



**PONTÍFICA UNIVERSIDADE CATÓLICA DE
SÃO PAULO**

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

**UMA ANÁLISE COMPARATIVA DE DESENVOLVIMENTO
POR ESTADO NO BRASIL A PARTIR DE ALGUNS
INDICADORES DO IBGE**

Aluno: Daniel Troise

Prof. Arnaldo José de Hoyos Guevara

1º Semestre 2012

INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise exploratória dos dados e outras análises estatísticas usando como base de dados os 27 Estados Brasileiros.

Nas páginas a seguir, teremos a oportunidade de analisar e tentar entender os resultados das análises de algumas variáveis sobre estes Estados, tais como: PIB, IDH, nº de habitantes, taxa de analfabetismo, dentre outros.

As simulações de dados e os gráficos referentes aos resultados obtidos foram tabulados e analisados com a ferramenta MINITAB 16.

Todas as bases de dados foram obtidas através do site do IBGE, instituto Brasileiro de Geografia e Estatística, com acesso de 18 a 27 de junho. Todos os dados e informações se referem ao ano de 2009, segundo o site do IBGE.

<http://www.ibge.gov.br/home/>

The screenshot shows the IBGE website homepage with the following content:

- Navigation:** "Acesso à Informação" and "BRASIL" buttons.
- Menu:** Indadores, População, Economia, Geociências, Canais, Download, Pesquisas, Sala de Imprensa.
- Calendários:** Indadores, Pesquisas Estruturais, Censo 2010, Geociências.
- Canais:** Banco de Dados, Canais temáticos, Produtos e Serviços, Conteúdo Histórico, Projetos e Entidades, IBGE Interativo, Artigos e Apresentações.
- Últimos Resultados:**
 - PME Mai/2012 (taxa de desocupação): 5,8%
 - IPCA-15 Jun/2012: 0,18%
 - PMC Abr/2012 (comércio varejista): 6,0%
 - PIMES Abr/2012 (Pessoal ocupado assalariado): -0,3%
- Variação dos Indicadores:** Line graph showing the variation of the unemployment rate (%) for the total of 6 areas from December to May. The y-axis ranges from -3 to 9. The x-axis shows months: Dez, Jan, Fev, Mar, Abr, Mai.
- Destques:** Sistema de Contas Trimestrais; Nota de esclarecimento; Nota técnica do SNIPC; Resultados do IPCA/INPC.
- Loja Virtual:** Indadores de desenvolvimento sustentável 2012 - Brasil.
- Resultados do Censo 2010:** nova divulgação.
- ENCE:** Escola Nacional de Ciências Estatísticas.

<http://seriesestatisticas.ibge.gov.br/Default.aspx>

The screenshot shows the IBGE website interface. At the top, there is a navigation bar with 'Página inicial', 'Fale conosco', and 'Pesquisar'. Below this is a search bar with the text 'Busca:' and a 'Pesquisa Avançada' link. The main content area displays a table of registered series. On the left side, there is a sidebar menu with various categories like 'Glossário', 'Nível Geográfico', 'Fontes', and 'Temas'. The table lists series with columns for 'Código', 'Séries Cadastradas', and 'Período'.

Código	Séries Cadastradas	Período
SEE35	Abandono escolar por Dependência Administrativa - Ensino Fundamental (série nova)	2007-2010
SEE15	Abandono escolar por nível de ensino (série encerrada)	1999-2005
M14	Abandono escolar por Série - Ensino Fundamental de 8 e 9 anos (série nova)	2007-2010
M15	Abandono escolar por série - Ensino Médio (série nova)	2007-2010
ST91	Abertura da economia	1995-2011**
SCN44	Abertura da economia	1947-2011
IU28	Acesso ao esgotamento sanitário - área urbana	1992-2009
IU27	Acesso a esgotamento sanitário - área rural	1992-2009
IU46	Acesso à Internet	2005-2009
IU22	Acesso ao sistema de abastecimento de água - área rural	1992-2009
IU25	Acesso ao sistema de abastecimento de água - área urbana	1992-2009
IU37	Adequação da moradia	1992-2009
IU38	Adequação da moradia, por critérios de adequação	1992-2009
SEE13	Aprovação por nível de ensino (série encerrada)	1999-2005
SEE30	Aprovação por Série - Ensino Fundamental de 8 e 9 anos (série nova)	2007-2010
M13	Aprovação por Série - Ensino Médio (série nova)	2007-2010
M101	Aprovação, reprovação e abandono - Ensino Fundamental (série nova)	2007-2010
M12	Aprovação, reprovação e abandono - Ensino Médio (série nova)	2007-2010
AGRO02	Área dos estabelecimentos (ha)	1920-2006
AGRO26	Área dos estabelecimentos por condição legal do responsável (ha)	1920-2006
CAT5	Área média de lavouras por trator (ha)	1920-2006

Agradecimento:

“ Deixo aqui um agradecimento aos colegas de turma, pelo companheirismo durante as aulas e um especial ao Professor Doutor Arnoldo Jose de Hoyos, pelos ensinamentos, além do entusiasmo e dedicação com que conduziu a disciplina.”

A BASE DE DADOS.

Abaixo, elencamos a tabela da dados obtidos no site do IBGE e que fundamentou o trabalho.

A Tabela de Dados

Estados	% de participação no PIB nacional	% da População com acesso à internet	Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	Taxa de mortalidade infantil para cada 1.000 nascidos vivos	IDH	PIB por Estado em R\$ mil	População Por Estado	PIB percapita em R\$ mil	Área em (Km²)	Nº Municípios por Estado
Acre	0,22	22,40	13,77	8,38	23,39	0,78	7.386.000	732.793	10,08	164.122,28	22
Alagoas	0,67	12,30	25,74	11,77	21,53	0,72	21.235.000	3.120.922	6,80	27.779,34	102
Amapá	0,23	12,70	4,11	16,96	28,91	0,80	7.404.000	668.689	11,07	142.827,90	16
Amazonas	1,58	13,20	8,28	25,98	21,92	0,79	49.614.000	3.480.937	14,25	1.559.161,68	62
Bahia	4,12	17,10	17,30	9,56	23,30	0,76	137.075.000	14.021.432	9,78	564.830,86	417
Ceará	1,89	12,70	19,06	11,18	19,12	0,74	65.704.000	8.488.055	7,74	148.920,54	184
Distrito Federal	3,76	53,00	4,02	18,22	11,89	0,90	131.487.000	2.562.963	51,30	5.787,78	1
Espírito Santo	2,27	30,50	8,81	17,84	14,46	0,82	66.763.000	3.512.672	19,01	46.098,57	78
Goiás	2,45	21,60	9,31	12,23	16,89	0,82	85.615.000	6.004.045	14,26	340.103,47	246
Maranhão	1,19	9,30	19,46	12,50	24,03	0,72	39.855.000	6.569.683	6,07	331.935,51	217
Mato Grosso	1,60	23,30	9,55	18,43	21,76	0,80	57.294.000	19.595.309	2,92	586.520,37	853
Mato Grosso do Sul	1,06	24,80	8,12	19,56	16,54	0,83	36.368.000	2.449.341	14,85	357.145,84	78
Minas Gerais	9,07	26,90	8,65	13,07	17,38	0,82	287.055.000	3.033.991	94,61	903.329,70	141
Pará	1,86	10,30	11,86	17,25	23,64	0,78	58.402.000	7.588.078	7,70	1.247.950,00	143
Paraíba	0,83	15,50	23,49	9,86	21,18	0,75	28.719.000	3.766.834	7,62	56.469,47	223
Paraná	6,07	32,50	6,57	17,07	13,09	0,84	189.992.000	10.439.601	18,20	199.316,69	399
Pernambuco	2,34	15,40	17,86	15,10	19,31	0,74	78.428.000	8.796.032	8,92	98.146,32	185
Piauí	0,53	10,60	24,33	9,52	23,62	0,74	19.033.000	3.119.015	6,10	251.576,84	224
Rio de Janeiro	11,15	36,10	4,44	26,27	14,31	0,85	353.878.000	3.168.133	111,70	52.810,70	167
Rio Grande do Norte	0,85	15,20	19,99	8,82	21,08	0,75	27.905.000	10.695.532	2,61	268.781,90	496
Rio Grande do Sul	6,64	31,00	4,97	41,17	12,76	0,84	215.864.000	15.993.583	13,50	43.780,16	92
Rondônia	0,56	21,20	9,17	17,27	21,58	0,78	20.236.000	1.560.501	12,97	237.590,86	52
Roraima	0,16	15,80	9,26	31,74	20,73	0,78	5.593.000	451.227	12,40	224.301,04	15
Santa Catarina	3,93	36,50	4,41	29,59	11,69	0,86	129.806.000	6.249.682	20,77	95.703,49	293
São Paulo	33,92	39,50	4,74	19,56	12,59	0,85	1.084.353.000	41.252.160	26,29	248.196,96	645
Sergipe	0,63	17,80	16,87	13,20	20,40	0,77	19.767.000	2.068.031	9,56	21.918,35	75
Tocantins	0,42	13,80	14,38	10,62	21,83	0,78	14.571.000	1.383.453	10,53	277.621,86	139

Estados	Total de Municípios	Municípios com a Lei Geral Regulamentada	%
Total	5.566	3.820	68,63%
AC	22	22	100,00%
AL	102	88	86,27%
AM	62	47	75,81%
AP	16	15	93,75%
BA	417	306	73,38%
CE	184	139	75,54%
DF	1	1	100,00%
ES	78	78	100,00%
GO	246	158	64,23%
MA	217	105	48,39%
MG	853	420	49,24%
MS	78	60	76,92%
MT	141	141	100,00%
PA	143	74	51,75%
PB	223	106	47,53%
PE	185	110	59,46%
PI	224	121	54,02%
PR	399	353	88,47%
RJ	92	92	100,00%
RN	167	150	89,82%
RO	52	43	82,69%
RR	15	15	100,00%
RS	497	426	85,71%
SC	293	293	100,00%
SE	75	34	45,33%
SP	645	310	48,06%
TO	139	113	81,29%

AS VARIÁVEIS

São 11 variáveis quantitativas que estaremos utilizando para a base de cálculo deste trabalho.

Assim, todos os exercícios serão feitos com base nestas variáveis, as quais iremos elencar a seguir. As variáveis serão, na sua maioria, quantitativas..

Tabela 1. As Variáveis

Variável	Significado	Tipo	Unidade de Medida
% de participação no PIB nacional	Refere-se ao % de participação no PIB do Estado.	Variável Quantitativa	Percentual
% da População com acesso à internet	Refere-se ao % da população com acesso à internet.	Variável Quantitativa	Percentual
Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	Refere-se ao % da população de pessoas analfabetas com 15 ou mais anos de idade.	Variável Quantitativa	Percentual
Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	Refere-se ao % da população para cada 100.000 habitantes que contraíram o vírus HIV.	Variável Quantitativa	Percentual
Taxa de mortalidade infantil para cada 1.000 nascidos vivos	Refere-se ao % da população para cada 1.000 nascidos vivos que vem a falecer.	Variável Quantitativa	Percentual
IDH	Refere-se ao índice de desenvolvimento Humano.	Variável Quantitativa	Percentual
PIB por Estado em R\$ mil	Refere-se ao Produto Interno		

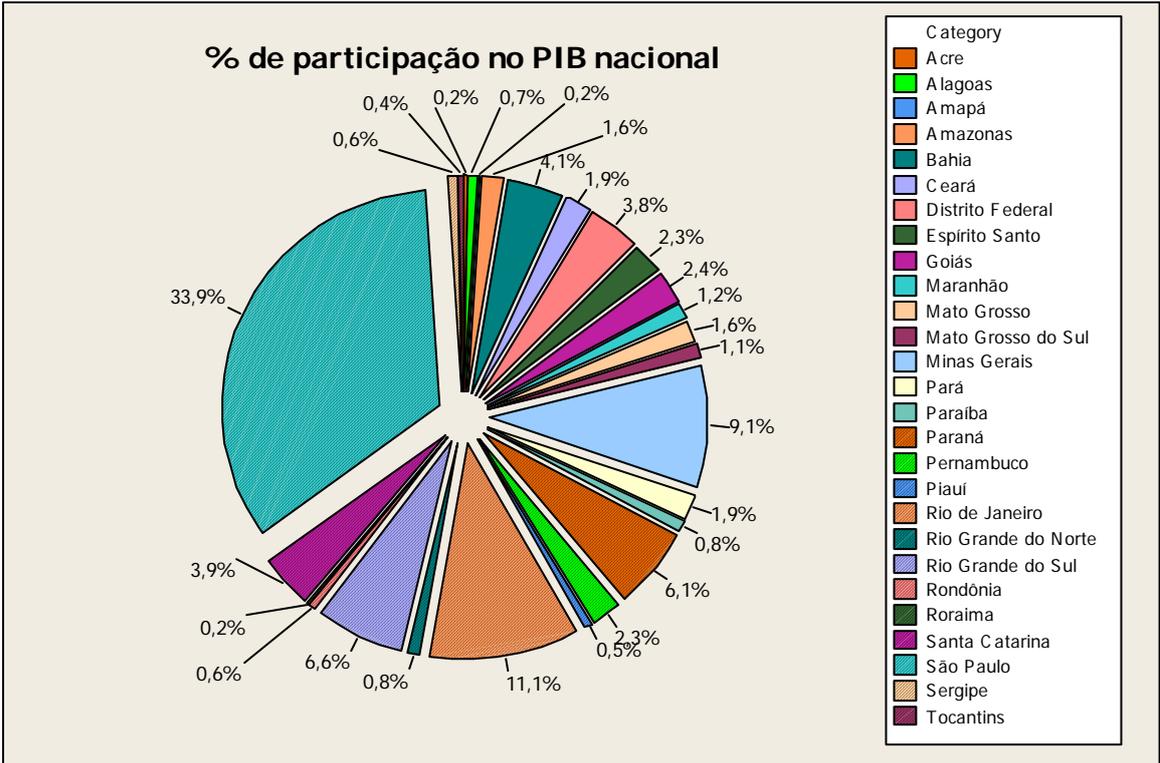
	Bruto do Estado.	Variável Quantitativa	R\$
População Por Estado	Refere-se à numero absoluto da habitantes do Estado.	Variável Quantitativa	Número Absoluto
PIB percapita em R\$ mil	Refere-se ao valor do PIB do Estado dividido pelo número de habitantes do Estado.	Variável Quantitativa	R\$
Área em (Km ²)	Refere-se à área do Estado em Km ² .	Variável Quantitativa	Km ²
Nº Municípios por Estado	Refere-se ao número de municípios do Estado.	Variável Quantitativa	Número Absoluto

Variáveis	Nomenclatura
% de participação no PIB nacional	C2
% da População com acesso à internet	C3
Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	C4
Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	C5
Taxa de mortalidade infantil para cada 1.000 nascidos vivos	C6
IDH	C7
PIB por Estado em R\$ mil	C8
População Por Estado	C9
PIB percapita em R\$ mil	C10
Área em (Km ²)	C11
Nº Municípios por Estado	C12

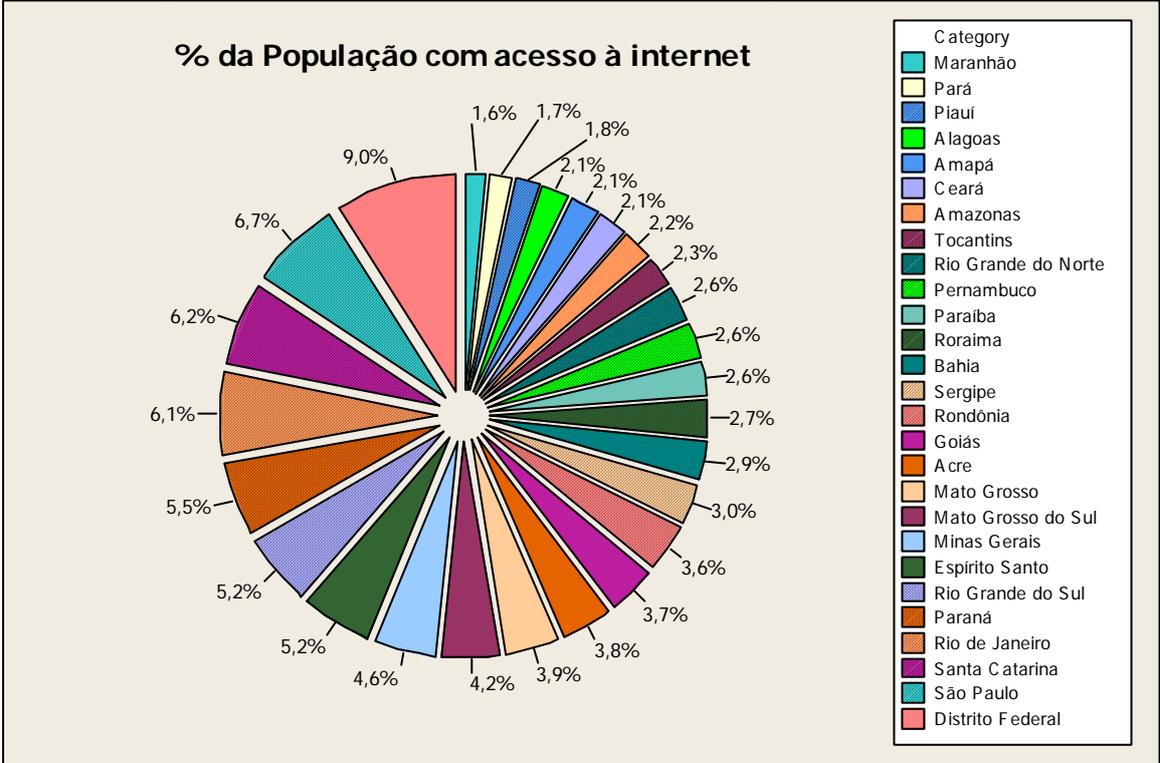
Pie Chart

A seguir, elencamos os gráficos de Pie Chart de algumas variáveis.

