

**Gilberto Martins dos Santos**

**Evolução Espaço-temporal das Matrículas  
na Modalidade de Ensino Educação de  
Jovens e Adultos no Estado do Rio de  
Janeiro**

Niterói - RJ, Brasil

13 de Julho de 2017

**Gilberto Martins dos Santos**

**Evolução Espaço-temporal das  
Matrículas na Modalidade de Ensino  
Educação de Jovens e Adultos no  
Estado do Rio de Janeiro**

**Trabalho de Conclusão de Curso**

Monografia apresentada para obtenção do grau de Bacharel em  
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. Ludmilla da Silva Viana Jacobson

Niterói - RJ, Brasil

13 de Julho de 2017

**Gilberto Martins dos Santos**

**Evolução Espaço-temporal das Matrículas na  
Modalidade de Ensino Educação de Jovens e  
Adultos no Estado do Rio de Janeiro**

Monografia de Projeto Final de Graduação sob o título “*Evolução Espaço-temporal das Matrículas na Modalidade de Ensino Educação de Jovens e Adultos no Estado do Rio de Janeiro*”, defendida por Gilberto Martins dos Santos e aprovada em 13 de Julho de 2017, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

---

**Profa. Dra. Ludmilla da Silva Viana Jacobson**  
Departamento de Estatística – UFF

---

**Prof. Dr. Licínio Esmeraldo da Silva**  
Departamento de Estatística – UFF

---

**Profa. Dra. Márcia Marques de Carvalho**  
Departamento de Estatística – UFF

Ficha catalográfica automática - SDC/BIME  
Gerada com informações fornecidas pelo autor

S237e Santos, Gilberto Martins dos  
Evolução espaço-temporal das matrículas na modalidade de ensino Educação de Jovens e Adultos no Estado do Rio de Janeiro / Gilberto Martins dos Santos ; Ludmilla da Silva Viana Jacobson, orientadora. Niterói, 2017.  
52 f. : il.

Trabalho de Conclusão de Curso (Graduação em Estatística)-Universidade Federal Fluminense, Instituto de Matemática e Estatística, Niterói, 2017.

1. Educação de Jovens e Adultos. 2. Análise espacial. 3. Regressão polinomial. 4. Produção intelectual. I. Jacobson, Ludmilla da Silva Viana, orientadora. II. Universidade Federal Fluminense. Instituto de Matemática e Estatística. III. Título.

CDD -

# Agradecimentos

Agradeço em primeiro lugar a Deus que iluminou o meu caminho e me deu forças durante esta longa caminhada.

Agradeço aos meus pais, meu irmão, minhas sobrinhas e a toda minha família que sempre acreditaram e me apoiaram durante toda essa jornada.

Agradeço às minhas amigas Martha, Tamiris e Juliana pelo apoio e carinho, mesmo eu estando tão ausente da vida delas nesses últimos anos.

Agradeço a minha amiga Nayara pela imensa ajuda e contribuição na minha vida acadêmica e profissional.

Agradeço ao curso de estatística da UFF, e as pessoas com quem convivi nesse espaço de tempo, especialmente Fernanda, Igor, Denise, Luciana, Eduardo China, Tamiris, Tayanie, Bruno, Jack, Cíntia e Evellyn.

Agradeço a minha orientadora Jaqueline Ventura do projeto de desenvolvimento acadêmico no qual participei por dois anos e me possibilitou a maturidade para trabalhar com o tema dessa monografia .

Agradeço à minha orientadora, Professora Ludmilla por toda ajuda, contribuição, conselhos e paciência não só durante a monografia, mas como em todo curso.

# Resumo

A Educação de Jovens e Adultos é uma modalidade de ensino da educação básica nos níveis fundamental e médio que tem como objetivo possibilitar pessoas adultas, que não tiveram a oportunidade de frequentar a escola na idade convencional, a retomar seus estudos. O objetivo principal do trabalho é avaliar a tendência e fatores associados ao número de matrículas na modalidade Educação para Jovens e Adultos no Rio de Janeiro. Através da análise de tendência, chegamos ao modelo quártico como aquele que melhor representa a curva dessa equação, e concluímos que o número de matrículas vem apresentando uma tendência de queda. O indicador da taxa de atendimento nos municípios sinaliza que grande parte deles ficou muito aquém do ideal, com muitos chegando a absurdos 0% de atendimento. Na análise de regressão, temos que as variáveis significativas "Taxa de Urbanização" e o "Número de Escolas", conseguem explicar a grande parte da variação na taxa de atendimento. Em linhas gerais, podemos constatar o grave cenário da Educação de Jovens e Adultos no estado do Rio de Janeiro e a necessidade de políticas públicas para conseguir alcançar essa grande demanda potencial existente no estado hoje.

