

UNIVERSIDADE CANDIDO MENDES – UCAM  
PROGRAMA DE PÓS-GRADUAÇÃO EM PLANEJAMENTO REGIONAL E GESTÃO  
DA CIDADE  
CURSO DE DOUTORADO EM PLANEJAMENTO REGIONAL E GESTÃO DA  
CIDADE

FABIO FREITAS DA SILVA

**RELAÇÃO ENTRE CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO  
HUMANO:  
UMA ANÁLISE PARA OS MUNICÍPIOS DO ESTADO DO RIO DE JANEIRO**

CAMPOS DOS GOYTACAZES, RJ  
Setembro de 2020

UNIVERSIDADE CANDIDO MENDES – UCAM  
PROGRAMA DE PÓS-GRADUAÇÃO EM PLANEJAMENTO REGIONAL E GESTÃO  
DA CIDADE  
CURSO DE DOUTORADO EM PLANEJAMENTO REGIONAL E GESTÃO DA  
CIDADE

FABIO FREITAS DA SILVA

**RELAÇÃO ENTRE CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO  
HUMANO:  
UMA ANÁLISE PARA OS MUNICÍPIOS DO ESTADO DO RIO DE JANEIRO**

Tese de doutorado apresentada ao  
Programa de Doutorado em Planejamento  
Regional e Gestão de Cidades da  
Universidade Candido Mendes –  
Campos/RJ, como requisito para obtenção  
do grau de DOUTORA EM  
PLANEJAMENTO REGIONAL E GESTÃO  
DA CIDADE

Orientador: Prof. Ítalo de Oliveira Matias, D.Sc.  
Coorientadora: Prof. Lia Hasenclever, D.Sc.

CAMPOS DOS GOYTACAZES, RJ  
Setembro de 2020

## Catálogo na Fonte

Preparada pela Biblioteca da **UCAM – CAMPOS** 004/2021

Silva, Fábio Freitas da.

Relação entre crescimento econômico e desenvolvimento humano: uma análise para os municípios do estado do Rio de Janeiro. / Fábio Freitas da Silva. – 2020. 89 f.

Orientador(a): Ítalo de Oliveira Matias.  
Coorientador(a): Lia Hasenclever.

Tese de Doutorado em Planejamento Regional e Gestão da Cidade – Universidade Candido Mendes – Campos. Campos dos Goytacazes, RJ, 2020. Referências: f. 79-83.

1. Desenvolvimento econômico. 2. Desenvolvimento humano. 3. Rio de Janeiro - Municípios. I. Matias, Ítalo de Oliveira, orient. II. Hasenclever, Lia, coorient. III. Universidade Candido Mendes – Campos. IV. Título.

CDU – 330.34 (815.3)

Bibliotecária Responsável: Flávia Mastrogirolamo CRB 7ª-6723

FABIO FREITAS DA SILVA

**RELAÇÃO ENTRE CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO  
HUMANO:  
UMA ANÁLISE PARA OS MUNICÍPIOS DO ESTADO DO RIO DE JANEIRO**

Tese de doutorado apresentada ao  
Programa de Doutorado em Planejamento  
Regional e Gestão de Cidades da  
Universidade Candido Mendes –  
Campos/RJ, como requisito para obtenção  
do grau de DOUTORA EM  
PLANEJAMENTO REGIONAL E GESTÃO  
DA CIDADE

Avaliado em 25 de setembro de 2020

*BANCA EXAMINADORA*

---

Prof. Ítalo de Oliveira Matias, D.Sc.  
Universidade Candido Mendes

---

Prof. Lia Hasenclever, D.Sc.  
Universidade Candido Mendes

---

Prof. Graciela Profeta, D.Sc.  
Universidade Federal Fluminense

---

Prof. Yves-A. Fauré, D.Sc.  
Professeur IRD / Université Paris 1, FR

---

Prof. Eduardo Shimoda, D.Sc.  
Universidade Candido Mendes

CAMPOS DOS GOYTACAZES, RJ  
2020

### ***DEDICATORIA***

Dedico este trabalho a Deus que sempre esteve ao meu lado, a minha família, e a todos os professores que foram tão importantes na minha vida acadêmica e no desenvolvimento desta tese.

### ***AGRADECIMENTOS***

Agradeço a CAPES, por tornar possível a conclusão do curso através da bolsa concedida. Agradeço a Universidade Candido Mendes, por fornecer conhecimento e confiança nos estudos oferecidos. Assim como, todos os funcionários a todo corpo docente da universidade que oferecem um ótimo ambiente de trabalho. Não posso deixar de agradecer, ao meu orientador Professor Ítalo Matias e minha coorientadora Professora Lia Hasenclever, por todo conhecimento transmitido ao longo dessa jornada de aprendizado.

## RESUMO

### RELAÇÃO ENTRE CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO HUMANO: UMA ANÁLISE DOS MUNICÍPIOS DO RIO DE JANEIRO

Na literatura é apontada por diversos autores a relação dinâmica entre crescimento econômico e desenvolvimento humano. Quando essas duas vertentes se reforçam mutuamente no sentido positivo, surge um círculo virtuoso de desenvolvimento. Por outro lado, quando ambos caminham em direção contrária, um círculo vicioso é formado. A partir dessa problemática, o objetivo da presente tese foi analisar o comportamento desses círculos nos municípios fluminenses, que sofreram significativas mudanças institucionais e econômicas ao longo dos séculos XX e XXI, e identificar os fatores que levaram ao círculo vicioso da pobreza. Os métodos utilizados na análise foram de cunho tanto exploratório quanto inferencial. Na análise exploratória, técnicas de estatísticas descritiva e de redução de dimensionalidade, análise fatorial múltipla, foram abordadas. Ainda na parte exploratória, elaborou-se de mapas temáticos. Além disso, na análise inferencial utilizou-se o modelo de regressão logística para verificar o impacto dos indicadores socioeconômicos na composição do círculo vicioso. Os dados da pesquisa foram referentes ao último censo demográfico disponível (2010), delimitados geograficamente em função dos municípios do Estado do Rio de Janeiro. Os principais resultados mostraram que as variáveis esperança de vida (32,63%); renda *per capita* (0,12%); probabilidade de sobreviver até os 60 anos (22,52%); taxa de indivíduos com 18 anos ou mais com ensino fundamental completo (3,87%); porcentagem de pessoas com 25 anos ou mais com superior completo (9,05%); e expectativa de anos de estudo (64,62%) reduzem a chance ( $P < 0,5$ ) de um município permanecer ou adentrar no círculo vicioso da pobreza a cada unidade incrementada. Sob outra perspectiva, a taxa de mortalidade infantil e de pessoas vulneráveis à pobreza aumentam essas chances em 23,99% e 3,63%, respectivamente. Isso demonstra que os municípios que apresentaram altas taxas de mortalidade infantil também tendem a apresentar altos índices de analfabetismo e pobreza, concomitantemente, baixos índices de escolaridade – tanto no ensino superior quanto no fundamental – menor renda per capita e longevidade.

**Palavras-chave:** Índice de desenvolvimento humano. Crescimento econômico. Saúde. Educação. Municípios do Rio de Janeiro.



## **ABSTRACT**

### **RELATIONSHIP BETWEEN ECONOMIC GROWTH AND HUMAN DEVELOPMENT: AN ANALYSIS OF THE MUNICIPALITIES OF RIO DE JANEIRO**

In the literature, the dynamic relationship between economic growth and human development is visible. When these two strands reinforce each other in the positive sense, a virtuous circle of development appears, on the other hand, when both go in the opposite direction, a vicious circle is formed. Based on this problem, the objective of this Thesis was to analyze the behavior of the circles in the municipalities of Rio de Janeiro, which underwent institutional changes and throughout the 20th and 21st centuries, and to identify the factors that led to the vicious circle of poverty. The methods used in the analysis were both exploratory and inferential. In exploratory analysis, descriptive statistics technique and dimensionality reduction, multiple factor analysis, were addressed. Still in the exploratory part, thematic maps were elaborated. On the other hand, in the inferential analysis, the logistic regression model was used to verify the impact of socioeconomic indicators on the composition of the vicious circle. The survey data were for the last available demographic census (2010), geographically delimited according to the municipalities of the State of Rio de Janeiro. The main determinant results are that the life variables (32.63%); per capita income (0.12%); possibility of surviving up to 60 years (22.52%); rate of owners aged 18 or over with complete elementary school (3.87%); percentage of people aged 25 or over with a college degree (9.05%); and expectation of years of study (64.62%) to give a chance,  $\alpha = 5\%$ , for a municipality to remain or enter the vicious circle of poverty at each incremented unit. From another perspective, the infant mortality rate and that of people vulnerable to poverty increase these chances by 23.99% and 3.63%, respectively. This demonstrates that the parallels that arise high rates of infant mortality also tend to present high rates of illiteracy and poverty, concomitantly, low levels of education - both in higher education and in elementary education - lower per capita income and longevity.

**Keywords:** Human development index. Economic growth. Health. Education. municipalities of Rio de Janeiro.

## LISTA DE TABELAS

Tabela 1. Grupos de variáveis que compõem o modelo .....	29
Tabela 2. Classificação dos municípios quanto ao desenvolvimento.....	34
Tabela 3. Estatísticas descritivas das variáveis de estudo.....	48
Tabela 4. Média dos indicadores dos municípios do estado do Rio de Janeiro, ano 2010. ....	53
Tabela 5. Classificação dos municípios do ERJ quanto ao nível de desenvolvimento .....	55
Tabela 6. Peso/coeficiente (autovetores) das variáveis quantitativas para as dimensões 1 e 2 .....	58
Tabela 7. Peso (autovetores) da variável categórica para as dimensões 1 e 2 .....	60
Tabela 8. Comparação das médias das categorias de desenvolvimento (classificação ponderada) em relação à média geral das variáveis avaliadas.....	65
Tabela 9. Modelo de regressão binomial com regularização de Ridge .....	69

## LISTA DE FIGURAS

Figura 1. Representação do círculo entre crescimento econômico (CE) e desenvolvimento humano (DH).....	13
Figura 2. Relação de casualidade entre saúde, educações e crescimento econômico .....	20
Figura 3. Diagrama de caixas e médias das variáveis avaliadas.....	46
Figura 4. Histograma e mediana das variáveis da pesquisa .....	49
Figura 5. Matriz de correlação de Pearson das variáveis estudadas.....	50
Figura 6. Matriz do diagrama de dispersão das variáveis de estudo.....	51
Figura 7. Dendograma resultante da análise de agrupamento, utilizando como medida de dissimilaridade o MDS de uma matriz r (r: correlação de Pearson). .....	52
Figura 8: Número de municípios classificados de acordo com seu nível de desenvolvimento. ....	54
Figura 9. Cartograma dos municípios do ERJ conforme seu nível de desenvolvimento .....	56
Figura 10. Proporção de variância explicada pelos principais autovalores .....	57
Figura 11. O gráfico de correlação entre os grupos e as duas primeiras dimensões, média ponderada. ....	58
Figura 12. Círculo de correlação entre as variáveis quantitativas; Dim1 - eixo x, Dim2 - eixo y.....	59
Figura 13. Mapa de fator da variável categórica ND; Dim1 - eixo x, Dim2 - eixo y....	60
Figura 14. Mapa de fatores dos indivíduos (municípios), média ponderada .....	63
Figura 15. Diagrama de dispersão e média das variáveis (padronizadas, z-score) do estudo, separadas conforme as classes do modelo binomial .....	67
Figura 16. Histograma das variáveis (padronizadas, z-score) do estudo, categorizadas conforme as classes do modelo binomial.....	68
Figura 17. Medidas de ajuste do modelo de regressão binomial .....	69
Figura 18. Cartograma dos atributos significativos no modelo de regressão .....	72
Figura 19. Cartograma da agregação das variáveis significativas no modelo de regressão .....	73

## LISTA DE ABREVIATURAS

CE	Crescimento econômico
CVic	círculo vicioso
CVir	círculo virtuoso
DH	Desenvolvimento humano
EAE	Expectativa de anos de estudo
EV	Esperança de vida ao nascer
FC18	% de 18 anos ou mais com fundamental completo
Gi	Índice de Gini
IDH	Índice de desenvolvimento humano
MDS	Escalonamento Multidimensional
MFA	Análise Fatorial Múltipla
Mi	Mortalidade infantil
P	% de pobres
PCA	Análise de Componentes Principais
PDA	% da população em domicílios com água encanada
PDBA	% da população em domicílios com banheiro e água encanada;
PDCL	% da população em domicílios com coleta de lixo
PIB (Mil)	Produto interno bruto
PNUD	Programa das Nações Unidas para o Desenvolvimento
PS60	Probabilidade de sobrevivência até 60 anos
QL	Quociente Locacional
RP	Renda per capita
SC25	% de 25 anos ou mais com superior completo;
SVD	Decomposição do Valor Singular
T18	Taxa de analfabetismo - 18 anos ou mais
TC	Tendendo ao crescimento
TD	Tendendo ao desenvolvimento
VP	% de vulneráveis à pobreza

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
1.1	CONTEXTUALIZAÇÃO	12
1.2	OBJETIVO	14
1.2.1	OBJETIVO GERAL	15
1.2.2	OBJETIVO ESPECÍFICOS	15
1.3	JUSTIFICATIVA	15
1.4	QUESTÕES DE PESQUISA	16
1.5	DELIMITAÇÃO DO TRABALHO	16
1.6	ESTRUTURA DO TRABALHO	16
<b>2</b>	<b>REVISÃO DE LITERATURA</b>	<b>18</b>
2.1	ESTADO DA ARTE	19
2.2	DESENVOLVIMENTO HUMANO	22
2.3	CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO HUMANO	25
<b>3</b>	<b>METODOLOGIA</b>	<b>28</b>
3.1	CLASSIFICAÇÃO DOS MUNICÍPIOS	30
3.2	ANÁLISE FATORIAL MÚLTIPLA (MFA)	34
3.2.1	Análise de Componentes Principais (PCA)	35
3.2.2	Descrição da MFA	37
3.3	MENSURAÇÃO DO IMPACTO DAS VARIÁVEIS SOCIOECONÔMICAS NA COMPOSIÇÃO DO CÍRCULO VICIOSO	38
3.3.1	Descrição do modelo de regressão penalizada	39
3.4	MAPAS TEMÁTICOS	43
<b>4</b>	<b>RESULTADOS</b>	<b>45</b>
4.1	ANÁLISE EXPLORATÓRIA	45
4.2	CLASSIFICAÇÃO DOS MUNICÍPIOS	53
4.3	ANÁLISE FATORIAL	56
4.4	MODELO DE REGRESSÃO	66
4.5	MAPAS	71
<b>5</b>	<b>DISCUSSÃO</b>	<b>74</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>77</b>
6.1	SUGESTÃO DE TRABALHOS FUTUROS	78
	REFERÊNCIAS	79
	APÊNDICE A: Cartograma das variáveis quantitativas	84

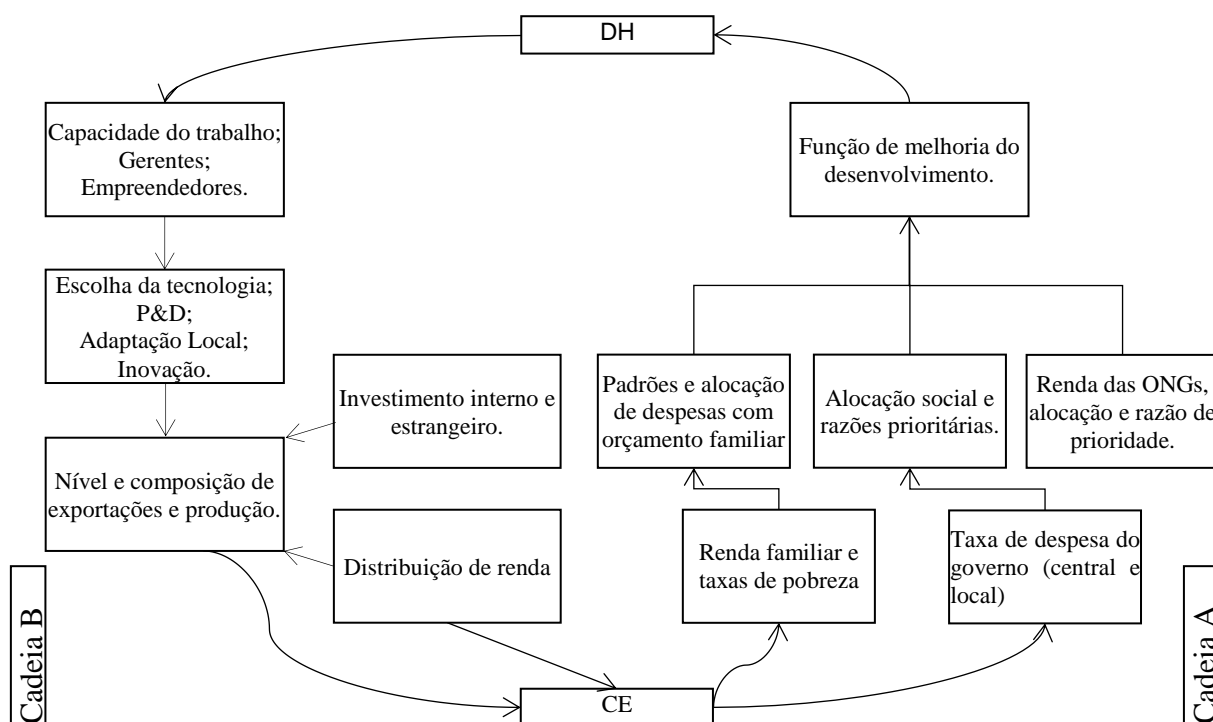
# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

De acordo com RANIS; STEWART; RAMIREZ. (2000) na literatura existem duas linhas de pensamento referentes à relação entre desenvolvimento humano (DH) e crescimento econômico (CE). A primeira linha de pensamento, denominada cadeia A, acredita que o CE promove o DH. Por outro lado, a segunda linha de pensamento, chamada de cadeia B, que melhorias no DH implicarão em CE (RANIS; STEWART, 2005). Segundo os autores, muitos estudos mostraram separadamente a existência dessas linhas de pensamento: a cadeia A está inserida na literatura sobre necessidades básicas (UNDP, 1991; BEHRMAN, 1996; HADDAD *et al.*, 2003); enquanto a cadeia B tem sido explorada na literatura sobre capital humano (PSACHAROPOULOS, 1994, STRAUSS; THOMAS, 1995; ROSENZWEIG, 1995; BAYRAKTAR-SAGLAM, 2017).

No entanto, existem poucos trabalhos que estudaram a interdependência dessas duas linhas de pensamento, ou seja, a relação dinâmica entre CE e DH (RANIS; STEWART, 2005; STEWART, 2019). Segundo RANIS; STEWART; RAMIREZ. (2000,) a conexão entre CE e DH é evidente, uma vez que o CE disponibiliza os recursos necessários para as melhorias no DH; conseqüentemente, uma força de trabalho capacitada impacta positivamente o CE. Para os autores essa relação, apresentada na Figura 1, não é automática e depende da força de cada elo e suas conexões.

Figura 1. Representação do círculo entre crescimento econômico (CE) e desenvolvimento humano (DH).



Fonte: Adaptado de Ranis e Stewart (2005).

Ambas as cadeias têm importância na causalidade que leva ao DH, sendo os resultados intensificados, na cadeia A, pela ampliação do gasto social, ampliação do número de matrículas escolares femininas, melhor distribuição de renda e baixos índices de pobreza; na cadeia B, pela maior taxa de investimento interno e externo (RANIS; STEWART, 2005). De acordo com os autores a existência dessas duas cadeias que conectam CE e DH forma um círculo que é reforçado mutuamente. Assim, CE leva a DH que, por sua vez, promove maior CE e vice-versa. Sobre outra perspectiva, efeitos negativos também podem levar a baixo CE e, por sua vez, a um menor índice de DH e assim por diante.

Chikalipah e Okafor (2019) reforçam as ideias de Ranis e Stewart (2005) ao afirmarem que a casualidade entre CE e DH pode ser obtida de duas formas. A primeira, relativa à cadeia A, pode ser alcançada de duas maneiras: com o crescimento da renda familiar e a propensão de gastos com saúde e educação; e com o aumento do investimento do governo em gastos sociais, principalmente na saúde e educação. A segunda, concernente à cadeia B, pode ser atingida da seguinte forma: por meio da elevação da produtividade econômica e com a redução de gastos dado

uma população saudável e instruída; aumento das exportações e competitividade decorrente de uma força de trabalho altamente qualificada.

Nessa perspectiva, Ranis e Stewart (2005) propõem classificar o desempenho dos países em quatro categorias: (i) círculo virtuoso, em que mutuamente CE sustenta DH; (ii) círculo vicioso, no qual CE não proporciona DH que, por sua vez, não promove CE e assim por diante; (iii) tendendo ao crescimento (CE desequilibrado) em que se tem um CE relativamente elevado e baixo DH e; (iv) tendendo ao desenvolvimento (DH desequilibrado) em que se tem um alto índice relativo de DH e um baixo CE.

Nesse contexto, o espaço geopolítico fluminense sofreu significativas mudanças institucionais e econômicas ao longo dos séculos XX e XXI. Olhando para o século XX, seu território abrigou a Capital Federal até 1960, depois essa cidade transformou-se no Estado da Guanabara e em 1975 fundiram-se os territórios carioca e fluminense, formando o Estado do Rio de Janeiro (ERJ). A fusão de dois territórios com origens históricas, produtivas e institucionais distintas engendrou diferenças marcantes. Nas terras cariocas havia sido construído um sólido sistema industrial e comercial muito superior ao interior fluminense, que, na maior parte do século XX, permaneceu pobre e estagnado. Apesar disso, a economia fluminense persistiu em segundo lugar no *ranking* do Produto Interno Bruto (PIB) nacional.

A chegada do século XXI trouxe novas mudanças na estrutura econômica do estado, deslocando a indústria para a indústria extrativa e aumentando sua representatividade no interior do Estado. O petróleo anuncia um período de potencial expansão e crescimento econômico fluminense, mas também de deslocamento da atividade industrial em seu território, da metrópole para o interior. Sua demanda de logística, equipamentos e serviços tem papel importante de indução da demanda efetiva para o desenvolvimento local. Diante dessas mudanças os círculos de desenvolvimento dos municípios se alteraram ao longo do tempo. A questão que se coloca é em que sentido (positivo ou negativo) e quais fatores influenciaram diretamente essas transformações. Mais especificamente, em que medida o CE proporcionado pela expansão da indústria de petróleo e gás natural (P&G) resultou em um aumento do DH?

## 1.2 OBJETIVO



### 1.2.1 OBJETIVO GERAL

O objetivo desta tese foi identificar os fatores que levaram alguns municípios fluminenses ao círculo vicioso da pobreza em 2010.

### 1.2.2 OBJETIVO ESPECÍFICOS

1. Classificar os municípios do Estado do Rio de Janeiro quanto ao seu nível de desenvolvimento em 2010;
2. Identificar os fatores que mais contribuíram para a formação do círculo vicioso nesses municípios;
3. Verificar o impacto de cada fator na contribuição do círculo vicioso; e
4. Elaborar mapas temáticos com as variáveis que contribuíram para composição do círculo vicioso da pobreza e criar um mapa com a combinação desses atributos.

## 1.3 JUSTIFICATIVA

As condições de subdesenvolvimento humano são agravadas temporalmente em regiões que possuem elementos negativos profundos, tais como níveis baixos de renda, saúde e educação. Para romper o círculo vicioso é necessário conhecer os fatores que aumentam a chance de um município permanecer nessa condição, de tal forma que auxilie na promoção de políticas públicas eficazes para o seu rompimento (OLIVEIRA; LIMA; BARRINHA, 2019). A ruptura do círculo vicioso inicia-se a partir do mapeamento das regiões que estão inseridas nesse processo de subdesenvolvimento (RAIHER; LIMA, 2014).

Esta tese pretende contribuir para o conhecimento de duas maneiras. Em primeiro lugar, por estudar uma região específica que tem crescido bastante com a exploração do petróleo, mas paradoxalmente mantém níveis de desenvolvimento

humano muito aquém do que poderia obter segundo as teorias tradicionais que indicam que primeiro dever-se-ia crescer para depois distribuir e ampliar o desenvolvimento humano. Estudos dessa natureza têm sido feitos para outras regiões (SILVA, 2019), mas para a região fluminense, objeto desta tese, ainda foram realizados poucos estudos, por exemplo, Silva *et al.* (2019). Em segundo lugar, por trazer uma contribuição metodológica com a aplicação de métodos estatísticos multivariados que auxiliaram na compreensão das interrelações entre os indicadores socioeconômicos; e, por meio da utilização de Sistema de Informação geográfica (SIG), entender distribuição espacial dessas variáveis, avançando, assim, na contribuição de Raiher e Lima (2014).

#### **1.4 QUESTÕES DE PESQUISA**

O CE proporcionado pela expansão da indústria de petróleo e gás natural (P&G) resultou em um aumento do DH?

Quais fatores influenciaram significativamente a formação do círculo vicioso?

#### **1.5 DELIMITAÇÃO DO TRABALHO**

O estudo está centrado nos municípios do Estado do Rio de Janeiro. O Índice de Desenvolvimento Humano (IDH) foi utilizado como *proxy* do desenvolvimento humano. Apesar de não refletir totalmente a qualidade de vida das pessoas, mensura os pilares essenciais (renda, educação e longevidade) para o desenvolvimento da economia de um país ou região. O recorte temporal foi referente a 2010, por corresponder ao último censo demográfico disponível.

#### **1.6 ESTRUTURA DO TRABALHO**

O trabalho está estruturado em seis capítulos e um apêndice. A presente seção, capítulo 1, apresenta a Introdução da pesquisa.

O capítulo 2 compõe a revisão da literatura, dividida em três subitens: (i) estado da arte, que aborda a relação de trabalhos empíricos que analisam a dinâmica entre CE e DH; (ii) desenvolvimento humano, que traz o cenário histórico até a culminada do conceito de DH; (iii) crescimento econômico e desenvolvimento humano, que retrata os aspectos teóricos da relação CE e DH.

No capítulo 3, encontra-se a metodologia empregada, subdividida em quatro abordagens principais: classificação dos municípios; análise fatorial múltipla; modelo de regressão; e elaboração de mapas.

No capítulo 4, os resultados foram apresentados na mesma estruturação dos métodos utilizados, além de uma análise exploratória dos dados. No capítulo 5, tem-se a discussão e, por fim, no capítulo 6, a conclusão.

No Apêndice A desta tese constam os mapas elaborados.

## 2 REVISÃO DE LITERATURA

A pesquisa para a revisão da literatura foi realizada utilizando a seguinte sintaxe na base Scopus: ( "*Human Development index*" AND ( "*Economic Growth*" OR "*Health*" OR "*Education*" ) ). A base *Scopus*, disponível no portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), foi utilizada como mecanismo de busca. Essa base é o maior banco de dados de resumos e citações da literatura revisado por pares. Estão inclusas revistas científicas, livros e anais de congressos que abrangem temas de diversas áreas do conhecimento. São mais de 70 milhões de registros atualizados diariamente, 1.4 bilhões de referências citadas desde 1970, aproximadamente 70 mil instituições cadastradas e 16 milhões de autores (ELSEVIER, 2019).

A busca foi realizada entre o período mais antigo disponível pela base (1ª publicação sobre o tema em 1991) até 28 de maio de 2019. Os termos utilizados foram pesquisados no título, resumo e palavras-chave dos artigos de periódicos.

Para leitura dos artigos foi realizada uma pré-seleção no conjunto dos trabalhos encontrados na busca (1.169). Filtraram-se os estudos pelas seguintes áreas de pesquisa disponíveis na base *Scopus*: economia; econometria e finanças; e ciências sociais. Assim, um total de 450 estudos foram selecionados para verificação. Desses, apenas sete abordaram a causalidade entre CE e DH. Além desses artigos da base *Scopus* outros trabalhos aderentes ao tema foram acrescentados, por sugestão dos orientadores, formando o núcleo inicial de pesquisa.

