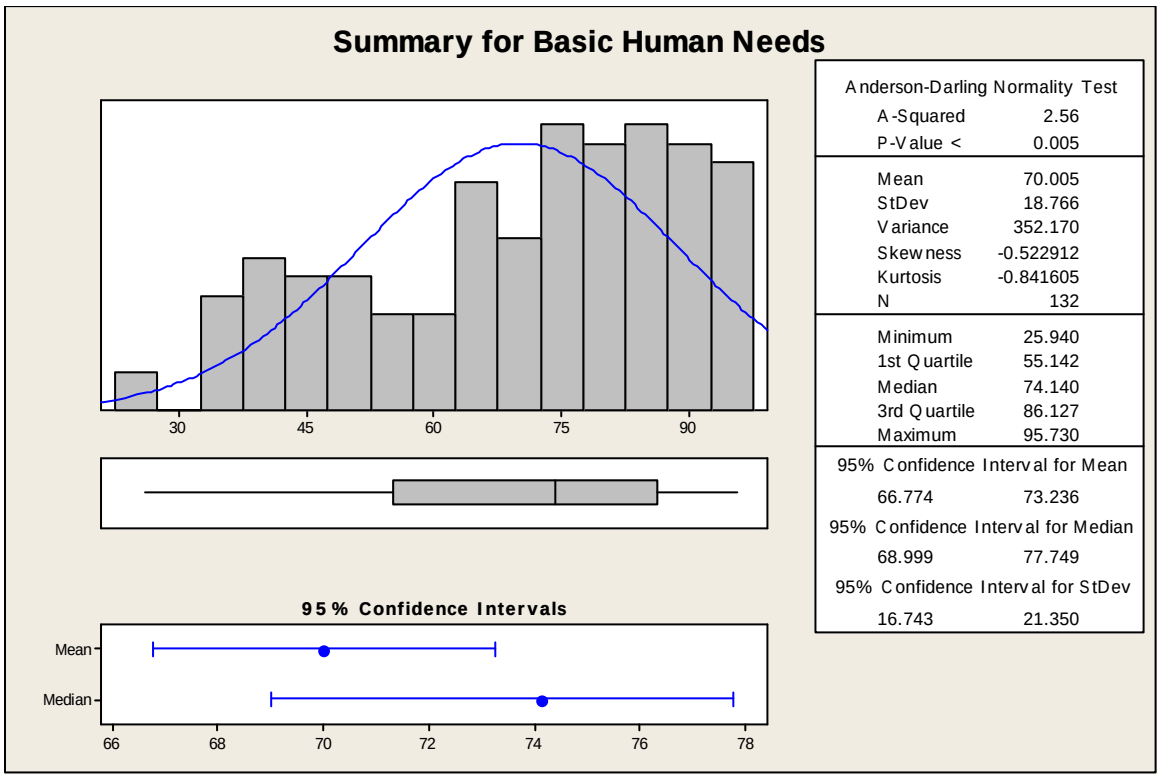


Figure 05: graph - Basic Human Needs

Definição:
Necessidades básicas do ser humano (alimentação básica e cuidados médicos)

Análise da variável:
Pela distribuição dos dados no histograma, vemos que esta distribuição não é uma normal.
A média (70.005) e a mediana (74.140) apresentam valores relativamente altos, corroborando com o que nos mostra o gráfico acima, 73 países acima da média.
Não houve observações no estrato entre 27.5 a 32.5

4.1.4 Foundations of Wellbeing - Fundamentos do Bem-Estar

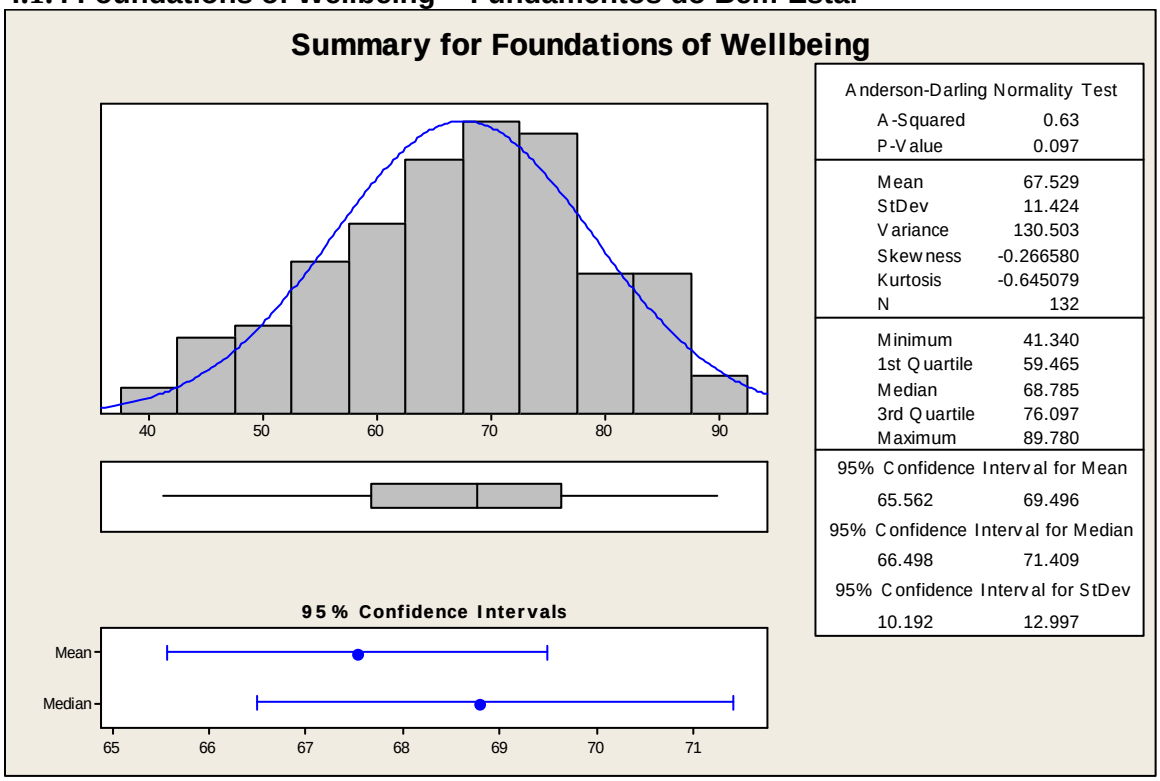


Figure 06: graph - Foundations of Wellbeing

Análise da variável:

O gráfico dos Fundamentos do Bem-Estar e melhoria de vida através da educação, mostra que a distribuição aproxima-se da normal, considerando os 132 países. A média está com o valor de 67.529, a mediana com 68.785. O estrato 67.5 e 72.5 representa a maior concentração das observações, com 23 países.

4.1.5 Opportunity

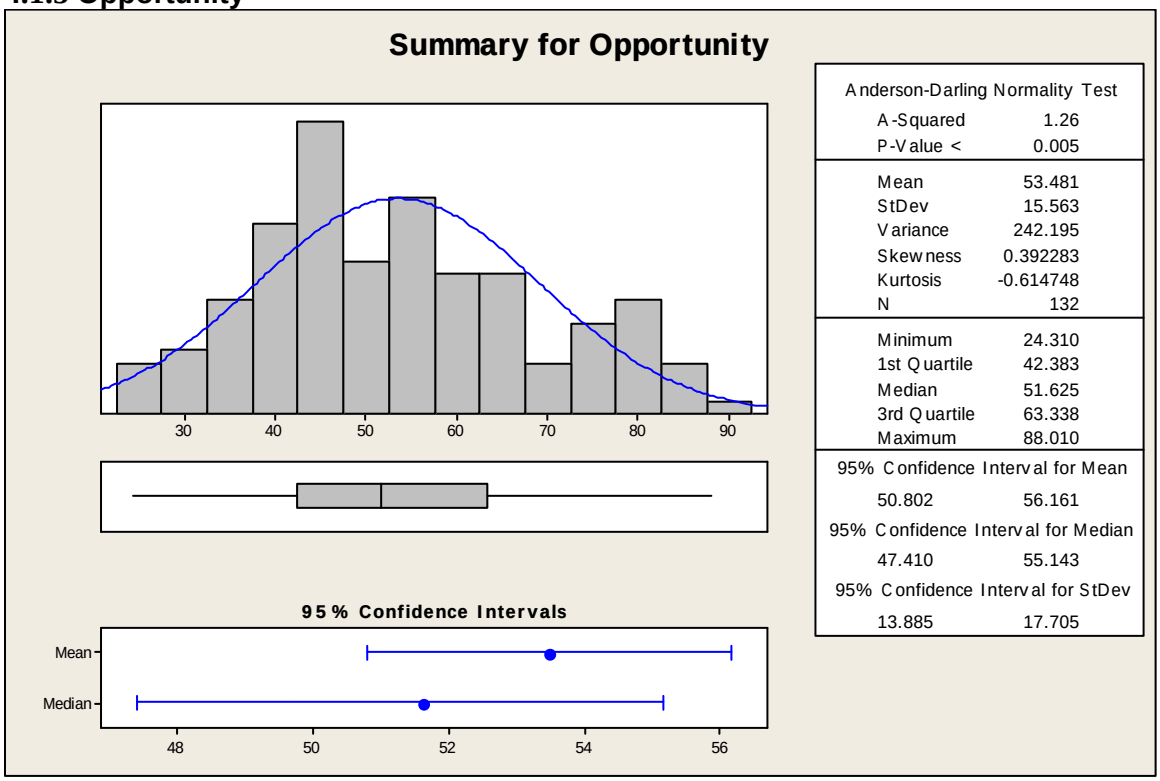


Figure 07: graph - Opportunity

Análise da variável:

Este indicador é composto por: Direitos individuais, Liberdades individuais, Tolerância e inclusão e Acesso à educação superior. E busca mostrar o quanto as pessoas são livres para exercerem seus direitos. O histograma mostra uma distribuição próxima da normal e concentra um total de 95 países no estrato entre 37.5 e 67.5. Apenas 1 país figura no estrato mais elevado do gráfico cujos valores estão entre 87.5 e 92.5, este país é a Nova Zelândia. A maior concentração de observações se encontra no estrato de 42.5 a 47.5, com 23 países.

4.1.6 Health and Wellness – Saúde e Bem-Estar

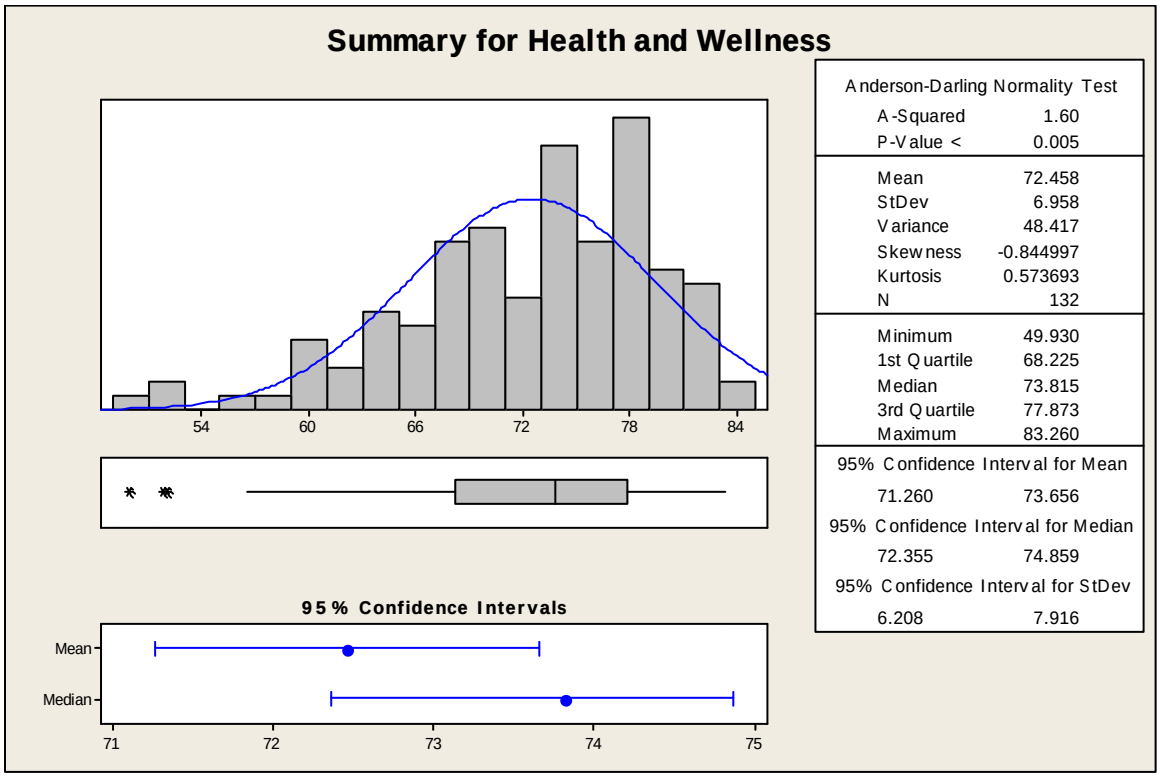


Figure 07: graph - Health and Wellness

Análise da variável:

Este indicador busca evidenciar se as pessoas vivem vidas longas e saudáveis em seus países. O histograma mostra uma distribuição próxima da normal, porém deslocada para a direita do gráfico, onde se concentra a maior parte dos dados, com média de 72.458 e mediana com valor de 73.815. O terceiro quartil com valor de 77.873, é representado pelo estrato com 21 países.

A análise nos revela 3 observações apontadas como *outliers* na parte esquerda do histograma, são eles: Cazaquistão, Rússia e Ucrânia.

4.1.7 Personal Freedom and Choice - Liberdade na tomada de decisão

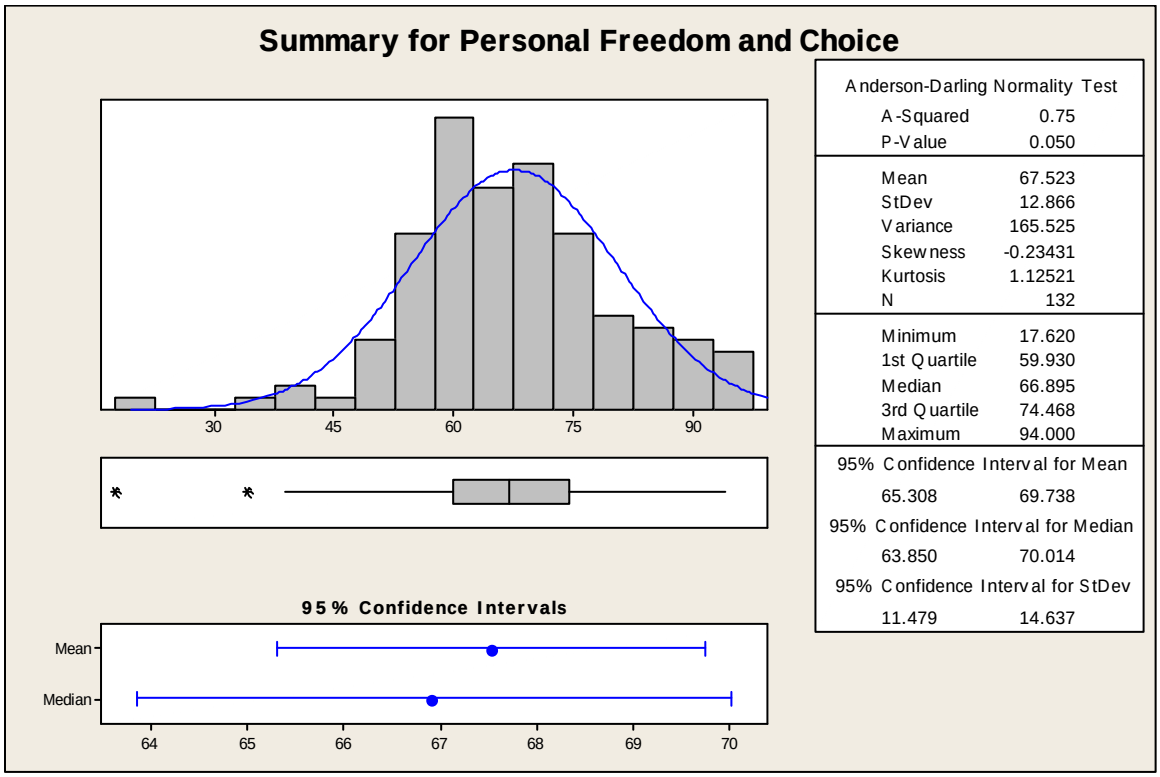


Figure 08: graph - Personal Freedom and Choice

Análise da variável:

Esta variável visa medir a liberdade na tomada de decisão e escolhas de vida do indivíduo e da população de uma país.

A distribuição aproxima-se da normal. Percebe-se uma concentração dos valores na parte direita do gráfico. A média tem valor 67.523 e a mediana com 66.895

A análise revela a presença de 2 *outliers* no conjunto de 132 países analisados, são eles: Sudão (34.05) e Mauritânia (17.62). A maior concentração de observações, se encontra no estrato entre 52.5 e 77.5, com 95 países.

4.1.8 Greenhouse gas emissions - Emissões de gases de efeito estufa

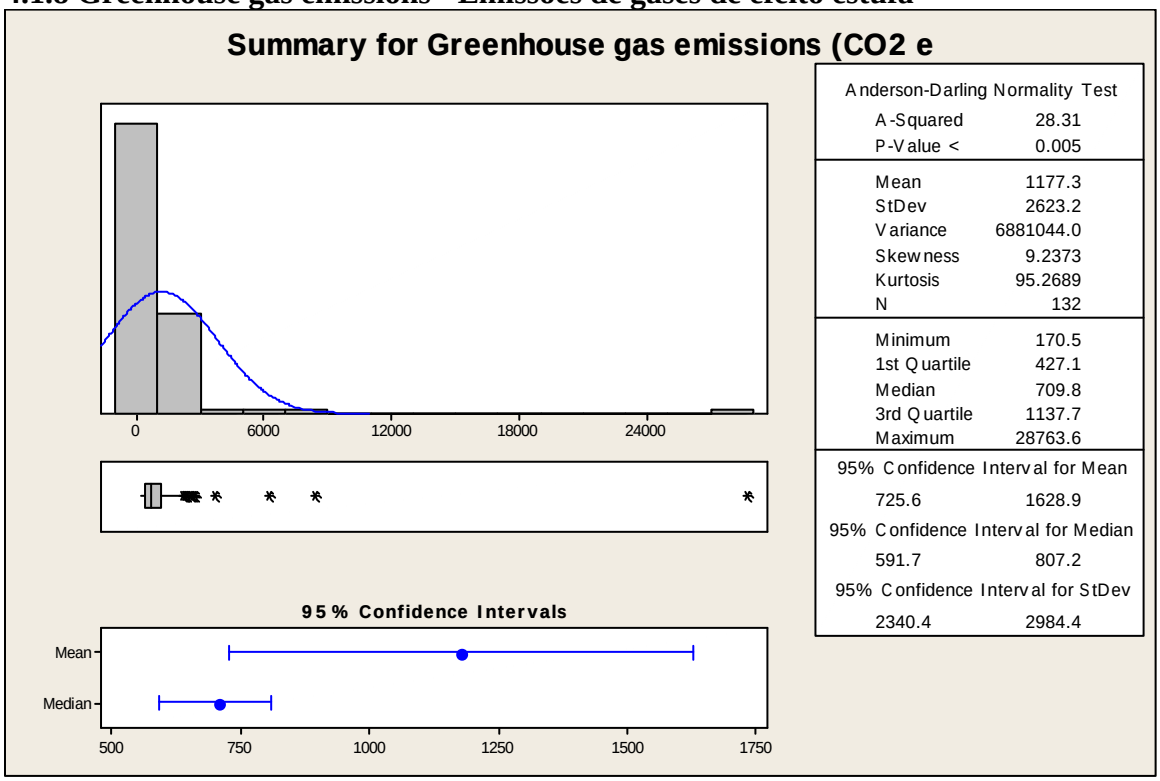


Figure 09: graph - Greenhouse gas emissions (CO2 equivalents per GDP)

Análise da variável:

As emissões de dióxido de carbono (CO2), metano (CH4), óxido nitroso (N2O), hidrofluorcarbonos (HFCs), perfluorcarbonos (PFCs) e hexafluoreto de enxofre (SF6), expressas em equivalentes de CO2 usando 100 anos potenciais de aquecimento global encontrados no Painel Intergovernamental de Mundaça de Clima - Segundo Relatório de Avaliação por PIB-PPP. No modelo SPI, os dados são escalados 0-4: 4: <100 3: 100-200 2: 200-1000 1: 1000-2000 0:> 2000. Este indicador captura a quantidade de emissão de poluentes na atmosfera.

A média apresenta valor de 1177.3 e a mediana de 709.8. Dos 132 países observados, 95 deles estão no estrato entre -1000 e 1000. Observa-se a presença de 4 *outliers* nesta análise, sendo: Burundi (8442.23), República Africana Central (28763.64), Montenegro (6220.0) e Zâmbia (3734.91).

4.1.9 Water withdrawals as a percentage of resources - Uso de água como porcentagem dos recursos

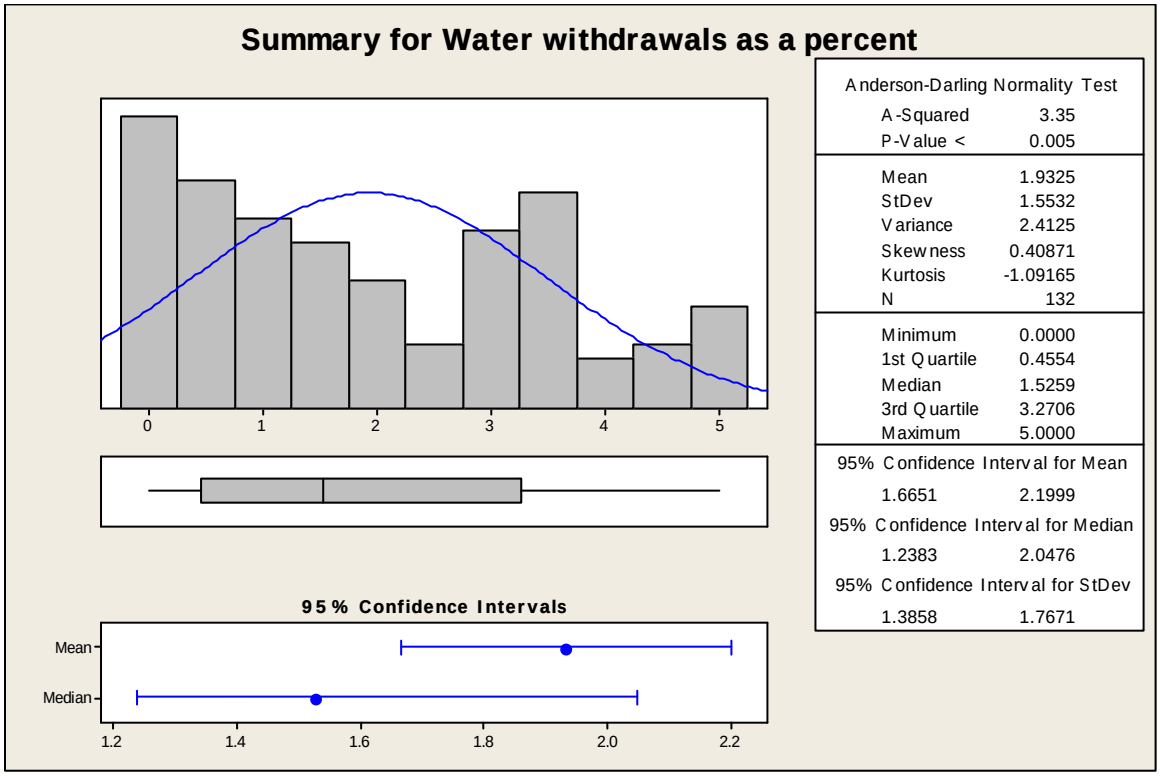


Figure 10: graph - Water withdrawals as a percentage of resources

Análise da variável:

Esta variável demonstra o uso de água como porcentagem dos recursos naturais. E apresenta uma escala de 0 a 5, sendo: De 4–5: Extremamente alto (>80%), de 3–4: Alto (40–80%), de 2–3: Médio-Alto (20–40%), de 1–2: Médio-Baixo (10–20%) e de 0–1: Baixo uso (<10%). A curva de distribuição aproxima-se da normal. E 41 países estão abaixo do coeficiente 1, denotando que aproximadamente um terço dos países analisados mostram eficiência na gestão do recurso. Enquanto 8 países figuram na faixa “Extremamente alto (>80%)”, o maior valor atingido nesta escala é 5, onde figuram 3 países, como segue: Jamaica, Trinidade e Tobago e União dos Emirados Árabes. A média é de 1.9325 e a mediana tem valor de 1.5259.

4.1.10 Modern slavery, human trafficking and child marriage - privação da liberdade e trabalho forçado, o tráfico de pessoas e exploração, e o casamento forçado de indivíduos menores de 18 anos

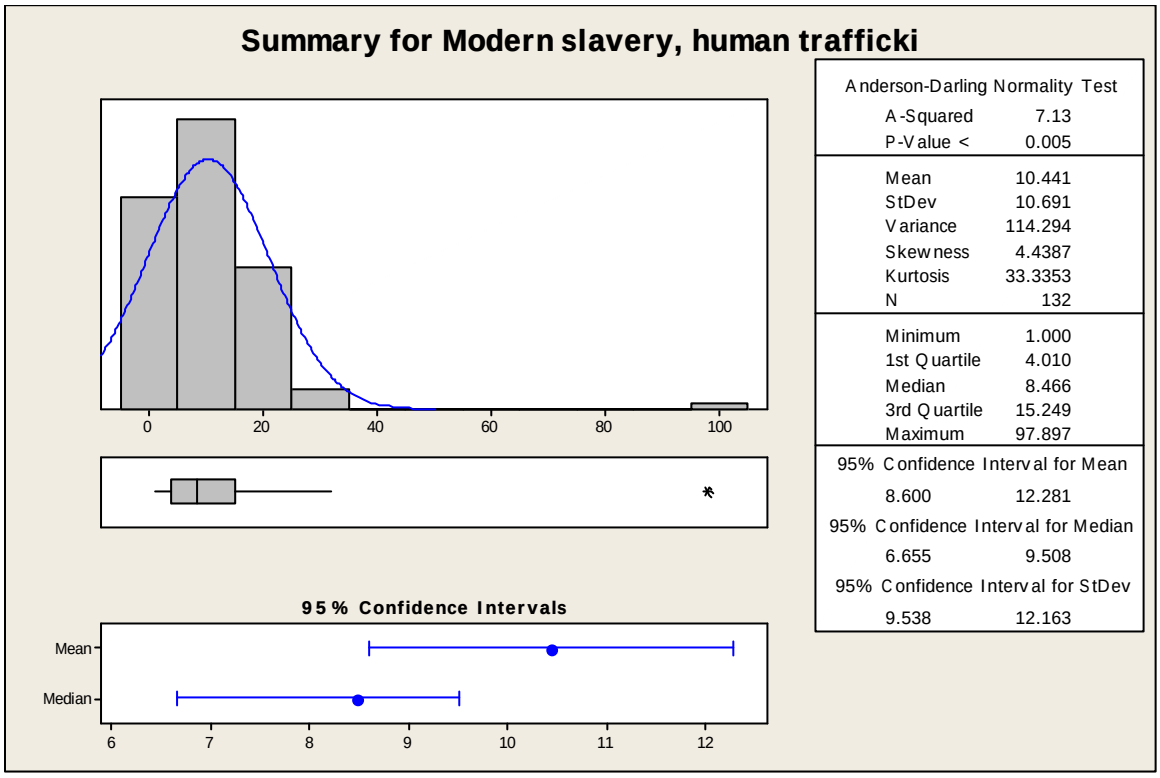


Figure 11: graph - Modern slavery, human trafficking and child marriage (1=low; 100=high)

Análise da variável:

Esta variável é a combinação de 3 variáveis que medem: a escravidão nos tempos atuais (privação da liberdade e trabalho forçado), o tráfico de pessoas e exploração, e o casamento forçado de indivíduos menores de 18 anos.

Os dados em sua maioria, total de 127 observações , se concentram na parte esquerda do gráfico. Há a presença de 1 *outlier*, com valor de 97.90 cujo país correspondente é a Mauritânia. A média é de 10.441 e a mediana de 8.466.

4.1.11 Satisfied demand for contraception (% of women) - Demanda por contracepção atendida

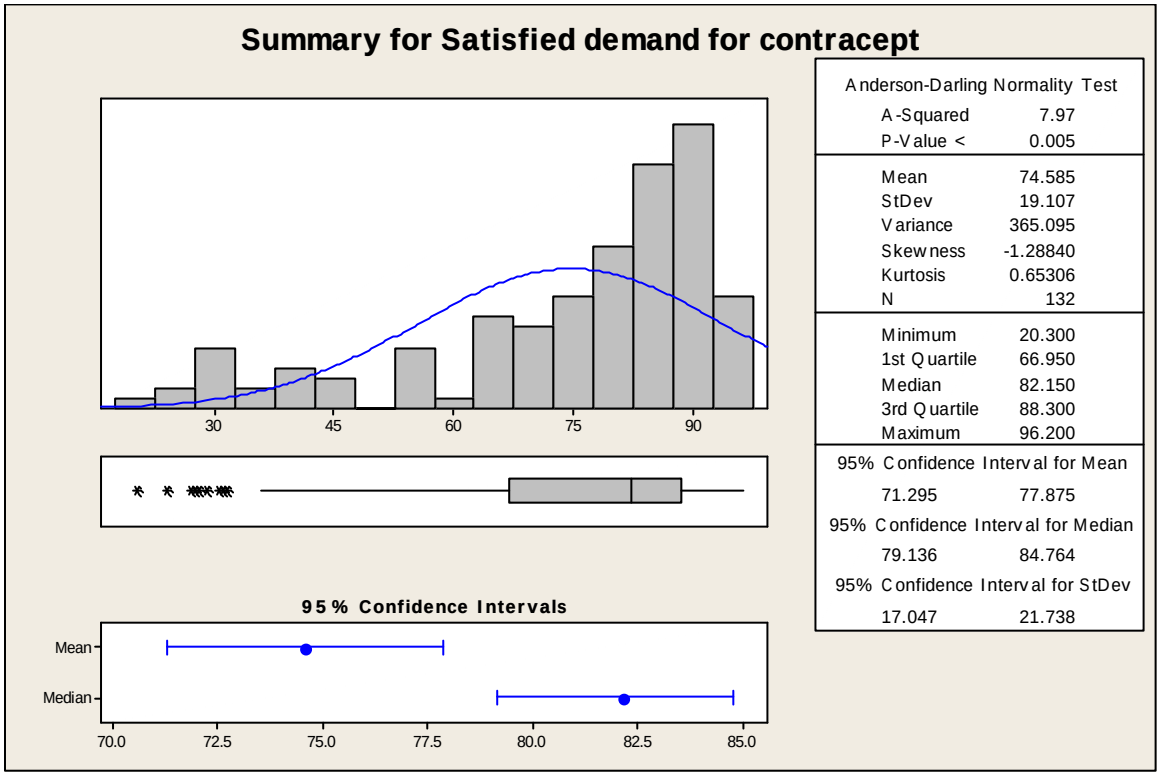


Figure 12: graph - Satisfied demand for contraception (% of women)

Análise da variável:

A variável “Demanda por contracepção atendida”, mede o percentual de demanda total de planeamento familiar entre mulheres casadas ou em união com idades entre 15 e 49 anos que estão satisfeitas com métodos contraceptivos disponíveis.

Podemos observar que 87 países figuram acima da média que tem valor de 74.585, isso indica que esta demanda está sendo bem atendida.

A mediana tem valor de 82.150 e o histograma destaca a presença de 9 países como *outliers* nas observações, são eles: Guinea (30.50), Togo (31.60), Angola (31.20), Sudão (29.00), Mauritânia (28.10), Senegal (27.70), Libéria (27.00), Mali (24.20) e Chade (20.30).

4.1.12 Women treated with respect – Mulheres tratadas com respeito e dignidade

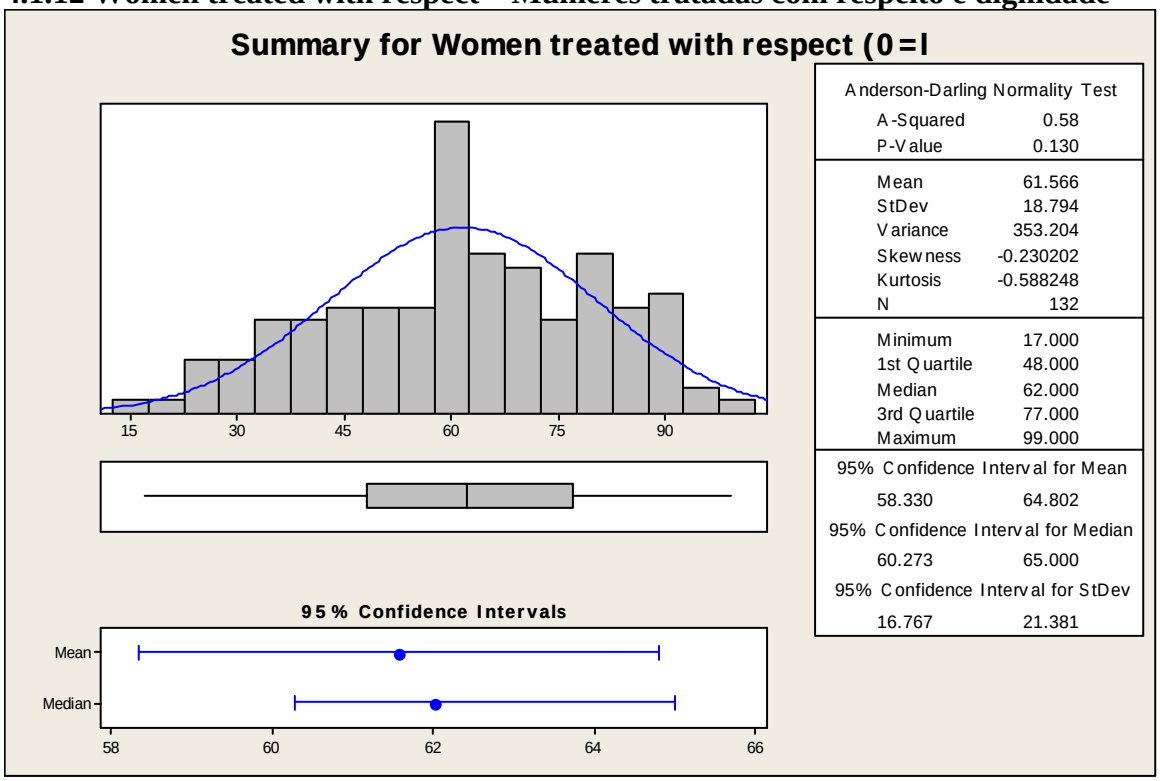


Figure 13: graph - Women treated with respect (0=low; 100=high)

Análise da variável:

Variável que contabiliza respostas SIM (positivas) de mulheres respondentes a respeito de sua percepção sobre serem tratadas com respeito e dignidade em seus países.

A curva de distribuição se aproxima de uma normal. A média tem valor de 61.566, a mediana é 62.000, onde se encontra a maior quantidade de observações agrupadas, totalizando 22 países.

4.1.13 Tolerance for immigrants – Tolerância a Imigrantes de outros países

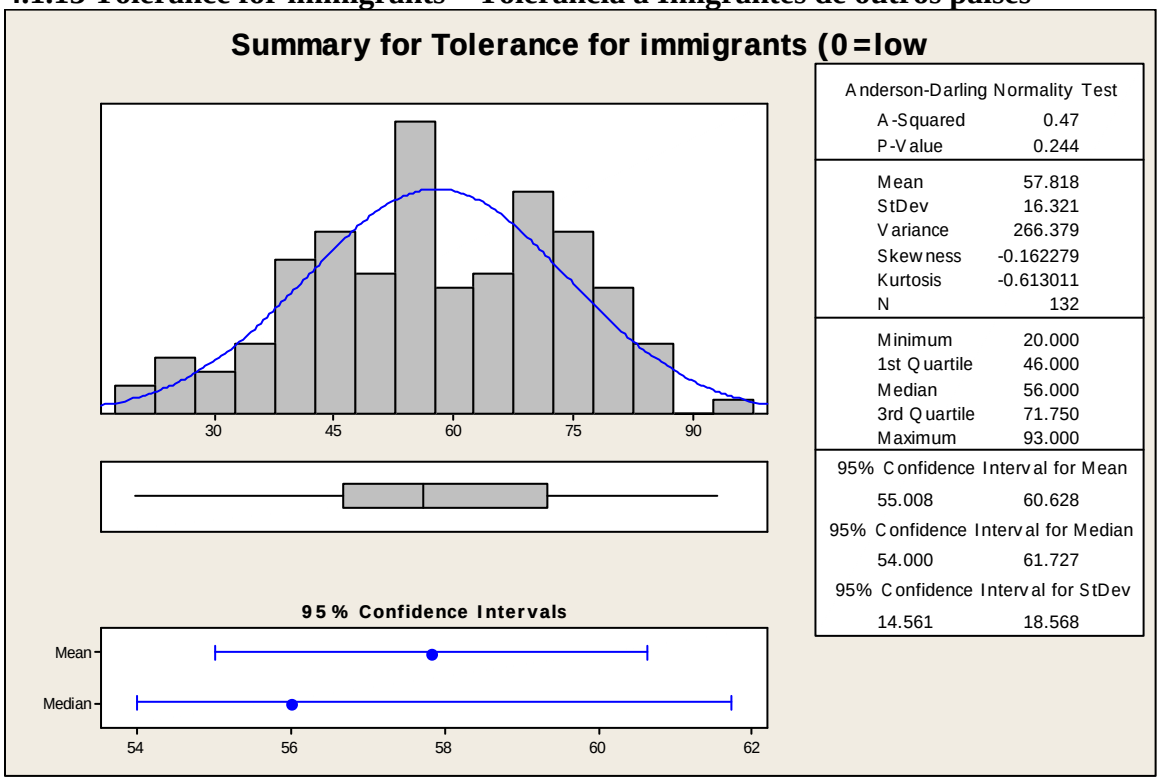


Figure 14: graph - Tolerance for immigrants (0=low; 100=high)

Análise da variável:

Tolerância a imigrantes: O percentual de entrevistados respondendo sim para a pergunta: "Será que a cidade ou área onde você vive é um bom lugar ou não é um bom lugar para viver para imigrantes de outros países?"

Variável que contabiliza respostas SIM (positivas) da população sobre a percepção de onde vivem ser um bom lugar para imigrantes de outros países.

A curva se aproxima de uma distribuição normal. A média apresenta valor de 57.818 e a mediana de 56.000, região esta do gráfico onde concentra o maior valor observado, composto por 21 países.

Nota-se a ausência de observações no estrato entre 87.5 e 92.5. E somente uma observação na barra de valor máximo do gráfico, ocupada pelo país Canadá (93.00). Em contrapartida, temos 2 observações no estrato de 17.5 a 22.5, denotando países menos tolerantes ao imigrante, são eles: Indonésia (22.00) e Malásia (20.00).

4.1.14 Tolerance for homosexuals - Tolerância a homossexuais

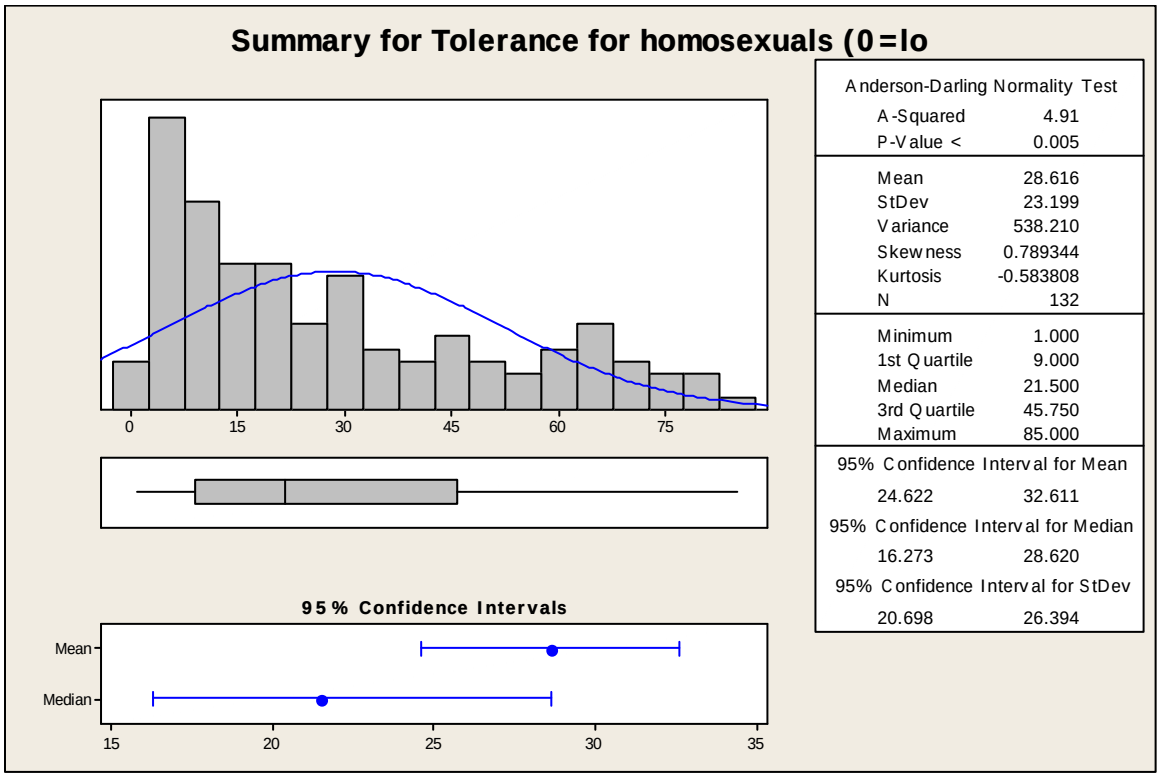


Figure 15: graph - Tolerance for homosexuals (0=low; 100=high)

Análise da variável:

Tolerância a homossexuais: O percentual de entrevistados que respondem sim à pergunta: "Será que a cidade ou área onde você vive é um bom lugar ou não é um bom lugar para se viver sendo gays ou lésbicas?"

Esta variável que contabiliza respostas SIM (positivas) da população sobre a percepção de onde vivem ser um bom lugar para a comunidade LGBT (gays, lésbicas,...).

A média tem valor de 28.616 e a mediana de 21.500, o que é mostrado no gráfico com a maior quantidade de observações figurarem no lado esquerdo do gráfico, cujo maior valor está no estrato entre 2.5 e 7.5, com 24 países que demonstram baixa tolerância. Do outro lado, no estrato de 82.5 e 87.5, com apenas um país, figura a Holanda (85.00) com o maior índice de tolerância encontrada nas observações.

4.1.15 Discrimination and violence against minorities - Discriminação e violência contra minorias

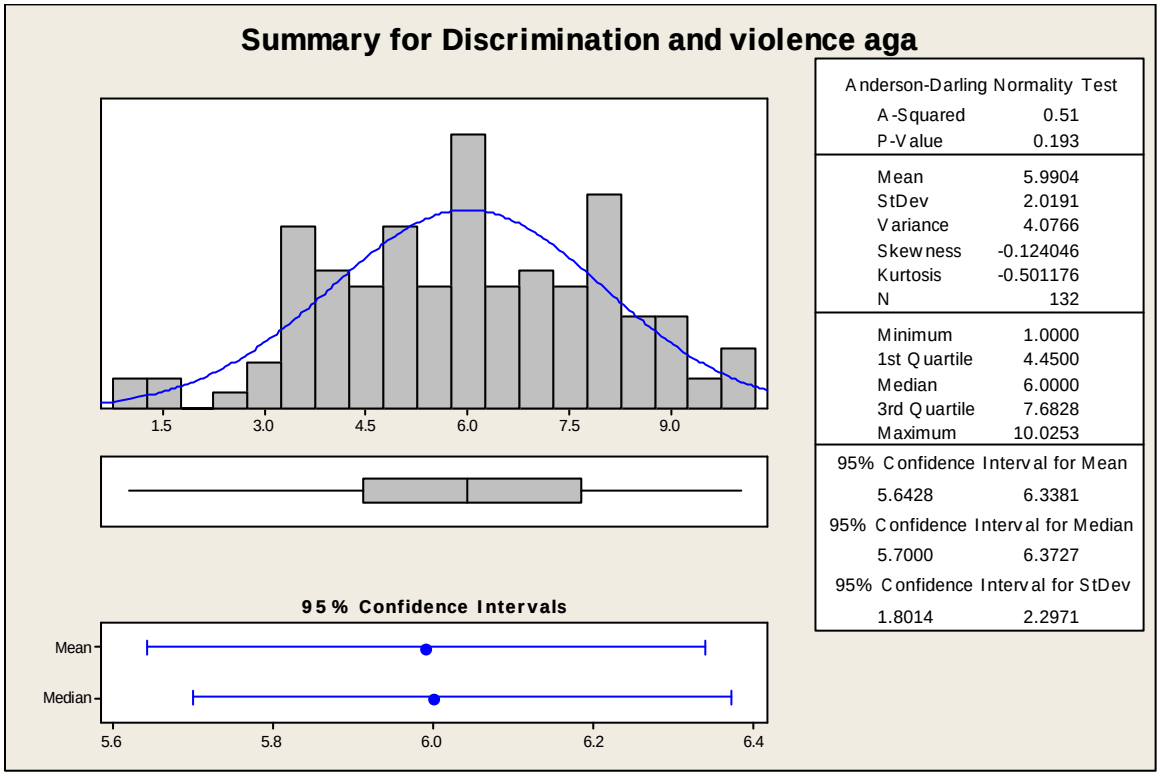


Figure 16: graph - Discrimination and violence against minorities (0=low; 10=high)

Análise da variável:

Discriminação e violência contra minorias: Indicador de Reclamações em Grupo. A discriminação, falta de poder, violência étnica, a violência comunitária, a violência sectária e violência religiosa, medido em uma escala de 0 (baixa pressão) a 10 (pressões muito altas). A variável mede o grau de intolerância, discriminação e violência contra as minorias e desfavorecidos (éticas, religiosas,...).

A curva se aproxima de uma distribuição normal. A média tem valor de 5.9904 e a mediana de 6.0000, onde se encontra o maior agrupamento de países, totalizando 18 observações. Não há dados no estrato entre 1.75 e 2.25, portanto há 4 países no estrato entre 0.75 e 1.75, são eles: Suécia (1.00), Irlanda (1.60), Islândia (1.00) e Finlândia (1.40), representando o grupo que apresenta o menor indicador de discriminação e violência contra minorias. E na outra extremidade na faixa entre 9.75 e 10.25, países como Iraque (10.03), Israel (9.75), Nigéria (9.80) e Sudão (9.96), apresentam os maiores índices.

4.1.16 Religious tolerance - Tolerância religiosa

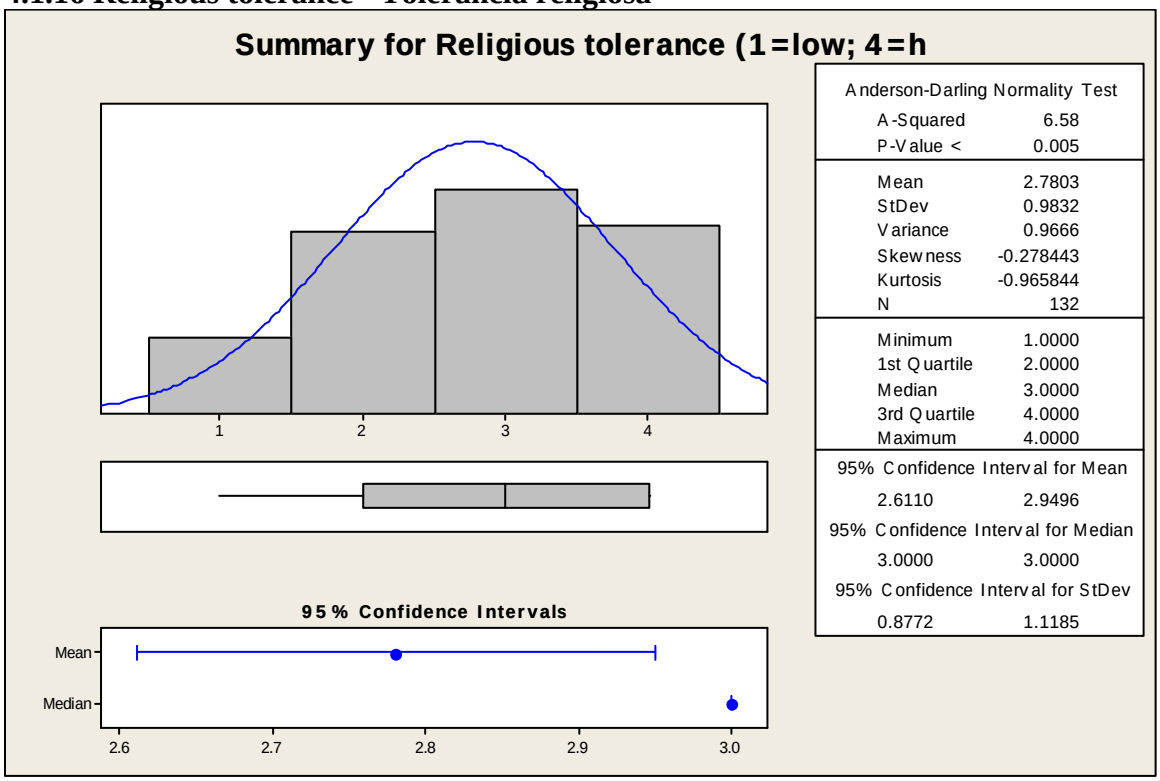


Figure 17: graph - Religious tolerance (1=low; 4=very high)

Análise da variável:

Tolerância religiosa: Uma medida de 13 tipos de hostilidade religiosa por indivíduos, organizações ou grupos da sociedade, incluindo conflitos relacionados com religião armados ou terrorismo, multidão ou a violência sectária, o assédio sobre traje ou por razões religiosas ou outras intimação ou abuso relacionadas com a religião. No modelo SPI, os pontos variam de 1 (baixo) a 4 (muito alto).

Esta variável é a medida de 13 tipos de hostilidades causadas por indivíduos, organizações ou sociedade (assédio, abuso, intimidação,..)

A curva de distribuição se aproxima de uma normal. A média tem valor de 2.7803 e a mediana de 3.0000. O terceiro quartil tem valor 4.0000, o mesmo do valor máximo.

Das observações, notamos que há 81 países nas faixas que representam tolerância “alta” e “muito alta”, ao passo que 15 países estão na faixa de “baixa” tolerância às expressões religiosas.

4.1.17 Community safety net - Apoio familiar e comunitário

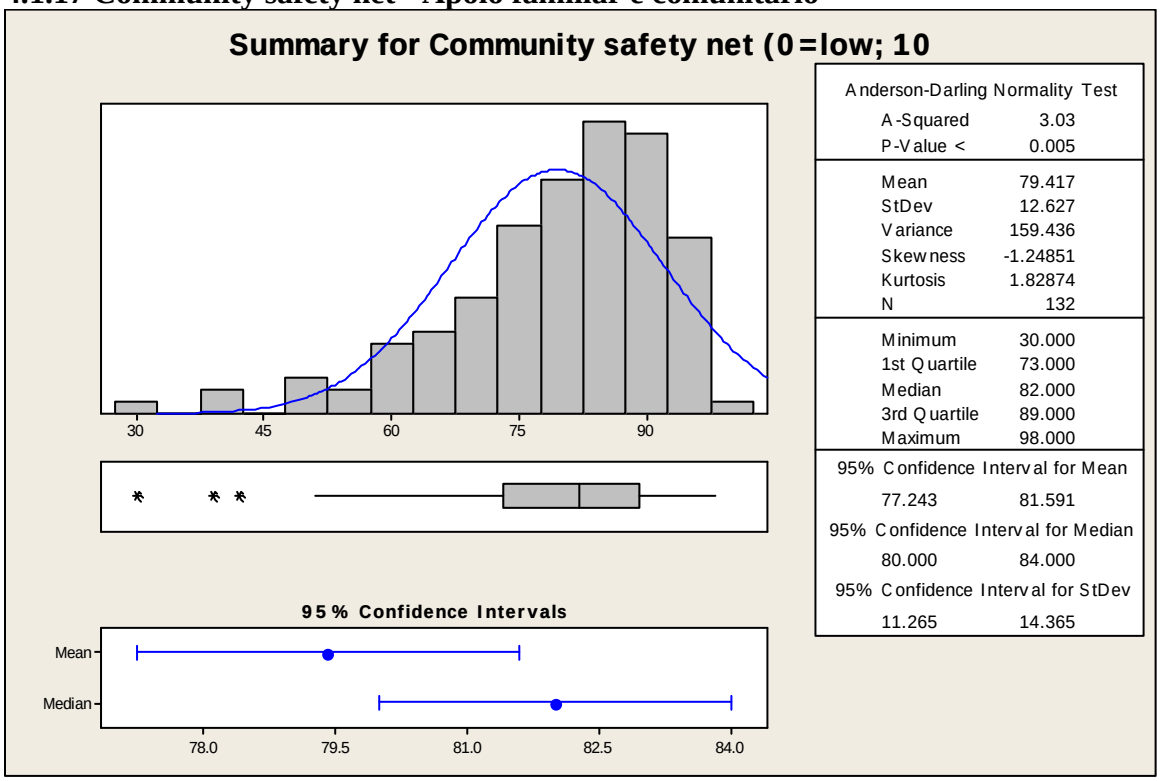


Figure 18: graph - Community safety net (0=low; 100=high)

Análise da variável:

Apoio familiar e comunitário: O percentual de entrevistados que respondem sim à pergunta: "se você estivesse em apuros, você tem parentes ou amigos que você pode contar para ajudá-lo sempre que você precisar deles, ou não?"

Variável que contabiliza respostas SIM (positivas) da população sobre uma situação problemática, onde possam contar com seus familiares ou amigos onde estiverem.

A curva aproxima-se de uma normal. A média tem valor de 79.417 e a mediana de 82.000. A maior concentração de observações está a direita no histograma. Dos 132 países analisados, 100 deles se encontram no estrato entre 72.5 e 97.5.

Não há observações na faixa entre 42.5 e 47.5, no entanto aparecem 3 *outliers* situados abaixo do valor 42.5, são eles: Togo (30.00), República da África Central (39.00) e Burundi (42.00), representando os países com menor índice de apoio familiar e comunitário.

4.1.18 Years of tertiary schooling - Anos de escolaridade superior

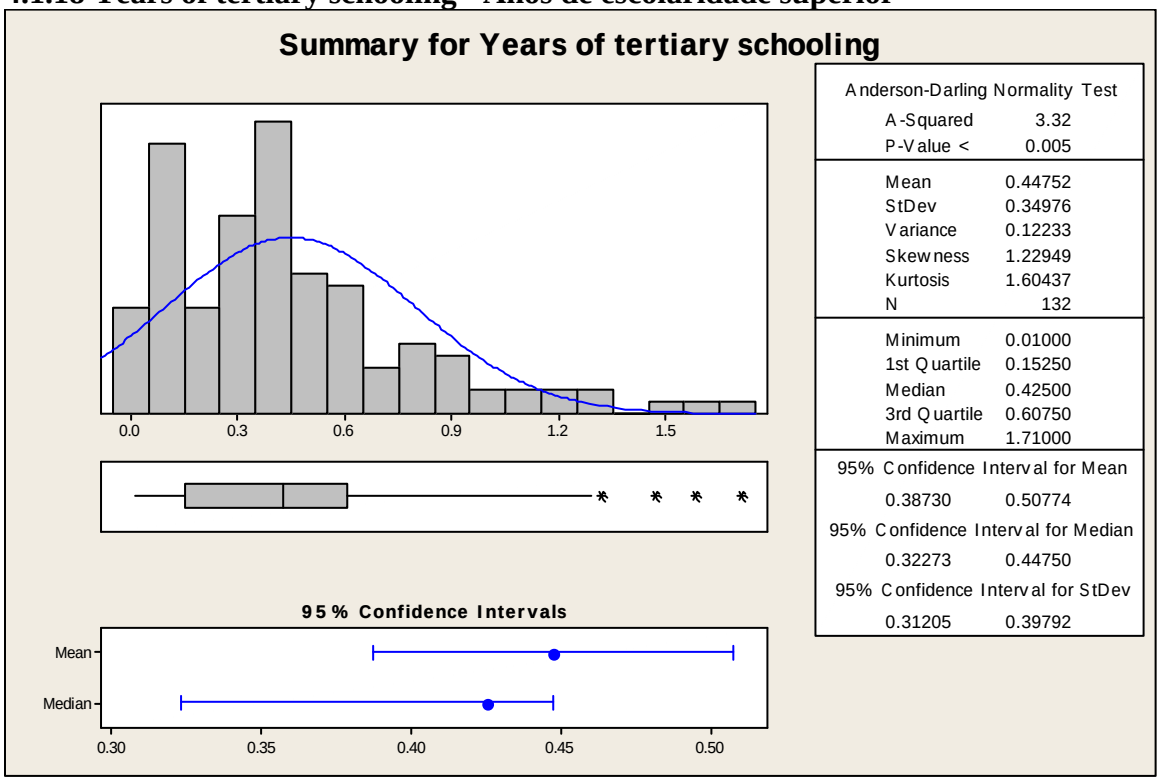


Figure 19: graph - Years of tertiary schooling

Análise da variável:

Anos de escolaridade superior

Esta variável retrata a média de anos de ensino superior concluído entre pessoas com mais de 25 anos de idade.

Nota-se a maior concentração de observações do lado esquerdo do gráfico. O que resulta em 106 países na faixa abaixo de 0.65. A média tem valor 0.44752, a mediana é 0.42500 e o valor máximo é de 1.71000. Há 4 *outliers* neste gráfico, representados pelos países Canadá (1.32), Nova Zelândia (1.47), Rússia (1.58) e USA (1.71) representanto o grupo de países com maior quantidade de anos de escolaridade superior.

4.1.19 Women's average years in school - Taxa média de escolarização das mulheres

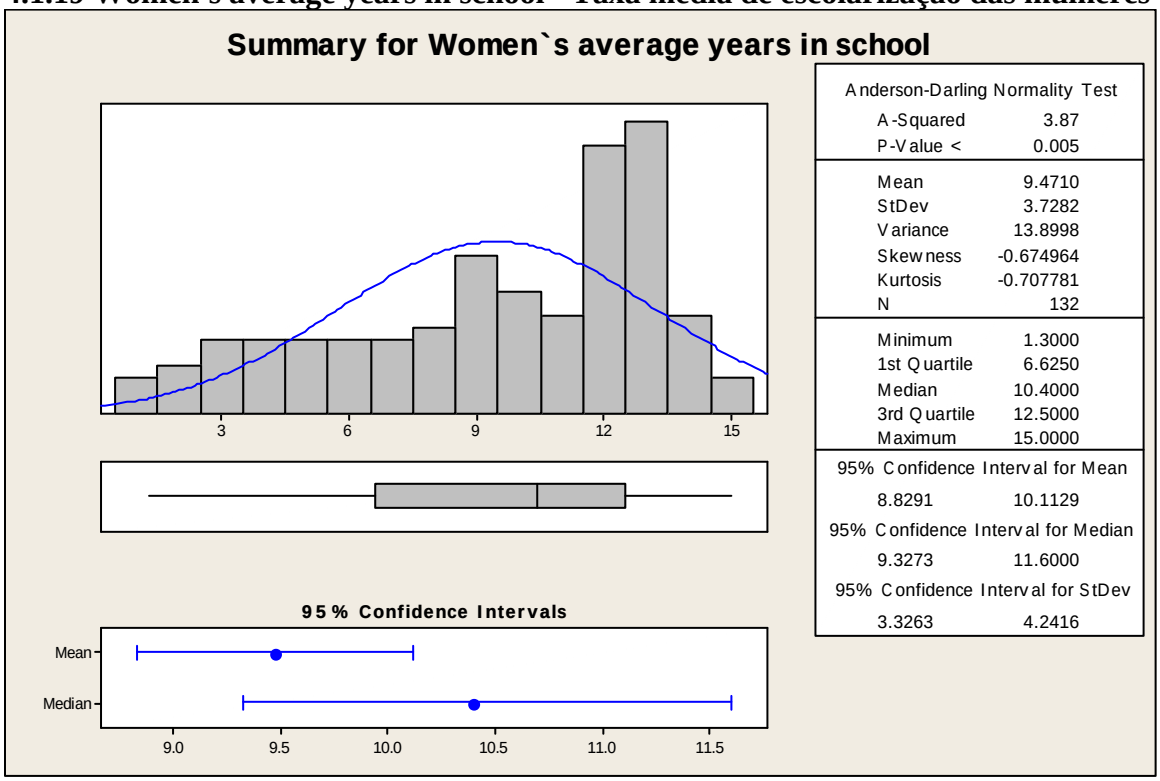


Figure 20: graph - Women's average years in school

Análise da variável:

Taxa média de escolarização das mulheres

O número médio de anos de escola frequentada por mulheres entre 25 e 34 anos de idade, incluindo fundamental, médio e ensino superior.

A curva se aproxima de uma normal. A média tem valor 9.4710 e a mediana de 10.4000

Dos 132 países estudados, vemos 48 países no estrato entre 11.5 e 13.5.

4.1.20 Inequality in the attainment of education - Desigualdade no desempenho educacional

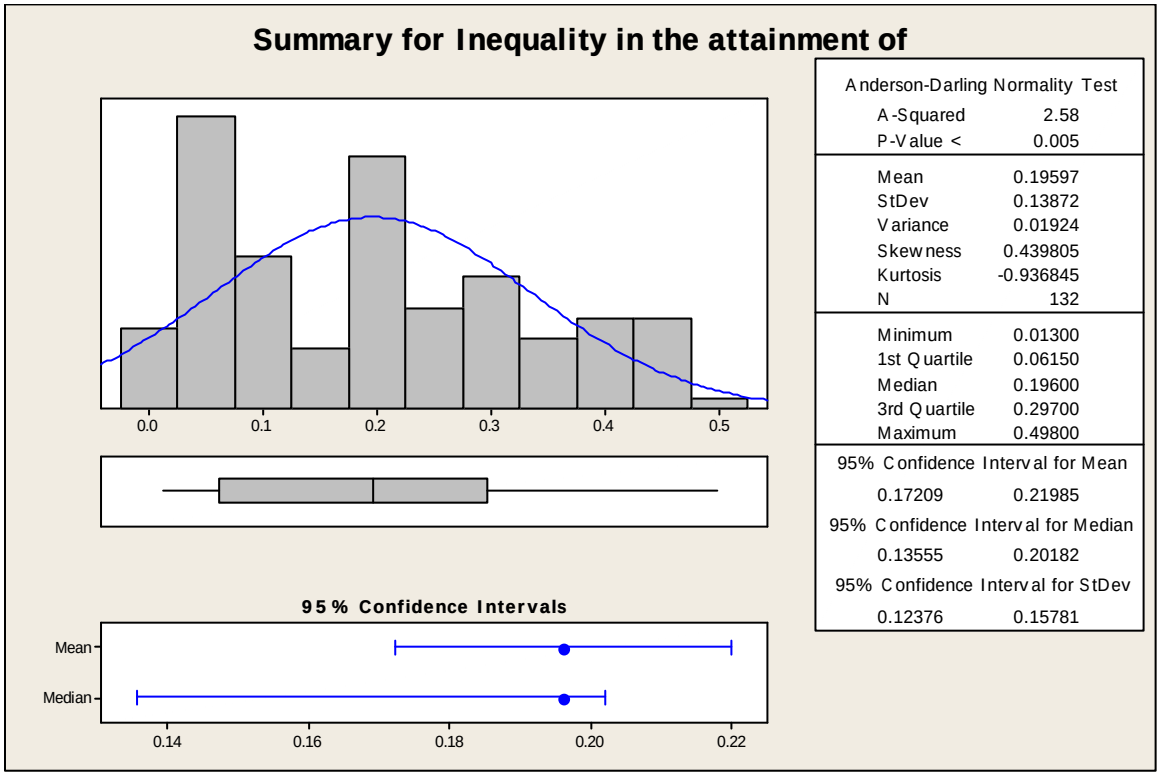


Figure 21: graph - Inequality in the attainment of education (0=low; 1=high)

Análise da variável:

Desigualdade no desempenho educacional

A perda na educação potencial devido à desigualdade, calculada como a diferença percentual entre o Índice de Educação do Índice de Desenvolvimento Humano, que compreende a média de anos de escolaridade e anos de escolaridade esperados, e o Índice de Educação Ajustado à Desigualdade.

A curva se aproxima de uma distribuição normal. A média é de 0.19597, a mediana tem valor de 0.19600, região do gráfico onde se encontra a segunda maior concentração de observações, com 25 países. Menor apenas que o estrato entre 0.025 e 0.075, que apresenta 29 países nesta faixa.

4.1.21 Índice de Desenvolvimento Humano - IDH – 2013

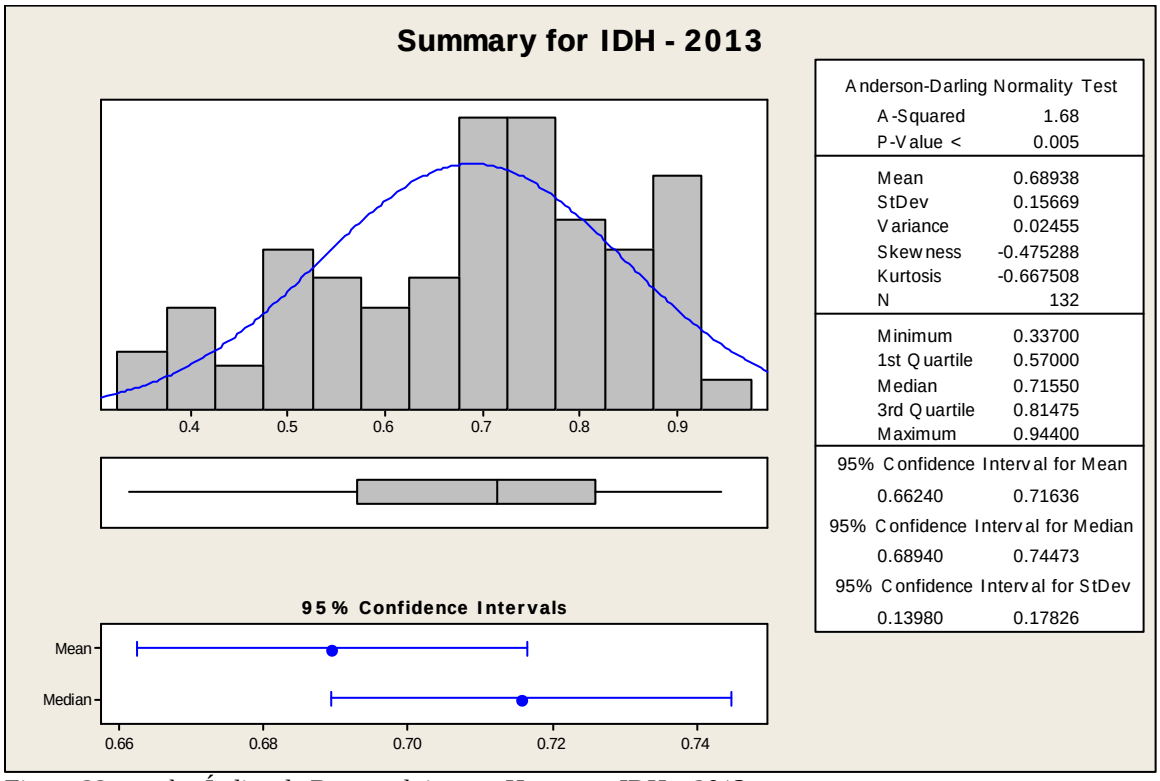


Figure 22: graph - Índice de Desenvolvimento Humano - IDH – 2013

Análise da variável:

IDH – 2013: O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.

O trabalho preliminar publicado por PNUD - Programa de las Naciones Unidas para el Desarrollo sobre o desenvolvimento humano em 2014 aponta que o IDH é progresso real da população em relação a expectativa de vida, educação, saúde, habitação, segurança e condições para um desenvolvimento sustentável.

Fonte: http://hdr.undp.org/sites/default/files/hdr14-summary-es.pdf

O histograma acima aproxima-se de uma distribuição normal. A média tem valor de 0.68938 e a mediana de 0.71550. A parte mais expressiva das observações se encontra mais a direita do gráfico e da média. Existem 40 países entre as faixas de 0.675 e 0.775.

Do lado esquerdo do gráfico, estão Niger, República do Congo, República da África Central (0.34) e Chade (0.37), considerados pelo índice como países de baixo desenvolvimento ou subdesenvolvidos.

4.1.22 Governança

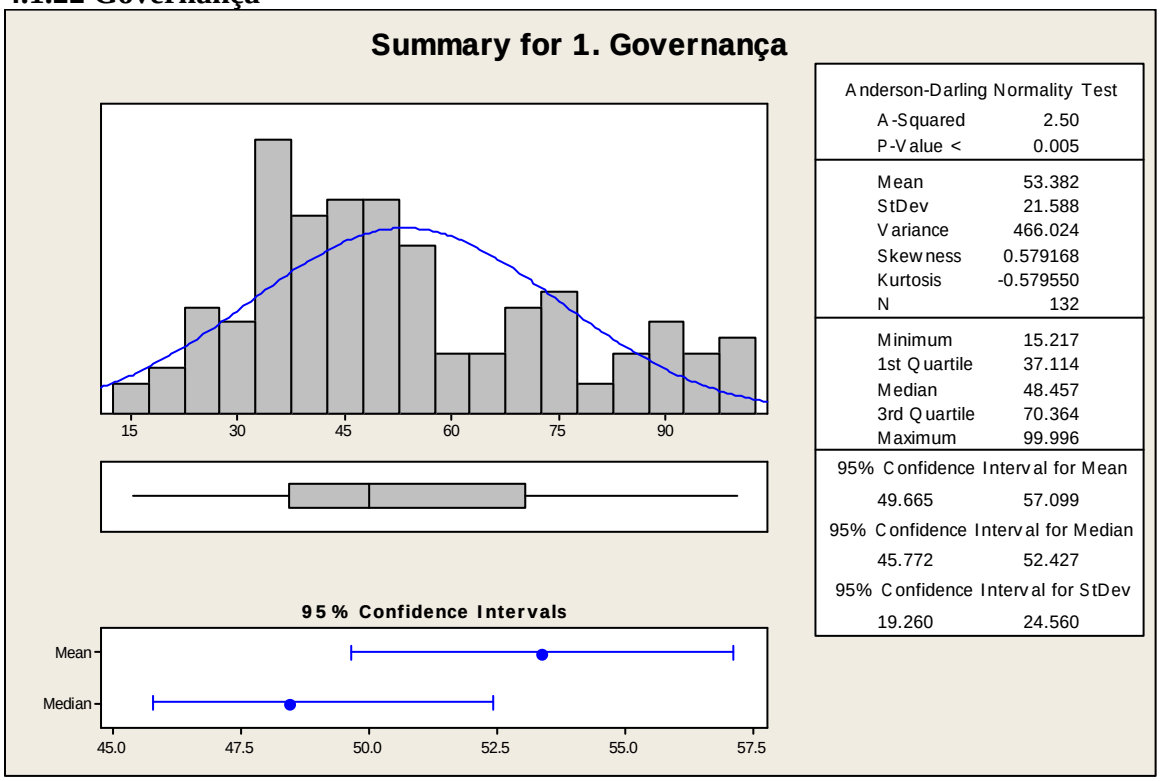


Figure 23: graph - Governança

Análise da variável:

Governança: O World Economic Fórum ressalta as estratégias de atuação: Simplificar os processos administrativos e fazer uso de alianças regionais e de tecnologia para coibir a corrupção e os abusos; Aprimorar os instrumentos e mecanismos regionais que possibilitem uma ação coletiva capaz de prevenir violações da democracia e garantir a preservação da ordem democrática (http://www3.weforum.org/docs/LA11/WEF_LA11_Report_PT.pdf).

Indicador que mede a capacidade de um país de eleger, monitorar e substituir seus governantes eleitos. Mede também nível de estabilidade política, medidas contra corrupção, entre outros.

A curva da distribuição se aproxima da normal. A média tem valor de 53.382, a mediana de 48.457 e o terceiro quartil tem valor de 70.364

A maior quantidade de observações para os 132 países analisados, está compreendida entre 32.5 e 37.5 com 18 países. Somente a Finlândia atingiu o índice 100.00 (máximo). No extremo inferior, 2 países figuram no estrato entre 12.5 e 17.5 , são eles: República da África Central (15.99) e Sudão (15.22).

4.1.23 EPI Score

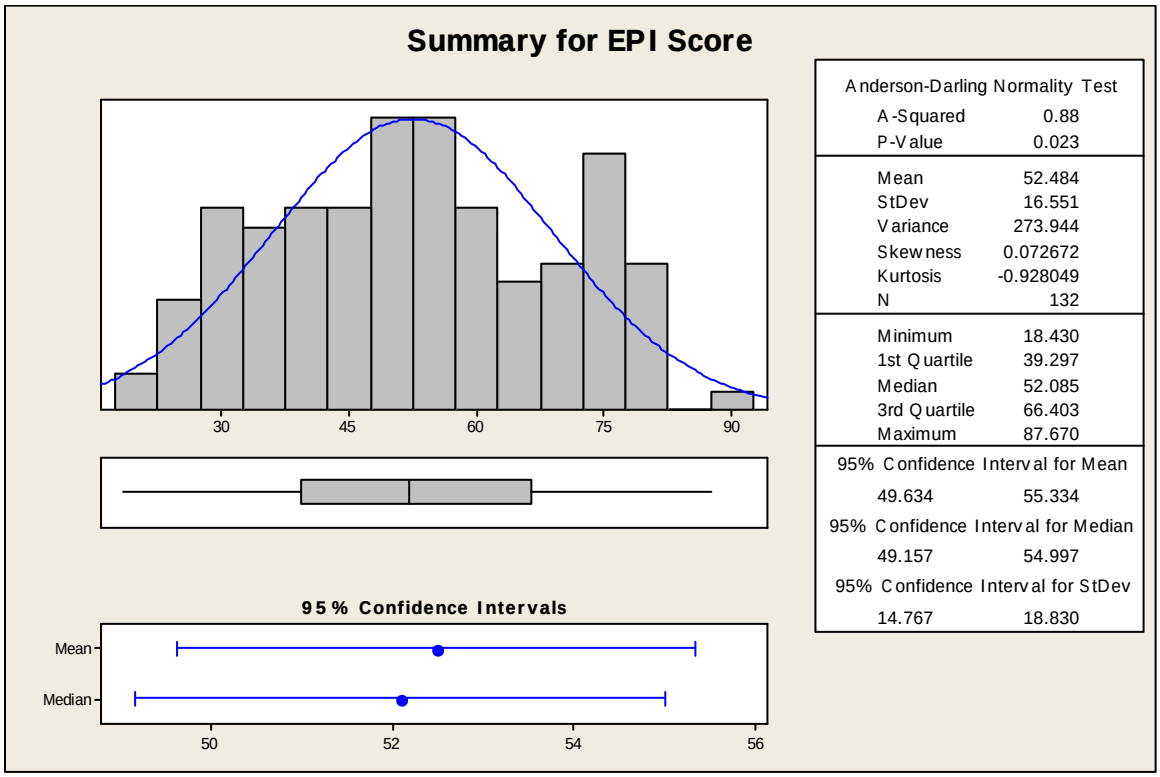


Figure 24: graph - EPI Score

Análise da variável:

EPI Score: Índice de Performance Ambiental, composto de: proteção da saúde do ser humano decorrente de ameaças ambientais e proteção do ecossistema.

A curva da distribuição se aproxima da normal, a média é de 52.484, a mediana 52.085, onde se concentram 32 países neste quartil. Os países se distribuem uniformemente no histograma. A linha se encontra praticamente no meio da caixa do box plot. Enquanto 1 país figura no limite máximo do gráfico (Suiça 87.67), 2 países se encontram no limite mínimo em relação aos 132 países analisados (Mali 18.43 e Lesotho 20.81)

4.1.24 EV - Water Resources

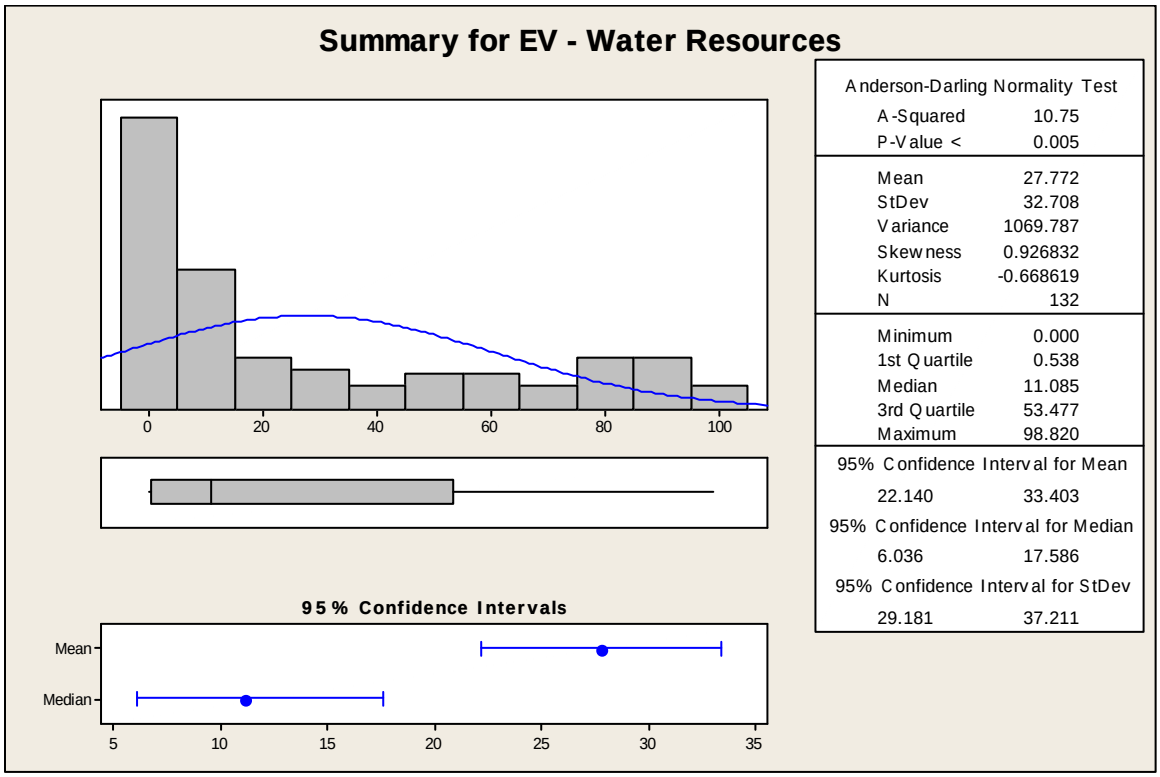


Figure 25: graph - EV - Water Resources

Análise da variável:

EV - Water Resources

Wastewater treatment level weighted by connection to wastewater treatment rate.

A curva da distribuição se aproxima da normal, a média tem valor de 27.772 e a mediana de 11.085, isso confirma a concentração das observações na parte esquerda do gráfico. Cerca de 90 observações encontram-se abaixo da média.

4.1.25 EV – Agriculture - subsídios destinados à Agricultura

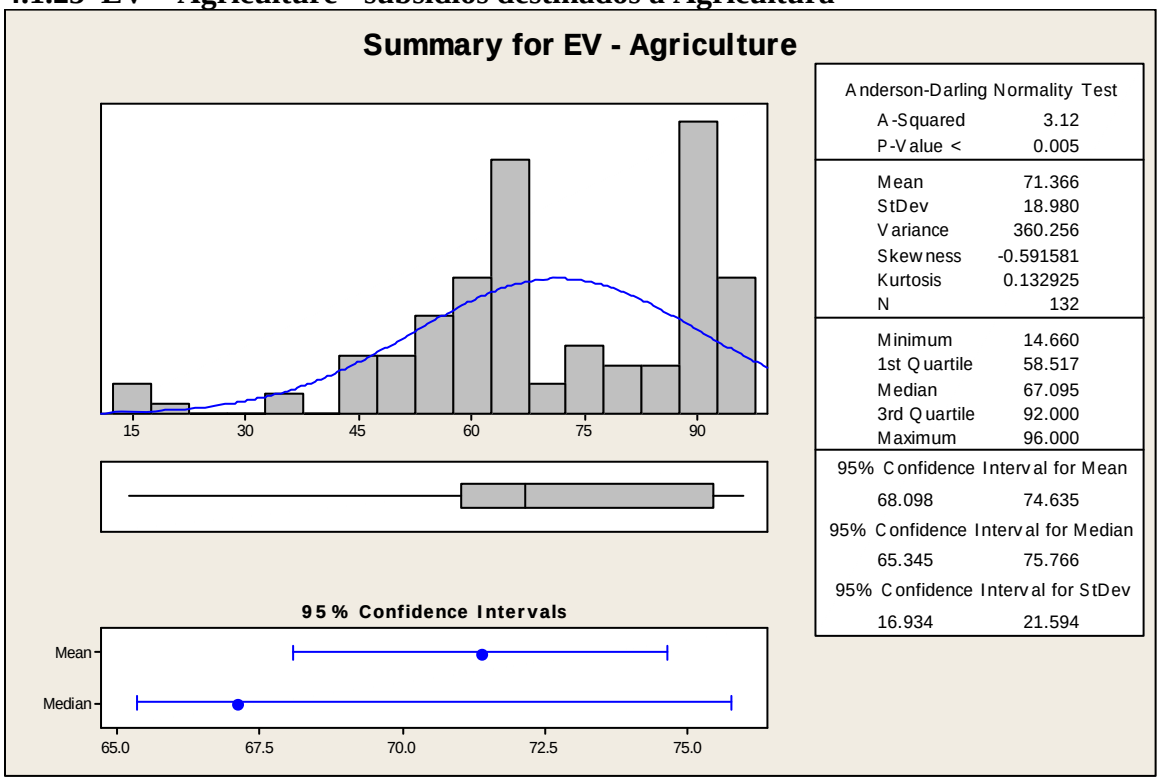


Figure 26: graph - EV – Agriculture

Análise da variável:

EV – Agriculture

Variável que mede os subsídios destinados à Agricultura e a Regulação no uso de pesticidas. A média tem valor 71.366, a mediana de 67.095 e nota-se os dados no histograma mais concentrados a direita do gráfico.

A maior concentração de países está situada no estrato entre 87.5 e 92.5.No canto esquerdo do gráfico com baixo resultado neste índice, encontram-se Israel (14.66), Rússia (16.93), Senegal (17.45) e Botswana (20.00).

4.1.26 EV - Forests

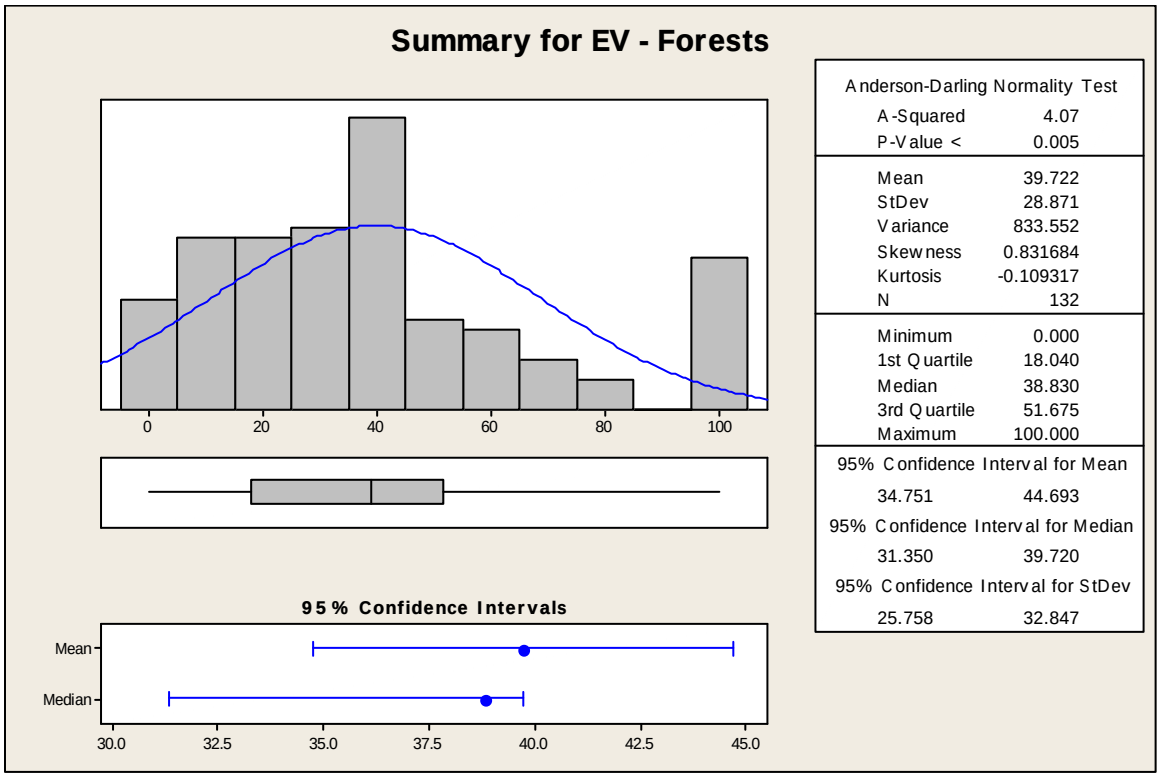


Figure 27: graph - EV - Forests

Análise da variável:

EV - Forests

Variável que demonstra a variação das áreas cobertas por floresta (desmatamento, queima e exploração).

A média tem valor de 39.722, a mediana de 38.830. Vê-se que 29 países estão localizados na faixa entre 35 e 45 pontos, a maior concentração deste gráfico. Dos 132 países analisados, somente 15 deles atingem o valor máximo no índice.

4.1.27 EV – Fisheries

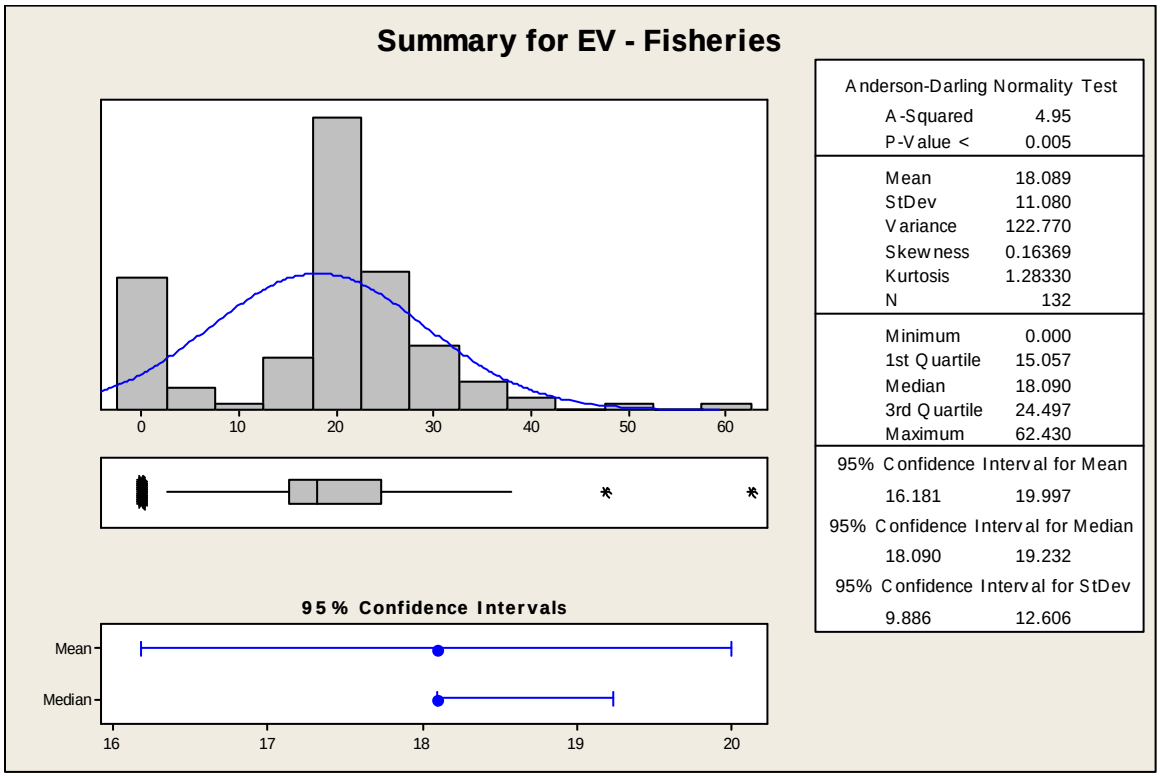


Figure 28: graph - EV – Fisheries

Análise da variável:

EV – Fisheries

Coastal Shelf Fishing Pressure and Fish Stocks

A média é 18.089, a mediana 18.090 e 51 observações no estrato entre 17.5 e 22.5.

Quanto aos *outliers*, não serão computados neste momento os casos com valor 0 (zero), pois esta revisão serão efetuadas nas próximas análises. Porém, na outra extremidade, temos 2 *outliers*, sendo Mauritius (47.50) e Sri Lanka (62.43).

4.1.28 Happy Planet Index - Bem Estar Sustentável

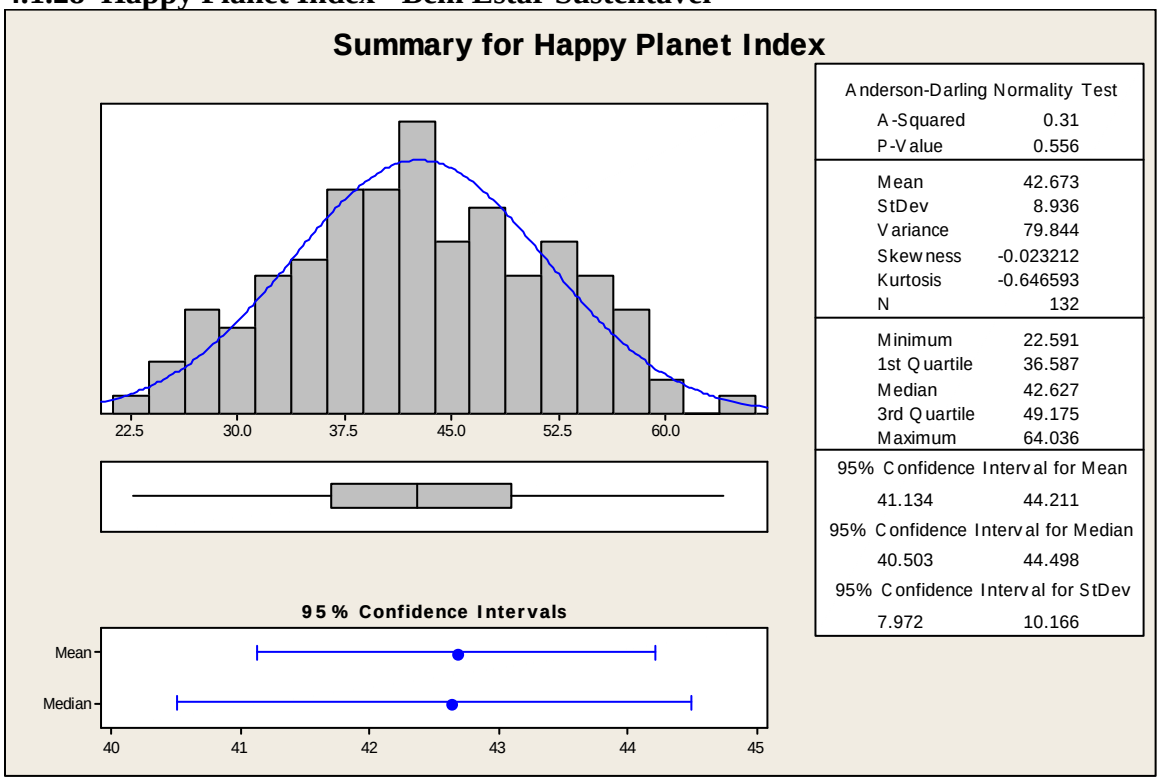


Figure 29: graph - Happy Planet Index

Análise da variável:

Happy Planet Index

HPI, medida do Bem Estar Sustentável (Longevidade, Felicidade e Vida sustentável para a população do país).

A curva da distribuição se aproxima de uma normal. A média tem valor 42.673 e a mediana com 42.627. OS dados se encontram distribuídos uniformemente seguindo a curva normal. No estrato entre 41.25 e 43.75 encontramos 17 países. O país com o maior índice de felicidade considerando os 132 indivíduos estudos, é a Costa Rica (64.04) e com menor índice de felicidade é a Botswana (22.59).

4.1.29 GINI Index

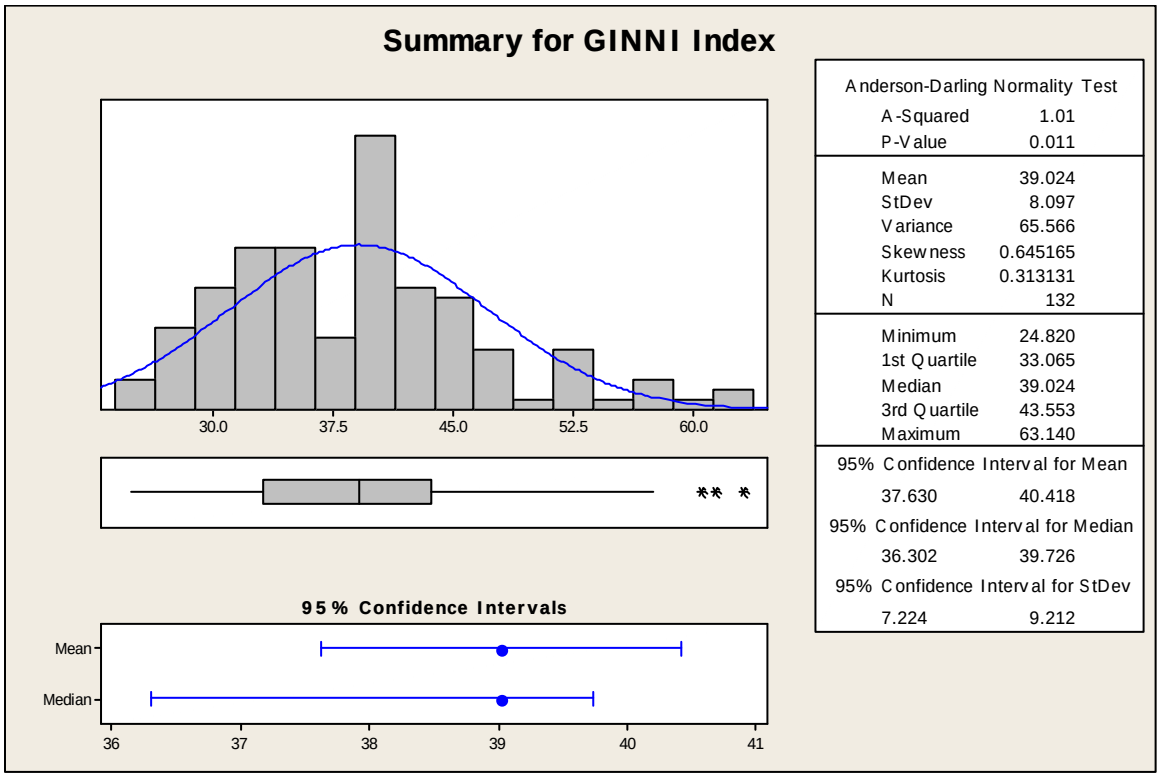


Figure 30: graph - GINI Index

Análise da variável:

GINI Index

Medida de desigualdade social utilizada para medir a desigualdade na distribuição de renda, ou desigualdade da riqueza.

A média tem valor de 39.024, a mediana de 39.024, e apresentam valores idênticos. A concentração dos dados no gráfico se encontra levemente a esquerda, o que denota uma situação aparentemente positiva.

Temos 3 *outliers*, são eles: Botswana (60.46), Namíbia (61.32) e África do Sul (63.14), e na outra extremidade os 3 países com resultados: Ucrânia (24.82), Suécia (26.08) e Eslovênia (24.87).

4.1.30 Democracy Index – 2014

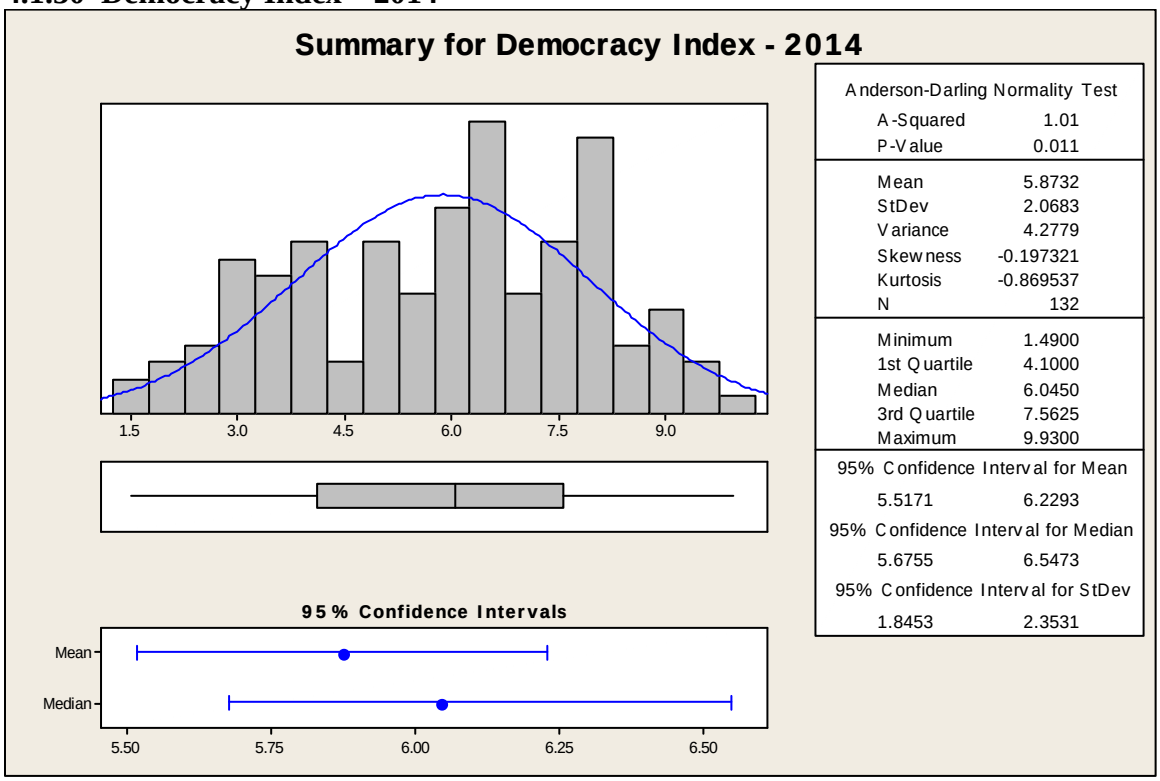


Figure 31: graph - Democracy Index – 2014

Análise da variável:
Democracy Index – 2014
A variável Índice de Demogracia é formada por 5 categorias: Processo eleitoral e pluralismo, Liberdade civil, Funcionamento do governo, Participação política e Cultura política.
A curva da distribuição se aproxima de uma normal. A média tem valor de 5.8732, a mediana de 6.0450. No estrato entre 6.25 e 6.75, encontramos 17 países. A Noruega (9.93) é o país que apresenta o maior índice de Democracia.

Considerações sobre a estatística descritiva das principais variáveis

Os índices analisados a partir de dados de pesquisa de 132 países que refletem o grau de evolução e engajamento dos países no quesito humano, sustentável e natural. Os resultados indicam os países que de uma forma geral estão alinhados com seus discursos e com os relatórios de organismos como a UNICEF, o Banco Mundial, WEF – World Econômica Forum, FAO e outros órgãos preocupados em identificar a evolução dos povos do globo.

Indicadores como o IDH – Índice de Desenvolvimento Humano apresentado nesse trabalho, por exemplo, ao ser se analisa dados do IDH de países da África Sub-saariana, República centro-africana e Lesotho, vemos resultados das variáveis que fazem parte deste indicador como saúde, educação, uso de água potável, instalações sanitária que indicam a qualidade de vida, estão baixos quando comparados com outros países.

2. RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS

2.1 Correlação linear

A matriz de correlação linear apresenta o teste de significância P-Value. Para a correlação foi utilizado o índice de Pearson. Conforme Las Casas e Guevara (2010)[1], a correlação (valores entre -1 e 1) é expressada pelo grau de relação linear que existe entre elas. Valores próximos a zero significam nenhuma relação, e próximos a 1 e -1 mostram alta relação positiva ou negativa, respectivamente. Vale ressaltar que o índice de correlação entre as variáveis não requer, necessariamente, que exista uma relação de causa-efeito entre as variáveis para dar sentido às relações. A seguir, apresentamos a matriz de correlação das variáveis estudadas no tema Biodiversidade - Capital Natural e Social:

Correlations: Social Progress, Basic Human Needs, Foundations Wellbeing, Opportu...

	Social Progress	Basic Human Needs	Foundations of W
Basic Human Needs	0.934 0.000		
Foundations of W	0.939 0.000	0.821 0.000	
Opportunity	0.922 0.000	0.748 0.000	0.847 0.000
Health and Welln	0.437 0.000	0.323 0.000	0.503 0.000
Personal Freedom	0.865 0.000	0.715 0.000	0.817 0.000
Greenhouse gas e	-0.342 0.000	-0.355 0.000	-0.319 0.000
Water withdrawal	0.140 0.109	0.288 0.001	0.001 0.994
Biodiversity and	0.353 0.000	0.184 0.034	0.566 0.000
Modern slavery,	-0.578 0.000	-0.564 0.000	-0.505 0.000
Satisfied demand	0.743 0.000	0.765 0.000	0.695 0.000
Women treated wi	0.162 0.064	0.201 0.021	0.109 0.215
Tolerance for im	0.243 0.005	0.066 0.454	0.205 0.018
Tolerance for ho	0.721	0.593	0.650

[1] LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisas de Marketing. São Paulo: Atlas, 2010.

	0.000	0.000	0.000
Discrimination a	-0.656	-0.496	-0.588
	0.000	0.000	0.000
Religious tolera	0.196	0.053	0.198
	0.024	0.547	0.023
Community safety	0.681	0.600	0.630
	0.000	0.000	0.000
Years of tertiary	0.625	0.601	0.501
	0.000	0.000	0.000
Women`s average	0.829	0.827	0.748
	0.000	0.000	0.000
Inequality in th	-0.772	-0.755	-0.682
	0.000	0.000	0.000
IDH - 2013	0.868	0.901	0.775
	0.000	0.000	0.000
1. Governança	0.890	0.757	0.827
	0.000	0.000	0.000
EPI Score	0.917	0.888	0.879
	0.000	0.000	0.000
EV - Water Resources	0.765	0.725	0.675
	0.000	0.000	0.000
EV - Agriculture	-0.215	-0.217	-0.173
	0.013	0.012	0.048
EV - Forests	0.124	0.257	-0.015
	0.155	0.003	0.867
EV - Fisheries	-0.027	-0.028	0.002
	0.758	0.746	0.978
Happy Planet Ind	0.296	0.331	0.293
	0.001	0.000	0.001
GINNI Index	-0.308	-0.415	-0.217
	0.000	0.000	0.013
Democracy Index	0.789	0.604	0.768
	0.000	0.000	0.000

	Opportunity	Health and Welln	Personal Freedom
Health and Welln	0.438		
	0.000		
Personal Freedom	0.907	0.504	
	0.000	0.000	
Greenhouse gas e	-0.273	-0.202	-0.235
	0.002	0.020	0.007
Water withdrawal	0.036	0.039	0.079
	0.685	0.658	0.368
Biodiversity and	0.330	0.308	0.348
	0.000	0.000	0.000
Modern slavery,	-0.531	-0.264	-0.682
	0.000	0.002	0.000

Satisfied demand	0.602 0.000	0.218 0.012	0.688 0.000
Women treated wi	0.121 0.168	0.175 0.045	0.155 0.076
Tolerance for im	0.435 0.000	0.333 0.000	0.391 0.000
Tolerance for ho	0.781 0.000	0.518 0.000	0.746 0.000
Discrimination a	-0.767 0.000	-0.264 0.002	-0.674 0.000
Religious tolera	0.328 0.000	0.117 0.183	0.285 0.001
Community safety	0.680 0.000	0.245 0.005	0.605 0.000
Years of tertiary	0.618 0.000	0.165 0.058	0.516 0.000
Women`s average	0.723 0.000	0.084 0.340	0.652 0.000
Inequality in th	-0.701 0.000	-0.047 0.593	-0.594 0.000
IDH - 2013	0.719 0.000	0.276 0.001	0.688 0.000
1. Governança	0.916 0.000	0.436 0.000	0.844 0.000
EPI Score	0.796 0.000	0.361 0.000	0.755 0.000
EV - Water Resou	0.724 0.000	0.332 0.000	0.655 0.000
EV - Agriculture	-0.200 0.021	0.044 0.615	-0.152 0.083
EV - Forests	0.042 0.634	-0.223 0.010	0.006 0.942
EV - Fisheries	-0.042 0.635	0.185 0.034	0.014 0.869
Happy Planet Ind	0.196 0.024	0.390 0.000	0.283 0.001
GINNI Index	-0.185 0.034	-0.008 0.924	-0.086 0.325
Democracy Index	0.869 0.000	0.450 0.000	0.758 0.000

	Greenhouse gas e	Water withdrawal	Biodiversity and
Water withdrawal	-0.173 0.048		
Biodiversity and	0.049 0.574	-0.201 0.021	
Modern slavery,	0.171 0.050	-0.275 0.001	-0.141 0.107

Satisfied demand	-0.247	0.314	0.172
	0.004	0.000	0.048
Women treated wi	-0.043	0.102	0.048
	0.623	0.246	0.584
Tolerance for im	0.074	-0.203	0.173
	0.398	0.020	0.047
Tolerance for ho	-0.195	0.067	0.272
	0.025	0.444	0.002
Discrimination a	0.190	0.064	-0.299
	0.029	0.469	0.000
Religious tolera	-0.052	-0.184	0.208
	0.551	0.035	0.017
Community safety	-0.392	0.129	0.243
	0.000	0.139	0.005
Years of tertiar	-0.186	0.211	0.104
	0.033	0.015	0.233
Women`s average	-0.249	0.278	0.159
	0.004	0.001	0.068
Inequality in th	0.201	-0.142	-0.173
	0.021	0.105	0.048
IDH - 2013	-0.332	0.332	0.183
	0.000	0.000	0.036
1. Governança	-0.283	0.012	0.355
	0.001	0.892	0.000
EPI Score	-0.204	0.155	0.455
	0.019	0.076	0.000
EV - Water Resou	-0.203	0.176	0.216
	0.019	0.043	0.013
EV - Agriculture	0.131	-0.012	-0.063
	0.133	0.893	0.471
EV - Forests	0.004	0.184	-0.309
	0.967	0.035	0.000
EV - Fisheries	-0.053	0.025	0.040
	0.544	0.777	0.647
Happy Planet Ind	-0.278	0.242	-0.024
	0.001	0.005	0.781
GINNI Index	0.190	-0.037	0.120
	0.029	0.672	0.172
Democracy Index	-0.307	-0.059	0.300
	0.000	0.500	0.000

	Modern slavery,	Satisfied demand	Women treated wi
Satisfied demand	-0.580		
	0.000		
Women treated wi	0.037	-0.063	
	0.672	0.472	
Tolerance for im	-0.080	-0.099	0.091

	0.360	0.256	0.301
Tolerance for ho	-0.431	0.487	0.091
	0.000	0.000	0.300
Discrimination a	0.325	-0.313	-0.137
	0.000	0.000	0.117
Religious tolera	0.003	0.045	-0.129
	0.970	0.609	0.140
Community safety	-0.409	0.501	0.054
	0.000	0.000	0.542
Years of tertiary	-0.436	0.441	0.115
	0.000	0.000	0.190
Women`s average	-0.554	0.773	0.067
	0.000	0.000	0.448
Inequality in th	0.483	-0.634	-0.141
	0.000	0.000	0.108
IDH - 2013	-0.580	0.765	0.111
	0.000	0.000	0.207
1. Governança	-0.437	0.514	0.283
	0.000	0.000	0.001
EPI Score	-0.553	0.683	0.205
	0.000	0.000	0.018
EV - Water Resources	-0.456	0.498	0.326
	0.000	0.000	0.000
EV - Agriculture	-0.028	-0.153	-0.175
	0.751	0.080	0.045
EV - Forests	-0.028	0.146	0.125
	0.754	0.095	0.152
EV - Fisheries	-0.017	0.035	-0.206
	0.849	0.689	0.018
Happy Planet Ind	-0.291	0.502	-0.242
	0.001	0.000	0.005
GINNI Index	0.141	-0.159	-0.323
	0.106	0.069	0.000
Democracy Index	-0.372	0.511	-0.007
	0.000	0.000	0.938

	Tolerance for im	Tolerance for ho	Discrimination a
Tolerance for ho	0.500		
	0.000		
Discrimination a	-0.428	-0.559	
	0.000	0.000	
Religious tolera	0.324	0.210	-0.506
	0.000	0.015	0.000
Community safety	0.212	0.606	-0.454
	0.015	0.000	0.000
Years of tertiary	0.151	0.482	-0.330
	0.084	0.000	0.000

Women`s average	-0.018	0.469	-0.464
	0.837	0.000	0.000
Inequality in th	-0.065	-0.443	0.483
	0.460	0.000	0.000
IDH - 2013	0.035	0.590	-0.470
	0.693	0.000	0.000
1. Governança	0.330	0.678	-0.737
	0.000	0.000	0.000
EPI Score	0.202	0.644	-0.560
	0.020	0.000	0.000
EV - Water Resou	0.201	0.601	-0.542
	0.021	0.000	0.000
EV - Agriculture	0.153	-0.072	0.113
	0.080	0.414	0.198
EV - Forests	0.047	-0.013	-0.006
	0.591	0.886	0.943
EV - Fisheries	-0.018	0.040	0.057
	0.839	0.650	0.516
Happy Planet Ind	-0.114	0.323	0.031
	0.195	0.000	0.720
GINNI Index	0.120	-0.061	0.139
	0.169	0.488	0.111
Democracy Index	0.264	0.637	-0.629
	0.002	0.000	0.000

	Religious tolera	Community safety	Years of tertiary
Community safety	0.135		
	0.124		
Years of tertiary	0.002	0.442	
	0.984	0.000	
Women`s average	0.085	0.632	0.621
	0.332	0.000	0.000
Inequality in th	-0.122	-0.595	-0.484
	0.165	0.000	0.000
IDH - 2013	0.013	0.626	0.608
	0.884	0.000	0.000
1. Governança	0.208	0.574	0.553
	0.017	0.000	0.000
EPI Score	0.096	0.611	0.605
	0.273	0.000	0.000
EV - Water Resou	0.020	0.527	0.603
	0.823	0.000	0.000
EV - Agriculture	0.066	-0.215	-0.275
	0.453	0.013	0.001
EV - Forests	-0.114	-0.023	0.098
	0.192	0.791	0.263
EV - Fisheries	0.008	0.170	-0.077
	0.928	0.051	0.378

Happy Planet Ind	-0.053	0.210	0.164
	0.543	0.016	0.060
GINNI Index	0.206	-0.135	-0.309
	0.018	0.122	0.000
Democracy Index	0.192	0.546	0.438
	0.028	0.000	0.000

	Women`s average	Inequality in th	IDH - 2013
Inequality in th	-0.852		
	0.000		
IDH - 2013	0.844	-0.736	
	0.000	0.000	
1. Governança	0.650	-0.637	0.697
	0.000	0.000	0.000
EPI Score	0.772	-0.722	0.824
	0.000	0.000	0.000
EV - Water Resou	0.623	-0.560	0.690
	0.000	0.000	0.000
EV - Agriculture	-0.188	0.186	-0.207
	0.030	0.033	0.017
EV - Forests	0.210	-0.203	0.201
	0.016	0.020	0.021
EV - Fisheries	-0.057	0.118	-0.024
	0.518	0.177	0.785
Happy Planet Ind	0.274	-0.135	0.333
	0.001	0.124	0.000
GINNI Index	-0.310	0.406	-0.297
	0.000	0.000	0.001
Democracy Index	0.559	-0.542	0.577
	0.000	0.000	0.000

	1. Governança	EPI Score	EV - Water Resou
EPI Score	0.811		
	0.000		
EV - Water Resources	0.783	0.840	
	0.000	0.000	
EV - Agriculture	-0.305	-0.188	-0.297
	0.000	0.031	0.001
EV - Forests	0.096	0.200	0.127
	0.272	0.022	0.147
EV - Fisheries	-0.073	-0.017	-0.111
	0.404	0.849	0.206
Happy Planet Ind	0.089	0.197	0.086
	0.312	0.024	0.325
GINNI Index	-0.247	-0.365	-0.421
	0.004	0.000	0.000
Democracy Index	0.849	0.633	0.577
	0.000	0.000	0.000

	EV - Agriculture	EV - Forests	EV - Fisheries
EV - Forests	-0.014 0.873		
EV - Fisheries	0.064 0.467	-0.117 0.181	
Happy Planet Ind	0.057 0.518	-0.035 0.692	0.192 0.028
GINNI Index	0.072 0.413	-0.175 0.045	0.137 0.116
Democracy Index	-0.249 0.004	-0.022 0.803	0.026 0.770

	Happy Planet Ind	GINNI Index
GINNI Index	0.051 0.562	
Democracy Index	0.231 0.008	-0.118 0.176

Cell Contents: Pearson correlation
P-Value

2.2 Gráficos de dispersão (*Scatterplot*)

Para efeito de análise prévia, foram construídos gráficos de dispersão evidenciando a relação entre as variáveis com maior grau de correlação. Como critério de escolha para elaboração dos gráficos, foi assumido os casos cujos resultados de correlação foram maior que 0.8, maior que -0.8 (negativo) e p-Value menor que 0.05.

Pelo tema central deste trabalho, buscamos destacar as correlações entre a variável *EPI Score* e outras de relevância. O índice *Environmental Performance Index (EPI)*, foi criado pela ONU – Organização das Nações Unidas, para atender o programa de Objetivos de Desenvolvimento do Milênio, em parceria com os centros de pesquisas das universidades de Yale e Columbia. O EPI pode ser considerado um método para quantificar a efetividade das políticas e ações voltadas às questões do Meio-Ambiente e seus impactos na saúde do ser humano, biodiversidade, sociedade e economia.

2.2.1 Variável de interesse *EPI Score* – Índice de Performance Ambiental

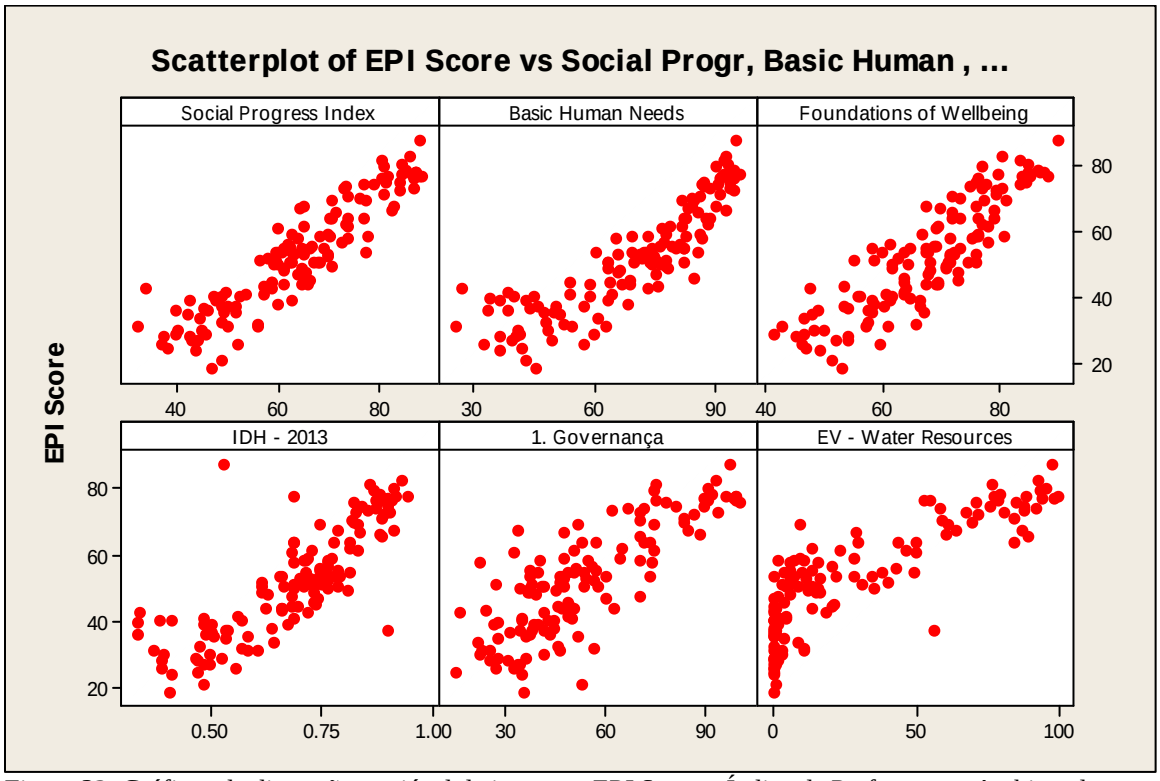


Figura 32: Gráficos de dispersão, variável de interesse *EPI Score* – Índice de Performance Ambiental

No gráfico acima (figura 32), percebe-se correlação linear positiva entre a variável EPI e as demais variáveis do estudo. O maior destaque foi com a variável Índice de Progresso Social (Social Progress Index), que apresentou correlação de 0.917 e p-Value de 0.000.

2.2.2 Variável de interesse *Social Progress Index* – Índice de Progresso Social

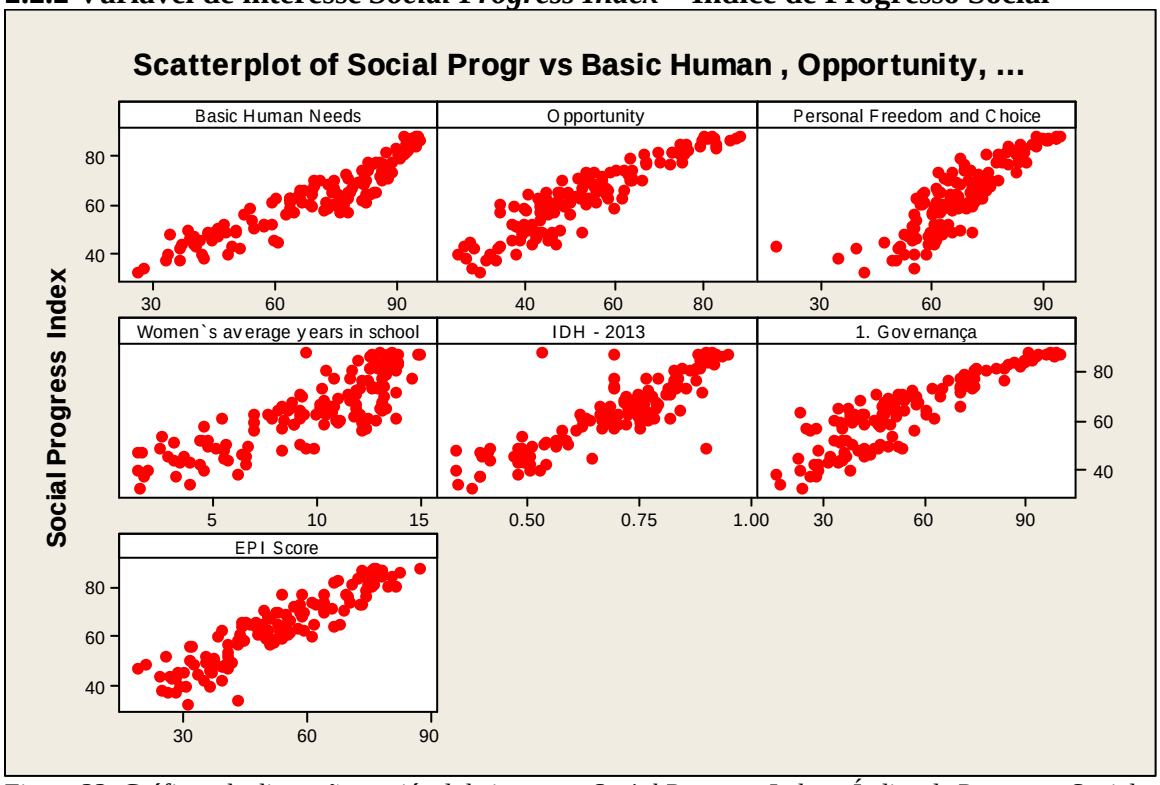


Figura 33: Gráficos de dispersão, variável de interesse *Social Progress Index* – Índice de Progresso Social

O gráfico baseado na variável de Índice de Progresso Social (*Social Progress Index*), mostra uma correlação positiva entre as variáveis. Com destaques para Necessidades Humanas Básicas (*Basic Human Needs*) com correlação de 0.934 e a variável Oportunidade (*Opportunity*) que apresentou correlação de 0.922 com a variável de referência.

5.2.3 Variável de interesse *Basic Human Needs – Necessidades Humanas Básicas*

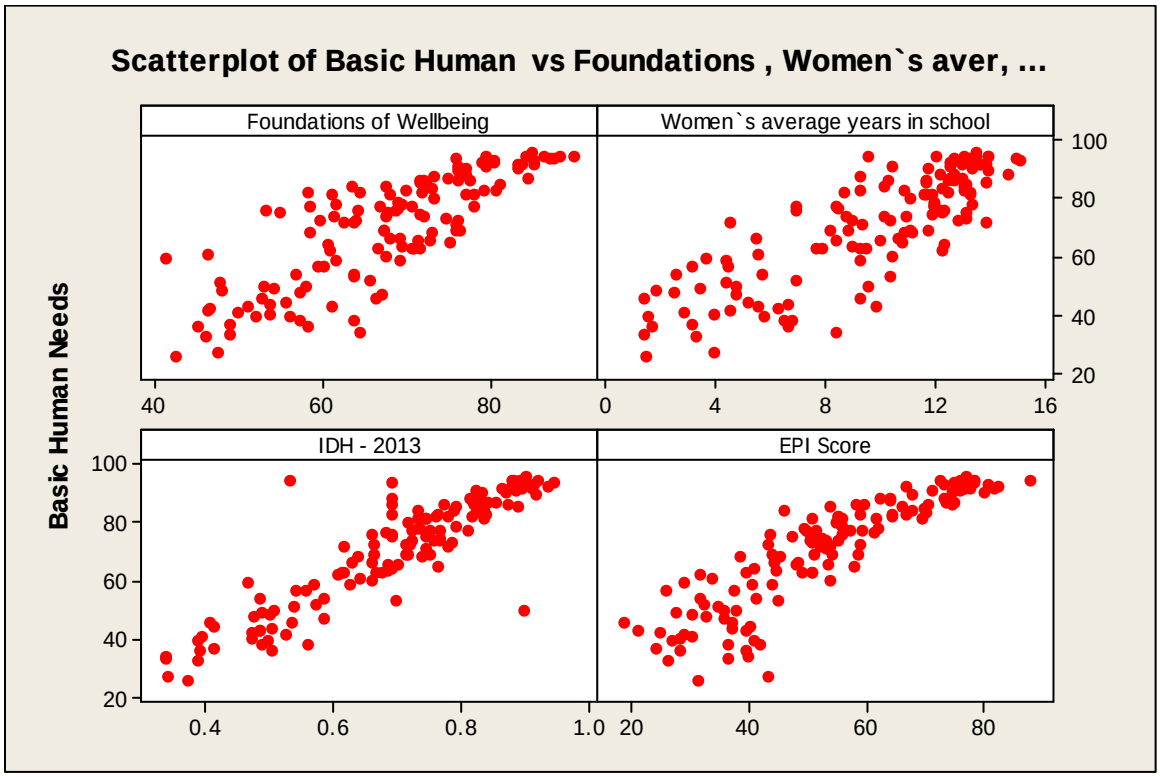


Figura 34: Gráficos de dispersão, variável de interesse *Basic Human Needs* – Necessidades Humanas Básicas

Acima nota-se correlação positiva entre a variável Necessidades Humanas Básicas (*Basic Human Needs*) e as demais verificadas. Destaca-se a variável IDH-2013, Índice de Desenvolvimento Humano, que resultou em 0.901 de grau de relação linear.

2.2.4 Variável de interesse *Foundations of Wellbeing –Fundamentos do Bem-Estar*

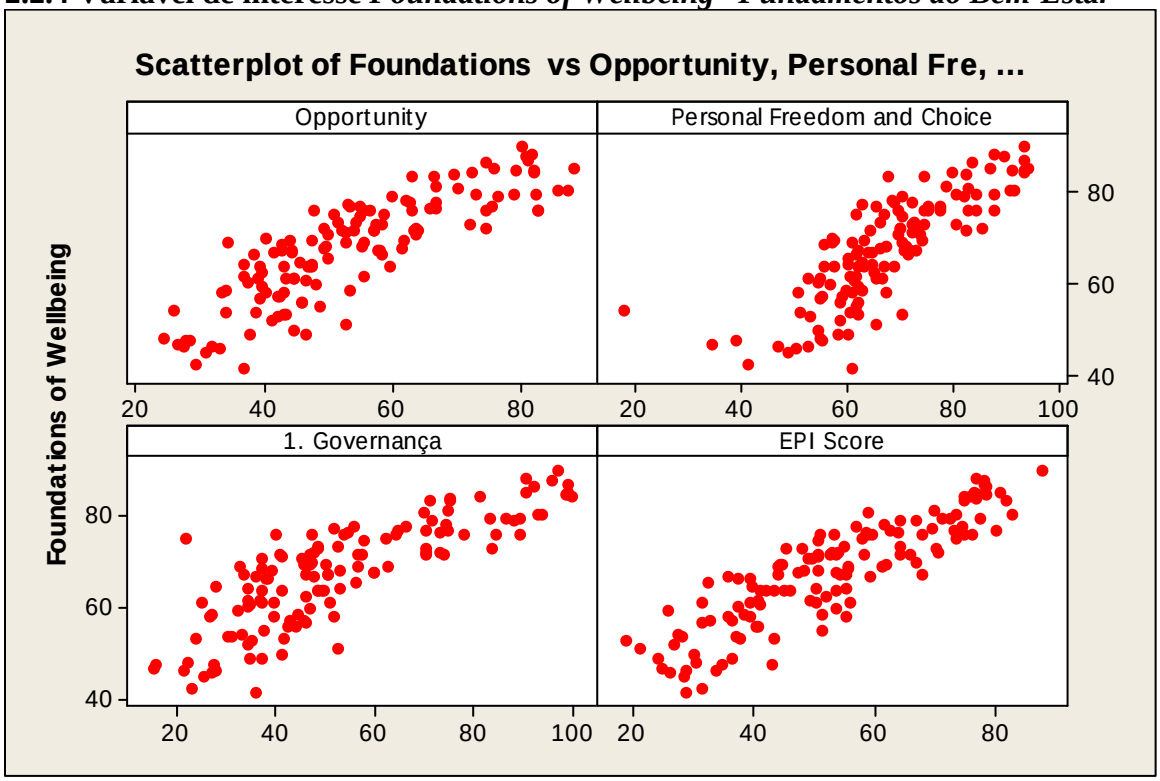


Figura 35: Gráficos de dispersão, variável de interesse *Foundations of Wellbeing* –Fundamentos do Bem-Estar

A variável Fundamentos do Bem-Estar (*Foundations of Wellbeing*), mostra no gráfico acima correlação positiva entre as variáveis em destaque no estudo. A que apresentou maior grau de correlação foi o Índice de Performance Ambiental (*EPI Score*) com grau de 0.879.

2.3 Linhas de tendência

Para esta análise, utilizamos a relação entre as variáveis com maior grau de correlação.

2.3.1 Variável de interesse *EPI Score* – Índice de Performance Ambiental

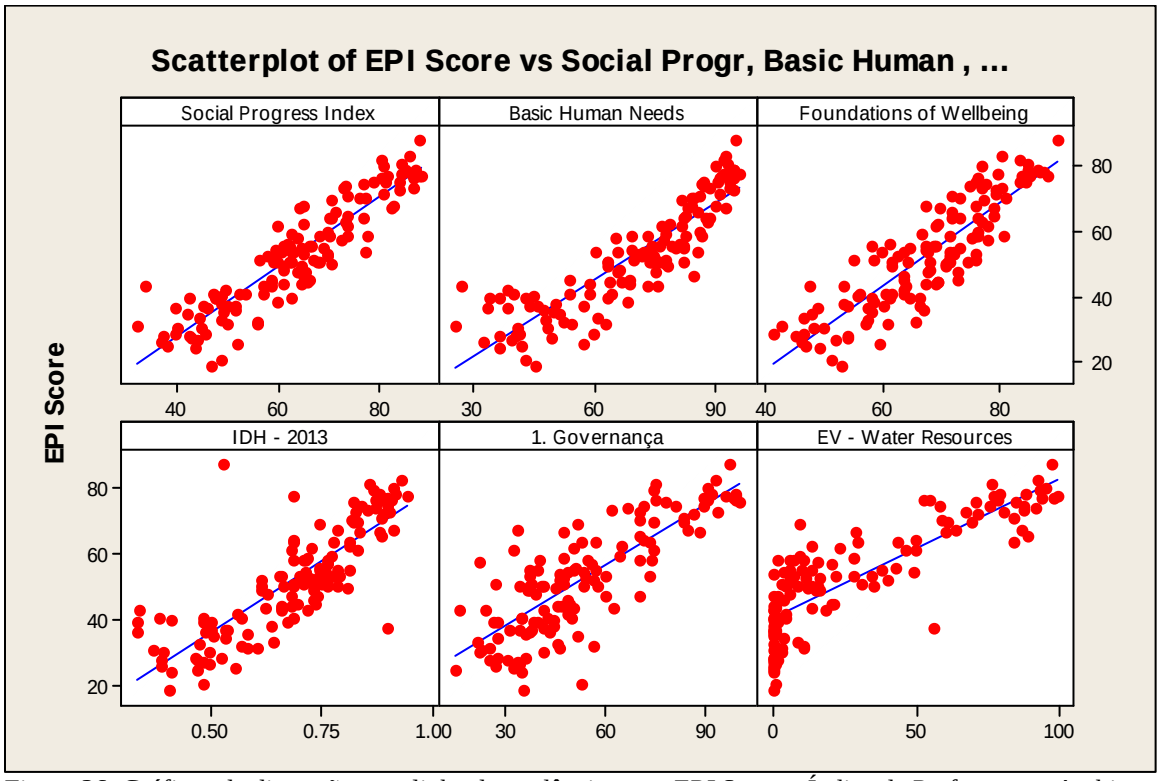


Figura 36: Gráficos de dispersão com linha de tendência, para *EPI Score* – Índice de Performance Ambiental

No gráfico acima (figura 36), percebe-se correlação linear positiva entre a variável EPI e as demais variáveis do estudo. O maior destaque foi com a variável Índice de Progresso Social (*Social Progress Index*), que apresentou correlação de 0.917 e *p-Value* de 0.000.

2.3.2 Variável de interesse *Social Progress Index* – Índice de Progresso Social

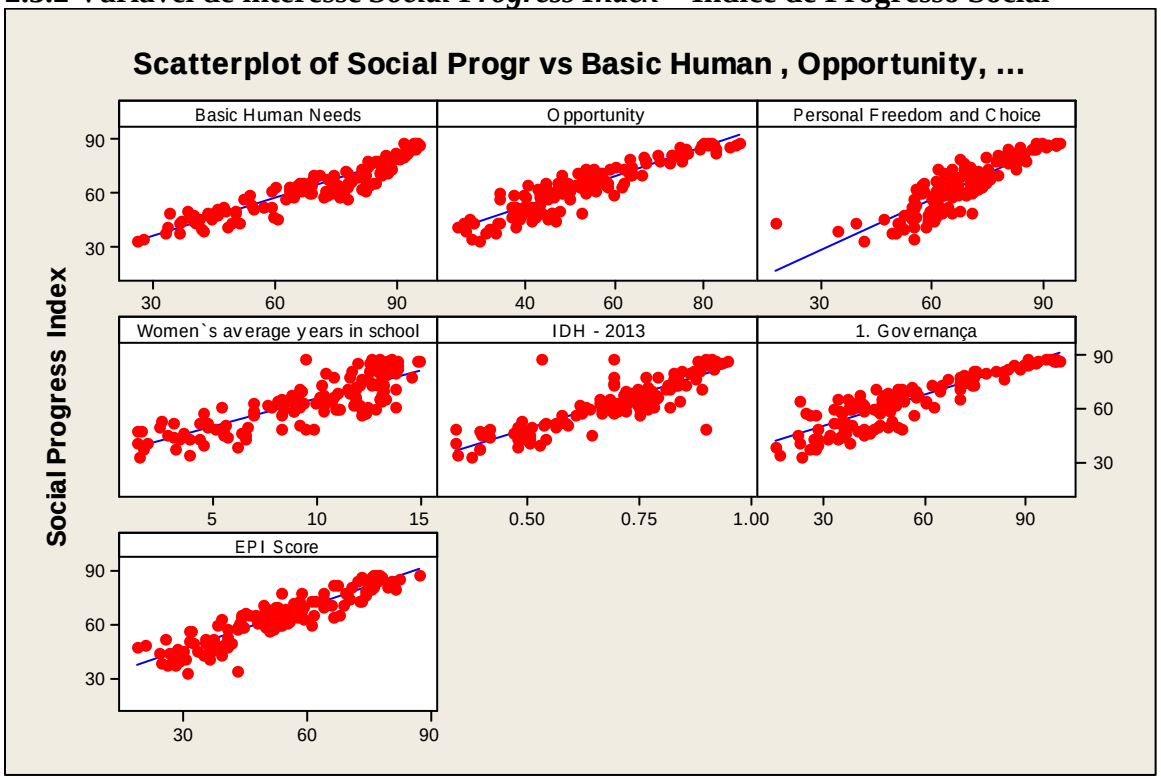


Figura 37: Gráficos de dispersão com linha de tendência para *Social Progress Index* – Índice de Progresso Social

O gráfico baseado na variável de Índice de Progresso Social (*Social Progress Index*), mostra uma correlação positiva entre as variáveis. Com destaques para as variáveis Necessidades Humanas Básicas (*Basic Human Needs*) com correlação de 0.934 e Oportunidade (*Opportunity*) que apresentou correlação de 0.922 com a variável de referência.

2.3.3 Variável de interesse *Basic Human Needs* – Necessidades Básicas Humanas

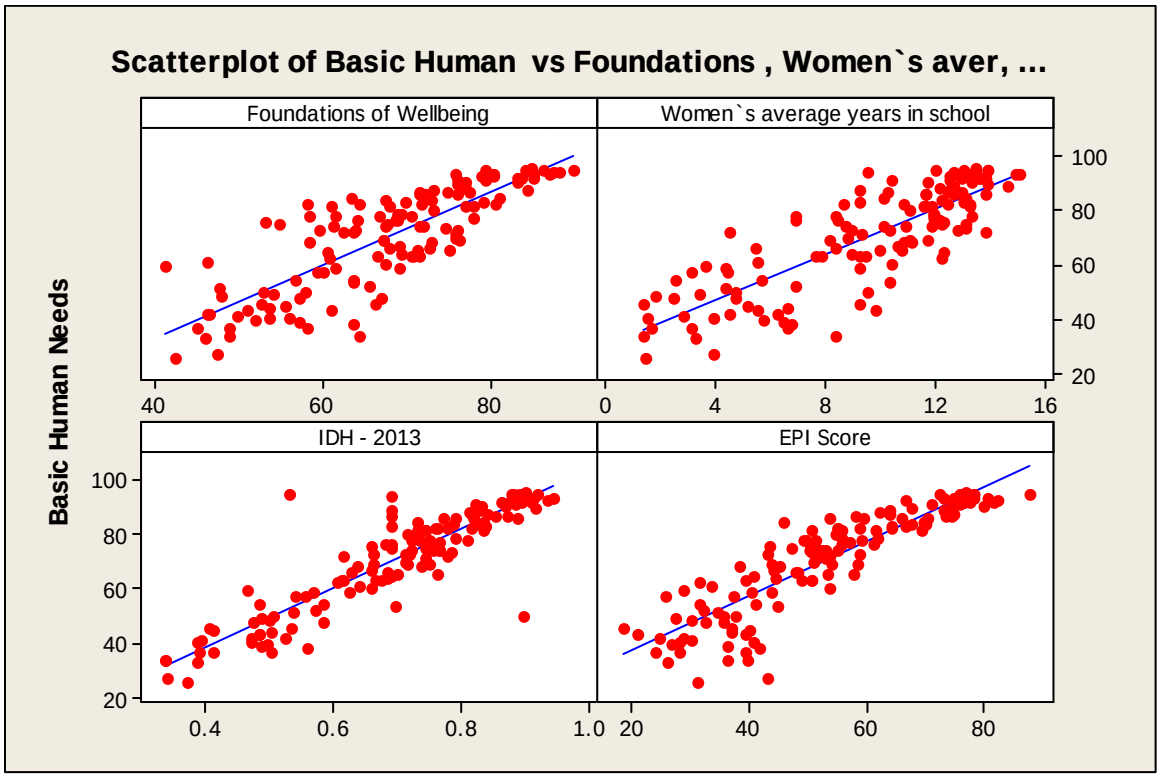


Figura 38: Gráficos de dispersão com linha de tendência para *Basic Human Needs* – Necessidades Básicas Humanas

Acima nota-se correlação positiva entre a variável Necessidades Humanas Básicas (*Basic Human Needs*) e as demais verificadas. Destaca-se a variável IDH-2013, Índice de Desenvolvimento Humano, que resultou em 0.901 de grau de relação linear.

2.3.4 Variável de interesse *Foundations of Wellbeing* – Fundamentos para o Bem-Estar

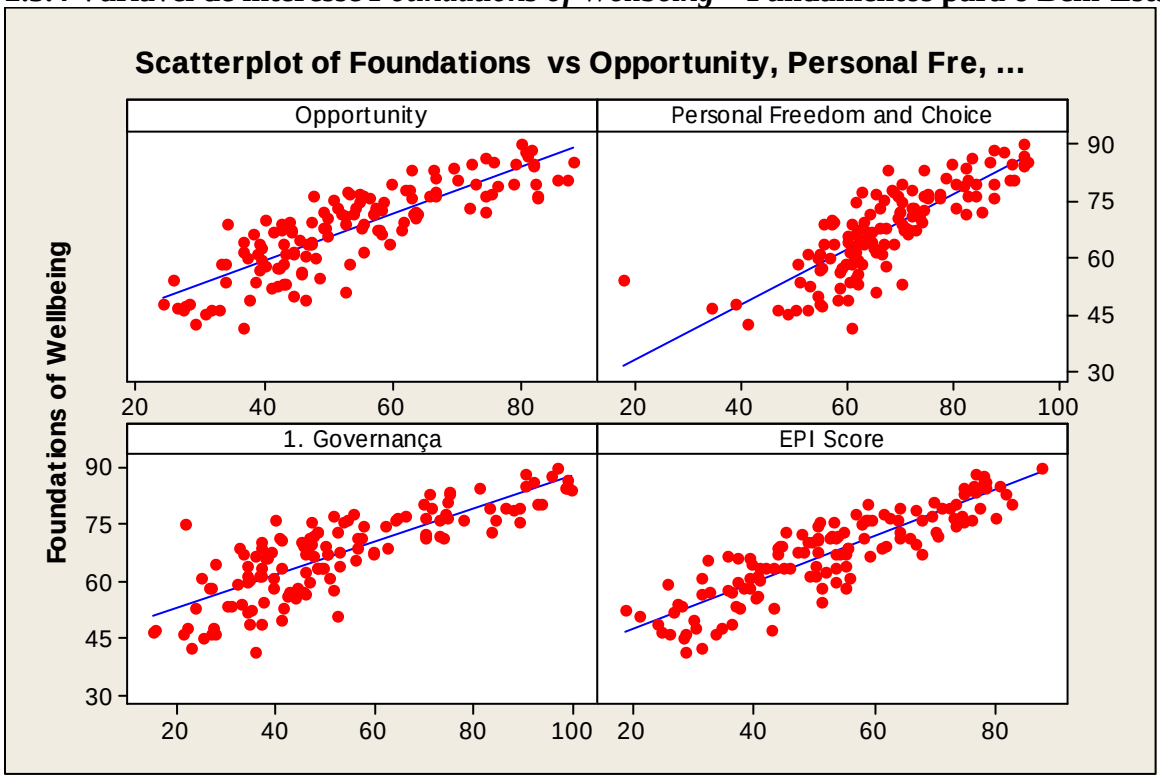


Figura 39: Gráficos de dispersão com linha de tendência para *Foundations of Wellbeing* – Fundamentos para o Bem-Estar

A variável Fundamentos do Bem-Estar (*Foundations of Wellbeing*), mostra no gráfico acima correlação positiva entre as variáveis em destaque no estudo. A que apresentou maior grau de correlação foi o Índice de Performance Ambiental (*EPI Score*) com grau de 0.879.

2.4 Dendrogram (Dendrograma ou Dendograma) – Diagrama de Árvore

Um Dendrograma (dendr(o) = árvore) é um tipo específico de diagrama ou representação icônica que organiza determinados fatores e variáveis. É um diagrama de similaridade. A interpretação de um dendrograma de similaridade entre amostras fundamenta-se na intuição: duas amostras próximas devem ter também valores semelhantes para as variáveis medidas. Ou seja, elas devem ser próximas matematicamente no espaço multidimensional. Portanto, quanto maior a proximidade entre as medidas relativas às amostras, maior a similaridade entre elas. O dendrograma hierarquiza esta similaridade de modo que podemos ter uma visão bidimensional da similaridade ou dissimilaridade de todo o conjunto de amostras utilizado no estudo. Segue abaixo o Dendrograma das variáveis analisadas:

Um dendrograma ilustra as informações na tabela de amalgamação na forma de um diagrama de árvore.

Por padrão, o nível de similaridade é medido no eixo vertical (também é possível exibir o nível de distância) e as diferentes observações são listadas ao longo do eixo horizontal. O gráfico mostra como os agrupamentos são formados: unindo duas observações individuais ou pareando uma observação individual com um agrupamento existente. É possível ver em que nível de similaridade os agrupamentos são formados, e a composição dos agrupamentos da partição final.

Para alguns conjuntos de dados, os métodos média, centróide, mediana e Ward podem não gerar um dendrograma hierárquico. Ou seja, as distâncias de amalgamação nem sempre aumentam com cada etapa. No dendrograma, essa etapa geraria uma junção para baixo no lugar de para cima.

O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação ou alto grau de similaridade.

2.4.1 Análise de cluster das variáveis - *Dendrogram*

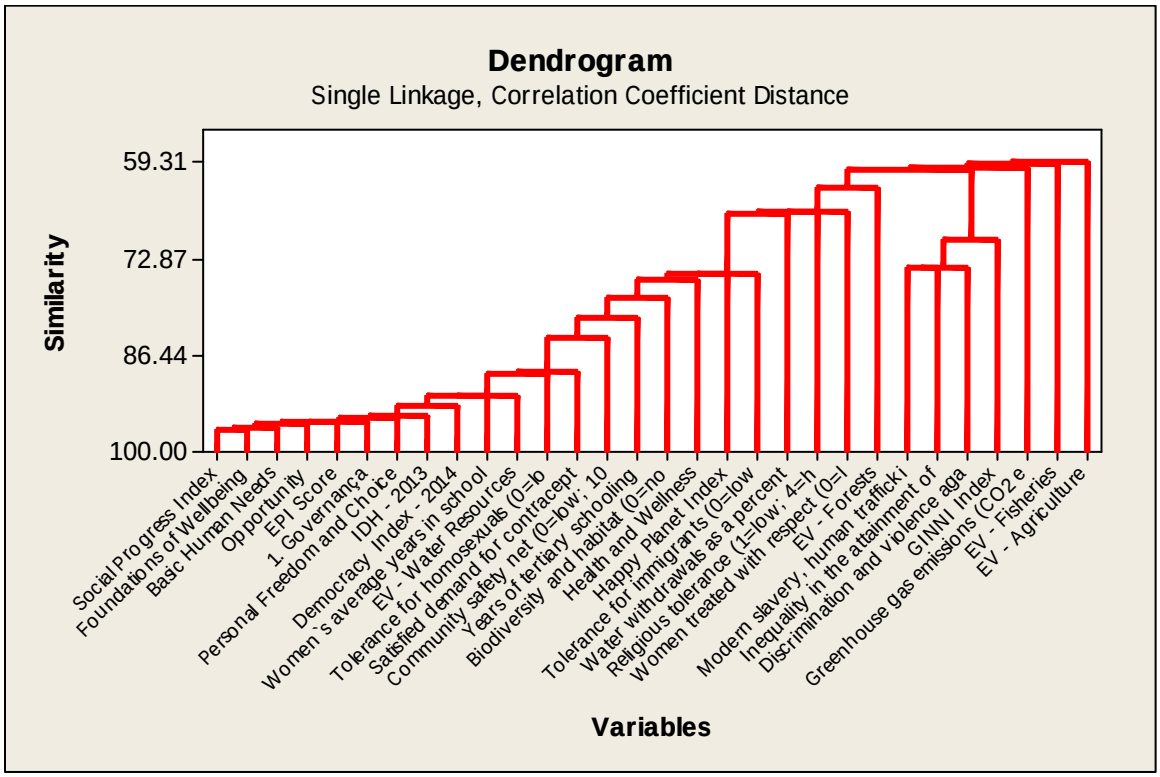


Gráfico (acima) gerado considerando apenas 1 cluster.

Em seguida, segue análise de cluster das variáveis deste estudo:

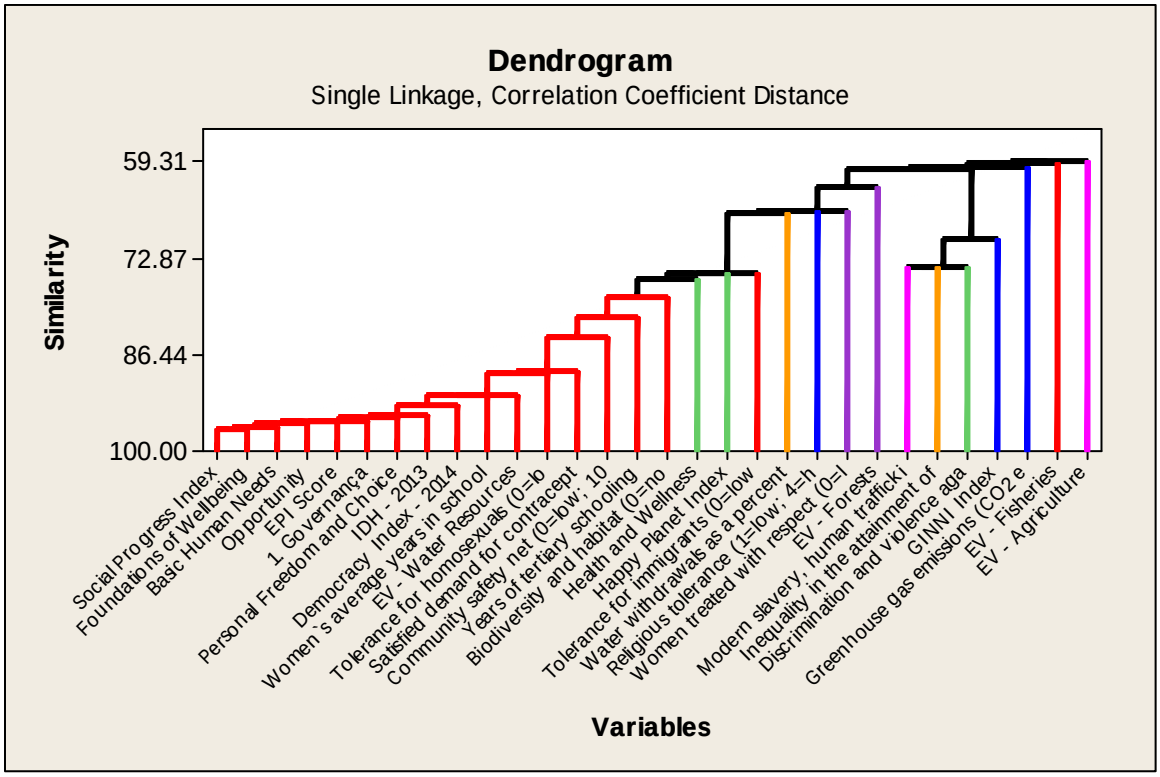
Cluster Analysis of Variables: Social Progress, Basic Human Needs, ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

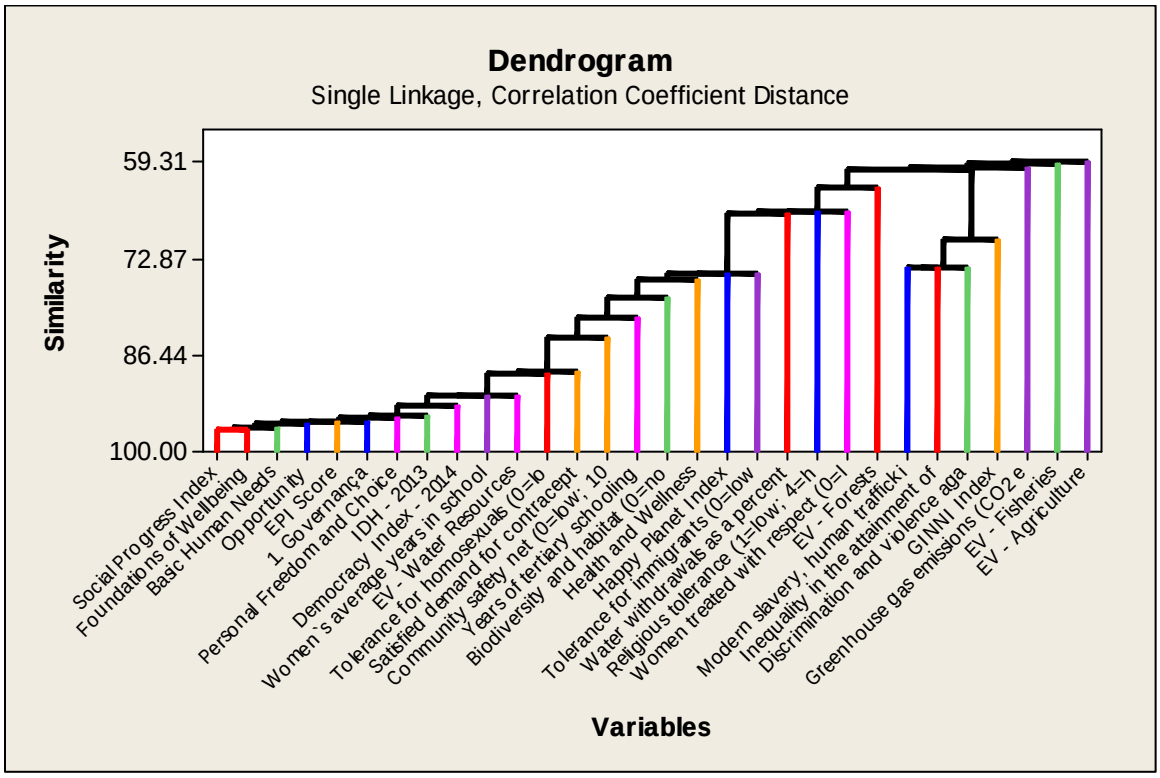
Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	29	96.9544	0.060912	1	3	1	2
2	28	96.6963	0.066073	1	2	1	3
3	27	96.0972	0.078056	1	4	1	4
4	26	95.8701	0.082598	1	23	1	5
5	25	95.8063	0.083875	1	22	1	6
6	24	95.3580	0.092840	1	6	1	7
7	23	95.0464	0.099071	1	21	1	8
8	22	93.4716	0.130567	1	30	1	9
9	21	92.1988	0.156024	1	19	1	10
10	20	92.0028	0.159944	1	24	1	11
11	19	89.0470	0.219061	1	14	1	12
12	18	88.6723	0.226554	1	11	1	13
13	17	84.0626	0.318747	1	17	1	14
14	16	81.2353	0.375293	1	18	1	15
15	15	78.2833	0.434334	1	9	1	16
16	14	75.8883	0.482233	1	5	1	17
17	13	75.0977	0.498045	1	28	1	18
18	12	74.9784	0.500432	1	13	1	19
19	11	74.1538	0.516925	10	20	10	2
20	10	74.1254	0.517492	10	15	10	3
21	9	70.3083	0.593835	10	29	10	4
22	8	66.5831	0.668338	1	8	1	20
23	7	66.3931	0.672139	1	16	1	21
24	6	66.3052	0.673897	1	12	1	22
25	5	62.8294	0.743413	1	26	1	23
26	4	60.3164	0.793671	1	10	1	27
27	3	60.0636	0.798727	1	7	1	28
28	2	59.5788	0.808425	1	27	1	29
29	1	59.3122	0.813756	1	25	1	30

Dendrogram

O gráfico abaixo utiliza um total de 15 clusters, divide as observações no ponto médio nível de similaridade (78.2833):



O seguinte gráfico foi construído considerando um total de 29 clusters, que segundo o software estatístico utilizado neste trabalho, é considerado como o maior nível de similaridade entre as variáveis (96.9544):



3. REGRESSÃO LINEAR MÚLTIPLA

3.1 Normalização dos Dados

Antes de partir para as análises de regressão, vamos proceder com a normalização de nossa base de dados.

A normalização dos dados permite atribuir uma nova escala para que todas as variáveis sejam expressas no mesmo padrão.

Neste trabalho foi utilizada a Normalização Min-Max, que se baseia nos valores mínimos e máximos das observações. Este método preserva o relacionamento entre os valores dos dados originais.

Todas as variáveis quantitativas neste trabalho foram normalizadas. A equação utilizada para normalização da base de dados foi:

(c4-MIN(c4))/(MAX(c4)-MIN(c4))*100

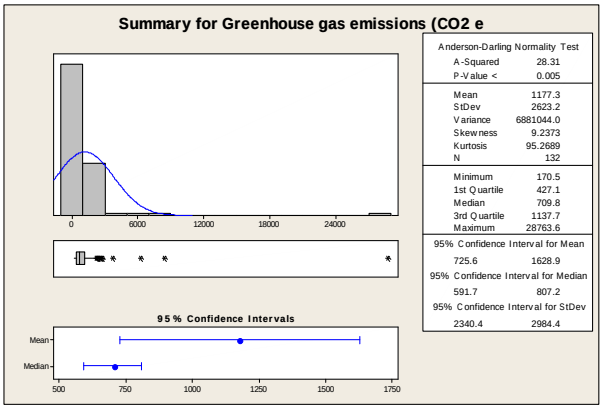
Onde: c4 refere-se a coluna dos dados que serão normalizados pelo Método Min-Mix

O procedimento acima, foi aplicado nas 30 variáveis quantitativas deste estudo. Adicionalmente, algumas variáveis tiveram ajustes em suas escalas (positivadas). Nas variáveis que demonstraram algum tipo de agravamento (negativo) a medida de seu aumento, por exemplo, o Índice de Emissão de Gases Poluentes que quanto maior seu resultado, menor é seu benefício ao meio ambiente e a saúde humana, foi aplicado a fórmula:
(100 - "valor da célula"), para todas observações daquela variável. Desta maneira, todos os indicadores referentes a este estudo passam a refletir um melhor desempenho a medida que se aproximam de 100 e o inverso também é verdadeiro.

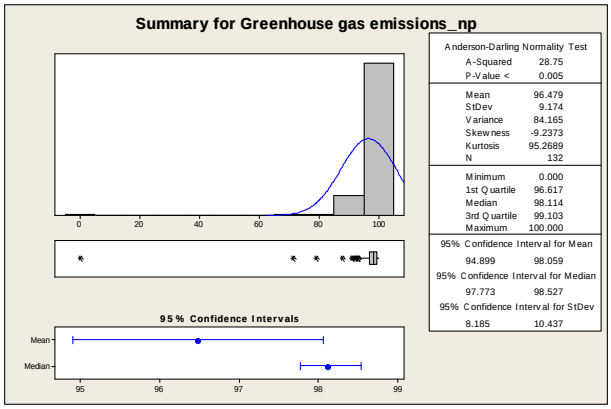
As seguintes variáveis foram "positivadas" para que permitissem o entendimento de quanto maior o seu valor, melhor (ou mais positiva) era a situação do indicador: *Greenhouse gas emissions_np, Water withdrawals_np, Modern slavery_np, Discrimination and violence_np, Inequality in the attainment_np, EV-Fisheries_np e GINI Index_np.*

Visualmente, estas variáveis são identificadas pelo sufixo "np", que denota ser uma variável normalizada e positivada.

Para efeitos de validação do método de "positivar" as variáveis deste estudo e garantir que não haveria perda de referência, foi gerado novo gráfico para a variável "*Greenhouse gas emissions_np*" e confrontado com o gráfico gerado na primeira fase deste trabalho.



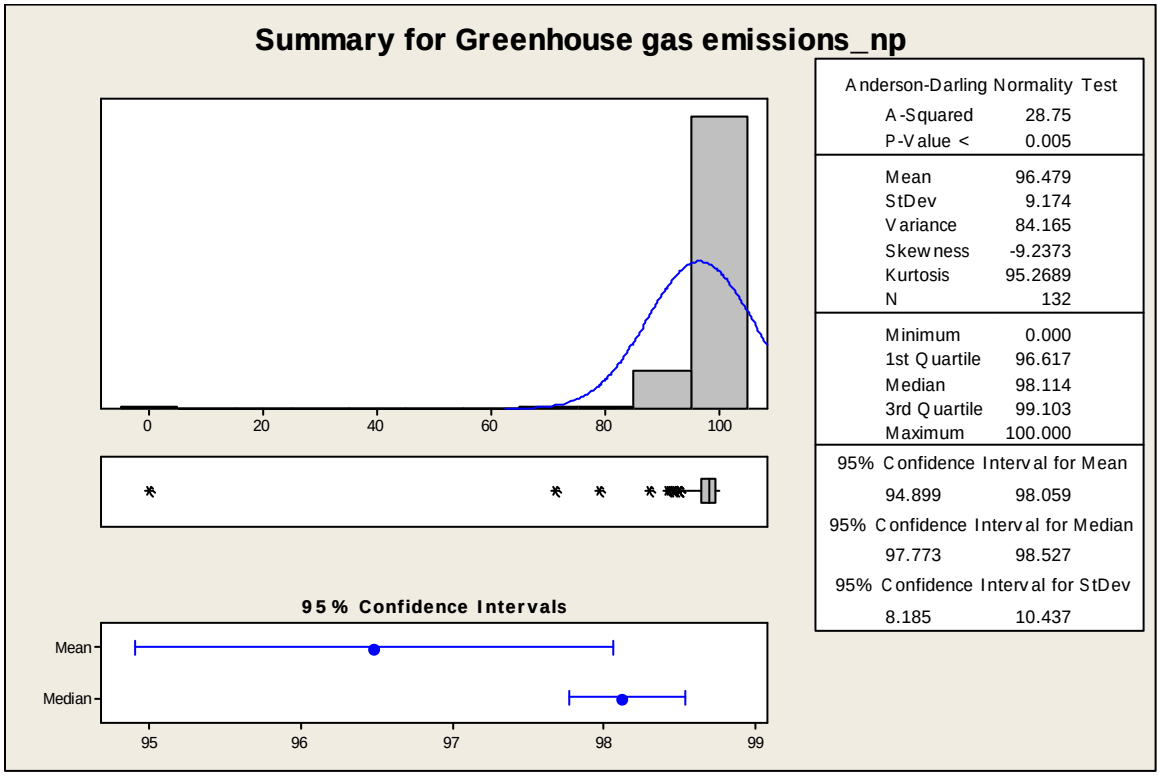
ANTES de normalizar e positivar

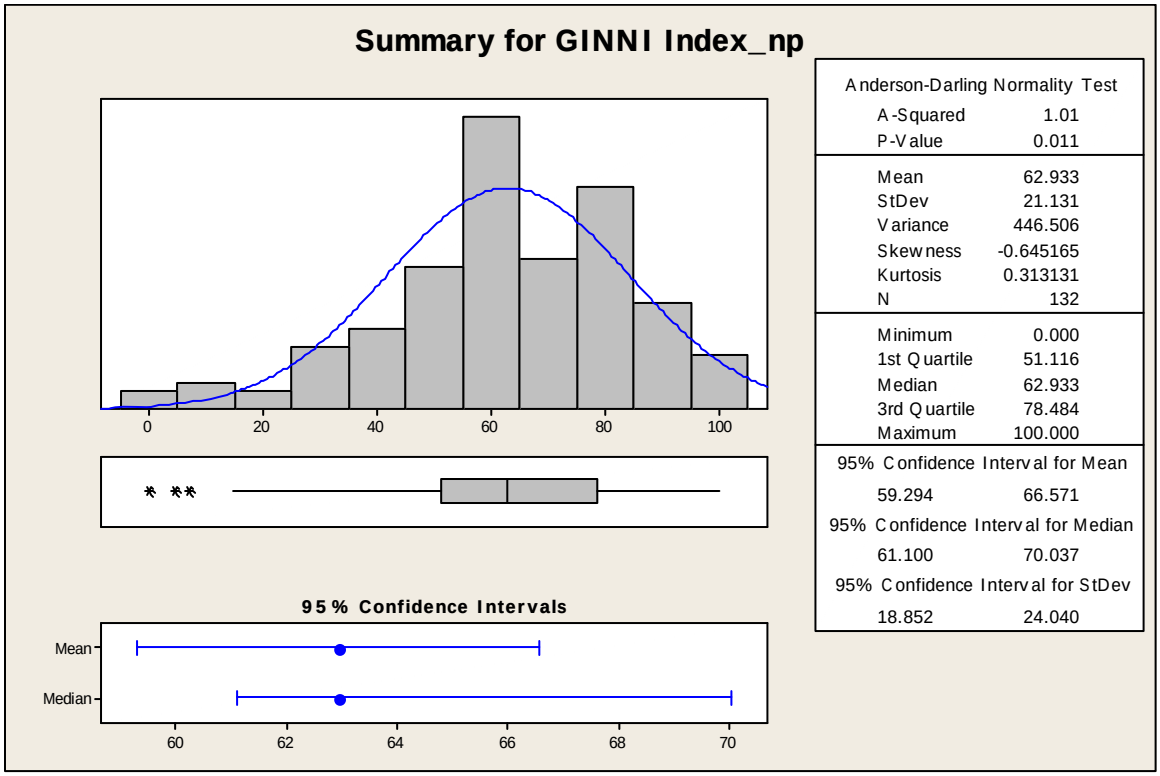


DEPOIS de normalizar e positivar

A comparação acima dos dois momentos, antes e depois de normalizar e positivar, mostra que foi mantido proporcionalmente o relacionamento na disposição dos dados.

Greenhouse gas emissions - variável NORMALIZADA e POSITIVADA – resultado





3.2 Análise de Regressão

Nesta etapa do trabalho, serão executadas as análises de regressão, para tanto será utilizado o software estatístico Minitab e a base dados contendo as variáveis referentes ao tema Biodiversidade – Capital Natural e Social, objeto deste estudo, já normalizadas e positivadas.

No software estatístico, a opção para executar a função de análise de regressão no menu é: *stat -> Regression -> Regression*

Será utilizada a variável *EPI Score* como variável independente (*response*), e como dependentes (preditoras), as variáveis: *Social Progress Index*, *Basic Human Needs*, *Foundations of Wellbeing*, *Opportunity*, 1. Governança, *Personal Freedom and Choice*, IDH – 2013, *Democracy Index – 2014*, *Women`s average years in school* e *EV - Water Resources.*, que apresentaram bom nível de similaridade, conforme dendrograma apresentado na seção anterior.

Abaixo, os resultados da regressão com as variáveis mencionadas:

Regression Analysis: EPI Score_n versus Social Progress, Basic Human Needs, ...

```
The regression equation is
EPI Score_n = 344 + 90 Social Progress Index_n - 37.5 Basic Human Needs_n
              - 25.8 Foundations of Wellbeing_n - 34.4 Opportunity_n
              + 0.0200 Governança_n - 0.0497 Personal Freedom and Choice_n
              - 0.0171 IDH-2013_n - 0.177 Democracy Index-2014_n
              + 0.0036 Women`s years in school_n + 0.257 EV-Water Resources_n

Predictor                          Coef  SE Coef      T      P
Constant                          343.6    409.5   0.84  0.403
Social Progress Index_n            90.4    109.9   0.82  0.412 <- descartar
Basic Human Needs_n               -37.54    45.96  -0.82  0.416 <- descartar
Foundations of Wellbeing_n        -25.77    31.90  -0.81  0.421 <- descartar
Opportunity_n                     -34.44    41.94  -0.82  0.413 <- descartar
Governança_n                     0.02005  0.07622   0.26  0.793 <- descartar
Personal Freedom and Choice_n  -0.04970  0.09268  -0.54  0.593 <- descartar
IDH-2013_n                     -0.01712  0.06032  -0.28  0.777 <- descartar
Democracy Index-2014_n         -0.17736  0.05849  -3.03  0.003
Women`s years in school_n       0.00359  0.04809   0.07  0.941 <- descartar
EV-Water Resources_n            0.25673  0.03308   7.76  0.000

S = 6.95727   R-Sq = 92.2%   R-Sq(adj) = 91.5%

Analysis of Variance

Source           DF       SS      MS       F      P
Regression       10  68997.8  6899.8  142.55  0.000
Residual Error  121   5856.8    48.4
Total           131  74854.7


Source                          DF   Seq SS
Social Progress Index_n          1  62999.6
Basic Human Needs_n              1    554.4
Foundations of Wellbeing_n       1    782.4
Opportunity_n                    1      3.4
Governança_n                     1    363.1
Personal Freedom and Choice_n    1      9.0
IDH-2013_n                       1     41.7
Democracy Index-2014_n           1   1327.9
Women`s years in school_n        1      0.2
EV-Water Resources_n             1   2916.1

Unusual Observations

       Social
     Progress
Obs   Index_n  EPI Score_n     Fit  SE Fit  Residual  St Resid
  9        35       10.370  27.320   1.736   -16.950     -2.52R
 23         3       35.399  12.282   2.268    23.117      3.51R
 44        42       19.700  35.687   1.507   -15.987     -2.35R
 77        26        0.000  16.864   2.112   -16.864     -2.54R
 78        19       12.652  23.061   3.942   -10.409     -1.82 X
 91        13       25.780  11.364   2.000    14.416      2.16R
106        68       73.224  56.749   1.627    16.475      2.44R
113        29       27.325  38.158   3.953   -10.832     -1.89 X
115       100      100.000  95.588   3.971     4.412      0.77 X
116        42       18.645  39.131   2.355   -20.486     -3.13R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

No resultado acima vemos que apesar do grau de similaridade no gráfico de Dendrograma e r-quadrado de 92.2% indicar uma boa relação entre as variáveis testadas, o resultado da análise mostrou o contrário, das 10 variáveis testadas, apenas 2 apresentaram p-value abaixo de 0.05 e destas, somente 1 com estatística t (de Student) com valor positivo.

Foi executado novamente a análise de regressão, desta vez mantido *EPI Score* como variável independente, descartado aquelas que apresentaram alto grau de imprecisão (p-value > 0.05) e considerado como dependentes (preditoras), as variáveis já testadas: *Democracy Index–2014* e *EV-Water Resources* e adicionado a variável: *Biodiversity and Habitat.*

Abaixo, os resultados com a nova constituição de variáveis testadas:

Regression Analysis: EPI Score_n versus Democracy Index, EV-Water Resources,...

```
The regression equation is
EPI Score_n = 14.0 + 0.151 Democracy Index-2014_n + 0.502 EV-Water Resources_n
              + 0.222 Biodiversity and habitat_n


Predictor                      Coef  SE Coef      T      P
Constant                     14.039    2.716   5.17  0.000
Democracy Index-2014_n      0.15105  0.04842   3.12  0.002
EV-Water Resources_n        0.50182  0.03503  14.33  0.000
Biodiversity and habitat_n  0.22174  0.03559   6.23  0.000


S = 10.8256   R-Sq = 80.0%   R-Sq(adj) = 79.5%


Analysis of Variance

Source           DF     SS     MS       F      P
Regression        3  59854  19951  170.24  0.000
Residual Error  128  15001    117
Total           131  74855


Source                       DF  Seq SS
Democracy Index-2014_n        1   30003
EV-Water Resources_n          1   25303
Biodiversity and habitat_n    1    4548


Unusual Observations

     Democracy     EPI
Obs  Index-2014_n  Score_n     Fit  SE Fit  Residual  St Resid
 54             6   47.155  24.540   2.065    22.615      2.13R
 77            51    0.000  24.905   2.047   -24.905     -2.34R
106            62   73.224  49.970   1.945    23.254      2.18R
113            19   27.325  49.008   2.732   -21.683     -2.07R
125            14   78.683  71.542   3.422     7.141      0.70 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Na segunda execução da análise de regressão, tomando as variáveis pré-selecionadas, temos bom nível de similiaridade nos resultados, com *p-value* próximo de 0 e e r-quadrado de 80%.. Todos os resultados da estatística t (de Student) apresentaram valor positivo.

3.3 Correção dos Dados (outliers)

A ferramenta de análise de regressão identificou um total de 5 observações em nossa base, as quais apresentaram valores inesperados (*outliers*) para a variável *EPI Score* , as referidas células foram modificadas manualmente, recebendo cada uma seu valor correspondente da coluna *FIT*, conforme abaixo:

Unusual Observations

Obs	Democracy Index-2014_n	EPI Score_n	Fit	SE Fit	Residual	St Resid
54	6	47.155	24.540	2.065	22.615	2.13R
77	51	0.000	24.905	2.047	-24.905	-2.34R
106	62	73.224	49.970	1.945	23.254	2.18R
113	19	27.325	49.008	2.732	-21.683	-2.07R
125	14	78.683	71.542	3.422	7.141	0.70 X

Aplicado os ajustes, executado novamente a análise, surgiram outras 5 observações com necessidade de correção. Inclusive repetindo a Obs nr.125

Unusual Observations

Obs	Democracy Index-2014_n	EPI Score_n	Fit	SE Fit	Residual	St Resid
9	51	10.370	30.912	1.368	-20.542	-2.08R
52	76	18.486	40.302	1.905	-21.816	-2.23R
70	61	3.437	25.429	2.437	-21.992	-2.28R
104	4	69.656	48.286	2.775	21.370	2.23R
125	14	71.542	70.164	3.150	1.378	0.15 X

3.4 Regressão Stepwise

A regressão stepwise é uma ferramenta automática usada nos estágios exploratórios da construção de modelos para identificar um subconjunto útil de preditores. O processo sistematicamente adiciona a variável mais significante ou remove a variável menos significante durante cada etapa.

Fonte:http://support.minitab.com/pt-br/minitab/17/topic-library/modeling-statistics/regression-and-correlation/basics/basics-of-stepwise-regression/

Devido ao problema de ajuste na fase anterior (seção 3.3), então somente para demonstrar o procedimento e fins de ilustração, será executado a ferramenta de regressão stepwise, utilizando a variável de interesse: *EPI Score*, e as seguintes variáveis como

dependentes: faixa de 'Social Progress Index_n'- até 'Governança_n' e a faixa de variáveis de 'EV-Water Resources_n'- até 'Democracy Index-2014_n'. Os resultados, seguem abaixo:

Stepwise Regression: EPI Score_n versus Social Progress, Basic Human Needs, ...

```
  Alpha-to-Enter: 0.15  Alpha-to-Remove: 0.15

Response is EPI Score_n on 29 predictors, with N = 132
```

Step	1	2	3	4	5	6
Constant	2.1011	9.6294	5.2423	0.6514	-2.1860	-1.4417
Social Progress Index_n	0.842	0.569	0.506	0.173	0.221	0.408
T-Value	25.96	14.36	13.55	2.38	3.15	4.09
P-Value	0.000	0.000	0.000	0.019	0.002	0.000
EV-Water Resources_n		0.276	0.289	0.288	0.287	0.314
T-Value		9.03	10.48	11.44	12.02	12.26
P-Value		0.000	0.000	0.000	0.000	0.000
Biodiversity and habitat_n			0.126	0.177	0.199	0.195
T-Value			5.57	7.74	8.91	8.93
P-Value			0.000	0.000	0.000	0.000
Basic Human Needs_n				0.320	0.253	0.160
T-Value				5.17	4.16	2.30
P-Value				0.000	0.000	0.023
EV-Forests_n					0.079	0.087
T-Value					3.95	4.41
P-Value					0.000	0.000
Governança_n						-0.139
T-Value						-2.58
P-Value						0.011
S	9.48	7.45	6.71	6.12	5.80	5.67
R-Sq	83.83	90.09	92.02	93.41	94.14	94.43
R-Sq(adj)	83.71	89.94	91.84	93.20	93.91	94.17

Step	7	8	9
Constant	11.68	10.45	14.25

4. ANÁLISE DE TENDÊNCIAS

Nesta sessão será efetuada análise de tendência e projeções das variáveis temporais e quantitativas *"Terrestrial and marine protected areas"* (áreas terrestres e marinhas sob proteção), *"Forest area"* (área florestal) e *"Population"* (população total de um país), relacionadas ao tema **Biodiversidade - Capital Natural e Social**. Os dados se referem aos países Brasil, México, Portugal e Espanha e foram extraídos do site ***Knoema***, que trata-se de um projeto de integração de bases de dados de diversos órgãos internacionais, como Banco Mundial, ONU, Unesco, entre outros.

A análise de cada variável está dividida em três partes. A primeira refere-se a análise do comportamento histórico através de gráficos. A segunda trata da análise de tendências, buscando encontrar a função estatística que melhor se adapte à cada uma das séries de variáveis analisadas. E a terceira busca a elaboração e análise de projeções através de extrapolações estatísticas. O software estatístico utilizado é o MINITAB.

4.1 Uma breve descrição do Projeto KNOEMA

O Projeto Knoema pode ser descrito como uma plataforma de conhecimento. A ideia de concepção reside no tripé: Disponibilização dos dados, Ferramentas de análise *online* e Ferramentas para elaboração de relatórios e apresentação dos resultados. Isso possibilita ao pesquisador selecionar um conjunto de dados de interesse, executar análises em busca de relevância e elaborar o relatório final com suas conclusões, diretamente na plataforma Knoema. fonte: http://knoema.com/about/us

4.2. Séries históricas

Os dados são séries históricas referentes ao Brasil, México, Portugal e Espanha, portanto, tratam-se de séries temporais. As séries vão de 1990 a 2012 para a variável "*Terrestrial and marine protected areas"* (áreas terrestres e marinhas sob proteção), e de 1990 a 2013 para as variáveis "*Forest area"* (área florestal) e "*Population"* (população total de um país).

Cabe mencionar que para a variável "*Terrestrial and marine protected areas"* (áreas terrestres e marinhas sob proteção), não havia dados para o ano de 2011, para nenhum dos 4 países listados neste estudo. Logo, para fim deste exercício, foi inserido o valor da média dos anos 2010 e 2012 de cada país em sua respectiva célula.

As três séries de dados podem ser consideradas satisfatórias para a realização desta pesquisa, uma vez que todas possuem dados de mais de 20 anos, sem interrupções.

4.2.1 As Variáveis

São 4 as variáveis desta pesquisa, incluindo o ano a que se referem os dados. Estas variáveis são melhor explicadas na Tabela 4.1.

Tabela 4.1. As Variáveis

Variável	Significado	Tipo	Unidade de Medida
Ano	É o ano a que se refere o dado de cada variável e de cada país.	Variável Categórica	Formato: AAAA
TerMarProt_XXX *Terrestrial and marine protected areas* (áreas terrestres e marinhas sob proteção) (unit: % of total territorial area) Período de 1990 a 2012	É a taxa que representa as áreas terrestres e marinhas que estão sob proteção em determinado país. É medida considerando a relação entre o total da área protegida pelo total da área territorial do país, em determinado ano. O sufixo XXX, no nome da variável, denota o nome abreviado do país em análise, por exemplo: BRA para Brasil, MEX para México, POR para Portugal e SPA para Espanha.	Variável Quantitativa	Percentual (%)
Forest_XXX *Forest area* (área florestal) (unit: sq. km) Período de 1990 a 2013	Refere-se a medida da área coberta por florestas que ocupam acima de 0.5 hectares, com árvores maiores que 5 metros de altura. Não inclui áreas destinadas a agricultura e zonas urbanas. É caracterizada pela presença de árvores e pela não utilização destas terras para qualquer outra finalidade. O sufixo XXX, no nome da variável, denota o nome abreviado do país em análise, por exemplo: BRA para Brasil, MEX para México, POR para Portugal e SPA para Espanha. .	Variável Quantitativa	Quilômetros quadrados (km^2)
Pop_XXX *Population* (população total de um país) (unit: Persons) Período de 1990 a 2013	Mede a quantidade total de indivíduos que vivem no país (residentes), independente da nacionalidade, exceto refugiados e asilados. O sufixo XXX, no nome da variável, denota o nome abreviado do país em análise, por exemplo: BRA para Brasil, MEX para México, POR para Portugal e SPA para Espanha.	Variável Quantitativa	Quantidade de indivíduos

4.3. ANÁLISE DAS VARIÁVEIS

Cada variável será analisada utilizando gráficos para demonstrar o comportamento histórico da série, linhas de tendência, funções, erros das funções, além de extrapolações estatísticas.

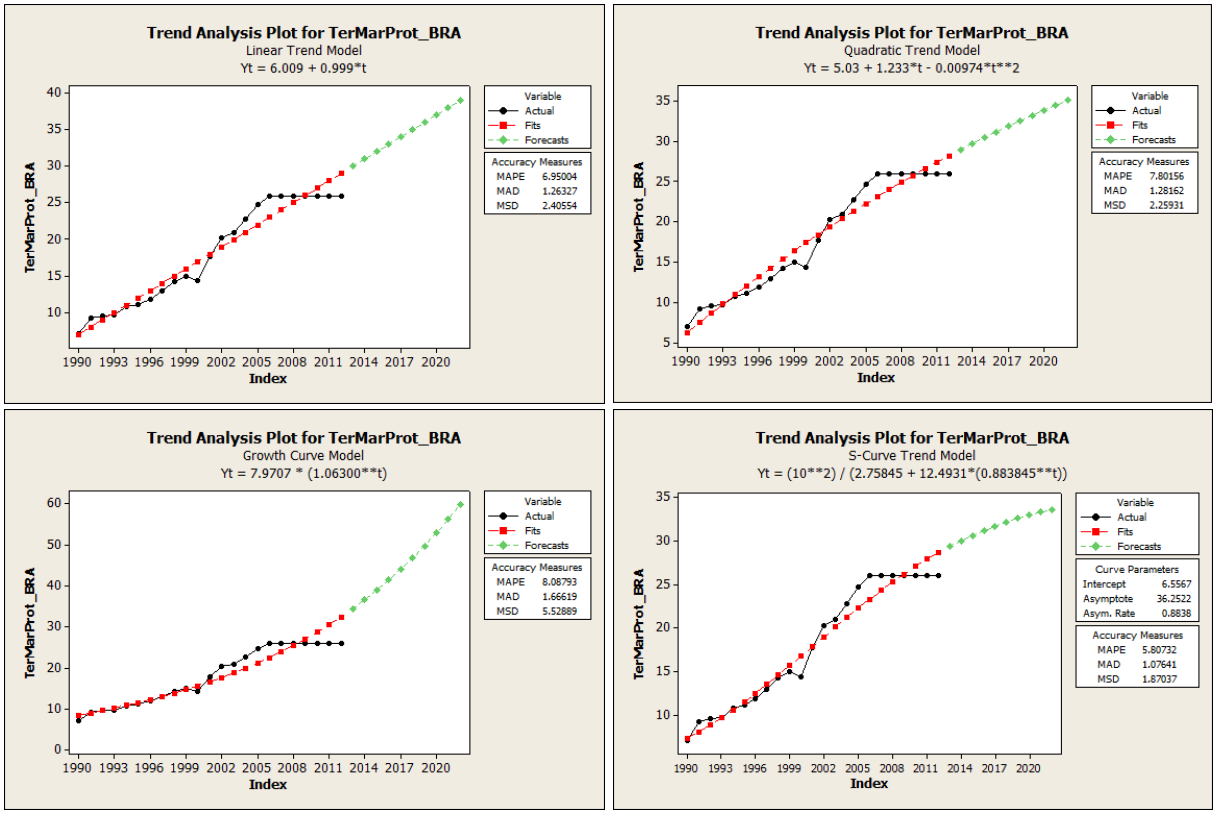
Na seção que segue, analisaremos a tendência da série de dados, buscando encontrar a função (linear, quadrática, exponencial ou s-curve) que melhor a represente. Para medir a eficiência da função ou da curva de tendência, analisaremos os erros: MAPE (*Mean Average Percentual Error)*, MAD (*Mean Absolute Deaviation)* e MSD (*Mean Standard Deviation).*

4.3.1 Variável: TerMarProt_BRA - Áreas terrestres e marinhas sob proteção, no Brasil

4.3.1.1. A análise de tendência

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

TerMarProt BRA



Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	6.95004	7.80156	8.08793	**5.80732**
MAD	1.26327	1.28162	1.66619	**1.07641**
MSD	2.40554	2.25931	5.52889	**1.87037**

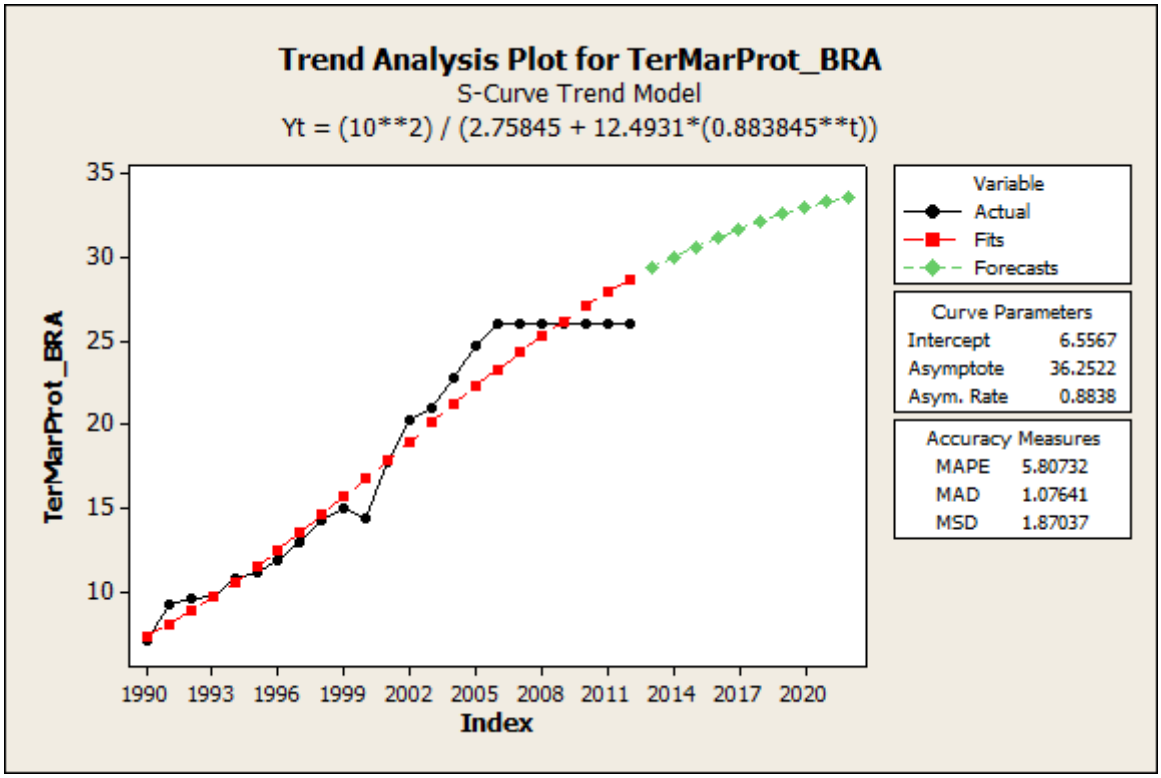
Percebemos que o modelo **S-Curve** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.1.2. As previsões

Utilizando a equação da função **S-Curve** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas terrestres e marinhas sob proteção, no Brasil.

Year	Forecast
2013	29.3802
2014	30.0416
2015	30.6516
2016	31.2117
2017	31.7240
2018	32.1911
2019	32.6155
2020	33.0000
2021	33.3475
2022	33.6608

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



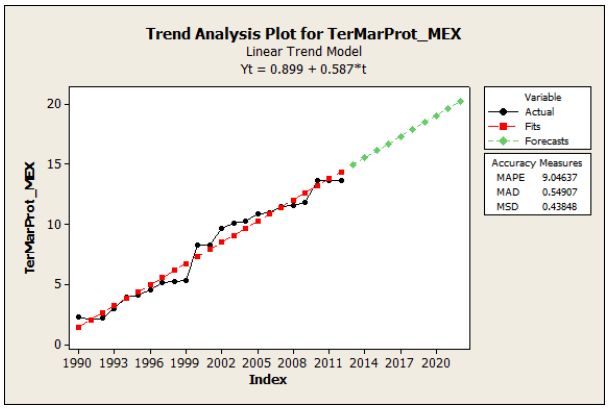
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

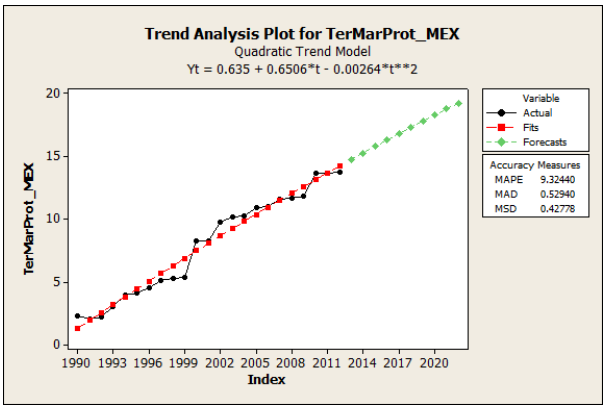
4.3.2 Variável: TerMarProt_MEX - Áreas terrestres e marinhas sob proteção, no México

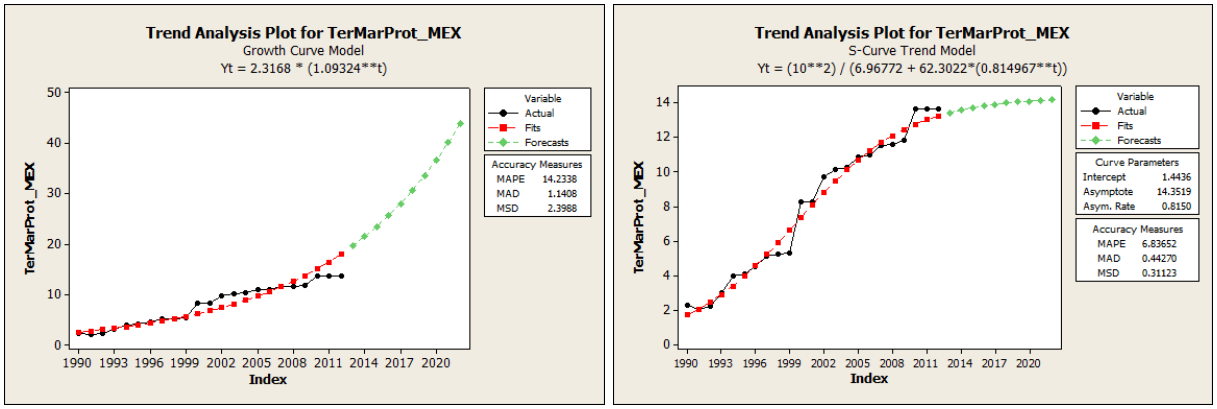
4.3.2.1. A análise de tendência

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

TerMarProt MEX







Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	9.04637	**9.32440**	14.2338	6.83652
MAD	0.54907	**0.52940**	1.1408	0.44270
MSD	0.43848	**0.42778**	2.3988	0.31123

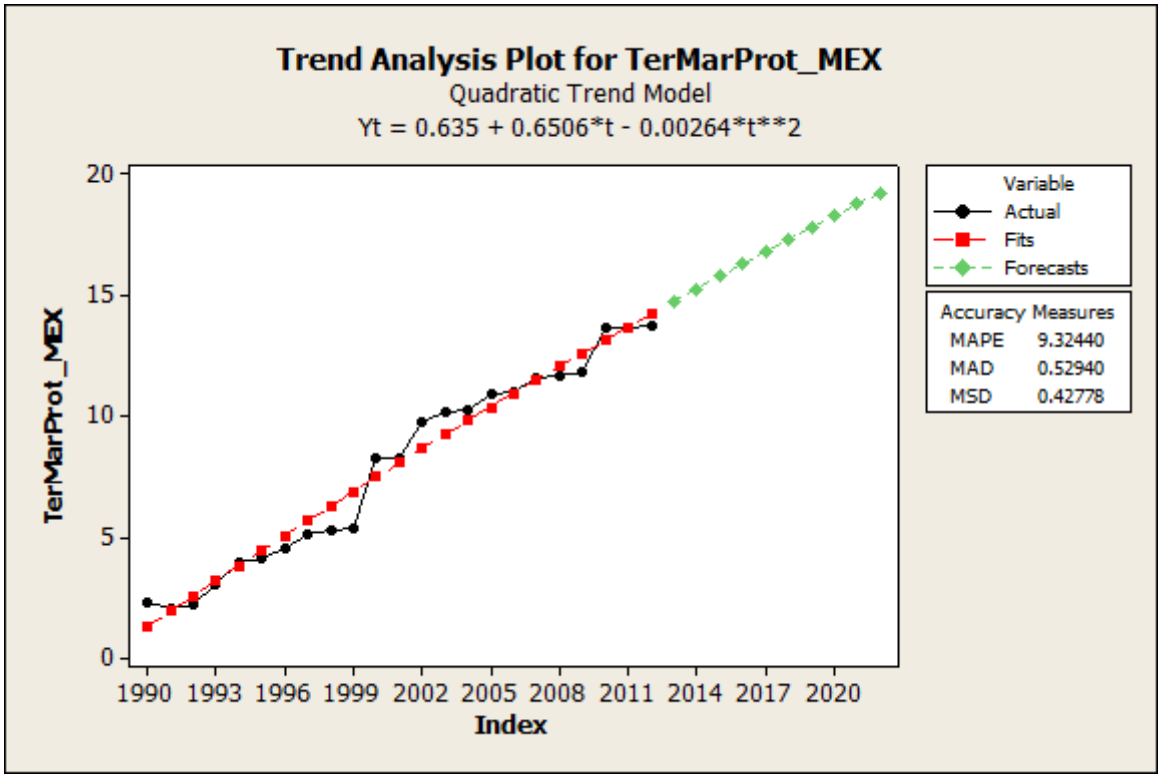
Notamos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.2.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas terrestres e marinhas sob proteção, no México.

