

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO**  
Programa de Pós Graduação em Economia e Administração FEA – PUC-SP



**GPS**

**TEMA 6. O FUTURO DA EDUCAÇÃO E O  
TRABALHO**

## 1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados elaborados:

Pelo Núcleo de Estudos do Futuro (NEF), do Guia para a Gestão Pública Sustentável o “GPS”, que servirá como “mapa do caminho” que orientará as equipes das secretarias responsáveis pelas gestões locais a elaborarem um diagnóstico, um prognóstico e um plano diretor com suas prioridades estratégicas e um plano de metas centrado no desenvolvimento sustentável, para e a partir do contexto local, mais sem esquecer o regional. Para tal, iniciamos com o entendimento dos dados, incluindo a definição dos temas e das variáveis, suas classificações em variáveis categóricas ou quantitativas, os significados e unidades de medida, além da apresentação da tabela de dados. Na sequência, analisamos cada uma das variáveis separadamente quanto a sua forma de distribuição, os valores atípicos, medidas de centro e dispersão. Para tal contamos com o auxílio de gráficos (*pie chart*, barras, histogramas, gráficos de ramos, box-plot, dot-plot e curvas de densidade) e de medidas numéricas (média, mediana, quartis, desvio-padrão, variância, intervalo de confiança e teste de normalidade de Anderson-Darling). No final, buscamos comparar as análises efetuadas para cada variável. O software estatístico utilizado é o **MINITAB**.

O presente trabalho tem por objetivo efetuar uma análise das correlações e dos componentes principais (análise multivariada) de dados dimensionadores do nível Adult literacy rate (% of pop.aged), Primary school enrollment(% of children), Lower secondary school enrollment(% of children) diversos países do mundo separados por OECD, Outros e Todos. Para tal, iniciamos com análise da estatística descritiva. Em seguida passamos para a análise das correlações e dendogramas. Na terceira parte utilizamos a análise dos componentes principais.

## **2. ENTENDENDO OS DADOS**

### **2.1 Os Temas**

Os temas desta análise são baseados no Observatório da Rede Ibero-americana de Prospectiva (ORIBER), discriminados na tabela abaixo:

O tema que será analisado por mim, é O Futuro da Educação e o Trabalho.

<b>TEMAS</b>	
1	<b>RECURSOS BÁSICOS: ÁGUA, ALIMENTO E ENERGÍA</b>
2	<b>MORADIA</b>
3	<b>SEGURANÇA E PAZ</b>
4	<b>SAÚDE E MEIO AMBIENTE</b>
5	<b>DESIGUALDADE E INCLUSÃO SOCIAL E GÊNERO</b>
6	<b>O FUTURO DA EDUCAÇÃO E O TRABALHO</b>
7	<b>TRANSFORMAÇÃO PRODUTIVA E INOVAÇÃO SUSTENTÁVEL</b>
8	<b>INTEGRAÇÃO E ALIANÇAS ESTRATÉGICAS</b>
9	<b>TENDENCIAS DE LONGEVIDADE</b>
10	<b>MUDANÇA CLIMÁTICA E ENERGÍAS RENOVÁVEIS</b>
11	<b>BIODIVERSIDADE, CAPITAL NATURAL E SOCIAL</b>

12	<b>RESILIÊNCIA</b>
13	<b>DEMOCRACIA E REDES SOCIAIS</b>
14	<b>GOVERNANÇA E EMPODERAMENTO CIDADÃO</b>
15	<b>BEM-ESTAR E QUALIDADE DE VIDA</b>

## 2.2 As Variáveis

<b>Variável</b>	<b>Significado</b>	<b>Tipo</b>	<b>Unidade de Medida Original</b>
PAÍS	É o nome do país.	Variável Categórica	N/A
PAÍS_COD	É o código de três letras atribuído a cada país.	Variável Categórica	N/A
REG	É a região em que o país está localizado, dentre 20 regiões do mundo.	Variável Categórica	N/A
SPI	Pontuação no Índice de Progresso Social	Variável Quantitativa	Pontuação do país, em escala centesimal.
Acess to information and Communications	Acesso a informação e a comunicações	Variável Quantitativa	(% Ocorrência)
Acess to Advanced Education	Acesso à Educação Avançada	Variável Quantitativa	(% Ocorrência)
Adult literacy rate (% of pop. Aged 15+)	Taxa de adultos alfabetizados	Variável Quantitativa	(% Ocorrência)

Primary school enrollment (% of children)	Percentual de crianças matriculadas no ensino fundamental ou equivalente	<i>Variável</i> <i>Quantitativa</i>	(% of children)
Lower secondary school enrollment (% of children)	Percentual baixo de crianças que estão estudando	<i>Variável</i> <i>Quantitativa</i>	(% of children)
Upper secondary school enrollment (% of children)	Percentual de crianças matriculadas no ensino médio ou equivalente)	<i>Variável</i> <i>Quantitativa</i>	(% of children)
Internet users	Percentual da população com acesso à Internet	<i>Variável</i> <b>Quantitativa</b>	(% of pop.)
Years of tertiary schooling	Anos que se permanece estudando	<i>Variável</i> <b>Quantitativa</b>	(% of children)
Inequality in the attainment of education (0=low; 1=high)	A desigualdade na condução da educação	<i>Variável</i> <i>Quantitativa</i>	Probabilidade (%)
Number of globally ranked universities (0=none; 5= >50)	A desigualdade na condução da educação	<i>Variável</i> <i>Quantitativa</i>	Probabilidade (%)
Unemployment Total (% of total Labor Force) 2013	A taxa de desemprego total em 2013	<i>Variável</i> <i>Quantitativa</i>	(% of children)
Total Labor Force (% Global)	Percentual do número de trabalhadores	<i>Variável</i> <i>Quantitativa</i>	(% of children)
Unemployment, youth total (% of	O desemprego, o percentual da juventude com idade de 15-24, que está trabalhando, estimado	<i>Variável</i> <i>Quantitativa</i>	(% of total labor force

total labor force ages 15-24) (modeled ILO estimate) 2013	conforme ILO – 2013		ages 15-24)
IDH2013	O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
Governança	Índice que avalia como o país, conduz questões éticas, ligadas corrupção, eficiência econômica, conflitos de interesse entre os acionistas	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
Knowledge and Technology	Conhecimento e Tecnologia	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
Creative Outputs	Saídas Criativas	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
EPI Score	Medição dos países que usam equipamentos de proteção individual, para segurança no trabalho	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
Happy Planet Index	Índice da felicidade está ligado ao bem-estar do indivíduo	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.
GINNI Index	Índice de desigualdade renda	<i>Variável Quantitativa</i>	Pontuação do país, em escala centesimal.

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados relativos a um conjunto de 21 (vinte e uma) variáveis de desenvolvimento humano selecionados

previamente, com o objetivo de explicar relações entre o futuro educação e trabalho num universo de 132 países, sendo vinte e umas variáveis quantitativas.

As variáveis numéricas são os 132 países, as variáveis categóricas e as variáveis quantitativas são os temas definidos no GPS, neste trabalho é o tema 6: O Futuro da Educação e do Trabalho, no qual serão analisadas todas as suas variáveis.

- 1) Social Progresso Index - Índice de Programa Social (variável quantitativa);
- 2) Acesso to information and Communications - Acesso a informação e comunicações (variável quantitativa);
- 3) Access to Advanced Education - Acesso à Educação Avançada (variável qualitativa);
- 4) Adult literacy rate (% of pop. Aged 15+) - Taxa de adultos alfabetizados (variável quantitativa);
- 5) Primary school enrollment (% of children) – Percentual de crianças matriculadas no ensino fundamental ou equivalente (variável quantitativa);
- 6) Lower secondary school enrollment (% of children) – Percentual baixo de crianças que estão estudando (variável quantitativa);
- 7) Upper secondary school enrollment (% of children) – Percentual de crianças matriculadas no ensino médio ou equivalente (variável quantitativa);
- 8) Internet users (% of pop.) – Percentual da população com acesso à Internet - (variável quantitativa);
- 9) Years of tertiary schooling – Anos que permanecem estudando - (variável quantitativa);
- 10) Inequality in the attainment of education (0=low; 1=high) – A desigualdade na condução da educação - (variável quantitativa);
- 11) Number of globally ranked universities (0=none; 5= >50) – Número de Universidades globalmente classificadas (variável quantitativa);
- 12) Unemployment Total (% of total Labor Force) 2013 - A taxa de desemprego total em 2013 - (variável quantitativa);
- 13) Total Labor Force (% Global) - Percentual do número de trabalhadores - (variável quantitativa);

- 14) Unemployment, youth total (% of total labor force ages 15-24) (modeled ILO estimate) 2013 – O desemprego, o percentual da juventude com idade de 15-24, que está trabalhando, estimado conforme ILO – 2013 - (variável quantitativa);
- 15) IDH - 2013 - Índice de Desenvolvimento Humano dos Países Subdesenvolvidos - (variável quantitativa);
- 16) Governança - Sistema que avalia como o país, conduz questões éticas, ligadas corrupção, eficiência econômica, conflitos de interesse entre os acionistas - (variável quantitativa);
- 17) Knowledge and Technology - Conhecimento e Tecnologia - (variável quantitativa);
- 18) Creative Outputs - Alternativas Criativas - (variável quantitativa);
- 19) EPI SCORE - Pontuação no Índice de Proteção Ambiental. Fonte: EPI trabalho - (variável quantitativa);
- 20) Happy Planet Index - Índice da felicidade está ligado ao bem-estar do indivíduo. (Variável quantitativa);
- 21) GINI Index – Índice de desigualdade renda - (variável quantitativa)

### **3. ANÁLISE DAS VARIÁVEIS**

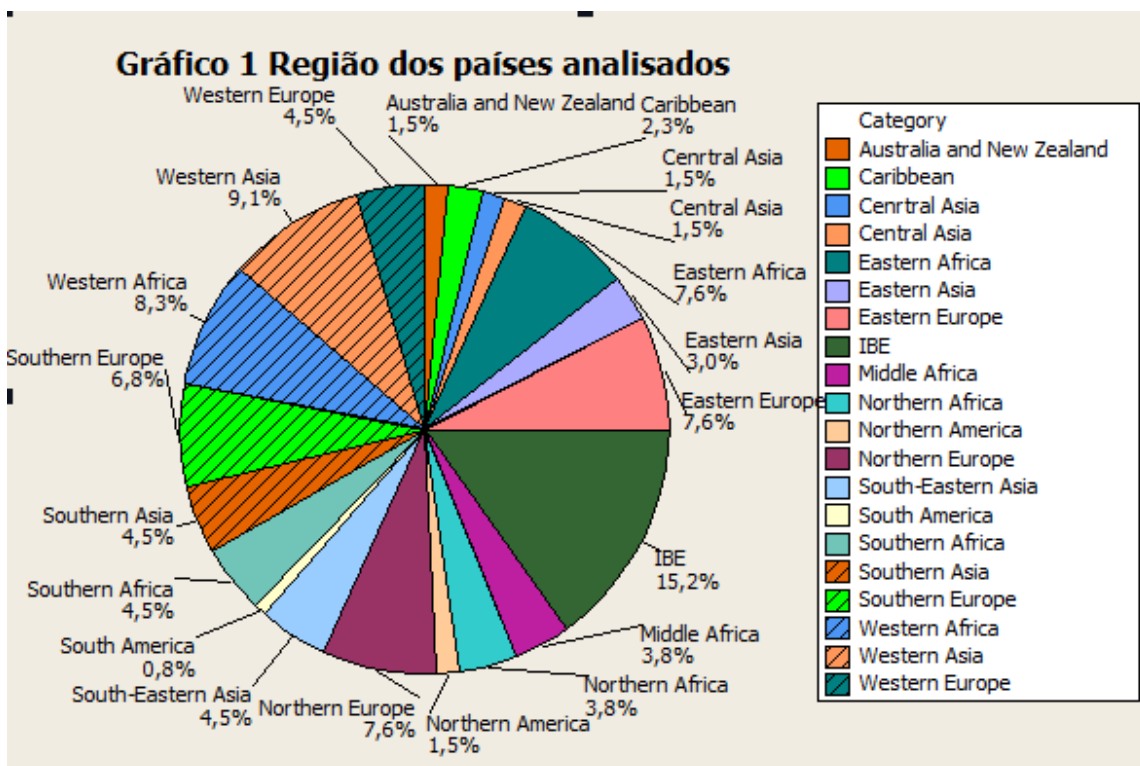
#### **3.1 Variáveis Categóricas**

Este tipo de variável indica que o foco de concentração deve ser a análise de gráficos do tipo *pie chart*, conforme abaixo.

##### **3.1.1 Variáveis: “PAÍS e Country Code” em função de “REGIÃO”**

Nossa amostra totaliza 132 países, que estão listados categoricamente em países e seus respectivos códigos, agrupados pela região do mundo em que estão localizados. O Gráfico I abaixo mostra a distribuição de países por região.





### 3.2 Variáveis Quantitativas

#### 3.2.1 Substituição de valores vazios pela média da categoria

A análise deste tipo de variável permite a utilização de uma maior gama de ferramentas de análise como histogramas, curvas de densidade, gráfico de ramos, box-plot e dot-plot, além de informações numéricas como média, desvio-padrão, mediana, quartis, 5 números, intervalo de confiança e teste de normalidade de Anderson-Darling.

O primeiro tratamento realizado foi a substituição de células vazias (marcadas com asterisco) pelo valor médio das variáveis em cada variável. Na análise abaixo (Tabela 2) foi possível identificar o número de valores faltantes na coluna “N\*” e a média de cada variável.

**Tabela 2:**

#### Descriptive Statistics: Social Progr; Access to In; Access to Ad; ...

Variable	Total			Mean	Minimum	Median	Maximum
	Count	N	N*				
Social Progress Index	132	132	0	63,67	32,60	63,86	88,24
Access to Information an	132	132	0	61,07	13,25	62,48	98,82
Access to Advanced Educa	132	132	0	39,78	4,72	41,74	89,37
Adult literacy rate (% o	132	132	0	84,91	25,31	93,73	100,00
Primary school enrollmen	132	127	5	90,724	40,621	93,716	100,000

Lower secondary school e	132	132	0	87,26	21,51	94,93	150,47
Upper secondary school e	132	132	0	72,71	6,76	76,36	174,48
Internet users (% of pop	132	132	0	41,96	1,22	41,87	96,21
Years of tertiary school	132	117	15	0,4475	0,0100	0,3500	1,7100
Inequality in the attain	132	117	15	0,1960	0,0130	0,1780	0,4980
Number of globally ranke	132	132	0	0,879	0,000	0,000	5,000
Unemployment Total (% of	132	131	1	9,011	0,300	7,200	31,000
Total Labor Force (% G	132	132	0	0,708	0,006	0,161	23,933
Unemployment, youth tota	132	131	1	19,19	0,70	15,80	60,40
IDH - 2013	132	125	7	0,6894	0,3370	0,7220	0,9440
1. Governança	132	132	0	53,38	15,22	48,46	100,00
1.5 Knowledge and Techno	132	123	9	29,63	2,40	26,60	60,90
1.6 Creative Outputs	132	123	9	32,57	0,60	32,60	66,10
EPI Score	132	132	0	52,48	18,43	52,09	87,67
Happy Planet Index	132	129	3	42,673	22,591	42,463	64,036
GINNI Index	132	119	13	39,024	24,820	38,160	63,140

Num segundo passo, substituímos os valores faltantes pela média de cada variável, resultando na Tabela 3, abaixo, na qual verificamos que os valores faltantes foram zerados sem impacto na média da distribuição:

**Tabela:3**

**Descriptive Statistics: Social Progr; Access to In; Access to Ad; ...**

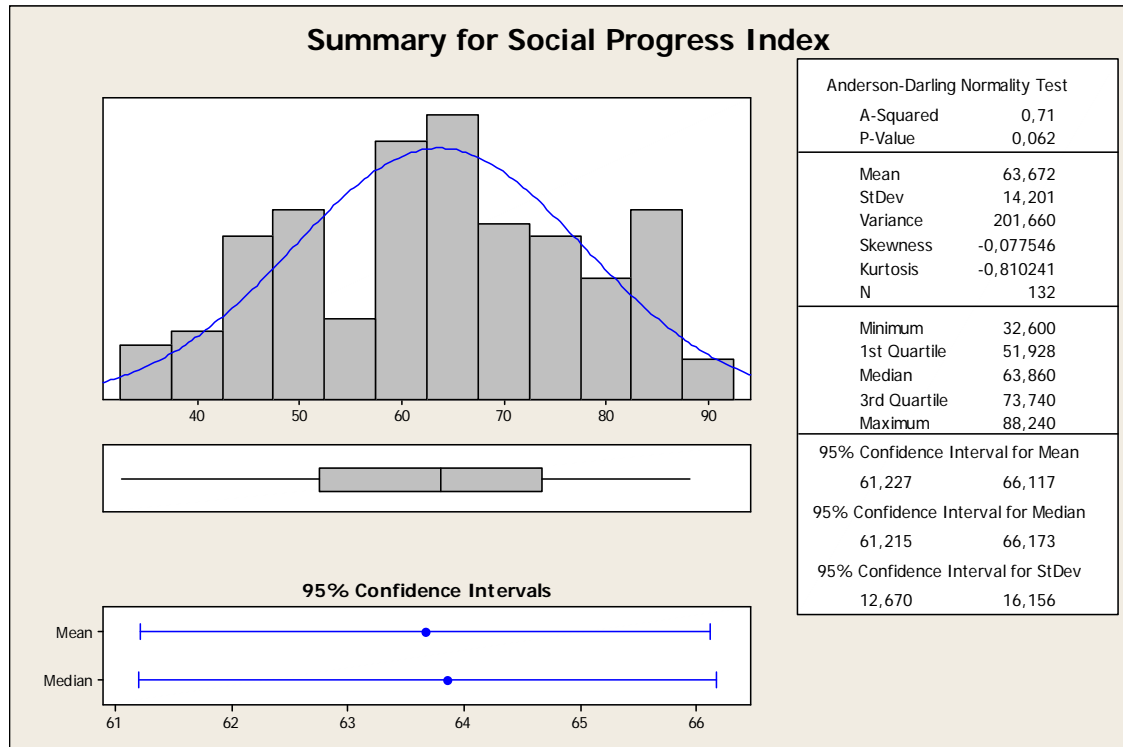
Variable	Total			Mean	Minimum	Median	Maximum
	Count	N	N*				
Social Progress Index	132	132	0	63,67	32,60	63,86	88,24
Access to Information an	132	132	0	61,07	13,25	62,48	98,82
Access to Advanced Educa	132	132	0	39,78	4,72	41,74	89,37
Adult literacy rate (% o	132	132	0	84,91	25,31	93,73	100,00
Primary school enrollmen	132	132	0	90,838	40,621	93,716	100,000
Lower secondary school e	132	132	0	87,26	21,51	94,93	150,47
Upper secondary school e	132	132	0	72,71	6,76	76,36	174,48
Internet users (% of pop	132	132	0	41,96	1,22	41,87	96,21
Years of tertiary school	132	132	0	0,4364	0,0100	0,3500	1,7100
Inequality in the attain	132	132	0	0,1942	0,0130	0,1800	0,4980
Number of globally ranke	132	132	0	0,879	0,000	0,000	5,000
Unemployment Total (% of	132	132	0	8,948	0,300	7,150	31,000
Total Labor Force (% G	132	132	0	0,708	0,006	0,161	23,933
Unemployment, youth tota	132	132	0	19,10	0,70	15,70	60,40
IDH - 2013	132	132	0	0,6911	0,3370	0,7220	0,9440
1. Governança	132	132	0	53,38	15,22	48,46	100,00
1.5 Knowledge and Techno	132	132	0	29,42	2,40	26,60	60,90
1.6 Creative Outputs	132	132	0	32,57	0,60	32,60	66,10
EPI Score	132	132	0	52,48	18,43	52,09	87,67
Happy Planet Index	132	132	0	42,668	22,591	42,463	64,036
GINNI Index	132	132	0	38,939	24,820	38,160	63,140

### 3.2.2 Análise das variáveis

Nesta seção utilizaremos as seguintes ferramentas estatísticas para analisar as 14 variáveis quantitativas: Histograma, Curva de Densidade, Box-Plot, Intervalo de confiança da média e mediana.

Também poderemos considerar na análise dados como Média, desvio-padrão, variância, quantidade de observações, valores mínimos, máximos, informações dos quartis e o teste de normalidade de Anderson-Darling (A-Squared e P-Value), para todas as variáveis.

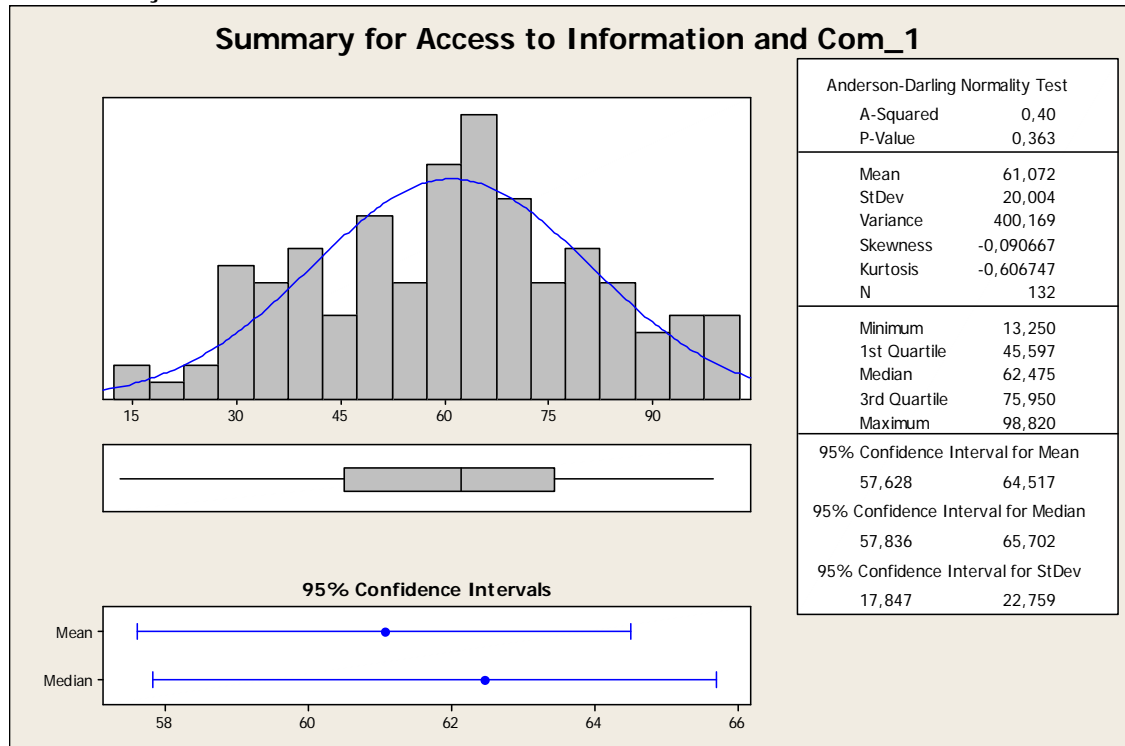
### Variável Social Progress Index - Índice de Progresso Social



A princípio a distribuição aproxima-se da normal. O gráfico registra a pontuação cardinal no conjunto de observações para os 132 países analisados.

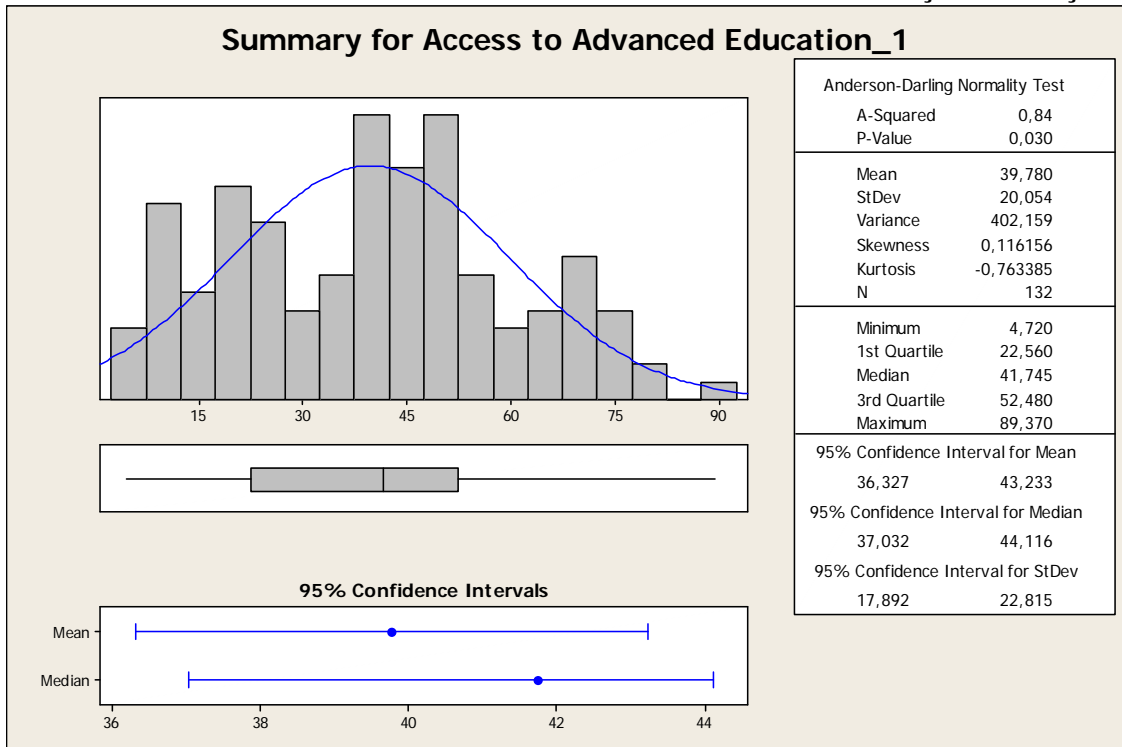
Percebe-se um número menor de países do que o esperado classificados entre 35 e 45 com apenas 7 observações, e o mesmo fenômeno se repete para a faixa de pontuação 75 à 85, que também apresenta apenas 7 observações. O maior número de observações encontra-se na faixa de 45 a 55, com 24 observações. Esta faixa corresponde a última faixa imediatamente inferior à mediana e corresponde a um nível de desenvolvimento considerado “médio” pelo índice. A mediana de 56,183 e o terceiro quartil de 73,94 indicam o poder discricionário do índice, que consegue separar o grupo de países em 11 estratos bem definidos, o que pode ser bastante útil para classificações futura de dados.

**Variável Access to information and Communications – Acesso a informação e Comunicações**



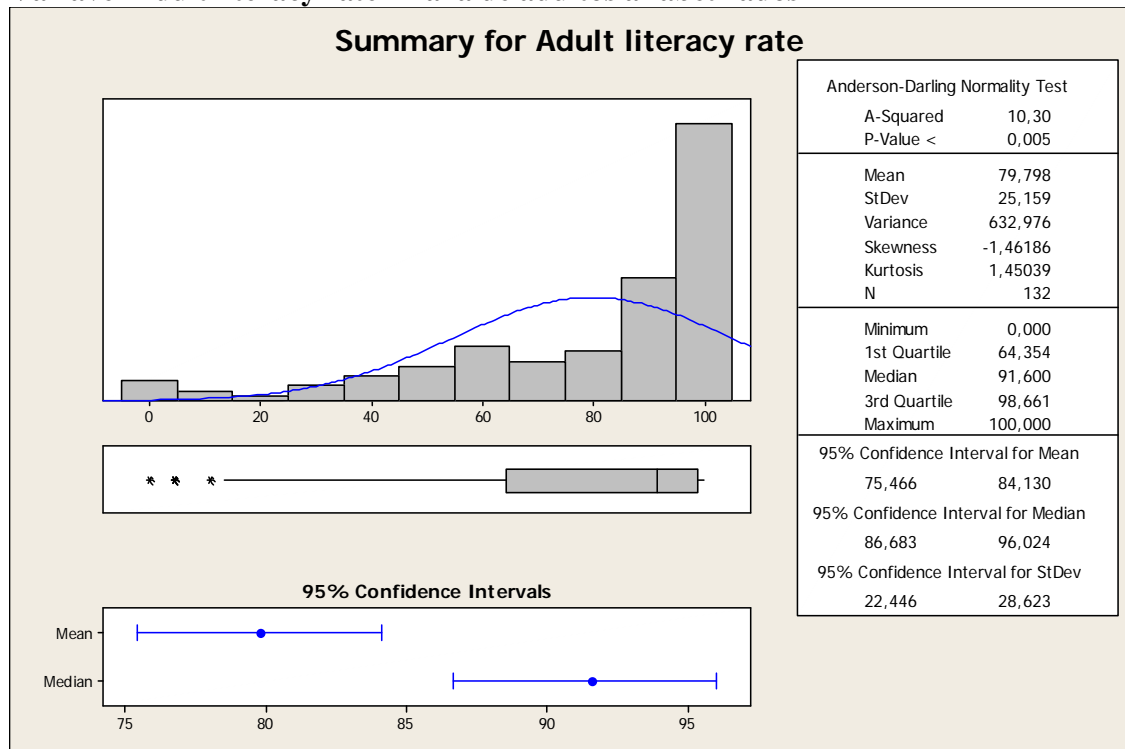
Nesse gráfico, temos a mesma informação do gráfico 1, países desenvolvidos com maior acesso à informação, países desenvolvidos com menor acesso à informação. Distribuição normal, com P-Value maior que 0,05. O país com o maior índice é a Irlanda, com o menor índice é Cuba com 13,25 %.

## Variável Access to Advanced Education – Acesso à Educação Avançada



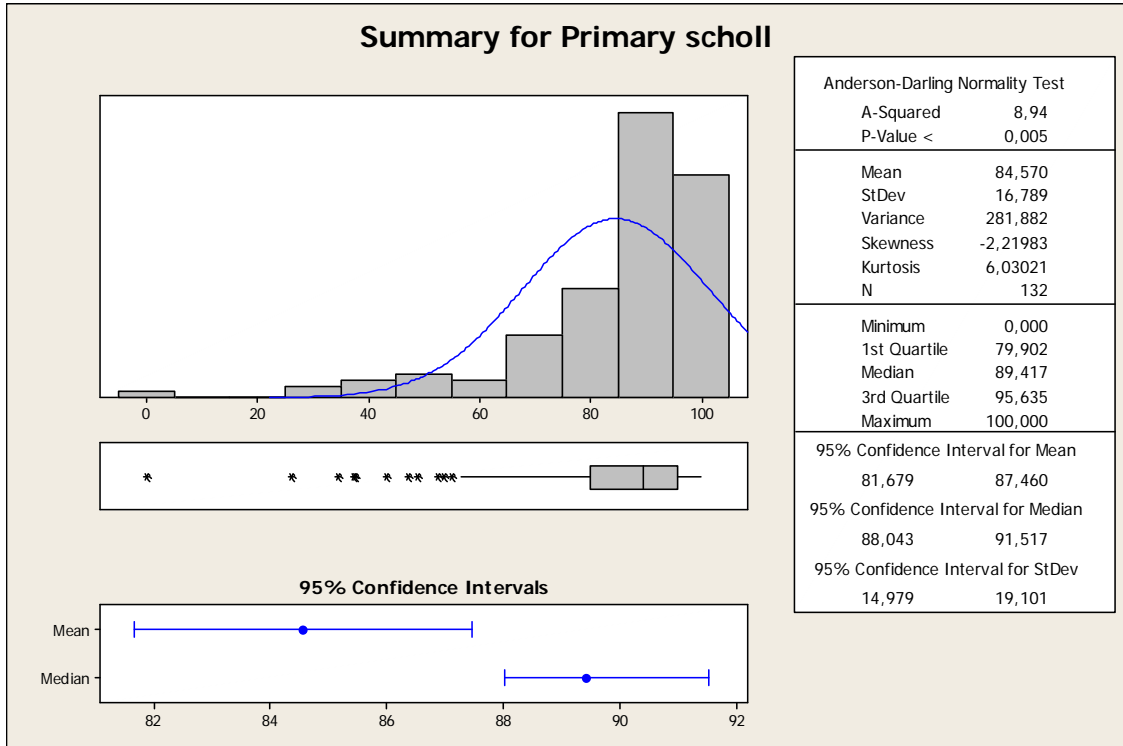
O P- Value maior que 0,05, também igual aos gráficos 1 e 2, países desenvolvidos melhor acesso à educação avançada e países em desenvolvimento com menor acesso. Posteriormente, pode ser feito um estudo mais detalhado, se existe uma congruência entre os países. O país com melhor índice é o USA com 89,37 %, o Iêmen com o menor índice de 4,72 %.

