



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

## **MÉTODOS**

### **TRABALHO FINAL**

### **PROJETO ORIBER**

## **Tema 10. MUDANÇAS CLIMÁTICAS E ENERGIAS RENOVÁVEIS**

Disciplina: Métodos Quantitativos

Professor Dr. Arnaldo José de Hoyos Guevara

**JORGE TENÓRIO FERNANDO**

**junho de 2015**

## SUMÁRIO

1.				
INTRODUÇÃO.....				4
2. ENTENDENDO OS DADOS				
2.1 – Os Indivíduos. ....				4
2.2 As Variáveis.....				4
2.3 A Tabela de Dados.....				6
3. ANÁLISE DAS VARIÁVEIS				
3.1 VARIÁVEIS CATEGÓRICAS				
3.1.1 Variável: “País” .....				7
3.2 VARIÁVEIS QUANTITATIVAS				
3.2.1 Variável: “Índice de Progresso Social” .....				8
3.2.2 Variável: “Acesso a Eletricidade” .....				9
3.2.3 Variável: “Qualidade do Fornecimento de Eletricidade”.....				10
3.2.4 Variável: “Mortes Atribuíveis à Poluição Atmosférica” .....				11
3.2.5 Variável: “Emissão de Gases Estufa (CO2 relativo ao PIB)” .....				12
3.2.6 Variável: Retirada de Água (sobre % de recursos disponíveis) .....				13
3.2.7 Variável: Biodiversidade e Habitat .....				14
3.2.8 Variável: IDH .....				15
3.2.9 Variável: Consumo de Energia Elétrica (kWh per capita) .....				16
3.2.10 Variável: Energia Nuclear e Alternativa .....				17
3.2.11 Variável: Combustíveis Renováveis e Gastos .....				18
3.2.12 Variável: Consumo de Energia de Combustíveis Fosséis .....				19
3.2.13 Variável: Índice Mundial de Risco a Desastres Naturais .....				20
3.2.14 Variável: Exposição .....				21
3.2.15 Variável: Vulnerabilidade .....				22
3.2.16 Variável: Suscetibilidade .....				23
3.2.17 Variável: Incapacidade para Enfrentar Situações Difíceis .....				24
3.2.18 Variável: Capacidade de Adaptação.....				25
3.2.19 Variável: Governança .....				26
3.2.20 Variável: EPI Score .....				27
3.2.21 Variável: Happy Planet Index .....				28
3.2.22 Variável: GINNI Index.....				29
4. TABELA COMPARATIVA DA ANÁLISE DESCRITIVA .....				30
5. RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS				
5.1 Dendrograma .....				31
5.2 Correlação Linear.....				32
5.3 Gráficos de Dispersão				
5.3.1 Dispersão entre as Variáveis Governança e Social Progress Index .....				33
5.3.2 Dispersão entre as Variáveis Social Governança e EPI Score .....				34
5.3.3 Dispersão entre as Variáveis EPI Score e Social Progress Index .....				35
5.3.4 Dispersão entre as Variáveis “Falta de Capacidade para Enfrentar Situações Difíceis” e “Falta de Capacidade de Adaptação” .....				36
5.3.5 Dispersão entre as Variáveis Suscetibilidade e Social Progress Index.....				37

5.3.6 Dispersão entre as Variáveis Suscetibilidade e Acesso à Eletricidade .....	38
5.3.7 Dispersão entre as Variáveis Suscetibilidade e IDH .....	39
5.4 Regressão .....	40

## 6. ANÁLISE DE VARIANCIA

6.1	Variável	Social	Progress	Index
.....	.....	.....	.....	.....
.....	.....	.....	.....	.....
6.2	Variável Governança.....	.....	.....	.....
6.3	Variável IDH .....	.....	.....	.....
6.4	Variável	Happy	Planet	Index.....
6.5	Variável	GINNI	Index	.....
6.6	Variável EPI Score .....	.....	.....	.....
6.7	Quadro-resumo – TESTE F.....	.....	.....	.....

## 7. ANALISE DE GRÁFICO RADAR REDUZIDO

7.1	Gráfico Radar - Todas as Variáveis (6) .....	50
7.2	Gráfico Radar - Variáveis “SPI”, “IDH” e “Governança” .....	51
7.3	Gráfico Radar - Variáveis “EPI Score”, “Happy Planet” e “GINNI” .....	51

## 8. PESQUISA POR AMOSTRAGEM

8.1	Variável Vulnerabilidade – Amostra de 25, 50, 100 e 132 Indivíduos .....	52
8.2	Variável Falta de Capacidade de Adaptação – Amostra de 25, 50, 100 e 132 Indivíduos .....	57
8.3.	Variável Índice de Risco a Desastres – Amostra de 25, 50, 100 Indivíduos .....	61

## 9. CORRELAÇÕES E DENDROGRAMAS .....

## 10. ANÁLISE DOS COMPONENTES PRINCIPAIS .....

## 11. ANÁLISE DISCRIMINANTE .....

## 12. REGRESSÃO LOGÍSTICA .....

## 13. ÁRVORES DE CLASSIFICAÇÃO .....

## 14. ANÁLISE DE CORRESPONDÊNCIA .....

## 15. UM OLHAR SOBRE OS PAÍSES IBEROAMERICANOS

### 15.1 ANÁLISE DE CLUSTERS DE VARIÁVEIS E PAÍSES

15.1.1.	Dendrogramas	das	Variáveis
.....	.....	.....	.....
.....	.....	.....	.....

15.1.2/3	Exame e Seleção	das	Variáveis	Sintéticas	e	Analíticas
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....

15.1.4	Reagrupamento de Países.....	96
--------	------------------------------	----

### 15.2 ANALISE DE VARIANCIA

15.2.1.	Análise de Variância com as Variáveis Sintéticas Seleccionadas .....	97
---------	--	----

15.2.2	Análise de Variância com as Variáveis Analíticas Seleccionadas .....	103
--------	--	-----

15.3.	REAGRUPAMENTO FINAL PAÍSES – AIBER, AVECO, OTHERS .....	106
-------	---	-----

### 15.3 ANALISE DE VARIANCIA

15.3.1.	Análise de Variância com as Variáveis Sintéticas Seleccionadas .....	107
---------	--	-----

15.3.2	Análise de Variância com as Variáveis Analíticas Seleccionadas	111
15.4	Quadro-Resumo: F Value	113
15.5.	ANÁLISE DISCRIMINANTE	114
15.6	REGRESSÃO LOGÍSTICA	115
15.7.	ÁRVORES DE CLASSIFICAÇÃO	117
15.8	ANÁLISE DE CORRESPONDÊNCIA	
15.8.1/2.	Médias e Desvios das Variáveis Sintéticas - 3 Regiões	119
15.8.3/4.	Médias e Desvios das Variáveis Analíticas - 3 Regiões	123
16.	DENDROGRAMA – AIBER	126
17.	ANÁLISE DISCRIMINANTE – AIBER	128
18.	ARVORE DE CLASSIFICAÇÃO – AIBER	129
19.	ANÁLISE DE VARIÂNCIA – AIBER	131
20.	CONSIDERAÇÕES FINAIS	135

## 1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise de **Mudanças Climáticas e Energias Renováveis** globalmente, a partir de uma série de indicadores, visando a identificar diferenças e semelhanças que auxiliem no entendimento do status do tratamento do tema pelo grupo de países da Ibero-América, bem como o apontamento dos avanços, destaques positivos e negativos em cotejo com os países de economia avançada e os demais países. Os softwares estatísticos utilizados são o **minitab** e o **SPSS**.

## 2. ENTENDENDO OS DADOS

### 2.1 Os Indivíduos

Os indivíduos desta análise são 132 países de todos os continentes, divididos em três grupos, quais sejam: **1) AIBER**, composto por 21 países, é basicamente o grupo de países latino-americanos mais Portugal e Espanha; **2) AVECO**, composto por 27 países, contempla os países de economia avançada localizados na Europa, somados ao Japão e Coreia do Sul (Ásia), Nova Zelândia e Austrália (Oceania), Estados Unidos e Canadá na América do Norte. **3) OUTROS**, composto por todos os demais países, em número de 87. Os dados analisados de cada país são as variáveis que descrevemos a seguir.

### 2.2 As Variáveis

São 24 as variáveis regulares desta pesquisa, além dos países. Elas são melhor explicadas na Tabela 1 apresentada a seguir.

**Tabela 1.** As Variáveis

	Variável	Significado	Tipo	Unid. Medida
--	----------	-------------	------	--------------

1	<b>Região</b>	Área geográfica/ política/ linguística comum	Categórica	Unidade
2	<b>Índice de Progresso Social</b>	Desempenho social por 1) Necessidades Básicas, 2) Bem-Estar, 3) Oportunidades	Quant	0 a 100
3	<b>Acesso a eletricidade</b>	População com acesso à eletricidade	Porcent	0 a 100
4	<b>Qualidade do fornecimento de eletricidade</b>	Nível de serviço de fornecimento de eletricidade para a população (1 baixa 7 alta)	Quant	1 a 7
5	<b>Mortes atribuídas à poluição atmosférica</b>	Mortes resultantes de emissões de Co2, expressas em taxa por 100,000 pessoas.	Quant	0 a 100
6	<b>Emissões de gases estufa</b>	CO2 por PIB 4:<100; 3:100–200; 2:200–1000; 1:1000–2000; 0:>2000	Quant	0 a 4
7	<b>Retirada de água</b>	Retirada de água em função da disponibilidade - Altíssima 4-5 (>80%); Alta 3-4 (40–80%); Méd.-alta 2-3 (20–40%); Méd.-baixa 1–2 (10–20%); Baixa 0–1(<10%)	Quant	0 a 5
8	<b>Biodiversidade e habitat</b>	Proteção do bioma em áreas terrestres/marítimas bem como de espécies ameaçadas (0=sem proteção; 100=alta proteção)	Quant	0 a 100
9	<b>IDH – 2013</b>	Progresso de uma nação a partir de três dimensões: renda, saúde e educação	Quant	0 a 100
10	<b>Consumo de Energia Elétrica</b>	Consumo de energia (kWh per capita)	Quant	0 a 100
11	<b>Uso de Energias alternativas e nuclear</b>	Não produzem dióxido de carbono (hidro, nuclear, geotérmica) % da energia total	Porcent	0 a 100
12	<b>Combustíveis renováveis e resíduos</b>	Biomassa, biogás, e resíduos, medidos como % da energia total.	Porcent	0 a 100
13	<b>Consumo energia de combustíveis fósseis</b>	Carvão, óleo, petróleo e gás natural. % do total	Porcent	0 a 100
14	<b>Índice mundial de risco a desastres naturais</b>	1) Exposição a desastres naturais; 2) Suscetibilidade; 3) Capacidade de enfrentar situações difíceis e 4) Capacidade de adaptação	Porcent	0 a 100
15	<b>Exposição</b>	Exposição a desastres naturais	Porcent	0 a 100
16	<b>Vulnerabilidade</b>	Fatores físicos, econômicos ou ambientais que levem a 1) Suscetibilidade; 2) capacidade de enfrentar situações difíceis 3) capacidade de adaptação	Porcent	0 a 100
17	<b>Suscetibilidade</b>	Probabilidade de sofrer perda, dano ou ruptura em evento extremo ou desastre natural em função da estrutura disponível (pública, moradia, nutrição, dist renda)	Porcent	0 a 100
18	<b>Capacidade de enfrentar situações difíceis</b>	Capacidades prontamente disponíveis para resolver/minimizar situações adversas	Porcent	0 a 100
19	<b>Capacidade de adaptação</b>	Medida da existência de estratégias de longo prazo para responder estruturalmente a situações adversas futuras	Porcent	0 a 100
20	<b>Governança</b>	Capacidade de conduzir os processos de formulação, execução e avaliação de políticas públicas integrando instrumentos e mecanismos de gestão	Quant	0 a 100
21	<b>EPI Score</b>	Mede 1) proteção da saúde frente a riscos ambientais e 2) proteção dos ecossistemas	Quant	0 a 100
22	<b>Happy Planet Index</b>	Capacidade do país proporcionar bem-estar sustentável aos seus cidadãos.	Quant	0 a 100
23	<b>GINNI Index</b>	Medida da distribuição igualitária de renda (0	Quant	0 a 100

		igualdade total; 100 desigualdade total)		
24	Ocean Health Index	Mede os fatores ecológicos, políticos, econômicos, políticos, e sociais do uso e preservação da fauna e flora marinhas, por país	Quant	0 a 100

## 2.3 A Tabela de Dados

A primeira visão da base de dados mostrou a seguinte configuração:

### Descriptive Statistics

Variable	N	N*	Mean	SE Mean	Min	Max
<i>Social Progress Index</i>	132	0	63,67	1,24	32,60	88,24
<i>Access to electricity (%)</i>	132	0	78,33	2,80	4,00	100,00
<i>Quality of electricity s</i>	121	11	4,515	0,142	1,273	6,754
<i>Outdoor air pollution at</i>	130	2	15,75	1,12	0,00	67,00
<i>Greenhouse gas emissions</i>	132	0	1177	228	170	28764
<i>Water withdrawals as a p</i>	131	1	1,933	0,136	0,000	5,000
<i>Biodiversity and habitat</i>	132	0	59,51	2,43	0,00	100,00
<i>IDH - 2013</i>	125	7	0,6894	0,0144	0,3370	0,9440
<i>Electric power consumpti</i>	110	22	4289	586	92	52374
<i>Alternative and nuclear</i>	111	21	10,14	1,31	0,00	83,81
<i>Combustible renewables a</i>	112	20	18,96	2,30	0,00	93,08
<i>Fossil fuel energy consu</i>	110	22	69,54	2,44	4,21	100,00
<i>Índice de risco mundial</i>	131	1	6,959	0,363	1,170	28,250
<i>Exposición (%)</i>	131	1	14,514	0,682	2,930	52,460
<i>Vulnerabilidad (%)</i>	131	1	47,51	1,21	22,27	77,12
<i>Suceptibilidad (%)</i>	131	1	30,31	1,36	10,47	69,89
<i>falta de capacidades pa</i>	130	2	68,36	1,35	37,61	93,05
<i>falta de capacidades de</i>	131	1	43,79	1,09	20,96	71,21
<i>1. Governança</i>	132	0	53,38	1,88	15,22	100,00
<i>EPI Score</i>	132	0	52,48	1,44	18,43	87,67
<i>Happy Planet Index</i>	129	3	42,673	0,796	22,591	64,036
<i>GINNI Index</i>	119	13	39,024	0,782	24,820	63,140

## 3. ANÁLISE DAS VARIÁVEIS

Notamos haver algumas variáveis que não continham dados registrados em alguns países, provavelmente pelo fato de não terem sido medidos ou disponibilizados na origem. Com isso,

para poder trabalhar na base completa inserimos os valores médios de cada variável nos países faltantes, e, portanto, a tabela corrigida ficou assim:

### Descriptive Statistics

Variable	N	N*	Mean	SE Mean	Minimum	Maximum
Social Progress Index	132	0	63,67	1,24	32,60	88,24
Access to electricity (%)	132	0	78,33	2,80	4,00	100,00
Quality of electricity s	132	0	4,515	0,131	1,273	6,754
Outdoor air pollution at	132	0	15,75	1,10	0,00	67,00
Greenhouse gás emissions	132	0	1177	228	170	28764
Water withdrawals as a p	132	0	1,933	0,135	0,000	5,000
Biodiversity and habitat	132	0	59,51	2,43	0,00	100,00
IDH - 2013	132	0	0,6894	0,0136	0,3370	0,9440
Electric power consumpti	132	0	4289	488	92	52374
Alternative and nuclear	132	0	8,69	1,14	0,00	83,81
Combustible renewables a	132	0	18,96	1,95	0,00	93,08
Fossil fuel energy consu	132	0	69,54	2,03	4,21	100,00
Índice de riego mundial	132	0	6,959	0,360	1,170	28,250
Exposición (%)	132	0	14,514	0,677	2,930	52,460
Vulnerabilidad (%)	132	0	47,51	1,20	22,27	77,12
Suceptibilidad (%)	132	0	30,31	1,35	10,47	69,89
falta de capacidades pa	132	0	68,38	1,33	37,61	93,05
falta de capacidades de	132	0	43,79	1,08	20,96	71,21
1. Governança	132	0	53,38	1,88	15,22	100,00
EPI Score	132	0	52,48	1,44	18,43	87,67
Happy Planet Index	132	0	42,673	0,778	22,591	64,036
GINNI Index	132	0	39,024	0,705	24,820	63,140

### 3.1 Variáveis Categóricas

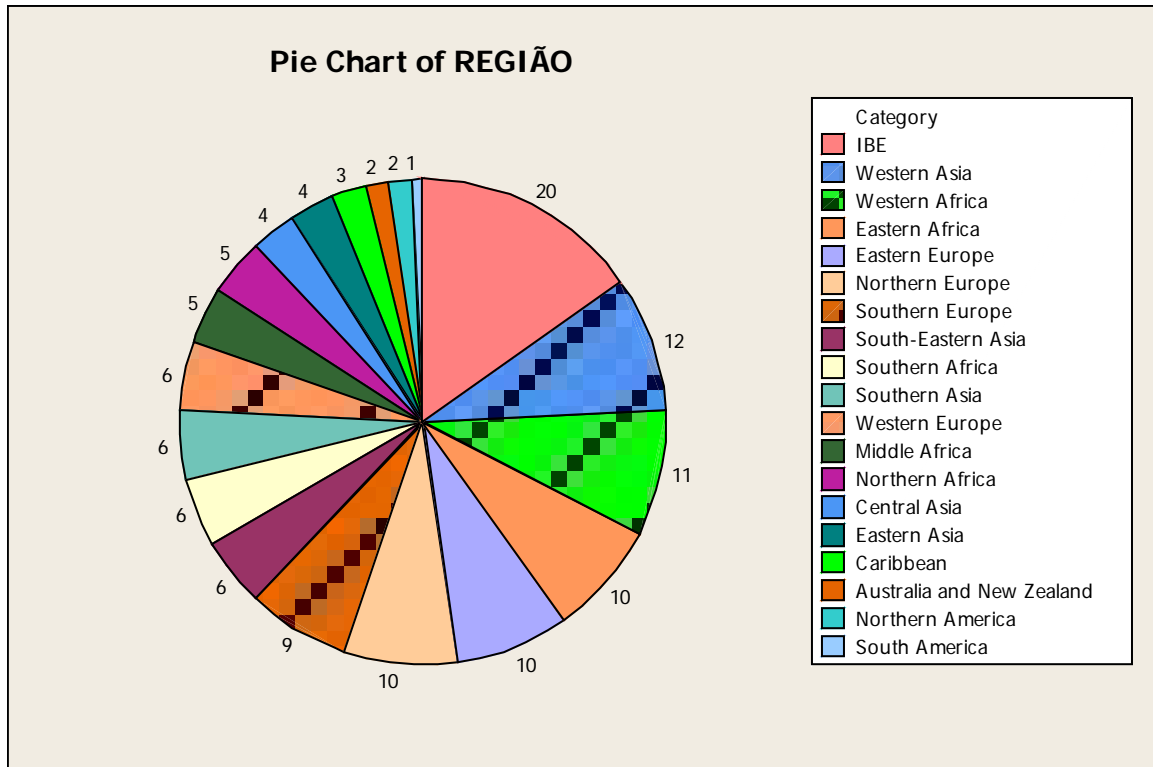
Este tipo de variável indica que o foco de concentração deve ser a análise de gráficos do tipo *pie chart* e barras.

#### 3.1.1 Variável: “Países”

Nossa amostra totaliza 132 países. Por se tratar do objeto de estudo, não se justifica a análise desta variável em si, mas em função das demais variáveis quantitativas.

#### 3.1.2 Variável: “Região”

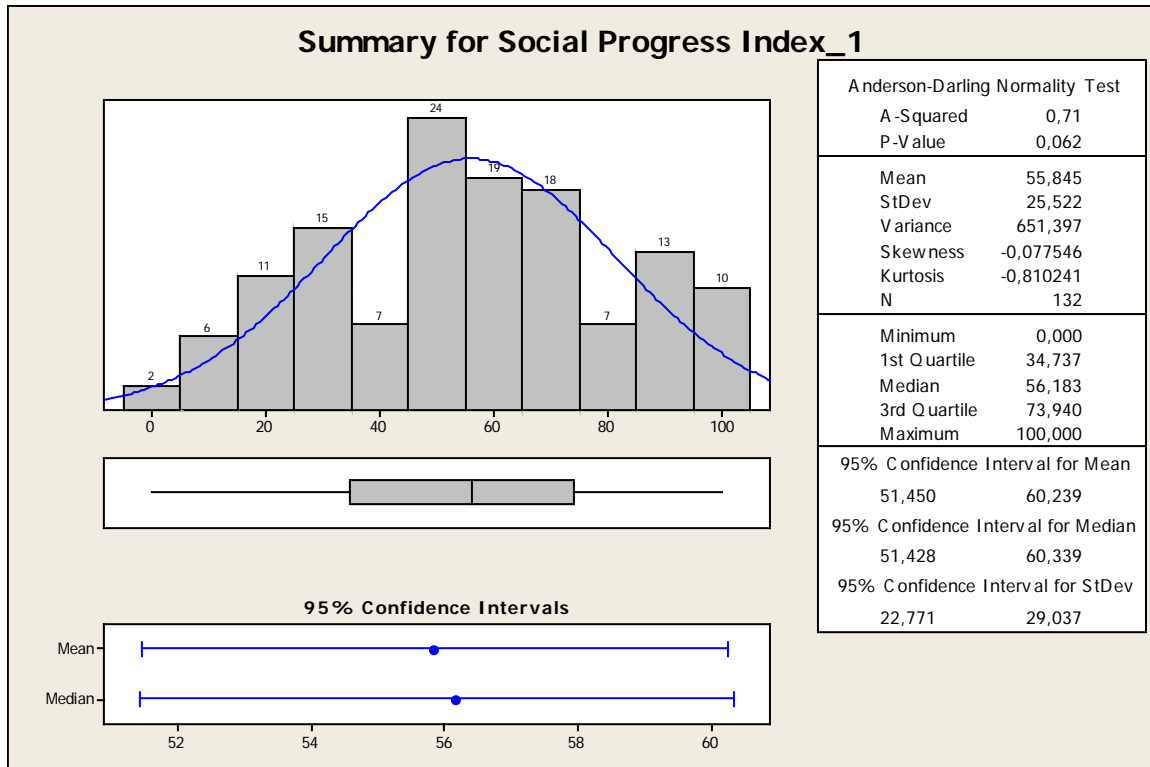




Com base na análise do gráfico acima podemos tecer alguns comentários sobre a distribuição dos países entre as regiões. Os países ibero-americanos são maioria (20) seguidos de países localizados nos continentes asiático e africano ocidentais, que juntos somam 23. Juntos, esses três grupos de países somam mais de um terço do total.

## 3.2 Variáveis Quantitativas

### 3.2.1 Variável: “Índice de Progresso Social”

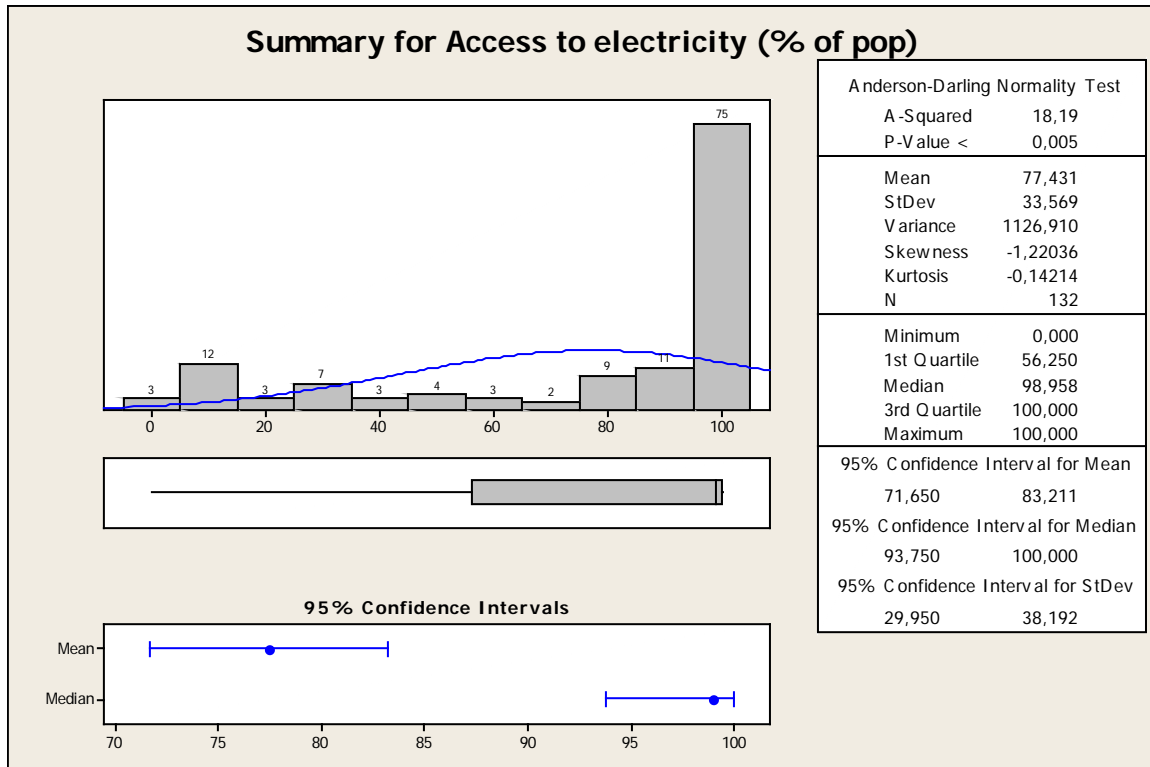


As principais observações que podemos fazer são:

No extremo positivo, encontramos um grupo de dez países, quase todos localizados no continente europeu – com exceção de Canadá e Austrália – com os índices mais altos, entre 95 e 100. No outro extremo encontramos 8 países, quase todos localizados no continente africano, exceção feita ao Iêmen, com os índices mais baixos, não acima de 14. Isso mostra o imenso abismo, em termos de desigualdade social, existente entre o continente europeu e o africano.

O Histograma nos mostra uma distribuição mais ao centro, com um número expressivo de países numa posição intermediária. Desse maior grupo, de 24 países, oito são oriundos do continente asiático, seis são africanos, outros seis são ibero-americanos, dois são da Europa do Leste e dois, da América do Sul e Caribe, e tem índice entre 45 e 55. O Brasil ocupa a posição 46 entre 132 países.

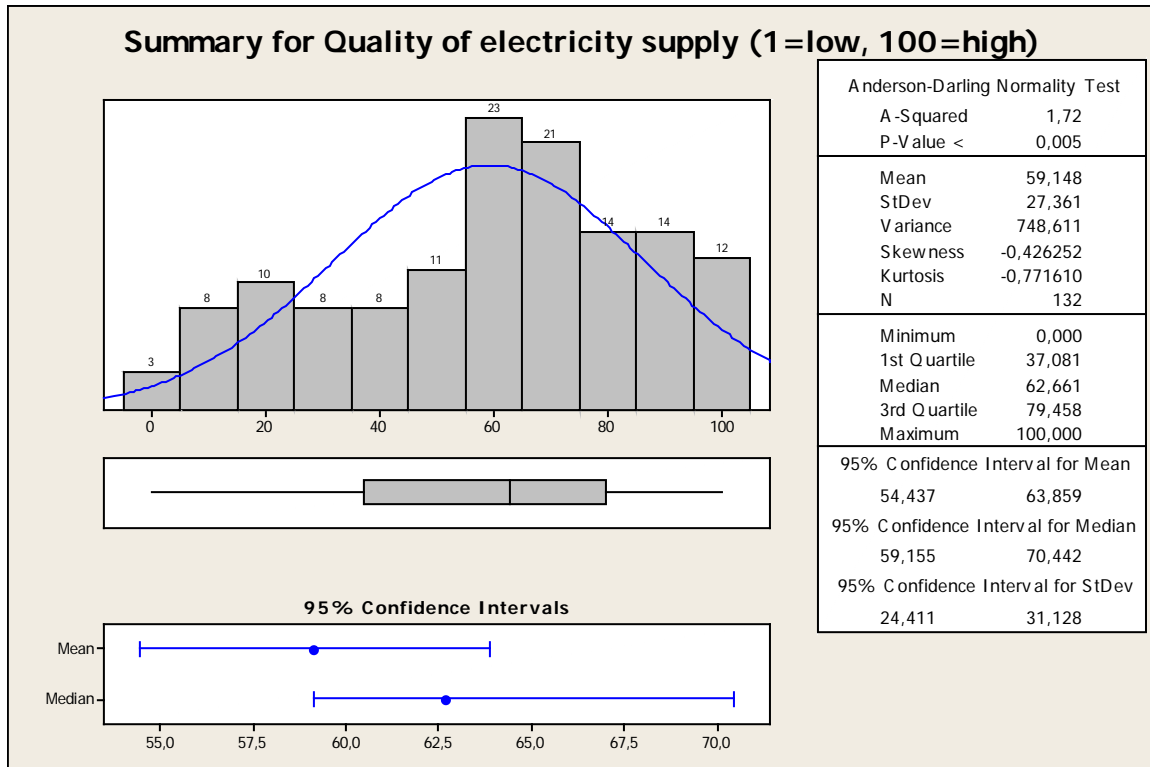
### 3.2.2 Variável: “Acesso a Eletricidade”



Com base nas informações acima, fazemos os seguintes comentários:

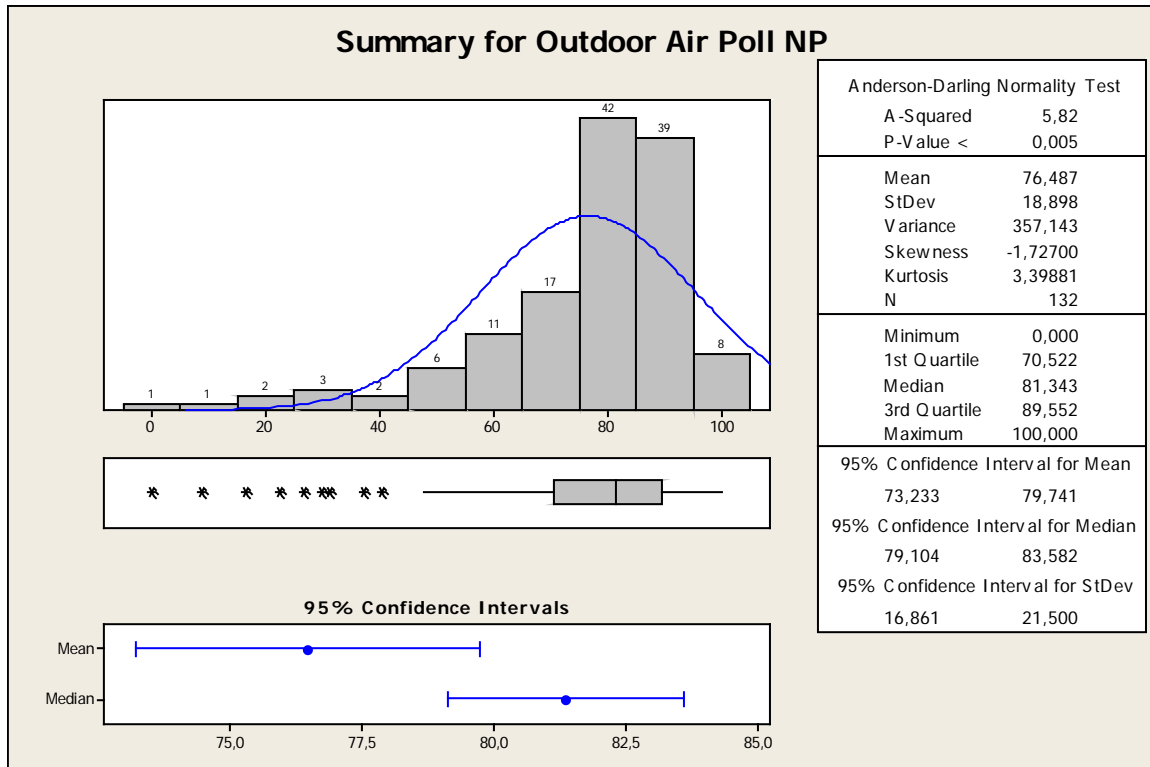
A distribuição é bastante assimétrica, com a mediana muito longe do centro, na extrema direita, concentrando 75 dos 132 países em um grande pico. O teste de normalidade de Anderson-Darling indica não tratar-se de uma distribuição Normal. Isso mostra que eletricidade está mais acessível para a população na maior parte dos países – 58 deles com 100%, inclusive Chile, Venezuela, Espanha e Portugal. Brasil aparece logo em seguida com mais de 98%, acompanhado de México, Uruguai e Costa Rica. O oposto está no continente africano, onde infelizmente ela não está disponível para a maioria da população (entre 0 e 1% na Liberia, Chade e Burundi). Considerando que no geral trata-se de países bastante populosos, pode-se dizer que ainda há muita gente literalmente vivendo no escuro naquela parte do mundo.

### 3.2.3 Variável: “Qualidade do Fornecimento de Eletricidade”



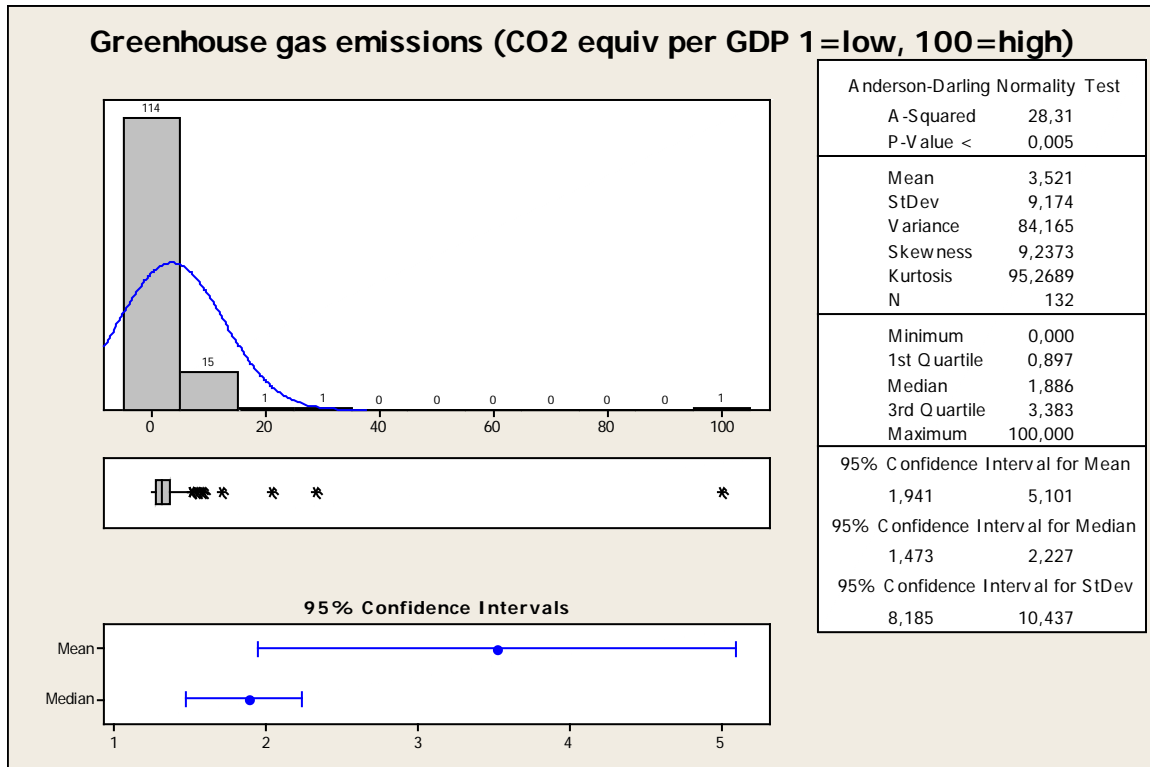
O gráfico mostra uma distribuição mais para a direita com importante concentração ao centro. Isso demonstra que quando há disponibilidade de energia elétrica ela tende a ser de boa qualidade. Novamente, onde isso ocorre com menos frequência é nos países africanos, sobretudo Nigéria, Angola, Chade, e Guiné, com índices menores de 10. Ressalte-se também a Venezuela como parte de este triste grupo. Por outro lado, na Europa predomina o melhor nível de serviço, com destaque para Suíça, Holanda, Áustria, Finlândia, Dinamarca, Islândia e Reino Unido, com índices entre 98 e 100. Brasil ocupa uma posição intermediária, na casa dos 60 pontos.

### 3.2.4 Variável: “Mortes Atribuíveis à Poluição Atmosférica”



Percebe-se que há uma concentração expressiva à direita, revelando que grande parte dos países agrupam-se em três blocos, dois deles representando picos. Juntos, os três grupos à esquerda perfazem 89 países, que se encontram com índice na faixa de 75 a 85. Há uma diferença importante das variáveis anteriores – aqui os países do continente africano têm um índice mais favorável comparado ao dos países desenvolvidos, precisamente devido a terem relativa baixa industrialização. Ressalte-se a presença dos *outliers*, países com índices bastante altos e distantes da média dos demais – grande parte países da Ásia ou Europa de Leste, tais como Rússia, Geórgia, Ucrânia, Cazaquistão e Armênia com índices abaixo de 20.

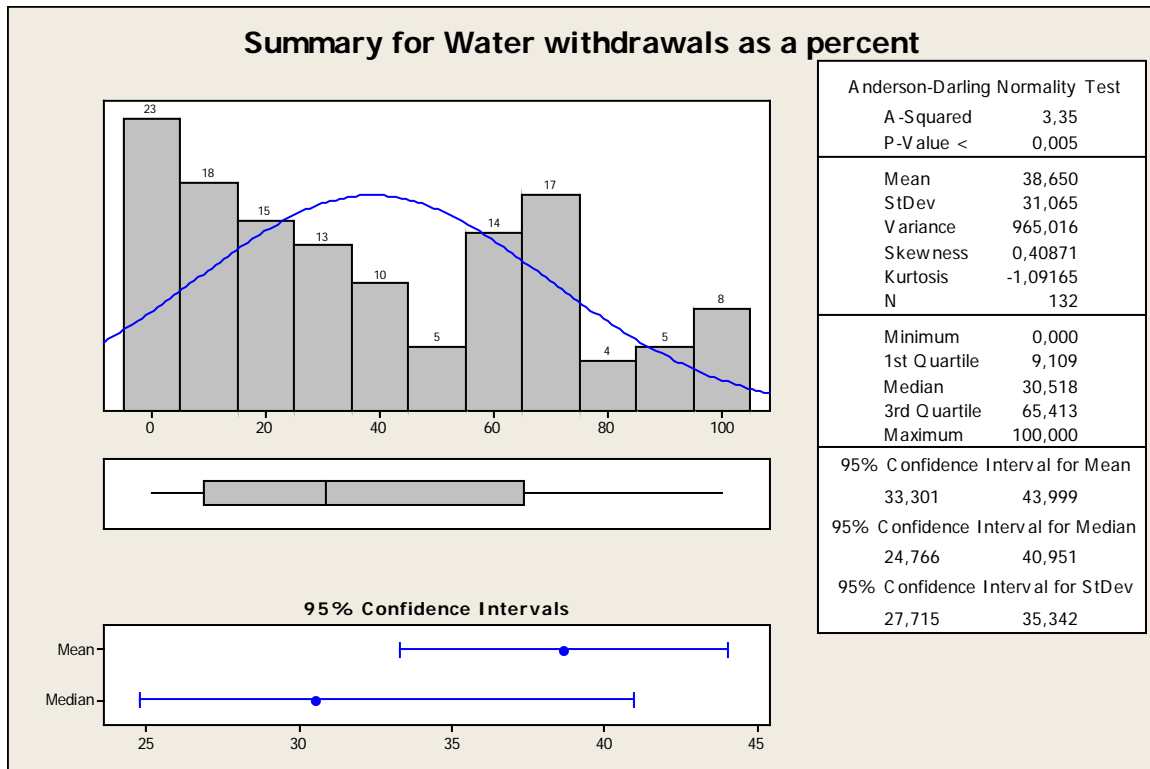
### 3.2.5 Variável: “Emissão de Gases Estufa (CO2 relativo ao PIB)”



Há uma concentração expressiva à esquerda, com pouca variação entre os países. Olhando mais no detalhe, primeiro há 39 países com registro menor que 1, com emissões próximas da marca de 2000 nos quais se encontram Suécia, Noruega, França e Reino Unido; logo, vem 32 países com registro entre 1 e 2, portanto com emissões de mais de 2000, com países tão heterogêneos quanto Brasil, Estados Unidos, Malásia, Nova Zelândia e Jamaica; em seguida, aparecem 25 países com registro entre 2 e 3, com emissões entre 200 e 1000, incluindo Austrália, Nigéria, Venezuela e Uruguai; depois vem 10 países com registro entre 3 e 4, com indicador de emissões CO<sub>2</sub> de até 200, incluindo China, Rússia, Ucrânia, Paraguai e África do Sul.

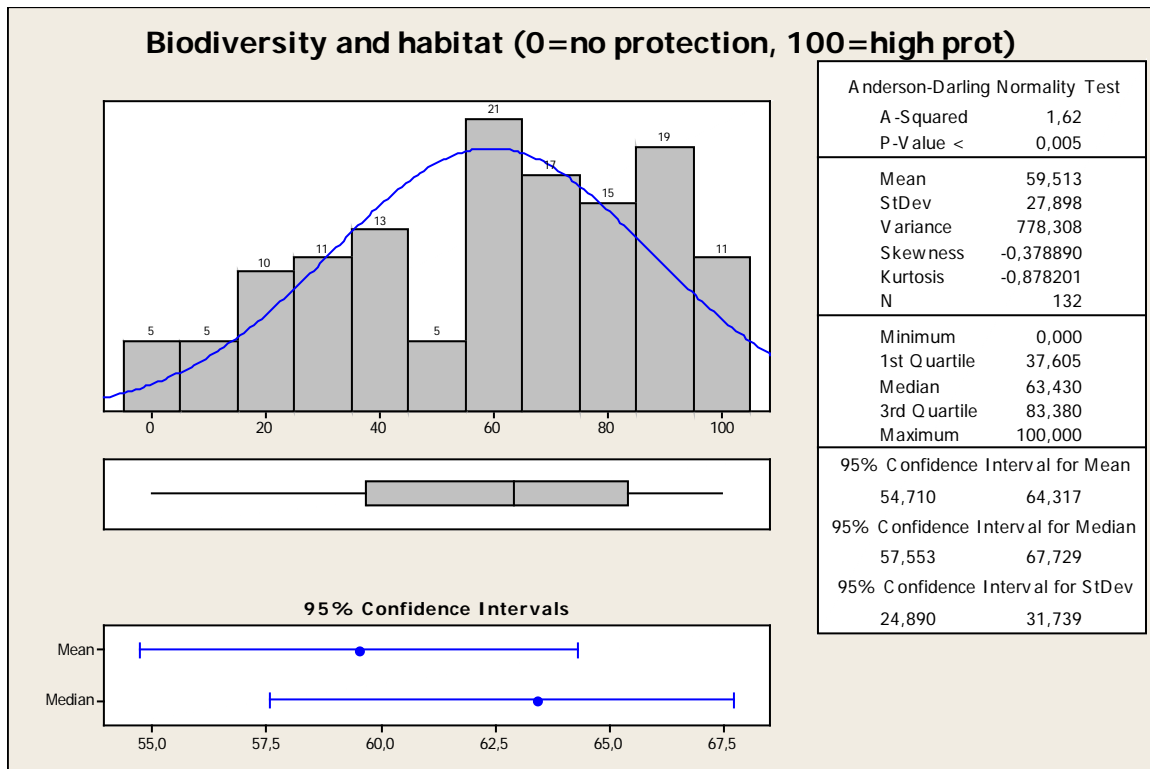
Penso que poderíamos rever a escala, ou talvez redefinir como utilizar essa variável, já que ela mostra a emissão de gases dividida pelo PIB, e isso possivelmente gere uma distorção, não permitindo que se veja o problema de forma isolada, sem vieses.

### 3.2.6 Variável: Retirada de Água (sobre % de recursos disponíveis)



O gráfico apresenta certa assimetria, com aproximadamente metade dos países mostrando índices de até 40%. Os primeiros estão localizados na África Central, região que tem uma grande bacia hidrográfica, e o Paraguai, em cujo território está disponível grande parte da água que gera a energia de Itaipu. Outro bloco importante, de 31 países, tem índices entre 55 e 75% e inclui países de vários continentes, incluindo México, Índia, Itália, Cuba, e Indonésia. O grupo do extremo direito do gráfico representa os mais problemáticos do ponto de vista da água disponível, e engloba principalmente países do Oriente Médio, como Arábia Saudita, Emirados Árabes e Kuwait, bem como ilhas das Américas como Jamaica e Trinidad e Tobago.

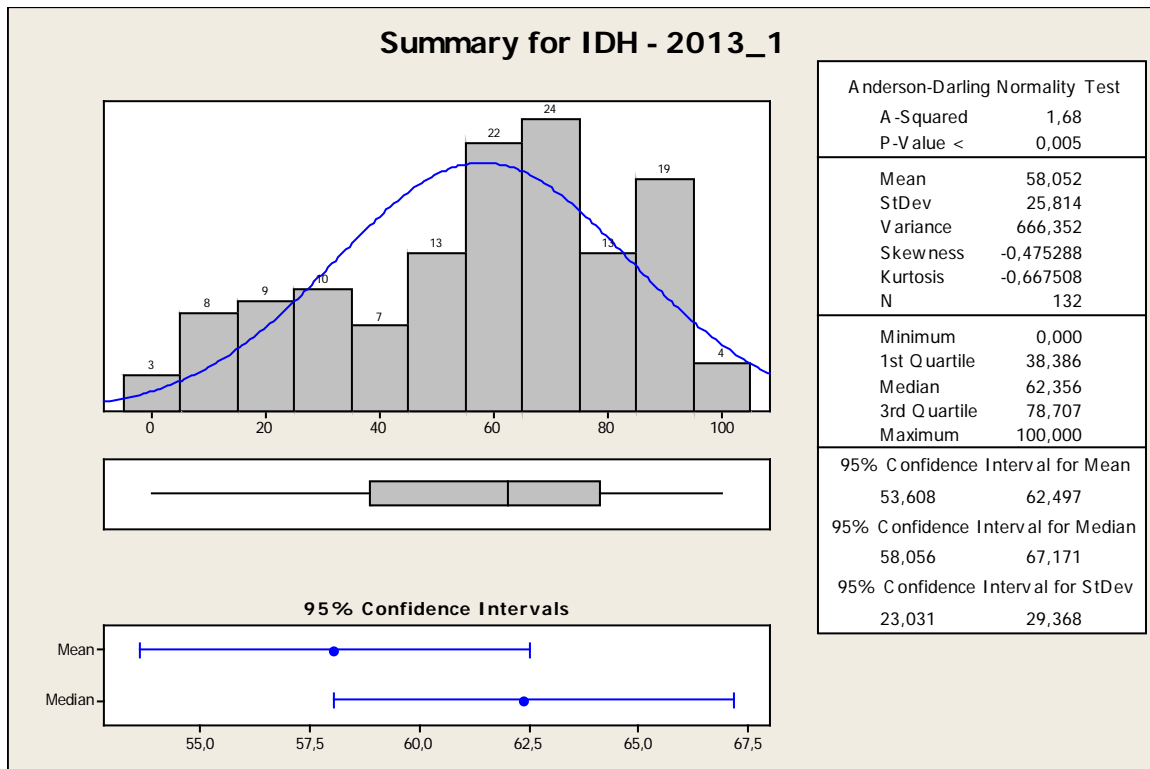
### 3.2.7 Variável: Biodiversidade e Habitat



O gráfico apresenta uma tendência à direita, reflexo do esforço dos países em procurar manter sua biodiversidade. Curiosamente, encontram-se nesse grupo mais preservacionista países tão distintos quanto Suíça e Alemanha, altamente desenvolvidos e que detêm uma política incisiva de preservação do meio ambiente, e Zâmbia e Botsuana, países desfavorecidos economicamente porem detentores “naturais” de um grande patrimônio ecológico. O Brasil, a despeito de sua fama de “santuário ecológico” e “pulmão do mundo”, ocupa uma posição intermediária, no grupo de 21 países na faixa de 55-65%, ocupando a modesta 57ª. posição entre 132.

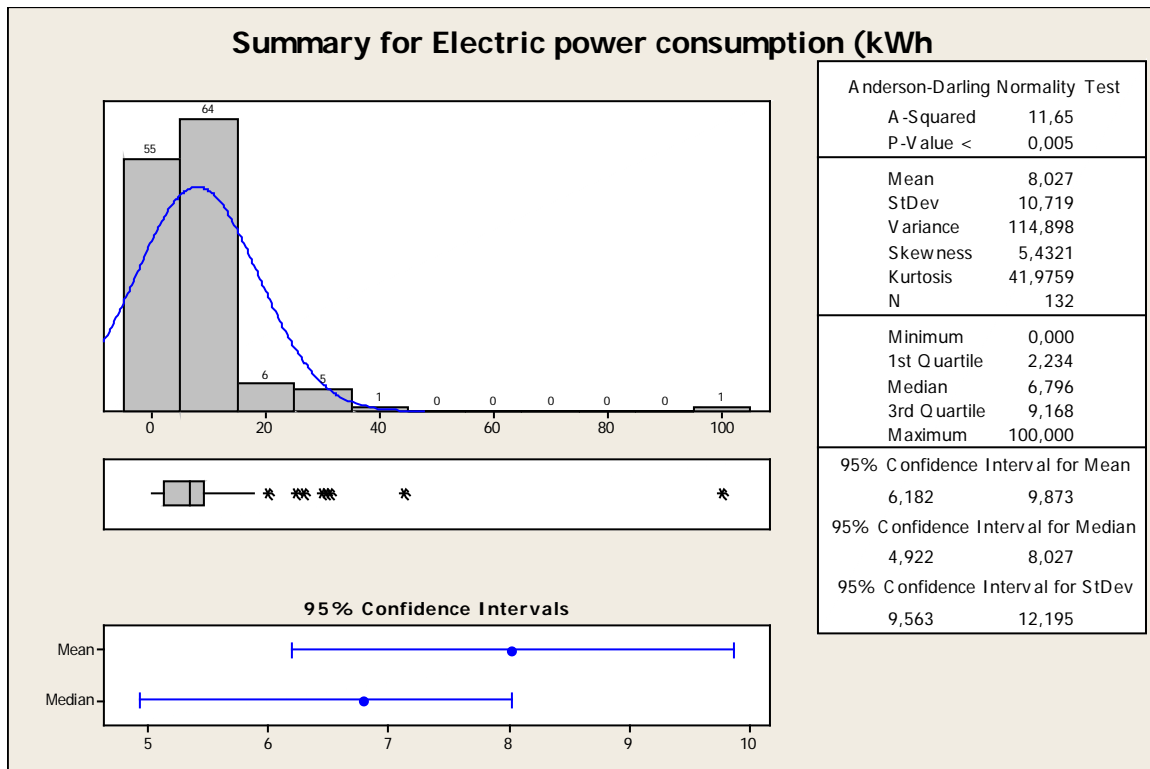
### 3.2.8 Variável: IDH 2013





A análise do gráfico mostra predominância de países em posição intermediária com tendência à direita, apresentando IDH acima de 50 (82 países), inclusive o Brasil. No extremo com índice 100 ou próximo aparecem países como Noruega, Austrália e Suécia. No outro extremo aparecem Congo, Níger e República Centro Africana com índice menor que 10. Revela-se grande disparidade entre os países da Europa e da África principalmente. Embora muitos destes países disponham de vastos recursos naturais, eles não se convertem em benefício da maioria da população, ao contrário do que ocorre na Europa.

