



**“ÍNDICE DE COMPETITIVIDADE URBANA – 2008 (ICUR):
ESTUDO DAS MELHORES CIDADES PARA SE FAZER
NEGÓCIOS NA AMÉRICA LATINA ”**

FABRICIO CÉSAR BASTOS e PAULA FERNANDA PRADO PEREIRA

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

1. INTRODUÇÃO

O Índice de Competitividade Urbana (ICUR) é elaborado pelo departamento de estudos *AméricaEconomia Intelligence*. Constitui-se em um ranking estabelecido por meio da análise de 50 variáveis em 42 cidades da América Latina e Miami, complementado por uma pesquisa feita com 1,2 mil executivos estrangeiros em toda a região.

Essas informações são organizadas de forma a se obter indicadores para cinco dimensões: poder de negócios, vantagem geoestratégica, serviços para os negócios, capital intelectual e serviços pessoais. O resultado dessa análise é o **Índice de Competitividade Urbana**.

Esse índice permite que seja realizada uma análise com as melhores cidades para se fazer negócios, o principal objetivo é sintetizar da melhor forma as variáveis que as empresas e os executivos pesquisados levam em consideração na hora de decidir qual cidade reúne as melhores características para se fazer negócios.

A tese central se baseia na oferta urbana de cada cidade para fazer e atrair negócios, bem como na estrutura proporcionada pelo país-sede.

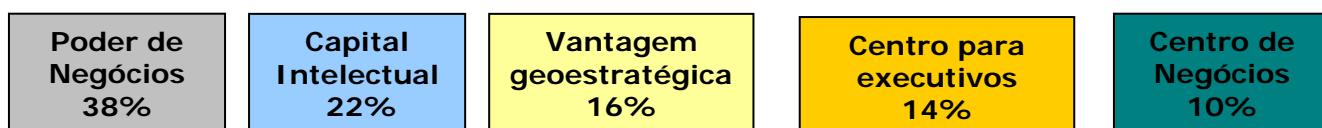
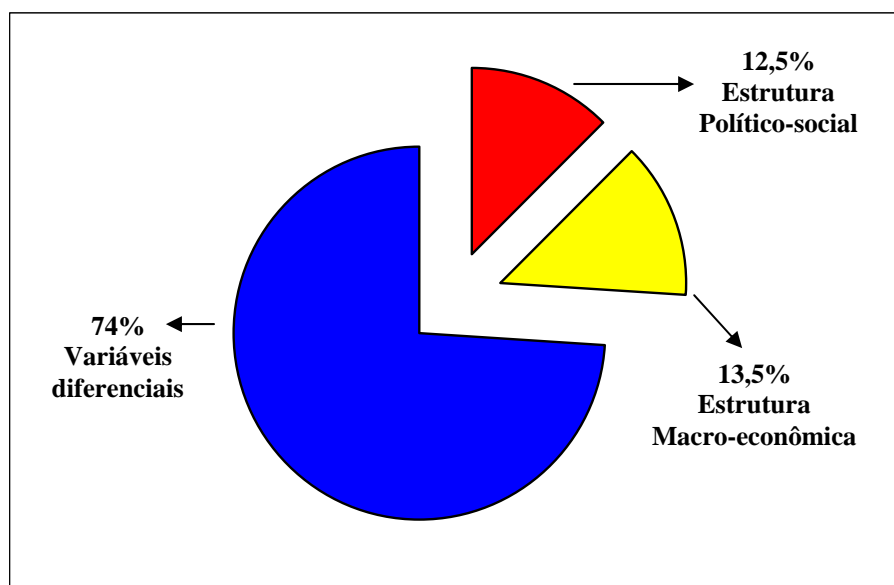
O **ICUR** encontra-se distribuído em duas grandes dimensões: variáveis de contexto e variáveis diferenciais.

As Variáveis de Contexto correspondem à plataforma que o país outorga às cidades consideradas, que define a escala dessas variáveis, agrupadas em dois indicadores: a estrutura macroeconômica e a político-social. Dentro destas, pode-se mencionar a liberdade para fazer negócios, o Índice de Desenvolvimento Humano (IDH), um indicador de governabilidade, entre outros.

A segunda dimensão, composta pelas Variáveis Diferenciais, se refere à oferta real que sustenta a competitividade própria das cidades e determina diferenças no índice final. As variáveis diferenciais são compostas por cinco sub-indicadores. O primeiro é o poder de negócios da cidade (PIB per capita, nível de crescimento, etc.). O segundo, os serviços oferecidos às empresas como centro de negócios (telecomunicações, custos de escritórios, etc.). O terceiro, a oferta urbana que a converte em centro para executivos (qualidade de vida, segurança e entretenimento). O quarto, o proveito que tira de sua localização, medido na vantagem geoestratégica (definida como movimento portuário e aeroportuário, pelo e-government e pelo poder de marca). E, finalmente, o capital intelectual, focado na criação e difusão de conhecimento (entre eles, universidades top 2000, escolas de MBA, patentes e *papers* ISI).

As ponderações, tanto das dimensões do índice como dos indicadores que o compõem, foram definidas por um painel de especialistas, tal como mostra o gráfico abaixo:

Ponderadores ICUR



As informações para a pesquisa procedem principalmente de organismos oficiais de cada cidade e país, tais como bancos centrais, bolsas de valores, institutos de pesquisa e portais web, assim como dados próprios da *AméricaEconomia Intelligence*. Também foram coletados dados em organismos internacionais como Banco Mundial, PNUD e ISI Web of Knowledge. Foram igualmente analisados e processados dados provenientes de companhias e órgãos internacionais como: ACI Aero, Cisco, CB Richard Ellies, The Heritage Foundation, Cybermetrics Lab, Economática e Mercer.

Como as variáveis são medidas em diferentes unidades, para efeito de comparação, elas são padronizadas em base 100, de forma a dar o maior valor à melhor prática, de acordo com a natureza da variável.

Algumas variáveis, como telecomunicações, por exemplo, foram sintetizadas em sub-índices, por meio da média das variáveis padronizadas. Finalmente, para cada indicador se designou um ponderador proposto pelo painel de especialistas, bem como para as duas dimensões do modelo, obtendo-se dessa forma o ICUR que permitiu determinar o ranking de cidades de 2008.

A seleção das cidades observa o alinhamento destas com os objetivos pesquisados pelas variáveis do ICUR (tamanho demográfico, importância da economia, etc.).

Em função da quantidade de variáveis, foram escolhidas para este estudo, 11 das 50. São elas: país, cidade, região, PIB US\$ milhões, PIB US\$ per capita, estabilidade política, liberdade de negociação, penetração de celulares e banda larga, periculosidade, poder de marca e ICUR.

Por meio destas variáveis, pode-se contemplar as principais dimensões do ICUR: estrutura macro-econômica e político social (liberdade de negociação e estabilidade política); competitividade das cidades (PIB em US\$ e per capita); o proveito que tira de sua localização – vantagem geoestratégica (poder de marca); e nas variáveis diferenciadas o foco foi específico nas telecomunicações (penetração de celulares e banda larga), promovendo oportunidade de informação e contato entre as pessoas.

Dado este cenário, o objetivo deste artigo é realizar análises estatísticas dos dados do ICUR das cidades incluídos no Estudo da América *EconomiaIntelligence* e verificar suas relações com os quatro indicadores escolhidos. Em resumo, as principais questões que buscamos explorar são:

- Como estão no ranking as cidades da América Latina em relação à Miami?
- Quais cidades melhoraram ou pioraram mais nos últimos anos ?
- Como relacionar a posição das cidades no ranking ICUR 2008, em relação ao índice de periculosidade?
- Como o índice (ICUR) pode ser aperfeiçoado considerando outras variáveis como qualidade de vida urbana?
- É possível criar um modelo que, a partir dos indicadores do ICUR 2008, possa explicar auxiliar as cidades a terem metas específicas para melhorar suas classificações?

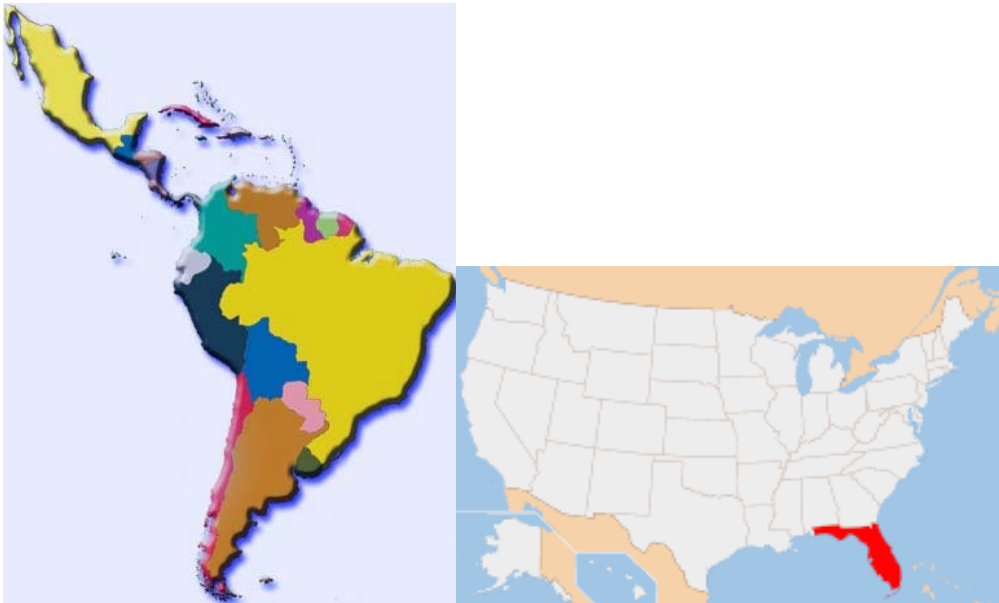
Importante ressaltar os limites deste estudo, pois em função do grande número de variáveis objetivou-se uma análise resumida destas descritas anteriormente.

Para que os objetivos deste trabalho fossem cumpridos, foram utilizadas ferramentas estatísticas disponíveis no pacote estatístico Minitab Statistical Software, versão 14.12.

2. ANÁLISE EXPLORATÓRIA DOS DADOS

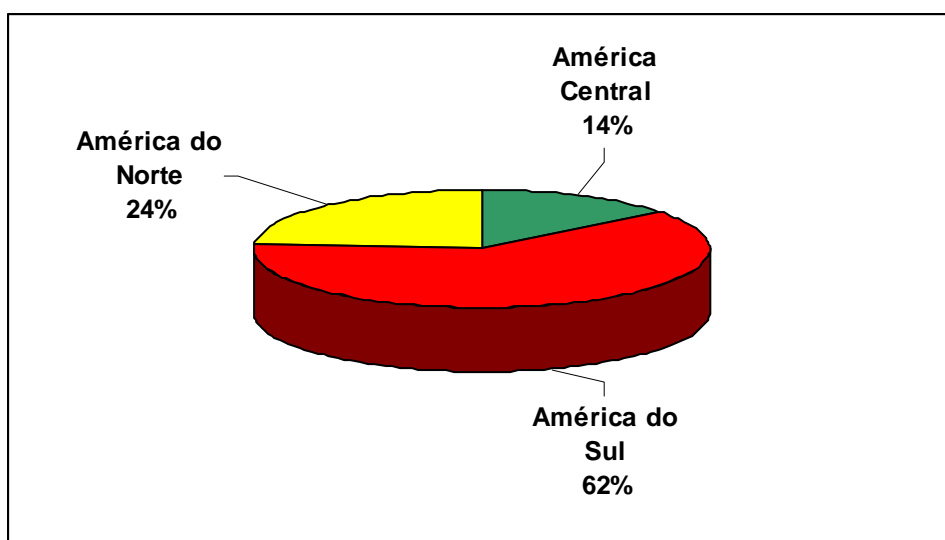
2.1 Os Indivíduos

Os indivíduos desta pesquisa são cidades da América Latina e Miami que compõem o Índice de Competitividade Urbana (ICUR) de 2008. As imagens a seguir indicam os países cujas cidades que compõem o índice (além de Miami):



O relatório é composto por 42 cidades (41 da América Latina, + Miami).

As cidades foram analisadas a partir de três regiões: América do Norte, América Central e América do Sul. O gráfico a seguir apresenta o percentual de cidades em cada região:



2.2 As Variáveis

São onze (11) as variáveis desta pesquisa, além do nome dos países. As variáveis estão melhor detalhadas na Tabela 1.

Tabela 1. Variáveis:

Variável	Significado	Tipo	Unidade de Medida
País	É o nome do país.	Variável Categórica	N/A
Cidade	Cidade a qual o país pertence.	Variável Categórica	N/A
Região	Região geográfica a qual o país pertence.	Variável Categórica	N/A
PIB US\$ Milhões	PIB em US\$ milhões.	Variável Quantitativa	US\$ (milhões)
PIB per capita US\$	PIB em US\$ per capita.	Variável Quantitativa	US\$ (milhões)
Estabilidade Política	Analisa a conjuntura do país referente as políticas internas e externas.	Variável Quantitativa	Varia entre 0 e 100.
Liberdade de Negociação	Grau de agilidade em aspectos negociais.	Variável Quantitativa	Varia entre 0 e 100.
Penetração (%) Celulares	% de Celulares nas cidades	Variável Quantitativa	Varia entre 0 e 100.
Penetração (%)_1 Banda Larga	% de Banda Larga nas cidades.	Variável Quantitativa	Varia entre 0 e 100.
Periculosidade	Analisa índices de violência	Variável Quantitativa	Varia entre 0 e 100.

	urbana.		
Poder de Marca	O quanto o nome da cidade tem impacto positivo nos negócios.	Variável Categórica	Varia entre muito baixo, baixo, médio, alto e muito alto.
ICUR	Índice de competitividade urbana. Baseia-se na oferta urbana de cada cidade para fazer e atrair negócios e, também, na estrutura proporcionada pelo país-sede.	Variável Quantitativa	Varia entre 0 e 100.

