



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE TRÊS RIOS
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS E EXATAS - DCEEX

CAIO OLIVEIRA AZEVEDO

**O POTENCIAL DE DESENVOLVIMENTO DOS MUNICÍPIOS DO ESTADO DO
RIO DE JANEIRO**

TRES RIOS - RJ

2016

CAIO OLIVEIRA AZEVEDO

**O POTENCIAL DE DESENVOLVIMENTO DOS MUNICÍPIOS DO ESTADO DO
RIO DE JANEIRO.**

Monografia de conclusão de curso apresentada ao curso de Ciências
Econômicas da Universidade Federal Rural do Rio de Janeiro,
Instituto Três Rios.

Orientador: José Leonardo Ribeiro Macrini.

Três Rios – RJ

2016

CAIO OLIVEIRA AZEVEDO

Monografia de conclusão de curso apresentada ao curso de Ciências
Econômicas da Universidade Federal Rural do Rio de Janeiro,
Instituto Três Rios.

Orientador: José Leonardo Ribeiro Macrini

Professor Orientador: José Leonardo Ribeiro Macrini

RESUMO:

AZEVEDO, Caio Oliveira. **O potencial de desenvolvimento dos municípios do Estado do Rio de Janeiro**. 31p. Monografia (Graduação em Ciências Econômicas). Instituto Três Rios, Universidade Federal Rural do rio de Janeiro, Três Rios, RJ, 2016.

O objetivo dessa pesquisa é identificar o potencial de desenvolvimento dos 92 (noventa e dois) municípios do Estado do Rio de Janeiro. Desse modo, pretende-se identificar e comparar o potencial de desenvolvimento dos municípios através das técnicas de estatística multivariada. Uma das técnicas utilizadas é conhecida como análise fatorial, que identifica um número reduzido de fatores através de um conjunto de variáveis correlacionadas. A outra técnica é denominada análise de agrupamentos, e visa explorar as similaridades entre os municípios, estabelecendo uma hierarquia entre os mesmos, ao passo que elenca os municípios com maiores ou menores potenciais de desenvolvimento industrial, comercial/serviços, desenvolvimento social e desenvolvimento agrícola. Através desse diagnóstico, busca-se facilitar o processo de tomada de decisão na área de promoção e alocação de investimentos.

PALAVRAS-CHAVE: Análise Fatorial; Análise de Agrupamentos; Estatística Multivariada; Potencial; Desenvolvimento; Rio de Janeiro.

SUMÁRIO

1. INTRODUÇÃO	6
2. A ECONOMIA FLUMINENSE: UM BREVE HISTÓRICO.	8
3. METODOLOGIA.....	12
4. DESCRIÇÃO DOS DADOS E ANÁLISE DOS RESULTADOS	16
4.1 ANÁLISE FATORIAL.....	16
4.2 ANÁLISE DE CONGLOMERADOS	20
5. CONCLUSÃO.....	24
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	26
7. ANEXOS.....	28

1. INTRODUÇÃO

Atualmente, muito se fala em desenvolvimento e em políticas que promovam desenvolvimento. Mas qual a melhor forma de promover o desenvolvimento? Ao longo da história do Brasil, vimos que o setor eleito como propulsor do desenvolvimento foi a indústria. No entanto, nem todos os municípios têm potencial industrial e assim como existe uma tendência de concentração nesse setor.

Desse modo é necessário investir em outros setores, entretanto a questão é: o que? Como? E onde? Aqui, depara-se com a necessidade de conhecer as especificidades de cada município, de forma a entender o seu potencial. E com base nisso facilitar o processo de tomada de decisão na área de promoção e alocação de investimentos.

Eis a razão do objetivo desta pesquisa, identificar o potencial de desenvolvimento dos municípios do Rio de Janeiro. Assim, podem ser propostos investimentos em diversas áreas e setores, estruturando um planejamento de desenvolvimento coerente, que leva em conta as características de cada município, buscando a eficácia das políticas regionais.

Esta pesquisa terá três capítulos. O primeiro capítulo vai debater a economia fluminense, passando um breve histórico que vai contemplar os fatos que arraigaram a desigualdade, principalmente entre regiões, no Estado e de que forma ocorreu o processo de reestruturação econômica e a consequente descentralização produtiva e transformação espacial. O segundo capítulo vai abordar a metodologia, destacando a importância da análise multivariada e também detalhando dois de seus instrumentais – Análise Fatorial e Análise de Conglomerados. E por fim o terceiro capítulo se destina a descrição dos dados e análise dos resultados, através da obtenção de um número de fatores que possam descrever um conjunto de variáveis dos noventa e dois municípios do Rio de Janeiro e assim classificar os municípios de acordo com suas similaridades em grupos distintos e, portanto, diagnosticar possíveis estratégias na alocação de investimentos.

A metodologia para determinar o potencial de desenvolvimento dos 92 municípios do Estado do Rio de Janeiro será a partir da Análise Exploratória de Dados e de Técnicas de Análise Multivariada, relacionando com um estudo bibliográfico.

O Banco de Dados será formado por variáveis econômicas relacionadas ao setor industrial, comercial/serviços, do setor agrícola, financeiro, das exportações, desenvolvimento urbano e social, da educação, saúde e do mercado de trabalho.

As fontes para a composição do Banco de Dados serão o Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto de Pesquisa Econômica Aplicada (IPEA), Secretária

do Tesouro Nacional / Ministério da Fazenda (STN / MF), Banco Central (BACEN), Departamento Nacional de Trânsito (DENATRAN), Cadastro Geral de Empregados e Desempregados (CAGED / MTE), dentre outros.

Após análise inicial sobre a relevância e redundância das informações coletadas, os dados serão padronizados devido às diferentes magnitudes das variáveis utilizadas e, posteriormente, serão feitas as técnicas de Estatística Multivariada já citadas, através do *software SPSS*.

2. A ECONOMIA FLUMINENSE: UM BREVE HISTÓRICO.

Desde o princípio, o Brasil mesmo com seu grande potencial de crescimento sempre se configurou como um país de extremas desigualdades regionais, seja por seu vasto território e riquezas espalhadas em distintas regiões ou também como apontam alguns autores, por exemplo, Furtado (1959) pela forma de colonização (exploratória) e toda a trajetória da formação econômica do Brasil. Alias, foi Furtado um dos primeiros a discutir questões ligadas ao desenvolvimento regional¹.

A este propósito muitos foram os debates relacionados às questões de desenvolvimento regional no Brasil, como forma de amenizar a grande concentração econômica no país e também de romper a histórica disparidade econômica entre os estados da federação e assim solucionar esse entrave ao crescimento e desenvolvimento econômico do país. Para Simões (2004) boa parte desses debates envolvem principalmente questões locacional e fatores aglomerativos das atividades econômicas, tendo nas atividades industriais o papel importante de polo gerador de emprego e renda de forma a superar ou talvez compensar as desigualdades econômicas, sociais e regionais. Myrdal (1960, p.41) alerta para o que ele denomina “princípio da causação circular cumulativa”:

A decisão de localizar uma indústria em determinada comunidade, por exemplo, impulsiona o seu desenvolvimento geral. Proporcionam-se possibilidades de emprego e renda elevadas àqueles que se encontram desempregados ou com empregos de baixo salário.[...].O estabelecimento de um novo negócio ou a ampliação de um existente expande o mercado para outros [...]. Os lucros em elevação aumentam as poupanças, ao mesmo tempo em que elevam, ainda mais, os investimentos; tal fato aumenta, ainda uma vez, a demanda e o nível de lucros. O processo de expansão cria economias externas favoráveis à sua continuidade.