## Variável Adult literacy rate - Taxa de adultos alfabetizados



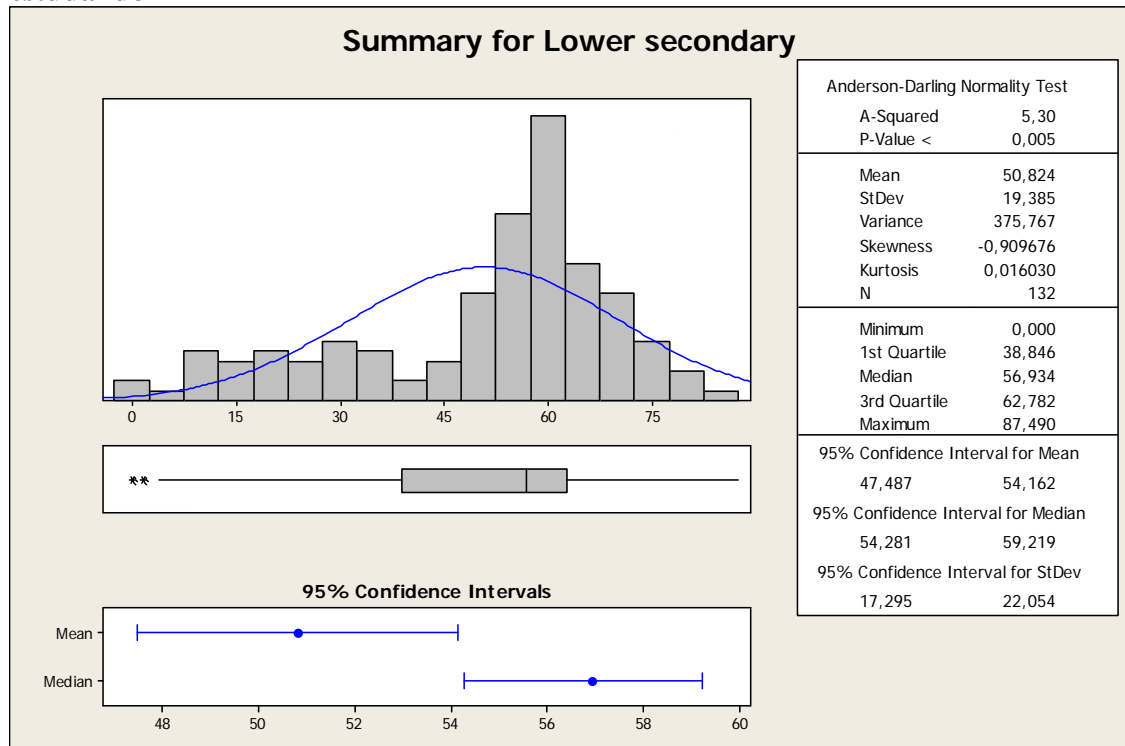
O P-Value não possui  $P > 0,05$  que é o nível de significância, portanto os dados no gráfico não apresentam distribuição normal, existe uma enorme variação entre o país que possui 100 % de alfabetização de adultos, com o país que tem somente 25,308 % de adultos alfabetizados. Existem vários outliers por definição, pois os países com melhores condições econômicas, políticas, sociais possuem mais adultos alfabetizados.

**Variável Primary school enrollment - Percentual de crianças matriculadas no ensino fundamental ou equivalente**



O P-Value não possui  $P > 0,05$  que é o nível de significância, portanto os dados no gráfico não apresentam distribuição normal, existe uma enorme variação entre o país que possui 100 % de crianças matriculadas no ensino fundamental ou equivalente, comparado com o país que tem somente 40,621 % de crianças matriculadas no ensino fundamental ou equivalente. Existem vários outliers por definição, pois os países com melhores condições econômicas, políticas, sociais possuem mais crianças matriculadas no ensino fundamental ou equivalente.

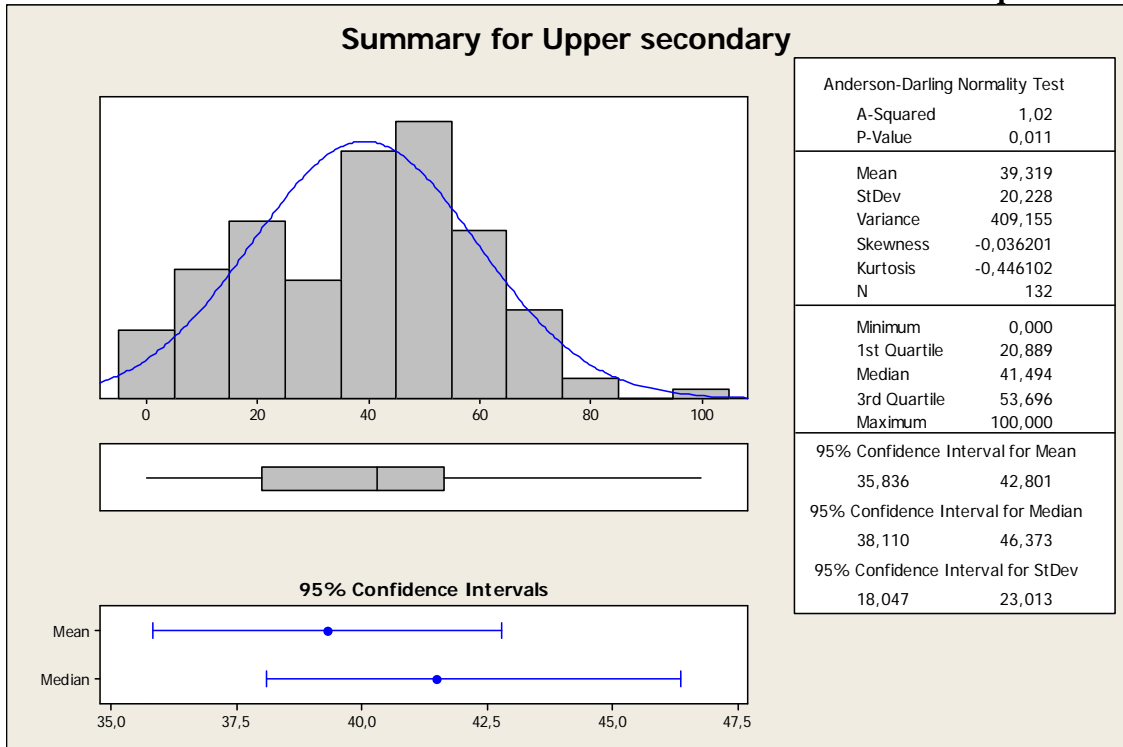
**Variável Lower secondary school enrollment – Percentual baixo de crianças estudando**



O P-Value não possui  $P > 0,05$  que é o nível de significância, portanto os dados no gráfico não apresentam distribuição normal, existe uma enorme variação entre o país que possui 150,46% de crianças com baixa taxa que estão estudando, contra o país que tem 21,50% de crianças com baixa taxa que estão estudando. Existem vários outliers por definição, pois os países com melhores condições econômicas, políticas, sociais não possuem crianças que não estejam estudando.

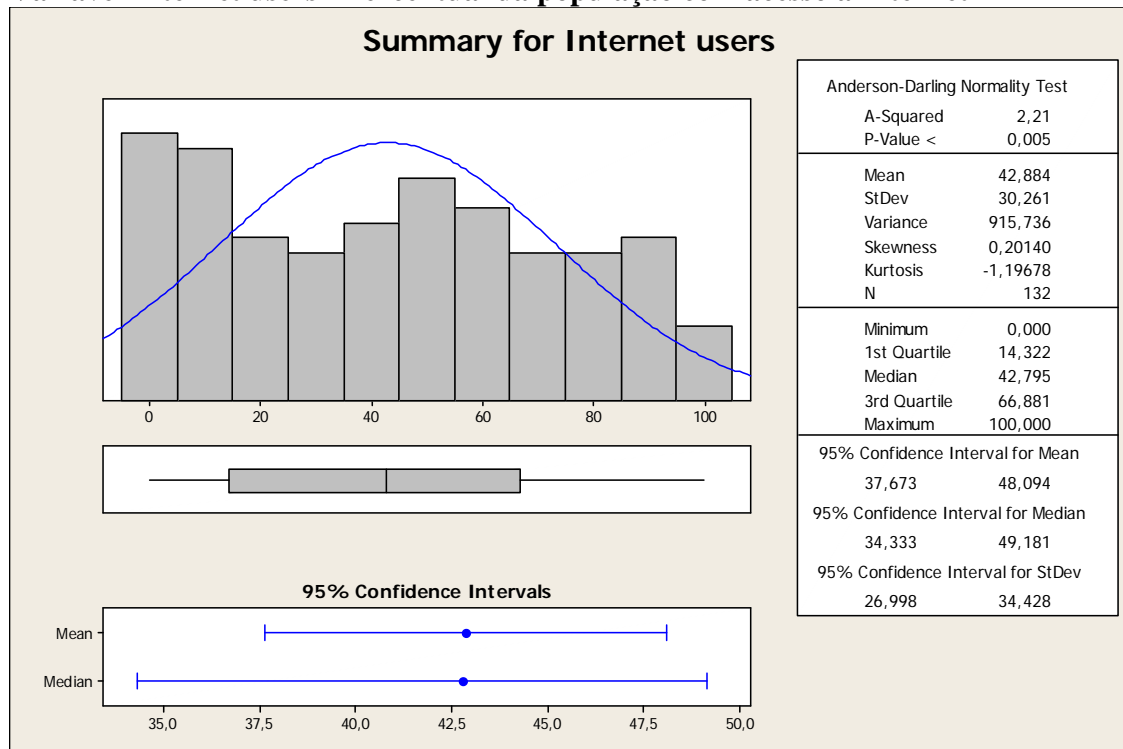


**Variável Upper secondary scholl enrollment - Percentual de crianças matriculadas no ensino médio ou equivalente**



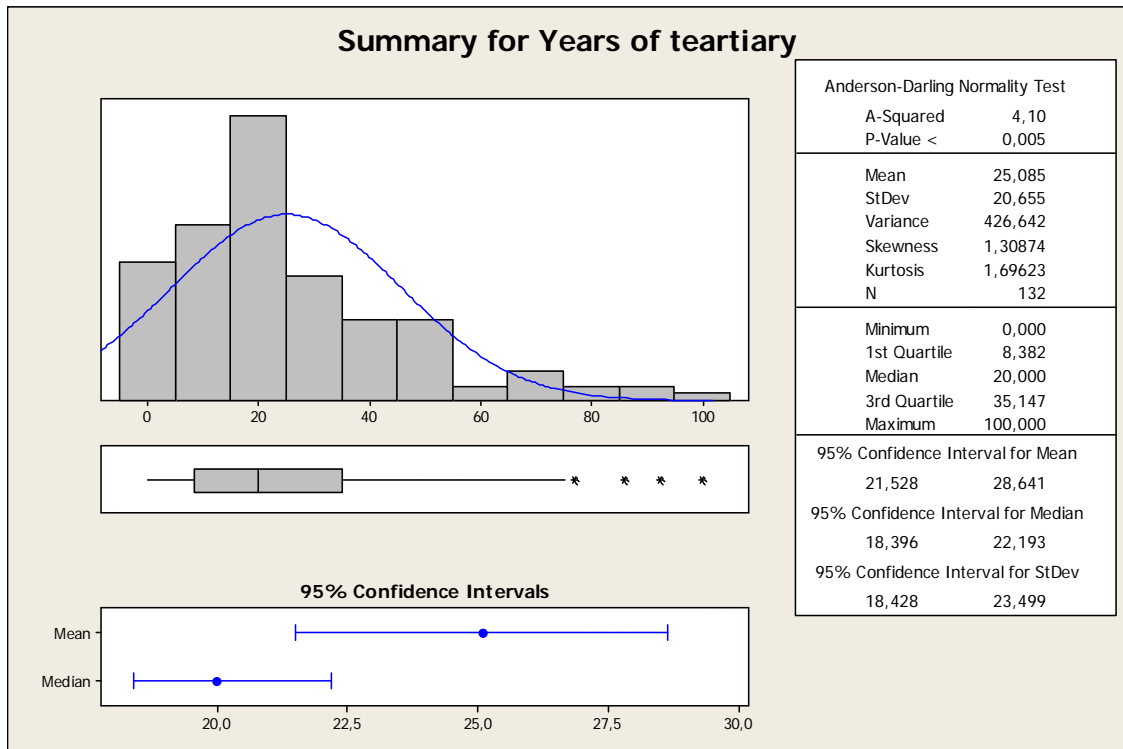
O P- Value  $> 0,05$  apresenta gráfico com distribuição normal, onde não existe uma discrepância entre os 132 países, significando que aparentemente o problema seria a pré-escola. Os outliers estão praticamente distribuídos, pois quando a criança consegue ingressar na escola, sem nenhuma interrupção ela termina o ensino médio ou equivalente.

## Variável Internet users - Percentual da população com acesso à Internet



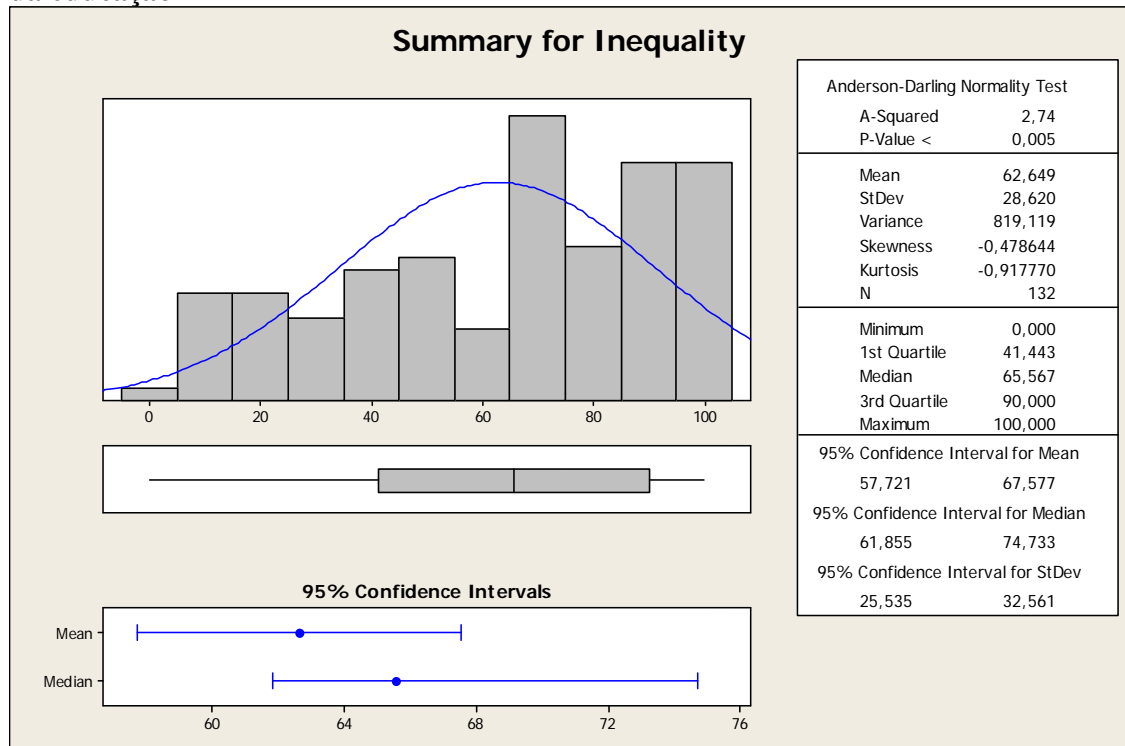
O P- Value não é  $> 0,05$ , portanto não é uma distribuição normal, existem vários países que são outliers, mas existem alguns países que estão abaixo dos outliers, por limitação a utilização da internet com restrições a sites, assim como existem países que não possuem tecnologia, outras restrições religiosas, essa variável é um mix a ser analisado, sendo que o mundo hoje é praticamente digital.

## Variável Years of teartary schooling - Anos que se permanece estudando



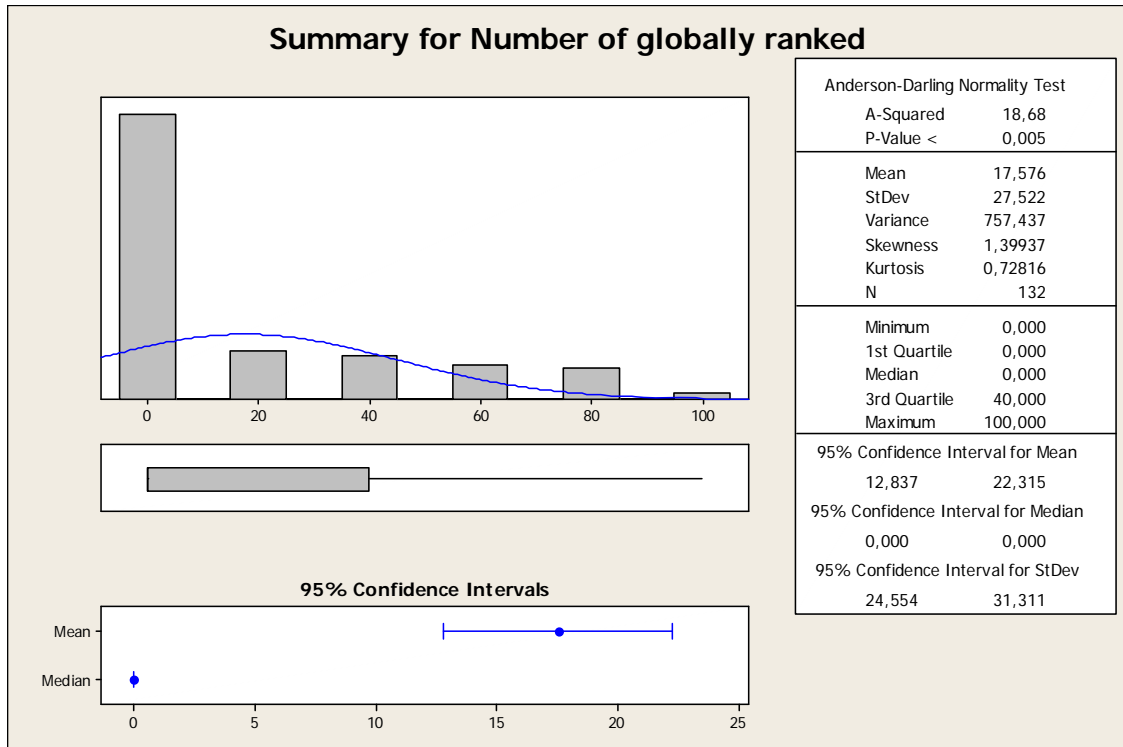
O P- Value não é maior que 0,05, portanto não é uma distribuição normal. Existem vários outliers, os países com melhores condições econômicas, sociais, políticas, possuem mais alunos nas escolas estudando por maior número de tempo, ao contrário dos países menos favorecidos do qual apresentam um enorme índice de alunos fora da escola.

## Variável Inequality in the attainment of education – A desigualdade na condução da educação



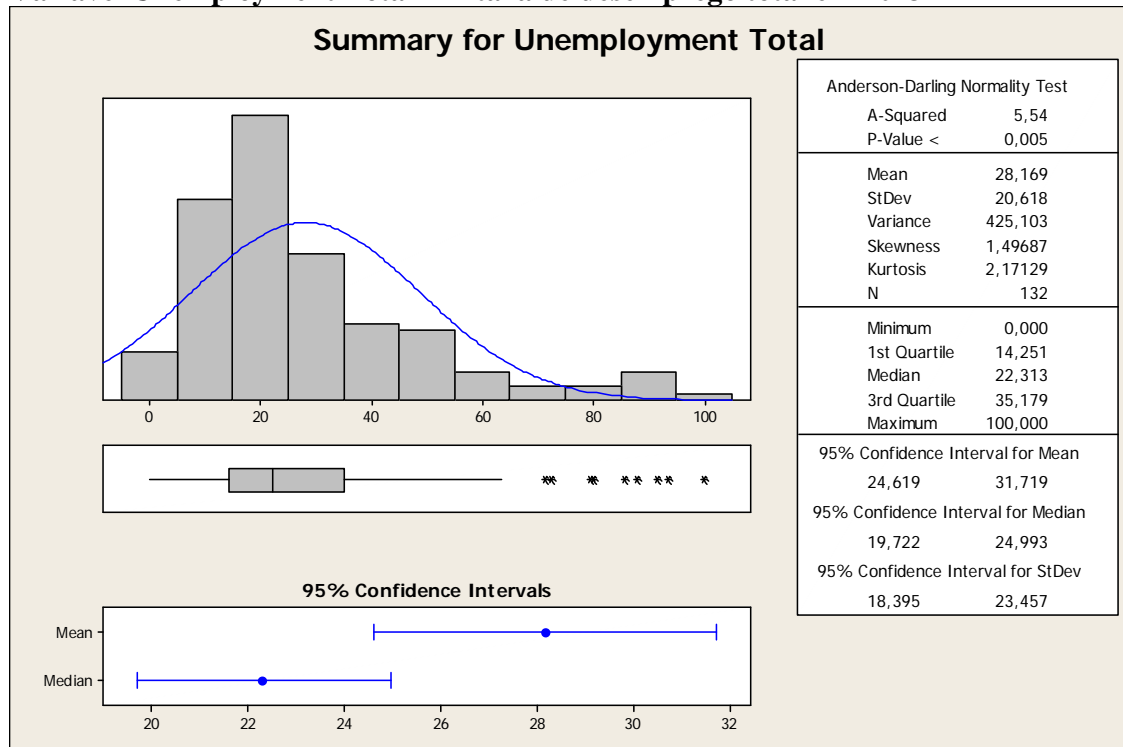
O P- Value não é maior 0,05, portanto não é uma distribuição normal. Existem vários países que são outliers, muitos países em condições muito melhores na condução da educação, não ocorrendo tanta desigualdade, de quem está estudando ou não. Por outro lado, existem países muito piores, onde existe uma desigualdade superior.

**Variável Number of globally ranked universities – Número de universidades globalmente classificadas**



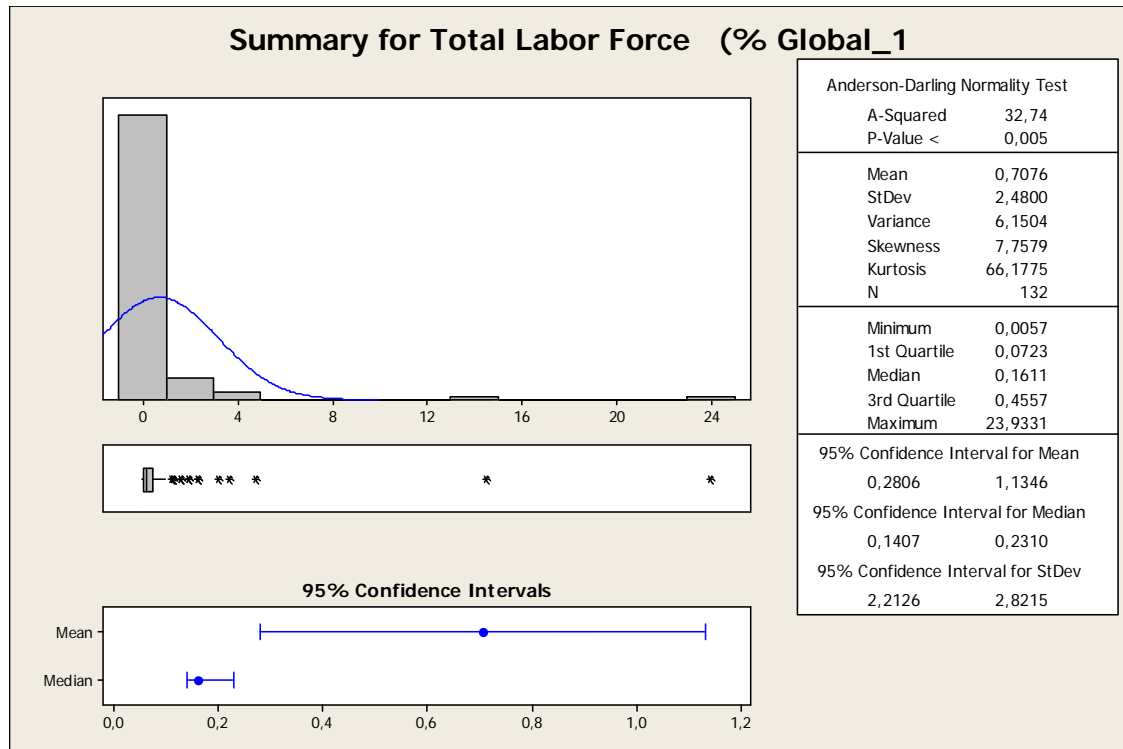
Nesse gráfico, o P-Value é menor que 0,05, é notável que é necessário investir muito nesse item, são poucos os países que possuem Universidades globalmente ranqueadas, além disso os países que tem 5 Universidades ranqueadas também são inexpressivos. Estamos em um mundo global, onde é possível se conectar com a grande maioria dos países, por que então não existe um investimento em 100% Universidades on-line gratuitas.

## Variável Unemployment Total – A taxa de desemprego total em 2013



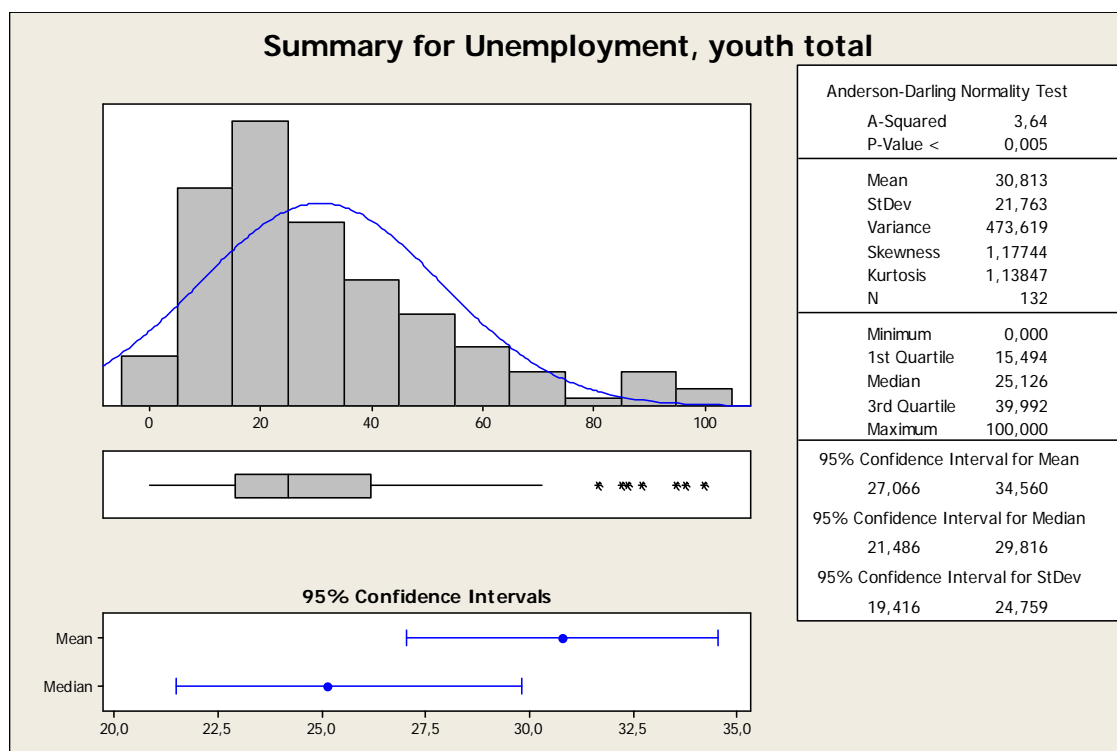
Nesse gráfico, o P-Value é menor que 0,05, é notável que é necessário investir muito nesse item, são poucos os países que possuem Universidades globalmente ranqueadas, além disso os países que tem 5 Universidades ranqueadas também são inexpressivos. Estamos em um mundo global, onde é possível se conectar com a grande maioria dos países, por que então não existe um investimento em 100% Universidades on-line gratuitas.

## Variável Total Force – Percentual global de números de trabalhadores



Esse índice demonstra o desequilíbrio mundial, não é uma distribuição normal P-Value é menor 0,05. Existem vários outliers, que estão retomando os negócios depois das últimas crises financeiras, mas não estão dando sinais de sustentabilidade, por outro lado temos muito países com índices expressivos do número de desempregados, seja pelas crises, pelas guerras civis, religiosas e etc.

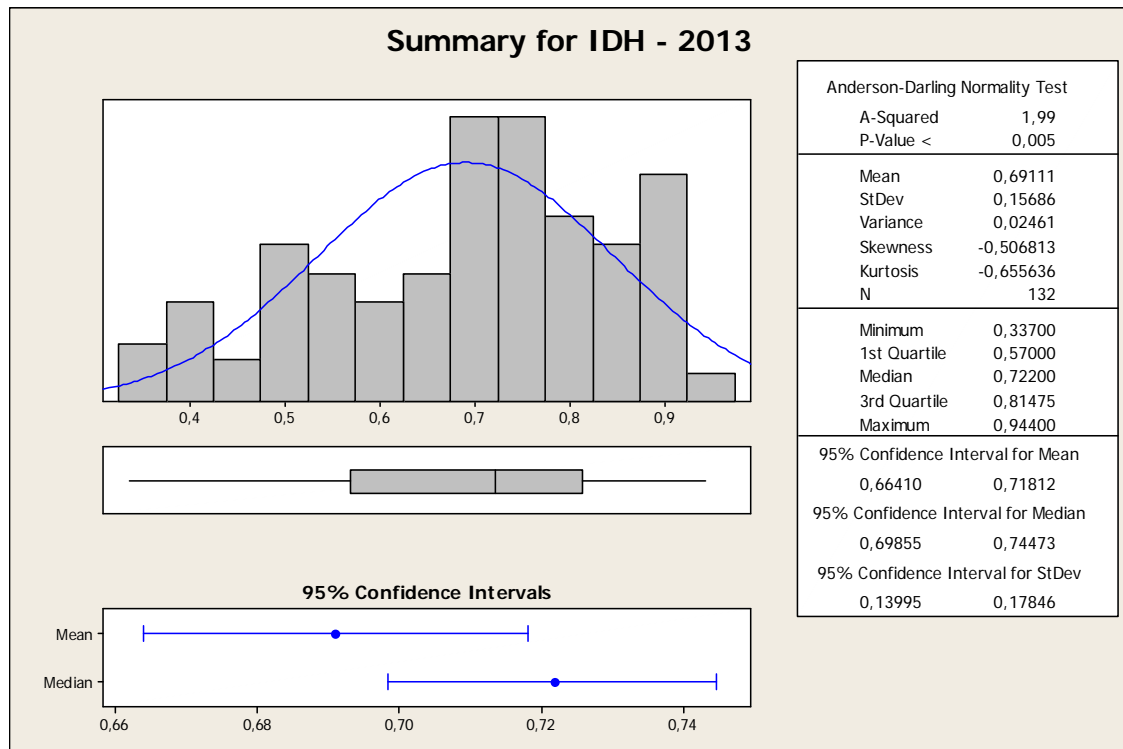
**Variável Unemployment, youth total - O desemprego, o percentual da juventude com idade de 15-24, que está trabalhando, estimado conforme ILO – 2013**



Esse item só reforça a análise anterior do gráfico 9, só que com um maior destaque ao número de jovens desempregados concentrados na região europeia, na análise do sumário o país com maior índice de jovens desempregados é a Bósnia Herzegovina, país com constantes guerras civis, por isso na minha análise foi descartado esse país, considerando países com índice iguais, mas com problemas econômicos financeiros graves.



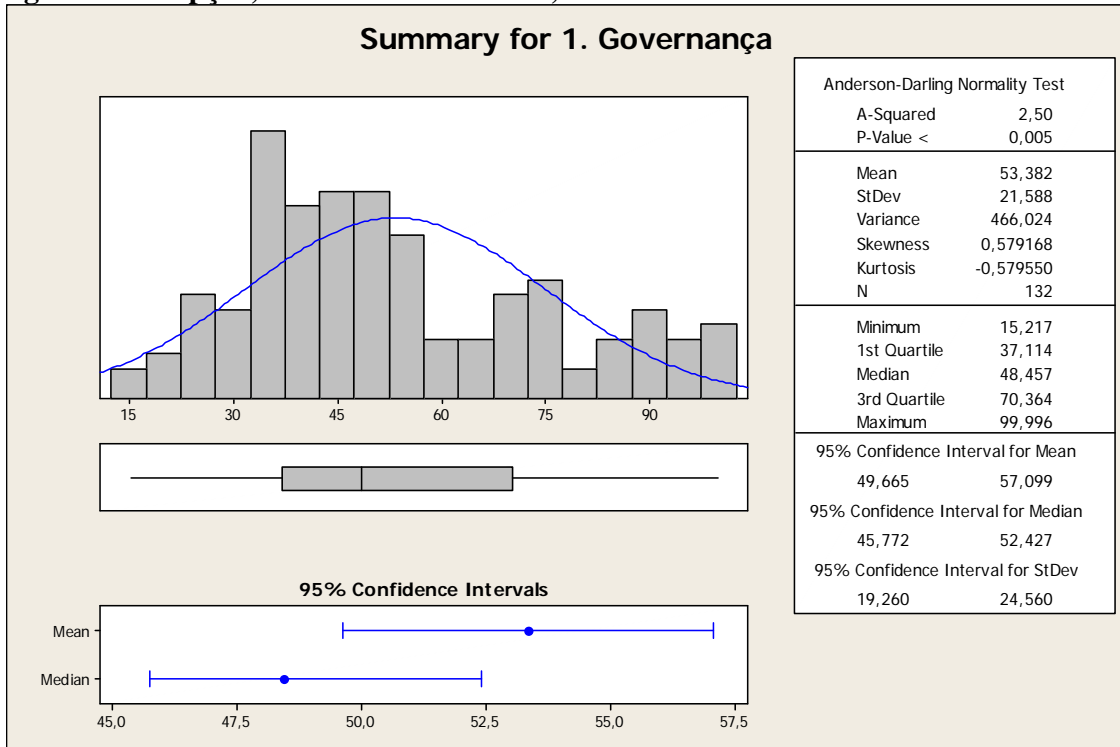
**Variável IDH 2013 - O Índice de Desenvolvimento Humano (IDH) mede o progresso de uma nação a partir de três dimensões: renda, saúde e educação. Fonte: PNUD, 2013.**



O gráfico para o IDH aproxima-se de uma distribuição normal, mas deslocada para a direita, com uma concentração maior de países dos 55 aos 75 pontos, totalizando 46 países. Isso pode ser constatado também pela mediana, em 62,35. No estrato inferior, com 55,84 pontos, encontra-se o Paraguai, com IDH de 0,676 (considerado “médio” nível de desenvolvimento humano) e no estrato superior, com 74,62 pontos encontra-se o Uruguai, com IDH de 0,79, que é considerado “alto desenvolvimento humano” pelo PNUD.

A partir destes dados é possível afirmar que 95, dos 132 países analisados tem nível de IDH “médio” ou superior.

