



PUC-SP

1946-2006

**PONTÍFICA UNIVERSIDADE CATÓLICA DE
SÃO PAULO**

**Faculdade de Economia, Administração,
Contabilidade e Atuariais**

**ANÁLISE DOS INDICADORES DO MEIO AMBIENTE DO
BANCO MUNDIAL-WDI**

Aluno: Antonio Thomaz P. Lessa Neto

Prof. Arnaldo José de Hoyos Guevara

1º Semestre 2012

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma **Árvore de Classificação** (que indica qual a variável que melhor separa os grupos e classifica as variáveis por ordem de importância na separação dos grupos) - ainda dentro da Análise Multivariada - usando os Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - que nos apresentam os últimos dados fundamentais para o desenvolvimento, extraídos do BANCO de DADOS/da BASE de DADOS principal do Banco Mundial - *the World Bank's* - e denominado Perfis dos países - *Country Profiles* - justamente com os dados de desenvolvimento comparáveis entre países.

Mais especificamente, iremos concentrar nosso estudo somente sobre os indicadores de **MEIO-AMBIENTE** (*Environment*) do Banco Mundial - *the World Bank's*.

(LINK: <http://data.worldbank.org/topic/environment>)

A **Árvore de Classificação** utilizará dados dimensionadores do **MEIO-AMBIENTE** (*Environment*) do Banco Mundial - *the World Bank's* - de 77 países do mundo, tendo como principal propósito a comparação entre 4 (quatro) regiões distintas do globo - a América Latina (com 23 países), a OECD (com 22 países), o Leste Europeu (com 22 países) e a África (com apenas 10 países) - para todos os 6 (seis) indicadores compilados pelo Banco Mundial e melhor detalhados no tópico seguinte.

O Banco Mundial (*The WORLD BANK*), com sede em Washington - DC e com mais de 9.000 funcionários em mais de 100 escritórios em todo o mundo, é uma fonte vital de assistência financeira e técnica aos países em desenvolvimento ao redor do mundo, desde sua fundação em 1944. Além de não ser um banco no sentido comum, mas uma parceria única para reduzir a pobreza e apoiar o desenvolvimento, compreende 2 (duas) instituições geridas por 188 países-membros: o Banco Internacional para Reconstrução e Desenvolvimento (BIRD) - *International Bank for Reconstruction and Development (IBRD)* - e a Associação Internacional de Desenvolvimento (AID) - *the International Development Association (IDA)* -. Enquanto o BIRD tem por objetivo reduzir a pobreza nos países mais pobres quanto à sua renda média e credibilidade, a AID já se concentra exclusivamente nos países mais pobres do mundo. Essas instituições são parte de um corpo maior conhecido como o Grupo do Banco Mundial (*World Bank Group*).

Para tal, iniciamos na **primeira parte** com o *entendimento dos dados, incluindo a definição dos indivíduos e das variáveis, suas classificações em variáveis categóricas ou quantitativas, os significados e unidades de medida*, além da apresentação da tabela de dados. Já na **segunda parte**, iniciamos com a *análise das medidas e gráficos da estatística descritiva*, onde analisamos cada uma das variáveis separadamente quanto a sua forma de distribuição, os valores atípicos, medidas de centro e dispersão, utilizando gráficos de barras, histogramas, *box-plot*, *dot-plot* e curvas de densidade, além de medidas numéricas como média, mediana, quartis, desvio-padrão, variância, intervalo de confiança. Na sequência - **terceira parte** - passamos para a *análise da relação entre variáveis, incluindo gráficos de dispersão, linhas de tendência, dendograma e matrizes de correlação*. Na **quarta parte** do trabalho buscamos *efetuar uma Árvore de Classificação* - ainda dentro da Análise Multivariada - destes dados de Meio-Ambiente coletados na BASE DE DADOS do Banco Mundial referente aos indicadores de Governança - *World Governance Indicators (WGI)* - do Banco Mundial (*World Data Bank*). Encerramos então com as considerações finais. Os softwares estatísticos utilizados são o **MINITAB 16** e o **XLSTAT**.

2. ENTENDENDO OS DADOS

Os indicadores de **MEIO-AMBIENTE** (*Environment*) do Banco Mundial - *the World Bank's* - têm como premissa que os recursos ambientais naturais e artificiais (criados pelo homem) - água fresca, ar limpo, florestas, pastagens, recursos marinhos e agro-ecossistemas - provêm o sustento e uma base para o desenvolvimento social e econômico. A necessidade de salvaguardar esses recursos cruza todas as fronteiras. Hoje, o Banco Mundial é um dos principais promotores e financiadores de adaptação ambiental em países em desenvolvimento. Os dados aqui cobrem florestas, biodiversidade, emissões e poluição. Outros indicadores relevantes para o ambiente são encontradas em páginas de dados para a Agricultura e Desenvolvimento Rural, Energia e Mineração, Infra-estrutura e Desenvolvimento Urbano.

Fonte: <http://data.worldbank.org/topic/environment>

2.1 Os Indivíduos

Os indivíduos desta análise são apenas 77 (setenta e sete) **países**, conforme já melhor explicado na fase introdutória acima. Os dados analisados de cada país são as variáveis que descrevemos a seguir. Os dados foram coletados no BANCO de DADOS do Banco Mundial referente aos Indicadores do Desenvolvimento Mundial - *World Development Indicators* (*WDI*) - do Banco Mundial (*World Data Bank*), e mais especificamente aos seus indicadores de **MEIO-AMBIENTE** (*Environment*).

Fontes: <http://data.worldbank.org/data-catalog/country-profiles> e mais especificamente o LINK: <http://data.worldbank.org/topic/environment>

2.2 As Variáveis

São **variáveis (indicadores)** desta pesquisa que constituem cada um dos respectivos 77 (setenta e sete) **países** selecionados. As mesmas serão bem melhor explicadas na Tabela 1 a seguir. Ressaltamos que todos os dados desta pesquisa são referentes ao ano de **2010**.

Fonte: <http://data.worldbank.org/topic/environment>

Tabela 1. As Variáveis

Variável	Significado	Tipo	Unidade de Medida
Country	É o nome do país.	Variável Categórica	N/A
Forest area (% of land area) Forest	Área de floresta é a terra sob grupo de árvores naturais ou plantadas de pelo menos 5 metros no local, se árvore produtiva ou não, e exclui grupo de árvores em sistemas de produção agrícola (por exemplo, em plantações de frutas e sistemas agroflorestais) e árvores em parques urbanos e jardins. Fontes: Food and Agriculture Organization, electronic files and web site. Catalog Sources World Development Indicators	Variável Quantitativa	% (percentua

<p><i>Bird species, threatened</i></p> <p><i>BirdsThreat</i></p>	<p>Os pássaros estão listados para os países incluídos dentro de sua reprodução ou faixas de invernada. Espécies ameaçadas são o número de espécies classificadas pela IUCN como ameaçadas de extinção, vulneráveis, raros, indeterminado, fora de perigo, ou insuficientemente conhecidos.</p> <p>Fontes: United Nations Environmental Program and the World Conservation Monitoring Centre, and International Union for Conservation of Nature, Red List of Threatened Species. Catalog Sources World Development Indicators.</p>	<p>Variável Quantitativa</p>	<p>Numeral inteiro (quantidade)</p>
<p><i>Plant species (higher), threatened</i></p> <p><i>PlantThreat</i></p>	<p>Plantas de maior porte são espécies nativas de plantas vasculares. Espécies ameaçadas são o número de espécies classificadas pela IUCN como ameaçadas de extinção, vulneráveis, raros, indeterminado, fora de perigo, ou insuficientemente conhecidos.</p> <p>Fontes: United Nations Environmental Program and the World Conservation Monitoring Centre, and International Union for Conservation of Nature, Red List of Threatened Species. Catalog Sources World Development Indicators.</p>	<p>Variável Quantitativa</p>	<p>Numeral inteiro (quantidade)</p>
<p><i>Mammal species, threatened</i></p> <p><i>MammalThreat</i></p>	<p>Espécies de mamíferos são mamíferos excluindo baleias e botos. Espécies ameaçadas são o número de espécies classificadas pela IUCN como ameaçadas de extinção, vulneráveis, raros, indeterminado, fora de perigo, ou insuficientemente conhecidos.</p> <p>Fontes: United Nations Environmental Program and the World Conservation Monitoring Centre, and International Union for Conservation of Nature, Red List of Threatened Species. Catalog Sources World Development Indicators.</p>	<p>Variável Quantitativa</p>	<p>Numeral inteiro (quantidade)</p>
<p><i>Terrestrial and marine protected areas (% of total territorial area)</i></p> <p><i>Ter&MarProtAreas</i></p>	<p>Áreas protegidas terrestres são áreas total ou parcialmente protegidas de pelo menos 1.000 hectares que são designados pelas autoridades nacionais, como reservas científicas com acesso público limitado, parques nacionais, monumentos naturais, reservas naturais ou santuários da vida selvagem, paisagens protegidas, e áreas de gestão, principalmente para o uso sustentável. Áreas marinhas protegidas são áreas de terreno entremarés ou sub-marés - e que cobrem de água e associados flora e fauna e as características históricas e culturais - que tenham sido reservados por lei ou por outros meios eficazes para proteger parte ou a totalidade do ambiente fechado. Sítios protegidos pela legislação local ou provincial, são excluídos.</p> <p>Fontes: United Nations Environmental Program and the World Conservation Monitoring Centre, as compiled by the World</p>	<p>Variável Quantitativa</p>	<p>% (percentual)</p>

	Resources Institute, based on data from national authorities, national legislation and international agreements. Catalog Sources World Development Indicators.		
Total natural resources rents (% of GDP) TotNatResRent	Receitas totais de recursos naturais – seria a soma das receitas do petróleo, de gás natural, receitas de carvão (duro e mole), receitas de minerais, e receitas florestais. Fontes: Estimates based on sources and methods described in "The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium" (World Bank, 2011). Catalog Sources World Development Indicators.	Variável Quantitativa	% (percentual)

2.3 A Tabela de Dados

Devido ao tamanho da tabela de dados, seria mais oportuno apresentarmos a planilha abaixo contendo todas as variáveis (indicadores) Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - do Banco Mundial (*World Data Bank*), e mais especificamente aos seus indicadores de **MEIO-AMBIENTE (Environment)**, de somente 77 (setenta e sete) **países**.

Os dados foram coletados no BANCO de DADOS do Banco Mundial referente aos Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - do Banco Mundial (*World Data Bank*), e mais especificamente aos seus indicadores de **MEIO-AMBIENTE (Environment)**.

Fontes: <http://data.worldbank.org/data-catalog/country-profiles> e mais especificamente o LINK:

<http://data.worldbank.org/topic/environment>



Os países foram inicialmente agrupados por região, a saber:

Região 1 = América Latina Região 2 = OECD Região 3 = Leste Asiático Região 4 = África					
ITEM	COUNTRY	REGIÃO	ITEM	COUNTRY	REGIÃO
1	Argentina	1	1	Denmark	2
2	Bolivia	1	2	Finland	2
3	Brazil	1	3	France	2
4	Chile	1	4	Germany	2
5	Colombia	1	5	Hungary	2
6	Costa Rica	1	6	Iceland	2

7	Cuba	1	7	Ireland	2
8	Dominican Republic	1	8	Israel	2
9	Ecuador	1	9	Italy	2
10	El Salvador	1	10	Japan	2
11	Guatemala	1	11	Korea, Rep.	2
12	Honduras	1	12	Netherlands	2
13	Jamaica	1	13	New Zealand	2
14	Mexico	1	14	Norway	2
15	Nicaragua	1	15	Poland	2
16	Panama	1	16	Portugal	2
17	Paraguay	1	17	Spain	2
18	Peru	1	18	Sweden	2
19	Puerto Rico	1	19	Switzerland	2
20	Suriname	1	20	Turkey	2
21	Trinidad and Tobago	1	21	United Kingdom	2
22	Uruguay	1	22	United States	2
23	Venezuela, RB	1			

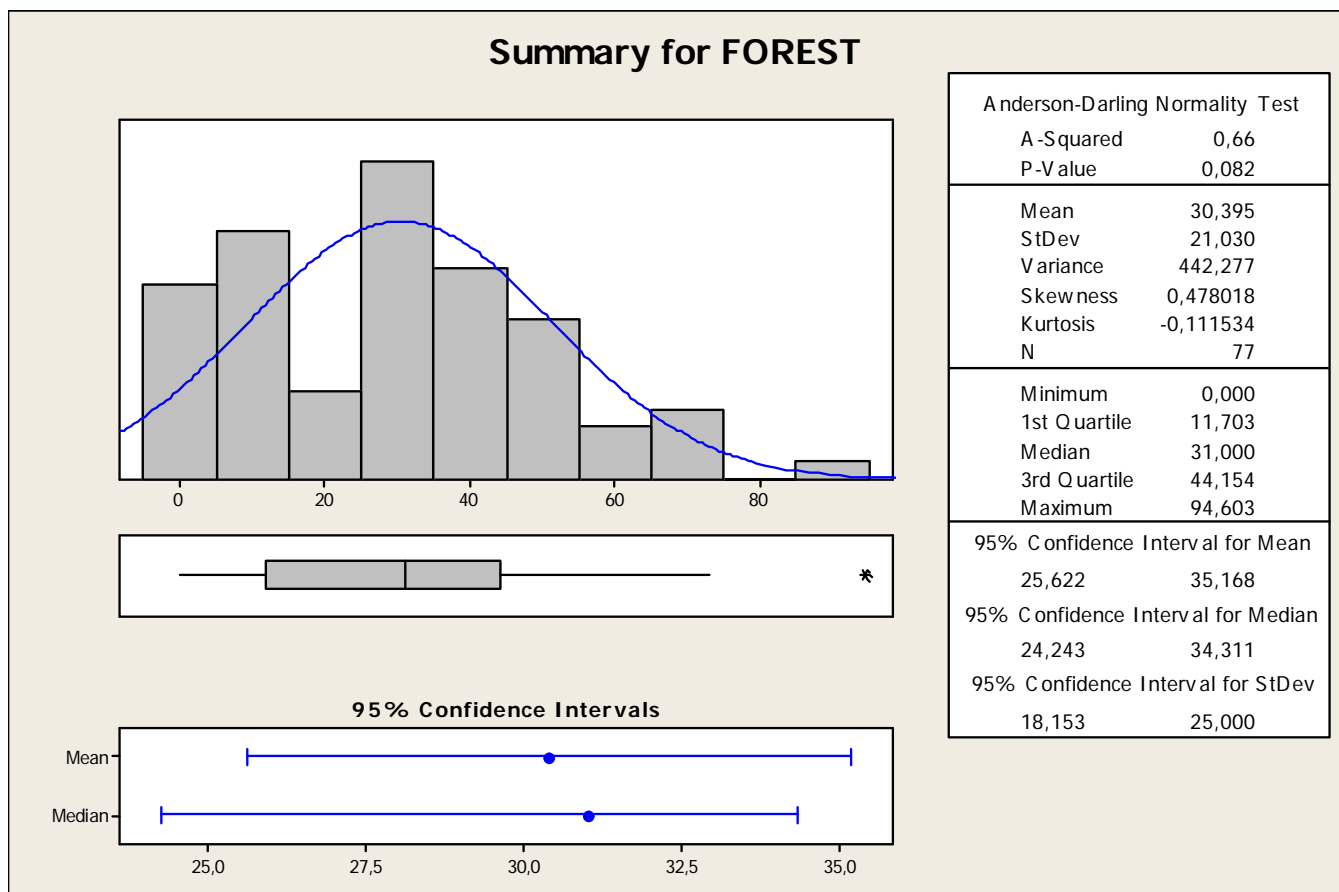
ITEM	COUNTRY	REGIÃO	ITEM	COUNTRY	REGIÃO
1	Bulgaria	3	1	Morocco	4
2	China	3	2	Mozambique	4
3	Egypt, Arab Rep.	3	3	Nigeria	4
4	India	3	4	Senegal	4
5	Indonesia	3	5	South Africa	4
6	Iran, Islamic Rep.	3	6	Sudan	4
7	Iraq	3	7	Tunisia	4
8	Jordan	3	8	Uganda	4
9	Lebanon	3	9	Zambia	4
10	Lithuania	3	10	Zimbabwe	4
11	Pakistan	3			
12	Philippines	3			
13	Romania	3			
14	Russian Federation	3			
15	Saudi Arabia	3			
16	Singapore	3			
17	Sri Lanka	3			
18	Thailand	3			
19	Ukraine	3			
20	United Arab Emirates	3			
21	Vietnam	3			
22	Yemen, Rep.	3			

3. ANÁLISE EXPLORATÓRIA

O nosso objetivo neste trabalho é encontrar regressões múltiplas que melhor expliquem os 6 (seis) indicadores de **MEIO-AMBIENTE** (*Environment*) disponíveis na BASE DE DADOS do Banco Mundial referente aos Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - do Banco Mundial (*World Data Bank*), e mais especificamente aos seus indicadores de **MEIO-AMBIENTE** (*Environment*), de somente **77 (setenta e sete) países**. Assim, focamos a análise em variáveis quantitativas utilizando Histograma, Box-Plot, Intervalo de Confiança, Média, Mediana, Mínimos e Máximos e Desvio-Padrão.