Nesse sentido, se desde a expansão cafeeira o estado de São Paulo já havia despontado como grande produtor de café, dando início a um processo de concentração econômica no Sudeste, foi justamente pós-crise de 29 que as desigualdades passaram a ser a tônica no Brasil. Isso se deu principalmente pelo início do processo de industrialização por substituição de importação (ISI), que ocorreu, especialmente, em torno da região metropolitana de São Paulo². Conforme São Paulo crescia industrialmente, criava um “atraso relativo” a alguns estados – como o Rio de Janeiro - pois acentuava suas diferenças, como visto no gráfico 2.1.

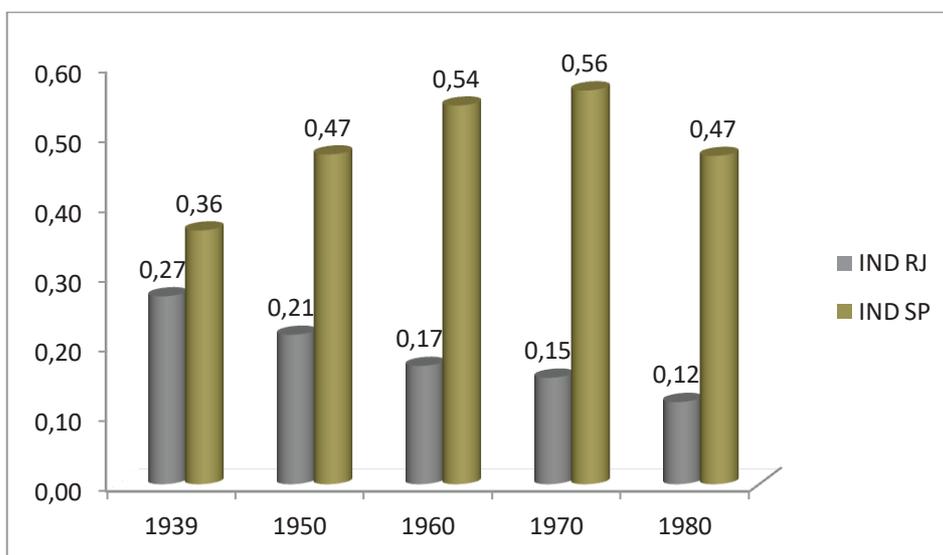
¹ Ver Furtado (1959)

² Ver Cano (1998)

Para Silva (2004) era de se esperar que a economia fluminense seguisse essa trajetória de perda de participação relativa³ por ser uma economia pautada pelo setor público, enquanto o perfil da economia paulista era industrial e que muito foi beneficiado pelo modelo de desenvolvimento da economia brasileira pós crise de 29, sendo o centro polarizador de recursos humanos e produtivos do país. Lessa (2000, p. 237-238) retrata bem esse período:

As décadas de 1920 a 1960 foram de prosperidade e de acumulação de prestígio no Rio de Janeiro. A cidade desdobrou-se em novos comportamentos e dimensões.[...] O Rio urbanizou-se em sintonia com esses novos tempos. Cabe sublinhar que foi sendo secundarizado, em termo de produção industrial, em relação a São Paulo. Desde a Primeira Guerra Mundial, São Paulo lidera a produção industrial e, apesar de crescer, o Rio vê a distância relativa das respectivas bases industriais ser ampliada, para não lembrar a espantosa diferença no campo agrícola.

Gráfico 2.1: Participação do setor industrial no PIB (em %)



FONTE: IPEA / Elaboração própria.

Diante de tal cenário, várias foram às medidas tomadas pelo governo a fim de realizar uma integração regional no Brasil, como foi o exemplo da transferência da capital federal da cidade do Rio de Janeiro para Brasília ocorrida em 1960, acarretando na descentralização política e administrativa, e posteriormente o II PND como ação para reduzir as desigualdades regionais, como está explícito no próprio Plano (1975), “na etapa que se inicia, o Governo define a estratégia de integração nacional dentro de suas preocupações principais: o melhor

³ Alguns autores utilizam o termo “esvaziamento econômico” para explicar a perda relativa na participação da economia fluminense no PIB. Porém para Silva(2012:43) o termo “esvaziamento” traz, a princípio, a ideia de perdas reais e redução/destruição de capacidade produtiva. Desde essa constatação, fica claro que o termo não é dos mais precisos para retratar a dinâmica econômica fluminense ao longo da industrialização brasileira, tendo em vista que sua trajetória foi claramente expansiva.

equilíbrio econômico-político entre as diferentes regiões seja dentro do Centro-Sul, seja das outras macrorregiões em relação ao Centro-Sul”.

É de fundamental importância analisar que o fato da capital fluminense ter sido a sede do governo federal durante quase dois séculos fez com que a administração pública tivesse um peso muito grande sobre a geração de valor no estado do Rio de Janeiro. Para Silva (2004, p.40-41) :

O setor governo adquiriu fundamental importância para a economia fluminense durante os quase dois séculos que a região abrigou a capital brasileira. Importância, diga-se de passagem, não observada em outro território nacional. Ademais funcionar como “componente autônoma” da renda estadual, este setor foi o principal responsável pela formação da estrutura urbana regional.[...]. Com a transferência da capital para Brasília, o Rio de Janeiro sofreu perdas de renda substanciais, tendo em vista o gasto público e a massa salarial que ali não mais se realizara.

Em suma, a cidade do Rio de Janeiro foi capital administrativa do Brasil entre os anos de 1763 e 1960, o que na época consagrou a cidade nacional e internacionalmente – por exemplo, foi o grande centro financeiro do Brasil⁴ - anos mais tarde se tornou um problema estrutural, já que durante quase dois séculos o interior do estado esteve abandonado.

Tal problema foi se intensificando décadas depois da transferência da Capital, visto que a criação de um novo estado - Guanabara⁵ - como meio de amenizar os impactos e as perdas obtidas pela cidade devido à transferência e posteriormente a fusão entre os Estados da Guanabara e do Rio de Janeiro (que tinha Niterói por capital), acabou por criar um aumento sucessivo na concentração econômica e populacional entre a cidade do Rio de Janeiro e cidades adjacentes, o que seria mais tarde conhecida por Região do Grande Rio.

A ideia que orienta a fusão não é de juntar duas coisas inviáveis para se alcançar uma terceira viável. O que se pretende é reunir duas economias e potencialidades perfeitamente viáveis para a criação de um polo de desenvolvimento de grandes dimensões, como o de São Paulo, dentro da política de diversificar os polos industriais, tendo em vista o objetivo de diminuir os desequilíbrios regionais (BRASILEIRO, 1979, apud MOTTA, 2001, p. 24).