Year	Forecast
2013	14.7313
2014	15.2528
2015	15.7690
2016	16.2799
2017	16.7855
2018	17.2858
2019	17.7809
2020	18.2707
2021	18.7553
2022	19.2346

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



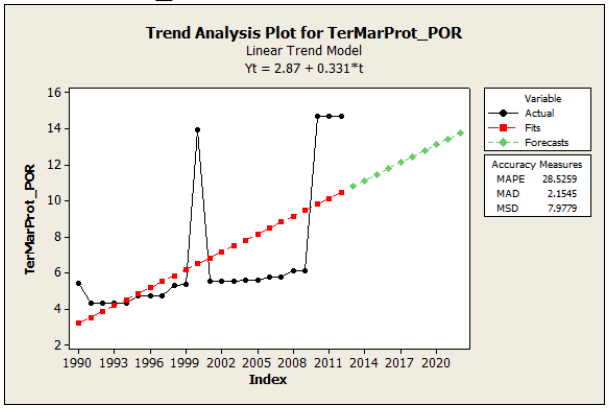
No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

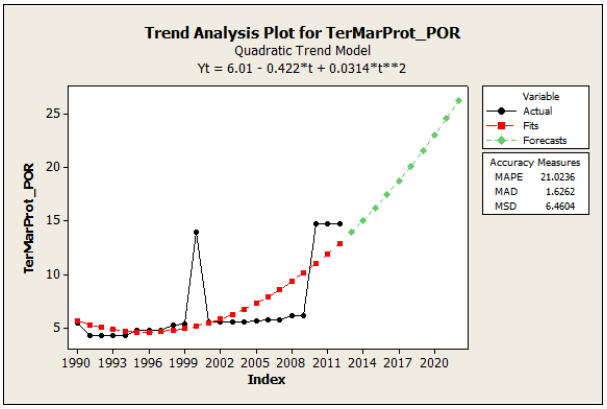
4.3.3 Variável: TerMarProt_POR - Áreas terrestres e marinhas sob proteção, em Portugal

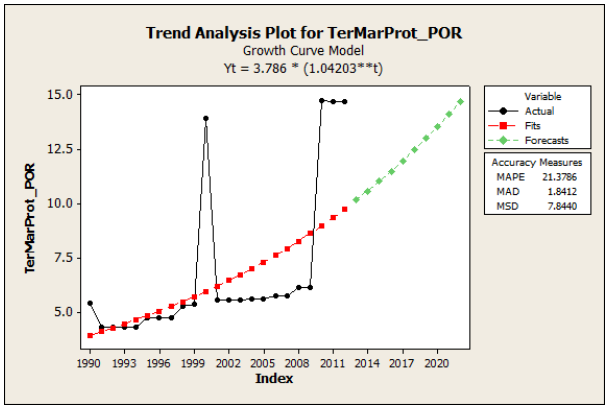
4.3.3.1. A análise de tendência

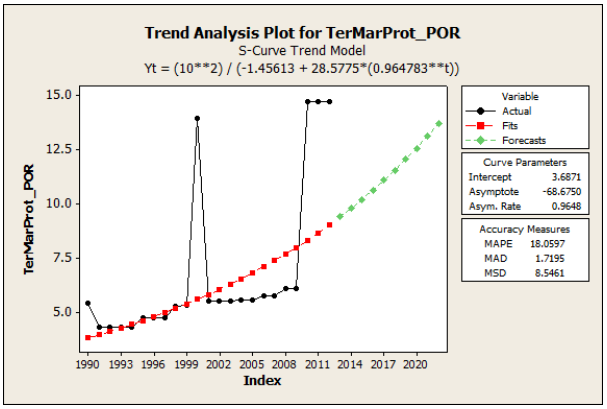
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

TerMarProt_POR









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	28.5259	21.0236	21.3786	**18.0597**
MAD	2.1545	**1.6262**	1.8412	1.7195
MSD	7.9779	**6.4604**	7.8440	8.5461

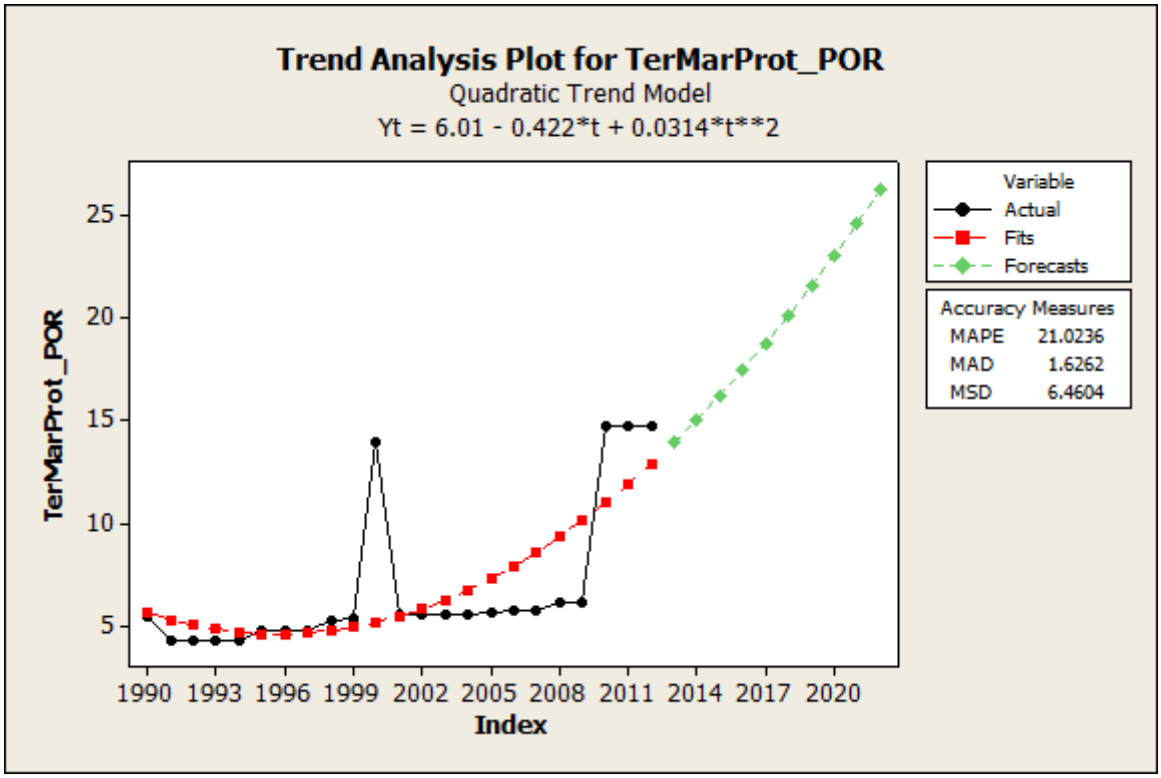
Percebemos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.3.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas terrestres e marinhas sob proteção, em Portugal.

Year	Forecast
2013	13.9527
2014	15.0684
2015	16.2469
2016	17.4882
2017	18.7922
2018	20.1591
2019	21.5887
2020	23.0811
2021	24.6363
2022	26.2543

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



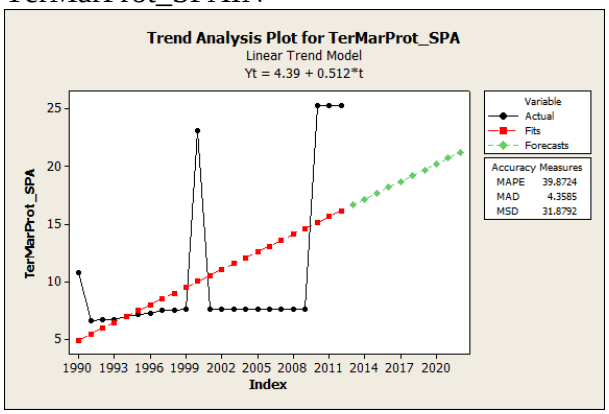
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

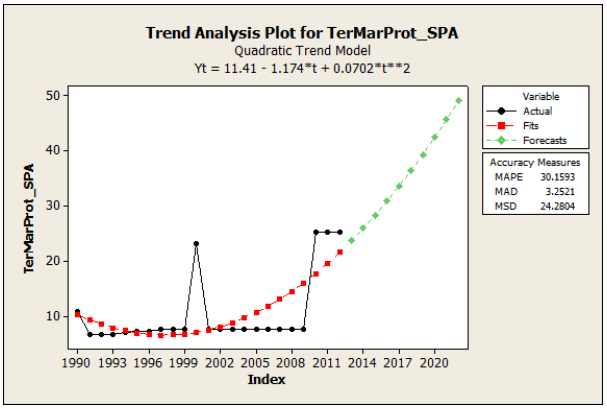
4.3.4 Variável: TerMarProt_SPA - Áreas terrestres e marinhas sob proteção, na Espanha

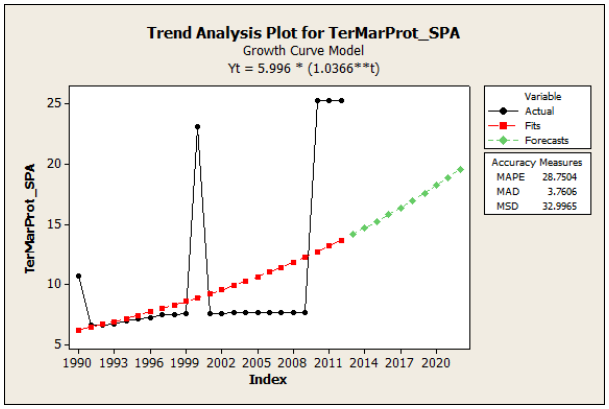
4.3.4.1. A análise de tendência

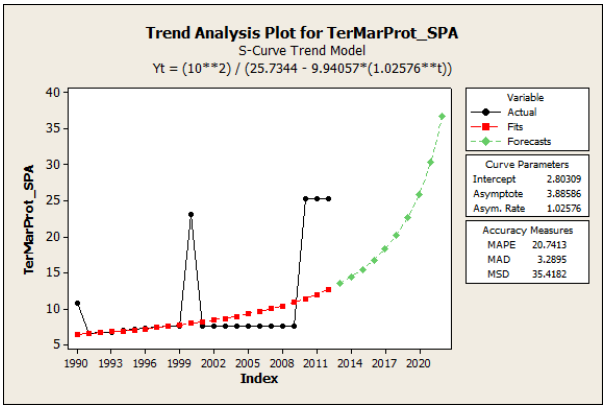
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

TerMarProt_SPAIN









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	39.8724	**30.1593**	28.7504	20.7413
MAD	4.3585	**3.2521**	3.7606	3.2895
MSD	31.8792	**24.2804**	32.9965	35.4182

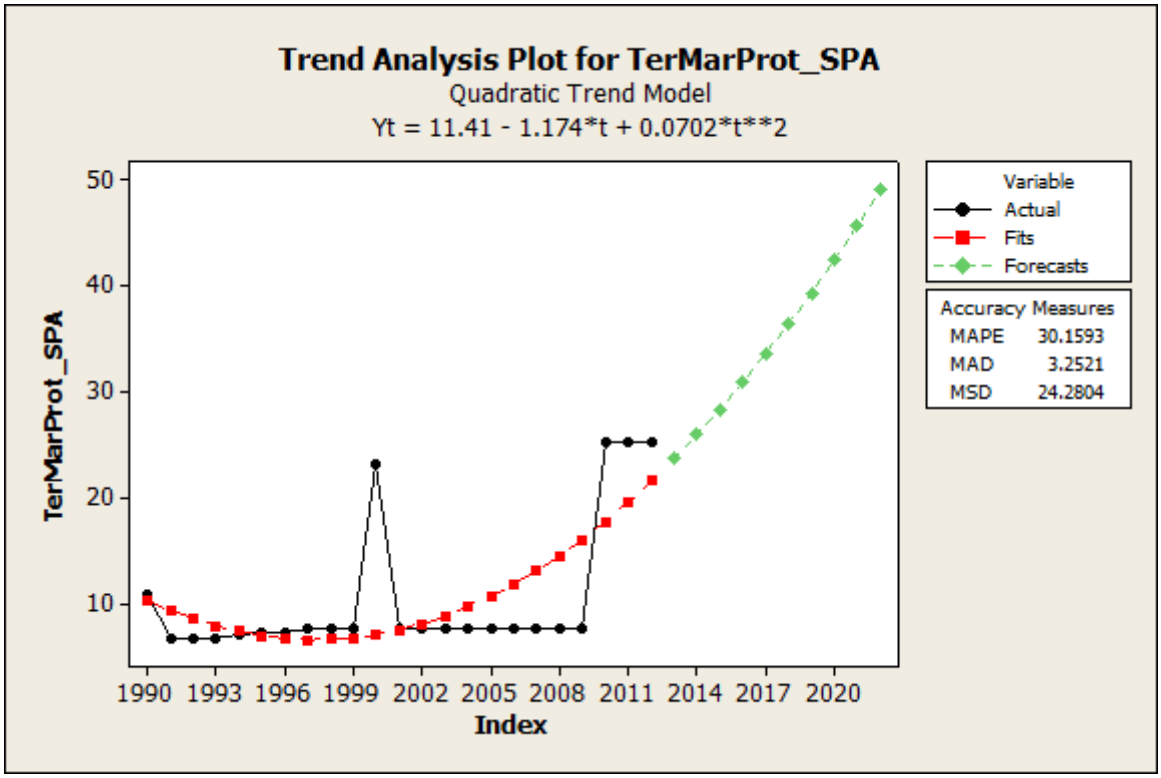
Notamos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.4.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas terrestres e marinhas sob proteção, na Espanha.

Year	Forecast
2013	23.7051
2014	25.9735
2015	28.3825
2016	30.9319
2017	33.6218
2018	36.4522
2019	39.4231
2020	42.5344
2021	45.7863
2022	49.1786

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



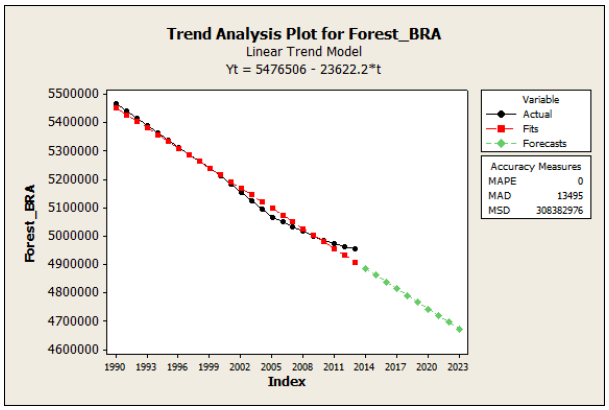
No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

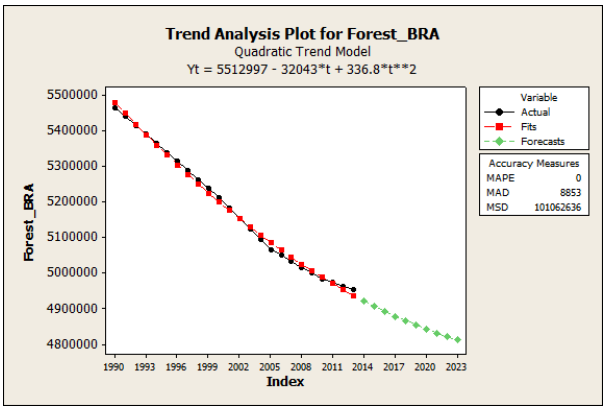
4.3.5 Variável: Forest_BRA - Áreas de florestas, no Brasil

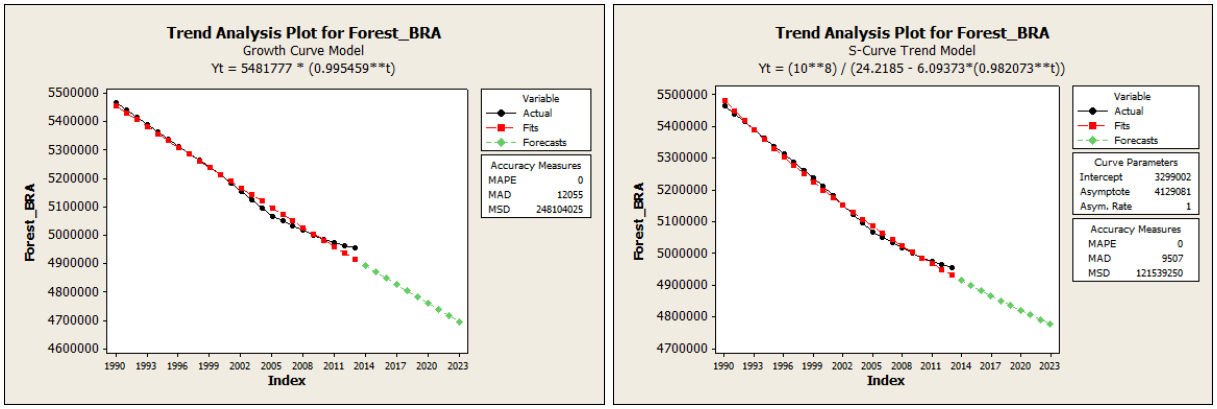
4.3.5.1. A análise de tendência

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Forest BRA







Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	0	**0**	0	0
MAD	13495	**8853**	12055	9507
MSD	308382976	**101062636**	248104025	121539250

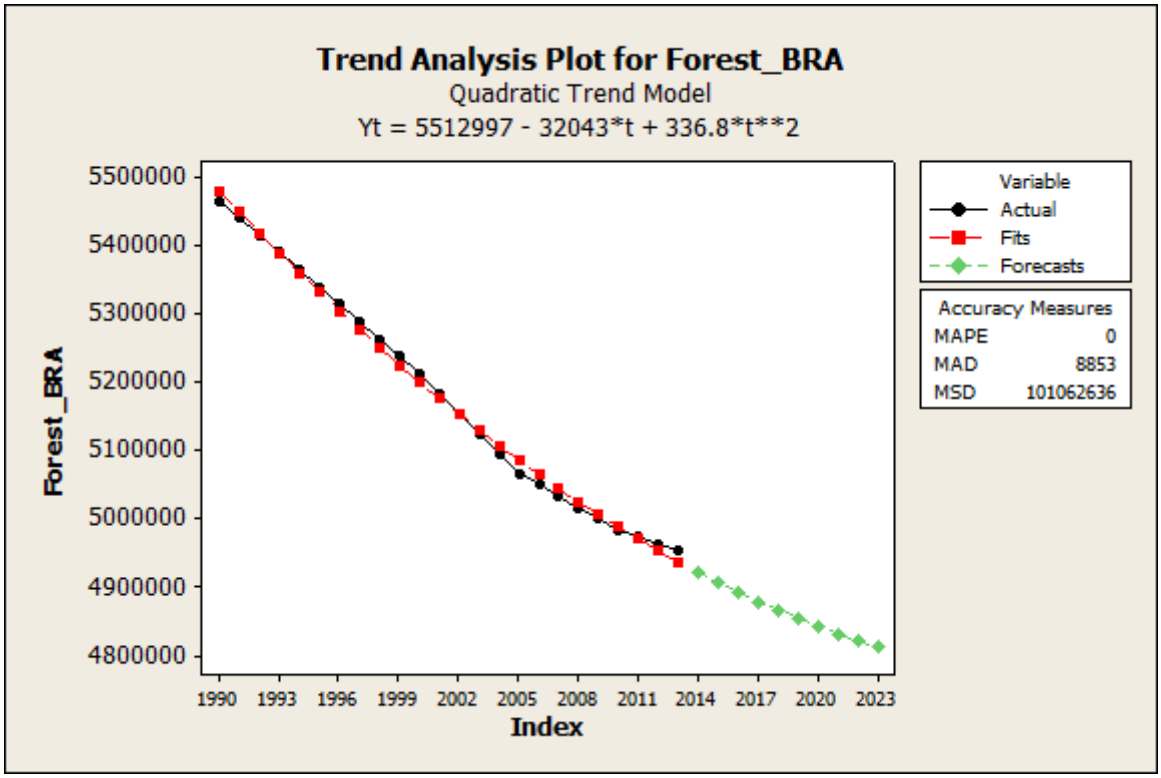
Percebemos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.5.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas de florestas sob proteção, no Brasil.

Year	Forecast
2014	4922441
2015	4907577
2016	4893386
2017	4879869
2018	4867026
2019	4854856
2020	4843360
2021	4832538
2022	4822389
2023	4812914

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



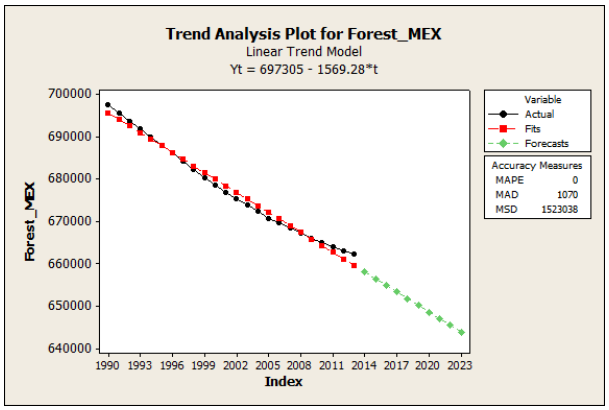
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

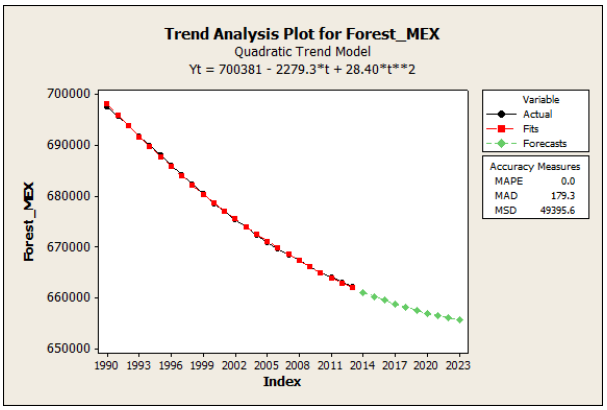
4.3.6 Variável: Forest_MEX - Áreas de florestas, no México

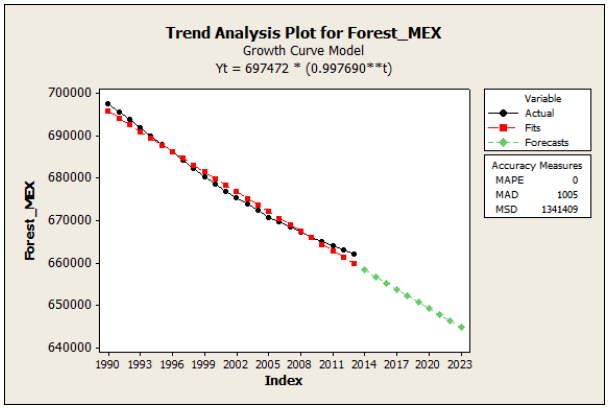
4.3.6.1. A análise de tendência

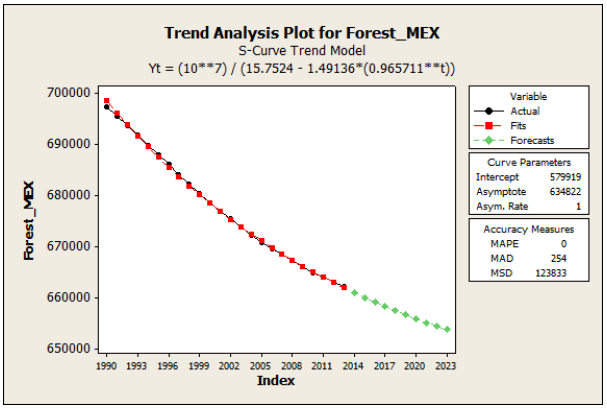
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Forest MEX









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	0	**0.0**	0	0
MAD	1070	**179.3**	1005	254
MSD	1523038	**49395.6**	1341409	123833

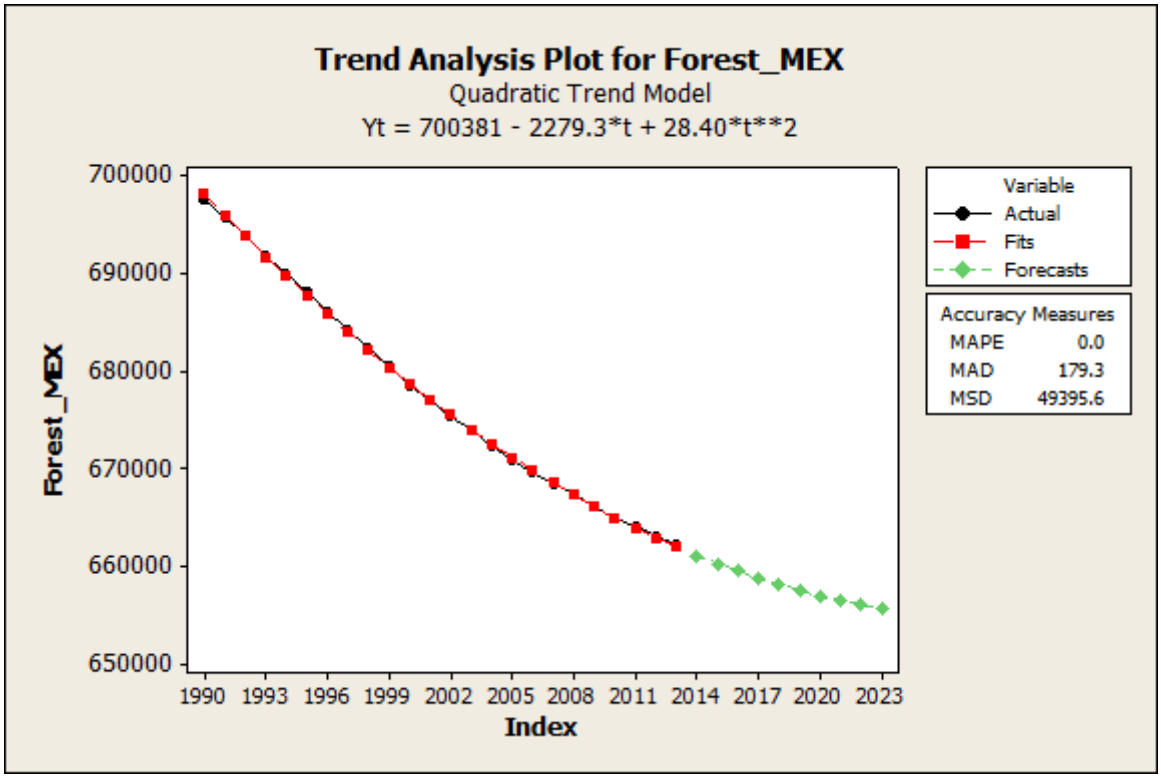
Notamos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.6.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas de florestas, no México.

Year	Forecast
2014	661149
2015	660319
2016	659544
2017	658827
2018	658167
2019	657563
2020	657016
2021	656526
2022	656092
2023	655716

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



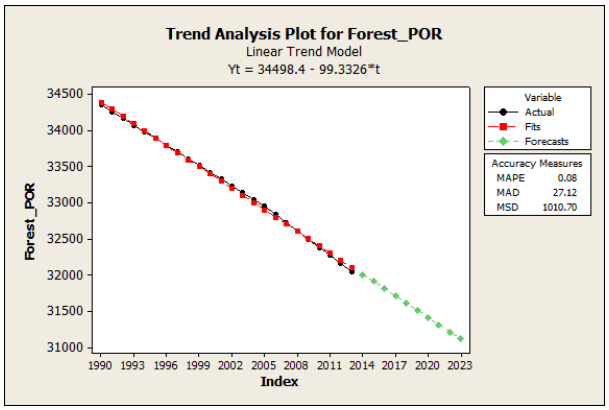
No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

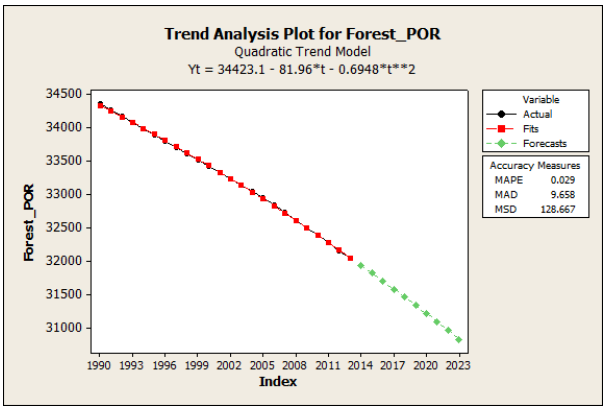
4.3.7 Variável: Forest_POR - Áreas de florestas, em Portugal

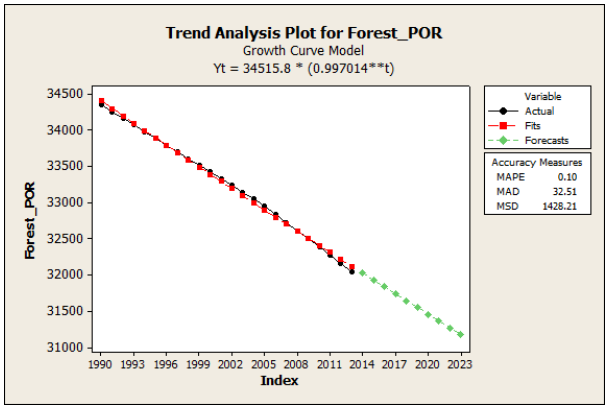
4.3.7.1. A análise de tendência

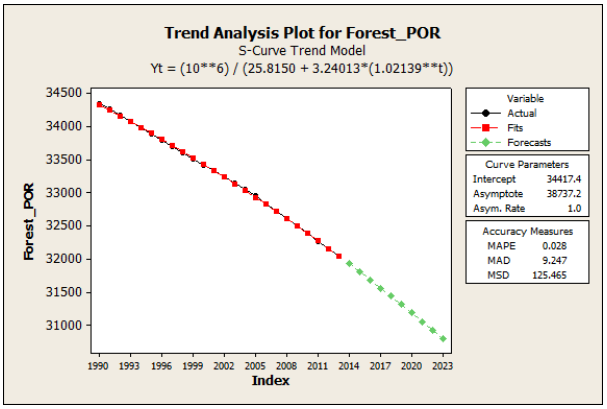
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Forest POR









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	0.08	0.029	0.10	**0.028**
MAD	27.12	9.658	32.51	**9.247**
MSD	1010.70	128.667	1428.21	**125.465**

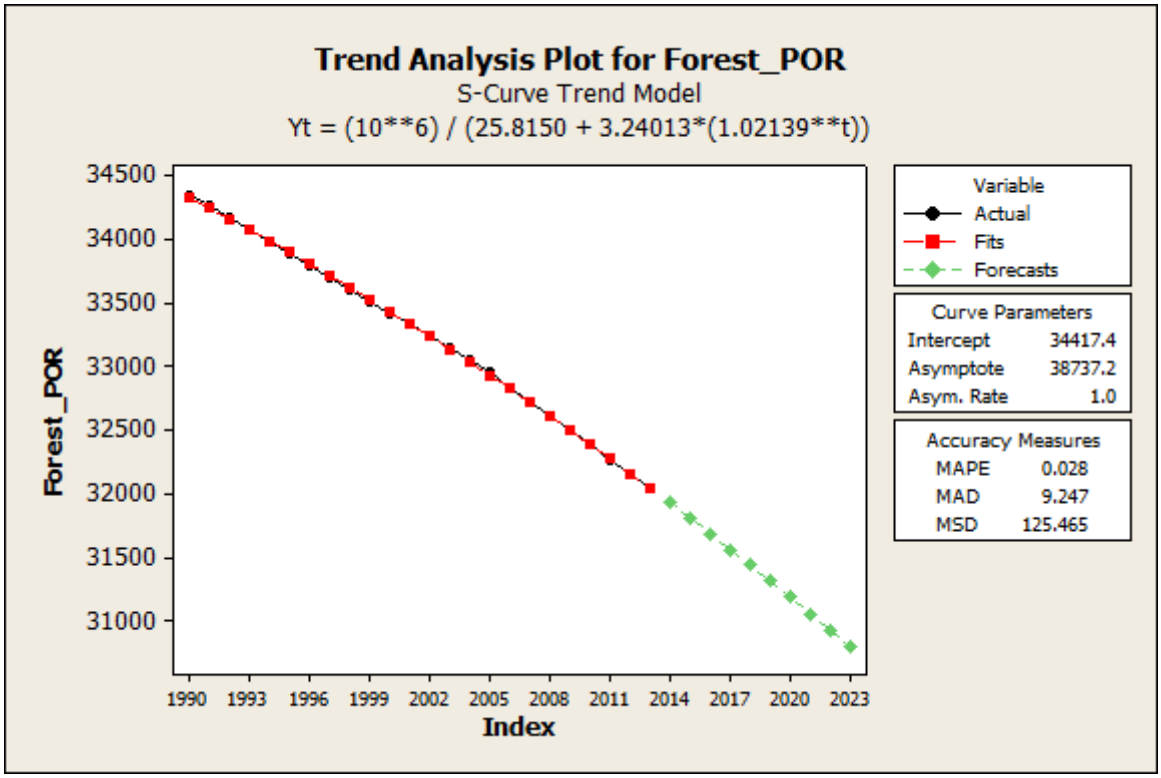
Percebemos que o modelo **S-Curve** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.7.2. As previsões

Utilizando a equação da função **S-Curve** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas de florestas sob proteção, em Portugal.

Year	Forecast
2014	31933.2
2015	31813.6
2016	31692.4
2017	31569.6
2018	31445.1
2019	31319.0
2020	31191.2
2021	31061.8
2022	30930.6
2023	30797.9

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



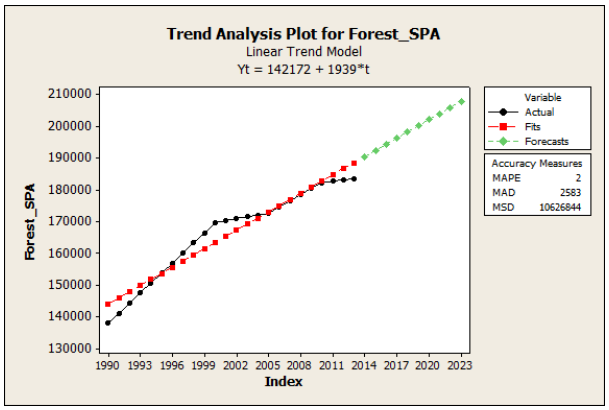
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

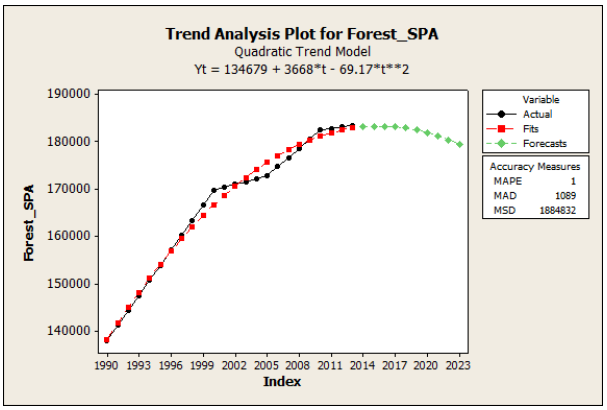
4.3.8 Variável: Forest_SPA - Áreas de florestas, na Espanha

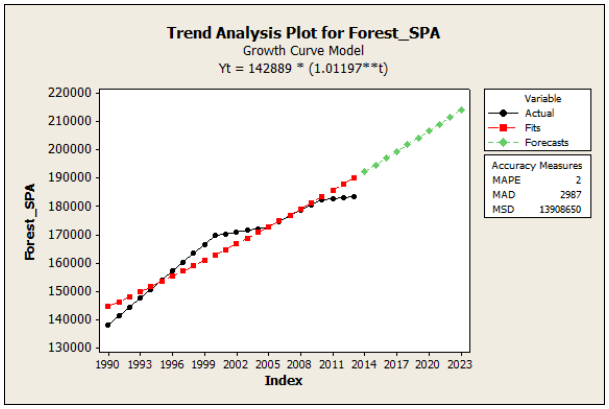
4.3.8.1. A análise de tendência

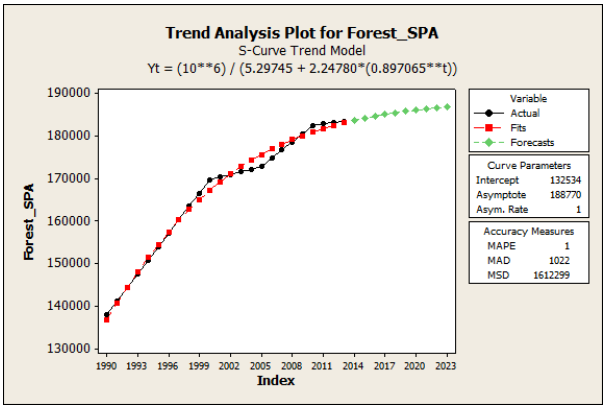
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Forest SPAIN









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	2	1	2	**1**
MAD	2583	1089	2987	**1022**
MSD	10626844	1884832	13908650	**1612299**

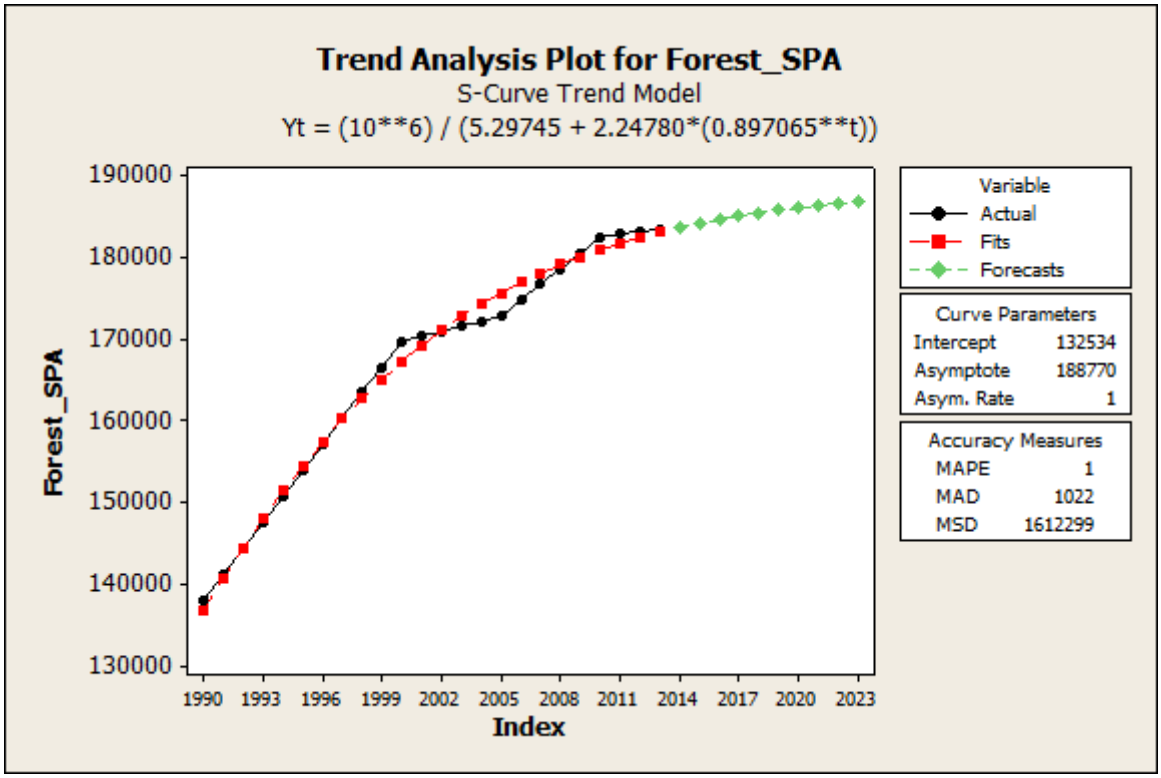
Notamos que o modelo **S-Curve** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.8.2. As previsões

Utilizando a equação da função **S-Curve** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação às áreas de florestas, na Espanha.

Year	Forecast
2014	183615
2015	184133
2016	184600
2017	185021
2018	185400
2019	185741
2020	186048
2021	186325
2022	186574
2023	186797

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



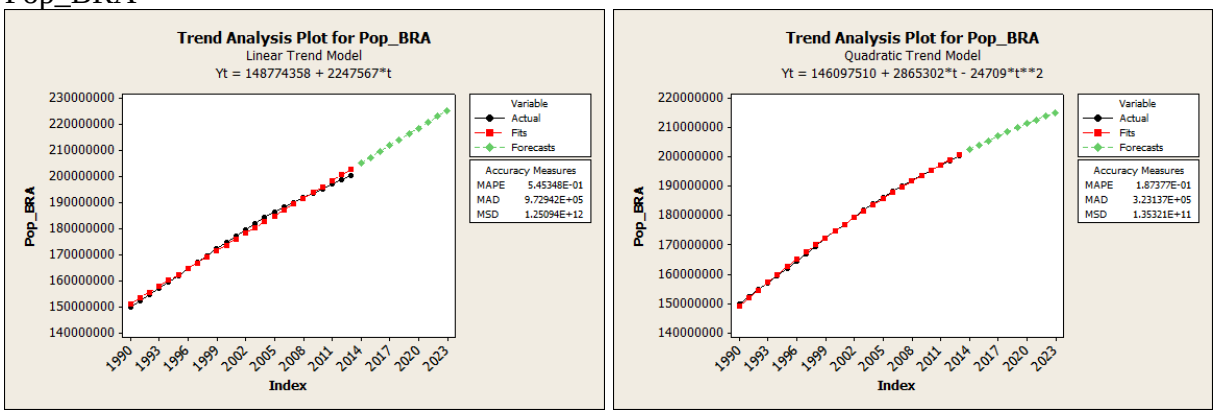
No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

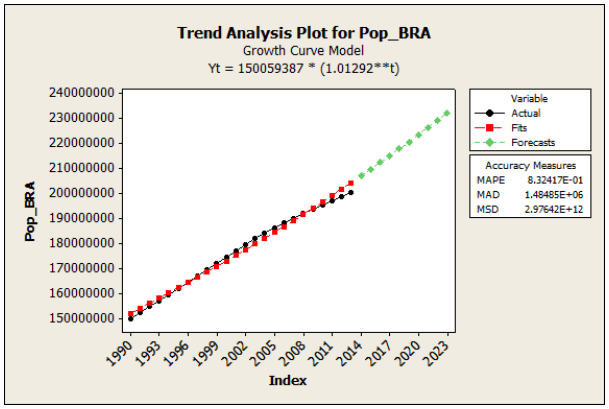
4.3.9 Variável: Pop_POR – População Total no Brasil

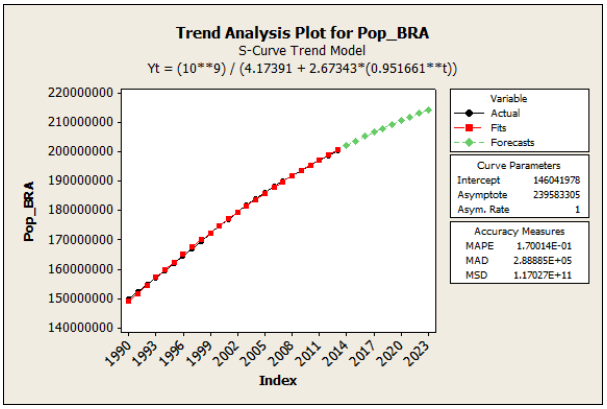
4.3.9.1. A análise de tendência

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Pop_BRA







Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	5.45348E-01	1.87377E-01	8.32417E-01	**1.70014E-01**
MAD	9.72942E+05	3.23137E+05	1.48485E+06	**2.88885E+05**
MSD	1.25094E+12	1.35321E+11	2.97642E+12	**1.17027E+11**

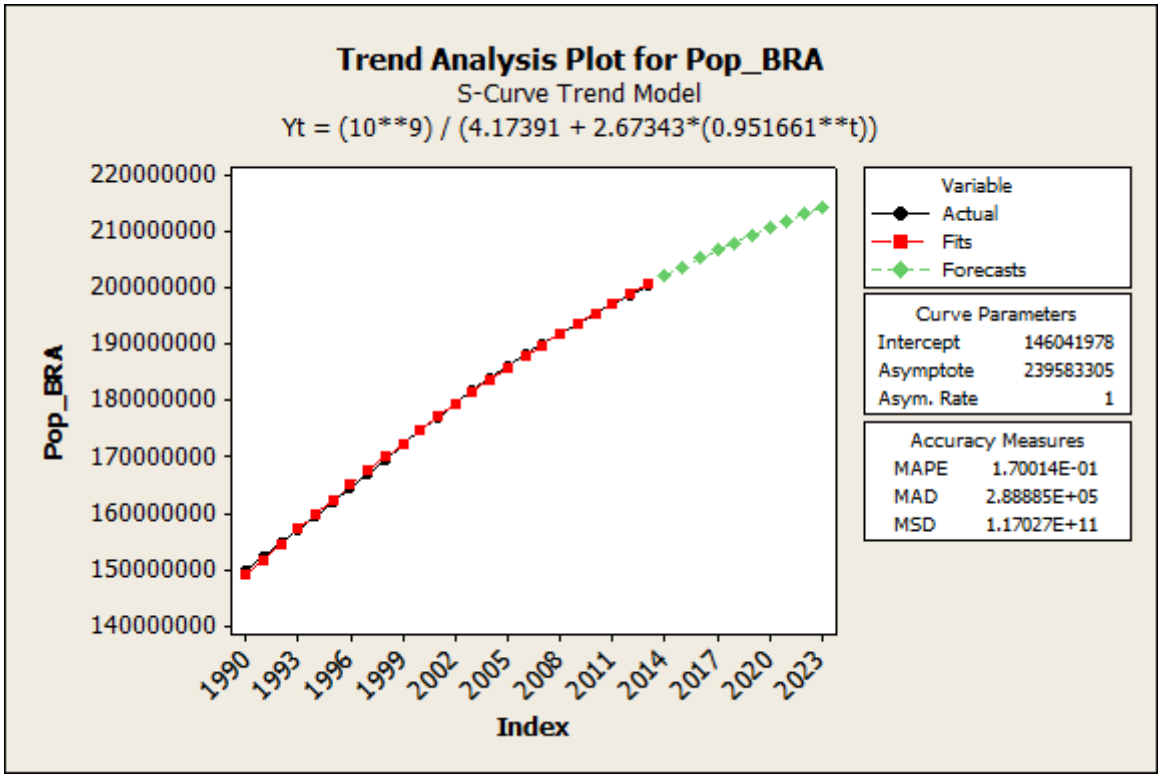
Percebemos que o modelo **S-Curve** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.9.2. As previsões

Utilizando a equação da função **S-Curve** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação população total, no Brasil.

Year	Forecast
2014	202077154
2015	203617994
2016	205106333
2017	206543075
2018	207929185
2019	209265680
2020	210553623
2021	211794119
2022	212988301
2023	214137331

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



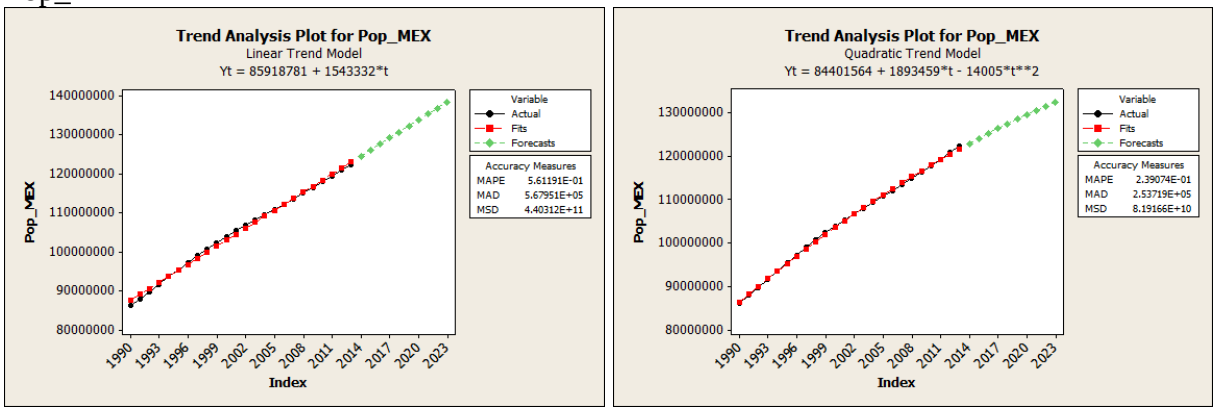
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

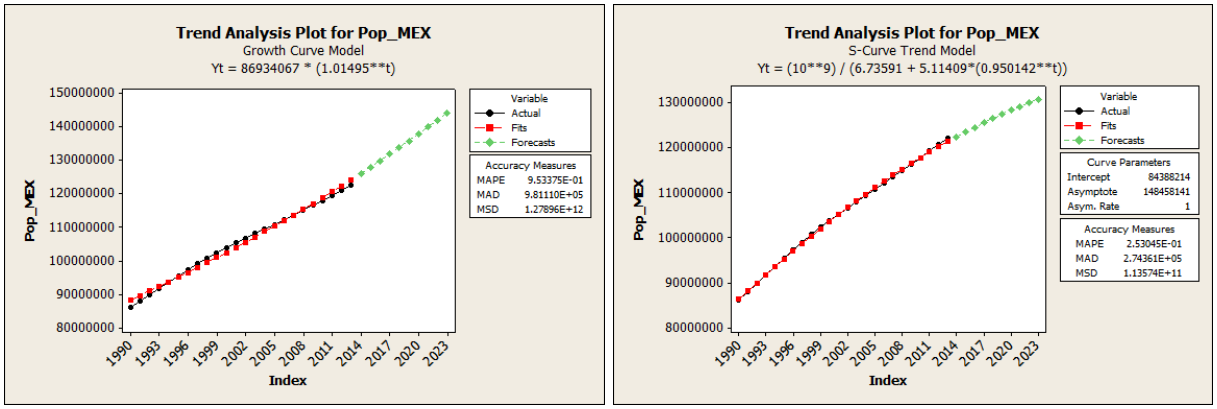
4.3.10 Variável: Pop_MEX – População Total no México

4.3.10.1. A análise de tendência

Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Pop_MEX





Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	5.61191E-01	**2.39074E-01**	9.53375E-01	2.53045E-01
MAD	5.67951E+05	**2.53719E+05**	9.81110E+05	2.74361E+05
MSD	4.40312E+11	**8.19166E+10**	1.27896E+12	1.13574E+11

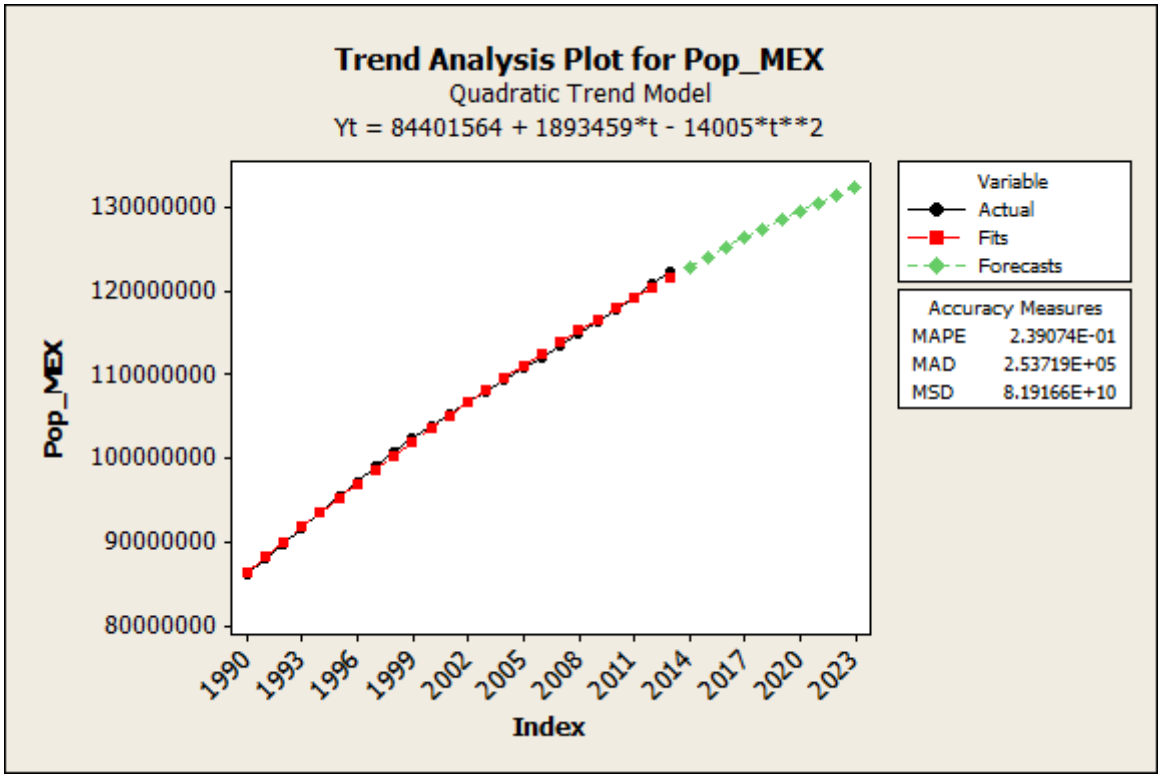
Notamos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.10.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação a população total no México.

Year	Forecast
2014	122984874
2015	124164074
2016	125315265
2017	126438445
2018	127533615
2019	128600774
2020	129639924
2021	130651063
2022	131634193
2023	132589312

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



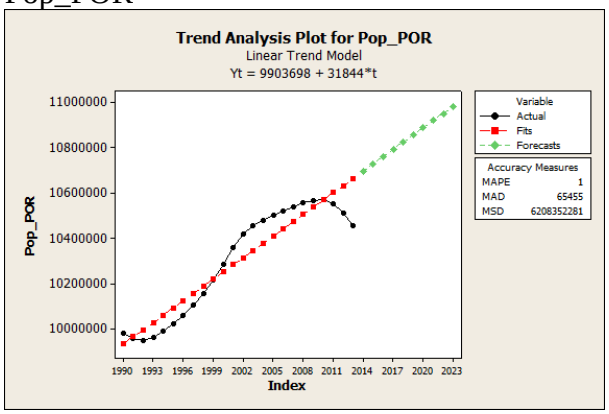
No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

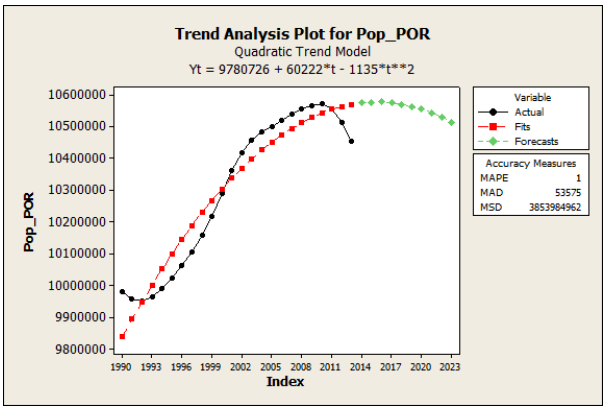
4.3.11 Variável: Pop_POR – População Total em Portugal

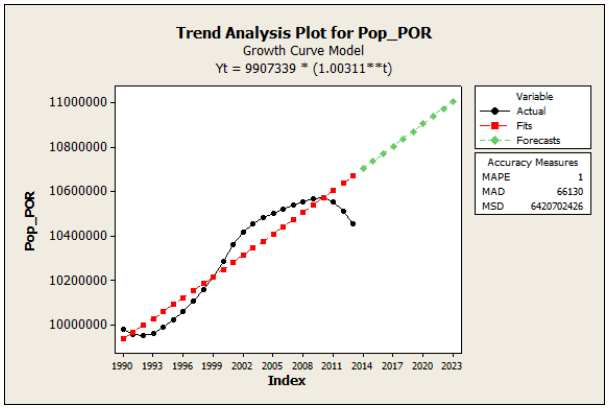
4.3.11.1. A análise de tendência

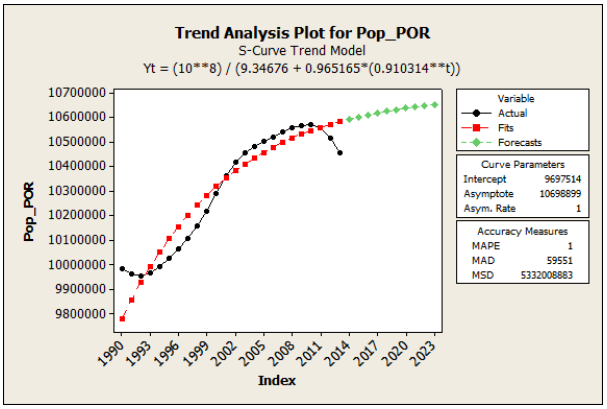
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Pop_POR









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	1	**1**	1	1
MAD	65455	**53575**	66130	59551
MSD	6208352281	**3853984962**	6420702426	5332008883

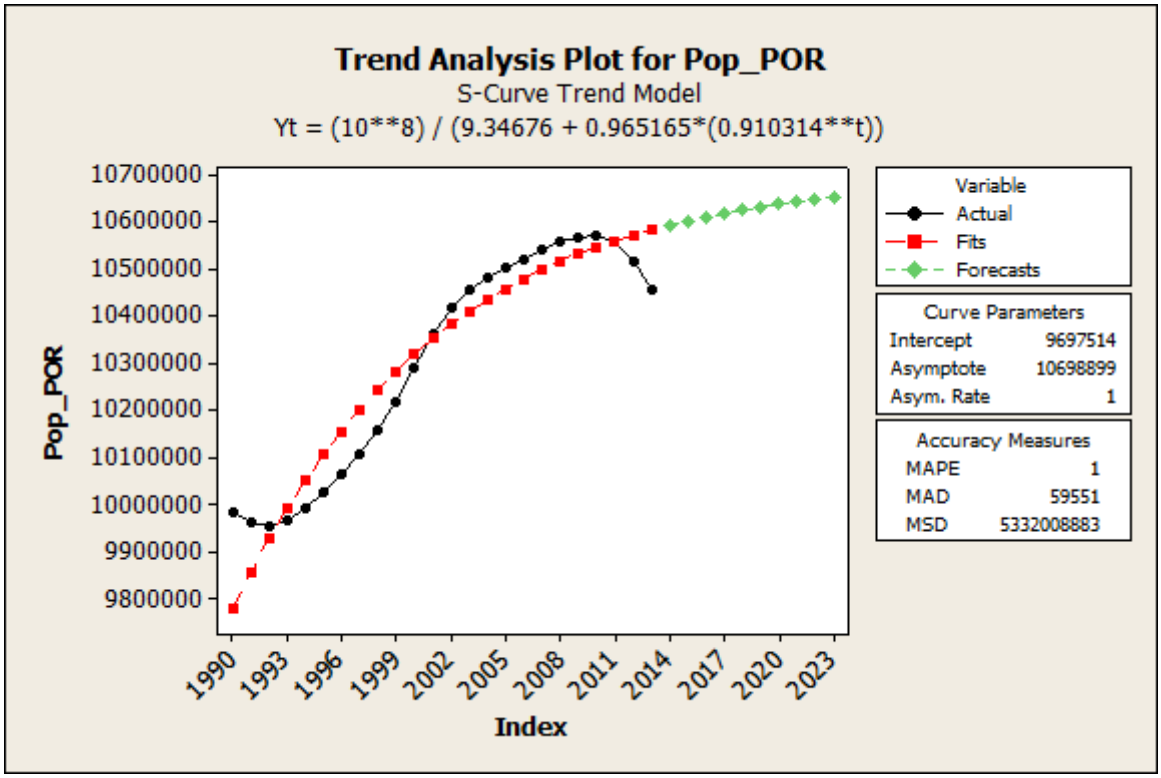
Percebemos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.11.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação população total, em Portugal.

Year	Forecast
2014	10594475
2015	10603757
2016	10612221
2017	10619937
2018	10626971
2019	10633383
2020	10639226
2021	10644551
2022	10649402
2023	10653823

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



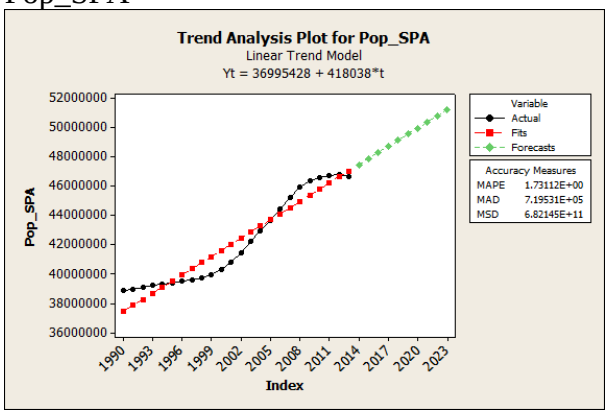
Ressaltamos que trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

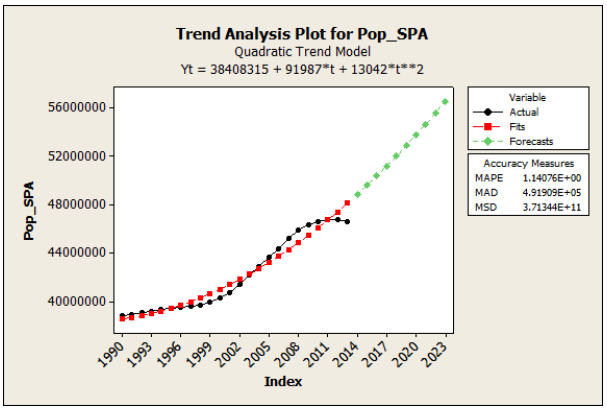
4.3.12 Variável: Pop_SPA – População Total na Espanha

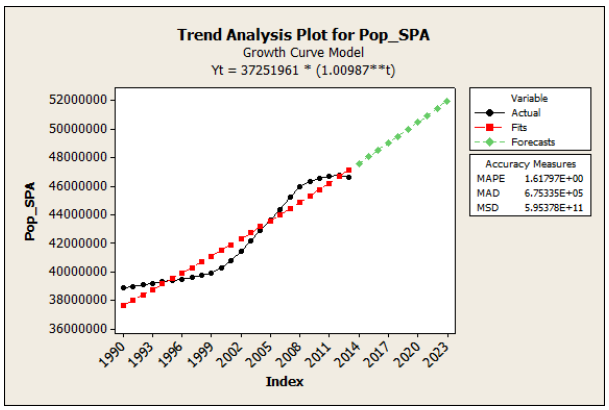
4.3.12.1. A análise de tendência

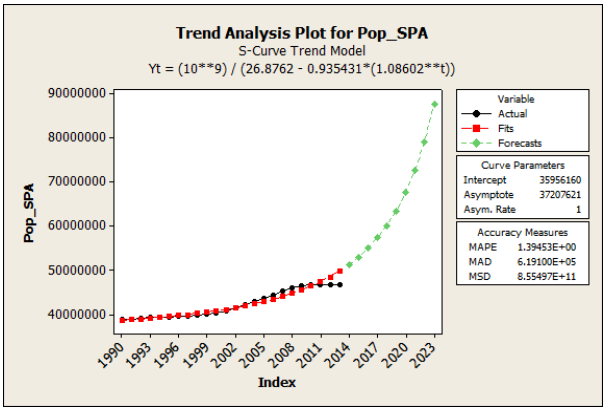
Segue abaixo gráficos incluindo a função, a linha de tendência que representa cada função e os erros mencionados acima.

Pop_SPA









Para podermos visualizar qual a melhor função para representar a série de dados acima, podemos analisar a tabela abaixo, que compara os três tipos de erros para cada uma das quatro funções:

Estatística	Linear	Quadrática	Exponencial	S-Curve
MAPE	1.73112E+00	**1.14076E+00**	1.61797E+00	1.39453E+00
MAD	7.19531E+05	**4.91909E+05**	6.75335E+05	6.19100E+05
MSD	6.82145E+11	**3.71344E+11**	5.95378E+11	8.55497E+11

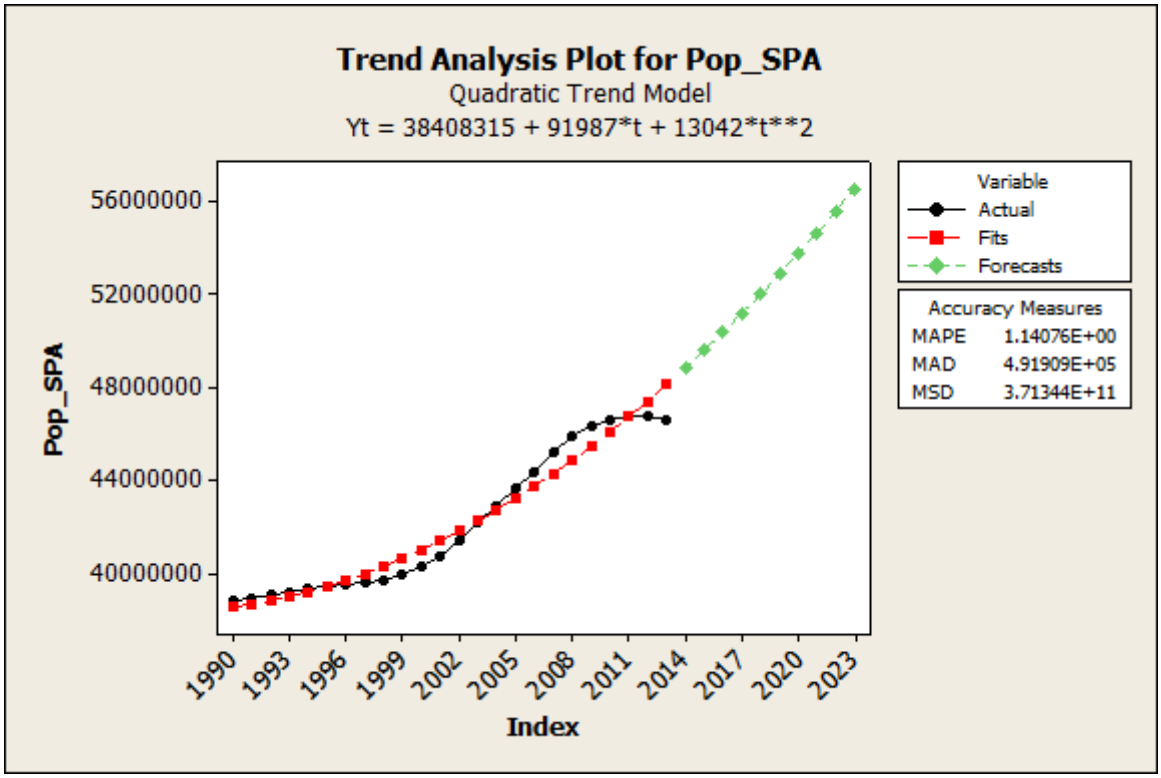
Notamos que o modelo **Quadrática** é o que melhor se adapta à nossa série de dados. Assim, esta será a função utilizada para as projeções (*forecasts*).

4.3.12.2. As previsões

Utilizando a equação da função **Quadrática** apresentada no gráfico acima, temos os seguintes valores previstos (*forecasts*) para os próximos 10 anos em relação a população total na Espanha.

Year	Forecast
2014	48859268
2015	49616399
2016	50399615
2017	51208914
2018	52044297
2019	52905764
2020	53793316
2021	54706951
2022	55646671
2023	56612474

Os valores acima podem ser melhor ilustrados de forma gráfica, conforme abaixo:



No gráfico acima, trata-se de uma previsão efetuada somente com base nos valores históricos da própria variável. Certamente esta metodologia apresenta deficiências ao assumir que o comportamento passado se repetirá. Do mesmo modo, talvez a utilização de regressões múltiplas poderiam ser mais eficientes neste trabalho de previsão.

5. COMPARAÇÕES

Nesta sessão será efetuada análise comparativa de médias e intervalos de confiança de dados dimensionadores. O principal propósito é comparar como os dados se comportarão quando agrupados em 3 principais regiões, os países Ibero-Americanos (AIBER), países de economias avançadas (AVECO) e os demais países constantes na base de dados deste estudo.

5.1. Categorização das variáveis – Siglas e Nomenclaturas

O objetivo é comparar a região Ibero-Americana (AIBER) com as demais regiões do mundo, para cada variável. Assim, cada variável será categorizada por região e receberá as siglas e nomenclaturas conforme segue:

Quadro - Siglas, Nomenclaturas e Quantidades de Países por Região (8)

Sigla da Região	Nomenclatura da Região	# Países
AIBER	Ibero-Americanos	21
AVECO	Economias Avançadas	27
CARLA	América Latina e o Caribe	3
CWEAS	Comunidade dos Estados Independentes	11
ESEAS	Emergentes e em Desenvolvimento da Ásia	12
EMDEU	Emergentes e em Desenvolvimento da Europa	12
NSWUA	Oriente Médio, Norte da África e Paquistão	14
SUSAF	África Subsariana	32

Total de países neste estudo: 132

Para que houvesse destaque visual e facilidade em identificar os agrupamentos das regiões, optou-se em colorir as células acima da seguinte forma:
-em azul, a região AIBER, contendo 21 países Ibero-Americanos;
-em verde, a região AVECO, que denota os 27 países considerados economias avançadas
-em vermelho, figuram 84 países das demais 6 regiões, que serão categorizados como OTHERS nas análises abaixo.
Obs: Conforme orientação recebida em sala, o país Cuba estava como CARLA (Caribbean and Latin America) e passou a figurar como AIBER, para fins deste estudo.

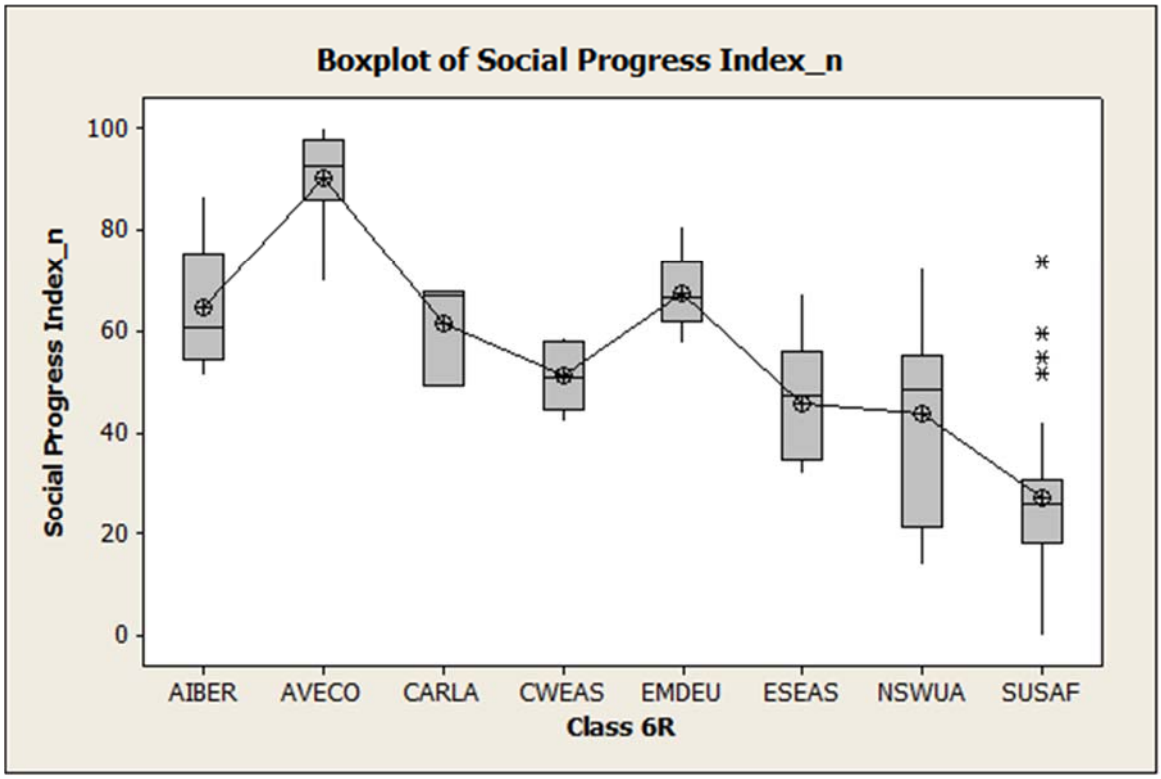
5.3. Comparação das regiões

Para esta análise comparativa será utilizada o modelo ANOVA. Esta análise de variância testa a hipótese de que as médias de duas ou mais populações são iguais.

One-way ANOVA: Social Progress Index_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  65372  9339  58.02  0.000
Error     124  19961   161
Total     131  85333

S = 12.69   R-Sq = 76.61%   R-Sq(adj) = 75.29%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  21  64.41  11.95                   (--*--)
AVECO  27  90.30   8.70                                (-*--)
CARLA   3  61.42  10.46              (-------*------)
CWEAS  11  51.27   6.12             (---*--)
EMDEU  12  67.34   7.04                      (---*--)
ESEAS  12  45.57  11.72          (---*--)
NSWUA  14  43.75  18.74          (--*--)
SUSAF  32  27.04  15.99  (--*-)
                          ---------+---------+---------+---------+
                                  40        60        80       100

Pooled StDev = 12.69
```

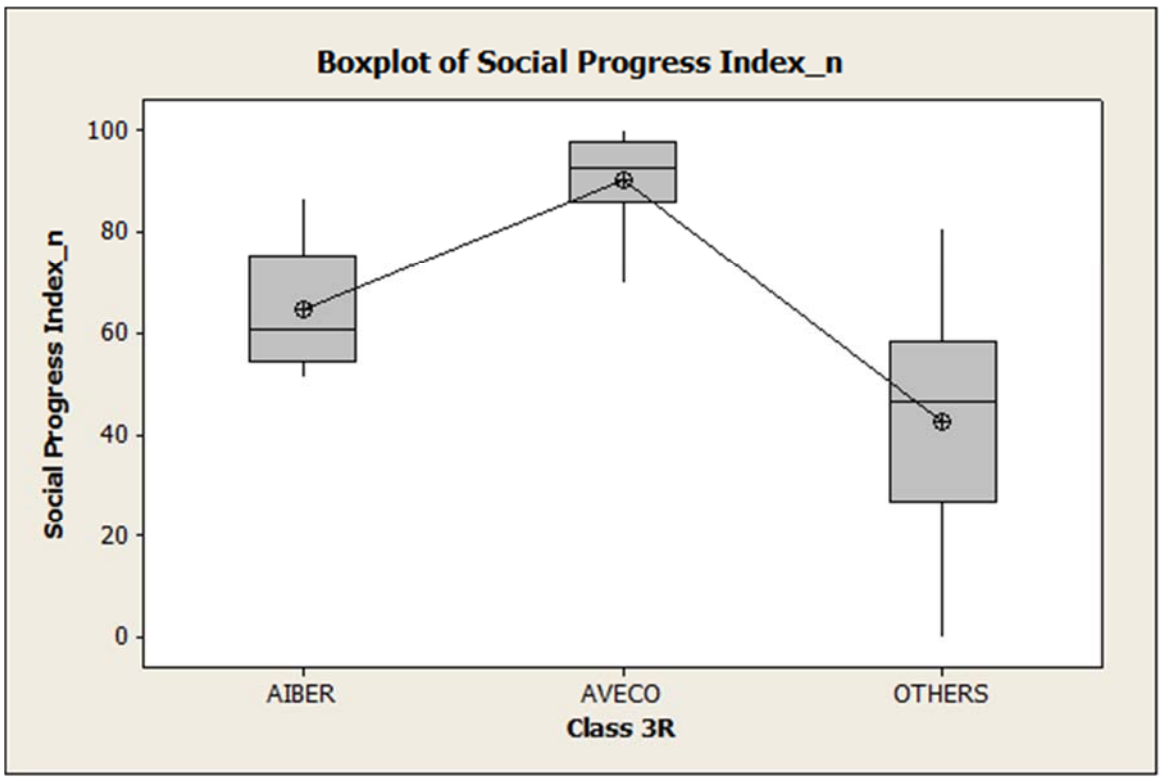
Boxplot of Social Progress Index_n



One-way ANOVA: Social Progress Index_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  48262  24131  83.97  0.000
Error     129  37071    287
Total     131  85333

S = 16.95   R-Sq = 56.56%   R-Sq(adj) = 55.88%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  64.41  11.95              (----*----)
AVECO   27  90.30   8.70                                (---*----)
OTHERS  84  42.63  19.71  (-*--)
                          ----+---------+---------+---------+-----
                             45        60        75        90

Pooled StDev = 16.95
```

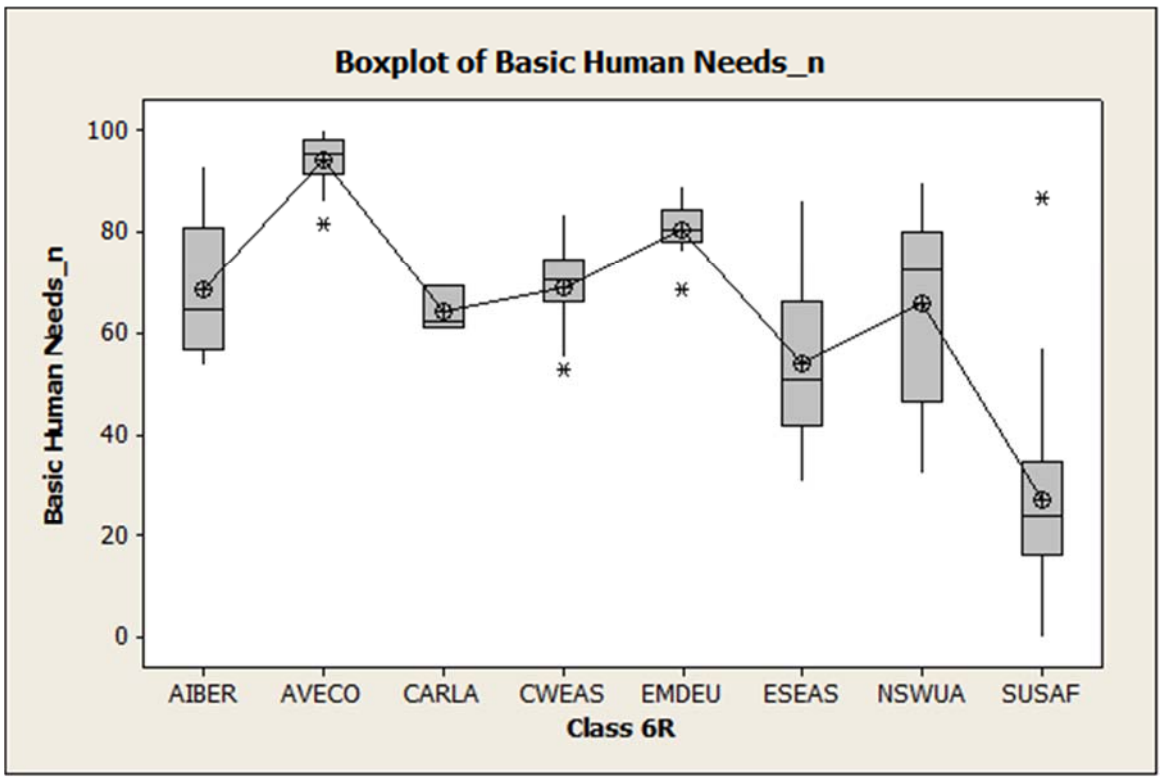
Boxplot of Social Progress Index_n



One-way ANOVA: Basic Human Needs_n versus Class 6R

```
Source     DF     SS     MS      F      P
Class 6R    7  72694  10385  58.47  0.000
Error     124  22025    178
Total     131  94719

S = 13.33   R-Sq = 76.75%   R-Sq(adj) = 75.43%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  21  68.42  13.13                    (--*--)
AVECO  27  94.10   4.73                                  (-*--)
CARLA   3  64.16   4.65              (-------*-------)
CWEAS  11  68.96   8.76                    (--*---)
EMDEU  12  80.53   5.32                         (---*---)
ESEAS  12  53.80  15.57             (---*---)
NSWUA  14  65.72  19.70                   (---*--)
SUSAF  32  27.31  17.19  (--*-)
                          ---------+---------+---------+---------+
                                  40        60        80       100

Pooled StDev = 13.33
```

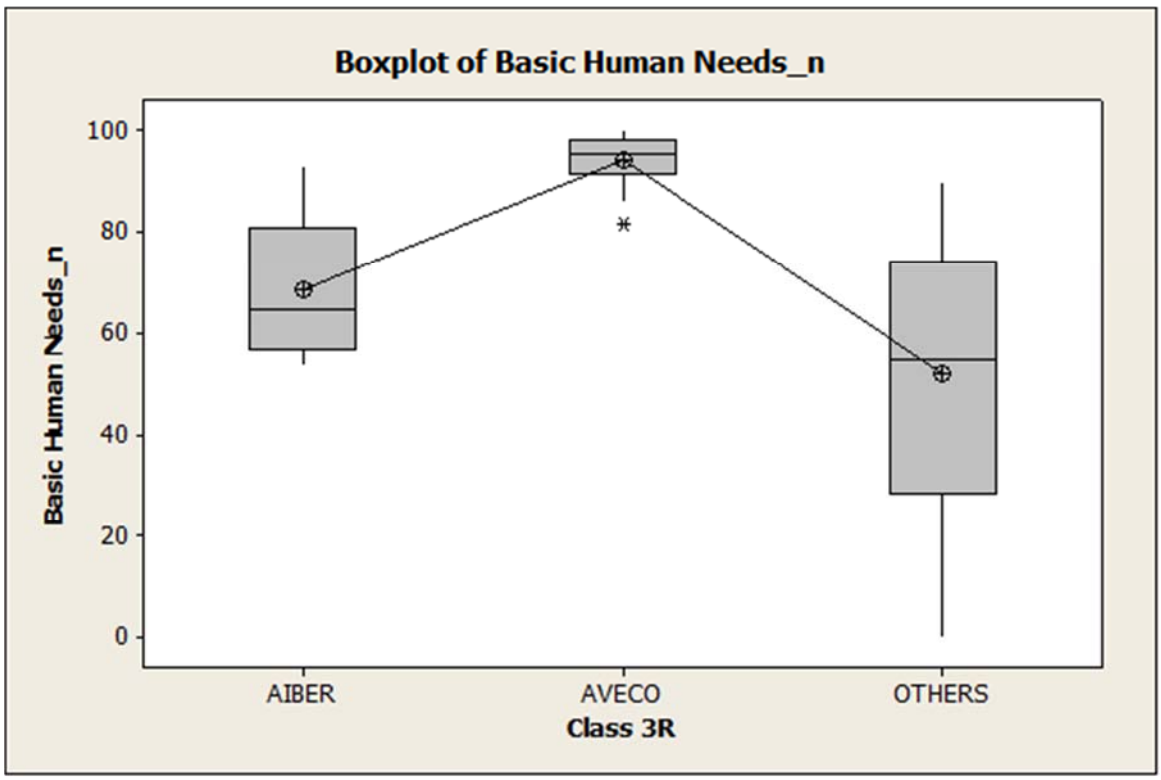
Boxplot of Basic Human Needs_n



One-way ANOVA: Basic Human Needs_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  37131  18566  41.59  0.000
Error     129  57588    446
Total     131  94719

S = 21.13   R-Sq = 39.20%   R-Sq(adj) = 38.26%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  68.42  13.13          (-----*-----)
AVECO   27  94.10   4.73                      (-----*----)
OTHERS  84  51.87  25.40  (--*--)
                          --------+---------+---------+---------+-
                                 60        75        90       105

Pooled StDev = 21.13
```

Boxplot of Basic Human Needs_n

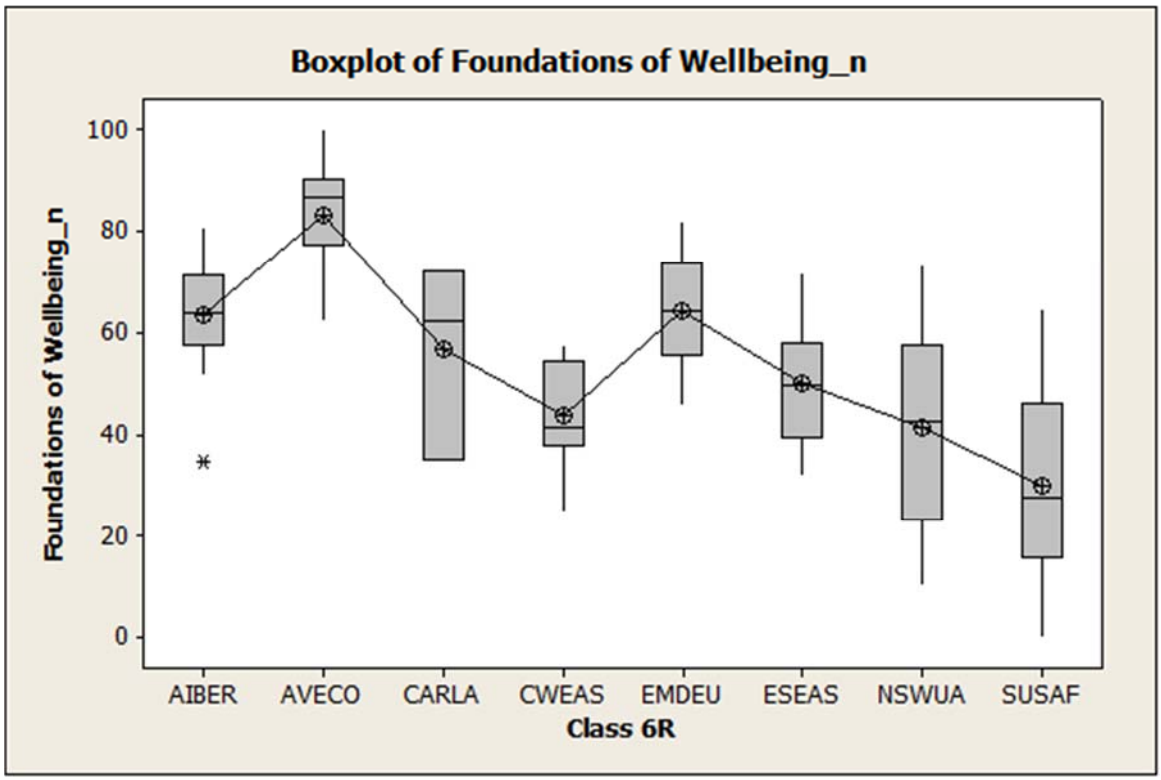


Obs: O país Latvia (Letônia) figura como *outlier* no agrupamento de países da região AVECO, para esta variável.

One-way ANOVA: Foundations of Wellbeing_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  48501  6929  35.27  0.000
Error     124  24358   196
Total     131  72859

S = 14.02   R-Sq = 66.57%   R-Sq(adj) = 64.68%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ----+---------+---------+---------+-----
AIBER  21  63.48  10.26                    (---*--)
AVECO  27  83.29   9.51                              (--*--)
CARLA   3  56.54  19.20          (---------*---------)
CWEAS  11  43.61  11.20       (----*----)
EMDEU  12  64.18  11.30                 (----*----)
ESEAS  12  49.91  11.65           (----*----)
NSWUA  14  41.33  19.82      (----*---)
SUSAF  32  29.93  17.88  (--*--)
                          ----+---------+---------+---------+-----
                             32        48        64        80

Pooled StDev = 14.02
```

Boxplot of Foundations of Wellbeing_n

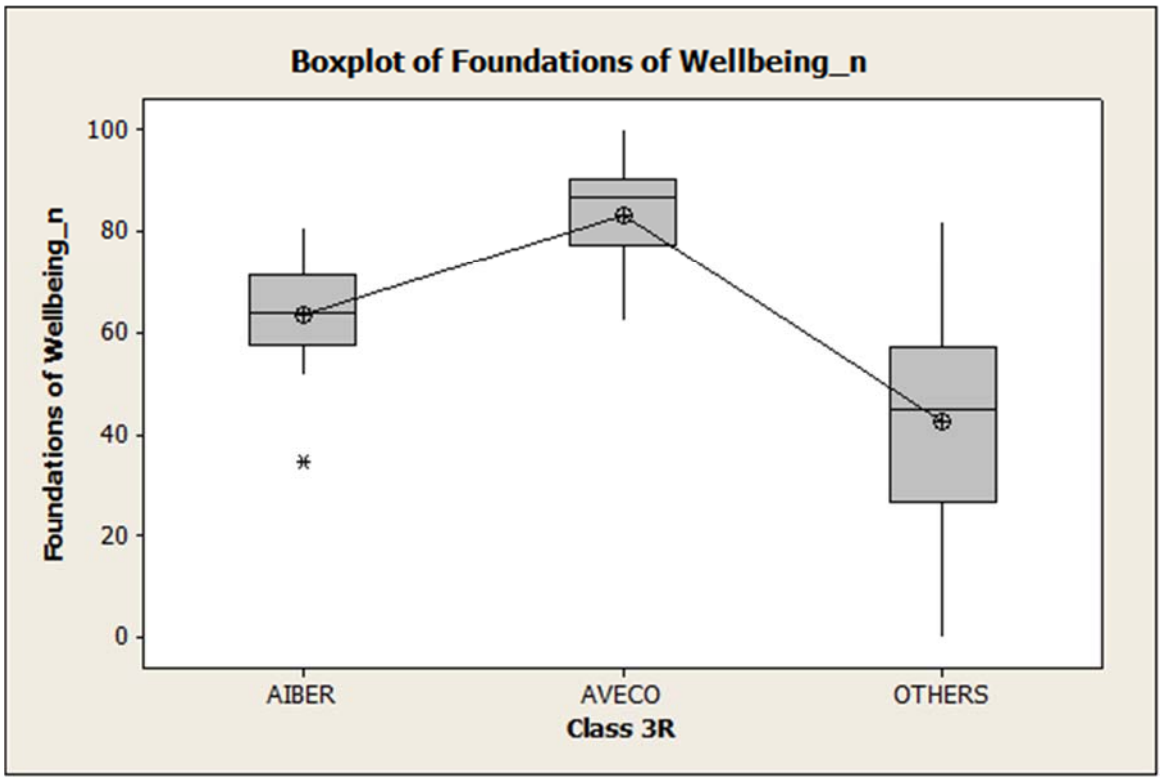


Obs: Os 3 países que compõem a região CARLA, são: Guiana, Jamaica e Trinidade e Tobago

One-way ANOVA: Foundations of Wellbeing_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  36523  18262  64.83  0.000
Error     129  36336    282
Total     131  72859

S = 16.78   R-Sq = 50.13%   R-Sq(adj) = 49.36%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  63.48  10.26            (----*----)
AVECO   27  83.29   9.51                       (----*---)
OTHERS  84  42.32  19.60  (-*--)
                          ----+---------+---------+---------+-----
                             45        60        75        90

Pooled StDev = 16.78
```

Boxplot of Foundations of Wellbeing_n

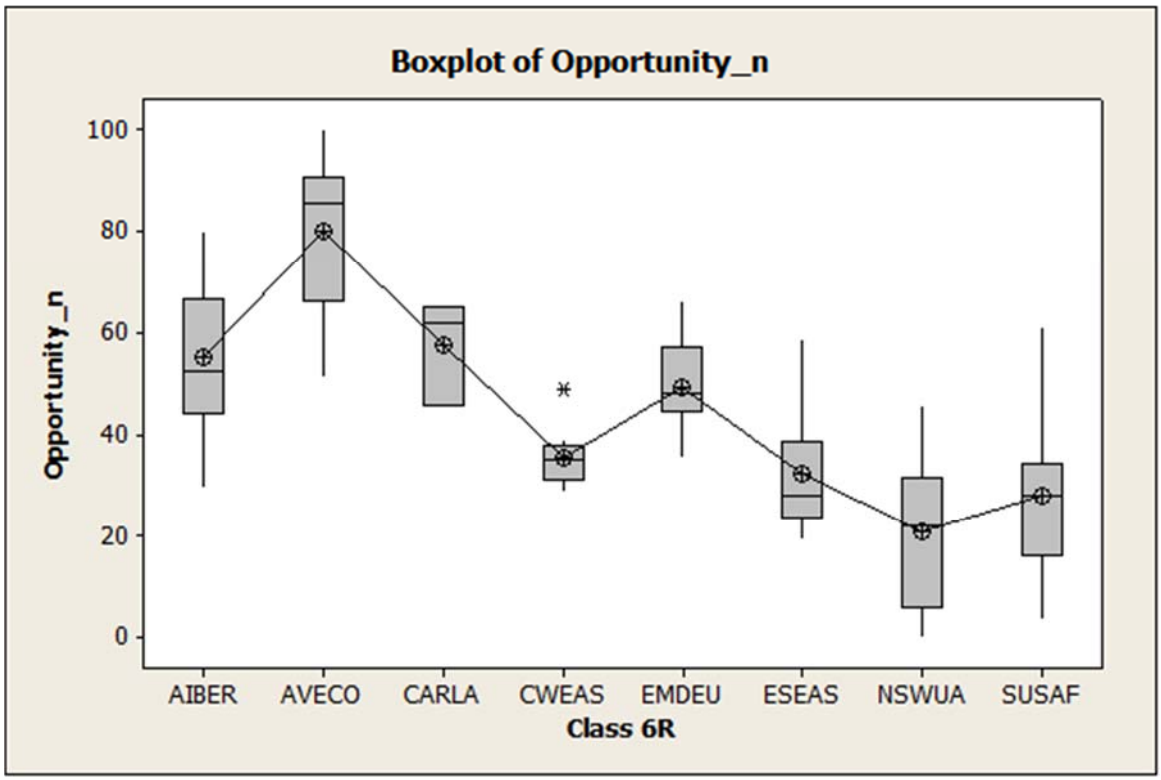


Obs: No gráfico acima, Cuba figura como *outlier* no agrupamento de países da região AIBER.

One-way ANOVA: Opportunity_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56157  8022  45.15  0.000
Error     124  22034   178
Total     131  78191

S = 13.33   R-Sq = 71.82%   R-Sq(adj) = 70.23%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  21  55.28  14.52                  (--*--)
AVECO  27  79.94  14.09                             (--*--)
CARLA   3  57.43  10.36              (-------*------)
CWEAS  11  35.37   5.63        (---*---)
EMDEU  12  49.18   9.28              (---*--)
ESEAS  12  32.27  13.01      (---*---)
NSWUA  14  20.97  14.49  (--*---)
SUSAF  32  27.92  14.57      (-*-)
                          ---+---------+---------+---------+------
                            20        40        60        80

Pooled StDev = 13.33
```

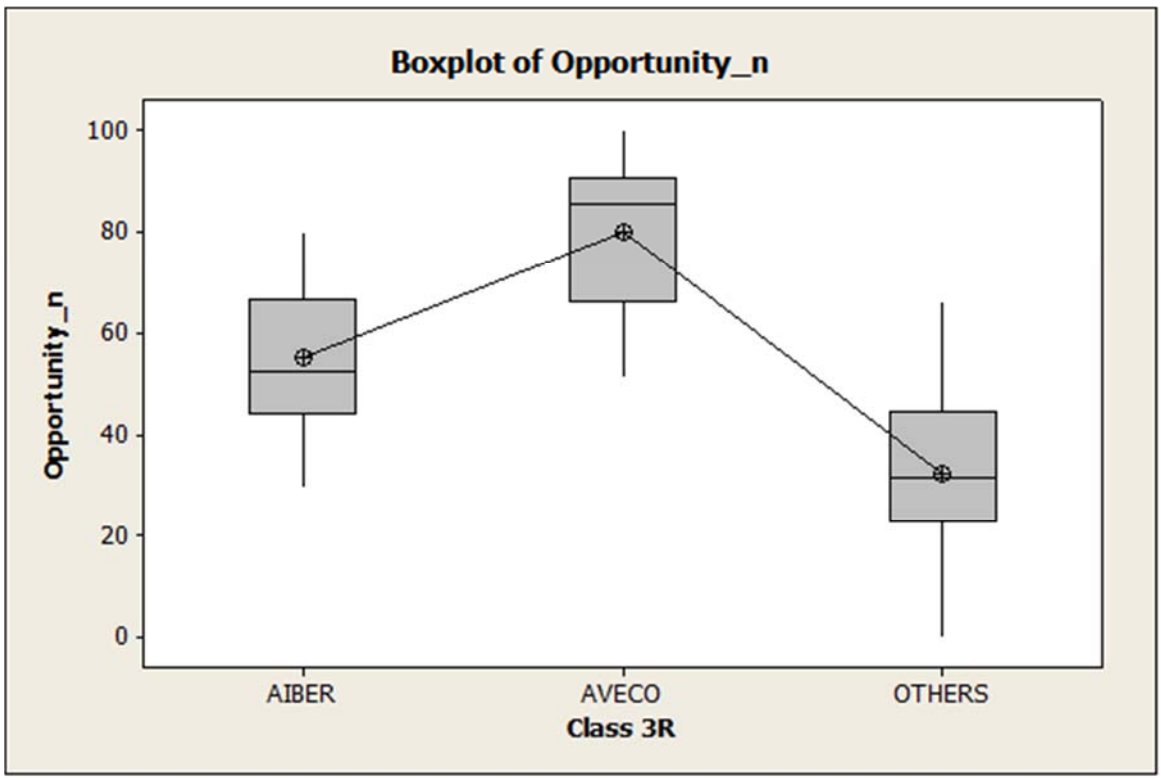
Boxplot of Opportunity_n



One-way ANOVA: Opportunity_n versus Class 3R

```
Source     DF     SS     MS       F      P
Class 3R    2  48330  24165  104.39  0.000
Error     129  29861    231
Total     131  78191

S = 15.21   R-Sq = 61.81%   R-Sq(adj) = 61.22%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  55.28  14.52              (----*---)
AVECO   27  79.94  14.09                                (---*---)
OTHERS  84  32.45  15.71  (--*-)
                          -+---------+---------+---------+--------
                          30        45        60        75

Pooled StDev = 15.21
```

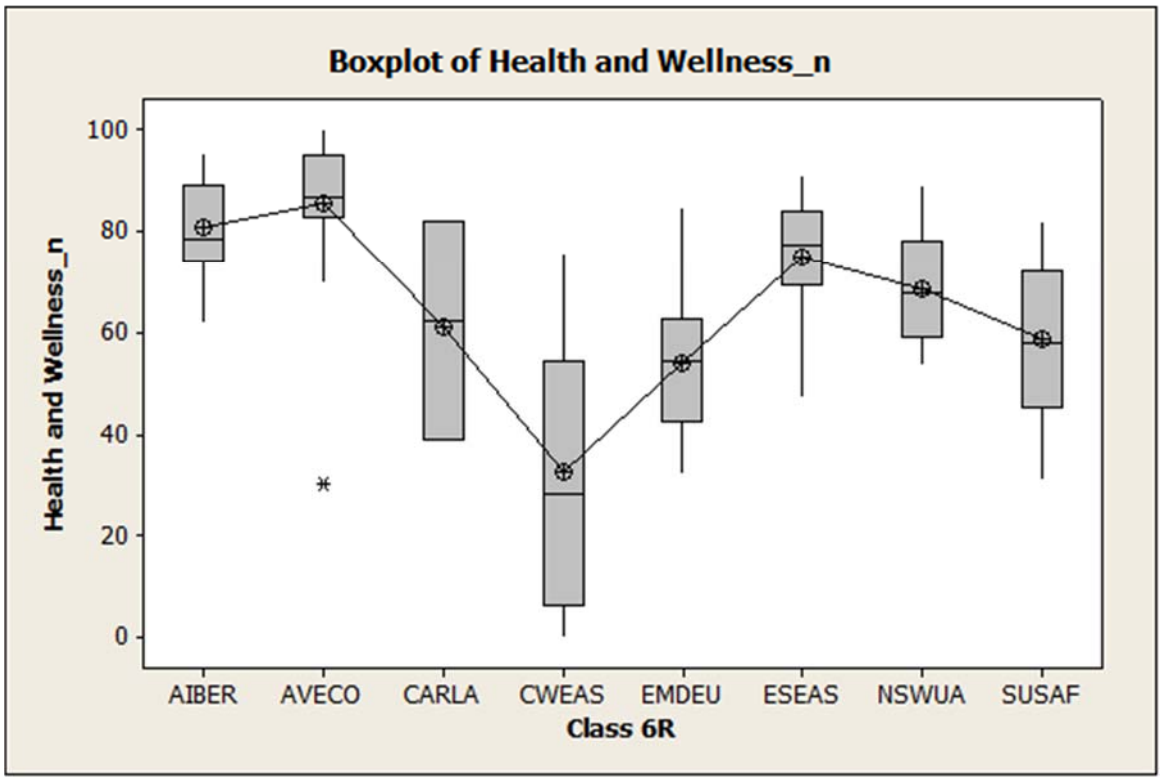
Boxplot of Opportunity_n



One-way ANOVA: Health and Wellness_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  31347  4478  21.57  0.000
Error     124  25748   208
Total     131  57095

S = 14.41   R-Sq = 54.90%   R-Sq(adj) = 52.36%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  21  80.66   9.28                          (--*--)
AVECO  27  85.46  14.00                             (--*-)
CARLA   3  61.04  21.44            (--------*-------)
CWEAS  11  32.48  24.22  (---*----)
EMDEU  12  54.08  14.08             (---*---)
ESEAS  12  74.94  12.49                       (---*----)
NSWUA  14  68.41  11.37                     (---*---)
SUSAF  32  58.57  14.55                  (-*--)
                          --------+---------+---------+---------+-
                                 40        60        80       100

Pooled StDev = 14.41
```

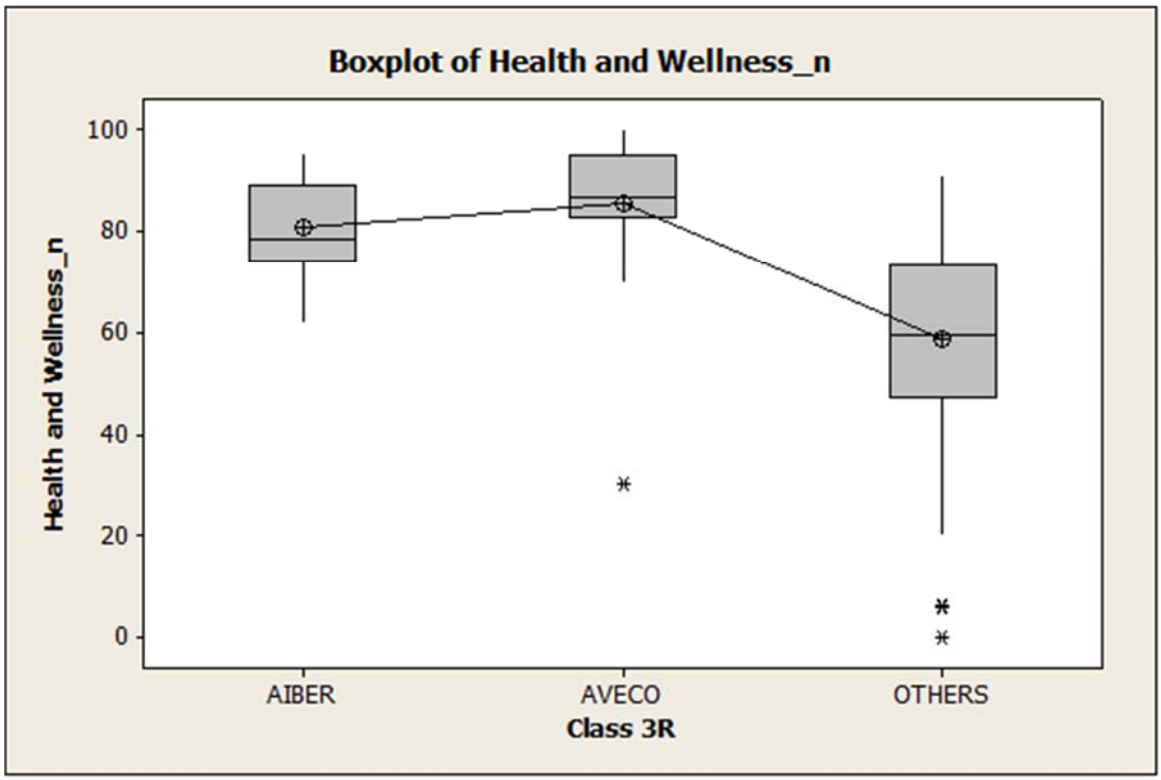
Boxplot of Health and Wellness_n



One-way ANOVA: Health and Wellness_n versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  19025  9513  32.23  0.000
Error     129  38070   295
Total     131  57095

S = 17.18   R-Sq = 33.32%   R-Sq(adj) = 32.29%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  80.66   9.28                   (-------*------)
AVECO   27  85.46  14.00                         (-----*------)
OTHERS  84  58.58  19.40  (---*--)
                          -----+---------+---------+---------+----
                              60        70        80        90

Pooled StDev = 17.18
```

Boxplot of Health and Wellness_n

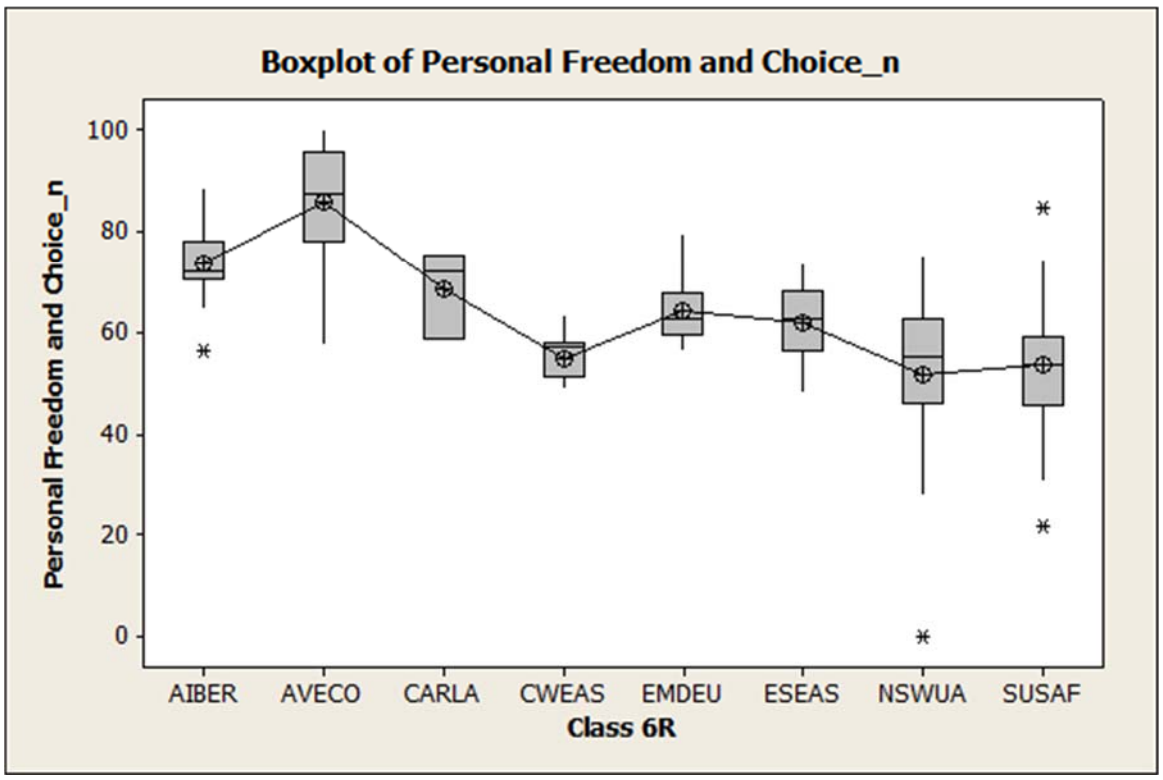


Obs: A Letônia (Latvia), aparece como *outlier* nesta avaliação, denotando resultado bem abaixo dos de países do agrupamento AVECO. E no grupo OTHERS, os dois *outliers* são Rússia e Cazaquistão (Kazakhstan).

One-way ANOVA: Personal Freedom and Choice_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  21207  3030  23.54  0.000
Error     124  15961   129
Total     131  37168

S = 11.35   R-Sq = 57.06%   R-Sq(adj) = 54.63%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
AIBER  21  73.74   7.32                    (---*----)
AVECO  27  85.82  11.92                              (---*--)
CARLA   3  68.60   8.78          (----------*----------)
CWEAS  11  54.87   4.54    (-----*----)
EMDEU  12  64.10   6.30           (----*-----)
ESEAS  12  61.66   8.22          (----*-----)
NSWUA  14  51.69  19.34  (----*----)
SUSAF  32  53.64  12.43     (---*--)
                          --+---------+---------+---------+-------
                           48        60        72        84

Pooled StDev = 11.35
```

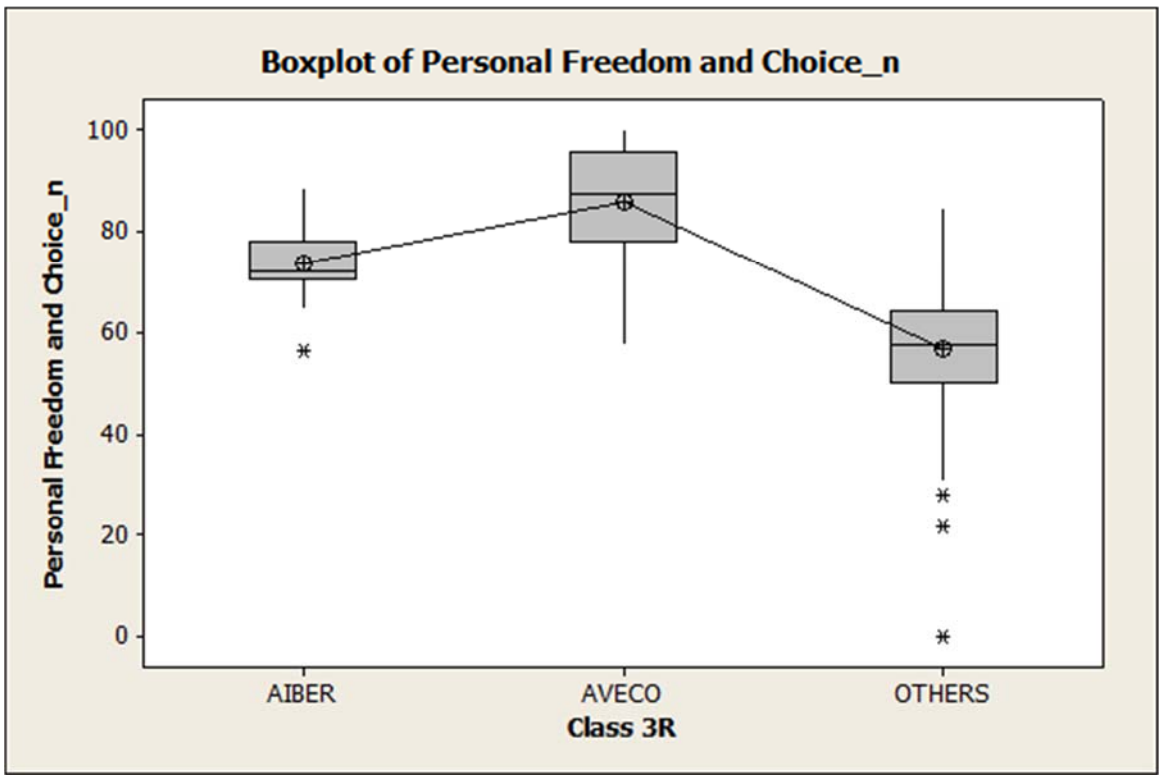
Boxplot of Personal Freedom and Choice_n



One-way ANOVA: Personal Freedom and Choice_n versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  19142  9571  68.49  0.000
Error     129  18027   140
Total     131  37168

S = 11.82   R-Sq = 51.50%   R-Sq(adj) = 50.75%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  73.74   7.32                (----*----)
AVECO   27  85.82  11.92                          (----*---)
OTHERS  84  56.65  12.64  (--*-)
                          ------+---------+---------+---------+---
                               60        70        80        90

Pooled StDev = 11.82
```

Boxplot of Personal Freedom and Choice_n

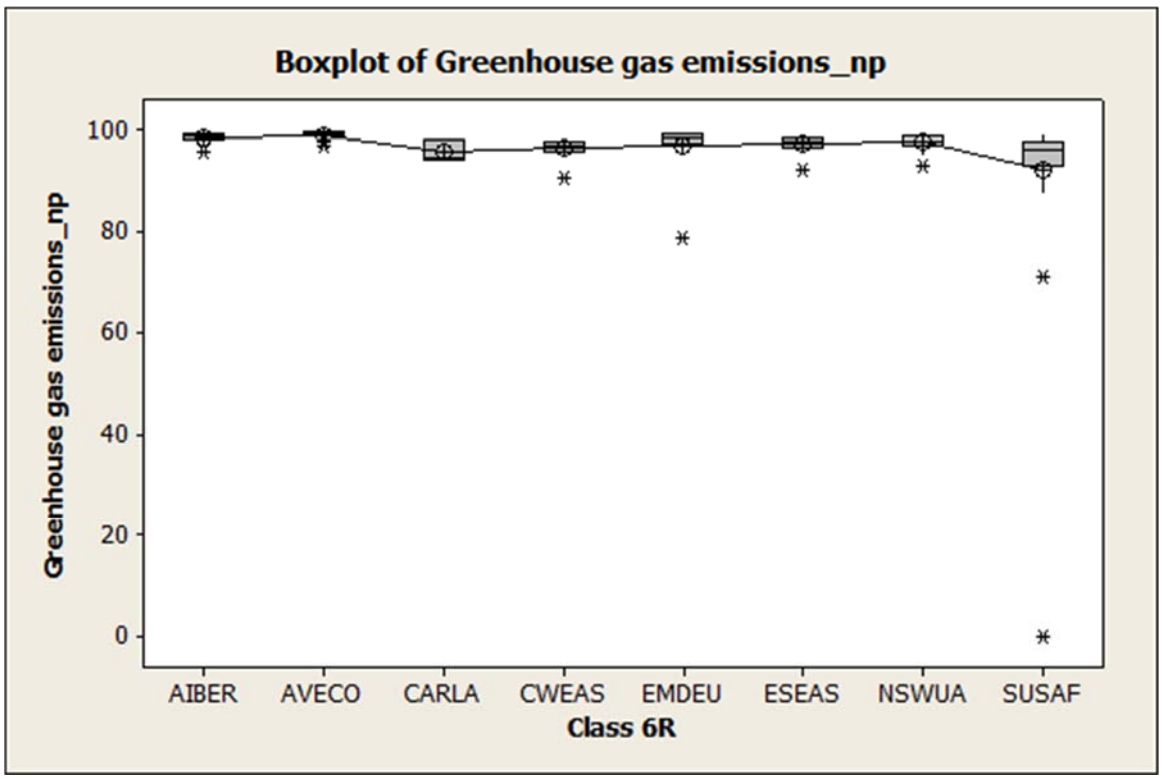


Obs: No agrupamento de países AIBER, Cuba aparece como *outlier* em relação aos demais. Os outliers do grupo OTHERS são respectivamente Paquistão, Sudão e Mauritânia.

One-way ANOVA: Greenhouse gas emissions_np versus Class 6R

```
Source     DF       SS     MS     F      P
Class 6R    7    933.2  133.3  1.64  0.131
Error     124  10092.4   81.4
Total     131  11025.6

S = 9.022   R-Sq = 8.46%   R-Sq(adj) = 3.30%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  21  98.53   1.02                (-----*------)
AVECO  27  99.09   0.62                 (-----*-----)
CARLA   3  95.64   2.23  (----------------*----------------)
CWEAS  11  96.53   2.07           (--------*--------)
EMDEU  12  96.93   5.75            (--------*-------)
ESEAS  12  97.26   1.89             (-------*--------)
NSWUA  14  97.60   1.83              (-------*-------)
SUSAF  32  92.04  17.56       (----*-----)
                          --------+---------+---------+---------+-
                               90.0      96.0     102.0     108.0

Pooled StDev = 9.02
```

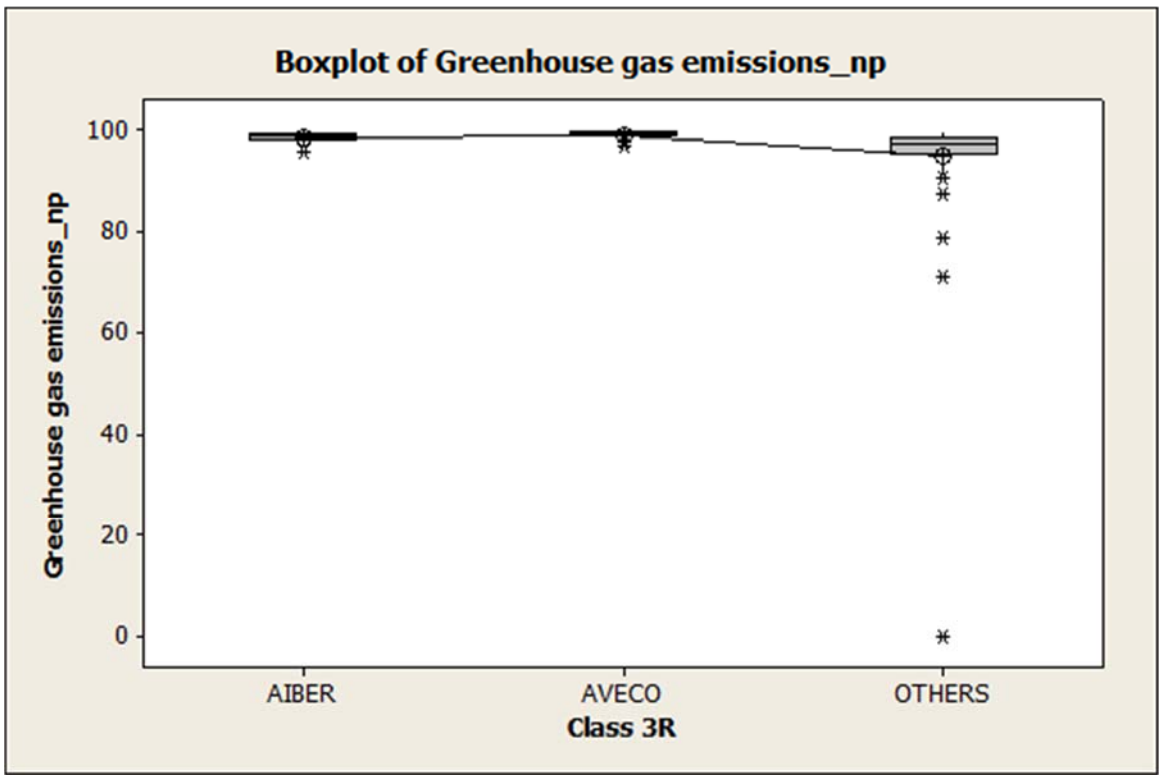
Boxplot of Greenhouse gas emissions_np



One-way ANOVA: Greenhouse gas emissions_np versus Class 3R

```
Source     DF       SS     MS     F      P
Class 3R    2    426.8  213.4  2.60  0.078
Error     129  10598.8   82.2
Total     131  11025.6

S = 9.064   R-Sq = 3.87%   R-Sq(adj) = 2.38%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  98.53   1.02       (---------------*---------------)
AVECO   27  99.09   0.62            (------------*-------------)
OTHERS  84  95.13  11.28  (-------*------)
                          -------+---------+---------+---------+--
                              95.0      97.5     100.0     102.5

Pooled StDev = 9.06
```

Boxplot of Greenhouse gas emissions_np

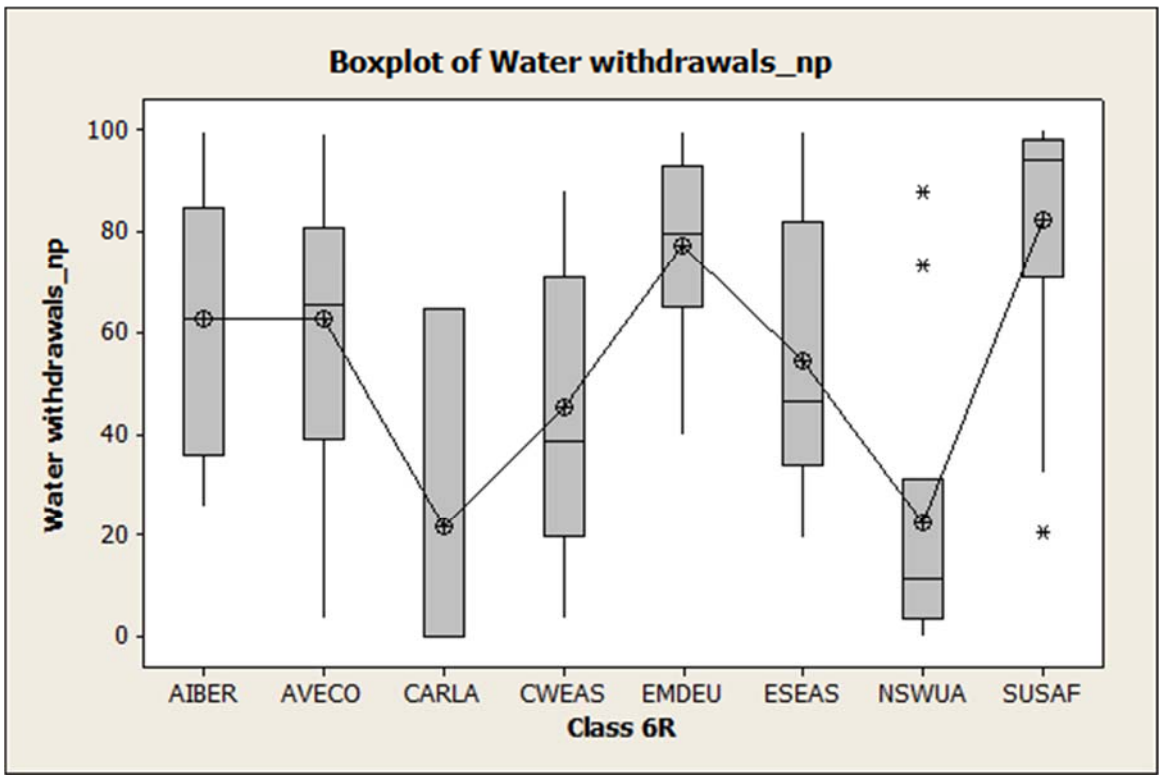


Obs:A República Centro-Africana (Central African Republic), no coração da África, aparece neste estudo como o maior emissor de gases de efeito estufa. Segundo o relatório da ONU, sobre os dados de 2010, a principal causa deste resultado estava ligado a Desmatamento e Exploração de madeira para comércio (89.46%), seguido Agricultura (5.26%), entre outros.

One-way ANOVA: Water withdrawals_np versus Class 6R

```
Source     DF      SS    MS      F      P
Class 6R    7   46991  6713  10.48  0.000
Error     124   79426   641
Total     131  126417

S = 25.31   R-Sq = 37.17%   R-Sq(adj) = 33.62%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---+---------+---------+---------+------
AIBER  21  62.78  25.68                          (---*---)
AVECO  27  62.44  26.28                          (---*---)
CARLA   3  21.47  37.19  (-----------*----------)
CWEAS  11  45.38  28.32                (-----*-----)
EMDEU  12  77.35  19.49                                (-----*-----)
ESEAS  12  54.21  26.59                     (-----*----)
NSWUA  14  22.29  27.26         (----*----)
SUSAF  32  82.49  22.63                                     (---*---)
                          ---+---------+---------+---------+------
                             0        25        50        75

Pooled StDev = 25.31
```

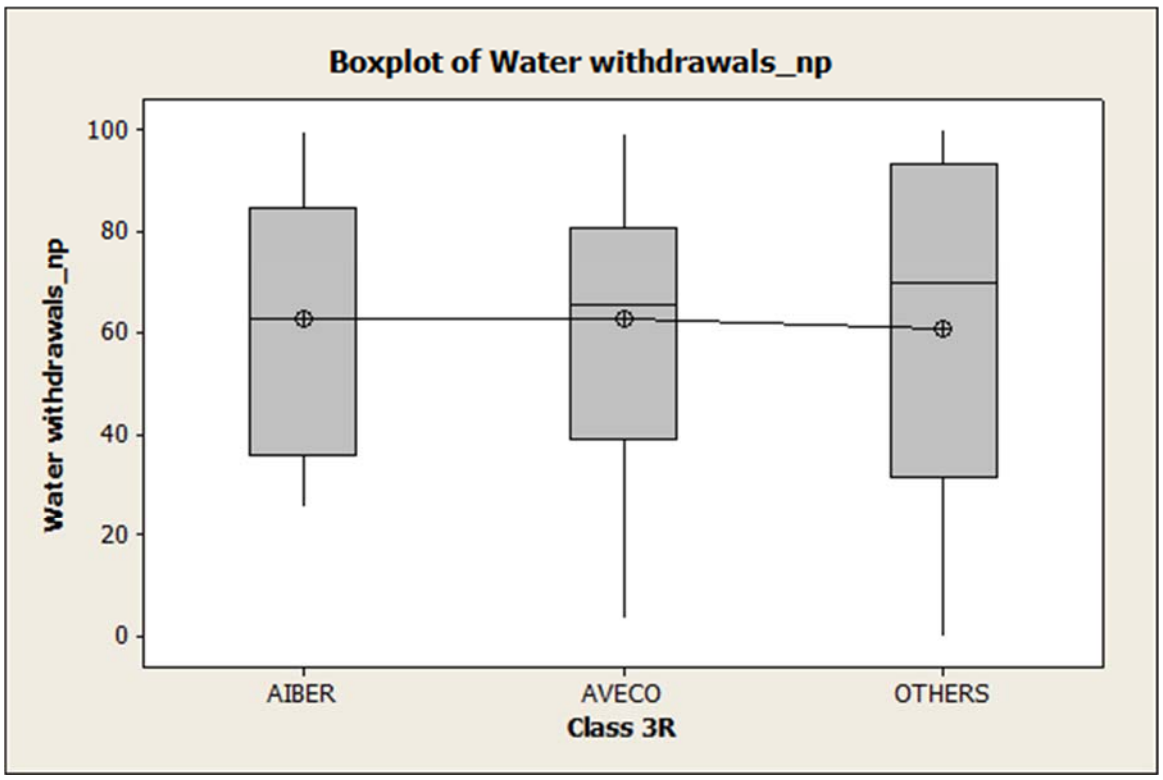
Boxplot of Water withdrawals_np



One-way ANOVA: Water withdrawals_np versus Class 3R

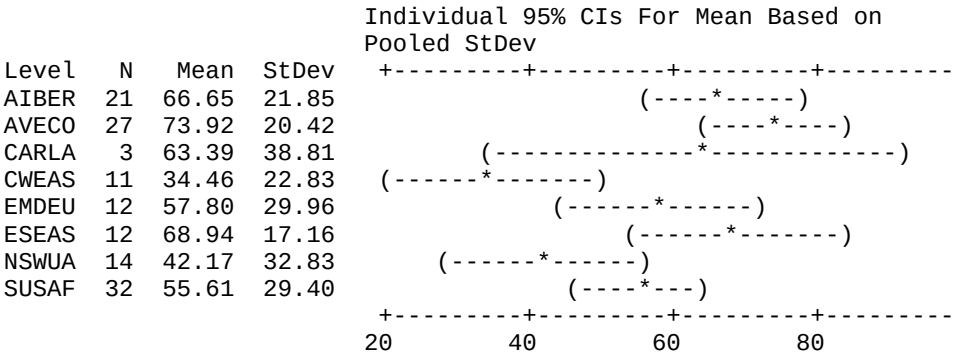
```
Source     DF      SS   MS     F      P
Class 3R    2     117   58  0.06  0.942
Error     129  126300  979
Total     131  126417

S = 31.29   R-Sq = 0.09%   R-Sq(adj) = 0.00%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev    +---------+---------+---------+---------
AIBER   21  62.78  25.68    (-------------------*-------------------)
AVECO   27  62.44  26.28      (----------------*----------------)
OTHERS  84  60.64  33.86           (---------*--------)
                             +---------+---------+---------+---------
                            49.0      56.0      63.0      70.0

Pooled StDev = 31.29
```

Boxplot of Water withdrawals_np



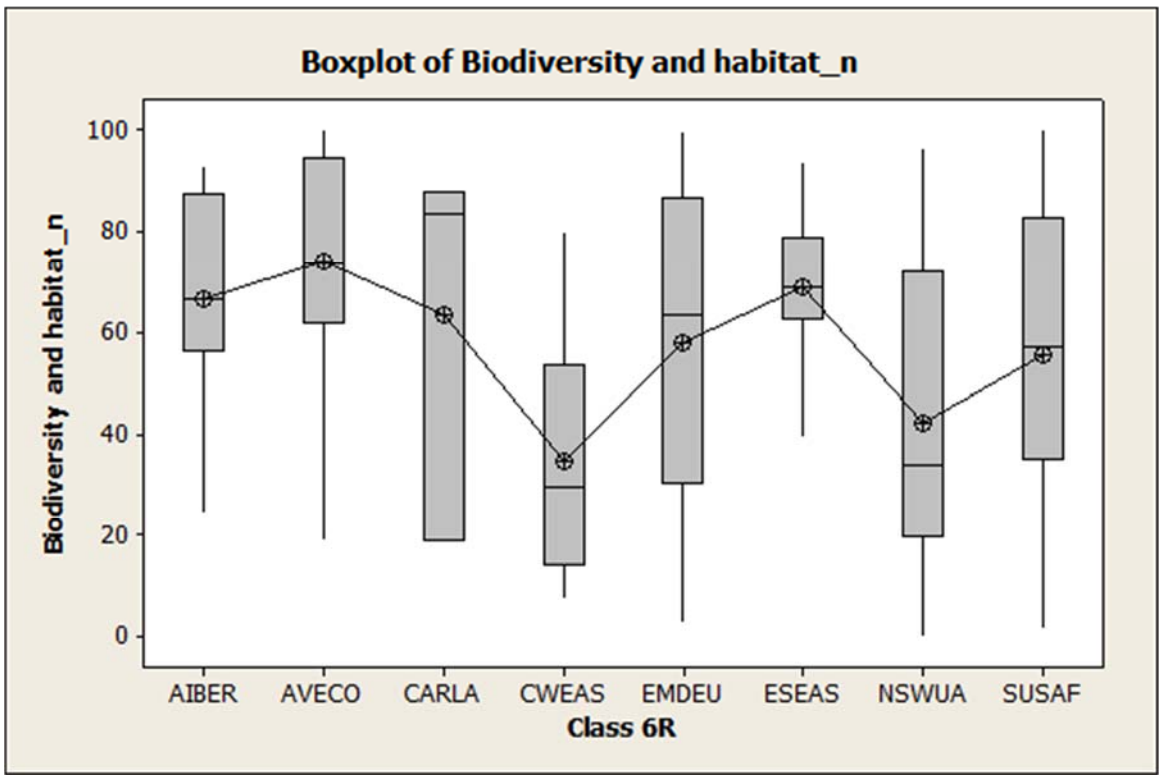
One-way ANOVA: Biodiversity and habitat_n versus Class 6R

```
Source     DF      SS    MS     F      P
Class 6R    7   19429  2776  4.17  0.000
Error     124   82530   666
Total     131  101958

S = 25.80   R-Sq = 19.06%   R-Sq(adj) = 14.49%
```




```
Pooled StDev = 25.80
```

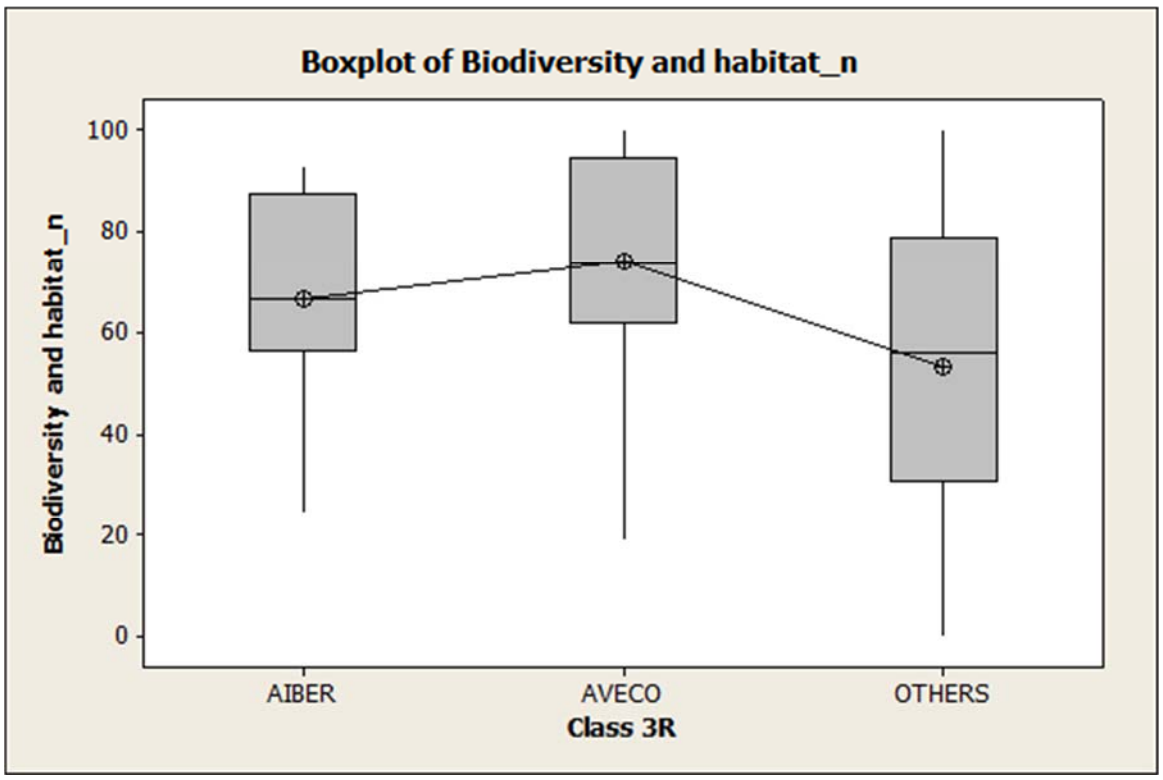
Boxplot of Biodiversity and habitat_n



One-way ANOVA: Biodiversity and habitat_n versus Class 3R

```
Source     DF      SS    MS     F      P
Class 3R    2   10137  5069  7.12  0.001
Error     129   91821   712
Total     131  101958

S = 26.68   R-Sq = 9.94%   R-Sq(adj) = 8.55%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  66.65  21.85          (-----------*----------)
AVECO   27  73.92  20.42                    (---------*---------)
OTHERS  84  53.10  29.34  (-----*-----)
                          ---+---------+---------+---------+------
                            50        60        70        80

Pooled StDev = 26.68
```

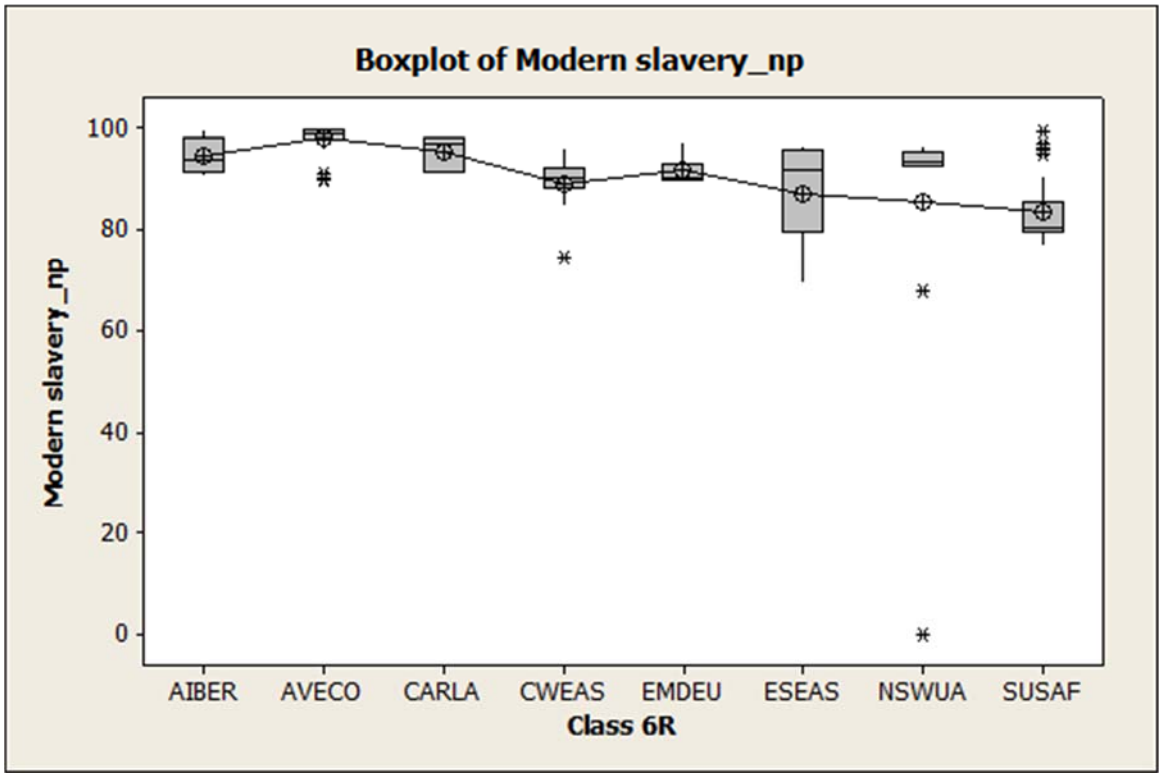
Boxplot of Biodiversity and habitat_n



One-way ANOVA: Modern slavery_np versus Class 6R

```
Source     DF       SS     MS     F      P
Class 6R    7   4022.0  574.6  5.97  0.000
Error     124  11925.0   96.2
Total     131  15947.0

S = 9.807   R-Sq = 25.22%   R-Sq(adj) = 21.00%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  21  94.54   3.29                (-----*-----)
AVECO  27  98.03   2.96                      (----*----)
CARLA   3  95.48   3.56       (---------------*---------------)
CWEAS  11  89.11   5.60      (-------*--------)
EMDEU  12  91.65   2.76          (-------*-------)
ESEAS  12  86.99   9.80    (-------*-------)
NSWUA  14  85.68  25.68   (------*-------)
SUSAF  32  83.49   6.79  (----*----)
                          ------+---------+---------+---------+---
                             84.0      91.0      98.0     105.0

Pooled StDev = 9.81
```

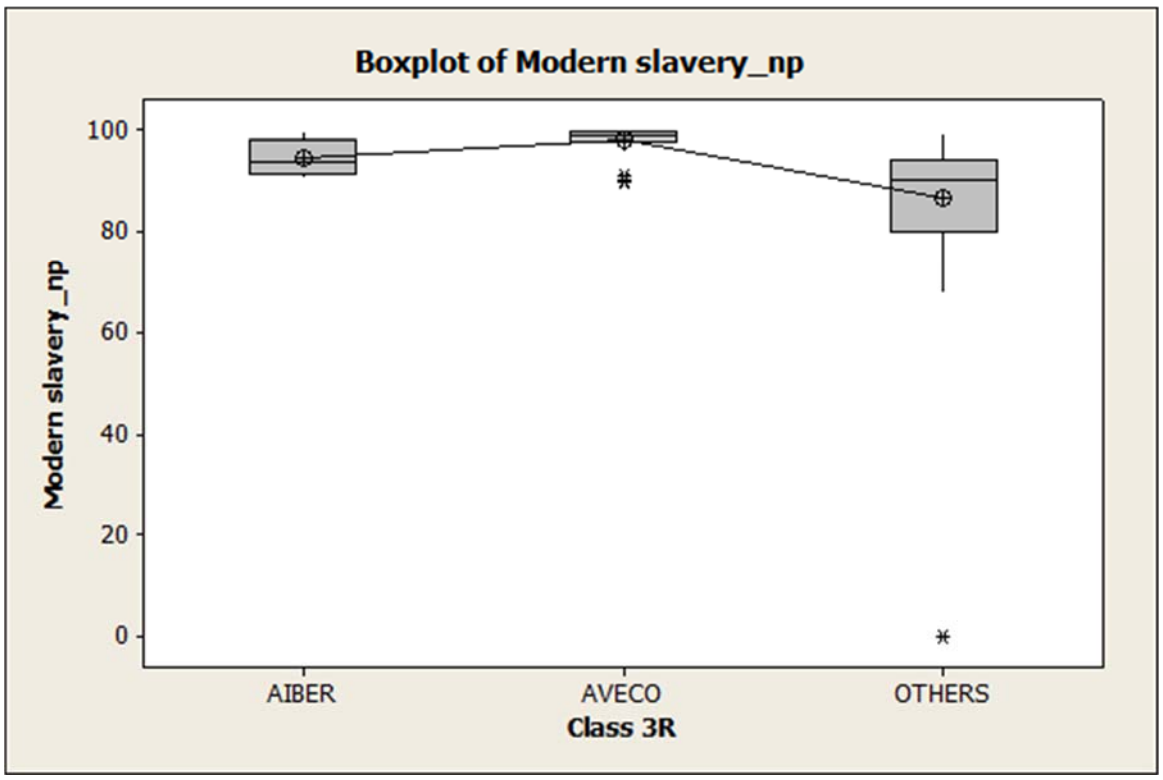
Boxplot of Modern slavery_np



One-way ANOVA: Modern slavery_np versus Class 3R

```
Source     DF       SS      MS      F      P
Class 3R    2   3088.6  1544.3  15.49  0.000
Error     129  12858.4    99.7
Total     131  15947.0

S = 9.984   R-Sq = 19.37%   R-Sq(adj) = 18.12%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev    -+---------+---------+---------+--------
AIBER   21  94.54   3.29               (--------*--------)
AVECO   27  98.03   2.96                       (-------*-------)
OTHERS  84  86.69  12.23    (---*----)
                            -+---------+---------+---------+--------
                          85.0      90.0      95.0     100.0

Pooled StDev = 9.98
```

Boxplot of Modern slavery_np

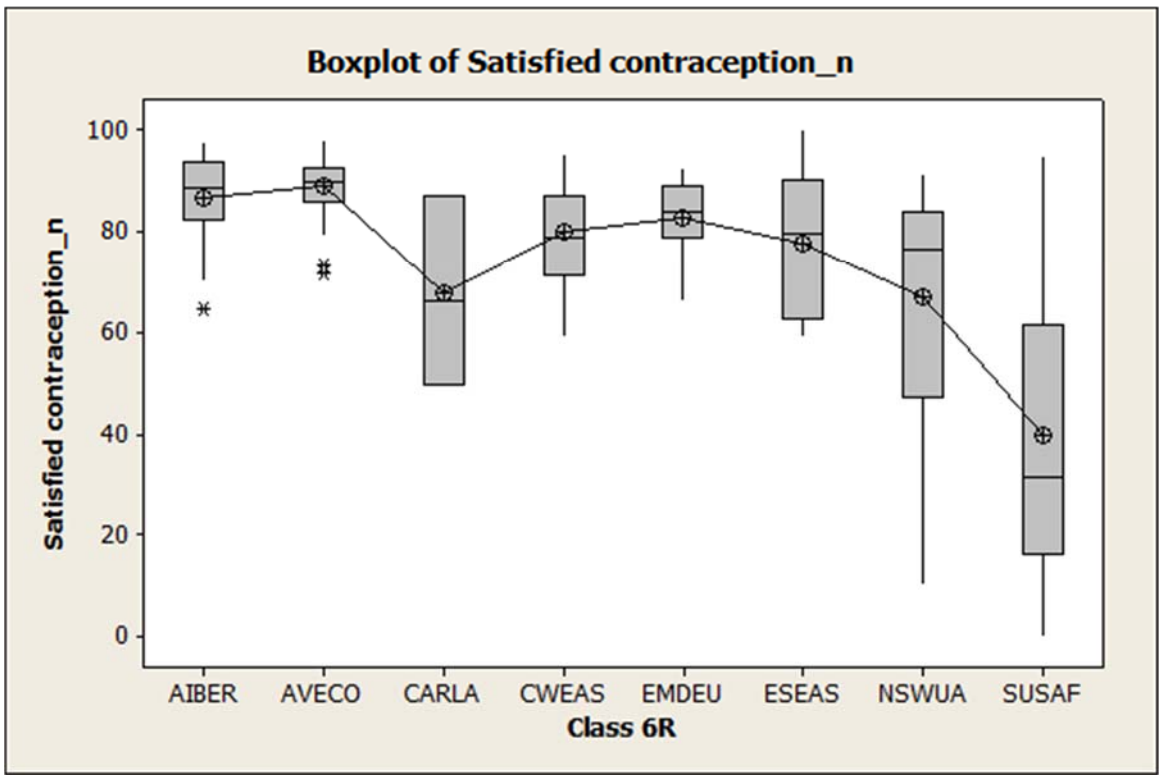


Obs: No agrupamento AVECO, vemos 3 *ouliers* que destoam do grupo, são eles: Eslovênia, Eslováquia e República Tcheca.
A Mauritânia lidera o ranking nesta categoria, com cerca de 155,600 indivíduos (4% da pop.), em situação de escravidão. fonte: http://www.globalslaveryindex.org/country/mauritania/

One-way ANOVA: Satisfied contraception (% of women) versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  48235  6891  24.56  0.000
Error     124  34788   281
Total     131  83022

S = 16.75   R-Sq = 58.10%   R-Sq(adj) = 55.73%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  21  86.69   8.82                                (---*----)
AVECO  27  88.99   6.48                                  (---*---)
CARLA   3  67.72  18.76          (-----------*-----------)
CWEAS  11  79.91  10.33                        (-----*-----)
EMDEU  12  82.55   8.25                          (-----*-----)
ESEAS  12  77.71  14.17                       (-----*-----)
NSWUA  14  66.98  22.77                (-----*----)
SUSAF  32  39.83  25.83  (---*---)
                          ---------+---------+---------+---------+
                                  48        64        80        96

Pooled StDev = 16.75
```

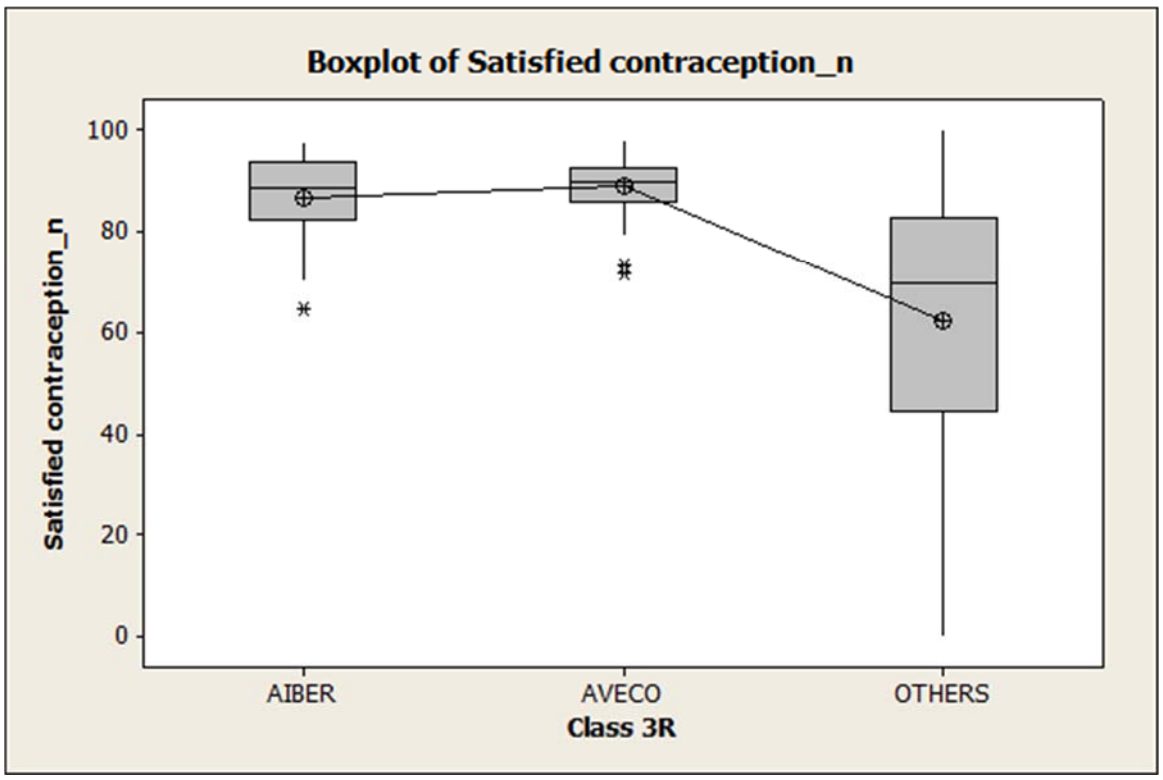
Boxplot of Satisfied contraception (% of women)



One-way ANOVA: Satisfied contraception (% of women) versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  20501  10251  21.15  0.000
Error     129  62521    485
Total     131  83022

S = 22.01   R-Sq = 24.69%   R-Sq(adj) = 23.53%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --+---------+---------+---------+-------
AIBER   21  86.69   8.82                 (-------*-------)
AVECO   27  88.99   6.48                    (------*------)
OTHERS  84  62.11  26.86  (---*---)
                          --+---------+---------+---------+-------
                           60        72        84        96

Pooled StDev = 22.01
```

Boxplot of Satisfied contraception (% of women)



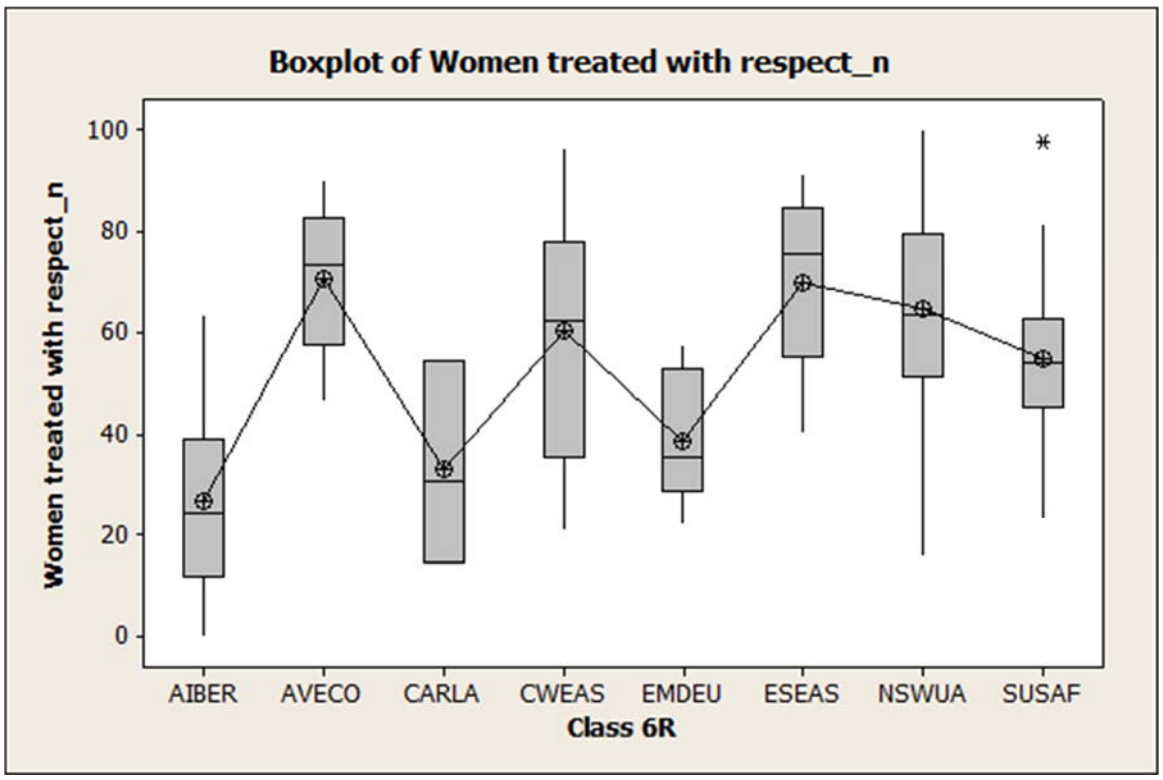
No agrupamento AIBER, figura como *outlier* a Guatemala, e no AVECO figuram Japão e Islândia respectivamente.

One-way ANOVA: Women treated with respect_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  32034  4576  15.43  0.000
Error     124  36779   297
Total     131  68813

S = 17.22   R-Sq = 46.55%   R-Sq(adj) = 43.53%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---+---------+---------+---------+------
AIBER  21  26.57  17.90      (--*---)
AVECO  27  70.33  13.59                            (--*--)
CARLA   3  33.16  19.99   (---------*--------)
CWEAS  11  60.42  23.44                     (----*----)
EMDEU  12  38.72  12.41          (----*----)
ESEAS  12  69.61  17.63                          (----*----)
NSWUA  14  64.55  22.05                        (---*----)
SUSAF  32  54.67  15.90                    (--*--)
                          ---+---------+---------+---------+------
                            20        40        60        80

Pooled StDev = 17.22
```

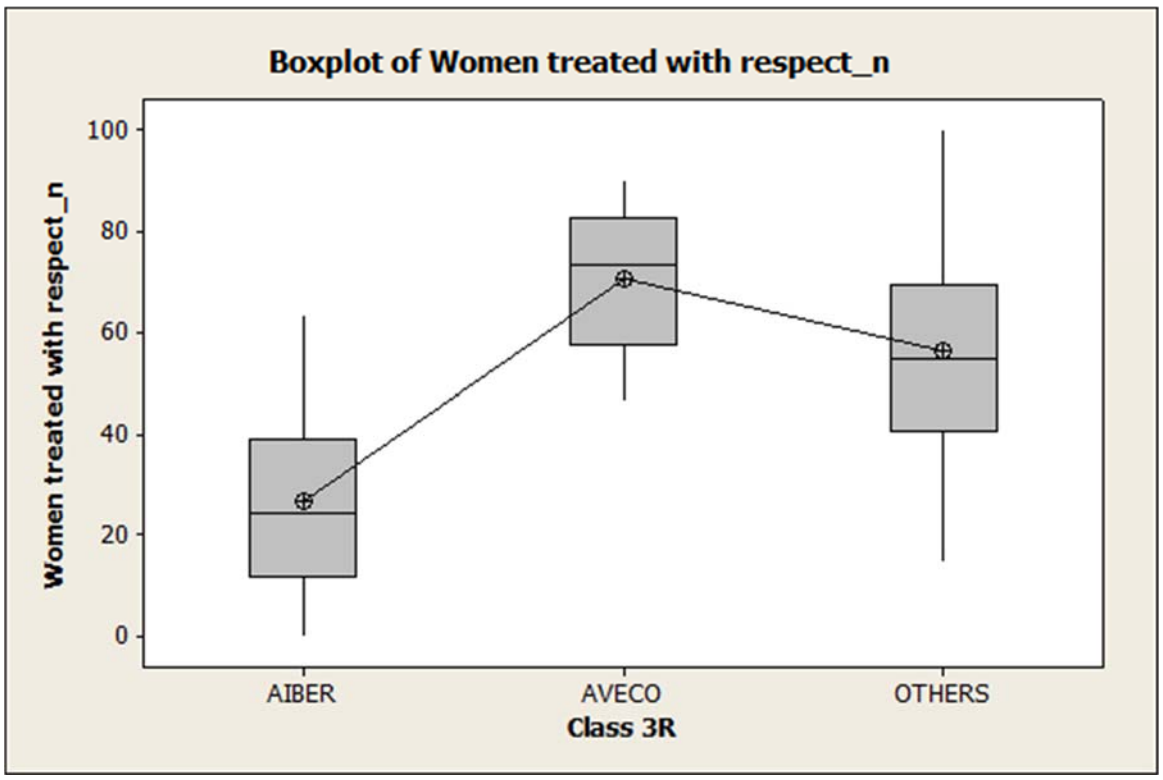
Boxplot of Women treated with respect_n



One-way ANOVA: Women treated with respect_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  23369  11685  33.17  0.000
Error     129  45444    352
Total     131  68813

S = 18.77   R-Sq = 33.96%   R-Sq(adj) = 32.94%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  26.57  17.90  (-----*----)
AVECO   27  70.33  13.59                                (----*----)
OTHERS  84  56.16  20.31                          (-*--)
                          --------+---------+---------+---------+-
                                 30        45        60        75

Pooled StDev = 18.77
```

Boxplot of Women treated with respect_n

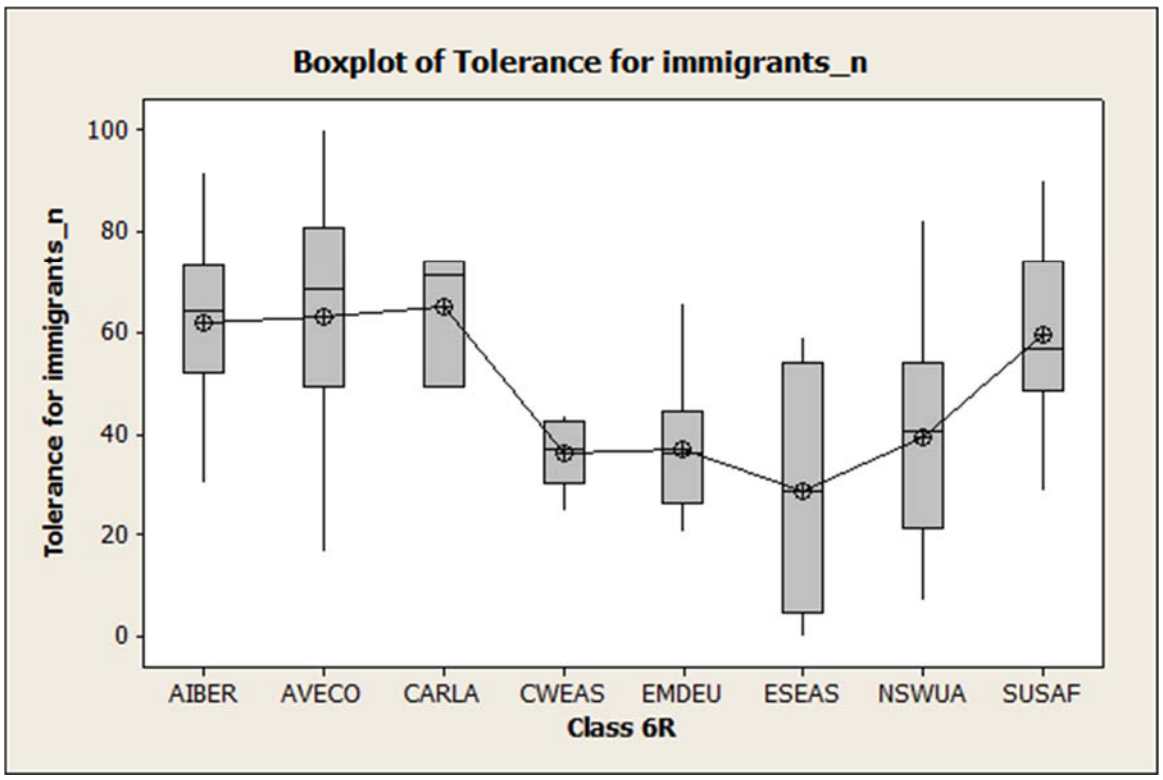


No agrupamento AIBER, os 3 países com os menores valores são: Perú com 24, República Dominicana com 18 e Colômbia com 17. O Brasil com 28, situa-se próximo destes.

One-way ANOVA: Tolerance for immigrants_n versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  21882  3126  8.89  0.000
Error     124  43600   352
Total     131  65483

S = 18.75   R-Sq = 33.42%   R-Sq(adj) = 29.66%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  21  61.97  16.25                   (---*---)
AVECO  27  62.86  24.27                    (--*---)
CARLA   3  64.84  13.51              (---------*----------)
CWEAS  11  36.11   6.46    (-----*-----)
EMDEU  12  37.10  12.67     (-----*----)
ESEAS  12  28.54  23.08  (----*-----)
NSWUA  14  39.43  22.78       (----*----)
SUSAF  32  59.63  15.91                   (--*--)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 18.75
```

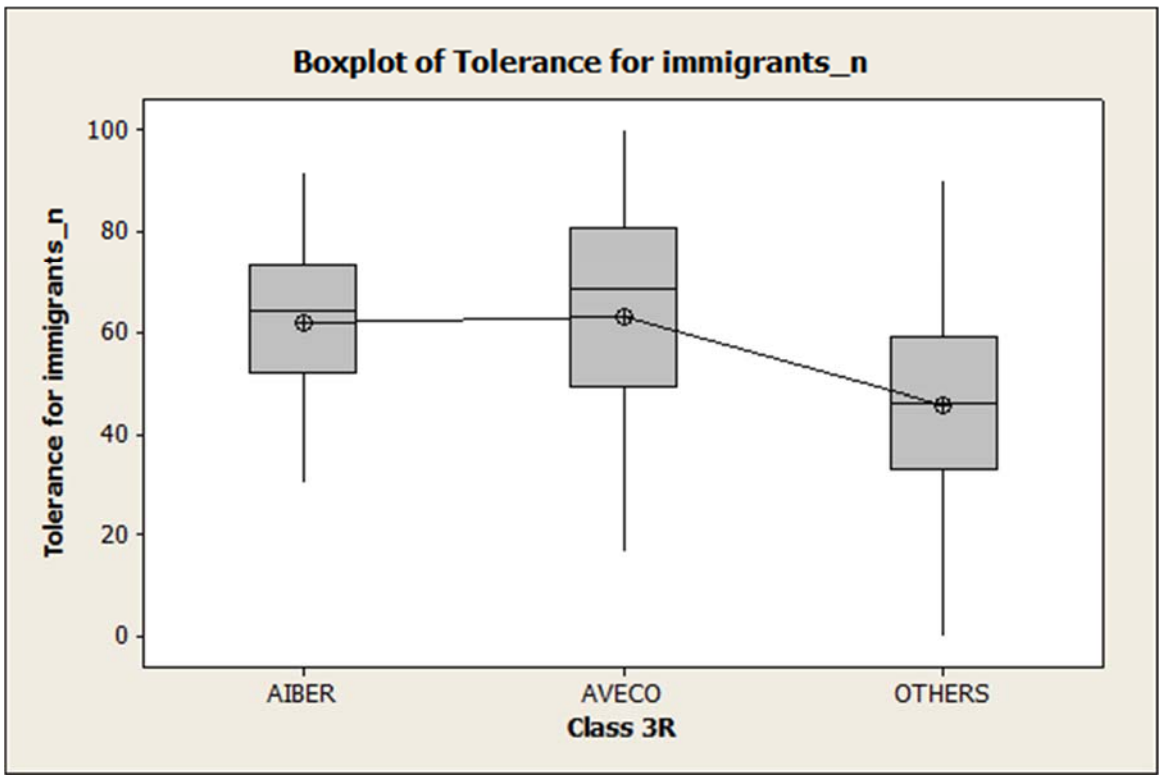
Boxplot of Tolerance for immigrants_n



One-way ANOVA: Tolerance for immigrants_n versus Class 3R

```
Source     DF     SS    MS     F      P
Class 3R    2   8590  4295  9.74  0.000
Error     129  56893   441
Total     131  65483

S = 21.00   R-Sq = 13.12%   R-Sq(adj) = 11.77%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  61.97  16.25                (----------*-----------)
AVECO   27  62.86  24.27                   (---------*---------)
OTHERS  84  45.71  20.91  (-----*-----)
                          ---------+---------+---------+---------+
                                48.0      56.0      64.0      72.0

Pooled StDev = 21.00
```

Boxplot of Tolerance for immigrants_n

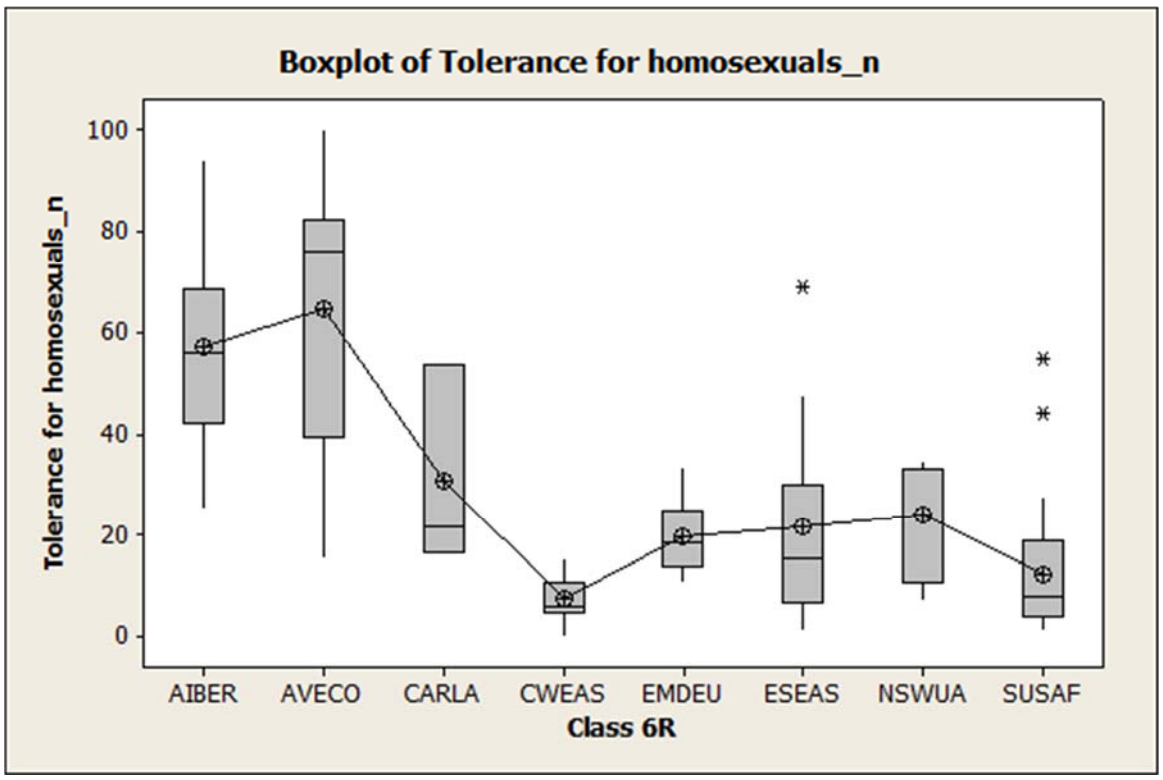


One-way ANOVA: Tolerance for homosexuals_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  64859  9266  32.77  0.000
Error     124  35063   283
Total     131  99923

S = 16.82   R-Sq = 64.91%   R-Sq(adj) = 62.93%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  21  57.14  17.43                            (---*--)
AVECO  27  64.51  25.29                               (--*--)
CARLA   3  30.56  20.07          (--------*---------)
CWEAS  11   7.25   4.59  (----*----)
EMDEU  12  19.74   6.69        (----*----)
ESEAS  12  21.73  19.95         (----*----)
NSWUA  14  24.01  11.54           (---*---)
SUSAF  32  12.28  12.37      (--*--)
                          -+---------+---------+---------+--------
                           0         20        40        60

Pooled StDev = 16.82
```

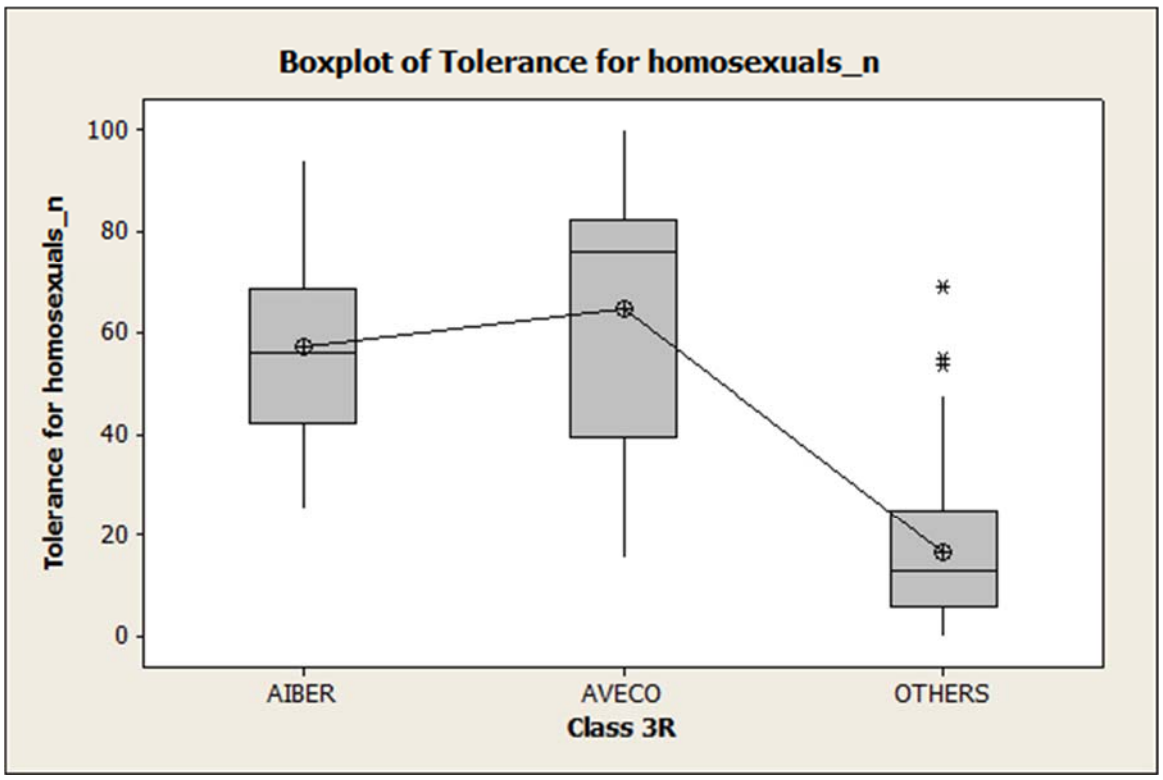
Boxplot of Tolerance for homosexuals_n



One-way ANOVA: Tolerance for homosexuals_n versus Class 3R

```
Source     DF     SS     MS       F      P
Class 3R    2  61514  30757  103.30  0.000
Error     129  38409    298
Total     131  99923

S = 17.26   R-Sq = 61.56%   R-Sq(adj) = 60.97%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  57.14  17.43                        (----*----)
AVECO   27  64.51  25.29                             (---*---)
OTHERS  84  16.64  13.76  (-*--)
                          -+---------+---------+---------+--------
                          15        30        45        60

Pooled StDev = 17.26
```

Boxplot of Tolerance for homosexuals_n

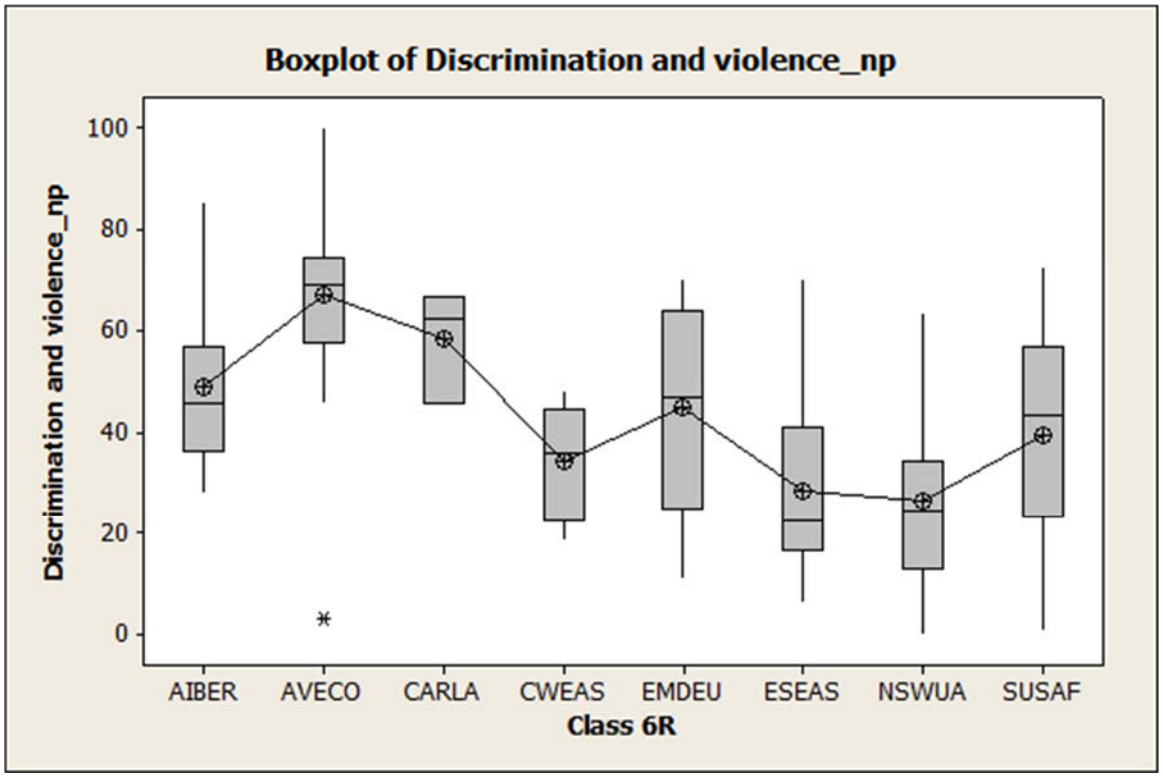


One-way ANOVA: Discrimination and violence against minorities_np versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  24562  3509  10.61  0.000
Error     124  40999   331
Total     131  65561

S = 18.18   R-Sq = 37.46%   R-Sq(adj) = 33.93%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  21  48.96  16.08                 (----*----)
AVECO  27  67.10  19.37                           (---*---)
CARLA   3  58.27  11.10               (------------*------------)
CWEAS  11  34.26  10.61       (-----*------)
EMDEU  12  44.86  20.39             (-----*------)
ESEAS  12  28.24  17.60    (------*-----)
NSWUA  14  26.21  18.64   (-----*-----)
SUSAF  32  39.55  19.84             (---*---)
                          +---------+---------+---------+---------
                          16        32        48        64

Pooled StDev = 18.18
```

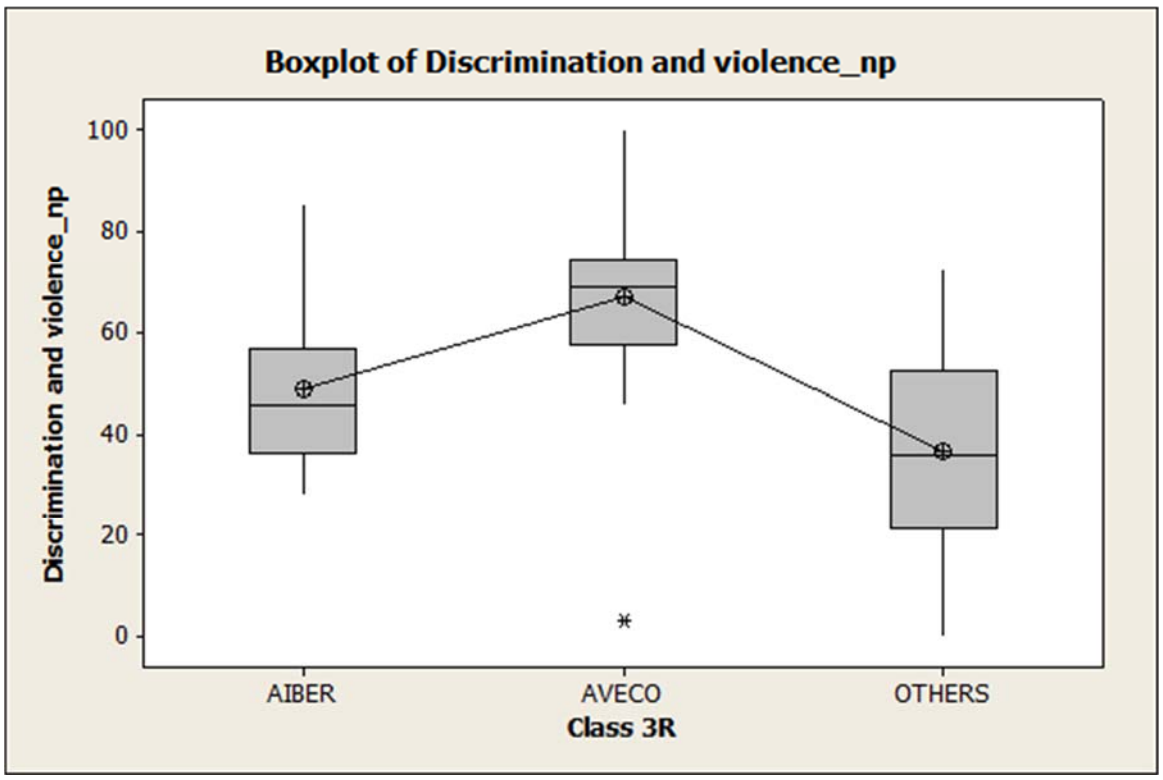
Boxplot of Discrimination and violence against minorities_np



One-way ANOVA: Discrimination and violence against minorities_np versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  19650  9825  27.61  0.000
Error     129  45911   356
Total     131  65561

S = 18.87   R-Sq = 29.97%   R-Sq(adj) = 28.89%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---+---------+---------+---------+------
AIBER   21  48.96  16.08        (------*------)
AVECO   27  67.10  19.37                       (-----*-----)
OTHERS  84  36.45  19.32  (--*---)
                          ---+---------+---------+---------+------
                            36        48        60        72

Pooled StDev = 18.87
```

Boxplot of Discrimination and violence against minorities_np

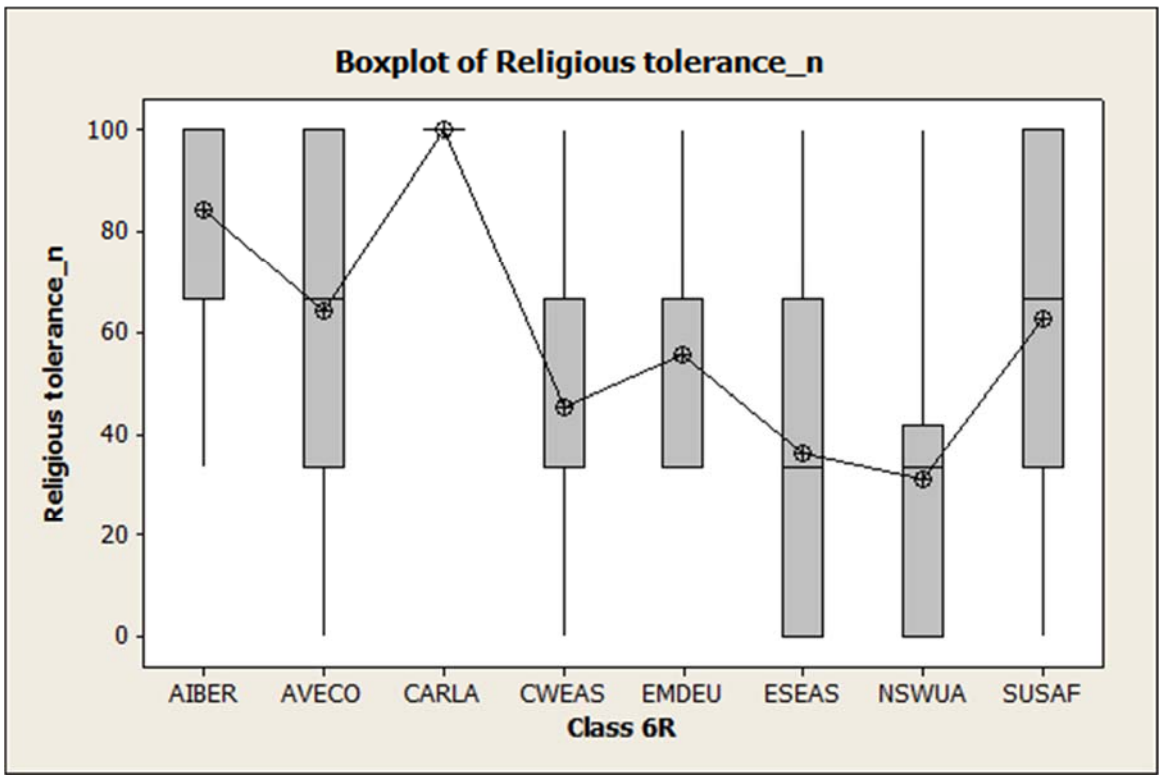


Obs: No agrupamento AVECO, Israel aparece como *outlier* destoando significativamente de seu grupo.

One-way ANOVA: Religious tolerance_n versus Class 6R

```
Source     DF      SS    MS     F      P
Class 6R    7   38868  5553  6.76  0.000
Error     124  101830   821
Total     131  140699

S = 28.66   R-Sq = 27.63%   R-Sq(adj) = 23.54%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  -----+---------+---------+---------+----
AIBER  21   84.13  22.65                    (---*---)
AVECO  27   64.20  27.62                (--*---)
CARLA   3  100.00   0.00                    (----------*----------)
CWEAS  11   45.45  26.97      (-----*-----)
EMDEU  12   55.56  21.71          (-----*----)
ESEAS  12   36.11  38.82    (----*----)
NSWUA  14   30.95  30.56  (----*----)
SUSAF  32   62.50  31.40              (---*--)
                          -----+---------+---------+---------+----
                              30        60        90       120

Pooled StDev = 28.66
```

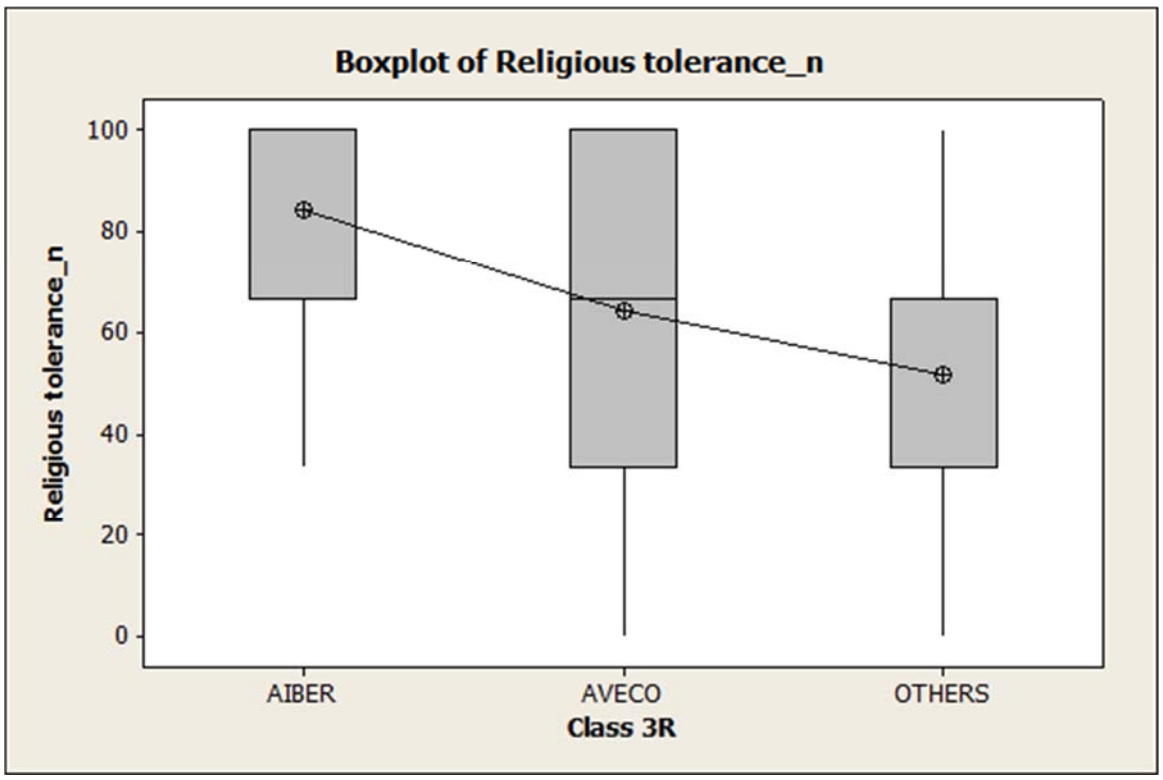
Boxplot of Religious tolerance_n



One-way ANOVA: Religious tolerance_n versus Class 3R

```
Source     DF      SS    MS     F      P
Class 3R    2   18588  9294  9.82  0.000
Error     129  122111   947
Total     131  140699

S = 30.77   R-Sq = 13.21%   R-Sq(adj) = 11.87%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  84.13  22.65                   (--------*--------)
AVECO   27  64.20  27.62         (-------*-------)
OTHERS  84  51.59  33.30   (---*----)
                            +---------+---------+---------+---------
                          45        60        75        90

Pooled StDev = 30.77
```

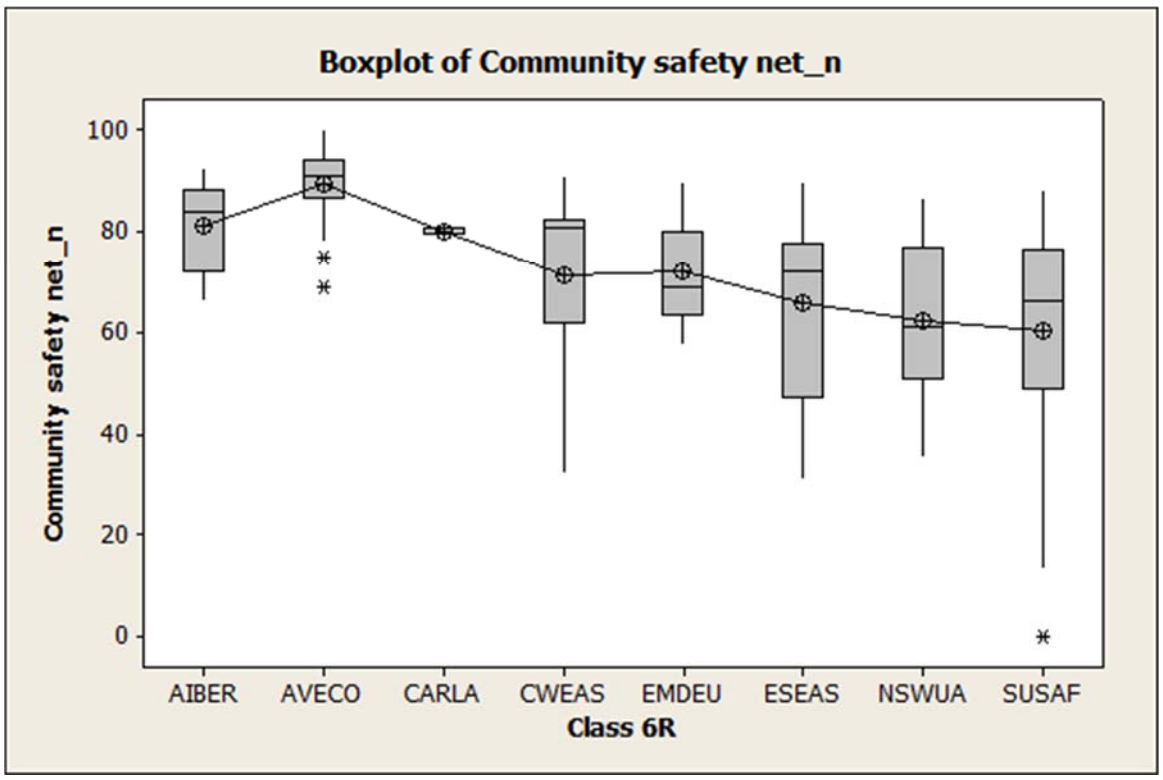
Boxplot of Religious tolerance_n



One-way ANOVA: Community safety net_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  16517  2360  10.21  0.000
Error     124  28652   231
Total     131  45169

S = 15.20   R-Sq = 36.57%   R-Sq(adj) = 32.99%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
AIBER  21  81.30   9.14                  (-----*----)
AVECO  27  89.43   6.99                        (----*---)
CARLA   3  79.90   0.85        (--------------*-------------)
CWEAS  11  71.39  17.23        (------*-------)
EMDEU  12  71.94  10.48         (------*------)
ESEAS  12  65.81  18.88    (------*------)
NSWUA  14  62.08  15.28  (------*-----)
SUSAF  32  60.11  21.68   (---*----)
                          -----+---------+---------+---------+----
                              60        72        84        96

Pooled StDev = 15.20
```

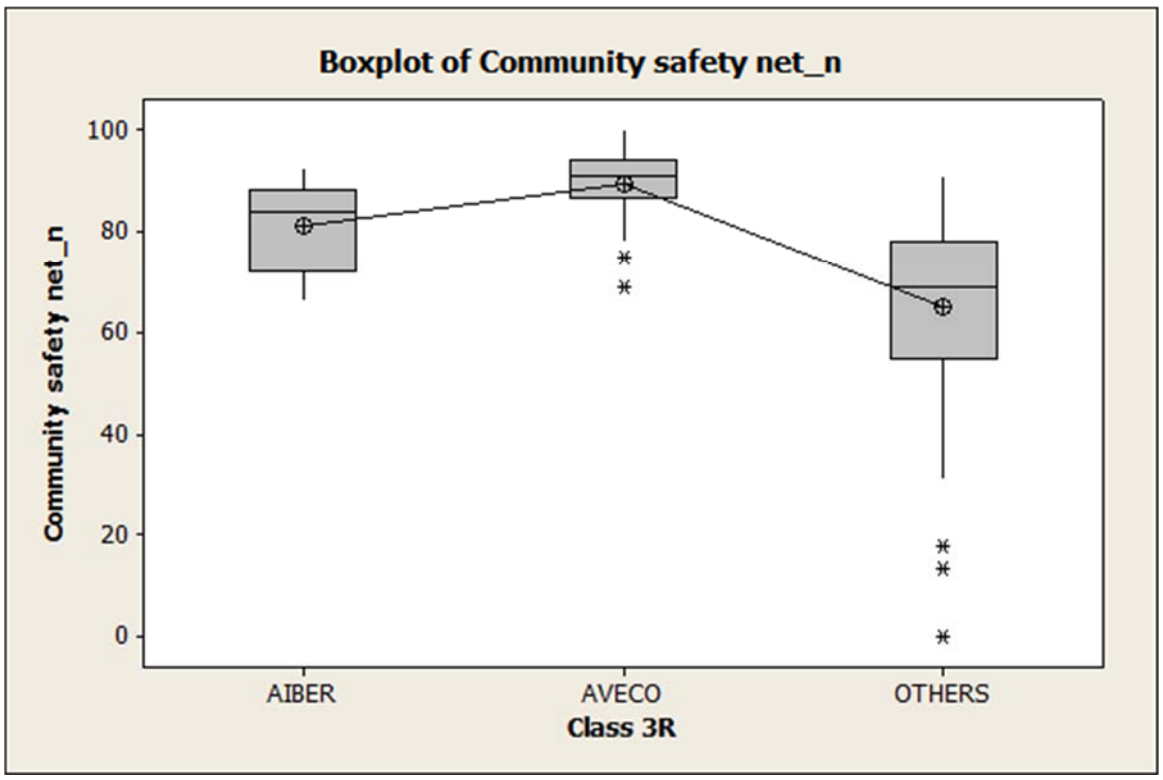
Boxplot of Community safety net_n



One-way ANOVA: Community safety net_n versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  13933  6966  28.77  0.000
Error     129  31236   242
Total     131  45169

S = 15.56   R-Sq = 30.85%   R-Sq(adj) = 29.77%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  81.30   9.14              (-----*------)
AVECO   27  89.43   6.99                       (----*-----)
OTHERS  84  65.13  18.46  (--*--)
                          --------+---------+---------+---------+-
                                 70        80        90       100

Pooled StDev = 15.56
```

Boxplot of Community safety net_n

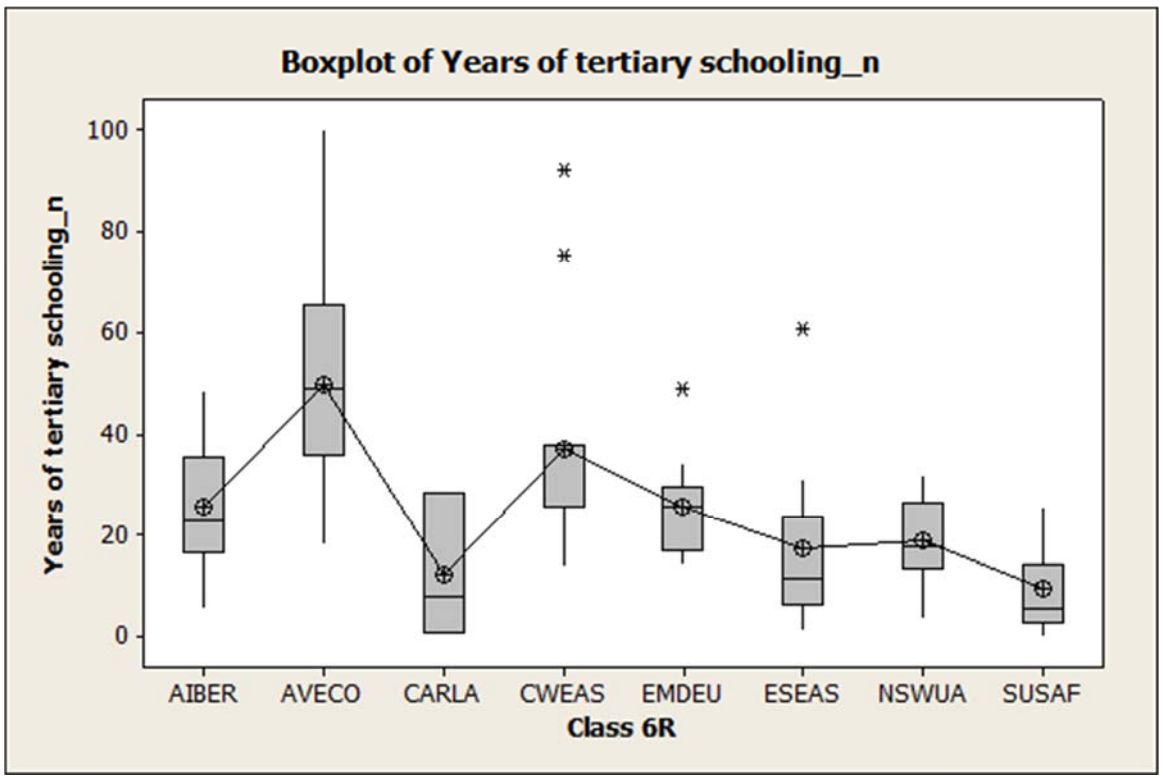


Obs: Os 2 outliers que aparecem no agrupamento AVECO, são: Grécia e Coreia do Sul (Republic of Korea).