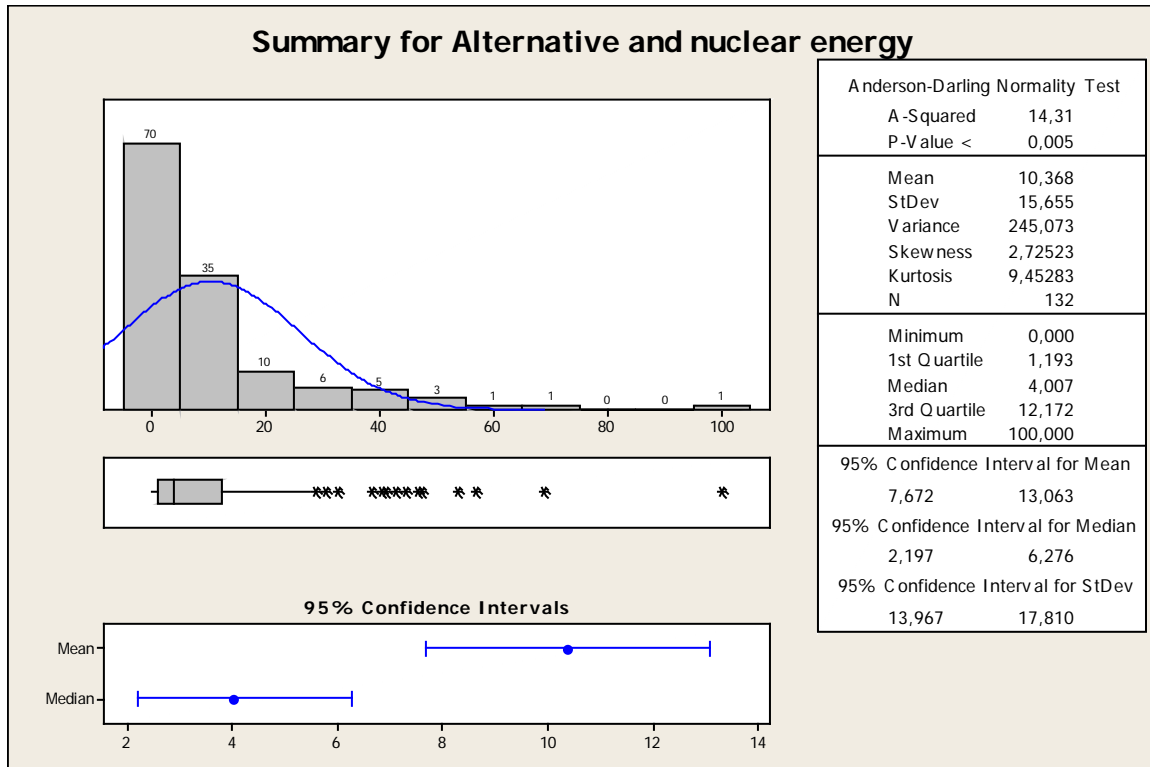
### 3.2.9 Variável: Consumo de Energia Elétrica (kWh per capita)



O gráfico mostra uma grade concentração a esquerda, dois grandes picos – o maior, com 64 países, indica um consumo de 2,500 a 7,500 kWh o segundo, de 55 países, indica consumo de até 2,500 kWh.

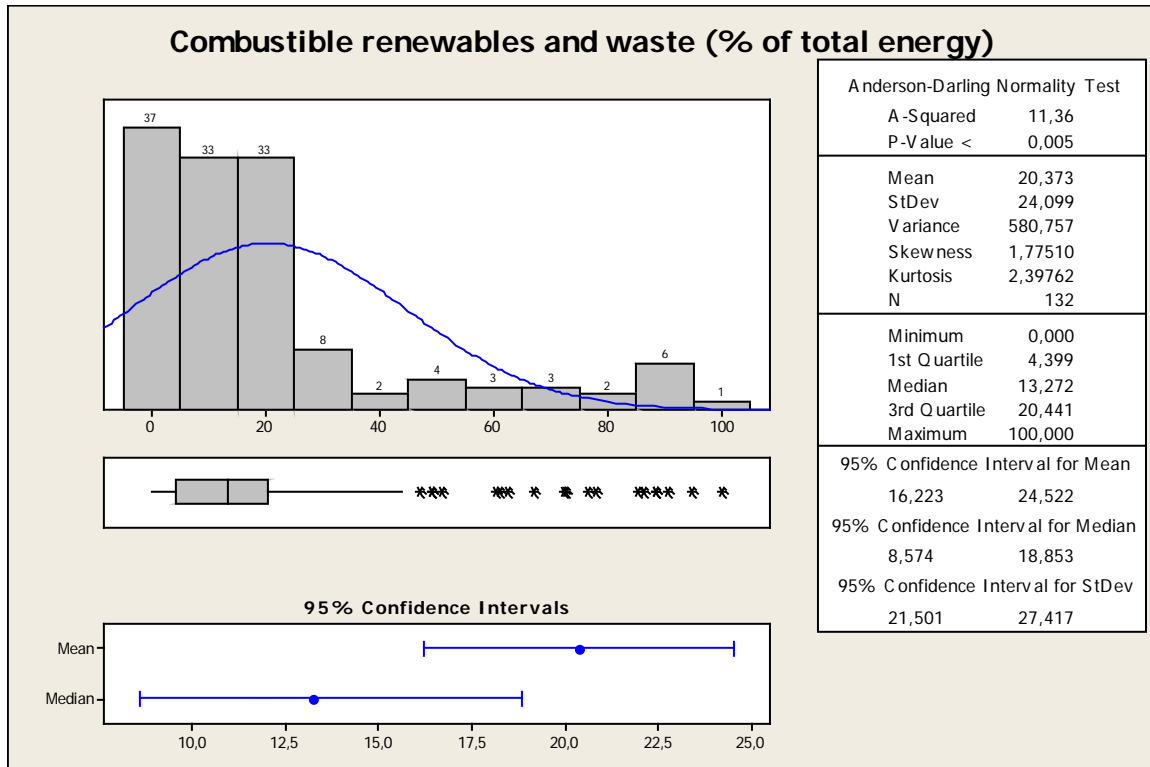
O grupo de outliers, que são os maiores usuários, inclui Islândia (o maior), Noruega, Canadá, Kuwait, Finlândia, Suécia e Estados Unidos. Interessante notar que o grupo possui tanto países quentes, do Oriente Médio como outros de clima baixíssimo, caso de Canadá, Islândia e Noruega. No outro extremo aparecem Tanzânia, Congo, Nepal, Sudão, Camboja, Nigéria, Quênia, Senegal e Iêmen, que tem consumo baixo pela não disponibilidade de energia elétrica para a maioria da população.

### 3.2.10 Variável: Energia Nuclear e Alternativa



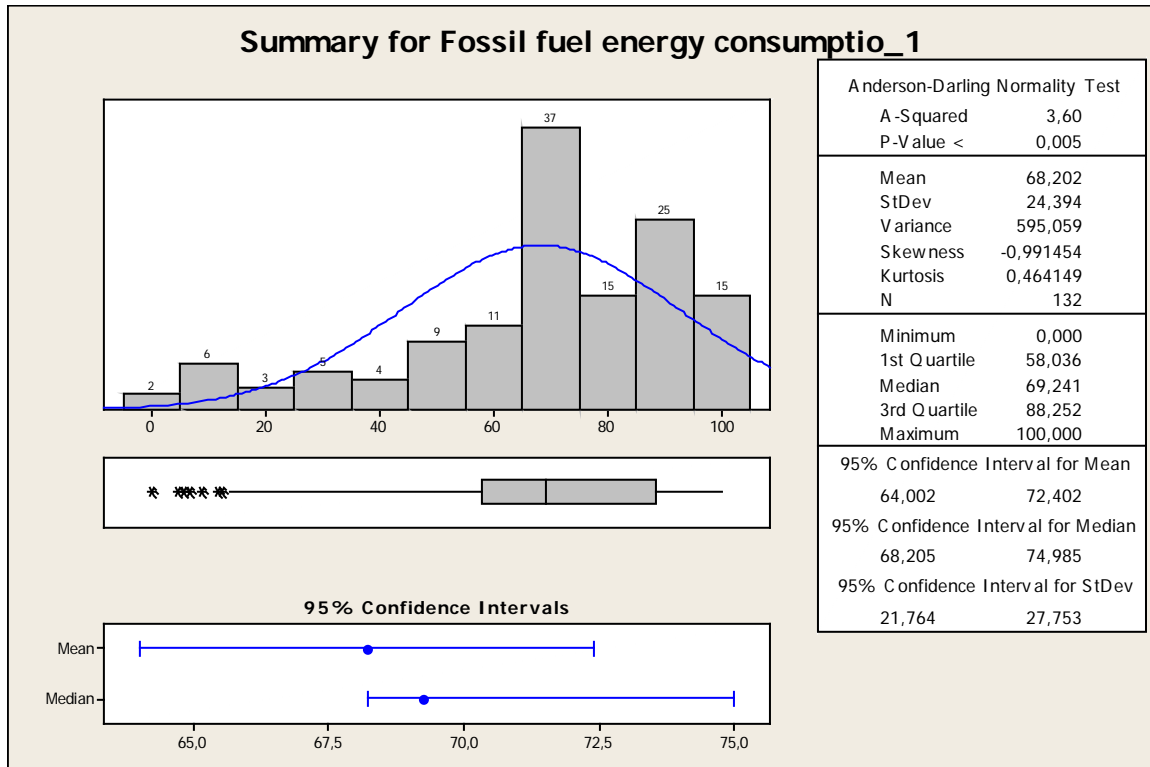
O gráfico mostra uma grande concentração à esquerda, revelando que pouco ainda se realiza em termos de energias alternativas. A exceção é justamente a Islândia, país gelado que curiosamente tem quase toda a energia que consome gerada por fonte geotérmica. Outros países que chamam a atenção no grupo dos outliers são a República Centro Africana e o Tajiquistão que figuram ao lado de França e Suíça. Na Europa ainda é alto o consumo de carvão como fonte de energia. O Brasil depende quase 100% das hidrelétricas para geração de energia, com algo de nuclear e algo de termelétrica, razão pela qual ocupa a 24ª posição entre os 132 países.

### 3.2.11 Variável: Combustíveis Renováveis e Gastos



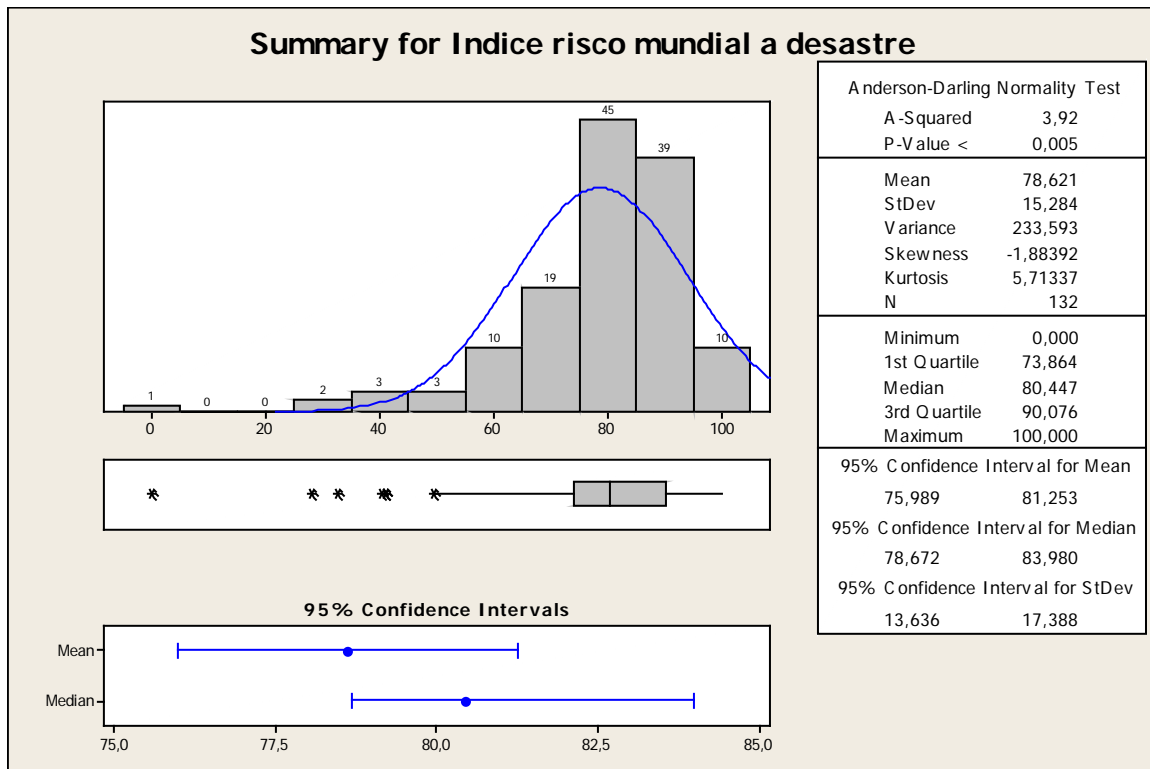
Analisando o gráfico, vemos que há uma grande concentração do lado esquerdo, com a maior parte das ocorrências (103 países) na faixa de 8 a 20%, outros 10 países na faixa até 50%, e os demais 19 no grupo dos outliers, dentre os quais se destacam Congo, Tanzânia, Nepal, Nigéria, Togo e Moçambique. No outro extremo encontram-se Kuwait, Tajiquistão, Arábia Saudita, e Uzbequistão, países com grandes reservas de petróleo, razão pela qual apresentam ocorrências relativamente baixas.

### 3.2.12 Variável: Consumo de Energia de Combustíveis Fósseis



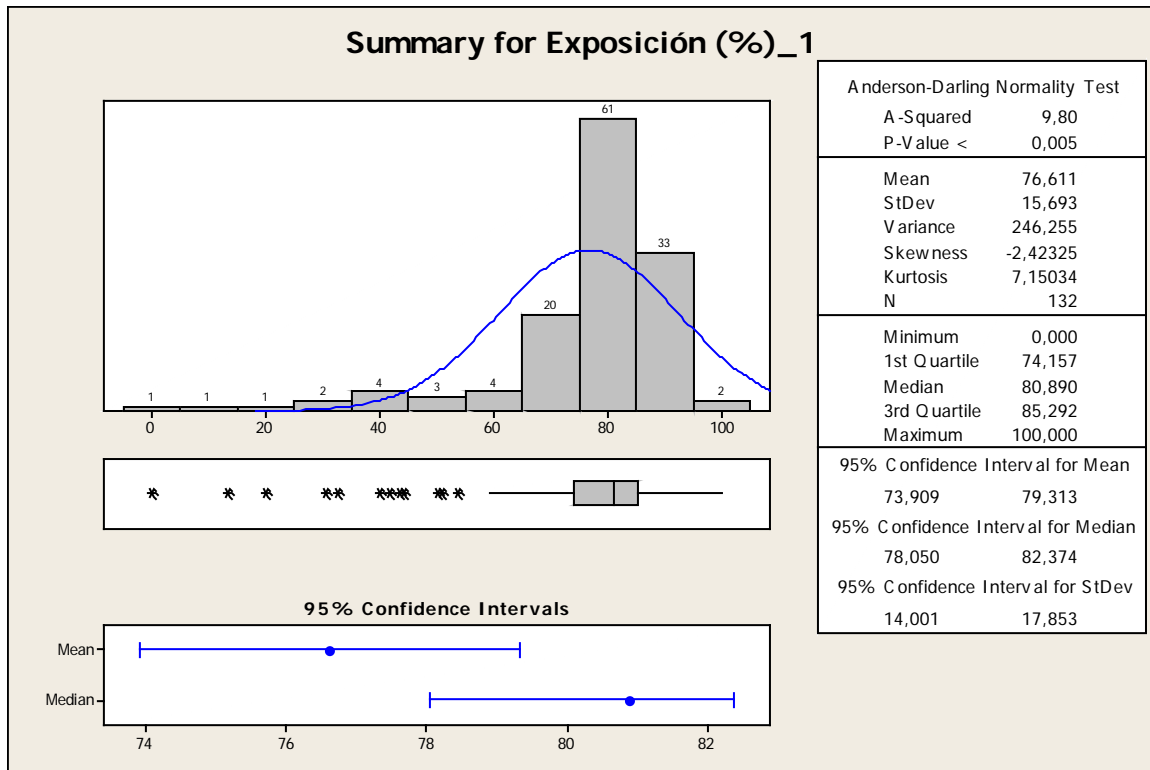
O gráfico mostra um pico de 37 países que, do total da energia que consomem, 65 a 75% são provenientes de combustíveis fósseis. No geral, há uma concentração desse lado do gráfico, revelando uma grade dependência dessa fonte de energia. Os mais dependentes são Arábia Saudita, Trinidad e Tobago, Argélia, Irã e Cazaquistão. As exceções, que figuram como outliers no outro extremo com índice de consumo menor que 12% do total, são Congo, Zâmbia, Moçambique, Tanzânia, Nepal, Togo e Islândia.

### 3.2.13 Variável: Índice Mundial de Risco a Desastres Naturais



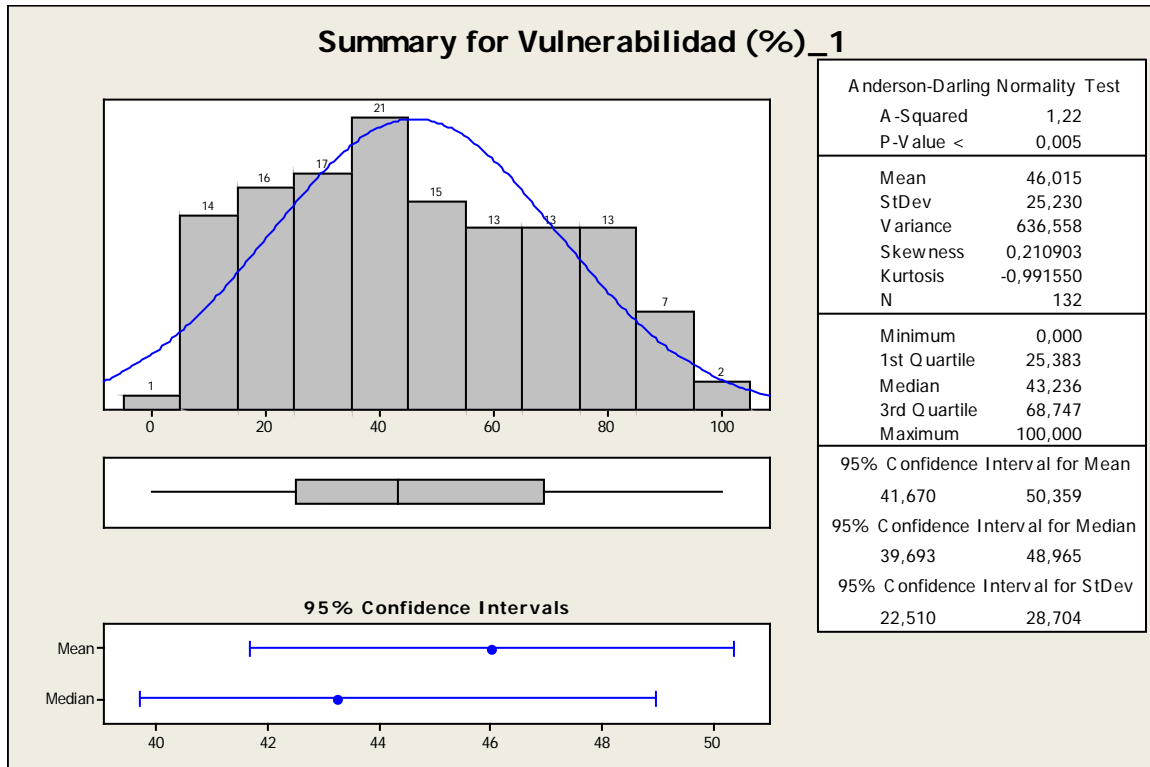
O gráfico apresenta uma concentração à direita (mediana com índice entre 78 e 84). Outliers, com alto índice de risco, são Guatemala, Filipinas, Bangladesh, Camboja, Costa Rica, El Salvador e Nicarágua, países propensos que não estão preparados para enfrentar catástrofes naturais. No outro extremo aparecem Arábia Saudita, Islândia, Emirados Árabes, Suécia e Finlândia, com índice entre 95 e 100, resultado provável de fatores como ações preventivas eventualmente combinadas com um favorecimento geográfico.

### 3.2.14 Variável: Exposição



O gráfico tem configuração parecida com o da variável Índice de Risco a Desastres Naturais, apresentando uma concentração expressiva do lado direito, com a mediana entre 78 e 82. A maior parte dos países (61) se encontra nesse pico, com índice entre 75 e 85. Destaque para a Arábia Saudita, com baixo índice de exposição a desastres naturais e os outliers Costa Rica, Japão e Filipinas, que aparecem com altíssimo índice relativo de exposição, provavelmente devido ao fato de que sejam ilhas ou tenham áreas relativamente reduzidas circundadas pelo oceano, sendo assim naturalmente mais sujeitas às intempéries climáticas que países que fazem parte de massas continentais mais amplas.

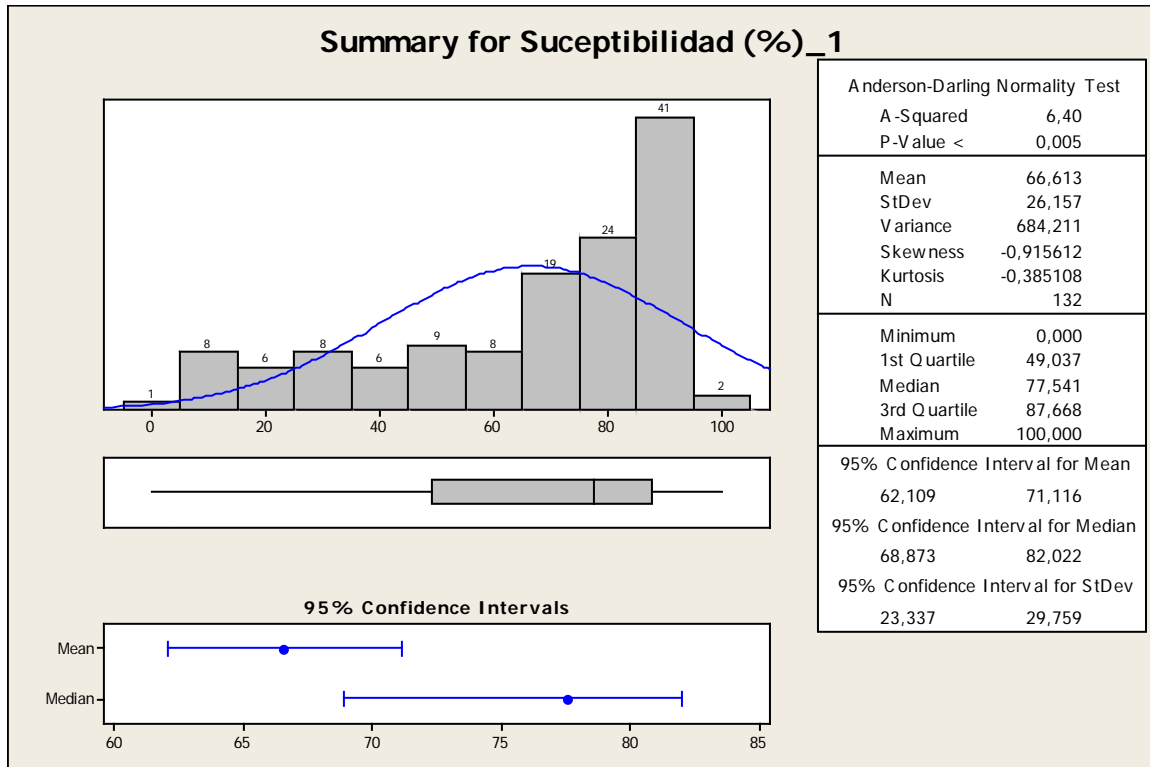
### 3.2.15 Variável: Vulnerabilidade



O gráfico mostra uma dispersão relativamente alta dos países. A vulnerabilidade parece estar associada ao grau de desenvolvimento, pois quanto maior o IDH, menor a vulnerabilidade. Sendo assim, como era de se esperar, os países mais pobres da África, como Nigéria, Chade, Libéria, República Centro-Africana, Guiné e Moçambique tem maior índice de vulnerabilidade, entre 88 e 100%, ao passo que os países mais economicamente estáveis na Europa e Oceania, como Nova Zelândia, Austrália, Suíça, Áustria, Alemanha e Noruega apresentam baixo risco, entre 0 e 8%.

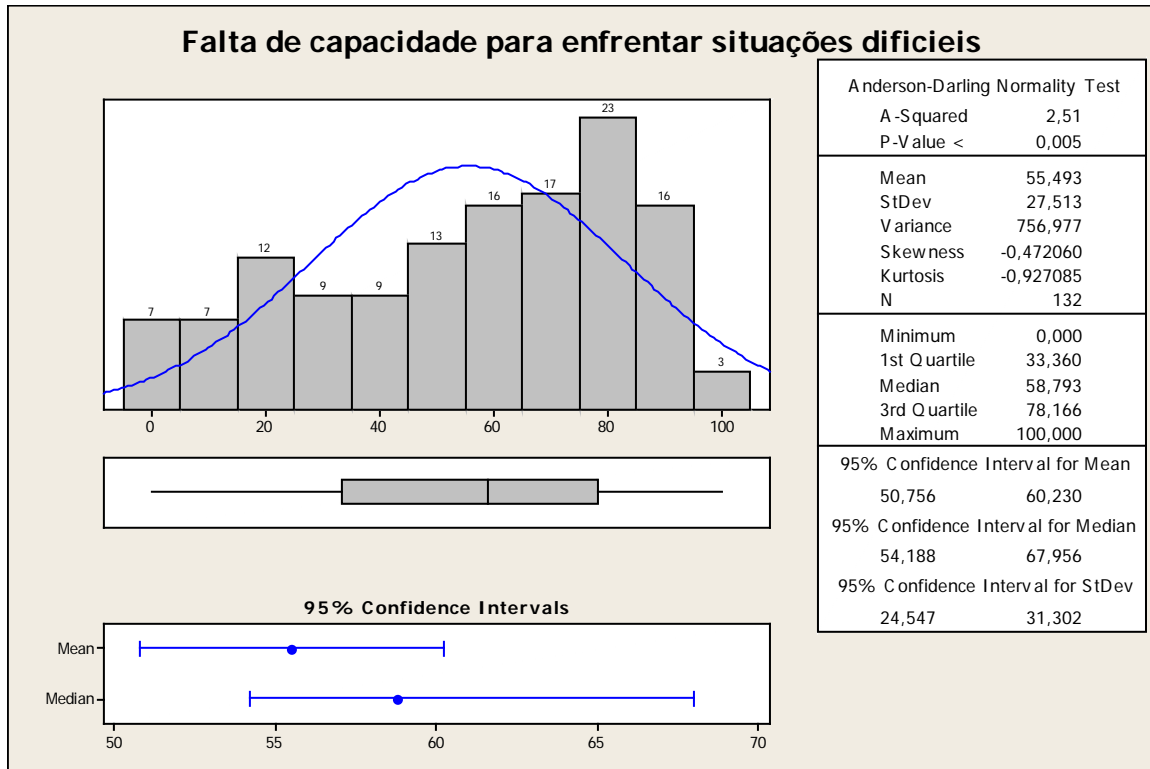


### 3.2.16 Variável: Suscetibilidade



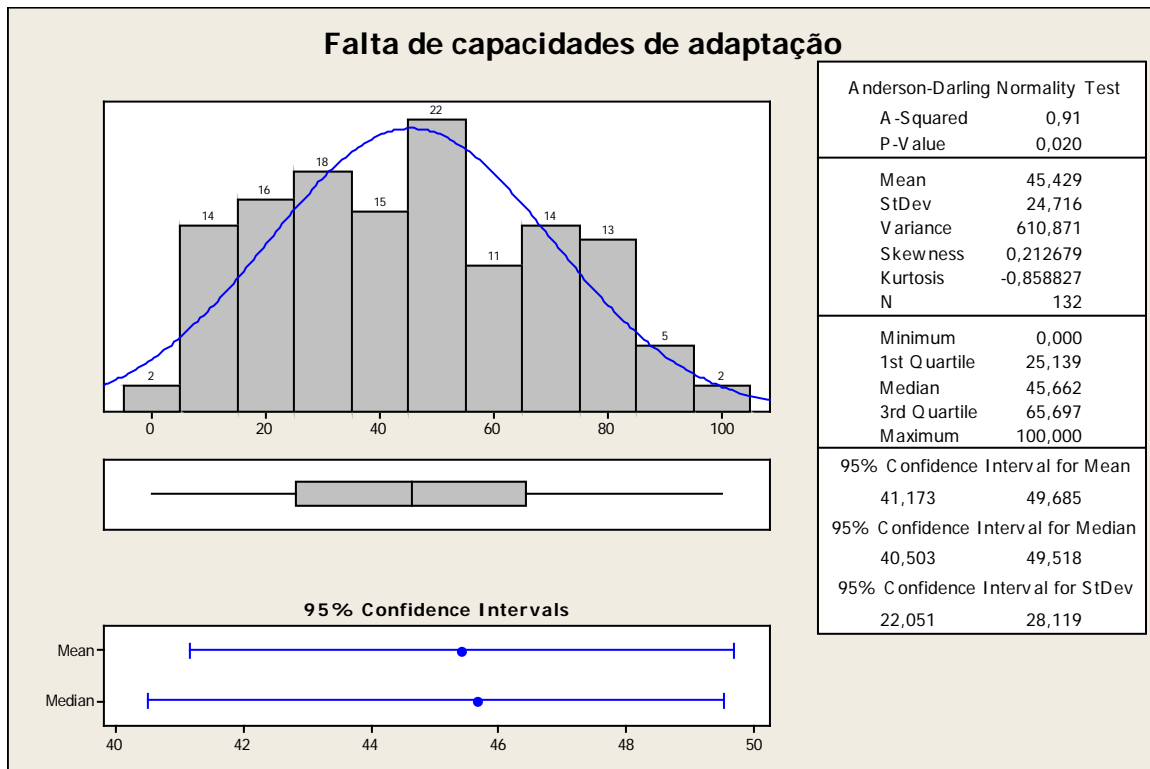
O gráfico apresenta um pico expressivo à direita, demonstrando que aproximadamente um terço dos países da base se encontra em situação relativa de baixa suscetibilidade, com um índice que varia entre 85 e 95. Os países mais à esquerda são justamente os que possuem histórico de pobreza e/ou vivem certa instabilidade civil, política ou econômica caso dos países africanos como Libéria, Moçambique, Chade e Burundi.

### 3.2.17 Variável: Incapacidade para Enfrentar Situações Difíceis



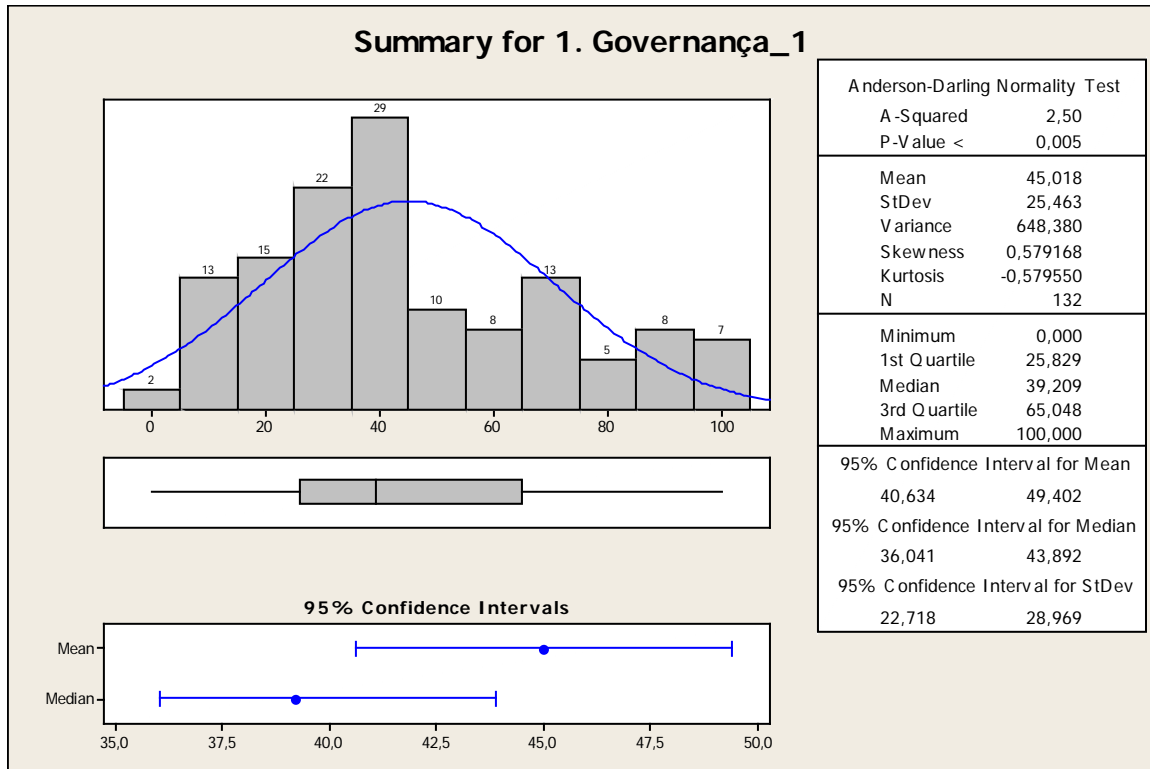
O gráfico mostra uma distribuição relativamente regular, porém com uma leve tendência para direita, ou seja, um número maior de países com falta de capacidade de enfrentar situações difíceis. Nesse grupo figuram Sudão com índice 100, seguido por Chade e Iêmen, com 97. Logo vem outros países do continente africano como Nigéria, Uganda, Burundi e Rep Centro Africana. Do outro lado aparecem países de alto grau de desenvolvimento e portanto menor dificuldade para enfrentar as intempéries, como Áustria, Alemanha e Suíça, com índice de 0 a 1, seguidos de Japão, Finlândia, Dinamarca e Noruega.

### 3.2.18 Variável: Capacidade de Adaptação



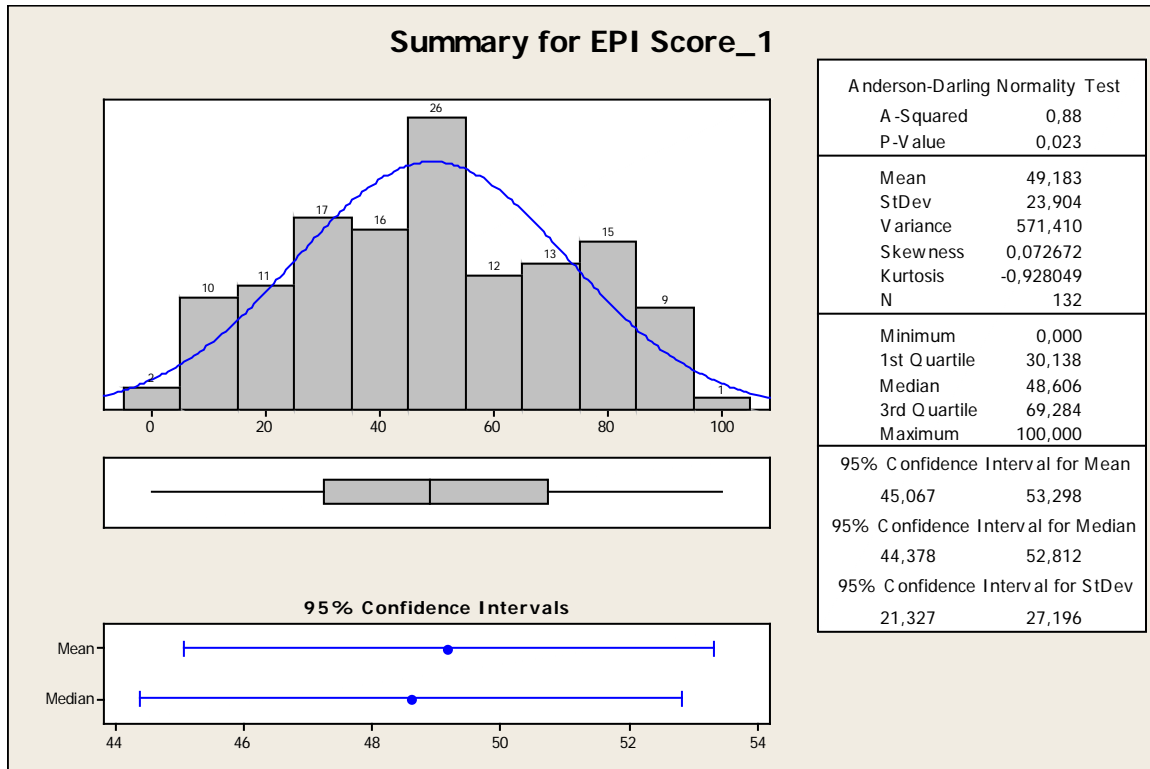
O gráfico apresenta dispersão bastante ampla, sem predominância expressiva. O maior grupo, de 22 países, tem um índice na faixa de 45 a 55, dentro da mediana. No lado direito, destaque para Austrália, Nova Zelândia, Holanda, Islândia e Suíça. No outro extremo aparecem países do continente africano como Chade, Libéria, República Centro-Africana. Esse resultado reforça outras variáveis como exposição, suscetibilidade, e vulnerabilidade, mostrando comportamento análogo entre os mesmos países, portanto pode-se supor que tais variáveis sejam correlacionadas.

### 3.2.19 Variável: Governança



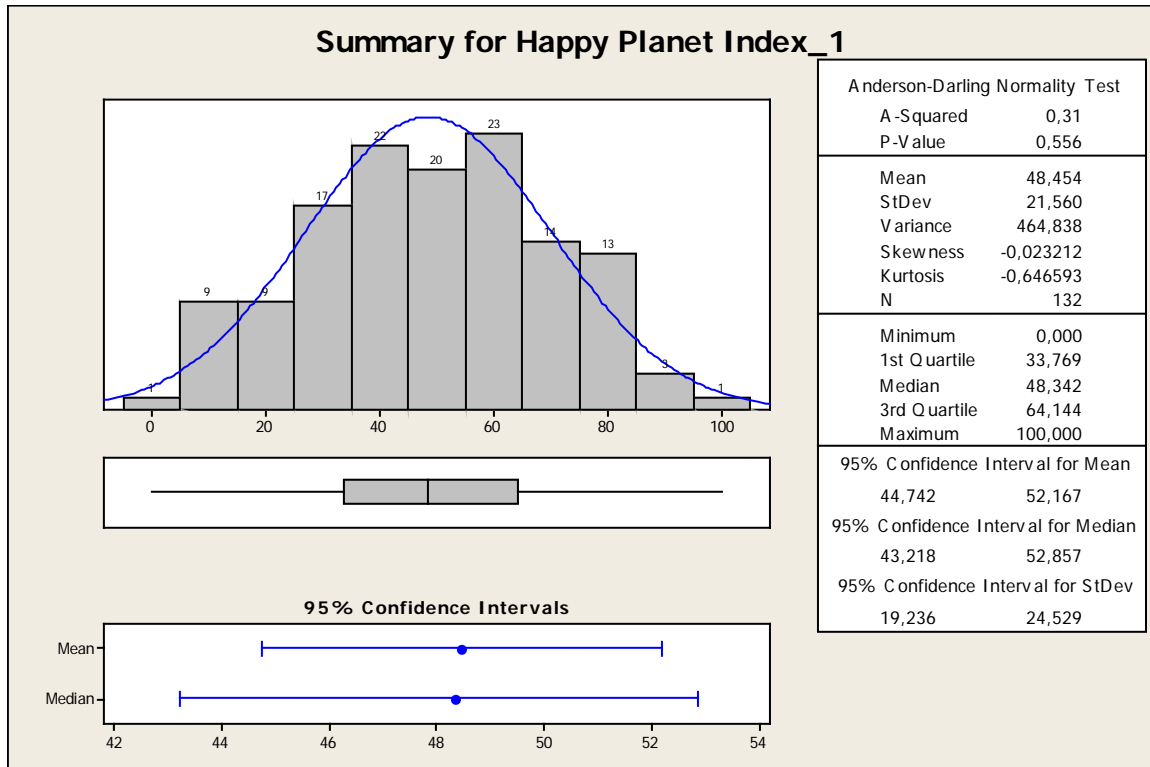
O gráfico apresenta uma distribuição assimétrica, mostrando que no geral há um índice relativamente baixo de governança entre os países – mediana entre 36 e 43 e um pico de 29 países com índice entre 35 e 45. Os destaques positivos estão quase todos na região nórdica da Europa – Finlândia, Suécia, Noruega, Dinamarca, com índices acima de 95, ao passo que o grupo à esquerda inclui Sudão, República Centro-Africana, Iraque, Venezuela e Iêmen, todos com índice menor que 15. Isso reforça os resultados apurados em outras variáveis, mostrando que países com índices positivos também vem acompanhados de alto destaque do ponto de vista de governança.

### 3.2.20 Variável: EPI Score



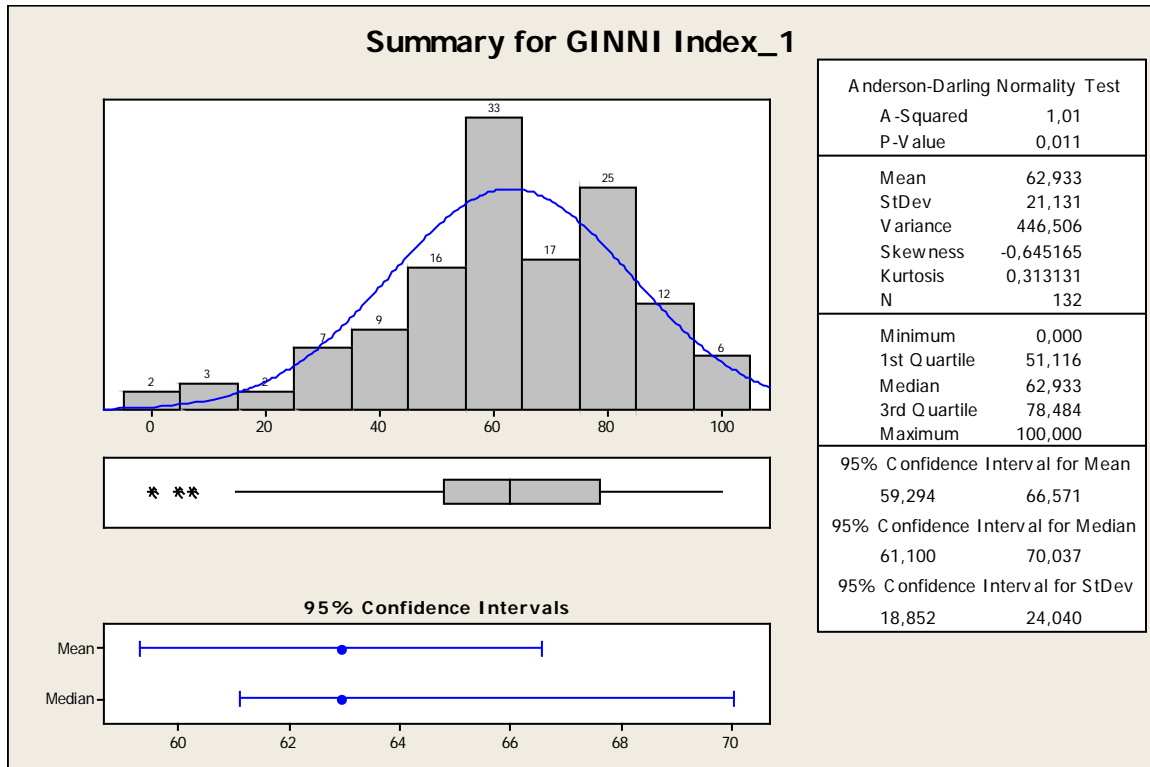
O gráfico apresenta dispersão relativamente alta dos países no índice EPI, com o maior grupo numa posição intermediária, entre 45 e 55. Os destaques positivos ficam por conta de países desenvolvidos como Austrália, Suíça, Alemanha e Rep. Tcheca, com índices entre 90 e 100. No outro extremo figuram países com baixo grau de desenvolvimento, como Mali, Lesoto, Sudão e Libéria, com índices de 0 a 10.

### 3.2.21 Variável: Happy Planet Index



O gráfico apresenta distribuição normal. Destaque para países que geralmente apresentam índices negativos ou pouco expressivos em indicadores socioeconômicos e de infraestrutura, e nesse quesito tem avaliação positiva, tais como Jamaica, Colômbia, El Salvador, Guatemala, Honduras, Nicarágua, Panamá. No entanto, ainda os países africanos apresentam os índices mais baixos, como Botsuana, Chade e República Centro-Africana, o que, se não desfaz totalmente, ao menos relativiza um certo mito que às vezes se divulga de maneira equivocada: de que pessoas pobres são mais felizes que pessoas ricas.

### 3.2.2 Variável: GINNI Index



Analisando o gráfico, nota-se uma certa assimetria, com um pico mostrando 33 países com índice entre 55 e 65. Chama a atenção no extremo direito do gráfico aparecerem países que não figuravam positivamente até então, como Ucrânia, Eslovênia e República Checa, com índice perto de entre 95 e 100. No outro extremo, aparecem países como África do Sul, que a despeito de seu relativo desenvolvimento, apresenta oportunidades de maneira desigual à sua população e junto com Namíbia e Botsuana, compõe o grupo dos outliers. O Brasil ocupa uma posição relativamente baixa, com índice 27.

#### 4. TABELA COMPARATIVA DA ANÁLISE DESCRITIVA

A seguir apresentamos tabelas comparativas de algumas variáveis abrangentes, indicando:

- as primeiras cinco posições gerais;
- as melhores posições dos países ibero-americanos;
- a classificação específica do Brasil;
- a posição mais baixa de um país ibero-americano; e
- as últimas cinco posições gerais.

País	SPI	País	IDH - 2013	País	Risk Desas	País	Governan	País	GINNI
1 New Zealand	100,0000	1 Norway	100,0000	1 Saudi Arabia	100,0000	1 Finland	100,0000	1 Ukraine	100,0000
2 Switzerland	99,9101	2 Australia	98,1878	2 Iceland	98,5598	2 Norway	98,9114	2 Slovenia	99,8695
3 Iceland	99,6945	3 Sweden	95,5519	3 Unit Arab Emir	97,2674	3 Sweden	98,8206	3 Sweden	96,7119
4 Netherlands	98,4364	4 United States	95,0577	4 Sweden	96,2334	4 New Zealand	98,7169	4 Czech Rep	95,9029
5 Norway	97,9871	5 Germany	94,5634	5 Finland	96,0487	5 Denmark	98,2836	5 Belarus	95,7203
21 Spain	86,57441	21 Spain	87,6442	19 Spain	92,5037	18 Chile	80,6744	49 Spain	71,4770
22 Portugal	86,07117	30 Chile	79,9012	28 Portugal	90,9897	21 Portugal	74,0262	62 Portugal	64,4311
25 Costa Rica	81,14666	31 Portugal	79,9012	30 Argentina	90,7312	25 Spain	70,1330	90 Uruguay	56,9415
26 Uruguay	80,71531	36 Argentina	77,5947	31 Paraguay	90,5096	28 Uruguay	68,4302	92 El Salvado	55,6889
30 Chile	78,54062	37 Uruguay	74,6293	36 Uruguay	89,5495	36 Costa Rica	64,5690	101 Argentina	51,0699
46 Brazil	67,16391	56 Brazil	67,05107	40 Brazil	88,4417	51 Brazil	45,42651	124 Brazil	27,3225
77 Honduras	51,54565	95 Nicaragua	45,6343	131 El Salvador	41,1004	97 Honduras	26,1149	125 Colombia	25,0783
128 Sudan	10,5140	128 Burkina Faso	8,4020	128 Cambodia	41,1004	128 Yemen	8,2782	128 Honduras	14,9791
129 Guínea	8,6449	129 Chad	5,7661	129 Costa Rica	40,3250	129 Venezuela	8,0062	129 Zambia	14,7443
130 Burundi	8,5011	130 Centr Africa	0,6590	130 Bangladesh	32,7917	130 Iraq	7,3225	130 Botswana	6,9937
131 Centr Africa	2,8217	131 Congo	0,1647	131 Guatemala	27,9542	131 Centr Africa	0,9088	131 Namibia	4,7495
132 Chad	0,0000	132 Niger	0,0000	132 Philippines	0,0000	132 Sudan	0,0000	132 South Afri	0,0000

De maneira geral, as posições melhores encontram-se nos países desenvolvidos da Europa e Oceania, sobretudo em variáveis como SPI, IDH e Governança. Em contrapartida, as posições mais baixas dessas mesmas variáveis estão nos países africanos, exceções feitas a Venezuela e Iraque na variável governança, acredito que relacionada às políticas de tais países ligadas ao petróleo.

No grupo de países ibero-americanos, destacam-se Portugal e Espanha, que se beneficiaram da inserção como membros da União Europeia, tendo que cumprir com uma agenda de desenvolvimento com forte relação com alguns desses índices. Em seguida, aparecem Uruguai em destaque, seguido por Chile, Argentina e Costa Rica. O Brasil no geral encontra-se mais abaixo, em posições intermediárias.

Uma diferença importante aparece na variável GINNI, com países do leste europeu surgindo entre os primeiros da lista positiva, algo inesperado. Esse país mais igualitário seria um efeito tardio da estrutura social do sistema capitalista? Seriam países de estrutura menos complexa e por isso mais controlável? Algo que poderia ser investigado para verificar o que fazem diferente dos demais para ganharem esse destaque.

Outro ponto que vale ressaltar nesta variável é a presença de vários países latino-americanos nas últimas posições, ou muito próximo delas, revelando a desigualdade abissal que ainda assola a região. Acredito que teremos oportunidade de nos debruçar mais atentamente sobre essa questão quanto nos detivermos mais profundamente sobre a análise desses países.



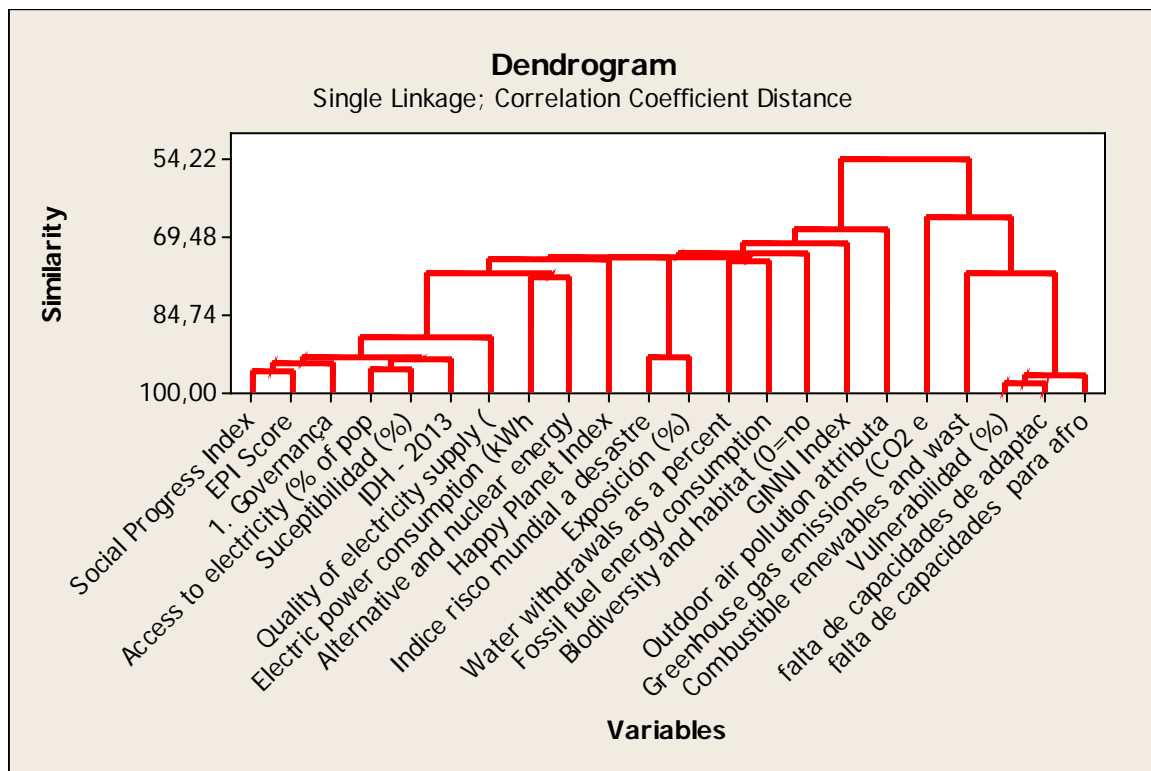
## 5. RELAÇÃO ENTRE AS VARIÁVEIS QUANTITATIVAS

### 5.1 Dendrograma

Um **Dendrograma** (dendr(o) = árvore) é um tipo específico de diagrama ou representação icônica que organiza determinados fatores e variáveis. É um diagrama de similaridade.

A interpretação de um dendrograma de similaridade entre amostras fundamenta-se na intuição: duas amostras próximas devem ter também valores semelhantes para as variáveis medidas. Ou seja, elas devem ser próximas matematicamente no espaço multidimensional. Portanto, quanto maior a proximidade entre as medidas relativas às amostras, maior a similaridade entre elas. O dendrograma hierarquiza esta similaridade de modo que podemos ter uma visão bidimensional da similaridade ou dissimilaridade de todo o conjunto de amostras utilizado no estudo.

Segue abaixo o Dendrograma das variáveis analisadas:



O primeiro comentário é que as variáveis “vulnerabilidade” e “falta de capacidade de adaptação” possuem alta similaridade – a razão é que a última é componente da primeira. Em seguida aparecem “Social Progress Index” e “EPI Score”, ambas as variáveis são índices sintéticos, que provavelmente incluam itens de análise similares. Vale também ressaltar a similaridade entre as variáveis “Susceptibilidade” e “Acesso à Eletricidade”. A análise de correlação irá apontar em mais detalhe, além das mencionadas, quais outras variáveis apresentam alto índice de correlação.

## 5.2 Correlação Linear

Segue abaixo a matriz de correlação incluindo o teste de significância p-value. Para a correlação foi utilizado o índice de Pearson. Vale ressaltar que o índice de correlação entre as variáveis não requer que exista uma relação de causa-efeito entre ambas.