Este capítulo foi dividido em três seções. Todas elas têm como fonte de pesquisa a seleção de supracitados. A primeira representa o estado da arte, que mostra a síntese dos artigos que tinham objetivos, resultados e estratégias empíricas mais próximas ao desenho desta Tese. A segunda apresenta o conceito de

desenvolvimento humano e sua origem histórica. Por fim, a terceira seção apresenta a relação do desenvolvimento humano com o crescimento econômico.

## 2.1 ESTADO DA ARTE

Existem poucos estudos que lidam com a relação dinâmica entre desenvolvimento humano e crescimento econômico (RANIS; STEWART; RAMIREZ., 2000; STEWART, 2019).

O primeiro trabalho a analisar a conexão entre crescimento econômico e desenvolvimento humano foi o de Ranis, Stewart e Ramirez (2000). A partir das observações de duas linhas de pensamento na literatura – uma que DH promove CE, e outra que o CE leva ao DH –, os autores formularam a hipótese sobre a interação dessas cadeias, ou seja, um *link* de causalidade. Os principais resultados mostraram: (i) a existência de uma relação significativa em ambas as direções; (II) e que se o CE não for precedido ou acompanhado por melhorias no DH não será sustentável por si só.

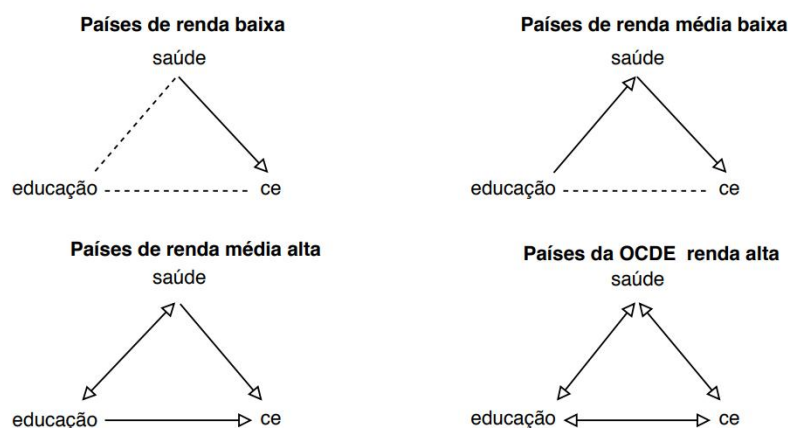
Na mesma linha de raciocínio Suri *et al.* (2011) investigaram essa relação de mão dupla entre CE e DH. Os resultados também sugerem um efeito de causalidade bidirecional, e que o DH deve preceder ou acompanhar o CE para que esse seja viável a longo prazo, o que contradiz a visão de que os investimentos em DH podem ser adiados até que a expansão da economia os torne acessíveis. O DH não é somente um fim para o bem-estar, é também um meio essencial para um crescimento sustentável.

De outra maneira, Mustafa, Rizov e Kernohan (2017) analisaram a relação dinâmica entre CE, DH e abertura comercial na economia asiática. Os principais achados mostraram que, embora o desenvolvimento humano contribua positivamente para o crescimento econômico, a relação inversa não foi observada. Para os autores isso ocorreu devido ao crescimento desigual e pelo atraso do desenvolvimento institucional, impossibilitando a formação do capital humano, o que pode ter restringido o DH no curto e médio prazo.

Já Bayraktar-Sağlam (2017) explorou a interação dinâmica entre saúde, educação e crescimento. Os principais achados, sintetizados na Figura 2, mostraram

uma relação complexa entre as interações desses elementos. Nos países de renda baixa, a saúde explica o crescimento econômico. No entanto, não há padrão preditivo entre educação e crescimento econômico e entre educação e saúde. Sendo assim, no estágio inicial do desenvolvimento a saúde desempenha um papel essencial na geração do crescimento econômico. Por outro lado, nos países de renda média baixa, a educação promove a saúde que, por sua vez, estimula o crescimento econômico. Mas, o círculo não é perfeito, uma educação de qualidade não tem impacto direto no crescimento econômico. Dessa maneira, a saúde ainda é um fator fundamental para a geração de crescimento econômico. Nos países de renda média alta, a educação e a saúde agem de forma positiva para o crescimento econômico, no qual ambas têm poder de predição sobre a economia e ainda possuem uma relação de reforço mútuo. Contudo, essa relação mútua não ocorre com o crescimento econômico. Sob outra perspectiva, os países de alta renda da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) apresentaram um círculo perfeito, também chamado de virtuoso, em que saúde, educação e crescimento econômico se reforçam mutuamente.

**Figura 2.** Relação de causalidade entre saúde, educação e crescimento econômico



Fonte: Adaptado de Bayraktar-Sağlam (2017).

Chikalipah e Okafor (2019) investigaram o link dinâmico entre crescimento econômico e desenvolvimento humano na Nigéria. Os principais resultados mostraram que, em longo prazo, o DH e o crescimento econômico foram cointegrados. Porém, apesar de ter uma relação no longo prazo, apenas o crescimento econômico exerceu um efeito significativo sobre o DH na Nigéria; portanto uma relação de

casualidade bilateral não foi estabelecida. Uma das possíveis explicações atribuídas pelos autores se dá pelo fato da não promoção do capital humano, devido a altas taxas de crescimento populacional, governança ruim, fracas qualidades institucionais e infraestrutura subdesenvolvidas. Na Zambia, Chikalipah e Makina (2019) encontraram resultados similares: uma cointegração entre CE e DH; efeito do crescimento econômico no desenvolvimento humano no longo prazo, sem casualidade reversa.

Finalmente, o artigo de Stewart (2019) é um estudo teórico sobre as questões do DH. Mostra como evoluiu o pensamento sobre o desenvolvimento – inicialmente pensado como desenvolvimento econômico em oposição a crescimento econômico – e a ampliação dos seus objetivos para aspectos sociais, o desenvolvimento humano. Também investigou a relação de mão dupla entre DH e CE.

A Síntese dos métodos, resultado principal e amostras utilizadas nos trabalhos que estudam a casualidade de mão dupla entre CE e DH foram descritos no Quadro 1.

Quadro 1: Síntese dos métodos, achados e amostras utilizadas nos trabalhos que estudaram a casualidade de mão dupla entre CE e DH

Autor	Método	Achados	Amostra
Ranis; Stewart; Ramirez (2000)	Ordinary Least Squares (OLS)	CE↔DH	35 a 70 países 1960-1992
Suri <i>et al.</i> (2011)	Panel data, OLS	CE↔DH	79 países 1960-2001
Bayraktar-Sağlam (2017)	Panel vector autoregression (VAR), Generalized Method of Moments (GMM)	CE↔DH	102 países 1975-210
Mustafa, Rizov Kernohan (2017)	System of simultaneous equations	DH→CE	20 países asiáticos 1970-2010
Chikalipah Okafor (2019)	Gregory–Hansen cointegration, Dynamic Ordinary Least Square (DOLS), Vector Error Correction Model (VECM)	CE→DH	Nigéria, n=55 1961- 2015,
Chikalipah e Makina (2019)	Gregory–Hansen cointegration, Johansen cointegration, Dynamic ordinary least squares (DOLS), Vector Error Correction Model (VECM)	CE→DH	Zambia 1970-2015

Fonte: Autor

Partindo do princípio da existência dessas relações dinâmicas entre CE e DH, alguns trabalhos classificam os países ou regiões conforme seu nível de desenvolvimento: círculo vicioso; círculo virtuoso; tendendo ao desenvolvimento; e

tendendo ao crescimento. O primeiro deles foi o de Ranis, Stewart e Ramirez (2000), que classificou os países com base no IDH (sem a componente renda) e PIB *per capita* em relação à média desses indicadores. Seguindo essa mesma metodologia, existem ainda o estudo de Suri *et al.* (2011) e Stewart (2019).

Há também adaptações dessa metodologia, em que, ao invés de se utilizar o índice IDH agregado (sem a componente renda) e o PIB *per capita*, faz-se uso das subdimensões educação, longevidade e renda *per capita*. O Programa das Nações Unidas para o Desenvolvimento - PNUD - (2003) do México foi o precursor dessa abordagem adaptada. Posteriormente os trabalhos de: Raiher e Lima (2014) analisaram a evolução dos níveis de desenvolvimento na região Sul do Brasil no período de 2000 a 2010; Oliveira, Lima e Raiher. (2017) investigaram a evolução dos níveis de desenvolvimento na região Nordeste do Brasil no período de 2000 a 2010; e Oliveira, Lima e Barrinha (2019) analisaram o desenvolvimento municipal das microrregiões da Bahia, referentes a 2000 e 2010.

## 2.2 DESENVOLVIMENTO HUMANO

O objetivo central do desenvolvimento na abordagem do DH é melhorar a vida das pessoas, sendo um processo de ampliação de suas escolhas (PNUD, 1990). De acordo com o relatório de DH, as escolhas mais críticas são: viver uma vida longa e saudável, ter uma educação de qualidade, assim como os recursos necessários para um padrão de vida digno.

A abordagem que tornou o bem-estar o objetivo central do desenvolvimento foi construída nas décadas de 1970 e 1980<sup>1</sup>, sendo consolidada no relatório de desenvolvimento humano de 1990.

No início da década de 1950, os tomadores de decisões adotaram o crescimento econômico como objetivo central do desenvolvimento, no qual era visto como mecanismo primordial para que os países em desenvolvimento alcançassem as nações economicamente avançadas (STEWART, 2019). Conforme colocado por Stewart, essa visão de desenvolvimento foi bem-sucedida em seus próprios termos –

---

<sup>1</sup> Embora tenha raízes na Grécia antiga, com o eudemonismo de Aristóteles, o debate aqui inicia-se na década de 1950.



principalmente nos países em desenvolvimento de renda média. Entretanto, apesar da redução das taxas de mortalidade infantil e aumento da alfabetização de adultos, os níveis de pobreza não se reduziram com essa abordagem, mas, sim, aumentaram à medida que a população crescia.

Na metade da década de 1960 observou-se um número crescente de evidências que demonstravam o desacoplamento entre o CE e o bem-estar da população em países do terceiro mundo. Embora o PIB aumentasse, os níveis de pobreza cresciam e as desigualdades sociais se sobressaíam em vários países. A *proxy* do desenvolvimento na época, PIB *per capita*, cada vez mais se tornava inapropriada para mensurar o bem-estar social (JANUZZI, 2003).

Para a *International Labour Organization* (ILO), tornou-se evidente que o rápido crescimento econômico, principalmente nos países em desenvolvimento, a nível nacional, não diminui a pobreza ou desigualdade de maneira automática nem gerou empregos produtivos suficientes (ILO, 1976).

Abordagens alternativas surgiram em detrimento às abordagens que privilegiavam o crescimento econômico. A primeira delas foi feita pelo economista britânico Dudley Seers, na qual o foco principal do desenvolvimento era orientado para o emprego (ILO, 1970); no entanto, apesar de importante para a dignidade humana e a renda, foi um objetivo muito restrito (STEWART, 2019). A segunda, *Redistribution with growth*, proposta por um grupo de economistas do Instituto de Economia de Sussex e do Banco Mundial, enfatizou a renda, porém, com foco na redistribuição dos resultados do crescimento econômico aos pobres na forma de ativos produtivos (CHENERY *et al.*, 1974). Essa estratégia nunca foi implementada devido a falhas políticas e econômicas. Uma abordagem mais tangível foi a de necessidades básicas (NB), cuja prioridade estava em garantir as necessidades básicas de cada pessoa, tais como comida, abrigo, educação, além de elementos imateriais, como direito político e emprego (ILO, 1976). Tal abordagem influenciou o Banco Mundial a redirecionar recursos para saúde, educação e abrigo. Nenhum país em desenvolvimento adotou diretamente a estratégia, embora elementos da abordagem da NB já estivessem presentes em vários países desenvolvidos.

Na década de 1980 houve uma mudança radical de orientação nas políticas para o desenvolvimento. Tanto o bem-estar como o crescimento econômico foram substituídos por políticas de estabilização e ajustes no desenvolvimento. Isso ocorreu devido a alterações políticas nos Estados Unidos e Reino Unido, combinadas com a

crise da dívida em muitos países em desenvolvimento. O Keynesianismo deu lugar a uma orientação filosófica pró-mercado, monetarista, e de *laissez-faire* nos países desenvolvidos, sendo essa orientação passada para as nações de terceiro mundo pelas condições impostas pelo Fundo Monetário Internacional (FMI) e Banco Mundial por meio do Consenso de Washington. Tais instituições foram capazes de ditar as regras econômicas dos países em desenvolvimento devido às altas dívidas econômicas com as quais eles se encontravam, o que os forçavam a buscar apoio financeiro (STEWART, 2019).

Em 1985 determinadas secretarias da ONU alertaram sobre os custos de deterioração da qualidade de vida humana que essas políticas envolviam (JOLLY, 2012). No decorrer dos anos a abordagem adotada pelo FMI e Banco Mundial sofreu várias críticas por terem efeitos adversos no padrão de vida do ser humano nos países de terceiro mundo (KAKWANI, 1995). Essas estratégias causaram retração na renda *per capita*, estagnação da saúde e educação, aumento da pobreza e desigualdade nos países em desenvolvimento (STEWART, 2019).

Os resultados negativos das estratégias adotadas enfatizaram a importância de se reorientar o objetivo central do desenvolvimento para o florescimento<sup>2</sup> humano. O desenvolvimento da abordagem das capacidades, proposto por Amartya Sen, forneceu a estrutura teórica necessária para essa mudança de visão (STEWART, 2019). O desenvolvimento é colocado como uma medida de possibilidades ou oportunidades que se apresentam objetivamente para que os indivíduos possam exercer sua liberdade de escolha. Seria um processo integrado de expansão de liberdades substantivas interligadas, em que o foco está nos fins que as tornam importantes, em vez de restringi-las a alguns dos meios, como o crescimento da renda (SEN; MOTTA; MENDES, 2000).

A abordagem do DH surgiu na década de 1970, inspirada, particularmente, nas diretrizes das NB e abordagem das capacidades; ganhou força na década de 1980 devido à deterioração da condição de vida dos pobres (STEWART, 2019). O objetivo básico do desenvolvimento “[...] *is to create an enabling environment for people to enjoy long, healthy and creative lives*” (PNUD, 1990, p. 9). A liberdade é sublinhada

---

<sup>2</sup> O florescimento humano, conforme definido na literatura internacional, é um estado em que as pessoas, na maioria das vezes, experimentam emoções positivas, funcionamento psicológico positivo e funcionamento social positivo. Em termos mais filosóficos, isso significa acesso à vida agradável, à vida comprometida ou boa e à vida significativa (NORRIS, 2010).

como um elemento essencial. A ênfase está em ampliar as possibilidades de escolhas, não apenas melhorar os seus resultados. A renda não é mais o objetivo principal do desenvolvimento, mas um meio para promover o DH, que, por sua vez, vai além das necessidades básicas por incluir aspectos do florescimento humano e não se limitar aos países em desenvolvimento (STEWART, 2019).

## **2.3 CRESCIMENTO ECONÔMICO E DESENVOLVIMENTO HUMANO**

A partir das novas perspectivas sobre o conceito de desenvolvimento que passa a incluir uma nova dimensão, o desenvolvimento humano, discutidas na seção anterior, fica evidente a necessidade de se revisitar a relação entre CE e DH. Ranis, Stewart e Ramirez (2000) afirmam que, apesar da conexão entre CE e DH ser evidente, a relação entre esses dois componentes é complexa. De acordo com os autores, a evidência pode ser assim expressa: o CE disponibiliza os recursos que possibilitam melhorias no DH, por outro lado, uma melhor força de trabalho impacta positivamente o CE. Essas alternativas refletem a dinâmica entre crescimento e desenvolvimento, que resulta em duas cadeias possíveis de causalidade: uma refere-se a que o CE leva ao DH (cadeia A); a outra (cadeia B), que melhorias no DH implicarão em CE (RANIS; STEWART, 2005).

Esse raciocínio, conhecido na literatura econômica como “causação circular cumulativa”, pode ser atribuído aos estudos de Myrdal (1968). Segundo o autor baixos níveis de escolaridade ou de saúde são associados a maiores níveis de desigualdade socioeconômica, na medida em que conduzem a salários mais baixos e menor produtividade, que acabam reforçando a incapacidade dos indivíduos e de suas famílias de melhorarem de vida.

Independente dessa influência causal circular, o que Ranis e Stewart (2005) concluíram sobre essa evidência é que a relação de causalidade entre elas não era unilateral, mas bilateral. Mostraram isso usando dados de 69 países em desenvolvimento e concluíram que o crescimento econômico é um contributo importante para a melhoria do desenvolvimento humano, mas não é em si sustentável sem investimentos em saúde e educação, tanto prévia como simultaneamente (RANIS; STEWART, 2005). Esse resultado mostra a importância de co-evolução das

variáveis econômica, de um lado, e de outro, as de desenvolvimento humano (saúde e educação).

Tal conclusão corrobora empiricamente a abordagem proposta anteriormente por Sen, Motta e Mendes (2000). Ao focar no conceito de capacitações (*capabilities*), eles colocam o desenvolvimento como uma medida de expansão de capacidades que se apresentam objetivamente para que o indivíduo possa exercer sua liberdade de escolha, que varia desde ser bem nutrido ou educado até atividades mais específicas, como andar de bicicleta. Para haver desenvolvimento dever-se-ia retirar as principais fontes de privação de liberdade, como: pobreza, carência de oportunidades econômicas, destituição social sistemática e negligência dos serviços públicos.

Ainda segundo os autores, não se poderia, por outro lado, desconsiderar a face política do desenvolvimento humano. De uma perspectiva interna do país, ele é um processo conflituoso, pois modificações no perfil produtivo e distributivo da economia alteram as estruturas políticas e sociais onde essas mudanças ocorrem. Quanto mais complexos e diversificados são os sistemas econômicos, maior o número de atores que se empenharão para apropriar-se da maior parte possível da renda, e galgar posições mais elevadas na estrutura de poder e bem-estar. Logo, a estrutura de apropriação e utilização dessa renda não é alheia à estrutura de estratificação social, o que reflete na própria produtividade do trabalho. Assim, a estrutura de poder enraizada (*embedded*) na estrutura social é que determina as formas de produção, utilização, apropriação da renda. Por essa visão, os determinantes internos do desenvolvimento humano de um país não são independentes de sua estrutura social, e qualquer mudança nas atividades econômicas ou no processo de acumulação de capital estará sujeita à resistência do sistema de dominação vigente.

As implicações de políticas para a promoção do desenvolvimento humano estão relacionadas a intencionalidade do Estado, requerida para resolução, inclusive de conflitos oriundos de interesses divergentes quanto às mudanças na estrutura produtiva. Tanto Myrdal (1968) e Ranis e Stewart (2005) quanto as escolas neoschumpeteriana e estruturalista convergem ao defender a intervenção do Estado na economia, pois reconhecem que o mercado não pode sozinho promover o desenvolvimento sob condições estruturais periféricas. O Estado deve ser flexível e capaz de se adequar aos desafios impostos pelas mudanças na conjuntura econômica mundial, promovendo e aprimorando o padrão de vida da população. Entretanto, o raio de manobra que o Estado possui para guiar e coordenar o desenvolvimento varia

ao longo do tempo, a depender de fatores externos, como a geopolítica internacional e sua inserção externa, e também de fatores internos, como a estrutura social e as disputas políticas e econômicas dos agentes pela renda nacional.

A interpretação institucionalista para o desenvolvimento humano credita às instituições a primazia no processo (NORTH, 1991). Entretanto, como destacado nesta corrente, existem as instituições boas e as más no sentido de gerar incentivos aos atores sociais, e as escolhas dos países, por um tipo ou outro de instituição, definem o sucesso ou o fracasso na empreitada do desenvolvimento humano. Segundo Acemoglu, Johnson e Robinson (2005, p. 397), as instituições "*[...] influence not only the size of aggregate pie, but how this pie is divided among different groups and individuals in society*". Por esse seu caráter duplo, as instituições podem ser consideradas endógenas ao sistema, uma vez que refletem a disputa de interesses dos agentes por uma maior fração do produto econômico.

Apesar de se reconhecer a importância das instituições no desenvolvimento humano, acredita-se que seria uma generalização afirmar que "*[...] differences in economic institutions are the fundamental cause of different patterns of economic growth*" (ACEMOGLU; JOHNSON; ROBINSON, 2005, p. 397). Isto, porque estaria se supondo que a maneira que os próprios humanos decidem por organizar suas sociedades determinaria se esta sociedade prosperaria ou não. A visão de desenvolvimento humano desta Tese não exclui a discussão das instituições e do papel do Estado no desenvolvimento, apenas não se faz dela o centro do debate, sendo mais uma questão de contorno do que sua causa fundamental.

### 3 METODOLOGIA

A taxonomia adotada segue o modelo de Vergara (1997), a partir da seguinte classificação de pesquisa: quanto aos fins de estudo; e quanto aos meios da investigação.

Quanto aos fins do estudo a pesquisa é tanto descritiva como explicativa. Descritiva no sentido de compreensão das características dos municípios do Estado do Rio de Janeiro em função de variáveis socioeconômicas, explicativa porque visa a elucidar quais fatores contribuem para o círculo vicioso da pobreza.

Em relação aos meios, trata-se de uma pesquisa, ao mesmo tempo, bibliográfica e empírica. Bibliográfica devido ao estudo sistematizado desenvolvido com bases em livros, artigos de periódicos, sítios de internet. Experimental, pois foi conduzida uma pesquisa empírica em que variáveis exploratórias foram controladas e os seus efeitos observados na variável dependente, que é o nível de desenvolvimento.

O universo do estudo refere-se aos municípios pertencentes ao Estado do Rio de Janeiro, no qual cada um deles foi considerado uma observação. O período da análise corresponde a 2010, pois há vários indicadores de interesse compilados pelo censo demográfico deste ano, último disponível. Os dados foram obtidos por meio do Atlas Brasil (2021) - que engloba o atlas de desenvolvimento humano dos municípios e regiões metropolitanas brasileiras. Também foi consultada a Relação Anual de Informações Sociais (RAIS), referente ao ano de 2010, para fins de cálculo do Quociente Locacional (QL)<sup>3</sup>.

---

<sup>3</sup> QL foi calculado com base na metodologia de Jorge Britto. Ver mais em: BRITTO, Jorge. **Arranjos produtivos locais**: perfil das concentrações de atividades econômicas no Estado do Rio de Janeiro. 2004. SEBRAE/RJ. Disponível em: <https://www.redetec.org.br/wp-content/uploads/2015/02/APLs.pdf>. Acesso em: 03 out. 2019.

Trabalhou-se no todo com 20 indicadores (Tabela 1), divididos nas áreas de desenvolvimento, saúde, renda, educação e habitação. Além disso, foi acrescentado o QL industrial para avaliar se o grau de concentração dessa atividade econômica influencia no desenvolvimento. A justificativa para introdução desse indicador é que os países desenvolvidos, em geral, apresentam uma maior densidade industrial.

Tabela 1. Grupos de variáveis que compõem o modelo

Grupos	Variável	Código
IDHM	IDHM Longevidade	IDHM_L
	IDHM Educação	IDHM_E
	IDHM Renda	IDHM_R
Desenvolvimento	Nível de desenvolvimento	ND
Saúde	Mortalidade infantil (%)	Mi
	Probabilidade de sobrevivência até 60 anos (%)	PS60
	Esperança de vida ao nascer	EV
QL	QL - Quociente Locacional	QL
Renda	PIB (Mil)	PIB
	Renda <i>per capita</i>	RP
	Índice de Gini	Gi
	% de pobres	P
	% de vulneráveis à pobreza	VP
Educação	% de 18 anos ou mais com fundamental completo	FC18
	% de 25 anos ou mais com superior completo	SC25
	Expectativa de anos de estudo	EAE
	Taxa de analfabetismo - 18 anos ou mais	T18
Habitação	% da população em domicílios com água encanada	PDA
	% da população em domicílios com banheiro e água encanada	PDBA
	% da população em domicílios com coleta de lixo	PDCL

Fonte: Autor.

As dimensões do IDHM (longevidade, saúde, educação) foram utilizadas somente para realizar a classificação da variável categórica Nível de Desenvolvimento. As demais foram utilizadas para verificar em que medida elas explicam o nível de desenvolvimento. As variáveis Mortalidade Infantil, Probabilidade de Sobreviver até 60 anos e Esperança de Vida funcionam como *proxy* para medir o impacto da saúde no nível de desenvolvimento dos municípios. O PIB e a Renda *per capita* mensuram a influência do fator econômico, enquanto o Índice Gini, % de pobres (renda domiciliar per capita igual ou inferior a R\$ 140,00 mensais) e % de vulneráveis

a pobreza (renda domiciliar *per capita* igual ou inferior a R\$ 255,00 mensais) captam a qualidade dos avanços na economia. A educação foi representada por quatro indicadores que podem influenciar o nível de desenvolvimento: Expectativa de anos de estudo, Taxa de analfabetismo de pessoas com 18 anos ou mais, % de 18 anos ou mais com fundamental completo e % de 25 anos ou mais com superior completo. No que diz respeito à estrutura habitacional, os seguintes indicadores refletem esse aspecto: % da população em domicílios com água encanada; % da população em domicílios com banheiro e água encanada; e % da população em domicílios com coleta de lixo.

As fontes de dados para todas as variáveis, exceto o QL, são baseadas no censo demográfico. Para o QL utilizou-se o número de empregos na atividade industrial fornecido pela RAIS.

As etapas da análise dos dados foram estruturadas em quatro fases principais: a primeira consiste na classificação dos municípios do Estado do Rio de Janeiro, conforme seu nível de desenvolvimento; a segunda diz respeito à análise exploratória fatorial, que visa a auxiliar na compreensão dos inter-relacionamentos dos vários indicadores selecionados; a terceira etapa consiste em quantificar o impacto dos indicadores para explicar o nível de desenvolvimento do círculo vicioso. Por fim, a quarta etapa refere-se à elaboração de mapas temáticos.

### **3.1 CLASSIFICAÇÃO DOS MUNICÍPIOS**

A metodologia adotada para analisar a evolução do desenvolvimento humano do ERJ segue a proposta por Raiher e Lima (2014). Os autores utilizaram o Índice de Desenvolvimento Humano Municipal (IDHM) de uma região em um período de tempo para analisar a sua evolução em termos de um comparativo intrarregional, comparação com a média dos municípios da região estudada, e inter-regional, comparação com a média dos municípios do Brasil. Embora existam inúmeras críticas de como é construído o IDH, ele permite identificar a real importância de cada dimensão no desenvolvimento humano e elencar as prioridades de políticas para se atingir um maior nível de bem-estar. Dessa forma, é um Índice que permite aderência à fundamentação teórica do desenvolvimento defendida nesta Tese.



De acordo com o Programa das Nações Unidas para o Desenvolvimento (PNUD), o IDH, criado em 1990, consiste em um índice com variação de 0 a 1, que mede de forma resumida o progresso do desenvolvimento humano a longo prazo a partir de três dimensões básicas: renda; educação; e saúde. Esse índice foi criado com o objetivo de oferecer uma alternativa em relação a um indicador muito utilizado: o Produto Interno Bruto (PIB) *per capita*, utilizado para medir o nível de desenvolvimento, que considerava apenas a dimensão econômica do desenvolvimento. Assim, o IDH atua como uma medida que amplia a perspectiva sobre o desenvolvimento humano sem, contudo, esgotar todos os aspectos de desenvolvimento. Os três pilares que constituem o IDH são mensurados da seguinte maneira, de acordo com o PNUD:

Uma vida longa e saudável (saúde) é medida pela expectativa de vida. O acesso ao conhecimento (educação) é medido por: i) média de anos de educação de adultos, que é o número médio de anos de educação recebidos durante a vida por pessoas a partir de 25 anos; e ii) a expectativa de anos de escolaridade para crianças na idade de iniciar a vida escolar, que é o número total de anos de escolaridade que uma criança na idade de iniciar a vida escolar pode esperar receber se os padrões prevaientes de taxas de matrículas específicas por idade permanecerem os mesmos durante a vida da criança. E o padrão de vida (renda) é medido pela Renda Nacional Bruta (RNB) *per capita* expressa em poder de paridade de compra (PPP) constante, em dólar, tendo 2005 como ano de referência (PNUD, 2020, p. 1).