Iniciamos, então, com a **análise das medidas e gráficos da estatística descritiva**, onde analisamos cada uma das variáveis separadamente quanto a sua **forma de distribuição**, os **valores atípicos**, **medidas de centro e dispersão**, utilizando **gráficos de barras**, **histogramas**, **box-plot**, **dot-plot** e **curvas de densidade**, além de medidas numéricas como **média**, **mediana**, **quartis**, **desvio-padrão**, **variância**, **intervalo de confiança**.

As principais observações que podemos fazer são:



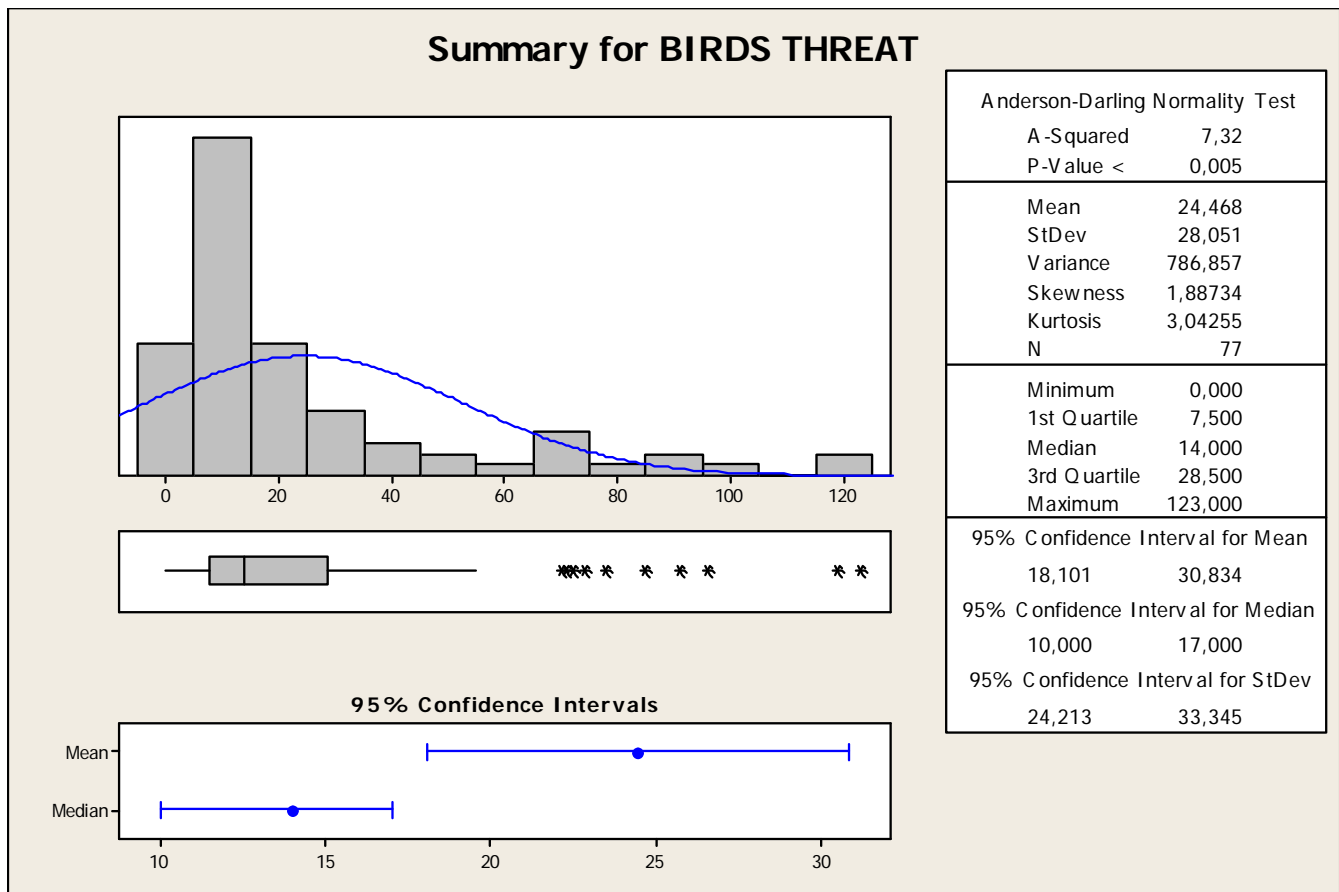
- **Forma:** A curva de densidade nos indica distribuição próxima da Normal para a variável **Forest area (% of land area) - Forest**, o que não está sendo corroborado pelo **P-Value** de 0,082 \Leftrightarrow 8,2% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da esquerda é relativamente longa, se comparada com a da direita.

A distribuição tem um ÚNICO pico mais destacado, e a posição da linha da mediana encontra-se relativamente próxima do centro dos quadrados principais do **Box-Plot**, apesar de estar mais próxima da respectiva linha do 3º Quartil.

- **Valores Atípicos:** Temos um ÚNICO valor desta variável que pode ser considerado como atípico no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países aqui analisados tem o indicador *Forest area (% of land area)* - **Forest** menor do que **31% (31,000%)** e a outra metade dos países têm este indicador maior do que este valor.

A média do indicador fica em **30,395% (aprox. 30%)**, com desvio-padrão (medida de dispersão) MUITO alto de **21,030% (aprox. 21%)**. A mediana está bem próxima da respectiva média.



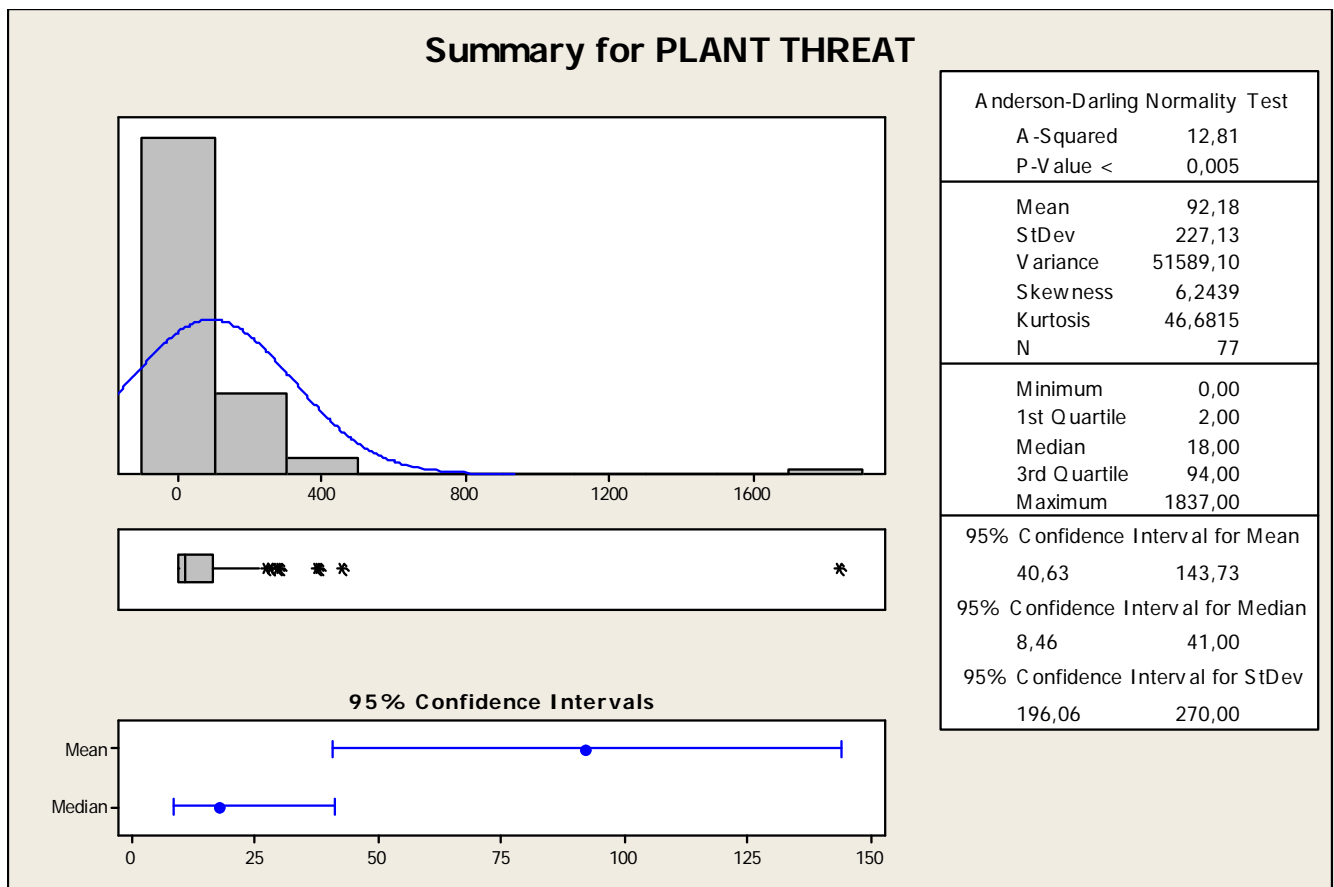
- **Forma:** A curva de densidade nos indica uma distribuição próxima da Normal para a variável *Bird species, threatened* - **BirdsThreat**, o que não é corroborado pelo **P-Value** de 0,084 \Leftrightarrow 8,4% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da esquerda é relativamente longa, se comparada com a da direita.

A distribuição tem um ÚNICO pico mais destacado, e a posição da linha da mediana encontra-se relativamente afastada do centro dos quadrados principais do **Box-Plot**, e está mais próxima da respectiva linha do 1º Quartil.

- **Valores Atípicos:** Temos 9 (nove) valores desta variável que podem ser considerados como atípicos no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países analisados tem o indicador *Bird species, threatened - BirdsThreat* menor do que **14** espécies de pássaro em risco de extinção e a outra metade dos países têm este indicador maior do que este valor.

A média do indicador fica em aproximadamente **24** espécies de pássaro ameaçadas de extinção com desvio-padrão (medida de dispersão) de aproximadamente **24** espécies para mais ou para menos, considerado MUITO ALTO. A mediana está bem afastada da respectiva média.



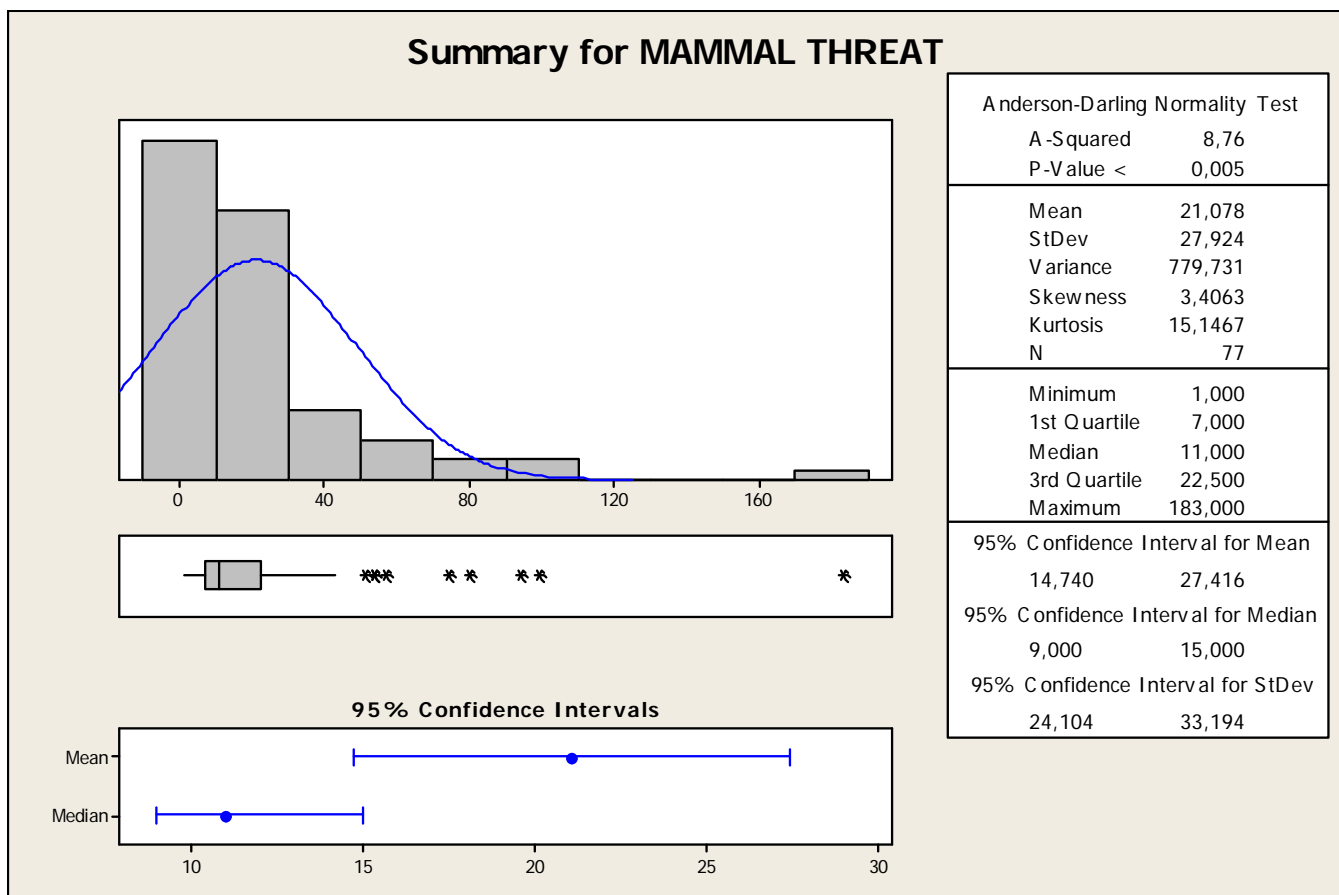
- **Forma:** A curva de densidade nos indica distribuição próxima da Normal para a variável *Plant species (higher), threatened - PlantThreat*, o que é corroborado pelo **P-Value** de 0,005 \Leftrightarrow 0,5% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da esquerda é visível e relativamente mais longa, se comparada com a da direita.