Nesse sentido, vale salientar que mudanças ocorridas no sistema capitalista, como a reestruturação produtiva e organizacional e a conseqüente descentralização, fez com que algumas das regiões do Estado fossem contempladas com investimentos públicos ou privados,

⁴ O setor financeiro brasileiro teve origem no Rio de Janeiro, começando com a fundação do Banco do Brasil, primeiro em 1808 e depois em 1852. Foi também no Rio de Janeiro que se fundou o Banco Comercial do Rio de Janeiro, em 1838, posteriormente fundido com o Banco do Brasil e depois a Caixa Econômica Federal, em 1861. Da mesma forma, foi no Rio de Janeiro que surgiu a primeira Bolsa de Valores do país, em 1845, assim como o Banco Nacional de Desenvolvimento em 1952 e posteriormente o Banco Nacional de Habitação. (PINHEIRO, 2012, p. 179-180)

⁵ Transformada em cidade-estado, ou seja, sem municípios, a Guanabara conservou a maior parte das funções de principal centro político do país, tornando-se o que se pode chamar de estado-capital.

como é o caso das Regiões Norte e Médio Paraíba, a primeira sendo um polo da indústria petrolífera e a segunda se destacando como polo industrial. A partir disso, o que se notou foi uma reestruturação econômica e transformação na espacialização do Estado do Rio de Janeiro.

Nos últimos anos o Rio de Janeiro começou, por assim dizer, a “dar a volta por cima”. O desenvolvimento do setor de petróleo e gás tem sido, sem dúvida, um motor importante dessa virada, seja pela ótica das contas públicas, seja pela criação de uma vocação econômica que tem atraído empresas, tecnologia e criado bons empregos. (PINHEIRO, 2012, p. 19)

Acontece apenas, que tais transformações não foram suficientes para solucionar anos de descaso, visto que regiões como a Noroeste e Centro Sul continuaram a ser carentes de investimentos, possuindo uma economia estagnada, que em boa parte ainda dependeriam de um setor comercial restrito.

Isso se perdurou até meados dos anos 2000, quando o governo do estado do Rio de Janeiro sancionou a lei 4.533/2005 com o objetivo de estimular o desenvolvimento de alguns municípios fluminenses, diminuindo as disparidades regionais, tornando-os mais atraentes para as empresas, através de um regime tributário especial para alguns municípios de baixa atividade econômica que reduziu de 19% para 2% a base de cálculo da alíquota do Imposto sobre Operações Relativas à Circulação de Mercadorias e Prestação de Serviços (ICMS), o que conseqüentemente propiciaria o aumento de emprego e renda, alterando a dinâmica econômica dos municípios contemplados pela lei. Essa medida surtiu efeito em vários municípios, porém se mostrou insuficiente para outros que não possuíam de um sistema logístico, mão de obra qualificada e infraestrutura.

As desigualdades constituem desafios para o Estado, sendo necessária a promoção de um desenvolvimento que atue em duas vertentes complementares. De um lado, uma sustentação e/ou dinamização das regiões com bom desempenho econômico e social; de outro, uma regeneração e/ou indução do processo de desenvolvimento das regiões com carências econômicas e sociais. (PEROBELLI, 1999, p.123)

Neste sentido, para as políticas de planejamento regional e urbano, torna-se fundamental a identificação das potencialidades de desenvolvimento dos municípios. E através dessa identificação, facilitar o processo de tomada de decisão na área de promoção e alocação de investimentos.

3. METODOLOGIA

De acordo com Ferreira (1989, p. 549):

A identificação de polos de desenvolvimento a serem ativados visando uma dada estratégia de desenvolvimento da estrutura do espaço econômico ou do sistema de cidades e distribuição das atividades, ou funções econômico-sociais, requer uma pesquisa da estrutura de seu espaço econômico atual, mostrando as interdependências diretas e indiretas entre as cidades de um subespaço e, quando necessário, suas relações com outras cidades de espaço maior. Torna-se recomendável, portanto, identificarem-se as cidades dominantes e suas áreas de influência a diversos níveis ou escalões de interdependência: regional, sub-regional, zonal, subzonal e local.

É importante observar que raríssimos são os ramos do conhecimento e as atividades humanas que podem dispensar o apoio de técnicas estatísticas em seu desenvolvimento. Um olhar mais acurado em torno de quase todos os fenômenos que nos cercam nos remete à conclusão de que tais técnicas estão participando cada vez mais do nosso cotidiano.

Diante tal constatação, esta pesquisa consiste no uso de técnicas de análise multivariada⁶ para tentar classificar o potencial de desenvolvimento econômico de um grupo de municípios.

E para isso, um dos métodos multivariados a serem utilizados será o de análise fatorial, que é uma técnica estatística usada no intuito de identificar um número reduzido de padrões de características latentes (fatores) para um conjunto de variáveis (observadas) correlacionadas entre si, ou seja, identificar dimensões de variabilidade comuns existentes, mas que não são observáveis diretamente, onde tais dimensões são denominadas fatores que, segundo Andrade (1989), permite identificar fatores que podem ser interpretados como estágio de desenvolvimento econômico.

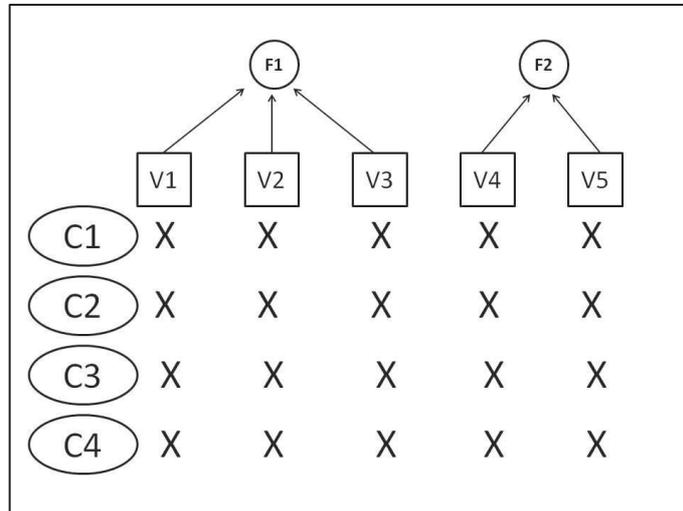
The essential purpose of factor analysis is to describe, if possible, the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors. Basically, the factor model is motivated by the following argument: Suppose variables can be grouped by their correlations. That is, suppose all variables within a particular group are highly correlated among themselves, but have relatively small correlations with variables in a different group. Then it is conceivable that each group of variables represents a single underlying construct, or factor, that is responsible for the observed correlations. (JOHNSON E WICHERN, 1988, p. 481)

A estrutura mais comum de Análise Fatorial está demonstrada na Figura 1, onde as colunas representam as variáveis – V_n (características analisadas) –, as linhas são os casos –

⁶ Multivariada no sentido que serão analisadas mais de duas variáveis simultaneamente.