De forma semelhante, o IDHM é composto por três indicadores do desenvolvimento humano: longevidade; educação; e renda. A metodologia aplicada ao cálculo desse índice adequa a metodologia do IDH global ao cenário brasileiro e à disponibilidade de dados nacionais. O IDHM também varia de 0 a 1 e, quanto mais próximo de 1, maior é o desenvolvimento humano. A Equação 1 mostra como esse índice é calculado.

$$IDHM = \sqrt[3]{IDHM_L * IDHM_E * IDHM_R} \quad (1)$$

em que:

IDHM<sub>L</sub>: Índice de Desenvolvimento Humano Municipal de Longevidade;

IDHM<sub>E</sub>: Índice de Desenvolvimento Humano Municipal de Educação;

IDHM<sub>R</sub>: Índice de Desenvolvimento Humano Municipal de Renda.

O IDHM pode, então, ser decomposto em suas três dimensões. Desse modo, a dimensão longevidade do IDHM é baseada na esperança de vida ao nascer e é demonstrada pela Equação 2.

$$IDHM_L = \frac{(V_{ob} - V_{mín})}{(V_{máx} - V_{mín})} \quad (2)$$

em que:

IDHM<sub>L</sub>: Índice de Desenvolvimento Humano Municipal de Longevidade;

V<sub>ob</sub>: valor da idade observada;

V<sub>mín</sub>: valor mínimo;

V<sub>máx</sub>: valor máximo.

A dimensão educação do IDHM é composta por indicadores de escolaridade da população adulta e de fluxo escolar da população jovem. O primeiro indicador refere-se ao percentual da população com idade igual ou superior a 18 anos, que possui o ensino fundamental completo. Já o segundo, é medido pela média aritmética do: i) percentual de crianças com 5 e 6 anos que frequenta a escola; ii) percentual dos jovens de 11 a 13 anos que frequentam os anos finais do ensino fundamental; iii) percentual dos jovens de 15 a 17 anos que possuem o ensino fundamental completo; iv) percentual de jovens de 18 a 20 anos que possuem o ensino médio completo. Assim, o IDHM Educação é demonstrado pela Equação 2.

$$IDHM_E = \sqrt[3]{(IE_{adulto} * IFE_{jovem} * IFE_{jovem})} \quad (3)$$

em que:

IDHM<sub>E</sub>: Índice de Desenvolvimento Humano Municipal de Educação;

IE<sub>adulto</sub>: índice de escolaridade da população adulta, obtido pela taxa de frequência e de conclusão dividida por 100;

IFE<sub>jovem</sub>: índice de fluxo escolar da população jovem, obtido pela média aritmética das quatro faixas etárias.

A dimensão renda do IDHM considera a renda *per capita* da população, isto é, a renda média mensal dos indivíduos, expressa em reais.

$$IDHM_R = \frac{[\ln(R_{per\ capita}) - \ln(V_{min})]}{[\ln(V_{max}) - \ln(V_{min})]} \quad (4)$$

em que:

IDHM<sub>R</sub>: Índice de Desenvolvimento Humano Municipal de Renda;

$R_{per\ capita}$ : renda *per capita* do local de referência;

$V_{min}$ : valor mínimo de referência;

$V_{max}$ : valor máximo de referência.

Raiher e Ferreira de Lima (2014) seguiram a metodologia do PNUD (2002). Nela propôs-se a classificação das unidades federativas do México, segundo seu perfil de desenvolvimento, tomando como parâmetro o valor médio obtido em cada dimensão do IDH. Assim, se uma região apresentasse os indicadores renda, longevidade e educação maior do que a média regional (e nacional), a região encontrar-se-ia em um círculo virtuoso (CVir) de desenvolvimento. Mas se essas três dimensões estivessem abaixo da média regional (e nacional), a região estaria em um círculo vicioso (CVic) de desenvolvimento.

Duas situações intermediárias seriam: região em que apenas a renda fosse superior à média regional (e nacional) e região em que apenas a longevidade e a educação fossem superiores à média regional (e nacional). Essas situações classificaram o desenvolvimento como tendendo ao crescimento (TC) e tendendo ao desenvolvimento (TD). Essa classificação mostra que o mecanismo de causalção circular cumulativa trabalha no sentido de reforçar o desenvolvimento (CVir), mas também no sentido de reforçar o subdesenvolvimento (CVic). Além disso, evidencia que crescimento sem investimento em educação e saúde não leva ao desenvolvimento e que indicadores de longevidade e educação acima dos indicadores de renda não leva ao desenvolvimento sustentável.

Foram coletados dados do IDHM para os municípios do Estado do Rio de Janeiro (RJ), no período de 2010, desagregados em três dimensões: renda; longevidade; e educação. Posteriormente, os municípios foram classificados quanto ao seu estágio de desenvolvimento, conforme Tabela 2. Além das quatro

classificações possíveis supracitadas, alguns municípios foram classificados como “outros” por não se encaixarem em nenhuma das situações. Para comparar os municípios que são bastante desiguais em termos de tamanho populacional, foi feita uma normalização dos dados pela população dos municípios.

**Tabela 2.** Classificação dos municípios quanto ao desenvolvimento.

CATEGORIA	REGRA (IDHM)
1. Círculo virtuoso (CVir)	IDHM-Renda, IDHM-Longevidade e IDHM-Educação > média do ERJ
2. Círculo vicioso (CVic)	IDHM-Renda, IDHM-Longevidade e IDHM-Educação < média do ERJ
3. Tendendo ao crescimento (TC)	IDHM-Renda > média do ERJ
4. Tendendo ao desenvolvimento (TD)	IDHM-Longevidade e IDHM-Educação > média do ERJ
5. Outros	Qualquer município que não se enquadre em uma das quatro categorias

Fonte: Adaptado de Raiher e Ferreira de Lima (2014).

### 3.2 ANÁLISE FATORIAL MÚLTIPLA (MFA)

Após a classificação dos municípios montou-se um banco de dados contendo vários indicadores (Tabela 1) separados por *clusters*. São no total 17 indicadores e seis grupos: o primeiro refere-se ao nível de desenvolvimento dos municípios; o segundo contém variáveis relacionadas à saúde; o terceiro refere-se ao Quociente Locacional; o quarto a indicadores de renda; o quinto a indicadores educacionais; o sétimo representa a habitação. A estruturação dos grupos teve como base um padrão já definido pelo Programa das Nações Unidas para o Desenvolvimento, Instituto de Pesquisa Econômica Aplicada e Fundação João Pinheiro, disponível no Atlas Brasil.

A MFA foi construída a partir desse banco de dados, seguindo os seguintes passos:

- I. Uso do método de Decomposição dos Valores Singulares (*singular-value decomposition*, SVD) para cada grupo da Tabela 2 e posteriormente divisão de cada indicador pelo respectivo valor singular de seus grupos;
- II. Concatenação de todos os *clusters* horizontalmente, obtendo:

$$\tilde{Z}_{MFA} = [\tilde{Z}_{desenvolvimento} || \tilde{Z}_{saúde} || \tilde{Z}_{QL} || \tilde{Z}_{renda} || \tilde{Z}_{educação} || \tilde{Z}_{habitação}]$$

Aplicação do PCA (não normalizada) sobre a matriz  $\tilde{Z}_{MFA}$  no qual resulta-se no chamado  $F_{MFA}$  de configuração de produto de consenso.

A metodologia empregada nesta Tese tem como base o trabalho de Husson, Lê e Pagès (2017) e Kassambara (2017). A operacionalização dos cálculos da MFA foi por meio do *software* R, bibliotecas FactoMineR e factoextra. A primeira biblioteca, desenvolvida gera os resultados, e a segunda extrai e visualiza os dados.

### 3.2.1 Análise de Componentes Principais (PCA)

A análise de componentes principais (PCA) é um método multivariado aplicado a uma tabela de dados, cujas observações são descritas por variáveis exploratórias quantitativas inter-relacionadas (ABDI; WILLIAMS, 2010). Tem como objetivo extrair padrões importantes dos dados e representá-los como um conjunto de novas variáveis ortogonais, chamadas de componentes principais. Segundo Hair *et al.* (2009), é uma técnica estatística utilizada para analisar as relações entre um grande número de variáveis e explicá-la em função de suas dimensões (fatores), condensando a informação contida nas variáveis originais em um conjunto menor (fatores) com perda mínima de informação.

As notações e prenotações a seguir foram oriundas do trabalho de Abdi e Williams (2010). Matrizes são indicadas por letras maiúsculas e em negrito, vetores em letras minúsculas em negrito, os elementos denotados em minúsculo e itálico. Matrizes, vetores e elementos de uma mesma matriz usam a mesma letra, por exemplo, (**A**, **a**, *a*). A transposta de uma matriz é indicada pelo subscrito <sup>T</sup>, e a matriz identidade por **I**.

A tabela de dados analisada pelo PCA compreende observações *I* descritas por variáveis *J*, sendo representadas pela matriz *I* x *J* (**X**), cujo elemento genérico é representado por  $x_{i,j}$ . A matriz **X** tem um posto *L*, onde  $L \leq \min\{I, J\}$ .

De maneira geral os dados são reprocessados para que cada coluna da matriz *X* tenha média 0 e desvio padrão igual a 1. A matriz *X* apresenta o seguinte valor de decomposição singular (SVD, ver Abdi e Williams, (2010), referências 11-13 e Apêndice B para introdução ao SVD):

$$X = P\Delta Q^T \quad (5)$$

onde  $P$  é uma  $I \times L$  matriz de vetores singulares a esquerda,  $Q$  é  $J \times L$  matriz de vetores singulares a direita e  $\Delta$  a matriz diagonal de valores singulares.

O PCA permite calcular novas variáveis chamadas *componentes principais*, obtidas pela combinação linear das variáveis originais, de maneira que a primeira componente tenha a maior variação/inércia possível, ou seja, irá explicar a maior parte da informação contida na tabela de dados. A segunda componente é calculada sob a restrição da ortogonalidade à primeira componente principal, de maneira a obter a maior inércia possível; os demais componentes são calculados da mesma forma. Os valores dessas novas variáveis são denominados de pontuações fatorais (*factor scores*) para as observações, sendo interpretadas de maneira geométrica, como a projeção das observações nos componentes principais.

As componentes são geradas pelo método SVD sobre os dados da tabela  $X$ . Como  $X = P\Delta Q^T$  a  $I \times L$  matriz dos *factor scores*, denotada por  $F$  é obtida:

$$F = XQ \quad (6)$$

A matriz  $Q$  fornece os coeficientes (autovetores) das variáveis para a combinação linear, também conhecida como matriz de carregamento *loading matrix*, ou matriz de projeção, pois a multiplicação de  $X$  por  $Q$  resulta na projeção das observações nos componentes principais. De acordo com Dunteman (1989), as componentes também podem ser expressas algebricamente por:

$$F_p = \sum_{j=1}^J q_{pj}x_j \quad (7)$$

As variáveis diferentes das observações não são representadas por suas projeções, mas, sim, por suas correlações. São plotadas sobre um círculo de correlações usando como coordenadas os *loadings*. Segundo Dunteman (1989), quando os elementos dos vetores latentes (peso das variáveis) são transformados em correlação das variáveis com a componente principal a interpretação fica mais fácil. Isso pode ser feito multiplicando os autovetores ( $q$ ) pela raiz quadrada do  $i$ -ésimo autovalor ( $\sqrt{\lambda_i}$ ). Desse modo as coordenadas das variáveis para cada componente são obtidas pela seguinte fórmula:  $\sqrt{\lambda_i}q_i$ ; ou matricialmente, a partir do SVD, expressa por:

$$C = \frac{Q\Delta}{\sqrt{I-1}} \quad (8)$$

A proporção de variância explicada por uma componente representa a sua importância. Esta pode ser mensurada dividindo o  $i$ -ésimo autovalor  $\lambda_i$  pelo seu somatório  $\sum \lambda$ , ou, se os dados foram padronizados, apenas dividir o autovalor pelo número de variáveis  $\lambda_i/J$  (DUNTEMAN, 1989).

Essa técnica pode ser generalizada para originar outros métodos: análise de correspondência simples (CA) e múltipla (MCA), assim como, escalonamento multidimensional (MDS) e análise de múltiplos fatores (MFA).

### 3.2.2 Descrição da MFA

A MFA é o método de análise de dados multivariada utilizado para resumir e visualizar tabelas de dados complexas na qual os indivíduos são descritos por diversos conjuntos de variáveis, seja essas quantitativas e/ou qualitativas. Para definir a distância dos indivíduos leva-se em consideração a contribuição de todos os grupos de variáveis (PAGÈS, 2014). Esses grupos, por sua vez, podem diferir em relação ao número e natureza (quantitativa ou qualitativa) das variáveis, no entanto, em cada agrupamento as variáveis devem ser da mesma natureza (ABDI; WILLIAMS, 2010).

A ideia principal da MFA se assemelha à normalização z-score, que torna as variáveis comparáveis pela divisão de cada elemento pelo desvio-padrão da respectiva variável. De forma análoga na Análise de Componentes Principais (PCA), o valor singular (raiz quadrada de um autovalor) funciona como um desvio padrão. Dessa forma, na MFA, cada tabela de dados – grupos de variáveis – é normalizada dividindo todos os seus elementos pelo primeiro valor singular dessa tabela. Isso assegura que nenhum grupo domine a solução por ter uma maior inércia na sua primeira dimensão (TOMIC; BERGET; NÆS, 2015). Uma vez normalizado cada conjunto de dados, realiza-se a combinação desses grupos em uma representação de observações em comum, que pode ser chamado de compromisso ou consenso (ABDI; WILLIAMS; VALENTIN, 2013). Na MFA esse consenso  $F_{MFA}$  pode ser obtido da seguinte forma (TOMIC; BERGET; NÆS, 2015):

- I. Usar o método SVD para cada bloco  $Z_k$  e posteriormente dividir cada  $Z_k$  pelo seu respectivo valor singular  $\gamma_{1,k}$ , obtendo:

$$\tilde{Z}_k = \gamma_{1,k}^{-1} * Z_k \quad (9)$$

II. Concatenar todos os  $\tilde{Z}_k$  horizontalmente, obtendo:

$$\tilde{Z}_{MFA} = [\tilde{Z}_1 | \tilde{Z}_2 | \dots | \tilde{Z}_k | \dots | \tilde{Z}_K] \quad (10)$$

III. Aplicar análise de componentes principais (PCA) - não normalizada - sobre a matriz  $\tilde{Z}_{MFA}$  no qual resulta-se no chamado  $F_{MFA}$  de configuração de produto de consenso.

Dessa forma, a PCA decompõe a variância do consenso em um conjunto de novas variáveis ortogonais, ditas, componentes principais (também chamadas de dimensões, eixos, fatores, até mesmo variáveis latentes), ordenadas de acordo com a sua quantidade de variância explicada. Por meio desses componentes principais são extraídas as coordenadas das observações (também conhecida como *scores*), usadas para traçar os mapas fatoriais, onde as observações são representadas como pontos e a distância entre os pontos refletem a similaridade entre as observações. As componentes são obtidas a partir da combinação das variáveis originais. Assim, cada variável “contribui” com uma determinada quantidade para formação dos eixos principais. Essa quantidade é chamada de carga fatorial (*loading*), que indica a importância da variável para um dado componente principal (ABDI; WILLIAMS; VALENTIN, 2013)<sup>4</sup>.

### 3.3 MENSURAÇÃO DO IMPACTO DAS VARIÁVEIS SOCIOECONÔMICAS NA COMPOSIÇÃO DO CÍRCULO VICIOSO

Para estimar a probabilidade de um município adentrar no círculo vicioso de desenvolvimento, bem como verificar o impacto das variáveis preditoras, foi utilizado o modelo de regressão binomial. No entanto, devido à natureza subjacente das variáveis analisadas, a multicolinearidade se faz presente (VIF acima de 100 para algumas variáveis). Para corrigir esse problema foi utilizado o método de penalidade de Ridge. Segundo Frank e Friedman (1993), essa técnica demonstrou desempenho equivalente entre as abordagens de regularização.

---

<sup>4</sup> Para mais detalhes da MFA ver Escofier e Pages (1994); e Abdi, Williams e Valentin. (2013).



No modelo a variável dependente foi a categórica ND, que representa o nível de desenvolvimento dos municípios. Essa variável possui cinco níveis, no entanto, é necessária uma conversão para adequação ao modelo logístico binomial. Assim, se um município foi classificado como vicioso, recebe valor “1” caso contrário “0”. A equação (11) representa o modelo teórico. As variáveis explicativas  $x^T$  foram aquelas apresentadas na Tabela 2 (exceto a ND e as subdimensões do IDH) e  $\beta$  os seus coeficientes.  $Pr(G = 1|x)$  foi a probabilidade condicional de um município ser do ciclo vicioso dado as variáveis independentes e  $Pr(G = 0|x)$  a probabilidade contrária a esse evento.

$$\ln \left[ \frac{Pr(G = 1|x)}{Pr(G = 0|x)} \right] = \beta_0 + x^T \beta. \quad (11)$$

Os coeficientes foram estimados pelo método da máxima verossimilhança regularizada.

$$\max_{(\beta_0, \beta) \in \mathbb{R}^{17}} \left[ \frac{1}{92} \sum_{i=1}^{92} \{y_i \cdot (\beta_0 + x_i^T \beta) - \log(1 + e^{(\beta_0 + x_i^T \beta)})\} - 0,037939 P_0(\beta) \right] \quad (12)$$

$P_\alpha$  é a penalização da rede elástica, sendo um compromisso entre a penalidade de Ridge ( $\alpha = 0$ ) e Lasso ( $\alpha = 1$ ). Os valores de  $\alpha$  e  $\lambda$  foram respectivamente 0 e 0,037939,  $\alpha$  igual a zero para indicar a penalidade L2, já a estimação do  $\lambda$  foi com base no método de Cule e Iorio (2013), técnica semiautomático para orientar a escolha do parâmetro de contração na regressão de ridge.

$$P_0(\beta) = \sum_{j=1}^{16} \left[ \frac{1}{2} (1 - 0) \beta_j^2 + 0 |\beta_j| \right] \quad (13)$$

A metodologia empregada nesta seção foi baseada no trabalho de Cule e Iorio (2013). A biblioteca Ridge disponíveis no *software* R, auxiliou nos cálculos matemáticos.

### 3.3.1 Descrição do modelo de regressão penalizada

Dados de alta dimensionalidade, que geralmente consistem em um número maior de preditores do que observações, além de uma estrutura correlacionada entre

variáveis, não se adequa ao modelo tradicional de regressão múltipla (CULE; IORIO, 2013).

A multicolinearidade ocorre quando há uma relação perfeita entre as variáveis, ou seja, quando os preditores são altamente correlacionados. Esse fenômeno faz com que os coeficientes se tornem estáveis e os erros-padrões inflem, dificultando a chegada de estimativas confiáveis, o que pode levar a conclusões incorretas de pesquisa (SCHREIBER-GREGORY, 2018). Não são um empecilho para modelos de predição, porém são especialmente problemáticos para modelos explicativos (GUJARATI; PORTER, 2011; SCHREIBER-GREGORY, 2018). Em relação à explicação, em que a ênfase está no efeito, torna-se difícil obter a significância dos diferentes parâmetros em modelos que contenham variáveis colineares, aumentando a chance de erro tipo II. Enquanto na previsão, independente das estimativas apresentarem significância estatística, provavelmente a soma dos parâmetros serão confiáveis para predição. Esse problema pode ser facilmente contornado em uma previsão de teste do modelo ao aumentar o tamanho da amostra. No entanto, se a colinearidade for encontrada nos modelos explicativos são necessárias medidas mais complexas para resolvê-la caso a exclusão das variáveis não seja uma opção, dado que uma forma simples de resolver o problema da multicolinearidade seria por meio da exclusão das variáveis redundantes do modelo (TABACHNICK; FIDELL, 2013)

Determinadas variáveis, apesar de altamente correlacionadas, precisam ser mantidas no modelo, pois são importantes para o objetivo da pesquisa. Nesse sentido, outros métodos devem ser explorados para corrigir esse problema (SCHREIBER-GREGORY, 2018). O uso de técnicas de regularização, também chamada de penalização, busca minimizar o problema da multicolinearidade (GOEMAN, 2010; SCHREIBER-GREGORY, 2018).

As regularizações L1 (Lasso) e L2 (Rigde) reduzem as estimativas dos coeficientes de regressão para zero em comparação aos parâmetros estimados pela máxima verossimilhança. A função desse encolhimento é prevenir excessos oriundos da colinearidade das variáveis ou alta dimensionalidade dos dados; embora os métodos L1 e L2 apresentem o mesmo objetivo, os efeitos da penalização são diferentes (GOEMAN, 2010).

A regularização L2 foi proposta originalmente para calcular os coeficientes com menor erro quadrático médio quando as variáveis explicativas são correlacionadas (HOERL; KENNARD, 1970). A aplicação da regularização de Rigde tende a reduzir

os coeficientes a valores assíntotos a zero. A magnitude quadrada do coeficiente é adicionada como termo de penalidade à função custo, equação (13) (SCHREIBER-GREGORY, 2018).

Se o valor da penalidade,  $\lambda$  – lambda, for zero a estimativa dos coeficientes será a mesma da regressão dos mínimos quadrados. Caso o valor de lambda seja muito grande, o peso adicionado ao modelo resultaria em um ajuste insuficiente. Ao tender o valor de  $\lambda$  ao infinito as estimativas dos parâmetros convergem para valores próximos a zero. (HOERL; KENNARD, 1970, LE CESSIE; VAN HOUWELINGEN, 1992).

$$\sum_{i=1}^n \left( Y_i - \sum_{j=1}^p X_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (13)$$

Enquanto para a regularização L1<sup>5</sup>, adiciona-se à função custo o valor absoluto da magnitude do coeficiente, equação (14). A diferença principal é que no método de Lasso as estimativas dos parâmetros menos importantes são reduzidas a zero, removendo, assim, variáveis menos importantes (TIBSHIRANI, 1996; GOEMAN, 2010; TIBSHIRANI, 2011). Tal método se torna muito adequado quando empregado para seleção de variáveis, todavia não é recomendada quando o objetivo for corrigir a multicolinearidade (FONTI, 2017; SCHREIBER-GREGORY, 2018).

$$\sum_{i=1}^n \left( Y_i - \sum_{j=1}^p X_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (14)$$

Ambos os métodos de penalização têm desvantagens. L1 fornece modelos interpretáveis, mas pode variar muito na seleção dos subconjuntos por ser um processo discreto em que pequenas alterações nos dados podem levar à seleção de modelos diferentes. Na L2 os coeficientes são reduzidos por um processo contínuo, logo, mais estável, todavia nenhum parâmetro chega a zero, o que torna a interpretação do modelo mais difícil (SCHREIBER-GREGORY, 2018).

---

<sup>5</sup> Para mais detalhes ver (TIBSHIRANI, 1996; TIBSHIRANI, 2011).

Há também a técnica de penalização rede elástica (*elastic-net penalty*). Trata-se da combinação entre a regularização de Lasso e Ridge, desenvolvido por Zou e Hastie (2005). A rede elástica realiza ao mesmo tempo a seleção das variáveis, assim como a regularização contínua dos coeficientes, podendo selecionar grupos de variáveis correlacionadas (ZOU; HASTIE, 2003; 2005).

Inicialmente os métodos de penalização foram aplicados à regressão linear e posteriormente estendidos para a regressão logística, que também apresenta problemas com dados de altas dimensões e variáveis altamente correlacionadas. A regressão logística é utilizada quando a variável resposta é dicotômica ou binária. Muitos fenômenos sociais são discretos ou qualitativos e podem ser representados como ocorrência ou não de um evento (PAMPEL, 2000). A regressão logística assume os seguintes pressupostos: relação linear entre uma variável contínua e os resultados transformados em logit; não influência de valores extremos em preditores contínuos; ausência de multicolinearidade (STOLTZFUS, 2011).

A relação entre um conjunto de variáveis explicativas em função de resultados categóricos pode ser dada pela equação (12). Segundo Zou e Hastie (2005),  $G$  é a variável resposta, assume valores  $G = \{0,1\}$ , 1 para ocorrência de um evento e 0 para não ocorrência;  $\beta_0$  constante escalar,  $\beta$  um vetor de coeficientes correspondentes às variáveis exploratórias  $x^T$ .

$$\ln \left[ \frac{Pr(G = 1|x)}{Pr(G = 0|x)} \right] = \beta_0 + x^T \beta. \quad (12)$$

Tomando como base o algoritmo de Zou e Hastie (2005) no modelo de regressão logística, os coeficientes são estimados pelo método da máxima verossimilhança regularizada, equação (13).

$$\max_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} \left[ \frac{1}{N} \sum_{i=1}^N \{y_i \cdot (\beta_0 + x_i^T \beta) - \ln(1 + e^{(\beta_0 + x_i^T \beta)})\} - \lambda P_\alpha(\beta) \right] \quad (13)$$

$P_\alpha$  é a penalização da rede elástica, sendo um compromisso entre a penalidade de Ridge ( $\alpha = 0$ ) e Lasso ( $\alpha = 1$ ). Caso adote o alfa igual 0 será aplicada a

regularização L2; se alfa igual a 1 a penalização L1 e; alfa em entre 0 e 1 a combinação de L1 e L2 (rede elástica).

$$P_{\alpha}(\beta) = \sum_{j=1}^p \left[ \frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta_j| \right] \quad (14)$$

Para detalhes da aplicação de técnicas de penalização em modelo binomial ver os trabalhos de Cessie e Van Houwelingen (1992); Zou e Hastie (2003; 2005); e Ravikumar *et al.* (2010).

### 3.4 MAPAS TEMÁTICOS

Os mapas são representações gráficas tanto do ambiente cultural como físico, os quais subdividem-se em duas subclasses: mapas de uso geral e mapas temáticos. O mapa temático mostra a distribuição espacial de um fenômeno geográfico (DENT; TORGUSON; HODLER, 2008). De acordo com Joly (2005, p. 75) mapas temáticos são “[...] todos os mapas que representam qualquer tema, além da representação do terreno

Esses mapas são subdivididos em dois grupos, qualitativos e quantitativos, ambos apresentados nesta Tese. Os mapas qualitativos representam a distribuição espacial de dados nominais; por outro lado, os quantitativos os aspectos espaciais de dados numéricos, com foco na variação das variáveis de um lugar para outro (DENT; TORGUSON; HODLER, 2008).

O mapa qualitativo, exibido no item 4.2, apresenta os níveis de desenvolvimento dos municípios fluminenses, variável categórica ND. Os mapas quantitativos ilustram a variação espacial dos indicadores socioeconômicos presentes na Tabela 2, exceto as subdimensões do IDH. Essas variáveis foram ordenadas em cinco classes, de acordo com o método de otimização de Jenks (quebras naturais), que minimiza a variância entre as faixas. As cores aumentam de intensidade conforme valores apresentados nas classes estabelecidas seguem o gradiente espectral, variando de vermelho a azul (a cor vermelha, aqui, sempre indica uma região crítica).

Além da elaboração de um mapa para cada indicador, foi gerado um mapa com a combinação linear das variáveis significativas no modelo de regressão, conforme

modelo proposto por Bonham-Carter (1994). Segundo o autor, a combinação espacial de dados de diversas fontes é o objetivo final de grande parte de projetos SIG, com a finalidade de descrever e analisar interações, fazer previsões e auxiliar no processo de tomada de decisão. O escore médio é definido por:

$$\bar{S} = \frac{\sum_i^n S_{ij} W_i}{\sum_i^n W_i} \quad (15)$$

onde  $\bar{S}$  é a pontuação ponderada de um objeto de área (pixel ou polígono);  $W_i$  é o peso do i-ésimo mapa de entrada;  $S_{ij}$  é a pontuação da j-ésima classe do i-ésimo mapa, o valor de j dependendo da classe que realmente ocorre no local atual.

Nesta pesquisa, o objeto de área foi um polígono que representa as divisões municipais;  $W_i$  os pesos dos indicadores, com valor 1, ou seja, igual ponderação;  $S_{ij}$  o valor (normalizado) do respectivo indicador para um dado município. Para não ocorrerem distorções devido a escalas diferentes, as variáveis foram normalizadas dentro de um intervalo de 0 a 1, antes da combinação dos mapas, de acordo com a seguinte fórmula:

$$V_i^* = \frac{V_i - V_{max}}{V_{max} - V_{min}} \quad (16)$$

onde  $V_i$  e  $V_i^*$  são respectivamente o i-ésimo valor original e o transformado da variável;  $V_{max}$  e  $V_{min}$  o valor máximo e mínimo respectivamente dos dados brutos.