A distribuição tem um ÚNICO pico mai destacado, e a posição da linha da mediana encontra-se relativamente longe do centro dos quadrados principais do **Box-Plot**, mas está mais próxima da respectiva linha do 1º Quartil.

- **Valores Atípicos:** Temos 5 (cinco) valores desta variável que podem ser considerados como atípicos no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países analisados neste trabalho tem o indicador *Plant species (higher), threatened - PlantThreat* menor do que **18** espécies de planta em risco de extinção e a outra metade dos países têm este indicador maior do que este valor.

A média do indicador fica em aproximadamente **92** espécies de planta ameaçadas de extinção com desvio-padrão (medida de dispersão) totalmente fora da curva e absurdamente alto por volta de **227** (!!?) espécies para mais ou para menos. A mediana está muito afastada da respectiva média.



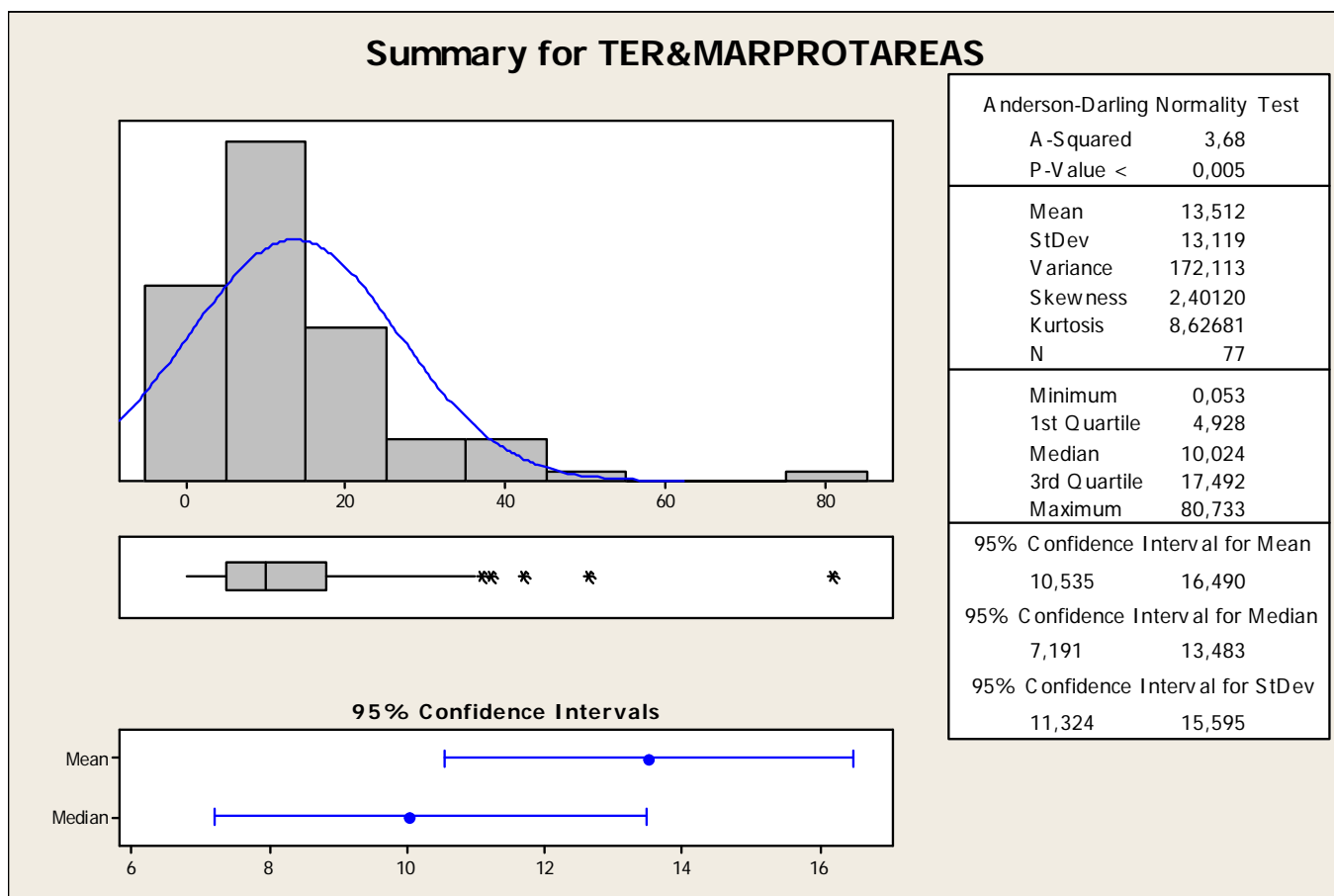
- **Forma:** A curva de densidade nos indica distribuição quase próxima da Normal para a variável *Mammal species, threatened - MammalThreat*, o que está sendo corroborado pelo **P-Value** de 0,005 \Leftrightarrow 0,5% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da esquerda está relativamente mais longa, se comparada com a da direita.

A distribuição tem um ÚNICO pico mais destacado, e a posição da linha da mediana encontra-se relativamente próxima do centro dos quadrados principais do **Box-Plot**, apesar de estar mais próxima da respectiva linha do 1º Quartil.

- **Valores Atípicos:** Temos 8 (oito) valores desta variável que podem ser considerados como atípicos no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países analisados neste estudo tem o indicador *Mammal species, threatened - MammalThreat* menor do que **11** espécies de mamífero sob risco de extinção e a outra metade dos países têm este indicador maior do que este valor.

A média do indicador fica em aproximadamente **21** espécies de mamífero ameaçadas de extinção com desvio-padrão (medida de dispersão) considerado bem alto de **28** espécies para mais ou para menos. A mediana está relativamente bem longe da respectiva média.



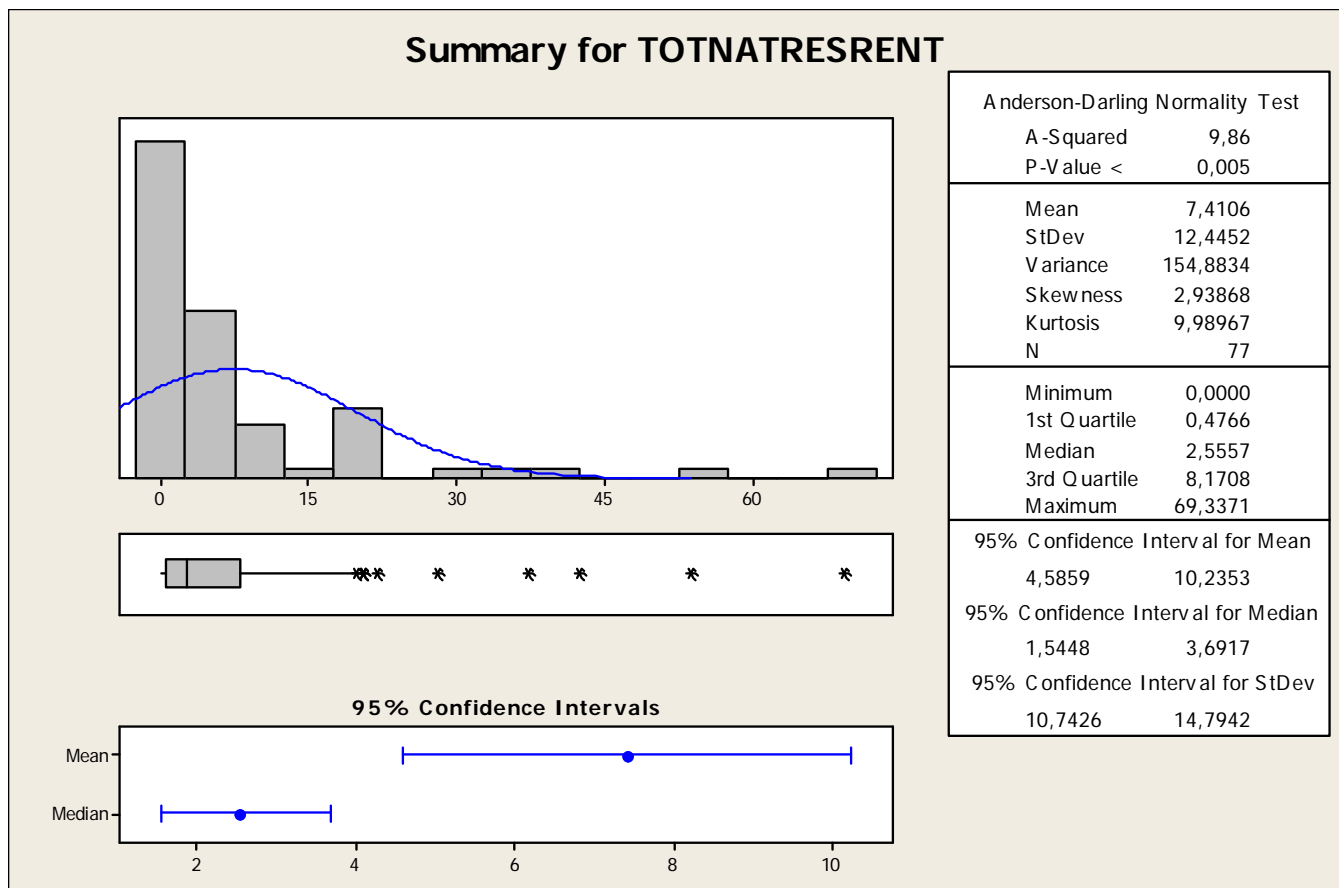
- **Forma:** A curva de densidade nos indica distribuição próxima da Normal para a variável *Terrestrial and marine protected areas (% of total territorial area) - Ter&MarProtAreas*, o que está sendo corroborado pelo **P-Value** de 0,005 \leq 0,5% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da esquerda é visível e relativamente mais longa, se comparada com a da direita.

A distribuição tem um ÚNICO pico mais destacado, e a posição da linha da mediana encontra-se longe do centro dos quadrados principais do **Box-Plot**, apesar de estar mais próxima da respectiva linha do 1º Quartil.

- **Valores Atípicos:** Temos 5 (cinco) valores desta variável que podem ser considerados como atípicos no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países têm *Terrestrial and marine protected areas (% of total territorial area) - Ter&MarProtAreas* menor do que **10,024%** e a outra metade dos países têm este indicador maior do que este valor.

A média do indicador fica em **13,512%**, com desvio-padrão (medida de dispersão) alto de também **13,119%**. A mediana é praticamente equivalente à respectiva média.



- **Forma:** A curva de densidade nos indica distribuição próxima da Normal para a variável *Total natural resources rents (% of GDP) - TotNatResRent*, o que é praticamente corroborado pelo **P-Value** de 0,005 \Leftrightarrow 0,5% (para este trabalho adotaremos o parâmetro de 5%) no teste de normalidade de Anderson-Darling.

A cauda da direita é relativamente mais longa, se comparada com a da esquerda.

A distribuição tem um ÚNICO pico mais destacado, e a posição da linha da mediana encontra-se relativamente longe do centro dos quadrados principais do **Box-Plot**, apesar de estar mais próxima da respectiva linha do 1º Quartil.

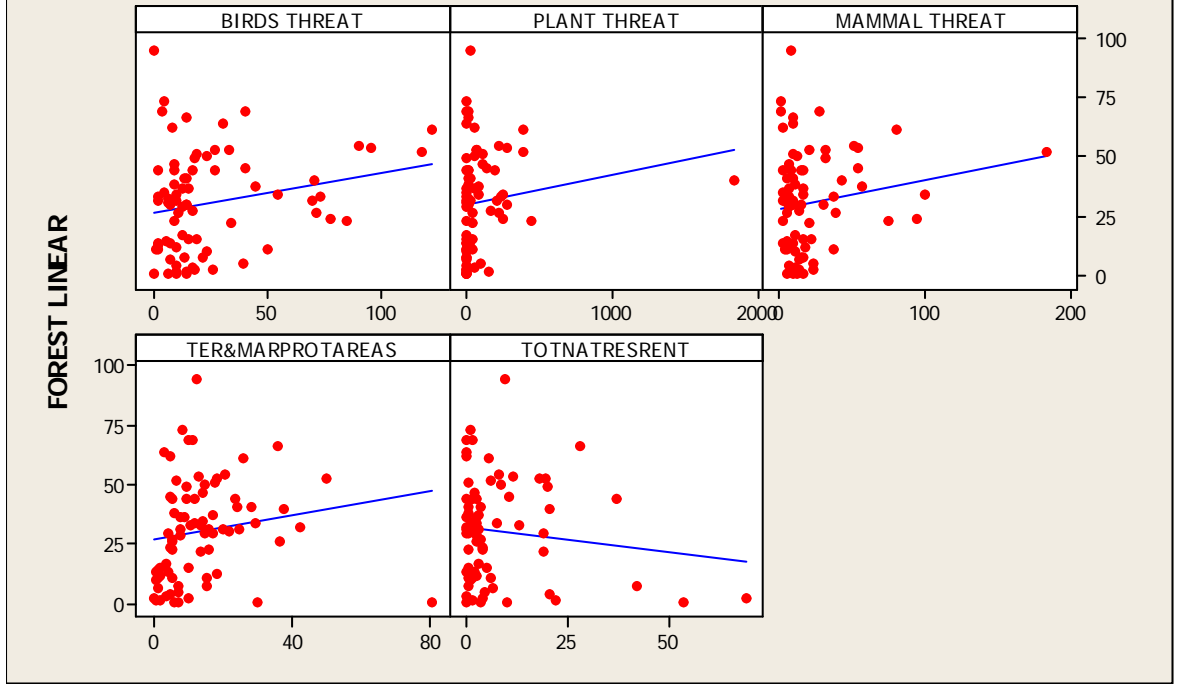
- **Valores Atípicos:** Temos 8 (oito) valores desta variável que podem ser considerados como atípicos no gráfico.
- **Centro e Dispersão:** A mediana nos indica que aproximadamente metade dos países tem *Total natural resources rents (% of GDP) - TotNatResRent* menor do que **2,5557%** e a outra metade dos países tem este indicador maior do que este valor.

A média do indicador fica em **7,4106%**, com desvio-padrão (medida de dispersão) bem alto de **12,4452%**. A mediana está relativamente próxima da respectiva média.