Com os gráficos Pie Chart poderemos visualizar melhor a participação de cada um dos Estados em cada uma das variáveis estudadas, dessa forma, teremos uma visualização melhor das variações de cada uma das variáveis.

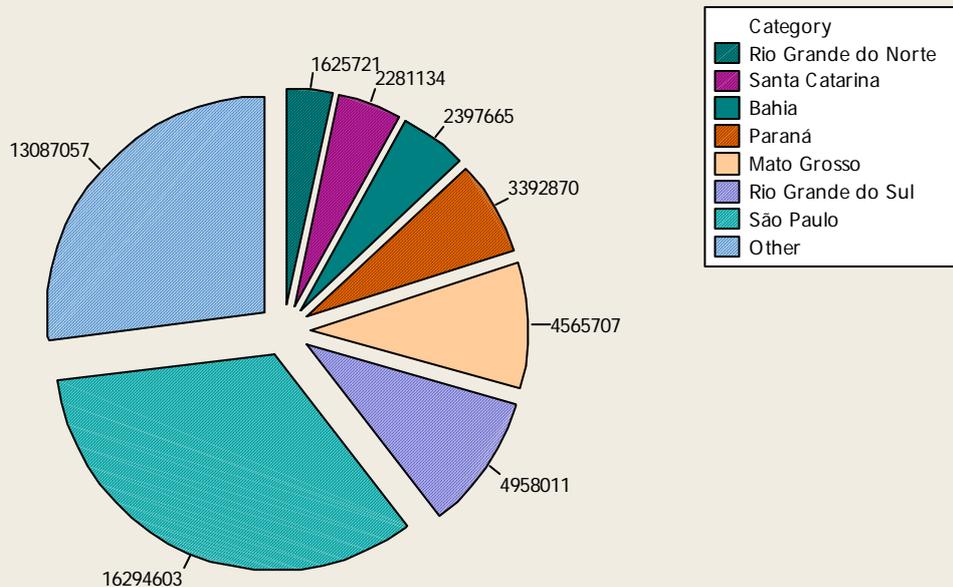


Neste caso, destaca-se o Estado de São Paulo como o maior participante do PIB nacional, seguido de Rio de Janeiro e Minas Gerais, respectivamente.



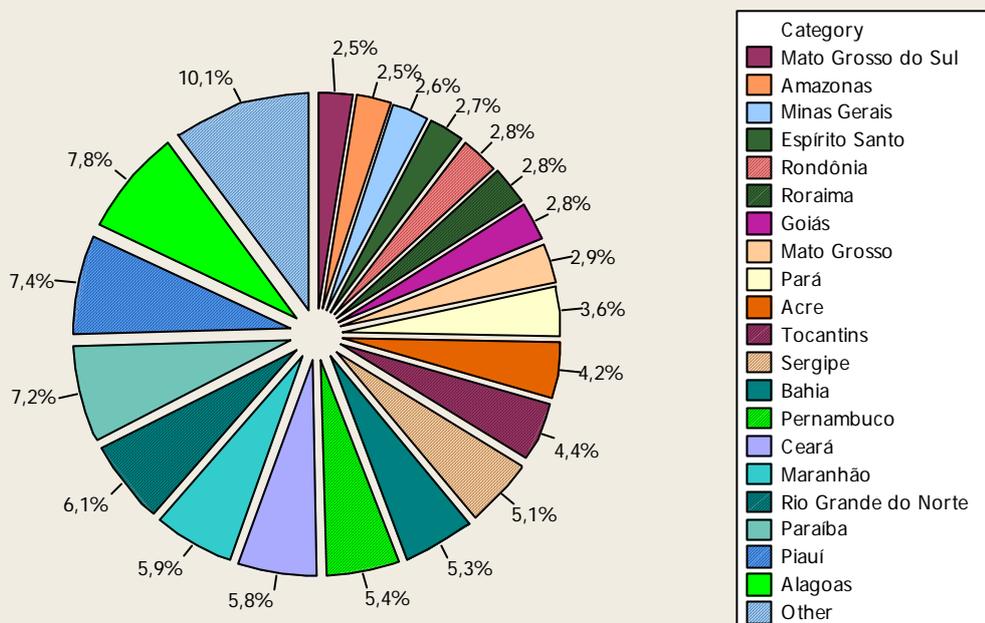
Em relação ao % da população com acesso a internet, destaca-se o Distrito Federal, Estado de São Paulo, seguido de Santa Catarina e Rio de Janeiro.

Número de pessoas com acesso à internet

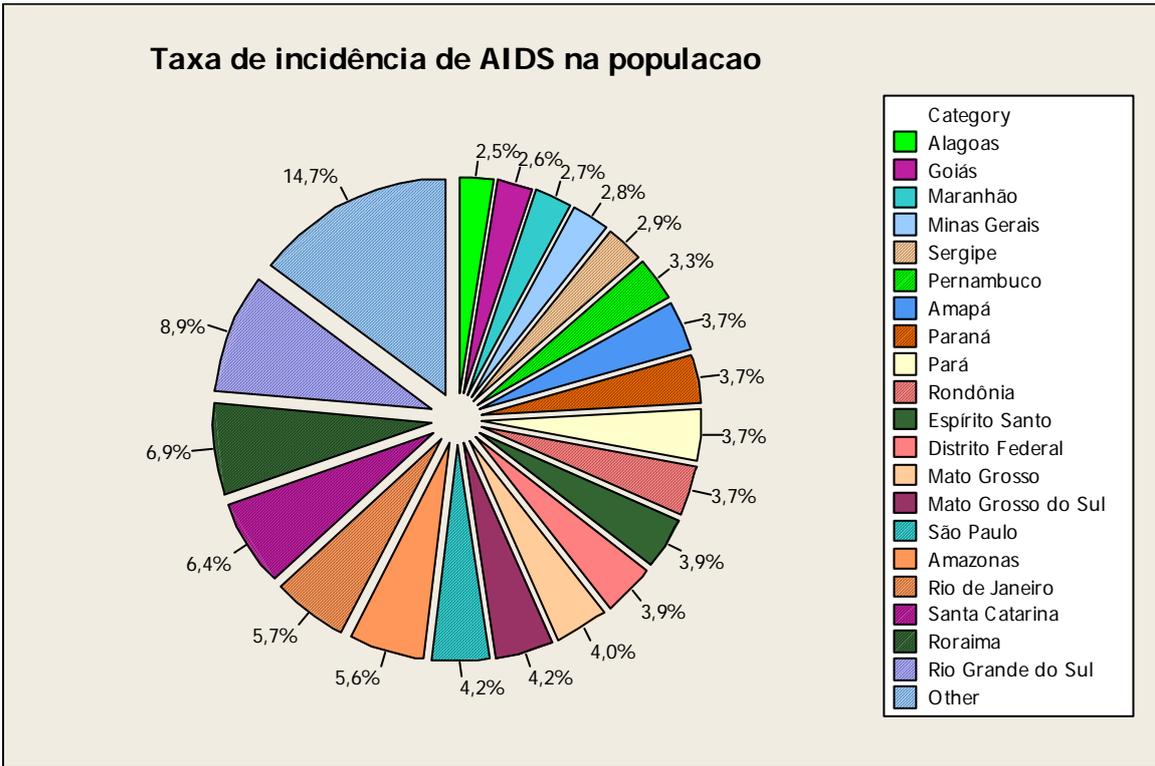


O número de pessoas com acesso a internet segue a mesma tendência, tendo São Paulo e Rio grande do Sul com o maior número de pessoas.

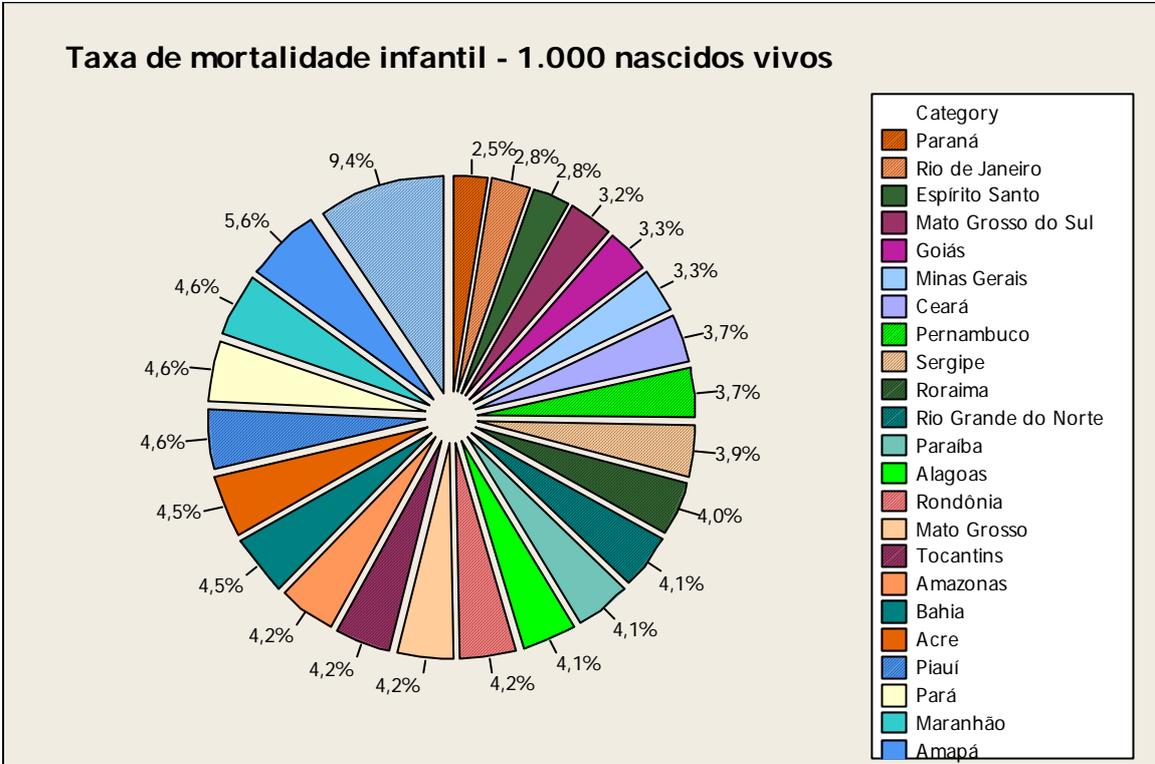
Taxa de analfabetismo absoluto



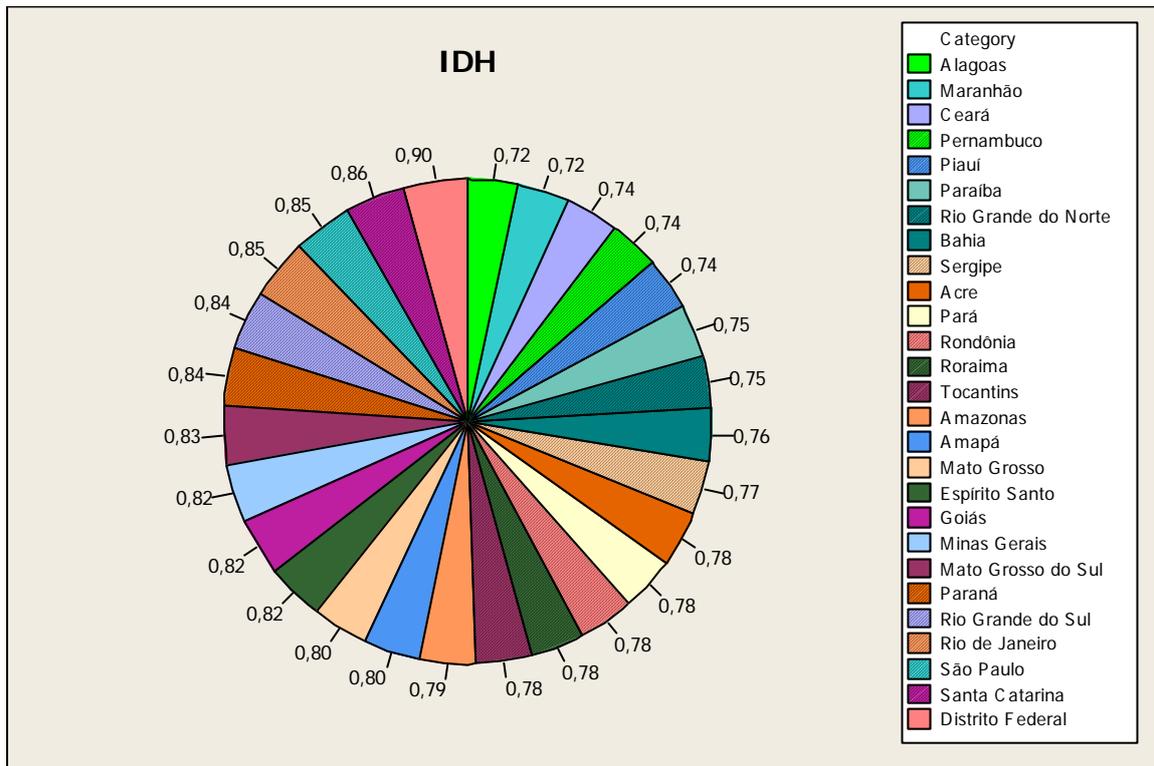
Liderando a taxa de analfabetismo temos os Estados de Alagoas, Piauí e Paraíba.



No caso do vírus HIV, os Estados que se destacam são os do Rio grande do Sul e Roraima.