Após a sobreposição obteve-se o mapa ponderado. Esse foi normalizado novamente e classificado de acordo com as faixas do IDH proposta pelo PNUD (2020).

O *software* utilizado foi o QGIS, um Sistema de Informação Geográfica (SIG) de código aberto licenciado segundo a Licença Pública Geral GNU (QGIS, 2020), desenvolvido a partir de *software* Livre e de Código Aberto, *Free and Open Source Software (FOSS)*. É um projeto oficial da *Open Source Geospatial Foundation (OSGeo)*, com suporte a uma variedade de formatos de vetores, rasters, base de dados e funcionalidades, e funciona nas plataformas Linux, Unix, Mac OSX, Windows e Android. O Projeto foi criado em maio de 2002, e é desenvolvido usando o kit de ferramentas Qt (<https://www.qt.io>) e C ++.

## 4 RESULTADOS

Os resultados foram divididos em cinco subseções. A primeira apresenta a análise descritiva dos dados; a segunda, as classificações dos municípios quanto ao nível de desenvolvimento; a terceira, a análise exploratória fatorial; a quarta, o modelo de regressão logística; e a quinta, os mapas temáticos.

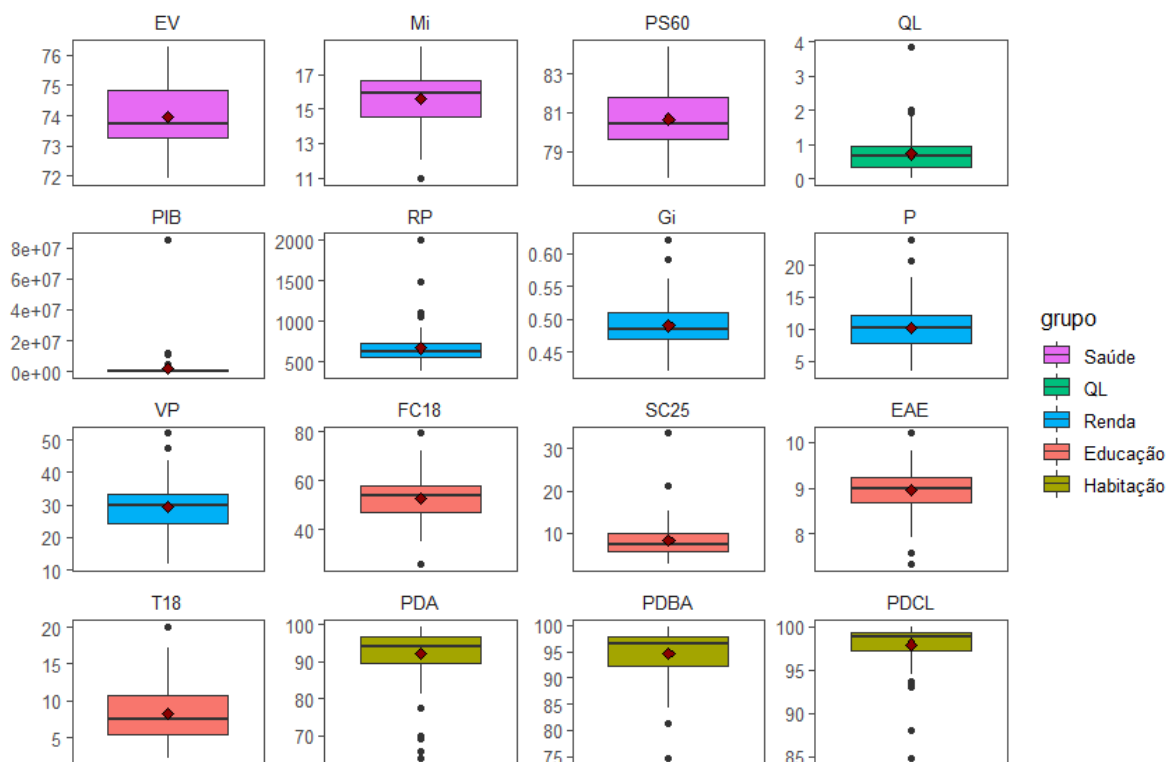
### 4.1 ANÁLISE EXPLORATÓRIA

Nesta subseção é apresentada uma análise descritiva das variáveis por meio de medida de tendência central, separatriz (quartis), valores máximos e mínimos. Além disso, também foi calculada a associação entre as variáveis quantitativas usando o coeficiente de correlação de Pearson ( $\alpha=0,05$ ), sendo complementado por uma análise de escalonamento multidimensional (MDS) e agrupamento hierárquico – cujo objetivo é facilitar a compreensão das interrelações entre os indicadores socioeconômicos.

O diagrama de caixa, assim como a média das variáveis analisadas, é apresentado na Figura 3. Nesse tipo de gráfico é possível visualizar os quartis, a medida separatriz, dos indicadores; isso possibilita a compreensão da dispersão dos dados. O primeiro quartil Q1 separa 25% dos dados; o Q2, mediana, divide os dados igualmente, 50%; o Q3 separa os dados em 75%. Respectivamente valores acima ou abaixo das hastes superior e inferior indicam pontos discrepantes (*outlier*). O limite

inferior é o  $\max\{\min(dados), Q_1 - 1,5(Q_3 - Q_1)\}$  e o superior  $\min\{\max(dados), Q_3 + 1,5(Q_3 - Q_1)\}$ .

Figura 3. Diagrama de caixas e médias das variáveis avaliadas.



Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda per capita; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Fonte: Autor

Por exemplo, observa-se que a mediana do indicador esperança de vida (EV) foi de 73,74 anos, isso significa que pelo menos 50% dos municípios do ERJ apresentaram o indicador abaixo desse valor, próximo a média de 73,96 anos. O primeiro quartil (Q1) indicou que ao menos 25% dos municípios têm índices de expectativa de vida de no máximo 73,26 anos, sendo assim, 75% estão acima desse valor. O terceiro quartil (Q3) apontou que pelo menos 75% das localidades apresentaram o indicador de EV inferior ou igual a 74,83 anos. De maneira complementar, 25% dos municípios têm uma expectativa de vida igual ou superior a Q3. Ao verificar a dispersão do conjunto de dados por meio do intervalo interquartil (Q3-Q1), constatou-se uma variação de 1,57 anos. Essa medida é menos sensível a



*outlier*, os valores mínimo e máximo foram respectivamente 71,93 e 76,27 anos. Na Tabela 3 encontra-se a estatística descritiva das variáveis.

Nota-se que as variáveis do grupo saúde não demonstraram valores extremos, exceto pelo indicador Mi com um *outlier* na parte inferior no valor de 10,96%, correspondente ao município de Rio das Ostras. Em relação à variável QL, três localidades apresentaram um índice de quociente locacional acima dos demais municípios, sendo estes Porto Real (3,85), Nova Friburgo (1,99) e Bom Jardim (1,93). No grupo renda todos os indicadores apontaram pontos extremos, com atenção para variável PIB com 14 pontos. Dentre eles destaca-se o município do Rio de Janeiro, com PIB de R\$ 84.888.946 (Mil), muito acima das outras localidades. Quatro municípios tiveram renda *per capita* acima do limite superior, Niterói (R\$ 2.000,29), Rio de Janeiro (R\$ 1.492,63), Macaé (R\$ 1.103,42) e Rio das Ostras (R\$ 1.051,19). Por outro lado, no índice Gini as cidades de Rio de Janeiro (0,62) e Niterói (0,59) apresentaram valores discrepantes. Dessa forma, observa-se que aquelas cidades com maiores RP foram as mais desiguais. No que diz respeito às variáveis relativas à pobreza (P e VP), ambas denotaram os mesmos municípios como *outlier*, São Francisco de Itabapoana (23,92%; 52,09%) e São José de Ubá (20,63%; 47,54%).

As variáveis que pertencem ao grupo educação também apresentaram valores extremos, tanto acima como abaixo do limite. O indicador que mensura a porcentagem de indivíduos com 18 anos ou mais que possuem fundamental completo foi composto por dois indivíduos discrepantes: um acima do limiar – Niterói (79,35) – e outro abaixo – Sumidouro (25,74). Os maiores níveis de educação superior completa (SC25) são de dois *outlier*, representados por Niterói (33,68%) e Rio de Janeiro (21,26%). Três municípios possuíram valores de expectativa de anos de estudo (EAE) fora dos limites inferior e superior, Sumidouro e Cardoso Moreira com a menor expectativa de anos de estudo, respectivamente 7,33 anos e 7,58 anos, e Natividade, com a maior (10,19 anos). São Francisco do Itabapoana foi considerado um ponto extremo ao se analisar o indicador taxa de analfabetismo (T18), com a maior taxa (20,8%) de pessoas com 18 anos ou mais analfabetas. Por fim, o grupo habitação foi composto por três variáveis, todas com valores discrepantes abaixo do limite inferior. Em relação à variável água encanada (PDA) foram encontrados cinco valores extremos. O menor deles foi de 63,9%, correspondente à localidade de Mangaratiba. Enquanto no indicador que representa banheiro e água encanada (PDBA), dois municípios estiveram abaixo do limiar inferior, Silva Jardim (74,75%) e Maricá (83,26%). Em

relação à variável coleta de lixo (PDCL), seis cidades foram identificadas como *outlier* inferior. Japeri (84,86%) teve a menor taxa.

Tabela 3. Estatísticas descritivas das variáveis de estudo

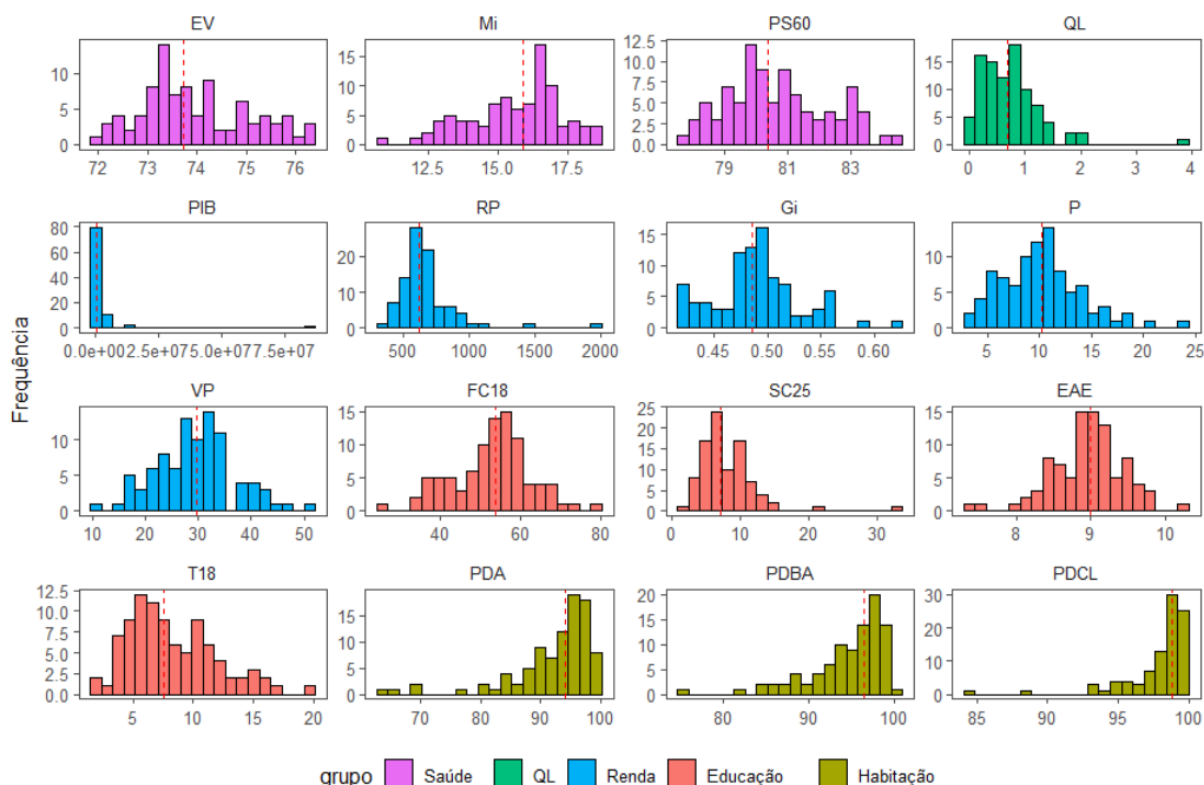
	<b>EV</b>	<b>Mi(%)</b>	<b>PS60(%)</b>	<b>QL</b>	<b>PIB (Mil)</b>	<b>RP(R\$)</b>	<b>Gi</b>	<b>P(%)</b>
Min.	71,93	10,96	77,59	0,01352	32.920	375,5	0,42	3,34
1st Qu.	73,26	14,56	79,61	0,35127	91.684	559,5	0,47	7,7
Median	73,74	15,9	80,38	0,67987	285.513	618,3	0,485	10,15
Mean	73,96	15,6	80,65	0,73518	1.974.541	666	0,4897	10,16
3rd Qu.	74,83	16,68	81,8	0,96709	935.132	718,1	0,51	12,2
Max.	76,27	18,6	84,39	3,85694	84.888.946	2000,3	0,62	23,92
	<b>VP(%)</b>	<b>FC18(%)</b>	<b>SC25(%)</b>	<b>EAE</b>	<b>T18(%)</b>	<b>PDA</b>	<b>PDBA</b>	<b>PDCL</b>
Min.	11,59	25,74	2,43	7,33	2,22	63,9	74,75	84,86
1st Qu.	24,23	47,19	5,53	8,695	5,397	89,65	92,33	97,17
Median	29,66	53,81	7,095	8,985	7,495	94,13	96,42	98,8
Mean	29,37	52,94	8,127	8,955	8,236	92,09	94,66	97,91
3rd Qu.	33,31	58,13	10,02	9,242	10,67	96,61	97,92	99,25
Max.	52,09	79,35	33,68	10,19	20,08	99,47	99,83	100

Fonte: Autor

No histograma, Figura 4, visualiza-se a distribuição de frequência dos 92 municípios do RJ para cada variável do estudo. Nesse gráfico observa-se a concentração dos atributos em determinadas classes. As variáveis EV, PS60, Gi, P, VP, FC18 e EAE aproximam-se de uma distribuição simétrica. Enquanto as variáveis Mi, PDA, PDBA e PDCL apresentam uma assimetria à esquerda, ou seja, há uma variação menor nas faixas mais altas (lado direito do histograma) e uma dispersão maior nas faixas mais baixas (lado esquerdo do histograma). Já os atributos PIB, RP, SC25 e T18 apresentam uma assimetria à direita, isto é, uma concentração de municípios nas classes mais baixas e uma variação maior nas categorias mais altas.

As variáveis que se aproximaram de um cenário ideal (mantendo os mesmos intervalos) foram as do grupo habitação, que possuem uma assimetria negativa/esquerda. Isso significa que a maioria dos municípios fluminenses (62) apresentam altas taxas ( $\geq 90\%$ ) de PDA, PDBA e PDCL. As demais variáveis, exceto T18, ou foram simétricas ou se comportaram o inverso da distribuição ideal.

Figura 4. Histograma e mediana das variáveis da pesquisa

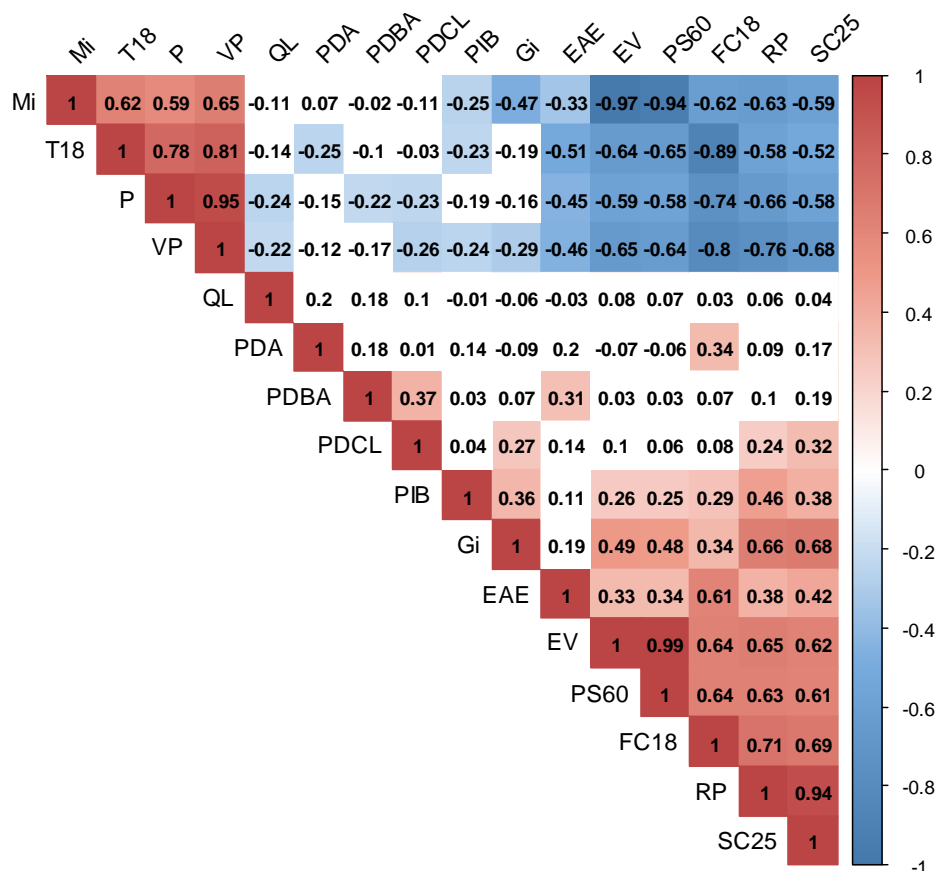


Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda *per capita*; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Fonte: Autor

A matriz de correlação de Pearson das variáveis analisadas pode ser vista na Figura 5. Os coeficientes que estão em branco não foram significativos ( $\alpha=5\%$ ); o gradiente azul indica uma correlação negativa, enquanto o vermelho, positiva; gradação em tons mais escuros mostram uma maior correlação, enquanto tons mais claros, menor correlação. As variáveis com correlação absoluta superior a 0,7 foram: Mi, com EV e PS60; T18, com P e FC18; P, com VP e FC18; VP, com FC18 e RP; FC18, com RP; e RP, com SC25.

Figura 5. Matriz de correlação de Pearson das variáveis estudadas.



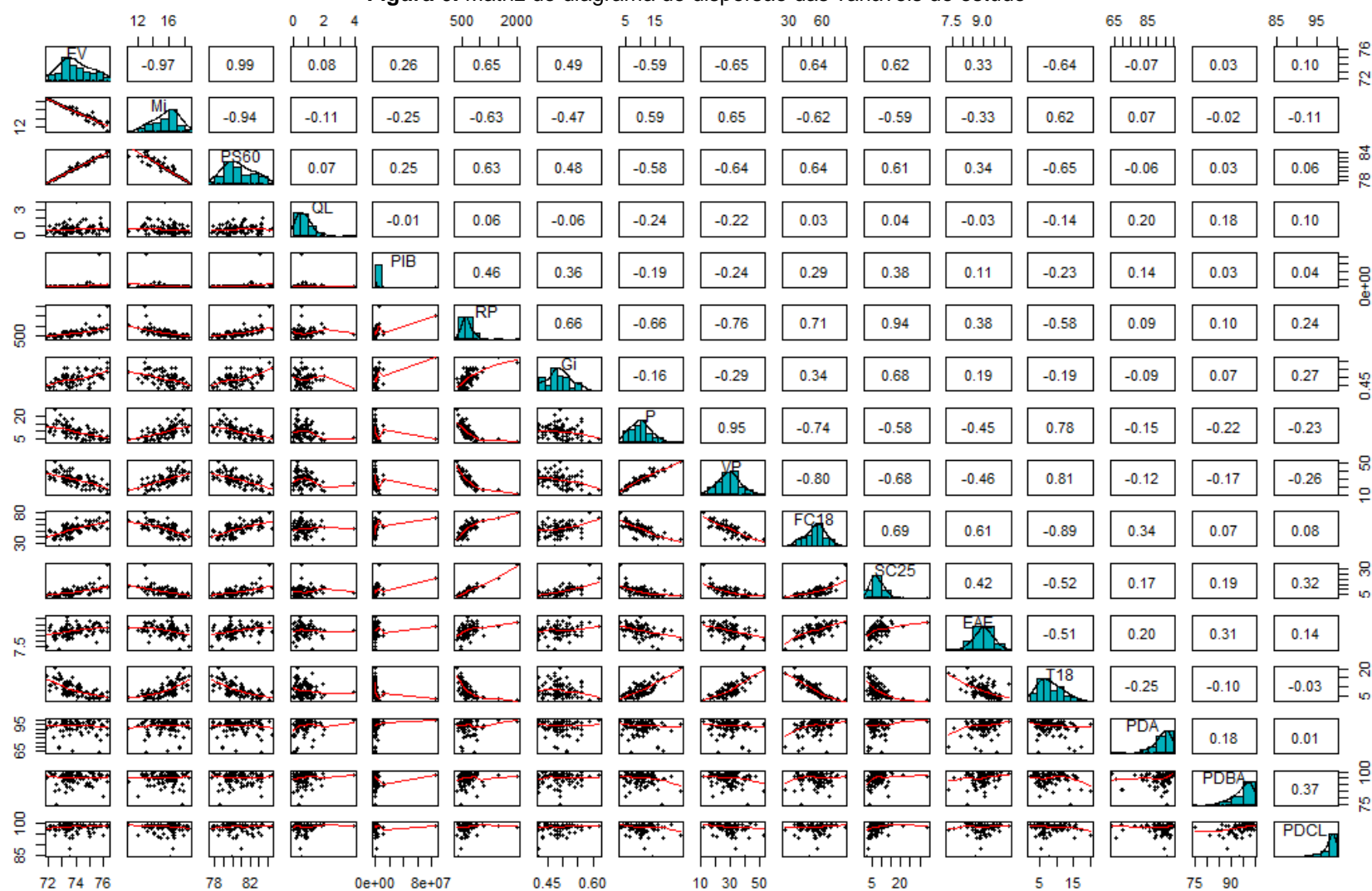
Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda *per capita*; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Nota. Escala de correlação  $r$ : 0 a 0.3  $\pm$  correlação desprezível; 0.3 a 0.5  $\pm$  correlação fraca; 0.5 a 0.7  $\pm$  correlação moderada; 0.7 a 0.9  $\pm$  correlação forte; 0.9  $\pm$  correlação muito forte (MUKAKA, 2012).

Fonte: Autor

O Diagrama de dispersão mostra a relação entre duas variáveis. Observa-se, na Figura 6, a matriz de dispersão contendo todas as combinações dos atributos. Na diagonal principal encontra-se o histograma de cada atributo. Essa diagonal também é o divisor que separa a intensidade das correlações (acima da diagonal) e o diagrama de dispersão (abaixo da diagonal). Nota-se como os valores extremos influenciaram a distribuição dos dados, a variável PIB e todas as suas combinações não apresentaram um padrão visível de dispersão, mascarado por um único ponto extremo.

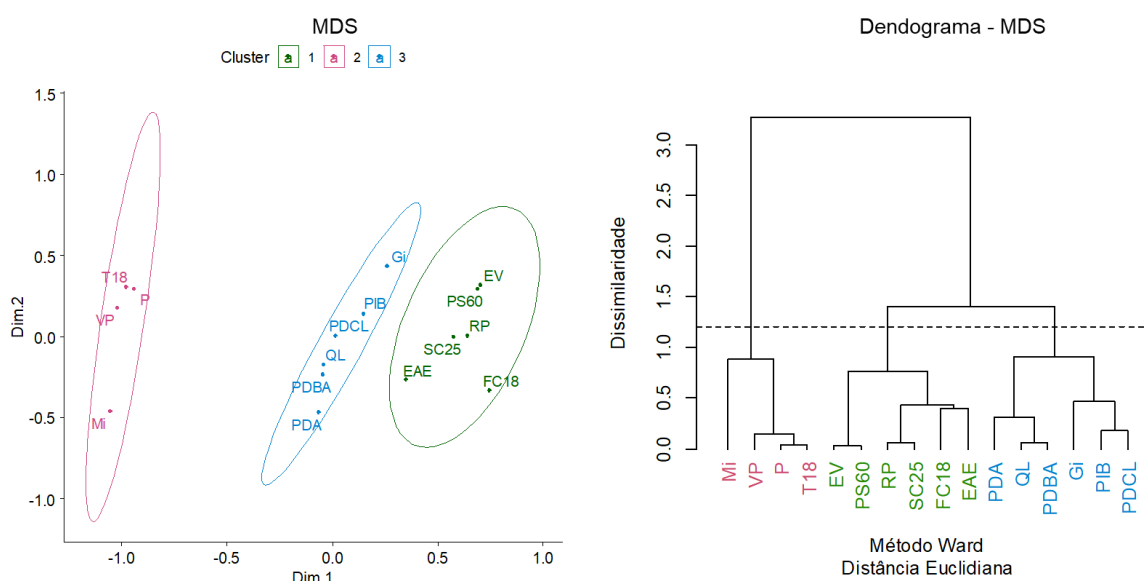
Figura 6. Matriz do diagrama de dispersão das variáveis de estudo



Fonte: Autor.

As interrelações das variáveis podem ser visualizadas na Figura 7, por meio do MDS e seu dendrograma. Se realizado um corte no diagrama de árvore um pouco acima do valor 1, destacam-se três agrupamentos – o nível de distância entre os grupos é mensurado ao longo do eixo vertical e os indicadores são listados ao longo do eixo horizontal. O grupo vermelho, localizado à esquerda, apresentou duas variáveis associadas à renda que medem a pobreza (P e VP); outra à educação, que mensura a taxa de analfabetismo em adultos (T18); por fim, uma relacionada à saúde, que afere a mortalidade infantil (Mi). O *cluster* verde foi composto por seis variáveis, das quais: três estão ligadas à educação (SC25, FC18, EAE); uma mede a renda *per capita* (RP); duas representam a saúde (EV, PS60) e dizem respeito à qualidade de vida. O grupo azul foi constituído de: três variáveis que representam a habitação (PDA, PDBA, PDCL); duas a renda (PIB e Gi); uma o grau de concentração da indústria (QL).

Figura 7. Dendrograma resultante da análise de agrupamento, utilizando como medida de dissimilaridade o MDS de uma matriz  $r$  ( $r$ : correlação de Pearson).



Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda *per capita*; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo

Fonte: Autor.

Os *clusters* vermelho e verde tiveram influência dos três grupos de indicadores: renda, saúde e educação. Já o grupo azul foi independente das variáveis de

educação. Isso demonstra que há interações entre os indicadores de diferentes temáticas, principalmente aqueles que apresentam associação conjunta com as variáveis de saúde, renda e educação.

Ao analisar a hierarquia do dendograma e a dispersão do MDS observa-se que nas relações intergrupos os *clusters* verde e azul são mais similares do que o vermelho. Esse último grupo tem características singulares. Analfabetismo e pobreza possuem uma correlação forte ( $0,7 < r < 0,9$ ); e moderada ( $0,5 < r < 0,7$ ) com mortalidade infantil. Quanto maiores os níveis desses indicadores, pior a condição do município; enquanto nos outros grupos a associação é contrária, quanto maior melhor a situação, exceto pela variável Gi. No grupo verde as relações internas mostram que as variáveis EV e PS60, RP e SC25 foram mais associadas ( $r > 0,9$ ). O primeiro par representa dois indicadores relacionados à saúde, correlacionados por natureza; o segundo apresenta um indicador que mensura a educação superior; e outro, a renda *per capita*, que indica que o aumento na renda está ligado a maiores níveis educacionais, e vice-versa. No grupo azul as correlações foram consideradas fracas ( $r = 0,5$ ). Tal grupo foi composto de indicadores predominantemente habitacionais, exceto pela variável QL, PIB e Gi.

## 4.2 CLASSIFICAÇÃO DOS MUNICÍPIOS

A classificação foi realizada em função da média dos indicadores ponderada pela população dos municípios, Tabela 4.