### Correlations: Social Progr; Access to el; Quality of e; Outdoor air ; ...

	Social Progress	Access to electr
IDH - 2013	0,868 0,000	0,823 0,000
Suceptibilidad	0,838 0,000	0,912 0,000
Governança	0,890 0,000	
EPI Score	0,917 0,000	

	Biodiversity and	IDH - 2013	Electric power
Suceptibilidad		0,872	
EPI Score		0,824 0,000	

	Indice risco mun	Vulnerabilidade
Exposición (%)	0,867 0,000	
falta de capacid		0,938 0,000
falta de capacid		0,965 0,000

	Suceptibilidad	falta de capacid
falta de capacid		0,897 0,000
EPI Score	0,811 0,000	

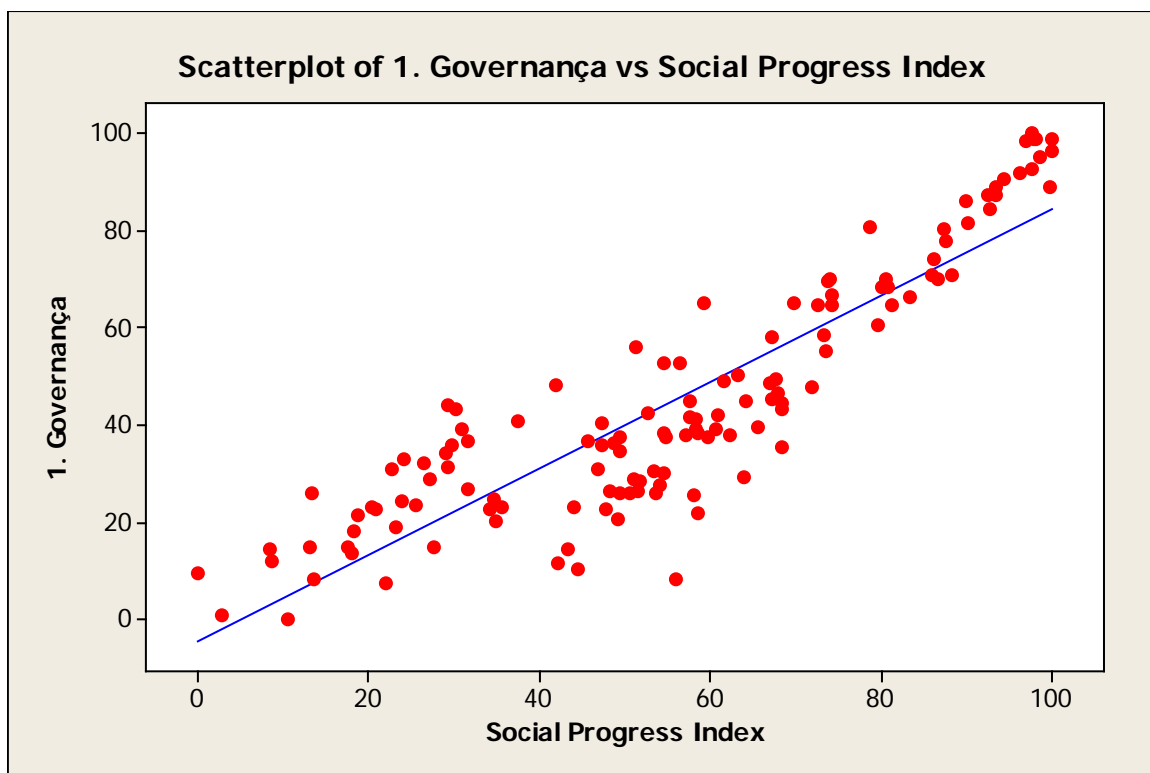
	Governança
EPI Score	0,811

0,000

Escolhemos, das variáveis acima, aquelas que apresentam a maior correlação para gerar gráficos de dispersão. Para se verificar qual o tipo de relação (linear, quadrática, cúbica, exponencial, etc.) existente entre as variáveis, adicionamos em cada gráfico de dispersão uma linha de tendência. Os gráficos sugerem que são relações lineares, pois os pontos estão relativamente próximos à linha de tendência.

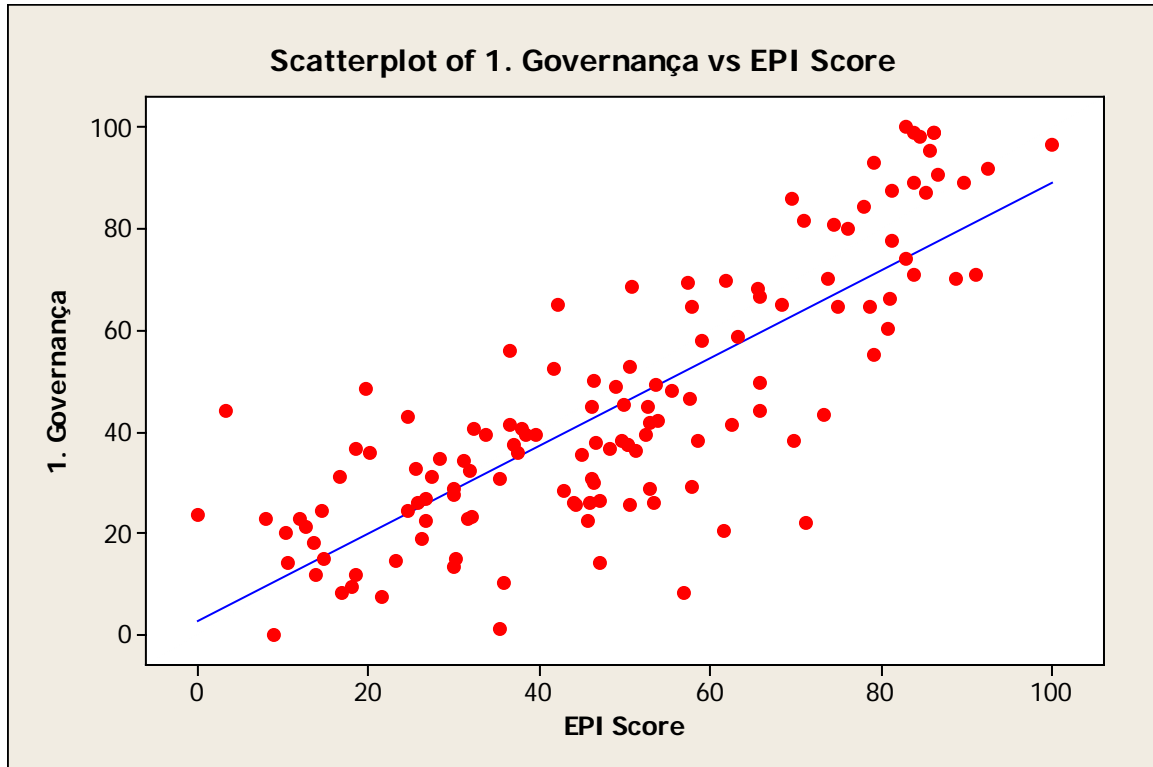
### 5.3 – Gráficos de Dispersão

#### 5.3.1 Dispersão entre as Variáveis Governança e Social Progress Index



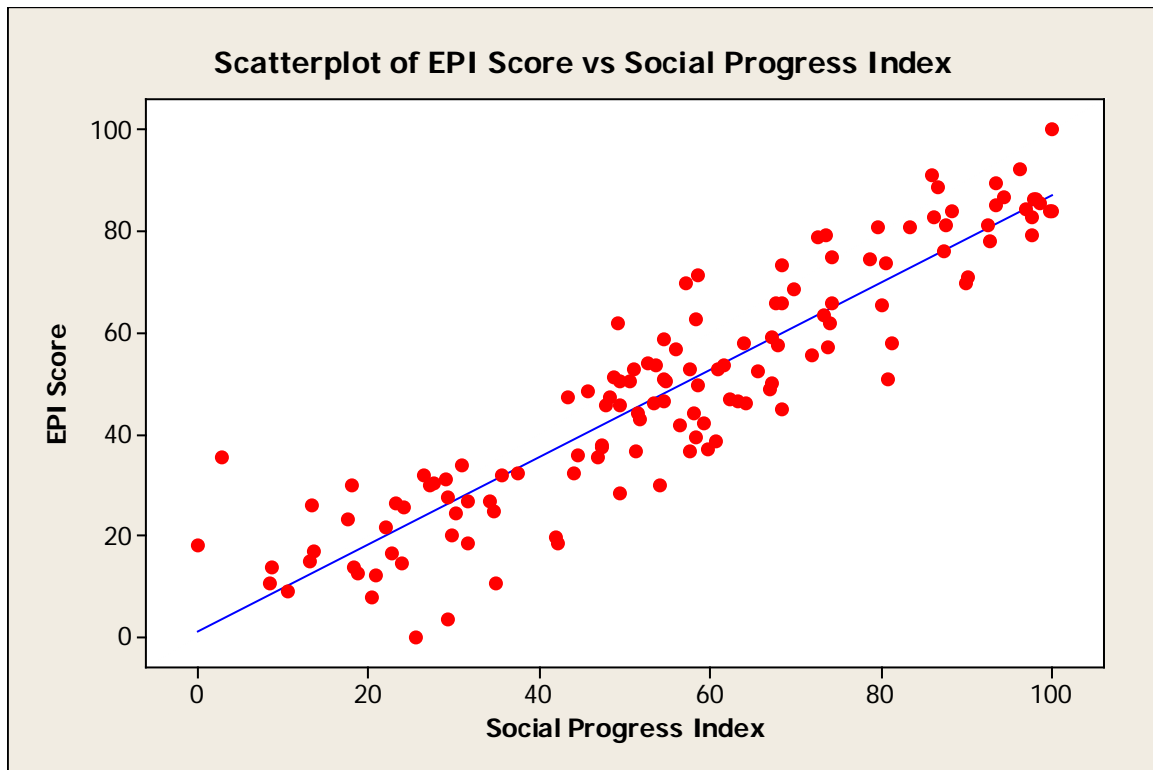
De forma análoga à relação com a variável EPI Score, também com a variável “Social Progress Index” a variável “Governança” apresenta forte relação. Nota-se claramente a concentração maior de três grupos, o primeiro com índice entre 20 e 40, outro que vai de 45 a 70 e o último que vai de 80 a 100. Uma exceção é a Venezuela, que apresenta SPI em torno de 57 e Governança 8, o que se configura coimo algo bastante raro. Acredito que tenha a ver com uma condição econômico-social relativamente boa (SPI) do país em detrimento das políticas governamentais relacionadas ao petróleo e ao uso dos recursos do país

### 5.3.2 Dispersão entre as Variáveis Social Governança e EPI Score



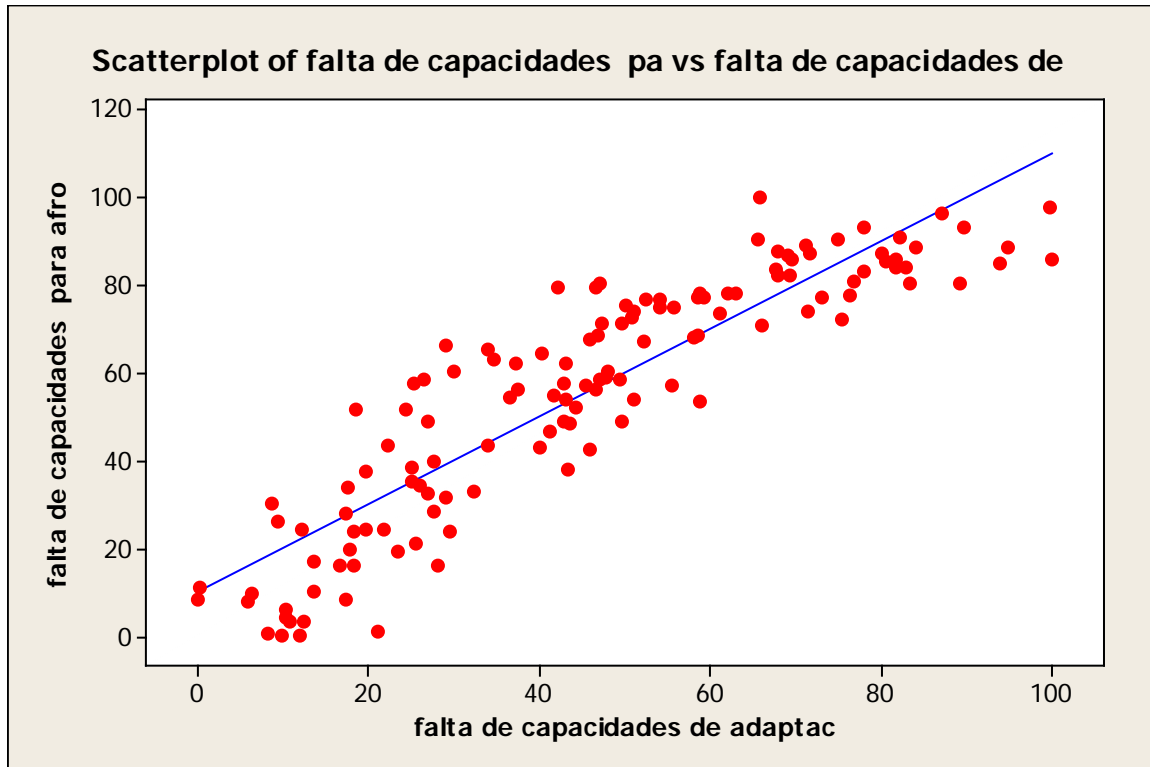
O gráfico de dispersão entre as variáveis “Governança” e “EPI Score” mostra uma relação forte entre elas, pois os países que tem altos índices em uma tem também na outra, como a Nova Zelândia, e também o oposto, como por exemplo no Chade, o que mostra, de certa forma, que as variáveis caminham juntas. Isso quer dizer que onde há boa governança provavelmente haja bom EPI Score, e onde baixo EPI Score deve-se esperar que haja governança sofrível. Entendo que não haja uma relação de subordinação entre as variáveis, no entanto me parece lógico, pela natureza das variáveis e suas definições que o EPI Score seja resultado de alguns fatores, dentre os quais governança.

### 5.3.3 Dispersão entre as Variáveis EPI Score e Social Progress Index



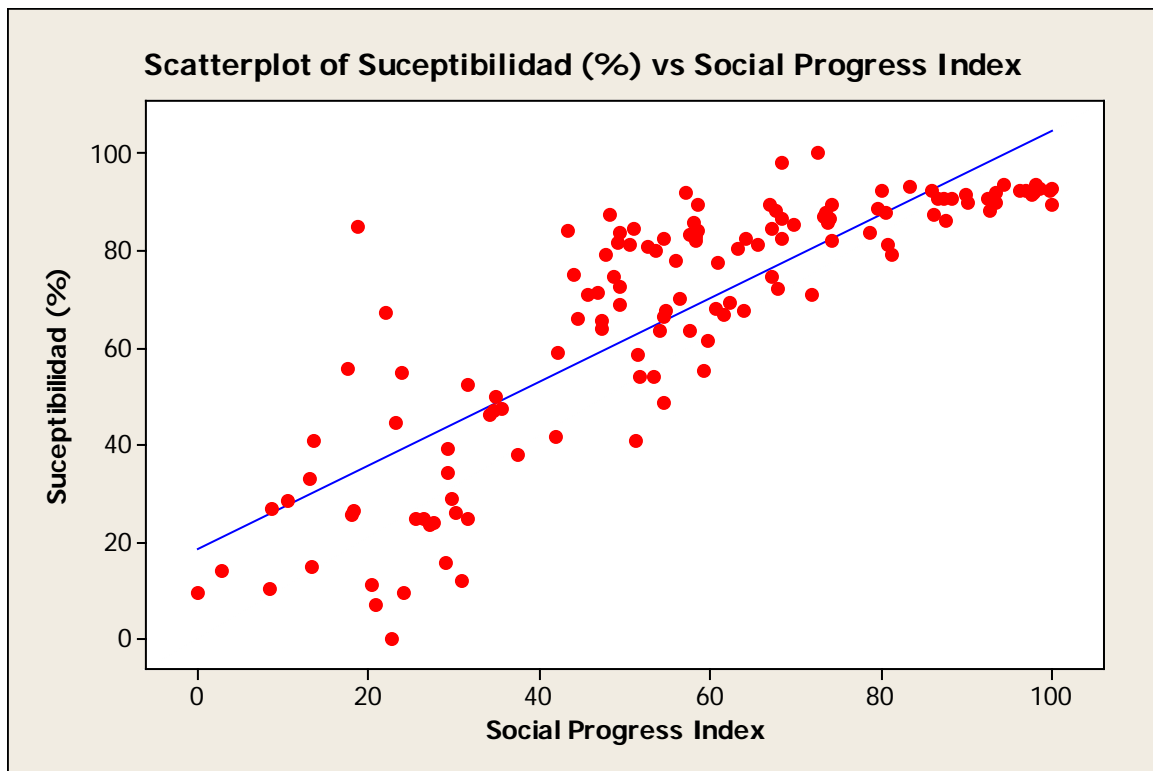
O gráfico de dispersão entre as variáveis “EPI Score” e “Social Progress Index” mostra que elas estão bastante relacionadas. No topo aparece a Suíça, com índice próximo de 100 em ambas variáveis, e no outro extremo aparece o Burundi, como índice aproximadamente 10 nos dois parâmetros. Chama a atenção a Republica Centro Africana como uma exceção, pois apresenta índice menor de 3 em SPI e acima de 35 em EPI. Nota-se um grupo expressivo de ocorrências no índice até 40, e uma concentração importante no nível intermediário, próximo de 60.

### 5.3.4 Dispersão entre as Variáveis “Falta de Capacidade para Enfrentar Situações Difíceis” e “Falta de Capacidade de Adaptação”



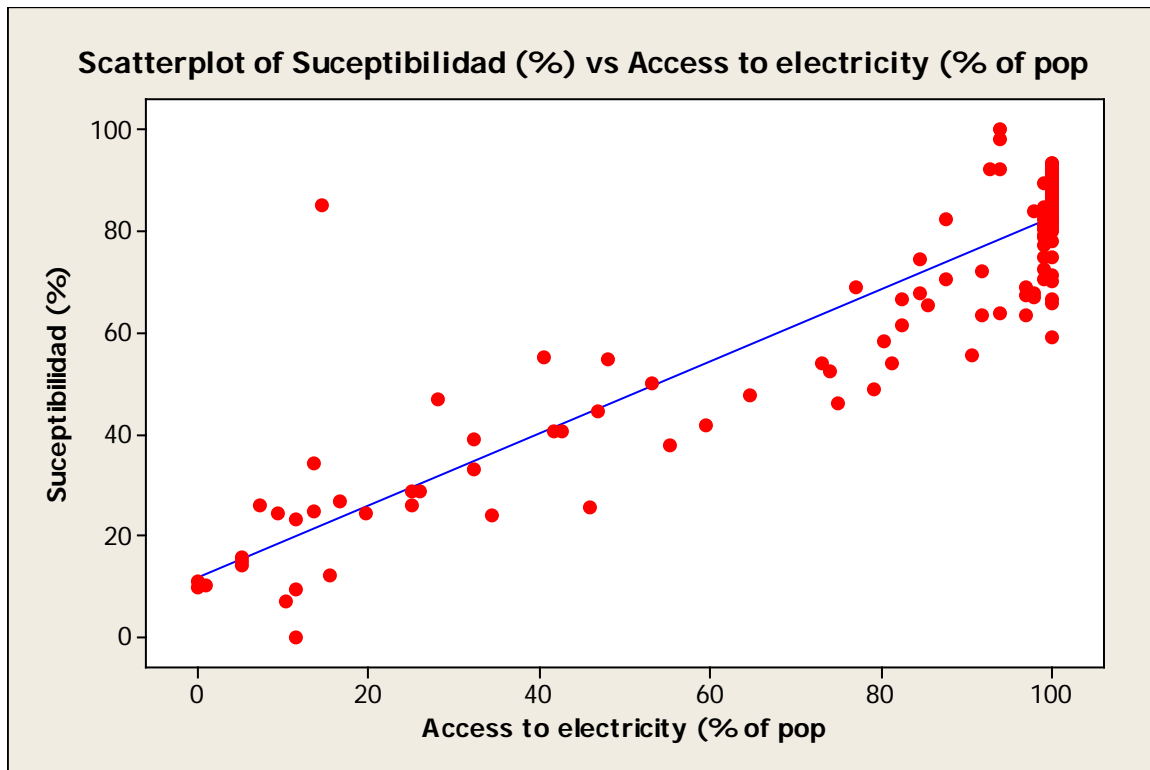
O gráfico mostra uma forte relação entre as variáveis “Falta de Capacidade para Enfrentar Situações Difíceis ” “Falta de Capacidade de Adaptação”, mostrando que normalmente um país apresenta as duas ou quando não, é débil em ambas. No extremo positivo temos países como Australia e Nova Zelandia, na Oceania, e no outro extremo Chade e Mali, no continente africano.

### 5.3.5 Dispersão entre as Variáveis Suscetibilidade e Social Progress Index



A dispersão mostra alta relação entre as variáveis “Suscetibilidade” e “Social Progress Index”. Num extremo aparecem Chad, com alta suscetibilidade e baixo progresso social, e do outro aparece a Islândia, com baixa suscetibilidade e alto progresso social. Uma exceção que demonstra não haver relação entre as variáveis é a Mauritânia, que tem SPI baixo mas não apresenta suscetibilidade alta, talvez devido ao fato de que é uma região desértica não sujeita a desastres naturais.

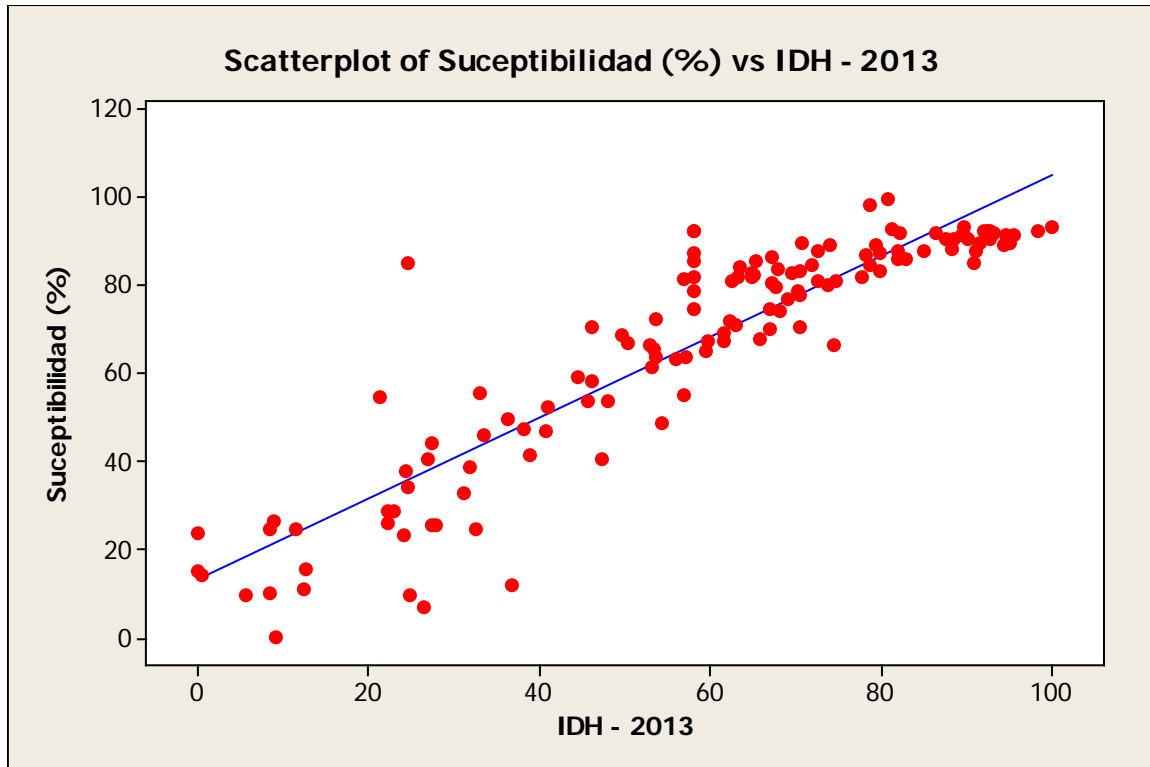
### 5.3.6 Dispersão entre as Variáveis Suscetibilidade e Acesso à Eletricidade



O gráfico apresenta uma certa relação entre as duas variáveis, embora, no meu entender, não faça muito sentido uma conexão direta entre acesso à eletricidade e suscetibilidade a desastres naturais. De qualquer forma, a maior parte dos países com 100% de acesso à eletricidade são também os que apresentam menor suscetibilidade. A exceção é o Tadjiquistão, cuja população tem alto índice de acesso à eletricidade mas apresenta índice de suscetibilidade em torno de 60, provavelmente devido ao fato de estar totalmente circundado por cordilheiras. Outra exceção, mas do lado oposto, é a Mauritânia, que possui baixíssimo acesso a eletricidade e também pouca suscetibilidade, provavelmente em função de seu território desértico, sem muitas variações climáticas ou riscos de acidentes naturais.



### 5.3.7 Dispersão entre as Variáveis Suscetibilidade e IDH



O gráfico mostra que parece haver relação entre as duas variáveis, e países com alto IDH apresentam baixa suscetibilidade, caso da Austrália e da Noruega, e do outro lado encontramos Níger e Rep Centro Africana, que tem baixíssimo IDH e apresentam alta suscetibilidade. Uma exceção interessante no continente africano é a Mauritânia, que apresenta índice de IDH baixo, o que era de se esperar, e índice de suscetibilidade também baixo, o que talvez se explique pelo fato de o país ser uma planície semidesértica, sem grandes riscos de desastres naturais.

Deste ponto do trabalho em diante, utilizaremos apenas a função StepWise para obtermos a melhor equação de regressão.

## 5.4 Regressão

Analisando o dendrograma, tomamos duas variáveis com alta similaridade e aplicamos a regressão escalonada (stepwise). Segue o resultado da função:

### Stepwise Regression: Social Progress versus EPI Score; 1. Governança

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

Response is Social Progress Index on 2 predictors, with N = 132

Step	1	2
Constant	7,670	6,602
EPI Score	0,980	0,611
T-Value	26,28	12,22
P-Value	0,000	0,000
1. Governança		0,427
T-Value		9,09
P-Value		0,000
S	10,2	7,99
R-Sq	84,16	90,35
R-Sq(adj)	84,04	90,20
Mallows Cp	83,7	3,0

Na regressão escalonada entre “Social Progress Index” em função das variáveis “EPI Score” e “Governança”, vemos que há um alto nível de correlação. Se levamos em conta apenas a variável “EPI Score”, temos um R-Quadrado de 84,16, e ao agregarmos a variável “Governança” o R-Quadrado sobe para 90,35, portanto a combinação ajuda a explicar melhor a variável SPI.

Para fins de comparação, tomamos Social Progress Index em função de duas variáveis que apresentam baixa similaridade a partir da análise do dendrograma:

### Stepwise Regression: Social Progr versus Outdoor air ; Alternative

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

Response is Social Progress Index on 2 predictors, with N = 132

Step	1	2
Constant	48,29	44,87
Alternative and nuclear energy	0,73	0,73
T-Value	5,70	5,71
P-Value	0,000	0,000
Outdoor air pollution attributa		0,16
T-Value		1,50
P-Value		0,137
S	22,9	22,8

R-Sq	20,00	21,37
R-Sq(adj)	19,38	20,15
Mallows Cp	3,2	3,0

Nesse caso, vemos que as variáveis em questão não ajudam a explicar Social Progress Index. Ao analisar o dendrograma, os gráficos de dispersão e a regressão, podemos identificar quais são as variáveis que apresentam maior relação, como é o caso entre “Susceptibilidade” e “IDH”, ou entre “Falta de Capacidade de Adaptação” e “Falta de Capacidade para Enfrentar Situações Difíceis”, ou ainda entre “Governança” e “Social Progress Index”. Vimos também quais são as variáveis que tem relações frágeis, pouca ou nenhuma relação entre si. Seja como for, fica bastante evidente que elas caminham juntas, e os países que apresentam os melhores índices em uma variável do par também terá índice alto na outra.

Como não era de se esperar, os países que mostram melhores desempenhos são países ricos, localizados na Europa e Oceania, enquanto os países do continente africano ocupam os últimos lugares. Os números são bastante claros, e na minha cabeça paira a seguinte dúvida: em que medida, por exemplo, melhorar a governança elevaria o progresso social num destes países periféricos?

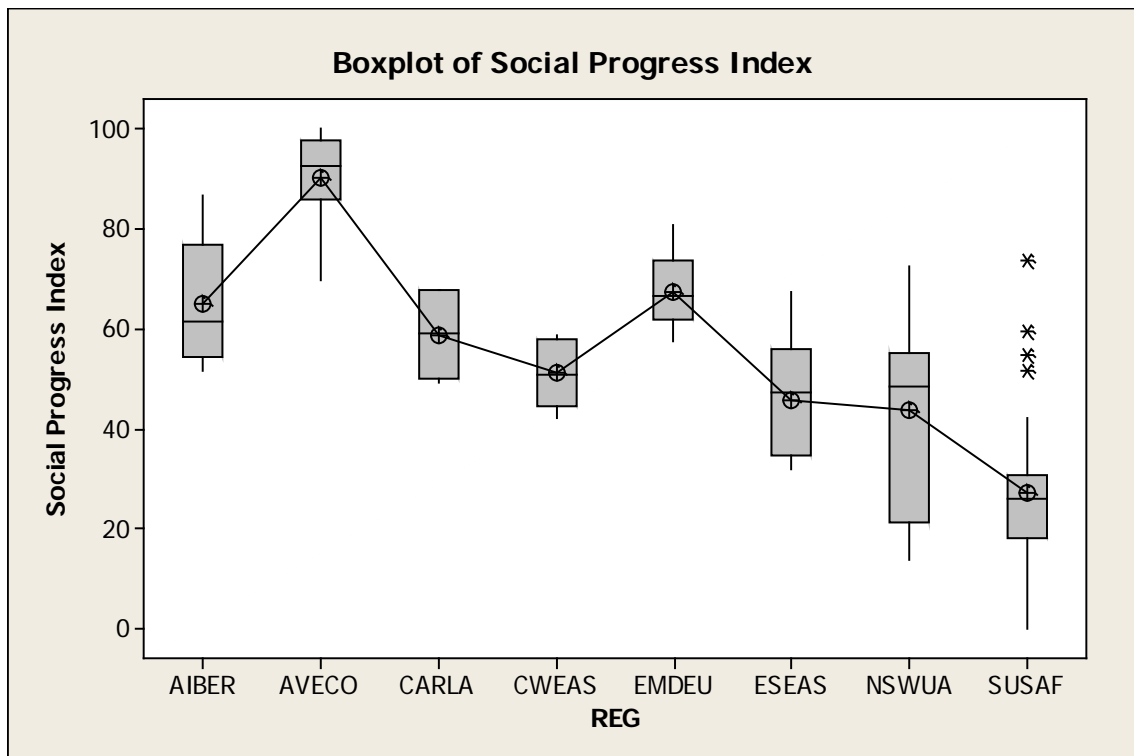
Cada vez que me debruço sobre esse grupo de variáveis fica clara sua interdependência, e sendo assim, atacar apenas uma delas não produziria o efeito desejado. Como deve ser quando se analisam políticas públicas, me parece fundamental lançar mão de uma série de medidas de longo prazo que impactem as variáveis simultaneamente, de forma tal que a melhora, ainda que gradual, ocorra sinergicamente, gerando efeito positivo em todas ou na maior parte das variáveis em questão.

O mais interessante, do ponto de vista metodológico, é ver como a estatística se converte em uma ferramenta altamente funcional. As conexões entre as variáveis apresentadas pelos modelos estatísticos nos levam a buscar entender os porquês de certas relações – algumas mais esperadas, por sua natureza similar, e outras menos óbvias, provocando reflexões interessantes. Para mim, justamente aí reside a importância desse trabalho.

## 6 – ANÁLISE DE VARIANCIA

### 6.1 – Variável Social Progress Index

Primeiramente iremos analisar os destaques positivos e a presença de outliers nas amostras através de gráficos Boxplot

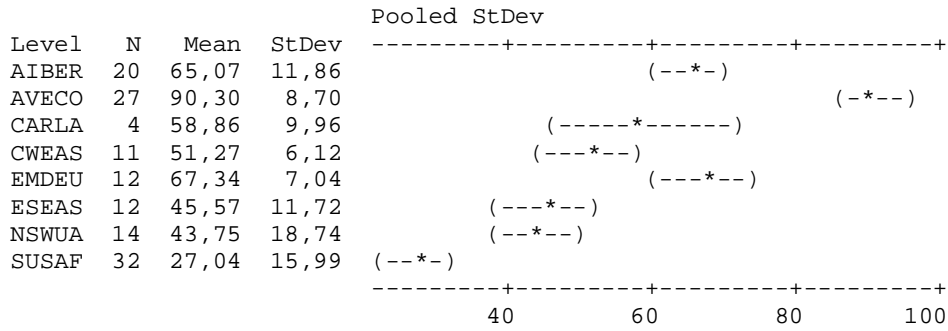


O destaque positivo fica por conta dos países de economia avançada, com média 90 em IDH, dentre os quais se destacam Nova Zelândia, Suíça, Islândia, Holanda e Noruega. No outro extremo encontram-se os países da África Subsaariana com média 27, na qual figuram países como Sudão, Guiné, Burundi, Republica Centro Africana e Chade. Vale ressaltar, ainda os quatro *outliers* no grupo da África subsaariana, que aparecem em condição melhor que a média da região: Ilhas Maurício, Botswana, África do Sul e Namíbia. No grupo da AIBER, destacam-se positivamente, além de Portugal e Espanha, como era de se esperar por estarem inseridos no contexto europeu, países como Uruguai, Costa Rica e Chile.

#### One-way ANOVA: Social Progress Index versus REG

Source	DF	SS	MS	F	P
REG	7	65477	9354	58,42	0,000
Error	124	19856	160		
Total	131	85333			

S = 12,65    R-Sq = 76,73%    R-Sq(adj) = 75,42%

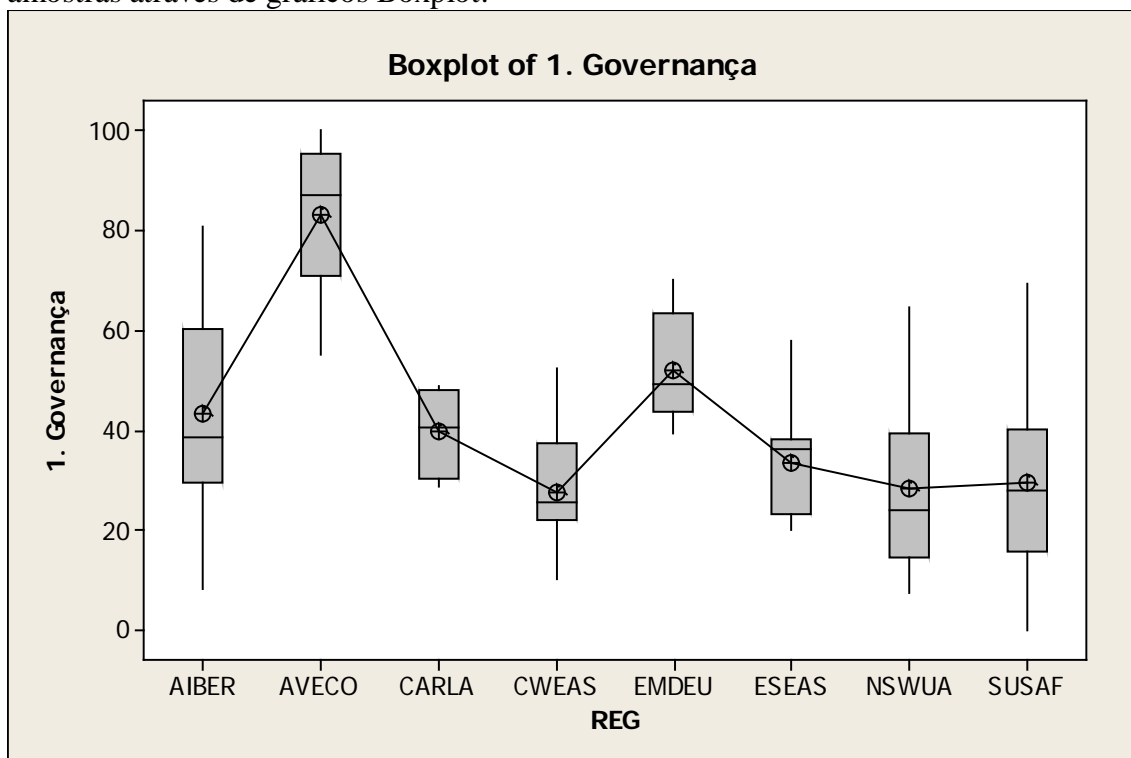


Pooled StDev = 12,65

Complementando o boxplot, a análise visual dos dados acima mostra haver alguma sobreposição entre as várias regiões, exceto os países contidos em SUSAF e AVECO que se encontram em posições extremas – de qualquer forma nota-se que as médias não são iguais o que se confirma pela análise de F (relativamente alto) e *P-value* baixo.

### 6.2 – Variável Governança

Primeiramente iremos analisar os destaques positivos e a eventual presença de outliers nas amostras através de gráficos Boxplot:

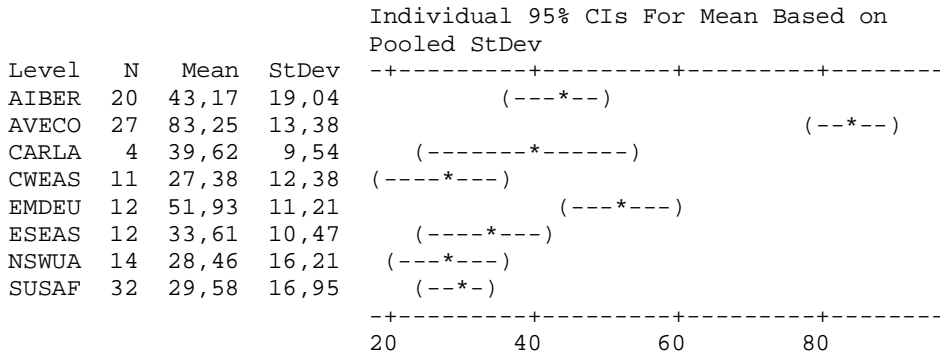


Não surpreende que os países do grupo da AVECO sejam também os que fazem melhor governança, com destaque para Finlândia, Noruega, Suécia, Nova Zelândia e Dinamarca que apresentam índices acima de 95 O segundo grupo de países mais bem ranqueados também estão quase todos na Europa, como Polônia, Lituânia, Hungria e Croácia, com índices entre 60 e 70. Os piores índices se encontram na Africa Subsaariana (Sudão, Rep Centro Africana e Chade) e no Oriente Médio como Iraque, Iêmen e Paquistão. Os países da América Latina ocupam uma posição intermediária.

### One-way ANOVA: 1. Governança versus REG

Source	DF	SS	MS	F	P
REG	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%

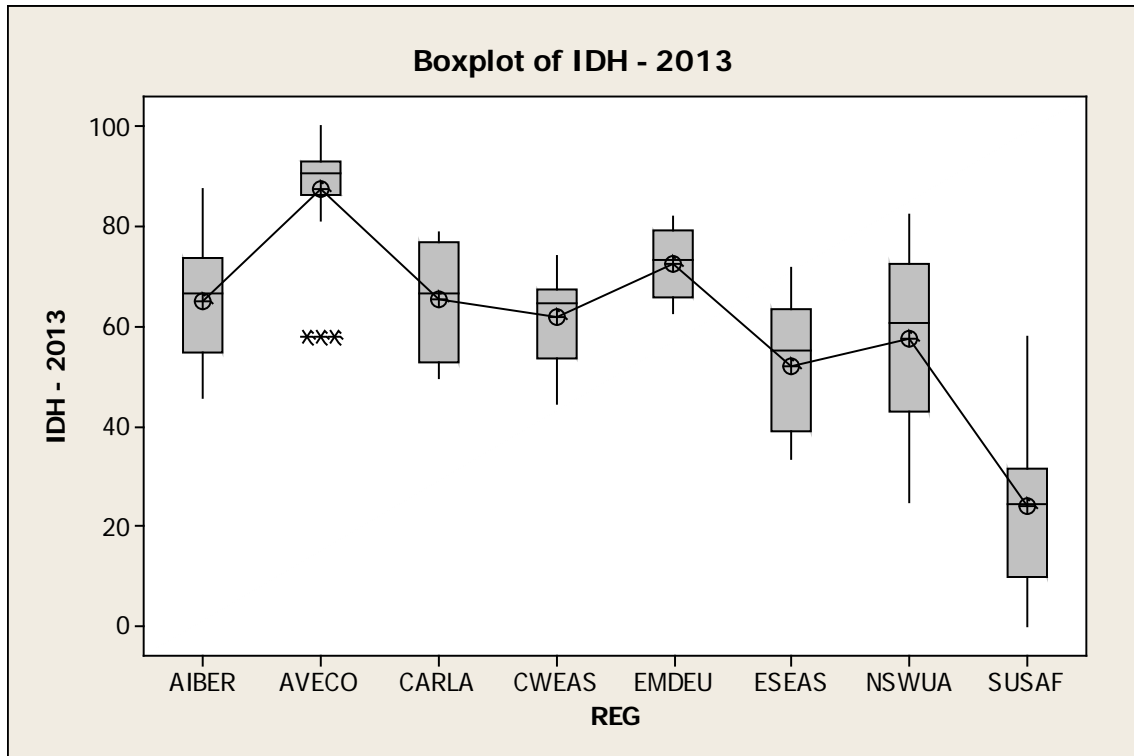


Pooled StDev = 15,10

Complementando o boxplot, a análise visual dos dados acima mostra haver sobreposição entre as regiões, exceto os países contidos na AVECO cujas medias se localizam bem distante dos demais grupos – de qualquer forma nota-se que as medias são mais próximas, relativamente, que as medias da variável SPI, o que se confirma pela análise de F e *P-value*.

### 6.3 – Variável IDH

Primeiramente iremos analisar os destaques positivos e a presença de outliers nas amostras através de gráficos Boxplot:



Em linha com as análises anteriores, os países da AVECO apresentam índices mais altos na variável IDH, com médias acima de 90, como é o caso da Noruega, Austrália, Suécia, Estados Unidos e Alemanha. Três países desse grupo (Coreia, Holanda e Letônia) figuram como *outliers*, com índices mais próximos da média dos outros grupos. Novamente, os países da África Subsaariana são os que apresentam os piores índices nessa variável, com destaque para Burkina Faso, Chade, Rep. Centro-Africana, Congo e Niger. No grupo da AIBER, o destaque fica por conta da Espanha, Chile, Portugal, Argentina e Uruguai.

### One-way ANOVA: IDH - 2013 versus REG

Source	DF	SS	MS	F	P
REG	7	65011	9287	51,69	0,000
Error	124	22281	180		
Total	131	87292			

S = 13,40    R-Sq = 74,48%    R-Sq(adj) = 73,03%

Individual 95% CIs For Mean Based on Pooled StDev

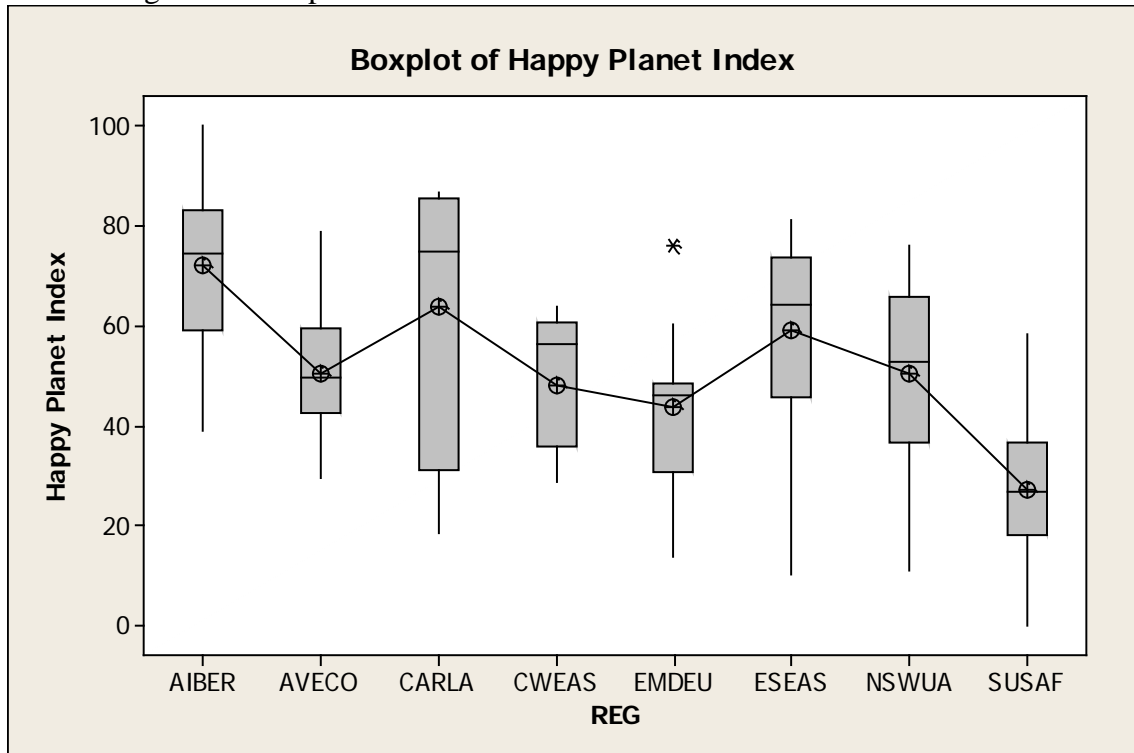
Level	N	Mean	StDev	CI Lower	CI Upper
AIBER	20	64,96	11,95	41,00	88,92
AVECO	27	87,36	11,35	74,66	100,06
CARLA	4	65,32	12,46	40,40	90,24
CWEAS	11	61,73	9,01	52,72	70,74
EMDEU	12	72,60	6,81	65,79	79,41
ESEAS	12	52,16	13,52	35,12	69,20
NSWUA	14	57,58	19,21	28,16	87,00
SUSAF	32	23,80	15,68	0,00	47,60

Pooled StDev = 13,40

Complementando o boxplot, a análise visual dos dados acima mostra haver alguma sobreposição entre as várias regiões, exceto os países contidos em SUSAF e AVECO que se encontram em posições extremas – de qualquer forma nota-se que as médias não são iguais o que se confirma pela análise de F (relativamente alto) e *P-value* baixo.

#### 6.4 – Variável Happy Planet Index

Primeiramente iremos analisar os destaques positivos e a presença de outliers nas amostras através de gráficos Boxplot



Essa é uma das poucas variáveis em que os grupo de países da AVECO não ocupa as primeiras posições. Acredito que isso se deva a alguns fatores, como a um certo pessimismo, ou realismo de sua população. Quem desponta nesse quesito são os países do grupo AIBER e do grupo CARLA, como Colômbia, Venezuela, Costa Rica, Jamaica, El Salvador e Panamá. Vale ressaltar, também, a presença da Albânia como outlier dentro do grupo EMDEU,

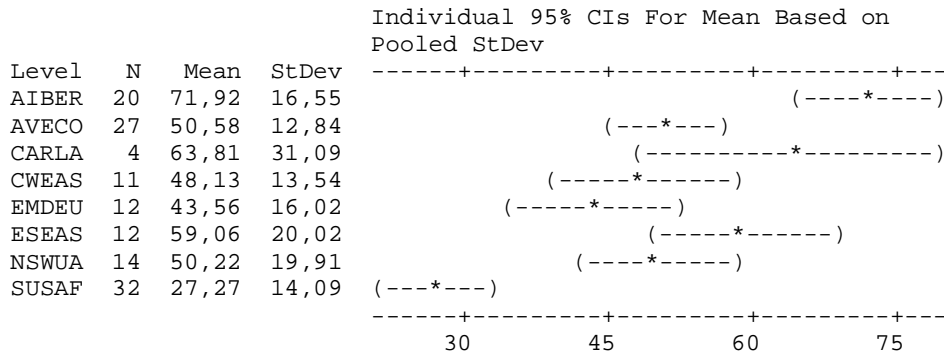


apresentando o índice mais alto entre os 12 países. Outro ponto que chama a atenção é o da África Subsaariana, que apresenta os índices mais baixos, contrariando uma visão distorcida de que os africanos, embora pobres, são felizes.

### One-way ANOVA: Happy Planet Index versus REG

Source	DF	SS	MS	F	P
REG	7	28129	4018	15,21	0,000
Error	124	32765	264		
Total	131	60894			

S = 16,26    R-Sq = 46,19%    R-Sq(adj) = 43,16%

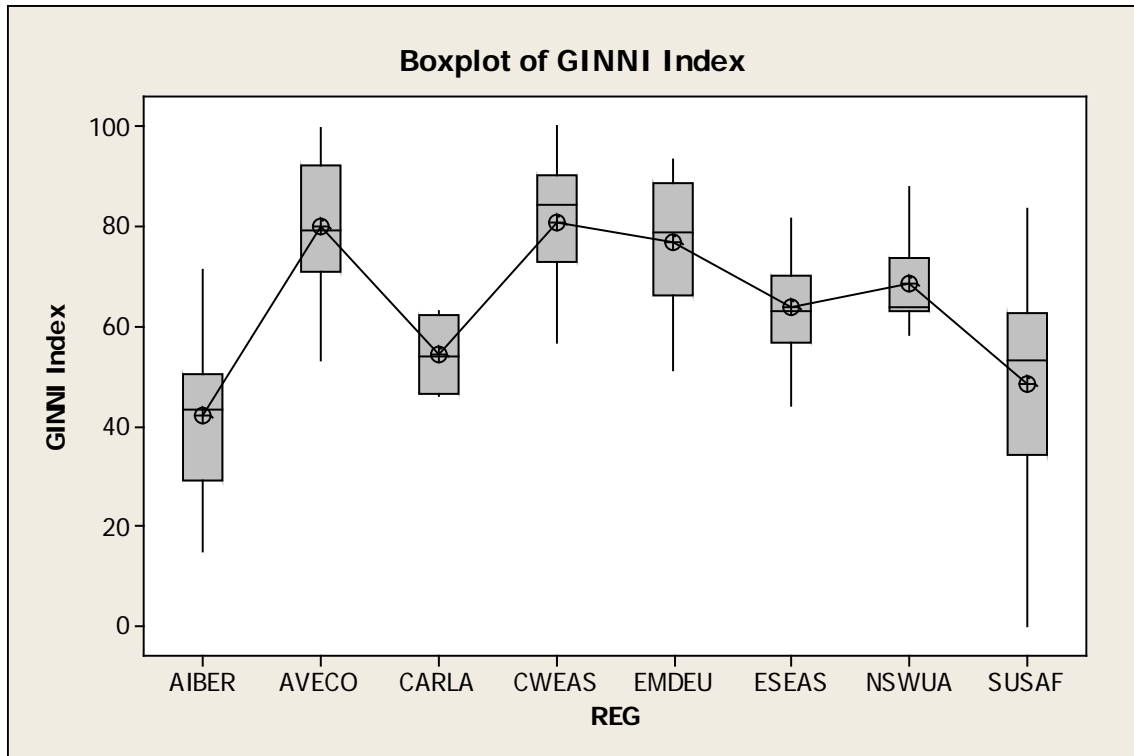


Pooled StDev = 16,26

Complementando o boxplot, a análise visual dos dados acima mostra haver alguma sobreposição entre as médias das várias regiões, exceto os países contidos em SUSAF, que se encontram num extremo – de qualquer forma nota-se que as médias relativamente próximas, o que se confirma pela análise de F e *P-value*.

## 6.5 – Variável GINI Index

Primeiramente iremos analisar os destaques positivos e a presença de eventuais outliers nas amostras através de gráficos Boxplot:



Nessa variável tampouco há predominância dos países da AVECO, como costuma acontecer em outras variáveis. Inclusive, nesse caso, o grupo de países da CWEAS - Comunidade de Estados Independentes (Ucrânia, Bielorrússia, Cazaquistão) aparecem em ligeira vantagem (média 80,7), seguidos pelos países do grupo AVECO (média 79,9) e EMDEU (média 76,9). No caso dos primeiros, seria um efeito de seu antigo regime comunista, que pregava maior igualdade entre os cidadãos? Outra constatação importante é que, na média, os países que compõem a AIBER (destaques: Honduras, Brasil e Colômbia, nessa ordem) apresentam índice mais baixo que os da África Subsaariana (Zâmbia, Botsuana, Namíbia e África do Sul, nessa ordem).

