

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

## Trabalho Final

# **“Análise dos países que compõem o Índice Planeta Vivo e a Pegada Ecológica”**

Disciplina: Métodos Quantitativos

Professor: Dr. Arnaldo Jose de Hoyos

**Marisa Gianetti**

## 1. INTRODUÇÃO

O presente trabalho tem por objetivo uma análise de regressões múltiplas, dos componentes principais, de correspondência, Regressão logística binária e análise discriminante com dados extraídos do Relatório Planeta Vivo de 2006, que mostra o estado do ambiente natural e o impacto exercido por atividades humanas até aquele ano. O principal propósito é identificar a situação dos países no contexto estudado. Para tal iniciamos com análise da estatística descritiva. Em seguida passamos para a análise da comparação de médias e variância através da One-Way Analysis of Variance. Na terceira parte utilizamos a análise discriminante, regressão logística e árvore de classificação. Com base nesses dados, será demonstrada qual modelo oferece melhor análise de classificação. Por fim, fazemos as considerações finais. Os softwares estatísticos utilizados são MINITAB e Xlstat.

## 2. ENTENDENDO OS DADOS

Neste trabalho utilizamos parte dos dados compilados pelo WWF<sup>1</sup> (<http://www.wwf.org.br/informacoes/biblioteca/?4420>) para a edição do Relatório Planeta Vivo 2006<sup>2</sup>. Esse relatório descreve “o estado da mudança da biodiversidade global e o nível de pressão na biosfera causado pelo consumo humano de recursos naturais. “É elaborado com base em dois indicadores: O Índice Planeta Vivo, que reflete a saúde dos ecossistemas do planeta e a Pegada Ecológica, que apresenta a extensão da exigência humana sobre esses ecossistemas” (WWF, 2006, p. 3). Agregamos também a variável PIB Per Capta, extraída do Relatório de Desenvolvimento Humano (<http://hdr.undp.org>)<sup>3</sup>, para complemento de análises econômicas.

### 2.1. Os Indivíduos

Os indivíduos desta análise referem-se aos países incluídos do Relatório Planeta Vivo 2006 (<http://www.wwf.org.br/informacoes/biblioteca/?4420>).

### 2.2. As Variáveis

São 7 as variáveis desta pesquisa, descritas na Tabela , abaixo.

**Tabela 1. As Variáveis**

Variável	Ano	Significado	Tipo	Unid Medida
População	2006	População	Variável Quantitativa	Milhões
Pegada Ecológica Total	2006	Medida do impacto humano sobre o planeta	Variável Quantitativa	hectares globais por pessoa, em 2003, hag
Captações água por pessoa	2006	Quantidade captações de água por pessoa, por país	Variável Quantitativa	1000 m3/ano

<sup>1</sup> (População, Pegada Ecológica Total, Captações água por pessoa, Biocapacidade total, Reserva ecológica ou déficit e IDH).

<sup>2</sup> Disponível sob licença Creative Commons

<sup>3</sup> Os valores de PIB Per Capta não encontrados no site <http://hdr.undp.org>, foram pesquisados no site <http://www.indexmundi.com/pt> (acesso em 07/06/2011).

Biocapacidade total	2006	Medida da quantidade de recursos produzidos pelos ecossistemas	Variável Quantitativa	hectares globais por pessoa, em 2003, hag
Reserva ecológica ou déficit	2006	Reserva ou déficit dos ativos ecológicos(ex: banco de pesca, florestas, zonas de cultivo, etc)	Variável Quantitativa	ha global/pessoa
PIB Per capita	2010	Soma do valor adicionado por todos os produtores residentes na economia mais quaisquer impostos sobre os produtos (menos subsídios) não incluídos na valoração da produção, calculado sem deduções por depreciação de ativos de capital fabricados ou para esgotamento e a degradação dos recursos naturais.	Variável Quantitativa	US\$ PPP
IDH	2006	mede o desenvolvimento através da combinação de indicadores de expectativa de vida, nível educacional e renda em um índice composto de desenvolvimento humano	Variável Quantitativa	Índice com Valores entre 0 e 1

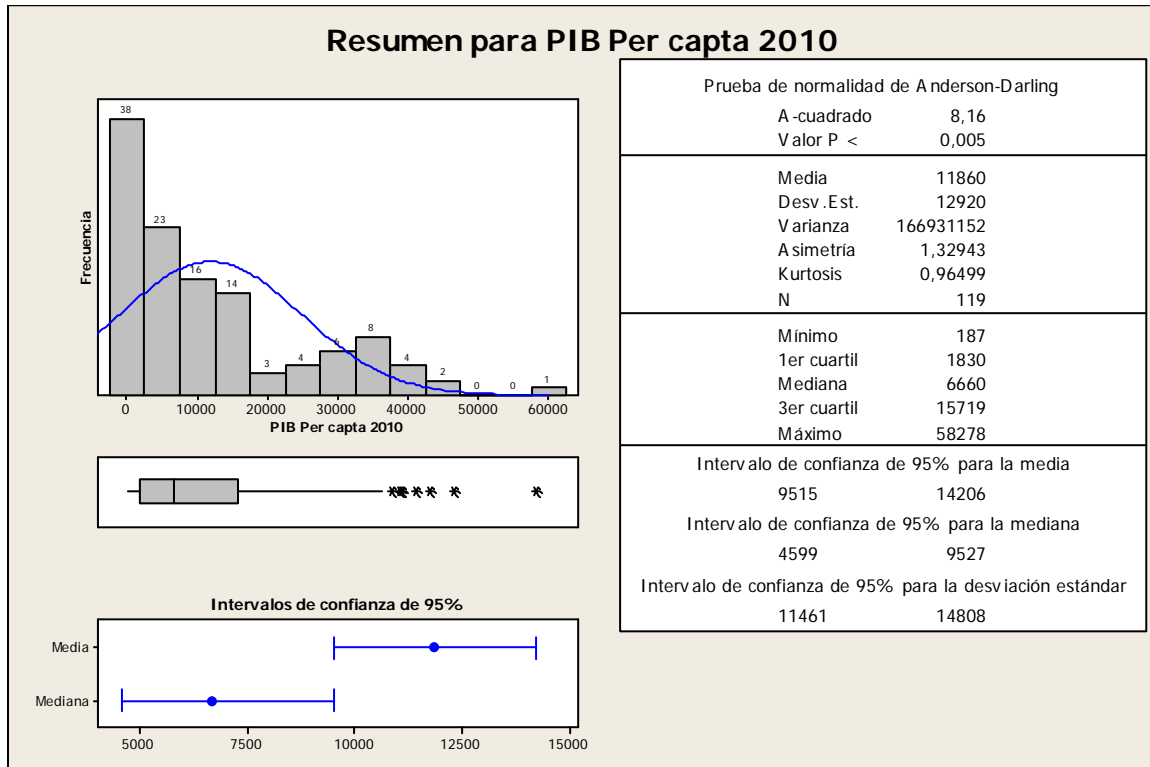
## A Tabela de Dados

Pais	População	Pegada Ecológica Total	Captações água por pessoa	Biocapacidade total	Reserva ecológica ou déficit	IDH 2003	PIB Per capita 2010
África do Sul	45	2,3	279	2	-0,3	0,66	10140
Albânia	3,2	1,4	544	0,9	-0,5	0,78	7737
Alemanha	82,5	4,5	571	1,7	-2,8	0,93	34743
Arábia Saudita	24,2	4,6	736	1	-3,7	0,77	24208
Argélia	31,8	1,6	194	0,7	-0,9	0,72	8477
Argentina	38,4	2,3	769	5,9	3,6	0,86	14930
Armênia	3,1	1,1	960	0,6	-0,5	0,76	5286
Áustria	8,1	4,9	260	3,4	-1,5	0,94	37602
Azerbaijão	8,4	1,7	2079	1,2	-0,5	0,73	9870
Bangladesh	146,7	0,5	552	0,3	-0,2	0,52	1458
Belarus	9,9	3,3	281	3,2	-0,1	0,79	13097
Bélgica/Luxemburgo	10,8	5,6	836	1,2	-4,4	0,95	34734
Benin	6,7	0,8	20	0,9	0,1	0,43	1506
Botsuana	1,8	1,6	110	4,5	3	0,57	13462
Brasil	178,5	2,1	336	9,9	7,8	0,79	10847
Bulgária	7,9	3,1	1318	2,1	-1	0,81	11547
Burkina Fasso	13	1	63	1	0	0,32	1217
Burundi	6,8	0,7	44	0,6	-0,1	0,38	403
Camarões	16	0,8	63	1,3	0,4	0,5	2219
Camboja	14,1	0,7	295	0,9	0,1	0,57	1952
Canadá	31,5	7,6	1470	14,5	6,9	0,95	39035
Cazaquistão	15,4	4	2263	4,1	0,1	0,76	12800
Chade	8,6	1	28	2,5	1,5	0,34	1339
Chile	15,8	2,3	804	5,4	3	0,85	14780
China	1311,7	1,6	484	0,8	-0,9	0,76	7206
Colômbia	44,2	1,3	246	3,6	2,3	0,79	8959
Congo	3,7	0,6	13	7,8	7,2	0,51	4583
Coreia do Sul	47,7	4,1	392	0,5	-3,5	0,9	29326
Costa do Marfim	16,6	0,7	57	1,2	1,2	0,42	1700
Costa Rica	4,2	2	655	1,5	-0,5	0,84	11143
Cuba	11,3	1,5	728	0,9	-0,7	0,82	9800
Dinamarca	5,4	5,8	237	3,5	-2,2	0,94	35736
Egito	71,9	1,4	969	0,5	-0,9	0,66	5840
El Salvador	6,5	1,4	200	0,6	-0,8	0,72	6660
Equador	13	1,5	1326	2,2	0,7	0,76	8170
Espanha	41,1	5,4	870	1,7	-3,6	0,93	30475
Estônia	1,3	6,5	118	5,7	-0,7	0,85	18355
EUA	294	9,6	1647	4,7	-4,8	0,94	46653
Filipinas	80	1,1	363	0,6	-0,5	0,76	3600
Finlândia	5,2	7,6	476	12	4,4	0,94	34107
França	60,1	5,6	668	3	-2,6	0,94	33103
Gâmbia	1,4	1,4	22	0,8	-0,5	0,47	1446
Gana	20,9	1	48	1,3	0,3	0,52	1533
Geórgia	5,1	0,8	697	1,2	0,5	0,73	4946
Grécia	11	5	708	1,4	-3,6	0,91	28608
Guatemala	12,3	1,3	167	1,3	0	0,66	4760
Guiné-Bissau	1,5	0,7	121	2,9	2,2	0,35	554
Haiti	8,3	0,6	120	0,3	-0,3	0,48	1040
Holanda	16,1	4,4	494	0,8	-3,6	0,94	41004
Honduras	6,9	1,3	127	1,8	0,5	0,67	3845
Hungria	9,9	3,5	770	2	-1,5	0,86	18986
Iêmen	20	0,8	343	0,4	-0,5	0,49	2600
Ilhas Maurício	1,2	1,9	504	1,2	-0,7	0,48	1310
Índia	1065,5	0,8	615	0,4	-0,4	0,6	3354
Indonésia	219,9	1,1	381	1	0	0,7	4394
Irã	68,9	2,4	1071	0,8	-1,6	0,74	11890
Irlanda	4	5	289	4,8	-0,2	0,95	38768
Israel	6,4	4,6	325	0,4	-4,2	0,92	28292
Itália	57,4	4,2	772	1	-3,1	0,93	30080
Jamaica	2,7	1,7	156	0,5	-1,3	0,74	7547
Japão	127,7	4,4	694	0,7	-3,6	0,94	33649
Jordânia	5,5	1,8	190	0,3	-1,5	0,75	570
Laos	5,7	0,9	543	1,3	0,4	0,55	2404
Lesoto	1,8	0,8	28	1,1	0,3	0,5	1605
Letônia	2,3	2,6	129	6,6	4	0,84	16300
Líbano	3,7	2,9	384	0,3	-2,6	0,76	13510
Lituânia	3,4	4,4	78	4,2	-0,2	0,85	15327
Madagascar	17,4	0,7	884	2,9	2,2	0,5	958
Malásia	24,4	2,2	376	3,7	1,5	0,8	14410
Malauí	12,1	0,6	85	0,5	-0,1	0,4	902
Mali	13	0,8	519	1,3	0,5	0,75	1207
Marrocos	30,6	0,9	419	0,8	-0,1	0,63	4638
Mauritânia	2,9	1,3	606	5,8	4,5	0,33	2037
México	103,5	2,6	767	1,7	-0,9	0,81	14192
Mianmar	49,5	0,9	680	1,3	0,4	0,58	1596
Moçambique	18,9	0,6	34	2,1	1,4	0,38	929
Moldávia	4,3	1,3	541	0,8	-0,5	0,67	2917
Mongólia	2,6	3,1	172	11,8	8,7	0,7	3710
Namíbia	2	1,1	153	4,4	3,3	0,63	6474
Nepal	25,2	0,7	414	0,5	-0,2	0,53	1189
Nicarágua	5,5	1,2	244	3,5	2,4	0,69	2632
Nigéria	124	1,2	66	0,9	-0,2	0,45	2289
Noruega	4,5	5,8	485	6,8	0,9	0,96	58278
Nova Zelândia	3,9	5,9	549	14,9	9	0,93	27520
Panamá	3,1	1,9	268	2,5	0,6	0,8	1321
Papua-Nova Guiné	5,7	2,4	13	2,1	-0,3	0,52	2395
Paquistão	153,6	0,6	1130	0,3	-0,3	0,53	2625
Paraguai	5,9	1,6	85	5,6	4	0,76	4629
Peru	27,2	0,9	752	3,8	3	0,76	9016
Polónia	38,6	3,3	419	1,8	-1,4	0,86	18406
Portugal	10,1	4,2	1121	1,6	-2,6	0,9	23019
Quirguistão	5,1	1,3	1989	1,4	0,1	0,7	2100
Reino Unido	59,5	5,6	161	1,6	-4	0,94	34342
República Democrática do Congo	52,8	0,6	7	1,5	0,9	0,39	327
República Dominicana	8,7	1,6	393	0,8	-0,8	0,75	8616
República Tcheca	10,2	4,9	252	2,6	-2,3	0,87	24419
Romênia	22,3	2,4	1035	2,3	-0,1	0,77	12910
Ruanda	8,4	0,7	18	0,5	-0,1	0,45	1102
Rússia	143,2	4,4	532	6,9	2,5	0,8	15719
Senegal	10,1	1,2	225	0,9	-0,3	0,46	1830
Serra Leoa	5	0,7	80	1,1	0,4	0,3	825
Síria	17,8	1,7	1148	0,8	-0,9	0,72	4857
Sri Lanka	19,1	1	667	0,4	-0,6	0,75	4999
Sudão	33,6	1	1135	1,8	0,8	0,51	230
Suécia	8,9	6,1	334	9,6	3,5	0,95	36139
Suíça	7,2	5,1	358	1,5	-3,6	0,95	43109
Tadjiquistão	6,2	0,6	1931	0,5	-0,1	0,65	2065
Tailândia	62,8	1,4	1400	1	-0,4	0,78	8328
Tanzânia	37	0,7	143	1,3	0,6	0,42	1426
Togo	4,9	0,9	35	0,8	-0,1	0,51	846
Trinidad e Tobago	1,3	3,1	239	0,4	-2,7	0,8	25162
Turcomenistão	4,9	3,5	5142	3,6	0,1	0,74	7627
Turquia	71,3	2,1	534	1,4	-0,7	0,75	13359
Ucrânia	48,5	3,2	767	1,7	-1,5	0,77	6590
Uganda	25,8	1,1	12	0,8	-0,2	0,51	1250
Uruguai	3,4	1,9	929	8	6,1	0,84	14022
Vietnã	81,4	0,9	889	0,8	-0,1	0,7	3097
Zâmbia	10,8	0,6	163	3,4	2,8	0,39	1497
Zimbábue	12,9	0,9	328	0,8	-0,1	0,5	187

### 3. ANÁLISE EXPLORATÓRIA

#### 3.1. ANÁLISE DAS VARIÁVEIS

##### 3.1.1. PIB Per Capta



Os dados de PIB Per Capta referem-se ao ano de 2010.

**Quanto à forma:** A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 1,37500, onde observamos dois blocos: No primeiro, representando os países mais pobres, estão 91 países com PIB per capita entre PPP US\$ 2.500 até PPP US\$ 17.500,00. No segundo temos os 28 países mais ricos com PIB per capita entre acima de PPP US\$ 17.500,00. No Brasil, esse valor é de PPP US\$ 10.847,00 para o ano de 2010.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor  $P = 0,005$  indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados têm PIB per capita até PPP US\$ 6.660,00 (mediana) e na outra metade o valor do PIB per capita está acima dessa medida. Verifica-se uma distância muito grande entre a média (PPP US\$ 11.860,00) e a mediana (PPP US\$ 6.660,00), o que significa que existem grandes distorções da média. Isso pode ser constatado pelo número de oito outliers observados, que representam os países mais ricos, a saber: Canadá, Irlanda, Suécia, Áustria, Holanda, Suíça, EUA e Noruega. Este último apresenta o maior PIB per capita, no valor de PPP US\$ 58.278,00.

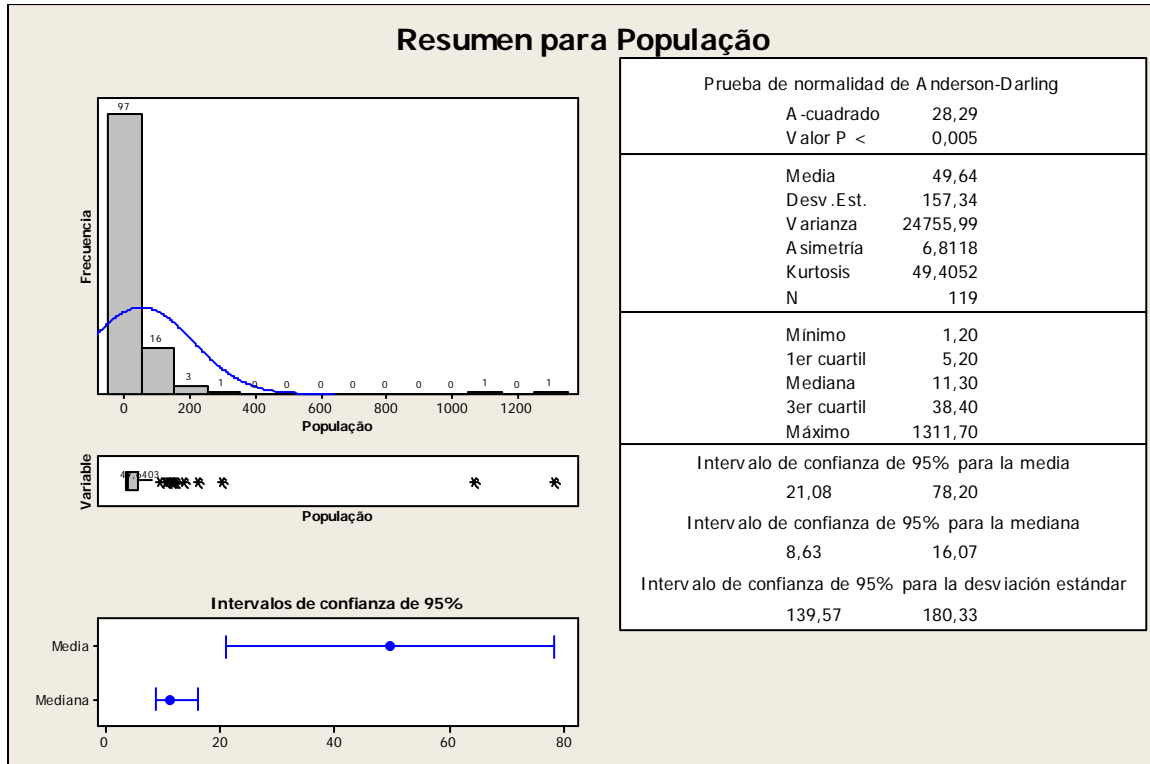
**Quanto à Dispersão:** pesquisa apresenta grande amplitude, na medida em que o menor valor de PIB per capita é de PPP US\$ 187,00 e o maior valor de PIB per capita é de PPP US\$ 58.278,00.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média entre PPP US\$ 9.515,00 e PPP US\$ 14.206,00.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor médio (PPP US\$ 11.860,00), podemos indicar um grau de dispersão forte, pois o desvio padrão (12920) é maior que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm um PIB per capita abaixo de PPP US\$ 1.830,00 e outros 25% têm um PIB per capita acima de PPP US\$ 15.719,00, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 3.1.2. População



Os dados de População referem-se ao ano de 2006.