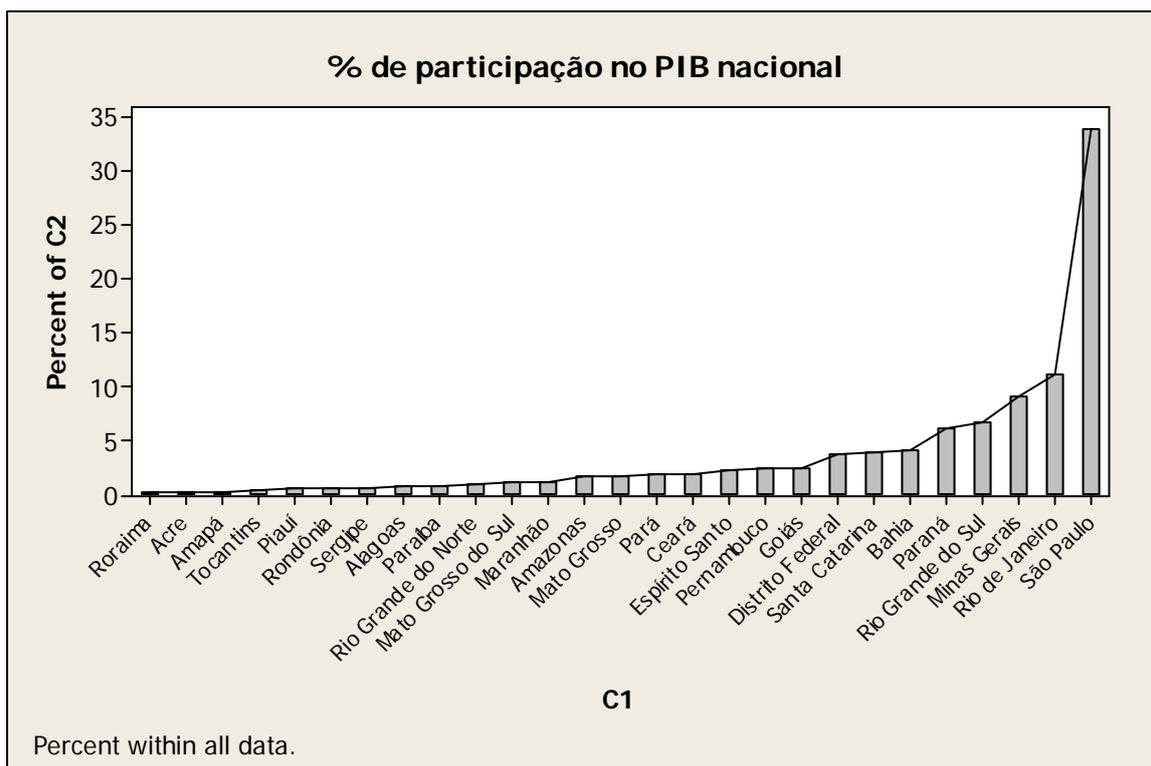
Para o índice de mortalidade infantil, o Estado que lidera é o do Amapá, seguido do Maranhão e Pará.



Em relação ao IDH, destacam-se Santa Catarina, São Paulo e Rio de Janeiro.

Bar Chart

A seguir, elencamos os gráficos de Bar Chart de algumas variáveis. Da mesma forma que o Pie Chart, o Bar Chart também permite ter uma visão mais destacada de do comportamento de cada uma das variáveis relativas a cada um dos Estados da Federação.

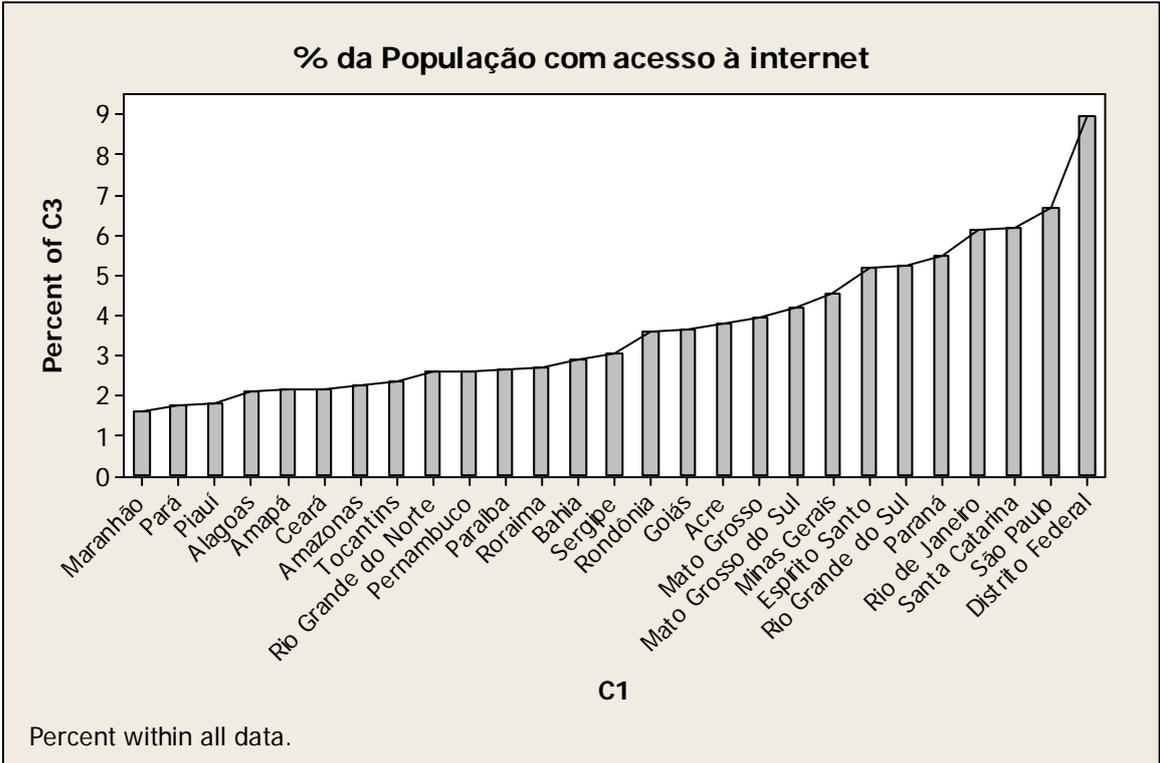


Na participação do PIB nacional, destacam-se São Paulo, Rio de Janeiro e Minas Gerais. Em último, temos os Estados de Roraima, Acre, Amapá e Tocantins.

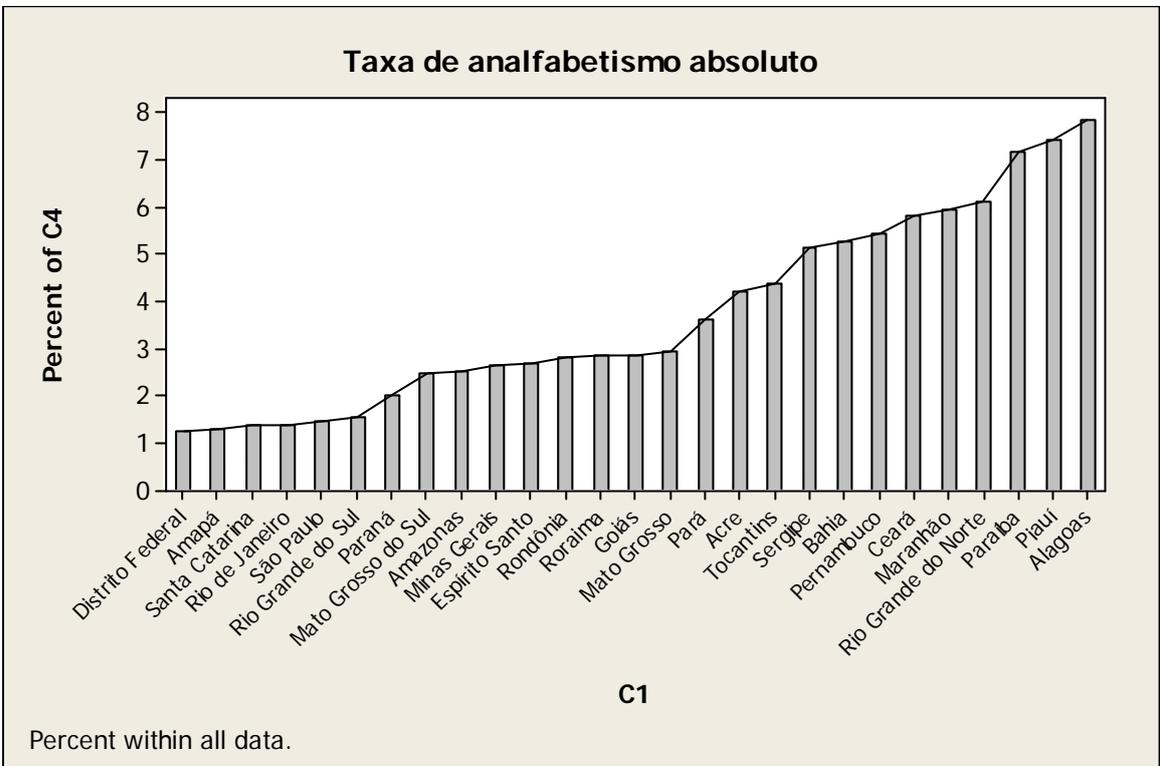
Vale lembrar que PIB significa, Produto Interno Bruto.

“O PIB (Produto Interno Bruto) é a soma de todos os bens e serviços produzidos em um país durante certo período. Isso inclui do pãozinho até o apartamento de luxo.”

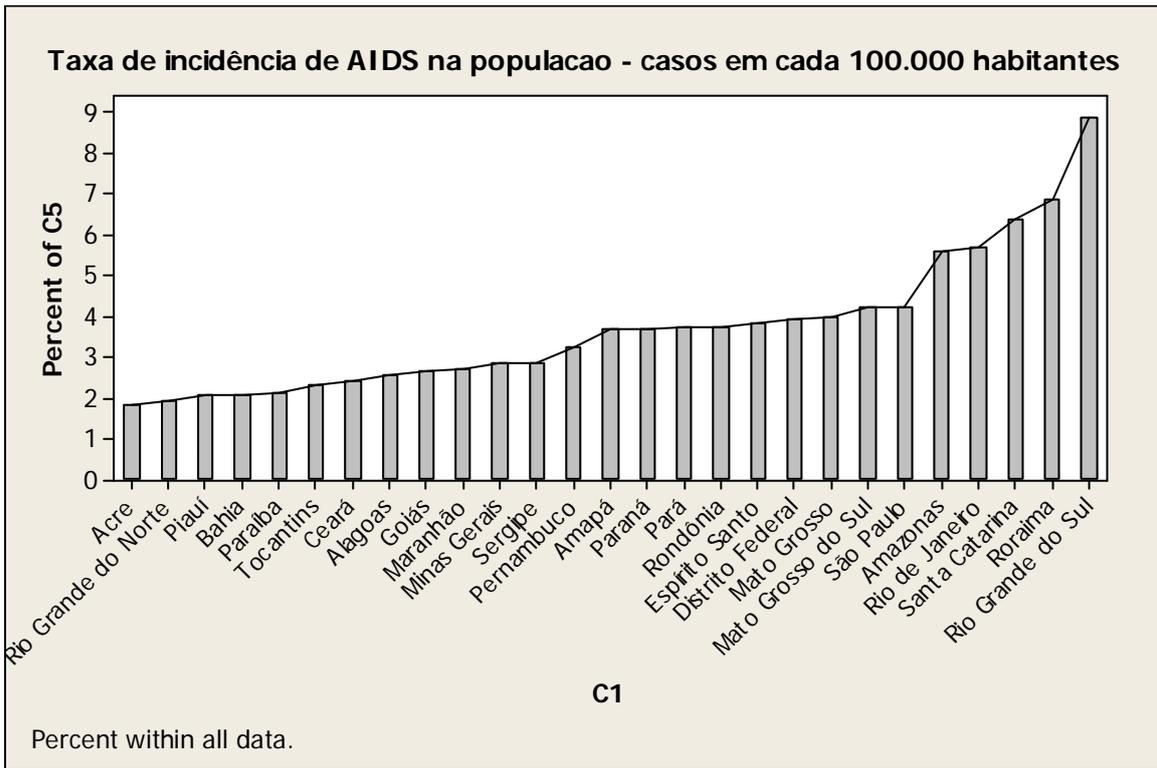
“O índice só considera os bens e serviços finais, de modo a não calcular a mesma coisa duas vezes. A matéria-prima usada na fabricação não é levada em conta. No caso de um pão, a farinha de trigo usada não entra na contabilidade”.



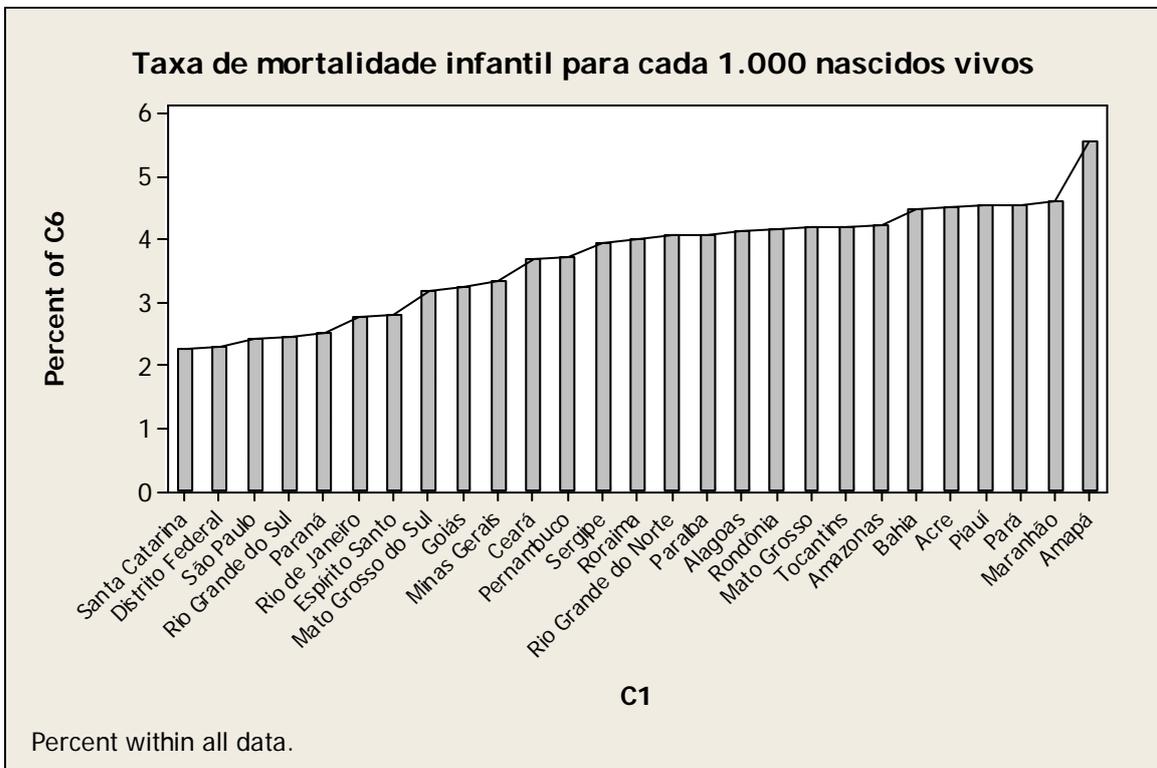
Nesta variável destacam-se o Distrito Federal, São Paulo e Santa Catarina.



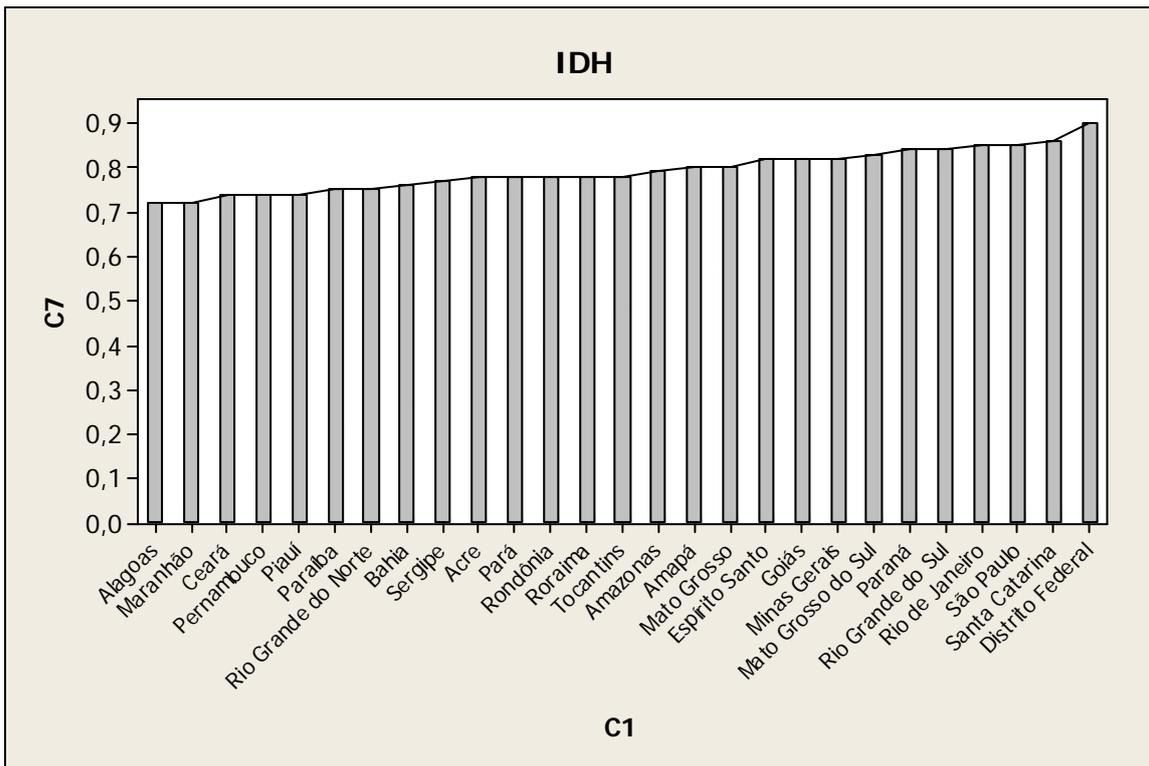
Nesta variável destacam-se os Estados de Alagoas, Piauí e Paraíba.



