



PONTÍFICIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

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

UMA ANÁLISE DE ÍNDICES DE DESENVOLVIMENTO BÁSICO DA REGIÃO DE ARARAS

**Aluna: Erika Camila Buzo Martins
Prof. Arnaldo José de Hoyos Guevara**

1° Semestre 2012

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise das principais cidades da região de Araras que “encaminham” alunos para o Centro Universitário Hermínio Ometto – Uniararas e suas relações com tamanho da população, índice de desenvolvimento humano e PIB per capita entre outros.

O software estatístico utilizado foi o MINITAB. Somente para a Árvore de Classificação foi utilizado o software SPSS.

2. ENTENDENDO OS DADOS

2.1 Os Indivíduos

Os indivíduos desta análise são as 10 cidades mais próximas da região de Araras, responsáveis pelo maior número de alunos matriculados anualmente no Centro Universitário Hermínio Ometto - Uniararas, incluindo a própria cidade de Araras e o Estado de São Paulo como fonte de comparação.

2.2 As Variáveis

São 8 as variáveis desta pesquisa:

Tabela nº 1 – Variáveis Escolhidas

| Variável | Significado | Unidade de medida |
|-----------------------|--|--------------------------|
| Área | Área 2012 | Quantidade |
| População | População 2011 | Quantidade |
| Natalidade | Taxa de Natalidade 2010 | Por mil habitantes |
| IDH | Índice de Desenvolvimento Humano – IDH 2000 | Indicador |
| Renda | Renda per Capita 2000 | Em salários mínimos |
| Analfabetismo | Taxa de Analfabetismo da População de 15 Anos e Mais 2000 | Em % |
| Média Estudos 15 a 64 | Média de Anos de Estudos da População de 15 a 64 Anos 2000 | Quantidade |
| PIB | PIB per Capita 2009 | Em reais correntes |

2.3 A Tabela de Dados

A seguir a tabela com as variáveis e as cidades escolhidas.

Tabela nº 2 – Tabela de Dados

| Variáveis | Área2012 | População2011 | Taxa de Natalidade (Por mil habitantes)2010 | Índice de Desenvolvimento Humano – IDH2000 | Renda per Capita (Em salários mínimos)2000 | Taxa de Analfabetismo da População de 15 Anos e Mais (Em %)2000 | Média de Anos de Estudos da População de 15 a 64 Anos2000 | PIB per Capita (Em reais correntes)2009 |
|------------------------|------------|---------------|---|--|--|---|---|---|
| Estado SP | 248.209,43 | 41.692.668 | 14,59 | 0,814 | 2,92 | 6,64 | 7,64 | 26.202,22 |
| Araras | 643,46 | 120.291 | 12,43 | 0,828 | 2,48 | 7,01 | 7,3 | 21.022,57 |
| Rio Claro | 498,01 | 188.019 | 12,48 | 0,825 | 3,1 | 5,1 | 7,95 | 23.570,31 |
| Leme | 403,08 | 92.841 | 14,49 | 0,796 | 2,1 | 9,86 | 6,29 | 15.044,04 |
| Limeira | 580,98 | 278.661 | 12,67 | 0,814 | 2,44 | 6,25 | 7,45 | 20.428,38 |
| Cordeirópolis | 137,34 | 21.435 | 14,16 | 0,835 | 2,44 | 6,72 | 6,88 | 65.931,43 |
| Mogi Guaçu | 813,14 | 138.520 | 13,48 | 0,813 | 2,34 | 7,12 | 7,2 | 17.997,83 |
| Mogi Mirim | 499,12 | 86.997 | 12,78 | 0,825 | 2,77 | 6,71 | 7,7 | 25.818,27 |
| Pirassununga | 726,94 | 70.584 | 12,14 | 0,839 | 2,66 | 6,05 | 7,85 | 18.276,15 |
| Conchal | 183,83 | 25.479 | 15,87 | 0,77 | 1,65 | 12,24 | 5,49 | 15.070,78 |
| Santa Gertrudes | 97,69 | 22.256 | 14,5 | 0,782 | 1,89 | 7,86 | 6,15 | 17.541,86 |

2.4 Fontes de Dados

Todos os dados desta pesquisa foram obtidos em: <http://www.seade.gov.br/>

2.4.1 Sobre o SEADE

O SEADE, fundação vinculada à Secretaria Estadual de Planejamento e Desenvolvimento Regional do Estado de São Paulo, é hoje um centro de referência nacional na produção e disseminação de análises e estatísticas socioeconômicas e demográficas. Para isso ela realiza pesquisas diretas e levantamentos de informações produzidas por outras fontes, compondo um amplo acervo, disponibilizado gratuitamente, que permite a caracterização de diferentes aspectos da realidade socioeconômica do estado, de suas regiões e municípios e de sua evolução histórica.

Ao longo de mais de 30 anos de atuação, o Seade tem se constituído em uma segura e sempre atualizada fonte de dados sobre o estado de São Paulo. As habilidades que a instituição desenvolve continuamente a capacitam para a criação e aprimoramento de metodologias e ferramentas para a formulação, monitoramento e avaliação de políticas públicas, o que tem levado instituições das mais diferentes áreas de atuação a contratá-la para a prestação de serviços técnicos.

Sua extensa e diversificada linha de produtos e serviços tem auxiliado cidadãos, gestores públicos, empresários e jornalistas à compreender melhor as características específicas da realidade paulista, as mudanças sociais, as transformações econômicas e os impactos das políticas públicas nos seus 645 municípios.

Para disseminar sua produção, a Fundação Seade orienta-se pelas seguintes diretrizes principais:

- Uso da internet como principal instrumento de disseminação;
- Gratuidade de acesso;

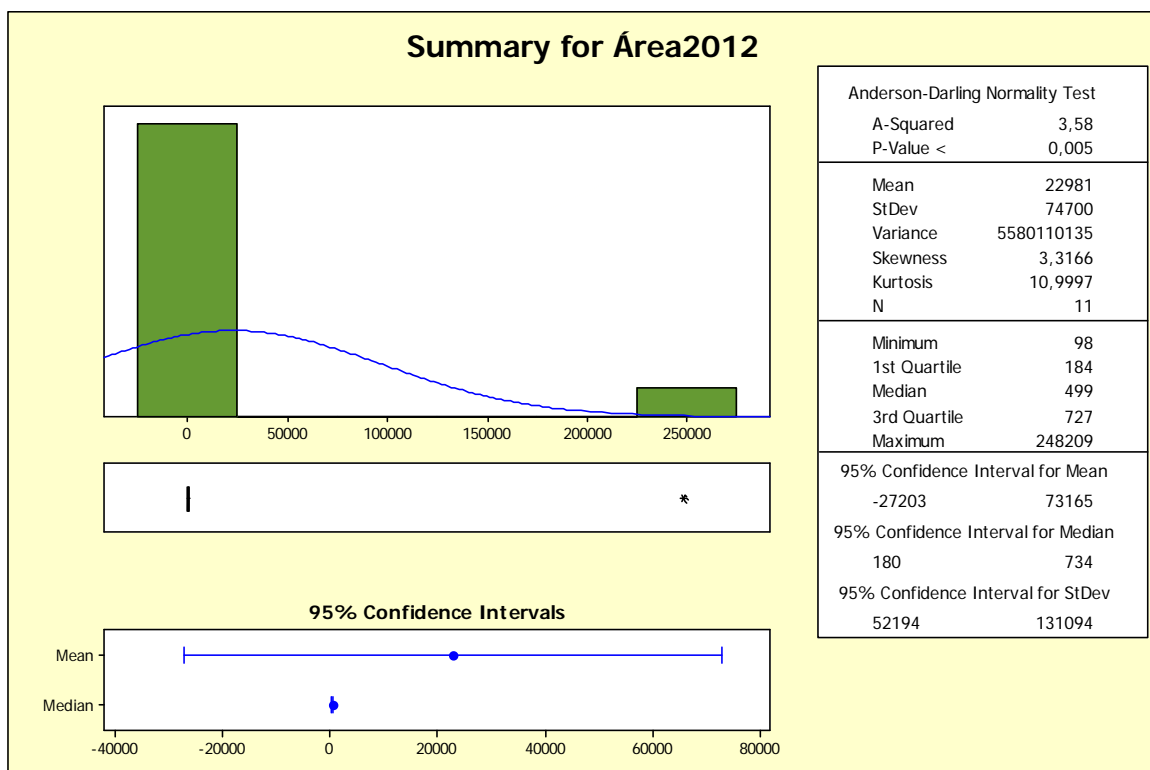
- Disponibilização de todo o acervo recente de informações;
- Disponibilização, sempre que possível, dos microdados resultantes das pesquisas;
- Fornecimento de metadados, como definições, classificações utilizadas, notas que sejam relevantes para a compreensão da informação;
- Pesquisa e desenvolvimento permanente de sistemas e ferramentas para apresentação de informações;
- Transferência de tecnologia e compartilhamento de soluções com instituições congêneres e parceiras.

3 – ANÁLISE EXPLORATÓRIA

Este tópico exibe análise baseada em variáveis quantitativas utilizando Histograma, Box-Plot, Intervalo de Confiança, Média, Mediana, Mínimos e Máximos e Desvio-Padrão.

3.1 Variável Área

Segue o gráfico resumo da variável:



- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição assimétrica, com um pico à esquerda. O P-Value igual a 0,005 indica tratar-se de uma distribuição normal

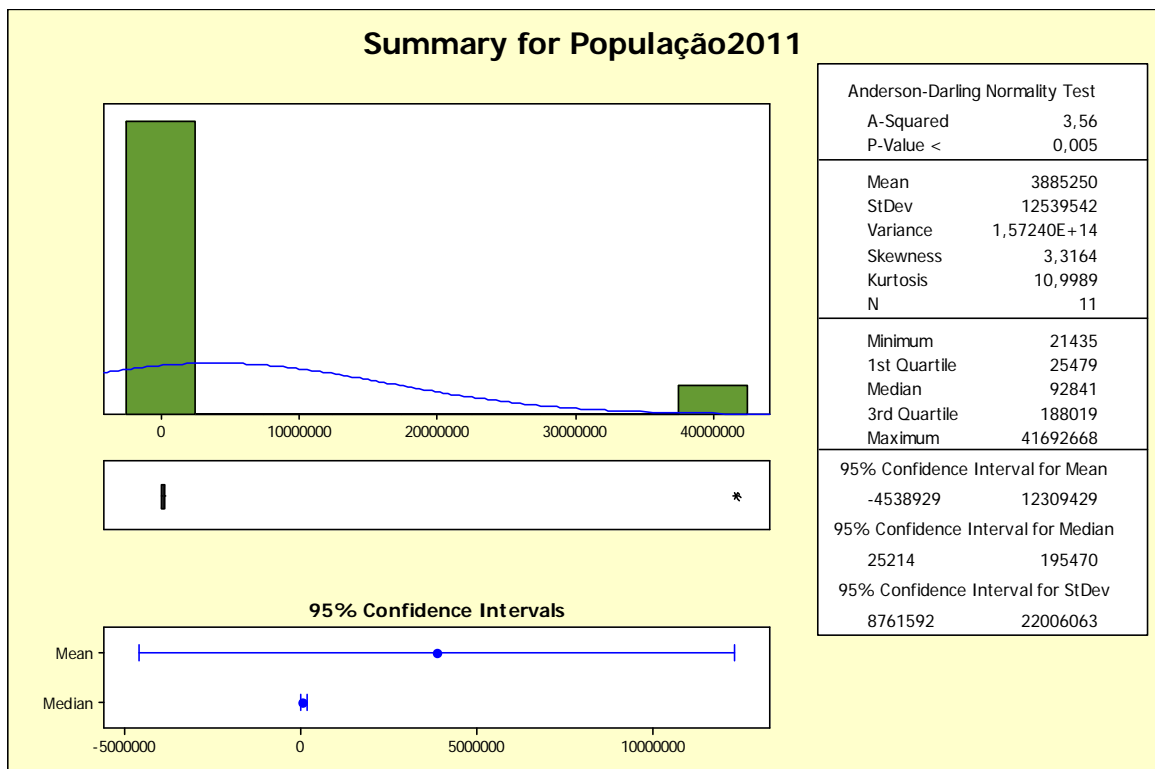
e significativa. O valor médio e a mediana estão muito distantes, confirmando a assimetria na distribuição. Os valores mínimo e máximo apresentam enorme distância, representando que existem cidades com valores significativamente opostos sobre a variável Área, mas isso acontece devido à inserção do Estado de São Paulo na análise.

- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 22981, e o valor da mediana é de 499. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 74700.