C_n (empresas, cidades, indivíduos, etc) – e as células são os valores assumidos pelas variáveis em cada caso (x).

Figura 3.1: Estrutura da Análise Fatorial



FONTE: Elaboração própria. (BEZERRA, 2009, p. 82)

Então com base na figura anterior, pode-se de uma maneira um tanto quanto simplificada ilustrar a ideia básica do modelo fatorial. Nesse exemplo anterior, o modelo fatorial identificou 2 (dois) fatores para as 5 (cinco) variáveis. Sendo assim cada uma dessas variáveis serão explicadas pelo fator comum (F) e pelo respectivo fator específico ou erro(ϵ). É importante frisar, que os coeficientes dos fatores (L) são as cargas fatoriais, e representam o grau de associação entre a variável e o fator.

$$V_1 = L_{11}F_1 + L_{12}F_2 + \epsilon_1$$

$$V_2 = L_{21}F_1 + L_{22}F_2 + \epsilon_2$$

$$V_3 = L_{31}F_1 + L_{32}F_2 + \epsilon_3$$

$$V_4 = L_{41}F_1 + L_{42}F_2 + \epsilon_4$$

$$V_5 = L_{51}F_1 + L_{52}F_2 + \epsilon_5$$

Para Haddad (1989) as etapas a serem seguidas para aplicação da análise fatorial são as seguintes:

- i. montagem da matriz de correlação e a partir disso adequação ou não do modelo fatorial aos dados;
- ii. extração dos fatores iniciais e por consequência as cargas fatoriais, a esse respeito, é importante salientar que existem alguns métodos para a extração dos

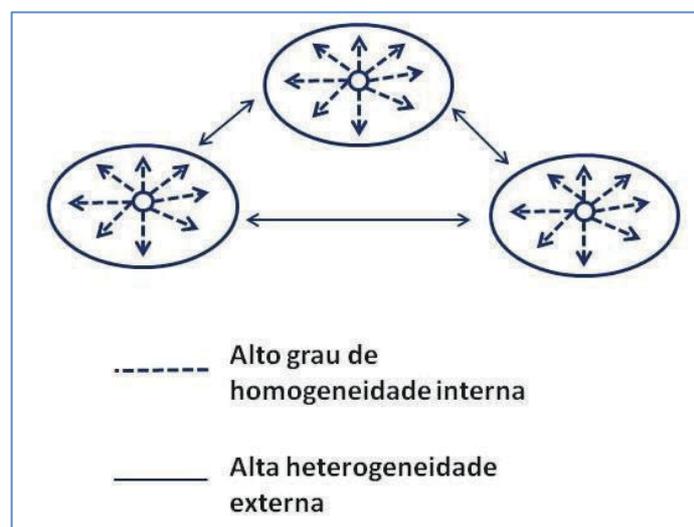
fatores iniciais, sendo os mais conhecidos⁷ o Método de Componentes Principais e o Método da Máxima Verossimilhança.

- iii. rotação dos fatores⁸ (quando necessário) no intuito de definir com mais clareza o a relação entre as variáveis e os fatores; e
- iv. cálculo dos escores fatoriais.

Sendo assim, a partir da obtenção dos fatores, os escores (valores numéricos) de cada elemento amostral serão obtidos, no intuito de serem utilizados no próximo método a ser aplicado que será o de análise de conglomerados.

A análise de conglomerados - também conhecida como análise de *cluster* - é uma técnica multivariada que tem como objetivo dividir os elementos da amostra (ou população) em grupos, de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si com respeito às variáveis (características) que neles forem medidas, e os elementos em grupos diferentes seja heterogêneos em relação a essas mesmas características, conforme a figura abaixo.

Figura 3.2: Agrupamentos com alta homogeneidade interna e alta heterogeneidade externa.



FONTE: POHLMANN, 2009, p. 324

⁷ Johnson e Wichern (1988) consideram os Métodos de Componentes Principais e o da Máxima Verossimilhança como os mais recomendados para análise fatorial. A respeito desses métodos, o que os diferencia é que o Método de Componentes Principais não exige a suposição de distribuição normal, e por isso é o mais utilizado, enquanto o Método da Máxima verossimilhança supõe distribuição normal das variáveis.

⁸ Existem dois tipos de rotação, ortogonais (mantém a perpendicularidade dos fatores) e oblíquas (permite a correlação entre os fatores comuns). E dentro da rotação ortogonal existem alguns métodos, como o Varimax, Quartimax e Orthomax.

Em linhas gerais, a análise de conglomerados consiste em uma técnica de procedimentos exploratórios que busca descobrir agrupamentos naturais de variáveis a partir dos dados observados, agrupando as mesmas com base na similaridade ou dissimilaridades.

A semelhança entre os indivíduos pode ser medida de acordo com suas similaridades ou, em sentido oposto, por suas dissimilaridades, chamadas de distâncias no espaço das variáveis, tais como, distância euclidiana, distância generalizada, distância de Minkowsky e métrica de Canberra.

Para Pohlmann (2009, p.325):

A Análise de Agrupamentos, ou *Cluster Analysis*, é uma das técnicas de análise multivariada cujo propósito primário é reunir objetos, baseando-se nas características dos mesmos. Ela classifica objetos (ex. municípios) segundo aquilo que cada elemento tem de similar em relação aos outros pertencentes a determinado grupo, considerando, é claro, um critério de seleção predeterminado.

Com relação as técnicas de agrupamentos, se dividem em Hierárquicas – que buscam identificar agrupamentos e o provável número de partições - e não hierárquicas, que já partem pré-estabelecido o número de partições. As técnicas de agrupamento hierárquicas mais conhecidas e usuais, são o *single linkage*, *complete linkage* e *average linkage*. Enquanto a técnica de agrupamento não hierárquica mais conhecida é a *k-means*.

Por fim, os critérios para a partição final, podem ser elencados dentre:

- i. Nível de fusão
- ii. Nível de similaridade
- iii. Coeficiente R^2
- iv. Estatística Pseudo F
- v. Correlação semiparcial
- vi. Estatística Pseudo T^2

Portanto pretende-se com o uso desses métodos, estabelecer uma hierarquia entre os municípios do Estado do Rio de Janeiro, assim como definir os municípios com maiores ou menores potenciais de desenvolvimento industrial, financeiro e social. E através desse diagnóstico facilitar o processo de tomada de decisão na área de promoção e alocação de investimentos e identificando o potencial de desenvolvimento dos noventa e dois municípios do Estado do Rio de Janeiro.

4. DESCRIÇÃO DOS DADOS E ANÁLISE DOS RESULTADOS

4.1 ANÁLISE FATORIAL

Para a realização do presente estudo, inicialmente foram selecionadas 18 variáveis (Anexo I) para os 92 municípios do Estado do Rio de Janeiro. Através dessa técnica, das variáveis selecionadas inicialmente constatou-se que sete variáveis apresentavam comunalidades⁹ abaixo de 50% e ou então não estavam correlacionadas significativamente, assim tais variáveis foram retiradas da análise. As 11 variáveis restantes, listadas a seguir apresentaram as seguintes comunalidades mostradas na tabela 4.1.