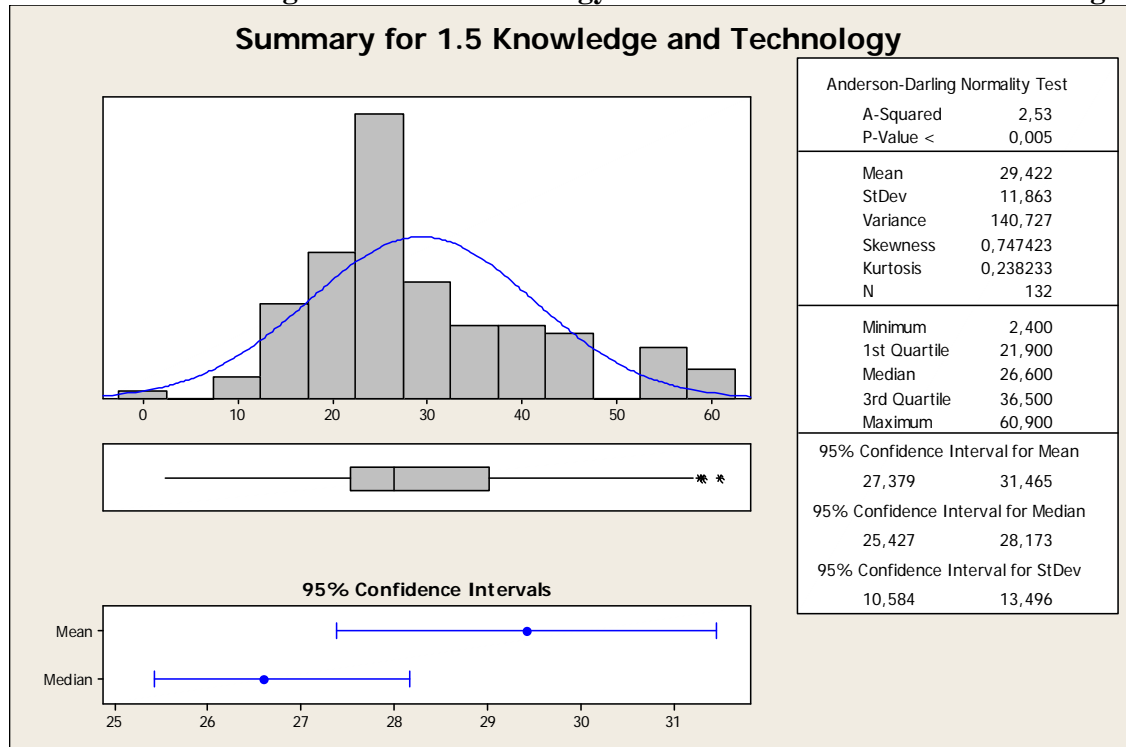
**Variável Governança - Índice que avalia como o país, conduz questões éticas, ligadas corrupção, eficiência econômica, conflitos de interesse entre os acionistas**



O gráfico para a variável GOV aproxima-se de uma distribuição normal deslocada para a esquerda, o primeiro quartil de 25,82 e a mediana de 39,21 confirmam esse deslocamento, indicando uma concentração crescente de países entre as faixas de 5 a 15 (13 países), de 15 a 25 (15 países), de 25 a 35 (22 países) até chegar ao pico – a faixa 35 a 45 (29 países). Essas faixas agrupam o equivalente a 60% dos países analisados.

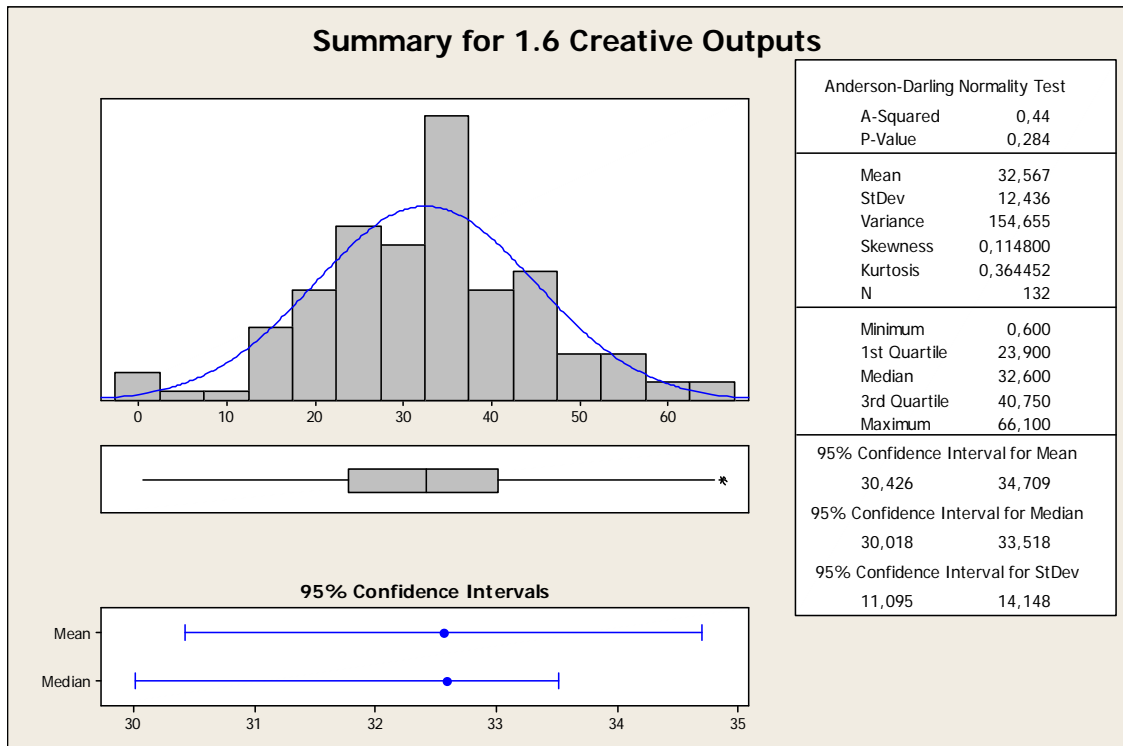
Em relação ao pico, os países mais bem colocados são a Macedônia (44,74), a Turquia (44,69) e o Kuwait (44,24) e na faixa inferior ficam Argentina (35,30), Benin (35,75) e Indonésia (35,96). Aparentemente trata-se de grupo bastante heterogêneo, não sendo possível estabelecer relações imediatas.

## Variável Knowledge and Techonology – Conhecimento e Tecnologia



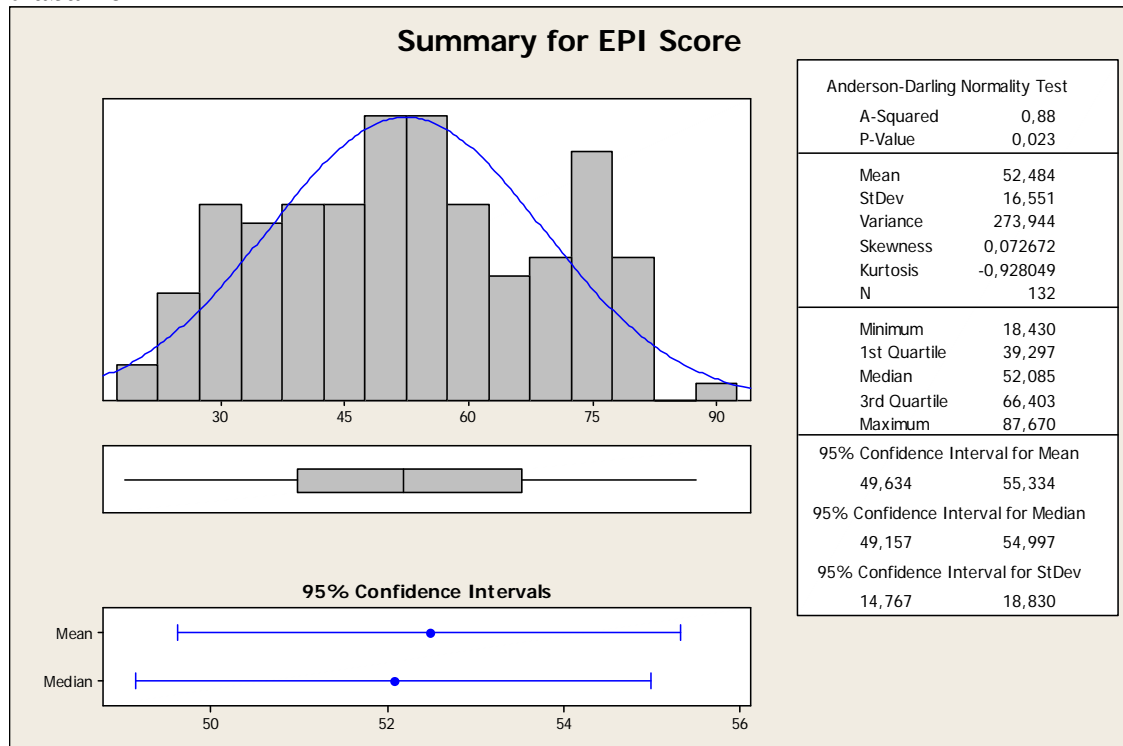
Esse gráfico, não é uma distribuição normal P- Value < 0,05, existe um outlier importante que é a Suíça com 60,9 %, além de outros países nórdicos terem índices semelhantes e a China ter o índice de 59,0 %, com um índice bem inexpressivo aparece o Sudão.

## Variável Creative Outputs – Saídas Criativas



O P-Value é maior que 0,05, portanto é uma distribuição normal. Os países com melhores condições econômicas, financeiras e sociais apresentam soluções mais criativas, são as chamadas “Ideias fora da Caixa”, o país que tem o melhor índice é a Irlanda e país com o menor índice é o Togo país subdesenvolvido da África.

**Variável EPI Score - Pontuação no Índice de Proteção Ambiental. Fonte: EPI trabalho**

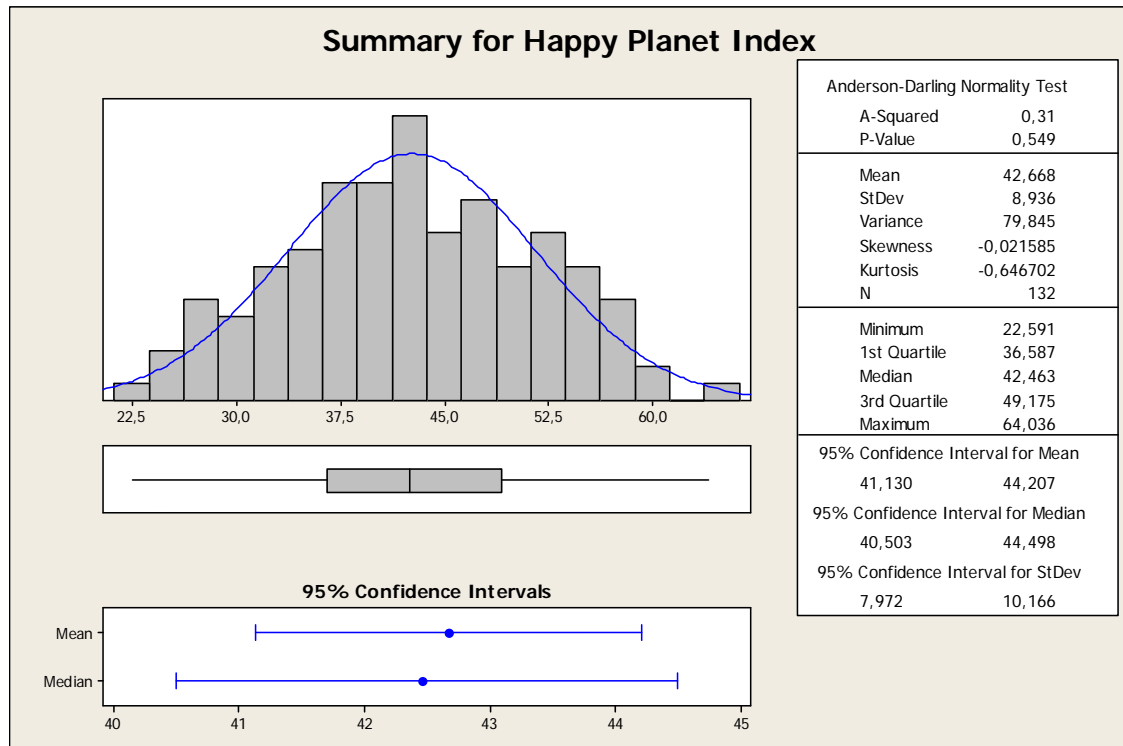


A análise dos dados do Índice de Proteção Ambiental apresenta uma distribuição normal, com média de 49,18 e mediana de 48,60, ou seja, um ligeiro deslocamento para a esquerda. Há apenas 1 país no estrato superior (nota normalizada acima de 95), que é a Suíça, com EPI de 87,67, seguida por 9 países no segundo estrato (de 85 a 95), sendo todos países de alto IDH e IPS (Austrália, República Tcheca, Alemanha, etc.) sendo o último país do estrato a Dinamarca com EPI de 76,92.

Já nos dois estratos inferiores temos a seguinte composição: no primeiro estrato (nota normalizada até 5) há dois países – Mali (EPI = 18,43) e Lesoto (EPI = 20,81), seguidos por um grupo de 10 países no segundo estrato que vai de 5 a 15. Destes 10 países, 9 encontram-se na África, sendo Bangladesh a única exceção do grupo. O país melhor colocado é Angola, com EPI de 39,93.

Já a faixa de pico, que vai de 45 a 55, é composta da seguinte maneira: na parte inferior, Algéria (EPI = 50,08) e Líbano (EPI = 50,15) seguindo por um grupo heterogêneo de 26 países nos quais Jordânia (EPI = 55,78) e Montenegro (EPI = 55,52).

## Variável Happy Planet – Medida de bem-estar sustentável. Fonte: HPI



A distribuição para o Índice de Felicidade Bruta (FIB) aproxima-se de uma normal, com maior concentração de países entre a marca do 35 até 65, que formam quase um plateau, com os três maiores picos da distribuição. Nas extremidades encontramos um pequeno número de países : Costa Rica ficou um primeiro lugar, com um FIB de 64,03, seguida por um estrato de notas normalizadas de 85 a 95, que é composto por três países : Colômbia (FIB=59,75), El Salvador (58,88) e Jamaica (58,53).

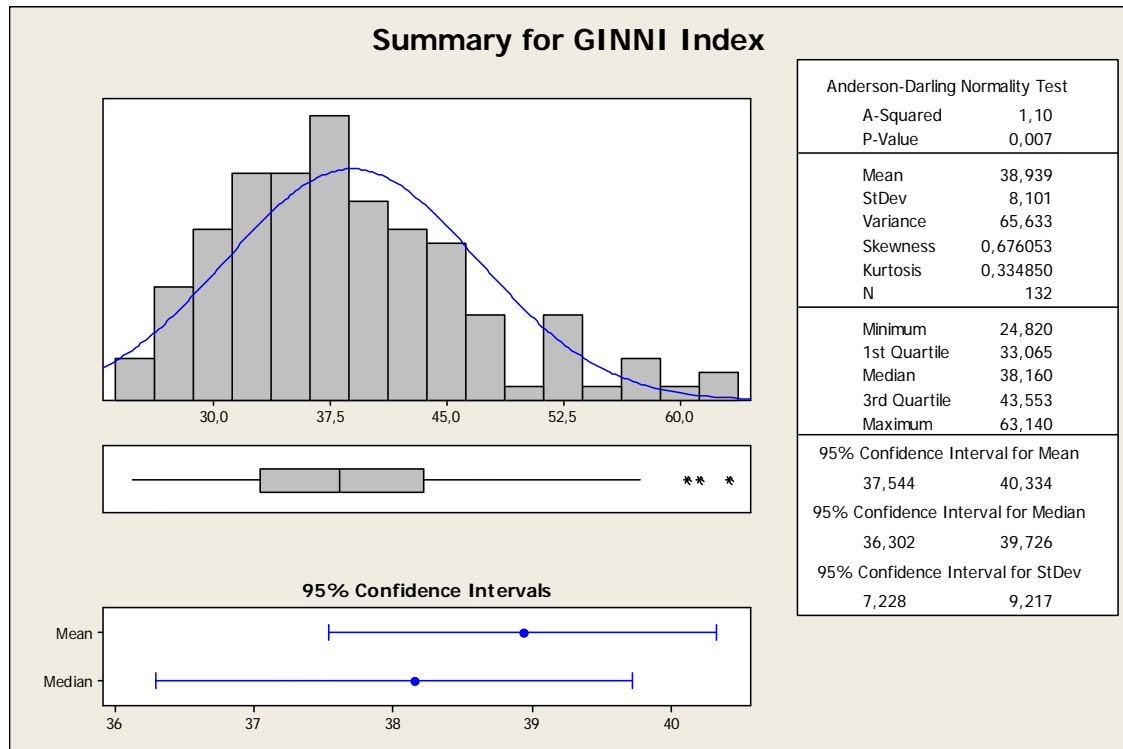
No estrato inferior, que vai até 5 pontos, figura também apenas um país :Botswana, com FIB = 22,59, seguida pelo segundo estrato, que vai de 5 a 15 pontos e compreende 9 países. A menor pontuação é do Chade (FIB=24,68) e a maior é da Macedônia (FIB=28,27). Chama a atenção neste grupo de países a presença da África do Sul, que figura com um FIB de 28,19, o oitavo pior colocado na classificação geral do FIB.

A maior concentração em número de países está alocada entre 35 e 65 pontos, sendo Djibouti (FIB=37,23), Estados Unidos (FIB=37,34) e Hungria (FIB=37,40) os piores colocados do grupo, e Sri Lanka (FIB=49,38), Iraque (FIB=49,19) e Laos (FIB=49,14) os melhores colocados do grupo.

O Índice de Felicidade Bruta mede aspectos subjetivos e objetivos da vida humana, evidenciando pesquisas que apontam que, a partir de certo nível de renda, o nível de felicidade médio reportado não aumenta na mesma proporção do aumento da renda, até decaindo com o tempo. Outro tópico relaciona é a questão da resiliência construída por pessoas em situação de privação em países de baixo IDH e IPS, que acabam adaptando-se às situações adversas.

Uma análise mais aprofundada poderá revelar o baixo nível de correlação deste indicadores com o PIB e inclusive com o IDH e o IPS, pois sua metodologia é bastante distinta.

## Variável GINNI Index – Medida de desigualdade de renda



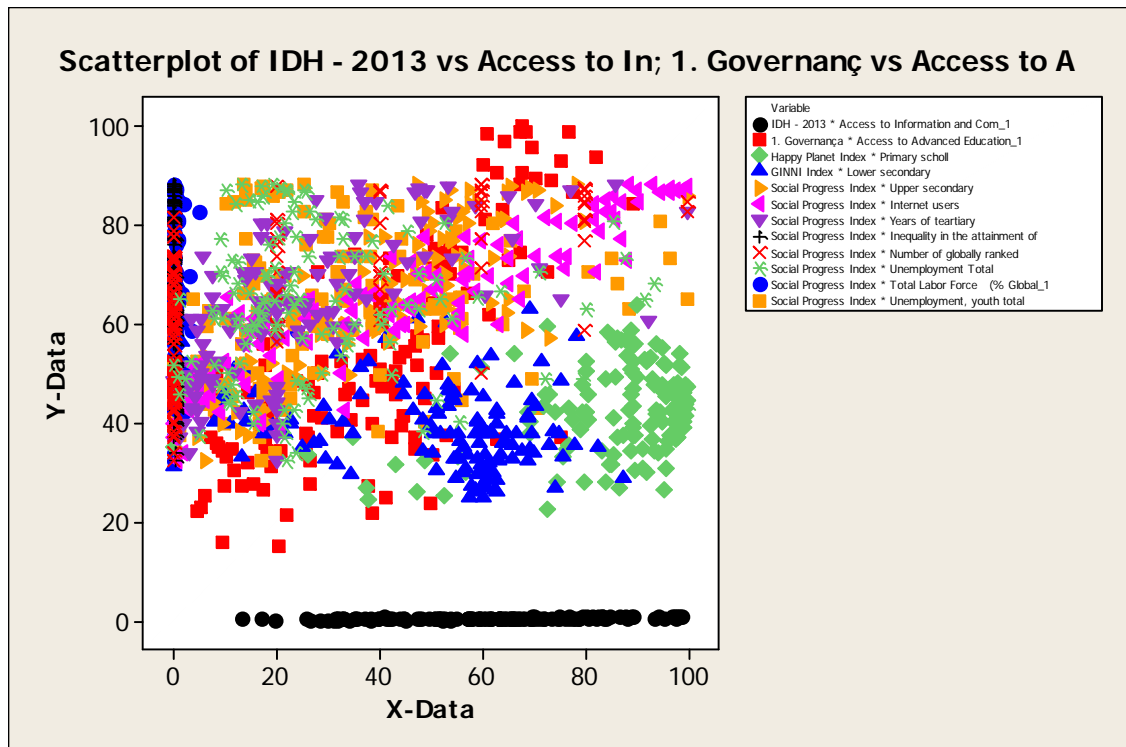
Para o Índice de Gini, que mede a desigualdade de renda, faltaram dados relativos a 13 países, que foram substituídos pela média da distribuição original, que foi de 39,02, que normalizado de 0 a 100 tornou-se 62,93, o que explica o pico registrado na coluna entre 55 e 65 – dos 33 países, na realidade apenas 20 tem sua pontuação original registrada neste intervalo.

Assim percebemos uma distribuição basicamente uniforme, com tendência para concentração de países mais a direita, como já demonstra o primeiro quartil, em 51,11. São 6 países no estrato superior, acima de 95 pontos na escala normalizada.

## 4. RELAÇÃO ENTRE AS VARIÁVEIS

### 4.1 – Gráficos de Dispersão

Na Tabela abaixo encontram-se os gráficos de dispersão entre a variável quantitativa Adult Literacy Index e as demais variáveis quantitativas.



Gráficos de dispersão devem ser inicialmente analisados quanto a seu padrão geral e seus desvios relativos ao padrão. A descrição do padrão geral pode ser feita pela verificação de sua forma, direção e intensidade.

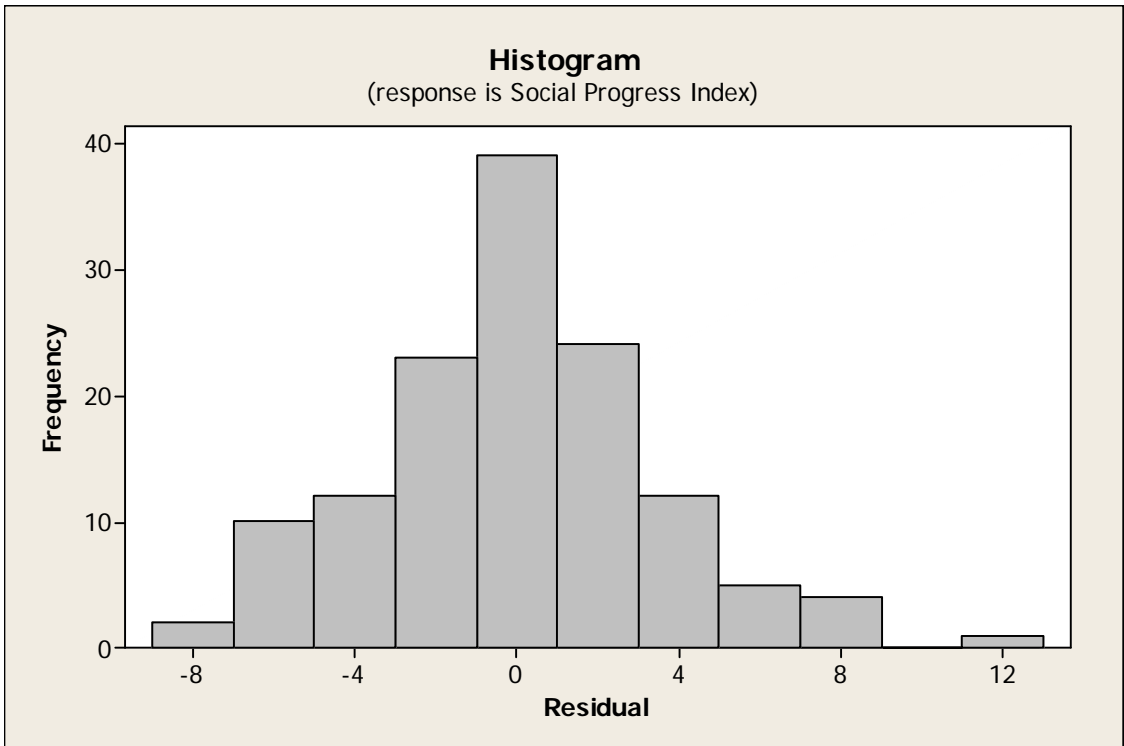
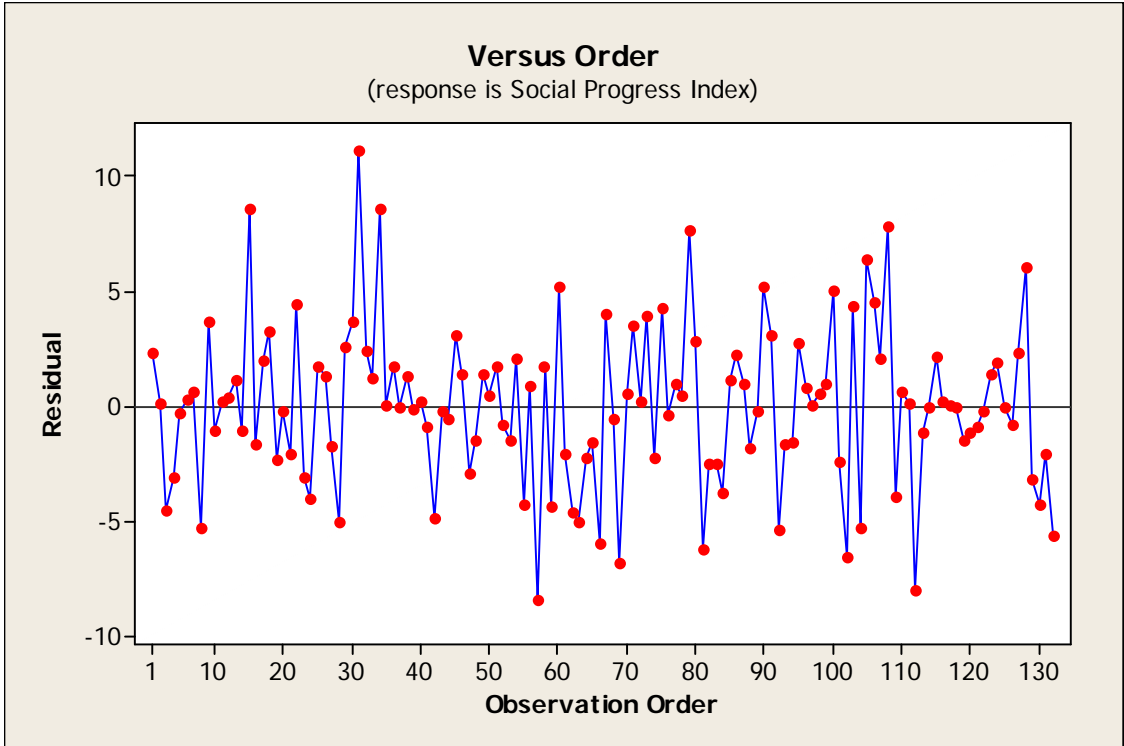
Direção: Da análise das correlações acima percebemos que quase todas possuem associações positivas, ou seja, o crescimento de uma variável é acompanhado do crescimento da outra. O que nos parece é que não há nenhuma associação negativa, ao menos de evidência visual.

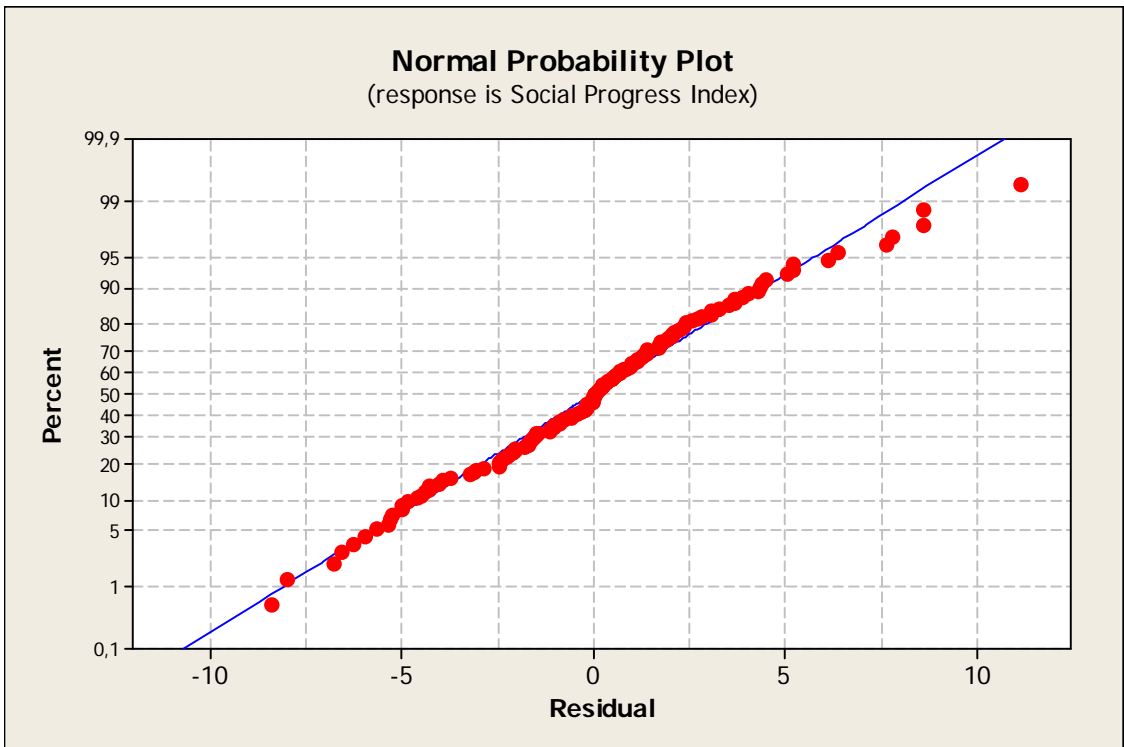
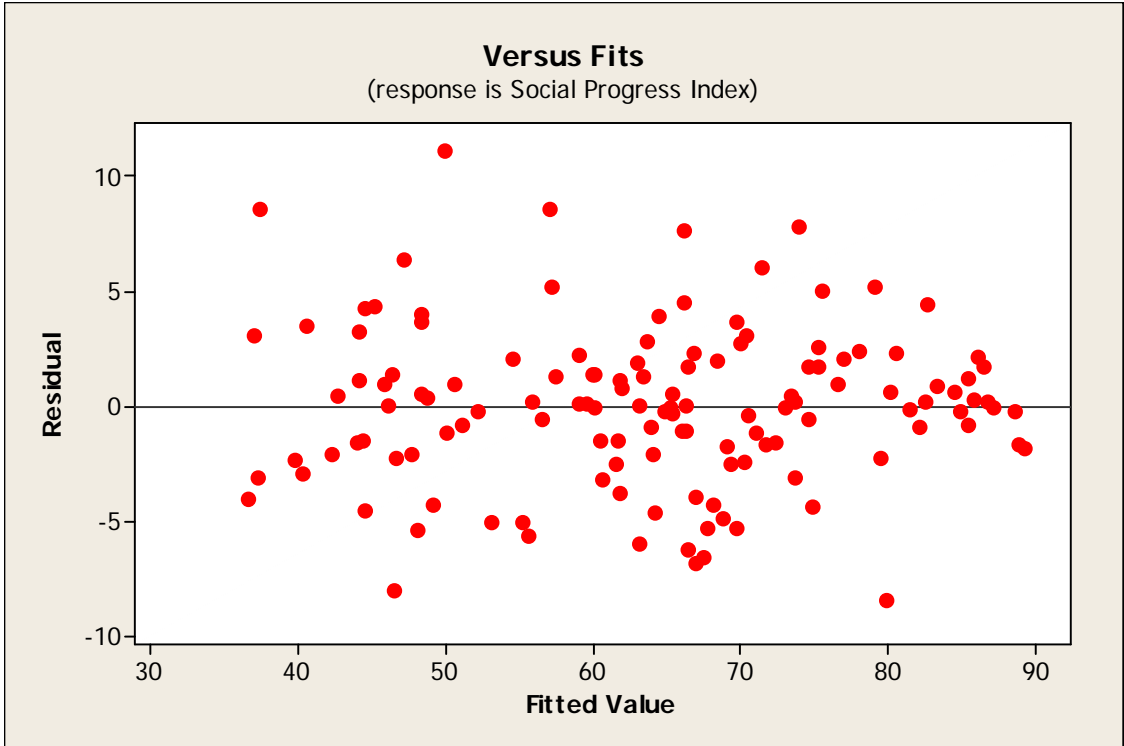
Intensidade: Alguns dos gráficos acima parecem indicar a existência de relações lineares, como Access to information and Communications, Access to Advanced Education, Primary school enrollment, Upper secondary school enrollment, Internet users, Years of tertiary, Inequality in the attainment of education, Number of globally ranked universities, Unemployment Total, Total Labor Force, Unemployment youth total.

Forma: Com exceção dos gráficos Social Progress Index, IDH, EPI, Happy Planet Index, GINI Index e, que têm uma distribuição que parece mais pulverizada, os demais apresentam conglomerados que sugerem relações lineares.

Valores Atípicos: Todos os gráficos indicam a existência de valores atípicos, ou seja, indivíduos ou empresas, que estão localizados longe dos demais.