3.2 Variável População

Segue o gráfico resumo da variável:

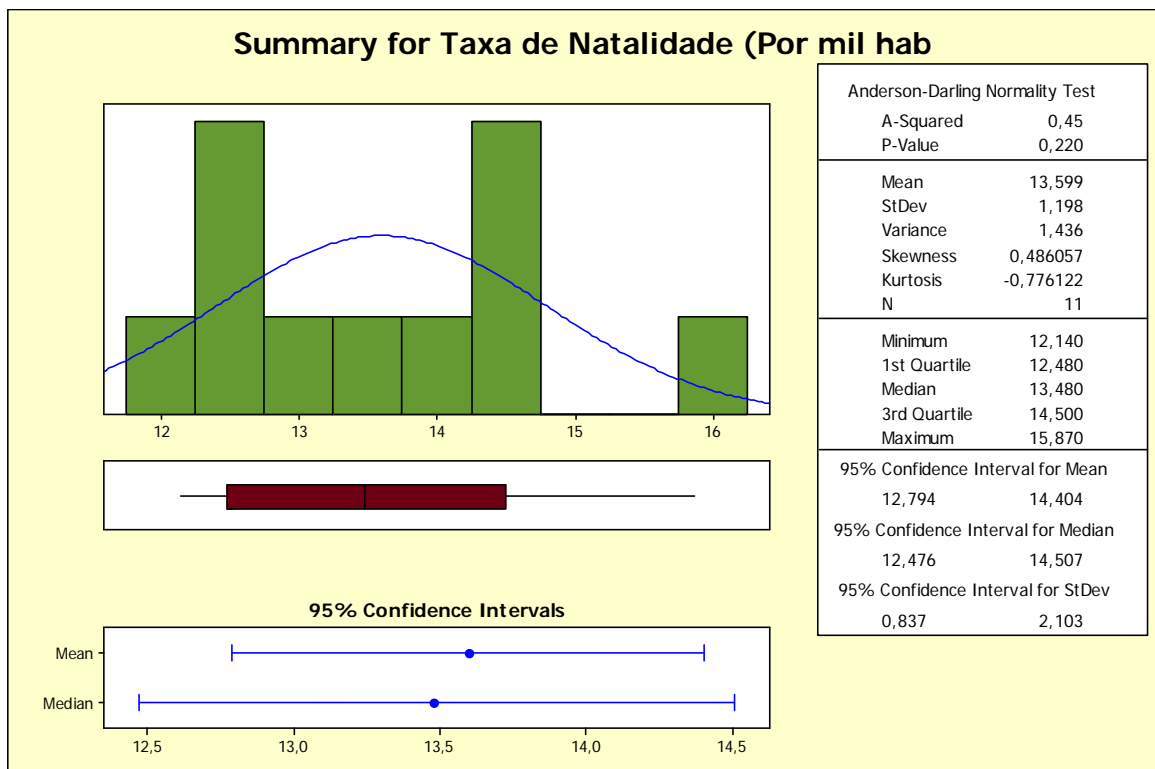


- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição assimétrica, com um pico à esquerda. O P-Value igual a 0,005 indica tratar-se de uma distribuição normal e significativa. O valor médio e a mediana estão muito distantes, confirmando a assimetria na distribuição. Os valores mínimo e máximo apresentam enorme distância, representando que existem cidades com valores significativamente opostos sobre a variável População, mas isso acontece devido à inserção do Estado de São Paulo na análise.

- **Valores Atípicos:** não são apresentados valores atípicos.
- **Centro e Dispersão:** o valor da média é igual a 3885250, e o valor da mediana é de 93841. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 12539542.

3.3 Variável Taxa de Natalidade

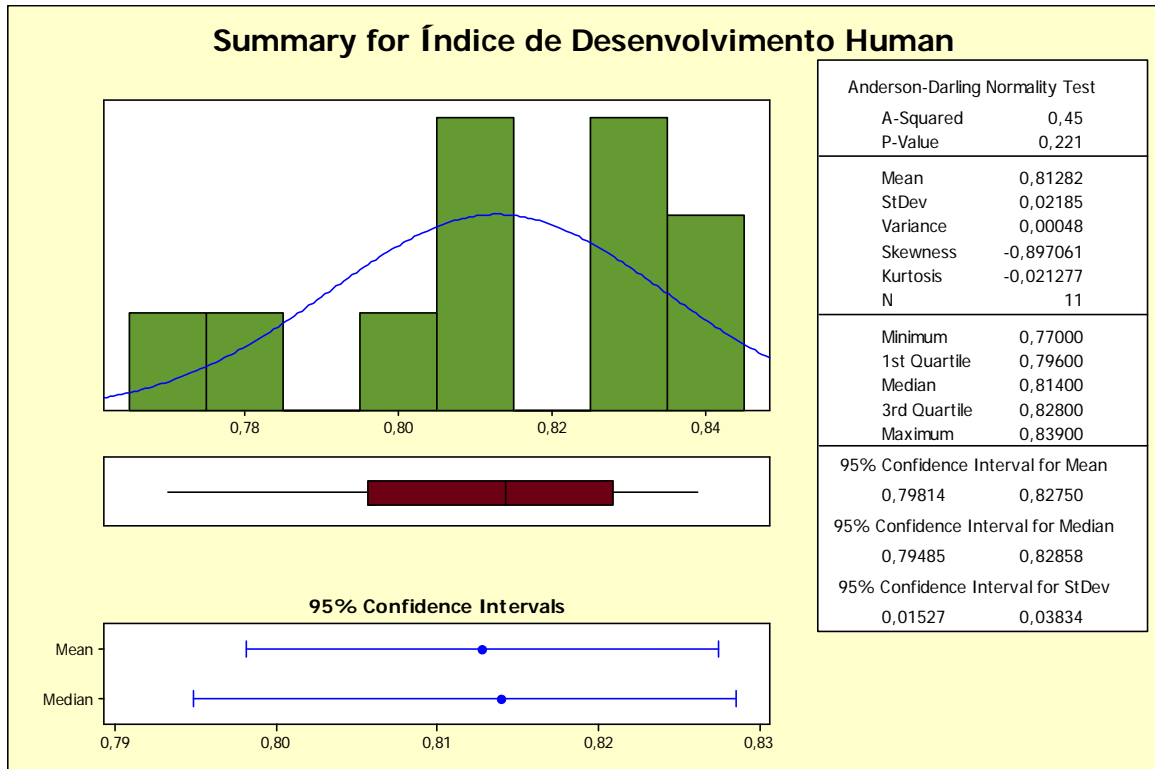
Segue o gráfico resumo da variável:



- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à esquerda. O P-Value igual a 0,220 indica tratar-se de uma distribuição anormal. O valor médio e a mediana estão muito próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam pequena distância, representando que existem cidades com valores significativamente opostos sobre a variável Taxa de Natalidade.
- **Valores Atípicos:** não são apresentados valores atípicos.
- **Centro e Dispersão:** o valor da média é igual a 13,599, e o valor da mediana é de 13,480. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 1,198.

3.4 Variável Índice de Desenvolvimento Humano

Segue o gráfico resumo da variável:



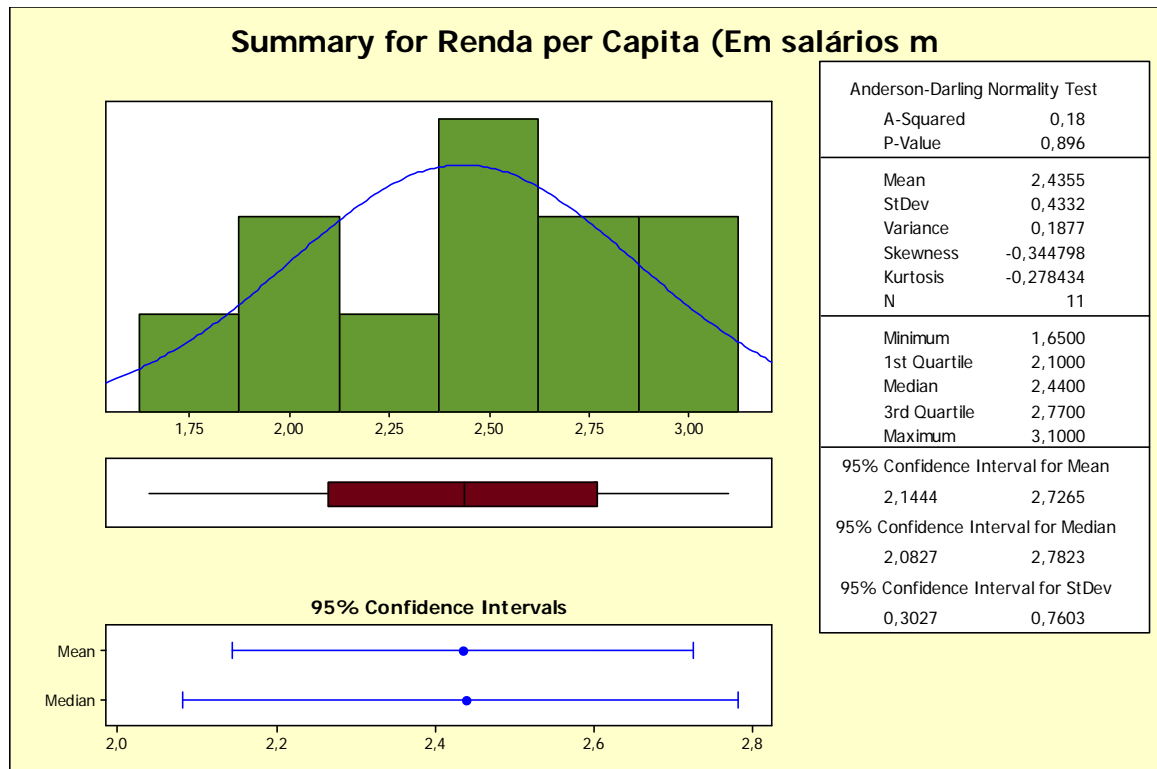
- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à direita. O P-Value igual a 0,221 indica tratar-se de uma distribuição anormal. O valor médio e a mediana estão muito próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam pequena distância, representando que existem cidades com valores significativamente opostos sobre a variável Taxa de Índice de Desenvolvimento Humano.

- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 0,81282, e o valor da mediana é de 0,81400. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 0,02185.

3.5 Variável Renda Per Capita

Segue o gráfico resumo da variável:



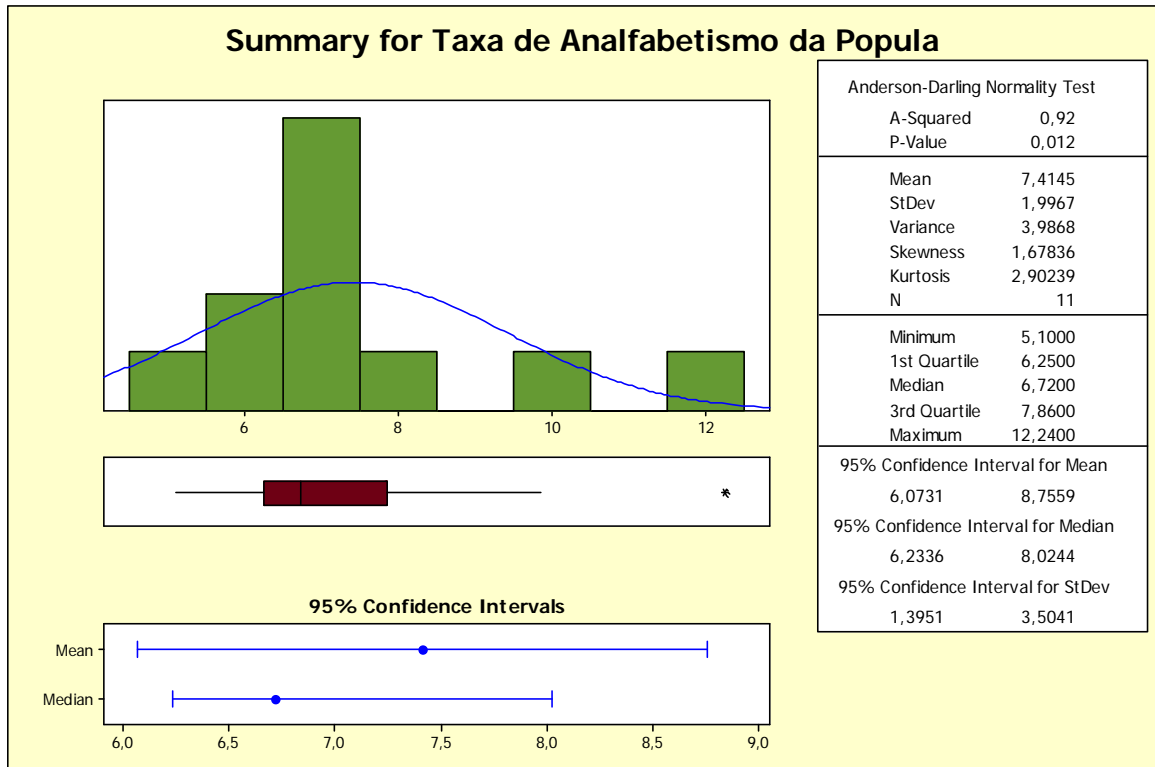
- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à direita. O P-Value igual a 0,896 indica tratar-se de uma distribuição anormal. O valor médio e a mediana estão muito próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam uma certa distância, representando que existem cidades com valores significativamente opostos sobre a variável Renda per Capita.

- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 2,4355, e o valor da mediana é de 2,4400. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 0,4332.

