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

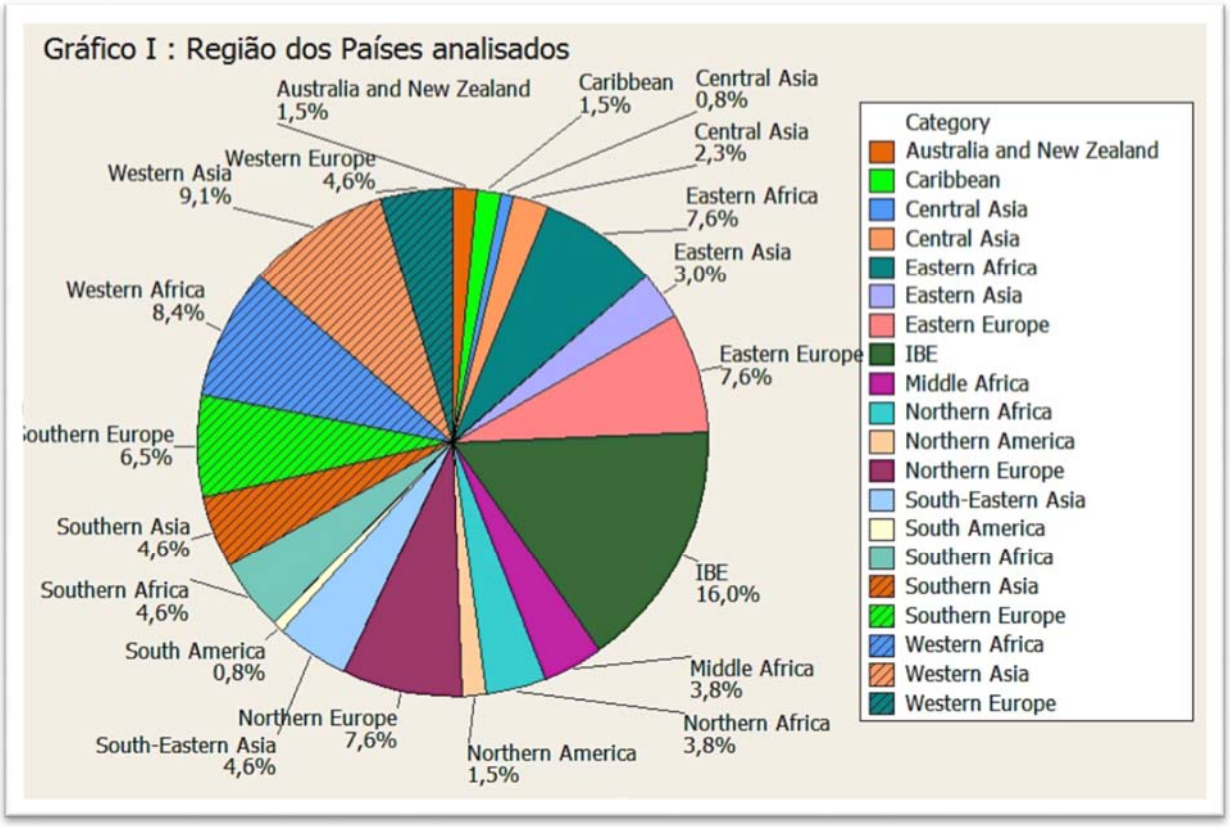
Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Social Progress Index	132	0	63,67	1,24	14,20	32,60	51,93	63,86	73,74	88,24
Maternal mortality rate	132	0	161,6	19,2	221,0	2,0	12,0	61,5	240,0	1100,0
Stillbirth rate (deaths/	132	0	12,878	0,844	9,696	2,000	4,000	10,000	20,750	47,000
Child mortality 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.

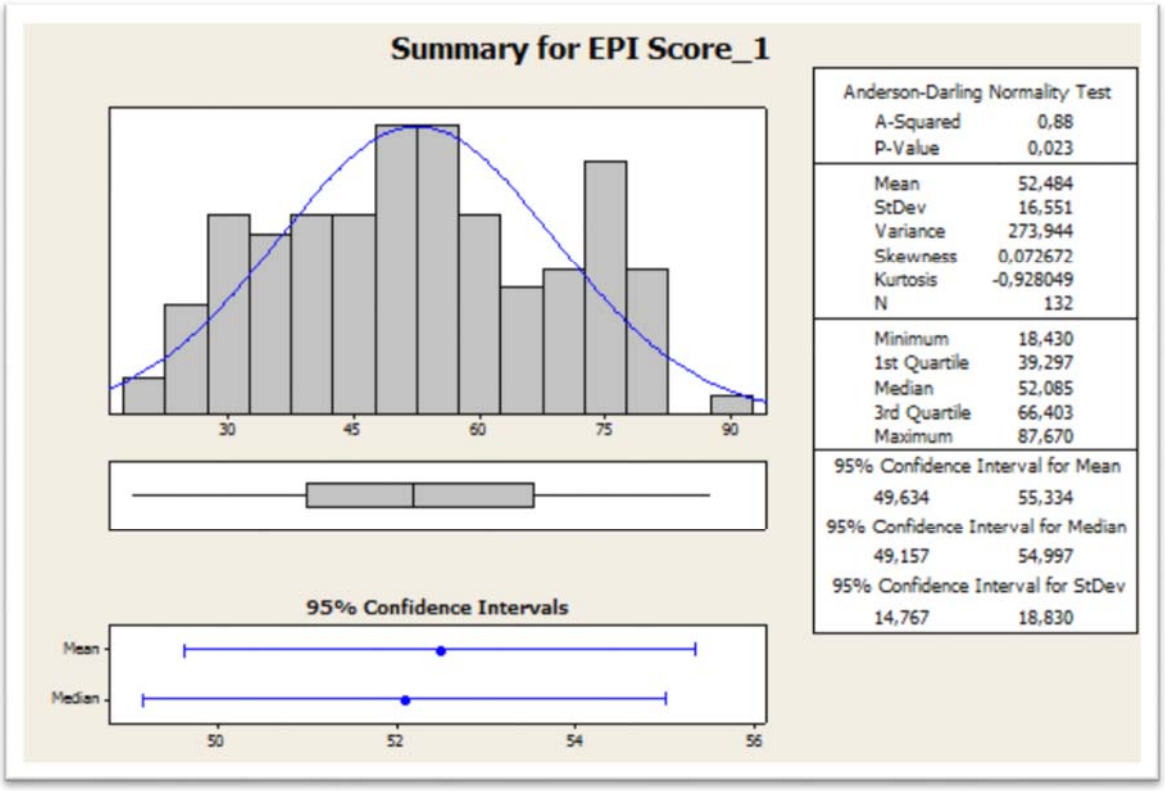
3.3 ANÁLISE EXPLORATÓRIA DAS VARIÁVEIS

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.

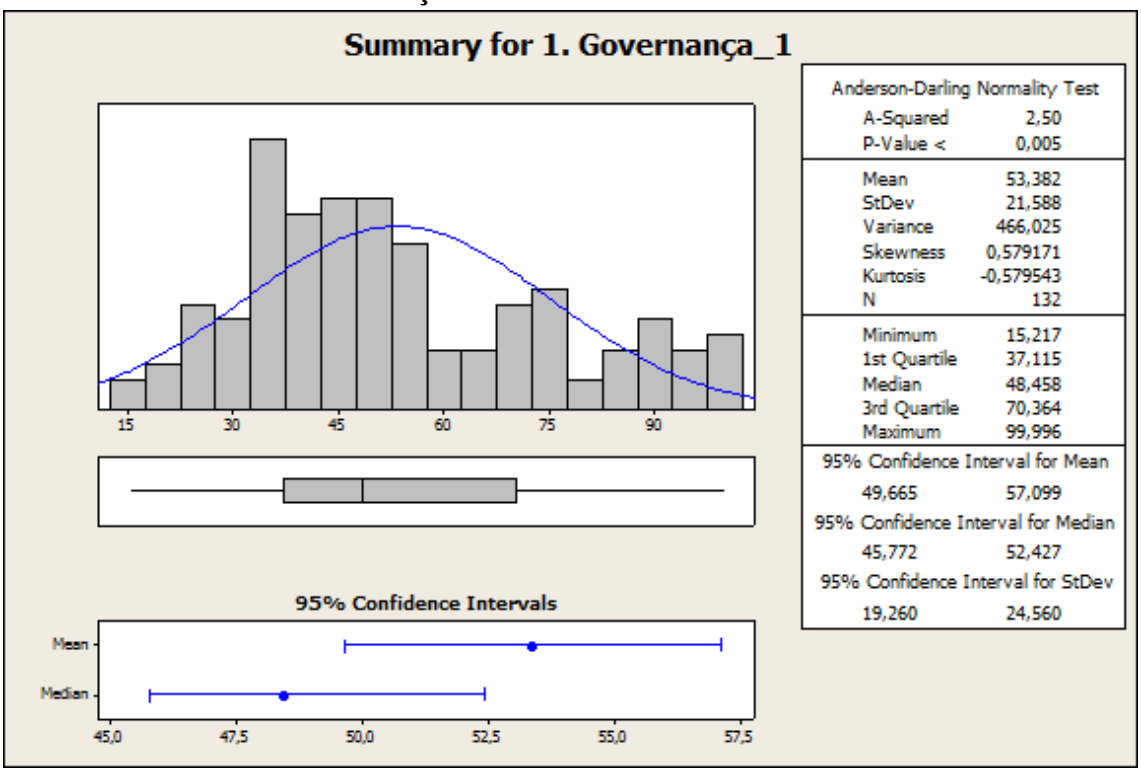
Os Indicadores de Governança Mundial segundo o relatório do Banco Mundial sobre seis grandes dimensões de governança para 215 países durante o período de 1996-2013 refere-se a:

- 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:

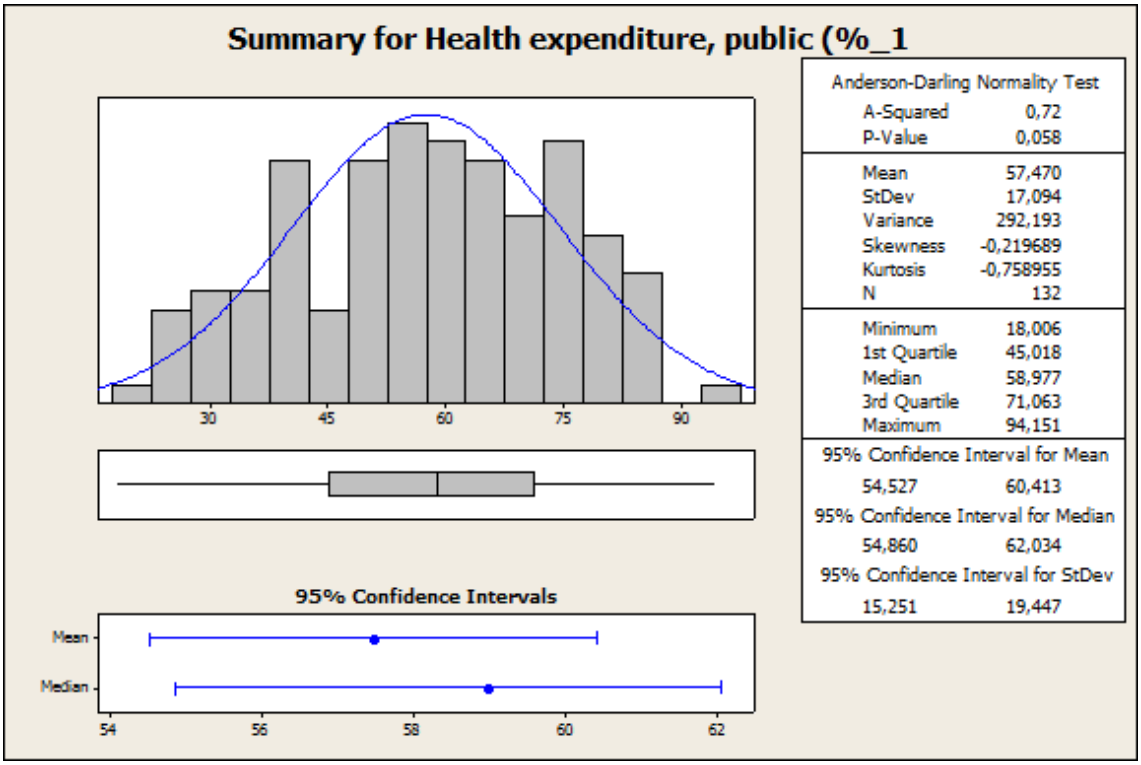
O Gráfico 3, mostra que índice em questão tem média de 53,382 (intervalo de confiança de 95%, entre 49,665 e 57,099) e desvio padrão de 21,588 (intervalo de confiança de 95%, entre 19,260 e 24,560). Pode-se observar pelo histograma que a maior parte dos países tem o índice de 0 a 50. Neste caso nota-se 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.

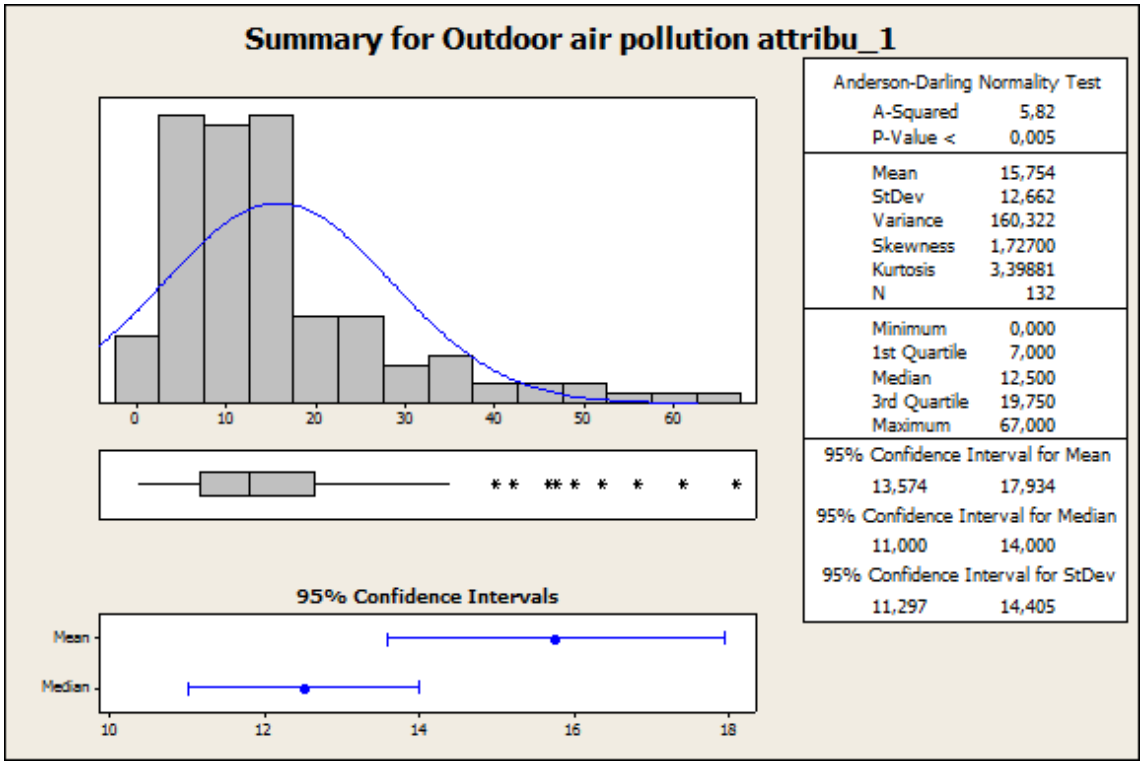
Pelo teste de Anderson-Darling para normalidade, os dados podem ser considerados próximos a normais (p-valor = 0,058), para o nível de significância de 0,05. Pela curva se percebe, porém com baixa frequência no substrato de 45 e um pouco mais alta no substrato de 40. Média e mediana muito próximas.

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.

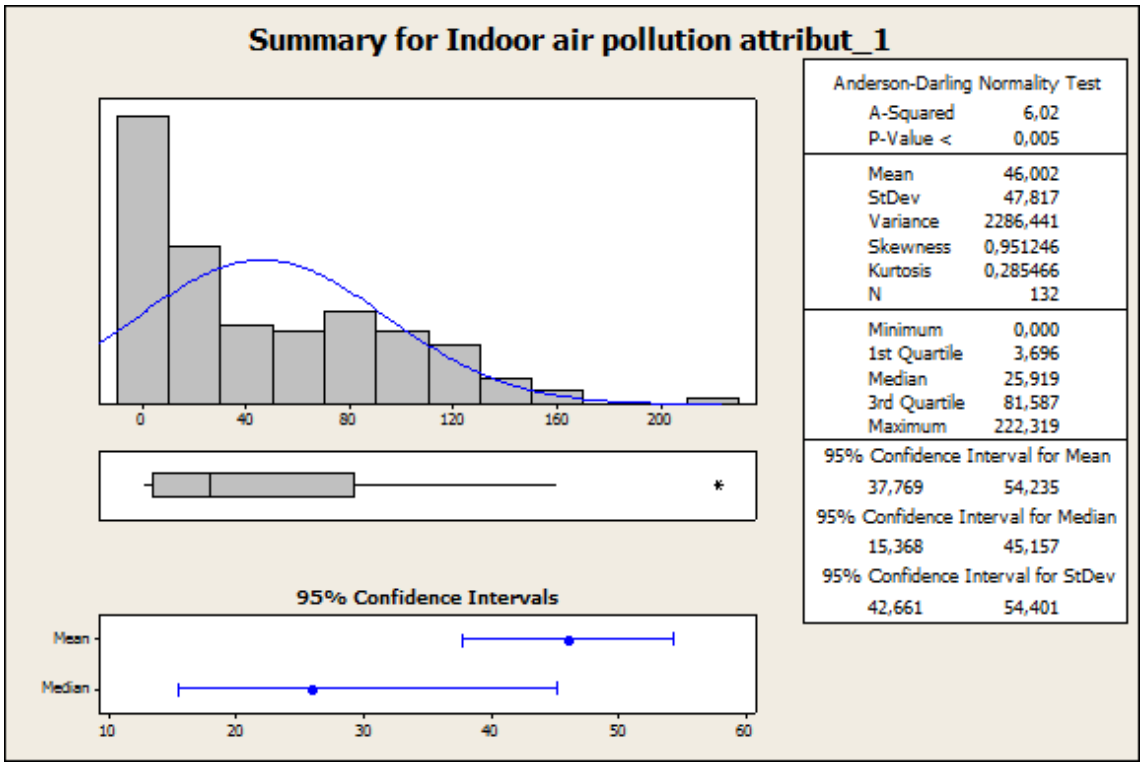
No caso, a mediana menor que a média, deslocando o gráfico para o lado esquerdo. A grande maioria dos países classificadas no substrato até 25, e países do Leste Europeu aparecendo com máximas.

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:

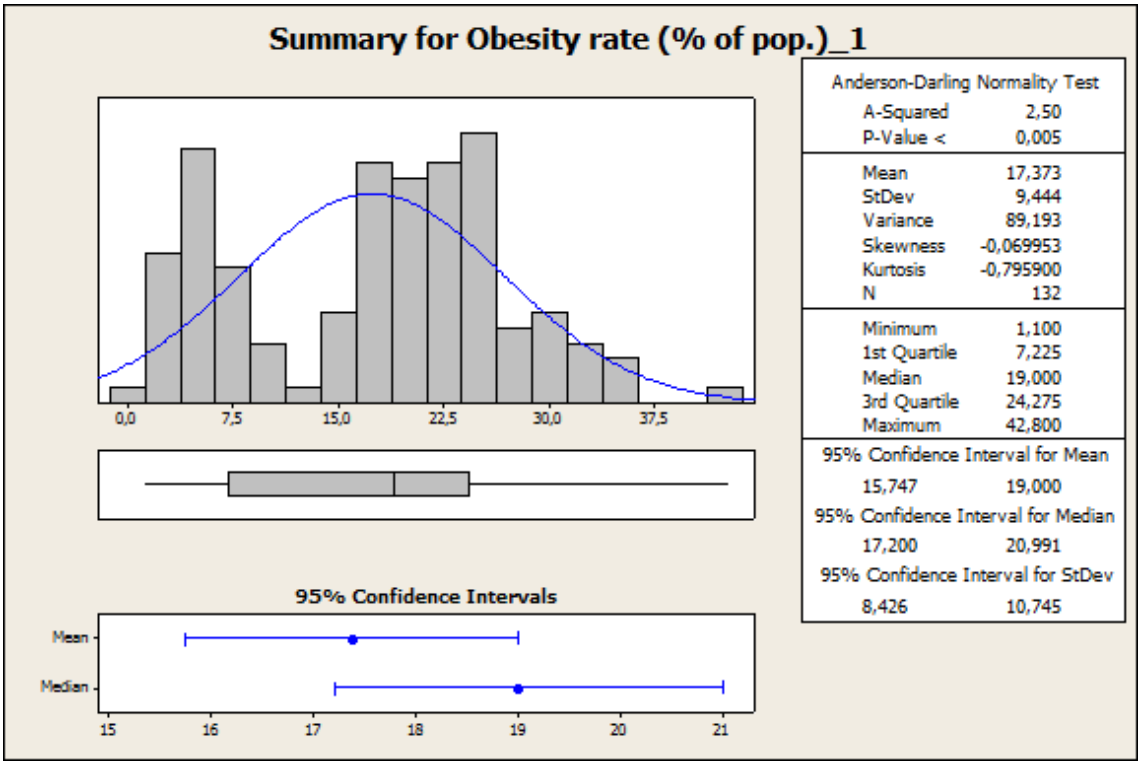
O valor do primeiro quartil de 3,696 indica que 25% da amostra possui valores entre 0 e 3,696. O índice Poluição do Ar Interior tem média 46,002 (intervalo de confiança de 95%, entre 37,769 e 54,235) e desvio padrão de 47,817 (intervalo de confiança de 95%, entre 42,661 e 54,401). Estes dados indicam 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:

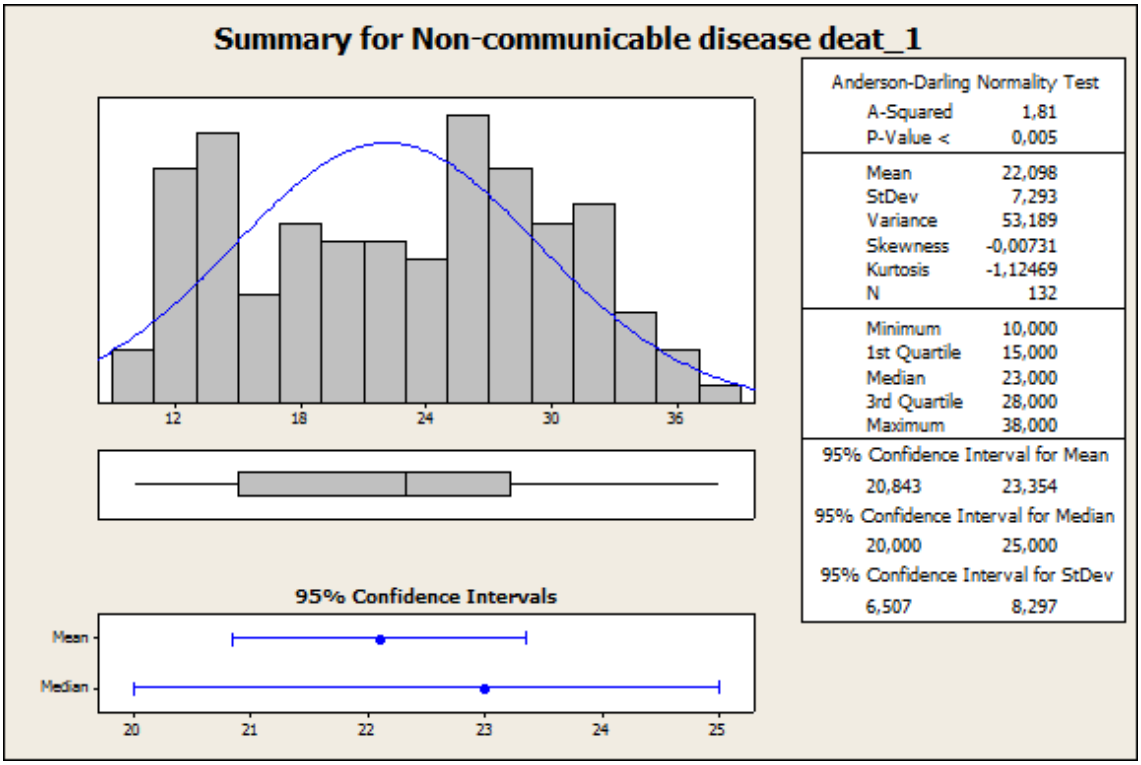
A média de Obesidade nos países pesquisados é de 17,373 (intervalo de confiança de 95%, entre 15,747 e 19) com desvio padrão de 9,444 (intervalo de confiança de 95%, entre 8,426 e 10,745), conforme apresentado no Gráfico 7.

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.

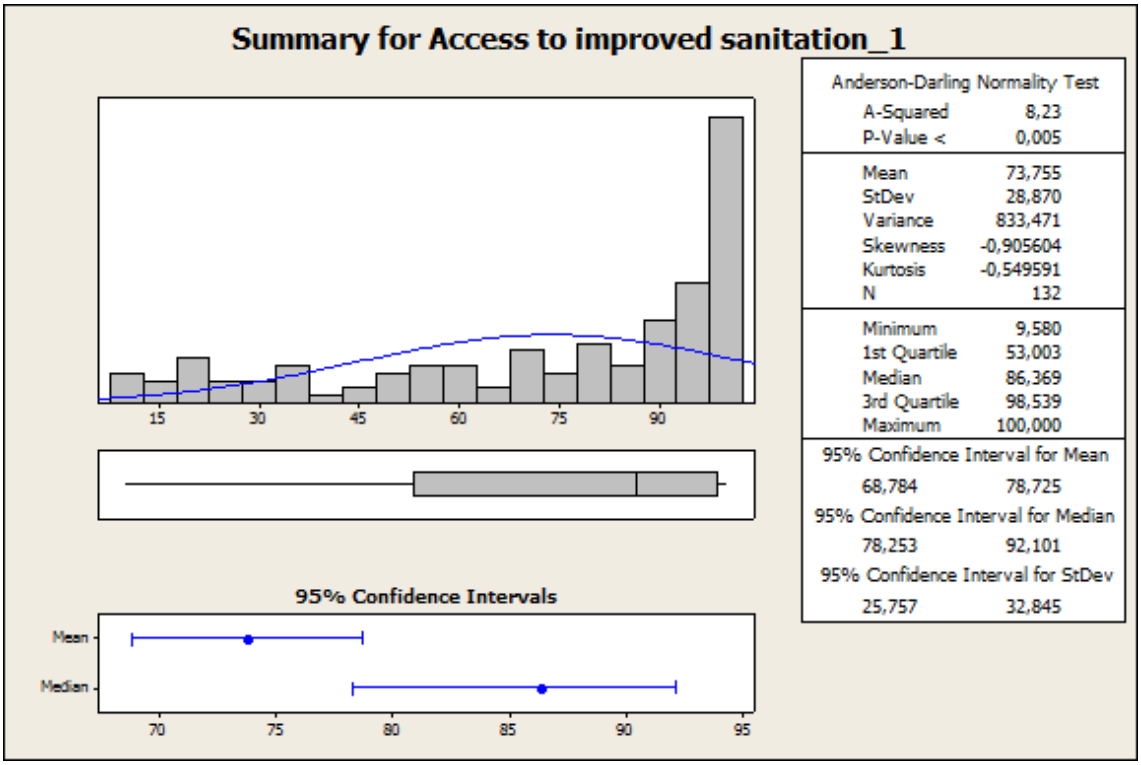
Análise:

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:

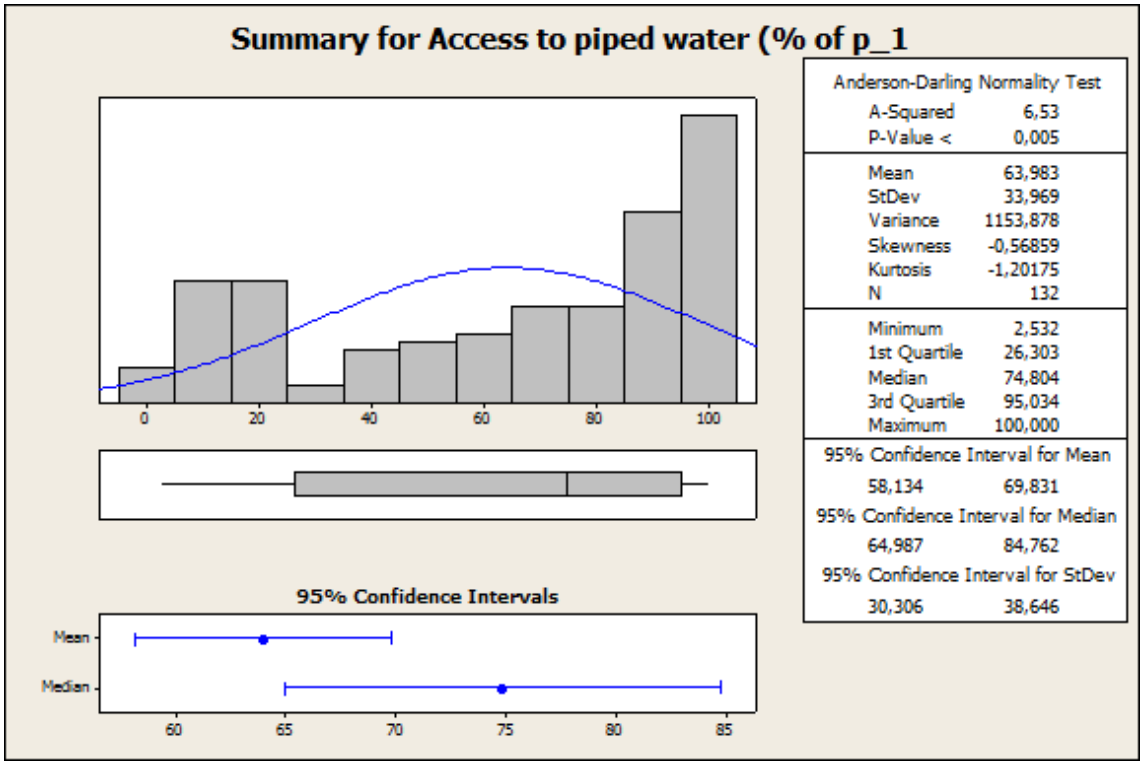
O índice em questão tem média de 73,755 (intervalo de confiança de 95%, entre 68,784 e 78,725) e desvio padrão de 28,870 (intervalo de confiança de 95%, entre 25,757 e 32,845). Esse é mais um índice que mostra a discrepância entre países de primeiro e de terceiro mundo, conforme mostra o Gráfico 8.

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:

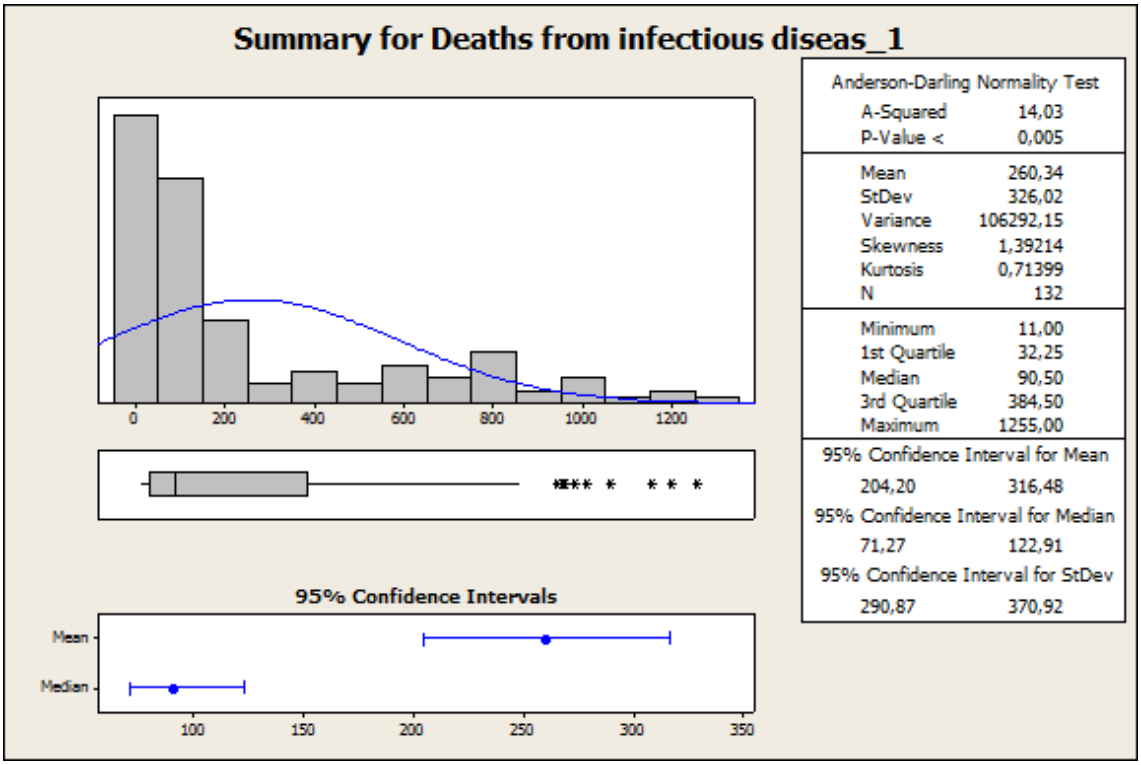
O índice de Acesso à Água Encanada médio é de 63,983 (intervalo de confiança de 95%, entre 58,134 e 69,831) e desvio padrão de 33,969 (intervalo de confiança de 95%, entre 30,306 e 38,646), conforme o Gráfico 10. Pode-se observar que o desvio-padrão desse índice é bem elevado, isso mostra que há uma 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).

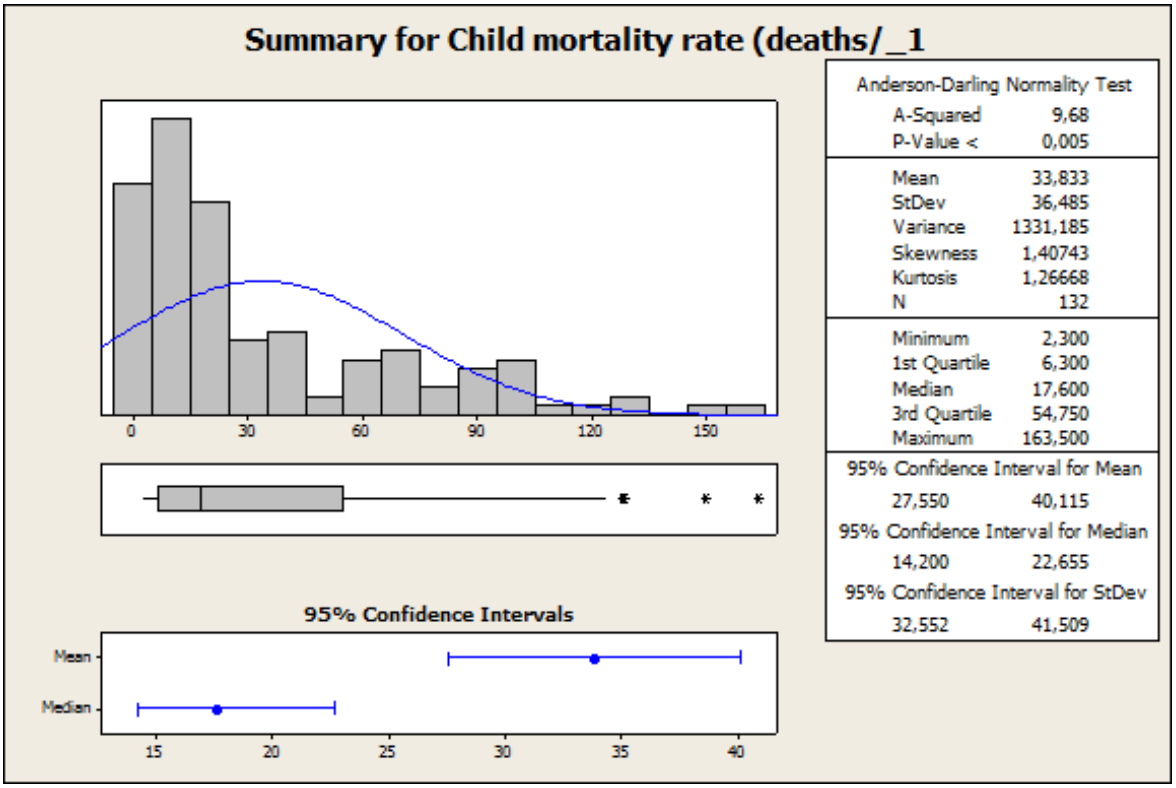
A Mediana (90,50) à Esquerda da Média (260,34 – que é quase 3 vezes maior) determinando intervalos de confiança bastante diferenciados. **A dispersão é muito alta**, com mínimo em 11 e máximo em 1255. A análise dos países nos substratos apresentados, com poucas **exceções são inversamente proporcionais 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:

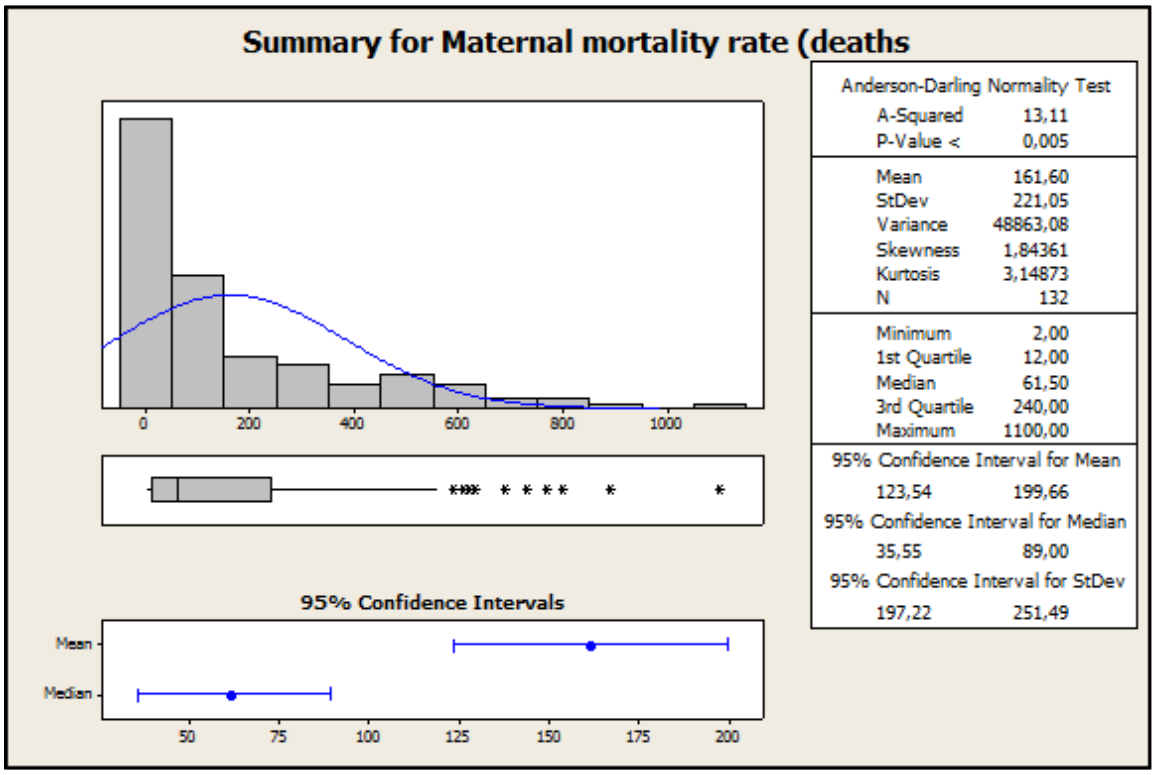
Conforme apresentado no Gráfico 12, a média para Mortalidade Infantil é de 33,833 (intervalo de confiança de 95%, entre 27,550 e 40,115) e desvio padrão de 36,485 (intervalo de confiança de 95%, entre 32,552 e 41,509).

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.