X₁: Valor adicionado bruto a preços correntes da agropecuária (Mil Reais) / Valor adicionado bruto a preços correntes total (Mil Reais);

X₂: Valor adicionado bruto a preços correntes da indústria (Mil Reais) / Valor adicionado bruto a preços correntes total (Mil Reais);

X₅: taxa de alfabetização;

X₇: população residente urbana / população residente total;

X₈: salário médio mensal;

X₉: operações de crédito / população economicamente ativa;

X₁₃: PIB per capita;

X₁₅: proporção de domicílios particulares com microcomputado e acesso a internet;

X₁₆: proporção de domicílios com rendimento nominal mensal de mais de 10 a 20 salários mínimos;

X₁₇: proporção de domicílios particulares permanentes do tipo apartamento;

X₁₈: número de famílias com bolsa família / população residente total.

⁹ Comunalidades representam o percentual de explicação que uma variável obteve pela Análise Fatorial, ou seja, quanto todos os fatores juntos são capazes de explicar uma variável. Quanto mais próximo de 1 estiverem as comunalidades, maior é o poder de explicação dos fatores.

Tabela 4.1: Comunalidades

Variable	Comunalities	Variable	Comunalities
x ₁	0.7783	x ₁₃	0.7688
x ₂	0.7518	x ₁₅	0.9018
x ₅	0.8962	x ₁₆	0.7837
x ₇	0.8223	x ₁₇	0.8615
x ₈	0.5792	x ₁₈	0.668
x ₉	0.7719		

É importante observar que as variáveis foram padronizadas através da conversão de cada variável para escores padrões (Z), isto é, variáveis com média zero e variância um, com o intuito de simplificar a pesquisa, eliminando a distorção introduzida pelas diferentes escalas de vários atributos usadas na análise, evitando assim enganos que possa surgir se analisando as variáveis em termos absolutos. Uma última observação é que os dados dos municípios foram ponderados pelas suas respectivas populações residentes ou pela quantidade de domicílios total, de forma a evitar distorções nas análises de municípios mais populosos, como é caso do município do Rio de Janeiro.

Alguns autores sugerem que, para que um modelo de análise fatorial possa ser adequadamente ajustado aos dados, é necessário que a matriz de correlação inversa seja próxima da matriz diagonal. Uma medida de adequacidade que é fundamental nesse princípio é o coeficiente KMO, que é classificado conforme tabela abaixo.

Tabela 4.2: Classificação do Critério de Kaiser Meyer Olkin (KMO)

KMO	Adequação da Análise Fatorial
0.90 – 1.00	Muito boa
0.80 – 0.90	Boa
0.70 – 0.80	Média
0.60 – 0.70	Razoável
0.50 – 0.60	Má
< 0.50	Inaceitável

Fonte: Adaptado de Mingoti (2005)

O teste de Kaiser-Meyer-Olkin ($KMO = 0.839$) indicou ser boa a execução da técnica de Análise Fatorial, com relação as variáveis analisadas. O teste de esfericidade de Bartlett

apresenta um *p-value* menor que 0.1%, concluindo que as variáveis estão correlacionadas significativamente.

Tabela 4.3: Teste *KMO*

Variable	Kmo	Variable	Kmo
x ₁	0.8494	x ₁₃	0.6814
x ₂	0.6625	x ₁₅	0.8624
x ₅	0.8242	x ₁₆	0.9019
x ₇	0.8636	x ₁₇	0.8138
x ₈	0.8405	x ₁₈	0.9323
x ₉	0.779		
Overall		0.8396	

A partir dessas 11 variáveis, foram estabelecidos - pelo método de componentes principais¹⁰ - 3 (três) fatores, que correspondem a aproximadamente 78% de toda variabilidade existente entre as variáveis utilizadas.

Tabela 4.4: Total de variância explicada pelas 11 variáveis

Factor analysis/correlation	Number of obs	=	92
Method: principal-component factors	Retained factors	=	3
Rotation: (unrotated)	Number of params	=	30

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	5.27477	3.46937	0.4795	0.4795
Factor2	1.8054	0.30222	0.1641	0.6437
Factor3	1.50318	0.89586	0.1367	0.7803
Factor4	0.60732	0.12734	0.0552	0.8355
Factor5	0.47998	0.07738	0.0436	0.8792
Factor6	0.4026	0.10013	0.0366	0.9158
Factor7	0.30247	0.07266	0.0275	0.9432
Factor8	0.2298	0.04592	0.0209	0.9641

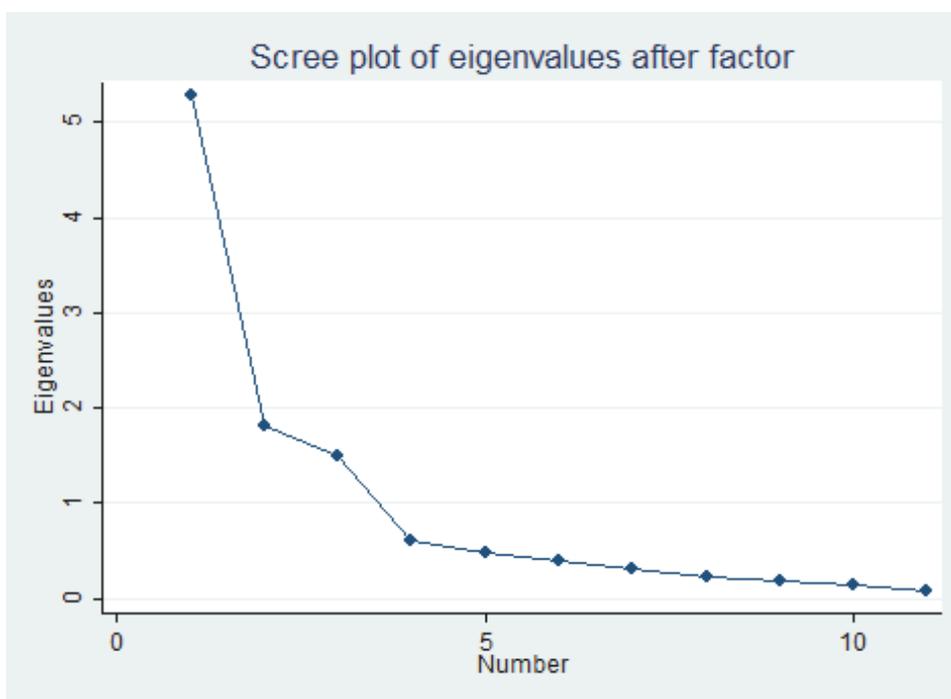
¹⁰ Ressalta-se que os fatores obtidos pelo método de máxima verossimilhança foram os mesmos obtidos pelo método de componentes principais, sendo assim não há perda de informação utilizando um ou outro método, dado que o intuito é a obtenção dos fatores.