**Scatterplot of Social Progress Index vs Adult literacy rate**

**Scatterplot of IDH - 2013 vs Access to Information and Com\_1**

**Scatterplot of 1. Governança vs Access to Advanced Education\_1**

**Scatterplot of Happy Planet Index vs Primary scholl**

**Scatterplot of GINNI Index vs Lower secondary**

**Scatterplot of Social Progress Index vs Upper secondary**

**Scatterplot of Social Progress Index vs Internet users**

**Scatterplot of Social Progress Index vs Years of teartary**

**Scatterplot of Social Progress Index vs Inequality in the attain**

**Scatterplot of Social Progress Index vs Number of globally ranked**

**Scatterplot of Social Progress Index vs Unemployment Total**

**Scatterplot of Social Progress Index vs Total Labor Force (% G**

**Scatterplot of Social Progress Index vs Unemployment, youth total**

**Scatterplot of IDH - 2013 vs Access to In; 1. Governanç vs Access to A**

**Regression Analysis: Social Progr versus Happy Planet; GINNI Index; ...**

The regression equation is

Social Progress Index = 16,9 + 0,146 Happy Planet Index + 0,0901 GINNI Index  
+ 0,330 Access to Information and Com\_1  
+ 0,015 Access to Advanced Education\_1  
+ 0,0134 Adult literacy rate + 0,0509 Primary scholl  
+ 0,0318 Lower secondary + 0,0686 Upper secondary  
+ 0,0840 Internet users + 0,0031 Years of teartary  
+ 0,0537 Inequality + 0,0502 Number of globally ranked  
- 0,0859 Unemployment Total  
- 0,223 Total Labor Force (% Global\_1  
+ 0,0426 Unemployment, youth total

Predictor	Coef	SE Coef	T	P
Constant	16,851	3,158	5,34	0,000
Happy Planet Index	0,14621	0,04331	3,38	0,001
GINNI Index	0,09008	0,04760	1,89	0,061
Access to Information and Com_1	0,32975	0,04017	8,21	0,000
Access to Advanced Education_1	0,0148	0,1064	0,14	0,890
Adult literacy rate	0,01341	0,03426	0,39	0,696
Primary scholl	0,05087	0,02659	1,91	0,058
Lower secondary	0,03184	0,03785	0,84	0,402
Upper secondary	0,06863	0,03962	1,73	0,086
Internet users	0,08400	0,03219	2,61	0,010
Years of teartary	0,00312	0,04225	0,07	0,941
Inequality	0,05368	0,03478	1,54	0,125
Number of globally ranked	0,05016	0,02510	2,00	0,048
Unemployment Total	-0,08588	0,04844	-1,77	0,079
Total Labor Force (% Global_1	-0,2233	0,1545	-1,45	0,151
Unemployment, youth total	0,04257	0,04768	0,89	0,374

S = 3,67988 R-Sq = 94,1% R-Sq(adj) = 93,3%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	15	24846,7	1656,4	122,32	0,000
Residual Error	116	1570,8	13,5		
Total	131	26417,5			

Source	DF	Seq SS
Happy Planet Index	1	2326,8
GINNI Index	1	2735,0
Access to Information and Com_1	1	17910,1
Access to Advanced Education_1	1	1231,2
Adult literacy rate	1	56,0
Primary scholl	1	126,1
Lower secondary	1	82,2
Upper secondary	1	52,0
Internet users	1	143,4
Years of teartary	1	13,6
Inequality	1	9,2
Number of globally ranked	1	32,8
Unemployment Total	1	87,9
Total Labor Force (% Global_1	1	29,6
Unemployment, youth total	1	10,8

#### Unusual Observations

Obs	Happy Planet Index	Social Progress Index	Fit	SE Fit	Residual	St Resid
15	22,6	65,600	56,998	1,584	8,602	2,59R
26	44,7	58,670	57,412	3,121	1,258	0,65 X
31	56,2	61,070	49,939	1,992	11,131	3,60R
34	37,2	45,950	37,346	1,799	8,604	2,68R
57	55,2	71,400	79,832	1,105	-8,432	-2,40R
79	36,6	73,680	66,066	0,830	7,614	2,12R
102	34,5	60,790	67,370	1,712	-6,580	-2,02R
108	40,2	81,650	73,868	0,879	7,782	2,18R
112	37,6	38,450	46,477	1,260	-8,027	-2,32R

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

## Normplot of Residuals for Social Progress Index

## Residuals vs Fits for Social Progress Index

## Residual Histogram for Social Progress Index

## Residuals vs Order for Social Progress Index

## Regression Analysis: Social Progr versus Happy Planet; GINNI Index; ...

The regression equation is

$$\begin{aligned} \text{Social Progress Index} = & 16,9 + 0,146 \text{ Happy Planet Index} + 0,0901 \text{ GINNI Index} \\ & + 0,330 \text{ Access to Information and Com}_1 \\ & + 0,015 \text{ Access to Advanced Education}_1 \\ & + 0,0134 \text{ Adult literacy rate} + 0,0509 \text{ Primary scholl} \\ & + 0,0318 \text{ Lower secondary} + 0,0686 \text{ Upper secondary} \\ & + 0,0840 \text{ Internet users} + 0,0031 \text{ Years of teartinary} \\ & + 0,0537 \text{ Inequality} + 0,0502 \text{ Number of globally ranked} \\ & - 0,0859 \text{ Unemployment Total} \\ & - 0,223 \text{ Total Labor Force (\% Global}_1 \\ & + 0,0426 \text{ Unemployment, youth total} \end{aligned}$$

Predictor	Coef	SE Coef	T	P
Constant	16,851	3,158	5,34	0,000
Happy Planet Index	0,14621	0,04331	3,38	0,001
GINNI Index	0,09008	0,04760	1,89	0,061
Access to Information and Com <sub>1</sub>	0,32975	0,04017	8,21	0,000
Access to Advanced Education <sub>1</sub>	0,0148	0,1064	0,14	0,890
Adult literacy rate	0,01341	0,03426	0,39	0,696
Primary scholl	0,05087	0,02659	1,91	0,058
Lower secondary	0,03184	0,03785	0,84	0,402
Upper secondary	0,06863	0,03962	1,73	0,086
Internet users	0,08400	0,03219	2,61	0,010
Years of teartinary	0,00312	0,04225	0,07	0,941
Inequality	0,05368	0,03478	1,54	0,125
Number of globally ranked	0,05016	0,02510	2,00	0,048
Unemployment Total	-0,08588	0,04844	-1,77	0,079
Total Labor Force (% Global <sub>1</sub> )	-0,2233	0,1545	-1,45	0,151
Unemployment, youth total	0,04257	0,04768	0,89	0,374

S = 3,67988    R-Sq = 94,1%    R-Sq(adj) = 93,3%

### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	15	24846,7	1656,4	122,32	0,000
Residual Error	116	1570,8	13,5		
Total	131	26417,5			

Source	DF	Seq SS
Happy Planet Index	1	2326,8
GINNI Index	1	2735,0
Access to Information and Com_1	1	17910,1
Access to Advanced Education_1	1	1231,2
Adult literacy rate	1	56,0
Primary scholl	1	126,1
Lower secondary	1	82,2
Upper secondary	1	52,0
Internet users	1	143,4
Years of teartinary	1	13,6
Inequality	1	9,2
Number of globally ranked	1	32,8
Unemployment Total	1	87,9
Total Labor Force (% Global_1	1	29,6
Unemployment, youth total	1	10,8

Obs	Happy Planet Index	Social Progress Index	Fit	SE Fit	Residual	St Resid
1	54,1	69,130	66,833	1,208	2,297	0,66
2	52,2	59,130	59,060	1,638	0,070	0,02
3	33,2	39,930	44,437	1,097	-4,507	-1,28
4	54,1	70,590	73,689	0,949	-3,099	-0,87
5	46,0	65,030	65,365	1,013	-0,335	-0,09
6	42,0	86,100	85,783	1,664	0,317	0,10
7	47,1	85,110	84,459	1,113	0,651	0,19
8	40,9	62,440	67,736	1,045	-5,296	-1,50
9	56,3	52,040	48,366	1,034	3,674	1,04
10	37,4	65,200	66,253	1,179	-1,053	-0,30
11	37,1	82,630	82,402	0,925	0,228	0,06
12	31,1	49,110	48,728	1,641	0,382	0,12
13	43,6	62,900	61,784	0,977	1,116	0,31
14	42,4	64,990	66,032	1,411	-1,042	-0,31
15	22,6	65,600	56,998	1,584	8,602	2,59R
16	52,9	69,970	71,653	1,087	-1,683	-0,48
17	34,1	70,240	68,297	0,924	1,943	0,55
18	31,8	47,330	44,093	1,479	3,237	0,96
19	30,5	37,330	39,705	2,167	-2,375	-0,80
20	40,3	51,890	52,140	1,168	-0,250	-0,07
21	33,7	45,510	47,631	0,897	-2,121	-0,59
22	43,6	86,950	82,553	1,108	4,397	1,25
23	25,3	34,170	37,317	1,326	-3,147	-0,92
24	24,7	32,600	36,624	1,282	-4,024	-1,17
25	53,9	76,300	74,542	0,988	1,758	0,50
26	44,7	58,670	57,412	3,121	1,258	0,65 X
27	59,8	67,240	68,957	1,220	-1,717	-0,49
28	34,5	47,990	53,012	1,117	-5,022	-1,43
29	64,0	77,750	75,219	1,268	2,531	0,73
30	40,6	73,310	69,669	1,326	3,641	1,06
31	56,2	61,070	49,939	1,992	11,131	3,60R
32	39,4	80,410	78,038	0,938	2,372	0,67
33	36,6	86,550	85,298	1,042	1,252	0,35
34	37,2	45,950	37,346	1,799	8,604	2,68R
35	50,7	63,030	63,035	0,822	-0,005	-0,00
36	52,5	68,150	66,390	0,764	1,760	0,49
37	39,6	59,970	60,026	1,569	-0,056	-0,02
38	58,9	64,700	63,408	1,100	1,292	0,37
39	34,9	81,280	81,457	0,971	-0,177	-0,05
40	42,7	86,910	86,738	0,888	0,172	0,05
41	46,5	81,110	82,020	0,993	-0,910	-0,26

42	46,0	63,940	68,809	0,956	-4,869	-1,37
43	47,2	84,610	84,799	1,156	-0,189	-0,05
44	40,3	55,960	56,560	1,165	-0,600	-0,17
45	40,5	73,430	70,377	1,332	3,053	0,89
46	56,9	61,370	60,004	1,233	1,366	0,39
47	30,0	37,410	40,310	1,369	-2,900	-0,85
48	51,2	60,060	61,577	1,388	-1,517	-0,45
49	56,0	61,280	59,881	1,218	1,399	0,40
50	37,4	73,870	73,382	0,717	0,488	0,14
51	40,2	88,070	86,381	1,212	1,689	0,49
52	50,9	50,240	51,055	1,991	-0,815	-0,26
53	55,5	58,980	60,460	1,188	-1,480	-0,42
54	41,7	56,650	54,585	1,149	2,065	0,59
55	49,2	44,840	49,127	1,206	-4,287	-1,23
56	42,4	84,050	83,207	0,893	0,843	0,24
57	55,2	71,400	79,832	1,105	-8,432	-2,40R
58	46,4	76,930	75,241	1,595	1,689	0,51
59	58,5	70,390	74,772	1,413	-4,382	-1,29
60	47,5	84,210	79,034	1,317	5,176	1,51
61	51,7	61,920	63,973	0,960	-2,053	-0,58
62	34,7	59,470	64,085	1,338	-4,615	-1,35
63	38,0	50,200	55,226	1,014	-5,026	-1,42
64	43,8	77,180	79,408	1,210	-2,228	-0,64
65	27,1	70,660	72,262	1,582	-1,602	-0,48
66	49,1	57,080	63,033	1,033	-5,953	-1,69
67	49,1	52,410	48,372	0,912	4,038	1,13
68	34,9	73,910	74,511	0,961	-0,601	-0,17
69	42,9	60,050	66,859	0,981	-6,809	-1,92
70	42,5	48,940	48,372	1,592	0,568	0,17
71	35,2	44,020	40,533	1,852	3,487	1,10
72	34,6	73,760	73,568	1,021	0,192	0,05
73	28,3	68,330	64,442	1,376	3,888	1,14
74	46,8	44,280	46,572	1,039	-2,292	-0,65
75	42,5	48,790	44,525	1,170	4,265	1,22
76	40,5	70,000	70,438	1,256	-0,438	-0,13
77	26,0	46,850	45,880	1,415	0,970	0,29
78	32,3	43,110	42,633	1,904	0,477	0,15
79	36,6	73,680	66,066	0,830	7,614	2,12R
80	52,9	66,410	63,613	1,000	2,797	0,79
81	48,0	60,120	66,392	1,111	-6,272	-1,79
82	26,8	58,970	61,464	1,683	-2,494	-0,76
83	42,5	66,800	69,296	0,943	-2,496	-0,70
84	47,9	58,010	61,773	1,525	-3,763	-1,12
85	35,7	45,230	44,105	1,512	1,125	0,34
86	38,9	61,190	58,972	1,366	2,218	0,65
87	45,6	51,580	50,586	1,207	0,994	0,29
88	43,1	87,370	89,193	1,315	-1,823	-0,53
89	51,6	88,240	88,476	1,408	-0,236	-0,07
90	57,1	62,330	57,144	0,948	5,186	1,46
91	26,8	40,100	37,045	1,411	3,055	0,90
92	33,6	42,650	48,031	1,383	-5,381	-1,58
93	51,4	87,120	88,831	1,108	-1,711	-0,49
94	54,1	42,400	43,978	1,118	-1,578	-0,45
95	57,8	72,580	69,882	1,105	2,698	0,77
96	45,8	62,650	61,876	0,990	0,774	0,22
97	52,4	66,290	66,284	0,873	0,006	0,00
98	52,4	65,860	65,326	1,108	0,534	0,15
99	42,6	77,440	76,523	0,730	0,917	0,25
100	38,7	80,490	75,441	0,869	5,049	1,41
101	42,2	67,720	70,155	1,038	-2,435	-0,69
102	34,5	60,790	67,370	1,712	-6,580	-2,02R
103	36,9	49,460	45,132	1,382	4,328	1,27
104	46,0	64,380	69,658	1,383	-5,278	-1,55

105	33,3	53,520	47,183	0,972	6,337	1,79
106	41,3	70,610	66,115	0,953	4,495	1,26
107	40,1	78,930	76,911	1,050	2,019	0,57
108	40,2	81,650	73,868	0,879	7,782	2,18R
109	28,2	62,960	66,917	1,703	-3,957	-1,21
110	44,1	80,770	80,110	1,499	0,660	0,20
111	49,4	59,710	59,593	1,136	0,117	0,03
112	37,6	38,450	46,477	1,260	-8,027	-2,32R
113	42,5	48,870	50,008	1,123	-1,138	-0,32
114	46,2	87,080	87,132	1,051	-0,052	-0,01
115	50,3	88,190	86,083	1,147	2,107	0,60
116	47,8	56,050	55,830	1,298	0,220	0,06
117	30,7	46,060	46,029	1,181	0,031	0,01
118	53,5	65,140	65,200	1,017	-0,060	-0,02
119	28,2	42,800	44,322	1,016	-1,522	-0,43
120	30,3	69,880	71,059	1,026	-1,179	-0,33
121	48,3	62,960	63,865	1,339	-0,905	-0,26
122	47,6	64,620	64,812	0,978	-0,192	-0,05
123	31,5	47,750	46,356	1,091	1,394	0,40
124	37,6	64,910	63,000	1,592	1,910	0,58
125	31,8	72,920	72,958	1,397	-0,038	-0,01
126	47,9	84,560	85,348	1,372	-0,788	-0,23
127	37,3	82,770	80,435	1,760	2,335	0,72
128	39,3	77,510	71,433	0,922	6,077	1,71
129	46,0	57,340	60,567	1,530	-3,227	-0,96
130	56,9	63,780	68,089	0,785	-4,309	-1,20
131	43,0	40,230	42,307	1,455	-2,077	-0,61
132	37,7	49,880	55,571	1,725	-5,691	-1,75

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

### Normplot of Residuals for Social Progress Index

### Residuals vs Fits for Social Progress Index

### Residual Histogram for Social Progress Index

### Residuals vs Order for Social Progress Index

## 4.2 Linhas de Tendência

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.



### 4.3 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: Access to In; Access to Ad; Adult litera; Primary scho; ...

Access to Advanc	Access to Inform	Access to Advanc	Adult literacy r
	0,794		
	0,000		
Adult literacy r	0,652	0,821	
	0,000	0,000	
Primary scholl	0,403	0,532	0,615
	0,000	0,000	0,000
Lower secondary	0,693	0,735	0,761
	0,000	0,000	0,000
Upper secondary	0,757	0,853	0,775
	0,000	0,000	0,000
Internet users	0,898	0,866	0,720
	0,000	0,000	0,000
Years of teartia	0,587	0,812	0,506
	0,000	0,000	0,000
Inequality	0,700	0,836	0,822
	0,000	0,000	0,000
Number of global	0,533	0,701	0,392
	0,000	0,000	0,000
Unemployment Tot	0,138	0,072	0,173
	0,114	0,411	0,047
Total Labor Forc	-0,120	0,039	-0,002
	0,171	0,653	0,982
Unemployment, yo	0,251	0,193	0,313
	0,004	0,027	0,000
	Primary scholl	Lower secondary	Upper secondary
Lower secondary	0,585		
	0,000		
Upper secondary	0,470	0,824	
	0,000	0,000	
Internet users	0,450	0,701	0,801
	0,000	0,000	0,000

Years of teartia	0,308 0,000	0,451 0,000	0,634 0,000
Inequality	0,530 0,000	0,689 0,000	0,786 0,000
Number of global	0,357 0,000	0,466 0,000	0,523 0,000
Unemployment Tot	-0,032 0,718	0,138 0,113	0,131 0,135
Total Labor Forc	0,092 0,297	0,060 0,493	-0,019 0,830
Unemployment, yo	0,099 0,258	0,319 0,000	0,278 0,001

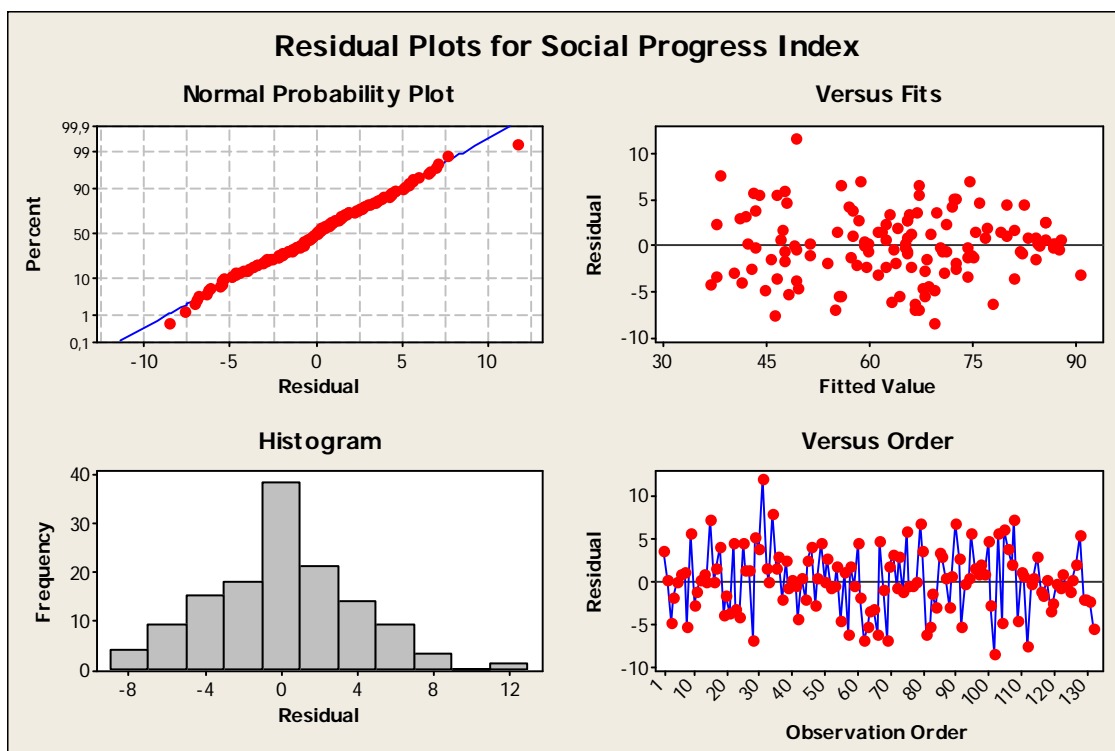
	Internet users	Years of teartia	Inequality
Years of teartia	0,662 0,000		
Inequality	0,731 0,000	0,492 0,000	
Number of global	0,633 0,000	0,590 0,000	0,383 0,000
Unemployment Tot	0,085 0,333	-0,071 0,418	0,186 0,033
Total Labor Forc	-0,033 0,709	-0,003 0,974	-0,110 0,210
Unemployment, yo	0,226 0,009	-0,004 0,967	0,297 0,001

	Number of global	Unemployment Tot	Total Labor Forc
Unemployment Tot	-0,045 0,607		
Total Labor Forc	0,335 0,000	-0,132 0,132	
Unemployment, yo	0,035 0,686	0,927 0,000	-0,114 0,193

Cell Contents: Pearson correlation  
P-Value

#### 4.4 Regressão de Mínimos Quadrados

A correlação mede a direção e a intensidade da relação linear (linha reta) entre duas variáveis quantitativas. Se um diagrama de dispersão mostra uma relação linear, é interessante resumirmos esse padrão geral traçando uma reta no diagrama de dispersão. Uma reta de regressão resume a relação entre duas variáveis, mas somente em um contexto específico: quando uma das variáveis ajuda a explicarmos ou predizermos a outra, ou seja, a regressão descreve uma relação entre uma variável explanatória e uma variável resposta.



#### Regression Analysis: Democracy In versus Social Progr; Political ri; ...

The regression equation is

Democracy Index = - 8,81 + 0,147 Social Progress Index  
+ 0,385 Political rights (1=full rights  
+ 0,0117 Freedom of speech (0=low; 2=hig  
- 0,0049 Freedom of assembly/association  
+ 0,0778 Freedom of movement (0=low; 4=h  
+ 0,120 Private property rights (0=none  
+ 0,0134 Freedom over life choices (% sa  
+ 0,0281 Freedom of religion (1=low; 4=h + 0,0687 IDH - 2013  
+ 0,150 Governança - 0,144 EPI Score  
+ 0,0765 Happy Planet Index - 0,0368 GINNI Index + 0,148 OHI

Predictor	Coef	SE Coef	T	P
Constant	-8,806	7,352	-1,20	0,233
Social Progress Index	0,1466	0,1570	0,93	0,352
Political rights (1=full rights	0,38521	0,04145	9,29	0,000
Freedom of speech (0=low; 2=hig	0,01174	0,03044	0,39	0,701
Freedom of assembly/association	-0,00488	0,02661	-0,18	0,855
Freedom of movement (0=low; 4=h	0,07775	0,03362	2,31	0,022
Private property rights (0=none	0,11962	0,06522	1,83	0,069
Freedom over life choices (% sa	0,01340	0,03769	0,36	0,723
Freedom of religion (1=low; 4=h	0,02806	0,02986	0,94	0,349
IDH - 2013	0,06871	0,09282	0,74	0,461
Governança	0,1497	0,1130	1,32	0,188
EPI Score	-0,14375	0,08008	-1,80	0,075
Happy Planet Index	0,07648	0,03752	2,04	0,044
GINNI Index	-0,03678	0,03670	-1,00	0,318
OHI	0,1480	0,1008	1,47	0,145

S = 7,43711 R-Sq = 91,8% R-Sq(adj) = 90,8%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	14	72200,9	5157,2	93,24	0,000

Residual Error	117	6471,3	55,3
Total	131	78672,3	

Source	DF	Seq SS
Social Progress Index	1	49033,1
Political rights (1=full rights	1	20993,2
Freedom of speech (0=low; 2=hig	1	39,9
Freedom of assembly/association	1	95,0
Freedom of movement (0=low; 4=h	1	313,6
Private property rights (0=none	1	825,8
Freedom over life choices (% sa	1	67,0
Freedom of religion (1=low; 4=h	1	40,9
IDH - 2013	1	43,3
Governança	1	26,7
EPI Score	1	285,1
Happy Planet Index	1	278,5
GINNI Index	1	39,6
OHI	1	119,3

#### Unusual Observations

##### Social

##### Progress Democracy

Obs	Index	Index	Fit	SE Fit	Residual	St Resid
15	59	75,592	60,302	3,161	15,290	2,27R
52	32	76,185	58,642	2,669	17,543	2,53R
77	26	50,948	33,512	3,241	17,436	2,60R
104	57	3,910	18,507	2,523	-14,597	-2,09R
116	42	10,427	26,034	2,985	-15,608	-2,29R
123	27	44,194	27,061	2,459	17,133	2,44R
125	72	13,626	29,545	3,183	-15,920	-2,37R

R denotes an observation with a large standardized residual.

### General Regression Analysis: Social Progr versus Access to In; Access to Ad; ..

Regression Equation

Social Progress Index = 24,2166 + 0,340012 Access to Information and Com\_1 + 0,0402113 Access to Advanced Education\_1 + 0,0496665 Adult literacy rate + 0,053172 Primary scholl + 0,0767042 Lower secondary + 0,0481312 Upper secondary + 0,0554064 Internet users - 0,000376823 Years of teartary + 0,0175204 Inequality + 0,0528492 Number of globally ranked - 0,0778691 Unemployment Total - 0,261848 Total Labor Force (% Global\_1 + 0,0267877 Unemployment, youth total

Coefficients

Term	Coef	SE Coef	T	P
Constant	24,2166	2,43649	9,93913	0,000
Access to Information and Com_1	0,3400	0,04214	8,06803	0,000
Access to Advanced Education_1	0,0402	0,11165	0,36016	0,719
Adult literacy rate	0,0497	0,03474	1,42960	0,155
Primary scholl	0,0532	0,02794	1,90280	0,060
Lower secondary	0,0767	0,03797	2,02006	0,046
Upper secondary	0,0481	0,04127	1,16621	0,246
Internet users	0,0554	0,03280	1,68942	0,094
Years of teartary	-0,0004	0,04431	-0,00851	0,993
Inequality	0,0175	0,03498	0,50089	0,617
Number of globally ranked	0,0528	0,02644	1,99863	0,048
Unemployment Total	-0,0779	0,05041	-1,54464	0,125
Total Labor Force (% Global_1	-0,2618	0,16227	-1,61370	0,109
Unemployment, youth total	0,0268	0,04981	0,53777	0,592

Summary of Model

S = 3,87814      R-Sq = 93,28%      R-Sq(adj) = 92,54%  
 PRESS = 2330,77      R-Sq(pred) = 91,18%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F
Regression	13	24642,8	24642,8	1895,60	126,037
Access to Information and Com_1	1	22182,0	979,0	979,00	65,093
Access to Advanced Education_1	1	1643,7	2,0	1,95	0,130
Adult literacy rate	1	155,3	30,7	30,74	2,044
Primary scholl	1	182,4	54,5	54,45	3,621
Lower secondary	1	154,8	61,4	61,37	4,081
Upper secondary	1	24,7	20,5	20,46	1,360
Internet users	1	94,5	42,9	42,93	2,854
Years of teartary	1	0,9	0,0	0,00	0,000
Inequality	1	3,6	3,8	3,77	0,251
Number of globally ranked	1	34,6	60,1	60,08	3,995
Unemployment Total	1	121,9	35,9	35,88	2,386
Total Labor Force (% Global_1	1	39,9	39,2	39,16	2,604
Unemployment, youth total	1	4,3	4,3	4,35	0,289
Error	118	1774,7	1774,7	15,04	
Total	131	26417,5			

Source	P
Regression	0,000000
Access to Information and Com_1	0,000000
Access to Advanced Education_1	0,719371
Adult literacy rate	0,155473

Primary scholl	0,059504
Lower secondary	0,045642
Upper secondary	0,245880
Internet users	0,093780
Years of teartinary	0,993228
Inequality	0,617385
Number of globally ranked	0,047947
Unemployment Total	0,125110
Total Labor Force (% Global_1	0,109263
Unemployment, youth total	0,591746
Error	
Total	

Fits and Diagnostics for Unusual Observations

Social Progress Index						
Obs	Index	Fit	SE Fit	Residual	St Resid	
19	37,33	41,2896	2,24230	-3,9596	-1,25138	X
26	58,67	57,5611	3,27629	1,1089	0,53439	X
31	61,07	49,3040	2,07653	11,7660	3,59227	R
34	45,95	38,2187	1,88084	7,7313	2,27960	R
102	60,79	69,3503	1,67315	-8,5603	-2,44675	R
112	38,45	46,1193	1,31531	-7,6693	-2,10218	R

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

### Residual Plots for Social Progress Index

A tabela acima nos mostra que, a regressão nada mais é do que a expressão numérica da reta de tendência que vimos nos itens acima. Esta equação tem um poder explicativo de 91,8%, que é o R-Quadrado.

### 4.5 Dendograma

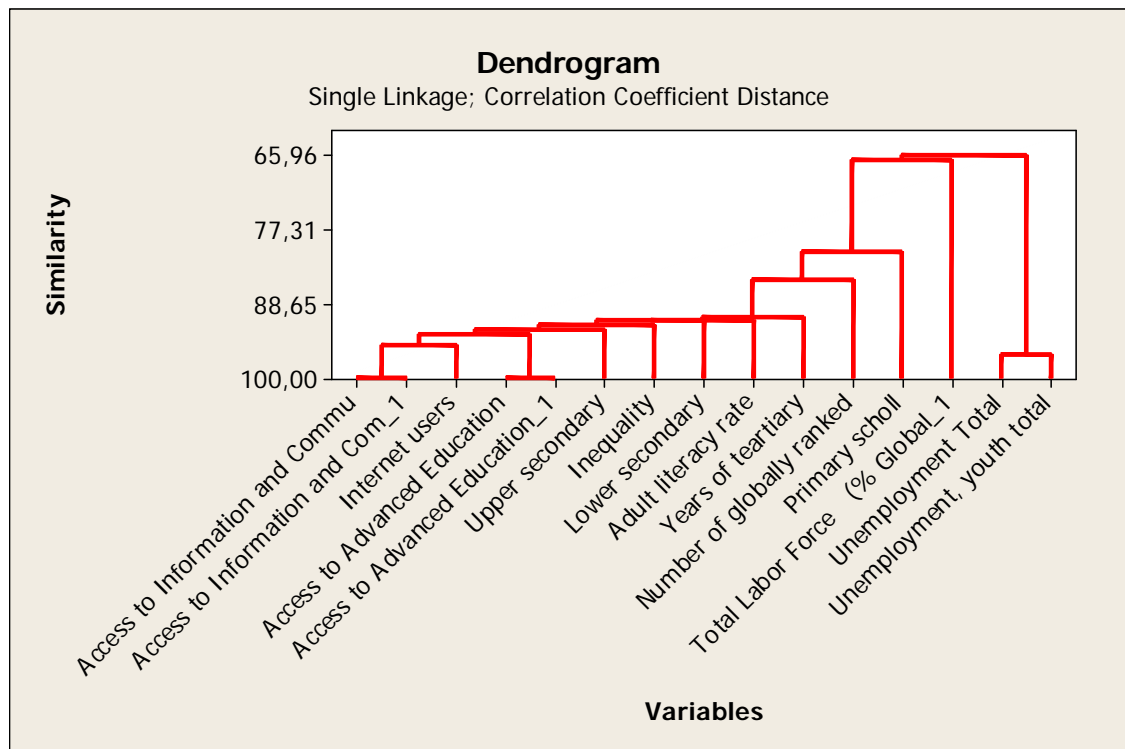
Um **Dendrograma** (dendr(o) = árvore) é um tipo específico de diagrama ou representação icônica que organiza determinados fatores e variáveis. É um diagrama de similaridade.

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.