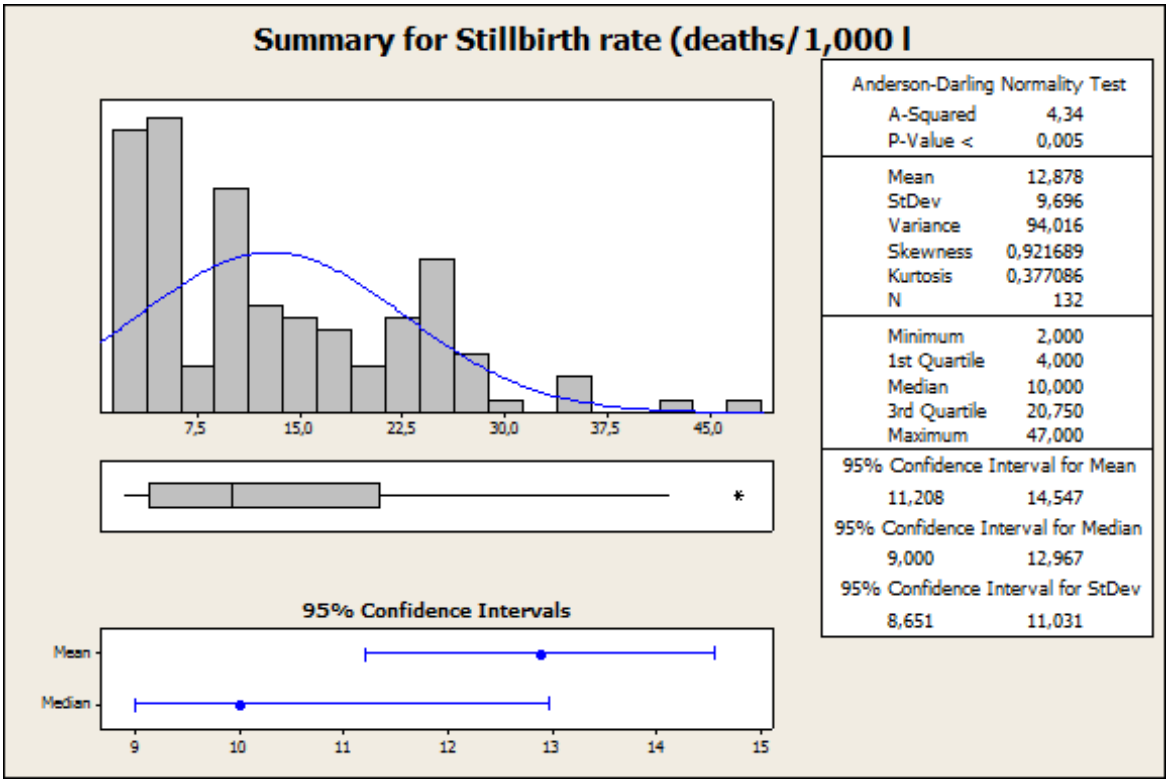
Análise:

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.

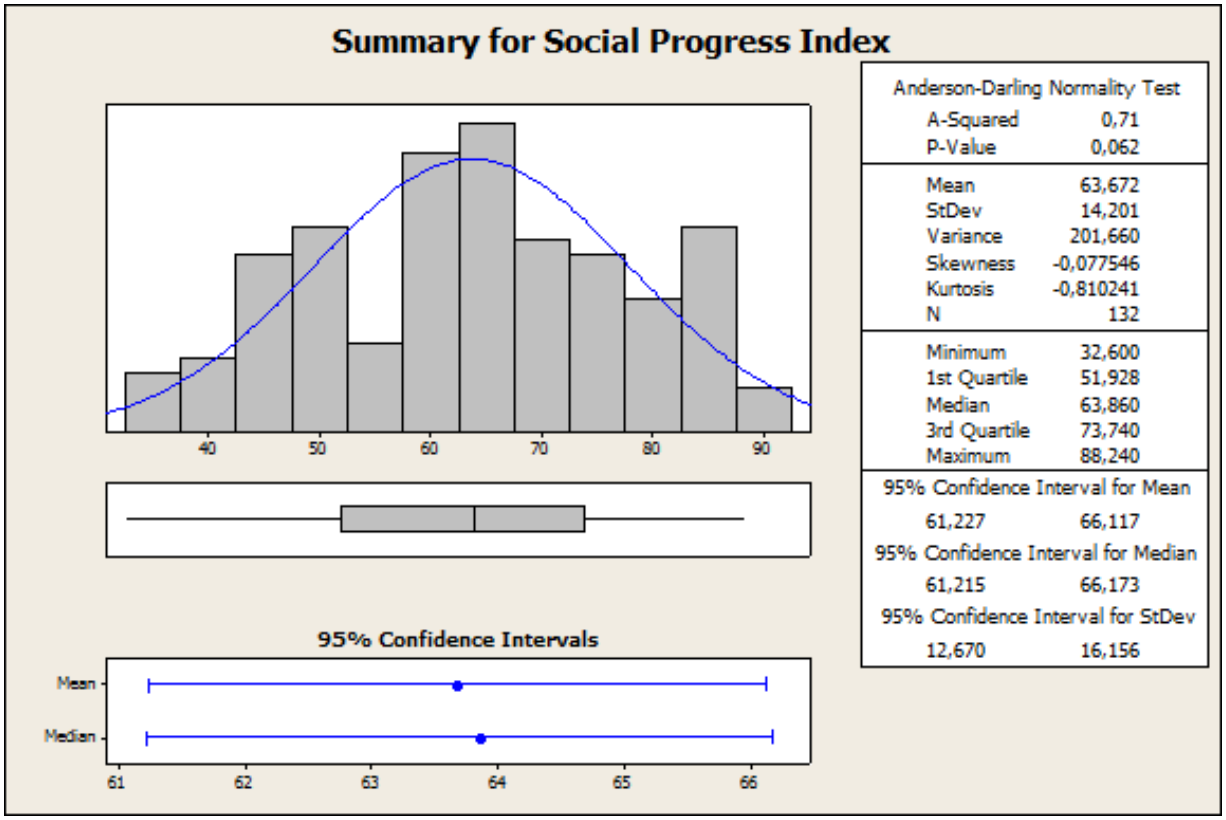
Análise:

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

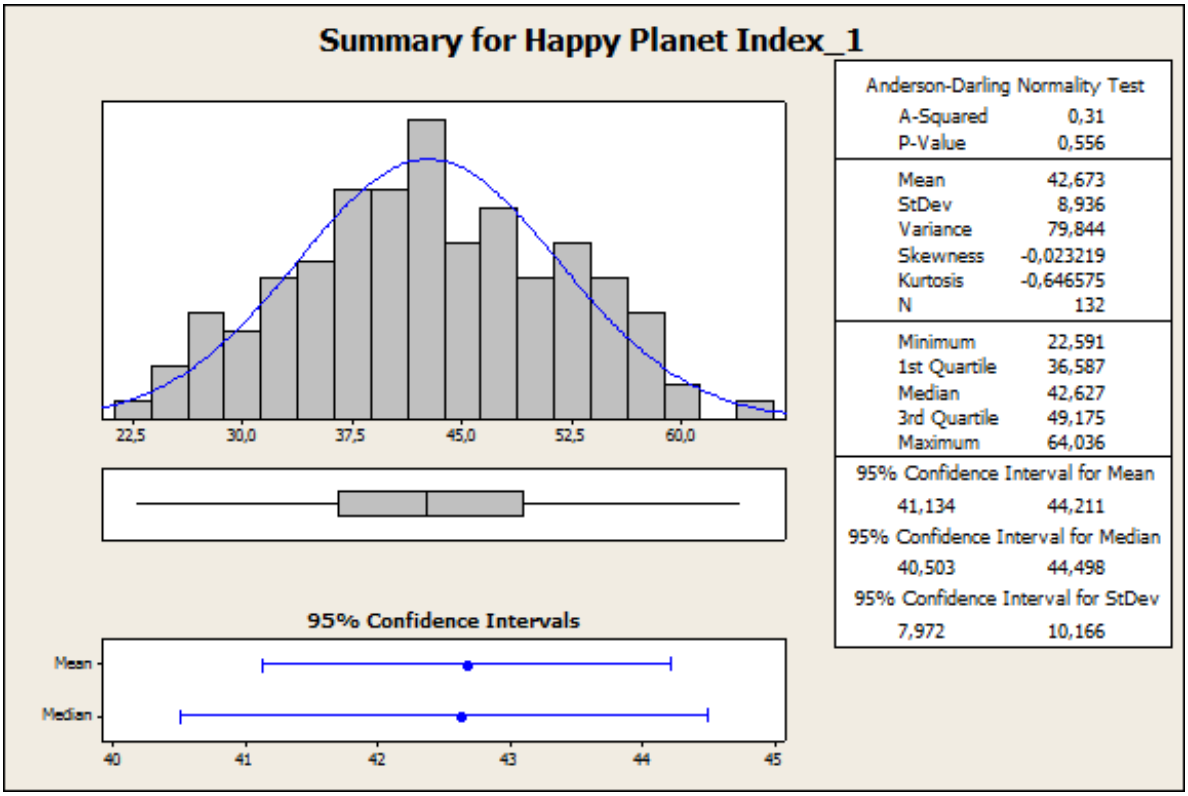
O Gráfico 15 demonstra o IPS médio de 63,672 (intervalo de confiança de 95%, entre 61,227 e 66,117) e desvio padrão de 14,201 (intervalo de confiança de 95%, entre 12,670 e 16,156).

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).

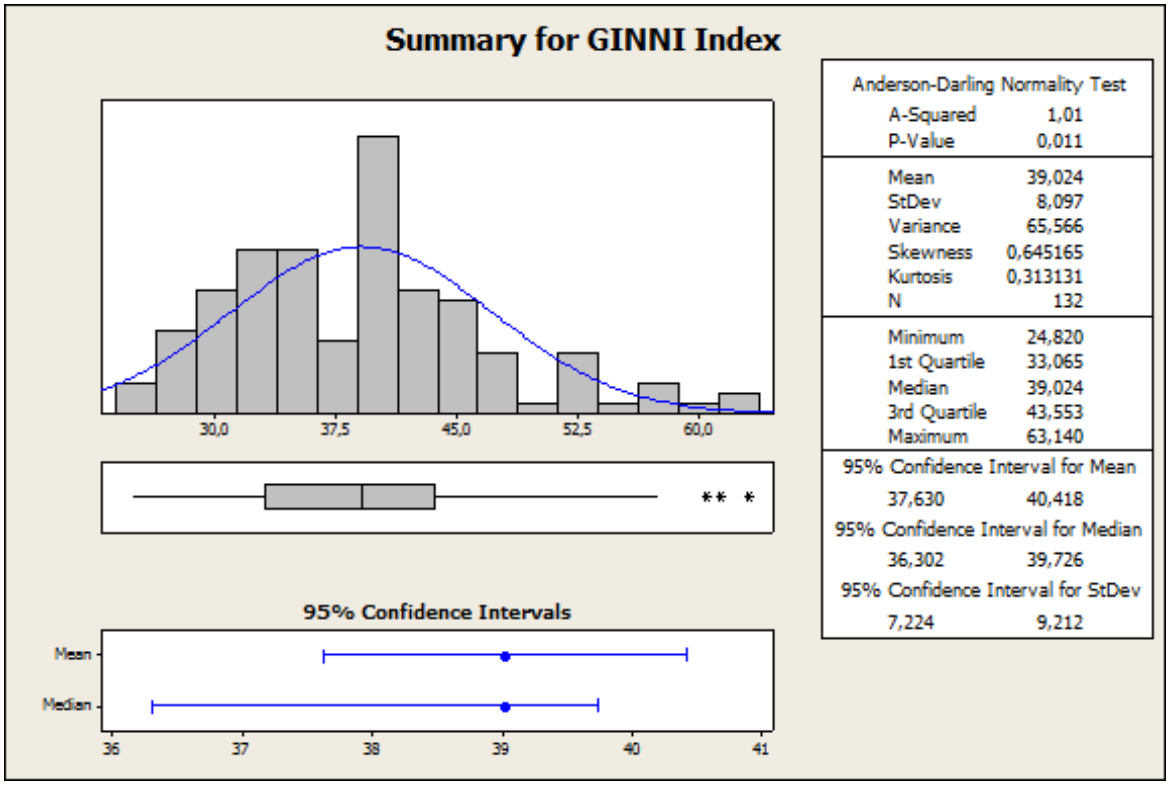
Pelo teste de Anderson-Darling para normalidade, os dados podem ser considerados normais (p-valor = 0,556), para o nível de significância de 0,05. Pode-se observar que a maioria dos dados se encontra em torno da média.

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

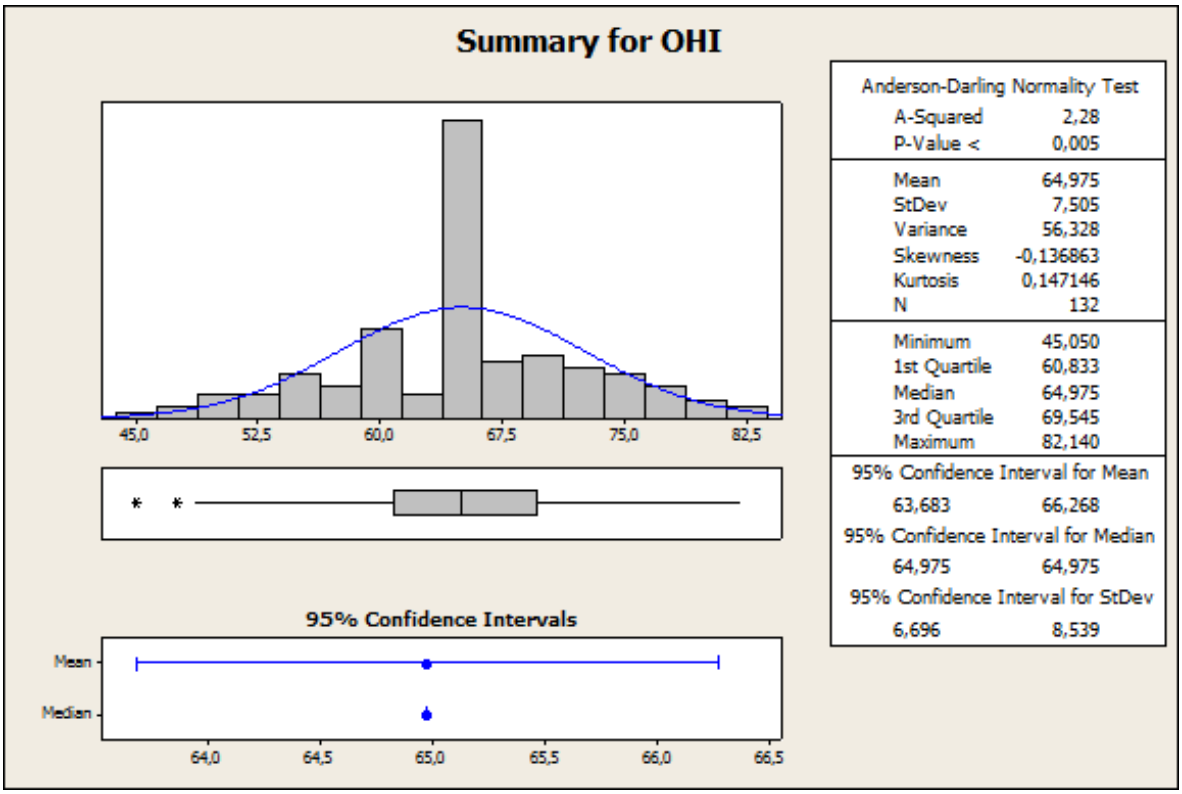
O Índice de Saúde do Oceano estabelece pontos de referência para atingir 10 objetivos sócio-ecológicos amplamente aceitos, e contagens dos oceanos adjacentes para 171 países e territórios. Avaliadas globalmente e por país, estas 10 metas públicas representam a ampla

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



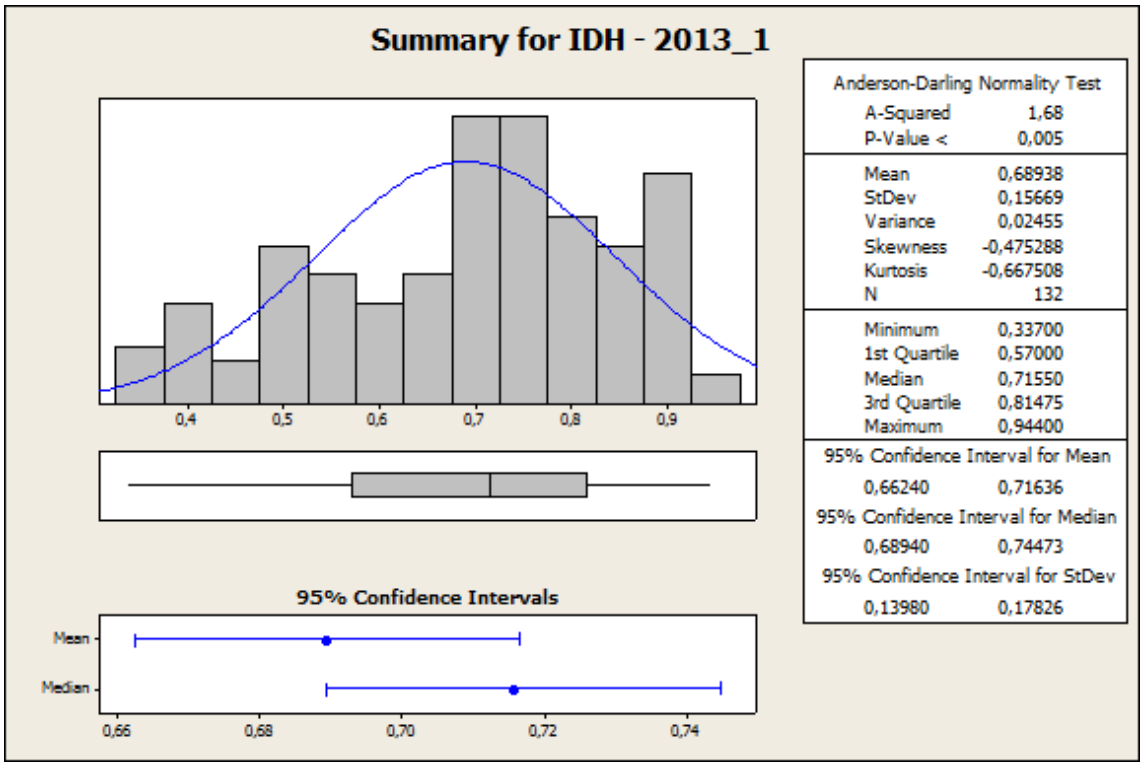
Fonte: Autor – Dados Guia GPS.

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



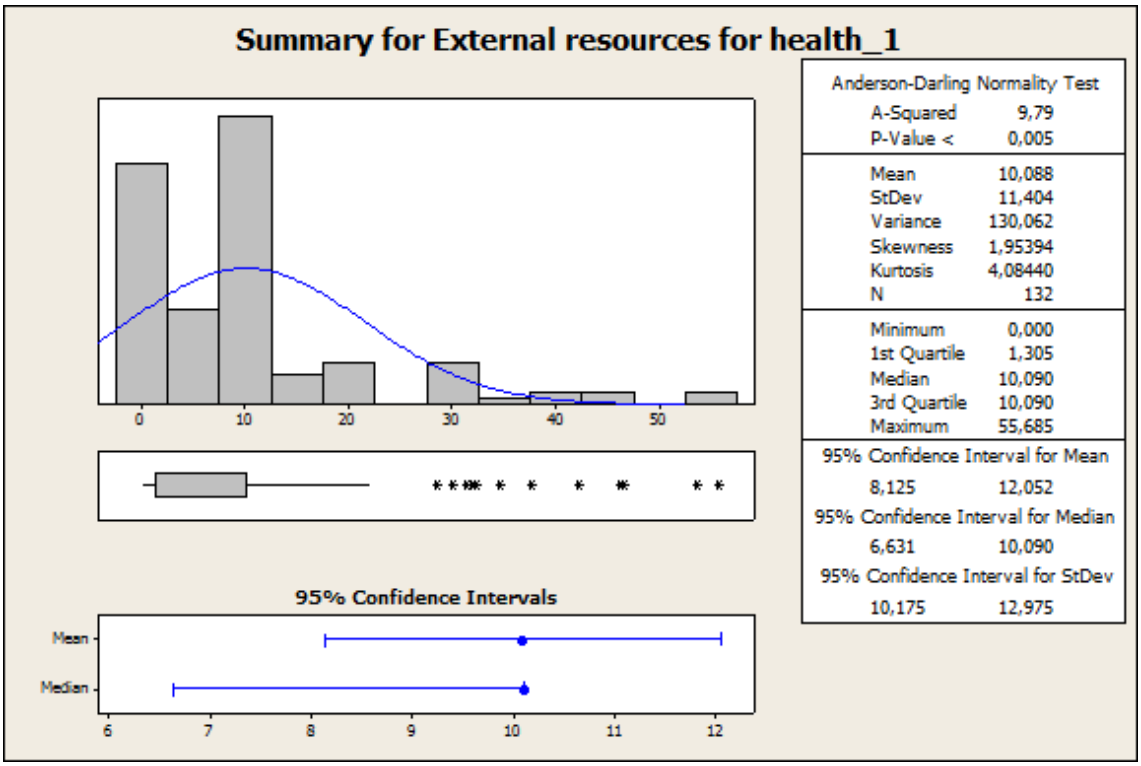
Fonte: Autor – Dados Guia GPS.

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*



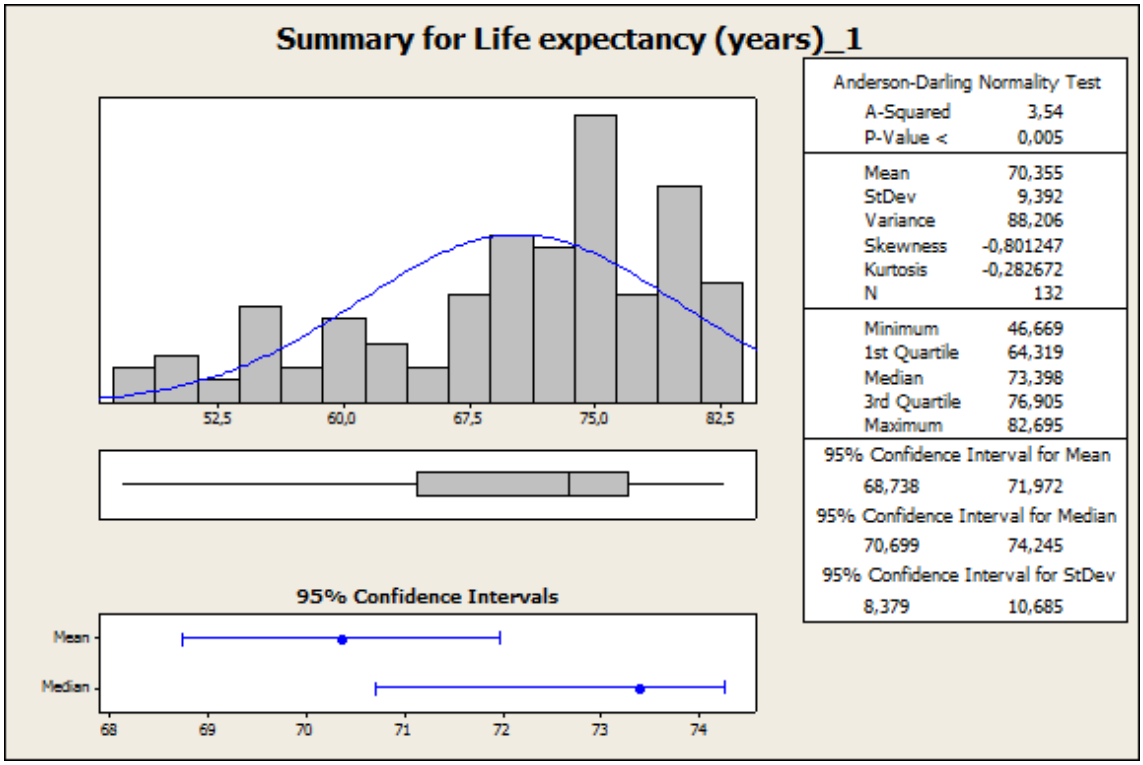
Fonte: Autor – Dados Guia GPS.

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

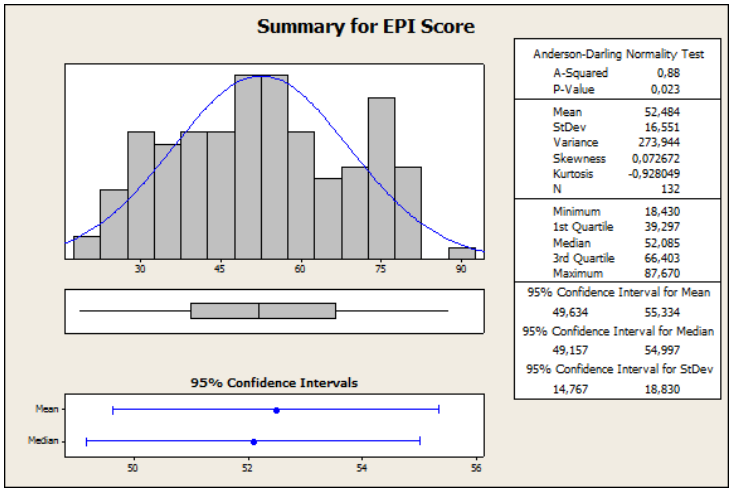
Fonte: Elaborado pela autora

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)



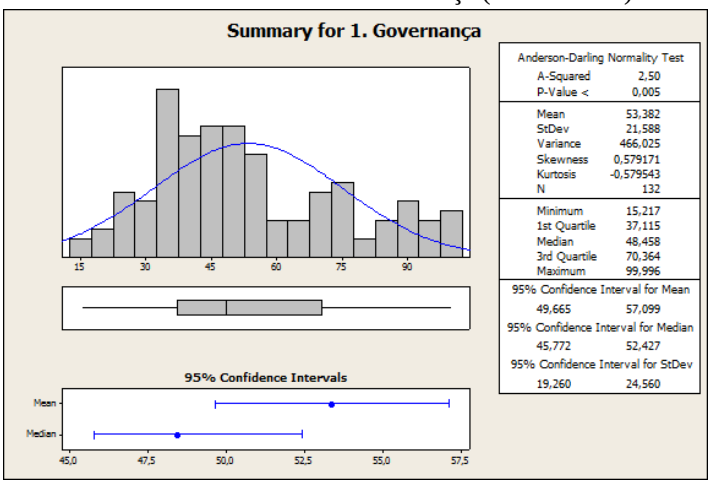
Fonte: Autor – Dados Guia GPS.

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)

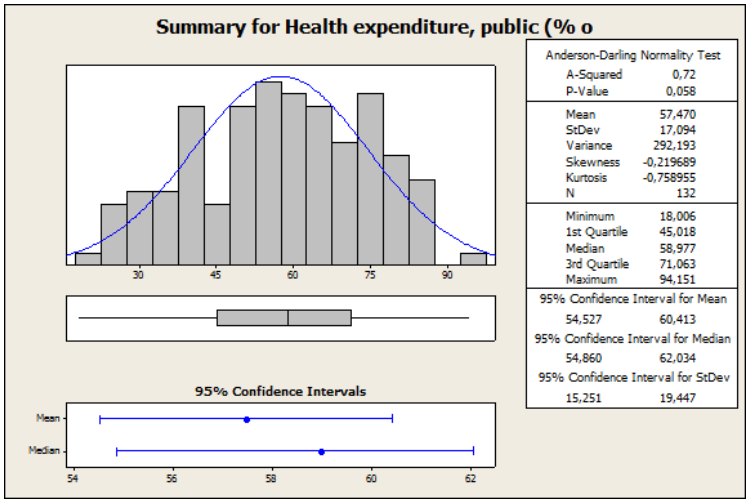


Fonte: Autor – Dados Guia GPS.

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)

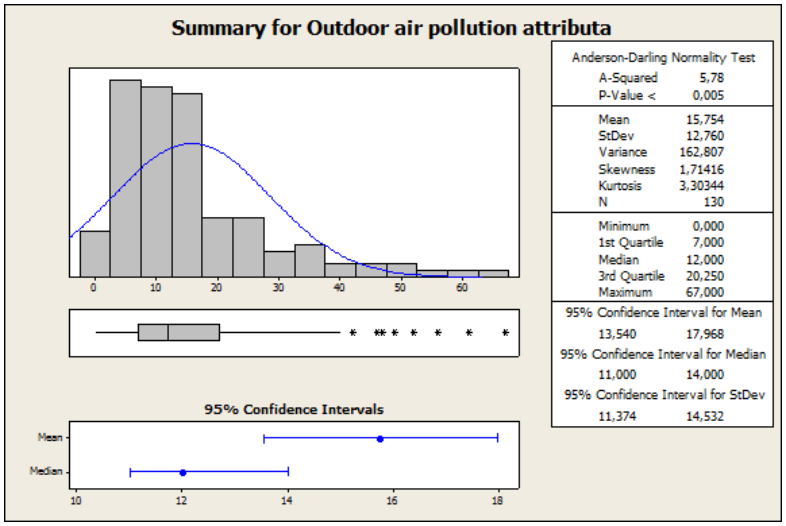


Fonte: Autor – Dados Guia GPS.

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)

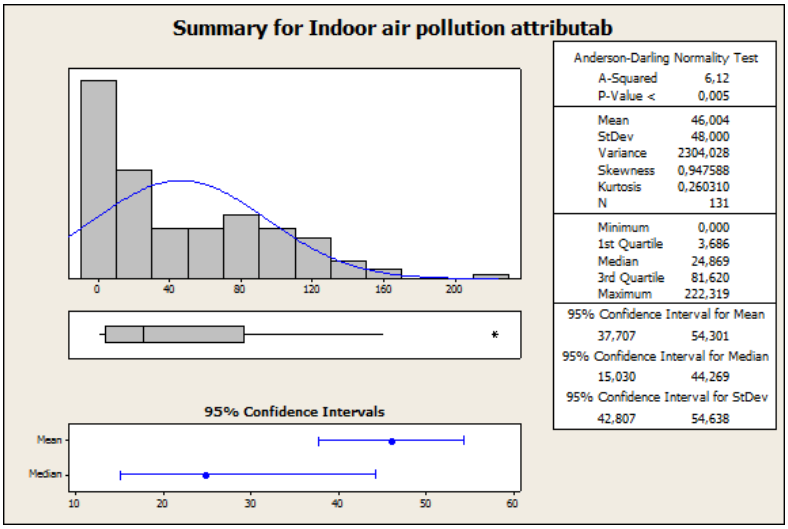


Fonte: Autor – Dados Guia GPS.

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)

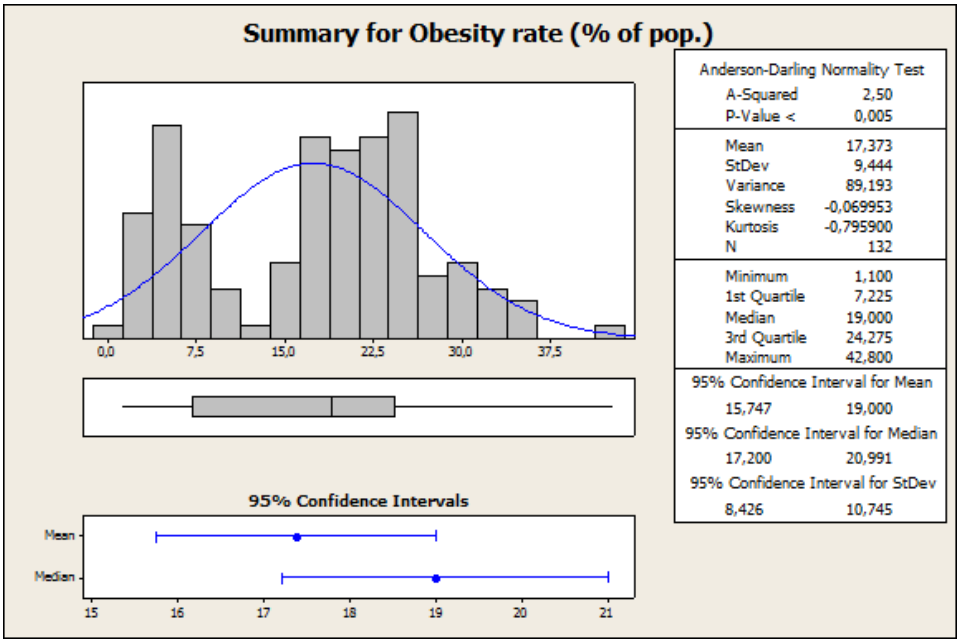


Fonte: Autor – Dados Guia GPS.

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)

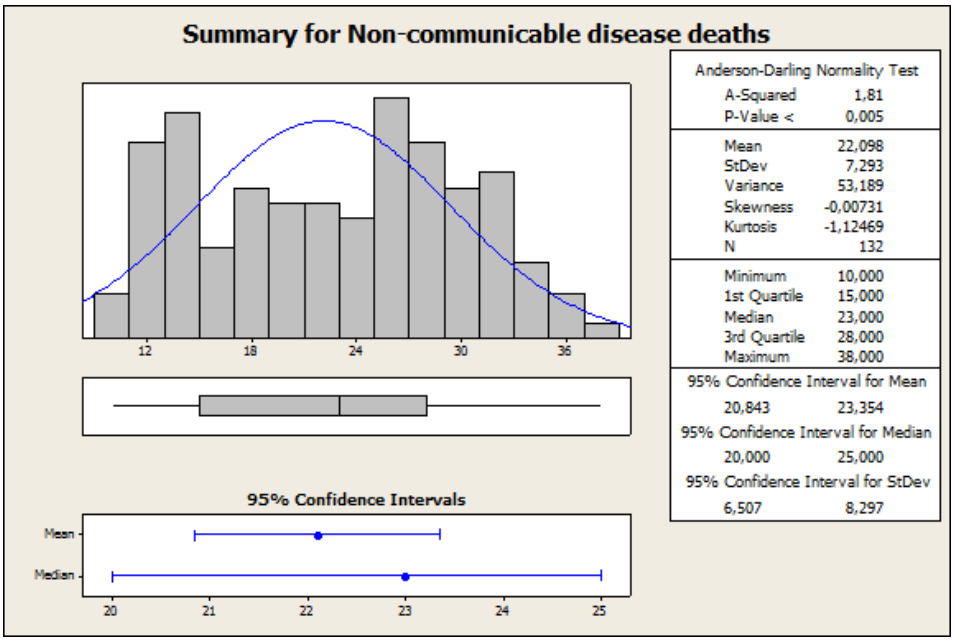


Fonte: Autor – Dados Guia GPS.

Análise:

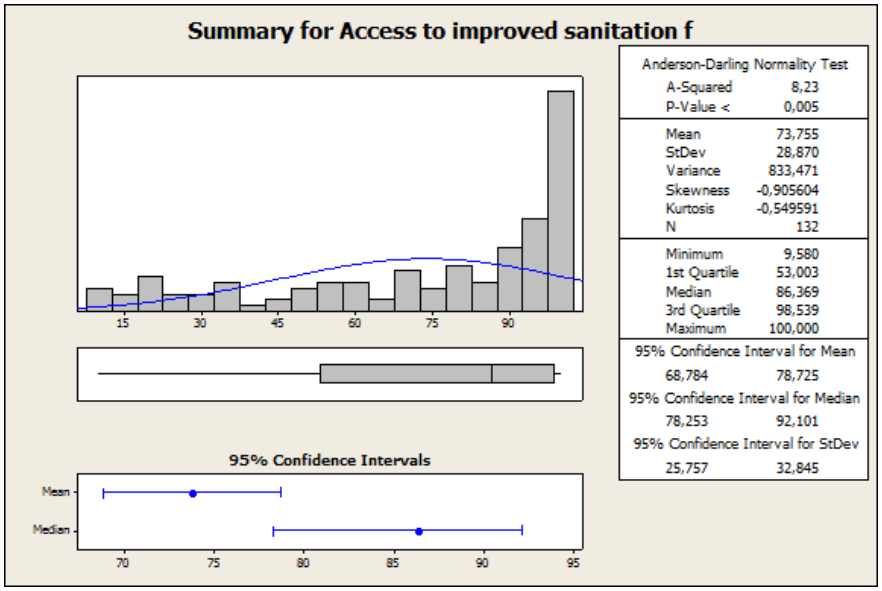
Aqui se apresenta uma análise interessante, onde praticamente podemos ver quase "duas distribuições" de dados... Um dos substratos de 0 a 14, representados basicamente por África e Ásia, e uma segunda com países de Europa, Austrália e Nova Zelandia, Ibéricos e Estados Unidos. Na parte superior voltam 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)

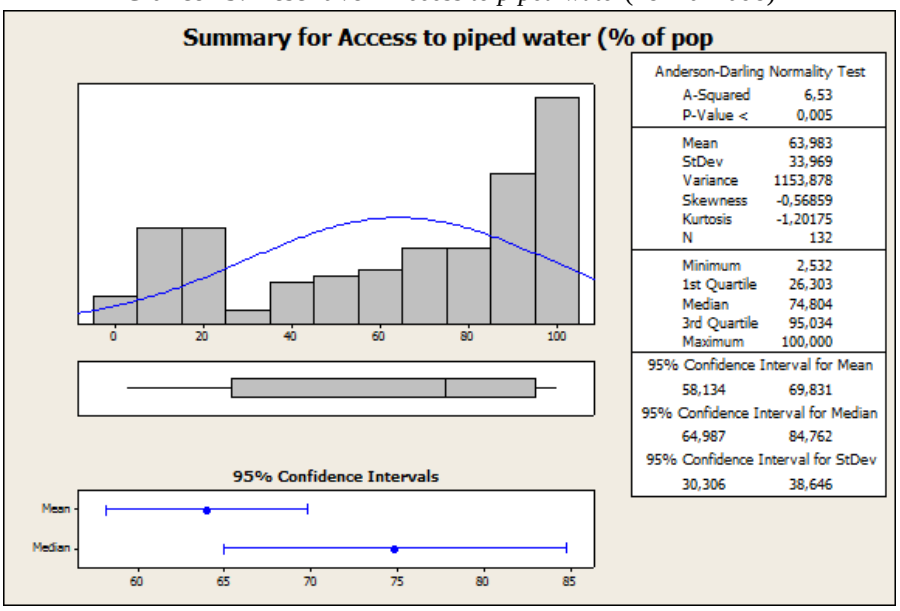


Fonte: Autor – Dados Guia GPS.

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)

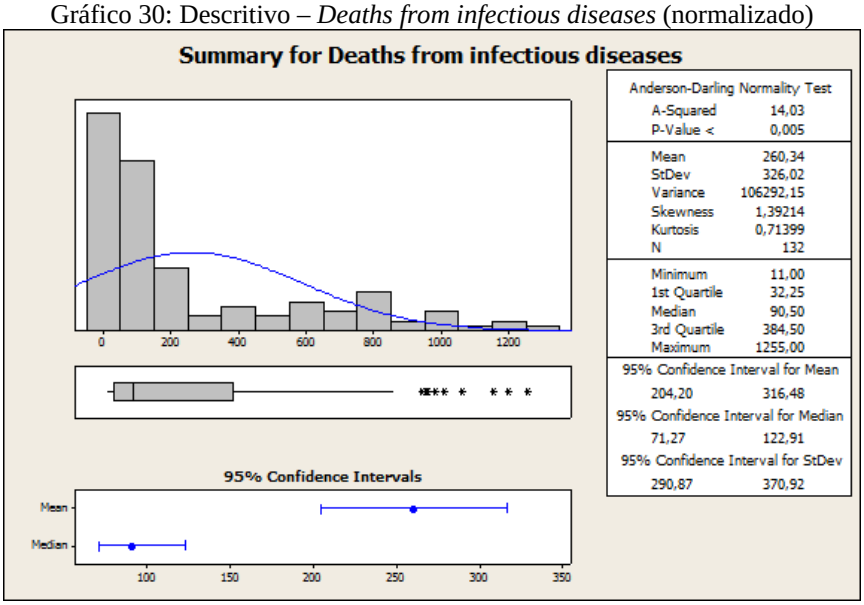


Fonte: Autor – Dados Guia GPS.

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)



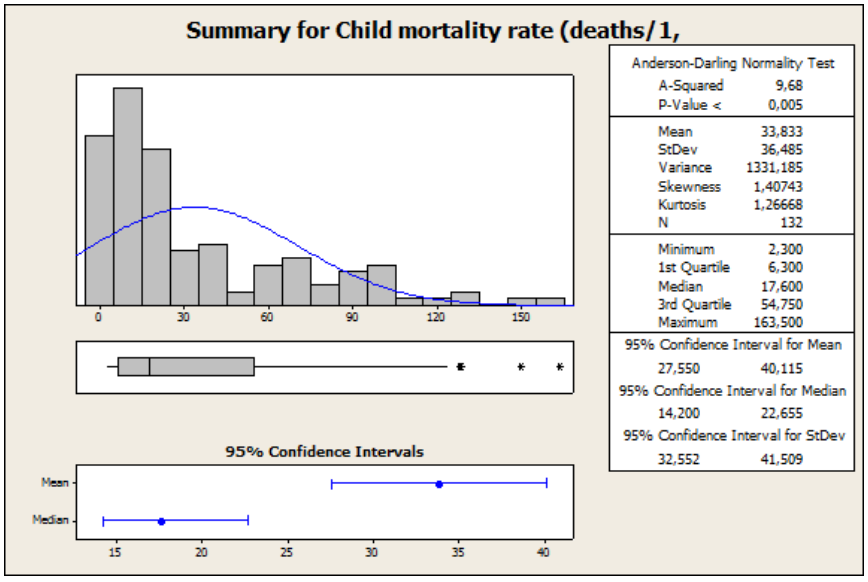
Fonte: Autor – Dados Guia GPS.