Palavras-chaves:

Educação de Jovens e Adultos; Análise Espacial; Regressão Polinomial

# Sumário

**Agradecimentos**

**Lista de Figuras**

**Lista de Tabelas**

<b>1</b>	<b>Introdução</b>	p. 10
1.1	A História da EJA no Brasil . . . . .	p. 10
1.2	A EJA no Nível Médio no Rio de Janeiro . . . . .	p. 11
1.3	Direito à Educação . . . . .	p. 11
1.4	Organização do trabalho . . . . .	p. 12
<b>2</b>	<b>Objetivos</b>	p. 13
2.1	Objetivo Geral . . . . .	p. 13
2.2	Objetivos Específicos . . . . .	p. 13
<b>3</b>	<b>Materiais e Métodos</b>	p. 14
3.1	Banco de Dados . . . . .	p. 14
3.2	Modelos de Regressão . . . . .	p. 16
3.2.1	Estimadores dos Parâmetros do Modelo . . . . .	p. 17
3.2.2	Inferência do Modelo . . . . .	p. 18
3.2.2.1	Significância individual dos parâmetros do modelo . . . . .	p. 18
3.2.2.2	Significância global do modelo . . . . .	p. 19
3.2.3	Análise de Resíduos . . . . .	p. 19

3.2.3.1	Diagnóstico de Normalidade . . . . .	p. 19
3.2.3.2	Diagnóstico de Homoscedasticidade . . . . .	p. 20
3.2.3.3	Diagnóstico de Independência . . . . .	p. 20
3.2.3.4	Análise de Multicolinearidade . . . . .	p. 21
3.2.4	Medidas de qualidade do ajuste . . . . .	p. 21
3.2.4.1	Coeficiente de Determinação ( $R^2$ ) . . . . .	p. 21
3.2.4.2	Coeficiente de Determinação Ajustado ( $R_{aj}^2$ ) . . . . .	p. 22
3.2.4.3	Critério de Informação de Akaike (AIC) . . . . .	p. 22
3.2.5	P-valor . . . . .	p. 22
3.2.6	Modelos de Regressão Polinomial . . . . .	p. 22
3.2.7	Modelo Teórico . . . . .	p. 23
3.2.8	Método de Seleção das Variáveis . . . . .	p. 25
3.3	Análise Espacial . . . . .	p. 25
3.3.1	Análise Exploratória . . . . .	p. 25
3.3.1.1	Matriz de Proximidade Espacial . . . . .	p. 25
3.3.1.2	Índice Global de Moran . . . . .	p. 26
3.3.1.3	Índice Local de Associação Espacial . . . . .	p. 26
3.3.1.4	Diagrama de Espalhamento de Moran . . . . .	p. 27
3.3.1.5	Box Map, Lisa Map e Moran Map . . . . .	p. 27
<b>4</b>	<b>Análise dos Resultados</b>	p. 28
4.1	Análise de Tendência Utilizando Modelos de Regressão Polinomial . . . . .	p. 28
4.2	Análise Exploratória de Dados Espaciais . . . . .	p. 31
4.3	Indicador de Taxa de atendimento para EJA por Município . . . . .	p. 38
4.4	Análise de Regressão . . . . .	p. 39
<b>5</b>	<b>Conclusão</b>	p. 44