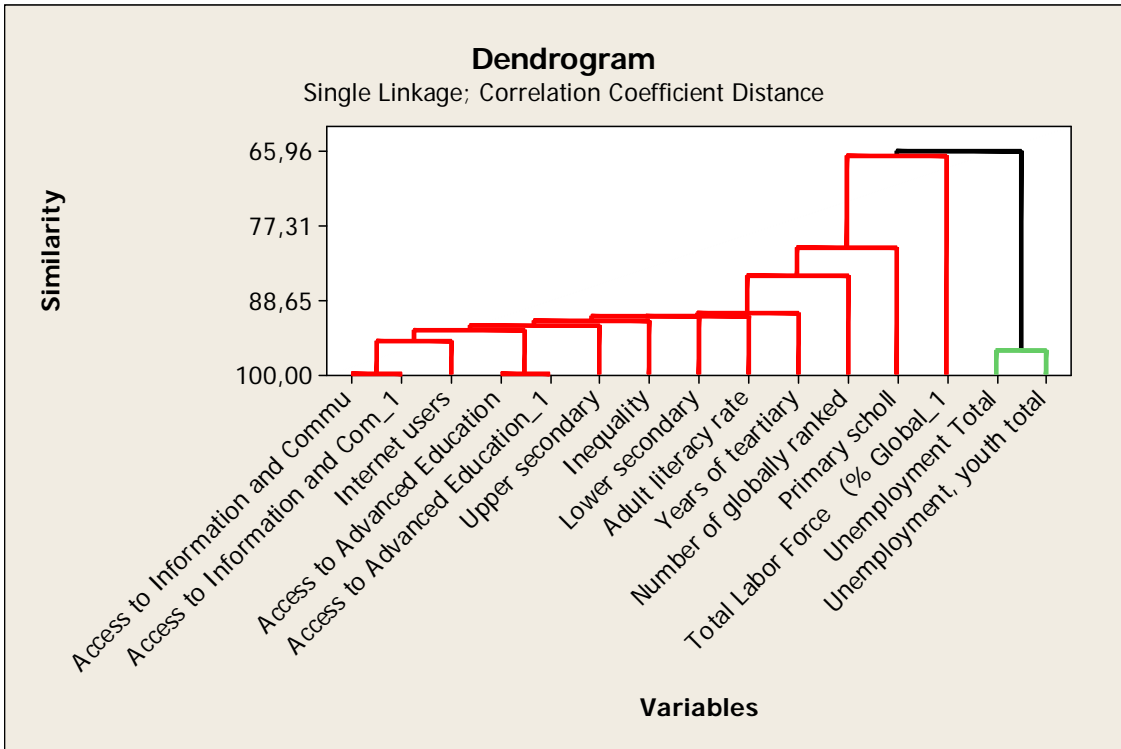
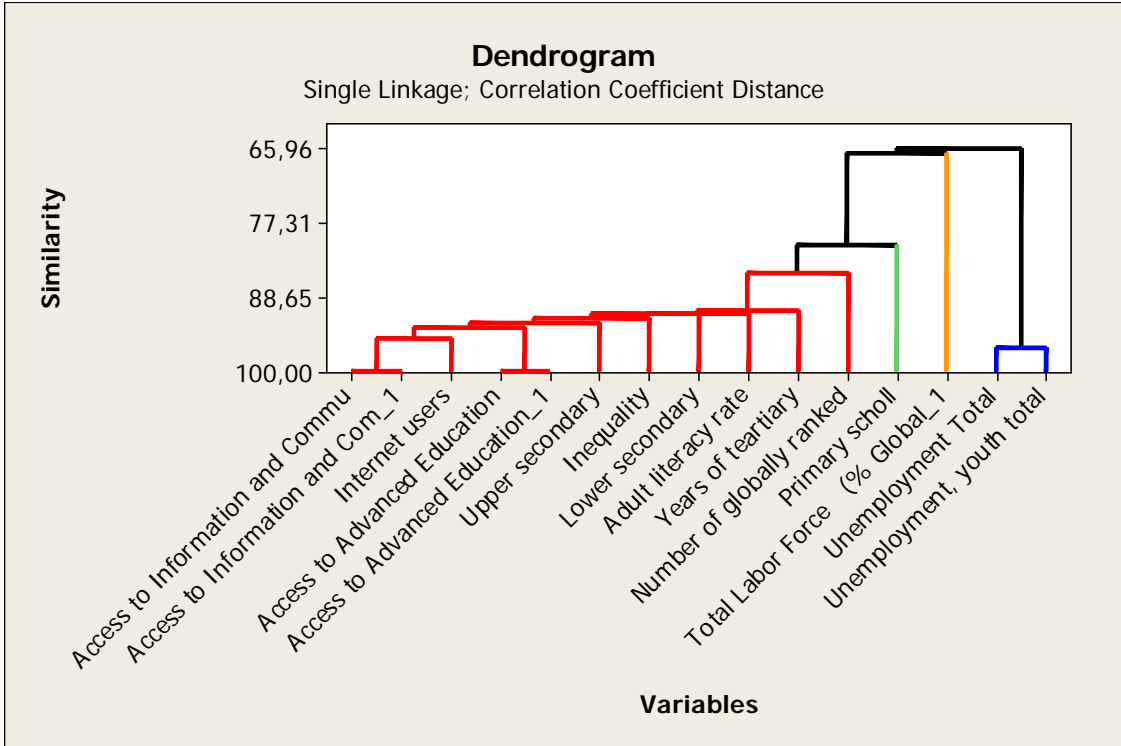
Segue abaixo o Dendograma de todas as variáveis analisadas:

### ANÁLISE DE CLUSTERS DE VARIÁVEIS E PAÍSES VARIÁVEIS



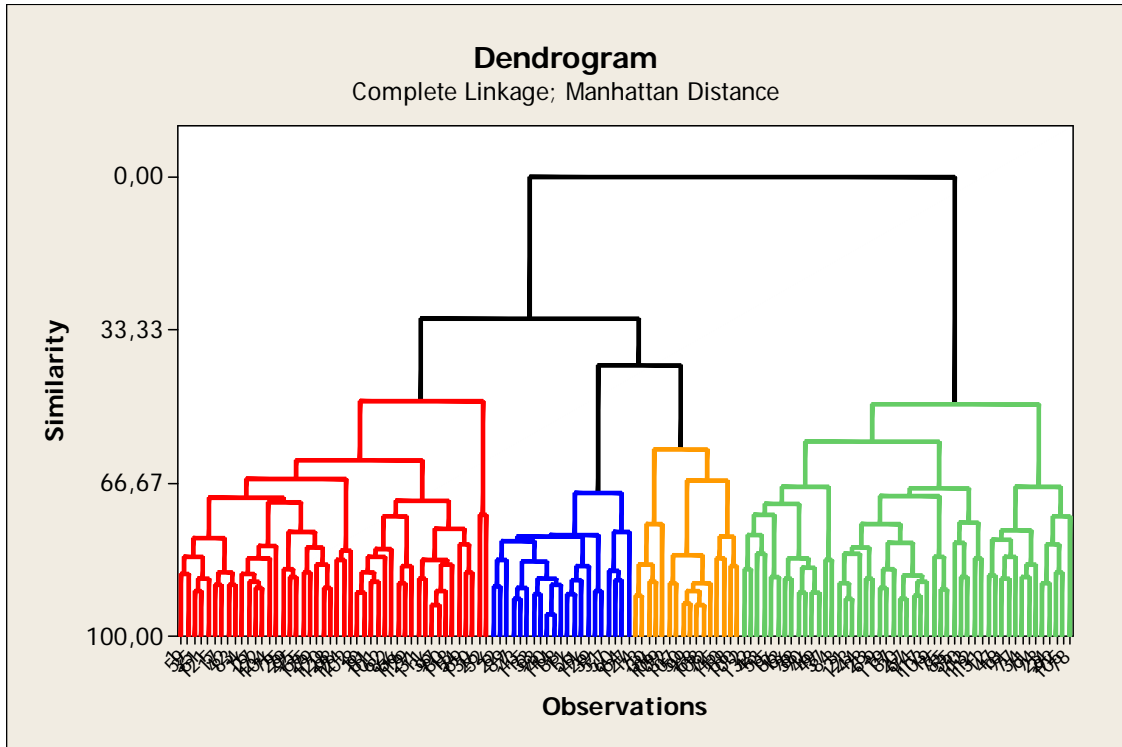
No gráfico acima verificamos uma grande correlação num primeiro grupo entre as variáveis Social Progress Index, EPI Score e IDH – 2013, no segundo grupo entre as variáveis Access to Advanced e Access to Advanced Educations e Governança, no terceiro grupo todas as demais, nota se também uma correlação entre o primeiro e o segundo grupo de variáveis e entre esse com o terceiro grupo, abaixo os níveis de similaridades entre essas correlações.





## 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.



## Cluster Analysis of Variables: Social Progr; IDH - 2013; 1. Governanç; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	20	96,3283	0,073435	19 21	19	2
2	19	95,8701	0,082598	1 6	1	2
3	18	95,8168	0,083664	1 9	1	3
4	17	95,4440	0,091120	1 15	1	4
5	16	94,4907	0,110186	1 3	1	5
6	15	93,9615	0,120770	1 10	1	6
7	14	93,5636	0,128728	1 2	1	7
8	13	92,6545	0,146910	1 14	1	8
9	12	91,8212	0,163576	1 17	1	9
10	11	91,1957	0,176087	1 11	1	10
11	10	91,1844	0,176312	1 13	1	11
12	9	90,6099	0,187802	1 16	1	12
13	8	89,4345	0,211311	1 5	1	13

14	7	87,9933	0,240134	4	18	4	2
15	6	87,0270	0,259460	1	4	1	15
16	5	80,7512	0,384975	1	12	1	16
17	4	68,9252	0,621497	1	7	1	17
18	3	66,7326	0,665347	1	20	1	18
19	2	65,9648	0,680705	1	19	1	20
20	1	52,9445	0,941111	1	8	1	21

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2 4	2	2
2	13	100,000	0,000000	1 3	1	2
3	12	96,328	0,073435	13 15	13	2
4	11	94,907	0,101850	1 9	1	3
5	10	93,290	0,134202	1 2	1	5
6	9	92,654	0,146910	1 8	1	6
7	8	91,821	0,163576	1 11	1	7
8	7	91,184	0,176312	1 7	1	8
9	6	91,118	0,177634	1 5	1	9
10	5	90,610	0,187802	1 10	1	10
11	4	85,045	0,299102	1 12	1	11
12	3	80,751	0,384975	1 6	1	12
13	2	66,733	0,665347	1 14	1	13
14	1	65,965	0,680705	1 13	1	15

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2 4	2	2
2	13	100,000	0,000000	1 3	1	2
3	12	96,328	0,073435	13 15	13	2
4	11	94,907	0,101850	1 9	1	3
5	10	93,290	0,134202	1 2	1	5
6	9	92,654	0,146910	1 8	1	6
7	8	91,821	0,163576	1 11	1	7
8	7	91,184	0,176312	1 7	1	8
9	6	91,118	0,177634	1 5	1	9
10	5	90,610	0,187802	1 10	1	10

11	4	85,045	0,299102	1	12	1	11
12	3	80,751	0,384975	1	6	1	12
13	2	66,733	0,665347	1	14	1	13
14	1	65,965	0,680705	1	13	1	15

Final Partition

Cluster 1

Access to Information and Commu Access to Advanced Education Access to  
Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate  
Primary scholl Lower secondary Upper secondary Internet users Years of  
teartinary Inequality Number of globally ranked Total Labor Force (%  
Global\_1

Cluster 2

Unemployment Total Unemployment, youth total

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2 4	2	2
2	13	100,000	0,000000	1 3	1	2
3	12	96,328	0,073435	13 15	13	2
4	11	94,907	0,101850	1 9	1	3
5	10	93,290	0,134202	1 2	1	5
6	9	92,654	0,146910	1 8	1	6
7	8	91,821	0,163576	1 11	1	7
8	7	91,184	0,176312	1 7	1	8
9	6	91,118	0,177634	1 5	1	9
10	5	90,610	0,187802	1 10	1	10
11	4	85,045	0,299102	1 12	1	11
12	3	80,751	0,384975	1 6	1	12
13	2	66,733	0,665347	1 14	1	13
14	1	65,965	0,680705	1 13	1	15

Final Partition

Cluster 1

Access to Information and Commu Access to Advanced Education Access to  
Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate  
Lower secondary Upper secondary Internet users Years of teartinary  
Inequality Number of globally ranked

Cluster 2

Primary scholl

Cluster 3

Unemployment Total Unemployment, youth total

Cluster 4

Total Labor Force (% Global\_1

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2	4	2	2
2	13	100,000	0,000000	1	3	1	2
3	12	96,328	0,073435	13	15	13	2
4	11	94,907	0,101850	1	9	1	3
5	10	93,290	0,134202	1	2	1	5
6	9	92,654	0,146910	1	8	1	6
7	8	91,821	0,163576	1	11	1	7
8	7	91,184	0,176312	1	7	1	8
9	6	91,118	0,177634	1	5	1	9
10	5	90,610	0,187802	1	10	1	10
11	4	85,045	0,299102	1	12	1	11
12	3	80,751	0,384975	1	6	1	12
13	2	66,733	0,665347	1	14	1	13
14	1	65,965	0,680705	1	13	1	15

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2	4	2	2
2	13	100,000	0,000000	1	3	1	2
3	12	96,328	0,073435	13	15	13	2
4	11	94,907	0,101850	1	9	1	3
5	10	93,290	0,134202	1	2	1	5
6	9	92,654	0,146910	1	8	1	6
7	8	91,821	0,163576	1	11	1	7
8	7	91,184	0,176312	1	7	1	8
9	6	91,118	0,177634	1	5	1	9
10	5	90,610	0,187802	1	10	1	10
11	4	85,045	0,299102	1	12	1	11
12	3	80,751	0,384975	1	6	1	12
13	2	66,733	0,665347	1	14	1	13
14	1	65,965	0,680705	1	13	1	15

Final Partition

Cluster 1  
 Access to Information and Commu Access to Advanced Education Access to  
 Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate  
 Lower secondary Upper secondary Internet users Years of teartary  
 Inequality Number of globally ranked

Cluster 2  
 Primary scholl

Cluster 3  
 Unemployment Total Unemployment, youth total

Cluster 4  
 Total Labor Force (% Global\_1

## Dendrogram

### Cluster Analysis of Variables: Access to In; Access to Ad; Access to In; ...

Correlation Coefficient Distance, Single Linkage  
 Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	14	100,000	0,000000	2 4	2	2
2	13	100,000	0,000000	1 3	1	2
3	12	96,328	0,073435	13 15	13	2
4	11	94,907	0,101850	1 9	1	3
5	10	93,290	0,134202	1 2	1	5
6	9	92,654	0,146910	1 8	1	6
7	8	91,821	0,163576	1 11	1	7
8	7	91,184	0,176312	1 7	1	8
9	6	91,118	0,177634	1 5	1	9
10	5	90,610	0,187802	1 10	1	10
11	4	85,045	0,299102	1 12	1	11
12	3	80,751	0,384975	1 6	1	12
13	2	66,733	0,665347	1 14	1	13
14	1	65,965	0,680705	1 13	1	15

Final Partition

Cluster 1  
 Access to Information and Commu Access to Advanced Education Access to  
 Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate  
 Primary scholl Lower secondary Upper secondary Internet users Years of  
 teartary Inequality Number of globally ranked Total Labor Force (%  
 Global\_1

Cluster 2  
 Unemployment Total Unemployment, youth total

## Dendrogram

### Cluster Analysis of Observations: Social Progr; IDH - 2013; 1. Governanç; ...

Manhattan Distance, Complete Linkage  
 Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,3416	48,54	40	114	40	2
2	130	93,3098	69,71	36	97	36	2
3	129	93,2910	69,90	68	108	68	2
4	128	92,9154	73,82	50	99	50	2
5	127	91,9796	83,57	21	123	21	2
6	126	91,8714	84,69	7	43	7	2
7	125	91,7476	85,98	20	67	20	2
8	124	91,2695	90,97	14	73	14	2
9	123	91,2494	91,17	74	117	74	2
10	122	91,0133	93,64	33	93	33	2
11	121	90,8761	95,06	11	56	11	2
12	120	90,6539	97,38	38	90	38	2
13	119	90,6148	97,79	8	10	8	2
14	118	90,5719	98,23	46	49	46	2
15	117	90,3032	101,03	50	68	50	4
16	116	90,1411	102,72	35	61	35	2
17	115	90,0444	103,73	39	51	39	2
18	114	90,0269	103,91	36	80	36	3
19	113	89,7995	106,28	75	85	75	2
20	112	89,5916	108,45	120	128	120	2
21	111	89,4673	109,74	9	119	9	2
22	110	89,4234	110,20	7	115	7	3
23	109	89,3325	111,15	27	122	27	2
24	108	89,2249	112,27	6	22	6	2
25	107	88,9248	115,40	42	83	42	2
26	106	88,8906	115,75	40	88	40	3
27	105	88,6831	117,91	5	17	5	2
28	104	88,4970	119,85	23	24	23	2
29	103	88,4790	120,04	66	116	66	2
30	102	88,4260	120,59	50	72	50	5
31	101	88,4194	120,66	3	21	3	3
32	100	88,3164	121,73	32	107	32	2
33	99	88,2211	122,73	74	103	74	3
34	98	88,2039	122,91	16	27	16	3
35	97	88,1848	123,11	30	106	30	2
36	96	88,0727	124,27	81	101	81	2
37	95	87,8548	126,54	11	41	11	3
38	94	87,7218	127,93	60	64	60	2
39	93	87,6218	128,97	13	96	13	2
40	92	87,6195	129,00	33	40	33	5
41	91	87,5809	129,40	31	111	31	2
42	90	87,4826	130,42	18	91	18	2
43	89	87,3722	131,57	28	63	28	2
44	88	87,3424	131,88	35	121	35	3
45	87	87,2894	132,43	34	71	34	2
46	86	87,2699	132,64	70	113	70	2
47	85	87,1348	134,05	29	95	29	2
48	84	87,1148	134,25	15	86	15	2
49	83	86,9258	136,22	20	74	20	5
50	82	86,7949	137,59	12	47	12	2
51	81	86,5029	140,63	1	59	1	2
52	80	86,4808	140,86	4	16	4	4
53	79	86,2541	143,22	65	125	65	2
54	78	86,2373	143,40	92	105	92	2
55	77	86,1446	144,36	5	42	5	4
56	76	86,0163	145,70	8	81	8	4
57	75	85,9991	145,88	69	130	69	2

58	74	85,9112	146,79	57	60	57	3
59	73	85,4211	151,90	9	87	9	3
60	72	85,4023	152,10	25	29	25	3
61	71	85,3734	152,40	2	132	2	2
62	70	85,2273	153,92	38	46	38	4
63	69	84,8643	157,70	66	129	66	3
64	68	84,6912	159,51	11	126	11	4
65	67	84,5929	160,53	58	100	58	2
66	66	84,5787	160,68	79	120	79	3
67	65	84,5772	160,69	36	118	36	4
68	64	84,4847	161,66	14	30	14	4
69	63	84,3845	162,70	34	94	34	3
70	62	83,8780	167,98	7	33	7	8
71	61	83,3880	173,08	6	89	6	3
72	60	83,3644	173,33	37	84	37	2
73	59	83,3621	173,35	31	36	31	6
74	58	83,3286	173,70	98	124	98	2
75	57	83,1404	175,66	13	38	13	6
76	56	83,1339	175,73	45	110	45	2
77	55	83,0778	176,32	4	104	4	5
78	54	82,9843	177,29	112	131	112	2
79	53	82,8446	178,75	8	62	8	5
80	52	82,7633	179,59	18	77	18	3
81	51	82,7011	180,24	1	35	1	5
82	50	82,6972	180,28	19	75	19	3
83	49	82,4240	183,13	32	50	32	7
84	48	82,1248	186,25	48	53	48	2
85	47	82,0777	186,74	3	44	3	4
86	46	81,2802	195,05	98	102	98	3
87	45	80,8688	199,33	8	82	8	6
88	44	80,8025	200,02	65	79	65	5
89	43	80,6455	201,66	3	28	3	6
90	42	80,3902	204,32	4	76	4	6
91	41	79,9455	208,95	54	69	54	3
92	40	79,8789	209,65	23	92	23	4
93	39	79,8123	210,34	55	70	55	3
94	38	79,5950	212,60	9	20	9	8
95	37	79,3686	214,96	6	7	6	11
96	36	78,8471	220,40	12	18	12	5
97	35	78,5697	223,29	1	5	1	9
98	34	78,4539	224,49	6	11	6	15
99	33	78,3176	225,91	45	58	45	4
100	32	78,0710	228,48	6	39	6	17
101	31	78,0330	228,88	2	48	2	4
102	30	77,3897	235,58	25	65	25	8
103	29	77,2957	236,56	57	127	57	4
104	28	77,0328	239,30	12	34	12	8
105	27	76,7027	242,74	31	54	31	9
106	26	75,7148	253,03	14	109	14	5
107	25	75,7010	253,18	3	9	3	14
108	24	75,1275	259,15	55	112	55	5
109	23	73,9606	271,31	8	66	8	9
110	22	73,8000	272,98	23	78	23	5
111	21	73,5456	275,63	26	52	26	2
112	20	73,4171	276,97	2	15	2	6
113	19	71,2737	299,31	2	13	2	12
114	18	70,9780	302,39	4	25	4	14
115	17	70,5098	307,27	8	31	8	18
116	16	69,8481	314,16	1	4	1	23
117	15	69,3549	319,30	3	19	3	17
118	14	68,9964	323,03	6	57	6	21
119	13	67,8150	335,34	3	55	3	22
120	12	67,6399	337,17	2	37	2	14



121	11	67,4416	339,23	12	23	12	13
122	10	66,2354	351,80	32	45	32	11
123	9	65,8374	355,95	1	98	1	26
124	8	61,8178	397,83	1	8	1	44
125	7	59,2693	424,38	14	32	14	16
126	6	57,5580	442,21	2	3	2	36
127	5	49,4312	526,89	2	12	2	49
128	4	48,9549	531,85	1	26	1	46
129	3	40,9923	614,82	6	14	6	37
130	2	30,8829	720,15	1	6	1	83
131	1	0,0000	1041,92	1	2	1	132

Final Partition

Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	46	123287	49,8505	105,379
Cluster2	49	183949	58,9466	100,922
Cluster3	21	35971	40,1548	66,468
Cluster4	16	46857	51,8979	84,805

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Social Progress Index	65,2963	49,7512	84,3567	74,4869
IDH - 2013	0,7429	0,5317	0,8635	0,8042
1. Governana	48,5352	37,6973	90,2136	67,0128
1.5 Knowledge and Technology	27,7370	20,7612	48,7762	35,3875
1.6 Creative Outputs	31,4326	23,9347	51,4095	37,5375
EPI Score	53,6489	36,7971	75,0100	67,6106
Happy Planet Index	46,1809	39,8753	44,5040	38,7123
GINNI Index	39,1041	42,7045	32,9990	34,7300
Access to Information and Com_1	61,1035	43,8039	89,7714	76,1981
Access to Advanced Education_1	44,6202	18,5810	69,7848	51,4081
Adult literacy rate	92,2443	54,2514	98,5845	97,5967
Primary scholl	88,9267	74,0729	95,4717	89,8802
Lower secondary	59,5805	32,3666	65,3765	63,0766
Upper secondary	45,7756	18,4434	62,3523	54,4533
Internet users	45,5774	12,8365	88,8802	66,7889
Years of teartary	27,0588	9,5558	55,3501	27,2426
Inequality	71,9767	34,1174	89,9951	87,3196
Number of globally ranked	12,6087	0,8163	63,8095	22,5000
Unemployment Total	24,7557	25,3754	21,0020	55,9446
Total Labor Force (% Global_1	1,2367	0,4046	0,6218	0,2270
Unemployment, youth total	30,1580	23,4130	23,5303	64,9183

Variable	Grand centroid
Social Progress Index	63,6721
IDH - 2013	0,6911
1. Governana	53,3824
1.5 Knowledge and Technology	29,4220
1.6 Creative Outputs	32,5674
EPI Score	52,4840
Happy Planet Index	42,6681
GINNI Index	38,9392

Access to Information and Com_1	61,0721
Access to Advanced Education_1	39,7804
Adult literacy rate	79,7983
Primary scholl	84,5696
Lower secondary	50,8242
Upper secondary	39,3186
Internet users	42,8837
Years of teartinary	25,0847
Inequality	62,6492
Number of globally ranked	17,5758
Unemployment Total	28,1690
Total Labor Force (% Global_1	0,7076
Unemployment, youth total	30,8132

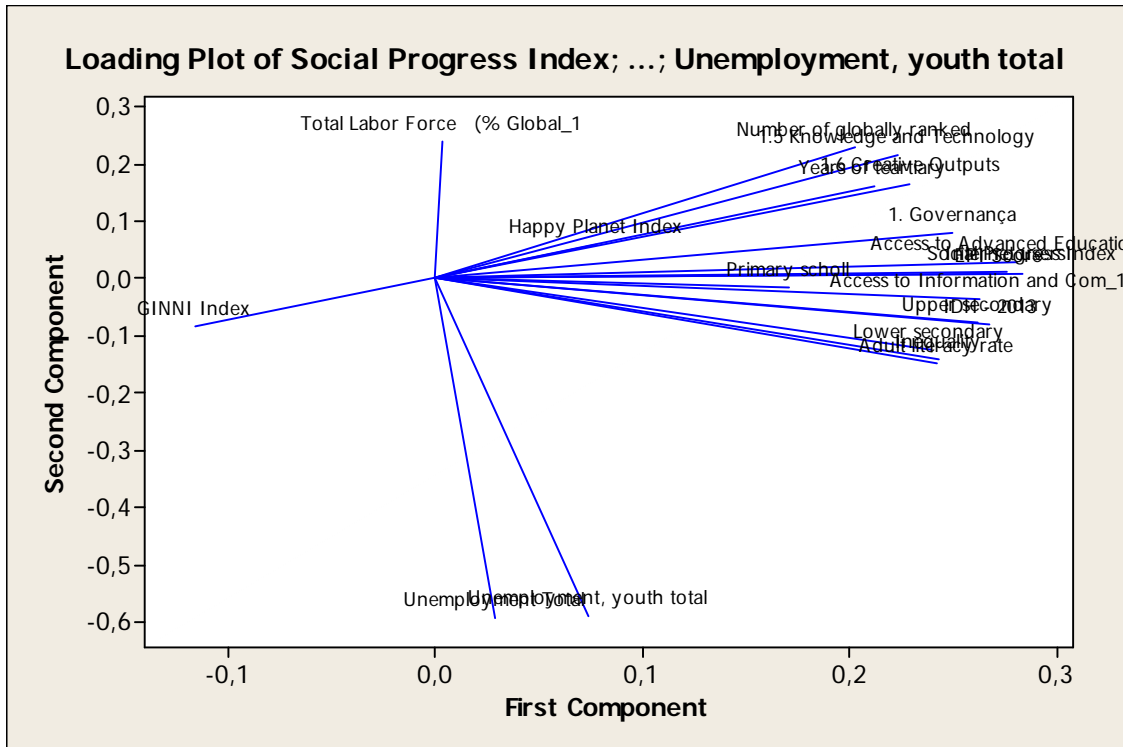
#### Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,000	89,034	104,715	64,403
Cluster2	89,034	0,000	179,545	136,242
Cluster3	104,715	179,545	0,000	87,443
Cluster4	64,403	136,242	87,443	0,000

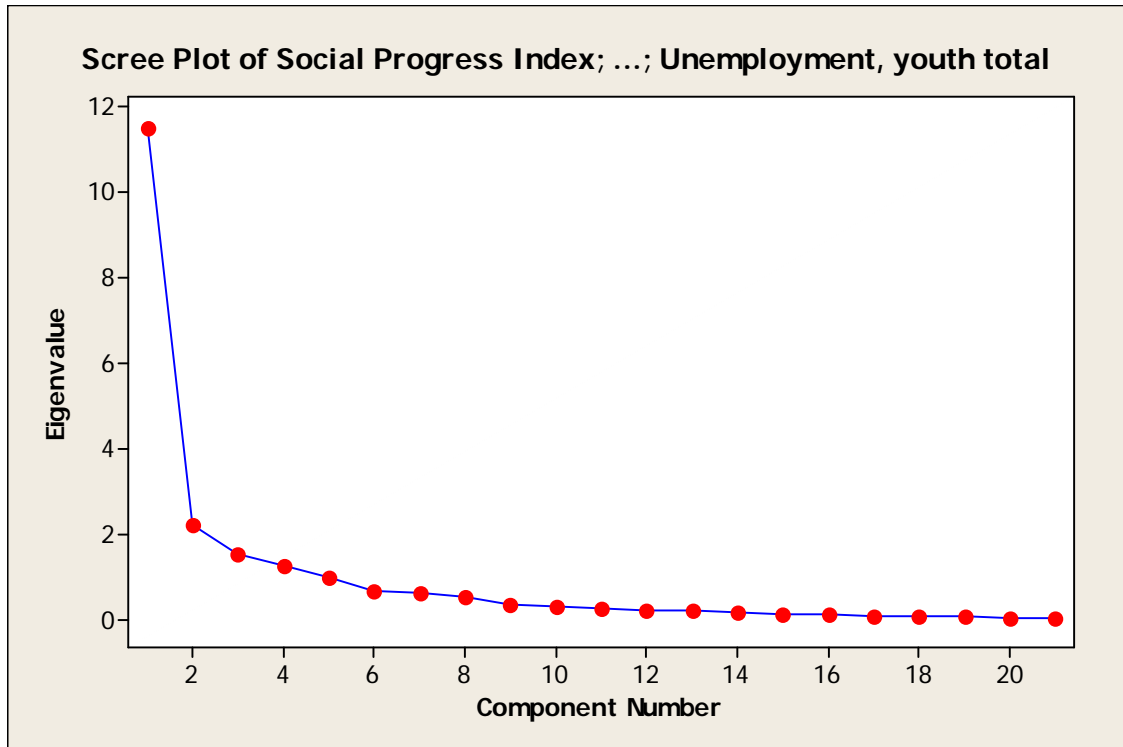
## Dendrogram

### 4.6 Análise do componente principal

Fazendo a análise do componente principal com as variáveis com maior grau de correlação temos o gráfico abaixo, onde podemos notar que todas as variáveis selecionadas apontam para a direita.



No gráfico abaixo vemos a relevância das variáveis, nota se que a primeira variável tem um alto grau de relevância e a segunda e a terceira estão com seus valores acima de um e também são importantes para a análise.



### Principal Component Analysis: Social Progr; IDH - 2013; 1. Governanç; 1.5 Knowl

Eigenanalysis of the Correlation Matrix

Eigenvalue	11,510	2,217	1,519	1,237	0,967	0,671	0,588	0,494	0,346
Proportion	0,548	0,106	0,072	0,059	0,046	0,032	0,028	0,024	0,016
Cumulative	0,548	0,654	0,726	0,785	0,831	0,863	0,891	0,914	0,931

Eigenvalue	0,275	0,241	0,187	0,174	0,143	0,121	0,095	0,075	0,062
Proportion	0,013	0,011	0,009	0,008	0,007	0,006	0,005	0,004	0,003
Cumulative	0,944	0,955	0,964	0,973	0,979	0,985	0,990	0,993	0,996

Eigenvalue	0,040	0,021	0,016
Proportion	0,002	0,001	0,001
Cumulative	0,998	0,999	1,000

Variable	PC1	PC2
Social Progress Index	0,284	0,009
IDH - 2013	0,268	-0,081
1. Governanç	0,250	0,079
1.5 Knowledge and Technology	0,223	0,214
1.6 Creative Outputs	0,229	0,165
EPI Score	0,273	0,006
Happy Planet Index	0,078	0,057
GINNI Index	-0,116	-0,086
Access to Information and Com_1	0,262	-0,036
Access to Advanced Education_1	0,280	0,028
Adult literacy rate	0,242	-0,148
Primary scholl	0,171	-0,018
Lower secondary	0,239	-0,124

Upper secondary	0,262	-0,078
Internet users	0,276	0,012
Years of teartinary	0,212	0,160
Inequality	0,243	-0,143
Number of globally ranked	0,203	0,228
Unemployment Total	0,029	-0,595
Total Labor Force (% Global_1	0,003	0,237
Unemployment, youth total	0,074	-0,591

### **Scree Plot of Social Progress Index; ...; Unemployment, youth total**

### **Loading Plot of Social Progress Index; ...; Unemployment, youth total**

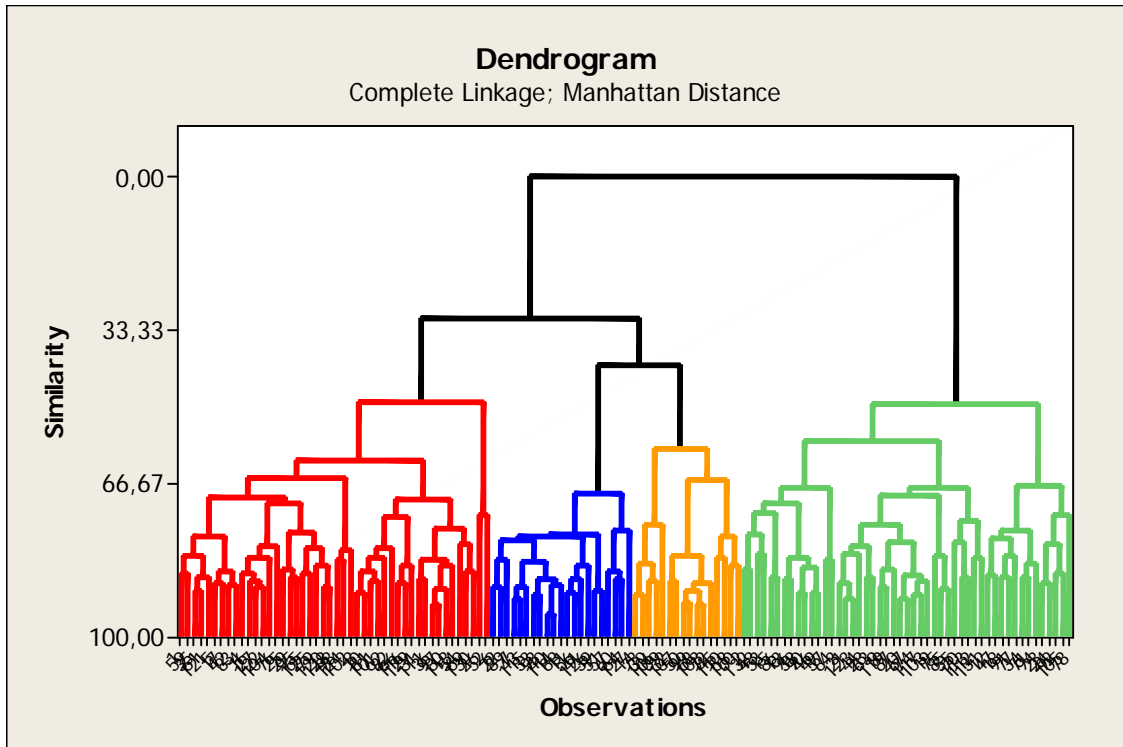
Na análise do componente principal acima vemos que o primeiro índice criado nos dá 54,8% de compreensão das variáveis, se usarmos o primeiro o segundo temos 65,4% de compreensão das variáveis, ao somarmos o terceiro índice chegamos a 72,6%, ou seja, podemos ter um número grande de variáveis, mas ao analisarmos o componente principal podemos criar novos índices e sintetizar a informação para facilitar o trabalho usando os novos índices ao invés de um número grande de variáveis.

## **5. ANÁLISE DE CONGLOMERADOS**

O Dendrograma permite uma análise do grau de similaridade entre os países com relação às variáveis estudadas.

**5.1. Segue** o Dendrograma com os 132 países da base de dados.

Abaixo o Dendrograma com 4 clusters.