2.3 Tabelas de dados

Segue abaixo a tabelas de dados utilizada nesta pesquisa:

Tabela 2 – Os valores de 2008

Ranking 2008	País	Região	Cidade	PIB US\$ Milhões	PIB per capita US\$	Estab. Política	Liberdade/Negociação	Penetração (%) Celulares	Penetração (%) 1 Banda Larga	Periculosidade	Poder de Marca	ICUR
1	BR	América do Sul	SÃO PAULO	257.830	12.922	51,2	55,7	74,1	6,2	91,7	Alto	68,2
2	US	América do Norte	MIAMI	262.387	79.934	67,2	84,8	84	23,8	41,7	Muito Alto	66,8
3	CL	América do Sul	SANTIAGO	74.679	11.153	88,3	80,6	80,4	11,9	41,7	Muito Alto	59,4
4	MX	América do Norte	C. DO MÉXICO	219.555	10.882	51,7	63,3	91,3	4,6	91,7	Médio	53,8
5	AR	América do Sul	BUENOS AIRES	129.133	10.414	38,4	53,1	115	5,2	75	Muito Alto	49,4
6	MX	América do Norte	MONTERREY	48.884	13.813	51,7	63,3	73,5	3,1	58,3	Alto	40
7	BR	América do Sul	RIO DE JANEIRO	96.430	8.152	51,2	55,7	79,4	4,2	100	Médio	39,5
8	CO	América do Sul	BOGOTÁ	43.246	5.992	40,3	63,1	70	4,7	83,3	Alto	39,2
9	PE	América do Sul	LIMA	36.034	4.293	37	59,8	81,1	3,8	66,7	Baixo	38,8
10	BR	América do Sul	CURITIBA	17.608	9.677	51,2	55,7	77	4,5	66,7	Médio	38,3
11	PA	América Central	C. DO PANAMÁ	12.920	9.653	56	65,2	66,2	0,7	66,7	Alto	38
12	MX	América do Norte	QUERÉTARO	12.990	13.991	51,7	63,3	62,4	2	50	Médio	37,8
13	BR	América do Sul	PORTO ALEGRE	16.476	11.337	51,2	55,7	82,8	4,5	66,7	Médio	37,7
14	BR	América do Sul	BELO HORIZONTE	16.181	6.674	51,2	55,7	88,1	4,2	75	Baixo	37,7
15	MX	América do Norte	GUADALAJARA	33.547	8.316	51,7	63,3	63,4	2,5	66,7	Alto	37,1
16	BR	América do Sul	FLORIANÓPOLIS	3.735	4.484	51,2	55,7	80,6	4,5	41,7	Baixo	36,1
17	BR	América do Sul	BRASÍLIA	46.251	18.996	51,2	55,7	96,6	3,9	66,7	Baixo	36
18	MX	América do Norte	CHIHUAHUA	5.903	12.521	51,7	63,3	65,8	2,3	66,7	Baixo	35,9
19	BR	América do Sul	FORTALEZA	12.303	5.003	51,2	55,7	71,9	0,7	41,7	Muito Baixo	35,5
20	UY	América do Sul	MONTEVIDÉU	11.750	7.998	67,9	62,6	80	0,2	50	Médio	34,5
21	PR	América Central	SAN JUAN	40.661	23.877	67,2	84,8	90,5	3,9	41,7	Muito Baixo	34,4
22	CO	América do Sul	MEDELLIN	25.529	13.046	40,3	63,1	70	2,3	91,7	Médio	34,3
23	MX	América do Norte	PUEBLA	17.245	8.640	51,7	63,3	39	1,6	50	Muito Baixo	34,2
24	BR	América do Sul	RECIFE	10.165	6.647	51,2	55,7	62,4	0,7	66,7	Baixo	33,6
25	CR	América Central	SAN JOSÉ	8.177	5.271	68,7	64,7	32,3	2,7	41,7	Médio	33,5
26	BR	América do Sul	SALVADOR	13.428	4.868	51,2	55,7	105,6	0,7	66,7	Muito Baixo	33,2
27	MX	América do Norte	LEÓN	6.905	1.951	51,7	63,3	43,2	1,7	58,3	Muito Baixo	32,6
28	EC	América do Sul	GUAYAQUIL	11.755	5.530	20,8	50	73,1	0,1	75	Baixo	31,5
29	AR	América do Sul	ROSARIO	10.430	7.448	38,4	53,1	79	1,6	66,7	Muito Baixo	31,4
30	AR	América do Sul	MENDOZA	8.195	9.607	38,4	53,1	80	3,6	66,7	Muito Baixo	31,3
31	AR	América do Sul	CÓRDOBA	9.437	6.693	38,4	53,1	86	3	66,7	Baixo	31,2
32	MX	América do Norte	TJUANA	12.841	10.032	51,7	63,3	78,7	2,6	91,7	Muito Baixo	30,5
33	MX	América do Norte	CIUDAD JUAREZ	12.561	7.972	51,7	63,3	65,8	2,3	91,7	Muito Baixo	30,3
34	VE	América do Sul	CARACAS	36.308	9.760	12,3	37,6	80,5	7,9	91,7	Muito Baixo	29
35	DO	América Central	SANTO DOMINGO	12.931	6.541	43,9	56,7	55,7	0,1	75	Muito Baixo	28,3
36	CO	América do Sul	CALI	17.711	7.745	40,3	63,1	70	0,7	91,7	Muito Baixo	28,3
37	EC	América do Sul	QUITO	9.574	4.388	20,8	50	73,1	0,1	75	Muito Baixo	28,1
38	PY	América do Sul	ASSUNÇÃO	3.208	1.928	22,4	55,7	53,8	0,3	66,7	Muito Baixo	23,2
39	SV	América Central	SAN SALVADOR	9.105	3.961	50,2	73,7	55,1	0,1	91,7	Muito Baixo	22,3
40	GT	América Central	C. DA GUATEMALA	9.996	3.554	28,8	61,6	75	0,3	66,7	Muito Baixo	22,1
41	BO	América do Sul	SANTA CRUZ	3.935	2.437	22,2	47,4	35,7	0,6	75	Muito Baixo	21,1
42	BO	América do Sul	LA PAZ	3.067	1.752	22,2	47,4	30,3	0,5	91,7	Muito Baixo	21

2.4 Fonte de dados

A tabela abaixo detalha as fontes de dados da pesquisa:

Tabela 3 – Fonte de dados

Dado	Fonte
ACI Aero	http://www.airports.org/
Mercer	http://www.mercer.com/
Economática	http://www.economica.com
Cybermetrics Lab	http://www.webometrics.info/

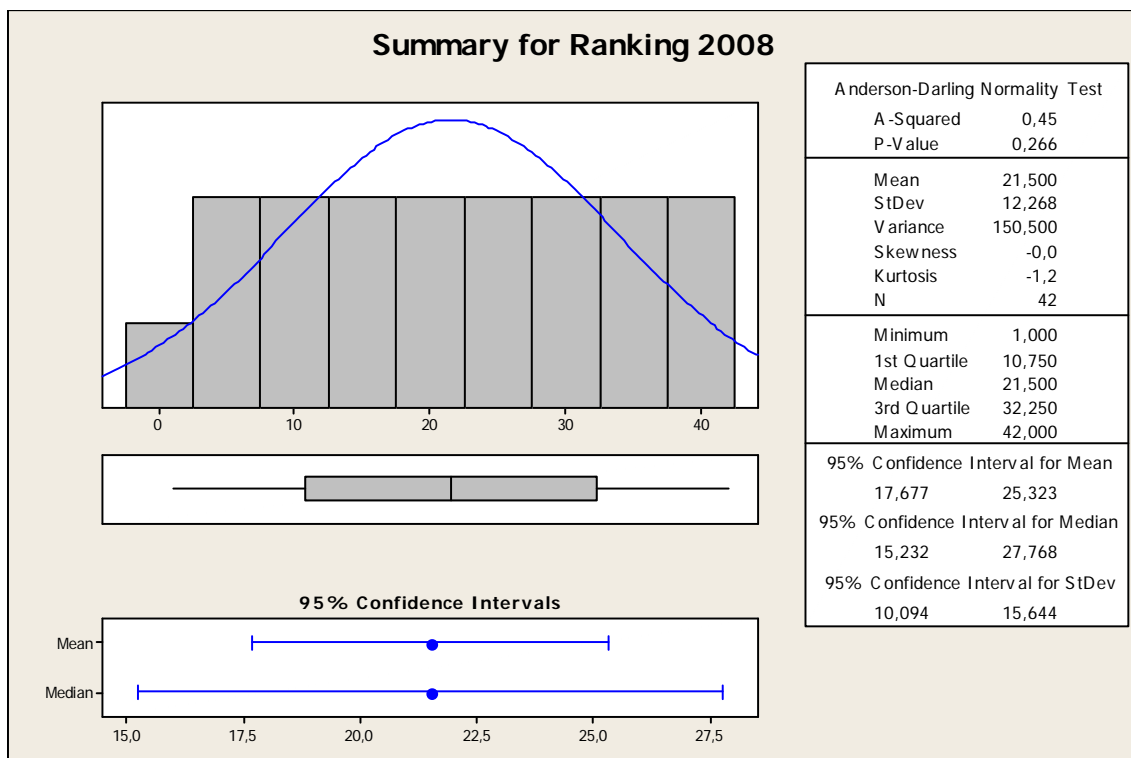
Cisco	http://www.cisco.com/web/BR/
CB Richard Ellis	http://www.cbrichardellis.com.br
The Heritage Foundation	http://www.heritage.org
Índice de Competitividade Urbana (ICUR)	http://www.americaintelligence.com
Bolsas de Valores	http://www.bovespa.com.br
PNUD	http://www.pnud.org.br/
ISI Web for Knowledge	http://www.isiknowledge.com/
Banco Mundial	http://www.worldbank.org/
<i>AméricaEconomia Intelligence</i>	http://www.americaintelligence.com
Banco Central do Brasil	http://www.bcb.gov.br/

3. Análise individual das variáveis quantitativas

A análise deste tipo de variável permite a utilização de uma grande gama de ferramentas como histogramas, curvas de densidade, *box-plot*, além de medidas numéricas como média, desvio-padrão, variância, quantidade de observações, valores mínimos e máximos, informações dos quartis e teste de normalidade. Assim, segue abaixo a análise individual de cada variável.

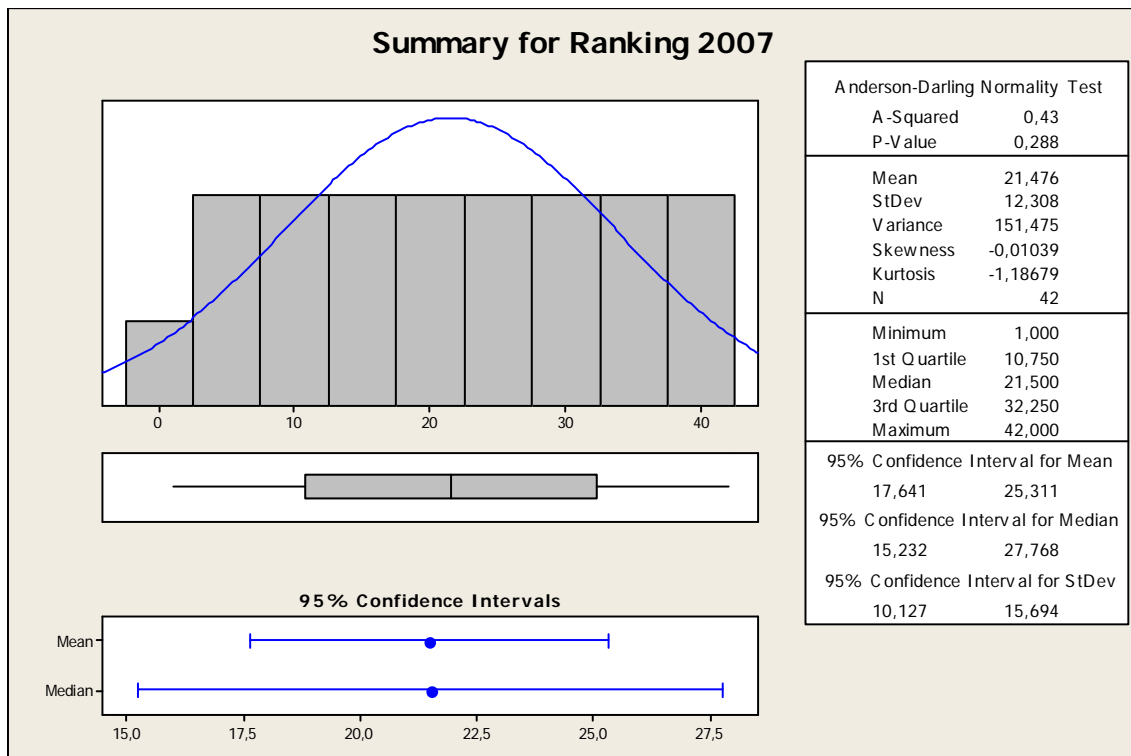
3.1 Graphical Summary (Teste de Normalidade Anderson Darling) para cada uma das variáveis.