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.**

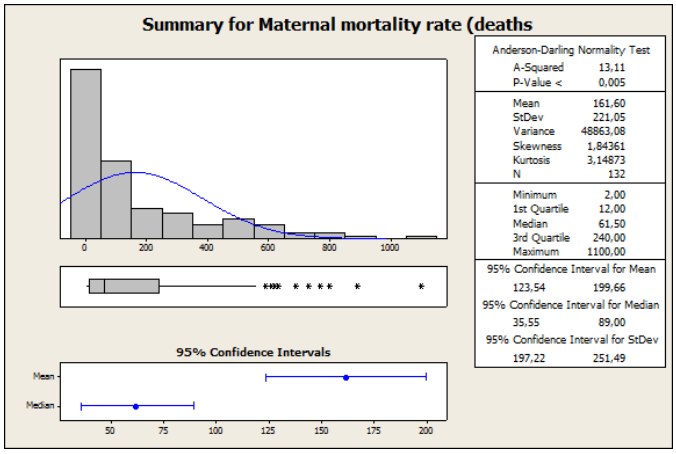
O MESMO VALE PARA OS PRÓXIMOS 3 GRÁFICOS – MORTALIDADE INFANTIL, MATERNAL MORTALITY RATE E STILLBIRTH RATE

Gráfico 31: Descritivo – *Child Mortality* (normalizado)



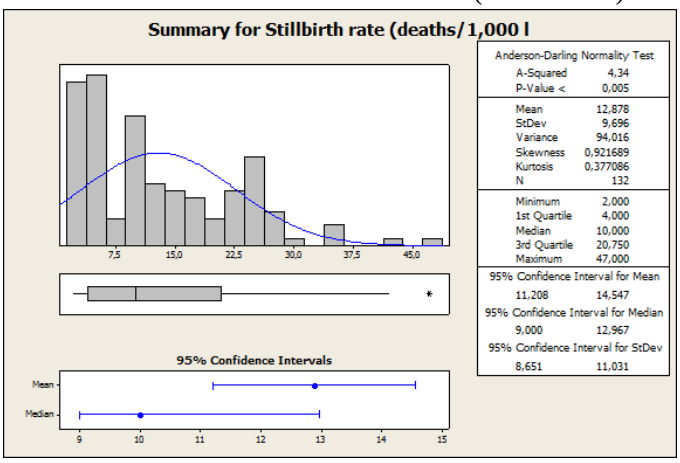
Fonte: Autor – Dados Guia GPS.

Gráfico 32: Descritivo – *Maternal Mortality* (normalizado)



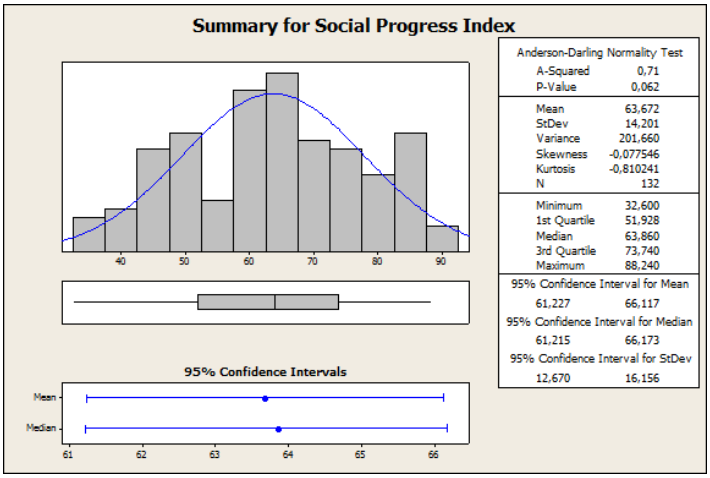
Fonte: Autor – Dados Guia GPS.

Gráfico 33: Descritivo – *Stillbirth rate* (normalizado)



Fonte: Autor – Dados Guia GPS.

Gráfico 34: Descritivo – *Social Progress Index* (normalizado)



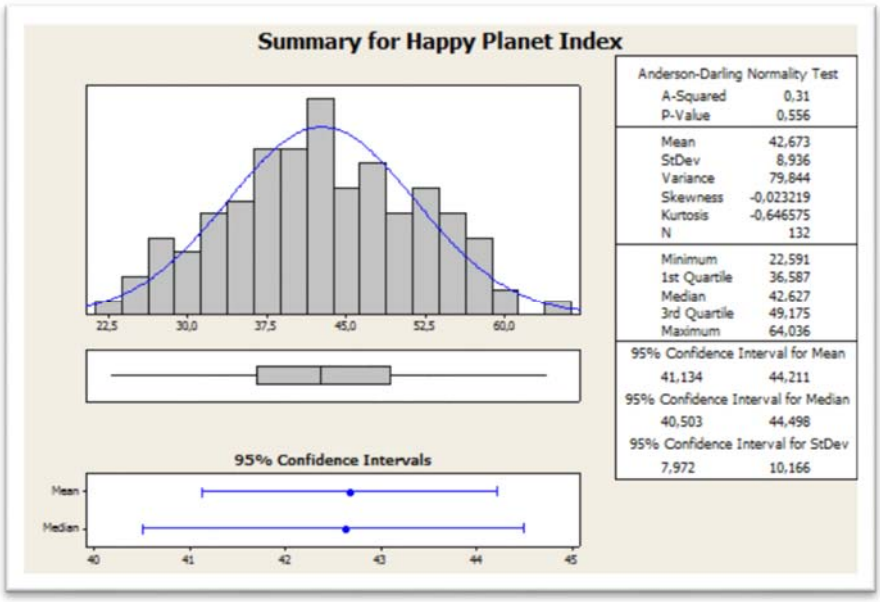
Fonte: Autor – Dados Guia GPS.

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

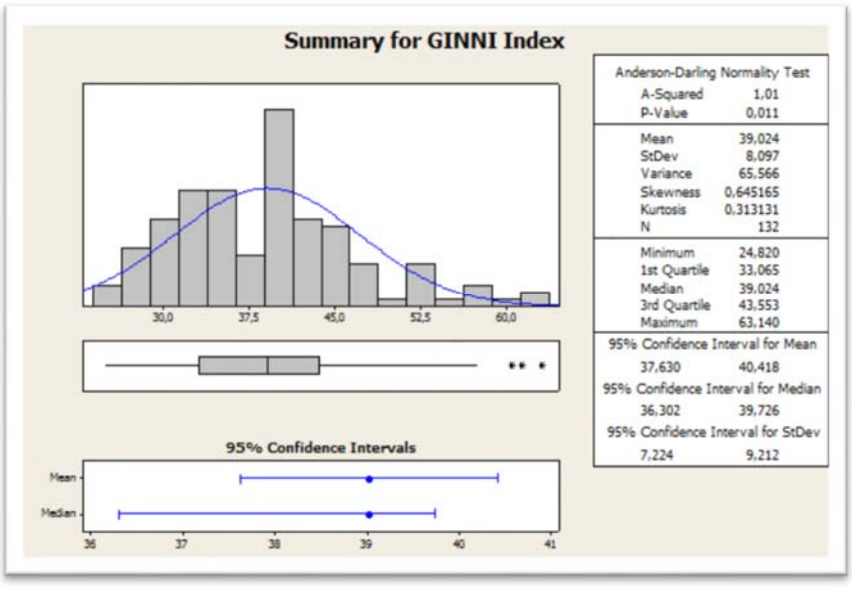
diferenciados (55 com poucas mostras e 85 com muitas), porém o Planeta feliz segue bastante a distribuição normal incluindo os substratos.

Gráfico 35: Descritivo – *Happy Planet Index* (normalizado)



Fonte: Autor – Dados Guia GPS.

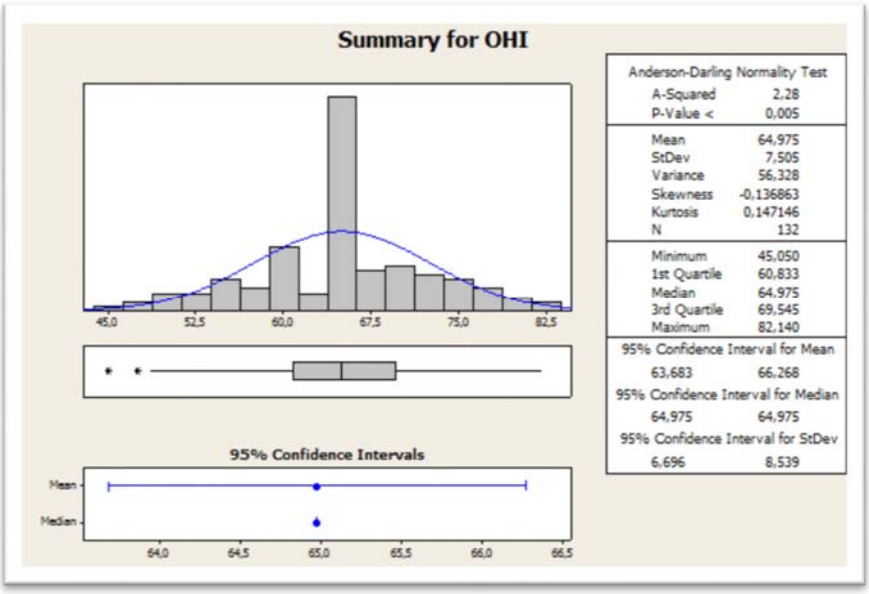
Gráfico 36: Descritivo – *Ginni* (normalizado)



Fonte: Autor – Dados Guia GPS.

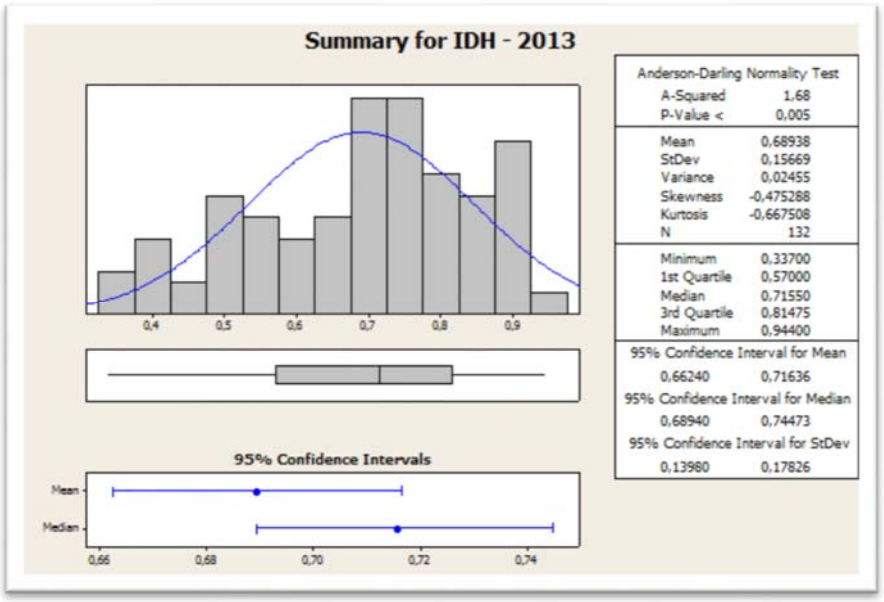
Análise: Para o Índice de Ginni, houve um pico cuja mediana foi de 39,024, equivalente à média. Tem uma distribuição bem normalizada, com uma maior quantidade de países à direita.

Gráfico 37: Descritivo – *OHI* (normalizado)



Fonte: Autor – Dados Guia GPS.

Gráfico 38: Descritivo – *IDH* (normalizado)

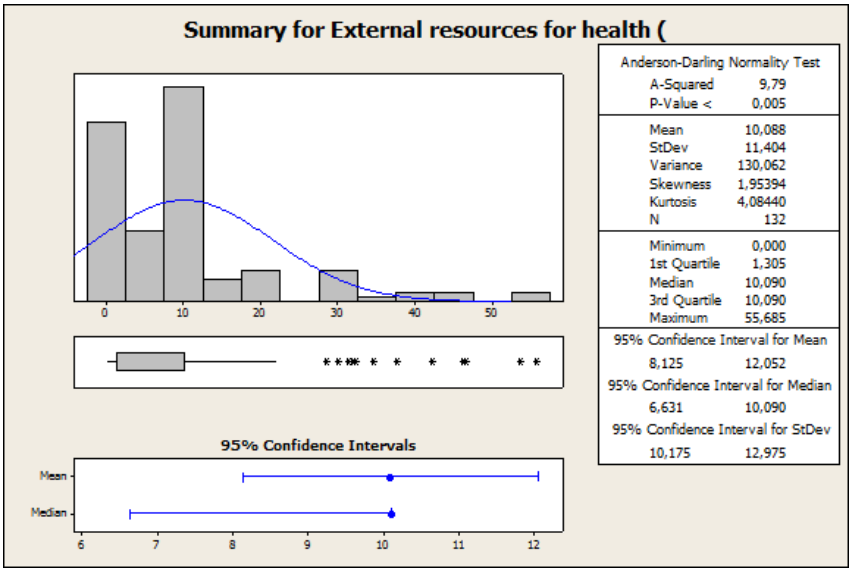


Fonte: Autor – Dados Guia GPS.

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)

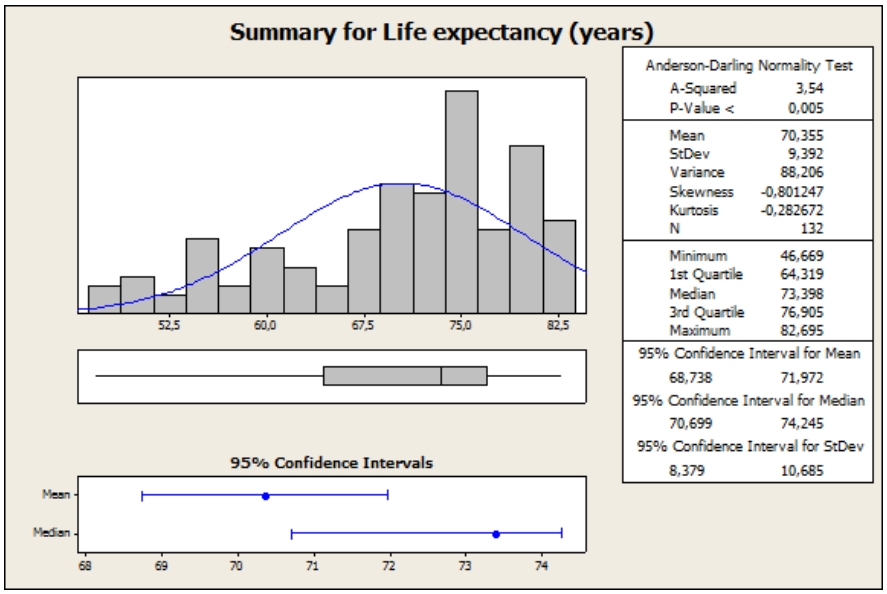


Fonte: Autor – Dados Guia GPS.

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

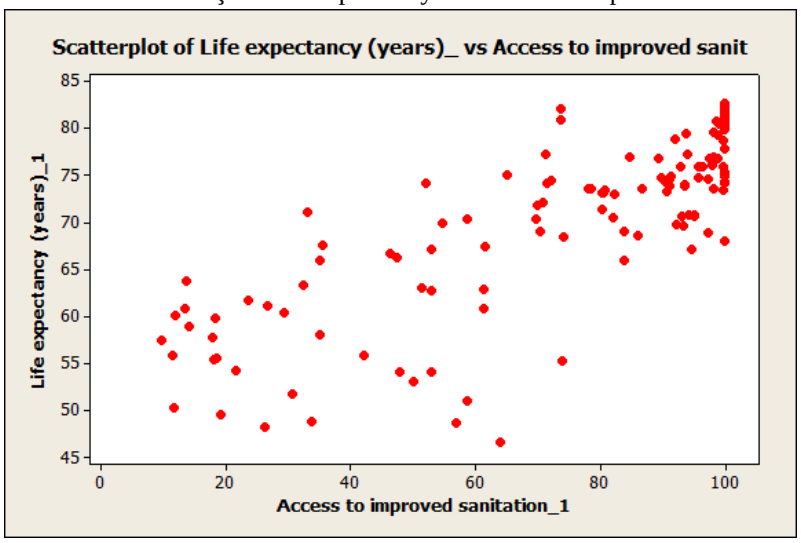
Fonte: Preparado pela autora

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.

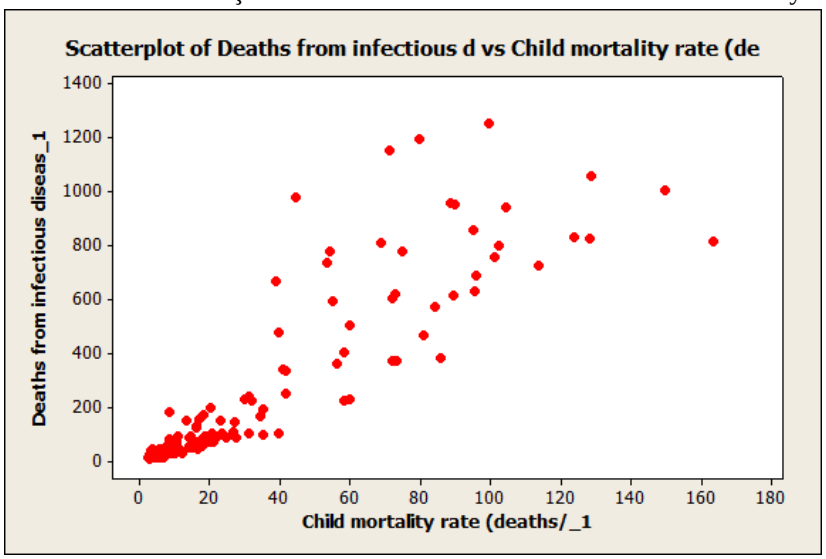
Gráfico 41: Correlação Life Expectancy vs Access to Improved Sanitation



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

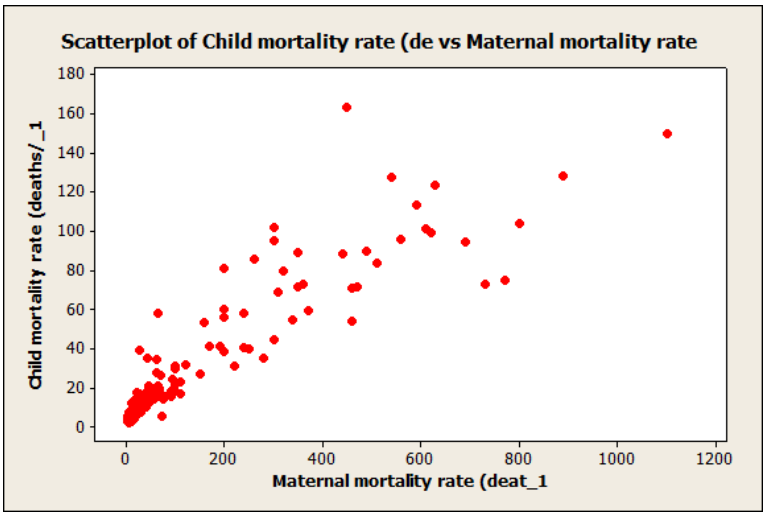


Fonte: Autor – Dados Guia GPS.

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.

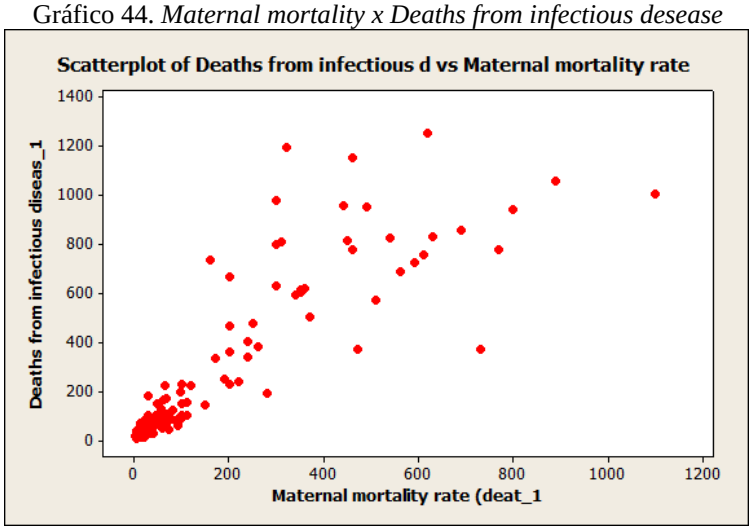
Gráfico 43. *Maternal mortality x child mortality*



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.

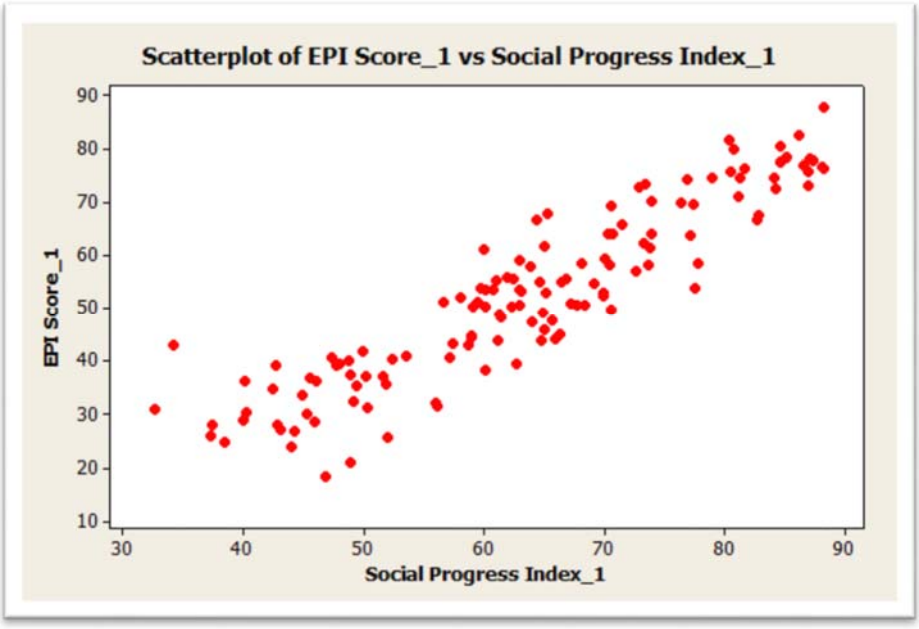
Gráfico 44. *Maternal mortality x Deaths from infectious desease*



Fonte: Autor – Dados Guia GPS.

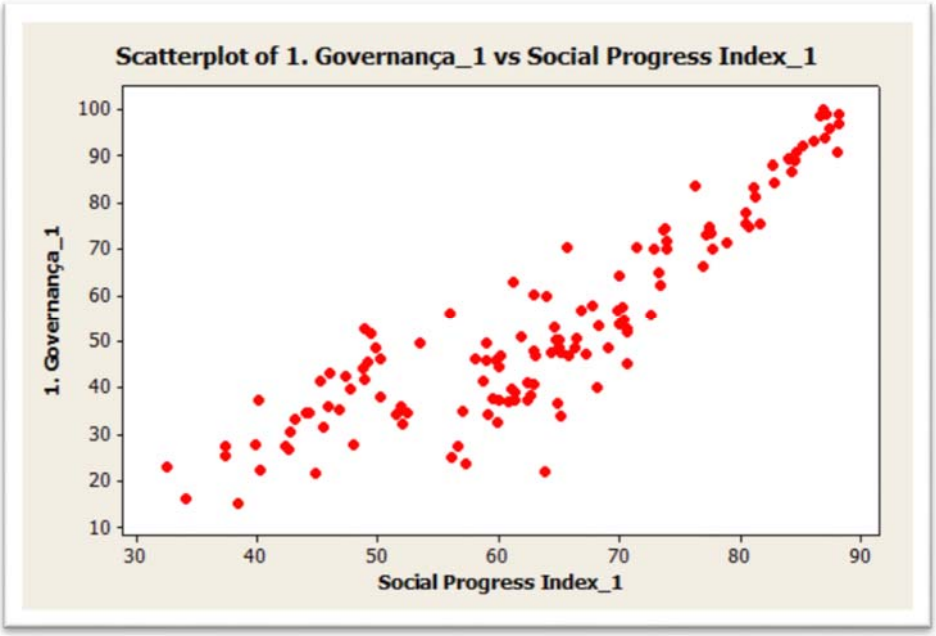
Para os próximos 4 gráficos (45, 46 e 47)(Social progress x EPI, Governança x Social Progress e IDH vs. Social PRogress) vemos uma forte correlação linear entre o Índice de progressão social e as variáveis IDH, governança e EPI score. Algumas com mais dispersão ou outliers que outras, porém todos demonstram claramente uma relação linear crescente.

Gráfico 45. *EPI x Social Progress*



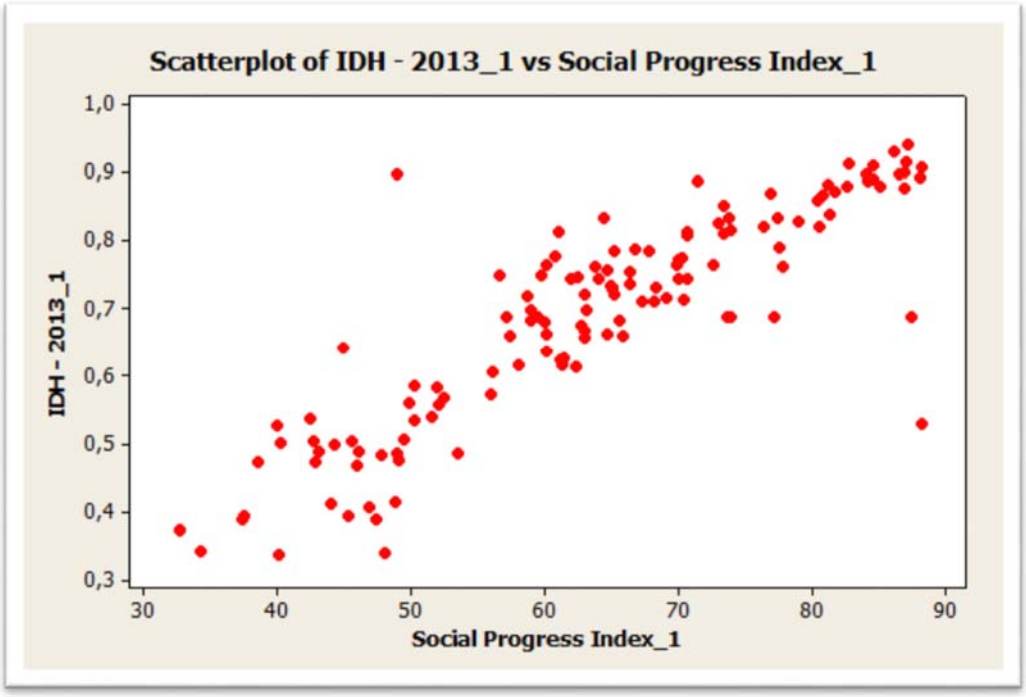
Fonte: Autor – Dados Guia GPS.

Gráfico 46. Governança *x Social Progress*



Fonte: Autor – Dados Guia GPS.

Gráfico 47. IDH *x Social Progress*



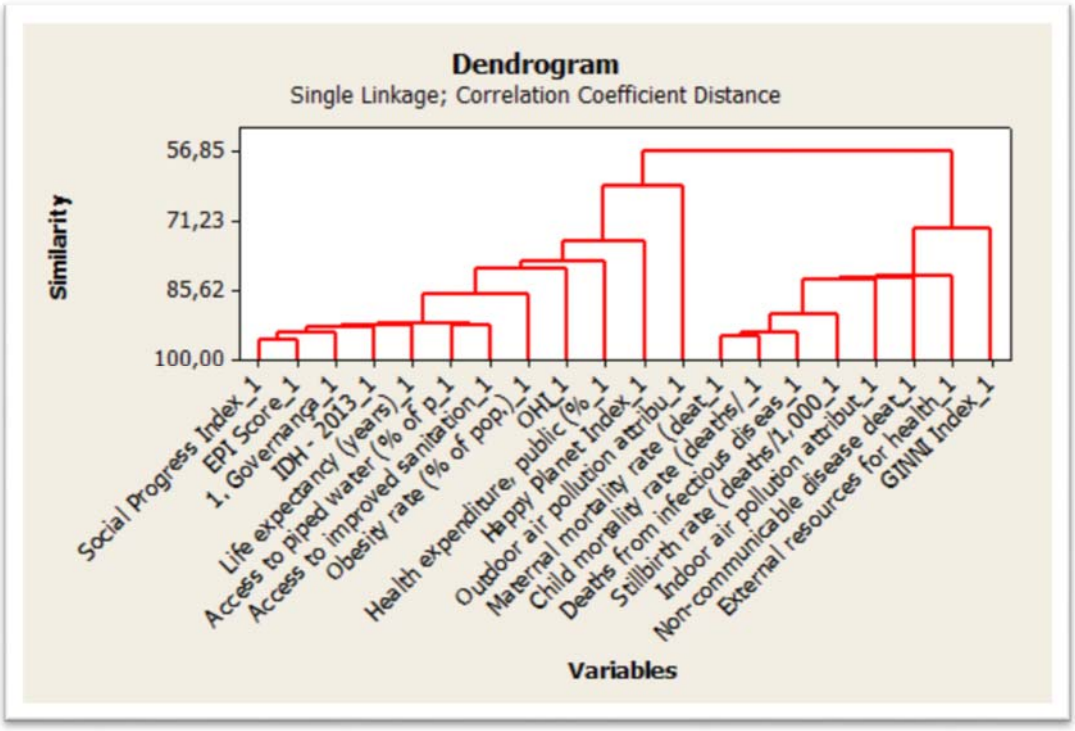
Fonte: Autor – Dados Guia GPS.

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.

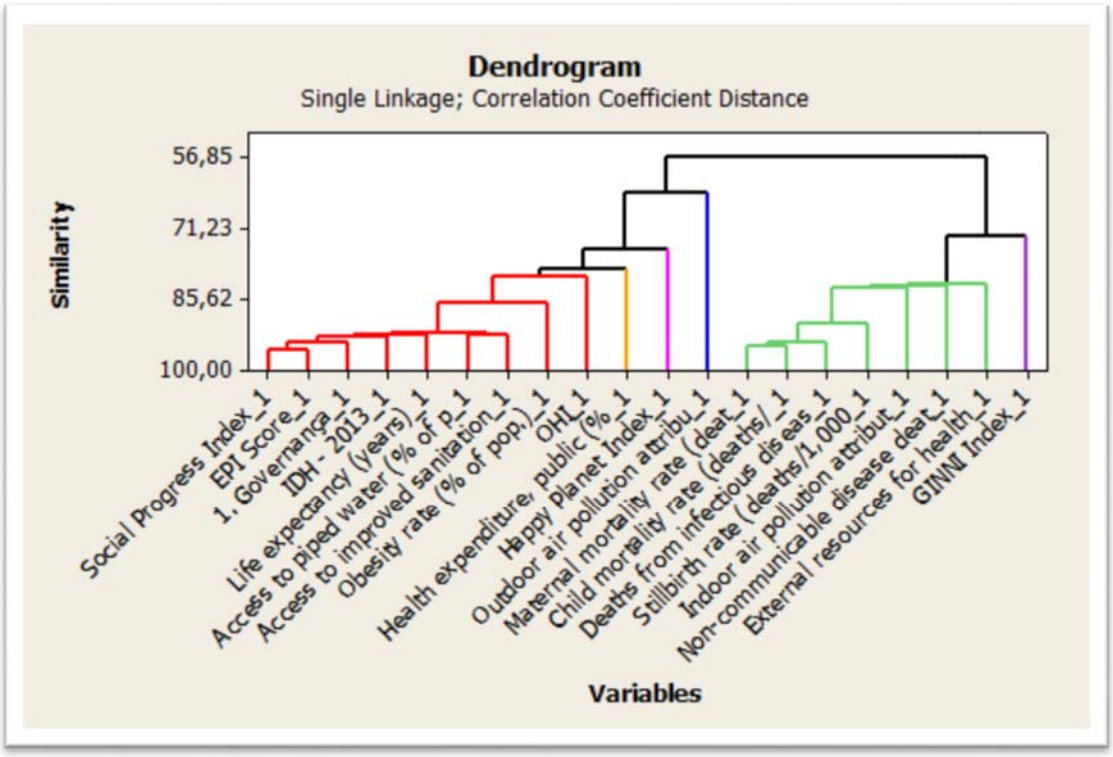
Gráfico 48 – Dendograma das Variáveis, Cluster Único



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.

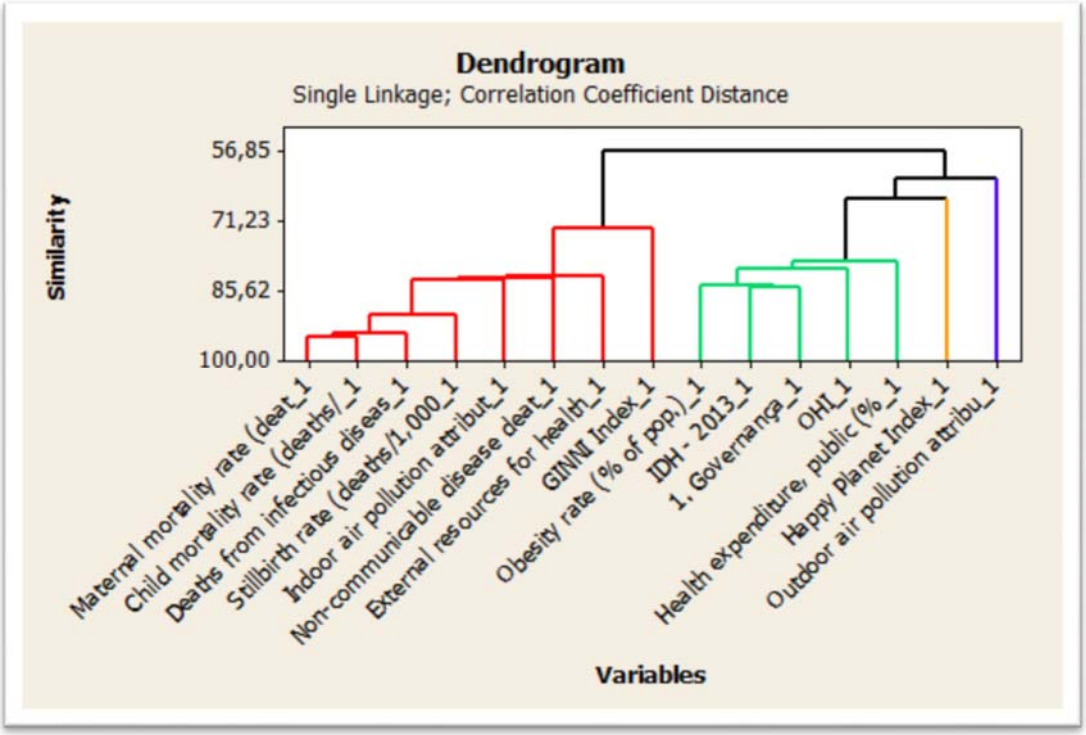
Gráfico 49 – Dendograma das Variáveis, 6 Clusters



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

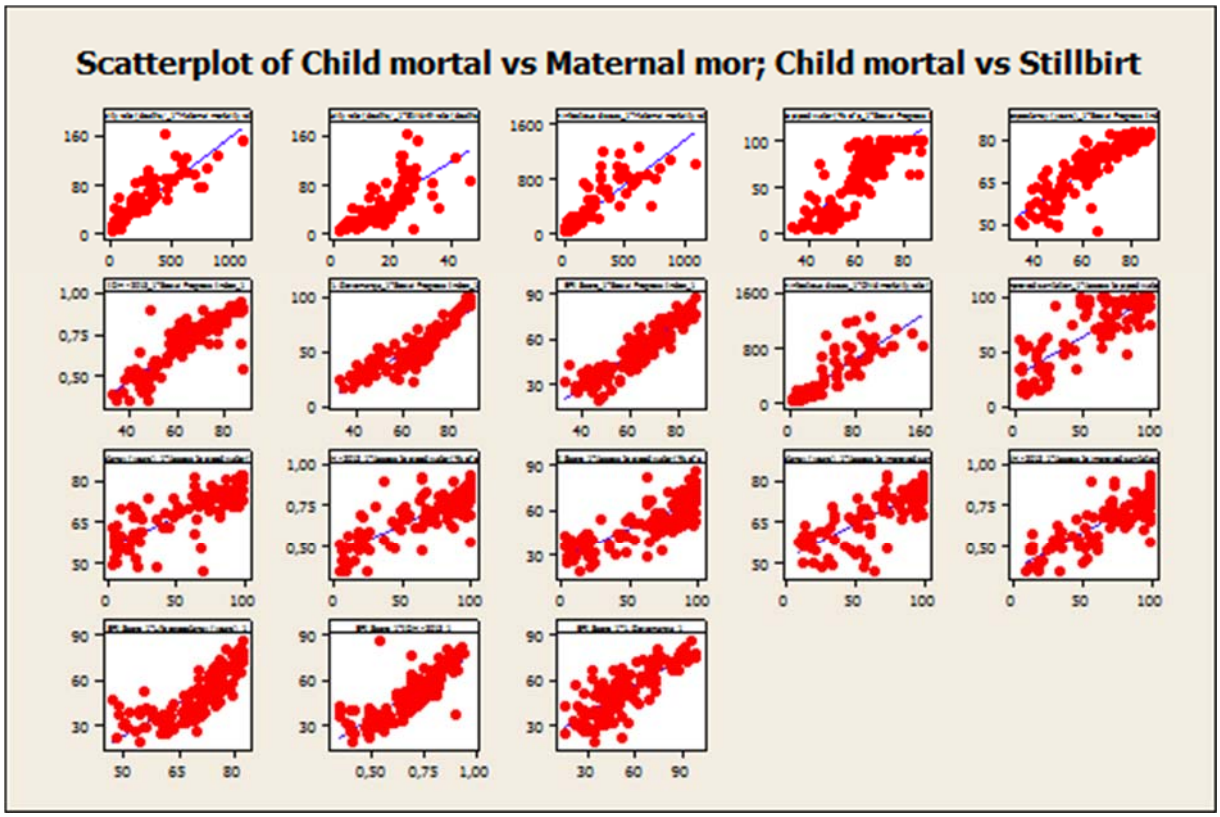
A correlação mede a direção e a intensidade da relação linear (linha reta) entre duas variáveis quantitativas. Se um diagrama de dispersão mostra uma relação linear, é interessante resumir esse padrão geral traçando uma reta no diagrama de dispersão. Uma reta de regressão resume a relação entre duas 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