**Cluster Analysis of Observations: Social Progr; IDH - 2013; 1. Governanç; ...**

Manhattan Distance, Complete Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	131	95,3416	48,54	40	114	40	2
2	130	93,3098	69,71	36	97	36	2
3	129	93,2910	69,90	68	108	68	2
4	128	92,9154	73,82	50	99	50	2
5	127	91,9796	83,57	21	123	21	2
6	126	91,8714	84,69	7	43	7	2
7	125	91,7476	85,98	20	67	20	2
8	124	91,2695	90,97	14	73	14	2
9	123	91,2494	91,17	74	117	74	2
10	122	91,0133	93,64	33	93	33	2
11	121	90,8761	95,06	11	56	11	2
12	120	90,6539	97,38	38	90	38	2
13	119	90,6148	97,79	8	10	8	2
14	118	90,5719	98,23	46	49	46	2
15	117	90,3032	101,03	50	68	50	4
16	116	90,1411	102,72	35	61	35	2
17	115	90,0444	103,73	39	51	39	2
18	114	90,0269	103,91	36	80	36	3
19	113	89,7995	106,28	75	85	75	2
20	112	89,5916	108,45	120	128	120	2

21	111	89,4673	109,74	9	119	9	2
22	110	89,4234	110,20	7	115	7	3
23	109	89,3325	111,15	27	122	27	2
24	108	89,2249	112,27	6	22	6	2
25	107	88,9248	115,40	42	83	42	2
26	106	88,8906	115,75	40	88	40	3
27	105	88,6831	117,91	5	17	5	2
28	104	88,4970	119,85	23	24	23	2
29	103	88,4790	120,04	66	116	66	2
30	102	88,4260	120,59	50	72	50	5
31	101	88,4194	120,66	3	21	3	3
32	100	88,3164	121,73	32	107	32	2
33	99	88,2211	122,73	74	103	74	3
34	98	88,2039	122,91	16	27	16	3
35	97	88,1848	123,11	30	106	30	2
36	96	88,0727	124,27	81	101	81	2
37	95	87,8548	126,54	11	41	11	3
38	94	87,7218	127,93	60	64	60	2
39	93	87,6218	128,97	13	96	13	2
40	92	87,6195	129,00	33	40	33	5
41	91	87,5809	129,40	31	111	31	2
42	90	87,4826	130,42	18	91	18	2
43	89	87,3722	131,57	28	63	28	2
44	88	87,3424	131,88	35	121	35	3
45	87	87,2894	132,43	34	71	34	2
46	86	87,2699	132,64	70	113	70	2
47	85	87,1348	134,05	29	95	29	2
48	84	87,1148	134,25	15	86	15	2
49	83	86,9258	136,22	20	74	20	5
50	82	86,7949	137,59	12	47	12	2
51	81	86,5029	140,63	1	59	1	2
52	80	86,4808	140,86	4	16	4	4
53	79	86,2541	143,22	65	125	65	2
54	78	86,2373	143,40	92	105	92	2
55	77	86,1446	144,36	5	42	5	4
56	76	86,0163	145,70	8	81	8	4
57	75	85,9991	145,88	69	130	69	2
58	74	85,9112	146,79	57	60	57	3
59	73	85,4211	151,90	9	87	9	3
60	72	85,4023	152,10	25	29	25	3
61	71	85,3734	152,40	2	132	2	2
62	70	85,2273	153,92	38	46	38	4
63	69	84,8643	157,70	66	129	66	3
64	68	84,6912	159,51	11	126	11	4
65	67	84,5929	160,53	58	100	58	2
66	66	84,5787	160,68	79	120	79	3
67	65	84,5772	160,69	36	118	36	4
68	64	84,4847	161,66	14	30	14	4
69	63	84,3845	162,70	34	94	34	3
70	62	83,8780	167,98	7	33	7	8
71	61	83,3880	173,08	6	89	6	3
72	60	83,3644	173,33	37	84	37	2
73	59	83,3621	173,35	31	36	31	6
74	58	83,3286	173,70	98	124	98	2
75	57	83,1404	175,66	13	38	13	6
76	56	83,1339	175,73	45	110	45	2
77	55	83,0778	176,32	4	104	4	5
78	54	82,9843	177,29	112	131	112	2
79	53	82,8446	178,75	8	62	8	5
80	52	82,7633	179,59	18	77	18	3
81	51	82,7011	180,24	1	35	1	5
82	50	82,6972	180,28	19	75	19	3
83	49	82,4240	183,13	32	50	32	7

84	48	82,1248	186,25	48	53	48	2
85	47	82,0777	186,74	3	44	3	4
86	46	81,2802	195,05	98	102	98	3
87	45	80,8688	199,33	8	82	8	6
88	44	80,8025	200,02	65	79	65	5
89	43	80,6455	201,66	3	28	3	6
90	42	80,3902	204,32	4	76	4	6
91	41	79,9455	208,95	54	69	54	3
92	40	79,8789	209,65	23	92	23	4
93	39	79,8123	210,34	55	70	55	3
94	38	79,5950	212,60	9	20	9	8
95	37	79,3686	214,96	6	7	6	11
96	36	78,8471	220,40	12	18	12	5
97	35	78,5697	223,29	1	5	1	9
98	34	78,4539	224,49	6	11	6	15
99	33	78,3176	225,91	45	58	45	4
100	32	78,0710	228,48	6	39	6	17
101	31	78,0330	228,88	2	48	2	4
102	30	77,3897	235,58	25	65	25	8
103	29	77,2957	236,56	57	127	57	4
104	28	77,0328	239,30	12	34	12	8
105	27	76,7027	242,74	31	54	31	9
106	26	75,7148	253,03	14	109	14	5
107	25	75,7010	253,18	3	9	3	14
108	24	75,1275	259,15	55	112	55	5
109	23	73,9606	271,31	8	66	8	9
110	22	73,8000	272,98	23	78	23	5
111	21	73,5456	275,63	26	52	26	2
112	20	73,4171	276,97	2	15	2	6
113	19	71,2737	299,31	2	13	2	12
114	18	70,9780	302,39	4	25	4	14
115	17	70,5098	307,27	8	31	8	18
116	16	69,8481	314,16	1	4	1	23
117	15	69,3549	319,30	3	19	3	17
118	14	68,9964	323,03	6	57	6	21
119	13	67,8150	335,34	3	55	3	22
120	12	67,6399	337,17	2	37	2	14
121	11	67,4416	339,23	12	23	12	13
122	10	66,2354	351,80	32	45	32	11
123	9	65,8374	355,95	1	98	1	26
124	8	61,8178	397,83	1	8	1	44
125	7	59,2693	424,38	14	32	14	16
126	6	57,5580	442,21	2	3	2	36
127	5	49,4312	526,89	2	12	2	49
128	4	48,9549	531,85	1	26	1	46
129	3	40,9923	614,82	6	14	6	37
130	2	30,8829	720,15	1	6	1	83
131	1	0,0000	1041,92	1	2	1	132

Final Partition

Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	46	123287	49,8505	105,379
Cluster2	49	183949	58,9466	100,922
Cluster3	21	35971	40,1548	66,468
Cluster4	16	46857	51,8979	84,805



## Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Social Progress Index	65,2963	49,7512	84,3567	74,4869
IDH - 2013	0,7429	0,5317	0,8635	0,8042
1. Governança	48,5352	37,6973	90,2136	67,0128
1.5 Knowledge and Technology	27,7370	20,7612	48,7762	35,3875
1.6 Creative Outputs	31,4326	23,9347	51,4095	37,5375
EPI Score	53,6489	36,7971	75,0100	67,6106
Happy Planet Index	46,1809	39,8753	44,5040	38,7123
GINNI Index	39,1041	42,7045	32,9990	34,7300
Access to Information and Com_1	61,1035	43,8039	89,7714	76,1981
Access to Advanced Education_1	44,6202	18,5810	69,7848	51,4081
Adult literacy rate	92,2443	54,2514	98,5845	97,5967
Primary scholl	88,9267	74,0729	95,4717	89,8802
Lower secondary	59,5805	32,3666	65,3765	63,0766
Upper secondary	45,7756	18,4434	62,3523	54,4533
Internet users	45,5774	12,8365	88,8802	66,7889
Years of teartary	27,0588	9,5558	55,3501	27,2426
Inequality	71,9767	34,1174	89,9951	87,3196
Number of globally ranked	12,6087	0,8163	63,8095	22,5000
Unemployment Total	24,7557	25,3754	21,0020	55,9446
Total Labor Force (% Global_1	1,2367	0,4046	0,6218	0,2270
Unemployment, youth total	30,1580	23,4130	23,5303	64,9183

Variable	Grand centroid
Social Progress Index	63,6721
IDH - 2013	0,6911
1. Governança	53,3824
1.5 Knowledge and Technology	29,4220
1.6 Creative Outputs	32,5674
EPI Score	52,4840
Happy Planet Index	42,6681
GINNI Index	38,9392
Access to Information and Com_1	61,0721
Access to Advanced Education_1	39,7804
Adult literacy rate	79,7983
Primary scholl	84,5696
Lower secondary	50,8242
Upper secondary	39,3186
Internet users	42,8837
Years of teartary	25,0847
Inequality	62,6492
Number of globally ranked	17,5758
Unemployment Total	28,1690
Total Labor Force (% Global_1	0,7076
Unemployment, youth total	30,8132

## Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,000	89,034	104,715	64,403
Cluster2	89,034	0,000	179,545	136,242
Cluster3	104,715	179,545	0,000	87,443
Cluster4	64,403	136,242	87,443	0,000

## Dendrogram

## 6. COMPARAÇÃO DE MÉDIA, ANÁLISE DE VARIÂNCIA E INTERVALO DE CONFIANÇA.

Para o presente trabalho faremos uma comparação para análise entre os países dos grupos AIBER, AVECO e OTHERS.

Abaixo a tabela com os países os grupos e as regiões.

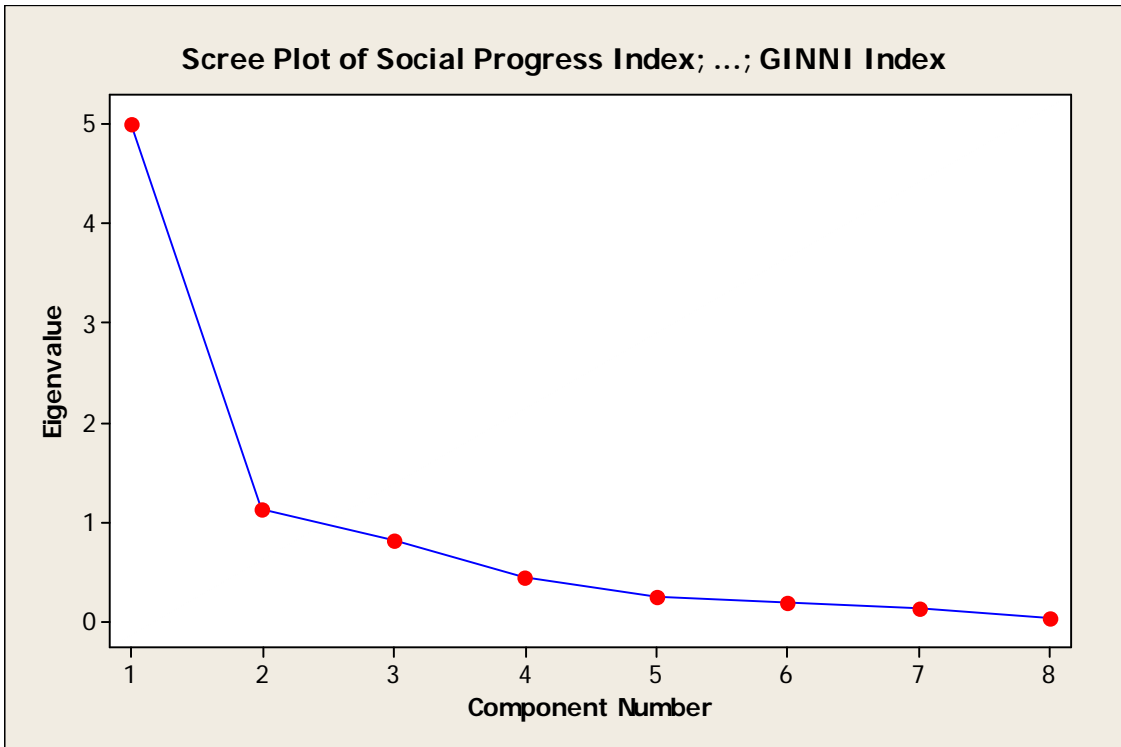
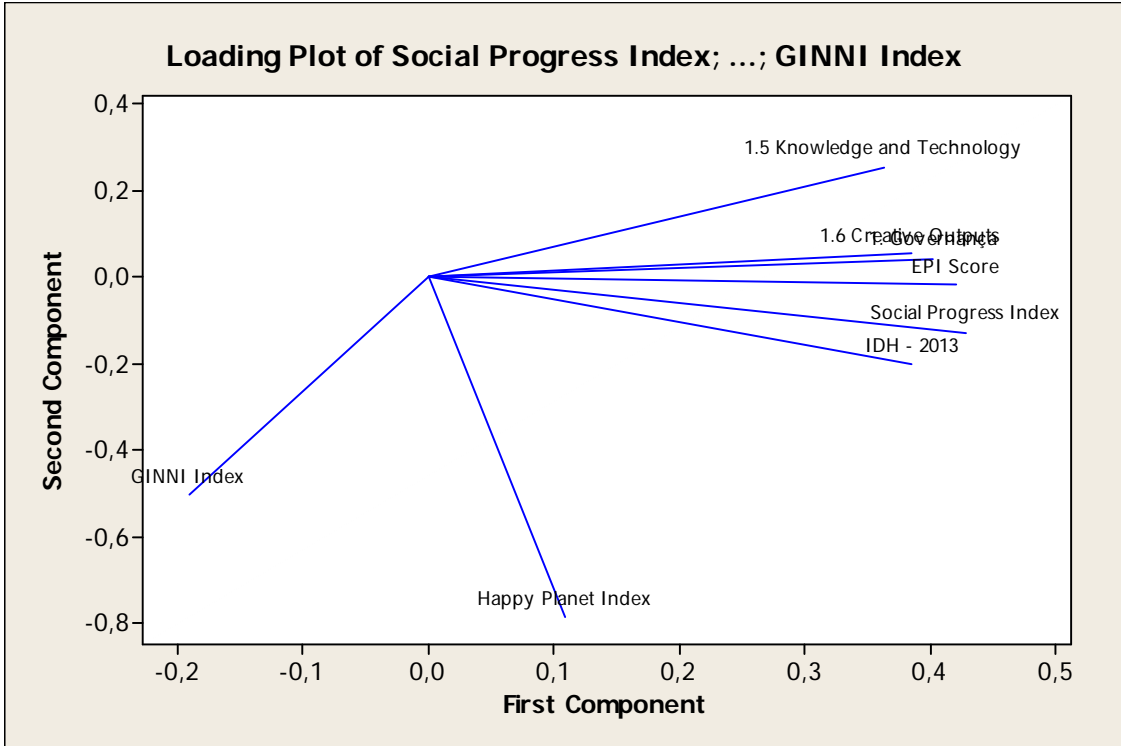
<b>País</b>	<b>Grupos</b>	<b>Região</b>
Albania	OTHERS	Southern Europe
Algeria	OTHERS	Northern Africa
Angola	OTHERS	Middle Africa
Argentina	AIBER	IBE
Armenia	OTHERS	Western Asia
Australia	AVECO	Australia and New Zealand
Austria	AVECO	Western Europe
Azerbaijan	OTHERS	Western Asia
Bangladesh	OTHERS	Southern Asia
Belarus	OTHERS	Eastern Europe
Belgium	AVECO	Western Europe
Benin	OTHERS	Western Africa
Bolivia	AIBER	IBE
Bosnia and Herzegovina	OTHERS	Southern Europe
Botswana	OTHERS	Southern Africa
Brazil	AIBER	IBE
Bulgaria	OTHERS	Eastern Europe
Burkina Faso	OTHERS	Western Africa
Burundi	OTHERS	Eastern Africa
Cambodia	OTHERS	South-Eastern Asia

Cameroon	OTHERS	Middle Africa
Canada	AVECO	Northern America
Central African Republic	OTHERS	Eastern Africa
Chad	OTHERS	Middle Africa
Chile	AIBER	IBE
China	OTHERS	Eastern Asia
Colombia	AIBER	IBE
Congo, Republic of	OTHERS	Middle Africa
Costa Rica	AIBER	IBE
Croatia	OTHERS	Southern Europe
Cuba	AIBER	Caribbean
Czech Republic	AVECO	Eastern Europe
Denmark	AVECO	Northern Europe
Djibouti	OTHERS	Eastern Africa
Dominican Republic	AIBER	IBE
Ecuador	AIBER	IBE
Egypt	OTHERS	Northern Africa
El Salvador	AIBER	IBE
Estonia	AVECO	Northern Europe
Finland	AVECO	Northern Europe
France	AVECO	Western Europe
Georgia	OTHERS	Western Asia
Germany	AVECO	Western Europe
Ghana	OTHERS	Western Africa
Greece	AVECO	Southern Europe
Guatemala	AIBER	IBE
Guinea	OTHERS	Western Africa
Guyana	OTHERS	South America

Honduras	AIBER	IBE
Hungary	OTHERS	Eastern Europe
Iceland	AVECO	Northern Europe
India	OTHERS	Southern Asia
Indonesia	OTHERS	South-Eastern Asia
Iran	OTHERS	Southern Asia
Iraq	OTHERS	Western Asia
Ireland	AVECO	Northern Europe
Israel	AVECO	Western Asia
Italy	AVECO	Southern Europe
Jamaica	OTHERS	Caribbean
Japan	AVECO	Eastern Asia
Jordan	OTHERS	Western Asia
Kazakhstan	OTHERS	Central Asia
Kenya	OTHERS	Eastern Africa
Korea, Republic of	AVECO	Eastern Asia
Kuwait	OTHERS	Western Asia
Kyrgyzstan	OTHERS	Central Asia
Laos	OTHERS	South-Eastern Asia
Latvia	AVECO	Northern Europe
Lebanon	OTHERS	Western Asia
Lesotho	OTHERS	Southern Africa
Liberia	OTHERS	Western Africa
Lithuania	OTHERS	Northern Europe
Macedonia	OTHERS	Southern Europe
Madagascar	OTHERS	Eastern Africa
Malawi	OTHERS	Southern Africa
Malaysia	OTHERS	South-Eastern Asia

Mali	OTHERS	Western Africa
Mauritania	OTHERS	Western Africa
Mauritius	OTHERS	Eastern Africa
Mexico	AIBER	IBE
Moldova	OTHERS	Eastern Europe
Mongolia	OTHERS	Eastern Asia
Montenegro	OTHERS	Southern Europe
Morocco	OTHERS	Northern Africa
Mozambique	OTHERS	Eastern Africa
Namibia	OTHERS	Southern Africa
Nepal	OTHERS	Southern Asia
Netherlands	AVECO	Western Europe
New Zealand	AVECO	Australia and New Zealand
Nicaragua	AIBER	IBE
Niger	OTHERS	Western Africa
Nigeria	OTHERS	Western Africa
Norway	AVECO	Northern Europe
Pakistan	OTHERS	Southern Asia
Panama	AIBER	IBE
Paraguay	AIBER	IBE
Peru	AIBER	IBE
Philippines	OTHERS	South-Eastern Asia
Poland	OTHERS	Eastern Europe
Portugal	AIBER	IBE
Romania	OTHERS	Eastern Europe
Russia	OTHERS	Eastern Europe
Rwanda	OTHERS	Eastern Africa
Saudi Arabia	OTHERS	Western Asia

Senegal	OTHERS	Western Africa
Serbia	OTHERS	Southern Europe
Slovakia	AVECO	Eastern Europe
Slovenia	AVECO	Southern Europe
South Africa	OTHERS	Southern Africa
Spain	AIBER	IBE
Sri Lanka	OTHERS	Southern Asia
Sudan	OTHERS	Northern Africa
Swaziland	OTHERS	Southern Africa
Sweden	AVECO	Northern Europe
Switzerland	AVECO	Western Europe
Tajikistan	OTHERS	Central Asia
Tanzania	OTHERS	Eastern Africa
Thailand	OTHERS	South-Eastern Asia
Togo	OTHERS	Western Africa
Trinidad and Tobago	OTHERS	Caribbean
Tunisia	OTHERS	Northern Africa
Turkey	OTHERS	Western Asia
Uganda	OTHERS	Eastern Africa
Ukraine	OTHERS	Eastern Europe
United Arab Emirates	OTHERS	Western Asia
United Kingdom	AVECO	Northern Europe
United States	AVECO	Northern America
Uruguay	AIBER	IBE
Uzbekistan	OTHERS	Central Asia
Venezuela	AIBER	IBE
Yemen	OTHERS	Western Asia
Zambia	OTHERS	Middle Africa



## Cluster Analysis of Variables: Social Progr; IDH - 2013; 1. Governanç; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	20	96,3283	0,073435	19 21	19	2
2	19	95,8701	0,082598	1 6	1	2
3	18	95,8168	0,083664	1 9	1	3
4	17	95,4440	0,091120	1 15	1	4
5	16	94,4907	0,110186	1 3	1	5
6	15	93,9615	0,120770	1 10	1	6
7	14	93,5636	0,128728	1 2	1	7
8	13	92,6545	0,146910	1 14	1	8
9	12	91,8212	0,163576	1 17	1	9
10	11	91,1957	0,176087	1 11	1	10
11	10	91,1844	0,176312	1 13	1	11
12	9	90,6099	0,187802	1 16	1	12
13	8	89,4345	0,211311	1 5	1	13
14	7	87,9933	0,240134	4 18	4	2
15	6	87,0270	0,259460	1 4	1	15
16	5	80,7512	0,384975	1 12	1	16
17	4	68,9252	0,621497	1 7	1	17
18	3	66,7326	0,665347	1 20	1	18
19	2	65,9648	0,680705	1 19	1	20
20	1	52,9445	0,941111	1 8	1	21

Final Partition

Cluster 1

Social Progress Index IDH - 2013 1. Governança 1.5 Knowledge and Technology 1.6 Creative Outputs EPI Score Happy Planet Index Access to Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate Primary scholl Lower secondary Upper secondary Internet users Years of teartertiary Inequality Number of globally ranked

Cluster 2

GINNI Index

Cluster 3

Unemployment Total Unemployment, youth total

Cluster 4

Total Labor Force (% Global\_1

## Dendrogram

## Cluster Analysis of Variables: Social Progr; IDH - 2013; 1. Governanç; ...

Correlation Coefficient Distance, Single Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	20	96,3283	0,073435	19 21	19	2



2	19	95,8701	0,082598	1	6	1	2
3	18	95,8168	0,083664	1	9	1	3
4	17	95,4440	0,091120	1	15	1	4
5	16	94,4907	0,110186	1	3	1	5
6	15	93,9615	0,120770	1	10	1	6
7	14	93,5636	0,128728	1	2	1	7
8	13	92,6545	0,146910	1	14	1	8
9	12	91,8212	0,163576	1	17	1	9
10	11	91,1957	0,176087	1	11	1	10
11	10	91,1844	0,176312	1	13	1	11
12	9	90,6099	0,187802	1	16	1	12
13	8	89,4345	0,211311	1	5	1	13
14	7	87,9933	0,240134	4	18	4	2
15	6	87,0270	0,259460	1	4	1	15
16	5	80,7512	0,384975	1	12	1	16
17	4	68,9252	0,621497	1	7	1	17
18	3	66,7326	0,665347	1	20	1	18
19	2	65,9648	0,680705	1	19	1	20
20	1	52,9445	0,941111	1	8	1	21

Final Partition

Cluster 1

Social Progress Index IDH - 2013 1. Governança 1.5 Knowledge and  
Technology 1.6 Creative Outputs EPI Score Happy Planet Index Access to  
Information and Com\_1 Access to Advanced Education\_1 Adult literacy rate  
Primary scholl Lower secondary Upper secondary Internet users Years of  
teartinary Inequality Number of globally ranked Unemployment Total  
Total Labor Force (% Global\_1 Unemployment, youth total

Cluster 2

GINNI Index

## Dendrogram

### 6.1 Abaixo ANOVAS dos grupos com as variáveis.

A análise das variâncias permite a verificação e visualização das médias e desvios padrões da variável a ser analisada.

### Principal Component Analysis: Social Progr; IDH - 2013; 1. Governanç; 1.5 Knowl

Eigenanalysis of the Correlation Matrix

Eigenvalue	4,9941	1,1374	0,8079	0,4486	0,2514	0,1938	0,1284	0,0384
Proportion	0,624	0,142	0,101	0,056	0,031	0,024	0,016	0,005
Cumulative	0,624	0,766	0,867	0,924	0,955	0,979	0,995	1,000

Variable	PC1	PC2	PC3
Social Progress Index	0,429	-0,128	-0,061
IDH - 2013	0,385	-0,203	0,093
1. Governança	0,403	0,044	-0,295
1.5 Knowledge and Technology	0,364	0,254	0,011
1.6 Creative Outputs	0,385	0,055	-0,213
EPI Score	0,421	-0,017	-0,037
Happy Planet Index	0,110	-0,787	0,468

GINNI Index -0,190 -0,503 -0,797

### Scree Plot of Social Progress Index; ...; GINNI Index

### Loading Plot of Social Progress Index; ...; GINNI Index

### Principal Component Analysis: Social Progr; IDH - 2013; 1. Governanç; 1.5 Knowl

Eigenanalysis of the Correlation Matrix

Eigenvalue	4,9941	1,1374	0,8079	0,4486	0,2514	0,1938	0,1284	0,0384
Proportion	0,624	0,142	0,101	0,056	0,031	0,024	0,016	0,005
Cumulative	0,624	0,766	0,867	0,924	0,955	0,979	0,995	1,000

Variable	PC1	PC2	PC3
Social Progress Index	0,429	-0,128	-0,061
IDH - 2013	0,385	-0,203	0,093
1. Governança	0,403	0,044	-0,295
1.5 Knowledge and Technology	0,364	0,254	0,011
1.6 Creative Outputs	0,385	0,055	-0,213
EPI Score	0,421	-0,017	-0,037
Happy Planet Index	0,110	-0,787	0,468
GINNI Index	-0,190	-0,503	-0,797

### Scree Plot of Social Progress Index; ...; GINNI Index

### Loading Plot of Social Progress Index; ...; GINNI Index

### Scatterplot of PC1 vs PC2

### Cluster Analysis of Observations: Access to In; Access to Ad; Adult litera; ...

Manhattan Distance, Complete Linkage  
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster
1	131	96,8845	26,613	68 108	68	2
2	130	96,0092	34,089	40 114	40	2
3	129	96,0058	34,119	36 97	36	2
4	128	95,7394	36,394	7 115	7	2
5	127	95,2243	40,794	9 119	9	2
6	126	95,2227	40,808	50 99	50	2
7	125	94,7592	44,767	14 73	14	2
8	124	94,3369	48,374	8 10	8	2
9	123	93,8490	52,542	74 117	74	2
10	122	93,6283	54,427	51 93	51	2

11	121	93,3064	57,177	20	67	20	2
12	120	93,2930	57,291	68	72	68	3
13	119	93,2088	58,011	27	122	27	2
14	118	92,7036	62,326	21	123	21	2
15	117	92,4994	64,071	36	95	36	3
16	116	92,3879	65,022	120	128	120	2
17	115	92,2195	66,461	38	90	38	2
18	114	92,1189	67,321	20	103	20	3
19	113	92,0554	67,863	24	91	24	2
20	112	92,0357	68,031	7	43	7	3
21	111	91,9187	69,031	66	81	66	2
22	110	91,8937	69,244	11	56	11	2
23	109	91,6593	71,247	29	130	29	2
24	108	91,4423	73,100	74	75	74	3
25	107	91,3286	74,072	65	125	65	2
26	106	91,2347	74,873	32	50	32	3
27	105	91,2263	74,945	8	101	8	3
28	104	91,0891	76,118	42	83	42	2
29	103	91,0707	76,274	44	46	44	2
30	102	91,0582	76,381	2	132	2	2
31	101	91,0338	76,590	35	61	35	2
32	100	90,8861	77,851	33	51	33	3
33	99	90,8718	77,974	18	47	18	2
34	98	90,5142	81,028	5	17	5	2
35	97	90,4029	81,979	15	86	15	2
36	96	90,3923	82,070	57	60	57	2
37	95	90,0057	85,371	23	94	23	2
38	94	89,9956	85,458	22	127	22	2
39	93	89,8691	86,538	82	111	82	2
40	92	89,8033	87,101	9	87	9	3
41	91	89,7241	87,777	30	106	30	2
42	90	89,5281	89,452	38	49	38	3
43	89	89,4844	89,825	3	21	3	3
44	88	89,3130	91,289	41	126	41	2
45	87	89,1615	92,582	25	62	25	2
46	86	89,1034	93,079	8	66	8	5
47	85	89,0620	93,433	11	40	11	4
48	84	88,8302	95,413	13	96	13	2
49	83	88,7046	96,485	57	64	57	3
50	82	88,6915	96,597	32	107	32	4
51	81	88,6061	97,327	23	105	23	3
52	80	88,6021	97,361	16	104	16	2
53	79	88,5438	97,860	70	113	70	2
54	78	88,3421	99,583	48	79	48	2
55	77	88,3384	99,613	5	59	5	3
56	76	88,1437	101,277	80	118	80	2
57	75	88,0431	102,136	4	76	4	2
58	74	87,7890	104,307	18	24	18	4
59	73	87,6301	105,664	19	85	19	2
60	72	87,6002	105,919	28	63	28	2
61	71	87,4445	107,249	27	69	27	3
62	70	87,3518	108,041	35	121	35	3
63	69	87,3203	108,310	33	39	33	4
64	68	87,2594	108,831	1	5	1	4
65	67	87,2347	109,042	29	36	29	5
66	66	87,2071	109,277	55	70	55	3
67	65	87,1895	109,427	82	116	82	3
68	64	86,9560	111,422	1	42	1	6
69	63	86,9453	111,513	3	74	3	6
70	62	86,7104	113,520	9	12	9	4
71	61	86,6658	113,901	98	124	98	2
72	60	86,6277	114,226	14	30	14	4
73	59	86,6135	114,348	6	88	6	2

74	58	86,2483	117,468	58	100	58	2
75	57	85,9696	119,848	32	68	32	7
76	56	85,7648	121,597	34	71	34	2
77	55	85,7486	121,736	8	120	8	7
78	54	85,5155	123,727	57	89	57	4
79	53	85,4103	124,626	38	44	38	5
80	52	85,3459	125,176	45	110	45	2
81	51	85,0969	127,303	18	77	18	5
82	50	84,8910	129,062	16	27	16	5
83	49	84,8177	129,687	31	82	31	4
84	48	84,8121	129,735	53	80	53	3
85	47	84,5993	131,554	37	84	37	2
86	46	84,3823	133,407	112	131	112	2
87	45	84,3671	133,537	7	41	7	5
88	44	84,3077	134,044	28	38	28	7
89	43	83,2648	142,952	4	16	4	7
90	42	83,2220	143,319	15	55	15	5
91	41	83,0227	145,021	34	92	34	3
92	40	82,9235	145,868	7	11	7	9
93	39	82,9190	145,907	9	20	9	7
94	38	82,8654	146,364	3	19	3	8
95	37	82,7277	147,541	35	54	35	4
96	36	82,4402	149,996	8	25	8	9
97	35	81,7436	155,946	13	53	13	5
98	34	81,6863	156,437	29	65	29	7
99	33	81,3530	159,283	18	23	18	8
100	32	80,9192	162,989	48	129	48	3
101	31	80,9073	163,090	98	102	98	3
102	30	80,4121	167,321	6	7	6	11
103	29	78,9848	179,513	1	14	1	10
104	28	78,6035	182,770	6	33	6	15
105	27	78,3420	185,003	13	31	13	9
106	26	77,7012	190,477	8	29	8	16
107	25	77,6292	191,092	2	35	2	6
108	24	77,2157	194,624	22	57	22	6
109	23	77,2119	194,656	18	34	18	11
110	22	76,8694	197,582	32	58	32	9
111	21	75,9579	205,369	26	52	26	2
112	20	75,5864	208,542	4	8	4	23
113	19	74,3008	219,524	15	112	15	7
114	18	73,3832	227,362	3	9	3	15
115	17	73,2952	228,114	13	48	13	12
116	16	72,9850	230,763	28	37	28	9
117	15	71,9073	239,969	6	22	6	21
118	14	70,2294	254,301	4	98	4	26
119	13	70,0912	255,482	1	32	1	19
120	12	69,0818	264,104	2	13	2	18
121	11	69,0219	264,616	45	109	45	3
122	10	67,2411	279,828	15	78	15	8
123	9	65,8469	291,737	3	18	3	26
124	8	62,1380	323,418	2	28	2	27
125	7	57,2534	365,143	1	4	1	45
126	6	57,2308	365,336	2	26	2	29
127	5	52,3493	407,034	3	15	3	34
128	4	47,6805	446,915	1	45	1	48
129	3	45,9216	461,940	1	6	1	69
130	2	33,5000	568,045	2	3	2	63
131	1	0,0000	854,203	1	2	1	132

Final Partition  
Number of clusters: 4

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	48	125422	47,9197	110,605
Cluster2	29	68838	46,8437	81,807
Cluster3	34	102056	52,0397	94,879
Cluster4	21	30454	36,6582	64,062

#### Cluster Centroids

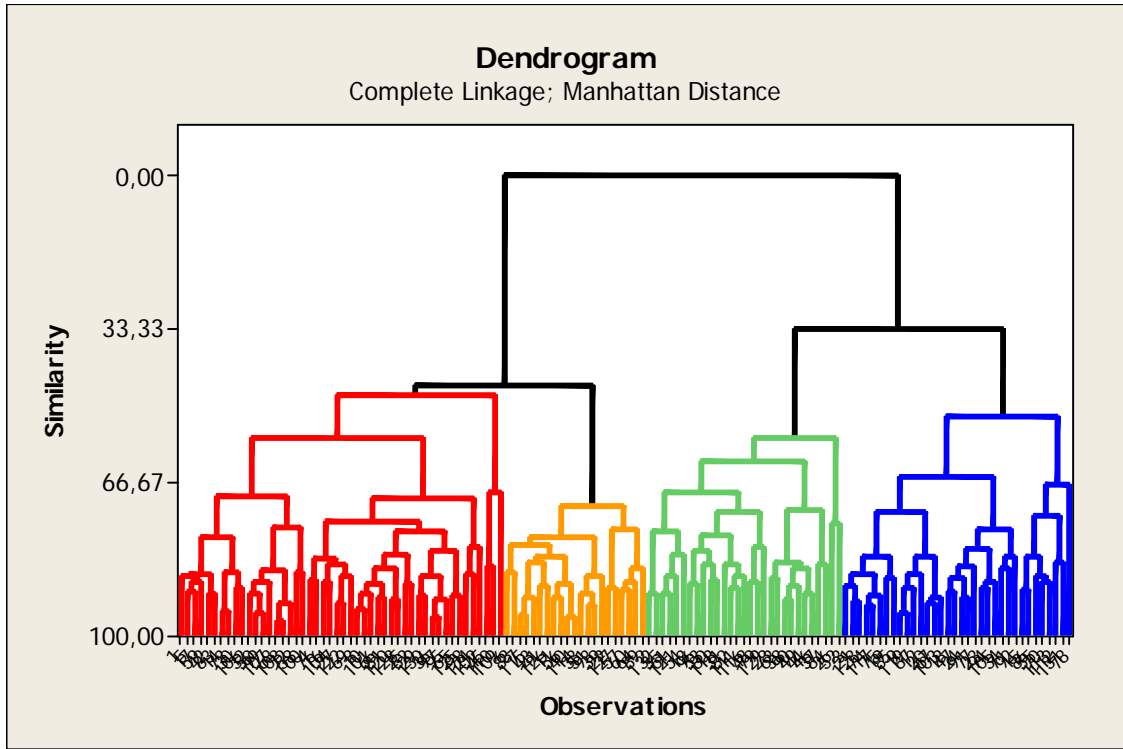
Variable	Cluster1	Cluster2	Cluster3	Cluster4
Access to Information and Com_1	69,9840	52,0166	38,4885	89,7714
Access to Advanced Education_1	48,3910	32,7166	15,1171	69,7848
Adult literacy rate	95,3526	79,6972	46,3223	98,5845
Primary scholl	88,0689	87,6750	70,2470	95,4717
Lower secondary	60,7372	56,4006	23,0850	65,3765
Upper secondary	48,9515	37,5190	13,0277	62,3523
Internet users	56,9730	27,4596	7,7393	88,8802
Years of teartinary	28,8848	16,4300	8,4083	55,3501
Inequality	79,2741	53,5300	30,0667	89,9951
Number of globally ranked	15,0000	8,9655	0,0000	63,8095
Unemployment Total	35,2063	23,9919	26,2236	21,0020
Total Labor Force (% Global_1	0,3572	1,7631	0,3549	0,6218
Unemployment, youth total	42,1308	27,3378	22,2978	23,5303
PC1	0,7127	-0,8066	-2,5702	3,6461
PC2	-0,1496	-0,6224	0,5284	0,3459
PC3	0,0863	0,2170	-0,2045	-0,1659