3.6 Variável Taxa de Analfabetismo

Segue o gráfico resumo da variável:



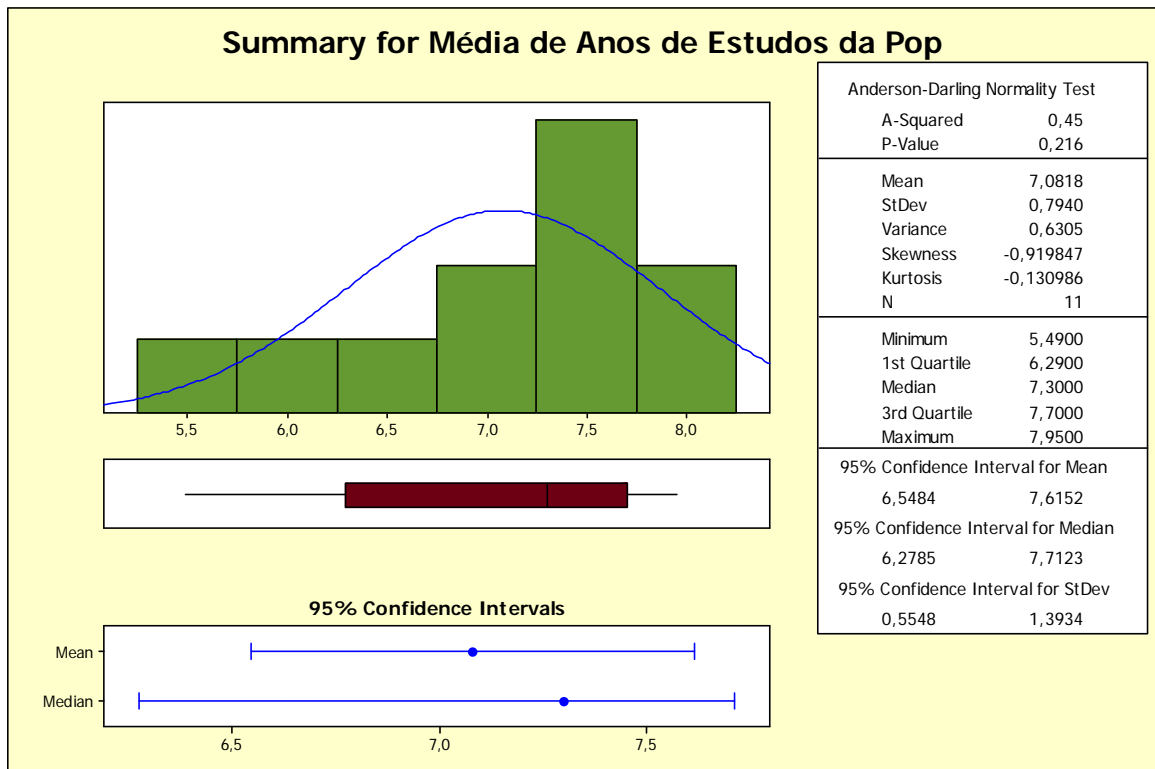
- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à esquerda. O P-Value igual a 0,012 indica tratar-se de uma distribuição anormal. O valor médio e a mediana estão próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam significativa distância, representando que existem cidades com valores significativamente opostos sobre a variável Taxa de Analfabetismo.

- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 7,4145, e o valor da mediana é de 6,7200. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 1,9967.

3.7 Variável Taxa de Anos de Estudo da População

Segue o gráfico resumo da variável:



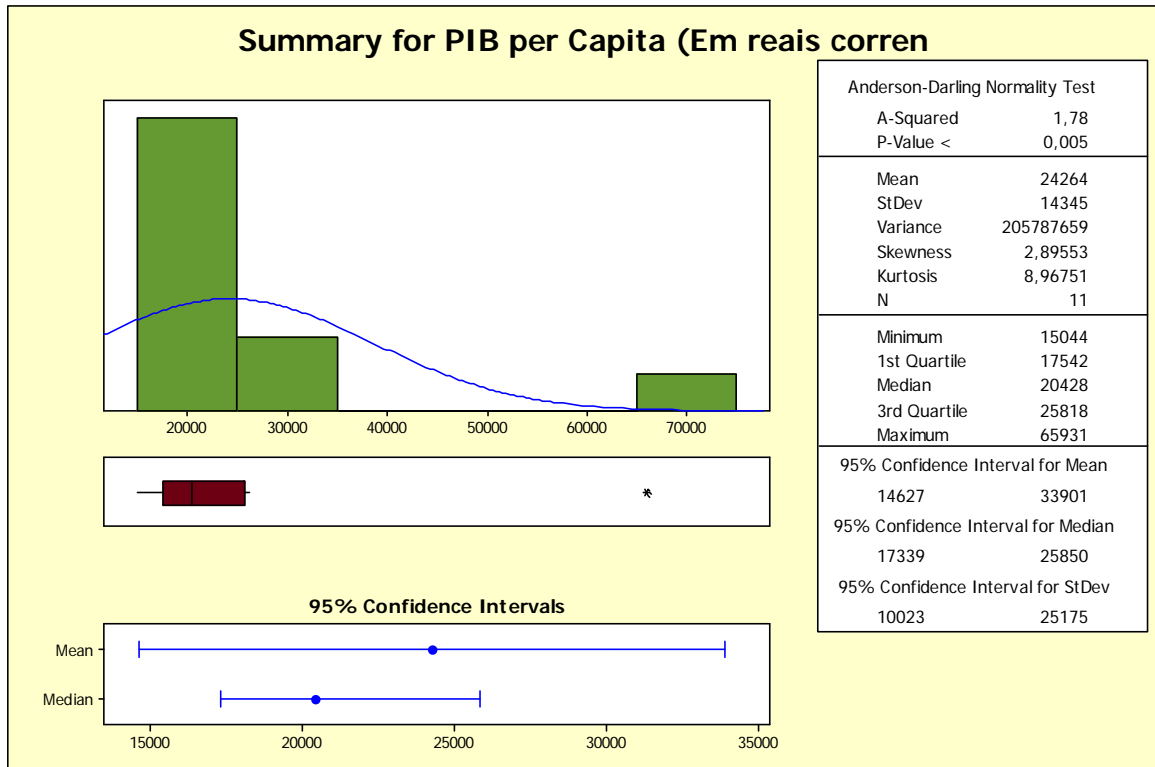
- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à esquerda. O P-Value igual a 0,216 indica tratar-se de uma distribuição anormal. O valor médio e a mediana estão muito próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam pequena distância, representando que existem cidades com valores significativamente opostos sobre a variável Anos de Estudos da População.

- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 7,0818, e o valor da mediana é de 7,3000. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 0,7940.

3.8 Variável PIB per capita

Segue o gráfico resumo da variável:



- **Forma:** O Histograma nos permite verificar que trata-se de uma distribuição simétrica, com um pico um pouco à esquerda. O P-Value igual a 0,005 indica tratar-se de uma distribuição normal. O valor médio e a mediana estão próximos, confirmando a simetria na distribuição mesmo com picos que indicam a diferença entre algumas cidades. Os valores mínimo e máximo apresentam significativa distância, representando que existem cidades com valores significativamente opostos sobre a variável PIB per Capita, mas isso acontece devido à inserção do Estado de São Paulo na análise.

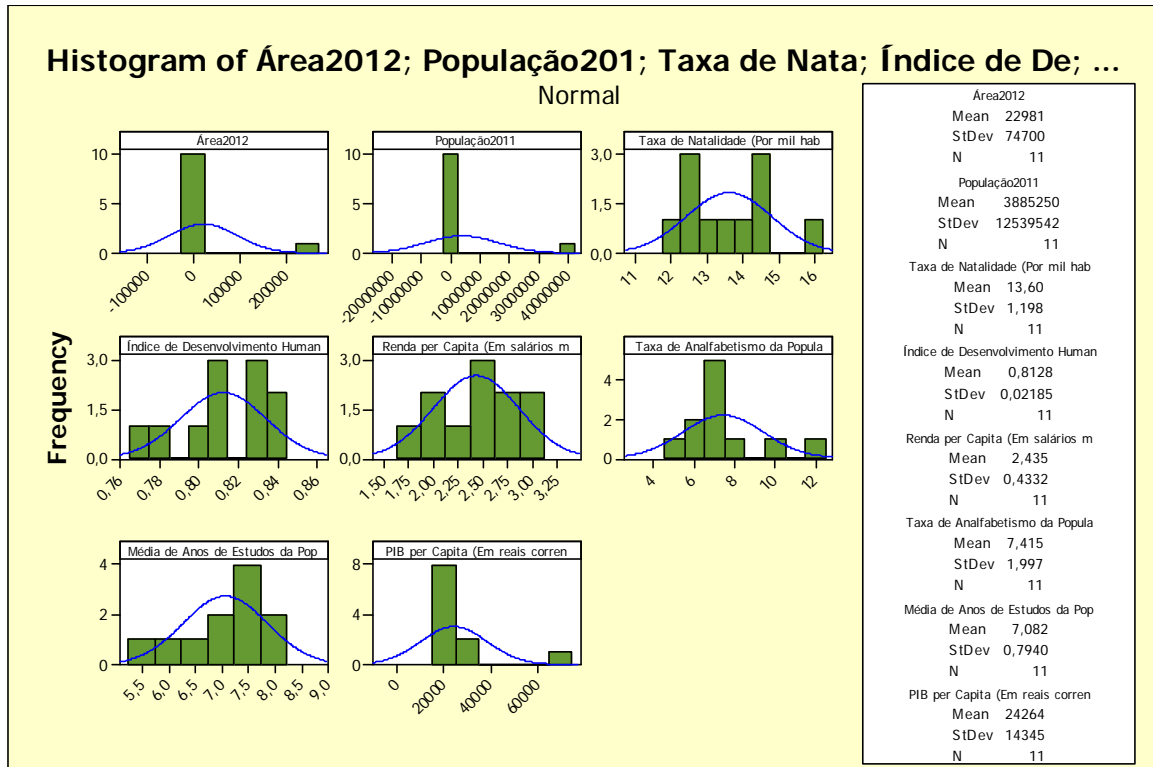
- **Valores Atípicos:** não são apresentados valores atípicos.

- **Centro e Dispersão:** o valor da média é igual a 24264, e o valor da mediana é de 20428. O valor mínimo é bastante distante do valor máximo, demonstrando uma grande dispersão. A medida do desvio-padrão (medida de dispersão) é de 14345.

| Descriptive Statistics: Área2012; População201; Taxa de Nata; Índice de De; ... | | | | | | | | | | |
|---|----|----|---------|---------|----------|---------|-------|--------|--------|----------|
| Variable | N | N* | Mean | SE Mean | StDev | Minimum | Q1 | Median | Q3 | Maximum |
| Área2012 | 11 | 0 | 22981 | 22523 | 74700 | 98 | 184 | 499 | 727 | 248209 |
| População2011 | 11 | 0 | 3885250 | 3780814 | 12539542 | 21435 | 25479 | 92841 | 188019 | 41692668 |
| Taxa de Natalidade (Por | 11 | 0 | 13,599 | 0,361 | 1,198 | 12,14 | 12,48 | 13,48 | 14,5 | 15,87 |
| Índice de Desenvolvement | 11 | 0 | 0,81282 | 0,00659 | 0,02185 | 0,77 | 0,796 | 0,814 | 0,828 | 0,839 |
| Renda per Capita (Em sal | 11 | 0 | 2,435 | 0,131 | 0,433 | 1,65 | 2,1 | 2,44 | 2,77 | 3,1 |
| Taxa de Analfabetismo da | 11 | 0 | 7,415 | 0,602 | 1,997 | 5,1 | 6,25 | 6,72 | 7,86 | 12,24 |
| Média de Anos de Estudos | 11 | 0 | 7,082 | 0,239 | 0,794 | 5,49 | 6,29 | 7,3 | 7,7 | 7,95 |
| PIB per Capita (Em reais | 11 | 0 | 24264 | 4325 | 14345 | 15044 | 17542 | 20428 | 25818 | 65931 |

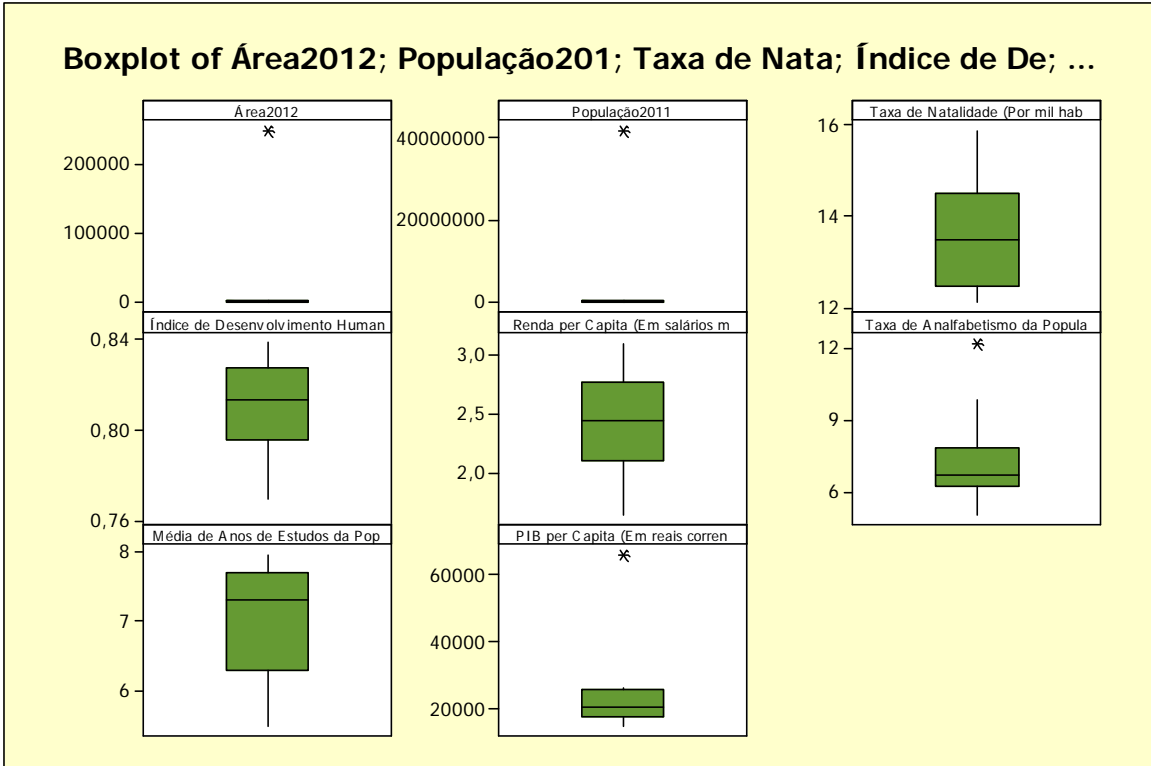
4. GRÁFICO DE HISTOGRAMA

O Histograma abaixo resume e possibilita compararmos graficamente as leituras feitas acima:



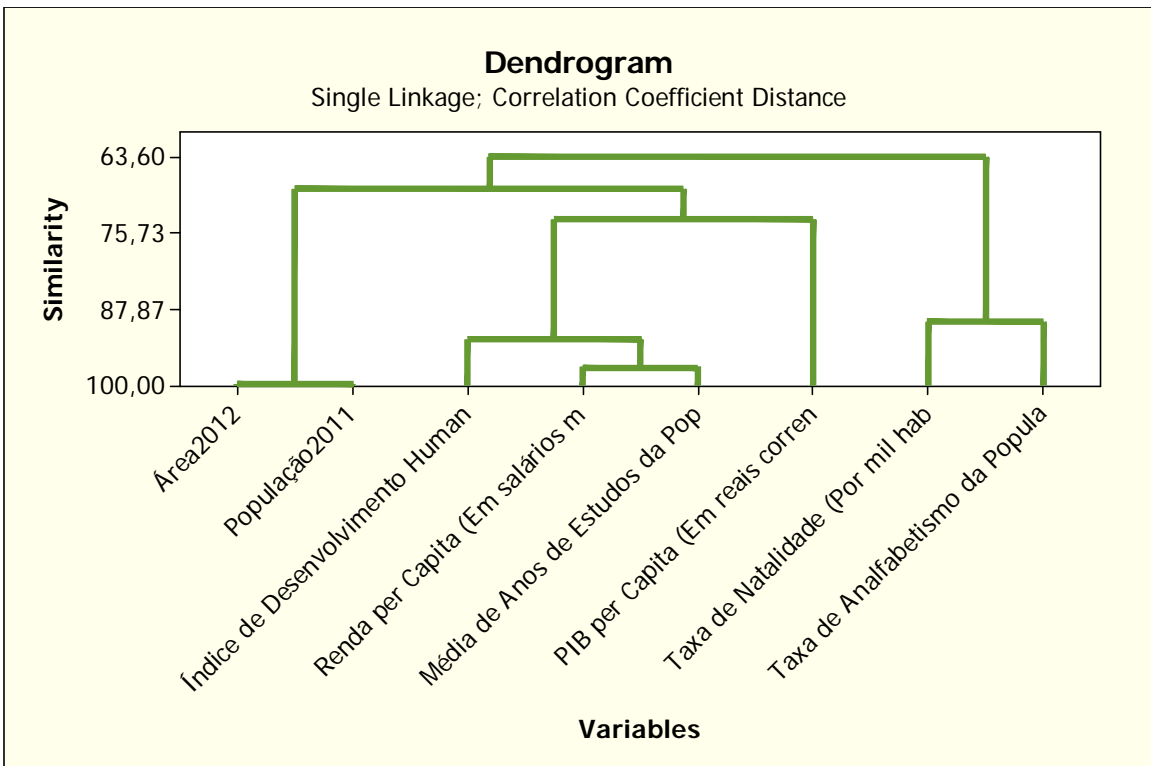
5. GRÁFICO BOXPLOT

O Boxplot abaixo demonstra a relação entre a média e a mediana, analisados acima.

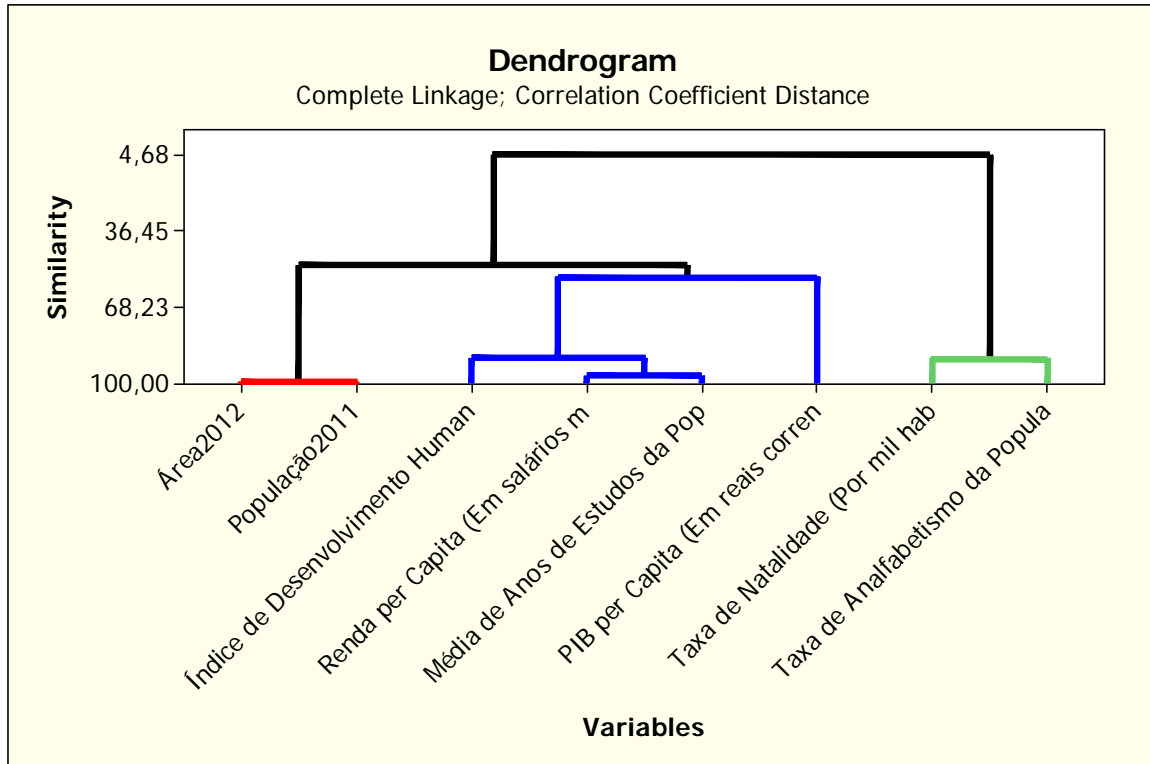


6 - ANÁLISE DE CORRELAÇÃO – DENDOGRAMA

O gráfico abaixo apresentará entre as possíveis variáveis explicativas, as que são mais próximas (similares).



O dendograma acima demonstra a maior similaridade entre as variáveis Área e População, bem como outro grupo que compreende o IDH, Renda e PIB per capita e Média de Anos de Estudos da População e, por fim, um terceiro grupo com forte relação se compõe pela taxa de natalidade com analfabetismo.



Ao dividi-las em 3 clusters, a visualização fica mais clara e é possível perceber os clusters da seguinte maneira:

Cluster Analysis of Variables: Taxa de Nata; Índice de De; Renda per Ca; ...

Correlation Coefficient Distance, Complete Linkage
Amalgamation Steps

| Step | Number of clusters | Similarity level | Distance level | Clusters joined | New cluster | Number of obs. in new cluster |
|------|--------------------|------------------|----------------|-----------------|-------------|-------------------------------|
| 1 | 5 | 97,2412 | 0,05518 | 3 | 5 | 2 |
| 2 | 4 | 89,9808 | 0,20038 | 1 | 4 | 2 |
| 3 | 3 | 89,8664 | 0,20267 | 2 | 3 | 3 |
| 4 | 2 | 56,0445 | 0,87911 | 2 | 6 | 4 |
| 5 | 1 | 4,6807 | 1,90639 | 1 | 2 | 6 |

Final Partition

Cluster 1

Taxa de Natalidade (Por mil hab Taxa de Analfabetismo da Popula

Cluster 2

Índice de Desenvolvimento Human Renda per Capita (Em salários m Média de
Anos de Estudos da Pop
Cluster 3
PIB per Capita (Em reais corren

7 – ANOVA one-way

“O termo ANOVA significa Análise de Variância (*Analysis Of Variance*). A ideia original vem de Ronald A. Fisher, um gênio interdisciplinar que quase sozinho criou as bases para a Estatística Aplicada, que num trabalho de 1921 estendeu o teste de comparação t de Gosset de 1908 para mais de dois grupos, o que deu origem ao chamado teste ANOVA ou teste F de Fisher”. (Las Casas/Guevara, 2010)

A seguir será apresentada a análise ANOVA para que sejam identificadas as similaridades entre as variáveis.

7.1. ANOVA Área 2012 versus Clusters

Ao realizar a ANOVA One-way da variável Área 2012 versus clusters, o resultado obtido foi:

One-way ANOVA: Área2012 versus Clusters

| Source | DF | SS | MS | F | P |
|----------|----|-------------|------------|------|-------|
| Clusters | 1 | 6765829247 | 6765829247 | 1,24 | 0,294 |
| Error | 9 | 49035272104 | 5448363567 | | |
| Total | 10 | 55801101351 | | | |

S = 73813 R-Sq = 12,12% R-Sq(adj) = 2,36%

Individual 95% CIs For Mean Based on Pooled StDev

| Level | N | Mean | StDev | CI Lower | CI Upper |
|-------|---|-------|--------|----------|----------|
| 2 | 5 | 50149 | 110719 | (-50000) | (100000) |
| 3 | 6 | 341 | 246 | (-50000) | (100000) |

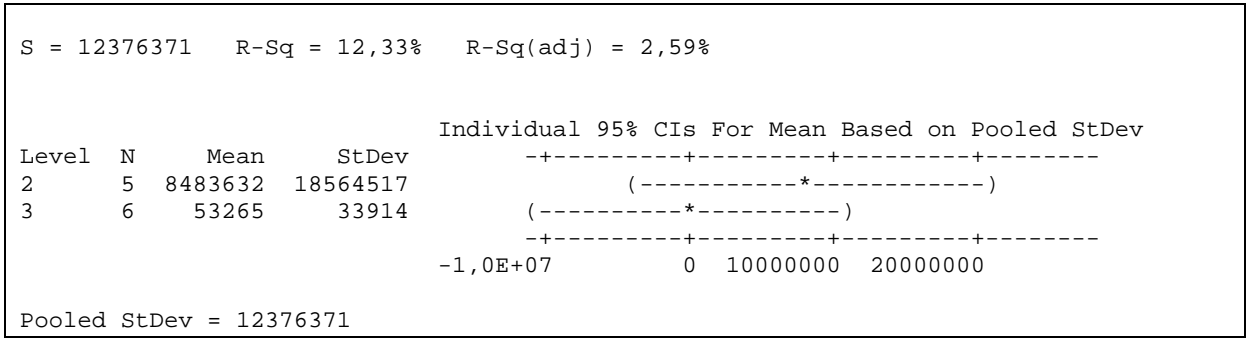
Pooled StDev = 73813

7.2. ANOVA População x Cluster

Ao realizar a ANOVA One-way da variável População versus clusters, o resultado obtido foi:

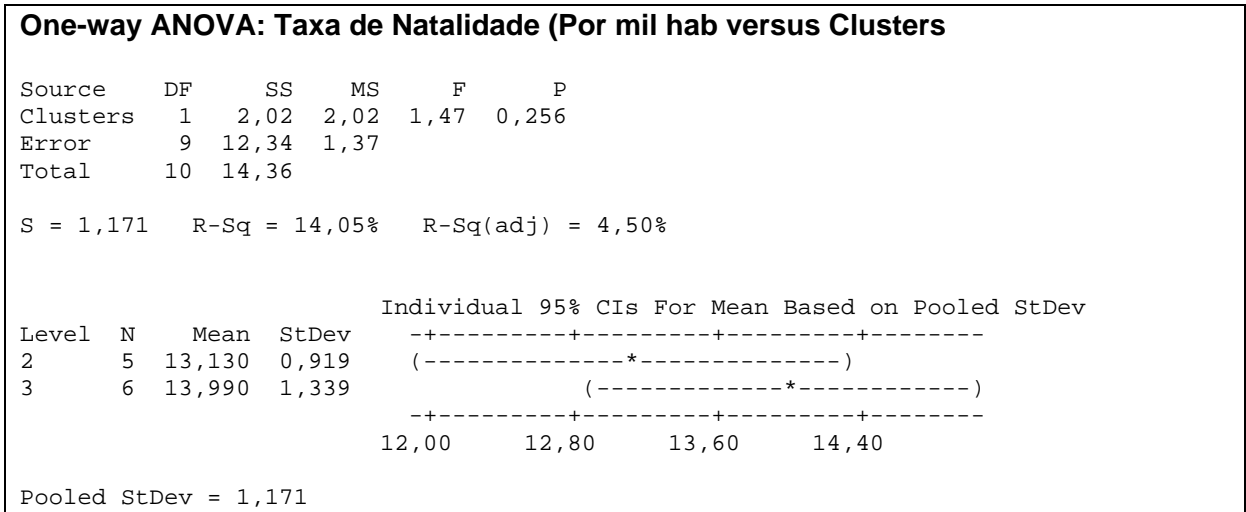
One-way ANOVA: População2011 versus Clusters

| Source | DF | SS | MS | F | P |
|----------|----|-------------|-------------|------|-------|
| Clusters | 1 | 1,93830E+14 | 1,93830E+14 | 1,27 | 0,290 |
| Error | 9 | 1,37857E+15 | 1,53175E+14 | | |
| Total | 10 | 1,57240E+15 | | | |



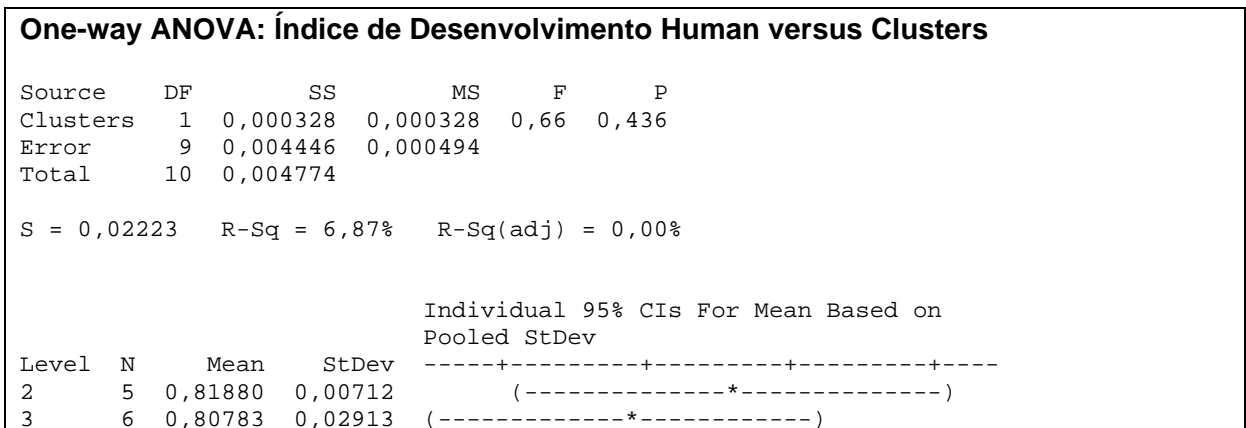
7.3. ANOVA Índice Taxa de Natalidades versus Clusters

Ao realizar a ANOVA One-way da variável Taxa de Natalidade versus clusters, o resultado obtido foi:



7.4. ANOVA Índice de Desenvolvimento Humano versus Cluster

Ao realizar a ANOVA One-way da variável Índice de Desenvolvimento Humano versus clusters, o resultado obtido foi:



| |
|--|
| <pre> -----+-----+-----+-----+----- 0,795 0,810 0,825 0,840 Pooled StDev = 0,02223 </pre> |
|--|

7.5. ANOVA Renda per capita versus Cluster

Ao realizar a ANOVA One-way da variável Renda per capita versus clusters, o resultado obtido foi:

| One-way ANOVA: Renda per Capita (Em salários m versus Clusters | | | | | |
|--|----|-------|-------|------|-------|
| Source | DF | SS | MS | F | P |
| Clusters | 1 | 0,446 | 0,446 | 2,80 | 0,128 |
| Error | 9 | 1,431 | 0,159 | | |
| Total | 10 | 1,877 | | | |

S = 0,3987 R-Sq = 23,76% R-Sq(adj) = 15,28%

| Individual 95% CIs For Mean Based on Pooled StDev | | | | | |
|---|---|--------|--------|---------------|--|
| Level | N | Mean | StDev | | |
| 2 | 5 | 2,6560 | 0,3333 | (-----*-----) | |
| 3 | 6 | 2,2517 | 0,4442 | (-----*-----) | |

```

-----+-----+-----+-----+-----
                2,10    2,40    2,70    3,00

```

Pooled StDev = 0,3987

7.6. ANOVA Taxa de Analfabetismo versus Cluster

Ao realizar a ANOVA One-way da variável Analfabetismo da População versus clusters, o resultado obtido foi:

| One-way ANOVA: Taxa de Analfabetismo da Popula versus Clusters | | | | | |
|--|----|-------|------|------|-------|
| Source | DF | SS | MS | F | P |
| Clusters | 1 | 8,99 | 8,99 | 2,62 | 0,140 |
| Error | 9 | 30,87 | 3,43 | | |
| Total | 10 | 39,87 | | | |

S = 1,852 R-Sq = 22,56% R-Sq(adj) = 13,96%

| Individual 95% CIs For Mean Based on Pooled StDev | | | | | |
|---|---|-------|-------|---------------|--|
| Level | N | Mean | StDev | | |
| 2 | 5 | 6,424 | 0,815 | (-----*-----) | |
| 3 | 6 | 8,240 | 2,376 | (-----*-----) | |

```

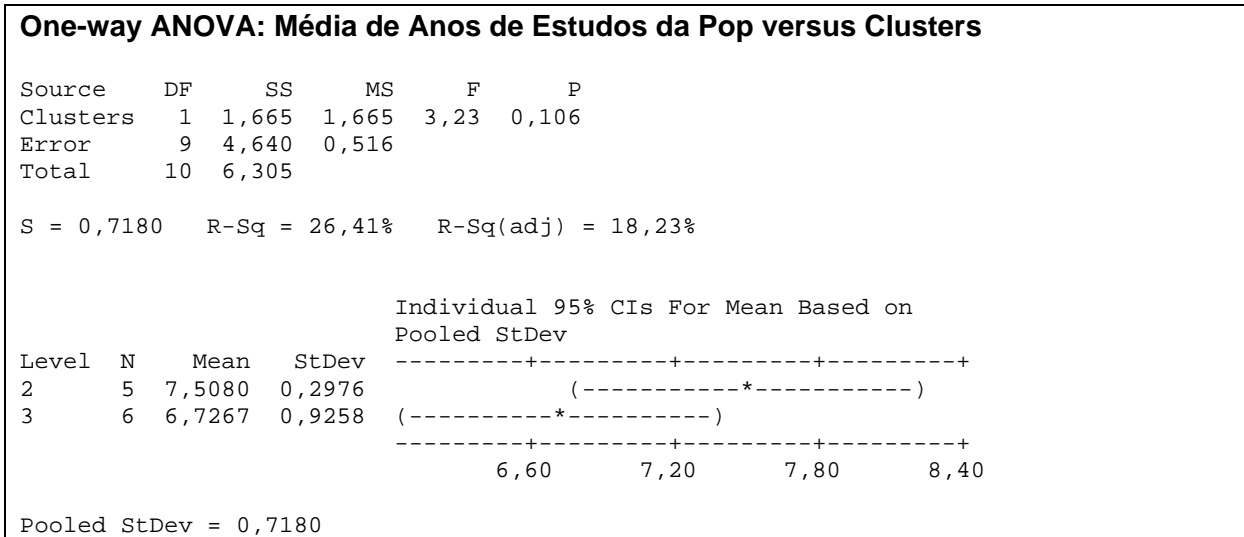
-----+-----+-----+-----+-----
                4,5    6,0    7,5    9,0

```

Pooled StDev = 1,852

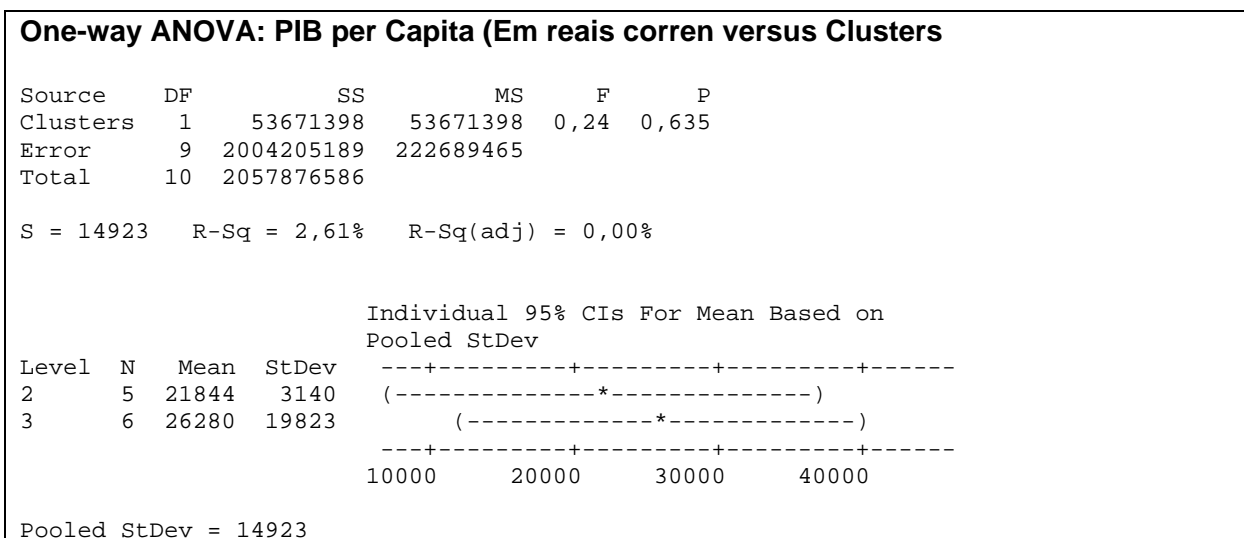
7.7. ANOVA Média de Estudos da População versus Cluster

Ao realizar a ANOVA One-way da variável Média de Estudos da População versus clusters, o resultado obtido foi:



7.8. ANOVA PIB per capita versus Cluster

Ao realizar a ANOVA One-way da variável PIB per capita versus clusters, o resultado obtido foi:



8 – MATRIZ DE CORRELAÇÃO

“As correlações entre as variáveis (valores em -1 e 1) mostram o grau de relação linear que existe entre elas. Valores próximos de zero significam nenhuma relação, e próximos de 1 ou -1 mostram alta relação positiva ou negativa respectivamente.” (Las Casas/Guevara 2010).

A matriz de correlação abaixo nos ajuda tirar conclusões mais precisas sobre a associação entre as variáveis.

| Correlations: Área2012; População201; Taxa de Nata; Índice de De; ... | | | |
|--|------------------|------------------|------------------|
| | Área2012 | População2011 | Taxa de Natalida |
| População2011 | 1,000 0,000 | | |
| Taxa de Natalida | 0,272 0,418 | 0,271 0,421 | |
| Índice de Desenv | 0,020 0,954 | 0,020 0,954 | -0,791 0,004 |
| Renda per Capita | 0,373 0,259 | 0,374 0,258 | -0,671 0,024 |
| Taxa de Analfabe | -0,130 0,703 | -0,132 0,700 | 0,800 0,003 |
| Média de Anos de | 0,235 0,486 | 0,237 0,484 | -0,834 0,001 |
| PIB per Capita (| 0,044 0,899 | 0,043 0,900 | 0,038 0,912 |
| | Índice de Desenv | Renda per Capita | Taxa de Analfabe |
| Renda per Capita | 0,797 0,003 | | |
| Taxa de Analfabe | -0,820 0,002 | -0,845 0,001 | |
| Média de Anos de | 0,855 0,001 | 0,945 0,000 | -0,906 0,000 |
| PIB per Capita (| 0,473 0,142 | 0,229 0,497 | -0,293 0,381 |
| | Média de Anos de | | |
| PIB per Capita (| 0,121 0,723 | | |
| Cell Contents: Pearson correlation P-Value | | | |