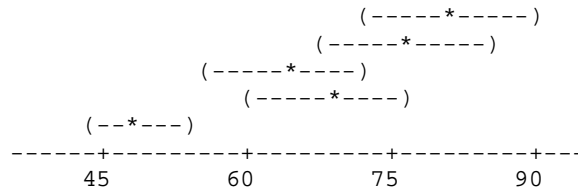
### One-way ANOVA: GINI Index versus REG

Source	DF	SS	MS	F	P
REG	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%

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

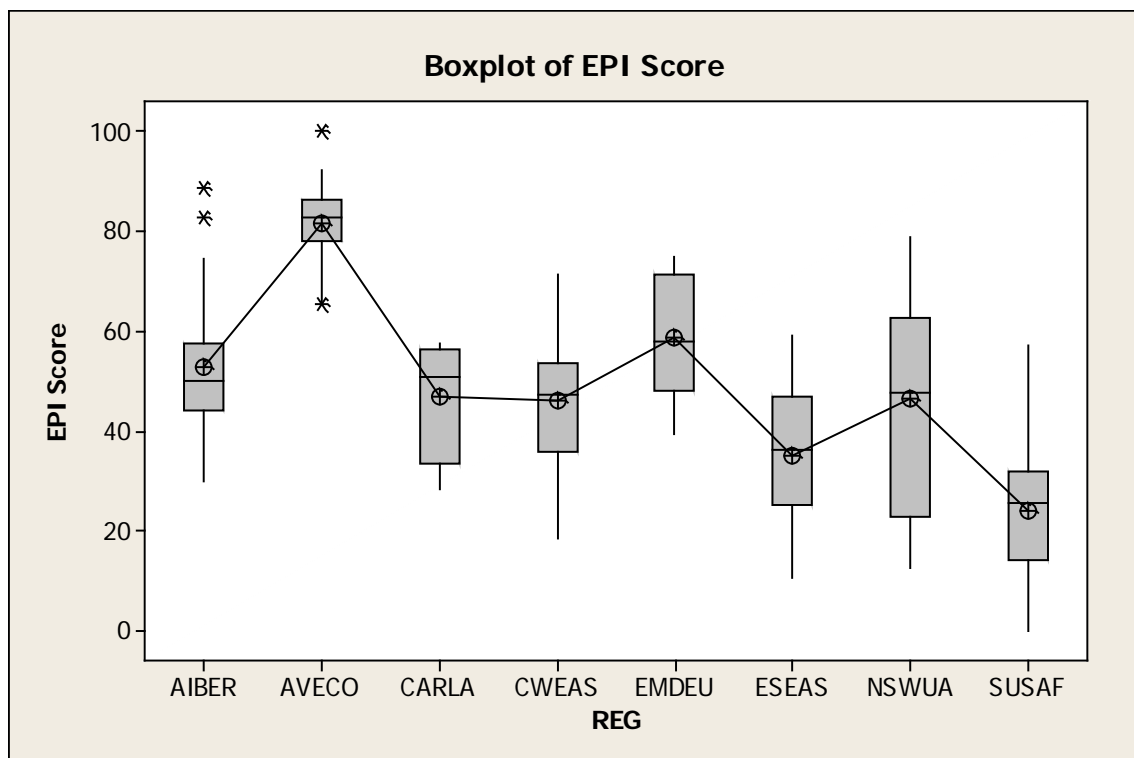


Pooled StDev = 15,23

Complementando o boxplot, a análise visual dos dados acima mostra haver alguma sobreposição entre as regiões (ESEAS e NSWUA; CWEAS e EMDEU) - de qualquer forma se nota que as médias relativamente próximas de algumas regiões o que se confirma pela análise de F e *P-value*.

## 6.6 – Variável EPI Score

Primeiramente iremos analisar os destaques positivos e a presença de outliers nas amostras através de gráficos Boxplot

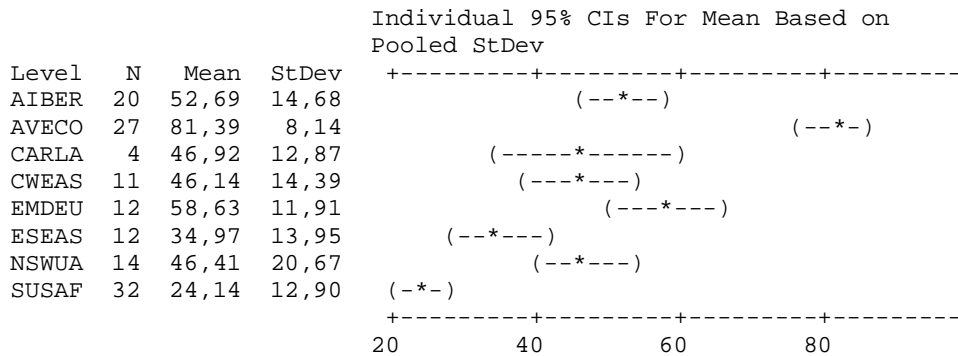


O primeiro comentário refere-se ao grupo AVECO, concentrados na média 80. Vale ressaltar dois outliers: de um lado a Suíça, que figura no topo da lista com índice 100, e de outro lado, a Coreia do Sul, com média 65. O grupo da AIBER tem média mais de 10 pontos abaixo, em torno de 52. Exceções nesse grupo, e que figuram como outliers, são Portugal e Espanha, como índices 82 e 88, respectivamente.

### One-way ANOVA: EPI Score versus REG

Source	DF	SS	MS	F	P
REG	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%



Complementando o boxplot, a análise visual dos dados acima mostra haver alguma sobreposição entre as médias das várias regiões, exceto os países contidos em AVECO e SUSAF, que se encontram em extremos opostos – de qualquer forma nota-se que as médias relativamente próximas (entre CARLA, CWEAS e NSWUA, por exemplo), o que se confirma pela análise de F e *P-value*.

## 6.7 – Quadro Resumo – Teste F

O teste F é uma ferramenta de análise de dados que consiste em comparar os resultados encontrados para duas variações de população. Um valor de "F" próximo da unidade é uma evidência de que as variações de população subjacentes são iguais. Segue um quadro resumo apresentando o valor de F para todas as variáveis analisadas:

Variável	F-value
1. Acesso a eletricidade	60,89
2. Suscetibilidade	60,01
3. Índice de Progresso Social	58,42
4. Vulnerabilidade	53,25
5. IDH - 2013	51,69
6. Capacidade de enfrentar situações difíceis	51,27
7. EPI Score	40,44
8. Capacidade de adaptação	37,82
9. Governança	35,53
10. GINNI Index	18,31
11. Qualidade do fornecimento de eletricidade	16,7
12. Happy Planet Index	15,21
13. Retirada de água	10,69
14. Combustíveis renováveis e resíduos	10,45
15. Mortes atribuídas à poluição atmosférica	9,87
16. Consumo de Energia Elétrica	7,31
17. Índice mundial de risco a desastres naturais	7,23
18. Uso de Energias alternativas e nuclear	6,4
19. Consumo energia de combustíveis fósseis	6,08

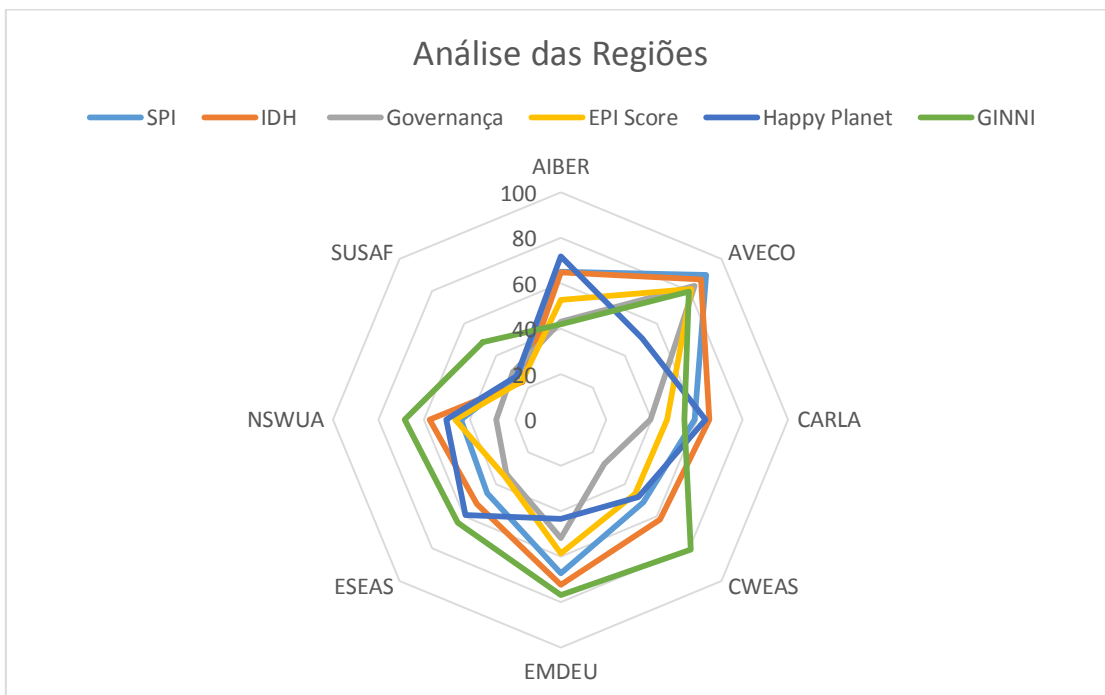
20. Biodiversidade e habitat	4,19
21. Exposição	2,94
22. Emissões de gases estufa	1,62

Analisando os valores de F, a comparação mostra haver maior proximidade das médias dos índices entre as regiões na variável “Acesso à Eletricidade” ( $F = 60,89$ ), e, no outro extremo, pouca sobreposição das médias das regiões na variável “Emissões de Gases Estufa” ( $F = 1,62$ ).

## 7 - ANÁLISE DE GRÁFICO RADAR REDUZIDO

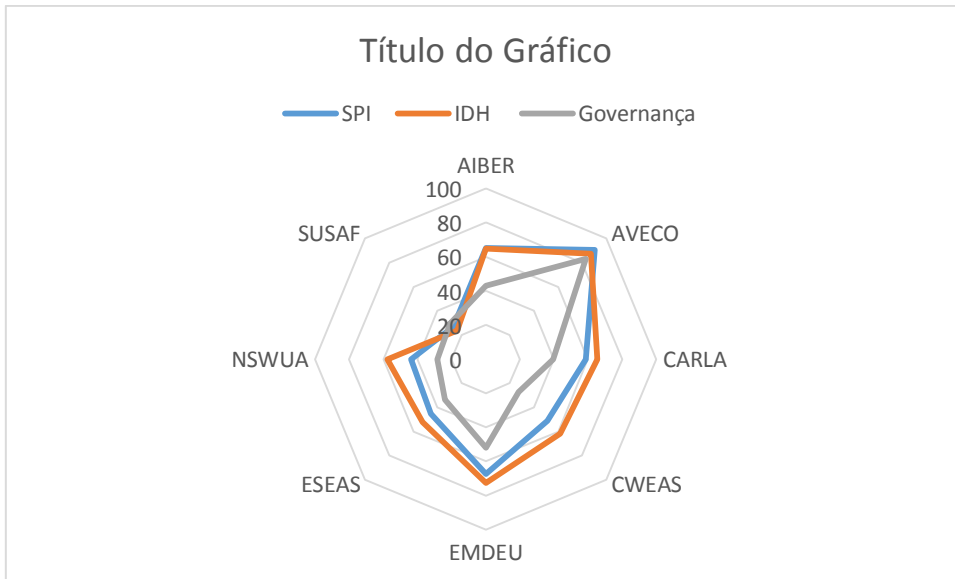
A seguir, faremos a análise das variáveis por meio de gráfico radar.

### 7.1 – Gráfico Radar - Todas as Variáveis (6)



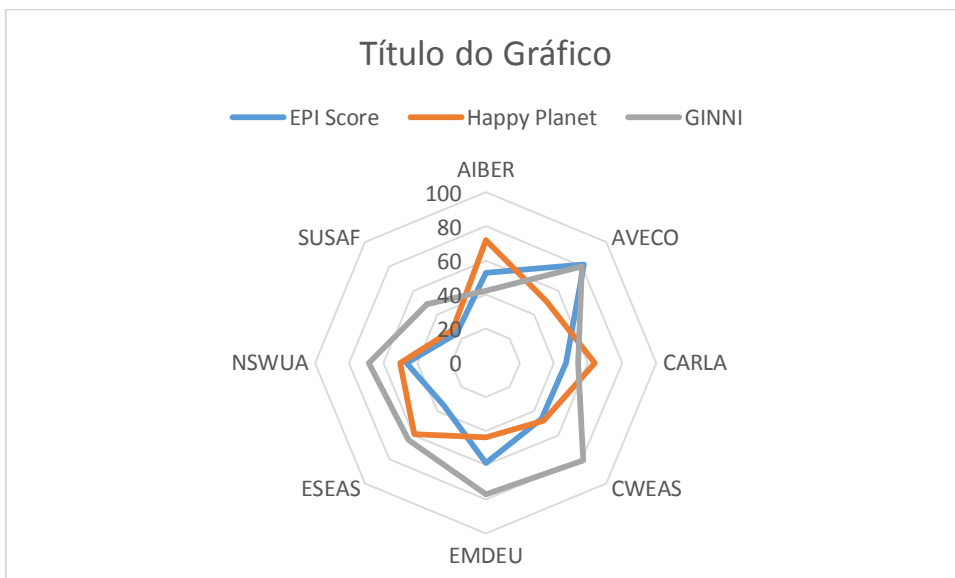
Analisando este primeiro gráfico radar, percebe-se que o grupo AVECO apresenta os melhores índices (média próxima a 80) em todas as variáveis, exceto em Happy Planet Index. No outro extremo estão os países da SUSAF.

### 7.2 – Gráfico Radar - Variáveis “SPI”, “IDH” e “Governança”



Olhando apenas três variáveis, vemos que SPI e IDH caminham bastante próximos na maior parte das regiões, enquanto governança fica mais deslocada dos outros dois. Novamente os destaques positivos ficam por conta dos países do grupo AVECO (média 85) e os negativos, pelos países da SUSAF.

### 7.3 – Gráfico Radar - Variáveis “EPI Score”, “Happy Planet” e “GINNI”



Nesse gráfico vemos um destaque positivo em CWEAS, EMDEU e AVECO na variável “GINNI” (média próxima de 80) e em AIBER, ESEAS e CARLA na variável “Happy Planet Index”

Essa etapa teve como objetivo comparar a média de alguns indicadores sintéticos que remetem ao bem-estar da população de uma dada nação. A comparação ocorrerá entre grupos de países classificados geograficamente e segundo seu nível de renda, visando cotejar a média dos indicadores dos grupos de países para apontar semelhanças e diferenças. Como destaques positivos aparecem os países do grupo AVECO (exceto em “Happy Planet Index”, e como

destaques negativos figuram os países da SUSAF. Os países da AIBER aparecem numa posição intermediária (com um destaque positivo em “Happy Planet Index”), com destaque positivo geral para Portugal e Espanha, que se beneficiam da inserção na União Europeia. Dentro desse grupo, o Brasil no geral figura em uma posição da linha intermediária para baixo. Exceções ficam por conta da Comunidade de Estados Independentes na variável GINNI, mostrando maior igualdade de renda, ao contrário dos países da AIBER, considerados os mais desiguais nesse quesito, em média, que os países do continente africano.

## **8. PESQUISA POR AMOSTRAGEM**

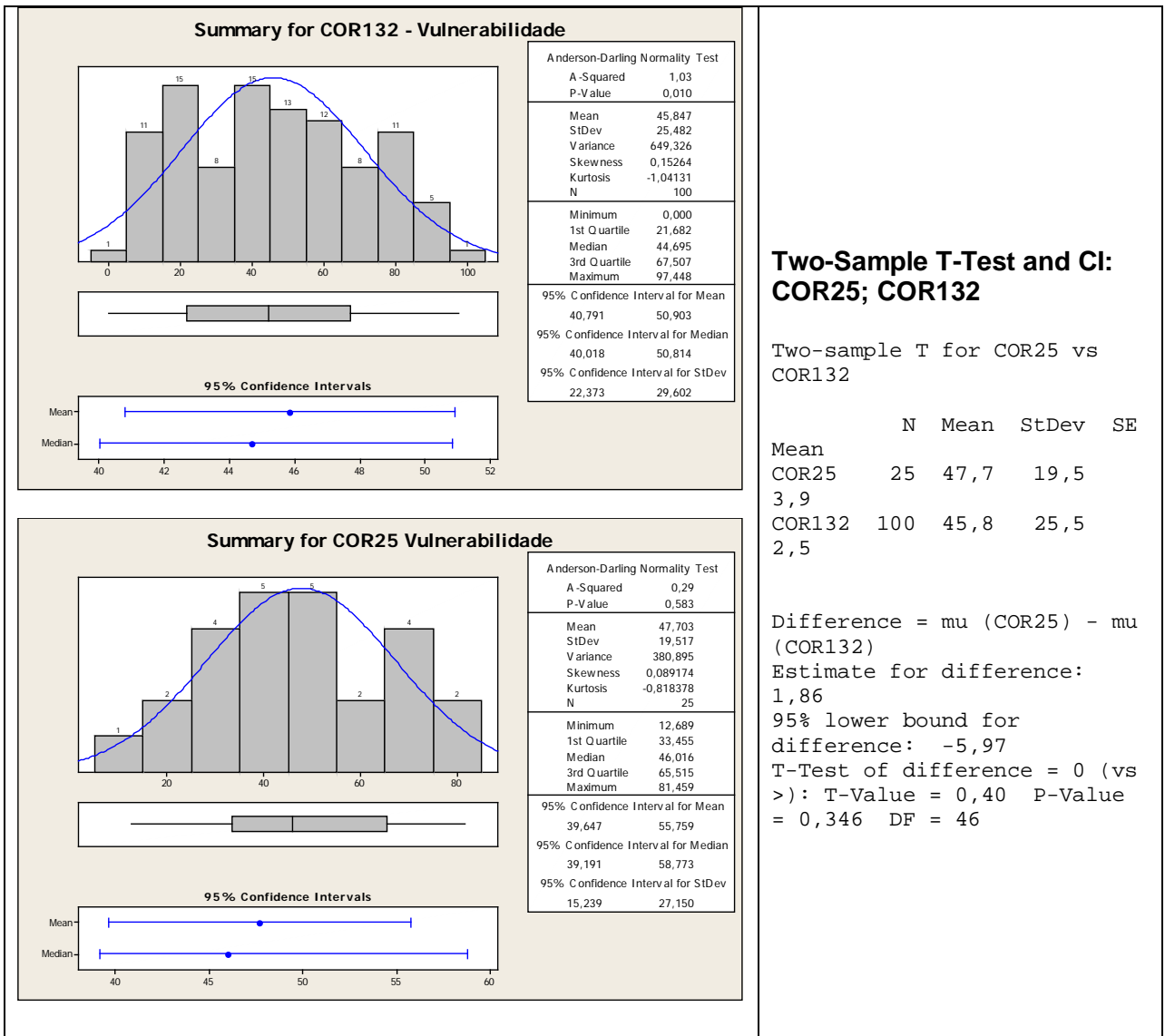
A pesquisa por amostragem foi feita em quatro amostras, uma de 25 indivíduos, ou países, a outra composta por 50 países, uma ainda com 100 indivíduos e a última com a totalidade, isto é, 132 países. Para cada amostra foram efetuadas análises exploratórias de dados.

### **8.1 Análise Exploratória para Amostra de 25, 50, 100 e 132 Indivíduos**

A amostragem dos indivíduos foi feita através da funcionalidade “Random Data” do MINITAB

#### **8.1.1 Variável Vulnerabilidade**

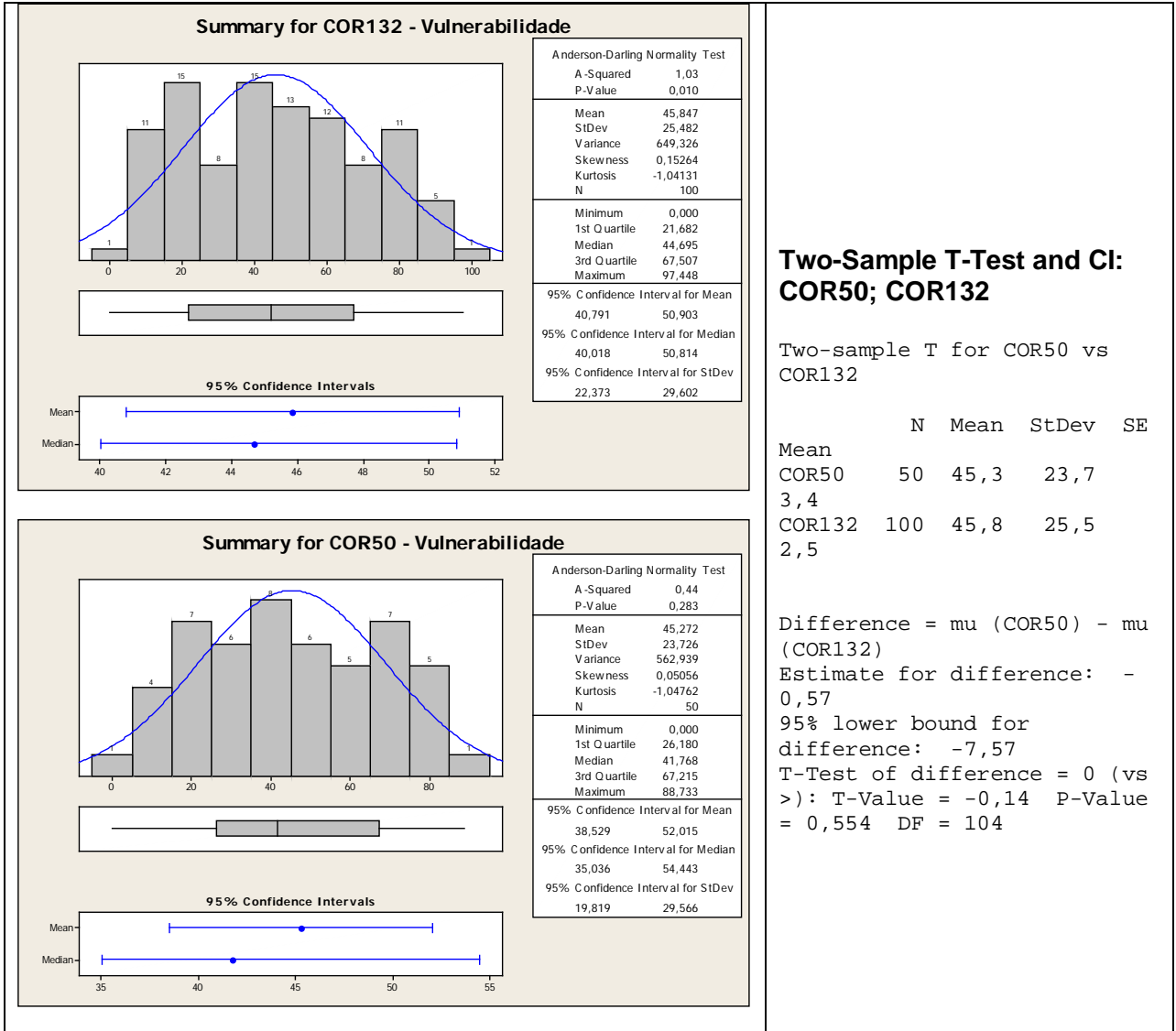
8.1.1.1 – Vulnerabilidade – Amostra de 25 Indivíduos



A amostra aleatória de 25 indivíduos apresentou média bastante próxima da população completa (47,7 contra 45,8), mostrando haver pouca variação entre os gráficos comparativamente; sendo assim, com base nisso podemos dizer que a amostra é significativa.

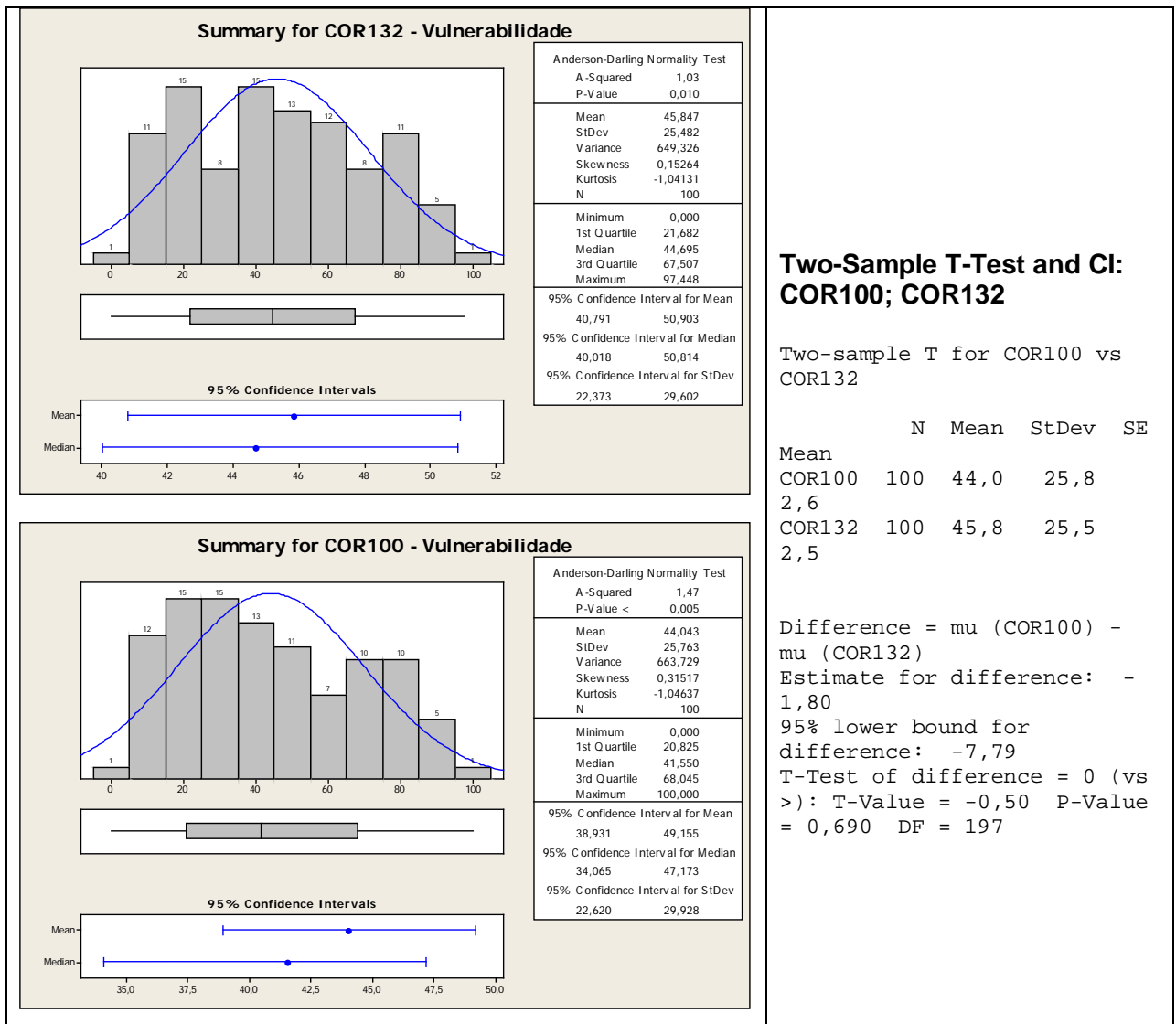
8.1.1.2 – Vulnerabilidade – Amostra de 50 Indivíduos





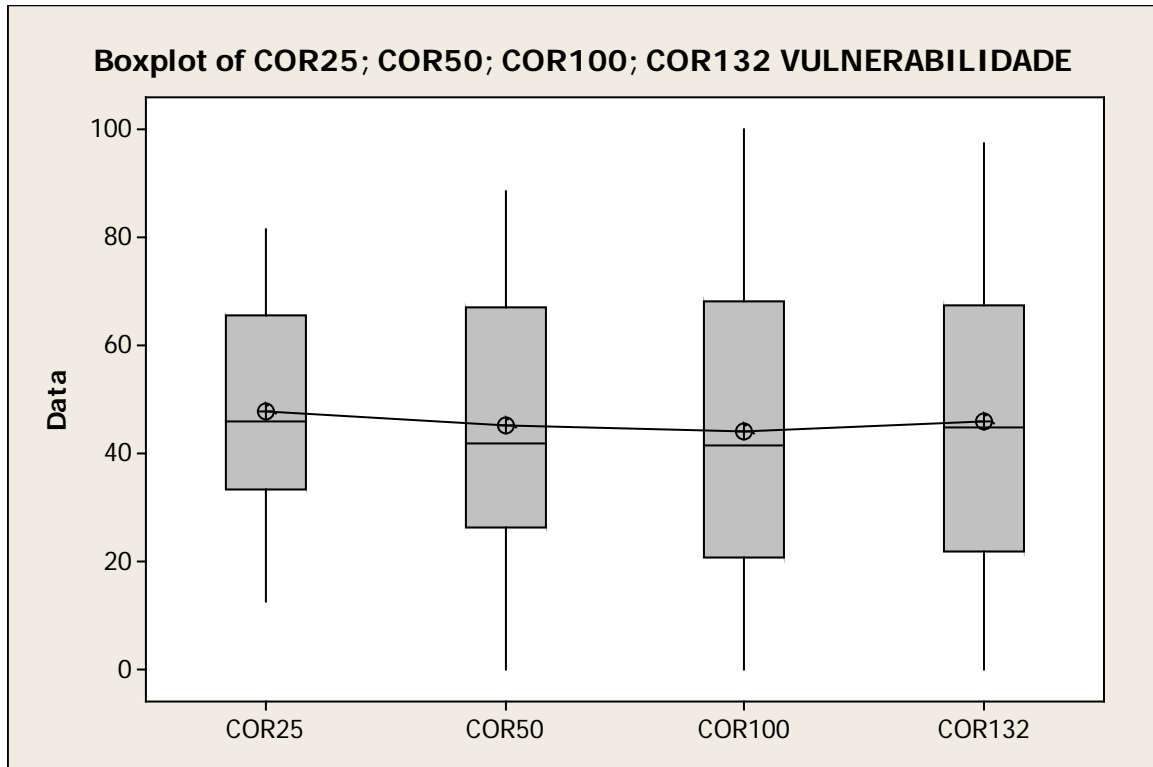
Na amostra aleatória com 50 indivíduos, notamos haver uma convergência ainda maior com a universo completo de 132 países, já que as médias de ambas as populações é praticamente o mesmo (45,3 e 45,8, respectivamente). Com isso, temos que poderia haver uma redução da amostra sem prejudicar o resultado da pesquisa.

### 8.1.1.3 – Vulnerabilidade – Amostra de 100 Indivíduos



A análise de uma amostra aleatória de 100 indivíduos apresenta uma média pouco diferente em comparação à da população total. Mais uma vez, isso parece revelar que não haveria necessidade de empregar uma amostra tão grande para realizar uma análise estatística com a mesma relevância e nível de acuracidade.

#### 8.1.1.4. Vulnerabilidade - Boxplot de Amostras Aleatórias com 25, 50, 100 Indivíduos e Amostra Completa



### One-way ANOVA: COR25; COR50; COR100; COR132

Source	DF	SS	MS	F	P
Factor	3	332	111	0,18	0,910
Error	271	166718	615		
Total	274	167050			

S = 24,80    R-Sq = 0,20%    R-Sq(adj) = 0,00%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
COR25	25	47,70	19,52
COR50	50	45,27	23,73
COR100	100	44,04	25,76
COR132	132	45,85	25,48

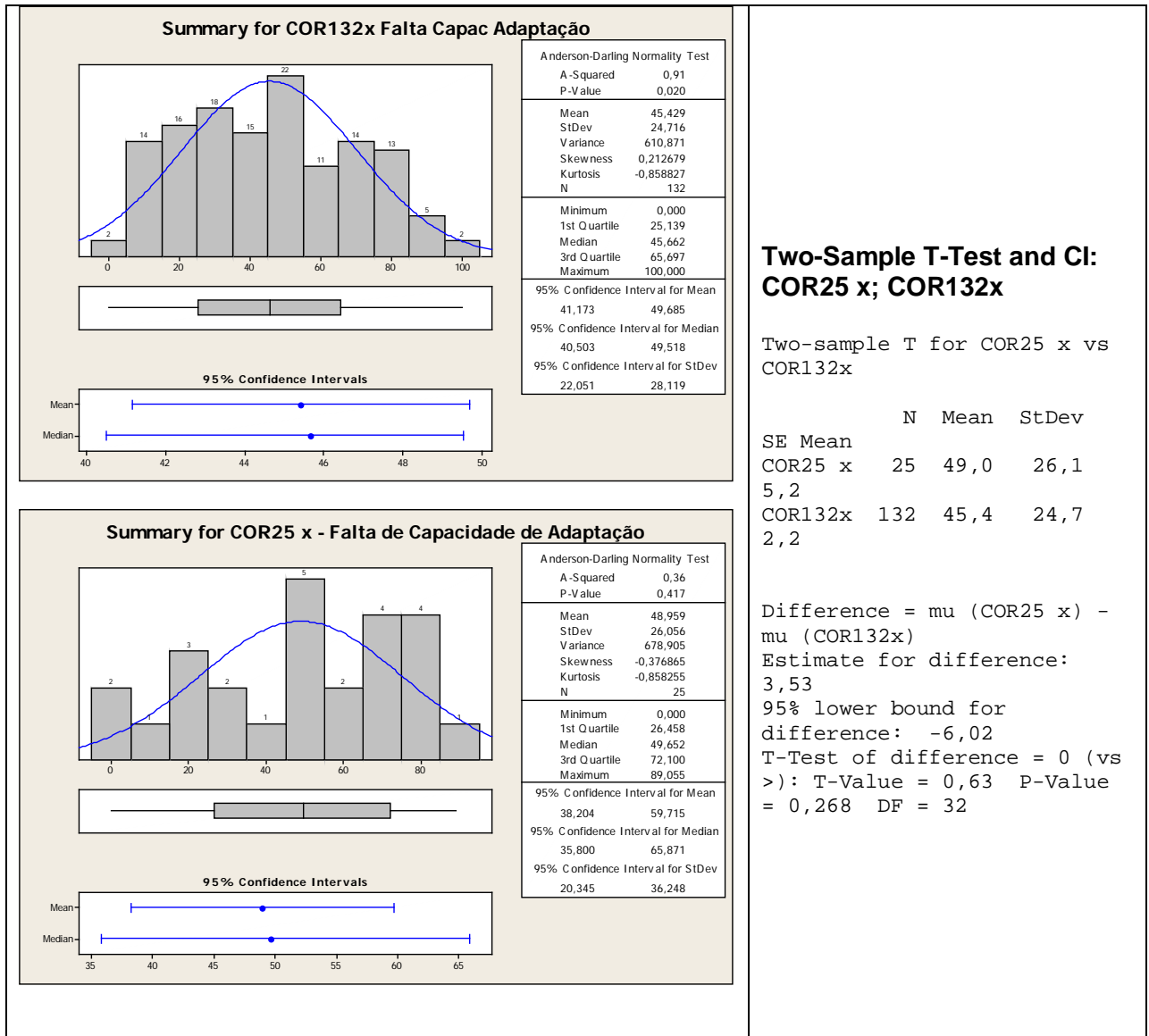
40,0    45,0    50,0    55,0

Pooled StDev = 24,80

A análise dos intervalos das médias de cada amostra populacional demonstra haver pouca diferença entre elas, o que é corroborado pelo valor de F.

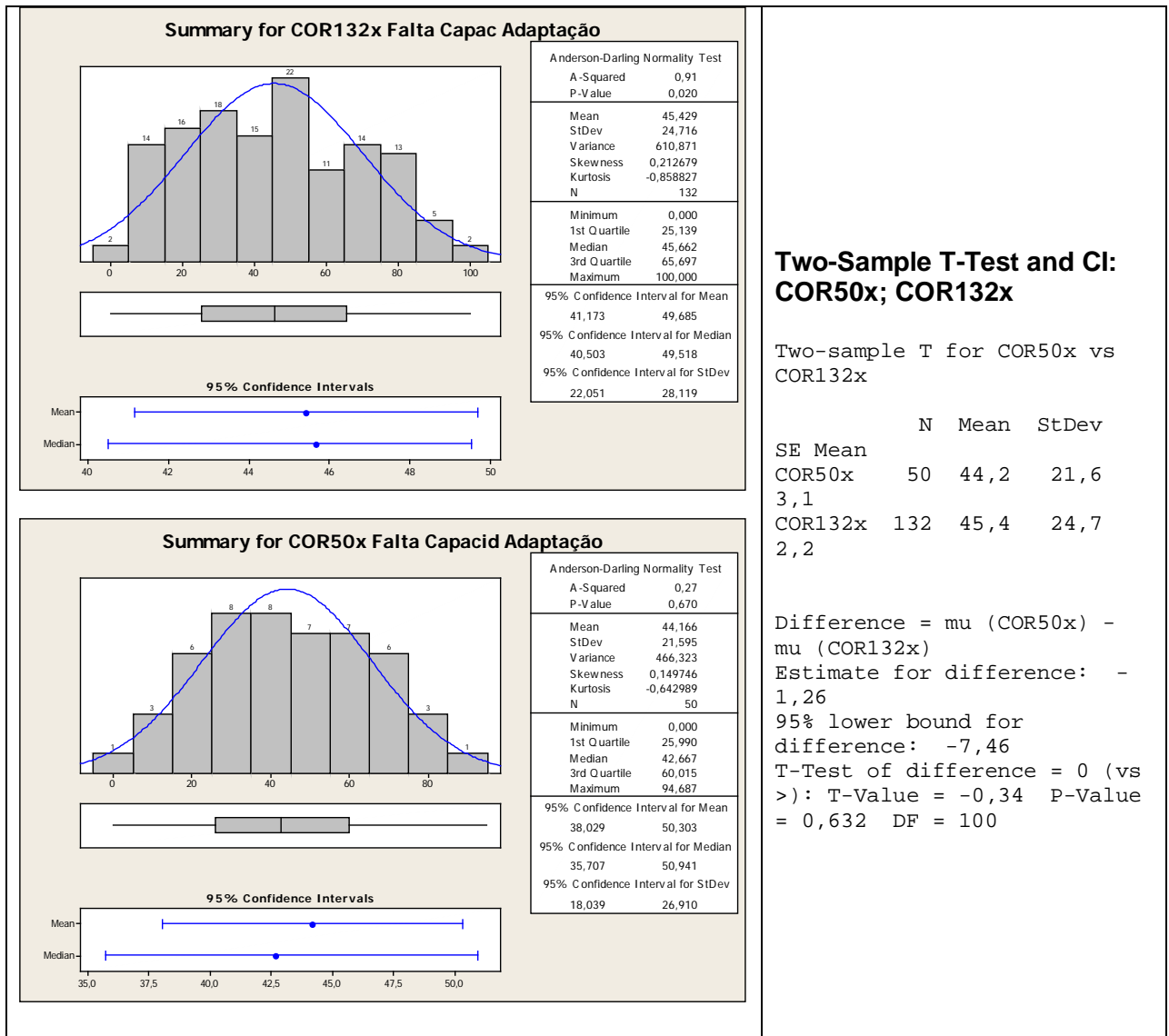
## 8.1.2 Variável Falta de Capacidade de Adaptação

### 8.1.2.1. Falta de Capacidade de Adaptação – Amostra de 25 Indivíduos

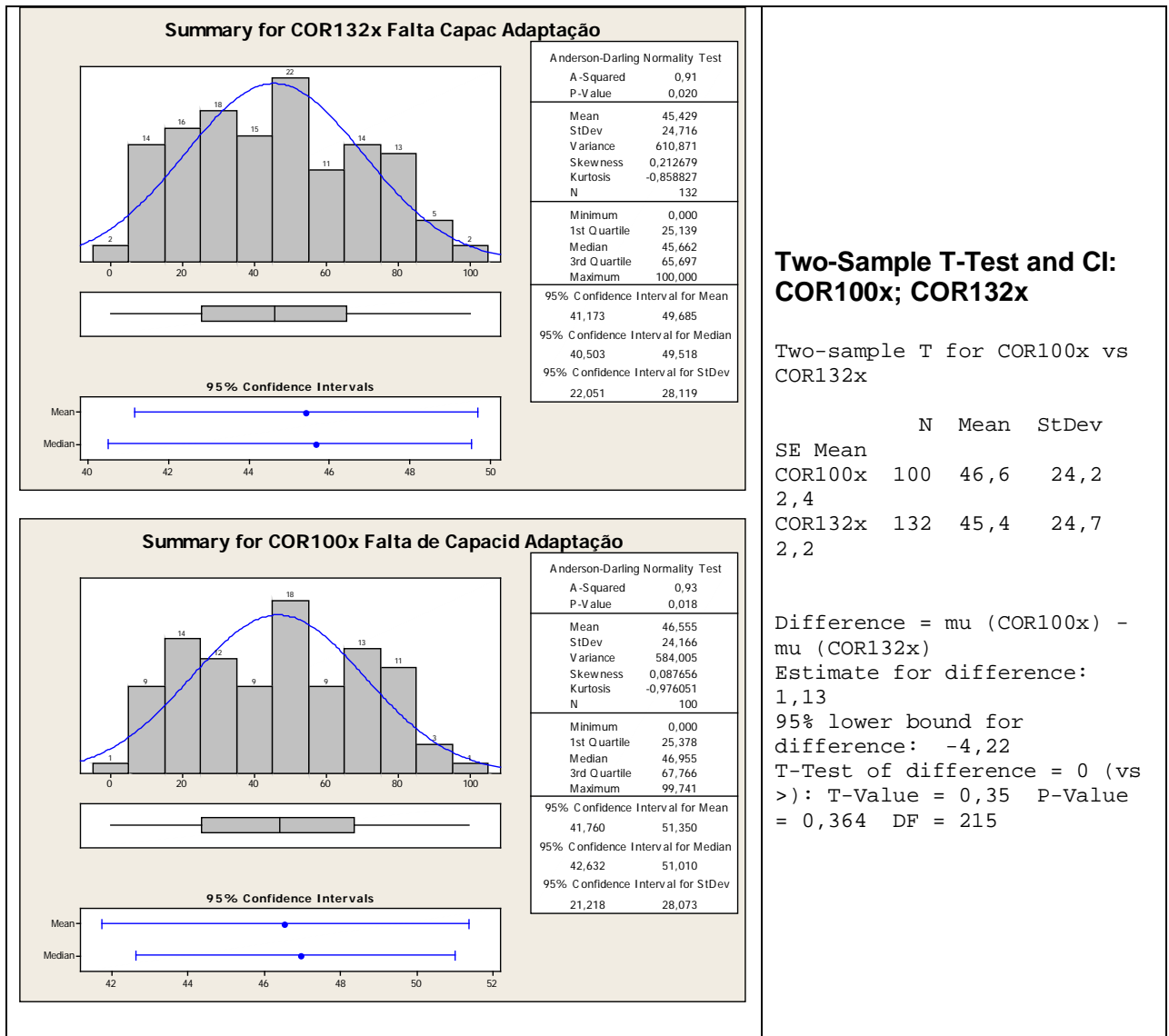


A análise da amostra aleatória de 25 indivíduos em comparação à população total mostra médias relativamente próximas (49 e 45, respectivamente). Com isso, a amostra de 25 elementos parece ser suficientemente válida para uma análise estatística confiável.

### 8.1.2.2. Falta de Capacidade de Adaptação – Amostra de 50 Indivíduos

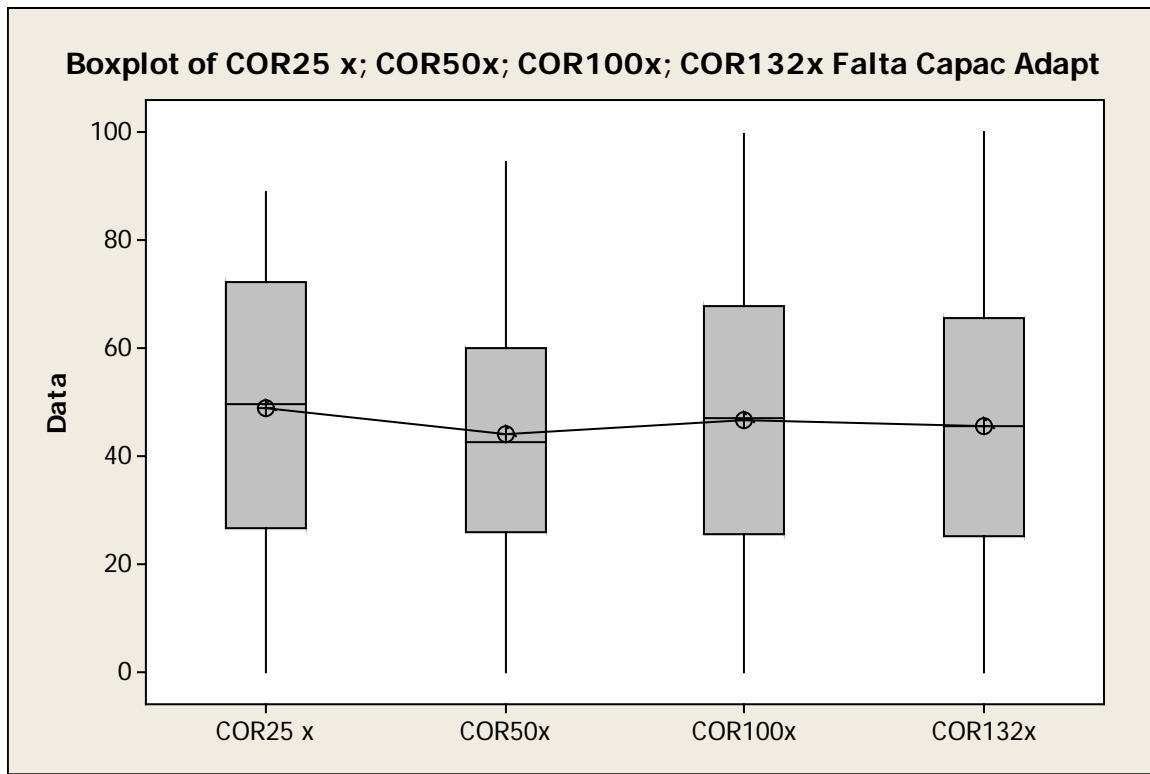


### 8.1.2.3. Falta de Capacidade de Adaptação – Amostra de 100 Indivíduos



Na amostragem aleatória com 100 indivíduos em comparação à população total, temos uma média relativamente mais alta, mas não de forma significativa (46,6 contra 45,4), o que corrobora os argumentos anteriores de que o aumento da amostragem, nesse caso em particular, pouco interfere no resultado estatístico.

### 8.1.2.4. Falta de Capacidade de Adaptação – Boxplot de Amostras Aleatórias com 25, 50, 100 Indivíduos e Amostra Completa



#### One-way ANOVA: COR25 x; COR50x; COR100x; COR132x

Source	DF	SS	MS	F	P
Factor	3	456	152	0,26	0,854
Error	303	176984	584		
Total	306	177440			

S = 24,17    R-Sq = 0,26%    R-Sq(adj) = 0,00%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
COR25 x	25	48,96	26,06
COR50x	50	44,17	21,59
COR100x	100	46,55	24,17
COR132x	132	45,43	24,72

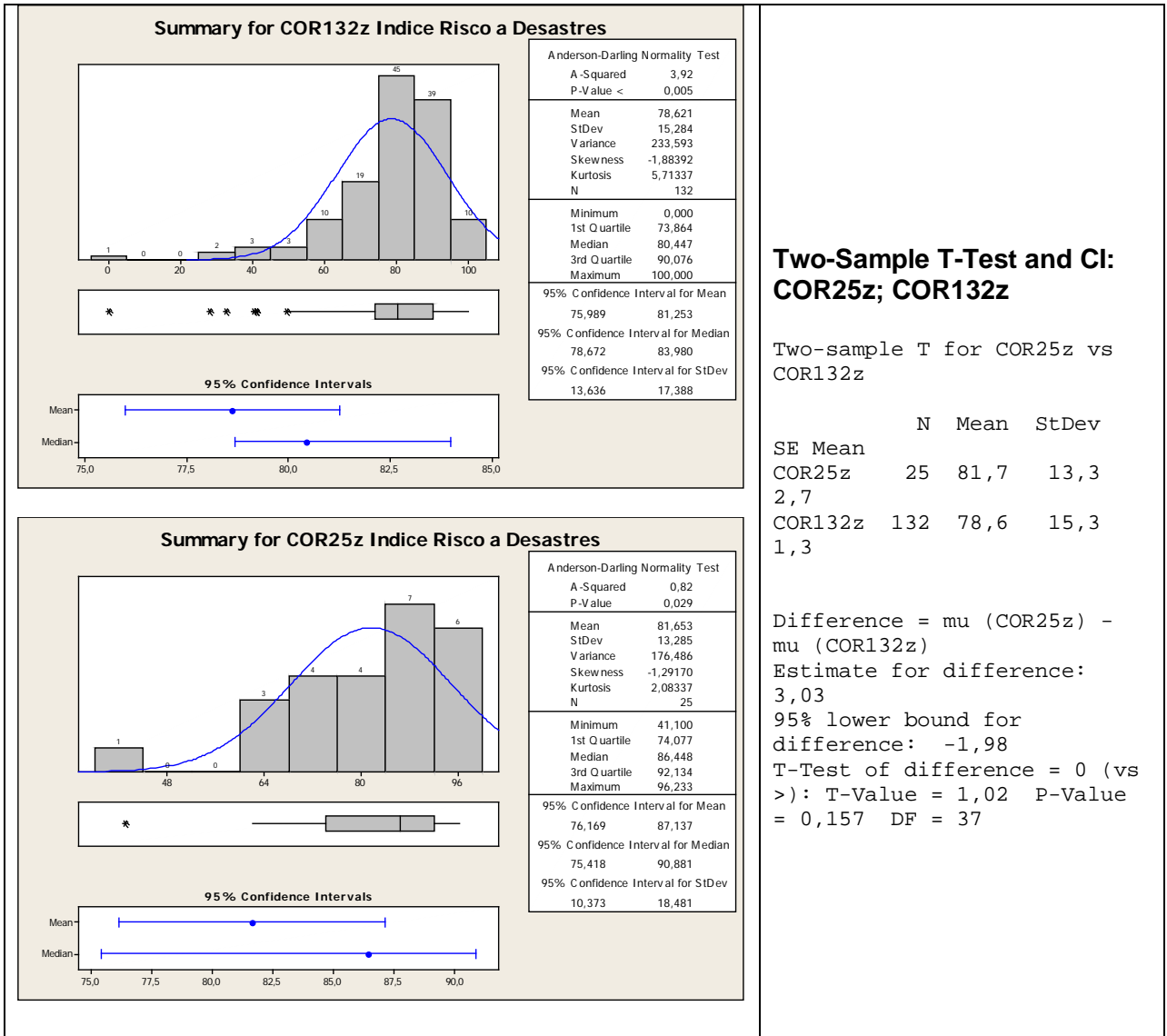
42,0    48,0    54,0    60,0

Pooled StDev = 24,17

A análise dos intervalos comparativamente mostra médias muito próximas e grande sobreposição, o que se confirma pelo baixo valor de F.