Variable	Grand centroid
Access to Information and Com_1	61,0721
Access to Advanced Education_1	39,7804
Adult literacy rate	79,7983
Primary scholl	84,5696
Lower secondary	50,8242
Upper secondary	39,3186
Internet users	42,8837
Years of teartinary	25,0847
Inequality	62,6492
Number of globally ranked	17,5758
Unemployment Total	28,1690
Total Labor Force (% Global_1	0,7076
Unemployment, youth total	30,8132
PC1	0,0000
PC2	0,0000
PC3	-0,0000

#### Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,000	55,116	116,206	76,734
Cluster2	55,116	0,000	68,857	116,554
Cluster3	116,206	68,857	0,000	172,411
Cluster4	76,734	116,554	172,411	0,000

#### Dendrogram



### One-way ANOVA: Social Progress Index versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	21869,9	7290,0	205,19	0,000
Error	128	4547,5	35,5		
Total	131	26417,5			

S = 5,961    R-Sq = 82,79%    R-Sq(adj) = 82,38%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev	CI Lower	CI Upper
1	48	69,658	6,272	61,114	78,202
2	29	59,560	5,396	52,768	66,352
3	34	45,953	6,851	38,251	53,655
4	21	84,357	4,089	79,660	89,054

-----+-----+-----+-----+-----  
 (\*)  
 (-\*)  
 (\*-)  
 (-\*-)  
 -----+-----+-----+-----+-----  
 48            60            72            84

Pooled StDev = 5,961

### One-way ANOVA: IDH - 2013 versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	2,27525	0,75842	102,39	0,000
Error	128	0,94807	0,00741		
Total	131	3,22332			

S = 0,08606 R-Sq = 70,59% R-Sq(adj) = 69,90%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
1	48	0,77271	0,05556
2	29	0,66086	0,09101
3	34	0,49521	0,10866
4	21	0,86352	0,09524

0,48      0,60      0,72      0,84

Pooled StDev = 0,08606

### One-way ANOVA: 1. Governança versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	42336	14112	96,53	0,000
Error	128	18713	146		
Total	131	61049			

S = 12,09 R-Sq = 69,35% R-Sq(adj) = 68,63%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
1	48	56,01	14,26
2	29	42,65	10,27
3	34	36,08	12,10
4	21	90,21	8,27

32      48      64      80

Pooled StDev = 12,09

### One-way ANOVA: 1.5 Knowledge and Technology versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	11248,8	3749,6	66,79	0,000
Error	128	7186,4	56,1		
Total	131	18435,2			

S = 7,493 R-Sq = 61,02% R-Sq(adj) = 60,10%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
1	48	30,071	7,672
2	29	24,928	8,259
3	34	20,385	6,506
4	21	48,776	7,463

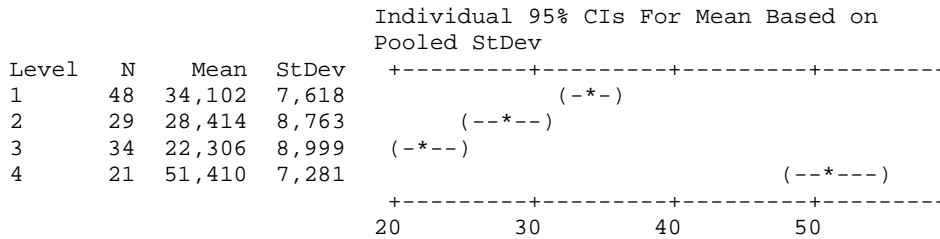
20      30      40      50

Pooled StDev = 7,493

### One-way ANOVA: 1.6 Creative Outputs versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	11649,1	3883,0	57,72	0,000
Error	128	8610,7	67,3		
Total	131	20259,8			

S = 8,202 R-Sq = 57,50% R-Sq(adj) = 56,50%

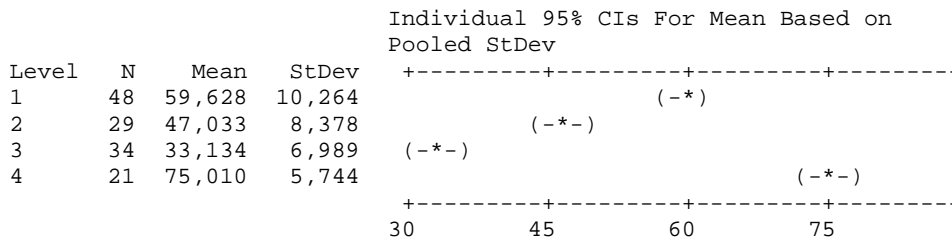


Pooled StDev = 8,202

### One-way ANOVA: EPI Score versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	26697,7	8899,2	123,96	0,000
Error	128	9189,0	71,8		
Total	131	35886,7			

S = 8,473 R-Sq = 74,39% R-Sq(adj) = 73,79%

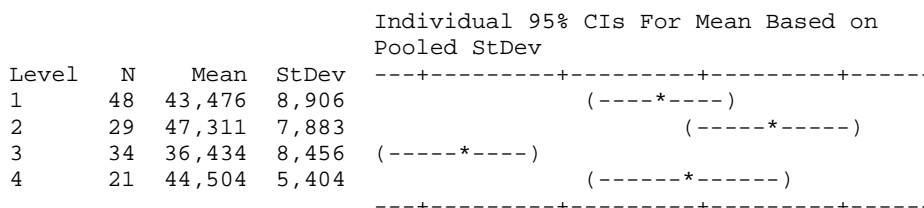


Pooled StDev = 8,473

### One-way ANOVA: Happy Planet Index versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	2048,5	682,8	10,39	0,000
Error	128	8411,1	65,7		
Total	131	10459,6			

S = 8,106 R-Sq = 19,59% R-Sq(adj) = 17,70%





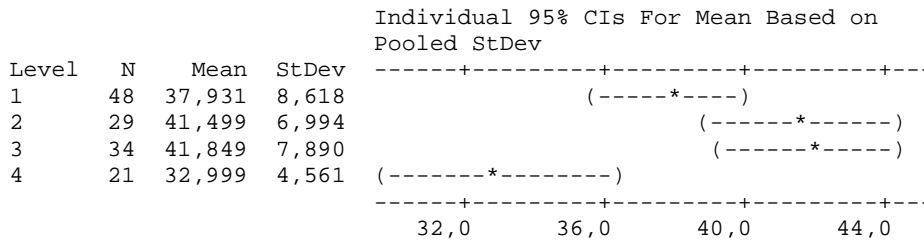
35,0      40,0      45,0      50,0

Pooled StDev = 8,106

### One-way ANOVA: GINI Index versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	1267,6	422,5	7,38	0,000
Error	128	7330,3	57,3		
Total	131	8597,9			

S = 7,568    R-Sq = 14,74%    R-Sq(adj) = 12,75%

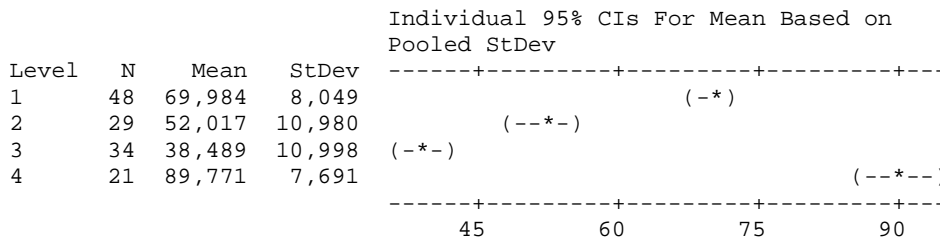


Pooled StDev = 7,568

### One-way ANOVA: Access to Information and Com\_1 versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	40827,6	13609,2	150,24	0,000
Error	128	11594,5	90,6		
Total	131	52422,1			

S = 9,517    R-Sq = 77,88%    R-Sq(adj) = 77,36%

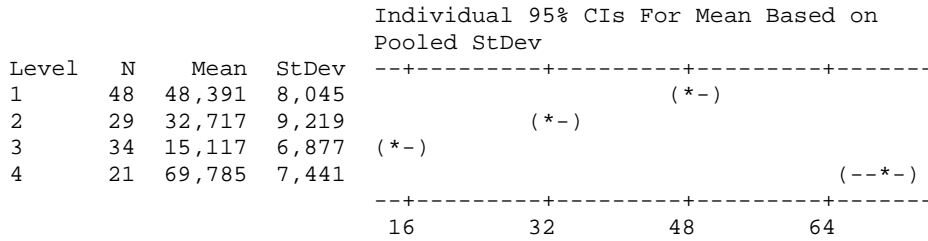


Pooled StDev = 9,517

### One-way ANOVA: Access to Advanced Education\_1 versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	44592,9	14864,3	235,19	0,000
Error	128	8089,9	63,2		
Total	131	52682,8			

S = 7,950    R-Sq = 84,64%    R-Sq(adj) = 84,28%

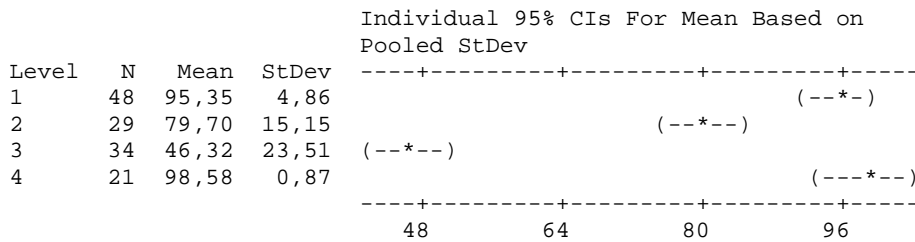


Pooled StDev = 7,950

### One-way ANOVA: Adult literacy rate versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	57126	19042	94,50	0,000
Error	128	25793	202		
Total	131	82920			

S = 14,20 R-Sq = 68,89% R-Sq(adj) = 68,16%

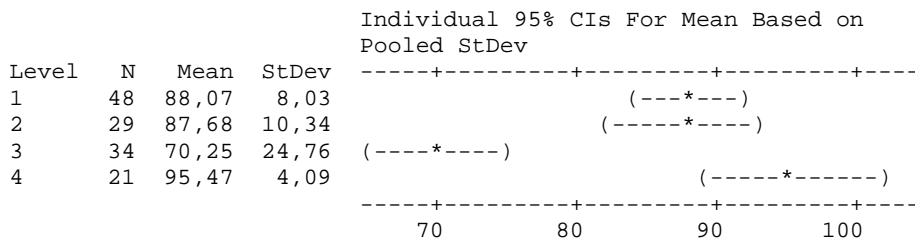


Pooled StDev = 14,20

### One-way ANOVA: Primary scholl versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	10338	3446	16,59	0,000
Error	128	26588	208		
Total	131	36926			

S = 14,41 R-Sq = 28,00% R-Sq(adj) = 26,31%

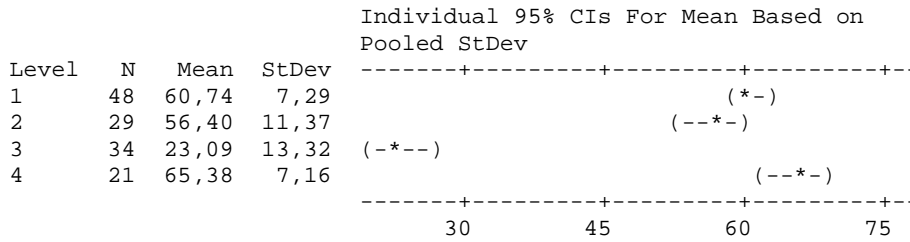


Pooled StDev = 14,41

### One-way ANOVA: Lower secondary versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	36228	12076	118,92	0,000
Error	128	12998	102		
Total	131	49226			

S = 10,08 R-Sq = 73,60% R-Sq(adj) = 72,98%

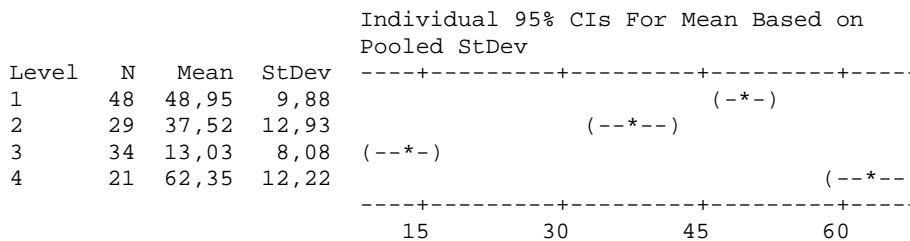


Pooled StDev = 10,08

### One-way ANOVA: Upper secondary versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	39191	13064	116,05	0,000
Error	128	14409	113		
Total	131	53599			

S = 10,61 R-Sq = 73,12% R-Sq(adj) = 72,49%

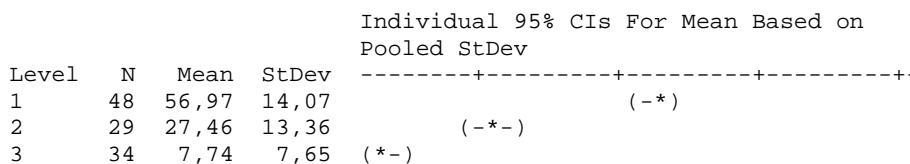


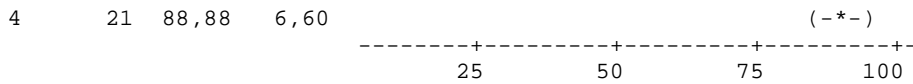
Pooled StDev = 10,61

### One-way ANOVA: Internet users versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	102851	34284	256,47	0,000
Error	128	17110	134		
Total	131	119961			

S = 11,56 R-Sq = 85,74% R-Sq(adj) = 85,40%



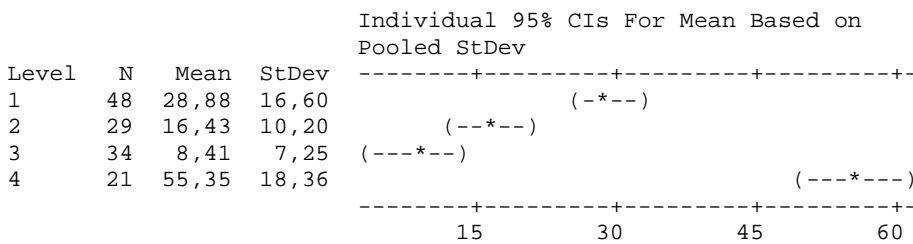


Pooled StDev = 11,56

### One-way ANOVA: Years of teartary versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	31557	10519	55,33	0,000
Error	128	24333	190		
Total	131	55890			

S = 13,79 R-Sq = 56,46% R-Sq(adj) = 55,44%

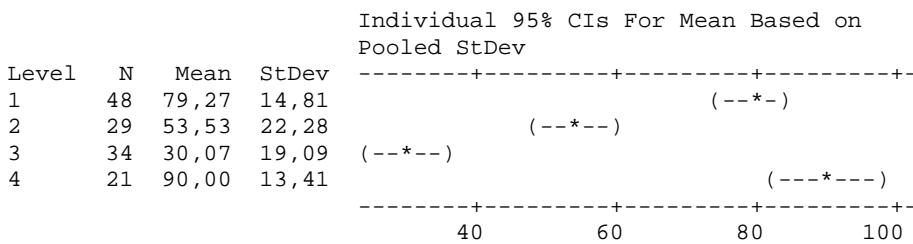


Pooled StDev = 13,79

### One-way ANOVA: Inequality versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	67477	22492	72,29	0,000
Error	128	39828	311		
Total	131	107305			

S = 17,64 R-Sq = 62,88% R-Sq(adj) = 62,01%

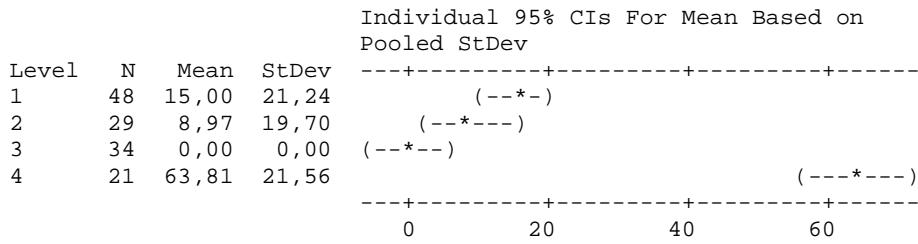


Pooled StDev = 17,64

### One-way ANOVA: Number of globally ranked versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	57860	19287	59,68	0,000
Error	128	41364	323		
Total	131	99224			

S = 17,98 R-Sq = 58,31% R-Sq(adj) = 57,34%

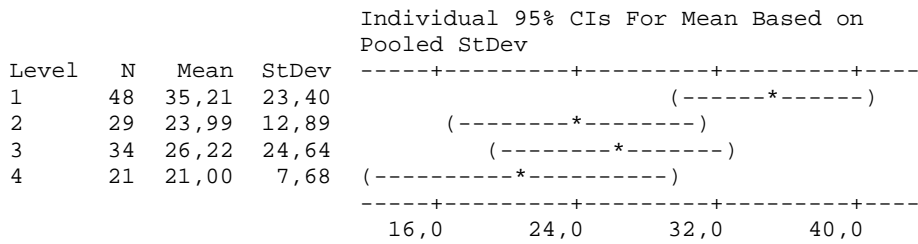


Pooled StDev = 17,98

### One-way ANOVA: Unemployment Total versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	4090	1363	3,38	0,020
Error	128	51598	403		
Total	131	55689			

S = 20,08 R-Sq = 7,35% R-Sq(adj) = 5,17%

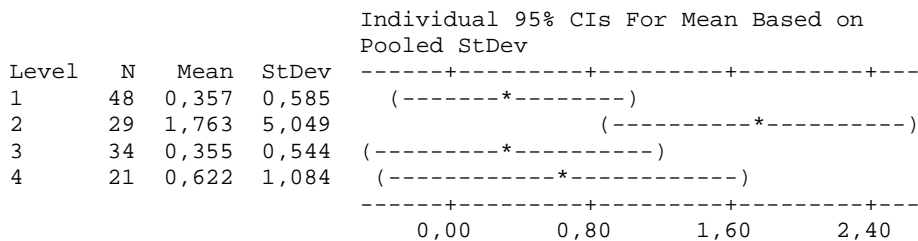


Pooled StDev = 20,08

### One-way ANOVA: Total Labor Force (% Global) versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	42,58	14,19	2,38	0,073
Error	128	763,11	5,96		
Total	131	805,70			

S = 2,442 R-Sq = 5,29% R-Sq(adj) = 3,07%

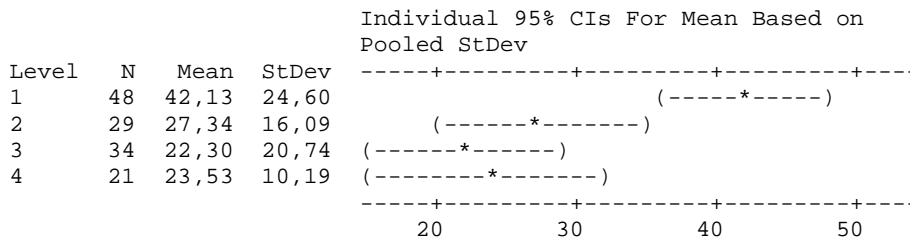


Pooled StDev = 2,442

## One-way ANOVA: Unemployment, youth total versus Região Nova

Source	DF	SS	MS	F	P
Região Nova	3	10078	3359	8,27	0,000
Error	128	51966	406		
Total	131	62044			

S = 20,15    R-Sq = 16,24%    R-Sq(adj) = 14,28%



Pooled StDev = 20,15

### Tabela do F'S superior a 60%

SPI 205,19	IDH 102,39	Governança 96,53	Know 66,79
EPI Score 123,96	Acess 150,24	Acess Adv 235,19	Adult 94,50
Lower 118,92	Upper 116,05	Internet 256,47	

Os F'S em vermelho não foram incluídos na tabela, pois não tem relação significativa com o tema.

## 7. ANÁLISE DISCRIMINANTE

A Análise Discriminante é usada para classificar observações em 2 ou mais grupos se há uma amostra com grupos conhecidos. Também pode ser usada para investigar como as variáveis contribuem para a separação dos grupos.

Através da “Análise Discriminante”, analisaremos qual modelo (linear ou quadrático) e quais agrupamento de variáveis podem explicar melhor e de maneira mais simples a população estudada.

Abaixo a análise discriminante linear, a qual mostrou 92,4% de acerto para essa classificação.

### Discriminant Analysis: Região Nova versus Social Progr; IDH - 2013; ...

Linear Method for Response: Região Nova

Predictors: Social Progress Index; IDH - 2013; 1. Governança;  
1.5 Knowledge and Technology; 1.6 Creative Outputs; EPI Score;

Happy Planet Index; GINNI Index; Access to Information and Com\_1;  
 Access to Advanced Education\_1; Adult literacy rate;  
 Primary scholl; Lower secondary; Upper secondary; Internet users;  
 Years of teartinary; Inequality; Number of globally ranked;  
 Unemployment Total; Total Labor Force (% Global\_1;  
 Unemployment, youth total

Group	1	2	3	4
Count	48	29	34	21

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	45	1	0	0
2	2	28	1	0
3	0	0	33	0
4	1	0	0	21
Total N	48	29	34	21
N correct	45	28	33	21
Proportion	0,938	0,966	0,971	1,000

N = 132

N Correct = 127

Proportion Correct = 0,962

Squared Distance Between Groups

	1	2	3	4
1	0,000	13,551	58,294	40,959
2	13,551	0,000	25,295	84,441
3	58,294	25,295	0,000	145,048
4	40,959	84,441	145,048	0,000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-166,69	-141,50	-106,75	-222,20
Social Progress Index	2,59	2,59	2,51	2,21
IDH - 2013	49,59	54,43	77,49	46,54
1. Governança	-1,38	-1,29	-1,07	-1,17
1.5 Knowledge and Technology	1,45	1,40	1,33	1,75
1.6 Creative Outputs	0,13	0,21	0,25	0,15
EPI Score	-0,24	-0,19	-0,31	-0,46
Happy Planet Index	0,45	0,38	0,19	0,68
GINNI Index	1,41	1,34	1,25	1,37
Access to Information and Com_1	0,68	0,45	0,21	0,97
Access to Advanced Education_1	-0,48	-0,61	-0,89	-0,62
Adult literacy rate	-0,19	-0,19	-0,30	-0,19
Primary scholl	0,35	0,39	0,42	0,42
Lower secondary	-0,07	-0,03	-0,34	-0,13
Upper secondary	0,20	0,19	0,11	0,35
Internet users	0,16	-0,06	-0,18	0,44
Years of teartinary	0,24	0,18	0,20	0,42
Inequality	0,37	0,29	0,30	0,47
Number of globally ranked	-0,08	-0,14	-0,08	0,21
Unemployment Total	0,21	0,30	0,44	0,42
Total Labor Force (% Global_1	-0,36	0,06	-0,34	-1,83
Unemployment, youth total	-0,23	-0,29	-0,36	-0,58

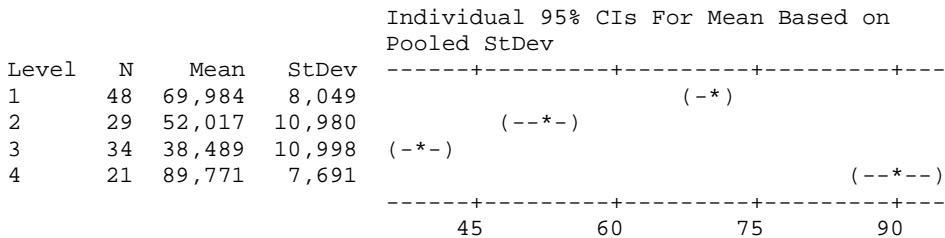
Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
36**	1	2	1	15,18	0,391
			2	14,29	0,609
			3	58,46	0,000
			4	83,75	0,000
86**	3	2	1	48,77	0,000
			2	26,78	0,731
			3	28,78	0,269
			4	142,05	0,000
110**	1	4	1	34,52	0,043
			2	66,61	0,000
			3	126,20	0,000
			4	28,33	0,957
122**	1	2	1	18,98	0,103
			2	14,66	0,897
			3	44,01	0,000
			4	55,97	0,000
129**	2	1	1	32,27	0,912
			2	36,95	0,088
			3	82,95	0,000
			4	93,80	0,000

**One-way ANOVA: Access to Information and Com\_1 versus Regia1**

Source	DF	SS	MS	F	P
Regial	3	40827,6	13609,2	150,24	0,000
Error	128	11594,5	90,6		
Total	131	52422,1			

S = 9,517    R-Sq = 77,88%    R-Sq(adj) = 77,36%

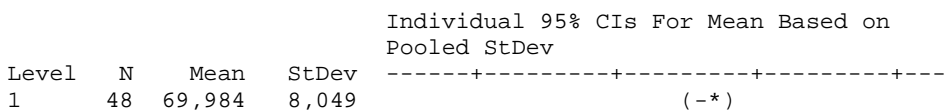


Pooled StDev = 9,517

**One-way ANOVA: Access to Information and Com\_1 versus Regia1**

Source	DF	SS	MS	F	P
Regial	3	40827,6	13609,2	150,24	0,000
Error	128	11594,5	90,6		
Total	131	52422,1			

S = 9,517    R-Sq = 77,88%    R-Sq(adj) = 77,36%





2	29	52,017	10,980	(--*-)
3	34	38,489	10,998	(--*-)
4	21	89,771	7,691	(--*--)

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

Pooled StDev = 9,517

## 8. REGRESSÃO LOGÍSTICA

A regressão logística é uma técnica estatística que tem como objetivo produzir, a partir de um conjunto de observações, um modelo que permita a predição de valores tomados por uma variável categórica, frequentemente binária, a partir de uma série de variáveis explicativas contínuas e/ou binárias<sup>1</sup>. A regressão logística é amplamente usada em ciências médicas e sociais, e tem outras denominações, como modelo logístico e classificador de máxima entropia.

### Discriminant Analysis: Regia1 versus Social Progress ; IDH - 2013; ...

Linear Method for Response: Regial

Predictors: Social Progress Index; IDH - 2013; 1. Governança;  
1.5 Knowledge and Technology; 1.6 Creative Outputs; EPI Score;  
Happy Planet Index; GINNI Index; Access to Information and Com\_1;  
Access to Advanced Education\_1; Adult literacy rate;  
Primary scholl; Lower secondary; Upper secondary; Internet users;  
Years of teartinary; Inequality; Number of globally ranked;  
Unemployment Total; Total Labor Force (% Global\_1;  
Unemployment, youth total

Group	1	2	3	4
Count	48	29	34	21

Summary of classification

		True Group			
Put into Group		1	2	3	4
1		45	1	0	0
2		2	28	1	0
3		0	0	33	0
4		1	0	0	21
Total N		48	29	34	21
N correct		45	28	33	21
Proportion		0,938	0,966	0,971	1,000

N = 132

N Correct = 127

Proportion Correct = 0,962

Squared Distance Between Groups

1	2	3	4
---	---	---	---

1	0,000	13,551	58,294	40,959
2	13,551	0,000	25,295	84,441
3	58,294	25,295	0,000	145,048
4	40,959	84,441	145,048	0,000

#### Linear Discriminant Function for Groups

	1	2	3	4
Constant	-166,69	-141,50	-106,75	-222,20
Social Progress Index	2,59	2,59	2,51	2,21
IDH - 2013	49,59	54,43	77,49	46,54
1. Governança	-1,38	-1,29	-1,07	-1,17
1.5 Knowledge and Technology	1,45	1,40	1,33	1,75
1.6 Creative Outputs	0,13	0,21	0,25	0,15
EPI Score	-0,24	-0,19	-0,31	-0,46
Happy Planet Index	0,45	0,38	0,19	0,68
GINNI Index	1,41	1,34	1,25	1,37
Access to Information and Com_1	0,68	0,45	0,21	0,97
Access to Advanced Education_1	-0,48	-0,61	-0,89	-0,62
Adult literacy rate	-0,19	-0,19	-0,30	-0,19
Primary scholl	0,35	0,39	0,42	0,42
Lower secondary	-0,07	-0,03	-0,34	-0,13
Upper secondary	0,20	0,19	0,11	0,35
Internet users	0,16	-0,06	-0,18	0,44
Years of teartinary	0,24	0,18	0,20	0,42
Inequality	0,37	0,29	0,30	0,47
Number of globally ranked	-0,08	-0,14	-0,08	0,21
Unemployment Total	0,21	0,30	0,44	0,42
Total Labor Force (% Global_1	-0,36	0,06	-0,34	-1,83
Unemployment, youth total	-0,23	-0,29	-0,36	-0,58

#### Summary of Misclassified Observations

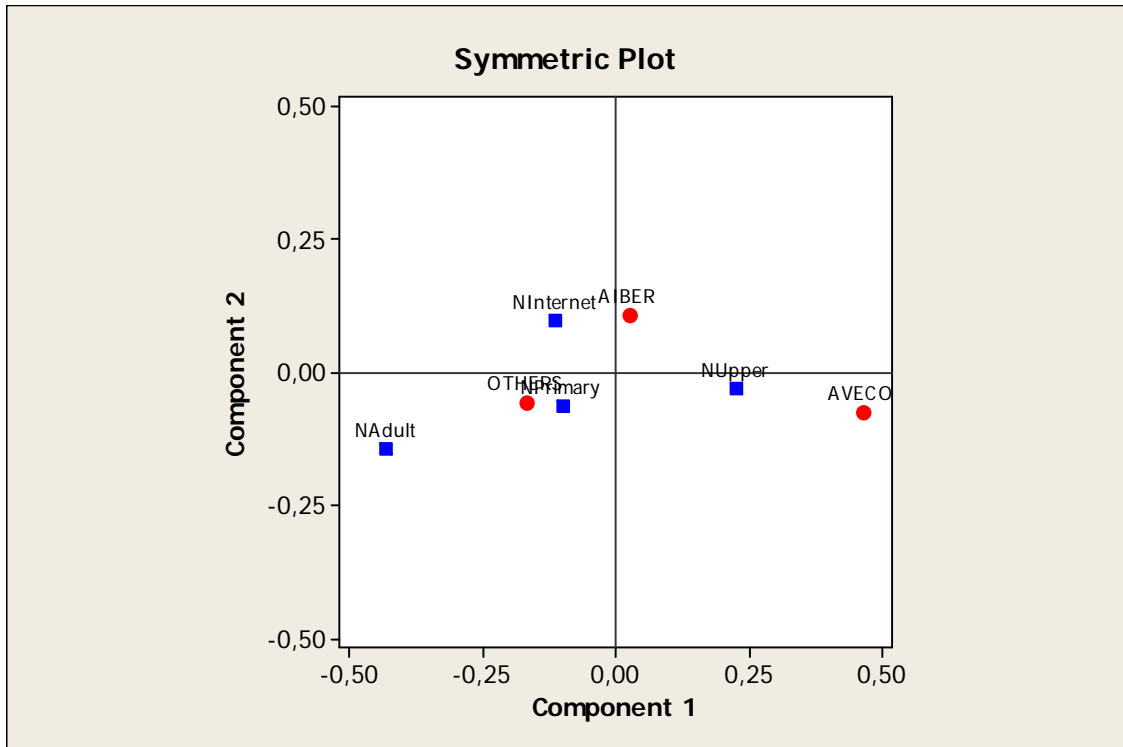
Observation	True Group	Pred Group	Group	Squared Distance	Probability
36**	1	2	1	15,18	0,391
			2	14,29	0,609
			3	58,46	0,000
			4	83,75	0,000
86**	3	2	1	48,77	0,000
			2	26,78	0,731
			3	28,78	0,269
			4	142,05	0,000
110**	1	4	1	34,52	0,043
			2	66,61	0,000
			3	126,20	0,000
			4	28,33	0,957
122**	1	2	1	18,98	0,103
			2	14,66	0,897
			3	44,01	0,000
			4	55,97	0,000
129**	2	1	1	32,27	0,912
			2	36,95	0,088
			3	82,95	0,000
			4	93,80	0,000

## 9. ANÁLISE DE CORRESPONDENCIA

Análise de correspondência é uma técnica de análise exploratória de dados adequada para analisar tabelas de duas entradas ou tabelas de múltiplas entradas, levando em conta algumas medidas de correspondência entre linhas e colunas. Consiste na conversão de uma matriz de dados não negativos em um tipo particular de representação gráfica em que as linhas e colunas da matriz são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico. Este método permite estudar as relações e semelhanças existentes entre as categorias de linhas e entre as categorias de colunas de uma tabela de contingência ou o conjunto de categorias de linhas e o conjunto categorias de colunas.