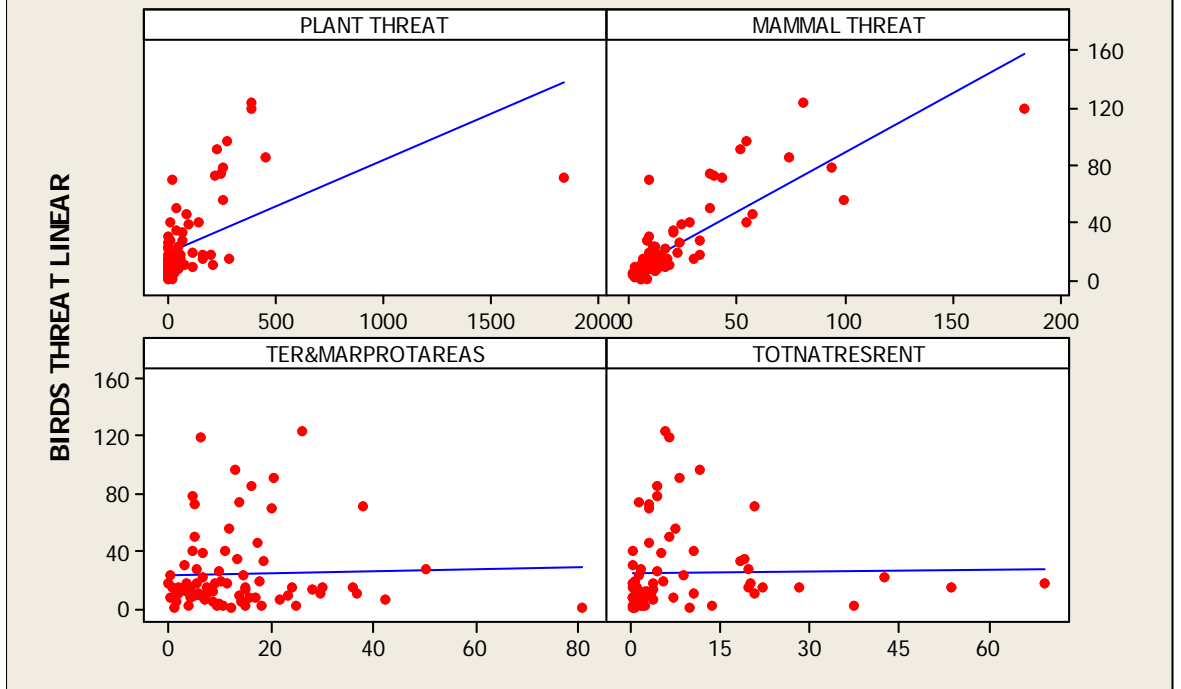
4. RELAÇÃO ENTRE VARIÁVEIS

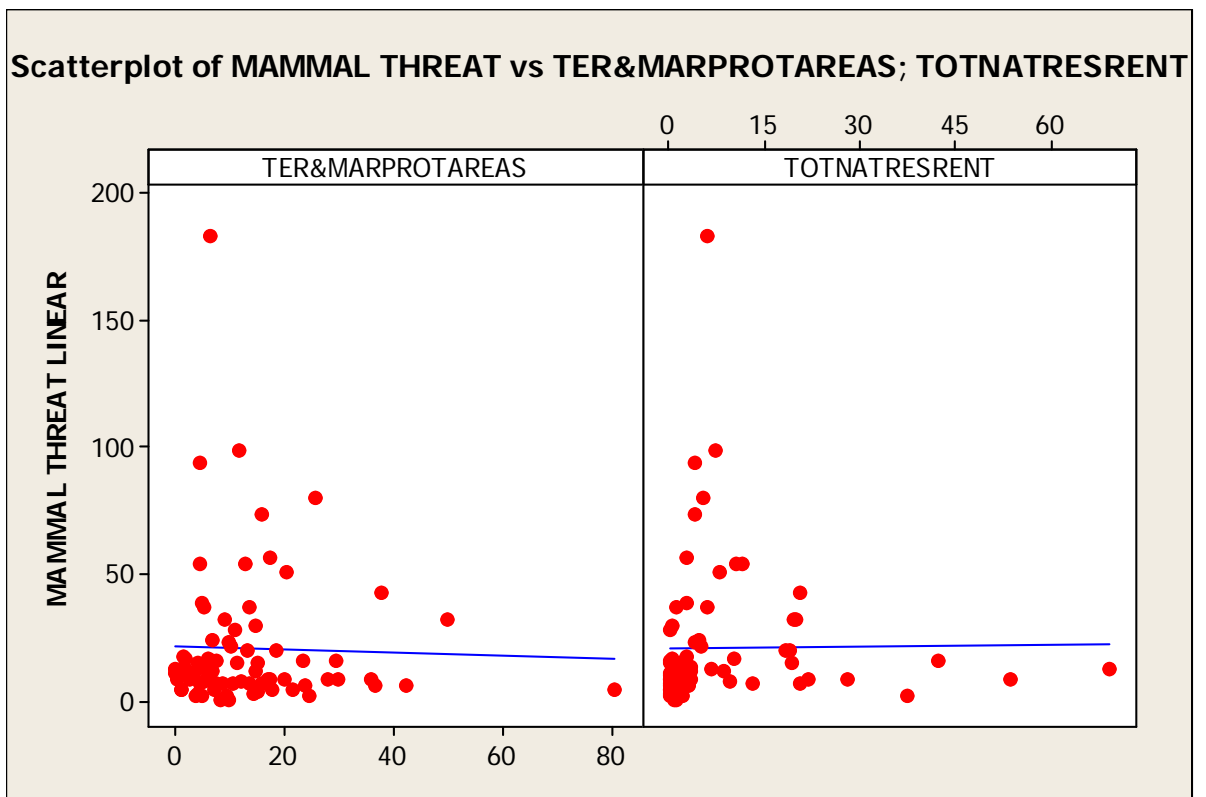
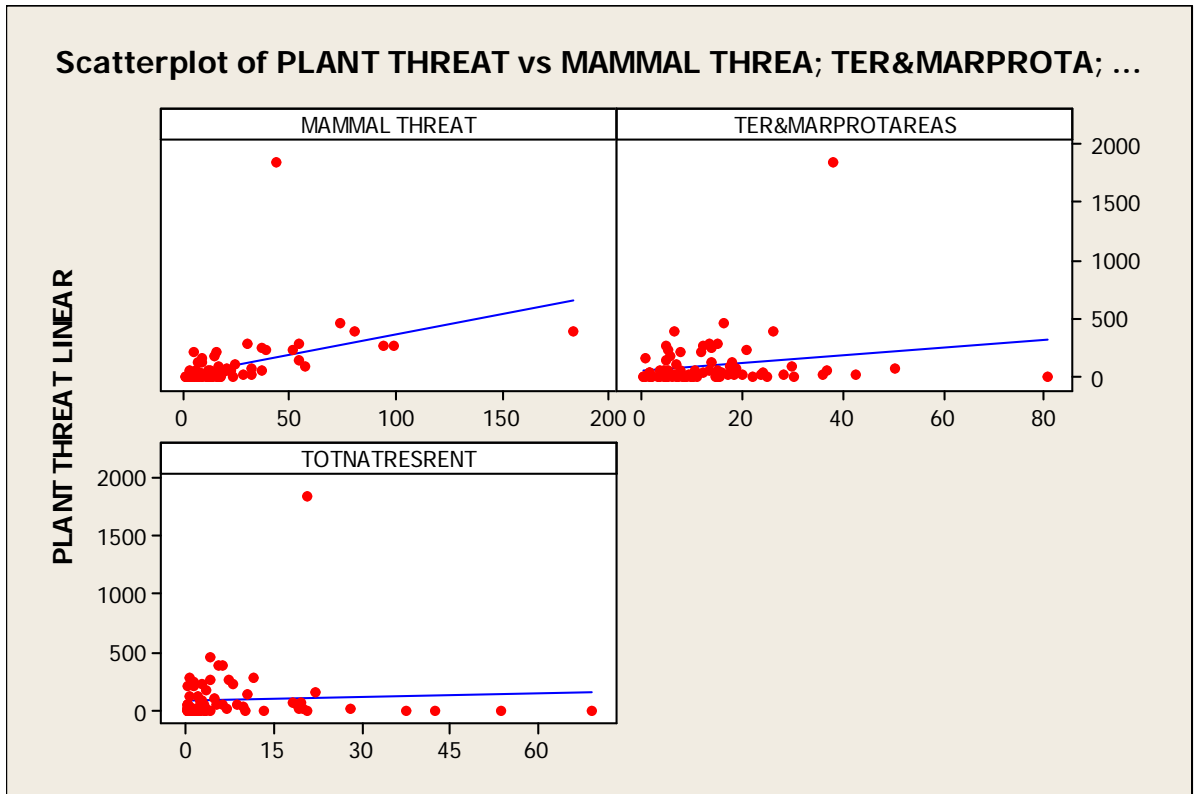
Passamos agora para a **análise da relação entre variáveis** onde analisaremos a seguir os **Gráficos de Dispersão, Linhas de Tendência, Dendograma e Matrizes de Correlação** das respectivas 6 (seis) variáveis (indicadores) de **MEIO-AMBIENTE (Environment)** - disponíveis na BASE DE DADOS do Banco Mundial referente aos Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - do Banco Mundial (*World Data Bank*), e mais especificamente aos seus indicadores de **MEIO-AMBIENTE (Environment)**, de somente 77 (setenta e sete) países.

Scatterplot of FOREST vs BIRDS THREAT; PLANT THREAT; MAMMAL THREA; ..



Scatterplot of BIRDS THREAT vs PLANT THREAT; MAMMAL THREA; ...





Forma: visualmente falando, os gráficos acima nos parecem denotar alguns tipos de formas características.

Direção: a análise dos diagramas acima nos permite fazer algumas poucas afirmações sobre as direções de associações, pois nos parece termos alguns poucos tipos de associações – quer positiva quer negativa – das variáveis analisadas.

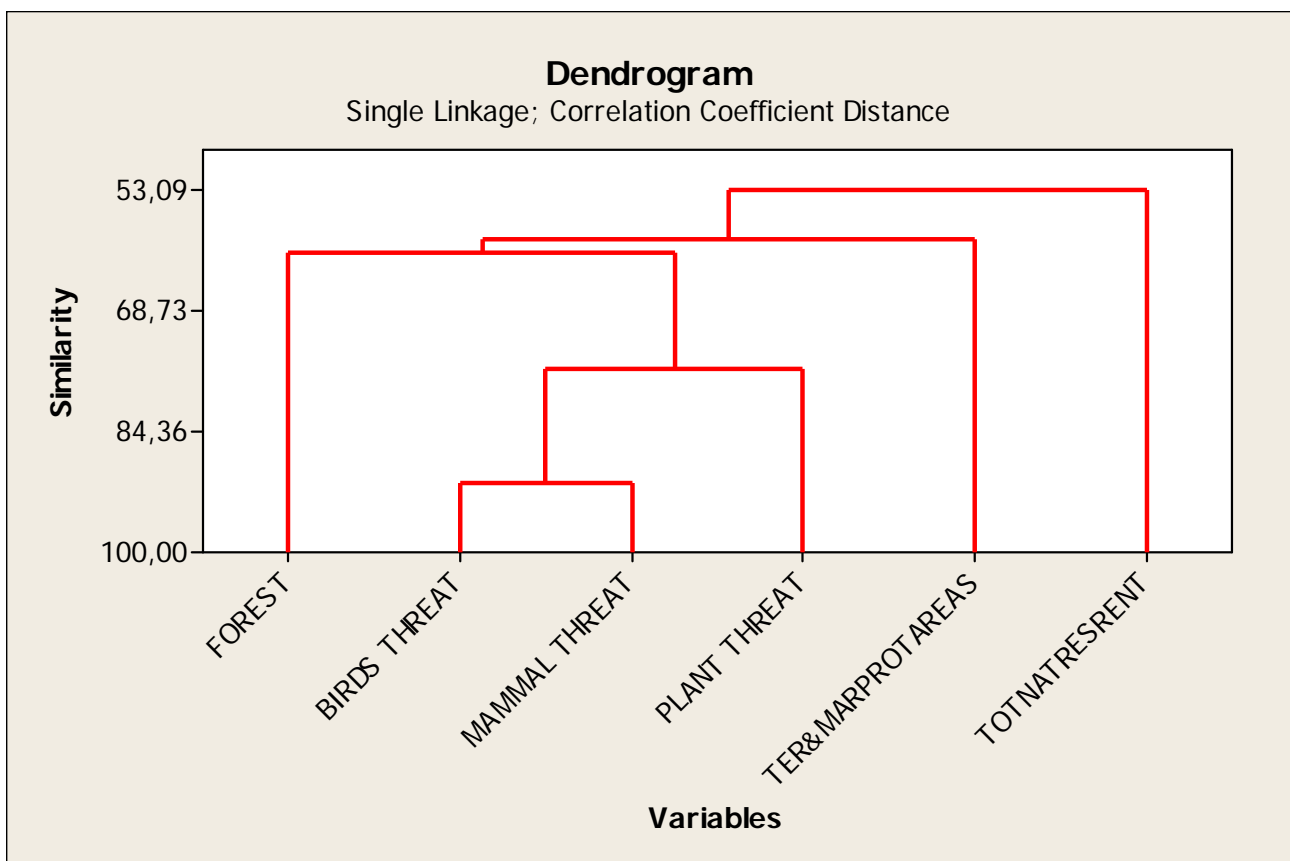
Intensidade: a intensidade de uma relação linear é mais facilmente observada mediante o cálculo de índices de correlação, que veremos mais adiante. Ocorre que muito pouco dos gráficos acima nos parecem indicar a existência de relações lineares.

Valores Atípicos: todos os gráficos indicam a existência de valores atípicos, ou seja, variáveis que estão localizadas longe das demais.

Os gráficos acima sugerem que algumas das combinações – tomadas 2 a 2 - possíveis entre estas 6 (seis) variáveis isoladamente pesquisadas neste projeto são efetivamente relações lineares, pois muitos pontos estão suficientemente próximos à linha de tendência.

Possíveis relações **LINEARES** visualizadas:

*Bird species, threatened - **BirdsThreat** e Mammal species, threatened - **MammalThreat***



Cluster Analysis of Variables: FOREST; BIRDS THREAT; PLANT THREAT; ...

Correlation Coefficient Distance, Single Linkage
Amalgamation Steps

Number
of obs.

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	in new cluster
1	5	90,9662	0,180675	2	4	2
2	4	76,2851	0,474297	2	3	2
3	3	61,1102	0,777795	1	2	1
4	2	59,4619	0,810762	1	5	1
5	1	53,0898	0,938204	1	6	1

Podemos concluir pelo **Dendograma** que as variáveis mais semelhantes são **BirdsThreat** e **MammalThreat**, confirmando a análise anterior. Seguidas pelas variáveis **PlantThreat** (unida às 2 anteriores), **Forest** (unida às 3 anteriores), e finalmente as 2 (duas) últimas **Ter&MarProtAreas** (unida às 4 anteriores), e **TotNatResRent** (unida às 5 anteriores).

Segue abaixo a **Matrizes 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: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREA; ...

	FOREST	BIRDS THREAT	PLANT THREAT
BIRDS THREAT	0,222 0,052		
PLANT THREAT	0,140 0,225	0,526 0,000	
MAMMAL THREAT	0,158 0,169	0,819 0,000	0,430 0,000
TER&MARPROTAREAS	0,159 0,167	0,026 0,825	0,189 0,099
TOTNATRESRENT	-0,121 0,294	0,019 0,868	0,062 0,593
TER&MARPROTAREAS		MAMMAL THREAT -0,027 0,814	TER&MARPROTAREAS 0,008 0,947
TOTNATRESRENT			0,029 0,802

Cell Contents: Pearson correlation
P-Value

- **$H_0: \rho = 0$, isto é, não há correlação entre as duas variáveis.**
- **$H_1: \rho \neq 0$, isto é, há correlação entre as duas variáveis.**

A correlação é sempre um número entre zero e um e mede a intensidade de relações lineares. A correlação entre as variáveis analisadas é positiva, mas de fraca intensidade.

Todas as correlações são significativas de acordo com o **P-Value** (para este trabalho consideraremos significativo **P-Value menor ou igual a 0,10**). Todas as 6 (seis) variáveis isoladamente possuem um coeficiente de correlação satisfatoriamente forte e positivo entre si,

indicando que, possivelmente, estas variáveis não podem entrar juntas no modelo de regressão, sob pena de diminuição do poder explicativo, bem como do poder de predição.