3.1.1 Variável Ranking 2008



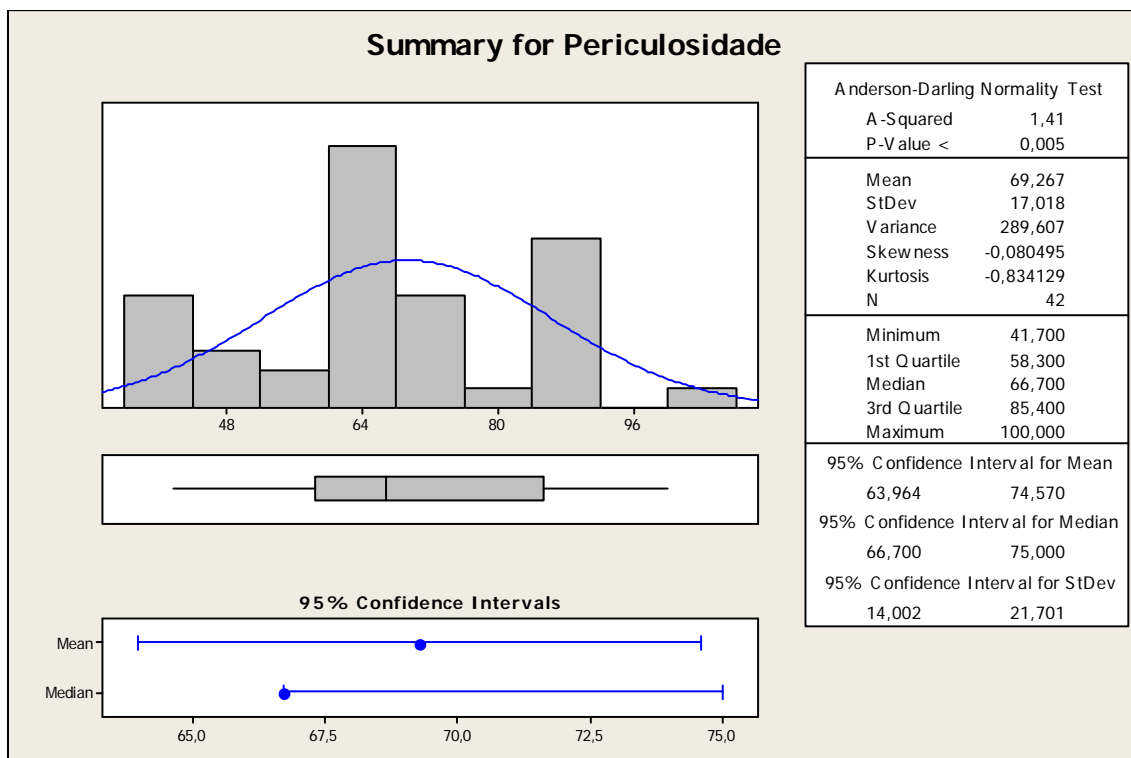
O teste de normalidade realizado para a variável Ranking 2008, representado a partir da tabela acima e do presente gráfico, aponta para uma curva normal, com média e mediana iguais, com o valor de 21,500 cada. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,266. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto. Enfim, esta variável pode ser considerada como uma boa variável.

3.1.2 Variável Ranking 2007

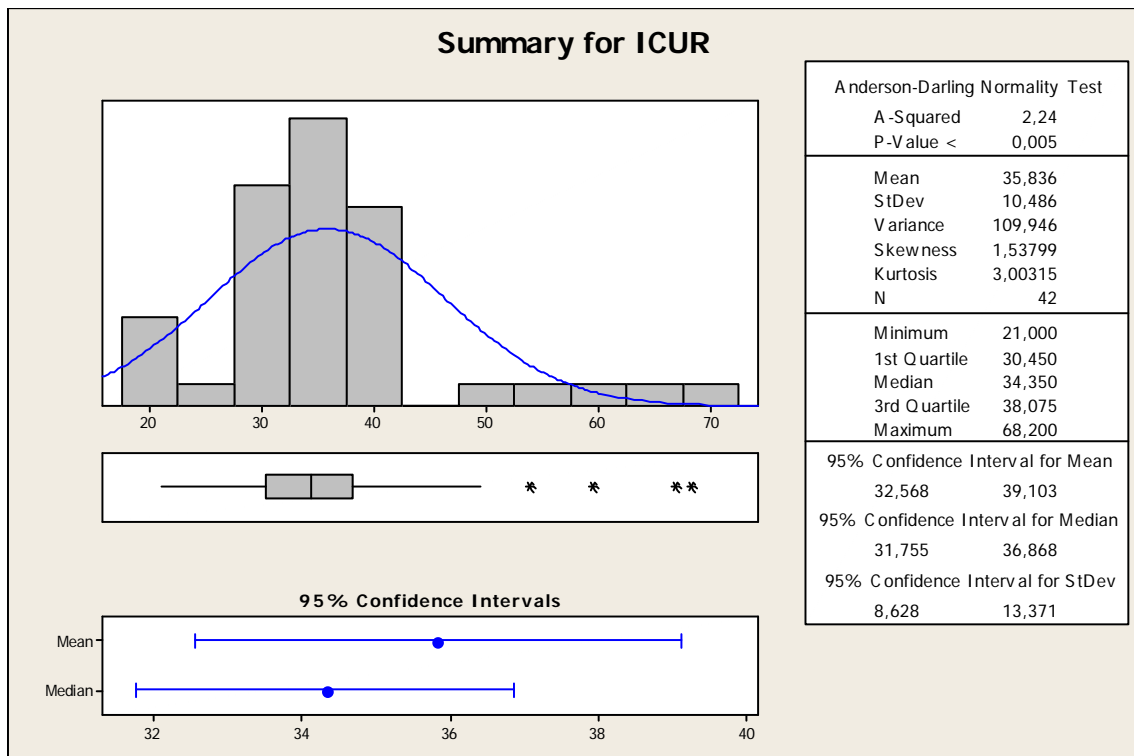


O teste de normalidade realizado para a variável Ranking 2007, representado a partir da tabela acima e do presente gráfico, aponta para uma curva normal, com média e mediana quase iguais, com o valor de aproximadamente 21,500 cada. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,288. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto. Enfim, esta variável pode ser considerada como uma boa variável.

3.1.3 Periculosidade

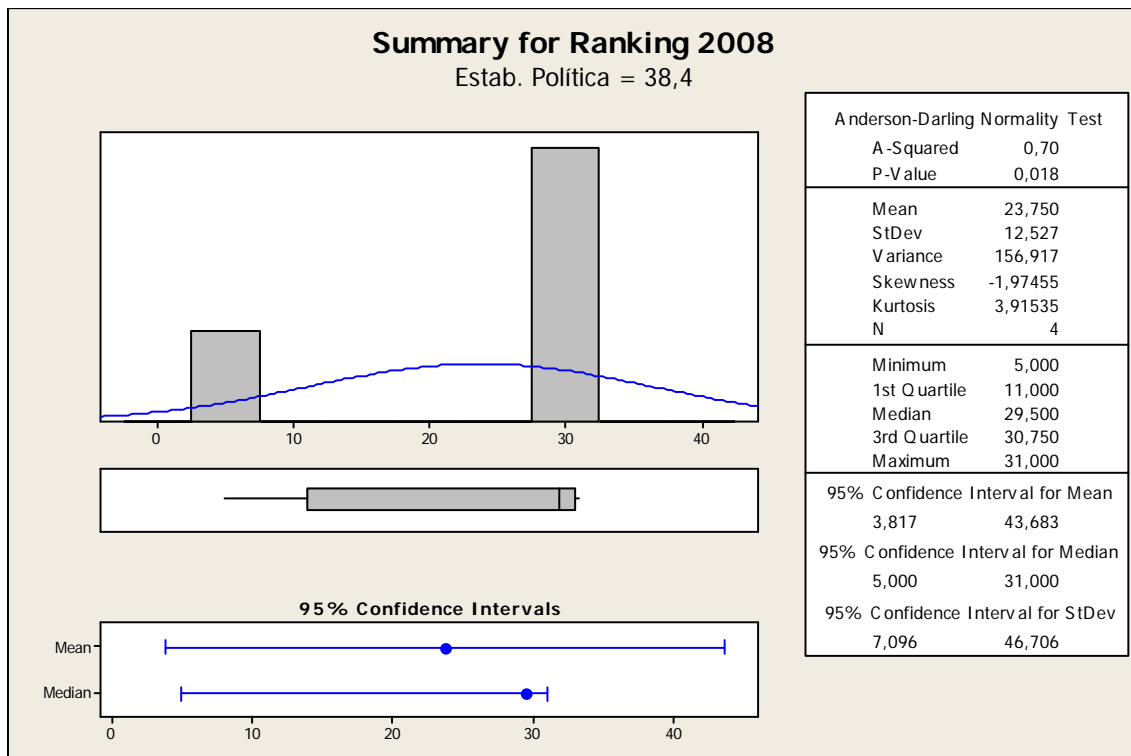


O teste de normalidade realizado para a variável Periculosidade, representado a partir da tabela acima e do presente gráfico, aponta para uma curva cujo formato lembra a normal, no entanto, a média e a mediana apresentam valores diferentes, de 69,267 e 66,700 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,005. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.



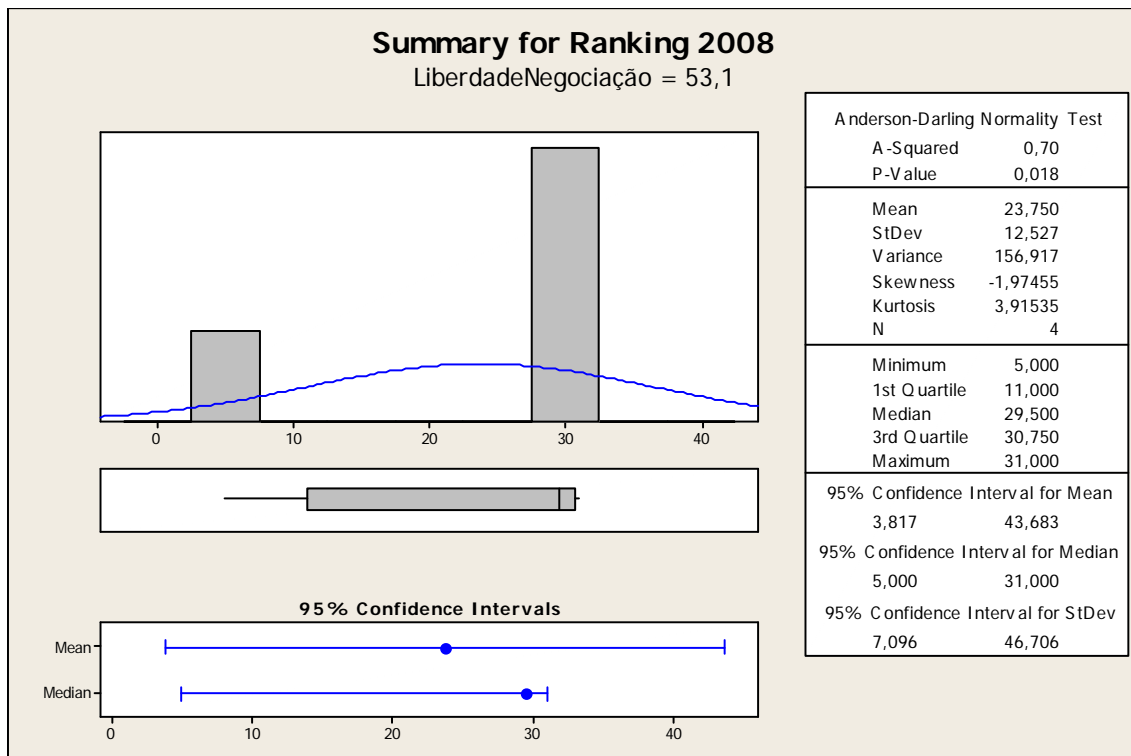
O teste de normalidade realizado para a variável ICUR, representado a partir da tabela acima e do presente gráfico, aponta para uma curva cuja média e mediana apresentam valores diferentes, de 35,836 e 34,350 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,005. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto. Alguns outliers podem ser observados, são eles: Cidade do México, Santiago, Miami e São Paulo.

