

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

**ARVORES DE DECISÃO E
INDICADORES DE GOVERANÇA MUNDIAL**

**Beatriz Marcos Telles
Maria Carolina Mirabella Belloque
Mestrado em Administração Empresas
Disciplina : Métodos Quantitativos
Prof. Dr. Arnoldo de Hoyos Guevara**

**SÃO PAULO
2010**

1. Introdução

O presente trabalho tem como objetivo efetuar uma análise da amostra comparando correlações de dados dimensionadores utilizando e comparando os métodos de análise discriminante, regressão logística e árvore de decisão. Os dados analisados serão uma amostra de 60 indivíduos retirada a partir dos Indicadores de Governança Mundiais (Worldwide Governance Indicators - WGI) e o software utilizado será o SPSS.

2. Dados

O projeto dos Indicadores de Governança Mundiais (Worldwide Governance Indicators - WGI) reportam indicadores de governança agregados e individuais de 212 países e territórios no período 1996–2008, para seis dimensões de governança:

- Voice and Accountability
(Voz e Contabilidade)
- Political Stability and Absence of Violence
(Estabilidade Política e Ausência de Violência)
- Government Effectiveness
(Efetividade Governamental)
- Regulatory Quality
(Qualidade Regulatória)
- Rule of Law
(Estado de Direito)
- Control of Corruption
(Controle da Corrupção)

Os indicadores agregados combinam as visões de um grande número de empresas, cidadãos e correspondentes especialistas em pesquisas de países industrializados e em desenvolvimento. As fontes de dados individuais fundamentam os indicadores agregados que são desenhados por uma grande variedade de institutos de pesquisas, *think tanks*, ong's e organizações internacionais.

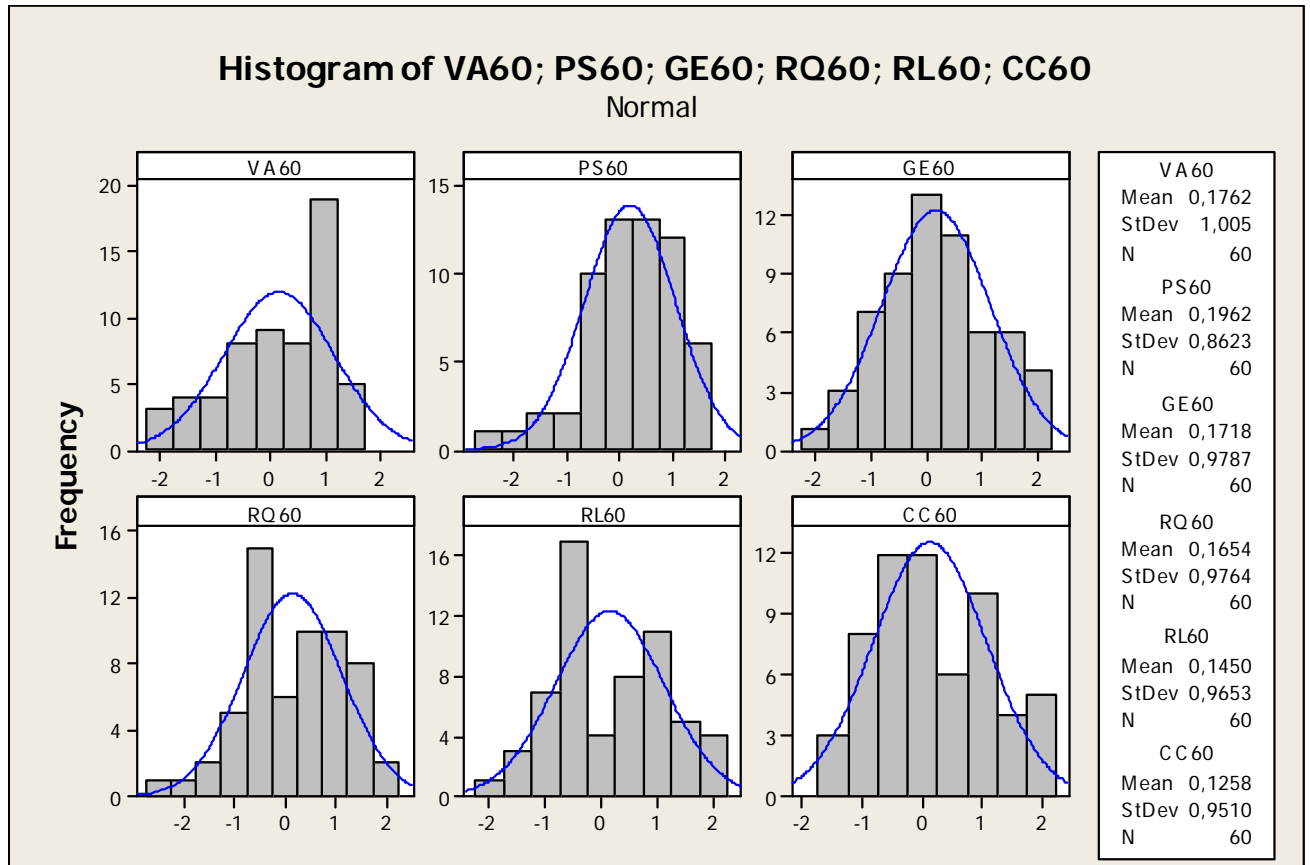
O que é Governança? Governança consiste das tradições e instituições pelas quais a autoridade num país é exercida, isto inclui o processo pelo qual cada governo é selecionado, monitorado e recolocado; a capacidade de um governo de efetivamente formular e implementar políticas sociais, o respeito pelos cidadãos e pelo estado para instituições que se relacionam econômica e socialmente com este.