**Quanto à forma:** A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 6,817, onde observamos três blocos: No primeiro, representando os países com menor população, estão 97 países com população até 50 milhões de pessoas. No segundo temos os 16 países com população variando entre 50 milhões e 150 milhões de pessoas e no terceiro bloco, temos 6 países com uma população acima de 150 milhões de pessoas. No Brasil, a população era de 178 milhões de pessoas para o ano de 2006.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor  $P = 0,005$  indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados têm população até 11,30 milhões de pessoas (mediana) e na outra a população está acima dessa medida. Verifica-se uma distância muito grande entre a média da população (49,64 milhões de pessoas) e a mediana (11,30 milhões de pessoas), o que significa que existem grandes distorções da média. Isso pode ser constatado pelo número de dez outliers observados, que representam os países mais populosos, a saber: México, Bangladesh, Nigéria, Rússia, Paquistão, Botsuwana, Indonésia, EUA, Índia e China. Este último apresenta a maior população, com 1.311,7 milhões de pessoas.

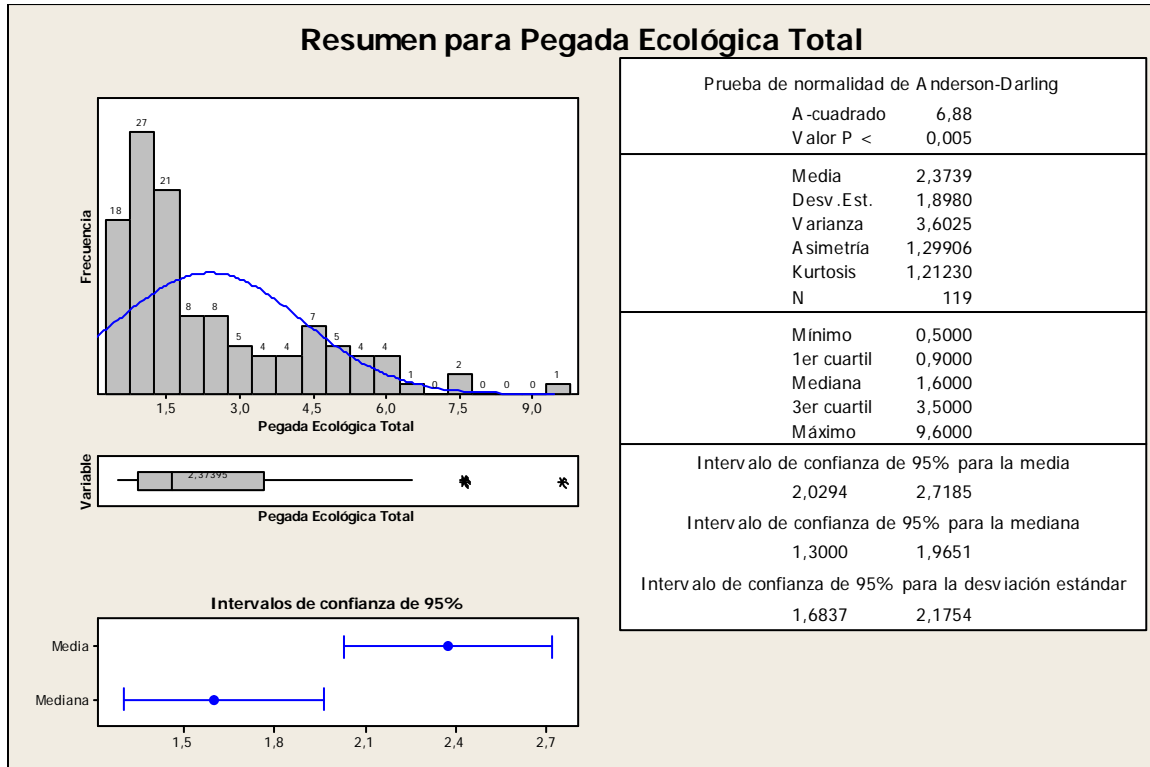
Quanto à Dispersão: pesquisa apresenta grande amplitude, na medida em que a menor população é de 1,2 milhões de pessoas (Ilhas Maurício) e a maior população com 1.311,7 milhões de pessoas.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média entre 21,08 e 78,20 milhões de pessoas.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor médio (49,64 milhões de pessoas), podemos indicar um alto grau de dispersão, pois o desvio padrão (157,34) é maior que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm uma população abaixo 5,20 milhões de pessoas e outros 25% têm uma população acima de 38,40 milhões de pessoas, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 1.1.1. Pegada Ecológica Total



Os dados de Pegada Ecológica Total referem-se ao ano de 2006.

**Quanto à forma:** A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 1,29906, onde observamos três blocos: No primeiro, representando os países com menor pegada ecológica, estão 66 países com Pegada Ecológica até 1,75 milhões de hectares globais por pessoa. No segundo temos os 49 países com pegada ecológica entre de 1,75 e 6,25 milhões de hectares globais por pessoa e no terceiro, temos os 4 países com maior pegada ecológica, com até 9,75 milhões de hectares globais por pessoa. No Brasil, esse valor é de 2,1 milhões de hectares globais por pessoa, para o ano de 2006.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor  $P = 0,005$  indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados têm pegada ecológica até 1,60 (mediana) e na outra metade o valor da pegada ecológica está acima dessa medida. Verifica-se uma distância muito grande entre a média (2,3739) e a mediana (1,60), o que significa que existem grandes distorções da média. Isso pode ser observado pelo número de 2 outliers observados, que representam os países com maior pegada ecológica, a saber: Canadá e EUA. Este último apresenta o maior valor de pegada ecológica total, igual a 7,6.

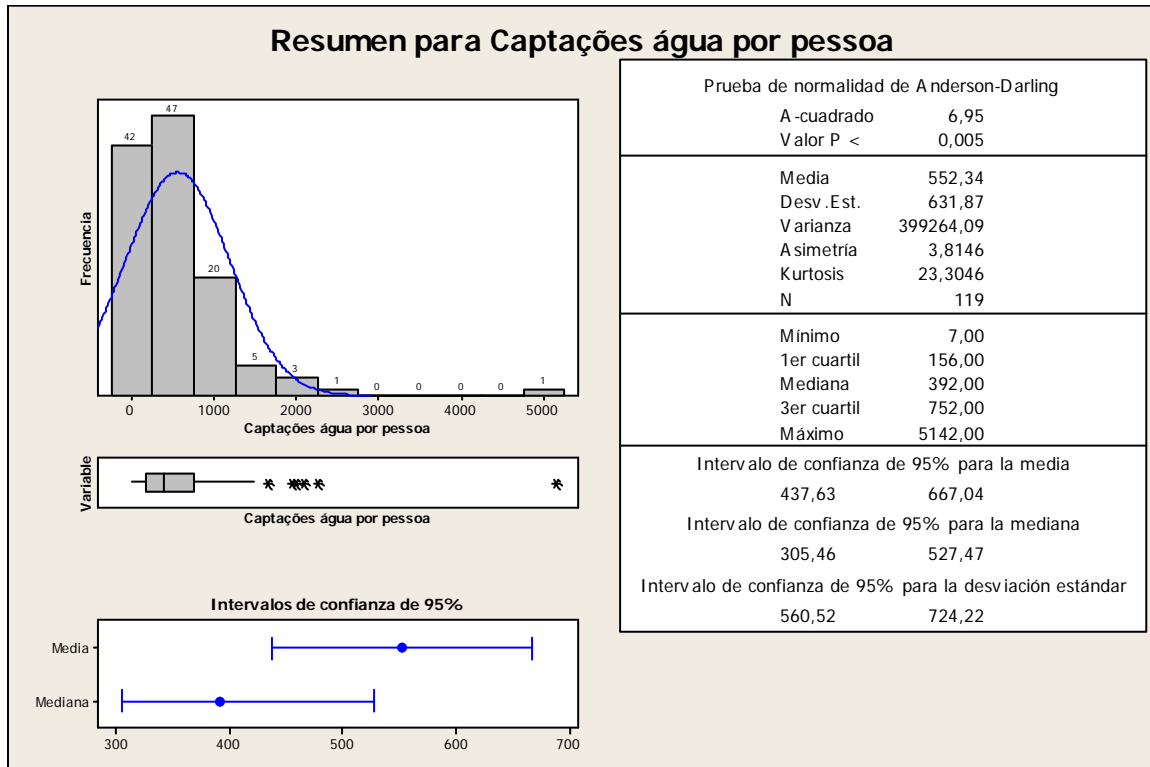
**Quanto à Dispersão:** pesquisa apresenta grande amplitude, na medida em que o menor valor de pegada ecológica é igual a 0,5 hectares globais por pessoa e o maior valor de pegada ecológica é igual a 7,6 hectares globais por pessoa.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média da pegada ecológica entre 2,0294 e 2,7185 hectares globais por pessoa.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor médio (2,3739), podemos indicar um grau de dispersão forte, pois o desvio padrão (1,8980) é maior que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm valor de pegada ecológica abaixo de 0,90 hectares globais por pessoa e outros 25% têm um valor de pegada ecológica acima de 3,50 hectares globais por pessoa, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 1.1.2. Captações água por pessoa



Os dados de Captação de Água por pessoa referem-se ao ano de 2006.

**Quanto à forma:** A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 3,8146, onde observamos dois blocos: No primeiro, representando os países com menor captação de água por pessoa, estão 89 países com Captação Global de Água por pessoa até 750 mil m<sup>3</sup> por ano. No segundo, estão 29 países com captação de água por pessoa entre 750 e 2750 mil m<sup>3</sup> por ano, e por último está 1 país com Captação Global de Água por pessoa de até 5.250 mil m<sup>3</sup> por ano. No Brasil, esse valor é de 336 mil m<sup>3</sup> para o ano de 2006.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor P = 0,005 indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados o valor de captação anual de água é de 392 mil m<sup>3</sup> (mediana) e na outra metade o valor de captação anual de água está acima dessa medida. Verifica-se uma distância muito grande entre a média (552,34) e a mediana (392 mil m<sup>3</sup>), o que significa que existem grandes distorções da média. Isso pode ser observado pelo número de seis outliers observados, que representam os países com maior captação de água por pessoa, a saber: EUA, Tadjiquistão, Quirgistão, Cazaquistão, Arzeibaijão e Turcomenistão. Este último apresenta o maior valor para Captação de Água por pessoa (5.142 mil m<sup>3</sup>) por pessoa.

Quanto à Dispersão: pesquisa apresenta grande amplitude, na medida em que o menor valor captação de água por pessoa é de 7 mil m<sup>3</sup> (República Democrática do Congo) e o maior é com 5.142 mil m<sup>3</sup> por pessoa.

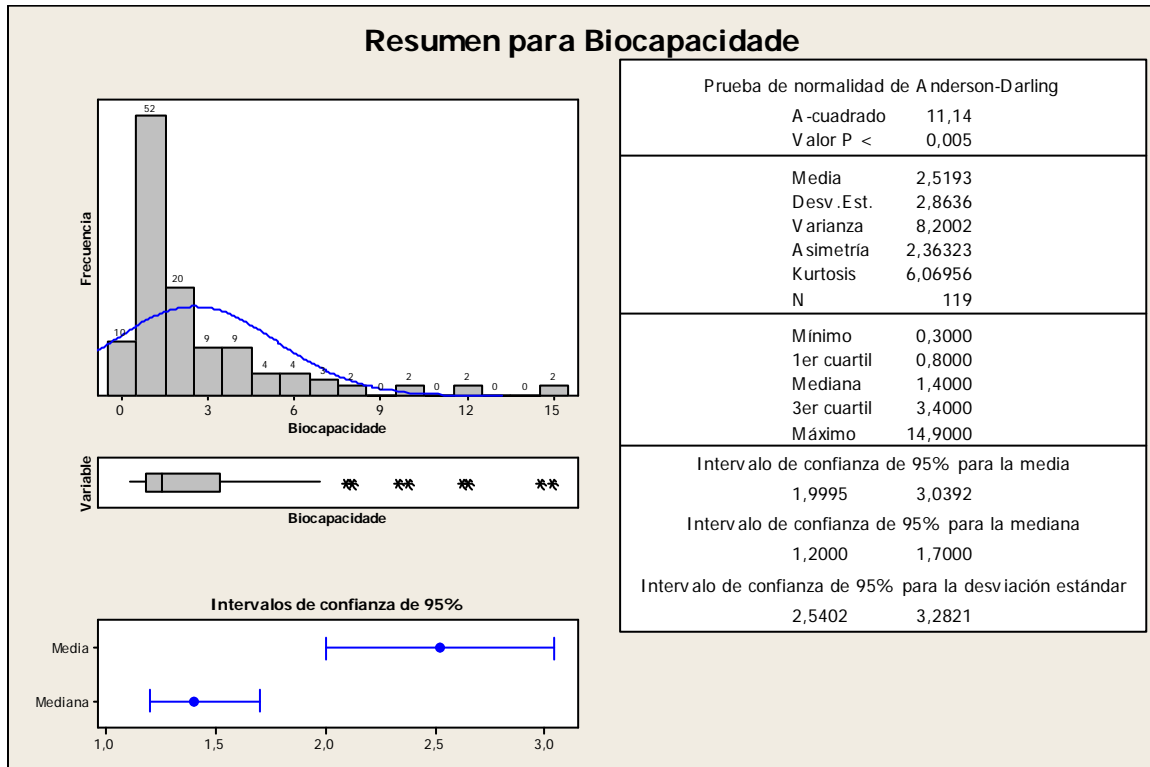
Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média entre 437,63 e 667,04 mil m<sup>3</sup> por pessoa.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor média (552,34), podemos indicar um grau de dispersão forte, pois o desvio padrão (631,87) é maior que 10%.



Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm um valor de captação de água por pessoa abaixo de 156 mil m<sup>3</sup> por pessoa outros 25% têm valor de captação de água por pessoa acima de 752 mil m<sup>3</sup> por pessoa, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 1.1.3. Biocapacidade total



Os dados de Biocapacidade total referem-se ao ano de 2006.

Quanto à forma: A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 2,36323, onde observamos dois blocos: No primeiro, 82 países com r biocapacidade total de até 2,5 hectares globais por pessoa. No segundo temos os 37 com biocapacidade total de até 15,5 hectares globais por pessoa. No Brasil, esse valor é de 9,9 hectares globais por pessoa.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor P = 0,005 indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

Quanto ao centro e dispersão: Pelo histograma, observa-se que na metade dos países estudados têm biocapacidade total de até 1,4 hectares globais por pessoa (mediana) e na outra metade o valor biocapacidade total está acima dessa medida. Verifica-se uma distância muito grande entre a média (2,5193 hectares globais por pessoa) e a mediana (1,4 hectares globais por pessoa), o que significa que existem grandes distorções da média. Isso pode ser observado pelo número de oito outliers observados, que representam os países com maior biocapacidade total, a saber: Congo, Uruguai, Suécia, Brasil, Mongolia, Finlândia, Canadá e Nova Zelândia. Este último apresenta o maior valor de biocapacidade total = 14,9 hectares globais por pessoa.

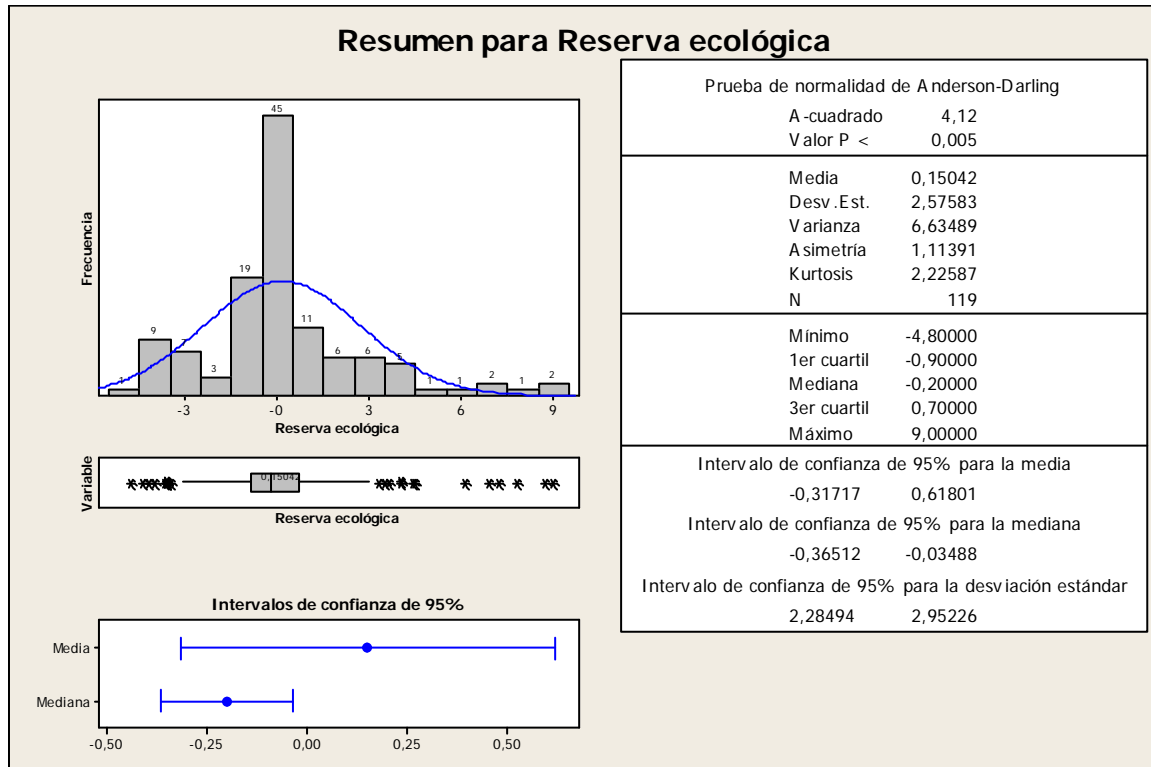
Quanto à Dispersão: pesquisa apresenta grande amplitude, na medida em que o menor valor de biocapacidade total é de Haiti (0,3 hectares globais por pessoa) e o maior valor de biocapacidade total é a Nova Zelândia com 14,9 hectares globais por pessoa.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média de biocapacidade total entre 1,9995 e 3,0392 hectares globais por pessoa.

Quanto ao desvio padrão, que mede o grau de dispersão em torno do valor médio (2,5193), podemos indicar um grau de dispersão baixo, pois o desvio padrão (2,8636) é menor que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm um valor de biocapacidade total abaixo de 0,80 e outros 25% têm um valor de biocapacidade total acima de 3,40 hectares globais por pessoa, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 1.1.4. Reserva ecológica ou déficit



Os dados de reserva ou déficit ecológicos são do ano de 2006..

**Quanto à forma:** A distribuição encontra-se assimétrica positiva à direita, com coeficiente de assimetria igual à 1,11391, onde observamos três blocos: No primeiro, estão 20 países com déficit ecológico de até -1,5 hectares globais por pessoa. No segundo, estão 64 países com reserva ecológica de -1,5 hectares globais por pessoa até 1,5 hectares globais por pessoa |Por último, estão 35 países com reserva ecológica maior do que 1,5 hectares globais por pessoa. No Brasil, esse valor é de 15 hectares globais por pessoa, para o ano de 2006.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor  $P = 0,005$  indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados têm reserva ou déficit ecológicos a até -0,20 hectares globais por pessoa (mediana) e outra metade têm reserva ou déficit ecológicos acima dessa medida. Verifica-se uma distância muito grande entre a média (15,042) e a mediana (-0,20 hectares globais por pessoa), o que significa que existem grandes distorções da média. Isso pode ser observado pelos vários outliers observados, a saber: EUA (-4,8), Bélgica, Israel, Reino Unido, Arábia Saudita, Espanha, Coréia do Sul, Namíbia, Suécia, Letônia, Finlândia, Uganda, Canadá, Congo, Brasil(7,8), Mongólia e Nova Zelândia (9,0).