A análise de correspondência mostra como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe. Embora seja considerada uma técnica descritiva e exploratória, esta análise simplifica dados complexos e produz análises exaustivas de informações que suportam conclusões a respeito das mesmas.

Possui diversos aspectos que a distingue de outras técnicas de análise de dados. A sua natureza multivariada permite revelar relações que não seriam detectadas em comparações aos pares das variáveis. É altamente flexível quanto a pressuposições sobre os dados: o único requisito é o de uma matriz retangular com entradas não negativas. É possível transformar qualquer característica quantitativa em qualitativa, realizando-se uma partição de seu domínio de variação em classes. A análise de correspondência é mais efetiva se a matriz de dados é bastante grande, de modo que a inspeção visual ou análise estatística simples não consegue revelar sua estrutura.



**Simple Correspondence Analysis: NAdult; NUpper; NPrimary; NInternet**

Relative Inertias

	NAdult	NUpper	NPrimary	NInternet	Total
AIBER	0,040	0,001	0,006	0,033	0,079
AVECO	0,190	0,327	0,006	0,118	0,640
OTHERS	0,178	0,089	0,013	0,002	0,281
Total	0,407	0,416	0,025	0,152	1,000

Analysis of Contingency Table

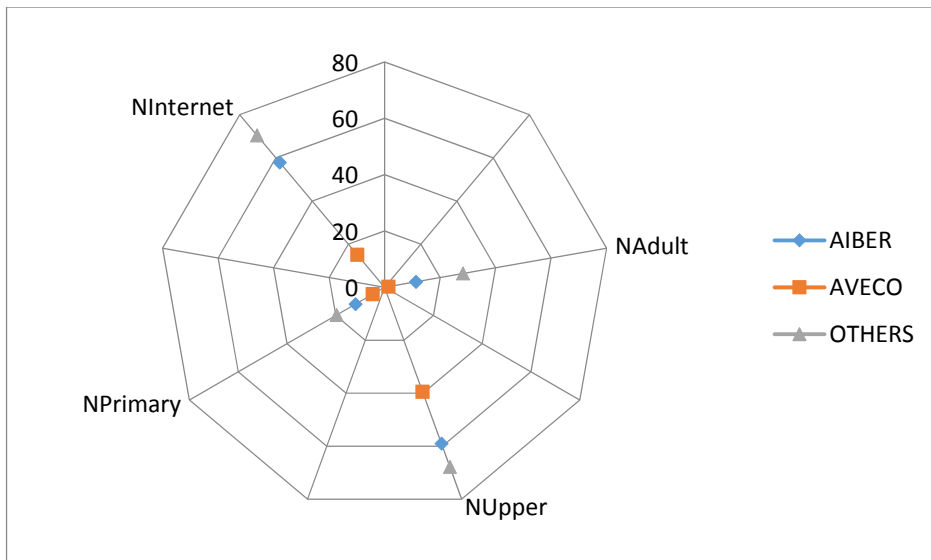
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0475	0,8806	0,8806	*****
2	0,0064	0,1194	1,0000	****
Total	0,0540			

Row Contributions

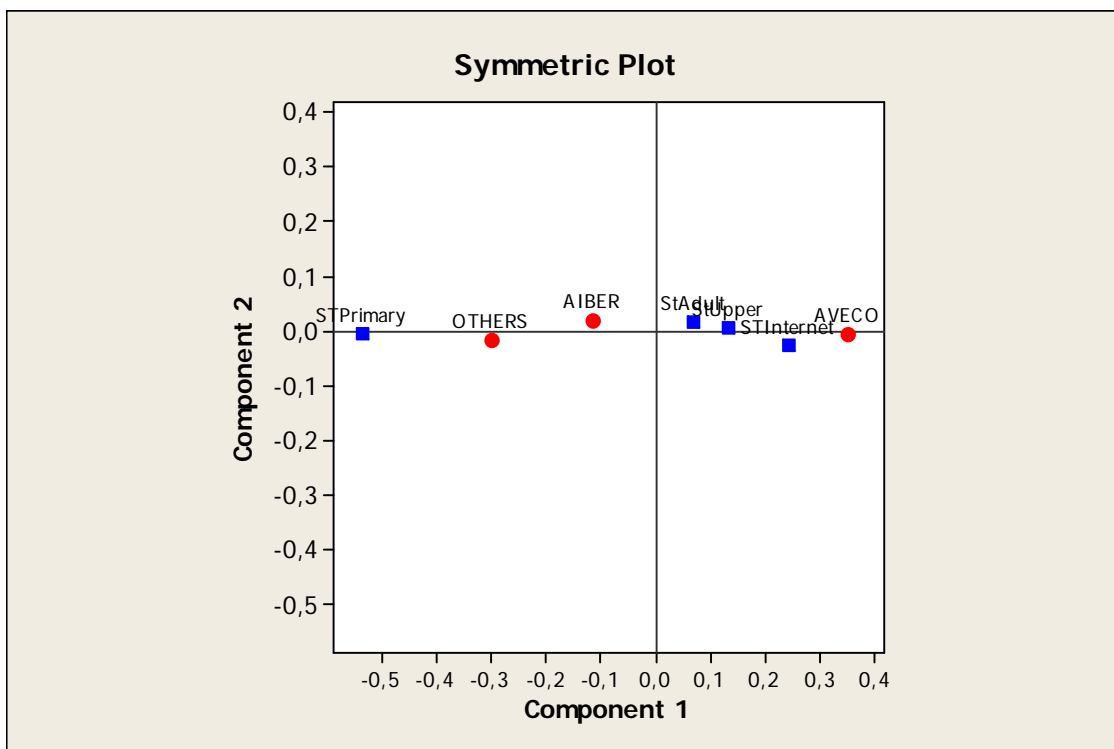
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	1,000	0,362	0,079	0,023	0,045	0,004	0,106	0,955	0,634
2	AVECO	1,000	0,157	0,640	0,463	0,975	0,709	-0,074	0,025	0,134
3	OTHERS	1,000	0,481	0,281	-0,168	0,901	0,287	-0,056	0,099	0,232

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	NAdult	1,000	0,106	0,407	-0,432	0,900	0,416	-0,144	0,100	0,341
2	NUpper	1,000	0,430	0,416	0,226	0,980	0,463	-0,033	0,020	0,071
3	NPrimary	1,000	0,094	0,025	-0,100	0,700	0,020	-0,065	0,300	0,062
4	NInternet	1,000	0,370	0,152	-0,114	0,586	0,101	0,096	0,414	0,526



## DESVIO DAS VARIÁVEIS analíticas 3 REGIÕES



## Simple Correspondence Analysis: NAdult; NUpper; NPrimary; NInternet

Relative Inertias

	StAdult	StUpper	STPrimary	STInternet	Total
AIBER	0,000	0,003	0,040	0,016	0,059
AVECO	0,010	0,039	0,442	0,095	0,586
OTHERS	0,010	0,026	0,275	0,043	0,355
Total	0,021	0,068	0,758	0,154	1,000

Analysis of Contingency Table

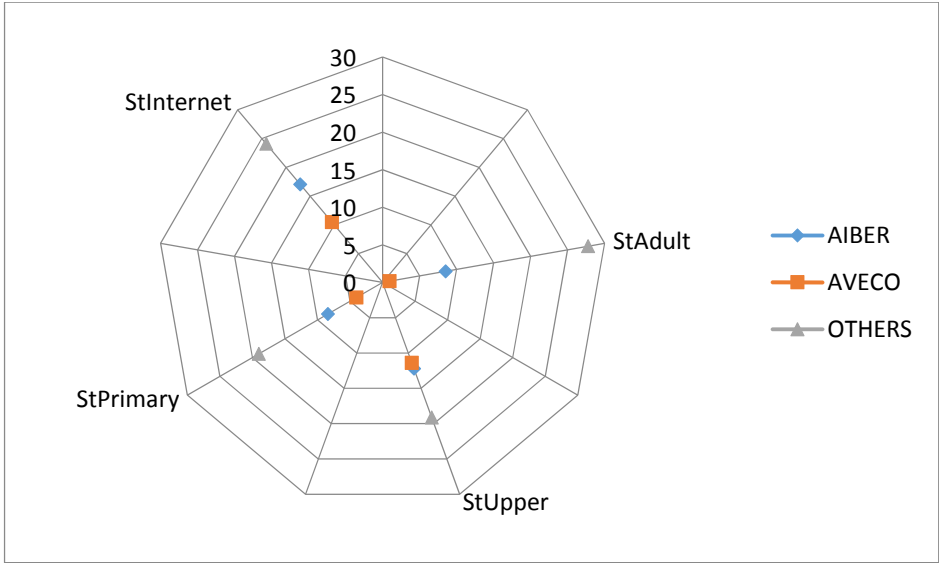
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0767	0,9971	0,9971	*****
2	0,0002	0,0029	1,0000	
Total	0,0769			

Row Contributions

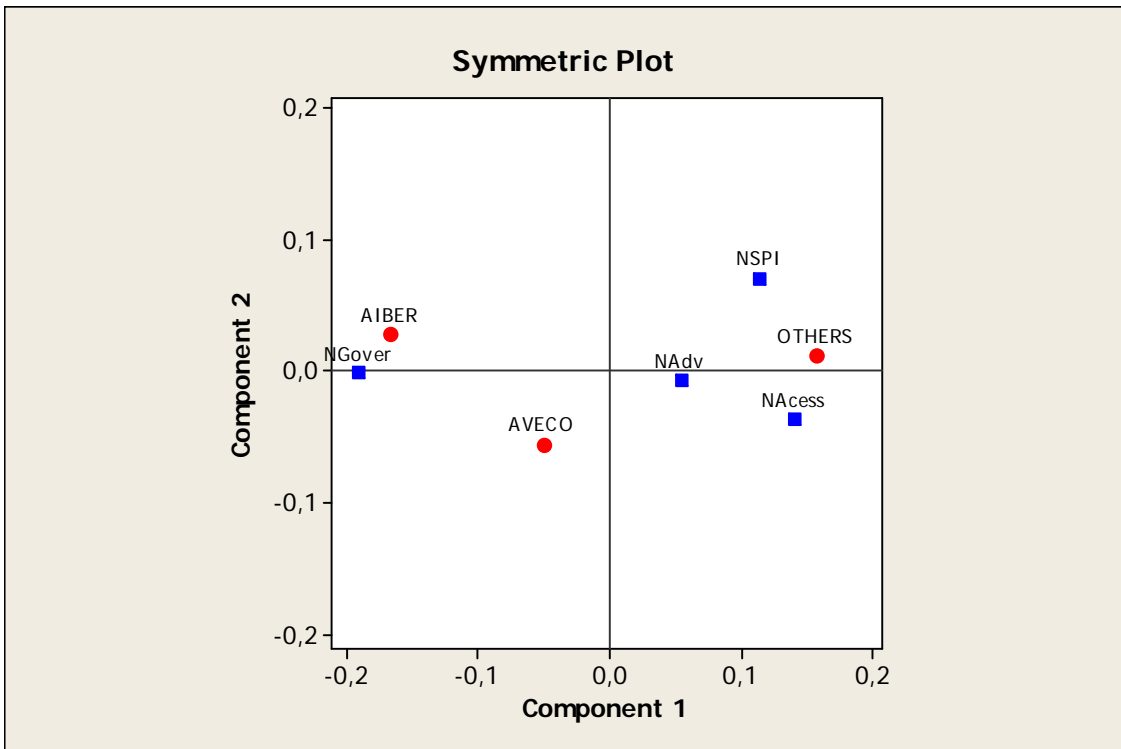
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	1,000	0,334	0,059	-0,115	0,970	0,058	0,020	0,030	0,609
2	AVECO	1,000	0,367	0,586	0,350	1,000	0,587	-0,005	0,000	0,045
3	OTHERS	1,000	0,299	0,355	-0,302	0,997	0,355	-0,016	0,003	0,346

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	StAdult	1,000	0,302	0,021	0,071	0,955	0,020	0,015	0,045	0,316
2	StUpper	1,000	0,296	0,068	0,133	0,998	0,068	0,006	0,002	0,050
3	STPrimary	1,000	0,203	0,758	-0,535	1,000	0,760	-0,006	0,000	0,036
4	STInternet	1,000	0,199	0,154	0,242	0,989	0,152	-0,026	0,011	0,598



**DESVIO DAS VARIÁVEIS sintéticas 3 REGIÕES**





## Simple Correspondence Analysis: StSPI; StAccess; StGover; StAdv

### Relative Inertias

	NSPI	NAccess	NGover	NAdv	Total
AIBER	0,014	0,140	0,263	0,019	0,436
AVECO	0,048	0,000	0,020	0,000	0,068
OTHERS	0,073	0,096	0,309	0,017	0,496
Total	0,134	0,237	0,592	0,037	1,000

### Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0200	0,9473	0,9473	*****
2	0,0011	0,0527	1,0000	*
Total	0,0211			

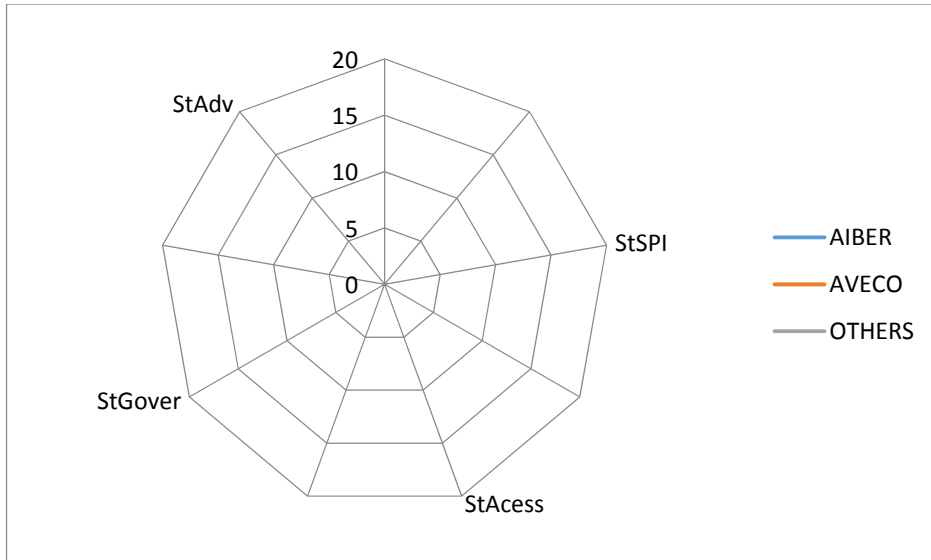
### Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AIBER	1,000	0,319	0,436	-0,167	0,972	0,447	0,029	0,028	0,234
2	AVECO	1,000	0,257	0,068	-0,050	0,454	0,033	-0,055	0,546	0,710
3	OTHERS	1,000	0,424	0,496	0,157	0,994	0,520	0,012	0,006	0,056

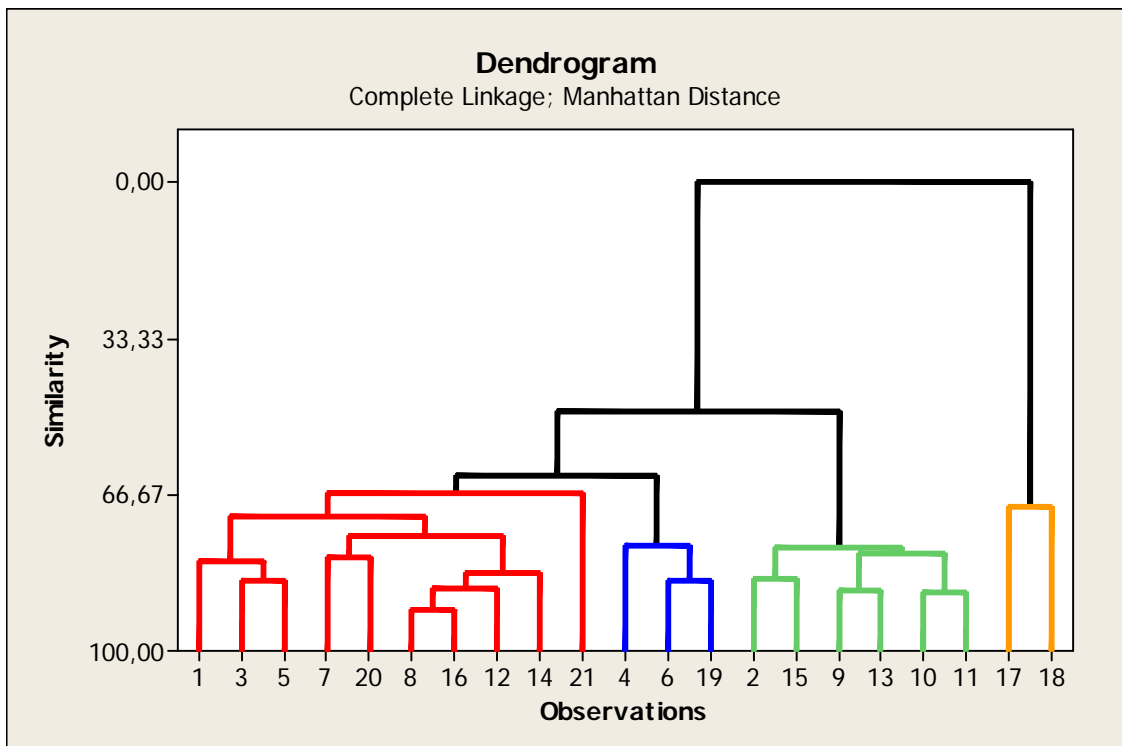
### Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	NSPI	1,000	0,158	0,134	0,114	0,727	0,103	0,070	0,273	0,698
2	NAccess	1,000	0,239	0,237	0,140	0,936	0,234	-0,037	0,064	0,288

3	NGover	1,000	0,343	0,592	-0,191	1,000	0,625	-0,001	0,000	0,000
4	NAdv	1,000	0,260	0,037	0,054	0,981	0,038	-0,008	0,019	0,013



## AIBER



**Results for: Worksheet 14**

**Cluster Analysis of Observations: Social Progr; Access to In; Access to Ad;**

Manhattan Distance, Complete Linkage

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters		Number of obs.	
				joined	new cluster	in new cluster	
1	20	91,1047	77,220	8	16	8	2
2	19	87,5144	108,387	10	11	10	2
3	18	86,9578	113,218	9	13	9	2
4	17	86,8422	114,222	8	12	8	3
5	16	85,2131	128,364	3	5	3	2
6	15	84,9064	131,027	6	19	6	2
7	14	84,6785	133,005	2	15	2	2
8	13	83,4225	143,909	8	14	8	4
9	12	81,1062	164,016	1	3	1	3
10	11	80,2515	171,436	7	20	7	2
11	10	79,4514	178,381	9	10	9	4
12	9	78,0499	190,547	2	9	2	6
13	8	77,6611	193,923	4	6	4	3
14	7	75,5248	212,468	7	8	7	6
15	6	71,3070	249,083	1	7	1	9
16	5	69,2923	266,572	17	18	17	2
17	4	66,4470	291,272	1	21	1	10
18	3	62,8215	322,745	1	4	1	13
19	2	49,1708	441,246	1	2	1	19
20	1	0,0000	868,095	1	17	1	21

Final Partition

Number of clusters: 4

		Average	Maximum
	Within	distance	distance
	Number of	from	from
	observations	centroid	centroid
	cluster	sum	
	of squares		
Cluster1	10	16161,1	38,7019
Cluster2	6	5541,6	30,2591
Cluster3	3	2218,5	26,9118
Cluster4	2	4746,0	48,7135

#### Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Social Progress Index	66,9110	62,538	77,1867	80,6300
Access to Information and Commu	58,5140	57,362	74,7667	78,9050
Access to Advanced Education	41,5330	27,682	45,9233	58,4950
1.5 Knowledge and Technology_1	24,0600	17,183	27,2333	37,9000
1.6 Creative Outputs_1	33,2700	27,100	36,3333	43,4000
EPI Score_1	53,4860	46,795	60,6900	77,7950
Happy Planet Index_1	54,5989	53,032	52,4134	41,3702
Adult literacy rate_1	8,5601	20,414	3,1755	4,5634
Primary scholl_1	11,7026	17,079	7,4965	1,0160
Lower secondary_1	39,5227	54,400	32,1454	23,9858
Upper secondary_1	58,6697	68,903	57,8977	31,5027
Internet users_1	56,4087	77,706	43,7226	29,6979
Years of teartinary_1	72,8824	81,863	68,2353	47,0496
Inequality_1	59,7526	42,337	75,0515	91,2371
Number of globally ranked_1	86,0000	100,000	93,3333	40,0000
Unemployment Total_1	79,2834	85,613	79,0445	30,7818
Unemployment, youth total_1	76,8007	86,879	71,6918	27,5455

IDH 2013_1	31,7792	49,423	25,0961	16,2273
Governança_1	64,9251	69,403	28,7755	27,9204
GINNI_1	42,0538	37,800	41,2578	67,9541

Grand

Variable	centroid
Social Progress Index	68,4362
Access to Information and Commu	62,4486
Access to Advanced Education	39,8181
1.5 Knowledge and Technology_1	23,8667
1.6 Creative Outputs_1	32,9095
EPI Score_1	54,9186
Happy Planet Index_1	52,5791
Adult literacy rate_1	10,7970
Primary scholl_1	11,6201
Lower secondary_1	41,2399
Upper secondary_1	58,8957
Internet users_1	58,1374
Years of teartinary_1	72,3241
Inequality_1	59,9607
Number of globally ranked_1	86,6667
Unemployment Total_1	76,4387
Unemployment, youth total_1	74,2593
IDH 2013_1	34,3846
Governança_1	57,5160
GINNI_1	43,1914

#### Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0,000	48,631	49,147	120,872
Cluster2	48,631	0,000	84,963	161,656
Cluster3	49,147	84,963	0,000	101,954

Cluster4 120,872 161,656 101,954 0,000

## Dendrogram

### Cluster Analysis of Observations: Social Progr; Access to In; Access to Ad;

Manhattan Distance, Complete Linkage

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters		Number of obs.	
				joined	cluster	New cluster	in new cluster
1	20	91,1047	77,220	8	16	8	2
2	19	87,5144	108,387	10	11	10	2
3	18	86,9578	113,218	9	13	9	2
4	17	86,8422	114,222	8	12	8	3
5	16	85,2131	128,364	3	5	3	2
6	15	84,9064	131,027	6	19	6	2
7	14	84,6785	133,005	2	15	2	2
8	13	83,4225	143,909	8	14	8	4
9	12	81,1062	164,016	1	3	1	3
10	11	80,2515	171,436	7	20	7	2
11	10	79,4514	178,381	9	10	9	4
12	9	78,0499	190,547	2	9	2	6
13	8	77,6611	193,923	4	6	4	3
14	7	75,5248	212,468	7	8	7	6
15	6	71,3070	249,083	1	7	1	9
16	5	69,2923	266,572	17	18	17	2
17	4	66,4470	291,272	1	21	1	10
18	3	62,8215	322,745	1	4	1	13
19	2	49,1708	441,246	1	2	1	19
20	1	0,0000	868,095	1	17	1	21

Final Partition

Number of clusters: 4

			Average	Maximum
		Within	distance	distance
	Number of	cluster sum	from	from
	observations	of squares	centroid	centroid
Cluster1	10	16161,1	38,7019	61,8654
Cluster2	6	5541,6	30,2591	35,2037
Cluster3	3	2218,5	26,9118	30,0036
Cluster4	2	4746,0	48,7135	48,7135

Cluster Centroids

Variable	Cluster1	Cluster2	Cluster3	Cluster4
Social Progress Index	66,9110	62,538	77,1867	80,6300
Access to Information and Commu	58,5140	57,362	74,7667	78,9050
Access to Advanced Education	41,5330	27,682	45,9233	58,4950
1.5 Knowledge and Technology_1	24,0600	17,183	27,2333	37,9000
1.6 Creative Outputs_1	33,2700	27,100	36,3333	43,4000
EPI Score_1	53,4860	46,795	60,6900	77,7950
Happy Planet Index_1	54,5989	53,032	52,4134	41,3702
Adult literacy rate_1	8,5601	20,414	3,1755	4,5634
Primary scholl_1	11,7026	17,079	7,4965	1,0160
Lower secondary_1	39,5227	54,400	32,1454	23,9858
Upper secondary_1	58,6697	68,903	57,8977	31,5027
Internet users_1	56,4087	77,706	43,7226	29,6979
Years of teartary_1	72,8824	81,863	68,2353	47,0496
Inequality_1	59,7526	42,337	75,0515	91,2371

Number of globally ranked_1	86,0000	100,000	93,3333	40,0000
Unemployment Total_1	79,2834	85,613	79,0445	30,7818
Unemployment, youth total_1	76,8007	86,879	71,6918	27,5455
IDH 2013_1	31,7792	49,423	25,0961	16,2273
Governança_1	64,9251	69,403	28,7755	27,9204
GINNI_1	42,0538	37,800	41,2578	67,9541

Grand

Variable	centroid
Social Progress Index	68,4362
Access to Information and Commu	62,4486
Access to Advanced Education	39,8181
1.5 Knowledge and Technology_1	23,8667
1.6 Creative Outputs_1	32,9095
EPI Score_1	54,9186
Happy Planet Index_1	52,5791
Adult literacy rate_1	10,7970
Primary scholl_1	11,6201
Lower secondary_1	41,2399
Upper secondary_1	58,8957
Internet users_1	58,1374
Years of teartinary_1	72,3241
Inequality_1	59,9607
Number of globally ranked_1	86,6667
Unemployment Total_1	76,4387
Unemployment, youth total_1	74,2593
IDH 2013_1	34,3846
Governança_1	57,5160
GINNI_1	43,1914

Distances Between Cluster Centroids

Cluster1 Cluster2 Cluster3 Cluster4



Cluster1	0,000	48,631	49,147	120,872
Cluster2	48,631	0,000	84,963	161,656
Cluster3	49,147	84,963	0,000	101,954
Cluster4	120,872	161,656	101,954	0,000

## Dendrogram

## Discriminant Analysis: CLASS versus Social Progr; Access to In;

Linear Method for Response: CLASS

Predictors: Social Progress Index; Access to Information and Commu;  
 Access to Advanced Education; Governança\_1; Adult literacy rate\_1;  
 Primary scholl\_1; Upper secondary\_1; Internet users\_1

Group	1	2	3	4
Count	10	6	3	2

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	10	0	0	0
2	0	6	0	0
3	0	0	3	0
4	0	0	0	2
Total N	10	6	3	2

N correct	10	6	3	2
Proportion	1,000	1,000	1,000	1,000

N = 21                      N Correct = 21                      Proportion Correct = 1,000

#### Squared Distance Between Groups

	1	2	3	4
1	0,000	18,320	29,839	51,842
2	18,320	0,000	57,848	109,273
3	29,839	57,848	0,000	25,462
4	51,842	109,273	25,462	0,000

#### Linear Discriminant Function for Groups

	1	2	3	4
Constant	-1492,6	-1458,3	-1499,5	-1544,1
Social Progress Index	37,9	36,4	38,6	39,7
Access to Information and Commu	-6,6	-6,1	-6,4	-6,8
Access to Advanced Education	2,8	2,8	2,2	2,5
Governança_1	7,2	6,9	6,6	7,0
Adult literacy rate_1	8,1	7,8	7,2	7,9
Primary scholl_1	4,9	5,0	4,7	4,6
Upper secondary_1	4,0	4,0	3,8	3,4
Internet users_1	-2,0	-1,2	-1,6	-2,2

## 10. ÁRVORE DE CLASSIFICAÇÃO

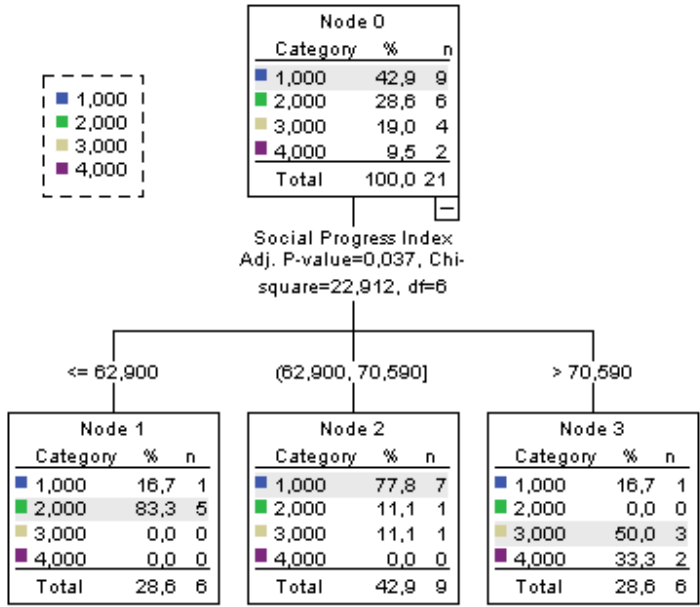
A árvore de classificação indica qual a variável melhor separa os grupos e classifica as variáveis por ordem de importância na separação dos grupos.

Abaixo a árvore de classificação com o método CHAID para os grupos com as variáveis.

**Model Summary**

Specifications	Growing Method	EXHAUSTIVE CHAID	
	Dependent Variable	CLASSIFICAÇÃO	
	Independent Variables	Adult literacy rate, Primary scholl, Upper secondary, Internet users, Social Progress Index, Access to Information and Commu , Access to Advanced Education, Governança	
	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	
	Number of Nodes		4
	Number of Terminal Nodes		3
	Depth		1

**CLASSIFICAÇÃO**



**Risk**

Estimate	Std. Error
,286	,099

Growing Method:

EXHAUSTIVE CHAID

Dependent Variable:

CLASSIFICAÇÃO

**Classification**

Observed	Predicted				
	1,00	2,00	3,00	4,00	Percent Correct
1,00	7	1	1	0	77,8%
2,00	1	5	0	0	83,3%
3,00	1	0	3	0	75,0%

4,00	0	0	2	0	,0%
Overall Percentage	42,9%	28,6%	28,6%	,0%	71,4%

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: CLASSIFICAÇÃO

ORDEM	PAÍS	CLASS
1	Argentina	1
2	Bolivia	2
3	Brazil	1
4	Chile	3
5	Colombia	1
6	Costa Rica	3
7	Dominican Republic	1
8	Ecuador	1
9	El Salvador	2
10	Guatemala	2
11	Honduras	2
12	Mexico	1
13	Nicaragua	2
14	Panama	1
15	Paraguay	2
16	Peru	1
17	Portugal	4
18	Spain	4
19	Uruguay	3
20	Venezuela	1

21	Cuba	1
----	------	---

Observed	Predicted				Percent Correct
	1,00	2,00	3,00	4,00	
1,00	7	1	1	0	77,8%
2,00	1	5	0	0	83,3%
3,00	1	0	3	0	75,0%
4,00	0	0	2	0	,0%
Overall Percentage	42,9%	28,6%	28,6%	,0%	71,4%

Growing Method: EXHAUSTIVE CHAID

Dependent Variable: CLASSIFICAÇÃO

A análise dos dados acima nos mostra 71,4% de acerto e usa na árvore apenas a variável Adult Literacy.

## CONSIDERAÇÕES FINAIS

O modelo de análise que melhor descreve a classificação nos três grupos estudados (AIBER, AVECO E OTHERS) temos a análise discriminante quadrática, a qual mostrou 100% de acerto para essa classificação, a que encontramos a menor taxa de acerto temos a árvore de classificação a qual a análise dos dados nos mostra 71,4% de acerto no método CHAID, o método de regressão logística nos mostrou 96,2% de precisão e mostrou que as variáveis **SPI, EPI SCORE, IDH,GOVERNANÇA,KNOW,ACCESS TO INFORMATION, ACCESS TO EDUCATION, ADULT, LOWER,UPPER,INTERNET**, estão com um valor de P elevado indicando que essas variáveis não contribuem tanto com o modelo, a análise discriminante linear, mostrou 92,4% de acerto para essa classificação, e a análise dos dados com a árvore de classificação com método EXHAUSTIVE CHAID mostrou 71,4% de acerto e mostra a relevância das variáveis Access to Education, Access to Information, Upper, Adult,Lower,Internet para o entendimento do modelo.

Observamos que os países desenvolvidos representados pelo grupo AVECO são os mais bem posicionados, enquanto que o grupo OTHERS estão bastante atrás.

Os países do grupo AIBER se encontram quase sempre em uma posição intermediária, ficando atrás nos indicadores SPI e GINI Index, e estando a frente com os indicadores Happy Planet Index e EPI Score, podemos notar pela análise discriminante linear que Portugal e Espanha seriam reclassificados para o grupo AVECO o que nos mostra que o grupo AIBER está ainda no geral em uma posição intermediária devido ao bom

desempenho de alguns países e que sem esses a maioria dos países seria reclassificados para o Grupo OTHERS.

ORDEM	PAÍS	CLASS
1	Argentina	1
2	Bolivia	2
3	Brazil	1
4	Chile	3
5	Colombia	1
6	Costa Rica	3
7	Dominican Republic	1
8	Ecuador	1
9	El Salvador	2
10	Guatemala	2
11	Honduras	2
12	Mexico	1
13	Nicaragua	2
14	Panama	1
15	Paraguay	2
16	Peru	1
17	Portugal	4
18	Spain	4
19	Uruguay	3
20	Venezuela	1
21	Cuba	1







Os países coloridos em vermelho apresentam pior desempenho no tema O futuro da Educação são os seguintes países Bolívia, El Salvador, Guatemala, Honduras, Nicarágua e Paraguai. Os países coloridos em azul tem o melhor desempenho da América Latina no tema, são os seguintes países Argentina, Brasil, Colômbia, República Dominicana, México, Panamá, Peru, Venezuela e Cuba. Em relação aos países amarelos que são um pouco piores do que os países em azul e melhores em relação aos países vermelhos, são os seguintes países Chile, Costa Rica e Uruguai. A Espanha e Portugal pertencem ao grupo AIBER, mas deveriam estar no grupo AVECO, porque esses países estão em um contexto bem melhor do que todos os demais países do grupo AIBER.