3. Histograma da Amostra

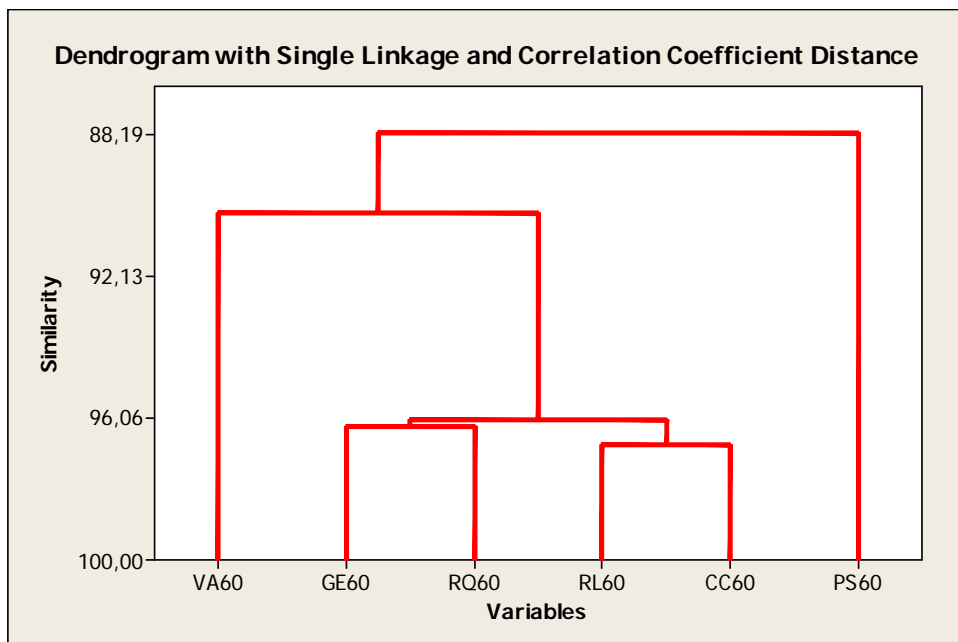
Descriptive Statistics: VA60; PS60; GE60; RQ60; RL60; CC60

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
VA60	60	0	0,176	0,130	1,005	-2,215	-0,393	0,511	1,001
PS60	60	0	0,196	0,111	0,862	-2,691	-0,270	0,290	0,853
GE60	60	0	0,172	0,126	0,979	-2,121	-0,660	0,190	0,758
RQ60	60	0	0,165	0,126	0,976	-2,277	-0,489	0,274	0,926
RL60	60	0	0,145	0,125	0,965	-1,870	-0,628	0,154	0,879
CC60	60	0	0,126	0,123	0,951	-1,738	-0,545	-0,0523	0,922

Na amostra de 60 indivíduos verificamos que os histogramas aparecem de forma muito parecidas aos da população o que significa que esta amostra está oferecendo uma boa representação da realidade. Assim como na população as curvas são todas normais.