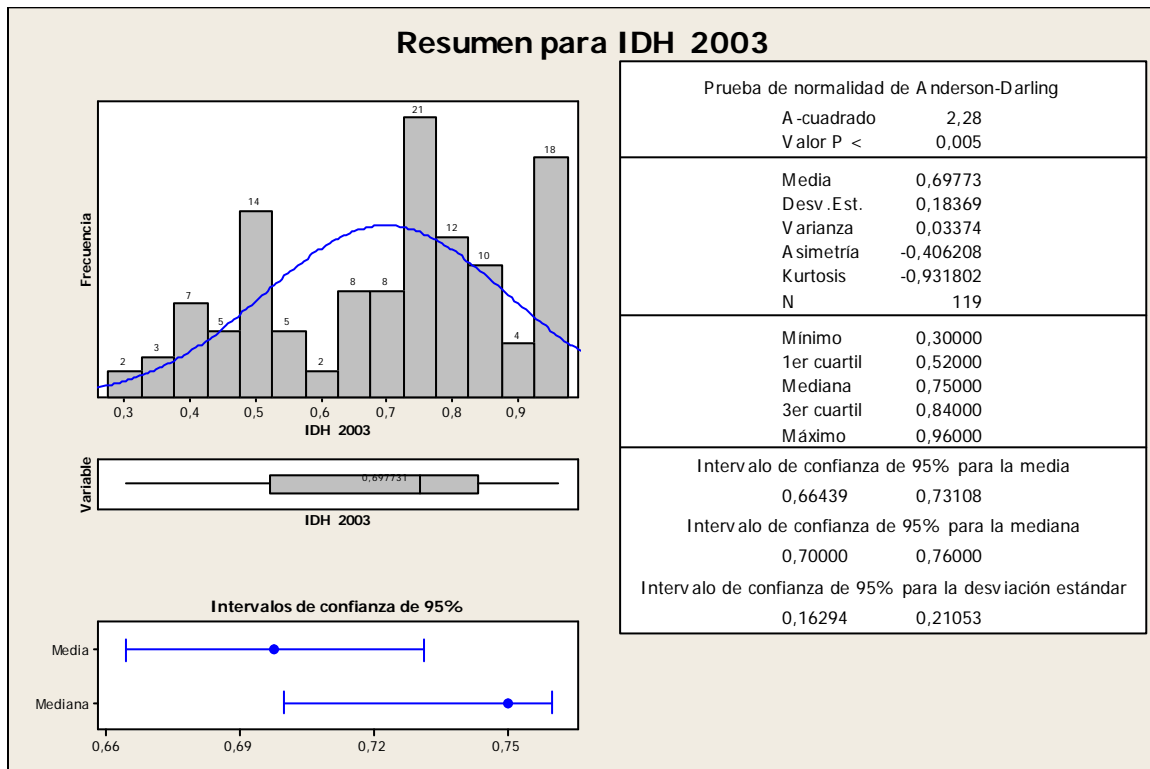
**Quanto à Dispersão:** pesquisa apresenta grande amplitude, na medida em que o menor valor Déficit ecológico é dos EUA (-4,8) e o maior valor de Reserva Ecológica é da Nova Zelândia (9,0) hectares globais por pessoa.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média entre -0,31717 e 0,61801 hectares globais por pessoa.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor médio (0,15042), podemos indicar um baixo grau de dispersão, pois o desvio padrão (2,57583) é menor que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm déficit ecológico abaixo -0,90 e outros 25% reserva ecológica acima 0,70, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

### 1.1.5. IDH



Os dados de IDH referem-se ao ano de 2006.

**Quanto à forma:** A distribuição encontra-se assimétrica negativa à esquerda, com coeficiente de assimetria igual à -0,406208, onde observamos dois blocos: No primeiro, representando os países mais pobres, estão 38 países com IDH até 0,625. No segundo temos os 81 países mais ricos com até 0,975. No Brasil, esse valor é de 0,79 para o ano de 2006.

No teste de Anderson-Darling, o valor de P representa um índice decrescente da confiabilidade de um resultado obtido. No caso estudado, o valor  $P = 0,005$  indica que existe forte evidência de que os dados podem não vir de uma distribuição Normal, pois o limite aceitável de erro é de 0,05.

**Quanto ao centro e dispersão:** Pelo histograma, observa-se que na metade dos países estudados têm um IDH até 0,75 (mediana) e na outra metade o valor IDH está acima dessa medida. Verifica-se que não existe uma distância muito grande entre a média (0,69773) e a mediana (0,75), o que significa que não existem grandes distorções da média. Não foram observados outliers.

**Quanto à Dispersão:** pesquisa apresenta pequena amplitude, na medida em que o menor valor do IDH é de 0,3 e o maior valor de IDH é de 0,96.

Considerando a probabilidade de ocorrência, pode-se dizer que com 95% de confiança existe a probabilidade de uma observação assumir um valor para a média entre 0,66439 e 0,73108.

**Quanto ao desvio padrão,** que mede o grau de dispersão em torno do valor médio 0,69773, podemos indicar um baixo grau de dispersão, pois o desvio padrão (0,18369) é menor que 10%.

Pelo gráfico **Boxplot** abaixo, identificamos que 25 % dos países estudados têm um IDH abaixo de 0,52 e outros 25% têm um IDH acima de 0,84, sendo que 50% dos países estudada encontram-se dentro dessa faixa.

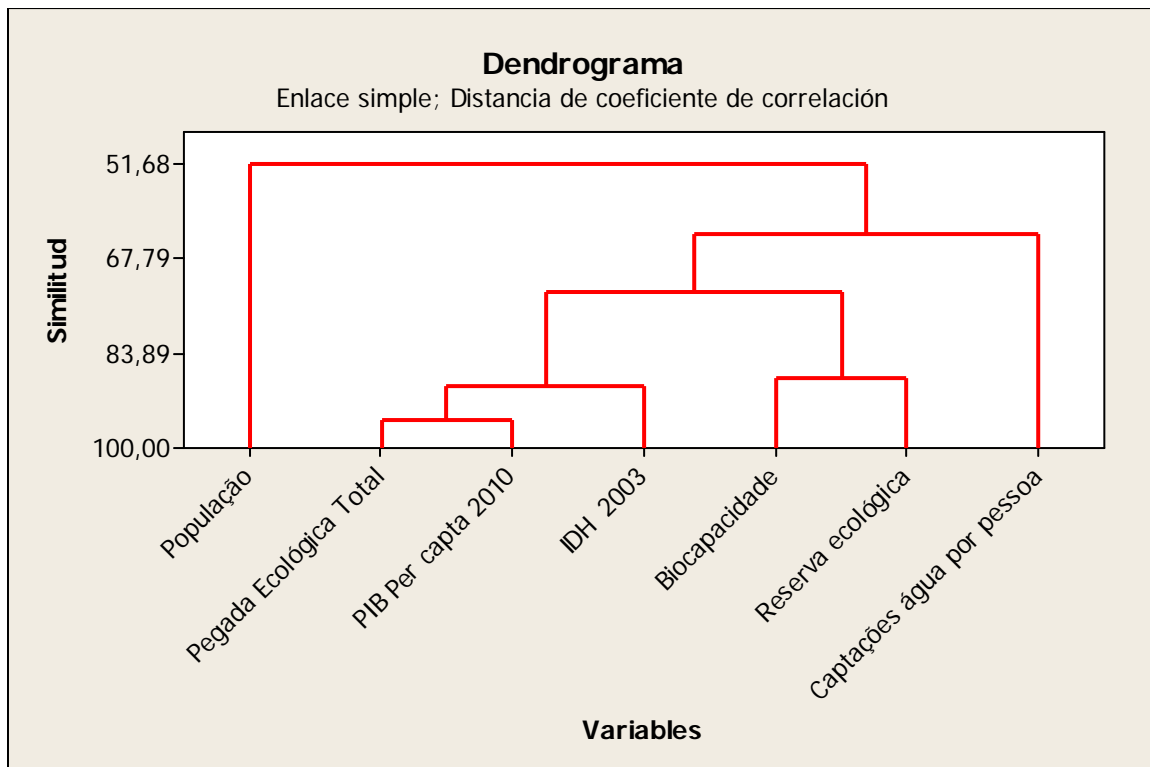
## 1. MATRIZ DE CORRELAÇÃO

**Correlaciones: Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010**

	População	Pegada Ecológica	Captações água p	Biocapacidade	Reserva ecológic	IDH
Pegada Ecológica	-0,026 0,777					
Captações água p	0,034 0,716	0,190 0,039				
Biocapacidade	-0,087 0,349	0,473 0,000	0,048 0,605			
Reserva ecológic	-0,079 0,396	-0,208 0,023	-0,086 0,350	<b>0,763</b> 0,000		
IDH	0,029 0,756	<b>0,760</b> 0,000	0,272 0,003	0,308 0,001	-0,216 0,018	
PIB Per capta	-0,020 0,827	<b>0,902</b> 0,000	0,112 0,226	0,353 0,000	-0,270 0,003	<b>0,792</b> 0,000

Contenido de la celda: Correlación de Pearson  
Valor P

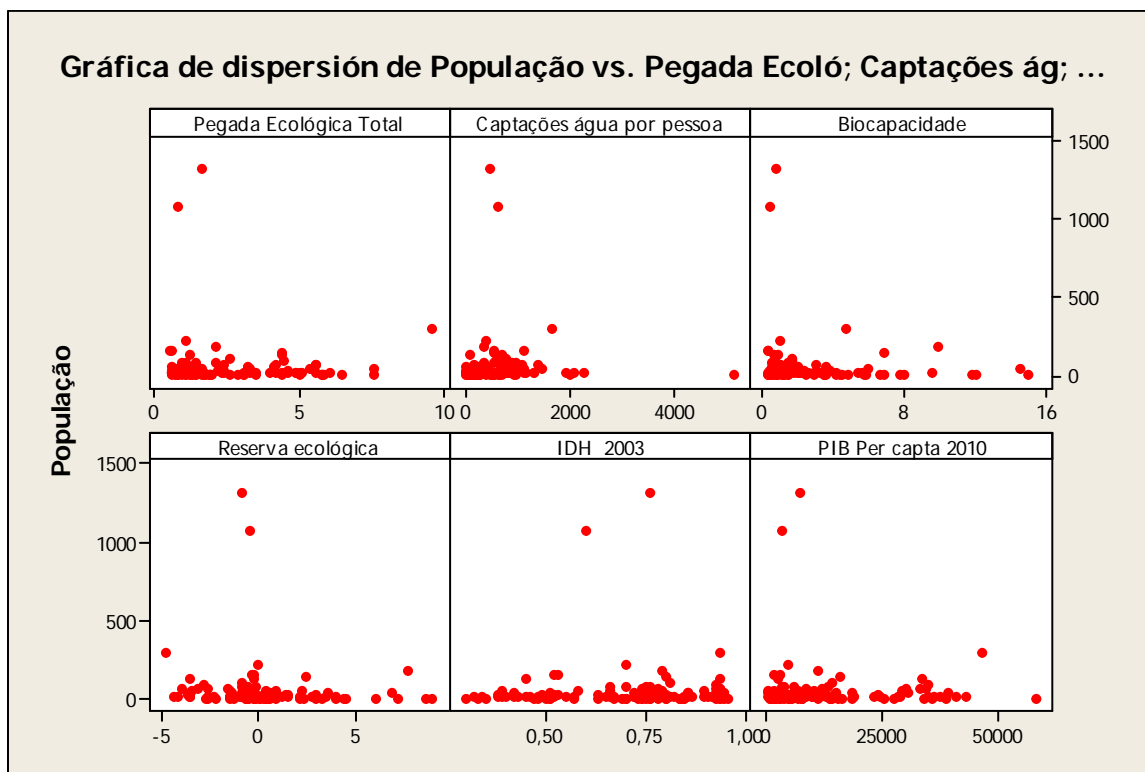
## 2. DENDOGRAMA



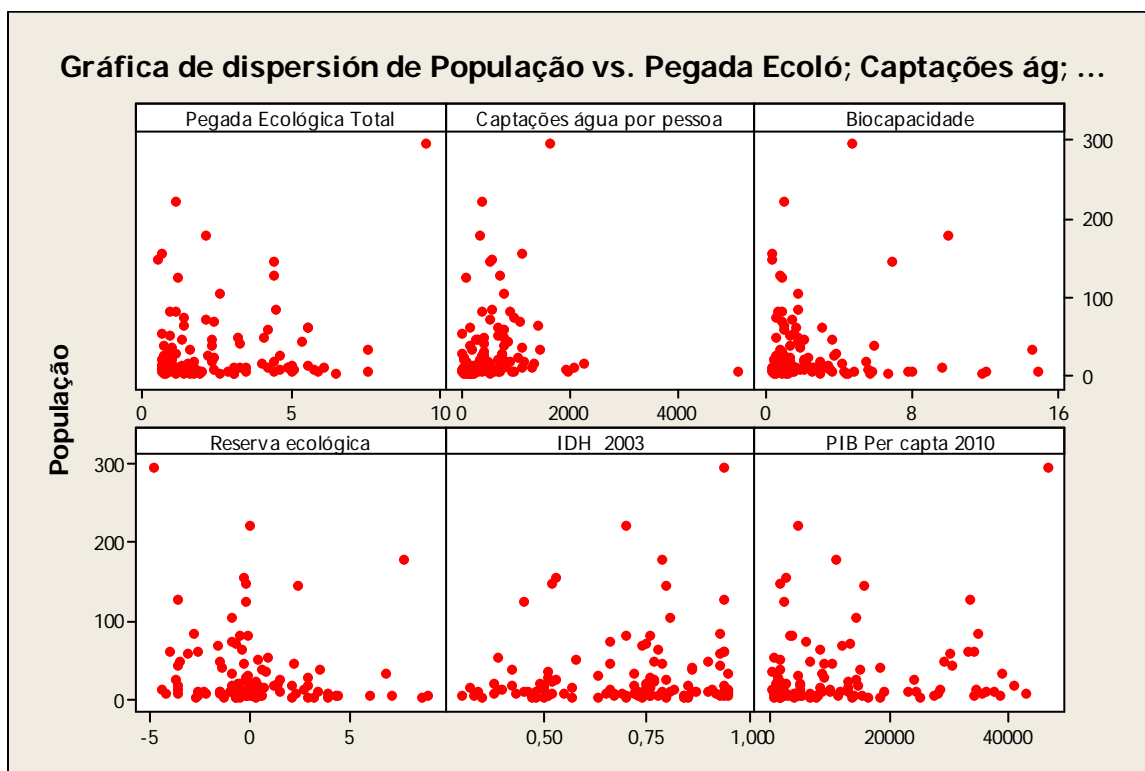
O gráfico acima refere-se ao universo total dos países incluídos do Relatório-Planeta-Vivo-2006 (<http://www.wwf.org.br/informacoes/biblioteca/?4420>). Nesse gráfico, identificamos quatro blocos de maior similaridade: **Bloco 1:** População. **Bloco 2:** Pegada Ecológica Total, PIB Per Capta e IDH. **Bloco 3:** Biocapacidade, Reserva Ecológica. e **Bloco 4:** Captações de água por pessoa, que confirmam os resultados obtidos na matriz de correlação.

### 3. GRAFICOS DE DISPERSÃO

#### 3.1. Análise População

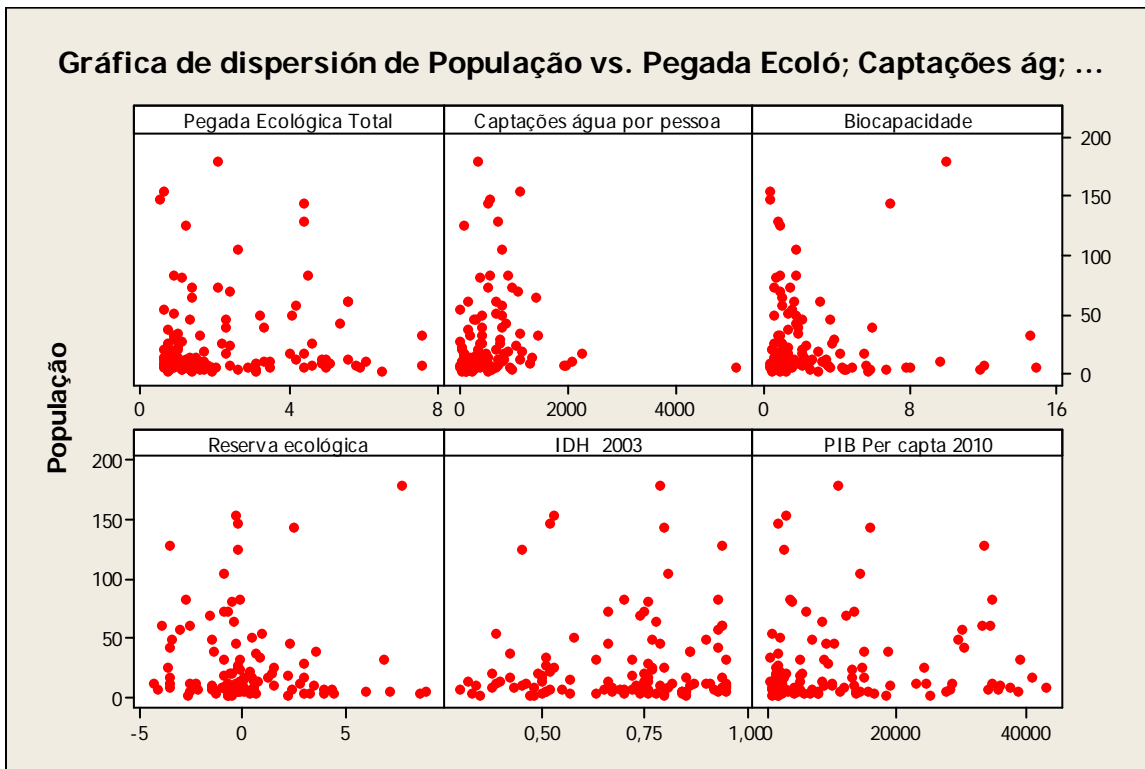


Os graficos de dispersão acima confirmam a tabela de correlações. Podemos verificar visualmente que não aparecem correlações significativas entre as variáveis estudadas: População x (Pegada Ecológica Total, Captações água por pessoa, Biocapacidade, Reserva ecológica, IDH e PIB Per capita). Destacam-se dois outliers: China e Índia.

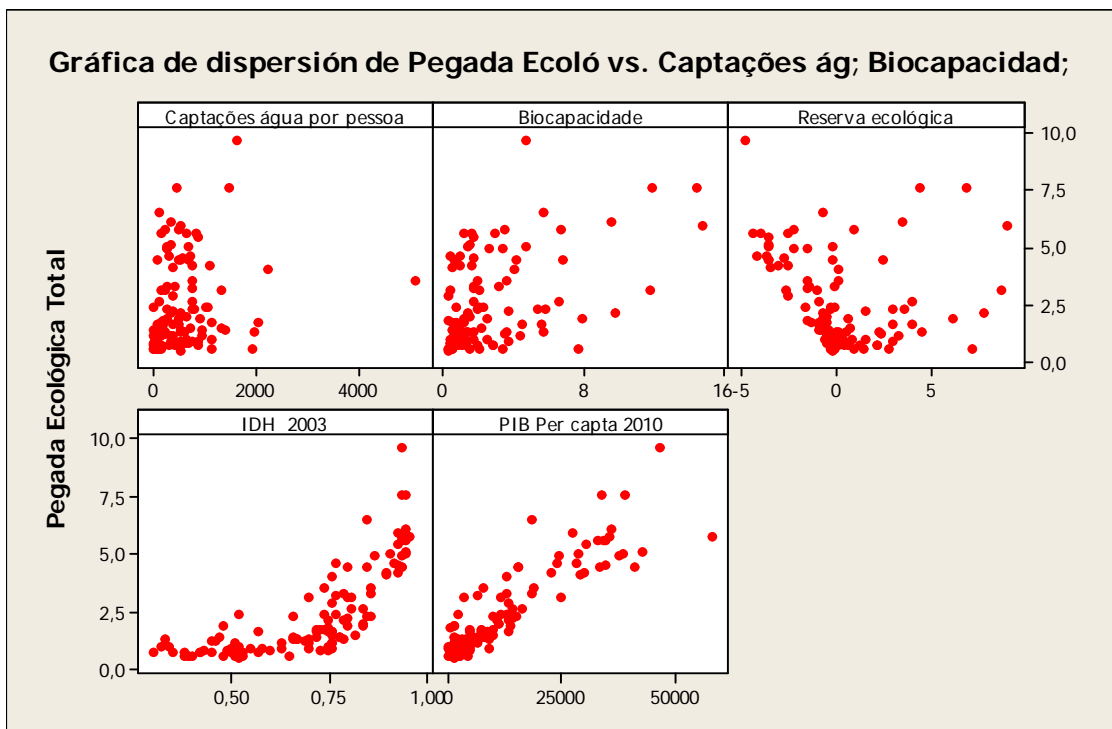


Excluindo-se Índia e China dos dados, vemos o gráfico abaixo, com o destaque para os EUA, e Indonésia.

Mesmo excluindo-se esses dois países, EUA (294 milhões de pessoas) e Indonésia (População = 219 milhões de pessoas), vemos pelo gráfico abaixo, que ainda não apresenta correlação entre população e as outras variáveis estudadas.