Tabela 4. Média dos indicadores dos municípios do estado do Rio de Janeiro, ano 2010.

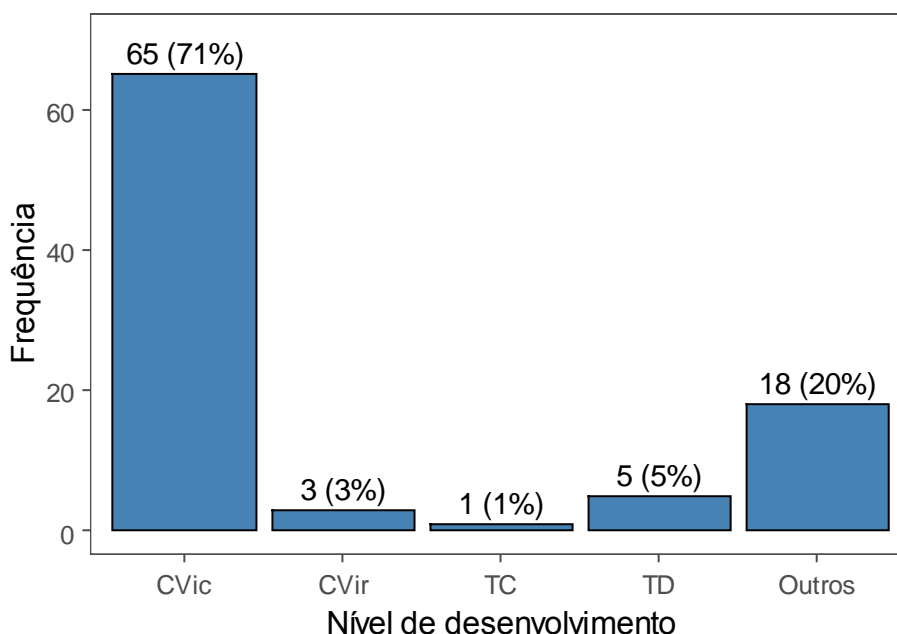
Média	IDHM		
	Renda	Longevidade	Educação
$\bar{x}_p$	0,767±0,071	0,834±0,015	0,675±0,052

Legenda. xx±yy: média±desvio-padrão.

Fonte: Autor

Após classificação dos municípios, nota-se que o número de localidades identificadas como círculo vicioso foi de 65 (71%), círculo virtuoso 3 (3%), tendendo ao crescimento 1 (1%), tendendo ao desenvolvimento 5 (5%) e outros 18 (20%), conforme Figura 8.

Figura 8: Número de municípios classificados de acordo com seu nível de desenvolvimento.



Legenda: CVic, círculo vicioso; CVir, círculo virtuoso; TC, tendendo ao crescimento; TD, tendendo ao desenvolvimento.

Fonte: Autor

Ao verificar a distribuição dos municípios por nível de desenvolvimento, percebeu-se que apenas o Rio de Janeiro (capital), Niterói e Rio das Ostras foram classificados como virtuosos. Em relação à categoria TC, somente Macaé; enquanto no nível TD as cidades de: Iguaba Grande, Mangaratiba, Maricá, Mesquita e Resende. Por outro lado, a classe que teve o maior número de municípios foi CVic, com 75% das classificações, seguido pela categoria “Outros”, com 20%.

Verifica-se que a concentração dos municípios na categoria círculo vicioso foi mais elevada (71%) em relação à soma de todas as demais categorias. A categoria tendendo ao crescimento, em longo prazo, tem uma chance maior de migrar para a categoria círculo vicioso, pois apresenta somente melhoria na componente econômica (STEWART, 2019). Nesse sentido, as classes TD e Outros provavelmente têm uma probabilidade maior de migrar para a categoria círculo virtuoso. No entanto, o rompimento de cada uma dessas situações depende da co-evolução das variáveis que influenciam tanto o CE quanto o DH (RANIS; STEWART; RAMIREZ, 2000, SURI *et al.*, 2011; STEWART, 2019).

Visualiza-se a categoria de desenvolvimento de cada município na Tabela 5. E a distribuição espacial das classificações dos municípios do ERJ no cartograma da Figura 9.



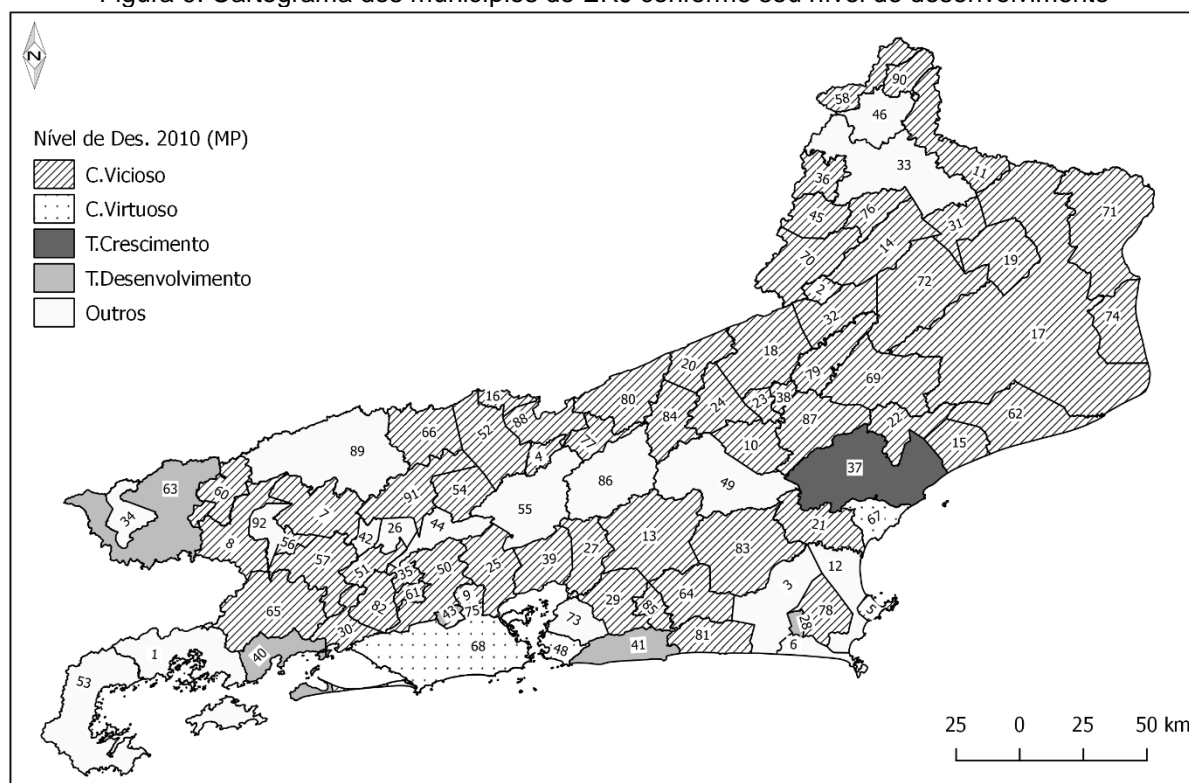
Tabela 5. Classificação dos municípios do ERJ quanto ao nível de desenvolvimento

ID	Município	2010	ID	Município	2010
1	Angra dos Reis	Outros	47	Nilópolis	Outros
2	Aperibé	CVic	48	Niterói	CVir
3	Araruama	Outros	49	Nova Friburgo	Outros
4	Areal	CVic	50	Nova Iguaçu	CVic
5	Armação dos Búzios	CVic	51	Paracambi	CVic
6	Arraial do Cabo	Outros	52	Paraíba do Sul	CVic
7	Barra do Piraí	CVic	53	Parati	Outros
8	Barra Mansa	CVic	54	Paty do Alferes	CVic
9	Belford Roxo	CVic	55	Petrópolis	Outros
10	Bom Jardim	CVic	56	Pinheiral	CVic
11	Bom Jesus do Itabapoana	CVic	57	Piraí	CVic
12	Cabo Frio	Outros	58	Porciúncula	CVic
13	Cachoeiras de Macacu	CVic	59	Porto Real	CVic
14	Cambuci	CVic	60	Quatis	CVic
15	Carapebus	CVic	61	Queimados	CVic
16	Comendador L. Gasparian	CVic	62	Quissamã	CVic
17	Campos dos Goytacazes	CVic	63	Resende	TD
18	Cantagalo	CVic	64	Rio Bonito	CVic
19	Cardoso Moreira	CVic	65	Rio Claro	CVic
20	Carmo	CVic	66	Rio das Flores	CVic
21	Casimiro de Abreu	CVic	67	Rio das Ostras	CVir
22	Conceição de Macabu	CVic	68	Rio de Janeiro	CVir
23	Cordeiro	CVic	69	Santa Maria Madalena	CVic
24	Duas Barras	CVic	70	Santo Antônio de Pádua	CVic
25	Duque de Caxias	CVic	71	São F. de Itabapoana	CVic
26	Engenheiro P. de Frontin	Outros	72	São Fidélis	CVic
27	Guapimirim	CVic	73	São Gonçalo	Outros
28	Iguaba Grande	TD	74	São João da Barra	CVic
29	Itaboraí	CVic	75	São João de Meriti	CVic
30	Itaguaí	CVic	76	São José de Ubá	CVic
31	Italva	CVic	77	São José do V. do R. Preto	CVic
32	Itaocara	CVic	78	São Pedro da Aldeia	CVic
33	Itaperuna	Outros	79	São Sebastião do Alto	CVic
34	Itatiaia	Outros	80	Sapucaia	CVic
35	Japeri	CVic	81	Squarema	CVic
36	Laje do Muriaé	CVic	82	Seropédica	CVic
37	Macaé	TC	83	Silva Jardim	CVic
38	Macuco	CVic	84	Sumidouro	CVic
39	Magé	CVic	85	Tanguá	CVic
40	Mangaratiba	TD	86	Teresópolis	Outros
41	Maricá	TD	87	Trajano de Moraes	CVic
42	Mendes	Outros	88	Três Rios	CVic
43	Mesquita	TD	89	Valença	Outros
44	Miguel Pereira	Outros	90	Varre-Sai	CVic
45	Miracema	CVic	91	Vassouras	CVic
46	Natividade	Outros	92	Volta Redonda	Outros

Fonte: Autor

Observa-se no cartograma que não houve um padrão específico de agrupamento espacial.

Figura 9. Cartograma dos municípios do ERJ conforme seu nível de desenvolvimento



Legenda: 1-Angra dos Reis; 2-Aperibé; 3-Araruama; 4-Areal; 5-Armação dos Búzios; 6-Arraial do Cabo; 7-Barra do Pirai; 8-Barra Mansa; 9-Belford Roxo; 10-Bom Jardim; 11-Bom Jesus do Itabapoana; 12-Cabo Frio; 13-Cachoeiras de Macacu; 14-Cambuci; 17-Campos dos Goytacazes; 18-Cantagalo; 15-Carapebus; 19-Cardoso Moreira; 20-Carmo; 21-Casimiro de Abreu; 16-Comendador L. Gasparian; 22-Conceição de Macabu; 23-Cordeiro; 24-Duas Barras; 25-Duque de Caxias; 26-Engenheiro P. de Frontin; 27-Guapimirim; 28-Iguaba Grande; 29-Itaboraí; 30-Itaguaí; 31-Italva; 32-Itaocara; 33-Itaperuna; 34-Itatiaia; 35-Japeri; 36-Laje do Muriaé; 37-Macaé; 38-Macuco; 39-Magé; 40-Mangaratiba; 41-Maricá; 42-Mendes; 43-Mesquita; 44-Miguel Pereira; 45-Miracema; 46-Natividade; 47-Nilópolis; 48-Niterói; 49-Nova Friburgo; 50-Nova Iguaçu; 51-Paracambi; 52-Paraíba do Sul; 53-Parati; 54-Paty do Alferes; 55-Petrópolis; 56-Pinheiral; 57-Pirai; 58-Porciúncula; 59-Porto Real; 60-Quatis; 61-Queimados; 62-Quissamã; 63-Resende; 64-Rio Bonito; 65-Rio Claro; 66-Rio das Flores; 67-Rio das Ostras; 68-Rio de Janeiro; 69-Santa Maria Madalena; 70-Santo Antônio de Pádua; 71-São F. de Itabapoana; 72-São Fidélis; 73-São Gonçalo; 74-São João da Barra; 75-São João de Meriti; 76-São José de Ubá; 77-São José do V. do R. Preto; 78-São Pedro da Aldeia; 79-São Sebastião do Alto; 80-Sapucaia; 81-Saquarema; 82-Seropédica; 83-Silva Jardim; 84-Sumidouro; 85-Tanguá; 86-Teresópolis; 87-Trajano de Moraes; 88-Três Rios; 89-Valença; 90-Varre-Sai; 91-Vassouras; 92-Volta Redonda.

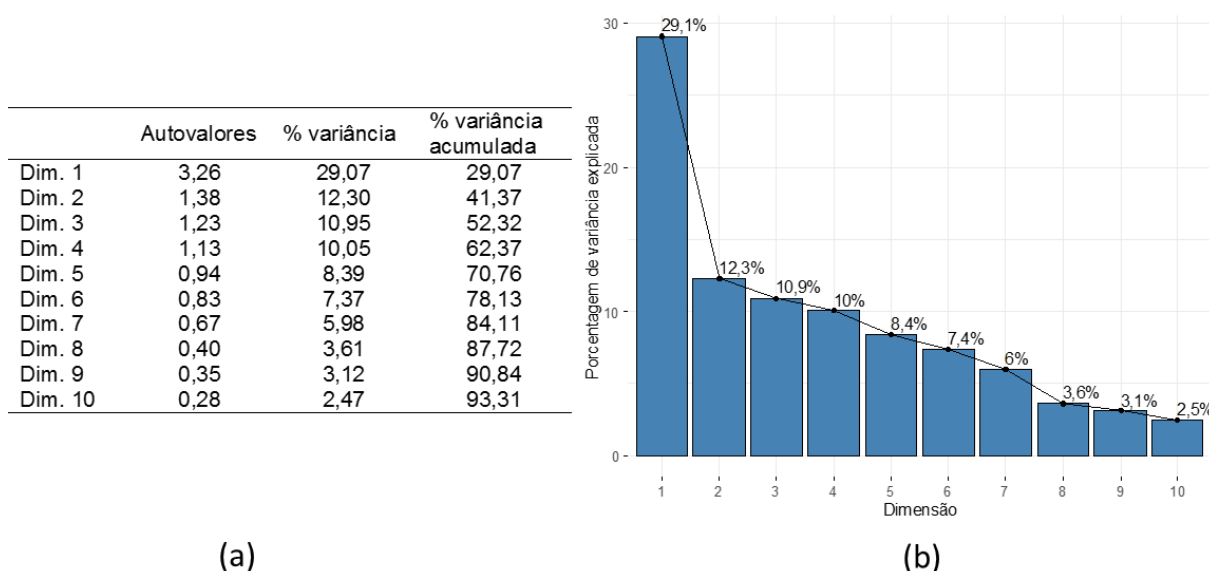
Fonte: Autor

#### 4.3 ANÁLISE FATORIAL

Os principais autovalores (variância), bem como a porcentagem de variância explicada e acumulada nas dez primeiras dimensões/eixos/componentes, podem ser vistos na Figura 10. No lado (a) da Figura observa-se a tabela contendo os dados dos autovalores, variância, e variância acumulada, enquanto o lado (b) exibe o gráfico do

percentual de variância retida por cada eixo. Essas informações auxiliam na determinação do número de dimensões a serem analisadas. Segundo Kassambara (2017), não existe um consenso na escolha dos números de dimensões a serem avaliadas. Isso dependerá da área de estudo e do conjunto de dados. Ainda segundo o autor, na prática tende-se a analisar os primeiros componentes principais. Neste estudo foram escolhidos os dois primeiros eixos que, por padrão, retêm a maior quantidade de variância, ou seja, informação. As dimensões 1 e 2 juntas representam 41,37% (29,07+12,30) da variância dos dados.

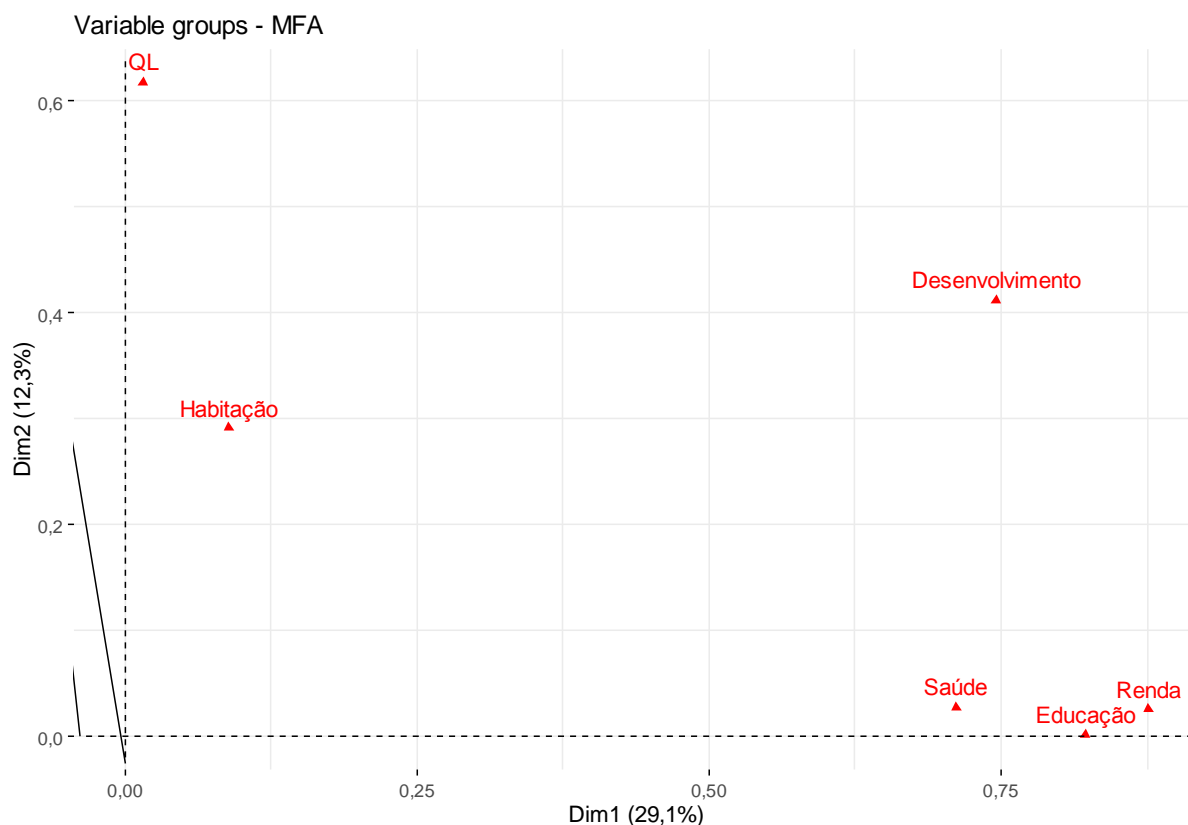
Figura 10. Proporção de variância explicada pelos principais autovalores



Fonte: Autor

Na Figura 11 pode-se visualizar o gráfico das relações entre as dimensões e os grupos avaliados. Os grupos desenvolvimento, habitação, educação e renda estão mais associados à primeira dimensão (esses grupos estão distantes da origem em relação a primeira dimensão ou eixo x), com contribuições semelhantes a essa componente. Os grupos QL e Habitação têm coordenadas mais altas na segunda dimensão, o que indica uma contribuição maior para essa componente. O grupo desenvolvimento está associado a ambas as dimensões, sendo um pouco mais representado pela dimensão 1.

Figura 11. O gráfico de correlação entre os grupos e as duas primeiras dimensões, média ponderada.



O círculo de correlação (mapa de fatores) das variáveis e dimensões foi apresentado na Figura 12. Variáveis que têm correlações positivas são agrupadas, enquanto as negativas são posicionadas em lados opostos. A qualidade das variáveis no mapa de fatores é mensurada pela sua distância em relação à origem – quanto mais longe melhor a representação. A correlação da variável com a dimensão se dá pelo seu distanciamento – nesse caso, quanto menor a distância de um ponto ao eixo maior será a associação. Essas questões podem ser visualizadas tanto graficamente (Figura 12) como numericamente (Tabela 6); quanto maior o peso mais intenso será a contribuição daquela variável para a dimensão.

Tabela 6. Peso/coeficiente (autovetores) das variáveis quantitativas para as dimensões 1 e 2

	EV	Mi	PS60	QL	PIB	RP	Gi	P	VP	FC18	SC25	EAE	T18	PDA	PDBA	PDCL
Dim.1	0,468	-0,460	0,458	0,068	0,228	0,492	0,313	-0,432	-0,475	0,469	0,477	0,299	-0,427	0,073	0,098	0,154
Dim.2	-0,146	0,117	-0,156	0,670	-0,022	-0,004	-0,096	-0,175	-0,130	0,004	0,006	-0,041	-0,058	0,398	0,313	0,209

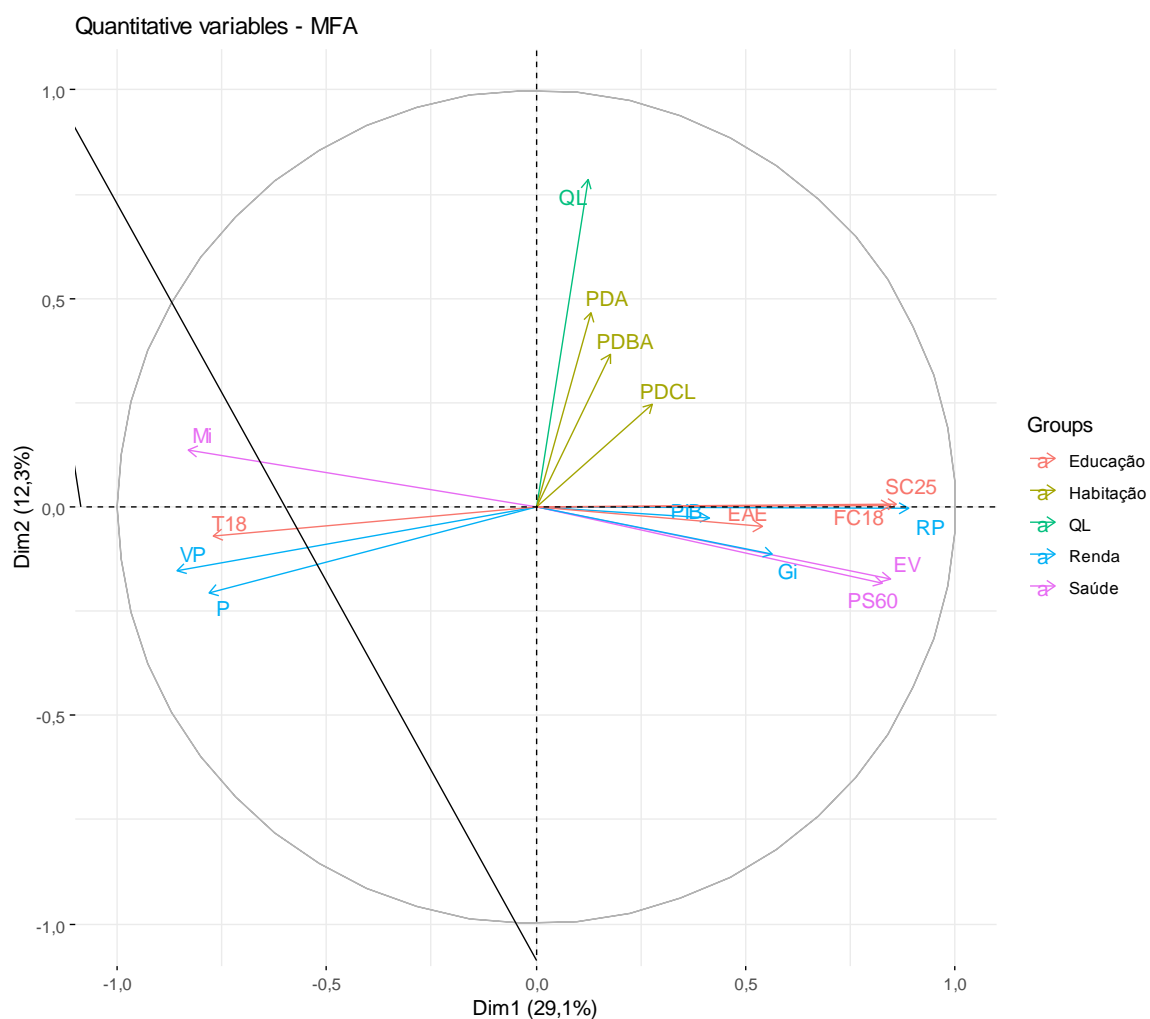
Fonte: Autor

As variáveis Mi, T18, VP e P foram relacionadas positivamente entre si, pois os pontos estão próximos no gráfico e possuem pesos semelhantes, com mesmo sinal

negativo (-) na dimensão 1. Essas variáveis ainda foram correlacionadas negativamente com SC25, FC18, RP, EV e PS60, que também demonstraram boa contribuição com a primeira dimensão. Nota-se que esses indicadores contribuíram pouco para a segunda componente (Figura 11, Tabela 6) quando comparados com a primeira dimensão. Apesar de EAE, Gi e PIB apresentarem algum tipo de relação com as variáveis já citadas, essas tiveram um peso menor (Tabela 6) nas dimensões avaliadas. Portanto é necessário ser mais parcimonioso quanto as suas interpretações.

As variáveis QL, PDA, PDBA e PDCL foram representadas pela segunda dimensão. Assim sendo, os pesos dos indicadores são maiores para essa componente. Ressalta-se que a variável QL foi a que obteve o maior peso (0,67), enquanto as outras três, respectivamente, pesos menores de 0,398; 0,313 e 0,209.

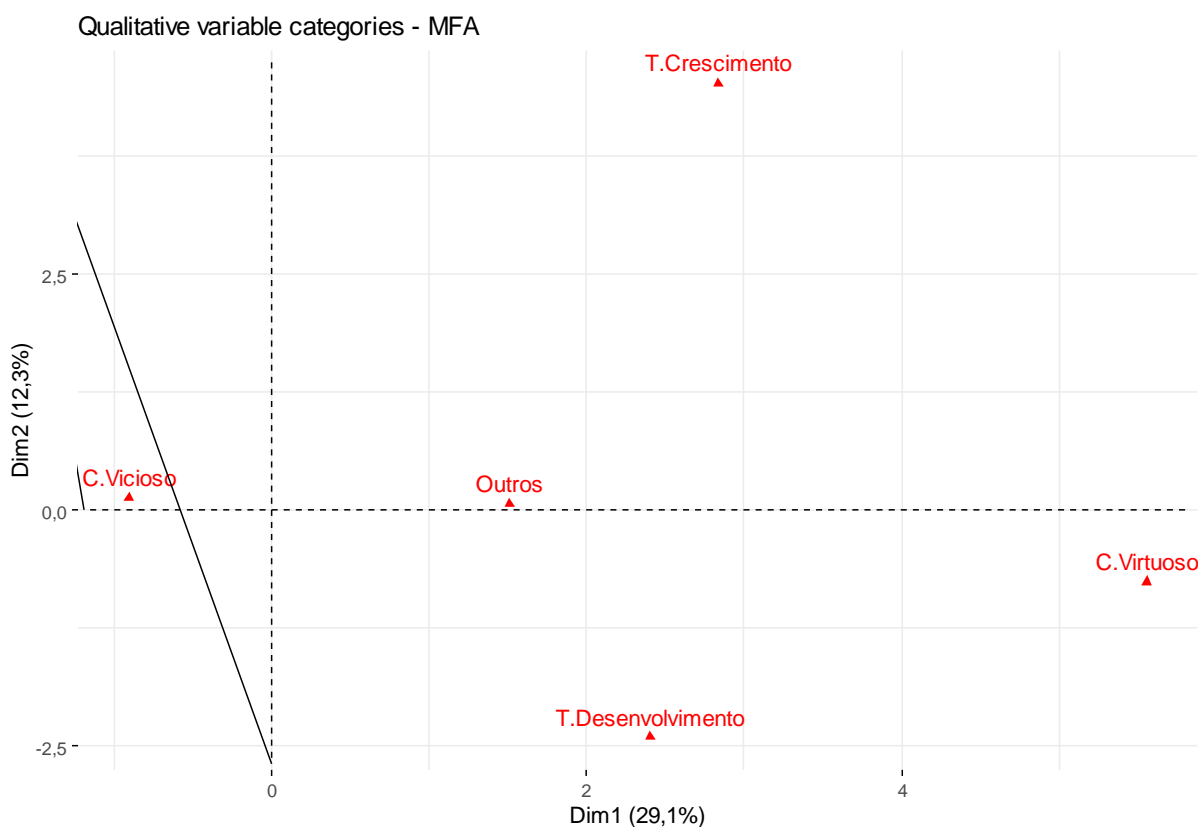
**Figura 12.** Círculo de correlação entre as variáveis quantitativas; Dim1 - eixo x, Dim2 - eixo y



Fonte: Autor

Na Figura 13 mostra-se a relação das categorias da variável ND, assim como a correlação destas categorias com as dimensões; sendo possível visualizar as ponderações das categorias na Tabela 7. Tanto no gráfico como na Tabela, observa-se que os níveis de desenvolvimento vicioso e virtuoso tiveram maior correlação com a primeira dimensão. Porém foram localizados em lados opostos, o que indica uma relação antagônica entre esses círculos de desenvolvimento. Dessa maneira, os municípios classificados como virtuosos tiveram atributos totalmente opostos aos viciosos. O mesmo ocorreu com as classes tendendo ao crescimento e desenvolvimento, sendo a relação mais forte com a segunda componente. A categoria “outros” mostrou-se mais associada à primeira dimensão.