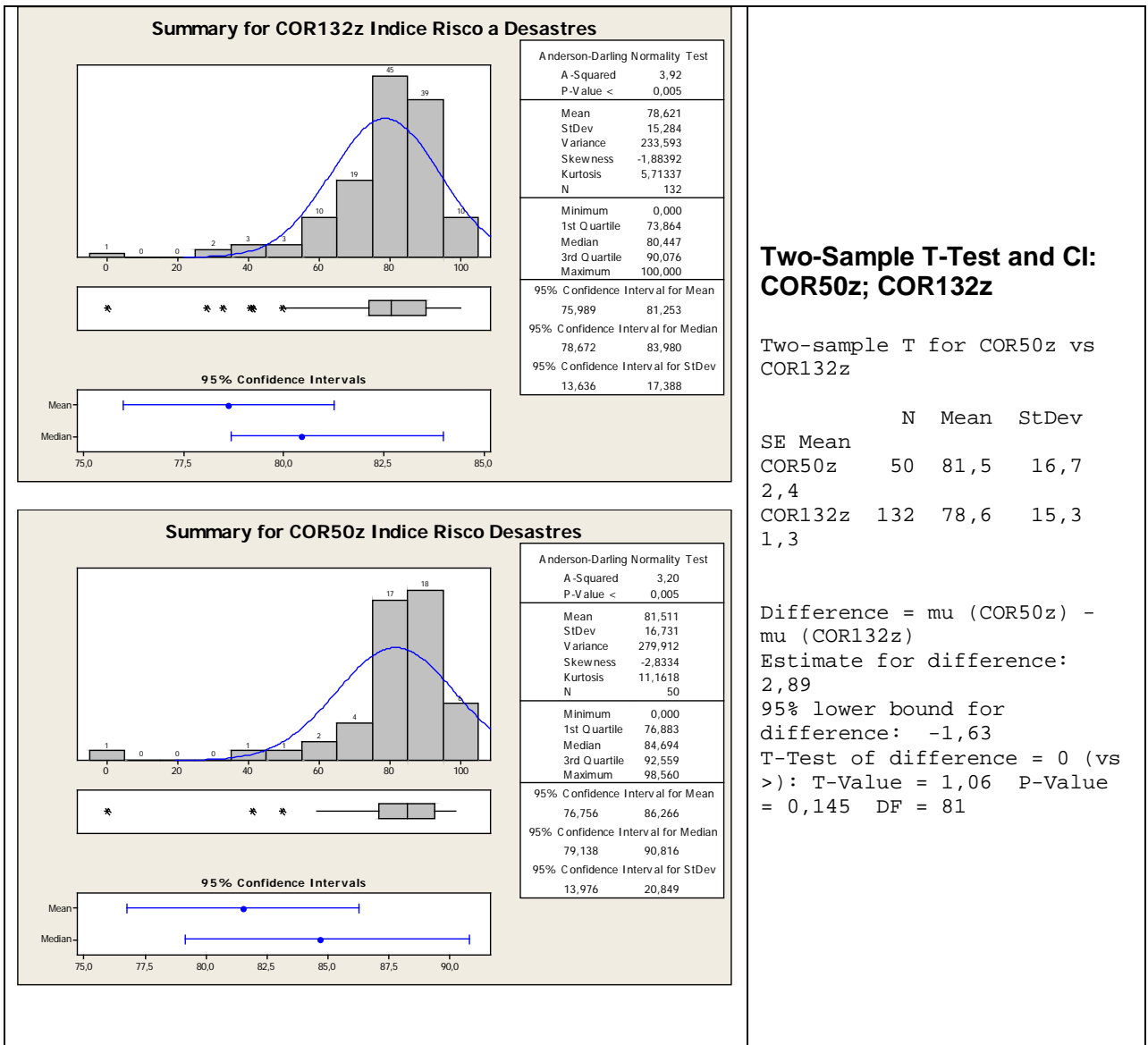
### 8.1.3 Variável Índice de Risco a Desastres

#### 8.1.3.1 Índice de Risco a Desastres – Amostra com 25 Indivíduos



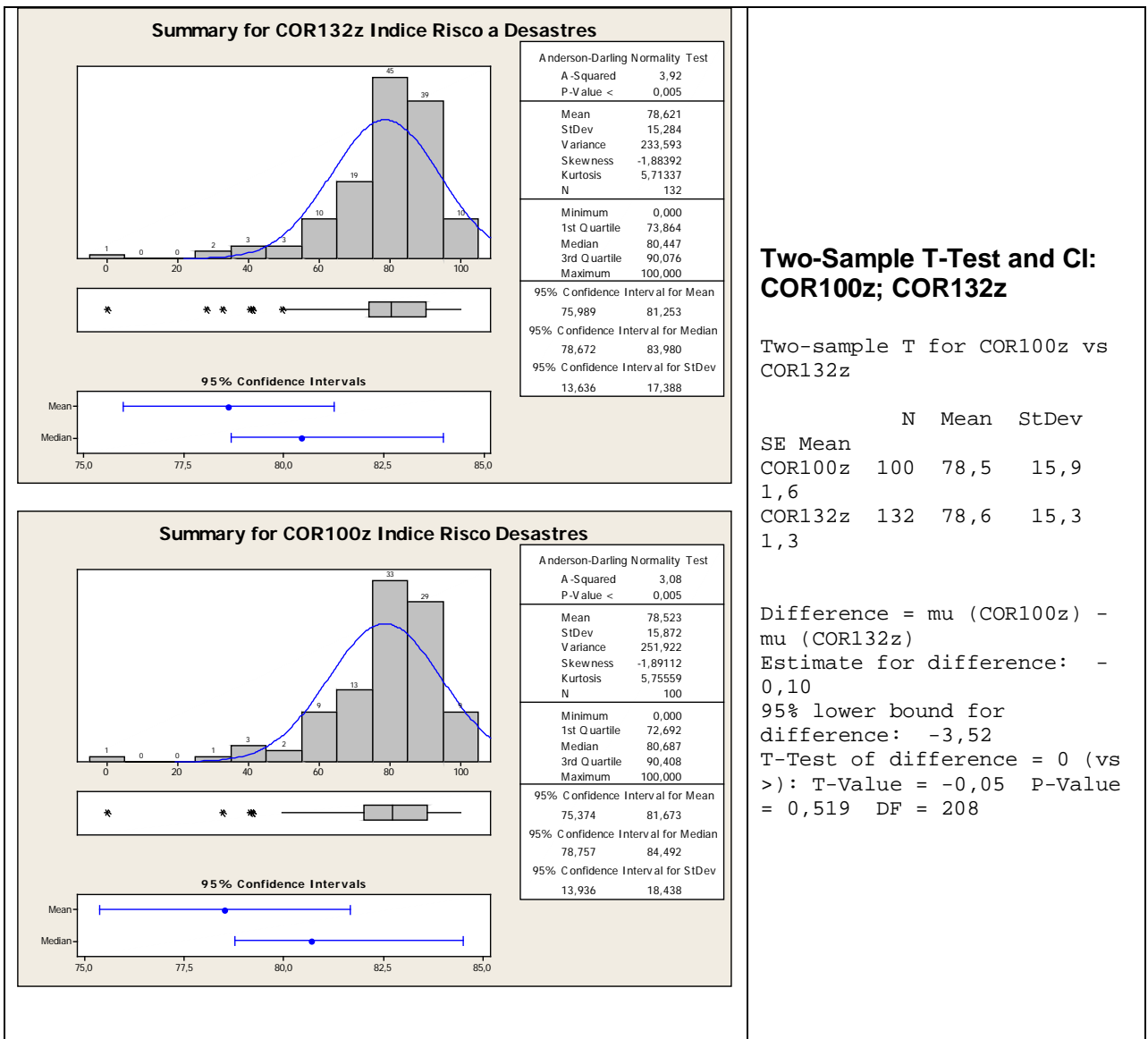


8.1.3.2 Índice de Risco a Desastres – Amostra com 50 Indivíduos



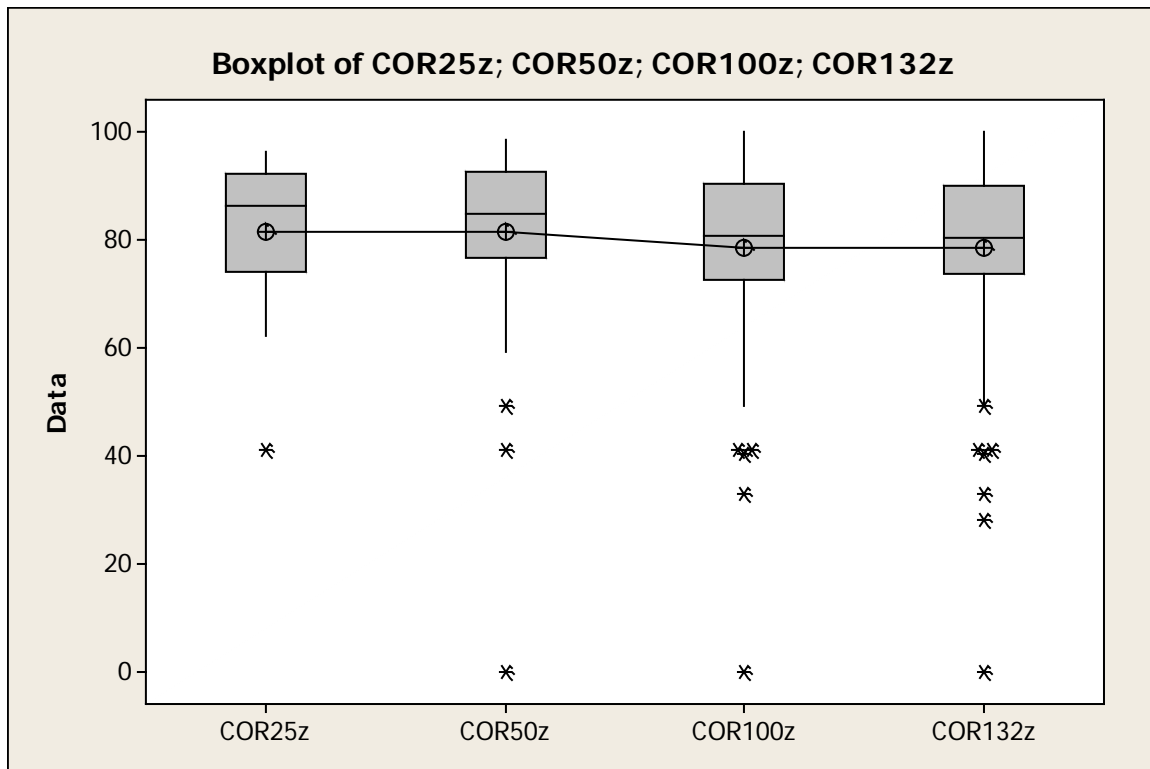
No caso de uma amostragem aleatória com 50 indivíduos comparada à população total de 132 elementos, também as médias são pouco diferentes (81 e 78 respectivamente), aliás a mesma da população aleatória de 25. Bangladesh, Chade e Guatemala aparecem no grupo dos outliers.

8.1.3.3 Índice de Risco a Desastres – Amostra com 100 Indivíduos



A amostragem aleatória de 100 indivíduos, comparativamente à população total, apresentou médias praticamente iguais (78). Chile, Egito, Azerbaijão e Coreia são os outliers dessa amostra.

#### 8.1.3.4. Índice de Risco a Desastres – Boxplot de Amostras Aleatórias com 25, 50, 100 Indivíduos e Amostra Completa



**One-way ANOVA: COR25z; COR50z; COR100z; COR132z**

Source	DF	SS	MS	F	P
Factor	3	504	168	0,69	0,557
Error	303	73492	243		
Total	306	73996			

S = 15,57    R-Sq = 0,68%    R-Sq(adj) = 0,00%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
COR25z	25	81,65	13,28	(-----*-----)
COR50z	50	81,51	16,73	(-----*-----)
COR100z	100	78,52	15,87	(-----*-----)
COR132z	132	78,62	15,28	(-----*-----)

-----+-----+-----+-----+-----  
77,0      80,5      84,0      87,5

Pooled StDev = 15,57

Analisando os intervalos das médias, percebe-se haver pouca diferença entre eles, portanto o tamanho da amostra, neste caso específico, não foi estatisticamente significativo para a análise da variável, o que se confirma pela sobreposição os valores das médias e pelo baixo valor de F. A única diferença evidente na análise do gráfico boxplot foi no número de outliers, que aumentou à medida que aumentou o tamanho da população.

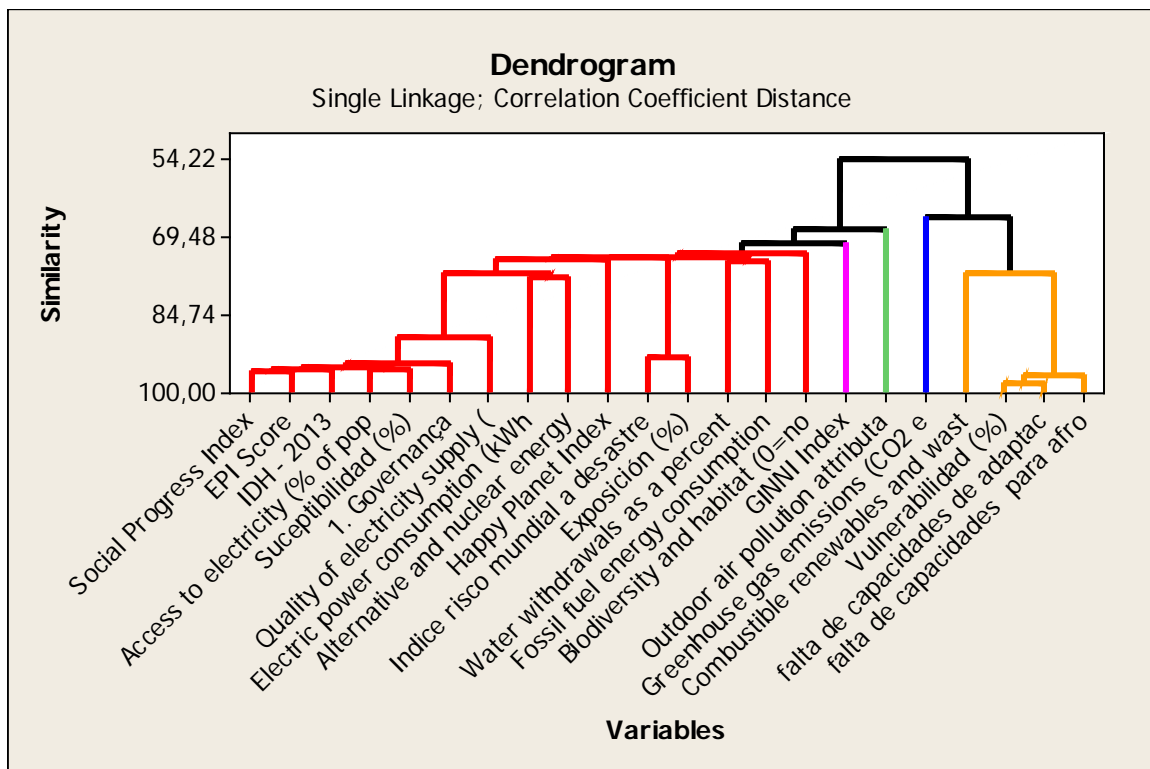
À medida que o tamanho da amostra aleatória aumenta, o valor da análise exploratória de dados também aumenta. No entanto, apesar do acréscimo de indivíduos na amostra trazer mais representatividade na amostra, é preciso comparar se este acréscimo é acompanhado de um acréscimo significativo de representatividade da população.

Pelo teste-t de diferença de médias foi possível observar que até mesmo amostras aleatórias de 25 indivíduos tiveram baixo P-Value, mostrando que tais amostras tiveram representatividade da população, o que se confirma quando se observam as médias comparativamente, revelando não haver diferenças significativas, o que nos leva a concluir que uma amostra aleatória de 25 elementos seria suficiente para realizar a análise nesse caso específico de 132 elementos.

Por fim, quando levamos em consideração que, em uma pesquisa de intenções de voto, são entrevistados aproximadamente 2 mil pessoas em uma população de mais de 140 milhões de eleitores, isso parece mostrar que 25 elementos de um universo de 132 é um número bastante significativo, o que de fato se confirma pelas análises acima.

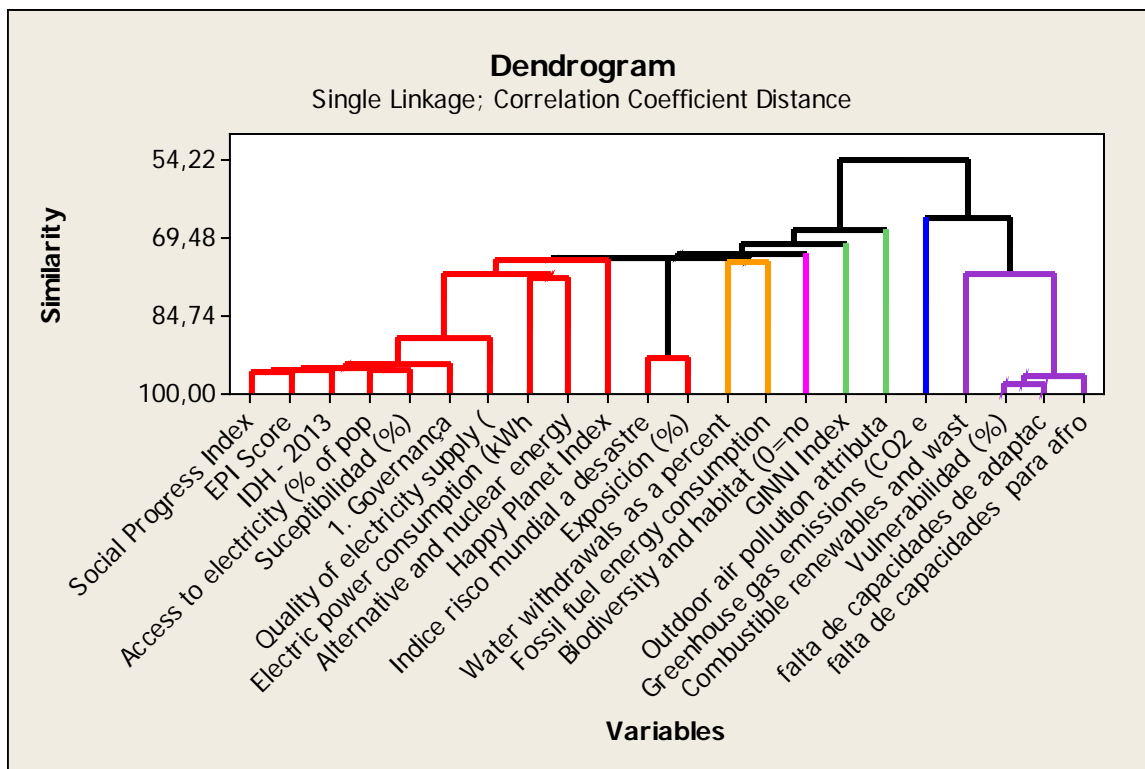
## 9. CORRELAÇÕES E DENDROGRAMAS

Seguem as matrizes de correlações e os dendrogramas entre as todas as variáveis:

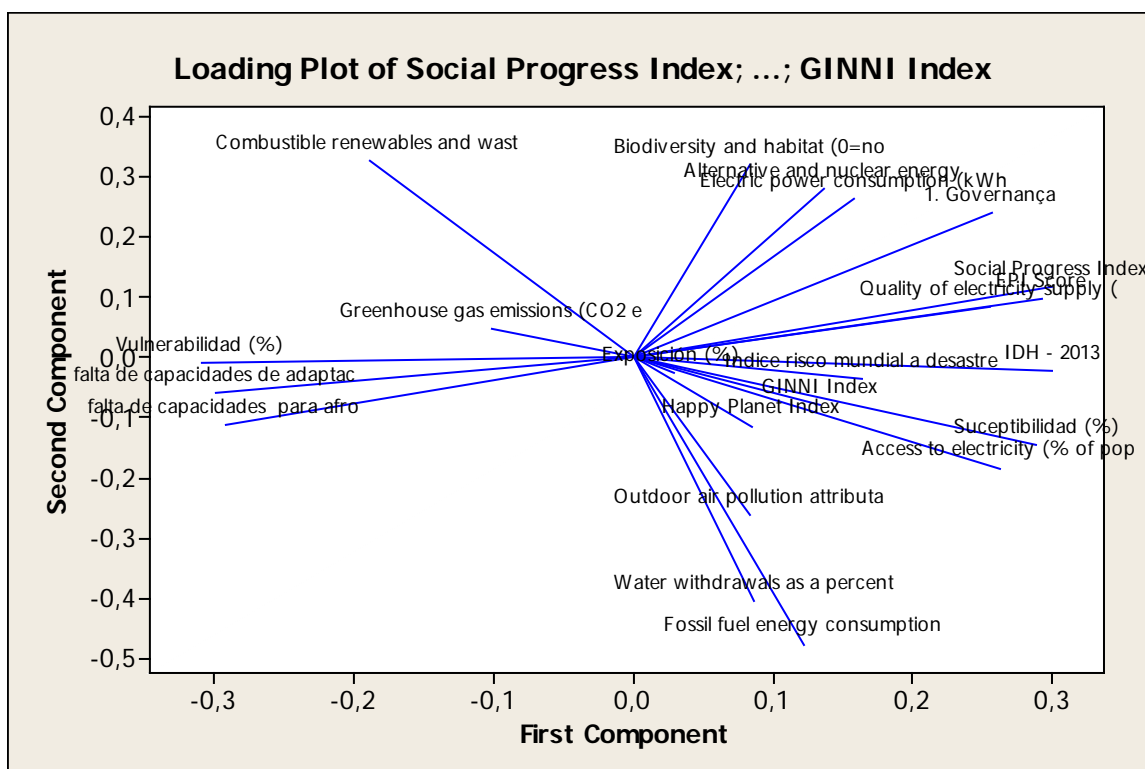


Analisando-se da esquerda para a direita, nota-se que o há uma grande similaridade entre as variáveis sintéticas como IDH, Social Progress Index, EPI Score e Governança, o que me parece lógico tendo em vista que tais índices são derivados de uma série de outros indicadores. O segundo grupo, que vem a seguir, inclui um grande número de variáveis. Em seguida aparecem variáveis isoladas ou pouco associadas com as demais.

Assim, decidimos aumentar os clusters de 5 para 8, para ter melhor agrupamento. Segue o outro dendrograma:

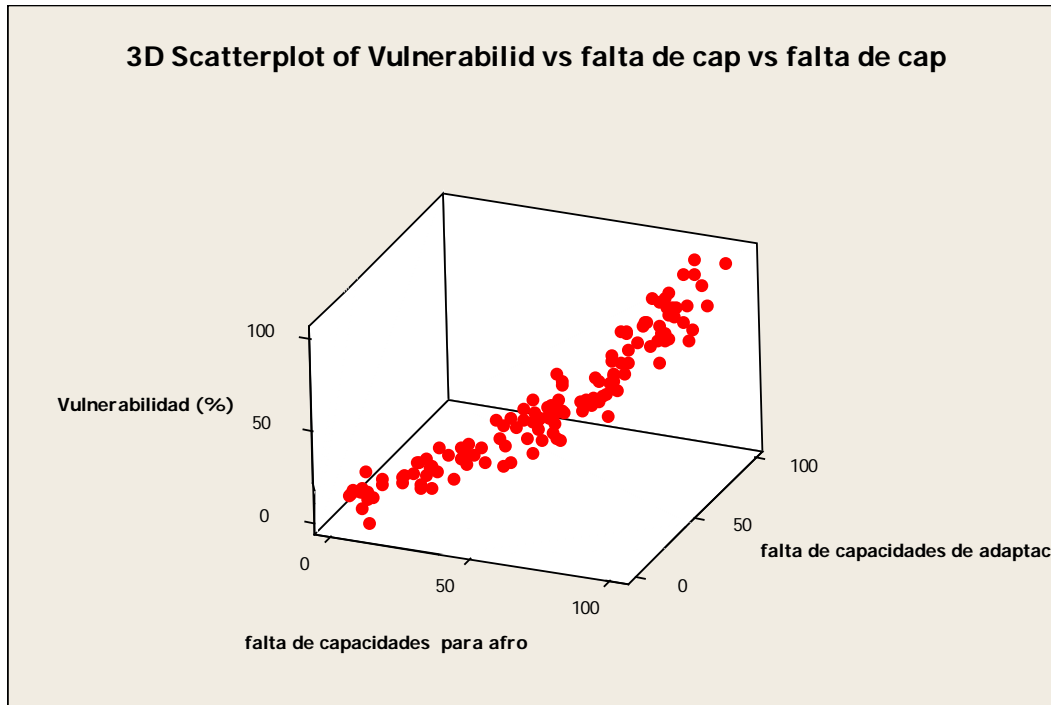


Outra forma de ver os agrupamentos de variáveis e as relações entre seus principais componentes é pelo gráfico *loading plot* abaixo:



Vemos os agrupamentos de forma condensada em algumas variáveis, por exemplo, **“vulnerabilidade”, “falta de capacidade de adaptação” e “falta de capacidade de enfrentar situações difíceis”**, que, por definição guardam um importante grau de relação

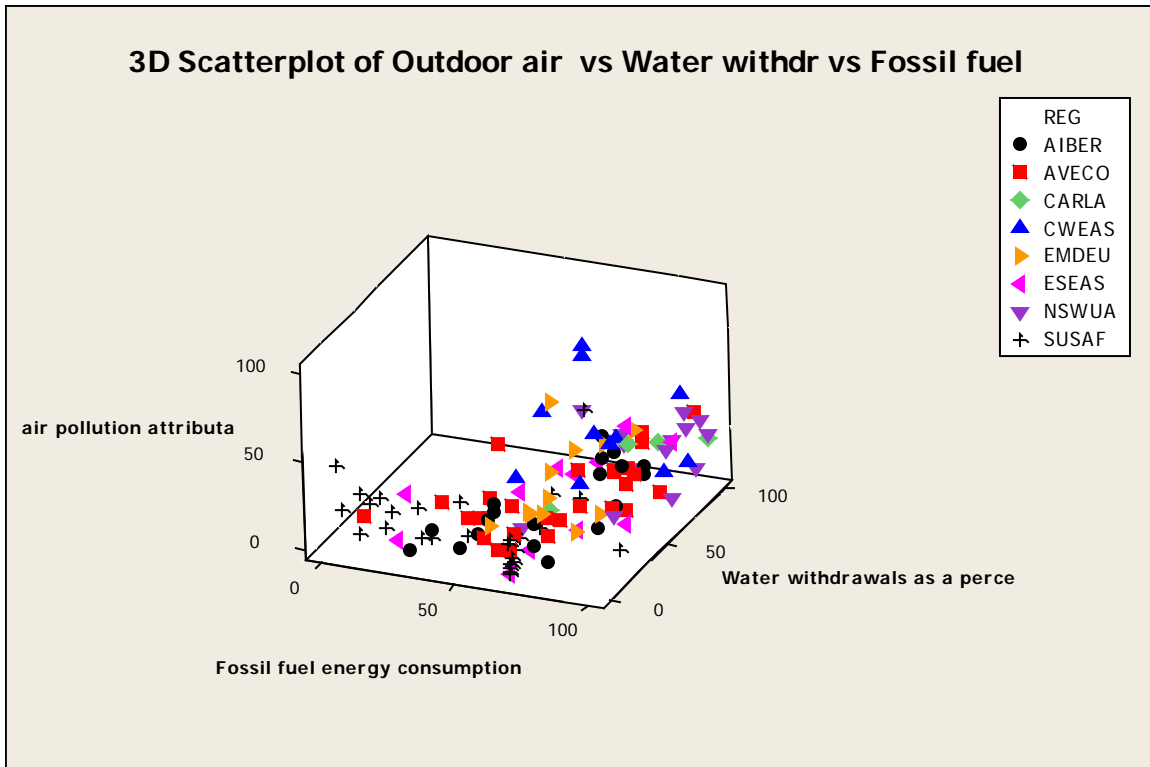
entre si. O mesmo ocorre com as variáveis “mortes atribuídas à poluição atmosférica”, “consumo de água” “consumo de energias fosseis”. Iremos realizar o teste de correlação com ambos os grupos.



**Correlations: Vulnerabilidad  
; falta de capacid adap; falta de capacid confront**

	Vulnerabilidad (	falta de capacid
falta de capacid	0,965 0,000	
falta de capacid	0,938 0,000	0,897 0,000

Cell Contents: Pearson correlation  
P-Value



**Correlations: Outdoor air poll; Water withdrawal; Fossil fuel ener**

	Outdoor air poll	Water withdrawal
Water withdrawal	0,143 0,102	
Fossil fuel ener	0,289 0,001	0,486 0,000

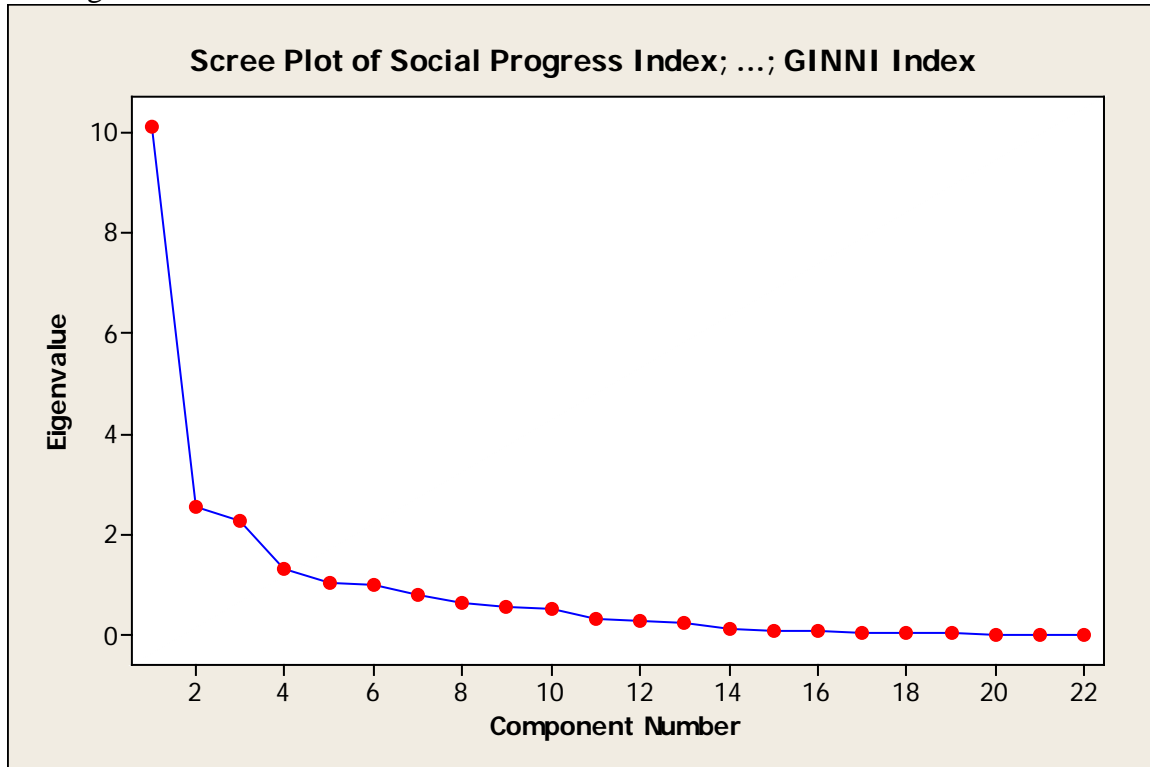
Cell Contents: Pearson correlation  
P-Value

Comparando visualmente os dois gráficos, bem como os números de correlação, notamos que o primeiro agrupamento é mais correlato, já que apresenta índices entre 0,897 e 0,965, enquanto a segunda apresenta um agrupamento mais frouxo, visto que mostras índices que vão de 0,143 a 0,486.

## 10. ANÁLISE DOS COMPONENTES PRINCIPAIS

O objetivo deste tópico é, através da análise dos componentes principais, tentarmos reduzir o número de variáveis, ou seja, percebermos as relações entre as variáveis e a possibilidade de agruparmos as mesmas. Certamente a análise de correlações e dendogramas acima já nos dão uma idéia de que a possibilidade de agrupamento é grande pelos elevados índices de correlação entre todas as variáveis:

Segue abaixo o resultado das análises dos componentes principais juntamente com o gráfico de “EigenValue”:



### Principal Component Analysis: Social Progr; Access to el; Quality of e; Outdoor

Eigenanalysis of the Correlation Matrix

Eigenvalue	10,098	2,541	2,260	1,304	1,042	0,990	0,805	0,646	0,567
Proportion	0,459	0,116	0,103	0,059	0,047	0,045	0,037	0,029	0,026
Cumulative	0,459	0,574	0,677	0,737	0,784	0,829	0,865	0,895	0,921

Eigenvalue	0,518	0,315	0,292	0,239	0,111	0,066	0,058	0,051	0,036
Proportion	0,024	0,014	0,013	0,011	0,005	0,003	0,003	0,002	0,002
Cumulative	0,944	0,958	0,972	0,983	0,988	0,991	0,993	0,996	0,997

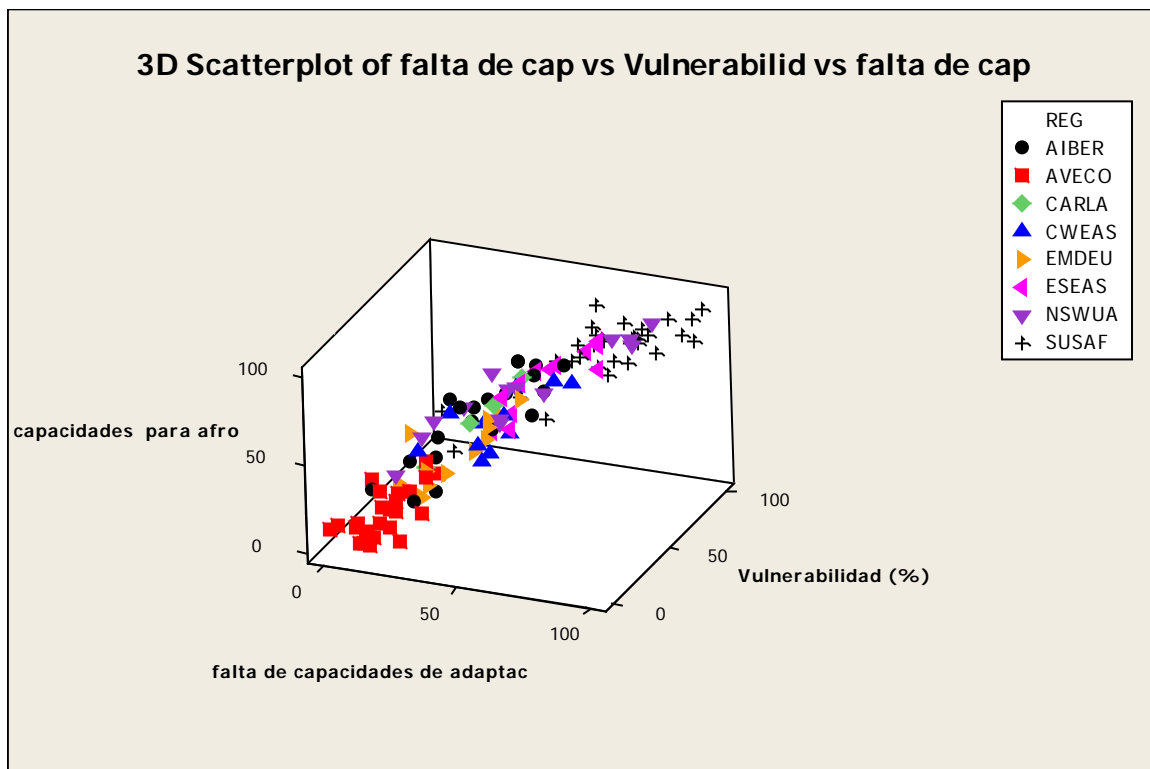
Eigenvalue	0,022	0,018	0,016	0,007
Proportion	0,001	0,001	0,001	0,000
Cumulative	0,998	0,999	1,000	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6
Social Progress Index	0,300	0,116	0,101	0,046	0,051	0,022
Access to electricity (% of pop	0,263	-0,184	0,137	-0,095	0,095	-0,036
Quality of electricity supply (	0,255	0,084	-0,003	0,038	-0,052	0,098
Outdoor air pollution attributa	0,084	-0,264	-0,193	-0,267	0,201	0,441
Greenhouse gas emissions (CO2 e	-0,102	0,046	-0,198	0,347	-0,448	0,292
Water withdrawals as a percent	0,087	-0,406	0,101	0,069	-0,083	-0,329
Biodiversity and habitat (0=no	0,084	0,320	0,042	0,507	0,167	0,042
IDH - 2013	0,301	-0,024	0,058	0,011	0,059	-0,080
Electric power consumption (kWh	0,158	0,264	-0,187	-0,112	-0,477	-0,118
Alternative and nuclear energy	0,137	0,278	0,059	-0,426	-0,415	-0,149



Combustible renewables and wast	-0,190	0,325	0,063	-0,131	0,459	-0,016
Fossil fuel energy consumption	0,122	-0,477	-0,056	0,310	-0,172	0,098
Indice risco mundial a desastre	0,164	-0,036	-0,502	0,016	0,156	-0,310
Exposición (%)	0,028	-0,025	-0,580	-0,030	0,136	-0,411
Vulnerabilidad (%)	-0,310	-0,010	-0,019	-0,015	-0,071	-0,036
Suceptibilidad (%)	0,289	-0,145	0,054	-0,038	0,044	0,020
falta de capacidades para afro	-0,292	-0,112	0,052	-0,009	-0,039	-0,123
falta de capacidades de adaptac	-0,299	-0,059	-0,033	-0,083	-0,084	0,012
1. Governança	0,257	0,240	0,024	0,078	0,035	0,053
EPI Score	0,293	0,097	-0,001	0,150	0,068	0,073
Happy Planet Index	0,085	-0,115	0,461	-0,113	-0,044	-0,338
GINNI Index	0,134	-0,079	-0,158	-0,420	0,004	0,378

Pela análise dos detalhes e gráfico acima percebemos a ocorrência de 5 componentes com Eigenvalue acima de 1, os quais, juntos, perfazem 78%.



A análise scatterplot dessas três variáveis por região mostra os melhores números do grupo da AVECO (quadrados vermelhos) com baixos índices negativos, enquanto no outro extremo aparecem os países da SUSAF (cruzes pretas). Já os países da AIBER (losangos verdes) encontram-se numa posição intermediária.

De acordo com todas as análises acima, percebemos claramente que o agrupamento de variáveis é bastante pertinente no caso das variáveis sintéticas, bem como de algumas variáveis específicas que apresentam maior similaridade. Isto pôde ser observado inicialmente pelas matrizes de correlação e dendrogramas e depois comprovados pela análise dos componentes principais. Foi bastante ilustrativo “fugir” das variáveis sintéticas – posto que, no fim, acabam mostrando similaridade por serem compostas por vários outros índices – e buscar relações entre variáveis mais específicas do meu tema, pois essas são as que realmente podem ajudar a explicar os problemas, bem com suas possíveis soluções, em maior detalhe.

## 11. ANÁLISE DISCRIMINANTE

### Discriminant Analysis: Clusters\_1 versus Outdoor air ; Combustible ; ...

Linear Method for Response: Clusters\_1

Predictors: Outdoor air pollution attributa; Combustible renewables and wast;  
Fossil fuel energy consumption

Group	1	2	3
Count	41	39	52

Summary of classification

	True Group		
Put into Group	1	2	3
1	28	1	8
2	8	19	11
3	5	19	33
Total N	41	39	52
N correct	28	19	33
Proportion	0,683	0,487	0,635

N = 132

N Correct = 80

Proportion Correct = 0,606

Squared Distance Between Groups

	1	2	3
1	0,00000	2,62113	2,41230
2	2,62113	0,00000	0,67399
3	2,41230	0,67399	0,00000

Linear Discriminant Function for Groups

	1	2	3
Constant	-26,595	-21,006	-26,267
Outdoor air pollution attributa	0,005	0,028	0,049
Combustible renewables and wast	0,594	0,480	0,536
Fossil fuel energy consumption	0,562	0,511	0,566

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	3	2	1	4,650	0,130
			2	1,339	0,680
			3	3,891	0,190
9**	3	1	1	0,5947	0,428
			2	1,0960	0,333
			3	1,7535	0,240
12**	3	2	1	1,8584	0,210
			2	0,4021	0,435
			3	0,8073	0,355
15**	3	2	1	3,204	0,203
			2	1,747	0,421
			3	1,978	0,375
17**	3	2	1	5,308	0,179
			2	2,642	0,679
			3	5,766	0,142
19**	3	1	1	1,974	0,937
			2	9,003	0,028
			3	8,558	0,035

20**	3	2	1	0,9304	0,314
			2	0,6597	0,360
			3	0,8605	0,326
21**	3	1	1	0,2308	0,771
			2	4,3332	0,099
			3	3,8027	0,129
22**	3	1	1	0,3968	0,434
			2	1,3654	0,268
			3	1,1521	0,298
27**	3	2	1	6,284	0,060
			2	1,293	0,724
			3	3,713	0,216
28**	3	2	1	6,544	0,058
			2	1,459	0,735
			3	3,998	0,207
34**	3	2	1	1,7911	0,229
			2	0,3569	0,469
			3	1,2438	0,301
36**	3	1	1	0,2364	0,744
			2	3,6493	0,135
			3	3,8676	0,121
37**	3	2	1	2,255	0,285
			2	1,653	0,386
			3	1,973	0,329
38**	3	1	1	1,411	0,748
			2	4,382	0,169
			3	5,812	0,083
40**	3	2	1	2,3056	0,213
			2	0,4695	0,533
			3	1,9565	0,254
45**	3	1	1	0,6784	0,846
			2	5,6863	0,069
			3	5,2885	0,084
46**	3	1	1	1,873	0,419
			2	2,953	0,244
			3	2,304	0,338
52**	3	2	1	3,688	0,121
			2	1,036	0,457
			3	1,201	0,421
54**	2	3	1	1,5641	0,237
			2	1,1503	0,291
			3	0,1838	0,472
56**	2	3	1	4,2268	0,076
			2	0,8849	0,402
			3	0,3619	0,522
58**	2	3	1	2,882	0,254
			2	3,198	0,217
			3	1,409	0,530
59**	2	3	1	5,6696	0,062
			2	2,7302	0,272
			3	0,9354	0,666
61**	2	3	1	5,4796	0,062
			2	2,3699	0,293
			3	0,7921	0,645
62**	2	3	1	2,8064	0,135
			2	0,6626	0,395
			3	0,3211	0,469
63**	2	3	1	3,8849	0,094
			2	1,4055	0,324
			3	0,2305	0,583
64**	2	3	1	9,555	0,112
			2	8,379	0,202
			3	5,938	0,685
65**	2	3	1	2,2715	0,153
			2	0,3762	0,395
			3	0,1071	0,452
66**	2	3	1	4,048	0,213
			2	3,182	0,328
			3	2,513	0,459
67**	2	3	1	1,0769	0,264

			2	0,4657	0,359
			3	0,3651	0,377
68**	2	3	1	3,2866	0,166
			2	2,2797	0,275
			3	0,8588	0,559
69**	2	3	1	3,0162	0,132
			2	1,1199	0,339
			3	0,2328	0,529
74**	2	3	1	3,2642	0,103
			2	0,3609	0,440
			3	0,2810	0,457
75**	2	1	1	0,8815	0,369
			2	1,2531	0,306
			3	1,1311	0,325
76**	2	3	1	4,173	0,195
			2	3,023	0,347
			3	2,471	0,458
79**	2	3	1	1,1158	0,297
			2	1,2199	0,282
			3	0,4117	0,422
85**	2	3	1	3,5912	0,104
			2	0,9178	0,396
			3	0,4520	0,500
86**	2	3	1	3,4758	0,137
			2	1,7408	0,326
			3	0,7457	0,537
91**	2	3	1	3,7177	0,087
			2	0,6152	0,410
			3	0,2066	0,503
92**	1	2	1	1,0153	0,348
			2	0,9596	0,358
			3	1,3506	0,294
93**	1	2	1	0,9502	0,293
			2	0,5361	0,360
			3	0,6099	0,347
94**	1	2	1	2,4563	0,230
			2	0,7830	0,532
			3	2,3887	0,238
95**	1	3	1	2,5285	0,199
			2	1,7116	0,300
			3	0,6814	0,501
103**	1	2	1	0,9304	0,314
			2	0,6597	0,360
			3	0,8605	0,326
105**	1	3	1	4,841	0,118
			2	2,922	0,307
			3	1,664	0,576
107**	1	2	1	0,9829	0,282
			2	0,4971	0,360
			3	0,5076	0,358
109**	1	3	1	5,0777	0,073
			2	2,2676	0,299
			3	0,7870	0,627
113**	1	2	1	0,9717	0,337
			2	0,8444	0,359
			3	1,1720	0,305
114**	1	3	1	5,154	0,197
			2	4,066	0,339
			3	3,437	0,464
117**	1	2	1	0,9829	0,282
			2	0,4971	0,360
			3	0,5076	0,358
127**	1	2	1	18,28	0,025
			2	11,04	0,928
			3	17,02	0,047
131**	1	3	1	4,860	0,182
			2	3,653	0,333
			3	2,900	0,485

Esse grupo de variáveis apresenta baixa similaridade, já que, de 132 países apenas 80 foram corretamente agrupados, sendo necessários muitos ajustes para que possam ser vistas como relevantes para análise juntas.

### Discriminant Analysis: Clusters\_1 versus Alternative ; Fossil fuel ; ...

Linear Method for Response: Clusters\_1

Predictors: Alternative and nuclear energy; Fossil fuel energy consumption;  
Access to electricity (% of pop)

Group	1	2	3
Count	41	39	52

Summary of classification

Put into Group	True Group		
	1	2	3
1	35	0	0
2	1	18	10
3	5	21	42
Total N	41	39	52
N correct	35	18	42
Proportion	0,854	0,462	0,808

N = 132

N Correct = 95

Proportion Correct = 0,720

Squared Distance Between Groups

	1	2	3
1	0,0000	16,6555	14,5131
2	16,6555	0,0000	0,5411
3	14,5131	0,5411	0,0000

Linear Discriminant Function for Groups

	1	2	3
Constant	-5,314	-23,796	-22,196
Alternative and nuclear energy	0,099	0,156	0,109
Fossil fuel energy consumption	0,133	0,173	0,174
Access to electricity (% of pop)	0,094	0,324	0,314

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	3	2	1	16,3794	0,000
			2	0,3898	0,584
			3	1,0691	0,416
4**	3	2	1	21,327	0,000
			2	2,344	0,767
			3	4,732	0,233
9**	3	2	1	15,821	0,000
			2	1,024	0,560
			3	1,509	0,440
10**	3	2	1	18,9172	0,000
			2	0,7393	0,676
			3	2,2132	0,324
12**	3	2	1	15,3860	0,000
			2	0,0699	0,523
			3	0,2575	0,476
17**	3	2	1	16,442	0,001

			2	2,973	0,776
			3	5,461	0,224
28**	3	2	1	24,108	0,000
			2	4,458	0,830
			3	7,623	0,170
40**	3	2	1	10,333	0,008
			2	1,714	0,619
			3	2,730	0,372
41**	3	2	1	16,9118	0,000
			2	0,0980	0,519
			3	0,2542	0,480
50**	3	2	1	18,3127	0,000
			2	0,3908	0,615
			3	1,3287	0,385
54**	2	3	1	16,6201	0,000
			2	0,8311	0,422
			3	0,2029	0,578
56**	2	3	1	16,9610	0,000
			2	0,8367	0,414
			3	0,1438	0,586
58**	2	3	1	17,8377	0,000
			2	1,6531	0,356
			3	0,4681	0,644
59**	2	3	1	17,9498	0,000
			2	1,2626	0,392
			3	0,3814	0,608
61**	2	3	1	18,8604	0,000
			2	1,4918	0,411
			3	0,7703	0,589
62**	2	3	1	17,2980	0,000
			2	0,6734	0,435
			3	0,1517	0,565
63**	2	3	1	14,4730	0,000
			2	0,9900	0,409
			3	0,2573	0,590
64**	2	3	1	16,404	0,000
			2	1,658	0,428
			3	1,083	0,571
65**	2	3	1	16,8753	0,000
			2	1,8900	0,365
			3	0,7833	0,635
66**	2	3	1	17,9663	0,000
			2	1,7926	0,356
			3	0,6095	0,644
67**	2	3	1	16,897	0,000
			2	2,312	0,363
			3	1,190	0,636
68**	2	3	1	18,1287	0,000
			2	1,7476	0,354
			3	0,5469	0,646
69**	2	3	1	16,5636	0,000
			2	0,3564	0,471
			3	0,1227	0,529
73**	2	3	1	14,018	0,001
			2	2,440	0,336
			3	1,079	0,663
76**	2	3	1	18,4929	0,000
			2	1,6764	0,371
			3	0,6250	0,628
77**	2	3	1	16,2267	0,000
			2	0,4859	0,487
			3	0,3795	0,513
79**	2	3	1	16,5623	0,000
			2	1,0153	0,414
			3	0,3180	0,586
84**	2	3	1	17,5359	0,000
			2	1,1364	0,392
			3	0,2626	0,607
85**	2	3	1	17,9719	0,000
			2	0,7471	0,457

			3	0,4012	0,543
86**	2	3	1	18,1427	0,000
			2	1,5127	0,374
			3	0,4843	0,626
91**	2	3	1	17,6609	0,000
			2	0,4226	0,492
			3	0,3598	0,508
95**	1	3	1	10,2857	0,006
			2	2,1319	0,340
			3	0,8200	0,655
108**	1	3	1	6,323	0,065
			2	3,206	0,307
			3	1,774	0,628
109**	1	3	1	17,9370	0,000
			2	1,9867	0,354
			3	0,7819	0,646
111**	1	3	1	3,843	0,399
			2	5,532	0,171
			3	3,691	0,430
122**	1	3	1	12,215	0,003
			2	2,188	0,387
			3	1,277	0,610
127**	1	2	1	31,88	0,000
			2	11,90	0,932
			3	17,14	0,068

Esse grupo de variáveis apresenta ligeira melhora em relação ao grupo anterior, haja vista ter havido um índice maior de acertos, 95 países de 132, ou seja, 72% corretos. No entanto, ainda não me parece adequado para analisar o agrupamento realizado.

### Discriminant Analysis: Clusters\_1 versus IDH - 2013; Social Progr; ...

Linear Method for Response: Clusters\_1

Predictors: IDH - 2013; Social Progress Index; Access to electricity (% of pop

Group	1	2	3
Count	41	39	52

Summary of classification

Put into Group	True Group		
	1	2	3
1	38	0	0
2	0	37	1
3	3	2	51
Total N	41	39	52
N correct	38	37	51
Proportion	0,927	0,949	0,981

N = 132

N Correct = 126

Proportion Correct = 0,955

Squared Distance Between Groups

	1	2	3
1	0,0000	40,7538	17,6201
2	40,7538	0,0000	9,3691
3	17,6201	9,3691	0,0000

Linear Discriminant Function for Groups

1 2 3

Constant	-4,249	-43,103	-24,706
IDH - 2013	0,026	0,126	0,087
Social Progress Index	0,197	0,650	0,357
Access to electricity (% of pop	0,077	0,195	0,250

#### Summary of Misclassified Observations

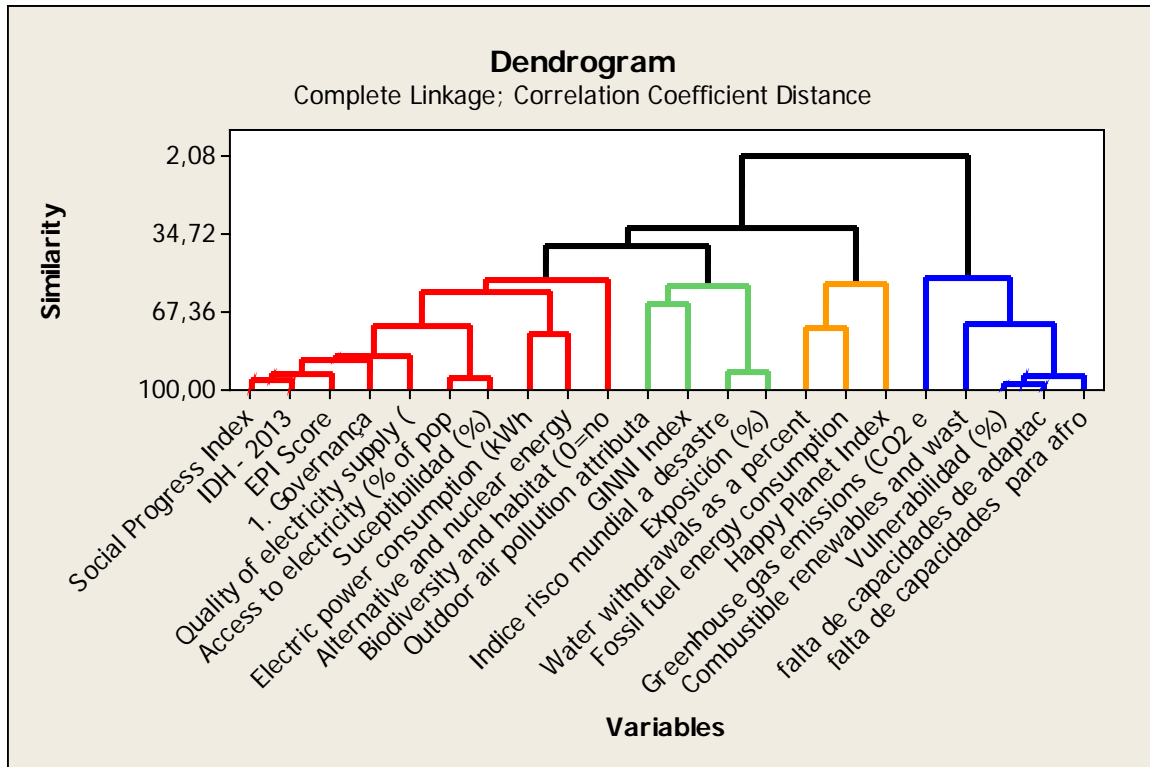
Observation	True Group	Pred Group	Group	Squared Distance	Probability
37**	3	2	1	24,178	0,000
			2	2,262	0,634
			3	3,362	0,366
66**	2	3	1	24,623	0,000
			2	3,955	0,200
			3	1,180	0,800
67**	2	3	1	32,496	0,000
			2	8,321	0,493
			3	8,261	0,507
95**	1	3	1	12,652	0,003
			2	11,116	0,007
			3	1,310	0,989
109**	1	3	1	19,97	0,048
			2	44,79	0,000
			3	13,98	0,952
127**	1	3	1	17,415	0,002
			2	24,636	0,000
			3	4,968	0,998

Dos 132 países, 126 estavam corretos em seus números referentes aos três grupos que foram criados anteriormente - no caso 95,5% estão exatos. Observa-se que nenhum dos grupos apresenta muita proximidade em seus índices, mostrando não haver grau de similaridade relevante. O agrupamento em três clusters de países mostra boa distribuição e a existência de duas extremidades – uma positiva e outra negativa – e um grupo mais central. Desse grupo central faz parte a maioria dos países ibero-americanos, embora haja algumas exceções no grupo mais bem posicionado, geralmente Chile, Uruguai e Costa Rica, além da Portugal e Espanha. Fica claro que tal divisão funciona bem para explicar algumas combinações de variáveis, enquanto outras não, motivo pelo qual outros testes e agrupamentos, tanto de países quanto de variáveis, poderiam ser feitos para comparar os distintos resultados.