One-way ANOVA: Years of tertiary schooling_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  27592  3942  17.54  0.000
Error     124  27861   225
Total     131  55453

S = 14.99   R-Sq = 49.76%   R-Sq(adj) = 46.92%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  21  25.74  12.61                (---*---)
AVECO  27  49.63  20.36                          (---*---)
CARLA   3  12.16  14.36  (----------*---------)
CWEAS  11  37.06  24.20                     (----*-----)
EMDEU  12  25.36   9.60               (-----*----)
ESEAS  12  17.40  16.31          (----*----)
NSWUA  14  18.86   8.65           (----*----)
SUSAF  32   9.23   9.37      (---*--)
                          ---+---------+---------+---------+------
                             0        16        32        48

Pooled StDev = 14.99
```

Boxplot of Years of tertiary schooling_n

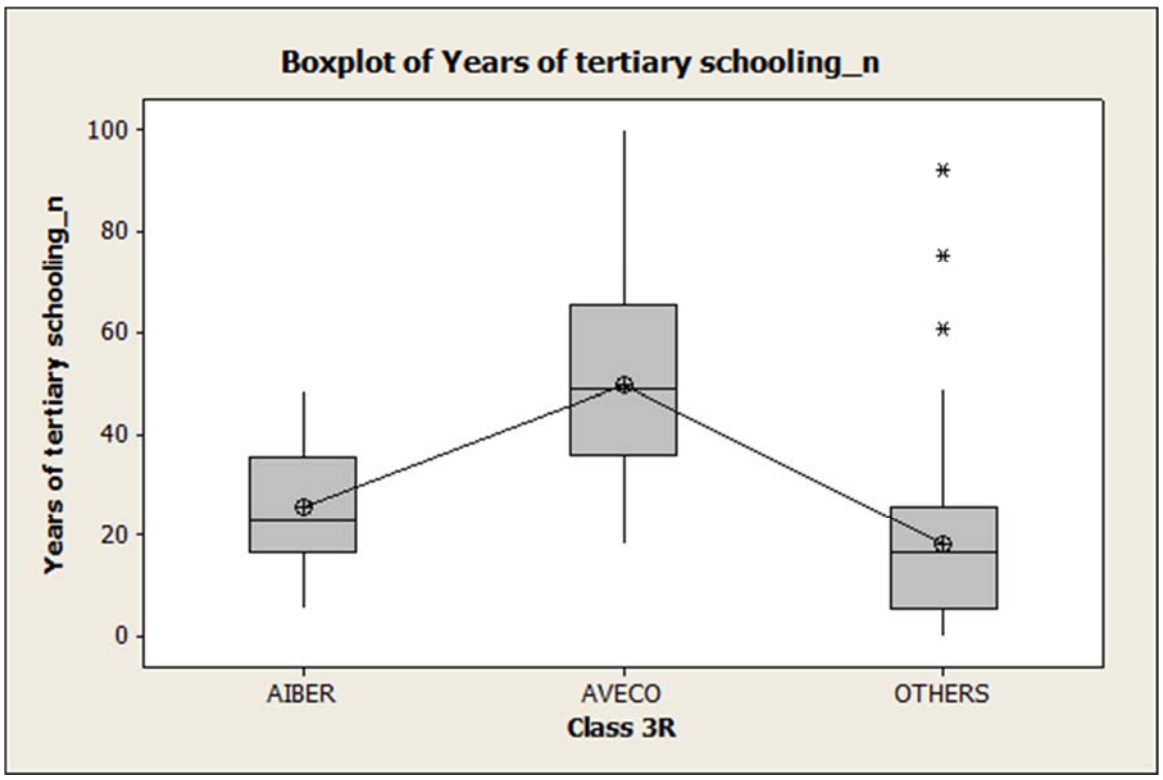


Years of tertiary schooling = 15 *missing values*

One-way ANOVA: Years of tertiary schooling_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  20370  10185  37.45  0.000
Error     129  35082    272
Total     131  55453

S = 16.49   R-Sq = 36.73%   R-Sq(adj) = 35.75%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  25.74  12.61      (----*-----)
AVECO   27  49.63  20.36                          (----*-----)
OTHERS  84  18.05  15.95  (--*--)
                          --------+---------+---------+---------+-
                                 24        36        48        60

Pooled StDev = 16.49
```

Boxplot of Years of tertiary schooling_n



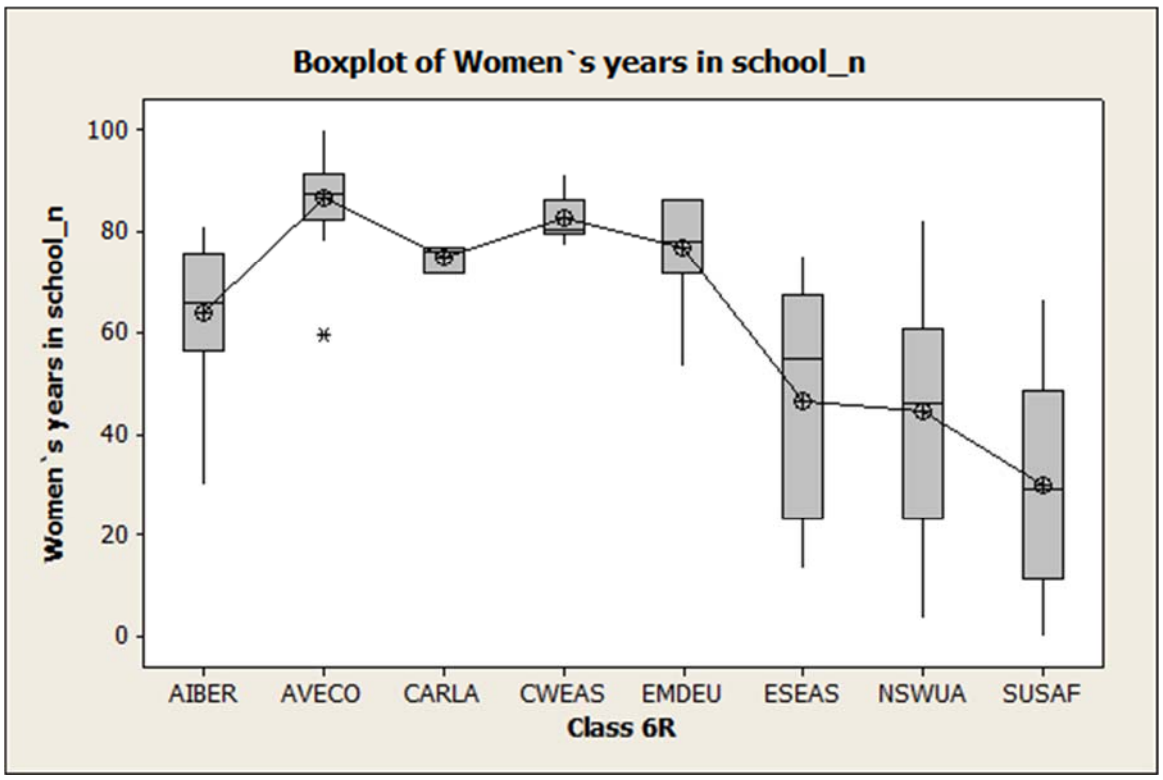
Years of tertiary schooling = 15 *missing values*

One-way ANOVA: Women`s years in school_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  64076  9154  34.46  0.000
Error     124  32939   266
Total     131  97015

S = 16.30   R-Sq = 66.05%   R-Sq(adj) = 64.13%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  21  63.61  13.02                 (---*--)
AVECO  27  86.80   7.72                           (--*---)
CARLA   3  74.70   2.76                 (--------*---------)
CWEAS  11  82.75   4.38                         (---*----)
EMDEU  12  76.89  10.31                       (---*----)
ESEAS  12  46.53  22.24        (---*----)
NSWUA  14  44.58  23.65       (---*----)
SUSAF  32  29.81  21.19  (--*--)
                          --------+---------+---------+---------+-
                                 40        60        80       100

Pooled StDev = 16.30
```

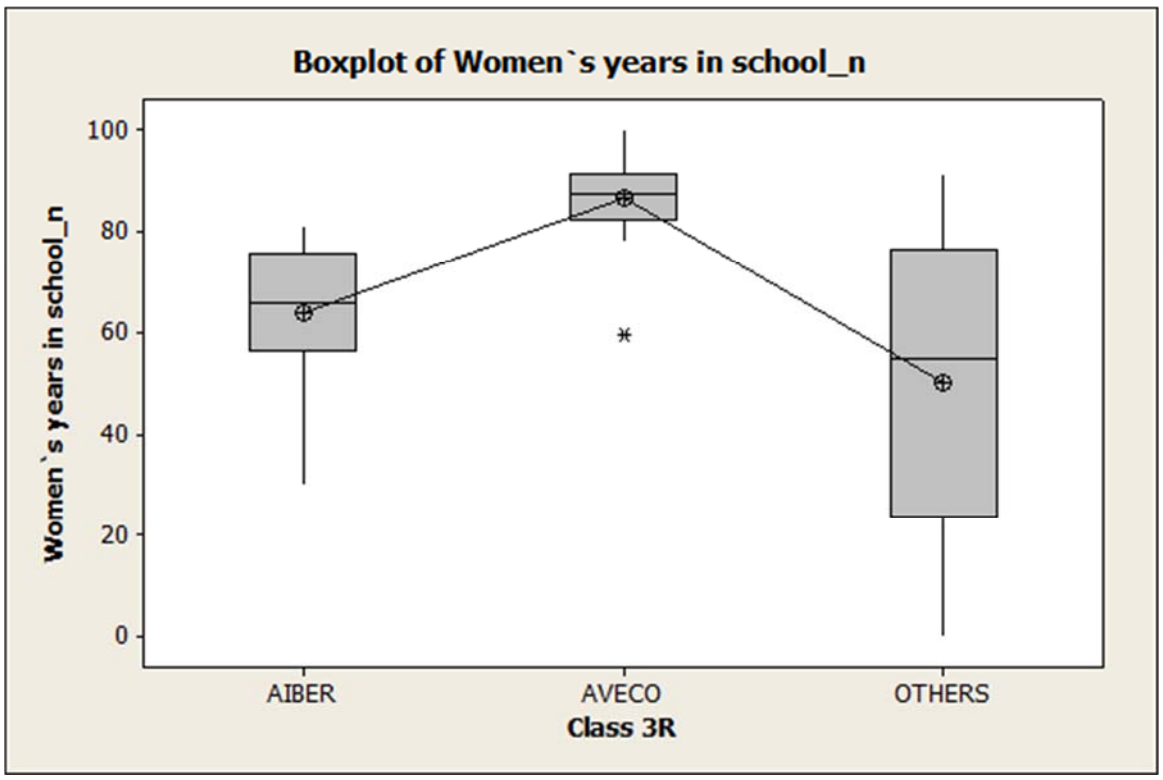
Boxplot of Women`s years in school_n



One-way ANOVA: Women`s years in school_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  28180  14090  26.41  0.000
Error     129  68835    534
Total     131  97015

S = 23.10   R-Sq = 29.05%   R-Sq(adj) = 27.95%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  63.61  13.02         (-----*------)
AVECO   27  86.80   7.72                      (-----*-----)
OTHERS  84  49.92  27.75   (--*---)
                           +---------+---------+---------+---------
                          45        60        75        90

Pooled StDev = 23.10
```

Boxplot of Women`s years in school_n

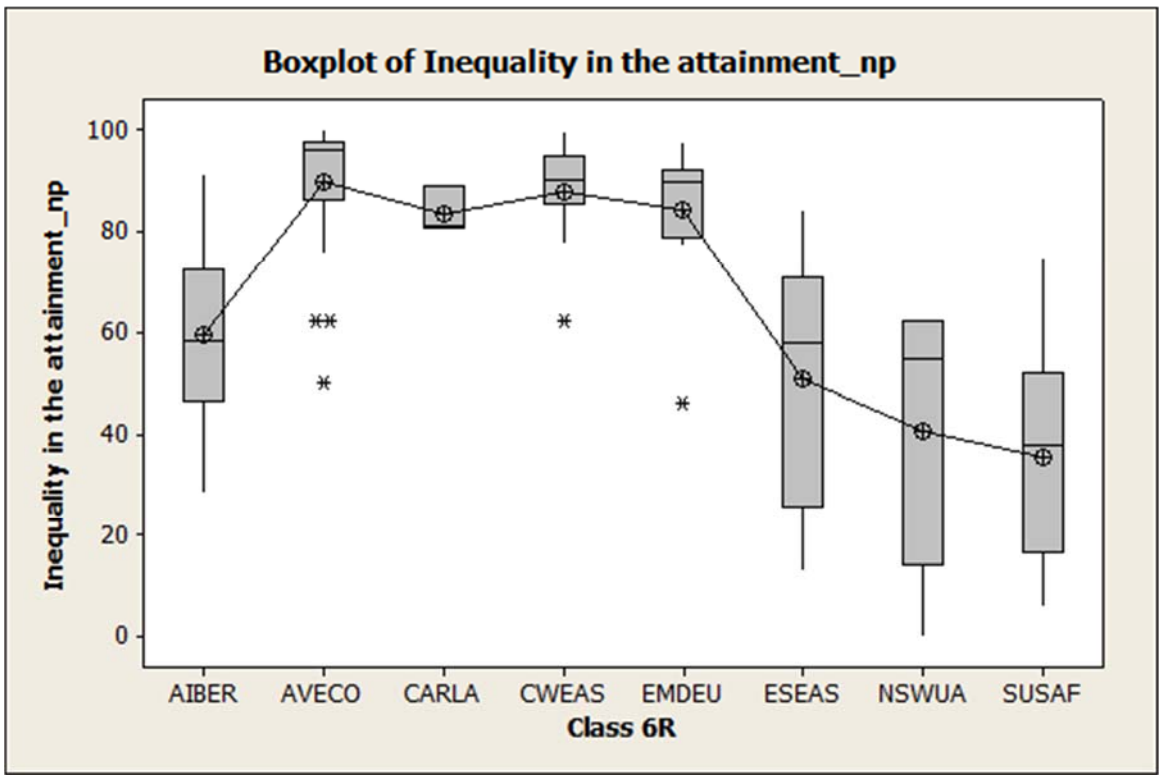


Obs: No agrupamento AVECO, a Islândia figura como *outlier* .

One-way ANOVA: Inequality in the attainment of education versus Class 6R

```
Source     DF      SS    MS      F      P
Class 6R    7   66599  9514  29.09  0.000
Error     124   40562   327
Total     131  107161

S = 18.09   R-Sq = 62.15%   R-Sq(adj) = 60.01%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
AIBER  21  59.65  17.47            (---*---)
AVECO  27  89.97  13.18                       (--*--)
CARLA   3  83.64   4.70                  (----------*---------)
CWEAS  11  87.78  10.25                        (-----*----)
EMDEU  12  84.45  13.74                       (----*----)
ESEAS  12  50.60  24.13       (----*----)
NSWUA  14  40.69  24.85   (---*----)
SUSAF  32  35.37  19.95  (--*--)
                          -----+---------+---------+---------+----
                              40        60        80       100

Pooled StDev = 18.09
```

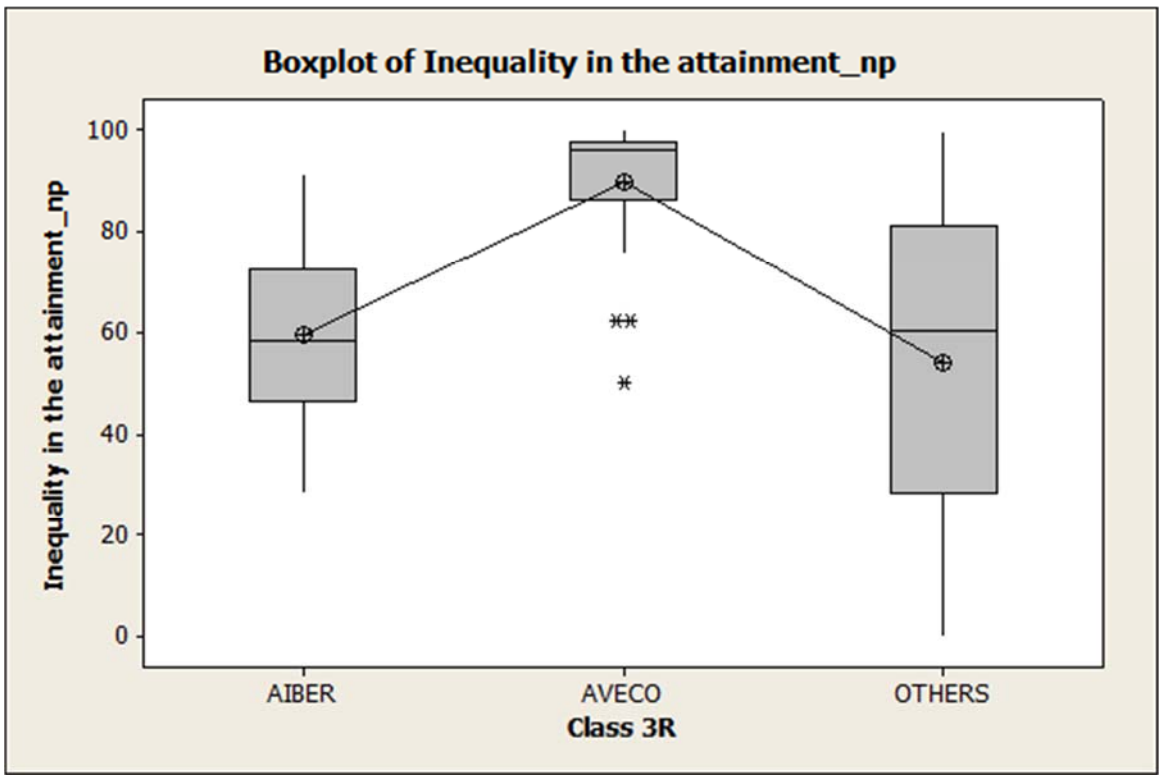
Boxplot of Inequality in the attainment of education



One-way ANOVA: Inequality in the attainment of education versus Class 3R

```
Source     DF      SS     MS      F      P
Class 3R    2   26558  13279  21.25  0.000
Error     129   80603    625
Total     131  107161

S = 25.00   R-Sq = 24.78%   R-Sq(adj) = 23.62%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  --------+---------+---------+---------+-
AIBER   21  59.65  17.47   (------*------)
AVECO   27  89.97  13.18                        (-----*-----)
OTHERS  84  54.03  29.04  (---*---)
                          --------+---------+---------+---------+-
                                 60        75        90       105

Pooled StDev = 25.00
```

Boxplot of Inequality in the attainment of education



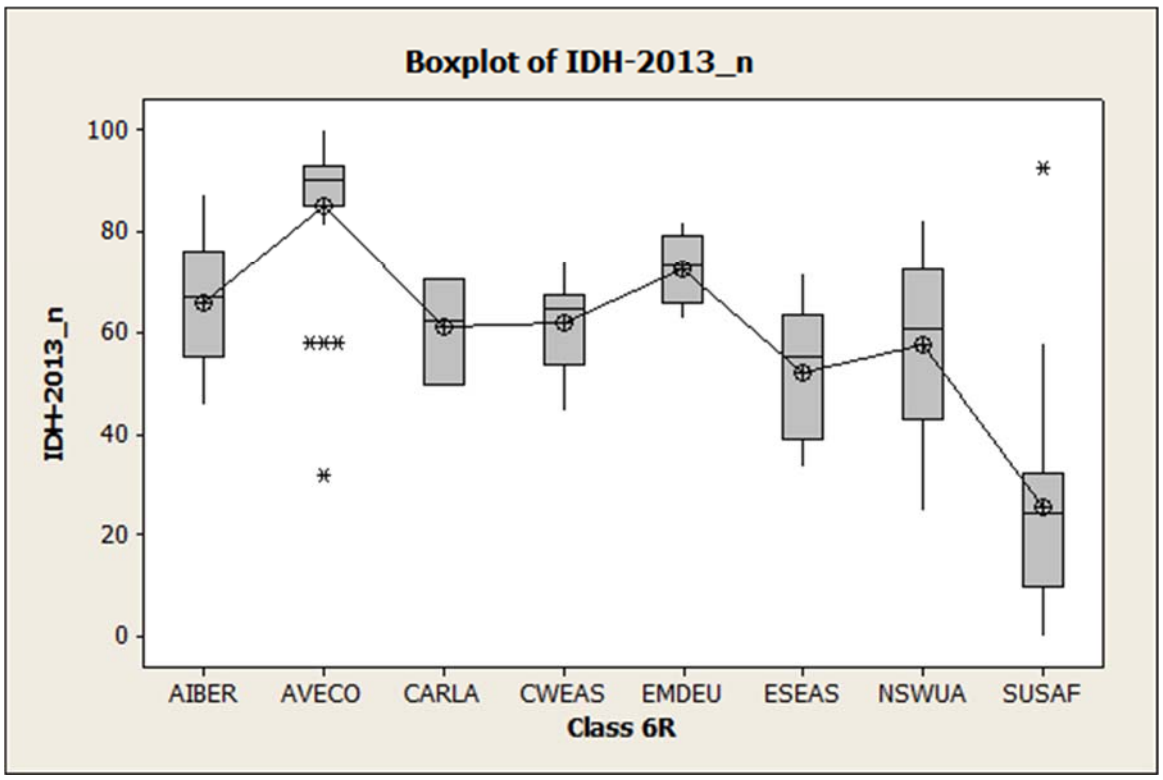
Obs: Japão, Nova Zelândia e Coreia do Sul
Inequality in the attainment of education = 15 *missing values*

One-way ANOVA: IDH-2013_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  57615  8231  34.39  0.000
Error     124  29677   239
Total     131  87292

S = 15.47   R-Sq = 66.00%   R-Sq(adj) = 64.08%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  21  65.62  12.03                    (---*--)
AVECO  27  85.11  15.53                              (--*--)
CARLA   3  60.85  10.62             (-------*--------)
CWEAS  11  61.73   9.01                 (----*---)
EMDEU  12  72.60   6.81                          (---*----)
ESEAS  12  52.16  13.52             (---*---)
NSWUA  14  57.58  19.21                (---*---)
SUSAF  32  25.69  19.80   (--*--)
                          +---------+---------+---------+---------
                          20        40        60        80

Pooled StDev = 15.47
```

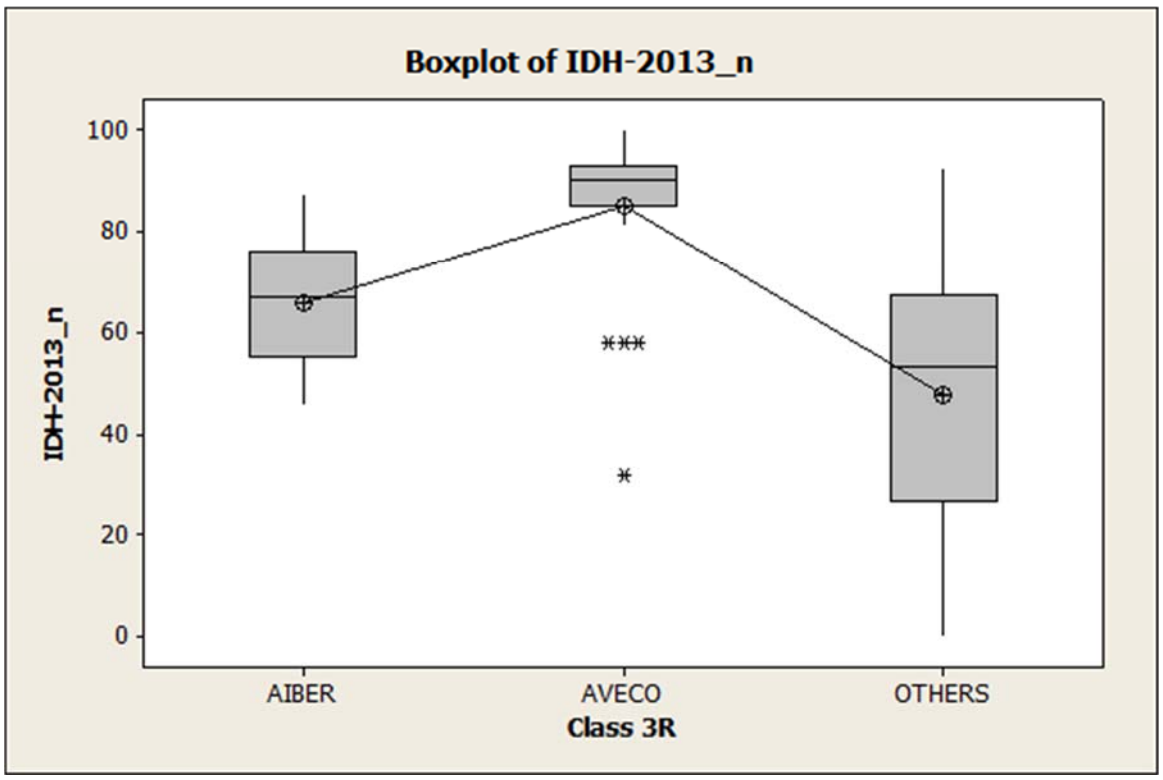
Boxplot of IDH-2013_n



One-way ANOVA: IDH-2013_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  30392  15196  34.45  0.000
Error     129  56900    441
Total     131  87292

S = 21.00   R-Sq = 34.82%   R-Sq(adj) = 33.81%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65.62  12.03           (-----*-----)
AVECO   27  85.11  15.53                       (-----*----)
OTHERS  84  47.46  23.98  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 21.00
```

Boxplot of IDH-2013_n



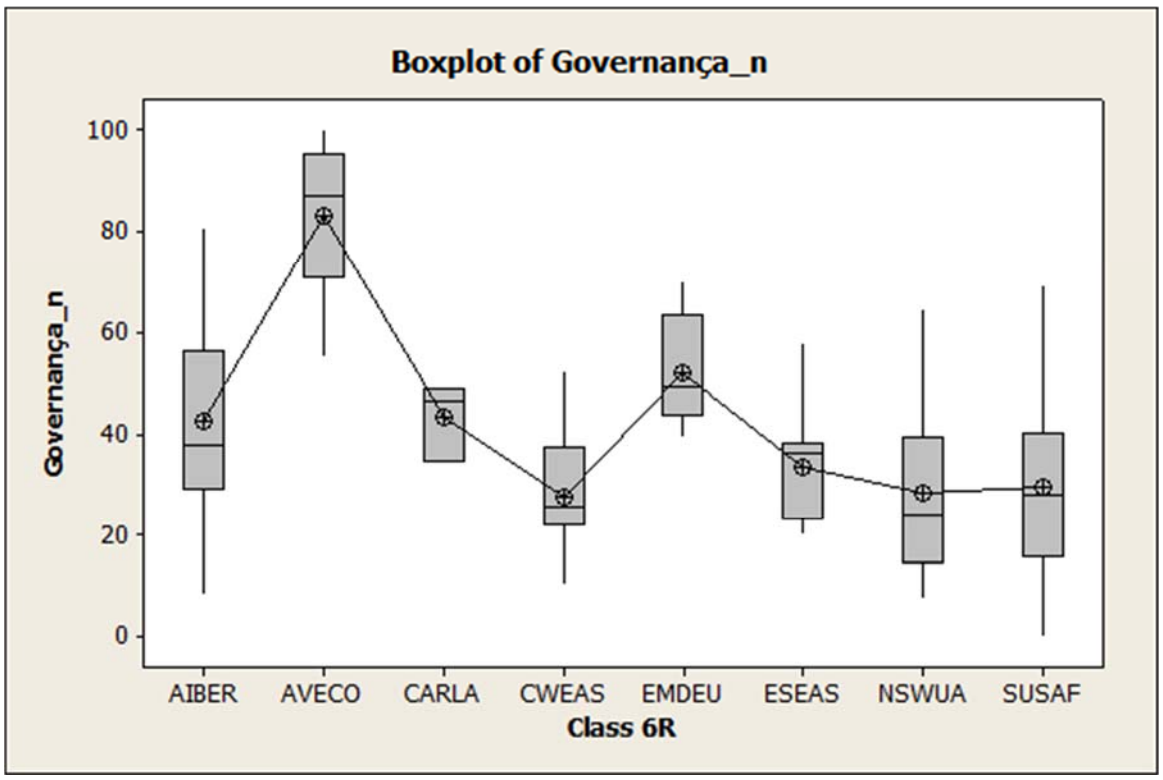
Obs: Coreia do Sul (Republic of Korea), Letônia (Latvia), Holanda e Suíça figura como *outliers* no gráfico acima.
IDH-2013 = 7 *missing values*
*Coreia, Holanda e Letônia estavam sem valores na planilha. E Suíça com 0.53 na planilha original

One-way ANOVA: Governança_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56636  8091  35.45  0.000
Error     124  28301   228
Total     131  84938

S = 15.11   R-Sq = 66.68%   R-Sq(adj) = 64.80%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  21  42.48  18.82            (--*---)
AVECO  27  83.25  13.38                                (--*--)
CARLA   3  43.25   7.61       (--------*-------)
CWEAS  11  27.38  12.38   (----*---)
EMDEU  12  51.93  11.21                 (---*---)
ESEAS  12  33.61  10.47      (----*---)
NSWUA  14  28.46  16.21    (---*---)
SUSAF  32  29.58  16.95      (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15.11
```

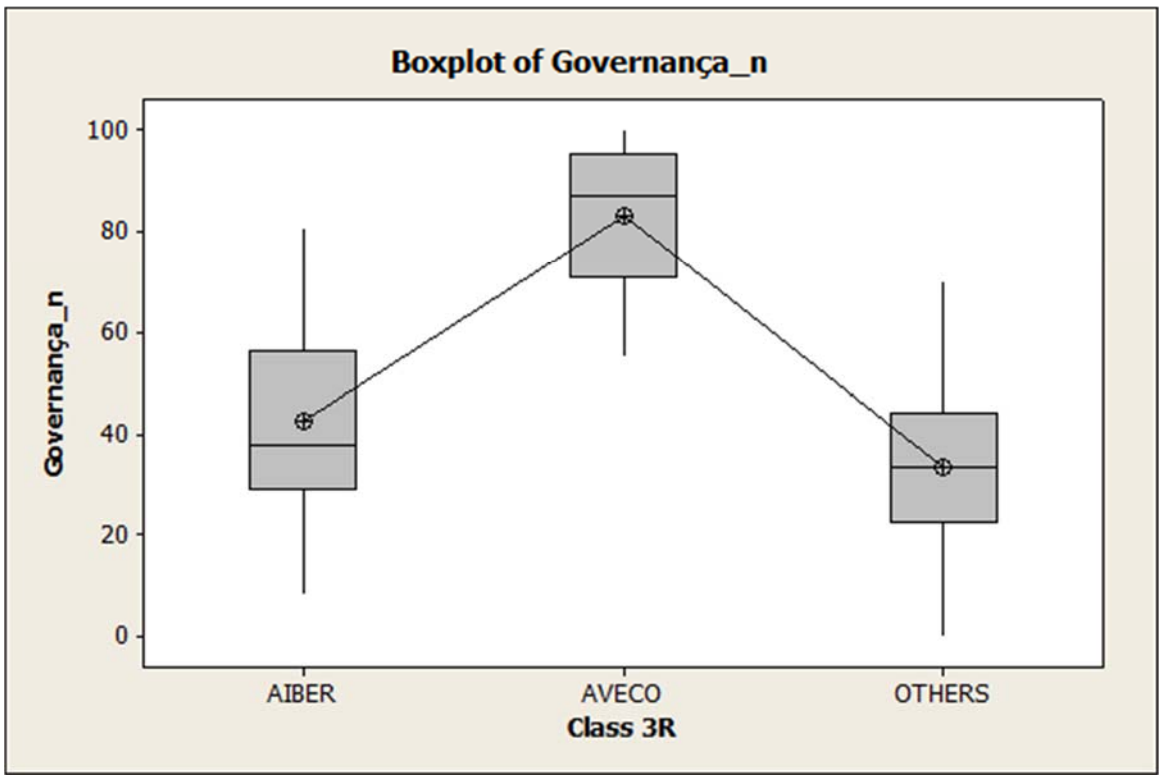
Boxplot of Governança_n



One-way ANOVA: Governança_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  51019  25510  97.02  0.000
Error     129  33919    263
Total     131  84938

S = 16.22   R-Sq = 60.07%   R-Sq(adj) = 59.45%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42.48  18.82       (---*----)
AVECO   27  83.25  13.38                                  (----*---)
OTHERS  84  33.36  16.35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75

Pooled StDev = 16.22
```

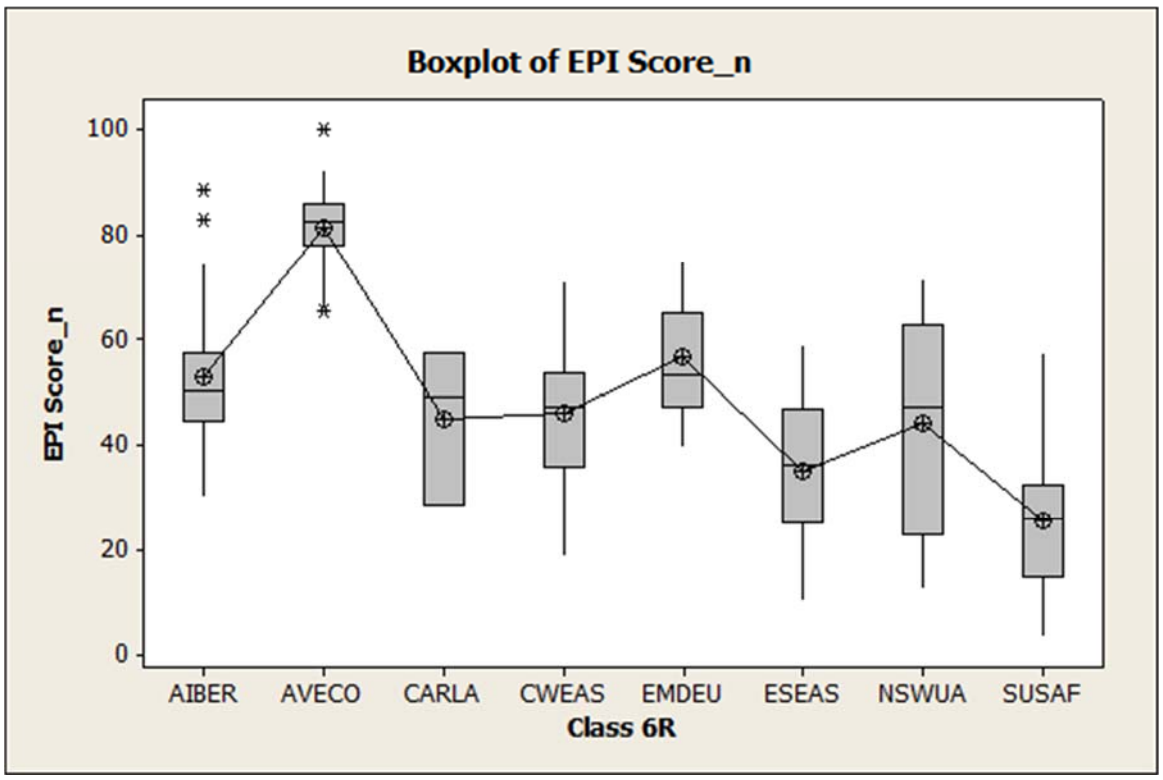
Boxplot of Governança_n



One-way ANOVA: EPI Score_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  49668  7095  39.04  0.000
Error     124  22535   182
Total     131  72203

S = 13.48   R-Sq = 68.79%   R-Sq(adj) = 67.03%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  21  52.70  14.31               (--*--)
AVECO  27  81.39   8.14                                (--*-)
CARLA   3  44.93  14.98         (------*-------)
CWEAS  11  46.14  14.39             (---*---)
EMDEU  12  56.69  11.20                  (---*---)
ESEAS  12  34.97  13.95       (--*---)
NSWUA  14  44.29  20.68             (--*---)
SUSAF  32  25.60  12.85   (--*-)
                          +---------+---------+---------+---------
                          20        40        60        80

Pooled StDev = 13.48
```

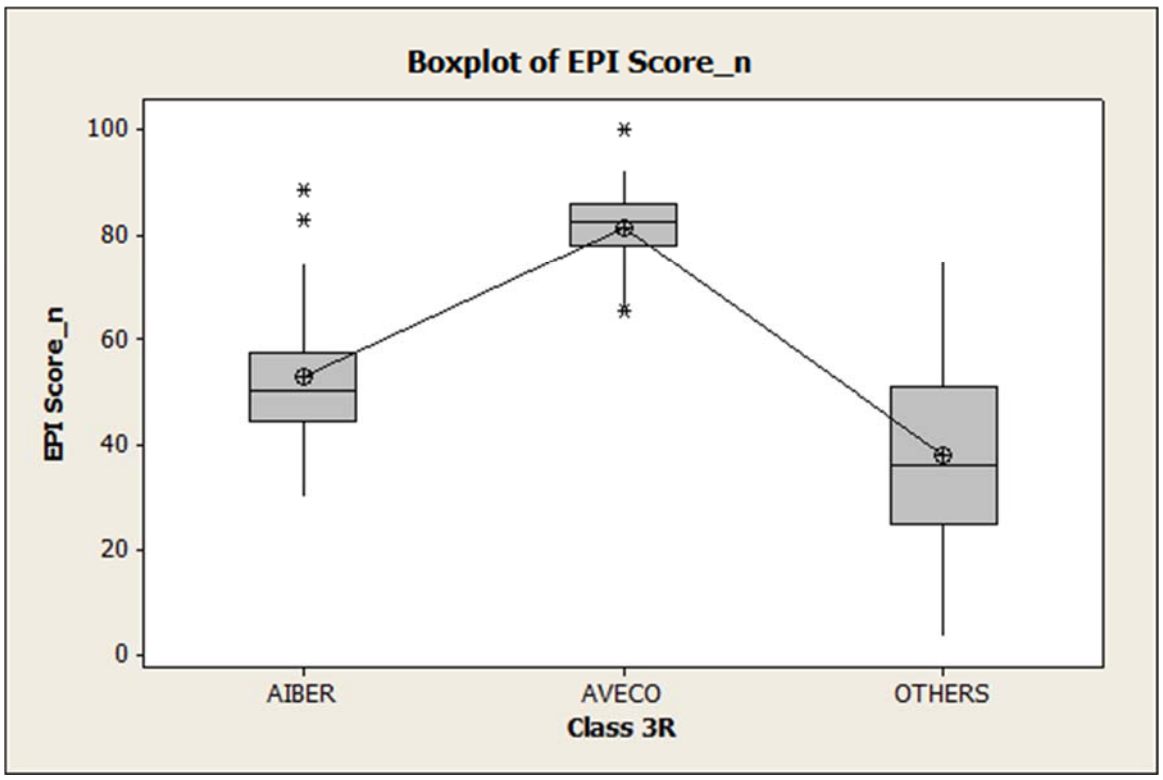
Boxplot of EPI Score_n



One-way ANOVA: EPI Score_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  39017  19508  75.83  0.000
Error     129  33186    257
Total     131  72203

S = 16.04   R-Sq = 54.04%   R-Sq(adj) = 53.33%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  52.70  14.31         (---*----)
AVECO   27  81.39   8.14                            (---*---)
OTHERS  84  37.87  18.16  (-*--)
                          -------+---------+---------+---------+--
                                45        60        75        90

Pooled StDev = 16.04
```

Boxplot of EPI Score_n

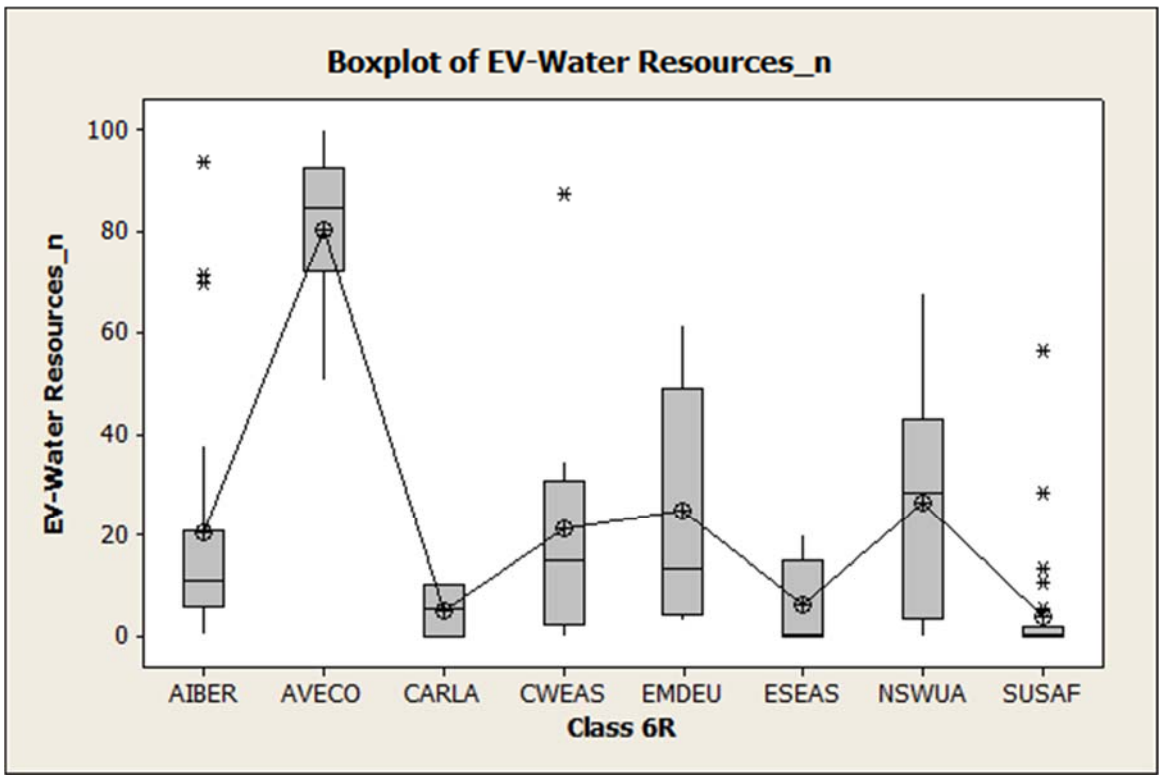


Obs: Destacaram-se positiviamente como *outliers* os países Portugal e Espanha no agrupamento AIBER e Suíça no AVECO. E negativamente como *outlier* a Coreia do Sul (Republic of Korea).

One-way ANOVA: EV-Water Resources_n versus Class 6R

```
Source     DF      SS     MS      F      P
Class 6R    7  101460  14494  42.74  0.000
Error     124   42049    339
Total     131  143509

S = 18.41   R-Sq = 70.70%   R-Sq(adj) = 69.05%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
AIBER  21  20.39  26.00          (--*-)
AVECO  27  80.29  14.88                              (--*-)
CARLA   3   5.14   5.06  (------*------)
CWEAS  11  21.11  25.13         (---*---)
EMDEU  12  24.74  23.16           (--*---)
ESEAS  12   6.21   8.14      (--*---)
NSWUA  14  26.19  21.49           (---*--)
SUSAF  32   3.99  11.05      (-*-)
                          -----+---------+---------+---------+----
                               0        30        60        90

Pooled StDev = 18.41
```

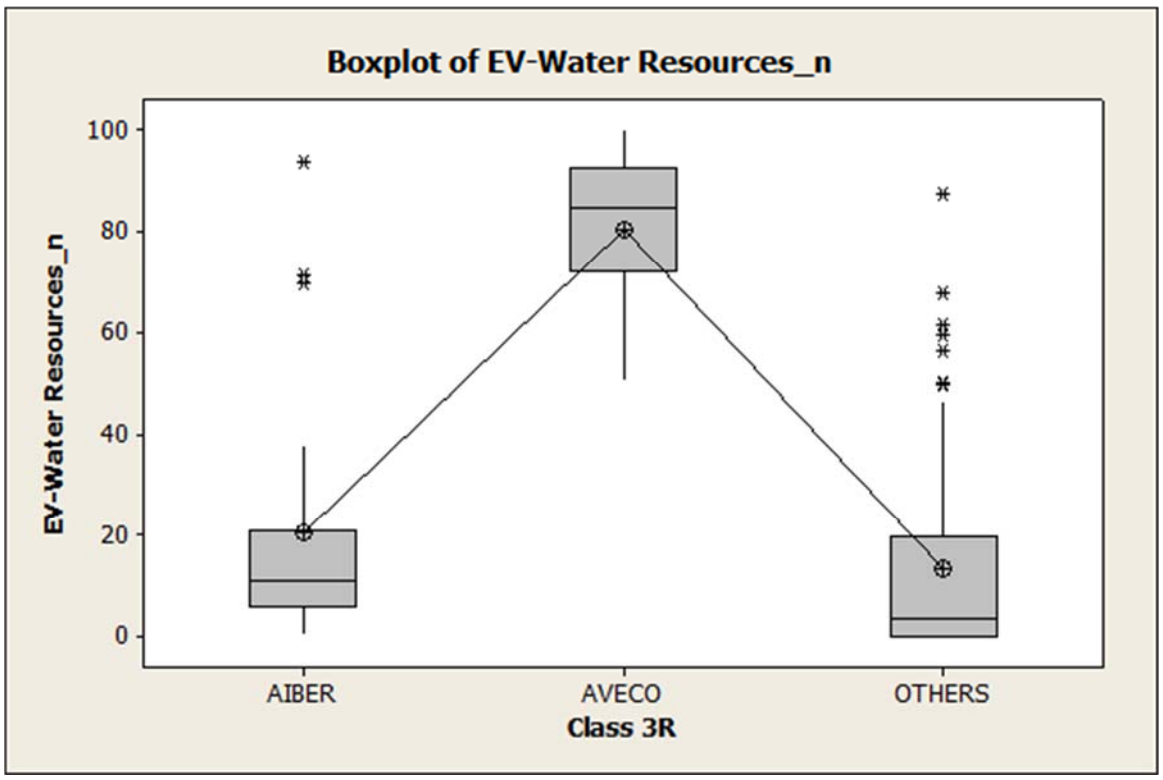
Boxplot of EV-Water Resources_n



One-way ANOVA: EV-Water Resources_n versus Class 3R

```
Source     DF      SS     MS       F      P
Class 3R    2   93314  46657  119.91  0.000
Error     129   50195    389
Total     131  143509

S = 19.73   R-Sq = 65.02%   R-Sq(adj) = 64.48%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  20.39  26.00    (---*---)
AVECO   27  80.29  14.88                                (---*---)
OTHERS  84  13.26  19.30  (--*-)
                          ------+---------+---------+---------+---
                               20        40        60        80

Pooled StDev = 19.73
```

Boxplot of EV-Water Resources_n



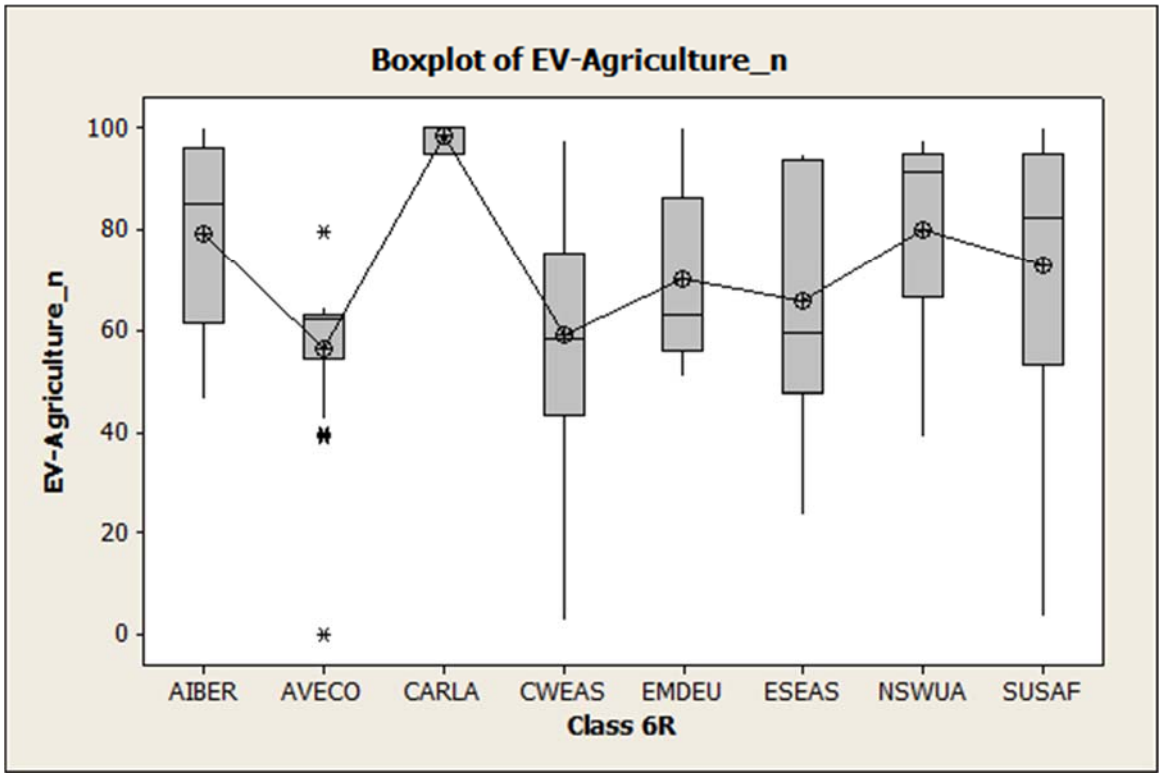
Obs: *Outliers* que se destacaram positivamente foram Chile, Portugal e Espanha no agrupamento AIBER

One-way ANOVA: EV-Agriculture_n versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  12631  1804  3.81  0.001
Error     124  58700   473
Total     131  71330

S = 21.76   R-Sq = 17.71%   R-Sq(adj) = 13.06%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
AIBER  21  79.27  19.33              (----*---)
AVECO  27  56.11  14.68   (---*---)
CARLA   3  98.36   2.84                (-----------*------------)
CWEAS  11  58.88  27.84  (-----*------)
EMDEU  12  70.12  17.49        (-----*-----)
ESEAS  12  65.71  24.75      (-----*-----)
NSWUA  14  79.92  20.98             (-----*-----)
SUSAF  32  72.85  26.66           (--*---)
                          -------+---------+---------+---------+--
                                60        80       100       120

Pooled StDev = 21.76
```

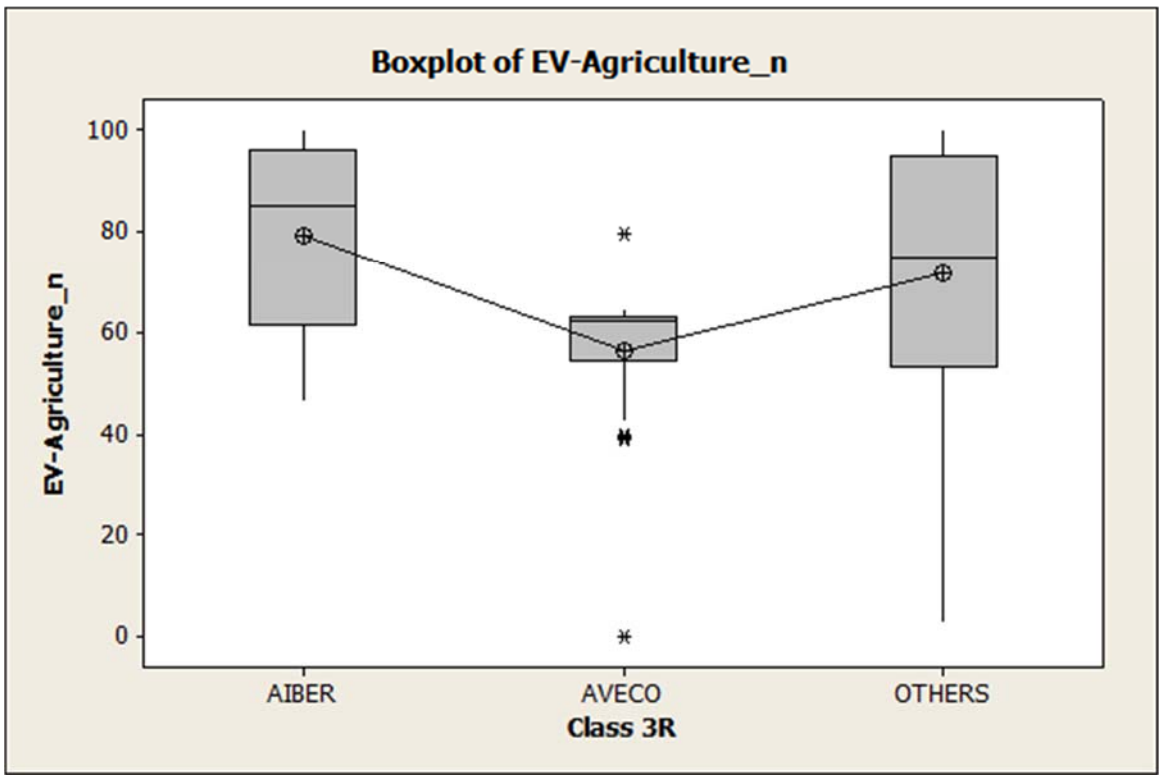
Boxplot of EV-Agriculture_n



One-way ANOVA: EV-Agriculture_n versus Class 3R

```
Source     DF     SS    MS     F      P
Class 3R    2   7241  3621  7.29  0.001
Error     129  64089   497
Total     131  71330

S = 22.29   R-Sq = 10.15%   R-Sq(adj) = 8.76%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  79.27  19.33                    (-------*-------)
AVECO   27  56.11  14.68   (------*------)
OTHERS  84  71.70  24.79                  (---*---)
                           +---------+---------+---------+---------
                          48        60        72        84

Pooled StDev = 22.29
```

Boxplot of EV-Agriculture_n

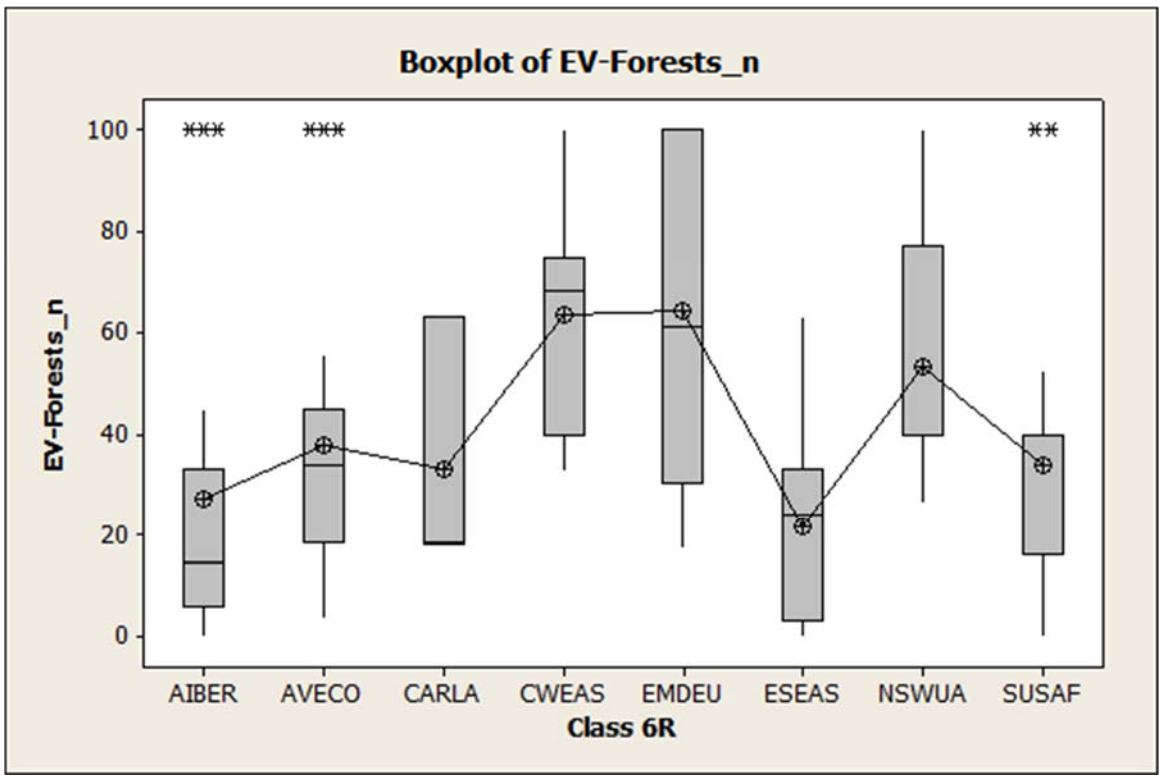


Obs: Os *outliers* no agrupamento AVECO em ordem crescente são: Israel, Japão, Coreia do Sul (Republic of Korea) e Nova Zelândia.

One-way ANOVA: EV-Forests_n versus Class 6R

```
Source     DF      SS    MS     F      P
Class 6R    7   24394  3485  5.10  0.000
Error     124   84801   684
Total     131  109195

S = 26.15   R-Sq = 22.34%   R-Sq(adj) = 17.96%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  21  27.24  32.87        (-----*----)
AVECO  27  37.70  26.49              (----*----)
CARLA   3  33.17  25.94  (--------------*--------------)
CWEAS  11  63.28  21.11                       (-------*------)
EMDEU  12  64.15  32.22                        (------*-------)
ESEAS  12  21.63  18.51   (-------*------)
NSWUA  14  53.29  24.94                     (------*------)
SUSAF  32  33.83  22.63            (----*---)
                          --------+---------+---------+---------+-
                                  20        40        60        80

Pooled StDev = 26.15
```

Boxplot of EV-Forests_n

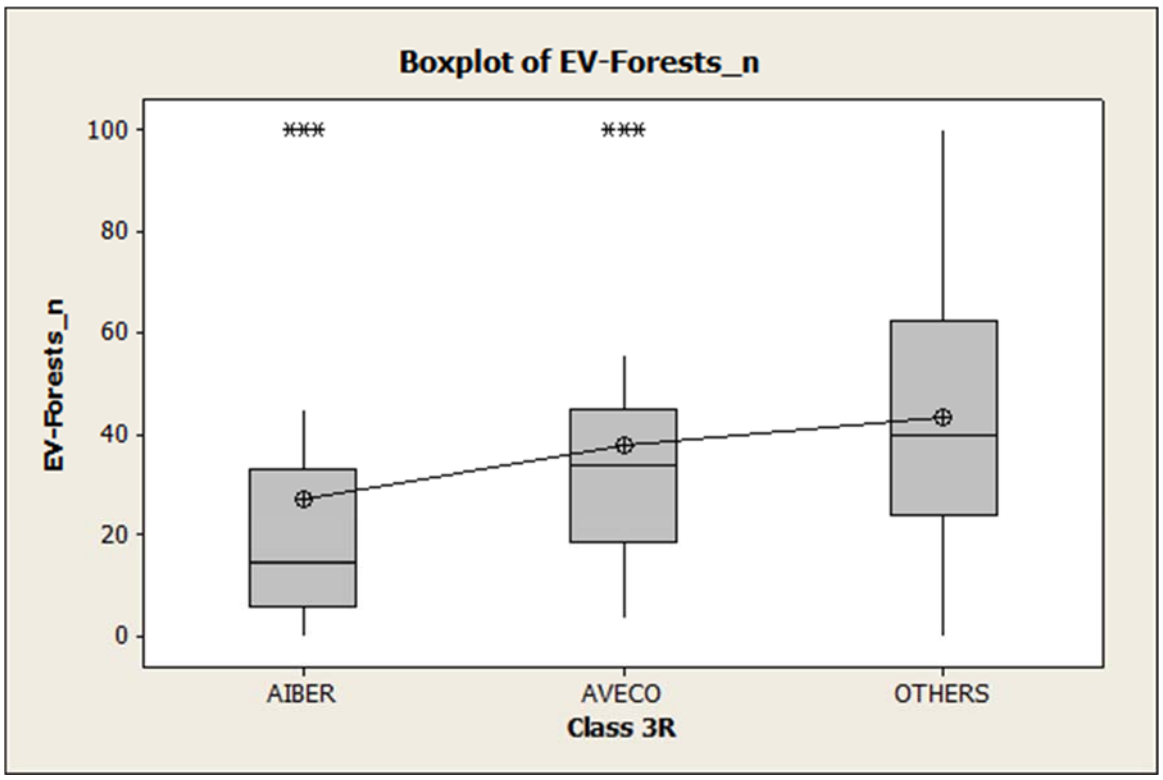


EV-Forests = 22 *missing values*

One-way ANOVA: EV-Forests_n versus Class 3R

```
Source     DF      SS    MS     F      P
Class 3R    2    4577  2288  2.82  0.063
Error     129  104619   811
Total     131  109195

S = 28.48   R-Sq = 4.19%   R-Sq(adj) = 2.71%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  27.24  32.87  (-----------*------------)
AVECO   27  37.70  26.49             (----------*----------)
OTHERS  84  43.49  27.93                       (-----*------)
                          -----+---------+---------+---------+----
                              20        30        40        50

Pooled StDev = 28.48
```

Boxplot of EV-Forests_n

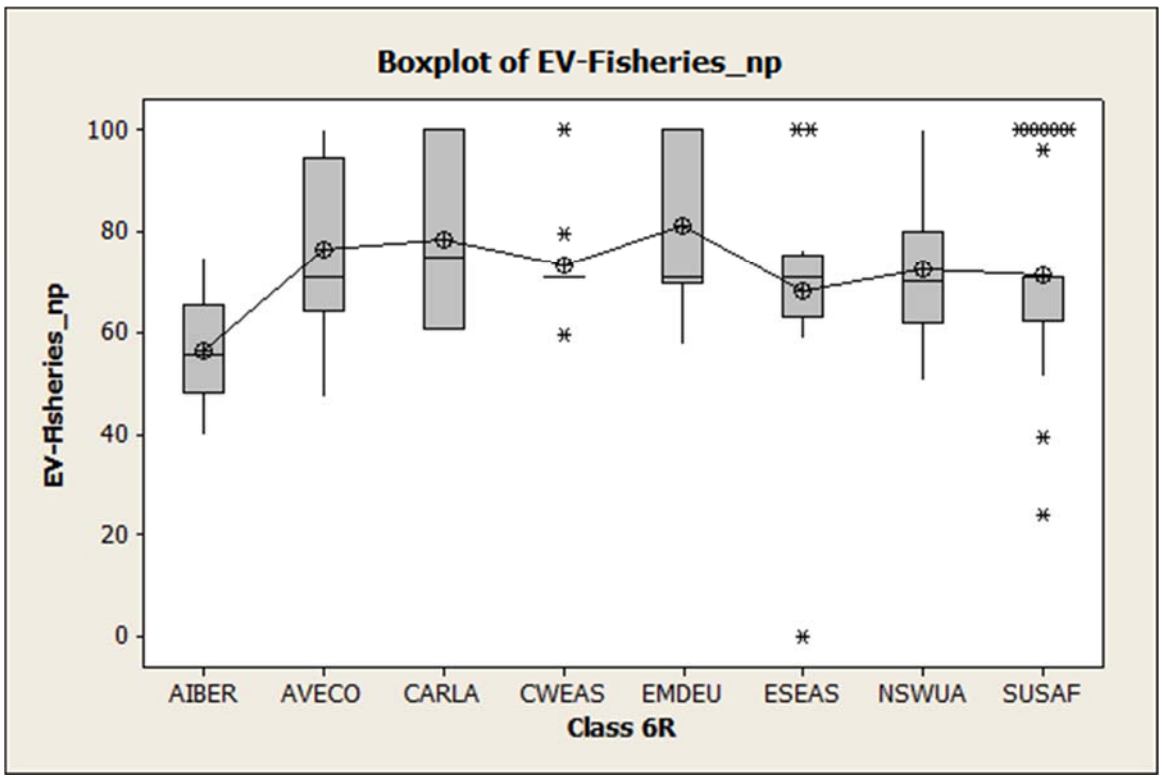


EV-Forests = 22 *missing values*

One-way ANOVA: EV-Fisheries_np versus Class 6R

```
Source     DF     SS   MS     F      P
Class 6R    7   6872  982  3.54  0.002
Error     124  34393  277
Total     131  41265

S = 16.65   R-Sq = 16.65%   R-Sq(adj) = 11.95%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -------+---------+---------+---------+--
AIBER  21  56.39  10.71   (----*---)
AVECO  27  76.24  16.42               (---*---)
CARLA   3  78.45  19.95         (-----------*-----------)
CWEAS  11  73.39   9.93           (------*------)
EMDEU  12  81.30  16.93                 (-----*------)
ESEAS  12  68.01  25.08        (-----*------)
NSWUA  14  72.57  14.82            (----*-----)
SUSAF  32  71.32  18.30             (---*--)
                          -------+---------+---------+---------+--
                                60        75        90       105

Pooled StDev = 16.65
```

Boxplot of EV-Fisheries_np

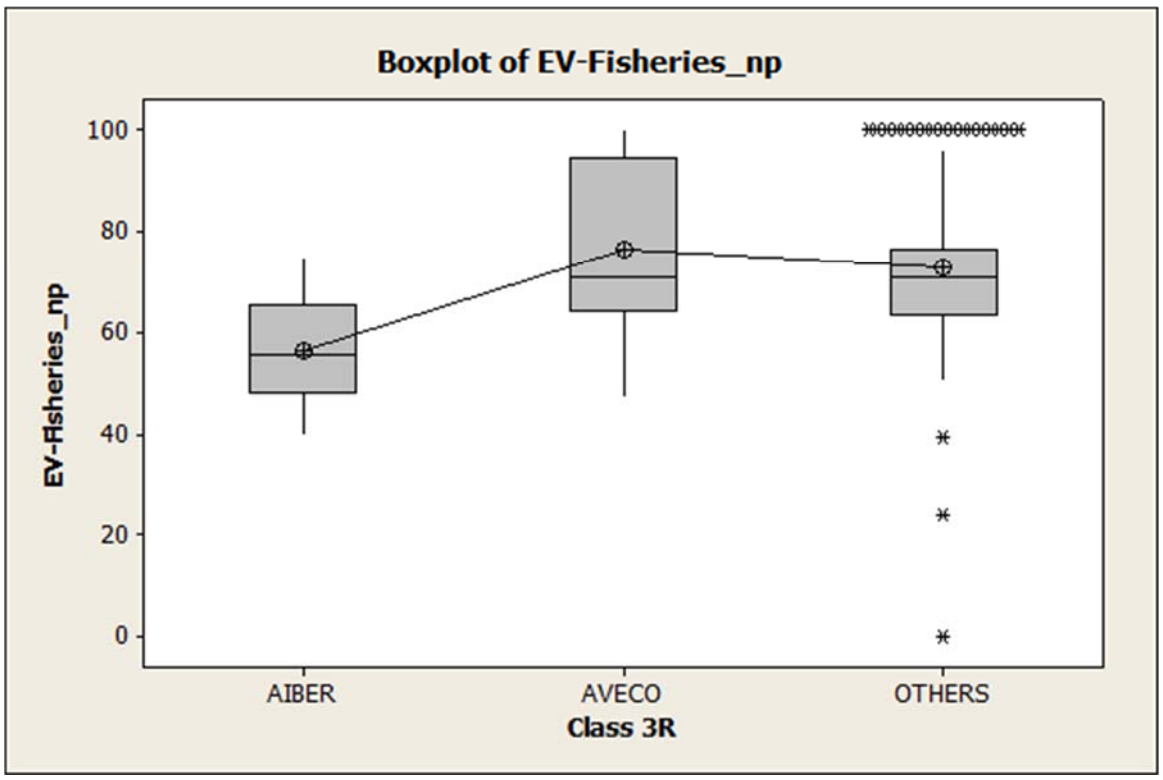


EV-Fisheries = 37 *missing values*

One-way ANOVA: EV-Fisheries_np versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2   5563  2781  10.05  0.000
Error     129  35702   277
Total     131  41265

S = 16.64   R-Sq = 13.48%   R-Sq(adj) = 12.14%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  56.39  10.71  (------*-------)
AVECO   27  76.24  16.42                      (-----*------)
OTHERS  84  73.01  17.83                     (---*---)
                          -+---------+---------+---------+--------
                          50        60        70        80

Pooled StDev = 16.64
```

Boxplot of EV-Fisheries_np

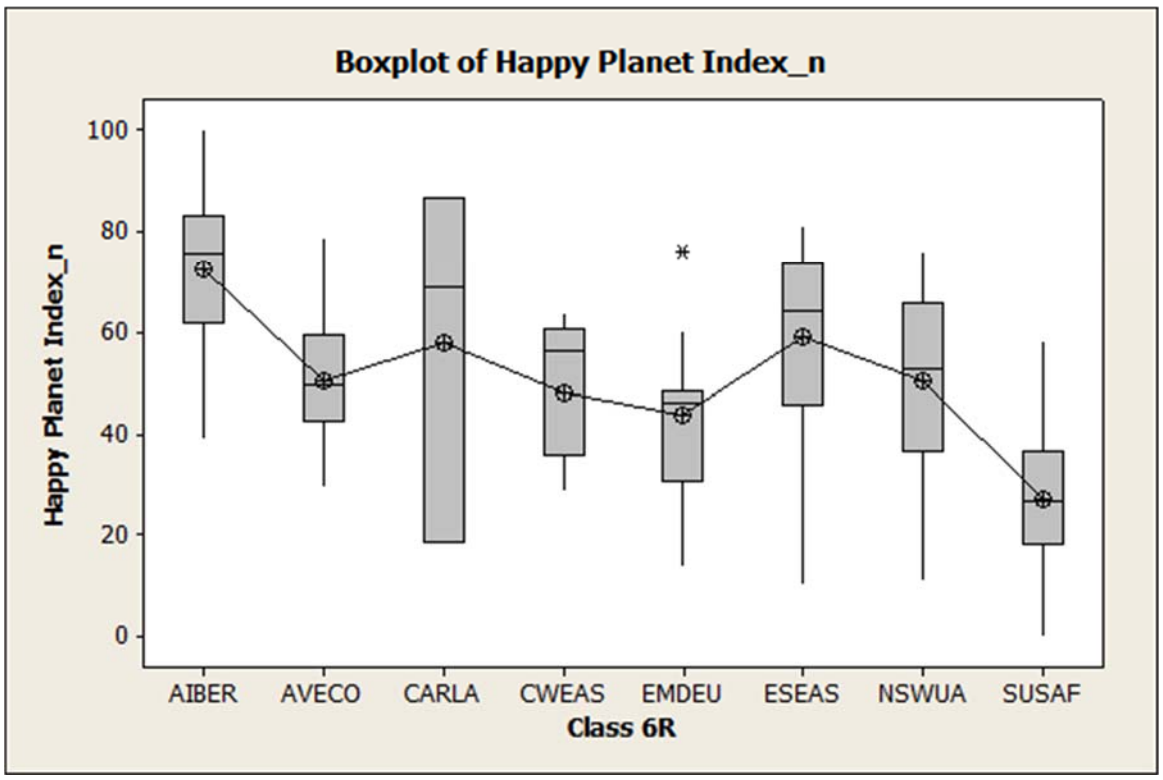


EV-Fisheries = 37 *missing values*

One-way ANOVA: Happy Planet Index_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  28446  4064  15.53  0.000
Error     124  32448   262
Total     131  60894

S = 16.18   R-Sq = 46.71%   R-Sq(adj) = 43.71%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  21  72.36  16.25                                (---*----)
AVECO  27  50.58  12.84                   (---*---)
CARLA   3  58.07  35.38              (------------*-----------)
CWEAS  11  48.13  13.54              (-----*------)
EMDEU  12  43.56  16.02           (-----*-----)
ESEAS  12  59.06  20.02                     (-----*------)
NSWUA  14  50.22  19.91                (----*-----)
SUSAF  32  27.27  14.09  (---*---)
                          ------+---------+---------+---------+---
                               30        45        60        75

Pooled StDev = 16.18
```

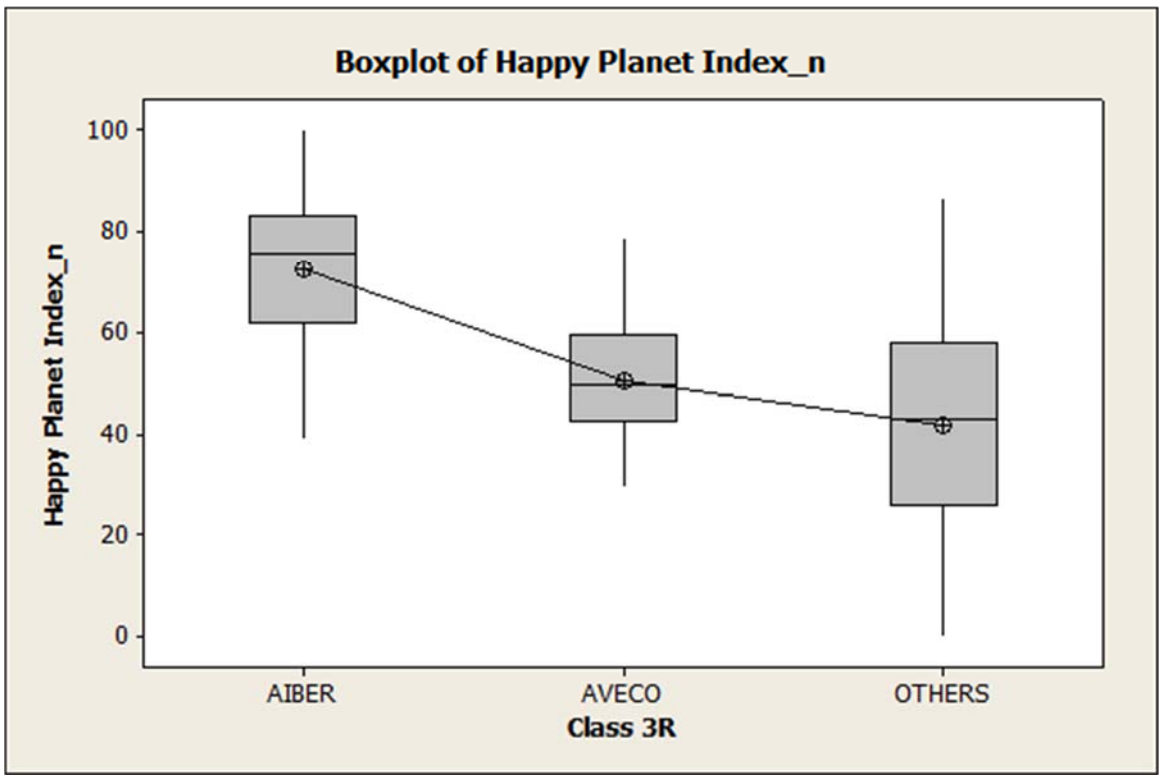
Boxplot of Happy Planet Index_n



One-way ANOVA: Happy Planet Index_n versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  15846  7923  22.69  0.000
Error     129  45048   349
Total     131  60894

S = 18.69   R-Sq = 26.02%   R-Sq(adj) = 24.87%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72.36  16.25                        (-----*------)
AVECO   27  50.58  12.84       (-----*-----)
OTHERS  84  41.79  20.68  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18.69
```

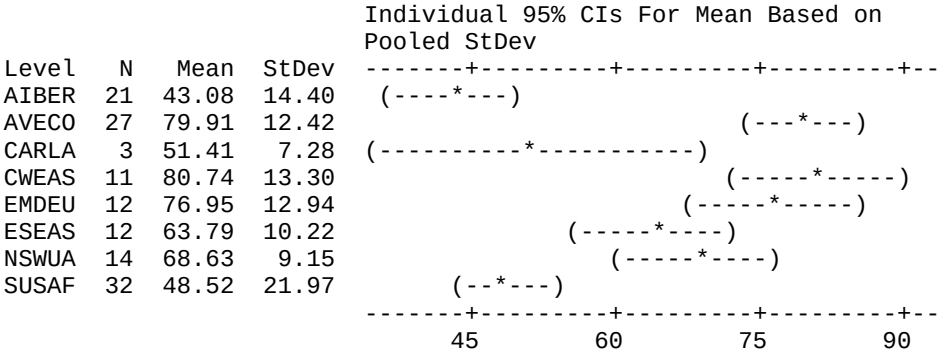
Boxplot of Happy Planet Index_n



Obs: O *outlier* desapareceu e a conformação dos 3 agrupamentos permaneceram bem similares em relação ao gráfica anterior (class 6R).

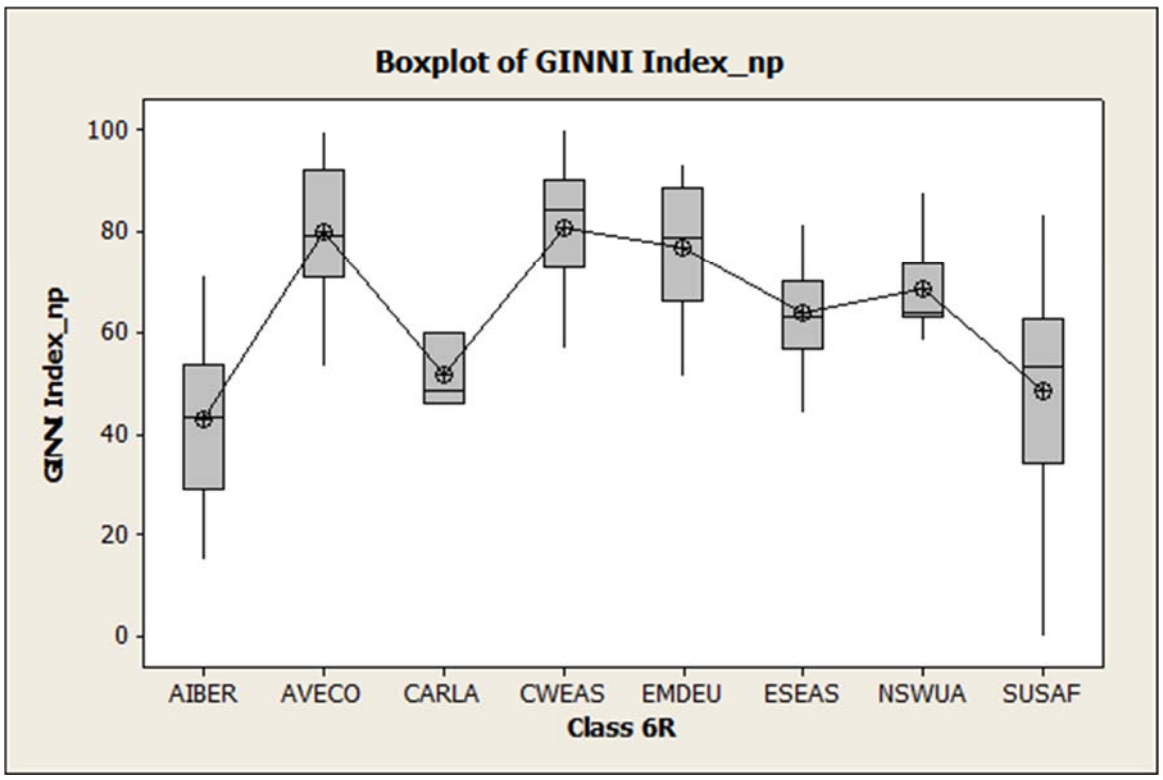
One-way ANOVA: GINNI Index_np versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  29417  4202  17.92  0.000
Error     124  29075   234
Total     131  58492

S = 15.31   R-Sq = 50.29%   R-Sq(adj) = 47.49%
```




```
Pooled StDev = 15.31
```

Boxplot of GINNI Index_np

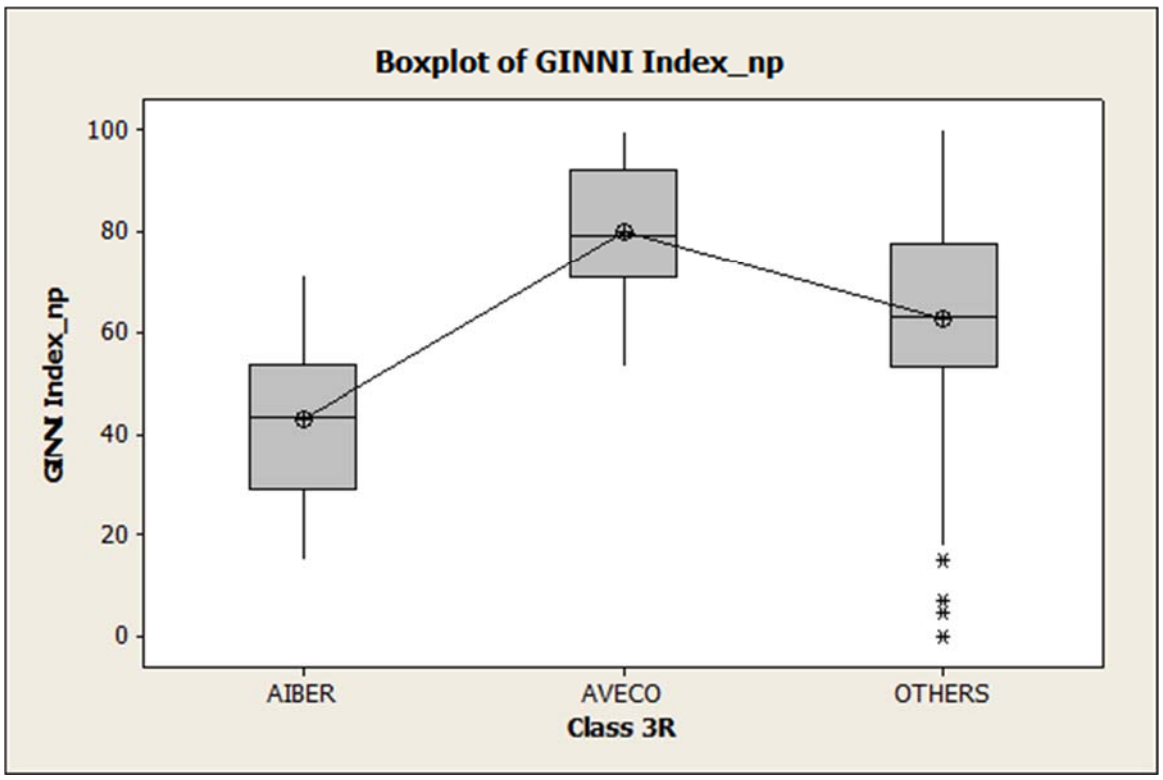


Obs: GINNI Index = 13 *missing values*

One-way ANOVA: GINNI Index_np versus Class 3R

```
Source     DF     SS    MS      F      P
Class 3R    2  16078  8039  24.45  0.000
Error     129  42414   329
Total     131  58492

S = 18.13   R-Sq = 27.49%   R-Sq(adj) = 26.36%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  43.08  14.40  (----*----)
AVECO   27  79.91  12.42                       (---*----)
OTHERS  84  62.44  20.32             (--*-)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 18.13
```

Boxplot of GINNI Index_np

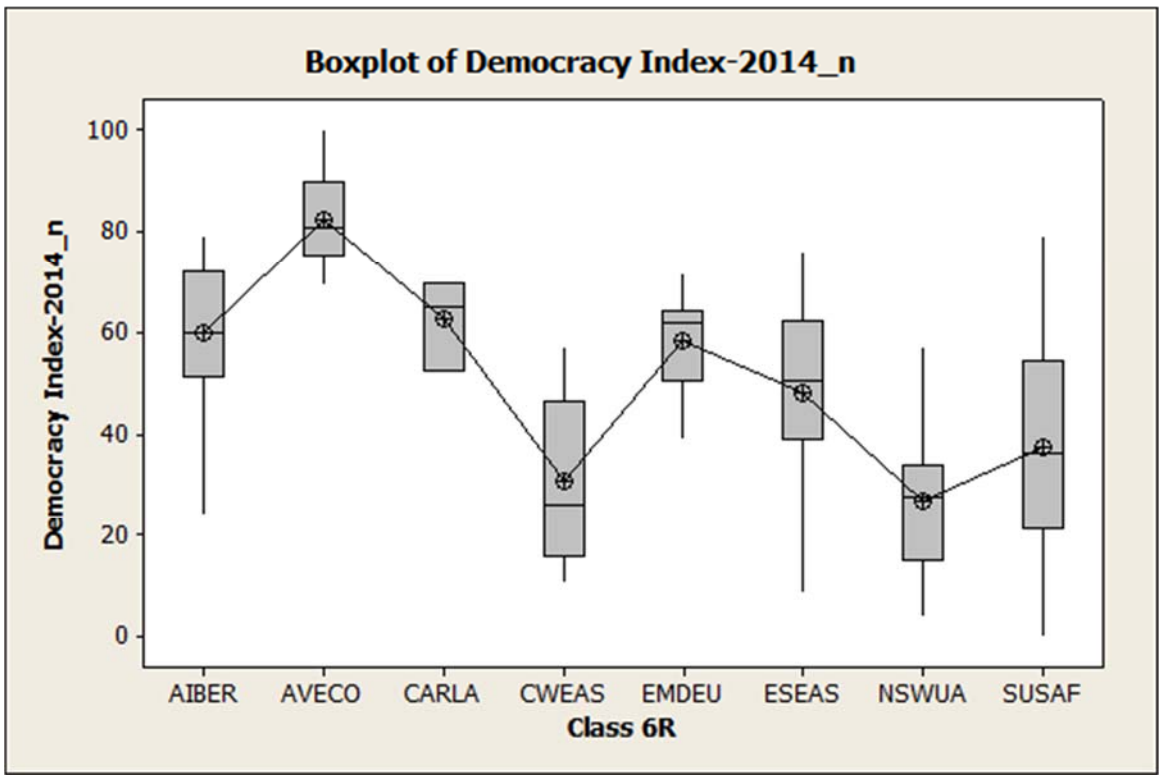


Obs: GINNI Index = 13 *missing values*

One-way ANOVA: Democracy Index-2014_n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  48182  6883  27.99  0.000
Error     124  30490   246
Total     131  78672

S = 15.68   R-Sq = 61.24%   R-Sq(adj) = 59.06%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -+---------+---------+---------+--------
AIBER  21  59.98  13.59                   (--*--)
AVECO  27  82.48   8.87                              (--*--)
CARLA   3  62.48   9.07               (--------*--------)
CWEAS  11  30.63  16.67    (---*----)
EMDEU  12  58.10  10.28                  (---*----)
ESEAS  12  47.88  19.58            (----*---)
NSWUA  14  26.55  14.34  (---*---)
SUSAF  32  37.53  21.06        (--*--)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15.68
```

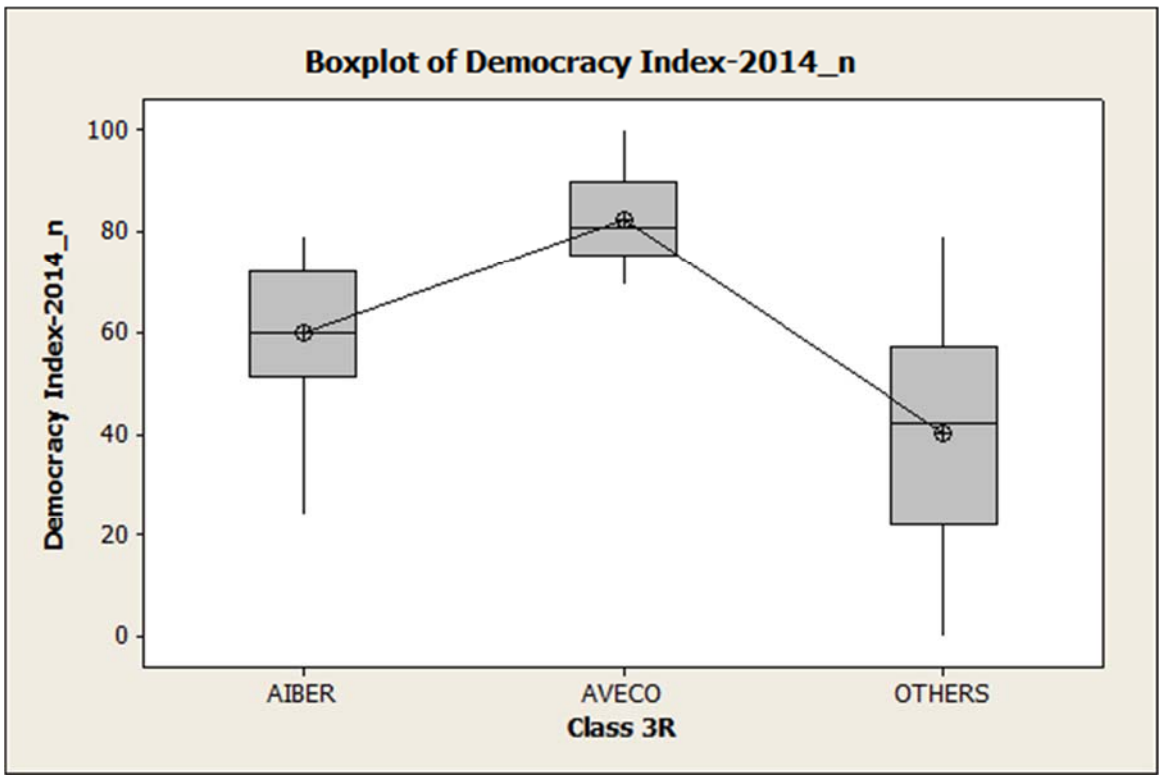
Boxplot of Democracy Index-2014_n



One-way ANOVA: Democracy Index-2014_n versus Class 3R

```
Source     DF     SS     MS      F      P
Class 3R    2  38300  19150  61.19  0.000
Error     129  40372    313
Total     131  78672

S = 17.69   R-Sq = 48.68%   R-Sq(adj) = 47.89%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ------+---------+---------+---------+---
AIBER   21  59.98  13.59            (----*----)
AVECO   27  82.48   8.87                       (----*---)
OTHERS  84  40.11  20.43  (--*-)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 17.69
```

Boxplot of Democracy Index-2014_n



5.4. Análise do Grau de discriminação e das Médias

Nesta etapa, todos os resultados da função One-Way ANOVA, foram ordenados considerando do maior F (grau de discriminação) para o menor. Em seguida, foram separados nas categorias: Indicador Sintético e Indicador Analítico. Selecionados 7 indicadores mais relevantes (maior F) de cada categoria e elaborado as tabelas e gráficos a seguir:

Quadro: Médias e F dos Indicadores Sintéticos

Regiões	**evwater**	**oppor**	**gover**	**spi**	**episc**	**fowell**	**democ**
AIBER	**20.39**	**55.28**	**42.48**	**64.41**	**52.70**	**63.48**	**59.98**
AVECO	**80.29**	**79.94**	**83.25**	**90.30**	**81.39**	**83.29**	**82.48**
OTHERS	**13.26**	**32.45**	**33.36**	**42.63**	**37.87**	**42.32**	**40.11**
F	**119.91**	**104.39**	**97.02**	**83.97**	**75.83**	**64.83**	**61.19**

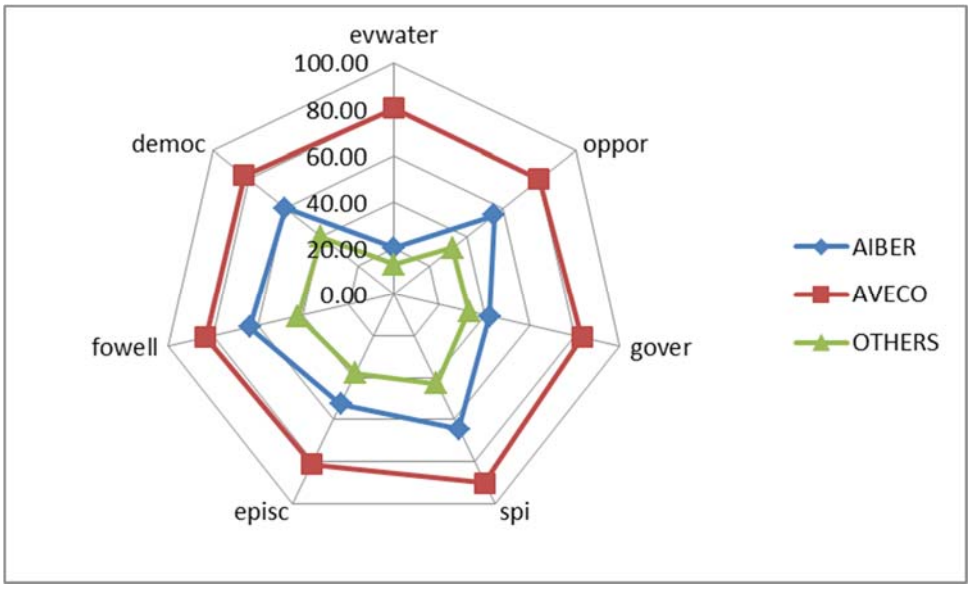


Gráfico com as Médias e F dos Indicadores Sintéticos

LEGENDA:
evwater: EV-Water Resources
oppor: Opportunity
gover: Governança
spi:Social Progress Index
episc: EPI Score
fowell: Foundations of Wellbeing
democ: Democracy Index-2014

Quadro: Médias e F dos Indicadores Analíticos

Regiões	**tolhomo**	**freedom**	**yr3rds**	**wtreres**	**safenet**	**wyrsco**	**contra**
AIBER	**57.14**	**73.74**	**25.74**	**26.57**	**81.30**	**63.61**	**86.69**
AVECO	**64.51**	**85.82**	**49.63**	**70.33**	**89.43**	**86.80**	**88.99**
OTHERS	**16.64**	**56.65**	**18.05**	**56.16**	**65.13**	**49.92**	**62.11**
F	**103.30**	**68.49**	**37.45**	**33.17**	**28.77**	**26.41**	**21.15**

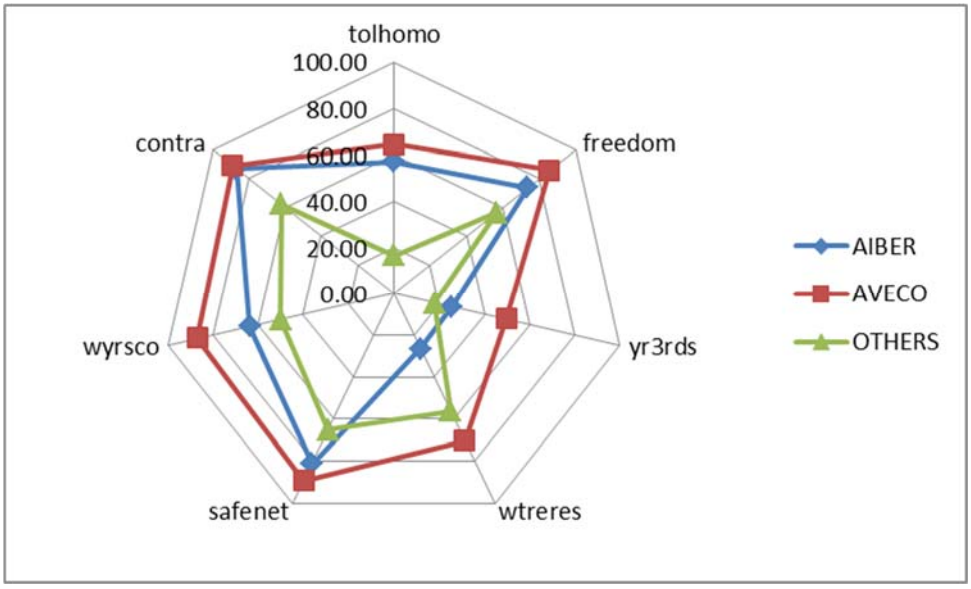
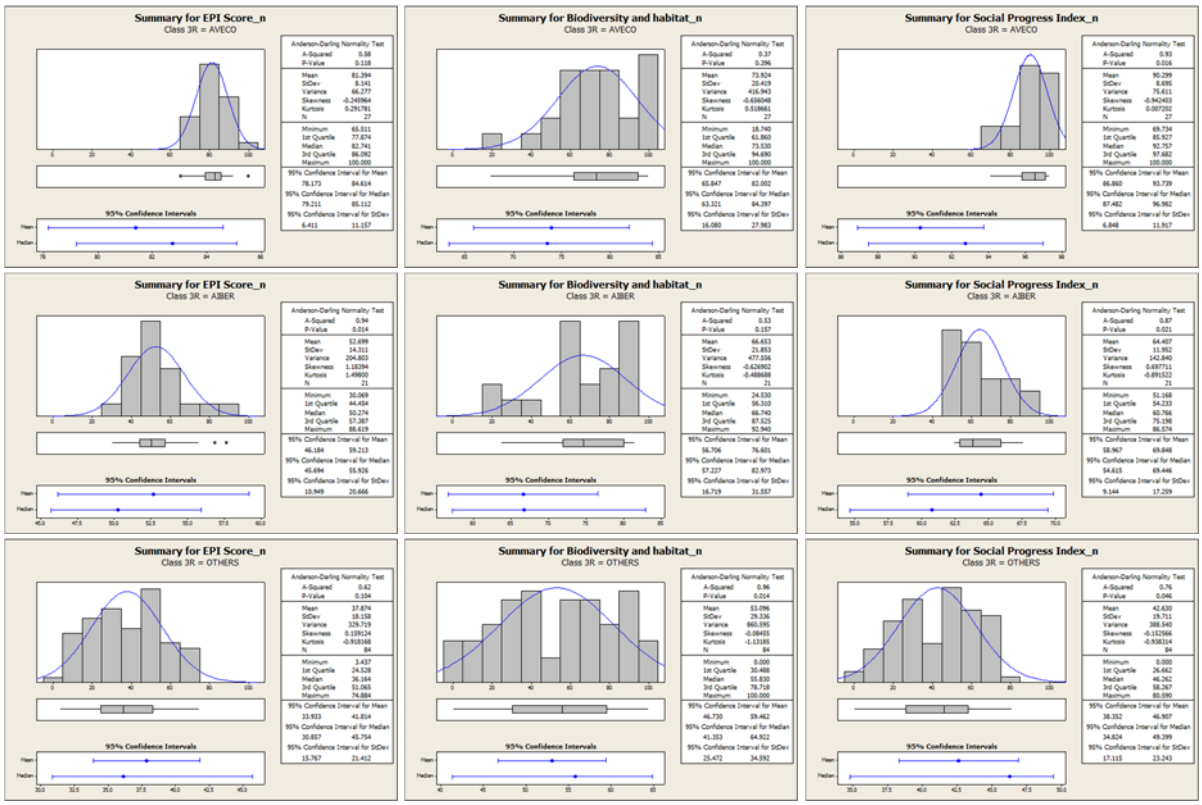


Gráfico com as Médias e F dos Indicadores Analíticos

LEGENDA:
tolhomo: Tolerance for homosexuals
freedom: Personal Freedom and Choice
yr3rds: Years of tertiary schooling
wtreres: Women treated with respect
safenet : Community safety net
wyrsco: Women`s years in school
contra: Satisfied contraception (% of women)

5.5. Análise das novas variáveis (normalizadas e positivadas)

Os gráficos abaixo são resultados de análise descritiva para 3 (três) variáveis importantes referentes ao tema Biodiversidade – Capital Natural e Social. A análise foi feita por agrupamento dos países nas 3 (três) regiões objeto deste estudo, AIBER, AVECO e OTHERS. Foram dispostos verticalmente afim de tentar evidenciar tendências.



5.6. CONSIDERAÇÕES

O objetivo desta sessão foi analisar comparativamente as variáveis deste estudo quando agrupadas em 3 grandes blocos, a saber, AIBER com os países da região Ibero-

Americana, AVECO, os países de economias avançadas e OTHERS, os demais países constantes na base de dados, com ênfase no primeiro bloco.

Cuba foi reclassificada e passou a figurar como um país do bloco AIBER.

Na primeira fase foi utilizado o modelo ANOVA (one-way), para efetuar as comparações. Esta análise de variância testa a hipótese de que as médias de duas ou mais populações são iguais.

Das variáveis que apresentaram relevância, destacamos a “Necessidades Humanas Básicas”, com F de 41.59 e o país Latvia (Letônia) figura como *outlier* no agrupamento de países da região AVECO, para esta variável. No indicador “Fundamentos de Bem-Estar”, Cuba figura como *outlier* no agrupamento de países da região AIBER e F de 64.83 para esta variável.

A Letônia (Latvia), aparece como outlier na avaliação da variável “Saúde e Bem-Estar”, denotando resultado bem abaixo dos de países do agrupamento AVECO. E no grupo OTHERS, os dois outliers são Rússia e Cazaquistão (Kazakhstan). Com F relativamente baixo, de 32.23. Para o indicador “Liberdade Pessoal e de Escolha”, o F é de 68.49 e no agrupamento de países AIBER, Cuba aparece como outlier em relação aos demais. Os outliers do grupo OTHERS são respectivamente Paquistão, Sudão e Mauritânia.

Para a variável que mede a “Emissão de Gases de Efeito Estufa”, a República Centro-Africana (Central African Republic), aparece neste estudo como o maior emissor de gases de efeito estufa. Segundo o relatório da ONU, sobre os dados de 2010, a principal causa deste resultado estava ligado a Desmatamento e Exploração de madeira para comércio (89.46%), seguido Agricultura (5.26%), entre outros. O F de 2.60, um dos mais baixos encontrados neste trabalho. Para o indicador de “Escravidão nos tempos atuais”, vemos no agrupamento AVECO, 3 *outliers* que destoam do grupo, são eles: Eslovênia, Eslováquia e República Tcheca. A Mauritânia lidera o ranking nesta categoria, com cerca de 155,600 indivíduos (4% da pop.), em situação de escravidão.
fonte:http://www.globalslaveryindex.org/country/mauritania/

Com F de 21.15, o indicador que mede o “% de Mulheres Satisfeitas com os Métodos Contraceptivos” oferecidos em seus países, identificamos no agrupamento AIBER, a Guatemala como *outlier*, e no AVECO figuram Japão e Islândia respectivamente.

Para a variável que mede a percepção das “Mulheres que são tratadas com respeito” em seus países, listamos em AIBER, os 3 países com os menores valores, são eles: Perú com 24, República Dominicana com 18 e Colômbia com 17. O Brasil com 28, situa-se próximo destes. O F é de 33.17. Na análise de países para o índice de “Discriminação e Violência”,

apresentando F de 27.61, Israel aparece como outlier destoando significativamente de seu grupo (AVECO).

No indicador de “Segurança em Comunidade”, com F de 28.77, os 2 outliers que aparecem no agrupamento AVECO, são: Grécia e Coreia do Sul (Republic of Korea). A variável que representa a quantidade de “Anos das Mulheres na Escola”, a Islândia figura como outlier no agrupamento AVECO, o F desta análise é de 26.41

O “Índice de Desenvolvimento Humano-IDH”, com F de 34.45, destacou negativamente a Coreia do Sul (Republic of Korea), Letônia (Latvia), Holanda e Suíça como *outliers* na análise. No entanto, as considerações sobre estes resultados (*outliers*), não deverão ser levadas em conta, neste momento, haja vista que foi necessário ajuste nos valores de 3 países. Para os próximos trabalhos, estes casos passarão a refletir a as correções efetuadas.

O indicador EPI – Índice de Performance Ambiental, apresentou F de 75,83 e se destacaram positiviamente como *outliers* os países Portugal e Espanha, os únicos pertencentes ao continente europeu dos países Ibero-Americanos e Suíça no AVECO. E negativamente como outlier a Coreia do Sul (Republic of Korea).

A variável *EV-Water Resources*, que reflete a preocupação quanto ao tratamento de águas residuais antes que sejam devolvidas ao meio-ambiente, apresentou um F de 119.91, o mais alto deste estudo em relação às outras variáveis. Como o*utliers* que se destacaram positivamente temos os países Chile, Portugal e Espanha no agrupamento AIBER. Por sua vez, o indicador *EV-Agriculture*, que retrata oa países que recebem subsídios para sua produção, despontou como *outliers* no agrupamento AVECO em ordem crescente são: Israel, Japão, Coreia do Sul (Republic of Korea) e Nova Zelândia.

O Índice de Felicidade do Planeta – HPI, resultou em um F de 22.69 e foi o único indicador em que AIBER performou melhor que AVECO.

Para algumas variáveis não foi possível uma análise mas robusto devido a falta de dados, para citar algumas, temos *EV-Fisheries_np* (37 *missing values*), *EV-Forests* (22 *missing values*), *Years of tertiary schooling* (15 *missing values*) e *GINI Index* (13 *missing values*), entre outras.

Na sessão de Análise do grau de discriminação e das médias, onde foi utilizado os resultados da função *One-Way ANOVA*, foi possível visualizar graficamente de maneira sintetizada o que os boxplots já vinham indicando individualmente para cada variável.

Em linhas gerais, os resultados deste trabalho de comparações com agrupamentos por regiões, veio reforçar o desempenho superior dos países AVECO em relação ao AIBER e demais países obejtos deste estudo.

6. AMOSTRAGEM

Os objetivos desta sessão são, realizar uma pesquisa por amostragem dos dados dimensionadores para as variáveis SPI-2014 – Índice de Progresso Social, IDH-2014 – Índice de Desenvolvimento Humano e EPI-2014 – Índice de Performance Ambiental, efetuar análises ANOVA para verificar variância entre AIBER e AVECO e testes de amostragens (two-sample) e gráficos boxplots.

O dados serão categorizados em 3 (três) principais regiões, os países Ibero-Americanos (AIBER), países de economias avançadas (AVECO) e os demais países constantes na base de dados deste estudo.

6.1 As variáveis

Este estudo trata de 4 (quatro) variáveis, sendo 3 (três) quantitativas – SPI-2014, IDH-2014 e SPI-2014, e 1 (um) categórica – Class 3R (região) descritas abaixo no Quadro 2:

Quadro 2. As 4 variáveis do eixo temático para este estudo

Variável	Significado	Tipo	Unidade de Medida
SPI-2014	O Índice de Progresso Social é composto por três dimensões do progresso social: Necessidades Humanas Básicas, Fundamentos de Bem-Estar e Oportunidades. fonte: www.socialprogressimperative.org	Variável quantitativa	Indicador nominal
IDH-2014	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável quantitativa	Indicador nominal
EPI-2014	Índice de Performance Ambiental, composto por: Saúde Ambiental (saúde humana e dos ecossistemas), Impactos na Saúde (medido pela mortalidade infantil – idade entre 1 e 5 anos), Qualidade do Ar (poluição do ar, partículas sólidas e seu impacto na saúde) e Nível de Água tratada e Sanitização adequada.	Variável quantitativa	Indicador nominal
Class 3R	Variável utilizada para agrupar os 132 países em 3 (três) regiões, AIBER (Ibero-Americana, AVECO (Economias Avançadas) e OTHERS (demais países)	Variável categórica	N/A

Fonte: O autor a partir dos dados da planilha estatística, do GPS p.16 e do site do Banco Mundial

6.2. PESQUISA POR AMOSTRAGEM

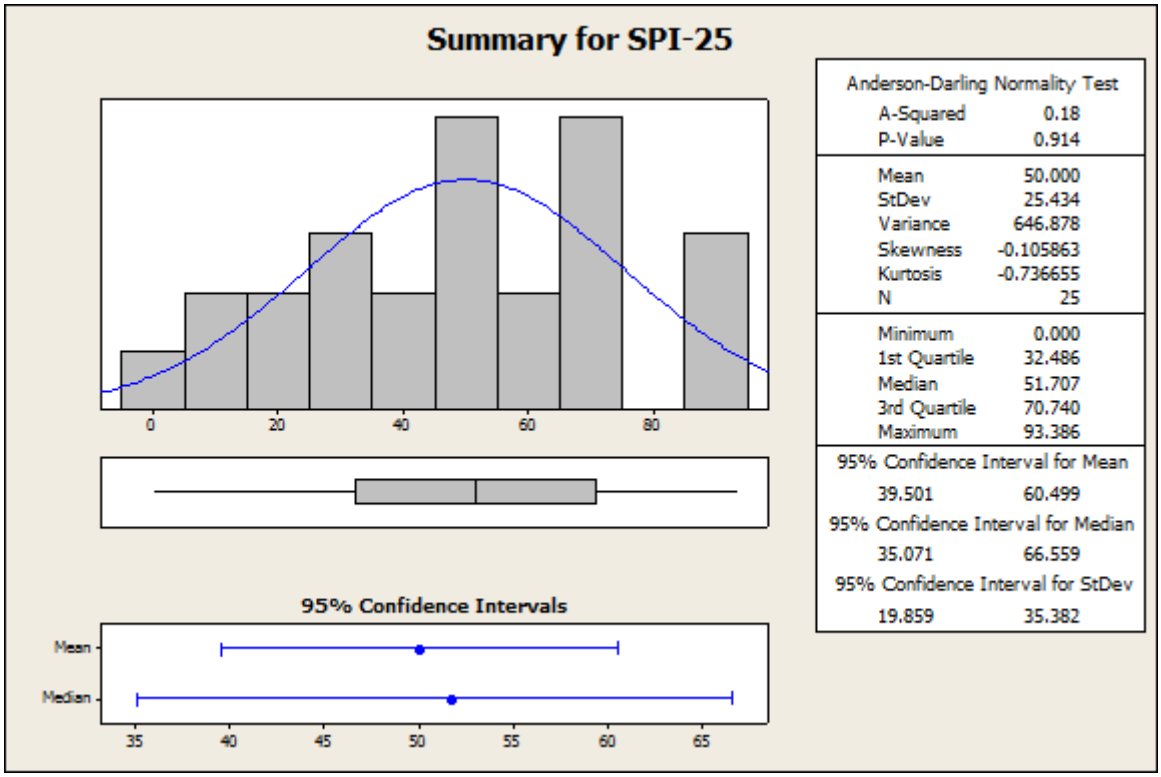
A pesquisa por amostragem foi feita em três amostras extraídas aleatoriamente da base de dados. Uma com 25 países, outra com 50 e por fim, uma amostra com 100 países. Estas amostras foram geradas a partir da funcionalidade "*Random Data*" do software Minitab. Para cada amostra extraída foram efetuadas análises exploratórias dos dados, análise de variância (ANOVA), testes de amostragens (two-sample) e gráficos boxplots.

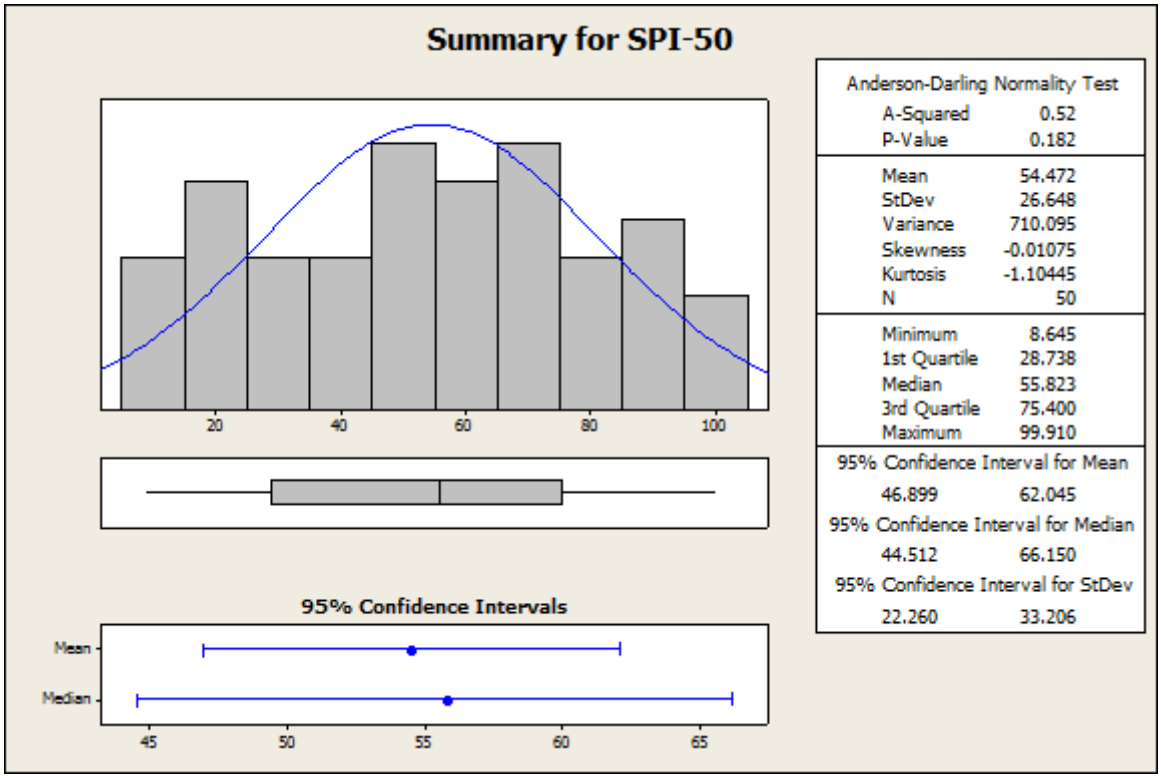
6.2.1 SPI-2014 – Índice de Progresso Social

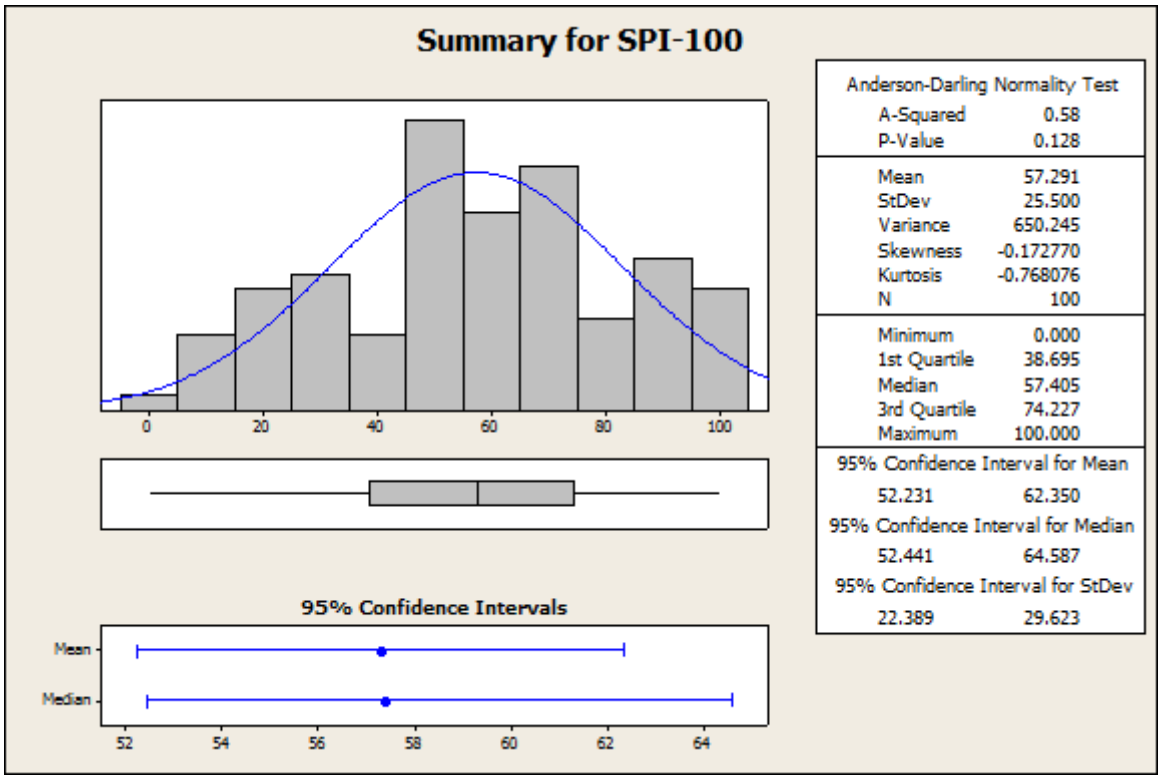
Nas sessões que seguem, serão demonstrados os resultados da análise exploratória, análise de variância e teste de amostragem entre duas regiões para a variável SPI-2014.

6.2.1.1 Análise Exploratória

Abaixo os gráficos boxplots de análise exploratória com as amostragens de 25, 50 e 100 países, escolhidas aleatoriamente entre as 3 (três) regiões.





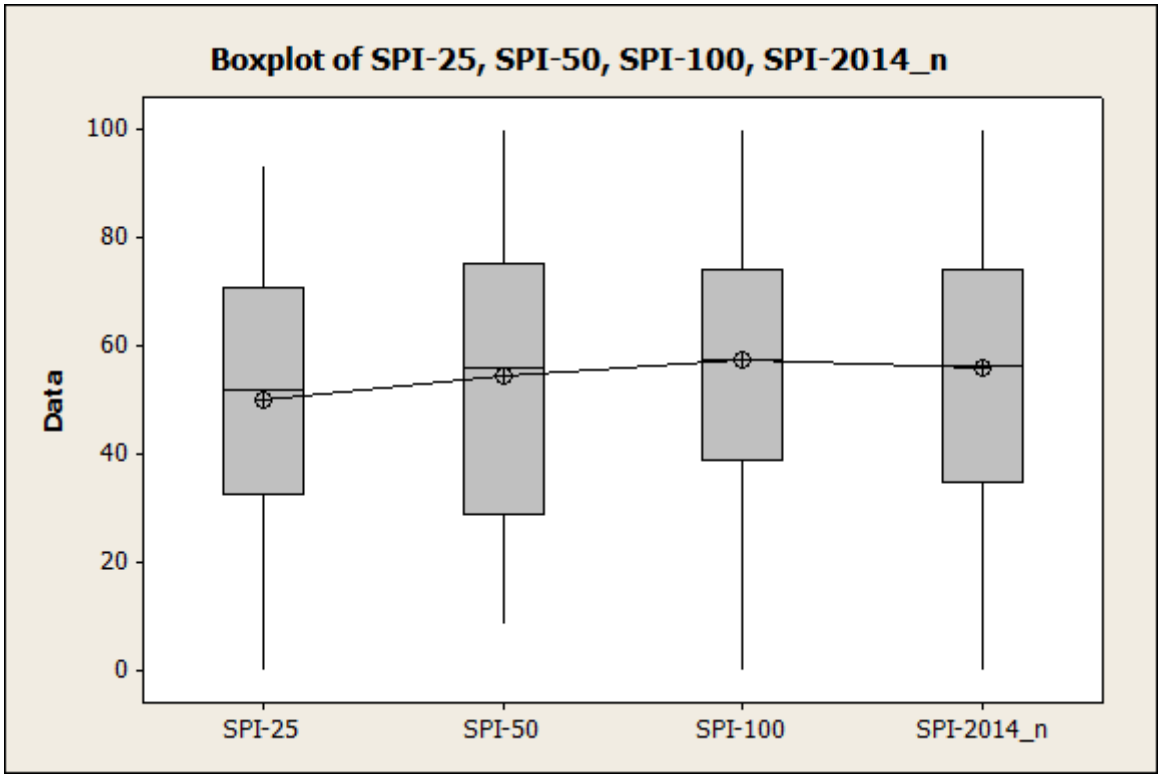


6.2.1.2 Análise de variância (ANOVA unstacked)

Os resultados para a ANOVA, considerando as amostras com 25, 50, 100 e 132 países, são demonstrados abaixo.

One-way ANOVA: SPI-25, SPI-50, SPI-100, SPI-2014_n

```
Source   DF      SS   MS     F      P
Factor    3    1141  380  0.58  0.631
Error   303  200027  660
Total   306  201168

S = 25.69   R-Sq = 0.57%   R-Sq(adj) = 0.00%

                              Individual 95% CIs For Mean Based on
                              Pooled StDev
Level         N   Mean  StDev  ----+---------+---------+---------+-----
SPI-25       25  50.00  25.43  (----------------*----------------)
SPI-50       50  54.47  26.65              (-----------*-----------)
SPI-100     100  57.29  25.50                      (-------*--------)
SPI-2014_n  132  55.84  25.52                     (------*------)
                              ----+---------+---------+---------+-----
                               42.0      48.0      54.0      60.0
Pooled StDev = 25.69
```



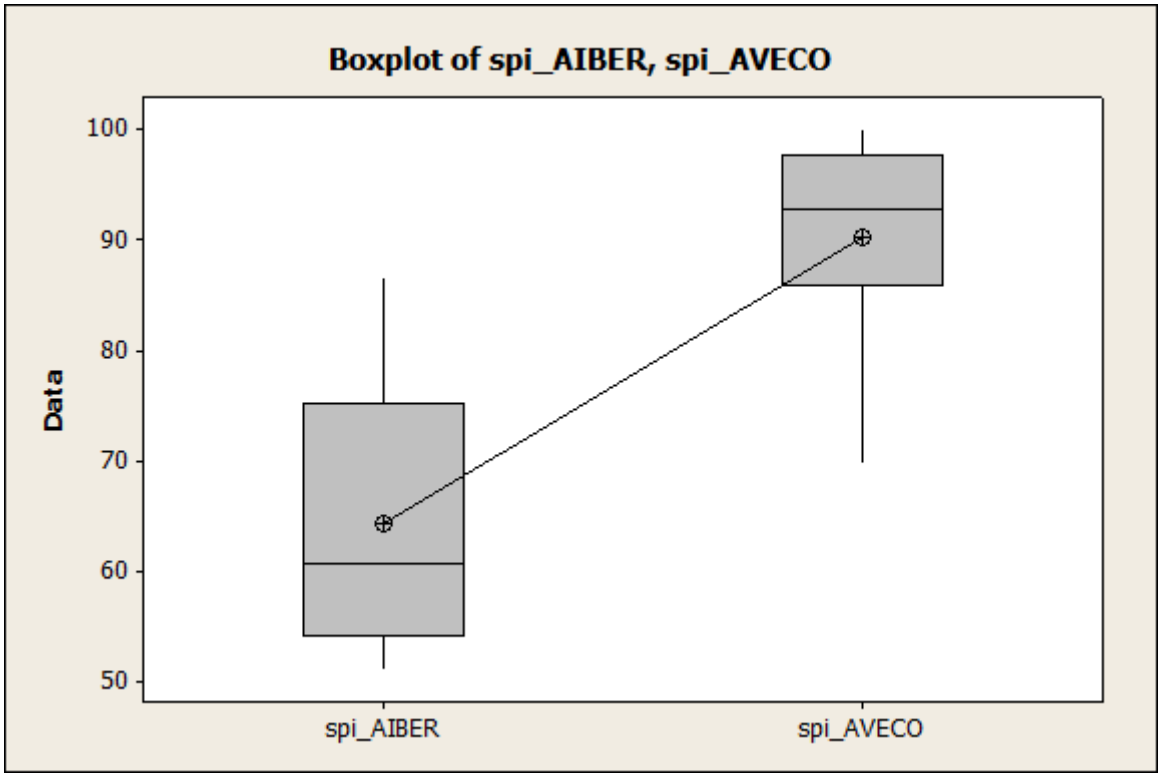
Obs: No gráfico acima, a amostra com 100 países apresentou o menor intervalo, denotando representar melhor a totalidade da população, com 95% de confiança.

Comparação entre as regiões AIBER e AVECO (ANOVA unstacked)

One-way ANOVA: spi_AIBER, spi_AVECO

```
Source  DF     SS    MS      F      P
Factor   1   7919  7919  75.53  0.000
Error   46   4823   105
Total   47  12742

S = 10.24   R-Sq = 62.15%   R-Sq(adj) = 61.33%

                              Individual 95% CIs For Mean Based on
                              Pooled StDev
```

```
Level       N   Mean  StDev   +---------+---------+---------+---------
spi_AIBER  21  64.41  11.95   (---*----)
spi_AVECO  27  90.30   8.70                          (---*---)
                              +---------+---------+---------+---------
                             60        70        80        90
Pooled StDev = 10.24
```

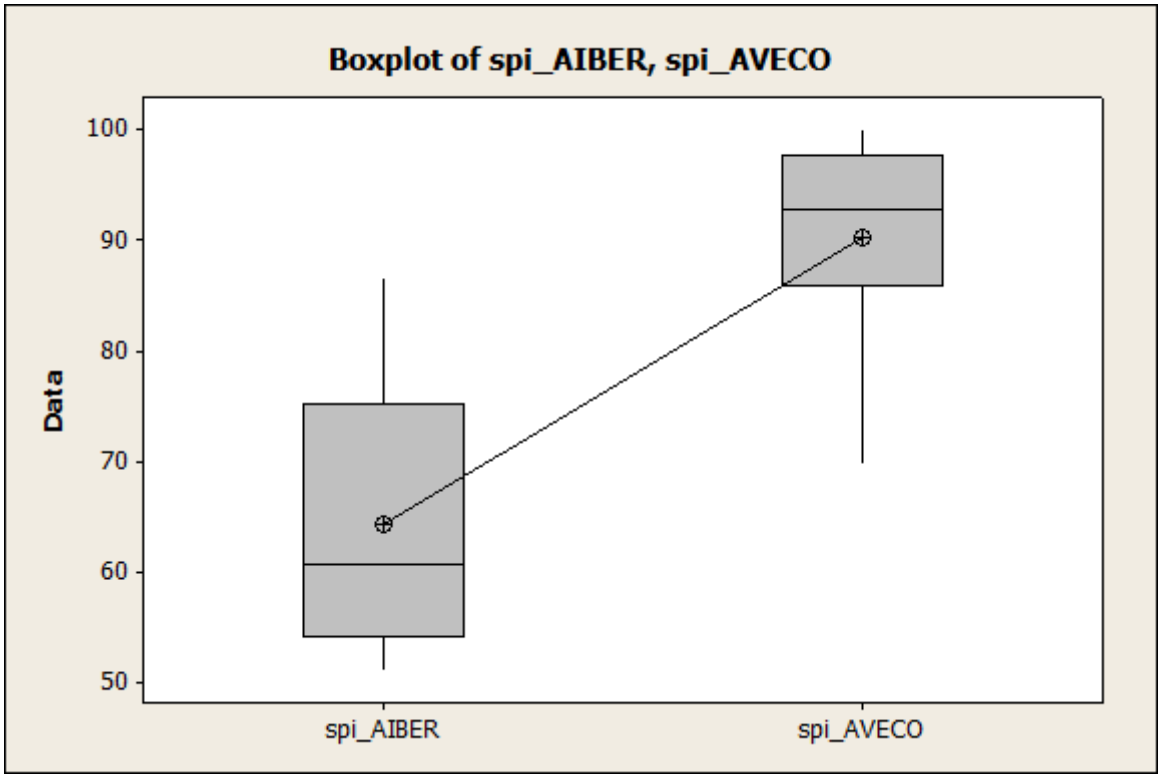


6.2.1.3 Testes de amostragens (two-sample)

Teste de amostragens (T-Test).

Two-Sample T-Test and CI: spi_AIBER, spi_AVECO

```
Two-sample T for spi_AIBER vs spi_AVECO

            N   Mean  StDev  SE Mean
spi_AIBER  21   64.4   12.0      2.6
spi_AVECO  27  90.30   8.70      1.7

Difference = mu (spi_AIBER) - mu (spi_AVECO)
Estimate for difference:  -25.89
95% upper bound for difference:  -20.66
T-Test of difference = 0 (vs <): T-Value = -8.36  P-Value = 0.000  DF = 35
```

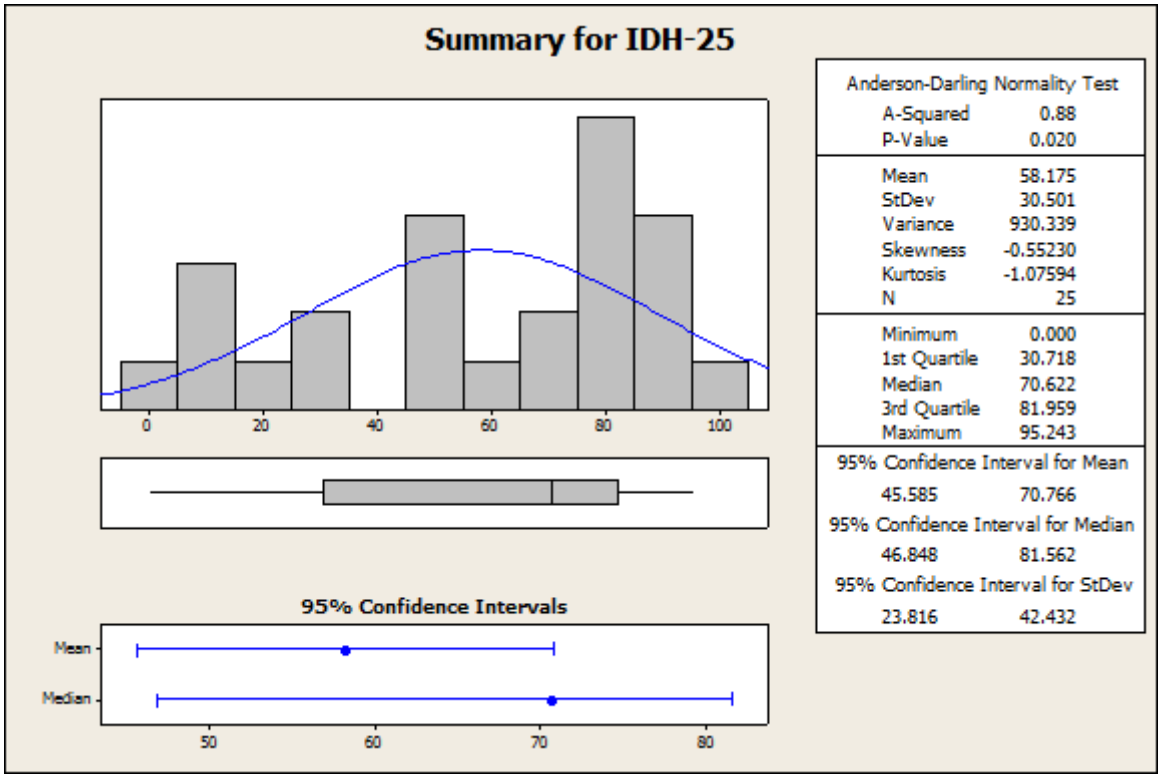


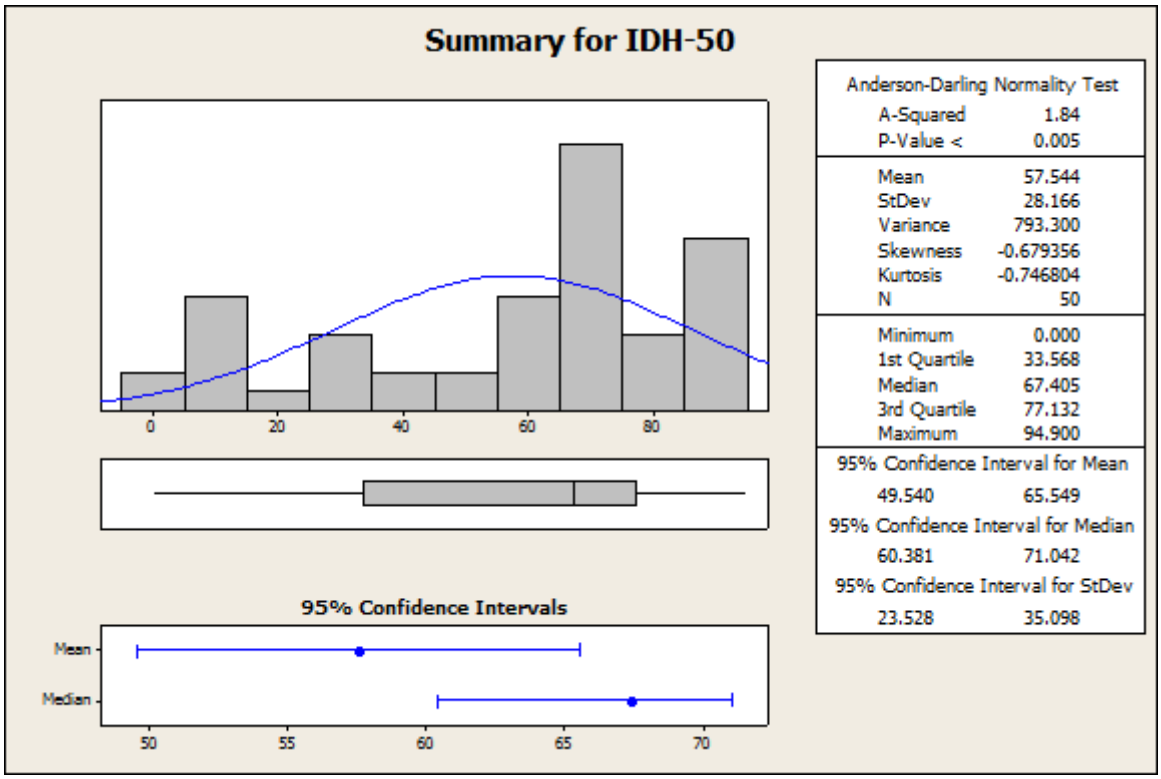
6.3 IDH-2014 – Índice de Desenvolvimento Humano

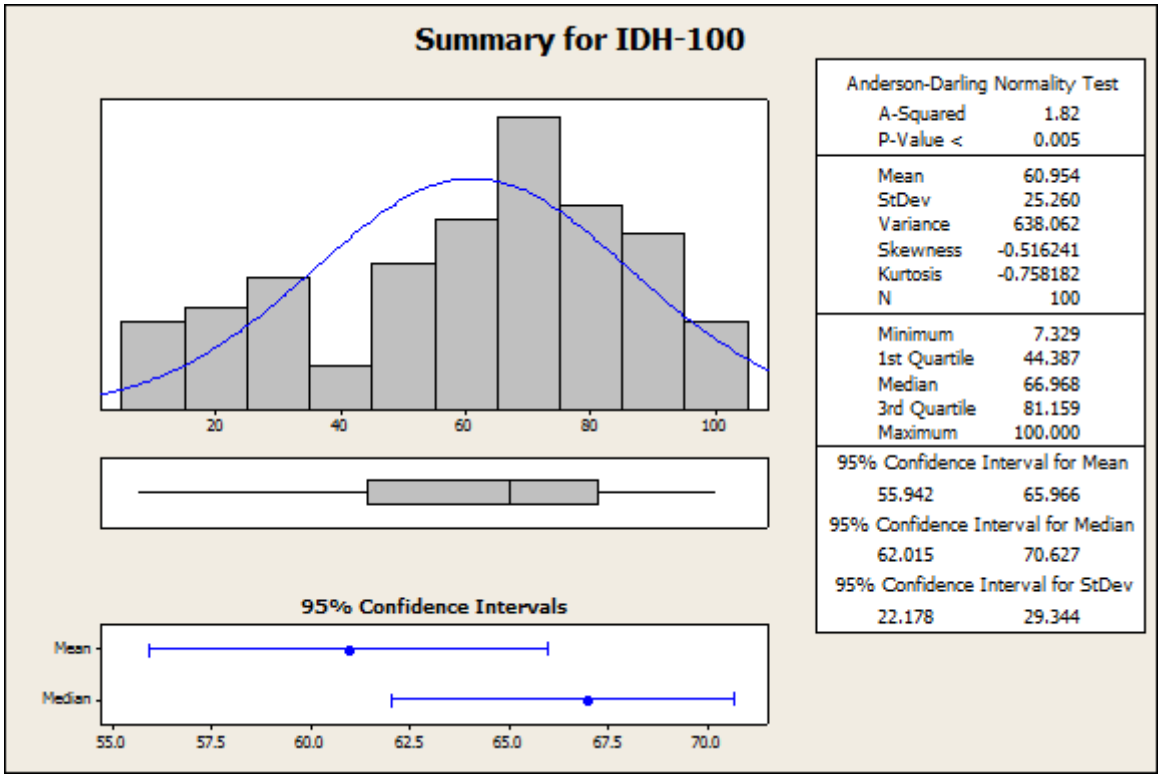
Nas sessões que seguem, serão demonstrados os resultados da análise exploratória, análise de variância e teste de amostragem entre duas regiões para a variável IDH-2014.

6.3.1.1 Análise Exploratória

Abaixo os gráficos boxplots de análise exploratória com as amostragens de 25, 50 e 100 países, escolhidas aleatoriamente entre as 3 (três) regiões.


Summary for IDH-25
0
20
40
60
80
100
95% Confidence Intervals
Mean
Median
50
60
70
80
Anderson-Darling Normality Test
A-Squared 0.88
P-Value 0.020
Mean 58.175
StDev 30.501
Variance 930.339
Skewness -0.55230
Kurtosis -1.07594
N 25
Minimum 0.000
1st Quartile 30.718
Median 70.622
3rd Quartile 81.959
Maximum 95.243
95% Confidence Interval for Mean
45.585 70.766
95% Confidence Interval for Median
46.848 81.562
95% Confidence Interval for StDev
23.816 42.432


Summary for IDH-50
0
20
40
60
80
95% Confidence Intervals
Mean
Median
50
55
60
65
70
Anderson-Darling Normality Test
A-Squared 1.84
P-Value < 0.005
Mean 57.544
StDev 28.166
Variance 793.300
Skewness -0.679356
Kurtosis -0.746804
N 50
Minimum 0.000
1st Quartile 33.568
Median 67.405
3rd Quartile 77.132
Maximum 94.900
95% Confidence Interval for Mean
49.540 65.549
95% Confidence Interval for Median
60.381 71.042
95% Confidence Interval for StDev
23.528 35.098

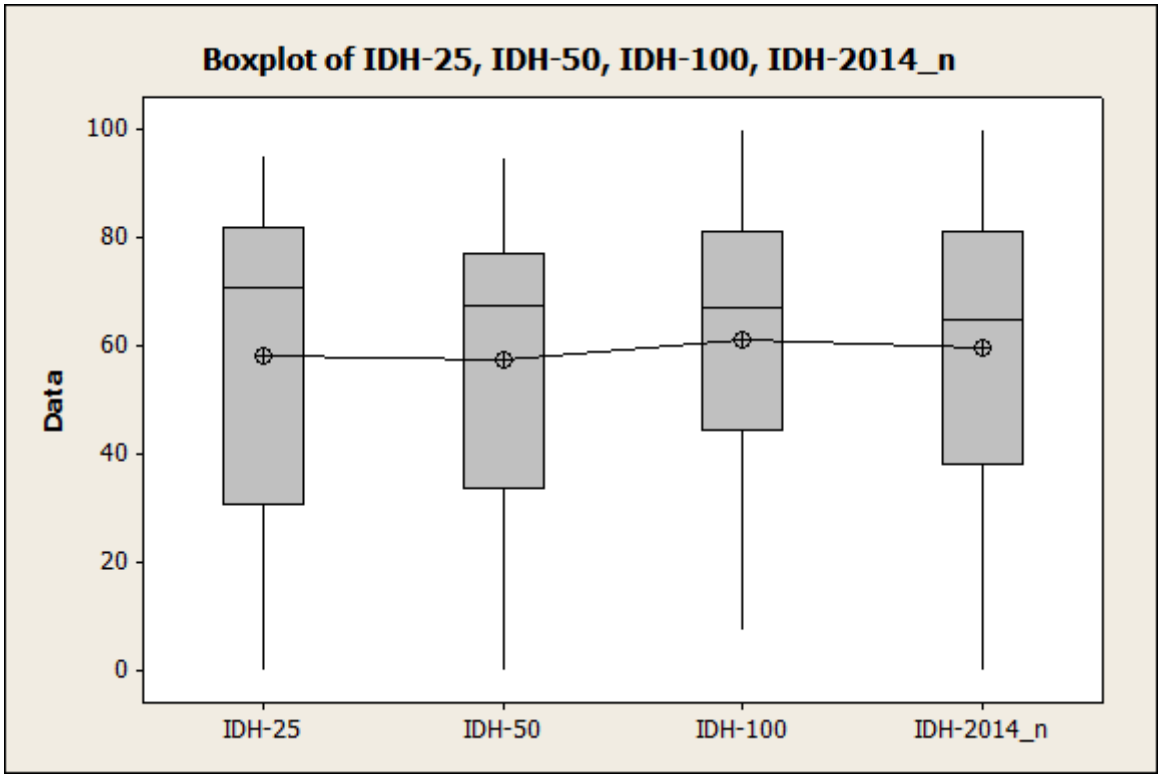


6.3.1.2 Análise de variância (ANOVA unstacked)

Os resultados para a ANOVA, considerando as amostras com 25, 50, 100 e 132 países, são demonstrados abaixo.

One-way ANOVA: IDH-25, IDH-50, IDH-100, IDH-2014_n

```
Source   DF      SS   MS     F      P
Factor    3     446  149  0.21  0.891
Error   303  215937  713
Total   306  216382

S = 26.70   R-Sq = 0.21%   R-Sq(adj) = 0.00%

Level         N   Mean  StDev
IDH-25       25  58.18  30.50
IDH-50       50  57.54  28.17
IDH-100     100  60.95  25.26
IDH-2014_n  132  59.64  26.44

            Individual 95% CIs For Mean Based on Pooled StDev
Level         -+---------+---------+---------+--------
IDH-25        (-----------------*----------------)
IDH-50             (-----------*-----------)
IDH-100                   (--------*-------)
IDH-2014_n                (------*-------)
              -+---------+---------+---------+--------
            48.0      54.0      60.0      66.0

Pooled StDev = 26.70
```

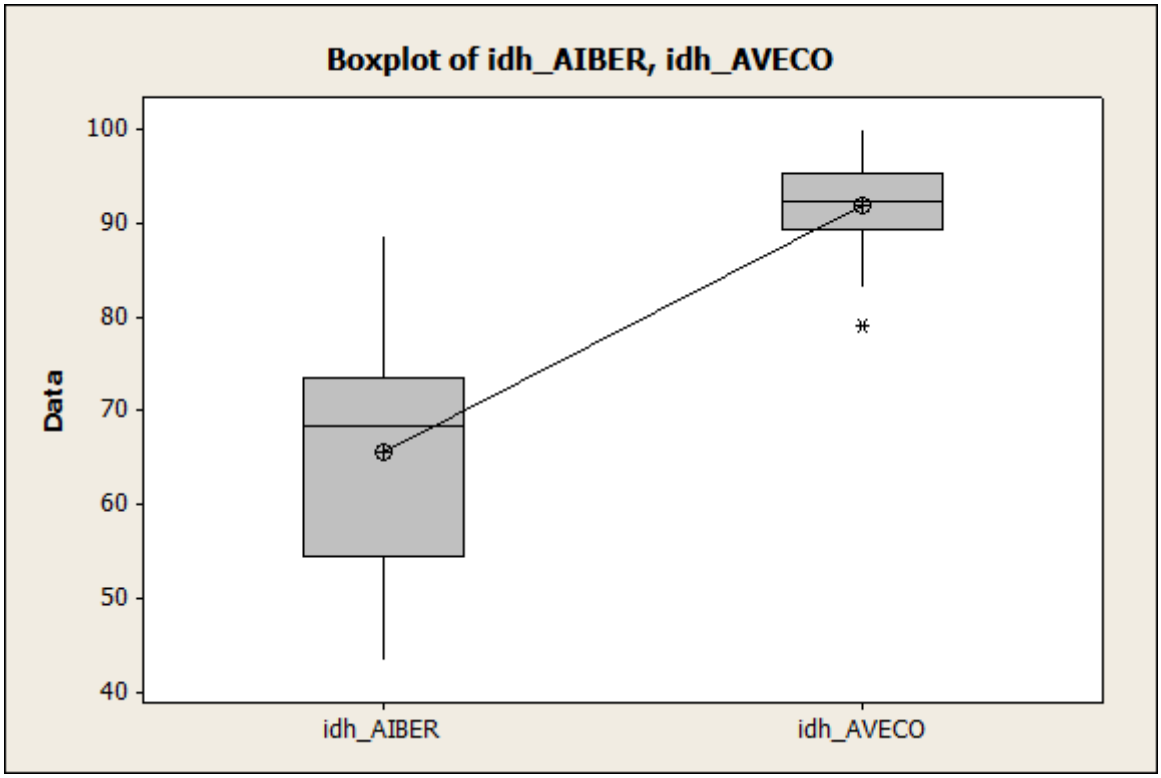


Obs: No gráfico acima, a amostra com 100 países apresentou o menor intervalo, denotando representar melhor a totalidade da população, com 95% de confiança.

Comparação entre as regiões AIBER e AVECO (ANOVA unstacked)

One-way ANOVA: idh_AIBER, idh_AVECO

```
Source  DF       SS      MS       F      P
Factor   1   8122.1  8122.1  101.08  0.000
Error   46   3696.1    80.4
Total   47  11818.2

S = 8.964   R-Sq = 68.73%   R-Sq(adj) = 68.05%

                              Individual 95% CIs For Mean Based on
                              Pooled StDev
Level       N    Mean   StDev  --------+---------+---------+---------+-
idh_AIBER  21  65.671  12.452  (---*---)
idh_AVECO  27  91.892   4.784                          (---*--)
                              --------+---------+---------+---------+-
                                     70        80        90       100
Pooled StDev = 8.964
```



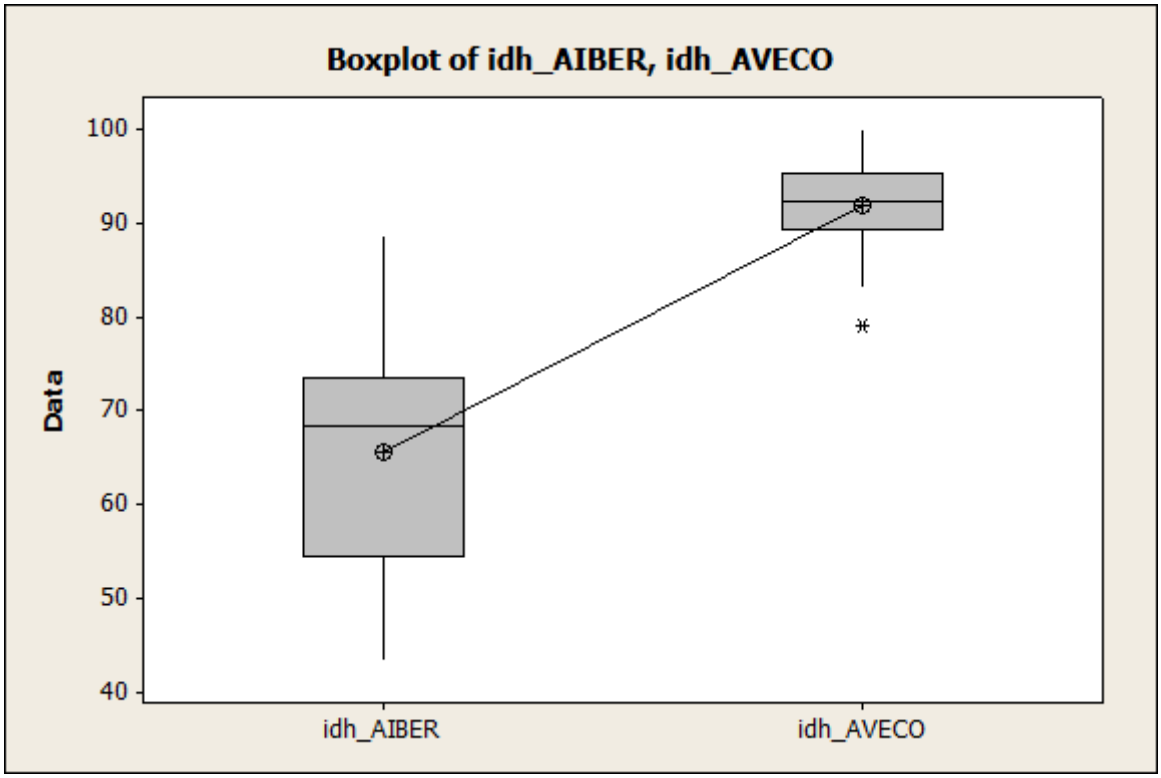
Obs: A Letônia (Latvia) figura como *outlier* no boxplot acima.

6.3.1.3 Testes de amostragens (two-sample)

Teste de amostragens (T-Test).

Two-Sample T-Test and CI: idh_AIBER, idh_AVECO

```
Two-sample T for idh_AIBER vs idh_AVECO

            N   Mean  StDev  SE Mean
idh_AIBER  21   65.7   12.5      2.7
idh_AVECO  27  91.89   4.78     0.92

Difference = mu (idh_AIBER) - mu (idh_AVECO)
Estimate for difference:  -26.22
95% upper bound for difference:  -21.31
T-Test of difference = 0 (vs <): T-Value = -9.14  P-Value = 0.000  DF = 24
```



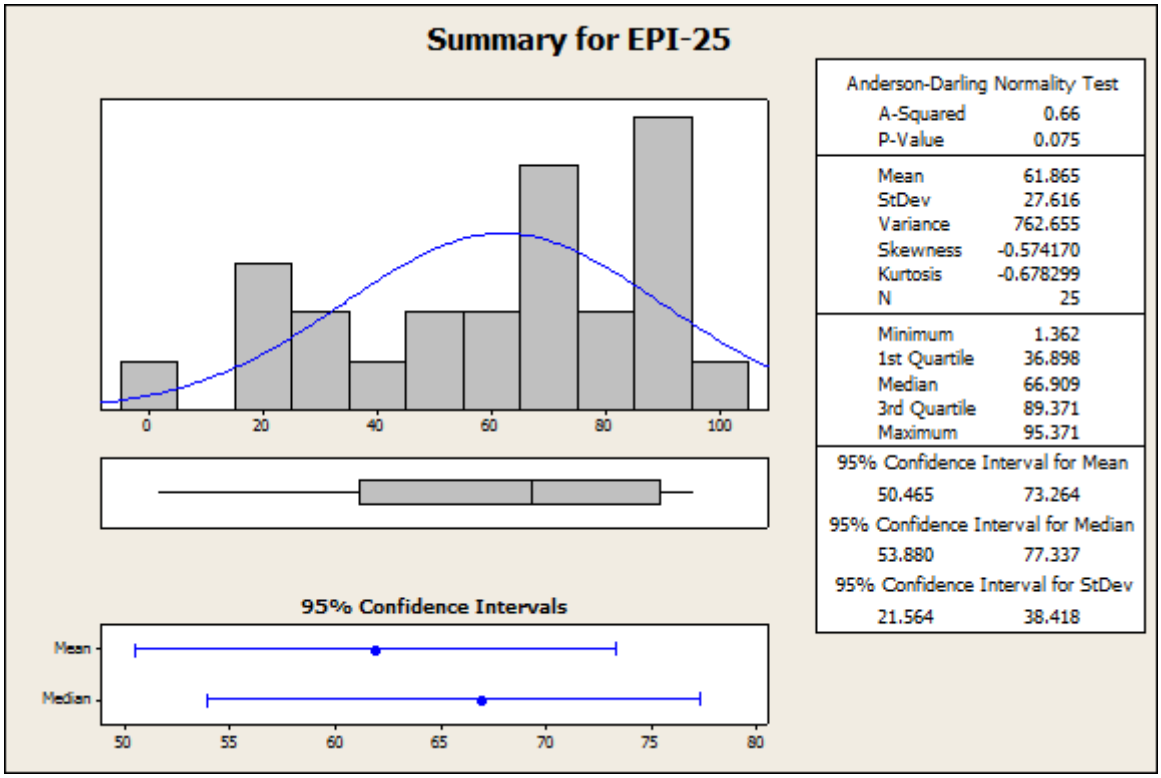
Obs: Visualmente, o teste de 2-Sample apresentou resultado muito próximo da ANOVA. Aqui a Letônia (Latvia) também figura como *outlier*.

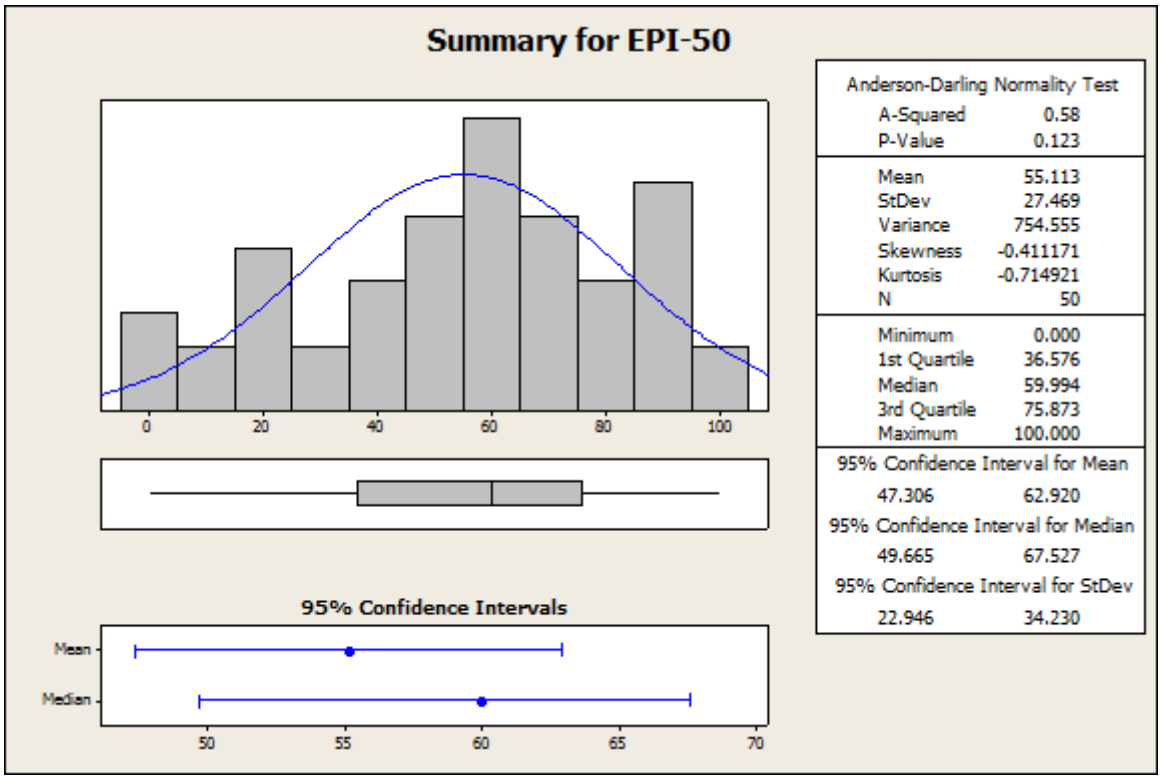
6.4 EPI-2014 – Índice de Performance Ambiental

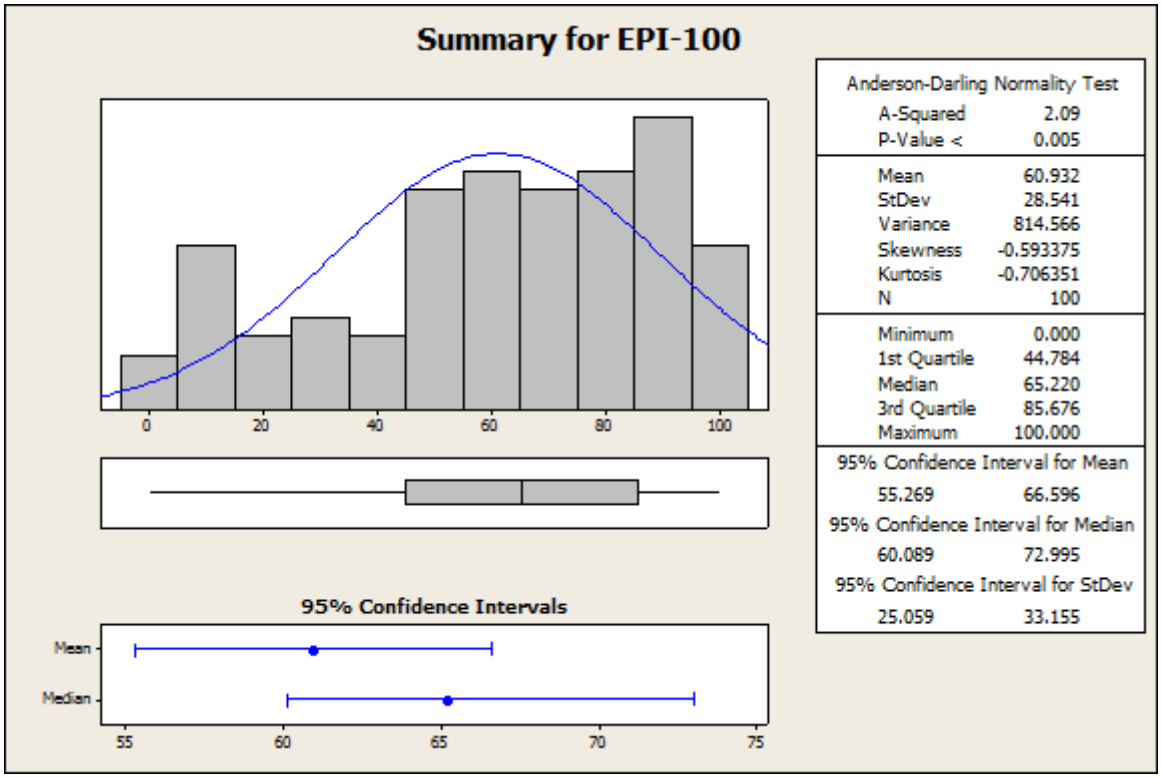
Nas sessões que seguem, serão demonstrados os resultados da análise exploratória, análise de variância e teste de amostragem entre duas regiões para a variável EPI-2014.

6.4.1 Análise Exploratória

Abaixo os gráficos boxplots de análise exploratória com as amostragens de 25, 50 e 100 países, escolhidas aleatoriamente entre as 3 (três) regiões.


Summary for EPI-25
0
20
40
60
80
100
95% Confidence Intervals
Mean
Median
50
55
60
65
70
75
80
Anderson-Darling Normality Test
A-Squared 0.66
P-Value 0.075
Mean 61.865
StDev 27.616
Variance 762.655
Skewness -0.574170
Kurtosis -0.678299
N 25
Minimum 1.362
1st Quartile 36.898
Median 66.909
3rd Quartile 89.371
Maximum 95.371
95% Confidence Interval for Mean
50.465 73.264
95% Confidence Interval for Median
53.880 77.337
95% Confidence Interval for StDev
21.564 38.418


Summary for EPI-50
0
20
40
60
80
100
95% Confidence Intervals
Mean
Median
50
55
60
65
70
Anderson-Darling Normality Test
A-Squared 0.58
P-Value 0.123
Mean 55.113
StDev 27.469
Variance 754.555
Skewness -0.411171
Kurtosis -0.714921
N 50
Minimum 0.000
1st Quartile 36.576
Median 59.994
3rd Quartile 75.873
Maximum 100.000
95% Confidence Interval for Mean
47.306 62.920
95% Confidence Interval for Median
49.665 67.527
95% Confidence Interval for StDev
22.946 34.230

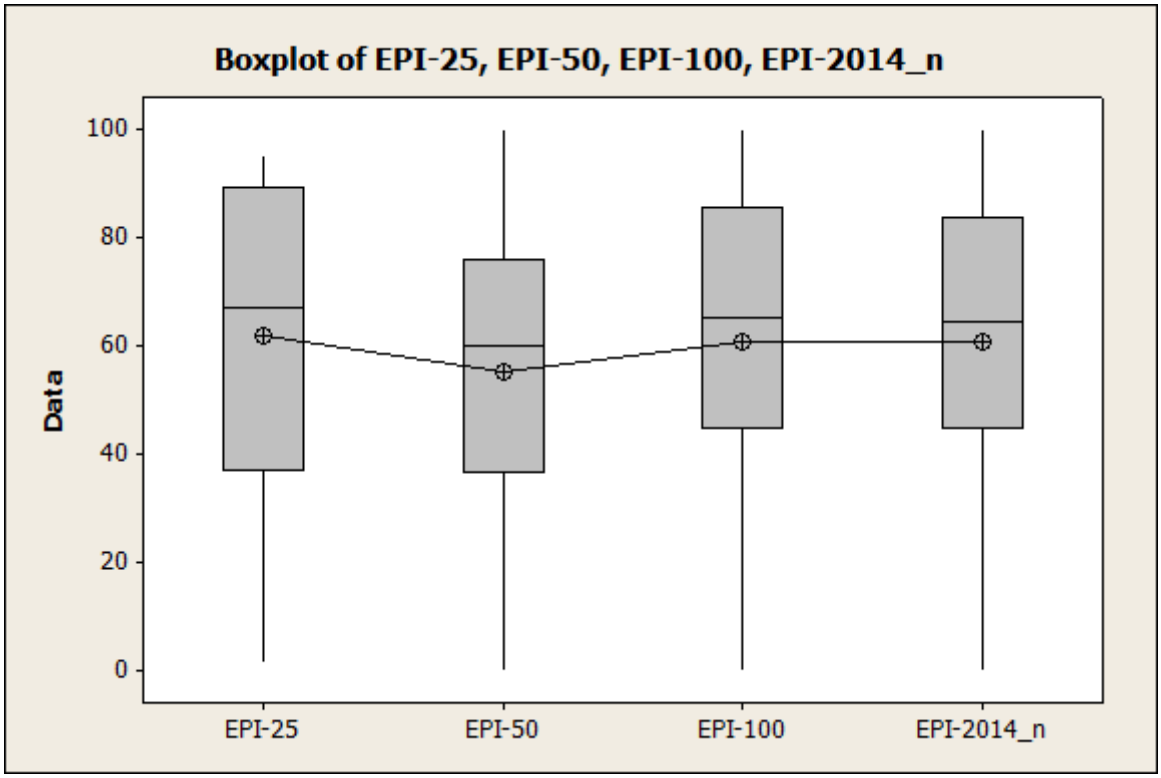


6.4.2 Análise de variância (ANOVA unstacked)

Os resultados para a ANOVA, considerando as amostras com 25, 50, 100 e 132 países, são demonstrados abaixo.

One-way ANOVA: EPI-25, EPI-50, EPI-100, EPI-2014_n

```
Source   DF      SS   MS     F      P
Factor    3    1454  485  0.63  0.598
Error   303  234390  774
Total   306  235844

S = 27.81   R-Sq = 0.62%   R-Sq(adj) = 0.00%

                                 Individual 95% CIs For Mean Based on
                                 Pooled StDev
Level         N   Mean  StDev   --+---------+---------+---------+-------
EPI-25       25  61.86  27.62        (--------------*--------------)
EPI-50       50  55.11  27.47   (----------*----------)
EPI-100     100  60.93  28.54             (-------*-------)
EPI-2014_n  132  60.81  27.42              (------*------)
                                 --+---------+---------+---------+-------
                                 49.0      56.0      63.0      70.0
Pooled StDev = 27.81
```

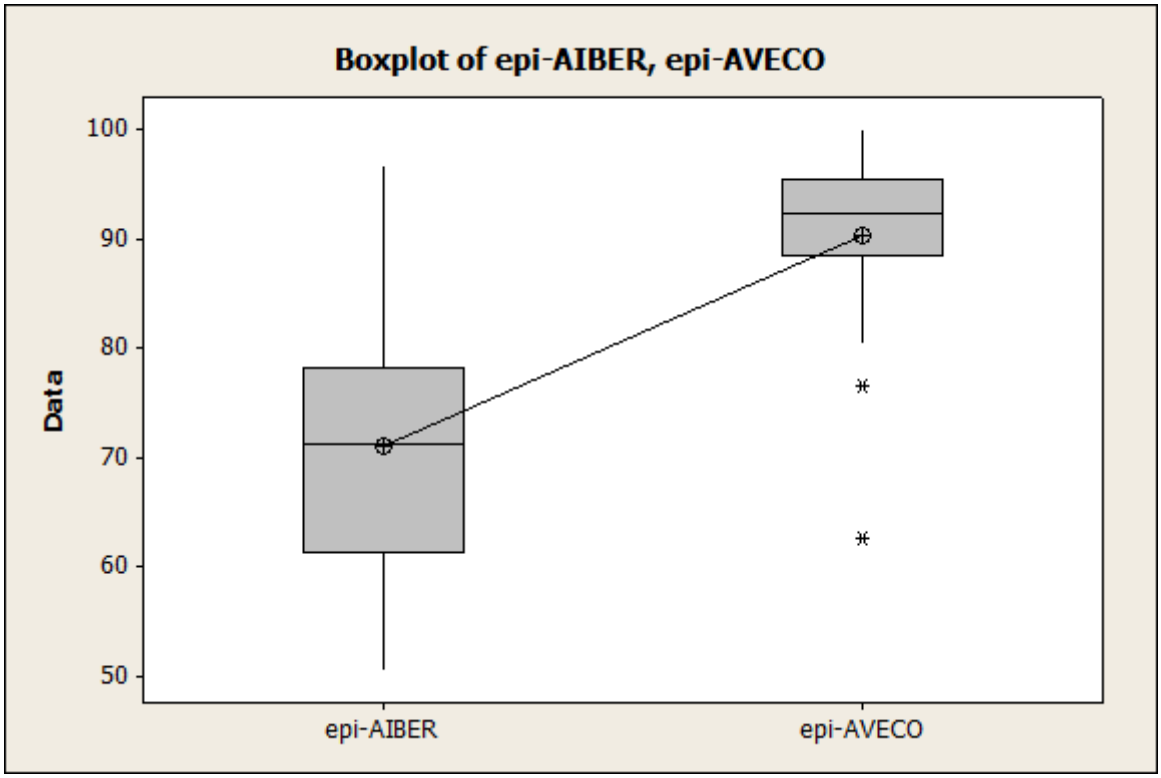


Obs: No gráfico acima, a amostra com 100 países apresentou o menor intervalo, denotando representar melhor a totalidade da população, com 95% de confiança.

Comparação entre as regiões AIBER e AVECO (ANOVA unstacked)

One-way ANOVA: epi-AIBER, epi-AVECO

```
Source  DF      SS      MS      F      P
Factor   1  4371.7  4371.7  44.48  0.000
Error   46  4520.9    98.3
Total   47  8892.6

S = 9.914   R-Sq = 49.16%   R-Sq(adj) = 48.06%

                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level       N    Mean   StDev  -----+---------+---------+---------+----
epi-AIBER  21  71.054  11.913  (------*-----)
epi-AVECO  27  90.292   8.045                           (----*----)
                                -----+---------+---------+---------+----
                                  70.0      77.0      84.0      91.0
Pooled StDev = 9.914
```



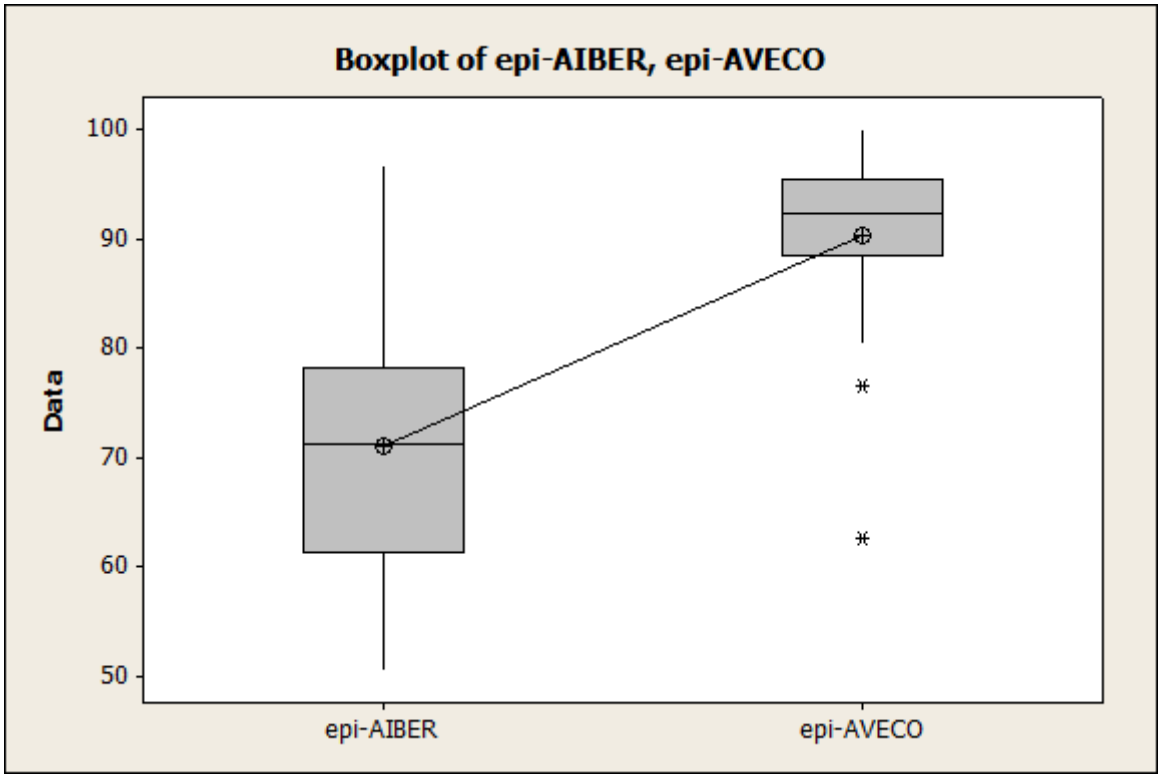
Obs: Dois países do agrupamento AVECO se mostraram como *outliers* , são eles Israel e Republic of Korea.

6.4.3.3 Testes de amostragens (two-sample)

Teste de amostragens (T-Test).

Two-Sample T-Test and CI: epi-AIBER, epi-AVECO

```
Two-sample T for epi-AIBER vs epi-AVECO

            N   Mean  StDev  SE Mean
epi-AIBER  21   71.1   11.9      2.6
epi-AVECO  27  90.29   8.04      1.5

Difference = mu (epi-AIBER) - mu (epi-AVECO)
Estimate for difference:  -19.24
95% upper bound for difference:  -14.12
T-Test of difference = 0 (vs <): T-Value = -6.36  P-Value = 0.000  DF = 33
```



Obs: O teste de "*2-Sample t*" se apresentou praticamente idêntico ao boxplot da sessão anterior com ANOVA (unstacked). Inclusive repetindo Israel e Republic of Korea como *outliers* .

6.5. CONSIDERAÇÕES

O objetivo central deste trabalho foi comparar a região dos países Ibero-Americanos – AIBER, com a região AVECO – Economias Avançadas, em relação às variáveis SPI, IDH e EPI, todas do ano 2014. Para apoiar estas conclusões foram efetuados análises exploratórias dos dados, pesquisas por amostragens, análise de variância (ANOVA), testes de amostragens (two-sample) e gráficos boxplots. Como resultado, obtivemos que a região AVECO apresenta melhores índices de performance que a região AIBER.

Como conclusão secundária, para as pesquisas de amostragens realizadas em agrupamentos de 25, 50 e 100 indivíduos (países), nos casos estudados, a amostra com 100 países apresentou menor intervalo, portanto, representando melhor o comportamento do todo.

7. COMPONENTES PRINCIPAIS

Esta sessão se dedica a apresentar os reultados de análise das correlações e dos componentes principais (*principal components analysis*). Para tal, iniciamos com a análise das correlações e dendrogramas (*dendrogram*). Em seguida partimos para os componentes principais, análise e criação de novos índices .

Importante: Os índices SPI, IDH e EPI foram atualizados para o ano base de 2014. E os valores da variável IDH para os países Suíça, Suécia e Swaziland também foram corrigidos.

7.1. ANÁLISE DE COMPONENTES PRINCIPAIS, DENDROGRAMAS E CORRELAÇÕES

Nesta sessão faremos as análises de componentes principais, dendrogramas (matriz de dissimilaridade) e as correlações e regressões *stepwise*.

7.1.1 ANÁLISES DE COMPONENTES PRINCIPAIS

Nesta primeira análise foram consideradas todas as variáveis deste estudo (total de 30), abaixo os resultados da matrix de correlação e as variáveis e seus gráficos.

Results for: METODOS_new SPI HDI and EPI_version 2014_MINITAB.MTW

Principal Component Analysis: SPI-2014_n, Basic Human , Foundations , Opportuni

Eigenanalysis of the Correlation Matrix

Eigenvalue	13.148	2.856	2.117	1.419	1.344	1.110	0.964	0.897	0.841
Proportion	0.438	0.095	0.071	0.047	0.045	0.037	0.032	0.030	0.028
Cumulative	0.438	0.533	0.604	0.651	0.696	0.733	0.765	0.795	0.823

Eigenvalue	0.794	0.627	0.581	0.515	0.467	0.373	0.319	0.294	0.253
Proportion	0.026	0.021	0.019	0.017	0.016	0.012	0.011	0.010	0.008
Cumulative	0.850	0.871	0.890	0.907	0.923	0.935	0.946	0.955	0.964

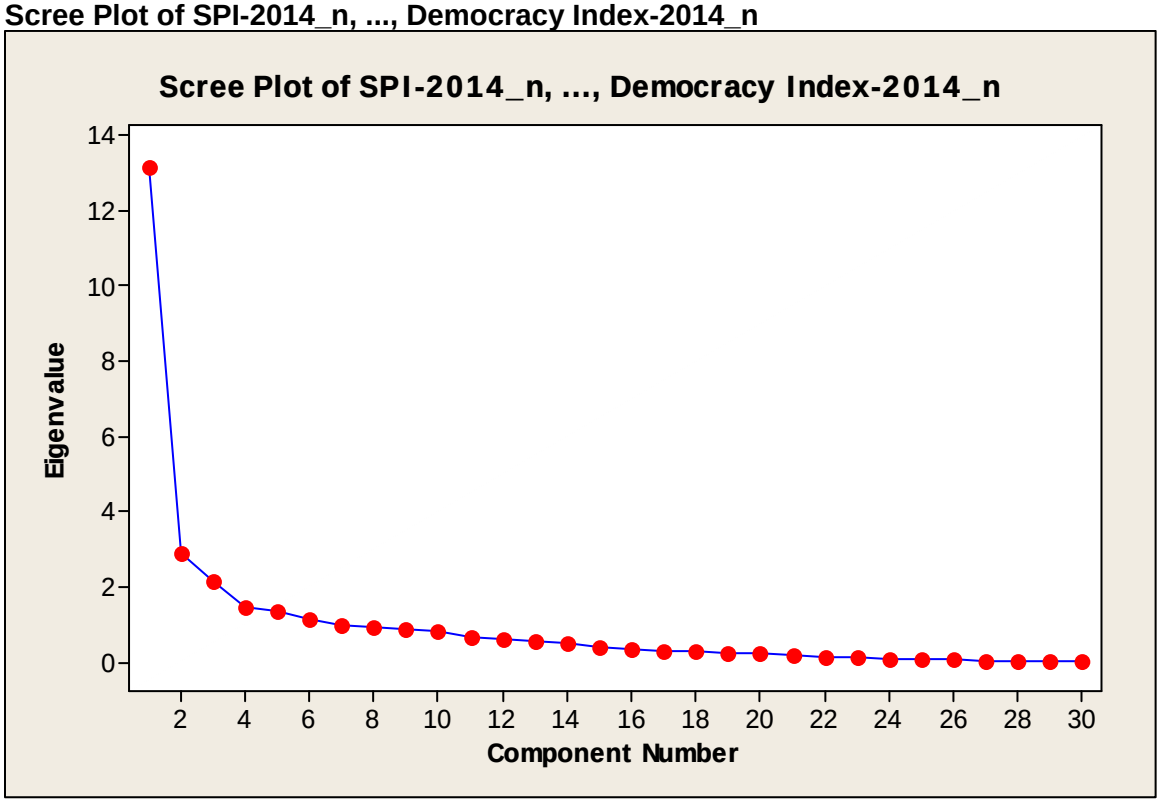
Eigenvalue	0.224	0.203	0.181	0.139	0.100	0.079	0.059	0.038	0.025
Proportion	0.007	0.007	0.006	0.005	0.003	0.003	0.002	0.001	0.001
Cumulative	0.971	0.978	0.984	0.989	0.992	0.995	0.997	0.998	0.999

Eigenvalue	0.020	0.014	0.000
Proportion	0.001	0.000	0.000
Cumulative	1.000	1.000	1.000

Variable	PC1	PC2	PC3	PC4
SPI-2014_n	0.272	-0.002	-0.009	-0.006
Basic Human Needs_n	0.250	0.163	-0.004	0.023
Foundations of Wellbeing_n	0.253	-0.079	0.045	0.006
Opportunity_n	0.259	-0.145	-0.054	-0.050
Health and Wellness_n	0.119	-0.245	0.199	0.505
Personal Freedom and Choice_n	0.245	-0.142	0.047	0.015
Greenhouse gas emissions_np	0.097	0.085	0.131	0.207
Water withdrawals_np	-0.048	-0.311	-0.166	-0.156
Biodiversity and habitat_n	0.093	-0.276	0.025	-0.005
Modern slavery_np	0.172	0.084	0.174	-0.037

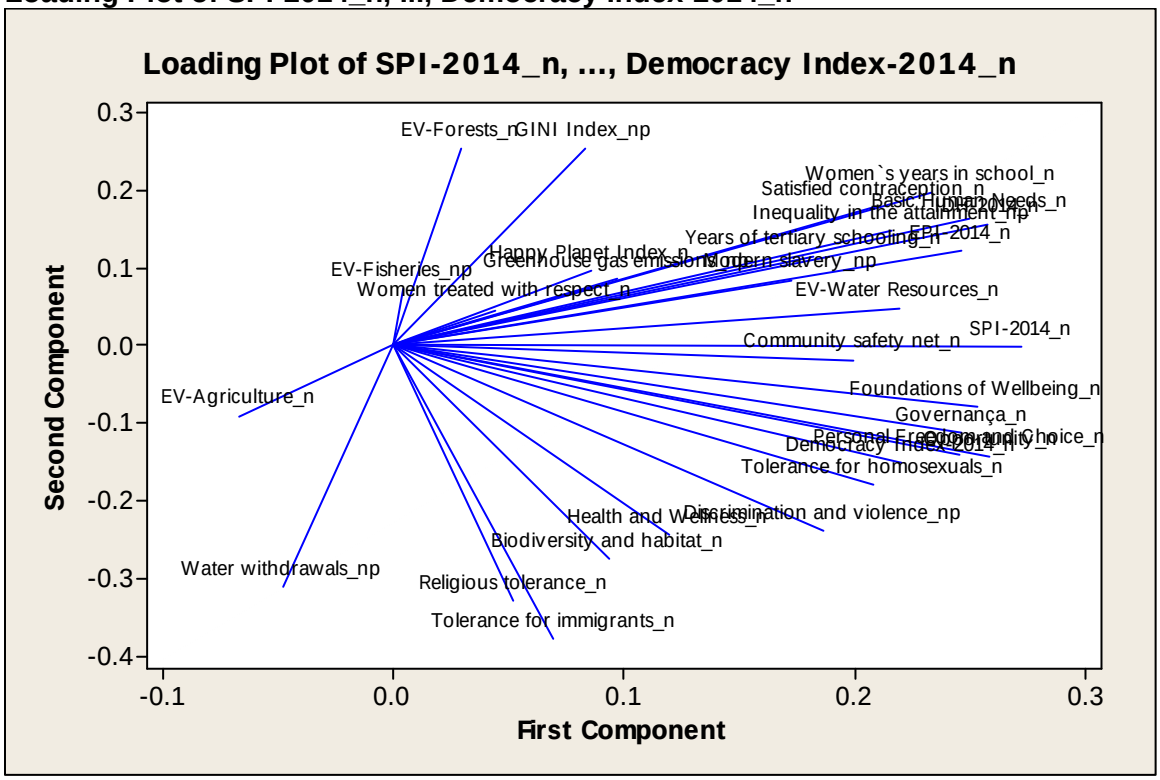
Satisfied contraception_n	0.209	0.179	0.226	-0.132
Women treated with respect_n	0.044	0.045	-0.405	0.465
Tolerance for immigrants_n	0.070	-0.380	-0.139	0.033
Tolerance for homosexuals_n	0.208	-0.181	0.067	0.122
Discrimination and violence_np	0.186	-0.239	-0.170	-0.149
Religious tolerance_n	0.052	-0.330	-0.009	-0.370
Community safety net_n	0.200	-0.019	0.077	-0.023
Years of tertiary schooling_n	0.182	0.113	-0.081	0.040
Women`s years in school_n	0.233	0.197	0.009	-0.221
Inequality in the attainment_np	0.216	0.148	-0.107	-0.239
IDH-2014_n	0.258	0.155	0.016	-0.007
Governança_n	0.246	-0.113	-0.161	0.071
EPI-2014_n	0.246	0.121	0.045	-0.139
EV-Water Resources_n	0.219	0.046	-0.190	0.175
EV-Agriculture_n	-0.067	-0.092	0.184	-0.080
EV-Forests_n	0.029	0.253	-0.194	-0.163
EV-Fisheries_np	0.004	0.071	-0.353	-0.094
Happy Planet Index_n	0.085	0.096	0.495	0.147
GINI Index_np	0.083	0.253	-0.283	0.196
Democracy Index-2014_n	0.220	-0.152	0.015	0.003

Scree Plot of SPI-2014_n, ..., Democracy Index-2014_n



O gráfico mostra visualmente o que vimos mais acima na matriz de correlação, onde denota que a variável SPI explica 43.8% dos resultados e se por convenção considerarmos as componentes com contribuição igual ou maio que 1, teríamos as 6 (seis) primeiras variáveis explicando 73.3% da variabilidade dos dados.

Loading Plot of SPI-2014_n, ..., Democracy Index-2014_n



Este gráfico mostra uma expressiva concentração dos dados na parte superior direita, onde as componentes apresentaram resultados positivos. Apenas duas variáveis apresentaram valores negativos nos dois eixos.

Conforme, Las Casas e Guevara (2010)[2], por convenção somente deve ser consideradas as componentes com uma contribuição igual ou superior a 1. Partindo dessa convenção, listamos abaixo as 6 (seis) variáveis as quais considerando o acumulado, explicam 73.3% da variabilidade dos dados.

Variable	PC1	PC2	PC3	PC4
SPI-2014_n	0.272	-0.002	-0.009	-0.006
Basic Human Needs_n	0.250	0.163	-0.004	0.023
Foundations of Wellbeing_n	0.253	-0.079	0.045	0.006
Opportunity_n	0.259	-0.145	-0.054	-0.050
Health and Wellness_n	0.119	-0.245	0.199	0.505
Personal Freedom and Choice_n	0.245	-0.142	0.047	0.015

[2] LAS CASAS, Alexandre Luzzi; GUEVARA, **Arnoldo José de Hoyos. Pesquisa de Marketing**. São Paulo: Atlas, 2010.

7.2 DENDROGRAMA (Dendrogram)

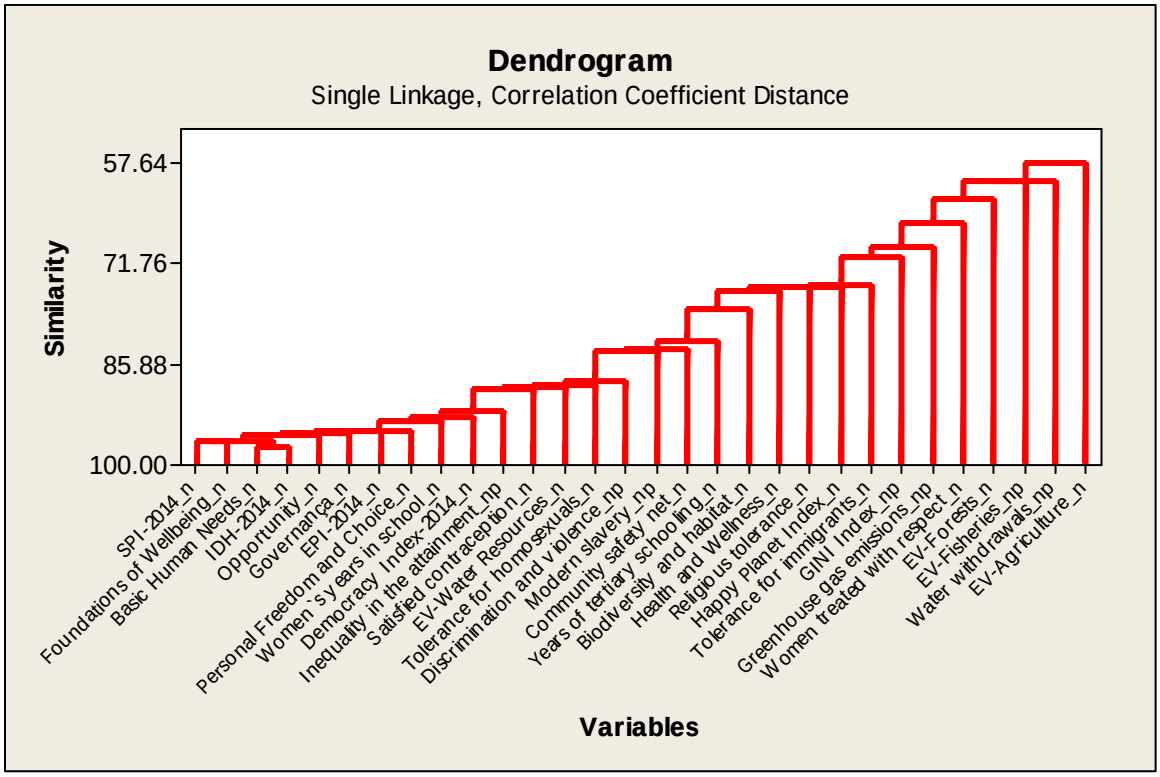
Abaixo os resultados das análises de *clusters*, considerando o grau de similaridade entre as 30 variáveis deste estudo.

Cluster Analysis of Variables: SPI-2014, Basic Human Needs, Foundations of Wellbeing, ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	29	97.8566	0.042868	2	21	2	2
2	28	96.9544	0.060912	1	3	1	2
3	27	96.8730	0.062539	1	2	1	4
4	26	96.0972	0.078056	1	4	1	5
5	25	95.8063	0.083875	1	22	1	6
6	24	95.4891	0.090218	1	23	1	7
7	23	95.3580	0.092840	1	6	1	8
8	22	94.2163	0.115674	1	19	1	9
9	21	93.4716	0.130567	1	30	1	10
10	20	92.5817	0.148366	1	20	1	11
11	19	89.6856	0.206288	1	11	1	12
12	18	89.1677	0.216646	1	24	1	13
13	17	89.0470	0.219061	1	14	1	14
14	16	88.3252	0.233497	1	15	1	15
15	15	84.0927	0.318146	1	10	1	16
16	14	84.0626	0.318747	1	17	1	17
17	13	82.9000	0.342001	1	18	1	18
18	12	78.2833	0.434334	1	9	1	19
19	11	75.8883	0.482233	1	5	1	20
20	10	75.2897	0.494205	1	16	1	21
21	9	75.0977	0.498045	1	28	1	22
22	8	74.9784	0.500432	1	13	1	23
23	7	71.0320	0.579361	1	29	1	24
24	6	69.6158	0.607684	1	7	1	25
25	5	66.3052	0.673897	1	12	1	26
26	4	62.8294	0.743413	1	26	1	27
27	3	60.2922	0.794155	1	27	1	28
28	2	60.1415	0.797170	1	8	1	29
29	1	57.6359	0.847283	1	25	1	30

Dendrogram



O gráfico de Dendrograma (matriz de dissimilaridade) reforça os resultados da tabela acima de análise de clusters, onde seriam necessários 29 clusters para ter similaridade de 97.85%, 13 para chegar a 82.90% e 7 clusters para atingir 71.03% de similaridade entre as variáveis.

7.3 CORRELAÇÕES

As correlações abaixo apresentaram resultados não significativos ($p > 0.05$). E os coeficientes de correlação de Pearson com valores próximos a 0 (zero) indica uma relação fraca ou inexistente entre as duas variáveis.

Correlations: PC1, PC2, PC3, PC4

	PC1	PC2	PC3
PC2	0.000		
	1.000		
PC3	0.000	-0.000	
	1.000	1.000	
PC4	0.000	-0.000	0.000
	1.000	1.000	1.000

Cell Contents: Pearson correlation
P-Value

7.4 Regressões *Stepwise*

Nesta sessão apresentamos os resultados da análise de regressão *stepwise*, utilizando como referência cada variável PC (principal componente) criada, contra as outras 30 variáveis deste trabalho como preditoras.

Stepwise Regression: PC1 versus SPI-2014_n, Basic Human Needs_n, ...

```
  Alpha-to-Enter: 0.15  Alpha-to-Remove: 0.15

Response is PC1 on 30 predictors, with N = 132

Step                        1       2       3       4        5       6
Constant                -7.832  -7.737  -7.874  -9.449  -10.026  -9.872

SPI-2014_n              0.1402  0.1162  0.0572  0.0598   0.0620  0.0631
T-Value                  70.42   25.14    6.89    7.95     9.09   10.16
P-Value                  0.000   0.000   0.000   0.000    0.000   0.000

Opportunity_n                   0.0272  0.0500  0.0470   0.0416  0.0375
T-Value                           5.64   10.25   10.58    10.04    9.74
P-Value                          0.000   0.000   0.000    0.000   0.000

IDH-2014_n                              0.0401  0.0347   0.0309  0.0260
T-Value                                   7.99    7.47     7.26    6.53
P-Value                                  0.000   0.000    0.000   0.000

Modern slavery_np                                0.0209   0.0211  0.0215
T-Value                                            5.42     6.04    6.76
P-Value                                           0.000    0.000   0.000

Community safety net_n                                     0.0126  0.0130
T-Value                                                      5.40    6.15
P-Value                                                     0.000   0.000

EV-Water Resources_n                                               0.0071
T-Value                                                              5.22
P-Value                                                             0.000

S                       0.582   0.523   0.429   0.388    0.351   0.319
R-Sq                    97.45   97.95   98.63   98.89    99.10   99.26
R-Sq(adj)               97.43   97.92   98.60   98.85    99.06   99.22
```

Para o PC1, temos as variáveis SPI-2014, Opportunity e IDH-2014 que explicam 98.63% dos dados.

Stepwise Regression: PC2 versus SPI-2014_n, Basic Human Needs_n, ...

```
  Alpha-to-Enter: 0.15  Alpha-to-Remove: 0.15

Response is PC2 on 30 predictors, with N = 132

Step                            1       2       3       4       5
Constant                    2.512   1.529   2.514   1.056   1.705
```

Tolerance for immigrants_n	-0.0485	-0.0501	-0.0448	-0.0419	-0.0363
T-Value	-9.54	-12.24	-12.46	-13.63	-13.29
P-Value	0.000	0.000	0.000	0.000	0.000
EV-Forests_n		0.0269	0.0231	0.0200	0.0188
T-Value		8.48	8.33	8.36	9.19
P-Value		0.000	0.000	0.000	0.000
Water withdrawals_np			-0.0181	-0.0185	-0.0170
T-Value			-6.89	-8.30	-8.94
P-Value			0.000	0.000	0.000
GINI Index_np				0.0231	0.0200
T-Value				7.18	7.20
P-Value				0.000	0.000
Religious tolerance_n					-0.0132
T-Value					-7.03
P-Value					0.000
Health and Wellness_n					
T-Value					
P-Value					
S	1.30	1.05	0.897	0.760	0.647
R-Sq	41.17	62.22	72.44	80.40	85.91
R-Sq(adj)	40.71	61.63	71.79	79.78	85.36

Para o PC2, as variáveis Tolerância a Imigrantes, Índice de Cobertura de Florestas e Extração de Água, explicam 72.44% dos dados.

Stepwise Regression: PC3 versus SPI-2014_n, Basic Human Needs_n, ...

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15

Response is PC3 on 30 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2.3539	-0.4865	1.3870	2.0243	2.1238
Happy Planet Index_n	0.0486	0.0414	0.0381	0.0387	0.0388
T-Value	11.82	12.35	13.71	16.14	17.44
P-Value	0.000	0.000	0.000	0.000	0.000
Women treated with respect_n		-0.0279	-0.0244	-0.0196	-0.0190
T-Value		-8.87	-9.33	-8.24	-8.61
P-Value		0.000	0.000	0.000	0.000
EV-Fisheries_np			-0.0268	-0.0252	-0.0242
T-Value			-8.02	-8.70	-8.97
P-Value			0.000	0.000	0.000
GINI Index_np				-0.0166	-0.0151
T-Value				-6.68	-6.50
P-Value				0.000	0.000
EV-Forests_n					-0.0076
T-Value					-4.69
P-Value					0.000
Water withdrawals_np					
T-Value					
P-Value					

S	1.01	0.802	0.657	0.568	0.526
R-Sq	51.82	70.06	80.07	85.25	87.44
R-Sq(adj)	51.45	69.60	79.60	84.78	86.94

Para o PC3, temos as variáveis Índice de Felicidade do Planeta, Índice de Respeito à Mulher e Extração de Peixes na Costa, as quais explicam 80.07% dos dados.

Stepwise Regression: PC4 versus SPI-2014_n, Basic Human Needs_n, ...

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15

Response is PC4 on 30 predictors, with N = 132

Step	1	2	3	4	5
Constant	-2.319	-1.436	-2.381	-1.787	-4.187
Health and Wellness_n	0.0343	0.0377	0.0334	0.0335	0.0311
T-Value	8.58	12.20	13.51	18.05	19.95
P-Value	0.000	0.000	0.000	0.000	0.000
Religious tolerance_n		-0.01880	-0.01664	-0.01497	-0.01500
T-Value		-9.54	-10.63	-12.63	-15.34
P-Value		0.000	0.000	0.000	0.000
Women treated with respect_n			0.0204	0.0231	0.0232
T-Value			9.02	13.45	16.45
P-Value			0.000	0.000	0.000
Inequality in the attainment_np				-0.0135	-0.0152
T-Value				-10.04	-13.45
P-Value				0.000	0.000
Greenhouse gas emissions_np					0.0275
T-Value					7.79
P-Value					0.000
GINI Index_np					
T-Value					
P-Value					
S	0.955	0.735	0.576	0.432	0.356
R-Sq	36.16	62.56	77.12	87.24	91.39
R-Sq(adj)	35.66	61.98	76.58	86.84	91.05

Para o PC4, as variáveis Saúde e Bem-Estar, Tolerância à Religião e Índice de Respeito à Mulher, explicam 77.12% dos dados.

7.5 Criando indicadores a partir dos PC1, PC2, PC3 e PC4

Aproveitamos os resultados da sessão anterior, que através da análise de regressão *stepwise*, relacionamos cada componente principal (PC), com as 30 variáveis afim de obter a mais alta correlação com as variáveis de resposta.

Abaixo, listamos as 3 (três) variáveis que apresentaram esta correlação, uma sugestão para o nome do indicador proposto e uma sugestão para o nome da variável.

```
PC1 = SPI + Opportunity + IDH
Nome do Indicador = Growth or Development (crescimento ou
desenvolvimento)   = socoppi

PC2 = Immigrants Tol. + EV-Forests + Water withdrawals
Nome do Indicador = Careful (consciência ou atento)  = immigepi

PC3 = HPI + Women treated w/respect + EV-Fisheries
Nome do Indicador = Happiness (Felicidade)  = hpiwfish

PC4 = Health & Wellness + Religious Tol. + Women treated w/respect
Nome do Indicador = Well-being or Take care (cuidado ou estar bem)
    = wellreliwo
```

7.6. CONSIDERAÇÕES

Nesta sessão foram realizadas análises de componentes principais para a construção de índices com alto poder explicativo, utilizando o menor número possível de variáveis.

De acordo com as análises realizadas, notamos que o agrupamento de variáveis é pertinente, porém, por alguma característica dos dados, aparentemente há um viés mais social do que natural (ecológico) nos indicadores gerados, ou seja, das 11 variáveis que se despontaram na análise, temos 8 delas ligadas a índices que medem dimensões ligadas ao indivíduo ou seu ambiente social e apenas 3 variáveis representando o uso de algum recurso natural do meio ambiente.