Figura 13. Mapa de fator da variável categórica ND; Dim1 - eixo x, Dim2 - eixo y



Fonte: Autor

Tabela 7. Peso (autovetores) da variável categórica para as dimensões 1 e 2

	C.Vicioso	C.Virtuoso	Outros	T.Crescimento	T.Desenvolvimento
Dim.1	-0,4300	0,3128	0,2286	0,0912	0,1768
Dim.2	0,1486	-0,1010	0,0243	0,3434	-0,4186

Fonte: Autor

A relação entre as variáveis categóricas e quantitativas se dá por meio da comparação dos gráficos das Figuras 12, 13 e também Tabelas 6 e 7. Assim, as variáveis, independentemente do tipo, relacionam-se em função das dimensões a que estão mais associadas. Ao avaliar as duas Figuras e Tabelas, identificaram-se os seguintes padrões: a categoria círculo vicioso está relacionada com as variáveis quantitativas Mi(+), T18(+), VP(+) e P(+) positivamente e negativamente com SC25 (-), FC18(-), RP(-), EV(-) e PS60(-); a classe círculo virtuoso relaciona-se com as mesmas variáveis da categoria anterior, no entanto, de maneira inversa. A classe tendendo ao crescimento mostrou-se associada ao QL(+), PDA(+), PDBA(+); a categoria tendendo ao desenvolvimento mostrou uma relação inversa à categoria tendendo ao crescimento. A categoria “outros” não mostrou um padrão específico.

Esses resultados mostram que o círculo vicioso é muito influenciado por saúde (Mi), educação (T18) e pobreza e sua vulnerabilidade (P e VP). Todas essas variáveis cerceiam o indivíduo em suas escolhas, tornam-no incapacitado, seja porque nem sequer sobrevivem à infância, seja porque não sabem ler, seja porque se situam à margem do sistema. Mais do que isso, verifica-se que apenas a melhoria desses indicadores não é capaz de romper o círculo vicioso. Para tanto é necessário melhorar outros indicadores também relacionados com saúde e educação (EV; PS60; SC25; FC18), mas também uma melhor distribuição de renda (RP).

Na Figura 14 observa-se a similaridade, indicada pela distância, entre os municípios à luz de todas as variáveis analisadas. Além disso, os municípios foram classificados em função do seu nível de desenvolvimento. Observa-se que os municípios identificados pelas categorias de desenvolvimento formaram um padrão de agrupamento que corrobora a análise da Figura 13. Dessa forma, os municípios localizados à esquerda estão relacionados com o círculo vicioso, à direita com o virtuoso, acima tendendo ao crescimento e abaixo ao desenvolvimento, já os centralizados são classificados como “outros”.

As características de cada um desses grupos de desenvolvimento podem ser identificadas relacionando a análise dos municípios (Figura 14) com as anteriores: MFA variáveis quantitativas (Figura 12) e MFA variáveis qualitativas (Figura 13). Desse modo, os municípios localizados tendendo à esquerda estão mais relacionados a altos índices de Mi, T18, VP e P, e baixos valores de SC25, FC18, RP, EV e PS60, enquanto os indivíduos à direita associam-se com essas variáveis de maneira inversa.

O município de Macaé, parte superior, foi o único classificado como tendendo ao crescimento e tem uma alta relação com o QL, que mede a concentração da indústria. Nota-se que Porto Real também está associado ao quociente locacional, apesar de se encontrar na categoria círculo vicioso. A diferença entre esses dois municípios refere-se aos indicadores essenciais do desenvolvimento humano: quanto mais abaixo da média, caso de Porto Real, mais próximo do círculo vicioso. Esses dois casos ilustram bem a ideia de que não só a concentração industrial leva ao círculo virtuoso<sup>6</sup>.

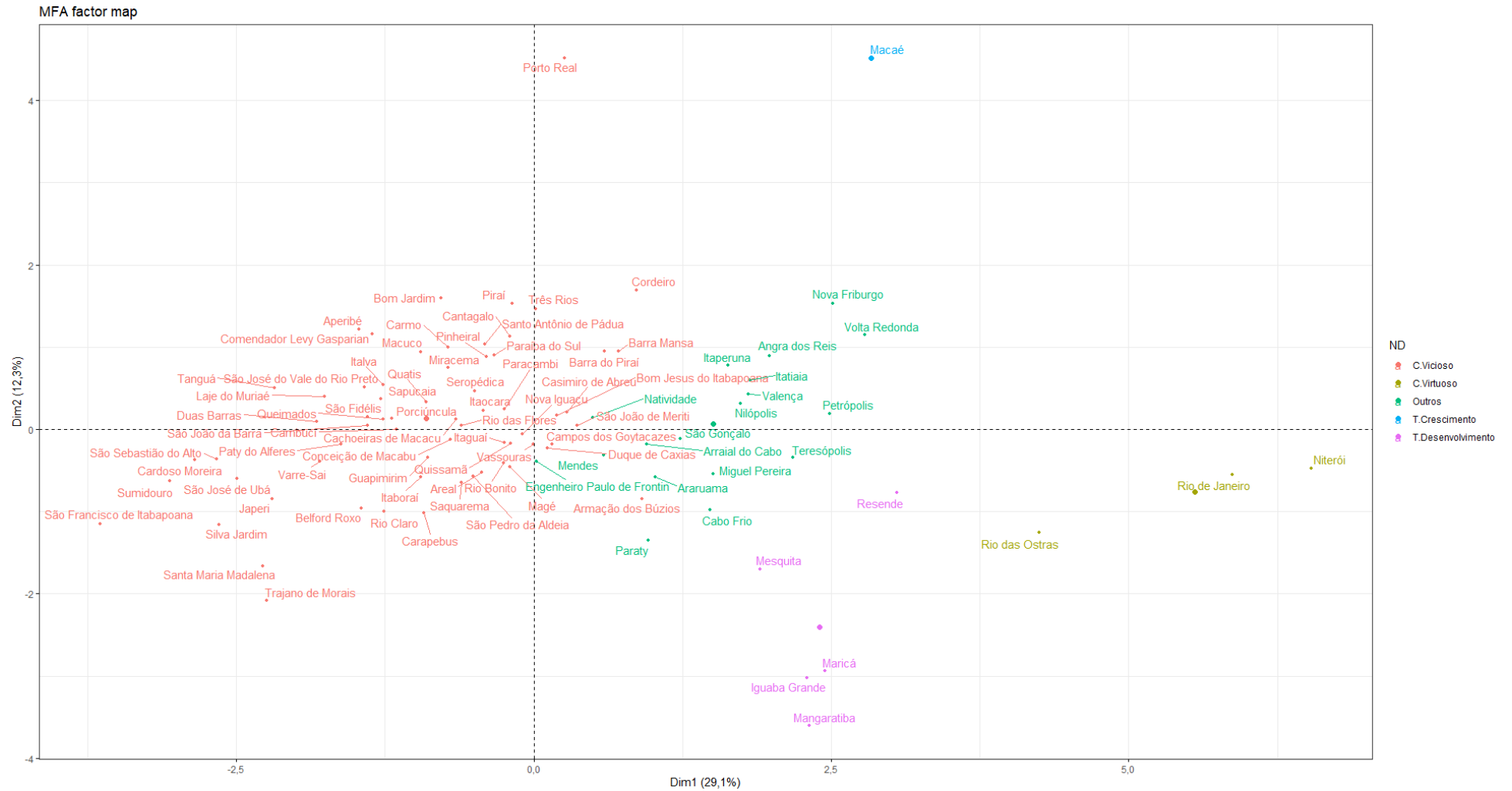
As localidades situadas na parte inferior têm características inversas àqueles classificados como tendendo ao crescimento, portanto apresentam baixo QL. Os municípios classificados como “outros” não apresentaram uma relação visível com nenhuma variável do estudo.

---

<sup>6</sup> Ver a esse respeito na crítica que Cepal e outros organismos internacionais fizeram ao desenvolvimento não inclusivo na década de 1990.



Figura 14. Mapa de fatores dos indivíduos (municípios), média ponderada



Sobre outra perspectiva, observa-se na Tabela 8 a comparação das médias das variáveis dentro de cada grupo com a média geral. Essa comparação foi realizada por meio do *test v*, adaptação do teste *t* por Lebart *et al.* (1997). Nota-se que nessa Tabela só foram apresentadas as variáveis com diferença significativa a 95% de confiabilidade. No círculo vicioso as variáveis *Mi*, *VP*, *P*, *T18* possuem médias acima da geral, dessa forma, contribuem positivamente para formação desse círculo. Por outro lado, as variáveis *Gi*, *EAE*, *RP*, *FC18*, *SC25*, *PS60* e *EV* apresentaram média abaixo da referência, assim contribuindo de forma negativa para esse grupo. Além da comparação das médias, o sinal (+/-) do *v test* indica a direção da contribuição de cada variável. Os municípios agrupados no círculo virtuoso possuem relações opostas ao círculo vicioso, conforme pode ser observado nos sinais do teste estatístico. A interpretação dos demais grupos é análoga à anterior. Nota-se que as características dos grupos Vicioso e Virtuoso estão de acordo com os padrões encontrados na análise da MFA, exceto as variáveis *Gi* e *EAE* que, apesar de estarem correlacionadas na MFA, não tiveram representatividade nas dimensões avaliadas. Nessa mesma linha, a categoria tendendo ao crescimento corroborou os resultados anteriores, exceto pela variável *RP*, que não mostrou nenhuma relação significativa. Por outro lado, as características das categorias “outros” e tendendo ao desenvolvimento foram mais perceptíveis nessa análise de comparação de médias. As duas dimensões avaliadas na MFA não foram suficientes para encontrar todas as relações dessa classe.

**Tabela 8.** Comparação das médias das categorias de desenvolvimento (classificação ponderada) em relação à média geral das variáveis avaliadas

Círculo Vicioso						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Mi	6,3646	16,2915	15,5977	1,1032	1,6135	0,0000
VP	5,6297	32,2118	29,3728	6,3113	7,4641	0,0000
P	4,8811	11,4149	10,1559	3,6211	3,8179	0,0000
T18	4,8216	9,4326	8,2355	3,5257	3,6747	0,0000
Gi	-3,8387	0,4789	0,4897	0,0354	0,0415	0,0001
EAE	-3,9432	8,8283	8,9551	0,4157	0,476	0,0001
RP	-5,5964	584,1875	666,0147	93,7887	216,4145	0,0000
FC18	-5,5996	49,3483	52,9425	8,0256	9,5004	0,0000
SC25	-5,6442	6,5238	8,1267	2,2589	4,2033	0,0000
PS60	-6,4684	79,9568	80,6528	1,105	1,5927	0,0000
EV	-6,6246	73,4749	73,9593	0,722	1,0823	0,0000
Círculo Virtuoso						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
RP	6,8683	1514,703	666,0147	387,7827	216,4145	0,0000
SC25	6,1153	22,8033	8,1267	8,3226	4,2033	0,0000
PIB	5,6776	30874694	1974541	38205108	8915039	0,0000
Gi	3,8163	0,58	0,4897	0,0374	0,0415	0,0001
FC18	3,6398	72,6867	52,9425	5,2496	9,5004	0,0003
EV	3,3992	76,06	73,9593	0,2619	1,0823	0,0007
PS60	3,1088	83,48	80,6528	0,4685	1,5927	0,0019
T18	-2,4572	3,08	8,2355	0,6489	3,6747	0,0140
P	-2,5777	4,5367	10,1559	0,8523	3,8179	0,0099
VP	-3,3662	15,0267	29,3728	2,4454	7,4641	0,0008
Mi	-3,5832	12,2967	15,5977	0,9462	1,6135	0,0003
Outros						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
PS60	4,0615	82,0278	80,6528	1,4383	1,5927	0,0000
EV	3,9899	74,8772	73,9593	0,9576	1,0823	0,0001
FC18	2,6732	58,3406	52,9425	5,2396	9,5004	0,0075
EAE	2,5305	9,2111	8,9551	0,5127	0,476	0,0114
SC25	2,4543	10,3194	8,1267	2,0642	4,2033	0,0141
T18	-2,769	6,0728	8,2355	2,1586	3,6747	0,0056
P	-3,0072	7,7156	10,1559	2,1821	3,8179	0,0026
VP	-3,2893	24,1544	29,3728	4,8535	7,4641	0,0010
Mi	-3,507	14,395	15,5977	1,3858	1,6135	0,0005
Tendendo ao Crescimento						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
QL	2,0502	1,8654	0,7352	0	0,5513	0,0403
RP	2,0211	1103,42	666,0147	0	216,4145	0,0433
Tendendo ao Desenvolvimento						
	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
EV	3,3652	75,552	73,9593	0,2632	1,0823	0,0008
PS60	3,2754	82,934	80,6528	0,2194	1,5927	0,0011
FC18	3,0026	65,416	52,9425	1,2543	9,5004	0,0027
EAE	2,57	9,49	8,9551	0,2355	0,476	0,0102
VP	-2,167	22,3	29,3728	3,576	7,4641	0,0302
T18	-2,4418	4,312	8,2355	0,3817	3,6747	0,0146
Mi	-3,2737	13,288	15,5977	0,7027	1,6135	0,0011

Fonte: Autor.

Em suma, os indicadores de pobreza, mortalidade infantil e analfabetismo de pessoas com 18 anos ou mais foram correlacionados positivamente entre si, e associados ao círculo vicioso. Já os indicadores: educacionais (SC25, FC18); renda *per capita*; esperança de vida; e probabilidade de sobreviver até os 60 anos são indicativos importantes para se romper o círculo vicioso, entretanto, há também a

necessidade de uma melhora da renda *per capita* para que esse círculo seja vencido. Dessa forma, observou-se o processo de causação circular em 74% (Figura 7) dos municípios, tanto negativo como positivo, que se reforça cumulativamente em que pobreza gera pobreza e riqueza gera riqueza. Todavia não se pode negar o fato de que existem situações de desequilíbrio, TC e TD, cujo círculo é interrompido de forma que o crescimento econômico não gera melhorias para desenvolvimento humano, assim como o desenvolvimento humano não promove o crescimento econômico.

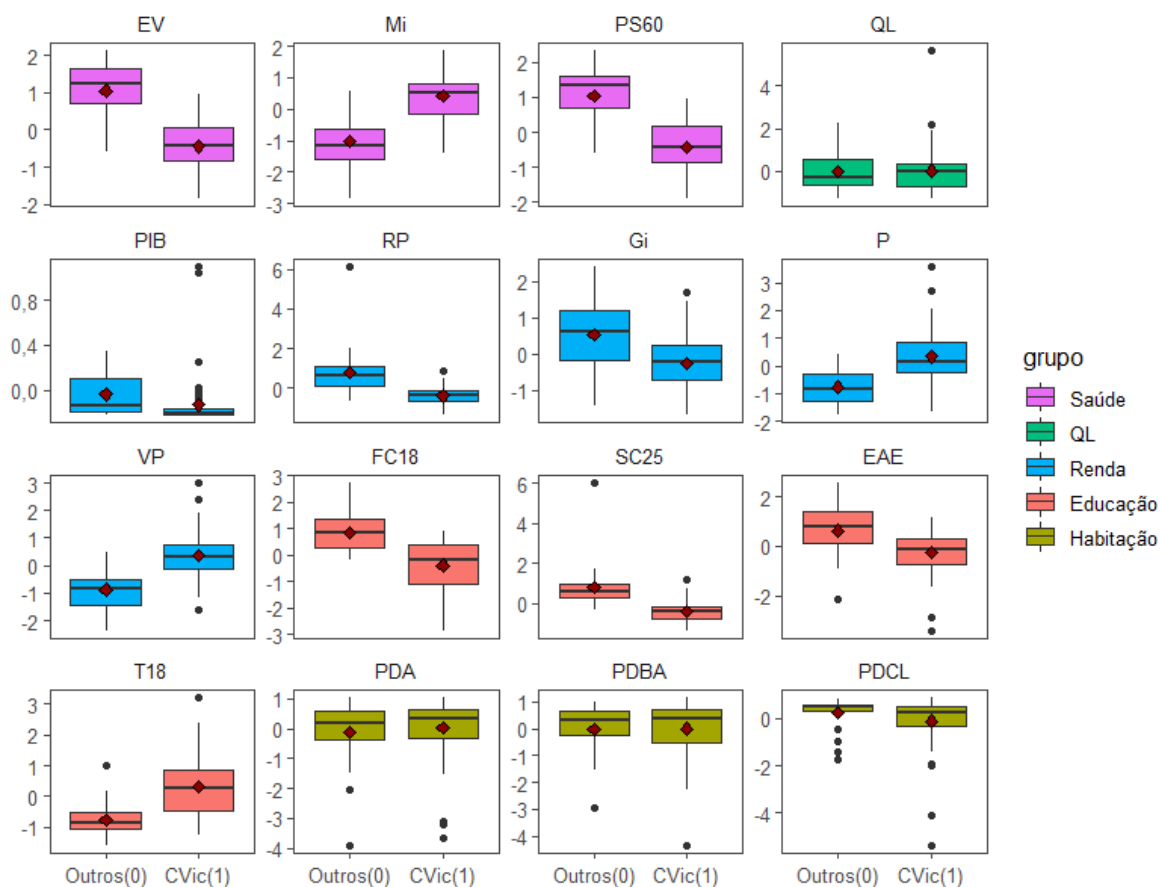
#### 4.4 MODELO DE REGRESSÃO

O modelo de regressão logístico foi o binomial, devido às distribuições das categorias (ver Figura 8), sendo comparada à categoria círculo vicioso em relação às demais.

Para interpretar o modelo quanto à razão de chances de estar no círculo vicioso tomou-se o antilogaritmo do  $j$ -ésimo parâmetro, posteriormente subtraiu-se o resultado de 1 e multiplicou-se por cem. Resultando, assim, na variação percentual das chances a cada unidade incrementada no  $j$ -ésimo regressor (GUJARATI, 2011).

O impacto das variáveis em função da categoria de desenvolvimento pode ser percebido visualmente por meio da média e do diagrama de caixa das variáveis desagregadas por classe de desenvolvimento (Figura 15). Nota-se que a maioria das variáveis apresentou diferença entre as classes. As exceções foram QL, PDA, PDBA e PDCL, que demonstraram baixa variação entre as categorias em relação às outras variáveis.

Figura 15. Diagrama de dispersão e média das variáveis (padronizadas, z-score) do estudo, separadas conforme as classes do modelo binomial

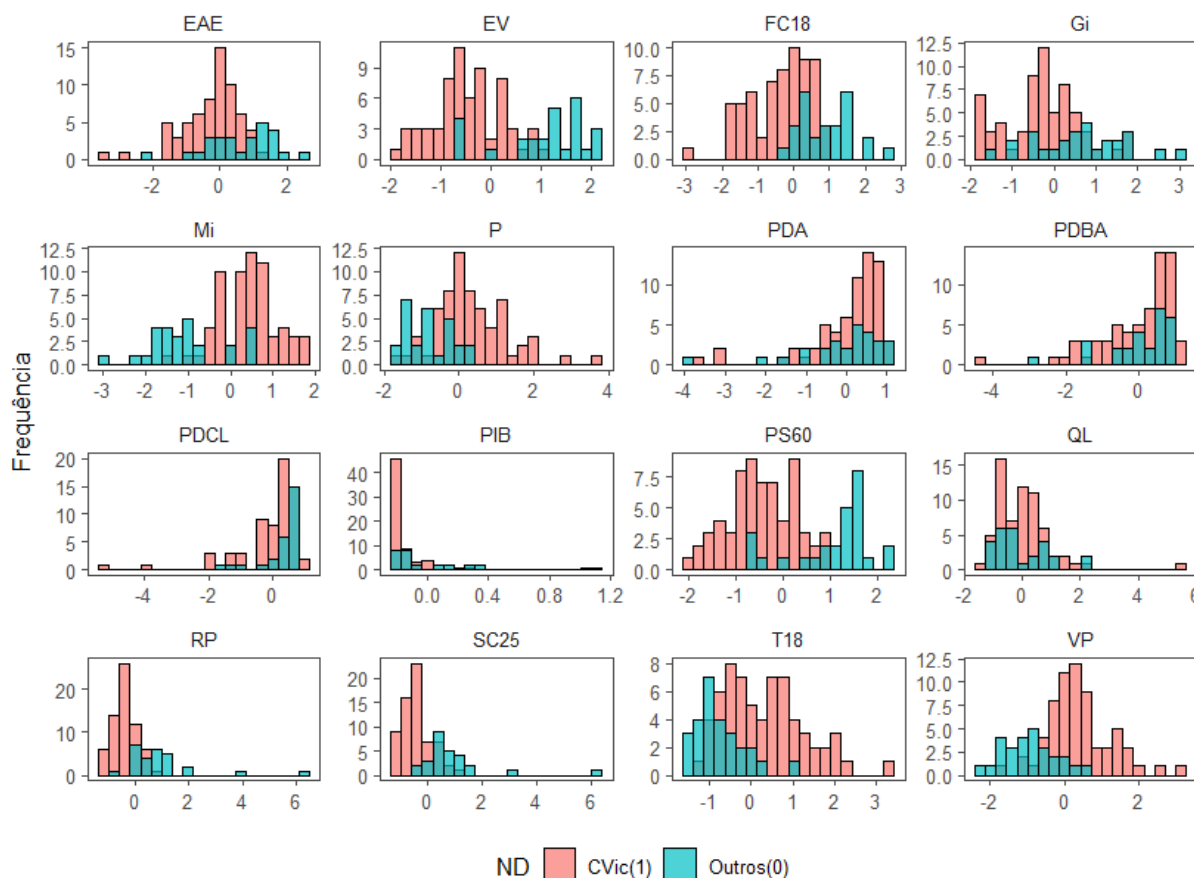


Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda per capita; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Fonte: Autor

No histograma apresentado na Figura 16 nota-se que os municípios classificados como “Outros” tendem aos extremos da distribuição e que as variáveis QL, PDA, PDBA e PDCL apresentam praticamente uma sobreposição completa das faixas de variação. Dessa maneira, não há uma discriminação das categorias CVic e “Outros” para essas variáveis.

Figura 16. Histograma das variáveis (padronizadas, z-score) do estudo, categorizadas conforme as classes do modelo binomial.



Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda *per capita*; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Fonte: Autor

Os coeficientes do modelo de regressão são visualizados na Tabela 9, enquanto os testes de ajuste do modelo na Figura 17. Observou-se que as variáveis EV (32,63%), PS60 (22,52%), RP (0,12%), FC18 (3,87%), SC25 (9,05%) e EAE (64,62%) reduzem a chance,  $\alpha=5\%$ , de um município passar para o CVic a cada unidade incrementada. Por outro lado, Mi e VP aumentam essa chance em 23,99% e 3,63%, respectivamente.

As medidas de qualidade do modelo (Figura 16) indicam: uma precisão de  $\cong 99\%$ ; um bom ajuste do modelo pelo teste de Hosmer e Lemeshow<sup>7</sup> ( $p\text{-value} > 0,05$ ).

<sup>7</sup> Teste de qualidade de ajuste, H0: as probabilidades estimadas não desviam das probabilidades observadas.

Os pseudos  $R^2$  foram aproximadamente 0,83 (McFadden); 0,63 (CoxSnell); 0,89 (Nagelkerke). O último está em uma escala de 0 a 1, logo 89% da variabilidade de  $y$  foi explicada pelo modelo.

Tabela 9. Modelo de regressão binomial com regularização de Ridge

	Estimate	Scaled estimate	Std. Error (scaled)	t value (scaled)	Pr(> t )	
(Intercept)	59,8100	NA	NA	NA	NA	
EV	-0,3950	-4,1010	0,9847	-4,1640	3,12E-05	***
Mi	0,2150	3,3280	1,2260	2,7140	0,006645	**
PS60	-0,2552	-3,8980	1,1410	-3,4170	0,000632	***
QL	0,3050	1,6130	1,7670	0,9130	0,361389	
PIB	0,0000	0,5942	1,3710	0,4330	0,664739	
RP	-0,0012	-2,5340	0,8952	-2,8310	0,004645	**
Gi	-3,9840	-1,5840	1,5460	-1,0250	0,305481	
P	0,0412	1,5100	1,2350	1,2230	0,221508	
VP	0,0357	2,5530	1,1310	2,2580	0,023969	*
FC18	-0,0395	-3,5950	1,2070	-2,9790	0,002894	**
SC25	-0,0949	-3,8250	1,2180	-3,1410	0,001685	**
EAE	-1,0390	-4,7430	1,7120	-2,7700	0,005597	**
T18	0,0429	1,5110	1,3230	1,1420	0,253544	
PDA	0,0189	1,2860	1,6800	0,7660	0,443887	
PDBA	0,0292	1,2650	1,6800	0,7530	0,451603	
PDCL	-0,0343	-0,7903	1,6560	-0,4770	0,633157	

Signif. codes: 0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 '.' 0,1 ' ' 1

Legenda: Mi - Mortalidade infantil (%); PS60 - Probabilidade de sobrevivência até 60 anos (%); EV - Esperança de vida ao nascer; QL - Quociente Locacional; PIB (Mil) – Produto interno bruto; RP - Renda *per capita*; Gi - Índice de Gini; P - % de pobres; VP - % de vulneráveis à pobreza; FC18 - % de 18 anos ou mais com fundamental completo; SC25 - % de 25 anos ou mais com superior completo; EAE - Expectativa de anos de estudo; T18 - Taxa de analfabetismo - 18 anos ou mais; PDA - % da população em domicílios com água encanada; PDBA - % da população em domicílios com banheiro e água encanada; PDCL - % da população em domicílios com coleta de lixo.

Fonte: Autor.

Figura 17. Medidas de ajuste do modelo de regressão binomial

Matriz de Confusão			Hosmer and Lemeshow test (binary model) X-squared = 11.753, df = 8, p-value = 0.1625
	CVic	Outros	
CVic	64	0	Pseudo $R^2$ McFadden= 0,828799 CoxSnell = 0,6333103 Nagelkerke = 0,9022235
Outros	1	27	
Precisão=(27+64)/92=0.9891304			

Fonte: Autor.