A correlação mais forte parece estar presente entre as variáveis - confirmando as análises anteriores no **Dendograma** acima -:

⇒ **BirdsThreat** e **MammalThreat** - (índice de Pearson = **0,819**);

⇒ **PlantThreat** e **BirdsThreat** - (índice de Pearson = **0,526**);

⇒ **PlantThreat** e **MammalThreat** - (índice de Pearson = **0,430**)

4.1 Análise EXTRA de Regressão + ANOVA + Stepwise

Regression Analysis: BIRDS THREAT versus FOREST; PLANT THREAT; ...

The regression equation is

$$\text{BIRDS THREAT} = 3,61 + 0,110 \text{ FOREST} + 0,0255 \text{ PLANT THREAT} + 0,721 \text{ MAMMAL THREAT} - 0,016 \text{ TER\&MARPROTAREAS} + 0,025 \text{ TOTNATRESRENT}$$

Predictor	Coef	SE Coef	T	P
Constant	3,612	3,835	0,94	0,349
FOREST	0,10995	0,08774	1,25	0,214
PLANT THREAT	0,025480	0,008928	2,85	0,006
MAMMAL THREAT	0,72052	0,07170	10,05	0,000
TER&MARPROTAREAS	-0,0156	0,1409	-0,11	0,912
TOTNATRESRENT	0,0254	0,1446	0,18	0,861

S = 15,5080 **R-Sq = 71,4%** R-Sq(adj) = 69,4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	42725,8	8545,2	35,53	0,000
Residual Error	71	17075,4	240,5		
Total	76	59801,2			

Source	DF	Seq SS
FOREST	1	2952,7
PLANT THREAT	1	14923,2
MAMMAL THREAT	1	24839,9
TER&MARPROTAREAS	1	2,6
TOTNATRESRENT	1	7,4

Unusual Observations

Obs	BIRDS		Fit	SE Fit	Residual	St Resid
	FOREST	THREAT				
3	61,4	123,00	77,60	4,87	45,40	3,08R
7	54,5	91,00	52,02	3,23	38,98	2,57R
12	39,7	71,00	85,70	14,69	-14,70	-2,95RX
21	0,3	0,00	5,99	10,51	-5,99	-0,53 X
23	52,1	119,00	151,26	10,73	-32,26	-2,88RX
25	1,9	18,00	14,95	9,35	3,05	0,25 X
35	33,3	55,00	85,11	5,38	-30,11	-2,07R
39	31,4	70,00	13,84	2,26	56,16	3,66R
46	53,1	96,00	55,42	3,31	40,58	2,68R
47	25,7	72,00	40,18	2,69	31,82	2,08R
53	0,5	14,00	11,12	7,56	2,88	0,21 X
71	33,2	74,00	39,98	2,42	34,02	2,22R

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

O **R-Quadrado** considera-se relativamente **ALTO**, de **71,4%** (maior do que 50%, porém ainda menor do que 100%). Todos os **Betas** apresentam valores positivos e maiores do que **zero** – **exceto a variável *Ter&MarProtAreas***, indicando algum poder explicativo da variável ***BirdsThreat***.

Tendo como parâmetro o **P-value** menor ou igual a **5% (0,050)** para a rejeição da hipótese nula - nenhuma das variáveis como fatores explicativos da variável ***BirdsThreat*** - vemos que apenas o **P-value** das variáveis ***MammalThreat*** e ***PlantThreat*** não nos permite descartá-las do modelo, pois são de fato significativas para a regressão.

Por outro lado, vemos que as variáveis ***Forest*** => ***P=0,214***, ***Ter&MarProtAreas*** => ***P=0,912*** e ***Voice and TotNatResRent*** => ***P=0,861*** nos permite descartá-las do modelo, pois nos parece que não seriam significativas para a regressão.

One-way ANOVA: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREA; ...

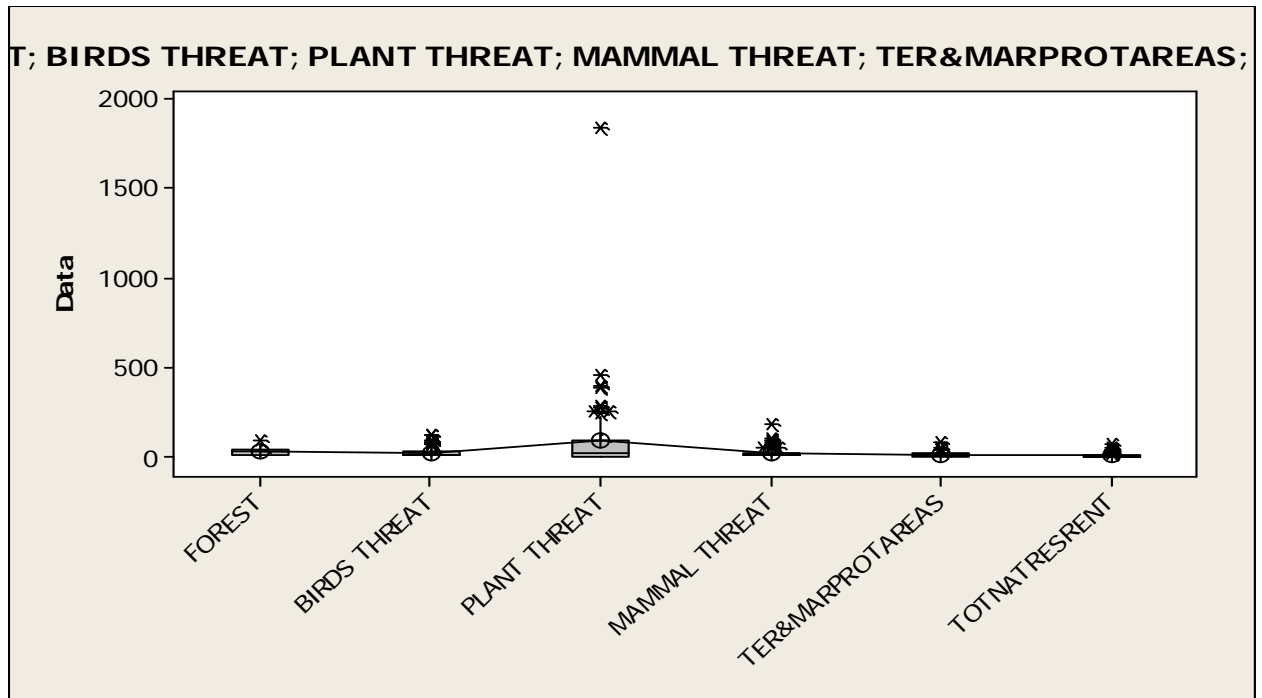
Source	DF	SS	MS	F	P
Factor	5	365399	73080	8,13	0,000
Error	456	4098297	8987		
Total	461	4463695			

S = 94,80 R-Sq = 8,19% R-Sq(adj) = 7,18%

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
FOREST	77	30,40	21,03	(-----*-----)
BIRDS THREAT	77	24,47	28,05	(-----*-----)
PLANT THREAT	77	92,18	227,13	(-----*-----)
MAMMAL THREAT	77	21,08	27,92	(-----*-----)
TER&MARPROTAREAS	77	13,51	13,12	(-----*-----)
TOTNATRESRENT	77	7,41	12,45	(-----*-----)

-----+-----+-----+-----+-----
0 35 70 105

Pooled StDev = 94,80



Stepwise Regression: BIRDS THREAT versus FOREST; PLANT THREAT; ...

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

Response is BIRDS THREAT on 5 predictors, with N = 77

Step	1	2
Constant	7,119	6,635
MAMMAL THREAT	0,823	0,731
T-Value	12,38	10,46
P-Value	0,000	0,000
PLANT THREAT		0,0263
T-Value		3,06
P-Value		0,003
S	16,2	15,4
R-Sq	67,13	70,81
R-Sq(adj)	66,69	70,02
Mallows Cp	8,7	1,6

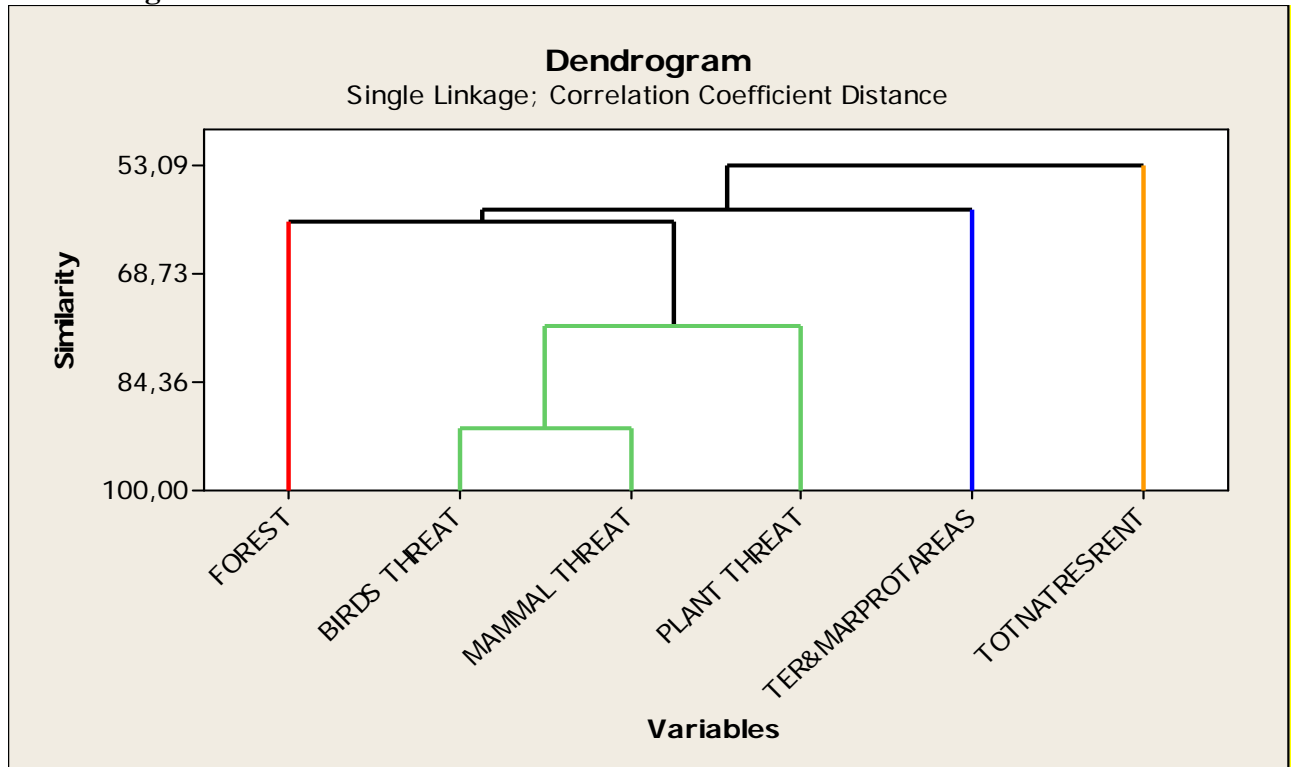
Da análise do **Stepwise Regression** acima, vemos claramente que a variável quantitativa independente **MammalThreat** responde - como *predictor* - por praticamente **67%** (**67,13**) do impacto sobre a variável resposta **BirdsThreat**. Também encontramos a 2ª mais significativa variável quantitativa independente, como sendo a **PlantThreat** e respondendo - também como *predictor* - por aproximadamente **4%** (**3,68**) do impacto sobre a variável resposta **BirdsThreat**.

5. ARVORE DE CLASSIFICAÇÃO

Primeiramente é feito uma **Análise de Cluster - com Dendograma (MÉTODO 8)** das 6 (seis) variáveis (indicadores) de Governança para identificarmos suas respectivas diferenciações. Em seguida efetuamos uma **Análise Discriminante (MÉTODO 9)** para vermos se as **REGIÕES** pré-definidas neste trabalho com nossa na **BASE de DADOS** do

Banco Mundial (*World Data Bank*) contendo os **77** (setenta e sete) **países**, realmente estão adequadas para as ditas 6 (seis) variáveis (indicadores) de **MEIO-AMBIENTE** (*Environment*) - disponíveis na BASE DE DADOS do Banco Mundial referente aos Indicadores do Desenvolvimento Mundial - *World Development Indicators (WDI)* - do Banco Mundial (*World Data Bank*). Antes de montarmos a **Árvore de Classificação** - via a outra ferramenta estatística o **XLSTAT** - iremos também apresentar uma **Regressão Logística Nominal** apenas das variáveis (indicadores) de **MEIO-AMBIENTE** (*Environment*) identificadas na **Análise de Cluster** com o maior nível de distanciamento.

5.1 Dendrograma



Cluster Analysis of Variables: FOREST; BIRDS THREAT; PLANT THREAT; ...

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	5	90,9662	0,180675	2	4	2
2	4	76,2851	0,474297	2	3	3
3	3	61,1102	0,777795	1	2	4
4	2	59,4619	0,810762	1	5	5
5	1	53,0898	0,938204	1	6	6

Final Partition

Cluster 1

FOREST

Cluster 2

BIRDS THREAT PLANT THREAT MAMMAL THREAT

Cluster 3

TER&MARPROTAREAS

Cluster 4

TOTNATRESRENT

Pelo **Dendograma** anterior iremos separar as variáveis pela diferenciação e estas terão seus respectivos indicadores verificados nas análises de classificação: **BirdsThreat**, **PlantThreat** e **MammalThreat** !

5.2 Análise Discriminante - dos 77 países (já agrupados por regiões) e seus respectivos 6 indicadores de Governança

Região 1 = América Latina
Região 2 = OECD
Região 3 = Leste Asiático
Região 4 = África

Discriminant Analysis: REGIÃO versus FOREST; BIRDS THREAT; ...