A variável *Biodiversity and Habitat* por exemplo, não se destacou em nenhuma das análises aqui realizadas.

8. ANÁLISE DE CONGLOMERADOS

O objetivo desta sessão é efetuar análise de conglomerados, agrupar dados de acordo com sua similaridade a partir de pontos definidos pelas médias das variáveis para os indivíduos de cada agrupamento (centróides). Será realizado também análise de dendrogramas e clusters. Portanto, iniciamos com a análise dos agrupamentos e dendrogramas

(*dendrogram*). Em seguida partimos para categorização e separação por região, análises com a nova amostra e gráficos de dispersão (s*catterplot 3D*).

8.1. METODOLOGIA

Nesta sessão faremos alguns ensaios (análise multivariada) a fim de identificar agrupamentos por grau de similaridade e a partir de então a amostra será categorizada e separada por região, para que possamos comparar AIBER e AVECO. A saber, serão considerados para este estudo somente países integrantes das regiões AIBER e AVECO, os demais países (região OTHERS) não serão contemplados nas análises.

As variáveis para este estudo, serão: EPI-2014, Governança, Biodiversity and habitat, Democracy Index-2014, SPI-2014, Opportunity, IDH-2014, Tolerance for immigrants, EV-Forests, Water withdrawals, HPI, Women treated with respect, EV-Fisheries, Health and Wellness, Religious tolerance. E os índices PC1, PC2, PC3 e PC4, gerados a partir da funcionalidade de components principais. Estes PCs foram gerados considerando as 30 variáveis referentes ao tema, como preditoras e apresentaram os seguintes resultados em sua composição:

Para o PC1, as variáveis SPI-2014, Opportunity e IDH-2014 que explicam 98.63% dos dados.

Para o PC2, Tolerância a Imigrantes, Índice de Cobertura de Florestas e Extração de Água, explicam 72.44% dos dados.

Para o PC3, as variáveis Índice de Felicidade do Planeta, Índice de Respeito à Mulher e Extração de Peixes na Costa, as quais explicam 80.07% dos dados.

E para o PC4, Saúde e Bem-Estar, Tolerância à Religião e Índice de Respeito à Mulher, explicam 77.12% dos dados.

Em tempo, optou-se em adicionar a este trabalho as variáveis Governança, Democracy Index-2014 e Biodiversity and Habitat, sendo as duas primeiras devido ao forte viés social nos índices PCs, e a terceira por fazer parte desta temática mas que não apresentou relevância até o momento nas análises.

8.2 Dendrogramas e Análise de Clusters (por variáveis e por observações)

Para encontrar a melhor combinação de agrupamento para as variáveis deste estudo, serão realizadas análises com 1, 10, 20 e 30 clusters, abaixo estão listadas apenas algumas para exemplificação.

8.2.1 Análise com 20 Clusters (por variáveis e por observações)

Análise com 20 Clusters (por variáveis e por observações) e Dendrogramas

---

Cluster Analysis of Observations: EPI-2014_n, Governança_n, Biodiversity, ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	89.5016	28.1919	32	108	32	2
2	130	88.8134	30.0399	43	88	43	2
3	129	88.7094	30.3193	32	99	32	3
4	128	88.3161	31.3755	49	90	49	2
5	127	87.2973	34.1112	11	127	11	2
6	126	86.7834	35.4914	46	49	46	3
7	125	86.0066	37.5771	7	93	7	2
8	124	85.7307	38.3182	18	91	18	2
9	123	85.6685	38.4851	7	115	7	3
10	122	85.3638	39.3034	43	126	43	3
11	121	85.0209	40.2243	21	119	21	2
12	120	84.7665	40.9073	116	129	116	2
13	119	84.4339	41.8006	21	28	21	3
14	118	84.3370	42.0607	41	43	41	4
15	117	84.0043	42.9541	7	40	7	4
16	116	83.9718	43.0414	13	46	13	4
17	115	83.8607	43.3397	7	114	7	5
18	114	83.8089	43.4789	12	18	12	3
19	113	83.5603	44.1464	19	21	19	4
20	112	83.4068	44.5586	63	92	63	2
21	111	83.3470	44.7193	7	33	7	6
22	110	83.3451	44.7244	68	72	68	2
23	109	83.2021	45.1083	13	36	13	5
24	108	83.0136	45.6146	13	97	13	6
25	107	83.0057	45.6358	13	35	13	7
26	106	82.6211	46.6685	5	8	5	2
27	105	82.2042	47.7881	7	32	7	9
28	104	82.1788	47.8562	60	110	60	2
29	103	81.9167	48.5602	13	38	13	8
30	102	81.8817	48.6541	22	51	22	2
31	101	81.8590	48.7150	75	132	75	2
32	100	81.8290	48.7956	7	41	7	13
33	99	81.8213	48.8163	7	30	7	14
34	98	81.8149	48.8335	7	22	7	16
35	97	81.7963	48.8833	19	78	19	5
36	96	81.6484	49.2806	53	118	53	2
37	95	81.6243	49.3452	117	123	117	2
38	94	81.5449	49.5585	4	16	4	2
39	93	81.3389	50.1117	7	107	7	17
40	92	81.1378	50.6517	58	60	58	3
41	91	81.0669	50.8420	12	105	12	4
42	90	81.0510	50.8848	17	83	17	2
43	89	81.0269	50.9495	45	80	45	2
44	88	80.9641	51.1181	47	63	47	3
45	87	80.8449	51.4381	75	117	75	4
46	86	80.7638	51.6559	12	75	12	8
47	85	80.7365	51.7294	15	86	15	2
48	84	80.7012	51.8242	14	81	14	2
49	83	80.5688	52.1798	4	27	4	3
50	82	80.3391	52.7965	13	95	13	9
51	81	80.3181	52.8530	61	66	61	2
52	80	80.2338	53.0792	94	131	94	2
53	79	80.1247	53.3723	7	11	7	19
54	78	80.0838	53.4820	12	47	12	11
55	77	80.0127	53.6729	19	103	19	6
56	76	79.6069	54.7628	102	124	102	2

57	75	79.5361	54.9527	71	77	71	2
58	74	79.4987	55.0532	4	13	4	12
59	73	78.5698	57.5477	7	101	7	20
60	72	78.5578	57.5798	3	12	3	12
61	71	78.5390	57.6303	3	112	3	13
62	70	78.5047	57.7224	4	29	4	13
63	69	78.4412	57.8930	65	104	65	2
64	68	78.3884	58.0349	3	85	3	14
65	67	78.3749	58.0710	3	19	3	20
66	66	78.3673	58.0916	52	94	52	3
67	65	78.2565	58.3891	61	69	61	3
68	64	78.2292	58.4623	64	100	64	2
69	63	78.1194	58.7571	3	24	3	21
70	62	78.1078	58.7883	7	58	7	23
71	61	78.0899	58.8365	7	64	7	25
72	60	78.0445	58.9583	2	122	2	2
73	59	78.0431	58.9622	5	26	5	3
74	58	77.9357	59.2505	7	68	7	27
75	57	77.9085	59.3235	45	98	45	3
76	56	77.7467	59.7579	7	45	7	30
77	55	77.4969	60.4289	2	61	2	5
78	54	77.4808	60.4721	7	17	7	32
79	53	77.4142	60.6510	84	121	84	2
80	52	77.2200	61.1723	2	7	2	37
81	51	77.1776	61.2864	5	54	5	4
82	50	77.0991	61.4970	2	53	2	39
83	49	76.8489	62.1689	5	84	5	6
84	48	76.8296	62.2207	2	50	2	40
85	47	76.7724	62.3743	2	5	2	46
86	46	76.7586	62.4114	2	130	2	47
87	45	76.7416	62.4571	2	31	2	48
88	44	76.6390	62.7326	65	125	65	3
89	43	76.6126	62.8034	4	96	4	14
90	42	76.5211	63.0491	2	39	2	49
91	41	76.4917	63.1280	2	37	2	50
92	40	76.4707	63.1846	2	4	2	64
93	39	76.3227	63.5820	1	2	1	65
94	38	76.1063	64.1630	1	76	1	66
95	37	76.0938	64.1966	1	15	1	68
96	36	75.5654	65.6156	6	89	6	2
97	35	75.5221	65.7319	1	128	1	69
98	34	75.3933	66.0778	82	120	82	2
99	33	75.2795	66.3832	1	42	1	70
100	32	75.2023	66.5906	1	116	1	72

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	85	695709	88.6533	154.971
Cluster2	25	100090	62.5368	85.858
Cluster3	3	5129	41.1018	44.547
Cluster4	1	0	0.0000	0.000
Cluster5	1	0	0.0000	0.000
Cluster6	1	0	0.0000	0.000
Cluster7	2	2354	34.3084	34.308
Cluster8	1	0	0.0000	0.000
Cluster9	1	0	0.0000	0.000
Cluster10	1	0	0.0000	0.000
Cluster11	1	0	0.0000	0.000
Cluster12	1	0	0.0000	0.000
Cluster13	1	0	0.0000	0.000
Cluster14	1	0	0.0000	0.000
Cluster15	1	0	0.0000	0.000
Cluster16	1	0	0.0000	0.000

Cluster17	1	0	0.0000	0.000
Cluster18	1	0	0.0000	0.000
Cluster19	2	1499	27.3814	27.381
Cluster20	1	0	0.0000	0.000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
EPI-2014_n	73.1199	23.7148	93.642	8.716	84.360
Governança_n	52.3117	24.0948	92.669	19.989	21.841
Biodiversity and habitat_n	64.2159	57.5424	59.290	39.680	41.250
Democracy Index-2014_n	58.5475	32.5640	88.942	50.829	26.066
Opportunity_n	53.3460	22.1821	95.892	23.752	31.115
IDH-2014_n	71.7066	19.5627	96.215	37.246	75.570
Tolerance for immigrants_n	49.6696	61.7534	90.411	50.685	35.616
EV-Forests_n	40.8304	28.1260	100.000	22.830	68.230
Water withdrawals_np	54.6389	93.0493	48.131	87.049	88.172
HPI_n	54.5616	24.7487	54.824	81.315	35.767
Women treated with respect_n	53.0723	55.3171	82.114	78.049	60.976
EV-Fisheries_np	70.4019	70.0554	77.959	100.000	71.024
Health and Wellness_n	72.1929	59.9820	90.859	79.958	19.892
Religious tolerance_n	60.0000	57.3333	77.778	0.000	100.000
SPI-2014_n	66.9087	21.6542	96.208	34.939	58.591

Variable	Cluster6	Cluster7	Cluster8	Cluster9	Cluster10
EPI-2014_n	26.390	29.6006	63.531	50.1493	76.596
Governança_n	24.541	27.7585	34.564	7.3225	65.156
Biodiversity and habitat_n	78.930	16.5700	18.660	0.0000	39.340
Democracy Index-2014_n	38.981	18.3649	52.370	32.4645	72.749
Opportunity_n	26.719	24.6075	45.589	4.7724	51.115
IDH-2014_n	34.670	25.5671	48.262	51.3087	91.672
Tolerance for immigrants_n	6.849	48.6301	49.315	23.2877	16.438
EV-Forests_n	0.000	46.0350	63.120	39.7200	39.720
Water withdrawals_np	91.154	35.0405	64.413	30.3363	3.492
HPI_n	42.784	41.8981	68.954	64.1798	78.689
Women treated with respect_n	91.463	64.6341	54.354	15.8537	53.659
EV-Fisheries_np	100.000	85.5118	74.708	71.0235	100.000
Health and Wellness_n	83.858	36.4686	39.064	54.7555	83.228
Religious tolerance_n	100.000	83.3333	100.000	0.0000	0.000
SPI-2014_n	34.669	26.6175	49.353	21.9986	69.734

Variable	Cluster11	Cluster12	Cluster13	Cluster14
EPI-2014_n	74.505	67.5439	24.6174	18.794
Governança_n	46.453	26.4472	22.8075	44.183
Biodiversity and habitat_n	88.040	14.1100	93.8500	1.300
Democracy Index-2014_n	69.905	19.9052	8.5308	61.256
Opportunity_n	64.819	37.9278	19.3721	44.364
IDH-2014_n	62.166	73.8266	38.0762	24.895
Tolerance for immigrants_n	73.973	43.8356	58.9041	32.877
EV-Forests_n	18.520	74.8200	13.2800	39.720
Water withdrawals_np	0.000	19.6290	99.7169	20.547
HPI_n	86.724	29.2259	64.0348	48.454
Women treated with respect_n	14.634	68.2927	86.5854	53.659
EV-Fisheries_np	100.000	71.0235	71.0235	71.024
Health and Wellness_n	81.908	0.0000	74.8275	44.614
Religious tolerance_n	100.000	66.6667	66.6667	100.000
SPI-2014_n	67.919	48.2926	35.6039	29.367

Variable	Cluster15	Cluster16	Cluster17	Cluster18
EPI-2014_n	76.3718	0.000	62.990	24.4681
Governança_n	44.7385	22.606	69.407	22.5662
Biodiversity and habitat_n	27.5400	37.080	43.700	61.7900
Democracy Index-2014_n	56.3981	34.716	79.147	38.8626
Opportunity_n	48.5400	26.515	61.068	20.6907
IDH-2014_n	66.9884	27.133	71.908	33.4578

Tolerance for immigrants_n	42.4658	46.575	80.822	54.7945
EV-Forests_n	25.3400	7.750	100.000	63.1200
Water withdrawals_np	47.0850	75.049	61.340	51.9583
HPI_n	13.7121	58.475	33.748	55.5695
Women treated with respect_n	57.3171	25.610	68.293	40.2439
EV-Fisheries_np	71.0235	100.000	23.915	71.0235
Health and Wellness_n	61.5662	78.938	80.408	71.5872
Religious tolerance_n	66.6667	66.667	66.667	33.3333
SPI-2014_n	64.2164	20.992	73.832	34.1121

Variable	Cluster19	Cluster20	Grand centroid
EPI-2014_n	83.0627	53.0982	60.8064
Governança_n	25.5564	36.2520	45.0180
Biodiversity and habitat_n	47.4250	67.7000	59.5132
Democracy Index-2014_n	34.5379	49.7630	51.9334
Opportunity_n	41.8289	29.0267	45.7947
IDH-2014_n	71.2168	68.5676	59.6370
Tolerance for immigrants_n	27.3973	32.8767	51.8057
EV-Forests_n	33.7950	33.7600	39.7221
Water withdrawals_np	66.6924	39.8425	61.3499
HPI_n	32.4759	64.6437	48.4542
Women treated with respect_n	34.1463	67.0732	54.3488
EV-Fisheries_np	69.5419	0.0000	71.0251
Health and Wellness_n	5.9256	74.0774	67.5908
Religious tolerance_n	16.6667	0.0000	59.3434
SPI-2014_n	54.3674	48.7239	55.8449

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0.000	111.928	114.173	127.244	100.199	125.431	117.111
Cluster2	111.928	0.000	201.747	99.358	120.286	98.246	87.759
Cluster3	114.173	201.747	0.000	207.816	165.596	210.950	195.919
Cluster4	127.244	99.358	207.816	0.000	165.912	127.157	127.741
Cluster5	100.199	120.286	165.596	165.912	0.000	136.457	107.542
Cluster6	125.431	98.246	210.950	127.157	136.457	0.000	122.288
Cluster7	117.111	87.759	195.919	127.741	107.542	122.288	0.000
Cluster8	83.402	110.326	147.540	140.111	71.715	131.373	78.896
Cluster9	137.179	132.577	228.126	120.014	155.286	177.371	116.016
Cluster10	101.543	180.615	153.314	145.266	173.208	184.458	164.015
Cluster11	100.562	169.621	149.909	182.536	165.418	167.758	156.466
Cluster12	117.635	140.148	175.839	169.148	87.443	172.895	86.641
Cluster13	118.360	74.472	204.746	104.908	134.147	81.789	121.142
Cluster14	120.452	117.649	188.487	147.192	132.095	135.951	63.822
Cluster15	60.528	110.143	141.390	141.484	95.012	128.669	101.596
Cluster16	124.917	73.465	217.871	93.589	148.809	101.635	93.745
Cluster17	93.819	149.498	90.438	174.437	132.320	184.928	153.976
Cluster18	93.048	72.535	180.152	88.857	127.666	127.986	89.405
Cluster19	98.843	121.433	188.501	140.776	99.318	160.240	124.420
Cluster20	105.303	135.688	185.527	131.596	158.050	167.521	152.309

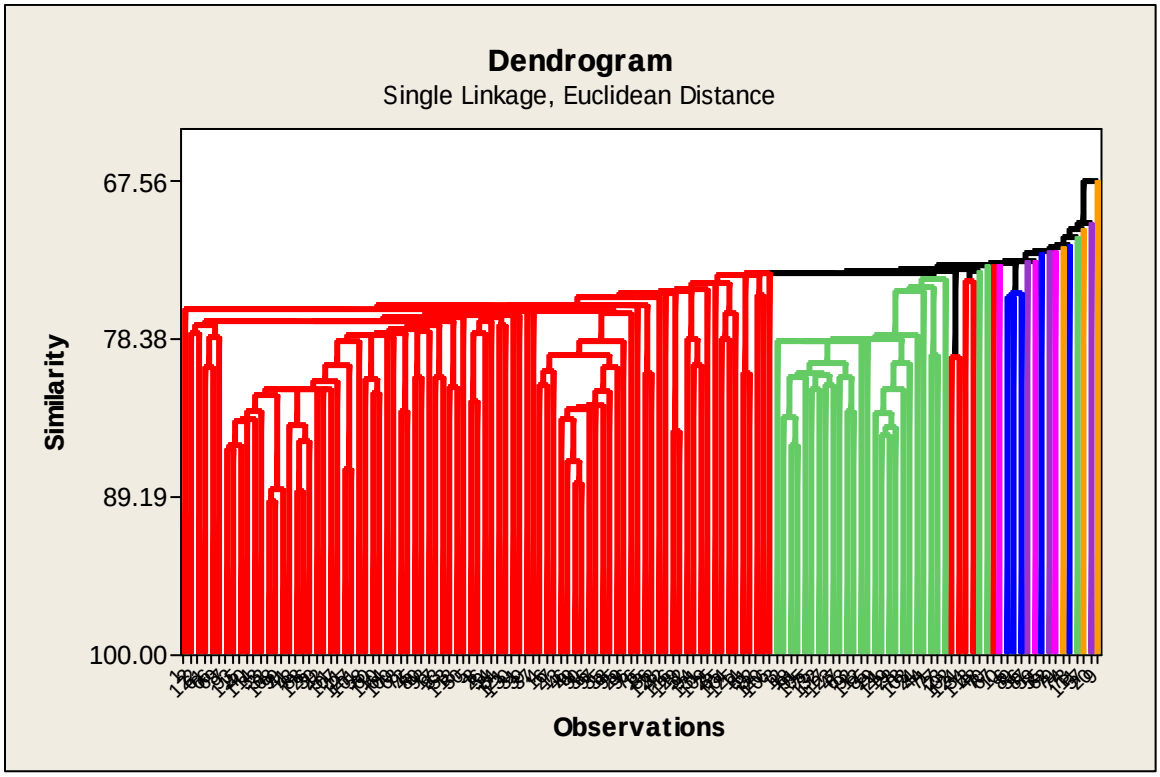
	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	83.402	137.179	101.543	100.562	117.635	118.360
Cluster2	110.326	132.577	180.615	169.621	140.148	74.472
Cluster3	147.540	228.126	153.314	149.909	175.839	204.746
Cluster4	140.111	120.014	145.266	182.536	169.148	104.908
Cluster5	71.715	155.286	173.208	165.418	87.443	134.147
Cluster6	131.373	177.371	184.458	167.758	172.895	81.789
Cluster7	78.896	116.016	164.015	156.466	86.641	121.142
Cluster8	0.000	135.056	149.349	131.931	92.010	132.277
Cluster9	135.056	0.000	131.635	182.030	128.476	165.763
Cluster10	149.349	131.635	0.000	138.408	151.909	187.218
Cluster11	131.931	182.030	138.408	0.000	170.197	168.610
Cluster12	92.010	128.476	151.909	170.197	0.000	166.436
Cluster13	132.277	165.763	187.218	168.610	166.436	0.000
Cluster14	81.415	132.133	159.060	147.580	111.344	153.847
Cluster15	86.080	130.287	119.344	130.094	98.453	138.216
Cluster16	117.932	121.437	170.247	152.934	160.328	101.466

Cluster17	114.685	182.459	156.109	160.908	141.031	168.718
Cluster18	104.226	97.677	137.606	143.928	125.100	100.478
Cluster19	115.447	108.923	133.128	167.404	97.502	147.341
Cluster20	148.101	124.851	127.559	172.857	150.406	138.286

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	120.452	60.528	124.917	93.819	93.048	98.843
Cluster2	117.649	110.143	73.465	149.498	72.535	121.433
Cluster3	188.487	141.390	217.871	90.438	180.152	188.501
Cluster4	147.192	141.484	93.589	174.437	88.857	140.776
Cluster5	132.095	95.012	148.809	132.320	127.666	99.318
Cluster6	135.951	128.669	101.635	184.928	127.986	160.240
Cluster7	63.822	101.596	93.745	153.976	89.405	124.420
Cluster8	81.415	86.080	117.932	114.685	104.226	115.447
Cluster9	132.133	130.287	121.437	182.459	97.677	108.923
Cluster10	159.060	119.344	170.247	156.109	137.606	133.128
Cluster11	147.580	130.094	152.934	160.908	143.928	167.404
Cluster12	111.344	98.453	160.328	141.031	125.100	97.502
Cluster13	153.847	138.216	101.466	168.718	100.478	147.341
Cluster14	0.000	103.215	106.740	147.689	112.843	146.185
Cluster15	103.215	0.000	126.824	110.384	111.043	92.822
Cluster16	106.740	126.824	0.000	179.616	85.828	144.673
Cluster17	147.689	110.384	179.616	0.000	130.809	153.760
Cluster18	112.843	111.043	85.828	130.809	0.000	113.753
Cluster19	146.185	92.822	144.673	153.760	113.753	0.000
Cluster20	158.304	124.272	158.085	132.158	105.496	120.359

	Cluster20
Cluster1	105.303
Cluster2	135.688
Cluster3	185.527
Cluster4	131.596
Cluster5	158.050
Cluster6	167.521
Cluster7	152.309
Cluster8	148.101
Cluster9	124.851
Cluster10	127.559
Cluster11	172.857
Cluster12	150.406
Cluster13	138.286
Cluster14	158.304
Cluster15	124.272
Cluster16	158.085
Cluster17	132.158
Cluster18	105.496
Cluster19	120.359
Cluster20	0.000

Dendrogram



O gráfico e tabela acima mostram a alta similaridade dos dados, tanto que 83% das observações estão concetradas em apenas 2 clusters (cluster 1 com 85 e cluster 2 com 25 obs), portanto não justifica considerar 20 clusters nesta análise.

8.2.2 Análise utilizando somente os PCs (por variáveis e por observações)

Análise com os 4 PCs com 20 clusters (por variáveis e por observações) e Dendrogramas

---

Cluster Analysis of Observations: PC1, PC2, PC3, PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98.7552	0.19128	114	115	114	2
2	130	98.4141	0.24369	103	123	103	2
3	129	98.1862	0.27871	93	114	93	3
4	128	97.2262	0.42623	60	110	60	2
5	127	97.1299	0.44103	44	67	44	2
6	126	97.0271	0.45682	4	59	4	2
7	125	96.8351	0.48633	41	60	41	3
8	124	96.7902	0.49322	99	107	99	2
9	123	96.7335	0.50193	2	53	2	2
10	122	96.5366	0.53219	36	38	36	2
11	121	96.2815	0.57138	40	93	40	4
12	120	96.0868	0.60131	32	108	32	2
13	119	96.0476	0.60734	35	36	35	3
14	118	95.9672	0.61968	21	28	21	2
15	117	95.8917	0.63128	35	97	35	4

16	116	95.7692	0.65011	68	72	68	2
17	115	95.6201	0.67302	43	126	43	2
18	114	95.6163	0.67361	27	35	27	5
19	113	95.5936	0.67709	16	95	16	2
20	112	95.5590	0.68241	82	109	82	2
21	111	95.4974	0.69187	13	96	13	2
22	110	95.4384	0.70094	22	40	22	5
23	109	95.3932	0.70789	26	121	26	2
24	108	95.3510	0.71436	6	56	6	2
25	107	95.2380	0.73174	76	98	76	2
26	106	95.0313	0.76349	6	22	6	7
27	105	94.8966	0.78420	8	129	8	2
28	104	94.7924	0.80020	30	106	30	2
29	103	94.7830	0.80165	84	116	84	2
30	102	94.6836	0.81693	33	88	33	2
31	101	94.5041	0.84451	11	127	11	2
32	100	94.4841	0.84758	12	18	12	2
33	99	94.4555	0.85198	50	68	50	3
34	98	94.4372	0.85479	7	43	7	3
35	97	94.3747	0.86439	94	131	94	2
36	96	94.3608	0.86653	26	104	26	3
37	95	94.3421	0.86940	6	7	6	10
38	94	94.2682	0.88075	39	64	39	2
39	93	94.2146	0.88899	6	33	6	12
40	92	94.1221	0.90320	26	84	26	5
41	91	94.0396	0.91588	6	89	6	13
42	90	94.0362	0.91640	75	85	75	2
43	89	94.0224	0.91852	102	124	102	2
44	88	93.9572	0.92854	26	37	26	6
45	87	93.9294	0.93281	5	14	5	2
46	86	93.9125	0.93542	2	111	2	3
47	85	93.8288	0.94828	13	86	13	3
48	84	93.8125	0.95077	30	101	30	3
49	83	93.8012	0.95252	39	99	39	4
50	82	93.7902	0.95421	8	69	8	3
51	81	93.7757	0.95643	3	34	3	2
52	80	93.7153	0.96572	11	41	11	5
53	79	93.6151	0.98112	3	92	3	3
54	78	93.5526	0.99071	81	102	81	3
55	77	93.5380	0.99296	5	42	5	3
56	76	93.5034	0.99828	20	103	20	3
57	75	93.3348	1.02418	32	39	32	6
58	74	93.3290	1.02508	70	113	70	2
59	73	93.3196	1.02652	20	117	20	4
60	72	93.2910	1.03092	8	54	8	4
61	71	93.2750	1.03337	27	130	27	6
62	70	93.2267	1.04080	5	81	5	6
63	69	93.2009	1.04476	6	11	6	18
64	68	93.1637	1.05047	20	74	20	5
65	67	93.0182	1.07284	1	73	1	2
66	66	93.0037	1.07506	5	8	5	10
67	65	92.9176	1.08829	3	87	3	4
68	64	92.8834	1.09355	5	61	5	11
69	63	92.8260	1.10236	12	105	12	3
70	62	92.8178	1.10363	5	62	5	12
71	61	92.8125	1.10444	27	90	27	7
72	60	92.7887	1.10810	48	70	48	3
73	59	92.7442	1.11494	30	82	30	5
74	58	92.6035	1.13655	6	32	6	24
75	57	92.6033	1.13659	6	51	6	25
76	56	92.5251	1.14860	20	44	20	7
77	55	92.5114	1.15070	20	75	20	9
78	54	92.4756	1.15620	10	17	10	2
79	53	92.4256	1.16390	16	27	16	9
80	52	92.3502	1.17547	5	10	5	14
81	51	92.3228	1.17968	5	50	5	17
82	50	92.2246	1.19478	13	16	13	12
83	49	92.1583	1.20496	13	49	13	13
84	48	92.0808	1.21687	1	30	1	7
85	47	92.0455	1.22230	5	66	5	18

86	46	92.0184	1.22646	13	46	13	14
87	45	91.9568	1.23593	2	118	2	4
88	44	91.9346	1.23934	6	25	6	26
89	43	91.8775	1.24811	5	26	5	24
90	42	91.7722	1.26429	6	58	6	27
91	41	91.7204	1.27226	20	21	20	11
92	40	91.6648	1.28080	4	13	4	16
93	39	91.6170	1.28814	77	91	77	2
94	38	91.4558	1.31291	3	9	3	5
95	37	91.4309	1.31674	6	100	6	28
96	36	91.3382	1.33099	24	47	24	2
97	35	91.2670	1.34192	5	122	5	25
98	34	91.2150	1.34991	1	48	1	10
99	33	91.2103	1.35064	2	5	2	29
100	32	91.1761	1.35589	12	20	12	14

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	60	475.981	2.71447	4.20087
Cluster2	6	11.755	1.34764	1.82405
Cluster3	28	81.312	1.60967	2.81252
Cluster4	17	42.485	1.50234	2.35890
Cluster5	1	0.000	0.00000	0.00000
Cluster6	3	2.319	0.86690	0.97743
Cluster7	1	0.000	0.00000	0.00000
Cluster8	1	0.000	0.00000	0.00000
Cluster9	1	0.000	0.00000	0.00000
Cluster10	1	0.000	0.00000	0.00000
Cluster11	1	0.000	0.00000	0.00000
Cluster12	1	0.000	0.00000	0.00000
Cluster13	2	0.950	0.68905	0.68905
Cluster14	1	0.000	0.00000	0.00000
Cluster15	2	1.208	0.77702	0.77702
Cluster16	2	0.374	0.43220	0.43220
Cluster17	1	0.000	0.00000	0.00000
Cluster18	1	0.000	0.00000	0.00000
Cluster19	1	0.000	0.00000	0.00000
Cluster20	1	0.000	0.00000	0.00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	-0.034390	-4.26267	5.16890	-3.96409	0.22869	-6.77750	-8.13369
PC2	0.823283	0.42291	-0.57848	-1.68652	-2.26910	-1.20122	-2.67723
PC3	0.711425	0.06524	-0.79642	-0.93628	-1.28844	-1.19452	-2.01699
PC4	-0.545134	1.09189	0.56372	0.45429	-1.25497	-0.19663	-2.36384

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
PC1	3.38342	2.80635	-4.12551	2.98336	-2.89933	1.31972
PC2	-2.46450	1.51104	3.46120	3.55180	-0.39931	-0.06594
PC3	3.56668	0.12352	1.26895	0.39320	1.94931	-1.39402
PC4	-0.47128	0.88649	0.61847	1.62904	0.50840	1.15299

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
PC1	-7.00312	1.67211	-5.45267	-6.25648	-5.86159	3.01773
PC2	-1.30180	-0.96676	1.92967	1.25796	-2.34410	-1.32301
PC3	-2.70399	0.47422	1.11733	-0.37244	-1.76740	-0.41917
PC4	0.93113	-0.90993	2.36377	0.89572	-0.28686	-1.82718

Variable	Cluster20	Grand centroid
PC1	-2.58605	-0.0000000
PC2	-2.25956	0.0000000
PC3	-0.40501	-0.0000000
PC4	-1.23919	-0.0000000

Distances Between Cluster Centroids

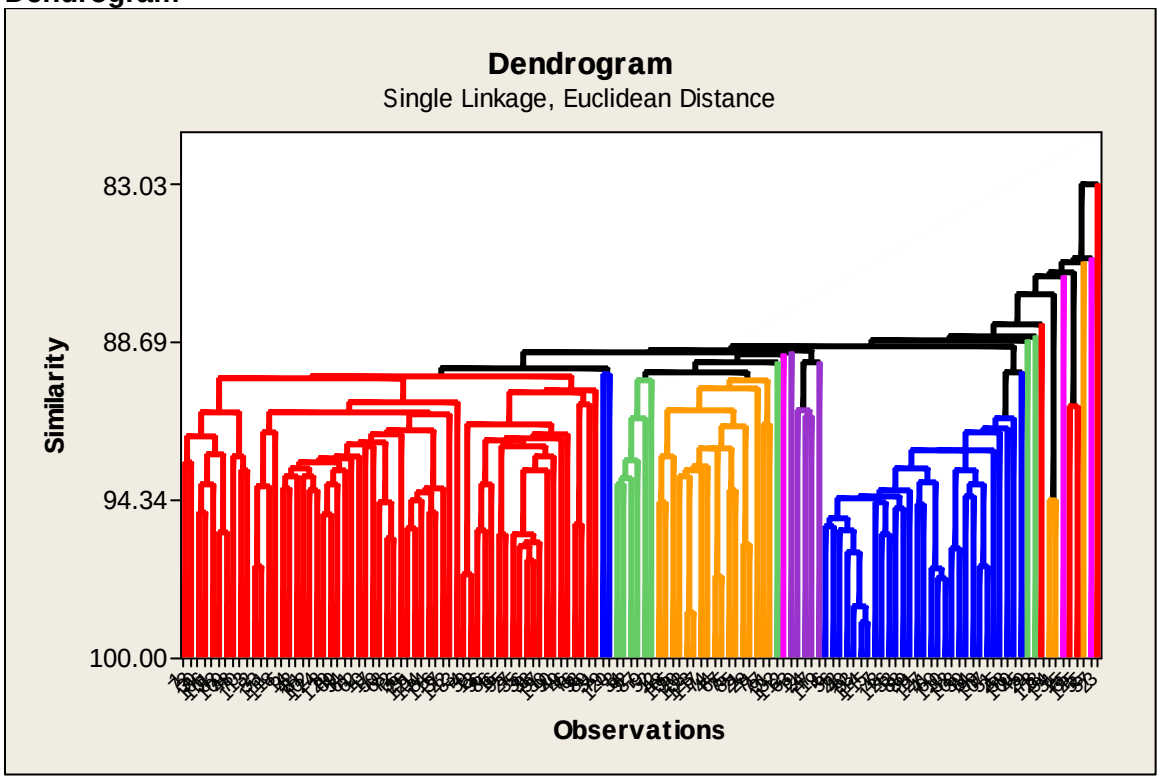
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0.00000	4.59739	5.7046	5.04533	3.75970	7.3022	9.4130
Cluster2	4.59739	0.00000	9.5383	2.43894	5.89571	3.4942	6.3932
Cluster3	5.70459	9.53828	0.0000	9.20167	5.55100	11.9934	13.8356
Cluster4	5.04533	2.43894	9.2017	0.00000	4.57868	2.9396	5.2418
Cluster5	3.75970	5.89571	5.5510	4.57868	0.00000	7.1663	8.4768
Cluster6	7.30221	3.49420	11.9934	2.93959	7.16631	0.0000	3.0645
Cluster7	9.41298	6.39323	13.8356	5.24184	8.47681	3.0645	0.0000
Cluster8	5.53614	9.02793	5.1820	8.70198	5.84610	11.2953	12.9402
Cluster9	3.30726	7.15546	3.3012	7.57451	5.24531	10.1053	12.3439
Cluster10	5.03594	3.30500	10.3428	5.60491	7.86418	5.9585	8.5696
Cluster11	4.62384	7.91776	4.9382	8.88001	7.25372	11.1229	13.5699
Cluster12	3.51353	2.53474	8.5247	3.33470	5.18393	5.1053	7.5212
Cluster13	3.15291	5.79096	3.9728	5.58951	3.44292	8.2895	10.4375
Cluster14	8.18069	4.26371	12.3473	3.56874	7.74710	1.9004	3.8077
Cluster15	2.51112	6.42863	4.0205	6.01126	2.64680	8.6455	10.3634
Cluster16	6.26165	2.53199	11.2255	4.81208	8.29400	4.8433	7.7837
Cluster17	6.49262	2.21419	11.5846	3.79972	7.74352	2.8615	5.6867
Cluster18	7.08519	3.93350	11.2452	2.29630	6.18577	1.5753	3.1064
Cluster19	4.10413	8.05034	3.3227	7.37227	3.12360	9.9610	11.3591
Cluster20	4.21225	3.95750	8.1467	2.31894	2.95018	4.5165	5.9003

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	5.5361	3.3073	5.0359	4.6238	3.51353	3.1529
Cluster2	9.0279	7.1555	3.3050	7.9178	2.53474	5.7910
Cluster3	5.1820	3.3012	10.3428	4.9382	8.52470	3.9728
Cluster4	8.7020	7.5745	5.6049	8.8800	3.33470	5.5895
Cluster5	5.8461	5.2453	7.8642	7.2537	5.18393	3.4429
Cluster6	11.2953	10.1053	5.9585	11.1229	5.10530	8.2895
Cluster7	12.9402	12.3439	8.5696	13.5699	7.52122	10.4375
Cluster8	0.0000	5.4623	9.8977	7.1301	6.87849	6.1040
Cluster9	5.4623	0.0000	7.2964	2.1955	6.29926	2.6591
Cluster10	9.8977	7.2964	0.0000	7.2341	4.10877	7.0334
Cluster11	7.1301	2.1955	7.2341	0.0000	7.34129	4.3905
Cluster12	6.8785	6.2993	4.1088	7.3413	0.00000	5.4318
Cluster13	6.1040	2.6591	7.0334	4.3905	5.43184	0.0000
Cluster14	12.2687	10.5894	6.8446	11.5485	6.28391	8.5184
Cluster15	3.8636	3.2827	7.4958	5.3469	5.04059	2.9465
Cluster16	10.5557	8.4591	2.6788	8.6523	4.00977	7.5908
Cluster17	11.1432	9.0799	3.4880	9.5792	4.42240	7.7628
Cluster18	10.6757	9.7441	6.8377	11.0152	5.19650	7.6794
Cluster19	4.3775	3.9667	9.0965	6.0308	6.85050	3.7809
Cluster20	7.2139	6.9349	6.4304	8.5822	3.48649	5.1738

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	8.1807	2.5111	6.2617	6.4926	7.0852	4.1041
Cluster2	4.2637	6.4286	2.5320	2.2142	3.9335	8.0503
Cluster3	12.3473	4.0205	11.2255	11.5846	11.2452	3.3227
Cluster4	3.5687	6.0113	4.8121	3.7997	2.2963	7.3723
Cluster5	7.7471	2.6468	8.2940	7.7435	6.1858	3.1236
Cluster6	1.9004	8.6455	4.8433	2.8615	1.5753	9.9610
Cluster7	3.8077	10.3634	7.7837	5.6867	3.1064	11.3591
Cluster8	12.2687	3.8636	10.5557	11.1432	10.6757	4.3775
Cluster9	10.5894	3.2827	8.4591	9.0799	9.7441	3.9667
Cluster10	6.8446	7.4958	2.6788	3.4880	6.8377	9.0965
Cluster11	11.5485	5.3469	8.6523	9.5792	11.0152	6.0308
Cluster12	6.2839	5.0406	4.0098	4.4224	5.1965	6.8505
Cluster13	8.5184	2.9465	7.5908	7.7628	7.6794	3.7809
Cluster14	0.0000	9.4267	5.4315	3.5422	2.1795	10.6417
Cluster15	9.4267	0.0000	8.3835	8.4728	8.0042	1.8913
Cluster16	5.4315	8.3835	0.0000	2.3392	5.8120	10.1120
Cluster17	3.5422	8.4728	2.3392	0.0000	4.0590	10.0044
Cluster18	2.1795	8.0042	5.8120	4.0590	0.0000	9.1693
Cluster19	10.6417	1.8913	10.1120	10.0044	9.1693	0.0000
Cluster20	5.5157	4.5480	6.4083	5.5140	3.6742	5.7119

```
           Cluster20
Cluster1     4.21225
Cluster2     3.95750
Cluster3     8.14672
Cluster4     2.31894
Cluster5     2.95018
Cluster6     4.51649
Cluster7     5.90034
Cluster8     7.21391
Cluster9     6.93493
Cluster10    6.43040
Cluster11    8.58220
Cluster12    3.48649
Cluster13    5.17375
Cluster14    5.51575
Cluster15    4.54804
Cluster16    6.40827
Cluster17    5.51398
Cluster18    3.67415
Cluster19    5.71186
Cluster20    0.00000
```

Dendrogram



Neste caso, mesmo considerando apenas os 4 PCs, os resultados foram muito parecidos com o anterior, com todas as variáveis, onde 84% das observações se apresentaram concentradas praticamente nos 4 primeiros clusters, portanto este modelo é igualmente insuficiente para explicar os dados.

8.2.3 Análise utilizando somente os PCs (por variáveis e por observações)

Análise com os 4 PCs com 30 clusters (por variáveis e por observações) e Dendrogramas

---

Cluster Analysis of Observations: PC1, PC2, PC3, PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	98.7552	0.19128	114	115	114	2
2	130	98.4141	0.24369	103	123	103	2
3	129	98.1862	0.27871	93	114	93	3
4	128	97.2262	0.42623	60	110	60	2
5	127	97.1299	0.44103	44	67	44	2
6	126	97.0271	0.45682	4	59	4	2
7	125	96.8351	0.48633	41	60	41	3
8	124	96.7902	0.49322	99	107	99	2
9	123	96.7335	0.50193	2	53	2	2
10	122	96.5366	0.53219	36	38	36	2
11	121	96.2815	0.57138	40	93	40	4
12	120	96.0868	0.60131	32	108	32	2
13	119	96.0476	0.60734	35	36	35	3
14	118	95.9672	0.61968	21	28	21	2
15	117	95.8917	0.63128	35	97	35	4
16	116	95.7692	0.65011	68	72	68	2
17	115	95.6201	0.67302	43	126	43	2
18	114	95.6163	0.67361	27	35	27	5
19	113	95.5936	0.67709	16	95	16	2
20	112	95.5590	0.68241	82	109	82	2
21	111	95.4974	0.69187	13	96	13	2
22	110	95.4384	0.70094	22	40	22	5
23	109	95.3932	0.70789	26	121	26	2
24	108	95.3510	0.71436	6	56	6	2
25	107	95.2380	0.73174	76	98	76	2
26	106	95.0313	0.76349	6	22	6	7
27	105	94.8966	0.78420	8	129	8	2
28	104	94.7924	0.80020	30	106	30	2
29	103	94.7830	0.80165	84	116	84	2
30	102	94.6836	0.81693	33	88	33	2
31	101	94.5041	0.84451	11	127	11	2
32	100	94.4841	0.84758	12	18	12	2
33	99	94.4555	0.85198	50	68	50	3
34	98	94.4372	0.85479	7	43	7	3
35	97	94.3747	0.86439	94	131	94	2
36	96	94.3608	0.86653	26	104	26	3
37	95	94.3421	0.86940	6	7	6	10
38	94	94.2682	0.88075	39	64	39	2
39	93	94.2146	0.88899	6	33	6	12
40	92	94.1221	0.90320	26	84	26	5
41	91	94.0396	0.91588	6	89	6	13
42	90	94.0362	0.91640	75	85	75	2
43	89	94.0224	0.91852	102	124	102	2
44	88	93.9572	0.92854	26	37	26	6
45	87	93.9294	0.93281	5	14	5	2
46	86	93.9125	0.93542	2	111	2	3
47	85	93.8288	0.94828	13	86	13	3
48	84	93.8125	0.95077	30	101	30	3
49	83	93.8012	0.95252	39	99	39	4
50	82	93.7902	0.95421	8	69	8	3
51	81	93.7757	0.95643	3	34	3	2
52	80	93.7153	0.96572	11	41	11	5
53	79	93.6151	0.98112	3	92	3	3
54	78	93.5526	0.99071	81	102	81	3
55	77	93.5380	0.99296	5	42	5	3
56	76	93.5034	0.99828	20	103	20	3
57	75	93.3348	1.02418	32	39	32	6
58	74	93.3290	1.02508	70	113	70	2
59	73	93.3196	1.02652	20	117	20	4
60	72	93.2910	1.03092	8	54	8	4
61	71	93.2750	1.03337	27	130	27	6

62	70	93.2267	1.04080	5	81	5	6
63	69	93.2009	1.04476	6	11	6	18
64	68	93.1637	1.05047	20	74	20	5
65	67	93.0182	1.07284	1	73	1	2
66	66	93.0037	1.07506	5	8	5	10
67	65	92.9176	1.08829	3	87	3	4
68	64	92.8834	1.09355	5	61	5	11
69	63	92.8260	1.10236	12	105	12	3
70	62	92.8178	1.10363	5	62	5	12
71	61	92.8125	1.10444	27	90	27	7
72	60	92.7887	1.10810	48	70	48	3
73	59	92.7442	1.11494	30	82	30	5
74	58	92.6035	1.13655	6	32	6	24
75	57	92.6033	1.13659	6	51	6	25
76	56	92.5251	1.14860	20	44	20	7
77	55	92.5114	1.15070	20	75	20	9
78	54	92.4756	1.15620	10	17	10	2
79	53	92.4256	1.16390	16	27	16	9
80	52	92.3502	1.17547	5	10	5	14
81	51	92.3228	1.17968	5	50	5	17
82	50	92.2246	1.19478	13	16	13	12
83	49	92.1583	1.20496	13	49	13	13
84	48	92.0808	1.21687	1	30	1	7
85	47	92.0455	1.22230	5	66	5	18
86	46	92.0184	1.22646	13	46	13	14
87	45	91.9568	1.23593	2	118	2	4
88	44	91.9346	1.23934	6	25	6	26
89	43	91.8775	1.24811	5	26	5	24
90	42	91.7722	1.26429	6	58	6	27
91	41	91.7204	1.27226	20	21	20	11
92	40	91.6648	1.28080	4	13	4	16
93	39	91.6170	1.28814	77	91	77	2
94	38	91.4558	1.31291	3	9	3	5
95	37	91.4309	1.31674	6	100	6	28
96	36	91.3382	1.33099	24	47	24	2
97	35	91.2670	1.34192	5	122	5	25
98	34	91.2150	1.34991	1	48	1	10
99	33	91.2103	1.35064	2	5	2	29
100	32	91.1761	1.35589	12	20	12	14

Final Partition
Number of clusters: 30

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	10	22.959	1.41578	2.58997
Cluster2	29	158.006	2.17945	4.05434
Cluster3	5	7.763	1.17763	1.70476
Cluster4	16	31.694	1.30496	2.34332
Cluster5	28	81.312	1.60967	2.81252
Cluster6	14	28.034	1.35581	1.85350
Cluster7	1	0.000	0.00000	0.00000
Cluster8	3	2.319	0.86690	0.97743
Cluster9	1	0.000	0.00000	0.00000
Cluster10	1	0.000	0.00000	0.00000
Cluster11	1	0.000	0.00000	0.00000
Cluster12	1	0.000	0.00000	0.00000
Cluster13	1	0.000	0.00000	0.00000
Cluster14	1	0.000	0.00000	0.00000
Cluster15	1	0.000	0.00000	0.00000
Cluster16	1	0.000	0.00000	0.00000
Cluster17	2	0.950	0.68905	0.68905
Cluster18	1	0.000	0.00000	0.00000
Cluster19	2	0.268	0.36587	0.36587
Cluster20	2	0.830	0.64407	0.64407
Cluster21	1	0.000	0.00000	0.00000
Cluster22	1	0.000	0.00000	0.00000

Cluster23	1	0.000	0.00000	0.00000
Cluster24	1	0.000	0.00000	0.00000
Cluster25	2	0.374	0.43220	0.43220
Cluster26	1	0.000	0.00000	0.00000
Cluster27	1	0.000	0.00000	0.00000
Cluster28	1	0.000	0.00000	0.00000
Cluster29	1	0.000	0.00000	0.00000
Cluster30	1	0.000	0.00000	0.00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	-0.17771	-0.36273	-4.41708	0.43718	5.16890	-3.71355	0.22869
PC2	0.45663	2.24372	0.14422	-1.49870	-0.57848	-1.65660	-2.26910
PC3	0.01435	-0.09745	-0.05626	2.57561	-0.79642	-0.78641	-1.28844
PC4	-1.44337	-0.24755	0.96224	-0.68371	0.56372	0.20966	-1.25497

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
PC1	-6.77750	-8.13369	3.38342	0.53971	2.80635	-3.49060
PC2	-1.20122	-2.67723	-2.46450	1.20791	1.51104	1.81639
PC3	-1.19452	-2.01699	3.56668	1.80326	0.12352	0.67273
PC4	-0.19663	-2.36384	-0.47128	-0.73527	0.88649	1.74014

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
PC1	-4.12551	2.98336	-2.89933	1.31972	-4.88770	0.96182
PC2	3.46120	3.55180	-0.39931	-0.06594	-2.03148	0.24482
PC3	1.26895	0.39320	1.94931	-1.39402	-0.53012	0.75147
PC4	0.61847	1.62904	0.50840	1.15299	1.40659	1.09514

Variable	Cluster20	Cluster21	Cluster22	Cluster23	Cluster24	Cluster25
PC1	-5.25606	-7.00312	2.00160	0.86966	-0.09497	-5.45267
PC2	-1.72348	-1.30180	-0.73249	0.66688	1.37770	1.92967
PC3	-2.18845	-2.70399	0.30798	1.94412	-1.09186	1.11733
PC4	1.69054	0.93113	-0.26753	0.58363	-2.19451	2.36377

Variable	Cluster26	Cluster27	Cluster28	Cluster29	Cluster30	Grand centroid
PC1	-6.25648	-5.86159	1.34261	3.01773	-2.58605	-0.0000000
PC2	1.25796	-2.34410	-1.20103	-1.32301	-2.25956	0.0000000
PC3	-0.37244	-1.76740	0.64046	-0.41917	-0.40501	-0.0000000
PC4	0.89572	-0.28686	-1.55233	-1.82718	-1.23919	-0.0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0.00000	2.16111	4.88486	3.36727	5.8603	4.51017	3.05409
Cluster2	2.16111	0.00000	4.72345	4.68839	6.3016	5.20808	4.81130
Cluster3	4.88486	4.72345	0.00000	5.99157	9.6499	2.19941	5.81735
Cluster4	3.36727	4.68839	5.99157	0.00000	6.0135	5.41800	3.98675
Cluster5	5.86033	6.30163	9.64988	6.01354	0.0000	8.95465	5.55100
Cluster6	4.51017	5.20808	2.19941	5.41800	8.9546	0.00000	4.27943
Cluster7	3.05409	4.81130	5.81735	3.98675	5.5510	4.27943	0.00000
Cluster8	7.02293	7.36363	3.16550	8.16034	11.9934	3.15069	7.16631
Cluster9	8.83698	9.63155	6.05648	9.93798	13.8356	5.35878	8.47681
Cluster10	5.89734	7.04818	9.10131	3.26197	5.1820	8.39242	5.84610
Cluster11	2.18649	2.39538	5.66044	2.81700	5.7503	5.82193	4.69202
Cluster12	3.93148	3.45178	7.35420	4.81142	3.3012	7.33683	5.24531
Cluster13	4.83657	3.80923	2.18885	5.99277	9.1796	4.07220	6.58340
Cluster14	5.51705	4.27292	3.60024	6.98726	10.3428	5.54556	7.86418
Cluster15	5.39960	4.08286	8.18687	6.48863	4.9382	8.68228	7.25372
Cluster16	3.96148	4.26387	2.61291	3.76222	8.5247	3.13324	5.18393
Cluster17	3.35260	3.43624	5.89755	4.68648	3.9728	5.39658	3.44292
Cluster18	6.06576	6.45570	2.31887	6.53093	10.1994	1.73710	5.82176
Cluster19	2.88630	2.87637	5.44177	3.13160	4.5887	5.35011	4.06713
Cluster20	6.72431	6.91461	3.04448	7.79723	10.6395	2.55784	6.31397
Cluster21	7.91871	8.05288	3.97367	9.26708	12.3473	3.89164	7.74710
Cluster22	2.76265	3.82265	6.60402	2.88960	3.4592	5.91120	3.00466
Cluster23	2.99565	2.97727	5.68927	2.62330	5.2484	5.83095	4.78129
Cluster24	1.62577	2.36668	5.58924	4.92846	6.2634	5.30795	3.78487

Cluster25	6.76061	5.85668	2.75714	7.60643	11.2255	4.91429	8.29400
Cluster26	6.57378	6.09021	2.17444	7.97432	11.5846	3.95009	7.74352
Cluster27	6.68302	7.35363	3.57304	7.70769	11.2452	2.50911	6.18577
Cluster28	2.33732	4.12570	6.46469	2.32545	4.6444	5.55993	2.48810
Cluster29	3.70312	5.17182	8.08342	4.11902	3.3227	7.05017	3.12360
Cluster30	3.65996	5.12841	3.75483	4.34873	8.1467	1.96963	2.95018

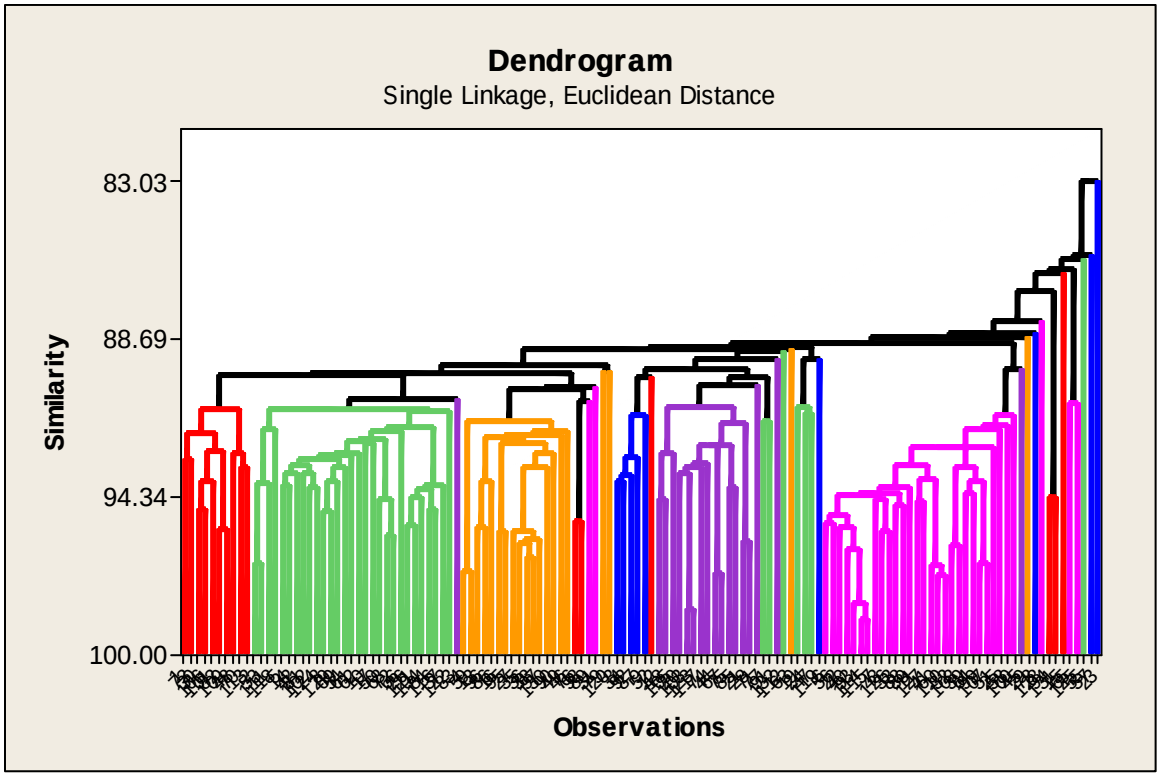
	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	7.0229	8.8370	5.8973	2.1865	3.9315	4.83657
Cluster2	7.3636	9.6315	7.0482	2.3954	3.4518	3.80923
Cluster3	3.1655	6.0565	9.1013	5.6604	7.3542	2.18885
Cluster4	8.1603	9.9380	3.2620	2.8170	4.8114	5.99277
Cluster5	11.9934	13.8356	5.1820	5.7503	3.3012	9.17959
Cluster6	3.1507	5.3588	8.3924	5.8219	7.3368	4.07220
Cluster7	7.1663	8.4768	5.8461	4.6920	5.2453	6.58340
Cluster8	0.0000	3.0645	11.2953	8.2839	10.1053	5.21031
Cluster9	3.0645	0.0000	12.9402	10.3715	12.3439	8.11345
Cluster10	11.2953	12.9402	0.0000	4.9752	5.4623	8.87939
Cluster11	8.2839	10.3715	4.9752	0.0000	3.2682	4.90096
Cluster12	10.1053	12.3439	5.4623	3.2682	0.0000	6.38554
Cluster13	5.2103	8.1135	8.8794	4.9010	6.3855	0.00000
Cluster14	5.9585	8.5696	9.8977	5.3814	7.2964	2.17305
Cluster15	11.1229	13.5699	7.1301	4.3639	2.1955	6.70927
Cluster16	5.1053	7.5212	6.8785	3.9973	6.2993	2.89927
Cluster17	8.2895	10.4375	6.1040	4.0024	2.6591	5.59451
Cluster18	2.6967	5.2326	9.4292	7.0698	8.5115	4.27973
Cluster19	8.2124	10.5307	4.8563	2.3585	2.3331	4.76614
Cluster20	2.6715	5.0654	10.6294	8.0001	9.0253	4.88225
Cluster21	1.9004	3.8077	12.2687	9.2890	10.5894	5.84103
Cluster22	8.9194	10.7845	3.9459	2.8908	2.6546	6.38943
Cluster23	8.5105	10.7991	4.4576	1.4700	2.8053	4.82569
Cluster24	7.4370	9.0525	7.1787	3.3080	4.4051	5.50618
Cluster25	4.8433	7.7837	10.5557	6.8194	8.4591	2.10929
Cluster26	2.8615	5.6867	11.1432	7.3201	9.0799	3.12528
Cluster27	1.5753	3.1064	10.6757	8.1575	9.7441	5.74406
Cluster28	8.4345	9.9850	3.9361	2.9099	3.9640	6.58074
Cluster29	9.9610	11.3591	4.3775	4.3218	3.9667	8.13217
Cluster30	4.5165	5.9003	7.2139	5.1888	6.9349	5.24114

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	5.5170	5.3996	3.96148	3.3526	6.0658	2.8863
Cluster2	4.2729	4.0829	4.26387	3.4362	6.4557	2.8764
Cluster3	3.6002	8.1869	2.61291	5.8976	2.3189	5.4418
Cluster4	6.9873	6.4886	3.76222	4.6865	6.5309	3.1316
Cluster5	10.3428	4.9382	8.52470	3.9728	10.1994	4.5887
Cluster6	5.5456	8.6823	3.13324	5.3966	1.7371	5.3501
Cluster7	7.8642	7.2537	5.18393	3.4429	5.8218	4.0671
Cluster8	5.9585	11.1229	5.10530	8.2895	2.6967	8.2124
Cluster9	8.5696	13.5699	7.52122	10.4375	5.2326	10.5307
Cluster10	9.8977	7.1301	6.87849	6.1040	9.4292	4.8563
Cluster11	5.3814	4.3639	3.99728	4.0024	7.0698	2.3585
Cluster12	7.2964	2.1955	6.29926	2.6591	8.5115	2.3331
Cluster13	2.1730	6.7093	2.89927	5.5945	4.2797	4.7661
Cluster14	0.0000	7.2341	4.10877	7.0334	5.8829	6.0598
Cluster15	7.2341	0.0000	7.34129	4.3905	9.6968	3.9289
Cluster16	4.1088	7.3413	0.00000	5.4318	3.6840	4.1355
Cluster17	7.0334	4.3905	5.43184	0.0000	6.5731	2.1980
Cluster18	5.8829	9.6968	3.68401	6.5731	0.0000	6.4139
Cluster19	6.0598	3.9289	4.13551	2.1980	6.4139	0.0000
Cluster20	6.4235	10.1186	5.08194	6.8490	1.7496	7.1787
Cluster21	6.8446	11.5485	6.28391	8.5184	3.1558	8.8204
Cluster22	7.5390	4.7878	5.23700	2.4133	7.2564	2.0223
Cluster23	5.7634	4.0359	3.91762	3.4938	6.8722	1.3677
Cluster24	5.8372	5.5703	5.25134	3.9221	6.9193	4.0767
Cluster25	2.6788	8.6523	4.00977	7.5908	4.4317	6.7622
Cluster26	3.4880	9.5792	4.42240	7.7628	3.6028	7.3779
Cluster27	6.8377	11.0152	5.19650	7.6794	2.3334	7.8432
Cluster28	7.5329	5.9551	4.95949	3.5703	7.0450	3.0425
Cluster29	9.0965	6.0308	6.85050	3.7809	8.5713	4.0737
Cluster30	6.4304	8.5822	3.48649	5.1738	3.5165	5.0642

	Cluster20	Cluster21	Cluster22	Cluster23	Cluster24	Cluster25
Cluster1	6.7243	7.9187	2.7626	2.9957	1.62577	6.7606
Cluster2	6.9146	8.0529	3.8227	2.9773	2.36668	5.8567
Cluster3	3.0445	3.9737	6.6040	5.6893	5.58924	2.7571
Cluster4	7.7972	9.2671	2.8896	2.6233	4.92846	7.6064
Cluster5	10.6395	12.3473	3.4592	5.2484	6.26339	11.2255
Cluster6	2.5578	3.8916	5.9112	5.8310	5.30795	4.9143
Cluster7	6.3140	7.7471	3.0047	4.7813	3.78487	8.2940
Cluster8	2.6715	1.9004	8.9194	8.5105	7.43701	4.8433
Cluster9	5.0654	3.8077	10.7845	10.7991	9.05250	7.7837
Cluster10	10.6294	12.2687	3.9459	4.4576	7.17866	10.5557
Cluster11	8.0001	9.2890	2.8908	1.4700	3.30798	6.8194
Cluster12	9.0253	10.5894	2.6546	2.8053	4.40513	8.4591
Cluster13	4.8822	5.8410	6.3894	4.8257	5.50618	2.1093
Cluster14	6.4235	6.8446	7.5390	5.7634	5.83717	2.6788
Cluster15	10.1186	11.5485	4.7878	4.0359	5.57026	8.6523
Cluster16	5.0819	6.2839	5.2370	3.9176	5.25134	4.0098
Cluster17	6.8490	8.5184	2.4133	3.4938	3.92205	7.5908
Cluster18	1.7496	3.1558	7.2564	6.8722	6.91926	4.4317
Cluster19	7.1787	8.8204	2.0223	1.3677	4.07675	6.7622
Cluster20	0.0000	2.0181	7.9826	7.8449	7.24916	4.9765
Cluster21	2.0181	0.0000	9.5874	9.3585	8.20188	5.4315
Cluster22	7.9826	9.5874	0.0000	2.5770	3.81069	8.3805
Cluster23	7.8449	9.3585	2.5770	0.0000	4.28614	6.7394
Cluster24	7.2492	8.2019	3.8107	4.2861	0.00000	7.3938
Cluster25	4.9765	5.4315	8.3805	6.7394	7.39378	0.0000
Cluster26	3.7175	3.5422	8.6008	7.5230	6.93149	2.3392
Cluster27	2.1998	2.1795	8.2906	8.3011	7.15550	5.8120
Cluster28	7.8952	9.3281	1.5540	3.1583	3.48278	8.4581
Cluster29	9.1717	10.6417	2.0839	4.4689	4.19167	10.1120
Cluster30	4.3795	5.5157	4.9831	5.4173	4.56284	6.4083

	Cluster26	Cluster27	Cluster28	Cluster29	Cluster30
Cluster1	6.5738	6.6830	2.33732	3.7031	3.65996
Cluster2	6.0902	7.3536	4.12570	5.1718	5.12841
Cluster3	2.1744	3.5730	6.46469	8.0834	3.75483
Cluster4	7.9743	7.7077	2.32545	4.1190	4.34873
Cluster5	11.5846	11.2452	4.64439	3.3227	8.14672
Cluster6	3.9501	2.5091	5.55993	7.0502	1.96963
Cluster7	7.7435	6.1858	2.48810	3.1236	2.95018
Cluster8	2.8615	1.5753	8.43453	9.9610	4.51649
Cluster9	5.6867	3.1064	9.98498	11.3591	5.90034
Cluster10	11.1432	10.6757	3.93608	4.3775	7.21391
Cluster11	7.3201	8.1575	2.90987	4.3218	5.18885
Cluster12	9.0799	9.7441	3.96395	3.9667	6.93493
Cluster13	3.1253	5.7441	6.58074	8.1322	5.24114
Cluster14	3.4880	6.8377	7.53287	9.0965	6.43040
Cluster15	9.5792	11.0152	5.95514	6.0308	8.58220
Cluster16	4.4224	5.1965	4.95949	6.8505	3.48649
Cluster17	7.7628	7.6794	3.57027	3.7809	5.17375
Cluster18	3.6028	2.3334	7.04499	8.5713	3.51645
Cluster19	7.3779	7.8432	3.04252	4.0737	5.06418
Cluster20	3.7175	2.1998	7.89521	9.1717	4.37954
Cluster21	3.5422	2.1795	9.32815	10.6417	5.51575
Cluster22	8.6008	8.2906	1.55404	2.0839	4.98307
Cluster23	7.5230	8.3011	3.15826	4.4689	5.41729
Cluster24	6.9315	7.1555	3.48278	4.1917	4.56284
Cluster25	2.3392	5.8120	8.45815	10.1120	6.40827
Cluster26	0.0000	4.0590	8.41497	10.0044	5.51398
Cluster27	4.0590	0.0000	7.78501	9.1693	3.67415
Cluster28	8.4150	7.7850	0.00000	2.0048	4.21259
Cluster29	10.0044	9.1693	2.00480	0.0000	5.71186
Cluster30	5.5140	3.6742	4.21259	5.7119	0.00000

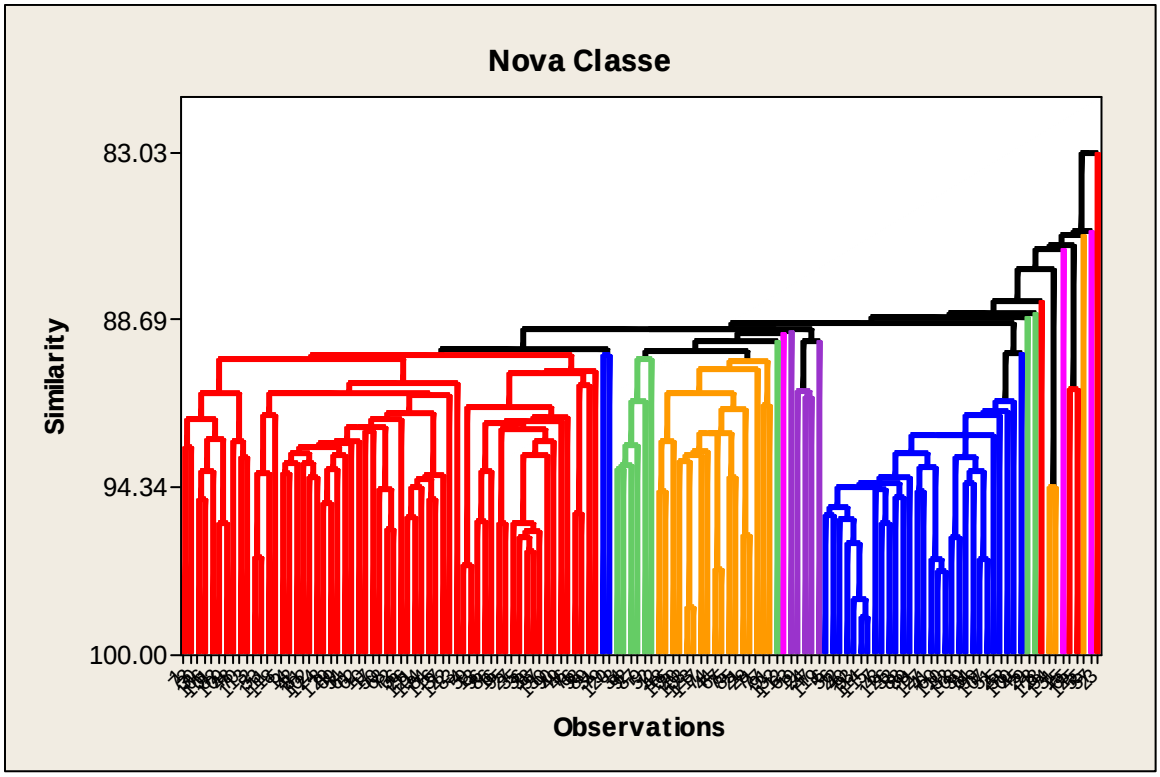
Dendrogram



O teste com 30 clusters também não se mostrou eficiente, 72% das observações ficaram concentradas em apenas 6 clusters.

8.2.4 Categorização (all regions)

Nesta fase será criada uma nova coluna na base de dados para comportar os membros do cluster.



O gráfico acima reflete um bom grau de similaridade entre as observações. Apesar de destacar 3 grandes blocos (vermelho, laranja e azul), vemos que as observações em média estão acima de 90% de similaridade.

8.2.5 Categorização (regions: AIBER e AVECO)

A partir daqui, os dados estão ordenados por bloco e separados para análise. Somente as regiões AIBER e AVECO constarão das análises. A região OTHERS, foi eliminada.

Realizado testes de similaridade com AIBER e AVECO, utilizando 1, 10, 12, 20 e 30 clusters, dos quais o de 12 clusters, aparentemente, apresentou o melhor grau de similaridade e será utilizado nas análises das próximas sessões.

8.2.6 Análise AIBER com 12 clusters (por variáveis e por observações)

Análise AIBER com 12 clusters (por variáveis e por observações) e Dendrogramas

---

Cluster Analysis of Variables: EPI-2014_n_1, Governança_n, Biodiversity, ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

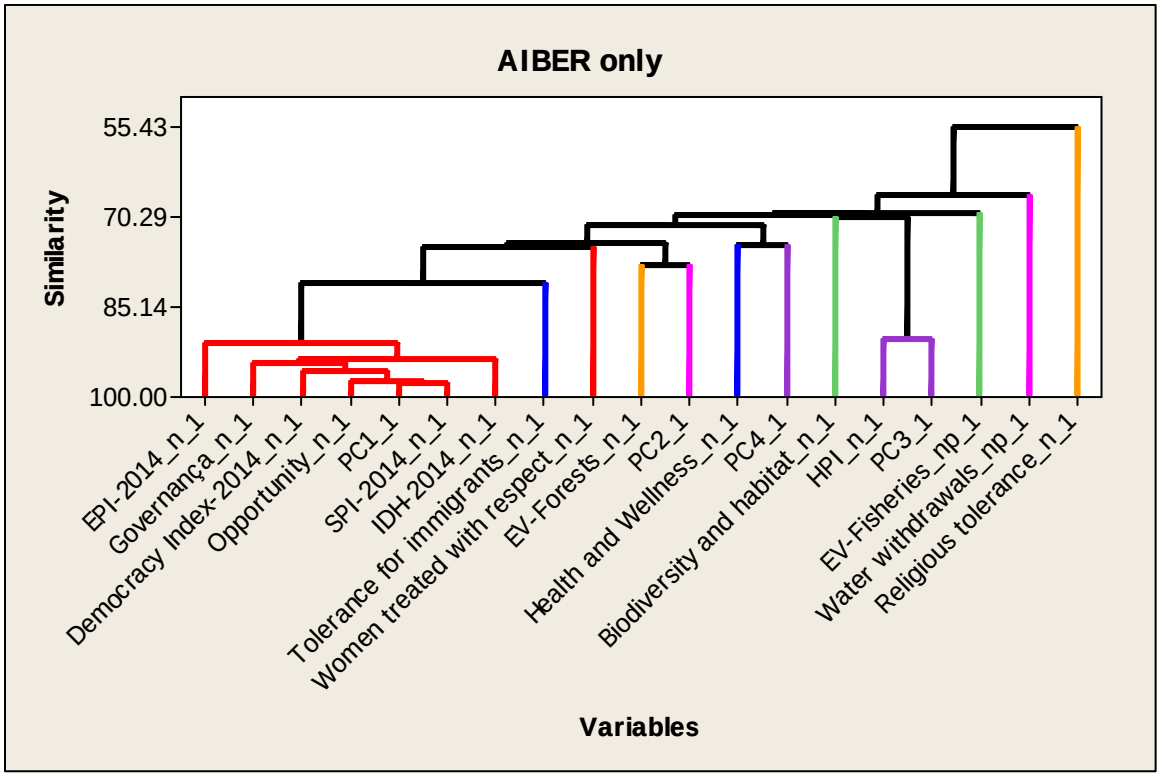
Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	18	98.0980	0.038040	15	19	15	2
2	17	97.7194	0.045613	5	15	5	3

3	16	95.9319	0.081361	4	5	4	4
4	15	94.7694	0.104612	2	4	2	5
5	14	93.8810	0.122379	2	6	2	6
6	13	91.5211	0.169577	1	2	1	7
7	12	90.6321	0.187359	10	17	10	2
8	11	81.3907	0.372187	1	7	1	8
9	10	78.3403	0.433194	8	16	8	2
10	9	75.2882	0.494235	1	11	1	9
11	8	75.0539	0.498921	13	18	13	2
12	7	74.8844	0.502312	1	8	1	11
13	6	71.5895	0.568210	1	13	1	13
14	5	70.3565	0.592871	3	10	3	3
15	4	70.0104	0.599791	1	3	1	16
16	3	69.7355	0.605290	1	12	1	17
17	2	66.6408	0.667184	1	9	1	18
18	1	55.4307	0.891386	1	14	1	19

Final Partition

Cluster 1
EPI-2014_n_1 Governança_n_1 Democracy Index-2014_n_1 Opportunity_n_1 IDH-2014_n_1 PC1_1 SPI-2014_n_1
Cluster 2
Biodiversity and habitat_n_1
Cluster 3
Tolerance for immigrants_n_1
Cluster 4
EV-Forests_n_1
Cluster 5
Water withdrawals_np_1
Cluster 6
HPI_n_1 PC3_1
Cluster 7
Women treated with respect_n_1
Cluster 8
EV-Fisheries_np_1
Cluster 9
Health and Wellness_n_1
Cluster 10
Religious tolerance_n_1
Cluster 11
PC2_1
Cluster 12
PC4_1

Dendrogram



No gráfico e tabela acima, notamos que mais de 50% das observações estão concentradas nos primeiros 4 clusters. O cluster 1 é a partição que abriga a maior quantidade de variáveis nesta análise, são elas: EPI-2014, Governança, Democracy Index-2014, Opportunity, IDH-2014, PC1 e SPI-2014.

Cluster Analysis of Observations: EPI-2014_n_1, Governança_n, Biodiversity, ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	79.3730	31.3986	12	14	12	2
2	19	76.6704	35.5125	11	12	11	3
3	18	71.6977	43.0820	2	11	2	4
4	17	70.3577	45.1218	2	9	2	5
5	16	70.0311	45.6189	2	17	2	6
6	15	70.0169	45.6405	2	8	2	7
7	14	68.0779	48.5921	2	10	2	8
8	13	67.4285	49.5807	1	3	1	2
9	12	65.7116	52.1941	1	5	1	3
10	11	65.2982	52.8234	2	15	2	9
11	10	63.8306	55.0574	1	2	1	12
12	9	62.0624	57.7490	1	6	1	13
13	8	61.3253	58.8710	18	19	18	2
14	7	58.7263	62.8272	1	16	1	14
15	6	58.4844	63.1955	1	21	1	15
16	5	50.1223	75.9243	1	13	1	16
17	4	46.6872	81.1532	4	18	4	3
18	3	46.2333	81.8442	4	20	4	4
19	2	45.1611	83.4763	1	4	1	20
20	1	36.8882	96.0693	1	7	1	21

Final Partition
Number of clusters: 12

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	3	4165.0	35.2947	44.2407
Cluster2	8	13691.7	40.7214	52.5045
Cluster3	1	0.0	0.0000	0.0000
Cluster4	1	0.0	0.0000	0.0000
Cluster5	1	0.0	0.0000	0.0000
Cluster6	1	0.0	0.0000	0.0000
Cluster7	1	0.0	0.0000	0.0000
Cluster8	1	0.0	0.0000	0.0000
Cluster9	1	0.0	0.0000	0.0000
Cluster10	1	0.0	0.0000	0.0000
Cluster11	1	0.0	0.0000	0.0000
Cluster12	1	0.0	0.0000	0.0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
EPI-2014_n_1	76.7513	60.816	75.719	80.123
Governança_n_1	39.5052	32.756	80.674	64.569
Biodiversity and habitat_n_1	64.0067	72.874	60.160	89.550
Democracy Index-2014_n_1	64.3760	53.984	74.763	77.488
Opportunity_n_1	57.5249	45.962	74.961	71.790
IDH-2014_n_1	70.8654	54.296	81.247	70.093
Tolerance for immigrants_n_1	68.0365	54.623	65.753	78.082
EV-Forests_n_1	12.3500	11.788	100.000	19.870
Water withdrawals_np_1	74.9674	67.975	35.893	61.173
HPI_n_1	79.5956	74.538	75.503	100.000
Women treated with respect_n_1	8.1301	16.616	24.390	24.390
EV-Fisheries_np_1	61.1191	51.514	62.598	39.917
Health and Wellness_n_1	76.0376	82.741	74.437	95.350
Religious tolerance_n_1	55.5556	100.000	66.667	100.000
PC1_1	1.3179	-0.078	3.546	3.383
PC2_1	-1.2851	-1.598	-0.369	-2.465
PC3_1	2.8997	2.886	1.004	3.567
PC4_1	-0.7688	-0.572	-0.848	-0.471
SPI-2014_n_1	65.8998	55.996	78.541	81.147

Variable	Cluster5	Cluster6	Cluster7	Cluster8
EPI-2014_n_1	78.275	68.1038	76.334	62.0754
Governança_n_1	28.758	41.7995	47.875	27.3252
Biodiversity and habitat_n_1	56.060	62.3200	86.960	31.2000
Democracy Index-2014_n_1	24.052	61.4929	66.232	56.5166
Opportunity_n_1	29.325	52.3705	59.922	53.1397
IDH-2014_n_1	70.641	68.4919	72.432	55.5570
Tolerance for immigrants_n_1	41.096	30.1370	45.205	91.7808
EV-Forests_n_1	100.000	19.8700	14.350	0.0000
Water withdrawals_np_1	36.156	29.6789	91.560	99.7770
HPI_n_1	81.060	73.1163	84.950	56.0612
Women treated with respect_n_1	54.354	41.4634	36.585	18.2927
EV-Fisheries_np_1	47.589	59.4105	55.550	71.0235
Health and Wellness_n_1	71.737	75.4575	91.059	85.2985
Religious tolerance_n_1	100.000	33.3333	100.000	66.6667
PC1_1	0.540	0.8697	2.056	-0.1236
PC2_1	1.208	0.6669	-1.852	-1.9864
PC3_1	1.803	1.9441	2.611	2.0242
PC4_1	-0.735	0.5836	-0.600	-0.8218
SPI-2014_n_1	51.168	60.7656	71.855	54.0079

Variable	Cluster9	Cluster10	Cluster11	Cluster12
EPI-2014_n_1	96.174	96.6965	68.832	73.0310
Governança_n_1	74.026	70.1330	68.430	8.0062

Biodiversity and habitat_n_1	64.360	56.5600	24.620	92.9200
Democracy Index-2014_n_1	74.645	77.7251	79.147	42.4171
Opportunity_n_1	78.681	79.8744	78.885	41.6797
IDH-2014_n_1	80.897	88.6249	74.629	69.5067
Tolerance for immigrants_n_1	75.342	83.5616	84.932	64.3836
EV-Forests_n_1	7.750	45.0500	100.000	33.7600
Water withdrawals_np_1	33.234	25.4877	82.736	54.0016
HPI_n_1	38.814	51.8078	40.367	82.7132
Women treated with respect_n_1	50.000	63.4146	36.585	51.2195
EV-Fisheries_np_1	51.834	62.7743	67.660	70.3668
Health and Wellness_n_1	78.008	92.7093	66.157	73.5374
Religious tolerance_n_1	100.000	66.6667	100.000	66.6667
PC1_1	4.094	4.7767	3.018	0.5190
PC2_1	-1.403	-0.2670	-1.323	-0.5172
PC3_1	0.222	-0.0469	-0.419	2.0777
PC4_1	-0.342	1.1515	-1.827	-0.1809
SPI-2014_n_1	86.071	86.5744	80.715	56.0388

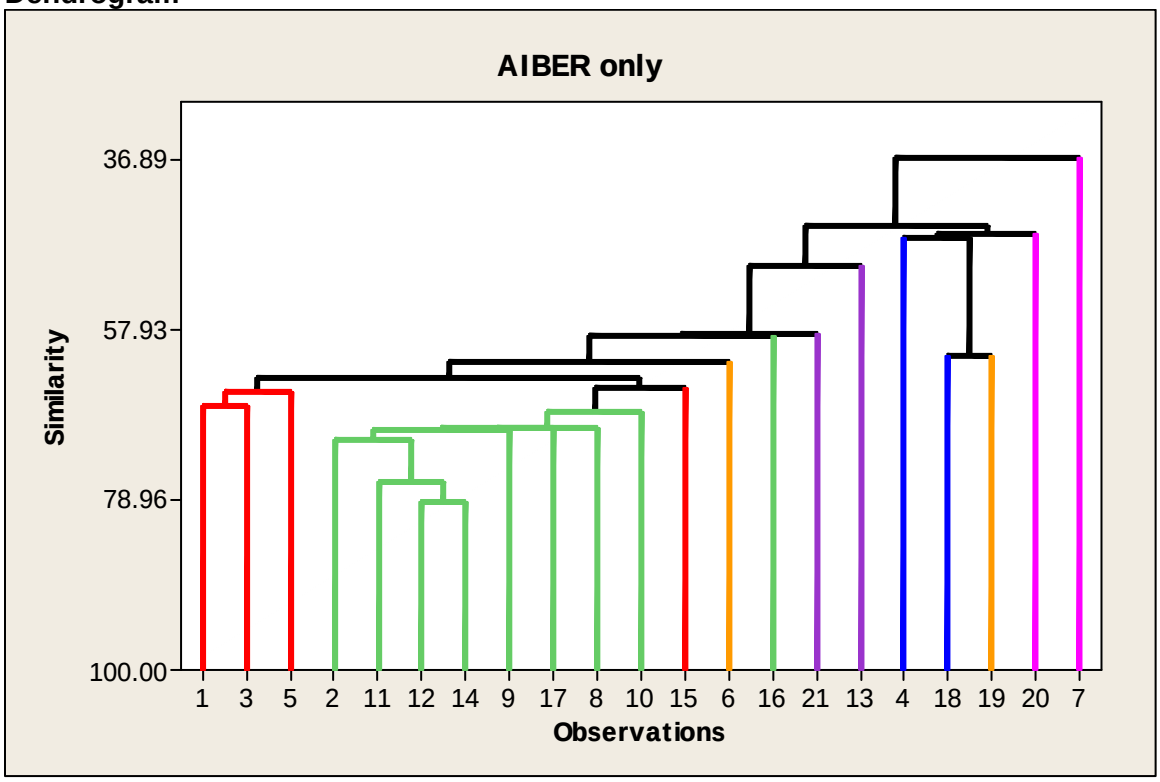
Variable	Grand centroid
EPI-2014_n_1	71.0544
Governança_n_1	42.4840
Biodiversity and habitat_n_1	66.6533
Democracy Index-2014_n_1	59.9752
Opportunity_n_1	55.2807
IDH-2014_n_1	65.6706
Tolerance for immigrants_n_1	61.9700
EV-Forests_n_1	27.2381
Water withdrawals_np_1	62.7811
HPI_n_1	72.3563
Women treated with respect_n_1	26.5720
EV-Fisheries_np_1	56.3900
Health and Wellness_n_1	80.6566
Religious tolerance_n_1	84.1270
PC1_1	1.2384
PC2_1	-1.1879
PC3_1	2.2177
PC4_1	-0.5225
SPI-2014_n_1	64.4072

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0.000	58.598	109.652	75.717	130.605	73.148	67.202
Cluster2	58.598	0.000	124.517	73.769	112.500	88.809	53.311
Cluster3	109.652	124.517	0.000	105.722	105.486	109.076	121.470
Cluster4	75.717	73.769	105.722	0.000	135.194	109.545	57.749
Cluster5	130.605	112.500	105.486	135.194	0.000	117.280	125.191
Cluster6	73.148	88.809	109.076	109.545	117.280	0.000	99.462
Cluster7	67.202	53.311	121.470	57.749	125.191	99.462	0.000
Cluster8	64.780	78.776	146.020	114.404	152.105	112.879	97.361
Cluster9	99.712	101.123	111.373	83.476	142.590	107.096	95.610
Cluster10	104.060	120.863	81.153	96.323	129.412	98.118	112.369
Cluster11	125.962	132.095	81.844	129.270	124.690	146.748	129.357
Cluster12	75.296	71.682	121.328	100.037	96.069	77.368	82.530

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12
Cluster1	64.780	99.712	104.060	125.962	75.296
Cluster2	78.776	101.123	120.863	132.095	71.682
Cluster3	146.020	111.373	81.153	81.844	121.328
Cluster4	114.404	83.476	96.323	129.270	100.037
Cluster5	152.105	142.590	129.412	124.690	96.069
Cluster6	112.879	107.096	98.118	146.748	77.368
Cluster7	97.361	95.610	112.369	129.357	82.530
Cluster8	0.000	121.283	129.696	128.277	103.699
Cluster9	121.283	0.000	58.871	118.642	117.462
Cluster10	129.696	58.871	0.000	105.050	112.281
Cluster11	128.277	118.642	105.050	0.000	143.296
Cluster12	103.699	117.462	112.281	143.296	0.000

Dendrogram



A análise por observações corrobora os resultados da análise por variáveis, mostrando que cerca de 52% das observações pertencem aos clusters 1 e 2.

8.2.7 Análise AVECO com 12 clusters (por variáveis e por observações)

Análise AVECO com 12 clusters (por variáveis e por observações) e Dendrogramas

---

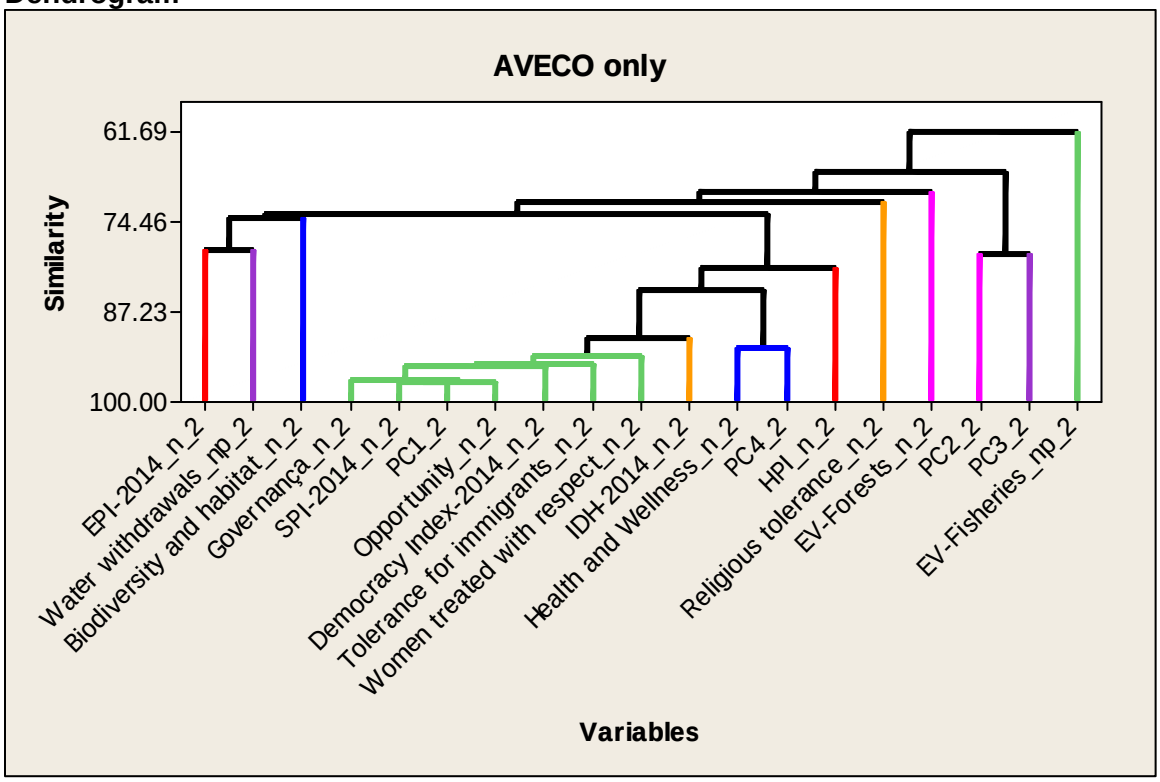
Cluster Analysis of Variables: EPI-2014_n_2, Governança_n, Biodiversity, ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	18	97.3647	0.052706	5	16	5	2
2	17	97.3147	0.053706	5	6	5	3
3	16	97.1696	0.056608	2	5	2	4
4	15	95.2458	0.095085	2	4	2	5
5	14	94.8526	0.102949	2	8	2	6
6	13	93.6875	0.126250	2	12	2	7
7	12	92.6547	0.146906	14	19	14	2
8	11	91.2472	0.175056	2	7	2	8
9	10	84.2351	0.315297	2	14	2	10
10	9	81.2592	0.374815	2	11	2	11
11	8	79.2532	0.414935	17	18	17	2
12	7	78.6780	0.426440	1	10	1	2
13	6	73.8758	0.522483	1	3	1	3
14	5	73.5142	0.529715	1	2	1	14
15	4	71.7020	0.565960	1	15	1	15
16	3	70.1845	0.596310	1	9	1	16
17	2	67.4238	0.651524	1	17	1	18

```
  18          1     61.6878  0.766243   1    13        1       19

Final Partition

Cluster 1
  EPI-2014_n_2
Cluster 2
  Governança_n_2  Democracy Index-2014_n_2  SPI-2014_n_2  Opportunity_n_2
     Tolerance for immigrants_n_2  Women treated with respect_n_2  PC1_2
Cluster 3
  Biodiversity and habitat_n_2
Cluster 4
  IDH-2014_n_2
Cluster 5
  EV-Forests_n_2
Cluster 6
  Water withdrawals_np_2
Cluster 7
  HPI_n_2
Cluster 8
  EV-Fisheries_np_2
Cluster 9
  Health and Wellness_n_2  PC4_2
Cluster 10
  Religious tolerance_n_2
Cluster 11
  PC2_2
Cluster 12
  PC3_2
```

Dendrogram



No gráfico e tabela acima, identificamos o Cluster 2 como o mais representativo nesta análise, tanto pelo alto grau de similiaridade entre as variáveis, quanto pela representatividade, concentrando cerca de 55% das observações. O cluster 2 é composto pelas variáveis:

Governança, Democracy Index-2014, SPI-2014, Opportunity, Tolerance for immigrants, Women treated with respect e PC1.

Cluster Analysis of Observations: EPI-2014_n_2, Governança_n, Biodiversity, ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	26	84.4131	28.1983	5	23	5	2
2	25	83.3805	30.0665	10	19	10	2
3	24	81.1390	34.1216	3	27	3	2
4	23	79.2092	37.6128	2	21	2	2
5	22	78.7265	38.4861	2	25	2	3
6	21	78.2612	39.3279	10	26	10	3
7	20	76.7414	42.0773	9	10	9	4
8	19	76.2541	42.9590	2	8	2	4
9	18	76.0414	43.3436	2	24	2	5
10	17	75.2713	44.7369	2	6	2	6
11	16	73.5693	47.8161	2	5	2	8
12	15	73.0987	48.6674	4	12	4	2
13	14	73.0212	48.8076	2	9	2	12
14	13	72.9790	48.8839	2	4	2	14
15	12	72.2946	50.1222	2	22	2	15
16	11	71.5692	51.4344	15	16	15	2
17	10	70.4866	53.3930	2	3	2	17
18	9	67.2331	59.2789	2	18	2	18
19	8	66.9545	59.7829	11	15	11	3
20	7	66.0333	61.4496	11	17	11	4
21	6	65.1424	63.0613	2	7	2	19
22	5	65.0717	63.1891	2	11	2	23
23	4	63.7257	65.6242	1	20	1	2
24	3	63.1701	66.6293	1	13	1	3
25	2	60.1227	72.1424	1	2	1	26
26	1	55.2126	81.0252	1	14	1	27

Final Partition
Number of clusters: 12

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	1	0.0	0.0000	0.0000
Cluster2	15	35580.4	47.3089	74.5527
Cluster3	2	582.1	17.0608	17.0608
Cluster4	1	0.0	0.0000	0.0000
Cluster5	1	0.0	0.0000	0.0000
Cluster6	1	0.0	0.0000	0.0000
Cluster7	1	0.0	0.0000	0.0000
Cluster8	1	0.0	0.0000	0.0000
Cluster9	1	0.0	0.0000	0.0000
Cluster10	1	0.0	0.0000	0.0000
Cluster11	1	0.0	0.0000	0.0000
Cluster12	1	0.0	0.0000	0.0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
EPI-2014_n_2	93.542	93.4179	84.6118	96.099
Governança_n_2	91.784	88.2965	83.8541	77.714
Biodiversity and habitat_n_2	83.080	79.4987	60.4550	100.000

Democracy Index-2014_n_2	89.100	85.7109	77.3697	74.052
SPI-2014_n_2	96.154	94.1313	90.0431	87.491
Opportunity_n_2	96.122	82.9105	86.5463	75.369
IDH-2014_n_2	98.503	92.7521	93.0718	86.058
Tolerance for immigrants_n_2	91.781	68.4932	72.6027	20.548
EV-Forests_n_2	100.000	30.7040	18.9950	9.020
Water withdrawals_np_2	29.770	79.0249	39.4717	44.994
HPI_n_2	46.782	52.0484	35.2859	29.809
Women treated with respect_n_2	80.488	75.8537	71.9512	53.659
EV-Fisheries_np_2	68.973	75.4573	97.3250	55.694
Health and Wellness_n_2	95.590	87.8208	77.4077	79.358
Religious tolerance_n_2	66.667	62.2222	66.6667	100.000
PC1_2	6.396	5.5815	5.1284	4.043
PC2_2	-0.557	-0.8508	-0.2685	-0.102
PC3_2	-1.423	-1.0524	-1.3354	0.032
PC4_2	0.928	0.7866	0.3353	-0.326

Variable	Cluster5	Cluster6	Cluster7	Cluster8
EPI-2014_n_2	90.9108	92.385	76.596	88.4285
Governança_n_2	55.2976	87.505	65.156	60.3270
Biodiversity and habitat_n_2	66.4900	18.740	39.340	79.7700
Democracy Index-2014_n_2	70.6161	85.664	72.749	75.3555
SPI-2014_n_2	73.3825	92.469	69.734	79.6729
Opportunity_n_2	53.5950	91.554	51.115	66.3579
IDH-2014_n_2	86.7886	95.243	91.672	88.1042
Tolerance for immigrants_n_2	27.3973	87.671	16.438	68.4932
EV-Forests_n_2	22.8300	100.000	39.720	55.4100
Water withdrawals_np_2	34.5486	41.659	3.492	32.9946
HPI_n_2	43.2719	47.801	78.689	57.3312
Women treated with respect_n_2	57.3171	86.585	53.659	57.3171
EV-Fisheries_np_2	58.5776	73.586	100.000	60.0673
Health and Wellness_n_2	86.1686	93.339	83.228	99.7300
Religious tolerance_n_2	33.3333	66.667	0.000	33.3333
PC1_2	2.8063	6.016	2.983	3.6299
PC2_2	1.5110	-0.080	3.552	0.0864
PC3_2	0.1235	-1.758	0.393	0.2691
PC4_2	0.8865	0.767	1.629	1.3935

Variable	Cluster9	Cluster10	Cluster11	Cluster12
EPI-2014_n_2	81.168	62.542	90.724	94.998
Governança_n_2	84.285	68.293	66.588	98.717
Biodiversity and habitat_n_2	73.530	50.400	95.170	76.050
Democracy Index-2014_n_2	78.081	77.844	70.972	92.062
SPI-2014_n_2	92.757	80.122	74.245	100.000
Opportunity_n_2	85.338	66.358	55.793	100.000
IDH-2014_n_2	91.054	92.354	78.995	94.900
Tolerance for immigrants_n_2	52.055	49.315	19.178	91.781
EV-Forests_n_2	55.410	33.760	3.300	100.000
Water withdrawals_np_2	38.914	29.124	92.968	72.964
HPI_n_2	60.121	51.128	29.626	69.890
Women treated with respect_n_2	54.878	46.341	47.561	79.268
EV-Fisheries_np_2	59.411	64.376	100.000	91.318
Health and Wellness_n_2	100.000	84.098	30.123	83.648
Religious tolerance_n_2	66.667	100.000	100.000	100.000
PC1_2	4.793	3.631	2.368	6.579
PC2_2	0.115	0.470	0.331	-1.223
PC3_2	-0.203	-0.229	-1.207	-1.304
PC4_2	1.047	0.134	-2.270	0.129

Variable	Grand centroid
EPI-2014_n_2	90.2921
Governança_n_2	83.2528
Biodiversity and habitat_n_2	73.9244
Democracy Index-2014_n_2	82.4776
SPI-2014_n_2	90.2994
Opportunity_n_2	79.9389
IDH-2014_n_2	91.8924

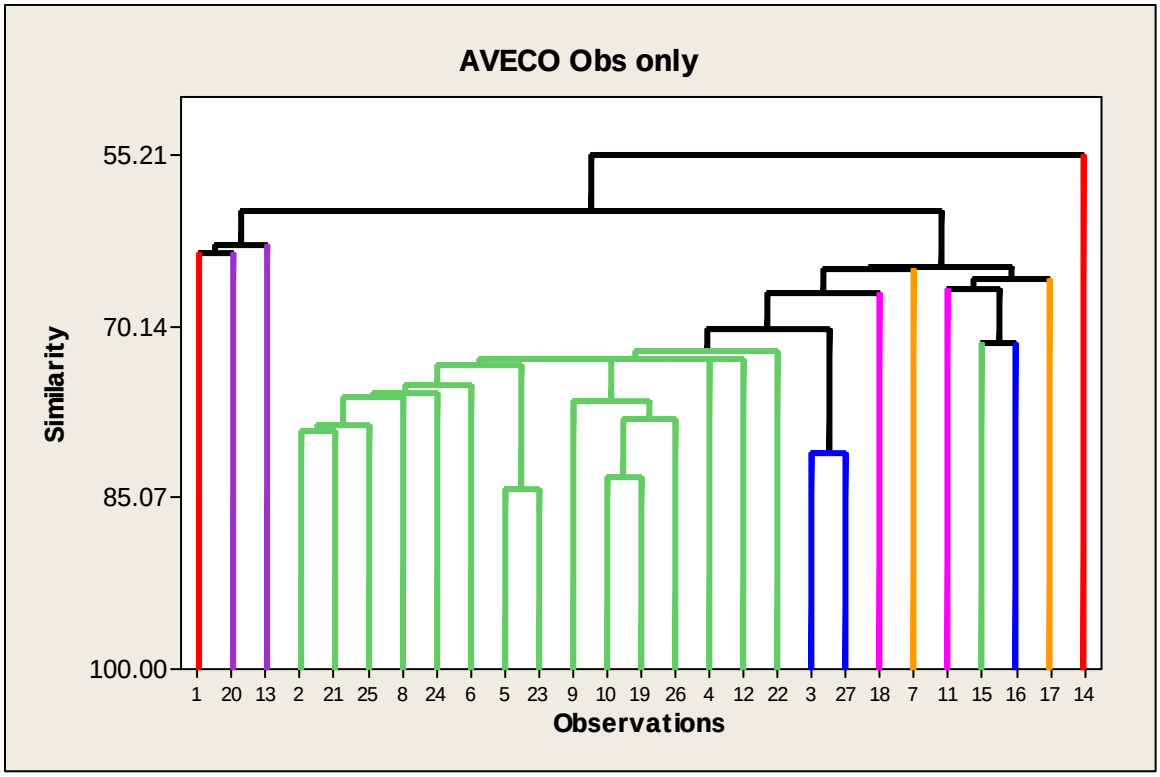
Tolerance for immigrants_n_2	62.8615
EV-Forests_n_2	37.7037
Water withdrawals_np_2	62.4350
HPI_n_2	50.5832
Women treated with respect_n_2	70.3252
EV-Fisheries_np_2	76.2412
Health and Wellness_n_2	85.4597
Religious tolerance_n_2	64.1975
PC1_2	5.0824
PC2_2	-0.3406
PC3_2	-0.8801
PC4_2	0.6217

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0.000	90.503	96.519	132.398	128.138	66.629	152.571
Cluster2	90.503	0.000	56.404	87.350	90.024	102.621	135.754
Cluster3	96.519	56.404	0.000	89.336	86.027	99.505	118.984
Cluster4	132.398	87.350	89.336	0.000	85.930	151.841	149.070
Cluster5	128.138	90.024	86.027	85.930	0.000	131.692	81.025
Cluster6	66.629	102.621	99.505	151.841	131.692	0.000	144.844
Cluster7	152.571	135.754	118.984	149.070	81.025	144.844	0.000
Cluster8	83.000	75.762	81.078	105.670	59.783	100.789	97.134
Cluster9	72.142	60.772	69.305	82.405	75.691	89.010	113.095
Cluster10	111.981	89.677	73.190	78.503	82.585	109.798	120.687
Cluster11	173.165	111.003	113.724	86.964	120.309	180.911	167.802
Cluster12	65.624	88.900	107.531	140.978	153.931	81.524	177.449

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12
Cluster1	83.000	72.142	111.981	173.165	65.624
Cluster2	75.762	60.772	89.677	111.003	88.900
Cluster3	81.078	69.305	73.190	113.724	107.531
Cluster4	105.670	82.405	78.503	86.964	140.978
Cluster5	59.783	75.691	82.585	120.309	153.931
Cluster6	100.789	89.010	109.798	180.911	81.524
Cluster7	97.134	113.095	120.687	167.802	177.449
Cluster8	0.000	51.434	85.676	145.029	117.741
Cluster9	51.434	0.000	61.450	132.836	93.292
Cluster10	85.676	61.450	0.000	117.218	121.081
Cluster11	145.029	132.836	117.218	0.000	159.206
Cluster12	117.741	93.292	121.081	159.206	0.000

Dendrogram



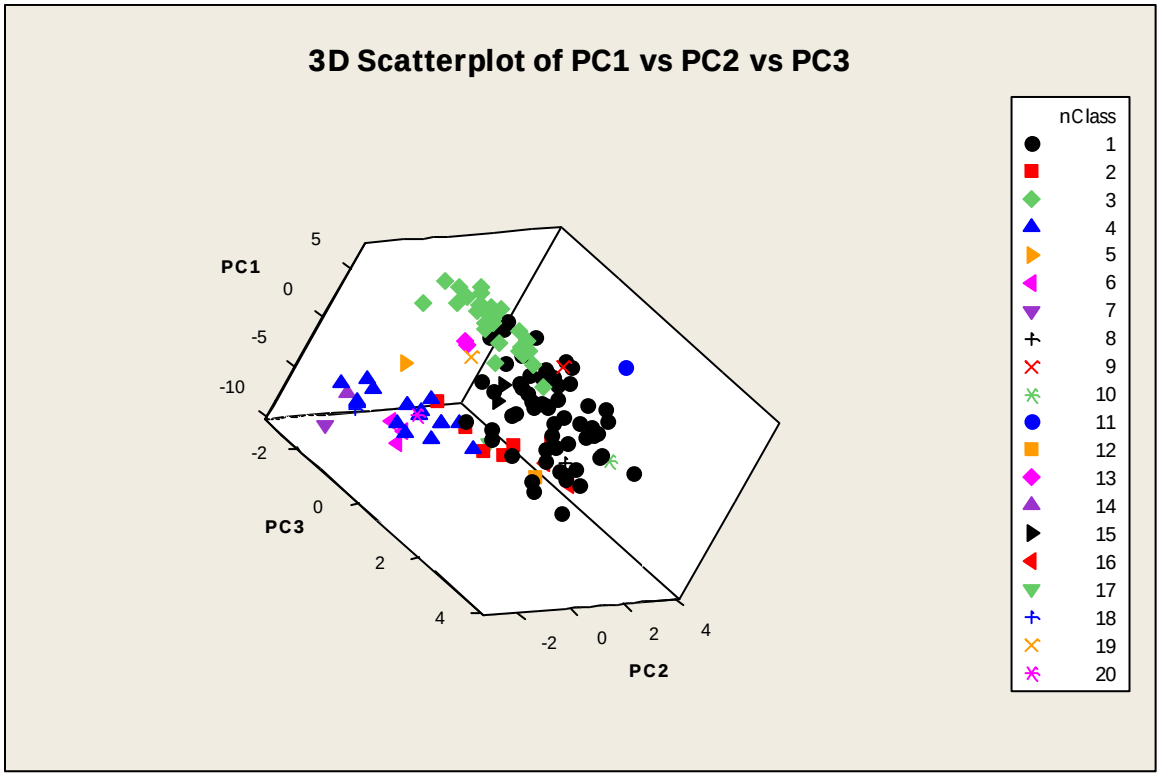
No gráfico e tabela acima, que representa os países do AVECO, identificamos o Cluster 2 como o mais representativo nesta análise, tanto pelo alto grau de similiaridade entre as variáveis, quanto pela representatividade, concentrando cerca de 55% das observações. O cluster 2 é composto pelas variáveis: Governança, Democracy Index-2014, SPI-2014, Opportunity, Tolerance for immigrants, Women treated with respect e PC1.

8.3 Gráficos de Dispersão (*Scatterplot 3D*)

Análise em formato de diagramas de dispersão (*scatterplot*), tomando como referência os 4 PCs e amostras de todos os países, somente AIBER e somente AVECO

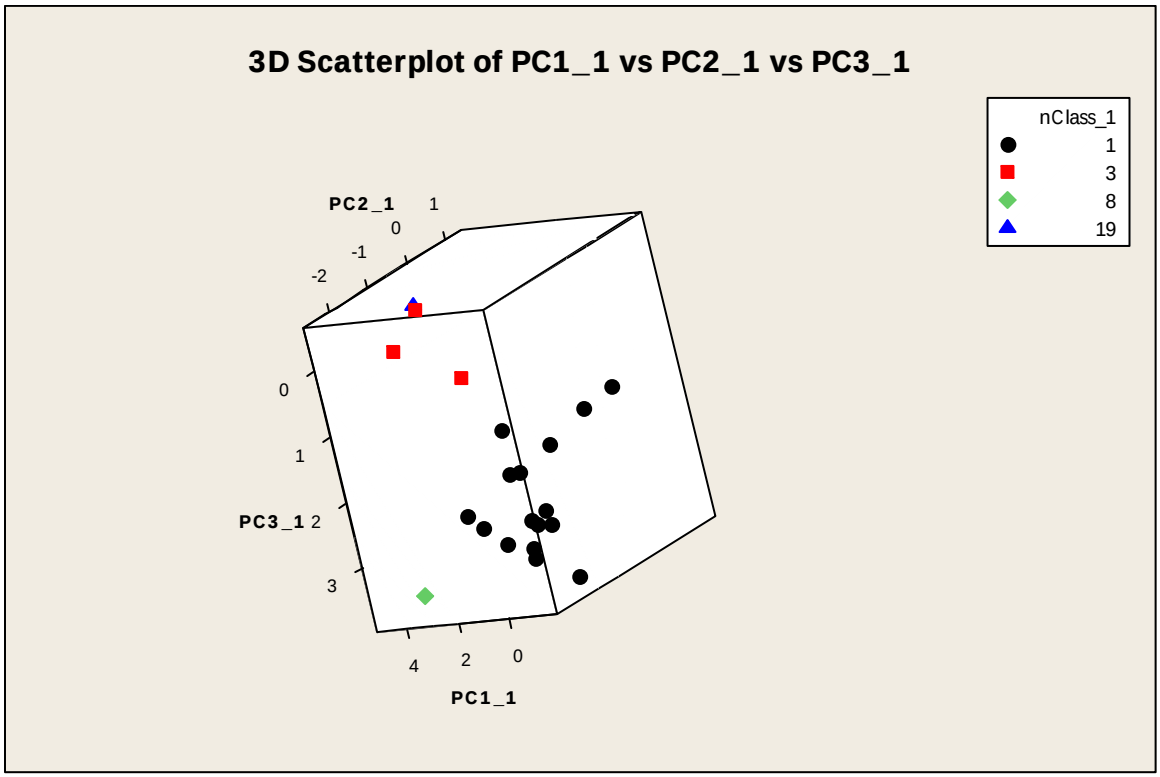
8.3.1 Gráficos de Dispersão (*scatterplot*) – PCs e todos os países e regiões

Análise dos gráficos de dispersão tomando como base os PCs e todos os países e regiões.



8.3.2 Gráficos de Dispersão (*scatterplot*) – PCs e a região AIBER

Análise dos gráficos de dispersão tomando como base os PCs e a região AIBER.

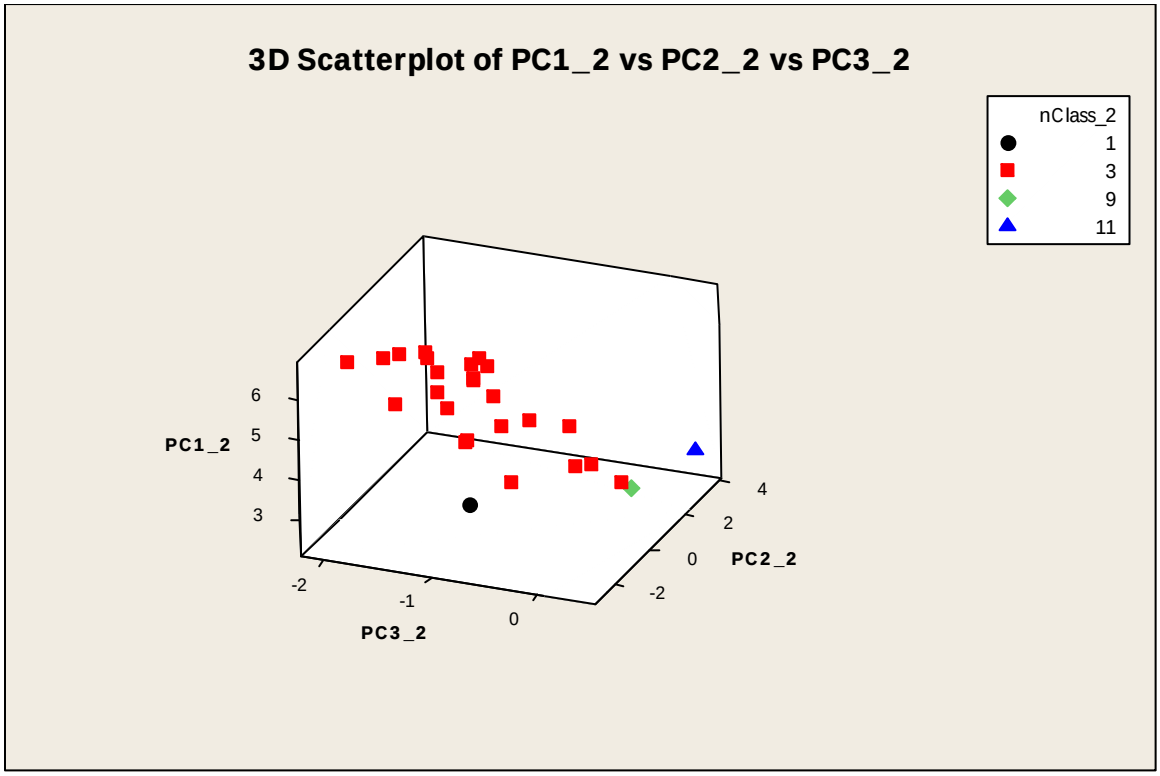


Neste gráfico de dispersão utilizando a coluna Nova Classe (nClass_1) como variável categórica, dos 21 países AIBER analisados destacaram-se Costa Rica (verde), Uruguai

(azul), Portugal, Espanha e Chile (vermelho). Enquanto os demais permaneceram agrupados (cor preta).

8.3.3 Gráficos de Dispersão (*scatterplot*) – PCs e a região AVECO

Análise dos gráficos de dispersão tomando como base os PCs e a região AVECO.



Neste gráfico de dispersão utilizando a coluna Nova Classe (nClass_2) como variável categórica obtivemos, dos 27 países AVECO analisados destacaram-se Letônia (preto), Grécia (verde) e Israel (azul). Enquanto os demais permaneceram agrupados (cor vermelha).

8.4. CONSIDERAÇÕES

Nesta sessão foiram realizadas análises de conglomerados, buscando identificar agrupamentos por similaridade. E análise de dendrogramas e clusters dos dados.

Para estabelecer a quantidade de clusters que melhor explicasse os dados deste estudo, testou-se as quantidades de 1, 10, 12, 20 e 30 clusters entre todas as regiões e somente AIBER e somente AVECO e com todas as variáveis e somente com os PCs. A configuração de 12 clusters refeltiu melhor a situação das três, entre todas estas combinações, por isso foi a adotada.

Durante a execução das análises de cada região e dos PCs, notou-se que o modelo de clusterização escolhido não produziu o efeito esperado, tanto que na região AIBER mais de 50% das variáveis se concentraram no cluster 1, restando apenas 1 variável para os demais clusters. E na quase totalidade dos casos, percebemos uma grande concentração do número de observações numa pequena quantidade de clusters. Não foi possível identificar tendências através dos centróides, uma vez que não apresentaram proporcionalidade nos resultados das colunas de cluster.

9. ANÁLISE DISCRIMINANTE

Esta sessão tratará de algumas ferramentas e técnicas como análise de regressão, a fim de identificar as variáveis que melhor explicam os dados deste estudo, agrupamentos, de maneira que se tenha uma quantidade reduzida de variáveis com o melhor grau de explicação dos resultados. Para em seguida realizar análise discriminante para identificar o grau de acerto na distribuição das observações em seus respectivos grupos.

9.1. Análises de Regressão, de Agrupamentos (cluster) e Discriminante

As variáveis para este trabalho, conforme detalhadas na sessão anterior, serão: EPI-2014, Governança, Biodiversity and habitat, Democracy Index-2014, SPI-2014, Opportunity, IDH-2014, Tolerance for immigrants, EV-Forests, Water withdrawals, HPI, Women treated with respect, EV-Fisheries, Health and Wellness, Religious tolerance.

Utilizamos da função “General Regression” tendo SPI como “Response”, para analisar a equação abaixo:

General Regression Analysis: SPI-2014_n versus Opportunity, Water withdrawals, ...

```
Regression Equation

SPI-2014_n  =  -3.18942 + 0.507796 Opportunity_n - 0.00943243 Water
               withdrawals_np + 0.0491839 Biodiversity and habitat_n +
               0.0507202 Women treated with respect_n - 0.0470843 Tolerance for
               immigrants_n - 0.00713546 Religious tolerance_n + 0.238379
               Governança_n + 0.339865 EPI-2014_n + 0.0235241 EV-Forests_n +
               0.0104549 EV-Fisheries_np + 0.118185 HPI_n - 0.101345 Democracy
               Index-2014_n


Coefficients

Term                            Coef  SE Coef         T      P
Constant                     -3.18942  4.03399  -0.79064  0.431
Opportunity_n                 0.50780  0.08149   6.23110  0.000
Water withdrawals_np         -0.00943  0.01780  -0.52984  0.597
```

```
Biodiversity and habitat_n     0.04918  0.02088   2.35608  0.020
Women treated with respect_n   0.05072  0.02702   1.87730  0.063
Tolerance for immigrants_n    -0.04708  0.03036  -1.55091  0.124
Religious tolerance_n         -0.00714  0.01767  -0.40391  0.687
Governança_n                   0.23838  0.06194   3.84867  0.000
EPI-2014_n                     0.33987  0.03549   9.57751  0.000
EV-Forests_n                   0.02352  0.01962   1.19916  0.233
EV-Fisheries_np                0.01045  0.02846   0.36735  0.714
HPI_n                          0.11818  0.02606   4.53513  0.000
Democracy Index-2014_n        -0.10134  0.04998  -2.02765  0.045


Summary of Model

S = 5.48267      R-Sq = 95.81%        R-Sq(adj) = 95.39%
PRESS = 4471.62  R-Sq(pred) = 94.76%


Analysis of Variance

Source                           DF   Seq SS   Adj SS   Adj MS        F
Regression                       12  81755.9  81755.9  6812.99  226.649
  Opportunity_n                   1  72531.5   1167.1  1167.12   38.827
  Water withdrawals_np            1    983.2      8.4     8.44    0.281
  Biodiversity and habitat_n      1    527.0    166.9   166.86    5.551
  Women treated with respect_n    1    124.3    105.9   105.94    3.524
  Tolerance for immigrants_n      1   2065.2     72.3    72.30    2.405
  Religious tolerance_n           1    351.9      4.9     4.90    0.163
  Governança_n                    1    237.6    445.3   445.25   14.812
  EPI-2014_n                      1   4214.3   2757.3  2757.33   91.729
  EV-Forests_n                    1     34.2     43.2    43.22    1.438
  EV-Fisheries_np                 1      0.1      4.1     4.06    0.135
  HPI_n                           1    562.9    618.2   618.25   20.567
  Democracy Index-2014_n          1    123.6    123.6   123.59    4.111
Error                           119   3577.1   3577.1    30.06
Total                           131  85333.0

Source                                 P
Regression                      0.000000
  Opportunity_n                 0.000000
  Water withdrawals_np          0.597211
  Biodiversity and habitat_n    0.020104
  Women treated with respect_n  0.062925
  Tolerance for immigrants_n    0.123578
  Religious tolerance_n         0.687006
  Governança_n                  0.000193
  EPI-2014_n                    0.000000
  EV-Forests_n                  0.232850
  EV-Fisheries_np               0.714007
  HPI_n                         0.000014
  Democracy Index-2014_n        0.044829
Error
Total


Fits and Diagnostics for Unusual Observations

Obs  SPI-2014_n      Fit   SE Fit  Residual  St Resid
 14     58.2135  44.0759  1.95875   14.1376   2.76080  R
 37     49.1912  38.0216  1.57593   11.1696   2.12703  R
 64     80.1222  69.1822  1.35236   10.9400   2.05900  R
 65     68.4040  52.5324  2.02685   15.8717   3.11559  R
113     29.2416  40.2637  1.37335  -11.0222  -2.07657  R
132     31.0568  45.5732  1.37424  -14.5164  -2.73500  R

R denotes an observation with a large standardized residual.
```

Alterando variável de Response de SPI para IDH:

General Regression Analysis: IDH-2014_n versus Opportunity, Water withdrawals, ...

Regression Equation

```
IDH-2014_n  =  14.6144 + 0.365813 Opportunity_n - 0.0977047 Water
               withdrawals_np + 0.00925913 Biodiversity and habitat_n +
               0.0142142 Women treated with respect_n - 0.106155 Tolerance for
               immigrants_n - 0.0776389 Religious tolerance_n + 0.255818
               Governança_n + 0.539926 EPI-2014_n + 0.0204428 EV-Forests_n +
               0.0533983 EV-Fisheries_np + 0.0927233 HPI_n - 0.200201 Democracy
               Index-2014_n
```

Coefficients

Term	Coef	SE Coef	T	P
Constant	14.6144	6.31033	2.31594	0.022
Opportunity_n	0.3658	0.12748	2.86958	0.005
Water withdrawals_np	-0.0977	0.02785	-3.50847	0.001
Biodiversity and habitat_n	0.0093	0.03266	0.28354	0.777
Women treated with respect_n	0.0142	0.04226	0.33633	0.737
Tolerance for immigrants_n	-0.1062	0.04749	-2.23529	0.027
Religious tolerance_n	-0.0776	0.02763	-2.80945	0.006
Governança_n	0.2558	0.09689	2.64032	0.009
EPI-2014_n	0.5399	0.05551	9.72664	0.000
EV-Forests_n	0.0204	0.03069	0.66617	0.507
EV-Fisheries_np	0.0534	0.04452	1.19944	0.233
HPI_n	0.0927	0.04077	2.27457	0.025
Democracy Index-2014_n	-0.2002	0.07819	-2.56059	0.012

Summary of Model

```
S = 8.57648      R-Sq = 90.44%        R-Sq(adj) = 89.48%
PRESS = 11007.1  R-Sq(pred) = 87.98%
```

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F
Regression	12	82815.4	82815.4	6901.28	93.8235
Opportunity_n	1	56512.3	605.7	605.70	8.2345
Water withdrawals_np	1	7027.5	905.4	905.43	12.3094
Biodiversity and habitat_n	1	215.3	5.9	5.91	0.0804
Women treated with respect_n	1	41.8	8.3	8.32	0.1131
Tolerance for immigrants_n	1	5698.9	367.5	367.52	4.9965
Religious tolerance_n	1	2082.4	580.6	580.58	7.8930
Governança_n	1	183.6	512.8	512.78	6.9713
EPI-2014_n	1	10122.0	6959.0	6958.95	94.6075
EV-Forests_n	1	52.0	32.6	32.64	0.4438
EV-Fisheries_np	1	106.6	105.8	105.82	1.4387
HPI_n	1	290.9	380.6	380.56	5.1737
Democracy Index-2014_n	1	482.3	482.3	482.28	6.5566
Error	119	8753.2	8753.2	73.56	
Total	131	91568.6			

Source	P
Regression	0.000000
Opportunity_n	0.004865
Water withdrawals_np	0.000637
Biodiversity and habitat_n	0.777253
Women treated with respect_n	0.737218
Tolerance for immigrants_n	0.027264
Religious tolerance_n	0.005803
Governança_n	0.009393
EPI-2014_n	0.000000
EV-Forests_n	0.506592
EV-Fisheries_np	0.232741

```
  HPI_n                          0.024723
  Democracy Index-2014_n         0.011700
Error
Total

Fits and Diagnostics for Unusual Observations

Obs  IDH-2014_n      Fit   SE Fit  Residual  St Resid
 64     92.3537  68.6928  2.11548   23.6609   2.84678  R
 78     26.4618  10.3259  3.36911   16.1359   2.04588  R
 84     46.9591  65.8935  2.74933  -18.9344  -2.33071  R
105     19.7499  36.9712  2.70060  -17.2213  -2.11559  R
113     30.6198  49.6268  2.14832  -19.0070  -2.28916  R
116     46.3718  63.5185  3.11384  -17.1467  -2.14568  R

R denotes an observation with a large standardized residual.
```

Com a alteração da variável de "response"de SPI para IDH, houve uma redução no grau de explicação (r quadrado) de 95 para 90%, portanto utilizaremos o IDH para as análises a seguir.

Iniciar processo de retirar variáveis a ver se equilibra a proporcionalidade, mas mantém r2 próximo do inicial (+ou- 90%).

Excluindo as seguintes variáveis que apresentaram p-value > 0.05

```
Biodiversity and habitat_n     0.0093  0.03266   0.28354  0.777
Women treated with respect_n   0.0142  0.04226   0.33633  0.737
EV-Forests_n                   0.0204  0.03069   0.66617  0.507
EV-Fisheries_np                0.0534  0.04452   1.19944  0.233
Health and Wellness_n          0.0854  0.05288   1.61442  0.109
```

Executar a regressão novamente considerando as seguintes variáveis:

```
Term                             Coef  SE Coef         T      P
Constant                      14.6144  6.31033   2.31594  0.022
Opportunity_n                  0.3658  0.12748   2.86958  0.005
Water withdrawals_np          -0.0977  0.02785  -3.50847  0.001
Tolerance for immigrants_n    -0.1062  0.04749  -2.23529  0.027
Religious tolerance_n         -0.0776  0.02763  -2.80945  0.006
Governança_n                   0.2558  0.09689   2.64032  0.009
EPI-2014_n                     0.5399  0.05551   9.72664  0.000
HPI_n                          0.0927  0.04077   2.27457  0.025
Democracy Index-2014_n        -0.2002  0.07819  -2.56059  0.012
```

Resultados da regressão:

General Regression Analysis: IDH-2014_n versus Opportunity, Water withdrawals, ...

```
Regression Equation

IDH-2014_n  =  21.1823 + 0.368822 Opportunity_n - 0.0991802 Water
               withdrawals_np - 0.105761 Tolerance for immigrants_n - 0.0820697
               Religious tolerance_n + 0.290008 Governança_n + 0.539847
               EPI-2014_n + 0.0816496 HPI_n - 0.227971 Democracy Index-2014_n
```

Coefficients

Term	Coef	SE Coef	T	P
Constant	21.1823	3.89519	5.4381	0.000
Opportunity_n	0.3688	0.12223	3.0174	0.003
Water withdrawals_np	-0.0992	0.02685	-3.6943	0.000
Tolerance for immigrants_n	-0.1058	0.04561	-2.3186	0.022
Religious tolerance_n	-0.0821	0.02661	-3.0844	0.003
Governança_n	0.2900	0.08230	3.5239	0.001
EPI-2014_n	0.5398	0.05042	10.7065	0.000
HPI_n	0.0816	0.03966	2.0585	0.042
Democracy Index-2014_n	-0.2280	0.07112	-3.2056	0.002

Summary of Model

S = 8.51629 R-Sq = 90.26% R-Sq(adj) = 89.62%
PRESS = 10414.3 R-Sq(pred) = 88.63%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F
Regression	8	82647.7	82647.7	10331.0	142.443
Opportunity_n	1	56512.3	660.3	660.3	9.105
Water withdrawals_np	1	7027.5	989.8	989.8	13.648
Tolerance for immigrants_n	1	5646.3	389.9	389.9	5.376
Religious tolerance_n	1	2109.0	690.0	690.0	9.513
Governança_n	1	254.2	900.6	900.6	12.418
EPI-2014_n	1	10205.9	8313.8	8313.8	114.630
HPI_n	1	147.3	307.3	307.3	4.238
Democracy Index-2014_n	1	745.3	745.3	745.3	10.276
Error	123	8920.8	8920.8	72.5	
Total	131	91568.6			

Source	P
Regression	0.0000000
Opportunity_n	0.0030994
Water withdrawals_np	0.0003303
Tolerance for immigrants_n	0.0220676
Religious tolerance_n	0.0025192
Governança_n	0.0005979
EPI-2014_n	0.0000000
HPI_n	0.0416509
Democracy Index-2014_n	0.0017172
Error	
Total	

Fits and Diagnostics for Unusual Observations

Obs	IDH-2014_n	Fit	SE Fit	Residual	St Resid	
64	92.3537	69.3427	1.99433	23.0110	2.77928	R
65	78.5761	61.9678	2.66048	16.6084	2.05294	R
78	26.4618	10.3997	3.01358	16.0622	2.01653	R
84	46.9591	65.2777	2.42454	-18.3186	-2.24387	R
105	19.7499	36.8387	2.62333	-17.0888	-2.10916	R
113	30.6198	50.2481	1.67926	-19.6282	-2.35094	R
116	46.3718	63.2430	2.52604	-16.8712	-2.07441	R

R denotes an observation with a large standardized residual.

Aparentemente obtivemos um resultado satisfatório com esta composição de variáveis. O coeficiente de determinação (r-quadrado) se manteve na ordem de 90% e o p-value < 0.05.

Mesmo com várias combinações, não foi possível melhorar a proporcionalidade (peso) das variáveis na equação final.

Utilizaremos as mesmas variáveis da regressão para criar agrupamentos de similaridade.

Executar a função Cluster Obs, com as variáveis:
'Opportunity_n' 'Water withdrawals_np' 'Tolerance for immigrants_n' 'Religious tolerance_n' 'Governança_n' 'EPI-2014_n' 'HPI_n' 'Democracy Index-2014_n' 'IDH-2014_n'

Utilizar os parâmetros:
Linkage Method: single
Distance measure: Manhattan
Nr of cluster: 4

Cluster Analysis of Observations: Opportunity, Water withdrawals, Tolerance for Im, ...

Manhattan Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95.3222	31.821	35	97	35	2
2	130	95.2452	32.345	56	127	56	2
3	129	95.0434	33.717	6	56	6	3
4	128	94.5520	37.060	68	107	68	2
5	127	94.4723	37.602	93	115	93	2
6	126	94.3038	38.748	33	40	33	2
7	125	94.1522	39.780	71	85	71	2
8	124	93.9747	40.987	43	88	43	2
9	123	93.7851	42.277	46	49	46	2
10	122	93.7736	42.356	43	126	43	3
11	121	93.7357	42.613	72	99	72	2
12	120	93.7288	42.660	32	72	32	3
13	119	93.6954	42.887	17	101	17	2
14	118	93.5863	43.630	11	60	11	2
15	117	93.5066	44.171	36	48	36	2
16	116	93.3775	45.050	33	93	33	4
17	115	93.3023	45.561	22	89	22	2
18	114	93.0428	47.327	19	119	19	2
19	113	92.6992	49.664	41	43	41	4
20	112	92.5858	50.435	46	90	46	3
21	111	92.5359	50.775	41	114	41	5
22	110	92.2519	52.707	32	108	32	4
23	109	92.2292	52.861	32	50	32	5
24	108	92.0678	53.959	18	71	18	3
25	107	92.0107	54.348	17	83	17	3
26	106	92.0011	54.413	1	36	1	3
27	105	91.9677	54.640	12	18	12	4
28	104	91.5875	57.227	19	21	19	3
29	103	91.5277	57.633	5	122	5	2
30	102	91.3053	59.146	5	42	5	3
31	101	91.2231	59.705	6	11	6	5
32	100	91.0282	61.031	116	129	116	2

33	99	91.0262	61.045	12	105	12	5
34	98	90.9845	61.328	7	33	7	5
35	97	90.9715	61.417	12	132	12	6
36	96	90.9698	61.429	1	35	1	5
37	95	90.9201	61.766	1	46	1	8
38	94	90.9053	61.867	1	38	1	9
39	93	90.8843	62.010	5	121	5	4
40	92	90.8471	62.263	1	86	1	10
41	91	90.7767	62.742	1	98	1	11
42	90	90.4941	64.664	6	25	6	6
43	89	90.4914	64.683	32	68	32	7
44	88	90.3854	65.404	16	106	16	2
45	87	90.2789	66.129	19	78	19	4
46	86	90.2225	66.512	2	26	2	2
47	85	90.1438	67.047	5	8	5	5
48	84	89.9559	68.326	30	32	30	8
49	83	89.8824	68.826	5	81	5	6
50	82	89.8404	69.111	5	111	5	7
51	81	89.7982	69.398	1	13	1	12
52	80	89.6163	70.636	6	41	6	11
53	79	89.5484	71.097	4	16	4	3
54	78	89.4107	72.034	19	28	19	5
55	77	89.3852	72.207	6	110	6	12
56	76	89.2439	73.169	6	7	6	17
57	75	89.2382	73.208	84	116	84	3
58	74	89.2323	73.248	5	80	5	8
59	73	89.2307	73.259	23	47	23	2
60	72	89.1611	73.732	2	5	2	10
61	71	89.1454	73.839	12	91	12	7
62	70	88.7711	76.385	12	19	12	12
63	69	88.7404	76.594	6	30	6	25
64	68	88.6017	77.537	12	75	12	13
65	67	88.5329	78.006	1	4	1	15
66	66	88.4367	78.660	73	109	73	2
67	65	88.4294	78.710	6	100	6	26
68	64	88.3745	79.083	6	58	6	27
69	63	88.2875	79.675	2	53	2	11
70	62	88.2731	79.773	2	14	2	12
71	61	88.1950	80.304	12	103	12	14
72	60	88.1569	80.563	12	44	12	15
73	59	88.1411	80.671	63	92	63	2
74	58	88.1360	80.706	22	51	22	3
75	57	88.1044	80.920	2	118	2	13
76	56	87.9110	82.236	12	77	12	16
77	55	87.8902	82.378	117	123	117	2
78	54	87.8509	82.645	1	17	1	18
79	53	87.8318	82.775	1	6	1	45
80	52	87.8009	82.985	2	52	2	14
81	51	87.7930	83.039	1	45	1	46
82	50	87.7001	83.671	2	12	2	30
83	49	87.6920	83.726	1	2	1	76
84	48	87.5094	84.968	1	22	1	79
85	47	87.4574	85.322	1	61	1	80
86	46	87.3431	86.099	55	94	55	2
87	45	87.1832	87.187	3	117	3	3
88	44	87.1611	87.337	1	69	1	81
89	43	87.0065	88.389	1	39	1	82
90	42	86.9615	88.695	1	66	1	83
91	41	86.8457	89.483	1	95	1	84
92	40	86.6991	90.480	1	124	1	85
93	39	86.4427	92.224	3	23	3	5
94	38	86.3904	92.580	54	104	54	2
95	37	86.3558	92.816	79	128	79	2
96	36	86.3346	92.960	1	84	1	88
97	35	86.1765	94.035	3	63	3	7
98	34	86.0566	94.851	59	120	59	2
99	33	85.9077	95.864	1	24	1	89
100	32	85.8825	96.035	1	96	1	90

Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	129	802305	76.1081	126.548
Cluster2	1	0	0.0000	0.000
Cluster3	1	0	0.0000	0.000
Cluster4	1	0	0.0000	0.000

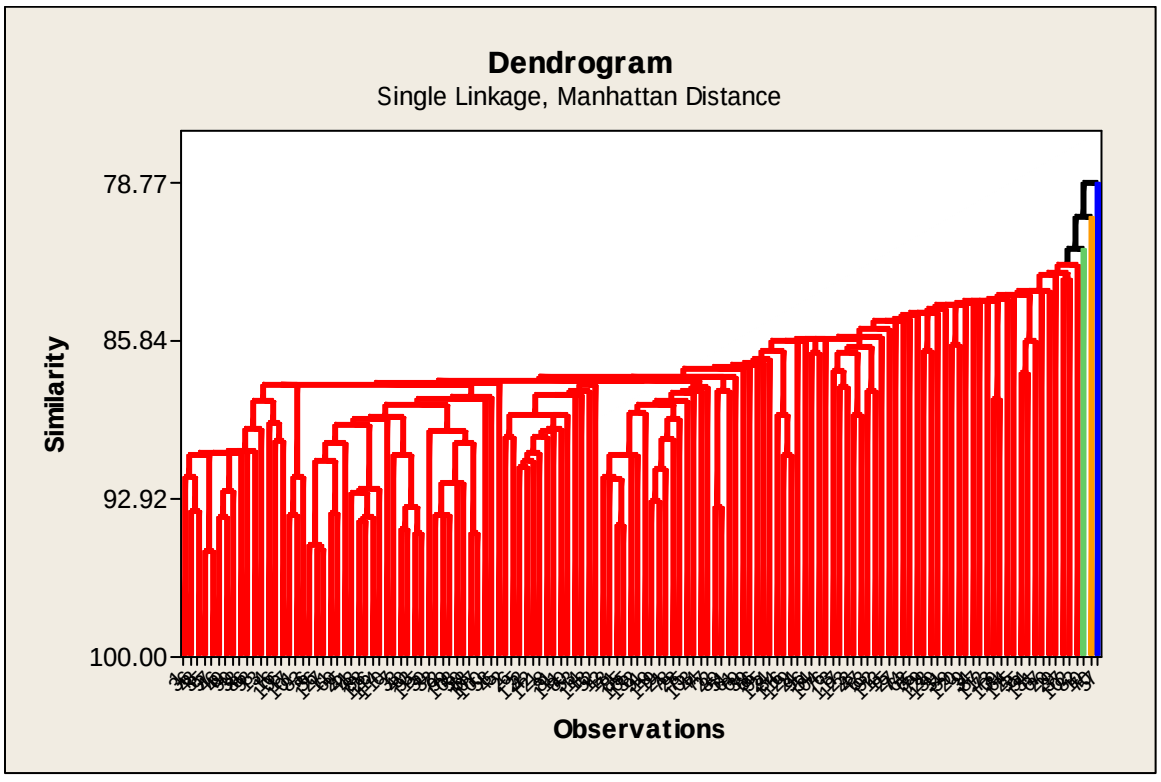
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
Opportunity_n	45.8783	31.115	51.1146	44.364	45.7947
Water withdrawals_np	61.9068	88.172	3.4923	20.547	61.3499
Tolerance for immigrants_n	52.3521	35.616	16.4384	32.877	51.8057
Religious tolerance_n	59.1731	100.000	0.0000	100.000	59.3434
Governança_n	45.0480	21.841	65.1564	44.183	45.0180
EPI-2014_n	60.8271	84.360	76.5957	18.794	60.8064
HPI_n	48.3182	35.767	78.6888	48.454	48.4542
Democracy Index-2014_n	51.9003	26.066	72.7488	61.256	51.9334
IDH-2014_n	59.5345	75.570	91.6719	24.895	59.6370

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0.000	70.916	106.178	82.547
Cluster2	70.916	0.000	155.421	116.282
Cluster3	106.178	155.421	0.000	141.042
Cluster4	82.547	116.282	141.042	0.000

Dendrogram



No dendrograma acima, percebemos uma grande concentração de observações agrupadas no cluster 1 (total de 129 obs), portanto o método “single” não conseguiu distribuir os dados de maneira satisfatória.

Executar novamente, porém utilizando “Linkage Method” como “Complete”, demais parâmetros permanecem os mesmos. Distance measure: Manhattan and Nr of clusters: 4. Gerar coluna de “Cluster Membership column”.

Resultados:

Cluster Analysis of Observations: Opportunity, Water withdrawals, Tolerance for Im, ...

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95.3222	31.821	35	97	35	2
2	130	95.2452	32.345	56	127	56	2
3	129	94.5520	37.060	68	107	68	2
4	128	94.4723	37.602	93	115	93	2
5	127	94.3038	38.748	33	40	33	2
6	126	94.1522	39.780	71	85	71	2
7	125	93.9747	40.987	43	88	43	2
8	124	93.7851	42.277	46	49	46	2
9	123	93.7357	42.613	72	99	72	2
10	122	93.6954	42.887	17	101	17	2
11	121	93.5863	43.630	11	60	11	2
12	120	93.5066	44.171	36	48	36	2
13	119	93.3023	45.561	22	89	22	2
14	118	93.0428	47.327	19	119	19	2
15	117	92.6692	49.868	32	72	32	3
16	116	91.9677	54.640	12	18	12	2
17	115	91.8450	55.475	1	36	1	3
18	114	91.5277	57.633	5	122	5	2
19	113	91.1426	60.253	50	108	50	2
20	112	91.0282	61.031	116	129	116	2
21	111	90.9715	61.417	105	132	105	2
22	110	90.8835	62.015	12	71	12	4
23	109	90.8815	62.029	46	90	46	3
24	108	90.7784	62.730	43	126	43	3
25	107	90.5882	64.024	6	56	6	3
26	106	90.3854	65.404	16	106	16	2
27	105	90.2225	66.512	2	26	2	2
28	104	90.0139	67.931	35	98	35	3
29	103	89.8824	68.826	42	81	42	2
30	102	89.7863	69.479	43	114	43	4
31	101	89.4581	71.712	33	93	33	4
32	100	89.4058	72.068	19	28	19	3
33	99	89.3806	72.239	5	121	5	3
34	98	89.2307	73.259	23	47	23	2
35	97	89.0005	74.824	11	25	11	3
36	96	88.4367	78.660	73	109	73	2
37	95	88.4294	78.710	100	110	100	2
38	94	88.3745	79.083	41	58	41	2
39	93	88.3179	79.468	17	83	17	3
40	92	88.2875	79.675	53	111	53	2
41	91	88.2049	80.237	30	68	30	3
42	90	88.1950	80.304	21	103	21	2
43	89	88.1411	80.671	63	92	63	2

44	88	87.9994	81.635	2	8	2	3
45	87	87.8902	82.378	117	123	117	2
46	86	87.8150	82.889	1	38	1	4
47	85	87.7001	83.671	14	44	14	2
48	84	87.3431	86.099	55	94	55	2
49	83	86.9522	88.758	13	46	13	4
50	82	86.8878	89.197	84	116	84	3
51	81	86.8479	89.468	12	91	12	5
52	80	86.7637	90.041	7	33	7	5
53	79	86.6600	90.746	32	50	32	5
54	78	86.5977	91.170	61	66	61	2
55	77	86.3904	92.580	54	104	54	2
56	76	86.3558	92.816	79	128	79	2
57	75	86.2958	93.224	21	78	21	3
58	74	86.2930	93.242	53	118	53	3
59	73	86.0951	94.589	22	51	22	3
60	72	86.0566	94.851	59	120	59	2
61	71	85.2915	100.056	15	86	15	2
62	70	85.1771	100.834	75	105	75	3
63	69	85.1433	101.064	12	77	12	6
64	68	85.0879	101.440	42	76	42	3
65	67	84.9464	102.403	45	80	45	2
66	66	84.5826	104.877	4	130	4	2
67	65	84.3974	106.138	19	21	19	6
68	64	84.0719	108.352	74	87	74	2
69	63	83.8958	109.550	16	27	16	3
70	62	83.8768	109.679	84	113	84	4
71	61	83.8137	110.108	65	125	65	2
72	60	83.5811	111.690	1	13	1	8
73	59	83.3732	113.105	2	5	2	6
74	58	83.1494	114.627	42	124	42	4
75	57	83.1051	114.928	37	102	37	2
76	56	82.4508	119.379	30	32	30	8
77	55	82.2748	120.576	11	41	11	5
78	54	82.0321	122.227	62	84	62	5
79	53	81.9344	122.892	17	42	17	7
80	52	81.7334	124.259	29	95	29	2
81	51	81.6949	124.521	10	31	10	2
82	50	81.5872	125.254	64	100	64	3
83	49	81.3959	126.556	2	61	2	8
84	48	80.8416	130.326	67	74	67	3
85	47	80.7901	130.676	55	131	55	3
86	46	80.7764	130.769	7	43	7	9
87	45	80.6964	131.314	3	112	3	2
88	44	80.6192	131.839	23	24	23	3
89	43	80.5349	132.412	73	82	73	3
90	42	80.4268	133.147	16	96	16	4
91	41	80.3770	133.486	2	69	2	9
92	40	79.8127	137.325	34	70	34	2
93	39	79.5640	139.017	14	117	14	4
94	38	79.4405	139.857	4	35	4	5
95	37	78.6634	145.143	6	7	6	12
96	36	78.4887	146.332	45	57	45	3
97	35	78.3819	147.058	3	63	3	4
98	34	78.0425	149.367	64	79	64	5
99	33	77.8832	150.451	52	53	52	4
100	32	77.6015	152.367	15	73	15	5

Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	28	74592	49.0583	78.925
Cluster2	34	96312	50.7613	87.392
Cluster3	36	131979	57.7593	109.353
Cluster4	34	77881	46.7241	69.964

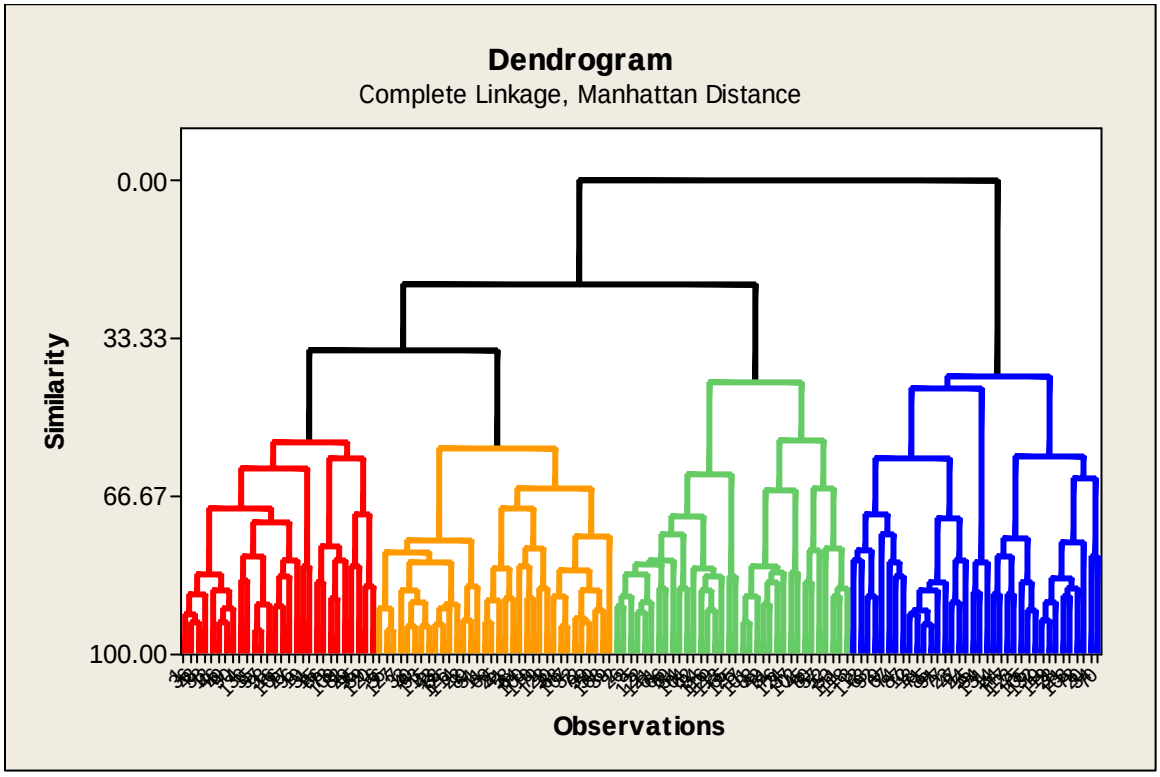
Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
Opportunity_n	50.4906	34.9307	21.7168	78.2856	45.7947
Water withdrawals_np	60.2911	36.7576	81.5041	65.4743	61.3499
Tolerance for immigrants_n	56.1155	34.3674	53.9193	63.4569	51.8057
Religious tolerance_n	88.0952	31.3725	53.7037	69.6078	59.3434
Governança_n	38.8391	36.7271	23.6891	80.9810	45.0180
EPI-2014_n	68.1924	65.6242	23.9045	88.9785	60.8064
HPI_n	61.4019	52.7796	34.4957	48.2457	48.4542
Democracy Index-2014_n	57.2571	38.9044	33.0338	80.5896	51.9334
IDH-2014_n	61.4657	66.1259	24.3879	88.9648	59.6370

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0.0000	70.2162	85.685	69.834
Cluster2	70.2162	0.0000	84.138	98.975
Cluster3	85.6851	84.1377	0.000	133.935
Cluster4	69.8341	98.9752	133.935	0.000

Dendrogram



Resultado visualmente mais uniforme entre as observações dos quatro agrupamentos. Esta constatação é confirmada acima na quantidade de observações distribuídas da seguinte maneira: Cluster1 com 28 obs, Cluster2 com 34 obs, Cluster3 com 36 e Cluster4 com 34 obs.

Abaixo será executada a análise discriminante (Discriminant Analysis), e será mantida as variáveis que já vêm sendo utilizadas nos resultados acima ('Opportunity_n' 'Water withdrawals_np' 'Tolerance for immigrants_n' 'Religious tolerance_n' 'Governança_n' 'EPI-2014_n' 'HPI_n' 'Democracy Index-2014_n' 'IDH-2014_n').

Utilizar como preditora (Groups), a coluna C45, recém criada para armazenar os “Cluster memberships” e nomeada como “ClusterM”.

Resultados:

Discriminant Analysis: C45 versus Opportunity_n, Water withdrawals, ...

```
Linear Method for Response: C45

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n,
            Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n,
            Democracy Index-2014_n, IDH-2014_n


Group         1         2         3         4
Count        28        34        36        34


Summary of classification

                        True Group
Put into Group      1      2      3      4
1                  27      0      0      0
2                   1     32      2      0
3                   0      2     34      0
4                   0      0      0     34
Total N            28     34     36     34
N correct          27     32     34     34
Proportion      0.964  0.941  0.944  1.000

N = 132           N Correct = 127           Proportion Correct = 0.962


Squared Distance Between Groups

         1        2        3        4
1   0.0000  12.8992  27.5292  15.5903
2  12.8992   0.0000  20.3814  19.2151
3  27.5292  20.3814   0.0000  51.6822
4  15.5903  19.2151  51.6822   0.0000


Linear Discriminant Function for Groups

                                 1        2        3        4
Constant                   -54.035  -33.655  -23.064  -66.764
Opportunity_n               -0.265   -0.299   -0.426   -0.139
Water withdrawals_np         0.087    0.064    0.121    0.100
Tolerance for immigrants_n   0.244    0.175    0.247    0.191
Religious tolerance_n        0.299    0.177    0.179    0.245
Governança_n                -0.336   -0.093   -0.090   -0.055
EPI-2014_n                   0.262    0.227    0.045    0.268
HPI_n                        0.163    0.133    0.116    0.121
Democracy Index-2014_n       0.438    0.231    0.322    0.297
IDH-2014_n                   0.590    0.550    0.372    0.669
```

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
14**	3	2	1	21.86	0.004
			2	10.70	0.932
			3	16.07	0.064
			4	26.11	0.000
52**	2	3	1	38.00	0.000
			2	25.75	0.065
			3	20.41	0.935
			4	57.81	0.000
55**	3	2	1	27.83	0.001
			2	13.48	0.951
			3	19.45	0.048
			4	55.62	0.000
73**	1	2	1	10.006	0.469
			2	9.997	0.471
			3	31.029	0.000
			4	14.091	0.061
113**	2	3	1	26.81	0.001
			2	13.81	0.491
			3	13.74	0.508
			4	42.60	0.000

Nos resultados acima, algumas observações foram identificadas como não adequadas para seus agrupamentos, são elas a Bosnia and Herzegovina (linha 14) que está no grupo 3, e suggest to 2, a Índia (52) no 2, mas sugere no 3. O Iraq (55) no 3, mas sugere em 2, a Macedonia (73) no 1 e sugere no 2 e finalmente Swaziland (113) no 2, sugerindo no 3.
Dentre os grupos formados, os que estão mais próximos são os grupos 1 e 2, e os que estão mais distantes são os grupos 3 e 4.

Para o novo teste, foi alterado o “discriminant function” field de Linear para Quadratic, com os seguintes resultados:

Discriminant Analysis: ClusterM versus Opportunity, Water withdrawals, ...

Quadratic Method for Response: ClusterM

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n, Democracy Index-2014_n, IDH-2014_n

Group	1	2	3	4
Count	28	34	36	34

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	28	0	0	0
2	0	34	1	0
3	0	0	35	0
4	0	0	0	34
Total N	28	34	36	34
N correct	28	34	35	34
Proportion	1.000	1.000	0.972	1.000

```
N = 132          N Correct = 131          Proportion Correct = 0.992

From  Generalized Squared Distance to Group

Group       1      2       3       4
1       43.31  69.38   77.58   70.99
2       85.81  46.02   70.76   81.29
3      114.10  74.87   46.25  183.92
4       71.44  72.33  120.05   39.96

Summary of Misclassified Observations

                True   Pred          Squared
Observation    Group  Group  Group  Distance  Probability
         14**      3      2      1    110.70        0.000
                                2     61.45        0.696
                                3     63.11        0.304
                                4    112.75        0.000
```

A quantidade de elementos corretos em seus devidos grupos (grau de acerto), aumentou de 127 para 131, given N 132.

A única observação que apresentou divergência de classificação foi a Bosnia and Herzegovina (linha 14) que está no grupo 3, e sugerido ir para 2. Portanto, a função discriminante quadrática na análise discriminante apresentou melhores resultados que a função discriminante Linear.

9.2. CONSIDERAÇÕES

O objetivo desta sessão foi realizar análises diversas a fim de reduzir a quantidade de variáveis do estudo, porém sem perder seu grau de explicação, significância nos resultados e a relevância do tema.

Seguindo a tendência das sessões anteriores, as variáveis que se destacaram, apresentaram um viés mais voltado ao capital social, que o natural. E mesmo reduzindo a quantidade para 8 (oito) variáveis, obtivemos um resultado satisfatório com a composição. O r-quadrado se manteve na ordem de 90% e o p-value < 0.05.

As observações foram distribídas uniformemente nos 4 (quatro) agrupamentos, com alto grau de acerto (N 131). A única observação que apresentou divergência de classificação foi a Bosnia and Herzegovina.

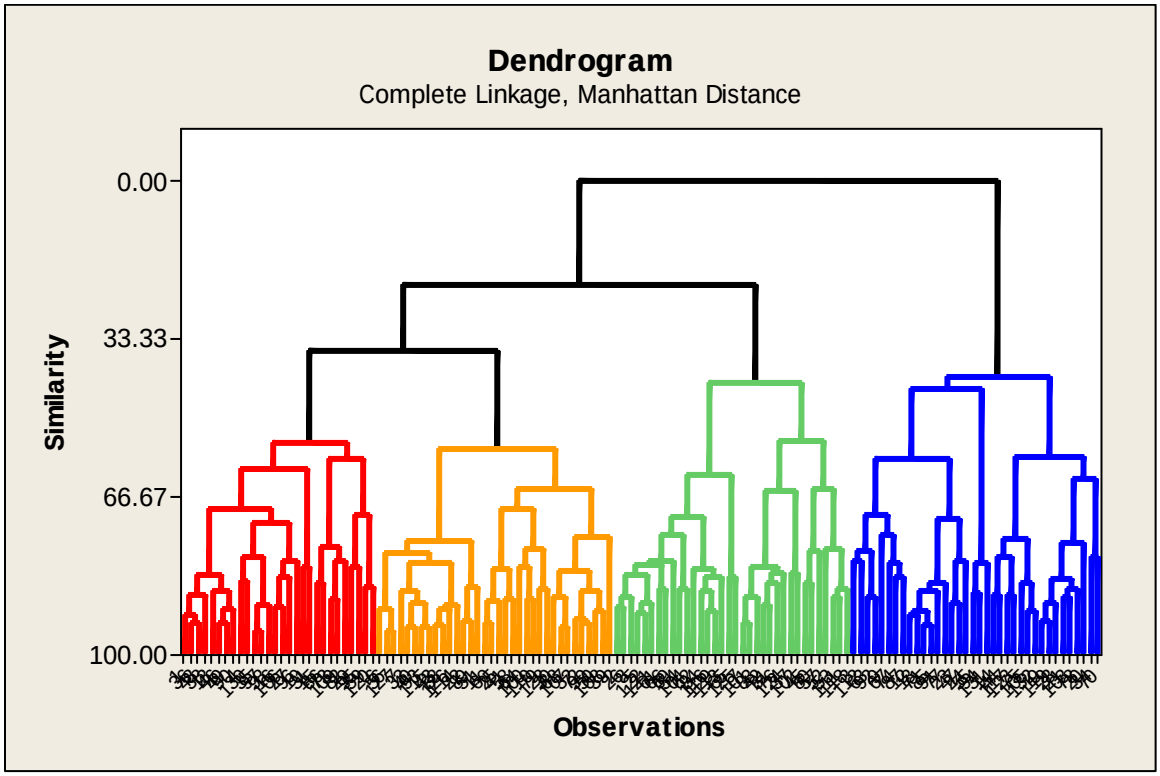
10. REGRESSÃO LOGÍSTICA

Serão executadas análises de regressão logística (*norminal e ordinal*) e discriminante (*linear e quadratic*) de um conjunto de dados para encontrar o modelo que melhor relacione as variáveis escolhidas, e com o maior o grau de acerto na distribuição das observações em seus respectivos grupos. Será mostrada uma tabela com os resultados das regressões logística e discriminante para efeito de comparação de efetividade dos modelos aqui analisados. Do total de 33, serão utilizadas apenas 9 (nove) variáveis previamente identificadas como as de melhor grau de explicação dos resultados.

10.1. Análises de Regressão Logística e Discriminante

Nesta sessão serão realizadas análises de regressão logística (*ordinal and nominal*) e discriminante (*linear and quadratic*), a fim de estabelecer o modelo que melhor represente os dados desta amostra.

As variáveis de trabalho para este estudo, são: EPI-2014, Governança, Democracy Index-2014, Opportunity, IDH-2014, Tolerance for immigrants, Water withdrawals, HPI (Happy Planet Index), Religious tolerance. Além destas 9 (nove) variáveis mencionadas, extraídas da base de dados deste estudo, 2 (duas) variáveis adicionais foram criadas para auxiliar nesta análise. A variável "**3Reg_reOrd**", que representa as 3 (três) regiões AIBER, AVECO e OTHERS ordenadas pelos resultados de uma ANOVA (análise de variância), classificando os resultados dos mais baixos (piores) para os mais altos (melhores). E a variável "**ClusterM_reOrd**", que é resultado de 4 (quatro) agrupamentos por similaridade, através do método *Complete* e medida de distância *Manhattan*, classificados em ordem crescente de média.



10.2 Análises de Regressão Logística (*nominal and ordinal*)

A análise de regressão logística tem como função evidenciar um modelo estatístico que melhor explique os dados e possa permitir predição de valores. É útil para modelar a probabilidade de um evento ocorrer como função de outros fatores. É comparável às técnicas supervisionadas utilizadas em aprendizagem automática, como árvores de decisão, redes neurais, etc... *fonte: Advanced Econometrics Harvard University Press via internet*

Neste trabalho, as análises de regressão logística terão como *response* as novas variáveis "ClusterM_reOrd" e "ClusterM_reOrd", criadas para este estudo e como *predictors*, as 9 (nove) variáveis definidas na sessão anterior.

Iniciando com a regressão logística ordinal, temos:

Ordinal Logistic Regression: ClusterM_reOrd versus Opportunity, Water withdrawals,..

```
Link Function: Logit

Response Information

Variable        Value  Count
ClusterM_reOrd  1         36
                2         28
                3         34
                4         34
                Total    132

Logistic Regression Table
                                                                        95%
                                                                Odds     CI
Predictor                          Coef    SE Coef      Z      P  Ratio  Lower
```

```
Const(1)                        3.72950    1.38597   2.69  0.007
Const(2)                        7.32211    1.55484   4.71  0.000
Const(3)                        11.2389    1.81172   6.20  0.000
Opportunity_n                -0.0744932  0.0402693  -1.85  0.064   0.93   0.86
Water withdrawals_np          0.0123907  0.0079966   1.55  0.121   1.01   1.00
Tolerance for immigrants_n    0.0404440  0.0148498   2.72  0.006   1.04   1.01
Religious tolerance_n         0.0199327  0.0090083   2.21  0.027   1.02   1.00
Governança_n                 -0.0887115  0.0269099  -3.30  0.001   0.92   0.87
EPI-2014_n                   -0.0541366  0.0228767  -2.37  0.018   0.95   0.91
HPI_n                        -0.0090037  0.0115265  -0.78  0.435   0.99   0.97
Democracy Index-2014_n        0.0667537  0.0229583   2.91  0.004   1.07   1.02
IDH-2014_n                   -0.0586925  0.0270379  -2.17  0.030   0.94   0.89

Predictor                    Upper
Const(1)
Const(2)
Const(3)
Opportunity_n                 1.00
Water withdrawals_np          1.03
Tolerance for immigrants_n    1.07
Religious tolerance_n         1.04
Governança_n                  0.96
EPI-2014_n                    0.99
HPI_n                         1.01
Democracy Index-2014_n        1.12
IDH-2014_n                    0.99

Log-Likelihood = -71.351
Test that all slopes are zero: G = 222.156, DF = 9, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      218.385  384  1.000
Deviance     142.702  384  1.000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    6209     95.3  Somers' D              0.91
Discordant     301      4.6  Goodman-Kruskal Gamma  0.91
Ties             6      0.1  Kendall's Tau-a        0.68
Total         6516    100.0
```

Apesar do nível de concordância chegar a 95.3%, duas variáveis apresentaram p-value acima do esperado e serão eliminadas da análise, são elas:

```
Water withdrawals_np        p-value = 0.121
HPI_n                       p-value = 0.435
```

Após executado novamente, seguem os reultados:

Ordinal Logistic Regression: ClusterM_reOrd versus Opportunity, Tolerance, ..

```
Link Function: Logit

Response Information

Variable        Value  Count
ClusterM_reOrd  1         36
                2         28
                3         34
                4         34
```

```
                Total    132

Logistic Regression Table
                                                                          95%
                                                                  Odds    CI
Predictor                         Coef    SE Coef      Z      P  Ratio  Lower
Const(1)                       3.94355    1.22689   3.21  0.001
Const(2)                       7.29842    1.41989   5.14  0.000
Const(3)                       11.2859    1.68620   6.69  0.000
Opportunity_n               -0.0798587  0.0401193  -1.99  0.047   0.92   0.85
Tolerance for immigrants_n   0.0389503  0.0145887   2.67  0.008   1.04   1.01
Religious tolerance_n        0.0207656  0.0088759   2.34  0.019   1.02   1.00
Governança_n                -0.0783726  0.0247985  -3.16  0.002   0.92   0.88
EPI-2014_n                  -0.0442065  0.0217034  -2.04  0.042   0.96   0.92
Democracy Index-2014_n       0.0683465  0.0220906   3.09  0.002   1.07   1.03
IDH-2014_n                  -0.0707820  0.0261333  -2.71  0.007   0.93   0.89

Predictor                   Upper
Const(1)
Const(2)
Const(3)
Opportunity_n                1.00
Tolerance for immigrants_n   1.07
Religious tolerance_n        1.04
Governança_n                 0.97
EPI-2014_n                   1.00
Democracy Index-2014_n       1.12
IDH-2014_n                   0.98

Log-Likelihood = -72.832
Test that all slopes are zero: G = 219.194, DF = 7, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      237.835  386  1.000
Deviance     145.664  386  1.000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    6184     94.9  Somers' D              0.90
Discordant     328      5.0  Goodman-Kruskal Gamma  0.90
Ties             4      0.1  Kendall's Tau-a        0.68
Total         6516    100.0
```

Todos com p-value dentro da faixa esperada e manteve-se o nível de concordância próximo ao resultado anterior, ou seja em 94.9%.

Agora executando novamente, mas tendo a coluna das 3 regiões ordenadas (3Reg_reOrd) como *response* e todas as variáveis, temos:

Ordinal Logistic Regression: 3Reg_reOrd versus Opportunity, Water withdrawals, ...

```
Link Function: Logit

Response Information

Variable    Value  Count
3Reg_reOrd  1         84
            2         21
            3         27
            Total    132
```

```
Logistic Regression Table
                                                                                  95%
                                                                          Odds    CI
Predictor                          Coef    SE Coef      Z      P  Ratio  Lower
Const(1)                        19.0782    3.72232   5.13  0.000
Const(2)                        21.8214    3.97668   5.49  0.000
Opportunity_n                -0.0636396  0.0562010  -1.13  0.257   0.94   0.84
Water withdrawals_np         -0.0180067  0.0110094  -1.64  0.102   0.98   0.96
Tolerance for immigrants_n    0.0104075  0.0185100   0.56  0.574   1.01   0.97
Religious tolerance_n        -0.0241964  0.0117145  -2.07  0.039   0.98   0.95
Governança_n                  0.0364954  0.0386019   0.95  0.344   1.04   0.96
EPI-2014_n                    0.0205287  0.0357756   0.57  0.566   1.02   0.95
HPI_n                        -0.0565586  0.0174538  -3.24  0.001   0.95   0.91
Democracy Index-2014_n       -0.0505792  0.0438318  -1.15  0.249   0.95   0.87
IDH-2014_n                    -0.153374  0.0515304  -2.98  0.003   0.86   0.78

Predictor                    Upper
Const(1)
Const(2)
Opportunity_n                 1.05
Water withdrawals_np          1.00
Tolerance for immigrants_n    1.05
Religious tolerance_n         1.00
Governança_n                  1.12
EPI-2014_n                    1.09
HPI_n                         0.98
Democracy Index-2014_n        1.04
IDH-2014_n                    0.95

Log-Likelihood = -47.494
Test that all slopes are zero: G = 143.849, DF = 9, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      101.921  253  1.000
Deviance      94.988  253  1.000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    4414     96.0  Somers' D              0.92
Discordant     176      3.8  Goodman-Kruskal Gamma  0.92
Ties             9      0.2  Kendall's Tau-a        0.49
Total         4599    100.0
```

As seguintes variáveis apresentaram p-value acima do esperado e serão retiradas para realizar nova análise:

```
Opportunity_n                p-value = 0.257
Tolerance for immigrants_n   p-value = 0.574
Governança_n                 p-value = 0.344
EPI-2014_n                   p-value = 0.566
Democracy Index-2014_n       p-value = 0.249
```

Por questão de parcimônia, a variável "**`Water withdrawals_np`**", que apresentou p-value de **`0.102`** , será mantida para a análise que segue.

Ordinal Logistic Regression: 3Reg_reOrd versus Water withdrawals, Religious, ...

```
Link Function: Logit

Response Information
```

```
Variable    Value  Count
3Reg_reOrd  1         84
            2         21
            3         27
            Total    132

Logistic Regression Table
                                                                 Odds     95% CI
Predictor                     Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)                   20.2490    3.22592   6.28  0.000
Const(2)                   22.6211    3.47550   6.51  0.000
Water withdrawals_np   -0.0223421  0.0096957  -2.30  0.021   0.98   0.96   1.00
Religious tolerance_n  -0.0327494  0.0098736  -3.32  0.001   0.97   0.95   0.99
HPI_n                  -0.0653890  0.0170031  -3.85  0.000   0.94   0.91   0.97
IDH-2014_n              -0.185861  0.0289818  -6.41  0.000   0.83   0.78   0.88

Log-Likelihood = -52.948
Test that all slopes are zero: G = 132.942, DF = 4, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      115.006  258  1.000
Deviance     105.895  258  1.000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    4378     95.2  Somers' D              0.90
Discordant     217      4.7  Goodman-Kruskal Gamma  0.91
Ties             4      0.1  Kendall's Tau-a        0.48
Total         4599    100.0
```

A quantidade de variáveis foi reduzida para 4 (quatro), todas apresentaram um p-value dentro da faixa esperada, inclusive a variável "**`Water withdrawals_np`**". O nível de concordância ficou em 95.2% contra os 96.0% no resultado anterior quando executado com as 9 (nove) variáveis.

Abaixo iremos executar a função "*nominal logistic regression*" com a nova coluna dos clusters ordenados (ClusterM_reOrd) e com a nova coluna das regiões ordenadas (3Reg_reOrd), como variáveis *response*.

Nominal Logistic Regression: ClusterM_reOrd versus Opportunity, Water withdrawals,

```
* WARNING * Algorithm has not converged after 20 iterations.
* WARNING * Convergence has not been reached for the parameter estimates
            criterion.
* WARNING * The results may not be reliable.
* WARNING * Try increasing the maximum number of iterations.

Response Information

Variable        Value  Count
ClusterM_reOrd  4         34  (Reference Event)
                3         34
                2         28
                1         36
                Total    132

Logistic Regression Table
                                                                      95%
```

Predictor	Coef	SE Coef	Z	P	Odds Ratio	CI Lower
Logit 1: (3/4)						
Constant	283.644	26330.3	0.01	0.991		
Opportunity_n	-0.739784	472.076	-0.00	0.999	0.48	0.00
Water withdrawals_np	-0.457341	73.6768	-0.01	0.995	0.63	0.00
Tolerance for immigrants_n	-0.646618	173.300	-0.00	0.997	0.52	0.00
Religious tolerance_n	-1.06655	161.395	-0.01	0.995	0.34	0.00
Governança_n	1.05558	407.672	0.00	0.998	2.87	0.00
EPI-2014_n	-0.0587769	302.559	-0.00	1.000	0.94	0.00
HPI_n	0.230493	113.474	0.00	0.998	1.26	0.00
Democracy Index-2014_n	-1.00389	332.920	-0.00	0.998	0.37	0.00
IDH-2014_n	-1.72852	496.859	-0.00	0.997	0.18	0.00
Logit 2: (2/4)						
Constant	78.2069	44733.8	0.00	0.999		
Opportunity_n	-1.04789	575.002	-0.00	0.999	0.35	0.00
Water withdrawals_np	-0.126672	66.9063	-0.00	0.998	0.88	0.00
Tolerance for immigrants_n	0.451843	203.967	0.00	0.998	1.57	0.00
Religious tolerance_n	0.495841	235.964	0.00	0.998	1.64	0.00
Governança_n	-1.88784	442.534	-0.00	0.997	0.15	0.00
EPI-2014_n	0.856805	205.541	0.00	0.997	2.36	0.00
HPI_n	0.0711650	58.2793	0.00	0.999	1.07	0.00
Democracy Index-2014_n	1.21463	526.663	0.00	0.998	3.37	0.00
IDH-2014_n	-1.59111	618.811	-0.00	0.998	0.20	0.00
Logit 3: (1/4)						
Constant	344.971	28900.9	0.01	0.990		
Opportunity_n	-4.03676	641.182	-0.01	0.995	0.02	0.00
Water withdrawals_np	0.493440	90.1376	0.01	0.996	1.64	0.00
Tolerance for immigrants_n	0.317522	210.905	0.00	0.999	1.37	0.00
Religious tolerance_n	-0.100582	224.256	-0.00	1.000	0.90	0.00
Governança_n	-0.536349	515.493	-0.00	0.999	0.58	0.00
EPI-2014_n	-2.02539	333.734	-0.01	0.995	0.13	0.00
HPI_n	-0.536653	175.247	-0.00	0.998	0.58	0.00
Democracy Index-2014_n	0.294046	313.633	0.00	0.999	1.34	0.00
IDH-2014_n	-0.974498	506.512	-0.00	0.998	0.38	0.00

Predictor	Upper
Logit 1: (3/4)	
Constant	
Opportunity_n	*
Water withdrawals_np	3.27456E+62
Tolerance for immigrants_n	1.70687E+147
Religious tolerance_n	8.25362E+136
Governança_n	*
EPI-2014_n	3.25846E+257
HPI_n	4.89379E+96
Democracy Index-2014_n	8.83108E+282
IDH-2014_n	*
Logit 2: (2/4)	
Constant	
Opportunity_n	*
Water withdrawals_np	7.86551E+56
Tolerance for immigrants_n	6.50543E+173
Religious tolerance_n	1.17005E+201
Governança_n	*
EPI-2014_n	2.13108E+175
HPI_n	4.34834E+49
Democracy Index-2014_n	*
IDH-2014_n	*
Logit 3: (1/4)	
Constant	
Opportunity_n	*
Water withdrawals_np	8.70133E+76
Tolerance for immigrants_n	4.57446E+179
Religious tolerance_n	6.96726E+190
Governança_n	*
EPI-2014_n	1.56828E+283
HPI_n	8.65975E+148
Democracy Index-2014_n	1.23780E+267
IDH-2014_n	*

```
Log-Likelihood = -0.000
Test that all slopes are zero: G = 364.858, DF = 27, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson    0.0000018  366  1.000
Deviance   0.0000036  366  1.000
```

O método de regressão logística *Nominal* não conseguiu analisar adequadamente os dados de nossa base.

Tomando a coluna das 3 (três) regiões ordenadas, e executando a função *nominal logistic regression*, temos os seguintes resultados

Nominal Logistic Regression: 3Reg_reOrd versus Opportunity, Water withdrawals, ...

```
* WARNING * Algorithm has not converged after 20 iterations.
* WARNING * Convergence has not been reached for the parameter estimates
            criterion.
* WARNING * The results may not be reliable.
* WARNING * Try increasing the maximum number of iterations.

Response Information

Variable    Value  Count
3Reg_reOrd  3         27  (Reference Event)
            2         21
            1         84
            Total    132

Logistic Regression Table
                                                                          95%
                                                                           CI
Predictor                          Coef  SE Coef      Z      P  Odds Ratio  Lower
Logit 1: (2/3)
Constant                        1787.37  26631.6   0.07  0.946
Opportunity_n                   5.60972  462.674   0.01  0.990      273.07   0.00
Water withdrawals_np          -0.411068  31.4881  -0.01  0.990        0.66   0.00
Tolerance for immigrants_n    0.968739  97.7937   0.01  0.992        2.63   0.00
Religious tolerance_n         -1.67529  75.8557  -0.02  0.982        0.19   0.00
Governança_n                  0.252903  394.961   0.00  0.999        1.29   0.00
EPI-2014_n                     1.59830  274.719   0.01  0.995        4.94   0.00
HPI_n                         -1.66043  172.874  -0.01  0.992        0.19   0.00
Democracy Index-2014_n        -12.3069  1412.70  -0.01  0.993        0.00   0.00
IDH-2014_n                    -14.7970  1014.67  -0.01  0.988        0.00   0.00
Logit 2: (1/3)
Constant                        1813.46  26631.6   0.07  0.946
Opportunity_n                   5.51325  462.674   0.01  0.990      247.96   0.00
Water withdrawals_np          -0.448119  31.4881  -0.01  0.989        0.64   0.00
Tolerance for immigrants_n    0.914868  97.7937   0.01  0.993        2.50   0.00
Religious tolerance_n         -1.71150  75.8557  -0.02  0.982        0.18   0.00
Governança_n                  0.368590  394.961   0.00  0.999        1.45   0.00
EPI-2014_n                     1.62405  274.719   0.01  0.995        5.07   0.00
HPI_n                         -1.77813  172.874  -0.01  0.992        0.17   0.00
Democracy Index-2014_n        -12.3383  1412.70  -0.01  0.993        0.00   0.00
IDH-2014_n                    -14.9646  1014.67  -0.01  0.988        0.00   0.00

Predictor                         Upper
Logit 1: (2/3)
Constant
Opportunity_n                         *
Water withdrawals_np          4.20912E+26
Tolerance for immigrants_n    4.60105E+83
```

```
Religious tolerance_n        6.93363E+63
Governança_n                            *
EPI-2014_n                  3.43044E+234
HPI_n                       2.68669E+146
Democracy Index-2014_n                  *
IDH-2014_n                              *
Logit 2: (1/3)
Constant
Opportunity_n                           *
Water withdrawals_np         4.05603E+26
Tolerance for immigrants_n   4.35970E+83
Religious tolerance_n        6.68700E+63
Governança_n                            *
EPI-2014_n                  3.51992E+234
HPI_n                       2.38838E+146
Democracy Index-2014_n                  *
IDH-2014_n                              *

Log-Likelihood = -13.036
Test that all slopes are zero: G = 212.766, DF = 18, P-Value = 0.000

Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      43.6505  244  1.000
Deviance     26.0716  244  1.000
```

O método de regressão logística *Nominal* não conseguiu analisar adequadamente os dados de nossa base.

10.3 Análise Discriminante (*linear and quadratic*)

Nesta sessão será executada a função de Análise Discriminante (*discriminant analysis*), tendo como *response* primeiramente a coluna “'ClusterM_reOrd'” e na sequência a coluna “3Reg_reOrd”. As variáveis preditoras serão as mesmas utilizadas nas regressões da sessão anterior, ou seja, 'Opportunity_n' 'Water withdrawals_np' 'Tolerance for immigrants_n' 'Religious tolerance_n' 'Governança_n' 'EPI-2014_n' 'HPI_n' 'Democracy Index-2014_n' 'IDH-2014_n'.

A análise discriminante é uma técnica da estatística multivariada utilizada para discriminar e classificar objetos. Segundo KHATTREE & NAIK (2000) é uma técnica da estatística multivariada que estuda a separação de objetos de uma população em duas ou mais classes. *fonte: apostila ANÁLISE MULTIVARIADA APLICADA AS CIÊNCIAS AGRÁRIAS - PÓS-GRADUAÇÃO EM AGRONOMIA CIÊNCIA DO SOLO: CPGA – Prof. Carlos Alberto Alves Varella, da Universidade Federal Rural do Rio de Janeiro.*

Iniciamos com o método linear da análise discriminante, cujos resultados são:

Discriminant Analysis: ClusterM_reOrd versus Opportunity, Water withdrawals, ...

```
Linear Method for Response: ClusterM_reOrd

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n,
            Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n,
            Democracy Index-2014_n, IDH-2014_n

Group         1         2         3         4
```

```
Count        36        28        34        34

Summary of classification

                        True Group
Put into Group      1      2      3      4
1                  34      0      2      0
2                   0     27      0      0
3                   2      1     32      0
4                   0      0      0     34
Total N            36     28     34     34
N correct          34     27     32     34
Proportion      0.944  0.964  0.941  1.000

N = 132           N Correct = 127           Proportion Correct = 0.962

Squared Distance Between Groups

         1        2        3        4
1   0.0000  27.5292  20.3814  51.6822
2  27.5292   0.0000  12.8992  15.5903
3  20.3814  12.8992   0.0000  19.2151
4  51.6822  15.5903  19.2151   0.0000


Linear Discriminant Function for Groups

                                    1        2        3        4
Constant                      -23.064  -54.035  -33.655  -66.764
Opportunity_n                  -0.426   -0.265   -0.299   -0.139
Water withdrawals_np            0.121    0.087    0.064    0.100
Tolerance for immigrants_n      0.247    0.244    0.175    0.191
Religious tolerance_n           0.179    0.299    0.177    0.245
Governança_n                   -0.090   -0.336   -0.093   -0.055
EPI-2014_n                      0.045    0.262    0.227    0.268
HPI_n                           0.116    0.163    0.133    0.121
Democracy Index-2014_n          0.322    0.438    0.231    0.297
IDH-2014_n                      0.372    0.590    0.550    0.669

Summary of Misclassified Observations

                True   Pred          Squared
Observation    Group  Group  Group  Distance  Probability
         14**      1      3      1     16.07        0.064
                                 2     21.86        0.004
                                 3     10.70        0.932
                                 4     26.11        0.000
         52**      3      1      1     20.41        0.935
                                 2     38.00        0.000
                                 3     25.75        0.065
                                 4     57.81        0.000
         55**      1      3      1     19.45        0.048
                                 2     27.83        0.001
                                 3     13.48        0.951
                                 4     55.62        0.000
         73**      2      3      1    31.029        0.000
                                 2    10.006        0.469
                                 3     9.997        0.471
                                 4    14.091        0.061
        113**      3      1      1     13.74        0.508
                                 2     26.81        0.001
                                 3     13.81        0.491
                                 4     42.60        0.000
```

Em seguida, passamos para o método quadrático da análise discriminante, para as variáveis ClusterM_reOrd e 3Reg_reOrd, cujos resultados são:

Discriminant Analysis: ClusterM_reOrd versus Opportunity, Water withdrawals, ...

```
Quadratic Method for Response:   ClusterM_reOrd

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n,
            Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n,
            Democracy Index-2014_n, IDH-2014_n

Group         1         2         3         4
Count        36        28        34        34

Summary of classification

                         True Group
Put into Group      1      2      3      4
1                  35      0      0      0
2                   0     28      0      0
3                   1      0     34      0
4                   0      0      0     34
Total N            36     28     34     34
N correct          35     28     34     34
Proportion      0.972  1.000  1.000  1.000

N = 132           N Correct = 131           Proportion Correct = 0.992

From  Generalized Squared Distance to Group

Group       1       2      3       4
1       46.25  114.10  74.87  183.92
2       77.58   43.31  69.38   70.99
3       70.76   85.81  46.02   81.29
4      120.05   71.44  72.33   39.96

Summary of Misclassified Observations

                True   Pred          Squared
Observation    Group  Group  Group  Distance  Probability
         14**      1      3      1     63.11        0.304
                                  2    110.70        0.000
                                  3     61.45        0.696
                                  4    112.75        0.000
```

A análise discriminante utilizando o método *Linear*, apresentou bons resultados com 96.2% de acertos, e 5 (cinco) países mal classificados do total de 132. Já no método *Quadratic*, a proporção de resultados corretos passou a 99.2% , apresentando apenas 1 país como mal classificado. Portanto, vamos utilizar os resultados do método *Quadratic* para nossas considerações.
A observação na linha 14 refere-se o país Bosnia and Herzegovina, que inicialmente foi categorizado como 1 e pela análise do software estatístico, deveria ser realocado no grupo 3.

Executando a análise tendo a coluna '3Reg_reOrd' como *response*, obtivemos os seguintes resultados

Discriminant Analysis: 3Reg_reOrd versus Opportunity, Water withdrawals, ...

```
Linear Method for Response: 3Reg_reOrd

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n,
            Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n,
            Democracy Index-2014_n, IDH-2014_n

Group         1         2         3
```

Count	84	21	27

Summary of classification

	True Group		
Put into Group	1	2	3
1	75	1	2
2	5	16	0
3	4	4	25
Total N	84	21	27
N correct	75	16	25
Proportion	0.893	0.762	0.926

N = 132 N Correct = 116 Proportion Correct = 0.879

Squared Distance Between Groups

	1	2	3
1	0.0000	9.5087	11.5092
2	9.5087	0.0000	11.3303
3	11.5092	11.3303	0.0000

Linear Discriminant Function for Groups

	1	2	3
Constant	-22.171	-42.275	-39.069
Opportunity_n	-0.415	-0.328	-0.291
Water withdrawals_np	0.117	0.139	0.145
Tolerance for immigrants_n	0.251	0.315	0.263
Religious tolerance_n	0.116	0.165	0.110
Governança_n	-0.059	-0.221	0.021
EPI-2014_n	0.013	0.032	-0.027
HPI_n	0.196	0.315	0.229
Democracy Index-2014_n	0.221	0.285	0.210
IDH-2014_n	0.367	0.424	0.441

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	1	2	1	9.652	0.040
			2	3.291	0.957
			3	14.583	0.003
25**	2	3	1	14.361	0.007
			2	11.650	0.028
			3	4.547	0.965
45**	3	1	1	6.440	0.771
			2	14.386	0.015
			3	9.000	0.214
48**	1	2	1	8.973	0.040
			2	2.628	0.959
			3	16.915	0.001
50**	1	3	1	7.248	0.185
			2	12.287	0.015
			3	4.315	0.800
59**	1	2	1	23.790	0.001
			2	9.238	0.999
			3	26.074	0.000
68**	3	1	1	13.04	0.498
			2	20.28	0.013
			3	13.07	0.489
72**	1	3	1	9.885	0.343
			2	21.011	0.001
			3	8.590	0.656
79**	1	3	1	10.830	0.170
			2	14.678	0.025
			3	7.715	0.806
80**	2	1	1	7.230	0.542
			2	7.921	0.383

			3	11.187	0.075
98**	1	2	1	12.975	0.028
			2	5.886	0.970
			3	17.925	0.002
99**	1	3	1	9.932	0.071
			2	15.201	0.005
			3	4.802	0.924
100**	2	3	1	12.507	0.057
			2	10.594	0.149
			3	7.253	0.793
106**	1	2	1	10.124	0.256
			2	8.077	0.712
			3	14.274	0.032
110**	2	3	1	13.657	0.021
			2	8.447	0.284
			3	6.654	0.695
128**	2	3	1	14.037	0.010
			2	7.623	0.246
			3	5.405	0.744

Executando o método: *Quadratic*

Discriminant Analysis: 3Reg_reOrd versus Opportunity, Water withdrawals, ...

Quadratic Method for Response: 3Reg_reOrd

Predictors: Opportunity_n, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, Governança_n, EPI-2014_n, HPI_n, Democracy Index-2014_n, IDH-2014_n

Group	1	2	3
Count	84	21	27

Summary of classification

	True Group		
Put into Group	1	2	3
1	78	0	0
2	2	21	0
3	4	0	27
Total N	84	21	27
N correct	78	21	27
Proportion	0.929	1.000	1.000

N = 132 N Correct = 126 Proportion Correct = 0.955

From Generalized Squared Distance to Group

Group	1	2	3
1	50.39	73.20	425.97
2	59.05	42.27	183.60
3	63.63	60.17	37.67

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	1	2	1	58.70	0.004
			2	47.65	0.996
			3	164.95	0.000
30**	1	3	1	58.89	0.024
			2	121.98	0.000
			3	51.50	0.976
48**	1	2	1	57.85	0.009
			2	48.35	0.991
			3	394.52	0.000

```
50**      1      3      1      56.56        0.147
                        2      78.84        0.000
                        3      53.03        0.853
72**      1      3      1      58.43        0.001
                        2     106.85        0.000
                        3      44.86        0.999
99**      1      3      1      58.68        0.048
                        2      76.39        0.000
                        3      52.70        0.952
```

No método *linear*, obtivemos 87.9% de acerto, resultando em 16 erros, ou 16 países mal alocados conforme análise do *software*. No método *quadratic* a quantidade de erros caiu para 6 e a proporção de acertos subiu para 95.5%, portanto este será o método escolhido para este estudo.
Os 6 países que figuraram como mal classificados pela ferramenta, pertencem ao agrupamento OTHERS, são eles Albânia (obs 1), Croácia (obs 30), Guyana (obs 48), Hungria (obs 50), Lituânia (obs 72) e Filipinas (obs 99).

10.4 Resultados das análises Logística e Discriminante

Nesta sessão, apresentamos uma tabela contendo o conjunto dos resultados dos testes de modelos realizados.

Modelo	ClusterM_reOrd	3Reg_reOrd
AD (ANÁLISE DISCRIMINANTE)	99.2% (Q)	95.5% (Q)
RL (REGRESSÃO LOGÍSTICA)	94.9% (O)	95.2% (O)

legenda:
ClusterM_reOrd = 4 clusters ordenados, sendo de 1 (pior) a 4 (melhor)
3Reg_reOrd = 3 regiões ordenadas, sendo 1 para Others, 2 AIBER e 3 para AVECO
Q = metódo *quadratic* na análise discriminante
O = método *ordinal* na regressão logística

Na tabela vemos que o modelo de análise discriminante apresentou melhores resultados que o de regressão logística.

Apesar dos resultados se apresentarem relativamente próximos, cabe ressaltar que no modelo de análise discriminante (AD), foi possível trabalhar com todas as variáveis em todos os métodos, já para a regressão logística (RL), foram consideradas como relevantes 7 (sete) variáveis quando analisando a ClusterM_reOrd, e 5 (cinco) para a variável 3Reg_reOrd, mostrando portanto robustez superior no primeiro modelo.

10.5. CONSIDERAÇÕES

Como resultado das análises efetuadas, vemos na tabela da sessão 10.4, que o modelo de análise discriminante quadrática foi o que se apresentou mais robustamente para classificar e discriminar os objetos deste estudo.

Em relação ao tema Biodiversidade – Capital Natural e Social, objetivo central desta análise, percebemos nos últimos estudos que as variáveis de maior significância apresentam um viés mais voltado ao capital social que o natural. Talvez a mensagem que está tentando emergir na evolução deste estudo é a de que o indivíduo, assim como na pirâmide de Maslow, necessita primeiramente tratar dos temas que o impacta diretamente (comunidade local, respeito e oportunidades), para depois se dedicar ao entorno (proteção e preservação dos ativos naturais, qualidade de vida e humanidade).

11. ÁRVORES DE CLASSIFICAÇÃO

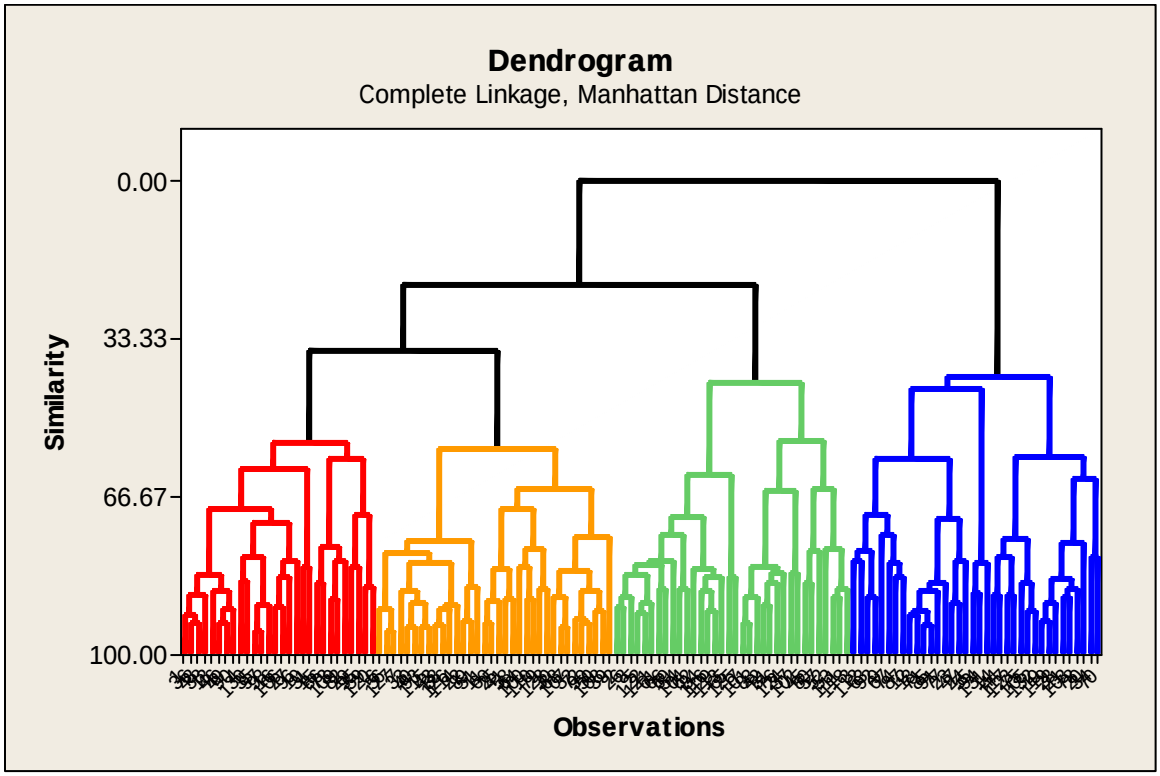
Nesta sessão serão realizadas análises discriminante e de árvores de classificação (árvores de decisão), para classificar hierarquicamente os dados via teste lógico da ferramenta SPSS e encontrar o modelo que melhor relacione as variáveis escolhidas. Ao fim será mostrada uma tabela com os resultados da análises discriminantes efetuadas nos *softwares* Minitab e SPPS para fins comparativos de efetividade.

11.1. Tratamento dos dados, análises discriminante e de árvores de classificação

Os dados originalmente armazenados na base do Minitab, foram exportados para o MS-Excel, tratados e importados no *software* SPSS para a realização das análises constantes neste estudo.

Nesta sessão serão realizadas análises discriminante e de árvores de classificação (árvores de decisão), a fim de estabelecer o modelo que melhor represente os dados desta amostra.

As variáveis de trabalho para este estudo, conforme detalhadas na sessão anterior (quadro 1), são: EPI-2014, Governança, Democracy Index-2014, Opportunity, IDH-2014, Tolerance for immigrants, Water withdrawals, HPI (Happy Planet Index), Religious tolerance. Além destas 9 (nove) variáveis mencionadas, extraídas da base de dados deste estudo, 2 (duas) variáveis adicionais foram criadas para auxiliar nesta análise. A variável "**3Reg_reOrd**", que representa as 3 (três) regiões AIBER, AVECO e OTHERS ordenadas pelos resultados de uma ANOVA (análise de variância), classificando os resultados dos mais baixos (piores) para os mais altos (melhores). E a variável "**ClusterM_reOrd**", que é resultado de 4 (quatro) agrupamentos por similaridade, através do método *Complete* e medida de distância *Manhattan*, classificados em ordem crescente de média.



11.2 Análise Discriminante (utilizando o *software SPSS*)

Nesta sessão será executada a função de Análise Discriminante (*discriminant analysis*), tendo como *response* (*dependent variable*) primeiramente a coluna "'ClusterM_reOrd'" e na sequência a coluna "3Reg_reOrd". As variáveis preditoras (*independent variables*) serão: 'Opportunity_n' 'Water withdrawals_np' 'Tolerance for immigrants_n' 'Religious tolerance_n' 'Governança_n' 'EPI-2014_n' 'HPI_n' 'Democracy Index-2014_n' 'IDH-2014_n'.