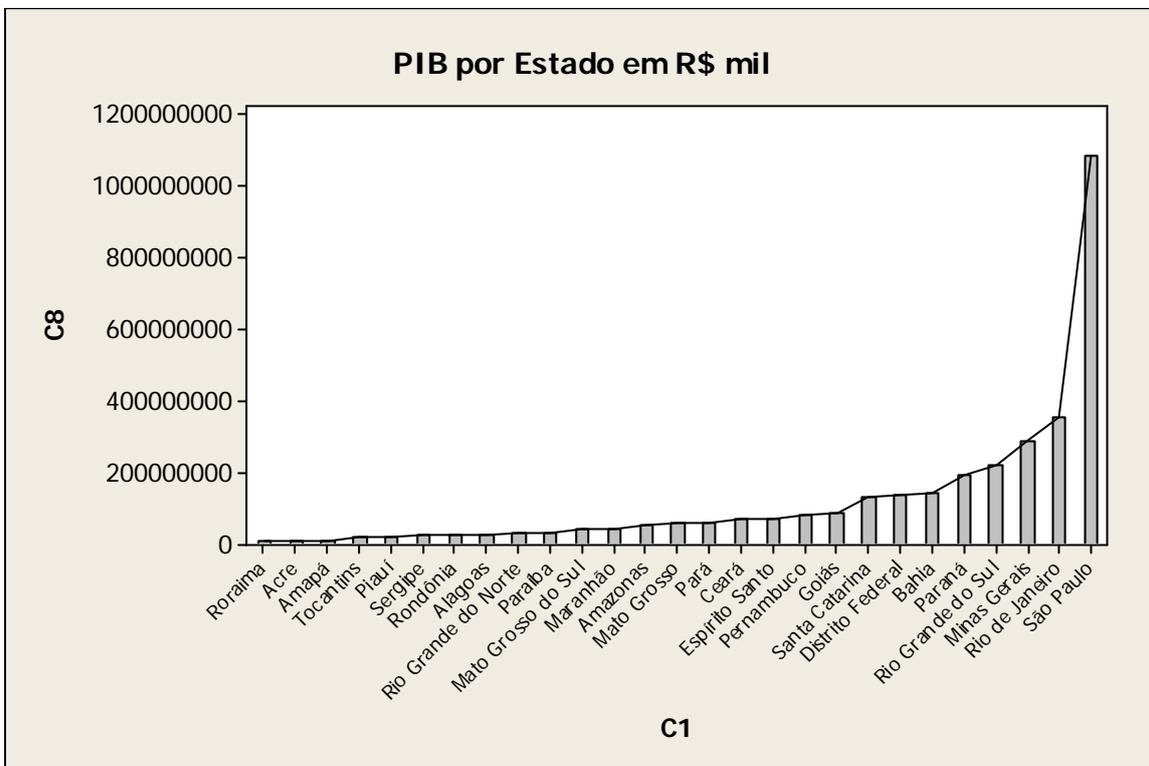
O Rio Grande do Sul, destaca-se nos índices de HIV, seguido de Roraima e Santa Catarina.



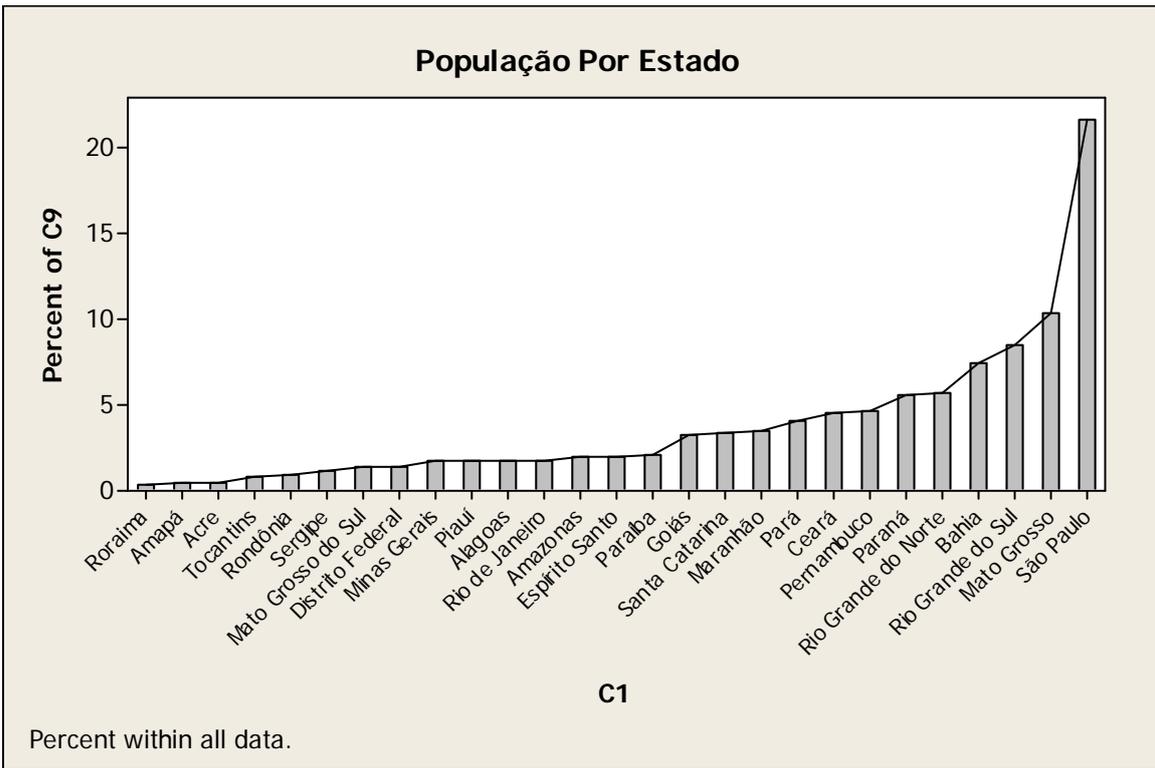
Na taxa de mortalidade, lideram os Estados do Amapá e Maranhão.



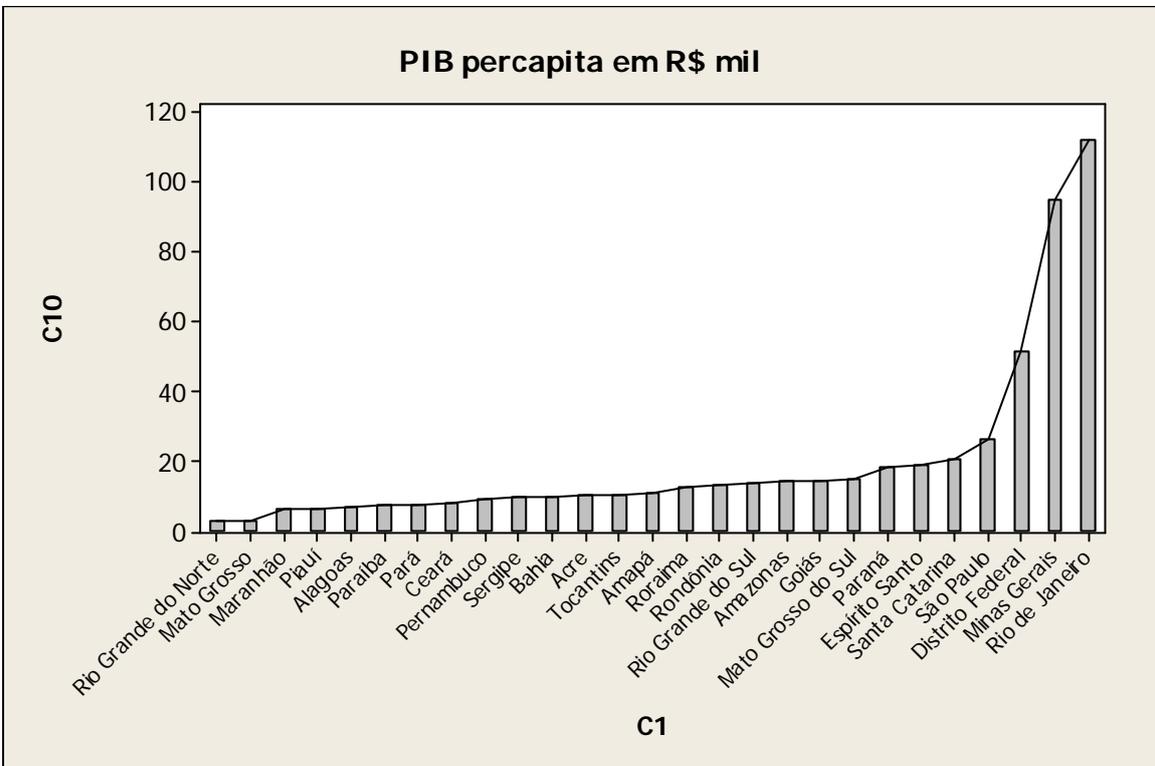
No IDH, o Distrito Federal, Santa Catarina e São Paulo, lideram o ranking.



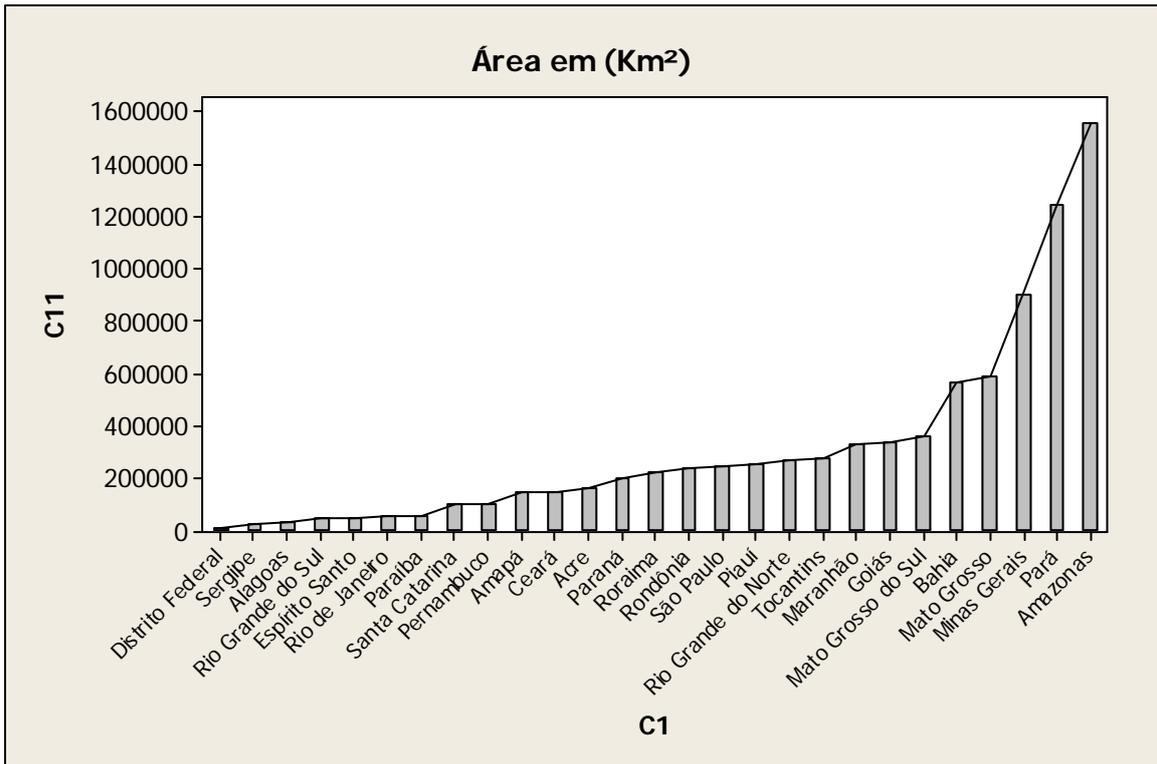
O PIB de São Paulo tem um destaque muito grande frente aos demais Estados, confirmando-se como Estado mais rico e com maior PIB do País, seguida de Rio de Janeiro e Minas Gerais.



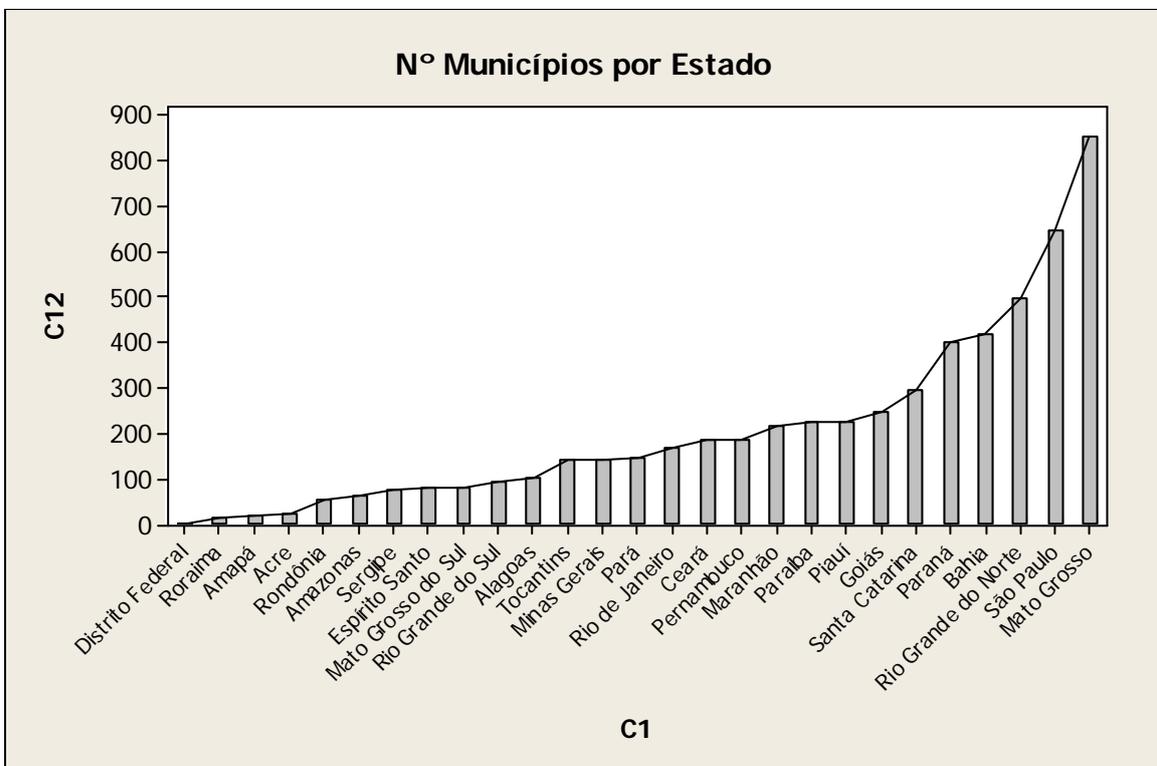
Em São Paulo, concentra-se também a maior população do País.