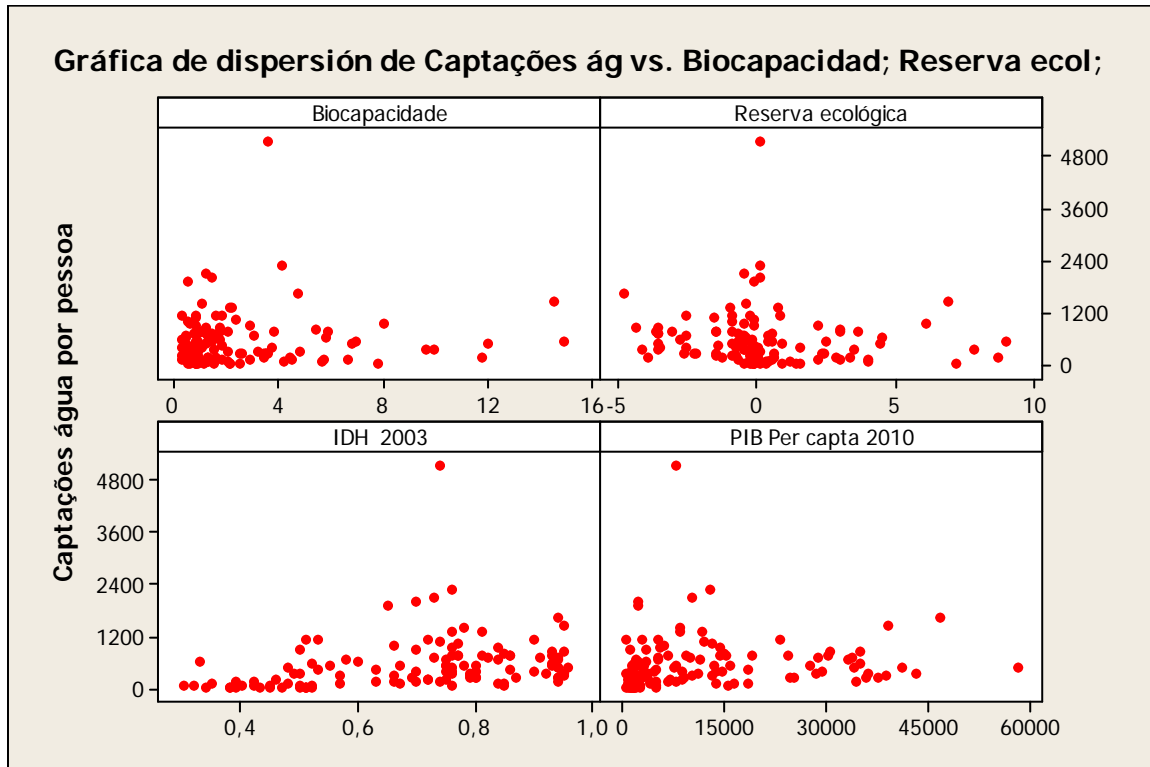


### 3.2. Análise Pegada Ecológica



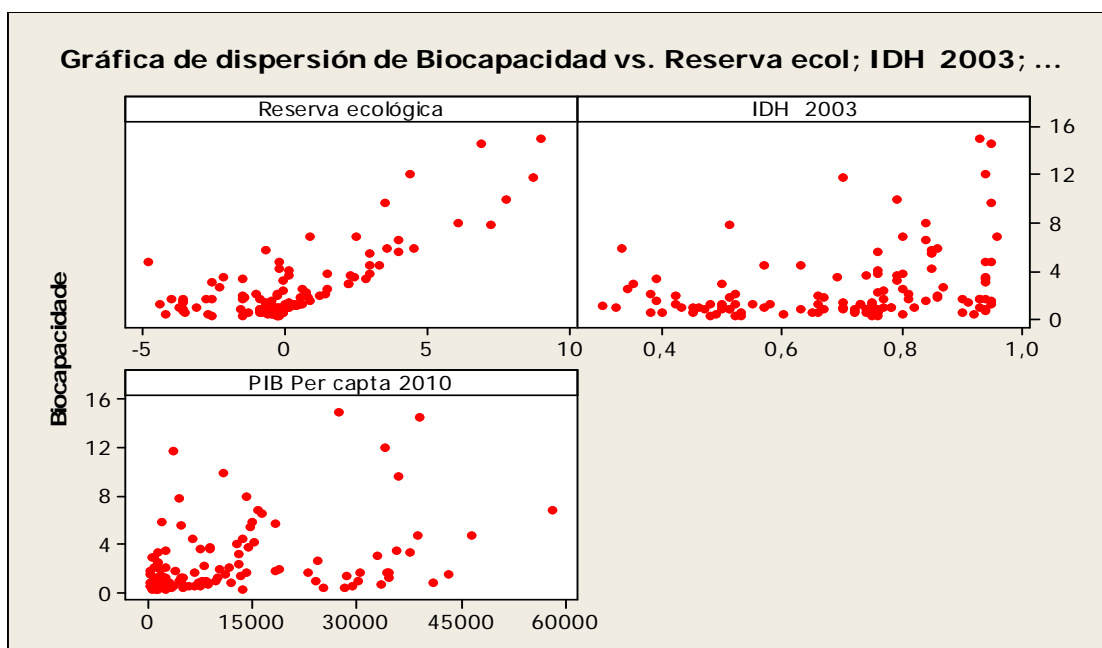
Novamente, confirmando os resultados da matriz de correlações, vemos visualmente nos gráficos de dispersão acima, uma forte correlação entre Pegada Ecológica e (IDH e PIB Per Capta). As outras variáveis apresentam baixa correlação com Pegada Ecológica.

### 3.3. Análise Captações Água por pessoa



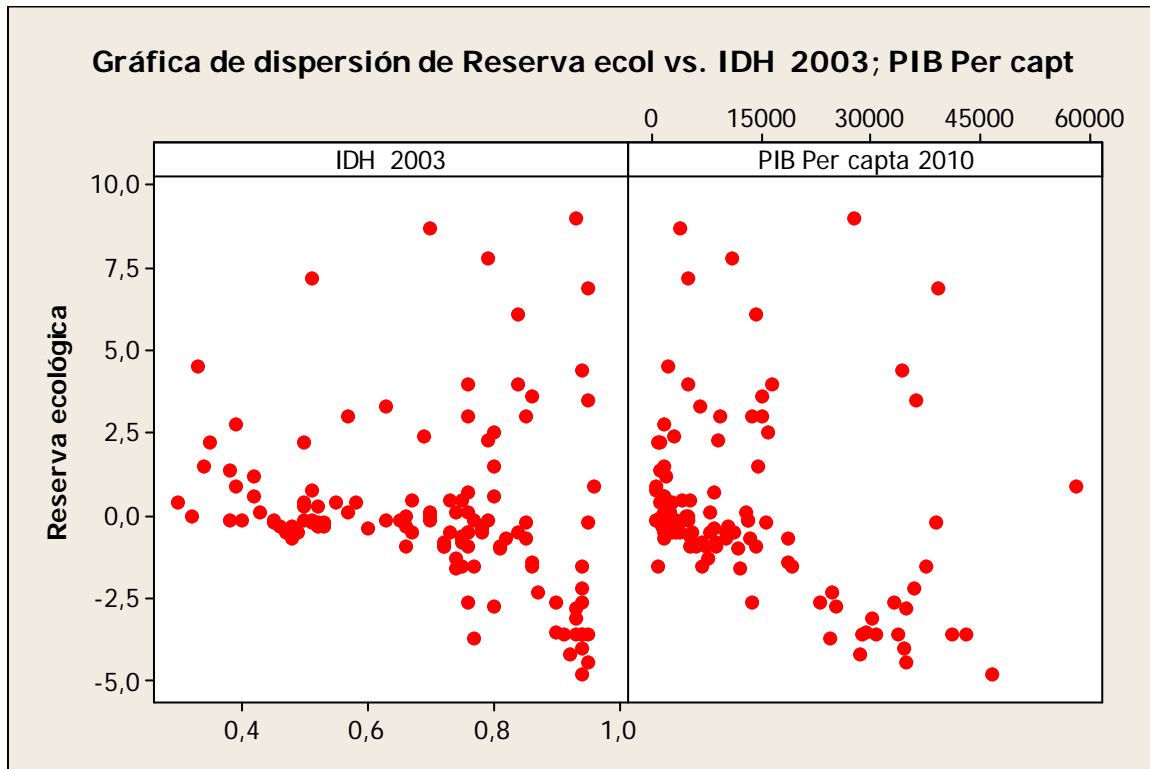
Novamente, os graficos de dispersão acima confirmam a tabela de correlações. Podemos verificar visualmente que não aparecem correlações significativas entre as variáveis estudadas: Captações água por pessoa x (Biocapacidade, Reserva ecológica, IDH e PIB Per capta). Destaca-se um outliers: Turcomenistão.

### 3.4. Análise Biocapacidade



Novamente, confirmando os resultados da matriz de correlações, vemos visualmente nos gráficos de dispersão acima, uma forte correlação entre Reserva Ecológica e Biocapacidade. As outras variáveis analisadas (IDH e PIB per capita) apresentam baixa correlação.

### 3.5. ReservaEcológica



Novamente, os gráficos de dispersão acima confirmam a tabela de correlações. Podemos verificar visualmente que não aparecem correlações significativas entre as variáveis estudadas: Reserva ecológica e (IDH e PIB Per capita).

## 4. ANÁLISE DOS COMPONENTES PRINCIPAIS

O objetivo deste tópico é criar um indicador utilizando uma ferramenta estatística, através da análise dos componentes principais, por meio da percepção das relações entre as variáveis e a possibilidade de agrupamento das mesmas.

Apresenta-se abaixo o resultado das análises dos componentes principais juntamente com o gráfico de “Dispersão”

### Análisis de componente principal: População; Pegada Ecoló; Captações ág; Biocapacidad; Reserva ecol; IDH 2003; PIB Per capt

Análisis de los valores y vectores propios de la matriz de correlación

Valor propio	2,9364	1,7693	1,0041	0,9338	0,2631	0,0932	0,0001
Proporción	0,419	0,253	0,143	0,133	0,038	0,013	0,000
Acumulada	0,419	0,672	0,816	0,949	0,987	1,000	1,000

Variable	PC1	PC2
População	-0,012	-0,152



Pegada Ecológica Total	0,556	-0,017
Captações água por pessoa	0,170	-0,099
Biocapacidade	0,289	0,646
Reserva ecológica	-0,087	0,730
IDH 2003	0,521	-0,095
PIB Per capita 2010	0,547	-0,088

Pela análise acima vê-se que se juntarmos as 6 variáveis em apenas 1 (PC1) teremos uma explicação de 42%, com 2 (PC1 e PC2) chegamos a 67% e assim por diante.

Em outras palavras, 42% destas seis variáveis estudadas estão sendo explicadas pela variável número 1 (amarelo). Se agregarmos a segunda variável, (azul) veremos que a ela explica mais 25%, chegando-se a um acumulado de 67%. Se agregarmos a terceira variável (verde), teremos um acumulado de 82% de explicação e por fim se agregarmos a quarta variável (cinza), chegaremos a um acumulado de 95% de explicação. Ou seja, precisamos de 4 componentes para este estudo.

Neste caso serão criados dois indicadores (PC1 e PC2) que mostram informações independentes e complementares.

Pela análise dos valores / participação de cada variável no índice PC1, poderíamos denominá-lo **“Influência econômica na Pegada Ecológica”**, uma vez que as participações das variáveis Pegada Ecológica Total, IDH 2003 e PIB Per capita 2010 são semelhantes.

Pela análise dos valores / participação de cada variável no índice PC2, poderíamos denominá-lo **“Possibilidade de Recuperação”**, uma vez que as participações mais significativas são Biocapacidade e Reserva ecológica.

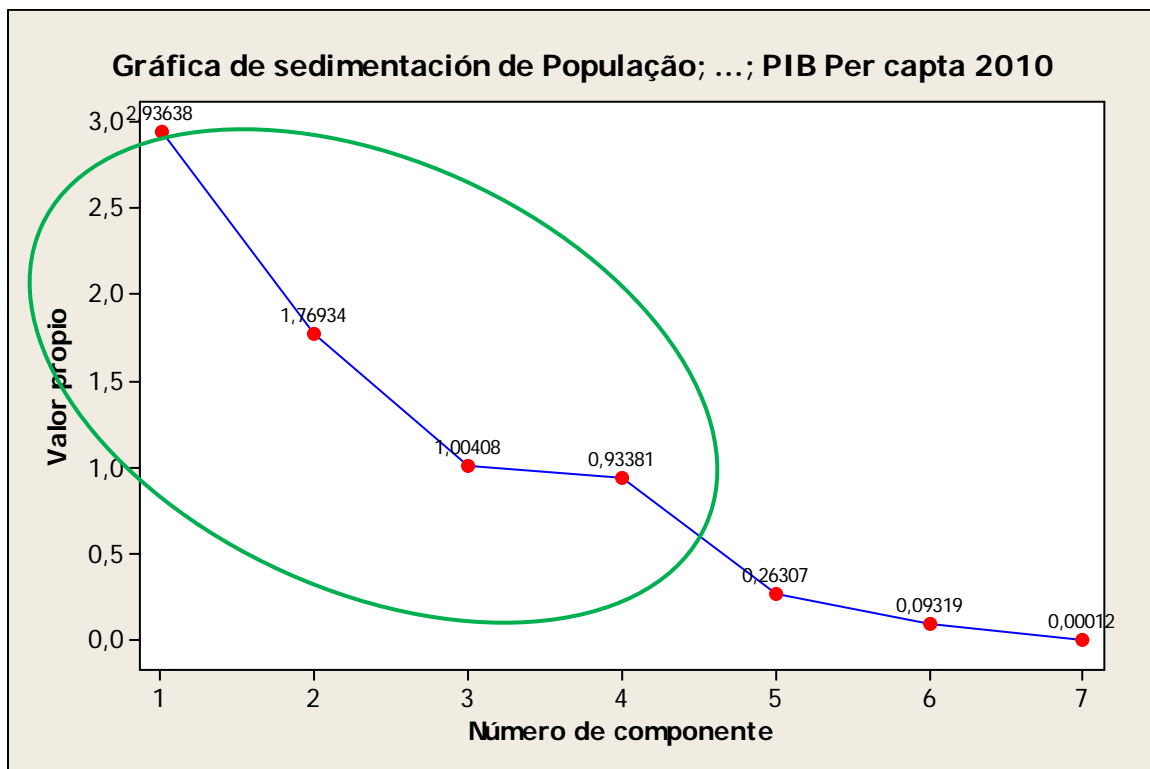
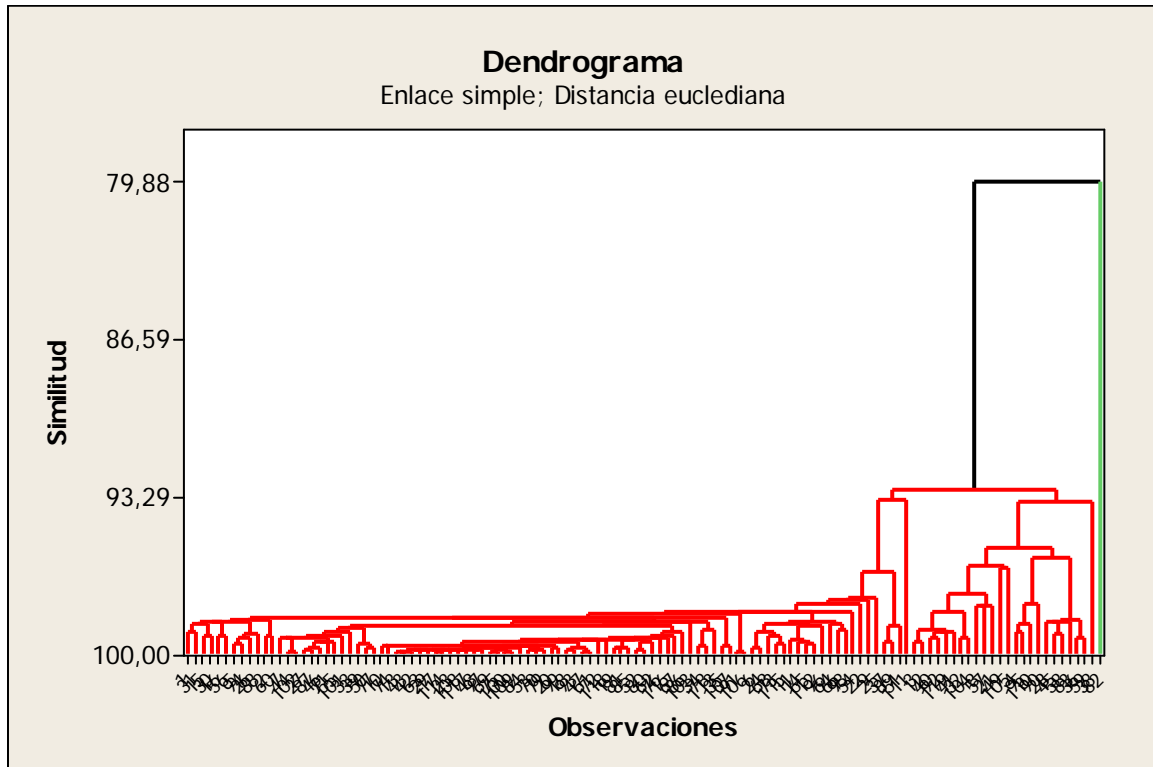


Gráfico mostra os componentes e quanto eles estão contribuindo para os dois indicadores que estão sendo criados: PC1 e PC2.

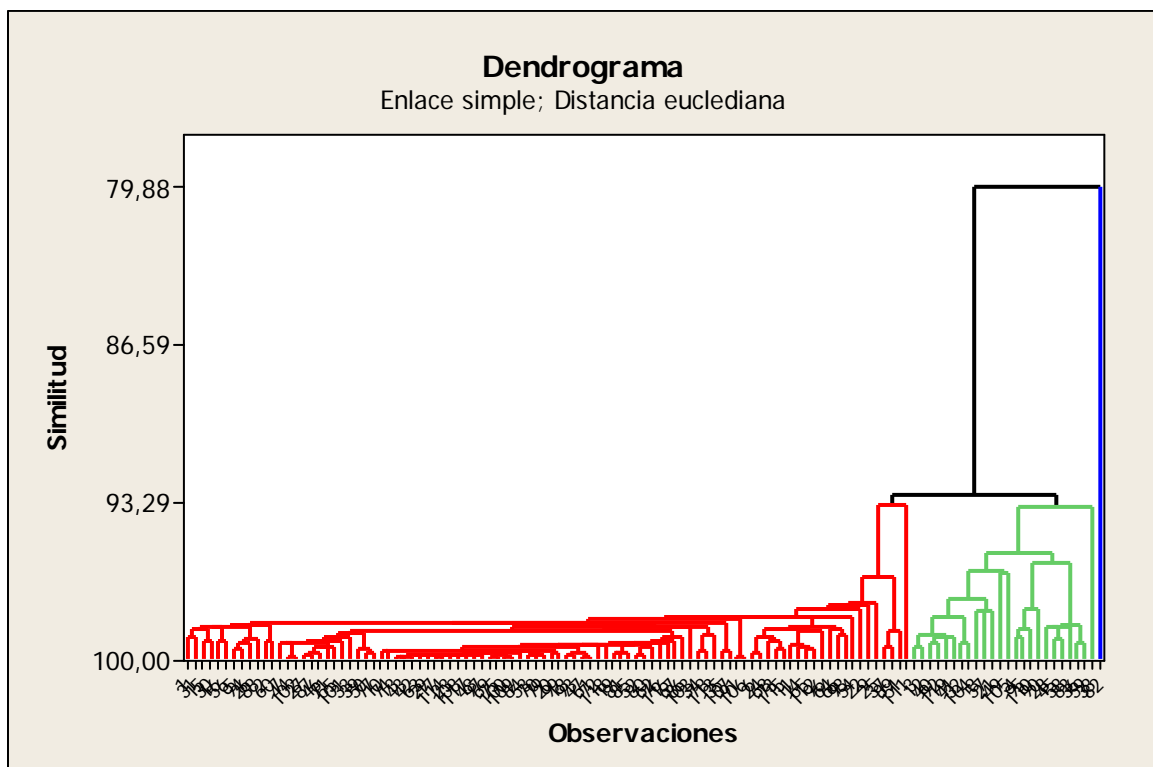
## 5. COMPARAÇÃO DE MÉDIA, ANÁLISE DE VARIÂNCIA E INTERVALO DE CONFIANÇA.