Factor9	0.18388	0.05121	0.0167	0.9809
Factor10	0.13267	0.05475	0.0121	0.9929
Factor11	0.07792	.	0.0071	1

LR test: independent vs. saturated: $\chi^2(55) = 736.14$ Prob> $\chi^2 = 0.0000$

O *Scree Plot* (Figura 4.1) confirma a seleção do número de fatores, destacando que de acordo com esse critério o número de fatores também poderia ser 4 (quatro), porém optou-se trabalhar com o número de fatores igual ao número de autovalores maiores ou iguais a 1, como mostra a tabela 4.4.

Figura 4.1: Critério do gráfico de declive



As variáveis que compõe cada fator foram estabelecidas utilizando a rotação *Varimax*, visando encontrar variáveis altamente correlacionadas com o fator e de baixa correlação com outro e ainda assim não se concentrar com cargas altas apenas no 1º fator como é o caso do *Quartimax*. A composição de cada fator ficou assim estabelecida:

- Fator 1 - Nível de Urbanização e desenvolvimento social: Valor adicionado da agropecuária, taxa de alfabetização; população urbana, cobertura do acesso a internet pelos domicílios e proporção de famílias do Programa Bolsa Família.
- Fator 2 - Nível de atividade econômica: operações de crédito, rendimento mensal de 10 a 20 salários e proporção de domicílios do tipo apartamento.

- Fator 3 – Nível de Industrialização: valor adicionado da indústria, salário médio mensal e PIB per capita.

A tabela 4.5 abaixo discrimina as variáveis por fatores, assim como suas respectivas cargas fatorais.

Tabela 4.5: Matriz após a rotação dos fatores

Variable	Factor1	Factor2	Factor3	Variable	Factor1	Factor2	Factor3
x1	-0.8493			x13			0.8763
x2			0.8398	x15	0.6882	0.6472	
x5	0.8959			x16	0.4046	0.7637	
x7	0.8996			x17		0.9021	
x8		0.4443	0.6023	x18	-0.632	-0.4937	
x9		0.8785					

4.2 ANÁLISE DE CONGLOMERADOS

O próximo passo foi realizar outra técnica de análise multivariada denominada análise de agrupamentos, também conhecida como análise de conglomerados ou *cluster*, tem como objetivo dividir os elementos da amostra, ou população, em grupos de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si com respeito às variáveis (características) que neles forem medidas, e os elementos em grupos diferentes sejam heterogêneos em relação a essas mesmas características.

Para a construção dos clusters foi utilizado dois métodos: não hierárquicos e hierárquicos. A análise se iniciou pelo método hierárquico, na medida que o número de grupos era desconhecido. Optou-se pela utilização de três técnicas: *single linkage*, *average linkage* e *Ward's Method*. E por fim, para determinar a partição final, além do dendograma foram utilizados dois critérios: Estatística Pseudo-F e Pseudo T², tais valores segue abaixo:

Tabela 4.6: Estatística Pseudo-F (*Calinski/Harabasz*)

<i>Number of clusters</i>	<i>Single Linkage</i>	<i>Average Linkage</i>	<i>Wards</i>
2	23.01	23.01	27.49
3	26.65	26.65	35.48
4	21.45	25.21	48.31
5	16.23	25.96	48.87
6	13.2	21.43	52.7
7	15.46	37.91	52.79
8	13.8	34.04	54.88
9	13	30.53	56.62
10	11.47	27.52	56.03
11	11.42	35.94	54.71
12	10.29	42.25	54.27
13	9.57	39.61	53.98
14	9.11	43.47	53.67
15	8.65	45.05	53.65

Tabela 4.7: Estatística Pseudo-T²

<i>Number of Clusters</i>	<i>Single Linkage</i>		<i>Average Linkage</i>		<i>Wards</i>	
	<i>Duda/Hart</i>		<i>Duda/Hart</i>		<i>Duda/Hart</i>	
	<i>Je(2)/Je(1)</i>	<i>Pseudo-T squared</i>	<i>Je(2)/Je(1)</i>	<i>Pseudo-T squared</i>	<i>Je(2)/Je(1)</i>	<i>Pseudo-T squared</i>
1	0.7964	23.01	0.7964	23.01	0.766	27.49
2	0.7839	24.25	0.7839	24.25	0.5604	60.41
3	0.922	7.19	0.8569	14.2	0.3824	17.77
4	0	.	0.8438	15.36	0.339	17.55
5	0.1372	6.29	0.3108	4.44	0.6951	24.57
6	0.8448	15.43	0.594	53.99	0.5676	25.13
7	0.9726	2.31	0.8428	4.85	0.4989	19.08
8	0.9542	3.89	0.1372	6.29	0.6001	13.99
9	0	.	0	.	0.4586	7.08
10	0.9371	5.37	0.6418	28.46	0.443	13.83
11	0	.	0.6104	23.61	0.2309	13.33

12	0.9842	1.26	0.0945	9.58	0.5633	6.98
13	0.9742	2.04	0.6529	13.29	0.6828	9.29
14	0.9786	1.66	0.5241	12.71	0.4039	7.38
15	0.9742	1.98	0.6495	14.03	0.1372	6.29

Pela análise das estatísticas se conclui que tanto pela técnica de *single linkage e average linkage* o número de clusters seria 7, enquanto pelo método de *Ward* foi de 9 clusters. Assim optou-se pela utilização da Estatística Pseudo F como critério, e portanto, os municípios foram classificados em 9 (quatro) *clusters*, de acordo com os escores fatoriais obtidos pelo Método de Regressão na Análise Fatorial¹¹ para os 92 municípios do Estado do Rio de Janeiro.

Através dessa constatação, foi também utilizado o método não hierárquico (*K-means*), para 9 partições obtendo sua respectiva Estatística Pseudo-F, que foi de 57.15. Portanto, o método de *k-means* possui um melhor resultado e se procedeu com a utilização do modelo não hierárquico. As classificações seguem abaixo:

Tabela 4.8: Número de municípios por *clusters*

Cluster	Número de municípios
1	8
2	20
3	16
4	16
5	6
6	3
7	2
8	13
9	8
Total	92

¹¹ Obtidos pelo Método de Regressão.

Tabela 4.9: Classificação dos municípios por *clusters*.

<i>Cluster 1:</i> Angra dos Reis, Cabo Frio, Campos dos Goytacazes, Macaé, Piraí, Resende, Rio das Ostras e Volta Redonda.
<i>Cluster 2:</i> Aperibé, Araruama, Arraial do Cabo, Belford Roxo, Conceição de Macabu, Engenheiro Paulo de Frontin, Guapimirim, Iguaba Grande, Itaboraí, Japeri, Macuco, Magé, Mendes, Mesquita, Nova Iguaçu, Pinheiral, São Gonçalo, São João de Meriti, Saquarema e Tanguá.
<i>Cluster 3:</i> Barra do Piraí, Barra Mansa, Cordeiro, Itaperuna, Mangaratiba, Maricá, Miguel Pereira, Nilópolis, Nova Friburgo, Paracambi, Paraíba do Sul, Petrópolis, São Pedro da Aldeia, Teresópolis, Três Rios e Valença.
<i>Cluster 4:</i> Bom Jardim, Bom Jesus do Itabapoana, Cambuci, Cantagalo, Casimiro de Abreu, Duas Barras, Itaocara, Miracema, Natividade, Porciúncula, Rio Bonito, Rio Claro, Santo Antônio de Pádua, São Fidélis, Sapucaia e Vassouras.
<i>Cluster 5:</i> São Francisco de Itabapoana, São José de Ubá, São Sebastião do Alto, Sumidouro, Trajano de Moraes e Varre-Sai.
<i>Cluster 6:</i> Porto Real, Quissamã e São João da Barra.
<i>Cluster 7:</i> Rio de Janeiro e Niterói.
<i>Cluster 8:</i> Areal, Armação de Búzios, Cachoeiras de Macacu, Carapebus, Cardoso Moreira, Carmo, Duque de Caxias, Itaguaí, Itatiaia, Parati, Quatis, Queimados e Seropédica.
<i>Cluster 9:</i> Comendador Levy Gasparian, Italva, Laje do Muriaé, Paty do Alferes, Rio das Flores, Santa Maria Madalena, São José do Vale do Rio Preto e Silva Jardim.