Ao confrontar os resultados do modelo com o boxplot e a média das variáveis desagregadas por classe (Figura 15) verificou-se que: tanto pela medida de tendência central quanto pelos quartis que as variáveis EV, Mi e PS60 estão coerentes com o modelo, conforme pode ser visto nos municípios classificados como CVic tem uma EV inferior, uma taxa de Mi superior e uma taxa de PS60 abaixo dos demais municípios. Já as variáveis QL e PIB não mostraram nenhum tipo de padrão, tanto que não foram significativas. Na variável RP as cidades categorizadas como CVic demonstram uma RP abaixo das outras localidades – essa diferença foi significativa na regressão. A variável Gi e P não contribuíram estatisticamente para o modelo, apesar de existir uma pequena diferença entre as categorias graficamente. Em relação às variáveis VP, FC18, SC25 e EAE a diferença entre as classes foi captada pelo modelo. Os municípios do círculo vicioso tiveram maiores taxas de VP, e baixos índices de indivíduos com FC18, SC25 e EAE. A variável T18 apresentou uma diferença entre as classes, mas não apresentou diferença significativa. PDA, PDBA e PDCL quase não apresentaram variações entre as classes, tanto que seus coeficientes não foram estatisticamente significativos.

O modelo final da regressão, considerando os coeficientes estatisticamente significativos ( $\alpha=0,05$ ), foi:

$$\log \left[ \frac{Pr(G = 1|x)}{Pr(G = 0|x)} \right] = 59,81 - 0,595EV + 0,215Mi - 0,255PS60 - 0,0012RP + 0,036VP - 0,039FC18 - 0,095SC25 - 1,039EAE$$

Por meio do modelo de regressão binomial penalizado, pode-se estabelecer um viés confirmatório entre o círculo vicioso da pobreza e a contribuição dos baixos indicadores socioeconômicos para sua formação. Das 16 variáveis preditoras, oito apresentaram significância estatística. Todas pertencem ao conjunto de indicadores representado pela renda, saúde e educação. No grupo renda os indicadores de  $RP_{(-)}$  e  $VP_{(+)}$  contribuíram para probabilidade de inserção no círculo vicioso – o primeiro com efeito negativo e o segundo positivo. No grupo de indicadores de saúde estatisticamente significativos  $EV_{(-)}$  e  $PS60_{(-)}$  diminuem as chances de inserção no círculo da pobreza, enquanto  $Mi_{(+)}$  aumenta essa probabilidade. No que diz respeito àquelas variáveis que mensuram a educação,  $FC18_{(-)}$ ,  $SC25_{(-)}$ ,  $EAE_{(-)}$  apresentaram efeitos significativos no sentido negativo.



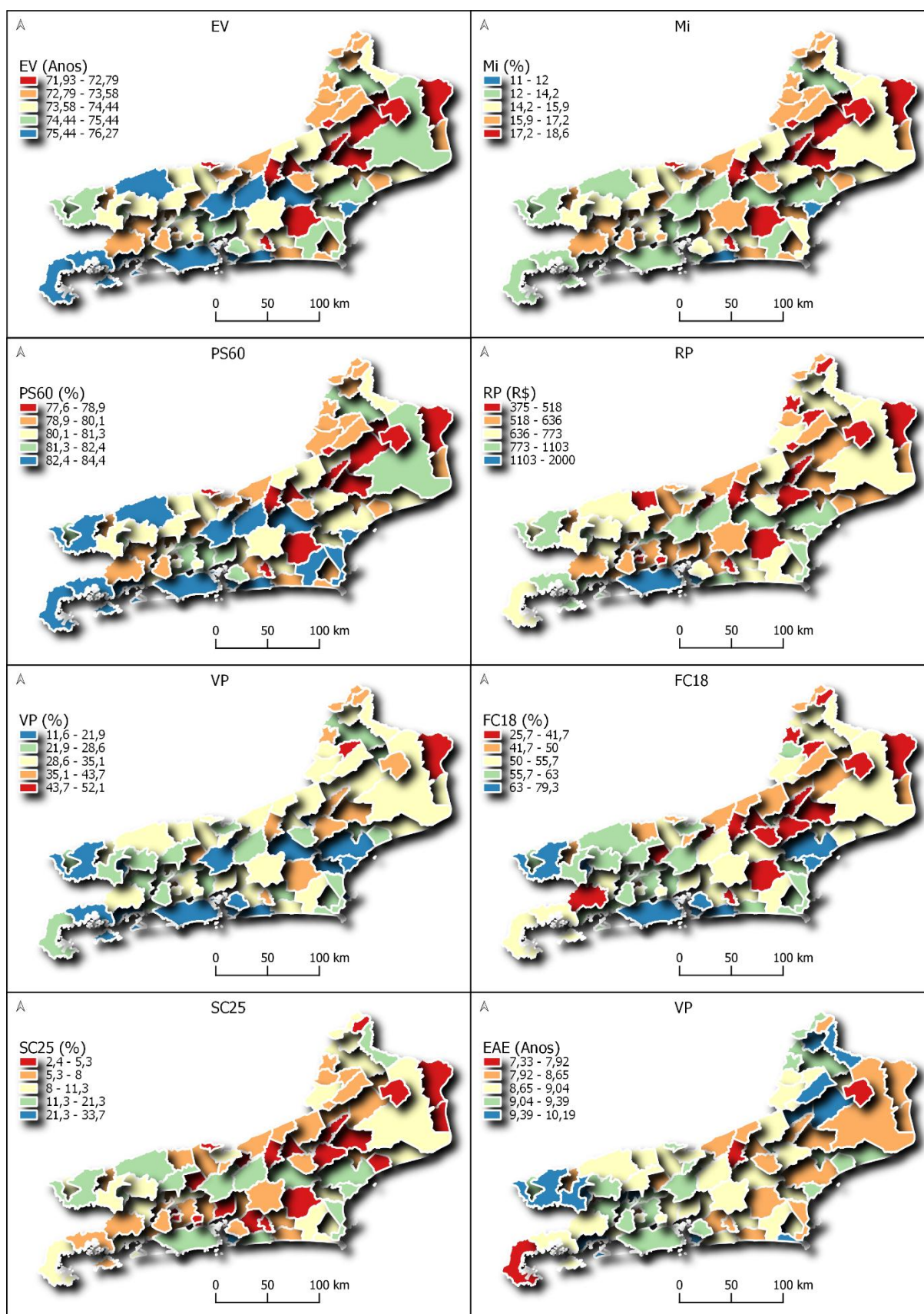
## 4.5 MAPAS

Os cartogramas das variáveis significativas do modelo foram apresentados na Figura 18. Cinco faixas foram estabelecidas de acordo com o método de otimização de Jenks (quebras naturais), que minimiza a variância entre as classes. As cores seguem o gradiente espectral, variando de vermelho a azul, a cor vermelha, aqui, sempre indica uma região crítica. Ao observar os mapas percebe-se uma preponderância de regiões mais críticas (amarelo, laranja e vermelho), principalmente no norte e noroeste fluminense. No Apêndice A encontram-se os mapas de todas as variáveis quantitativas

Na Figura 19 encontra-se o cartograma com a combinação linear das variáveis significantes no modelo de regressão. Vale ressaltar que as variáveis foram normalizadas (escala de 0 a 1), antes e depois da combinação, para a criação do indicador. As faixas foram definidas de acordo com a escala do IDH estabelecidas pelo PNUD. Dessa maneira foi possível observar os municípios mais vulneráveis ao círculo vicioso da pobreza em relação a todas variáveis significativas.

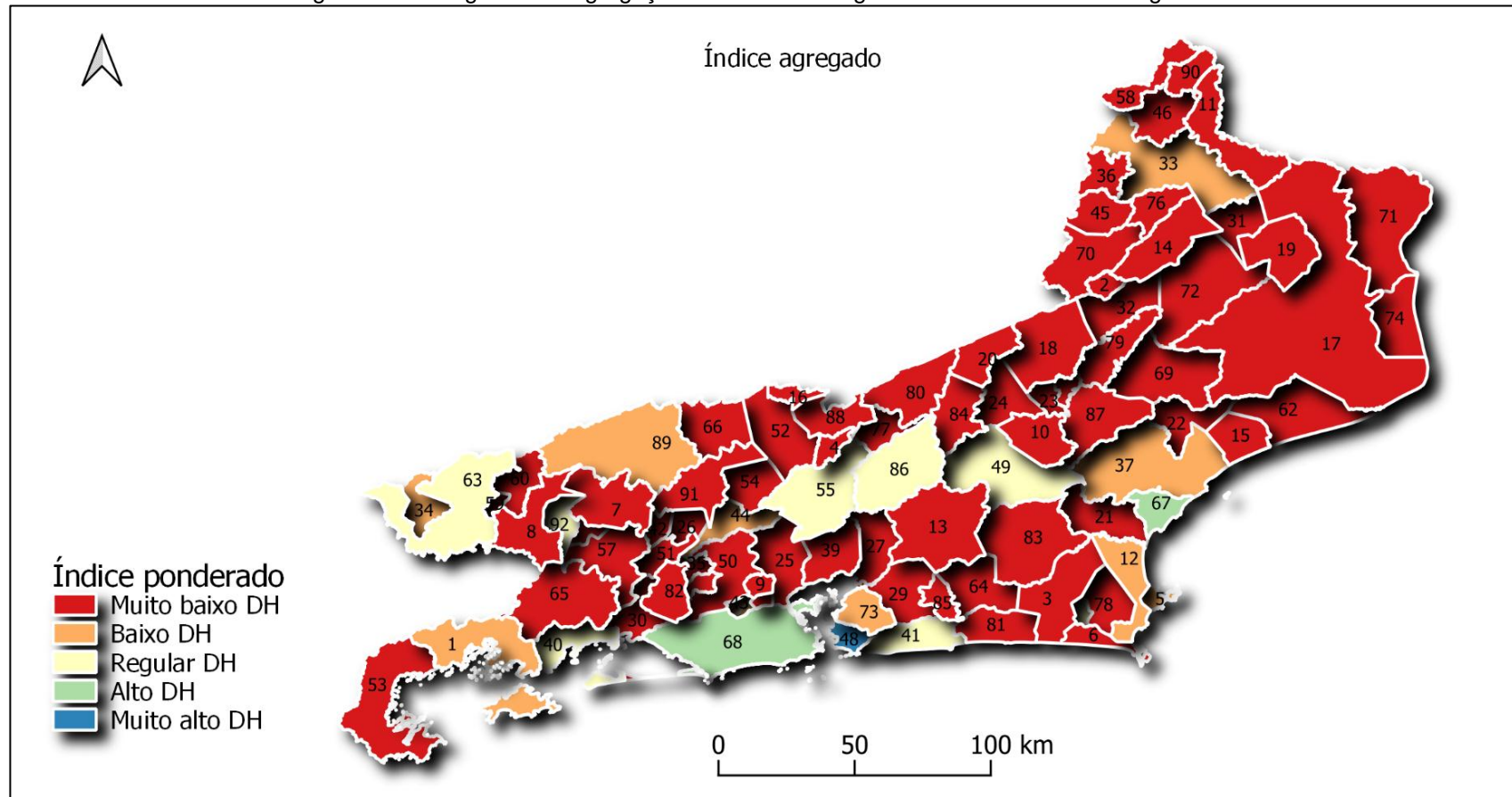
Dos 92 municípios, 70 (76%) estão na categoria de muito baixo DH; 11 (12%) na classe baixo DH; 8 (9%) na faixa intermediária; 2 (2%) com alto DH; e 1 com altíssimo DH. Tomando como base a classificação anterior dos municípios (seção 4.2) as classes extremas tendem a convergir; a categoria círculo vicioso representa 71% contra 76% para baixíssimo DH; e a classe virtuoso 3% contra 2% alto e 1% altíssimo DH. Ambos os mapas auxiliam no processo de tomada de decisão ao apontar as regiões mais vulneráveis, todavia os cartogramas apresentados nesta seção fornecem uma visão holística do problema, de modo a permitir um direcionamento pontual das políticas públicas de acordo com as necessidades de cada município.

Figura 18. Cartograma dos atributos significativos no modelo de regressão



Fonte: Autor

Figura 19. Cartograma da agregação das variáveis significativas no modelo de regressão



**Legenda:** 1-Angra dos Reis; 2-Aperibé; 3-Araruama; 4-Areal; 5-Armação dos Búzios; 6-Arraial do Cabo; 7-Barra do Pirai; 8-Barra Mansa; 9-Belford Roxo; 10-Bom Jardim; 11-Bom Jesus do Itabapoana; 12-Cabo Frio; 13-Cachoeiras de Macacu; 14-Cambuci; 17-Campos dos Goytacazes; 18-Cantagalo; 15-Carapebus; 19-Cardoso Moreira; 20-Carmo; 21-Casimiro de Abreu; 16-Comendador L. Gasparian; 22-Conceição de Macabu; 23-Cordeiro; 24-Duas Barras; 25-Duque de Caxias; 26-Engenheiro P. de Frontin; 27-Guapimirim; 28-Iguaba Grande; 29-Itaboraí; 30-Itaguaí; 31-Italva; 32-Itaocara; 33-Itaperuna; 34-Itatiaia; 35-Japeri; 36-Laje do Muriaé; 37-Macaé; 38-Macuco; 39-Magé; 40-Mangaratiba; 41-Maricá; 42-Mendes; 43-Mesquita; 44-Miguel Pereira; 45-Miracema; 46-Natividade; 47-Nilópolis; 48-Niterói; 49-Nova Friburgo; 50-Nova Iguaçu; 51-Paracambi; 52-Paraíba do Sul; 53-Parati; 54-Paty do Alferes; 55-Petrópolis; 56-Pinheiral; 57-Pirai; 58-Porciúncula; 59-Porto Real; 60-Quatis; 61-Queimados; 62-Quissamã; 63-Resende; 64-Rio Bonito; 65-Rio Claro; 66-Rio das Flores; 67-Rio das Ostras; 68-Rio de Janeiro; 69-Santa Maria Madalena; 70-Santo Antônio de Pádua; 71-São F. de Itabapoana; 72-São Fidélis; 73-São Gonçalo; 74-São João da Barra; 75-São João de Meriti; 76-São José de Ubá; 77-São José do V. do R. Preto; 78-São Pedro da Aldeia; 79-São Sebastião do Alto; 80-Sapucaia; 81-Saquarema; 82-Seropédica; 83-Silva Jardim; 84-Sumidouro; 85-Tanguá; 86-Teresópolis; 87-Trajano de Moraes; 88-Três Rios; 89-Valença; 90-Varre-Sai; 91-Vassouras; 92-Volta Redonda.

Fonte: Autor

## 5 DISCUSSÃO

O primeiro trabalho a analisar a conexão entre crescimento econômico e desenvolvimento humano foi o de RANIS; STEWART; RAMIREZ (2000). Os autores mostram resultados tanto empíricos como teóricos sobre a existência da conexão de causalidade entre CE e DH em ambos os sentidos. Suri *et al.* (2011) reafirmam essas conclusões. Ambos os trabalhos deixam claro que a conexão entre CE e DH não é automática, ou seja, entre o círculo virtuoso e vicioso há situações de desequilíbrios: tendendo ao crescimento e tendendo ao desenvolvimento.

O trabalho de Bayraktar-Sağlam (2017) explorou esses desequilíbrios ao estudar a interação entre saúde, educação e crescimento. Os principais achados mostraram uma relação complexa entre as interações desses elementos e que somente nos países de renda alta da OCDE houve uma relação de reforço mútuo entre saúde, educação e crescimento econômico.

Mustafa, Rizov e Kernohan (2017) investigaram a causalidade entre CE e DH na economia asiática, os resultados mostraram que embora o DH contribuísse para o CE a relação inversa não foi observada. Similarmente, Chikalipah e Okafor (2019) e Chikalipah e Makina (2019) não encontraram a relação de mão dupla entre CE e DH, apenas o efeito do crescimento econômico no desenvolvimento humano no longo prazo, respectivamente na Nigéria e Zâmbia. Para as três pesquisas, o denominador comum das possíveis causas da não causalidade bilateral se dá pela não promoção do capital humano e pelo atraso do desenvolvimento institucional.

No Estado do Rio de Janeiro 71% (65) dos municípios encontraram-se no círculo vicioso da pobreza. Isso indica uma conexão causal entre CE e DH no sentido

negativo para a maioria dos municípios fluminenses. Essa proporção foi semelhante aos achados de Raiher e Lima (2014) na região Sul do Brasil, com 83% e 73% dos municípios no círculo vicioso, respectivamente a média brasileira e da Região Sul; aos resultados de Oliveira, Lima e Raiher, (2017), na Região Nordeste, 84% dos municípios no círculo vicioso quando comparado com a média do Brasil, todavia, 24% quando comparado com a média da própria região; e Oliveira, Lima e Barrinha, (2019), no Estado da Bahia, 83% dos municípios do círculo da pobreza, média do Brasil, e 18% em relação à média do próprio estado. Uma vez iniciado o processo de subdesenvolvimento ou de desenvolvimento humano, a chance de regressão é menor dado o processo de causação circular cumulativa (RAIHER; LIMA, 2014; OLIVEIRA; LIMA; RAIHER, 2017; OLIVEIRA; LIMA; BARRINHA, 2019).

Ao analisar os indicadores socioeconômicos dos municípios fluminenses constatou-se que as variáveis mortalidade infantil (Mi), taxa de analfabetismo adulto (T18), % de vulneráveis à pobreza (VP) e % de pobres (P) foram relacionadas positivamente entre si. Essas variáveis ainda foram correlacionadas negativamente com % de 25 anos ou mais com superior completo (SC25), % de 18 anos ou mais com fundamental completo (FC18), renda *per capita* (RP), esperança de vida ao nascer (EV) e probabilidade de sobrevivência até 60 anos (PS60). As variáveis EV (32,63%), PS60 (22,52%), RP (0,12%), FC18 (3,87%), SC25 (9,05%) e EAE (64,62%) reduzem a chance,  $\alpha=5\%$ , de um município passar para o CVic a cada unidade incrementada; por outro lado, Mi e VP aumentam essas chances em 23,99% e 3,63%, respectivamente.

Isso demonstra que os municípios que apresentaram altas taxas de mortalidade infantil também tendem a apresentar altos índices de analfabetismo e pobreza, concomitantemente, baixos índices de escolaridade – tanto no ensino superior quanto no fundamental – menor renda per capita e longevidade.

Os resultados encontrados foram corroborados pelos trabalhos de Raiher e Lima (2014) e Oliveira, Lima e Raiher. (2017) em relação às variáveis que funcionam como *proxy* para educação, e para a variável PIB *per capita*. O PIB, por si, não influenciou estatisticamente a probabilidade de inserção ou manutenção do ciclo da pobreza. Tal resultado também foi corroborado pela literatura, que observou-se o desacoplamento do PIB com o desenvolvimento humano. Todavia, os resultados desses autores divergem quanto ao índice Gini, tanto nos efeitos como na significância. A desigualdade medida pelo coeficiente de Gini não mostrou evidências

significativas no modelo desta Tese. O trabalho de Stewart (2019) também demonstrou resultados semelhantes em que a desigualdade não teve efeito no DH, embora o próprio autor esperasse o contrário, amparado por outros estudos. Conforme colocado por Castells-Quintana, Royuela, Thiel, (2019), a relação entre desigualdade e DH é complexa, com diferentes relações no curto prazo entre desigualdade e as componentes do IDH. Os autores identificaram, no longo prazo, um efeito negativo da desigualdade no crescimento econômico; já no curto prazo um efeito contrário, encontrado principalmente em países em desenvolvimento.

Em suma, pode-se observar, no século XXI, que as mudanças na estrutura econômica do estado, deslocando a indústria de fabricação para a indústria extrativa e aumentando sua representatividade no interior do Estado, não levou os municípios ao círculo virtuoso. O CE proporcionado pela expansão da indústria de petróleo e gás natural (P&G) não resultou em um aumento do DH como um todo. Observação semelhante foi constatada por Acruche *et al.* (2019), mostrando que os municípios do Estado do Rio de Janeiro que recebiam maiores rendas petrolíferas não foram aqueles com melhores desempenhos no DH. Não basta ter apenas um bom desempenho econômico ou aspectos sociais positivos para estabelecer um círculo virtuoso de desenvolvimento (RAIHER; LIMA, 2014).

Isso demonstra que, independente da influência causal circular, o crescimento econômico, um contributo importante para a melhoria do desenvolvimento humano, não é em si sustentável sem investimentos em saúde e educação, tanto prévia como simultaneamente (RANIS; STEWART, 2005; SURI *et al.*, 2011). Esse resultado mostra a importância de coevolução das variáveis econômica, de um lado, e de outro as de desenvolvimento humano (saúde e educação).

## 6 CONCLUSÃO

O trabalho teve como objetivo compreender a relação entre o crescimento econômico e o desenvolvimento humano nos municípios do Estado do Rio de Janeiro e verificar quais os fatores influenciaram significativamente na formação do círculo vicioso da pobreza.

No Estado do Rio de Janeiro 71% (65), dos municípios encontraram-se no círculo vicioso da pobreza. Isso indica uma conexão causal entre CE e DH não virtuosa para a maioria dos municípios fluminenses, ou seja, apesar de eles terem apresentado CE isso não foi captado pelos indicadores de DH.

De fato, ao analisar os indicadores socioeconômicos dos municípios fluminenses constatou-se que as variáveis mortalidade infantil (Mi), taxa de analfabetismo adulto (T18), % de vulneráveis à pobreza (VP) e % de pobres (P) foram relacionadas positivamente entre si. Essas variáveis ainda foram correlacionadas negativamente com % de 25 anos ou mais com superior completo (SC25), % de 18 anos ou mais com fundamental completo (FC18), renda *per capita* (RP), esperança de vida ao nascer (EV) e probabilidade de sobrevivência até 60 anos (PS60). As variáveis EV (32,63%), PS60 (22,52%), RP (0,12%), FC18 (3,87%), SC25 (9,05%) e EAE (64,62%) reduzem a chance,  $\alpha=5\%$ , de um município passar para o CVic a cada unidade incrementada; por outro lado, Mi e VP aumentam essas chances em 23,99% e 3,63%, respectivamente.

Isso demonstra que os municípios que apresentaram altas taxas de mortalidade infantil também tendem a apresentar altos índices de analfabetismo e pobreza,

concomitantemente, baixos índices de escolaridade – tanto no ensino superior quanto no fundamental – menor renda per capita e longevidade.

Em suma, pode-se observar, no século XXI, que as mudanças na estrutura econômica do estado – deslocando a indústria de fabricação para a indústria extrativa e aumentando sua representatividade no interior – não levou os municípios ao círculo virtuoso. O CE proporcionado pela expansão da indústria de petróleo e gás natural (P&G) não resultou em um aumento do DH como um todo. Esse resultado mostra a importância de coevolução das variáveis econômica, de um lado, e de outro as de desenvolvimento humano (saúde e educação).

A metodologia utilizada, principalmente a exploratória, forneceu uma visão holística das variáveis e seus inter-relacionamentos, o que ajudou a entender a natureza subjacente dos dados. O modelo de regressão logística penalizada foi utilizado para confirmar os *insights* obtidos na análise exploratória dos dados. A técnica de penalização foi empregada devido à alta multicolinearidade encontrada nas variáveis.

## 6.1 SUGESTÃO DE TRABALHOS FUTUROS

A pesquisa foi realizada nos municípios do Estado do Rio de Janeiro. Como trabalhos futuros sugere-se a extensão para outras regiões e uma análise comparativa que permita compreender a dinâmica intrarregional e inter-regional. Também se sugere a utilização de outros modelos (Quadro 1) de regressão que visam a determinar a direção empírica de causalidade entre CE e DH, ora unidirecional ora bidirecional. Esses modelos demandam dados estruturados em painel, sendo necessário observar os indicadores ao longo do tempo, o que ainda não foi possível nesta Tese devido a indisponibilidade pública de dados.



## REFERÊNCIAS

ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Wiley interdisciplinary reviews: computational statistics**, Hoboken, New Jersey, EUA, v. 2, n. 4, p. 433-459, 2010.

ABDI, H.; WILLIAMS, L. J.; VALENTIN, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. **Wiley Interdisciplinary reviews: computational statistics**, Hoboken, New Jersey, EUA, v. 5, n. 2, p. 149-179, 2013.

ACEMOGLU, D.; JOHNSON, S.; ROBINSON, J. A. Institutions as a fundamental cause of long-run growth. **Handbook of economic growth**, Elsevier, Amsterdã, NLD v. 1, p. 385-472, 2005.

ACRUCHE, R. *et al.* Relação entre renda petrolífera e melhoria na qualidade de vida por municípios confrontantes da Bacia de Campos. **Cadernos do Desenvolvimento Fluminense**, Rio de Janeiro, v. 16, p. 27, 2019.

Atlas Brasil. Notas metodológicas: IDHM. 2021. Disponível em: <http://www.atlasbrasil.org.br/acervo/biblioteca>. Acesso em: 04 set. 2021.

BAYRAKTAR-SAĞLAM, B. Re-Examining Vicious Circles of Development: A Panel Var Approach. **Social Indicators Research**, Nature, Switzerland, AG, v. 137, n. 1, p. 231-256, 2017.

BEHRMAN, J. R. The impact of health and nutrition on education. **The World Bank Research Observer**, Oxford, UK, v. 11, n. 1, p. 23-37, 1996.

BONHAM-CARTER, G.F. **Geographic Information Systems for Geoscientists**. Oxford: Pergamon, 1994.

CASTELLS-QUINTANA, D.; ROYUELA, V.; THIEL, F. Inequality and sustainable development: Insights from an analysis of the human development index. **Sustainable Development**, Hoboken, New Jersey, EUA, v. 27, n. 3, p. 448–460, 2019.

CHENERY, H. *et al.* **Redistribution with growth; policies to improve income distribution in developing countries in the context of economic growth**. Oxford: Oxford University Press, 1974.

CHIKALIPAH, S.; OKAFOR, G. Dynamic Linkage Between Economic Growth and Human Development: Time Series Evidence from Nigeria. **Journal of International Development**, Hoboken, New Jersey, EUA, v. 31, n. 1, p. 22-38, jan. 2019.

**Computational statistics & data analysis**, Amsterdam, NLD, v. 18, n. 1, p. 121-140, 1994.

CULE, E.; IORIO, M. Ridge regression in prediction problems: automatic choice of the ridge parameter. **Genetic epidemiology**, Hoboken, New Jersey, EUA, v. 37, n. 7, p. 704-714, 2013.

DENT, B.; TORGUSON, J.; HODLER, T. **Thematic map design**. New York, EUA: McGraw-Hill, 2008.

DUNTEMAN, George H. **Principal components analysis**. Chicago, Illinois, EUA: Sage, 1989.

ELSEVIER. Sobre a solução Scopus. 2019. Disponível em: <https://www.elsevier.com/solutions/scopus/how-scopus-works>. Acesso em: 17 maio 2019

ESCOFIER, B.; PAGES, J. Multiple factor analysis (AFMULT package).

FONTI, Valeria. **Feature Selection using LASSO**. Amsterdam: Vrije, 2017.

FRANK, L. E.; FRIEDMAN, J. H. A statistical view of some chemometrics regression tools. **Technometrics**, Boston, Massachusetts, EUA, v. 35, n. 2, p. 109-135, 1993.

GOEMAN, J. J. L1 penalized estimation in the Cox proportional hazards model. **Biometrical journal**, Hoboken, New Jersey, EUA, v. 52, n. 1, p. 70-84, 2010.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. Porto Alegre: Amgh Editora, 2011.

HADDAD, L. *et al.* Reducing child malnutrition: How far does income growth take us? **The World Bank Economic Review**, Oxford, UK, v. 17, n. 1, p. 107-131, 2003.

HAIR, J. F. *et al.* **Análise multivariada de dados**. USA: Bookman, 2009.

HOERL, E.; KENNARD, R. W. Ridge regression: Biased estimation for nonorthogonal problems. **Technometrics**, Boston, Massachusetts, EUA, v. 12, n. 1, p. 55-67, 1970.

HUSSON, F.; LÊ, S.; PAGÈS, J. **Exploratory multivariate analysis by example using R**. United Kingdom: CRC press, 2017.

INTERNATIONAL LABOUR OFFICE - ILO. **Employment, growth and basic needs: a one world problem.** Report of the Director-General of the International Labour Office. Geneva: ILO, 177 p. 1976.

JANUZZI, P. M. **Indicadores sociais no Brasil.** São Paulo: Alínea, 2003.

JOLLY, R. Adjustment with a Human Face. In: RICHARD JOLLY (UK). **IDS Companions to Development** (ed.). **Milestones and Turning Points in Development Thinking.** London: Palgrave Macmillan, 2012. p. 172-185.

JOLY, F. **A Cartografia.** São Paulo: Papirus, 2005.