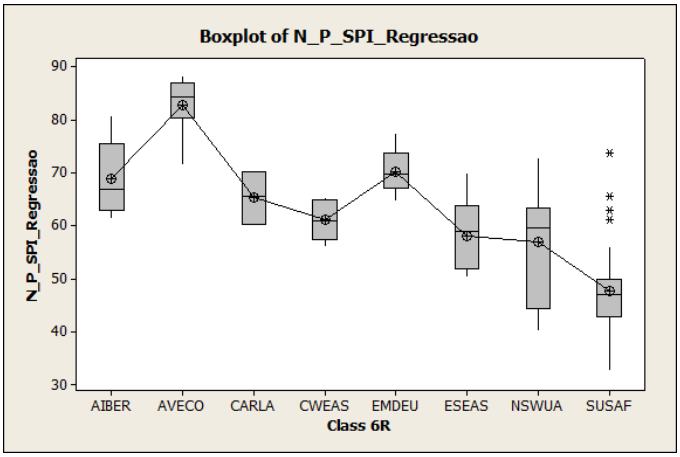
Nova Sigla	Nomenclatura da Região
AIBER	Ibero-Americanos
AVECO	Economias Avançadas
CARLA	América Latina e o Caribe
CWEAS	Comunidade dos Estados Independentes
ESEAS	Emergentes e em Desenvolvimento da Ásia
EMDEU	Emergentes e em Desenvolvimento da Europa
NSWUA	Oriente Médio, Norte da África e Paquistão
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
```

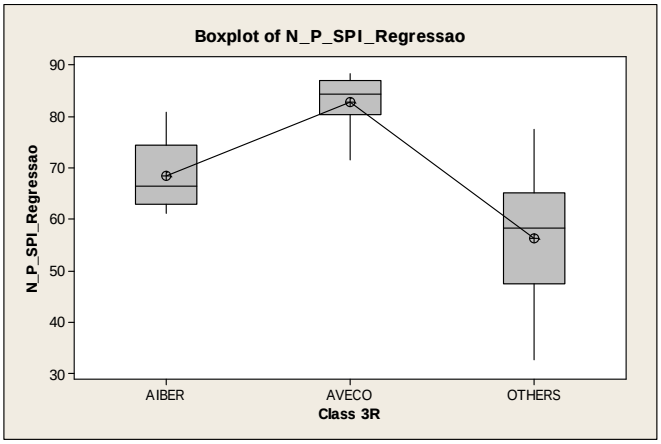
Análise:

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).

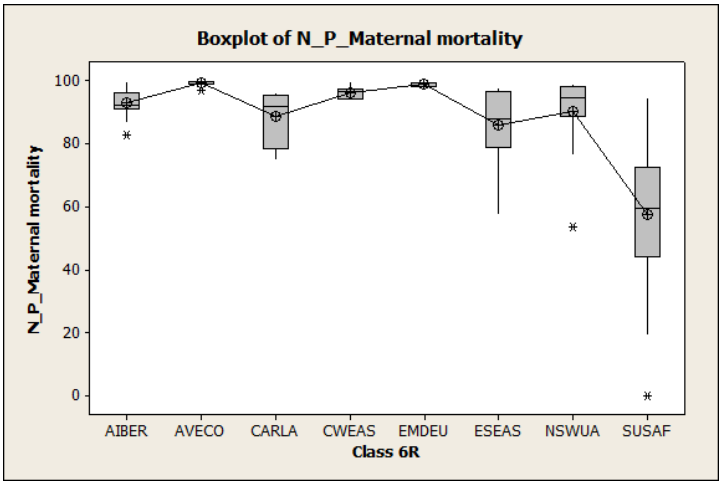
A análise gráfica e a análise por intervalo de confiança sugerem que a região que possui o maior índice é a AVECO e a menor é a OTHERS. Porém, é necessária uma análise mais específica para comparar tal fato.

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.

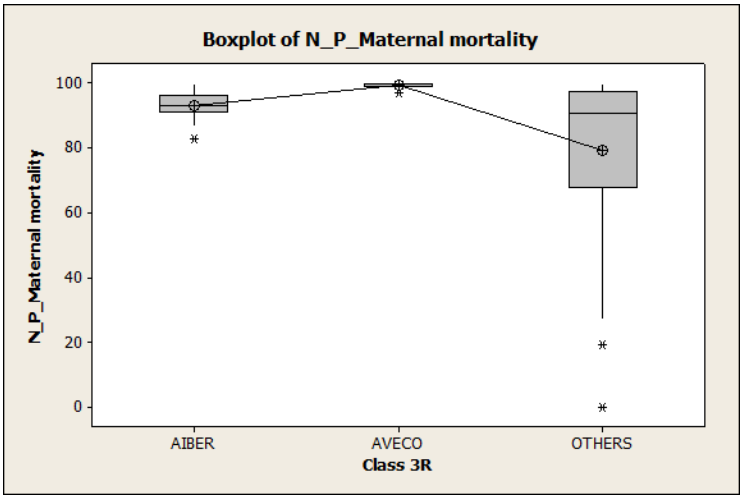
Menor índice: SUSAF.

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.

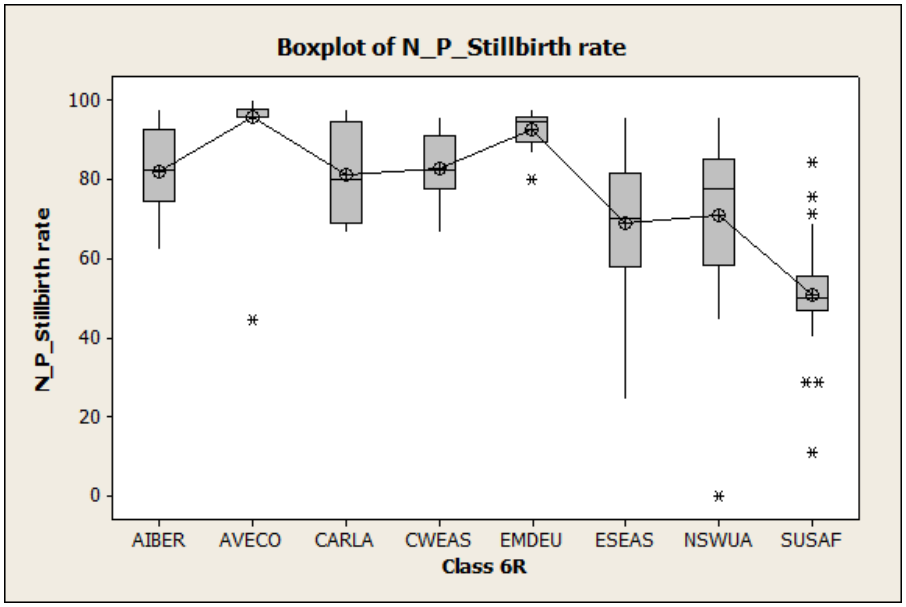
Há outliers nas 3 regiões.

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%
```

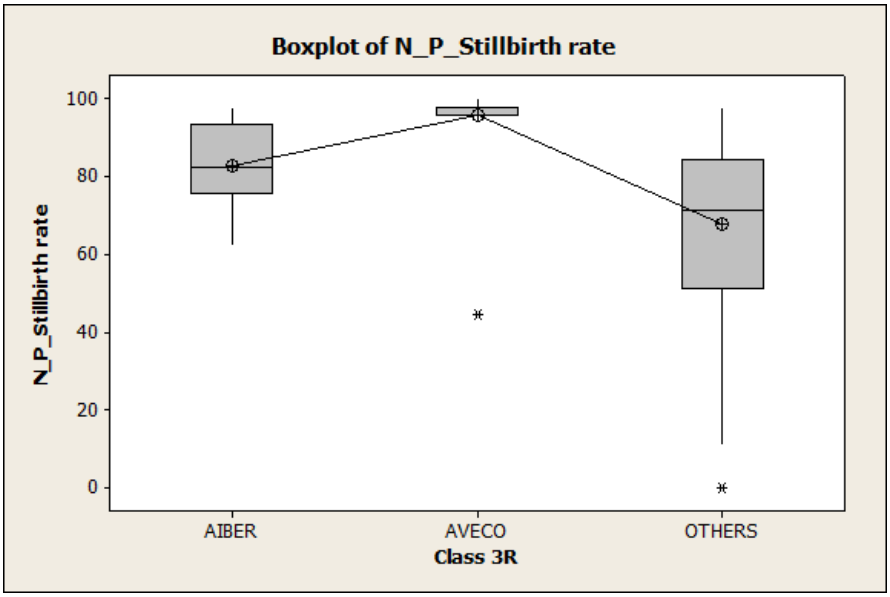
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   ---------+---------+---------+---------+
AIBER  20  82,00  10,52                     (---*---)
AVECO  27  95,64  10,34                               (---*--)
CARLA   4  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.

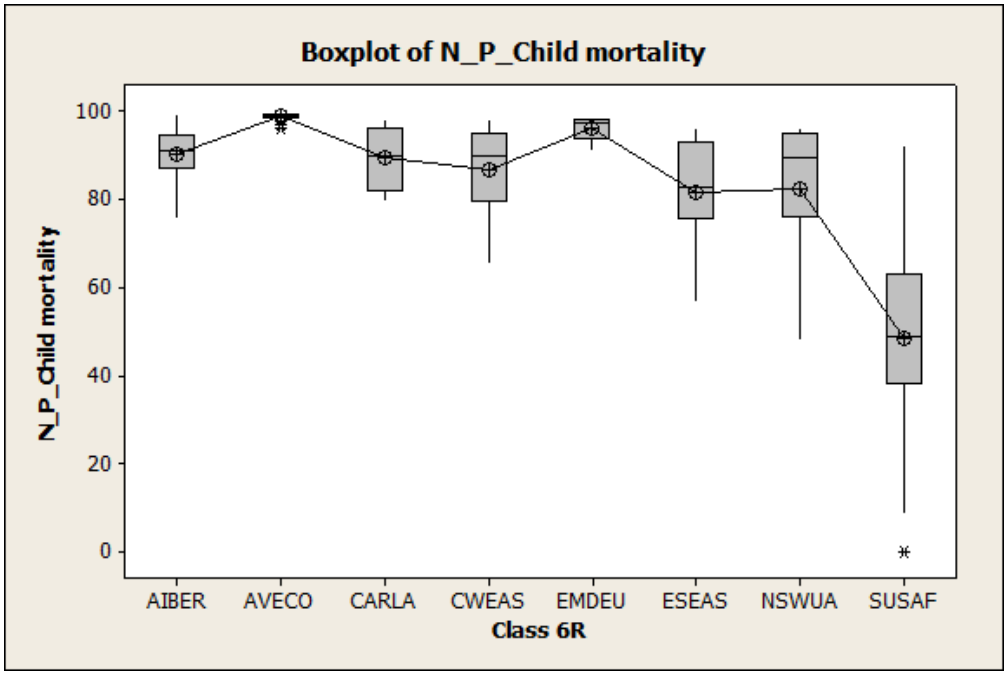
Não há outlier na AIBER.

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.

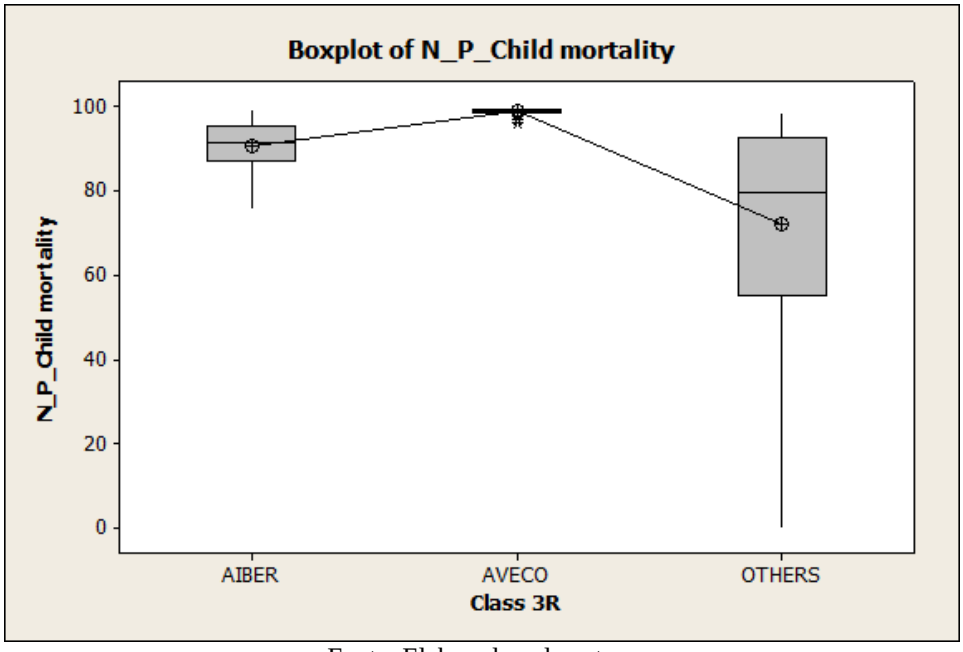
Menor índice: SUSAF.

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.

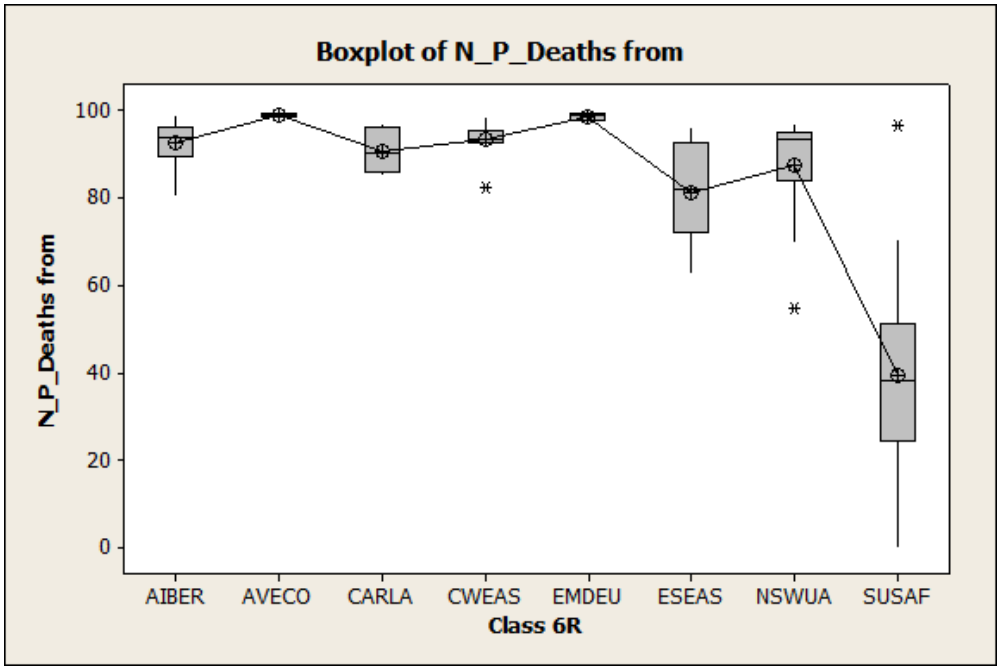
Há outlier na AVECO.

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.

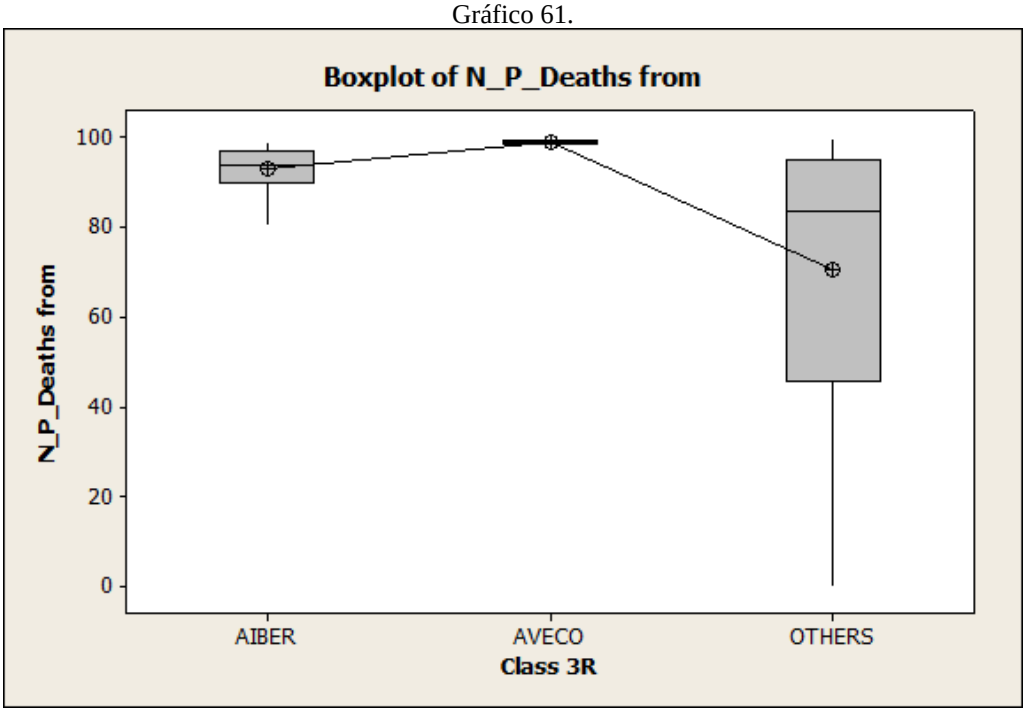
Menor índice: SUSAF.

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.

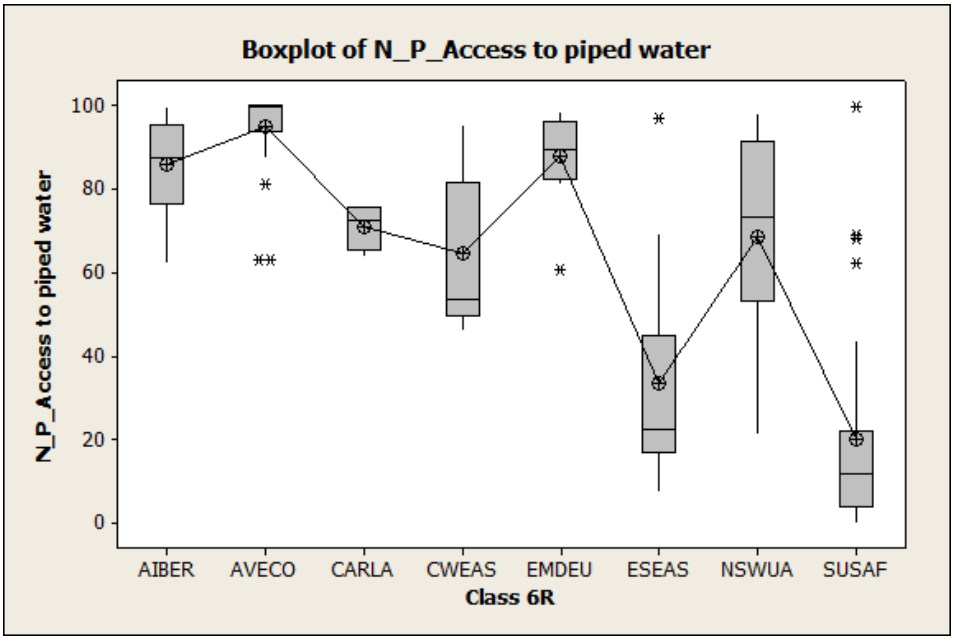
Menor índice: OTHERS.

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

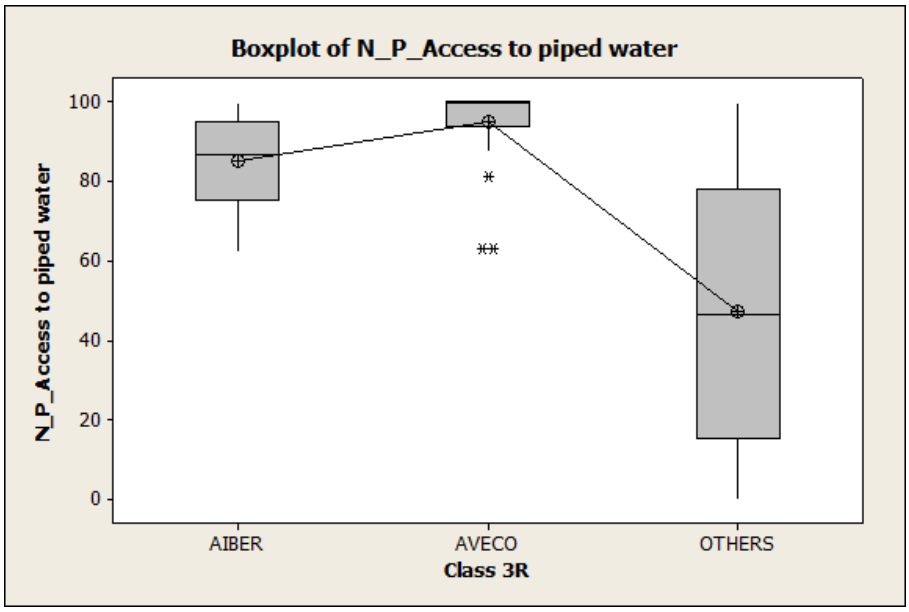
```
Source     DF      SS     MS      F      P
Class 6R    7  115589  16513  47,04  0,000
Error     124   43525    351
Total     131  159113

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.

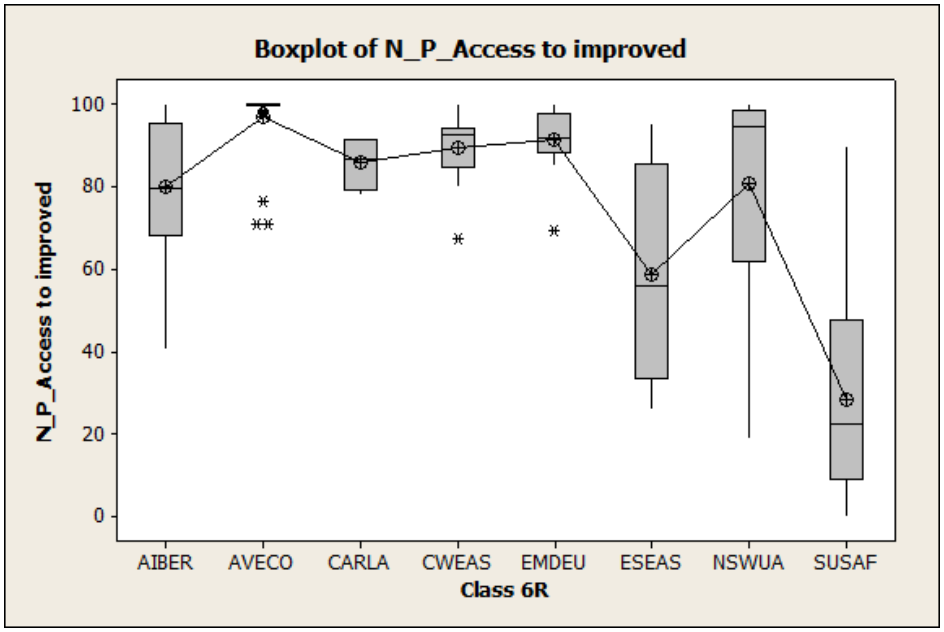
Há outliers na AVECO.

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.

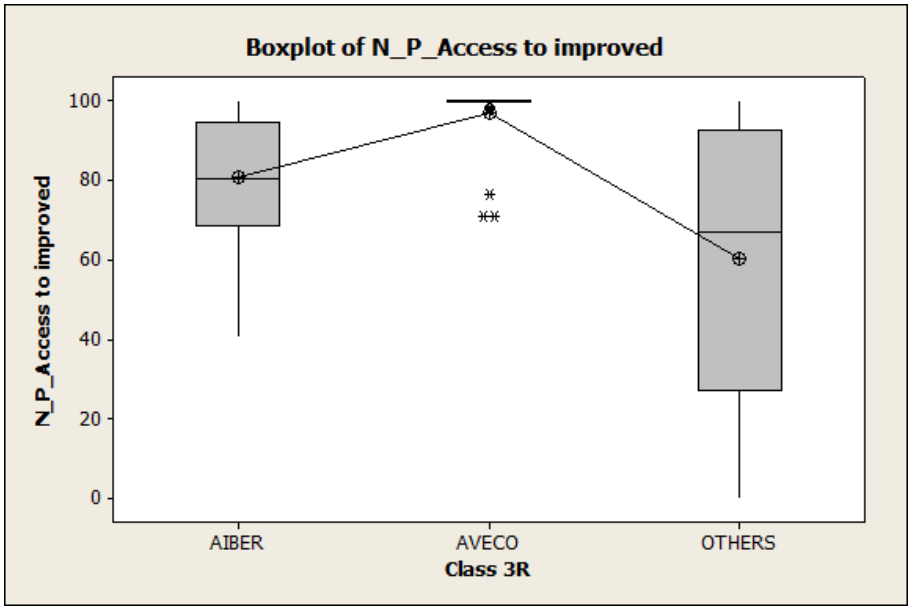
Menor índice: SUSAF.

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.

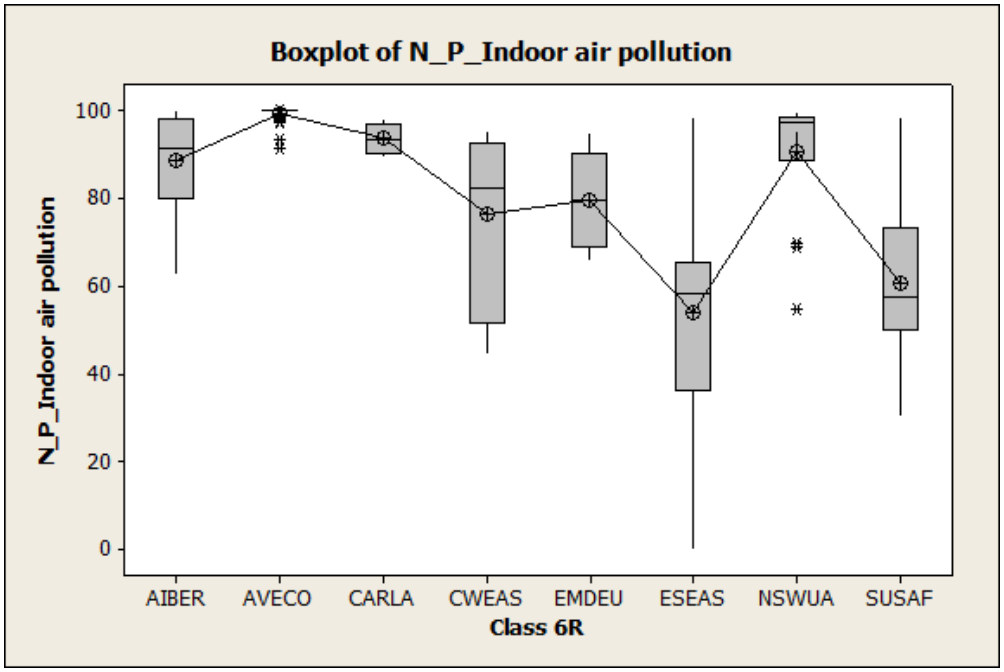
Há outliers na AVECO.

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.

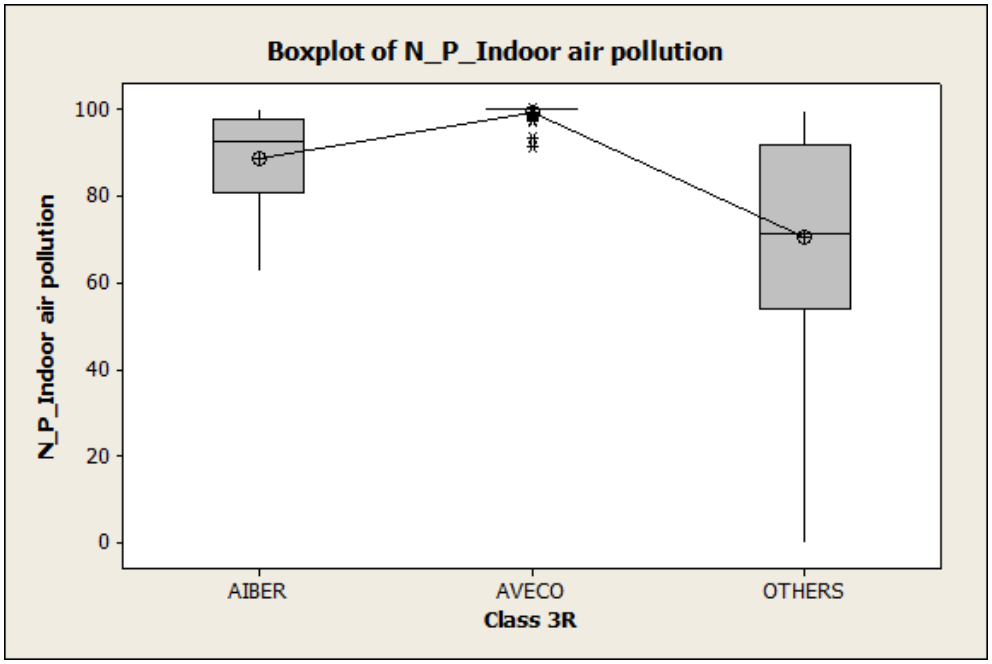
Menor índice: Não definido.

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

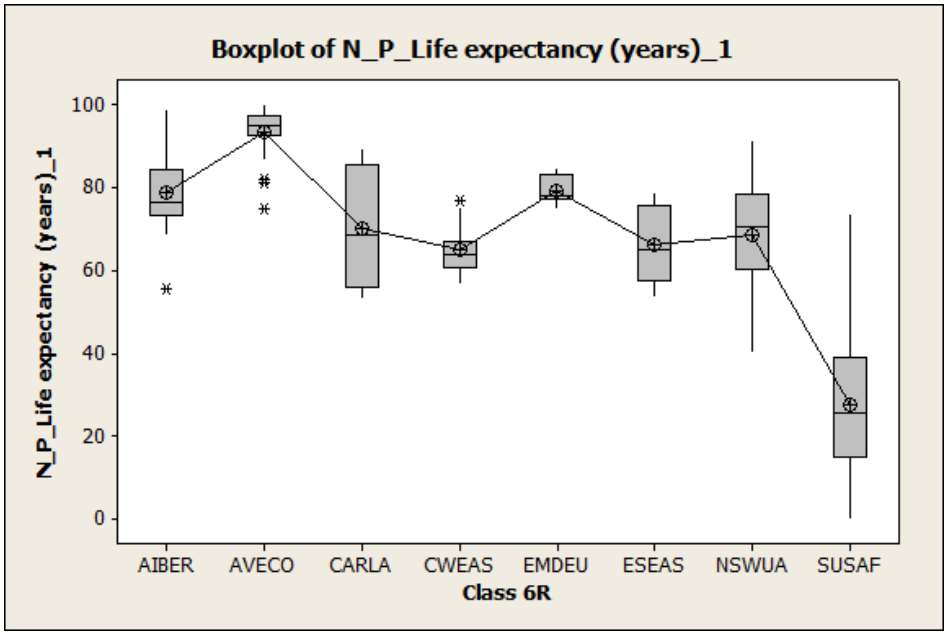
```
Source     DF     SS    MS      F      P
Class 3R    2  18982  9491  29,42  0,000
Error     129  41619   323
Total     131  60601

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.

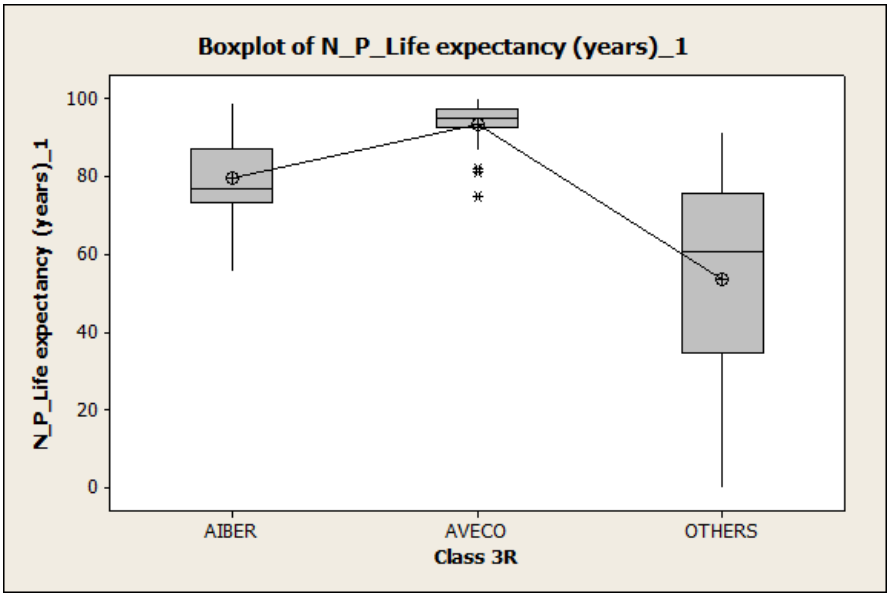
Menor índice: SUSAF.

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.

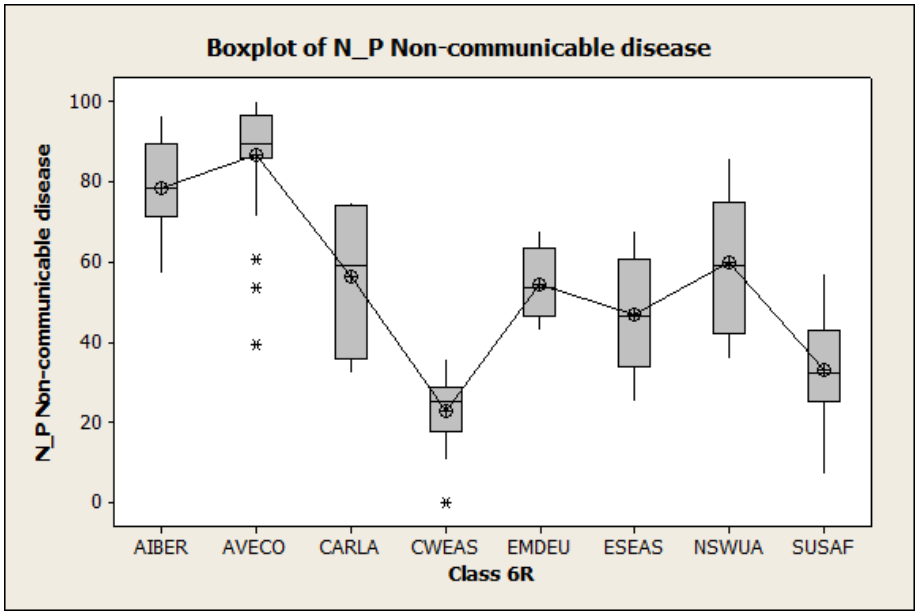
Há outliers na AVECO.

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.

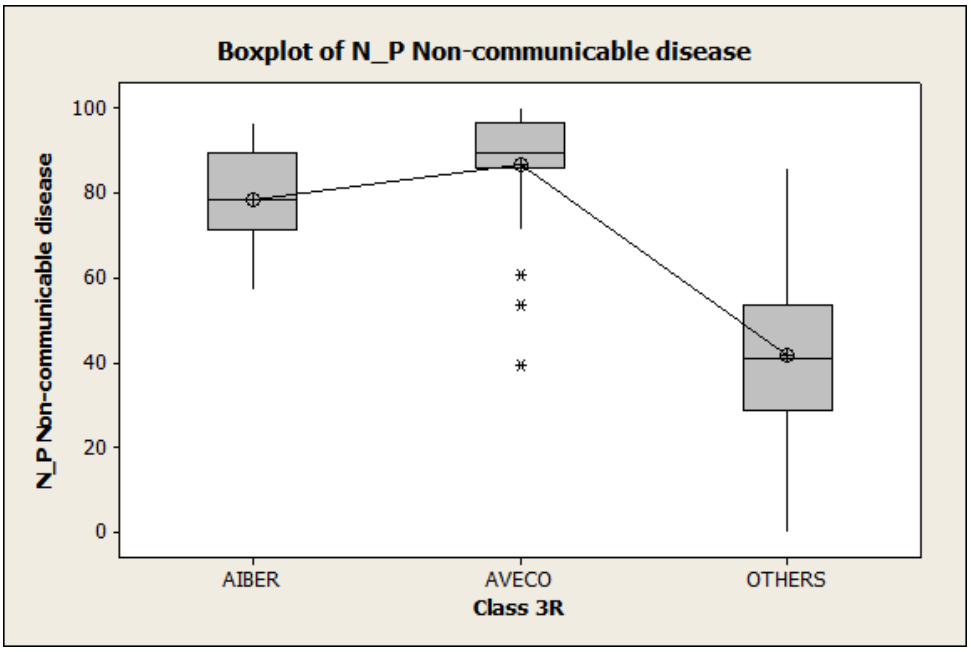
Menor índice: CWEAS.

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.

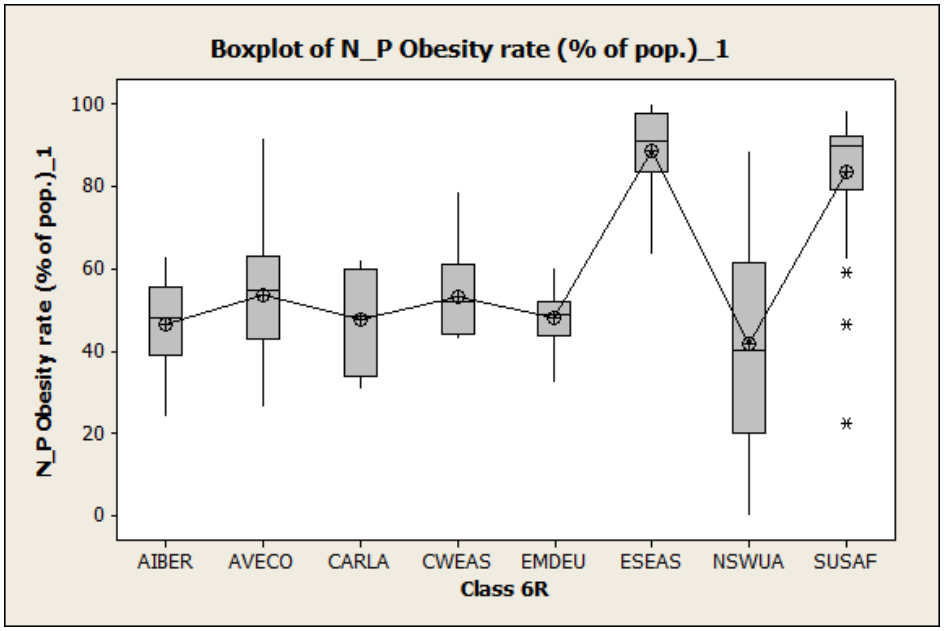
Há outliers na AVECO.

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.

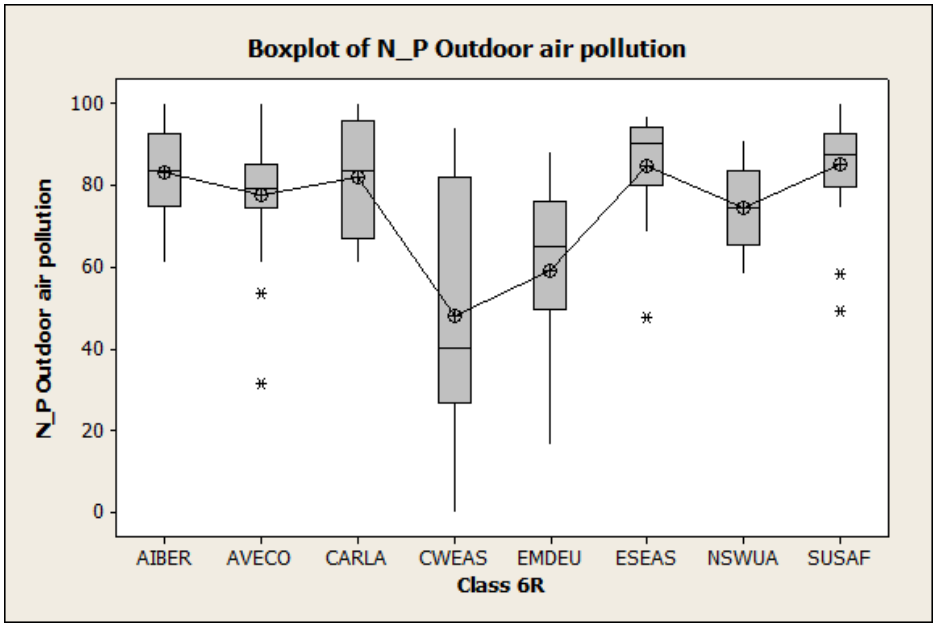
Menor índice: Não definido.

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.

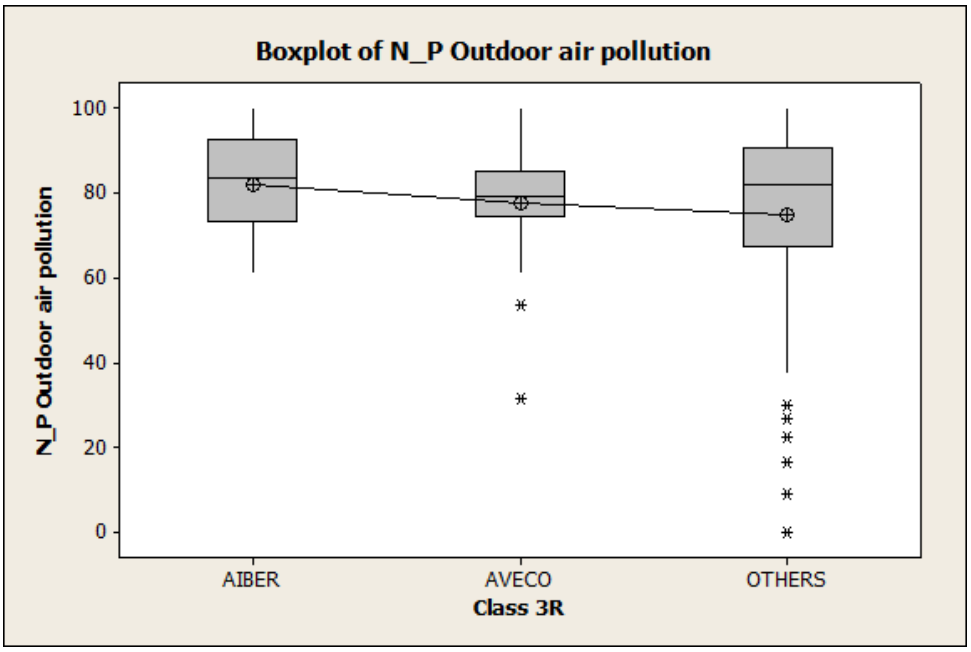
Menor índice: Não definido.