**Referências**

p. 46

**Anexo A - Número de Matrículas nos 92 municípios entre os anos de  
2007 e 2015**

p. 49

# Lista de Figuras

1	Diagrama com as variáveis que influenciam na taxa de atendimento . . .	p. 24
2	Número de matrículas em cada ano centralizado. . . . .	p. 28
3	Número de matrículas em cada ano centralizado, ajustado pelo modelo quártico com correção de Prais-Winsten . . . . .	p. 31
4	Mapa do Estado do Rio de Janeiro segundo o número de matrículas (por 1000 hab.) da EJA no período de 2007 a 2009. . . . .	p. 32
5	Mapa do Rio de Janeiro com as conexões entre os municípios considerando os 5 vizinhos mais próximos . . . . .	p. 33
6	Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 1º Triênio, 2007-2009 . . . . .	p. 34
7	Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 2º Triênio, 2010-2012 . . . . .	p. 35
8	Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 3º Triênio, 2013-2015 . . . . .	p. 36
9	Moran Map . . . . .	p. 37
10	Taxa de Atendimento nos Municípios do Estado do Rio de Janeiro em 2010 . . . . .	p. 38
11	Histograma dos resíduos (modelo com "outliers") . . . . .	p. 40
12	Boxplot dos resíduos (modelo com "outliers") . . . . .	p. 41
13	Histograma dos resíduos (modelo sem "outliers") . . . . .	p. 42
14	Gráfico resíduos versus valores ajustados (modelo sem "outliers") . . .	p. 43

# Lista de Tabelas

1	Estimativas dos parâmetros dos modelos polinomiais . . . . .	p. 29
2	P-valor dos testes de qualidade de ajuste dos modelos estimados . . . .	p. 29
3	Estimativas dos parâmetros dos modelos com a correção de Prais-Winsten	p. 30
4	P-valor dos testes de qualidade de ajuste dos modelos com correção de Prais-Winsten estimados . . . . .	p. 30
5	Índice Global de Moran e o nível descritivo . . . . .	p. 33
6	Estatísticas descritivas das variáveis resposta e explicativas-RJ-2010 . .	p. 39
7	Legenda das variáveis resposta e explicativas . . . . .	p. 39
8	Estimativas dos parâmetros do modelo sem intercepto . . . . .	p. 40
9	Estimativas dos parâmetros do modelo sem "outliers" e sem intercepto	p. 42

# 1 Introdução

A Educação de Jovens e Adultos (EJA) é uma modalidade de ensino voltada para uma parcela da população que não teve a oportunidade de concluir o ensino básico na idade recomendada e representa para esses cidadãos um ensejo para que possam regressar à sala de aula em busca do conhecimento, autoestima e de uma formação crítica, além de poder finalmente conquistar o seu sonhado diploma, no qual seguramente, representará uma nova oportunidade de conquistas pessoais, sejam elas sociais ou profissionais, em um mercado de trabalho cada vez mais concorrido.

## 1.1 A História da EJA no Brasil

O histórico da EJA no Brasil se inicia no período do Brasil Colônia, de uma maneira assistemática, porém, só no começo da República, movido pelo desenvolvimento industrial, que iniciativas governamentais mais efetivas começam a aparecer. Com a Constituição de 1934, a União passa a ter a responsabilidade de fixar o Plano Nacional de Educação, no qual define algumas normas, entre elas: o ensino primário integral, gratuito e de frequência obrigatória, sendo direito de todos (VIEIRA[1]).

No final dos anos 40, começa uma campanha de Educação de Jovens e Adultos, visando atender a população rural, focando principalmente na alfabetização em poucos meses e com o curso primário dividido em dois períodos de sete meses. Dando sequência, a próxima etapa seria a “Ação em profundidade”, destinada ao desenvolvimento comunitário e a capacitação profissional (RIBEIRO[2]).

As reflexões e discussões sobre o analfabetismo ganham um novo paradigma com Paulo Freire, que desenvolveu uma metodologia própria de trabalho, com uma nova pedagogia trabalhada para mudar essa relação no qual estava inserido o educando, aonde tinha a concepção de que a educação era um ato político, podendo servir tanto para a submissão como para a libertação do povo (RAUBER [3]).

Com o começo da ditadura militar em 1964, houve um rompimento com uso das propostas e métodos sugeridos por Freire. A educação passa a ser voltada para a formação de mão de obra, tendo em vista o mercado e é criado o Mobral. Ele tinha como objetivos erradicar o analfabetismo, incorporar os analfabetos na sociedade e desenvolver a alfabetização funcional (BELUZO[4]).

Já no período de redemocratização nos anos 80, com a promulgação da Constituição Federal em 1988, a EJA passa a ter políticas educacionais mais expressivas, passando a ser dever do Estado garantir o ensino fundamental, obrigatório e gratuito a todos aqueles que não tiveram acesso na idade apropriada(SOUZA[5]).

Na década de 90, o governo passa a transferir a responsabilidade da EJA também para os municípios e estimula as parcerias na esfera privada, entre empresas e Organizações Não Governamentais (ONGs), sendo criticada por alguns especialistas, vista como uma política com base na solidariedade (MACHADO [6]).