3.1.5 Estabilidade Política



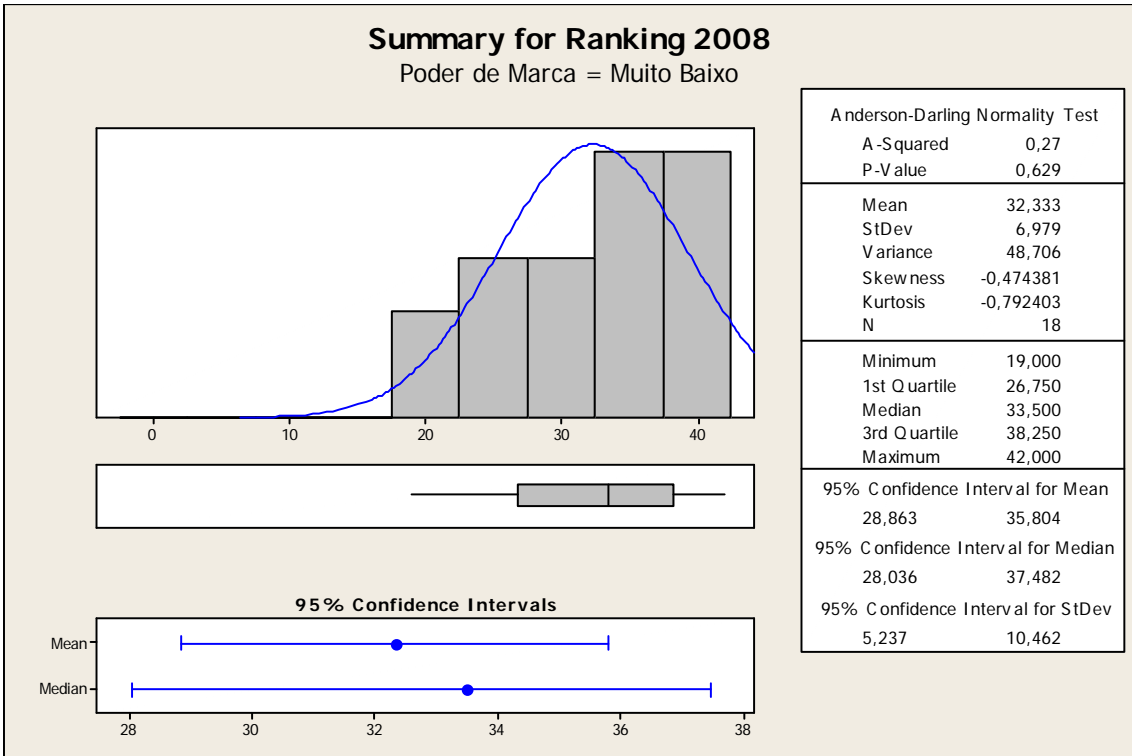
O teste de normalidade realizado para a variável Estabilidade Política, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. Aponta para uma curva baixa, próxima da base do gráfico. A média e a mediana apresentam valores diferentes, de 23,750 e 29,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,018. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto. Enfim, esta variável pode ser considerada como uma boa variável.

3.1.6 Liberdade de Negociação

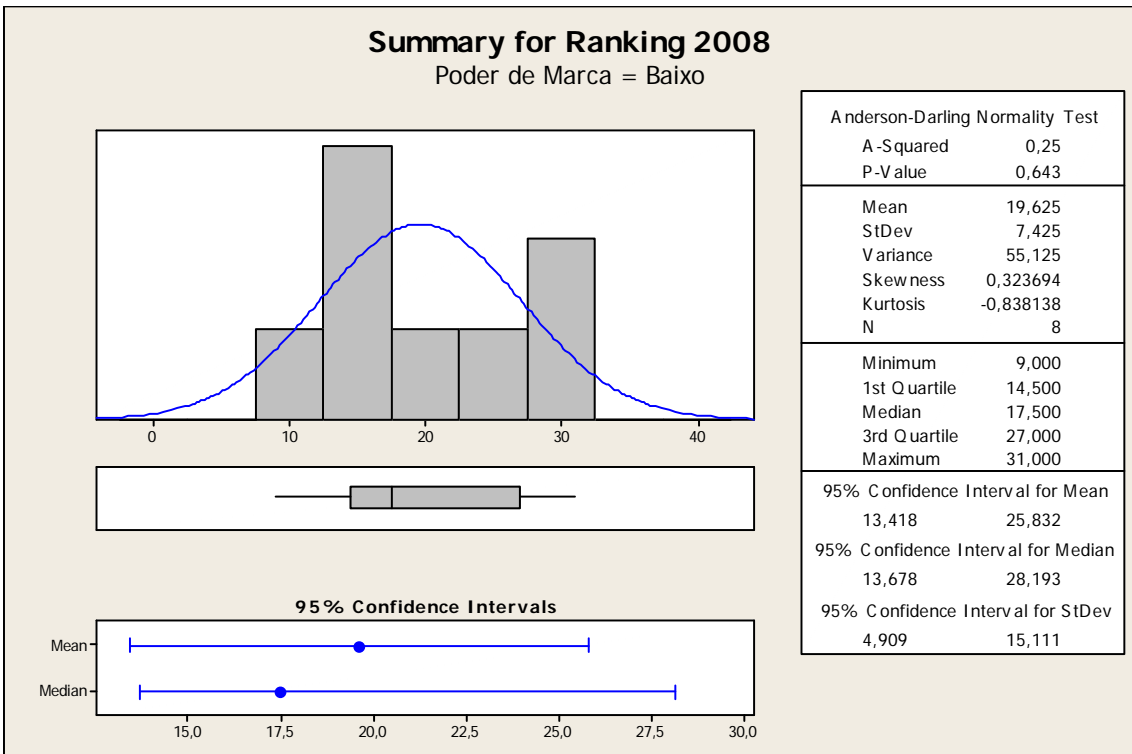


O teste de normalidade realizado para a variável Liberdade de Negociação, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. Aponta para uma curva baixa, próxima da base do gráfico. A média e a mediana apresentam valores diferentes, de 23,750 e 29,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,018. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.

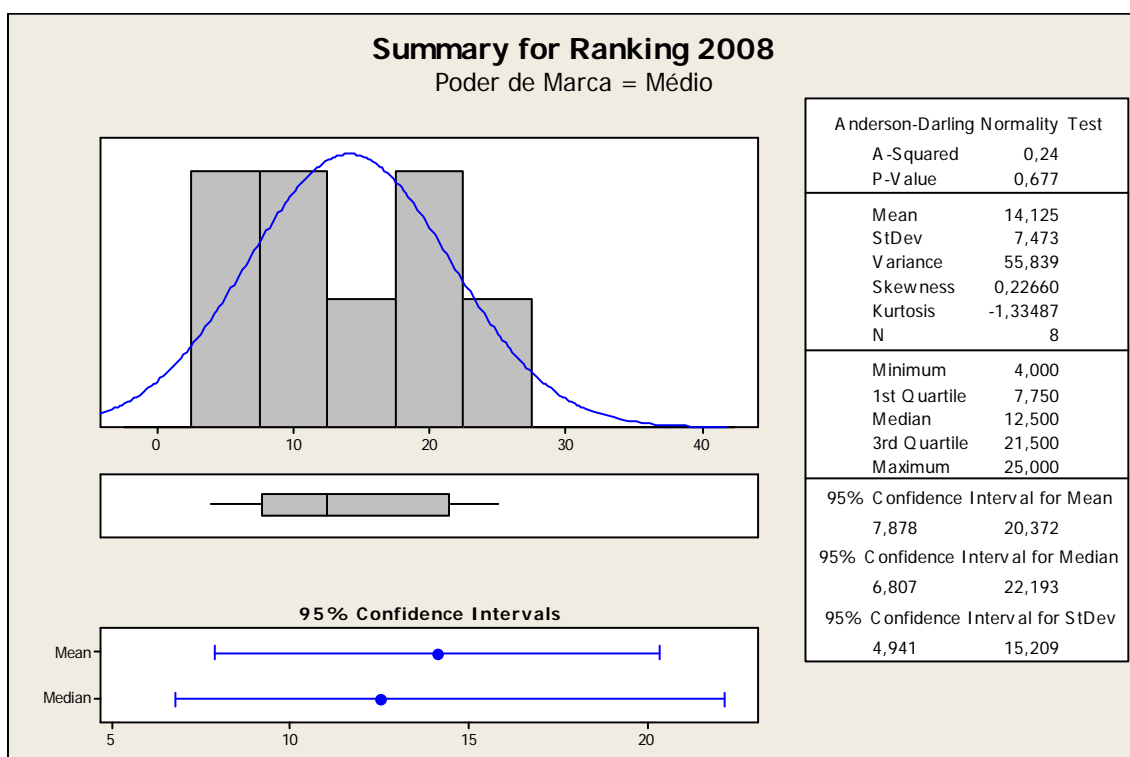
3.1.7 Poder de Marca



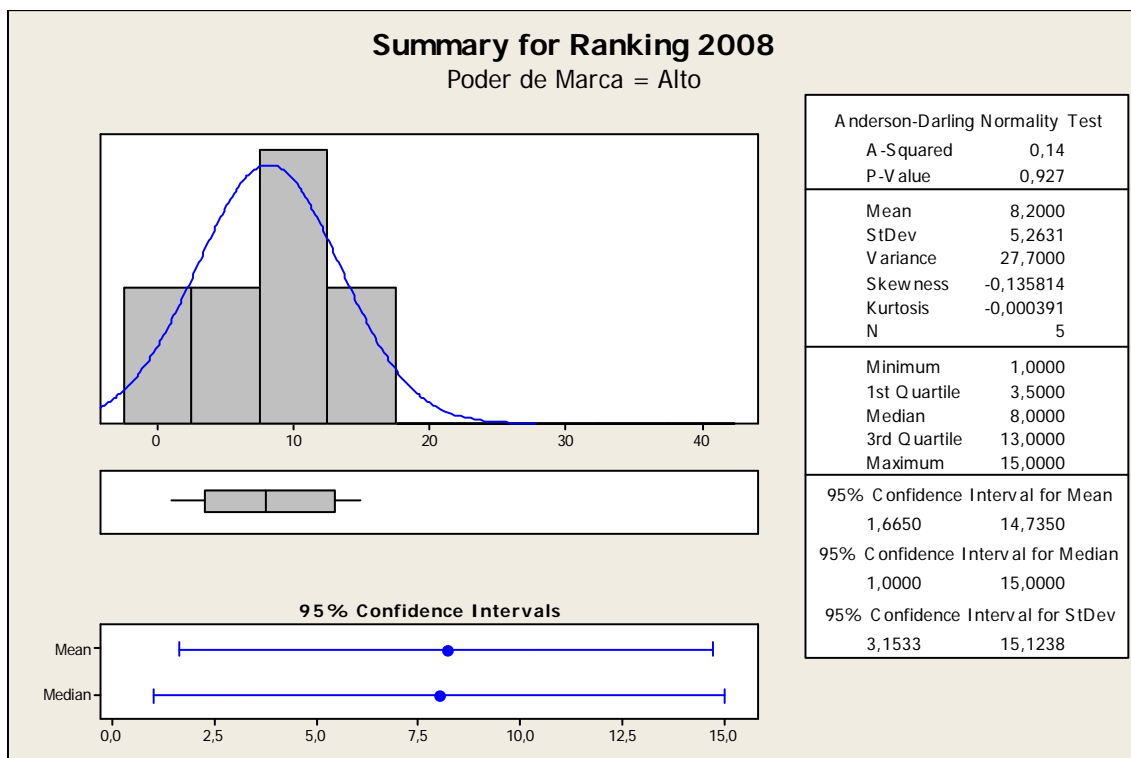
O teste de normalidade realizado para a variável Poder de Marca, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. Aponta para uma curva mais alta. A média e a mediana apresentam valores diferentes, de 32,333 e 33,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,629. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.



O teste de normalidade realizado para a variável Poder de Marca, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. O presente gráfico aponta para um poder de marca baixo. A média e a mediana apresentam valores diferentes, de 19,625 e 17,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,643. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.



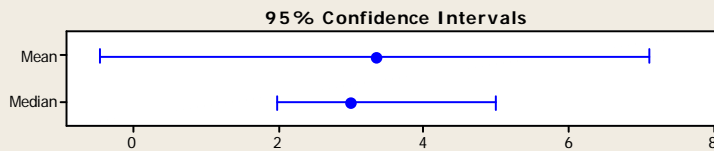
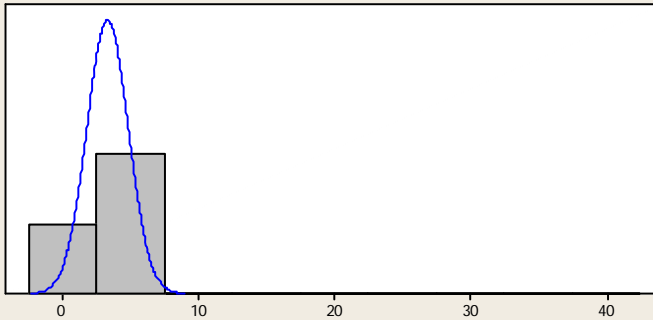
O teste de normalidade realizado para a variável Poder de Marca, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. O presente gráfico aponta para um poder de marca médio. A média e a mediana apresentam valores diferentes, de 14,125 e 12,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,677. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.



O teste de normalidade realizado para a variável Poder de Marca, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. O presente gráfico aponta para um poder de marca alto. A média e a mediana apresentam valores próximos, de 8,200 e 8,000 respectivamente. O p-value apresentado pode ser considerado alto, na medida em que tem por valor, 0,927. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.