A análise discriminante é uma técnica da estatística multivariada utilizada para discriminar e classificar objetos. Segundo KHATTREE & NAIK (2000) é uma técnica da estatística multivariada que estuda a separação de objetos de uma população em duas ou mais classes. *fonte: apostila ANÁLISE MULTIVARIADA APLICADA AS CIÊNCIAS AGRÁRIAS - PÓS-GRADUAÇÃO EM AGRONOMIA CIÊNCIA DO SOLO: CPGA – Prof. Carlos Alberto Alves Varella, da Universidade Federal Rural do Rio de Janeiro.*

RESULTADOS: Análise Discriminante: by GROUPS=**ClusterM_reOrd**

```
GET
  FILE='C:\Users\Arvores de Classificacao.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
DISCRIMINANT
  /GROUPS=ClusterM_reOrd(1 3)
  /VARIABLES=Opportunity_n Waterwithdrawals_np Toleranceforimmigrants_n Religioustolerance_n
IDH2014_n Governança_n EPI2014_n HPI_n DemocracyIndex2014_n
  /ANALYSIS ALL
```

```
/METHOD=WILKS
/FIN=3.84
/FOUT=2.71
/PRIORS EQUAL
/HISTORY
/STATISTICS=TABLE
/CLASSIFY=NONMISSING POOLED.
```

Discriminant

Notes

Output Created		22-May-2016 15:24:13
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing in the analysis phase.
	Cases Used	In the analysis phase, cases with no user- or system-missing values for any predictor variable are used. Cases with user-, system-missing, or out-of-range values for the grouping variable are always excluded.
Syntax		DISCRIMINANT /GROUPS=ClusterM_reOrd(1 3) /VARIABLES=Opportunity_n Waterwithdrawals_np Toleranceforimmigrants_n Religioustolerance_n IDH2014_n Governança_n EPI2014_n HPI_n DemocracyIndex2014_n /ANALYSIS ALL /METHOD=WILKS /FIN=3.84 /FOUT=2.71 /PRIORS EQUAL /HISTORY /STATISTICS=TABLE /CLASSIFY=NONMISSING POOLED.
Resources	Processor Time	00:00:00.047

Notes

Output Created		22-May-2016 15:24:13
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing in the analysis phase.
	Cases Used	In the analysis phase, cases with no user- or system-missing values for any predictor variable are used. Cases with user-, system-missing, or out-of-range values for the grouping variable are always excluded.
Syntax		DISCRIMINANT /GROUPS=ClusterM_reOrd(1 3) /VARIABLES=Opportunity_n Waterwithdrawals_np Toleranceforimmigrants_n Religioustolerance_n IDH2014_n Governança_n EPI2014_n HPI_n DemocracyIndex2014_n /ANALYSIS ALL /METHOD=WILKS /FIN=3.84 /FOUT=2.71 /PRIORS EQUAL /HISTORY /STATISTICS=TABLE /CLASSIFY=NONMISSING POOLED.
Resources	Processor Time	00:00:00.047
	Elapsed Time	00:00:00.077

[DataSet1] C:\Users\Arvores de Classificacao.sav

Analysis Case Processing Summary

Unweighted Cases	N	Percent
Valid	98	74.2

Excluded	Missing or out-of-range group codes	34	25.8
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	34	25.8
Total		132	100.0

Group Statistics

ClusterM_reOrd		Valid N (listwise)	
		Unweighted	Weighted
1	Opportunity	36	36.000
	Water withdrawals_np	36	36.000
	Tolerance for immigrants_n	36	36.000
	Religious tolerance_n	36	36.000
	IDH-2014_n	36	36.000
	Governance	36	36.000
	EPI-2014_n	36	36.000
	HPI	36	36.000
	Democracy Index-2014_n	36	36.000
2	Opportunity	28	28.000
	Water withdrawals_np	28	28.000
	Tolerance for immigrants_n	28	28.000
	Religious tolerance_n	28	28.000
	IDH-2014_n	28	28.000
	Governance	28	28.000
	EPI-2014_n	28	28.000
	HPI	28	28.000
	Democracy Index-2014_n	28	28.000
3	Opportunity	34	34.000
	Water withdrawals_np	34	34.000
	Tolerance for immigrants_n	34	34.000
	Religious tolerance_n	34	34.000
	IDH-2014_n	34	34.000
	Governance	34	34.000
	EPI-2014_n	34	34.000
	HPI	34	34.000

	Democracy Index-2014_n	34	34.000
Total	Opportunity	98	98.000
	Water withdrawals_np	98	98.000
	Tolerance for immigrants_n	98	98.000
	Religious tolerance_n	98	98.000
	IDH-2014_n	98	98.000
	Governance	98	98.000
	EPI-2014_n	98	98.000
	HPI	98	98.000
	Democracy Index-2014_n	98	98.000

Analysis 1

Stepwise Statistics

Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda			
		Statistic	df1	df2	df3
1	IDH-2014_n	.285	1	2	95.000
2	Religious tolerance_n	.158	2	2	95.000
3	Democracy Index-2014_n	.128	3	2	95.000
4	Governance	.101	4	2	95.000
5	EPI-2014_n	.087	5	2	95.000
6	Tolerance for immigrants_n	.078	6	2	95.000
7	Water withdrawals_np	.071	7	2	95.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 18.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables Entered/Removed[a,b,c,d]

Step	Wilks' Lambda			
	Exact F			
	Statistic	df1	df2	Sig.
1	119.257	2	95.000	.000

2	71.420	4	188.000	.000
3	55.553	6	186.000	.000
4	49.257	8	184.000	.000
5	43.576	10	182.000	.000
6	38.877	12	180.000	.000
7	35.116	14	178.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 18.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	IDH-2014_n	1.000	119.257	
2	IDH-2014_n	.932	127.404	.585
	Religious tolerance_n	.932	37.989	.285
3	IDH-2014_n	.914	105.993	.421
	Religious tolerance_n	.904	37.974	.233
	Democracy Index-2014_n	.935	10.601	.158
4	IDH-2014_n	.785	92.030	.304
	Religious tolerance_n	.712	55.619	.224
	Democracy Index-2014_n	.471	25.133	.157
	Governance	.429	12.240	.128
5	IDH-2014_n	.570	11.295	.108
	Religious tolerance_n	.712	52.212	.186
	Democracy Index-2014_n	.465	22.051	.129
	Governance	.428	11.500	.109
	EPI-2014_n	.650	7.614	.101
6	IDH-2014_n	.556	10.758	.096
	Religious tolerance_n	.709	39.398	.145
	Democracy Index-2014_n	.441	25.218	.121
	Governance	.406	13.882	.101
	EPI-2014_n	.649	7.542	.091
	Tolerance for immigrants_n	.887	5.388	.087
7	IDH-2014_n	.523	6.599	.081
	Religious tolerance_n	.709	38.959	.133
	Democracy Index-2014_n	.439	23.554	.108

	Governance	.406	13.678	.092
	EPI-2014_n	.616	9.254	.085
	Tolerance for immigrants_n	.872	4.387	.078
	Water withdrawals_np	.893	4.317	.078

Variables Not in the Analysis

Step		Tolerance	Min. Tolerance	F to Enter	Wilks' Lambda
0	Opportunity	1.000	1.000	56.981	.455
	Water withdrawals_np	1.000	1.000	23.286	.671
	Tolerance for immigrants_n	1.000	1.000	14.241	.769
	Religious tolerance_n	1.000	1.000	33.762	.585
	IDH-2014_n	1.000	1.000	119.257	.285
	Governance	1.000	1.000	13.585	.778
	EPI-2014_n	1.000	1.000	116.304	.290
	HPI	1.000	1.000	13.512	.779
	Democracy Index-2014_n	1.000	1.000	16.196	.746
1	Opportunity	.931	.931	25.626	.184
	Water withdrawals_np	.976	.976	6.027	.252
	Tolerance for immigrants_n	.979	.979	9.951	.235
	Religious tolerance_n	.932	.932	37.989	.158
	Governance	.903	.903	.718	.281
	EPI-2014_n	.665	.665	11.645	.228
	HPI	.980	.980	3.094	.267
	Democracy Index-2014_n	.965	.965	10.449	.233
2	Opportunity	.867	.834	7.160	.137
	Water withdrawals_np	.976	.910	5.019	.142
	Tolerance for immigrants_n	.943	.898	2.924	.148
	Governance	.850	.809	.251	.157
	EPI-2014_n	.662	.643	9.823	.130
	HPI	.952	.906	4.478	.144
	Democracy Index-2014_n	.935	.904	10.601	.128
3	Opportunity	.431	.431	1.933	.123
	Water withdrawals_np	.959	.887	3.689	.119
	Tolerance for immigrants_n	.935	.876	3.299	.120
	Governance	.429	.429	12.240	.101
	EPI-2014_n	.650	.641	8.269	.109
	HPI	.944	.883	2.953	.121
4	Opportunity	.384	.374	3.961	.093
	Water withdrawals_np	.958	.428	3.665	.094

	Tolerance for immigrants_n	.888	.407	5.437	.091
	EPI-2014_n	.650	.428	7.614	.087
	HPI	.871	.396	1.005	.099
5	Opportunity	.382	.372	2.759	.082
	Water withdrawals_np	.908	.428	5.316	.078
	Tolerance for immigrants_n	.887	.406	5.388	.078
	HPI	.866	.395	1.261	.084
6	Opportunity	.310	.310	2.986	.073
	Water withdrawals_np	.893	.406	4.317	.071
	HPI	.851	.381	1.378	.075
7	Opportunity	.310	.310	2.895	.066
	HPI	.847	.381	1.165	.069

Wilks' Lambda

Step	Number of Variables	Lambda	df1	df2	df3
1	1	.285	1	2	95
2	2	.158	2	2	95
3	3	.128	3	2	95
4	4	.101	4	2	95
5	5	.087	5	2	95
6	6	.078	6	2	95
7	7	.071	7	2	95

Wilks' Lambda

Step	Exact F			
	Statistic	df1	df2	Sig.
1	119.257	2	95.000	.000
2	71.420	4	188.000	.000
3	55.553	6	186.000	.000
4	49.257	8	184.000	.000
5	43.576	10	182.000	.000
6	38.877	12	180.000	.000
7	35.116	14	178.000	.000

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation

1	3.945[a]	67.9	67.9	.893
2	1.862[a]	32.1	100.0	.807

a. First 2 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.071	243.789	14	.000
2	.349	96.739	6	.000

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
Water withdrawals_np	-.274	.246
Tolerance for immigrants_n	.065	.391
Religious tolerance_n	.512	.831
IDH-2014_n	.552	.076
Governance	-.347	-.861
EPI-2014_n	.582	-.120
Democracy Index-2014_n	.385	1.015

Structure Matrix

	Function	
	1	2
EPI-2014_n	.774*	-.216
IDH-2014_n	.753*	-.382
Opportunity[a]	.369*	.303
Governance	.267*	-.045
Democracy Index-2014_n	.246*	.234
HPI[a]	.110*	-.036
Religious tolerance_n	.129	.589*
Tolerance for immigrants_n	-.071	.388*
Water withdrawals_np	-.257	.351*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable
and any discriminant function
a. This variable not used in the analysis.

Functions at Group Centroids

ClusterM_reOrd	Function	
	1	2
1	-2.494	.417
2	2.088	1.567
3	.921	-1.731

Unstandardized canonical discriminant
functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed		132
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		132

Prior Probabilities for Groups

ClusterM_reOrd		Cases Used in Analysis	
	Prior	Unweighted	Weighted
1	.333	36	36.000
2	.333	28	28.000
3	.333	34	34.000
Total	1.000	98	98.000

Classification Results[a]

		ClusterM_reOrd	Predicted Group Membership			
			1	2	3	Total
Original	Count	1	33	1	2	36
		2	0	28	0	28

	3	2	0	32	34
	Ungrouped cases	0	20	14	34
%	1	91.7	2.8	5.6	100.0
	2	.0	100.0	.0	100.0
	3	5.9	.0	94.1	100.0
	Ungrouped cases	.0	58.8	41.2	100.0

a. 94.9% of original grouped cases correctly classified.

RESULTADOS: Análise Discriminante: by GROUPS=**3Reg_reOrd**

```
DISCRIMINANT
  /GROUPS=@3Reg_reOrd(1 3)
  /VARIABLES=Opportunity_n Waterwithdrawals_np Toleranceforimmigrants_n Religioustolerance_n
IDH2014_n Governança_n EPI2014_n HPI_n DemocracyIndex2014_n
  /ANALYSIS ALL
  /METHOD=WILKS
  /FIN=3.84
  /FOUT=2.71
  /PRIORS EQUAL
  /HISTORY
  /STATISTICS=TABLE
  /CLASSIFY=NONMISSING POOLED.
```

Discriminant

Notes

Output Created		22-May-2016 15:35:59
Comments		
Input	Data	C:\Users\ Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing in the analysis phase.
	Cases Used	In the analysis phase, cases with no user- or system-missing values for any predictor variable are used. Cases with user-, system-missing, or out-of-range values for the grouping variable are always excluded.

Syntax		DISCRIMINANT /GROUPS=@3Reg_reOrd(1 3) /VARIABLES=Opportunity_n Waterwithdrawals_np Toleranceforimmigrants_n Religioustolerance_n IDH2014_n Governança_n EPI2014_n HPI_n DemocracyIndex2014_n /ANALYSIS ALL /METHOD=WILKS /FIN=3.84 /FOUT=2.71 /PRIORS EQUAL /HISTORY /STATISTICS=TABLE /CLASSIFY=NONMISSING POOLED.
Resources	Processor Time	00:00:00.046
	Elapsed Time	00:00:00.128

[DataSet1] C:\Users\Arvores de Classificacao.sav

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		132	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		132	100.0

Group Statistics

3Reg_reOrd		Valid N (listwise)	
		Unweighted	Weighted
1	Opportunity	84	84.000
	Water withdrawals_np	84	84.000
	Tolerance for immigrants_n	84	84.000

	Religious tolerance_n	84	84.000
	IDH-2014_n	84	84.000
	Governance	84	84.000
	EPI-2014_n	84	84.000
	HPI	84	84.000
	Democracy Index-2014_n	84	84.000
2	Opportunity	21	21.000
	Water withdrawals_np	21	21.000
	Tolerance for immigrants_n	21	21.000
	Religious tolerance_n	21	21.000
	IDH-2014_n	21	21.000
	Governance	21	21.000
	EPI-2014_n	21	21.000
	HPI	21	21.000
	Democracy Index-2014_n	21	21.000
3	Opportunity	27	27.000
	Water withdrawals_np	27	27.000
	Tolerance for immigrants_n	27	27.000
	Religious tolerance_n	27	27.000
	IDH-2014_n	27	27.000
	Governance	27	27.000
	EPI-2014_n	27	27.000
	HPI	27	27.000
	Democracy Index-2014_n	27	27.000
Total	Opportunity	132	132.000
	Water withdrawals_np	132	132.000
	Tolerance for immigrants_n	132	132.000
	Religious tolerance_n	132	132.000
	IDH-2014_n	132	132.000
	Governance	132	132.000
	EPI-2014_n	132	132.000
	HPI	132	132.000
	Democracy Index-2014_n	132	132.000

Analysis 1

Stepwise Statistics

Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda			
		Statistic	df1	df2	df3
1	Opportunity	.382	1	2	129.000
2	Governance	.274	2	2	129.000
3	HPI	.222	3	2	129.000
4	Religious tolerance_n	.199	4	2	129.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 18.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables Entered/Removed[a,b,c,d]

Step	Wilks' Lambda			
	Exact F			
	Statistic	df1	df2	Sig.
1	104.392	2	129.000	.000
2	58.351	4	256.000	.000
3	47.451	6	254.000	.000
4	39.140	8	252.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 18.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	Opportunity	1.000	104.392	
2	Opportunity	.288	29.405	.399
	Governance	.288	25.327	.382
3	Opportunity	.288	21.490	.298
	Governance	.288	20.409	.294
	HPI	.999	14.655	.274
4	Opportunity	.272	12.934	.240
	Governance	.285	14.460	.244
	HPI	.905	19.357	.260

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	Opportunity	1.000	104.392	
2	Opportunity	.288	29.405	.399
	Governance	.288	25.327	.382
3	Opportunity	.288	21.490	.298
	Governance	.288	20.409	.294
	HPI	.999	14.655	.274
4	Opportunity	.272	12.934	.240
	Governance	.285	14.460	.244
	HPI	.905	19.357	.260
	Religious tolerance_n	.832	7.434	.222

Variables Not in the Analysis

Step		Tolerance	Min. Tolerance	F to Enter	Wilks' Lambda
0	Opportunity	1.000	1.000	104.392	.382
	Water withdrawals_np	1.000	1.000	.060	.999
	Tolerance for immigrants_n	1.000	1.000	9.739	.869
	Religious tolerance_n	1.000	1.000	9.818	.868
	IDH-2014_n	1.000	1.000	51.614	.555
	Governance	1.000	1.000	97.018	.399
	EPI-2014_n	1.000	1.000	40.281	.616
	HPI	1.000	1.000	22.688	.740
	Democracy Index-2014_n	1.000	1.000	61.189	.513
1	Water withdrawals_np	.992	.992	.519	.379
	Tolerance for immigrants_n	.914	.914	1.332	.374
	Religious tolerance_n	.927	.927	6.975	.344
	IDH-2014_n	.675	.675	1.355	.374
	Governance	.288	.288	25.327	.274
	EPI-2014_n	.674	.674	.285	.380
	HPI	1.000	1.000	19.201	.294
	Democracy Index-2014_n	.475	.475	.152	.381
2	Water withdrawals_np	.988	.286	.570	.271
	Tolerance for immigrants_n	.898	.267	.188	.273
	Religious tolerance_n	.919	.274	3.340	.260
	IDH-2014_n	.654	.274	.784	.270
	EPI-2014_n	.668	.265	.595	.271
	HPI	.999	.288	14.655	.222
	Democracy Index-2014_n	.432	.261	1.497	.267

3	Water withdrawals_np	.900	.285	2.973	.212
	Tolerance for immigrants_n	.795	.264	2.694	.213
	Religious tolerance_n	.832	.272	7.434	.199
	IDH-2014_n	.580	.273	.868	.219
	EPI-2014_n	.630	.264	.020	.222
	Democracy Index-2014_n	.421	.260	.449	.221
4	Water withdrawals_np	.880	.267	2.024	.193
	Tolerance for immigrants_n	.791	.252	1.939	.193
	IDH-2014_n	.524	.246	.111	.198
	EPI-2014_n	.599	.241	.248	.198
	Democracy Index-2014_n	.412	.241	1.012	.196

Wilks' Lambda

Step	Number of Variables	Lambda	df1	df2	df3
1	1	.382	1	2	129
2	2	.274	2	2	129
3	3	.222	3	2	129
4	4	.199	4	2	129

Wilks' Lambda

Step	Exact F			
	Statistic	df1	df2	Sig.
1	104.392	2	129.000	.000
2	58.351	4	256.000	.000
3	47.451	6	254.000	.000
4	39.140	8	252.000	.000

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.770[a]	68.5	68.5	.799
2	.815[a]	31.5	100.0	.670

a. First 2 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.199	205.939	8	.000
2	.551	76.013	3	.000

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
Opportunity	.737	.789
Religious tolerance_n	.025	.531
Governance	.256	-1.169
HPI	.273	.687

Structure Matrix

	Function	
	1	2
Opportunity	.955*	-.066
Governance	.875*	-.427
Democracy Index-2014_n[a]	.744*	-.182
IDH-2014_n[a]	.630*	-.106
EPI-2014_n[a]	.603*	-.070
Tolerance for immigrants_n[a]	.179*	-.068
Water withdrawals_np[a]	-.156*	-.124
HPI	.246	.548*
Religious tolerance_n	.188	.332*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

a. This variable not used in the analysis.

Functions at Group Centroids

3Reg_reOrd	Function	
	1	2
1	-.934	-.231
2	.789	1.981

3	2.293	-.823

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed		132
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		132

Prior Probabilities for Groups

3Reg_reOrd	Prior	Cases Used in Analysis	
		Unweighted	Weighted
1	.333	84	84.000
2	.333	21	21.000
3	.333	27	27.000
Total	1.000	132	132.000

Classification Results[a]

		3Reg_reOrd	Predicted Group Membership			Total
			1	2	3	
Original	Count	1	74	6	4	84
		2	0	17	4	21
		3	2	0	25	27
	%	1	88.1	7.1	4.8	100.0
		2	.0	81.0	19.0	100.0
		3	7.4	.0	92.6	100.0

a. 87.9% of original grouped cases correctly classified.

11.3 Resultados das análises discriminantes (Minitab e SPSS)

Nesta sessão, apresentamos uma tabela contendo o conjunto dos resultados dos testes de modelos realizados na sessão 10 e os resultados obtidos neste trabalho, onde foi utilizado o *software* SPSS, para fins de comparação.

Modelo	ClusterM_reOrd	3Reg_reOrd
AD (Análise Discriminante) **Minitab**	99.2% (Q)	95.5% (Q)
RL (Regressão Logística) **Minitab**	94.9% (O)	95.2% (O)
AD (Análise Discriminante) **SPSS**	94.9%	87.9%

legenda:
ClusterM_reOrd = Agrupamento de 4 clusters ordenados, sendo de 1 (pior) a 4 (melhor)
3Reg_reOrd = As 3 regiões ordenadas, sendo 1 para Others, 2 AIBER e 3 para AVECO
Q = metódo *quadratic* na análise discriminante
O = método *ordinal* na regressão logística
Minitab = resultados obtidos através do *software* Minitab
SPSS = resultados obtidos através do *software* SPSS

Na tabela acima vemos que o modelo de análise discriminante realizado no *software* Minitab, apresentou melhores resultados que o realizado no SPSS.

11.4 Árvores de Classificação

Nesta sessão, apresentaremos análises de árvores de classificação, ou árvores de decisão.

A árvore de decisão consiste de uma hierarquia de nós internos e externos que são conectados por ramos. O nó interno, também conhecido como nó decisório ou nó intermediário, é a unidade de tomada de decisão que avalia através de teste lógico qual será o próximo nó descendente ou filho. Em contraste, um nó externo (não tem nó descendente), também conhecido como folha ou nó terminal, está associado a um rótulo ou a um valor.

Em geral, o procedimento de uma árvore de decisão é o seguinte: apresenta-se um conjunto de dados ao nó inicial (ou nó raiz que também é um nó interno) da árvore; dependendo do resultado do teste lógico usado pelo nó, a árvore ramifica-se para um dos nós filhos e este procedimento é repetido até que um nó terminal é alcançado. A repetição deste procedimento caracteriza a recursividade da árvore de decisão. *fonte: Sistema Maxwell – PUC-Rio. http://www.maxwell.vrac.puc-rio.br/3710/3710_4.PDF*

11.4.1 Árvores de Classificação – utilizando ClusterM_reOrd

Nesta sessão, apresentaremos análises de árvores de classificação, ou árvores de decisão, tendo ClusterM_reOrd como variável dependente.

Os métodos explorados serão CHAID, Exhaustive CHAID e CRT, conforme abaixo:

Método: CHAID

* Decision Tree.

```
TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=CHAID
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO ADJUST=BONFERRONI
INTERVALS=10.
```

Classification Tree

Notes

Output Created		22-May-2016 16:30:25
Comments		
Input	Data	C:\Users\ Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.

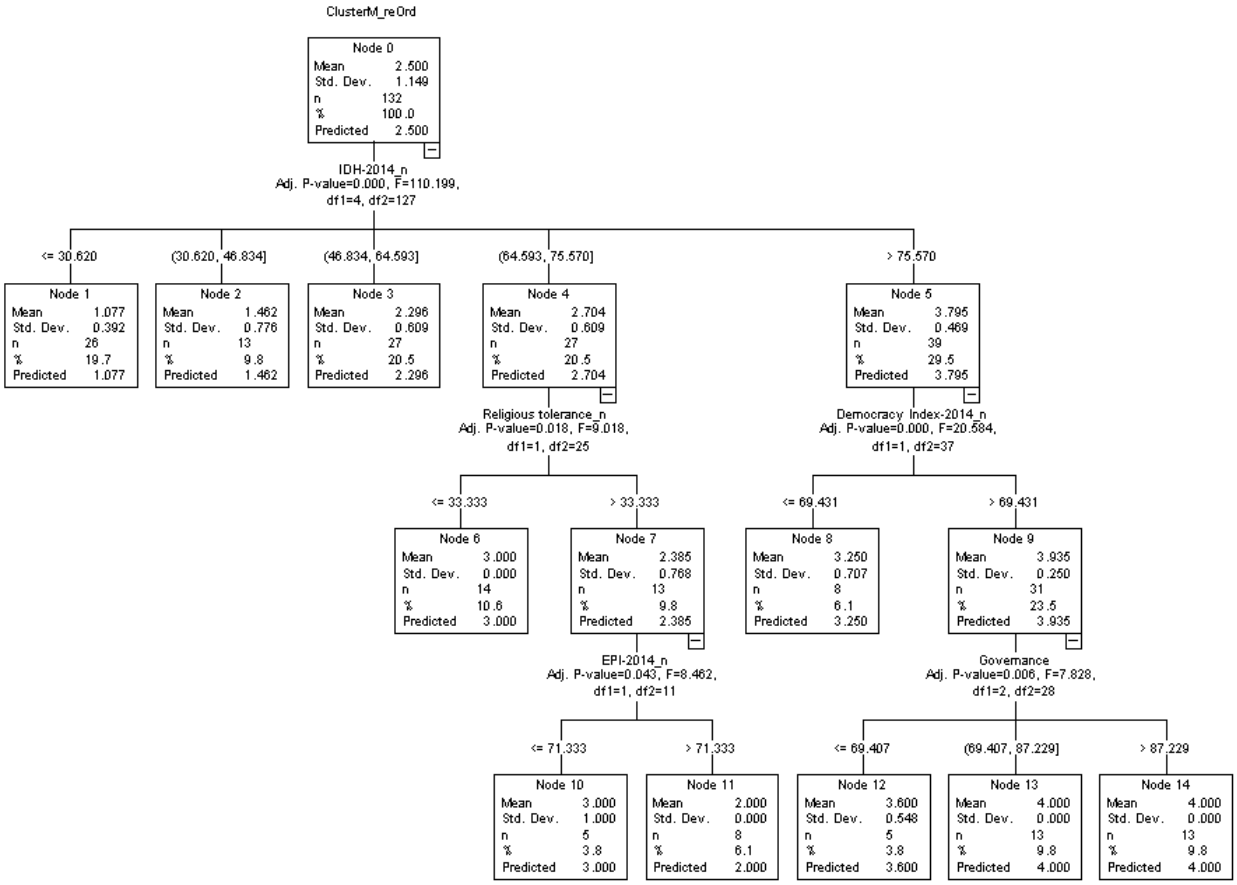
Syntax		TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=CHAID /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
Resources	Processor Time	00:00:00.515
Files Saved	Rules File	

[DataSet1] C:\Users\Arvores de Classificacao.sav

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	ClusterM_reOrd
	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5

Results	Independent Variables Included	IDH-2014_n, Religious tolerance_n, EPI-2014_n, Democracy Index-2014_n, Governance
	Number of Nodes	15
	Number of Terminal Nodes	10
	Depth	3



Gain Summary for Nodes

Node	N	Percent	Mean
13	13	9.8%	4.00
14	13	9.8%	4.00
12	5	3.8%	3.60
8	8	6.1%	3.25
6	14	10.6%	3.00
10	5	3.8%	3.00
3	27	20.5%	2.30
11	8	6.1%	2.00
2	13	9.8%	1.46
1	26	19.7%	1.08

Gain Summary for Nodes

Node	N	Percent	Mean
13	13	9.8%	4.00
14	13	9.8%	4.00
12	5	3.8%	3.60
8	8	6.1%	3.25
6	14	10.6%	3.00
10	5	3.8%	3.00
3	27	20.5%	2.30
11	8	6.1%	2.00
2	13	9.8%	1.46
1	26	19.7%	1.08

Growing Method: CHAID

Dependent Variable: ClusterM_reOrd

Risk

Estimate	Std. Error
.223	.045

Growing Method: CHAID

Dependent Variable:

ClusterM_reOrd

Método: Exhaustive CHAID

```
* Decision Tree.
TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=EXHAUSTIVECHAID
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
```

Classification Tree

Notes

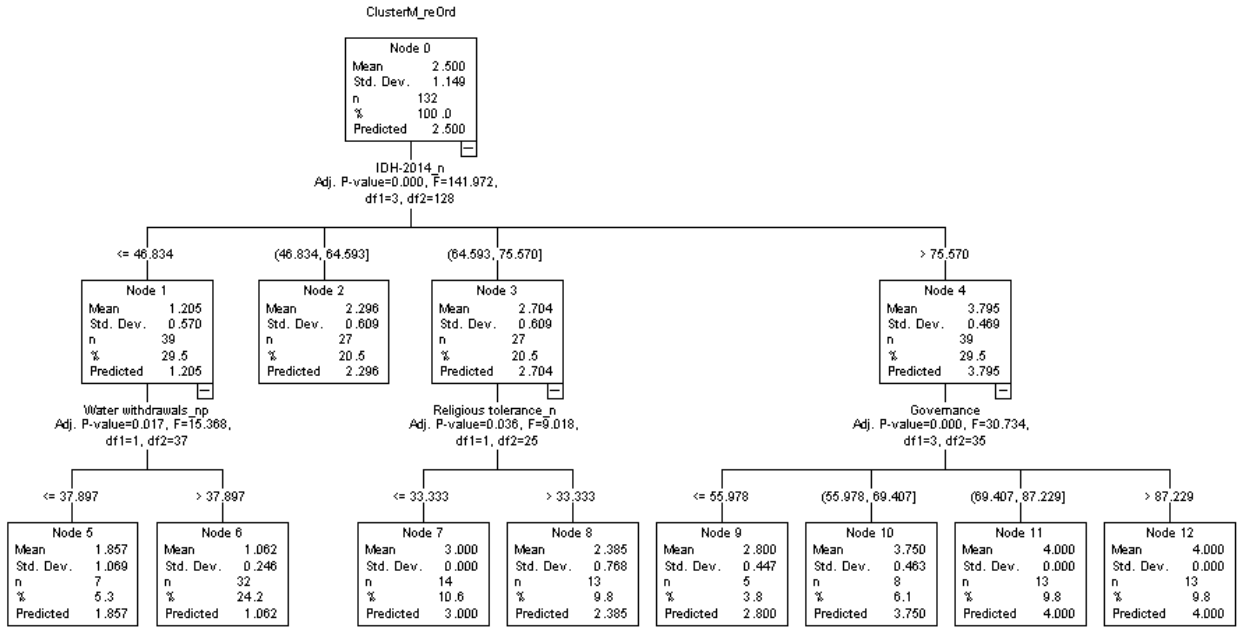
Output Created		22-May-2016 16:30:57
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav
	Active Dataset	DataSet1

	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.
Syntax		TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=EXHAUSTIVECHAID /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CHAID ALPHASPLIT=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
Resources	Processor Time	00:00:00.717
	Elapsed Time	00:00:01.986
Files Saved	Rules File	

[DataSet1] C:\Users\Arvores de Classificacao.sav

Model Summary

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	ClusterM_reOrd
	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH-2014_n, Water withdrawals_np, Religious tolerance_n, Governance
	Number of Nodes	13
	Number of Terminal Nodes	9
	Depth	2



Gain Summary for Nodes

Node	N	Percent	Mean
11	13	9.8%	4.00
12	13	9.8%	4.00
10	8	6.1%	3.75

7	14	10.6%	3.00
9	5	3.8%	2.80
8	13	9.8%	2.38
2	27	20.5%	2.30
5	7	5.3%	1.86
6	32	24.2%	1.06

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: ClusterM_reOrd

Risk

Estimate	Std. Error
.210	.039

Growing Method:

EXHAUSTIVE CHAID

Dependent Variable:

ClusterM_reOrd

Método: CRT

```
* Decision Tree.
TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CRT MINIMPROVEMENT=0.0001.
```

Classification Tree

Notes

Output Created		22-May-2016 16:31:31
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132

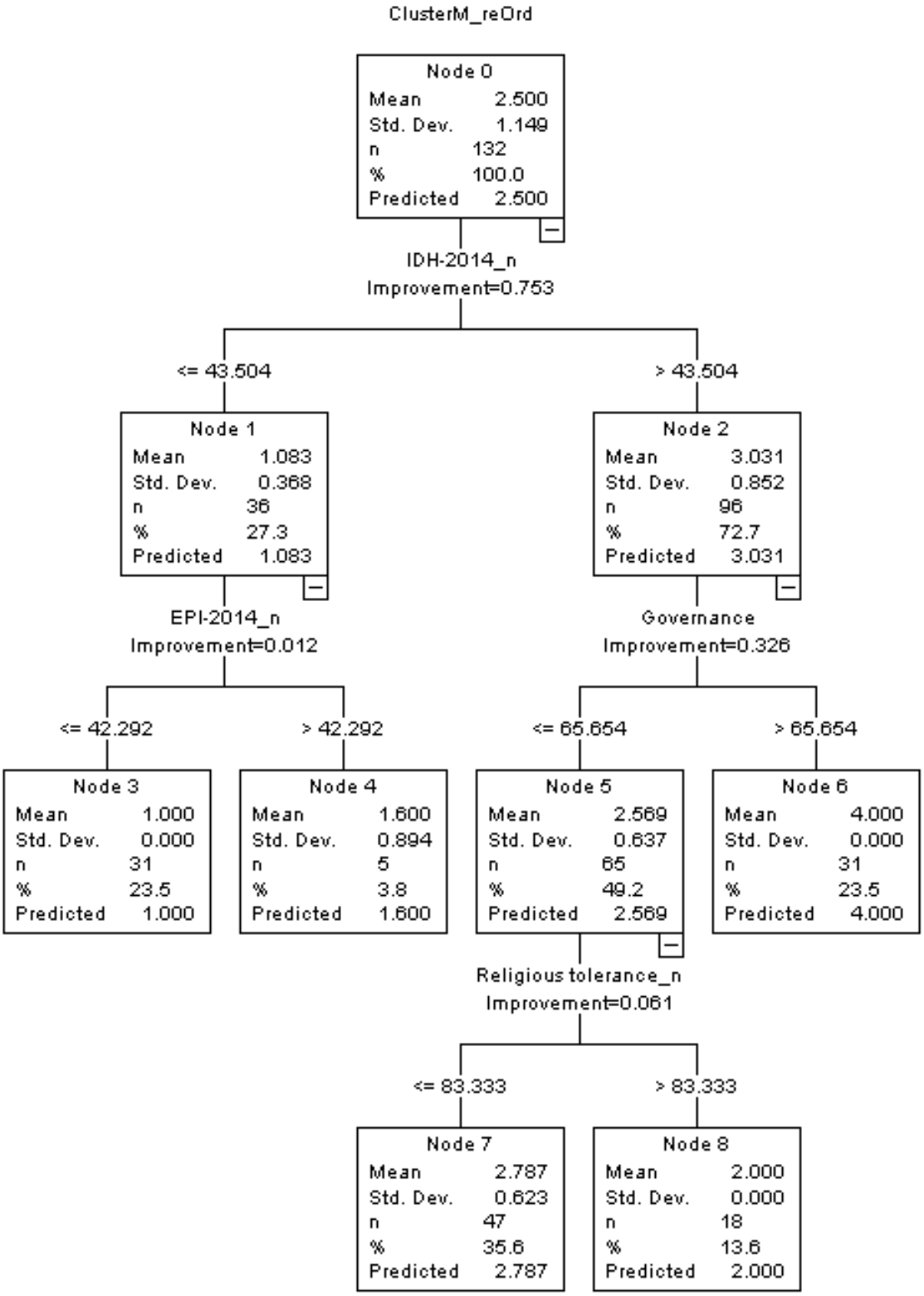
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.
Syntax		TREE ClusterM_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CRT MINIMPROVEMENT=0.0001.
Resources	Processor Time	00:00:00.531
	Elapsed Time	00:00:01.763
Files Saved	Rules File	

[DataSet1] C:\Users\Arvores de Classificacao.sav

Model Summary

Specifications	Growing Method	CRT
	Dependent Variable	ClusterM_reOrd

	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH-2014_n, EPI-2014_n, Opportunity, Water withdrawals_np, Governance, HPI, Democracy Index-2014_n, Tolerance for immigrants_n, Religious tolerance_n
	Number of Nodes	9
	Number of Terminal Nodes	5
	Depth	3


ClusterM_reOrd
Node 0
Mean 2.500
Std. Dev. 1.149
n 132
% 100.0
Predicted 2.500
IDH-2014_n
Improvement=0.753
<= 43.504
> 43.504
Node 1
Mean 1.083
Std. Dev. 0.368
n 36
% 27.3
Predicted 1.083
Node 2
Mean 3.031
Std. Dev. 0.852
n 96
% 72.7
Predicted 3.031
EPI-2014_n
Improvement=0.012
Governance
Improvement=0.326
<= 42.292
> 42.292
<= 65.654
> 65.654
Node 3
Mean 1.000
Std. Dev. 0.000
n 31
% 23.5
Predicted 1.000
Node 4
Mean 1.600
Std. Dev. 0.894
n 5
% 3.8
Predicted 1.600
Node 5
Mean 2.569
Std. Dev. 0.637
n 65
% 49.2
Predicted 2.569
Node 6
Mean 4.000
Std. Dev. 0.000
n 31
% 23.5
Predicted 4.000
Religious tolerance_n
Improvement=0.061
<= 83.333
> 83.333
Node 7
Mean 2.787
Std. Dev. 0.623
n 47
% 35.6
Predicted 2.787
Node 8
Mean 2.000
Std. Dev. 0.000
n 18
% 13.6
Predicted 2.000

Gain Summary for Nodes

Node	N	Percent	Mean
6	31	23.5%	4.00
7	47	35.6%	2.79
8	18	13.6%	2.00
4	5	3.8%	1.60
3	31	23.5%	1.00

Growing Method: CRT

Dependent Variable: ClusterM_reOrd

Risk

Estimate	Std. Error
.160	.042

Growing Method: CRT

Dependent Variable:

ClusterM_reOrd

11.4.2 Árvores de Classificação – utilizando "3Reg_reOrd"

Nesta sessão, apresentaremos análises de árvores de classificação, ou árvores de decisão, tendo 3Reg_reOrd como variável dependente. Os métodos explorados serão CHAID, Exhaustive CHAID e CRT, conforme abaixo:

Método: CHAID

```
* Decision Tree.
TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=CHAID
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO ADJUST=BONFERRONI
INTERVALS=10.
```

Classification Tree

Notes

Output Created		22-May-2016 16:32:42
Comments		
Input	Data	C:\Users\Cla\Arvores de Classificacao.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.

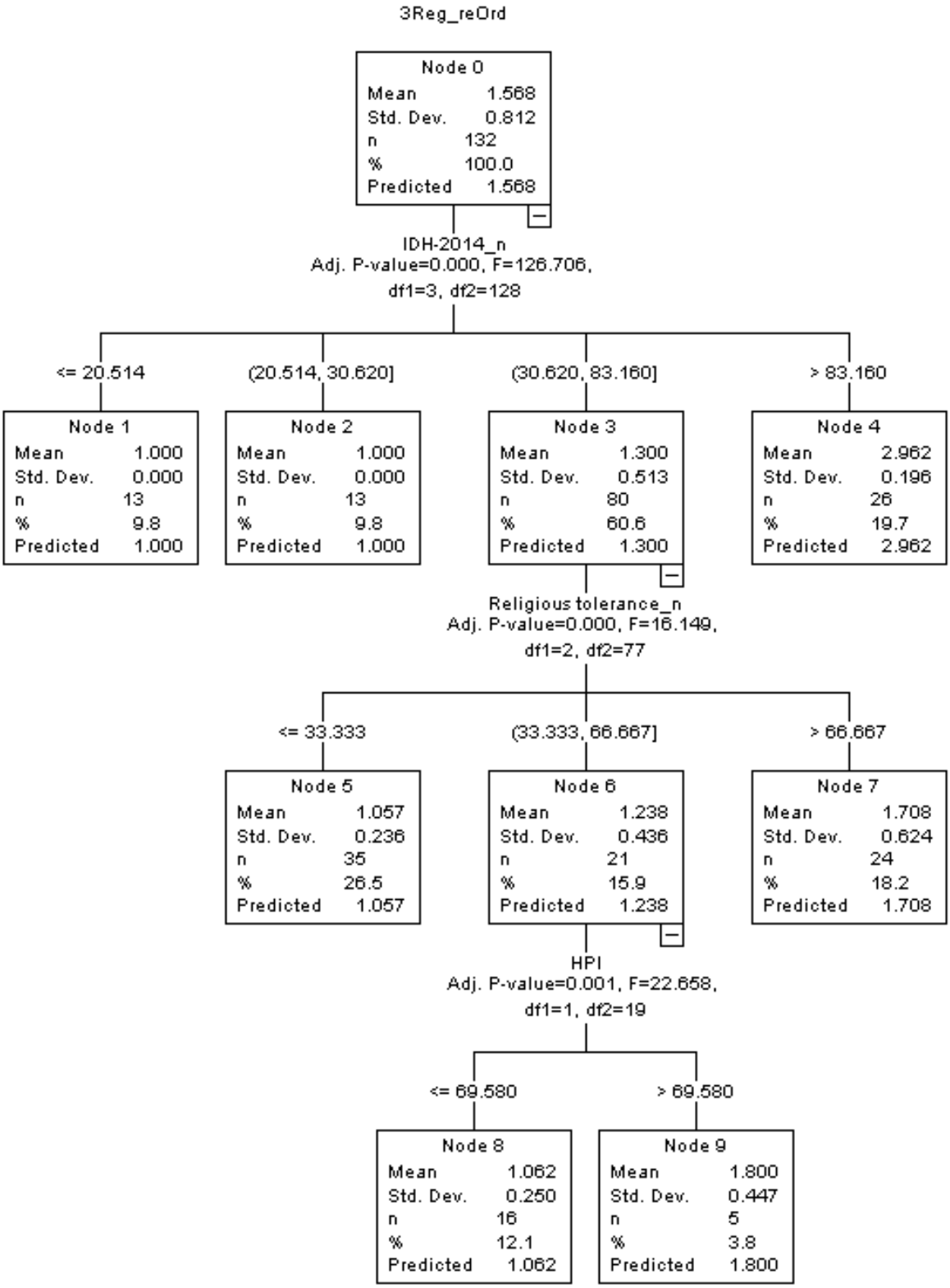
Syntax		TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=CHAID /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
Resources	Processor Time	00:00:00.593
	Elapsed Time	00:00:01.580
Files Saved	Rules File	

[DataSet1] C:\Users\Arvores de Classificacao.sav

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	3Reg_reOrd
	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10

	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH-2014_n, Religious tolerance_n, HPI
	Number of Nodes	10
	Number of Terminal Nodes	7
	Depth	3


3Reg_reOrd
Node 0
Mean 1.568
Std. Dev. 0.812
n 132
% 100.0
Predicted 1.568
IDH-2014_n
Adj. P-value=0.000, F=126.706,
df1=3, df2=128
<= 20.514
(20.514, 30.620]
(30.620, 83.160]
> 83.160
Node 1
Mean 1.000
Std. Dev. 0.000
n 13
% 9.8
Predicted 1.000
Node 2
Mean 1.000
Std. Dev. 0.000
n 13
% 9.8
Predicted 1.000
Node 3
Mean 1.300
Std. Dev. 0.513
n 80
% 60.6
Predicted 1.300
Node 4
Mean 2.962
Std. Dev. 0.196
n 26
% 19.7
Predicted 2.962
Religious tolerance_n
Adj. P-value=0.000, F=16.149,
df1=2, df2=77
<= 33.333
(33.333, 66.667]
> 66.667
Node 5
Mean 1.057
Std. Dev. 0.236
n 35
% 26.5
Predicted 1.057
Node 6
Mean 1.238
Std. Dev. 0.436
n 21
% 15.9
Predicted 1.238
Node 7
Mean 1.708
Std. Dev. 0.624
n 24
% 18.2
Predicted 1.708
HPI
Adj. P-value=0.001, F=22.658,
df1=1, df2=19
<= 69.580
> 69.580
Node 8
Mean 1.062
Std. Dev. 0.250
n 16
% 12.1
Predicted 1.062
Node 9
Mean 1.800
Std. Dev. 0.447
n 5
% 3.8
Predicted 1.800

Gain Summary for Nodes

Node	N	Percent	Mean
4	26	19.7%	2.96
9	5	3.8%	1.80
7	24	18.2%	1.71
8	16	12.1%	1.06
5	35	26.5%	1.06
1	13	9.8%	1.00
2	13	9.8%	1.00

Growing Method: CHAID

Dependent Variable: 3Reg_reOrd

Risk

Estimate	Std. Error
.103	.024

Growing Method: CHAID

Dependent Variable:

3Reg_reOrd

Método: Exhaustive CHAID

```
* Decision Tree.
TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=EXHAUSTIVECHAID
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
```

Classification Tree

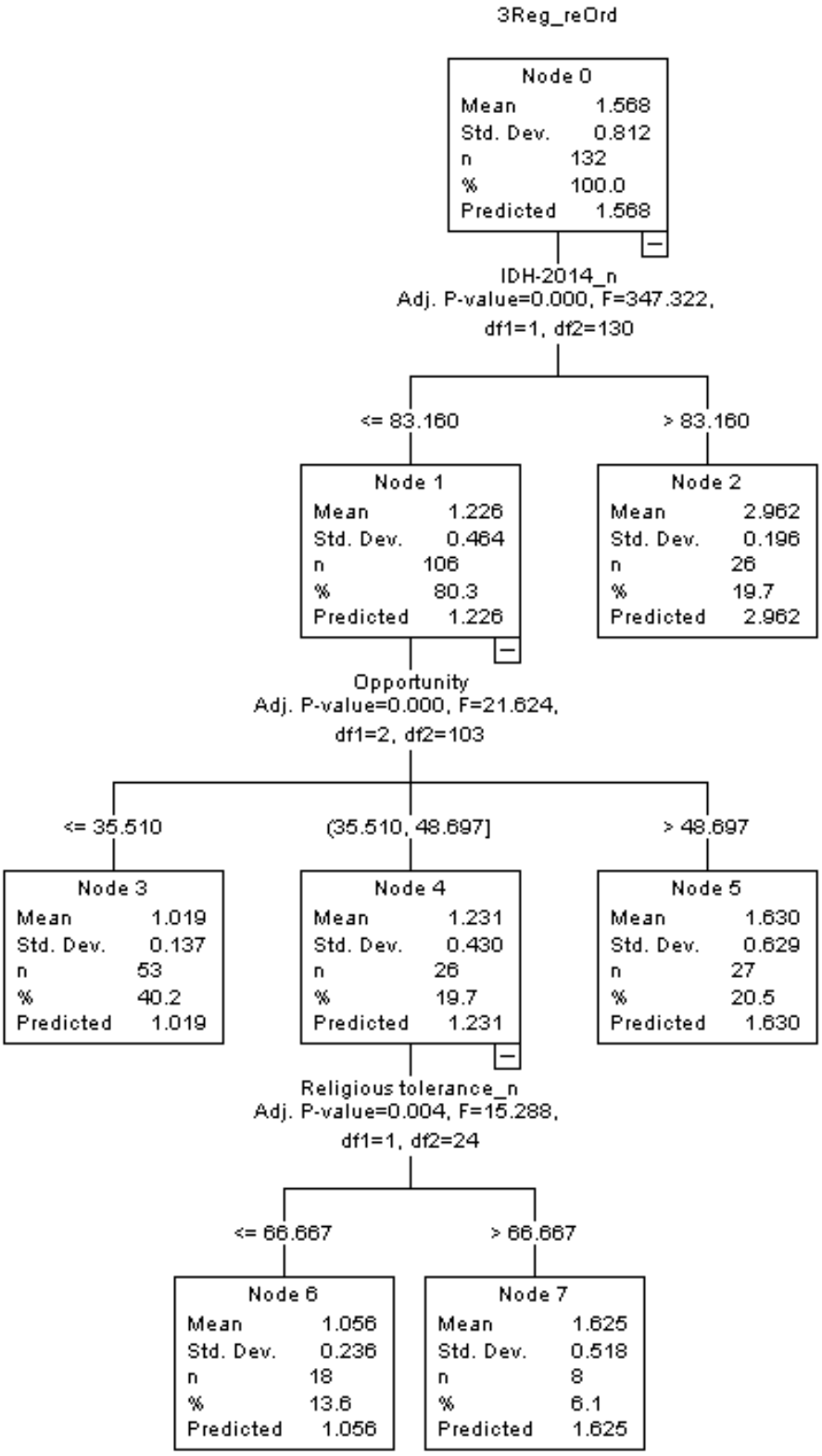
Notes

Output Created		22-May-2016 16:33:04
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav

	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.
Syntax		TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=EXHAUSTIVECHAID /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CHAID ALPHASPLIT=0.05 SPLITMERGED=NO ADJUST=BONFERRONI INTERVALS=10.
Resources	Processor Time	00:00:00.562
	Elapsed Time	00:00:01.556
Files Saved	Rules File	

Model Summary

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	3Reg_reOrd
	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH-2014_n, Opportunity, Religious tolerance_n
	Number of Nodes	8
	Number of Terminal Nodes	5
	Depth	3


3Reg_reOrd
Node 0
Mean 1.568
Std. Dev. 0.812
n 132
% 100.0
Predicted 1.568
IDH-2014_n
Adj. P-value=0.000, F=347.322,
df1=1, df2=130
<= 83.160
> 83.160
Node 1
Mean 1.226
Std. Dev. 0.464
n 106
% 80.3
Predicted 1.226
Node 2
Mean 2.962
Std. Dev. 0.196
n 26
% 19.7
Predicted 2.962
Opportunity
Adj. P-value=0.000, F=21.624,
df1=2, df2=103
<= 35.510
(35.510, 48.697]
> 48.697
Node 3
Mean 1.019
Std. Dev. 0.137
n 53
% 40.2
Predicted 1.019
Node 4
Mean 1.231
Std. Dev. 0.430
n 26
% 19.7
Predicted 1.231
Node 5
Mean 1.630
Std. Dev. 0.629
n 27
% 20.5
Predicted 1.630
Religious tolerance_n
Adj. P-value=0.004, F=15.288,
df1=1, df2=24
<= 66.667
> 66.667
Node 6
Mean 1.056
Std. Dev. 0.236
n 18
% 13.6
Predicted 1.056
Node 7
Mean 1.625
Std. Dev. 0.518
n 8
% 6.1
Predicted 1.625

Gain Summary for Nodes

Node	N	Percent	Mean
2	26	19.7%	2.96
5	27	20.5%	1.63
7	8	6.1%	1.63
6	18	13.6%	1.06
3	53	40.2%	1.02

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: 3Reg_reOrd

Risk

Estimate	Std. Error
.114	.025

Growing Method:

EXHAUSTIVE CHAID

Dependent Variable:

3Reg_reOrd

Método: CRT

```
* Decision Tree.
TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s]
Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES
SCALE=AUTO
  /PRINT MODELSUMMARY RISK
  /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE
  /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CRT MINIMPROVEMENT=0.0001.
```

Classification Tree

Notes

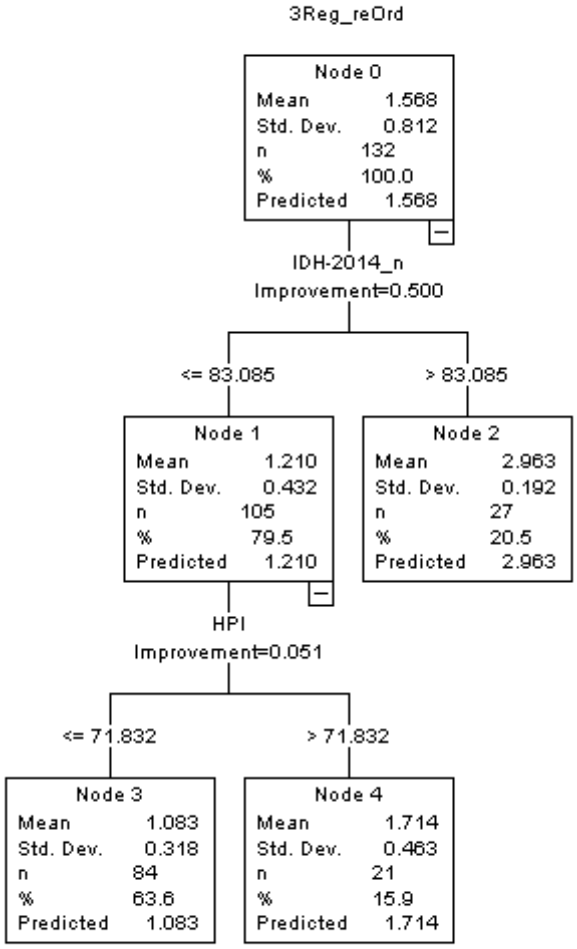
Output Created		22-May-2016 16:33:27
Comments		
Input	Data	C:\Users\Arvores de Classificacao.sav

	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	132
Missing Value Handling	Definition of Missing	Handling of user-defined missing values of nominal independent variables depends on the growing method.
	Cases Used	Only cases with valid data for the dependent variable and some or all independent variables are used in computing any statistics.
Syntax		TREE @3Reg_reOrd [s] BY Opportunity_n [s] Waterwithdrawals_np [s] Toleranceforimmigrants_n [s] Religioustolerance_n [s] IDH2014_n [s] Governança_n [s] EPI2014_n [s] HPI_n [s] DemocracyIndex2014_n [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /PRINT MODELSUMMARY RISK /GAIN SUMMARYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE /GROWTHLIMIT MAXDEPTH=5 MINPARENTSIZE=10 MINCHILDSIZE=5 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CRT MINIMPROVEMENT=0.0001.
Resources	Processor Time	00:00:00.624
	Elapsed Time	00:00:01.442
Files Saved	Rules File	

[DataSet1] C:\Users\Arvores de Classificacao.sav

Model Summary

Specifications	Growing Method	CRT
	Dependent Variable	3Reg_reOrd
	Independent Variables	Opportunity, Water withdrawals_np, Tolerance for immigrants_n, Religious tolerance_n, IDH-2014_n, Governance, EPI-2014_n, HPI, Democracy Index-2014_n
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	IDH-2014_n, Opportunity, Governance, Democracy Index-2014_n, EPI-2014_n, Tolerance for immigrants_n, HPI
	Number of Nodes	5
	Number of Terminal Nodes	3
	Depth	2



Gain Summary for Nodes

Node	N	Percent	Mean
2	27	20.5%	2.96
4	21	15.9%	1.71
3	84	63.6%	1.08

Growing Method: CRT

Dependent Variable: 3Reg_reOrd

Risk

Estimate	Std. Error
.104	.032

Growing Method: CRT

Dependent Variable:

3Reg_reOrd

Analisando os resultados acima, vemos que o método CHAID apresenta o melhor resultado no crescimento das árvores (ramificação), resultando em 15 nós para a variável "ClusterM_reOrd", e 10 nós para "3Reg_reOrd".

11.5. CONSIDERAÇÕES

As análises discriminantes efetuadas no *software* Minitab, apresentaram maior grau de acerto que no SPSS, conforme mostra tabela comparativa na sessão anterior.

Com relação às árvores de classificação, o método CHAID apresentou o melhor resultado no crescimento das árvores (ramificação), resultando em 15 nós para a variável "ClusterM_reOrd", e 10 nós para "3Reg_reOrd", conforme resultados nas sessões 3.2.1 e 3.2.2.

12. ANÁLISE DE CORRESPONDÊNCIA

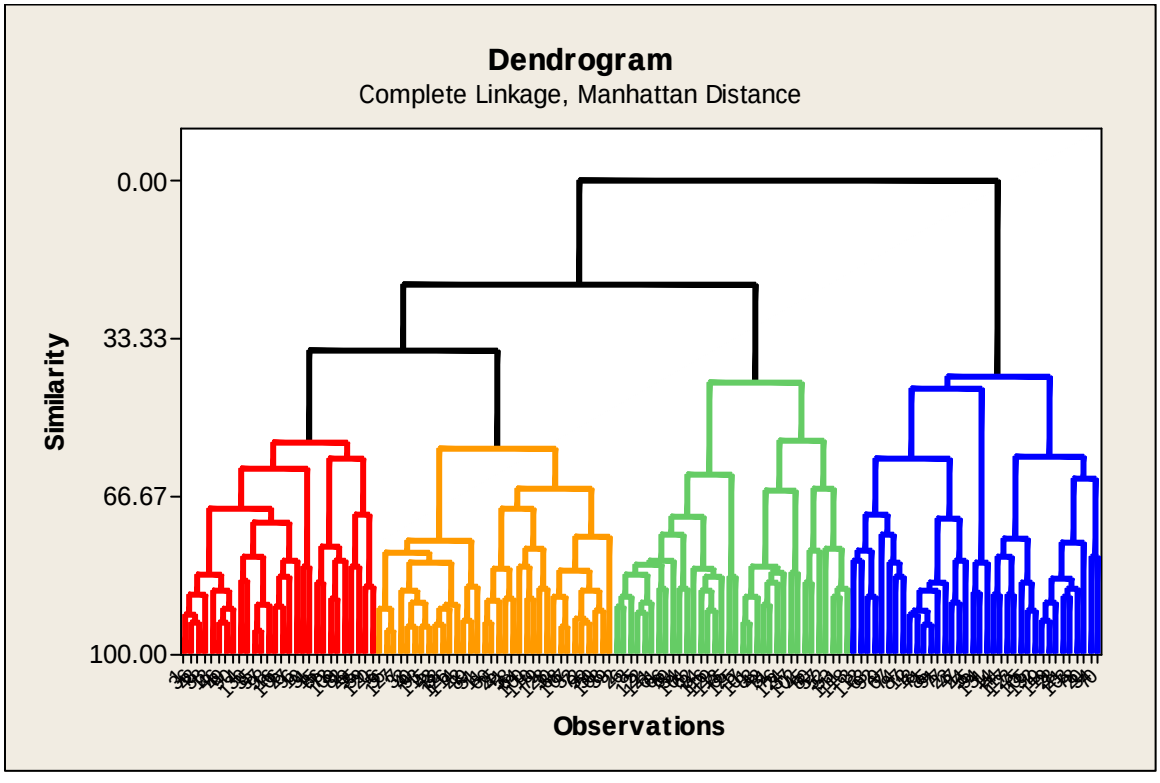
O objetivo de realizar análises de correspondência é obter graficamente a distribuição dos resultados das variáveis e analisar o grau de relacionamento entre as variáveis de interesse entre si, com as regiões assignadas, e com os agrupamentos de *clusters* selecionados. As análises serão efetuadas sobre as médias e os desvios-padrão de cada variável selecionada. Adicionalmente serão apresentados os coeficientes de determinação (r-quadrado) para auxiliar no entendimento da representatividade de cada variável e seu poder explicativo no modelo.

Obs: Conforme estabelecido, os países listados nesta tabela que não pertençam às regiões AIBER e AVECO, serão agrupados e classificados como OTHERS nas análises e resultados das sessões seguintes.

12.1. TRATAMENTO DOS DADOS, ANÁLISES DE CORRESPONDÊNCIA, AGRUPAMENTOS E RELACIONAMENTOS

Nesta sessão serão realizadas análises de correspondência dos dados dimensionadores do eixo temático Biodiversidade – Capital Natural e Social. Para apoiar estas análises, serão apresentados também os resultados em forma de gráfico radar relacionando os indicadores (variáveis) com as regiões e os agrupamentos de *clusters* pelas médias, dispersão (desvios-padrão) e grau de explicação dos modelos (r-quadrado).

As variáveis de trabalho para este estudo, conforme detalhadas na sessão anterior (quadro 1), são: EPI-2014, Governança, Democracy Index-2014, Opportunity, IDH-2014, Tolerance for immigrants, Water withdrawals, HPI (Happy Planet Index) e Religious tolerance. Além destas 9 (nove) variáveis, extraídas da base de dados deste estudo, 2 (duas) variáveis adicionais foram criadas para auxiliar nas análises. A variável "**3Reg_reOrd**", que representa as 3 (três) regiões AIBER, AVECO e OTHERS ordenadas pelos resultados de uma ANOVA (análise de variância), classificando os resultados dos mais baixos (piores) para os mais altos (melhores). E a variável "**ClusterM_reOrd**", que é resultado de 4 (quatro) agrupamentos por similaridade, através do método *Complete* e medida de distância *Manhattan*, classificados em ordem crescente de média.



12.2 Análises de correspondência

Nesta sessão será executada a função de Análise de correspondência (*simple correspondence analysis*) da ferramenta Minitab, relacionando primeiramente as regiões e os agrupamentos (*Row Plot*), para em seguida mostrar os relacionamentos das regiões, agrupamentos e as 9 (nove) variáveis deste estudo (*Symmetric Plot*).

A análise de correspondência é um método de análise fatorial para variáveis categóricas. Basicamente, converte uma tabela de dados não negativos de duas ou múltiplas entradas em um tipo de representação gráfica em que as linhas e as colunas são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico. Este método permite mostrar como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe.

As análises foram efetuadas para cruzamento e conhecer a correspondência, de média e desvio padrão, entre as 9 variáveis selecionadas, as 3 regiões e os 4 clusters.

12.2.1 Análise de Correspondência das Médias

Nas análises a seguir, algumas variáveis tiveram seus nomes reduzidos para melhor acomodação nos gráficos e tabelas. A variável *Opportunity*, tornou-se “opp”, *Water withdrawals* = “waterw”, *Tolerance for immigrants* = “immi”, *Religious tolerance* = “relig”, IDH = “idh”, *Governance* = “gov”, EPI = “epi”, HPI = “hpi” e *Democracy* = “democ”.

Simple Correspondence Analysis: opp, water, immi, relig, idh, gov, epi, hpi, democ

Contingency Table

	opp	waterw	immig	relig	idh	gov	epi	hpi
others	32.450	60.640	45.710	51.590	47.760	33.360	48.770	41.790
aiber	55.280	62.780	61.970	84.130	65.670	42.480	71.050	72.360
aveco	79.940	62.440	62.860	64.200	91.890	83.250	90.290	50.580
g1	21.720	81.500	53.920	53.700	24.390	23.690	23.900	34.500
g2	50.490	60.290	56.120	88.100	61.470	38.840	68.190	61.400
g3	34.930	36.760	34.370	31.370	66.130	36.730	65.620	52.780
g4	78.290	65.470	63.460	69.610	88.960	80.980	88.980	48.250
Total	353.100	429.880	378.410	442.700	446.270	339.330	456.800	361.660

	democ	Total
others	40.110	402.180
aiber	59.980	575.700
aveco	82.480	667.930
g1	33.030	350.350
g2	57.260	542.160
g3	38.900	397.590
g4	80.590	664.590
Total	392.350	3600.500

Chi-Square Distances

	opp	waterw	immig	relig	idh	gov	epi	hpi	democ
others	1.239	3.318	0.280	0.093	0.088	0.545	0.100	0.048	0.315
aiber	0.025	0.516	0.035	2.516	0.453	2.556	0.054	3.652	0.121
aveco	3.182	3.756	0.767	3.913	1.001	6.547	0.363	4.064	1.291
g1	4.649	37.622	7.940	2.619	8.344	2.636	9.500	0.014	0.694
g2	0.135	0.305	0.013	6.895	0.488	2.940	0.005	0.885	0.056
g3	0.423	2.416	1.316	6.276	5.761	0.015	4.567	4.130	0.452
g4	2.639	2.427	0.584	1.793	0.527	5.373	0.258	5.130	0.921
Total	12.291	50.360	10.936	24.104	16.662	20.611	14.847	17.923	3.851

	Total
others	6.025
aiber	9.929
aveco	24.884
g1	74.018
g2	11.722
g3	25.357
g4	19.653
Total	171.586

Relative Inertias

	opp	waterw	immig	relig	idh	gov	epi	hpi	democ	Total
others	0.007	0.019	0.002	0.001	0.001	0.003	0.001	0.000	0.002	0.035
aiber	0.000	0.003	0.000	0.015	0.003	0.015	0.000	0.021	0.001	0.058
aveco	0.019	0.022	0.004	0.023	0.006	0.038	0.002	0.024	0.008	0.145
g1	0.027	0.219	0.046	0.015	0.049	0.015	0.055	0.000	0.004	0.431
g2	0.001	0.002	0.000	0.040	0.003	0.017	0.000	0.005	0.000	0.068
g3	0.002	0.014	0.008	0.037	0.034	0.000	0.027	0.024	0.003	0.148
g4	0.015	0.014	0.003	0.010	0.003	0.031	0.002	0.030	0.005	0.115
Total	0.072	0.293	0.064	0.140	0.097	0.120	0.087	0.104	0.022	1.000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0.0331	0.6951	0.6951	******************************
2	0.0097	0.2037	0.8987	********
3	0.0045	0.0949	0.9937	****
4	0.0003	0.0063	0.9999	
5	0.0000	0.0001	1.0000	
Total	0.0477			

Row Contributions

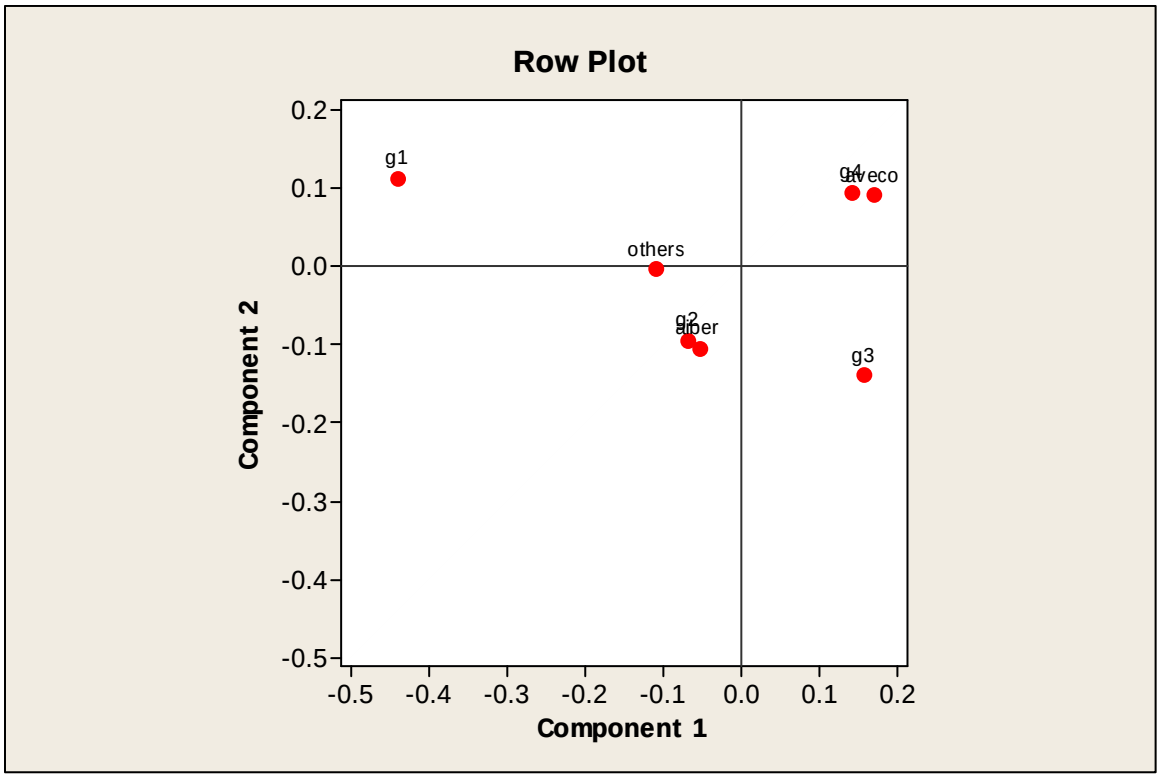
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	others	0.806	0.112	0.035	-0.110	0.806	0.041	-0.002	0.000	0.000
2	aiber	0.802	0.160	0.058	-0.055	0.176	0.015	-0.104	0.625	0.178
3	aveco	0.996	0.186	0.145	0.169	0.765	0.160	0.093	0.231	0.164
4	g1	0.984	0.097	0.431	-0.441	0.921	0.572	0.115	0.063	0.134
5	g2	0.625	0.151	0.068	-0.070	0.224	0.022	-0.093	0.401	0.134
6	g3	0.675	0.110	0.148	0.155	0.377	0.080	-0.138	0.298	0.216
7	g4	0.983	0.185	0.115	0.141	0.674	0.111	0.096	0.309	0.174

Column Contributions

					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	opp	0.834	0.098	0.072	0.160	0.731	0.075	0.060	0.103	0.036
2	waterw	0.938	0.119	0.293	-0.317	0.857	0.362	0.097	0.081	0.117
3	immig	0.986	0.105	0.064	-0.163	0.920	0.084	0.044	0.066	0.021
4	relig	0.667	0.123	0.140	-0.181	0.603	0.122	-0.059	0.064	0.044
5	idh	0.895	0.124	0.097	0.178	0.844	0.118	-0.044	0.051	0.024
6	gov	0.997	0.094	0.120	0.195	0.624	0.108	0.151	0.374	0.220
7	epi	0.960	0.127	0.087	0.161	0.801	0.100	-0.072	0.159	0.068
8	hpi	0.945	0.100	0.104	-0.069	0.096	0.014	-0.205	0.849	0.435

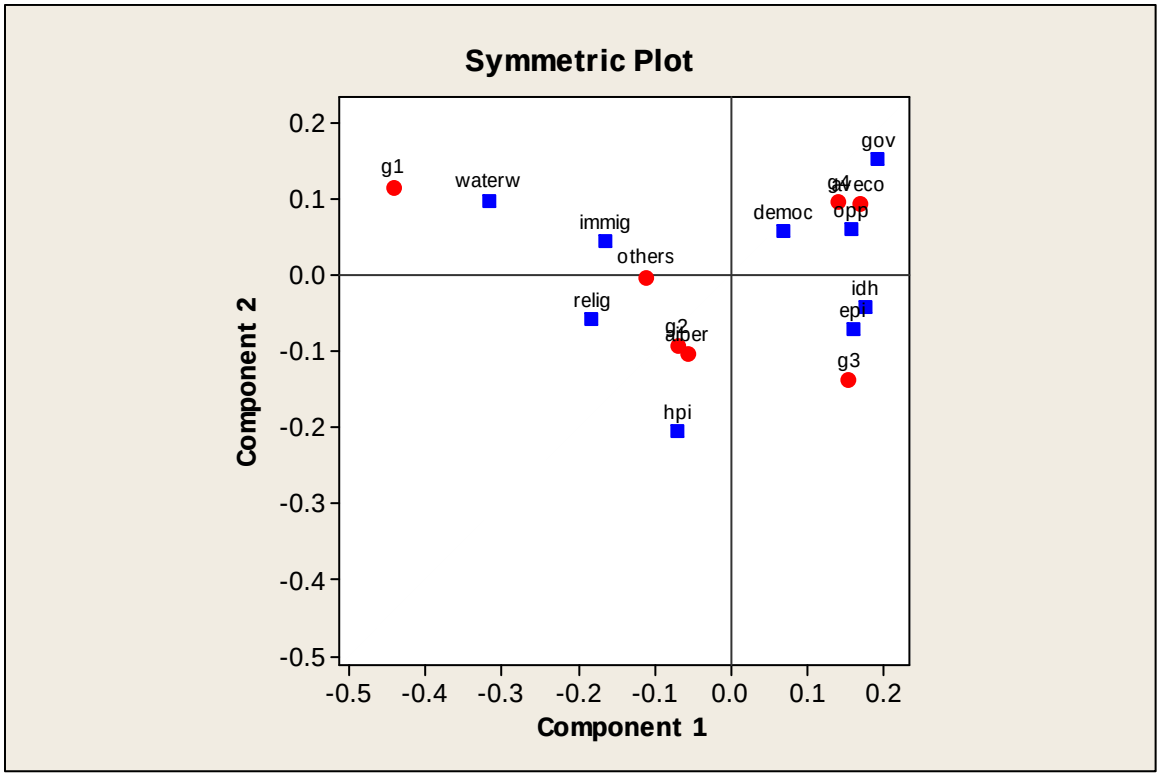
9	democ	0.830	0.109	0.022	0.071	0.512	0.017	0.056	0.318	0.035

Gráfico das médias, relacionando as 3 regiões e os 4 agrupamentos - Row Plot



O gráfico acima mostra a grande proximidade entre a região AVECO e o agrupamento G4, denotando forte relação entre estes dois grupos. Situação similar acontece entre AIBER e G2, apesar de não figurar positivamente no gráfico. E por fim, a região OTHERS que aparece num ponto médio entre G1 e G3. As proximidades e relações mostradas no gráfico acima, era de certa forma esperada já que quase a totalidade dos países de AVECO, foram categorizados no agrupamento G4, o mesmo com AIBER em relação ao G2 e os países da região OTHERS que foram distribuídos entre os G1 e G3, e figura num ponto médio entre estes dois, tanto no eixo horizontal, como no vertical.

Gráfico das médias, relacionando as 3 regiões, os 4 agrupamentos e as 9 variáveis - Symmetric Plot



Quando adicionamos as variáveis no gráfico, AIBER se destaca pela proximidade com os indicadores relig (tolerância às diferentes crenças religiosas) e hpi (índice de felicidade do planeta). Enquanto a região AVECO demonstra forte relação com os indicadores opp (oportunidades), democ (índice de democracia) e gov (governança), no quadrante positivo do gráfico . Pela grande dispersão de OTHERS no gráfico, os países desta região classificados como G1, se destacam negativamente pelo indicador waterw (retirada de água), enquanto G3, positivamente com idh (índice de desenvolvimento humano) e epi (índice de desempenho ambiental) em relação ao *component 1*.

12.2.2 Análise de Correspondência dos Desvios-Padrão

Nesta sessão, o processo será repetido utilizando as mesmas variáveis e metodologias da sessão anterior, porém utilizando os resultados de seus desvios-padrão para evidenciar distanciamentos em relação às médias. Quanto maior o desvio padrão, maior a dispersão e mais afastados da média estarão os eventos extremos.

Simple Correspondence Analysis: opp, water, immi, relig, idh, gov, epi, hpi, de

Contingency Table

	opp	waterw	immig	relig	idh	gov	epi	hpi
others	15.710	33.860	20.910	33.300	23.840	16.350	26.000	20.680
aiber	14.520	25.680	16.250	22.650	12.450	18.820	11.910	16.250
aveco	14.090	26.280	24.270	27.620	4.780	13.380	8.040	12.840
g1	11.650	27.200	20.380	35.880	13.480	12.090	15.370	19.040
g2	9.760	28.150	15.090	18.620	9.980	12.490	9.080	28.320
g3	10.380	27.050	18.040	21.620	12.710	14.030	14.190	16.560
g4	13.390	24.120	23.630	22.270	7.150	12.730	9.540	12.500
Total	89.500	192.340	138.570	181.960	84.390	99.890	94.130	126.190

	democ	Total

others	20.430	211.080
aiber	13.590	152.120
aveco	8.870	140.170
g1	16.310	171.400
g2	12.650	144.140
g3	21.270	155.850
g4	9.160	134.490
Total	102.280	1109.250

Chi-Square Distances

	opp	waterw	immig	relig	idh	gov	epi	hpi	democ	Total
others	0.102	0.205	1.130	0.051	3.771	0.372	3.652	0.463	0.048	9.793
aiber	0.411	0.018	0.399	0.213	0.066	1.915	0.077	0.064	0.014	3.177
aveco	0.684	0.160	2.609	0.931	3.247	0.045	1.249	0.605	1.272	10.803
g1	0.343	0.214	0.050	2.144	0.015	0.725	0.047	0.011	0.016	3.564
g2	0.301	0.399	0.472	1.068	0.089	0.019	0.812	8.669	0.031	11.858
g3	0.383	0.000	0.105	0.609	0.061	0.000	0.070	0.077	3.313	4.619
g4	0.594	0.027	2.776	0.002	0.928	0.032	0.307	0.512	0.847	6.026
Total	2.818	1.024	7.541	5.017	8.177	3.107	6.215	10.401	5.540	49.840

Relative Inertias

	opp	waterw	immig	relig	idh	gov	epi	hpi	democ	Total
others	0.002	0.004	0.023	0.001	0.076	0.007	0.073	0.009	0.001	0.196
aiber	0.008	0.000	0.008	0.004	0.001	0.038	0.002	0.001	0.000	0.064
aveco	0.014	0.003	0.052	0.019	0.065	0.001	0.025	0.012	0.026	0.217
g1	0.007	0.004	0.001	0.043	0.000	0.015	0.001	0.000	0.000	0.072
g2	0.006	0.008	0.009	0.021	0.002	0.000	0.016	0.174	0.001	0.238
g3	0.008	0.000	0.002	0.012	0.001	0.000	0.001	0.002	0.066	0.093
g4	0.012	0.001	0.056	0.000	0.019	0.001	0.006	0.010	0.017	0.121
Total	0.057	0.021	0.151	0.101	0.164	0.062	0.125	0.209	0.111	1.000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0.0222	0.4932	0.4932	**********
2	0.0130	0.2903	0.7836	******
3	0.0049	0.1098	0.8934	**
4	0.0031	0.0696	0.9630	*
5	0.0016	0.0348	0.9978	*
6	0.0001	0.0022	1.0000	
Total	0.0449			

Row Contributions

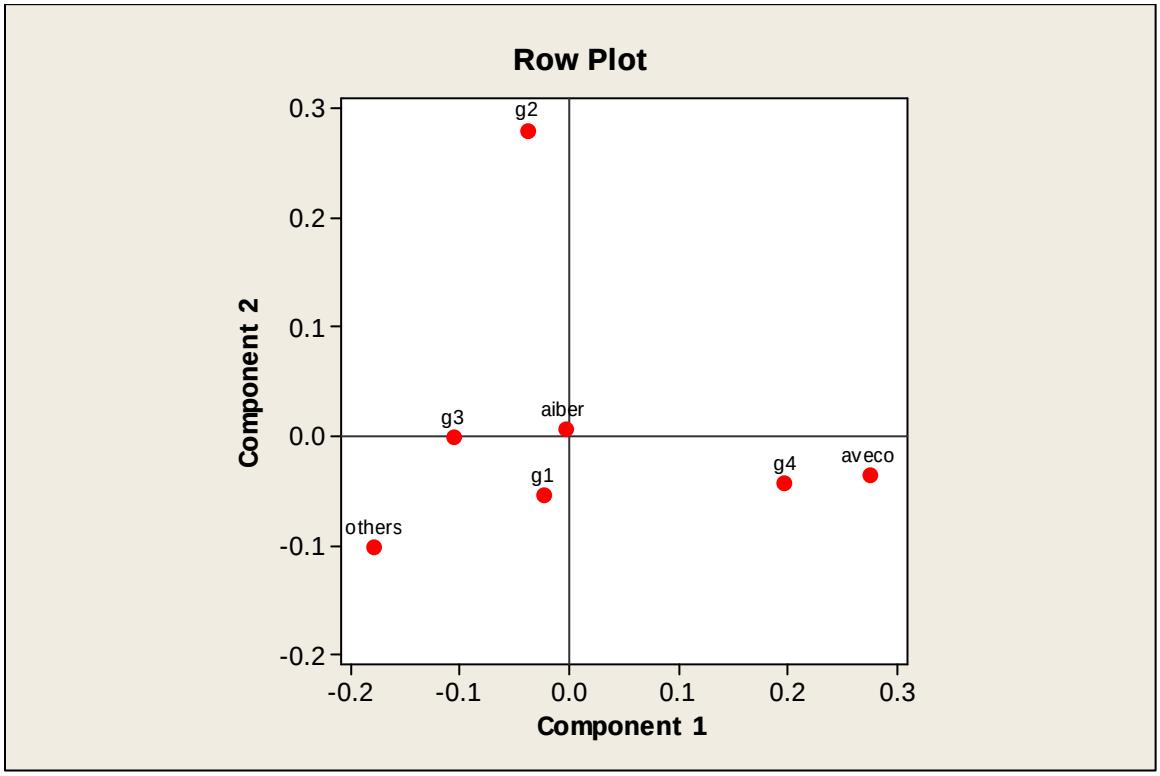
						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	others	0.915	0.190	0.196	-0.179	0.693	0.276	-0.101	0.222	0.150
2	aiber	0.003	0.137	0.064	-0.004	0.001	0.000	0.007	0.002	0.001
3	aveco	0.990	0.126	0.217	0.274	0.973	0.428	-0.036	0.017	0.012
4	g1	0.168	0.155	0.072	-0.024	0.028	0.004	-0.054	0.140	0.035
5	g2	0.977	0.130	0.238	-0.039	0.018	0.009	0.281	0.958	0.785
6	g3	0.385	0.141	0.093	-0.107	0.385	0.072	-0.001	0.000	0.000
7	g4	0.901	0.121	0.121	0.196	0.861	0.211	-0.042	0.040	0.017

Column Contributions

						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	opp	0.676	0.081	0.057	0.139	0.611	0.070	-0.045	0.065	0.013
2	waterw	0.843	0.173	0.021	0.039	0.291	0.012	0.054	0.552	0.039
3	immig	0.906	0.125	0.151	0.217	0.864	0.265	-0.048	0.042	0.022
4	relig	0.493	0.164	0.101	0.073	0.195	0.040	-0.091	0.298	0.103
5	idh	0.946	0.076	0.164	-0.293	0.885	0.294	-0.077	0.061	0.035

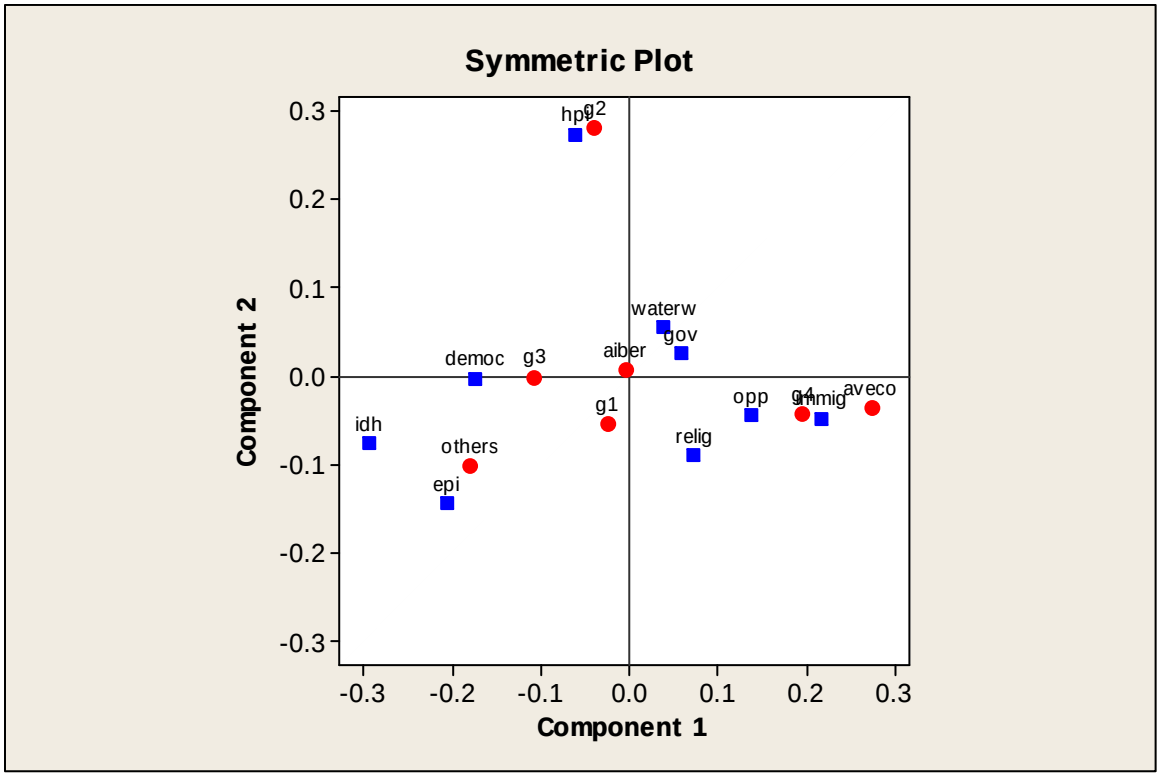
6	gov	0.135	0.090	0.062	0.060	0.114	0.014	0.026	0.022	0.005
7	epi	0.946	0.085	0.125	-0.205	0.635	0.161	-0.143	0.311	0.134
8	hpi	0.948	0.114	0.209	-0.060	0.044	0.018	0.273	0.904	0.650
9	democ	0.556	0.092	0.111	-0.174	0.556	0.125	-0.003	0.000	0.000

Gráfico dos desvios-padrão, relacionando as 3 regiões e os 4 agrupamentos - Row Plot



No gráfico acima destaca-se o agrupamento G2, com a maior dispersão (0.281) em relação ao eixo vertical (Component 2). O G2 é composto em sua quase totalidade por países do AIBER, que aparece no gráfico bem próximo de 0,0 para os dois eixos, provavelmente o modelo tenha capturado algum ruído que justifique este comportamento, e que aparentemente está relacionado ao fato que de os países Espanha, Portugal, Chile e Uruguai de AIBER, estejam no G4. AVECO e OTHERS também apresentaram certa dispersão, mas não ficaram tão distantes de seus agrupamentos.
AIBER destacou-se positivamente por situar-se na parte mais central do gráfico, mostrando a menor dispersão em relação aos dois eixos (*component* 1 e 2)

Gráfico dos desvios-padrão, relacionando as 3 regiões, os 4 agrupamentos e as 9 variáveis - Symmetric Plot



Como vimos na análise do gráfico anterior, o agrupamento G2 é composto em sua maioria por países de AIBER, com exceção de Espanha, Portugal, Chile e Uruguai, e como vínhamos notando nos trabalhos anteriores, este grupo apresentava uma relação bem forte com o hpi (índice de felicidade do planeta), o que percebemos também neste gráfico. As outras variáveis também apresentaram certo nível de dispersão, em sua maioria no eixo horizontal (*component 1*), mas mantiveram certa proximidade com suas regiões e agrupamentos de *clusters*.

12.3 Agrupamentos e Relacionamentos

Nesta sessão serão apresentados gráficos do tipo radar e tabelas de resumo para mostrar como as variáveis se agrupam, se relacionam e seus graus de explicação.

Conforme mencionado na sessão anterior, algumas variáveis tiveram seus nomes reduzidos para melhor acomodação nos gráficos e tabelas. A variável *Opportunity*, tornou-se "opp", *Water withdrawals* = "waterw", *Tolerance for immigrants* = "immi", *Religious tolerance* = "relig", IDH = "idh", *Governance* = "gov", EPI = "epi", HPI = "hpi" e *Democracy* = "democ".

As análises foram efetuadas para conhecer a correspondência, de média e desvio padrão, entre as 9 variáveis selecionadas, as 3 regiões e os 4 clusters.

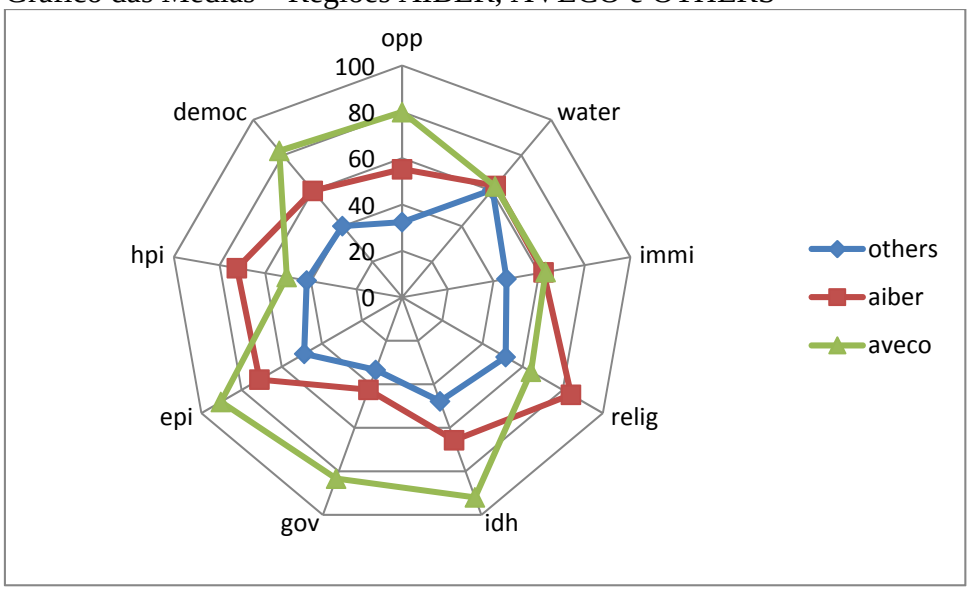
12.3.1 Tabela com as médias e gráficos radar

A tabela abaixo mostra as médias das variáveis contemplando as 3 regiões (others, aiber e aveco) e os 4 agrupamentos (clusters). E o coeficiente de determinação (r2).

MÉDIA									
Mean	**opp**	**waterw**	**immi**	**relig**	**idh**	**gov**	**epi**	**hpi**	**democ**
others	**32.45**	60.64	45.71	51.59	47.76	**33.36**	48.77	41.79	**40.11**
aiber	55.28	62.78	61.97	**84.13**	65.67	42.48	71.05	72.36	59.98
aveco	79.94	62.44	62.86	64.20	**91.89**	83.25	**90.29**	50.58	82.48
R-Sq	**61.81%**	**0.09%**	13.12%	13.21%	44.45%	60.07%	38.44%	26.02%	48.68%
g1	**21.72**	81.50	53.92	53.70	24.39	**23.69**	**23.90**	34.50	33.03
g2	50.49	60.29	56.12	**88.10**	61.47	38.84	68.19	61.40	57.26
g3	34.93	36.76	34.37	31.37	66.13	36.73	65.62	52.78	38.90
g4	78.29	65.47	63.46	69.61	**88.96**	80.98	**88.98**	48.25	80.59
R-Sq	78.52%	28.32%	23.88%	38.72%	**82.45%**	75.06%	79.54%	**20.27%**	60.18%

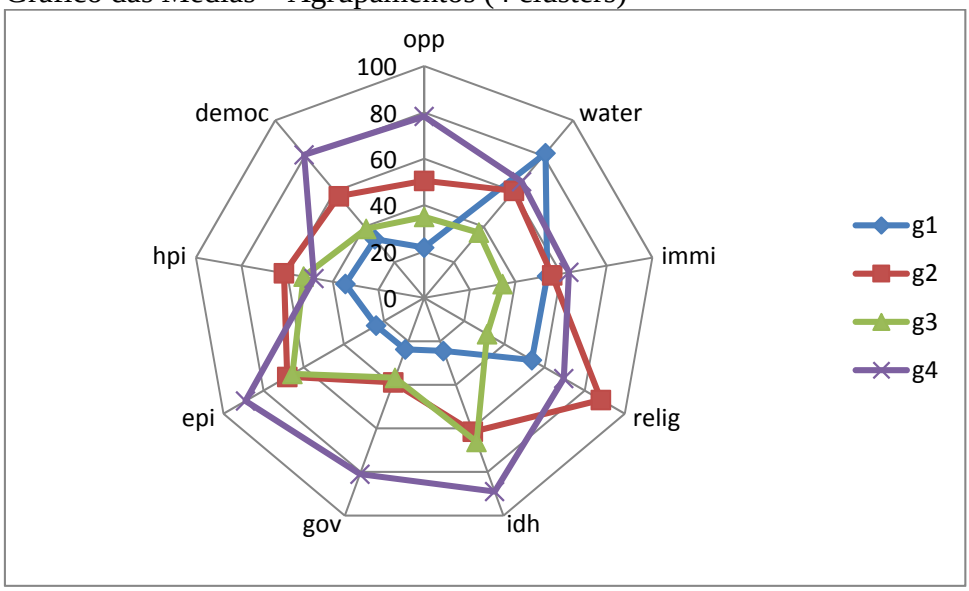
Nesta tabela figuram em cores de destaque as variáveis que apresentaram as 3 (três) maiores médias (cor verde) e as 3 menores (cor laranja), e sua respectiva região e agrupamento.
Mostra também em destaque as variáveis que apresentaram o maior r-quadrado por região e por agrupamento (cor azul) e o menor (cor vermelho).
Pelos resultados acima, vemos que a região OTHERS destacou-se negativamente com as menores médias para as variáveis opp, gov e democ. Comportamento similar nos agrupamentos, onde G1 resultou nas menores médias para opp, gov e epi. No lado positivo, as maiores médias foram para as variáveis idh e epi para AVECO e relig para AIBER.
Desempenho análogo nos agrupamentos, com destaque para idh e epi no G4, e relig para G2 (ver r-quadrado).

Gráfico das Médias – Regiões AIBER, AVECO e OTHERS



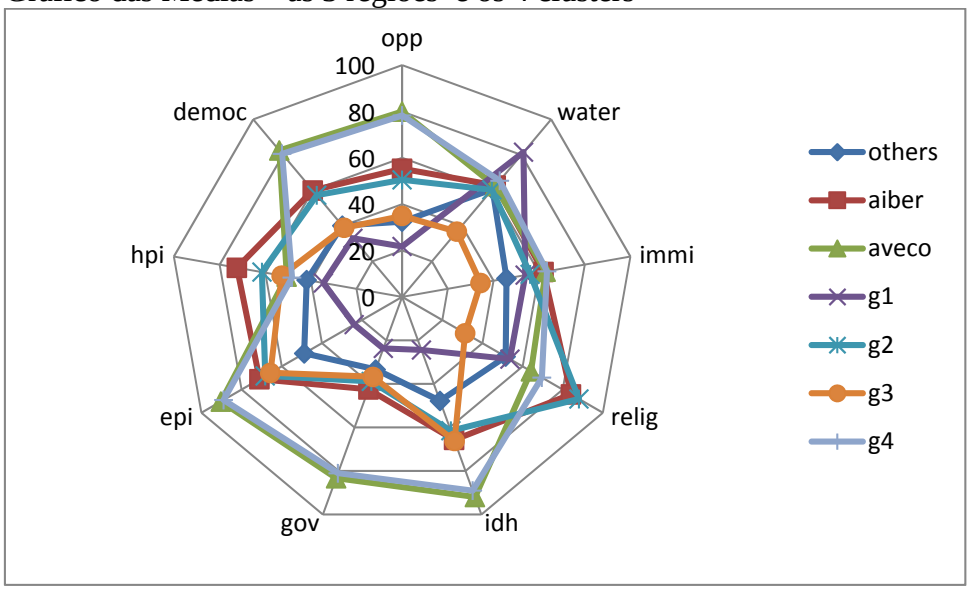
Os maiores destaques para AIBER neste gráfico estão relacionados às variáveis hpi e relig.

Gráfico das Médias – Agrupamentos (4 clusters)



O agrupamento G2, formado predominantemente por países AIBER se destaca positivamente pelo índice relig (*religious tolerance*). E o G1, positivamente pelo water (retirada de água/bombeamento).

Gráfico das Médias – as 3 regiões e os 4 clusters



Quando sobrepostos os gráficos, o indicador relig continua destacando AIBER e G2.

12.3.2 Tabela com os desvios-padrão e gráfico radar

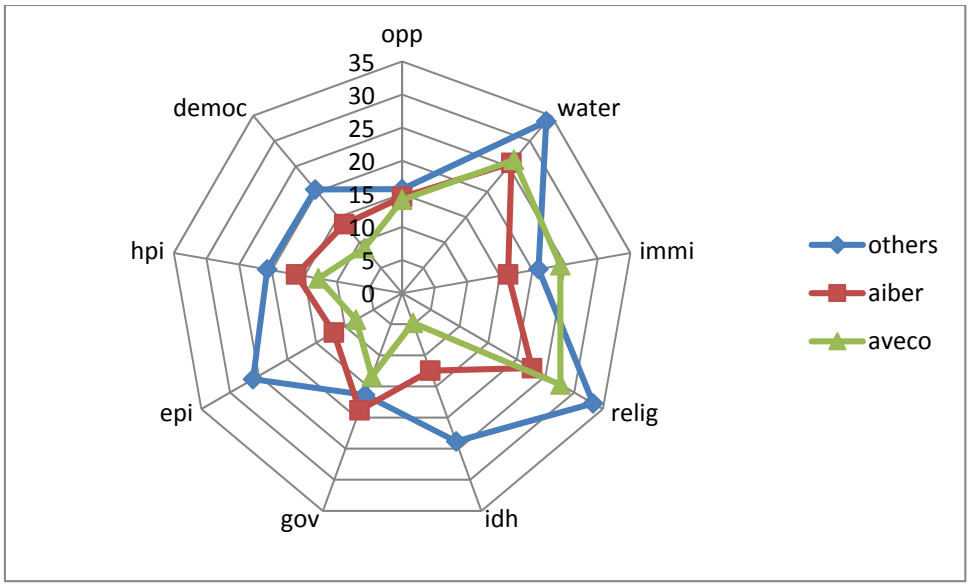
A tabela e gráficos abaixo mostram os desvios-padrão das variáveis contemplando as 3 regiões (others, aiber e aveco) e os 4 agrupamentos (clusters). E o coeficiente de determinação (r-quadrado).

Desvio-Padrão									
StDev	**opp**	**waterw**	**immi**	**relig**	**idh**	**gov**	**epi**	**hpi**	**democ**

others	15.71	**33.86**	20.91	**33.30**	23.84	16.35	26.00	20.68	20.43
aiber	14.52	25.68	16.25	22.65	12.45	18.82	11.91	16.25	13.59
aveco	14.09	26.28	24.27	**27.62**	**4.78**	13.38	**8.04**	12.84	**8.87**
R-Sq	**61.81%**	**0.09%**	13.12%	13.21%	44.45%	60.07%	38.44%	26.02%	48.68%
g1	11.65	27.20	20.38	**35.88**	13.48	12.09	15.37	19.04	16.31
g2	9.76	**28.15**	15.09	18.62	9.98	12.49	**9.08**	**28.32**	12.65
g3	10.38	27.05	18.04	21.62	12.71	14.03	14.19	16.56	21.27
g4	13.39	24.12	23.63	22.27	**7.15**	12.73	9.54	12.50	**9.16**
R-Sq	78.52%	28.32%	23.88%	38.72%	**82.45%**	75.06%	79.54%	**20.27%**	60.18%

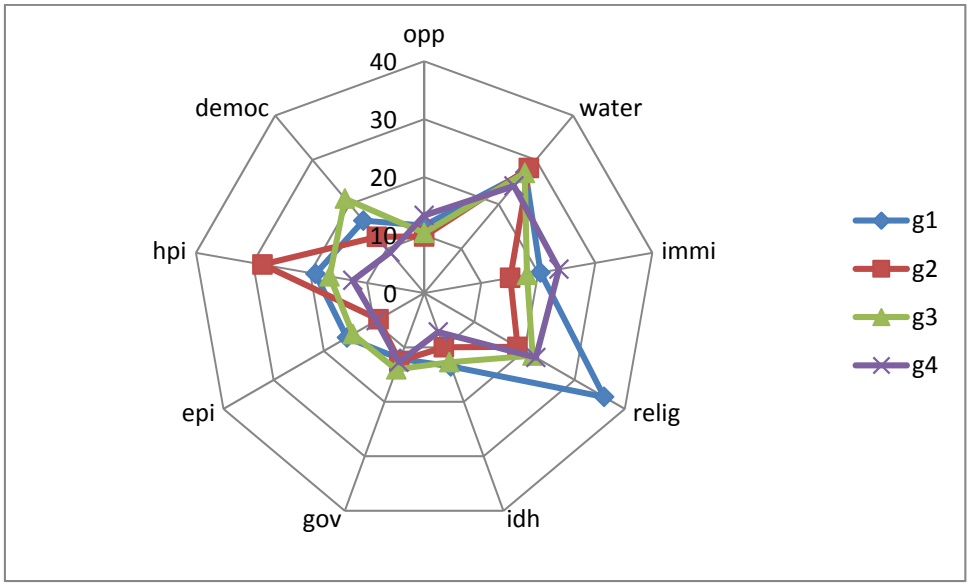
Nesta tabela figuram em cores de destaque as variáveis que apresentaram as 3 (três) maiores médias (cor verde) e as 3 menores (cor laranja), e sua respectiva região e agrupamento. Mostra também em destaque as variáveis que apresentaram o maior r-quadrado por região e por agrupamento (cor azul) e o menor (cor vermelho).
Os resultados acima mostram os maiores desvios-padrão para waterw e relig para a região OTHERS e relig para AVECO (ver r2 muito baixo para as três). E waterw e hpi para G2, e relig para G1. Nos destaques positivos, ou seja, com os menores valores temos idh, epi e democ para AVECO. E nos agrupamentos, idh e democ para G4 e epi para G2.

Desvio-padrão – 3 regiões



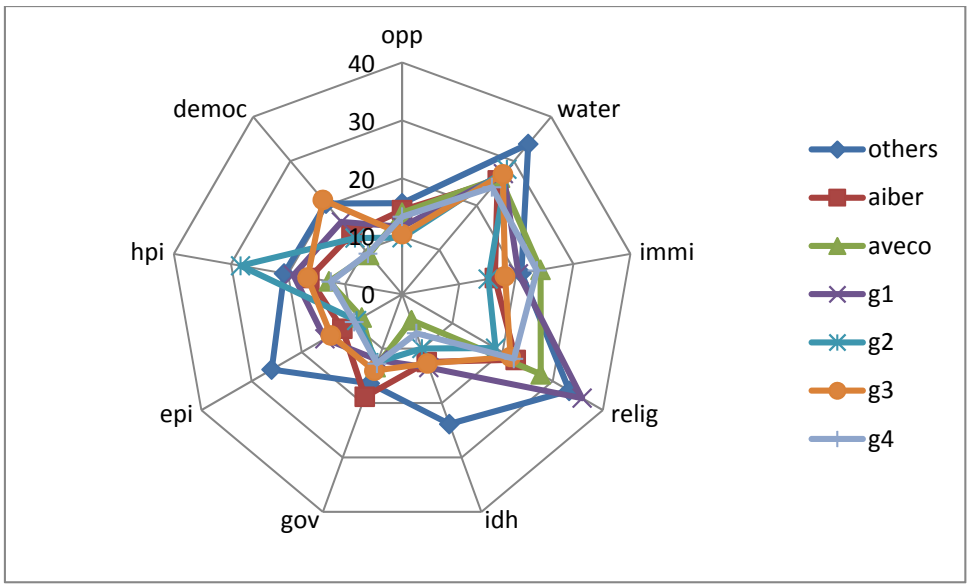
Em AIBER, waterw e relig apresentaram os maiores afastamentos de suas médias.

Desvio-padrão – 4 clusters



Similar ao gráfico da sessão anterior, hpi e waterw apresentaram as maiores dispersões no G2.

Desvio-padrão – as 3 regiões e os 4 clusters



Nesta sobreposição, AIBER destaca-se positivamente em epi. E G2 negativamente em hpi.

12.4. CONSIDERAÇÕES

Em média os resultados da região AVECO estiveram muito próximos aos do agrupamento G4, que se destacaram positivamente em quase todos os indicadores. Similarmente AIBER se relacionando com G2, figurando numa faixa intermediária nos resultados. E OTHERS com países nos agrupamentos G1 e G3, apresentando resultados médios e baixos.

Houve destaque positivo para AIBER em relação aos indicadores relig (tolerância às crenças religiosas) e hpi (índice de felicidade do planeta), mesmo quando a região foi

analisada como agrupamento G2, ou seja, sem os países Espanha, Portugal, Chile e Uruguai. Apesar destes resultados, AIBER (e G2) ainda tem um difícil caminho a trilhar, haja vista os baixos resultados em índices importantes ao desenvolvimento como gov, idh, democ, opp e epi.

13. Análises dos países da região Ibero-Americana (AIBER)

Este sessao se dedicará unicamente a analisar os países da região **Ibero-Americana (AIBER)**. Para tanto serão efetuadas análises de correspondência, para obter graficamente a distribuição dos resultados das variáveis do eixo temático **Biodiversidade - Capital Natural e Social** e analisar o grau de relacionamento entre as variáveis de interesse. Análises de variância (ANOVA), regressões, de agrupamentos (clusters) e de componentes principais também serão efetuadas. E por fim, ao se identificar as variáveis de maior relevância na composição do modelo, será criado um indicador de performance, através do qual os países serão agrupados e ordenados. E um mapa indicando suas posições no ranking através das cores para destaque.

13.1. ENTENDENDO OS DADOS

Os dados desta análise compreende os 21 países pertencentes à região Ibero-Americana (AIBER). Estes dados são representados pelos indicadores sintéticos e analíticos de maior relevância para permitir análises e conclusões sobre o tema **Biodiversidade - Capital Natural e Social** e auxiliar os envolvidos no Projeto ORIBER – Observatório RIBER.

13.1.2 As variáveis do tema Biodiversidade - Capital Natural e Social

Este tema de pesquisa apresenta um total de 33 (trinta e três) variáveis, sendo três categóricas – País, Country code e REGIÃO e trinta variáveis quantitativas, porém neste trabalho utilizaremos somente as variáveis identificadas como de alto poder explicativo, conforme mostra o Quadro 1.

Quadro 1. As variáveis de interesse para este trabalho

Variável	Significado	Tipo	Unidade de Medida
opp (Opportunity)	Indicador composto por: Direitos individuais, Liberdades individuais, Tolerância e inclusão e	Variável Quantitativa	Indicador nominal

	Acesso à educação superior		
waterw (Water withdrawals as a percent of resources)	Uso de água retirada como porcentagem dos recursos. Obedece uma escala de 0 a 5, sendo: 4–5: Extremely high stress (>80%), 3–4: High stress (40–80%), 2–3: Medium-high stress (20–40%), 1–2: Low-medium stress (10–20%) e 0–1: Low stress (<10%) (variável positivada)	Variável Quantitativa	Indicador nominal que varia de 0 a 5
immi (Tolerance for immigrants)	Variável que contabiliza respostas SIM (positivas) da população sobre a percepção do local onde vivem ser um bom lugar para imigrantes de outros países.	Variável Quantitativa	Percentual
relig (Religious tolerance)	Esta variável é a medida de 13 tipos de hostilidades causadas por indivíduos, organizações ou sociedade (assédio, abuso, intimidação,..). A escala varia de 1 (baixa tolerância) a 4 (alta tolerância).	Variável Quantitativa	Indicador nominal que varia de 1 a 4
idh (IDH-2014)	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: Saúde (vida longa e saudável; expectativa de vida), Educação (anos de escolaridade para adultos e crianças) e Renda (padrão de vida medido pela Renda Nacional Bruta – RNB).	Variável Quantitativa	Indicador nominal
gov (Governança)	Indicador que mede a capacidade de um país de eleger, monitorar e substituir seus governantes eleitos. Mede também nível de estabilidade política, medidas contra corrupção, entre outros.	Variável Quantitativa	Indicador nominal
epi (EPI-2014)	Índice de Performance Ambiental (*Environmental Performance Index*), composto por: Saúde Ambiental (saúde humana e dos ecossistemas), Impactos na Saúde (medido pela mortalidade infantil – idade entre 1 e 5 anos), Qualidade do Ar (poluição do ar, partículas sólidas e seu impacto na saúde) e Nível de Água tratada e Sanitização adequada.	Variável Quantitativa	Indicador nominal
hpi (HPI-Happy Planet Index)	HPI, Índice de Felicidade do Planeta, medida do Bem Estar Sustentável (Longevidade, Felicidade e Vida sustentável percebida pela população do país).	Variável Quantitativa	Indicador nominal
democ (Democracy Index – 2014)	A variável Índice de Democracia é formada por 5 categorias: Processo eleitoral e pluralismo, Liberdade civil, Funcionamento do governo, Participação política e Cultura política.	Variável Quantitativa	Indicador nominal

Fonte: O autor a partir dos dados da planilha estatística, do GPS p.16 e do site do Banco Mundial. A pontuação foi padronizada para uma escala de 0 a 100 e as variáveis foram positivadas.

13.1.3 Países analisados

Este trabalho consiste em fazer análises com os indicadores mencionados na sessão anterior abrangendo os 21 países da região Ibero-Americana, conforme o Quadro 2.

Quadro 2. Países contemplados neste estudo

Ord.	PAÍS	COD. PAÍS	REGIÃO
1	Argentina	ARG	AIBER
2	Bolivia	BOL	AIBER
3	Brazil	BRA	AIBER
4	Chile	CHL	AIBER
5	Colombia	COL	AIBER
6	Costa Rica	CRI	AIBER
7	Cuba	CUB	AIBER
8	Dominican Republic	DOM	AIBER
9	Ecuador	ECU	AIBER
10	El Salvador	SLV	AIBER
11	Guatemala	GTM	AIBER
12	Honduras	HND	AIBER
13	Mexico	MEX	AIBER
14	Nicaragua	NIC	AIBER
15	Panama	PAN	AIBER
16	Paraguay	PRY	AIBER
17	Peru	PER	AIBER
18	Portugal	PRT	AIBER
19	Spain	ESP	AIBER
20	Uruguay	URY	AIBER
21	Venezuela	VEN	AIBER

Obs: Conforme estabelecido, Cuba que estava originalmente classificado como um país da região do Caribe, passou a figurar como AIBER para as análises e resultados deste trabalho.

13.2. TRATAMENTO DOS DADOS, AGRUPAMENTOS E RELACIONAMENTOS

Nesta sessão serão realizadas análises de agrupamentos (*clusters*), regressões, análise da variância (ANOVA) e de componentes principais.

As variáveis de trabalho para este estudo, conforme detalhadas na sessão anterior (quadro 1), são: EPI-2014, Governança, Democracy Index-2014, Opportunity, IDH-2014, Tolerance for immigrants, Water withdrawals, HPI (Happy Planet Index) e Religious tolerance.

Tomando como base os resultados das componentes principais, será criado um indicador com as variáveis de maior poder de explicação do modelo e uma tabela de ranking,

ordenando os países pela sua performance e um mapa para mostrar graficamente estes resultados.

13.2.1 Análises de agrupamentos (*clusters*)

Nesta sessão será efetuada análises de agrupamentos a fim de evidenciar o grau de relacionamento das variáveis estudadas. Algumas variáveis tiveram seus nomes reduzidos para melhor acomodação nos gráficos e tabelas. A saber, a variável *Opportunity*, tornou-se "opp", *Water withdrawals* = "waterw", *Tolerance for immigrants* = "immi", *Religious tolerance* = "relig", IDH = "idh", *Governance* = "gov", EPI = "epi", HPI = "hpi" e *Democracy* = "democ".

Quanto a quantidade de *clusters*, foram testadas de 3 a 8 *clusters*, porém a que melhor agrupamento formava foi a de 4 clusters, portanto, a que será utilizada no transcorrer das análises que se seguem.

Abaixo a análise de clusters, utilizando o método completo (*complete*), medida de distância pelo sistema *Manhattan*, e quantidade de clusters igual a 4 (quatro).

Cluster Analysis of Observations: opp, waterw, immi, relig, idh, gov, epi, ...

Manhattan Distance, Complete Linkage
Amalgamation Steps

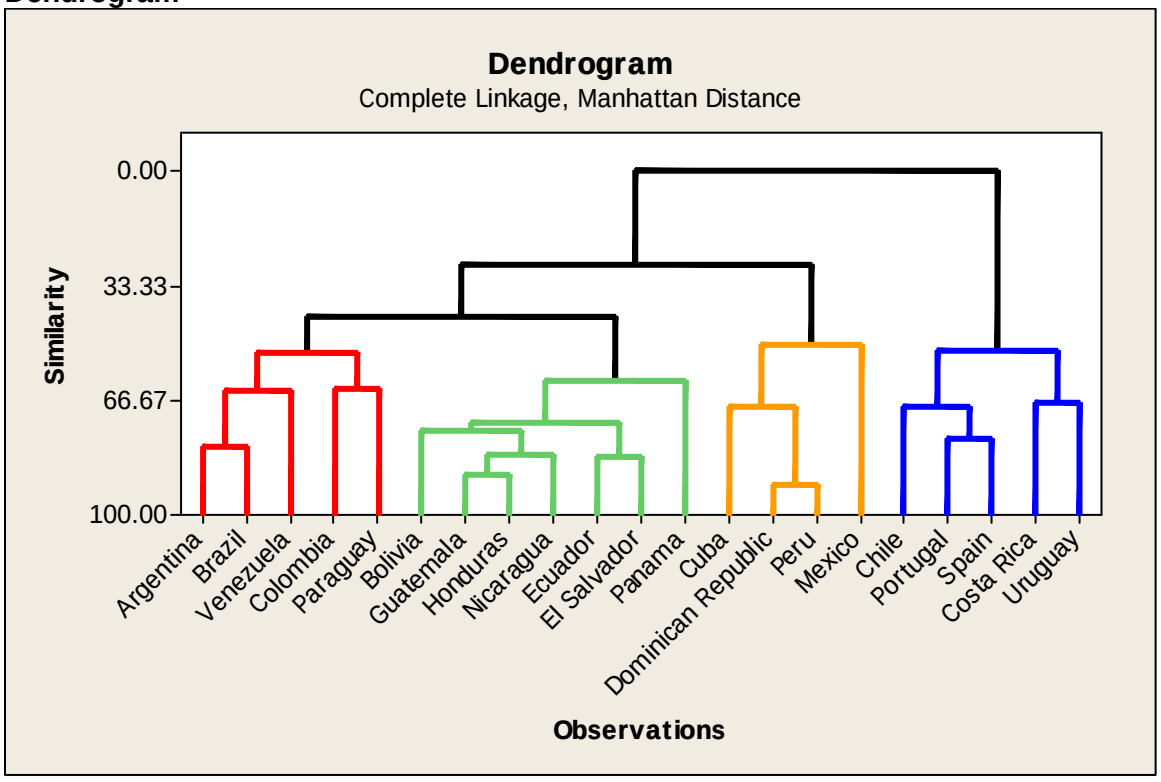
Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	20	91.1800	31.821	8	17	8	2
2	19	88.2816	42.277	11	12	11	2
3	18	82.8518	61.867	9	10	9	2
4	17	82.8070	62.029	11	14	11	3
5	16	80.2934	71.097	1	3	1	2
6	15	78.1834	78.710	18	19	18	2
7	14	75.3982	88.758	2	11	2	4
8	13	73.0828	97.112	2	9	2	6
9	12	68.6452	113.122	7	8	7	3
10	11	68.5229	113.563	4	18	4	3
11	10	67.7118	116.489	6	20	6	2
12	9	63.9665	130.001	1	21	1	3
13	8	63.0945	133.147	5	16	5	2
14	7	60.9999	140.704	2	15	2	7
15	6	52.9883	169.608	1	5	1	5
16	5	52.4140	171.680	4	6	4	5
17	4	50.6913	177.895	7	13	7	4
18	3	42.1679	208.646	1	2	1	12
19	2	27.4809	261.634	1	7	1	16
20	1	0.0000	360.779	1	4	1	21

Final Partition
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	5	6175.88	34.5654	40.8627

```
Cluster2             7      4606.30   24.9371   36.1484
Cluster3             5      7585.58   38.4766   46.4003
Cluster4             4      5594.84   35.1330   54.6214

Cluster Centroids
                                                    Grand
Variable  Cluster1  Cluster2  Cluster3  Cluster4  centroid
opp        53.4788    47.515   76.8383   44.1758   55.2807
waterw     75.7362    81.306   47.7046   33.0142   62.7811
immi       72.0548    52.446   77.5342   46.5753   61.9700
relig      60.0000   100.000   86.6667   83.3333   84.1270
idh        67.5320    54.346   79.0982   66.3773   65.6706
gov        30.7694    33.326   71.5665   36.8007   42.4840
epi        73.0720    60.660   83.5088   71.1553   71.0544
hpi        75.5122    77.386   61.2983   73.4319   72.3563
democ      58.4123    53.842   76.7536   51.6884   59.9752

Distances Between Cluster Centroids

          Cluster1  Cluster2  Cluster3  Cluster4
Cluster1   0.0000   49.0913   67.2170   56.5356
Cluster2  49.0913    0.0000   78.5934   54.2470
Cluster3  67.2170   78.5934    0.0000   67.4596
Cluster4  56.5356   54.2470   67.4596    0.0000
```

Dendrogram



Este dendrograma mostra os agrupamentos feitos pelos países por similaridade. Os países que figuram na cor azul, são os de melhores resultados na análise. Os de cor vermelha e laranja, apresentaram resultados intermediários. E os de core verde apresentaram os valores mais baixos nesta análise.

Pelas conexões das linhas pretas, percebe-se de certa forma relação entre os grupos de cores vermelha e verde, e estes com o de cor laranja.

Ao compararmos a distância entre os coeficientes de correlação (cluster de variáveis), temos os seguintes resultados:

Cluster Analysis of Variables: opp, waterw, immi, relig, idh, gov, epi, democ,…

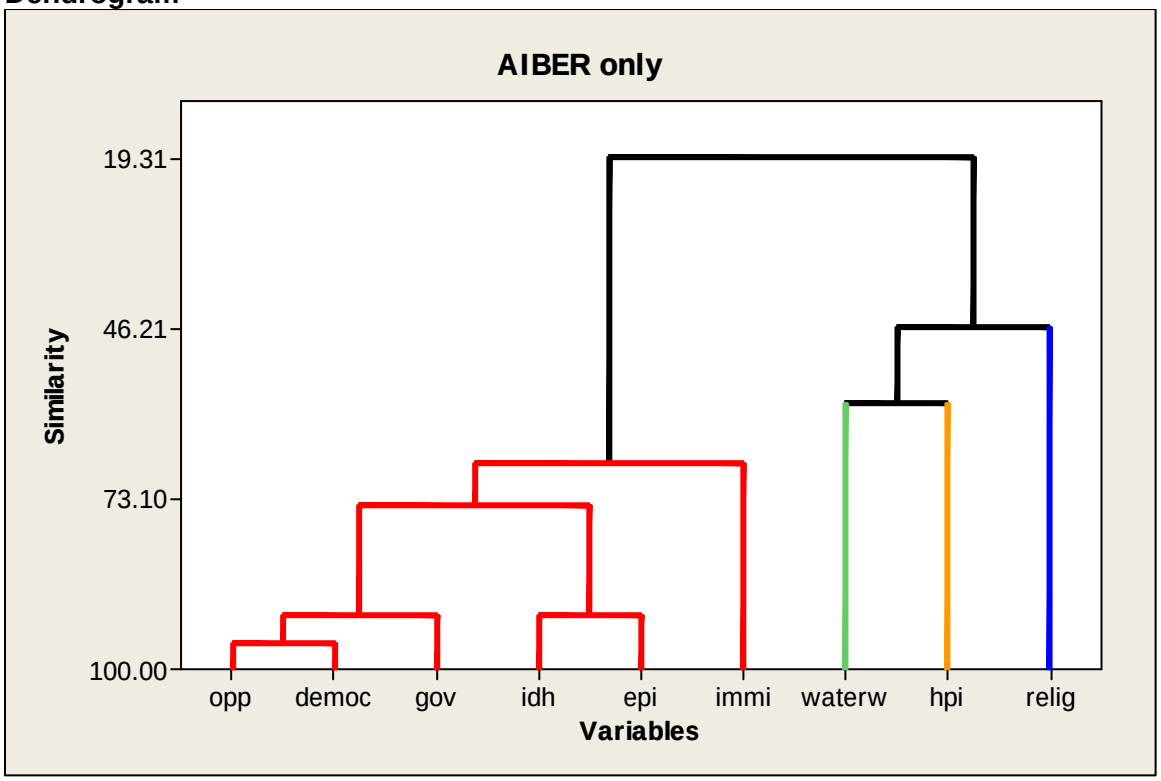
Correlation Coefficient Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	8	95.9319	0.08136	1	9	1	2
2	7	91.5211	0.16958	5	7	5	2
3	6	91.4732	0.17054	1	6	1	3
4	5	73.9669	0.52066	1	5	1	5
5	4	67.3468	0.65306	1	3	1	6
6	3	57.8964	0.84207	2	8	2	2
7	2	46.1059	1.07788	2	4	2	3
8	1	19.3136	1.61373	1	2	1	9

Final Partition

Cluster 1
  opp immi idh gov epi democ
Cluster 2
  waterw
Cluster 3
  relig
Cluster 4
  hpi

Dendrogram



Pela similaridade encontrada na análise da ferramenta estatística, as 6 (seis) primeiras variáveis (cor vermelha), formaram um agrupamento, enquanto as demais não apresentaram similaridade suficiente para formar outros agrupamentos.

13.2.2 Análise de variância - ANOVA

Com os agrupamentos definidos e estabelecido cada país como membro de seu respectivo *cluster*, como próximo passo será executado a ANOVA para entender a variabilidade entre os grupos e auxiliar na classificação para as análises nas sessões posteriores.

One-way ANOVA: idh versus ClusterM

```
Source    DF      SS     MS     F      P
ClusterM   3  1818.5  606.2  8.04  0.001
Error     17  1282.5   75.4
Total     20  3101.0

S = 8.686   R-Sq = 58.64%   R-Sq(adj) = 51.35%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N    Mean   StDev  ---+---------+---------+---------+------
1      5  67.532   9.720              (---------*-------)
2      7  54.346  10.462  (------*------)
3      5  79.098   7.065                          (-------*-------)
4      4  66.377   4.008            (---------*---------)
                          ---+---------+---------+---------+------
                            50        60        70        80

Pooled StDev = 8.686
```

Na análise acima, vemos o que havia sido descrito na sessão anterior, análise de dendrograma, que o grupo 3, no gráfico representado pela cor azul, apresenta a maior média entre todos. Os grupos 1 e 4, com médias entre 66 e 67.5, bem próximas e quase se sobrepondo completamente, representados no gráfico da sessão anterior pelas cores vermelha e laranja, numa posição intermediária. E finalmente o grupo 2, com média de 54.346, os resultados mais baixos da amostra e representados no gráfico pela cor verde.

A coluna de ordenamento ficou com a seguinte característica de grupos:

De 2 para 1
De 4 para 2
De 1 para 3
De 3 para 4

Feito o ordenamento (*cluster members*), é executado novamente a ANOVA com os seguintes resultados:

One-way ANOVA: idh versus ClusterM_reOrd

```
Source          DF      SS     MS     F      P
ClusterM_reOrd   3  1818.5  606.2  8.04  0.001
Error           17  1282.5   75.4
Total           20  3101.0

S = 8.686   R-Sq = 58.64%   R-Sq(adj) = 51.35%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N    Mean   StDev  ---+---------+---------+---------+------
1      7  54.346  10.462  (------*------)
2      4  66.377   4.008            (--------*---------)
3      5  67.532   9.720              (--------*-------)
4      5  79.098   7.065                        (-------*-------)
                          ---+---------+---------+---------+------
                            50        60        70        80

Pooled StDev = 8.686
```

Esta disposição reflete no grupo 1 com os resultados mais baixos, os grupos 2 e 3 os intermediários e o grupo 4, com os melhores.

Apesar de seguir com os 4 (quarto) clusters para as análises das próximas sessões, como indica os resultados da ferramenta Minitab, estamos alterando as classificações dos grupos 2 e 3, para que representem a categoria dos países com indicadores de nível “MÉDIO” devido sua proximidade de valores e sua quase sobreposição demonstrado no histograma da ANOVA (Análise de Variância).

Portanto, a nova classificação será:

De 1 para 1 - BAIXO

De 2 para 2 - MÉDIO

De 3 para 2 - MÉDIO

De 4 para 3 - ALTO

A nova classificação proposta será utilizada na elaboração do *ranking* dos países de AIBER.

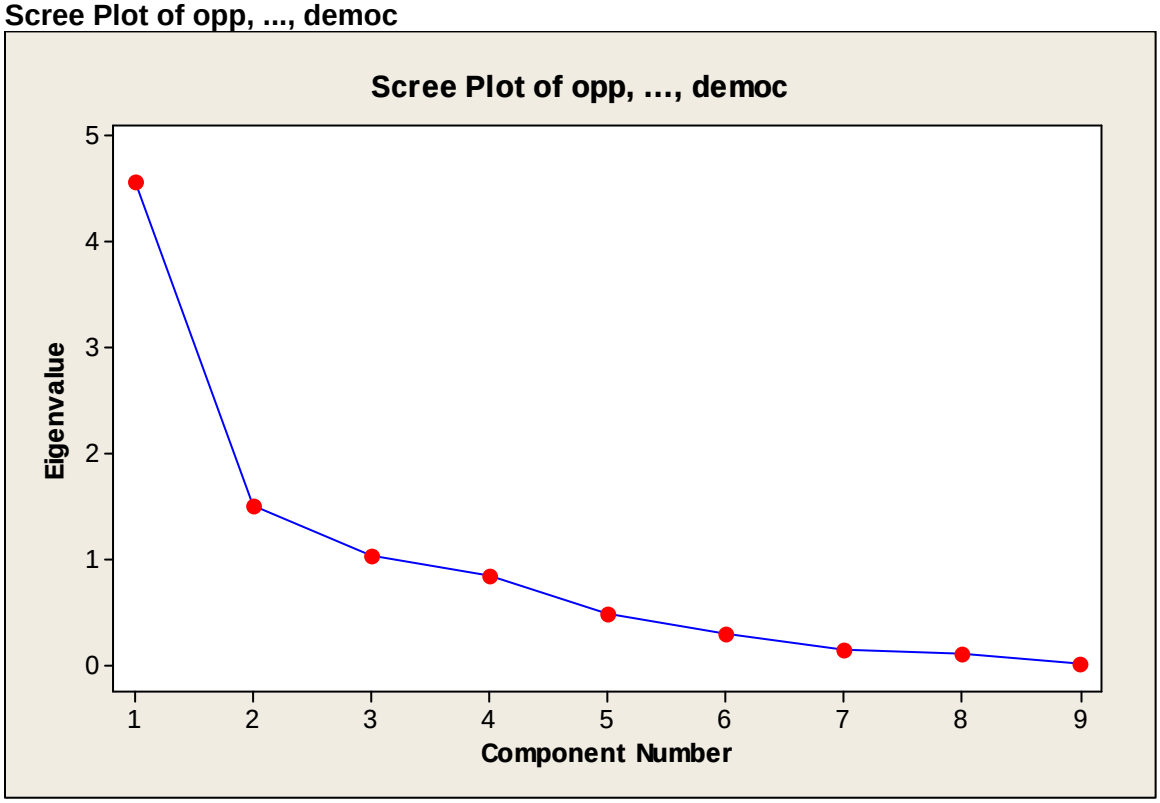
13.3 Análise de Componentes Principais (PC)

Com os agrupamentos definidos e estabelecido cada país como membro de seu respectivo *cluster*, como próximo passo será executado a ANOVA para entender a variabilidade entre os grupos e auxiliar na classificação para as análises das sessões posteriores.

Principal Component Analysis: opp, waterw, immi, relig, idh, gov, epi, hpi, democ.

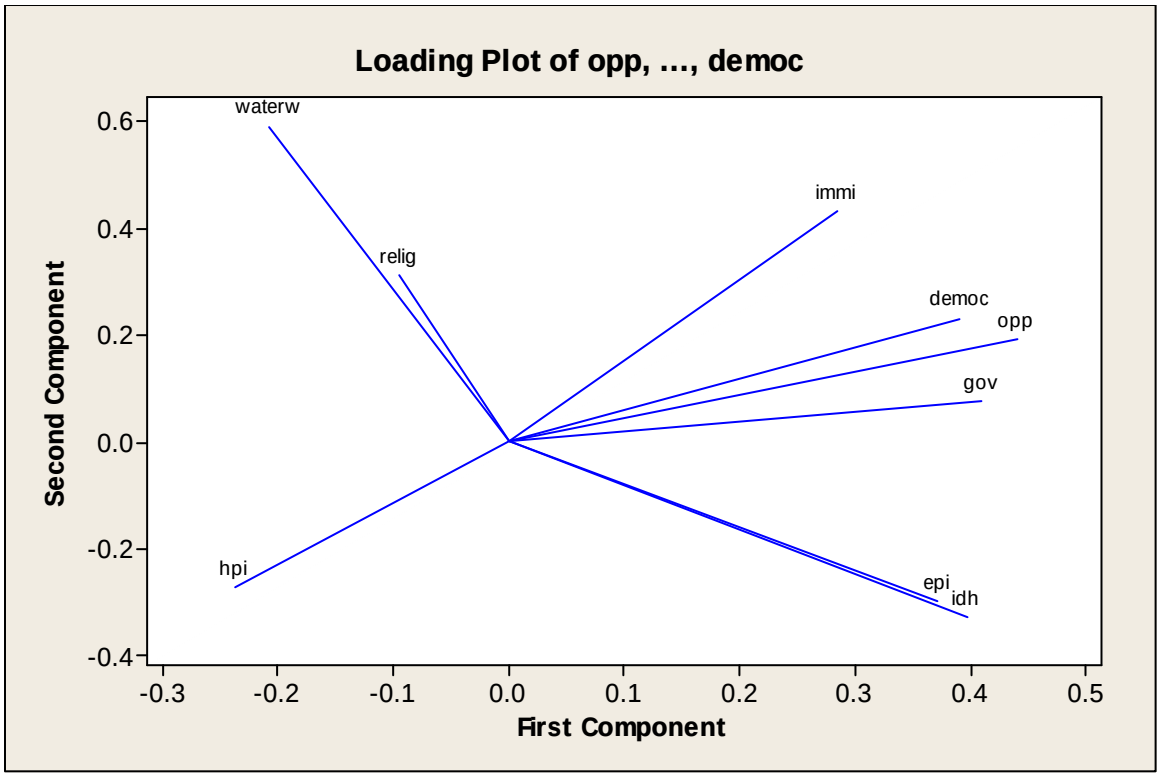
```
Eigenanalysis of the Correlation Matrix

Eigenvalue  4.5595  1.4955  1.0310  0.8446  0.4804  0.3004  0.1520  0.1154
Proportion   0.507   0.166   0.115   0.094   0.053   0.033   0.017   0.013
Cumulative   0.507   0.673   0.787   0.881   0.935   0.968   0.985   0.998

Eigenvalue  0.0212
Proportion   0.002
Cumulative   1.000


Variable      PC1      PC2      PC3      PC4
opp         0.440    0.192    0.075   -0.193
waterw     -0.207    0.591    0.377   -0.075
immi        0.284    0.431    0.155    0.380
relig      -0.095    0.311   -0.814   -0.293
idh         0.397   -0.328   -0.002    0.088
gov         0.410    0.076   -0.105   -0.389
epi         0.371   -0.300   -0.042    0.105
hpi        -0.238   -0.273    0.351   -0.658
democ       0.391    0.231    0.170   -0.351
```

Scree Plot of opp, ..., democ



Na análise, vemos que as componentes PC1, PC2, PC3 e PC4, explicam 88.1% da variabilidade dos dados. O que é percebido visualmente pelo gráfico acima.

Loading Plot of opp, ..., democ



As proximidades mais relevantes, encontram-se no quadrante superior direito e ficam por conta das variáveis democ (*Democracy*), opp (*Opportunity*) e gov (*Governance*), que foi demonstrado na coluna PC1 da tabela de análise de componentes. A variável immi (*Tolerance of immigrants*), se encontra no mesmo quadrante, mas um pouco mais afastada. As demais apresentam contribuição negativa ou quase nula no modelo em pelo menos um dos dois componentes (eixos).

13.4 Novo índice de performance

A nova PC1 criada na sessão anterior será utilizada para a criação de um novo índice de performance.

Para tanto, iniciamos executando a análise de variância ANOVA tomando como base a PC1 e a coluna de Cluster Re-Ordenados.

One-way ANOVA: pc1 versus ClusterM_reOrd

```
Source          DF     SS     MS      F      P
ClusterM_reOrd   3  72.66  24.22  22.22  0.000
Error           17  18.53   1.09
Total           20  91.19

S = 1.044   R-Sq = 79.68%   R-Sq(adj) = 76.09%


                        Individual 95% CIs For Mean Based on
                        Pooled StDev
Level  N    Mean  StDev  ---+---------+---------+---------+------
1      7  -1.753  1.047  (---*---)
2      4  -0.715  0.914      (----*-----)
3      5  -0.106  1.089          (---*----)
4      5   3.132  1.084                       (----*----)
                        ---+---------+---------+---------+------
                        -2.0       0.0       2.0       4.0

Pooled StDev = 1.044
```

Os resultados acima, mostram pelo histograma, que manteve a mesma disposição da análise anterior (idh vs Cluster Re-Ord), porém houve aumento do coeficiente de determinação (r-quadrado), passou a 79.68% contra 58.64% descrito na sessão anterior ao ser comparado idh versus ClusterM_reOrd.

Para determinar o peso da contribuição das variáveis na PC1, será executada regressão *stepwise* utilizando como variável de interesse (*response*) a PC1 e como variáveis independentes (*predictors*), as 9 (nove) variáveis deste estudo, são elas: “opp”, “epi”, “waterw”, “hpi”, “relig”, “immi”, “gov”, “idh” e “democ”.

Stepwise Regression: pc1 versus opp, waterw, epi, gov, immi, hpi, relig, democ,...

```
  Alpha-to-Enter: 0.15  Alpha-to-Remove: 0.15

Response is pc1 on 9 predictors, with N = 21

Step            1        2        3        4        5        6
Constant   -7.636  -10.450   -8.771   -7.469   -6.761   -7.036

opp        0.1381   0.1065   0.1098   0.1034   0.1030   0.0979
T-Value     11.93    13.01    16.75    18.69    20.32    19.00
P-Value     0.000    0.000    0.000    0.000    0.000    0.000

epi                 0.0642   0.0485   0.0484   0.0461   0.0445
T-Value               6.44     5.31     6.68     6.84     7.30
P-Value              0.000    0.000    0.000    0.000    0.000

waterw                      -0.0119  -0.0113  -0.0113  -0.0132
T-Value                       -3.42    -4.10    -4.50    -5.44
P-Value                       0.003    0.001    0.000    0.000

hpi                                  -0.0136  -0.0148  -0.0121
T-Value                                -3.32    -3.90    -3.35
P-Value                                0.004    0.001    0.005

relig                                           -0.0051  -0.0048
T-Value                                           -2.03    -2.14
```

P-Value					0.061	0.050
immi						0.0093
T-Value						2.14
P-Value						0.050
S	0.752	0.425	0.336	0.267	0.244	0.219
R-Sq	88.23	96.44	97.89	98.75	99.02	99.26
R-Sq(adj)	87.61	96.04	97.52	98.44	98.69	98.94

Step	7	8	9
Constant	-6.623	-7.036	-7.803
opp	0.07718	0.06659	0.03030
T-Value	8.40	7.37	*
P-Value	0.000	0.000	*
epi	0.04316	0.03507	0.03118
T-Value	8.32	6.27	*
P-Value	0.000	0.000	*
waterw	-0.01215	-0.00927	-0.00807
T-Value	-5.79	-4.29	*
P-Value	0.000	0.001	*
hpi	-0.01254	-0.01302	-0.01464
T-Value	-4.09	-4.94	*
P-Value	0.001	0.000	*
relig	-0.00650	-0.00573	-0.00420
T-Value	-3.22	-3.26	*
P-Value	0.007	0.007	*
immi	0.01352	0.01434	0.01750
T-Value	3.35	4.12	*
P-Value	0.005	0.001	*
gov	0.01573	0.01961	0.02180
T-Value	2.56	3.55	*
P-Value	0.024	0.004	*
idh		0.01748	0.03190
T-Value		2.39	*
P-Value		0.034	*
democ			0.02878
T-Value			*
P-Value			*
S	0.186	0.159	0.000000
R-Sq	99.51	99.67	100.00
R-Sq(adj)	99.24	99.44	100.00

Com os resultados desta regressão, nota-se, analisando o coeficiente de determinação (r-quadrado), que a variável “opp” (*opportunity*) explica 88.23% da PC1 e somada com “epi” (*environmental performance index*), juntas explicam 96.44% do comportamento da variável PC1.

Ao executar a análise de regressão do PC1 e as duas variáveis de maior peso, “opp” e “epi”, temos:

Regression Analysis: pc1 versus opp, epi

```
The regression equation is
pc1 = - 10.5 + 0.106 opp + 0.0642 epi

Predictor      Coef   SE Coef       T      P
Constant   -10.4504    0.5748  -18.18  0.000
opp        0.106486  0.008183   13.01  0.000
epi        0.064230  0.009974    6.44  0.000

S = 0.424866   R-Sq = 96.4%   R-Sq(adj) = 96.0%

Analysis of Variance

Source          DF      SS      MS       F      P
Regression       2  87.941  43.970  243.59  0.000
Residual Error  18   3.249   0.181
Total           20  91.190

Source  DF  Seq SS
opp      1  80.455
epi      1   7.486

Unusual Observations

Obs   opp      pc1      Fit  SE Fit  Residual  St Resid
  8  43.4  -0.3423  -1.2516  0.1358    0.9093      2.26R

R denotes an observation with a large standardized residual.
```

Que gera a seguinte equação de regressão:

pc1 = - 10.5 + 0.106 opp + 0.0642 epi

Por alguma razão o modelo acima identificou a República Dominicana como uma observação “não usual”. Não houve ajustes de valor na base de dados por não ser o foco deste trabalho.

Para trabalhos futuros, este novo índice de performance será identificado como:

MeAmO = índice de performance do Meio Ambiente (epi) e de Oportunidade Social (opp).

Os países AIBER foram ordenados em ordem decrescente pelos valores do PC1 (indicador MeAmO), o qual gerou o ranking:

countries_ord	ranking	pc1_normaliz
Spain	ALTO	100.000
Portugal	ALTO	93.662
Chile	ALTO	77.400
Uruguay	ALTO	75.803
Costa Rica	ALTO	62.569
Argentina	MÉDIO	55.208
Brazil	MÉDIO	49.868
Panama	MÉDIO	39.209
Mexico	MÉDIO	36.565

Peru	MÉDIO	34.660
Dominican Republic	MÉDIO	33.901
Colombia	MÉDIO	31.318
Paraguay	MÉDIO	30.862
Bolivia	BAIXO	20.620
El Salvador	BAIXO	19.783
Ecuador	BAIXO	18.680
Venezuela	BAIXO	18.599
Cuba	BAIXO	9.896
Nicaragua	BAIXO	2.564
Honduras	BAIXO	0.045
Guatemala	BAIXO	0.000

Devido a proximidade dos resultados na análise de variância (ANOVA), agrupou-se os países dos clusters 2 e 3 que na tabela acima figuram como “MÉDIO” na escala de ranking dos países de AIBER.

Importante pontuar que a Venezuela figurava na categoria “MÉDIO” quando na análise baseada pelo IDH e passou figurar como “BAIXO” ao efetuar as análises pelo PC1. Enquanto o México que estava como “BAIXO”, passou a “MÉDIO”, e Cuba que estava categorizada como “MÉDIO” passou a “BAIXO”. O que denota que o PC1 (Componentes Principais) conseguiu capturar um cenário um pouco distinto ao do IDH.

13.5 MAPA AIBER com ranking

Mapa com os países da região Ibero-Americana - AIBER

Legenda:
Verde escuro = ALTO
Verde médio = MÉDIO
Verde claro = BAIXO
Branco = NÃO pertencem à AIBER

Obs: A cor verde foi escolhida devido padrão do GPS que trata com a cor “verde” a sessão referente ao tema Biodiversidade – Capital Natural e Social.

13.6. CONSIDERAÇÕES

Como resultado das análises foi criado o indicador **MeAmO**, que reflete um índice de performance do Meio Ambiente (*EPI*) e de Oportunidade Social (*Opportunity*).

Cabe ressaltar que a variável *Opportunity* explica 88.23% do indicador MeAmO, e quando somada com "*EPI*" (*environmental performance index*), juntas explicam 96.44% do comportamento do indicador.

Da listagem dos países ordenados pelo indicador MeAmO, foi gerado um ranking de performance e um mapa contendo os países da região AIBER, o qual foi colorido refletindo com seus tons, o posicionamento em relação aos resultados de cada país.

Como já era esperado, países como Espanha e Portugal ficaram no topo da lista com os melhores resultados, e seguidos por estes, estão Chile, Uruguai e Costa Rica. Houveram 7 (sete) países na faixa intermediária da tabela, com destaque para Argentina, Brasil, Panamá e México. E na parte inferior da lista, com as performances mais baixas, foram destaque, Guatemala, Honduras, Nicarágua e Cuba.

14. CONSIDERAÇÕES FINAIS

Em relação ao tema Biodiversidade – Capital Natural e Social, objetivo central desta análise, percebemos no desenrolar das análises que as variáveis de maior significância apresentam um viés mais voltado ao capital social que o natural. Talvez a mensagem que está tentando emergir na evolução deste estudo é a de que o indivíduo, fazendo uma analogia à pirâmide de Maslow, necessite primeiramente tratar dos temas que impacta diretamente suas necessidades mais individuais (comunidade local, respeito e oportunidades), para só a partir daí, poder se dedicar ao entorno (proteção e preservação dos ativos naturais, qualidade de vida e humanidade).

No transcorrer das análises, uma das variáveis que mais se destacou foi *opportunity* composta pelos indicadores de Direitos individuais, Liberdades individuais, Tolerância e inclusão e Acesso à educação superior, que essencialmente demonstra o grau de liberdade dos indivíduos para exercerem seus direitos. Seguida pela variável *epi*, composta por itens como Saúde Ambiental (humana e dos ecossistemas), Impactos na Saúde, Qualidade do Ar e Nível de Água tratada e Sanitização adequada, que juntos formam o Índice de Performance Ambiental.

Com relação às análises dos países, ficou evidente o superior desempenho dos países classificados como AVECO (países de economias avançadas). Já a região Ibero-Americana (AIBER), em média apresentou resultados intermediários na maioria das análises, porém em algumas se destacou positivamente, como no caso do indicador HPI (Índice de Felicidade do Planeta).

Ainda tratando de AIBER, muitos dos resultados foram alavancados pelas presenças de Espanha e Portugal neste grupo. Retirando estes dois da amostra, houve destaque para Chile, Uruguai e Costa Rica com os resultados mais elevados. Seguidos por Argentina, Brasil, Panamá e México numa posição intermediária do *ranking*. Os piores resultados ficaram para Guatemala, Honduras e Nicarágua.

Segundo relatórios da ONU e do Banco Mundial, os países da América Latina vêm melhorando seus indicadores gradativamente, ano após ano, e mostram grande potencial para o futuro. E que parte dos resultados negativos de maneira geral, se devem ao fato de serem países “jovens”, com cerca de 500 anos de idade e baixa qualidade na educação.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

FEA - Faculdade de Economia e Administração

Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA: Integração e Alianças Estratégicas

TRABALHO FINAL

Disciplina: Métodos Quantitativos

Professor: Dr. Arnoldo Jose de Hoyos

Fábia de Mattos Spadoto

JUNHO /2016

1 – INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados relativos ao tema que possui um conjunto de vinte variáveis previamente selecionadas que refletem o desenvolvimento humano considerando 132 países, conforme mostrado no Quadro 1 que apresenta as Variáveis e sua categorização em três e dezessete quantitativas, enquanto a Tabela 1 que apresenta algumas variáveis relacionadas com os 132 países.

Iniciamos o trabalho apresentando o objetivo do Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP – Pontifícia Universidade Católica. Depois partimos para o entendimento dos dados apresentando: média, mínimo, máximo, padronização e normalização dos dados e, por fim, a análise exploratória dos dados sobre o tema: i***ntegração e alianças estratégicas*** empregando software estatístico MINITAB, em especial seus recursos: Display Descriptive Statistics que está no link principal Basic Statistics, ele permite obter a média, o valor mínimo e o valor máximo do objeto estudado.

1.1 O tema de pesquisa: integração e alianças estratégicas

A análise dos conceitos que compõem o tema foi extraída do material de aula Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica. Nesse estudo são apresentadas as informações teóricas de outros indicadores, cuja análise e interpretação estão sob responsabilidade de outros colegas desse curso.

Vale destacar o desafio citado no GPS e sua proposta de solução:

Desafio: *"A grande maioria dos países Ibero-americanos se depara com o seguinte desafio: - como equilibrar a necessária dinâmica econômica com a sustentabilidade ambiental e o equilíbrio social, no contexto de uma gestão aberta, democrática e participativa?"*

Solução: *"GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."*

1.2 ENTENDENDO OS DADOS

Os dados são referentes ao ano de 2014. As variáveis são indicadores agregados ***de integração e aliança estratégica*** de cada país.

1.3 Os Indivíduos

Os indivíduos desta análise são países que constam na base de dados do Banco Mundial quanto a seus indicadores de ***integração e aliança estratégica***. São no total 132 países, sendo desses 21 classificados como Ibero-Americanos.

1.4 As variáveis

Este tema da pesquisa apresenta vinte variáveis, sendo três categóricas e dezessete quantitativas. Elas ainda podem ser divididas em: índices sintéticos e variáveis-componentes ou indicadores, a saber:

- **Índices sintéticos:** São sete: Índice de Progresso Social, o Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD – ONU, Índice de Governança, o Índice de Proteção Ambiental – EPI (*Environmental Protection Index*), o *Happy Planet Index* (HPI), o Índice de Saúde dos Oceanos – OHI (*Ocean Health Index*) e o Índice de GINNI para os países selecionados.

- **Indicadores ou variáveis componentes: Instituições, Cooperação Internacional, Utilização eficaz do apoio, Credibilidade, Cooperação Regional** conforme apresentado no Quadro 1 que classifica o tipo de variável e a unidade de medida

Quadro 1. As **11** Variáveis do Eixo temático: **integração e alianças estratégicas**

Variável	*Significado*	**Tipo**	**Unidade de Medida**
Instituição	**Pilar 1:** Instituições "Fortalecer um quadro institucional que atrai negócios e promove o crescimento, fornecendo a boa governação e os níveis corretos de proteção e incentivos é essencial para a inovação O pilar Instituições capta o quadro institucional de um país (Quadro 1a)..	*Variável Quantitativa*	Percentual
Cooperação internacional	A liderança política está disposta e capaz de cooperar com os apoiantes externos e organizações	*Variável Quantitativa*	Percentual

Utilização eficaz do apoio	Esta questão avalia a forma como a liderança política faz uso de projetos de assistência-major internacionais de cooperação, em sua própria agenda de desenvolvimento técnico ou pessoal. O foco aqui é sobre a capacidade de aprender com know-how internacional, para se adaptar aconselhamento externo às realidades nacionais e para integrar a assistência internacional para uma estratégia consistente e de longo prazo do desenvolvimento.	*Variável Quantitativa*	Percentual
Credibilidade	Esta questão aborda o nível de confiança do governo tem sido capaz de alcançar com a comunidade internacional por suas políticas de reforma no caminho para a democracia ea economia de mercado. As políticas não relacionadas com as reformas democráticas e de mercado (por exemplo, uma cooperação de confiança em alianças militares internacionais ou missões de paz) não devem ser considerados.	*Variável Quantitativa*	Percentual
Cooperação Regional	Esta questão visa avaliar a disposição e capacidade da liderança política: · Desenvolver relações de boa vizinhança · Cooperar com os países vizinhos em organizações internacionais e regionais · Apoiar a integração regional ou internacional.	*Variável Quantitativa*	Percentual
IDH – índice Desenvolvimento humano	Mede o progresso de um país por indicadores de qualidade de vida: renda, saúde e educação são os principais.	*Variável Quantitativa*	Indicador nominal que varia de 0 a 1
GOV - Governança	Indica como os governos são indicados. Como a autoridade do país é exercida e como as politicas são implementadas e qual a capacidade de formulá-las.	*Variável Quantitativa*	Percentual
EPI- Indice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	*Variável Quantitativa*	Percentual

HPI – Bem estar sustentável do país	Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas.	*Variável Quantitativa*	Percentual
OHI – Índice de saúde do Oceano	Pontos de referência para a realização de dez objetivos sócio ecológicos e como os países colocam em prática.	*Variável Quantitativa*	Percentual
GINNI index	Mede a distribuição de renda e de despesas das famílias	*Variável Quantitativa*	Percentual

Fonte: autor a partir dos dados da planilha estatística e do GPS p.16

1.5 Países analisados

Este trabalho consiste em fazer análises com os indicadores mencionados na sessão anterior abrangendo os 132 países representados, conforme o Quadro 2.

Quadro 2. Países contemplados neste estudo

QUANT.	PAÍS	COD. PAÍS	REGIÃO
1	Albania	ALB	EMDEU
2	Algeria	DZA	NSWUA
3	Angola	AGO	SUSAF
4	Argentina	ARG	AIBER
5	Armenia	ARM	CWEAS
6	Australia	AUS	AVECO
7	Austria	AUT	AVECO
8	Azerbaijan	AZE	CWEAS
9	Bangladesh	BGD	ESEAS
10	Belarus	BLR	CWEAS
11	Belgium	BEL	AVECO
12	Benin	BEN	SUSAF

13	Bolivia	BOL	AIBER
14	Bosnia and Herzegovina	BIH	EMDEU
15	Botswana	BWA	SUSAF
16	Brazil	BRA	AIBER
17	Bulgaria	BGR	EMDEU
18	Burkina Faso	BFA	SUSAF
19	Burundi	BDI	SUSAF
20	Cambodia	KHM	ESEAS
21	Cameroon	CMR	SUSAF
22	Canada	CAN	AVECO
23	Central African Republic	CAF	SUSAF
24	Chad	TCD	SUSAF
25	Chile	CHL	AIBER
26	China	CHN	ESEAS
27	Colombia	COL	AIBER
28	Congo, Republic of	COG	SUSAF
29	Costa Rica	CRI	AIBER
30	Croatia	HRV	EMDEU
31	Cuba	CUB	AIBER
32	Czech Republic	CZE	AVECO
33	Denmark	DNK	AVECO
34	Djibouti	DJI	SUSAF
35	Dominican Republic	DOM	AIBER
36	Ecuador	ECU	AIBER
37	Egypt	EGY	NSWUA
38	El Salvador	SLV	AIBER
39	Estonia	EST	AVECO

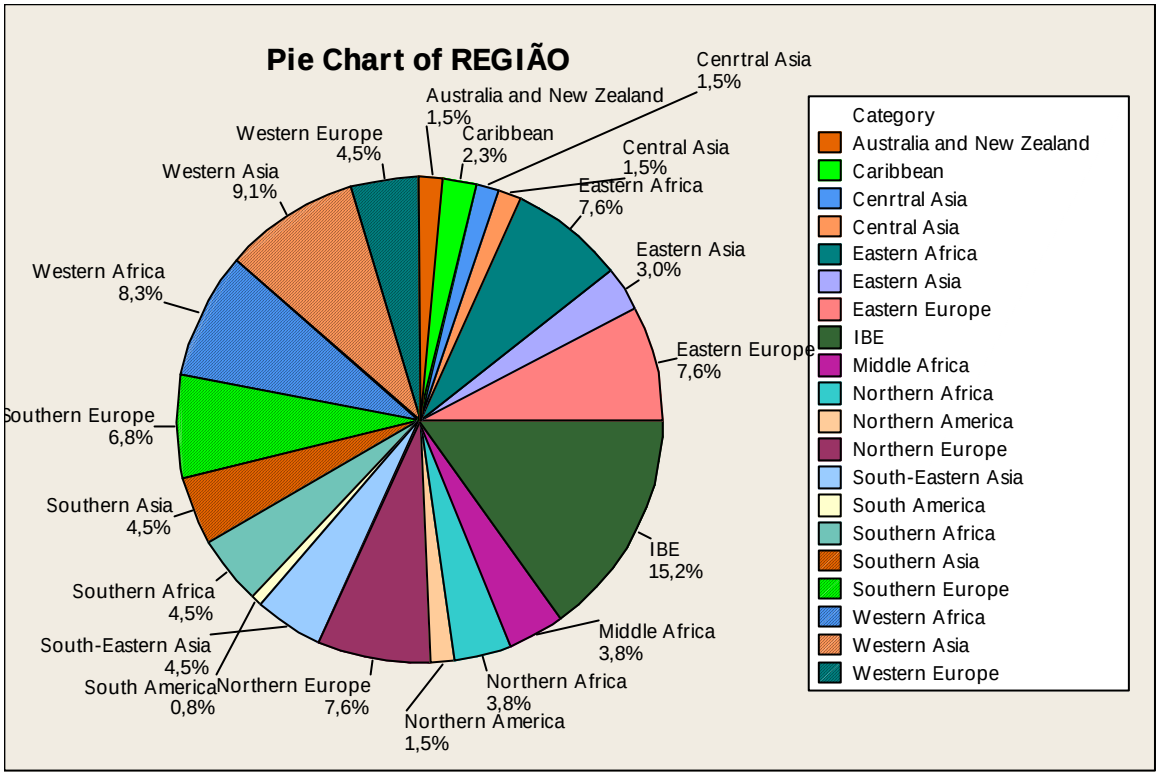
40	Finland	FIN	AVECO
41	France	FRA	AVECO
42	Georgia	GEO	CWEAS
43	Germany	DEU	AVECO
44	Ghana	GHA	SUSAF
45	Greece	GRC	AVECO
46	Guatemala	GTM	AIBER
47	Guinea	GIN	SUSAF
48	Guyana	GUY	CARLA
49	Honduras	HND	AIBER
50	Hungary	HUN	EMDEU
51	Iceland	ISL	AVECO
52	India	IND	ESEAS
53	Indonesia	IDN	ESEAS
54	Iran	IRN	NSWUA
55	Iraq	IRQ	NSWUA
56	Ireland	IRL	AVECO
57	Israel	ISR	AVECO
58	Italy	ITA	AVECO
59	Jamaica	JAM	CARLA
60	Japan	JPN	AVECO
61	Jordan	JOR	NSWUA
62	Kazakhstan	KAZ	CWEAS
63	Kenya	KEN	SUSAF
64	Korea, Republic of	KOR	AVECO
65	Kuwait	KWT	NSWUA
66	Kyrgyzstan	KGZ	CWEAS

67	Laos	LAO	ESEAS
68	Latvia	LVA	AVECO
69	Lebanon	LBN	NSWUA
70	Lesotho	LSO	SUSAF
71	Liberia	LBR	SUSAF
72	Lithuania	LTU	EMDEU
73	Macedonia	MKD	EMDEU
74	Madagascar	MDG	SUSAF
75	Malawi	MWI	SUSAF
76	Malaysia	MYS	ESEAS
77	Mali	MLI	SUSAF
78	Mauritania	MRT	NSWUA
79	Mauritius	MUS	SUSAF
80	Mexico	MEX	AIBER
81	Moldova	MDA	CWEAS
82	Mongolia	MNG	ESEAS
83	Montenegro	MNE	EMDEU
84	Morocco	MAR	NSWUA
85	Mozambique	MOZ	SUSAF
86	Namibia	NAM	SUSAF
87	Nepal	NPL	ESEAS
88	Netherlands	NLD	AVECO
89	New Zealand	NZL	AVECO
90	Nicaragua	NIC	AIBER
91	Niger	NER	SUSAF
92	Nigeria	NGA	SUSAF
93	Norway	NOR	AVECO

94	Pakistan	PAK	NSWUA
95	Panama	PAN	AIBER
96	Paraguay	PRY	AIBER
97	Peru	PER	AIBER
98	Philippines	PHL	ESEAS
99	Poland	POL	EMDEU
100	Portugal	PRT	AIBER
101	Romania	ROU	EMDEU
102	Russia	RUS	CWEAS
103	Rwanda	RWA	SUSAF
104	Saudi Arabia	SAU	NSWUA
105	Senegal	SEN	SUSAF
106	Serbia	SRB	EMDEU
107	Slovakia	SVK	AVECO
108	Slovenia	SVN	AVECO
109	South Africa	ZAF	SUSAF
110	Spain	ESP	AIBER
111	Sri Lanka	LKA	ESEAS
112	Sudan	SDN	SUSAF
113	Swaziland	SWZ	SUSAF
114	Sweden	SWE	AVECO
115	Switzerland	CHE	AVECO
116	Tajikistan	TJK	CWEAS
117	Tanzania	TZA	SUSAF
118	Thailand	THA	ESEAS
119	Togo	TGO	SUSAF
120	Trinidad and Tobago	TTO	CARLA

121	Tunisia	TUN	NSWUA
122	Turkey	TUR	EMDEU
123	Uganda	UGA	SUSAF
124	Ukraine	UKR	CWEAS
125	United Arab Emirates	ARE	NSWUA
126	United Kingdom	GBR	AVECO
127	United States	USA	AVECO
128	Uruguay	URY	AIBER
129	Uzbekistan	UZB	CWEAS
130	Venezuela	VEN	AIBER
131	Yemen	YEM	NSWUA
132	Zambia	ZMB	SUSAF

Gráfico 1: Região dos Países Selecionados



1.5 Variáveis Quantitativas

Como apresentado terceira coluna e corroborado pelo indicador de medida na quarta coluna do Quadro 1, as variáveis são denominadas quantitativas. Essa informação aparece na Tabela 1 com uma seleção de países e indicadores com os números em percentual.

Com isso, a variável é denominada quantitativa pois mostra as que ela pode ser medida em uma escala quantitativa, isto é, com números (SANTOS; PARRA FILHO, 2011).

Elas são Variáveis Continuas porque assumem valores fracionados já que estão em percentual **(IDH e GINNI).**

1.5.1 Média, Mínimo e Máximo das variáveis com uso do MINITAB

Para facilitar o acompanhamento das análises vamos expor os dados do MINITAB citando com títulos que representem a situação de estudo.

- Valor de N é o total de elementos da amostra que nota caso ela é total: 132 países
- **N*** indica o total de dados ausentes na tabela em cada variável. Por exemplo, a variável **International Cooperation** mostra que não tem uma informação que a Austria.
- Mean (média) mostra o valor médio da variável para os 132 países.
- SE Mean mostra o erro padrão da amostra, isto é, ele mostra o quanto a média varia em relação as outra da amostra.
- StDev é o desvio padrão mostra a distância estimativa entre a média e os valores individuais da amostra. Quanto maior for o desvio padrão maior a dispersão ou distância dos dados em relação a média.
- Minimum mostra o valor mínimo dos dados dos 132 países
- Maximum mostra o valor máximo dos dados dos 132 países

Tabela com dados originais da base de dados do tema: **Integração e alianças estratégicas**

Descriptive Statistics: Social Progr; IDH - 2013; 1. Governanç; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93
IDH - 2013	125	7	0,6894	0,0144	0,1611	0,3370	0,5595
1. Governança	132	0	53,38	1,88	21,59	15,22	37,11
1.1 Institutions	123	9	62,14	1,41	15,69	21,10	50,70
1 | International Cooper	105	27	6,963	0,169	1,736	2,000	6,000
1.1.| Effective use of s	105	27	6,790	0,180	1,843	1,000	6,000
1.3.| Regional cooperati	105	27	7,390	0,159	1,632	3,000	6,000
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30
Happy Planet Index	129	3	42,673	0,796	9,040	22,591	36,163
OHI	98	33	64,975	0,881	8,722	45,050	59,645
GINNI Index	119	13	39,024	0,782	8,532	24,820	32,780

Variable	Median	Q3	Maximum
Social Progress Index	63,86	73,74	88,24
IDH - 2013	0,7220	0,8200	0,9440
1. Governança	48,46	70,36	100,00
1.1 Institutions	59,80	73,20	95,30
1 | International Cooper	7,000	8,300	10,000
1.1.| Effective use of s	7,000	8,000	10,000
1.3.| Regional cooperati	8,000	9,000	10,000
EPI Score	52,09	66,40	87,67
Happy Planet Index	42,463	49,286	64,036
OHI	65,140	71,497	82,140
GINNI Index	38,160	44,550	63,140

Análise: Selecionamos o **Indice de Progresso Social** para esse comentário que mostra a **Média de 63,86** em uma escala até 100, indicando que de uma maneira geral os países estão dentro da escala normal. Embora o valor mínimo observado é de 32,60 cujo país desse número é o Chad no centro da África com 12,300 milhões de pessoas e tem seus vizinhos: Líbia, Egito, Nigeria e Sudão. E o valor máximo observado foi 88,24 para o país Nova Zelândia que é um país formado por um conjunto de ilhas sendo as maiores denominadas: Ilha do Norte e Ilha do Sul. Por ter uma característica de economia desenvolvida esse indicador confirma essa qualidade.

1.5.2 Usando o link principal Data, acessando o Code – Numeric to Numeric para substituir os dados ausentes que aparecem na tabela com o símbolo asterisco (*) pela média encontrada.

Foram substituídos:

IDH 2013 – incluímos a média 0,7220 nos países: Argélia, Kazaquistão, República da Korea, Kigisquistão, Latvia, Mauritius e Holanda.

Instituições – incluímos a média 59,80 nos países: Republica Central Africana, Chad, Cuba, Djibouiti, Iraque, Laos, Líberia e Mauritania.

Cooperação Internacional – incluímos a média 7 nos países: Austrália, Áustria, Bélgica, Canadá, Dinamarca, Djibouti, Finlândia, França, Alemanha, Grécia, Guiana, Islândia, Irlanda, Israel, Itália, Japão, Holanda, Nova Zelândia, Noruega, Portugal, Espanha, Swaziland, Suécia, Suiça, Trinidad e Tobago, Reino Unido, Estados Unidos.

Uso Efetivo de suporte – incluímos a média 7 nos países: Austrália, Áustria, Bélgica, Canadá, Dinamarca, Djibouti, Finlândia, França, Alemanha, Grécia, Guiana, Islândia, Irlanda, Israel, Itália, Japão, Holanda, Nova Zelândia, Noruega, Portugal, Espanha, Swaziland, Suécia, Suiça, Trinidad e Tobago, Reino Unido, Estados Unidos.

Credibilidade – incluímos a média 7 nos países: Austrália, Áustria, Bélgica, Canadá, Dinamarca, Djibouti, Finlândia, França, Alemanha, Grécia, Guiana, Islândia, Irlanda, Israel, Itália, Japão, Holanda, Nova Zelândia, Noruega, Portugal, Espanha, Swaziland, Suécia, Suiça, Trinidad e Tobago, Reino Unido, Estados Unidos.

Cooperação regional – incluímos a média 8 nos países: Austrália, Áustria, Bélgica, Canadá, Dinamarca, Djibouti, Finlândia, França, Alemanha, Grécia, Guiana, Islândia, Irlanda, Israel, Itália, Japão, Holanda, Nova Zelândia, Noruega, Portugal, Espanha, Swaziland, Suécia, Suiça, Trinidad e Tobago, Reino Unido, Estados Unidos.

HPI – incluímos a média 42,463 nos países: Lesotho, Montenegro e Swaziland.

OHI – incluímos a média 65,140 nos países: Armênia, Áustria, Azerbaijão, Belarus, Bolívia, Botswana, Burkina Faso, Burundi, Republica Central Africana, Chad, República Tcheca, Hungria, Kazakhstan, Kyrgyzstan, Laos, Lesotho, Macedônia, Malawi, Mali, Moldova, Mongólia, Nepal, Nigéria, Paraguai, Ruanda, Servia, Slovênia, Sri Lanka, Swaziland, Suiça, Tajikistan, Uganda, Uzbekistan, Zambia.

GINNI – incluímos a média 38,160 nos países: Bangladesh, Cuba, Guinea, Islândia, Kuwait, Líbano, Mauritius, Montenegro, Nepal, Paquistão, Arábia Saudita, Sri Lanka, Emirados Árabes Unidos,

Tabela Confirmando a eliminação dos dados ausentes depois que foram substituídos pela Média de cada uma das variáveis.

Descriptive Statistics: Social Progr; Undernourish; Depth of foo; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93
IDH - 2013	132	0	0,6911	0,0137	0,1569	0,3370	0,5700
1. Governança	132	0	53,38	1,88	21,59	15,22	37,11
1.1 Institutions	132	0	61,98	1,32	15,15	21,10	52,15
1 | International Cooper	132	0	6,970	0,135	1,547	2,000	6,075
1.1.| Effective use of s	132	0	6,833	0,143	1,645	1,000	6,000
1.2.| Credibility	132	0	6,765	0,163	1,877	1,000	6,000

1.3.| Regional cooperati	132	0	7,515	0,128	1,475	3,000	7,000
EPI Score	132	0	52,48	1,44	16,55	18,43	39,30
Happy Planet Index	132	0	42,668	0,778	8,936	22,591	36,587
OHI	132	0	65,018	0,653	7,506	45,050	60,833
GINNI Index	132	0	38,939	0,705	8,101	24,820	33,065

Variable	Median	Q3	Maximum
Social Progress Index	63,86	73,74	88,24
IDH - 2013	0,7220	0,8147	0,9440
1. Governança	48,46	70,36	100,00
1.1 Institutions	59,80	71,65	95,30
1 | International Cooper	7,000	8,000	10,000
1.1.| Effective use of s	7,000	8,000	10,000
1.2.| Credibility	7,000	8,000	10,000
1.3.| Regional cooperati	8,000	8,000	10,000
EPI Score	52,09	66,40	87,67
Happy Planet Index	42,463	49,175	64,036
OHI	65,140	69,545	82,140
GINNI Index	38,160	43,553	63,140

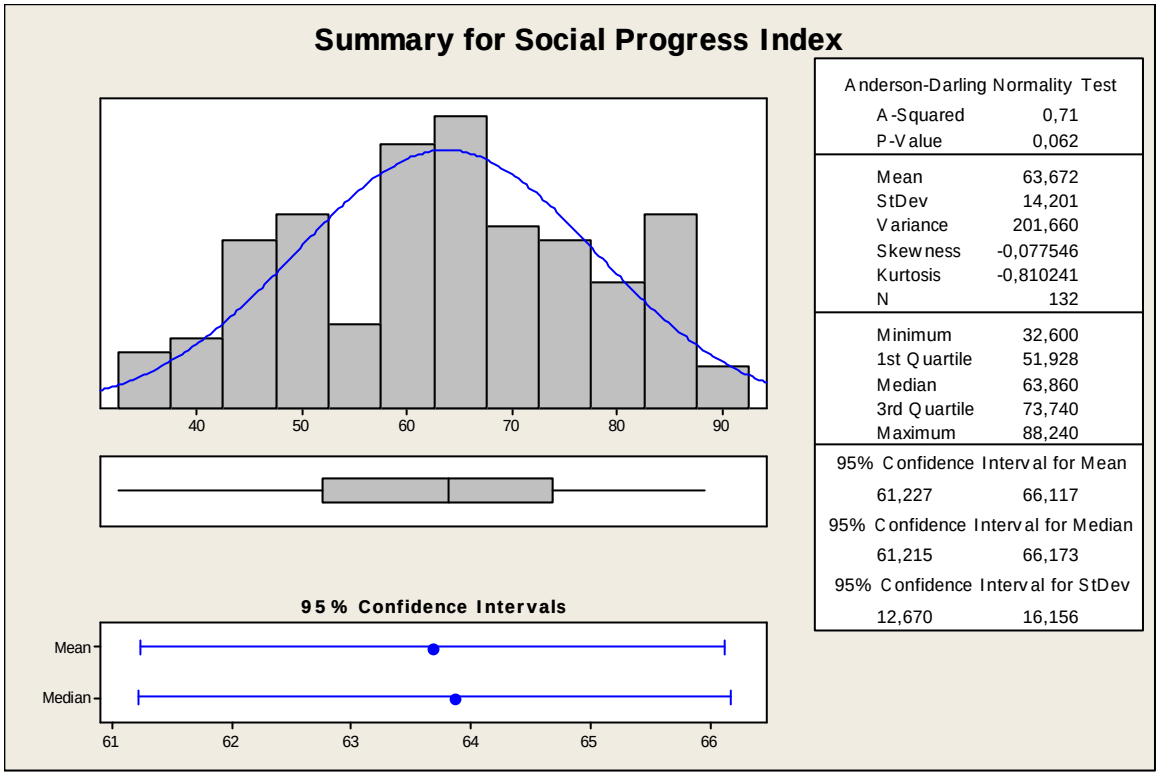
Análise: Após a substituição dos dados ausentes identificados por asterisco (*) pela média de cada variável foi rodado novamente os indicadores de número de elementos (N), dados ausentes (N*), média, erro da média (SE mean), desvio padrão (Stdev), valor mínimo e máximo; apenas para confirmar se foram lançados corretamente. Portanto, sem alteração nos respectivos valores.

1.6. ANÁLISE COMPARATIVA DAS VARIAVEIS

1.6.1 SPI – Social Progress Index - (Índice de Progresso Social)

O índice de progresso social de um país é obtido pela média de 15 indicadores de três grandes temas: **a) Necessidades Humanas Básicas** (Saúde básica e nutrição, Saneamento de águas, Moradia e Segurança pessoal); **b) Fundamentos de Bem Estar** (Acesso à educação básica,

acesso à informação e comunicação, Saúde e bem estar; Sustentabilidade ambiental e Direitos pessoais); e **c) Oportunidade** (Liberdade e escolha pessoal, Tolerância e inclusão e Acesso à educação superior.



Análise dos dados:

Demonstra que a distribuição aproxima-se da normal, considerando os 132 países. Percebe-se um número menor de países do que o esperado classificados entre 32,5 e 37,5 com apenas 4 observações, e o mesmo fenômeno se repete para a faixa de pontuação 87,5 à 92,5, que também apresenta apenas 3 observações. O maior número de observações encontra-se na faixa de 62,5 a 67,5 com 21 observações. Esta faixa corresponde a faixa à mediana e corresponde a um nível de desenvolvimento considerado "médio" pelo índice.

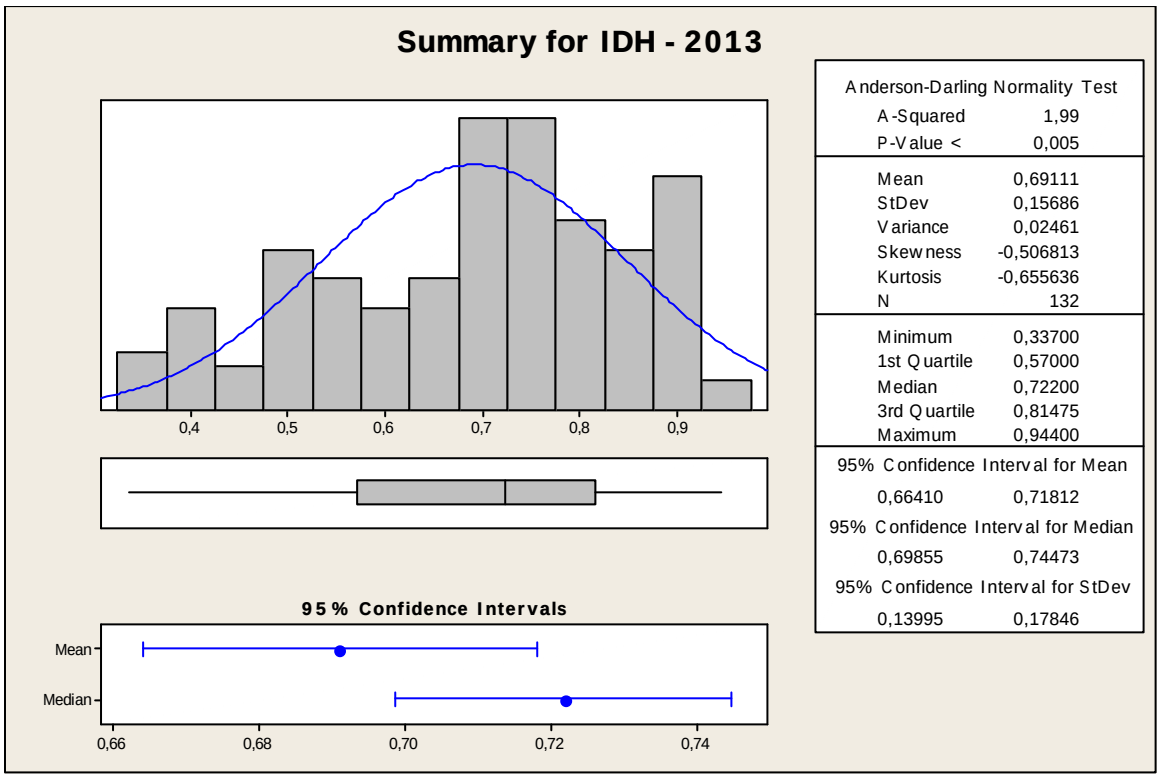
A mediana de 63,86 e o terceiro quartil de 73,74 indicam o poder discricionário do índice, que consegue separar o grupo de países em 12 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados.

1.6.2 IDH – Índice de Desenvolvimento humano

O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.

O trabalho publicado preliminar publicado por PNUD - Programa de las Naciones Unidas para el Desarrollo sobre o desenvolvimento humano em 2014 aponta que o IDH é progresso real da população em relação a expectativa de vida, educação, saúde, habitação, segurança e condições para um desenvolvimento sustentável.

http://hdr.undp.org/sites/default/files/hdr14-summary-es.pdf



Análise dos dados

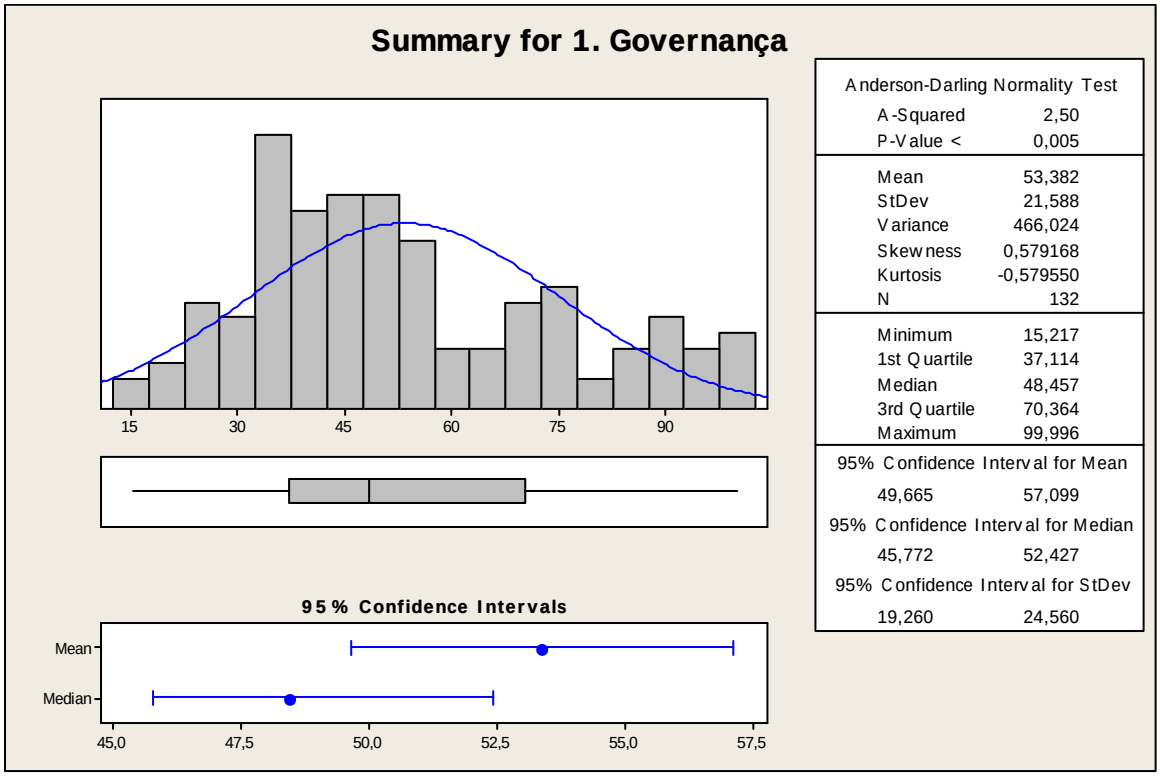
O gráfico para o IDH aproxima-se de uma distribuição normal, mas deslocada para a direita, com uma concentração maior de países dos 0,675 a 0,725 pontos, totalizando 40 países. Isso pode ser constatado também pela mediana, em 0,7220.

Do lado extremo esquerdo do gráfico estão 4 paíss com 0,325 a 0,375 no estrato de -5 e 5 pontos, sendo a Nigéria, Republica do Congo, República Central Africana e Chad. Estes são considerados países de baixo desenvolvimento ou subdesenvolvidos.

1.6.3 – Governança

Governança: World Economic Fórum ressalta que as estratégias de atuação: Simplificar os processos administrativos e fazer uso de alianças regionais e de tecnologia para coibir a corrupção e os abusos; aprimorar os instrumentos e mecanismos regionais que possibilitem

uma ação coletiva capaz de prevenir violações da democracia e garantir a preservação da ordem democrática (http://www3.weforum.org/docs/LA11/WEF_LA11_Report_PT.pdf).



Anderson-Darling Normality Test	
A-Squared	2,50
P-Value <	0,005
Mean	53,382
StDev	21,588
Variance	466,024
Skewness	0,579168
Kurtosis	-0,579550
N	132
Minimum	15,217
1st Quartile	37,114
Median	48,457
3rd Quartile	70,364
Maximum	99,996
95% Confidence Interval for Mean	
49,665	57,099
95% Confidence Interval for Median	
45,772	52,427
95% Confidence Interval for StDev	
19,260	24,560

Análise dos dados

O gráfico para a variável GOV aproxima-se de uma distribuição normal deslocada para a esquerda, o primeiro quartil de 12,5 e a mediana de 48,457 confirmam esse deslocamento, indicando uma concentração crescente de países entre as faixas de 12 a 17 (2 países), de 17 a 22 (3 países), de 22 a 37 (7 países) até chegar ao pico – a faixa 32 a 37 (18 países). Essas faixas agrupam o equivalente a 27% dos países analisados.

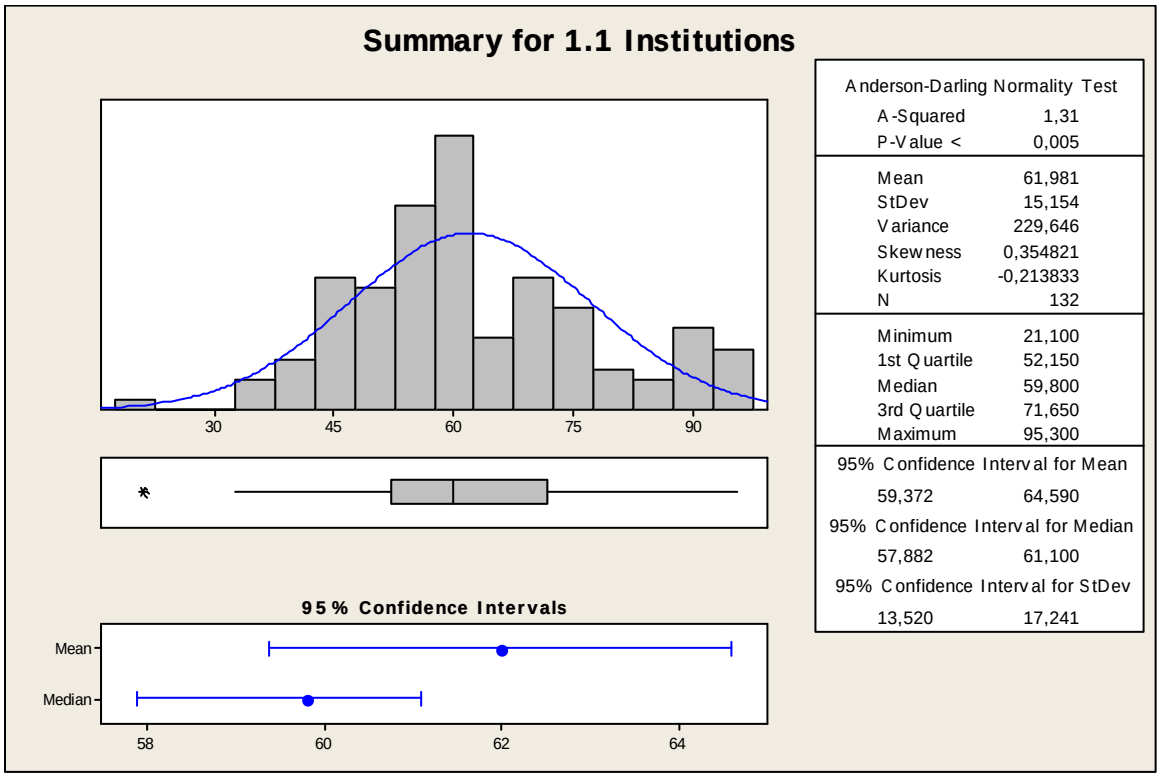
Em relação ao pico, os países mais bem colocados são a Honduras (37,36), a Azerbaijao (37,16) e o Líbano (37,15) e na faixa inferior ficam Bangladesh (32,16), Camarões (31,34) e Togo (30,46). Aparentemente trata-se de grupo bastante heterogêneo, não sendo possível estabelecer relações imediatas.

1.6.4 – Institutions (Instituições)

O ambiente político sub pilar inclui três índices que refletem percepções de a probabilidade Que um governo pode ser desestabilizado; a qualidade do público e dos serviços públicos, a formulação de políticas e implementação; e as percepções de violações à liberdade de imprensa.

O ambiente regulatório sub-pilar baseia-se em dois índices que visam captar as percepções sobre a capacidade do governo de formular e implementar políticas coesas que promovem o desenvolvimento do sector privado e avaliar a medida em que o Estado de Direito prevalece (em aspectos tais como o cumprimento de contratos, direitos de propriedade, a polícia e os tribunais).

O terceiro indicador avalia o custo de demissão redundância como a soma, em semanas de salário, do custo dos requisitos de aviso prévio adicionado ao pagamento de indenizações devidas na demissão de um trabalhador despedido. (https://www.globalinnovationindex.org/content.aspx?page=GII-Home (pag 42 e 392))



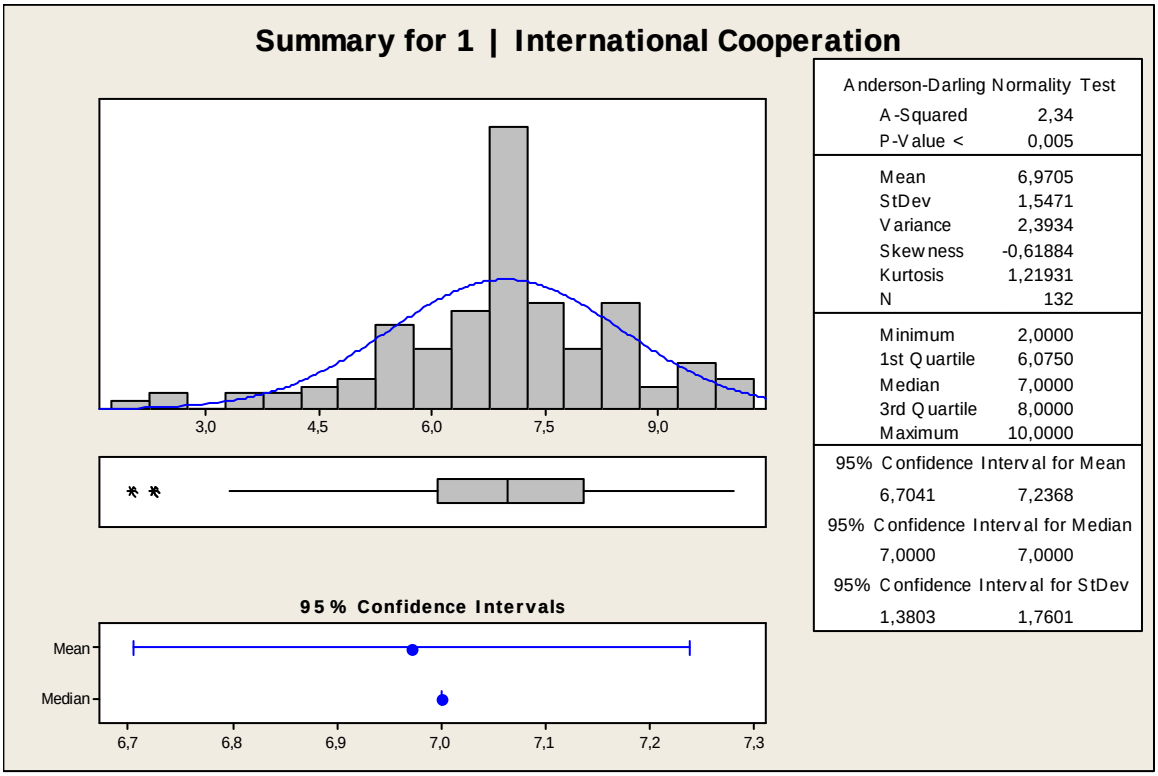
Análise dos dados

O gráfico deste índice tem uma maior concentração na direita como média 59,80 e tendo 27 países nesta posição, sendo 20% do total dos países.

O gráfico apresenta 1 outlier que é a Venezuela que tem o valor do mínimo que 21,1, sendo o máximo deste gráfico com 95,30 composto por 6 países sendo Canadá (92,7), Holanda (93,3), Dinamarca (93,6), Noruega (94,1), Nova Zelândia (94,3) e Finlândia (95,3).

1.6.5 International Cooperation (Cooperação Internacional)

A liderança política está disposta e capaz de cooperar com os apoiantes externos e organizações. http://www.bti-project.org/bti-home/



Análise dos dados

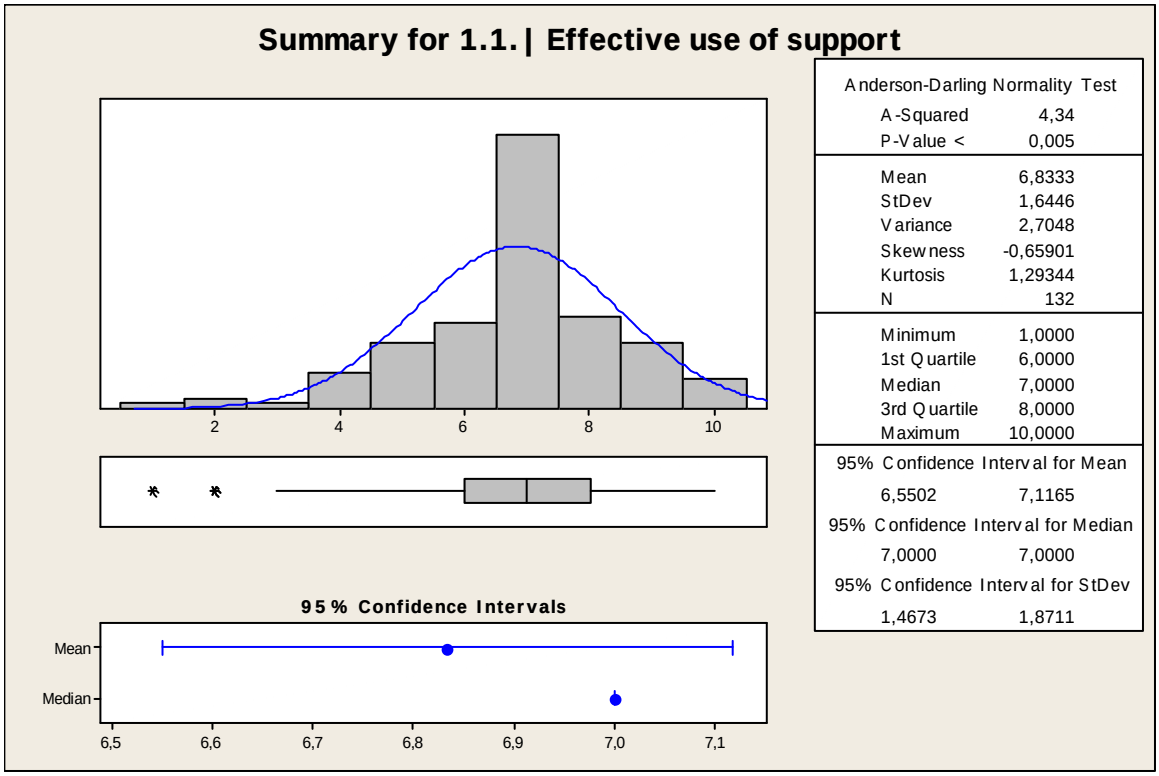
Neste gráfico a média e mediana estão muito próximas, sendo 6,9705 e 7 respectivamente. E a distribuição pode ser classificada como assimétrica negativa.

O gráfico apresenta 2 outliers, sendo o mais ao extremo o Iran (2,0) e depois em seguida o Sudão (2,3). O pico do gráfico é bastante acentuado e é composto por 37 países entre 6,75 e 7,25, sendo 28% do total de países.

1.6.6. Effective use support (Utilização eficaz do apoio)

A liderança política faz uso bem focado de assistência internacional, a fim de implementar a sua estratégia de longo prazo de desenvolvimento.

A liderança política tenta usar a assistência internacional para a sua própria agenda de desenvolvimento, mas tem CITS defi claras na elaboração de uma estratégia consistente de longo prazo que poderia integrar esse apoio de forma eficaz. http://www.bti-project.org/uploads/tx_itao_download/BTI2014_Codebook_01.pdf



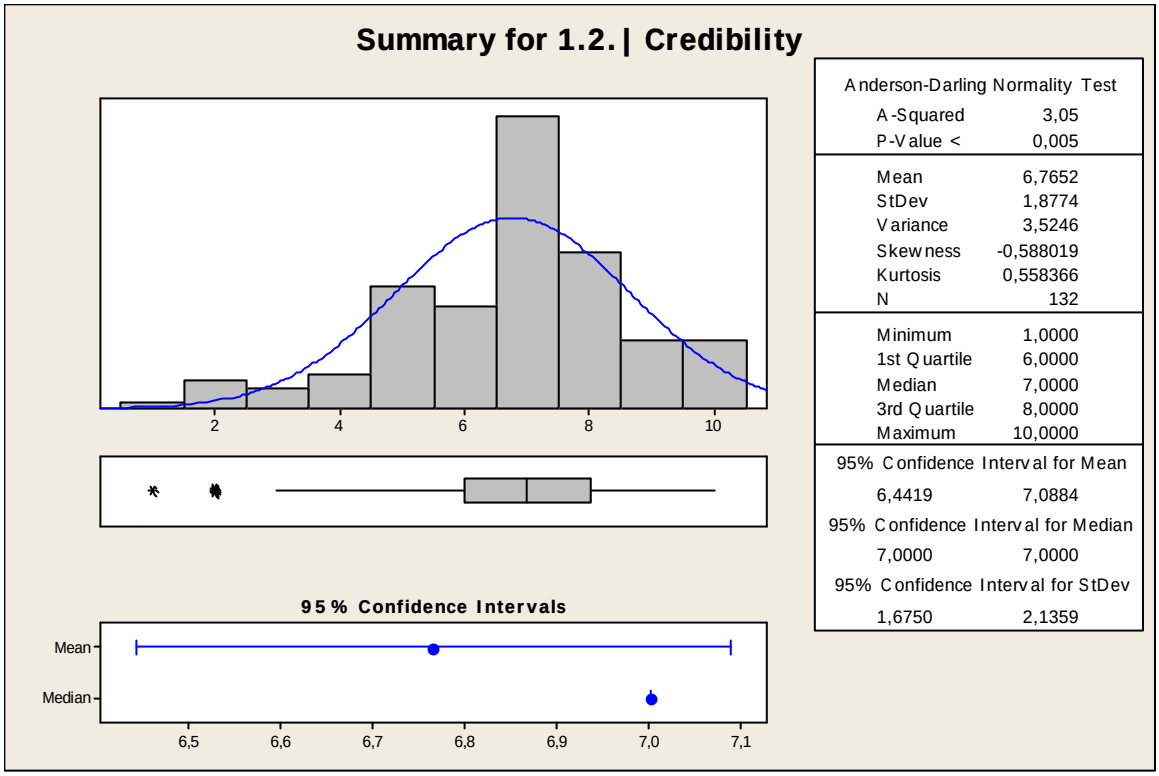
Análise dos dados

O gráfico apresenta uma distribuição mais acentuada para a direita. Há 2 outliers no lado esquerdo, sendo o mais ao extremo o Iran (1,0) e depois em seguida com dois países Sudão (2,0) e Uzbekistan (2,0). O pico tem uma concentração acentuada representando 41% do total de países.

O ponto máximo tem valor 10 é composto por 6 países, sendo eles República da Coreia, Estônia, Brasil, Uruguai, Chile e Polônia, todos com 10.

1.6.7. Credibility (Credibilidade)

Esta questão aborda o nível de confiança do governo tem sido capaz de alcançar com a comunidade internacional por suas políticas de reforma no caminho para a democracia ea economia de mercado. As políticas não relacionadas com as reformas democráticas e de mercado (por exemplo, uma cooperação de confiança em alianças militares internacionais ou missões de paz) não devem ser considerados.
A comunidade internacional inclui organizações internacionais, doadores bilaterais e multilaterais, governos e investidores estrangeiros e ONGs internacionais. O governo é considerado um parceiro credível e fiável pela comunidade internacional. http://www.bti-project.org/uploads/tx_itao_download/BTI2014_Codebook_01.pdf



Análise dos dados

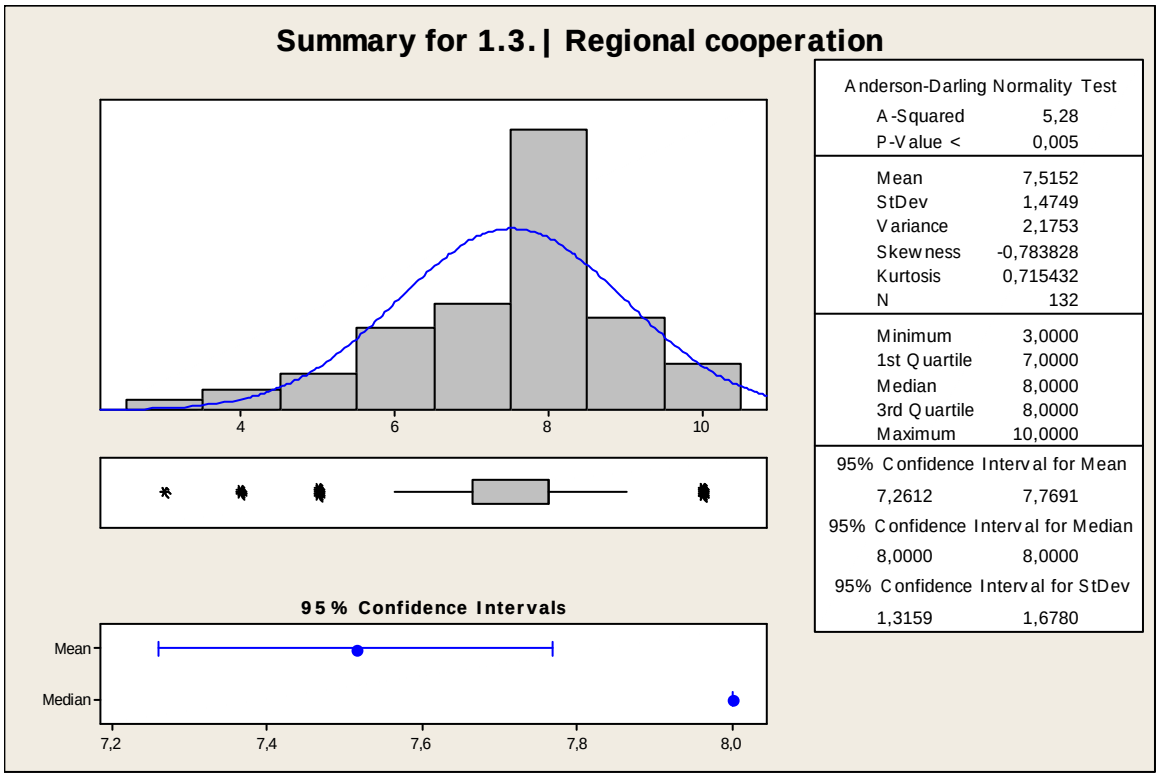
O gráfico apresenta uma distribuição mais acentuada para a direita. Há 2 outliers no lado esquerdo, sendo o mais ao extremo o Iran (1,0) e depois em seguida com 4 países Sudão (2,0) e Uzbekistan (2,0), Venezuela (2,0) e Belarus (2,0). O pico tem uma concentração acentuada representando 33% do total de países. O ponto máximo tem valor 10 é composto por 10 países.