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

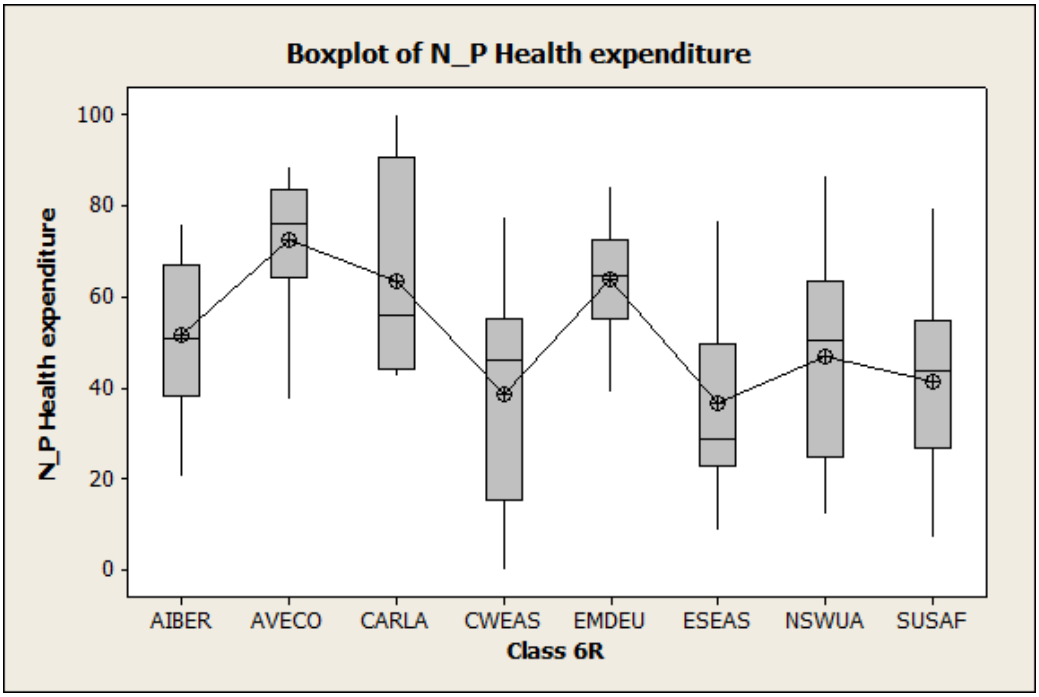
```
Source     DF     SS   MS     F      P
Class 3R    2    937  469  1,32  0,271
Error     129  45848  355
Total     131  46786

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).

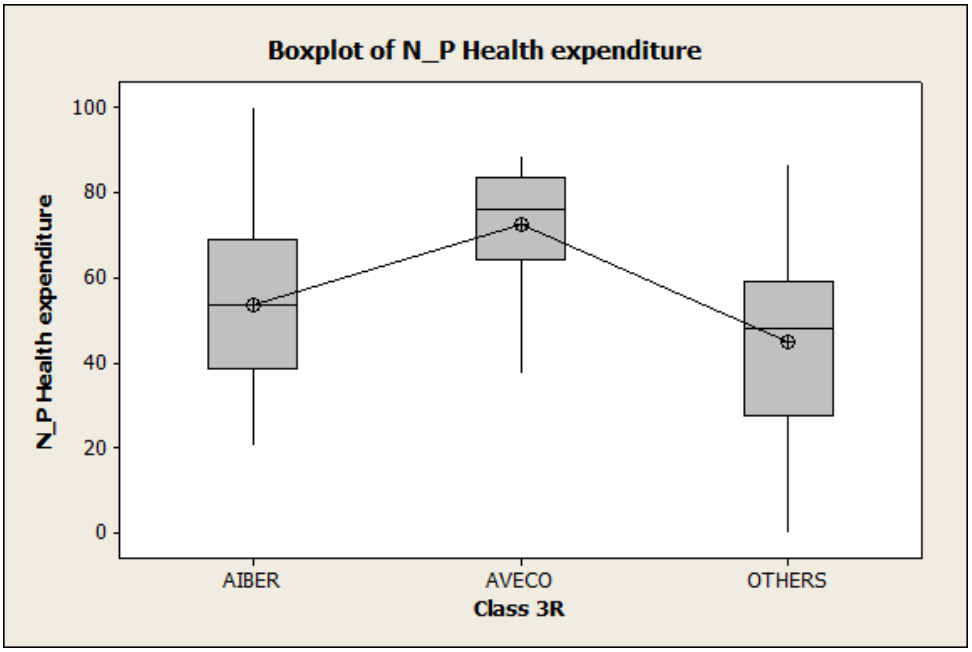
Maior índice: AVECO mas parece não definido.

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

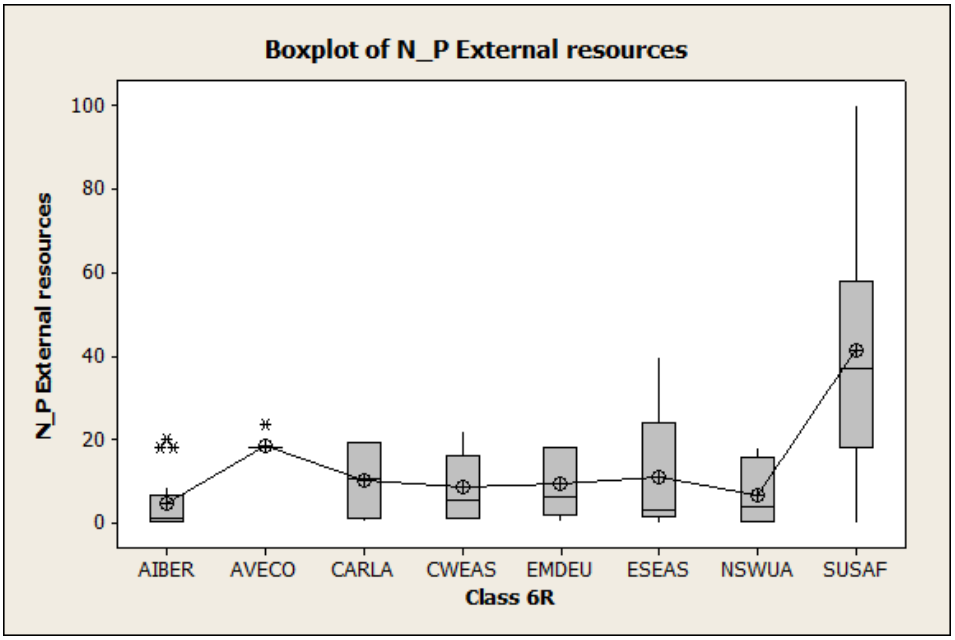
```
Source     DF     SS    MS      F      P
Class 3R    2  15901  7951  20,47  0,000
Error     129  50116   388
Total     131  66017

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

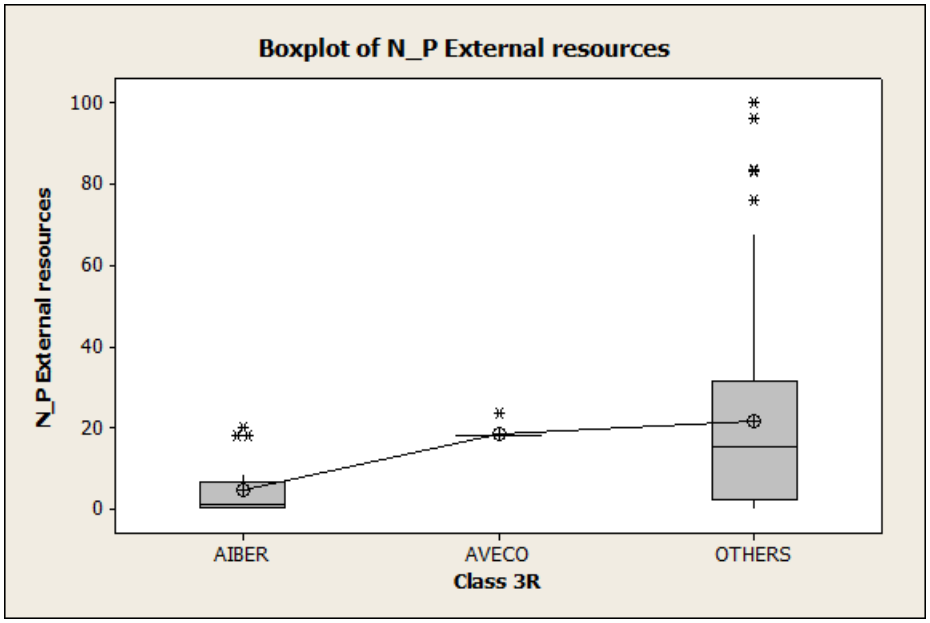
```
Source     DF     SS    MS      F      P
Class 6R    7  25703  3672  15,57  0,000
Error     124  29244   236
Total     131  54947

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).

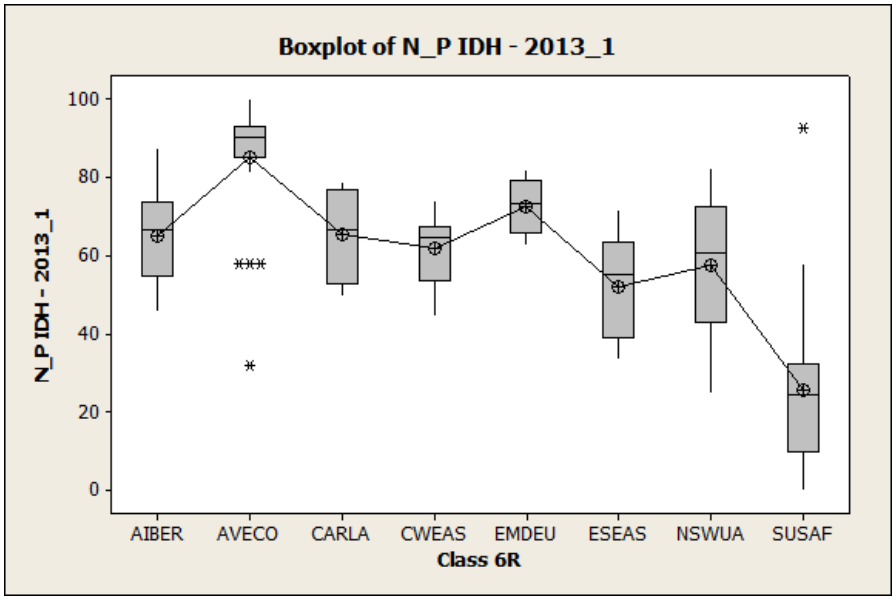
Menor índice: AIBER.

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%
```

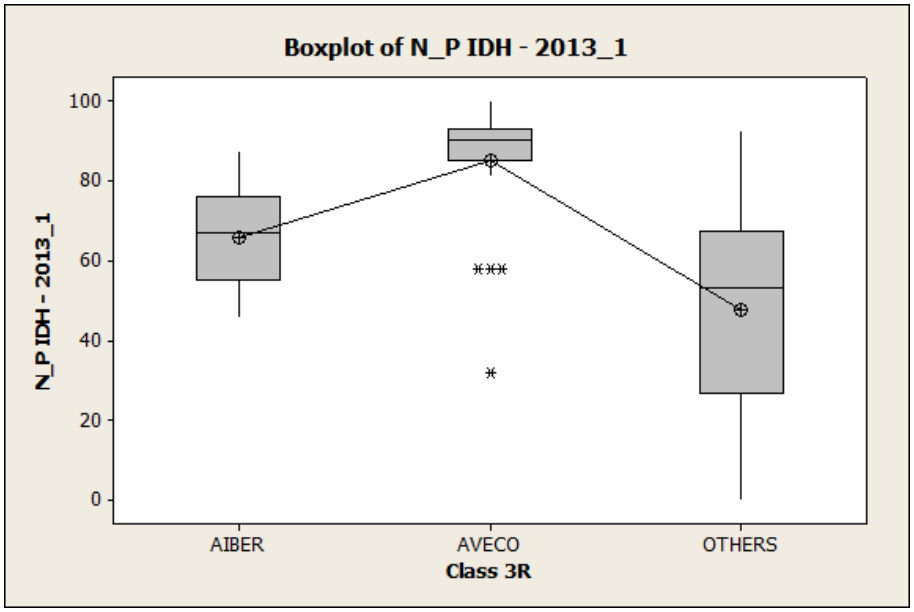
```
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev   +---------+---------+---------+---------
AIBER  20  64,96  11,95                    (--*---)
AVECO  27  85,11  15,53                              (--*--)
CARLA   4  65,32  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.

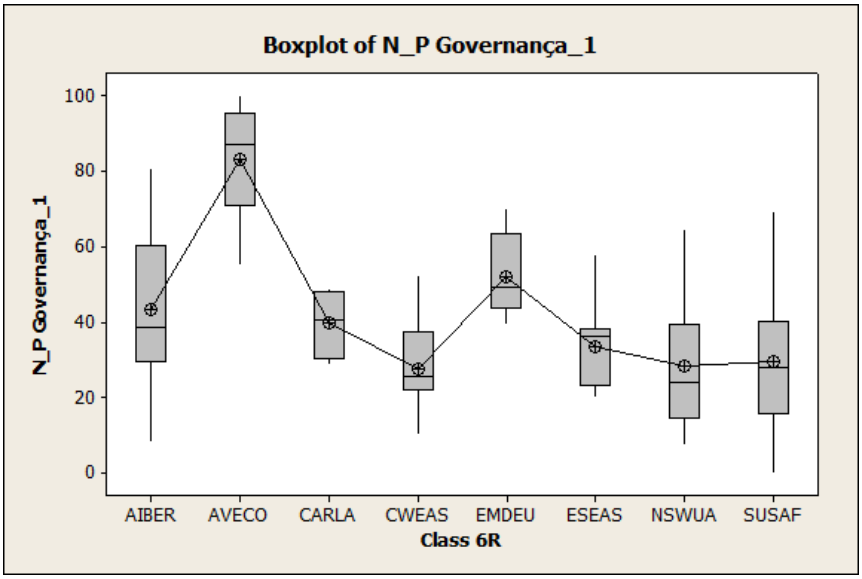
Menor índice: OTHERS.

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).

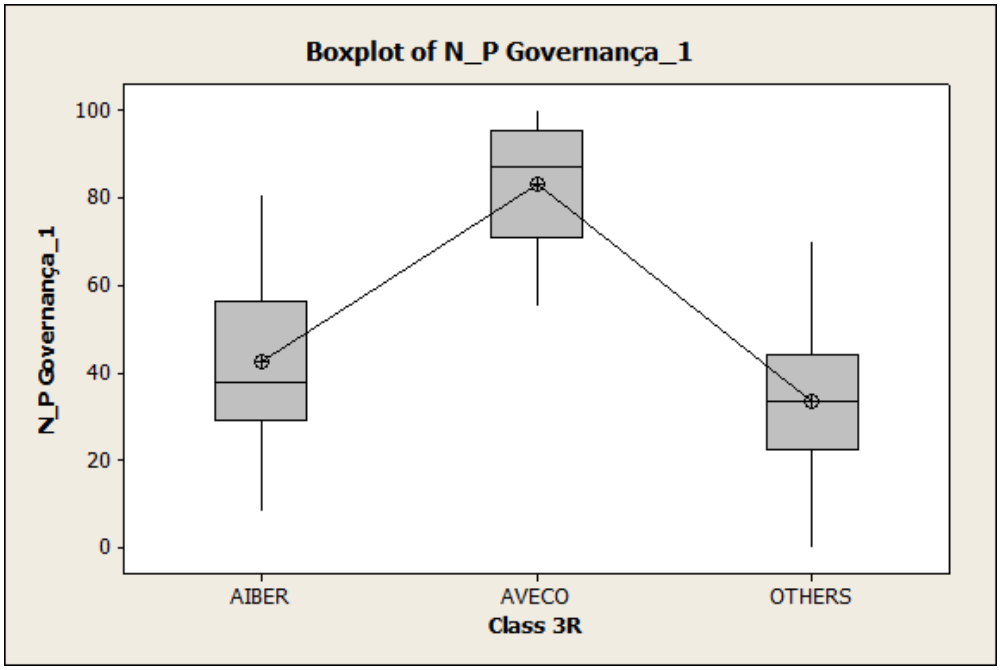
Maior índice: AVECO.

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

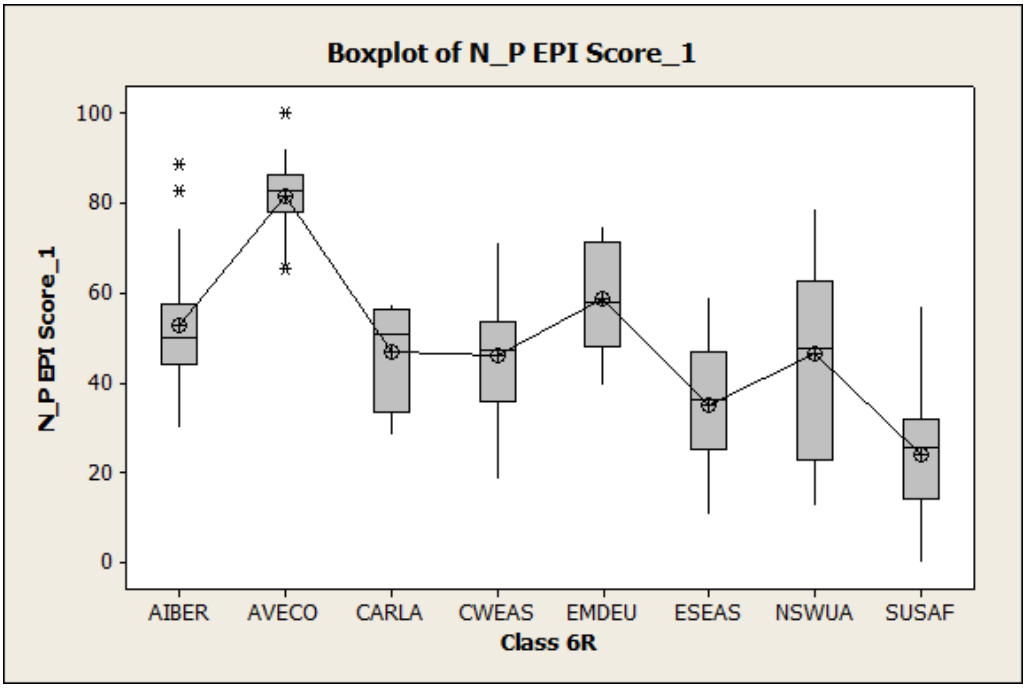
```
Source     DF     SS     MS      F      P
Class 3R    2  51019  25510  97,02  0,000
Error     129  33919    263
Total     131  84938

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.

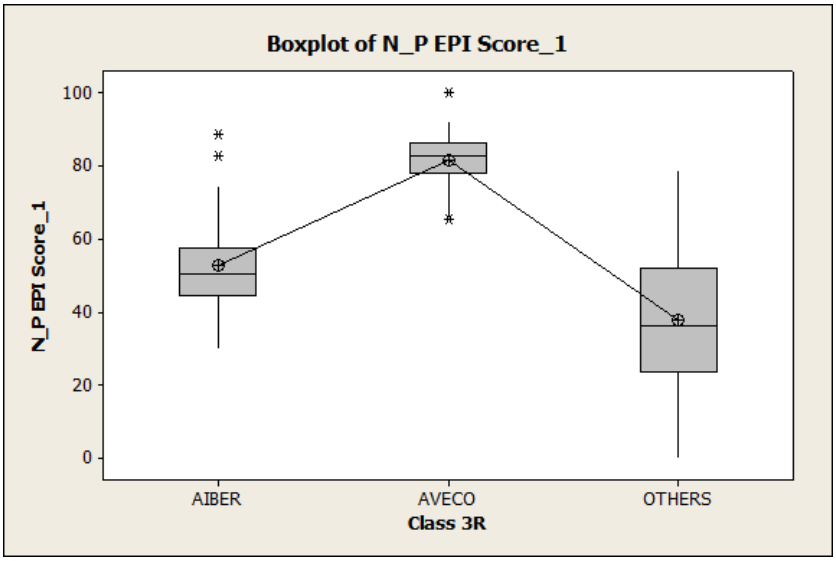
Menor índice: SUSAF.

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.

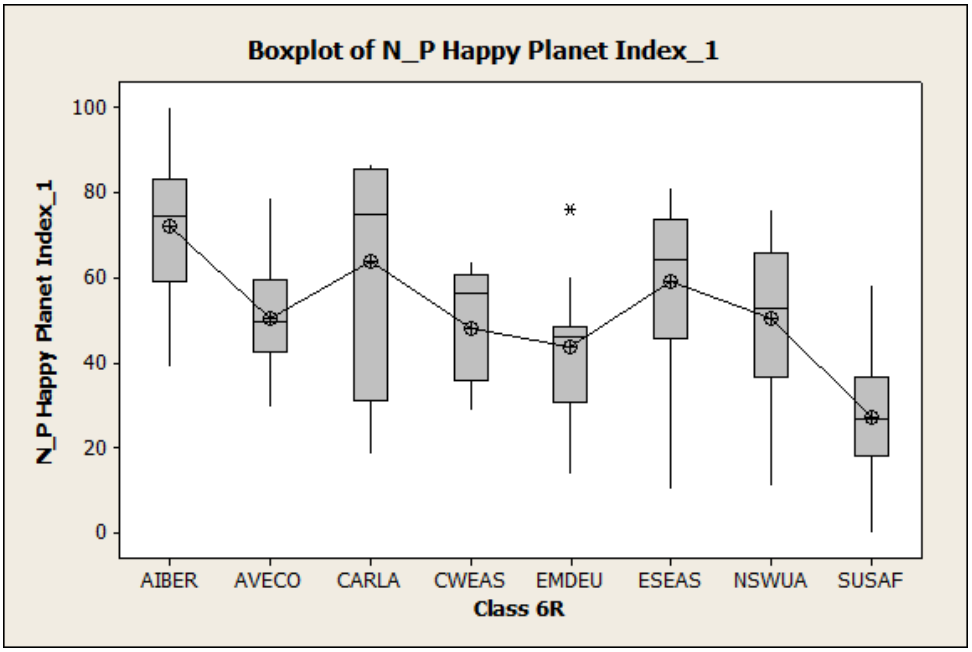
Há outliers em AIBER e AVECO.

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.

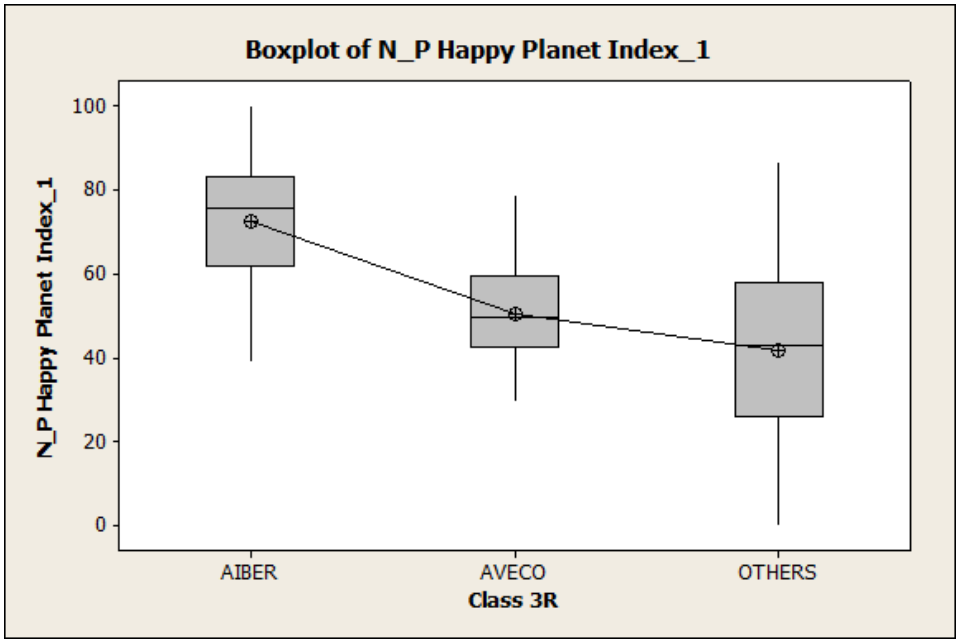
Menor índice: SUSAF.

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.

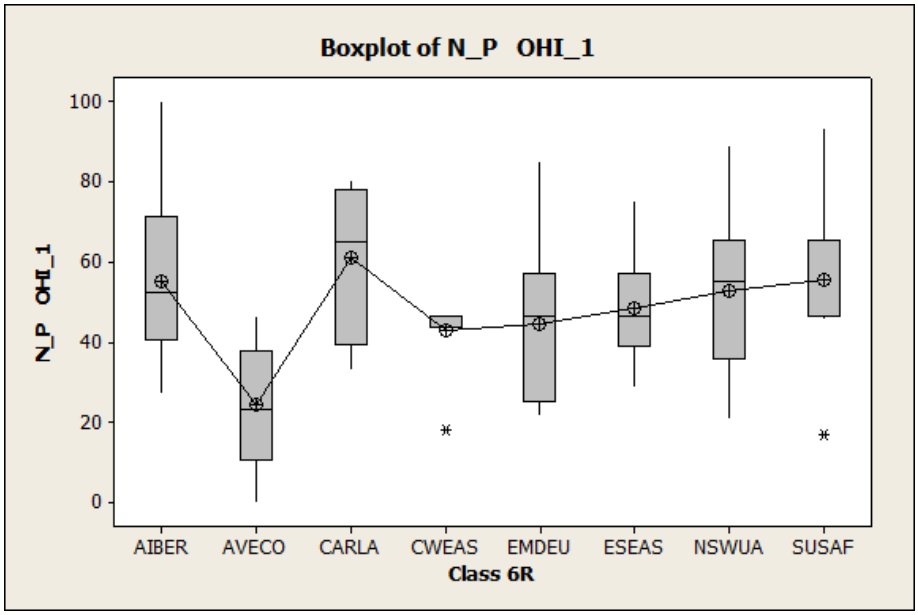
Menor índice: OTHERS.

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.

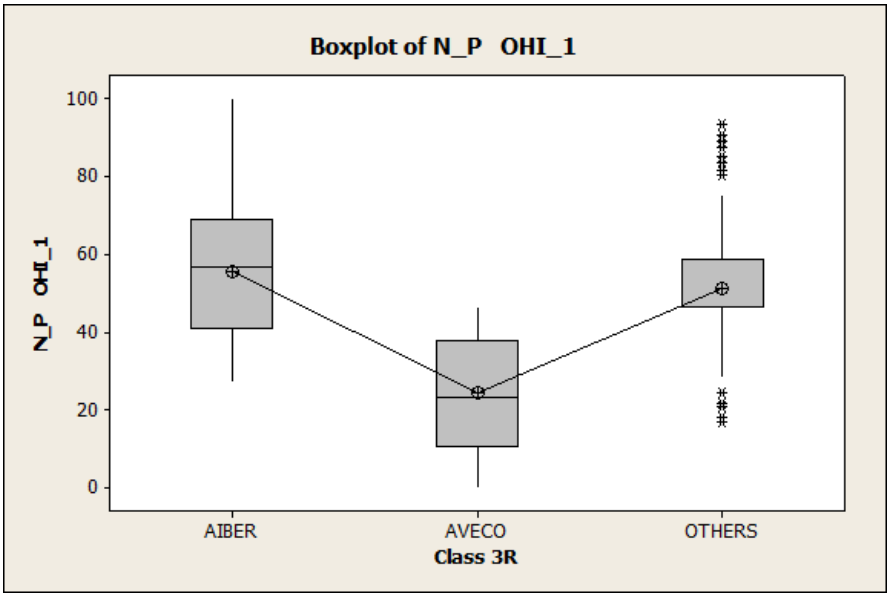
Menor índice: AVECO.

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.

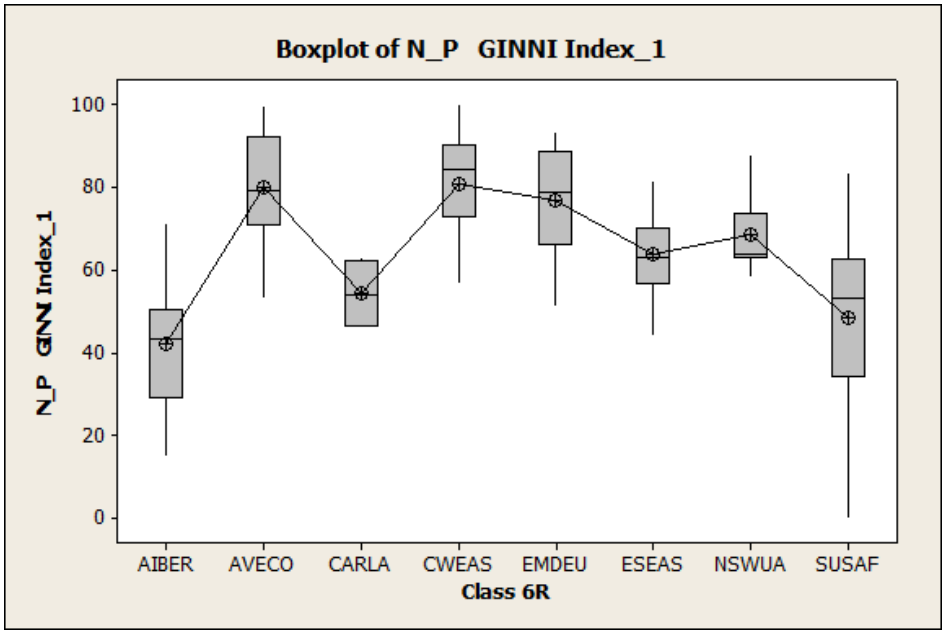
Há vários outliers em OTHERS.

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).

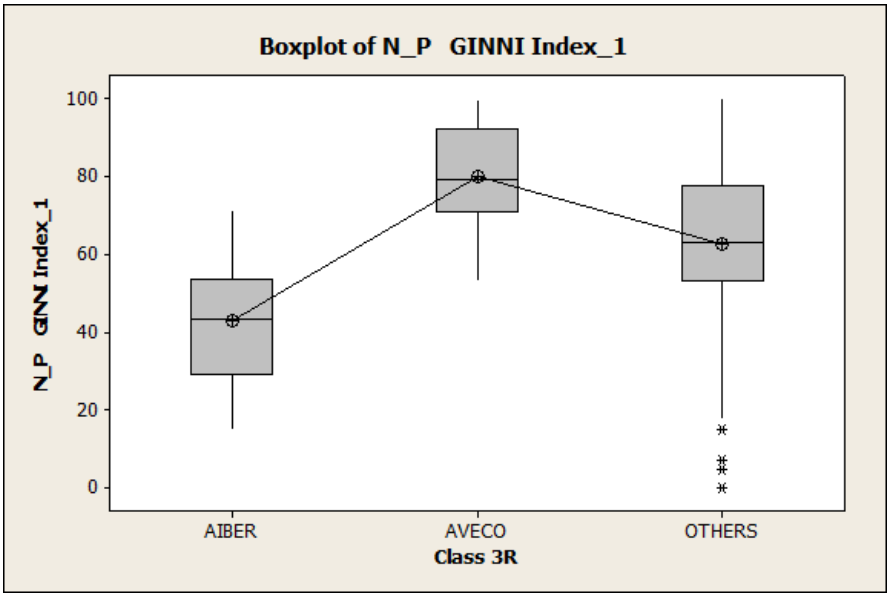
Maior índice: CWEAS (Mas parece não definido).

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.

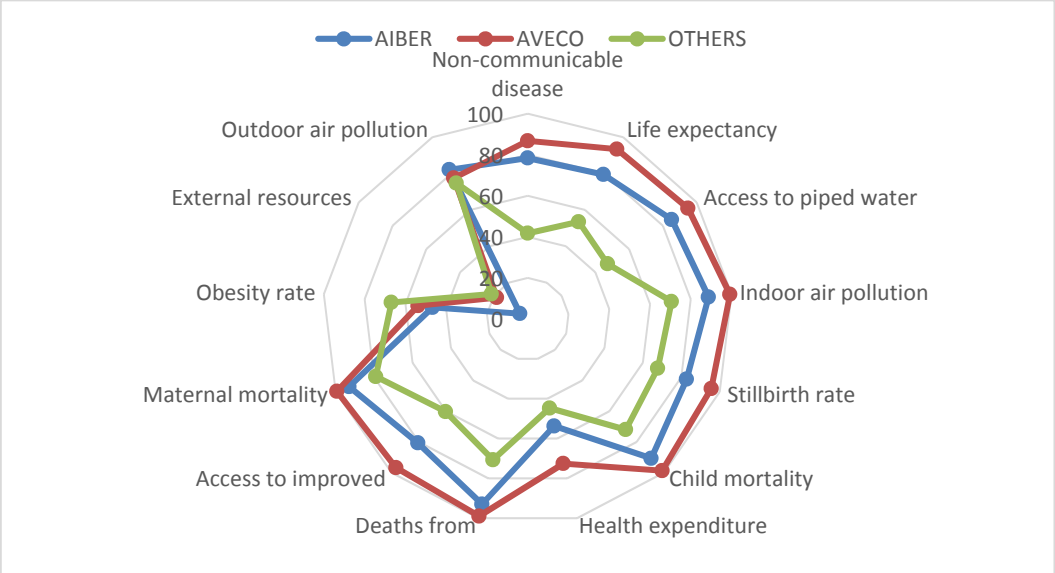
Variável	3R
N_P GovernanÁa_1	97,02
N_P Non-communicable disease	95,63
N_P_SPI_Regressao	83,97
N_P EPI Score_1	69,68
N_P_Life expectancy (years)_1	46,67
N_P_Access to piped water	37,63
N_P IDH - 2013_1	34,45
N_P OHI_1	29,45
N_P_Indoor air pollution	29,42
N_P_Stillbirth rate	25,25
N_P GINNI Index_1	24,45
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:

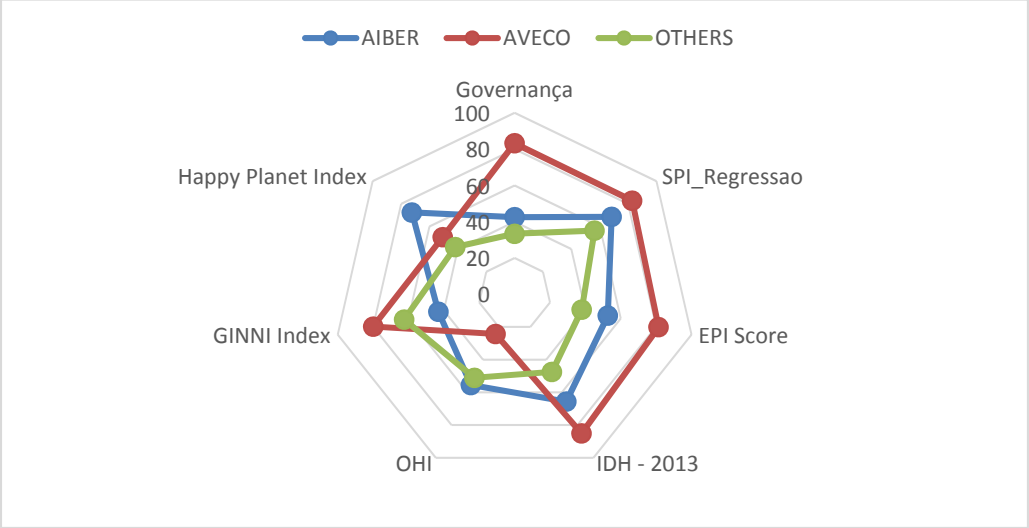
Exceto para a variável N_P Outdoor air pollution da tabela com 3 regiões, todos os outros valores de F foram significativos. Ou seja, para cada variável, as médias das regiões foram consideradas estatisticamente diferentes.

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.

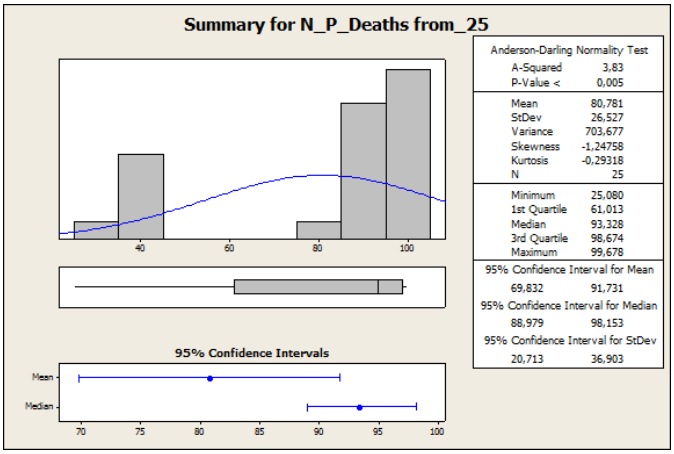
3.10.1 *Deaths for infectious desease* para Amostra de 25 países. Ver Quadro 4

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

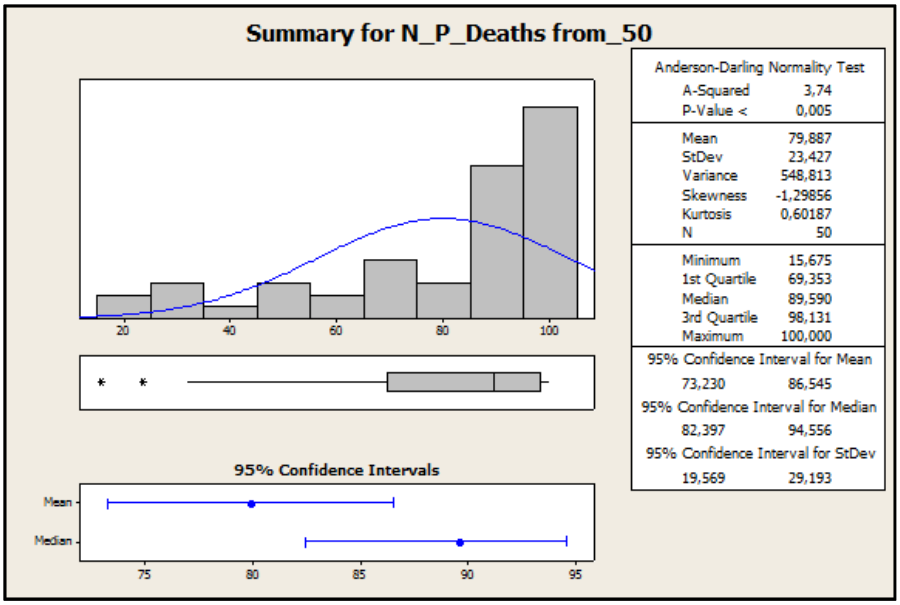
Quadro 5.

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.