Summary for Ranking 2008

Poder de Marca = Muito Alto



Anderson-Darling Normality Test

A-Squared	0,23
P-Value	0,487

Mean	3,3333
StDev	1,5275
Variance	2,3333
Skewness	0,935220
Kurtosis	*
N	3

Minimum	2,0000
1st Quartile	2,0000
Median	3,0000
3rd Quartile	5,0000
Maximum	5,0000

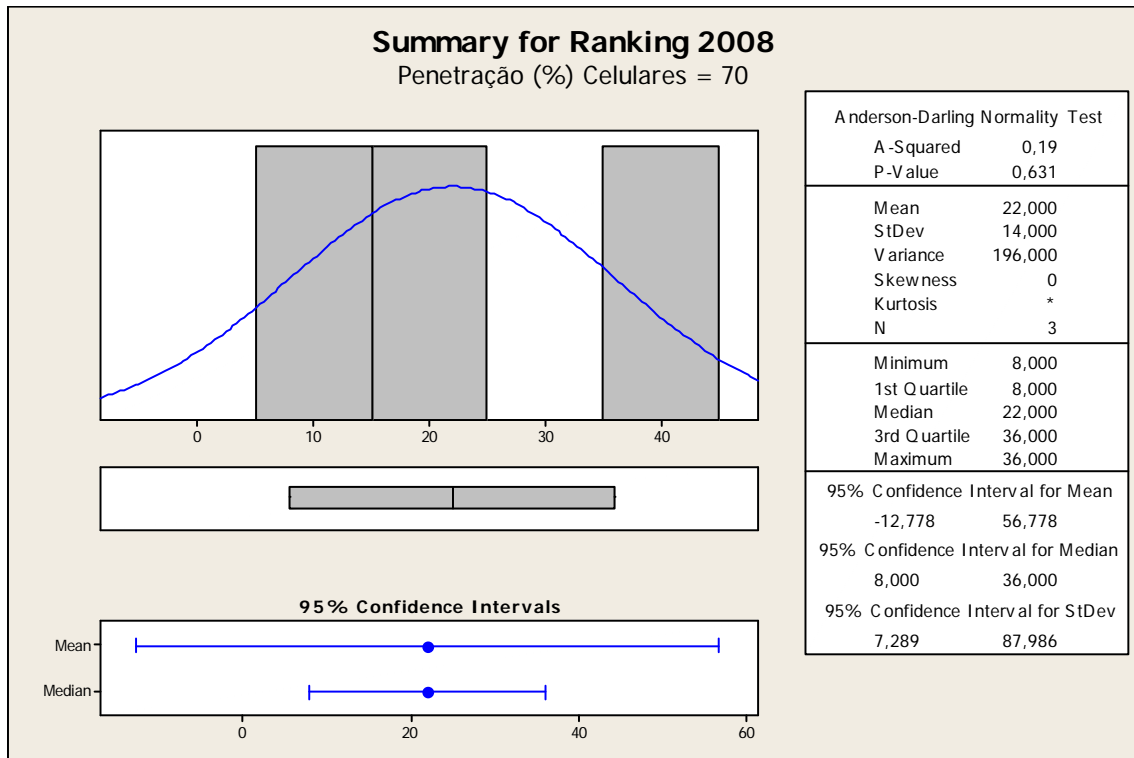
95% Confidence Interval for Mean	
-0,4612	7,1279

95% Confidence Interval for Median	
2,0000	5,0000

95% Confidence Interval for StDev	
0,7953	9,6001

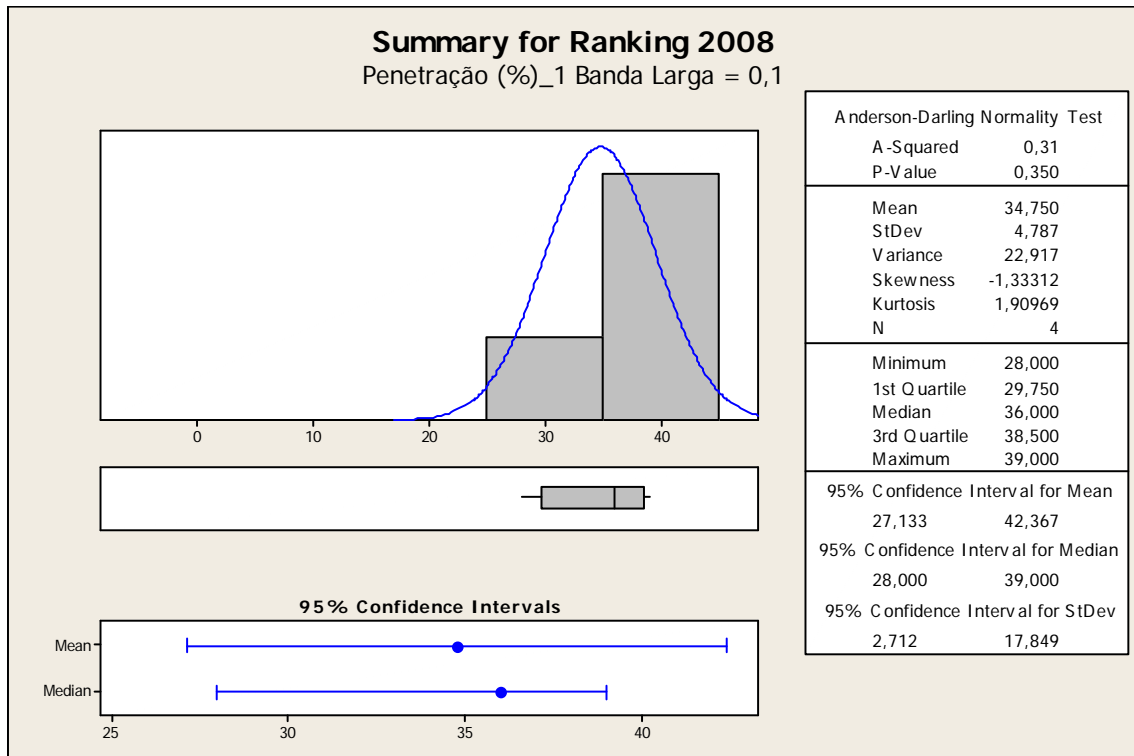
O teste de normalidade realizado para a variável Poder de Marca, representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. O presente gráfico aponta para um poder de marca muito alto. A média e a mediana apresentam valores próximos, de 3,333 e 3,000 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,487. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.

3.1.8 Penetração (%) Celulares



O teste de normalidade realizado para a variável Penetração Celulares (%), representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. Aponta para uma curva cuja média e mediana apresentam valores iguais, de 22,000. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,631. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.

3.1.9 Penetração (%) Banda Larga



O teste de normalidade realizado para a variável Penetração Banda Larga (%), representado a partir da tabela acima e do presente gráfico, teve de ser relacionado à variável Ranking 2008. Aponta para uma curva alta, cuja média e mediana apresentam valores diferentes, de 23,750 e 29,500 respectivamente. O p-value apresentado pode ser considerado baixo, na medida em que tem por valor, 0,350. O intervalo de confiança para a média, mediana e desvio-padrão encontra-se próximo de 95%, um valor que pode ser considerado alto.

4. Relação entre as variáveis

Esta seção do trabalho é dedicada à análise das relações entre as variáveis, incluindo as variáveis quantitativas, além da criação de modelos que possam explicar o ICUR das cidades.

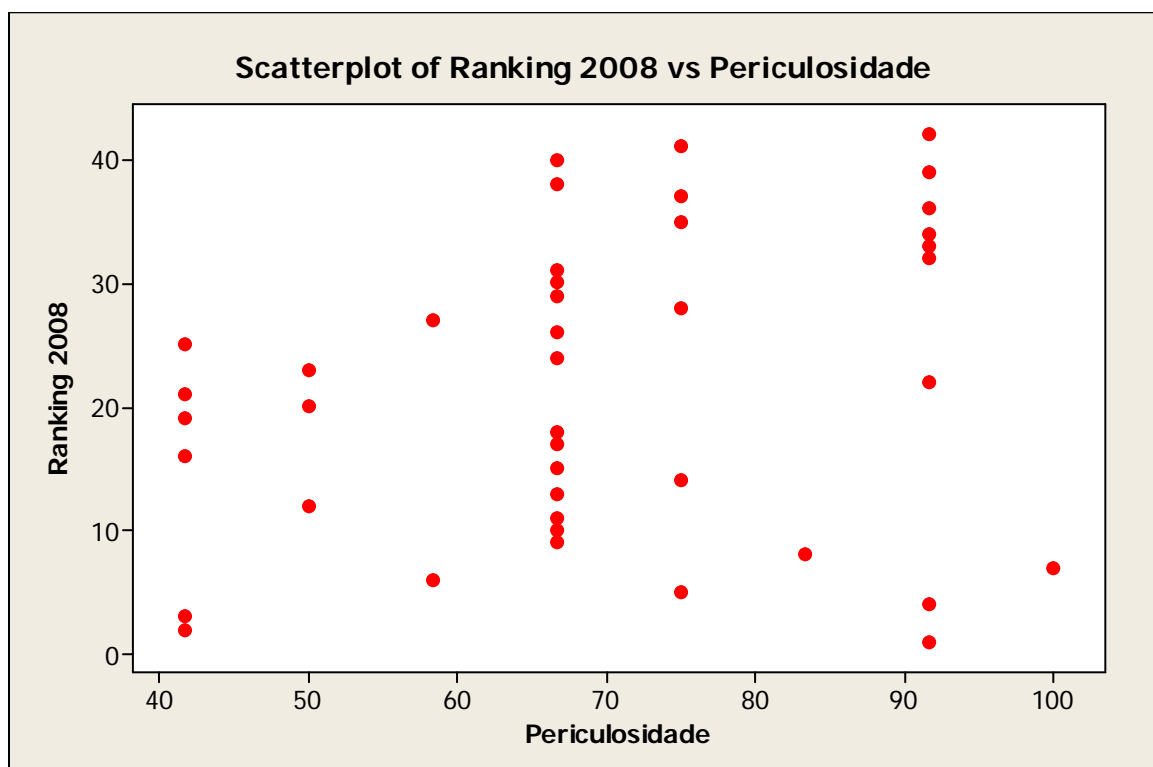
4.1 Correlations: Ranking 2008; Periculosidade; ICUR; Ranking 2007

	Ranking 2008	Periculosidade	ICUR
Periculosidade	0,246 0,116		
ICUR	-0,866 0,000	-0,186 0,239	
Ranking 2007	0,881 0,000	0,271 0,083	-0,777 0,000

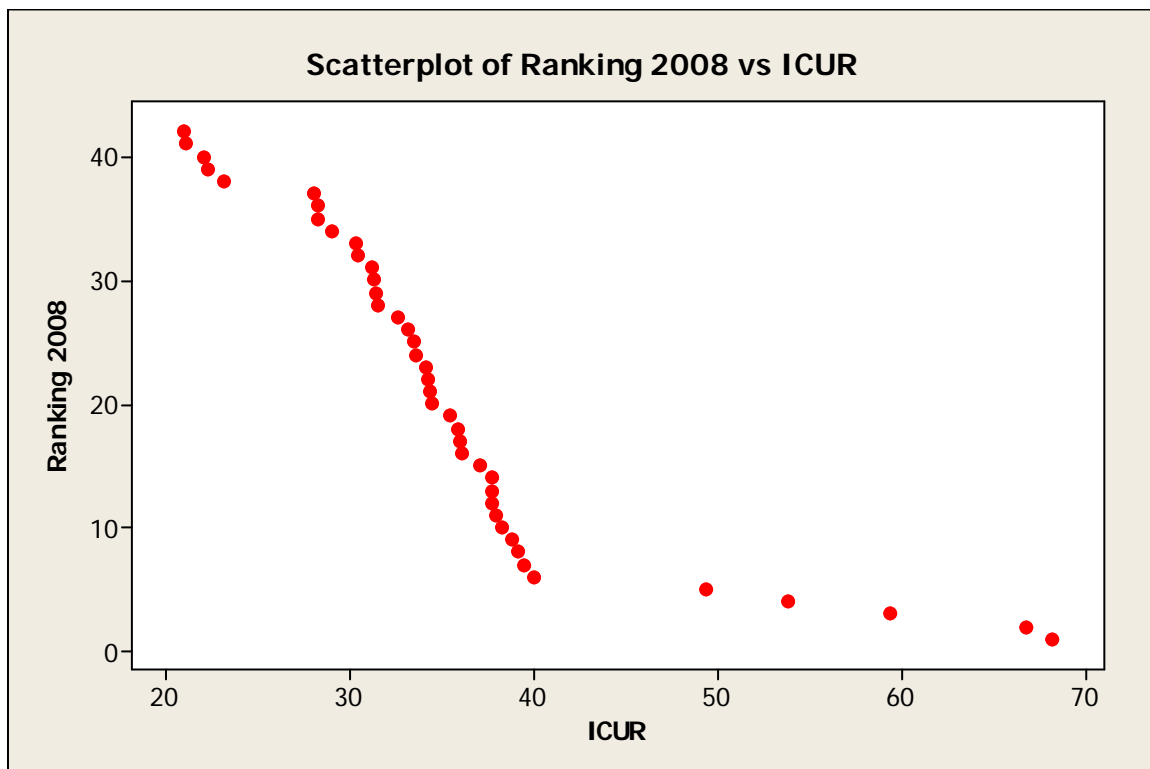
Cell Contents: Pearson correlation
P-Value

A maior correlação observada ocorre entre as variáveis Ranking 2008 e Ranking 2007, representada pelo valor: 0,881. A segunda maior correlação observada ocorre entre as variáveis Periculosidade e Ranking 2007, representada pelo valor 0,271. E por fim, a terceira maior correlação observada ocorre entre as variáveis Ranking 2008 e Periculosidade, representada pelo valor 0,246. É importante ressaltar que estas duas últimas correlações são baixas, quando comparadas a outros tipos de correlações possíveis, no entanto, para o presente exercício, foram destacadas por representar um valor mais alto quando comparada às demais.