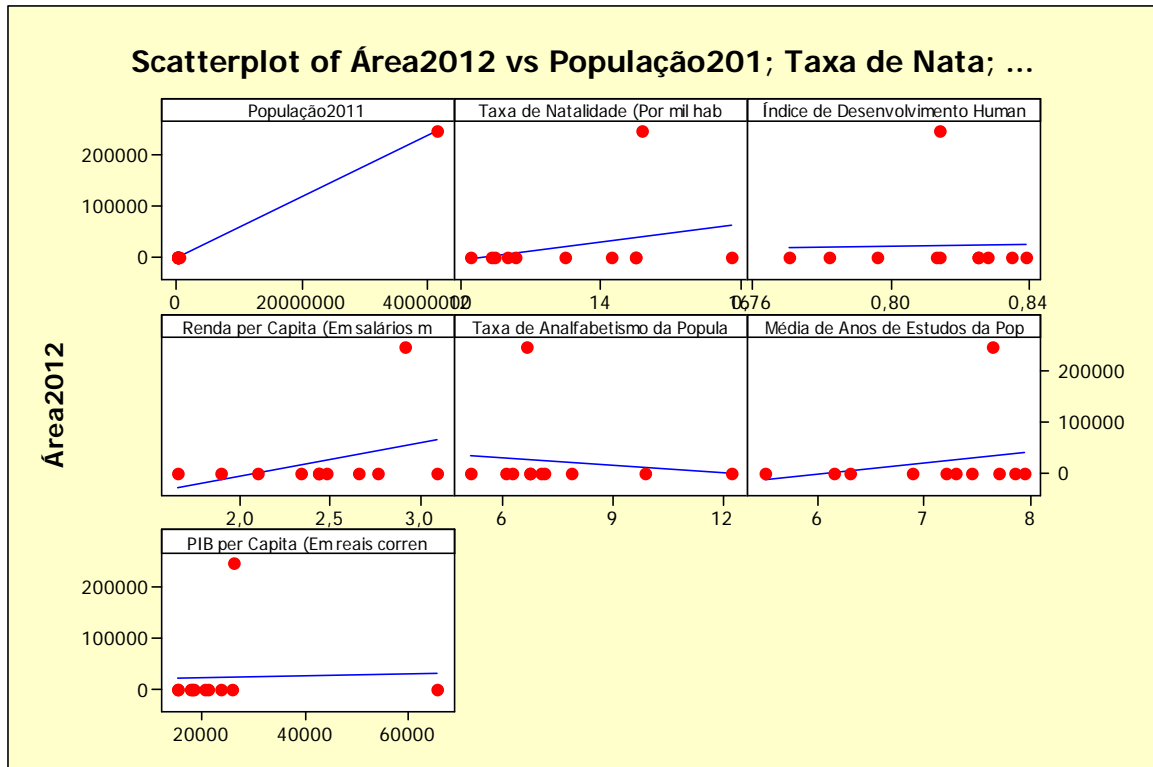
A matriz de correlação acima confirma a forte relação entre as variáveis relacionadas à Taxa de Analfabetismo e Natalidade, bem como a média dos Anos de Estudo da População.

| Principal Component Analysis: Área2012; População201; Taxa de Nata; Índice de D | | | | | | | | |
|---|--------|--------|--------|--------|--------|--------|--------|--------|
| Eigenanalysis of the Correlation Matrix | | | | | | | | |
| Eigenvalue | 4,4288 | 2,2018 | 1,0788 | 0,1564 | 0,1093 | 0,0136 | 0,0112 | 0 |
| Proportion | 0,554 | 0,275 | 0,135 | 0,02 | 0,014 | 0,002 | 0,001 | 0 |
| Cumulative | 0,554 | 0,829 | 0,964 | 0,983 | 0,997 | 0,999 | 1 | 1 |
| Variable | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 |
| Área2012 | 0,1 | 0,656 | 0,056 | -0,034 | 0,203 | 0,113 | -0,01 | 0,708 |
| População2011 | 0,101 | 0,656 | 0,057 | -0,029 | 0,2 | 0,132 | 0,005 | -0,706 |
| Taxa de Natalidade (Por mil hab | -0,388 | 0,325 | -0,27 | 0,071 | -0,328 | -0,742 | -0,084 | -0,01 |
| Índice de Desenvolvimento Human | 0,438 | -0,12 | -0,216 | -0,576 | 0,366 | -0,378 | -0,373 | -0,012 |
| Renda per Capita (Em salários m | 0,448 | 0,129 | 0,065 | -0,11 | -0,789 | 0,176 | -0,337 | 0,001 |
| Taxa de Analfabetismo da Popula | -0,448 | 0,048 | 0,003 | -0,798 | -0,194 | 0,234 | 0,26 | 0,003 |
| Média de Anos de Estudos da Pop | 0,464 | 0,015 | 0,177 | -0,075 | -0,121 | -0,344 | 0,784 | 0,004 |
| PIB per Capita (Em reais corren | 0,143 | 0,025 | -0,916 | 0,086 | -0,045 | 0,27 | 0,241 | 0,006 |

Por seus valores, P1, que explica 55% das variáveis, pode ser denominado como Índice de Relação entre o Índice de Desenvolvimento Humano, Renda Per capita e a Média de Estudos da População. Já P2, que explica 27%, pode ser denominado como Demonstrativo de Relevância apenas da Área e População.

9 – SCATTERPLOT

9.1 - Scatterplot de Área versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

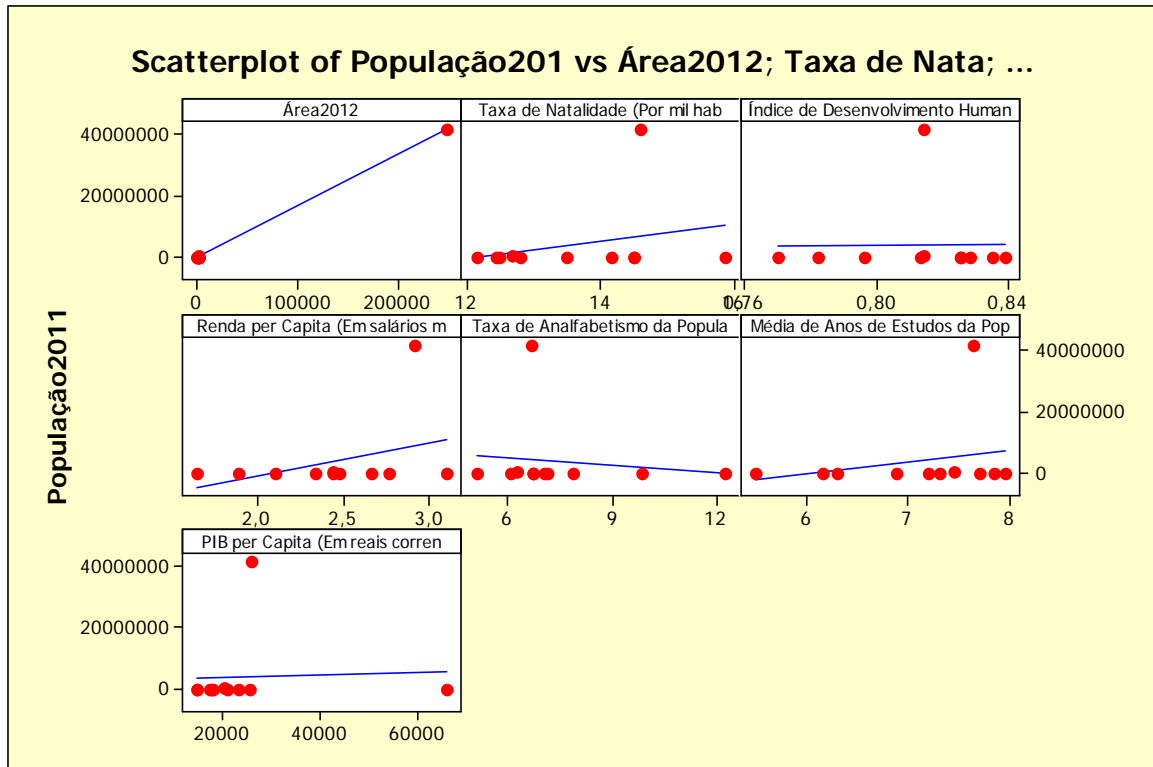
Direção: Grande parte das dispersões acima nos permite verificar um padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Área com as demais. Todos os gráficos apresentam relativas dispersões com exceção do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima parecem denotar forma crescente.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.2 - Scatterplot de População versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

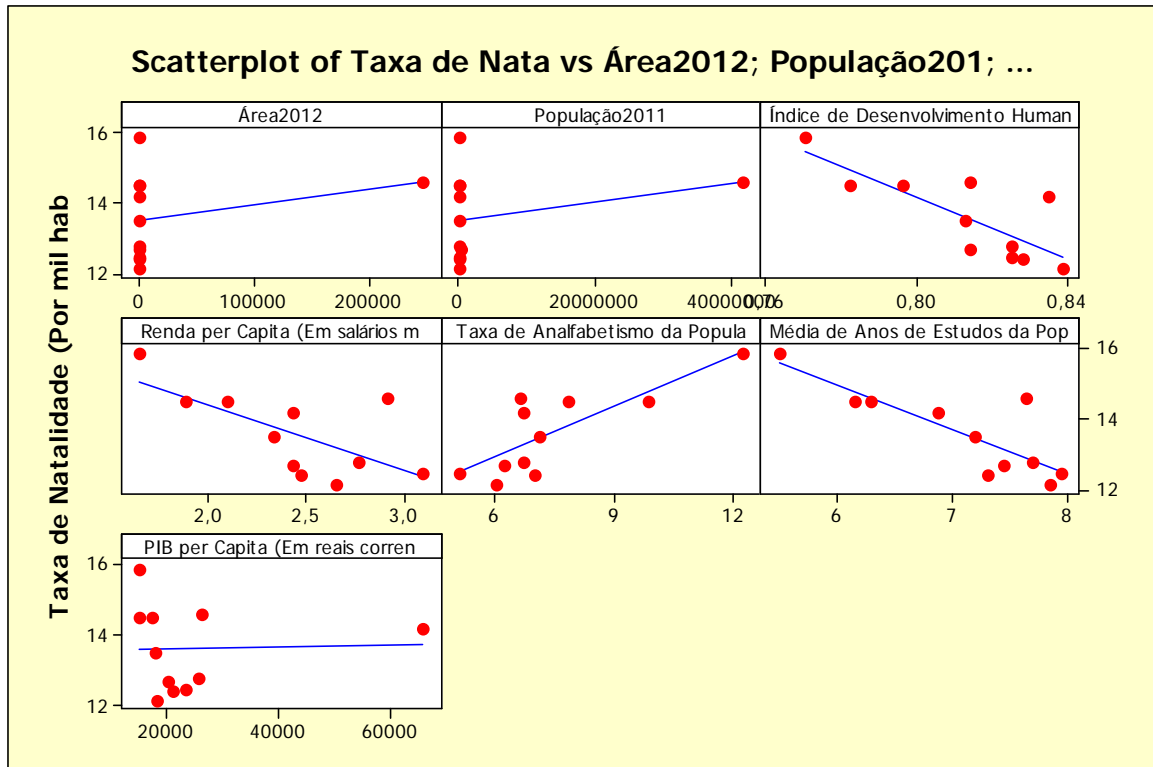
Direção: Grande parte das dispersões acima nos permite verificar um padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável População com as demais. Todos os gráficos apresentam relativas dispersões com exceção do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima parecem denotar forma crescente.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.3 - Scatterplot de Taxa de Natalidade versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

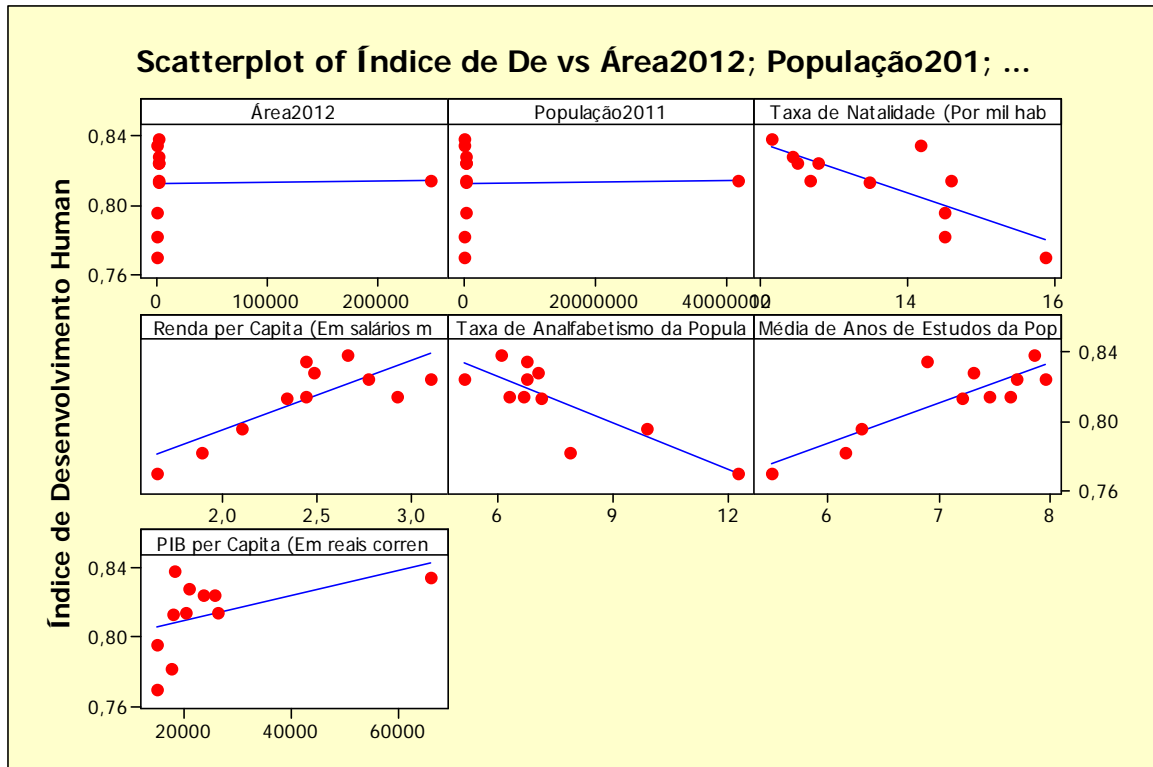
Direção: Grande parte das dispersões acima nos permite verificar uma falta de padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima apresentam uma variância quanto à forma.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.4 - Scatterplot de Índice de Desenvolvimento Humano versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