3. Dendrograma



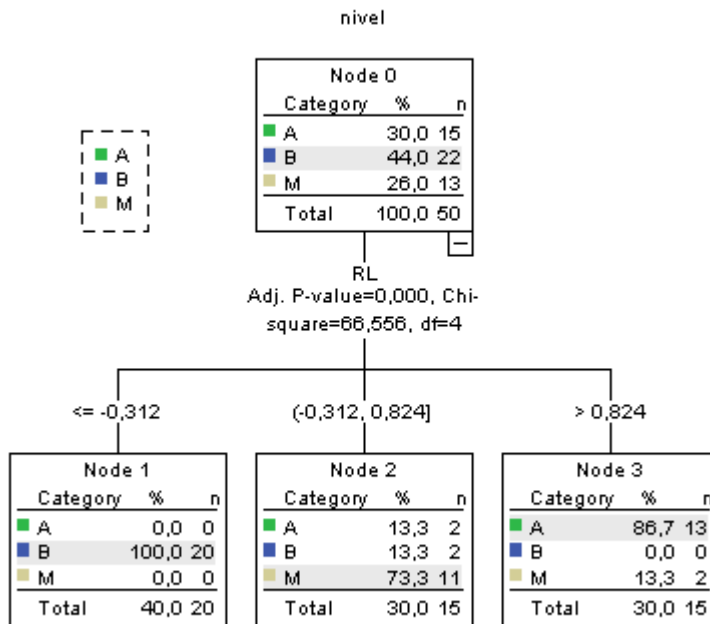
O Dendrograma mostra que as variáveis GE, RQ, RL e CC apresentam um alto grau de similaridade. Enquanto que a PS e a VA se apresentam de forma bem distinta das demais.

4. Regressão Logística Nominal

Classification				
Observed	Predicted			Percent Correct
	A	B	M	
A	15	0	0	100,0%
B	0	22	0	100,0%
M	0	0	13	100,0%
Overall Percentage	30,0%	44,0%	26,0%	100,0%

A Classificação utilizando o método de regressão logística apresentou um índice de acerto de 100% para as três variáveis.

5. Árvore de Decisão 1



Classification

Observed	Predicted			Percent Correct
	A	B	M	
A	13	0	2	86,7%
B	0	20	2	90,9%
M	2	0	11	84,6%
Overall Percentage	30,0%	40,0%	30,0%	88,0%