Linear Method for Response: REGIÃO

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT;
TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	23	22	22	10

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	10	5	1	2
2	6	11	3	2
3	3	0	14	5
4	4	6	4	1
Total N	23	22	22	10
N correct	10	11	14	1
Proportion	0,435	0,500	0,636	0,100

N = 77	N Correct = 36	Proportion Correct =
0,468		

Squared Distance Between Groups

	1	2	3	4
1	0,00000	1,20172	2,65336	1,16068
2	1,20172	0,00000	3,23292	0,67552
3	2,65336	3,23292	0,00000	1,13910
4	1,16068	0,67552	1,13910	0,00000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-3,3374	-2,0419	-1,9086	-1,7054
FOREST	0,0977	0,0761	0,0387	0,0638
BIRDS THREAT	0,0359	0,0258	0,0073	0,0086
PLANT THREAT	0,0007	-0,0019	-0,0014	-0,0018
MAMMAL THREAT	-0,0146	-0,0170	0,0419	0,0105
TER&MARPROTAREAS	0,0771	0,0939	0,0328	0,0713
TOTNATRESRENT	0,0554	0,0037	0,1037	0,0612

Summary of Misclassified Observations

Observation	True Group	Pred Group	Squared Distance	Probability
1**	1	3	5,120	0,120
		2	4,867	0,136
		3	2,269	0,498

			4	3,672	0,247
4**	3	2	1	1,2599	0,242
			2	0,4480	0,363
			3	3,1335	0,095
			4	0,8306	0,300
5**	1	3	1	2,646	0,206
			2	3,509	0,134
			3	1,567	0,354
			4	1,856	0,306
9**	1	4	1	1,600	0,254
			2	1,638	0,250
			3	2,148	0,193
			4	1,253	0,303
10**	2	4	1	3,760	0,115
			2	1,872	0,295
			3	2,473	0,219
			4	1,412	0,371
11**	1	2	1	1,8134	0,270
			2	0,6754	0,477
			3	6,1294	0,031
			4	2,2107	0,221
14**	1	4	1	3,961	0,122
			2	2,238	0,288
			3	2,651	0,234
			4	1,807	0,357
15**	2	1	1	5,470	0,452
			2	5,921	0,361
			3	12,019	0,017
			4	7,416	0,171
18**	1	2	1	2,778	0,232
			2	1,627	0,413
			3	5,440	0,061
			4	2,303	0,294
19**	1	2	1	1,300	0,349
			2	1,246	0,359
			3	5,171	0,050
			4	2,036	0,242
26**	2	4	1	4,807	0,096
			2	2,709	0,273
			3	2,801	0,260
			4	2,092	0,371
27**	2	4	1	4,341	0,096
			2	1,916	0,323
			3	2,680	0,220
			4	1,690	0,361
29**	1	2	1	2,051	0,310
			2	2,004	0,318
			3	4,287	0,101
			4	2,323	0,271
30**	2	1	1	3,321	0,507
			2	4,352	0,303
			3	8,708	0,034
			4	5,689	0,155
32**	2	1	1	4,280	0,486
			2	5,031	0,334
			3	9,961	0,028
			4	6,600	0,152
33**	3	4	1	4,337	0,104
			2	2,508	0,260
			3	2,368	0,279
			4	1,872	0,357
34**	3	2	1	1,5726	0,214
			2	0,2667	0,411
			3	3,7792	0,071
			4	0,8616	0,305
35**	1	3	1	15,400	0,051
			2	17,539	0,017
			3	9,837	0,818

			4	13,773	0,114
36**	4	3	1	4,547	0,089
			2	3,123	0,181
			3	1,597	0,388
			4	1,849	0,342
37**	4	1	1	0,7263	0,411
			2	1,4571	0,285
			3	4,3967	0,066
			4	1,8187	0,238
38**	2	4	1	3,767	0,108
			2	1,312	0,369
			3	3,081	0,152
			4	1,308	0,370
40**	1	2	1	4,726	0,194
			2	2,799	0,508
			3	8,030	0,037
			4	4,128	0,261
41**	4	3	1	5,842	0,063
			2	3,545	0,200
			3	2,286	0,375
			4	2,357	0,362
42**	2	4	1	2,0046	0,203
			2	1,9148	0,212
			3	2,0624	0,197
			4	0,7158	0,387
45**	1	2	1	2,015	0,350
			2	1,930	0,365
			3	5,421	0,064
			4	2,926	0,222
47**	3	1	1	5,486	0,394
			2	6,923	0,192
			3	6,490	0,238
			4	7,093	0,176
51**	3	2	1	1,5702	0,200
			2	0,4586	0,349
			3	2,4648	0,128
			4	0,6188	0,322
54**	4	2	1	2,227	0,251
			2	1,301	0,399
			3	5,161	0,058
			4	1,928	0,292
55**	3	4	1	5,089	0,101
			2	3,207	0,259
			3	2,834	0,311
			4	2,724	0,329
56**	4	3	1	4,711	0,132
			2	3,986	0,189
			3	2,486	0,400
			4	3,206	0,279
58**	3	4	1	3,627	0,255
			2	3,792	0,235
			3	4,130	0,199
			4	3,232	0,311
59**	4	3	1	2,981	0,170
			2	3,963	0,104
			3	1,288	0,395
			4	1,639	0,332
61**	2	1	1	4,609	0,421
			2	4,853	0,373
			3	10,717	0,020
			4	6,240	0,186
64**	1	4	1	8,405	0,258
			2	11,306	0,061
			3	7,888	0,335
			4	7,816	0,347
65**	4	3	1	5,134	0,071
			2	3,708	0,145
			3	1,472	0,444

			4	2,011	0,339
66**	2	4	1	3,750	0,111
			2	2,188	0,242
			3	1,717	0,306
			4	1,494	0,342
68**	3	4	1	2,9724	0,129
			2	1,5853	0,258
			3	1,5792	0,259
			4	0,9613	0,353
71**	2	1	1	5,064	0,511
			2	6,529	0,246
			3	8,308	0,101
			4	7,627	0,142
72**	1	4	1	4,513	0,124
			2	2,975	0,268
			3	2,832	0,288
			4	2,625	0,319
76**	4	1	1	8,091	0,553
			2	10,441	0,171
			3	13,391	0,039
			4	9,789	0,237
77**	4	2	1	2,133	0,264
			2	1,090	0,446
			3	5,989	0,038
			4	2,235	0,251

Discriminant Analysis: REGIÃO versus FOREST; BIRDS THREAT; ...

Quadratic Method for Response: REGIÃO

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT;
TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	23	22	22	10

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	9	2	3	0
2	8	19	7	0
3	5	1	11	0
4	1	0	1	10
Total N	23	22	22	10
N correct	9	19	11	10
Proportion	0,391	0,864	0,500	1,000

N = 77	N Correct = 49	Proportion Correct = 0,636
---------------	-----------------------	---

From Generalized Squared Distance to Group

Group	1	2	3	4
1	39,11	50,85	41,23	722,97
2	40,16	30,27	39,20	31,01
3	41,78	61,20	35,86	71,37
4	40,00	38,54	37,23	23,34

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	1	3	1	42,95	0,341
			2	54,91	0,001

			3	41,64	0,658
			4	537,72	0,000
2**	1	4	1	41,00	0,005
			2	74,19	0,000
			3	46,43	0,000
			4	30,24	0,995
4**	3	2	1	39,91	0,010
			2	30,76	0,973
			3	38,91	0,017
			4	68,46	0,000
5**	1	3	1	42,82	0,122
			2	76,66	0,000
			3	38,86	0,878
			4	93,50	0,000
6**	3	1	1	44,47	0,703
			2	123,20	0,000
			3	46,19	0,297
			4	3003,41	0,000
8**	1	2	1	40,50	0,434
			2	40,18	0,509
			3	44,58	0,056
			4	364,42	0,000
9**	1	3	1	40,43	0,395
			2	47,21	0,013
			3	39,62	0,592
			4	1009,50	0,000
11**	1	2	1	41,21	0,007
			2	31,59	0,888
			3	44,71	0,001
			4	35,88	0,104
13**	3	4	1	44,69	0,004
			2	48,05	0,001
			3	38,78	0,082
			4	33,96	0,913
14**	1	2	1	42,29	0,011
			2	33,50	0,894
			3	37,98	0,095
			4	66,73	0,000
18**	1	2	1	42,21	0,031
			2	35,30	0,967
			3	47,49	0,002
			4	346,08	0,000
30**	2	1	1	42,59	0,595
			2	43,37	0,402
			3	53,54	0,002
			4	672,73	0,000
31**	3	2	1	44,07	0,012
			2	35,59	0,836
			3	39,14	0,142
			4	44,41	0,010
33**	3	2	1	42,45	0,009
			2	33,40	0,853
			3	38,19	0,078
			4	38,74	0,059
34**	3	2	1	40,24	0,008
			2	30,90	0,876
			3	39,36	0,013
			4	35,17	0,103
35**	1	3	1	58,11	0,000
			2	211,81	0,000
			3	42,69	1,000
			4	1021,64	0,000
40**	1	2	1	45,46	0,002
			2	33,33	0,998
			3	55,52	0,000
			4	140,54	0,000
42**	2	3	1	40,80	0,146
			2	49,69	0,002

			3	38,55	0,452
			4	38,79	0,399
45**	1	2	1	40,80	0,011
			2	31,89	0,987
			3	44,59	0,002
			4	260,53	0,000
47**	3	1	1	43,07	0,795
			2	46,90	0,117
			3	47,47	0,088
			4	97,10	0,000
50**	1	2	1	42,88	0,064
			2	37,51	0,934
			3	50,15	0,002
			4	127,77	0,000
51**	3	2	1	40,20	0,009
			2	30,87	0,959
			3	37,66	0,032
			4	62,35	0,000
55**	3	2	1	43,65	0,010
			2	34,83	0,833
			3	38,17	0,157
			4	56,10	0,000
58**	3	1	1	42,28	0,952
			2	78,35	0,000
			3	48,26	0,048
			4	3825,25	0,000
64**	1	3	1	52,57	0,032
			2	204,62	0,000
			3	45,75	0,968
			4	87,22	0,000
68**	3	2	1	41,38	0,012
			2	32,71	0,887
			3	37,05	0,101
			4	60,22	0,000
71**	2	1	1	42,62	0,945
			2	49,15	0,036
			3	50,40	0,019
			4	183,48	0,000
72**	1	2	1	42,91	0,015
			2	34,71	0,888
			3	39,12	0,098
			4	216,25	0,000

O método de Análise de Distância apresenta no modelo **QUADRÁTICO** a melhor proporção de acerto: **63,6%** contra **46,8%** do **LINEAR**.

Segue abaixo a TABELA dos 77 países e respectivas REGIÕES - original e **NOVA** sugerida pela **Análise Discriminante** obtida pelos seus respectivos 6 indicadores de **MEIO-AMBIENTE (Environment)** através do método **QUADRÁTICO** com proporção de acerto de **63,6%**:

ITEM	COUNTRY	REGIÃO	NOVA REGIÃO LINEAR (46,8% acerto)	NOVA REGIÃO QUADRÁTICA (63,6% acerto)
1	Argentina	1	3	3
2	Bolivia	1	1	4
3	Brazil	1	1	1
4	Bulgaria	3	2	2
5	Chile	1	3	3
6	China	3	3	1
7	Colombia	1	1	1

8	Costa Rica	1	1	2
9	Cuba	1	4	3
10	Denmark	2	4	2
11	Dominican Republic	1	2	2
12	Ecuador	1	1	1
13	Egypt, Arab Rep.	3	3	4
14	El Salvador	1	4	2
15	Finland	2	1	2
16	France	2	2	2
17	Germany	2	2	2
18	Guatemala	1	2	2
19	Honduras	1	2	1
20	Hungary	2	2	2
21	Iceland	2	2	2
22	India	3	3	3
23	Indonesia	3	3	3
24	Iran, Islamic Rep.	3	3	3
25	Iraq	3	3	3
26	Ireland	2	4	2
27	Israel	2	4	2
28	Italy	2	2	2
29	Jamaica	1	2	1
30	Japan	2	1	1
31	Jordan	3	3	2
32	Korea, Rep.	2	1	2
33	Lebanon	3	4	2
34	Lithuania	3	2	2
35	Mexico	1	3	3
36	Morocco	4	3	4
37	Mozambique	4	1	4
38	Netherlands	2	4	2
39	New Zealand	2	2	2
40	Nicaragua	1	2	2
41	Nigeria	4	3	4
42	Norway	2	4	3
43	Pakistan	3	3	3
44	Panama	1	1	1
45	Paraguay	1	2	1
46	Peru	1	1	1
47	Philippines	3	1	1
48	Poland	2	2	2
49	Portugal	2	2	2
50	Puerto Rico	1	1	2
51	Romania	3	2	2
52	Russian Federation	3	3	3
53	Saudi Arabia	3	3	3

54	Senegal	4	2	4
55	Singapore	3	4	2
56	South Africa	4	3	4
57	Spain	2	2	2
58	Sri Lanka	3	4	1
59	Sudan	4	3	4
60	Suriname	1	1	1
61	Sweden	2	1	2
62	Switzerland	2	2	2
63	Thailand	3	3	3
64	Trinidad and Tobago	1	4	3
65	Tunisia	4	3	4
66	Turkey	2	4	2
67	Uganda	4	4	4
68	Ukraine	3	4	2
69	United Arab Emirates	3	3	3
70	United Kingdom	2	2	2
71	United States	2	1	1
72	Uruguay	1	4	2
73	Venezuela, RB	1	1	1
74	Vietnam	3	3	3
75	Yemen, Rep.	3	3	3
76	Zambia	4	1	4
77	Zimbabwe	4	2	4

5.3 Regressão Logística Nominal - dos 77 países (agrupamento por regiões ORIGINAL) e seus respectivos e somente 3 indicadores de **MEIO-AMBIENTE (Environment)** - **BirdsThreat**, **PlantThreat** e **MammalThreat** ! - obtidos acima pela **Análise de Cluster - com Dendograma (MÉTODO 8)**:

⇒ **NOMINAL (Nominal)** - quando não tem ordem entre os *clusters*

Nominal Logistic Regression: REGIÃO versus BIRDS THREAT; PLANT THREAT; ...

Response Information

Variable	Value	Count	
REGIÃO	4	10	(Reference Event)
	3	22	
	2	22	
	1	23	
	Total	77	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (3/4)							
Constant	0,244310	0,573607	0,43	0,670			
BIRDS THREAT	0,0000960	0,0426916	0,00	0,998	1,00	0,92	1,09
PLANT THREAT	0,0084547	0,0094059	0,90	0,369	1,01	0,99	1,03
MAMMAL THREAT	0,0050692	0,0423017	0,12	0,905	1,01	0,93	1,09
Logit 2: (2/4)							
Constant	1,41129	0,621506	2,27	0,023			

BIRDS THREAT	0,0505048	0,0463796	1,09	0,276	1,05	0,96	1,15
PLANT THREAT	0,0001557	0,0107654	0,01	0,988	1,00	0,98	1,02
MAMMAL THREAT	-0,118095	0,0623795	-1,89	0,058	0,89	0,79	1,00
Logit 3: (1/4)							
Constant	0,296393	0,584402	0,51	0,612			
BIRDS THREAT	0,0230172	0,0425965	0,54	0,589	1,02	0,94	1,11
PLANT THREAT	0,0138214	0,0092997	1,49	0,137	1,01	1,00	1,03
MAMMAL THREAT	-0,0434399	0,0451993	-0,96	0,337	0,96	0,88	1,05

Log-Likelihood = -89,436

Test that all slopes are zero: G = 27,778, DF = 9, P-Value = 0,001

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	218,548	216	0,439
Deviance	178,872	216	0,969

Agora apresentamos a mesma **Regressão Logística Nominal** - dos 77 países (agrupamento por regiões ORIGINAL, mas cobrindo TODOS e seus respectivos 6 (seis) indicadores de **MEIO-AMBIENTE (Environment)** - **BirdsThreat**, **PlantThreat**, **MammalThreat**, **TotNatResRent** e **Ter&MarProtAreas**.

Nominal Logistic Regression: REGIÃO versus FOREST; BIRDS THREAT; ...

Response Information

Variable	Value	Count	
REGIÃO	4	10	(Reference Event)
	3	22	
	2	22	
	1	23	
	Total	77	

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1: (3/4)							
Constant	0,893372	0,862086	1,04	0,300			
FOREST	-0,0286292	0,0246693	-1,16	0,246	0,97	0,93	1,02
BIRDS THREAT	-0,0038976	0,0449862	-0,09	0,931	1,00	0,91	1,09
PLANT THREAT	0,0119746	0,0103434	1,16	0,247	1,01	0,99	1,03
MAMMAL THREAT	0,0087425	0,0449961	0,19	0,846	1,01	0,92	1,10
TER&MARPROTAREAS	-0,0560683	0,0496923	-1,13	0,259	0,95	0,86	1,04
TOTNATRESRENT	0,0374798	0,0340555	1,10	0,271	1,04	0,97	1,11
Logit 2: (2/4)							
Constant	1,11713	0,982047	1,14	0,255			
FOREST	0,0143187	0,0226553	0,63	0,527	1,01	0,97	1,06
BIRDS THREAT	0,0436747	0,0489241	0,89	0,372	1,04	0,95	1,15
PLANT THREAT	-0,0040988	0,0119232	-0,34	0,731	1,00	0,97	1,02
MAMMAL THREAT	-0,0595672	0,0636021	-0,94	0,349	0,94	0,83	1,07
TER&MARPROTAREAS	0,0179708	0,0321530	0,56	0,576	1,02	0,96	1,08
TOTNATRESRENT	-0,303617	0,171512	-1,77	0,077	0,74	0,53	1,03
Logit 3: (1/4)							
Constant	-0,621540	0,945103	-0,66	0,511			
FOREST	0,0281992	0,0215039	1,31	0,190	1,03	0,99	1,07
BIRDS THREAT	0,0131580	0,0438438	0,30	0,764	1,01	0,93	1,10
PLANT THREAT	0,0144496	0,0102141	1,41	0,157	1,01	0,99	1,04
MAMMAL THREAT	-0,0361189	0,0458459	-0,79	0,431	0,96	0,88	1,06
TER&MARPROTAREAS	-0,0014939	0,0350447	-0,04	0,966	1,00	0,93	1,07
TOTNATRESRENT	-0,0004658	0,0396022	-0,01	0,991	1,00	0,92	1,08

Log-Likelihood = -72,278

Test that all slopes are zero: G = 62,094, DF = 18, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	208,128	210	0,524
Deviance	144,556	210	1,000

5.4 Árvore de Classificação – via XLSTAT

Esse aplicativo indica qual a variável que melhor separa os grupos e classifica as variáveis por ordem de importância na separação dos grupos.