KAKWANI, N. Structural adjustment and performance in living standards in developing countries. **Development and Change**, Hoboken, New Jersey, EUA, v. 26, n. 3, p. 469-502, 1995.

KASSAMBARA, A. **Practical Guide to Principal Component Methods in R: PCA, M (CA), FAMD, MFA, HCPC.** France: DATANOVIA, 2017.

LE CESSIE, S.; VAN HOUWELINGEN, J. C. Ridge estimators in logistic regression. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, Hoboken, Nova Jersey, EUA, v. 41, n. 1, p. 191-201, 1992.

MUKAKA, M.M. A guide to appropriate use of correlation coefficient in medical research. **Malawi medical journal**, Zomba, MW, v. 24, n. 3, p. 69-71, 2012.

MUSTAFA, G.; RIZOV, M.; KERNOHAN, D. Growth, human development, and trade: The Asian experience. **Economic Modelling**. [S.l.] v. 61, p. 93–101, 2017.

MYRDAL, G. **Teoría económica e regiones subdesarrolladas.** México: Fondo de Cultura Económica, 1968.

NORRIS, H. Flourishing, positive mental health and well-being: how can they be increased? **International Journal of Leadership in Public Services**, Bingley, UK, v. 6, n. 4, p. 46-50, 2010.

NORTH, D. C. Institutions. **Journal of economic perspectives**, Pittsburgh,

OLIVEIRA, N. S. M. N.; LIMA, J.; RAIHER, A. P. Convergência do desenvolvimento humano municipal no Nordeste do Brasil. **Revista Brasileira de Gestão e Desenvolvimento Regional**, Taubaté, SP, v. 13, n. 3, p. 164-184, 2017.

OLIVEIRA, N. S. M. N.; LIMA, J. F.; BARRINHA, R. N. Desenvolvimento humano municipal no Estado da Bahia. **Revista Brasileira de Desenvolvimento Regional**, Blumenau, SC, v. 7, n. 1, p. 137-164, 2019.

PAGÈS, J. **Multiple factor analysis by example using R.** United Kingdom: Chapman and Hall, 2014.

PAMPEL, F. C. **Logistic regression: a primer**. Chicago, Illinois, EUA: Sage, 2000

Pennsylvania, EUA, v. 5, n. 1, p. 97-112, 1991.

PNUD. **Human Development Report 1990**. New work: United Nations Development Programme, 1990.

PNUD. **Human Development Report Mexico**. Mexico City: Mundi-Prensa, 2002.

PNUD. **O que é o IDH**. 2020. Disponível em: <https://www.br.undp.org/content/brazil/pt/home/idh0/conceitos/o-que-e-o-idh.html>. Acesso em: 30 mar. 2020.

PSACHAROPOULOS, G. Returns to investment in education: A global update. **World development**, Amsterdam, NLD, v. 22, n. 9, p. 1325-1343, 1994.

QGIS. **A Free and Open Source Geographic Information System**. 2020. Disponível em: <https://qgis.org/en/site/>. Acesso em: 17 jun. 2020.

RAIHER, A. P.; LIMA, J. F. Desenvolvimento humano municipal no Sul do Brasil: evolução recente e o círculo vicioso da pobreza. **Acta Scientiarum. Human and Social Sciences**, v. 36, n. 2, p. 147-154, 2014.

RANIS, G.; STEWART, F. **Dynamic Links between the Economy and Human Development**. New York: Department Of Economic And Social Affairs, 2005.

RANIS, G.; STEWART, F.; RAMIREZ, A. Economic growth and human development. **World development**, Amsterdam, NLD, v. 28, n. 2, p. 197–219, 2000.

RAVIKUMAR, P. *et al.* High-dimensional Ising model selection using  $\ell_1$ -regularized logistic regression. **The Annals of Statistics**, EUA, v. 38, n. 3, p. 1287-1319, 2010.

ROSENZWEIG, M.R. Why are there returns to schooling? **The American Economic Review**, Pittsburgh, Pensilvânia, EUA, v. 85, n. 2, p. 153-158, 1995.

SCHREIBER-GREGORY, D. N. Ridge Regression and multicollinearity: An in-depth review. **Model Assisted Statistics and Applications**, Amsterdam, NLD, v. 13, n. 4, p. 359-365, 2018.

SEN, A.; MOTTA, L. T.; MENDES, R. D. **Desenvolvimento como liberdade**. São Paulo: Companhia das Letras, 2000.

SILVA, F. F. *et al.* Evolução do desenvolvimento humano dos municípios do Rio de Janeiro entre 2005 e 2016. **Revista Tecnologia e Sociedade (online)**, Curitiba, PR, v. 16, p. 76, 2019.

SILVA, J. C. A. L. **Saúde e Desenvolvimento: uma análise econométrica, 1995 a 2014**. 2019. 178 f. Tese (Doutorado) - Curso de Programa de Pós-Graduação em Economia, Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

STEWART, F. The Human Development Approach: An Overview. **Oxford Development Studies**, Oxfordshire, UK v. 47, n. 2, p. 135-153, 2019.

STOLTZFUS, J. C. Logistic regression: a brief primer. **Academic Emergency Medicine**, Hoboken, New Jersey, EUA, v. 18, n. 10, p. 1099-1104, 2011.

STRAUSS, J.; THOMAS, D. Human resources: Empirical modeling of household and family decisions. **Handbook of development economics**, Amsterdam, NLD, v. 3, p. 1883-2023, 1995.

SURI, T. *et al.* Paths to success: The relationship between human development and economic growth. **World Development**, Amsterdam, NLD, v. 39, n. 4, p. 506-522, 2011.

TABACHNICK, B. G.; FIDELL, L. S. **Using multivariate statistics**. 6. ed. Boston: Pearson, 2013.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, Hoboken, New Jersey, EUA, v. 58, n. 1, p. 267-288, 1996.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso: a retrospective. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Hoboken, New Jersey, EUA, v. 73, n. 3, p. 273-282, 2011.

TOMIC, O.; BERGET, I.; NÆS, T. A comparison of generalised procrustes analysis and multiple factor analysis for projective mapping data. **Food quality and preference**, Amsterdam, NLD, v. 43, p. 34-46, 2015.

UNDP. **Human Development Report**: 1991. New York: Oxford University Press, 1990.

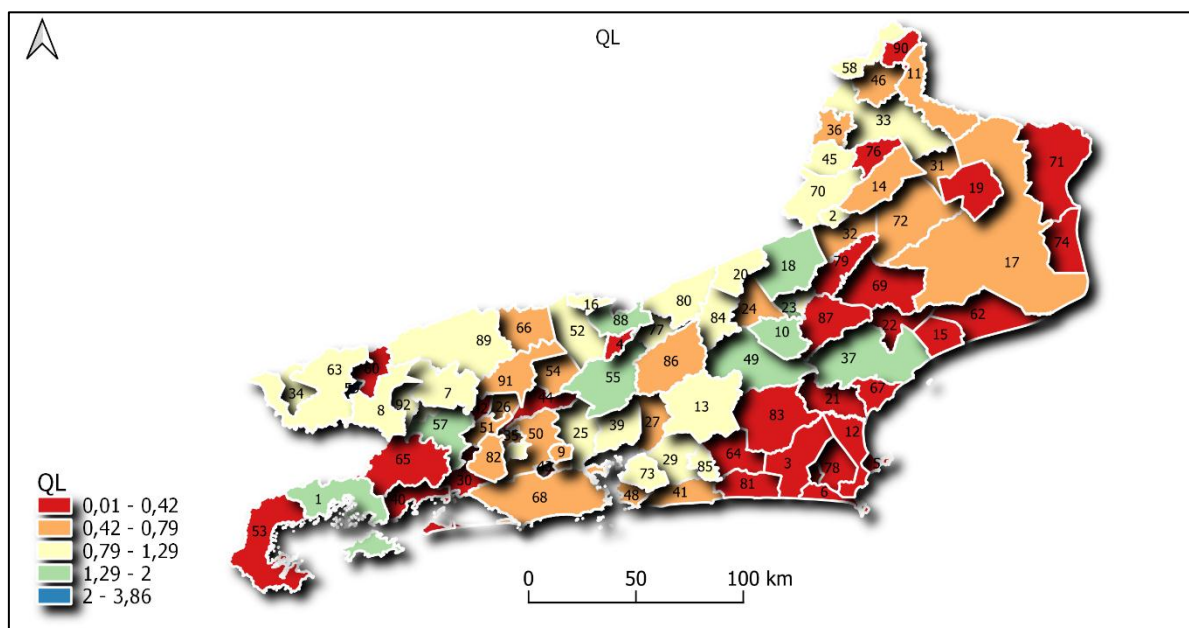
VERGARA, Sylvia Constant. **Projetos e relatórios de pesquisa em administração**. São Paulo: Atlas, 1997

ZOU, H.; HASTIE, T. Regression shrinkage and selection via the elastic net, with applications to microarrays. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Hoboken, New Jersey, EUA, v. 67, p. 301-20, 2003.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. **Journal of the royal statistical society: series B (statistical methodology)**, Amsterdam, NLD, v. 67, n. 2, p. 301-320, 2005.

## APÊNDICE A: Cartograma das variáveis quantitativas

Figura A1. Variável QL



Fonte: Autor

Figura A2. Variáveis Mi, SC25, RP

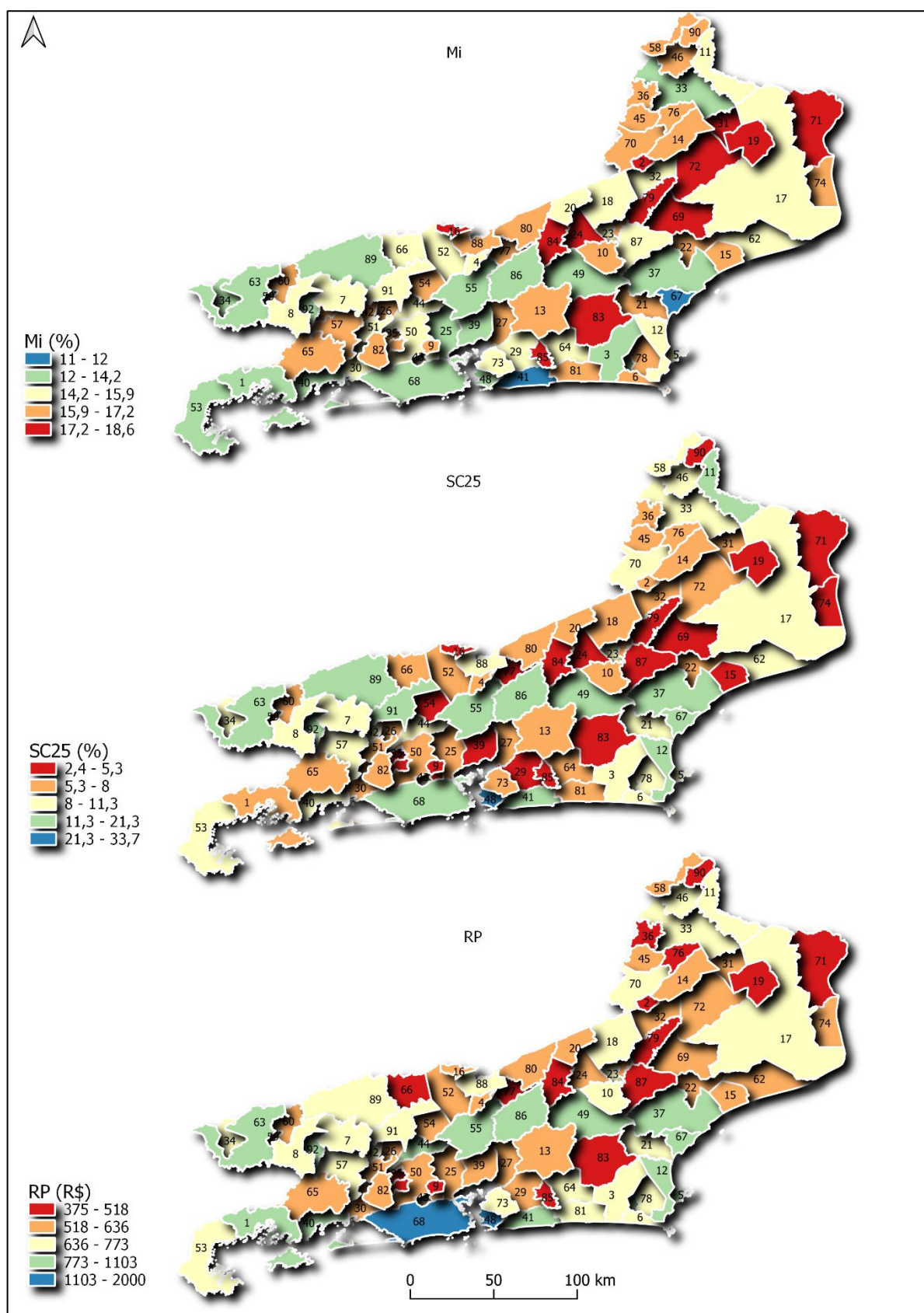
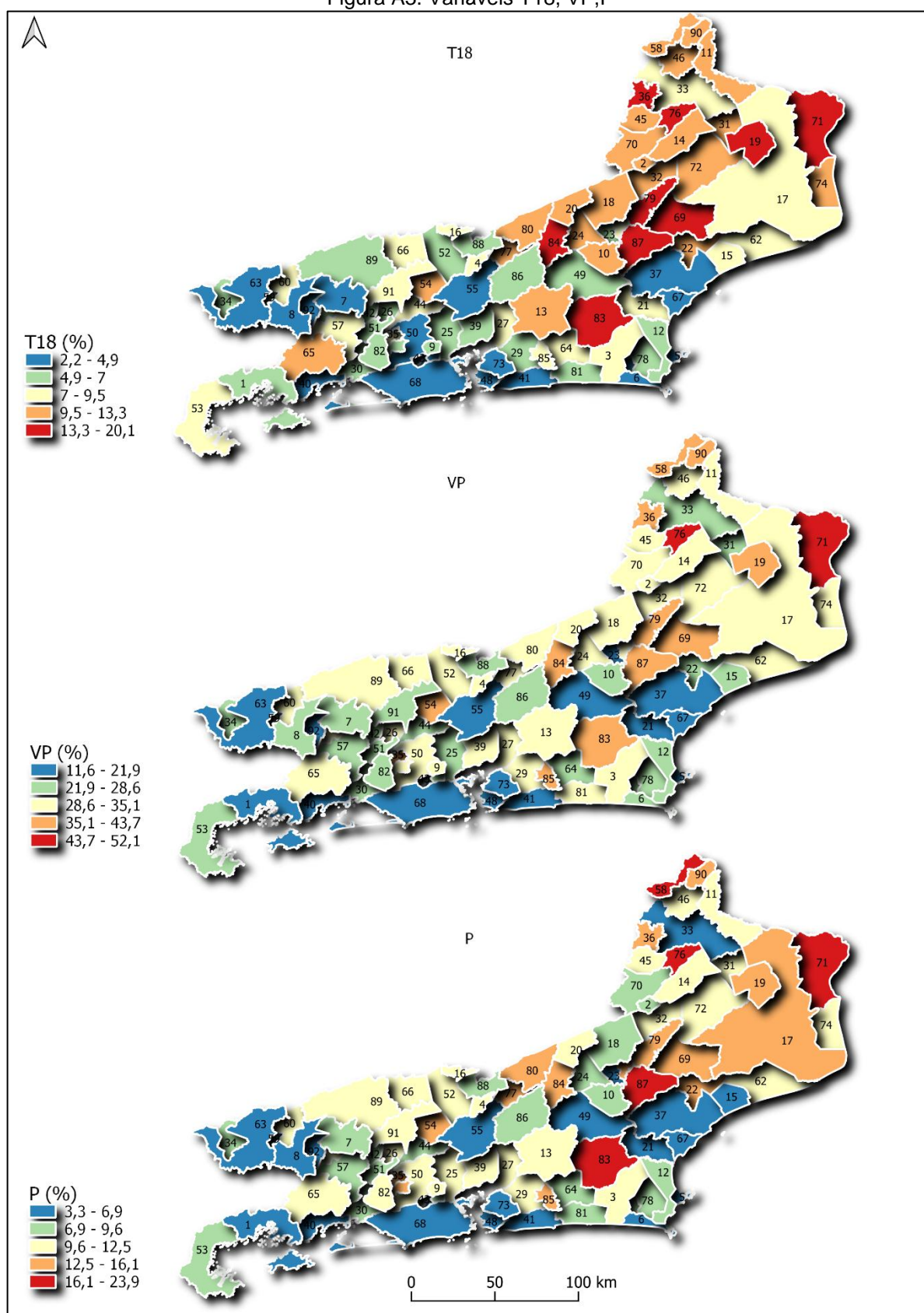




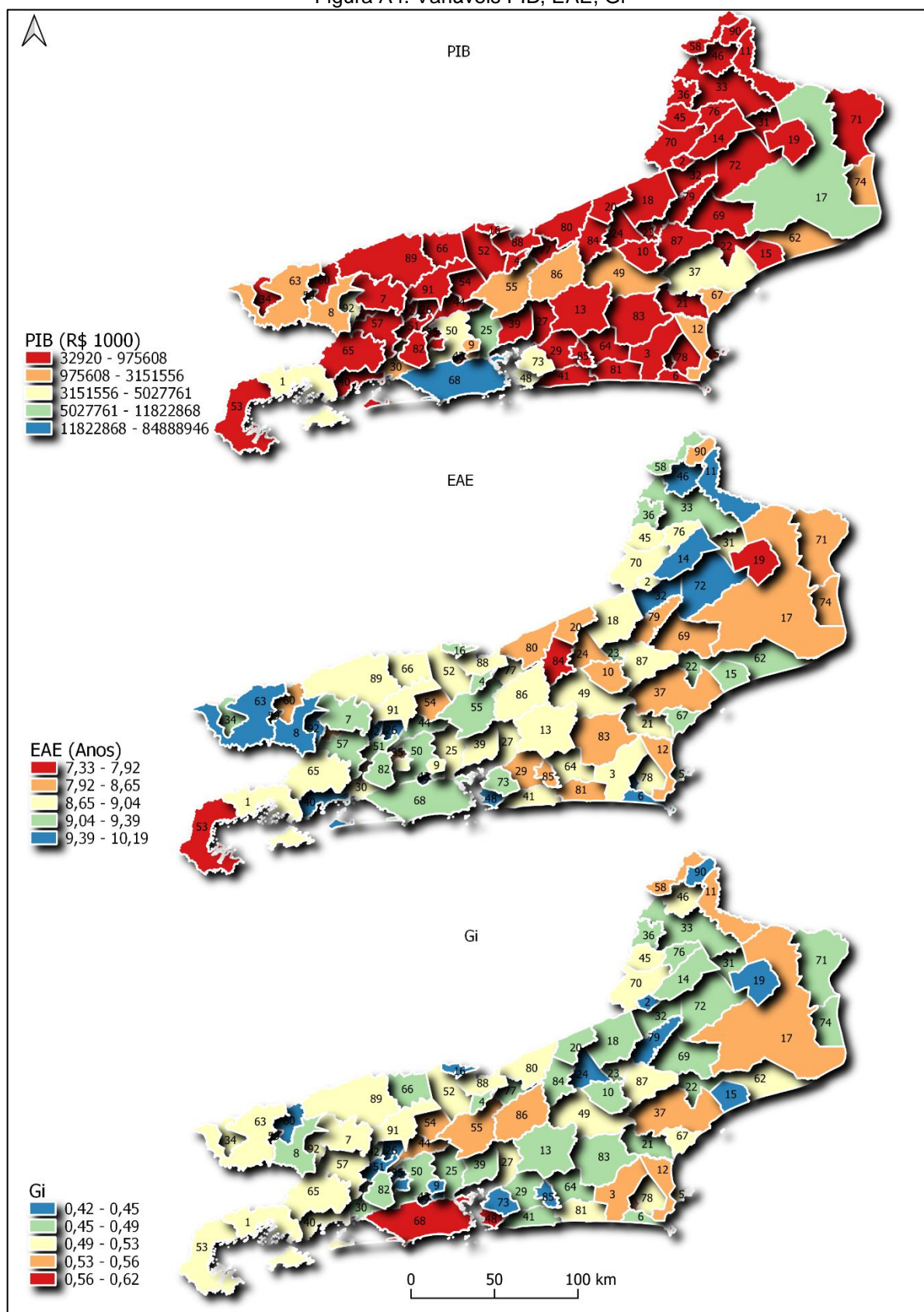
Figura A3. Variáveis T18, VP,P



Fonte: Autor

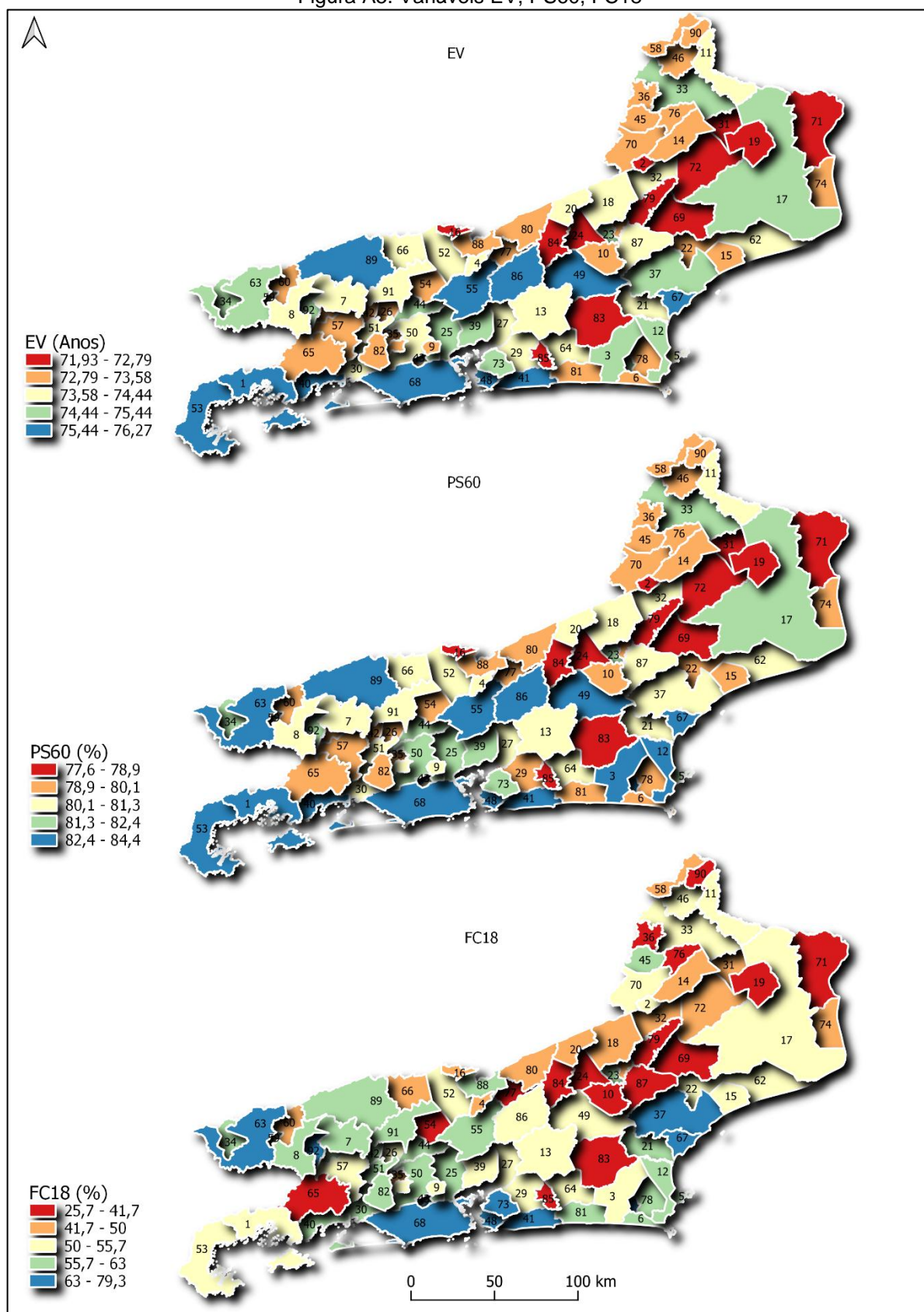


Figura A4. Variáveis PIB, EAE, Gi



Fonte: Autor

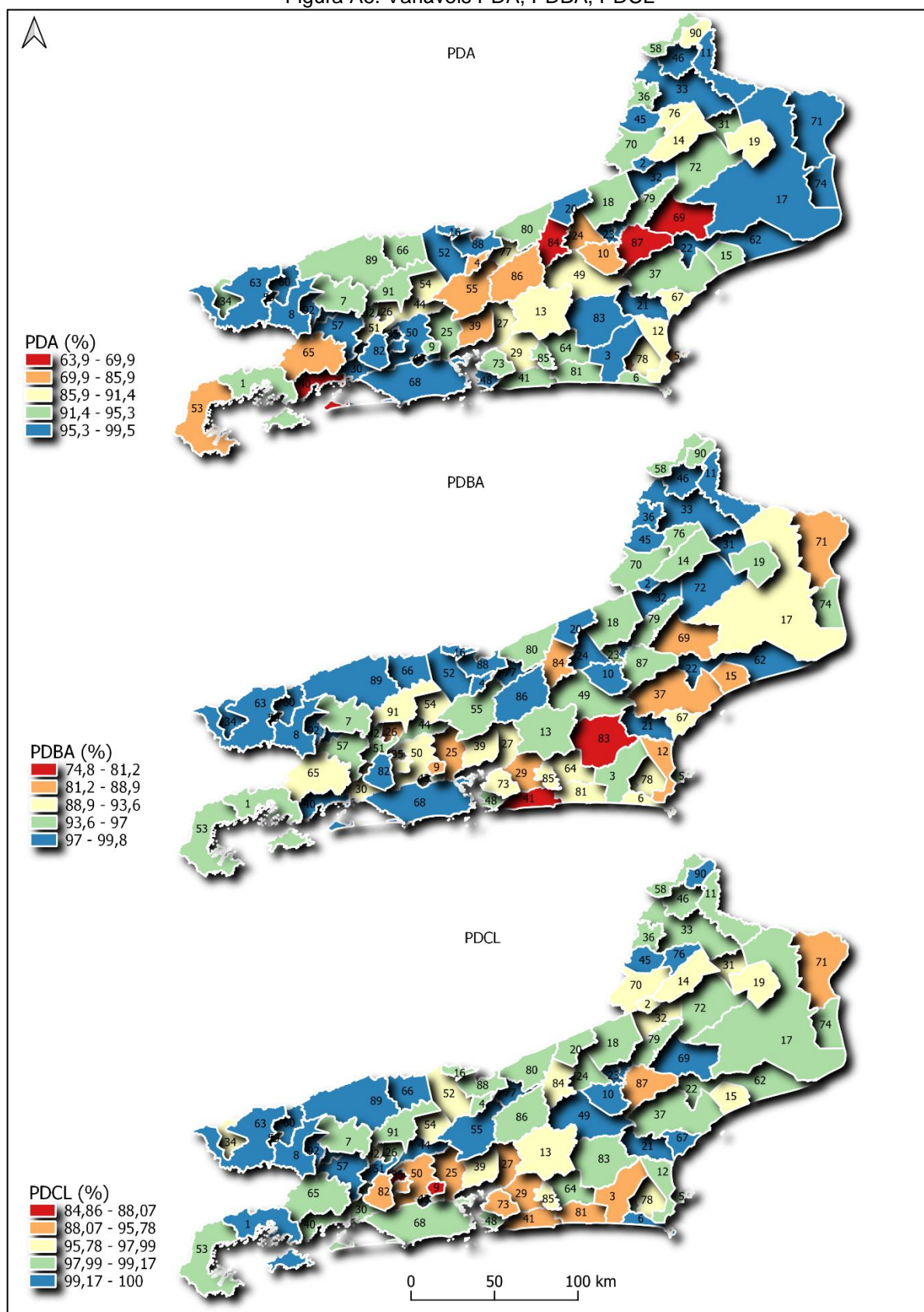
Figura A5. Variáveis EV, PS60, FC18



Fonte: Autor



Figura A6. Variáveis PDA, PDBA, PDCL



Fonte: Autor