1.6.8 Regional Cooperation (Cooperação Regional)

Esta questão visa avaliar a disposição e capacidade da liderança política:

- Desenvolver relações de boa vizinhança
- Cooperar com os países vizinhos em organizações internacionais e regionais
- Apoiar a integração regional ou internacional.

http://www.bti-project.org/uploads/tx_itao_download/BTI2014_Codebook_01.pdf

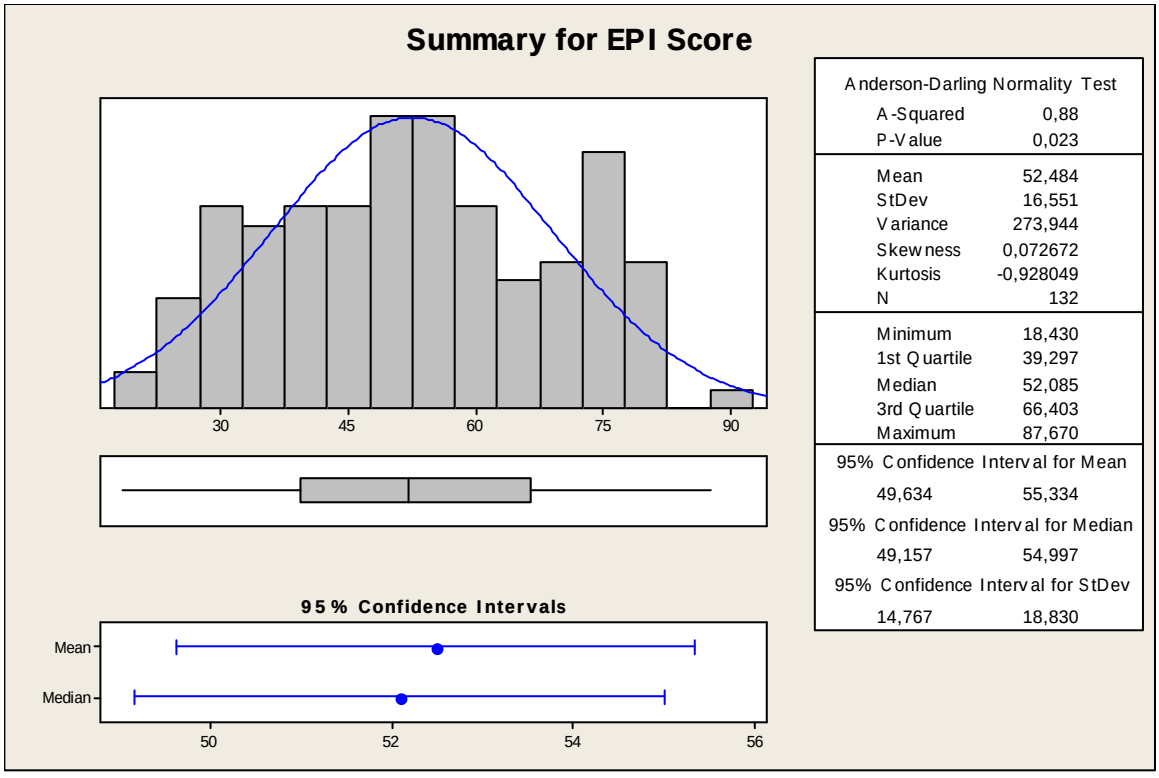


Análise dos dados

O gráfico apresenta uma distribuição mais acentuada para a direita. Há 3 outliers no lado esquerdo, sendo o mais ao extremo o Sudão (2,0) e Uzbekistan (2,0). Depois em seguida com 4 países Iran (4,0), Belarus (4,0), Rússia (4,0) e Camarões (4,0). No canto direito outro outlier com 9 paises sendo Malasia, República Tcheca, Latvia, El Salvador, Eslovaquia, Brasil, Uruguai, Chile e Polônia. O pico tem uma concentração acentuada com 55 países que representa 42% do total.

1.6.9. EPI Score

Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental. http://epi.yale.edu/epi



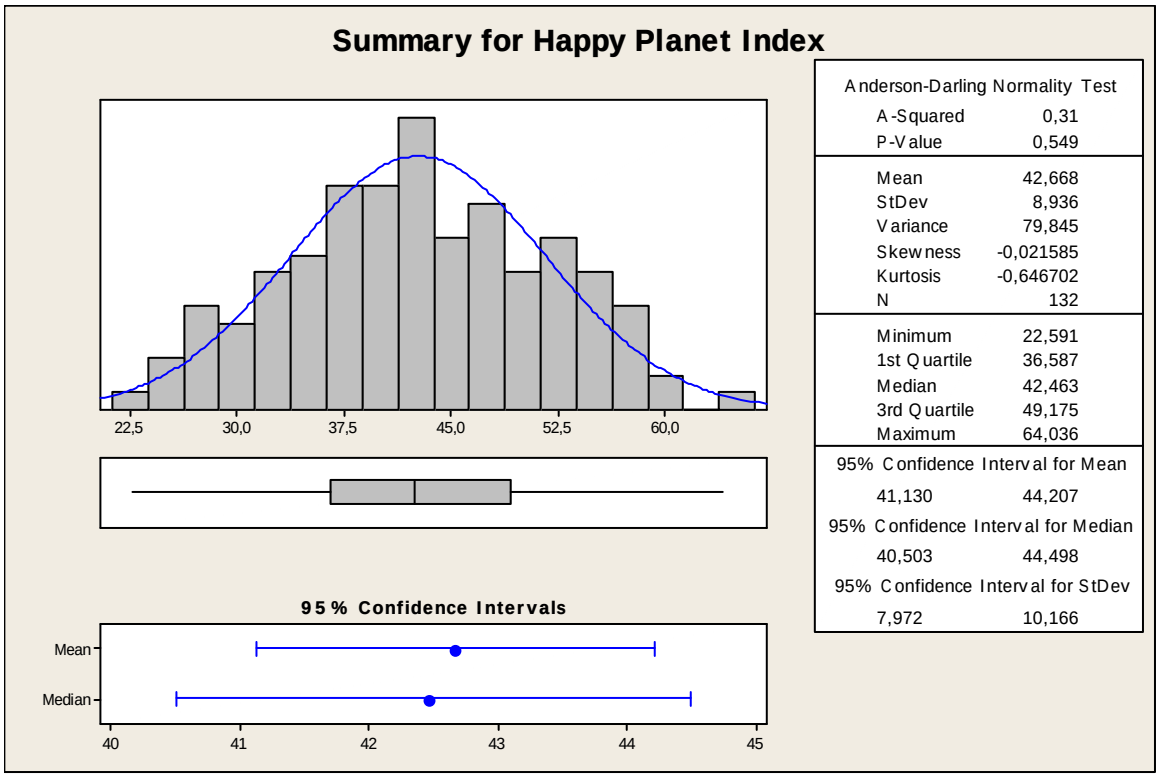
Análise dos dados

A análise dos dados do Índice de Proteção Ambiental apresenta uma distribuição normal, com média de 52,48 e mediana de 52,08, ou seja, um ligeiro deslocamento para a esquerda. Há 2 estratos extremos, totalizando 32 países.

Já nos dois estratos inferiores temos a seguinte composição: no primeiro estrato há dois países – Mali (EPI = 18,43) e Lesoto (EPI = 20,81) , seguidos por um grupo de 6 países. Destes 6 países, 5 encontram-se na África, sendo Bangladesh a única exceção do grupo. O país melhor colocado é Mauritânia, com EPI de 27,190.

1.6.10 – HPI

Medida de bem estar sustentável país oferece uma vida feliz sustentável a longo prazo para as pessoas que vivem nelas. http://www.happyplanetindex.org/data/



Anderson-Darling Normality Test	
A-Squared	0,31
P-Value	0,549
Mean	42,668
StDev	8,936
Variance	79,845
Skewness	-0,021585
Kurtosis	-0,646702
N	132
Minimum	22,591
1st Quartile	36,587
Median	42,463
3rd Quartile	49,175
Maximum	64,036
95% Confidence Interval for Mean	
41,130	44,207
95% Confidence Interval for Median	
40,503	44,498
95% Confidence Interval for StDev	
7,972	10,166

Análise dos dados

A distribuição para o Índice de Felicidade Bruta (FIB) aproxima-se de uma normal, com maior concentração de países entre a marca do 36 até 48, que formam quase um plateau, com os 4 maiores picos da distribuição. No estrato inferior, que vai até 5 pontos, figura também apenas um país :Botswana, com FIB = 22,59, seguida pelo segundo estrato, que

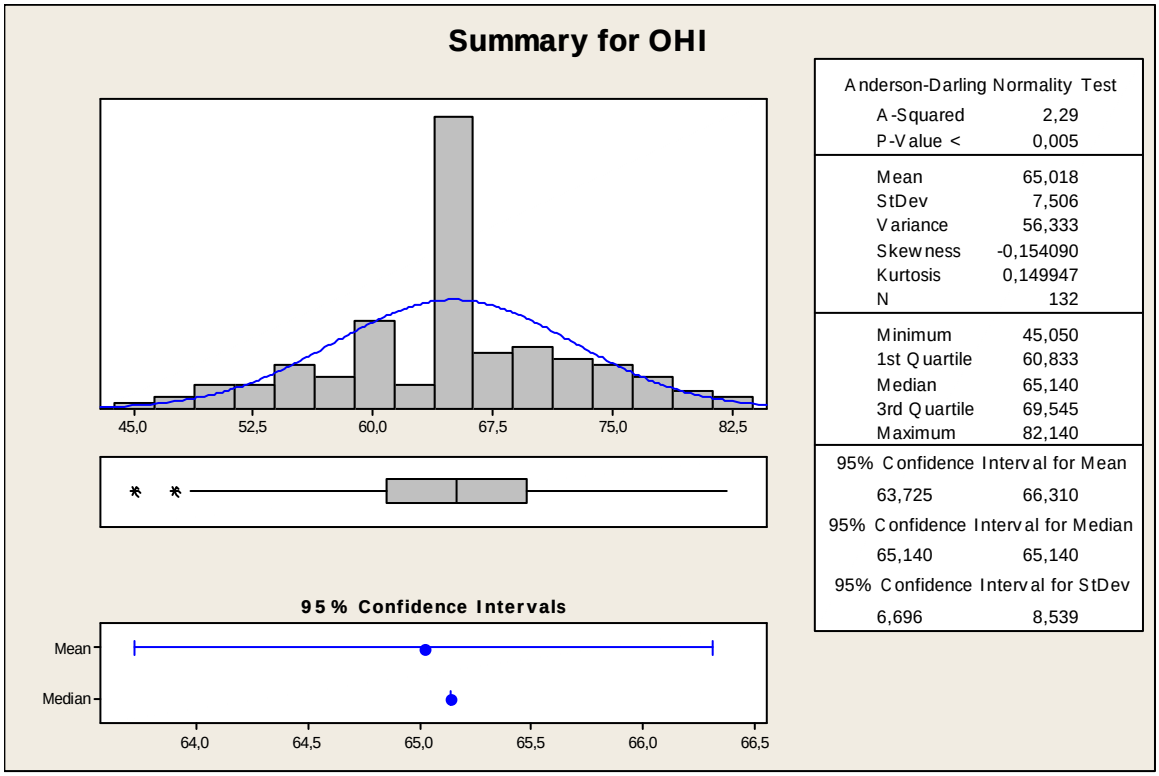
vai de 5 a 15 pontos e compreende 3 países. A menor pontuação é do Chade (FIB=24,68) e a maior é da Mali (FIB=26,04).

A maior concentração em número de países está alocada entre 36 e 48 pontos, sendo Mauritius (FIB=36,58), Dinamarca (FIB=36,61) e Ruanda (FIB=36,85) os piores colocados do grupo, e Tunisia (FIB=48,30), Moldova (FIB=47,96) e Reino Unido (FIB=47,93) os melhores colocados do grupo.

O Índice de Felicidade Bruta mede aspectos subjetivos e objetivos da vida humana, evidenciando pesquisas que apontam que, a partir de certo nível de renda, o nível de felicidade médio reportado não aumenta na mesma proporção do aumento da renda, até decaindo com o tempo. Outro tópico relaciona é a questão da resiliência construída por pessoas em situação de privação em países de baixo IDH e IPS, que acabam adaptando-se às situações adversas. Uma análise mais aprofundada poderá revelar o baixo nível de correlação deste indicados com o PIB e inclusive com o IDH e o IPS, pois sua metodologia é bastante distinta.

1.6.11 OHI

Pontos de referência para a realização de dez objetivos sócio ecológicos e como os países colocam em prática.



Análise dos dados

O gráfico apresenta um pico acentuado que contém concentração de 47 países o que representa 36% do total.

Há 2 outliers no ponto esquerdo gráfico, sendo o primeiro com apenas 1 país, a Nicarágua (45,05) e em seguida com 2 países sendo Libéria (47,54) e Angola (48,54).

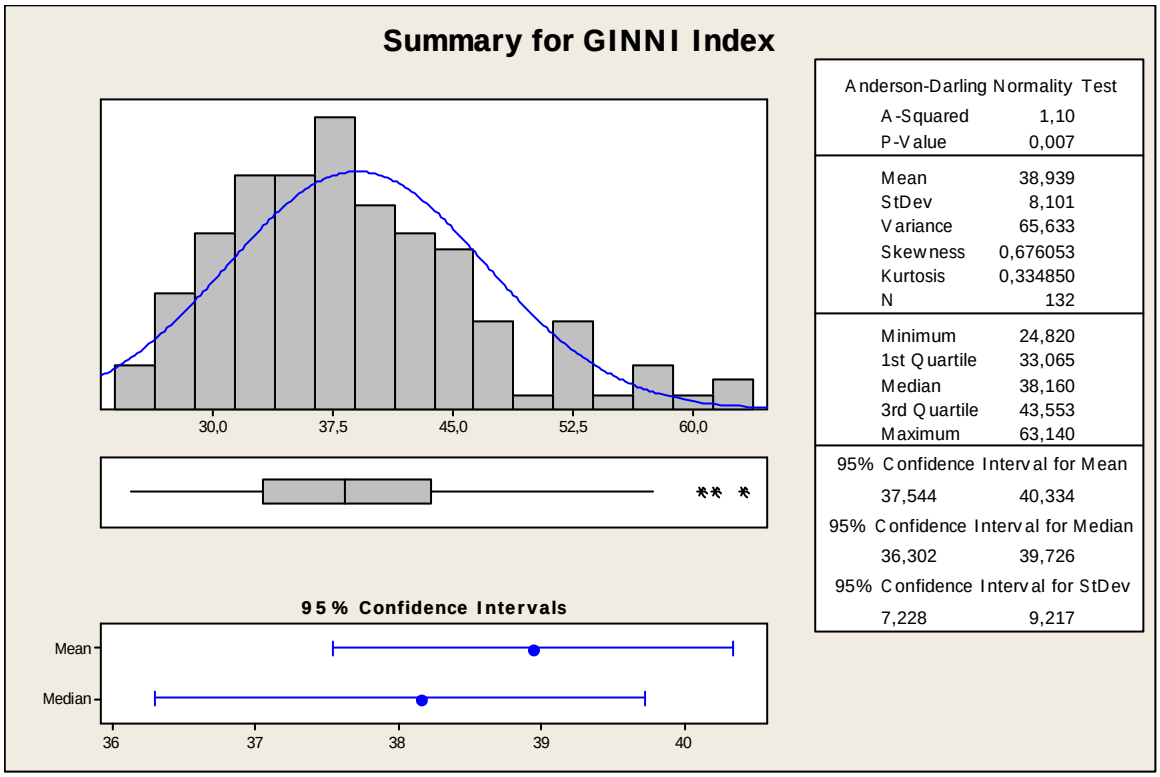
O gráfico para o Índice de Saúde dos Oceanos (OHI) chama atenção pela grande concentração de países entre o marco do 63 e 66 na escala normalizada (47 países) , o que equivaleria as pontuações de Kenia (64,12), Brasil (64,30) e Tunísia (64,38) no estrato inferior e Espanha (65,76) e Ucrânia (65,75) no estrato superior.

É preciso registrar, contudo, que faltavam dados para 35 países do OHI, e que conforme relatado acima, aos valores faltantes, foi atribuída a média das pontuações do OHI dos países disponível no momento de elaboração deste estudo. A média foi de 65,14.

Os melhores colocados foram a Dinamarca (82,14), Finlândia (81,4) e Noruega (80,11).

1.6.12 GINNI

Mede a distribuição de renda e de despesas das famílias



Análise dos dados

Para o Índice de Gini, que mede a desigualdade de renda, faltaram dados relativos a 13 países, que foram substituídos pela média da distribuição original, que foi de 38,94. Há uma concentração mais a esquerda com um pico que registra 20 países. O gráfico apresenta 3 outliers a direita, o mais extremo com 2 paises sendo África do Sul (63,140) e Namíbia (61,32) e seguido por Botswana (60,46).

2.RELAÇÃO ENTRE AS VARIÁVEIS

2.1 Correlação linear

A matriz de correlação linear apresenta o teste de significância *P-Value*. Para a correlação foi utilizado o índice de *Pearson*. Conforme Las Casas e Guevara (2010)[1], a correlação (valores entre -1 e 1) é expressada pelo grau de relação linear que existe entre elas. Valores próximos a zero significam nenhuma relação, e próximos a 1 e -1 mostram alta relação positiva ou negativa, respectivamente. Vale ressaltar que o índice de correlação entre as variáveis não requer, necessariamente, que exista uma relação de causa-efeito entre as variáveis para dar sentido às relações.

A seguir, apresentamos a matriz de correlação das variáveis estudadas, foram destacadas as correlações maiores de 0,8 para análise:

Correlations: Social Progr; IDH - 2013; 1. Governanç; 1.1 Institut; ...

	Social Progress	IDH - 2013	1. Governança
IDH - 2013	0,871		
	0,000		
1. Governança	0,890	0,699	
	0,000	0,000	
1.1 Institutions	0,781	0,604	0,910
	0,000	0,000	0,000
1 | Internationa	0,366	0,208	0,478
	0,000	0,017	0,000
1.1.| Effective	0,387	0,221	0,500
	0,000	0,011	0,000

[1] LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisas de Marketing. São Paulo: Atlas, 2010.

1.2.| Credibilit	0,366	0,216	0,499
	0,000	0,013	0,000
1.3.| Regional c	0,415	0,273	0,495
	0,000	0,002	0,000
EPI Score	0,917	0,826	0,811
	0,000	0,000	0,000
Happy Planet Ind	0,297	0,332	0,089
	0,001	0,000	0,311
OHI	0,559	0,487	0,623
	0,000	0,000	0,000
GINNI Index	-0,307	-0,305	-0,245
	0,000	0,000	0,005

	1.1 Institutions	1 | Internationa	1.1.| Effective
1 | Internationa	0,391		
	0,000		
1.1.| Effective	0,430	0,937	
	0,000	0,000	
1.2.| Credibilit	0,413	0,954	0,865
	0,000	0,000	0,000
1.3.| Regional c	0,405	0,883	0,741
	0,000	0,000	0,000
EPI Score	0,745	0,269	0,295
	0,000	0,002	0,001
Happy Planet Ind	-0,044	0,007	-0,005
	0,619	0,932	0,953
OHI	0,635	0,218	0,254
	0,000	0,012	0,003
GINNI Index	-0,293	0,122	0,120
	0,001	0,164	0,171

	1.2.| Credibilit	1.3.| Regional c	EPI Score
1.3.| Regional c	0,777		
	0,000		
EPI Score	0,264	0,337	
	0,002	0,000	
Happy Planet Ind	-0,004	0,051	0,197
	0,961	0,563	0,024
OHI	0,238	0,222	0,544
	0,006	0,011	0,000
GINNI Index	0,118	0,031	-0,364
	0,178	0,723	0,000

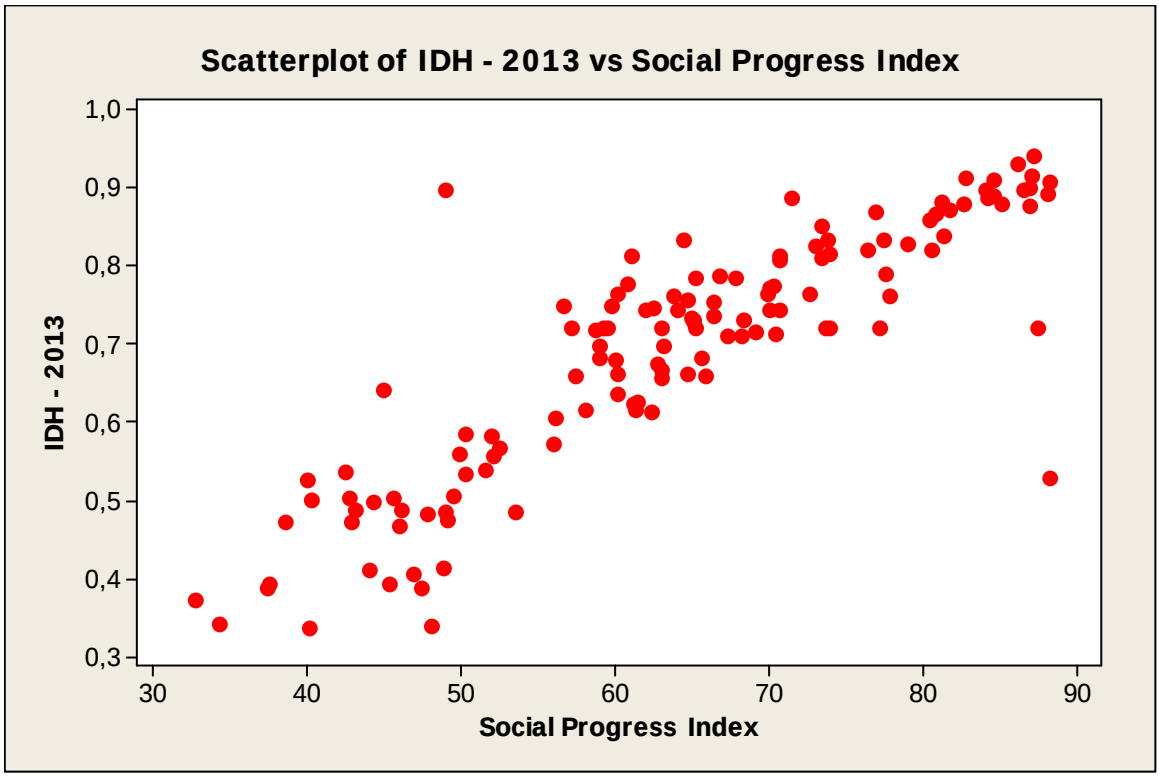
	Happy Planet Ind	OHI
OHI	-0,145	
	0,098	
GINNI Index	0,050	-0,289
	0,570	0,001

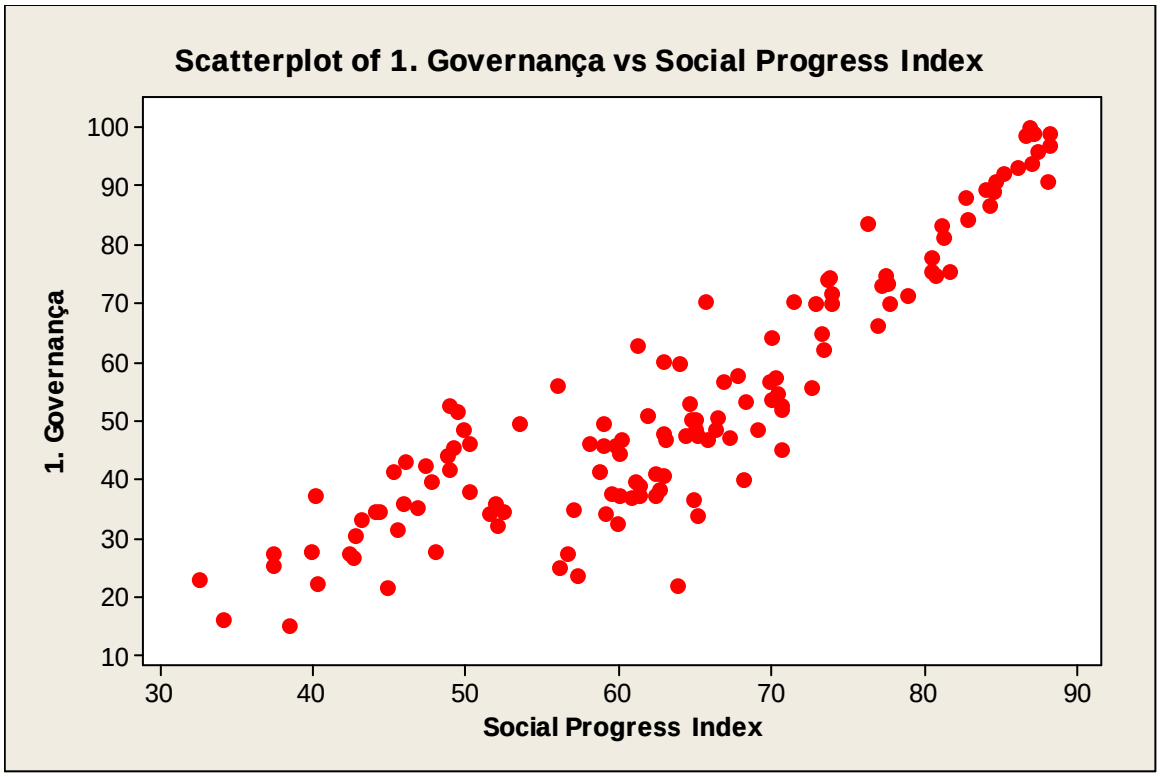
```
Cell Contents: Pearson correlation
               P-Value
```

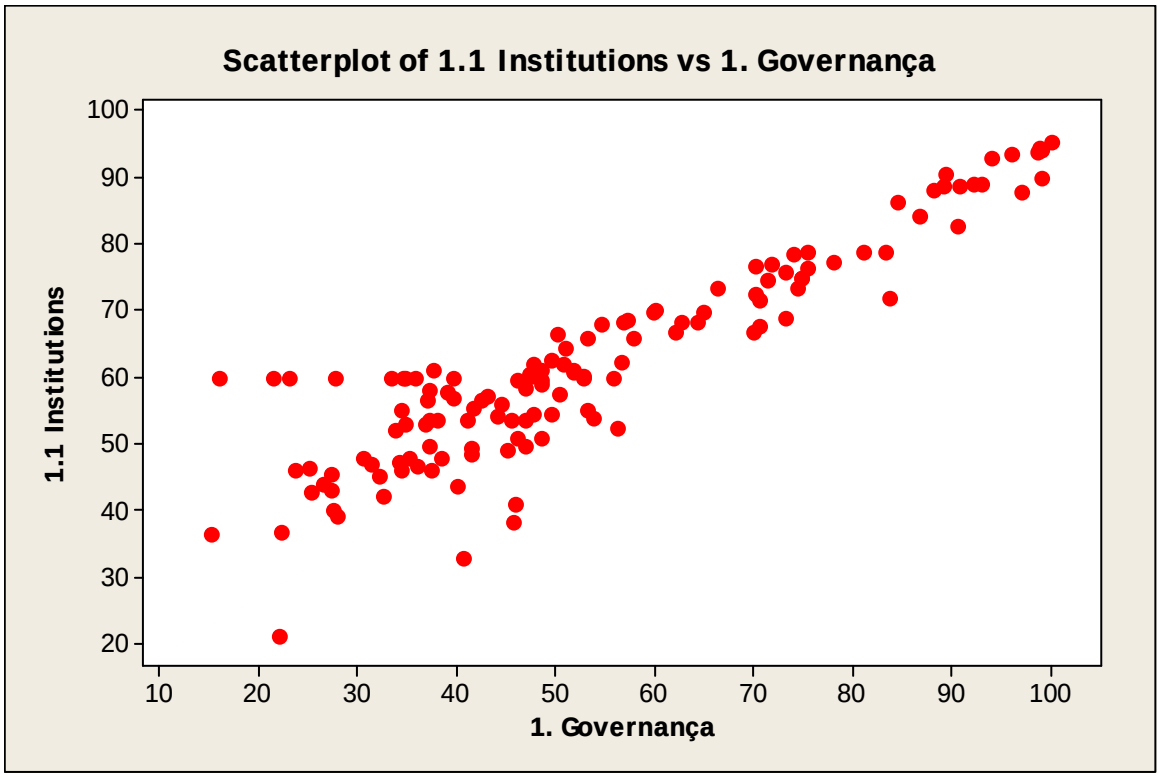
2.2 Gráficos de dispersão (Scatterplot)

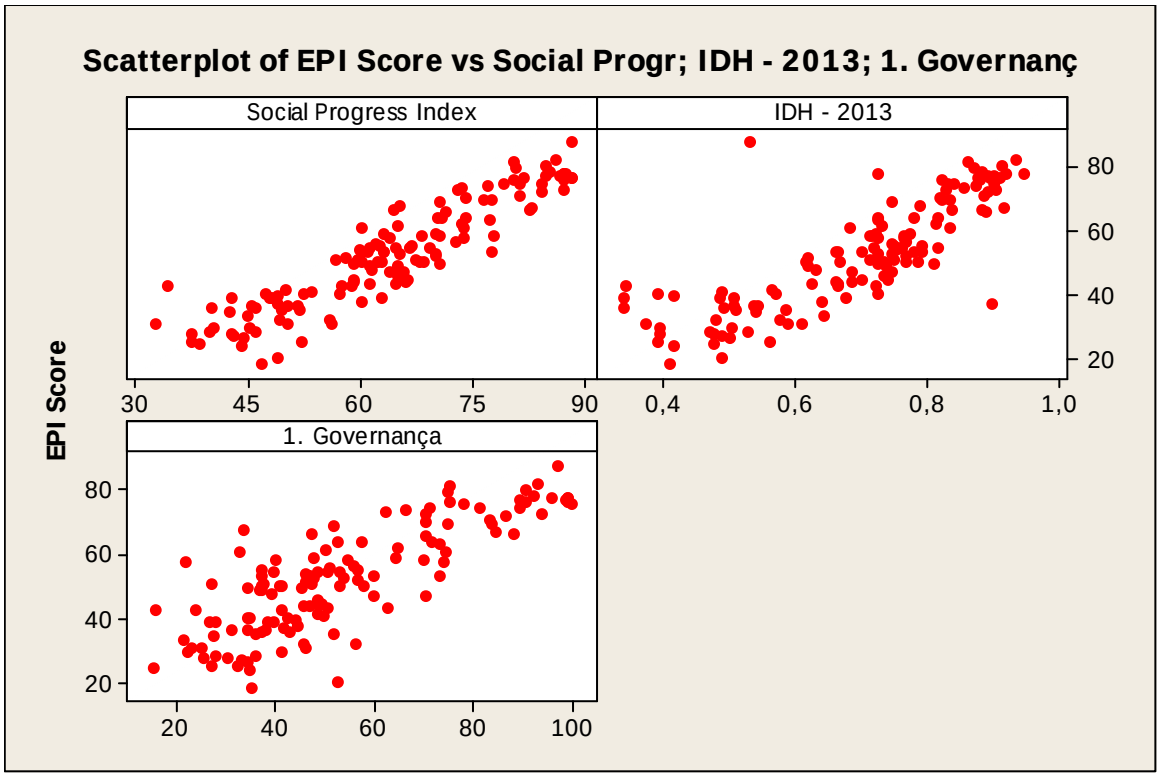
Para efeito de análise foi realizada a relação entre as variáveis com maior grau de correlação.

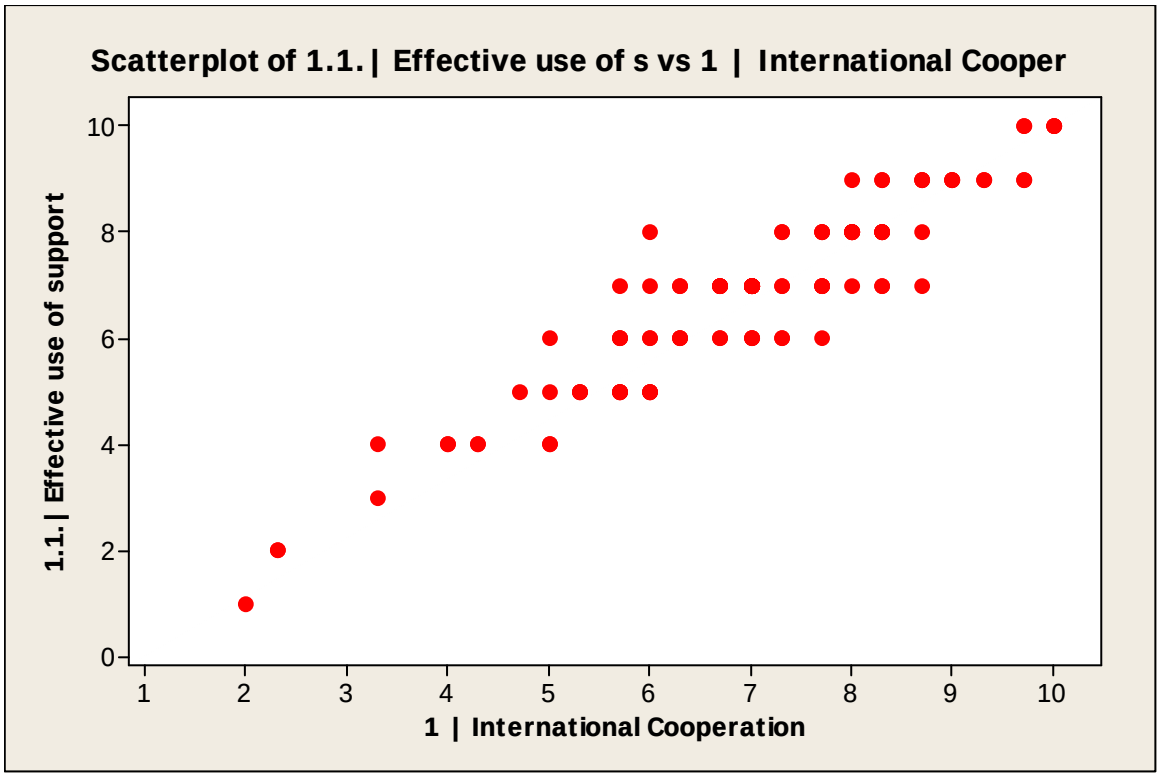
Todos os gráficos abaixo apresentam correlação linear positiva.

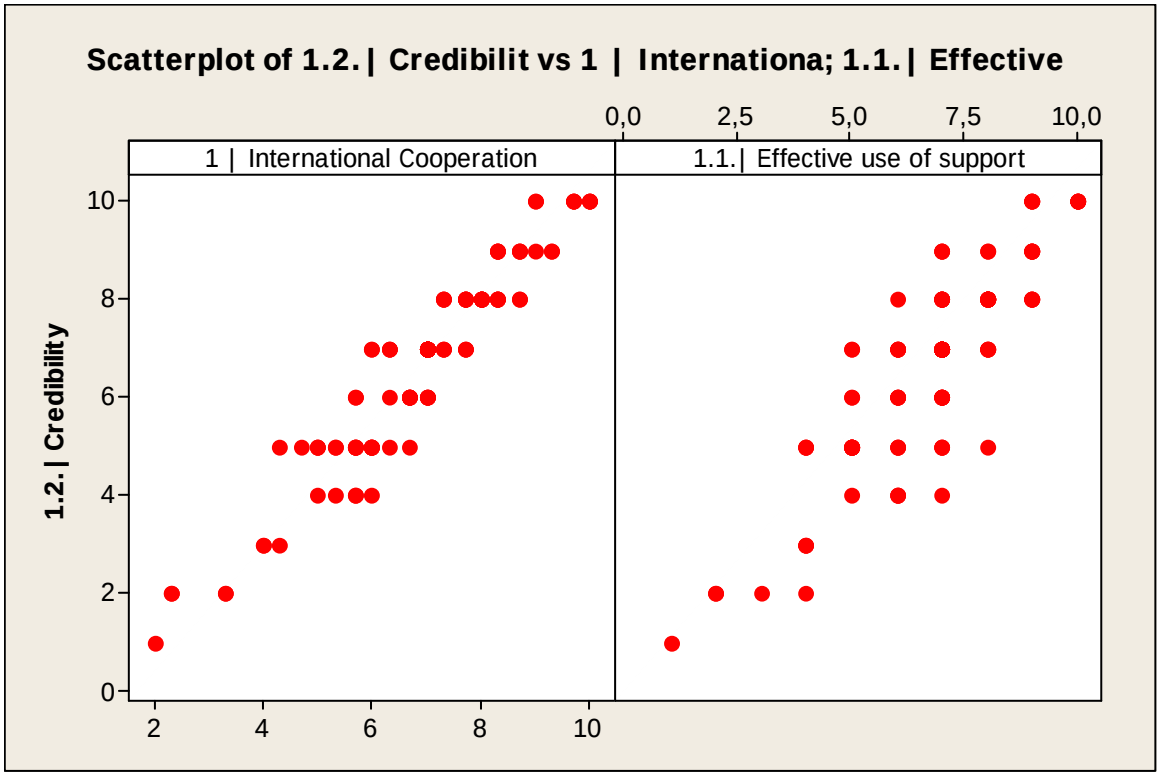



Scatterplot of 1. Governança vs Social Progress Index
1. Governança
Social Progress Index
100
90
80
70
60
50
40
30
20
10
30
40
50
60
70
80
90


Scatterplot of 1.1 Institutions vs 1. Governança
1.1 Institutions
1. Governança
100
90
80
70
60
50
40
30
20
10
20
30
40
50
60
70
80
90
100

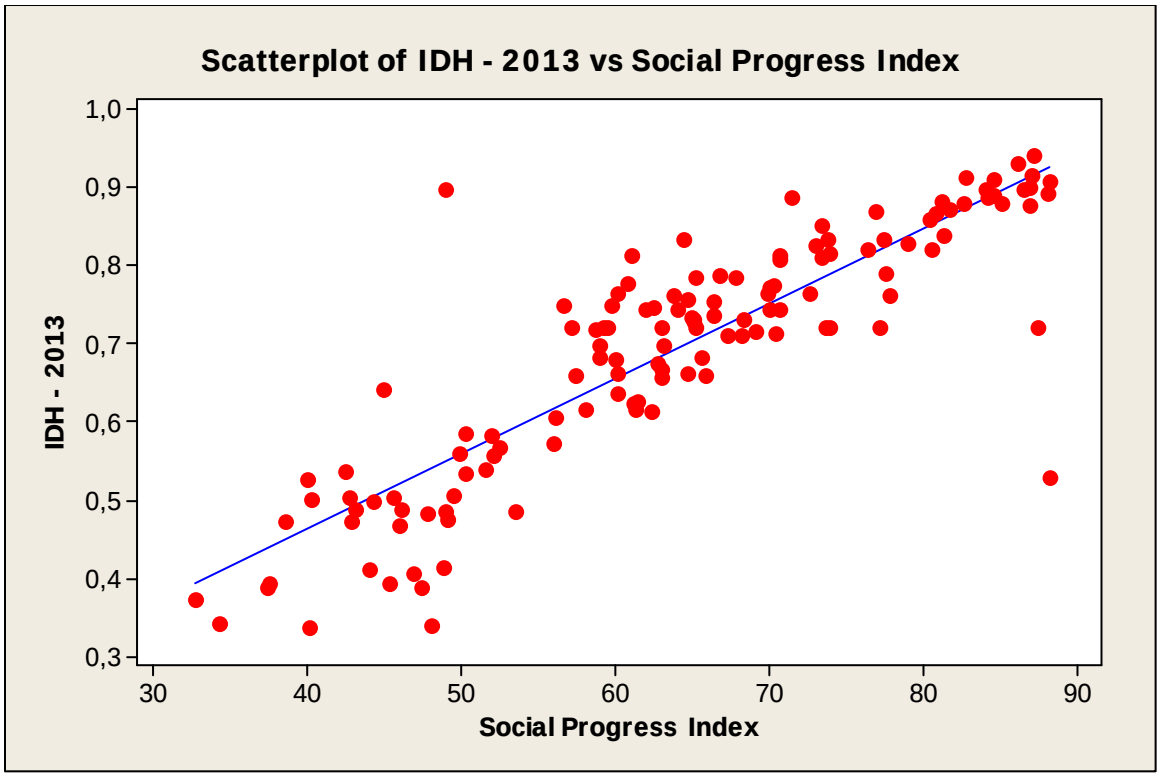

Scatterplot of EPI Score vs Social Progr; IDH - 2013; 1. Governanç
Social Progress Index
IDH - 2013
1. Governança
EPI Score
30
45
60
75
90
0,4
0,6
0,8
1,0
20
40
60
80
100


Scatterplot of 1.1. | Effective use of s vs 1 | International Cooper
1.1. | Effective use of support
1 | International Cooperation
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10



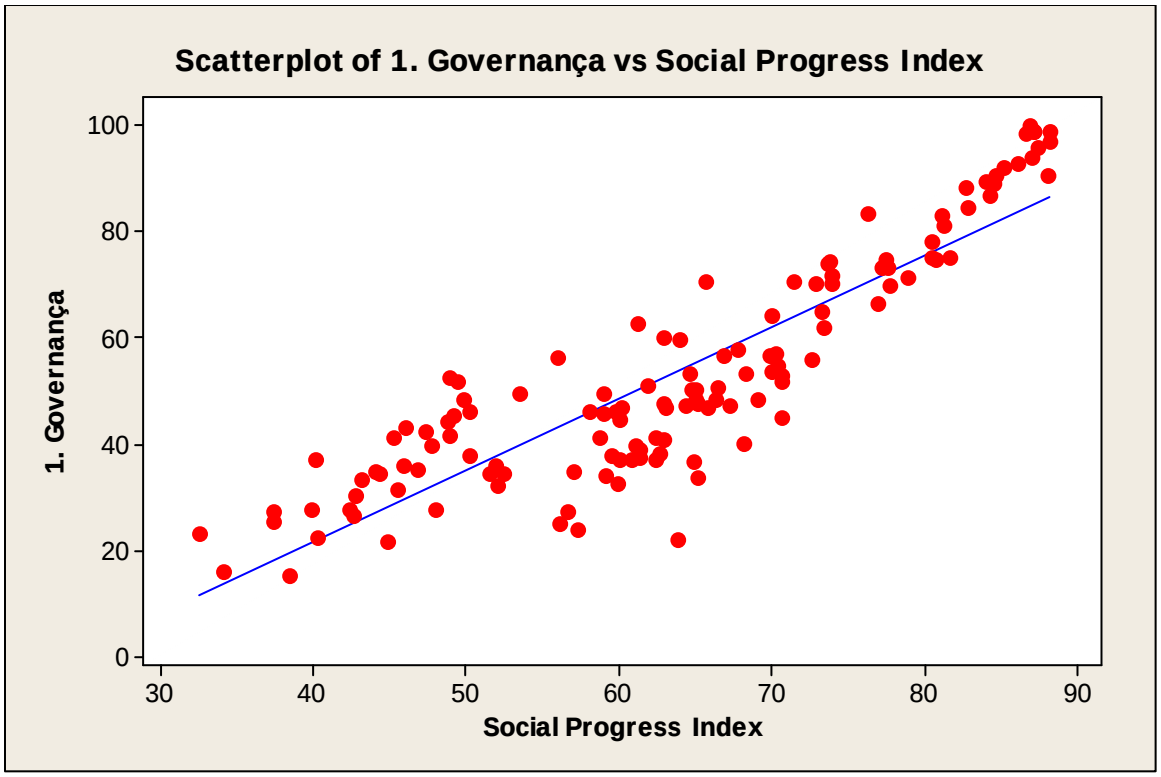
2.3 Linhas de Tendência

Gráfico1 – IDH 2013 vs SPI



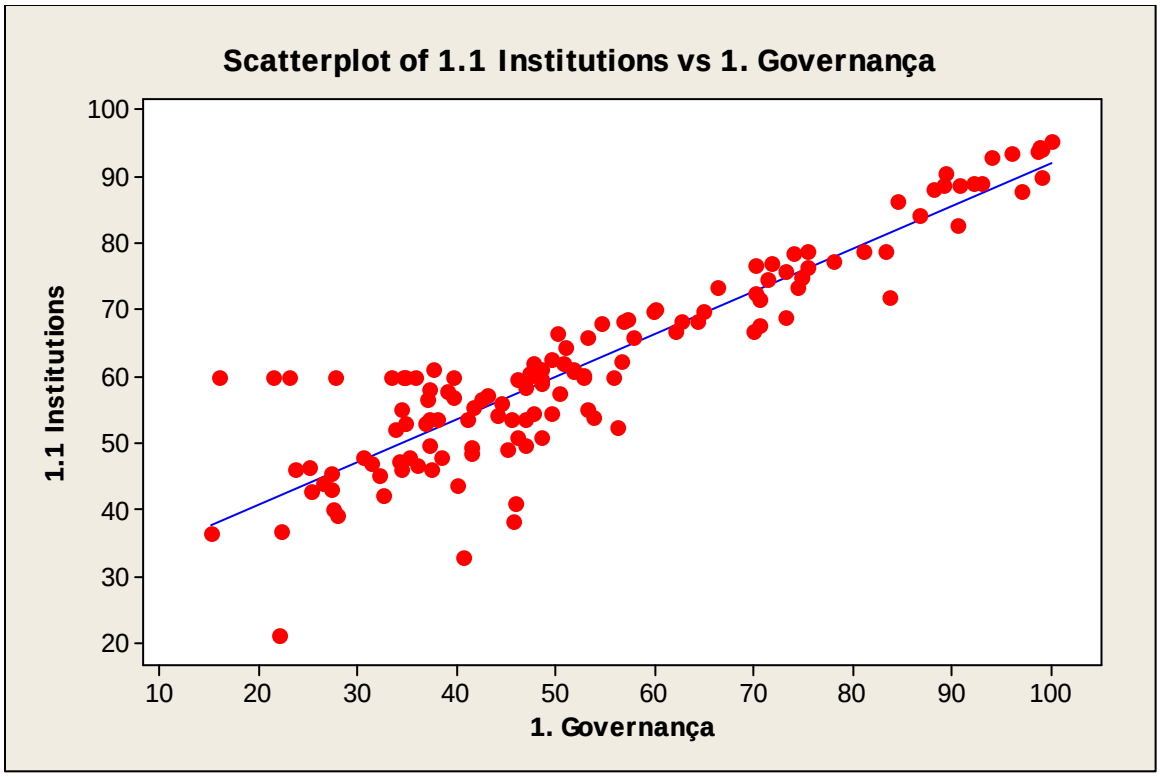
No gráfico 1 o ponto acima da reta é o país Swaziland e abaixo da reta são Holanda e Suiça, mais ao extremo à direita.

Gráfico 2 – Governança vs SPI

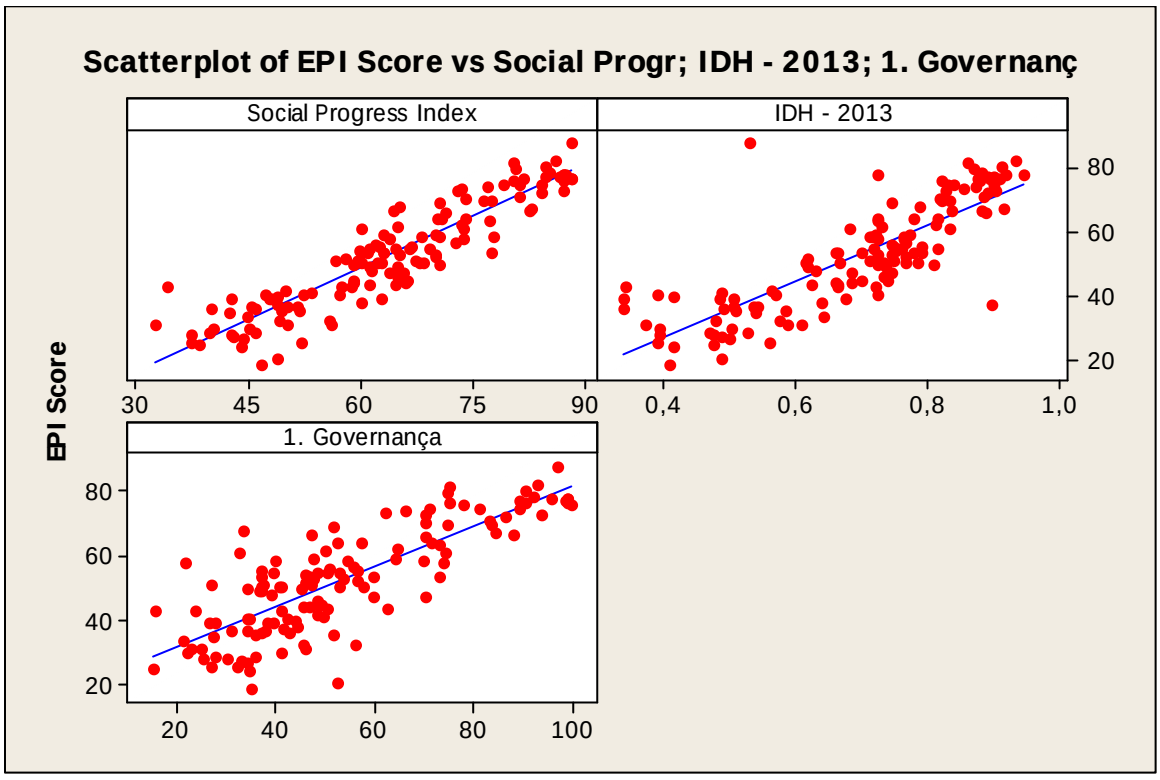


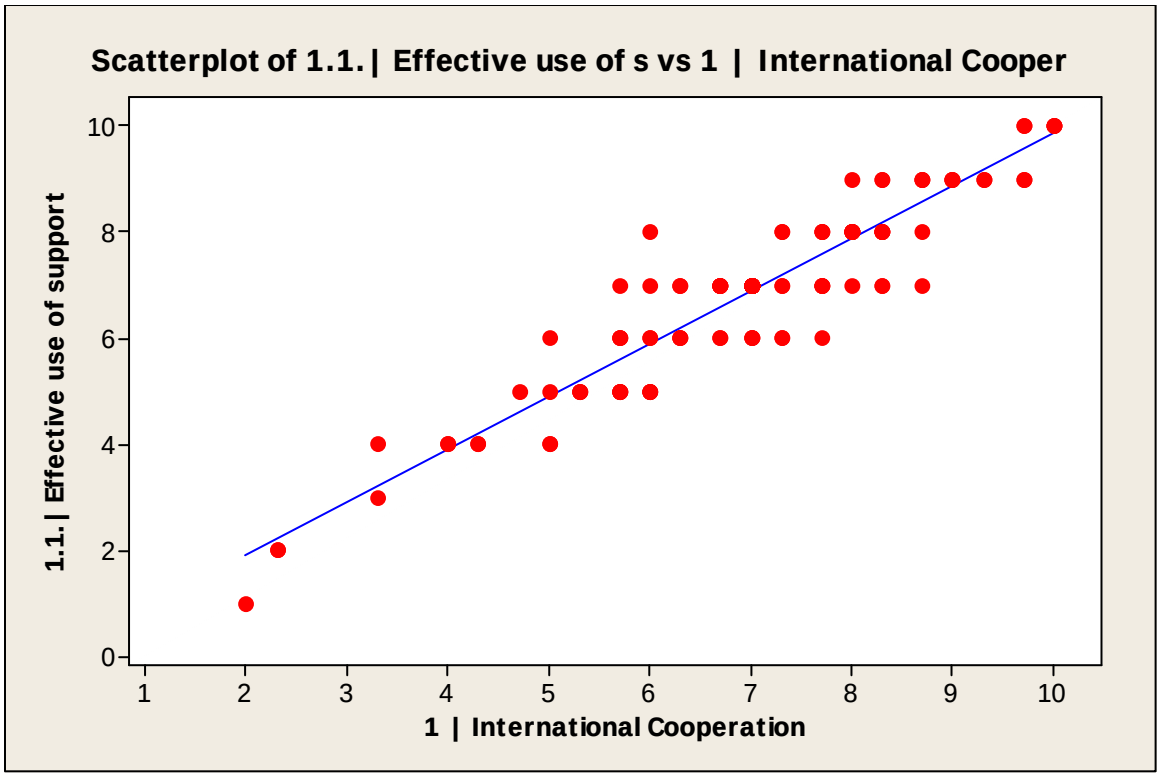
No gráfico 2 a dispersão aparece mais concentrada perto da reta, somente um destaque para o ponto abaixo da reta para a Venezuela.

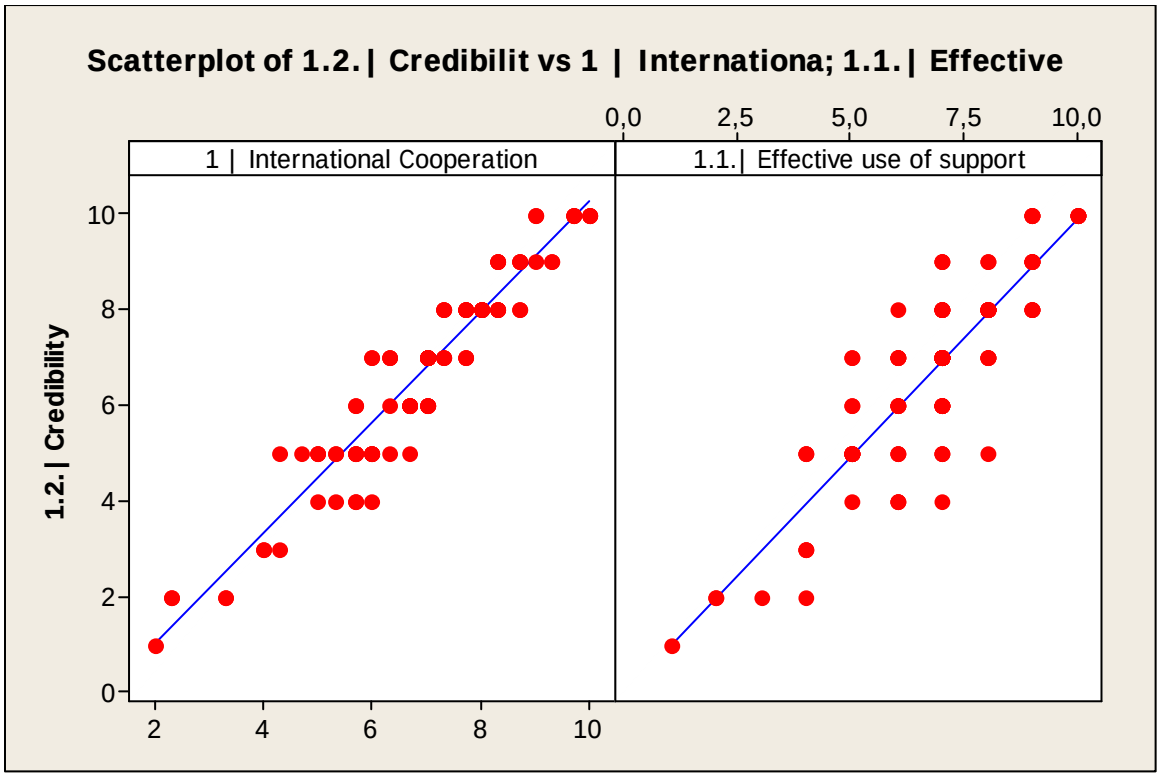
Gráfico 3 – Institutions vs Governança



No gráfico 3 os pontos acima da reta são República Central da África, Iraque, Chad e República do Congo (respectivamente). Abaixo da reta próximo ao zero, encontra-se respectivamente a Venezuela, Bolívia, Indonésia e Sri Lanka.



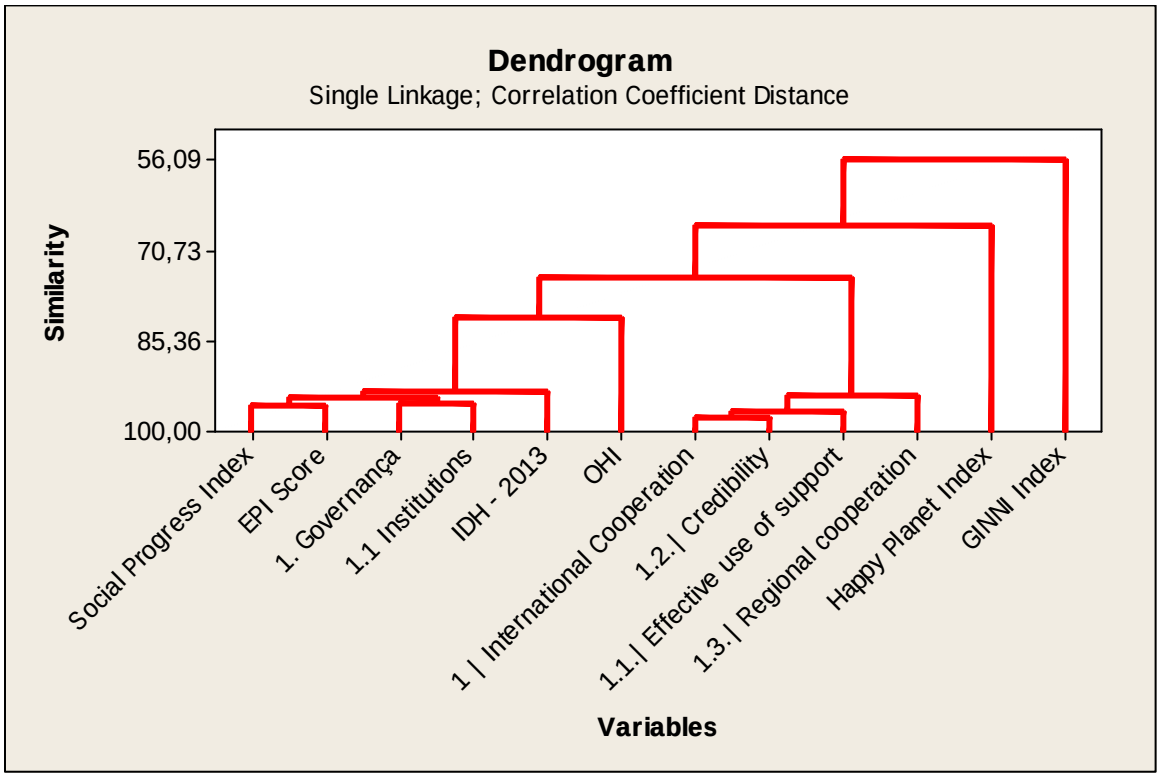

Scatterplot of 1.1.| Effective use of s vs 1 | International Cooper
1.1.| Effective use of support
1 | International Cooperation
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10


Scatterplot of 1.2.| Credibilit vs 1 | Internationa; 1.1.| Effective
1 | International Cooperation
1.1.| Effective use of support
1.2.| Credibility
0,0
2,5
5,0
7,5
10,0
0
2
4
6
8
10

2.2. Dendograma

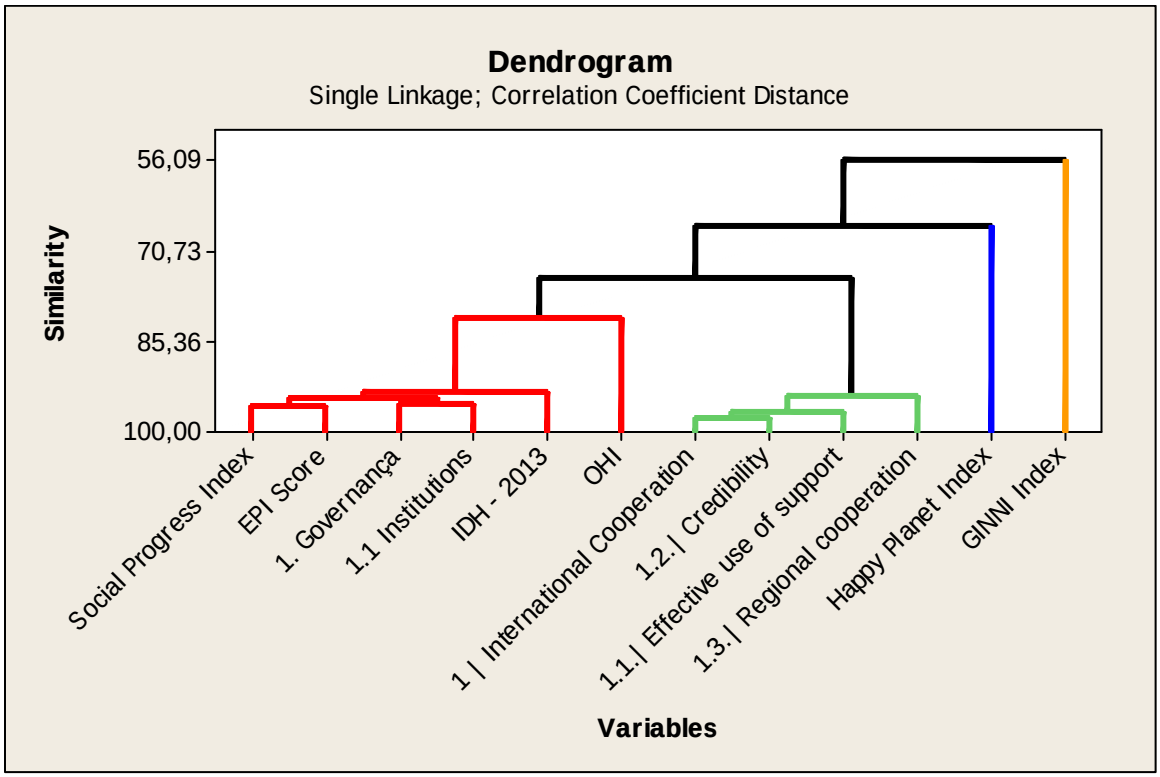
Um dendrograma ilustra as informações na tabela de amalgamação na forma de um diagrama de árvore.

Dendograma no 1 – 1 Cluster



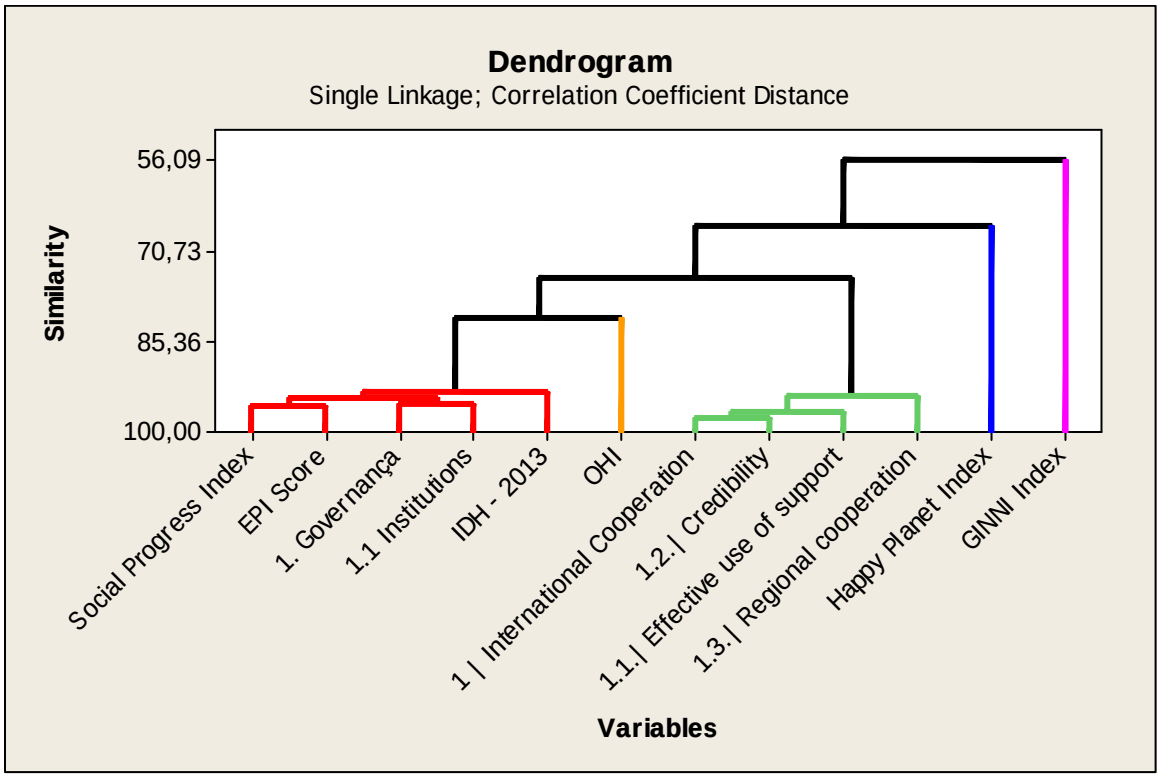
Separando as variáveis em *clusters*, como visualiza-se no dendograma 2 abaixo, e possível classificar as variáveis em 4 grupos de acordo com as suas similaridades. Mas a variável OHI está distante de grupo em vermelho que compreende as variáveis: ***SPI, EPI Score, Governança, Institutions e IDH 2013.***

Dendograma no 2 – 4 Clusters



Fazendo uma nova separação em 5 clusters, talvez demonstre melhor a similaridades entre as variáveis. Fica nítido a similaridades das variáveis que se encontram nos grupos vermelho e verde abaixo. As variáveis ***OHI, HPI e GINNI Index*** não apresentam um nível desejado de correlação com as demais variáveis, não oferecendo um poder explicativo

Dendograma no 5 – 5 Clusters

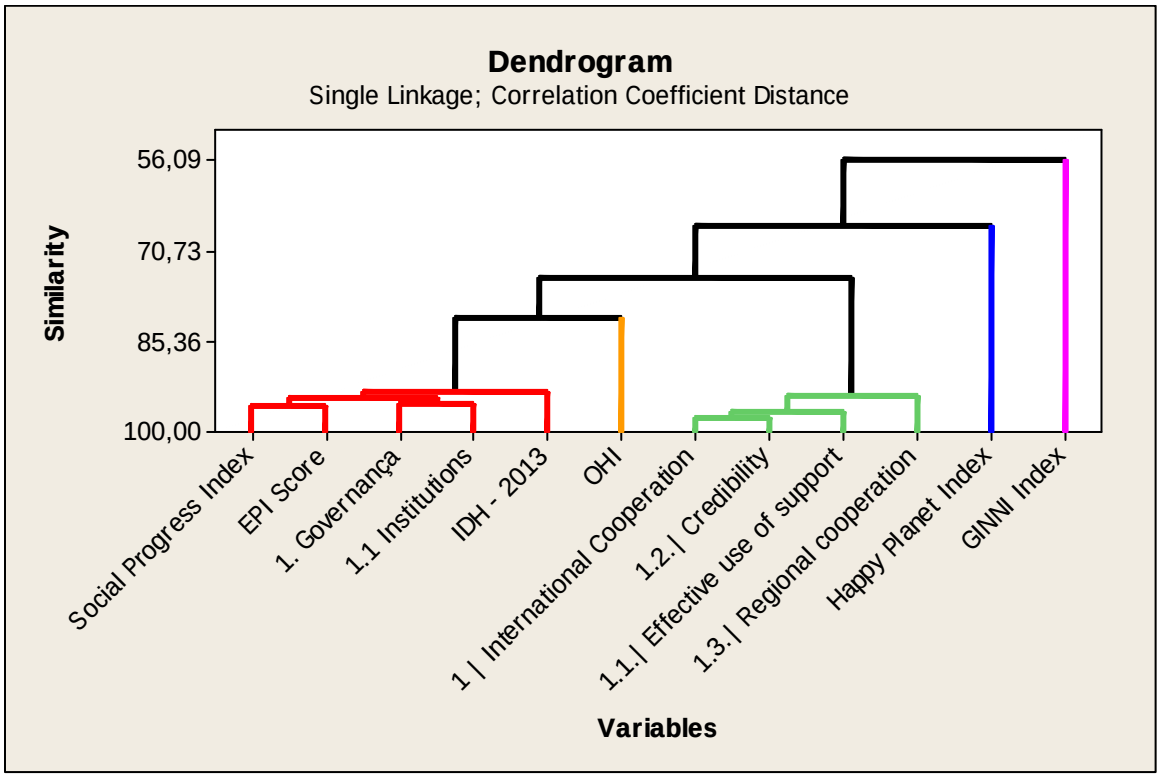


3. REGRESSÃO

O estudo que será feito é uma análise de regressão, que consiste em entender se há relação entre as variáveis e, se afirmativo, se a relação é fraca ou forte. Se a relação existir, será estabelecido um modelo que interprete a relação existente entre as variáveis e o modelo terá como objetivo a predição (FONSECA).[2]

Com base no dendograma de 5 clusters feito no tópico anterior, utilizaremos o agrupamento **vermelho** que há um grupo maior de variáveis. No grupo vermelho escolhido encontram-se as seguintes variáveis: Índice de Progresso Social, EPI (Índice de proteção ambiental), Governança, Instituições e IDH.

A variável escolhida é a **Institutions**, apresenta o quadro institucional de um país.



Assim será gerada uma primeira regressão com as variáveis citadas, conforme demonstrado abaixo:

[2] FONSECA, Jairo Simon,et al. Estatistica Aplicada. 2.ed. São Paulo: Atlas, 1995.

3.1 – Análise da Regressão

Regression Analysis: 1.1 Institut versus Social Progr; IDH - 2013; ...

```
The regression equation is
1.1 Institutions = 35,9 - 0,348 Social Progress Index - 1,36 IDH - 2013
                   + 0,727 1. Governança + 0,198 EPI Score


Predictor                 Coef  SE Coef      T      P
Constant                35,925    3,329  10,79  0,000
Social Progress Index  -0,3482   0,1496  -2,33  0,022
IDH - 2013              -1,363    7,503  -0,18  0,856
1. Governança          0,72653  0,05809  12,51  0,000
EPI Score              0,19786  0,08256   2,40  0,018


S = 6,16181   R-Sq = 84,0%   R-Sq(adj) = 83,5%


Analysis of Variance

Source           DF       SS      MS       F      P
Regression        4  25261,7  6315,4  166,34  0,000
Residual Error  127   4821,9    38,0
Total           131  30083,6


Source                 DF   Seq SS
Social Progress Index   1  18331,0
IDH - 2013              1    720,9
1. Governança           1   5991,8
EPI Score               1    218,1


Unusual Observations

       Social
     Progress           1.1
Obs     Index  Institutions     Fit  SE Fit  Residual  St Resid
 13      62,9        32,700  52,647   0,908   -19,947     -3,27R
 23      34,2        59,800  43,674   2,301    16,126      2,82RX
 24      32,6        59,800  46,928   1,744    12,872      2,18R
 28      48,0        59,800  46,723   1,870    13,077      2,23R
 53      59,0        38,100  56,439   0,691   -18,339     -3,00R
 55      44,8        59,800  41,609   1,286    18,191      3,02R
111      59,7        40,900  58,157   0,959   -17,257     -2,84R
113      48,9        55,300  55,319   2,918    -0,019     -0,00 X
115      88,2        87,600  92,339   3,211    -4,739     -0,90 X
130      63,8        21,100  40,099   1,849   -18,999     -3,23R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Analisando as variáveis preditoras a IDH apresenta um grau alto de imprecisão (0,856), sendo que para este estudo é aceitável até 10%. Assim, uma nova regressão será rodada retirando está variável.

Regression Analysis: 1.1 Institut versus Social Progr; 1. Governanç; ...

```
The regression equation is
1.1 Institutions = 35,9 - 0,364 Social Progress Index + 0,730 1. Governança
                   + 0,196 EPI Score


Predictor                Coef  SE Coef      T      P
Constant               35,903    3,314  10,83  0,000
Social Progress Index  -0,3639   0,1216  -2,99  0,003
1. Governança          0,73010  0,05446  13,41  0,000
EPI Score              0,19579  0,08145   2,40  0,018


S = 6,13849   R-Sq = 84,0%   R-Sq(adj) = 83,6%


Analysis of Variance

Source           DF       SS      MS       F      P
Regression        3  25260,4  8420,1  223,46  0,000
Residual Error  128   4823,2    37,7
Total           131  30083,6


Source                 DF   Seq SS
Social Progress Index   1  18331,0
1. Governança           1   6711,8
EPI Score               1    217,7


Unusual Observations

       Social
     Progress           1.1
Obs     Index  Institutions     Fit  SE Fit  Residual  St Resid
 13      62,9        32,700  52,585   0,838   -19,885     -3,27R
 23      34,2        59,800  43,547   2,184    16,253      2,83RX
 24      32,6        59,800  46,917   1,737    12,883      2,19R
 28      48,0        59,800  46,424   0,884    13,376      2,20R
 53      59,0        38,100  56,493   0,622   -18,393     -3,01R
 55      44,8        59,800  41,764   0,960    18,036      2,97R
111      59,7        40,900  58,270   0,727   -17,370     -2,85R
130      63,8        21,100  40,073   1,837   -18,973     -3,24R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Nesta regressão o R2, que é o coeficiente de explicação ou de determinação não foi alterado, sendo 84%. Quanto mais próximo de 1 for R2, mais próximos da reta de regressão as variáveis estiverem, mais serão explicadas. Neste caso, 84% das variações de **Institutions** são explicadas pelas outras variáveis.

Uma outra análise dessa nova regressão foi a mudança do T, que significa a precisão dessas variáveis, neste caso a que a apresentou uma maior T e um aumento de uma regressão para a outra é a variável de **Governança.** Talvez pela similaridade que ela tem em relação a variável escolhida.

Agora será realizada uma regressão stepwise, ou seja, este tipo de análise permite demonstrar as variáveis preditoras mais uteis para esta relação.

3.2 – Análise da Regressão Stepwise

Stepwise Regression: 1.1 Institut versus Social Progr; 1. Governanç; ...

```
  Alpha-to-Enter: 0,15  Alpha-to-Remove: 0,15


Response is 1.1 Institutions on 3 predictors, with N = 132


Step                        1       2       3
Constant                27,88   32,74   35,90

1. Governança           0,639   0,727   0,730
T-Value                 25,05   13,11   13,41
P-Value                 0,000   0,000   0,000

Social Progress Index          -0,150  -0,364
T-Value                         -1,78   -2,99
P-Value                         0,078   0,003

EPI Score                               0,196
T-Value                                  2,40
P-Value                                 0,018

S                        6,30    6,25    6,14
R-Sq                    82,83   83,24   83,97
R-Sq(adj)               82,70   82,98   83,59
Mallows Cp                9,1     7,8     4,0
```

Somente o *P-value* de **Governança** foi de 0 (zero), isto demonstra que os dados são estatisticamente significativos.

Todas as variáveis possuem um R2 alto (acima de 82,5), isto demonstra um alto grau de explicação entre elas e não sendo necessário uma nova regressão.

3.3 – Ajuste das bases

No início deste trabalho alguns países tiveram que ter variáveis preenchidas com a média geral por causa da falta de informação de tais dados, para poder dar continuidade nas análises.

Agora após as regressões será feito um ajuste das bases para depois serem normalizadas e assim ter uma base mais consistente para outras análises.

Unusual Observations

Obs	Social Progress Index	1.1 Institutions	Fit	SE Fit	Residual	St Resid
13	62,9	32,700	52,585	0,838	-19,885	-3,27R
23	34,2	59,800	43,547	2,184	16,253	2,83RX
24	32,6	59,800	46,917	1,737	12,883	2,19R
28	48,0	59,800	46,424	0,884	13,376	2,20R
53	59,0	38,100	56,493	0,622	-18,393	-3,01R
55	44,8	59,800	41,764	0,960	18,036	2,97R
111	59,7	40,900	58,270	0,727	-17,370	-2,85R
130	63,8	21,100	40,073	1,837	-18,973	-3,24R

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large leverage.

Conforme a tabela acima os ajustes foram feitos nos seguintes países na ordem apresentada pela tabela acima sendo: Bolívia, República Central da África, Chad, República do Congo, Indonésia, Iraque, Sri Lanka e Venezuela.

4. TESTE DE COMPARAÇÕES

O trabalho em questão tem como objetivo realizar um teste de comparações com o conjunto de dados sobre o tema: **integração e alianças estratégicas** empregando software estatístico MINITAB.

As variáveis já foram normalizadas e será utilizado a análise de variância ANOVA (Analise of variance) do MINITAB.

A análise de variância (ANOVA) testa a hipótese de que as médias de duas ou mais populações são iguais. Análises ANOVA testam a importância de um ou mais fatores comparando as médias das variáveis de resposta em diferentes níveis dos fatores. A hipótese nula afirma que todas as médias das populações (médias dos níveis dos fatores) são iguais, enquanto a hipótese alternativa afirma que pelo menos uma é diferente (MINITAB 17).

Para efetuar uma ANOVA, é necessário haver uma variável de resposta contínua e pelo menos um fator categórico com dois ou mais níveis. As análises ANOVA exigem dados de populações aproximadamente normalmente distribuídas com variâncias iguais entre fatores.

No MINITAB será utilizado o ANOVA – one way, que significa analisar se as médias de dois ou mais grupos se diferentes.

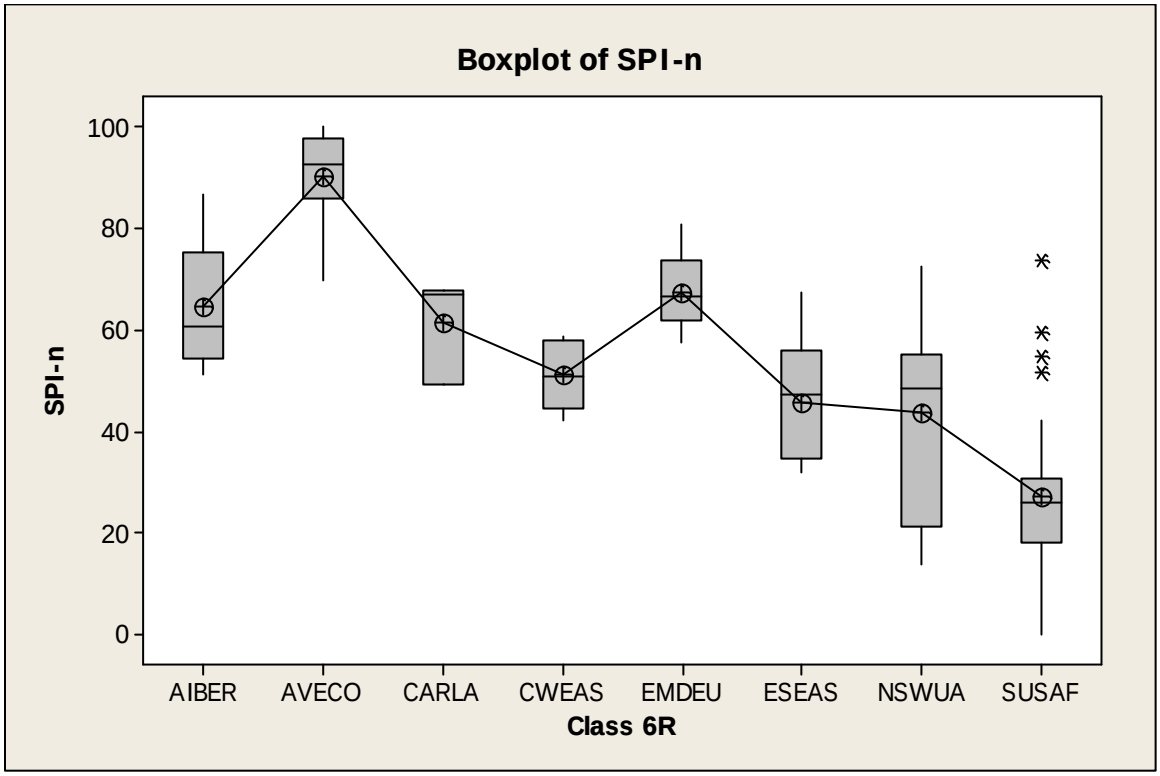
Serão feitas duas análises de variância, com diferentes fatores, sendo:

- **1ª análise - Fator Classificação 6R:** 6 regiões
- **2ª análise – Fator Classificação 3R:** 3 regiões

As representações gráficas destas análises serão feitas através de gráficos tipo Boxplot. Um boxplot é um sumário gráfico da distribuição de uma amostra que exibe sua forma, tendência central e variabilidade.

4.1 ANOVA – FATOR Classificação 6R e 3R

4.2 Análise SPI



Analisando o Boxplot do SPI, temos a mediana (ponto de destaque de cada caixa) sendo apresentada de forma diferente para cada região destacada. O retângulo compreende em 50% dos valores dos conjuntos dos dados.

A região AVECO é a melhor colocada, pois está mais próximo de 100. Esta região está representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 3º quartil, neste caso é **negativamente assimétrico.**

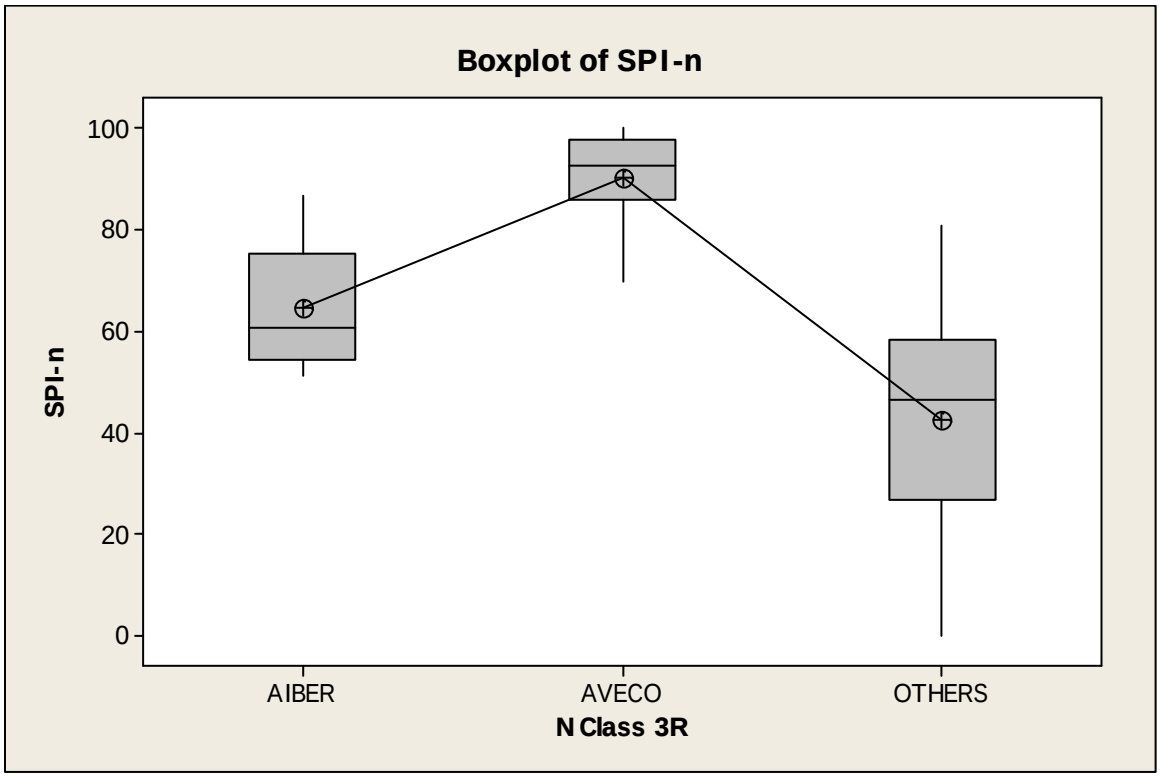
No caso da região SUSAF há 4 outliers, sendo os seguintes países de cima para baixo respectivamente Mauritius, Botswana, Africa do Sul e Namíbia, destacando com o SPI mais alto do que os demais deste grupo.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o SPI tem significância sobre a análise, sendo o modelo explica 76,61% da variância.

Results for: CODING REGIONS MINITAB.MTW

One-way ANOVA: SPI-n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  65372  9339  58,02  0,000
Error     124  19961   161
Total     131  85333

S = 12,69   R-Sq = 76,61%   R-Sq(adj) = 75,29%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---------+---------+---------+---------+
AIBER  21  64,41  11,95                   (--*--)
AVECO  27  90,30   8,70                                (-*--)
CARLA   3  61,42  10,46             (-------*------)
CWEAS  11  51,27   6,12            (---*--)
EMDEU  12  67,34   7,04                     (---*--)
ESEAS  12  45,57  11,72          (---*--)
NSWUA  14  43,75  18,74          (--*--)
SUSAF  32  27,04  15,99  (--*-)
                          ---------+---------+---------+---------+
                                  40        60        80       100

Pooled StDev = 12,69
```

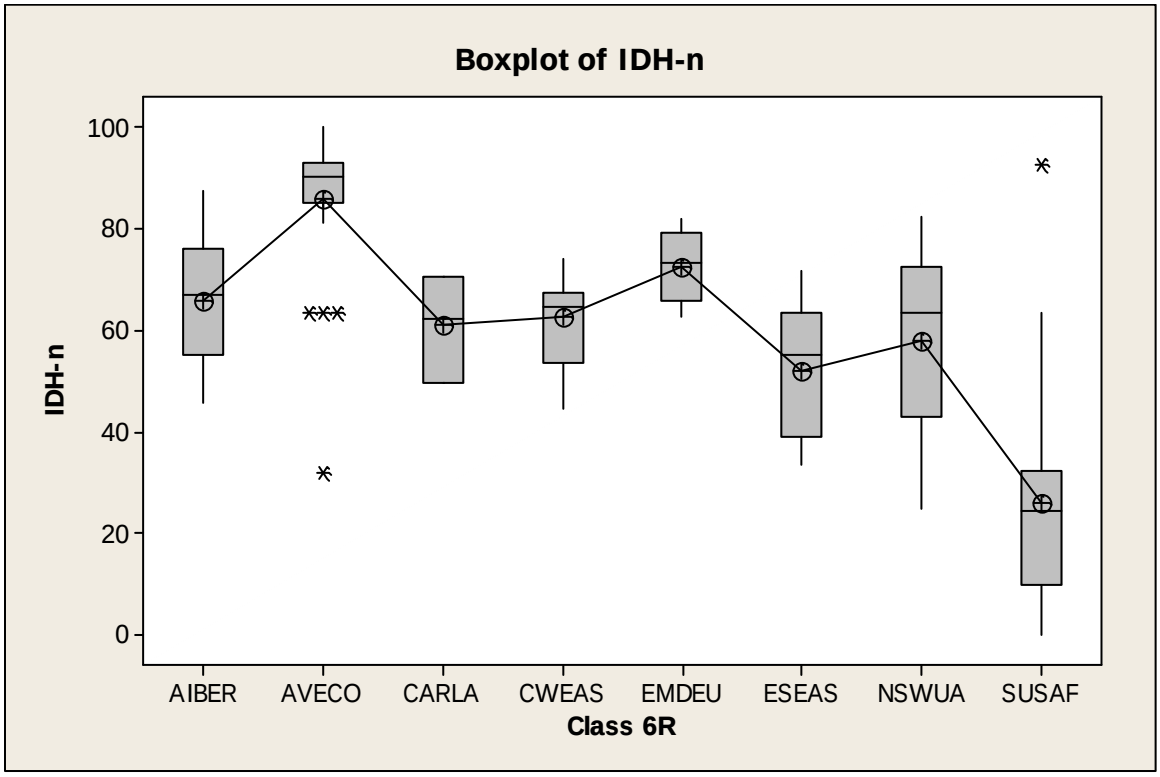


One-way ANOVA: SPI-n versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  48262  24131  83,97  0,000
Error       129  37071    287
Total       131  85333

S = 16,95   R-Sq = 56,56%   R-Sq(adj) = 55,88%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ----+---------+---------+---------+-----
AIBER   21  64,41  11,95              (----*----)
AVECO   27  90,30   8,70                                (---*----)
OTHERS  84  42,63  19,71  (-*--)
                          ----+---------+---------+---------+-----
                             45        60        75        90

Pooled StDev = 16,95
```

4.3 Análise IDH



A região AVECO é a melhor colocada, pois está mais próximo de 100. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**. E há 4 outliers abaixo do 3º quartil sendo: República da Korea, Latvia, Holanda e abaixo Suiça.

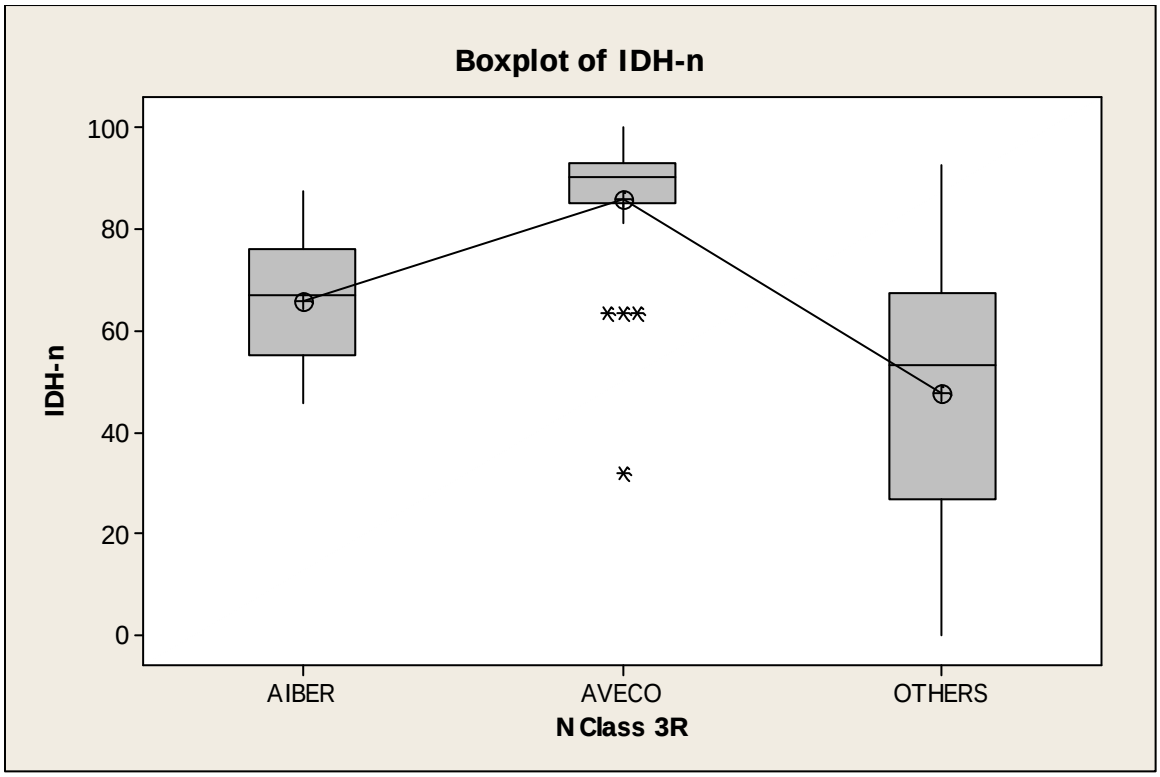
No caso da região SUSAF há 1 outlier acima do 1º quartil a Swaziland.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o IDH tem significância sobre a análise, sendo o modelo explica 66,56% da variância.

One-way ANOVA: IDH-n versus Class 6R

```
Source      DF      SS     MS      F      P
```

```
Class 6R    7  58225  8318  35,25  0,000
Error     124  29258   236
Total     131  87484

S = 15,36   R-Sq = 66,56%   R-Sq(adj) = 64,67%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  21  65,62  12,03                   (---*--)
AVECO  27  85,71  14,51                            (--*--)
CARLA   3  60,85  10,62             (-------*--------)
CWEAS  11  62,71   8,84                  (---*----)
EMDEU  12  72,60   6,81                       (---*----)
ESEAS  12  52,16  13,52             (---*---)
NSWUA  14  57,97  19,28                (---*---)
SUSAF  32  25,86  20,10   (--*--)
                           +---------+---------+---------+---------
                          20        40        60        80

Pooled StDev = 15,36
```

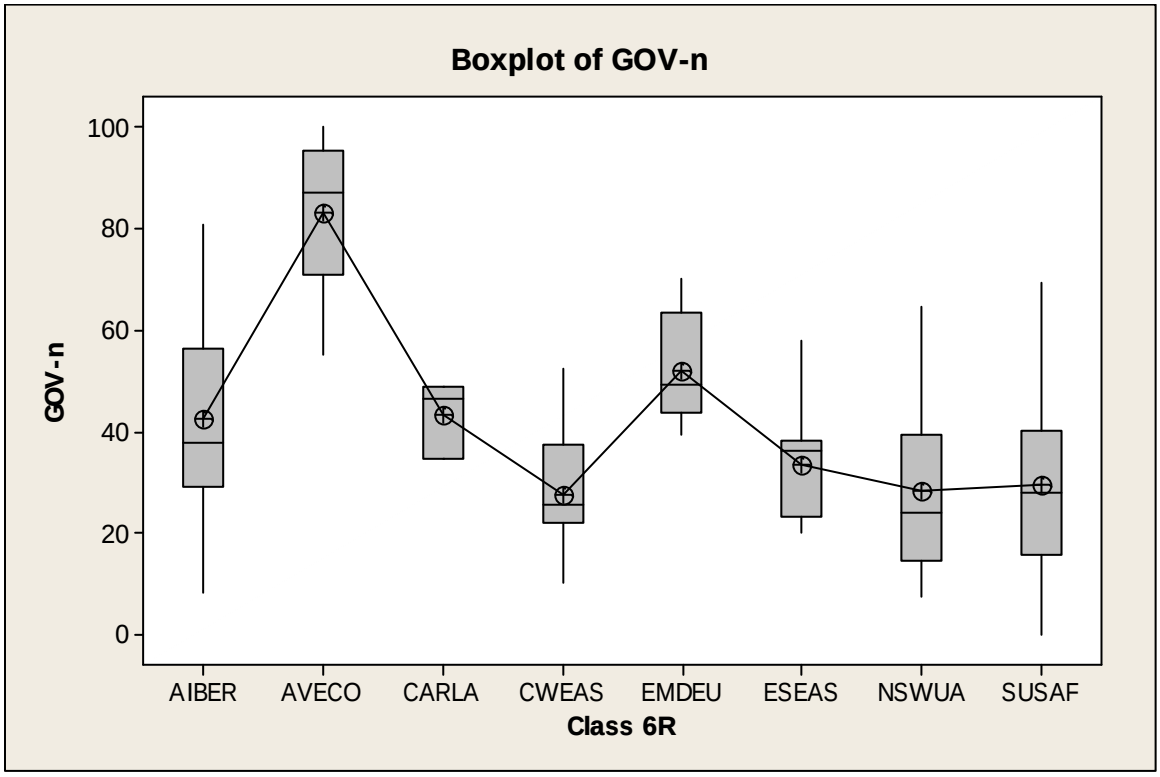


One-way ANOVA: IDH-n versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  30813  15407  35,07  0,000
Error       129  56670    439
Total       131  87484

S = 20,96   R-Sq = 35,22%   R-Sq(adj) = 34,22%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
```

```
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  65,62  12,03           (-----*-----)
AVECO   27  85,71  14,51                      (----*----)
OTHERS  84  47,72  24,12  (--*--)
                          -+---------+---------+---------+--------
                          45        60        75        90

Pooled StDev = 20,96
```

4.4 Análise GOVERNANÇA



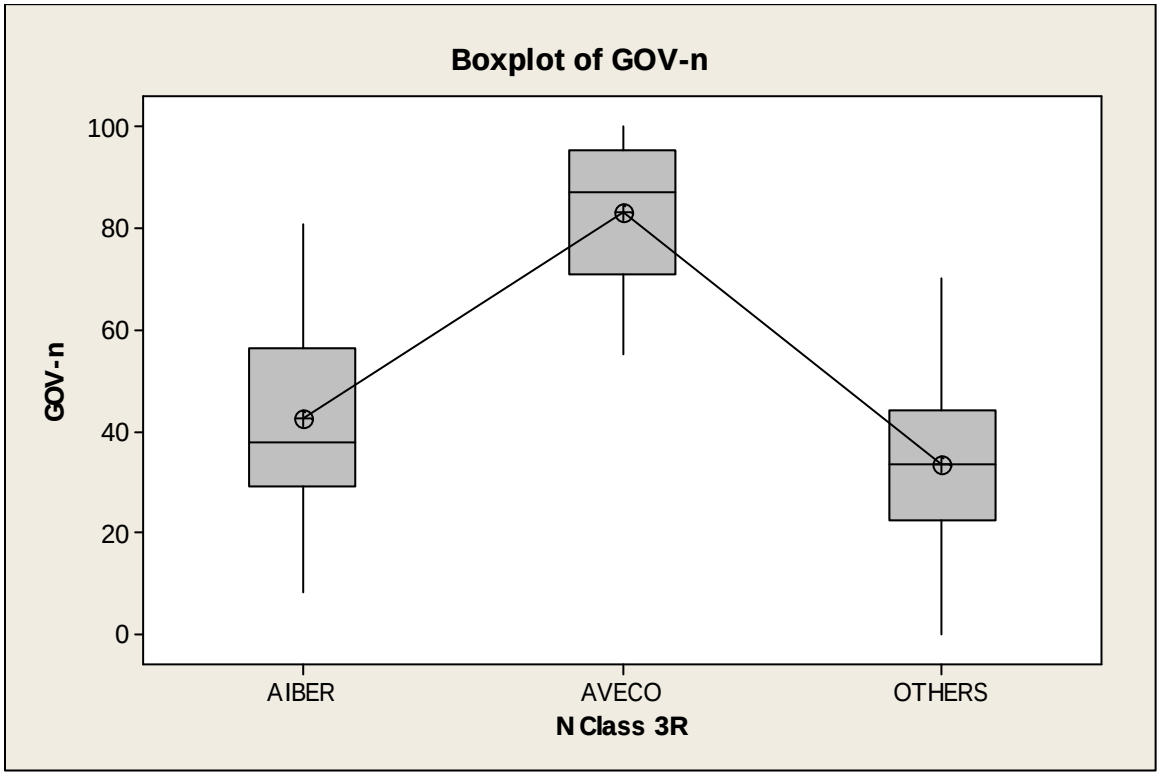
A região AVECO é a melhor colocada, pois está mais próximo de 100. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o GOV tem significância sobre a análise, sendo o modelo explica 66,68% da variância.

One-way ANOVA: GOV-n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  56636  8091  35,45  0,000
Error     124  28301   228
Total     131  84938

S = 15,11   R-Sq = 66,68%   R-Sq(adj) = 64,80%


                          Individual 95% CIs For Mean Based on
```

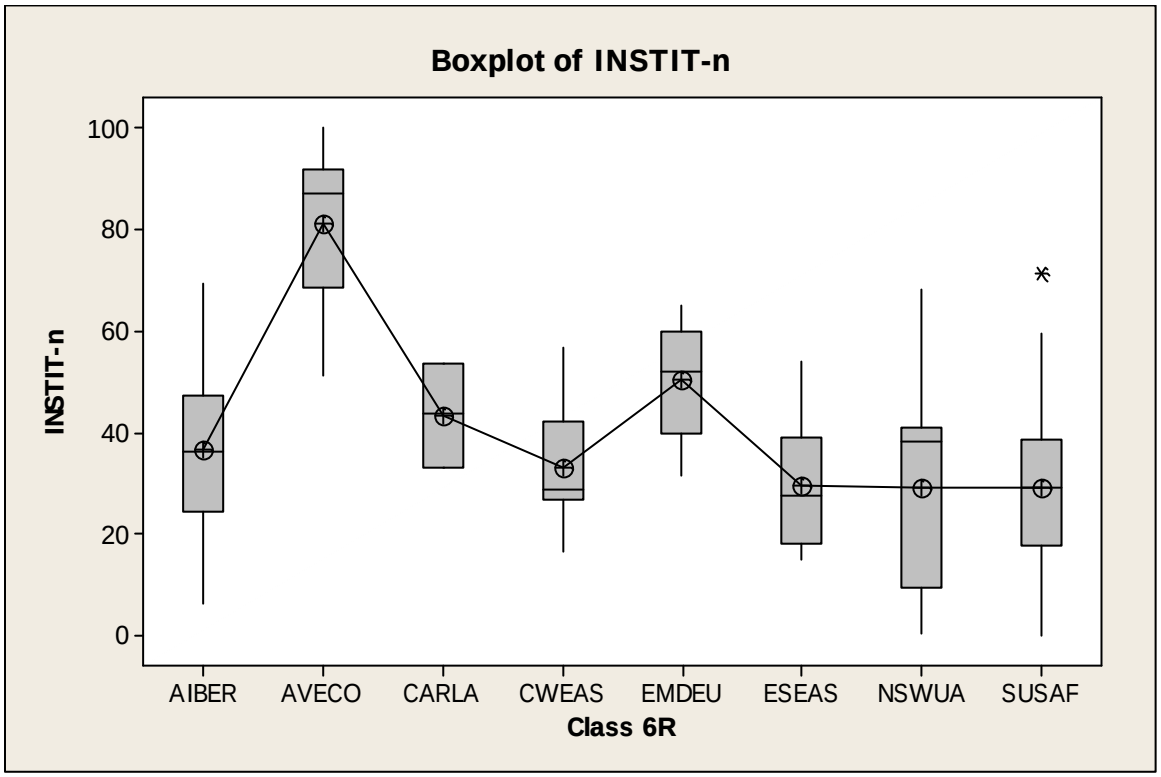
```
                          Pooled StDev
Level   N   Mean  StDev  -+---------+---------+---------+--------
AIBER  21  42,48  18,82           (--*---)
AVECO  27  83,25  13,38                                (--*--)
CARLA   3  43,25   7,61       (--------*-------)
CWEAS  11  27,38  12,38  (----*---)
EMDEU  12  51,93  11,21               (---*---)
ESEAS  12  33,61  10,47     (----*---)
NSWUA  14  28,46  16,21   (---*---)
SUSAF  32  29,58  16,95     (--*-)
                          -+---------+---------+---------+--------
                          20        40        60        80

Pooled StDev = 15,11
```



One-way ANOVA: GOV-n versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  51019  25510  97,02  0,000
Error       129  33919    263
Total       131  84938

S = 16,22   R-Sq = 60,07%   R-Sq(adj) = 59,45%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  42,48  18,82       (---*----)
AVECO   27  83,25  13,38                                  (----*---)
OTHERS  84  33,36  16,35   (-*--)
                           +---------+---------+---------+---------
                          30        45        60        75

Pooled StDev = 16,22
```

4.5 Análise INSTITUIÇÕES



A região AVECO é a melhor colocada, pois está mais próximo de 100. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**.

No caso das regiões ESEAS, NSWUA e SUSAF estão com a distribuições próximas e um há 1 outlier acima do 1º quartil da SUSAF que é a Mauritius.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o INSTIT tem significância sobre a análise, sendo o modelo explica 64,60% da variância.

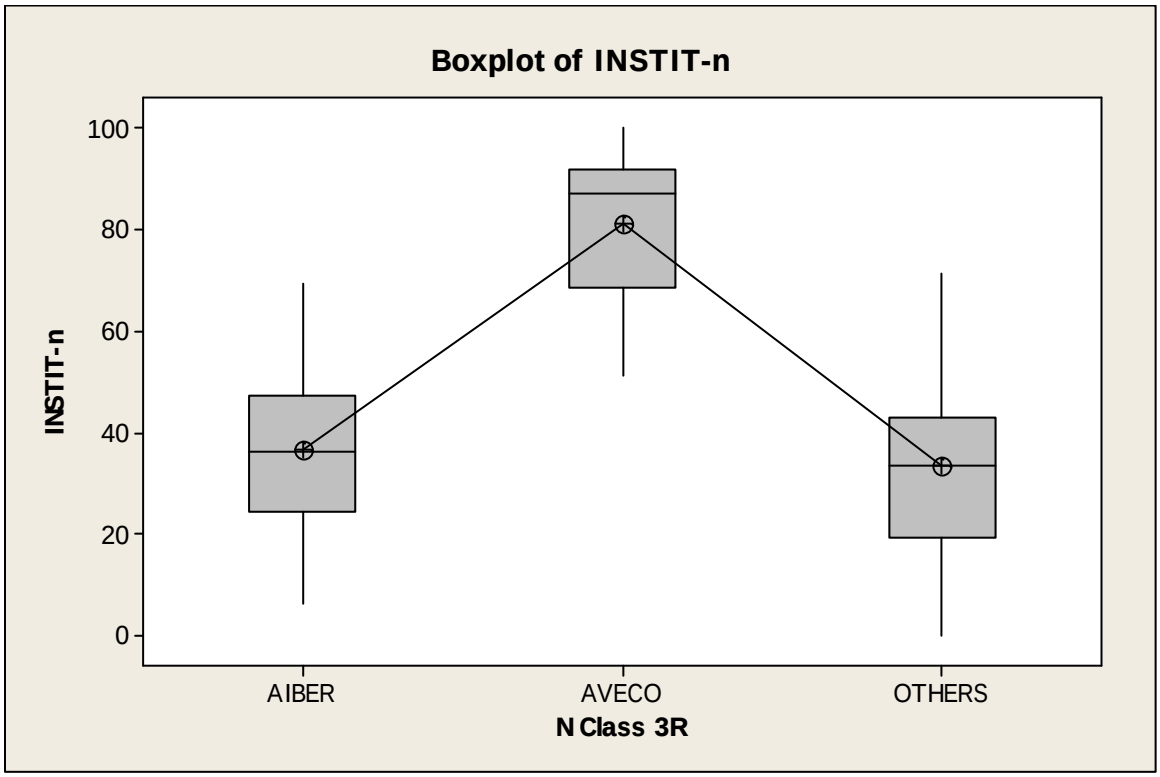
One-way ANOVA: INSTIT-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	52789	7541	32,32	0,000
Error	124	28928	233		
Total	131	81717			

S = 15,27 R-Sq = 64,60% R-Sq(adj) = 62,60%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
AIBER	21	36,61	17,02	(--*---)
AVECO	27	81,16	14,24	(--*-)
CARLA	3	43,41	10,19	(--------*-------)

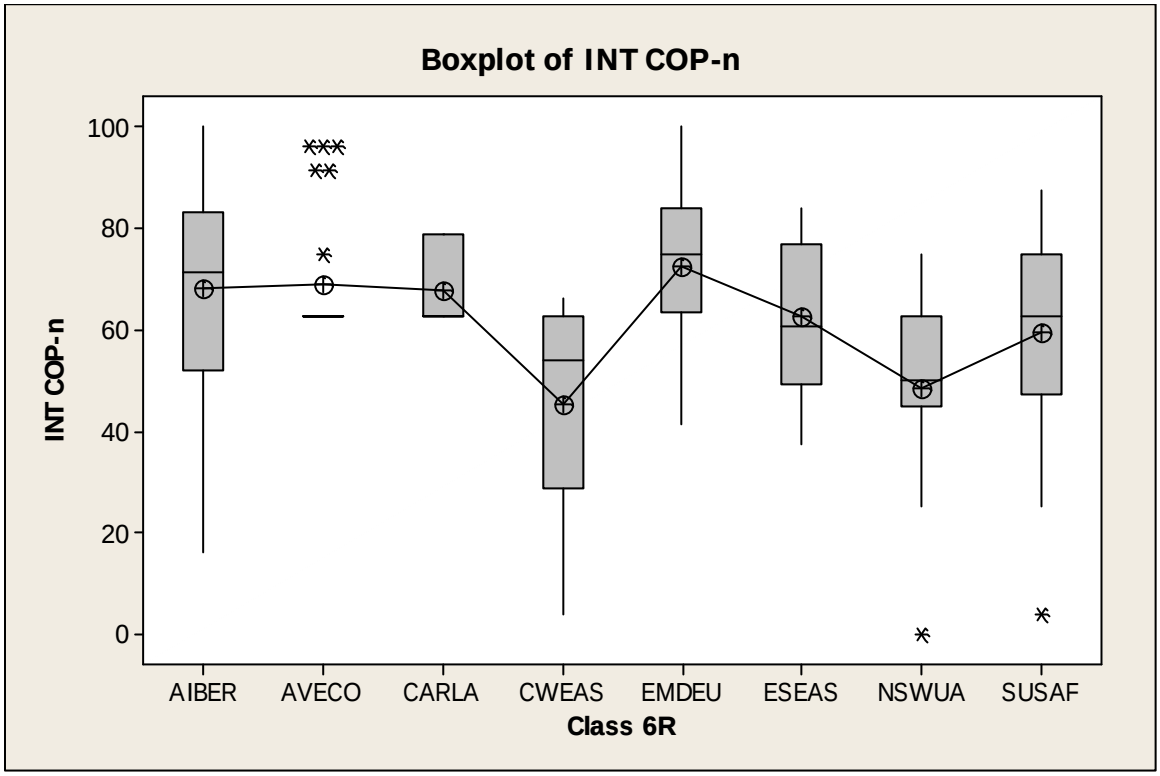
```
CWEAS   11   33,22   12,74      (----*---)
EMDEU   12   50,35   10,71                (---*----)
ESEAS   12   29,64   12,42    (----*---)
NSWUA   14   29,29   19,90     (---*---)
SUSAF   32   29,10   15,99      (--*-)
                             +---------+---------+---------+---------
                            20        40        60        80

Pooled StDev = 15,27
```



One-way ANOVA: INSTIT-n versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  48041  24020  92,01  0,000
Error       129  33676    261
Total       131  81717

S = 16,16   R-Sq = 58,79%   R-Sq(adj) = 58,15%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level    N   Mean  StDev   +---------+---------+---------+---------
AIBER   21  36,61  17,02   (---*----)
AVECO   27  81,16  14,24                                 (---*---)
OTHERS  84  33,29  16,51   (-*--)
                           +---------+---------+---------+---------
                           30        45        60        75

Pooled StDev = 16,16
```

4.6 Análise COOPERAÇÃO INTERNACIONAL



A região AVECO não apresenta o Boxplot, isto pode significar que o conjunto de observações encontram-se com resultados próximos. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**. E há 6 outliers acima do 3º quartil sendo respectivamente (cima para baixo): Estônia, República da Coreia, Slovakia, abaixo República Tcheca e Latvia e abaixo Slovenia.

Nesta análise as regiões de AIBER, AVECO e CARLA apresentam o conjunto de observações com a distribuição próxima.

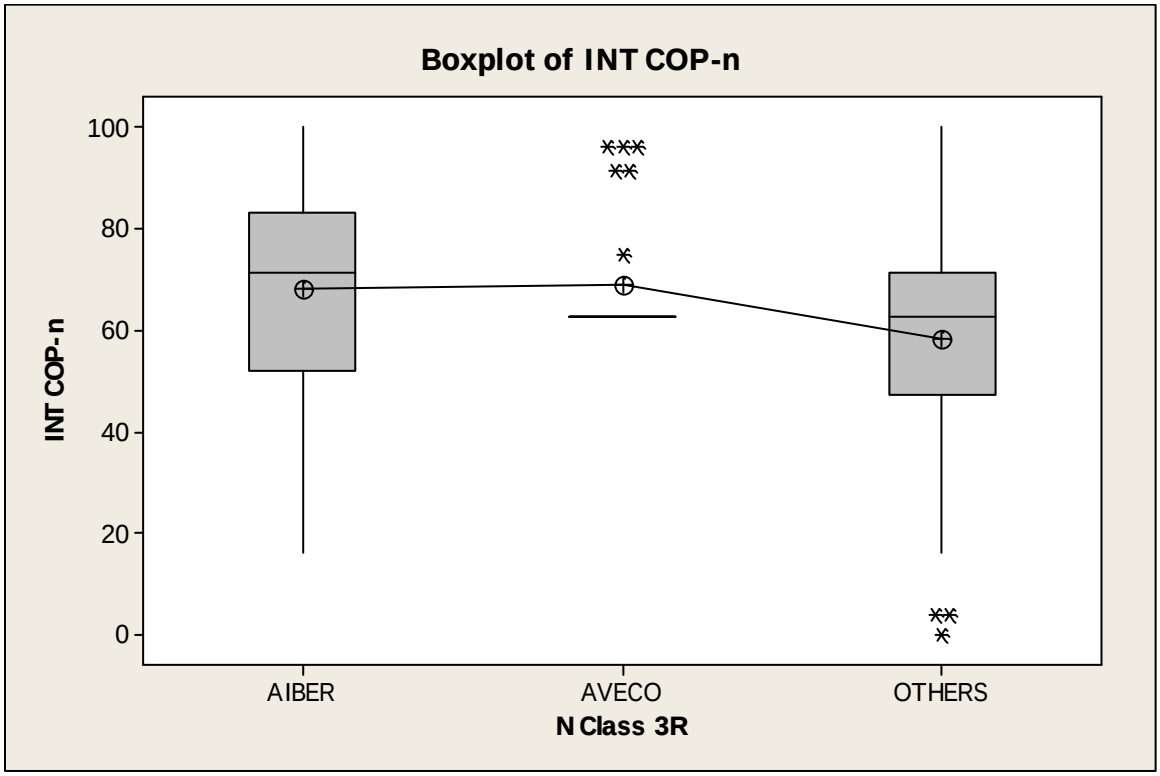
Na região NSWUA há um outlier abaixo do 1º quartil que é o Irã e no caso da região SUSAF há 1 outlier abaixo do 1º quartil a Sudão..

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o INT COP tem significância sobre a análise, apesar do modelo explicar apenas 19,04% da variância.

One-way ANOVA: INT COP-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	9329	1333	4,17	0,000
Error	124	39661	320		
Total	131	48990			

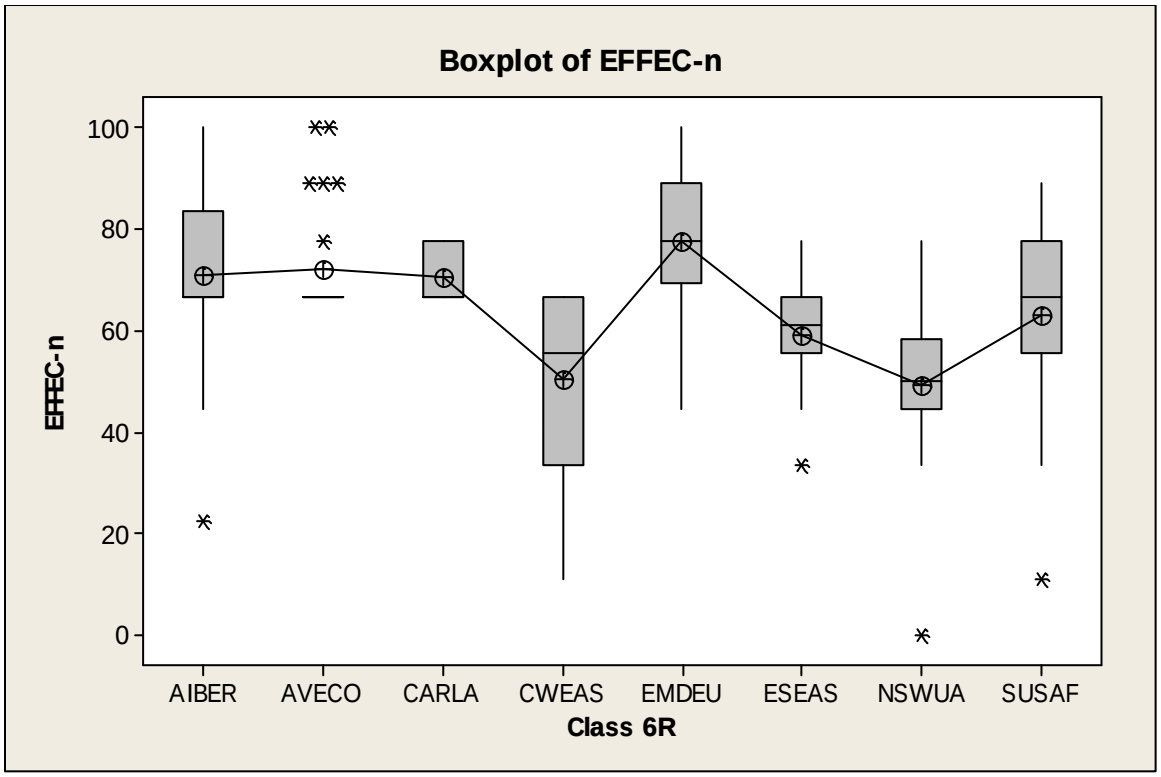
```
S = 17,88   R-Sq = 19,04%   R-Sq(adj) = 14,47%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   -------+---------+---------+---------+--
AIBER  21  68,27  21,82                  (-----*----)
AVECO  27  68,84  12,62                   (----*---)
CARLA   3  67,92   9,38            (------------*-------------)
CWEAS  11  45,34  20,95  (------*------)
EMDEU  12  72,50  16,45                     (-----*------)
ESEAS  12  62,71  14,55               (------*------)
NSWUA  14  48,48  18,42     (-----*------)
SUSAF  32  59,53  19,44               (---*---)
                          -------+---------+---------+---------+--
                                45        60        75        90

Pooled StDev = 17,88
```



One-way ANOVA: INT COP-n versus N Class 3R

```
Source       DF     SS    MS     F      P
N Class 3R    2   3155  1577
4,44  0,014
Error       129  45835   355
Total       131  48990

S = 18,85   R-Sq = 6,44%   R-Sq(adj) = 4,99%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level    N   Mean  StDev   ---------+---------+---------+---------+
AIBER   21  68,27  21,82            (-------------*------------)
AVECO   27  68,84  12,62               (-----------*-----------)
OTHERS  84  58,44  19,69  (-----*------)
                           ---------+---------+---------+---------+
                                 60,0      66,0      72,0      78,0

Pooled StDev = 18,85
```

4.7 Análise UTILIZAÇAO EFICAZ DO APOIO



A região AVECO não apresenta o Boxplot, isto pode significar que o conjunto de observações encontram-se com resultados próximos. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**. E há 6 outliers acima do 3º quartil sendo respectivamente (cima para baixo): Estônia, República da Coreia, abaixo República Tcheca, Latvia e Slovakia e abaixo Slovenia.

Nas outras regiões como AIBER, ESEAS, NSWUA e SUSAF apresentam os seguintes outliers respectivamente: Venezuela, Cambodia, Irã e Sudão. Todos abaixo do 1º quartil.

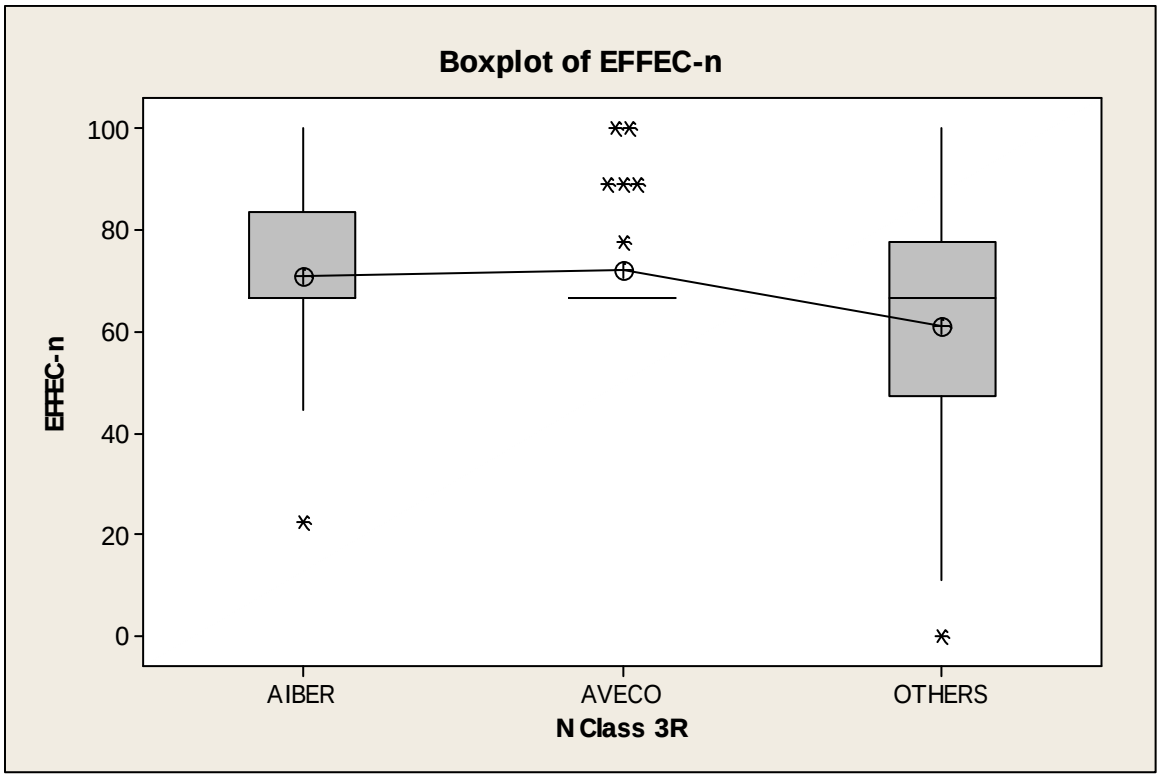
Nesta análise as regiões de AIBER, AVECO e CARLA apresentam o conjunto de observações com a distribuição próxima.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o EFFEC tem significância sobre a análise, apesar do modelo explicar apenas 23,78% da variância.

One-way ANOVA: EFFEC-n versus Class 6R

```
Source     DF     SS    MS     F      P
Class 6R    7  10404  1486  5,53  0,000
Error     124  33340   269
Total     131  43745

S = 16,40   R-Sq = 23,78%   R-Sq(adj) = 19,48%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
AIBER  21  70,90  19,40                (---*----)
```

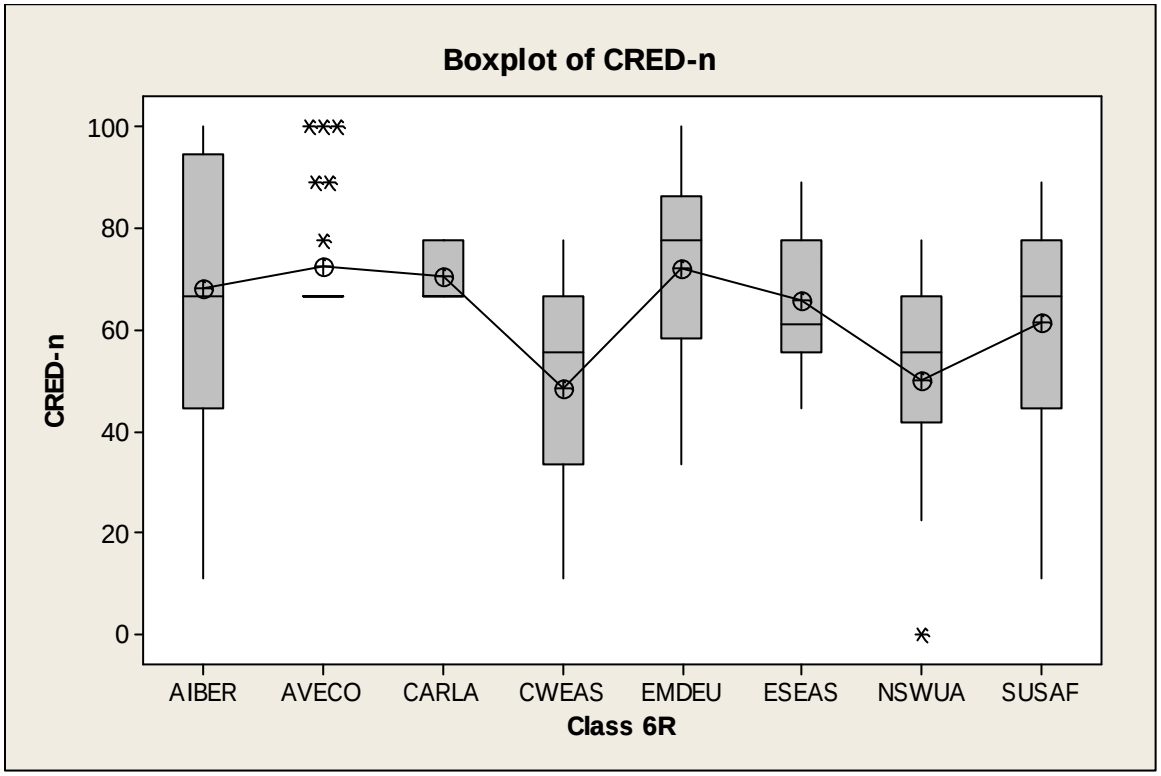
```
AVECO   27  72,02  10,84                    (---*---)
CARLA    3  70,37   6,42          (------------*-----------)
CWEAS   11  50,51  18,16  (------*-----)
EMDEU   12  77,78  15,71                      (-----*-----)
ESEAS   12  59,26  11,92         (------*-----)
NSWUA   14  49,21  18,34  (-----*-----)
SUSAF   32  63,19  18,60             (---*---)
                          ---+---------+---------+---------+------
                             45        60        75        90

Pooled StDev = 16,40
```



One-way ANOVA: EFFEC-n versus N Class 3R

```
Source       DF     SS    MS     F      P
N Class 3R    2   3414  1707  5,46  0,005
Error       129  40331   313
Total       131  43745

S = 17,68   R-Sq = 7,80%   R-Sq(adj) = 6,37%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -----+---------+---------+---------+----
AIBER   21  70,90  19,40           (------------*------------)
AVECO   27  72,02  10,84              (----------*----------)
OTHERS  84  60,98  18,93  (------*-----)
                          -----+---------+---------+---------+----
                            60,0      66,0      72,0      78,0

Pooled StDev = 17,68
```

4.8 Análise CREDIBILIDADE



A região AIBER possui a distribuição de seus dados bem dispersas, isto pode representar que para esta variável não há muita semelhança entre os países desta região.

A região AVECO não apresenta o Boxplot, isto pode significar que o conjunto de observações encontram-se com resultados próximos. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**. E há 6 outliers acima do 3º quartil sendo respectivamente (cima para baixo): Estônia, República da Coreia e Slovakia abaixo República Tcheca, Latvia e bem próximo da mediana Slovenia.

Na região NSWUA apresenta o Irã como outlier, bem abaixo do primeiro quartil, quase em zero.

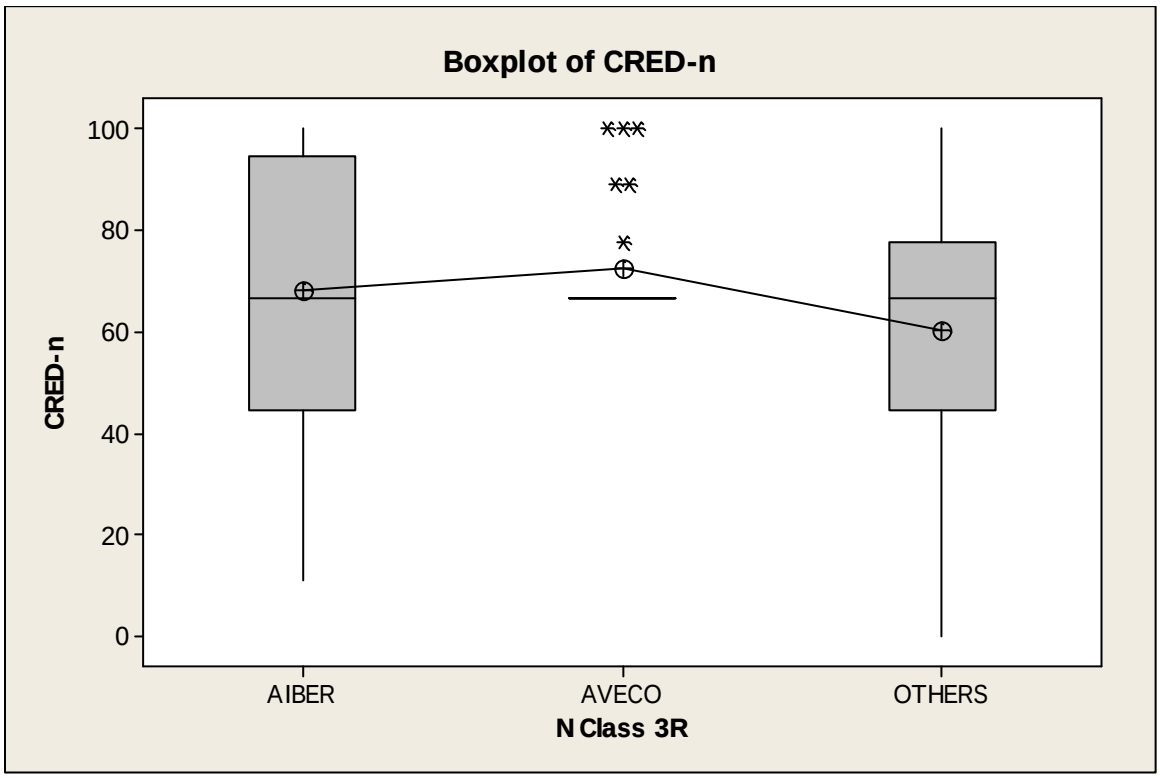
Abaixo no GLM (Modelo Linear Geral) o ***p-value*** nãoé zero, portanto há pouca evidencia que o CRED tem significância sobre a análise, apesar do modelo explicar apenas 15,55% da variância.

One-way ANOVA: CRED-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	8865	1266	3,26	0,003
Error	124	48137	388		
Total	131	57002			

S = 19,70 R-Sq = 15,55% R-Sq(adj) = 10,79%

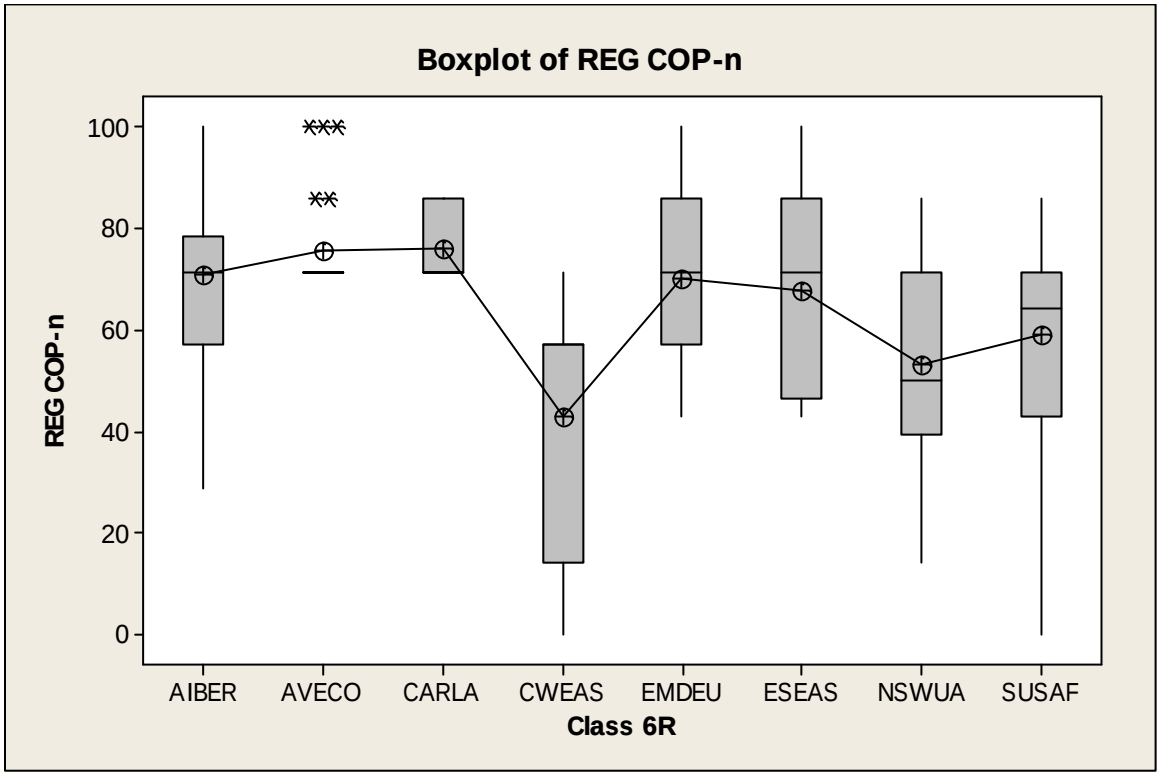
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ------+---------+---------+---------+---
AIBER  21  68,25  25,89                 (-----*----)
AVECO  27  72,43  11,68                    (----*----)
CARLA   3  70,37   6,42          (--------------*--------------)
CWEAS  11  48,48  22,37  (-------*-------)
EMDEU  12  72,22  20,38                  (------*-------)
ESEAS  12  65,74  16,04              (-------*------)
NSWUA  14  50,00  20,79    (------*------)
SUSAF  32  61,46  20,54              (----*----)
                          ------+---------+---------+---------+---
                               45        60        75        90

Pooled StDev = 19,70
```



One-way ANOVA: CRED-n versus N Class 3R

```
Source       DF     SS    MS     F      P
N Class 3R    2   3437  1718  4,14  0,018
Error       129  53566   415
Total       131  57002

S = 20,38   R-Sq = 6,03%   R-Sq(adj) = 4,57%


                           Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev     +---------+---------+---------+---------
AIBER   21  68,25  25,89          (-----------*-----------)
AVECO   27  72,43  11,68                (-----------*-----------)
OTHERS  84  60,32  21,00     (-----*-----)
                             +---------+---------+---------+---------
                            56,0      63,0      70,0      77,0

Pooled StDev = 20,38
```

4.9 Análise COOPERAÇÃO REGIONAL



A região AVECO não apresenta o Boxplot, isto pode significar que o conjunto de observações encontram-se com resultados próximos. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. Sendo que a sua mediana está mais próxima do 1º quartil, neste caso é **positivamente assimétrico**. E há 5 outliers acima do 3º quartil sendo respectivamente (cima para baixo): República Tcheca, Latvia e Slovakia, abaixo Estonia e República da Coreia.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o REG COP tem significância sobre a análise, apesar do modelo explicar apenas 22,53% da variância.

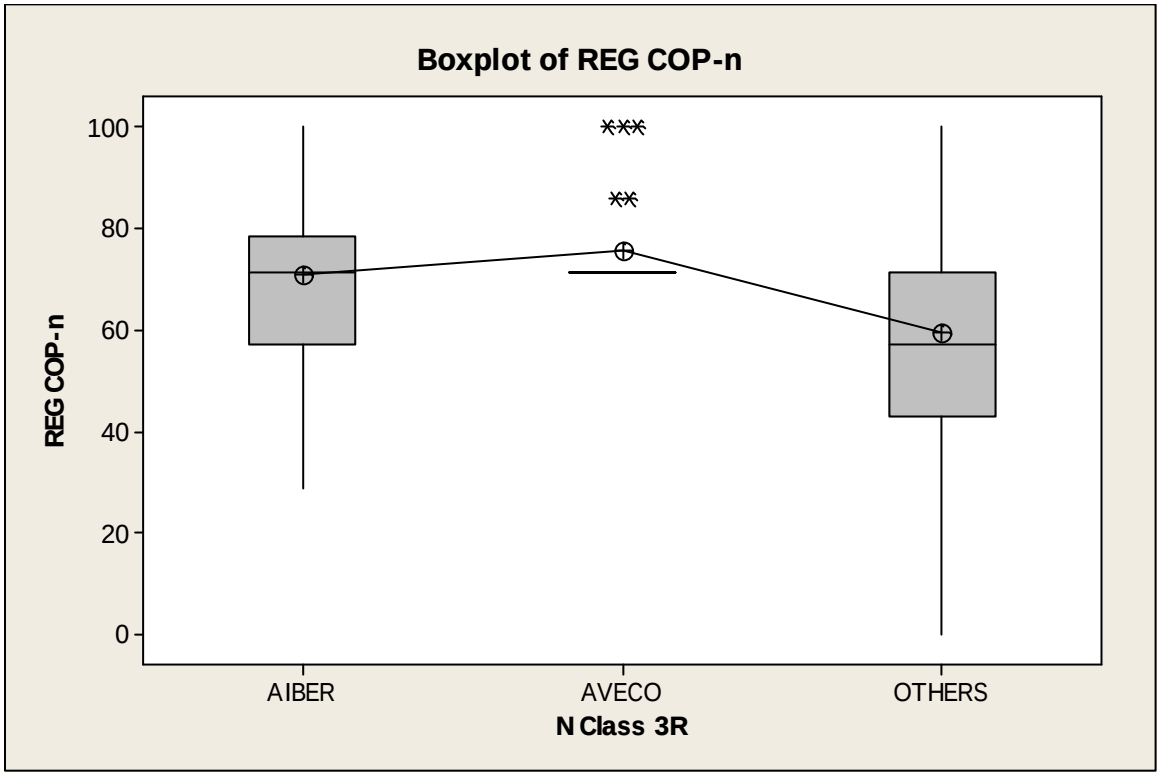
One-way ANOVA: REG COP-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	13101	1872	5,15	0,000
Error	124	45056	363		
Total	131	58157			

S = 19,06 R-Sq = 22,53% R-Sq(adj) = 18,15%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev

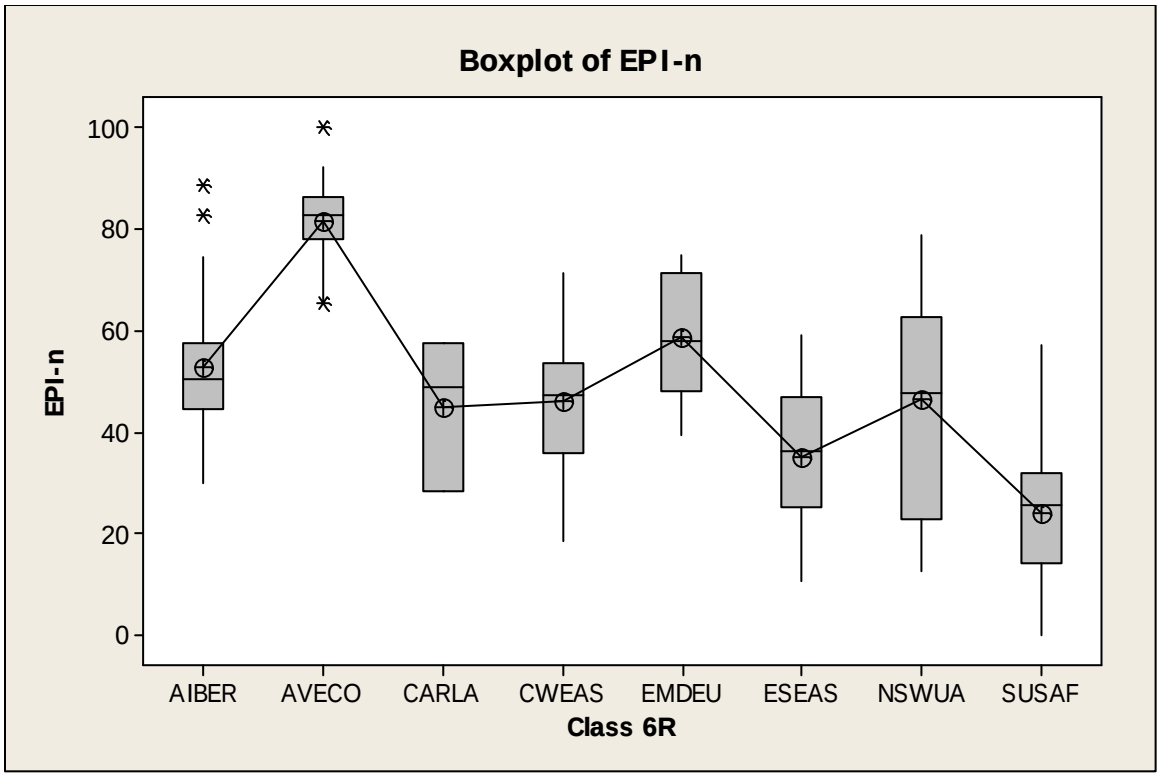
```
AIBER   21  70,75  19,42                  (---*---)
AVECO   27  75,66   9,55                     (---*--)
CARLA    3  76,19   8,25               (----------*----------)
CWEAS   11  42,86  24,74  (----*-----)
EMDEU   12  70,24  18,73                  (----*-----)
ESEAS   12  67,86  19,38                (-----*----)
NSWUA   14  53,06  22,01       (-----*----)
SUSAF   32  58,93  21,70             (--*---)
                           ----+---------+---------+---------+-----
                              40        60        80       100

Pooled StDev = 19,06
```



One-way ANOVA: REG COP-n versus N Class 3R

```
Source       DF     SS    MS     F      P
N Class 3R    2   6408  3204  7,99  0,001
Error       129  51749   401
Total       131  58157

S = 20,03   R-Sq = 11,02%   R-Sq(adj) = 9,64%


                          Individual 95% CIs For Mean Based on Pooled StDev
Level    N   Mean  StDev    -+---------+---------+---------+--------
AIBER   21  70,75  19,42             (---------*----------)
AVECO   27  75,66   9,55                    (---------*--------)
OTHERS  84  59,35  22,45    (----*-----)
                            -+---------+---------+---------+--------
                           56,0      64,0      72,0      80,0

Pooled StDev = 20,03
```

4.10 Análise EPI



A região AIBER possui um outlier bem acima do 3º quartil sendo (de cima para baixo) a Espanha e abaixo Portugal, estão bem acima da caixa e longe da mediana.

A região AVECO está destacada das demais e sua caixa pequena, ou seja, sua distribuição está concentrada. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc. E há 2 outliers bem acima do 3º quartil sendo a Suiça e o outro bem abaixo do 1º quartil a República da Coreia.

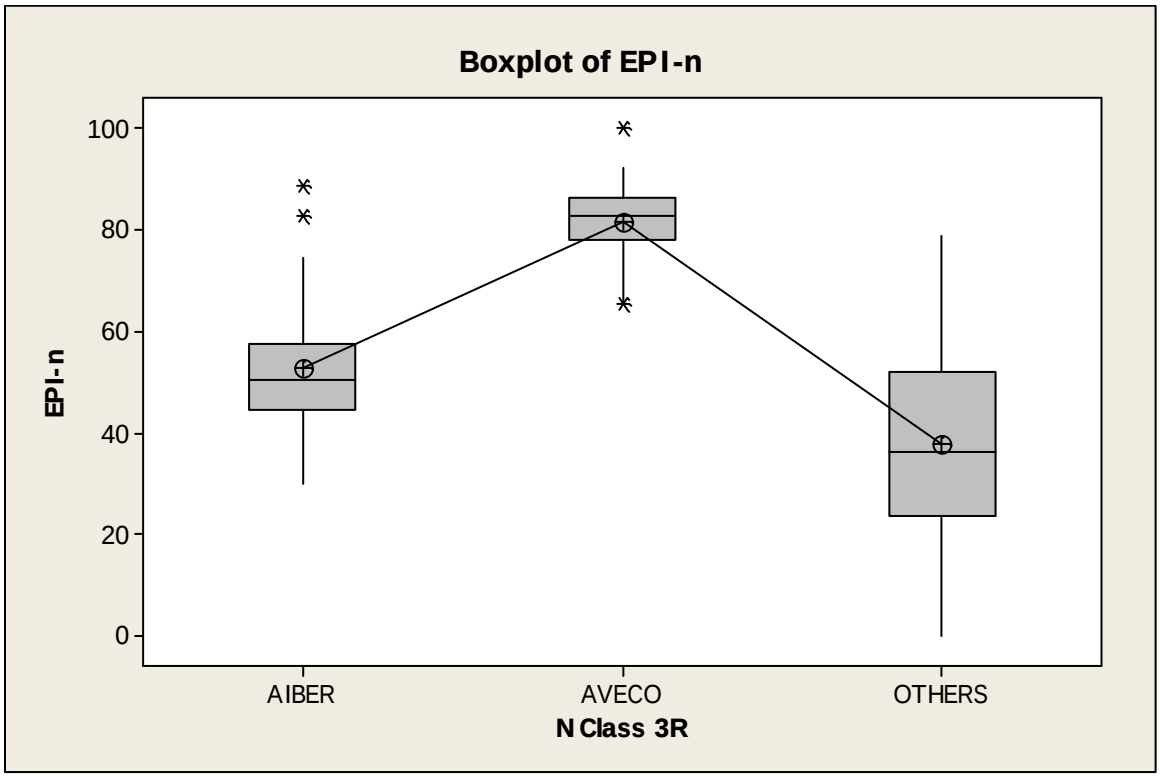
A região SUSAF está mais abaixo de todas, ou seja, com o EPI bem diferenciado das demais regiões.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o EPI tem significância sobre a análise, seu F tem 40,56 e seu modelo explica 69,60% da variância.

One-way ANOVA: EPI-n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  52100  7443  40,56  0,000
Error     124  22755   184
Total     131  74855

S = 13,55   R-Sq = 69,60%   R-Sq(adj) = 67,89%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  21  52,70  14,31               (--*--)
AVECO  27  81,39   8,14                                  (--*-)
CARLA   3  44,93  14,98        (------*-------)
```

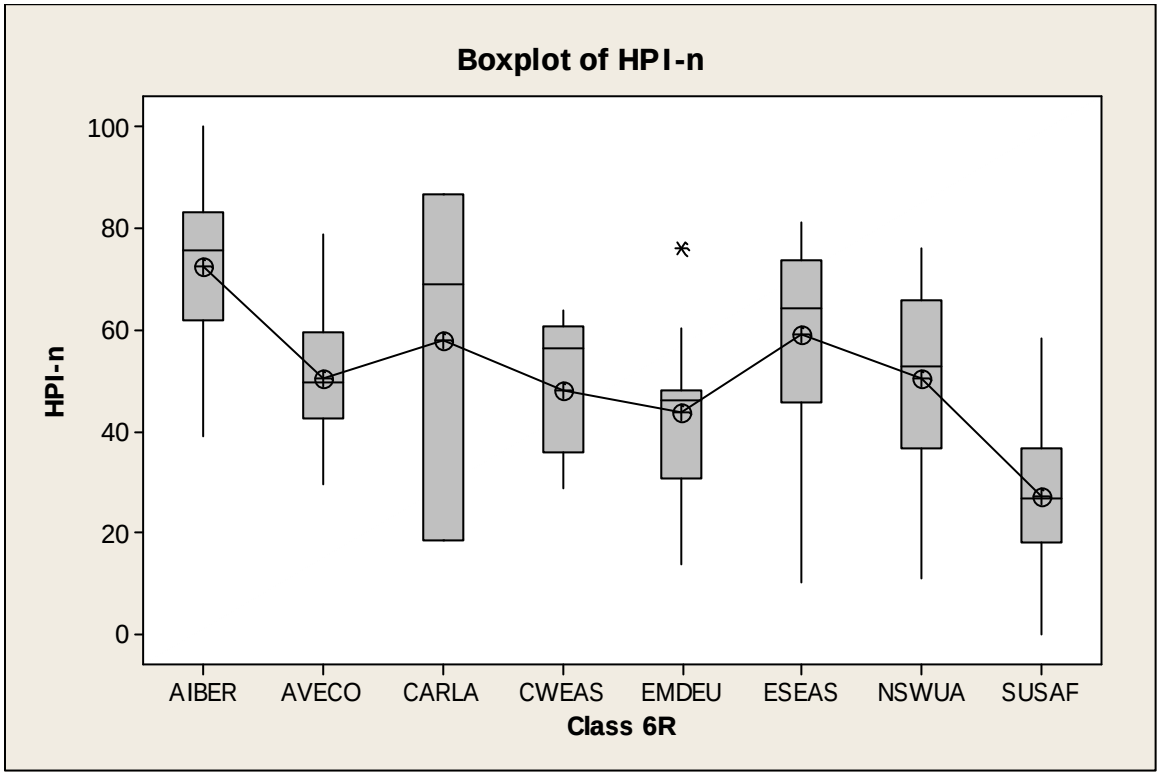
```
CWEAS   11   46,14   14,39            (---*---)
EMDEU   12   58,63   11,91                 (---*---)
ESEAS   12   34,97   13,95       (--*---)
NSWUA   14   46,41   20,67             (--*---)
SUSAF   32   24,14   12,90   (-*-)
                             +---------+---------+---------+---------
                            20        40        60        80

Pooled StDev = 13,55
```



One-way ANOVA: EPI-n versus N Class 3R

```
Source       DF     SS     MS      F      P
N Class 3R    2  38872  19436  69,68  0,000
Error       129  35982    279
Total       131  74855

S = 16,70   R-Sq = 51,93%   R-Sq(adj) = 51,18%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level    N   Mean  StDev  -------+---------+---------+---------+--
AIBER   21  52,70  14,31         (----*----)
AVECO   27  81,39   8,14                              (---*----)
OTHERS  84  37,95  19,06  (-*--)
                           -------+---------+---------+---------+--
                                 45        60        75        90

Pooled StDev = 16,70
```

4.11 Análise HPI



A região AIBER possui uma melhor colocação que as demais, apesar da região CARLA ter um 3º quartil próximo, mas seus dados estão bem distribuídos. A região é representada por apenas 3 países sendo eles: Guinea (África), Jamaica e Trinidad e Tobago.

A região SUSAF está mais abaixo de todas, ou seja, com o HPI bem diferenciado das demais regiões.

A região de EMDEU possui sua mediana muito próxima do 3º quartil. E ainda acima do 3º quartil um outlier sendo Malawi. Esta região é representada pelos países como Lituânia, Macedônia, Montenegro, Polônia, Servia e etc.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o HPI tem significância sobre a análise, apesar de seu modelo explicar apenas 46,79%% da variância.

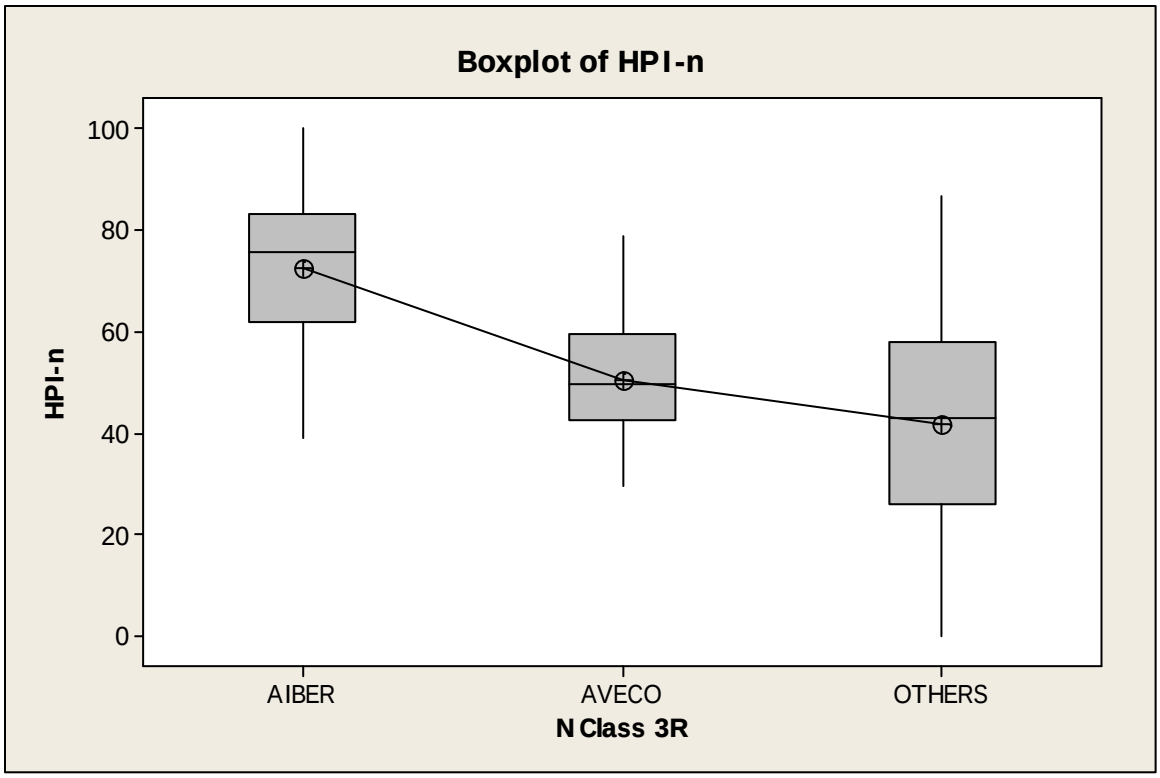
One-way ANOVA: HPI-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	28494	4071	15,58	0,000
Error	124	32401	261		
Total	131	60894			

S = 16,16 R-Sq = 46,79% R-Sq(adj) = 43,79%

Individual 95% CIs For Mean Based on

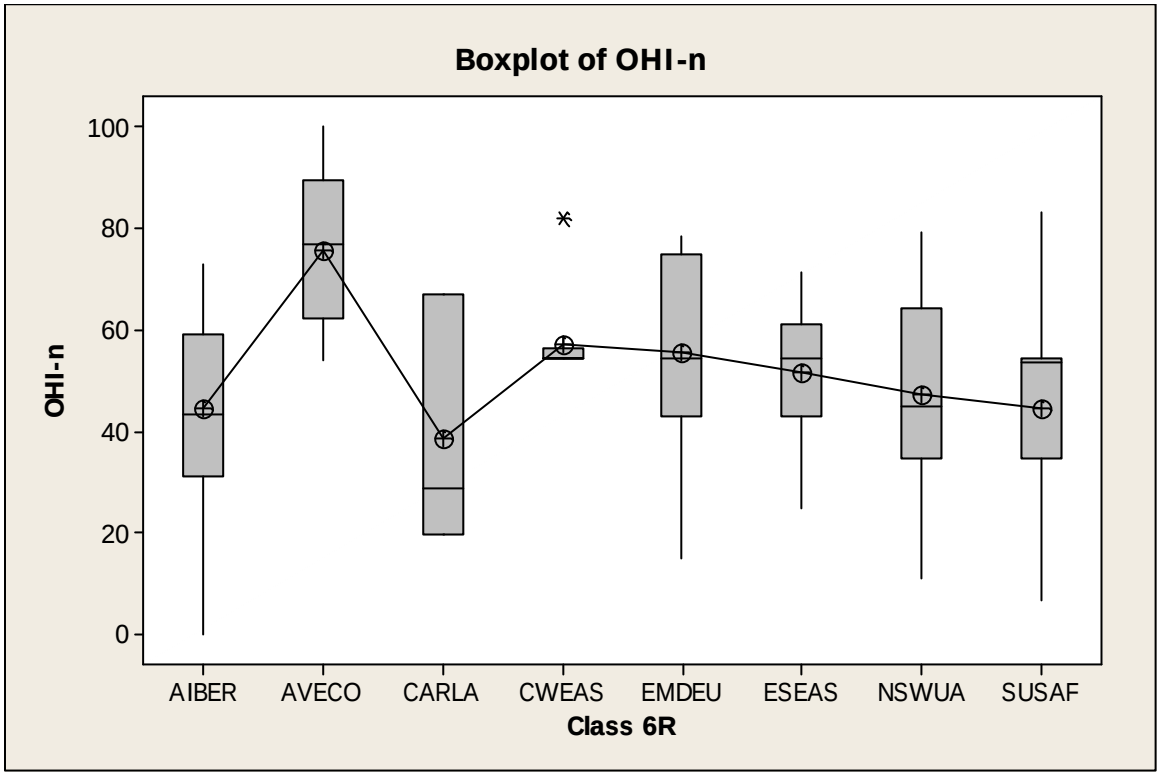
```
                          Pooled StDev
Level   N   Mean  StDev  ------+---------+---------+---------+---
AIBER  21  72,36  16,25                              (---*----)
AVECO  27  50,58  12,84                   (---*---)
CARLA   3  58,07  35,38              (------------*-----------)
CWEAS  11  48,13  13,54              (-----*------)
EMDEU  12  43,52  16,01           (-----*-----)
ESEAS  12  59,06  20,02                     (-----*------)
NSWUA  14  50,22  19,91                (----*-----)
SUSAF  32  27,23  14,04  (---*---)
                          ------+---------+---------+---------+---
                               30        45        60        75

Pooled StDev = 16,16
```



One-way ANOVA: HPI-n versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  15866  7933  22,73  0,000
Error       129  45029   349
Total       131  60894

S = 18,68   R-Sq = 26,05%   R-Sq(adj) = 24,91%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  72,36  16,25                         (-----*------)
AVECO   27  50,58  12,84       (-----*-----)
OTHERS  84  41,78  20,67  (---*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 18,68
```

4.12 Análise OHI



A região AVECO está destacada das demais e sua caixa pequena, ou seja, sua distribuição está concentrada. Esta região representada por países como: Bélgica, Finlândia, Canadá, França, Alemanha, Grécia, etc.

A região CARLA tem uma caixa maior, ou seja, seus dados estão mais distribuídos. A região é representada por apenas 3 países sendo eles: Guinea (África), Jamaica e Trinidad e Tobago.

A distribuição dos dados da região CWEAS está concentrando, bem próximo da média de todas as regiões, há um outlier que é a França. Está região é composta por 11 países sendo Azerbaijão, Balarus, Georgia, Moldova e etc.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o OHI tem significância sobre a análise, seu F tem 9,81 apesar do seu modelo explicar apenas 35,63% da variância.

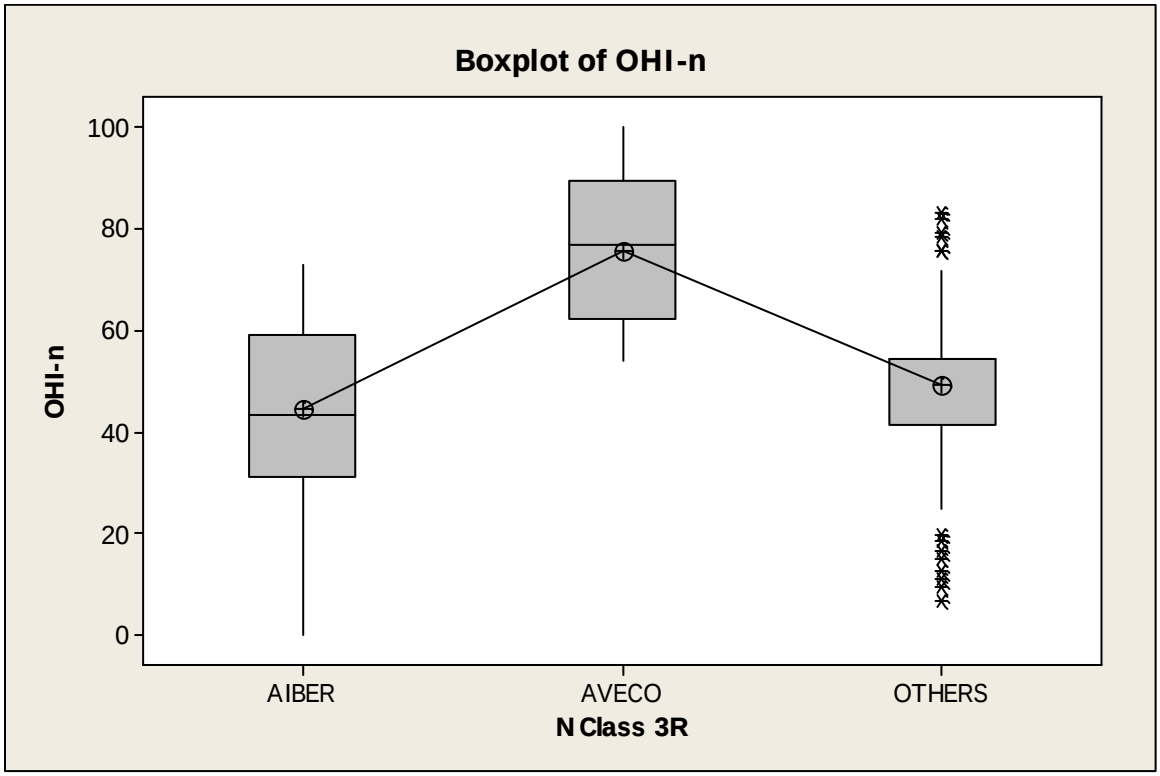
One-way ANOVA: OHI-n versus Class 6R

Source	DF	SS	MS	F	P
Class 6R	7	19115	2731	9,81	0,000
Error	124	34529	278		
Total	131	53644			

S = 16,69 R-Sq = 35,63% R-Sq(adj) = 32,00%

Individual 95% CIs For Mean Based on

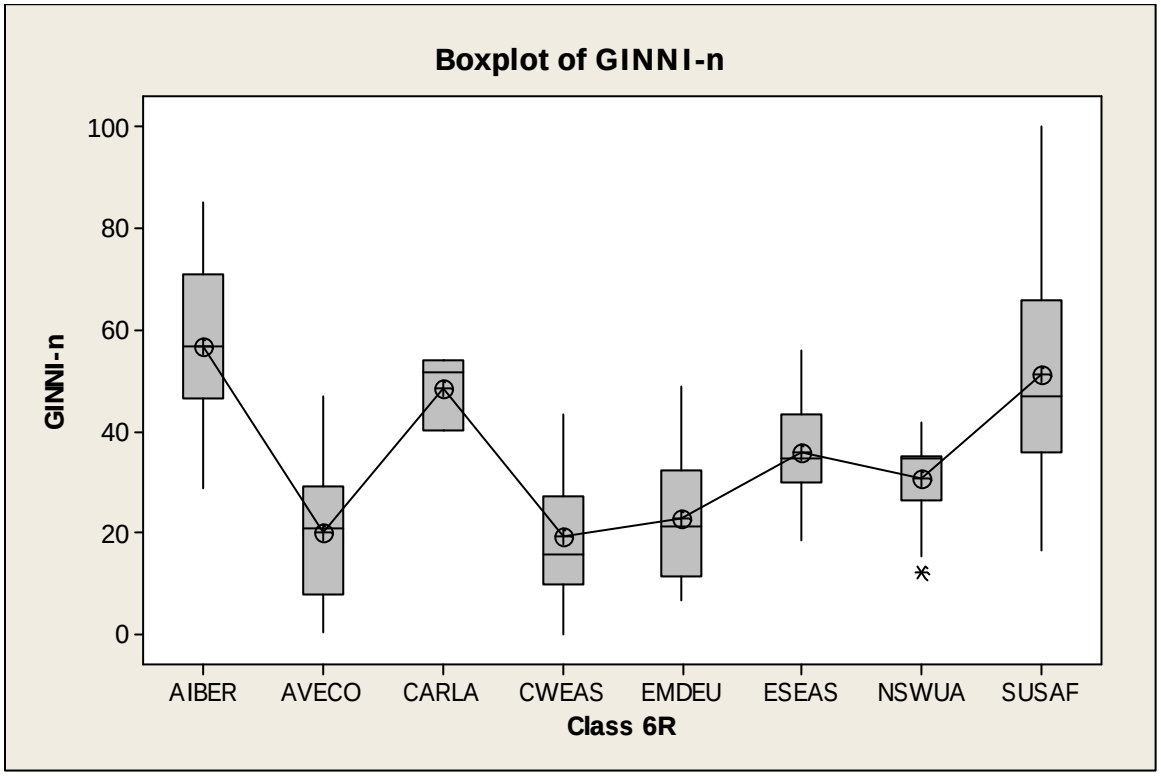
```
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
AIBER  21  44,59  19,05            (----*---)
AVECO  27  75,83  15,38                                 (---*---)
CARLA   3  38,47  25,19  (-----------*-----------)
CWEAS  11  57,29   8,30                   (-----*-----)
EMDEU  12  55,61  18,57                  (-----*-----)
ESEAS  12  51,76  13,06                (-----*-----)
NSWUA  14  47,10  18,68               (----*-----)
SUSAF  32  44,66  17,00               (---*---)
                          --------+---------+---------+---------+-
                                 32        48        64        80

Pooled StDev = 16,69
```



One-way ANOVA: OHI-n versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  16759  8379  29,31  0,000
Error       129  36885   286
Total       131  53644

S = 16,91   R-Sq = 31,24%   R-Sq(adj) = 30,17%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  ---------+---------+---------+---------+
AIBER   21  44,59  19,05  (-----*-----)
AVECO   27  75,83  15,38                           (----*-----)
OTHERS  84  49,08  16,82        (--*--)
                          ---------+---------+---------+---------+
                                  48        60        72        84

Pooled StDev = 16,91
```

4.13 Análise GINNI



A região AIBER está em destaque sobre as demais, sendo seguida pela região SUSAF, as demais estão bem abaixo da média das distribuição dos dados.

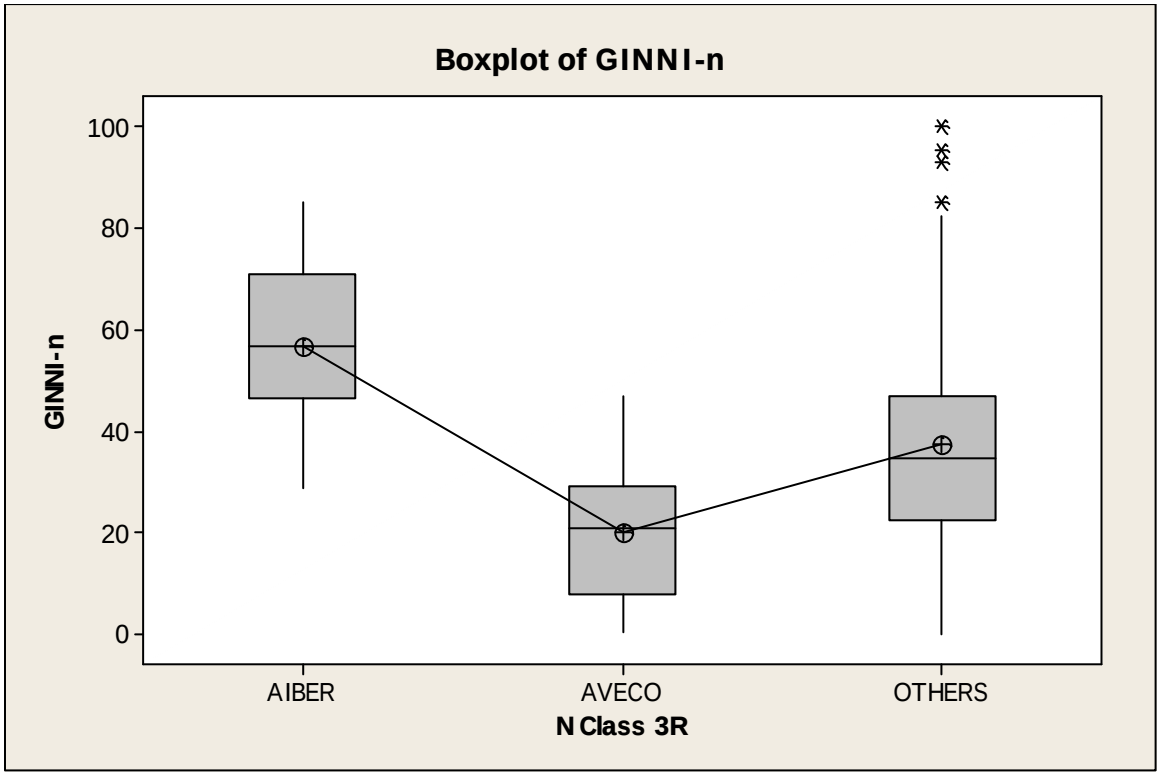
A região NSWUA tem um outlier abaixo do 1º quartil sendo o país Iraque. Esta região é composta por 14 países sendo: Argélia, Egito, Irã, Iraque, Jordânia, Kwuait, Líbano e etc.

Abaixo no GLM (Modelo Linear Geral) o ***p-value*** é zero, portanto há evidencia que o GINNI tem significância sobre a análise, seu F tem 17,97 apesar do seu modelo explicar 50,36% da variância.

One-way ANOVA: GINNI-n versus Class 6R

```
Source     DF     SS    MS      F      P
Class 6R    7  29485  4212  17,97  0,000
Error     124  29067   234
Total     131  58552

S = 15,31   R-Sq = 50,36%   R-Sq(adj) = 47,55%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---+---------+---------+---------+------
AIBER  21  56,81  14,56                            (----*---)
AVECO  27  20,00  12,31     (---*---)
CARLA   3  48,59   7,28                 (----------*-----------)
CWEAS  11  19,26  13,30   (-----*-----)
EMDEU  12  22,86  12,74     (-----*-----)
ESEAS  12  35,65  10,22               (-----*-----)
NSWUA  14  30,56   8,67           (----*-----)
```

```
SUSAF   32   51,34   22,07                          (--*---)
                          ---+---------+---------+---------+------
                            15        30        45        60

Pooled StDev = 15,31
```



One-way ANOVA: GINNI-n versus N Class 3R

```
Source       DF     SS    MS      F      P
N Class 3R    2  16042  8021  24,34  0,000
Error       129  42510   330
Total       131  58552

S = 18,15   R-Sq = 27,40%   R-Sq(adj) = 26,27%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level    N   Mean  StDev  -+---------+---------+---------+--------
AIBER   21  56,81  14,56                          (----*----)
AVECO   27  20,00  12,31  (---*----)
OTHERS  84  37,27  20,34              (--*-)
                          -+---------+---------+---------+--------
                          15        30        45        60

Pooled StDev = 18,15
```

4.14 Análise – FATOR Classificação 6R e 3R

Analisando os resultados da classificação 6R e 3R temos o quadro abaixo em ordem crescente considerando o “F”.

Observa-se que os 5 primeiros colocados além um alto F tem um grau de explicação da R-sqd acima de 60%.

		6R	3R
R	**Variável**	**F**	**F**
1	**SPI**	**58,02**	
2	**EPI**	**40,56**	
3	**GOVERNANÇA**	**35,45**	
4	**IDH**	**35,25**	
5	**INSTITUIÇÕES**	**32,32**	
6	GINNI	17,97	
7	HPI	15,58	
8	OHI	9,81	
9	UTILIZAÇÃO EFICAZ DO APOIO	5,53	
10	COOP REGIONAL	5,15	
11	COOP INTERN	4,17	
12	CREDIBILIDADE	3,26	

5. PESQUISA POR AMOSTRAGEM

A pesquisa por amostragem foi feita em três amostras, uma de 25 indivíduos, ou países, a outra composta por 50 indivíduos e por fim, com 100 países. Para cada amostra foram efetuadas análises exploratória de dados, amostragens e bloxpots.

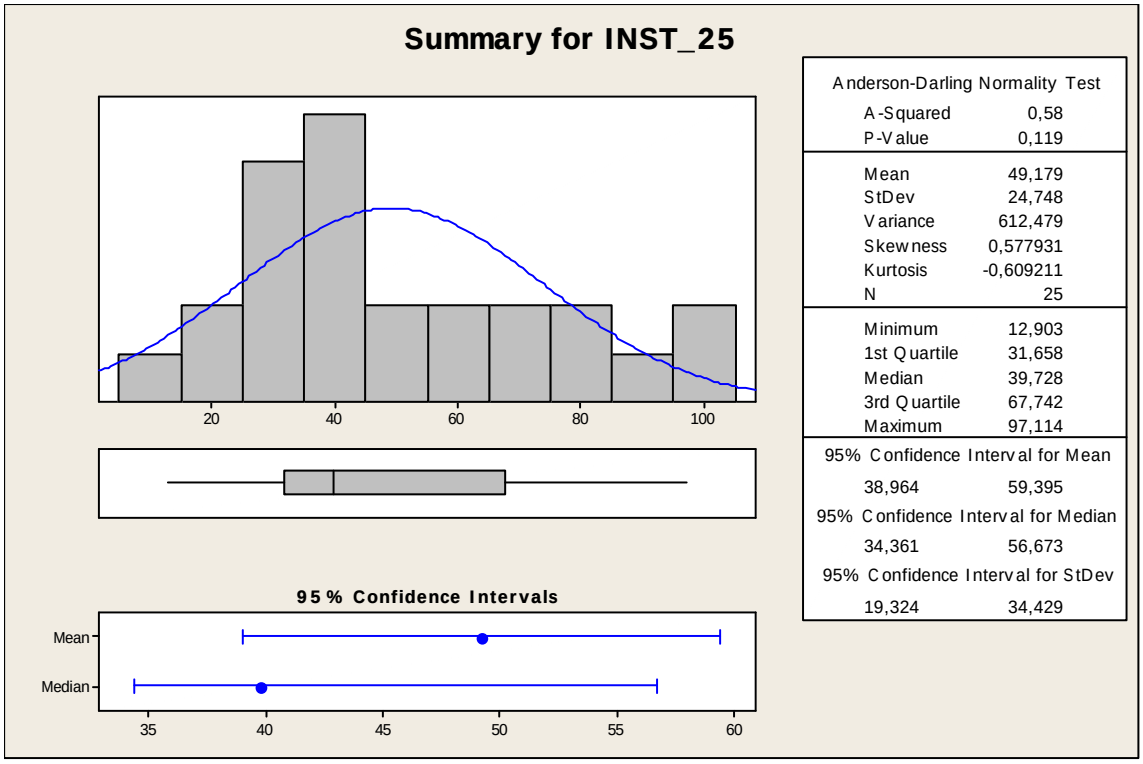
5.1 Análise Exploratória para Amostra de 25 Indivíduos

A amostra de 25, 50 e 100 indivíduos foram feitas de seguinte maneira:

- **Amostra Aleatória**: foi feita através a funcionalidade "Random Data" do MINITAB.

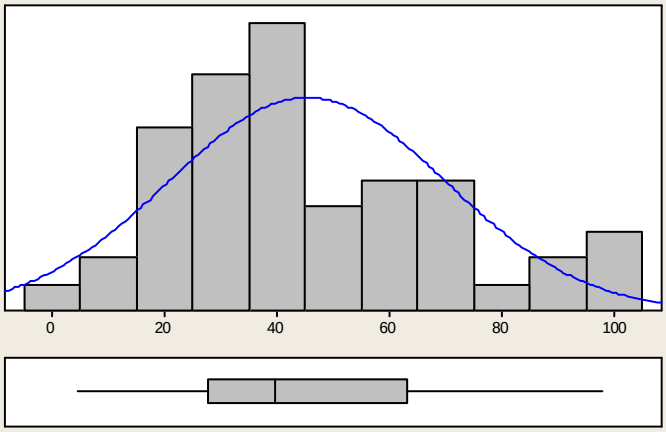
5.2 INSTITUTIONS

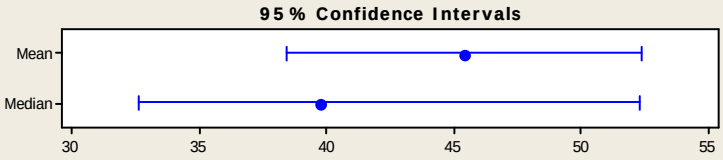
Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países. Em seguida o Bloxpots que faz o comparativo das regiões.



Anderson-Darling Normality Test	
A-Squared	0,58
P-Value	0,119
Mean	49,179
StDev	24,748
Variance	612,479
Skewness	0,577931
Kurtosis	-0,609211
N	25
Minimum	12,903
1st Quartile	31,658
Median	39,728
3rd Quartile	67,742
Maximum	97,114
95% Confidence Interval for Mean	
38,964	59,395
95% Confidence Interval for Median	
34,361	56,673
95% Confidence Interval for StDev	
19,324	34,429

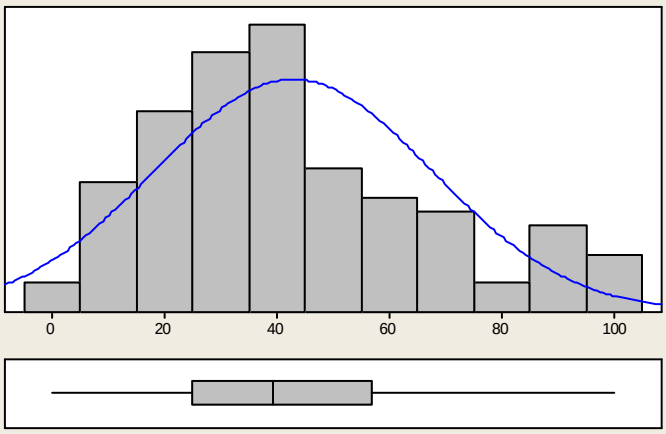
Summary for INST_50

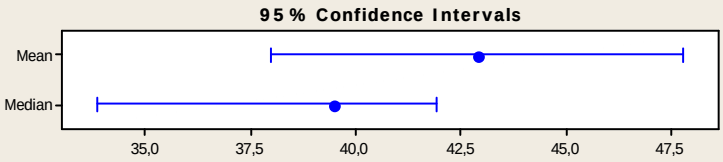




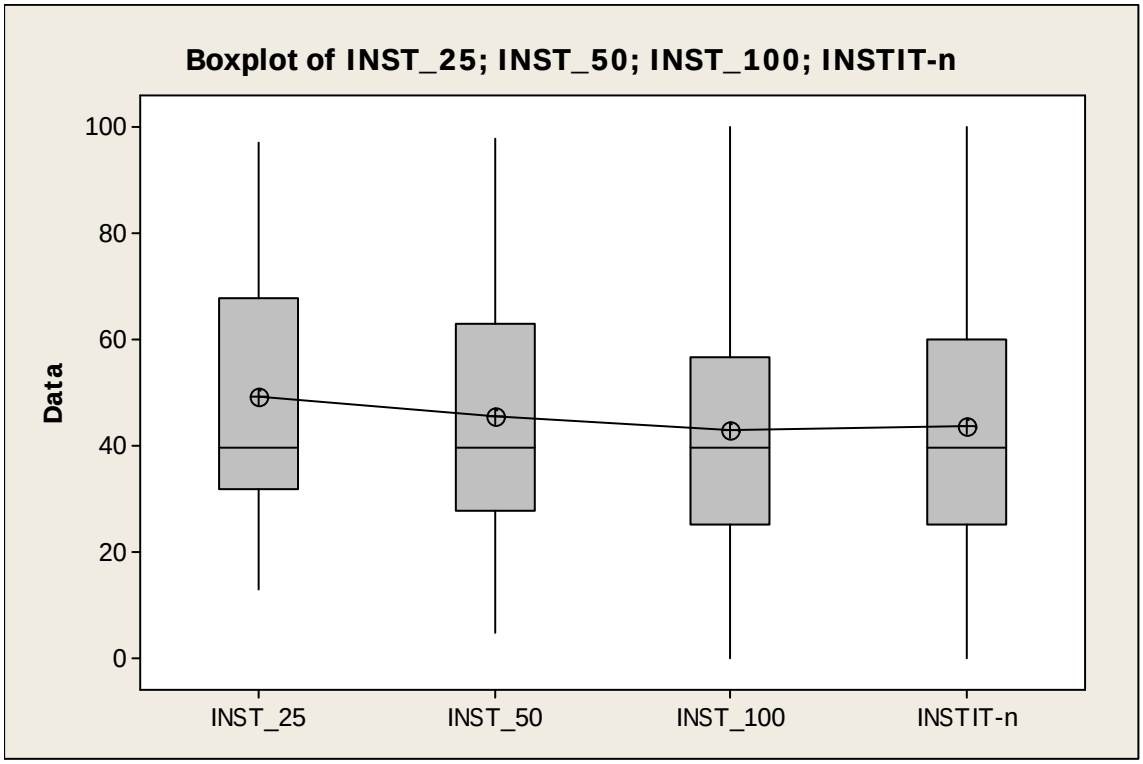
Anderson-Darling Normality Test	
A-Squared	0,76
P-Value	0,045
Mean	45,369
StDev	24,510
Variance	600,761
Skewness	0,544969
Kurtosis	-0,471918
N	50
Minimum	4,584
1st Quartile	27,880
Median	39,728
3rd Quartile	63,031
Maximum	97,963
95% Confidence Interval for Mean	
38,403	52,335
95% Confidence Interval for Median	
32,661	52,279
95% Confidence Interval for StDev	
20,474	30,543

Summary for INST_100





Anderson-Darling Normality Test	
A-Squared	1,45
P-Value <	0,005
Mean	42,891
StDev	24,670
Variance	608,599
Skewness	0,615182
Kurtosis	-0,348834
N	100
Minimum	0,000
1st Quartile	25,000
Median	39,474
3rd Quartile	56,834
Maximum	100,000
95% Confidence Interval for Mean	
37,996	47,786
95% Confidence Interval for Median	
33,854	41,935
95% Confidence Interval for StDev	
21,660	28,658



Diante do Bloxpots apresentado acima é a amostra de 25 é a que está melhor colocada, mais próxima de 100. Sua mediana está mais próxima do 3º quartil, neste caso é **negativamente assimétrico.** A média de todas as caixas estão bem próximas e não nenhum outlier.

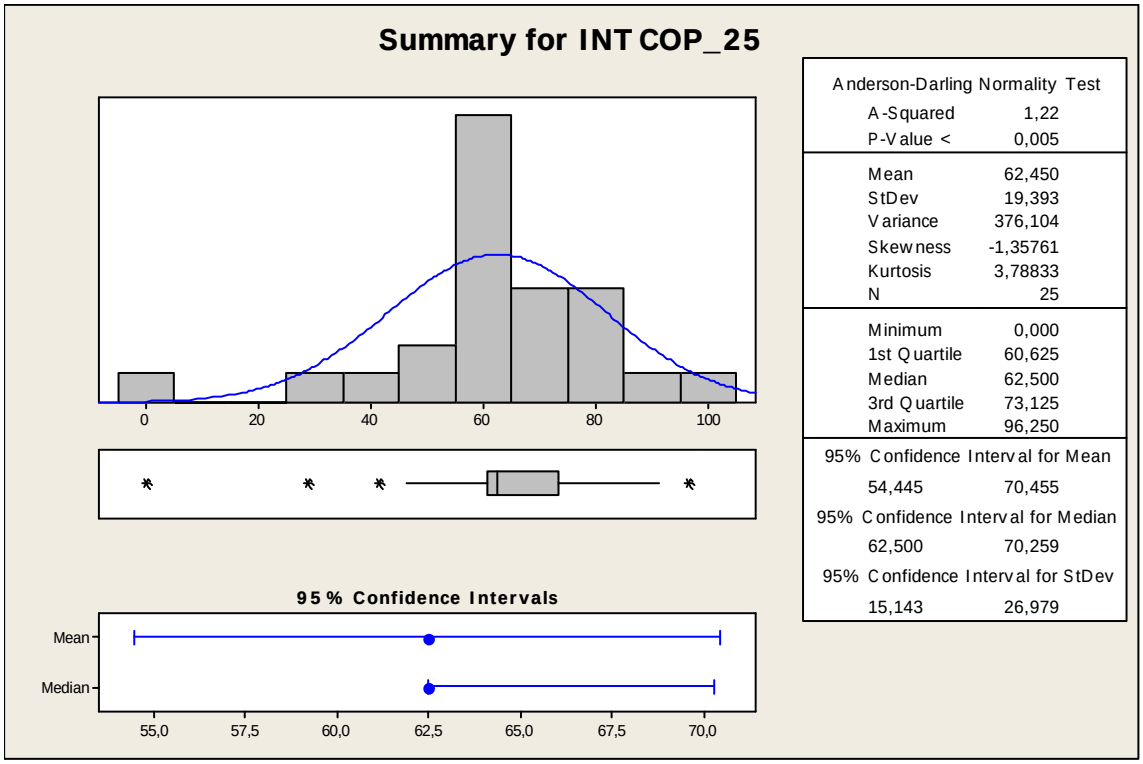
One-way ANOVA: INST_25; INST_50; INST_100; INSTIT-n

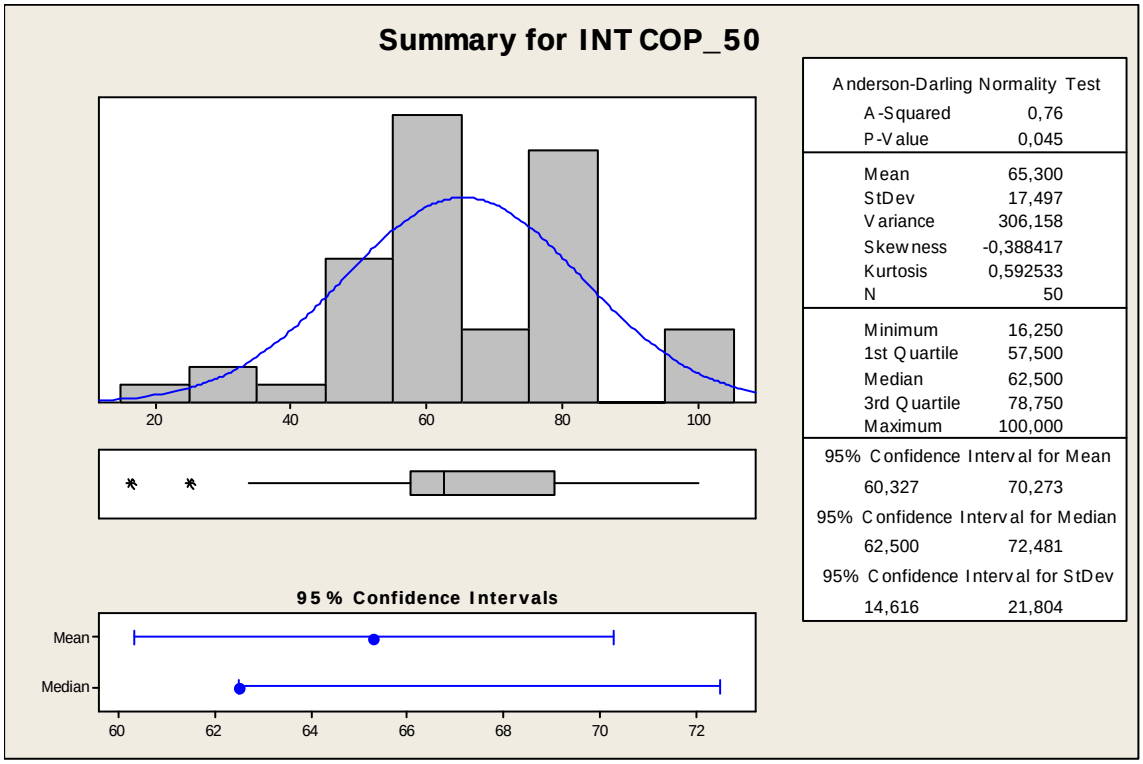
```
Source   DF      SS   MS     F      P
Factor    3     903  301  0,49  0,689
Error   303  186105  614
Total   306  187008

S = 24,78   R-Sq = 0,48%   R-Sq(adj) = 0,00%


                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level       N   Mean  StDev  -------+---------+---------+---------+--
INST_25    25  49,18  24,75     (---------------*---------------)
INST_50    50  45,37  24,51   (-----------*----------)
INST_100  100  42,89  24,67  (-------*--------)
INSTIT-n  132  43,61  24,98     (------*------)
                            -------+---------+---------+---------+--
                                42,0      48,0      54,0      60,0

Pooled StDev = 24,78
```

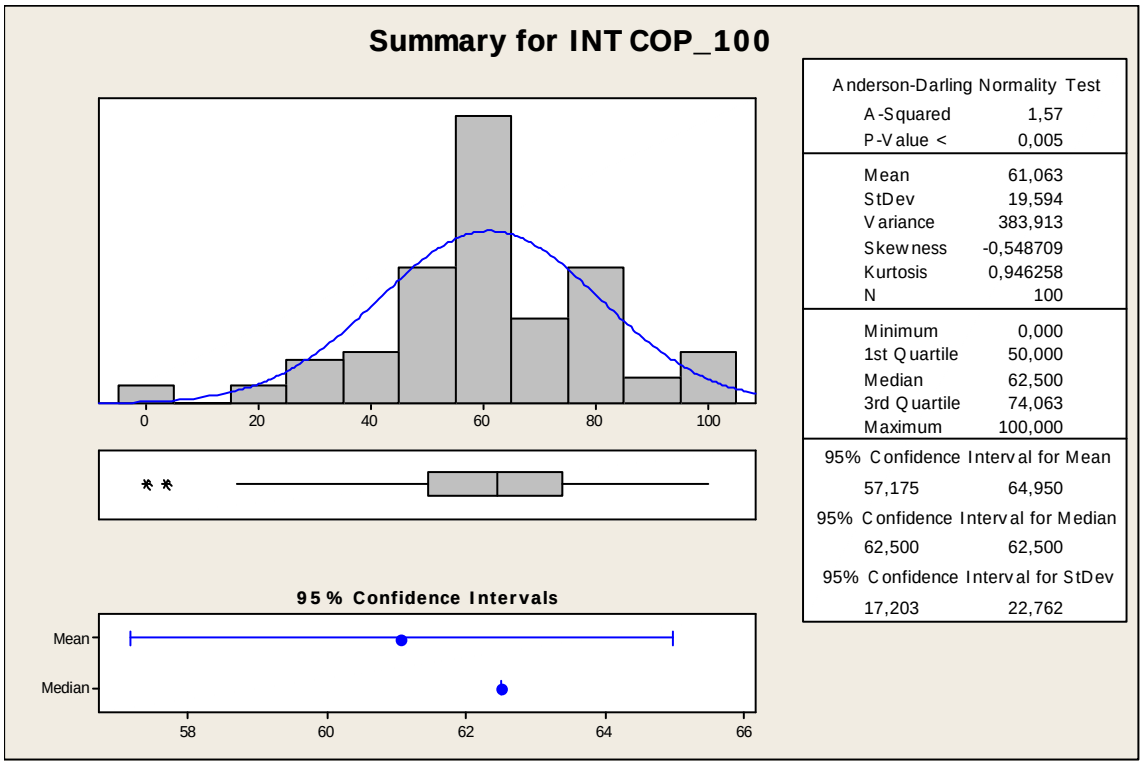
Diante da analise ANOVA que foi apresentada acima observa-se que a maioria das regiões estão sobrepostas, tendo as médias parecidas. Na amostragem com 100 países, está ficou com a média bem próxima da base de 132 países.

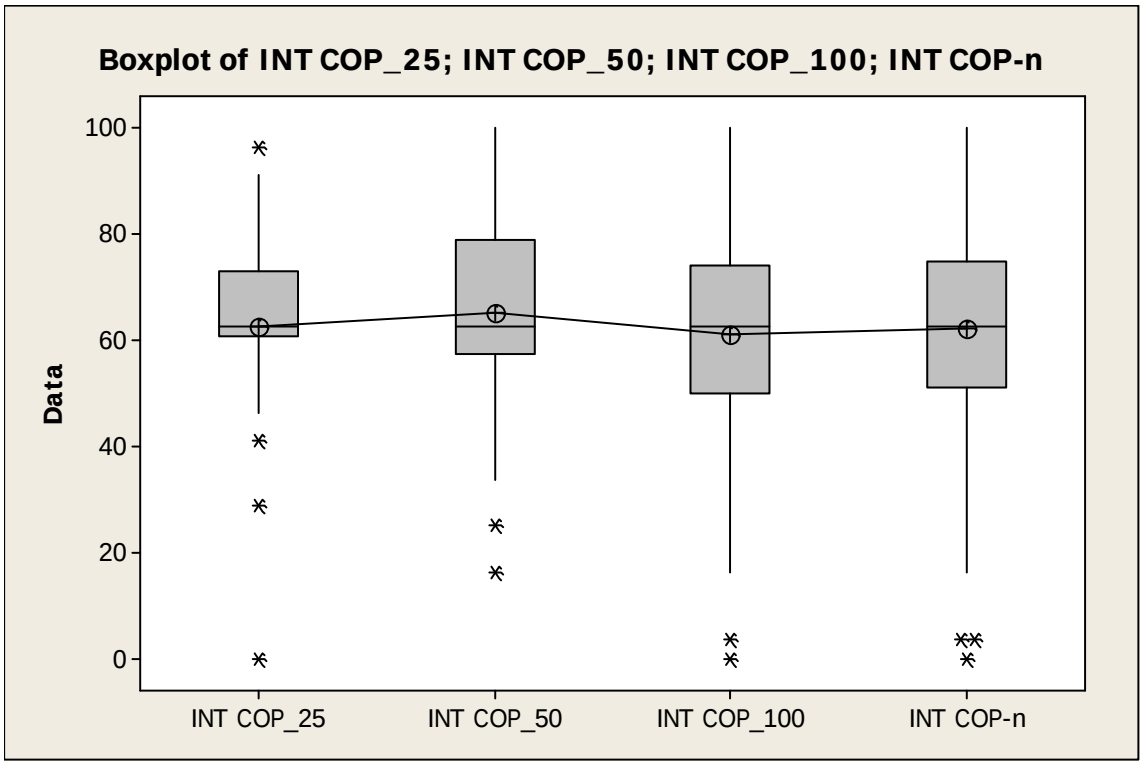
5.3 INTERNATIONAL COOPERATION

Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países. Em seguida o Bloxpots que faz o comparativo das regiões.









Na análise apresentada no Bloxpots acima há outliers em todas as amostras e no total de 132 países. A média de todas estão bem parecidas.

Na amostra de 25, a média e mediana estão muito próximas. Sua mediana está mais próxima do 1º quartil, neste caso é negativamente assimétrica.

One-way ANOVA: INT COP_25; INT COP_50; INT COP_100; INT COP-n

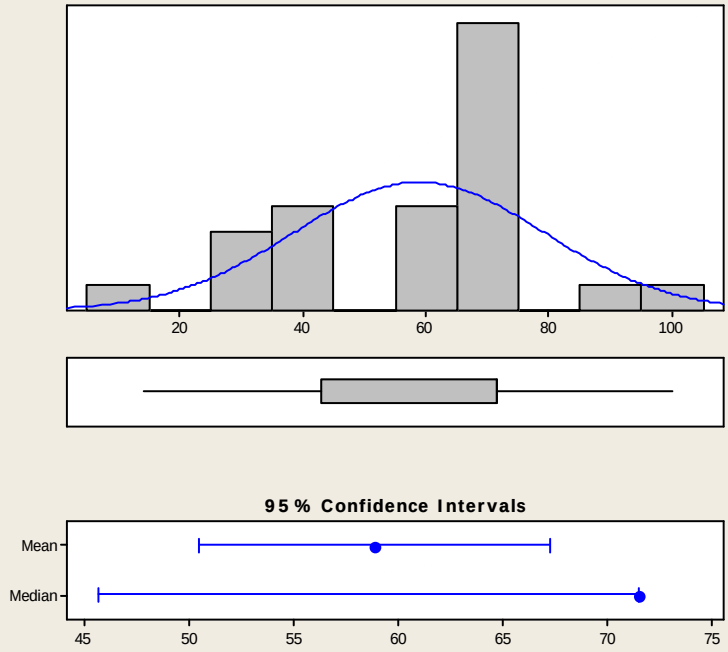
```
Source   DF      SS   MS     F      P
Factor    3     607  202  0,55  0,647
Error   303  111025  366
Total   306  111633

S = 19,14   R-Sq = 0,54%   R-Sq(adj) = 0,00%


                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level          N   Mean  StDev  ---+---------+---------+---------+------
INT COP_25    25  62,45  19,39  (------------------*------------------)
INT COP_50    50  65,30  17,50              (------------*-------------)
INT COP_100  100  61,06  19,59       (---------*--------)
INT COP-n    132  62,13  19,34           (-------*--------)
                                ---+---------+---------+---------+------
                                56,0      60,0      64,0      68,0

Pooled StDev = 19,14
```

Diante da analise ANOVA que foi apresentada acima observa-se que a maioria das regiões estão sobrepostas, tendo as médias parecidas. A amostra de 100 é que mais se assemelha ao total de 132 países. Na análise do MINITAB F apresenta o número 0,55 que significa que as regiões não são todas iguais e o gráfico mostra que os intervalos não estão totalmente separados uns dos outros.