Pelo **Dendograma** obtido no item 5.1 acima identificamos as variáveis pela diferenciação e estas possuem seus respectivos indicadores que seriam verificados nas análises de classificação: **BirdsThreat**, **PlantThreat** e **MammalThreat**

A seguir é demonstrado o teste desse modelo não só para as 3 (três) indicadores de **MEIO-AMBIENTE (Environment)**, mas sim para TODOS os 6 (seis): **BirdsThreat**, **PlantThreat**, **MammalThreat**, **TotNatResRent** e **Ter&MarProtAreas**.

Classification Tree

Warnings

Gain summary Tables are not displayed because profits are undefined.

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

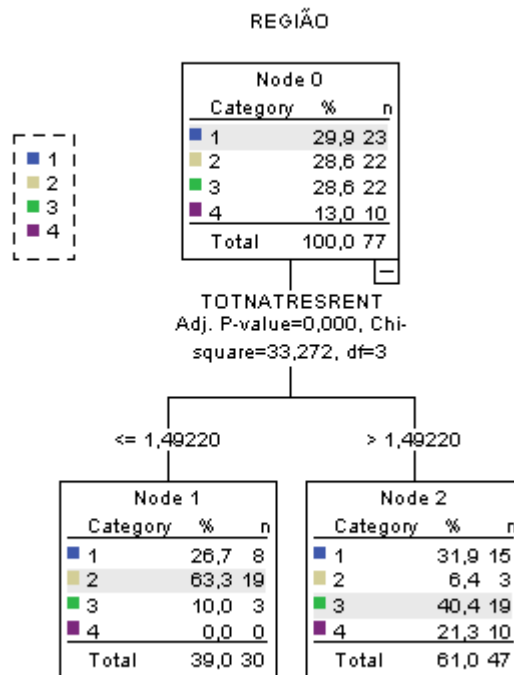
Model Summary

Specifications	Growing Method	CHAID		
	Dependent Variable	REGIÃO		
	Independent Variables	FOREST, TERMARPROTAREAS, TOTNATRESRENT, BIRDSTHREAT, PLANTTHREAT, MAMMALTHREA		
	Validation	None		
	Maximum Tree Depth		3	
	Minimum Cases in Parent Node		10	
	Minimum Cases in Child Node		5	
	Results	Independent Variables	TOTNATRESRENT	
		Included		
		Number of Nodes		3
Number of Terminal Nodes			2	
Depth			1	

Tree Table

Node	1		2		3		4		Total	
	N	Percent	N	Percent	N	Percent	N	Percent	N	
0	23	29,9%	22	28,6%	22	28,6%	10	13,0%	77	
1	8	26,7%	19	63,3%	3	10,0%	0	,0%	30	
2	15	31,9%	3	6,4%	19	40,4%	10	21,3%	47	

Growing Method: CHAID
 Dependent Variable: REGIÃO
 a. Bonferroni adjusted



Risk

Estimate	Std. Error
,506	,057

Growing Method: CHAID
 Dependent Variable:
 REGIÃO

Região 1 = América Latina

Região 2 = OECD

Região 3 = Leste Asiático

Região 4 = África

Classification					
Observed	Predicted				Percent Correct
	1	2	3	4	
1	0	8	15	0	,0%
2	0	19	3	0	86,4%
3	0	3	19	0	86,4%
4	0	0	10	0	,0%
Overall Percentage	,0%	39,0%	61,0%	,0%	49,4%

Growing Method: CHAID

Dependent Variable: REGIÃO

6. CONSIDERAÇÕES FINAIS

Dada a diversidade dos 77 (setenta e sete) países observados para cada uma das 6 (seis) variáveis (indicadores) eleitos para este estudo, parece haver boas oportunidades de agrupamento das observações.

Nos cenários em que nos dedicamos ao estudo específico de índices, dados e até estatísticas utilizadas para posicionarmos alguns países dentro de algumas poucas regiões num *ranking*, não é possível, com análises estatísticas mais estruturadas como Análise Exploratória de Dados, Análise Discriminante e Regressão, termos sucesso na determinação exata dos agrupamentos que nos ajudem a explicar, com clareza, os possíveis pontos em comum existentes entre esses ditos países e os seus respectivos índices.

Senão, vejamos como exemplo os tópicos acima, onde, depois de reordenarmos os países por regiões mais afins sugeridas pela **Análise Discriminante** e buscarmos identificar pela **Análise de Regressão (Stepwise)**, poderíamos chegar a uma conclusão de que os indicadores *Bird species, threatened* - **BirdsThreat**, *Mammal species, threatened* - **MammalThreat** e *Plant species (higher), threatened* - **PlantThreat** identificariam melhor e mais adequadamente países/regiões aqui analisadas.

Contudo, observamos que, através da **Árvore de Classificação**, foi possível fazer um agrupamento bem mais claro, nos permitindo assim entender melhor os pontos em comum dos países, como segue:

⇒ Indicador *Total natural resources rents (% of GDP)* - **TotNatResRent**

a. **Nó 1:** OECD (63% dos 30 países neste nó) - 19 países com **TotNatResRent** < = 1,49220

b. **Nó 2:** Leste Asiático (40% dos 47 países neste nó) - 19 países com **TotNatResRent** > 1,49220

Tal resultado apresentado pela **Árvore de Classificação**, nos mostra a verdadeira dependência de todos os países destas regiões aqui analisadas das Receitas Totais de Recursos Naturais (**TotNatResRen**) - a soma das receitas do petróleo, de gás natural, receitas de carvão (duro e mole), receitas de minerais, e receitas florestais - e que a preocupação da

PRESERVAÇÃO AMBIENTAL e dos seus RECURSOS NATURAIS ainda disponíveis neste nosso pequeno, já LOTADO e quase esgotado planeta TERRA, merece uma enorme e imediata atenção, não só com a criação de uma AGENDA GLOBAL para o MEIO-AMBIENTE agressiva e plenamente mensurável através de métricas realistas, mas também aproveitando os exemplos disponíveis de muita INOVAÇÃO SUSTENTÁVEL já desenvolvida ao redor do mundo para os mais diferentes problemas AMBIENTAIS da nossa humanidade.

Como exemplo, apresento aqui uma excelente reportagem apresentada durante o programa do FANTÁSTICO de 17/06/12: veja o vídeo pelo link (vale a pena !!!)

<http://fantastico.globo.com/Jornalismo/FANT/0,,MUL1681006-15605,00-CHINA+VAI+CONSTRUIR+UMA+CIDADE+DO+TAMANHO+DO+RJ+E+SP+POR+AN+O+ATE.html>



Outra fonte muito boa de notícias e informação pode-se encontrar no JORNAL VALOR ECONÔMICO, tanto no formato impresso quanto o digital (<http://www.valor.com.br/>)

Artigos interessantes:

- (1) <http://www.valor.com.br/opiniao/2675994/orientando-para-sustentabilidade>
- (2) <http://www.valor.com.br/empresas/2727610/bb-empresta-r-1-bi-para-agricultura-de-baixo-carbono>
- (3) <http://www.valor.com.br/rio20/2726128/sem-acoes-conferencia-delineou-agenda>
- (4) <http://www.valor.com.br/rio20/2725294/pnuma-precisa-de-maior-autonomia-diz-diretor-da-entidade>
- (5) <http://www.valor.com.br/rio20/2723820/agronegocio-ja-pode-crescer-sem-desmatar>
- (6) <http://www.valor.com.br/rio20/2723812/preservacao-pode-ser-uma-fonte-de-novos-negocios>
- (7) <http://www.valor.com.br/rio20/2723822/sp-tem-catalogo-de-biodiversidade>
- (8) <http://www.valor.com.br/rio20/2724270/mundo-pode-poupar-us-100-bi-com-luz-verde>
- (9) <http://www.valor.com.br/rio20/2724200/o-futuro-que-queremos-continua-incerto-em-meio-promessas-vagas>
- (10) <http://www.valor.com.br/rio20/2723716/secretario-geral-da-ocde-quer-taxar-emissores-de-carbono>

APÊNDICE

Análise Discriminante - dos 77 países (já agrupados por regiões) e seus respectivos 6 indicadores de **MEIO-AMBIENTE** (*Environment*).

Discriminant Analysis: NOVA REGIÃO L versus FOREST; BIRDS THREAT; ...

Linear Method for Response: NOVA REGIÃO L - Definida pelo MODELO LINEAR(46,80% de acerto no item 5.2 acima)

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT; TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	18	22	22	15

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	16	2	0	0
2	2	18	0	0
3	0	0	14	1
4	0	2	8	14
Total N	18	22	22	15
N correct	16	18	14	14
Proportion	0,889	0,818	0,636	0,933

N = 77

N Correct = 62

Proportion Correct = 0,805

Squared Distance Between Groups

	1	2	3	4
1	0,0000	6,7453	26,3960	18,3869
2	6,7453	0,0000	13,6682	5,5505
3	26,3960	13,6682	0,0000	3,5099
4	18,3869	5,5505	3,5099	0,0000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-16,667	-6,908	-3,024	-1,267
FOREST	0,448	0,270	0,053	0,117
BIRDS THREAT	0,256	0,136	0,026	0,046
PLANT THREAT	0,004	0,000	-0,002	0,000
MAMMAL THREAT	-0,285	-0,171	0,051	-0,045
TER&MARPROTAREAS	0,238	0,238	0,028	0,069
TOTNATRESRENT	-0,041	-0,076	0,153	0,016

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
20**	2	4	1	12,684	0,002
			2	2,878	0,295
			3	7,840	0,025
			4	1,211	0,678
31**	3	4	1	34,289	0,000
			2	14,722	0,002
			3	3,939	0,344
			4	2,657	0,654
36**	3	4	1	28,479	0,000
			2	11,744	0,004
			3	2,775	0,321
			4	1,289	0,675
39**	2	1	1	12,44	0,966
			2	19,15	0,034

			3	43,71	0,000
			4	34,03	0,000
41**	3	4	1	34,821	0,000
			2	14,256	0,002
			3	3,933	0,362
			4	2,807	0,636
43**	3	4	1	29,162	0,000
			2	11,709	0,005
			3	2,725	0,477
			4	2,559	0,518
44**	1	2	1	4,160	0,265
			2	2,262	0,685
			3	16,359	0,001
			4	7,544	0,049
45**	2	1	1	2,873	0,616
			2	3,889	0,371
			3	19,473	0,000
			4	10,601	0,013
47**	1	2	1	9,841	0,332
			2	9,585	0,377
			3	13,536	0,052
			4	10,496	0,239
56**	3	4	1	23,004	0,000
			2	10,040	0,014
			3	3,464	0,388
			4	2,603	0,597
59**	3	4	1	16,052	0,001
			2	8,405	0,030
			3	2,873	0,469
			4	2,741	0,501
63**	3	4	1	13,706	0,006
			2	6,017	0,261
			3	5,460	0,345
			4	5,221	0,389
64**	4	3	1	15,41	0,082
			2	14,18	0,152
			3	12,07	0,435
			4	12,62	0,331
65**	3	4	1	32,217	0,000
			2	14,231	0,001
			3	2,567	0,444
			4	2,121	0,555
70**	2	4	1	19,922	0,000
			2	4,178	0,195
			3	6,770	0,053
			4	1,482	0,751

Discriminant Analysis: NOVA REGIÃO L versus FOREST; BIRDS THREAT; ...

Quadratic Method for Response:NOVA REGIÃO L—Definida pelo MODELO LINEAR(46,80% de acerto no item 5.2 acima)

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT;
TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	18	22	22	15

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	17	0	0	0
2	1	22	0	0
3	0	0	20	0
4	0	0	2	15
Total N	18	22	22	15
N correct	17	22	20	15
Proportion	0,944	1,000	0,909	1,000

N = 77	N Correct = 74	Proportion Correct = 0,961
---------------	-----------------------	-----------------------------------

From Generalized Squared Distance to Group

Group	1	2	3	4
1	35,19	77,38	56,16	160,57
2	43,24	25,49	44,99	80,74
3	115,03	297,75	34,92	67,84
4	60,19	48,15	36,99	25,60

Summary of Misclassified Observations

Observation	True Group	Pred Group	Squared Distance	Probability	
8**	1	2	1	39,49	0,111
			2	35,32	0,889
			3	53,44	0,000
			4	158,86	0,000
36**	3	4	1	88,24	0,000
			2	57,76	0,000
			3	37,15	0,427
59**	3	4	4	36,56	0,573
			1	81,15	0,000
			2	354,90	0,000
			3	38,10	0,048
			4	32,14	0,952

Discriminant Analysis: NOVA REGIÃO Q versus FOREST; BIRDS THREAT; ...