Interessante que no PIB per capita, São Paulo deixa de liderar e fica atrás de Rio de Janeiro, Minas Gerais e Distrito Federal.



Na variável relativa às dimensões territoriais, destacam-se o Amazonas e o Pará como os maiores Estados da Federação.



Mato Grosso e São Paulo lideram como cidades com o maior número de municípios.

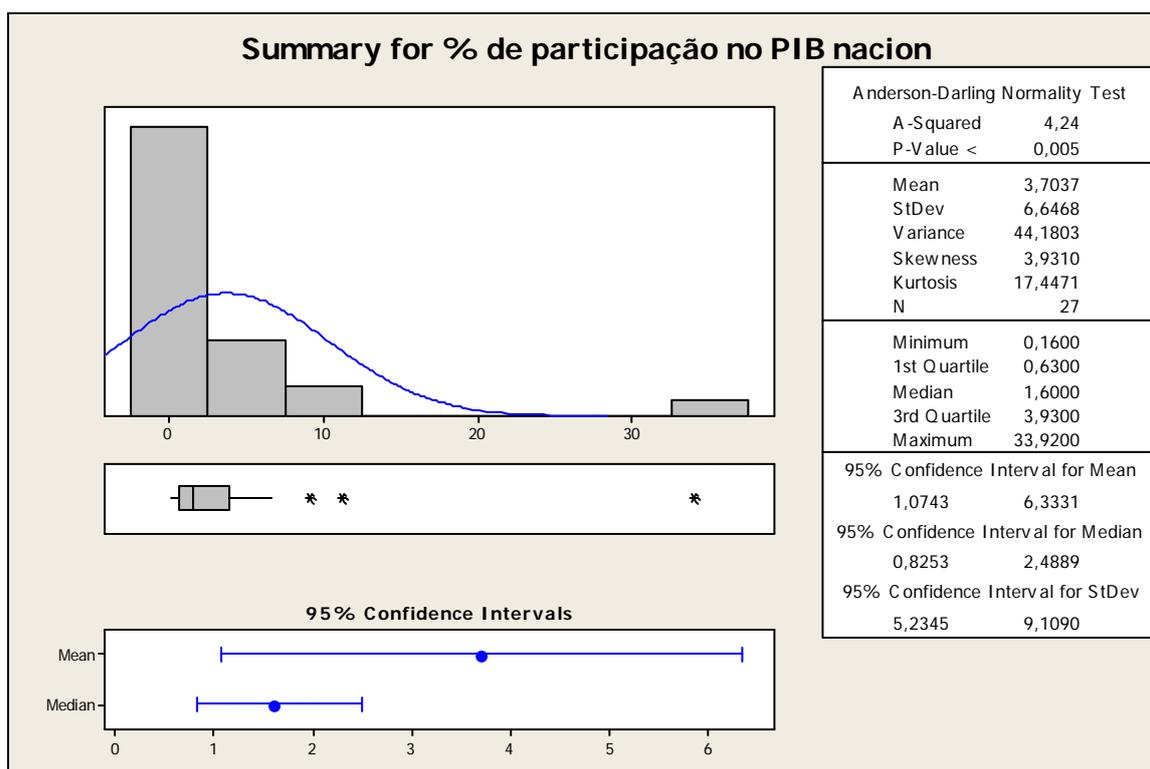
ANÁLISE DAS VARIÁVEIS – Análise Gráfica

Variáveis Quantitativas

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

Com o auxílio do MINITAB os dados foram tabulados e produzidos gráficos que facilitam o entendimento das tendências e concentrações das respostas de cada uma das variáveis.

A seguir, teremos os gráficos com as médias, medianas, histogramas e intervalos de confiança de todas as variáveis.



Pelo gráfico acima, é possível verificar que é uma curva assimétrica para a esquerda, com um ponto marcante de assimetria, demonstrando que não se trata de uma curva normal. Ainda o desenvolvimento da Curva não acompanha os histogramas, mostrando falta de similaridade com o que seria uma curva normal.

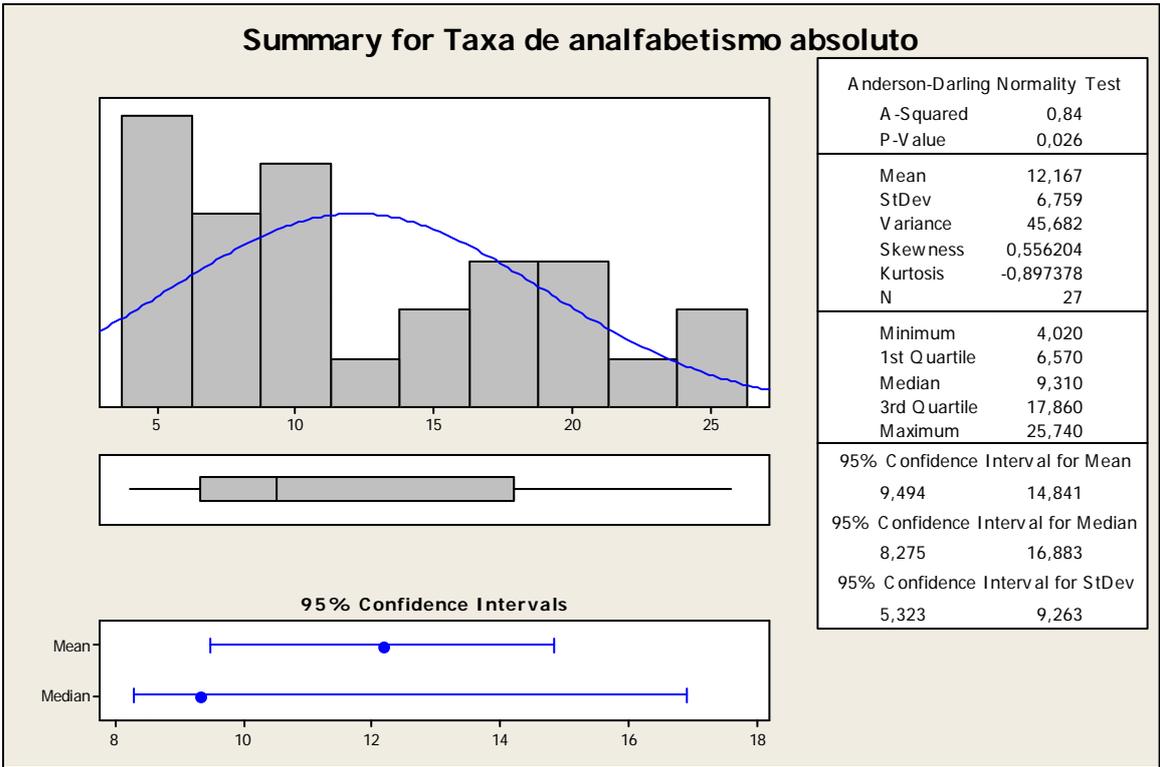
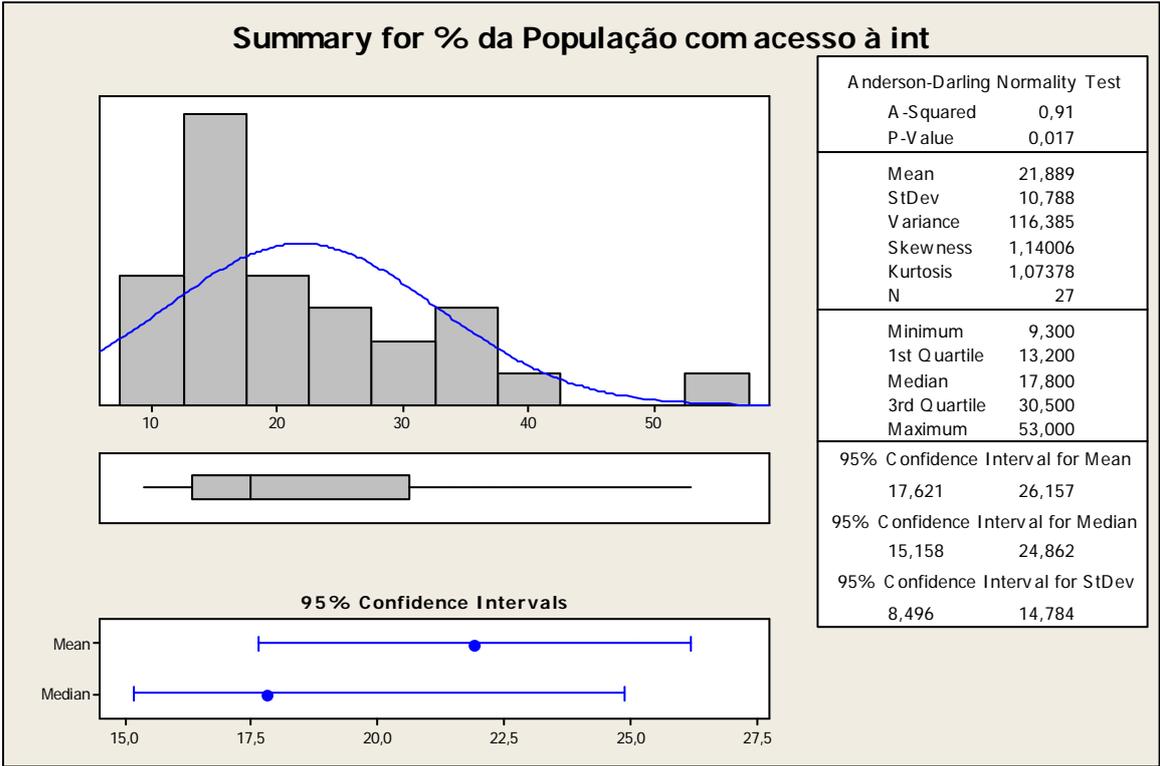


Gráfico 1: curva assimétrica para a esquerda, ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

Gráfico 2: curva assimétrica central com leve similaridade a uma curva normal. Mesmo assim os histogramas não acompanham a curva.

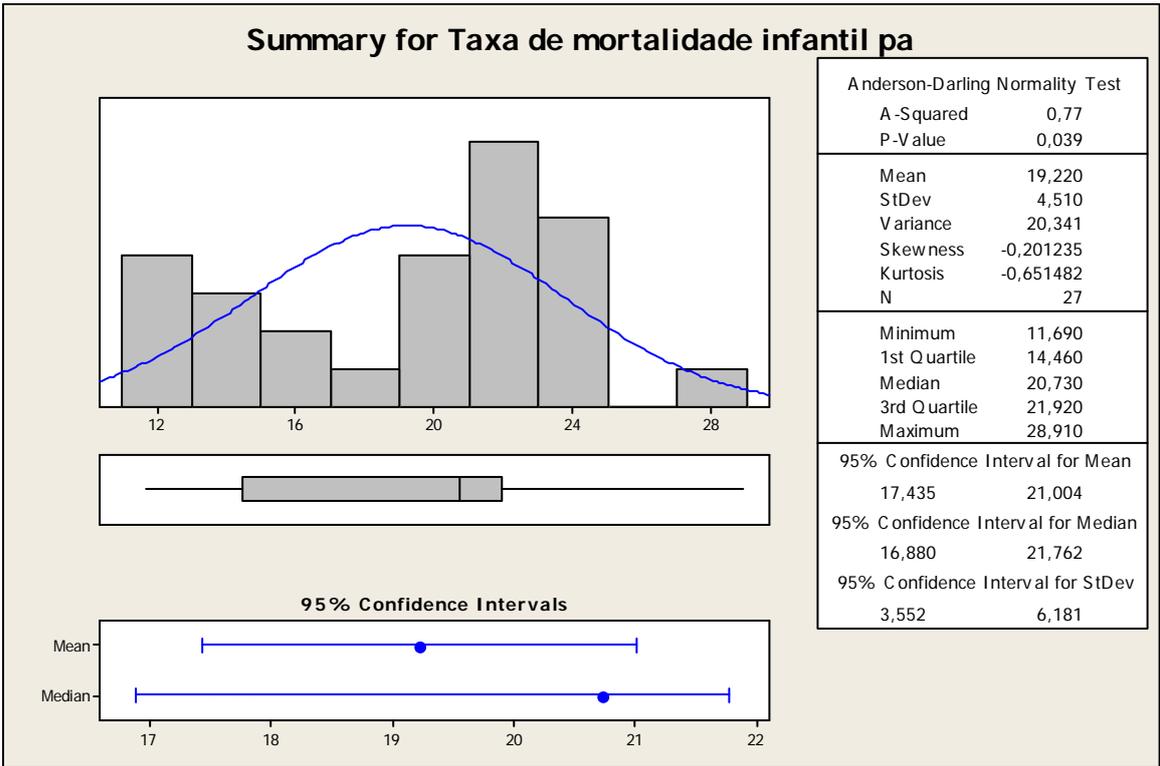
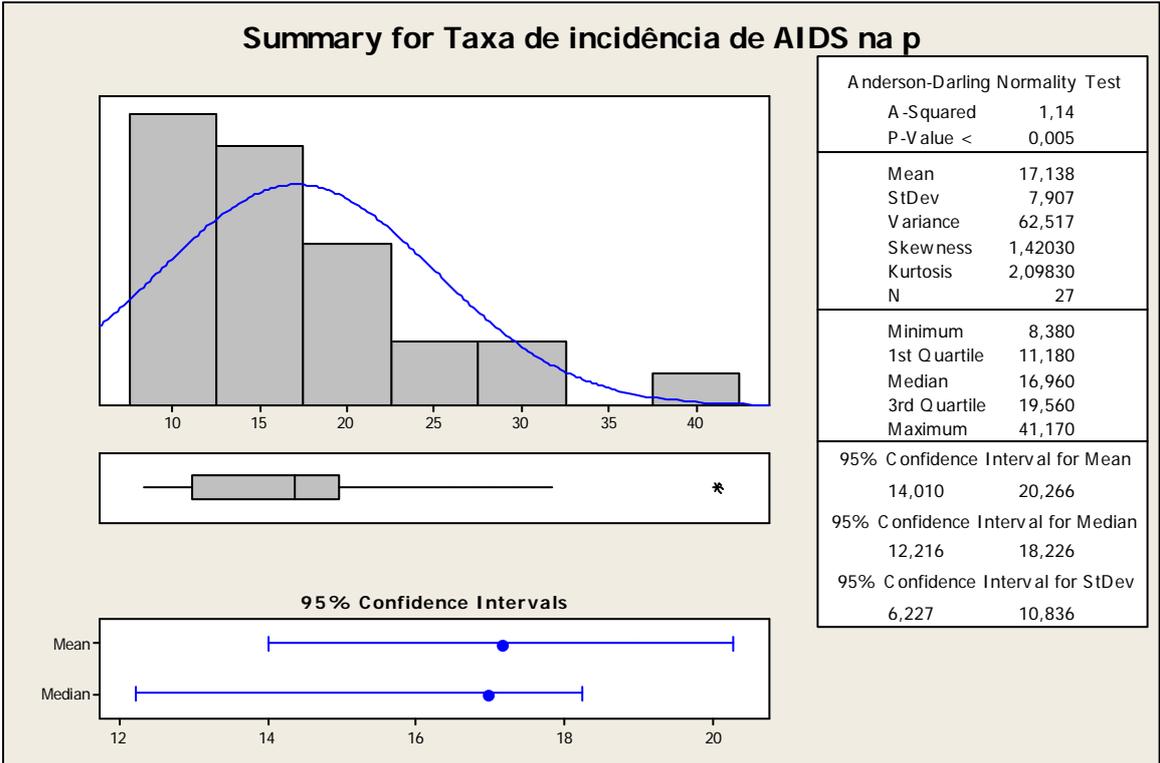


Gráfico 1: curva assimétrica para a esquerda, ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

Gráfico 2: curva assimétrica central com leve similaridade a uma curva normal. Mesmo assim os histogramas não acompanham a curva.