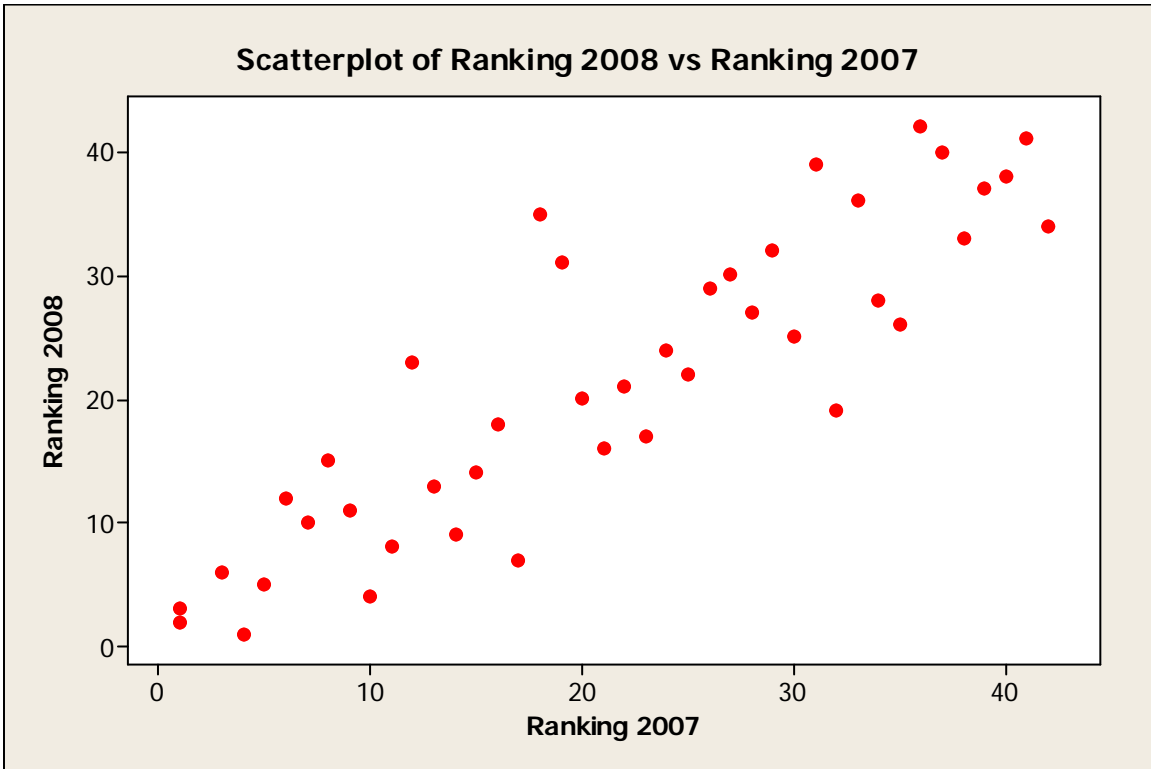
4.2 Gráficos Scatterplot



A partir das correlações observadas anteriormente, decidiu-se fazer análises utilizando o gráfico tipo “Scatterplot”. A primeira correlação se dá entre as variáveis Ranking 2008 e Periculosidade. Conforme pode-se observar, o presente gráfico apresenta uma leve concentração que aponta para a verticalidade.

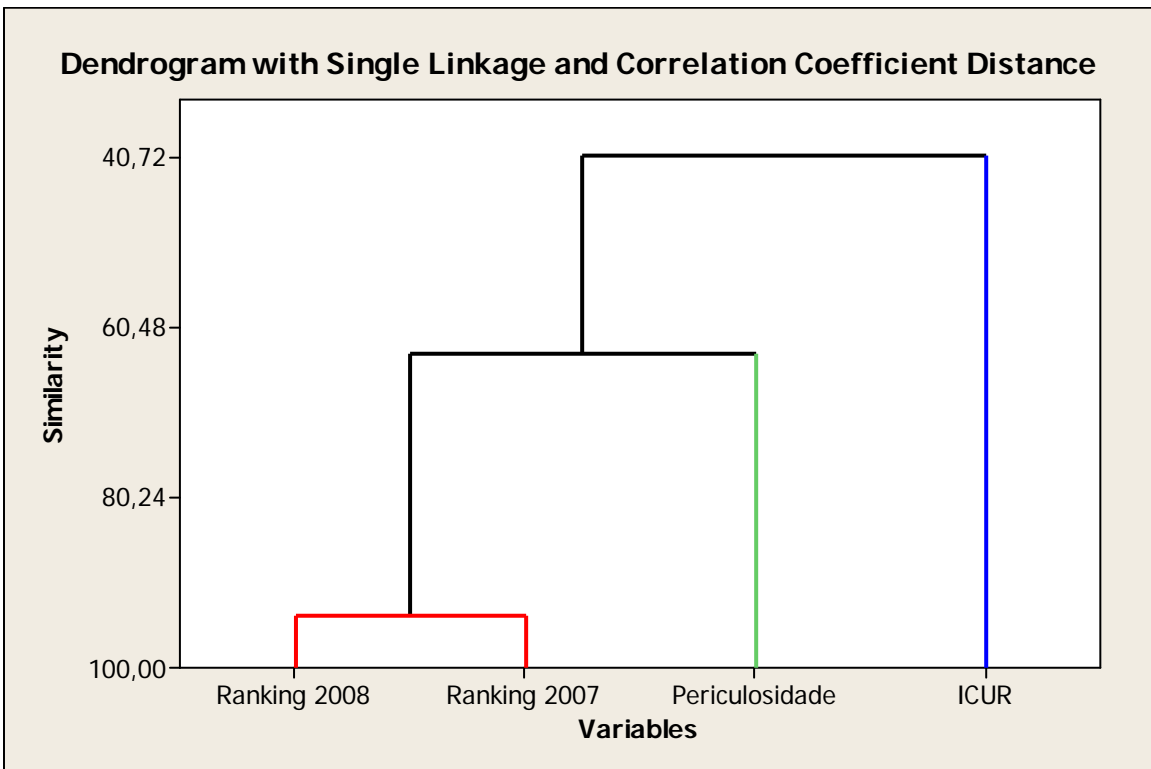


A segunda correlação se dá entre as variáveis Ranking 2008 e ICUR. Conforme pode-se observar, o presente gráfico apresenta uma certa concentração descendente entre as cidades analisadas com uma visível tendência à esquerda, sinalizando uma alta correlação entre as variáveis.



A terceira correlação se dá entre as variáveis Ranking 2008 e Ranking 2007. Conforme pode-se observar, o presente gráfico apresenta uma nuvem de tendência que aponta para cima, demonstrando uma dispersão ascendente, com $r > 0$.

4.3 Dendrograma



O dendrograma acima encontra-se dividido em três clusters, que aglomera as variáveis “Ranking 2008” e “Ranking 2007” e singulariza as variáveis Periculosidade e ICUR. Conforme observado, dentre as quatro variáveis, apenas duas apresentam maior similaridade.

4.4 Análise de Variância

Regression Analysis: Ranking 2008 versus Periculosidade; ICUR

The regression equation is

$$\text{Ranking 2008} = 52,7 + 0,0637 \text{ Periculosidade} - 0,994 \text{ ICUR}$$

Predictor	Coef	SE Coef	T	P
Constant	52,698	5,776	9,12	0,000
Periculosidade	0,06369	0,05791	1,10	0,278
ICUR	-0,99369	0,09398	-10,57	0,000

S = 6,20037 R-Sq = 75,7% R-Sq(adj) = 74,5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	4671,2	2335,6	60,75	0,000
Residual Error	39	1499,3	38,4		
Total	41	6170,5			

Source	DF	Seq SS
Periculosidade	1	373,4
ICUR	1	4297,8

Unusual Observations

Obs	Periculosidade	Ranking				
		2008	Fit	SE Fit	Residual	St Resid
1	92	1,000	-9,231	3,650	10,231	2,04RX
2	42	2,000	-11,025	3,195	13,025	2,45RX
7	100	7,000	19,816	2,104	-12,816	-2,20R

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large influence.

4.5 Stepwise Regression: Ranking 2008 versus Periculosidade; ICUR

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

Response is Ranking 2008 on 2 predictors, with N = 42

Step	1
Constant	57,80
ICUR	-1,013
T-Value	-10,94
P-Value	0,000

S	6,22
R-Sq	74,95
R-Sq(adj)	74,32
Mallows C-p	2,2

CO	0	0	0	0	0	0	0	0	0
CR	0	0	0	0	0	0	0	0	0
DO	0	0	0	0	0	0	0	0	0
EC	0	0	0	0	0	0	0	1	0
GT	0	0	0	0	0	0	0	0	0
MX	0	1	0	0	1	1	0	0	0
PA	0	0	0	0	0	0	0	0	0
PE	0	0	0	0	0	0	0	0	0
PR	0	0	0	0	0	0	0	0	0
PY	0	0	0	0	0	0	0	0	0
SV	0	0	0	0	0	0	0	0	0
US	1	0	0	0	0	0	0	0	0
UY	0	0	1	0	0	0	0	0	0
VE	0	0	0	0	0	0	0	0	0
All	1	1	1	1	1	1	1	1	1

RIO DE SANTO JANEIRO SAN JOSÉ SAN JUAN SAN SALVADOR SANTA CRUZ SANTA SANTIAGO

AR	0	1	0	0	0	0	0	0
BO	0	0	0	0	0	0	1	0
BR	1	0	1	0	0	0	0	0
CL	0	0	0	0	0	0	0	1
CO	0	0	0	0	0	0	0	0
CR	0	0	0	1	0	0	0	0
DO	0	0	0	0	0	0	0	0
EC	0	0	0	0	0	0	0	0
GT	0	0	0	0	0	0	0	0
MX	0	0	0	0	0	0	0	0
PA	0	0	0	0	0	0	0	0
PE	0	0	0	0	0	0	0	0
PR	0	0	0	0	1	0	0	0
PY	0	0	0	0	0	0	0	0
SV	0	0	0	0	0	1	0	0
US	0	0	0	0	0	0	0	0
UY	0	0	0	0	0	0	0	0
VE	0	0	0	0	0	0	0	0
All	1	1	1	1	1	1	1	1

SÃO PAULO TIJUANA All

AR	0	0	4
BO	0	0	2
BR	1	0	10
CL	0	0	1
CO	0	0	3

CR	0	0	1
DO	0	0	1
EC	0	0	2
GT	0	0	1
MX	0	1	9
PA	0	0	1
PE	0	0	1
PR	0	0	1
PY	0	0	1
SV	0	0	1
US	0	0	1
UY	0	0	1
VE	0	0	1
All	1	1	42

Cell Contents: Count

6. Análise de *clusters*

Esta seção dedica-se a análise de *clusters* ou agrupamentos. Estes agrupamentos podem se dar em relação às variáveis explicativas ou às observações.

O principal objetivo de agrupar as variáveis explicativas é possibilitar a criação de índices para simplificar as análises e modelos. Em nosso caso, o objetivo é tentar agrupar as quatro variáveis em um ou mais índices que bem representem o comportamento das cidades em relação às variáveis resposta. Já o objetivo de agrupar observações é tentar unir em grupos cidades com características semelhantes em relação a todas as variáveis, dependentes e explicativas.

6.1 Análise de Cluster

Cluster Analysis of Observations: Ranking 2008; Periculosida; ICUR; ...

Standardized Variables, Euclidean Distance, Single Linkage
Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined	New cluster	Number of obs. in new cluster	
1	41	98,3051	0,11549	29	30	29	2
2	40	97,2991	0,18404	10	11	10	2
3	39	95,4293	0,31145	38	40	38	2
4	38	94,9102	0,34682	10	15	10	3
5	37	94,8338	0,35202	9	13	9	2
6	36	94,7567	0,35727	33	34	33	2
7	35	94,6473	0,36473	9	10	9	5
8	34	93,4532	0,44609	16	21	16	2
9	33	92,8623	0,48636	24	29	24	3
10	32	92,8071	0,49012	39	42	39	2
11	31	92,5891	0,50497	9	18	9	6
12	30	92,5751	0,50593	32	36	32	2
13	29	92,4991	0,51110	32	33	32	4
14	28	92,3612	0,52051	9	14	9	7
15	27	92,3599	0,52059	16	20	16	3
16	26	91,9974	0,54529	26	28	26	2
17	25	91,9351	0,54954	19	25	19	2
18	24	91,8325	0,55653	24	27	24	4
19	23	91,5665	0,57465	9	17	9	8
20	22	91,3876	0,58684	38	41	38	3

21	21	91,3124	0,59197	24	31	24	5
22	20	90,8989	0,62014	9	24	9	13
23	19	90,5613	0,64315	32	39	32	6
24	18	90,4069	0,65367	9	35	9	14
25	17	89,9441	0,68520	37	38	37	4
26	16	89,8125	0,69417	6	9	6	15
27	15	89,7992	0,69508	16	23	16	4
28	14	89,5739	0,71043	2	3	2	2
29	13	89,2536	0,73225	16	19	16	6
30	12	89,0014	0,74944	6	26	6	17
31	11	88,8183	0,76191	6	12	6	18
32	10	88,6022	0,77664	6	8	6	19
33	9	86,8049	0,89911	6	37	6	23
34	8	86,0663	0,94943	22	32	22	7
35	7	85,2375	1,00591	6	16	6	29
36	6	84,6809	1,04383	6	22	6	36
37	5	83,8691	1,09915	6	7	6	37
38	4	83,1978	1,14489	4	5	4	2
39	3	82,1363	1,21722	4	6	4	39
40	2	78,3141	1,47766	1	4	1	40
41	1	69,0036	2,11207	1	2	1	42

Final Partition

Number of clusters: 2

Um pequeno número de Clusters (quatro no presente exemplo), permite um nível de similaridade confiável, de 83%.