Linear Method for Response: NOVA REGIÃO Q-Definida pelo MODELO QUADRÁTICO(63,60% de acerto no item 5.2 acima)

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT; TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	15	33	17	12

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	9	2	0	1
2	5	29	1	2
3	0	0	11	1
4	1	2	5	8
Total N	15	33	17	12
N correct	9	29	11	8
Proportion	0,600	0,879	0,647	0,667

N = 77	N Correct = 57	Proportion Correct = 0,740
---------------	-----------------------	-----------------------------------

Squared Distance Between Groups

	1	2	3	4
1	0,00000	4,85013	9,55933	4,63495
2	4,85013	0,00000	9,23495	1,26286
3	9,55933	9,23495	0,00000	3,77060
4	4,63495	1,26286	3,77060	0,00000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-5,9598	-1,5704	-5,5148	-2,0406
FOREST	0,1142	0,0663	0,0472	0,0612
BIRDS THREAT	0,1047	0,0296	-0,0106	0,0153
PLANT THREAT	0,0041	-0,0008	-0,0027	-0,0017
MAMMAL THREAT	-0,0413	-0,0163	0,1168	0,0277

TER&MARPROTAREAS	0,0631	0,0756	0,0108	0,0561
TOTNATRESRENT	0,0647	0,0002	0,2890	0,1152

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1**	3	4	1	6,357	0,114
			2	5,341	0,190
			3	5,159	0,208
			4	3,462	0,487
2**	4	1	1	3,252	0,401
			2	6,063	0,098
			3	5,490	0,131
			4	3,418	0,369
5**	3	4	1	5,504	0,082
			2	5,428	0,086
			3	2,505	0,369
			4	2,051	0,463
9**	3	2	1	3,785	0,133
			2	1,370	0,444
			3	7,274	0,023
			4	1,577	0,400
19**	1	2	1	3,687	0,171
			2	1,341	0,552
			3	11,057	0,004
			4	2,745	0,273
29**	1	2	1	4,318	0,160
			2	1,770	0,571
			3	11,436	0,005
			4	3,308	0,265
31**	2	4	1	10,676	0,010
			2	3,022	0,449
			3	7,625	0,045
			4	2,824	0,496
32**	2	1	1	4,778	0,584
			2	6,151	0,294
			3	16,876	0,001
			4	7,931	0,121
39**	2	1	1	12,84	0,874
			2	17,18	0,100
			3	30,63	0,000
			4	19,86	0,026
42**	3	4	1	6,8618	0,029
			2	2,6314	0,238
			3	3,8973	0,126
			4	0,7571	0,607
43**	3	4	1	8,582	0,022
			2	3,043	0,357
			3	6,005	0,081
			4	2,214	0,540
44**	1	2	1	2,859	0,277
			2	1,770	0,477
			3	10,902	0,005
			4	3,134	0,241
45**	1	2	1	3,245	0,313
			2	2,490	0,456
			3	12,089	0,004
			4	3,883	0,227
54**	4	2	1	5,741	0,065
			2	1,537	0,534
			3	9,272	0,011
			4	2,165	0,390
58**	1	2	1	6,403	0,102
			2	3,523	0,431
			3	8,399	0,038
			4	3,527	0,430
59**	4	3	1	7,977	0,021
			2	5,924	0,058

			3	1,549	0,514
			4	2,015	0,407
63**	3	4	1	5,828	0,118
			2	5,308	0,153
			3	4,114	0,277
			4	3,137	0,452
68**	2	4	1	6,806	0,028
			2	1,176	0,466
			3	6,536	0,032
			4	1,140	0,474
73**	1	4	1	12,68	0,164
			2	13,33	0,118
			3	12,32	0,195
			4	10,35	0,523
77**	4	2	1	4,926	0,101
			2	1,416	0,585
			3	11,210	0,004
			4	2,688	0,310

Discriminant Analysis: NOVA REGIÃO Q versus FOREST; BIRDS THREAT; ...

Quadratic Method for Response:NOVA REGIÃO Q-Definida pelo MODELO QUADRÁTICO(63,60% de acerto no item 5.2 acima)

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT; TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	15	33	17	12

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	14	0	0	0
2	1	33	0	0
3	0	0	17	0
4	0	0	0	12
Total N	15	33	17	12
N correct	14	33	17	12
Proportion	0,933	1,000	1,000	1,000

N = 77	N Correct = 76	Proportion Correct = 0,987
---------------	-----------------------	-----------------------------------

From Generalized Squared Distance to Group

Group	1	2	3	4
1	36,45	177,92	81,30	1531,65
2	40,21	25,38	38,10	30,72
3	93,74	537,18	34,62	86,62
4	46,25	114,42	36,22	23,40

Summary of Misclassified Observations

Observation	True Group	Pred Group	Group	Squared	Probability
				Distance	
45**	1	2	1	39,03	0,004
			2	28,08	0,996
			3	45,80	0,000
			4	254,53	0,000

Vemos que, usando a distribuição dos países por região sugerida pelo modelo

QUADRÁTICO com **63,6%** de acerto do item 5.2 mais acima, temos pelo método de Análise de Distância apresenta no modelo **QUADRÁTICO** logo acima, uma proporção de acerto de **98,7%** contra **74,0%** do **LINEAR**.

Segue abaixo a TABELA dos 77 países e respectivas REGIÕES - original e + NOVA sugerida pela **Análise Discriminante** obtida pelos seus respectivos 6 indicadores de **MEIO-AMBIENTE (Environment)** através do método **QUADRÁTICO** com proporção de acerto de **98,7%** :

+ NOVA REGIÃO QUADRÁTICA	ITEM	COUNTRY	NOVA REGIÃO QUADRÁTICA
3	1	Argentina	3
4	2	Bolivia	4
1	3	Brazil	1
2	4	Bulgaria	2
3	5	Chile	3
1	6	China	1
1	7	Colombia	1
2	8	Costa Rica	2
3	9	Cuba	3
2	10	Denmark	2
2	11	Dominican Republic	2
1	12	Ecuador	1
4	13	Egypt, Arab Rep.	4
2	14	El Salvador	2
2	15	Finland	2
2	16	France	2
2	17	Germany	2
2	18	Guatemala	2
1	19	Honduras	1
2	20	Hungary	2
2	21	Iceland	2
3	22	India	3
3	23	Indonesia	3
3	24	Iran, Islamic Rep.	3
3	25	Iraq	3
2	26	Ireland	2
2	27	Israel	2
2	28	Italy	2
1	29	Jamaica	1
1	30	Japan	1
2	31	Jordan	2
2	32	Korea, Rep.	2
2	33	Lebanon	2
2	34	Lithuania	2
3	35	Mexico	3
4	36	Morocco	4
4	37	Mozambique	4
2	38	Netherlands	2
2	39	New Zealand	2
2	40	Nicaragua	2

4	41	Nigeria	4
3	42	Norway	3
3	43	Pakistan	3
1	44	Panama	1
2	45	Paraguay	1
1	46	Peru	1
1	47	Philippines	1
2	48	Poland	2
2	49	Portugal	2
2	50	Puerto Rico	2
2	51	Romania	2
3	52	Russian Federation	3
3	53	Saudi Arabia	3
4	54	Senegal	4
2	55	Singapore	2
4	56	South Africa	4
2	57	Spain	2
1	58	Sri Lanka	1
4	59	Sudan	4
1	60	Suriname	1
2	61	Sweden	2
2	62	Switzerland	2
3	63	Thailand	3
3	64	Trinidad and Tobago	3
4	65	Tunisia	4
2	66	Turkey	2
4	67	Uganda	4
2	68	Ukraine	2
3	69	United Arab Emirates	3
2	70	United Kingdom	2
1	71	United States	1
2	72	Uruguay	2
1	73	Venezuela, RB	1
3	74	Vietnam	3
3	75	Yemen, Rep.	3
4	76	Zambia	4
4	77	Zimbabwe	4

Discriminant Analysis: + NOVA REGIÃO Q versus FOREST; BIRDS THREAT; ...

Linear Method for Response: + NOVA REGIÃO Q

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT;
TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	14	34	17	12

Summary of classification

	True Group			
Put into Group	1	2	3	4
1	9	2	0	0
2	3	30	1	2
3	0	0	11	1
4	2	2	5	9
Total N	14	34	17	12
N correct	9	30	11	9
Proportion	0,643	0,882	0,647	0,750

N = 77	N Correct = 59	Proportion Correct = 0,766
---------------	-----------------------	-----------------------------------

Squared Distance Between Groups

	1	2	3	4
1	0,00000	5,24609	9,57039	4,95874
2	5,24609	0,00000	9,33992	1,27884
3	9,57039	9,33992	0,00000	3,81188
4	4,95874	1,27884	3,81188	0,00000

Linear Discriminant Function for Groups

	1	2	3	4
Constant	-6,3253	-1,6005	-5,6085	-2,0644
FOREST	0,1125	0,0672	0,0468	0,0607
BIRDS THREAT	0,1016	0,0309	-0,0114	0,0144
PLANT THREAT	0,0048	-0,0007	-0,0024	-0,0016
MAMMAL THREAT	-0,0339	-0,0167	0,1194	0,0295
TER&MARPROTAREAS	0,0717	0,0753	0,0139	0,0581
TOTNATRESRENT	0,0727	0,0013	0,2923	0,1169

Summary of Misclassified Observations

	True	Pred	Squared		
Observation	Group	Group	Group	Distance	Probability
1**	3	4	1	6,668	0,099
			2	5,302	0,196
			3	5,175	0,209
			4	3,452	0,495
5**	3	4	1	5,710	0,075
			2	5,415	0,087
			3	2,503	0,372
			4	2,050	0,466
9**	3	2	1	4,157	0,112
			2	1,347	0,457
			3	7,314	0,023
			4	1,573	0,408
19**	1	2	1	4,113	0,141
			2	1,297	0,575
			3	11,104	0,004
			4	2,738	0,280
29**	1	2	1	4,728	0,133
			2	1,737	0,592
			3	11,484	0,005
			4	3,303	0,271
31**	2	4	1	11,226	0,008
			2	3,070	0,447
			3	7,745	0,043
			4	2,837	0,502
32**	2	1	1	5,453	0,460
			2	5,765	0,394

			3	16,821	0,002
			4	7,775	0,144
39**	2	1	1	13,26	0,820
			2	16,71	0,146
			3	30,44	0,000
			4	19,63	0,034
42**	3	4	1	7,2474	0,024
			2	2,6472	0,239
			3	3,9562	0,124
			4	0,7585	0,613
43**	3	4	1	8,941	0,019
			2	3,084	0,354
			3	6,063	0,080
			4	2,209	0,548
44**	1	2	1	3,178	0,243
			2	1,729	0,502
			3	10,919	0,005
			4	3,123	0,250
54**	4	2	1	6,042	0,057
			2	1,573	0,534
			3	9,312	0,011
			4	2,162	0,398
58**	1	4	1	6,388	0,106
			2	3,648	0,416
			3	8,387	0,039
			4	3,540	0,439
59**	4	3	1	8,306	0,018
			2	5,908	0,059
			3	1,579	0,511
			4	2,011	0,412
63**	3	4	1	5,838	0,119
			2	5,415	0,147
			3	4,114	0,282
			4	3,177	0,451
68**	2	4	1	7,395	0,021
			2	1,183	0,472
			3	6,656	0,031
			4	1,160	0,477
73**	1	4	1	12,40	0,193
			2	13,63	0,105
			3	12,34	0,199
			4	10,49	0,503
77**	4	2	1	5,238	0,088
			2	1,438	0,591
			3	11,250	0,004
			4	2,688	0,316

Discriminant Analysis: + NOVA REGIÃO Q versus FOREST; BIRDS THREAT; ...

Quadratic Method for Response: + NOVA REGIÃO Q

Predictors: FOREST; BIRDS THREAT; PLANT THREAT; MAMMAL THREAT;
TER&MARPROTAREAS; TOTNATRESRENT

Group	1	2	3	4
Count	14	34	17	12

Summary of classification

Put into Group	True Group			
	1	2	3	4
1	14	0	0	0
2	0	34	0	0
3	0	0	17	0
4	0	0	0	12
Total N	14	34	17	12
N correct	14	34	17	12
Proportion	1,000	1,000	1,000	1,000

N = 77	N Correct = 77	Proportion Correct = 1,000
--------	----------------	----------------------------

From Generalized Squared Distance to Group

Group	1	2	3	4
1	36,68	206,14	86,51	1841,23
2	40,67	25,28	38,02	31,36
3	89,93	552,55	34,62	86,62
4	46,30	116,95	36,22	23,40

Vemos que, usando a distribuição dos países por região sugerida pelo modelo

QUADRÁTICO com **98,7%** de acerto do item acima, temos pelo método de Análise de Distância apresenta no modelo **QUADRÁTICO** logo acima, uma proporção de acerto de **100,0%** contra **76,6%** do **LINEAR**.

ITEM	COUNTRY	REGIÃO	ITEM	COUNTRY	REGIÃO
1	Argentina	1	1	Denmark	2
2	Bolivia	1	2	Finland	2
3	Brazil	1	3	France	2
4	Chile	1	4	Germany	2
5	Colombia	1	5	Hungary	2
6	Costa Rica	1	6	Iceland	2
7	Cuba	1	7	Ireland	2
8	Dominican Republic	1	8	Israel	2
9	Ecuador	1	9	Italy	2
10	El Salvador	1	10	Japan	2
11	Guatemala	1	11	Korea, Rep.	2
12	Honduras	1	12	Netherlands	2
13	Jamaica	1	13	New Zealand	2
14	Mexico	1	14	Norway	2
15	Nicaragua	1	15	Poland	2
16	Panama	1	16	Portugal	2
17	Paraguay	1	17	Spain	2
18	Peru	1	18	Sweden	2
19	Puerto Rico	1	19	Switzerland	2
20	Suriname	1	20	Turkey	2
21	Trinidad and Tobago	1	21	United Kingdom	2
22	Uruguay	1	22	United States	2
23	Venezuela, RB	1			

.- 9 PAÍSES			. + 12 PAÍSES		
ITEM	COUNTRY	.+ NREGQ	ITEM	COUNTRY	.+ NREGQ
1	Brazil	1	1	Bulgaria	2
2	China	1	2	Costa Rica	2
3	Colombia	1	3	Denmark	2
4	Ecuador	1	4	Dominican Republic	2

5	Honduras	1	5	El Salvador	2
6	Jamaica	1	6	Finland	2
7	Japan	1	7	France	2
8	Panama	1	8	Germany	2
9	Peru	1	9	Guatemala	2
10	Philippines	1	10	Hungary	2
11	Sri Lanka	1	11	Iceland	2
12	Suriname	1	12	Ireland	2
13	United States	1	13	Israel	2
14	Venezuela, RB	1	14	Italy	2
			15	Jordan	2
			16	Korea, Rep.	2
			17	Lebanon	2
			18	Lithuania	2
			19	Netherlands	2
			20	New Zealand	2
			21	Nicaragua	2
			22	Paraguay	2
			23	Poland	2
			24	Portugal	2
			25	Puerto Rico	2
			26	Romania	2
			27	Singapore	2
			28	Spain	2
			29	Sweden	2
			30	Switzerland	2
			31	Turkey	2
			32	Ukraine	2
			33	United Kingdom	2
			34	Uruguay	2
Região 1 = América Latina					
Região 2 = OECD					
Região 3 = Leste Asiático					
Região 4 = África					

ITEM	COUNTRY	REGIÃO	ITEM	COUNTRY	REGIÃO
1	Bulgaria	3	1	Morocco	4
2	China	3	2	Mozambique	4
3	Egypt, Arab Rep.	3	3	Nigeria	4
4	India	3	4	Senegal	4
5	Indonesia	3	5	South Africa	4
6	Iran, Islamic Rep.	3	6	Sudan	4
7	Iraq	3	7	Tunisia	4
8	Jordan	3	8	Uganda	4
9	Lebanon	3	9	Zambia	4
10	Lithuania	3	10	Zimbabwe	4

11	Pakistan	3	
12	Philippines	3	
13	Romania	3	
14	Russian Federation	3	
15	Saudi Arabia	3	
16	Singapore	3	
17	Sri Lanka	3	
18	Thailand	3	
19	Ukraine	3	Região 1 = América Latina
20	United Arab Emirates	3	Região 2 = OECD
21	Vietnam	3	Região 3 = Leste Asiático
22	Yemen, Rep.	3	Região 4 = África

.- 5 PAÍSES			.+ 2 PAÍSES		
ITEM	COUNTRY	.+ NREGQ	ITEM	COUNTRY	.+ NREGQ
1	Argentina	3	1	Bolivia	4
2	Chile	3	2	Egypt, Arab Rep.	4
3	Cuba	3	3	Morocco	4
4	India	3	4	Mozambique	4
5	Indonesia	3	5	Nigeria	4
6	Iran, Islamic Rep.	3	6	Senegal	4
7	Iraq	3	7	South Africa	4
8	Mexico	3	8	Sudan	4
9	Norway	3	9	Tunisia	4
10	Pakistan	3	10	Uganda	4
11	Russian Federation	3	11	Zambia	4
12	Saudi Arabia	3	12	Zimbabwe	4
13	Thailand	3			
14	Trinidad and Tobago	3			
15	United Arab Emirates	3			
16	Vietnam	3			
17	Yemen, Rep.	3			