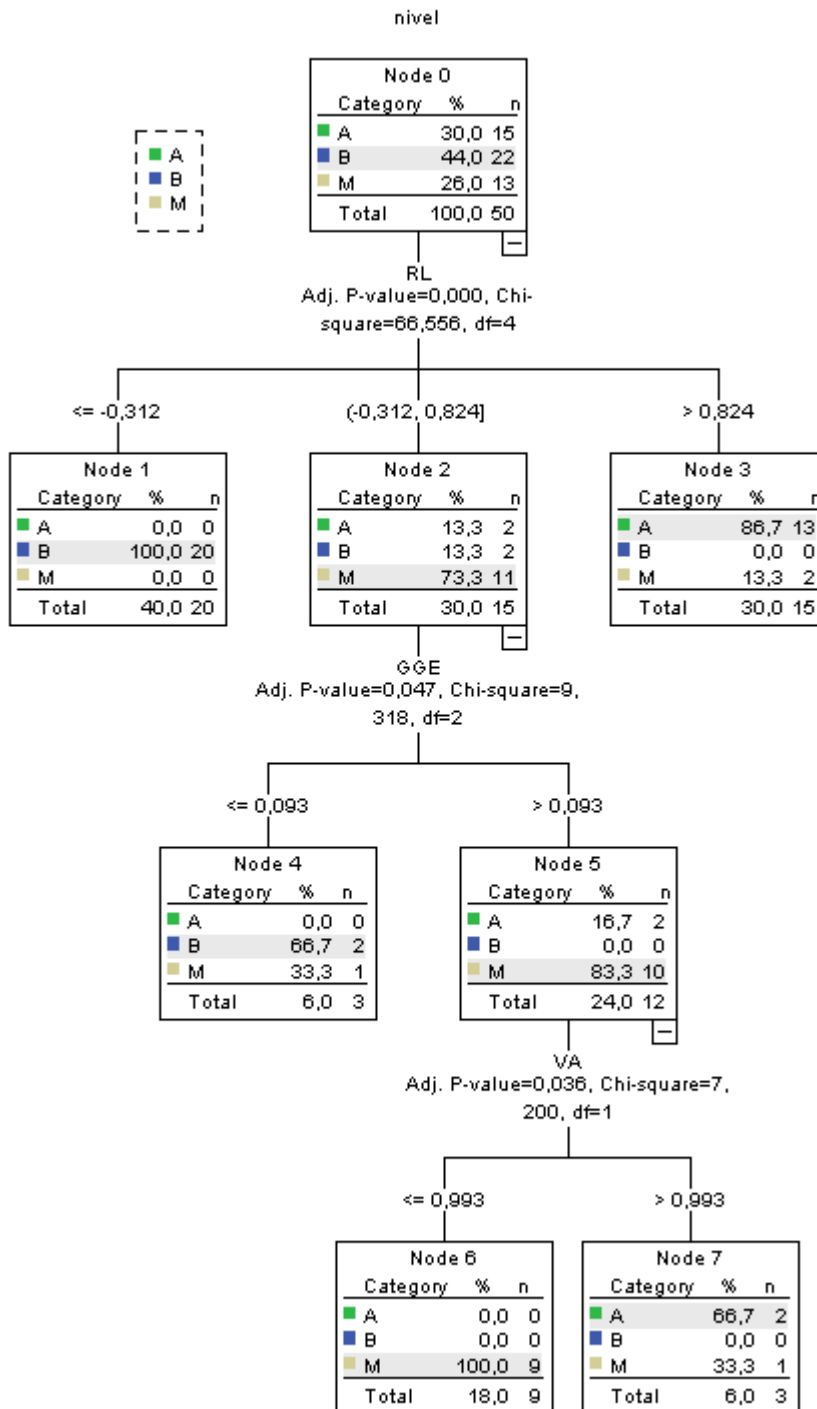
Growing Method: CHAID

Dependent Variable: nivel

Na primeira análise utilizando o método de árvore de decisão, o parent node utilizado foi de 30 e o child node foi de 5. Assim, a árvore apresentada utilizou a variável RL (Rule of Law) e se dividiu em três grupos. O primeiro grupo conseguiu concentrar 100% de B. O segundo já não foi tão bem, ficou dividido em 73,3; 13,3 e 13,3 por cento e no terceiro grupo foi bom, conseguiu concentrar 86,7% de A e apenas 13,3% de M.

Essa árvore apresentou um índice de 88% de acerto de acordo com a tabela de classificação.

6. Árvore de Decisão 2



Classification

Observed	Predicted			
	A	B	M	Percent Correct
A	15	0	0	100,0%
B	0	22	0	100,0%
M	3	1	9	69,2%
Overall Percentage	36,0%	46,0%	18,0%	92,0%

Growing Method: CHAID

Dependent Variable: nivel

No segundo teste utilizando o método da árvore de decisão, o parent node utilizado foi de 5 e o child node de 3. Assim a árvore ficou maior, mais detalhada. A primeira parte da árvore é idêntica a do teste 1, já descrita acima. Porém ela continua e melhora o resultado do segundo grupo, que não tinha sido bom. O segundo grupo é subdividido utilizando a variável GE (Government Effectiveness). A partir dessa subdivisão, foi possível agrupar o M em 83,3% para um valor acima de 0,093 desta variável. A análise continua e utiliza a variável VA (Voice and Accountability) para re-agrupar pela terceira vez o nível, finalmente chegando a 100% de M para um valor abaixo de 0,993 desta variável.

O índice de acerto desta análise foi de 92% de acordo com a tabela de classificação.

7. Conclusão Final

Apesar do método de árvore de decisão ter apresentado uma metodologia muito interessante e um índice de acerto significativo, de 92%. O método de regressão linear apresentou um índice de acerto de 100%, o que significa que é um teste mais preciso do que o outro.

Ambos apresentaram um alto índice de acerto mostrando que o trabalho de classificação foi realizado de uma forma ótima.