## 12. REGRESSÃO LOGÍSTICA

Esse estudo tem como objetivo efetuar uma Regressão Logística usando variáveis combinadas para 132 países. A variável dependente de nossa análise serão os 132 países e para tentar explicar em qual grupo um determinado país se encaixa, utilizamos variáveis selecionadas conforme agrupamento realizado via dendrograma.





Segue o resultado da regressão logística:

### Ordinal Logistic Regression: Clusters\_1 versus Outdoor air ; GINNI Index; ...

Link Function: Logit

#### Response Information

Variable	Value	Count
Clusters_1	1	41
	2	39
	3	52
Total		132

#### Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	-0,634617	0,983388	-0,65	0,519	
Const(2)	0,893307	0,987518	0,90	0,366	
Outdoor air pollution attributa	-0,0427137	0,0119497	-3,57	0,000	0,96
GINNI Index	0,0062264	0,0087834	0,71	0,478	1,01
Indice risco mundial a desastre	-0,0827525	0,0272930	-3,03	0,002	0,92
Exposición (%)	0,0885882	0,0272498	3,25	0,001	1,09

Predictor	95% CI	
	Lower	Upper
Const(1)		
Const(2)		
Outdoor air pollution attributa	0,94	0,98
GINNI Index	0,99	1,02
Indice risco mundial a desastre	0,87	0,97
Exposición (%)	1,04	1,15

Log-Likelihood = -126,194

Test that all slopes are zero: G = 35,471, DF = 4, P-Value = 0,000

## Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	262,827	258	0,405
Deviance	252,388	258	0,587

## Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	4243	73,7	Somers' D 0,48
Discordant	1502	26,1	Goodman-Kruskal Gamma 0,48
Ties	14	0,2	Kendall's Tau-a 0,32
Total	5759	100,0	

Analisando o tratamento dos dados acima verifica-se o P próximo de 0 de significa que existe um grau de confiabilidade. Além disso é possível constatar que há uma concordância 73,7% nas variáveis tratadas, o que torna essa combinação relevante. Vamos ao próximo grupo de variáveis.

**Ordinal Logistic Regression: Clusters\_1 versus Water withdr; Fossil fuel ; ...**

Link Function: Logit

## Response Information

Variable	Value	Count
Clusters_1	1	41
	2	39
	3	52
Total		132

## Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	4,08674	0,812488	5,03	0,000	
Const(2)	5,86799	0,901801	6,51	0,000	
Water withdrawals as a percent	-0,0049725	0,0066646	-0,75	0,456	1,00
Fossil fuel energy consumption	-0,0389302	0,0098721	-3,94	0,000	0,96
Happy Planet Index	-0,0479800	0,0094626	-5,07	0,000	0,95

Predictor	95% CI	
	Lower	Upper
Const(1)		
Const(2)		
Water withdrawals as a percent	0,98	1,01
Fossil fuel energy consumption	0,94	0,98
Happy Planet Index	0,94	0,97

Log-Likelihood = -113,850

Test that all slopes are zero: G = 60,160, DF = 3, P-Value = 0,000

## Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	274,548	259	0,242
Deviance	227,700	259	0,920

Measures of Association:  
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	4502	78,2	Somers' D 0,57
Discordant	1237	21,5	Goodman-Kruskal Gamma 0,57
Ties	20	0,3	Kendall's Tau-a 0,38
Total	5759	100,0	

Analisando o tratamento dos dados acima verifica-se o P foi abaixo de 10% de significa que existe um grau de confiabilidade relativo. Além disso é possível constatar que há uma concordância de 78% nas variáveis tratadas. Vamos ao próximo grupo de variáveis:

### Ordinal Logistic Regression: Clusters\_1 versus Greenhouse g; Combustible ; ...

Link Function: Logit

Response Information

Variable	Value	Count
Clusters_1	1	41
	2	39
	3	52
Total		132

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	-2,63997	0,507122	-5,21	0,000	
Const(2)	-0,220646	0,421863	-0,52	0,601	
Greenhouse gas emissions (CO2 e	-0,0199518	0,0290245	-0,69	0,492	0,98
Combustible renewables and wast	0,0212867	0,0127544	1,67	0,095	1,02
Vulnerabilidad (%)	0,350411	0,0687332	5,10	0,000	1,42
falta de capacidades para afro	-0,224244	0,0359012	-6,25	0,000	0,80
falta de capacidades de adaptac	-0,0569235	0,0372410	-1,53	0,126	0,94

Predictor	95% CI	
	Lower	Upper
Const(1)		
Const(2)		
Greenhouse gas emissions (CO2 e	0,93	1,04
Combustible renewables and wast	1,00	1,05
Vulnerabilidad (%)	1,24	1,62
falta de capacidades para afro	0,74	0,86
falta de capacidades de adaptac	0,88	1,02

Log-Likelihood = -89,119

Test that all slopes are zero: G = 109,621, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	261,861	257	0,404
Deviance	178,239	257	1,000

Measures of Association:  
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	4956	86,1	Somers' D	0,72
Discordant	789	13,7	Goodman-Kruskal Gamma	0,73
Ties	14	0,2	Kendall's Tau-a	0,48
Total	5759	100,0		

Esse grupo de variáveis é o que apresenta o maior índice de concordância: 86,1%, mostrando haver uma boa combinação nesse conjunto em relação ao agrupamento de três países.

Vimos, por meio da regressão logística, a possibilidade de agrupamento entre certas variáveis devido ao grau de concordância entre elas. Isso nos ajuda a pensar em tais grupos conjuntamente, o que pode facilitar a análise e interpretação dos fenômenos, tendo em vista que guardam relação estatística relevante.

### 13. ÁRVORES DE CLASSIFICAÇÃO

Este aplicativo indica qual a variável que melhor separa os grupos e classifica as variáveis por ordem de importância na separação dos grupos. A seguir é demonstrado o teste deste modelo.

#### Classification Tree

##### Warnings

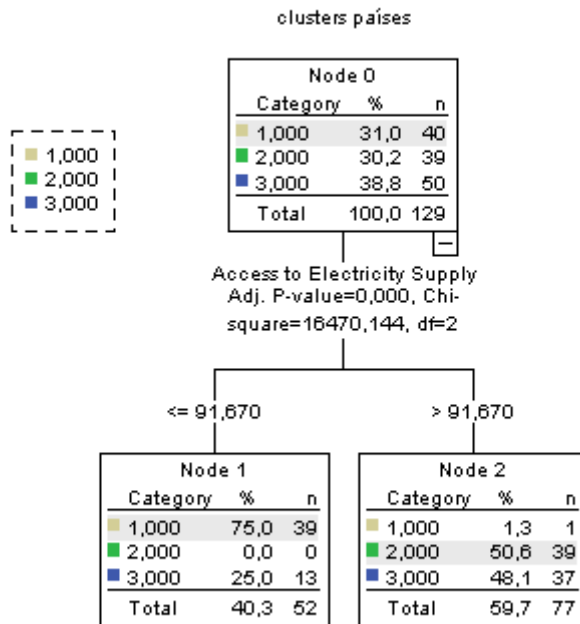
Gain summary Tables are not displayed because profits are undefined.

Target category gains tables are not displayed because target categories are undefined.

##### Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	clusters países
	Independent Variables	IDH, Access to Electricity Supply, Social Progress Index, Alternative and Nuclear Energy, Water Withdrawals, Greenhouse Gas Emissions
	Validation	None
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	100
	Minimum Cases in Child Node	50

Results	Independent Variables	Access to Electricity Supply
	Included	
	Number of Nodes	3
	Number of Terminal Nodes	2
	Depth	1



**Risk**

Estimate	Std. Error
,395	,043

Growing Method: CHAID

Dependent Variable:

clusters países

**Classification**

Observed	Predicted			
	1,00	2,00	3,00	Percent Correct
1,00	39	1	0	97,5%
2,00	0	39	0	100,0%
3,00	13	37	0	,0%
Overall Percentage	40,3%	59,7%	,0%	60,5%

Classification				
Observed	Predicted			
	1,00	2,00	3,00	Percent Correct
1,00	39	1	0	97,5%
2,00	0	39	0	100,0%
3,00	13	37	0	,0%
Overall Percentage	40,3%	59,7%	,0%	60,5%

Growing Method: CHAID

Dependent Variable: clusters países

A classificação em árvore não pode ser considerada significativa, uma vez que houve apenas 60% de acertos. Parte disso deve-se ao fato de haver um número relativamente pequeno de dados, razão pela qual para esta base o modelo não se revela adequado nem relevante.

### Nova Árvore de Classificação

Vimos que, em razão de termos uma base de dados numericamente baixa (132 países), é necessário rever os critérios para geração de uma árvore significativamente mais robusta: com isso, mudamos o número mínimo de casos de 100 para 10 no *parent node* e de 50 para 5 no *child node*; ademais, mudamos a profundidade máxima da árvore de 3 para 5. Por fim, alteramos o método de crescimento de CHAID para Exhaustive CHAID. Eis os resultados:

### Classification Tree

[DataSet1] C:\Users\jtfem\Desktop\material da aula Hoyos\_ Metodos\trabalhos Jorge\variaveis completas 8jun2015.sav

#### Warnings

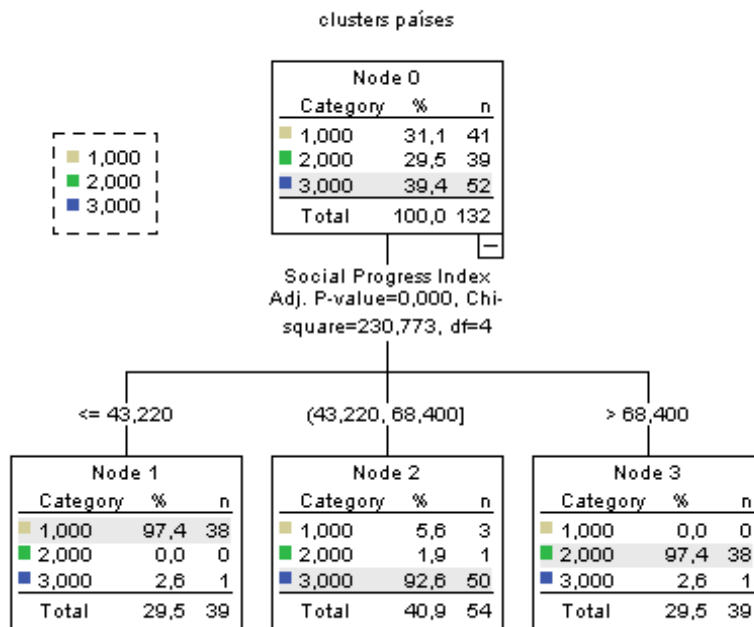
Gain summary Tables are not displayed because profits are undefined.

Target category gains tables are not displayed because target categories are undefined.

#### Model Summary

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	clusters países
	Independent Variables	Social Progress Index, Access to Electricity Supply, Alternative and Nuclear Energy, IDH, Water Withdrawals, Greenhouse Gas Emissions
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10

Results	Minimum Cases in Child Node	Social Progress Index	5
	Independent Variables Included		
	Number of Nodes		4
	Number of Terminal Nodes		3
	Depth		1



**Risk**

Estimate	Std. Error
,045	,018

Growing Method:

EXHAUSTIVE CHAID

Dependent Variable:

clusters países

**Classification**

Observed	Predicted			
	1,00	2,00	3,00	Percent Correct
1,00	38	0	3	92,7%
2,00	0	38	1	97,4%
3,00	1	1	50	96,2%
Overall Percentage	29,5%	29,5%	40,9%	95,5%

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: clusters países

Vimos que, com essa mudança nos critérios, e mantidas as mesmas variáveis, houve melhora significativa na distribuição dos dados, com índice de acertos acima de 95%, contra 59% da árvore anterior. A árvore de classificação foi uma ferramenta interessante para verificar o agrupamento os países de forma a entender a variável escolhida para efetuar a divisão (Social Progress Index foi a utilizada), bem como para confirmar a assertividade do agrupamento realizado anteriormente, com índice de acerto acima de 90% em cada cluster individualmente.

#### 14. ANÁLISE DE CORRESPONDÊNCIA

A análise de correspondência é um método de análise fatorial para variáveis categóricas. A AC, basicamente, converte uma tabela de dados não negativos de duas ou múltiplas entradas em um tipo de representação gráfica em que as linhas e as colunas são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico. Este método permite mostrar como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe.

A seguir, é apresentado o resultado da análise de correspondência para a tabela abaixo que no minitab foi agrupada considerando-se as variáveis com valor de F mais alto na ANOVA, bem como região, média e desvio-padrão.

One-way ANOVA: Access to electricity (% of pop versus REG

Level	Mean	StDev
AIBER	92,45	8,69
AVECO	99,73	1,4
CARLA	91,93	10,57
CWEAS	99,91	0,31
EMDEU	99,91	0,3
ESEAS	78,3	21,49
NSWUA	87,28	25,61
SUSAF	26,63	23,99

One-way ANOVA: Social Progress Index versus REG

Level	Mean	StDev
AIBER	65,07	11,86
AVECO	90,3	8,7
CARLA	58,86	9,96
CWEAS	51,27	6,12
EMDEU	67,34	7,04
ESEAS	45,57	11,72



NSWUA	43,75	18,74
SUSAF	27,04	15,99

One-way ANOVA: IDH - 2013 versus REG

Level	Mean	StDev
AIBER	64,96	11,95
AVECO	87,36	11,35
CARLA	65,32	12,46
CWEAS	61,73	9,01
EMDEU	72,6	6,81
ESEAS	52,16	13,52
NSWUA	57,58	19,21
SUSAF	23,8	15,68

One-way ANOVA: falta de capacic para afro versus REG

Level	Mean	StDev
AIBER	57,14	19,07
AVECO	14,48	11,17
CARLA	57,34	16,74
CWEAS	55,45	12,55
EMDEU	43,63	14,48
ESEAS	72,7	12,89
NSWUA	69,92	17,76
SUSAF	80,53	12,25

One-way ANOVA: Vulnerabilidad (%) versus REG

Level	Mean	StDev	StDev
AIBER	42,59	14,24	14,24
AVECO	13,74	6,58	6,58
CARLA	40,69	11,53	11,53
CWEAS	42,03	10,3	10,3
EMDEU	32,53	9,36	9,36
ESEAS	57,17	11,61	11,61
NSWUA	49,76	19,35	19,35
SUSAF	76,66	15,1	15,1

**Simple Correspondence Analysis: MAccess; MSPI; MIDH; MFaltCap; MVuln**

Relative Inertias

	Acess		Falta Cap		Vuln		Total
	Elet	SPI	IDH				
AIBER	0,000	0,002	0,000	0,001	0,002	0,005	
AVECO	0,010	0,080	0,045	0,129	0,088	0,352	
CARLA	0,001	0,000	0,000	0,000	0,003	0,004	
CWEAS	0,008	0,003	0,000	0,000	0,001	0,013	
EMDEU	0,006	0,005	0,005	0,016	0,017	0,050	
ESEAS	0,003	0,009	0,006	0,017	0,014	0,048	
NSWUA	0,000	0,013	0,001	0,011	0,002	0,027	
SUSAF	0,094	0,026	0,047	0,125	0,210	0,502	

Total 0,122 0,138 0,104 0,300 0,336 1,000

Analysis of Contingency Table

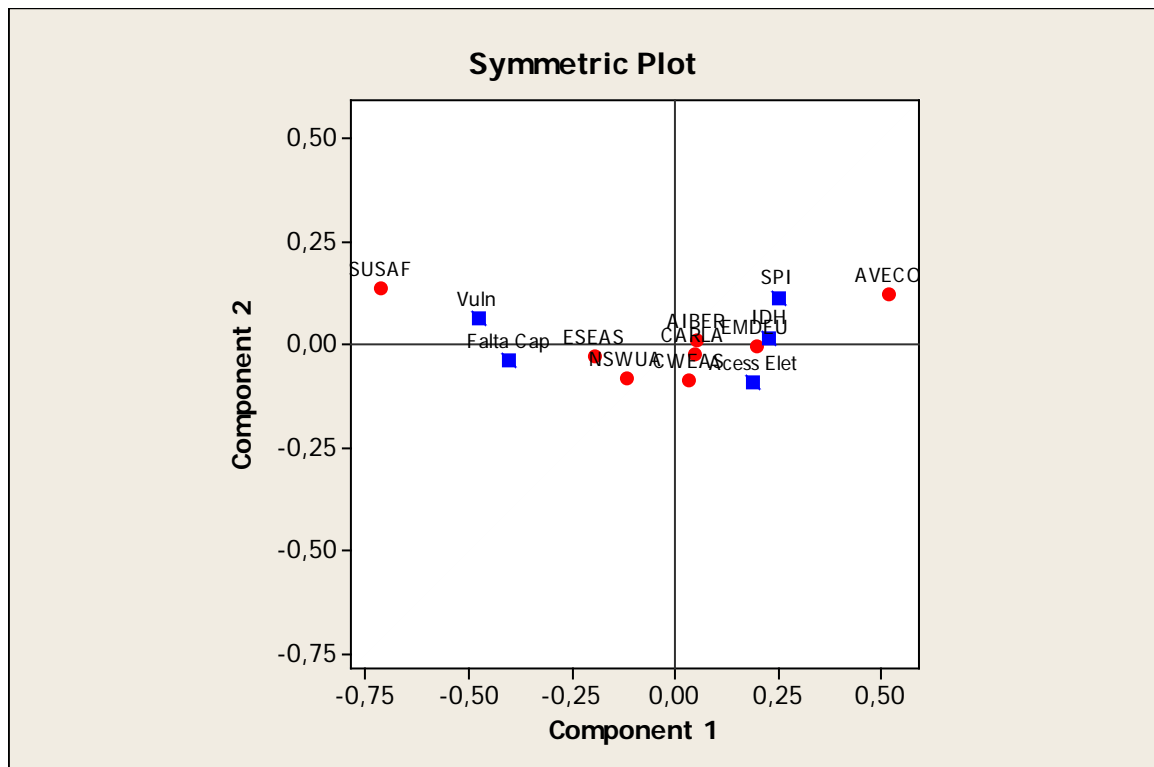
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0957	0,9410	0,9410	*****
2	0,0057	0,0558	0,9968	*
3	0,0002	0,0020	0,9988	
4	0,0001	0,0012	1,0000	
Total	0,1017			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	0,813	0,133	0,005	0,053	0,791	0,004	0,009	0,022	0,002
2	AVECO	1,000	0,126	0,352	0,518	0,948	0,355	0,121	0,052	0,326
3	CARLA	0,907	0,130	0,004	0,047	0,721	0,003	-0,024	0,187	0,013
4	CWEAS	0,908	0,128	0,013	0,036	0,129	0,002	-0,088	0,778	0,175
5	EMDEU	0,999	0,131	0,050	0,197	0,998	0,053	-0,006	0,001	0,001
6	ESEAS	1,000	0,127	0,048	-0,195	0,978	0,050	-0,029	0,022	0,019
7	NSWUA	0,978	0,128	0,027	-0,118	0,652	0,019	-0,084	0,327	0,158
8	SUSAF	1,000	0,097	0,502	-0,713	0,966	0,515	0,134	0,034	0,307

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Access Elet	0,997	0,280	0,122	0,189	0,807	0,104	-0,092	0,190	0,415
2	SPI	0,998	0,186	0,138	0,250	0,827	0,121	0,114	0,171	0,423
3	IDH	0,992	0,201	0,104	0,228	0,987	0,109	0,016	0,005	0,010
4	Falta Cap	0,997	0,187	0,300	-0,402	0,987	0,315	-0,041	0,010	0,056
5	Vuln	0,998	0,147	0,336	-0,478	0,982	0,350	0,061	0,016	0,097



Verificando o gráfico das médias das variáveis selecionadas, notamos, em primeiro lugar, a localização oposta das regiões AVECO e SUSAF, polarização que foi uma constante ao longo desse trabalho. AIBER e CARLA, encontram-se muito próximos, em uma posição intermediária. Quanto às variáveis, vemos **SPI, IDH e Acesso à Eletricidade** de um lado e com índices relativamente positivos, ao passo que **Vulnerabilidade e Falta de Capacidade de Adaptação** encontram-se mais à esquerda, demonstrando serem esses problemas mais recorrentes para uma parte considerável dos países.

### Simple Correspondence Analysis: StAcess; StSPI; StIDH; StFaltCap; StVuln

#### Relative Inertias

	Acess						Total
	Elet	SPI	IDH	Falta Cap	Vuln		
AIBER	0,021	0,000	0,003	0,017	0,003	0,044	
AVECO	0,099	0,007	0,032	0,009	0,004	0,150	
CARLA	0,001	0,002	0,000	0,008	0,001	0,012	
CWEAS	0,135	0,002	0,005	0,029	0,021	0,192	
EMDEU	0,135	0,000	0,002	0,072	0,010	0,218	
ESEAS	0,108	0,002	0,001	0,018	0,009	0,138	
NSWUA	0,053	0,000	0,001	0,031	0,000	0,086	
SUSAF	0,099	0,001	0,001	0,056	0,002	0,160	
Total	0,651	0,016	0,044	0,240	0,048	1,000	

#### Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0875	0,9057	0,9057	*****
2	0,0061	0,0629	0,9686	**
3	0,0020	0,0204	0,9890	
4	0,0011	0,0110	1,0000	
Total	0,0966			

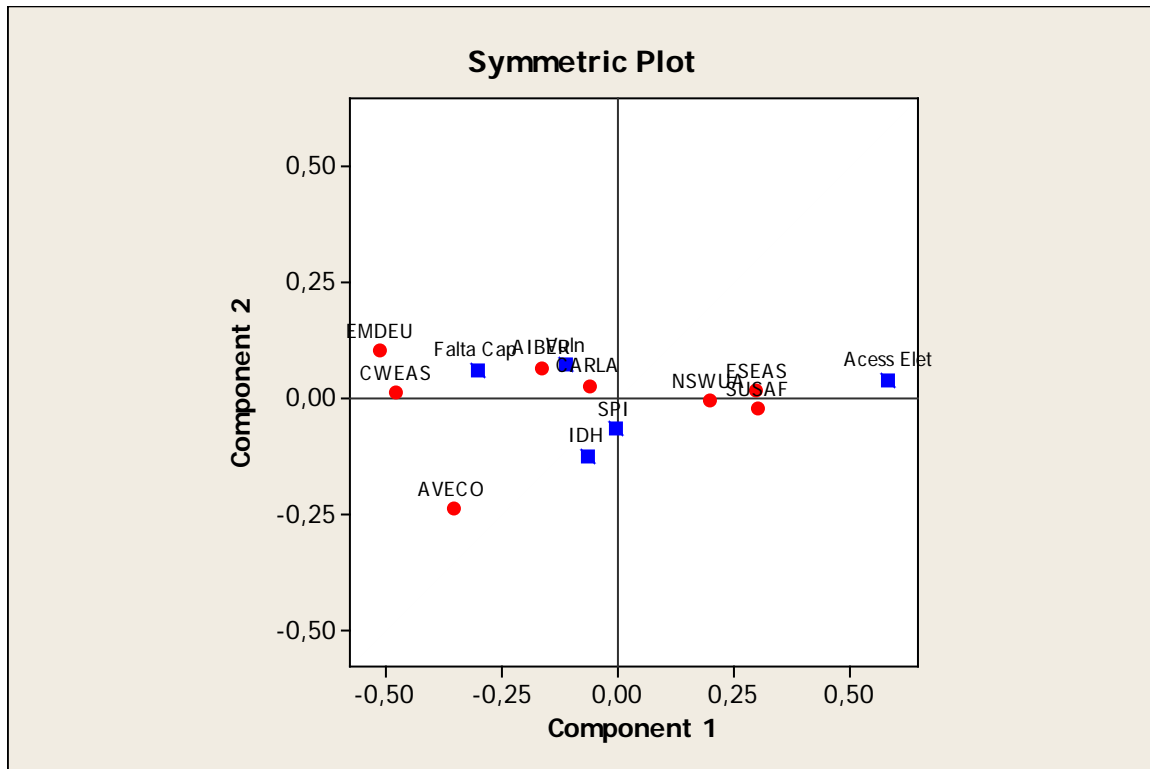
#### Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	0,977	0,132	0,044	-0,164	0,842	0,041	0,066	0,135	0,094
2	AVECO	0,994	0,079	0,150	-0,355	0,684	0,114	-0,239	0,309	0,739
3	CARLA	0,477	0,123	0,012	-0,062	0,402	0,005	0,027	0,076	0,015
4	CWEAS	0,949	0,077	0,192	-0,478	0,948	0,201	0,014	0,001	0,002
5	EMDEU	0,990	0,076	0,218	-0,512	0,951	0,229	0,102	0,038	0,132
6	ESEAS	0,963	0,143	0,138	0,299	0,960	0,146	0,016	0,003	0,006
7	NSWUA	0,970	0,202	0,086	0,199	0,970	0,092	-0,004	0,000	0,001
8	SUSAF	0,980	0,167	0,160	0,301	0,975	0,172	-0,021	0,005	0,012

#### Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Acess Elet	0,998	0,186	0,651	0,580	0,994	0,715	0,040	0,005	0,049
2	SPI	0,495	0,181	0,016	-0,005	0,004	0,000	-0,064	0,491	0,123
3	IDH	0,929	0,201	0,044	-0,067	0,211	0,010	-0,123	0,718	0,504
4	Falta Cap	0,969	0,235	0,240	-0,303	0,930	0,247	0,062	0,039	0,148

5 Vuln 0,757 0,197 0,048 -0,112 0,529 0,028 0,073 0,228 0,175



Quanto ao desvio-padrão, novamente **AVECO** se destaca de forma positiva, pois comparativamente apresenta o menor desvio padrão, ou seja, tem menor dispersão e maior uniformidade dentre os grupos de países por região. **AIBER** e **CARLA** apresentam posições intermediárias. Quanto às variáveis, estatisticamente falando **IDH** e **SPI** são as menos dispersas e com relação de equilíbrio nos dois componentes, e chama a atenção **Acesso à Eletricidade**, provavelmente devido à precariedade em termos de infraestrutura e disponibilidade, nos países africanos e em alguns asiáticos.

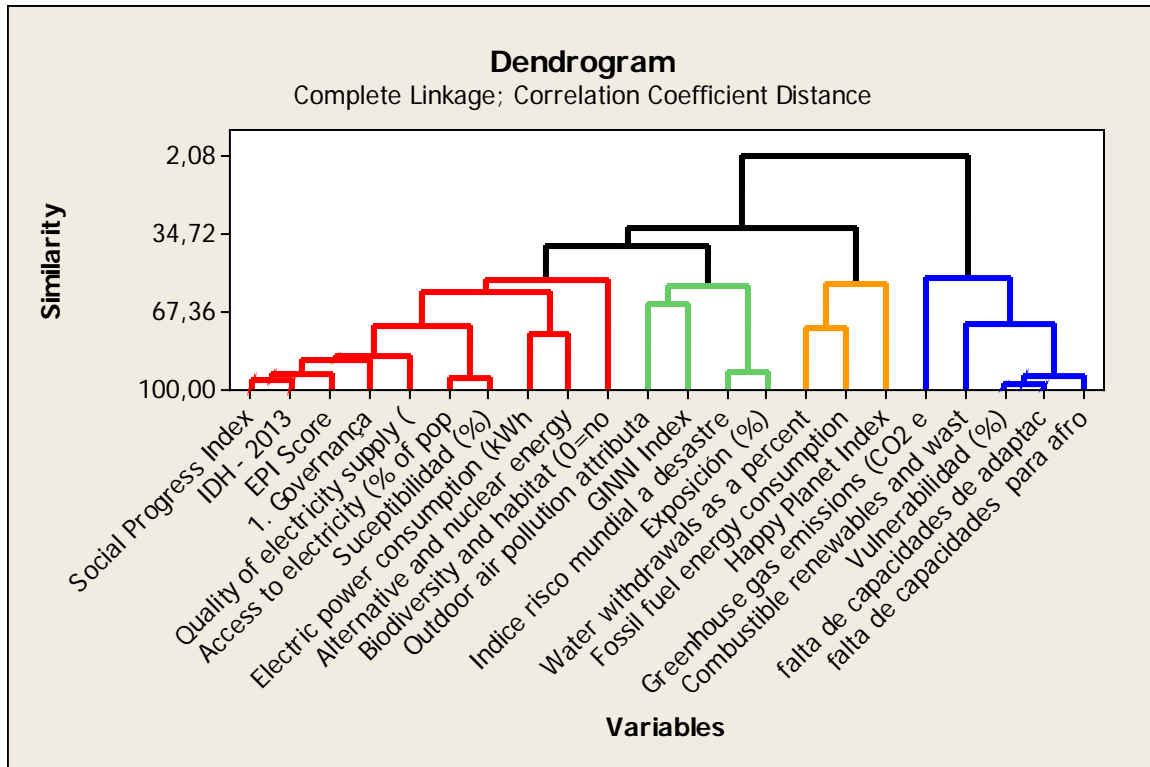
Vemos que os países desenvolvidos, sobretudo da Europa Ocidental são os mais bem posicionados, enquanto os países africanos estão bastante atrás. Os países ibero-americanos se encontram quase sempre em uma posição intermediária, muitas vezes puxada por Portugal, Espanha, Chile, Costa Rica Uruguai – excluídos tais países, muito provavelmente a região estaria em um patamar bem pior. É o que veremos em breve, ao explorar em detalhe e profundidade apenas essa região na análise da base de dados.

## **15. UM OLHAR SOBRE OS PAÍSES IBEROAMERICANOS**

Passaremos, agora, a uma análise mais acurada dos países do grupo ibero-americano, com o objetivo de verificar, mais do que a posição específica desse grupo junto aos demais – o que já foi identificado nas análises anteriores – ver principalmente diferenças e semelhanças marcantes e que aproximam alguns de seus países do grupo AVECO, enquanto outros se aproximam mais do grupo do SUFAF.

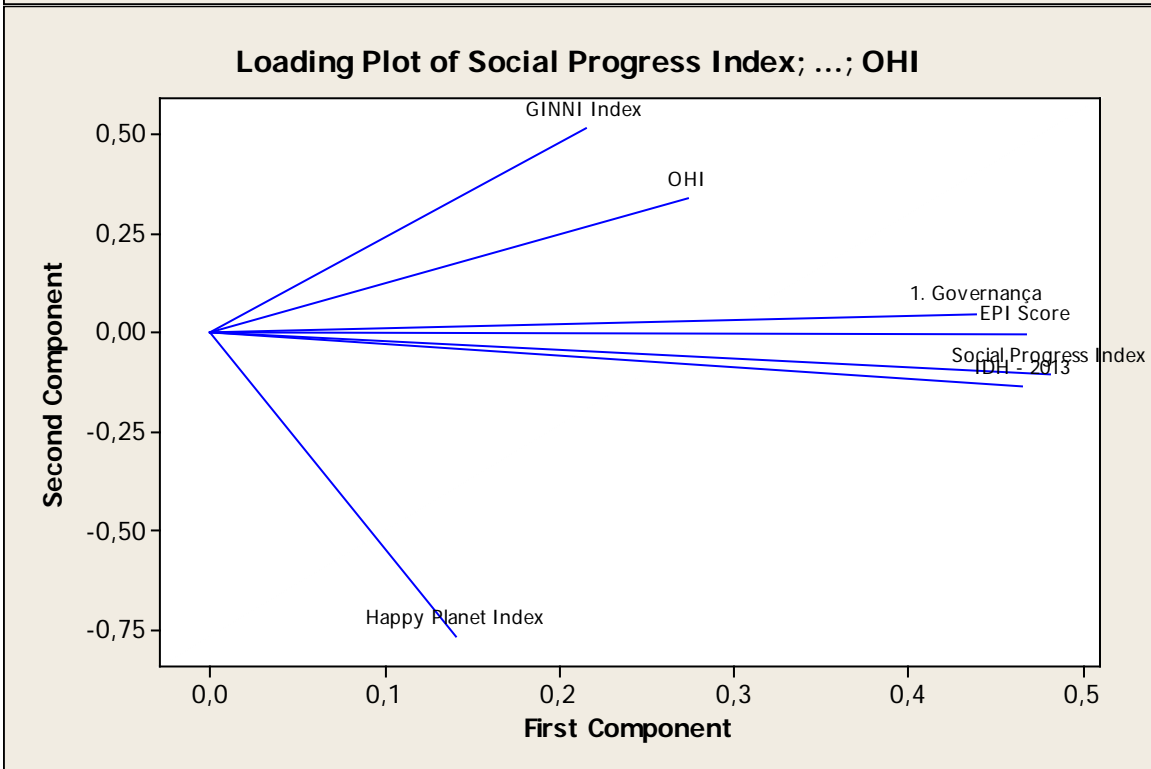
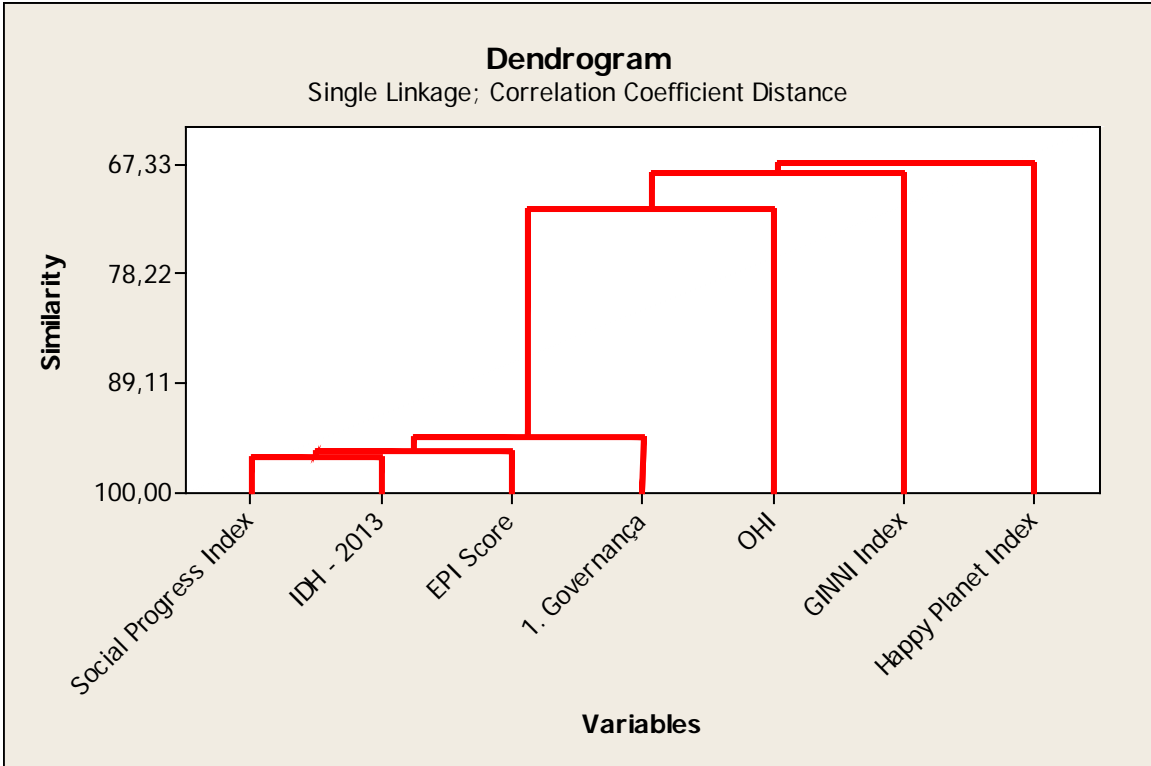
### **15.1 ANÁLISE DE CLUSTERS DE VARIÁVEIS E PAÍSES**

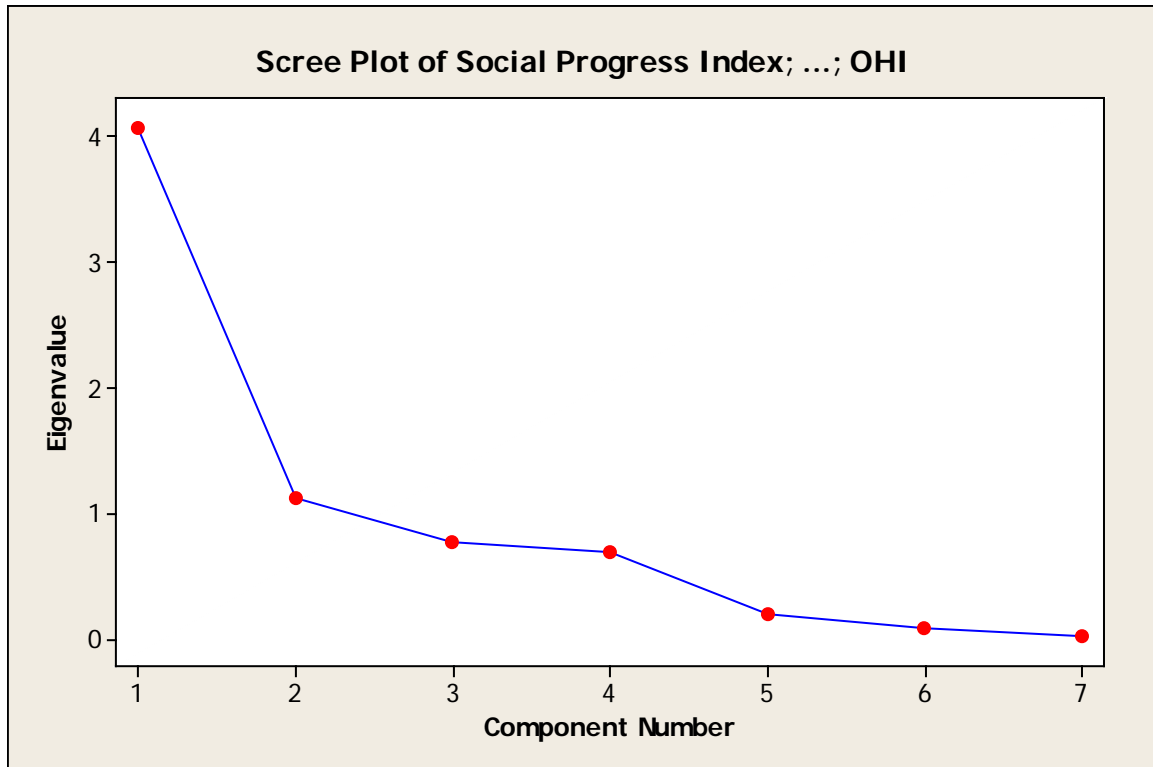
#### **15.1.1. DENDROGRAMAS DAS VARIÁVEIS**



Analisando-se da esquerda para a direita, nota-se que há uma grande similaridade entre as variáveis sintéticas como IDH, Social Progress Index, EPI Score e Governança, o que me parece lógico tendo em vista que tais índices são derivados de uma série de outros indicadores. O segundo grupo, que vem a seguir, inclui um grande número de variáveis. Em seguida aparecem variáveis isoladas ou pouco associadas com as demais. Vale ressaltar que as variáveis “vulnerabilidade” e “falta de capacidade de adaptação” possuem alta similaridade – a razão é que a última é componente da primeira. Digno de nota é a similaridade entre as variáveis “Susceptibilidade” e “Acesso à Eletricidade”. A análise de correlação irá apontar em mais detalhe, além das mencionadas, quais outras variáveis apresentam alto índice de correlação.

### 15.1.2 EXAME E SELEÇÃO DAS VARIÁVEIS SINTÉTICAS





Eigenanalysis of the Correlation Matrix

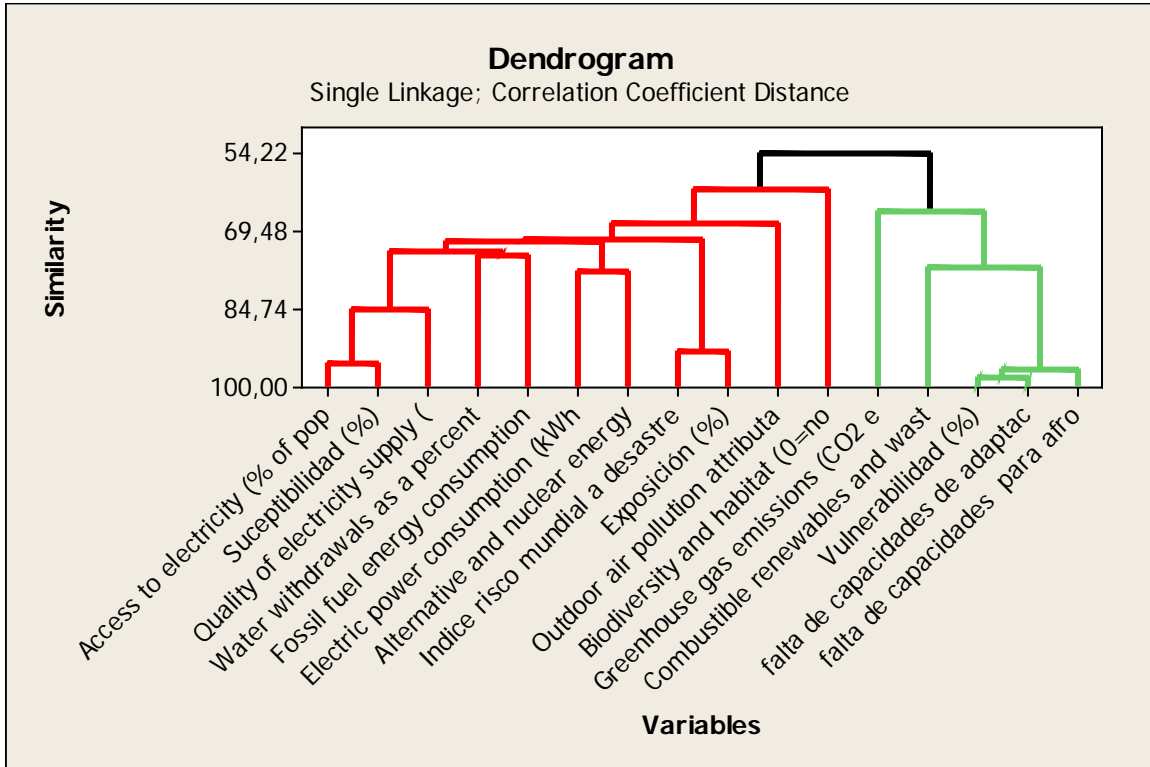
Eigenvalue	4,0653	1,1208	0,7781	0,7013	0,2031	0,0985	0,0327
Proportion	0,581	0,160	0,111	0,100	0,029	0,014	0,005
Cumulative	0,581	0,741	0,852	0,952	0,981	0,995	1,000

Variable	PC1	PC2
Social Progress Index	0,481	-0,107
IDH - 2013	0,465	-0,138
1. Governança	0,439	0,048
EPI Score	0,467	-0,003
Happy Planet Index	0,141	-0,766
GINNI Index	0,215	0,515
OHI	0,274	0,340

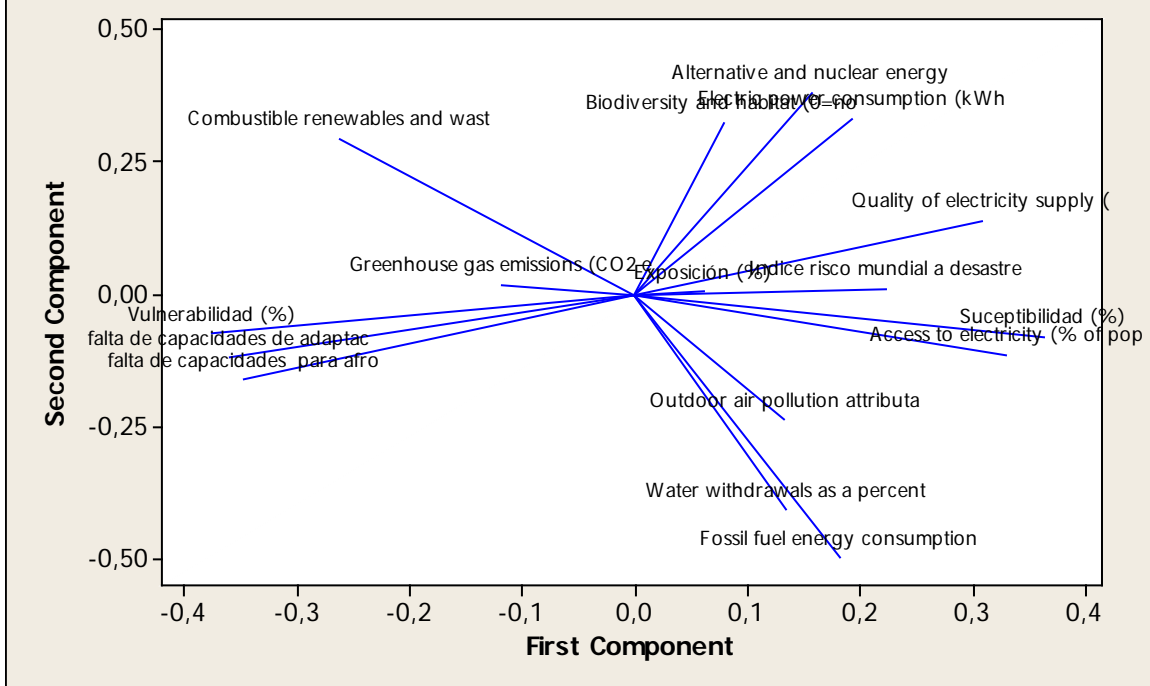
A análise de componentes principais e o loading plot mostram as mesmas variáveis em proximidade, o que também se confirma pelo dendrograma, o qual inclui as variáveis mais representativas do grupo das sintéticas, quais sejam: **Social Progress Index, IDH,,Governança, EPI Score, Happy Planet Index, GINI Index e Ocean Health Index.**

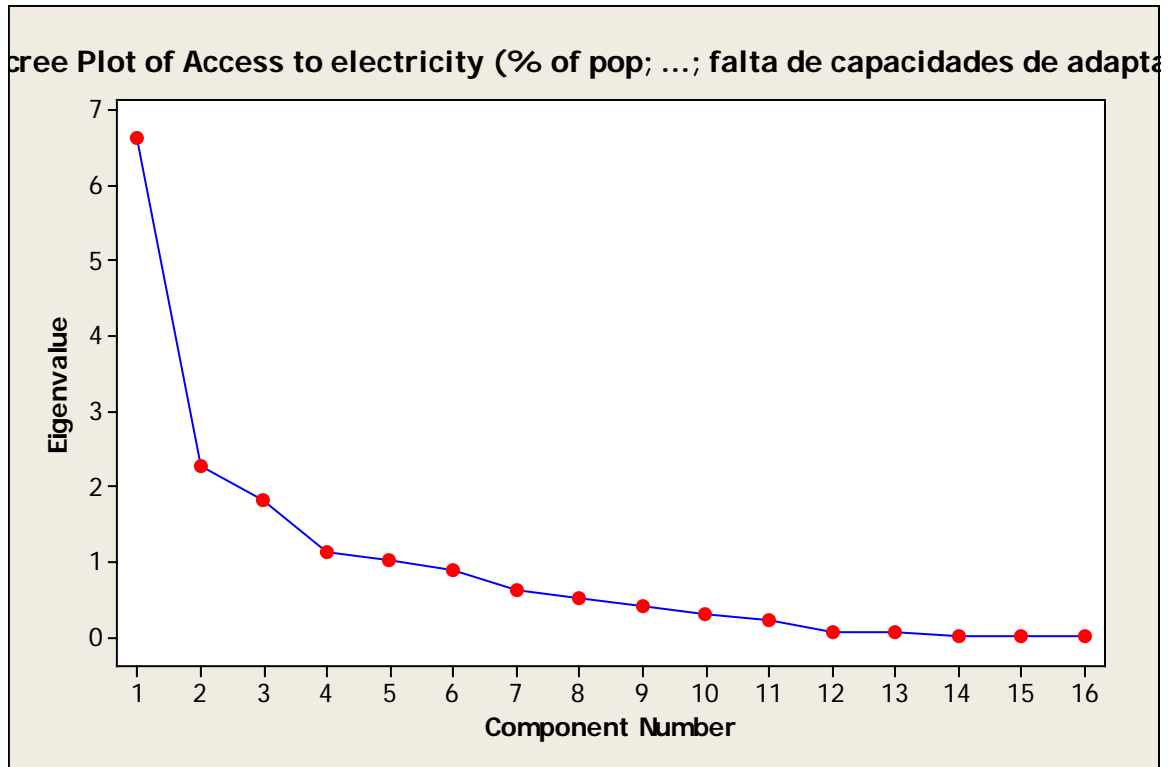


### 15.1.3 EXAME E SELEÇÃO DAS VARIÁVEIS ANALÍTICAS



loading Plot of Access to electricity (% of pop; ...; falta de capacidades de adap





### Principal Component Analysis: Vulnerabilid; Fossil fuel ; Electric pow; Biodive

Eigenanalysis of the Correlation Matrix

Eigenvalue	6,0327	2,2390	1,8219	1,1206	1,0145	0,8773	0,6081	0,4601
Proportion	0,402	0,149	0,121	0,075	0,068	0,058	0,041	0,031
Cumulative	0,402	0,551	0,673	0,748	0,815	0,874	0,914	0,945

Eigenvalue	0,4035	0,2392	0,0749	0,0607	0,0214	0,0187	0,0074
Proportion	0,027	0,016	0,005	0,004	0,001	0,001	0,000
Cumulative	0,972	0,988	0,993	0,997	0,998	1,000	1,000

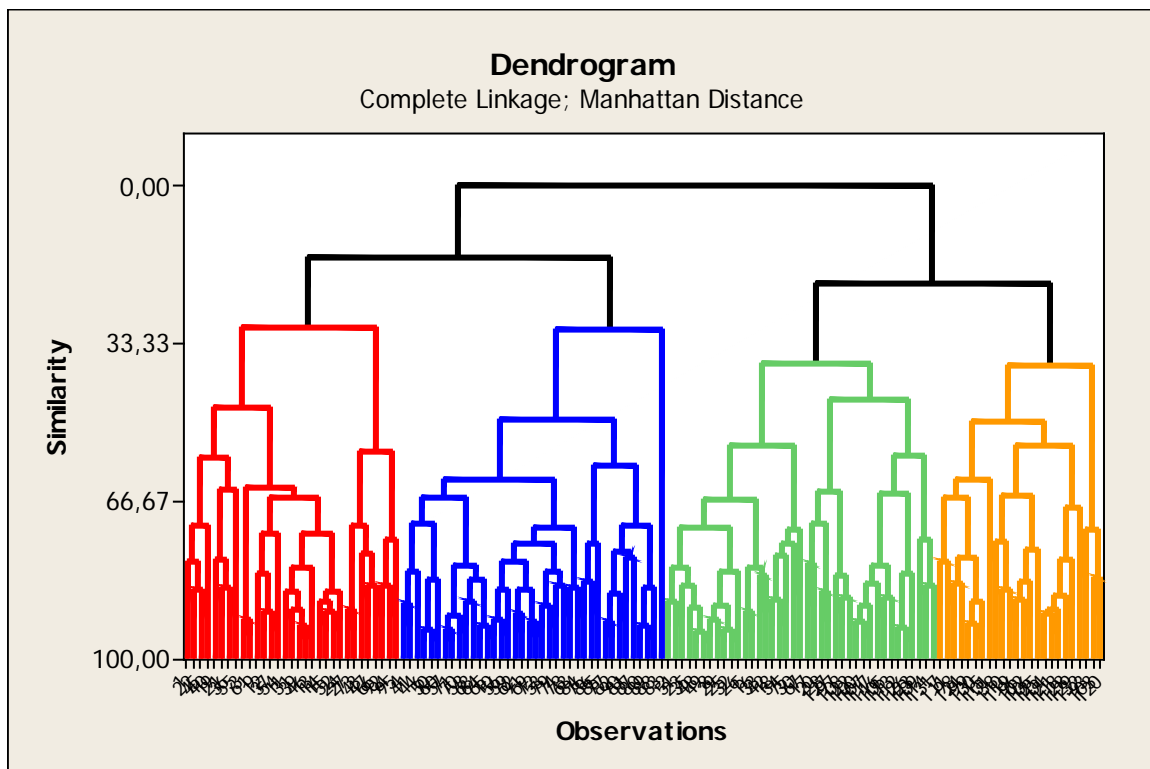
Variable	PC1	PC2
Vulnerabilidade (%)	-0,391	-0,095
Fossil fuel energy consumption	0,202	-0,489
Electric power consumption (kWh	0,197	0,355
Biodiversity and habitat (0=no	0,075	0,329
Exposición (%)	0,070	0,033
Access to electricity (% of pop	0,349	-0,095
Outdoor air pollution attributa	0,147	-0,224
Greenhouse gas emissions (CO2 e	-0,130	0,005
Water withdrawals as a percent	0,151	-0,402
Alternative and nuclear energy	0,159	0,400
Combustible renewables and wast	-0,283	0,277
Indice risco mundial a desastre	0,238	0,045
Suceptibilidad (%)	0,381	-0,059
falta de capacidades para afro	-0,358	-0,183
falta de capacidades de adaptac	-0,372	-0,141

Realizamos o mesmo teste incluindo todas as variáveis analíticas, o que mostrou agrupamentos mais dispersos entre as variáveis, conforme mostra o loading plot dos

componentes. De qualquer forma, foram identificadas as variáveis analíticas mais representativas, quais sejam: “Vulnerabilidade”, “Fossil fuel energy consumption”, “Electric power consumption”, “Biodiversidade e habitat” e “Exposition”.