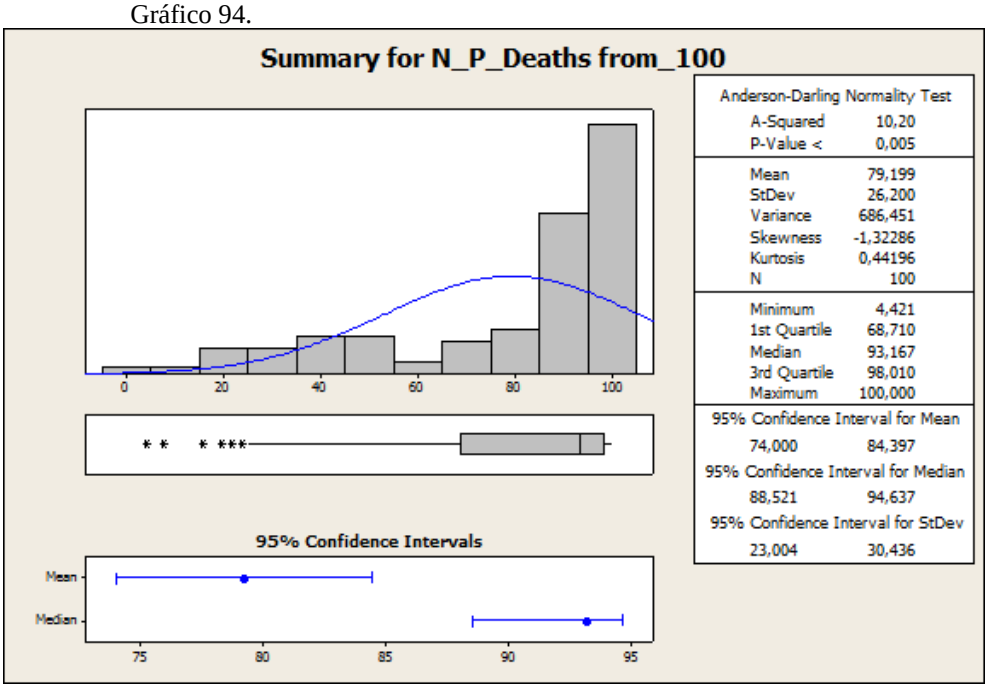
N_P_Deaths from_100	
Malaysia	86,013
Iran	94,293
United States	98,151
Sweden	99,277
Indonesia	81,270
Honduras	88,666
Central African Republic	15,675
Laos	70,659

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:

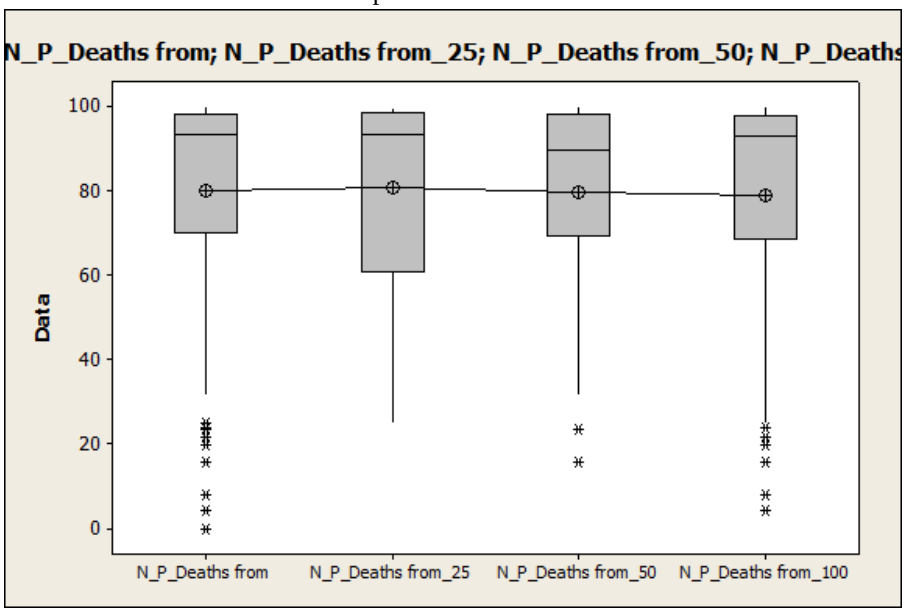
Numa amostra de 100, observa-se que a taxa de sobrevivência por doenças infecciosas diminuiu, e a taxa de sobrevida em alguns países permaneceu mais ou menos a mesma, o que pode ser observado pelo último quartil, na parte mais lateralizada à direita do gráfico 94.

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

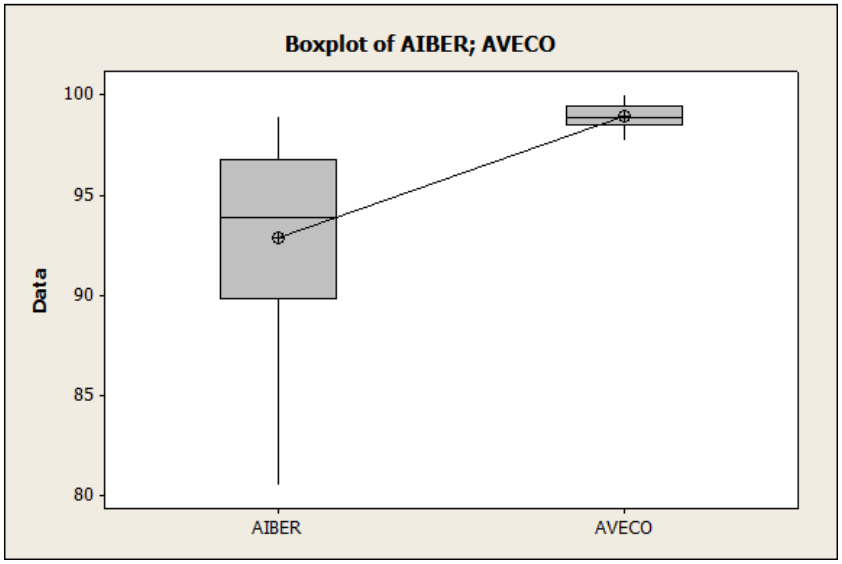
Foram retirados, portanto, os outros países e comparados apenas Aiber e Aveco para o mesmo indicador.

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:

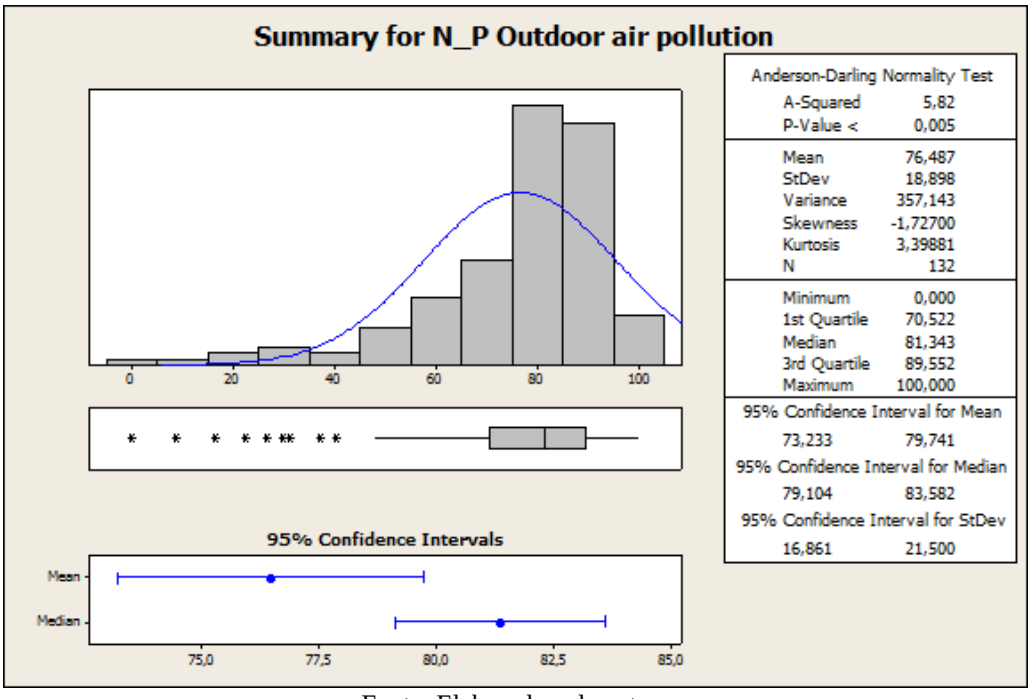
Aqui, o p-value = 0, já no gráfico boxplot gerado pela ANOVA da comparação das amostras, o p-value era de 0,992, o que significa que há uma maior confiança, menor p-value.

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.

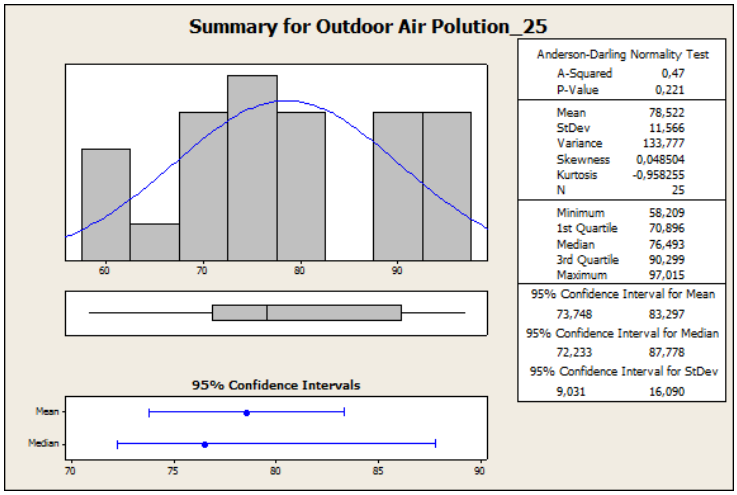
Depois, se seguirá o mesmo procedimento acima, comparando as amostras e depois analisando e comparando pelas regiões dadas.

Gráfico 97.



Fonte: Elaborado pela autora

Gráfico 98.

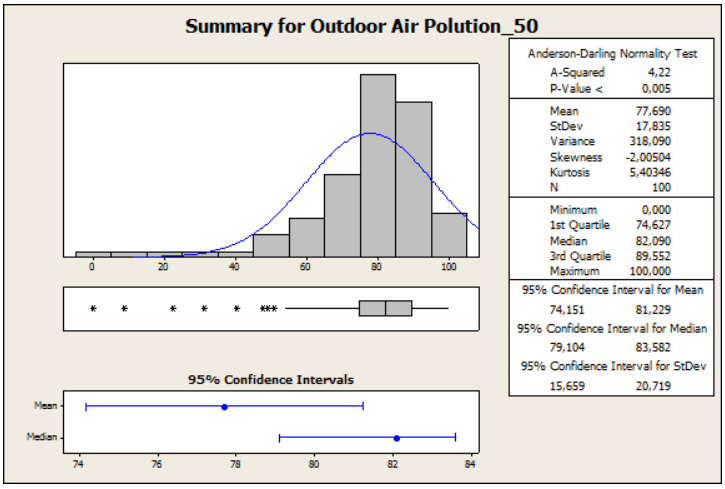


Fonte: Elaborado pela autora

Análise:

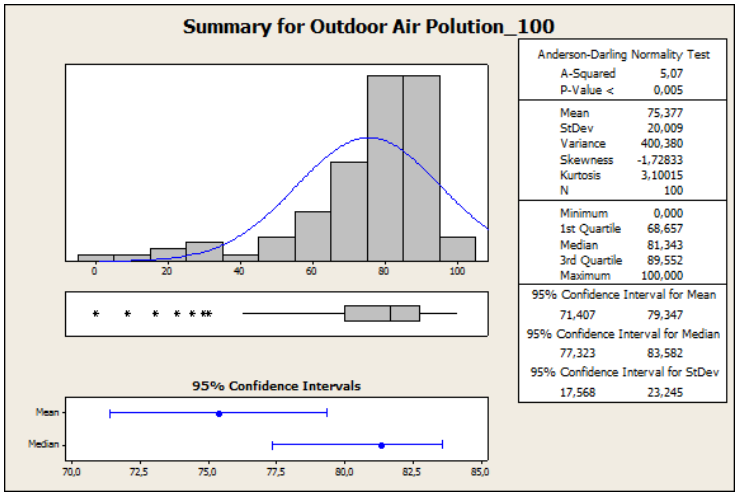
Interessante que para os 25 países, a curva normal é também evidenciada como nas amostras seguintes para de 50 e 100 países.

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

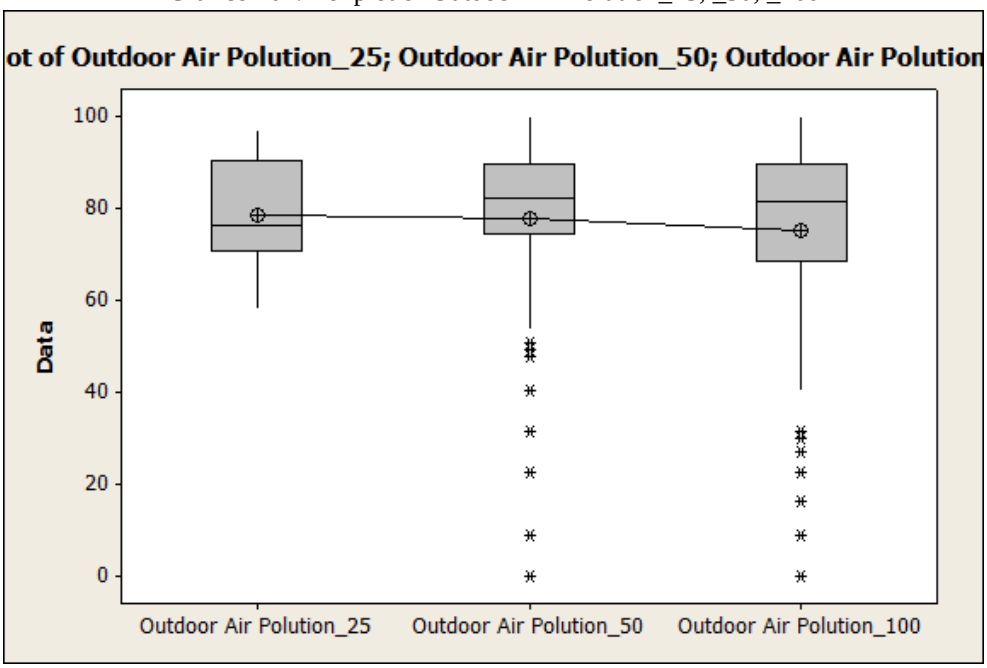
Análise:
A melhor amostra é a de 100.

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			

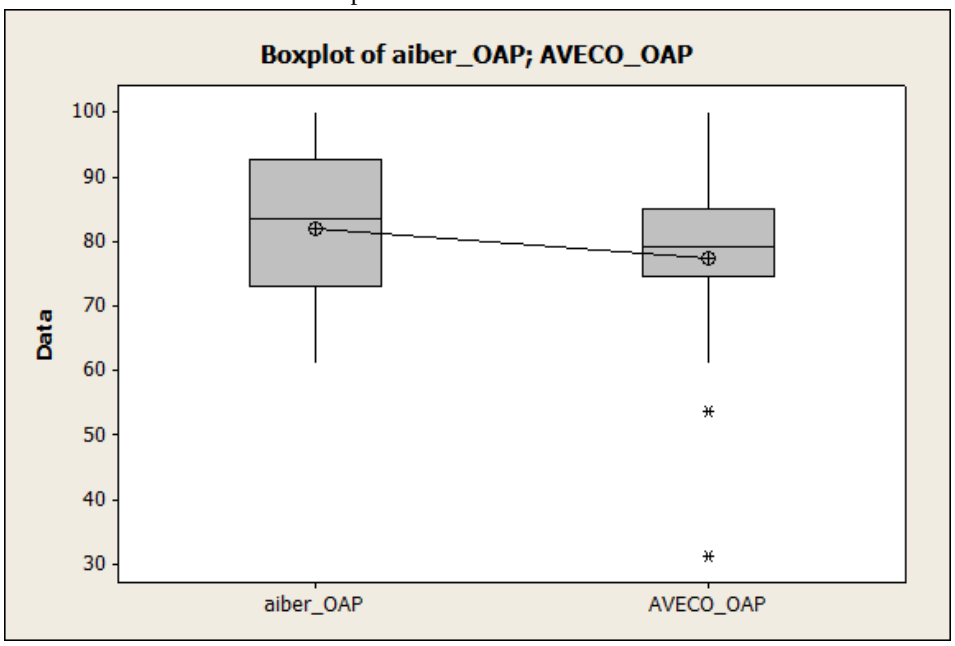
S = 12,55 R-Sq = 3,32% R-Sq(adj) = 1,22%