De acordo com a Análise realizada pelo software estatístico, dois clusters já seriam suficientes para que houvesse um nível de similaridade confiável, de aproximadamente 78%.

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	40	133,189	1,67179	4,14653
Cluster2	2	0,252	0,35521	0,35521

Cluster Centroids

Variable	Cluster1	Cluster2	Grand centroid
Ranking 2008	0,077438	-1,54876	-0,000000
Periculosidade	0,080993	-1,61987	0,000000
ICUR	-0,130009	2,60018	-0,000000
Ranking 2007	0,083186	-1,66371	-0,000000

Distances Between Cluster Centroids

	Cluster1	Cluster2
Cluster1	0,00000	4,00538
Cluster2	4,00538	0,00000

6.2 Dendrograma e Distância Euclidiana

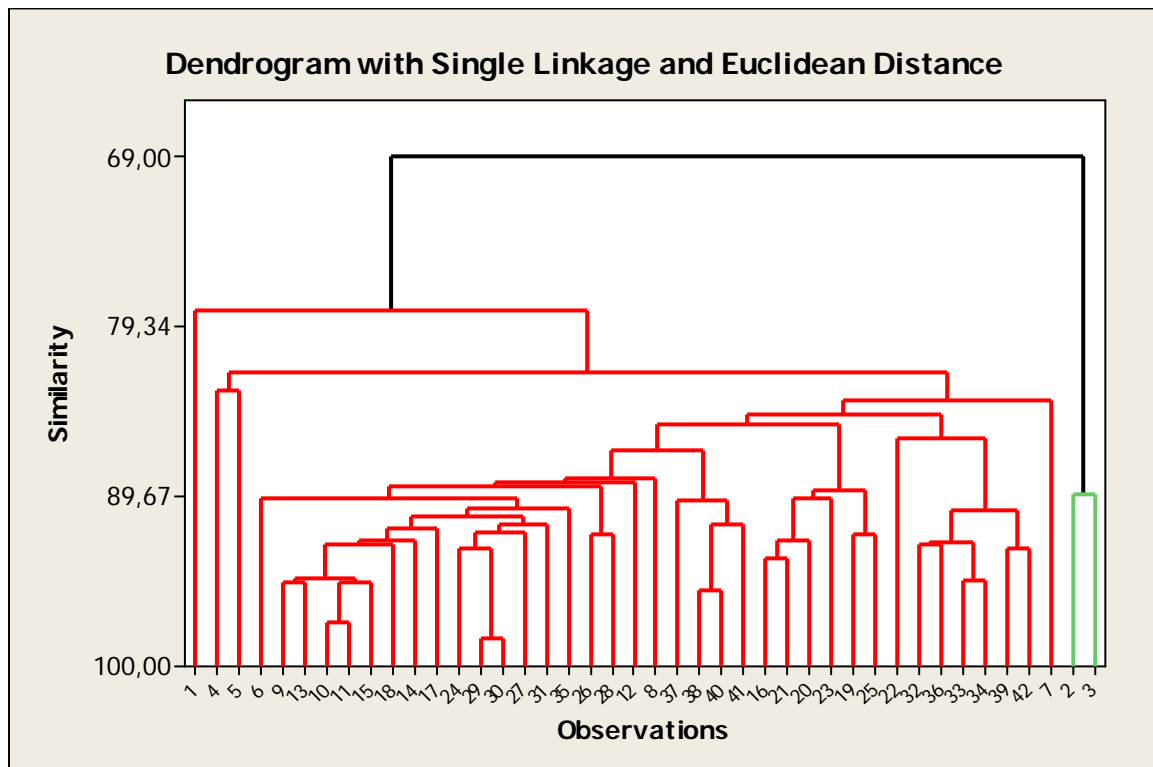
Quando se trabalha com variáveis quantitativas, a Distância Euclidiana comumente soma distâncias não comparáveis, como cm, kg, anos, milhões, etc.; muito embora a mudança de uma das unidades possa alterar completamente o significado e o valor do coeficiente.

A seguir, demonstra-se a fórmula para o cálculo da Distância Euclidiana:

$$d = \sqrt{\sum_{q=1}^M (g_q(l,p) - mg_q)^2}$$

onde:

- d representa a distância Euclidiana computada;
- $g_q(l,p)$ é o valor no nível de cinza no ponto de coordenadas l e p , na imagem q ;
- M é o número de componentes no vetor;
- mg_q é o valor médio que caracteriza a classe na imagem q .



O dendrograma acima foi feito a partir de dois clusters. A similaridade pode ser encontrada entre as 40 cidades de diferentes países e entre as 2 cidades de diferentes países: Miami (Estados Unidos da América) e Santiago (Chile).

6.3 Descriptive Statistics: Ranking 2008; Periculosidade; ICUR; Ranking 2007

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
Ranking 2008	42	0	21,50	1,89	12,27	1,00	10,75	21,50	32,25
Periculosidade	42	0	69,27	2,63	17,02	41,70	58,30	66,70	85,40
ICUR	42	0	35,84	1,62	10,49	21,00	30,45	34,35	38,08
Ranking 2007	42	0	21,48	1,90	12,31	1,00	10,75	21,50	32,25

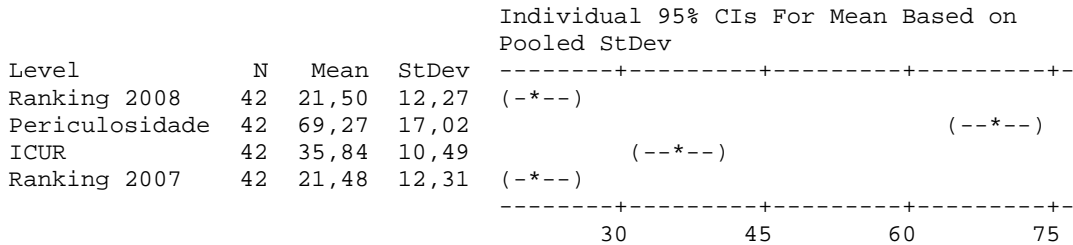
Variable	Maximum
Ranking 2008	42,00
Periculosidade	100,00
ICUR	68,20
Ranking 2007	42,00

Na presente análise, decidiu-se avaliar quatro índices: Ranking 2008, Periculosidade, ICUR e Ranking 2007. Pode-se afirmar que as médias possuem valores distintos, sendo que apenas o Ranking 2008 e Ranking 2007 apresentam valores mais próximos. O desvio-padrão gira em torno de 1,62 (valor mínimo), até 2,63 (valor máximo). O comportamento dos valores das medianas é similar as médias, valores distintos, porém idênticos para as variáveis Ranking 2008 (21,50) e Ranking 2007 (21,50).

6.4 One-way ANOVA: Ranking 2008; Periculosidade; ICUR; Ranking 2007

Source	DF	SS	MS	F	P
Factor	3	63997	21332	121,63	0,000
Error	164	28763	175		
Total	167	92759			

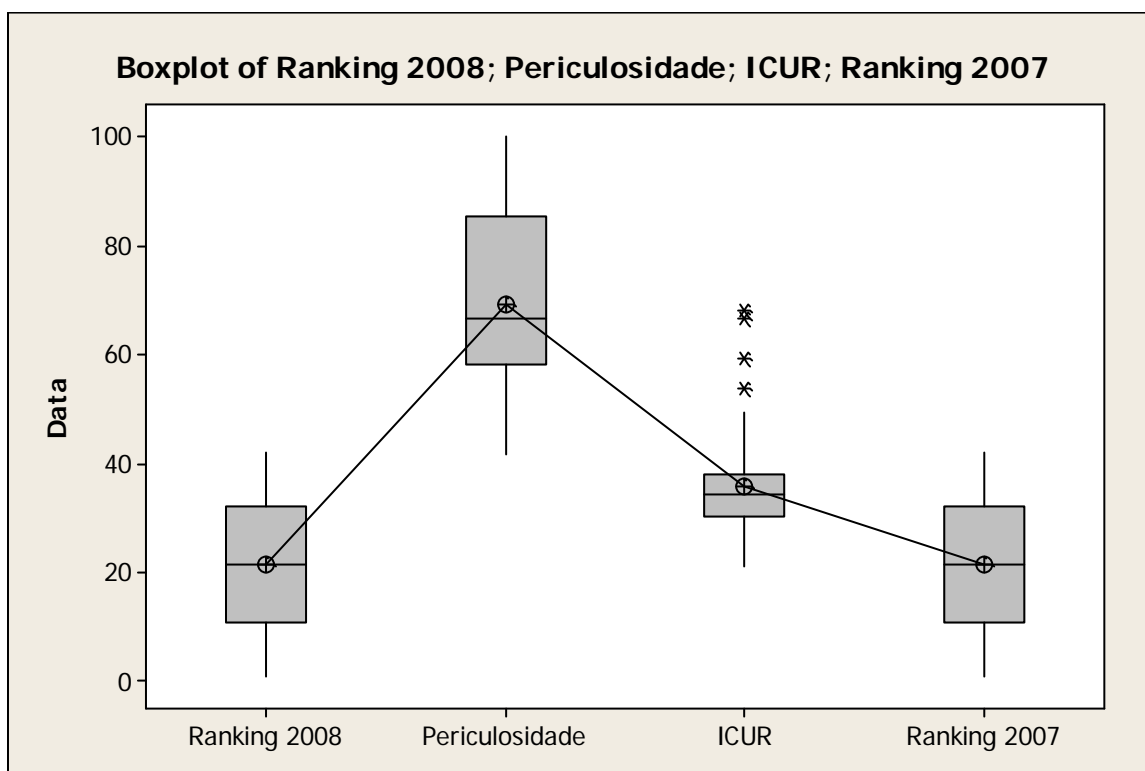
S = 13,24 R-Sq = 68,99% R-Sq(adj) = 68,42%



Pooled StDev = 13,24

A Análise One Way Anova permite uma visualização do que foi apontado anteriormente no item Estatística Descritiva, as distâncias entre as variáveis de acordo com a média, mediana e desvio-padrão.

6.5 Boxplot of Ranking 2008; Periculosidade; ICUR; Ranking 2007

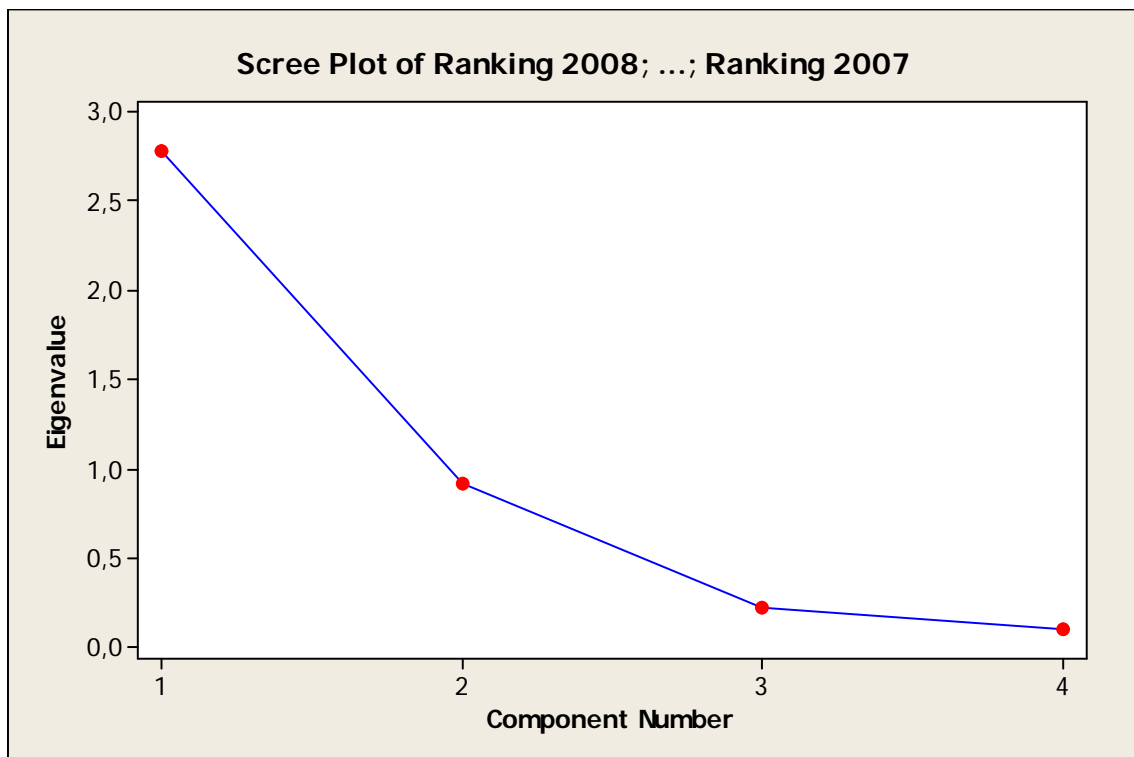


O presente Boxplot foi feito a partir de quatro variáveis: Ranking 2008, Ranking 2007, Periculosidade e ICUR. Foram observados três *outliers* na variável ICUR, são eles: Cidade do México, Santiago, Miami e São Paulo.