### 5.1. Dividindo em dois conglomerados



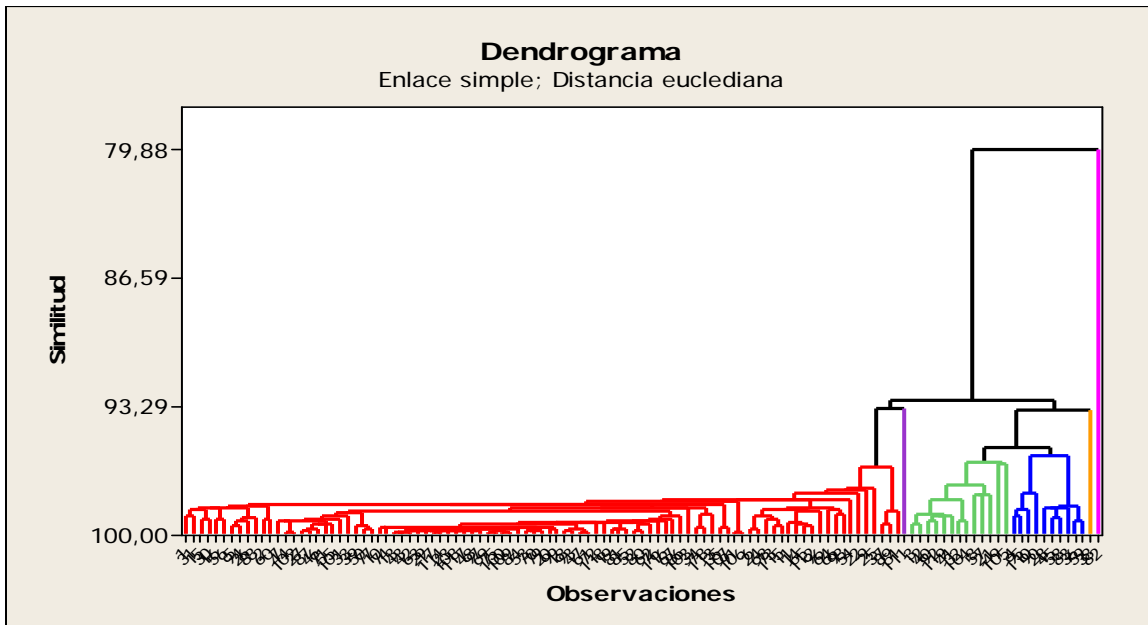
Apresenta todos os países em um único bloco, com Noruega como outlier.

### 5.2. Dividindo em três conglomerados



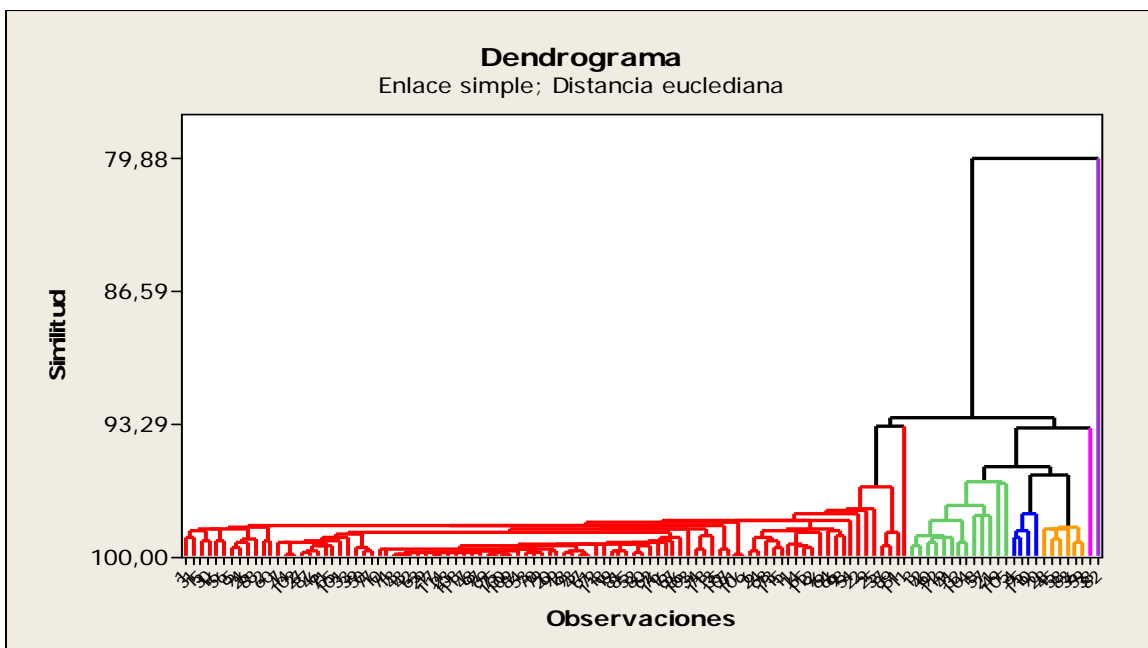
Apresenta dois blocos, o da esquerda com 94 países, que possuem a menor pegada ecológica. O segundo bloco é formado por 24 países, que possuem o maior valor de pegada ecológica. O Brasil encontra-se no primeiro bloco, com uma pegada ecológica de 2,1 hectares globais por pessoa. A Noruega aparece como outlier .

### 5.3. Dividindo em seis conglomerados



A divisão com seis conglomerados continua tendo a Noruega como outlier, além do Turcomenistão e EUA. Os valores para pegada ecológica variam entre 4 e 5 hectares globais por pessoa

### 5.4. Dividindo em sete conglomerados



A divisão com sete conglomerados continua tendo a Noruega, Turcomenistão e EUA como outlier. Neste ensaio, os países República Tcheca, Arábia Saudita, Portugal e Trinidad e Tobago destacados em azul, foram colocados em um bloco separados, e seus valores para pegada ecológica variam em torno de 4 hectares globais por pessoa.

País	População	Pegada Ecológica Total	Captações água por pessoa	Biocapacidade	Reserva ecológica	IDH 2003	PIB Per capita 2010	3 conglomerados	6 conglomerados	7 conglomerados
EUA	294	9,6	1647	4,7	-4,8	0,94	46653	2	4	5
Canadá	31,5	7,6	1470	14,5	6,9	0,95	39035	2	2	2
Finlândia	5,2	7,6	476	12	4,4	0,94	34107	2	2	2
Estônia	1,3	6,5	118	5,7	-0,7	0,85	18355	1	1	1
Suécia	8,9	6,1	334	9,6	3,5	0,95	36139	2	2	2
Nova Zelândia	3,9	5,9	549	14,9	9	0,93	27520	2	3	4
Dinamarca	5,4	5,8	237	3,5	-2,2	0,94	35736	2	2	2
Noruega	4,5	3,8	485	6,8	0,9	0,96	58278	3	5	6
Bélgica/Luxemburgo	10,8	5,6	836	1,2	-4,4	0,95	34734	2	2	2
França	60,1	5,6	668	3	-2,6	0,94	33103	2	2	2
Reino Unido	59,5	5,6	161	1,6	-4	0,94	34342	2	2	2
Espanha	41,1	5,4	870	1,7	-3,6	0,95	30475	2	3	4
Suíça	7,2	5,1	358	1,5	-3,6	0,95	43109	2	2	2
Grécia	11	5	708	1,4	-3,6	0,91	28608	2	3	4
Irlanda	4	5	289	4,8	-0,2	0,95	38768	2	2	2
Áustria	8,1	4,9	260	3,4	-1,5	0,94	37602	2	2	2
República Tcheca	10,2	4,9	252	2,6	-2,3	0,87	24419	2	3	3
Arábia Saudita	24,2	4,6	736	1	-3,7	0,77	24208	2	3	3
Israel	6,4	4,6	325	0,4	-4,2	0,92	28292	2	3	4
Alemanha	82,5	4,5	571	1,7	-2,8	0,93	34743	2	2	2
Holanda	16,1	4,4	494	0,8	-3,6	0,94	41004	2	2	2
Lituânia	3,4	4,4	78	4,2	-0,2	0,85	15327	1	1	1
Rússia	143,2	4,4	532	6,9	2,5	0,8	15719	1	1	1
Japão	127,7	4,4	694	0,7	-3,6	0,94	33649	2	2	2
Itália	57,4	4,2	772	1	-3,1	0,93	30080	2	3	4
Portugal	10,1	4,2	1121	1,6	-2,6	0,9	23019	2	3	3
Coreia do Sul	47,7	4,1	392	0,5	-3,5	0,9	29326	2	3	4
Cazaquistão	15,4	4	2263	4,1	0,1	0,76	12800	1	1	1
Hungria	9,9	3,5	770	2	-1,5	0,86	18986	1	1	1
Turcomenistão	4,9	3,5	5142	3,6	0,1	0,74	7627	1	6	7
Belarus	9,9	3,3	281	3,2	-0,1	0,79	13097	1	1	1
Polônia	38,6	3,3	419	1,8	-1,4	0,86	18406	1	1	1
Ucrânia	48,5	3,2	767	1,7	-1,5	0,77	6590	1	1	1
Bulgária	7,9	3,1	1318	2,1	-1	0,81	11547	1	1	1
Mongólia	2,6	3,1	172	11,8	8,7	0,7	3710	1	1	1
Trinidad e Tobago	1,3	3,1	239	0,4	-2,7	0,8	25162	2	3	3
Líbano	3,7	2,9	384	0,3	-2,6	0,76	13510	1	1	1
Letônia	2,3	2,6	129	6,6	4	0,84	16300	1	1	1
México	103,5	2,6	767	1,7	-0,9	0,81	14192	1	1	1
Irã	68,9	2,4	1071	0,8	-1,6	0,74	11890	1	1	1
Papua-Nova Guiné	5,7	2,4	13	2,1	-0,3	0,52	2395	1	1	1
Romênia	22,3	2,4	1035	2,3	-0,1	0,77	12910	1	1	1
África do Sul	45	2,3	279	2	-0,3	0,66	10140	1	1	1
Argentina	38,4	2,3	769	5,9	3,6	0,86	14930	1	1	1
Chile	15,8	2,3	804	5,4	3	0,85	14780	1	1	1
Malásia	24,4	2,2	376	3,7	1,5	0,8	14410	1	1	1
Brasil	178,5	2,1	336	9,9	7,8	0,79	10847	1	1	1
Turquia	71,3	2,1	534	1,4	-0,7	0,75	13359	1	1	1
Costa Rica	4,2	2	655	1,5	-0,5	0,84	11143	1	1	1
Ilhas Maurício	1,2	1,9	504	1,2	-0,7	0,48	1310	1	1	1
Panamá	3,1	1,9	268	2,5	0,6	0,8	1321	1	1	1
Uruguai	3,4	1,9	929	8	6,1	0,84	14022	1	1	1
Jordânia	5,5	1,8	190	0,3	-1,5	0,75	570	1	1	1
Azerbaijão	8,4	1,7	2079	1,2	-0,5	0,73	9870	1	1	1
Jamaica	2,7	1,7	156	0,5	-1,3	0,74	7547	1	1	1
Síria	17,8	1,7	1148	0,8	-0,9	0,72	4857	1	1	1
Argélia	31,8	1,6	194	4,7	-0,9	0,72	8477	1	1	1
Botsuana	1,8	1,6	110	4,5	3	0,57	13462	1	1	1
China	1311,7	1,6	484	0,8	-0,9	0,76	7206	1	1	1
Paraguai	5,9	1,6	85	5,6	4	0,76	4629	1	1	1
República Dominicana	8,7	1,6	393	0,8	-0,8	0,75	8616	1	1	1
Cuba	11,3	1,5	728	0,9	-0,7	0,82	9800	1	1	1
Equador	13	1,5	1326	2,2	0,7	0,76	8170	1	1	1
Albânia	3,2	1,4	544	0,9	-0,5	0,78	7737	1	1	1
Egito	71,9	1,4	969	0,5	-0,9	0,66	5840	1	1	1
El Salvador	6,5	1,4	200	0,6	-0,8	0,72	6660	1	1	1
Gâmbia	1,4	1,4	22	0,8	-0,5	0,47	1446	1	1	1
Tailândia	62,8	1,4	1400	1	-0,4	0,78	8328	1	1	1
Colômbia	44,2	1,3	246	3,6	2,3	0,79	8959	1	1	1
Guatemala	12,3	1,3	167	1,3	0	0,66	4760	1	1	1
Honduras	6,9	1,3	127	1,8	0,5	0,67	3845	1	1	1
Mauritânia	2,9	1,3	606	5,8	4,5	0,33	2037	1	1	1
Moldávia	4,3	1,3	541	0,8	-0,5	0,67	2917	1	1	1
Quirguistão	5,1	1,3	1989	1,4	0,1	0,7	2100	1	1	1
Nicarágua	5,5	1,2	244	3,5	2,4	0,69	2632	1	1	1
Nigéria	124	1,2	66	0,9	-0,2	0,45	2289	1	1	1
Senegal	10,1	1,2	225	0,9	-0,3	0,46	1830	1	1	1
Armênia	3,1	1,1	960	0,6	-0,5	0,76	5286	1	1	1
Filipinas	80	1,1	363	0,6	-0,5	0,76	3600	1	1	1
Indonésia	219,9	1,1	381	1	0	0,7	4394	1	1	1
Namíbia	2	1,1	153	4,4	3,3	0,63	6474	1	1	1
Uganda	25,8	1,1	12	0,8	-0,2	0,51	1250	1	1	1
Burkina Fasso	13	1	63	1	0	0,32	1217	1	1	1
Chade	8,6	1	28	2,5	1,5	0,34	1330	1	1	1
Gana	20,9	1	48	1,3	0,3	0,52	1533	1	1	1
Sri Lanka	19,1	1	667	0,4	-0,6	0,75	4999	1	1	1
Sudão	33,6	1	1135	1,8	0,8	0,51	230	1	1	1
Laos	5,7	0,9	543	1,3	0,4	0,55	2404	1	1	1
Marrocos	30,6	0,9	419	0,8	-0,1	0,63	4638	1	1	1
Mianmar	49,5	0,9	680	1,3	0,4	0,58	1596	1	1	1
Peru	27,2	0,9	752	3,8	3	0,76	9016	1	1	1
Togo	4,9	0,9	35	0,8	-0,1	0,51	846	1	1	1
Vietnã	81,4	0,9	889	0,8	-0,1	0,7	3097	1	1	1
Zimbábue	12,9	0,9	328	0,8	-0,1	0,5	187	1	1	1
Benin	6,7	0,8	20	0,9	0,1	0,43	1506	1	1	1
Camarões	16	0,8	63	1,3	0,4	0,5	2219	1	1	1
Geórgia	5,1	0,8	697	1,2	0,5	0,73	4946	1	1	1
Emen	20	0,8	343	0,4	-0,5	0,49	2600	1	1	1
Índia	1065,5	0,8	615	0,4	-0,4	0,6	3354	1	1	1
Lesoto	1,8	0,8	28	1,1	0,3	0,5	1605	1	1	1
Mali	13	0,8	519	1,3	0,5	0,75	1207	1	1	1
Burundi	6,8	0,7	44	0,6	-0,1	0,38	403	1	1	1
Camboja	14,1	0,7	295	0,9	0,1	0,57	1952	1	1	1
Costa do Marfim	16,6	0,7	57	2	1,2	0,42	1700	1	1	1
Guiné-Bissau	1,5	0,7	121	2,9	2,2	0,35	554	1	1	1
Madagascar	17,4	0,7	884	2,9	2,2	0,5	958	1	1	1
Nepal	25,2	0,7	414	0,5	-0,2	0,53	1189	1	1	1
Ruanda	8,4	0,7	18	0,5	-0,1	0,45	1102	1	1	1
Serra Leoa	5	0,7	80	1,1	0,4	0,3	825	1	1	1
Tanzânia	37	0,7	143	1,3	0,6	0,42	1426	1	1	1
Congo	3,7	0,6	13	7,8	7,2	0,51	4583	1	1	1
Haiti	8,3	0,6	120	0,3	-0,3	0,48	1040	1	1	1
Malauí	12,1	0,6	85	0,5	-0,1	0,4	902	1	1	1
Moçambique	18,9	0,6	34	2,1	1,4	0,38	929	1	1	1
Paquistão	153,6	0,6	1130	0,3	-0,3	0,53	2625	1	1	1
República Democrática do Congo	52,8	0,6	7	1,5	0,9	0,39	327	1	1	1
Tadjiquistão	6,2	0,6	1931	0,5	-0,1	0,65	2065	1	1	1
Zâmbia	10,8	0,6	163	3,4	2,8	0,39	1497	1	1	1
Bangladesh	146,7	0,5	552	0,3	-0,2	0,52	1458	1	1	1

Analisando-se a tabela e os gráficos de conglomerados acima identificamos que a melhor divisão em conglomerados para este estudo é representada pelo gráfico com três conglomerados, apresentando dois blocos, sendo o primeiro com 94 países, que possuem a menor pegada ecológica e o segundo com 24 países, que possuem o maior valor de pegada ecológica. Em todas as simulações, a Noruega apareceu como outlier .

## Dendrograma

Análisis de observaciones de conglomerado: População; Pegada Ecoló; Captações ág; Biocapacidad; Reserva ecol; IDH 2003; ...

Distancia eucladiana, Enlace simple  
Pasos de amalgamación

Paso	Número de conglomerados	Nivel de semejanza	Nivel de distancia	Conglomerados incorporados	Nuevo conglomerado	Número de obs. en el conglomerado nuevo	
1	117	99,9109	41,41	13	43	13	2
2	116	99,8932	49,66	99	108	99	2
3	115	99,8749	58,15	69	75	69	2
4	114	99,8703	60,27	13	42	13	3
5	113	99,8665	62,08	17	113	17	2
6	112	99,8655	62,51	44	101	44	2
7	111	99,8542	67,76	90	105	90	2
8	110	99,8471	71,10	47	61	47	2
9	109	99,8378	75,42	69	99	69	4
10	108	99,8341	77,13	13	63	13	4
11	107	99,8315	78,34	107	116	107	2
12	106	99,8206	83,41	17	23	17	3
13	105	99,8159	85,56	27	86	27	2
14	104	99,7930	96,24	18	92	18	2
15	103	99,7839	100,43	13	29	13	5
16	102	99,7746	104,77	53	70	53	2
17	101	99,7738	105,14	52	80	52	2
18	100	99,7693	107,23	53	79	53	3
19	99	99,7496	116,41	13	17	13	8
20	98	99,7432	119,37	48	96	48	2
21	97	99,7398	120,97	13	107	13	10
22	96	99,7342	123,58	13	48	13	12
23	95	99,7231	128,74	19	81	19	2
24	94	99,6973	140,71	20	98	20	2
25	93	99,6970	140,84	13	69	13	16
26	92	99,6925	142,96	50	77	50	2
27	91	99,6670	154,79	27	46	27	3
28	90	99,6651	155,68	6	24	6	2
29	89	99,6413	166,74	13	83	13	17
30	88	99,6398	167,46	19	84	19	3
31	87	99,6351	169,61	18	47	18	4
32	86	99,6097	181,45	35	106	35	2
33	85	99,5870	191,98	34	78	34	2
34	84	99,5739	198,08	13	53	13	20
35	83	99,5443	211,84	10	74	10	2
36	82	99,5428	212,53	13	20	13	22
37	81	99,5417	213,03	10	13	10	24
38	80	99,5196	223,33	65	111	65	2
39	79	99,4965	234,06	39	50	39	3
40	78	99,4755	243,84	5	93	5	2
41	77	99,4505	255,45	73	114	73	2
42	76	99,4103	274,12	10	18	10	28
43	75	99,4091	274,68	3	12	3	2
44	74	99,4014	278,27	14	65	14	3
45	73	99,3968	280,40	52	62	52	3
46	72	99,3964	280,58	27	71	27	4
47	71	99,3383	307,60	37	88	37	2
48	70	99,3307	311,15	27	55	27	5
49	69	99,3059	322,67	10	72	10	29
50	68	99,2880	330,97	19	52	19	6
51	67	99,2418	352,47	10	117	10	30
52	66	99,2391	353,72	10	19	10	36
53	65	99,2054	369,36	11	14	11	4
54	64	99,1953	374,09	10	76	10	37
55	63	99,1936	374,88	5	26	5	3
56	62	99,1460	396,96	40	91	40	2