5. CONCLUSÃO

Neste artigo foram utilizados dois métodos da Estatística Multivariada: a Análise Fatorial, que buscou condensar toda a informação contida nas diversas variáveis, de forma a obter um conjunto menor de variáveis estatísticas (fatores) e a Análise de Agrupamentos que teve por intuito reunir os municípios em grupos, de maneira que os municípios com características mais parecidas estejam em um mesmo grupo.

Tais técnicas foram aplicadas num conjunto de 11 variáveis para caracterizar o potencial de desenvolvimento dos municípios do Estado do Rio de Janeiro. Na análise fatorial, o objetivo principal foi alcançado. Apresentou-se a redução do número de variáveis analisadas, passando de 11 para 3 fatores (Fator 1 – Nível de Urbanização e desenvolvimento social; Fator 2 – Nível de atividade econômica; Fator 3 – Nível de industrialização) implicando em uma condensação no contingente de variáveis.

Na análise de *cluster* foram utilizados os 3 (três) fatores obtidos anteriormente para formar os *clusters*, obtendo-se um total de 9 (nove) *clusters*.

Os resultados foram bastante interessantes e coerentes com toda a revisão de literatura realizada, já que *clusters* como o de número 7, que contempla somente dois municípios, o Rio de Janeiro e Niterói, reforça toda o ponto de visto já levantado da importância da cidade do Rio de Janeiro como capital e os grandes investimentos feitos enquanto capital federal e posteriormente como cidade-estado e Niterói que como destacado, se tornou durante alguns anos capital do Estado do Rio de Janeiro. Assim como o *cluster* 1, corroborou para demonstrar a eficácia dos investimentos públicos e privados em regiões do Estado como o caso das Regiões Norte e Médio Paraíba, onde boa parte dos municípios que compõe esse *cluster* são oriundos dessas regiões.

Acontece apenas, que como já relatado algumas regiões ainda não passaram por transformações suficientes para solucionar anos de esquecimento político e essa conjectura acaba sendo confirmada pelo *cluster* 5, que possui os municípios com os piores índices de desenvolvimento humano do Estado, todos localizados na Região Noroeste. O *cluster* 9, é composto em grande parte de municípios da região Noroeste e Centro Sul. Portanto, empiricamente podemos afirmar que esses *clusters* possuem os menores potenciais de desenvolvimento, onde boa parte da atividade econômica desses municípios é originada pelo setor comercial/serviços, diminuindo a participação da indústria na formação do valor bruto dessas economias. O mesmo vale para o *cluster* 4, também composto por municípios do Noroeste e Centro Sul fluminense.

Neste sentido, os resultados obtidos na pesquisa demonstraram ainda falta de dinamismo econômico no Estado do Rio de Janeiro, onde apenas 2 municípios conseguem de certa forma se diferenciar totalmente dos demais (Rio de Janeiro e Niterói).

6. REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, Thompson A., Métodos Estatísticos e Econométricos Aplicados à Análise Regional. In: HADDAD, Paulo Roberto. (org.). **Economia Regional: teoria e métodos de análise.** Fortaleza: BNB, 1989.

BEZERRA, Francisco Antônio. Análise Fatorial. In: CORRAR, Luiz J.; PAULO, Edilson; DIAS FILHO, José Maria. (coord.). **Análise Multivariada: para os curso de Administração, Ciências Contábeis e Economia.** 1. ed. São Paulo: Atlas, 2009.

CANO, Wilson. **Raízes da concentração industrial em São Paulo.** Campinas: UNICAMP/IE, 1998.

CRUZ, Bruno Oliveira. et all. **Economia regional e urbana: teorias e métodos com ênfase no Brasil.** Brasília: Ipea, 2011.

FERREIRA, Carlos Mauricio de C., Métodos de Regionalização. In: HADDAD, Paulo Roberto. (org.). **Economia Regional: teoria e métodos de análise.** Fortaleza: BNB, 1989.

HADDAD, Paulo Roberto. (org.). **Economia Regional: teoria e métodos de análise.** Fortaleza: BNB, 1989.

JOHNSON, R. A. & WICHERN, W. W. **Applied multivariate statistical analysis.** 4. Ed. : New Jersey. Prentice Hall. 1998.

LESSA, Carlos. **O Rio de todos os Brasis: uma reflexão em busca de auto-estima.** Rio de Janeiro: Record, 2000.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada.** Belo Horizonte: Editora UFMG, 2005.

MOTTA, Marly Silva da. A fusão da Guanabara com o Estado do Rio: desafios e desencantos. In FREIRE, Américo; SARMENTO, Carlos Eduardo; MOTTA, Marly Silva da; (org.). **Um estado em questão: os 25 anos do Rio de Janeiro**. Rio de Janeiro: FGV, 2001.

MYRDAL, G. Tendências para desigualdades econômicas regionais em um país. In: TEORIA ECONÔMICA E REGIÕES SUBDESENVOLVIDAS. Rio de Janeiro: ISEB, 1960, p.39-56.

PEROBELLI, Fernando Salgueiro. et all. Planejamento regional e potenciais de desenvolvimento dos municípios de Minas Gerais na região em torno de Juiz de Fora: uma aplicação de análise fatorial. **Revista Nova Economia**. Belo Horizonte: v. 9, n.1, p. 121-150. Jul./1999.

PINHEIRO, Armando Castelar. VELOSO, Fernando. **Rio de Janeiro: um estado em transição**. 1. ed. Rio de Janeiro: FGV, 2012.

POHLMANN, M. C. Análise de Conglomerado. In: CORRAR, Luiz J.; PAULO, Edilson; DIAS FILHO, José Maria. (coord.). **Análise Multivariada: para os curso de Administração, Ciências Contábeis e Economia**. 1. ed. São Paulo: Atlas, 2009.

SIMÕES, Rodrigo. et all. **Aglomerados industriais e desenvolvimento socioeconômico: uma análise multivariada para Minas Gerais**. Ensaio FEE, Porto Alegre, v.25, n.1, p.203-232, abr.2004. Disponível em:
<<http://revistas.fee.tche.br./index.php/ensaio/article/view/2057/2439>>

SILVA, Robson Dias. Indústria e desenvolvimento regional no Rio de Janeiro. 1.ed. Rio de Janeiro: FGV, 2012.