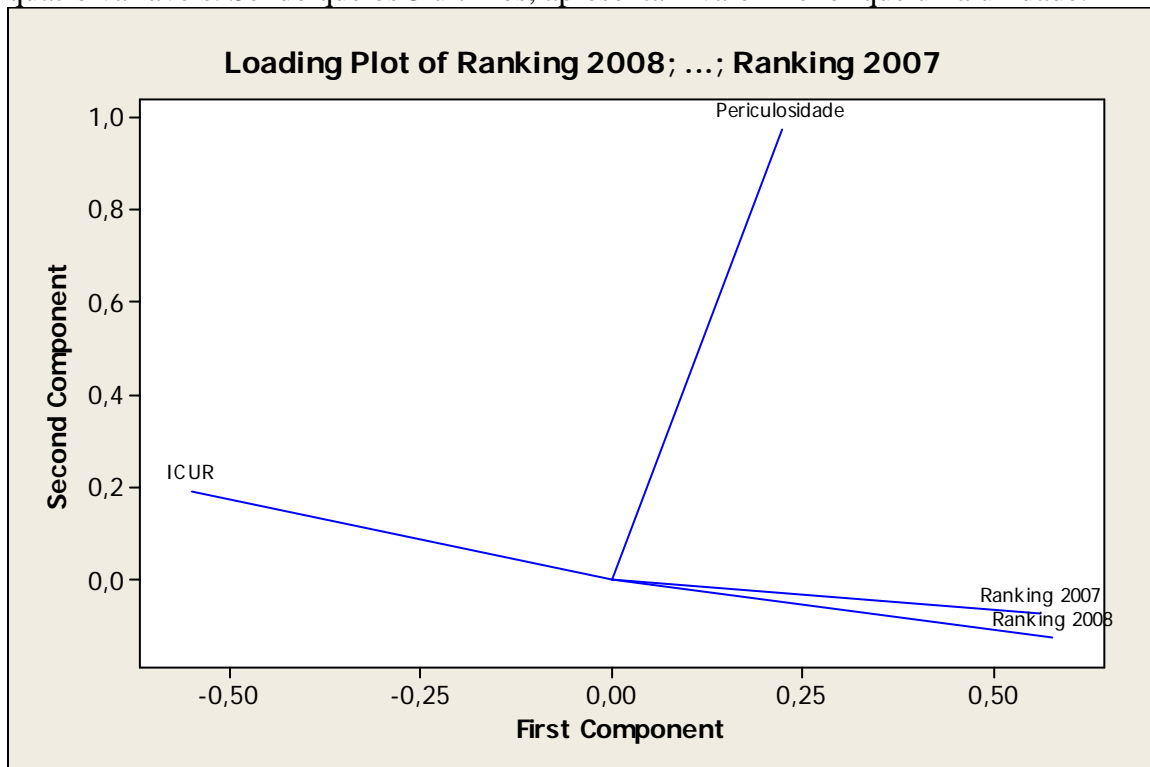
Pode-se observar que as variáveis comparadas, “Ranking 2008”, “Periculosidade”, “ICUR” e “Ranking 2007” apresentam resultados distintos. Considerando-se que estas variáveis possuem baixa correlação, com exceção das variáveis “Ranking 2008” e “Ranking 2007”, uma cidade pode estar bem posicionada no ICUR, porém, pode apresentar um alto índice de periculosidade, como é o caso de São Paulo, Cidade do México e Rio de Janeiro. Os *outliers* que mais se repetiram foram: Cidade do México, Santiago, Miami e São Paulo.

6.6 Principal Component Analysis

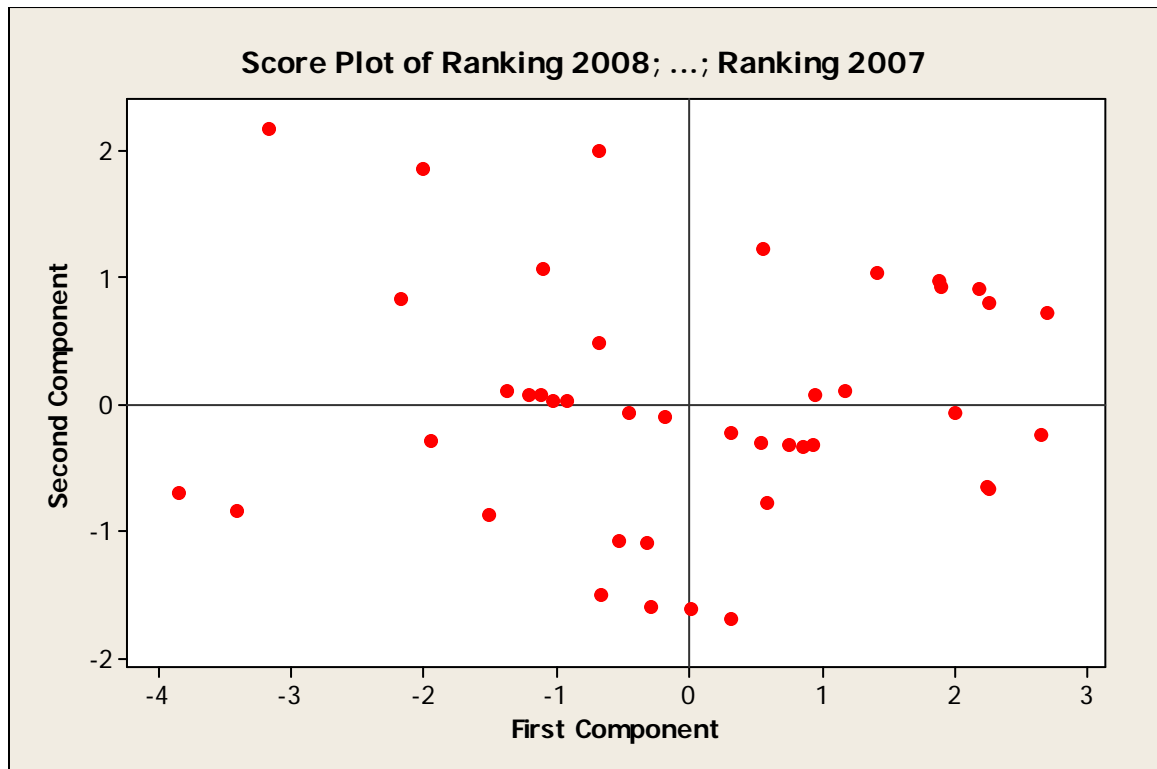
Ao se realizar a Análise de Componente Principal, optou-se pelos três diferentes tipos de gráficos: “Scree Plot”, “Loading Plot” e “Score Plot”



No gráfico acima, do tipo Scree Plot, pode-se observar que as 4 variáveis, Ranking 2008, Periculosidade, ICUR e Ranking 2007, apresentam um comportamento descendente, os quatro pontos vermelhos apontam para, respectivamente, cada uma das quatro variáveis. Sendo que os 3 últimos, apresentam valor menor que uma unidade.



No gráfico acima, do tipo Loading Plot, pode-se observar que novamente foram avaliadas as quatro variáveis citadas anteriormente: Ranking 2008, Periculosidade, ICUR e Ranking 2007. O gráfico aponta para uma alta similaridade entre as variáveis Ranking 2008 e Ranking 2007 na primeira componente e para um grande distanciamento entre as variáveis ICUR e Periculosidade, sendo que esta última, encontra-se perto do 1,0 na segunda componente.



No gráfico acima, do tipo Score Plot, pode-se observar que novamente foram avaliadas as quatro variáveis citadas anteriormente: Ranking 2008, Periculosidade, ICUR e Ranking 2007. O gráfico aponta para uma nuvem de maior concentração próxima do valor zero, tanto na primeira quanto na segunda componente. Chegando no máximo até a unidade um. Existe também uma grande concentração nos valores negativos, podendo-se observar duas cidades que encontram-se muito próximas do valor -4 para a primeira componente.

7. Considerações Finais

O objetivo deste artigo foi realizar análises estatísticas dos dados referentes ao ICUR 2008, composto por cidades da América Latina para se fazer negócios (totalizando 41 cidades), além de Miami (Estados Unidos da América) incluídos nos estudos da *AméricaEconomia Intelligence*. Também foram verificadas suas relações com os quatro indicadores escolhidos para que se pudesse analisar a competitividade urbana das cidades. As principais evidências encontradas como respostas às questões formuladas estão resumidas abaixo:

- Percebe-se que das cidades do estudo referentes à América Latina, apenas São Paulo, Santiago, Cidade do México e Buenos Aires aparecem com ICUR próximo a Miami. Interessante observar que em alguns quesitos há uma grande disparidade, como exemplo o índice de periculosidade, no qual todas apresentam valor elevado, enquanto que Miami não;
- Todas as cidades do México apresentaram ligeira melhora em comparação ao ICUR 2007 e 2006, enquanto que a maioria das cidades da América Central permaneceram no mesmo nível, e, com exceção a São Paulo e Buenos Aires, as cidades da América do Sul demonstraram um queda no índice;
- No quesito periculosidade o índice mostra alguns aspectos relevantes. Partindo da premissa que a periculosidade pode influenciar de modo significativo a vida social das pessoas e, conseqüentemente afetar os negócios, algumas cidades que tem uma posição de destaque deveriam observar que isso pode impactar de forma significativa. As cidades que apresentaram valor entre 91,7 e 100 foram: Rio de Janeiro, Cidade do México e São Paulo. Já Miami, Santiago e Monterrey apresentaram respectivamente 41,7; 41,7 e 58,3. Seria de extrema relevância que Rio de Janeiro, São Paulo e Cidade do México observassem esse fator como preponderante para tentar atingir níveis parecidos com as cidades de Miami, Santiago e Monterrey;
- Algumas contribuições para aumentar o poder explicativo e também o potencial o ICUR, seria considerar algumas outras dimensões que possam influenciar no índice de competitividade urbana, como transportes públicos, índice de poluição urbana, trânsito, etc. (para obter mais informações veja Observatório Nossa São Paulo – www.nossasaopaulo.org.br);

Apesar dos resultados favoráveis, vale ressaltar as limitações desta pesquisa. A principal delas é que o trabalho foi baseado em técnicas estatísticas que, como todas, possuem margem de erro e requerem cautela quanto à sua interpretação e, principalmente, extrapolação das conclusões. Um exemplo disso é a base de dados utilizada para a criação da maioria dos modelos, considerando apenas o ano de 2008, o que não permite a utilização dos mesmos para uma série histórica de dados para outros anos.

Desta forma, as evidências encontradas não podem ser interpretadas como respostas definitivas às questões levantadas neste trabalho, mas, sem dúvida, contribuíram para a exploração do tema e para o desenvolvimento de pesquisas futuras, dentre as quais podemos citar:

- A utilização de série de dados temporais do ICUR, em relação a outros indicadores (PIB países, PIB cidades, etc.), com o intuito de construir modelos estatísticos temporais.
- A inclusão de novas variáveis aos modelos, como exemplo, a taxa de crescimento populacional, mudança de classe social (Classe E para D; Classe D para C; Classe C para B; e, Classe B para A).

- A exploração de outros indicadores de riqueza ou de “bem-estar” social, como o FIB (Felicidade Interna Bruta) e também indicadores de Qualidade de Vida Urbana (indicadores do observatório Nossa São Paulo), com o intuito de enriquecer o conteúdo do estudo, partindo do pressuposto que para uma cidade ter alto índice de competitividade urbana, a população local deve possuir condições apropriadas e adequadas para viver e trabalhar.

8. BIBLIOGRAFIA

AMÉRICA ECONOMIA INTELLIGENCE, ICUR, 2008. Acesso <<http://www.americaeconomia.com.br>> em junho de 2008.

BANCO MUNDIAL. Acesso <www.worldbank.org> em junho de 2008.

BARTHOLOMEW, D. J.; STEELE, F.; MOUSTAKI, I.; GALBRAITH, J. I. *The Analysis and Interpretation of Multivariate Data for Social Scientists*. New York: Chapman & Hall/CRC, 2002.

ISI WEB OF KNOWLEDGE, 2008. Acesso <www.isiwebofknowledge.com> em junho de 2008.

MINITAB – MINITAB Statistical Software – Release 14.12 da Minitab Inc. (2004)

MOORE, D. S.; MCCABE, G. P. *Introdução à Prática da Estatística*. 3a. Edição, Rio de Janeiro: LTC, 2002.

NOSSA SÃO PAULO 2008. Acesso <www.nossasaopaulo.org.br/> em maio de 2008.

PNUD, BRASIL 2008. Acesso <www.pnud.org.br/> em junho de 2008.

RELATÓRIO DE DESENVOLVIMENTO HUMANO, ONU, 2008. Acesso <<http://hdr.undp.org/hdr2008/statistics/>> em junho de 2008.