57	61	99,1410	399,31	10	115	10	38
58	60	99,1238	407,30	36	59	36	2
59	59	99,1170	410,46	7	44	7	3
60	58	99,1081	414,61	32	103	32	2
61	57	99,1058	415,69	7	27	7	8
62	56	99,0813	427,06	16	56	16	2
63	55	99,0706	432,02	2	60	2	2
64	54	99,0494	441,87	10	67	10	39
65	53	99,0220	454,61	68	73	68	3
66	52	99,0107	459,88	7	100	7	9
67	51	98,9912	468,93	15	30	15	2
68	50	98,9318	496,56	45	58	45	2
69	49	98,9040	509,48	5	87	5	4
70	48	98,8770	522,01	40	118	40	3
71	47	98,8638	528,18	4	94	4	2
72	46	98,8494	534,86	10	85	10	40
73	45	98,8151	550,79	40	41	40	4
74	44	98,7989	558,33	7	33	7	10
75	43	98,7863	564,22	1	31	1	2
76	42	98,7828	565,83	6	68	6	5
77	41	98,7677	572,85	34	112	34	3
78	40	98,7653	573,96	3	40	3	6
79	39	98,6751	615,90	66	97	66	2
80	38	98,6204	641,32	7	39	7	13
81	37	98,5489	674,54	11	95	11	5
82	36	98,5403	678,55	37	51	37	3
83	35	98,4729	709,86	7	10	7	53
84	34	98,4490	720,99	64	66	64	3
85	33	98,4472	721,84	1	15	1	4
86	32	98,4013	743,17	4	109	4	3
87	31	98,3912	747,87	6	11	6	10
88	30	98,3431	770,23	7	102	7	54
89	29	98,3298	776,40	1	16	1	6
90	28	98,3283	777,08	7	34	7	57
91	27	98,3123	784,54	1	5	1	10
92	26	98,3106	785,32	28	45	28	3
93	25	98,2840	797,71	6	64	6	13
94	24	98,2702	804,08	28	82	28	4
95	23	98,2380	819,10	1	2	1	12
96	22	98,1835	844,40	28	36	28	6
97	21	98,0895	888,10	1	7	1	69
98	20	98,0770	893,93	1	35	1	71
99	19	97,8738	988,40	1	90	1	73
100	18	97,8021	1021,70	1	6	1	86
101	17	97,7487	1046,53	1	54	1	87
102	16	97,7402	1050,50	3	32	3	8
103	15	97,4909	1166,37	8	57	8	2
104	14	97,3945	1211,18	8	21	8	3
105	13	97,3477	1232,94	1	22	1	88
106	12	97,3113	1249,86	4	89	4	4
107	11	97,0898	1352,82	1	9	1	89
108	10	97,0058	1391,89	1	25	1	90
109	9	96,8487	1464,89	3	8	3	11
110	8	95,5792	2055,04	1	37	1	93
111	7	95,4622	2109,41	49	104	49	2
112	6	95,2722	2197,74	3	49	3	13
113	5	94,8836	2378,37	4	28	4	10
114	4	94,3298	2635,82	3	4	3	23
115	3	91,8641	3782,03	3	38	3	24
116	2	91,8331	3796,45	1	110	1	94
117	1	91,2914	4048,25	1	3	1	118

Partición final

Número de conglomerados: 2

	Número de observaciones	Dentro de la suma de cuadrados del conglomerado	Distancia promedio desde el centroide	Distancia máxima desde centroide
Conglomerado1	94	2620853701	4570,48	13038,4
Conglomerado2	24	885588574	4987,66	13619,1

Centroides de grupo

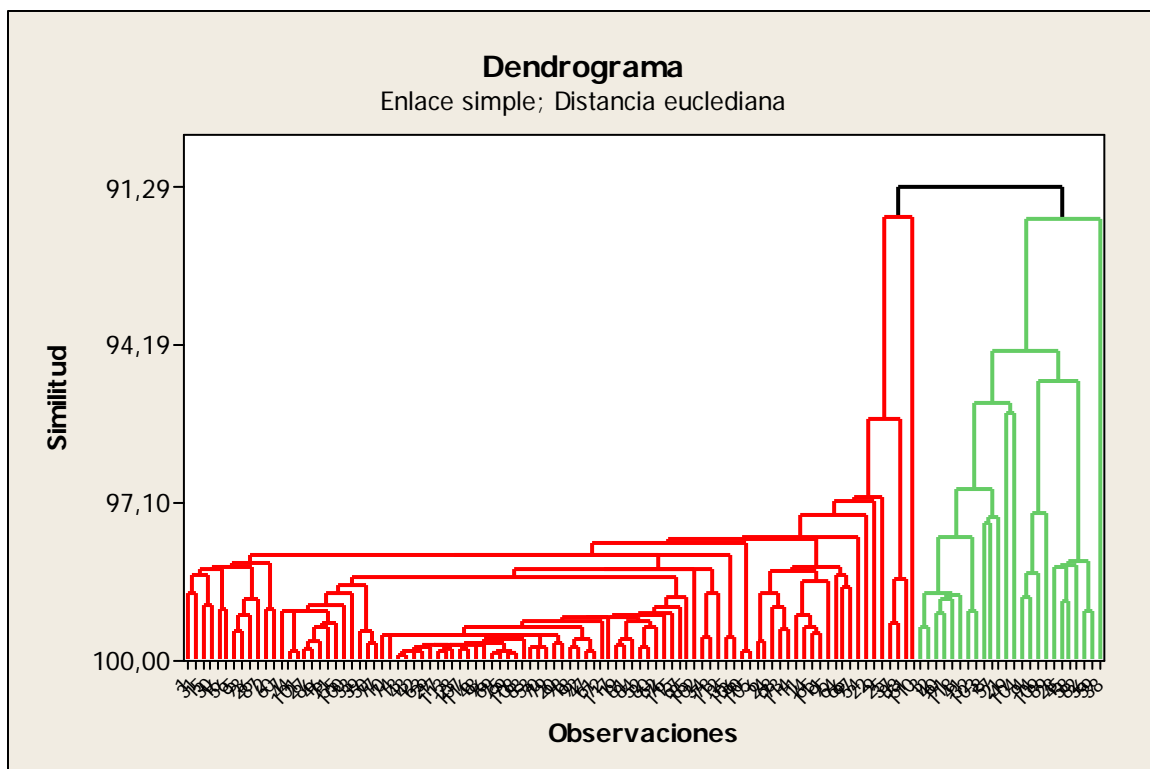
Centroide

Variable	Conglomerado1	Conglomerado2	principal
População	52,86	38,9	50,0
<b>Pegada Ecológica Total</b>	<b>1,58</b>	<b>5,3</b>	<b>2,3</b>
Captações água por pessoa	540,26	602,5	552,9
Biocapacidade	2,18	3,7	2,5
Reserva ecológica	0,59	-1,6	0,1
IDH 2003	0,64	0,9	0,7
<b>PIB Per capita 2010</b>	<b>5949,74</b>	<b>33076,4</b>	<b>11467,0</b>

Las distancias entre los centroides de conglomerados

	Conglomerado1	Conglomerado2
Conglomerado1	0,0	27126,7
Conglomerado2	27126,7	0,0

Dendrograma



Para esta análise foi retirado da base o outlier Noruega, para possibilitar um estudo mais aprofundado dos países que possuem características semelhantes.

Identificamos que o cluster 1 (**em vermelho**) representa o grupo de países com menor desenvolvimento sócio-econômico e menor pegada ecológica, enquanto que o cluster 1(**em verde**) compõe o grupo de países com maior desenvolvimento sócio-econômico e maior pegada ecológica.

### 5.5. Análise ANOVA

Procedemos a análise ANOVA para identificar as variáveis que se destacam nessa amostra.

#### ANOVA unidireccional: População vs. com 2 conglomerados

Fuente	GL	SC	CM	F	P
--------	----	----	----	---	---

com 2 conglomerados	1	3708	3708	0,15	0,702
Error	116	2915445	25133		
Total	117	2919152			

S = 158,5    R-cuad. = 0,13%    R-cuad.(ajustado) = 0,00%

ICs de 95% individuales para la media basados en Desv.Est. agrupada

Nivel	N	Media	Desv.Est.
1	94	52,9	174,3
2	24	38,9	62,6

Desv.Est. agrupada = 158,5

**ANOVA unidireccional: Pegada Ecológica Total vs. com 2 conglomerados**

Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	267,56	267,56	213,03	0,000
Error	116	145,69	1,26		
Total	117	413,25			

S = 1,121    R-cuad. = 64,75%    R-cuad.(ajustado) = 64,44%

ICs de 95% individuales para la media basados en Desv.Est. agrupada

Nivel	N	Media	Desv.Est.
1	94	1,584	1,050
2	24	5,325	1,368

Desv.Est. agrupada = 1,121

**ANOVA unidireccional: Captações água por pessoa vs. com 2 conglomerados**

Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	73974	73974	0,18	0,670
Error	116	47034616	405471		
Total	117	47108590			

S = 636,8    R-cuad. = 0,16%    R-cuad.(ajustado) = 0,00%

ICs de 95% individuales para la media basados en Desv.Est. agrupada

Nivel	N	Media	Desv.Est.
1	94	540,3	685,0
2	24	602,5	384,1

Desv.Est. agrupada = 636,8

**ANOVA unidireccional: Biocapacidade vs. com 2 conglomerados**

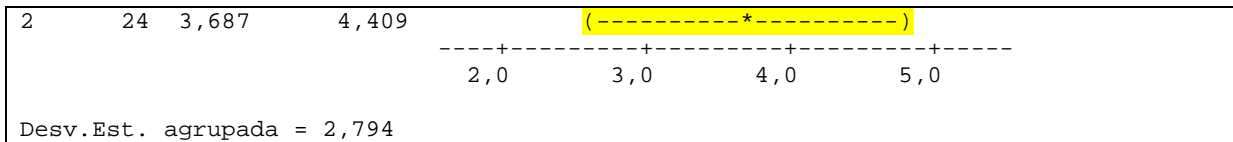
Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	43,71	43,71	5,60	0,020
Error	116	905,44	7,81		
Total	117	949,15			

S = 2,794    R-cuad. = 4,60%    R-cuad.(ajustado) = 3,78%

ICs de 95% individuales para la media basados en Desv.Est. agrupada

Nivel	N	Media	Desv.Est.
1	94	2,176	2,220

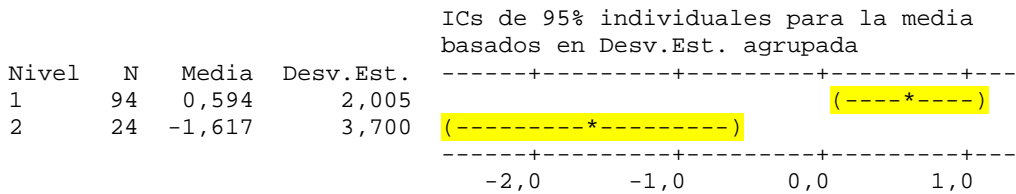




**ANOVA unidireccional: Reserva ecológica vs. com 2 conglomerados**

Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	93,40	93,40	15,73	0,000
Error	116	688,95	5,94		
Total	117	782,35			

S = 2,437 R-cuad. = 11,94% R-cuad.(ajustado) = 11,18%

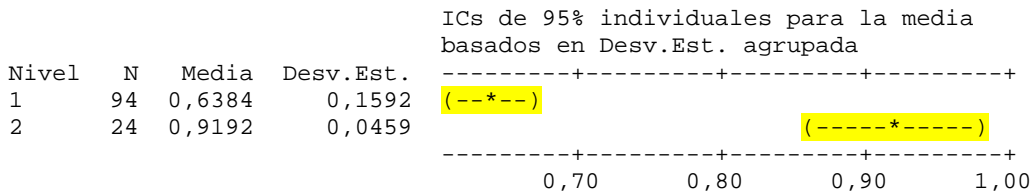


Desv.Est. agrupada = 2,437

**ANOVA unidireccional: IDH 2003 vs. com 2 conglomerados**

Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	1,5071	1,5071	72,69	0,000
Error	116	2,4050	0,0207		
Total	117	3,9121			

S = 0,1440 R-cuad. = 38,52% R-cuad.(ajustado) = 37,99%

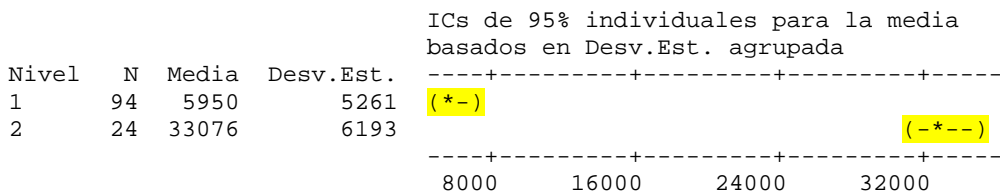


Desv.Est. agrupada = 0,1440

**ANOVA unidireccional: PIB Per capita 2010 vs. com 2 conglomerados**

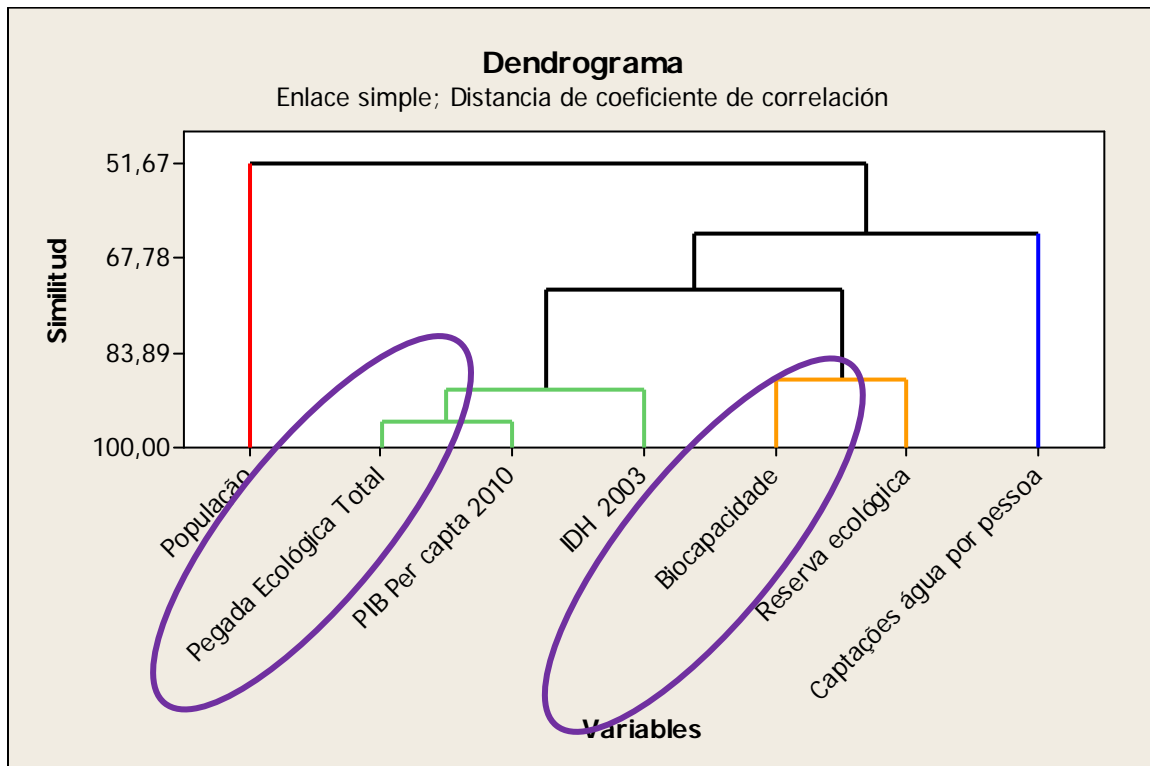
Fuente	GL	SC	CM	F	P
com 2 conglomerados	1	14068532097	14068532097	472,14	0,000
Error	116	3456490471	29797332		
Total	117	17525022569			

S = 5459 R-cuad. = 80,28% R-cuad.(ajustado) = 80,11%



Desv.Est. agrupada = 5459

Pela análise de variância acima e tendo como parâmetro o p-value menor ou igual a 5%, identificamos que as variáveis que têm maior relação e portanto, melhor se enquadram no nosso estudo são Pegada Ecológica Total, Biocapacidade, Reserva ecológica, HDI e PIB Per capita, pois as médias dessas variáveis são mais distantes nos clusters estudados. Esse é um indicador para a escolha dessas variáveis para a discriminante.



Como análise adicional, poderemos escolher uma variável de cada grupo apresentado no dendrograma acima para a análise discriminante: Pegada Ecológica Total e Biocapacidade.

## 6. ANÁLISE DISCRIMINANTE

**Análisis discriminante: com 2 conglo vs. Todas as variáveis: População; Pegada Ecoló;**

...

Método lineal para respuesta: com 2 conglomerados

**Predictores: População; Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capita 2010**

Grupo	1	2
Conteo	94	24

Resumen de clasificación

Colocar en un grupo	Grupo verdadero	
	1	2
1	94	0
2	0	24
N Total	94	24
N correcta	94	24
Proporción	1,000	1,000

N = 118

N Correcta = 118

Proporción Correcta = 1,000

Distancia cuadrada entre grupos

	1	2
1	0,0000	31,1651
2	31,1651	0,0000

Función discriminativa lineal para grupos

	1	2
Constante	-14,351	-23,997
População	0,000	-0,002
Pegada Ecológica Total	1,476	-5,678
Captações água por pessoa	-0,001	-0,002
Biocapacidade	-0,801	6,549
Reserva ecológica	0,935	-6,942
IDH 2003	52,044	31,507
PIB Per capta 2010	-0,001	0,000

### Análisis discriminante: com 2 conglo vs. População; Pegada Ecoló; ...

Método quadratico para respuesta: com 2 conglomerados

Predictores: População; Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010

Pegada Ecológica Total está altamente correlacionado con otros predictores del grupo 1.

Biocapacidade está altamente correlacionado con otros predictores del grupo 1.

Reserva ecológica está altamente correlacionado con otros predictores del grupo 1.

\* ERROR \*

Biocapacidade está altamente correlacionado con otros predictores del grupo 2.

\* ERROR \* No se pueden realizar los cálculos para análisis discriminante.

### Análisis discriminante: com 2 conglo vs. Pegada Ecoló; Captações ág; ...

Después de restar la media del grupo, Pegada Ecológica Total está altamente correlacionado con otros predictores.

Después de restar la media del grupo, Biocapacidade está altamente correlacionado con otros predictores.

Después de restar la media del grupo,  
Reserva ecológica está altamente correlacionado con otros predictores.

**Método lineal** para respuesta: com 2 conglomerados

**Predictores: Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010**

Grupo	1	2
Conteo	94	24

Resumen de clasificación

	Grupo	
Colocar en	verdadero	
un grupo	1	2
1	94	0
2	0	24
N Total	94	24
N correcta	94	24
Proporción	1,000	1,000

N = 118

N Correcta = 118

**Proporción Correcta = 1,000**

Distancia cuadrada entre grupos

	1	2
1	0,0000	31,0559
2	31,0559	0,0000

Función discriminativa lineal para grupos

	1	2
Constante	-14,349	-23,963
Pegada Ecológica Total	1,341	-5,159
Captações água por pessoa	-0,001	-0,002
Biocapacidade	-0,667	6,035
Reserva ecológica	0,799	-6,418
IDH 2003	52,076	31,386
PIB Per capta 2010	-0,001	0,000

### **Análisis discriminante: com 2 conglo vs. Pegada Ecoló; Captações ág; ...**

**Método quadratico** para respuesta: com 2 conglomerados

**Predictores: Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010**

Pegada Ecológica Total está altamente correlacionado con otros predictores del grupo 1.

Biocapacidade está altamente correlacionado con otros predictores del grupo 1.

Reserva ecológica está altamente correlacionado con otros predictores del grupo 1.

\* ERROR \*

Biocapacidade está altamente correlacionado con otros predictores del grupo 2.

\* ERROR \* No se pueden realizar los cálculos para análisis discriminante.

Método lineal para respuesta: com 2 conglomerados

Predictores: Pegada Ecológica Total; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010

Grupo	1	2
Conteo	94	24

Resumen de clasificación

	Grupo	
Colocar en	verdadero	
un grupo	1	2
1	94	0
2	0	24
N Total	94	24
N correcta	94	24
Proporción	1,000	1,000

N = 118

N Correcta = 118

Proporción Correcta = 1,000

Distancia cuadrada entre grupos

	1	2
1	0,0000	30,8515
2	30,8515	0,0000

Función discriminativa lineal para grupos

	1	2
Constante	-14,043	-23,201
Pegada Ecológica Total	1,947	-4,202
Biocapacidade	-1,474	4,762
Reserva ecológica	1,640	-5,091
IDH 2003	50,143	28,335
PIB Per capta 2010	-0,001	0,001

**Análisis discriminante: com 2 conglo vs. Pegada Ecoló; Biocapacidad; ...**

Método quadratico para respuesta: com 2 conglomerados

Predictores: Pegada Ecológica Total; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010

Pegada Ecológica Total está altamente correlacionado con otros predictores del grupo 1.