A EJA no novo século passa a ter uma visibilidade maior, porém ainda levanta muitos questionamentos referente aos esforços que o Estado deveria fazer para garantir os direitos e incentivos a retomada dos estudos desse grande contingente de pessoas que deveriam ser inseridas na sociedade.

## 1.2 A EJA no Nível Médio no Rio de Janeiro

A EJA, apesar de em suas determinações normativas, se referir ao direito a educação básica completa, somente em 2009, com a Lei N°12.060, o ensino médio passa a ser obrigatório no Brasil. Ela é destinada a jovens e adultos com 18 anos ou mais, que concluíram o ensino fundamental, porém não começaram ou concluíram o ensino médio e é responsabilidade da rede estadual de ensino, sendo ofertada predominantemente em cursos presenciais, oferecidos em colégios da rede estadual e nos cursos semipresenciais nos Centros de Educação de Jovens e Adultos (VENTURA[7]) . A estimativa é que exista somente no estado do Rio de Janeiro, em torno de dois milhões e duzentas mil pessoas dentro dessas condições.

## 1.3 Direito à Educação

Cabe ao Estado a responsabilidade de fomentar o interesse do retorno a sala de aula dessa parcela da população que não concluiu a educação básica, criando condições fa-

voráveis a permanência dos mesmo, conforme a lei de Diretrizes e Bases da Educação (LDB 9394/96), em seu artigo 37º § 1º. "Os sistemas de ensino assegurarão gratuitamente aos jovens e aos adultos, que não puderam efetuar os estudos na idade regular, oportunidades educacionais apropriadas, consideradas as características do alunado, seus interesses, condições de vida e de trabalho, mediante cursos e exames."

Mas será que o Estado vem fazendo sua parte ao longe dos últimos anos? A EJA está mais próxima e dentro da realidade dessa população que poderia se beneficiar de uma educação transformadora? O que vem sendo feito e quais são os frutos dessas ações? E elas têm sido suficientes para alcançar o público alvo? A intenção no decorrer desse trabalho é tentar responder alguns desses questionamentos, avaliando a tendência do número de matrículas da EJA no ensino médio , o alcance das políticas públicas através da taxa de atendimento e quais fatores podem interferir nela . E por fim, apontar a direção que essa modalidade tem tomado nesses últimos anos.

## 1.4 Organização do trabalho

Este trabalho de conclusão de curso se sustenta em cinco capítulos, além desse que introduz a história e importância da EJA como ação transformadora da sociedade e o dever do Estado de garantir o acesso a todos que a ela deseje. O segundo capítulo explana quais são os objetivos gerais e específicos desse trabalho. O terceiro capítulo demonstra os materiais e métodos utilizados. O capítulo quatro apresenta os resultados obtidos e no quinto, a conclusão .

## 2 Objetivos

### 2.1 Objetivo Geral

O objetivo geral deste trabalho é avaliar a tendência temporal e espacial da EJA no ensino médio no estado do Rio de Janeiro.

### 2.2 Objetivos Especificos

- Avaliar a tendência temporal do número de matrículas da EJA no período de 2007 a 2015, segundo modelos de regressão polinomial.
- Avaliar a dependência espacial do número de matrículas nos municípios do estado, em três períodos de tempo.
- Analisar o indicador da taxa de atendimento para EJA por município no ano censitário.
- Avaliar a associação entre a taxa de atendimento e fatores socioeconômicos e demográficos, segundo modelos de regressão.

## 3 Materiais e Métodos

### 3.1 Banco de Dados

Nesse trabalho foram utilizados os dados do Censo Escolar dos municípios do estado do Rio de Janeiro, para o período de 2007 a 2015. Os dados de todas as escolas públicas e privadas do estado são levantados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), junto às secretarias estaduais e municipais de educação. A coleta dos dados do Censo Escolar é feita através de preenchimento de um questionário disponível no sistema Educacenso, no endereço eletrônico [www.educacenso.inep.gov.br](http://www.educacenso.inep.gov.br), no qual o usuário de cada escola tem uma senha para acesso pessoal. Ele é realizado anualmente e tradicionalmente na última quarta-feira do mês de Maio. Do Censo Escolar foram utilizados o número de matrículas, turno e o número de escolas por ano e por município.