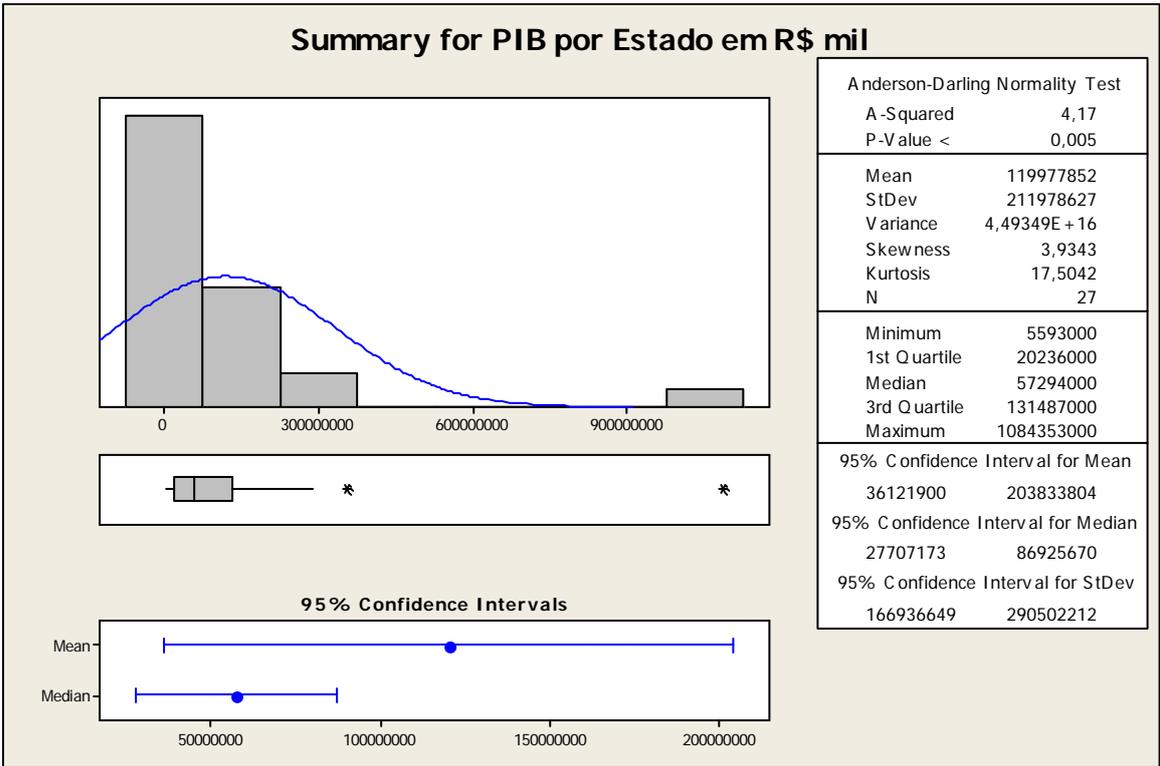
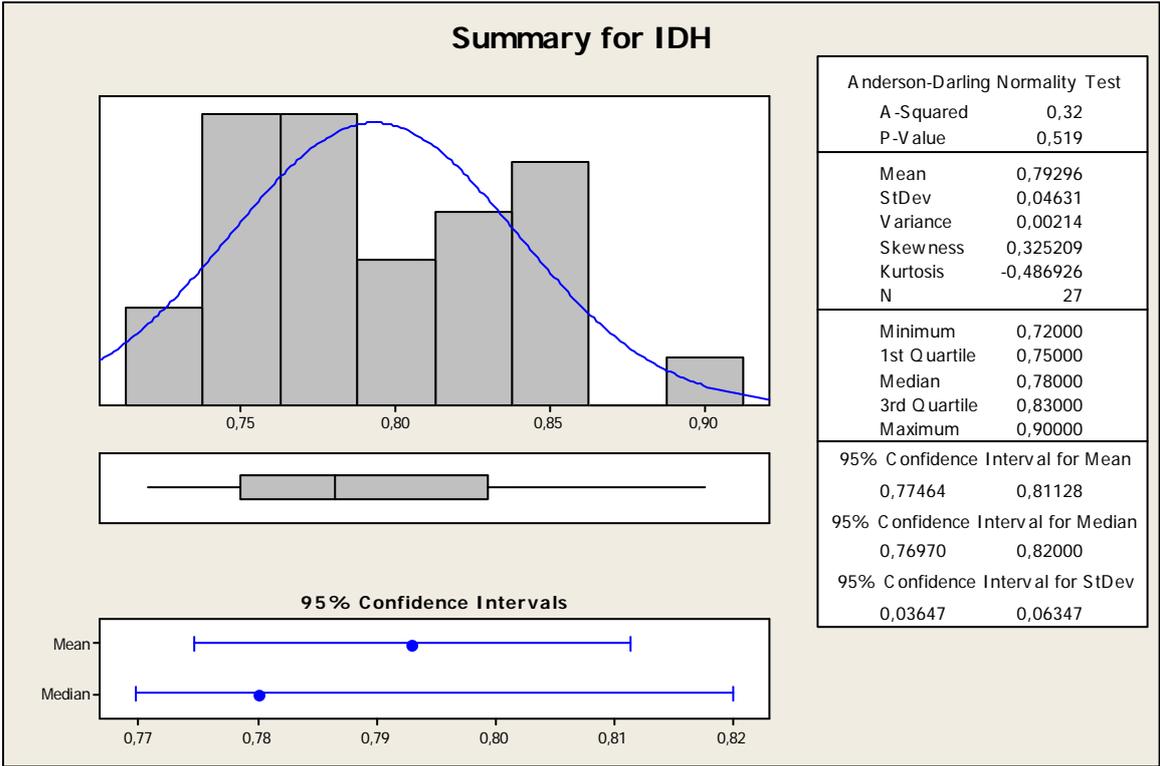


Gráfico 1: curva assimétrica central com leve similaridade a uma curva normal. Mesmo assim os histogramas não acompanham a curva.

Gráfico 2: curva assimétrica para a esquerda, ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

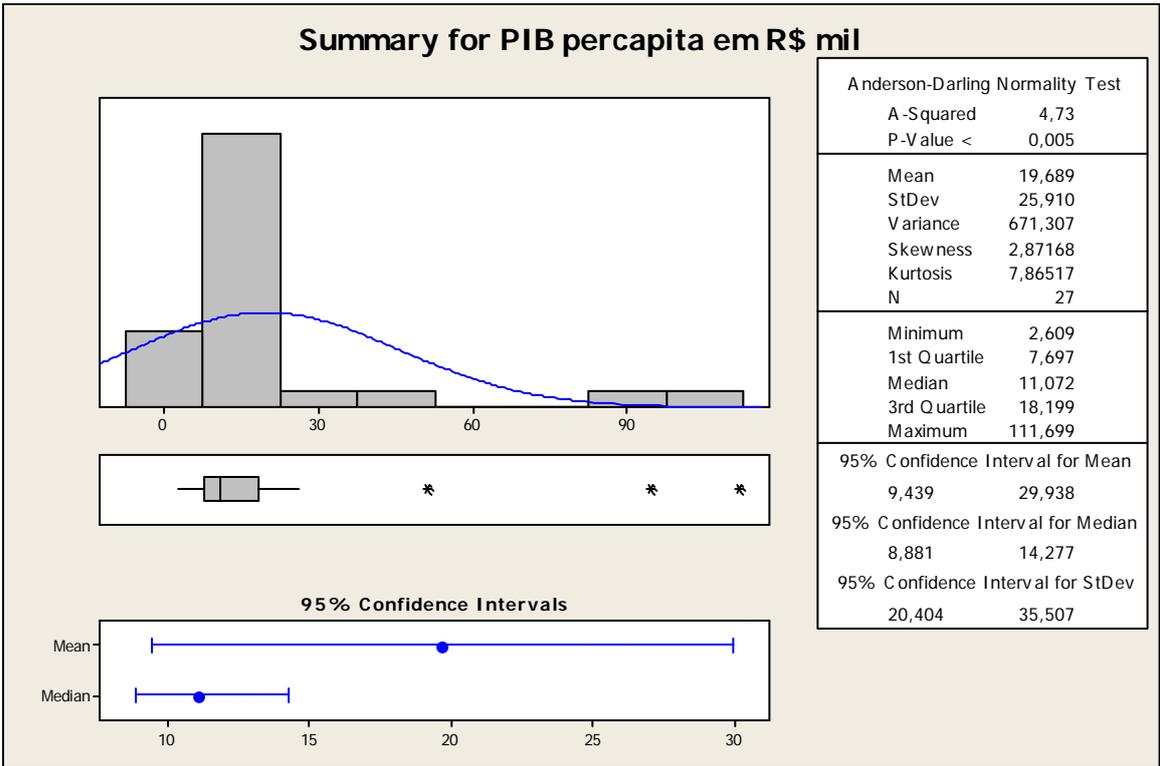
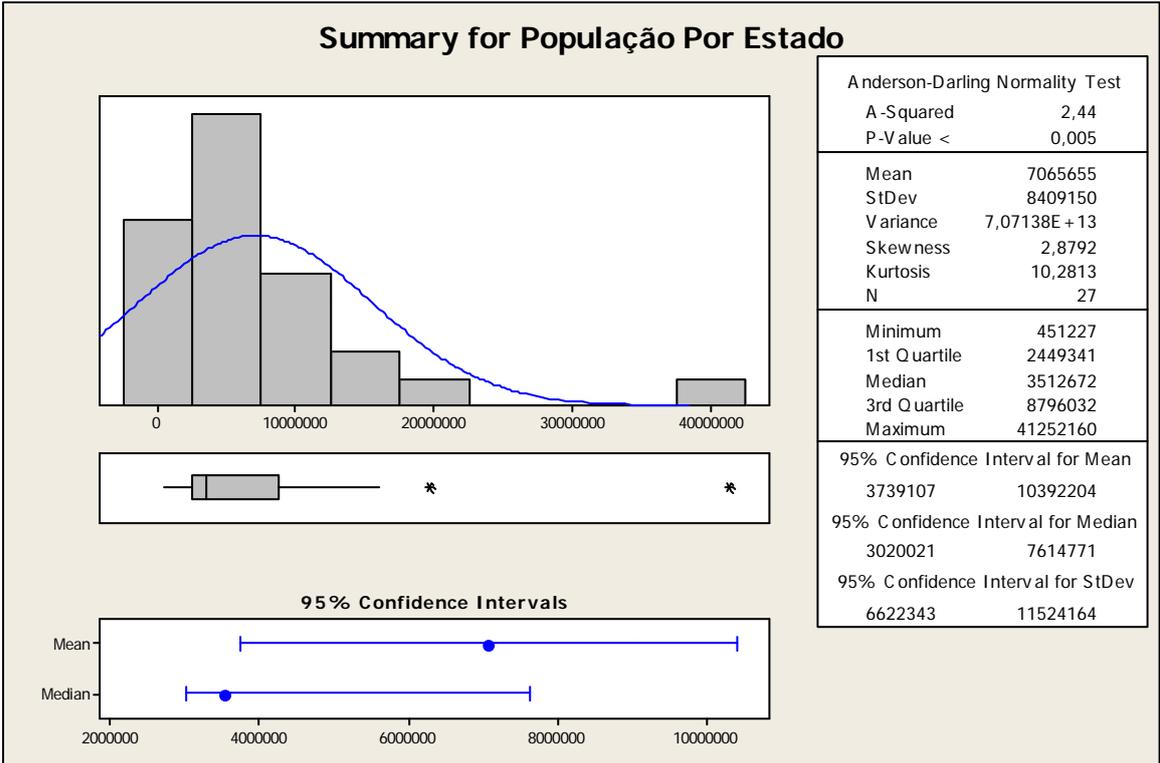


Gráfico 1: curva assimétrica para a esquerda, ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

Gráfico 2: curva assimétrica central com ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

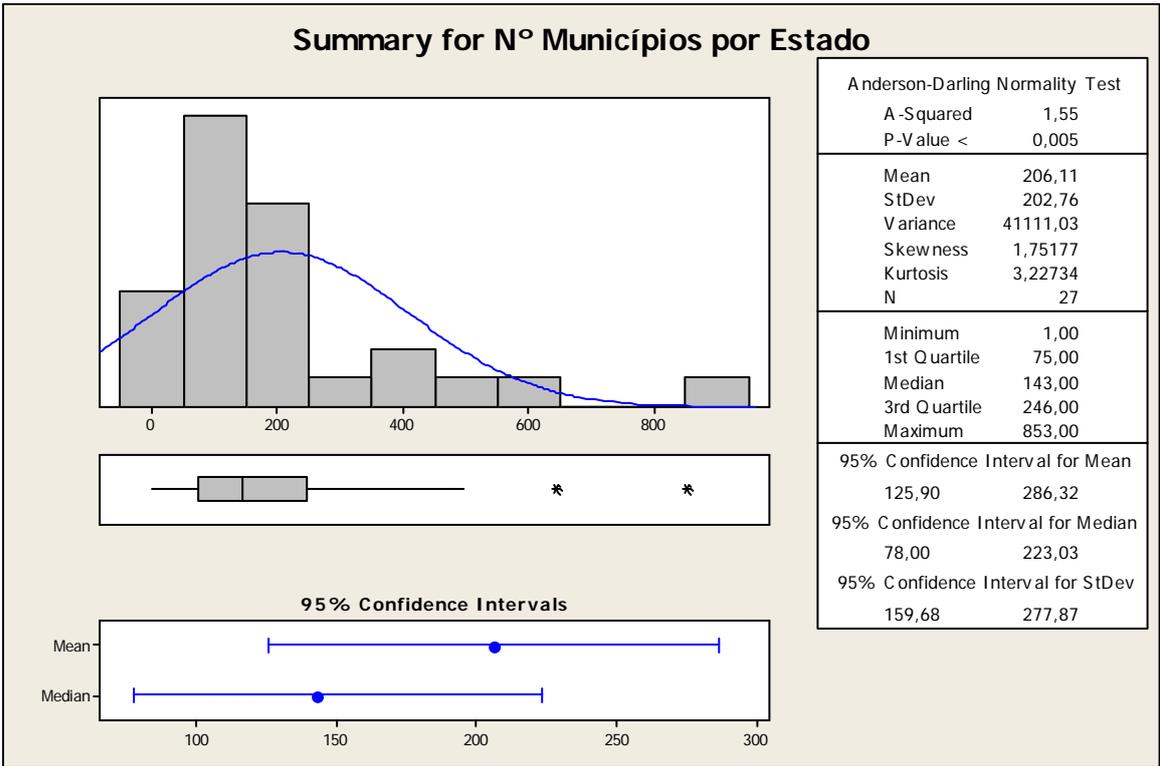
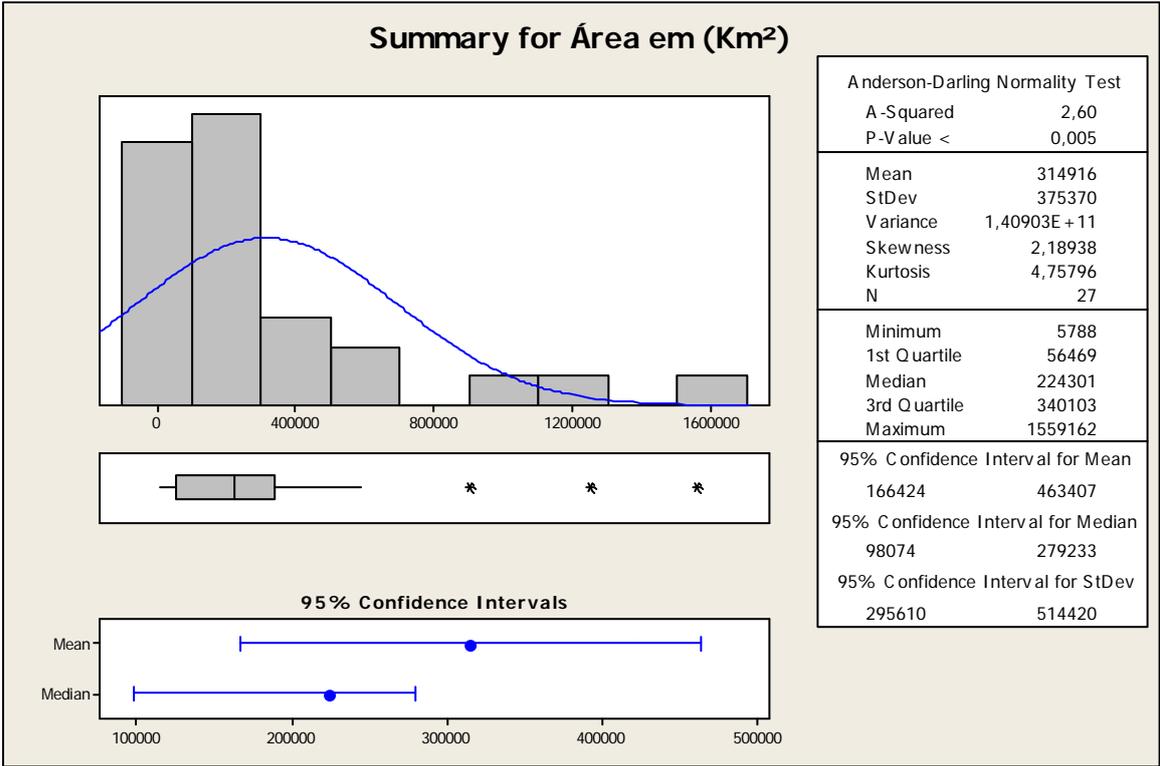


Gráfico 1: curva assimétrica para a esquerda, ausência de similaridade com uma curva normal. Histogramas não acompanham a curva.