Biocapacidade está altamente correlacionado con otros predictores del grupo 1.

Reserva ecológica está altamente correlacionado con otros predictores del grupo 1.

\* ERROR \*

Biocapacidade está altamente correlacionado con otros predictores del grupo 2.

\* ERROR \* No se pueden realizar los cálculos para análisis discriminante.

### **Análisis discriminante: com 2 conglo vs. Pegada Ecoló; PIB Per capt**

**Método lineal** para respuesta: com 2 conglomerados

**Predictores: Pegada Ecológica Total; PIB Per capta 2010**

Grupo	1	2
Conteo	94	24

Resumen de clasificación

Colocar en un grupo	Grupo verdadero	
	1	2
1	94	0
2	0	24
N Total	94	24
N correcta	94	24
Proporción	1,000	1,000

N = 118

N Correcta = 118

**Proporción Correcta = 1,000**

Distancia cuadrada entre grupos

	1	2
1	0,0000	24,7948
2	24,7948	0,0000

Función discriminativa lineal para grupos

	1	2
Constante	-1,005	-18,533
Pegada Ecológica Total	1,159	0,756
PIB Per capta 2010	0,000	0,001

### **Análisis discriminante: com 2 conglo vs. Pegada Ecoló; PIB Per capt**

**Método cuadrático de respuesta:** com 2 conglomerados

**Predictores: Pegada Ecológica Total; PIB Per capta 2010**

Grupo	1	2
Conteo	94	24

Resumen de clasificación

Colocar en un grupo	Grupo verdadero	
	1	2
1	93	0
2	1	24
N Total	94	24
N correcta	93	24
Proporción	0,989	1,000

N = 118

N Correcta = 117

Proporción Correcta = 0,992

De la distancia cuadrada generalizada al grupo

Grupo	1	2
1	16,33	36,95
2	43,34	17,69

Resumen da las observaciones clasificadas incorrectamente

Observación	Grupo verdadero		Grupo de predictores	Grupo	Distancia cuadrada	Probabilidad
	1	2				
37**	1	2	1	42,10	0,003	
			2	30,68	0,997	

**Análisis discriminante: com 2 conglo vs. Pegada Ecoló; Biocapacidad**

Método lineal para respuesta: com 2 conglomerados

Predictores: Pegada Ecológica Total; Biocapacidades

Grupo	1	2
Conteo	94	24

Resumen de clasificación

Colocar en un grupo	Grupo verdadero	
	1	2
1	88	0
2	6	24
N Total	94	24
N correcta	88	24
Proporción	0,936	1,000

N = 118

N Correcta = 112

Proporción Correcta = 0,949

Distancia cuadrada entre grupos

	1	2
1	0,0000	12,7935
2	12,7935	0,0000

Función discriminativa lineal para grupos

	1	2
Constante	-1,003	-12,004
Pegada Ecológica Total	1,216	4,850
Biocapacidades	0,037	-0,493

## Resumen da las observaciones clasificadas incorrectamente

Observación	Grupo verdadero	Grupo de predictores	Grupo	Distancia cuadrada	Probabilidad
22**	1	2	1	4,840	0,204
			2	2,113	0,796
37**	1	2	1	20,354	0,000
			2	1,152	1,000
51**	1	2	1	4,025	0,341
			2	2,707	0,659
66**	1	2	1	6,676	0,059
			2	1,148	0,941
88**	1	2	1	3,406	0,491
			2	3,330	0,509
97**	1	2	1	6,575	0,209
			2	3,908	0,791

**Análisis discriminante: com 2 conglo vs. Pegada Ecoló; Biocapacidad**

Método cuadrático de respuesta: com 2 conglomerados

Predictores: Pegada Ecológica Total; Biocapacidad

Grupo	1	2
Conteo	94	24

## Resumen de clasificación

Colocar en un grupo	Grupo verdadero	
	1	2
1	88	0
2	6	24
N Total	94	24
N correcta	88	24
Proporción	0,936	1,000

N = 118

N Correcta = 112

Proporción Correcta = 0,949

De la distancia cuadrada generalizada al grupo

Grupo	1	2
1	1,488	13,625
2	15,070	3,100

## Resumen da las observaciones clasificadas incorrectamente

Observación	Grupo verdadero	Grupo de predictores	Grupo	Distancia cuadrada	Probabilidad
22**	1	2	1	6,796	0,273
			2	4,837	0,727
37**	1	2	1	23,617	0,000
			2	3,848	1,000
51**	1	2	1	5,736	0,435
			2	5,211	0,565
66**	1	2	1	8,747	0,086
			2	4,032	0,914
77**	1	2	1	20,50	0,332
			2	19,11	0,668
97**	1	2	1	9,832	0,114
			2	5,728	0,886



Cenário:	Modelo Quadrático	Modelo linear
Todas as variáveis	Erro	1,000
Todas as variáveis, tirando população	Erro	1,000
Variáveis escolhidas tomando-se como base a análise ANOVA (Pegada Ecológica Total, Biocapacidade, Reserva ecológica, HDI e PIB Per capita (tirando população e captação de água))	Erro	1,000
Variáveis Pegada Ecológica Total e PIB Per capita (modelo parcimonioso – 2 variáveis)	0,992	1,000
Variáveis escolhidas pela análise do dendograma: Pegada Ecológica Total; Biocapacidade	0,949	0,949

Em três situações, análise discriminatória no modelo linear apresentou um percentual de acerto de 100%, a saber: 1) Todas as variáveis, 2) tirando-se a variável população e 3) Utilizando-se as variáveis escolhidas tomando-se como base a análise ANOVA – (todas as variáveis menos população e captação de água). Nestes três casos a análise discriminante pelo modelo quadrático apresentou erro, pelas variáveis estarem altamente correlacionadas.

Para os dois modelos (linear e quadrático), a análise discriminante para as variáveis escolhidas pelo dendograma apresentaram a mesma correlação (0,949), com quase os mesmos erros.

Na tentativa de um modelo parcimonioso utilizando apenas duas variáveis Pegada Ecológica Total e PIB Per capita, obtivemos 100% de acerto para o modelo linear e um nível de acerto menor para o modelo quadrático (0,992).

## 7. ANÁLISE LOGÍSTICA

Foi utilizada a regressão binária, pois temos apenas dois clusters.

regresión logística ordinal: amostra 1 vs. Todas as variáveis			
População; Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capita 2010			
Regresión logística binaria: com 2 conglo vs. População; Pegada Ecoló; ...			
* ADVERTENCIA * El algoritmo no ha convergido después de 20 iteraciones.			
* ADVERTENCIA * No se ha alcanzado la convergencia para el criterio de cálculos de parámetros.			
* ADVERTENCIA * Los resultados podrían no ser confiables.			
* ADVERTENCIA * Intente elevar el número máximo de iteraciones.			
Función de enlace: Logit			
Información de respuesta			
Variable	Valor	Conteo	
com 2 conglomerados	2	24	(Evento)
	1	94	
	Total	118	
Tabla de regresión logística			
		Relación de	IC de 95%

Predictor	Coef	SE Coef	Z	P	probabilidades	Inferior	Superior
Constante	59,0288	15620,4	0,00	0,997			
População	0,0424658	15,6677	0,00	0,998	1,04	0,00	2,26355E+13
Pegada Ecológica Total	-190,907	157821	-0,00	0,999	0,00	0,00	*
Captações água por pessoa	0,0126799	4,85306	0,00	0,998	1,01	0,00	13691,13
Biocapacidade	195,502	153441	0,00	0,999	8,04022E+84	0,00	*
Reserva ecológica	-197,622	152909	-0,00	0,999	0,00	0,00	*
IDH 2003	-300,146	29671,1	-0,01	0,992	0,00	0,00	*
PIB Per capta 2010	0,0081303	1,22949	0,01	0,995	1,01	0,09	11,22

Log-verosimilitud = -0,000  
 Probar que todas las pendientes son cero: G = 119,196, GL = 7, valor P = 0,000

Pruebas de bondad del ajuste

Método	Chi-cuadrada	GL	P
Pearson	0,0000005	110	1,000
Desviación	0,0000009	110	1,000
Hosmer-Lemeshow	0,0000005	8	1,000

Medidas de asociación:  
 (Entre la variable de respuesta y las probabilidades pronosticadas)

Pares	Número	Porcentaje	Medidas de resumen
Concordante	2256	100,0	D de Somers 1,00
Discordante	0	0,0	Gamma de Goodman-Kruskal 1,00
Empates	0	0,0	Tau-a de Kendall 0,33
Total	2256	100,0	

#### Regresión logística binaria: com 2 conglo vs. Pegada Ecoló; Captações ág; ...

Pegada Ecológica Total; Captações água por pessoa; Biocapacidade; Reserva ecológica; IDH 2003; PIB Per capta 2010

- \* ADVERTENCIA \* El algoritmo no ha convergido después de 20 iteraciones.
- \* ADVERTENCIA \* No se ha alcanzado la convergencia para el criterio de cálculos de parámetros.
- \* ADVERTENCIA \* Los resultados podrían no ser confiables.
- \* ADVERTENCIA \* Intente elevar el número máximo de iteraciones.

Función de enlace: Logit

Información de respuesta

Variable	Valor	Conteo
com 2 conglomerados	2	24 (Evento)
	1	94
Total		118

Tabla de regresión logística

Predictor	Coef	SE Coef	Z	P	Relación de probabilidades	IC de 95% Inferior	Superior
Constante	59,2858	14190,1	0,00	0,997			
Pegada Ecológica Total	-202,044	196639	-0,00	0,999	0,00	0,00	*
Captações água por pessoa	0,0127029	4,18919	0,00	0,998	1,01	0,00	3727,01
Biocapacidade	206,575	193432	0,00	0,999	5,18010E+89	0,00	*
Reserva ecológica	-208,700	193051	-0,00	0,999	0,00	0,00	*
IDH 2003	-299,262	27729,4	-0,01	0,991	0,00	0,00	*
PIB Per capta 2010	0,0081120	1,19869	0,01	0,995	1,01	0,10	10,56

Log-verosimilitud = -0,000

Probar que todas las pendientes son cero: G = 119,196, GL = 6, valor P = 0,000

Pruebas de bondad del ajuste

Método	Chi-cuadrada	GL	P
Pearson	0,0000005	111	1,000
Desviación	0,0000009	111	1,000
Hosmer-Lemeshow	0,0000005	8	1,000

Medidas de asociación:

(Entre la variable de respuesta y las probabilidades pronosticadas)

Pares	Número	Porcentaje	Medidas de resumen
Concordante	2256	100,0	D de Somers 1,00
Discordante	0	0,0	Gamma de Goodman-Kruskal 1,00
Empates	0	0,0	Tau-a de Kendall 0,33

Total	2256	100,0						
<b>Regresión logística binaria: com 2 conglo vs. Pegada Ecoló; Biocapacidad; ...</b>								
<b>Predictores: Pegada Ecológica Total; Biocapacidad; Reserva ecológica; IDH 2003; PIB Per capta 2010</b>								
* ADVERTENCIA * El algoritmo no ha convergido después de 20 iteraciones.								
* ADVERTENCIA * No se ha alcanzado la convergencia para el criterio de cálculos de parámetros.								
* ADVERTENCIA * Los resultados podrían no ser confiables.								
* ADVERTENCIA * Intente elevar el número máximo de iteraciones.								
Función de enlace: Logit								
Información de respuesta								
Variable	Valor	Conteo						
com 2 conglomerados	2	24	(Evento)					
	1	94						
	Total	118						
Tabla de regresión logística								
Predictor	Coef	SE Coef	Z	P	Relación de probabilidades	IC de 95%		
Constante	59,2294	18203,0	0,00	0,997		Inferior	Superior	
Pegada Ecológica Total	-293,324	96639,0	-0,00	0,998	0,00	0,00		*
Biocapacidad	303,767	100111	0,00	0,998	8,40157E+131	0,00		*
Reserva ecológica	-307,200	100425	-0,00	0,998	0,00	0,00		*
IDH 2003	-307,568	32552,6	-0,01	0,992	0,00	0,00		*
PIB Per capta 2010	0,0078234	1,08740	0,01	0,994	1,01	0,12	8,49	
Log-verosimilitud = -0,000								
Probar que todas las pendientes son cero: G = 119,196, GL = 5, valor P = 0,000								
Pruebas de bondad del ajuste								
Método	Chi-cuadrada	GL	P					
Pearson	0,0000005	112	1,000					
Desviación	0,0000011	112	1,000					
Hosmer-Lemeshow	0,0000005	8	1,000					
Medidas de asociación:								
(Entre la variable de respuesta y las probabilidades pronosticadas)								
Pares	Número	Porcentaje	Medidas de resumen					
Concordante	2256	100,0	D de Somers	1,00				
Discordante	0	0,0	Gamma de Goodman-Kruskal	1,00				
Empates	0	0,0	Tau-a de Kendall	0,33				
Total	2256	100,0						
<b>Regresión logística binaria: com 2 conglo vs. Pegada Ecoló; PIB Per capt</b>								
<b>Pegada Ecológica Total; PIB Per capta 2010</b>								
* ADVERTENCIA * El algoritmo no ha convergido después de 20 iteraciones.								
* ADVERTENCIA * No se ha alcanzado la convergencia para el criterio de cálculos de parámetros.								
* ADVERTENCIA * Los resultados podrían no ser confiables.								
* ADVERTENCIA * Intente elevar el número máximo de iteraciones.								
Función de enlace: Logit								
Información de respuesta								
Variable	Valor	Conteo						
com 2 conglomerados	2	24	(Evento)					
	1	94						
	Total	118						
Tabla de regresión logística								
Predictor	Coef	SE Coef	Z	P	Relación de probabilidades	IC de 95%		
Constante	-152,909	6917,01	-0,02	0,982		Inferior	Superior	
Pegada Ecológica Total	0,102170	2452,67	0,00	1,000	1,11	0,00		*
PIB Per capta 2010	0,0072660	0,510386	0,01	0,989	1,01	0,37	2,74	
Log-verosimilitud = -0,000								

Probar que todas las pendientes son cero:  $G = 119,196$ ,  $GL = 2$ , valor  $P = 0,000$

Pruebas de bondad del ajuste

Método	Chi-cuadrada	GL	P
Pearson	0,0000009	115	1,000
Desviación	0,0000017	115	1,000
Hosmer-Lemeshow	0,0000009	8	1,000

Medidas de asociación:  
(Entre la variable de respuesta y las probabilidades pronosticadas)

Pares	Número	Porcentaje	Medidas de resumen	
Concordante	2256	100,0	D de Somers	1,00
Discordante	0	0,0	Gamma de Goodman-Kruskal	1,00
Empates	0	0,0	Tau-a de Kendall	0,33
Total	2256	100,0		

**Regresión logística binaria: com 2 conglo vs. Biocapacidad; Reserva ecol**  
**Pegada Ecológica Total; Biocapacidade**  
Función de enlace: Logit

Información de respuesta

Variable	Valor	Conteo	
com 2 conglomerados	2	24	(Evento)
	1	94	
	Total	118	

Tabla de regresión logística

Predictor	Coef	SE Coef	Z	P	Relación de probabilidades	IC de 95%	
						Inferior	Superior
Constante	-9,62003	2,52019	-3,82	0,000			
Biocapacidade	2,24009	0,584047	3,84	0,000	9,39	2,99	29,51
Reserva ecológica	-2,70237	0,682955	-3,96	0,000	0,07	0,02	0,26

Log-verosimilitud = -12,817  
Probar que todas las pendientes son cero:  $G = 93,562$ ,  $GL = 2$ , valor  $P = 0,000$

Pruebas de bondad del ajuste

Método	Chi-cuadrada	GL	P
Pearson	166,562	101	0,000
Desviación	25,633	101	1,000
Hosmer-Lemeshow	3,675	8	0,885

Medidas de asociación:  
(Entre la variable de respuesta y las probabilidades pronosticadas)

Pares	Número	Porcentaje	Medidas de resumen	
Concordante	2229	98,8	D de Somers	0,98
Discordante	25	1,1	Gamma de Goodman-Kruskal	0,98
Empates	2	0,1	Tau-a de Kendall	0,32
Total	2256	100,0		

Cenário:Amostra 1	Análise Regressão Logística	
	Valor de P	Porcentagem concordante
Todas as variáveis	P variando de 0,999 a 0,992	100,0
Todas as variáveis, tirando população	P variando de 0,999 a 0,991	100,0
Variáveis escolhidas tomando-se como base a análise ANOVA(Pegada Ecológica Total,	P variando de 0,998 a 0,992	100,0

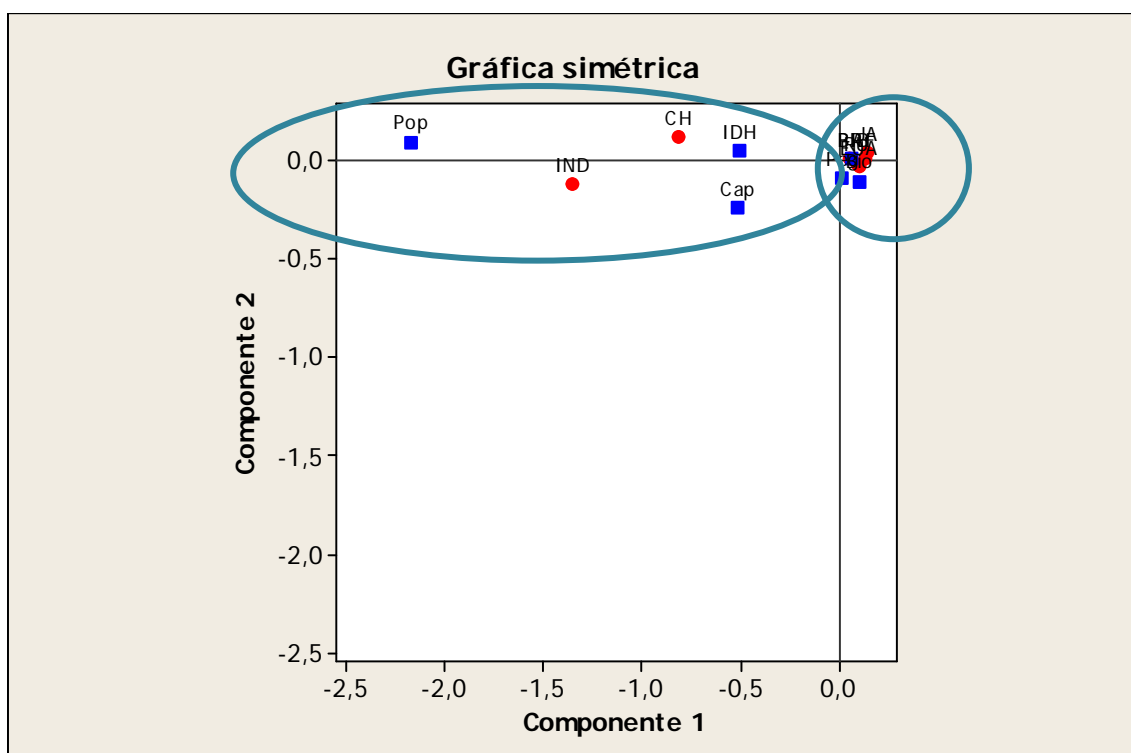
Biocapacidade, Reserva ecológica, HDI e PIB Per capta (tirando população e captação de água)		
Variáveis Pegada Ecológica Total e PIB Per capta (modelo parcimonioso – 2 variáveis)	P variando de 1,000 a 0,989	100,0
Variáveis escolhidas pela análise do dendograma: Pegada Ecológica Total; Biocapacidade	P = 0,000	98,8

A análise Regressão Logística, apresentou os mesmos resultados para os quatro primeiros testes ( 1) Todas as variáveis. 2) Todas as variáveis, tirando população. 3) Variáveis escolhidas tomando-se como base a análise ANOVA(Pegada Ecológica Total, Biocapacidade, Reserva ecológica, HDI e PIB Per capta (tirando população e captação de água) e 4)Variáveis Pegada Ecológica Total e PIB Per capta (modelo parcimonioso – 2 variáveis)). Nesses testes obtivemos valores de P que reprovam o teste, variando de 0,991 a 1,000. Igualmente a porcentagem de concordância (100,0) foi idêntica em todos os testes. Porém o teste realizado com as variáveis escolhidas pela análise do dendograma: Pegada Ecológica Total e Biocapacidade, pode ser aceito, pois o valor de P = zero. Neste teste, a porcentagem de concordância é 98,9%.