7. ANEXOS

ANEXO I

X₁: Valor adicionado bruto a preços correntes da agropecuária (Mil Reais) / Valor adicionado bruto a preços correntes total (Mil Reais);

X₂: Valor adicionado bruto a preços correntes da indústria (Mil Reais) / Valor adicionado bruto a preços correntes total (Mil Reais);

X₃: Valor adicionado bruto a preços correntes dos serviços, inclusive administração, saúde e educação públicas e seguridade social (Mil Reais) / Valor adicionado bruto a preços correntes total (Mil Reais);

X₄: pessoas ocupadas no comércio / pessoas ocupadas;

X₅: taxa de alfabetização;

X₆: total de veículos / população residente;

X₇: população residente urbana / população residente total;

X₈: salário médio mensal;

X₉: operações de crédito / população economicamente ativa;

X₁₀: leitos por mil habitantes;

X₁₁: proporção de domicílios particulares permanentes em áreas urbanas com ordenamento regular adequados;

X₁₂: IDHM

X₁₃: PIB per capita;

X₁₄: pessoas ocupadas na indústria / pessoas ocupadas;

X₁₅: proporção de domicílios particulares com microcomputado e acesso a internet;

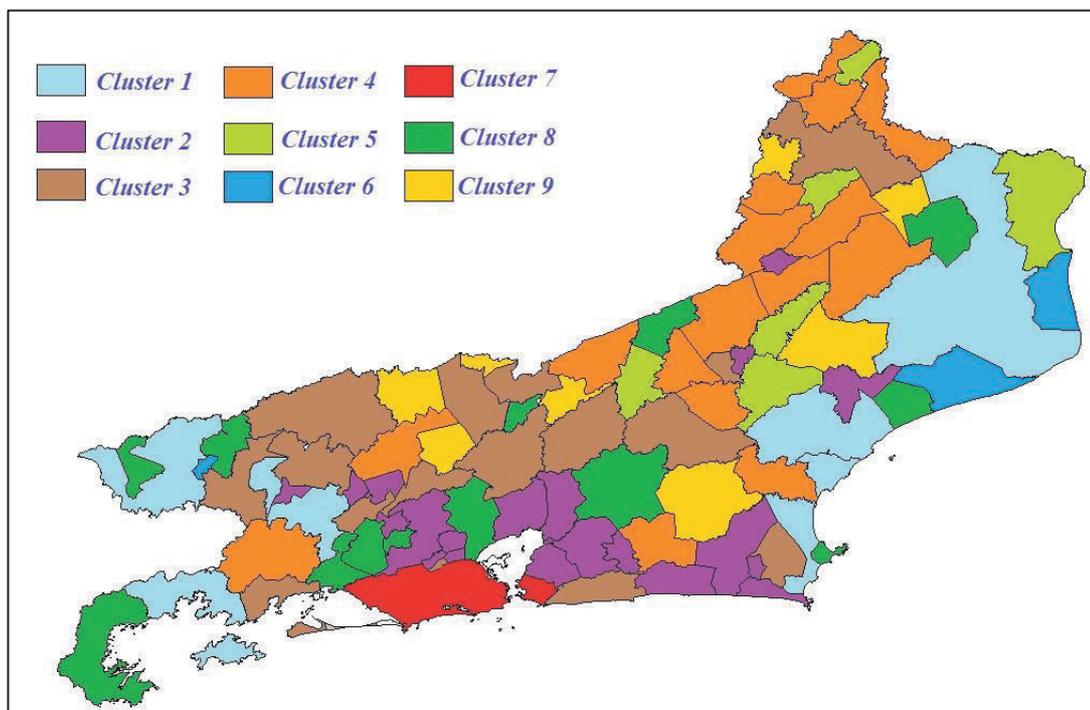
X₁₆: proporção de domicílios com rendimento nominal mensal de mais de 10 a 20 salários mínimos;

X₁₇: proporção de domicílios particulares permanentes do tipo apartamento;

X₁₈: número de famílias com bolsa família / população residente total.

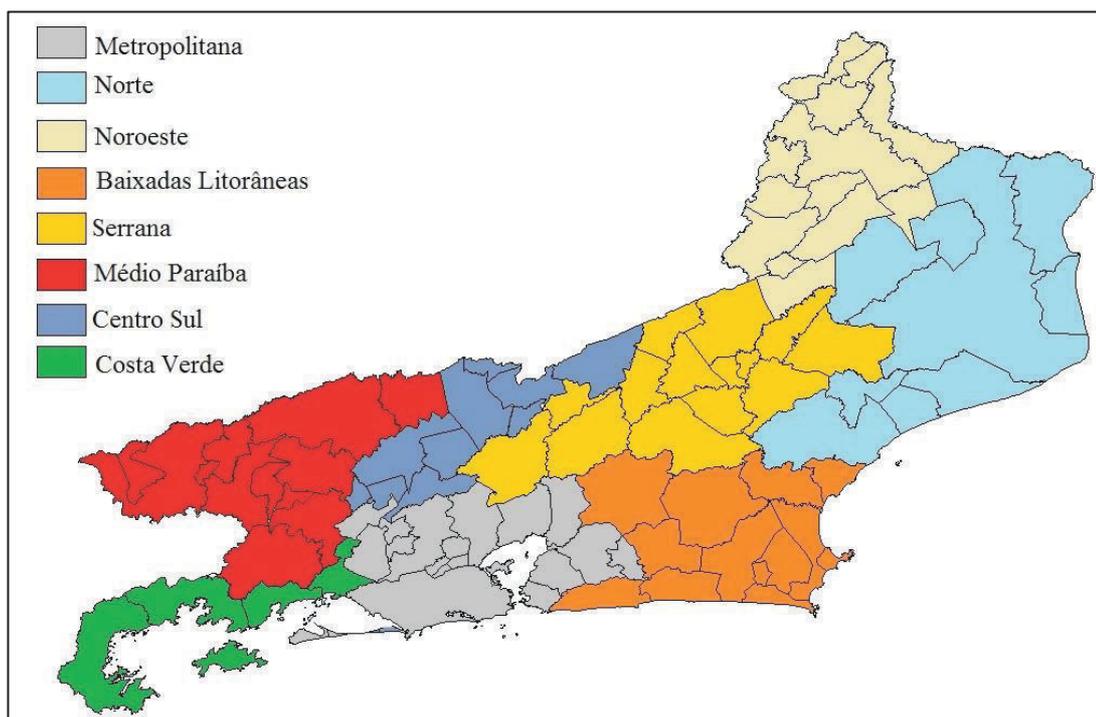
ANEXO II

Distribuição dos municípios do Estado do Rio de Janeiro de acordo com os *cluster* selecionados.



FONTE: Elaboração própria.

Distribuição dos municípios do Estado do Rio de Janeiro de acordo com suas Mesorregiões.



FONTE: Elaboração própria.