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

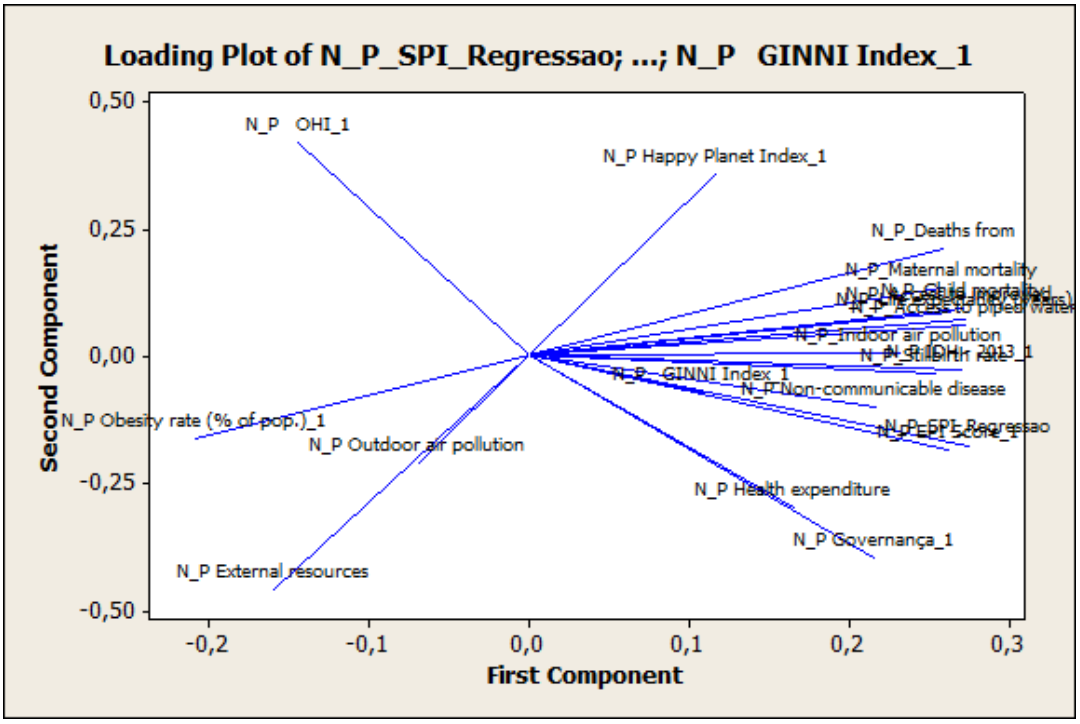
Fonte: Elaborado pela autora

```
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.



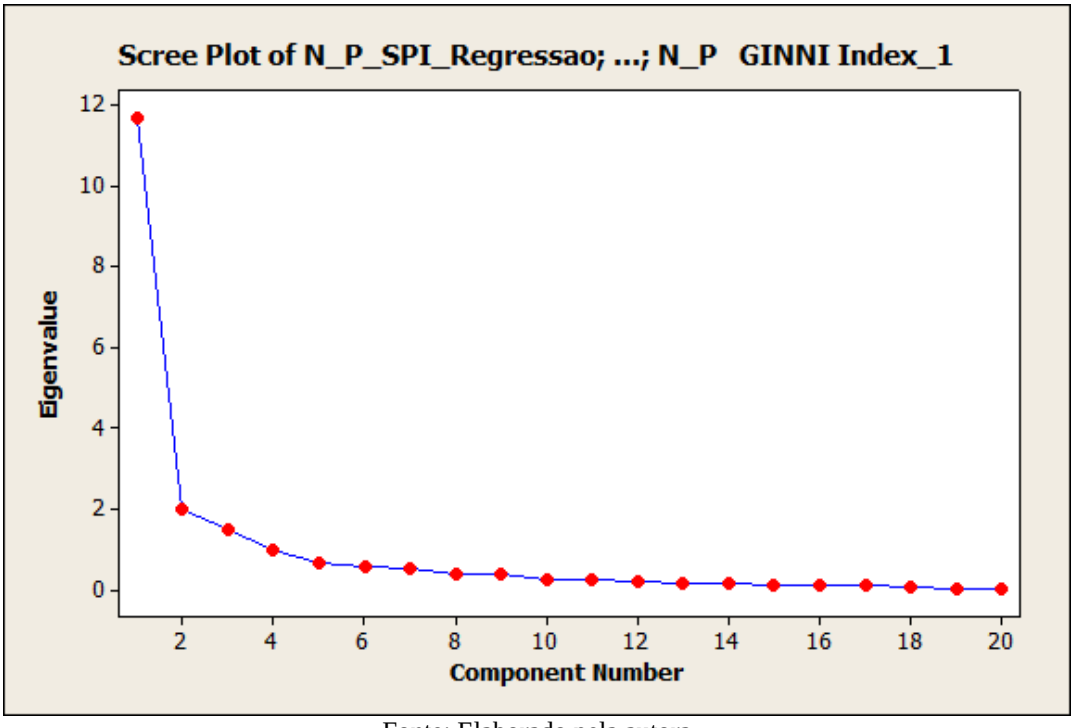
Fonte: Elaborado pela autora

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

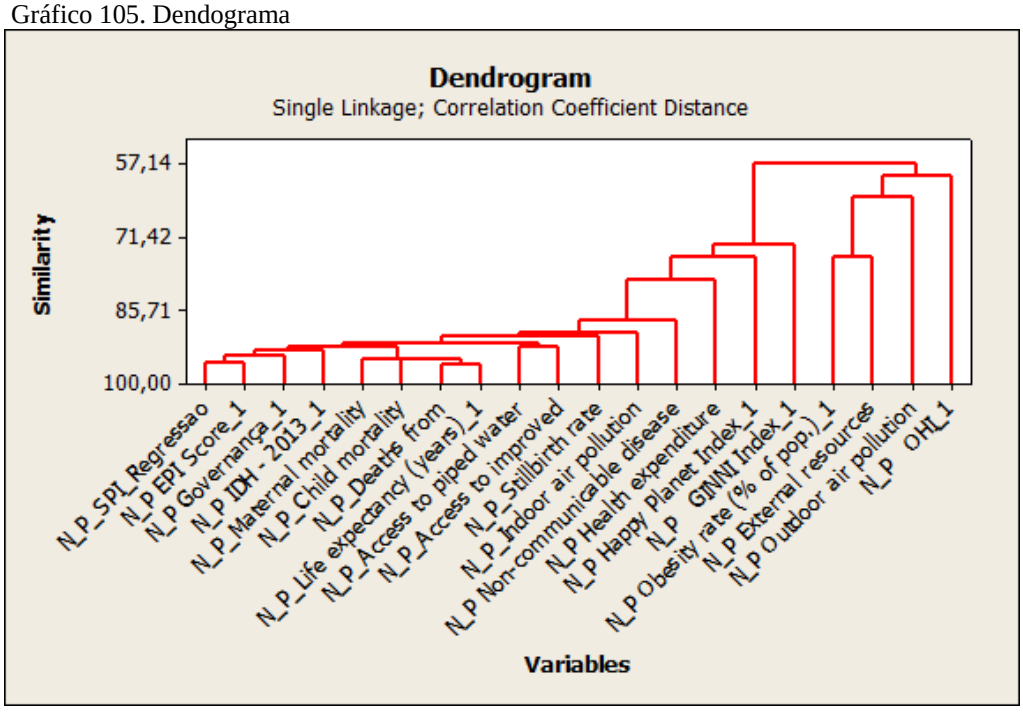
Análise:
PELO Stepwise feito a seguir serão descobertos quais itens correspondem às 4 primeiras bolas vermelhas, da esquerda para a direita, que respondem fundamentalmente pelo Modelo.

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



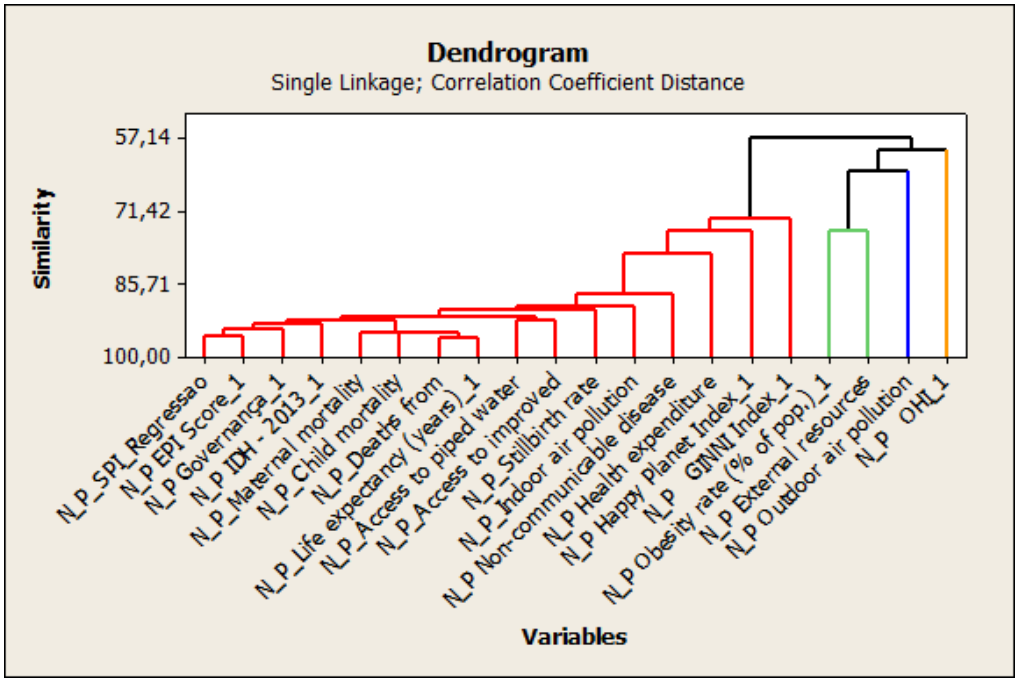
Fonte: Elaborado pela autora

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.

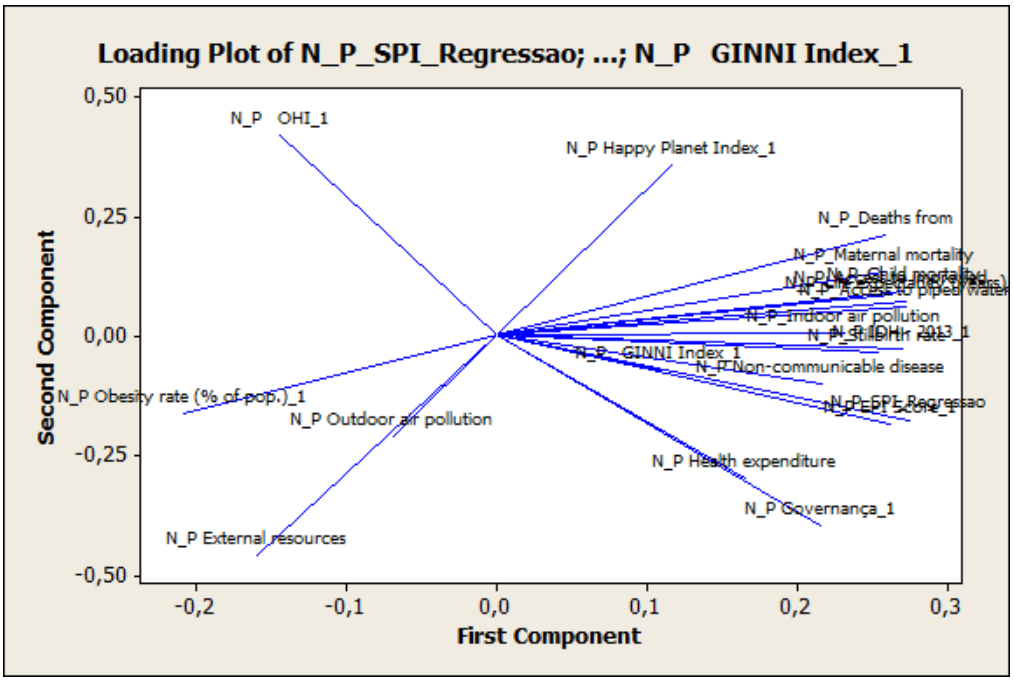


Fonte: Elaborado pela autora

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.

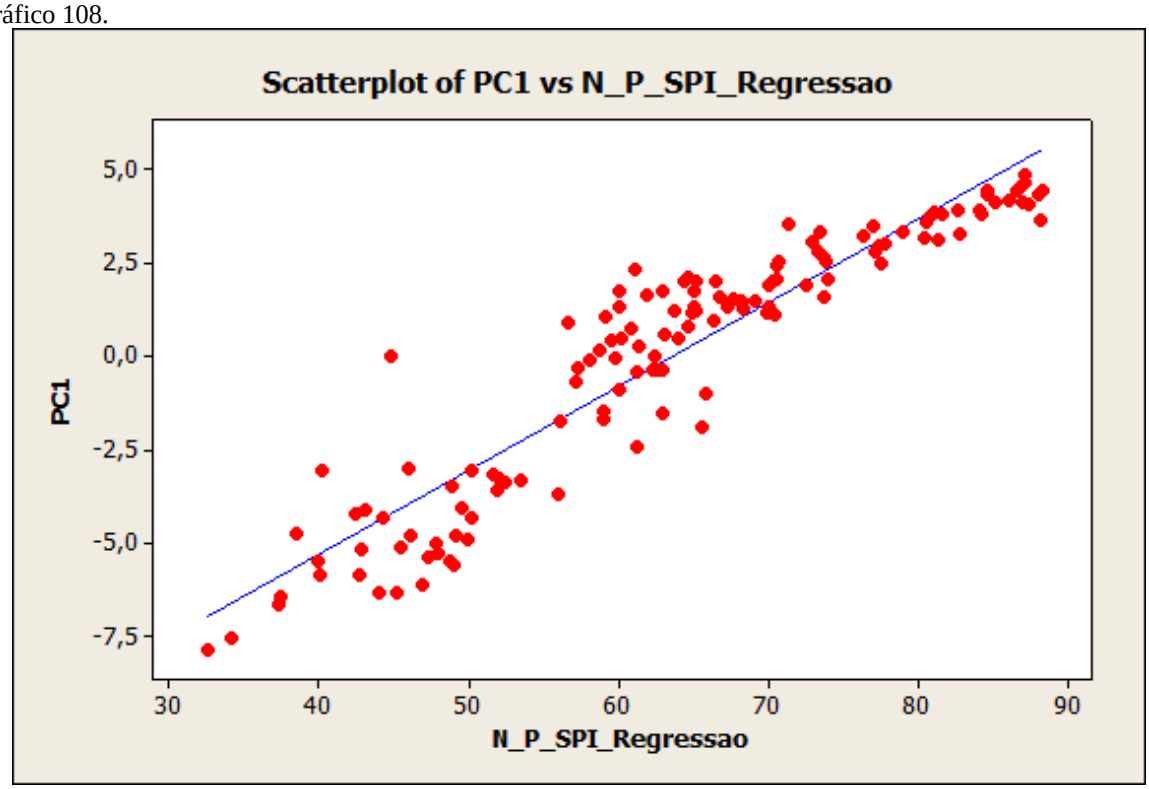
Fits and Diagnostics for Unusual Observations

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.



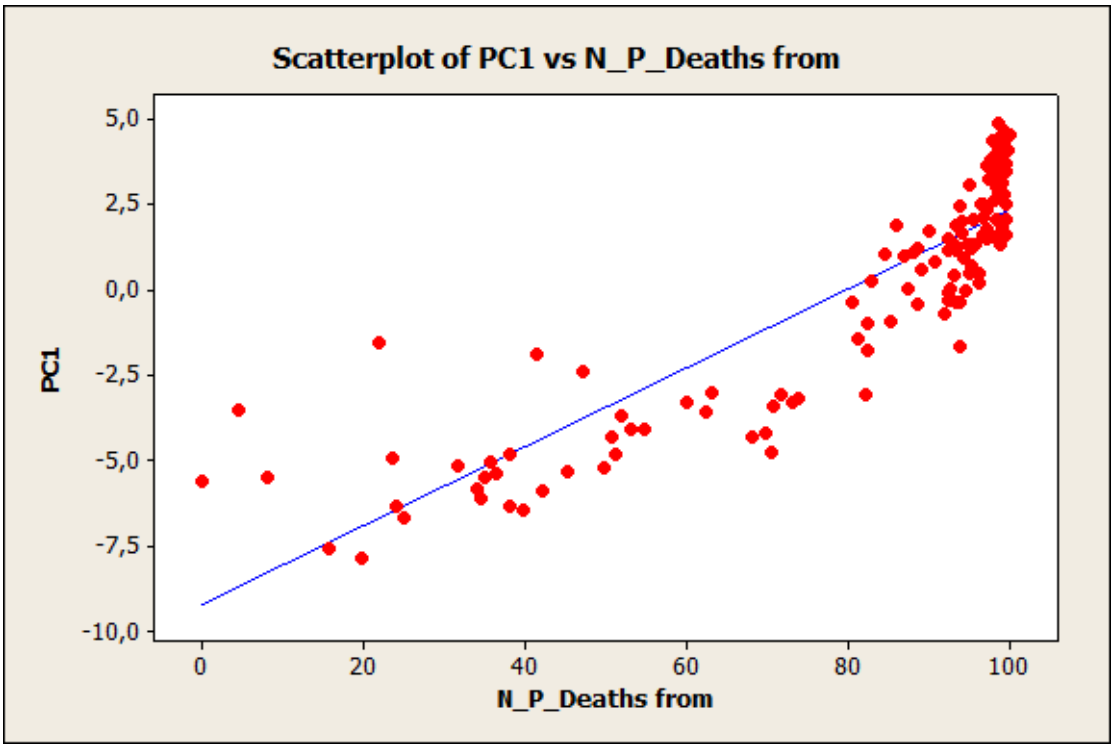
Fonte: Elaborado pela autora

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.

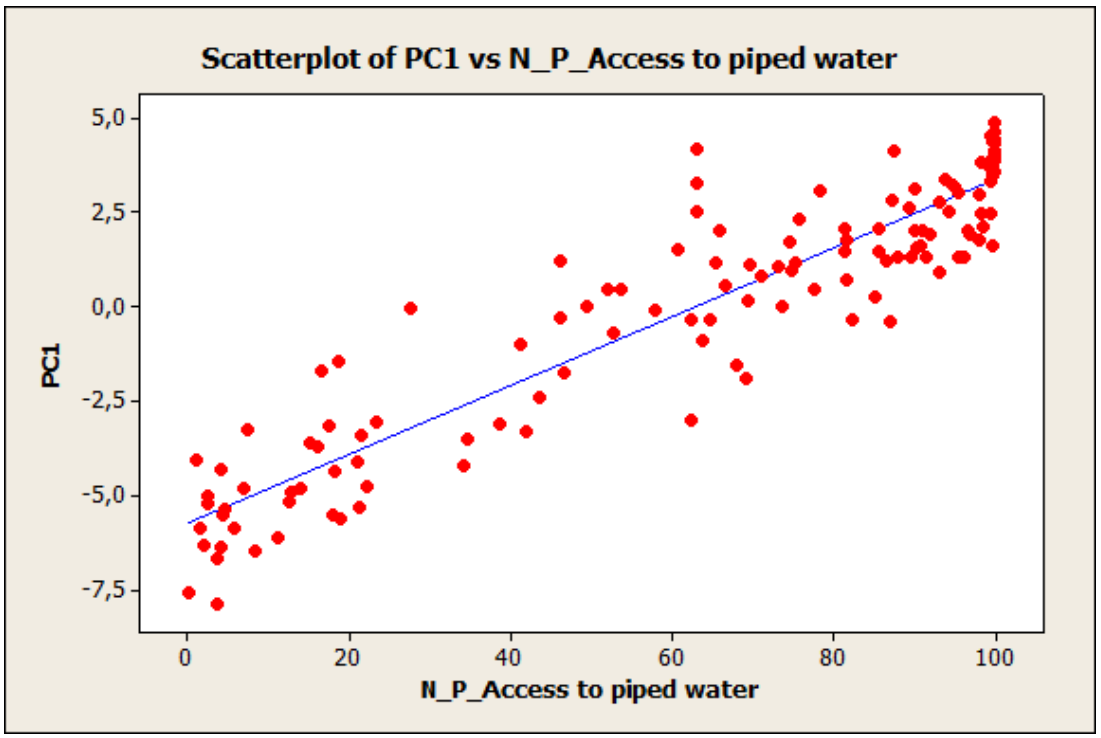


Fonte: Elaborado pela autora

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.

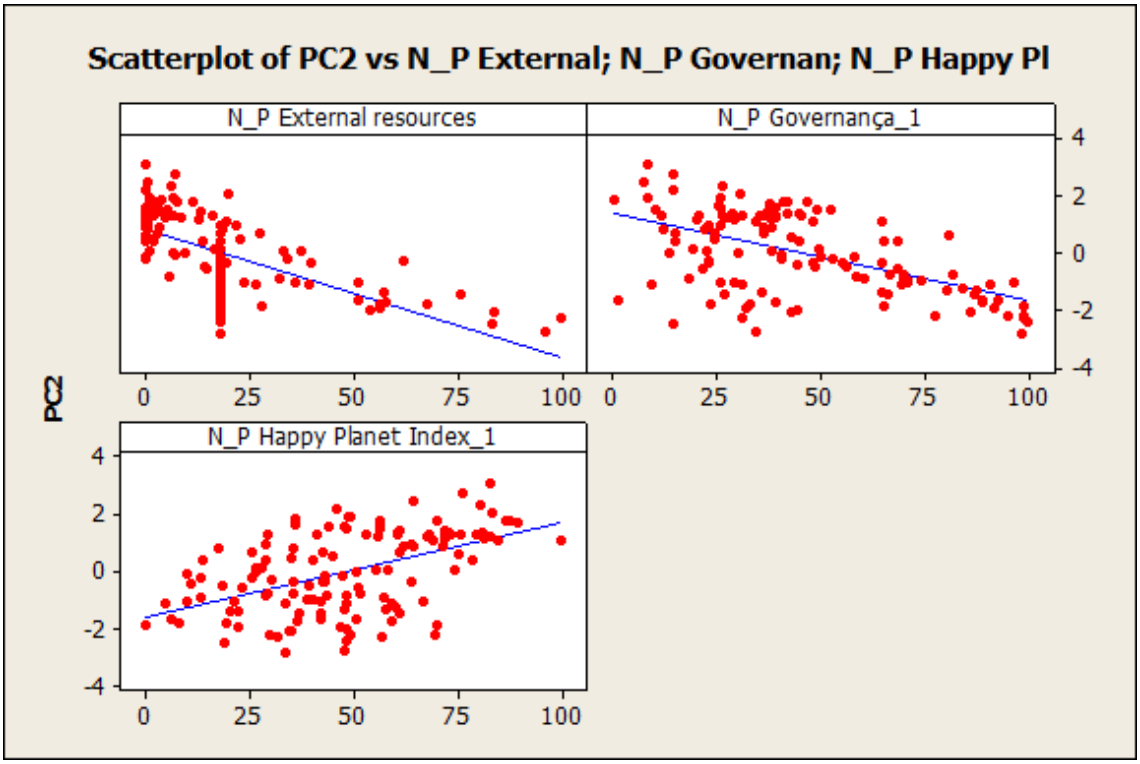
Fits and Diagnostics for Unusual Observations

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

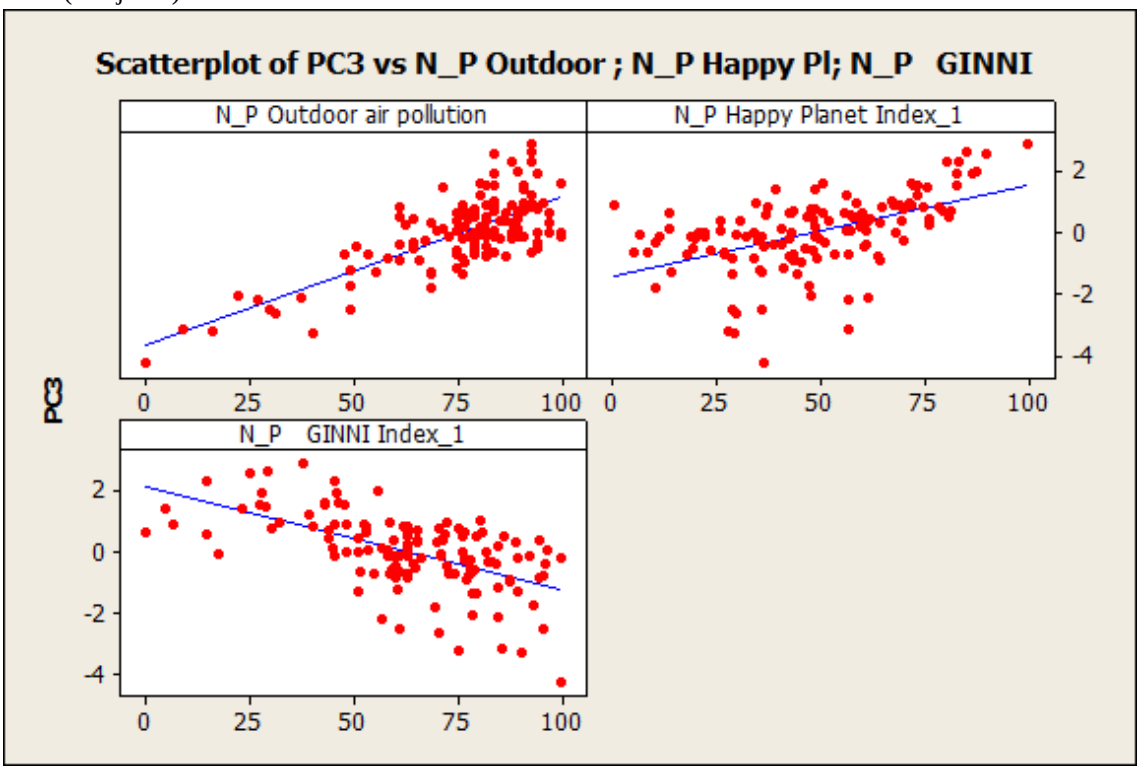
Fits and Diagnostics for Unusual Observations

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.

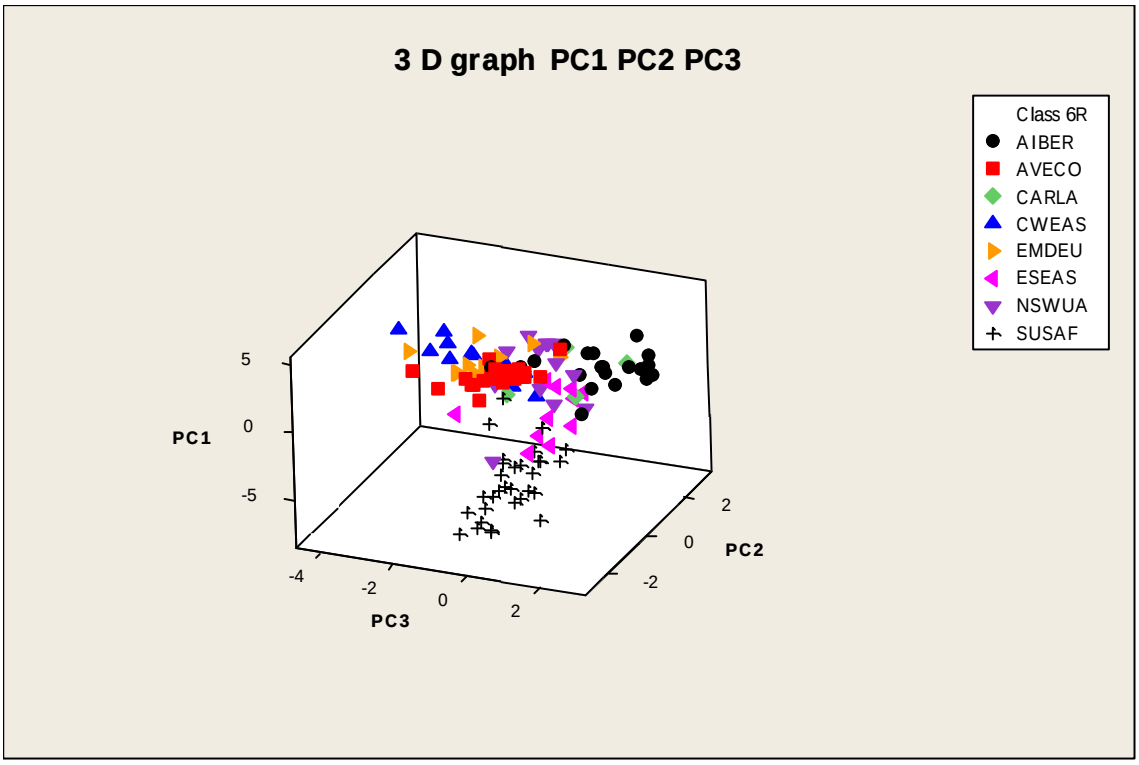
3.13 CLASSIFICAÇÃO NÃO SUPERVISIONADA, ANÁLISE DISCRIMINANTE

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.

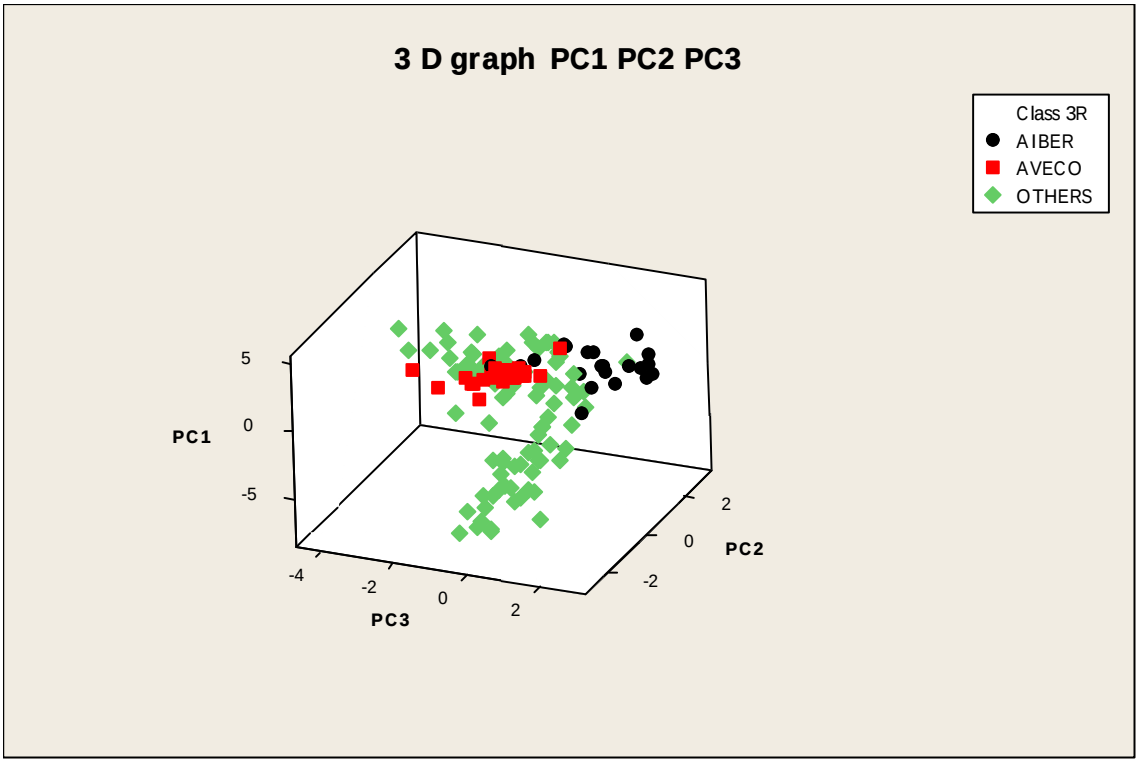
Os gráficos 113 (6 regiões) e 114 (3 regiões) mostram como estão distribuídos os países por região, de acordo com as componentes.

Gráfico 113.



Fonte: Elaborado pela autora

Gráfico 114.



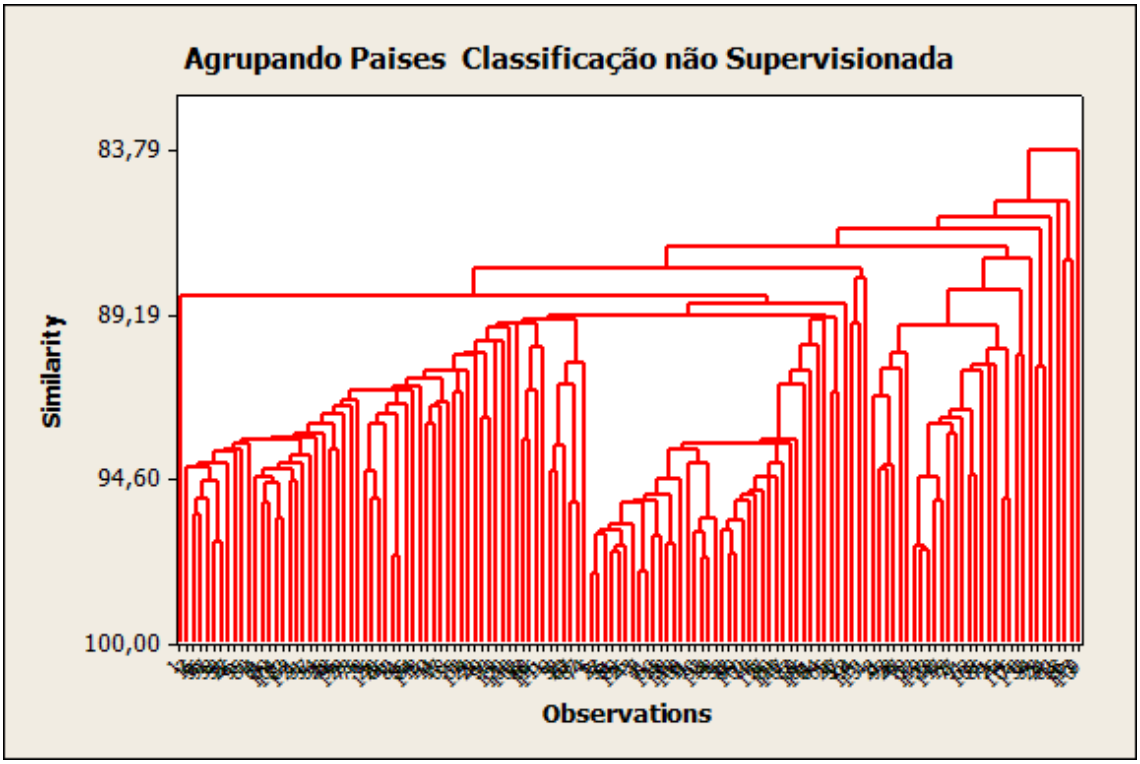
Fonte: Elaborado pela autora

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.*

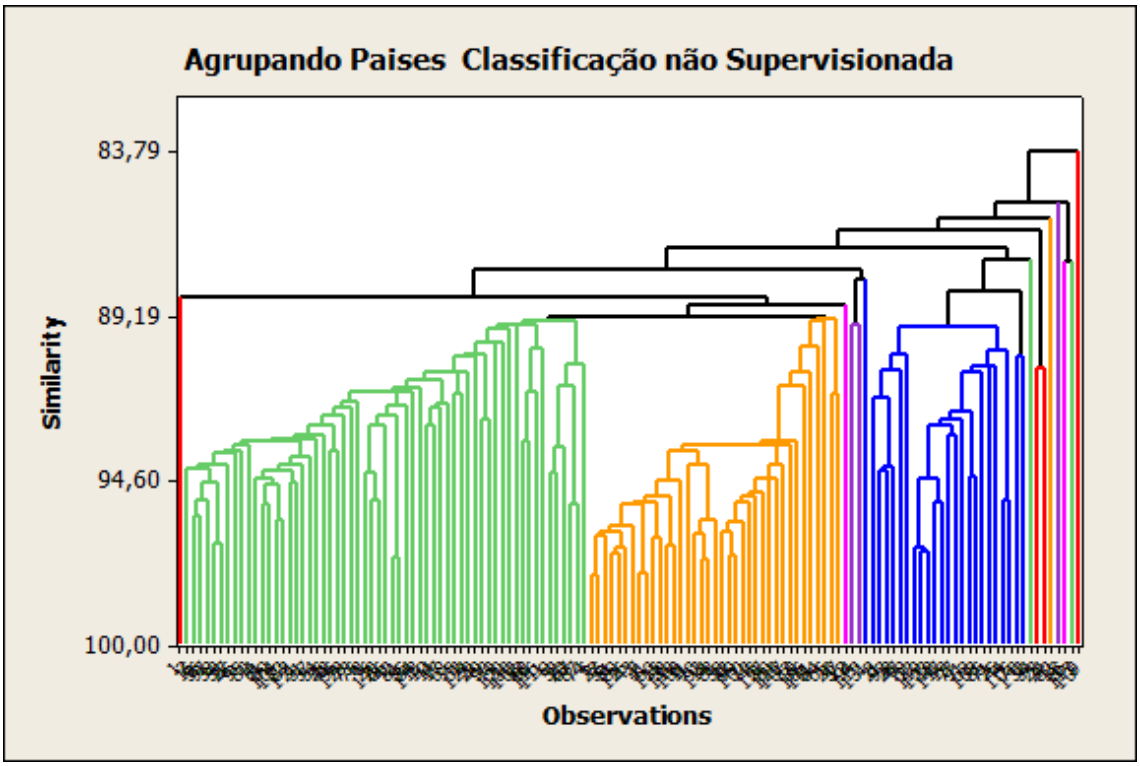
Será feita a observação dos *Clusters* PC1, PC2 e PC3 através de dendogramas. O primeiro deles foi apresentado com um único cluster, conforme pode ser visto no Dendograma monocromático do Gráfico 115. Posteriormente em vários mais (10), Gráfico 116.

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.

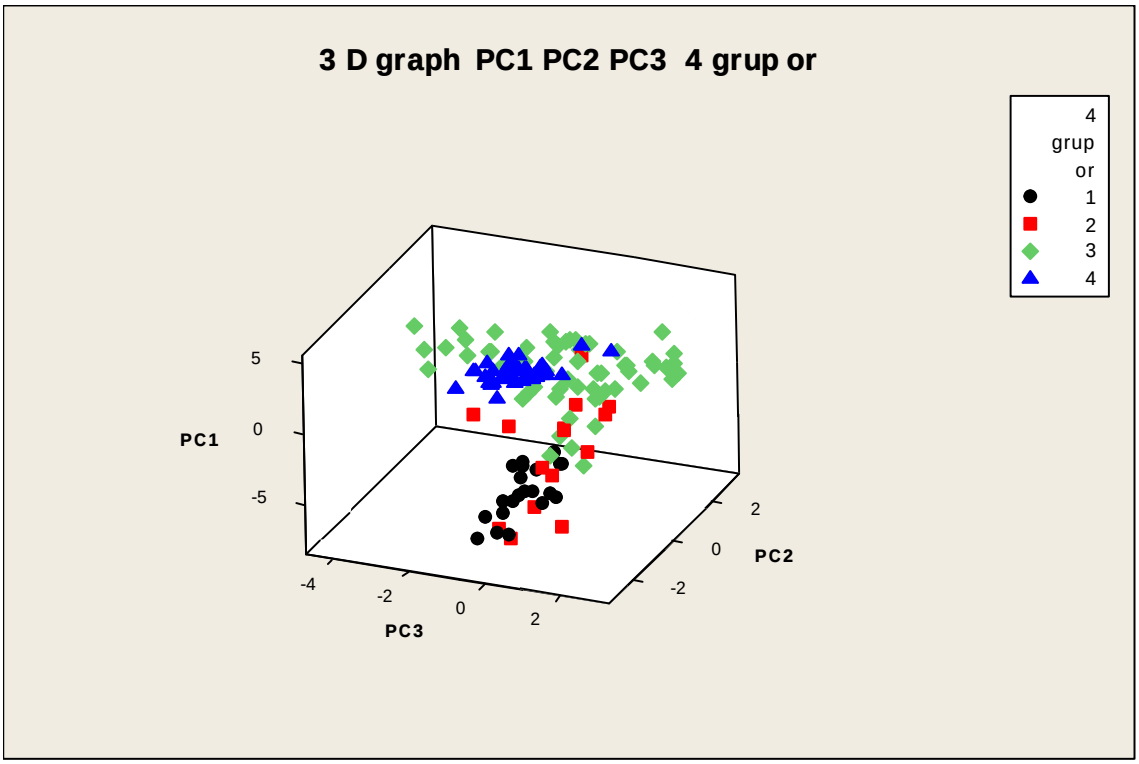
				Individual 95% CIs For Mean Based on Pooled StDev
Level	N	Mean	StDev	----+---------+---------+---------+-----
azul	21	45,767	4,829	(-*--)
laranja	37	80,559	6,002	(-*-)
o	15	50,816	11,947	(--*--)
verde	59	62,723	6,998	(*-)
				----+---------+---------+---------+-----
				48 60 72 84

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

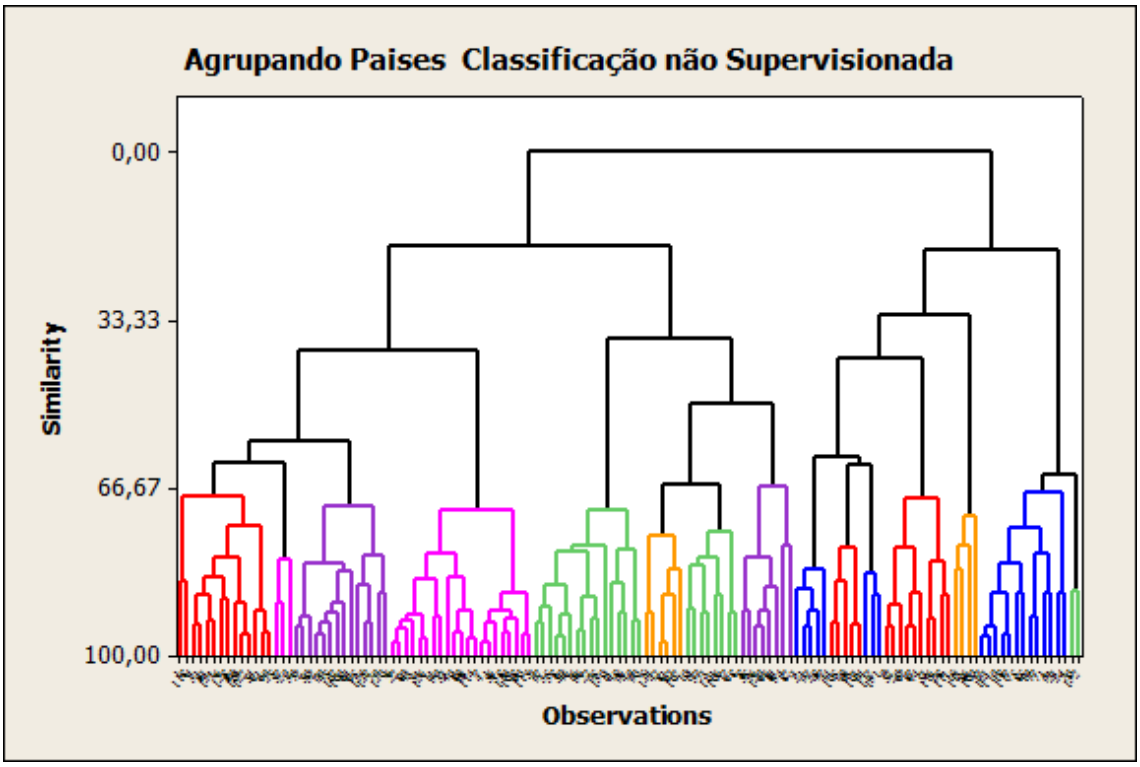
Rows: Class 3R	Columns: 4 grup or				
	1	2	3	4	All
AIBER	0	1	17	3	21
AVECO	0	0	1	26	27
OTHERS	21	14	41	8	84
All	21	15	59	37	132

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

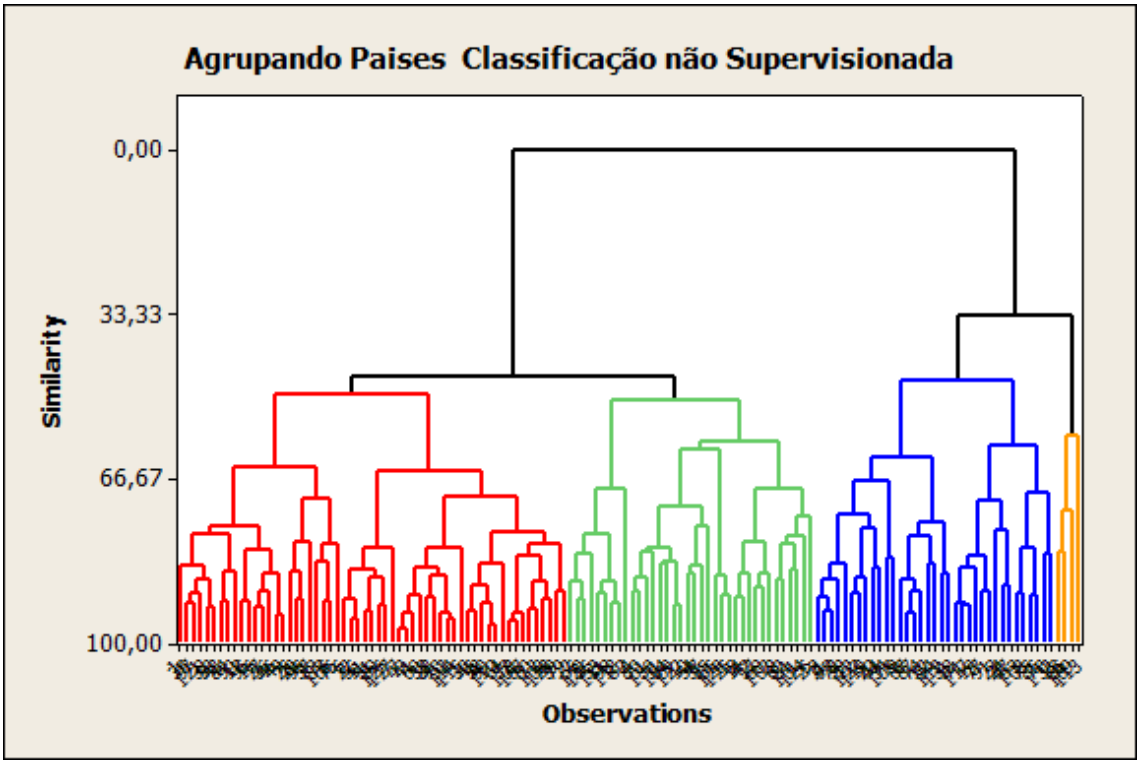
Fonte: Elaborado pela autora

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

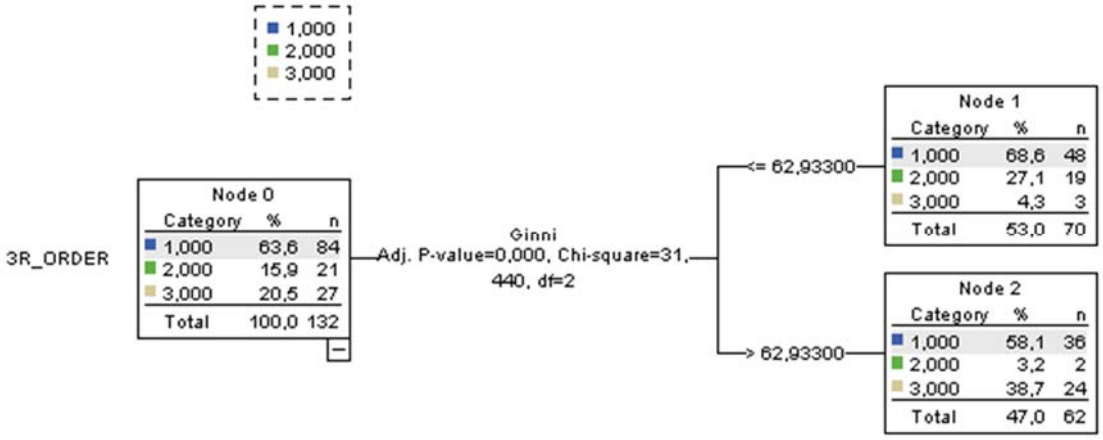
Gain summary Tables are not displayed because profits are undefined.
Target category gains tables are not displayed because target categories are undefined.

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)

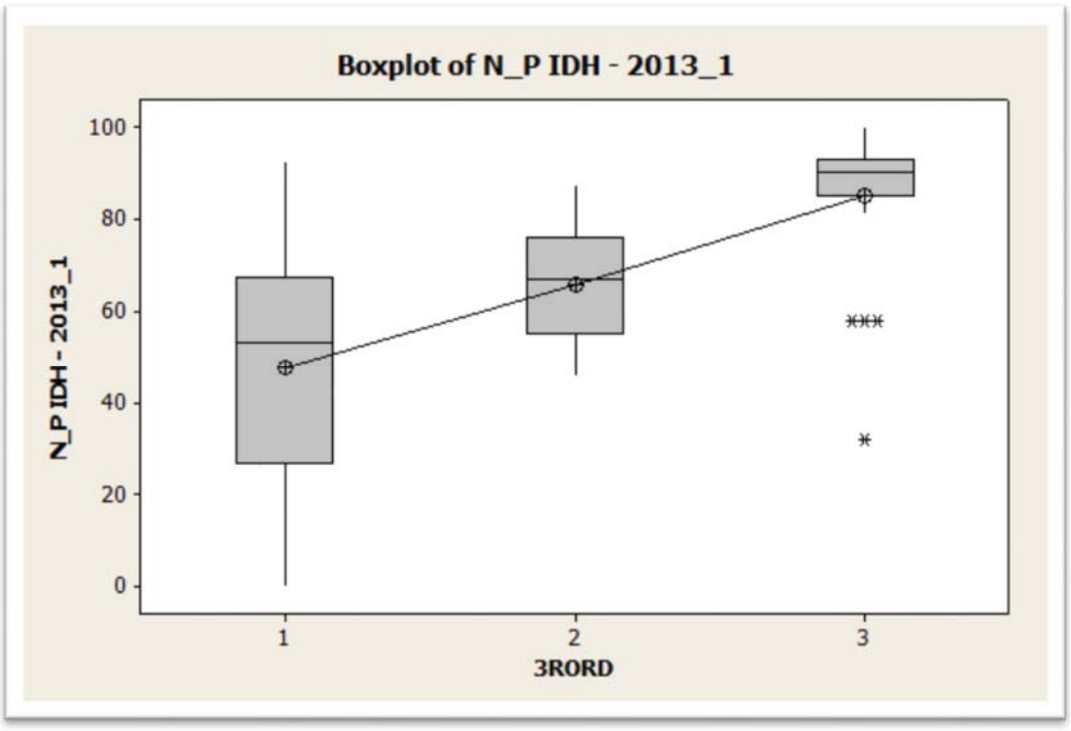
Source	DF	SS	MS	F	P
3RORD	2	30392	15196	34,45	0,000
Error	129	56900	441		
Total	131	87292			

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

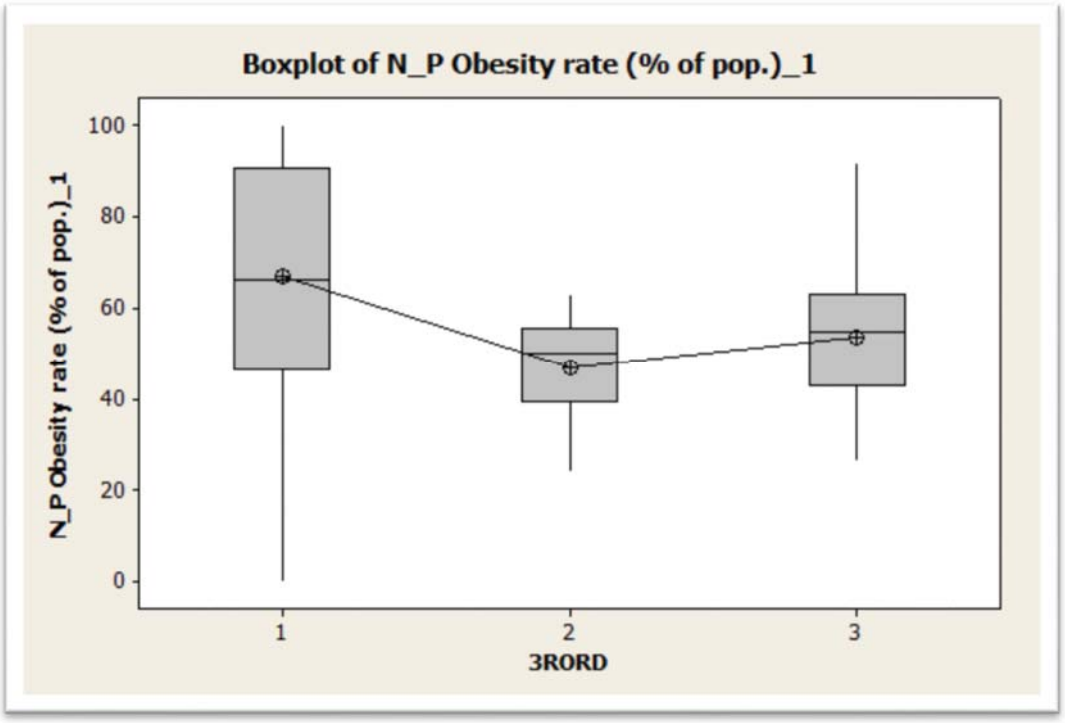
Source	DF	SS	MS	F	P
3RORD	2	8629	4315	9,50	0,000
Error	129	58565	454		
Total	131	67194			