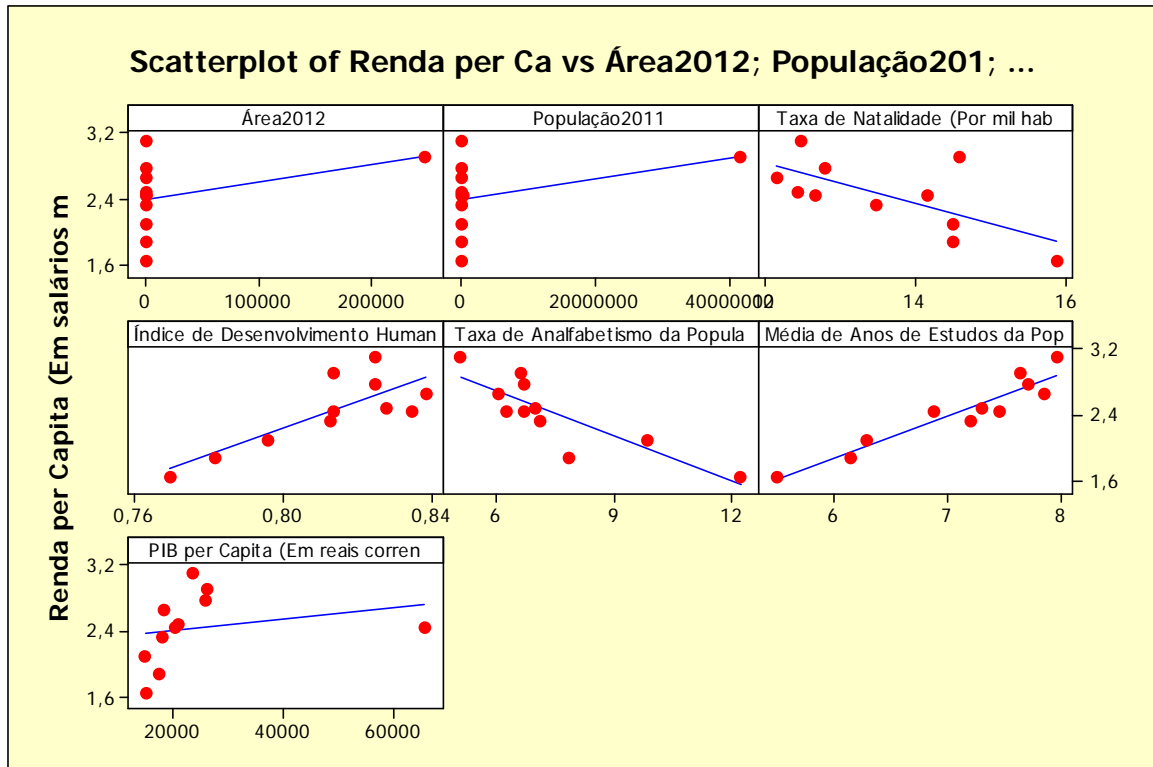
Direção: Grande parte das dispersões acima nos permite verificar uma falta de padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima apresentam uma variância quanto à forma.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.5 - Scatterplot de Renda per Capita versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

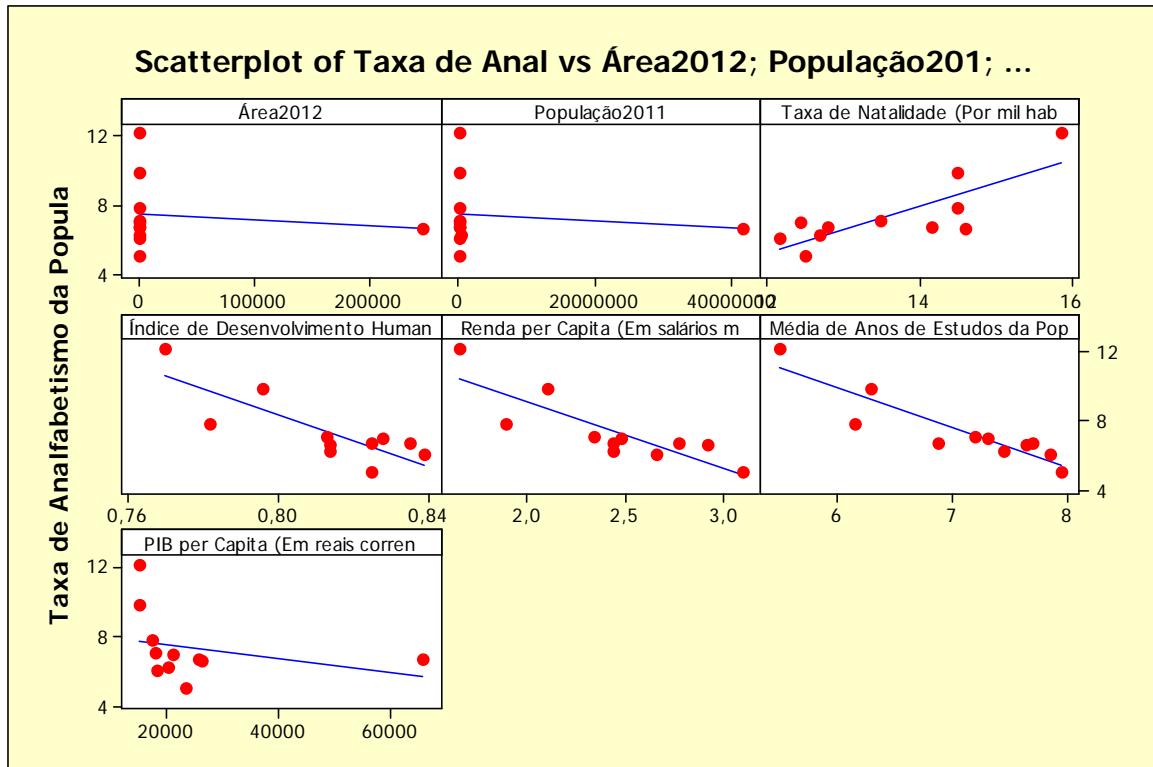
Direção: Grande parte das dispersões acima nos permite verificar uma falta de padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima apresentam uma variância quanto à forma.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.6 - Scatterplot de Taxa de Analfabetismo versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

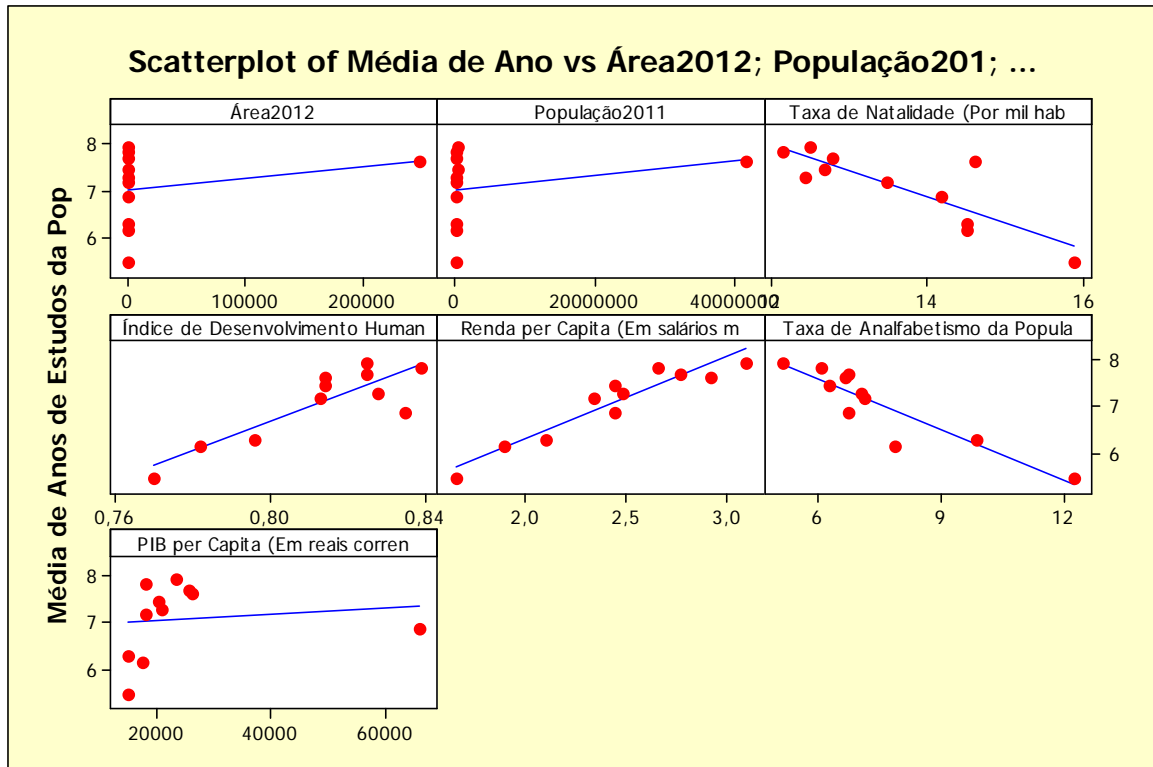
Direção: Grande parte das dispersões acima nos permite verificar uma falta de padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima parecem denotar forma decrescente em sua maioria.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.7 - Scatterplot de Média de Anos de Estudos versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

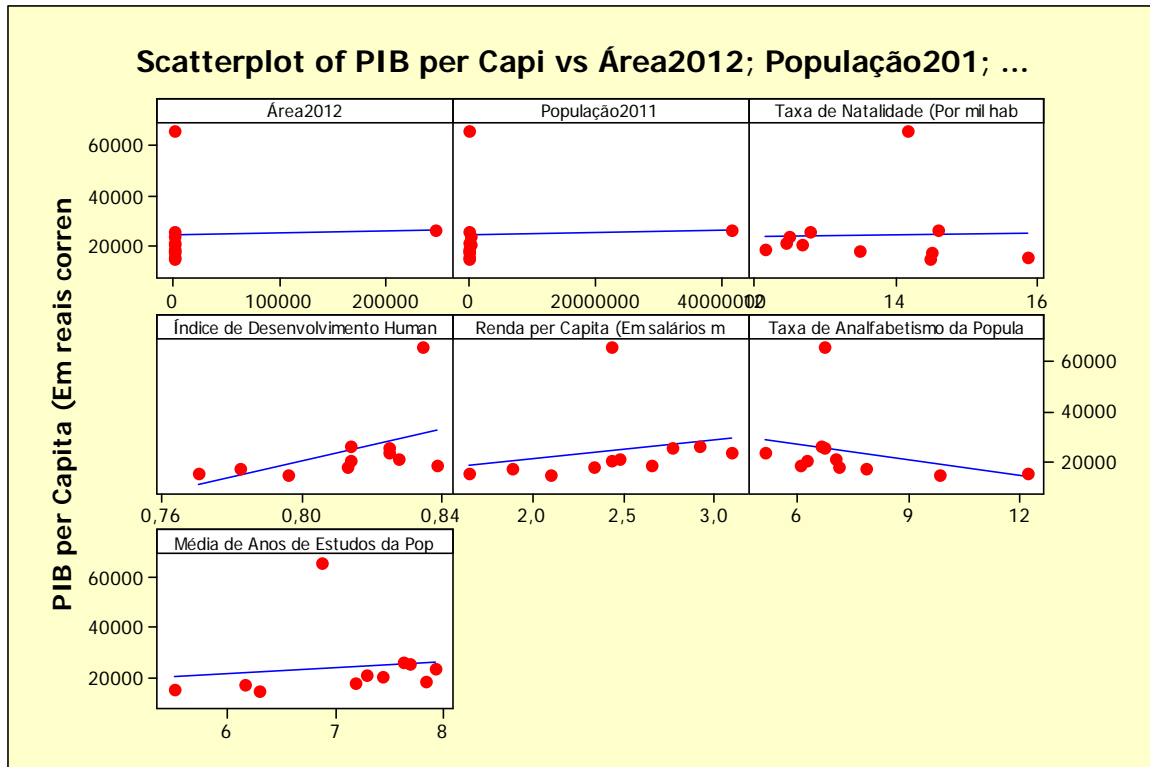
Direção: Grande parte das dispersões acima nos permite verificar uma falta de padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima apresentam uma variância quanto à forma.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

9.8 - Scatterplot de PIB per Capita versus as demais variáveis



Segue abaixo uma breve análise dos gráficos acima:

Direção: Grande parte das dispersões acima nos permite verificar um padrão linear na maioria dos gráficos, mostrando que há uma relação, ao menos de evidência visual, da variável Taxa de Natalidade com as demais. Todos os gráficos apresentam relativas dispersões com exceção de Área, População e do PIB per Capita que apresenta uma relação maior entre as cidades diferenciando apenas do Estado de SP por conta de seus valores.