Gráfico 2: curva assimétrica central com leve similaridade a uma curva normal. Mesmo assim os histogramas não acompanham a curva.

Médias

Apresentamos abaixo as médias de cada uma das variáveis.

Em seguida, apresentamos uma tabela com os códigos utilizados para identificar cada uma delas nas próximas análises.

Variáveis	Médias das variáveis nos Estados
% de participação no PIB nacional	3,70
% da População com acesso à internet	21,89
Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	12,17
Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	17,14
Taxa de mortalidade infantil para cada 1.000 nascidos vivos	19,22
IDH	0,79
PIB por Estado em R\$ mil	119.977.851,85
População Por Estado	7.065.655,33
PIB percapita em R\$ mil	19,69
Área em (Km ²)	314.915,86
Nº Municípios por Estado	206,11

Variáveis	Nomenclatura
% de participação no PIB nacional	C2
% da População com acesso à internet	C3
Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	C4
Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	C5
Taxa de mortalidade infantil para cada 1.000 nascidos vivos	C6
IDH	C7
PIB por Estado em R\$ mil	C8
População Por Estado	C9
PIB percapita em R\$ mil	C10
Área em (Km ²)	C11
Nº Municípios por Estado	C12

Médias, medianas e desvio padrão.

A seguir, apresentamos as médias, desvios padrão e medianas de cada uma das variáveis estudadas.

Descriptive Statistics: C2; C3; C4; C5; C6; C7; C8; C9; C10; C11; C12

Variable	N	N*	Percent	CumPct	Mean	SE Mean	TrMean	StDev
C2	27	0	100	100	3,70	1,28	2,64	6,65
C3	27	0	100	100	21,89	2,08	21,15	10,79
C4	27	0	100	100	12,17	1,30	11,95	6,76
C5	27	0	100	100	17,14	1,52	16,53	7,91
C6	27	0	100	100	19,220	0,868	19,133	4,510
C7	27	0	100	100	0,79296	0,00891	0,79160	0,04631
C8	27	0	100	100	119977852	40795306	85978240	211978627
C9	27	0	100	100	7065655	1618342	5962772	8409150
C10	27	0	100	100	19,69	4,99	16,69	25,91
C11	27	0	100	100	314916	72240	277511	375370
C12	27	0	100	100	206,1	39,0	188,4	202,8

Variable	Variance	CoefVar	Minimum	Q1	Median	Q3
C2	44,18	179,46	0,16	0,63	1,60	3,93
C3	116,38	49,29	9,30	13,20	17,80	30,50
C4	45,68	55,55	4,02	6,57	9,31	17,86
C5	62,52	46,14	8,38	11,18	16,96	19,56
C6	20,341	23,47	11,690	14,460	20,730	21,920
C7	0,00214	5,84	0,72000	0,75000	0,78000	0,83000
C8	4,49349E+16	176,68	5593000	20236000	57294000	131487000
C9	7,07138E+13	119,01	451227	2449341	3512672	8796032
C10	671,31	131,60	2,61	7,70	11,07	18,20
C11	1,40903E+11	119,20	5788	56469	224301	340103
C12	41111,0	98,37	1,0	75,0	143,0	246,0

Variable	Maximum
C2	33,92
C3	53,00
C4	25,74
C5	41,17
C6	28,910
C7	0,90000
C8	1084353000
C9	41252160
C10	111,70
C11	1559162
C12	853,0

CORRELAÇÃO ENTRE AS VARIÁVEIS

Em teoria da probabilidade e estatística, correlação, também chamada de coeficiente de correlação, indica a força e a direção do relacionamento linear entre duas variáveis aleatórias. No uso estatístico geral, correlação ou co-relação se refere a medida da relação entre duas variáveis, embora correlação não implique causalidade. Neste sentido geral, existem vários coeficientes medindo o grau de correlação, adaptados à natureza dos dados.

A seguir, apresentamos os índices de correlação das variáveis estudadas.

Correlations: C2; C3; C4; C5; C6; C7; C8; C9; C10; C11; C12

	C2	C3	C4	C5	C6	C7	C8	C9	C10
C3	0,531 0,004								
C4	-0,389 0,045	-0,690 0,000							
C5	0,212 0,289	0,407 0,035	-0,658 0,000						
C6	-0,513 0,006	-0,851 0,000	0,522 0,005	-0,482 0,011					
C7	0,450 0,018	0,900 0,000	-0,889 0,000	0,542 0,004	-0,750 0,000				
C8	1,000 0,000	0,535 0,004	-0,390 0,045	0,212 0,289	-0,516 0,006	0,453 0,018			
C9	0,796 0,000	0,290 0,142	-0,171 0,394	0,137 0,494	-0,320 0,104	0,194 0,333	0,800 0,000		
C10	0,393 0,043	0,547 0,003	-0,441 0,021	0,208 0,299	-0,442 0,021	0,528 0,005	0,390 0,044	-0,091 0,651	
C11	-0,020 0,920	-0,298 0,131	-0,101 0,617	0,024 0,904	0,291 0,141	-0,065 0,747	-0,023 0,908	0,039 0,845	0,025 0,903
C12	0,425 0,027	0,107 0,595	0,039 0,849	-0,138 0,492	-0,110 0,583	0,019 0,924	0,428 0,026	0,762 0,000	-0,135 0,503
C12		C11							
	0,084 0,677								

Regression Analysis: C2 versus C3; C4; C5; C6; C7

The regression equation is

$$C2 = 132 + 0,361 C3 - 0,743 C4 - 0,182 C5 - 0,670 C6 - 140 C7$$

Predictor	Coef	SE Coef	T	P
Constant	131,66	96,22	1,37	0,186
C3	0,3614	0,3450	1,05	0,307
C4	-0,7434	0,5499	-1,35	0,191
C5	-0,1820	0,2274	-0,80	0,433
C6	-0,6699	0,5821	-1,15	0,263
C7	-139,8	113,1	-1,24	0,230

S = 5,94064 R-Sq = 35,5% R-Sq(adj) = 20,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	407,57	81,51	2,31	0,081
Residual Error	21	741,11	35,29		
Total	26	1148,69			

Source	DF	Seq SS
C3	1	323,56
C4	1	1,18
C5	1	0,91
C6	1	28,07
C7	1	53,86

Unusual Observations

Obs	C3	C2	Fit	SE Fit	Residual	St Resid
25	39,5	33,92	11,62	2,69	22,30	4,21R

R denotes an observation with a large standardized residual.

General Regression Analysis: C2 versus C3; C4; C5; C6; C7; C8; C9; C10; ...

Regression Equation

$$C2 = 2,19594 - 0,00367373 C3 - 0,0115408 C4 + 0,000792129 C5 - 0,0105217 C6 - 2,29232 C7 + 3,2047e-008 C8 - 2,16931e-008 C9 - 0,00108381 C10 + 3,35177e-008 C11 + 0,000262494 C12$$

Coefficients

Term	Coef	SE Coef	T	P
Constant	2,19594	2,03338	1,0799	0,296
C3	-0,00367	0,00830	-0,4428	0,664
C4	-0,01154	0,01102	-1,0477	0,310
C5	0,00079	0,00485	0,1634	0,872
C6	-0,01052	0,01150	-0,9152	0,374
C7	-2,29232	2,44356	-0,9381	0,362
C8	0,00000	0,00000	80,9336	0,000
C9	-0,00000	0,00000	-1,6874	0,111
C10	-0,00108	0,00188	-0,5775	0,572
C11	0,00000	0,00000	0,4165	0,683
C12	0,00026	0,00025	1,0358	0,316

Summary of Model

S = 0,113744 R-Sq = 99,98% R-Sq(adj) = 99,97%
PRESS = 1,05323 R-Sq(pred) = 99,91%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	10	1148,48	1148,48	114,848	8876,98	0,000000
C3	1	323,56	0,00	0,003	0,20	0,663821
C4	1	1,18	0,01	0,014	1,10	0,310335
C5	1	0,91	0,00	0,000	0,03	0,872228
C6	1	28,07	0,01	0,011	0,84	0,373699
C7	1	53,86	0,01	0,011	0,88	0,362136
C8	1	740,83	84,75	84,745	6550,24	0,000000
C9	1	0,06	0,04	0,037	2,85	0,110916
C10	1	0,00	0,00	0,004	0,33	0,571647
C11	1	0,00	0,00	0,002	0,17	0,682602
C12	1	0,01	0,01	0,014	1,07	0,315706
Error	16	0,21	0,21	0,013		
Total	26	1148,69				

Fits and Diagnostics for Unusual Observations

Obs	C2	Fit	SE Fit	Residual	St Resid	
7	3,76	3,88410	0,0975927	-0,124098	-2,12408	R
8	2,27	2,02927	0,0510706	0,240727	2,36856	R
9	2,45	2,63539	0,0666698	-0,185390	-2,01168	R

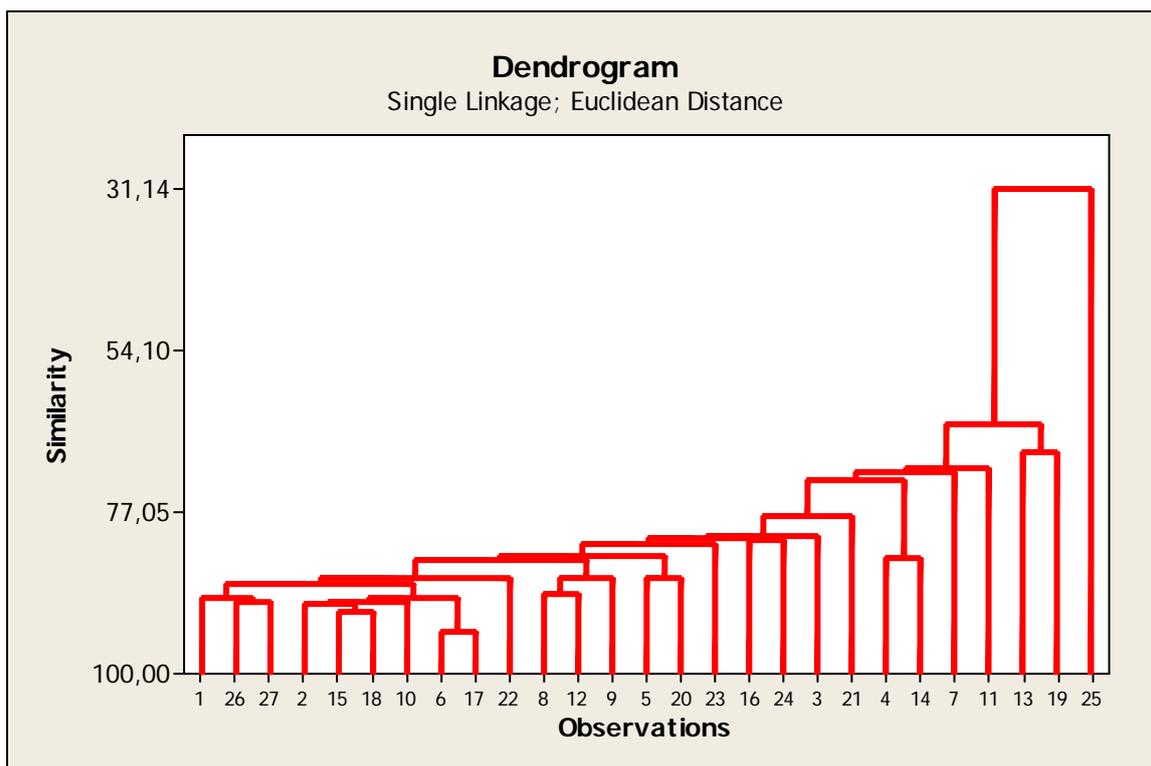
R denotes an observation with a large standardized residual.

DENDOGRAMAS

Um Dendograma (dendo = árvore) é um tipo específico de diagrama ou representação icónica que organiza determinados fatores e variáveis. Resulta de uma análise estatística de determinados dados, em que se emprega um método quantitativo que leva a agrupamentos e à sua ordenação hierárquica ascendente - o que em termos gráficos se assemelha aos ramos de uma árvore que se vão dividindo noutros sucessivamente. Isto é, ilustra o arranjo de agrupamentos derivado da aplicação de um "algoritmo de clustering" (ver agrupamento de dados).

Para facilitar a identificação, segue a seguir uma lista com os números correspondentes de cada Estado.

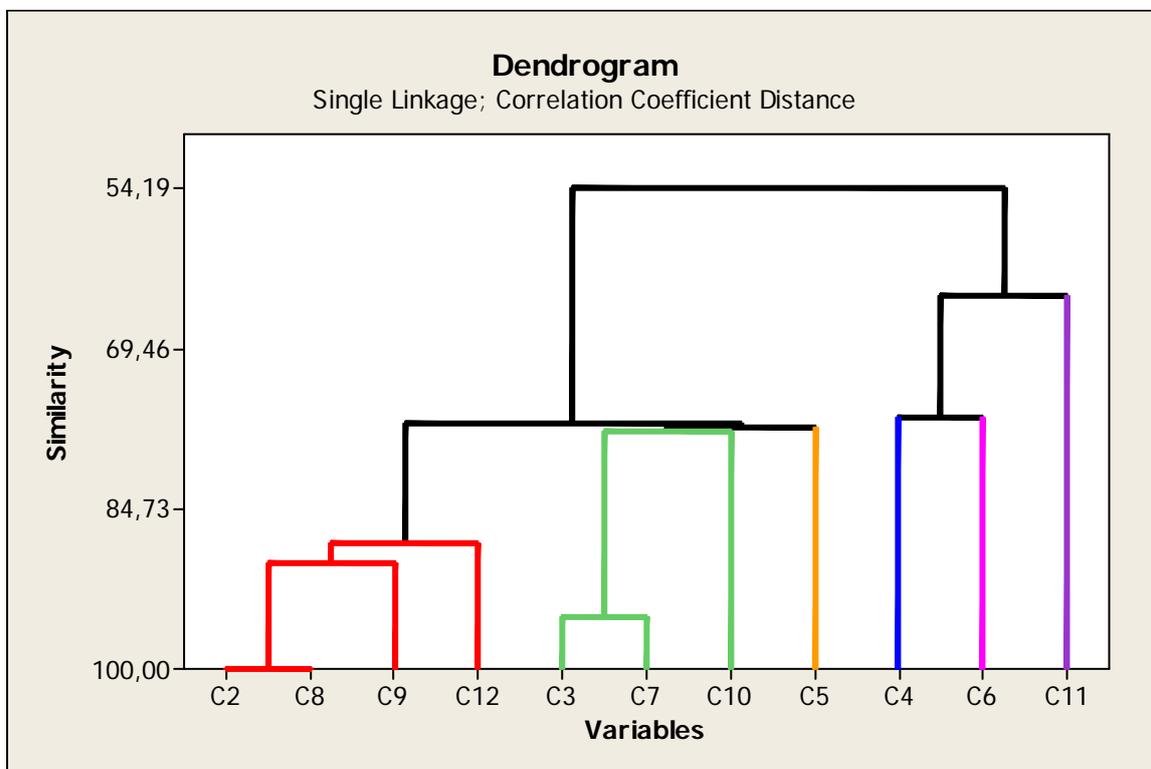
Estados	Número de referencia do Estado
Acre	1
Alagoas	2
Amapá	3
Amazonas	4
Bahia	5
Ceará	6
Distrito Federal	7
Espírito Santo	8
Goiás	9
Maranhão	10
Mato Grosso	11
Mato Grosso do Sul	12
Minas Gerais	13
Pará	14
Paraíba	15
Paraná	16
Pernambuco	17
Piauí	18
Rio de Janeiro	19
Rio Grande do Norte	20
Rio Grande do Sul	21
Rondônia	22
Roraima	23
Santa Catarina	24
São Paulo	25
Sergipe	26
Tocantins	27



Neste dendrograma podemos analisar a correlação entre os Estados de acordo com o comportamento de todas as variáveis. Destacam-se os Estados 6 e 17, que são Ceará e Pernambuco respectivamente.