#### 15.1.4 REAGRUPAMENTO DE PAÍSES

Neste estudo queremos usar o reagrupamento de forma tal que possamos entender melhor o comportamento das variáveis selecionadas (sintéticas + analíticas) no contexto dos novos grupos a serem criados.



**Final Partition Ngrupos**  
**Number of clusters: 4**

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	31	72407,5	45,0608	79,896
Cluster2	39	79707,1	43,6628	66,838
Cluster3	38	63255,4	37,4439	108,190
Cluster4	24	48433,4	43,0652	67,955

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Social Progress Index	61,1309	41,4823	84,6561	26,7389
Happy Planet Index	63,9708	48,0993	46,9124	31,4298
GINNI Index	44,0038	65,4968	81,6618	53,5606
Vulnerabilidad (%)	42,8582	56,8057	17,3703	77,9095

Fossil fuel energy consumption	73,5653	79,0917	72,5235	36,7365
Electric power consumption (kWh)	5,8985	5,3164	15,5407	3,2864
Biodiversity and habitat (0=no	76,5645	28,6913	72,5416	66,9458
Exposición (%)	69,7108	77,4028	79,6439	79,4375

Variable	Grand centroid
Social Progress Index	55,8449
Happy Planet Index	48,4542
GINNI Index	62,9325
Vulnerabilidad (%)	46,0146
Fossil fuel energy consumption	68,2021
Electric power consumption (kWh)	8,0274
Biodiversity and habitat (0=no	59,5132
Exposición (%)	76,6115

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,0000	60,6328	55,8667	71,4980
Cluster2	60,6328	0,0000	75,8774	65,9448
Cluster3	55,8667	75,8774	0,0000	97,5244
Cluster4	71,4980	65,9448	97,5244	0,0000

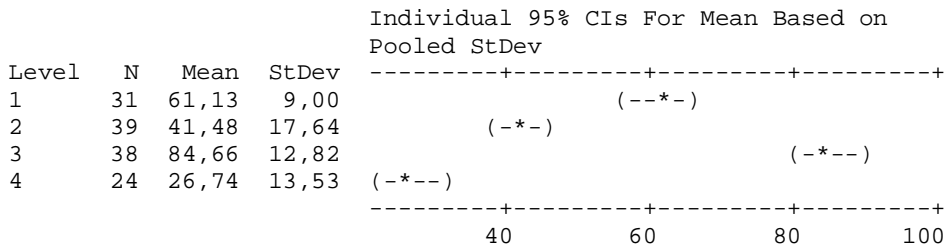
Vemos que há maior distanciamento entre os clusters 3 e 4 (97), seguidos dos grupos 2 e 3 (75), e maior aproximação entre os grupos 1 e 3 (55), seguidos dos grupos 1 e 2(60).

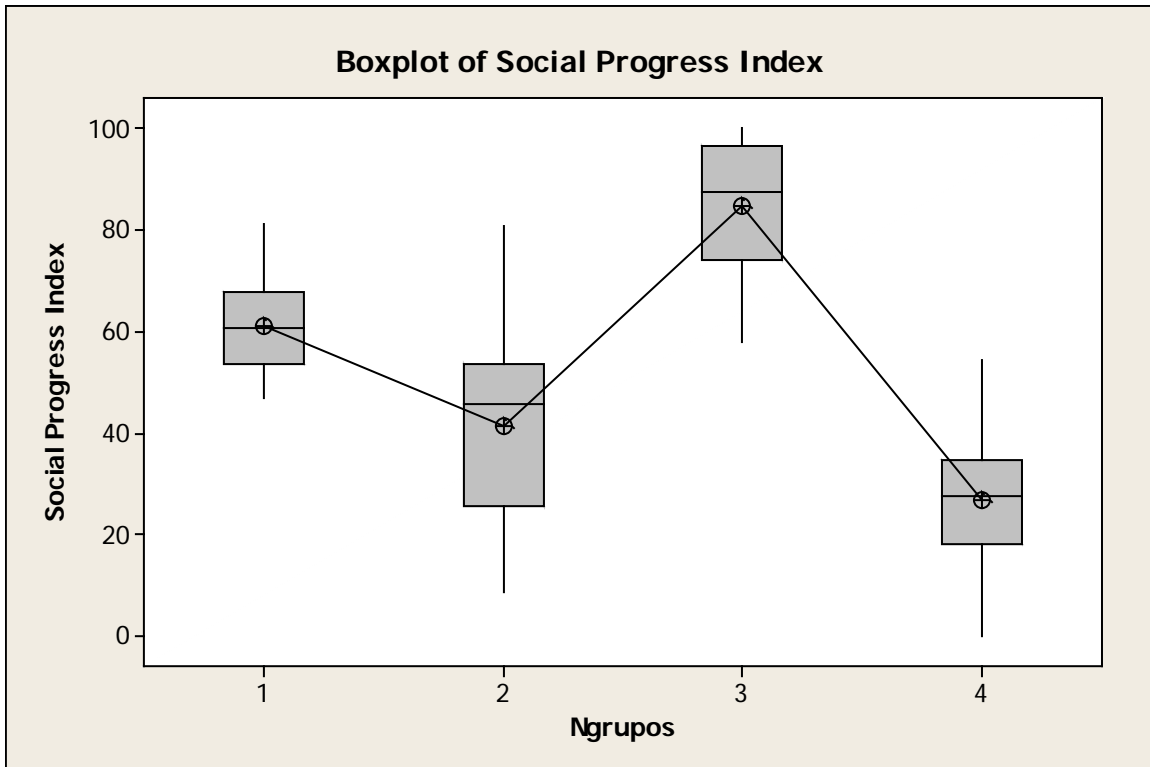
### 15.2.1. ANÁLISE DE VARIÂNCIA COM AS VARIÁVEIS SINTÉTICAS SELECIONADAS

#### One-way ANOVA : Social Progress Index versus Ngrupos

Source	DF	SS	MS	F	P
Ngrupos	3	60786	20262	105,66	0,000
Error	128	24547	192		
Total	131	85333			

S = 13,85    R-Sq = 71,23%    R-Sq(adj) = 70,56%





**One-way ANOVA: EPI Score versus Ngrupos**

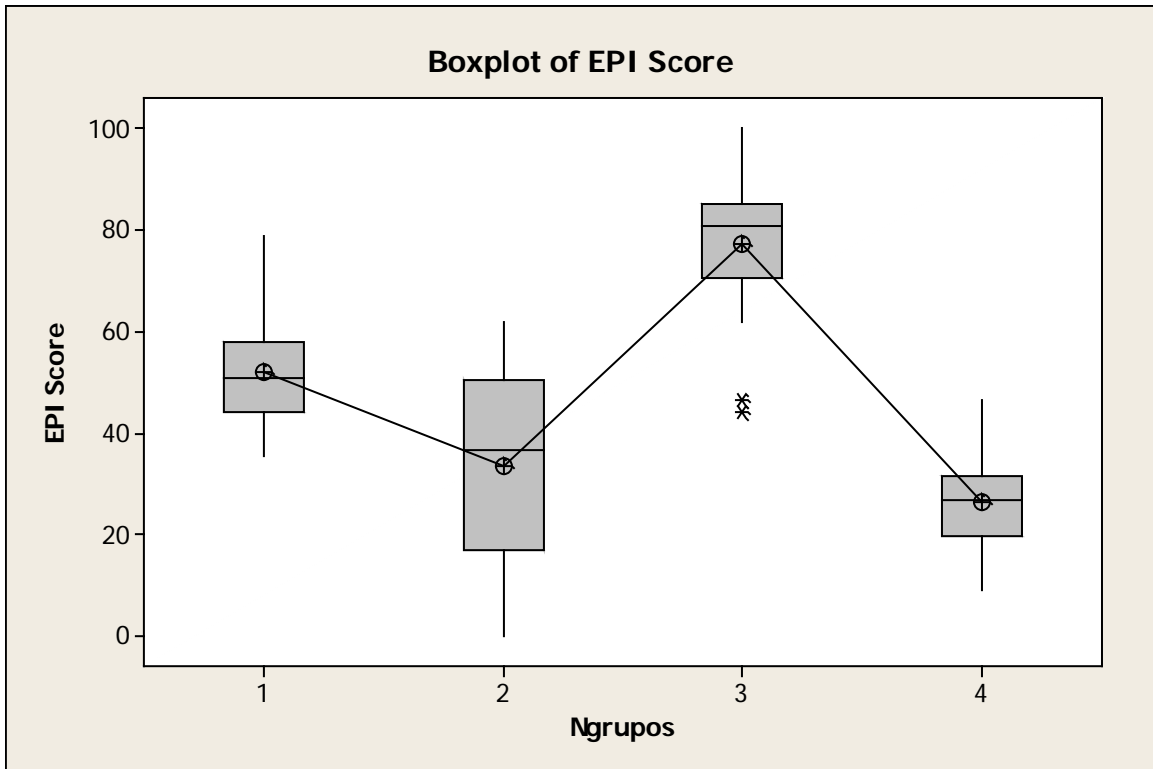
Source	DF	SS	MS	F	P
Ngrupos	3	52310	17437	99,00	0,000
Error	128	22545	176		
Total	131	74855			

S = 13,27    R-Sq = 69,88%    R-Sq(adj) = 69,18%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	CI Lower	CI Upper
1	31	51,99	11,09	32,00	71,98
2	39	33,63	17,87	10,00	57,26
3	38	77,31	11,82	64,00	90,62
4	24	26,30	8,21	10,00	42,60

Pooled StDev = 13,27



**One-way ANOVA: IDH - 2013 versus Ngrupos**

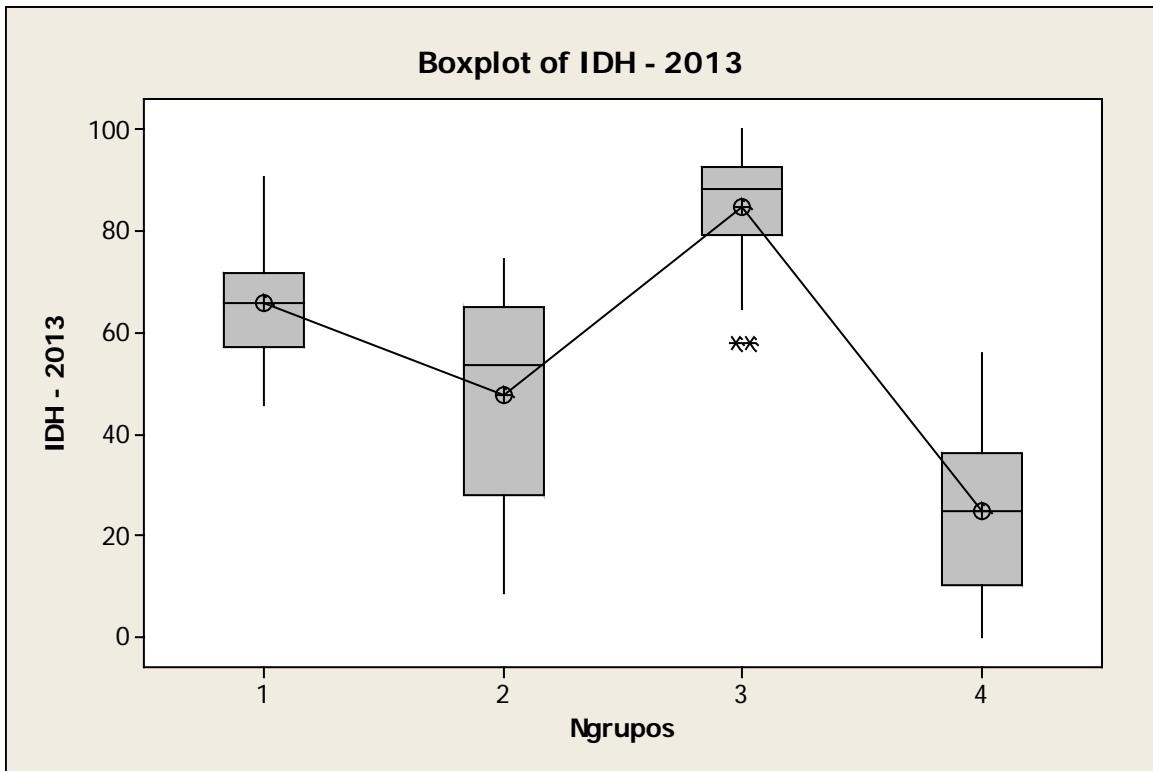
Source	DF	SS	MS	F	P
Ngrupos	3	59075	19692	85,22	0,000
Error	128	29576	231		
Total	131	88651			

S = 15,20    R-Sq = 66,64%    R-Sq(adj) = 65,86%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	CI Lower	CI Upper
1	31	65,59	11,49	42,61	88,57
2	39	47,69	20,19	7,31	88,07
3	38	84,51	11,00	73,51	95,51
4	24	24,80	15,67	-3,07	52,67

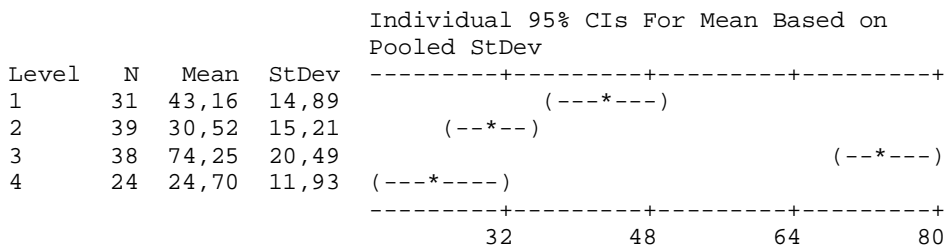
Pooled StDev = 15,20



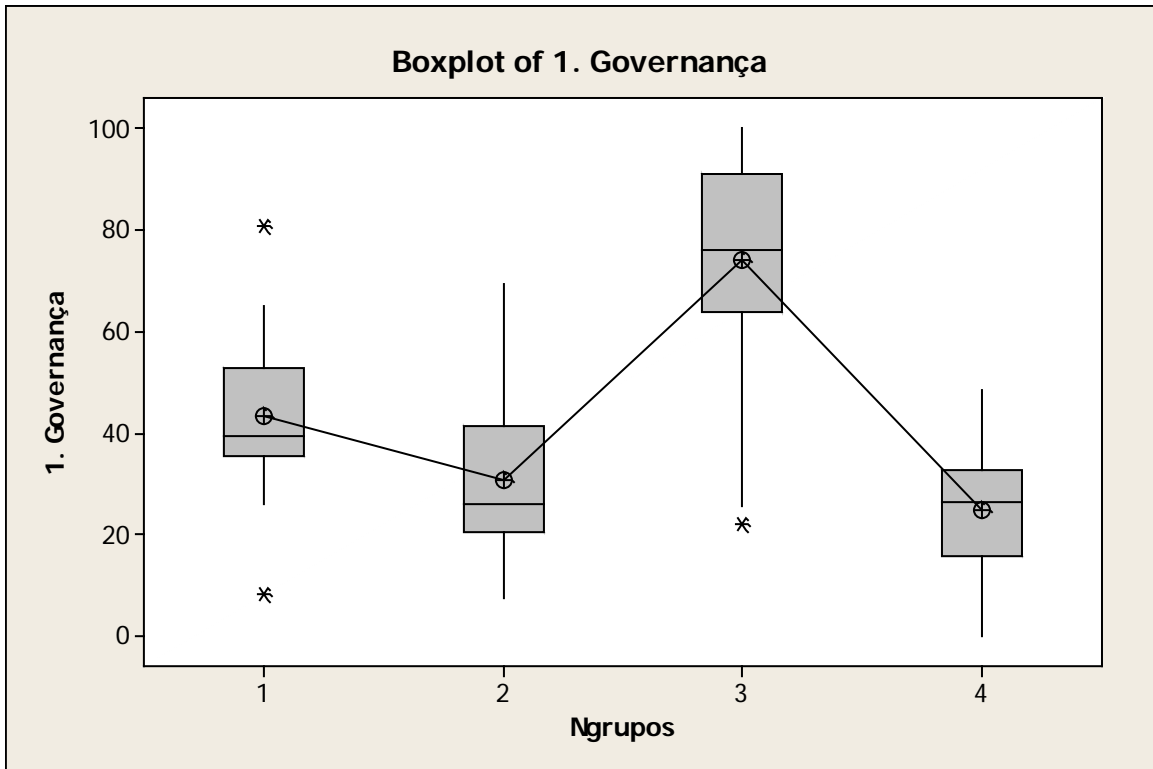
**One-way ANOVA: 1. Governança versus Ngrupos**

Source	DF	SS	MS	F	P
Ngrupos	3	50683	16894	63,13	0,000
Error	128	34254	268		
Total	131	84938			

S = 16,36    R-Sq = 59,67%    R-Sq(adj) = 58,73%



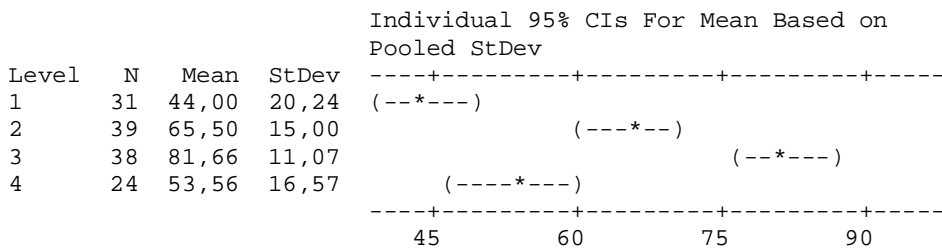
Pooled StDev = 16,36



**One-way ANOVA: GINNI Index versus Ngrupos**

Source	DF	SS	MS	F	P
Ngrupos	3	26801	8934	36,08	0,000
Error	128	31691	248		
Total	131	58492			

S = 15,73    R-Sq = 45,82%    R-Sq(adj) = 44,55%

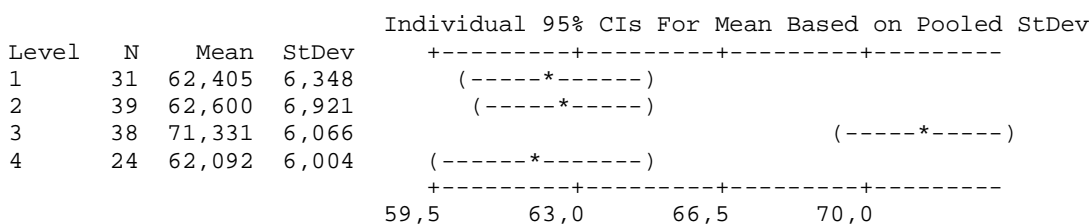




### One-way ANOVA: OHI versus Ngrupos

Source	DF	SS	MS	F	P
Ngrupos	3	2159,4	719,8	17,65	0,000
Error	128	5219,5	40,8		
Total	131	7379,0			

S = 6,386    R-Sq = 29,26%    R-Sq(adj) = 27,61%

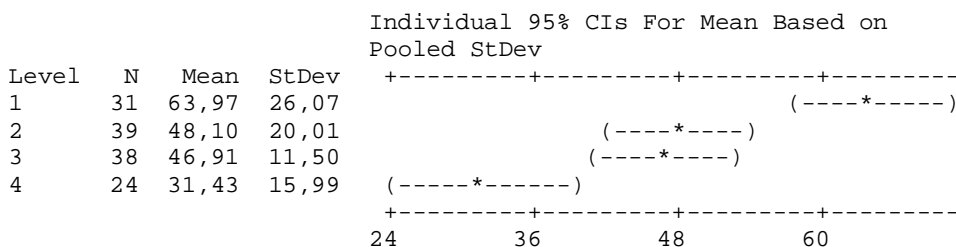


Pooled StDev = 6,386

### One-way ANOVA: Happy Planet Index versus Ngrupos

Source	DF	SS	MS	F	P
Ngrupos	3	14515	4838	13,35	0,000
Error	128	46379	362		
Total	131	60894			

S = 19,04    R-Sq = 23,84%    R-Sq(adj) = 22,05%



Pooled StDev = 19,04

Das variáveis sintéticas, as que apresentaram maior discriminação entre os grupos foram **“Social Progress Index”**, **“EPI Score”**, **“IDH”** e **“Governança”**, o que é corroborado pelos valores de F de **105**, **99**, **85**, e **63**, respectivamente. Em oposição, as variáveis **“GINI”**, **“Ocean Health Index”** e **“Happy Planet Index”** apresentaram F baixo (respectivamente **36**, **17**, e **13**) e alta sobreposição entre as médias dos grupos, relevando que não há discriminação significativa entre as variáveis.

## 15.2.2 ANÁLISE DE VARIÂNCIA COM AS VARIÁVEIS ANALÍTICAS SELECIONADAS

### One-way ANOVA: Vulnerabilidad (%) versus Ngrupos

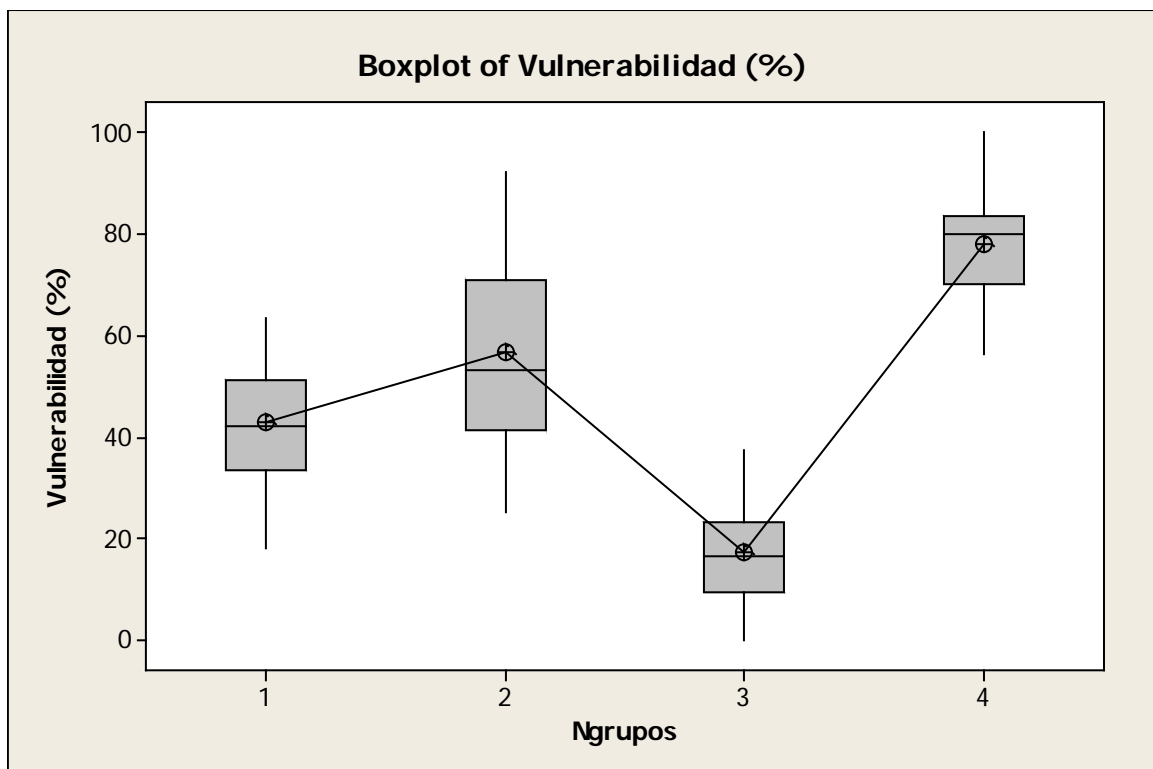
Source	DF	SS	MS	F	P
Ngrupos	3	60444	20148	112,40	0,000
Error	128	22945	179		
Total	131	83389			

S = 13,39    R-Sq = 72,48%    R-Sq(adj) = 71,84%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	95% CI
1	31	42,86	11,62	(- * - -)
2	39	56,81	18,70	(- * - -)
3	38	17,37	8,78	(- * -)
4	24	77,91	10,94	(- - * - -)

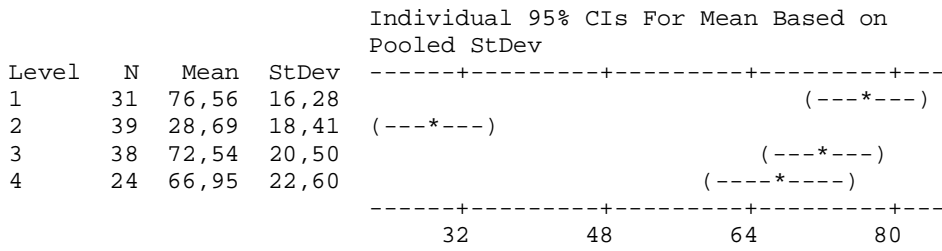
20                      40                      60                      80



**One-way ANOVA: Biodiversity and habitat (0=no versus Ngrupos**

Source	DF	SS	MS	F	P
Ngrupos	3	53839	17946	47,74	0,000
Error	128	48120	376		
Total	131	101958			

S = 19,39 R-Sq = 52,80% R-Sq(adj) = 51,70%

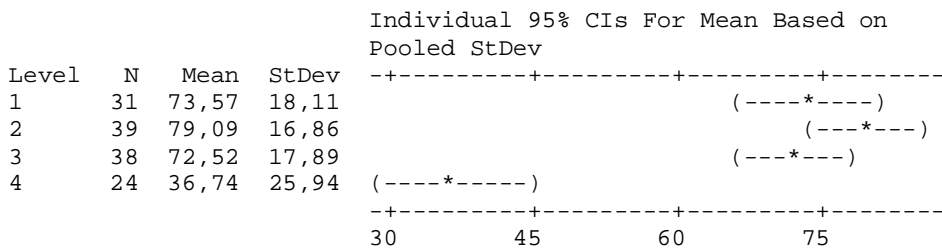


Pooled StDev = 19,39

**One-way ANOVA: Fossil fuel energy consumption versus Ngrupos**

Source	DF	SS	MS	F	P
Ngrupos	3	29988	9996	26,68	0,000
Error	128	47965	375		
Total	131	77953			

S = 19,36 R-Sq = 38,47% R-Sq(adj) = 37,03%



Pooled StDev = 19,36

### One-way ANOVA: Electric power consumption (kWh versus Ngrupos)

Source	DF	SS	MS	F	P
Ngrupos	3	3111,7	1037,2	11,12	0,000
Error	128	11939,9	93,3		
Total	131	15051,6			

S = 9,658    R-Sq = 20,67%    R-Sq(adj) = 18,81%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	CI
1	31	5,899	6,366	(-----*-----)
2	39	5,316	3,168	(-----*-----)
3	38	15,541	16,452	(-----*-----)
4	24	3,286	3,779	(-----*-----)

0,0                      5,0                      10,0                      15,0

Pooled StDev = 9,658

### One-way ANOVA: Exposición (%) versus Ngrupos

Source	DF	SS	MS	F	P
Ngrupos	3	2042	681	2,88	0,038
Error	128	30218	236		
Total	131	32259			

S = 15,36    R-Sq = 6,33%    R-Sq(adj) = 4,13%

Individual 95% CIs For Mean Based on Pooled StDev

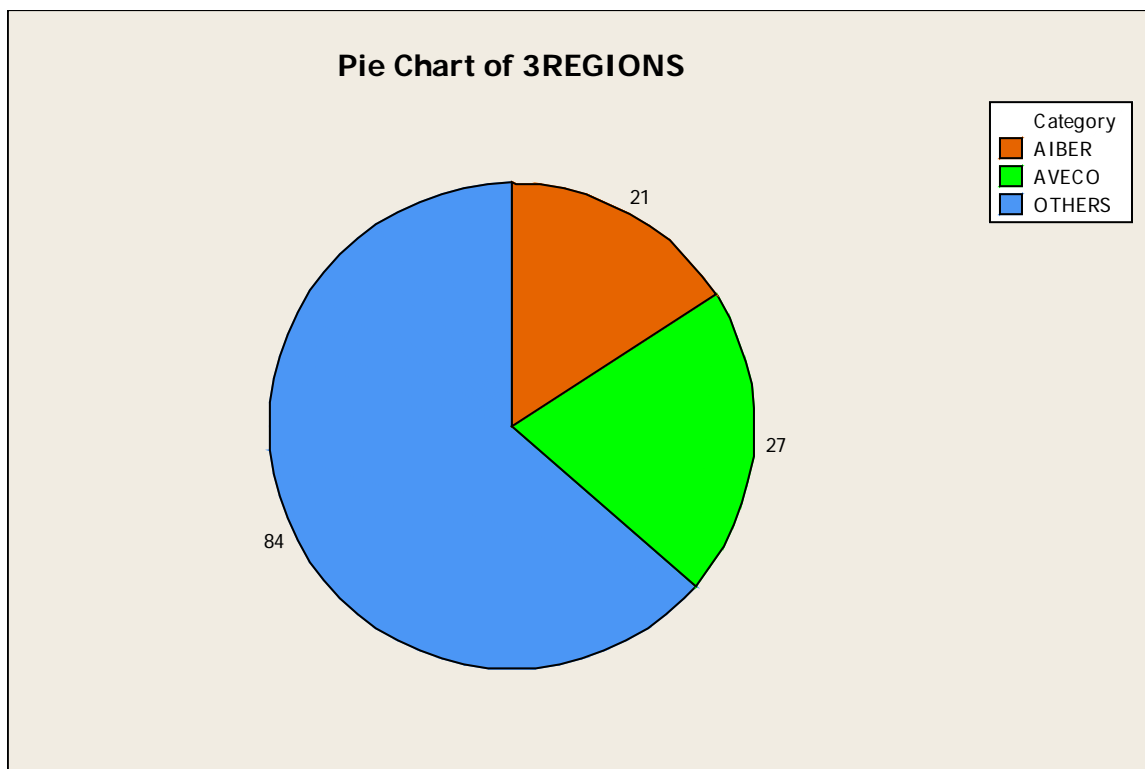
Level	N	Mean	StDev	CI
1	31	69,71	21,79	(-----*-----)
2	39	77,40	13,33	(-----*-----)
3	38	79,64	14,24	(-----*-----)
4	24	79,44	8,65	(-----*-----)

66,0                      72,0                      78,0                      84,0

Pooled StDev = 15,36

Das variáveis analíticas, as que apresentaram maior discriminação entre os grupos foi **“Vulnerabilidade”** e **“Biodiversidade e Habitat”**, o que é corroborado pelos valores de F de **112** e **47**, respectivamente. Em oposição, as variáveis **“Fossil Fuel Energy Consumption”**, **“Electric Power Consumption”** e **“Exposición”** e apresentaram F baixo (respectivamente **26**, **11**, e **2**) e alta sobreposição entre as médias dos grupos, relevando que não há discriminação significativa entre as variáveis.

### 15.3. REAGRUPAMENTO FINAL PAÍSES – AIBER, AVECO, OTHERS



Foi realizado o reagrupamento entre os países, de forma a que tivéssemos apenas três clusters, quais sejam: AIBER, composto por países da América Latina mais Portugal e Espanha; AVECO, composto por países de economia avançada e todos os demais países. O objetivo desse agrupamento é realizar uma análise comparativa entre o desempenho de AIBER com AVECO e com os demais países.

### 15.3.1. ANÁLISE DE VARIÂNCIA COM AS VARIÁVEIS SINTÉTICAS SELECIONADAS

#### One-way ANOVA: 1. Governança versus 3REGIONS

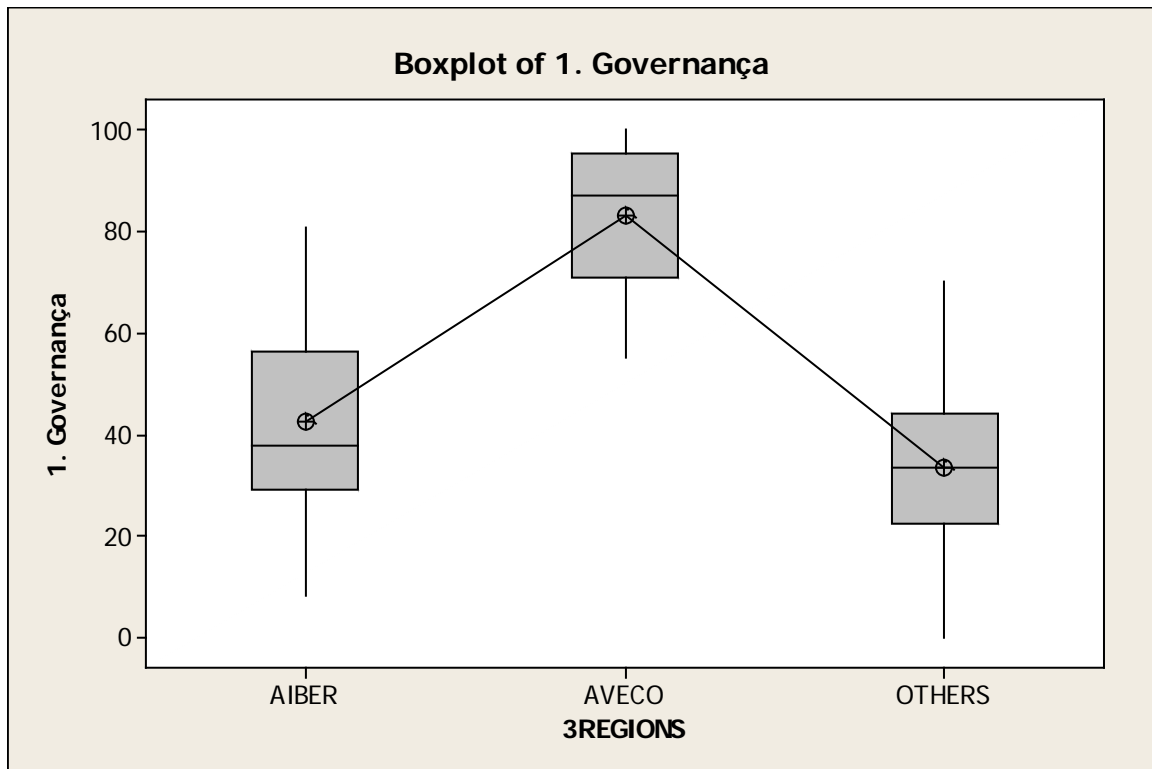
Source	DF	SS	MS	F	P
3REGIONS	2	51019	25510	97,02	0,000
Error	129	33919	263		
Total	131	84938			

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

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

+-----+-----+-----+-----+  
30          45          60          75

Pooled StDev = 16,22



**One-way ANOVA: Social Progress Index versus 3REGIONS**

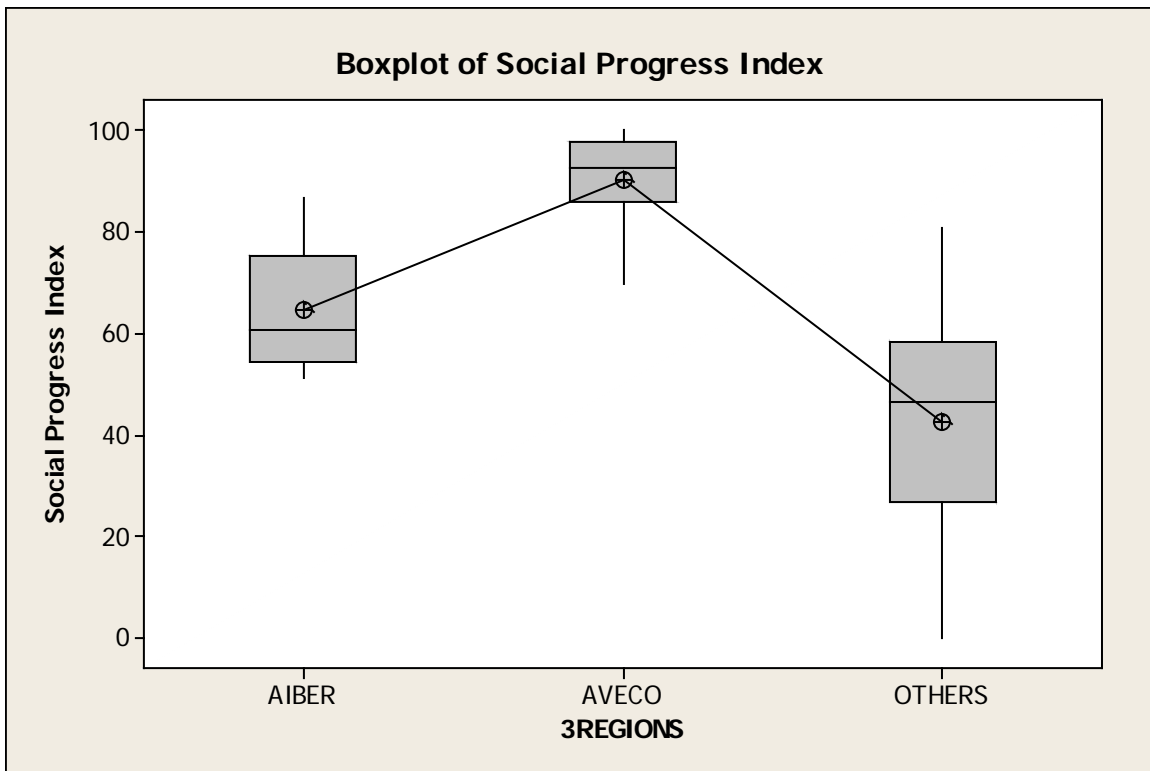
Source	DF	SS	MS	F	P
3REGIONS	2	48262	24131	83,97	0,000
Error	129	37071	287		
Total	131	85333			

S = 16,95    R-Sq = 56,56%    R-Sq(adj) = 55,88%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
AIBER	21	64,41	11,95
AVECO	27	90,30	8,70
OTHERS	84	42,63	19,71

Pooled StDev = 16,95



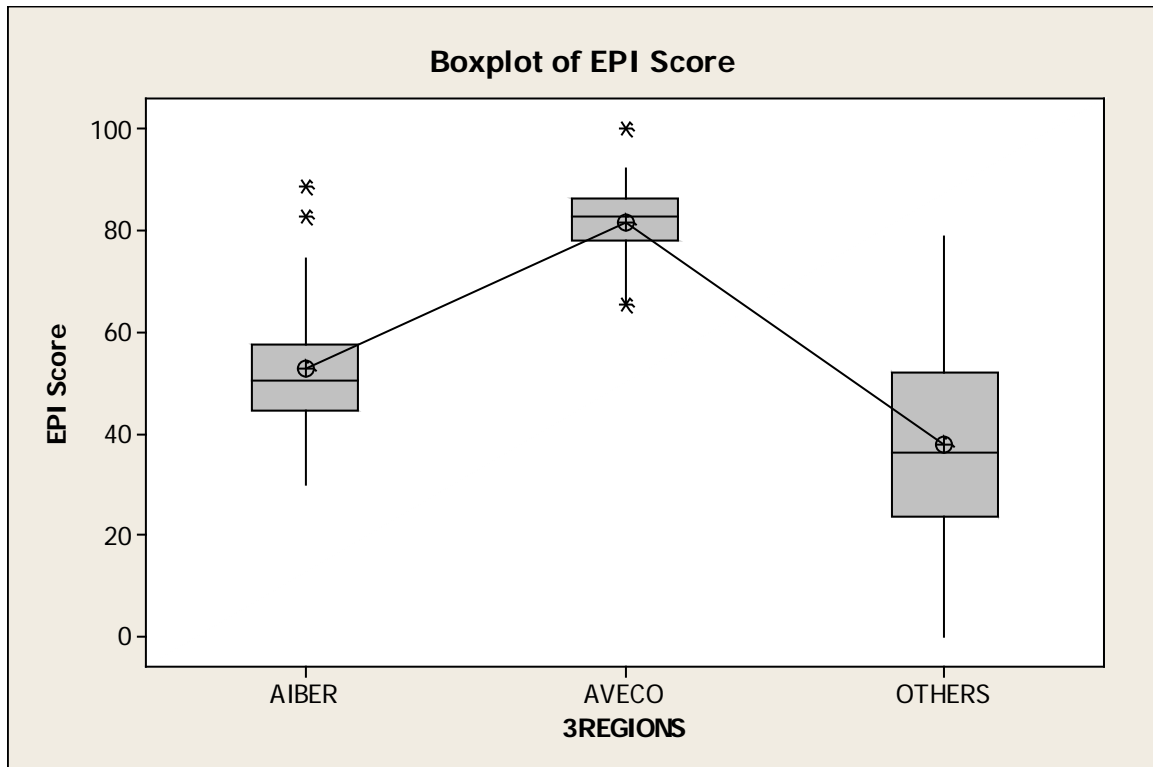
### One-way ANOVA: EPI Score versus 3REGIONS

Source	DF	SS	MS	F	P
3REGIONS	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

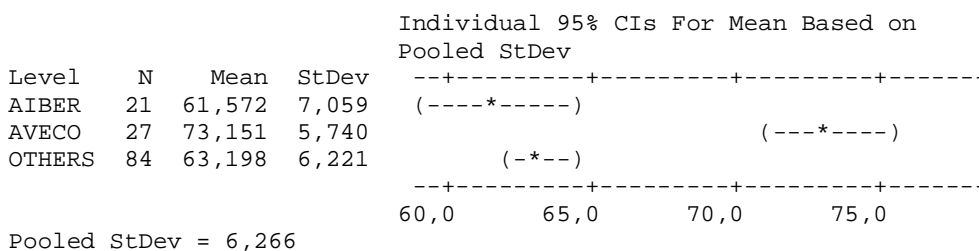




### One-way ANOVA: OHI versus 3REGIONS

Source	DF	SS	MS	F	P
3REGIONS	2	2313,3	1156,7	29,45	0,000
Error	129	5065,6	39,3		
Total	131	7379,0			

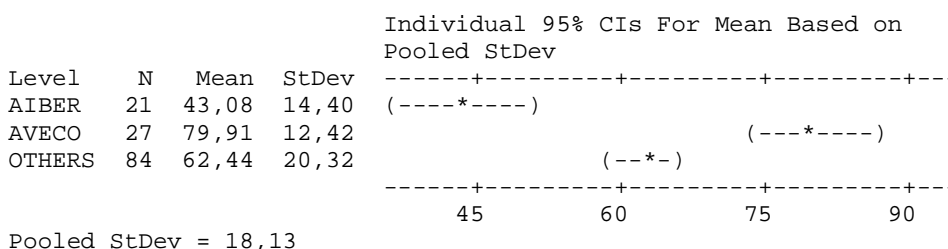
S = 6,266 R-Sq = 31,35% R-Sq(adj) = 30,29%



### One-way ANOVA: GINI Index versus 3REGIONS

Source	DF	SS	MS	F	P
3REGIONS	2	16078	8039	24,45	0,000
Error	129	42414	329		
Total	131	58492			

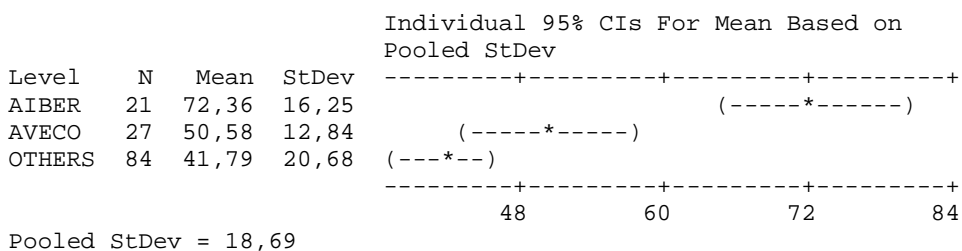
S = 18,13 R-Sq = 27,49% R-Sq(adj) = 26,36%



### One-way ANOVA: Happy Planet Index versus 3REGIONS

Source	DF	SS	MS	F	P
3REGIONS	2	15846	7923	22,69	0,000
Error	129	45048	349		
Total	131	60894			

S = 18,69 R-Sq = 26,02% R-Sq(adj) = 24,87%



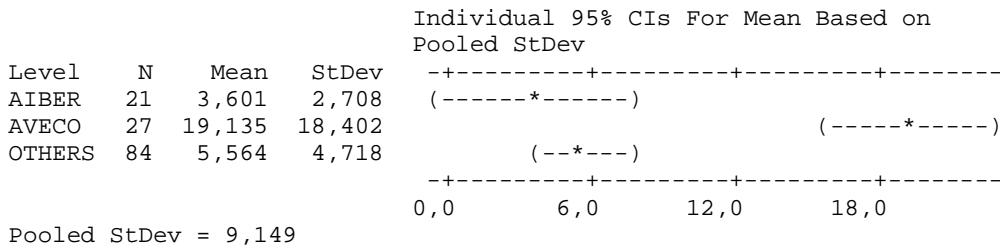
Na análise de variância percebe-se uma discriminação importante entre as médias das 3 regiões relativamente às variáveis sintéticas “Governança”, “Social Progress Index” e “EPI Score” (com valores F de 97, 83 e 69 respectivamente), ao passo as médias das variáveis



**One-way ANOVA: Electric power consumption (kWh versus 3REGIONS)**

Source	DF	SS	MS	F	P
3REGIONS	2	4252,7	2126,4	25,40	0,000
Error	129	10798,9	83,7		
Total	131	15051,6			

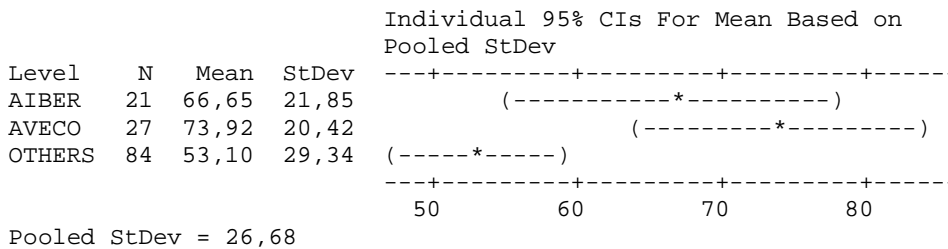
S = 9,149    R-Sq = 28,25%    R-Sq(adj) = 27,14%



**One-way ANOVA: Biodiversity and habitat (0=no versus 3REGIONS)**

Source	DF	SS	MS	F	P
3REGIONS	2	10137	5069	7,12	0,001
Error	129	91821	712		
Total	131	101958			

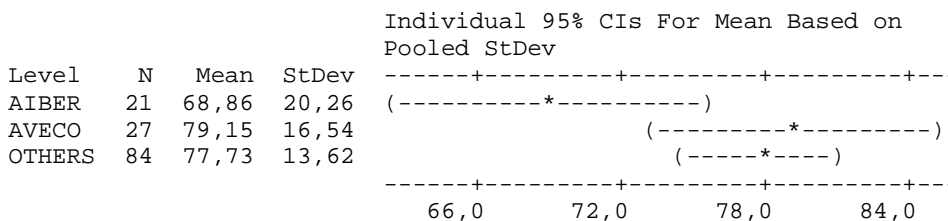
S = 26,68    R-Sq = 9,94%    R-Sq(adj) = 8,55%



**One-way ANOVA: Exposición (%) versus 3REGIONS**

Source	DF	SS	MS	F	P
3REGIONS	2	1543	771	3,24	0,042
Error	129	30716	238		
Total	131	32259			

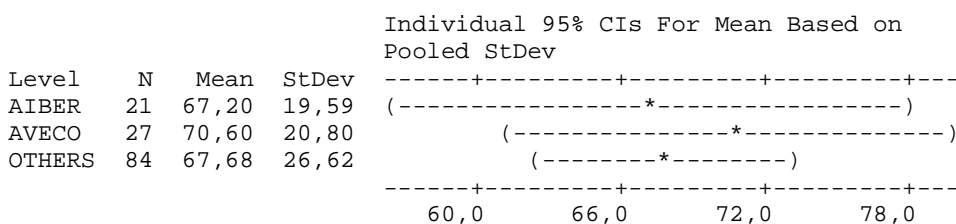
S = 15,43    R-Sq = 4,78%    R-Sq(adj) = 3,31%



Pooled StDev = 15,43 **One-way ANOVA: Fossil fuel energy consumption versus 3REGIONS**

Source	DF	SS	MS	F	P
3REGIONS	2	199	100	0,17	0,848
Error	129	77753	603		
Total	131	77953			

S = 24,55 R-Sq = 0,26% R-Sq(adj) = 0,00%



Pooled StDev = 24,55

Quanto às variáveis analíticas há um destaque com maior discriminação entre as médias dos 3 grupos – apenas a variável “**Vulnerabilidade**”, com valor de **F 57** e baixa sobreposição. Em seguida aparece “**Consumo de Energia Elétrica**”, porém já com um número F relativamente menor (**25**) e mais sobreposição devido à proximidade das médias entre AIBER e OUTROS.

#### 15.4 Quadro-resumo: F-Value

Variável	F-Value
<b>Governança</b>	<b>97,02</b>
<b>SPI</b>	<b>83,97</b>
<b>EPI</b>	<b>69,68</b>
<b>Vulnerabilidade</b>	<b>57,93</b>
<b>IDH</b>	<b>46,95</b>
<b>OHI</b>	<b>29,45</b>
<b>Electric power consumption</b>	<b>25,40</b>
<b>GINI</b>	<b>24,45</b>
<b>HPI</b>	<b>22,69</b>
<b>Biodiversity and habitat</b>	<b>7,12</b>
<b>Exposición</b>	<b>3,24</b>
<b>Fossil fuel energy consumption</b>	<b>0,17</b>

A análise comparativa de F mostra uma predominância de discriminação entre as variáveis sintéticas como “**Governança**”, “**Social Progress Index**”, “**EPI Score**” e “**IDH**”, ao passo que das variáveis analíticas apenas “**Vulnerabilidade**” aparece com maior discriminação nas médias dos três grupos de países. Isso dificulta a análise das diferenças de maneira relevante para o tema específico, para o qual as variáveis analíticas deveriam fazer toda a diferença.

## 15.5. ANÁLISE DISCRIMINANTE

### Discriminant Analysis: 3REGIONS versus Social Progr; Happy Planet; ...

#### Quadratic Method for Response: 3REGIONS

Predictors: Social Progress Index; Happy Planet Index; GINNI Index; OHI;  
Vulnerabilidad (%); Electric power consumption (kWh; Exposición (%))

Group	AIBER	AVECO	OTHERS
Count	21	27	84

#### Summary of classification

Put into Group	True Group		
	AIBER	AVECO	OTHERS
AIBER	20	0	5
AVECO	1	26	1
OTHERS	0	1	78
Total N	21	27	84
N correct	20	26	78
<b>Proportion</b>	<b>0,952</b>	<b>0,963</b>	<b>0,929</b>

N = 132

N Correct = 124

**Proportion Correct = 0,939**

#### From Generalized Squared Distance to Group

Group	AIBER	AVECO	OTHERS
AIBER	28,61	76,82	38,37
AVECO	103,84	30,73	49,30
OTHERS	60,92	94,63	33,02

#### Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
19**	AIBER	AVECO	AIBER	37,32	0,378
			AVECO	36,54	0,558
			OTHERS	40,85	0,065
26**	AVECO	OTHERS	AIBER	70,02	0,000
			AVECO	47,22	0,345
			OTHERS	45,93	0,655
50**	OTHERS	AIBER	AIBER	37,62	0,985
			AVECO	126,58	0,000
			OTHERS	45,97	0,015
70**	OTHERS	AIBER	AIBER	34,31	0,942
			AVECO	74,63	0,000
			OTHERS	39,87	0,058
71**	OTHERS	AVECO	AIBER	41,02	0,048
			AVECO	35,77	0,662
			OTHERS	37,42	0,290
74**	OTHERS	AIBER	AIBER	36,04	0,599
			AVECO	37,50	0,288
			OTHERS	39,37	0,113
78**	OTHERS	AIBER	AIBER	51,25	1,000
			AVECO	189,25	0,000
			OTHERS	66,53	0,000
80**	OTHERS	AIBER	AIBER	34,09	0,786
			AVECO	73,92	0,000
			OTHERS	36,69	0,214

A análise discriminante mostra um acerto relativamente alto, de 124 em 132 elementos totais, ou quase 94% correto. Ademais, nota-se uma distância quadrada considerável entre os três

grupos. AVECO apresenta distância sempre maior dos demais, enquanto AIBER e OUTROS apresentam maior proximidade relativa.