Intensidade: É possível notar grande intensidade na maioria dos gráficos, o que confirma a existência de relações entre as variáveis analisadas.

Forma: De maneira visual os gráficos acima apresentam uma variância quanto à forma, porém são mais lineares.

Valores Atípicos: Alguns gráficos indicam a existência de valores atípicos, ou seja, cidades que apresentam valores distantes dos demais.

10- ANÁLISE DE REGRESSÃO

10.1 – Stepwise

A análise por meio da tabela de regressão Stepwise, abaixo apresentada, fornecerá mais informações sobre essa hipótese. Para isso, foi feita uma relação entre a variável Índice de Desenvolvimento Humano e as demais variáveis:

Stepwise Regression: Índice de De versus Área2012; População201; ...

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

Response is Índice de Desenvolvimento Human on 7 predictors, with N = 11

| Step | 1 | 2 | 3 | 4 |
|---------------------------------|--------|---------|---------|---------|
| Constant | 0,6462 | 0,6411 | 0,8270 | 0,7570 |
| Média de Anos de Estudos da Pop | 0,0235 | 0,0223 | 0,0116 | 0,0198 |
| T-Value | 4,95 | 6,30 | 2,23 | 3,04 |
| P-Value | 0,001 | 0,000 | 0,061 | 0,023 |
| PIB per Capita (Em reais corren | | 0,00000 | 0,00000 | 0,00000 |
| T-Value | | 2,91 | 4,18 | 5,00 |
| P-Value | | 0,019 | 0,004 | 0,002 |
| Taxa de Natalidade (Por mil hab | | | -0,0083 | -0,0102 |
| T-Value | | | -2,41 | -3,19 |
| P-Value | | | 0,047 | 0,019 |
| Taxa de Analfabetismo da Popula | | | | 0,0048 |
| T-Value | | | | 1,76 |
| P-Value | | | | 0,129 |
| S | 0,0119 | 0,00882 | 0,00697 | 0,00611 |
| R-Sq | 73,14 | 86,97 | 92,89 | 95,31 |
| R-Sq(adj) | 70,15 | 83,72 | 89,84 | 92,19 |

11. ANÁLISE DISCRIMINANTE

A variável dependente dessa análise será o Cluster, e, para tentar explicar em qual conglomerado uma determinada cidade está, foram utilizadas 6 variáveis.

Segue o resultado da Análise Discriminante, com opção LINEAR:

Discriminant Analysis: Clusters versus Taxa de Nata; Índice de De; ...

Linear Method for Response: Clusters

Predictors: Taxa de Natalidade (Por mil hab; Índice de Desenvolvimento Human;
Renda per Capita (Em salários m; Taxa de Analfabetismo da Popula;
Média de Anos de Estudos da Pop; PIB per Capita (Em reais corren

| | | |
|-------|---|---|
| Group | 2 | 3 |
| Count | 5 | 6 |

Summary of classification

| Put into Group | True Group | | | | |
|---|------------|------------|-------|------------------|-------------|
| | 2 | 3 | | | |
| 2 | 4 | 2 | | | |
| 3 | 1 | 4 | | | |
| Total N | 5 | 6 | | | |
| N correct | 4 | 4 | | | |
| Proportion | 0,800 | 0,667 | | | |
| N = 11 N Correct = 8 Proportion Correct = 0,727 | | | | | |
| Squared Distance Between Groups | | | | | |
| | 2 | 3 | | | |
| 2 | 0,00000 | 2,13297 | | | |
| 3 | 2,13297 | 0,00000 | | | |
| Linear Discriminant Function for Groups | | | | | |
| | 2 | 3 | | | |
| Constant | -12390 | -12423 | | | |
| Taxa de Natalidade (Por mil hab | 352 | 351 | | | |
| Índice de Desenvolvimento Human | 30571 | 30638 | | | |
| Renda per Capita (Em salários m | -45 | -44 | | | |
| Taxa de Analfabetismo da Popula | -115 | -115 | | | |
| Média de Anos de Estudos da Pop | -464 | -467 | | | |
| PIB per Capita (Em reais corren | -0 | -0 | | | |
| Summary of Misclassified Observations | | | | | |
| Observation | True Group | Pred Group | Group | Squared Distance | Probability |
| 2** | 2 | 3 | 2 | 4,092 | 0,342 |
| | | | 3 | 2,782 | 0,658 |
| 8** | 3 | 2 | 2 | 2,917 | 0,698 |
| | | | 3 | 4,589 | 0,302 |
| 9** | 3 | 2 | 2 | 4,180 | 0,652 |
| | | | 3 | 5,434 | 0,348 |

O resultado pode ser considerado somente bom, já que o modelo realizou 8 opções corretas, em 11 opções possíveis:

| | | |
|--------|---------------|---------------------------|
| N = 11 | N Correct = 8 | Proportion Correct = 0,72 |
|--------|---------------|---------------------------|

A opção quadrática não trouxe resultados

Discriminant Analysis: Clusters versus Taxa de Nata; Índice de De; ...

* ERROR *

PIB per Capita (Em reais corren is highly correlated with other predictors in group 2.

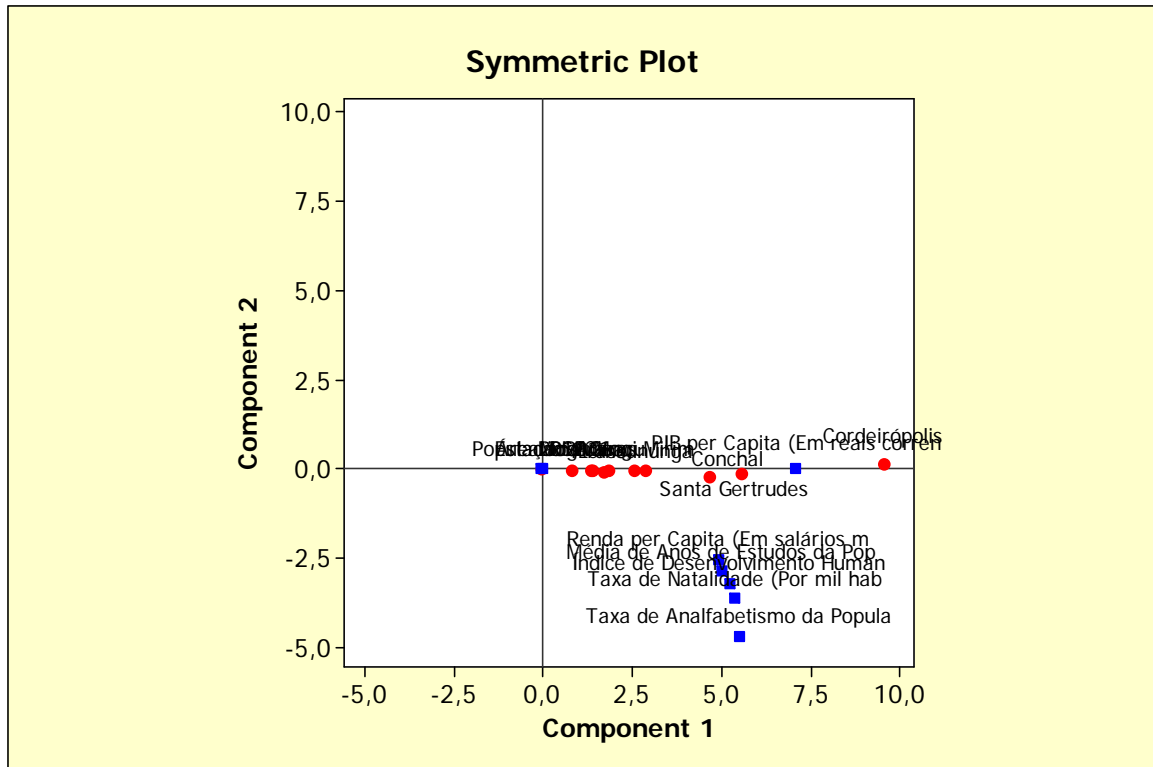
* ERROR * Calculations for discriminant analysis cannot be done.

12 – ANÁLISE DE CORRESPONDÊNCIAS

A análise de correspondência é um método de análise fatorial para variáveis categóricas. A AC, basicamente, converte uma tabela de dados não negativos de duas ou múltiplas entradas em um tipo de representação gráfica em que as linhas e as colunas são simultaneamente representadas em dimensão reduzida, isto é, por pontos no gráfico. Este método permite mostrar como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe. A seguir, são apresentados os resultados da análise de correspondência para cidades e variáveis.

12.1 – Análise de correspondência

| Simple Correspondence Analysis: Área2012; População201; Taxa de Nata; Índice de | | | | | | | |
|---|-----------------|------------|------------|-----------|--------|-------|-------|
| Analysis of Contingency Table | | | | | | | |
| Axis | Inertia | Proportion | Cumulative | Histogram | | | |
| 1 | 0,3101 | 0,9995 | 0,9995 | ***** | | | |
| 2 | 0,0001 | 0,0003 | 0,9999 | | | | |
| 3 | 0,0000 | 0,0001 | 1,0000 | | | | |
| 4 | 0,0000 | 0,0000 | 1,0000 | | | | |
| Total | 0,3102 | | | | | | |
| Row Contributions | | | | | | | |
| ID | Name | Qual | Mass | Inert | Coord | Corr | Contr |
| 1 | Estado SP | 1,000 | 0,970 | 0,016 | -0,071 | 1,000 | 0,016 |
| 2 | Araras | 1,000 | 0,003 | 0,035 | 1,813 | 1,000 | 0,035 |
| 3 | Rio Claro | 0,999 | 0,005 | 0,028 | 1,341 | 0,999 | 0,028 |
| 4 | Leme | 1,000 | 0,003 | 0,023 | 1,697 | 0,999 | 0,023 |
| 5 | Limeira | 0,997 | 0,007 | 0,014 | 0,792 | 0,997 | 0,014 |
| 6 | Cordeirópolis | 1,000 | 0,002 | 0,594 | 9,540 | 1,000 | 0,594 |
| 7 | Mogi Guaçu | 1,000 | 0,004 | 0,022 | 1,383 | 1,000 | 0,022 |
| 8 | Mogi Mirim | 1,000 | 0,003 | 0,068 | 2,832 | 1,000 | 0,068 |
| 9 | Pirassununga | 1,000 | 0,002 | 0,043 | 2,528 | 0,999 | 0,043 |
| 10 | Conchal | 1,000 | 0,001 | 0,066 | 4,649 | 0,998 | 0,066 |
| 11 | Santa Gertrudes | 1,000 | 0,001 | 0,091 | 5,537 | 0,999 | 0,091 |
| Component 2 | | | | | | | |
| ID | Name | Coord | Corr | Contr | | | |
| 1 | Estado SP | 0,001 | 0,000 | 0,004 | | | |
| 2 | Araras | -0,023 | 0,000 | 0,015 | | | |
| 3 | Rio Claro | -0,010 | 0,000 | 0,005 | | | |
| 4 | Leme | -0,064 | 0,001 | 0,095 | | | |
| 5 | Limeira | -0,013 | 0,000 | 0,010 | | | |
| 6 | Cordeirópolis | 0,133 | 0,000 | 0,331 | | | |
| 7 | Mogi Guaçu | -0,029 | 0,000 | 0,028 | | | |
| 8 | Mogi Mirim | -0,015 | 0,000 | 0,006 | | | |
| 9 | Pirassununga | -0,040 | 0,000 | 0,031 | | | |
| 10 | Conchal | -0,198 | 0,002 | 0,340 | | | |
| 11 | Santa Gertrudes | -0,127 | 0,001 | 0,137 | | | |
| Column Contributions | | | | | | | |
| ID | Name | Qual | Mass | Inert | Coord | Corr | Contr |
| 1 | Área2012 | 0,391 | 0,006 | 0,000 | -0,053 | 0,384 | 0,000 |
| 2 | População2011 | 1,000 | 0,988 | 0,006 | -0,044 | 1,000 | 0,006 |



No Symmetric Plot observa-se Cordeirópolis e PIB per Capita um pouco afastada das demais.

8. CONSIDERAÇÕES FINAIS

A pesquisa acima confirma a relação entre as variáveis de taxa de natalidade e analfabetismo com os anos de estudos da população em algumas cidades.

Demonstra que a cidade de Cordeirópolis, apesar de pequena, é consideravelmente mais “rica”, porém Rio Claro (uma das maiores dessa região) apresenta maior índice de anos de estudos e índice de desenvolvimento humanos, se destacando das demais cidades.

De qualquer maneira, todas as cidades da região estão um tanto próximas no que diz respeito às possibilidades de encaminharem alunos para o Centro Universitário Hermínio Ometto.

Como era de se esperar, as variáveis População e Área destoam das demais por conta dos números utilizados no Estado de São Paulo, porém as demais variáveis são bastante próximas dos índices do estado.

Assim, o presente trabalho permite constatar a conquista de patamares elevados de desenvolvimento em uma cidade. É essencial que seja efetivo o desenvolvimento dos índices relacionados à educação, desenvolvimento humano e condições de vida para aumento das

taxas de natalidade. O resultado será a elevação da renda média por emprego, resultando em elevação de PIB per Capita. Por fim, esse processo de desenvolvimento refletirá na elevação dos índices de desenvolvimento humano.