Variáveis	Nomenclatura
% de participação no PIB nacional	C2
% da População com acesso à internet	C3
Taxa de analfabetismo absoluto de pessoas com 15 anos ou mais de idade	C4
Taxa de incidência de AIDS na população - casos em cada 100.000 habitantes	C5
Taxa de mortalidade infantil para cada 1.000 nascidos vivos	C6
IDH	C7
PIB por Estado em R\$ mil	C8
População Por Estado	C9
PIB percapita em R\$ mil	C10
Área em (Km ²)	C11
Nº Municípios por Estado	C12

Um Dendrograma (dend(o) = árvore) é um tipo específico de diagrama ou representação icônica que organiza determinados fatores e variáveis. É um diagrama de similaridade. Segue o Dendrograma para as cinco variáveis selecionadas:



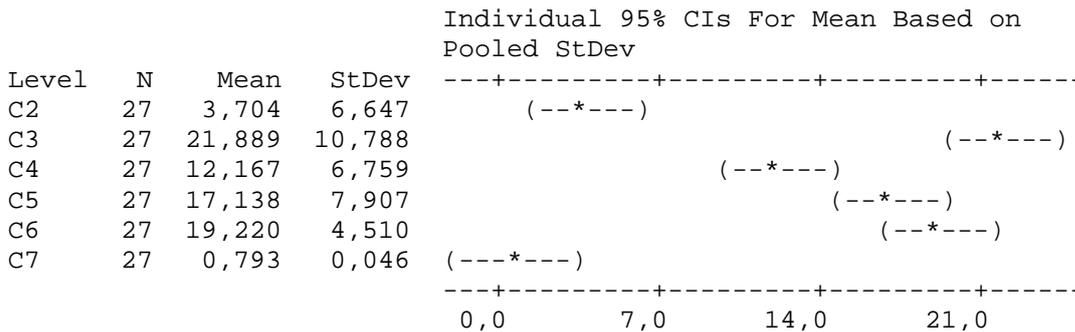
Neste dendrograma, podemos observar a similaridade entre as variáveis C2 e C8, ambas relacionadas ao PIB, C3 e C7 que são o % de pessoas com acesso a internet e o IDH.

Análise ANOVA, Análise Fatorial e Principais Componentes.

One-way ANOVA: C2; C3; C4; C5; C6; C7

Source	DF	SS	MS	F	P
Factor	5	9972,5	1994,5	41,39	0,000
Error	156	7516,8	48,2		
Total	161	17489,3			

S = 6,942 R-Sq = 57,02% R-Sq(adj) = 55,64%



Pooled StDev = 6,942

Nota-se que não há convergência entre as variáveis pois os as médias estão bem distantes e os intervalos pequenos e distantes.

Factor Analysis: C2; C3; C4; C5; C6; C7

Principal Component Factor Analysis of the Correlation Matrix

Unrotated Factor Loadings and Communalities

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Communality
C2	0,606	-0,599	0,516	0,083	0,006	1,000
C3	0,917	-0,199	-0,278	-0,012	0,195	0,996
C4	-0,860	-0,318	-0,057	-0,371	0,105	0,993
C5	0,661	0,585	0,359	-0,292	0,074	1,000
C6	-0,853	0,212	0,185	0,410	0,156	1,000
C7	0,949	0,069	-0,197	0,195	-0,008	0,982
Variance	4,0145	0,8923	0,5491	0,4361	0,0793	5,9713
% Var	0,669	0,149	0,092	0,073	0,013	0,995

Rotated Factor Loadings and Communalities

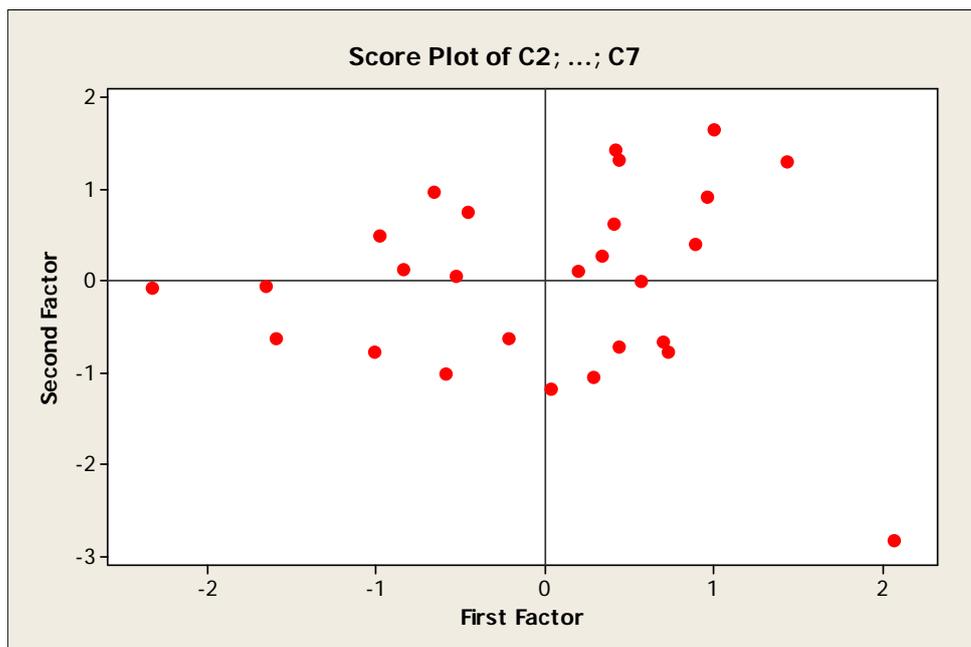
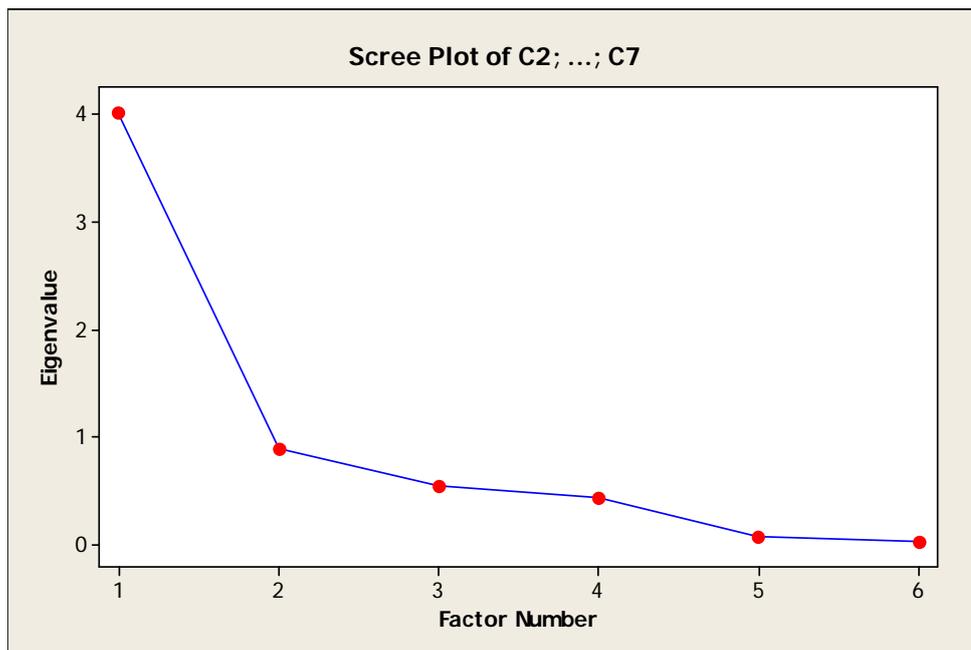
Equamax Rotation

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Communality
C2	0,128	0,189	0,065	0,957	0,164	1,000
C3	0,336	0,496	0,135	0,277	0,737	0,996
C4	-0,822	-0,149	-0,405	-0,188	-0,309	0,993
C5	0,243	0,181	0,942	0,068	0,123	1,000
C6	-0,161	-0,841	-0,241	-0,269	-0,369	1,000
C7	0,653	0,397	0,243	0,193	0,550	0,982

Variance	1,3164	1,2009	1,1921	1,1432	1,1187	5,9713
% Var	0,219	0,200	0,199	0,191	0,186	0,995

Factor Score Coefficients

Variable	Factor1	Factor2	Factor3	Factor4	Factor5
C2	-0,118	-0,263	0,016	1,135	-0,168
C3	-0,942	-0,850	0,158	-0,067	2,187
C4	-1,342	-0,136	0,212	0,004	0,898
C5	-0,629	-0,383	1,243	0,009	0,353
C6	-0,231	-1,801	0,220	0,146	1,274
C7	0,527	0,015	-0,227	-0,149	0,221

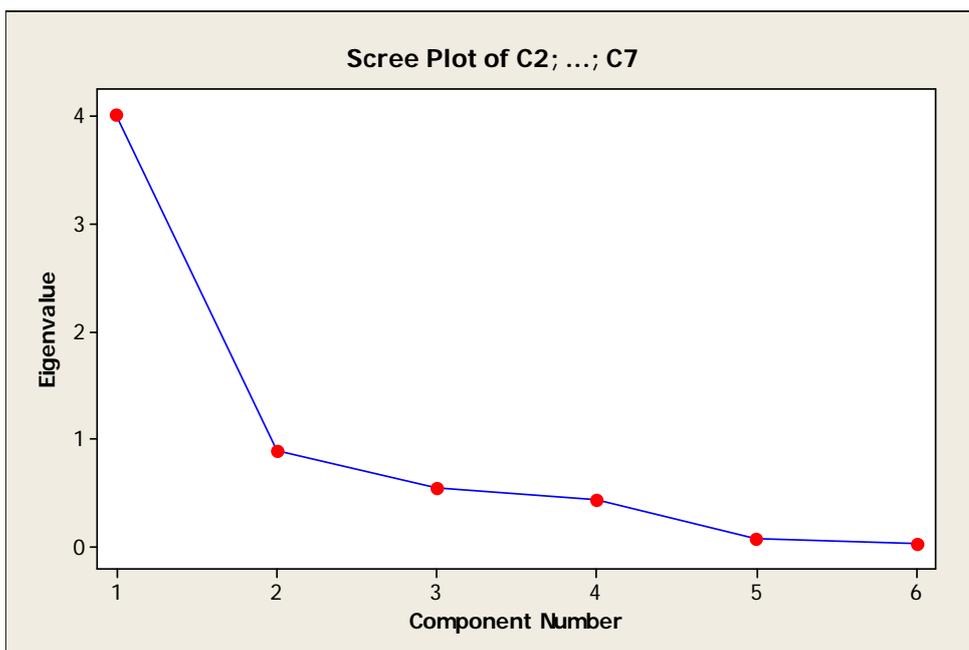
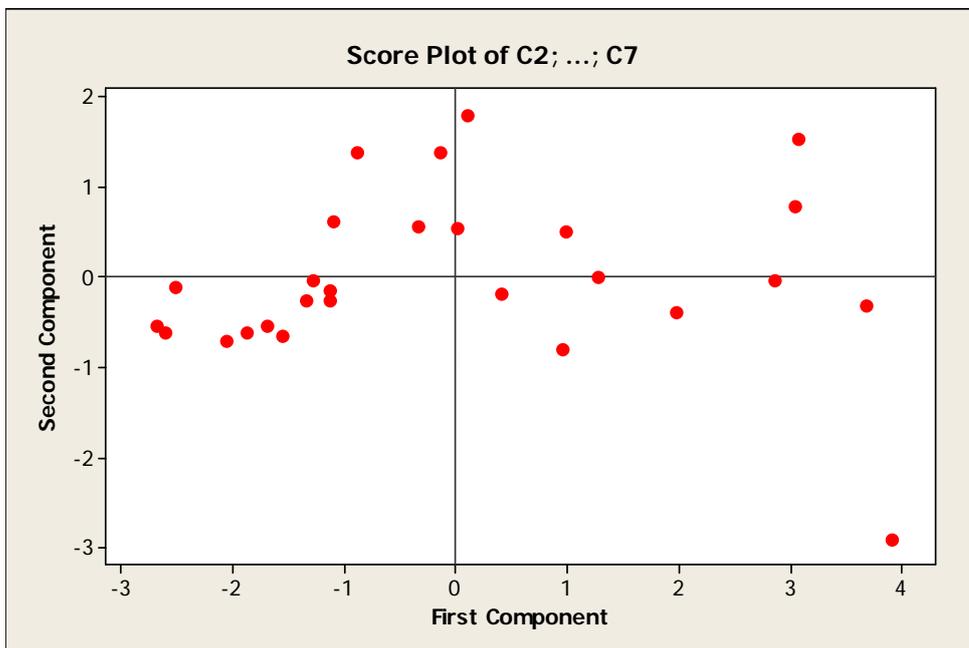


Principal Component Analysis: C2; C3; C4; C5; C6; C7

Eigenanalysis of the Correlation Matrix

Eigenvalue	4,0145	0,8923	0,5491	0,4361	0,0793	0,0287
Proportion	0,669	0,149	0,092	0,073	0,013	0,005
Cumulative	0,669	0,818	0,909	0,982	0,995	1,000

Variable	PC1	PC2	PC3	PC4	PC5
C2	0,302	-0,634	0,697	0,126	0,022
C3	0,458	-0,210	-0,376	-0,018	0,693
C4	-0,429	-0,337	-0,077	-0,561	0,374
C5	0,330	0,620	0,485	-0,442	0,264
C6	-0,426	0,224	0,249	0,621	0,555
C7	0,474	0,073	-0,265	0,295	-0,030



CONSIDERAÇÕES FINAIS

A ferramenta Minitab 16 foi fundamental para conseguirmos realizar as análises e gráficos deste trabalho. Com o uso da ferramenta, pudemos nos aprofundar no estudo de algumas das variáveis que explicam a situação econômico-social de nosso país.

Pudemos ver claramente quais são os Estados que se destacam em cada uma das variáveis estudadas e este material poderá servir como base para estudos mais profundos, buscando fundamentar políticas públicas que melhorem estes índices como : taxa de analfabetismo, taxa de mortalidade infantil, incidência de HIV e outros mais.

O estudo aprofundado dos índices de cada Estado, permitirá que os tomadores de decisão possam ter dados estatísticos concretos para que, com base nestes, possam desenvolver a sua gestão e melhorar as políticas públicas de desenvolvimento social no Brasil.

Todas as bases de dados foram obtidas através do site do IBGE, instituto Brasileiro de Geografia e Estatística, com acesso de 18 a 27 de junho. Todos os dados e informações se referem ao ano de 2009, segundo o site do IBGE.

Agradecimento:

“ Deixo aqui um agradecimento aos colegas de turma, pelo companheirismo durante as aulas e um especial ao Professor Doutor Arnoldo Jose de Hoyos, pelos ensinamentos, além do entusiasmo e dedicação com que conduziu a disciplina.”

São Paulo, junho de 2012.

Mestrado em administração de Empresas – PUC-SP

Trabalho Final da disciplina Métodos Quantitativos e Qualitativos da Pesquisa Empírica

Professor Doutor: Arnoldo Jose de Hoyos

Mestrando: Daniel Troise