5.4 REGIONAL COOPERATION

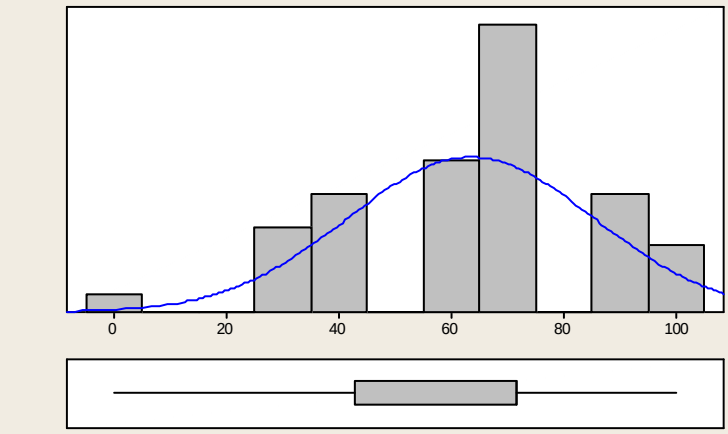
Abaixo apresenta-se a análise com as amostragens com 25, 50, 100 e com os 132 países.

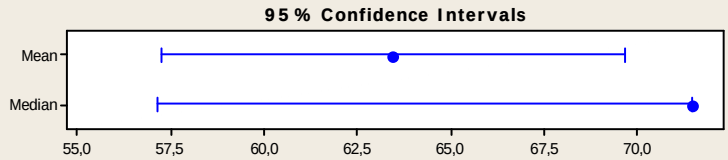
Summary for REG COP_25



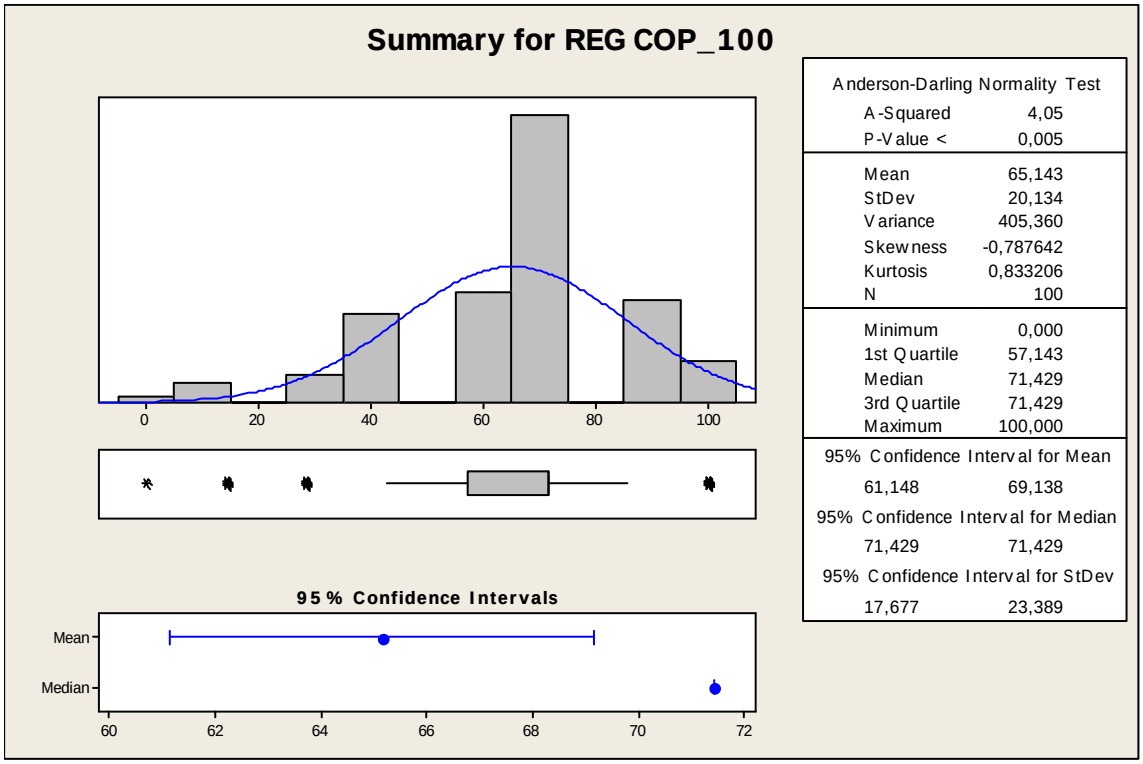
Anderson-Darling Normality Test	
A-Squared	1,23
P-Value <	0,005
Mean	58,857
StDev	20,337
Variance	413,605
Skewness	-0,416822
Kurtosis	-0,163221
N	25
Minimum	14,286
1st Quartile	42,857
Median	71,429
3rd Quartile	71,429
Maximum	100,000
95% Confidence Interval for Mean	
50,462	67,252
95% Confidence Interval for Median	
45,687	71,429
95% Confidence Interval for StDev	
15,880	28,292

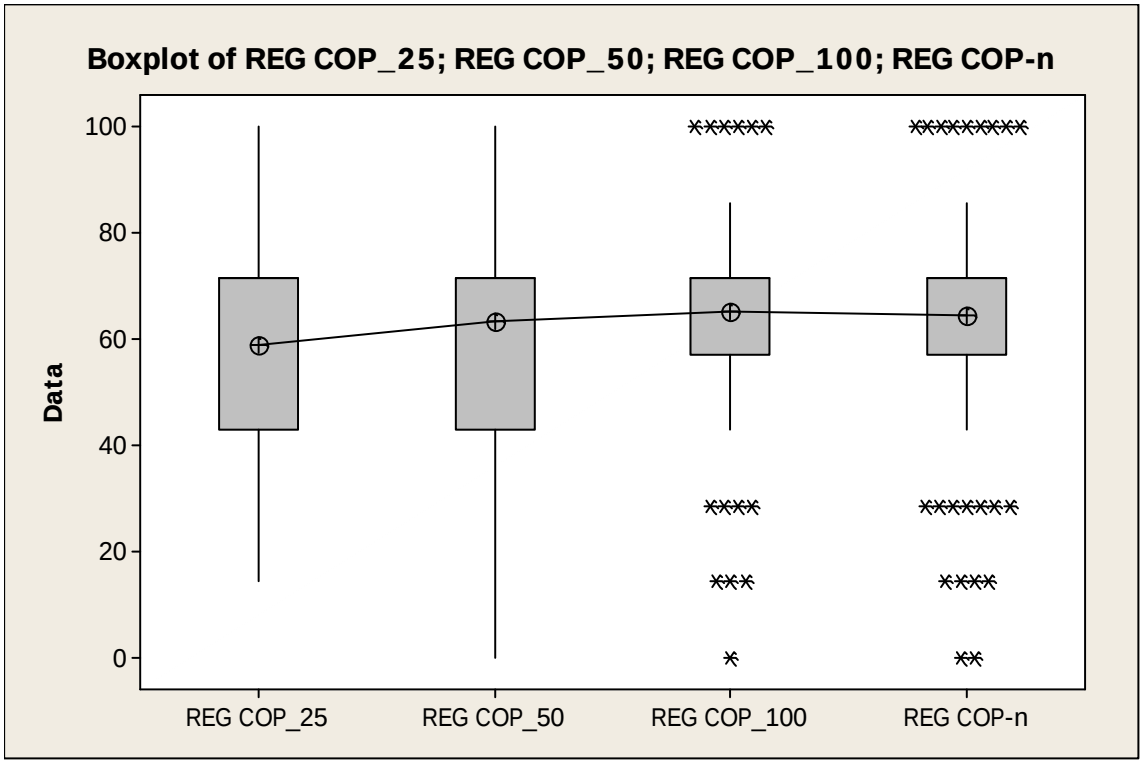
Summary for REG COP_50





Anderson-Darling Normality Test	
A-Squared	1,30
P-Value <	0,005
Mean	63,429
StDev	21,821
Variance	476,135
Skewness	-0,508028
Kurtosis	0,245050
N	50
Minimum	0,000
1st Quartile	42,857
Median	71,429
3rd Quartile	71,429
Maximum	100,000
95% Confidence Interval for Mean	
57,227	69,630
95% Confidence Interval for Median	
57,143	71,429
95% Confidence Interval for StDev	
18,227	27,191





Na análise apresentada no Bloxpots acima há outliers na amostra de 100 e no total de 132 países. A média de todas as caixas estão bem parecidas.

One-way ANOVA: REG COP_25; REG COP_50; REG COP_100; REG COP-n

```
Source   DF      SS   MS     F      P
Factor    3     840  280  0,64  0,587
Error   303  131545  434
Total   306  132384

S = 20,84   R-Sq = 0,63%   R-Sq(adj) = 0,00%


                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level          N   Mean  StDev  ---------+---------+---------+---------+
REG COP_25    25  58,86  20,34  (----------------*---------------)
REG COP_50    50  63,43  21,82               (-----------*----------)
REG COP_100  100  65,14  20,13                      (-------*-------)
REG COP-n    132  64,50  21,07                      (------*------)
                                ---------+---------+---------+---------+
                                      55,0      60,0      65,0      70,0

Pooled StDev = 20,84
```

Diante da analise ANOVA que foi apresentada acima observa-se que as amostras de 50 e 100 estão bem parecidas com a de 132 países. Na análise do MINITAB F apresenta o número 0,64 onde as médias são bem similares.

5.5. AMOSTRAGEM - CONSIDERAÇÕES FINAIS

Como as amostras tiveram médias semelhantes ao total de 132 países, isso demonstra que é possível utilizar-se de média para a representação de uma população.

6.COMPONENTES PRINCIPAIS

O presente trabalho tem por objetivo efetuar uma análise das correlações e dos componentes principais (análise multivariada) de dados dimensionadores do nível de ***Integração e Alianças Estratégicas*** dos países Ibero-Americanos. Para tal, iniciamos com análise da estatística descritiva. Em seguida passamos para a análise das correlações e dendogramas. Na terceira parte utilizamos a análise dos componentes principais. Por fim, fazemos as considerações finais. Os dados são compilados pelo Banco Mundial. O software estatístico utilizado é o **MINITAB**

6.1. ANÁLISES DE COMPONENTES PRINCIPAIS

Inicialmente, para análises de componentes principais, com base de todas as variáveis (12) pré-selecionadas, foram criadas 4 novas variáveis: PC1, PC2, PC3 e PC4, para analisar e decidir quais delas podem melhor representar as demais 12 variáveis.

Tabela 1 – Componentes Principais – 12 Variáveis

Variable	PC1	PC2	PC3	PC4
SPI-n	0,361	-0,219	-0,179	-0,196
IDH-n	0,301	-0,279	-0,229	-0,322
GOV-n	0,376	-0,124	-0,112	0,097
INSTIT-n	0,361	-0,164	-0,058	0,164
INT COP-n	0,286	0,430	0,119	-0,046
EFFEC-n	0,283	0,386	0,066	0,038
CRED-n	0,282	0,401	0,092	0,019
REG COP-n	0,282	0,339	0,148	-0,179
EPI-n	0,333	-0,270	-0,128	-0,222
OHI-n	0,256	-0,197	0,119	0,833
GINNI-n	-0,097	0,332	-0,906	0,208

Na tabela 1 apresenta o peso de cada variável nas componentes PC1, PC2, PC3 e PC4, analisando, vemos que a componente PC1 é a única componente que tem somente uma variável negativa (GINNI).

Tabela 2 – Matriz de Correlação – 12 Variáveis

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalue	6,0474	2,6514	1,2406	0,7203	0,4251	0,3006	0,2597	0,1534
Proportion	0,504	0,221	0,103	0,060	0,035	0,025	0,022	0,013
Cumulative	0,504	0,725	0,828	0,888	0,924	0,949	0,970	0,983

	PC9	PC10	PC11	PC12
Eigenvalue	0,1180	0,0520	0,0291	0,0023
Proportion	0,010	0,004	0,002	0,000
Cumulative	0,993	0,997	1,000	1,000

Na Tabela 2, pode-se verificar que a componente PC1 explica 0,504% da variabilidade dos dados, e agregada com as componentes PC2 e PC3, chegam a explicam 82,8%.

Conforme, Las Casas e Guevara (2010)[3], por convenção somente deve ser consideradas a componentes com uma contribuição igual ou superior a 1. Partindo dessa convenção, poder-se-á somente considerar as componentes PC1 a PC3, as quais acumuladas explicam 88,8% da variabilidade dos dados. O gráfico 1 abaixo confirma está afirmação, as 3 primeiras componentes são maiores que 1.

[3] [3] LAS CASAS, Alexandre Luzzi; GUEVARA, **Arnoldo José de Hoyos. Pesquisa de Marketing**. São Paulo: Atlas, 2010.

Gráfico 1 – Análise das componentes

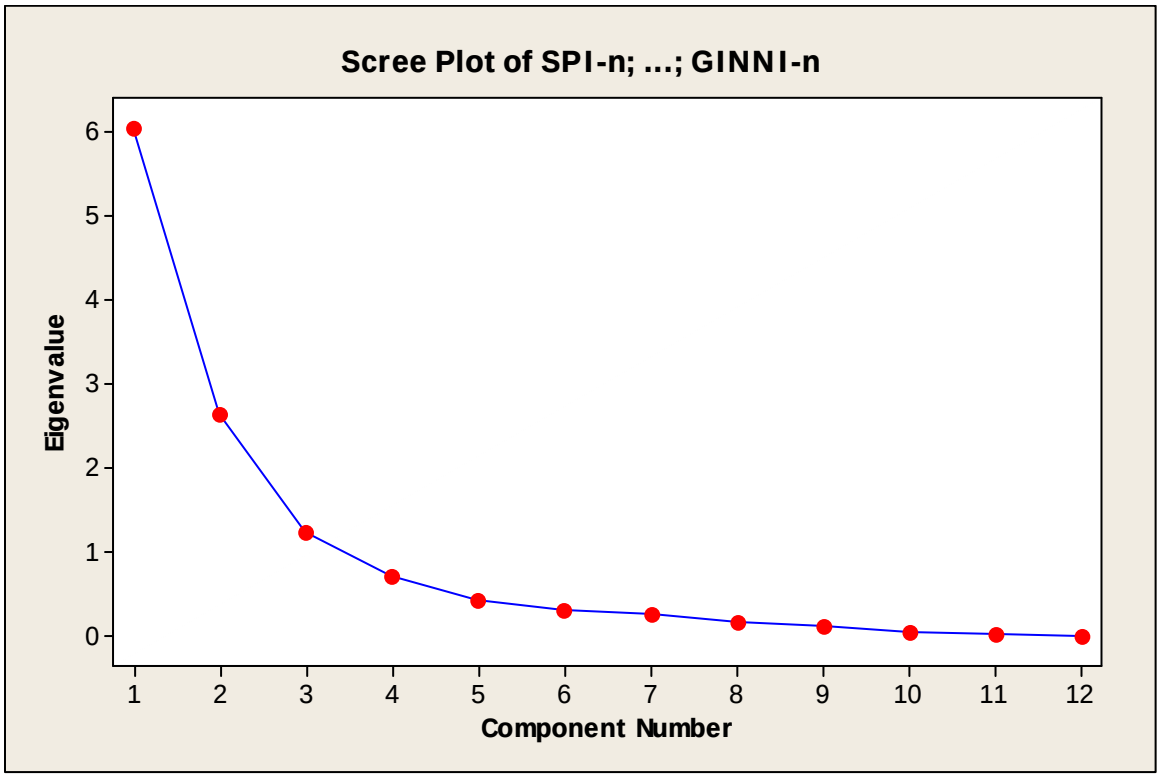
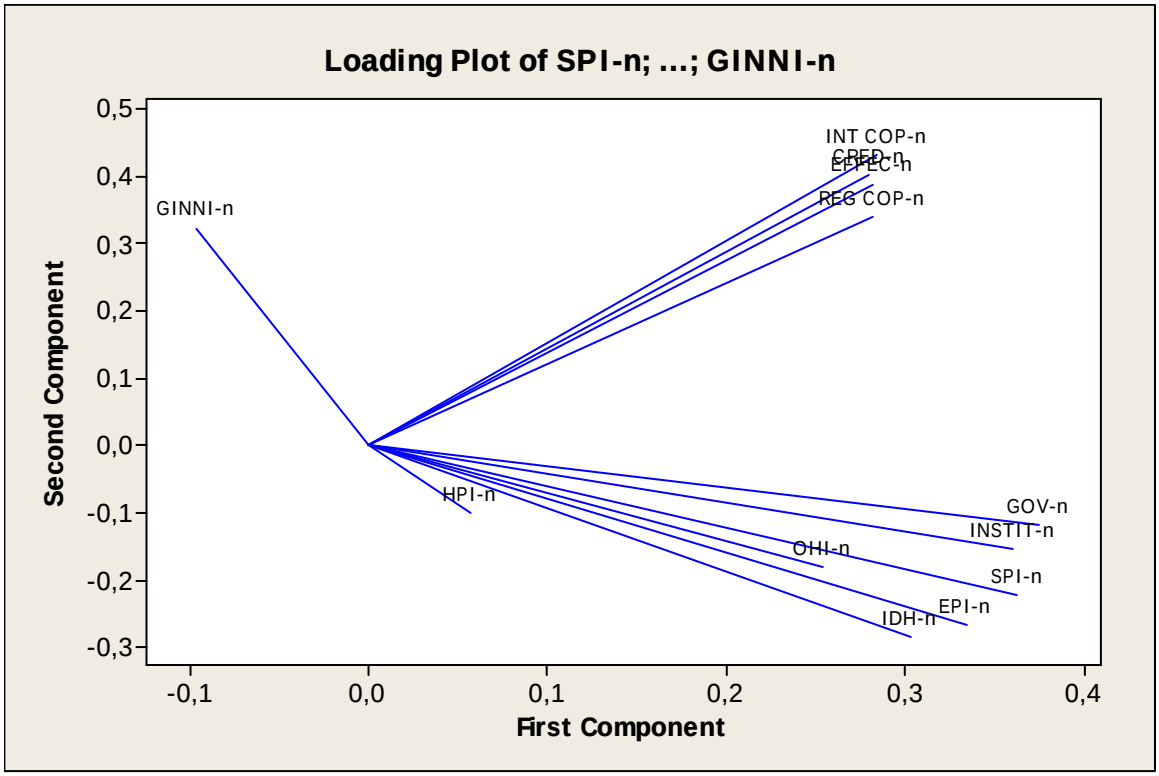


Gráfico 2 – Análise das 12 variáveis



No Gráfico 2 identifica-se claramente as 2 variáveis GINNI e HPI, todas elas concentradas no lado esquerdo do gráfico e OHI também evidenciando a baixa similaridade com as demais.

6.3 Análise dos clusters

A análise de clusters nos ajuda a compreender a similariedade dos agrupamentos das variáveis.

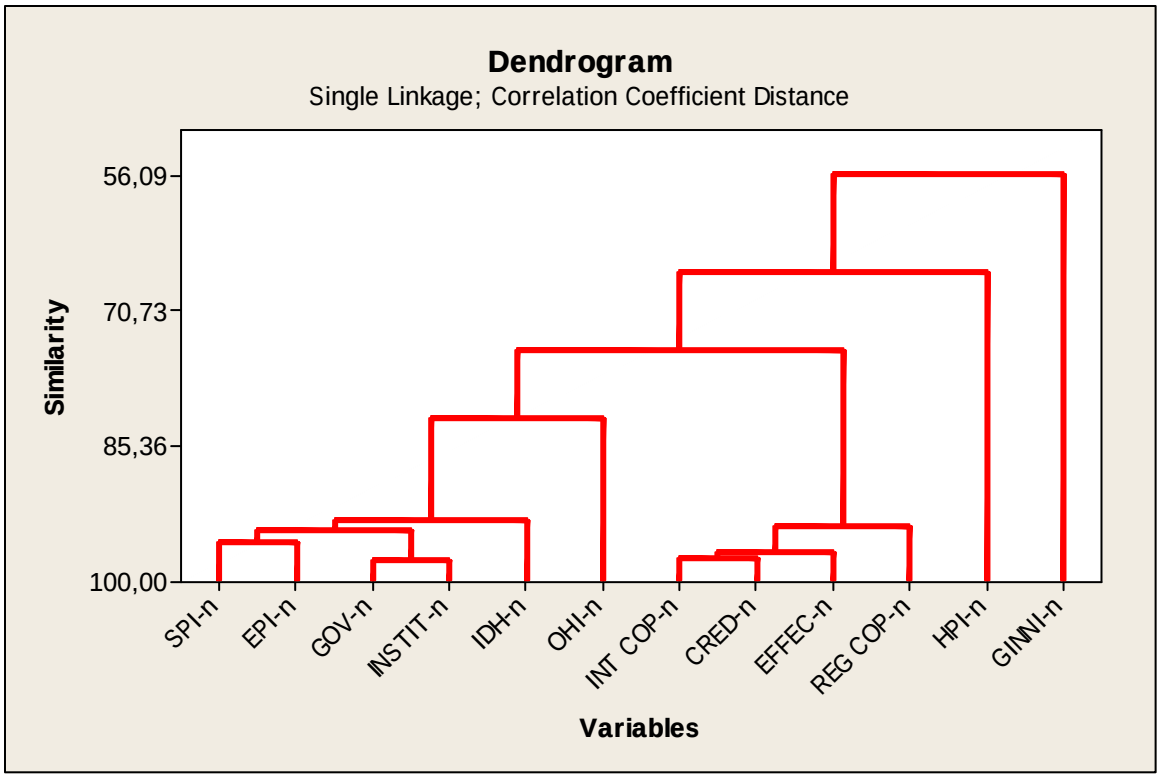
Cluster Analysis of Variables: SPI-n; IDH-n; GOV-n; INSTIT-n; INT COP-n; ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	11	97,8387	0,043227	3	4	3	2
2	10	97,6876	0,046247	5	7	5	2
3	9	96,8707	0,062585	5	6	5	3
4	8	95,8701	0,082598	1	9	1	2
5	7	94,4907	0,110186	1	3	1	4
6	6	94,1283	0,117433	5	8	5	4
7	5	93,5636	0,128728	1	2	1	5
8	4	82,3853	0,352293	1	11	1	6
9	3	74,9859	0,500283	1	5	1	10
10	2	66,6038	0,667924	1	10	1	11
11	1	56,0883	0,878234	1	12	1	12

Na tabela acima demonstra que são necessários 11 agrupamentos para se ter 97,84% de similariedade 3 agrupamentos para ter 74,9% de similariedade. O dendograma abaixo apresenta esta similariedades e ainda destaca as variáveis que estão fora destes agrupamentos.

Gráfico 3 – Dendograma – 12 Variáveis



6.4 CORRELAÇÕES

As correlações abaixo apresentaram resultados não significativos (p > 0.05). E os coeficientes de correlação de Pearson com valores próximos a 0 (zero) indica uma relação fraca ou inexistente entre as duas variáveis.

Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	0,000		
	1,000		
PC3	-0,000	-0,000	
	1,000	1,000	
PC4	-0,000	0,000	-0,000
	1,000	1,000	1,000

Cell Contents: Pearson correlation
P-Value

6.5 Regressões stepwise

Abaixo serão feiras as regressões stepwise para cada PC, com o objetivo de verificar as variáveis que mais se destaca.

Stepwise Regression: PC1 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-4,010	-5,962	-6,971	-7,125	-7,566	-7,583
GOV-n	0,0891	0,0736	0,0517	0,0415	0,0353	0,0228
T-Value	27,20	30,35	26,13	22,66	27,01	11,14
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
INT COP-n		0,04265	0,04873	0,05070	0,05190	0,05280
T-Value		13,36	25,60	34,80	53,30	63,68
P-Value		0,000	0,000	0,000	0,000	0,000
IDH-n			0,02769	0,01792	0,01730	0,01764
T-Value			15,82	10,79	15,64	18,92
P-Value			0,000	0,000	0,000	0,000
EPI-n				0,0216	0,0212	0,0205
T-Value				9,78	14,45	16,54
P-Value				0,000	0,000	0,000
OHI-n					0,01292	0,01159
T-Value					12,69	13,22
P-Value					0,000	0,000
INSTIT-n						0,0140

T-Value						7,26
P-Value						0,000
S	0,954	0,621	0,362	0,275	0,183	0,154
R-Sq	85,05	93,73	97,88	98,79	99,47	99,63
R-Sq(adj)	84,94	93,63	97,83	98,75	99,45	99,61

Para o PC1, temos as variáveis GOV, International Cooperation, IDH-2014 que explicam 97,88% dos dados.

Stepwise Regression: PC2 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-3,671	-2,493	-3,136	-2,823	-2,632	-2,569
INT COP-n	0,05908	0,08109	0,07600	0,07652	0,07469	0,07409
T-Value	11,23	39,67	59,66	75,96	98,51	126,04
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
SPI-n		-0,04557	-0,03978	-0,03612	-0,02600	-0,01822
T-Value		-29,42	-39,52	-40,32	-22,33	-14,88
P-Value		0,000	0,000	0,000	0,000	0,000
GINNI-n			0,01724	0,01596	0,01583	0,01502
T-Value			15,14	17,52	23,73	28,79
P-Value			0,000	0,000	0,000	0,000
OHI-n				-0,00933	-0,00931	-0,00891
T-Value				-8,85	-12,07	-14,94
P-Value				0,000	0,000	0,000
IDH-n					-0,01094	-0,01020
T-Value					-10,53	-12,67
P-Value					0,000	0,000
EPI-n						-0,0101
T-Value						-9,33
P-Value						0,000
S	1,16	0,421	0,253	0,200	0,146	0,113
R-Sq	49,23	93,41	97,64	98,54	99,22	99,54
R-Sq(adj)	48,84	93,31	97,58	98,49	99,19	99,52

Para o PC2, temos as variáveis International Cooperation, SPI e GINNI que explicam 97,64% dos dados.

Stepwise Regression: PC3 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	2,275	1,255	1,821	1,916	1,928	1,913

HPI-n	-0,04696	-0,04467	-0,04458	-0,03912	-0,03879	-0,03759
T-Value	-24,88	-34,05	-46,44	-94,46	-109,12	-182,40
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
OHI-n		0,01689	0,01365	0,02061	0,01926	0,01937
T-Value		12,08	12,80	43,25	42,93	79,10
P-Value		0,000	0,000	0,000	0,000	0,000
GINNI-n			-0,01075	-0,01322	-0,01313	-0,01325
T-Value			-10,63	-33,57	-39,13	-72,34
P-Value			0,000	0,000	0,000	0,000
IDH-n				-0,01103	-0,01252	-0,00904
T-Value				-27,63	-31,28	-30,46
P-Value				0,000	0,000	0,000
INSTIT-n					0,00295	0,00700
T-Value					7,03	21,37
P-Value					0,000	0,000
SPI-n						-0,00760
T-Value						-17,28
P-Value						0,000
S	0,466	0,320	0,234	0,0889	0,0756	0,0412
R-Sq	82,64	91,85	95,67	99,38	99,56	99,87
R-Sq(adj)	82,51	91,73	95,57	99,36	99,54	99,86

Para o PC3, temos as variáveis HPI, OHI, GINNI e IDH que explicam 99,38% dos dados.

Stepwise Regression: PC4 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-1,104	-2,430	-1,784	-1,669	-1,379	-1,343
GINNI-n	0,02997	0,03569	0,03858	0,04147	0,04169	0,04123
T-Value	12,79	20,64	31,91	59,38	85,61	113,64
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
OHI-n		0,02071	0,02484	0,01635	0,01487	0,01468
T-Value		11,46	19,34	18,60	23,77	31,71
P-Value		0,000	0,000	0,000	0,000	0,000
INT COP-n			-0,01570	-0,02110	-0,02127	-0,01434
T-Value			-12,11	-26,58	-38,44	-18,15
P-Value			0,000	0,000	0,000	0,000
INSTIT-n				0,01311	0,01409	0,01453
T-Value				16,78	25,58	35,48
P-Value				0,000	0,000	0,000
HPI-n					-0,00517	-0,00487
T-Value					-11,65	-14,76
P-Value					0,000	0,000
REG COP-n						-0,00735
T-Value						-10,27
P-Value						0,000

```
S           0,567    0,401     0,275     0,154     0,107    0,0792
R-Sq        55,72    78,05     89,77     96,82     98,47     99,17
R-Sq(adj)   55,38    77,71     89,53     96,72     98,41     99,13
```

Para o PC4, temos as variáveis GINNI, OHI, International Cooperation e Institution que explicam 96,82% dos dados.

6.6 Análise de regressão com as principais variáveis das PCs

Regression Analysis: PC1 versus GOV-n; INT COP-n; IDH-n

```
The regression equation is
PC1 = - 6,97 + 0,0517 GOV-n + 0,0487 INT COP-n + 0,0277 IDH-n


Predictor      Coef   SE Coef       T      P
Constant    -6,9711    0,1244  -56,03  0,000
GOV-n      0,051708  0,001979   26,13  0,000
INT COP-n  0,048730  0,001903   25,60  0,000
IDH-n      0,027695  0,001750   15,82  0,000


S = 0,362361   R-Sq = 97,9%   R-Sq(adj) = 97,8%


Analysis of Variance

Source           DF      SS      MS        F      P
Regression        3  775,40  258,47  1968,44  0,000
Residual Error  128   16,81    0,13
Total           131  792,21


Source     DF  Seq SS
GOV-n       1  673,80
INT COP-n   1   68,72
IDH-n       1   32,87


Unusual Observations

Obs  GOV-n      PC1      Fit  SE Fit  Residual  St Resid
  3     15  -4,4738  -3,5123  0,0573   -0,9614     -2,69R
 15     65   1,4435   2,2396  0,0563   -0,7961     -2,22R
 44     48  -0,1937   0,6849  0,0581   -0,8786     -2,46R
 54     14  -4,6790  -4,3575  0,1128   -0,3215     -0,93 X
 70     44  -2,1780  -1,3874  0,0705   -0,7906     -2,22R
113     31  -0,6388   0,2451  0,0873   -0,8839     -2,51R
115     97   2,7007   1,9453  0,1416    0,7554      2,26RX

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC1 com as principais variáveis tem um poder de explicação de 97,9%.

Regression Analysis: PC2 versus INT COP-n; SPI-n; GINNI-n

```
The regression equation is
PC2 = - 3,14 + 0,0760 INT COP-n - 0,0398 SPI-n + 0,0172 GINNI-n


Predictor       Coef   SE Coef       T      P
Constant    -3,13586   0,08894  -35,26  0,000
INT COP-n   0,075999  0,001274   59,66  0,000
SPI-n      -0,039778  0,001007  -39,52  0,000
GINNI-n     0,017245  0,001139   15,14  0,000


S = 0,253077   R-Sq = 97,6%   R-Sq(adj) = 97,6%


Analysis of Variance

Source           DF      SS      MS        F      P
Regression        3  339,13  113,04  1764,98  0,000
Residual Error  128    8,20    0,06
Total           131  347,33


Source     DF  Seq SS
INT COP-n   1  170,99
SPI-n       1  153,46
GINNI-n     1   14,68


Unusual Observations

       INT
Obs  COP-n      PC2      Fit  SE Fit  Residual  St Resid
 28     54   1,1208   0,5396  0,0332    0,5812      2,32R
 48     63   1,0601   0,5384  0,0269    0,5218      2,07R
 54      0  -4,2869  -4,2495  0,0774   -0,0374     -0,16 X
 77     38  -0,4021  -0,9356  0,0448    0,5336      2,14R
105     79   2,5670   2,0505  0,0389    0,5165      2,07R
113     63   0,9459   1,6511  0,0429   -0,7052     -2,83R
115     63  -1,3149  -2,0213  0,0465    0,7064      2,84R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC2 com as principais variáveis tem um poder de explicação de 97,6%.

Regression Analysis: PC3 versus HPI-n; OHI-n; GINNI-n

```
The regression equation is
PC3 = 1,82 - 0,0446 HPI-n + 0,0137 OHI-n - 0,0108 GINNI-n


Predictor        Coef    SE Coef       T      P
Constant      1,82084    0,09586   18,99  0,000
HPI-n      -0,0445821  0,0009600  -46,44  0,000
OHI-n        0,013653   0,001067   12,80  0,000
GINNI-n     -0,010753   0,001012  -10,63  0,000


S = 0,234395   R-Sq = 95,7%   R-Sq(adj) = 95,6%


Analysis of Variance
```

```
Source           DF       SS      MS       F      P
Regression        3  155,488  51,829  943,36  0,000
Residual Error  128    7,032   0,055
Total           131  162,520


Source   DF   Seq SS
HPI-n     1  134,309
OHI-n     1   14,974
GINNI-n   1    6,205


Unusual Observations

Obs  HPI-n      PC3      Fit  SE Fit  Residual  St Resid
  9     81  -0,8195  -1,3169  0,0399    0,4974      2,15R
 15      0   1,2070   1,5603  0,0766   -0,3532     -1,59 X
 65     11   0,8877   1,5588  0,0441   -0,6711     -2,92R
 70     48   0,0836  -0,4008  0,0452    0,4844      2,11R
109     14   0,6051   0,8765  0,0752   -0,2714     -1,22 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC3 com as principais variáveis tem um poder de explicação de 95,7%.

Regression Analysis: PC4 versus GINNI-n; OHI-n; INT COP-n

```
The regression equation is
PC4 = - 1,78 + 0,0386 GINNI-n + 0,0248 OHI-n - 0,0157 INT COP-n


Predictor       Coef   SE Coef       T      P
Constant     -1,7835    0,1070  -16,66  0,000
GINNI-n     0,038582  0,001209   31,91  0,000
OHI-n       0,024842  0,001285   19,34  0,000
INT COP-n  -0,015699  0,001297  -12,11  0,000


S = 0,274619   R-Sq = 89,8%   R-Sq(adj) = 89,5%


Analysis of Variance

Source           DF      SS      MS       F      P
Regression        3  84,708  28,236  374,41  0,000
Residual Error  128   9,653   0,075
Total           131  94,362


Source     DF  Seq SS
GINNI-n     1  52,574
OHI-n       1  21,080
INT COP-n   1  11,054


Unusual Observations

Obs  GINNI-n      PC4      Fit  SE Fit  Residual  St Resid
  9       35  -0,5032   0,2052  0,0273   -0,7084     -2,59R
 15       93   2,3620   1,7767  0,0734    0,5853      2,21R
 37       15  -0,9480  -0,3968  0,0371   -0,5512     -2,03R
 46       72   1,2534   1,8128  0,0600   -0,5594     -2,09R
131       29  -1,4146  -0,7010  0,0341   -0,7136     -2,62R
```

```
R denotes an observation with a large standardized residual.
```

A equação apresentada para a PC4 com as principais variáveis tem um poder de explicação de 89,8%.

6.8 Criando indicadores a partir dos PC1, PC2, PC3 e PC4

Aproveitamos os resultados da sessão anterior, que através da análise de regressão *stepwise*, relacionamos cada componente principal (PC), com as 12 variáveis afim de obter a mais alta correlação com as variáveis de resposta.

Abaixo, listamos as 3 (três) variáveis que apresentaram esta correlação, uma sugestão para o nome do indicador proposto e uma sugestão para o nome da variável.

PC1 = GOV + INT COP + IDH – Nome do indicador: Cooperação e desenvolvimento humano

PC2 = INT COP + SPI + GINNI – Nome do indicador: Progresso e cooperação do país

PC3 = HPI + OHI + GINNI – Nome do indicador: Desenvolvimento sustentável humano

PC4 = GINNI + OHI + INT COP – Nome do indicador:Cooperação para o bem-estar financeiro e dos oceanos

6.9. COMPONENTES PRINCIPAIS - CONSIDERAÇÕES FINAIS

O presente trabalho teve como objetivo realizar análises de componentes principais para construção de índice com alto poder explicativo com o menor número possível de variáveis.

O estudo foi iniciado com 12 variáveis e ao final, mediante as análises das componentes principais, chegamos a 4 componentes com 3 variáveis cada com alto poder de explicação. Sendo que os 4 componentes explicam 88,8%, sendo que são as seguintes variáveis: GOV + INT COP + IDH + OHI + GINNI + HPI + SPI

Para o tema em questão Integração e Aliança Estratégica, observou-se que nas PC´s apareceram variáveis a ligadas ao âmbito ecológico como OHI.

7. ANÁLISE DE CONGLOMERADOS

7.1. ANÁLISES DAS NOVAS VARIÁVEIS EXPLICATIVAS

Abaixo segue as variáveis mais significativas para cada PC e também o poder de explicação da equação gerada com elas para cada PC.

PC1 = GOV + INT COP + IDH – Nome do indicador: <u>Cooperação e desenvolvimento humano</u>

```
PC1 = - 6,97 + 0,0517 GOV-n + 0,0487 INT COP-n + 0,0277 IDH-n
R-Sq = 97,9%
```

PC2 = INT COP + SPI + GINNI – Nome do indicador: <u>Progresso e cooperação do país</u>

```
PC2 = - 3,14 + 0,0760 INT COP-n - 0,0398 SPI-n + 0,0172 GINNI-n

R-Sq = 97,6%
```

PC3 = HPI + OHI + GINNI – Nome do indicador: <u>Desenvolvimento sustentável humano</u>

```
PC3 = 1,82 - 0,0446 HPI-n + 0,0137 OHI-n - 0,0108 GINNI-n
R-Sq = 95,7%
```

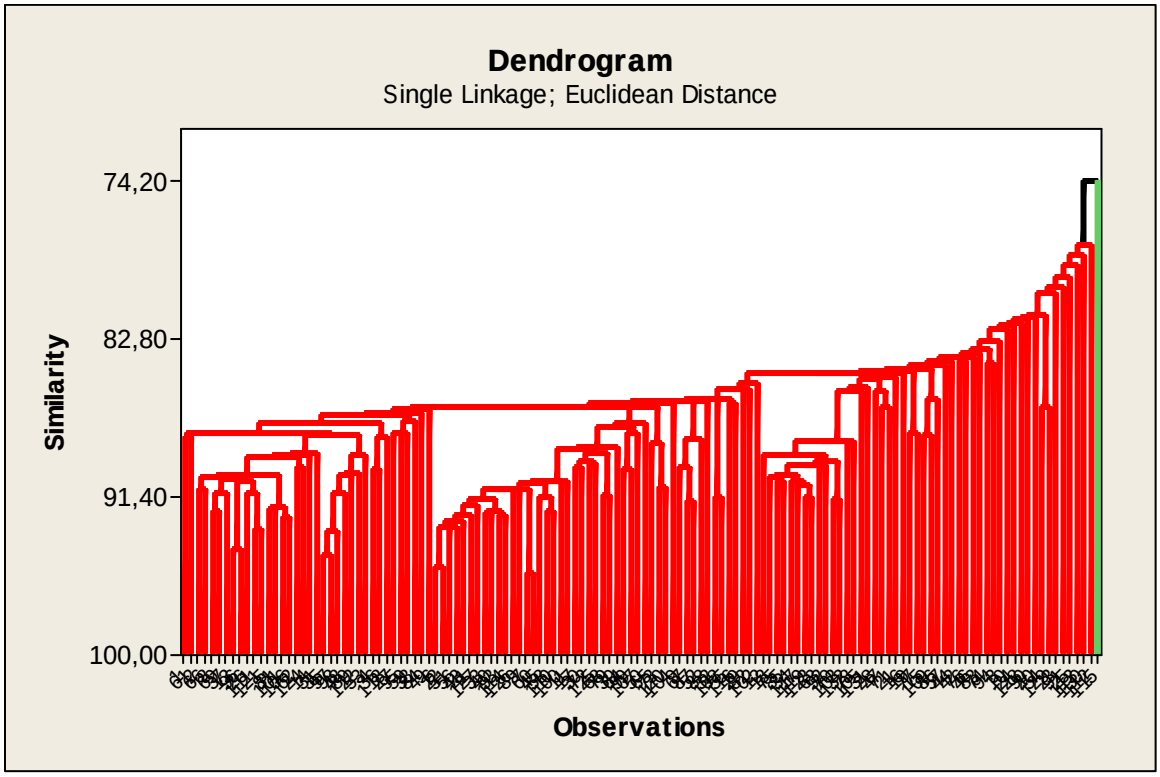
PC4 = GINNI + OHI + INT COP – Nome do indicador: <u>Cooperação para o bem-estar financeiro e dos oceanos</u>

```
PC4 = - 1,78 + 0,0386 GINNI-n + 0,0248 OHI-n - 0,0157 INT COP-n
R-Sq = 89,8%
```

7.2 ANÁLISES DE CLUSTERS

7.2.1 –ANÁLISE COM 2 CLUSTERS

A análise resultou em 2 clusters 1 com 131 observações e outro com 1, como mostra a tabela abaixo.



Cluster Analysis of Observations: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,6765	8,6821	7	60	7	2
2	130	95,2515	9,5355	6	22	6	2
3	129	94,6384	10,7666	35	97	35	2
4	128	94,3569	11,3318	13	96	13	2
5	127	93,3660	13,3218	35	98	35	3
6	126	93,1904	13,6743	111	121	111	2
7	125	93,1141	13,8276	51	56	51	2
8	124	93,0943	13,8673	6	51	6	4
9	123	92,7531	14,5525	6	43	6	5
10	122	92,5067	15,0474	101	106	101	2
11	121	92,4641	15,1328	6	41	6	6
12	120	92,4009	15,2598	93	114	93	2
13	119	92,2886	15,4854	33	40	33	2
14	118	92,2825	15,4976	33	93	33	4
15	117	92,2356	15,5917	8	69	8	2
16	116	92,1948	15,6736	58	110	58	2
17	115	92,0977	15,8686	5	81	5	2
18	114	92,0119	16,0409	5	101	5	4
19	113	91,9338	16,1978	6	127	6	7
20	112	91,6873	16,6928	67	87	67	2
21	111	91,5682	16,9320	70	103	70	2
22	110	91,5509	16,9666	6	33	6	11
23	109	91,4864	17,0962	119	123	119	2
24	108	91,4720	17,1252	32	108	32	2
25	107	91,4146	17,2404	45	58	45	3
26	106	91,3045	17,4615	68	79	68	2
27	105	91,2764	17,5179	8	37	8	3
28	104	91,2719	17,5270	26	111	26	3
29	103	91,2337	17,6037	35	48	35	4
30	102	91,0489	17,9748	2	66	2	2
31	101	90,9767	18,1197	6	126	6	12

32	100	90,9760	18,1212	6	89	6	13
33	99	90,9754	18,1223	6	7	6	15
34	98	90,9166	18,2404	73	120	73	2
35	97	90,8519	18,3704	117	119	117	3
36	96	90,6858	18,7039	6	45	6	18
37	95	90,6258	18,8243	6	100	6	19
38	94	90,6029	18,8705	75	85	75	2
39	93	90,5769	18,9226	6	11	6	20
40	92	90,5662	18,9441	13	26	13	5
41	91	90,5402	18,9962	34	117	34	4
42	90	90,4519	19,1736	8	13	8	8
43	89	90,4087	19,2603	2	8	2	10
44	88	90,4037	19,2704	18	75	18	3
45	87	90,3061	19,4664	18	34	18	7
46	86	90,2797	19,5194	2	5	2	14
47	85	90,2439	19,5912	35	80	35	5
48	84	90,1891	19,7013	35	122	35	6
49	83	89,9552	20,1711	4	36	4	2
50	82	89,9381	20,2054	64	99	64	2
51	81	89,8427	20,3968	9	67	9	3
52	80	89,8324	20,4176	17	30	17	2
53	79	89,8040	20,4746	62	124	62	2
54	78	89,7164	20,6506	18	63	18	8
55	77	89,6965	20,6905	18	92	18	9
56	76	89,5643	20,9559	72	125	72	2
57	75	89,4898	21,1055	18	70	18	11
58	74	89,4623	21,1608	17	72	17	4
59	73	89,2684	21,5503	2	62	2	16
60	72	89,1917	21,7042	2	14	2	17
61	71	89,1793	21,7292	35	53	35	7
62	70	89,1540	21,7799	12	18	12	12
63	69	89,0683	21,9520	2	31	2	18
64	68	89,0201	22,0487	17	68	17	6
65	67	88,7875	22,5159	6	17	6	26
66	66	88,7102	22,6711	6	39	6	27
67	65	88,5635	22,9657	65	73	65	3
68	64	88,4100	23,2739	12	132	12	13
69	63	88,3645	23,3652	12	78	12	14
70	62	88,3225	23,4498	9	52	9	4
71	61	88,1618	23,7724	4	118	4	3
72	60	88,1458	23,8045	1	61	1	2
73	59	88,1160	23,8644	2	35	2	25
74	58	88,0206	24,0560	15	109	15	2
75	57	87,9978	24,1016	19	77	19	2
76	56	87,9844	24,1285	64	107	64	3
77	55	87,9770	24,1434	27	95	27	2
78	54	87,9723	24,1530	27	59	27	3
79	53	87,9145	24,2690	1	2	1	27
80	52	87,6153	24,8697	6	64	6	30
81	51	87,4656	25,1704	6	50	6	31
82	50	87,4102	25,2817	1	4	1	30
83	49	87,2714	25,5604	27	38	27	4
84	48	87,0188	26,0676	1	27	1	34
85	47	86,8785	26,3493	1	84	1	35
86	46	86,6690	26,7701	1	49	1	36
87	45	86,5719	26,9650	1	6	1	67
88	44	86,5491	27,0108	47	71	47	2
89	43	86,5219	27,0654	1	65	1	70
90	42	86,4903	27,1289	54	129	54	2
91	41	86,3762	27,3581	1	104	1	71
92	40	86,3253	27,4603	55	116	55	2
93	39	86,2675	27,5763	1	9	1	75
94	38	86,2364	27,6388	1	88	1	76
95	37	86,1115	27,8896	1	32	1	78
96	36	86,0757	27,9615	1	55	1	80
97	35	86,0653	27,9824	15	86	15	3
98	34	85,6435	28,8293	12	105	12	15
99	33	85,6433	28,8298	28	47	28	3
100	32	85,5772	28,9626	1	29	1	81
101	31	85,5650	28,9871	1	82	1	82
102	30	85,4577	29,2025	12	131	12	16

103	29	85,2064	29,7070	1	102	1	83
104	28	84,9772	30,1674	12	28	12	19
105	27	84,8635	30,3958	12	21	12	20
106	26	84,6624	30,7996	1	12	1	103
107	25	84,6444	30,8357	1	16	1	104
108	24	84,5628	30,9996	1	19	1	106
109	23	84,4995	31,1266	1	15	1	109
110	22	84,2594	31,6088	1	57	1	110
111	21	84,1180	31,8928	74	94	74	2
112	20	84,0570	32,0152	1	44	1	111
113	19	84,0140	32,1016	42	76	42	2
114	18	83,8436	32,4439	1	42	1	113
115	17	83,7907	32,5499	1	46	1	114
116	16	83,5694	32,9945	1	83	1	115
117	15	83,3531	33,4287	1	74	1	117
118	14	82,9175	34,3035	1	3	1	118
119	13	82,3043	35,5349	1	10	1	119
120	12	82,0845	35,9762	1	24	1	120
121	11	81,9486	36,2492	1	128	1	121
122	10	81,6992	36,7500	1	90	1	122
123	9	81,6629	36,8228	1	20	1	123
124	8	81,5153	37,1193	1	54	1	125
125	7	80,3194	39,5207	1	23	1	126
126	6	79,9905	40,1812	1	91	1	127
127	5	79,4762	41,2139	1	25	1	128
128	4	78,7878	42,5962	1	113	1	129
129	3	78,2054	43,7659	1	130	1	130
130	2	77,6970	44,7868	1	112	1	131
131	1	74,1983	51,8125	1	115	1	132

Final Partition
Number of clusters: 2

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	131	473854	56,8278	109,924
Cluster2	1	0	0,0000	0,000

Cluster Centroids

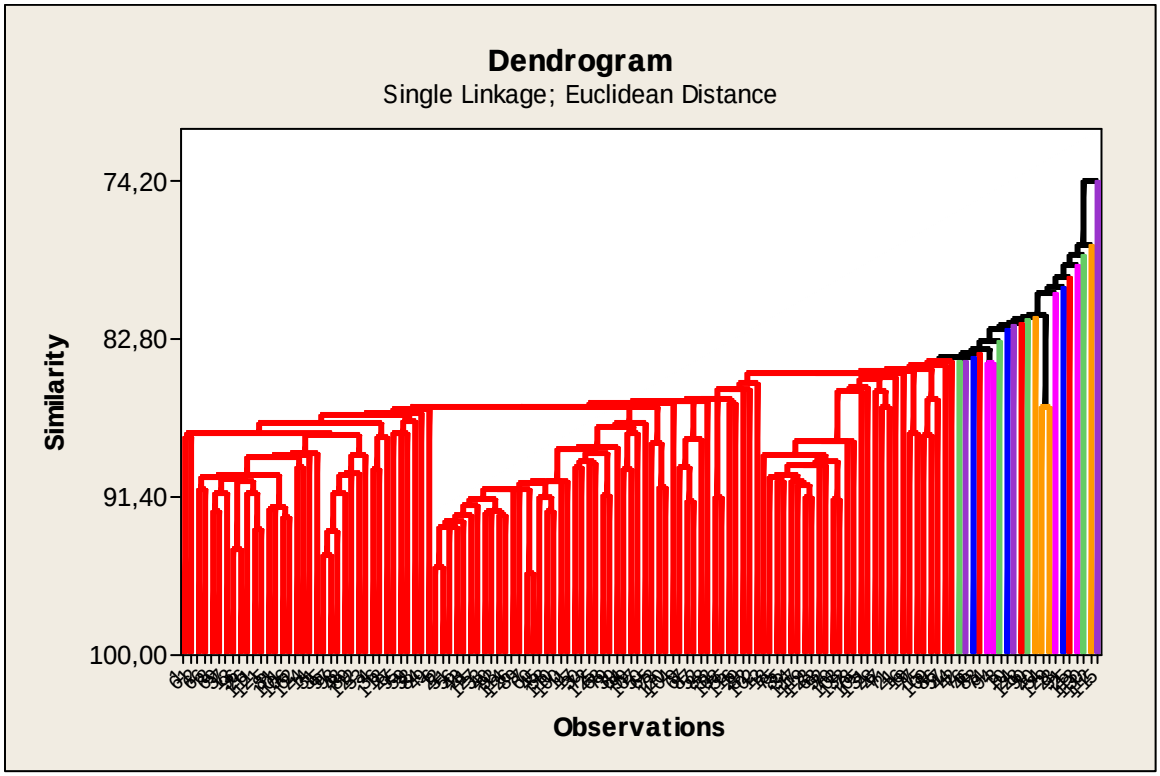
Variable	Cluster1	Cluster2	Grand centroid
GOV-n	44,6250	96,5059	45,0180
INT COP-n	62,1279	62,5000	62,1307
IDH-n	58,5397	31,7957	58,3371
SPI-n	55,5086	99,9101	55,8449
GINNI-n	36,9767	19,6503	36,8454
HPI-n	48,3014	66,9505	48,4427
OHI-n	53,8336	54,1655	53,8361

Distances Between Cluster Centroids

	Cluster1	Cluster2
Cluster1	0,0000	77,6312
Cluster2	77,6312	0,0000

7.1.2 –ANÁLISE COM 20 CLUSTERS

Foi realizada uma análise com 20 clusters que resultou em 1 clusters com 111 observações, 2 clusters com 2 observações cada e 18 com 1 observação cada.



Cluster Analysis of Observations: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; ...

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,6765	8,6821	7	60	7	2
2	130	95,2515	9,5355	6	22	6	2
3	129	94,6384	10,7666	35	97	35	2
4	128	94,3569	11,3318	13	96	13	2
5	127	93,3660	13,3218	35	98	35	3
6	126	93,1904	13,6743	111	121	111	2
7	125	93,1141	13,8276	51	56	51	2
8	124	93,0943	13,8673	6	51	6	4
9	123	92,7531	14,5525	6	43	6	5
10	122	92,5067	15,0474	101	106	101	2
11	121	92,4641	15,1328	6	41	6	6
12	120	92,4009	15,2598	93	114	93	2
13	119	92,2886	15,4854	33	40	33	2
14	118	92,2825	15,4976	33	93	33	4
15	117	92,2356	15,5917	8	69	8	2
16	116	92,1948	15,6736	58	110	58	2
17	115	92,0977	15,8686	5	81	5	2
18	114	92,0119	16,0409	5	101	5	4
19	113	91,9338	16,1978	6	127	6	7
20	112	91,6873	16,6928	67	87	67	2
21	111	91,5682	16,9320	70	103	70	2
22	110	91,5509	16,9666	6	33	6	11
23	109	91,4864	17,0962	119	123	119	2
24	108	91,4720	17,1252	32	108	32	2
25	107	91,4146	17,2404	45	58	45	3

26	106	91,3045	17,4615	68	79	68	2
27	105	91,2764	17,5179	8	37	8	3
28	104	91,2719	17,5270	26	111	26	3
29	103	91,2337	17,6037	35	48	35	4
30	102	91,0489	17,9748	2	66	2	2
31	101	90,9767	18,1197	6	126	6	12
32	100	90,9760	18,1212	6	89	6	13
33	99	90,9754	18,1223	6	7	6	15
34	98	90,9166	18,2404	73	120	73	2
35	97	90,8519	18,3704	117	119	117	3
36	96	90,6858	18,7039	6	45	6	18
37	95	90,6258	18,8243	6	100	6	19
38	94	90,6029	18,8705	75	85	75	2
39	93	90,5769	18,9226	6	11	6	20
40	92	90,5662	18,9441	13	26	13	5
41	91	90,5402	18,9962	34	117	34	4
42	90	90,4519	19,1736	8	13	8	8
43	89	90,4087	19,2603	2	8	2	10
44	88	90,4037	19,2704	18	75	18	3
45	87	90,3061	19,4664	18	34	18	7
46	86	90,2797	19,5194	2	5	2	14
47	85	90,2439	19,5912	35	80	35	5
48	84	90,1891	19,7013	35	122	35	6
49	83	89,9552	20,1711	4	36	4	2
50	82	89,9381	20,2054	64	99	64	2
51	81	89,8427	20,3968	9	67	9	3
52	80	89,8324	20,4176	17	30	17	2
53	79	89,8040	20,4746	62	124	62	2
54	78	89,7164	20,6506	18	63	18	8
55	77	89,6965	20,6905	18	92	18	9
56	76	89,5643	20,9559	72	125	72	2
57	75	89,4898	21,1055	18	70	18	11
58	74	89,4623	21,1608	17	72	17	4
59	73	89,2684	21,5503	2	62	2	16
60	72	89,1917	21,7042	2	14	2	17
61	71	89,1793	21,7292	35	53	35	7
62	70	89,1540	21,7799	12	18	12	12
63	69	89,0683	21,9520	2	31	2	18
64	68	89,0201	22,0487	17	68	17	6
65	67	88,7875	22,5159	6	17	6	26
66	66	88,7102	22,6711	6	39	6	27
67	65	88,5635	22,9657	65	73	65	3
68	64	88,4100	23,2739	12	132	12	13
69	63	88,3645	23,3652	12	78	12	14
70	62	88,3225	23,4498	9	52	9	4
71	61	88,1618	23,7724	4	118	4	3
72	60	88,1458	23,8045	1	61	1	2
73	59	88,1160	23,8644	2	35	2	25
74	58	88,0206	24,0560	15	109	15	2
75	57	87,9978	24,1016	19	77	19	2
76	56	87,9844	24,1285	64	107	64	3
77	55	87,9770	24,1434	27	95	27	2
78	54	87,9723	24,1530	27	59	27	3
79	53	87,9145	24,2690	1	2	1	27
80	52	87,6153	24,8697	6	64	6	30
81	51	87,4656	25,1704	6	50	6	31
82	50	87,4102	25,2817	1	4	1	30
83	49	87,2714	25,5604	27	38	27	4
84	48	87,0188	26,0676	1	27	1	34
85	47	86,8785	26,3493	1	84	1	35
86	46	86,6690	26,7701	1	49	1	36
87	45	86,5719	26,9650	1	6	1	67
88	44	86,5491	27,0108	47	71	47	2
89	43	86,5219	27,0654	1	65	1	70
90	42	86,4903	27,1289	54	129	54	2
91	41	86,3762	27,3581	1	104	1	71
92	40	86,3253	27,4603	55	116	55	2
93	39	86,2675	27,5763	1	9	1	75
94	38	86,2364	27,6388	1	88	1	76
95	37	86,1115	27,8896	1	32	1	78
96	36	86,0757	27,9615	1	55	1	80

97	35	86,0653	27,9824	15	86	15	3
98	34	85,6435	28,8293	12	105	12	15
99	33	85,6433	28,8298	28	47	28	3
100	32	85,5772	28,9626	1	29	1	81
101	31	85,5650	28,9871	1	82	1	82
102	30	85,4577	29,2025	12	131	12	16
103	29	85,2064	29,7070	1	102	1	83
104	28	84,9772	30,1674	12	28	12	19
105	27	84,8635	30,3958	12	21	12	20
106	26	84,6624	30,7996	1	12	1	103
107	25	84,6444	30,8357	1	16	1	104
108	24	84,5628	30,9996	1	19	1	106
109	23	84,4995	31,1266	1	15	1	109
110	22	84,2594	31,6088	1	57	1	110
111	21	84,1180	31,8928	74	94	74	2
112	20	84,0570	32,0152	1	44	1	111
113	19	84,0140	32,1016	42	76	42	2
114	18	83,8436	32,4439	1	42	1	113
115	17	83,7907	32,5499	1	46	1	114
116	16	83,5694	32,9945	1	83	1	115
117	15	83,3531	33,4287	1	74	1	117
118	14	82,9175	34,3035	1	3	1	118
119	13	82,3043	35,5349	1	10	1	119
120	12	82,0845	35,9762	1	24	1	120
121	11	81,9486	36,2492	1	128	1	121
122	10	81,6992	36,7500	1	90	1	122
123	9	81,6629	36,8228	1	20	1	123
124	8	81,5153	37,1193	1	54	1	125
125	7	80,3194	39,5207	1	23	1	126
126	6	79,9905	40,1812	1	91	1	127
127	5	79,4762	41,2139	1	25	1	128
128	4	78,7878	42,5962	1	113	1	129
129	3	78,2054	43,7659	1	130	1	130
130	2	77,6970	44,7868	1	112	1	131
131	1	74,1983	51,8125	1	115	1	132

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	111	362272	54,3046	96,2168
Cluster2	1	0	0,0000	0,0000
Cluster3	1	0	0,0000	0,0000
Cluster4	1	0	0,0000	0,0000
Cluster5	1	0	0,0000	0,0000
Cluster6	1	0	0,0000	0,0000
Cluster7	1	0	0,0000	0,0000
Cluster8	1	0	0,0000	0,0000
Cluster9	1	0	0,0000	0,0000
Cluster10	2	368	13,5645	13,5645
Cluster11	2	509	15,9464	15,9464
Cluster12	1	0	0,0000	0,0000
Cluster13	1	0	0,0000	0,0000
Cluster14	1	0	0,0000	0,0000
Cluster15	1	0	0,0000	0,0000
Cluster16	1	0	0,0000	0,0000
Cluster17	1	0	0,0000	0,0000
Cluster18	1	0	0,0000	0,0000
Cluster19	1	0	0,0000	0,0000
Cluster20	1	0	0,0000	0,0000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7

GOV-n	47,5759	14,8727	21,8412	24,5415	0,9088	9,2017	80,674
INT COP-n	64,9550	37,5000	16,2500	37,5000	46,2500	28,7500	100,000
IDH-n	60,2624	31,1367	73,9703	40,6919	0,6590	5,7661	79,901
SPI-n	58,3549	13,1740	58,5909	34,6693	2,8217	0,0000	78,541
GINNI-n	35,7330	46,5553	4,2797	18,2672	82,1503	48,2255	71,086
HPI-n	48,3634	25,6010	35,7667	42,7841	6,4309	5,0446	75,503
OHI-n	56,3285	9,4095	54,1655	24,6697	54,1655	54,1655	64,303

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
GOV-n	52,5262	28,1691	12,0976	18,5722	57,9103	49,0240
INT COP-n	66,2500	62,5000	1,8750	25,0000	83,7500	83,7500
IDH-n	67,0511	47,9407	60,6260	29,7364	71,8287	74,4646
SPI-n	56,3264	51,7074	43,8444	19,3027	67,2178	61,4666
GINNI-n	43,1367	71,8424	31,0934	38,0350	55,8194	34,8121
HPI-n	56,4148	82,6885	51,2897	67,2986	43,1981	47,9478
OHI-n	81,9628	72,6881	45,2278	22,5802	67,1879	15,0714

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
GOV-n	30,5296	25,8158	0,0000	31,1536	96,5059	68,430
INT COP-n	46,2500	83,7500	3,7500	62,5000	62,5000	100,000
IDH-n	45,6343	0,0000	22,4053	92,4217	31,7957	74,629
SPI-n	53,4328	13,4795	10,5140	29,2416	99,9101	80,715
GINNI-n	54,5668	16,5449	27,3225	69,5981	19,6503	43,058
HPI-n	83,1763	10,2355	36,1505	47,9478	66,9505	40,367
OHI-n	0,0000	31,1405	18,4147	54,1655	54,1655	43,138

Variable	Cluster20	Grand centroid
GOV-n	8,0062	45,0180
INT COP-n	16,2500	62,1307
IDH-n	70,3460	58,3371
SPI-n	56,0388	55,8449
GINNI-n	52,0616	36,8454
HPI-n	82,7132	48,4427
OHI-n	19,4662	53,8361

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,000	86,888	66,144	59,723	114,376	105,810	71,880
Cluster2	86,888	0,000	91,056	44,443	70,322	57,917	144,887
Cluster3	66,144	91,056	0,000	57,039	129,412	106,188	130,589
Cluster4	59,723	44,443	57,039	0,000	97,594	76,977	126,295
Cluster5	114,376	70,322	129,412	97,594	0,000	39,521	162,051
Cluster6	105,810	57,917	106,188	76,977	39,521	0,000	165,648
Cluster7	71,880	144,887	130,589	126,295	162,051	165,648	0,000
Cluster8	29,209	107,803	78,762	82,827	122,997	116,830	63,619
Cluster9	57,692	102,331	100,067	88,031	109,074	113,493	77,634
Cluster10	74,845	72,588	41,676	50,646	109,874	90,413	135,854
Cluster11	79,983	46,842	81,968	39,295	92,354	77,628	135,675
Cluster12	34,935	110,913	93,901	92,193	126,815	126,345	47,546
Cluster13	47,643	90,467	88,953	70,638	134,539	129,064	78,011
Cluster14	74,955	74,969	97,547	63,614	122,663	119,517	108,050
Cluster15	94,491	69,551	115,806	73,142	83,602	71,251	149,081
Cluster16	106,718	44,787	86,098	53,145	86,641	61,479	171,101
Cluster17	57,526	89,093	88,558	83,100	110,356	110,827	85,499
Cluster18	74,473	138,742	111,394	107,682	165,431	155,144	85,142
Cluster19	51,585	120,814	106,400	100,329	147,610	145,731	51,483
Cluster20	82,767	85,245	76,826	69,458	128,713	121,325	123,760

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	29,209	57,692	74,845	79,983	34,935	47,643
Cluster2	107,803	102,331	72,588	46,842	110,913	90,467
Cluster3	78,762	100,067	41,676	81,968	93,901	88,953
Cluster4	82,827	88,031	50,646	39,295	92,193	70,638
Cluster5	122,997	109,074	109,874	92,354	126,815	134,539
Cluster6	116,830	113,493	90,413	77,628	126,345	129,064
Cluster7	63,619	77,634	135,854	135,675	47,546	78,011

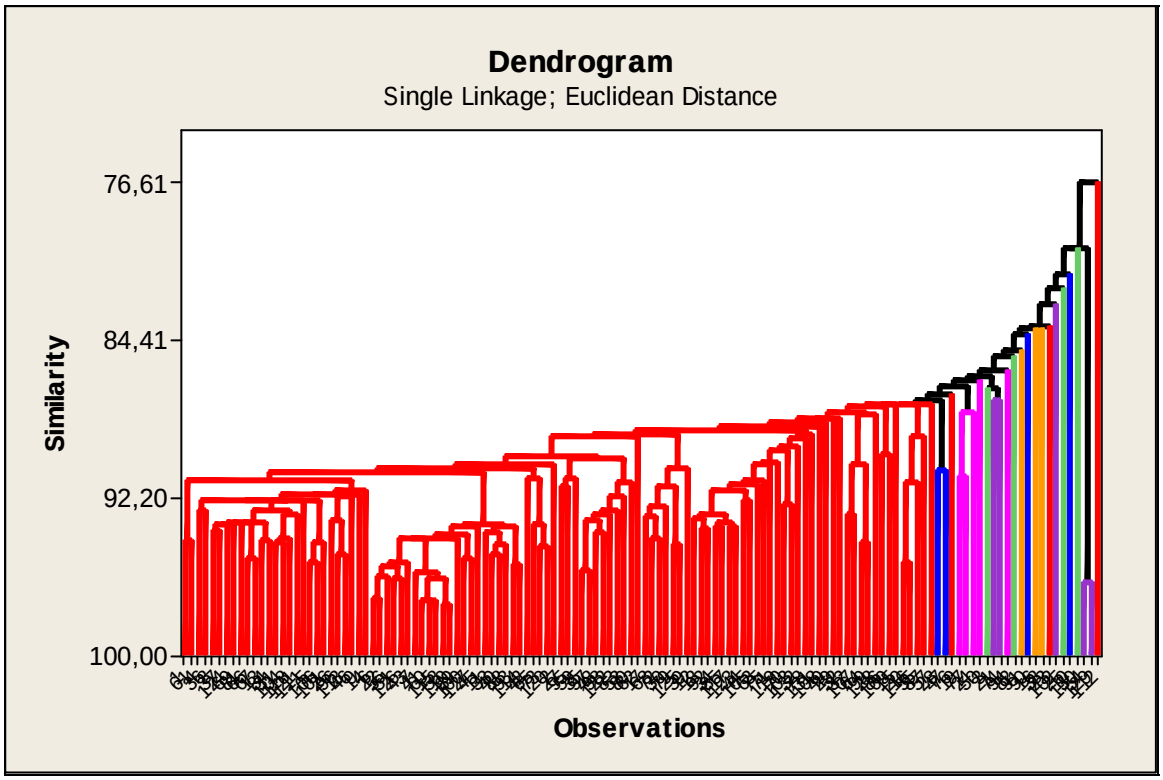
Cluster8	0,000	50,933	86,582	96,376	32,102	70,819
Cluster9	50,933	0,000	86,930	82,282	63,196	87,081
Cluster10	86,582	86,930	0,000	54,319	103,118	97,486
Cluster11	96,376	82,282	54,319	0,000	109,333	92,728
Cluster12	32,102	63,196	103,118	109,333	0,000	57,437
Cluster13	70,819	87,081	97,486	92,728	57,437	0,000
Cluster14	94,423	76,552	78,987	55,203	95,636	67,100
Cluster15	113,022	119,374	116,554	90,628	114,042	101,983
Cluster16	124,324	116,993	60,696	45,053	136,573	119,650
Cluster17	58,247	63,618	82,682	89,225	58,444	69,787
Cluster18	80,784	102,805	123,413	123,764	81,576	90,133
Cluster19	61,718	89,854	120,931	122,764	36,249	43,868
Cluster20	95,760	79,311	50,570	60,257	105,008	88,378

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	74,955	94,491	106,718	57,526	74,473	51,585
Cluster2	74,969	69,551	44,787	89,093	138,742	120,814
Cluster3	97,547	115,806	86,098	88,558	111,394	106,400
Cluster4	63,614	73,142	53,145	83,100	107,682	100,329
Cluster5	122,663	83,602	86,641	110,356	165,431	147,610
Cluster6	119,517	71,251	61,479	110,827	155,144	145,731
Cluster7	108,050	149,081	171,101	85,499	85,142	51,483
Cluster8	94,423	113,022	124,324	58,247	80,784	61,718
Cluster9	76,552	119,374	116,993	63,618	102,805	89,854
Cluster10	78,987	116,554	60,696	82,682	123,413	120,931
Cluster11	55,203	90,628	45,053	89,225	123,764	122,764
Cluster12	95,636	114,042	136,573	58,444	81,576	36,249
Cluster13	67,100	101,983	119,650	69,787	90,133	43,868
Cluster14	0,000	113,327	91,721	86,253	106,700	98,674
Cluster15	113,327	0,000	92,341	118,479	132,990	118,004
Cluster16	91,721	92,341	0,000	113,489	152,087	149,930
Cluster17	86,253	118,479	113,489	0,000	125,684	81,519
Cluster18	106,700	132,990	152,087	125,684	0,000	75,986
Cluster19	98,674	118,004	149,930	81,519	75,986	0,000
Cluster20	49,094	135,218	85,863	81,244	126,036	117,163

	Cluster20
Cluster1	82,767
Cluster2	85,245
Cluster3	76,826
Cluster4	69,458
Cluster5	128,713
Cluster6	121,325
Cluster7	123,760
Cluster8	95,760
Cluster9	79,311
Cluster10	50,570
Cluster11	60,257
Cluster12	105,008
Cluster13	88,378
Cluster14	49,094
Cluster15	135,218
Cluster16	85,863
Cluster17	81,244
Cluster18	126,036
Cluster19	117,163
Cluster20	0,000

7.1.3 –ANÁLISE COM 20 CLUSTERS COM PC1, PC2, PC3 e PC4

Uma outra análise é realizar o cluster com os 4 PC´s e resultou em 1 cluster com 108 observações, 1 com 3 observações, 3 com 2 observações cada e 15 com 1 observações cada.



Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	131	97,5587	0,28392	58 110	58	2
2	130	97,3575	0,30732	60 115	60	2
3	129	97,2948	0,31462	41 60	41	3
4	128	97,2119	0,32425	6 22	6	2
5	127	96,4049	0,41811	54 129	54	2
6	126	96,2316	0,43826	41 58	41	5
7	125	96,1845	0,44374	51 126	51	2
8	124	96,1733	0,44505	6 56	6	3
9	123	95,8883	0,47819	7 41	7	6
10	122	95,8696	0,48036	35 97	35	2
11	121	95,6056	0,51107	6 51	6	5
12	120	95,5527	0,51722	93 114	93	2
13	119	95,4715	0,52666	20 116	20	2
14	118	95,4306	0,53142	5 106	5	2
15	117	95,4141	0,53334	6 43	6	6
16	116	95,2613	0,55110	66 67	66	2
17	115	95,2312	0,55461	100 127	100	2
18	114	95,0610	0,57440	33 40	33	2
19	113	95,0107	0,58025	96 113	96	2

20	112	94,6594	0,62111	73	120	73	2
21	111	94,5902	0,62916	33	88	33	3
22	110	94,5890	0,62930	79	125	79	2
23	109	94,4936	0,64039	5	101	5	3
24	108	94,4759	0,64245	64	99	64	2
25	107	94,4418	0,64641	31	104	31	2
26	106	94,3491	0,65720	13	84	13	2
27	105	94,3195	0,66064	1	61	1	2
28	104	94,2815	0,66506	68	72	68	2
29	103	94,2727	0,66608	6	7	6	12
30	102	94,2639	0,66710	6	89	6	13
31	101	94,2221	0,67197	31	118	31	3
32	100	94,0426	0,69285	6	100	6	15
33	99	93,9328	0,70561	11	33	11	4
34	98	93,9198	0,70713	48	98	48	2
35	97	93,8689	0,71304	8	37	8	2
36	96	93,7638	0,72527	18	92	18	2
37	95	93,6832	0,73464	34	85	34	2
38	94	93,6632	0,73697	6	45	6	16
39	93	93,6424	0,73938	117	123	117	2
40	92	93,6330	0,74048	11	93	11	6
41	91	93,5649	0,74840	6	11	6	22
42	90	93,5510	0,75002	65	73	65	3
43	89	93,5505	0,75008	8	124	8	3
44	88	93,5491	0,75024	8	69	8	4
45	87	93,4825	0,75798	8	81	8	5
46	86	93,4639	0,76014	8	66	8	7
47	85	93,4315	0,76391	34	117	34	4
48	84	93,4273	0,76440	8	13	8	9
49	83	93,4005	0,76752	26	96	26	3
50	82	93,3759	0,77038	35	48	35	4
51	81	93,2813	0,78138	12	18	12	3
52	80	93,1508	0,79656	17	68	17	3
53	79	93,0743	0,80546	35	122	35	5
54	78	93,0663	0,80639	12	34	12	7
55	77	93,0517	0,80808	31	121	31	4
56	76	93,0376	0,80972	32	107	32	2
57	75	92,8883	0,82708	8	31	8	13
58	74	92,8687	0,82936	4	36	4	2
59	73	92,8121	0,83595	35	83	35	6
60	72	92,6541	0,85432	17	30	17	4
61	71	92,5849	0,86238	70	103	70	2
62	70	92,4085	0,88289	8	111	8	14
63	69	92,3804	0,88615	4	8	4	16
64	68	92,3682	0,88757	44	105	44	2
65	67	92,3472	0,89001	4	5	4	19
66	66	92,1972	0,90747	35	53	35	7
67	65	92,1949	0,90773	26	46	26	4
68	64	92,0691	0,92236	4	26	4	23
69	63	92,0619	0,92319	4	9	4	24
70	62	91,8986	0,94219	4	14	4	25
71	61	91,8837	0,94392	12	44	12	9
72	60	91,7106	0,96405	17	79	17	6
73	59	91,6785	0,96778	27	95	27	2
74	58	91,5850	0,97866	12	63	12	10
75	57	91,5036	0,98813	35	80	35	8
76	56	91,4253	0,99724	20	55	20	3
77	55	91,3510	1,00588	1	4	1	27
78	54	91,3341	1,00784	12	75	12	11
79	53	91,2800	1,01413	42	65	42	4
80	52	91,2380	1,01901	27	59	27	3
81	51	91,2177	1,02138	19	77	19	2
82	50	90,9749	1,04961	1	6	1	49
83	49	90,9195	1,05605	28	47	28	2
84	48	90,8130	1,06845	17	39	17	7
85	47	90,7558	1,07510	1	42	1	53
86	46	90,5827	1,09523	1	2	1	54
87	45	90,5800	1,09555	12	71	12	12
88	44	90,5600	1,09787	32	64	32	4
89	43	90,4777	1,10744	12	119	12	13
90	42	90,3052	1,12750	27	35	27	11

91	41	90,2050	1,13915	1	27	1	65
92	40	90,0643	1,15551	86	109	86	2
93	39	89,8505	1,18039	12	70	12	15
94	38	89,6528	1,20338	12	132	12	16
95	37	89,3375	1,24004	12	78	12	17
96	36	89,2179	1,25395	1	62	1	66
97	35	89,1972	1,25636	20	87	20	4
98	34	89,1030	1,26732	12	131	12	18
99	33	89,0862	1,26927	1	17	1	73
100	32	88,8692	1,29451	1	12	1	91
101	31	88,7024	1,31391	1	108	1	92
102	30	88,5173	1,33543	1	49	1	93
103	29	88,3398	1,35607	1	29	1	94
104	28	88,2522	1,36626	1	52	1	95
105	27	88,0642	1,38813	32	128	32	5
106	26	88,0054	1,39497	1	32	1	100
107	25	87,9729	1,39875	19	24	19	3
108	24	87,7480	1,42490	1	86	1	102
109	23	87,6481	1,43652	1	15	1	103
110	22	87,6261	1,43908	1	20	1	107
111	21	87,6219	1,43957	1	57	1	108
112	20	87,4384	1,46091	21	74	21	2
113	19	87,4202	1,46302	1	28	1	110
114	18	87,1207	1,49785	1	76	1	111
115	17	86,8483	1,52953	3	21	3	3
116	16	86,6789	1,54924	1	19	1	114
117	15	86,3881	1,58305	1	50	1	115
118	14	86,1929	1,60576	1	3	1	118
119	13	85,9217	1,63731	1	94	1	119
120	12	85,1772	1,72388	1	82	1	120
121	11	84,9076	1,75523	1	91	1	121
122	10	84,1266	1,84606	1	90	1	122
123	9	83,8395	1,87946	16	38	16	2
124	8	83,8152	1,88229	1	16	1	124
125	7	83,7773	1,88670	1	23	1	125
126	6	82,6428	2,01864	1	102	1	126
127	5	81,8506	2,11077	1	25	1	127
128	4	81,1943	2,18710	1	10	1	128
129	3	79,8909	2,33868	1	130	1	129
130	2	79,8509	2,34334	1	54	1	131
131	1	76,6147	2,71969	1	112	1	132

Final Partition
Number of clusters: 20

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	108	866,449	2,69833	4,45327
Cluster2	1	0,000	0,00000	0,00000
Cluster3	1	0,000	0,00000	0,00000
Cluster4	1	0,000	0,00000	0,00000
Cluster5	3	2,454	0,85114	1,13504
Cluster6	2	1,067	0,73046	0,73046
Cluster7	1	0,000	0,00000	0,00000
Cluster8	1	0,000	0,00000	0,00000
Cluster9	2	0,558	0,52803	0,52803
Cluster10	1	0,000	0,00000	0,00000
Cluster11	1	0,000	0,00000	0,00000
Cluster12	2	0,087	0,20906	0,20906
Cluster13	1	0,000	0,00000	0,00000
Cluster14	1	0,000	0,00000	0,00000
Cluster15	1	0,000	0,00000	0,00000
Cluster16	1	0,000	0,00000	0,00000
Cluster17	1	0,000	0,00000	0,00000
Cluster18	1	0,000	0,00000	0,00000
Cluster19	1	0,000	0,00000	0,00000
Cluster20	1	0,000	0,00000	0,00000

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
PC1	0,531356	-4,47375	-2,47931	2,01264	-4,02893	-3,99894	-3,96736
PC2	0,078486	0,15078	-4,18707	3,18193	-0,17042	-1,11968	1,26316
PC3	-0,032403	0,19714	0,62836	-1,71820	2,12571	0,31086	1,58351
PC4	-0,007305	-0,10877	-0,37448	0,40932	0,00191	0,43716	1,92409

Variable	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
PC1	3,79097	-3,29346	1,18203	0,79281	-4,59077	2,01832
PC2	2,16823	1,49810	2,95196	-2,17609	-4,16929	1,28189
PC3	-1,54436	0,82137	-2,13921	0,65883	-0,30342	0,01021
PC4	0,97342	-0,68015	-1,20695	-0,29887	0,55106	0,60053

Variable	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
PC1	0,52070	-2,22606	-1,36040	-4,71854	-2,24926	-6,61136
PC2	1,53794	-0,34009	3,14568	-1,38010	-2,47450	-2,54456
PC3	1,57837	-2,57453	1,91509	-1,56473	0,64761	0,34139
PC4	-0,45819	-0,14853	-1,82312	-0,58317	0,96612	-0,15828

Variable	Cluster20	Grand centroid
PC1	-3,76478	-0,0000000
PC2	-2,96259	0,0000000
PC3	-2,55464	0,0000000
PC4	0,45395	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7
Cluster1	0,00000	5,01192	5,27547	3,8524	5,05131	4,71959	5,28993
Cluster2	5,01192	0,00000	4,80118	7,4295	2,00816	1,46646	2,74743
Cluster3	5,27547	4,80118	0,00000	8,9778	4,57368	3,53238	6,17374
Cluster4	3,85241	7,42954	8,97778	0,0000	7,91710	7,66556	7,25519
Cluster5	5,05131	2,00816	4,57368	7,9171	0,00000	2,09407	2,45921
Cluster6	4,71959	1,46646	3,53238	7,6656	2,09407	0,00000	3,08375
Cluster7	5,28993	2,74743	6,17374	7,2552	2,45921	3,08375	0,00000
Cluster8	4,27083	8,75099	9,28676	2,1304	9,00186	8,67309	8,46748
Cluster9	4,22210	1,98104	5,75454	6,2150	2,34337	2,97649	2,80576
Cluster10	3,81550	6,81902	8,52782	1,8795	7,52026	7,21988	7,28219
Cluster11	2,39046	5,77931	3,84155	6,0289	5,43269	4,97391	6,34699
Cluster12	6,68319	4,40030	2,48664	9,9833	4,74435	3,16870	5,94524
Cluster13	2,00761	6,63057	7,17433	2,5757	6,59637	6,48784	6,32896
Cluster14	2,21991	5,37574	6,53340	4,0680	4,91211	5,46793	5,08856
Cluster15	3,77636	3,60234	5,01727	5,6050	5,03925	3,52412	5,21417
Cluster16	4,48066	4,95507	7,66652	5,4372	4,63602	5,73041	4,94890
Cluster17	5,68933	2,39432	4,21265	8,1933	3,98755	2,26816	4,87328
Cluster18	3,95721	3,63302	2,18710	7,4877	3,40449	2,30003	4,32616
Cluster19	7,61980	3,44348	4,46103	10,5703	3,93891	3,03488	5,23150
Cluster20	5,85485	4,25260	3,73758	8,4756	5,47503	3,41504	6,09783

	Cluster8	Cluster9	Cluster10	Cluster11	Cluster12	Cluster13
Cluster1	4,2708	4,22210	3,81550	2,39046	6,6832	2,00761
Cluster2	8,7510	1,98104	6,81902	5,77931	4,4003	6,63057
Cluster3	9,2868	5,75454	8,52782	3,84155	2,4866	7,17433
Cluster4	2,1304	6,21501	1,87946	6,02895	9,9833	2,57569
Cluster5	9,0019	2,34337	7,52026	5,43269	4,7443	6,59637
Cluster6	8,6731	2,97649	7,21988	4,97391	3,1687	6,48784
Cluster7	8,4675	2,80576	7,28219	6,34699	5,9452	6,32896
Cluster8	0,0000	7,67914	3,53959	5,85960	10,5894	2,54630
Cluster9	7,6791	0,00000	5,58447	5,51082	6,0484	5,52809
Cluster10	3,5396	5,58447	0,00000	5,92470	9,5131	3,37276
Cluster11	5,8596	5,51082	5,92470	0,00000	5,8825	3,83264
Cluster12	10,5894	6,04842	9,51305	5,88253	0,0000	8,57299
Cluster13	2,5463	5,52809	3,37276	3,83264	8,5730	0,00000
Cluster14	4,7847	3,89509	4,10096	3,83914	7,9536	2,42660
Cluster15	6,6945	4,04142	4,87467	4,79184	5,0894	5,28087
Cluster16	6,8761	2,99229	4,82892	6,07117	8,6315	4,93883
Cluster17	9,3502	4,00222	7,36903	6,00281	3,2670	7,50687

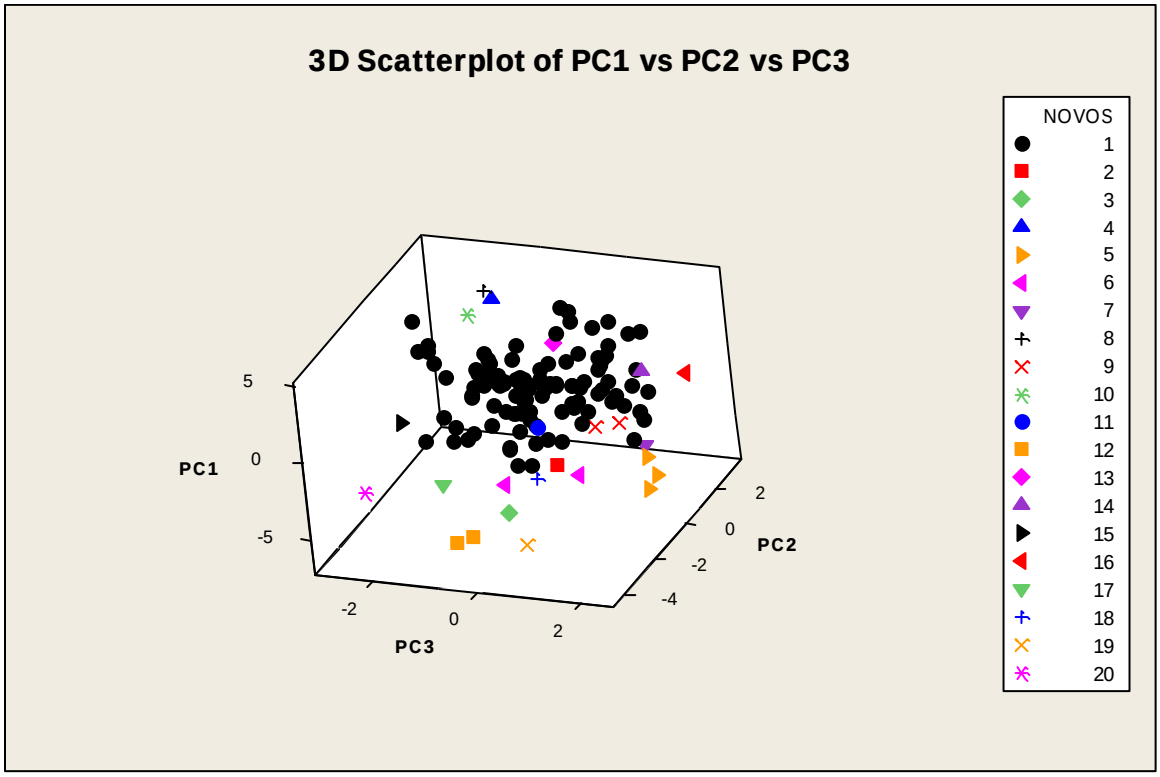
Cluster18	7,9274	4,42858	7,32863	3,30811	3,0711	5,73259
Cluster19	11,6300	5,27773	9,90969	7,42147	2,7643	9,47621
Cluster20	9,2035	5,72743	7,89835	5,68185	2,6862	7,61970

	Cluster14	Cluster15	Cluster16	Cluster17	Cluster18	Cluster19
Cluster1	2,21991	3,77636	4,48066	5,68933	3,95721	7,6198
Cluster2	5,37574	3,60234	4,95507	2,39432	3,63302	3,4435
Cluster3	6,53340	5,01727	7,66652	4,21265	2,18710	4,4610
Cluster4	4,06799	5,60497	5,43722	8,19326	7,48773	10,5703
Cluster5	4,91211	5,03925	4,63602	3,98755	3,40449	3,9389
Cluster6	5,46793	3,52412	5,73041	2,26816	2,30003	3,0349
Cluster7	5,08856	5,21417	4,94890	4,87328	4,32616	5,2315
Cluster8	4,78465	6,69450	6,87608	9,35018	7,92743	11,6300
Cluster9	3,89509	4,04142	2,99229	4,00222	4,42858	5,2777
Cluster10	4,10096	4,87467	4,82892	7,36903	7,32863	9,9097
Cluster11	3,83914	4,79184	6,07117	6,00281	3,30811	7,4215
Cluster12	7,95357	5,08940	8,63149	3,26700	3,07111	2,7643
Cluster13	2,42660	5,28087	4,93883	7,50687	5,73259	9,4762
Cluster14	0,00000	5,33049	2,84601	6,77195	5,16404	8,3158
Cluster15	5,33049	0,00000	5,98840	2,91594	4,02255	5,7091
Cluster16	2,84601	5,98840	0,00000	6,73843	6,46243	8,0746
Cluster17	6,77195	2,91594	6,73843	0,00000	3,81967	2,9585
Cluster18	5,16404	4,02255	6,46243	3,81967	0,00000	4,5156
Cluster19	8,31584	5,70906	8,07464	2,95848	4,51563	0,0000
Cluster20	7,51890	3,09976	8,26169	2,33868	3,61272	4,1279

	Cluster20
Cluster1	5,85485
Cluster2	4,25260
Cluster3	3,73758
Cluster4	8,47557
Cluster5	5,47503
Cluster6	3,41504
Cluster7	6,09783
Cluster8	9,20354
Cluster9	5,72743
Cluster10	7,89835
Cluster11	5,68185
Cluster12	2,68623
Cluster13	7,61970
Cluster14	7,51890
Cluster15	3,09976
Cluster16	8,26169
Cluster17	2,33868
Cluster18	3,61272
Cluster19	4,12790
Cluster20	0,00000

7.2. GRÁFICO 3D DO EXERCICIO COM 20 CLUSTERS – COMPONENTES PRINCIPAIS

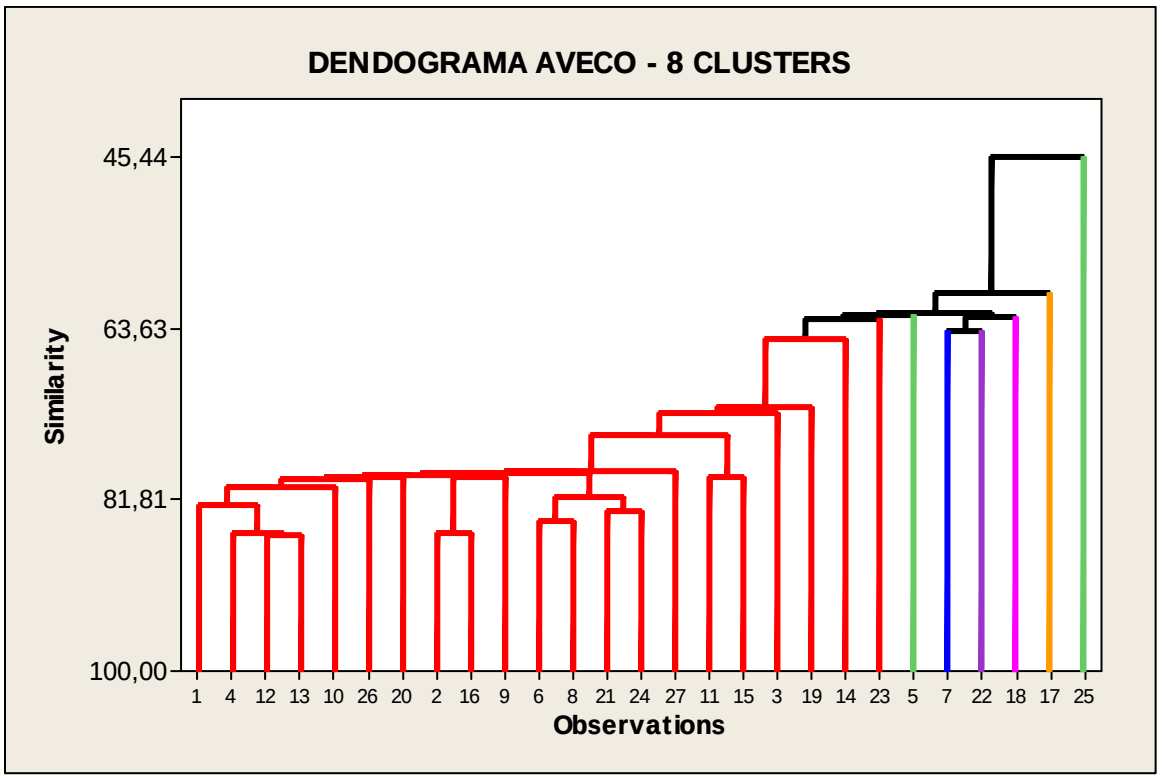
Observando o gráfico abaixo a concentração para estar no PC3. O país mais ao extremo esquerdo trata-se da Venezuela. Os dois países mais ao alto são Chile e Brasil, respectivamente. O país que está a extrema direita é a Niger. E o país mais ao sul é o Sudão.



7.3 AGRUPAMENTO POR REGIÕES AIBER E AVECO

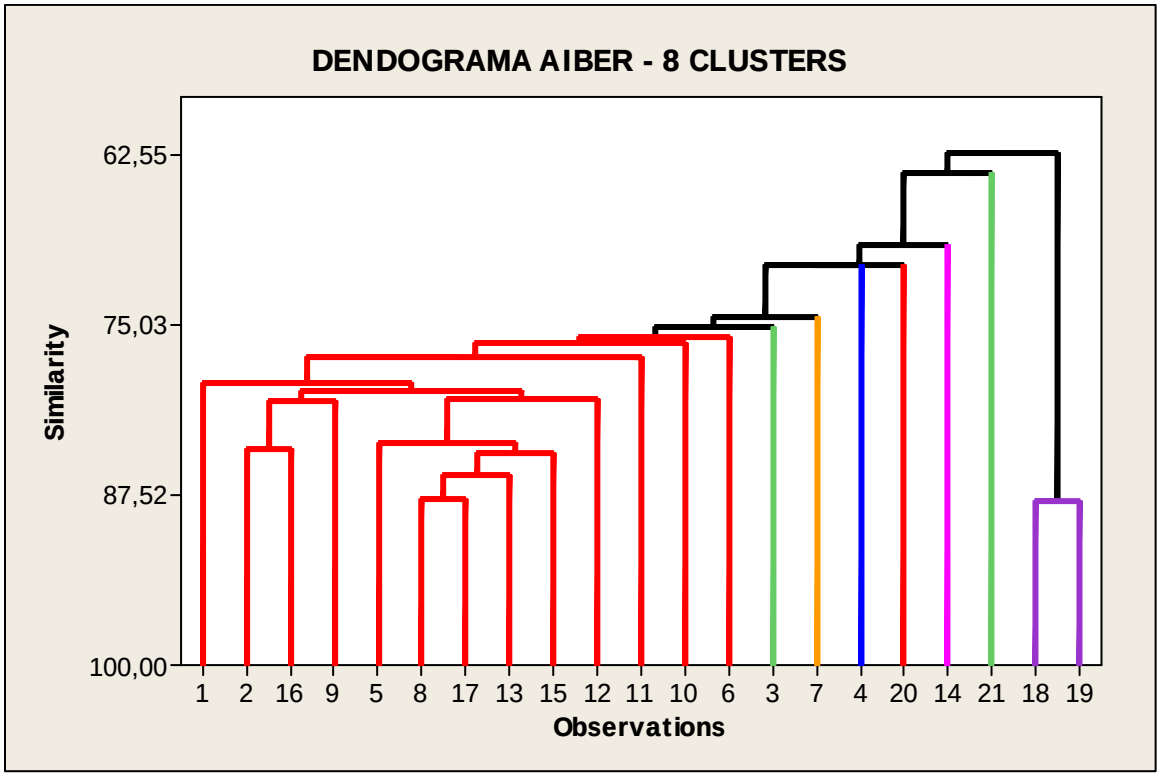
Primeiramente realiza-se agrupamento das observações em 20 clusters com as oito variáveis principais. O resultado é bastante fragmentado.

Analisando AVECO com 8 clusters formado temo o gráfico abaixo.



Análise:

- Cluster 1: Similarity = 79,02
- Cluster 2: Similarity = 62,24
- Cluster 3: Similarity = 64,07
- Cluster 4: Similarity = 59,95
- Cluster 5: Similarity = 62,49
- Cluster 6: Similarity = 64,07
- Cluster 7: Similarity = 62,61
- Cluster 8: Similarity = 45,44



Análise:

- **Cluster 1: Similarity = 75,92**
- **Cluster 2: Similarity = 75,29**
- **Cluster 3: Similarity = 70,77**
- **Cluster 4: Similarity = 74,53**
- **Cluster 5: Similarity = 69,17**
- **Cluster 6: Similarity = 88,04**
- **Cluster 7: Similarity = 70,69**
- **Cluster 8: Similarity = 64,03**

7.4. ANÁLISE DE CONGLOMERADOS - CONSIDERAÇÕES FINAIS

Análise de conglomerados teve como objetivo através do agrupamento das componentes principais, criar outra variáveis que correspondem e explicassem um grupo de variáveis, para facilitar assim as analyses.

8. ANÁLISE DISCRIMINANTE

8.1. ANÁLISES DAS COMPONENTES PRINCIPAIS

Inicialmente, para análises de componentes principais, com base de todas as variáveis (12) pré-selecionadas, foram criadas 4 novas variáveis: PC1, PC2, PC3 e PC4, para analisar e decidir quais delas podem melhor representar as demais 12 variáveis.

Tabela 1 – Componentes Principais – 12 Variáveis

Variable	PC1	PC2	PC3	PC4
SPI-n	0,361	-0,219	-0,179	-0,196
IDH-n	0,301	-0,279	-0,229	-0,322
GOV-n	0,376	-0,124	-0,112	0,097
INSTIT-n	0,361	-0,164	-0,058	0,164
INT COP-n	0,286	0,430	0,119	-0,046
EFFEC-n	0,283	0,386	0,066	0,038
CRED-n	0,282	0,401	0,092	0,019
REG COP-n	0,282	0,339	0,148	-0,179
EPI-n	0,333	-0,270	-0,128	-0,222
OHI-n	0,256	-0,197	0,119	0,833
GINNI-n	-0,097	0,332	-0,906	0,208

Análise:

- O PC1 contém apenas uma variável com contribuição negativa (GINNI -0,097), as demais variam entre 0,256 e 0,361.
- O PC2 contém 6 variáveis com contribuição negativa (GOV -0,124, INSTIT -0,164, OHI -0,197, SPI -0,219, EPI -0,270 e IDH -0,279) e as demais variam entre 0,332 e 0,430.
- O PC3 contém 6 variáveis com contribuição negativa (INSTIT -0,058, GOV -0,112, EPI -0,128, SPI -0,179, ID -0,229 e GINNI -0,906) e as demais variam entre 0,066 e 0,119
- O PC4 contém 5 variáveis com contribuição negativa (INT COP -0,046, REG COOP -0,179, SPI -0,196, EPI -0,22 e IDH -0,322) e as demais variam entre 0,038 e 0,833.

Tabela 2 – Matriz de Correlação – 12 Variáveis

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalue	6,0474	2,6514	1,2406	0,7203	0,4251	0,3006	0,2597	0,1534
Proportion	0,504	0,221	0,103	0,060	0,035	0,025	0,022	0,013
Cumulative	0,504	0,725	0,828	0,888	0,924	0,949	0,970	0,983

	PC9	PC10	PC11	PC12
Eigenvalue	0,1180	0,0520	0,0291	0,0023
Proportion	0,010	0,004	0,002	0,000
Cumulative	0,993	0,997	1,000	1,000

Análise:

Segundo Las Casas e Guevara (2010)[2], por convenção só se deve considerar componentes com contribuição >= 1, pois só acima ou igual a l é que ajuda a explicar o fenômeno. Neste caso, só PC1, PC2 e PC3 têm Eigenvalue que atendem a esse critério. Se se inclui o PC4 por uma questão de ampliar o espectro de análise, apesar de se perceber que não se altera significativamente (todos os cálculos foram feitos dessa forma anteriormente), também tais agrupamentos explicam, cumulativamente 82,8% das variáveis. Se se observa o peso de explicação de PC1 é de 50,4%. Quando se passa para o grupo de variáveis de PC2, há um acréscimo de 22,2% de poder de explicação; de PC2 para PC3, há um acúmulo acrescido de apenas 10,3% de poder de explicação; de PC3 para PC4, há um acúmulo acrescido de somente 6,0% não sendo relevante com o qual, essas 3 primeiras variáveis são suficientes para explicar os indicadores em questão de Integração e Alianças Estratégicas pesquisados para os 12 indicadores pré-selecionados. Para comprovar que os PC's sào complementares, calculamos a correlação entre eles. Ver tabela 3.

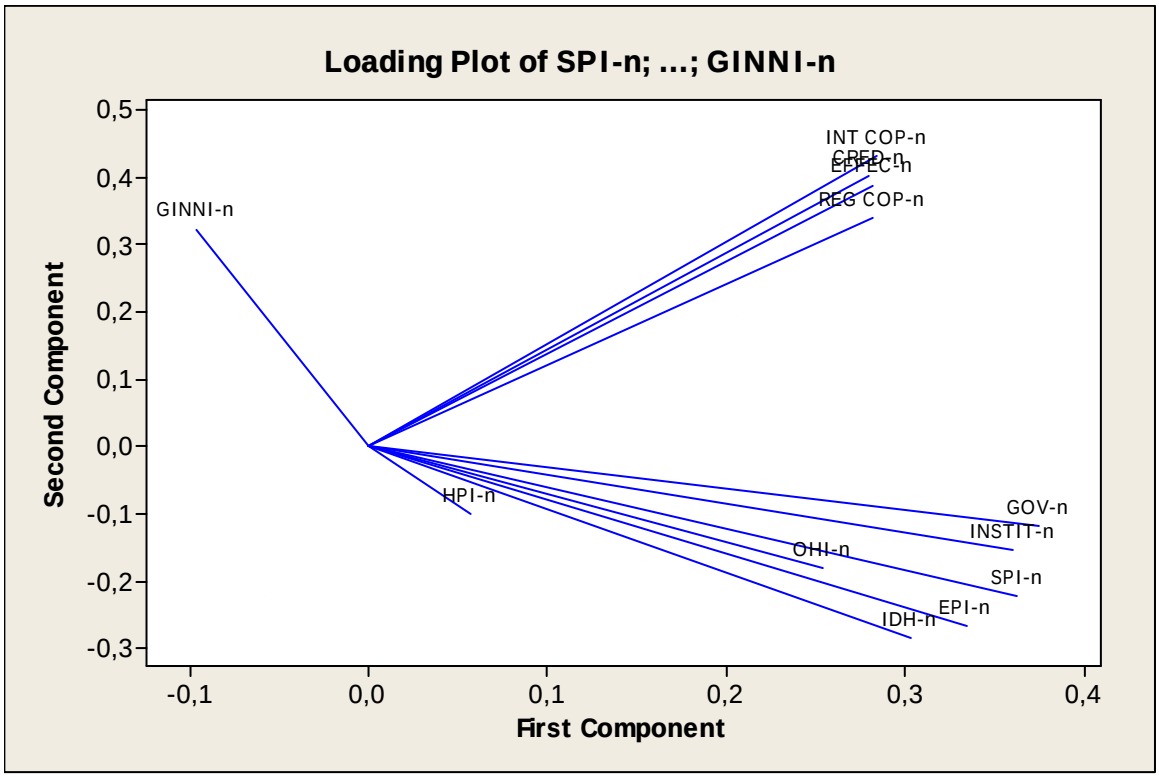
Tabela 3 – Correlação PC1; PC2; PC3; PC4

As correlações abaixo apresentaram resultados não significativos (p > 0.05). E os coeficientes de correlação de Pearson com valores próximos a 0 (zero) indica uma relação fraca ou inexistente entre as duas variáveis.

Correlations: PC1; PC2; PC3; PC4

	PC1	PC2	PC3
PC2	0,000		
	1,000		
PC3	-0,000	-0,000	
	1,000	1,000	
PC4	-0,000	0,000	-0,000
	1,000	1,000	1,000

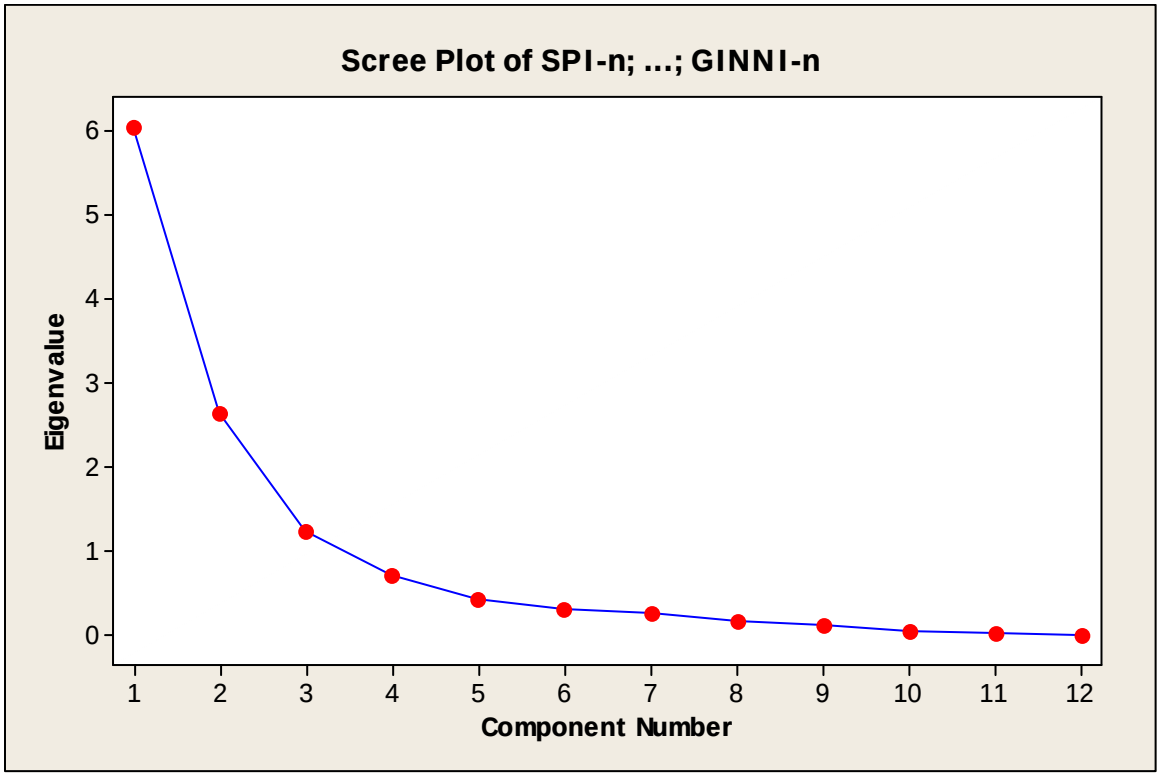
Gráfico 1 – Análise das 12 variáveis



Análise no gráfico 2, temos as seguintes análises:

- A primeira componente complementa a segunda componente e vice-versa, mas não se correlacionam.
- A variável GINNI está concentrada no lado esquerdo do gráfico e possui **Project line: First Component = -0,096116; Second Component = 0,322388**
- No lado direito do gráfico vemos 2 agrupamentos diferentes, podendo sugerir uma afinidade entre elas sendo o primeiro grupo com peso maior nas duas components e nenhum negativo (INT COP, CRED, EFFEC e REG COP) e abaixo com peso negative na primeira componente (GOV, INSTIT, OHI, SPI, EPI, IDH e HPI)
- A variável HPI tem sua reta bem menor que as demais e possui a seguinte projeção: **Project line: First Component = 0,056932; Second Component = -0,101946**

Gráfico 2 – Scree Plot das 12 variáveis



Análise:

- As 4 primeiras bolas vermelhas, da esquerda para a direita, são as que realmente importam: Social Progress é a principal, Com as 4 primeiras possui 88,8% de explicação. – ver a linha amarela - são os números no "Cumulative" na tabela 2.

8.2.1 Análise dos clusters

A análise de clusters nos ajuda a compreender a similariedade dos agrupamentos das variáveis.

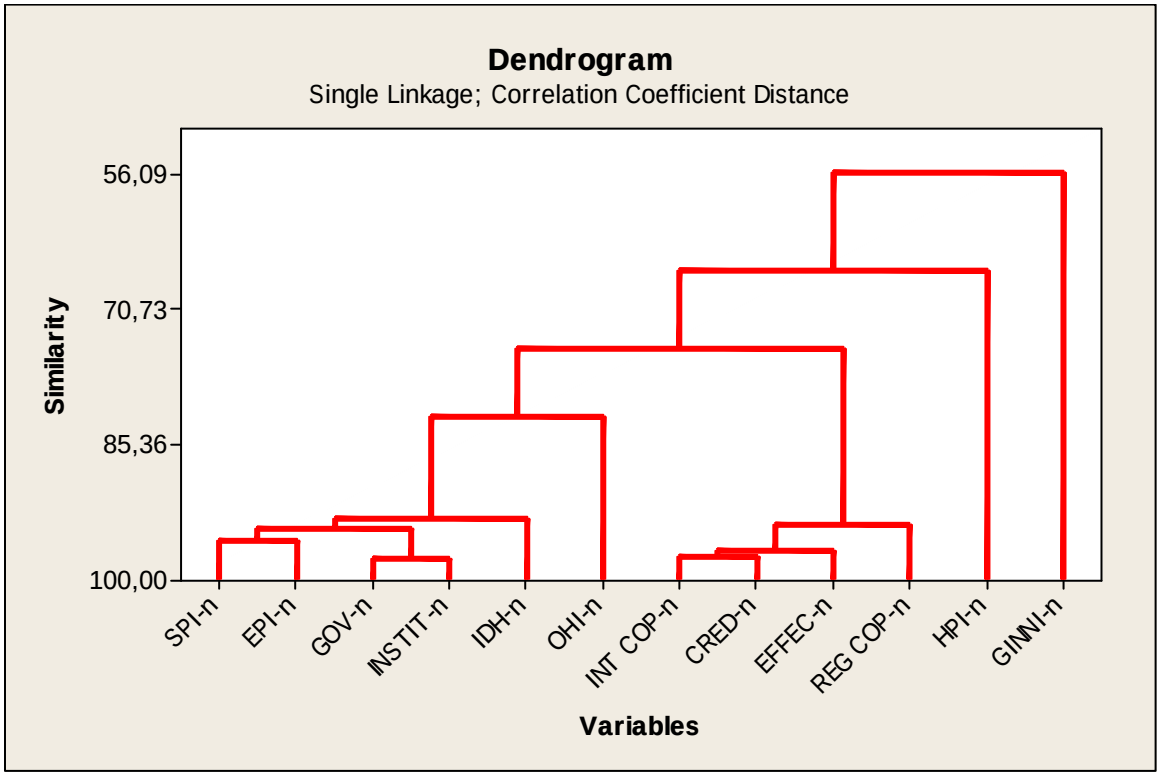
Cluster Analysis of Variables: SPI-n; IDH-n; GOV-n; INSTIT-n; INT COP-n; ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	11	97,8387	0,043227	3	4	3	2
2	10	97,6876	0,046247	5	7	5	2
3	9	96,8707	0,062585	5	6	5	3
4	8	95,8701	0,082598	1	9	1	2
5	7	94,4907	0,110186	1	3	1	4
6	6	94,1283	0,117433	5	8	5	4
7	5	93,5636	0,128728	1	2	1	5
8	4	82,3853	0,352293	1	11	1	6
9	3	74,9859	0,500283	1	5	1	10
10	2	66,6038	0,667924	1	10	1	11
11	1	56,0883	0,878234	1	12	1	12

Na tabela acima demonstra que são necessários 11 agrupamentos para se ter 97,84% de similariedade 3 agrupamentos para ter 74,9% de similariedade. O dendograma abaixo apresenta esta similariedades e ainda destaca as variáveis que estão fora destes agrupamentos.

Gráfico 3 – Dendograma – 12 Variáveis



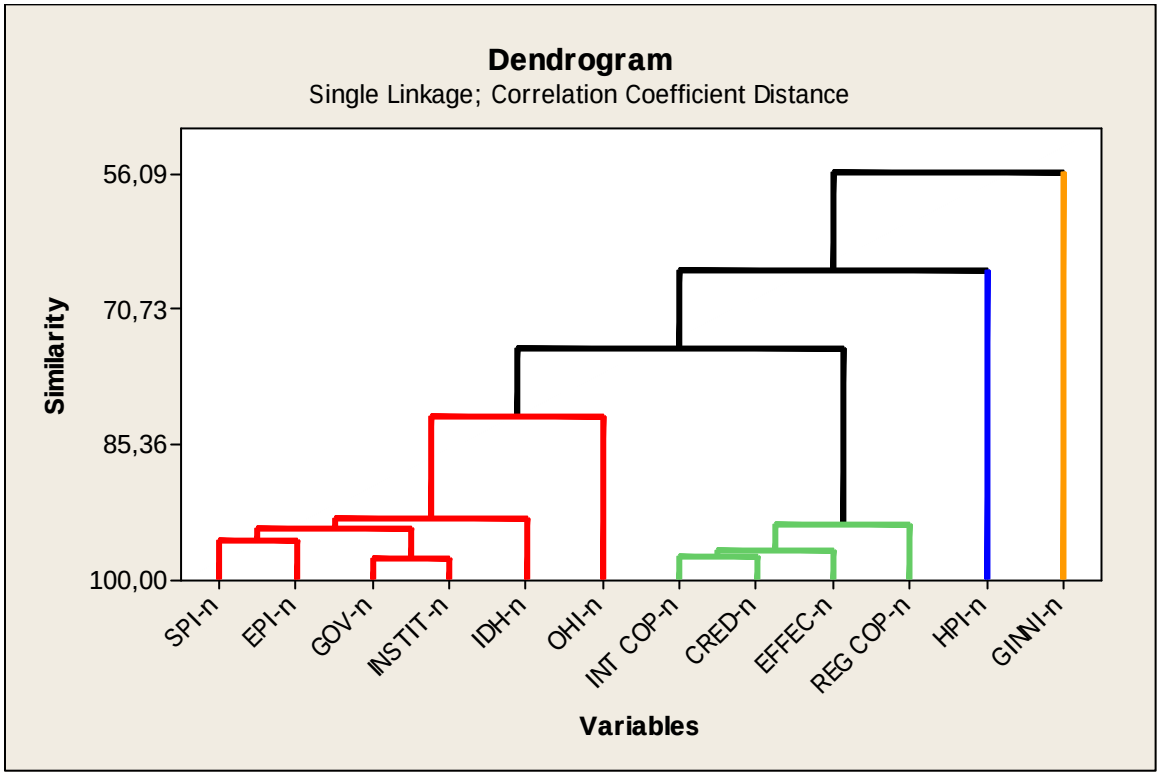
8.2.3. Reagrupando em 4 clusters

Cluster Analysis of Variables: SPI-n; IDH-n; GOV-n; INSTIT-n; INT COP-n; ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	11	97,8387	0,043227	3	4	3	2
2	10	97,6876	0,046247	5	7	5	2
3	9	96,8707	0,062585	5	6	5	3
4	8	95,8701	0,082598	1	9	1	2
5	7	94,4907	0,110186	1	3	1	4
6	6	94,1283	0,117433	5	8	5	4
7	5	93,5636	0,128728	1	2	1	5
8	4	82,3853	0,352293	1	11	1	6
9	3	74,9859	0,500283	1	5	1	10
10	2	66,6038	0,667924	1	10	1	11

```
  11          1     56,0883  0,878234  1     12        1       12


Final Partition

Cluster 1
  SPI-n  IDH-n  GOV-n  INSTIT-n  EPI-n  OHI-n
Cluster 2
  INT COP-n  EFFEC-n  CRED-n  REG COP-n
Cluster 3
  HPI-n
Cluster 4
  GINNI-n
```



Análise:

- Na análise dos PC´s as variaveis HPI e GINNI estavam separadas das demais e aqui cada uma compõem um novo cluster.

8.2.4 Regressões stepwise

Abaixo serão feiras as regressões stepwise para cada PC, com o objetivo de verificar as variáveis que mais se destacam.

Stepwise Regression: PC1 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-4,010	-5,962	-6,971	-7,125	-7,566	-7,583
GOV-n	0,0891	0,0736	0,0517	0,0415	0,0353	0,0228
T-Value	27,20	30,35	26,13	22,66	27,01	11,14
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
INT COP-n		0,04265	0,04873	0,05070	0,05190	0,05280
T-Value		13,36	25,60	34,80	53,30	63,68
P-Value		0,000	0,000	0,000	0,000	0,000
IDH-n			0,02769	0,01792	0,01730	0,01764
T-Value			15,82	10,79	15,64	18,92
P-Value			0,000	0,000	0,000	0,000
EPI-n				0,0216	0,0212	0,0205
T-Value				9,78	14,45	16,54
P-Value				0,000	0,000	0,000
OHI-n					0,01292	0,01159
T-Value					12,69	13,22
P-Value					0,000	0,000
INSTIT-n						0,0140
T-Value						7,26
P-Value						0,000
S	0,954	0,621	0,362	0,275	0,183	0,154
R-Sq	85,05	93,73	97,88	98,79	99,47	99,63
R-Sq(adj)	84,94	93,63	97,83	98,75	99,45	99,61

Análise:

- Para o PC1, temos as variáveis GOV, International Cooperation, IDH-2014 que explicam 97,88% dos dados. Sendo o aumento seguinte é de menos de 1%, não sendo relevante.

Stepwise Regression: PC2 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-3,671	-2,493	-3,136	-2,823	-2,632	-2,569
INT COP-n	0,05908	0,08109	0,07600	0,07652	0,07469	0,07409
T-Value	11,23	39,67	59,66	75,96	98,51	126,04
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
SPI-n		-0,04557	-0,03978	-0,03612	-0,02600	-0,01822
T-Value		-29,42	-39,52	-40,32	-22,33	-14,88
P-Value		0,000	0,000	0,000	0,000	0,000

GINNI-n			0,01724	0,01596	0,01583	0,01502
T-Value			15,14	17,52	23,73	28,79
P-Value			0,000	0,000	0,000	0,000
OHI-n				-0,00933	-0,00931	-0,00891
T-Value				-8,85	-12,07	-14,94
P-Value				0,000	0,000	0,000
IDH-n					-0,01094	-0,01020
T-Value					-10,53	-12,67
P-Value					0,000	0,000
EPI-n						-0,0101
T-Value						-9,33
P-Value						0,000
S	1,16	0,421	0,253	0,200	0,146	0,113
R-Sq	49,23	93,41	97,64	98,54	99,22	99,54
R-Sq(adj)	48,84	93,31	97,58	98,49	99,19	99,52

Análise:

- Para o PC2, temos as variáveis International Cooperation, SPI e GINNI que explicam 97,64% dos dados. Sendo que o aumento seguinte é de menos de 1%, não sendo relevante.

Stepwise Regression: PC3 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	2,275	1,255	1,821	1,916	1,928	1,913
HPI-n	-0,04696	-0,04467	-0,04458	-0,03912	-0,03879	-0,03759
T-Value	-24,88	-34,05	-46,44	-94,46	-109,12	-182,40
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
OHI-n		0,01689	0,01365	0,02061	0,01926	0,01937
T-Value		12,08	12,80	43,25	42,93	79,10
P-Value		0,000	0,000	0,000	0,000	0,000
GINNI-n			-0,01075	-0,01322	-0,01313	-0,01325
T-Value			-10,63	-33,57	-39,13	-72,34
P-Value			0,000	0,000	0,000	0,000
IDH-n				-0,01103	-0,01252	-0,00904
T-Value				-27,63	-31,28	-30,46
P-Value				0,000	0,000	0,000
INSTIT-n					0,00295	0,00700
T-Value					7,03	21,37
P-Value					0,000	0,000
SPI-n						-0,00760
T-Value						-17,28
P-Value						0,000
S	0,466	0,320	0,234	0,0889	0,0756	0,0412
R-Sq	82,64	91,85	95,67	99,38	99,56	99,87
R-Sq(adj)	82,51	91,73	95,57	99,36	99,54	99,86

Análise:

- Para o PC3, temos as variáveis HPI, OHI, GINNI e IDH que explicam 99,38% dos dados. Sendo o aumento seguinte é de menos de 1%, não sendo relevante.

Stepwise Regression: PC4 versus SPI-n; IDH-n; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4 on 12 predictors, with N = 132

Step	1	2	3	4	5	6
Constant	-1,104	-2,430	-1,784	-1,669	-1,379	-1,343
GINNI-n	0,02997	0,03569	0,03858	0,04147	0,04169	0,04123
T-Value	12,79	20,64	31,91	59,38	85,61	113,64
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
OHI-n		0,02071	0,02484	0,01635	0,01487	0,01468
T-Value		11,46	19,34	18,60	23,77	31,71
P-Value		0,000	0,000	0,000	0,000	0,000
INT COP-n			-0,01570	-0,02110	-0,02127	-0,01434
T-Value			-12,11	-26,58	-38,44	-18,15
P-Value			0,000	0,000	0,000	0,000
INSTIT-n				0,01311	0,01409	0,01453
T-Value				16,78	25,58	35,48
P-Value				0,000	0,000	0,000
HPI-n					-0,00517	-0,00487
T-Value					-11,65	-14,76
P-Value					0,000	0,000
REG COP-n						-0,00735
T-Value						-10,27
P-Value						0,000
S	0,567	0,401	0,275	0,154	0,107	0,0792
R-Sq	55,72	78,05	89,77	96,82	98,47	99,17
R-Sq(adj)	55,38	77,71	89,53	96,72	98,41	99,13

Análise:

- Para o PC4, temos as variáveis GINNI, OHI, International Cooperation e Institution que explicam 96,82% dos dados. Sendo o aumento seguinte é de 1,65%, não sendo relevante.

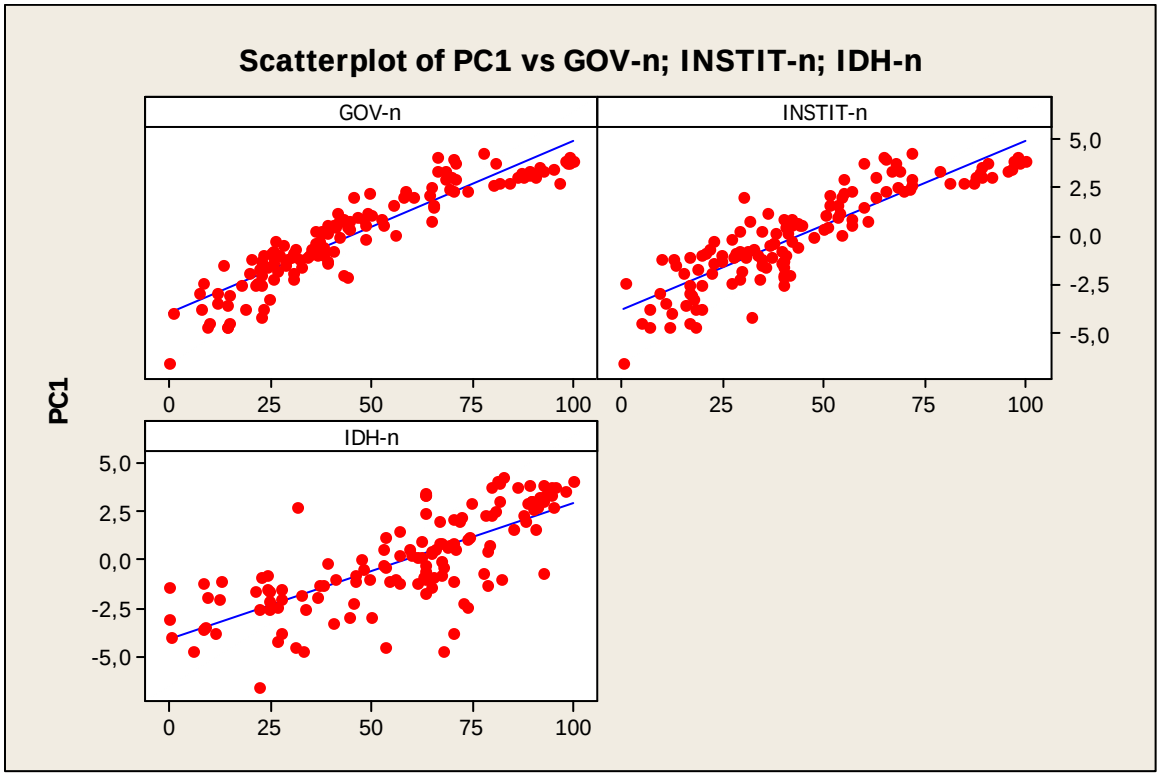
8.2.4 Análise de regressão com as principais variáveis das PCs

Antes de batizar as Componentes Principais é recomendado, logo apos realizadas as Regressões Stepwise, rodar uma Regressão Multiple para cada componente utilizando só as variáveis selecionadas no Stepwise. Dessa forma se pode avaliar o peso com o qual cada variável entra na Componente Principal.

Regression Analysis: PC1 versus GOV-n; INT COP-n; IDH-n

```
The regression equation is
PC1 = - 6,97 + 0,0517 GOV-n + 0,0487 INT COP-n + 0,0277 IDH-n


Predictor       Coef   SE Coef       T      P
Constant    -6,9711    0,1244  -56,03  0,000
GOV-n      0,051708  0,001979   26,13  0,000
INT COP-n  0,048730  0,001903   25,60  0,000
IDH-n      0,027695  0,001750   15,82  0,000


S = 0,362361   R-Sq = 97,9%   R-Sq(adj) = 97,8%


Analysis of Variance

Source           DF      SS      MS        F      P
Regression        3  775,40  258,47  1968,44  0,000
Residual Error  128   16,81    0,13
Total           131  792,21


Source     DF  Seq SS
GOV-n       1  673,80
INT COP-n   1   68,72
IDH-n       1   32,87


Unusual Observations

Obs  GOV-n      PC1      Fit  SE Fit  Residual  St Resid
  3     15  -4,4738  -3,5123  0,0573   -0,9614     -2,69R
 15     65   1,4435   2,2396  0,0563   -0,7961     -2,22R
 44     48  -0,1937   0,6849  0,0581   -0,8786     -2,46R
 54     14  -4,6790  -4,3575  0,1128   -0,3215     -0,93 X
 70     44  -2,1780  -1,3874  0,0705   -0,7906     -2,22R
113     31  -0,6388   0,2451  0,0873   -0,8839     -2,51R
115     97   2,7007   1,9453  0,1416    0,7554      2,26RX

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC1 com as principais variáveis tem um poder de explicação de 97,9%.



Regression Analysis: PC2 versus INT COP-n; SPI-n; GINNI-n

The regression equation is
PC2 = - 3,14 + 0,0760 INT COP-n - 0,0398 SPI-n + 0,0172 GINNI-n

Predictor	Coef	SE Coef	T	P
Constant	-3,13586	0,08894	-35,26	0,000
INT COP-n	0,075999	0,001274	59,66	0,000
SPI-n	-0,039778	0,001007	-39,52	0,000
GINNI-n	0,017245	0,001139	15,14	0,000

S = 0,253077 R-Sq = 97,6% R-Sq(adj) = 97,6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	339,13	113,04	1764,98	0,000
Residual Error	128	8,20	0,06		
Total	131	347,33			

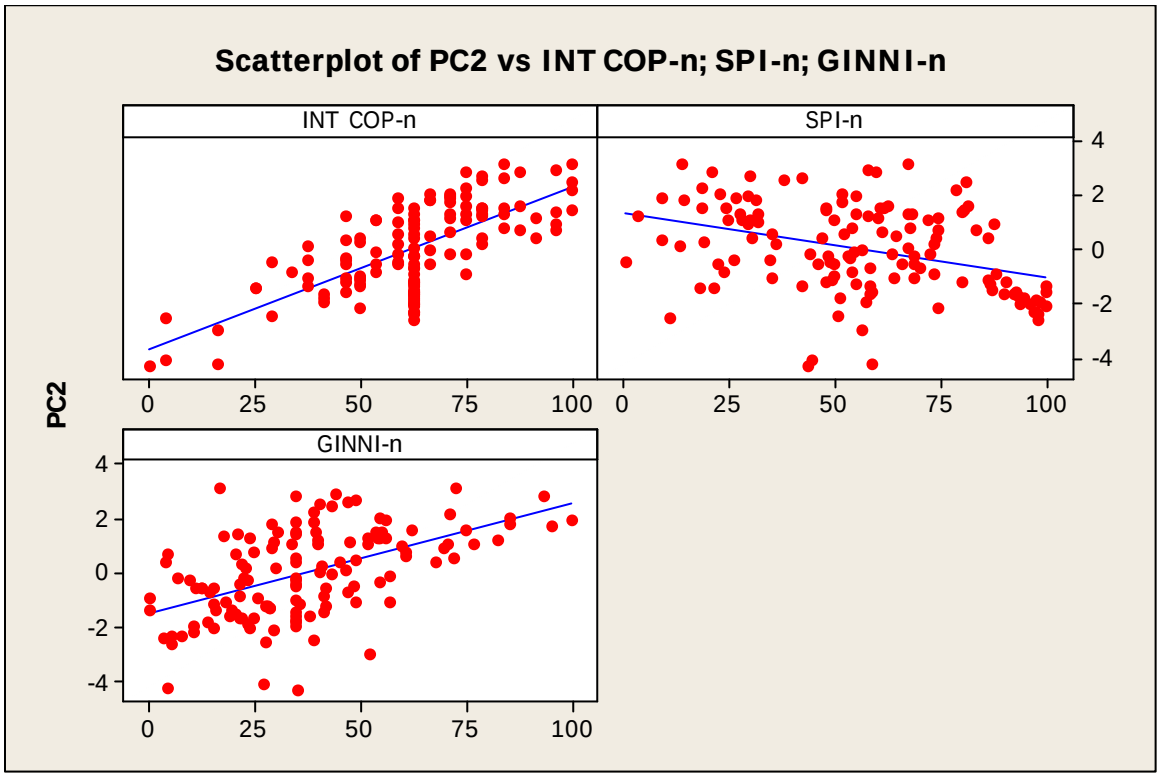
Source	DF	Seq SS
INT COP-n	1	170,99
SPI-n	1	153,46
GINNI-n	1	14,68

Unusual Observations

Obs	INT COP-n	PC2	Fit	SE Fit	Residual	St Resid
28	54	1,1208	0,5396	0,0332	0,5812	2,32R
48	63	1,0601	0,5384	0,0269	0,5218	2,07R
54	0	-4,2869	-4,2495	0,0774	-0,0374	-0,16 X

```
 77     38  -0,4021  -0,9356  0,0448    0,5336      2,14R
105     79   2,5670   2,0505  0,0389    0,5165      2,07R
113     63   0,9459   1,6511  0,0429   -0,7052     -2,83R
115     63  -1,3149  -2,0213  0,0465    0,7064      2,84R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC2 com as principais variáveis tem um poder de explicação de 97,6%.

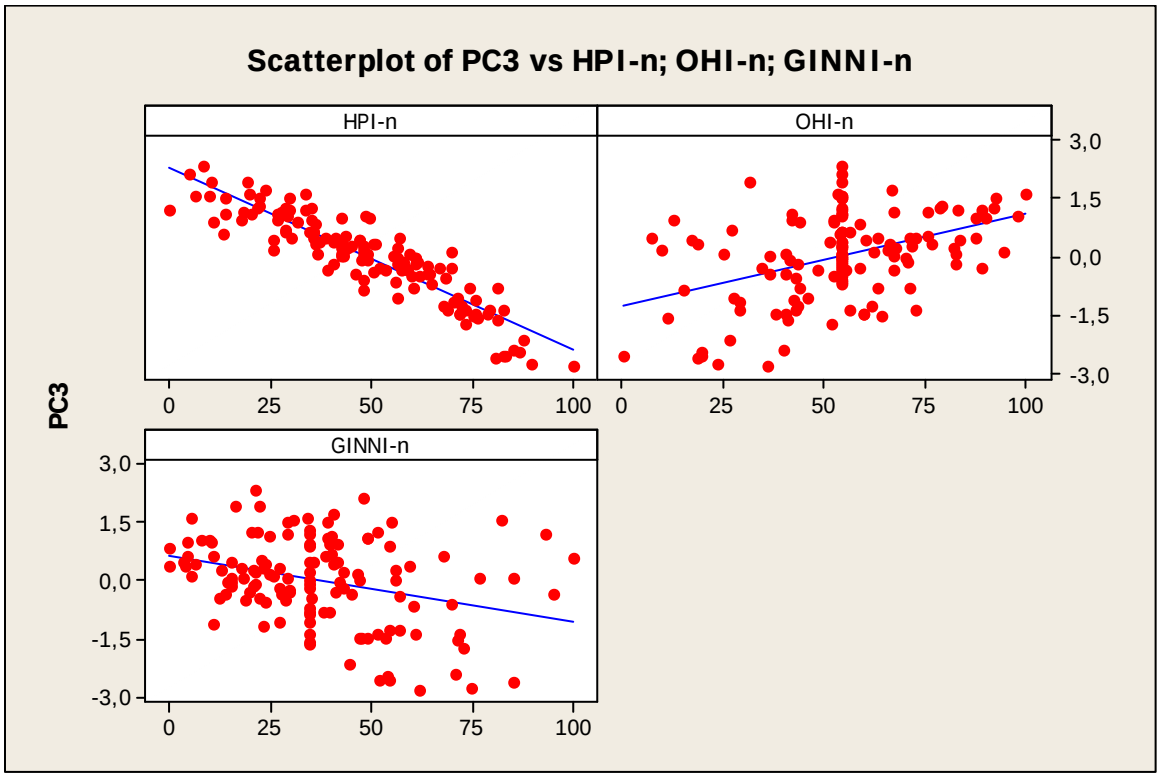


Regression Analysis: PC3 versus HPI-n; OHI-n; GINNI-n

```
The regression equation is
PC3 = 1,82 - 0,0446 HPI-n + 0,0137 OHI-n - 0,0108 GINNI-n


Predictor         Coef     SE Coef        T      P
Constant       1,82084    0,09586    18,99  0,000
HPI-n      -0,0445821  0,0009600   -46,44  0,000
OHI-n        0,013653   0,001067    12,80  0,000
GINNI-n     -0,010753   0,001012   -10,63  0,000


S = 0,234395   R-Sq = 95,7%   R-Sq(adj) = 95,6%


Analysis of Variance

Source           DF       SS      MS       F      P
Regression        3  155,488  51,829  943,36  0,000
Residual Error  128    7,032   0,055
Total           131  162,520
```

```
Source   DF   Seq SS
HPI-n     1  134,309
OHI-n     1   14,974
GINNI-n   1    6,205


Unusual Observations

Obs  HPI-n      PC3      Fit  SE Fit  Residual  St Resid
  9     81  -0,8195  -1,3169  0,0399    0,4974     2,15R
 15      0   1,2070   1,5603  0,0766   -0,3532    -1,59 X
 65     11   0,8877   1,5588  0,0441   -0,6711    -2,92R
 70     48   0,0836  -0,4008  0,0452    0,4844     2,11R
109     14   0,6051   0,8765  0,0752   -0,2714    -1,22 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

A equação apresentada para a PC3 com as principais variáveis tem um poder de explicação de 95,7%.

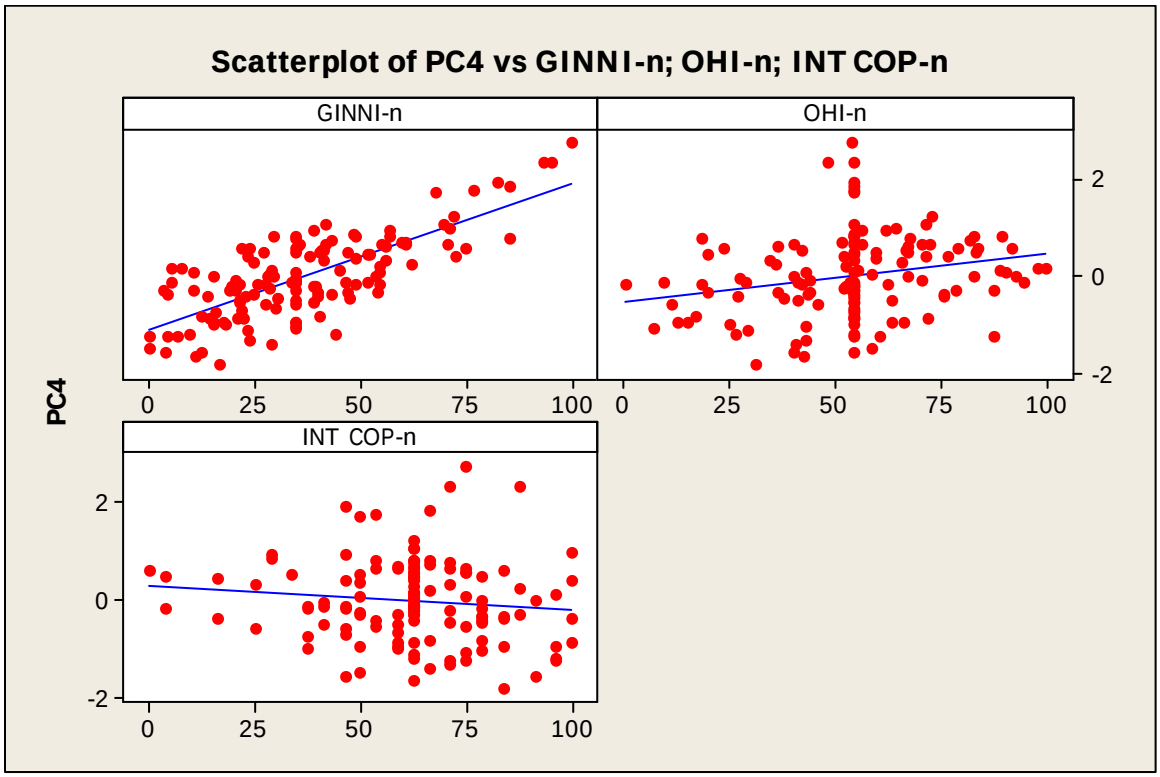


Regression Analysis: PC4 versus GINNI-n; OHI-n; INT COP-n

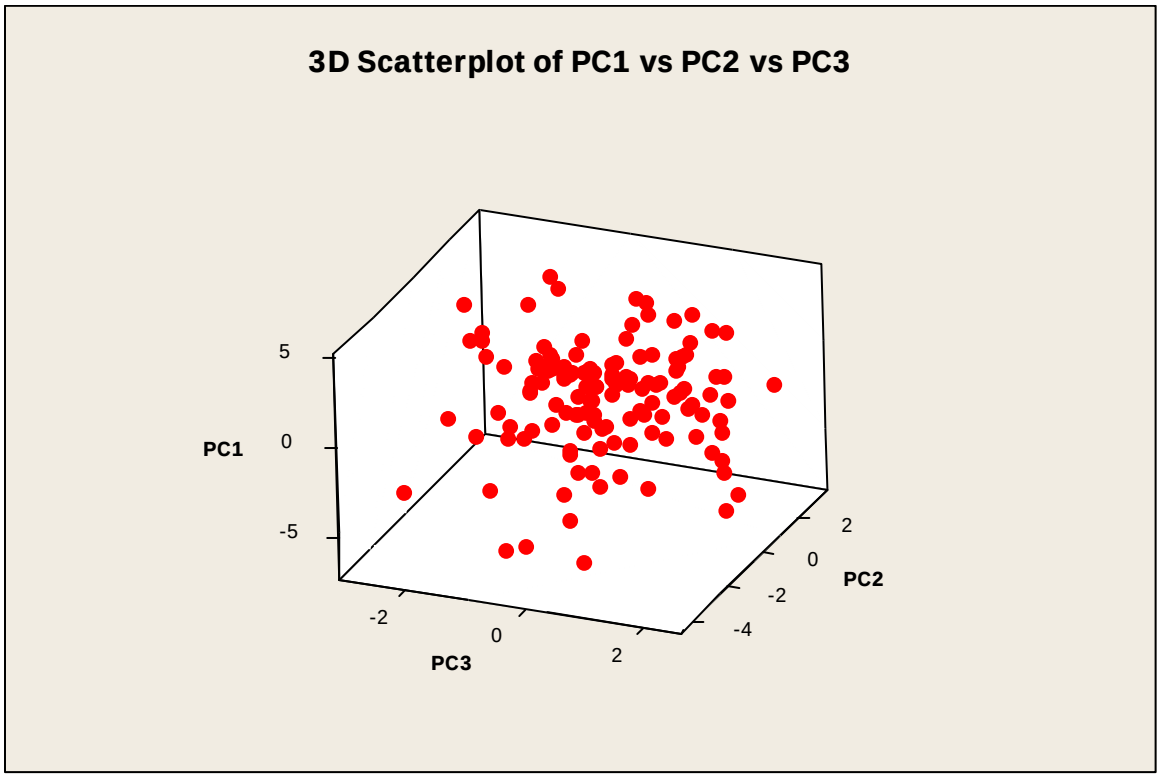
```
The regression equation is
PC4 = - 1,78 + 0,0386 GINNI-n + 0,0248 OHI-n - 0,0157 INT COP-n


Predictor       Coef   SE Coef       T      P
Constant     -1,7835    0,1070  -16,66  0,000
GINNI-n     0,038582  0,001209   31,91  0,000
OHI-n       0,024842  0,001285   19,34  0,000
```

```
INT COP-n   -0,015699   0,001297   -12,11   0,000


S = 0,274619    R-Sq = 89,8%    R-Sq(adj) = 89,5%


Analysis of Variance

Source            DF       SS       MS        F       P
Regression         3   84,708   28,236   374,41   0,000
Residual Error   128    9,653    0,075
Total            131   94,362


Source      DF  Seq SS
GINNI-n      1  52,574
OHI-n        1  21,080
INT COP-n    1  11,054


Unusual Observations

Obs  GINNI-n       PC4       Fit   SE Fit   Residual   St Resid
  9       35   -0,5032    0,2052   0,0273    -0,7084      -2,59R
 15       93    2,3620    1,7767   0,0734     0,5853       2,21R
 37       15   -0,9480   -0,3968   0,0371    -0,5512      -2,03R
 46       72    1,2534    1,8128   0,0600    -0,5594      -2,09R
131       29   -1,4146   -0,7010   0,0341    -0,7136      -2,62R

R denotes an observation with a large standardized residual.
```

A equação apresentada para a PC4 com as principais variáveis tem um poder de explicação de 89,8%.

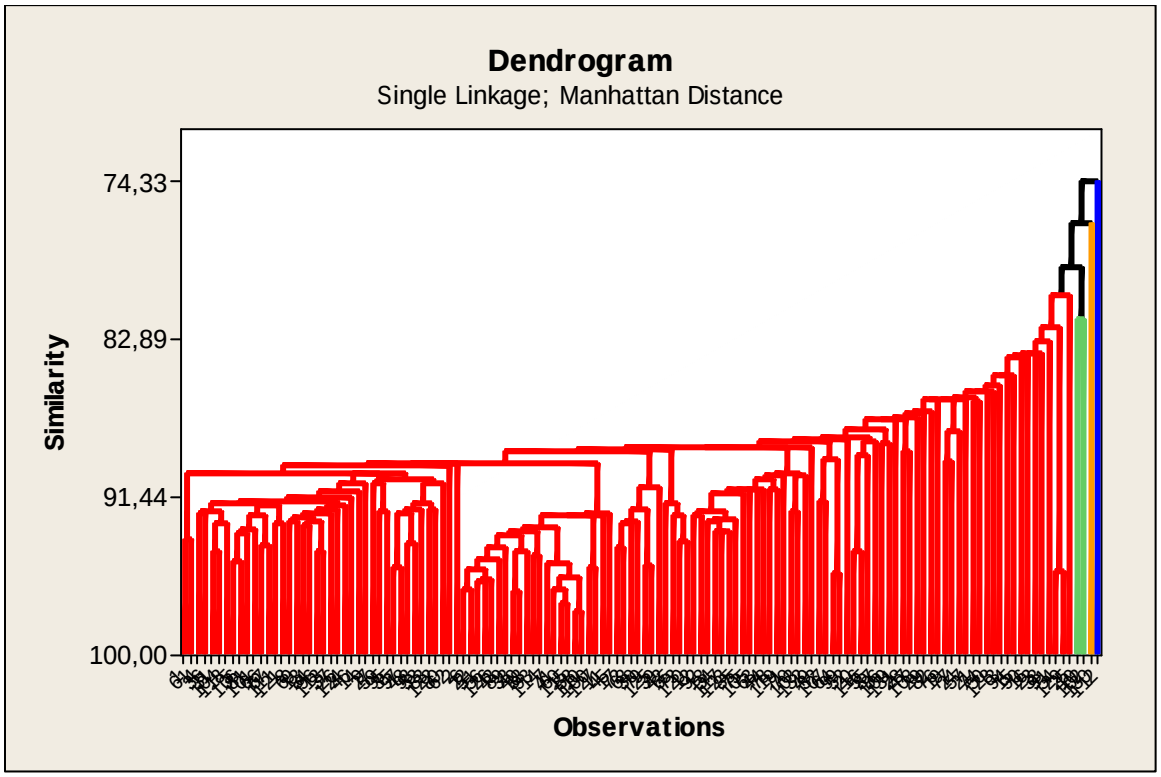


Apresentação em 3D das PC1, PC2 e PC3



8.4. Dendograma

Na análise do cluster para os 4 PCs temos o seguinte dendograma e agrupamentos:



Análise:

- Pela análise Manhanttan observa-se que a separação em 6 cluster

Cluster Analysis of Observations: PC1; PC2; PC3; PC4

Manhattan Distance, Complete Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	97,6620	0,4239	58	110	58	2
2	130	97,3312	0,4839	60	115	60	2
3	129	96,6679	0,6042	33	40	33	2
4	128	96,4661	0,6408	6	22	6	2
5	127	96,0856	0,7098	51	56	51	2
6	126	95,8715	0,7486	41	60	41	3
7	125	95,6266	0,7930	64	99	64	2
8	124	95,5502	0,8069	54	129	54	2
9	123	95,3232	0,8480	35	97	35	2
10	122	95,3040	0,8515	100	127	100	2
11	121	95,2040	0,8696	79	125	79	2
12	120	94,9582	0,9142	5	106	5	2
13	119	94,6257	0,9745	93	114	93	2
14	118	94,4895	0,9992	31	104	31	2
15	117	94,4490	1,0065	20	116	20	2
16	116	94,4288	1,0102	96	113	96	2
17	115	94,2407	1,0443	17	72	17	2
18	114	94,2041	1,0509	89	126	89	2
19	113	94,1717	1,0568	66	67	66	2
20	112	93,9905	1,0897	48	98	48	2
21	111	93,9433	1,0982	73	120	73	2
22	110	93,7748	1,1288	1	61	1	2
23	109	93,4121	1,1946	117	123	117	2
24	108	93,3524	1,2054	6	43	6	3
25	107	93,2301	1,2276	34	117	34	3
26	106	93,0492	1,2604	13	84	13	2
27	105	92,9073	1,2861	6	88	6	4
28	104	92,8753	1,2919	111	121	111	2
29	103	92,8110	1,3036	8	69	8	2
30	102	92,6969	1,3242	7	41	7	4
31	101	92,5195	1,3564	12	92	12	2
32	100	92,4235	1,3738	4	36	4	2
33	99	92,3135	1,3938	70	103	70	2
34	98	92,2359	1,4078	59	95	59	2
35	97	92,1844	1,4172	83	122	83	2
36	96	92,1608	1,4214	11	51	11	3
37	95	91,9741	1,4553	37	124	37	2
38	94	91,8066	1,4857	35	53	35	3
39	93	91,6843	1,5079	32	107	32	2
40	92	91,2905	1,5793	75	85	75	2
41	91	91,2359	1,5892	12	18	12	3
42	90	91,0962	1,6145	78	119	78	2
43	89	91,0665	1,6199	44	105	44	2
44	88	90,9620	1,6388	39	68	39	2
45	87	90,4991	1,7228	8	14	8	3
46	86	90,2439	1,7690	45	58	45	3
47	85	90,2383	1,7701	46	80	46	2
48	84	89,8626	1,8382	5	81	5	3
49	83	89,8032	1,8489	26	66	26	3
50	82	89,5392	1,8968	19	77	19	2
51	81	89,5084	1,9024	4	118	4	3
52	80	89,0140	1,9921	28	47	28	2
53	79	88,9506	2,0035	65	73	65	3
54	78	88,7963	2,0315	27	49	27	2
55	77	88,6468	2,0586	17	30	17	3

56	76	88,5960	2,0678	86	109	86	2
57	75	88,4937	2,0864	20	55	20	3
58	74	88,4687	2,0909	11	100	11	5
59	73	88,4333	2,0973	62	101	62	2
60	72	88,2555	2,1296	42	65	42	4
61	71	88,0930	2,1590	2	31	2	3
62	70	88,0537	2,1662	63	75	63	3
63	69	87,6741	2,2350	33	93	33	4
64	68	87,3297	2,2975	17	79	17	5
65	67	87,3060	2,3018	71	131	71	2
66	66	86,9954	2,3581	9	26	9	4
67	65	86,7667	2,3995	32	64	32	4
68	64	86,5231	2,4437	20	87	20	4
69	63	86,2748	2,4887	21	74	21	2
70	62	86,0556	2,5285	35	48	35	5
71	61	85,8512	2,5655	6	89	6	6
72	60	85,7108	2,5910	29	59	29	3
73	59	85,5974	2,6116	2	4	2	6
74	58	85,4524	2,6378	13	96	13	4
75	57	85,1181	2,6985	7	108	7	5
76	56	84,7047	2,7734	9	52	9	5
77	55	84,6006	2,7923	5	62	5	5
78	54	84,2887	2,8489	70	132	70	3
79	53	84,2382	2,8580	12	34	12	6
80	52	83,7673	2,9434	45	57	45	4
81	51	83,3364	3,0215	3	21	3	3
82	50	83,0961	3,0651	8	37	8	5
83	49	82,9155	3,0979	16	38	16	2
84	48	81,7870	3,3025	10	102	10	2
85	47	81,7069	3,3170	76	128	76	2
86	46	81,4664	3,3606	1	5	1	7
87	45	81,3868	3,3751	6	11	6	11
88	44	80,7960	3,4822	2	111	2	8
89	43	80,2733	3,5770	12	82	12	7
90	42	80,0581	3,6160	28	71	28	4
91	41	79,4646	3,7236	19	24	19	3
92	40	78,3519	3,9254	35	46	35	7
93	39	78,3207	3,9310	32	39	32	6
94	38	78,2512	3,9436	23	78	23	3
95	37	77,7086	4,0420	15	86	15	3
96	36	77,3006	4,1160	76	83	76	4
97	35	76,8681	4,1944	13	70	13	7
98	34	76,4399	4,2721	8	50	8	6
99	33	75,8565	4,3778	27	29	27	5
100	32	75,6966	4,4068	6	33	6	15
101	31	75,0984	4,5153	54	130	54	3
102	30	74,6376	4,5989	44	63	44	5
103	29	74,1916	4,6797	3	20	3	7
104	28	72,9249	4,9094	2	90	2	9
105	27	72,8802	4,9175	19	23	19	6
106	26	72,3966	5,0052	17	42	17	9
107	25	72,0546	5,0672	6	7	6	20
108	24	71,4140	5,1834	9	35	9	12
109	23	71,3747	5,1905	12	91	12	8
110	22	71,0911	5,2419	3	94	3	8
111	21	70,0978	5,4220	1	8	1	13
112	20	68,8685	5,6449	16	27	16	7
113	19	68,5988	5,6938	28	44	28	9
114	18	66,8811	6,0053	25	76	25	5
115	17	65,3787	6,2777	10	112	10	3
116	16	64,8590	6,3720	17	45	17	13
117	15	64,7495	6,3918	2	9	2	21
118	14	63,4976	6,6188	12	28	12	17
119	13	63,1653	6,6791	6	32	6	26
120	12	62,6482	6,7728	10	54	10	6
121	11	59,4515	7,3525	15	25	15	8
122	10	56,2052	7,9411	3	10	3	14
123	9	55,2669	8,1113	2	13	2	28
124	8	54,0408	8,3336	1	17	1	26
125	7	52,0071	8,7023	15	16	15	15
126	6	43,7918	10,1920	3	19	3	20

127	5	43,4330	10,2571	1	2	1	54
128	4	31,6985	12,3848	12	15	12	32
129	3	30,3414	12,6309	1	6	1	80
130	2	16,0320	15,2256	1	12	1	112
131	1	0,0000	18,1326	1	3	1	132

Final Partition
Number of clusters: 6

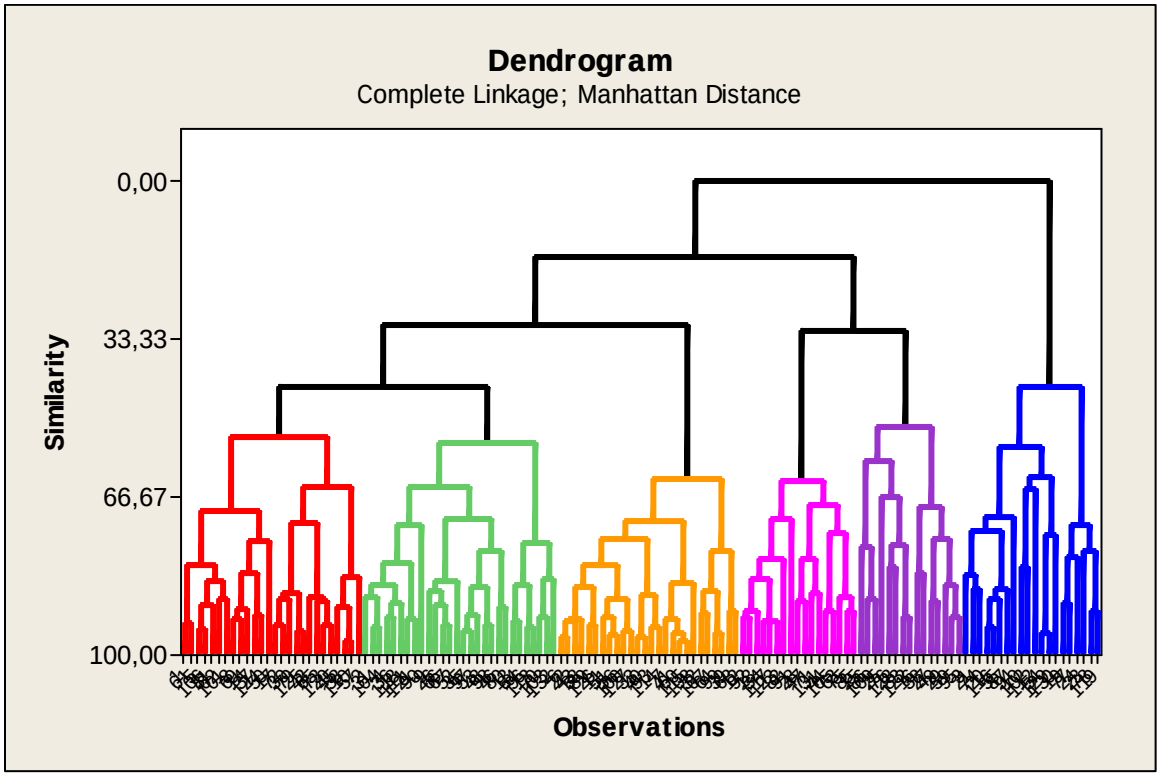
	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	26	82,752	1,69140	2,74467
Cluster2	28	78,998	1,59302	2,55184
Cluster3	20	119,137	2,29294	3,46317
Cluster4	26	68,124	1,50221	2,77945
Cluster5	17	34,019	1,30773	2,21073
Cluster6	15	69,699	2,08162	3,17981

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Grand centroid
PC1	0,728407	-0,897776	-3,71213	3,36418	-1,53471	1,27086	-0,0000000
PC2	-0,491567	0,145837	-1,23250	-1,12301	1,96820	1,93908	0,0000000
PC3	0,314055	-0,861093	0,42318	0,42705	0,88976	-1,24984	0,0000000
PC4	-0,352210	0,221182	0,09319	-0,02367	-0,40051	0,56831	0,0000000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6
Cluster1	0,00000	2,18186	4,52522	2,73254	3,39203	3,08147
Cluster2	2,18186	0,00000	3,38912	4,63611	2,67930	2,86187
Cluster3	4,52522	3,38912	0,00000	7,07813	3,93028	6,15742
Cluster4	2,73254	4,63611	7,07813	0,00000	5,82330	4,11349
Cluster5	3,39203	2,67930	3,93028	5,82330	0,00000	3,65904
Cluster6	3,08147	2,86187	6,15742	4,11349	3,65904	0,00000



8.5. ANÁLISE DISCRIMINANTE

A análise discriminante é utilizada para classificar observações em dois ou mais grupos de uma amostra com grupos conhecidos. Pode se usar a análise discriminante para investigar como variáveis contribuem para a separação do grupo e para colocar objetos ou indivíduos em grupos definidos.

Foi realizada uma Análise Discriminante com função linear e obteve-se 88,6% de acerto. O grupo 6 foi o melhor com 26 acertos.

Neste caso os países piores classificados foi um total de 15 (abaixo apresentado com *), sendo eles na ordem: Giorgia, India, Israel, Jordânia, Kenya, Líbano, Lituânia, Mauritius, Mongólia, Peru, Filipinas, Ruanda, Espanha, Togo e Venezuela.

Discriminant Analysis: 6 cluster ord versus GOV-n; INT COP-n; ...

```
Linear Method for Response: 6 cluster ord


Predictors: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; HPI-n; OHI-n


Group          1         2         3         4         5         6
Count         17        20        28        15        26        26


Summary of classification

                                  True Group
Put into Group      1      2      3      4      5      6
1                  15      1      2      0      0      0
2                   0     18      0      0      0      0
3                   1      1     24      0      4      0
```

4	0	0	2	15	0	0
5	1	0	0	0	19	0
6	0	0	0	0	3	26
Total N	17	20	28	15	26	26
N correct	15	18	24	15	19	26
Proportion	0,882	0,900	0,857	1,000	0,731	1,000

N = 132 N Correct = 117 Proportion Correct = 0,886

Squared Distance Between Groups

	1	2	3	4	5	6
1	0,0000	11,8892	10,4801	23,2256	14,5839	41,9352
2	11,8892	0,0000	13,7445	45,0450	20,5432	56,5291
3	10,4801	13,7445	0,0000	11,9813	8,4276	31,9822
4	23,2256	45,0450	11,9813	0,0000	17,4405	29,5899
5	14,5839	20,5432	8,4276	17,4405	0,0000	12,7421
6	41,9352	56,5291	31,9822	29,5899	12,7421	0,0000

Linear Discriminant Function for Groups

	1	2	3	4	5	6
Constant	-29,784	-18,385	-44,504	-68,923	-43,518	-66,437
GOV-n	-0,033	-0,126	-0,086	-0,020	-0,044	0,220
INT COP-n	0,508	0,261	0,480	0,675	0,492	0,507
IDH-n	0,059	0,125	0,175	0,164	0,171	0,106
SPI-n	0,161	0,116	0,205	0,378	0,348	0,458
GINNI-n	0,236	0,222	0,320	0,417	0,193	0,150
HPI-n	0,189	0,193	0,299	0,277	0,182	0,199
OHI-n	0,114	0,181	0,180	0,059	0,177	0,192

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
42**	5	3	1	18,962	0,001
			2	24,563	0,000
			3	6,295	0,663
			4	18,544	0,001
			5	7,671	0,333
			6	19,373	0,001
52**	3	1	1	9,735	0,500
			2	20,687	0,002
			3	9,802	0,483
			4	24,970	0,000
			5	16,821	0,014
			6	38,011	0,000
57**	5	3	1	32,45	0,000
			2	39,55	0,000
			3	11,63	0,665
			4	14,43	0,165
			5	14,48	0,160
			6	19,98	0,010
61**	5	3	1	13,197	0,027
			2	21,118	0,001
			3	6,921	0,620
			4	14,925	0,011
			5	8,120	0,341
			6	25,623	0,000
63**	1	3	1	4,353	0,463
			2	8,542	0,057
			3	4,294	0,477
			4	19,186	0,000
			5	14,224	0,003
			6	43,725	0,000
69**	5	3	1	14,412	0,005
			2	9,488	0,059
			3	4,603	0,678

			4	23,745	0,000
			5	6,539	0,258
			6	35,178	0,000
72**	5	6	1	29,619	0,000
			2	48,636	0,000
			3	25,346	0,000
			4	23,562	0,000
			5	7,613	0,372
			6	6,568	0,628
79**	5	6	1	24,369	0,000
			2	41,149	0,000
			3	19,741	0,000
			4	19,614	0,000
			5	6,961	0,278
			6	5,055	0,721
82**	1	5	1	10,368	0,104
			2	23,863	0,000
			3	15,741	0,007
			4	21,265	0,000
			5	6,076	0,888
			6	26,077	0,000
97**	3	4	1	19,306	0,000
			2	34,929	0,000
			3	6,220	0,132
			4	2,464	0,864
			5	13,234	0,004
			6	32,179	0,000
98**	3	4	1	14,369	0,005
			2	27,700	0,000
			3	5,037	0,482
			4	5,010	0,488
			5	10,931	0,025
			6	29,184	0,000
103**	3	1	1	9,325	0,578
			2	12,598	0,113
			3	10,587	0,308
			4	26,055	0,000
			5	21,157	0,002
			6	42,763	0,000
110**	5	6	1	31,079	0,000
			2	40,867	0,000
			3	19,195	0,000
			4	17,520	0,001
			5	5,797	0,280
			6	3,911	0,719
119**	2	1	1	4,720	0,757
			2	7,102	0,230
			3	12,868	0,013
			4	33,795	0,000
			5	20,581	0,000
			6	56,684	0,000
130**	2	3	1	39,87	0,000
			2	20,44	0,444
			3	19,99	0,556
			4	42,07	0,000
			5	34,38	0,000
			6	73,45	0,000

Foi realizada uma Análise Discriminante com função quadrática e obteve-se 97% de acerto. Nesta análise os grupos 5 e 6 foram os melhoroes com 26 acertos cada um.

Os países piores classificados foi para um total de 4 (abaixo apresentado com *), sendo eles na ordem: Jordânia, Kyrgyzstan, Peru e Tunísia. Os países Jordânia e Peru também apresentaram erro na análise linear.

Discriminant Analysis: 6 cluster ord versus GOV-n; INT COP-n; ...

Quadratic Method for Response: 6 cluster ord

Predictors: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; HPI-n; OHI-n

Group	1	2	3	4	5	6
Count	17	20	28	15	26	26

Summary of classification

	True Group					
Put into Group	1	2	3	4	5	6
1	17	0	0	0	0	0
2	0	20	0	0	0	0
3	0	0	25	0	1	0
4	0	0	1	15	0	0
5	0	0	2	0	25	0
6	0	0	0	0	0	26
Total N	17	20	28	15	26	26
N correct	17	20	25	15	25	26
Proportion	1,000	1,000	0,893	1,000	0,962	1,000

N = 132 N Correct = 128 Proportion Correct = 0,970

From Generalized Squared Distance to Group

Group	1	2	3	4	5	6
1	31,17	45,77	57,91	101,70	90,70	378,32
2	68,79	35,44	58,06	118,34	63,97	443,19
3	57,00	49,44	33,73	56,13	43,63	185,59
4	58,82	85,57	48,01	32,77	70,28	84,90
5	83,28	57,40	56,15	69,37	31,70	88,14
6	169,74	126,70	144,08	99,27	43,36	30,19

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
61**	5	3	1	101,48	0,000
			2	56,64	0,000
			3	41,40	0,747
			4	53,56	0,002
			5	43,58	0,251
			6	165,00	0,000
66**	3	5	1	79,27	0,000
			2	45,23	0,019
			3	39,00	0,422
			4	74,69	0,000
			5	38,43	0,559
			6	228,53	0,000
97**	3	4	1	66,84	0,000
			2	66,36	0,000
			3	39,29	0,151
			4	35,84	0,849
			5	62,54	0,000
			6	100,72	0,000
121**	3	5	1	96,34	0,000
			2	48,06	0,001
			3	38,00	0,211
			4	78,90	0,000
			5	35,36	0,788
			6	172,38	0,000

9. REGRESSÃO LOGÍSTICA

A regressão logística é uma técnica estatística apropriada quando a variável dependente é categórica (nominal ou não-métrica) e as variáveis independentes são métricas.

9.1 REGRESSÃO LOGÍSTICA 6 CLUSTERS

Neste caso a variável dependente será 6 CLUSTERS ORD e a relação a ser estudada será com as variáveis independentes: GOV, INT COP, IDH, SPI, GINNI, HPI e OHI.

Primeiramente será feita uma ANOVA One Way com cada uma das variáveis independentes e verificar que tem o maior "F", para que a partir da seja feita a ordenação e depois a regressão logística nominal e ordinal.

Esta ANOVA com IDH teve um ***F*** de **48,62**

One-way ANOVA: IDH-n versus 6 clusters

```
Source       DF     SS     MS      F      P
6 clusters    5  57618  11524  48,62  0,000
Error       126  29866    237
Total       131  87484

S = 15,40   R-Sq = 65,86%   R-Sq(adj) = 64,51%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  ---+---------+---------+---------+------
1      26  71,80   9,86                              (--*--)
2      28  56,67  16,22                        (--*--)
3      20  36,06  22,70           (--*--)
4      26  85,08  14,81                                    (--*--)
5      17  20,88  14,81  (--*---)
6      15  63,92  10,51                       (---*---)
                          ---+---------+---------+---------+------
                            20        40        60        80

Pooled StDev = 15,40
```

Esta ANOVA com GOV teve um ***F*** de **110,91**

One-way ANOVA: GOV-n versus 6 clusters

```
Source       DF     SS     MS       F      P
6 clusters    5  69212  13842  110,91  0,000
Error       126  15726    125
Total       131  84938

S = 11,17   R-Sq = 81,49%   R-Sq(adj) = 80,75%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
1      26  47,49  15,14                    (-*-)
2      28  33,43   6,49              (-*-)
3      20  15,20   7,63  (--*-)
4      26  85,05  11,30                                      (--*-)
5      17  27,77  11,03        (--*--)
6      15  52,29  13,65                    (--*--)
                          -----+---------+---------+---------+----
```

```
                              20        40        60        80

Pooled StDev = 11,17
```

Esta ANOVA com INT COP teve um ***F*** de 38,72

One-way ANOVA: INT COP-n versus 6 clusters

```
Source       DF     SS    MS      F      P
6 clusters    5  29675  5935  38,72  0,000
Error       126  19314   153
Total       131  48990

S = 12,38   R-Sq = 60,57%   R-Sq(adj) = 59,01%


                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
1      26  63,56  10,99                (--*-)
2      28  58,53  10,89              (-*--)
3      20  31,25  16,54  (--*-)
4      26  70,53  14,08                    (-*--)
5      17  69,85   9,27                   (--*--)
6      15  84,25  10,57                          (--*--)
                         -------+---------+---------+---------+--
                               40        60        80       100

Pooled StDev = 12,38
```

Esta ANOVA com SPI teve um ***F*** de 108,48

One-way ANOVA: SPI-n versus 6 clusters

```
Source       DF     SS     MS       F      P
6 clusters    5  69247  13849  108,48  0,000
Error       126  16086    128
Total       131  85333

S = 11,30   R-Sq = 81,15%   R-Sq(adj) = 80,40%


                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
1      26  64,42  10,35                      (-*-)
2      28  48,39  11,30              (-*-)
3      20  27,28  17,37    (--*-)
4      26  91,61   6,96                                    (-*-)
5      17  26,12  10,08   (--*--)
6      15  64,69   9,97                    (--*--)
                          +---------+---------+---------+---------
                         20        40        60        80

Pooled StDev = 11,30
```

Esta ANOVA com GINNI teve um ***F*** de 25,37

One-way ANOVA: GINNI-n versus 6 clusters

```
Source       DF     SS    MS      F      P
6 clusters    5  29374  5875  25,37  0,000
Error       126  29178   232
Total       131  58552

S = 15,22   R-Sq = 50,17%   R-Sq(adj) = 48,19%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --+---------+---------+---------+-------
1      26  24,24  12,83     (---*---)
2      28  48,11  16,81                    (--*---)
3      20  34,98  17,70           (---*---)
4      26  19,07  11,56  (---*---)
5      17  40,85  11,01              (----*---)
6      15  66,43  21,04                                (----*---)
                          --+---------+---------+---------+-------
                           16        32        48        64

Pooled StDev = 15,22
```

Esta ANOVA com HPI teve um ***F*** de 12,29

One-way ANOVA: HPI-n versus 6 clusters

```
Source       DF     SS    MS      F      P
6 clusters    5  19968  3994  12,29  0,000
Error       126  40927   325
Total       131  60894

S = 18,02   R-Sq = 32,79%   R-Sq(adj) = 30,12%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
1      26  43,30  17,75            (----*----)
2      28  65,07  13,22                              (---*----)
3      20  38,62  23,29         (-----*----)
4      26  48,98  11,79                 (----*---)
5      17  28,11  11,64  (-----*-----)
6      15  61,53  29,74                        (-----*-----)
                          -------+---------+---------+---------+--
                                30        45        60        75

Pooled StDev = 18,02
```

Esta ANOVA com OHI teve um ***F*** de 19,96.

One-way ANOVA: OHI-n versus 6 clusters

```
Source       DF     SS    MS      F      P
6 clusters    5  23712  4742  19,96  0,000
Error       126  29932   238
Total       131  53644

S = 15,41   R-Sq = 44,20%   R-Sq(adj) = 41,99%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  --------+---------+---------+---------+-
1      26  60,03  13,40               (---*---)
2      28  49,95  14,88         (---*---)
3      20  42,29  17,22    (---*----)
4      26  77,03  14,99                          (---*---)
5      17  40,77  16,36  (----*----)
6      15  40,34  16,76  (----*----)
                          --------+---------+---------+---------+-
                                 45        60        75        90

Pooled StDev = 15,41
```

Nestas análises feitas com a ANOVA a variável com o maior F foi GOV com 110,91. A partir daí será feita a ordenação com base nas médias.

ANOVA GOV

```
Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
1      26  47,49  15,14                  (-*-)
2      28  33,43   6,49            (-*-)
3      20  15,20   7,63  (--*-)
4      26  85,05  11,30                                    (--*-)
5      17  27,77  11,03        (--*--)
6      15  52,29  13,65                    (--*--)
                          -----+---------+---------+---------+----
                              20        40        60        80
```

A ordenação será feita da menor media para a maior média, sendo:

De	Para
3	1
5	2
2	3
1	4
6	5
4	6

Depois da ordenação foi realizada a REGRESSÃO LOGÍSTICA ORDINAL, com duas variáveis com P >0,10, sendo SPI e GOV.

Ordinal Logistic Regression: 6 clusters ord versus GOV-n; INT COP-n; ...

```
Link Function: Logit


Response Information

Variable        Value  Count
6 clusters ord  1         20
                2         17
                3         28
                4         26
                5         15
                6         26
                Total    132


Logistic Regression Table

                                                        Odds     95% CI
Predictor         Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)      7,39612    1,28259   5,77  0,000
Const(2)      9,72121    1,42226   6,84  0,000
Const(3)      12,5500    1,61462   7,77  0,000
Const(4)      15,6045    1,85402   8,42  0,000
Const(5)      17,8502    2,00778   8,89  0,000
GOV-n      -0,0685537  0,0220907  -3,10  0,002   0,93   0,89   0,98
INT COP-n  -0,0680568  0,0145182  -4,69  0,000   0,93   0,91   0,96
IDH-n      -0,0091318  0,0166176  -0,55  0,583   0,99   0,96   1,02
SPI-n      -0,0874241  0,0278435  -3,14  0,002   0,92   0,87   0,97
GINNI-n    -0,0002264  0,0097913  -0,02  0,982   1,00   0,98   1,02
HPI-n      -0,0089289  0,0104702  -0,85  0,394   0,99   0,97   1,01
OHI-n       0,0057659  0,0127002   0,45  0,650   1,01   0,98   1,03


Log-Likelihood = -113,892
Test that all slopes are zero: G = 238,431, DF = 7, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      334,398  648  1,000
Deviance     227,783  648  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    6724     93,6  Somers' D              0,87
Discordant     457      6,4  Goodman-Kruskal Gamma  0,87
Ties             6      0,1  Kendall's Tau-a        0,72
Total         7187    100,0
```

Foi gerada uma nova REGRESSÃO LOGÍSTICA ORDINAL, somente com GOVe SPI, que tiveram P >0,10.

O resultado foi de P =0 e a aceitação do modelo foi de 92,2%

Ordinal Logistic Regression: 6 clusters ord versus GOV-n; SPI-n

```
Link Function: Logit
```

```
Response Information

Variable        Value  Count
6 clusters ord  1         20
                2         17
                3         28
                4         26
                5         15
                6         26
                Total    132


Logistic Regression Table

                                                    Odds     95% CI
Predictor         Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)      3,59331   0,598148   6,01  0,000
Const(2)      5,34190   0,667607   8,00  0,000
Const(3)      7,78973   0,836765   9,31  0,000
Const(4)      10,2685    1,05804   9,71  0,000
Const(5)      12,3421    1,25761   9,81  0,000
GOV-n      -0,0943379  0,0184415  -5,12  0,000   0,91   0,88   0,94
SPI-n      -0,0723936  0,0161096  -4,49  0,000   0,93   0,90   0,96


Log-Likelihood = -127,473
Test that all slopes are zero: G = 211,269, DF = 2, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      374,715  653  1,000
Deviance     254,945  653  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    6623     92,2  Somers' D              0,84
Discordant     556      7,7  Goodman-Kruskal Gamma  0,85
Ties             8      0,1  Kendall's Tau-a        0,70
Total         7187    100,0
```

9.2 REGRESSÃO LOGÍSTICA N CLASS 3R

Foi realizada uma ordenação e reclassificação para os 3 grupos de países OTHERS, AIBER, AVECO, sendo do pior para o melhor e assim uma regressão logística ordinal foi gerada.

Ordinal Logistic Regression: 3 ORD versus GOV-n; INT COP-n; ...

```
Link Function: Logit


Response Information

Variable  Value  Count
3 ORD     1         84
          2         21
          3         27
          Total    132


Logistic Regression Table
```

```
                                                    Odds     95% CI
Predictor        Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)      16,4121    3,51145   4,67  0,000
Const(2)      18,9443    3,74570   5,06  0,000
GOV-n       0,0327965  0,0387751   0,85  0,398   1,03   0,96   1,11
INT COP-n   0,0131226  0,0179027   0,73  0,464   1,01   0,98   1,05
IDH-n       0,0059486  0,0354693   0,17  0,867   1,01   0,94   1,08
SPI-n       -0,219631  0,0588958  -3,73  0,000   0,80   0,72   0,90
GINNI-n    -0,0187461  0,0164842  -1,14  0,255   0,98   0,95   1,01
HPI-n      -0,0533658  0,0187751  -2,84  0,004   0,95   0,91   0,98
OHI-n      -0,0217625  0,0184851  -1,18  0,239   0,98   0,94   1,01


Log-Likelihood = -51,216
Test that all slopes are zero: G = 136,405, DF = 7, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      111,886  255  1,000
Deviance     102,433  255  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    4400     95,7  Somers' D              0,91
Discordant     198      4,3  Goodman-Kruskal Gamma  0,91
Ties             1      0,0  Kendall's Tau-a        0,49
Total         4599    100,0
```

Foi gerada uma nova REGRESSÃO LOGÍSTICA ORDINAL, somente com SPI e HPI, que tiveram P >0,10.

O resultado foi de P =0 e a aceitação do modelo foi de %

Ordinal Logistic Regression: 3 ORD versus SPI-n; HPI-n

```
Link Function: Logit


Response Information

Variable  Value  Count
3 ORD     1         84
          2         21
          3         27
          Total    132


Logistic Regression Table

                                                    Odds     95% CI
Predictor        Coef    SE Coef      Z      P  Ratio  Lower  Upper
Const(1)      14,3274    2,29742   6,24  0,000
Const(2)      16,7530    2,56841   6,52  0,000
SPI-n       -0,169770  0,0258633  -6,56  0,000   0,84   0,80   0,89
HPI-n      -0,0571402  0,0155884  -3,67  0,000   0,94   0,92   0,97


Log-Likelihood = -53,238
Test that all slopes are zero: G = 132,361, DF = 2, P-Value = 0,000
```

```
Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      130,067  260  1,000
Deviance     106,476  260  1,000


Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs       Number  Percent  Summary Measures
Concordant    4368     95,0  Somers' D              0,90
Discordant     230      5,0  Goodman-Kruskal Gamma  0,90
Ties             1      0,0  Kendall's Tau-a        0,48
Total         4599    100,0
```

REGRESSÃO LOGÍSTICA NOMINAL 3ORD

Nominal Logistic Regression: 3 ORD versus GOV-n; INT COP-n; ...

```
Response Information

Variable  Value  Count
3 ORD     3         27  (Reference Event)
          2         21
          1         84
          Total    132


Logistic Regression Table

                                                        Odds     95% CI
Predictor             Coef    SE Coef      Z      P  Ratio  Lower  Upper
Logit 1: (2/3)
Constant           3,88760    14,7272   0,26  0,792
GOV-n           -0,261855   0,161449  -1,62  0,105   0,77   0,56   1,06
INT COP-n        0,0270279  0,0561448   0,48  0,630   1,03   0,92   1,15
IDH-n           -0,0192609  0,0807374  -0,24  0,811   0,98   0,84   1,15
SPI-n            0,212191   0,185561   1,14  0,253   1,24   0,86   1,78
GINNI-n          0,200370  0,0873682   2,29  0,022   1,22   1,03   1,45
HPI-n           -0,0527986  0,0701692  -0,75  0,452   0,95   0,83   1,09
OHI-n            -0,127292  0,0967350  -1,32  0,188   0,88   0,73   1,06
Logit 2: (1/3)
Constant           35,1166    12,8364   2,74  0,006
GOV-n           -0,0170246   0,134004  -0,13  0,899   0,98   0,76   1,28
INT COP-n        0,0223337  0,0508707   0,44  0,661   1,02   0,93   1,13
IDH-n           -0,0716689  0,0631463  -1,13  0,256   0,93   0,82   1,05
SPI-n            -0,217301   0,160729  -1,35  0,176   0,80   0,59   1,10
GINNI-n          0,0569162  0,0814002   0,70  0,484   1,06   0,90   1,24
HPI-n            -0,129398  0,0680083  -1,90  0,057   0,88   0,77   1,00
OHI-n            -0,131668  0,0922017  -1,43  0,153   0,88   0,73   1,05


Log-Likelihood = -24,150
Test that all slopes are zero: G = 190,538, DF = 14, P-Value = 0,000


Goodness-of-Fit Tests

Method    Chi-Square   DF      P
Pearson      142,675  248  1,000
Deviance      48,299  248  1,000
```

9.4. REGERESSÃO LOGÍSTICA - CONSIDERAÇÕES FINAIS

Nas considerações finais a tabela abaixo, foi colocada para realizer a comparação entre a análise discriminante e a regressão logística.

	6 CLUSTERS ORD	**3 ORD**
AD (ANÁLISE DISCRIMINANTE)	**97%**	**88,6%**
RL (REGRESSÃO LOGISTICA)- ordinal	**92,2%**	**95%**

Obs.: As análises discriminantes estão no apêndice deste trabalho.

Segundo Hair, Jr: "Análise discriminante e regressão logística são as técnicas estatísticas apropriadas quando a variável dependente é categórica (nominal ou não métrica) e as variáveis independentes são métricas."

Nestas duas análises pode-se concluir que a análise discriminante teve melhor aceitação (97%) para o "6 CLUSTERS ORD" e a regressão logística (95%) para o "3 ORD".

9. APÊNDICE 1

Discriminant Analysis: 6 clusters ord versus GOV-n; INT COP-n; ...

Quadratic Method for Response: 6 clusters ord

Predictors: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; HPI-n; OHI-n

Group	1	2	3	4	5	6
Count	20	17	28	26	15	26

Summary of classification

	True Group					
Put into Group	1	2	3	4	5	6
1	20	0	0	0	0	0
2	0	17	0	0	0	0
3	0	0	25	1	0	0
4	0	0	2	25	0	0
5	0	0	1	0	15	0
6	0	0	0	0	0	26
Total N	20	17	28	26	15	26
N correct	20	17	25	25	15	26
Proportion	1,000	1,000	0,893	0,962	1,000	1,000

N = 132 N Correct = 128 Proportion Correct = 0,970

From Generalized Squared Distance to Group

Group	1	2	3	4	5	6
1	35,44	68,79	58,06	63,97	118,34	443,19
2	45,77	31,17	57,91	90,70	101,70	378,32
3	49,44	57,00	33,73	43,63	56,13	185,59
4	57,40	83,28	56,15	31,70	69,37	88,14
5	85,57	58,82	48,01	70,28	32,77	84,90
6	126,70	169,74	144,08	43,36	99,27	30,19

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
61**	4	3	1	56,64	0,000
			2	101,48	0,000
			3	41,40	0,747
			4	43,58	0,251
			5	53,56	0,002
			6	165,00	0,000
66**	3	4	1	45,23	0,019
			2	79,27	0,000
			3	39,00	0,422
			4	38,43	0,559
			5	74,69	0,000
			6	228,53	0,000
97**	3	5	1	66,36	0,000
			2	66,84	0,000
			3	39,29	0,151
			4	62,54	0,000
			5	35,84	0,849
			6	100,72	0,000
121**	3	4	1	48,06	0,001
			2	96,34	0,000
			3	38,00	0,211
			4	35,36	0,788
			5	78,90	0,000
			6	172,38	0,000

Discriminant Analysis: 3 ORD versus GOV-n; INT COP-n; ...

Quadratic Method for Response: 3 ORD

Predictors: GOV-n; INT COP-n; IDH-n; SPI-n; GINNI-n; HPI-n; OHI-n

Group	1	2	3
Count	84	21	27

Summary of classification

	True Group		
Put into Group	1	2	3
1	74	1	0
2	6	18	2
3	4	2	25
Total N	84	21	27
N correct	74	18	25
Proportion	0,881	0,857	0,926

N = 132 N Correct = 117 Proportion Correct = 0,886

From Generalized Squared Distance to Group

Group	1	2	3
1	37,62	52,97	110,18
2	44,72	33,95	53,64
3	55,05	46,42	31,78

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
17**	1	3	1	43,10	0,450
			2	49,74	0,016
			3	42,75	0,534
30**	1	3	1	41,87	0,072
			2	43,82	0,027
			3	36,83	0,900
31**	2	1	1	43,78	0,765
			2	46,14	0,235
			3	85,56	0,000
45**	3	2	1	41,00	0,315
			2	40,30	0,448
			3	41,57	0,237
48**	1	2	1	42,76	0,268
			2	40,76	0,732
			3	75,70	0,000
50**	1	2	1	51,33	0,179
			2	48,30	0,819
			3	60,82	0,002
58**	3	2	1	42,44	0,144
			2	39,42	0,652
			3	41,74	0,205
59**	1	2	1	52,70	0,001
			2	37,66	0,999
			3	59,96	0,000
76**	1	2	1	41,31	0,386
			2	40,43	0,602
			3	48,25	0,012
79**	1	3	1	46,34	0,295
			2	47,46	0,168
			3	45,13	0,537

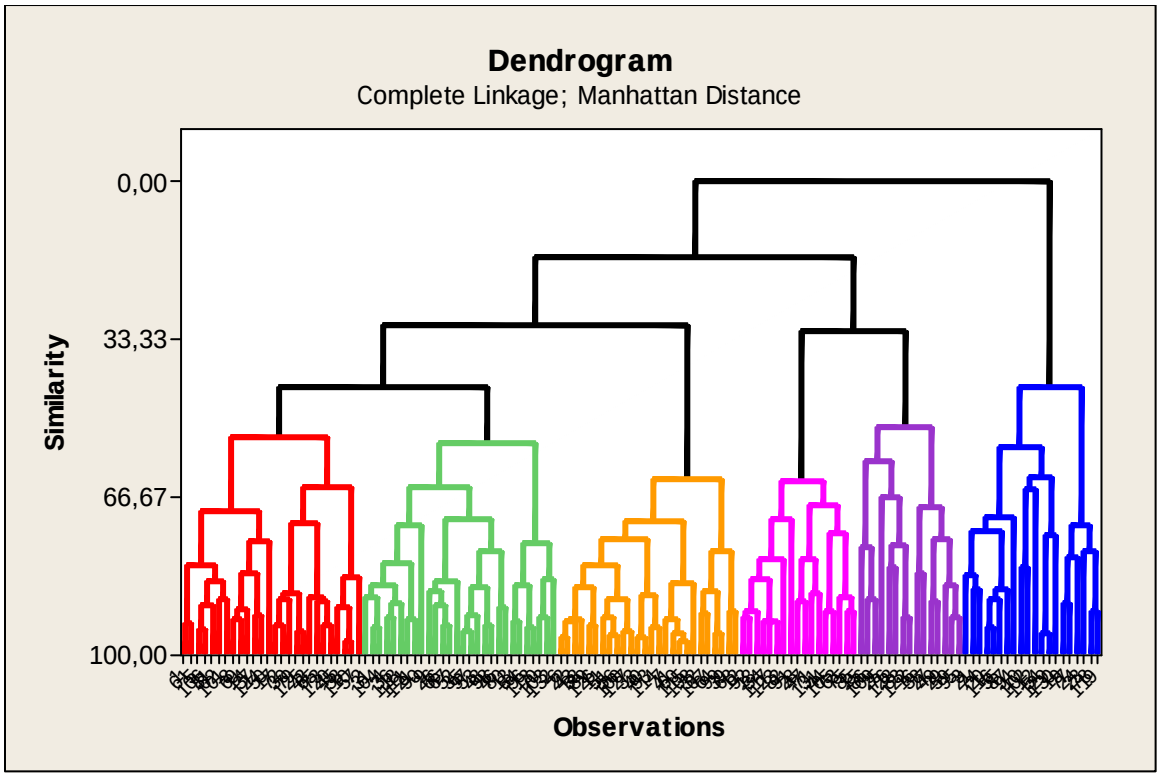
98**	1	2	1	45,58	0,025
			2	38,21	0,975
			3	56,21	0,000
99**	1	3	1	45,25	0,108
			2	48,51	0,021
			3	41,08	0,871
100**	2	3	1	49,02	0,002
			2	42,01	0,068
			3	36,79	0,929
110**	2	3	1	46,51	0,005
			2	39,74	0,157
			3	36,40	0,838
122**	1	2	1	40,59	0,464
			2	40,31	0,535
			3	53,51	0,001

10. ARVORE DE DECISÃO (que é o 11)

(falta terminar).

12. MÉTODO – ANÁLISE DE CORRESPONDÊNCIA

Primeiramente será realizada uma ANOVA *One-way* com cada umas variáveis mais importantes sendo: SPI, IDH, INT COP, HPI, OHI, GINNI e GOV, com o fator dos 6 grupos ordenados **(6 clusters ord)** pelo DENDOGRAMA MANHATTAN e depois uma nova ANOVA com fator das 3 regiões ordenadas **(3 ORD)**. A partir daí foi retirada a média de cada ANOVA e a classificação foi do pior para o melhor.



3.1 ANÁLISE DE CORRESPONDÊNCIA SIMPLES

A análise de correspondência simples, explora relações em uma classificação com 2 fatores. Você pode usar este procedimento com tabelas de 3 e 4 fatores porque o Minitab pode reduzi-las a tabelas de 2 fatores. A análise de correspondência simples decompõe uma tabela de

contingência de forma similar à análise de componentes principais em relação a dados contínuos multivariáveis. A análise de correspondência simples faz uma autoanálise dos dados, decompõe a variabilidade em dimensões subjacentes e associa variabilidade a linhas e/ou colunas.

A análise de correspondência foi realizada com as 7 variáveis, as 3 regiões e os 6 clusters.

3.1.1 Análise com as médias

Simple Correspondence Analysis: SPI; IDH; INT COP; HPI; OHI; GINNI; GOV

Contingency Table

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	42,630	47,720	58,440	41,780	44,590	20,000	33,360	288,520
AIBER	64,410	65,620	68,270	50,580	49,080	37,270	42,480	377,710
AVECO	90,300	85,710	68,840	72,360	75,830	56,810	83,250	533,100
G1	26,120	20,880	31,250	28,110	31,250	19,070	15,200	171,880
G2	27,280	36,060	58,530	38,620	58,530	24,240	27,770	271,030
G3	48,390	56,670	63,560	43,300	63,560	34,980	33,430	343,890
G4	64,420	63,920	69,850	48,980	69,850	40,850	47,490	405,360
G5	64,690	71,800	70,530	61,530	70,530	48,110	52,290	439,480
G6	91,610	85,080	84,250	65,070	84,250	66,430	85,050	561,740
Total	519,850	533,460	573,520	450,330	547,470	347,760	420,320	3392,710

Chi-Square Distances

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	0,056	0,122	1,916	0,317	0,083	3,099	0,159	5,753
AIBER	0,738	0,654	0,306	0,004	2,312	0,054	0,398	4,465
AVECO	0,909	0,042	5,024	0,036	1,208	0,086	4,482	11,787
G1	0,002	1,398	0,166	1,229	0,445	0,120	1,744	5,103
G2	4,889	1,009	3,528	0,194	5,005	0,451	1,005	16,081
G3	0,351	0,125	0,507	0,121	1,173	0,002	1,976	4,254
G4	0,086	0,001	0,026	0,433	0,301	0,012	0,148	1,006
G5	0,104	0,105	0,190	0,175	0,002	0,208	0,085	0,871
G6	0,356	0,119	1,208	1,208	0,451	1,360	3,433	8,136
Total	7,491	3,574	12,870	3,717	10,980	5,393	13,429	57,456

Relative Inertias

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	0,001	0,002	0,033	0,006	0,001	0,054	0,003	0,100
AIBER	0,013	0,011	0,005	0,000	0,040	0,001	0,007	0,078
AVECO	0,016	0,001	0,087	0,001	0,021	0,001	0,078	0,205
G1	0,000	0,024	0,003	0,021	0,008	0,002	0,030	0,089
G2	0,085	0,018	0,061	0,003	0,087	0,008	0,017	0,280
G3	0,006	0,002	0,009	0,002	0,020	0,000	0,034	0,074
G4	0,001	0,000	0,000	0,008	0,005	0,000	0,003	0,018
G5	0,002	0,002	0,003	0,003	0,000	0,004	0,001	0,015
G6	0,006	0,002	0,021	0,021	0,008	0,024	0,060	0,142
Total	0,130	0,062	0,224	0,065	0,191	0,094	0,234	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0111	0,6581	0,6581	**********
2	0,0028	0,1642	0,8224	*******
3	0,0013	0,0739	0,8963	***
4	0,0011	0,0621	0,9584	**
5	0,0005	0,0268	0,9852	*
6	0,0003	0,0148	1,0000	
Total	0,0169			

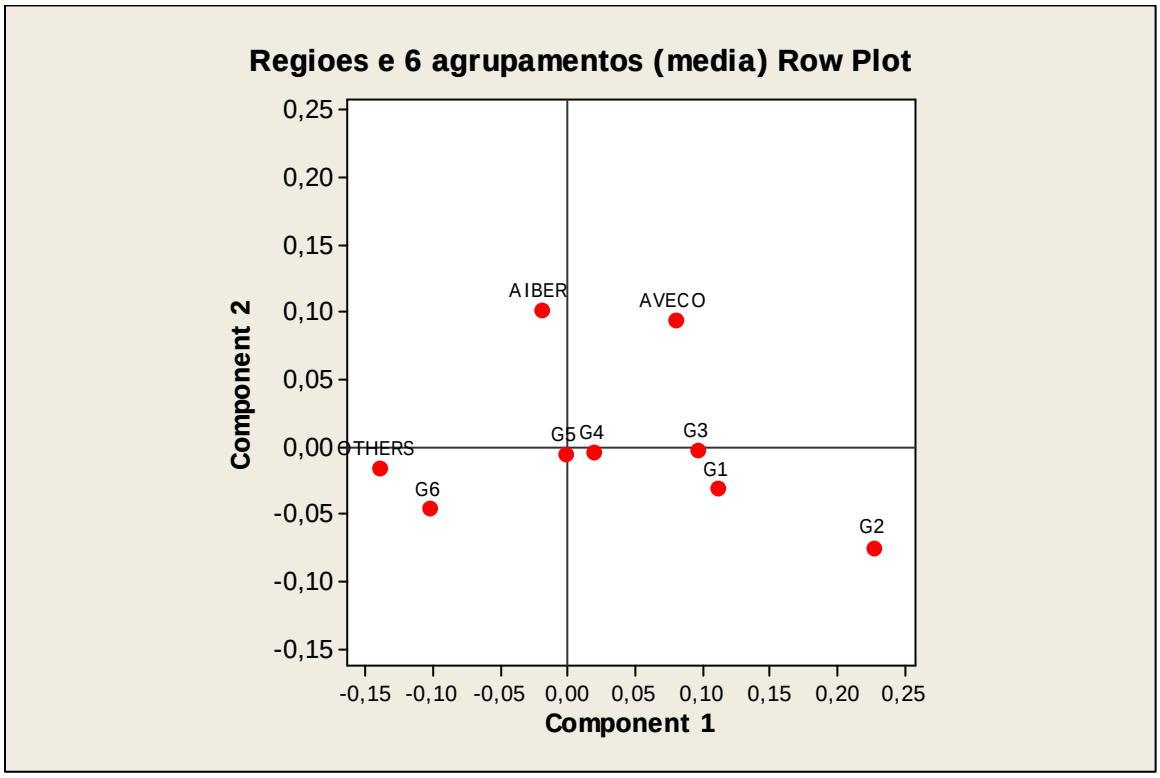
Row Contributions

ID	Name	Qual	Mass	Inert	Component 1 Coord	Component 1 Corr	Component 1 Contr	Component 2 Coord	Component 2 Corr	Component 2 Contr
1	OTHERS	0,758	0,085	0,100	0,079	0,314	0,048	0,094	0,444	0,271
2	AIBER	0,902	0,111	0,078	-0,021	0,038	0,004	0,101	0,864	0,409
3	AVECO	0,910	0,157	0,205	-0,141	0,898	0,280	-0,017	0,012	0,015
4	G1	0,444	0,051	0,089	0,111	0,413	0,056	-0,031	0,032	0,017
5	G2	0,954	0,080	0,280	0,226	0,861	0,366	-0,075	0,094	0,160
6	G3	0,728	0,101	0,074	0,095	0,727	0,082	-0,003	0,001	0,000
7	G4	0,131	0,119	0,018	0,018	0,124	0,003	-0,004	0,007	0,001
8	G5	0,023	0,130	0,015	-0,003	0,004	0,000	-0,006	0,018	0,002
9	G6	0,894	0,166	0,142	-0,104	0,749	0,161	-0,046	0,145	0,125

Column Contributions

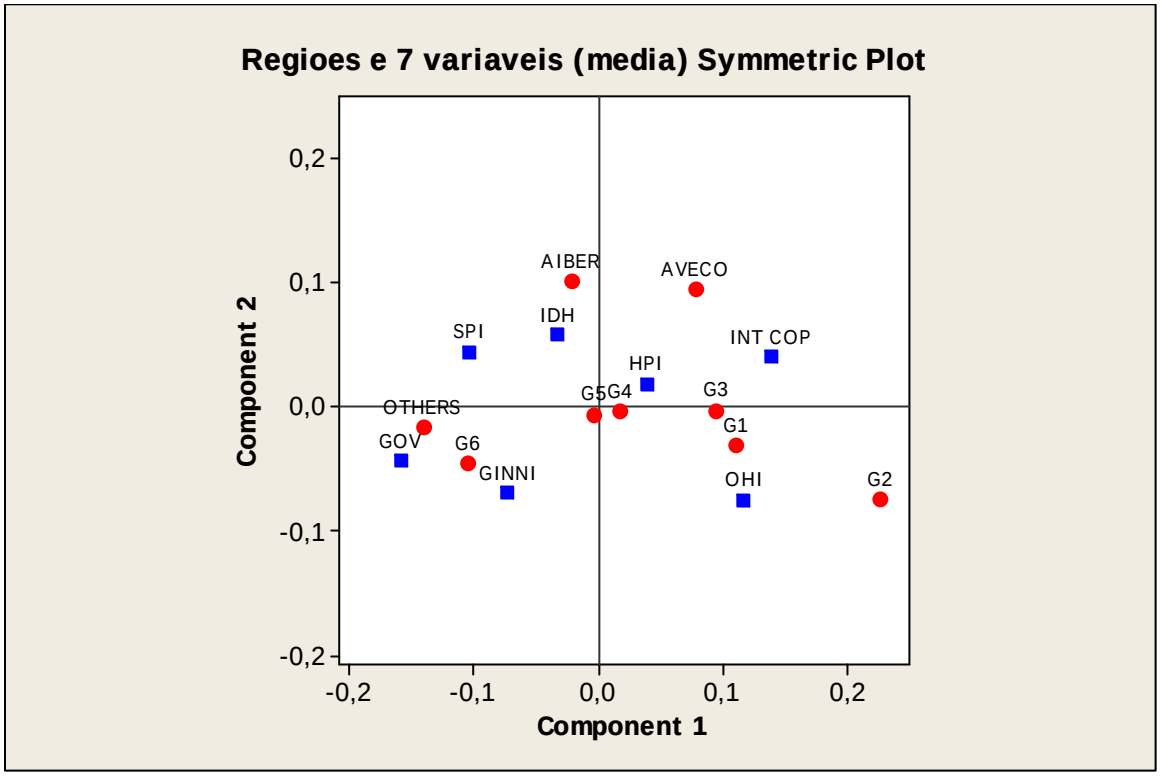
ID	Name	Qual	Mass	Inert	Component 1 Coord	Component 1 Corr	Component 1 Contr	Component 2 Coord	Component 2 Corr	Component 2 Contr
1	SPI	0,871	0,153	0,130	-0,103	0,737	0,146	0,044	0,134	0,106
2	IDH	0,648	0,157	0,062	-0,032	0,151	0,014	0,058	0,496	0,188
3	INT COP	0,936	0,169	0,224	0,140	0,868	0,295	0,039	0,068	0,093
4	HPI	0,232	0,133	0,065	0,040	0,193	0,019	0,018	0,039	0,016
5	OHI	0,967	0,161	0,191	0,117	0,685	0,199	-0,075	0,283	0,329
6	GINNI	0,650	0,103	0,094	-0,072	0,334	0,048	-0,070	0,316	0,181
7	GOV	0,847	0,124	0,234	-0,158	0,785	0,279	-0,044	0,062	0,088

Gráfico 2 das médias, relacionando as 3 regiões e os 6 agrupamentos - Row Plot



Analisando o gráfico 2 acima, uma primeira identificação é a distância entre as 3 regiões (OTHERS, AIBER, AVEC0. Isto demonstra heterogeneidade entre as 3 regiões e pode-se se dizer que as variáveis AIBER e OTHERS são homogêneas.