## 8. ANÁLISE DE CORRESPONDÊNCIA

Para a análise de correspondência, utilizou-se os países do BRICS (África do Sul, China, Índia, Rússia e Brasil), Japão e EUA e as seguintes variáveis População, Pegada Ecológica Total, Captações água por pessoa, Biocapacidade, IDH e PIB Per capta.

O Gráfico abaixo mostra a relação entre os países (linhas) e as variáveis (colunas):



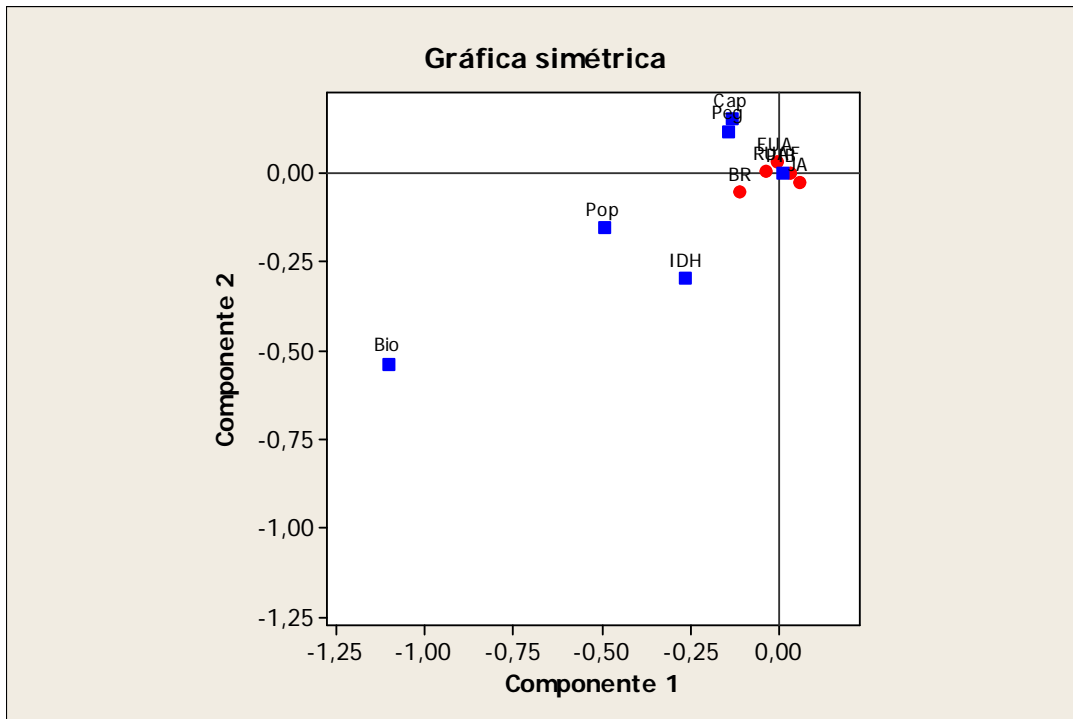
Resultado da análise de correspondência:

Pelo Gráfico acima nota-se uma proximidade entre os países analisados, com relação às variáveis estudadas, com um pequeno afastamento da Índia e China.

Também podemos observar um afastamento das seguintes variáveis: População, IDH e Captação..

Pela análise da tabela de contingência abaixo, vemos que 97,98% é contabilizada no primeiro componente, 1,77% é contabilizada no segundo componente.

O Gráfico abaixo vemos que tirando China e Índia, países ficam muito próximos, com a correspondência em torno do PIB Per Capta.



**Análisis de correspondencia simple: População; Pegada; Captações; Biocapacidad; IDH; PIB\_P\_capta**

La tabla de análisis de contingencia

Eje	Inercia	Proporción	Acumulada	Histograma
1	0,1239	0,9798	0,9798	*****
2	0,0022	0,0177	0,9975	
3	0,0003	0,0024	0,9999	
4	0,0000	0,0001	1,0000	
5	0,0000	0,0000	1,0000	
Total	0,1264			

Contribuciones de las filas

ID	Nombre	Cal	Total	Inercia	Componente 1			Componente 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AF	0,998	0,077	0,011	0,134	0,998	0,011	0,004	0,001	0,000
2	BR	0,561	0,084	0,004	0,056	0,552	0,002	0,007	0,009	0,002
3	CH	1,000	0,067	0,354	-0,812	0,980	0,354	0,116	0,020	0,402

4	EUA	0,995	0,359	0,038	0,110	0,922	0,035	-0,031	0,073	0,155
5	IND	1,000	0,037	0,537	-1,345	0,991	0,543	-0,125	0,009	0,260
6	RU	0,970	0,121	0,009	0,096	0,932	0,009	-0,020	0,039	0,021
7	JA	0,994	0,255	0,047	0,148	0,934	0,045	0,037	0,060	0,160

Contribuciones de columnas

ID	Nombre	Cal	Total	Inercia	Componente 1			Componente 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Pop	1,000	0,023	0,872	-2,169	0,998	0,888	0,092	0,002	0,088
2	Peg	0,166	0,000	0,000	0,021	0,008	0,000	-0,092	0,158	0,001
3	Cap	1,000	0,034	0,086	-0,513	0,817	0,072	-0,243	0,183	0,892
4	Bio	0,014	0,000	0,002	0,103	0,007	0,000	-0,109	0,007	0,001
5	IDH	0,615	0,000	0,000	-0,503	0,610	0,000	0,046	0,005	0,000
6	PIB	1,000	0,942	0,039	0,072	0,992	0,040	0,006	0,008	0,018

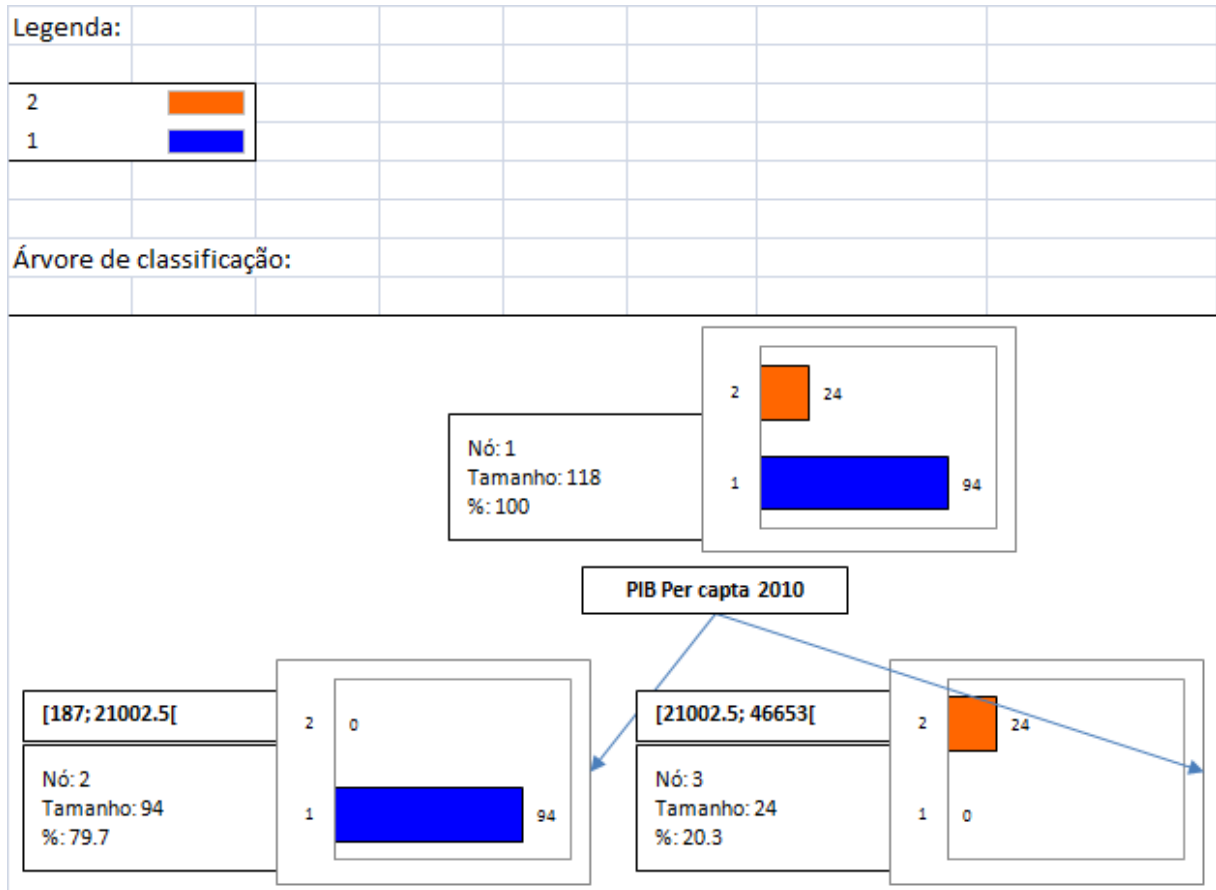
O país que mais contribui para o componente 1 é a Índia (0,543) e para o segundo componente é a China (0,402). A variável que mais contribui para o **componente 1** é População e para o **componente 2**, é Captação de água, sendo que Pegada Ecológica, Biodiversidade e IDH, contribuem com muito pouco para os dois componentes.

## 9. ÁRVORE DE CLASSIFICAÇÃO – XLSTAT

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

Estatísticas descritivas:			
Variável	Categorias	reqüência	%
com 2 con	1	94	79,661
	2	24	20,339

Nó	p-valor	Objetos	%	Nó pai	Filhos	Variável de separação	Valores	Pureza
1	1,000	118	100,00%		2; 3			79,66%
2	0,000	94	79,66%	1		PIB Per capita 2010	[187; 21002.5[	100,00%
3	0,000	24	20,34%	1		PIB Per capita 2010	[21002.5; 46653[	100,00%



Este conjunto de dados conta com 118 indivíduos assim distribuídos:

- 94 estão no cluster 1 que compõe o grupo de países com menor desenvolvimento sócio-econômico. Os países que compõem o BRICS estão nesse grupo.
- 24 países estão no cluster 2 que representa o grupo de países com maior desenvolvimento sócio-econômico.

Pelo gráfico acima foi possível observar que a separação dos grupos é feita pela variável PIB per capita.

Com 79,66% de acerto, 94 países com PIB per capita entre 187 e 21002,5 ficaram no segundo Nó, enquanto que os 24 países restantes ficaram no Nó 3.

Como análise complementar, apresentamos a árvore de classificação, sem a variável PIB Per Capta:

Estatísticas descritivas:			
Variável	Categorias	Frequências	%
com 2 conglomerado	1	94	79,661
	2	24	20,339



Estrutura da árvore:								
Nó	p-valor	Objetos	%	Nó pai	Filhos	Ável de separa	Valores	Pureza
1	0,921	118	100,00%		2, 3			79,66%
2	0,861	97	82,20%	1	4, 5	IDH 2003	[0.3, 0.885[	96,91%
3	0,000	21	17,80%	1		IDH 2003	[0.885, 0.95[	100,00%
4	0,577	4	3,39%	2	6, 7	Reserva ecológica	[-3.7, -1.95[	75,00%
5	0,000	93	78,81%	2		Reserva ecológica	[-1.95, 8.7[	100,00%
6	0,000	2	1,69%	4		Pegada Ecológica	[2.9, 3.85[	50,00%
7	0,000	2	1,69%	4		Pegada Ecológica	[3.85, 4.9[	100,00%

Sem o PIB Per Capta, o conjunto de dados foi distribuído da seguinte maneira:

- 97 ficaram no cluster 1 que compõe o grupo de países com menor desenvolvimento sócio-econômico. que possuem o IDH variando de 0,3 a 0,885.
- 21 países ficaram no cluster 2 que representa o grupo de países com maior desenvolvimento sócio-econômico, com IDH variando de 0,885 a 0,95.

Pelo gráfico abaixo foi possível observar que a separação dos grupos foi feita pela variável IDH.

Novamente, houve outra separação, agora pela Reserva Ecológica, com a seguinte configuração:

- 93 ficaram no nó 5, com Reserva Ecológica variando de -1,95 a 8,7 hectares globais por pessoa
- 4 países ficaram no nó 4,. com Reserva Ecológica variando de -3,7 a -1,95 hectares globais por pessoa

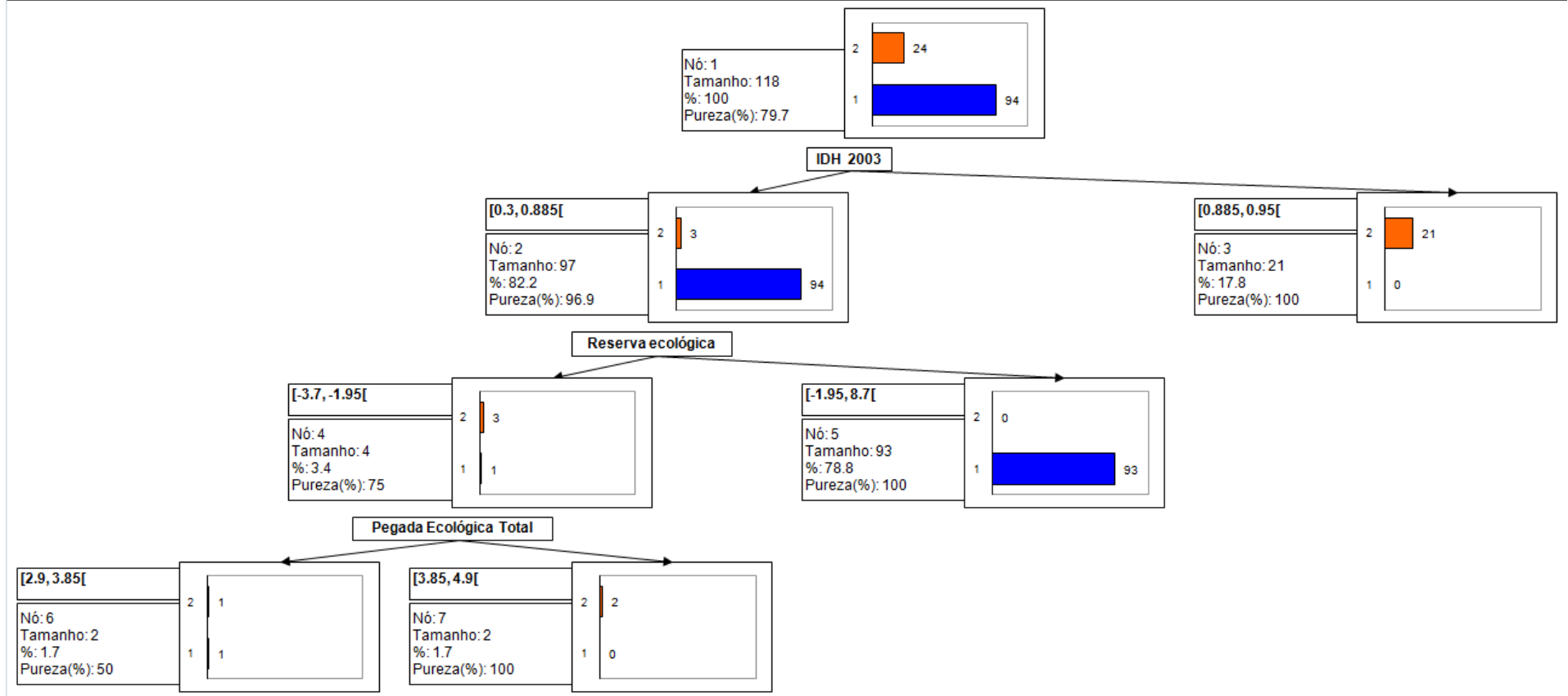
Este ramo foi novamente separado, agora pela variável Pegada Ecológica, com a seguinte configuração:

- 2 ficaram no nó 6, com Pegada Ecológica variando de 2,9 a 3,85 hectares globais por pessoa
- 2 países ficaram no nó 7,. com Reserva Ecológica variando de 3,85 a 4,9 hectares globais por pessoa

Legenda:



Árvore de classificação:



## 10. CONSIDERAÇÕES FINAIS

Pela análise de componente principal com as variáveis : Pegada Ecológica Total, Captações água por pessoa, Biocapacidade , Reserva ecológica , IDH e PIB Per capita 2010, percebemos que 42% destas cinco variáveis estudadas estão sendo explicadas pela variável número 1 (Pegada Ecológica Total). Se agregarmos a segunda variável, (Captações água por pessoa) veremos que a ela explica mais 25%, chegando-se a um acumulado de 67%.

Assim, ao invés de trabalharmos com um grupo grande de variáveis (5) poderíamos utilizar apenas o indicador PC1 ou os dois indicadores PC1 e PC2.

Cenário:	Modelo Quadrático	Modelo linear	Análise Regressão Logística	
			Valor de P	Porcentagem concordante
Todas as variáveis	Erro	1,000	P variando de 0,999 a 0,992	100,0
Todas as variáveis, tirando população	Erro	1,000	P variando de 0,999 a 0,991	100,0
Variáveis escolhidas tomando-se como base a análise ANOVA(Pegada Ecológica Total, Biocapacidade, Reserva ecológica, HDI e PIB Per capita (tirando população e captação de água)	Erro	1,000	P variando de 0,998 a 0,992	100,0
Variáveis Pegada Ecológica Total e PIB Per capita (modelo parcimonioso – 2 variáveis)	0,992	1,000	P variando de 1,000 a 0,989	100,0
Variáveis escolhidas pela análise do dendograma: Pegada Ecológica Total; Biocapacidade	0,949	0,949	P = 0,000	98,8

Foram realizados os mesmos testes para a **análise discriminante e análise Regressão logística**

Para a análise regressão logística, a conclusão é que mesmo no modelo parcimonioso em destaque, utilizando as Variáveis Pegada Ecológica Total e PIB Per capita, existe uma pequena diferença entre o modelo quadrático (para os testes que não apresentaram erro) e linear, não sendo significativa em termos estatísticos. Desta forma, opta-se pelo modelo linear, que envolve cálculos mais simples e que introduz menos erros no modelo.

Já para a análise Regressão logística, foram realizados os mesmos testes na, obtendo a porcentagem de concordância =100,0 em todos eles, inclusive no modelo parcimonioso em destaque. Porém esse valor deve ser desconsiderado, dado o resultado das análises dos valores de P que reprovam o teste para todas as amostras estudadas, variando de 0,989 a 1,000. Porém o teste realizado com as variáveis escolhidas pela análise do dendograma: Pegada Ecológica Total e Biocapacidade, pode ser aceito, pois o valor de P = zero. Neste teste, a porcentagem de concordância é 98,9%.

Desta forma, comparando as análises Regressão Logística e Discriminante, opta-se pela análise discriminante com o modelo linear, que envolve cálculos mais simples e com a possibilidade de introdução de menos erros no modelo.

Quanto à análise da árvore classificatória, concluímos que a melhor opção é a **análise discriminante**, que alcança 1,000% de proporção de acerto, enquanto que a “árvore classificatória” utiliza uma variável (PIB per Capita) para realizar a separação dos países, com 79,7% de acerto. Se tirarmos o PIB Per Capta, teremos a separação pela variável IDH, com 78,81% de acerto. A análise regressão logística apresenta erros que tornam os resultados não confiáveis.

Cenário:	Árvore classificatória PIB Per Capta	Árvore classificatória IDH	Análise DISCRIMINANTE - Modelo linear	Regressão Logística
Resultados	79,66	78,81	1,000	erro

Desta forma, continua-se a optar pela análise discriminante, utilizando o modelo linear, que envolve cálculos mais simples e que introduz menos erros no modelo.

Já na análise de correspondência, Para a análise de correspondência, utilizou-se os países do BRICS (África do Sul, China, Índia, Rússia e Brasil), Japão e EUA e as seguintes variáveis População, Pegada Ecológica Total, Captações água por pessoa, Biocapacidade, IDH e PIB Per capta.

Conclui-se que o país que mais contribui para o componente 1 é a Índia (0,543) e para o segundo componente é a China (0,402). A variável que mais contribui para o **componente 1** é População e para o **componente 2**, é Captação de água, sendo que Pegada Ecológica, Biodiversidade e IDH, contribuem com muito pouco para os dois componentes.

## Bibliografia

WWF–World Wide Fund For Nature. Relatório Planeta Vivo 2006 , Gland, Suíça.  
<http://www.wwf.org.br/informacoes/bliblioteca/?26162/Relatrio-Planeta-Vivo-2010> Acesso em 07/06/2011.