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			

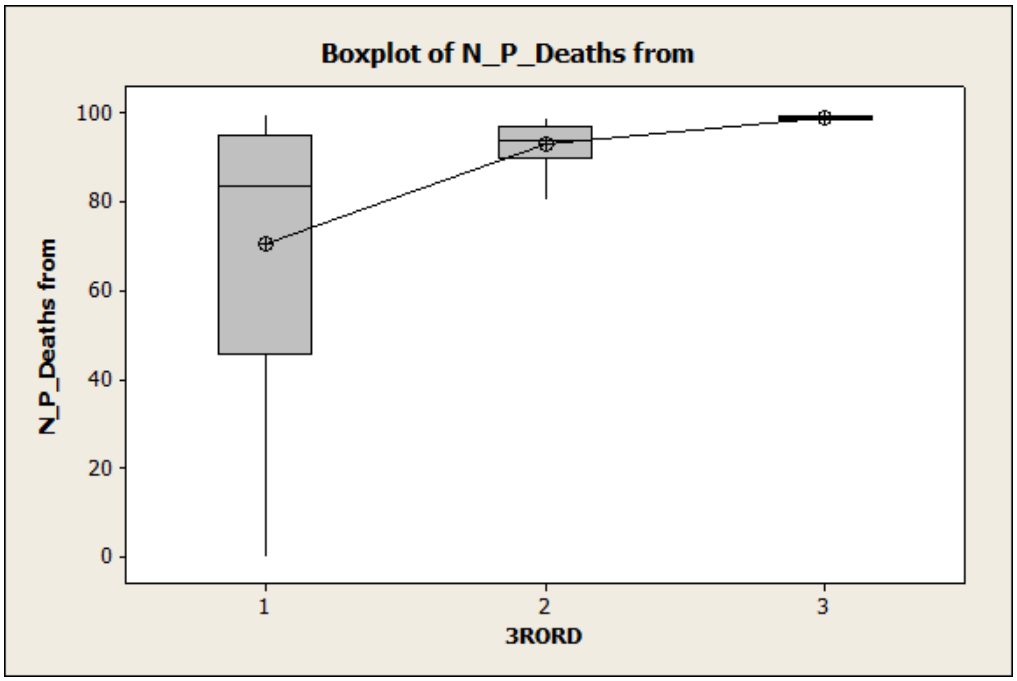
S = 23,21 R-Sq = 22,78% R-Sq(adj) = 21,59%

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*

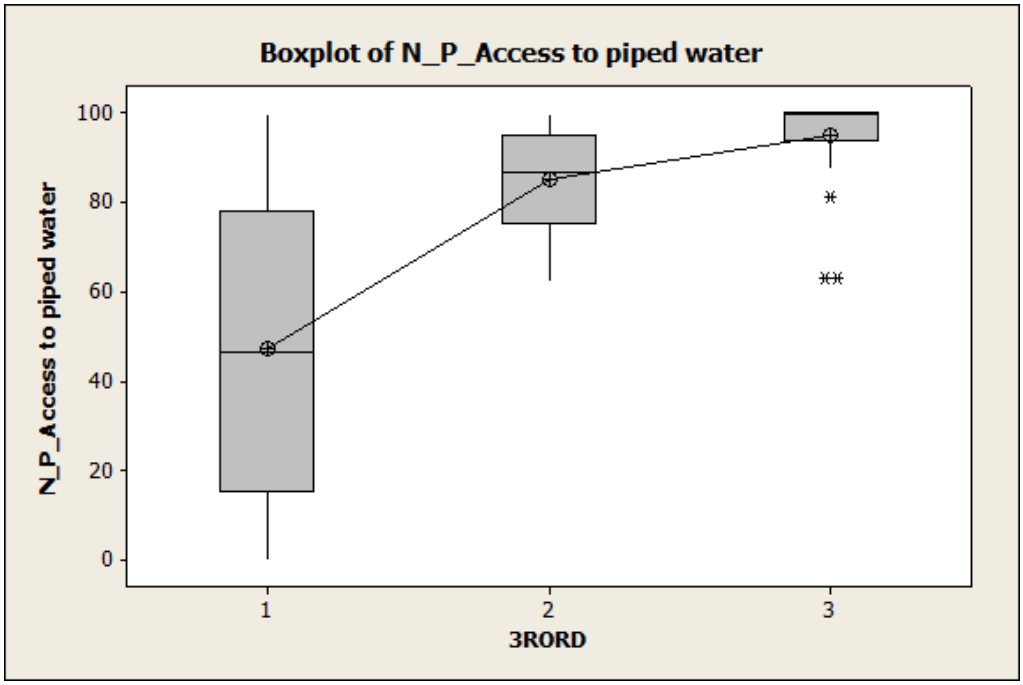
Source	DF	SS	MS	F	P
3RORD	2	58626	29313	37,63	0,000
Error	129	100488	779		
Total	131	159113			

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*

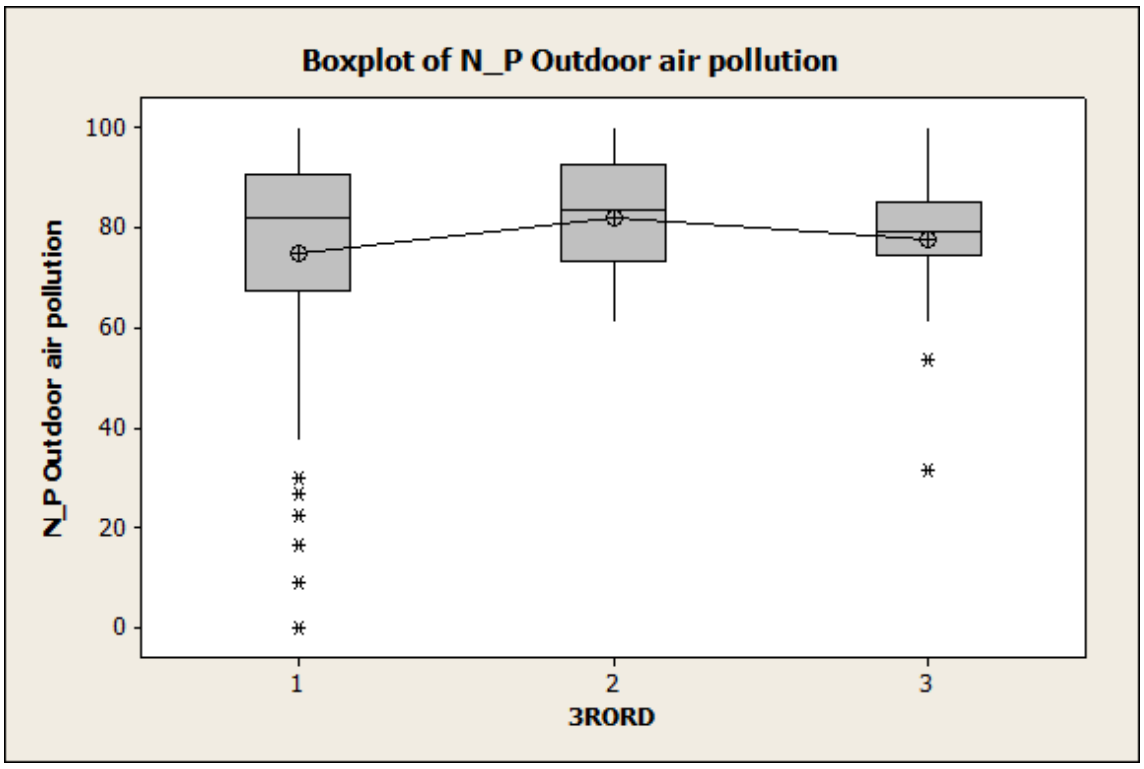
Source	DF	SS	MS	F	P
3RORD	2	937	469	1,32	0,271
Error	129	45848	355		
Total	131	46786			

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

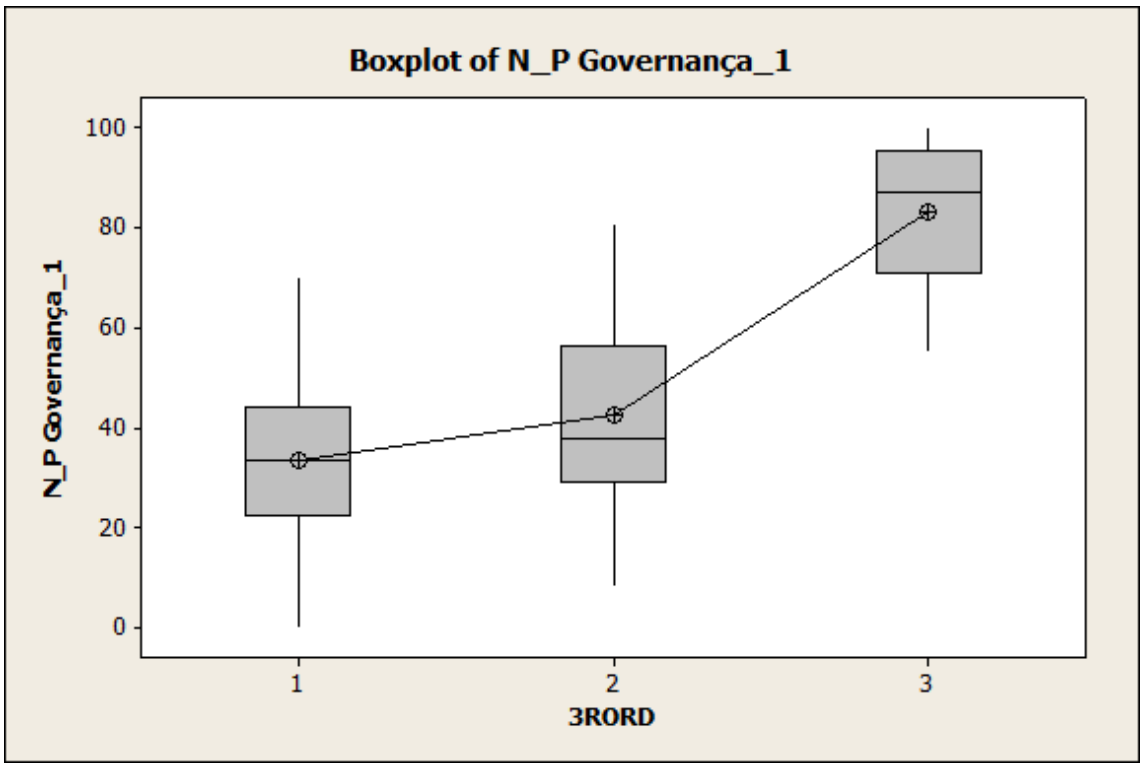
Source	DF	SS	MS	F	P
3RORD	2	51019	25510	97,02	0,000
Error	129	33919	263		
Total	131	84938			

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

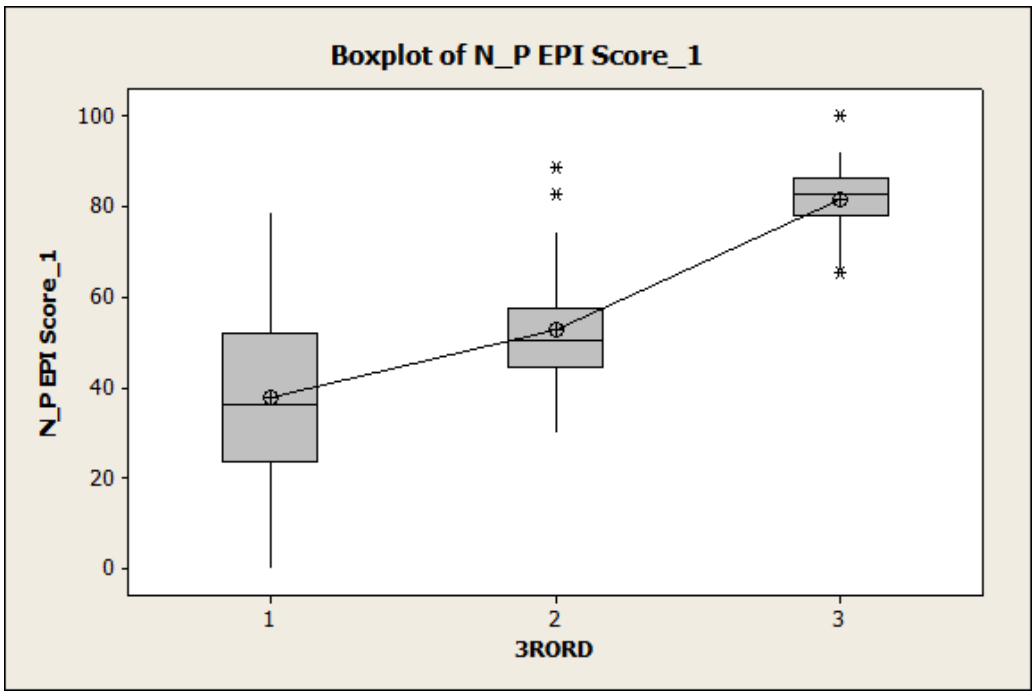
Source	DF	SS	MS	F	P
3RORD	2	38872	19436	69,68	0,000
Error	129	35982	279		
Total	131	74855			

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*

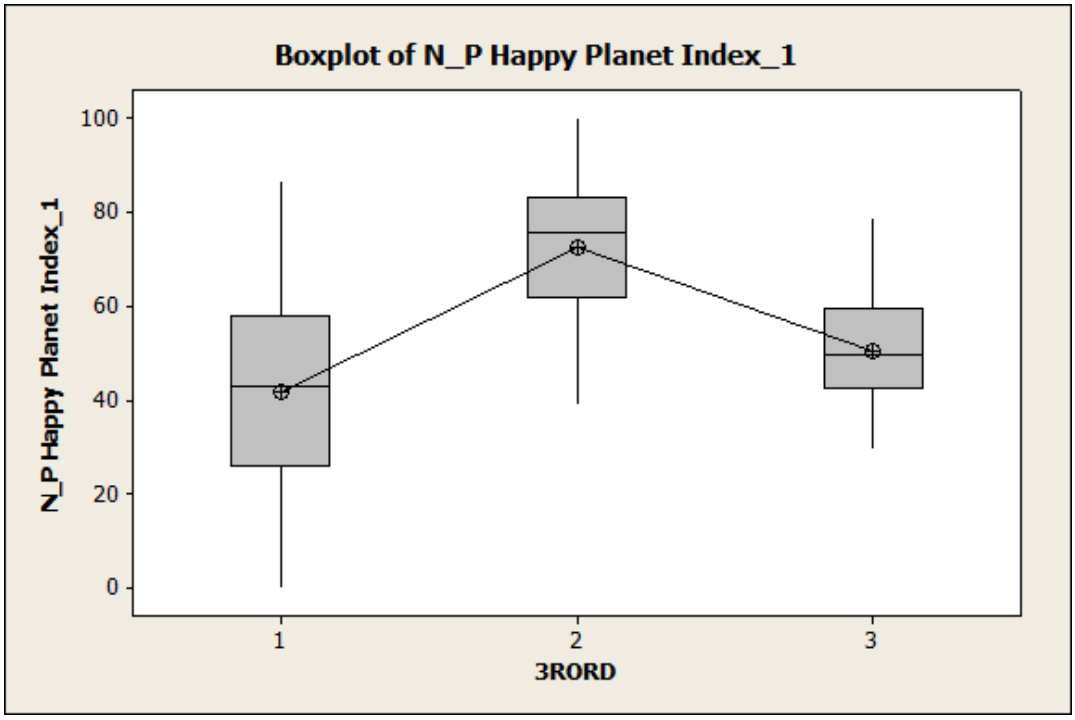
Source	DF	SS	MS	F	P
3RORD	2	15845	7923	22,69	0,000
Error	129	45048	349		
Total	131	60893			

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			

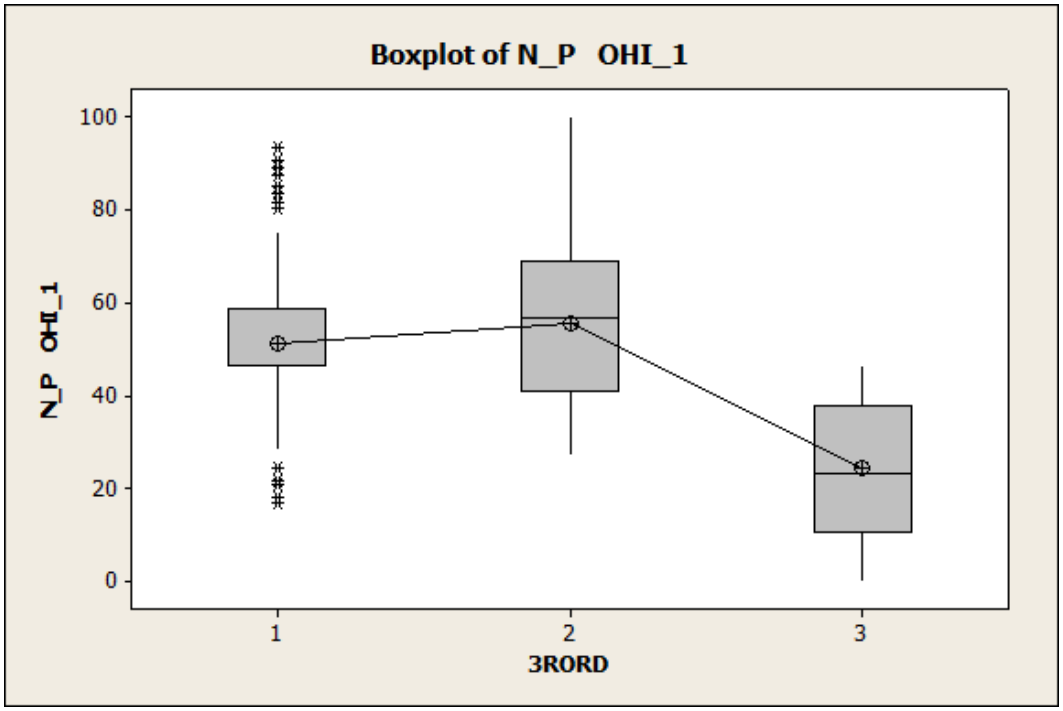
S = 16,90 R-Sq = 31,35% R-Sq(adj) = 30,29%

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			

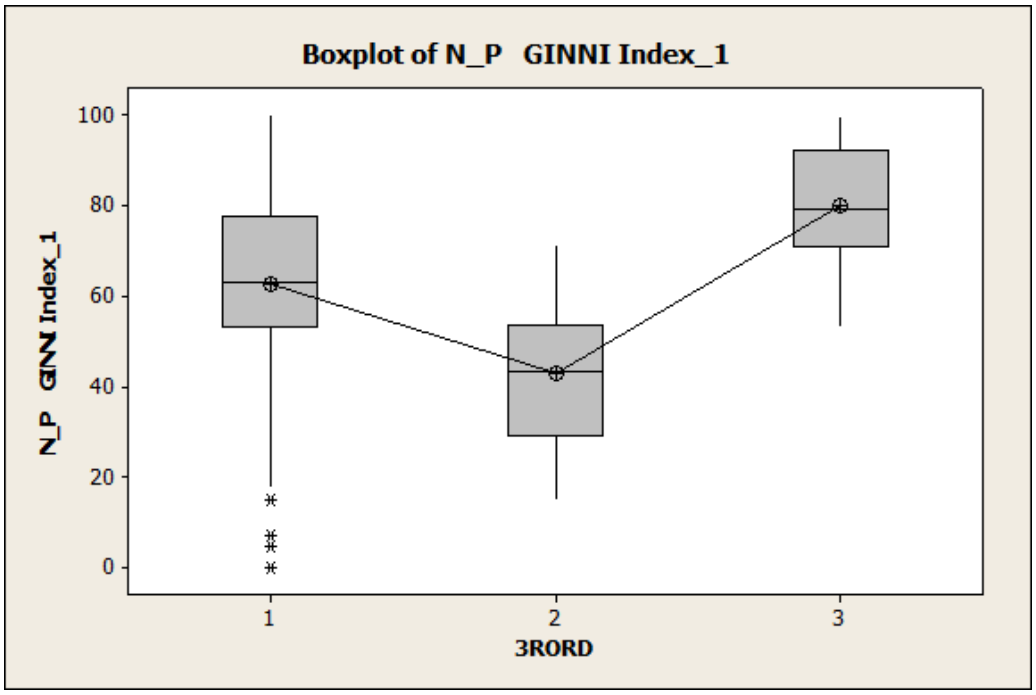
S = 18,13 R-Sq = 27,49% R-Sq(adj) = 26,36%

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

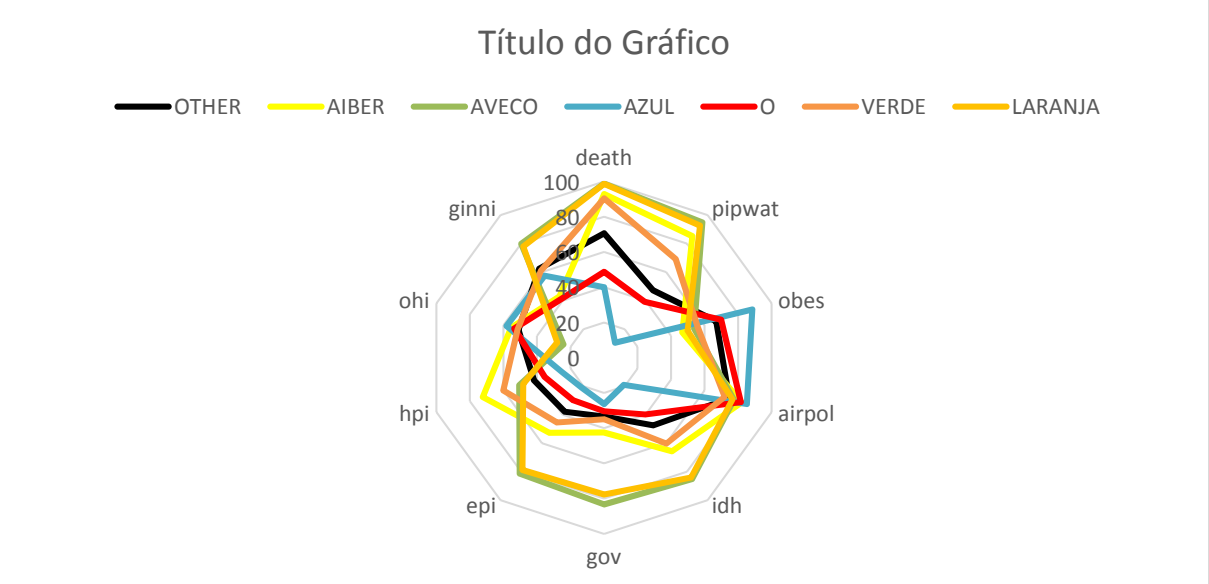
Após o cálculo das médias por região referente a cada variável, através do ANOVA, foi criada uma tabela com todas as médias, e foi possível fazer as relações entre as mesmas. Ver a seguir (Tabela 119).

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

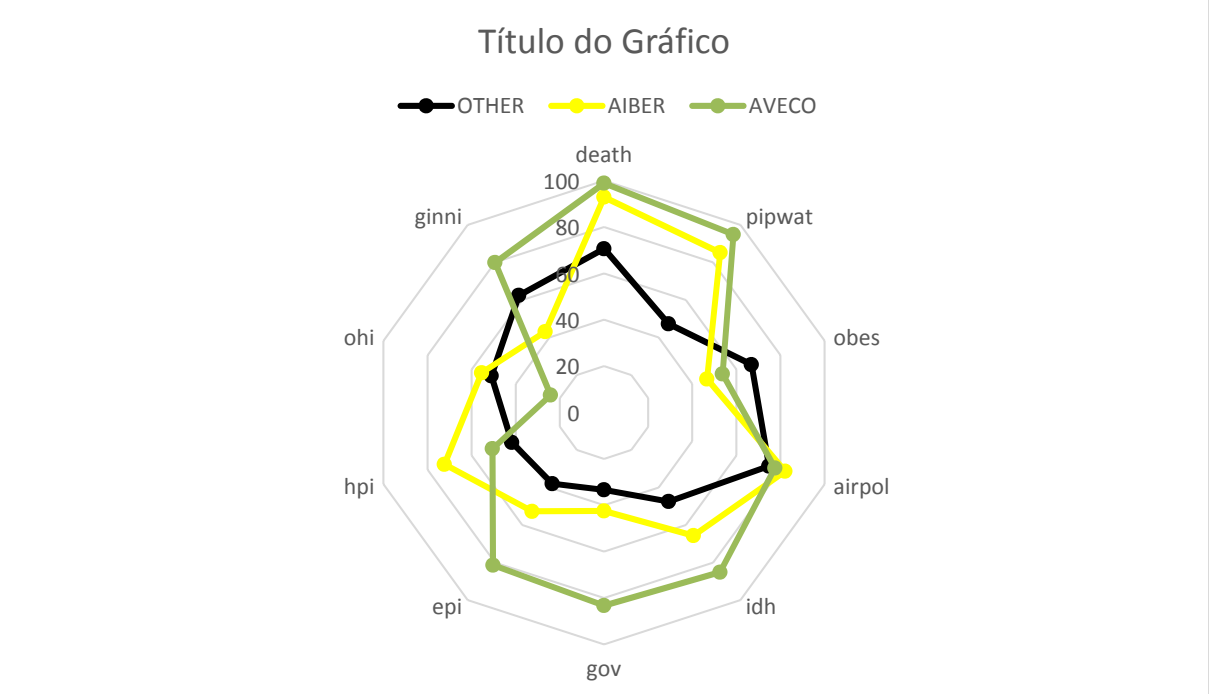
Fonte: Elaborado pela autora (Excel)

Gráfico 130



Fonte: Elaborado pela autora (Excel)

Gráfico 131



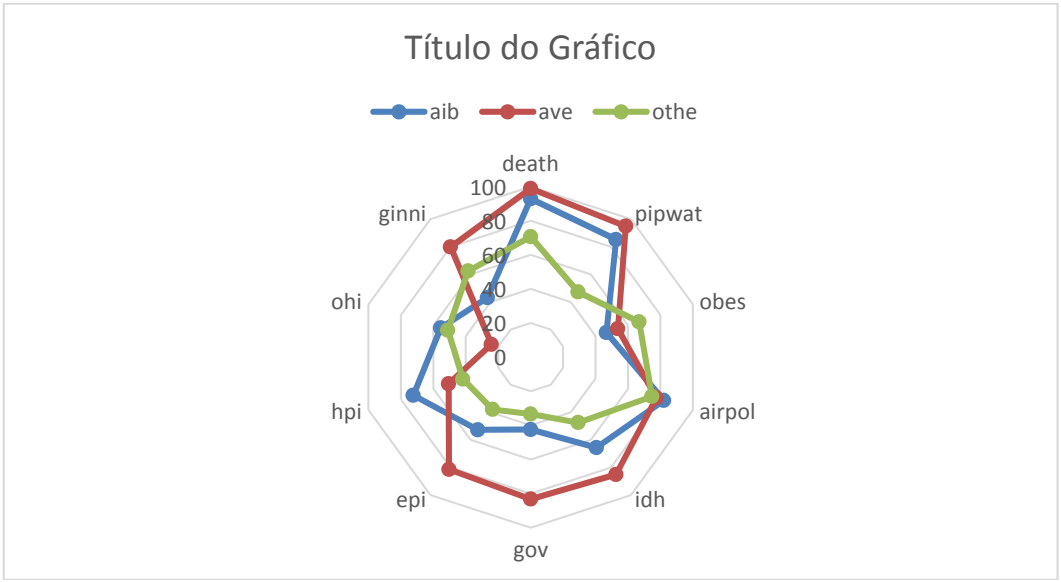
Fonte: Elaborado pela autora (Excel)

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

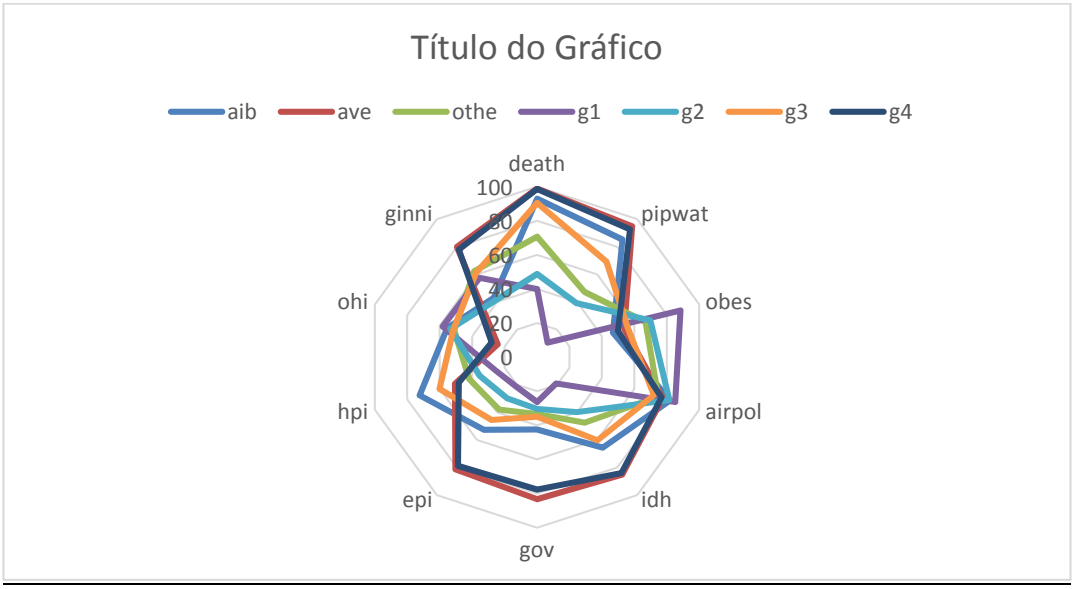
Fonte: Elaborado pela autora (Excel)

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

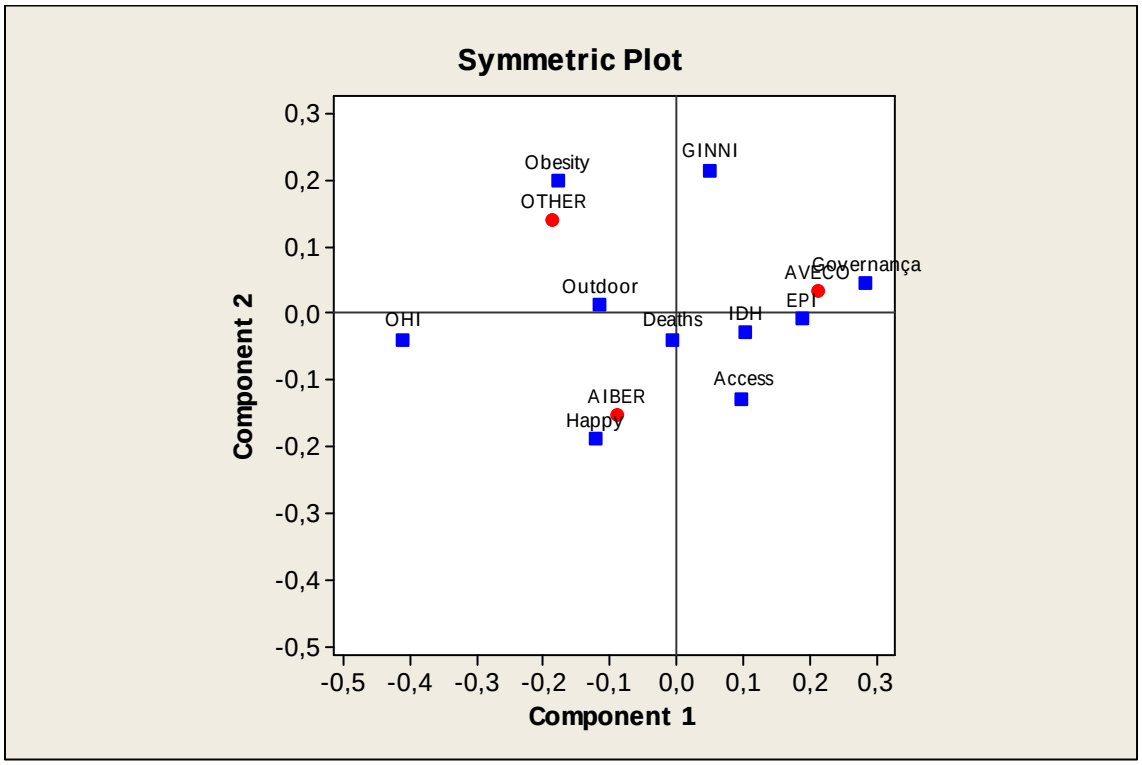
Tabela 124. *Column Contributions*

					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

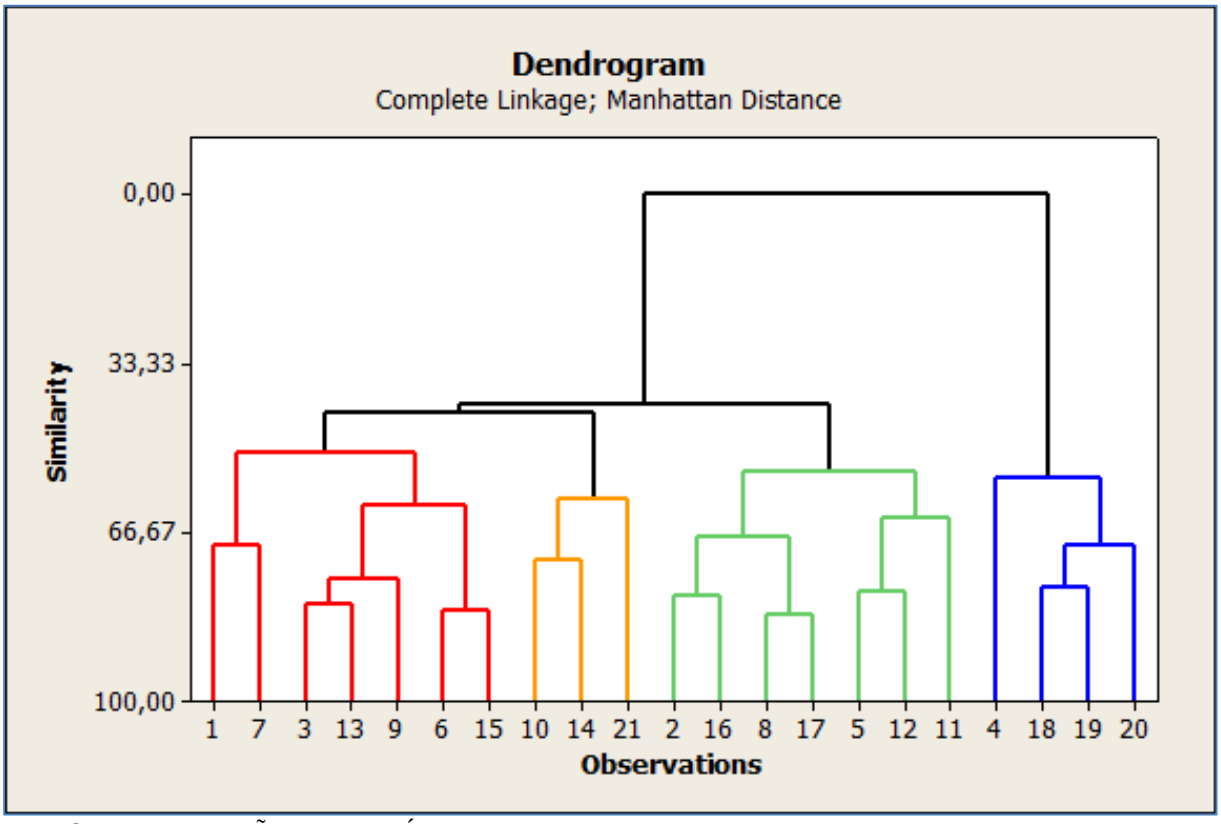
Fonte: Elaborado pela Autora

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

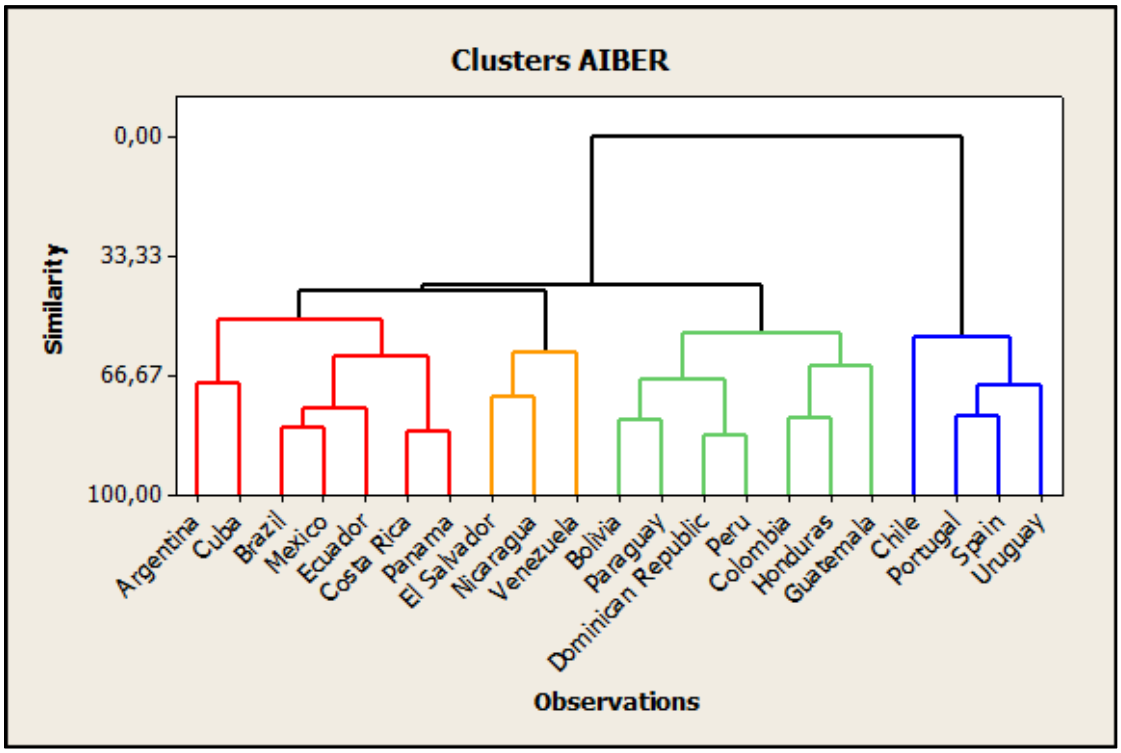
Gráfico 135. Dendrograma Manhattan



Legenda: EMBAIXO SÃO OS 21 PAÍSES: De 1 a 20.
Fonte: Elaborado pela Autora

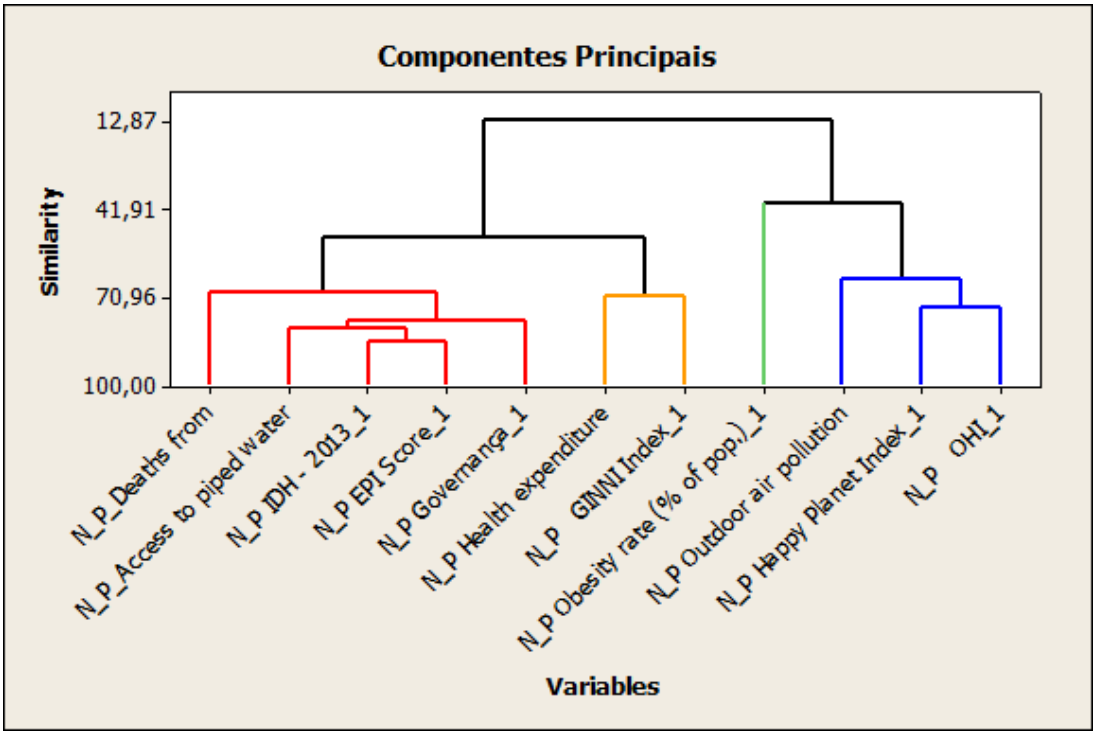
Agora, buscando analisar através do Comando STAT> MULTIVARIATE> CLUSTER VARIABLES SÓ COM OS PAÍSES AIBER, tem-se:

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

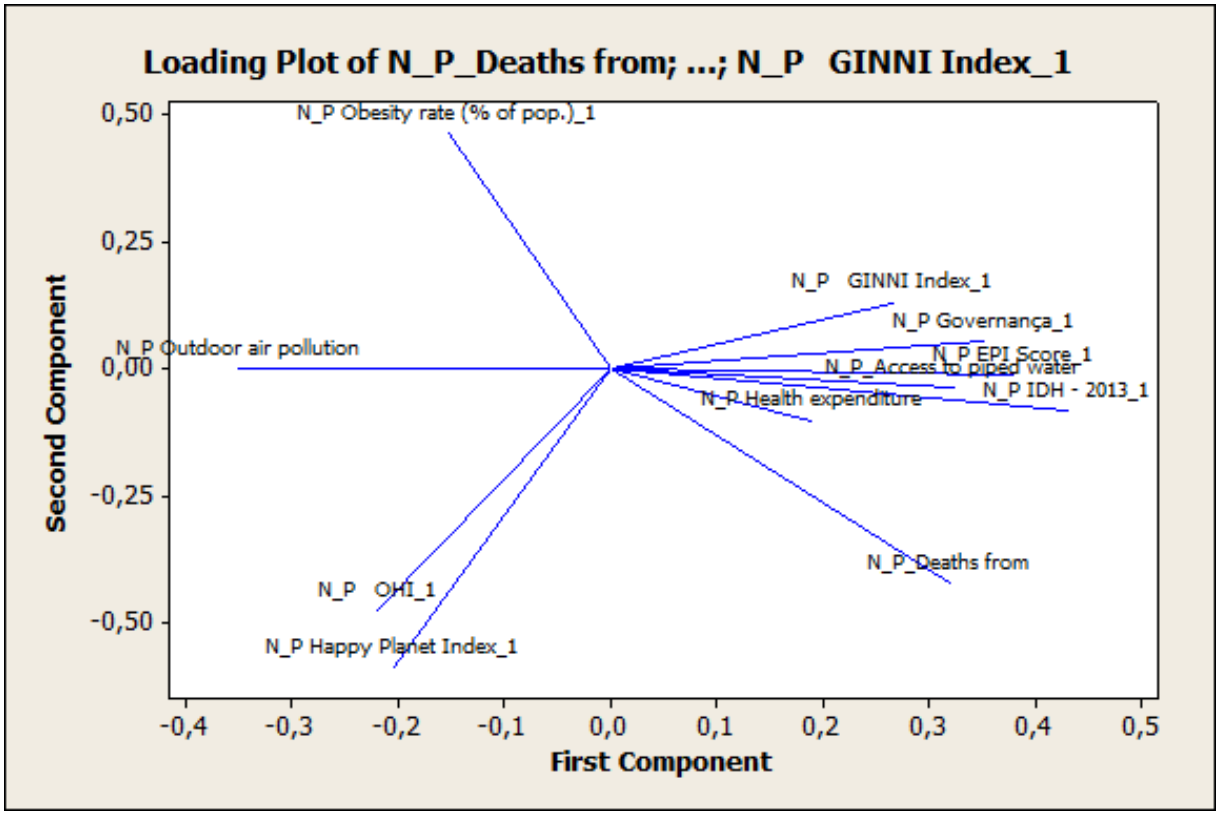
Tabela 141. Grand Centroid

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

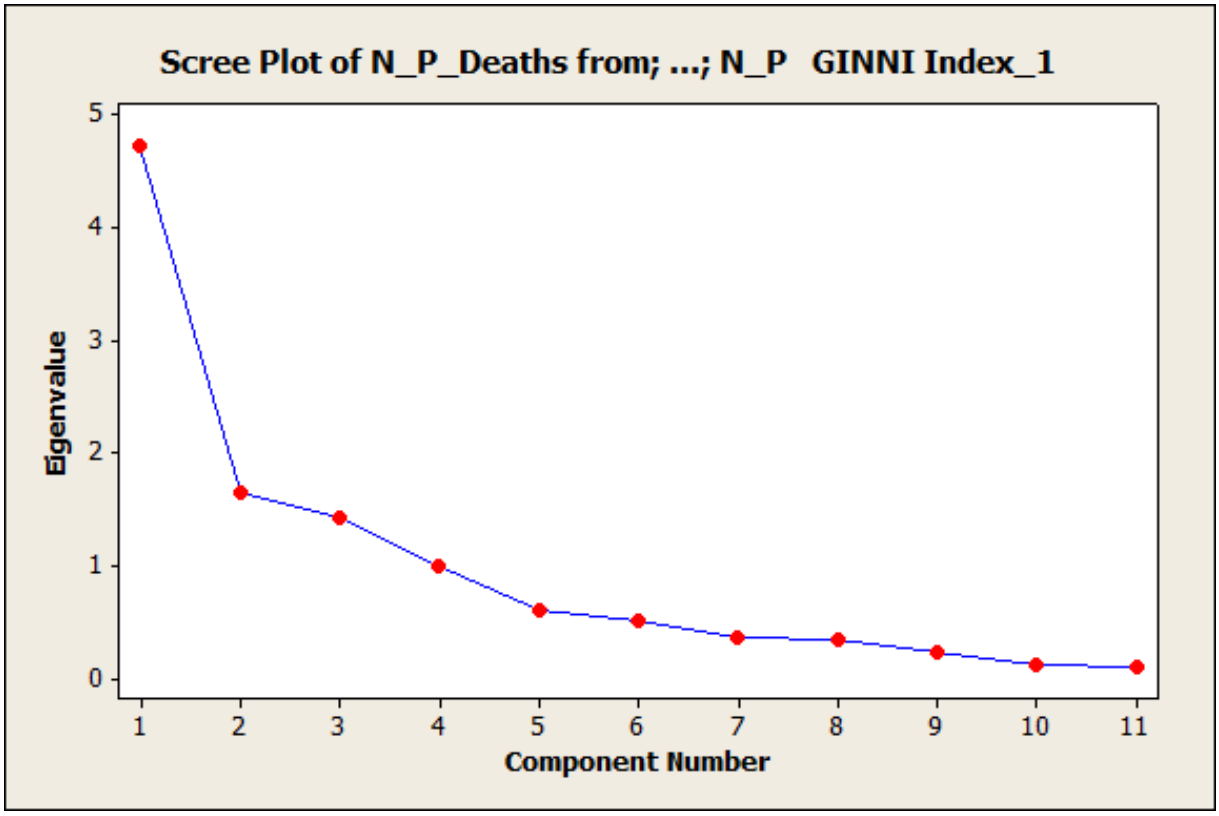
	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,0000	31,4315	53,1958	45,7710
Cluster2	31,4315	0,0000	72,9847	40,8099
Cluster3	53,1958	72,9847	0,0000	87,3415
Cluster4	45,7710	40,8099	87,3415	0,0000

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

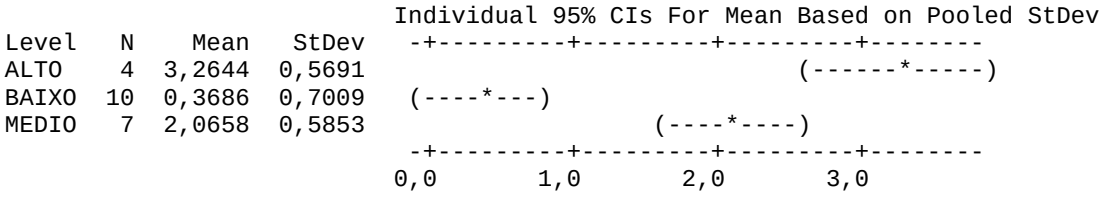
Através da Codificação, ordenar a planilha de países no MINITAB.
De 2 para 1 ALTO
De 4 para 2 MEDIO
De 1 para 3 BAIXO
De 3 para 4 MUITO BAIXO

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%

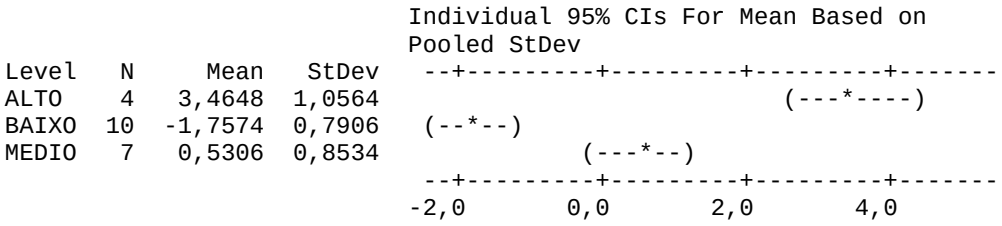


Pooled StDev = 0,6433

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).

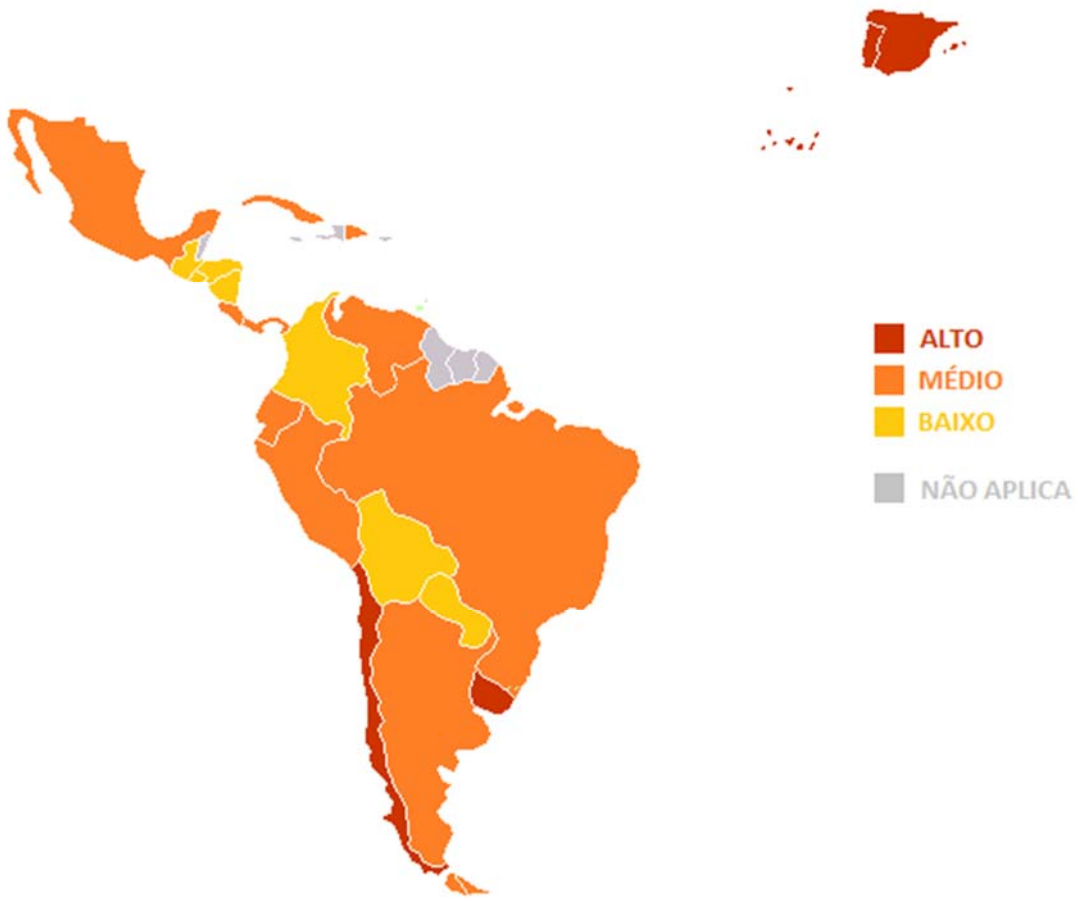
O Componente P1 será batizado de Indicador CRISMA (Cr_I_Saúde_Meio_Ambiente) – aceitando a sugestão do Prof. Dr. Hoyos.

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.