Gráfico 3 das médias, relacionando as 3 regiões e as 7 variáveis - Symmetric Plot



Analisando o gráfico com as variáveis, a grande maioria encontra-se na esfera positiva. G2 está isolado e somente próximo da variável OHI, países que se estão neste grupo são: Chile, Colombia, Costa Rica, Republica do Congp, etc. A região AVECO está próxima das variáveis INT COP e HPI e da mesma forma o grupo de países do G3, que tem países como China, Cuba, Equador, Republica Dominicana e etc.

3.1.2 Gráficos Radar e agrupamentos (médias)

Será demonstrado os gráficos do tipo radar e as tabelas com o objetivo de demonstrar os agrupamentos, como as variáveis se relacionam e seus graus de explicação.

Tabela 1 - Médias

SPI	IDH	INT COP	HPI	OHI	GINNI	GOV

OTHERS	42,63	47,72	58,44	41,78	44,59	20,00	33,36
AIBER	64,41	65,62	68,27	50,58	49,08	37,27	42,48
AVECO	90,30	85,71	68,84	72,36	75,83	56,81	83,25
R-q	**56,56%**	**35,22%**	**6,44%**	**26,05%**	**31,24%**	**27,40%**	**60,07%**
G1	26,12	20,88	31,25	28,11	31,25	19,07	15,20
G2	27,28	36,06	58,53	38,62	58,53	24,24	27,77
G3	48,39	56,67	63,56	43,30	63,56	34,98	33,43
G4	64,42	63,92	69,85	48,98	69,85	40,85	47,49
G5	64,69	71,80	70,53	61,53	70,53	48,11	52,29
G6	91,61	85,08	84,25	65,07	84,25	66,43	85,05
R-q	**81,15%**	**65,86%**	**60,57%**	**32,79%**	**44,20%**	**50,17%**	**81,49%**

Análise da tabela 1:

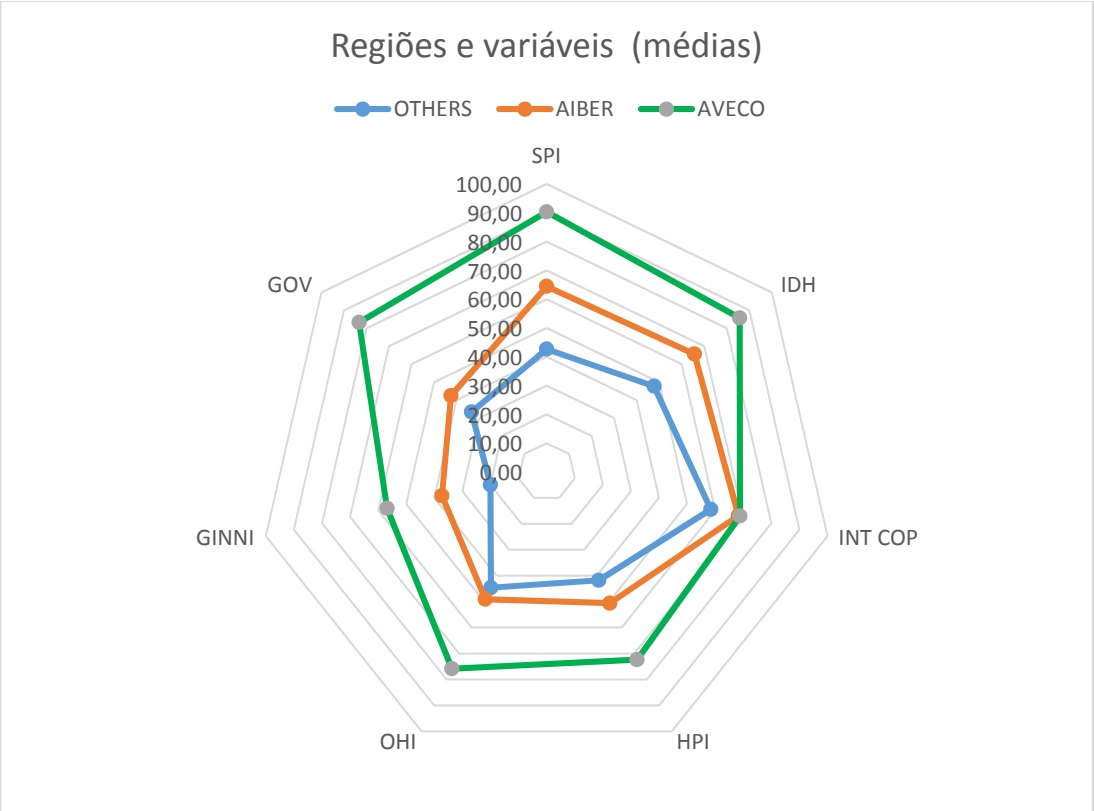
Regiões:

- Média mais alta: SPI
- Média mais baixa: GINNI
- R-q mais alto: GOV
- R-q mais baixo: INT COP

6 agrupamentos:

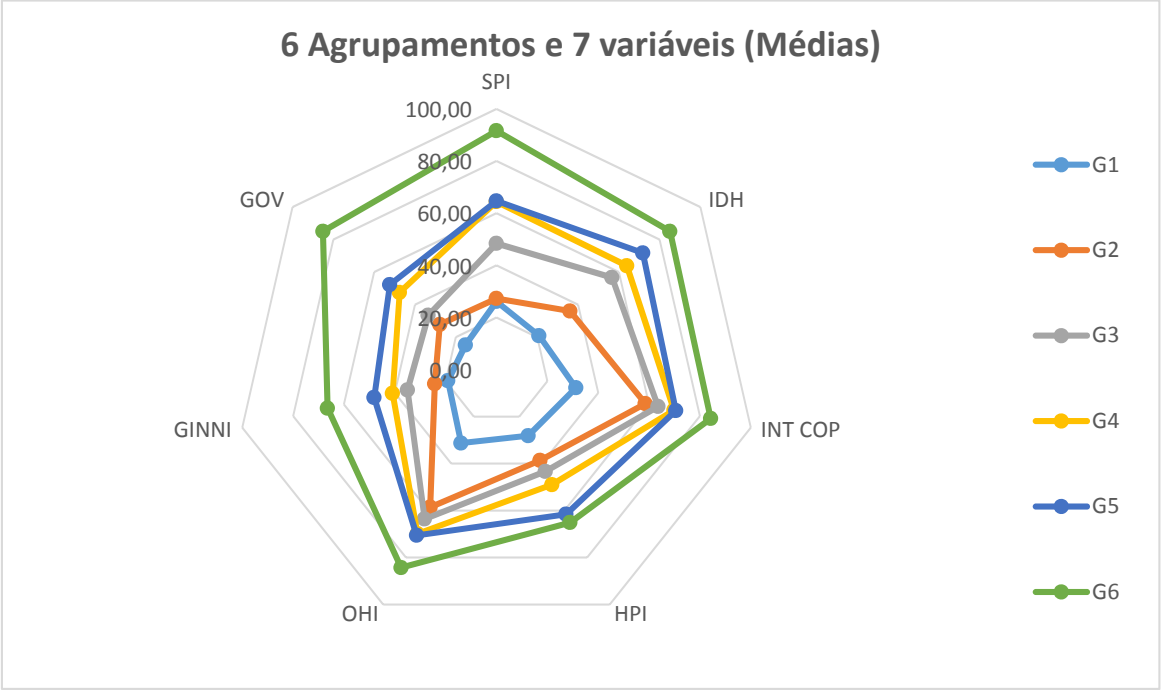
- Média mais alta: SPI
- Média mais baixa: GINNI
- R-q mais alto: GOV
- R-q mais baixo: HPI

Gráfico 1 – 3 Regiões e 7 variáveis (média)



Em 6 variáveis a região AVECO se destaca das demais, somente na variável INT COP ela fica na mesma posição que a AIBER e bem próxima de OTHERS.

Gráfico 2 – 6 agrupamentos e 7 variáveis



O agrupamento G6está mais distante em quase todas as variáveis, menos a HPI que está também próximo de G5. No grupo G6 encontram se países como: Áustria, Austrália, Bélgica, Republica Tcheca, Dinamarca e etc.

3.2 ANÁLISES COM OS DESVIOS PADRÃO

Simple Correspondence Analysis: SPI; IDH; INT COP; HPI; OHI; GINNI; GOV

Contingency Table

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	42,630	47,720	58,440	41,780	44,590	20,000	33,360	288,520
AIBER	64,410	65,620	68,270	50,580	49,080	37,270	42,480	377,710
AVECO	90,300	85,710	68,840	72,360	75,830	56,810	83,250	533,100
G1	26,120	20,880	31,250	28,110	31,250	19,070	15,200	171,880
G2	27,280	36,060	58,530	38,620	58,530	24,240	27,770	271,030
G3	48,390	56,670	63,560	43,300	63,560	34,980	33,430	343,890
G4	64,420	63,920	69,850	48,980	69,850	40,850	47,490	405,360
G5	64,690	71,800	70,530	61,530	70,530	48,110	52,290	439,480
G6	91,610	85,080	84,250	65,070	84,250	66,430	85,050	561,740
Total	519,850	533,460	573,520	450,330	547,470	347,760	420,320	3392,710

Chi-Square Distances

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	0,056	0,122	1,916	0,317	0,083	3,099	0,159	5,753
AIBER	0,738	0,654	0,306	0,004	2,312	0,054	0,398	4,465
AVECO	0,909	0,042	5,024	0,036	1,208	0,086	4,482	11,787
G1	0,002	1,398	0,166	1,229	0,445	0,120	1,744	5,103
G2	4,889	1,009	3,528	0,194	5,005	0,451	1,005	16,081
G3	0,351	0,125	0,507	0,121	1,173	0,002	1,976	4,254
G4	0,086	0,001	0,026	0,433	0,301	0,012	0,148	1,006
G5	0,104	0,105	0,190	0,175	0,002	0,208	0,085	0,871
G6	0,356	0,119	1,208	1,208	0,451	1,360	3,433	8,136
Total	7,491	3,574	12,870	3,717	10,980	5,393	13,429	57,456

Relative Inertias

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV	Total
OTHERS	0,001	0,002	0,033	0,006	0,001	0,054	0,003	0,100
AIBER	0,013	0,011	0,005	0,000	0,040	0,001	0,007	0,078
AVECO	0,016	0,001	0,087	0,001	0,021	0,001	0,078	0,205
G1	0,000	0,024	0,003	0,021	0,008	0,002	0,030	0,089
G2	0,085	0,018	0,061	0,003	0,087	0,008	0,017	0,280
G3	0,006	0,002	0,009	0,002	0,020	0,000	0,034	0,074
G4	0,001	0,000	0,000	0,008	0,005	0,000	0,003	0,018
G5	0,002	0,002	0,003	0,003	0,000	0,004	0,001	0,015
G6	0,006	0,002	0,021	0,021	0,008	0,024	0,060	0,142
Total	0,130	0,062	0,224	0,065	0,191	0,094	0,234	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0111	0,6581	0,6581	******************************
2	0,0028	0,1642	0,8224	*******
3	0,0013	0,0739	0,8963	***
4	0,0011	0,0621	0,9584	**
5	0,0005	0,0268	0,9852	*
6	0,0003	0,0148	1,0000	
Total	0,0169			

Row Contributions

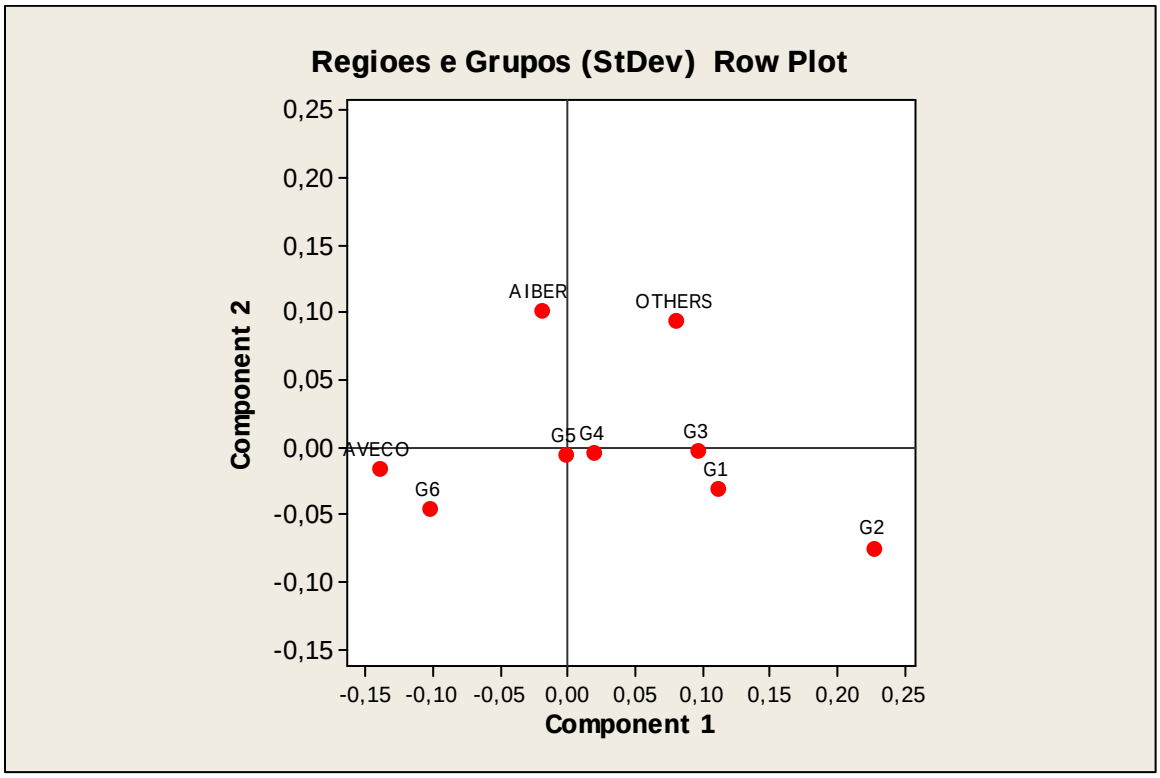
						Component 1			Component 2	
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	OTHERS	0,758	0,085	0,100	0,079	0,314	0,048	0,094	0,444	0,271
2	AIBER	0,902	0,111	0,078	-0,021	0,038	0,004	0,101	0,864	0,409

3	AVECO	0,910	0,157	0,205	-0,141	0,898	0,280	-0,017	0,012	0,015
4	G1	0,444	0,051	0,089	0,111	0,413	0,056	-0,031	0,032	0,017
5	G2	0,954	0,080	0,280	0,226	0,861	0,366	-0,075	0,094	0,160
6	G3	0,728	0,101	0,074	0,095	0,727	0,082	-0,003	0,001	0,000
7	G4	0,131	0,119	0,018	0,018	0,124	0,003	-0,004	0,007	0,001
8	G5	0,023	0,130	0,015	-0,003	0,004	0,000	-0,006	0,018	0,002
9	G6	0,894	0,166	0,142	-0,104	0,749	0,161	-0,046	0,145	0,125

Column Contributions

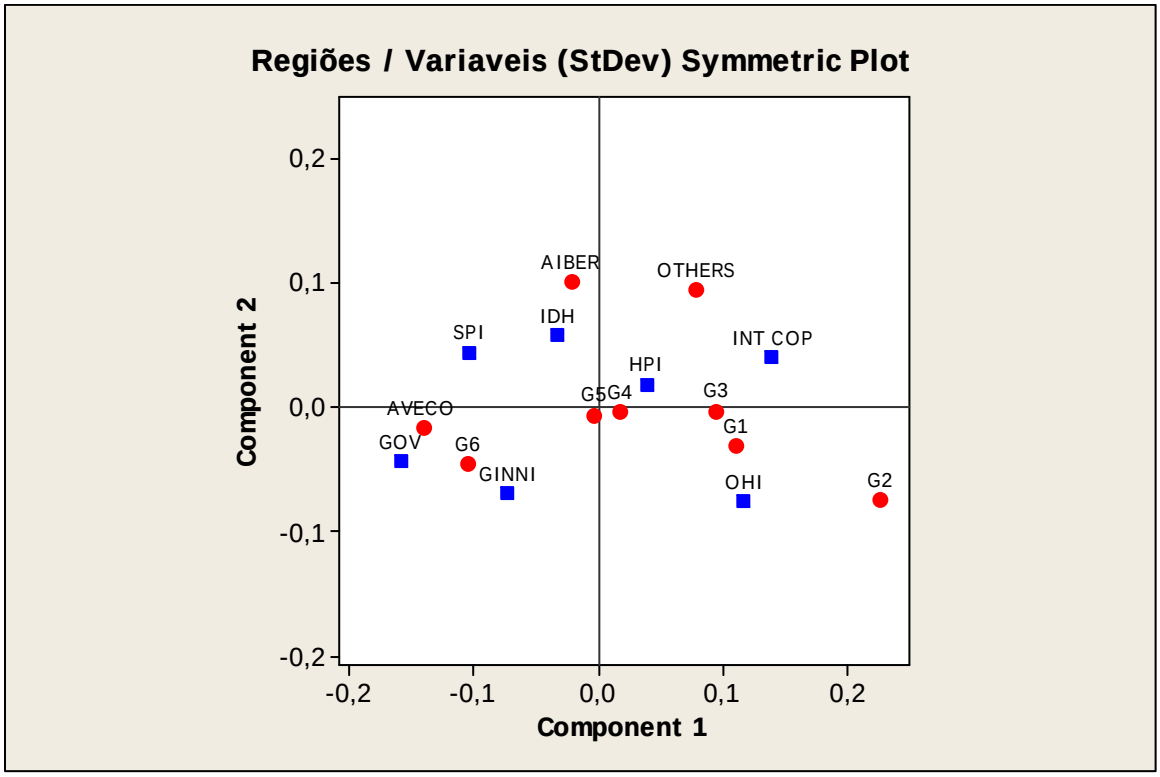
					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	SPI	0,871	0,153	0,130	-0,103	0,737	0,146	0,044	0,134	0,106
2	IDH	0,648	0,157	0,062	-0,032	0,151	0,014	0,058	0,496	0,188
3	INT COP	0,936	0,169	0,224	0,140	0,868	0,295	0,039	0,068	0,093
4	HPI	0,232	0,133	0,065	0,040	0,193	0,019	0,018	0,039	0,016
5	OHI	0,967	0,161	0,191	0,117	0,685	0,199	-0,075	0,283	0,329
6	GINNI	0,650	0,103	0,094	-0,072	0,334	0,048	-0,070	0,316	0,181
7	GOV	0,847	0,124	0,234	-0,158	0,785	0,279	-0,044	0,062	0,088

Gráfico 4 – 3 Regiões e 6 agrupamentos (desvio padrão)



Analisando o gráfico 4 acima, uma primeira identificação é a distância entre as 3 regiões (OTHERS, AIBER, AVECO), o mesmo acontece no gráfico 2 com as médias. Isto demonstra heterogeneidade entre as 3 regiões e pode-se se dizer que as variáveis AIBER e OTHERS são homogêneas, por estarem nos quadrantes positivos e próximas uma da outra.

Gráfico 5 – 3 Regiões e as 7 variáveis



Analisando o gráfico 5 com as variáveis, a grande maioria encontra-se na esfera positiva. G2 está isolado e somente próximo da variável OHI (muito semelhante ao gráfico 3 com as médias), países que se estão neste grupo são: Chile, Colombia, Costa Rica, Republica do Congp, etc. A região OTHERS (posição ocupada pela AVECO na análise das médias) está próxima das variáveis INT COP e HPI e da mesma forma o grupo de países do G3, que tem países como China, Cuba, Equador, Republica Dominicana e etc. E a região AVECO está no quadrante negativo.

3.2.2 Gráficos Radar e agrupamentos (desvio padrão)

Será demonstrado os gráficos do tipo radar e as tabelas com o objetivo de demonstrar os agrupamentos, como as variáveis se relacionam e seus graus de explicação.

Tabela 2 – Desvio padrão

	SPI	IDH	INT COP	HPI	OHI	GINNI	GOV
OTHERS	19,71	24,12	19,69	20,67	19,05	12,31	16,35
AIBER	11,95	12,03	21,82	12,84	16,82	20,34	18,82
AVECO	**8,70**	14,51	12,62	16,25	15,38	14,56	13,38
R-q	**56,56%**	**35,22%**	**6,44%**	**26,05%**	**31,24%**	**27,40%**	**60,07%**
G1	10,08	14,81	16,54	11,64	16,76	11,56	7,63
G2	17,37	22,70	10,89	23,29	16,36	12,83	11,03
G3	11,30	16,22	10,99	17,75	17,22	17,70	6,49
G4	10,35	10,51	9,27	11,79	14,88	11,01	15,14
G5	9,97	9,86	14,08	29,74	13,40	16,81	13,65
G6	6,96	14,81	10,57	13,22	14,99	21,04	11,30
R-q	**81,15%**	**65,86%**	**60,57%**	**32,79%**	**44,20%**	**50,17%**	**81,49%**

Análise da tabela 1:

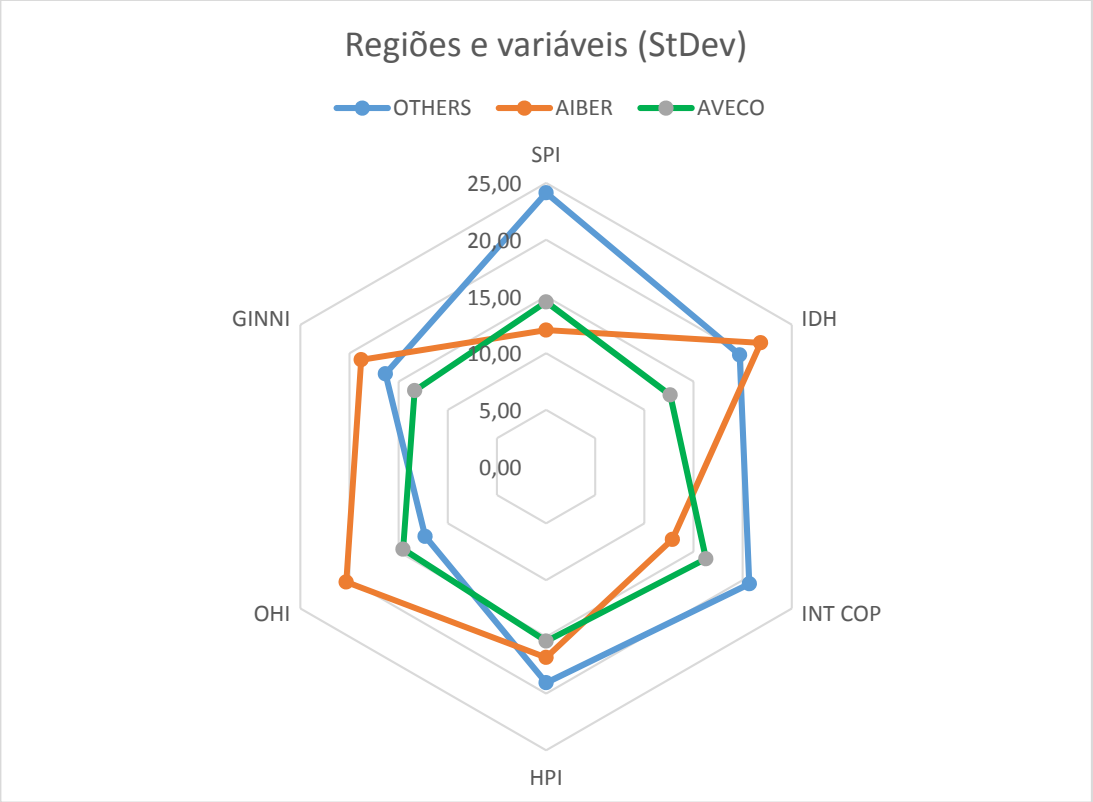
Regiões:

- StDev mais alto: IDH
- StDev mais baixo: SPI
- R-q mais alto: GOV
- R-q mais baixo: INT COP

6 agrupamentos:

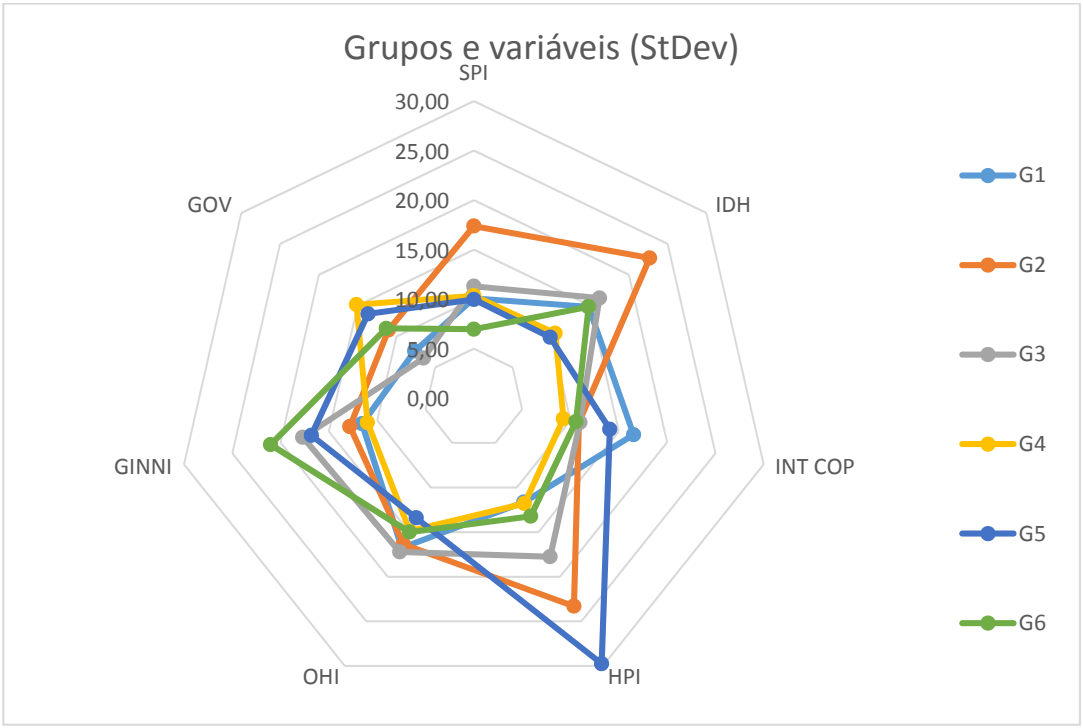
- StDev mais alto: HPI
- StDev mais baixo: SPI
- R-q mais alto: GOV
- R-q mais baixo: HPI

Gráfico 3 – 3 Regiões e 7 variáveis (desvio padrão)



A região AVECO é a que possui o menor desvio padrão na maioria, somente na OHI sendo maior que OTHERS, na SPI e INT COP maior que a AIBER.

Gráfico 4 – 6 agrupamentos e as 7 variáveis (desvio padrão)



No gráfico 4 o agrupamento G4 é o que parece que tem o menor desvio padrão na maioria das variáveis. E um destaque para o G5 com maior desvio padrão para a variável HPI

4. CONSIDERAÇÕES FINAIS

Um ponto que se destacou foi a variável SPI em todas as análises respondeu positivamente, na análise das médias teve a maior médio e o maior R-q, no caso do desvio padrão foram um dos menores.

Na análise de correspondência no, gráfico 3 está mais próxima da região AIBER, mas no gráfico de radar mais próxima da AVECO. Isto se repete na mesma análise com o desvio padrão.

Demonstra que o Índice de progresso social de um país, o SPI, é a variável mais que mais se destaca neste estudo.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FEA - Faculdade de Economia e Administração
Programa de Estudos Pós-Graduados em Administração

PROJETO ORIBER

TEMA

RECURSOS BASICOS: TRANSFORMAÇÃO PRODUTIVA E INOVAÇÃO SUTENTÁVEL

TRABALHO FINAL

Disciplina: METODOS QUANTITATIVOS e QUALITATIVOS DA PESQUISA EMPIRICA

Prof. Dr. Arnoldo José de Hoyos Guevara

MESTRANDA – Leda Honorato da Silva Reis

1. INTRODUÇÃO

O presente trabalho teve por objetivo efetuar uma análise exploratória dos dados relativos ao tema Transformação Produtiva e Inovação Sustentável, que possui um conjunto de dezoito variáveis previamente selecionadas que refletem o tema.

Para análise, selecionou-se uma amostra a partir dos dados de uma população com 132 países identificados pelo Programa Países Sustentáveis, pelo Guia Gestão Pública Sustentável uma Publicação do Núcleo de Estudos Futuros da PUC-SP, que a partir de agora chamaremos de GPS. Para a amostra objeto deste estudo selecionou-se os países da região Ibero Americana- AIBER, conforme indicado no Quadro 1.

Cabe ressaltar que incluiu-se Cuba como região do AIBER, e essa decisão foi tomada mediante análise não subordinada.

Cabe ressaltar que os dados da amostra foram normalizados e positivados e foram trabalhados a partir do programa de software estatístico MINITAB.

Para algumas das variáveis, o respectivo país não continha informação a respeito. Para minimizar a distorção dos dados, utilizamos a média dos demais países para completar o índice faltante. O quadro 2 apresenta os países e as variáveis com dados faltantes, às quais completou-se com a média dos demais países para essa mesma variável.

O trabalho inicia-se pelo entendimento dos dados a partir da análise de histogramas contendo: média, mínimo, máximo, mediana, variância.

Com o intuito de facilitar a leitura dos valores que envolvem cada uma das 18 variáveis, elaborou-se o Quadro 3, que resume a explicação dos dados que as compõem.

Para proceder a análise, efetuou-se regressão múltipla, analise discriminante, analise de correspondência.

O Objetivo é explorar os dados sobre **Transformação Produtiva e Inovação Sustentável (Basic Features: Productive Transformation and Sustainable Innovation)** de tal modo que possamos com os recursos citados anteriormente apresentar conclusões de correlações entre as variáveis que compõem o tema.

2 – ENTENDENDO A AMOSTRA

A partir de uma base de dados do GPS com 132 países, classificou-se os países Ibero Americano denominados neste trabalho como AIBER, e extraiu-se da base inicial de dados somente os países da região AIBER, que totalizaram 21 países, apresentados no Quadro 1 em ordem alfabética.

Incluiu-se Cuba na amostra, que originalmente estava classificado como CARLA- Países do Caribe.

Quadro 1 – Países classificados como Ibero Americano (AIBER)

Pais	Sigla	Sigla da regiao	Classificação de Renda	Classificação
Argentina	ARG	AIBER	Upper middle income	AIBER
Bolivia	BOL	AIBER	Lower middle income	AIBER
Brazil	BRA	AIBER	Upper middle income	AIBER
Chile	CHL	AIBER	High income: OECD	AIBER
Colombia	COL	AIBER	Upper middle income	AIBER
Costa Rica	CRI	AIBER	Upper middle income	AIBER
Cuba	CUB	Caribbean		AIBER
Dominican Republic	DOM	AIBER	Upper middle income	AIBER
Ecuador	ECU	AIBER	Upper middle income	AIBER
El Salvador	SLV	AIBER	Lower middle income	AIBER
Guatemala	GTM	AIBER	Lower middle income	AIBER
Honduras	HND	AIBER	Lower middle income	AIBER
Mexico	MEX	AIBER	Upper middle income	AIBER
Nicaragua	NIC	AIBER	Lower middle income	AIBER
Panama	PAN	AIBER	Upper middle income	AIBER
Paraguay	PRY	AIBER	Lower middle income	AIBER
Peru	PER	AIBER	Upper middle income	AIBER
Portugal	PRT	AIBER	High income: OECD	AIBER
Spain	ESP	AIBER	High income: OECD	AIBER
Uruguay	URY	AIBER	High income: nonOECD	AIBER
Venezuela	VEN	AIBER	Upper middle income	AIBER

Fonte: autor com base na Base de Dados do GPS

A amostra continha dados ausentes. Para validar os dados da amostra, fez-se necessário transformar os dados ausentes em dados válidos. Para tanto utilizou-se o critério de substituir os dados ausentes pela média, da mesma variável, correspondente aos demais países da população do GPS com 132 países.

Elaborou-se o Quadro 2 para deixar registrado os países que continham dados ausentes (faltantes), os quais foram substituídos pela média.

Quadro 2: Variáveis e Países com dados ausentes na seleção da amostra

Variável	***País (com informação ausente/faltante)***
Global Innovation Ind	Cuba
Institutions	Cuba
Human Capital and Research	Cuba
Infraestructure	Cuba
Market Sophistication	Cuba
Knowledge and Technology	Cuba
Creative Outputs	Cuba
International Cooperation	Portugal, Spain
Effective use of support	Portugal, Spain
Regional cooperation	Portugal, Spain

Fonte: autor com base na Base de Dados do GPS

2.1 -Normalização e Positivação dos Dados

A normalização dos dados é um ajuste nos dados para que os dados analisados fiquem dentro de uma certa normalidade. Por exemplo: a variável IDH é medida de 0 a 1, enquanto que a variável Suceptibilidad é medida em percentual. Desta forma é necessário apresentar os dados em valores simétricos (os dados das variáveis analisadas devem ser os mais próximos ou iguais possíveis) para que a análise e interpretação não sejam distorcidas.

Para realizar a Normalização dos dados usamos o link da barra de ferramentas do MINITAB denominada Calculator e aplicamos a fórmula desejada indicando uma coluna para que os novos valores (agora normalizados) sejam gerados.

A Positivação dos dados é necessária para que todas as variáveis indiquem seus dados no mesmo grau de compreensão. Por exemplo: a variável Suceptibilidad mede a probabilidade de desastres naturais prejudicarem a produção do país, então trata-se de um índice que quanto mais alto o seu valor, é pior para o país, enquanto que a variável Happy Planet mede a felicidade dos moradores de um país e quanto maior o seu valor melhor para o desenvolvimento do país. A Positivação faz com que haja a inversão do valor do índice , então a variável Suceptibilidad positivada faz com que o valor do índice indique que quanto maior melhor para o desenvolvimento do país, isto é quanto maior o índice, menor será a possibilidade de haver desastres naturais que venham a prejudicar a produtividade do país.

3. ENTENDENDO O TEMA: TRANSFORMAÇÃO PRODUTIVA E INOVAÇÃO SUSTENTÁVEL

A análise dos conceitos que compõem o tema foi extraída do material Programa Países Sustentáveis - Guia GPS – Gestão Pública Sustentável uma publicação do Núcleo de Estudos Futuros da PUC-SP –Pontifícia Universidade Católica. Nesse estudo são apresentadas as informações teóricas de outros indicadores, porém este trabalho contempla somente o tema Transformação Produtiva e Inovação Sustentável.

Vale destacar o desafio citado no GPS e sua proposta de solução:

> *"GPS - Guia para Gestão Pública Sustentável propõe então uma forma de contribuir na superação desses desafios, que consiste na promoção, a partir das secretarias de planejamento de cada país, de sinergias entre os setores científico-tecnológico, sociocultural e institucional, que harmonizem os processos e impactos do desenvolvimento ao nível local, tornando-o sustentável, procurando sempre estimular a participação dos cidadãos como forma de contribuir para a melhoria da qualidade de vida, e aproveitando de modo efetivo a troca de informações e experiências com outros países da região ibero-americana e suas redes de contatos."*

O tema Transformação Produtiva e Inovação Sustentável (Basic Features: Productive Transformation and Sustainable Innovation), é composto por dezoito variáveis quantitativas, divididas em 8 índices sintéticos e 10 indicadores ou variáveis componentes. " Com isso, a variável é denominada quantitativa pois mostra que ela pode ser medida em uma escala quantitativa, isto é, com números" (SANTOS; PARRA FILHO, 2011).
a saber:

- **Índices sintéticos:** Índice de Progresso Social (Social Progress Index), Índice de Sustentabilidade em Ecossistema (Ecosystem Sustainability), Combustible renewables

and waste (Combustíveis Renováveis e Resíduos), o Índice de Desenvolvimento Humano – IDH, publicado pelo PNUD – ONU, Suscetibilidade (Susceptibilidade), Índice de Proteção Ambiental – EPI (*Environmental Protection Index*), o *Happy Planet Index* (HPI), e o Índice de GINNI para os países selecionados.

- **Indicadores ou variáveis componentes:** Global Innovation Ind (inovação industrial global), Institutions (Instituições), Human Capital and Research (Recursos e Capital Humano), Infraestructure (Infraestrutura), Market Sophistication (Sofisticação de Mercado), Knowledge and Technology (Conhecimento e Tecnologia), Creative Outputs (Soluções Criativas), International Cooperation (Cooperação Internacional), Effective use of support (eficácia na utilização de apoio) e Regional Cooperation (Cooperação Regional)

Para melhor compeensão acerca das variáveis, resumiu-se no Quadro 3 essas variáveis com seu significado e sua unidade de medida.

Quadro 3. Variáveis do Eixo temático:Transformação Produtiva e Inovação Sustentável

Variável	Significado	Tipo	Unidade de Medida
Social Progress Index	O Índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidadas em três dimensões do Progresso Social: necessidades humanas básicas, Fundações de Bem-estar e oportunidade. A primeira dimensão, necessidades humanas básicas, avalia o quão bem um país prevê necessidades essenciais de seu povo através da medição se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se tiverem acesso à água potável, acesso a uma habitação condigna com serviços básicos, e se eles são seguros e protegidos.	Variável Quantitativa	Indicador nominal que varia de 0 a 100
Ecosystem Sustainability	A manutenção de um ecossistema sustentável depende da interação de 4 elementos: Produtividade, Diversidade, Resiliência, Distúrbios Naturais. A produtividade é a taxa de crescimento de cada parte viva de um ecossistema; os animais, plantas e outros organismos que compõem um ecossistema. Produtividade afeta o nível básico de recursos de um ecossistema pode proporcionar. Diversidade, o número de habitats, os tipos de espécies e a quantidade de variabilidade genética também influenciam a produtividade. Resiliência é a capacidade de um ecossistema de resistir e se recuperar de eventos de distúrbios naturais e é influenciada tanto pela produtividade e diversidade. Distúrbios Naturais, um sistema poder tornar-se insustentável caso o nível combinado de distúrbios naturais tornar-se muito grande. A sustentabilidade dos ecossistemas é determinada pela relação entre esses elementos, e juntos eles determinam o nível de recursos que podem ser tomadas a partir de um ambiente e ainda mantê-lo de forma sustentável.	Variável Quantitativa	Indicador nominal que varia de 0 a 100
IDH - 2013	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação.	Variável Quantitativa	Indicador nominal que varia de 0 a 1
Combustible renewables and waste	Compreendem a biomassa sólida, biomassa líquida, biogás, resíduos industriais e resíduos urbanos, medido em percentagem do consumo total de energia	Variável Quantitativa	Percentual
Suceptibilidad	Susceptibilidade refere-se à probabilidade de sofrer dano, perda e interrupção de um evento extremo ou desastres naturais. Assim, a susceptibilidade descreve as características estruturais e condições de enquadramento de uma sociedade.	Variável Quantitativa	Percentual
Global Innovation Ind	Reconhece o papel fundamental da inovação como motor do crescimento econômico e prosperidade, e reconhece a necessidade de uma ampla visão horizontal de inovação, que é aplicável à ambas as economias: desenvolvidas e emergentes, com a inclusão de indicadores que vão além das medidas tradicionais de inovação (como o nível de pesquisa e desenvolvimento em um determinado país).	Variável Quantitativa	Percentual
Institutions	É formado por três pilares: a) **Ambiente Político:** indica a percepção da probabilidade de que um país possa ser desestabilizado; da	Variável	Percentual

	qualidade dos serviços públicos e civis; da formulação e implementação de políticas; e a percepção de violações à liberdade de imprensa; b) **Ambiente Regulatório**; indica a percepção sobre a capacidade de um governo de formular e implementar políticas de coesão que promovam o desenvolvimento do setor privado c) **Ambiente de Negócios**: reflete a avaliação de três aspectos : a facilidade de começar um negócio, a facilidade de solução de insolvência e a facilidade de pagamento de impostos.	Quantitativa	
Human Capital and Research	Os quatro pilares do Índice são: (a) O pilar Educação contém indicadores relativos à aspectos quantitativos e qualitativos da educação do outro lado primário, secundário e terciário e contém informações tanto no presente como força de trabalho bem como a futura força de trabalho. (b) A Saúde e Bem-Estar pilar contém indicadores relativa ao bem-estar físico e mental de uma população sendo, desde a infância até a idade adulta. (c) O pilar da Força de Trabalho e Emprego destina-se a quantificar a experiência, talento, conhecimento e treinamento na população em idade ativa de um país. (d) O pilar Ambiente propício capta o legal quadro, infraestrutura e outros fatores que possibilitam retornos sobre o capital humano.	Variável Quantitativa	Percentual
Infraestructure	Infraestrutura ajuda a determinar o sucesso de fabricação e atividades agrícolas. Investimentos em água, saneamento, energia, habitação e transportes também melhorar a vida e ajudar a reduzir a pobreza. E as novas tecnologias de informação e comunicação para promover o crescimento, melhorar a prestação de saúde e outros serviços, ampliar o alcance da educação, e apoiar os avanços sociais e culturais. Dados aqui são compilados a partir de fontes tais como a Federação Internacional Road, Containerisation International, a Organização da Aviação Civil Internacional, a Associação Internacional de Energia, e da União Internacional das Telecomunicações.	Variável Quantitativa	Percentual
Market Sophistication	Este indicador é um derivado dos seguintes indicadores: - (a) Disponibilidade de serviços financeiros (b)A acessibilidade dos serviços financeiros (c) Financiamento através do mercado de ações local (d) Facilidade de acesso a empréstimos (e) A disponibilidade de capital de risco (f) Restrição sobre os fluxos de capital (g) Solidez dos bancos (h) A regulação das bolsas de valores e (i) índice de direitos legais.	Variável Quantitativa	Percentual
Knowledge and Technology	Direcionadas para a atividade de invenções e inovações, abrangendo: **(a) Criação do Conhecimento** que envolve: o resultado de Atividades Criativas e Inovadores. (**b) Sobre o Impacto do Conhecimento:** aumento da produtividade do trabalho, a entrada de novas empresas, os gastos com software e quantidade de certificação de qualidade ISSO 9001; **(c) Difusão do Conhecimento**: royalties e taxas de licenças; percentual de exportações de alta tecnologia; percentual das exportações de serviços de comunicação, computação e informação em relação ao total de exportação de serviços; e o percentual de saída de FDI (Foreign Direct Investment) em relação ao PIB (Produto Interno Bruto).	Variável Quantitativa	Percentual
Creative Outputs	A criatividade foi reconhecida como um motor econômico para gerar riqueza, emprego, o desenvolvimento sustentável das cidades do mundo, mudanças tecnológicas, inovação empresarial e reforço da competitividade das cidades e países individuais. Criatividade mostrada na produção econômica e nos domínios culturais, por exemplo, pode ser avaliado e apreciado pela comunidade, enquanto a inovação criativa em ciência ou ciência humana poderia ser reconhecida pelas comunidades científicas e acadêmicas. Isto significa "criatividade social" é mensurável e quantificável, embora "a criatividade da vida diária" é uma capacidade geral encontrada em outro lugar em nossa vida é mais difícil de medir.	Variável Quantitativa	Percentual
International Cooperation	Esta questão visa avaliar a disposição e capacidade da liderança política em desenvolver boas relações de vizinhança e cooperar com os países vizinhos em organizações internacionais e regionais • apoiar a integração regional ou internacional.	Variável Quantitativa	Indicador nominal que varia de 0 a 10
Effective use of support	Utilização eficáz de apoio : avalia se o Governo tem objetivos claros do desenvolvimento político e econômico, e sobre a forma como a liderança política faz uso de grandes projetos assistenciais internacionais, envolvendo cooperação técnica ou pessoal em sua própria agenda de desenvolvimento. O foco é sobre a capacidade de aprender a partir de know-how internacional, e se adaptar ao	Variável Quantitativa	Indicador nominal que varia de 0 a 10

	aconselhamento externo às realidades nacionais e ainda integrar essa assistência internacional em uma estratégia consistente de desenvolvimento de longo prazo.		
Regional Cooperation	Avalia a disposição e capacidade da liderança política em: Desenvolver relações de boa vizinhança; Cooperar com os países vizinhos em organizações internacionais e regionais; Apoiar a integração regional ou internacional; A liderança política ativamente e com êxito constrói e expande cooperativa vizinhança e internacional relacionamentos. Ela promove a integração regional e internacional; A liderança política coopera com muitos estados vizinhos e está em conformidade com as regras estabelecidas pela organizações regionais e internacionais.	Variável Quantitativa	Indicador nominal que varia de 0 a 10
EPI- Indice de desempenho ambiental	Classifica o desempenho dos países em questões ambientais principais: proteção à saúde humana e proteção do ecossistema ambiental	Variável Quantitativa	Percentual
HPI – Bem estar sustentável do país	Medida de bem estar sustentável país oferece uma vida feliz sustentável ao longo prazo para as pessoas que vivem nelas.	Variável Quantitativa	Percentual
GINNI index	Mede a desigualdade de distribuição de renda e de despesas das famílias. A medição do índice de Gini obedece a uma escala que vai de **0** (quando não há desigualdade) a **1** (com desigualdade máxima),	Variável Quantitativa	Indicador nominal de 0 a 1, quanto mais próximo de zero melhor

Fonte: autor a partir dos dados da planilha estatística e do GPS p.16

4 – ANÁLISE DE DADOS POR HISTOGRAMA

Elaborou-se o Histograma de todas as 18 variáveis do estudo e procedeu-se análise com base na média, desvio padrão, variância, valor mínimo e máximo de cada uma das variáveis.

A seguir apresentamos os dados da média, mínimo e máximo para todas as variáveis da amostra com os dados normalizados e positivados:

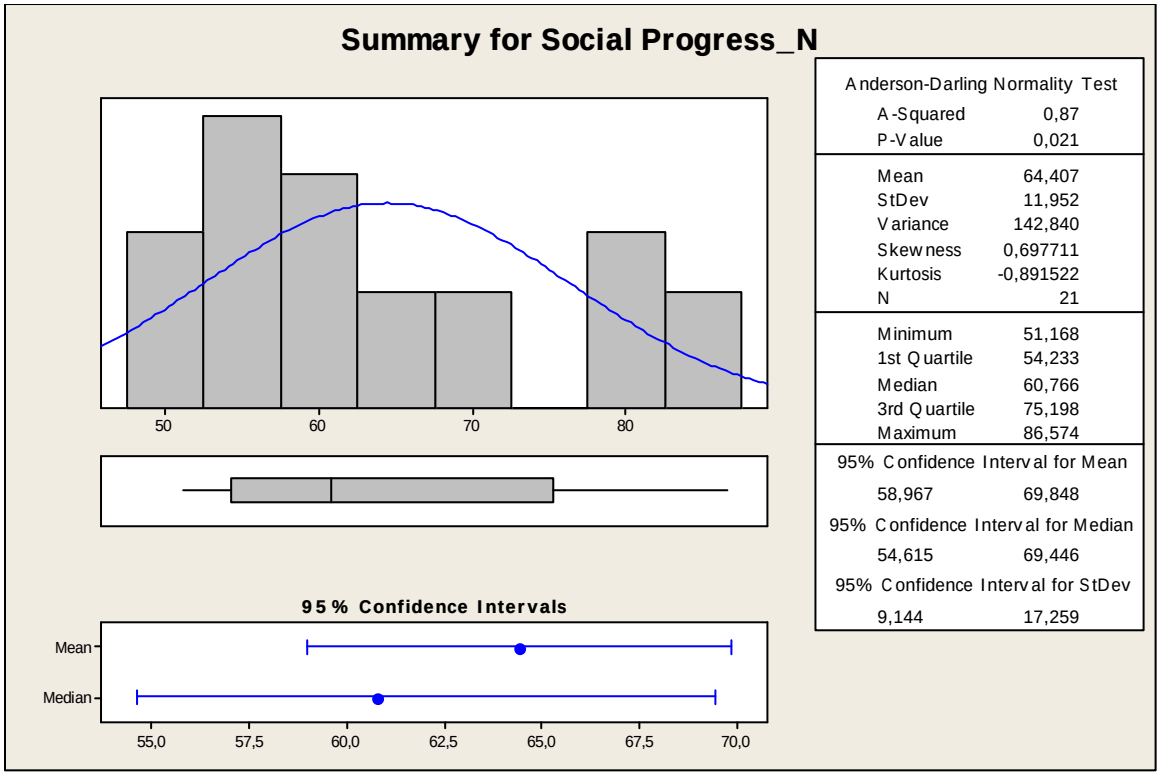
Tabela 1: Descriptive Statistics: Social Progr; Ecosystem Su; IDH - 2013_N; ...

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Maximum
Social Progress_N	21	0	64,41	2,61	11,95	51,17	86,57
Ecosystem Sustainability	21	0	65,62	2,93	13,42	46,90	85,31
IDH - 2013_N	21	0	65,62	2,63	12,03	45,63	87,64
Combustible renewables_N	21	0	21,50	3,82	17,52	1,00	66,83
Suceptibilidad (%)_N_P	21	0	71,56	2,57	11,78	48,77	90,56
Global Innovation Index_	21	0	41,19	2,62	12,00	24,57	70,25
Institutions_N	21	0	47,34	3,95	18,10	0,00	75,74
Human Capital and Resear	21	0	38,24	3,65	16,74	5,56	75,83
Infraestructure_N	21	0	44,55	3,61	16,53	22,81	85,34
Market Sophistication_N	21	0	32,98	2,99	13,69	0,00	64,76
Creative Outputs_N_1	21	0	49,33	2,26	10,38	31,30	67,79
International Cooperatio	21	0	68,23	4,76	21,83	16,25	100,00
Effective use of suppor_	21	0	70,68	4,25	19,46	22,22	100,00
Regional cooperation_N	21	0	69,92	4,27	19,56	28,57	100,00
EPI Score_N	21	0	52,70	3,12	14,31	30,07	88,62
Happy Planet Index_N	21	0	72,36	3,55	16,25	38,81	100,00
GINNI Index_N_P	21	0	43,08	3,14	14,40	14,98	71,48
Knowledge_N_R_N	21	0	36,78	2,67	12,22	16,58	69,57

Análise: Selecionamos o Índice de Progresso Social para esse comentário que mostra a Média de 64,41 em uma escala até 100, indicando que de uma maneira geral os países estão dentro da escala normal. Embora o valor mínimo observado é de 51,17 para Cuba país socialista com embargo de importação pelos Estados Unidos até o início de 2016. E o valor máximo observado foi 86,57 para a Espanha que é localizada na Europa e com desenvolvimento.

Devido a Espanha ter uma característica de economia desenvolvida esse indicador confirma essa qualidade.

4.1 SPI – Índice de Progresso Social



Análise:

Variável: Social Progress Index – índice de Progresso Social

Origem e definição dado pelo portal: http://pt.knoema.com/jklbfre/the-social-progress-index

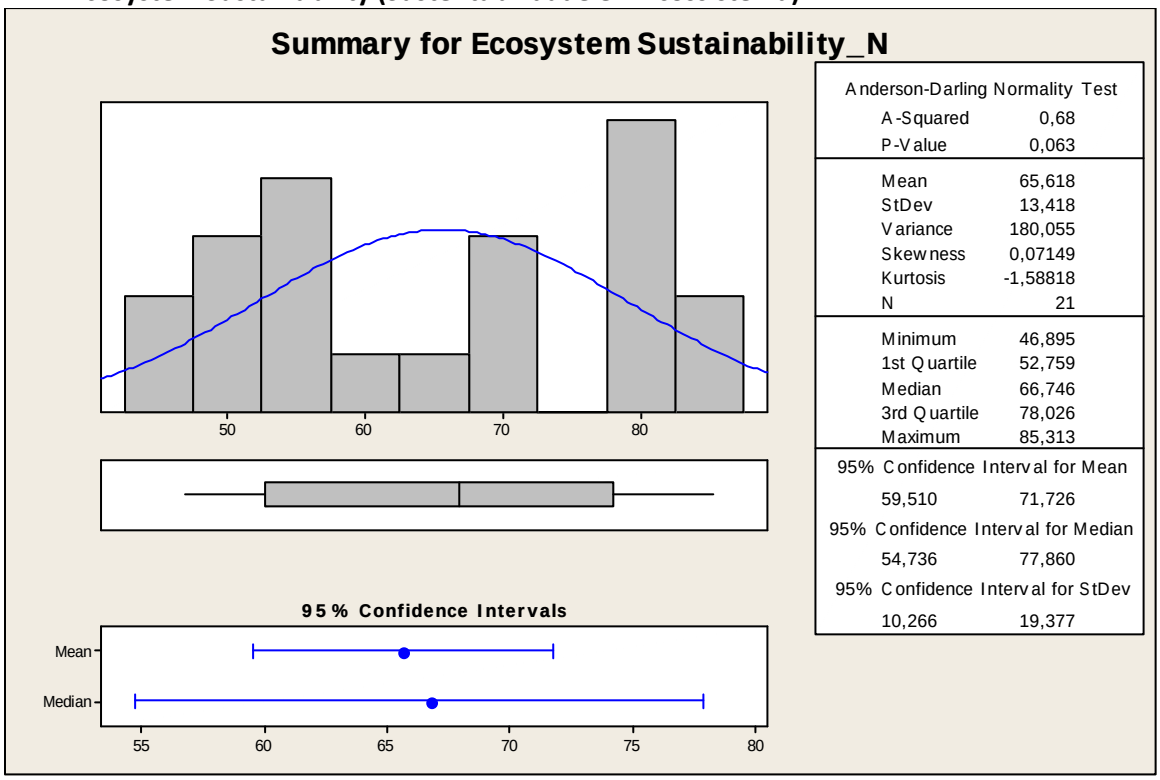
O Índice de Progresso Social é o resultado de um processo de dois anos de pesquisas envolvendo uma equipe renomada de especialistas que inclui, entre outros, os economistas Hernando de Soto e Michael Porter e do presidente da Fundação Rockefeller Dr. Judith Rodin. O índice sintetiza um total extenso de dados pesquisados para identificar as dimensões do desempenho das sociedades e medir o progresso social de forma abrangente e rigorosa. O Índice foi estruturado em torno de 12 componentes e 54 indicadores distintos consolidadas em três dimensões do Progresso Social: necessidades humanas básicas, Fundações de Bem-estar e oportunidade. A primeira dimensão, necessidades humanas básicas, avalia o quão bem um país prevê necessidades essenciais de seu povo através da medição se as pessoas têm comida suficiente e estão recebendo cuidados médicos básicos, se tiverem acesso à água potável, se tiverem acesso a uma habitação condigna com serviços básicos, e se eles são seguros e protegidos.

Análise dos dados da nossa pesquisa mostra que a distribuição aproxima-se da normal, considerando os 21 países.

Percebe-se um número menor de países do que o esperado classificados entre 65 e 75 com apenas 3 observações. O maior número de observações encontra-se na faixa de 55 a 65, com 6 observações. Esta faixa corresponde a última faixa imediatamente inferior à mediana e corresponde a um nível de desenvolvimento considerado "médio" pelo índice. A mediana de 64,4 e o terceiro quartil de 75,2 indicam o poder discricionário do índice, que consegue

separar o grupo de países em 5 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados.

4.2 – Ecosystem Sustainability (Sustentabilidade em Ecossistema)



Análise:

Variável: Ecosystem Sustainability - Sustentabilidade em Ecossistema

A sustentabilidade do ecossistema se correlaciona e integra de forma organizada os aspectos econômicos, sociais, culturais e ambientais da sociedade com relação aos ecossistema. O ecossistema pode ser definido pelo conjunto de seres vivos e do meio ambiente em que eles vivem, e todas as interações desses organismos com o meio e entre si. (http://www.blogfreeway.com.br/o-que-sao-ecossistemas-sustentabilidade-agenda-21-cidades-sustentaveis-aquecimento-global-produto-biodegradavel)

A manutenção de um ecossistema sustentável depende da interação de 4 elementos: Produtividade, Diversidade, Resiliência, Distúrbios Naturais.

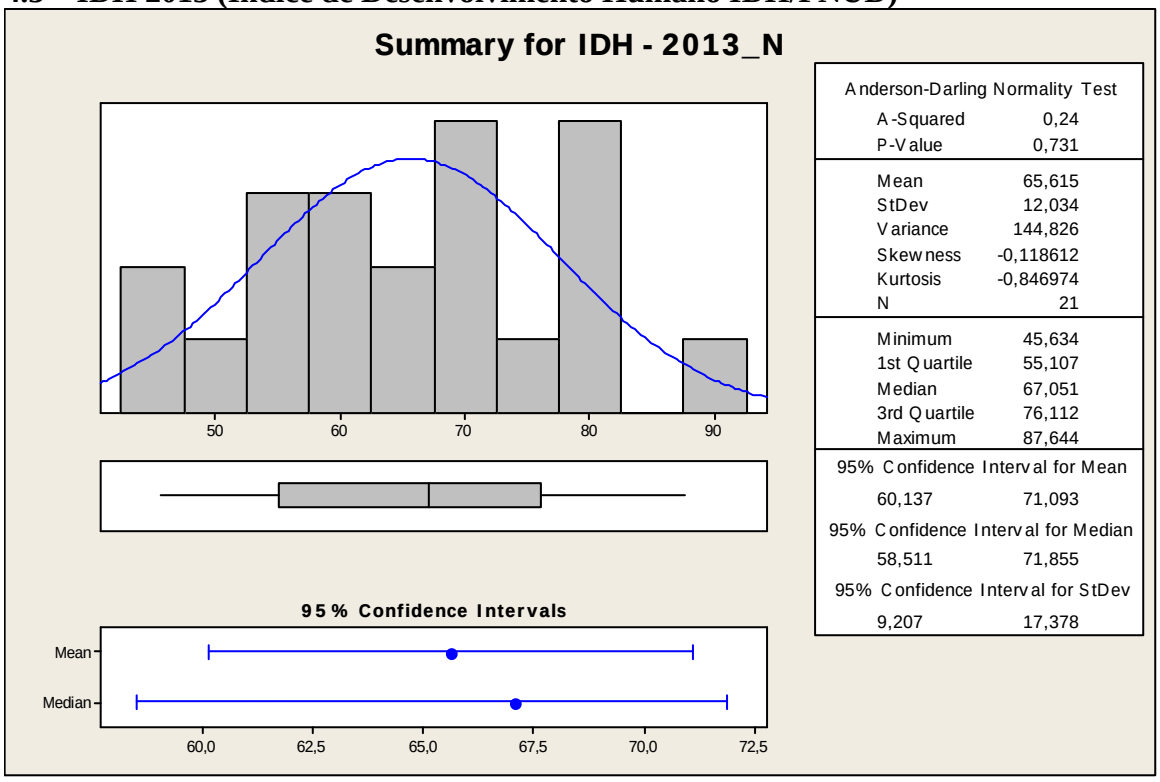
A sustentabilidade dos ecossistemas é determinada pela relação entre esses elementos, e juntos eles determinam o nível de recursos que podem ser tomadas a partir de um ambiente e ainda mantê-lo de forma sustentável.

A produtividade é a taxa de crescimento de cada parte viva de um ecossistema; os animais, plantas e outros organismos que compõem um ecossistema. Produtividade afeta o nível básico de recursos de um ecossistema pode proporcionar. Diversidade, o número de habitats, os tipos de espécies e a quantidade de variabilidade genética também influenciam a produtividade.
Uma quantidade moderada de distúrbios naturais, como tempestades ou outros eventos naturais, também aumenta a produtividade, mas dependendo do tipo do ecossistema, pode acarretar diminuição da produtividade (e assim a sustentabilidade). A preocupação em adicionar taxa de distúrbio natural é devido a um sistema poder tornar-se insustentável caso o nível combinado de distúrbios naturais tornar-se muito grande. Resiliência é a capacidade de um ecossistema de resistir e

se recuperar de eventos de distúrbios naturais e é influenciada tanto pela produtividade e diversidade.

Análise dos dados da nossa pesquisa mostra que a distribuição aproxima-se da normal, considerando os 21 países. Percebe-se a média e mediana estao muito proximas e que 11 países estão acima da média e somente 5 países apresentam-se no 3.quartil, isto é com índice superior a 78, indicando o poder discricionário do índice.

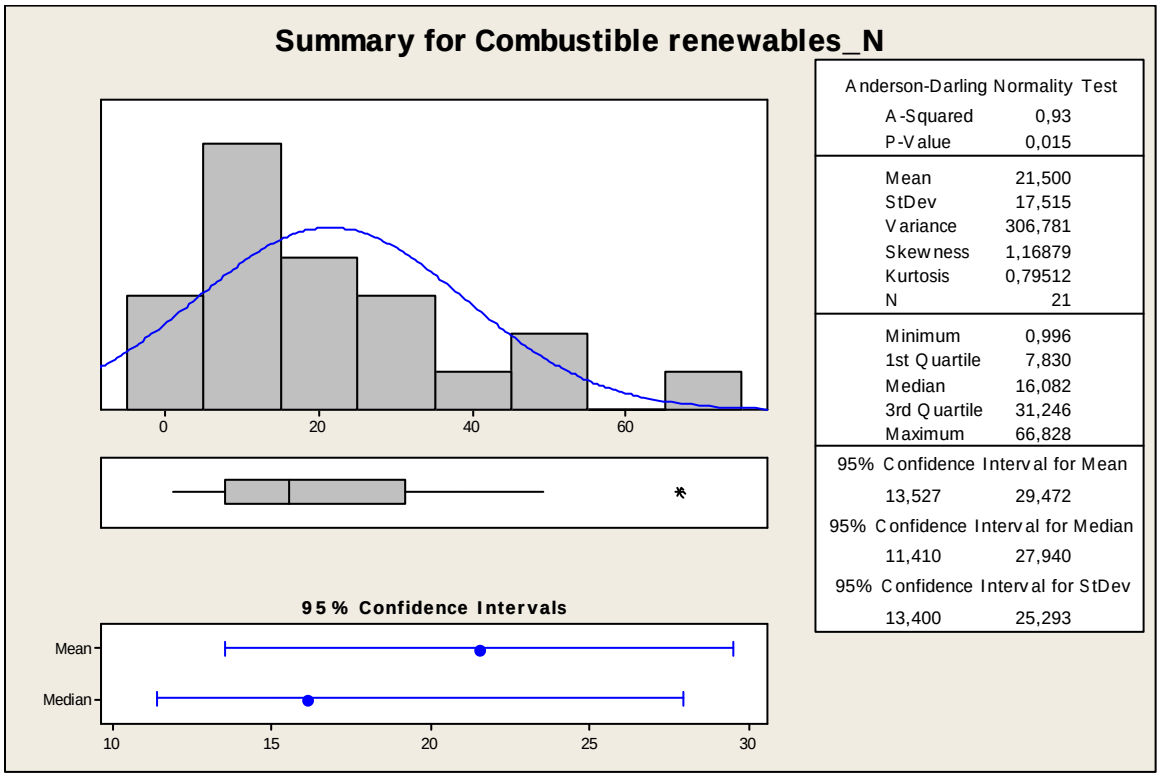
4.3 – IDH 2013 (Índice de Desenvolvimento Humano IDH/PNUD)



O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.
O trabalho publicado preliminar publicado por PNUD - Programa de las Naciones Unidas para el Desarrollo sobre o desenvolvimento humano em 2014 aponta que o IDH é progresso real da população em relação a expectativa de vida, educação, saúde, habitação, segurança e condições para um desenvolvimento sustentável.
http://hdr.undp.org/sites/default/files/hdr14-summary-es.pdf

Análise dos dados da nossa pesquisa: O gráfico para o IDH aproxima-se de uma distribuição normal, com uma concentração maior de países dos 52 aos 75 pontos, totalizando 13 países. Isso pode ser constatado também pela mediana, em 67,05. No estrato do 55 a 65 pontos, encontra-se o Paraguai, com IDH de 55,85 pontos. O Brasil está no 12. Lugar com indice de 67,05, igual a mediana.

4.4 – Combustible Renewables and Wast (Combustíveis Renováveis e Resíduos)

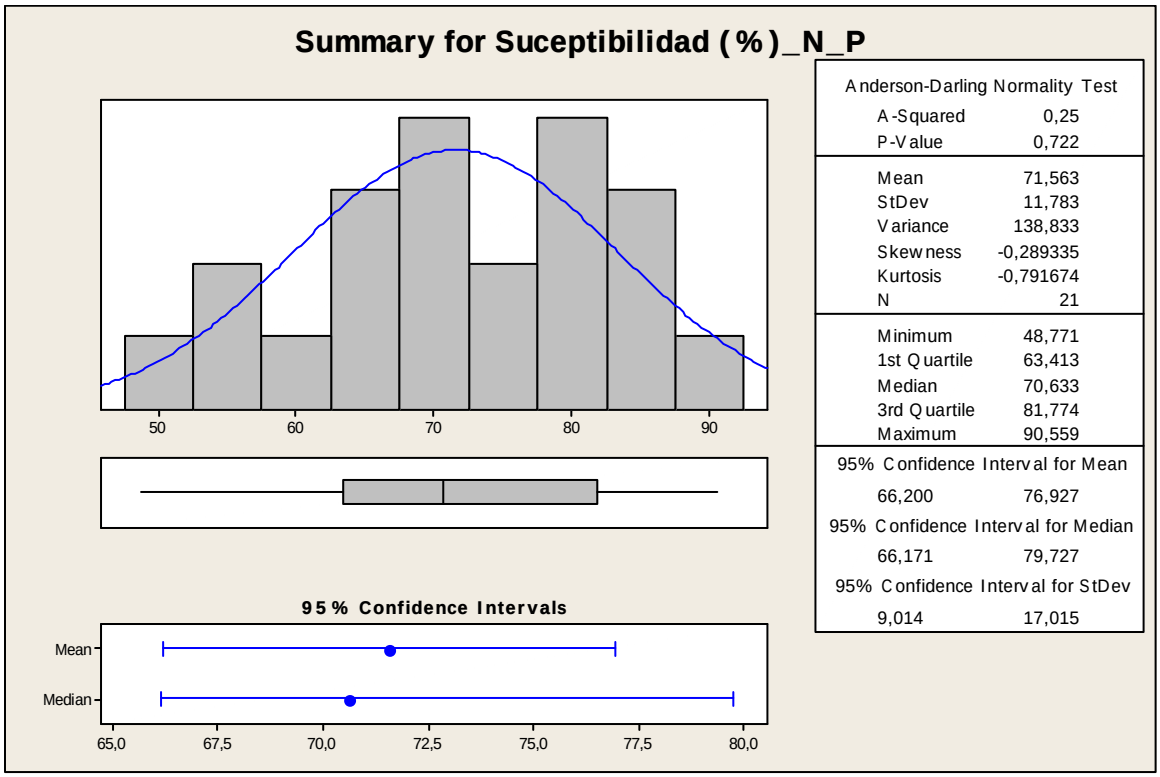


Definição: combustíveis renováveis e resíduos compreendem a biomassa sólida, biomassa líquida, biogás, resíduos industriais e resíduos urbanos, medido em percentagem do consumo total de energia. Os combustíveis renováveis e resíduos (% da energia total) no mundo foi de 9,81 a partir de 2011. O seu valor mais alto nos últimos 51 anos foi de 12,70 em 1983, enquanto o seu valor mais baixo foi de 2,12 em 1969.

Fonte: Agência Internacional de Energia (IEA Estatísticas © OECD / IEA, http://www.iea.org/stats/index.asp

Análise dos dados da nossa pesquisa: O resultado da amostra apresenta assimetria positiva ou à direita, com média de 21,5 e mediana de 16,082. A amostra representada por 21 países, somente 7 países apresentam-se acima da média. O terceiro quartil foi calculado em 31,24, que está bem distante do valor máximo de 100, indicando que poucos países têm altos índices de combustíveis renováveis e resíduos. A Venezuela, grande produtor de petróleo, teve índice igual a 1, seguido pela Argentina e México, ambos com índice igual a 4.

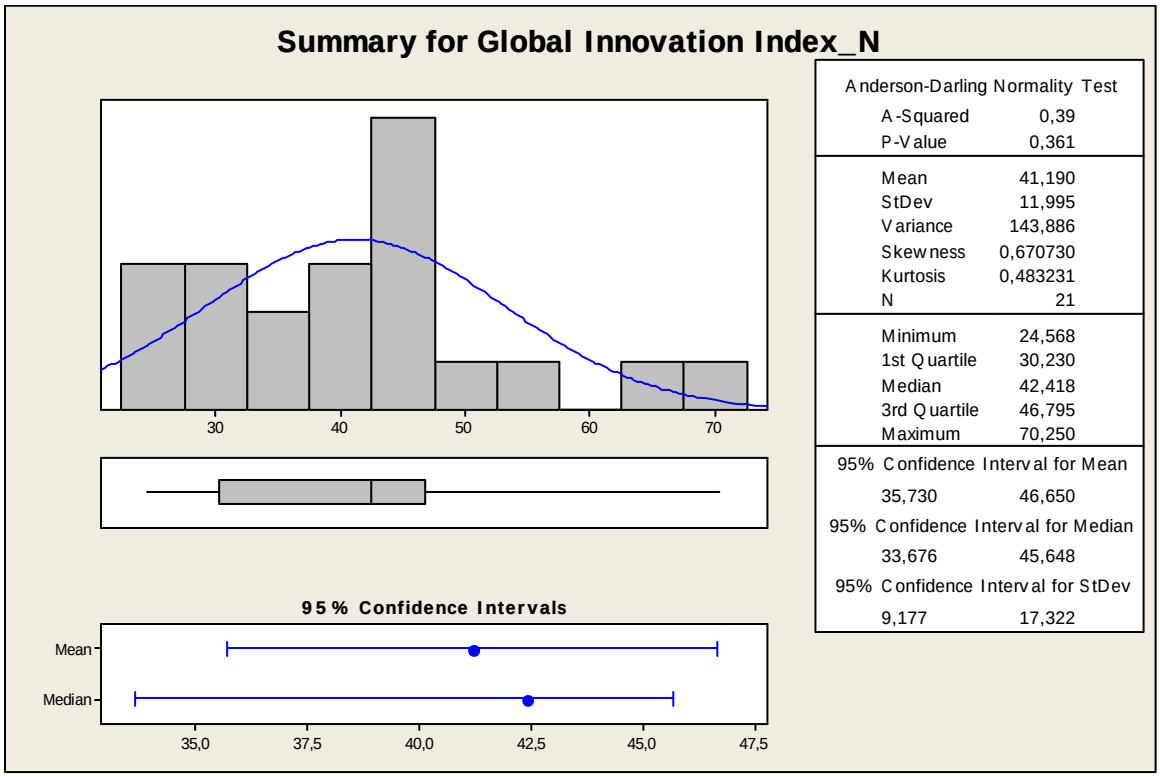
4.5 - Suceptibilidad (Sucetibilidade, Sensibilidade)



Suscetibilidade refere-se à probabilidade de sofrer dano, perda e interrupção de um evento extremo ou desastres naturais. Assim, a susceptibilidade descreve as características estruturais e condições de enquadramento de uma sociedade.Suscetibilidade refere-se à probabilidade de sofrer dano, perda e interrupção de um evento extremo ou desastres naturais. Assim, a susceptibilidade descreve as características estruturais e condições de enquadramento de uma sociedade.

Análise dos dados da nossa pesquisa – os dados do gráfico acima estão positivados, então a leitura dos dados deve considerar que quanto mais próximo de 100% menor será a probabilidade de ocorrência de eventos naturais que prejudiquem o país. O índice médio de 71,5 apresenta 10 países acima da média. No primeiro quartil tem-se 6 países, Bolívia, Guatemala, Nicarágua, Honduras, Paraguai e El Salvador, em ordem crescente, isto é o pior índice da amostra foi apresentado pela Bolivia. Observa-se que 10 países encontram-se acima da mediana, apesar da diferença de 42 pontos entre o índice mínimo de índice máximo. Os países com melhor desempenho nesta variável são Portugal e Espanha.

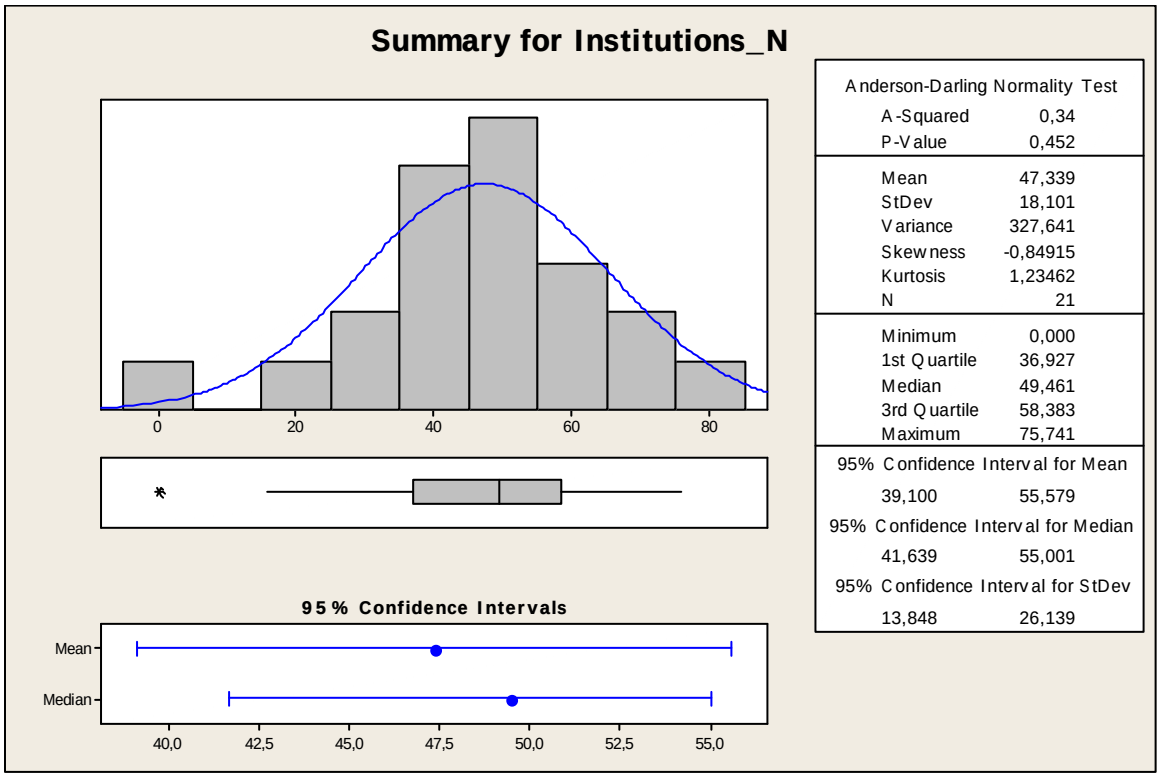
4.6 Global Innovation Ind (inovação industrial global)



O Índice Global de Inovação (GII) reconhece o papel fundamental da inovação como motor do crescimento económico e prosperidade, e reconhece a necessidade de uma ampla visão horizontal de inovação, que é aplicável a ambas as economias: desenvolvidas e emergentes, com a inclusão de indicadores que vão além das medidas tradicionais de inovação (como o nível de pesquisa e desenvolvimento em um determinado país). O GII evoluiu para uma ferramenta de benchmarking valioso para facilitar o diálogo público-privada, em que os responsáveis políticos, líderes empresariais e outros interessados podem avaliar o progresso em uma base contínua.

Análise dos dados da nossa pesquisa: observa-se que 12 países estão acima da mediana, todavia o maior índice apresentado foi pela Espanha com 70,1, seguido de Portugal com 63,1. Ambos estão ainda distantes de um nível ótimo de Inovação próximo a 100. Desta forma pode-se afirmar que 9 países estão abaixo da média para a amostra, média essa considerada baixa de 41,1 pontos. Os menores índices de inovação foram encontrados em: Nicarágua, Venezuela, Honduras, Equador e Bolivia. O Brasil ocupa a 7. posição no ranking com índice de 45.

4.7 - Institutions (Instituições)



Capta o quadro institucional de um país, por meio de suas instituições, medindo a capacidade de atrair negócios e promover o crescimento, fornecendo boa governança e níveis adequados de proteção e incentivo para inovações. O índice é formado por três pilares: a) **Ambiente Político:** indica a percepção da probabilidade de que um país possa ser desestabilizado; da qualidade dos serviços públicos e civis; da formulação e implementação de políticas; e a percepção de violações à liberdade de imprensa; b) **Ambiente Regulatório**; indica a percepção sobre a capacidade de um governo de formular e implementar suas políticas de coesão que promovam o desenvolvimento do setor privado; a prevalência do Estado de Direito; e avalia os custos de indenização por demissão de trabalhador; e c) **Ambiente de Negócios**: reflete a avaliação de três aspectos que afetam diretamente os empreendimentos privados, a facilidade de começar um negócio, a facilidade de solução de insolvência e a facilidade de pagamento de impostos.

Análise: A distribuição é normal, com média dos 21 países de 47,33, com desvio padrão de 18,1 e variância de 327. Cuba ocupou a 6. Posição com 55,15 pontos. Na tabela abaixo listou-se os países que se encontram no extremo superior e extremo inferior do gráfico:

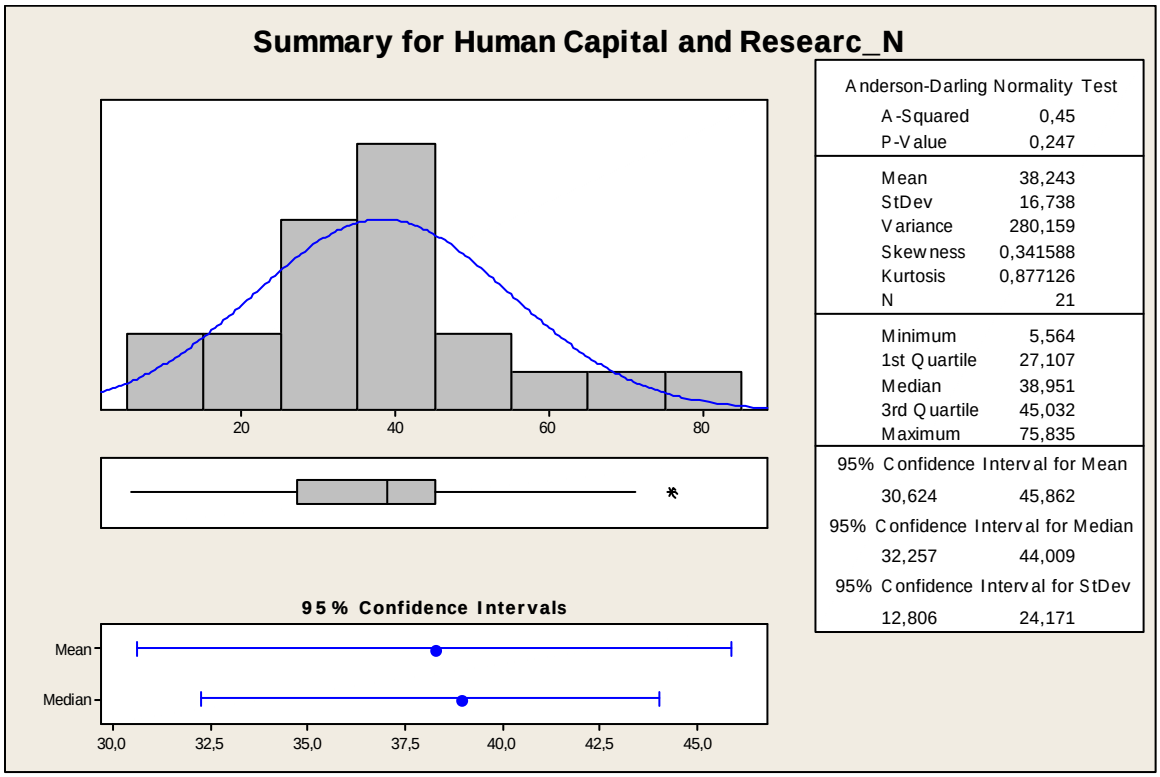
OS 10 PRIMEIROS COLOCADOS

Ranking	País	Índice
1	Portugal	75,74
2	Spain	72,37
3	Chile	68,19
4	Uruguay	64,15
5	Costa Rica	61,46

OS 10 ULTIMOS COLOCADOS

Ranking	País	Índice
17	Paraguay	36,12
18	Honduras	33,69
19	Ecuador	30,32
20	Bolívia	15,63
21	Venezuela	0,00

4.8 - Human Capital and Research (Recursos e Capital Humano)



O Índice de Capital Humano é uma nova medida para a captura e rastrear o estado de desenvolvimento do capital humano ao redor do mundo. Ele tem três características fundamentais:

. Em primeiro lugar, o índice mede um conjunto mais amplo de indicadores do que as definições tradicionais de capital humano.

. Em segundo lugar, o índice tem uma abordagem de longo prazo para capital humano. Além de fornecer um instantâneo do estado do capital humano de um país hoje procura desenvolver uma forte consciência em torno da necessidade de planejamento resultantes de práticas e decisões políticas afetando as crianças de hoje e que irão moldar a força de trabalho futuro.

. O Índice inclui, assim, medidas, indicando a qualidade de primeira infância. Além disso, o Índice de captura a extensão para que os investimentos feitos em anos anteriores na saúde e educação estão sendo realizados na população em idade ativa através da aprendizagem e formação ao longo da vida. Finalmente, na outra extremidade do contínuo, o índice leva em conta a saúde e produtividade da população idosa.

Então podemos resumir o Índice de Capital Humano e Pesquisa com base nos quatro pilares do Índice, que são:

• O pilar Educação contém indicadores relativos à aspectos quantitativos e qualitativos da educação do outro lado primário, secundário e terciário e contém informações tanto no presente como força de trabalho bem como a futura força de trabalho.

• A Saúde e Bem-Estar pilar contém indicadores relativa ao bem-estar físico e mental de uma população sendo, desde a infância até a idade adulta.

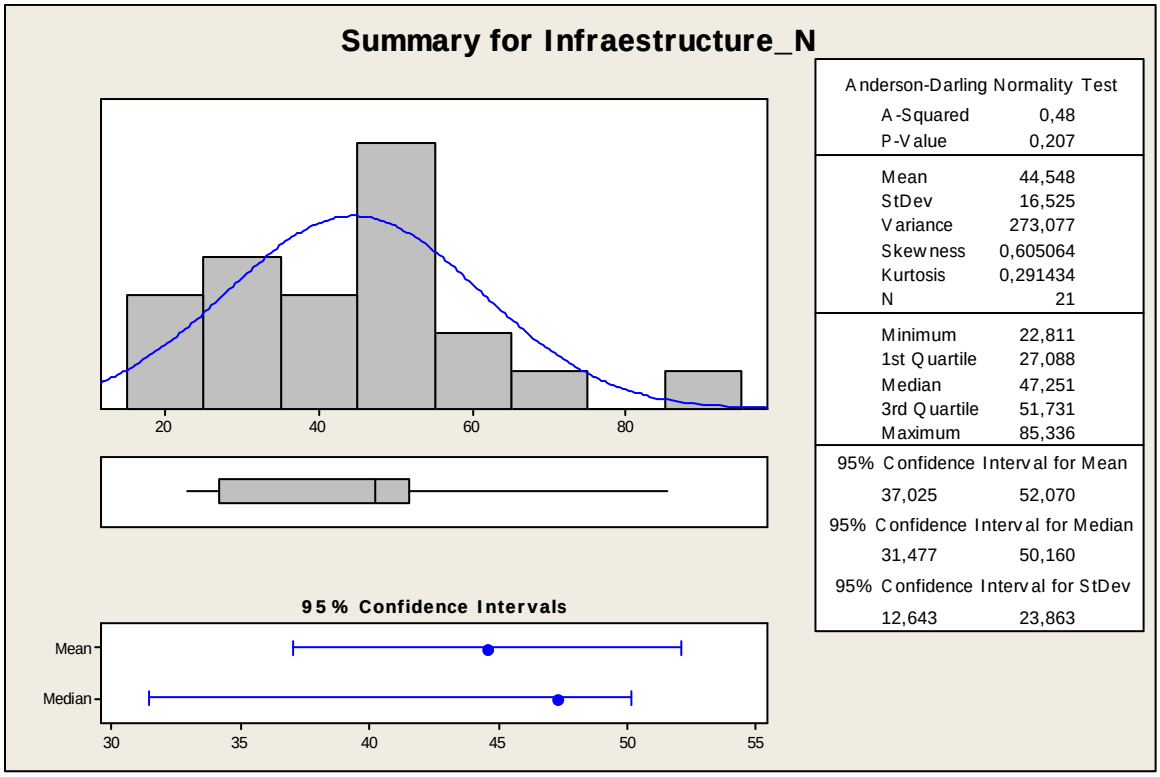
• O pilar da Força de Trabalho e Emprego destina-se a quantificar a experiência, talento, conhecimento e treinamento na população em idade ativa de um país.

• O pilar Ambiente propício capta o legal quadro, infra-estrutura e outros fatores que possibilitam retornos sobre o capital humano.

http://www3.weforum.org/docs/WEF_HumanCapitalReport_2013.pdf

Análise dos dados da nossa pesquisa: Pode-se observar que a média de 38,24 apresenta 11 países acima da média e que há uma enorme disparidade entre o máximo de 75,8 entre o mínimo de 5,5. Da amostra, 5países se encontram no primeiro quartil: Honduras, El Salvador, Guatemala, Nicarágua e República Dominicana, em ordem decrescente. Portugal e Espanha estão praticamente empatados em 1. Lugar no Ranking com 75 e 71, respectivamente. Todavia logo na sequencia vem a Argentina com 55 pontos.

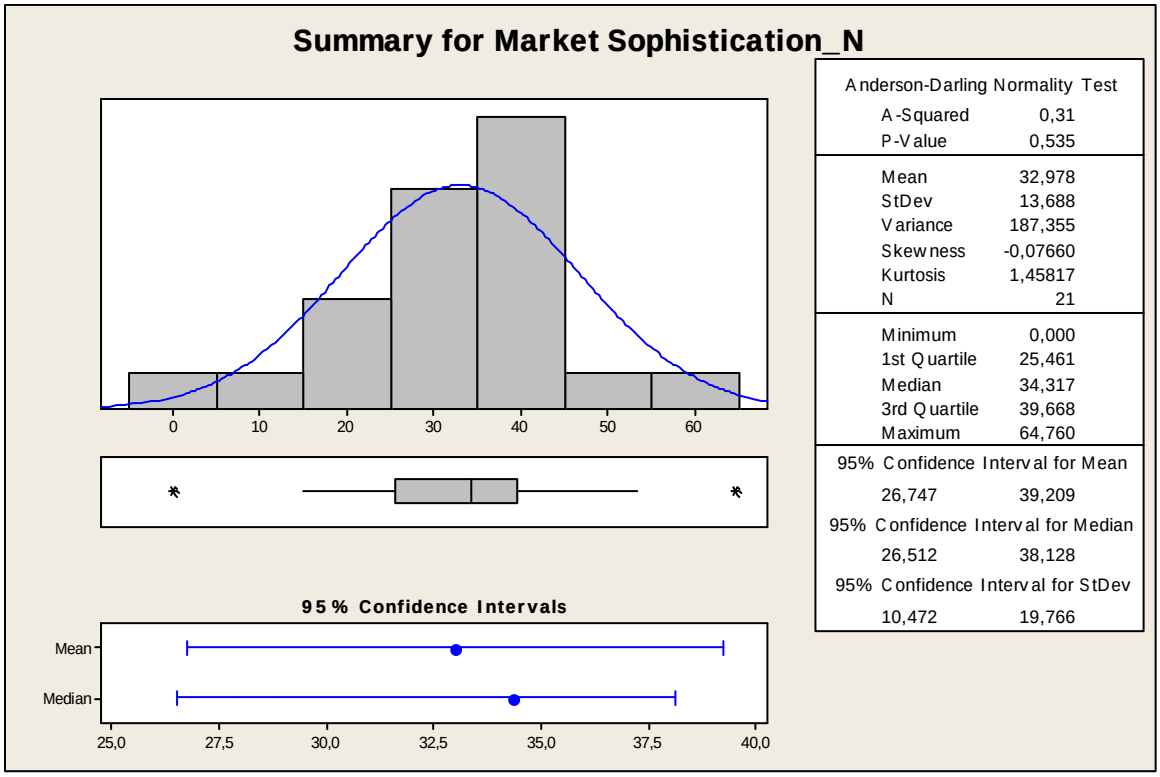
4.9 - Infraestructure (Infraestrutura)



Permite determinar o sucesso de fabricação e atividades agrícolas, bem como investimentos em água, saneamento, energia, habitação e transportes para melhorar a vida e ajudar a reduzir a pobreza. Determina ainda, as novas tecnologias de informação e comunicação para promover o crescimento, melhorar a prestação de saúde e outros serviços, ampliar o alcance da educação, e apoiar os avanços sociais e culturais. Dados aqui são compilados a partir de fontes tais como a Federação Internacional Road, Containerisation International, a Organização da Aviação Civil Internacional, a Associação Internacional de Energia, e da União Internacional das Telecomunicações.
http://data.worldbank.org/topic/infrastructure

Análise dos dados da nossa pesquisa: Os índices apontam para uma significativa parcela dos países (8 países) estarem próximo da média 44 e 55, indicado pelo primeiro quartil de 25,56 que está relativamente próximo à média de 44,877. Essa característica pode ser confirmada pela mediana de 44,633 estar muito próxima da média. Espanha e Chile ficam com a primeira e segunda posição no ranking com valores bem distintos, 85,3 e 68,0 respectivamente. O Brasil ficou em 7. Lugar com índice de 49,7.

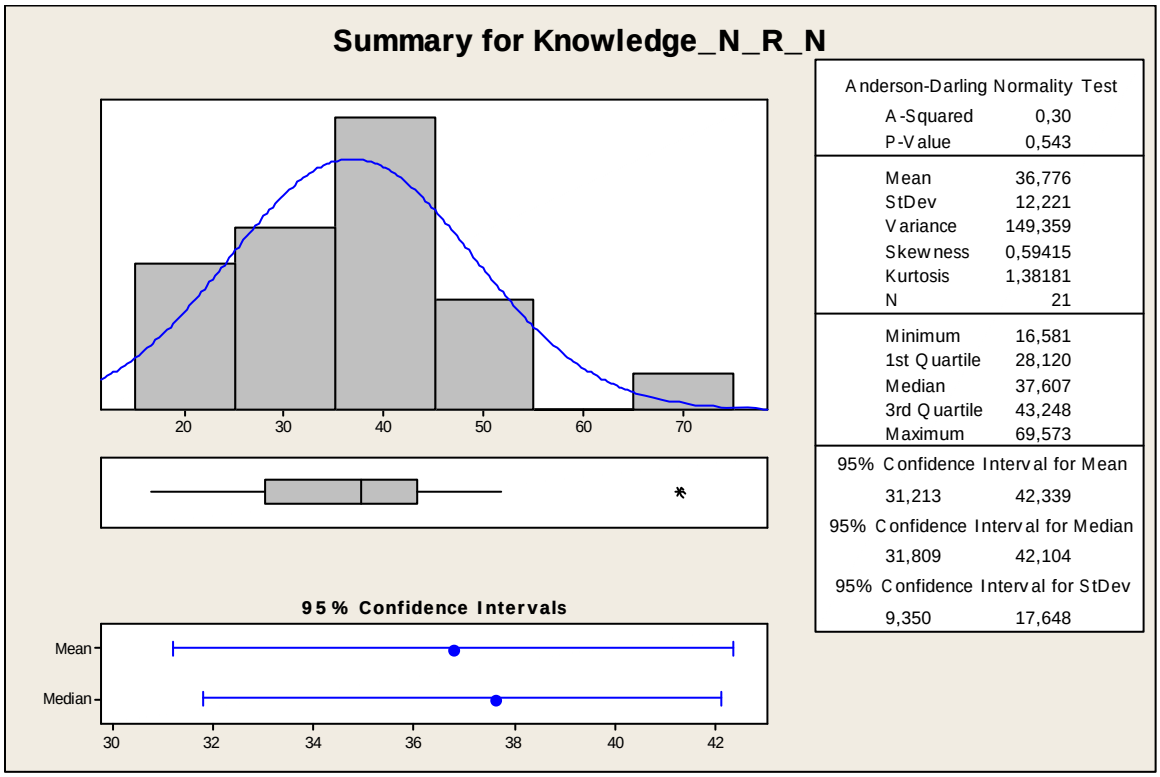
4.10 - Market Sophistication (Sofisticação de Mercado)



Este indicador é um derivado dos seguintes indicadores: - (a) disponibilidade de serviços financeiros (b) A acessibilidade dos serviços financeiros (c) financiamento através do mercado de ações local (d) Facilidade de acesso a empréstimos (e) a disponibilidade de capital de risco (f) restrição sobre os fluxos de capital (g) solidez dos bancos (h) a regulação das bolsas de valores e (i) índice de direitos legais.

Análise dos dados da nossa pesquisa: A mediana está próxima do valor da média, porém ligeiramente abaixo da média, indicando uma parcela significativa de 12 países acima da média. Todavia o primeiro quartil apresentando índice de 25,461, relativamente distante da média de 32,978 . O Brasil apresenta o 14. lugar no Ranking entre 21 países, com índice de 28,782. Estão também no 1. Quartil os seguintes países da Região AIBER : Costa Rica, Uruguai, Argentina e Venezuela, esta última com índice igual a zero.

4.11 - Knowledge and Technology (Conhecimento e Tecnologia)



Reflete todas as variáveis tradicionais direcionadas para a atividade de invenções e inovações, abrangendo:

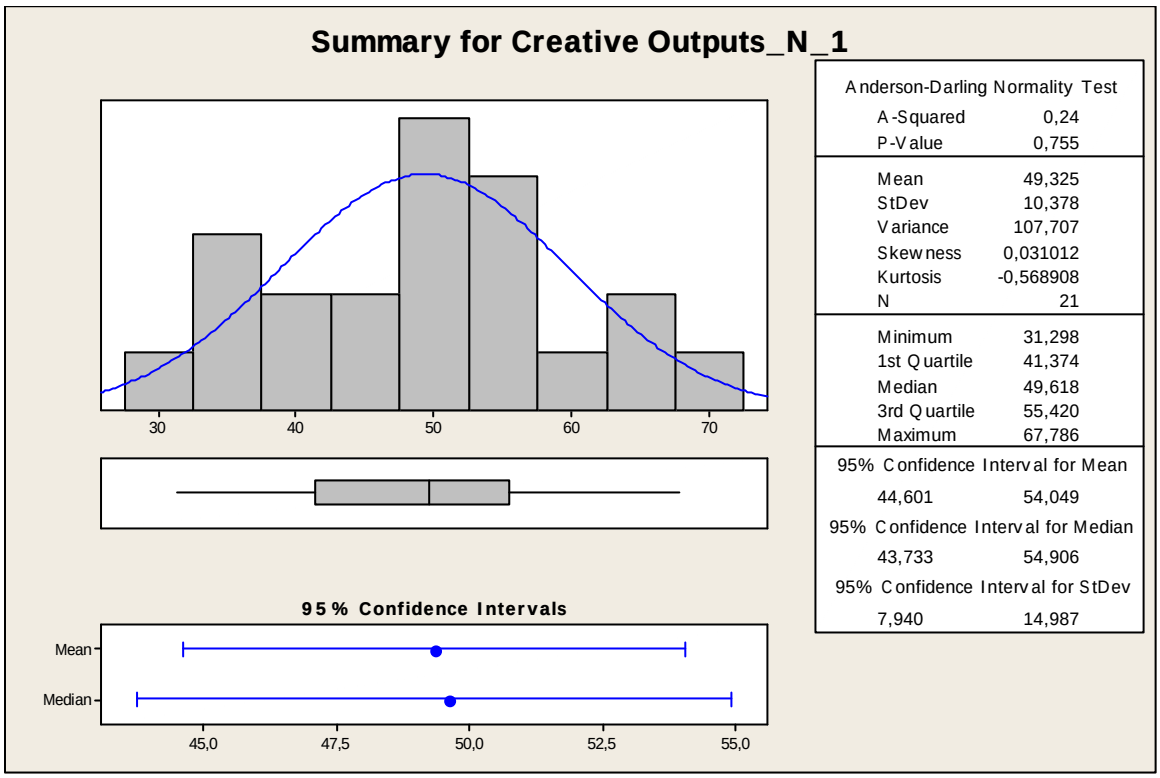
a) Criação do Conhecimento que envolve: o resultado de Atividades Criativas e Inovadores; pedidos de patentes; Aplicação de modelos de utilidade; artigos científicos e técnicos públicos; e quantidade de citações dos artigos publicados;

b) Sobre o Impacto do Conhecimento, envolvendo: estatísticas que representem o impacto das atividade de invenções e inovações no contexto micro e macroeconômico, aumento da produtividade do trabalho, a entrada de novas empresas, os gastos com software e quantidade de certificação de qualidade ISSO 9001; e

c) Difusão do Conhecimento medida por estatísticas ligadas ao setor de alta tecnologia ou que são fundamentais para a inovação, envolvendo quadro aspectos: royalties e taxas de licenças; percentual de exportações de alta tecnologia em relação ao total de exportações; percentual das exportações de serviços de comunicação, computação e informação em relação ao total de exportação de serviços; e o percentual de saída de FDI (Foreign Direct Investment) em relação ao PIB (Produto Interno Bruto).

Análise dos dados da nossa pesquisa: A máxima para esta variável foi considerada baixa 69,5 e foi alcaçada pela Espanha. Somente 5 países estão no 3. Quartil : Brasil, Cuba, Costa Rica, Portugal e Espanha. A média dos 21 países é de 36,7, considerada baixa e com desvio padrão de 12,22. A Venezuela se acomete na 11. Posição.

4.12 - Creative Outputs (Soluções Criativas)



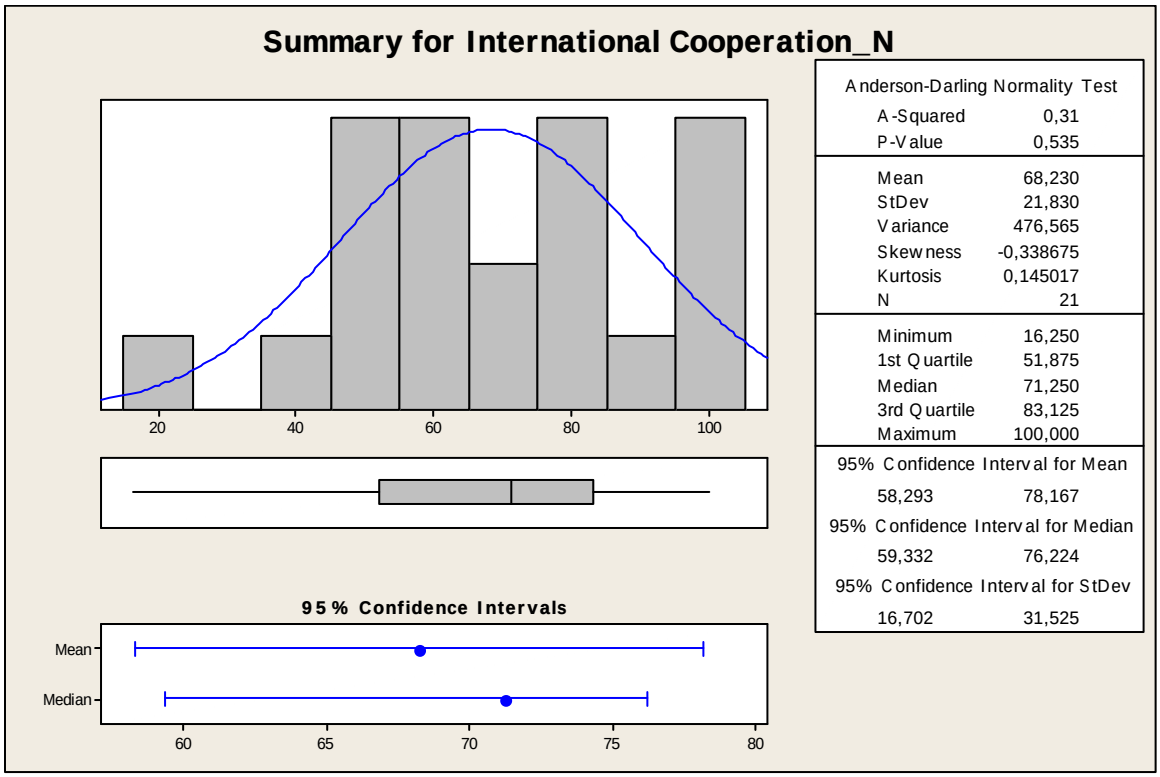
A criatividade foi reconhecida como um motor econômico para gerar riqueza, emprego, o desenvolvimento sustentável das cidades do mundo, mudanças tecnológicas, inovação empresarial e reforço da competitividade das cidades e países individuais.

Criatividade mostrada na produção econômica e nos domínios culturais, por exemplo, pode ser avaliado e apreciado pelos a comunidade, enquanto a inovação criativa em ciência ou ciência humana poderia ser reconhecida pelas comunidades científicas e acadêmicas. Isto significa "criatividade social" é mensurável e quantificável, embora "a criatividade da vida diária" é uma capacidade geral encontrada em outro lugar em nossa vida mais difícil de medir.

Em 2002, Richard Florida, de acordo com sua "teoria do capital criativo", o centros da classe criativa são "mais prováveis de serem vencedores econômicos", tendo sucesso na geração de ^^opostos de trabalho high-end e o crescimento econômico, assim, aumentando ainda mais a vantagem regional da indivíduo em sua região. Sua teoria para explicar o crescimento econômico regional constrói em torno da "3Ts" - talento, tecnologia e tolerância - que argumenta que pessoas criativas preferem lugares "que são diversos, tolerante e aberto a novas ideia ", e que a presença e concentração do capital criativo uma região "por sua vez levam a maiores taxas de inovação, alta tecnologia formação de negócios, geração de emprego e crescimento económico ". http://www.uis.unesco.org/culture/Documents/Hui.pdf

Análise dos dados da nossa pesquisa: Apesar da média e da mediana estarem próximas, indicando que praticamente metade dos 21 países analisados estão acima da média. Todavia, percebe-se uma certa hegemonia entre os países da região AIBER que tem máxima de 67 e mínima de 31. Os 3 primeiros países do ranking são: 1º. Panamá, 2.Portugal e 3.Espanha. Os três piores classificados foram: Honduras, Nicaragua, Venezuela e Bolivia, todos próximos de 31.

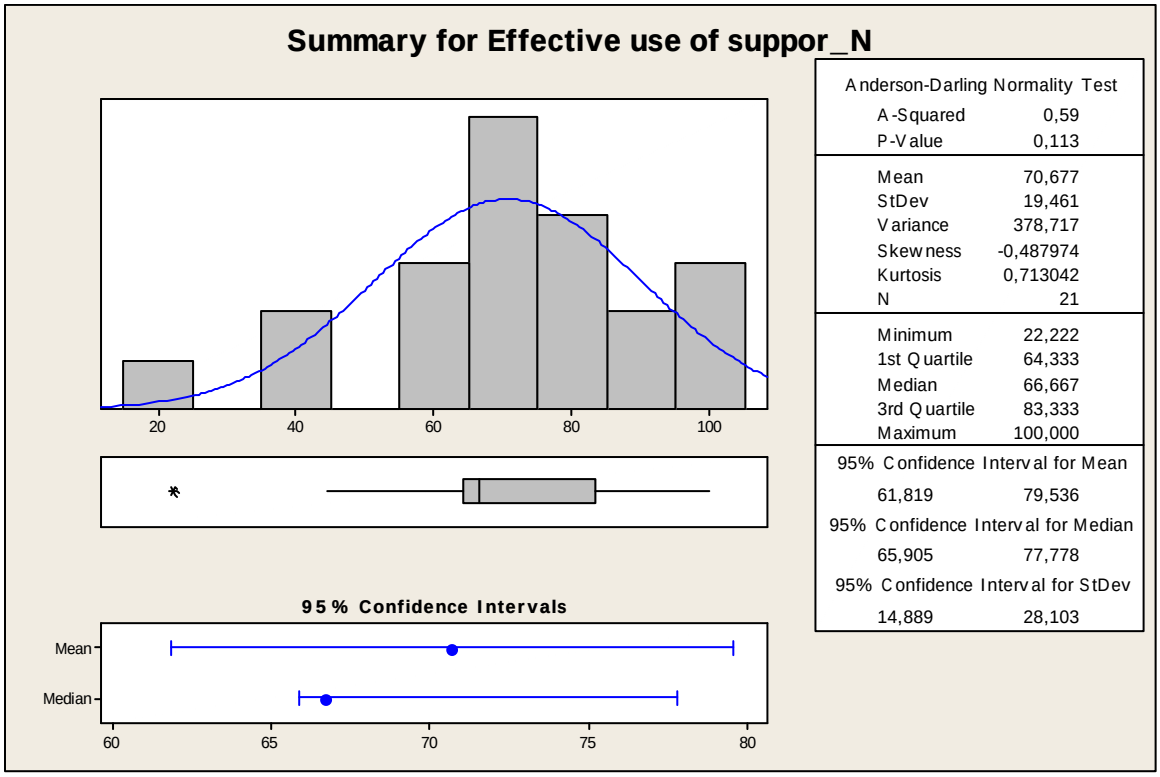
4.13 - International Cooperation (Cooperação Internacional)



Esta questão visa avaliar a disposição e capacidade da liderança política em desenvolver boas relações de vizinhança e cooperar com os países vizinhos em organizações internacionais e regionais • apoiar a integração regional ou internacional

Análise dos dados da nossa pesquisa: Apesar da média e da mediana estarem próximas, indicando que praticamente metade dos 21 países analisados estão acima da média, alguns países estão muito longe de apresentarem bons índices de Cooperação Internacional, onde a 21ª. Posição é ocupada pela Venezuela que apresenta índice de 16,25, bem distante do próximo colocado em 20º. Lugar Cuba com índice de 41,2. Enquanto que os países Brazil, Chile, Uruguai, apresentam índice máximo 100.

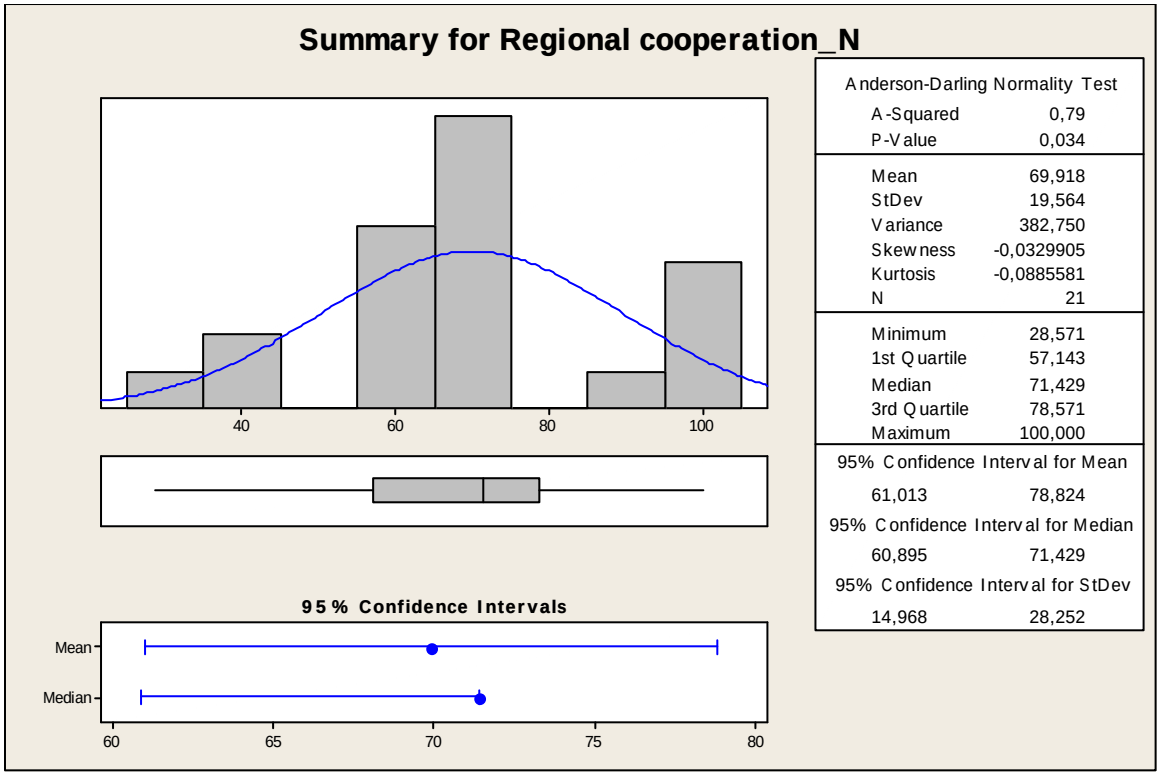
4.14 - Effective use of support (eficácia na utilização de apoio)



Este indicador mede se o Governo tem definidos de forma clara os objetivos políticos e econômicos e se consegue utilizar know-how de grandes projetos assistenciais internacionais, trazendo esses projetos para a realidade do seu país.

Análise dos dados da nossa pesquisa: Apesar da média e da mediana estarem próximas, indicando que praticamente metade dos 21 países analisados estão acima da média, a Venezuela com índice de 22,22 está muito aquém dos demais países, que apresentaram razoáveis índices entre 44,44 alcançado por Cuba que é o 20º no ranking e índice de 100 dos países: Brasil, Chile, Uruguai.

4.15 - Regional Cooperation



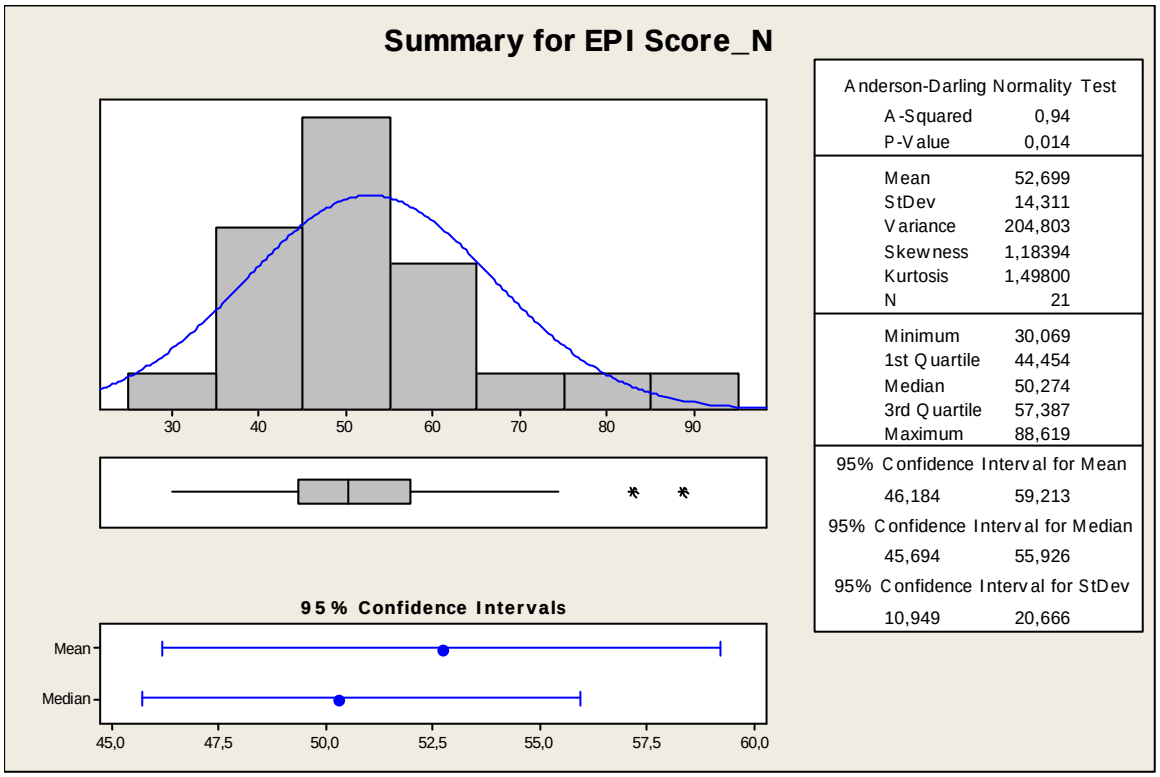
Esta questão visa avaliar a disposição e capacidade da liderança política:
· Desenvolver relações de boa vizinhança
· Cooperar com os países vizinhos em organizações internacionais e regionais
· Apoiar a integração regional ou internacional.

A liderança política ativamente e com êxito constrói e expande cooperativa vizinhança e internacional relacionamentos. Ela promove a integração regional e internacional.

A liderança política coopera com muitos estados vizinhos e está em conformidade com as regras estabelecidas pela organizações regionais e internacionais.

Análise dos dados da nossa pesquisa: A média é considerada alta 69,918, indicando que todos os países tiveram bons índices para essa variável. Os países que apresentam índice máximo 100, são: Brasil, Chile, El Salvador, Uruguai.

4.16 - Proteção Ambiental – EPI (*Environmental Protection Index*)



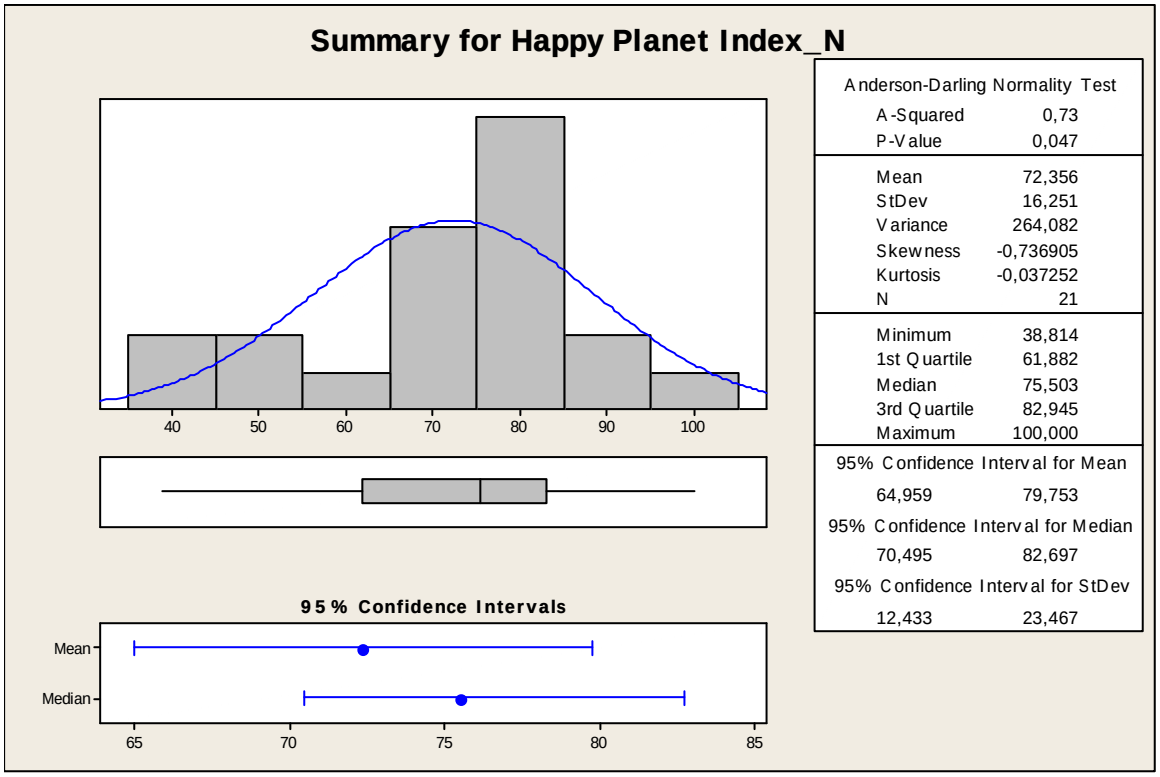
Anderson-Darling Normality Test	
A-Squared	0,94
P-Value	0,014
Mean	52,699
StDev	14,311
Variance	204,803
Skewness	1,18394
Kurtosis	1,49800
N	21
Minimum	30,069
1st Quartile	44,454
Median	50,274
3rd Quartile	57,387
Maximum	88,619
95% Confidence Interval for Mean	
46,184	59,213
95% Confidence Interval for Median	
45,694	55,926
95% Confidence Interval for StDev	
10,949	20,666

Unidade de medida utilizada proteção ambiental (EP) inclui todas as atividades intencionais diretamente com vista à prevenção, redução e eliminação da poluição ou qualquer outra degradação do ambiente resultante de processos de produção ou de consumo. O âmbito da proteção ambiental é definida de acordo com a classificação das atividades de proteção ambiental (CEPA), que distingue nove domínios ambientais diferentes.

Análise dos dados da nossa pesquisa: A análise dos dados do Índice de Proteção Ambiental apresenta uma distribuição normal, com média de 52,699 e mediana de 50,274. O Brasil ocupa a 12ª. Posição no Ranking com índice de 49,884, ligeiramente abaixo da média. O menor índice de 30,069 é do Paraguai.

4.17 - *Happy Planet Index (HPI)*

Reflete a felicidade de um país com base em três indicadores: Expectativa de Vida, Bem-estar e Pegada Ecológica (Ecological Footprint).



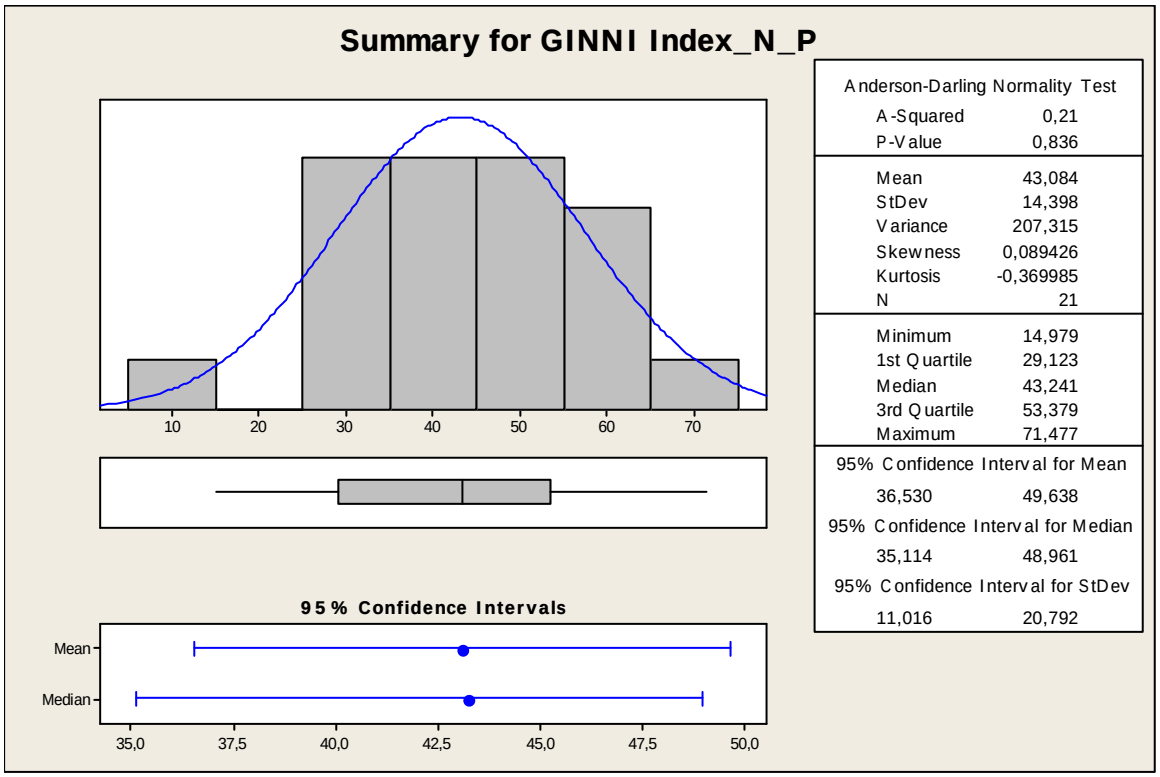
Análise dos dados da nossa pesquisa: A distribuição para o Índice de Felicidade Bruta (FIB) aproxima-se de uma normal, com maior concentração de países entre a marca do 65 até 85, que formam quase um plateau, com os dois maiores picos da distribuição. Nas extremidades encontramos um pequeno número de países : Costa Rica ficou um primeiro lugar, com um indice de 100, seguida por um estrato de notas normalizadas de 85 a 95, que é composto por três países : Colômbia (FIB=89,66), El Salvador (87,58) e `Panamá (84,95).

No estrato inferior, que vai até 50 pontos, figura Portugal = 38,81, seguida pelo segundo lugar Uruguay 40,37. Chama a atenção os paises da Europa Portugal e Espanha estarem entre 4 piores do ranking e Cuba em 9. Lugar.

O Índice de Felicidade Bruta mede aspectos subjetivos e objetivos da vida humana, evidenciando pesquisas que apontam que, a partir de certo nível de renda, o nível de felicidade médio reportado não aumenta na mesma proporção do aumento da renda, até decaindo com o tempo. Outro tópico relaciona é a questão da resiliência construída por pessoas em situação de privação em países de baixo IDH e IPS, que acabam adaptando-se às situações adversas.

Uma análise mais aprofundada poderá revelar o baixo nível de correlação deste indicados com o PIB e inclusive com o IDH e o IPS, pois sua metodologia é bastante distinta.

4.18 - Índice de Ginni



O Índice Ginni foi positivado, de forma que quanto maior o índice apresentado, menor é a desigualdade de renda e de classe social em um país.

Análise dos dados da nossa pesquisa: Para o Índice de Gini, que mede a desigualdade de renda, faltaram dados relativos a 13 países, que foram substituídos pela média da distribuição original, que foi de 39,02, que normalizado de 0 a 100.

Assim percebemos uma distribuição basicamente uniforme, com tendência para concentração de países mais a direita, como já demonstra o primeiro quartil, em 29,123.

Os países com maior desigualdade de renda é Honduras, Colombia, Brasil, Guatemala, Chile e Panamá, todos com indice inferior a 30. Os países com menor índice de desigualdade são: Cuba, Portugal e Espanha.

2. RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS

2.1 Correlação linear

A matriz de correlação linear apresenta o teste de significância *P-Value*. Para a correlação foi utilizado o índice de *Pearson*. Conforme Las Casas e Guevara (2010)[1], a correlação (valores entre -1 e 1) é expressada pelo grau de relação linear que existe entre elas. Valores próximos a zero significam nenhuma relação, e próximos a 1 e -1 mostram alta relação positiva ou negativa, respectivamente. Vale ressaltar que o índice de correlação entre as variáveis não requer, necessariamente, que exista uma relação de causa-efeito entre as variáveis para dar sentido às relações. Para seleção das correlações, escolheu-se as variáveis

[1] LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisas de Marketing. São Paulo: Atlas, 2010.

com correlação superior a 0,92 e PValue igual a zero. A seguir, apresentamos a Tabela 2 com a matriz de correlação das variáveis estudadas (identificadas em amarelo).

Tabela 2 - Correlations: Social Progr; Ecosystem Su; IDH - 2013_N; Combustible ; ...

	Social Progress_	Ecosystem Sustai	IDH - 2013_N
Ecosystem Sustai	-0,221 0,337		
IDH - 2013_N	0,744 0,000	-0,419 0,058	
Combustible rene	-0,361 0,108	0,061 0,794	-0,654 0,001
Suceptibilidad (	0,721 0,000	-0,415 0,061	0,949 0,000
Global Innovatio	0,790 0,000	-0,396 0,075	0,806 0,000
Institutions_N	0,658 0,001	-0,377 0,093	0,483 0,027
Human Capital an	0,659 0,001	-0,439 0,047	0,794 0,000
Infraestructure_	0,803 0,000	-0,271 0,234	0,789 0,000
Market Sophistic	0,155 0,501	-0,237 0,301	0,114 0,622
Creative Outputs	0,723 0,000	-0,381 0,088	0,709 0,000
International Co	0,392 0,078	-0,201 0,383	0,050 0,831
Effective use of	0,382 0,088	-0,166 0,473	0,013 0,957
Regional coopera	0,359 0,110	-0,224 0,329	0,050 0,828
EPI Score_N	0,750 0,000	-0,071 0,759	0,713 0,000
Happy Planet Ind	-0,348 0,122	0,493 0,023	-0,296 0,193
GINNI Index_N_P	0,321 0,155	-0,583 0,006	0,523 0,015
Knowledge_N_R_N	0,672 0,001	-0,265 0,246	0,823 0,000

	Combustible rene	Suceptibilidad (	Global Innovatio
Suceptibilidad (	-0,612 0,003		
Global Innovatio	-0,345 0,126	0,766 0,000	
Institutions_N	-0,035	0,525	0,802

	0,880	0,014	0,000
Human Capital an	-0,414	0,736	0,741
	0,062	0,000	0,000
Infraestructure_	-0,499	0,755	0,916
	0,021	0,000	0,000
Market Sophistic	0,120	0,044	0,534
	0,606	0,850	0,013
Creative Outputs	-0,356	0,663	0,858
	0,114	0,001	0,000
International Co	0,159	0,091	0,320
	0,492	0,696	0,158
Effective use of	0,178	0,039	0,254
	0,441	0,867	0,266
Regional coopera	0,177	0,073	0,275
	0,443	0,752	0,227
EPI Score_N	-0,441	0,684	0,725
	0,046	0,001	0,000
Happy Planet Ind	-0,028	-0,203	-0,328
	0,905	0,378	0,146
GINNI Index_N_P	-0,413	0,492	0,364
	0,063	0,023	0,105
Knowledge_N_R_N	-0,401	0,773	0,871
	0,071	0,000	0,000

	Institutions_N	Human Capital an	Infraestructure_
Human Capital an	0,365		
	0,104		
Infraestructure_	0,762	0,673	
	0,000	0,001	
Market Sophistic	0,565	0,216	0,509
	0,008	0,347	0,018
Creative Outputs	0,685	0,524	0,741
	0,001	0,015	0,000
International Co	0,578	-0,072	0,348
	0,006	0,758	0,122
Effective use of	0,488	-0,077	0,313
	0,025	0,740	0,168
Regional coopera	0,507	-0,052	0,303
	0,019	0,822	0,181
EPI Score_N	0,483	0,649	0,748
	0,027	0,001	0,000
Happy Planet Ind	-0,181	-0,465	-0,233
	0,434	0,034	0,309
GINNI Index_N_P	0,250	0,445	0,300
	0,274	0,043	0,186
Knowledge_N_R_N	0,517	0,756	0,759
	0,016	0,000	0,000

	Market Sophistic	Creative Outputs	International Co
Creative Outputs	0,322 0,154		
International Co	0,220 0,337	0,346 0,124	
Effective use of	0,221 0,335	0,262 0,252	0,960 0,000
Regional coopera	0,092 0,691	0,330 0,145	0,932 0,000
EPI Score_N	0,311 0,171	0,511 0,018	-0,033 0,888
Happy Planet Ind	-0,315 0,165	-0,311 0,170	-0,018 0,940
GINNI Index_N_P	0,123 0,595	0,322 0,155	-0,271 0,234
Knowledge_N_R_N	0,321 0,156	0,637 0,002	0,064 0,783

	Effective use of	Regional coopera	EPI Score_N
Regional coopera	0,856 0,000		
EPI Score_N	-0,024 0,917	-0,060 0,797	
Happy Planet Ind	-0,089 0,702	-0,002 0,994	-0,322 0,155
GINNI Index_N_P	-0,309 0,173	-0,265 0,247	0,384 0,086
Knowledge_N_R_N	0,023 0,921	0,029 0,901	0,752 0,000

	Happy Planet Ind	GINNI Index_N_P
GINNI Index_N_P	-0,492 0,024	
Knowledge_N_R_N	-0,262 0,250	0,371 0,098

Cell Contents: Pearson correlation
P-Value

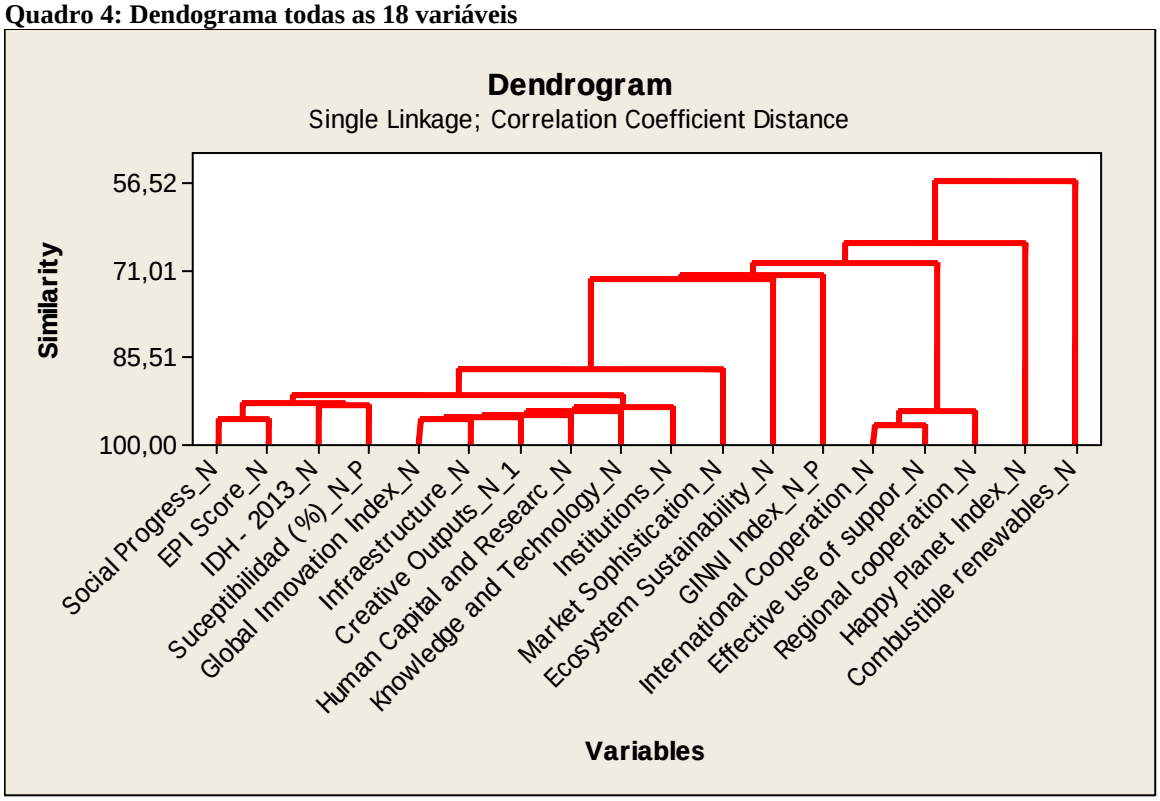
Em resumo, as análises efetuadas permitiu identificar 6 variáveis com maior correlação no tema Transformação Produtiva e Inovação Sustentável: Social Progress, Suceptibilidad, Effective Use of Support, Regional Cooperation e Market Sophisticate.

Com base no Dendograma (quadro 4) abaixo, pode-se perceber forte correlação entre as variáveis Social Progress, Suceptibilidad, Market Sophisticate e Knowledge and Technology. As variáveis Effective Use of Support e Regional Cooperation também tem forte correlação entre si.

2.2 Dendograma

O dendograma indica as variáveis mais próximas em relação ao coeficiente de correlação. Assim, verifica-se que as variáves Social Progress até Ginni Index apresentam alto grau de similaridade quando correlacionadas todas as variáveis do tema do estudo, conforme Quadro 4.

Quadro 4: Dendograma todas as 18 variáveis



Foi elaborado um novo Dendograma, (Quadro 9) onde apresentou-se as variáveis em 4 grupos (clusters) de acordo com as suas similaridades.

3 – Análise dos Componentes Principais

A análise dos componentes principais é normalmente utilizada quando a quantidade de variáveis dificulta a execução das análises. Utilizando-se os métodos estatístico multivariados é possível estudar o comportamento de poucas variáveis (normalmente um número próximo a 4 variáveis) que explicam todo o conjunto de variáveis de uma amostra.

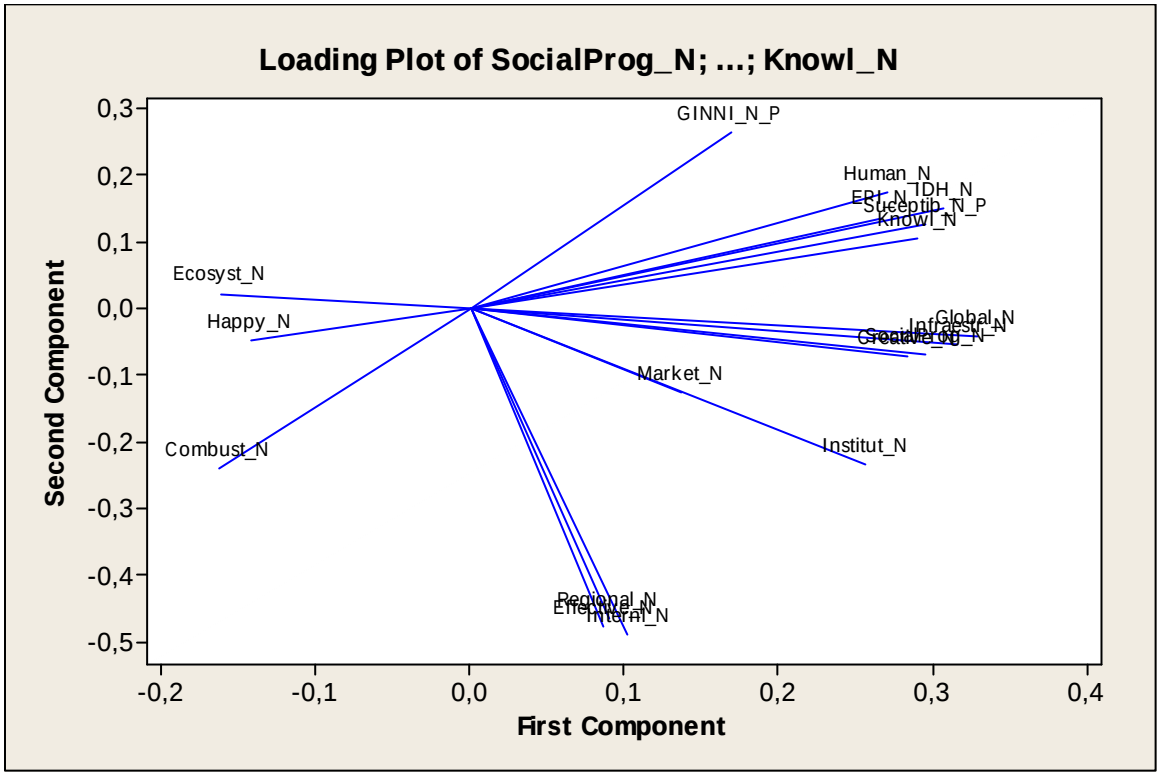
Os Métodos estatístico mais utilizados para a análise dos componentes principais são a Regressão Linear ou Multipla e a Análise Discriminante. Todavia também existe o método da Regressão Logística, Análise Multivariada de Variância (ANOVA), entre outras.

Segundo Bartholomew et al, pag.115 “O principal objetivo da análise de componentes principais (PCA) é substituir variáveis correlacionadas *p* métricos por um número muito menor de variáveis não correlacionadas que contêm mais informações no conjunto inicial. Isto simplifica a tarefa de compreender a estrutura de dados uma vez.” (tradução livre)

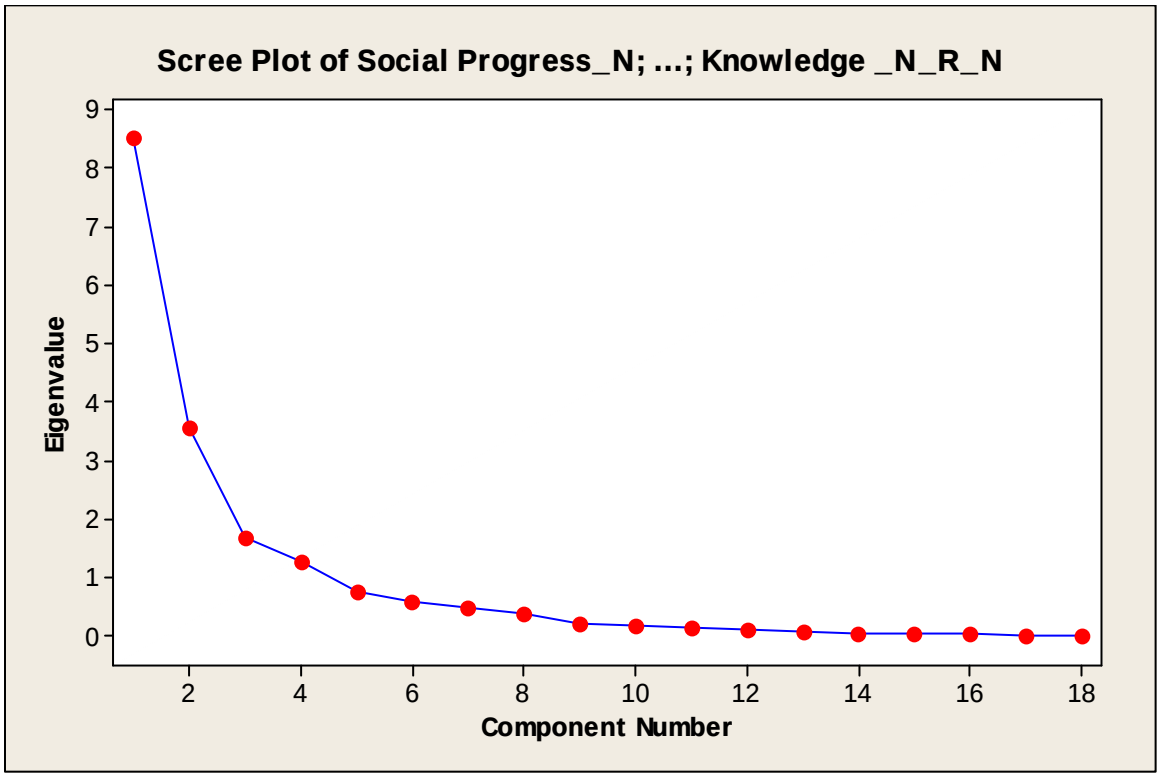
3.1 – Análise Multivariada

O gráfico apresentado no Quadro 5 demonstra a correlação entre as 18 variáveis para o tema do estudo. Verifica-se que as variáveis Ecosystem Sustaintable, Happy Planet index, Combustible renewables and waste se contrapõem às demais variáveis.

Quadro 5: Plot todas as variáveis do estudo



Quadro 6: Screen Plot das 18 variáveis do estudo:



Com base no Quadro 6, observa-se que quatro variáveis estão com valor superior a 1. Conforme, Las Casas e Guevara (2010)[2], "por convenção somente deve ser consideradas as componentes com uma contribuição igual ou superior a 1".

A partir dessa convenção, poder-se-á considerar somente 4 variáveis que acumuladas explicam 83% da variabilidade dos dados, conforme pode ser observado na Tabela 3.

Tabela 3: Principal Component Analysis: SocialProg_N; Ecosyst_N; IDH_N; Combust_N; Sucept

Eigenanalysis of the Correlation Matrix

	1	2	3	4	5	6	7	8
Eigenvalue	8,5314	3,5458	1,6629	1,2582	0,7676	0,5674	0,4693	0,3871
Proportion	0,474	0,197	0,092	0,070	0,043	0,032	0,026	0,022
Cumulative	0,474	0,671	0,763	0,833	0,876	0,907	0,933	0,955

	9	10	11	12	13	14	15	16
Eigenvalue	0,2220	0,1706	0,1426	0,1129	0,0633	0,0440	0,0277	0,0207
Proportion	0,012	0,009	0,008	0,006	0,004	0,002	0,002	0,001
Cumulative	0,967	0,977	0,985	0,991	0,995	0,997	0,998	1,000

	17	18
Eigenvalue	0,0064	0,0001
Proportion	0,000	0,000
Cumulative	1,000	1,000

Efetuou-se o agrupamento das variáveis em 4 componentes, e nas próximas páginas deste estudo estão demonstrados os testes de validação destes componentes que explicam o tema do estudo Transformação Produtiva e Inovação Sustentável. A Tabela 4 apresenta os Componentes Principais e seus índices e a Tabela 5 demonstra que há correlação entre as variáveis PC1, PC2, PC3 E PC4.

Tabela 4 - Componentes Principais – 18 Variáveis

Variable	**PC1**	**PC2**	**PC3**	**PC4**
SocialProg_N	0,296	-0,069	0,143	0,061
Ecosyst_N	-0,162	0,022	0,505	-0,393
IDH_N	0,308	0,149	0,119	0,173
Combust_N	-0,163	-0,241	-0,326	-0,252
Suceptib_N_P	0,296	0,126	0,157	0,215
Global_N	0,329	-0,043	0,002	-0,173
Institut_N	0,257	-0,235	-0,071	-0,173
Human_N	0,271	0,173	-0,065	0,024
Infraestr_N	0,316	-0,054	0,126	-0,163
Market_N	0,138	-0,128	-0,329	-0,605
Creative_N	0,283	-0,072	0,017	0,027
Internl_N	0,102	-0,490	0,050	0,154
Effective_N	0,087	-0,478	0,029	0,133
Regional_N	0,090	-0,467	0,057	0,241
EPI_N	0,267	0,136	0,133	-0,258
Happy_N	-0,142	-0,049	0,537	-0,053
GINNI_N_P	0,169	0,263	-0,353	0,237
Knowl_N	0,291	0,104	0,093	-0,156

Tabela 5- Correlations: PC1; PC2; PC3; PC4

```
          PC1      PC2      PC3
PC2   -0,000
       1,000

PC3   -0,000    0,000
       1,000    1,000

PC4   -0,000   -0,000   -0,000
       1,000    1,000    1,000


Cell Contents: Pearson correlation
               P-Value
```

3.2 – Regressão Múltipla

A fim de determinar as variáveis principais, elaborou-se o cálculo da Regressão Múltipla para o PC1, PC2, PC3 e PC4 como variável *response* (dependente), correlacionando-se com todas as 18 variáveis do estudo.

O resultado da Regressão Múltipla é explicado pelo Coeficiente de Determinação Múltiplo, chamado de R-sq pelo programa Minitab. A adição de variável ao modelo aumenta o resultado de R-sq.

Segundo SWEENEY, et al, pag. 565, "Lembre-se de que a variável que estamos prevendo ou explicando é chamada variável dependente e a variável usada para prever ou explicar a variável dependente é denominada variável independente. Análise de Regressão Multipla é o estudo de como a variável dependente *y* se relaciona com duas ou mais variáveis independentes. "

Para análise do presente estudo, efetuou-se o cálculo da Regressão Múltipla, cujo resumo está apresentado na Tabela 6.

Tabela 6: Resumo das Variaveis Dependentes PCs versus 18 variáveis do estudo

Dados apresentados na:	Variável Dependente	Variáveis Independentes	Explicam em percentual
Tabela 8	PC1	Global Innovation e IDH	97,6%
Tabela 9	PC2	Happy Planet, EPI e Ecosystem Sustaintable	92,5%
Tabela 10	PC3	Happy Planet, Ecosystem Sustaintable, International Cooperation e Market Sophisticate	91,2%
Tabela 11	PC4	International Cooperation, Market Sophisticate, Knowledge and Technology, Happy Planet, Suceptibilidad, Global Innovation	83,9%

Conforme Tabela 6, e excluídas as variáveis independentes que se repetem nos PCs, podemos afirmar que o resultado da regressão stepwise das variáveis dependente PC1, PC2, PC3 e PC4, contra todas as 18 variáveis da amostra podem ser explicadas com grau de certeza superior a 84%, pela utilização de somente 10 variáveis: Global Innovation, IDH, Happy Planet, EPI, Ecosystem Sustaintable, Market Sophisticate, Suceptibilidad, International Cooperation, Knowledge and Technology e Suceptibilidad.

O autor do presente estudo, em análise não supervisionada, percebeu que nem todas as variáveis listadas na Tabela 6, devido a sua composição, trarão ganhos significados na análise do tema do estudo. Assim, as principais variáveis que explicam o tema Transformação Produtiva e Inovação Sustentável são : **Global Innovation, IDH, Market Sophisticate, International Cooperation e Knowledge and Technology.**

Os cálculos que nos levaram a concluir sobre as variáveis principais geraram relatórios, os quais estão apresentados nas tabelas 7 a 10 abaixo apresentadas.

Tabela 7 - Stepwise Regression: PC1 versus SocialProg_N; Ecosyst_N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1 on 18 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-7,435	-10,534	-11,222	-9,990	-9,517	-9,152
Global_N	0,1805	0,1028	0,0860	0,0802	0,0956	0,0959
T-Value	12,25	8,69	7,80	8,64	9,77	11,17
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
IDH_N		0,0960	0,0867	0,0803	0,0801	0,0818
T-Value		8,14	8,60	9,38	10,99	12,72
P-Value		0,000	0,000	0,000	0,000	0,000
SocialProg_N			0,0309	0,0367	0,0397	0,0356
T-Value			3,16	4,43	5,56	5,47
P-Value			0,006	0,000	0,000	0,000
Ecosyst_N				-0,0144	-0,0157	-0,0112
T-Value				-3,02	-3,82	-2,76
P-Value				0,008	0,002	0,015
Creative_N					-0,0244	-0,0241
T-Value					-2,66	-3,00
P-Value					0,018	0,009
Happy_N						-0,0074
T-Value						-2,35
P-Value						0,034
S	0,790	0,375	0,307	0,252	0,215	0,188
R-Sq	88,76	97,60	98,49	99,04	99,35	99,53
R-Sq(adj)	88,17	97,33	98,22	98,79	99,13	99,33

Tabela 8: Stepwise Regression: PC2 versus SocialProg_N; Ecosyst_N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2 on 18 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	3,748	6,513	7,498	8,477	9,264	9,425
Happy_N	-0,0518	-0,0625	-0,0480	-0,0453	-0,0436	-0,0444
T-Value	-5,17	-8,30	-8,76	-15,52	-24,64	-30,18
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
EPI_N		-0,0378	-0,0347	-0,0123	-0,0053	-0,0037
T-Value		-4,42	-6,40	-2,80	-1,81	-1,53
P-Value		0,000	0,000	0,013	0,090	0,149
Ecosyst_N			-0,0334	-0,0463	-0,0513	-0,0502

T-Value			-5,31	-12,07	-20,84	-24,60
P-Value			0,000	0,000	0,000	0,000
Global_N				-0,0366	-0,0279	-0,0242
T-Value				-6,72	-7,71	-7,48
P-Value				0,000	0,000	0,000
IDH_N					-0,0201	-0,0179
T-Value					-5,48	-5,82
P-Value					0,000	0,000
SocialProg_N						-0,0086
T-Value						-2,91
P-Value						0,011
S	0,728	0,518	0,327	0,173	0,103	0,0840
R-Sq	58,45	80,07	92,50	98,04	99,35	99,59
R-Sq(adj)	56,26	77,85	91,17	97,54	99,13	99,42

Tabela 9: Stepwise Regression: PC3 versus SocialProg_N; Ecosyst_N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3 on 18 predictors, with N = 21

Step	1	2	3	4
Constant	-1,6241	0,2905	1,1393	1,4718
Happy_N	0,0224	0,0440	0,0455	0,0437
T-Value	2,27	8,00	10,46	10,14
P-Value	0,035	0,000	0,000	0,000
Ecosyst_N		-0,0530	-0,0571	-0,0575
T-Value		-7,95	-10,63	-11,15
P-Value		0,000	0,000	0,000
Internl_N			-0,0100	-0,0091
T-Value			-3,48	-3,22
P-Value			0,003	0,005
Market_N				-0,0073
T-Value				-1,58
P-Value				0,133
S	0,719	0,348	0,274	0,262
R-Sq	21,32	82,54	89,81	91,19
R-Sq(adj)	17,18	80,60	88,01	88,99

Tabela 10: Stepwise Regression: PC4 versus SocialProg_N; Ecosyst_N; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4 on 18 predictors, with N = 21

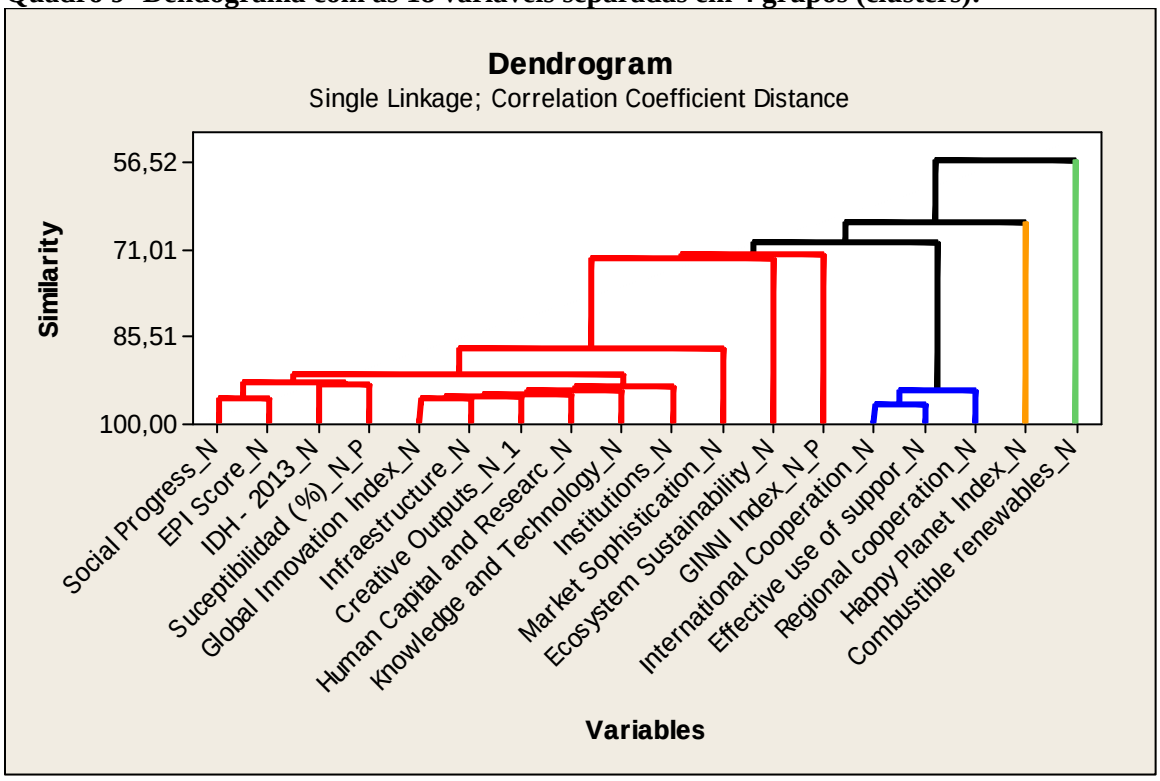
Step	1	2	3	4	5	6
Constant	-1,0676	-1,5036	-1,0404	-1,9828	-2,0338	-0,8689
Internl_N	0,0156	0,0131	0,0130	0,0126	0,0083	0,0064
T-Value	3,19	2,99	3,32	3,52	2,18	1,82
P-Value	0,005	0,008	0,004	0,003	0,046	0,090
Market_N		0,0185	0,0234	0,0267	0,0199	0,0094
T-Value		2,65	3,54	4,30	3,08	1,23

P-Value		0,016	0,003	0,001	0,008	0,240
Knowl_N			-0,0168	-0,0143	-0,0413	-0,0410
T-Value			-2,33	-2,13	-2,98	-3,28
P-Value			0,032	0,049	0,009	0,006
Happy_N				0,0106	0,0120	0,0118
T-Value				2,11	2,61	2,86
P-Value				0,051	0,020	0,013
Global_N					0,036	0,061
T-Value					2,16	3,19
P-Value					0,047	0,007
Suceptib_N_P						-0,024
T-Value						-2,11
P-Value						0,054
S	0,479	0,417	0,374	0,341	0,307	0,277
R-Sq	34,91	53,19	64,52	72,22	78,83	83,92
R-Sq(adj)	31,49	47,99	58,26	65,28	71,77	77,03

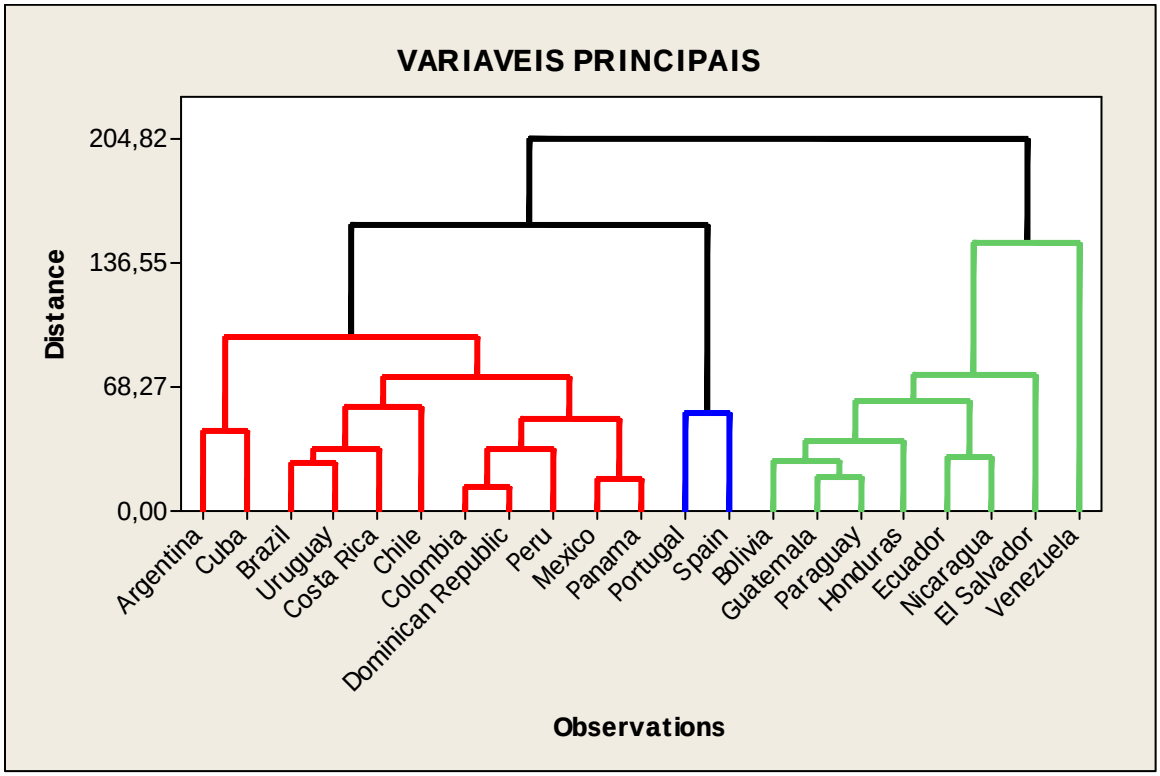
3.3 – DENDOGRAMA E CLUSTER ANÁLISE

A partir de um Dendograma é possível visualizar a similaridade dos países que compõem a amostra além de poder observar a correlação entre as variáveis.

Quadro 9- Dendograma com as 18 variáveis separadas em 4 grupos (clusters):



Quadro 7– Dendograma da Amostra com 3 Clusters.



Explicações adicionais:
Para efeito de análise, elaborou-se um Dendograma, agrupando os países em 3 grupos, segundo a similaridade, conforme Quadro 7.

3.4 ANÁLISE DISCRIMINANTE

Utilizou-se a Análise Discriminante para identificar e analisar os componentes principais, com base de todas as variáveis do tema do estudo, ou seja as 18 variáveis. Desta forma, as 4 novas variáveis: PC1, PC2, PC3, PC4, foram testadas para analisar e decidir se podem melhor representar as 18 variáveis iniciais da amostra.

Conforme (Hair. pag. 273) “análise discriminante e regressão logística são as técnicas estatísticas adequadas quando a variável dependente é uma variável categórica (nominal ou não métrico) e as variáveis independentes são variáveis métricas.” (tradução livre)

Tabela 11 - Discriminant Analysis: 3 Grups versus PC1; PC2; PC3; PC4

```
Linear Method for Response: 3 Grups

Predictors: PC1; PC2; PC3; PC4

Group          1         2         3
Count         11         8         2
```

```
Summary of classification

                     True Group
Put into Group       1       2       3
1                   10       0       0
2                    1       8       0
3                    0       0       2
Total N             11       8       2
N correct           10       8       2
Proportion       0,909   1,000   1,000

N = 21           N Correct = 20           Proportion Correct = 0,952

Squared Distance Between Groups
          1         2         3
1   0,0000   11,9171   26,5657
2  11,9171    0,0000   69,4009
3  26,5657   69,4009    0,0000

Linear Discriminant Function for Groups
                1         2         3
Constant   -0,424    -3,315   -17,066
PC1         0,748    -2,007     3,912
PC2         1,343    -4,278     9,722
PC3         0,394    -0,186    -1,423
PC4        -1,543     4,940   -11,275

Summary of Misclassified Observations

                 True   Pred            Squared
Observation     Group  Group  Group  Distance  Probability
          8**       1      2      1     4,486        0,251
                                  2     2,303        0,749
                                  3    47,441        0,000
```

O cálculo da Análise Discriminante, conforme Tabela 11 demonstra que temos 95,2% de proposições corretas e que a República Dominicana seria melhor classificada no grupo 2.Todavia, os países classificados no grupo 2 tem características de governabilidade que diferem da República Dominicana, o que influenciou o autor deste estudo a manter a República Dominicana no grupo 1.

3.5 – CÁLCULO DA ANOVA COM CLASSIFICAÇAO DOS GRUPOS DE PAISES DENTRO DA REGIÃO AIBER

O cálculo da ANOVA é necessário para identificar a classificação, em ordem decrescente, dos 4 Grupos de Países, conforme Quadro 7 e explicações adicionais . A ANOVA foi calculada a partir da variável Knowledge and Technology versus 3Grupos, apresentado o resultado do cálculo na Tabela 12.

Tabela 12– Cálculo da Anova para todas as variáveis principais:

One-way ANOVA: IDH versus 3 Grups

```
Source   DF      SS     MS      F      P
3 Grups   2  1917,9  959,0  17,64  0,000
Error    18   978,6   54,4
Total    20  2896,5

S = 7,373   R-Sq = 66,21%   R-Sq(adj) = 62,46%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  ---------+---------+---------+---------+
1      11  70,451  6,727               (---*---)
2       8  54,428  8,419  (---*----)
3       2  83,773  5,475                     (--------*--------)
                          ---------+---------+---------+---------+
                                  60        72        84        96
```

One-way ANOVA: Global versus 3 Grups

```
Source   DF      SS      MS      F      P
3 Grups   2  2550,3  1275,1  70,10  0,000
Error    18   327,4    18,2
Total    20  2877,7

S = 4,265   R-Sq = 88,62%   R-Sq(adj) = 87,36%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev  --------+---------+---------+---------+-
1      11  45,029  4,142              (--*-)MÉDIO
2       8  29,535  4,321  (--*-)BAIXO
3       2  66,699  5,022                           (-----*----)MELHOR
                          --------+---------+---------+---------+-
                                 36        48        60        72
```

One-way ANOVA: Market versus 3 Grups

```
Source   DF    SS   MS     F      P
3 Grups   2  1062  531  3,56  0,050
Error    18  2685  149
Total    20  3747

S = 12,21   R-Sq = 28,34%   R-Sq(adj) = 20,38%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
1      11  32,40  11,71     (-----*----)
2       8  28,48  12,47  (-----*-----)
3       2  54,15  15,00            (-----------*-----------)
                          -------+---------+---------+---------+--
                                30        45        60        75
```

One-way ANOVA: Internl versus 3 Grups

```
Source   DF    SS    MS     F      P
3 Grups   2  2077  1038  2,51  0,110
Error    18  7455   414
Total    20  9531

S = 20,35   R-Sq = 21,79%   R-Sq(adj) = 13,10%
```

```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
1      11  77,61  19,36                      (--------*-------)
2       8  56,88  23,01          (---------*--------)
3       2  62,04   0,00   (------------------*------------------)
                           +---------+---------+---------+---------
```

```
                        32        48        64        80
```

One-way ANOVA: Knowl versus 3 Grups

```
Source   DF      SS      MS      F      P
3 Grups   2  2149,1  1074,6  23,08  0,000
Error    18   838,1    46,6
Total    20  2987,2

S = 6,823   R-Sq = 71,94%   R-Sq(adj) = 68,83%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level   N    Mean   StDev  ------+---------+---------+---------+---
1      11  40,082   5,110            (--*--)
2       8  26,254   7,736  (---*--)
3       2  60,684  12,571                      (-----*------)
                           ------+---------+---------+---------+---
                                30        45        60        75
```

Segundo a análise da ANOVA, pode-se identificar a classificação dos grupos de países apresentados no Quadro 7- Dendograma. Abaixo apresenta-se a classificação dos países agrupados;

Quadro 8: Classificação dos grupo de países

Classificaçao	**n. original do Grupo (Anova)**	**Cor no Dendograma**	**Países**
Alto	g3	Azul	**Portugal e Espanha**
Médio	g1	Vermelho	**Brasil, México, Panamá, Argentina, Chile, Cuba, Uruguay, Peru, Colombia, Costa Rica, Republica Dominicana.**
Baixo	g2	Verde	**Bolivia, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Paraguai, Venezuela. `**

Desta forma, fez-se a seguinte reclassificação no programa Minitab:

De 3 para 1
De 1 para 2
De 2 para 3

3.6 Análise de Correspondência

Tabela 13 : Simple Correspondence Analysis: Knowl_M; IDH_M; Mrkt_M; Globl_M; Intl_M

```
Contingency Table
       Knowl_M    IDH_M   Mrkt_M  Globl_M   Intl_M    Total
g1      40,082   70,451   32,400   45,029   77,610  265,572
g2      26,254   54,428   28,480   29,535   56,880  195,577
g3      60,684   83,773   54,150   66,699   62,040  327,346
Total  127,020  208,652  115,030  141,263  196,530  788,495

Chi-Square Distances
       Knowl_M  IDH_M  Mrkt_M  Globl_M  Intl_M   Total
g1       0,170  0,000   1,039    0,137   1,969   3,315
g2       0,875  0,138   0,000    0,864   1,357   3,235
g3       1,199  0,094   0,856    1,106   4,684   7,939
Total    2,245  0,232   1,895    2,107   8,011  14,490

Relative Inertias
       Knowl_M  IDH_M  Mrkt_M  Globl_M  Intl_M  Total
g1       0,012  0,000   0,072    0,009   0,136  0,229
```

```
g2       0,060  0,010   0,000    0,060   0,094  0,223
g3       0,083  0,006   0,059    0,076   0,323  0,548
Total    0,155  0,016   0,131    0,145   0,553  1,000

Analysis of Contingency Table
 Axis  Inertia  Proportion  Cumulative  Histogram
    1   0,0172      0,9382      0,9382  ****************************
    2   0,0011      0,0618      1,0000  *
Total   0,0184
```

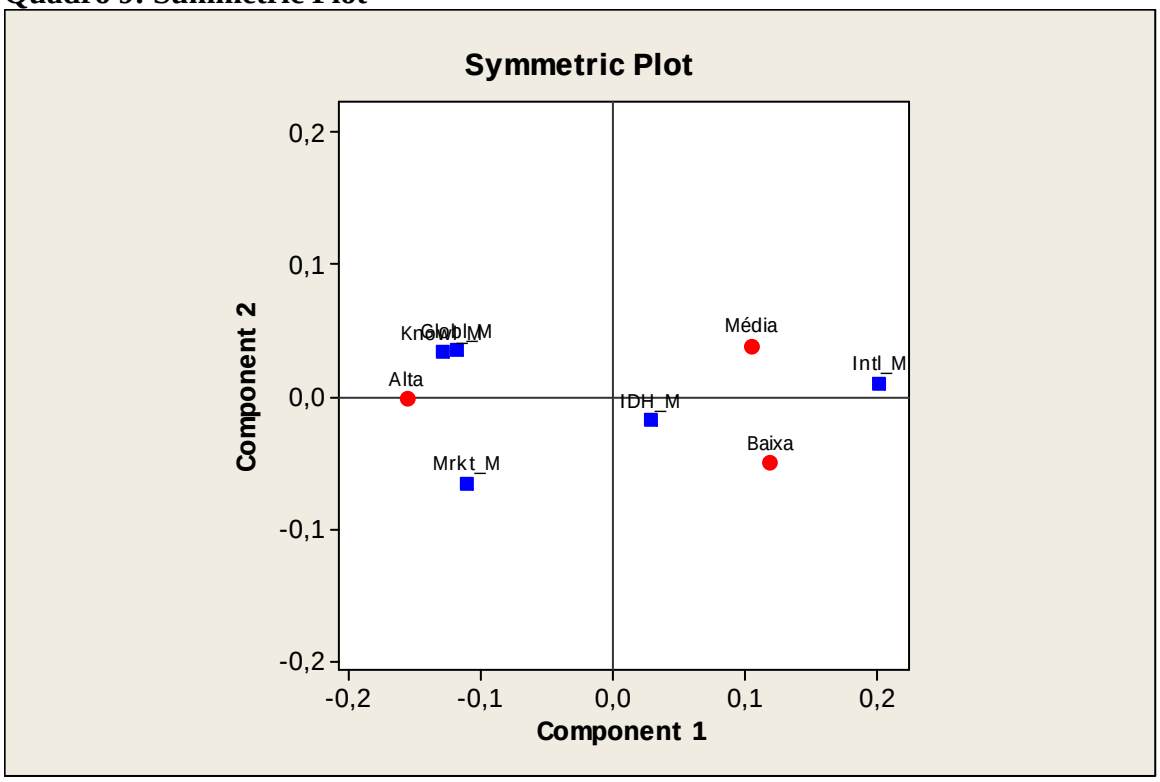
Row Contributions

ID	Name	Qual	Mass	Inert	Component 1 Coord	Corr	Contr	Component 2 Coord	Corr	Contr
1	g1	1,000	0,337	0,229	0,105	0,879	0,214	0,039	0,121	0,449
2	g2	1,000	0,248	0,223	0,118	0,848	0,202	-0,050	0,152	0,550
3	g3	1,000	0,415	0,548	-0,156	1,000	0,584	-0,002	0,000	0,001

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1 Coord	Corr	Contr	Component 2 Coord	Corr	Contr
1	Knowl_M	1,000	0,161	0,155	-0,128	0,934	0,154	0,034	0,066	0,165
2	IDH_M	1,000	0,265	0,016	0,028	0,726	0,012	-0,017	0,274	0,071
3	Mrkt_M	1,000	0,146	0,131	-0,110	0,739	0,103	-0,066	0,261	0,553
4	Globl_M	1,000	0,179	0,145	-0,117	0,920	0,143	0,035	0,080	0,188
5	Intl_M	1,000	0,249	0,553	0,202	0,998	0,588	0,010	0,002	0,022

Quadro 9: Summetric Plot



Quadro 10 : Row Plot

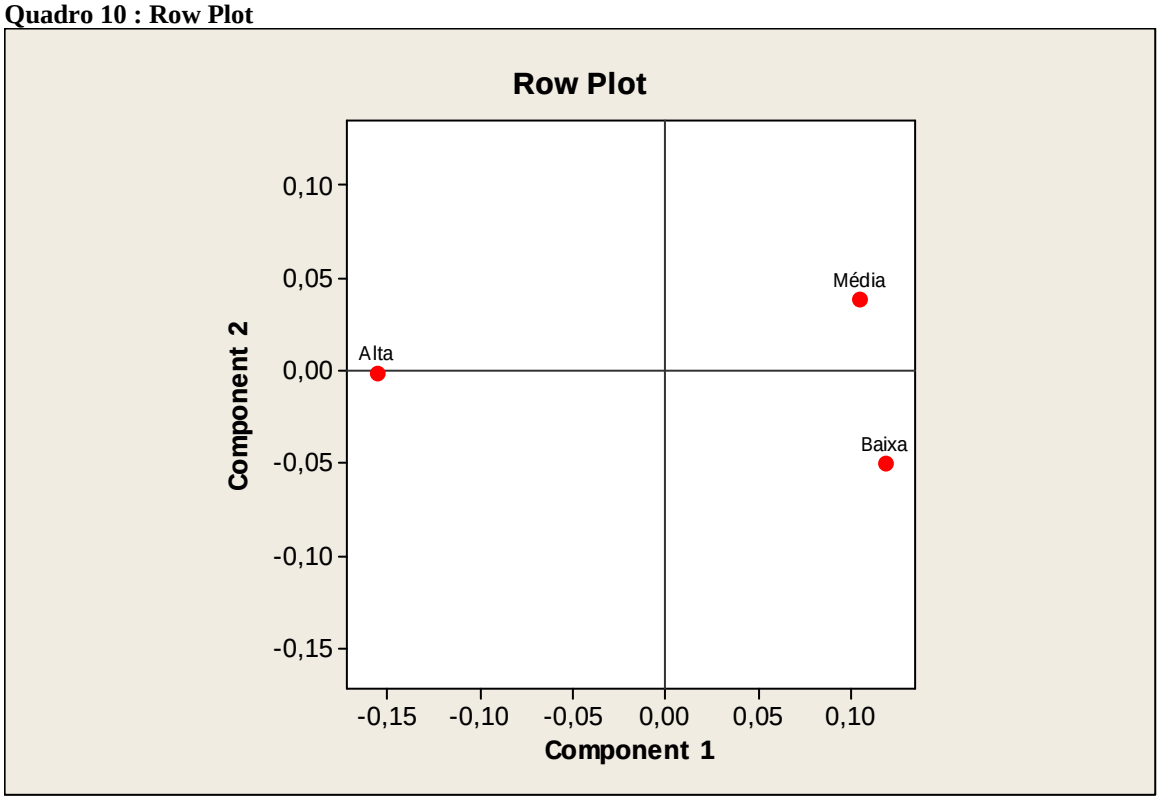
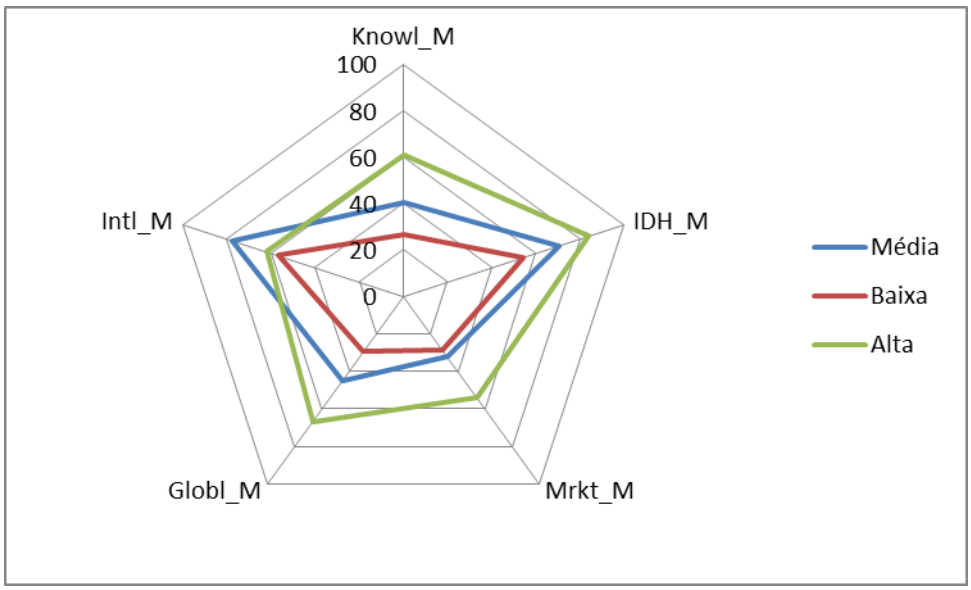


Tabela 14: Médias das principais variáveis:

	Knowl_M	IDH_M	Mrkt_M	Globl_M	Intl_M
Média	40,082	70,451	32,4	45,029	77,61
Baixa	26,254	54,428	28,48	29,535	56,88
Alta	60,684	83,773	54,15	66,699	62,04

Quadro 11: Gráfico Radar



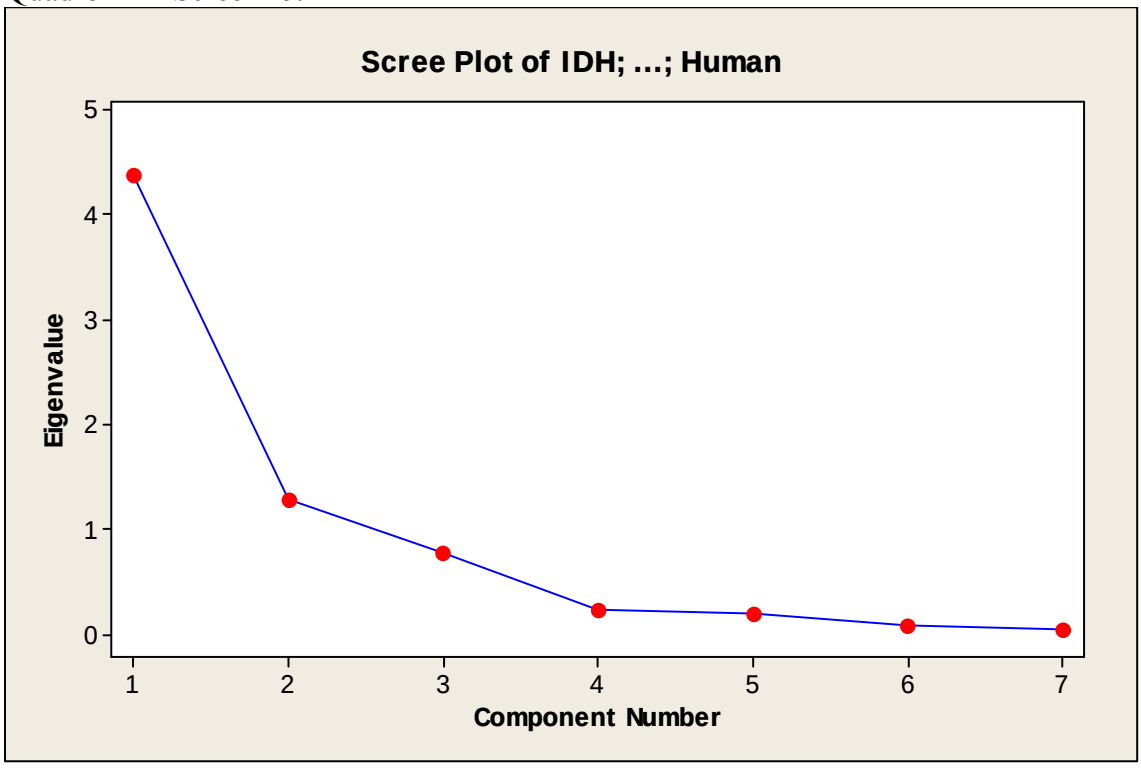
Verificando o gráfico das médias das variáveis selecionadas (quadro 10), notamos, em primeiro lugar, a localização oposta dos países Espanha e Portugal, formado pelos países com melhor posição no Ranking do AIBER. Os países na Média e Baixa classificação encontram-se muito próximos. Quanto às variáveis, vemos **Knowlegde and Technology, Global Innovation e International Cooperation** de um lado e com índices relativamente positivos, ao passo que **IDH e Market Sophisticate** encontram-se mais à esquerda, demonstrando serem esses problemas mais recorrentes para uma parte considerável dos países.

Com base no índice de Knowledge and Technology, elaboramos um ranking da classificação dos países que também estão classificados como Alta, Baixa e Média, de acordo com o grupo ao qual pertencem, conforme Dendograma do quadro 7.

4- DEMONSTRATIVO DO CÁLCULO DO RANKING DE PAÍSES

A fim de efetuar um ranking dos países do AIBER, classificados neste estudo em Alto, Médio e Baixo, elaborou-se o cálculo dos componentes principais utilizando as variáveis Knowledge and Technology, Market Sophisticate, Infraestructure, IDH, Global , International Cooperation e Human Capital. Como resultado apresentou-se os Quadros 12 e 13

Quadro 12 – Scree Plot



Quadro 13 – Loading Plot

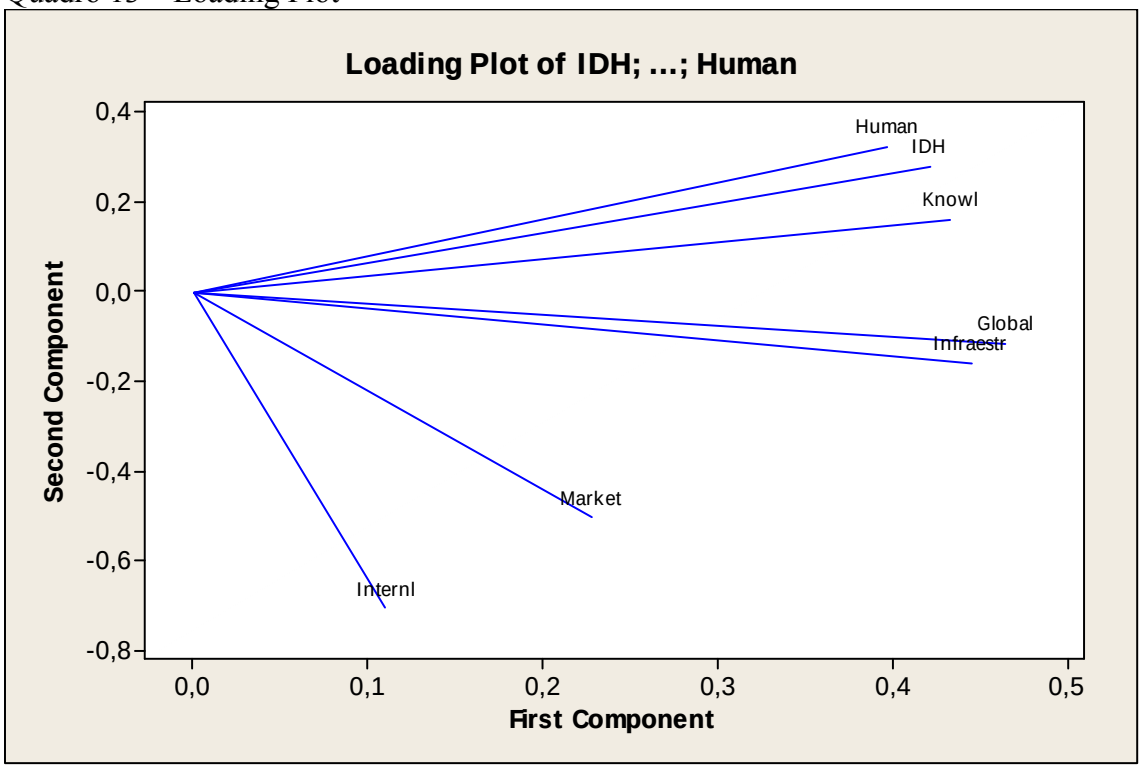


Tabela 15 - Principal Component Analysis: IDH; Global; Market; Internl; Knowl; Infraestr; H

Eigenanalysis of the Correlation Matrix

	PC1_NR	PC2_NR	PC3_NV	PC4_NV			
Eigenvalue	4,3916	1,2867	0,7716	0,2383	0,1934	0,0770	0,0413
Proportion	0,627	0,184	0,110	0,034	0,028	0,011	0,006
Cumulative	0,627	0,811	0,921	0,955	0,983	0,994	1,000

Variable	PC1	PC2	PC3	PC4
IDH	0,421	0,279	0,243	0,166
Global	0,464	-0,116	-0,007	0,107
Market	0,227	-0,503	-0,753	-0,068
Internl	0,110	-0,706	0,609	-0,192
Knowl	0,433	0,162	-0,005	0,481
Infraestr	0,445	-0,159	0,043	0,083
Human	0,397	0,325	-0,030	-0,826

Tabela 16 – Calculo da ANOVA para todos os PCs Novos
One-way ANOVA: PC1_NV versus 3 Grups_Or

Source	DF	SS	MS	F	P
3 Grups_Or	2	79,792	39,896	89,30	0,000
Error	18	8,041	0,447		
Total	20	87,833			

S = 0,6684 R-Sq = 90,84% R-Sq(adj) = 89,83%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	2	4,4471	1,3923	(----*----)

```
2      10   0,8391  0,5676                 (-*-)
3       9  -1,9206  0,6328   (-*--)
                             --+---------+---------+---------+-------
                            -2,0       0,0       2,0       4,0
```

One-way ANOVA: PC2_NV versus 3 Grups_Or

```
Source      DF     SS    MS     F      P
3 Grups_Or   2   0,36  0,18  0,13  0,882
Error       18  25,38  1,41
Total       20  25,73

S = 1,187   R-Sq = 1,38%   R-Sq(adj) = 0,00%


                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level   N    Mean  StDev  -----+---------+---------+---------+----
1       2   0,303  0,512  (------------------*-----------------)
2      10  -0,117  0,990        (-------*-------)
3       9   0,063  1,427         (--------*-------)
                            -----+---------+---------+---------+----
                              -1,0       0,0       1,0       2,0
```

One-way ANOVA: PC3_NV versus 3 Grups_Or

```
Source      DF      SS     MS     F      P
3 Grups_Or   2   4,818  2,409  4,09  0,034
Error       18  10,613  0,590
Total       20  15,431

S = 0,7679   R-Sq = 31,22%   R-Sq(adj) = 23,58%


                              Individual 95% CIs For Mean Based on
                              Pooled StDev
Level   N     Mean   StDev  -------+---------+---------+---------+--
1       2  -0,9793  0,6787  (--------------*-------------)
2      10   0,4606  0,8812                          (------*-----)
3       9  -0,2941  0,6289                 (-----*------)
                              -------+---------+---------+---------+--
                                  -1,60     -0,80     -0,00      0,80
```

One-way ANOVA: PC4_NV versus 3 Grups_Or

```
Source      DF     SS     MS     F      P
3 Grups_Or   2  0,106  0,053  0,20  0,817
Error       18  4,661  0,259
Total       20  4,767

S = 0,5089   R-Sq = 2,22%   R-Sq(adj) = 0,00%


                              Individual 95% CIs For Mean Based on
                              Pooled StDev
Level   N     Mean   StDev  ----+---------+---------+---------+-----
1       2  -0,2172  0,7794  (------------------*------------------)
2      10   0,0314  0,3707                (--------*--------)
3       9   0,0133  0,5934               (---------*---------)
                              ----+---------+---------+---------+-----
                               -0,80     -0,40     -0,00      0,40
```

Tabela 17 – Stepwise dos PCs Novos

Stepwise Regression: PC1_NV versus IDH; Global; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC1_NV on 7 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	-7,003	-8,579	-7,972	-7,519	-7,150	-7,754
Global	0,1700	0,1305	0,1223	0,0916	0,0666	0,0523
T-Value	18,46	11,77	12,87	7,82	6,60	8,03
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
IDH		0,0488	0,0328	0,0253	0,0151	0,0295
T-Value		4,42	3,13	2,96	2,35	6,36
P-Value		0,000	0,006	0,009	0,033	0,000
Human			0,0205	0,0232	0,0213	0,0208
T-Value			3,09	4,40	5,74	9,49
P-Value			0,007	0,000	0,000	0,000
Infraestr				0,0269	0,0343	0,0298
T-Value				3,40	5,91	8,46
P-Value				0,004	0,000	0,000
Knowl					0,0292	0,0302
T-Value					4,22	7,40
P-Value					0,001	0,000
Market						0,0131
T-Value						5,40
P-Value						0,000
S	0,494	0,352	0,290	0,227	0,159	0,0937
R-Sq	94,72	97,47	98,38	99,06	99,57	99,86
R-Sq(adj)	94,44	97,18	98,09	98,82	99,43	99,80

Stepwise Regression: PC2_NV versus IDH; Global; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC2_NV on 7 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	2,839	3,648	2,620	1,921	1,423	1,561
Internl	-0,04161	-0,03686	-0,03404	-0,03525	-0,03330	-0,03318
T-Value	-5,83	-6,64	-19,15	-30,54	-34,57	-81,60
P-Value	0,000	0,000	0,000	0,000	0,000	0,000
Market		-0,03436	-0,04312	-0,04173	-0,03700	-0,03847
T-Value		-3,88	-14,89	-22,44	-20,71	-49,70
P-Value		0,001	0,000	0,000	0,000	0,000
Human			0,02941	0,01927	0,02003	0,01858
T-Value			12,70	7,80	11,21	24,02
P-Value			0,000	0,000	0,000	0,000
IDH				0,0171	0,0272	0,0215
T-Value				5,11	7,80	13,28
P-Value				0,000	0,000	0,000

Infraestr					-0,0108	-0,0115
T-Value					-3,99	-10,06
P-Value					0,001	0,000
Knowl						0,0099
T-Value						8,38
P-Value						0,000
S	0,697	0,528	0,168	0,107	0,0767	0,0324
R-Sq	64,11	80,47	98,14	99,29	99,66	99,94
R-Sq(adj)	62,23	78,30	97,81	99,12	99,54	99,92

Stepwise Regression: PC3_NV versus IDH; Global; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC3_NV on 7 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	1,3997	-0,2311	-1,4676	-1,5246	-1,4120	-1,4205
Market	-0,04244	-0,05255	-0,05447	-0,05413	-0,05520	-0,05511
T-Value	-3,84	-12,33	-127,03	-173,63	-511,94	-1229,98
P-Value	0,001	0,000	0,000	0,000	0,000	0,000
Internl		0,02879	0,02850	0,02831	0,02786	0,02786
T-Value		10,77	106,59	146,33	479,38	1183,57
P-Value		0,000	0,000	0,000	0,000	0,000
IDH			0,02010	0,02204	0,01977	0,02011
T-Value			42,21	39,24	94,00	214,65
P-Value			0,000	0,000	0,000	0,000
Human				-0,00178	-0,00195	-0,00186
T-Value				-4,30	-18,11	-41,63
P-Value				0,001	0,000	0,000
Infraestr					0,00244	0,00248
T-Value					14,96	37,52
P-Value					0,000	0,000
Knowl						-0,00060
T-Value						-8,81
P-Value						0,000
S	0,676	0,254	0,0255	0,0179	0,00463	0,00187
R-Sq	43,74	92,45	99,93	99,97	100,00	100,00
R-Sq(adj)	40,78	91,61	99,92	99,96	100,00	100,00

Stepwise Regression: PC4_NV versus IDH; Global; ...

Alpha-to-Enter: 0,15 Alpha-to-Remove: 0,15

Response is PC4_NV on 7 predictors, with N = 21

Step	1	2	3	4	5	6
Constant	0,44970	-0,33843	0,06547	-0,67082	-0,56258	-0,35193
Human	-0,01176	-0,03954	-0,04184	-0,04828	-0,04879	-0,04858
T-Value	-1,92	-9,47	-13,32	-36,92	-44,05	-67,99
P-Value	0,070	0,000	0,000	0,000	0,000	0,000
Knowl		0,0503	0,0534	0,0427	0,0413	0,0425
T-Value		8,80	12,43	22,63	25,15	39,13
P-Value		0,000	0,000	0,000	0,000	0,000

Internl			-0,00630	-0,00686	-0,00773	-0,00800
T-Value			-3,99	-11,77	-13,41	-21,31
P-Value			0,001	0,000	0,000	0,000
IDH				0,0216	0,0194	0,0153
T-Value				10,48	10,32	10,24
P-Value				0,000	0,000	0,000
Infraestr					0,0036	0,0067
T-Value					2,81	6,36
P-Value					0,013	0,000
Market						-0,00337
T-Value						-4,71
P-Value						0,000
S	0,458	0,205	0,151	0,0556	0,0464	0,0299
R-Sq	16,25	84,20	91,84	98,96	99,32	99,74
R-Sq(adj)	11,85	82,45	90,41	98,70	99,09	99,63

Com base nas variáveis Global Innovation, International Cooperation, Market Sophisticate e Human Capital e utilizando a fórmula :
(0,95*0,63*'Global')+((0,64*'Internl')+(0,16*'Market')+(0,18*'Human'))*0,184, foi criada a nova variável chamada NOVAS, a qual foi normalizado e permitiu a geração do ranking dos países apresentado na Tabela 18.

TABELA 16 : RANKING AIBER

País	Classificaçao	Nível Relativo
Spain	Alta	100,00
Portugal	Alta	86,64
Chile	Média	80,23
Brazil	Média	64,80
Costa Rica	Média	62,28
Panama	Média	61,91
Uruguay	Média	58,87
Mexico	Média	54,71
Colombia	Média	54,62
Peru	Média	53,97
Cuba	Média	47,84
Argentina	Média	44,14
Dominican Republic	Média	40,67
El Salvador	Baixa	37,93
Paraguay	Baixa	36,91
Guatemala	Baixa	33,01
Honduras	Baixa	22,85
Bolivia	Baixa	21,74
Ecuador	Baixa	16,60
Nicaragua	Baixa	8,96
Venezuela	Baixa	0,00

4- MAPA DA REGIÃO AIBER

O mapa da Região AIBER foi pintado na cor laranja, que é utilizada pelo GPS para o tema do estudo Transformação Produtiva e Inovação Sustentável. Pintou-se em tom mais intenso os países classificados como ALTO no Ranking de países, com tom um pouco menos intenso para os países com classificação MEDIA e menos intenso os países considerados BAIXO. Os países pintos em cinza não fizeram parte da amostra do estudo.

CONSIDERAÇÕES FINAIS:

O presente trabalho teve como objetivo realizar análises de componentes principais para construção de índice com alto poder explicativo com o menor número possível de variáveis.

O estudo foi iniciado com 18 variáveis e ao final, mediante as análises das componentes principais, verificou-se que com apenas 5 das 18 variáveis é possível explicar 99 % da variabilidade.

Conclui-se que as principais variáveis que explicam o tema Transformação Produtiva e Inovação Sustentável utilizadas no estudo são: Global Innovation Index, Knowledge and Technology, IDH-2013, International Cooperation e Market Sophisticate.

REFERÊNCIAS BIBLIOGRÁFICAS

BARTHOLOMEW, David J.; STEELE, Fiona; MOUSTAKI, Irini; GALBRAITH, Jane I.. The Analysis and Interpretation of Multivariate Data for Social Scientists. 1.ed. London: : Chapman and Hall/CRC, 2002.

HAIR, Joseph F.; BLACK, William C.; BABIN, Barry J., ANDERSON, Ralph E.; TATHAM, Ronald L.. Multivariate Data Analysis. 6. Ed. Upper Saddle River: Pearson Prentice Hall, 2006.

LAS CASAS, Alexandre Luzzi; GUEVARA, Arnoldo José de Hoyos. Pesquisa de Marketing. São Paulo: Atlas, 2010.

SWEENEY, Dennis J.; WILLIAMS,Thomas A.; ANDERSON, David R.; Estatística Aplicada à Administração e Economia. 3. Ed. São Paulo: Cengage Learning, 2015.

http://pt.knoema.com/jklbfre/the-social-progress-index