## 15.6 REGRESSÃO LOGÍSTICA

### Ordinal Logistic Regression: C43 versus Social Progr; Happy Planet; ...

Link Function: Logit

Response Information

Variable	Value	Count
C43	1	27
	2	21
	3	84
Total		132

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio
Const(1)	-23,5606	4,44302	-5,30	0,000	
Const(2)	-20,7574	4,15184	-5,00	0,000	
Social Progress Index	0,171840	0,0284971	6,03	0,000	1,19
Happy Planet Index	0,0737095	0,0192621	3,83	0,000	1,08
Exposición (%)	0,0351076	0,0185031	1,90	0,058	1,04
Fossil fuel energy consumption	-0,0356742	0,0175907	-2,03	0,043	0,96
OHI	0,0841231	0,0393663	2,14	0,033	1,09

95% CI

Predictor	Lower	Upper
Const(1)		
Const(2)		
Social Progress Index	1,12	1,26
Happy Planet Index	1,04	1,12
Exposición (%)	1,00	1,07
Fossil fuel energy consumption	0,93	1,00
OHI	1,01	1,18

Log-Likelihood = -48,351

Test that all slopes are zero: G = 142,135, DF = 5, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	123,683	257	1,000
Deviance	96,702	257	1,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	4397	95,6	Somers' D 0,91
Discordant	198	4,3	Goodman-Kruskal Gamma 0,91
Ties	4	0,1	Kendall's Tau-a 0,49
Total	4599	100,0	

A regressão mostra um índice de concordância acima de 95%, com predominância da variável “Social Progress Index”, seguida de “Happy Planet Index”.

## 15.7. ÁRVORES DE CLASSIFICAÇÃO

### Classification Tree

[DataSet1] C:\Users\jtfem\Desktop\material da aula Hoyos\_ Metodos\trabalhos Jorge\variaveis completas 8jun2015.sav

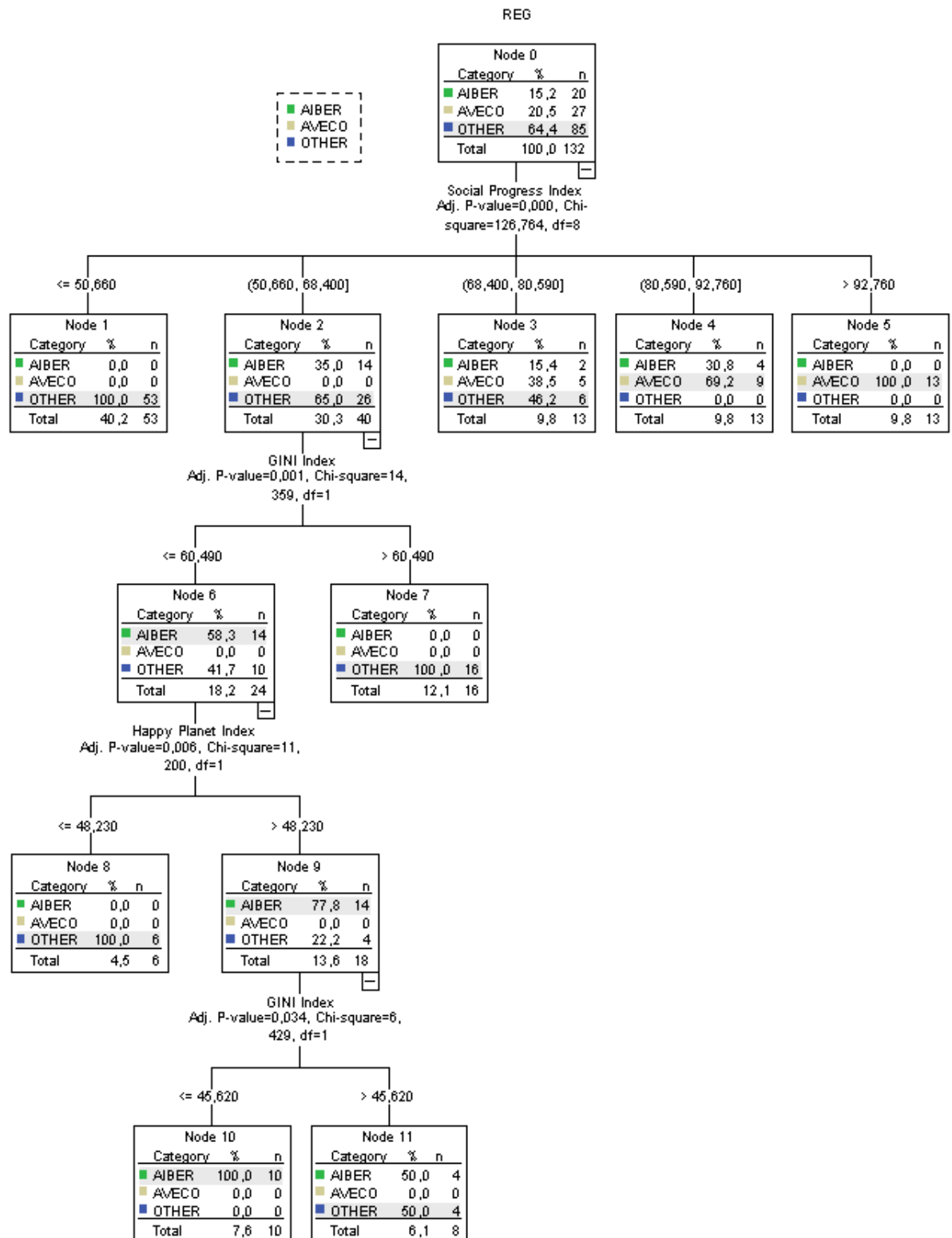
#### Warnings

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	REG	
	Independent Variables	Social Progress Index, Access to Electricity Supply, Biodiversidade e Habitat, Fossil Fuel Energy Consumption, Happy Planet Index, falta capac adap, GINI Index, Alternative and Nuclear Energy	
	Validation	None	
	Maximum Tree Depth		5
	Minimum Cases in Parent Node		10
	Minimum Cases in Child Node		5
Results	Independent Variables Included	Social Progress Index, GINI Index, Happy Planet Index	
	Number of Nodes		12
	Number of Terminal Nodes		8
	Depth		4



Risk



Estimate	Std. Error
,114	,028

Growing Method: CHAID

Dependent Variable: REG

#### Classification

Observed	Predicted			
	AIBER	AVECO	OTHER	Percent Correct
AIBER	10	4	6	50,0%
AVECO	0	22	5	81,5%
OTHER	0	0	85	100,0%
Overall Percentage	7,6%	19,7%	72,7%	88,6%

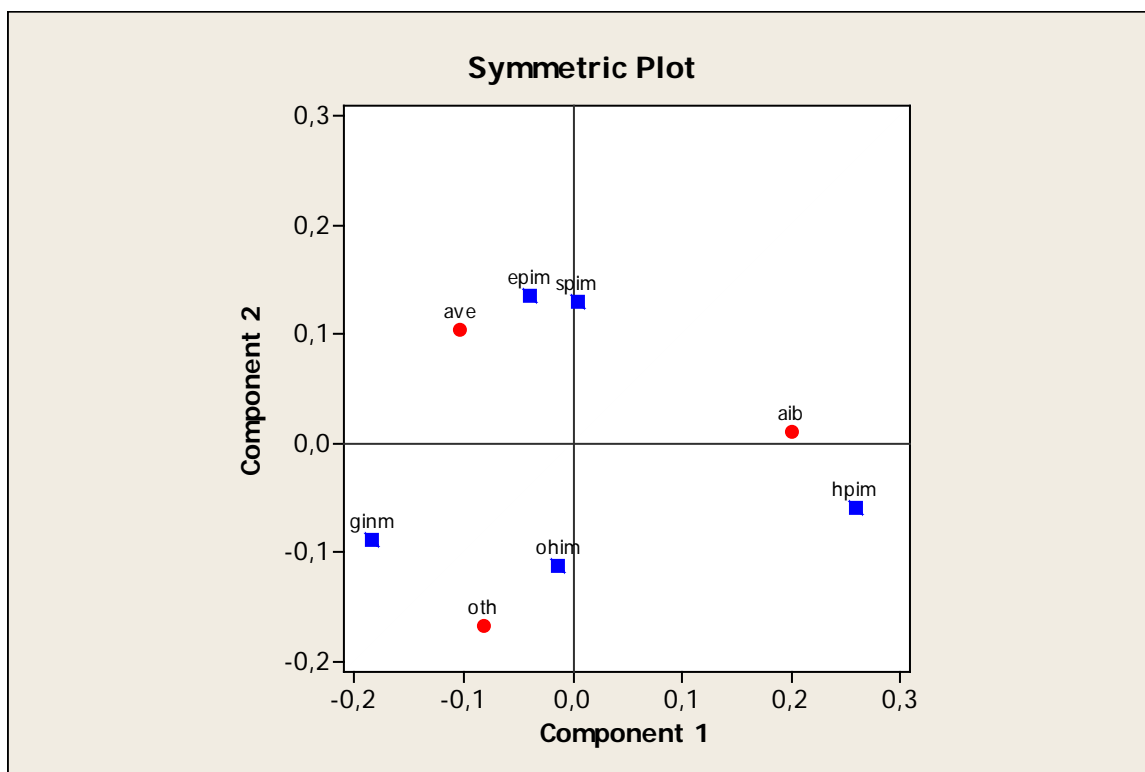
Growing Method: CHAID

Dependent Variable: REG

A árvore mostra um percentual de acerto considerável, acima de 88%, e novamente a variável que aparece no topo para ajudar a explicar o fenômeno é “**Social Progress Index**”. Na primeira ramificação vemos OTHERS predominando à esquerda (primeiro galho), em seguida dividindo com AIBER (segundo galho), e à medida que se avança em direção aos melhores índices ocorre uma predominância de AVECO, uma diminuição gradual de AIBER e o desaparecimento total de OTHERS. Fica patente a situação extrema vivida pelos países da AIBER, já que eles estão no segundo extremo melhor e pior da árvore, mostrando a enorme disparidade recorrente na região.

## 15.8 ANÁLISE DE CORRESPONDÊNCIA

### 15.8.1. Médias das Variáveis Sintéticas - 3 Regions



#### Simple Correspondence Analysis: SPIm; HPIm; GINIm; OHIm; EPIm

Analysis of Contingency Table

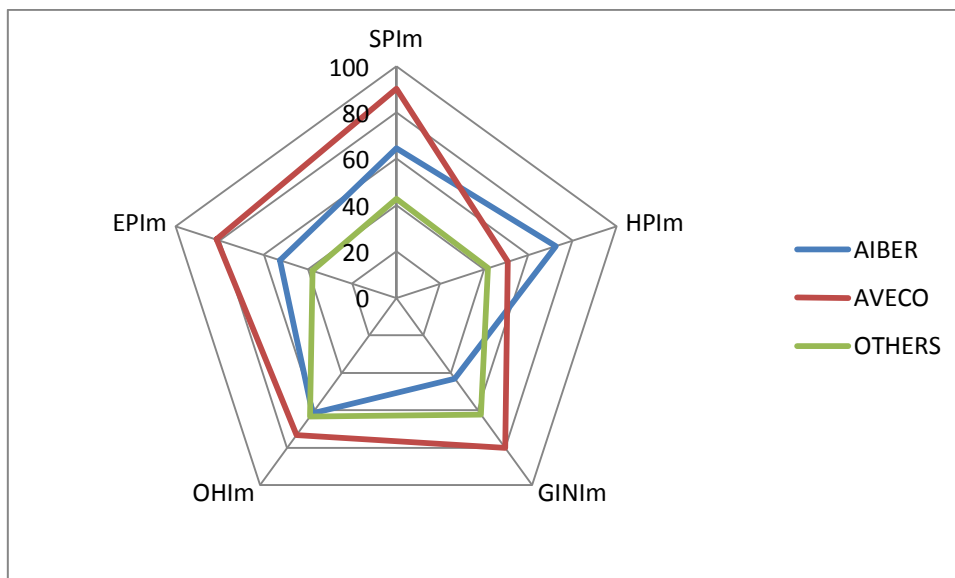
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0192	0,6142	0,6142	*****
2	0,0120	0,3858	1,0000	*****
Total	0,0312			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1		Component 2			
					Coord	Corr	Contr	Coord	Corr	Contr
1	aib	1,000	0,321	0,417	0,201	0,998	0,677	0,010	0,002	0,003
2	ave	1,000	0,409	0,280	-0,103	0,501	0,229	0,103	0,499	0,362
3	oth	1,000	0,270	0,303	-0,082	0,192	0,095	-0,168	0,808	0,635

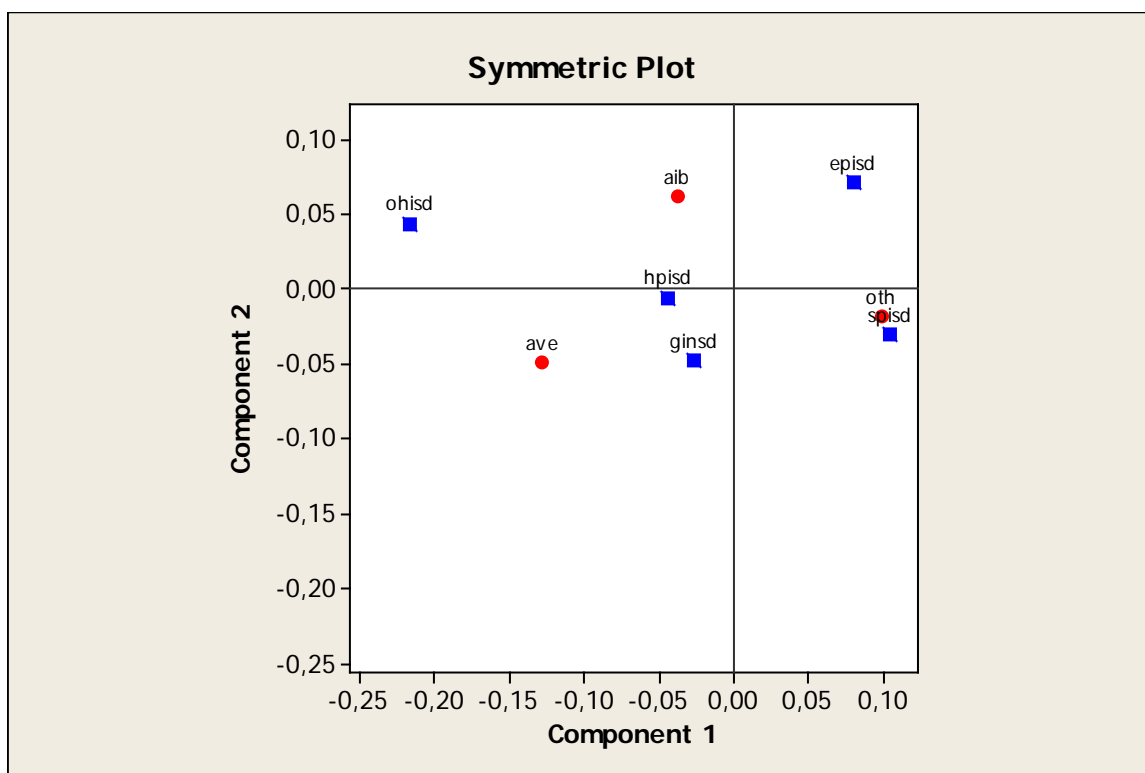
Column Contributions

ID	Name	Qual	Mass	Inert	Component 1		Component 2			
					Coord	Corr	Contr	Coord	Corr	Contr
1	spim	1,000	0,215	0,116	0,004	0,001	0,000	0,129	0,999	0,299
2	hpim	1,000	0,180	0,405	0,259	0,949	0,626	-0,060	0,051	0,053
3	ginm	1,000	0,202	0,271	-0,184	0,809	0,357	-0,089	0,191	0,134
4	ohim	1,000	0,216	0,090	-0,013	0,014	0,002	-0,113	0,986	0,229
5	epim	1,000	0,188	0,119	-0,039	0,077	0,015	0,135	0,923	0,284



A análise de correspondência mostra os países da AVECO associados aos índices mais robustos ligados a um importante desenvolvimento econômico e social, caso das variáveis **“EPI Score”** e **“Social Progress Index”**. AIBER aparece com a maior proximidade da variável **“Happy Planet Index”**, que é um índice composto por indicadores mais subjetivos, e tem o pior desempenho no **“GINI Index”**, relevando quanto a distribuição de renda é desigual na região.

### 15.8.2. Desvio Padrão das Variáveis Sintéticas - 3 Regiões



#### Simple Correspondence Analysis: SPISd; HPIsd; GINIsd; OHIsd; EPIsd

Analysis of Contingency Table

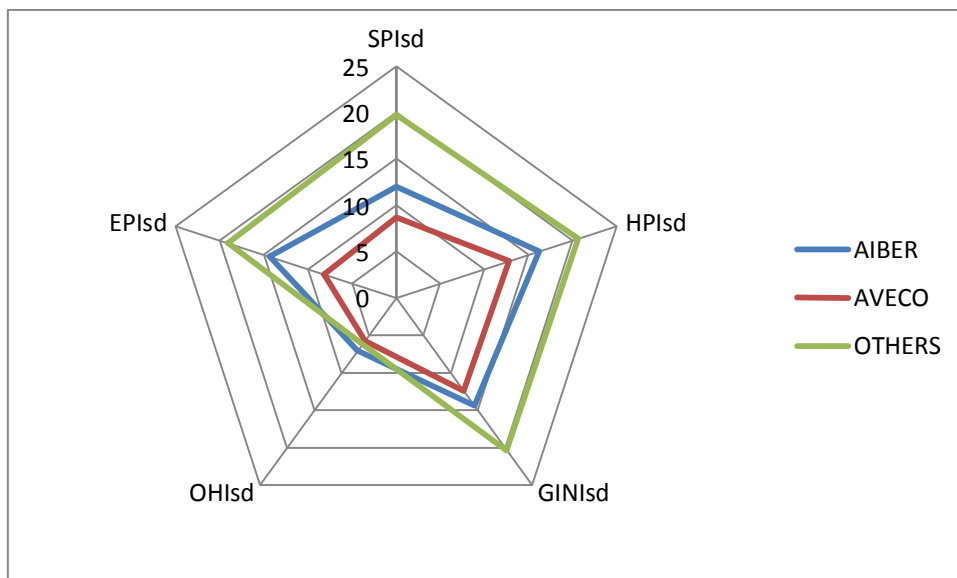
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0087	0,8152	0,8152	*****
2	0,0020	0,1848	1,0000	*****
Total	0,0107			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	aib	1,000	0,323	0,158	-0,038	0,271	0,053	0,062	0,729	0,624
2	ave	1,000	0,242	0,427	-0,128	0,869	0,456	-0,050	0,131	0,303
3	oth	1,000	0,435	0,414	0,099	0,967	0,492	-0,018	0,033	0,073

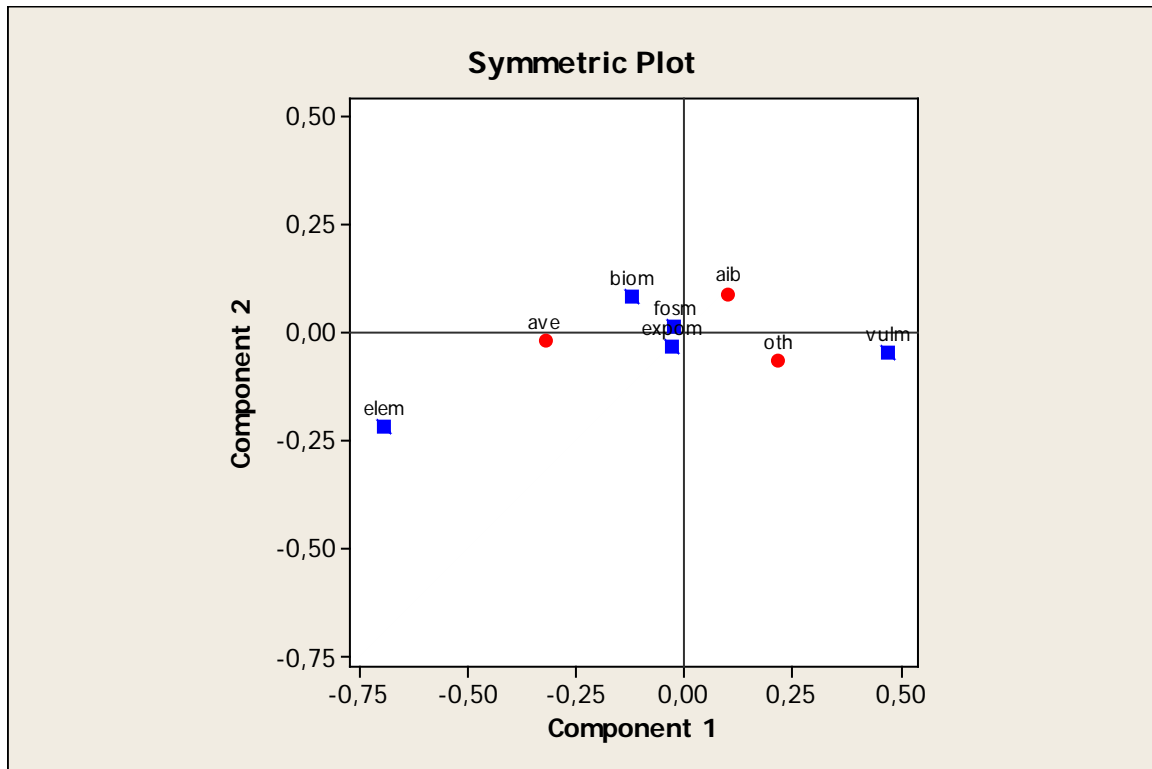
Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	spisd	1,000	0,204	0,225	0,104	0,921	0,254	-0,031	0,079	0,097
2	hpsid	1,000	0,252	0,046	-0,044	0,983	0,056	-0,006	0,017	0,004
3	ginsd	1,000	0,238	0,066	-0,026	0,235	0,019	-0,048	0,765	0,273
4	ohisd	1,000	0,096	0,438	-0,216	0,961	0,516	0,044	0,039	0,092
5	episd	1,000	0,210	0,225	0,080	0,562	0,155	0,071	0,438	0,533



Quanto aos desvios padrão, novamente há poucas surpresas: AVECO apresenta o menor desvio padrão em todas as variáveis sintéticas analisadas, AIBER ocupa uma posição intermediária, e OTHERS demonstra maior variabilidade, com exceção de “Ocean Health Index” na qual AIBER mostra variabilidade relativamente maior que OTHERS.

### 15.8.3. Médias de Variáveis Analíticas - 3 Regiões



#### Simple Correspondence Analysis: VULm; FOSFm; ELEPm; BIODm; EXPOM

Analysis of Contingency Table

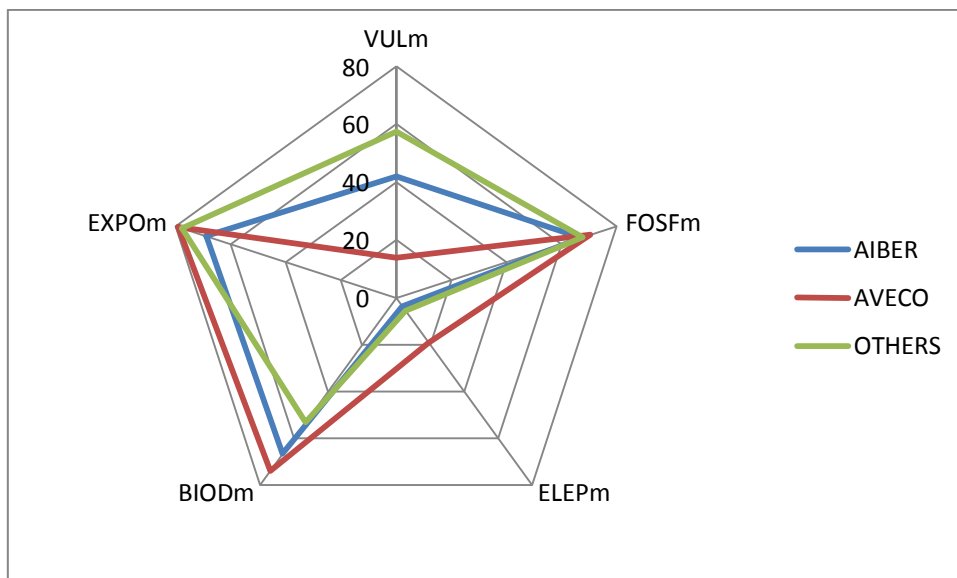
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0544	0,9276	0,9276	*****
2	0,0042	0,0724	1,0000	**
Total	0,0586			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	aib	1,000	0,324	0,103	0,103	0,569	0,063	0,090	0,431	0,613
2	ave	1,000	0,335	0,594	-0,322	0,997	0,638	-0,018	0,003	0,027
3	oth	1,000	0,341	0,303	0,218	0,914	0,299	-0,067	0,086	0,360

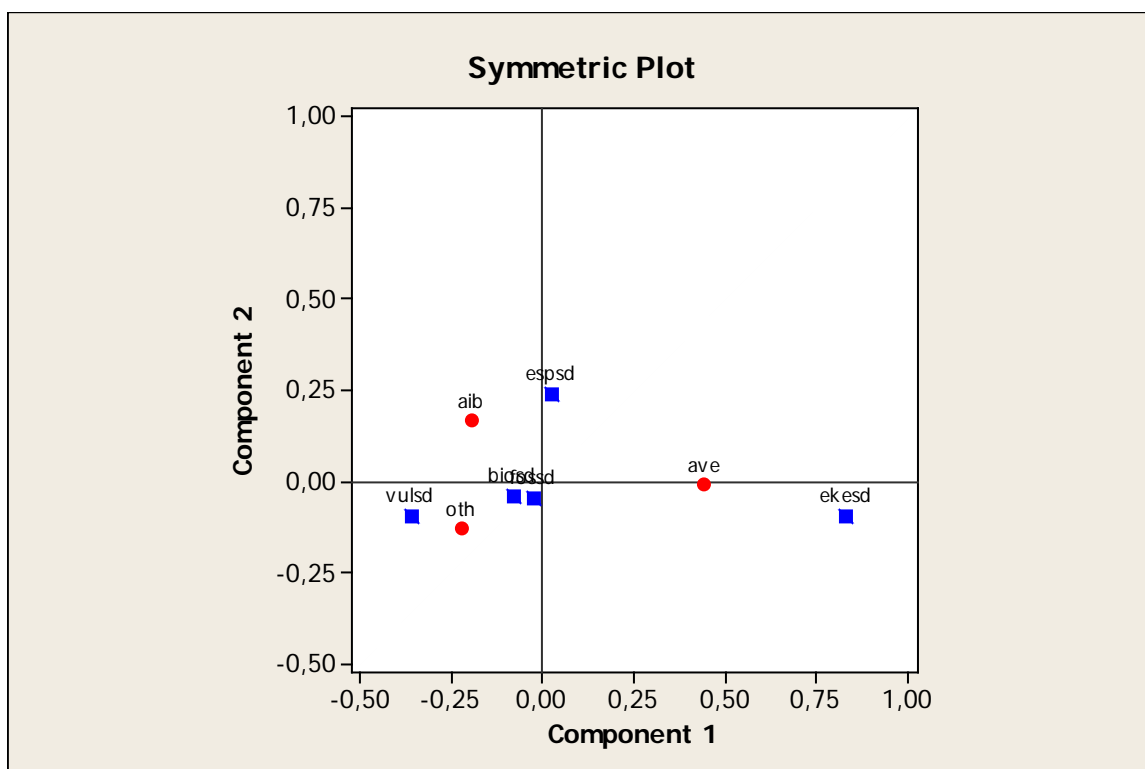
Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	vulm	1,000	0,147	0,564	0,471	0,990	0,602	-0,047	0,010	0,078
2	fosm	1,000	0,268	0,003	-0,022	0,705	0,002	0,014	0,295	0,013
3	elem	1,000	0,037	0,333	-0,693	0,909	0,327	-0,219	0,091	0,416
4	biom	1,000	0,253	0,091	-0,119	0,668	0,065	0,084	0,332	0,416
5	expom	1,000	0,295	0,009	-0,027	0,399	0,004	-0,033	0,601	0,078



No caso das variáveis analíticas, a análise de correspondência mostra novamente os países da AVECO mostrando os melhores índices relacionados a “**Fossil Fuel Energy Consumption**”, “**Biodiversidade e Habitat**”, “**Exposição**”. O grupo de países AIBER, juntamente com o grupo dos OTHERS aparecem associados à variável “**Vulnerabilidade**”, revelando a carência de infraestrutura dos países da região.

### 15.8.4 .Desvios Padrão de Variáveis Analíticas - 3 Regiões



### Simple Correspondence Analysis: VULsd; FOSFsd; ELEPsd; BIODsd; EXPOsd

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0934	0,8649	0,8649	*****

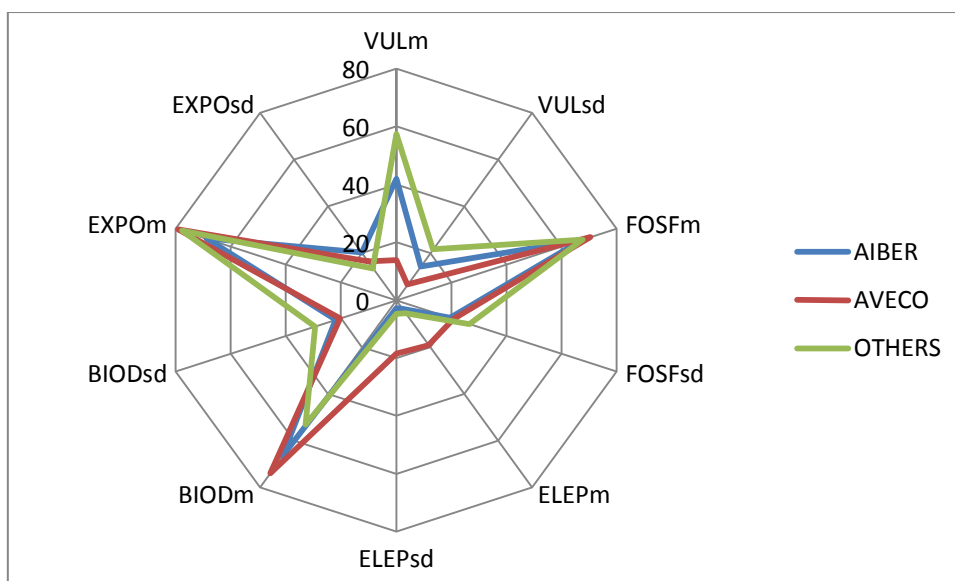
```

2      0,0146      0,1351      1,0000      ****
Total  0,1079

```

## Row Contributions

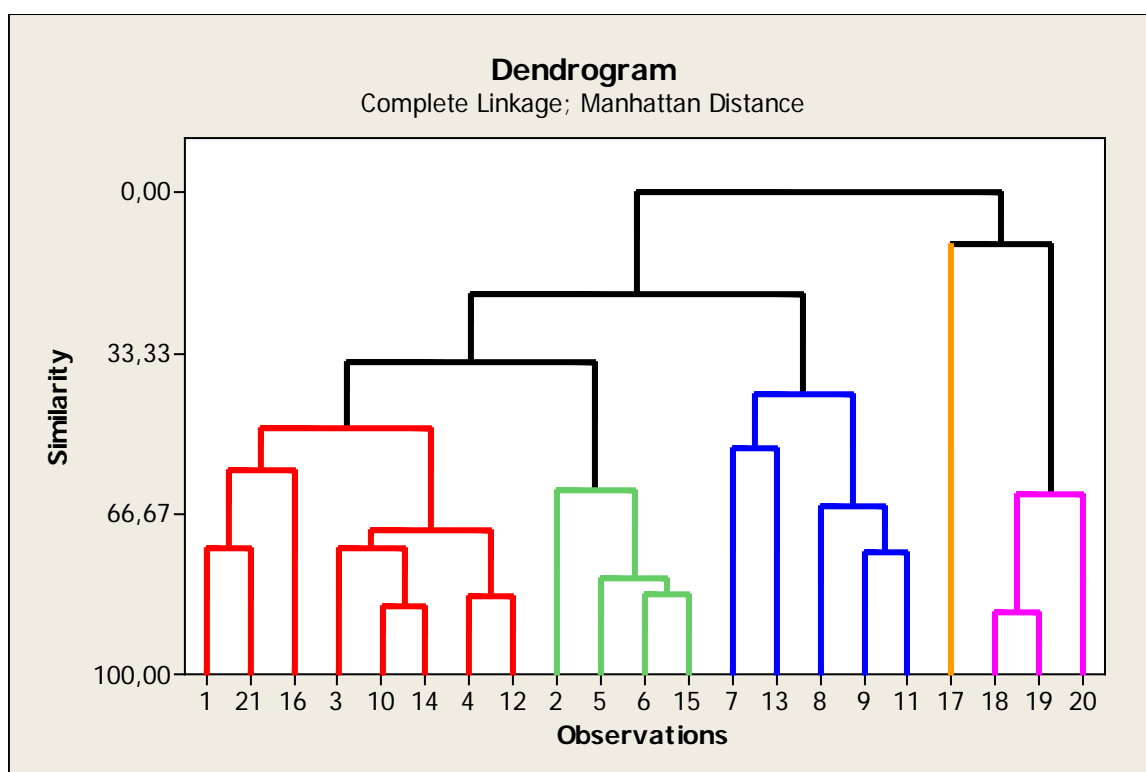
ID	Name	Qual	Mass	Inert	Component 1		Component 2			
					Coord	Corr	Contr	Coord	Corr	Contr
1	aib	1,000	0,306	0,184	-0,194	0,580	0,123	0,165	0,420	0,571
2	ave	1,000	0,321	0,586	0,444	1,000	0,677	-0,007	0,000	0,001
3	oth	1,000	0,373	0,230	-0,223	0,749	0,199	-0,129	0,251	0,428



A comparação das variáveis analíticas em cada região por meio de suas medias e desvios-padrão confirmam as análises anteriores, de que AIBER, quando não se encontra em posição intermediária para baixo está na pior posição dos três grupos. Isso ocorre em função de não haver uma obrigatoriedade ou compromisso de uniformidade nas políticas e ações públicas como existe, por exemplo, na maioria dos países da AVECO (União Europeia).



## 16. DENDROGRAMA - AIBER



	<b>PIBER</b>	<b>PIBERcl</b>	<b>PRIBERncl</b>	<b>PRIBERord</b>
<b>17</b>	<b>Costa Rica</b>	<b>4</b>	<b>3</b>	<b>1</b>
<b>18</b>	<b>Portugal</b>	<b>5</b>	<b>3</b>	<b>1</b>
<b>19</b>	<b>Spain</b>	<b>5</b>	<b>3</b>	<b>1</b>
<b>20</b>	<b>Uruguay</b>	<b>5</b>	<b>3</b>	<b>1</b>
<b>3</b>	<b>Brazil</b>	<b>1</b>	<b>2</b>	<b>2</b>
<b>4</b>	<b>Colombia</b>	<b>1</b>	<b>2</b>	<b>2</b>
<b>8</b>	<b>Guatemala</b>	<b>3</b>	<b>2</b>	<b>2</b>
<b>9</b>	<b>Honduras</b>	<b>3</b>	<b>2</b>	<b>2</b>
<b>12</b>	<b>Panama</b>	<b>1</b>	<b>2</b>	<b>2</b>
<b>16</b>	<b>Chile</b>	<b>1</b>	<b>2</b>	<b>2</b>
<b>1</b>	<b>Argentina</b>	<b>1</b>	<b>1</b>	<b>3</b>
<b>2</b>	<b>Bolivia</b>	<b>2</b>	<b>1</b>	<b>3</b>
<b>5</b>	<b>Domin Repub</b>	<b>2</b>	<b>1</b>	<b>3</b>
<b>6</b>	<b>Ecuador</b>	<b>2</b>	<b>1</b>	<b>3</b>
<b>7</b>	<b>El Salvador</b>	<b>3</b>	<b>1</b>	<b>3</b>
<b>10</b>	<b>Mexico</b>	<b>1</b>	<b>1</b>	<b>3</b>
<b>11</b>	<b>Nicaragua</b>	<b>3</b>	<b>1</b>	<b>3</b>

<b>13</b>	<b>Paraguay</b>	<b>3</b>	<b>1</b>	<b>3</b>
<b>14</b>	<b>Peru</b>	<b>1</b>	<b>1</b>	<b>3</b>
<b>15</b>	<b>Venezuela</b>	<b>2</b>	<b>1</b>	<b>3</b>
<b>21</b>	<b>Cuba</b>	<b>1</b>	<b>1</b>	<b>3</b>

O dendrograma dividiu a região em 5 grupos (PIBERcl), porém, usando o mesmo princípio empregado para o total dos 132 países, buscamos a divisão em três clusters (PIBERncl), e por fim, invertemos a numeração entre os grupos 1 e 3 (PIBERord), já que o grupo 3 é composto pelos países que costumam apresentar os melhores índices, comparáveis com o de seus pares no grupo AVECO. O mapa abaixo mostra também a divisão por cores, sendo

-  Grupo 1
-  Grupo 2
-  Grupo 3



## 17. ANÁLISE DISCRIMINANTE – AIBER

### Discriminant Analysis: PIBERncl versus Social Progr; Happy Planet; ...

Linear Method for Response: PIBERncl

Predictors: Social Progress Index; Happy Planet Index; GINNI Index; OHI;  
Vulnerabilidad (%); Electric power consumption (kWh; Exposición (%))

Group	1	2	3
Count	11	6	4

Summary of classification

Put into Group	True Group		
	1	2	3
1	11	0	0
2	0	6	0
3	0	0	4
Total N	11	6	4
N correct	11	6	4
Proportion	1,000	1,000	1,000

N = 21

N Correct = 21

Proportion Correct = 1,000

Squared Distance Between Groups

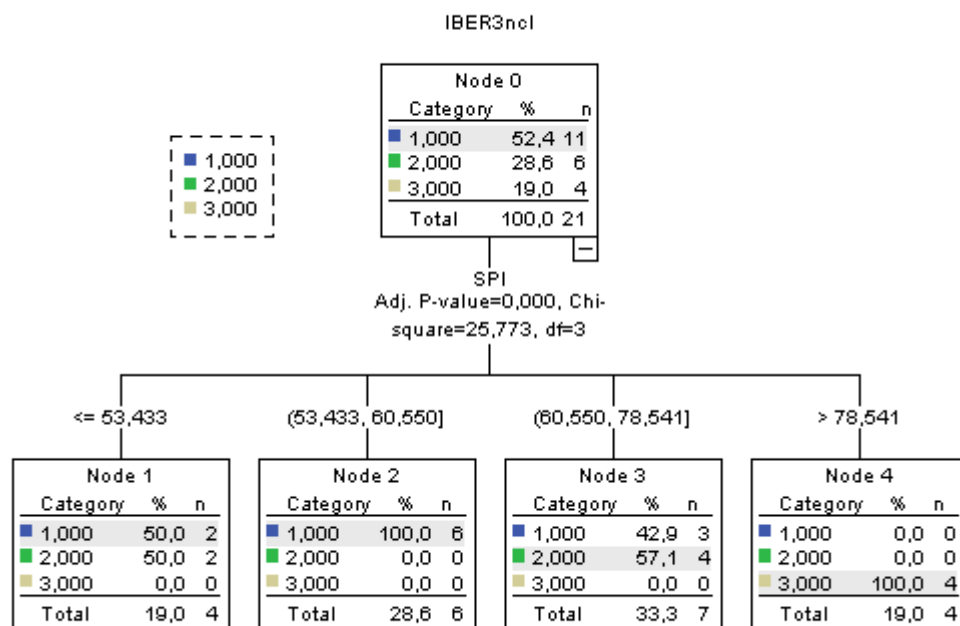
	1	2	3
1	0,0000	10,5979	68,1971
2	10,5979	0,0000	68,9282
3	68,1971	68,9282	0,0000

Linear Discriminant Function for Groups			
	1	2	3
Constant	-432,44	-419,21	-576,85
Social Progress Index	6,31	6,43	8,13
Happy Planet Index	1,83	1,80	1,73
GINNI Index	2,11	1,64	1,92
OHI	19,60	20,39	31,26
Vulnerabilidad (%)	3,21	3,16	3,27
Electric power consumption (kWh	-5,30	-4,59	-7,18
Exposición (%)	1,29	1,26	1,28

A análise discriminante mostra um acerto de 100% na distribuição dos países pelos três grupos. Nota-se também, pela distância quadrada entre os grupos, maior proximidade entre os grupos 1 e 2 (na reordenação, seriam 3 e 2) e maior distanciamento entre os grupos 1 e 3 e entre os grupos 2 e 3 (na reordenação, seriam 2 e 1). Em outras palavras, olhando a reordenação o grupo 1 está sempre isolado com os países mais bem posicionados em quase todos os índices, enquanto os grupos 2 e 3 estão mais similares em seus índices.

## 18. ARVORE DE CLASSIFICAÇÃO - AIBER

Model Summary			
Specifications	Growing Method	EXHAUSTIVE CHAID	
	Dependent Variable	IBER3ncl	
	Independent Variables	SPI, HPI, GINI, OHI, VUL, ELEC, EXPO	
	Validation	None	
	Maximum Tree Depth	5	
	Minimum Cases in Parent Node	5	
	Minimum Cases in Child Node	2	
	Results	Independent Variables Included	SPI
		Number of Nodes	5
		Number of Terminal Nodes	4
Depth		1	



#### Classification

Observed	Predicted			
	1,00	2,00	3,00	Percent Correct
1,00	8	3	0	72,7%
2,00	2	4	0	66,7%
3,00	0	0	4	100,0%
<b>Overall Percentage</b>	<b>47,6%</b>	<b>33,3%</b>	<b>19,0%</b>	<b>76,2%</b>

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: IBER3ncl

A árvore de classificação mostra os melhores países (somente 4) isolados na extrema direita, ao passo que os outros dois grupos se dividem entre as outras posições, revelando que há uma grande variação, ou seja baixa uniformidade do índice Social Progress Index entre os 21 países. Os outros dois grupos, comendo 17 países, se distribuem quase igualmente nos outros ramos da árvore, mostrando que não há uma predominância. Dos 21 países, 13 se encontram no intervalo entre 53 e 78. Os 8 restantes se dividem igualmente no valor abaixo de 53 (Nicarágua, Guatemala, Honduras e Cuba) ou acima de 78 (Espanha, Portugal, Costa Rica e Uruguai). Chile está no limite, mas acabou entrando no grupo 2.

## 19. ANÁLISE DE VARIÂNCIA – AIBER

### One-way ANOVA: Social Progress Index versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	1970,7	985,3	20,01	0,000
Error	18	886,1	49,2		
Total	20	2856,8			

S = 7,016    R-Sq = 68,98%    R-Sq(adj) = 65,53%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	4	83,627	3,125	(-----*-----)
2	6	63,845	10,883	(-----*-----)
3	11	57,725	5,144	(-----*-----)

-----+-----+-----+-----+-----  
60                      70                      80                      90

Pooled StDev = 7,016

Na análise de variância de “**Social Progress Index**” somente com o grupo da AIBER, vemos os mesmos quatros países mencionados acima (Espanha, Portugal, Costa Rica e Uruguai) com média comparável a dos países da AVECO; os outros 17 países apresentam médias relativamente próximas entre si, o que também se pode notar, visualmente, pela sobreposição dos intervalos.

### One-way ANOVA: Happy Planet Index versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	1315	658	2,98	0,076
Error	18	3967	220		
Total	20	5282			

S = 14,84    R-Sq = 24,90%    R-Sq(adj) = 16,55%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
1	4	57,75	28,76	(-----*-----)
2	6	81,09	6,08	(-----*-----)
3	11	72,90	11,41	(-----*-----)

-----+-----+-----+-----+-----  
45                      60                      75                      90

Pooled StDev = 14,84

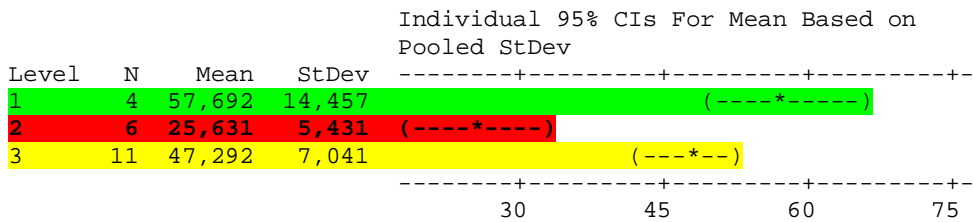
Na análise de variância de “**Happy Planet Index**” apenas com grupo de países da AIBER, notamos uma diferença importante: os quatro países que normalmente ocupam o topo dos resultados positivos, nessa variável apresentam o resultado mais baixo. Talvez isso se explique pela relativa subjetividade do índice ou pelo alto grau de exigência e visão crítica da população desses países. Os seis países mais “felizes”, segundo esse índice são Costa Rica,

Colômbia, El Salvador, Panamá, Nicarágua, Venezuela. Os quatro mais “tristes” são Espanha, Bolívia, Uruguai e Portugal.

**One-way ANOVA: GINI Index versus PRIBERord**

Source	DF	SS	MS	F	P
PRIBERord	2	2876,1	1438,0	20,38	0,000
Error	18	1270,2	70,6		
Total	20	4146,3			

S = 8,401 R-Sq = 69,36% R-Sq(adj) = 65,96%



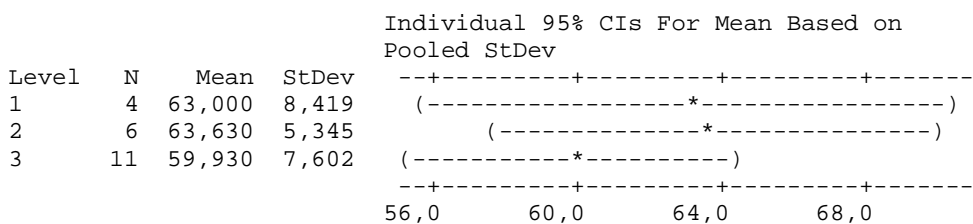
Pooled StDev = 8,401

Na análise de variância de “GINI Index” somente com os países da AIBER, temos algumas descobertas interessantes. No topo aparecem Espanha, Portugal, Cuba e Uruguai como os países mais igualitários, enquanto Panamá, Chile, Guatemala, Brasil, Colômbia e Honduras figuram no extremo oposto. Destaque positivo para Cuba nos países mais igualitários, provavelmente resultado de sua política social inclusiva, e Chile, Brasil e Colômbia como destaque negativo. O Brasil já era conhecido pela alta desigualdade social e de concentração de renda nas mãos de poucos. O Chile, porém, que muitas vezes é citado como exemplo do país mais desenvolvido da América Latina, não é motivo de orgulho nesse quesito.

**One-way ANOVA: OHI versus PRIBERord**

Source	DF	SS	MS	F	P
PRIBERord	2	63,2	31,6	0,61	0,554
Error	18	933,3	51,9		
Total	20	996,6			

S = 7,201 R-Sq = 6,34% R-Sq(adj) = 0,00%



Pooled StDev = 7,201

Na análise de variância de “**Ocean Health Index**” somente com os países da AIBER, vemos que ocorre muita sobreposição, ou seja, pouca discriminação entre os três grupos. De qualquer forma, vemos que os quatro primeiros incluem Mexico, Costa Rica, Colombia e Guatemala, enquanto os últimos incluem Honduras, Dominicana, Uruguai, Nicarágua e Paraguai. Vale ressaltar que Paraguai não tem acesso ao mar, razão pela qual ocupa o último lugar no ranking.

### One-way ANOVA: 1. Governança versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	4169	2084	12,87	0,000
Error	18	2915	162		
Total	20	7084			

S = 12,73    R-Sq = 58,85%    R-Sq(adj) = 54,27%

Individual 95% CIs For Mean Based on Pooled StDev					
Level	N	Mean	StDev		
1	4	69,29	3,92	(-----*-----)	
2	6	44,34	19,85	(-----*-----)	
3	11	31,72	9,48	(-----*-----)	

-----+-----+-----+-----+-----  
30            45            60            75

Pooled StDev = 12,73

A análise de variância de “**Governança**” somente com os 21 países da AIBER traz resultados interessantes. Mais da metade dos países se encontra na faixa mais baixa, o que significa que tem uma política instável de governança. São eles: Colômbia, Dominicana, Argentina, Nicarágua, Bolívia, Equador, Cuba, Guatemala, Paraguai, Honduras e Venezuela. No extremo positivo aparecem, em primeiro lugar, Chile, seguido de Portugal, Espanha e Uruguai. Brasil aparece no grupo intermediário, porém provavelmente entraria no grupo dos 11 após os escândalos de corrupção noticiados nos últimos meses.

### One-way ANOVA: Electric power consumption (kWh versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	58,20	29,10	5,92	0,011
Error	18	88,52	4,92		
Total	20	146,72			

S = 2,218    R-Sq = 39,67%    R-Sq(adj) = 32,97%

Individual 95% CIs For Mean Based on Pooled StDev					
Level	N	Mean	StDev		
1	4	7,012	3,293	(-----*-----)	
2	6	3,077	2,217	(-----*-----)	
3	11	2,647	1,773	(-----*-----)	

-----+-----+-----+-----+-----



2,5            5,0            7,5            10,0

Pooled StDev = 2,218

A análise de variância de **“Consumo de Eletricidade”** somente com os países da AIBER não apresenta grandes surpresas: os países com melhores condições econômicas são aqueles que dão mais acesso e, portanto, o consumo per capita é maior. Destacam-se positivamente Espanha, Portugal, Chile e Venezuela, uma surpresa; e no outro extremo, com baixo consumo, provavelmente devido ao pouco acesso disponível, encontram-se El Salvador, Honduras, Bolívia, Guatemala e Nicarágua.

### One-way ANOVA: Vulnerabilidad (%) versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	1631	816	5,95	0,010
Error	18	2468	137		
Total	20	4099			

S = 11,71    R-Sq = 39,79%    R-Sq(adj) = 33,10%

Level	N	Mean	StDev
1	4	23,67	7,52
2	6	46,59	13,32
3	11	45,83	11,88

Individual 95% CIs For Mean Based on Pooled StDev

Pooled StDev = 11,71

A análise de variância de **“Vulnerabilidade”** apenas com os países do grupo da AIBER mostra uma grande sobreposição de 17 países, e apenas 4 com uma certa discriminação. Esses quatro países, que apresentam um grau de vulnerabilidade mais baixo, são Espanha, Portugal, Uruguai e Chile. Curioso que Cuba viria logo em seguida. No outro extremo, com vulnerabilidade mais alta aparecem Guatemala, Bolívia, Nicarágua, Honduras e Paraguai.

### One-way ANOVA: Exposición (%) versus PRIBERord

Source	DF	SS	MS	F	P
PRIBERord	2	323	161	0,37	0,697
Error	18	7889	438		
Total	20	8211			

S = 20,93    R-Sq = 3,93%    R-Sq(adj) = 0,00%

Level	N	Mean	StDev
1	4	68,13	32,17
2	6	63,15	20,93
3	11	72,23	16,11

Individual 95% CIs For Mean Based on Pooled StDev

Pooled StDev = 20,93

A análise de variância de “**Exposição**” apenas no grupo de países da AIBER apresenta muita sobreposição, não permitindo fazer discriminação expressiva entre os três grupos. De toda forma, os países que aparecem como tendo menos índice de exposição são Paraguai, Bolívia, Brasil e Argentina, e no outro extremo aparecem Nicarágua, Chile, El Salvador, Guatemala e Costa Rica, provavelmente devido à sua posição costeira sujeita a vulcanismos, abalos sísmicos e outros desastres naturais.

## 20. CONSIDERAÇÕES FINAIS

Vemos que os países desenvolvidos, sobretudo da Europa Ocidental, são os mais bem posicionados, enquanto os países africanos estão bastante atrás tanto nos índices das variáveis sintéticas quanto no das analíticas. Em linhas gerais, os países ibero-americanos se encontram quase sempre em uma posição intermediária, muitas vezes puxada por Portugal, Espanha, Chile, Costa Rica e Uruguai – excluídos tais países, muito provavelmente a região estaria em um patamar bem pior. É o que vemos quando analisamos o grupo de países da AIBER em separado. De um lado, vemos países como Chile e Uruguai, que quase sempre apresentam índices econômicos e sociais similares aos dos países da AVECO – mesma situação de Portugal e Espanha, que no caso se beneficiam de serem membros da União Europeia, tendo que estar alinhados em relação a essas políticas. Na realidade, parece-nos que a AIBER como região guarda similaridades com o grupo de 132 países e replica, ela sozinha, como num microcosmo, as condições do macrocosmo – com isso, de um lado temos países bem posicionados, como os citados, mas de outro temos países em condições muito parecidas com as daqueles que figuram nas últimas posições da tabela, caso dos países da África subsaariana, por exemplo. Mesmo nos casos relativamente bem-sucedidos, como o do Chile e em menor grau, o Brasil, nota-se que alguns índices macroeconômicos positivos escondem problemas graves, como o da desigualdade de renda, enquanto Cuba e Venezuela, muitas vezes execradas, aparecem positivamente no topo do ranking em alguns casos. Isso torna mais evidente o fato de que os números são frios e para que possamos ter um parecer mais crítico é necessária a análise comparada, no detalhe, às vezes isolando um dado, ou agrupando alguns de forma a poder inferir os resultados, pois tais números, conforme o tratamento que lhe sejam aplicados, podem trazer a luz na mesma medida em que podem trazer as trevas, mais ainda na América Latina, cujas instituições são muitas vezes de independência, no mínimo, questionável.