Além dos dados do INEP, também foram utilizados dados do Instituto Brasileiro de Geografia e Estatística (IBGE) referente ao censo de 2010, o qual fornece dados censitários e amostrais, coletados em todas as unidades domiciliares do estado do Rio de Janeiro através de uma entrevista de um recenseador, utilizando dois diferentes tipos de questionários: o Básico, que é aplicado em todas as unidades domiciliares que não foram selecionadas para a amostra; e o da Amostra, que é aplicado em todas as unidades domiciliares selecionadas para compor a amostra e possuem informações de características complementares dos seus moradores. Do Censo 2010, foram utilizadas as variáveis índice de desenvolvimento humano (IDH), produto interno bruto (PIB), taxa de desemprego, grau de instrução e população.

O banco de dados também inclui indicadores demográficos e socioeconômicos coletados ou construídos a partir dos dados do INEP e IBGE, tais como:

- Taxa de atendimento: é um indicador construído com base na razão entre o número de matrículas na modalidade EJA no ensino médio e o número de pessoas com



18 anos ou mais que possuem o ensino fundamental completo e o ensino médio incompleto.

$$TA = \frac{\text{N}^\circ \text{ de pessoas com matrículas na EJA no ensino médio}}{\text{N}^\circ \text{ de pessoas com mais de 17 anos com fundamental e o médio incompleto}} \times 100 \quad (3.1)$$

- Taxa de matrículas por triênios (2007-2009,2010-2012,2013-2015): é a razão entre o número total de matrículas no triênio dividido por três vezes a população total do ano central do triênio.
- Taxa de urbanização: é a percentagem da população urbana em relação à população total de uma dada região.
- Razão de dependência: mede a participação relativa do contingente populacional potencialmente inativo, que deveria ser sustentado pela parcela da população potencialmente produtiva.
- Esperança de vida aos 60 anos de idade: número médio de anos de vida esperados para uma pessoa ao completar 60 anos de idade, mantido o padrão de mortalidade existente na população residente, em determinado espaço geográfico, no ano considerado.
- Taxa de desemprego: é a razão entre o número de pessoas que estão procurando emprego e o número de pessoas economicamente ativas.
- Produto Interno Bruto (PIB) per capita: é a soma de todos os bens produzidos em um município em um determinado período dividido pela quantidade de habitantes desse município.
- Renda média domiciliar per capita: é a média das rendas domiciliares per capita das pessoas residentes em determinado espaço geográfico, no ano considerado.
- Índice de Gini da renda domiciliar per capita: mede o grau de desigualdade da distribuição de renda domiciliar per capita de uma determinada população em um determinado espaço geográfico. O índice de Gini é utilizado para medir a desigualdade de distribuição de renda em uma região e varia entre 0 e 1, quanto mais próximo de 1, maior é a desigualdade .
- Número de escolas : Número de escolas com ensino médio por 10 mil habitantes .

## 3.2 Modelos de Regressão

A fim de buscar um modelo estatístico que explique a relação entre a variabilidade do número de matrículas efetivadas na EJA e suas variáveis dependentes, foram utilizadas nesse trabalho as ferramentas estatísticas de análise de regressão. Ela é uma técnica na qual nos permite estudar a relação entre uma variável dependente e outras várias variáveis independentes. A equação que representa esse modelo com  $k$  variáveis explicativas é a seguinte:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + e_i, \quad i = 1, 2, \dots, n \quad (3.2)$$

Onde :

- $Y_i$  é a variável dependente, também conhecida como resposta, da  $i$ -ésima observação;
- $X_{ki}$  é a  $k$ -ésima variável explicativa da  $i$ -ésima observação;
- $\beta_0$  é o intercepto;
- $\beta_i$  é o coeficiente parcial de regressão da  $i$ -ésima observação;
- $e_i$  é o erro aleatório da  $i$ -ésima observação

No que diz respeito à interpretação do modelo, o intercepto  $\beta_0$  é o valor médio de  $Y$  quando as demais variáveis são iguais à zero. O coeficiente de regressão  $\beta_1$  mede a variação no valor médio de  $Y$ , por unidade de variação em  $x_1$ , mantendo-se os valores das demais variáveis explicativas constantes. Da mesma forma,  $\beta_k$  segue a mesma interpretação anterior. E por fim,  $e_i$  é o erro associado a cada observação do modelo(GUJARATI[8]).

Algumas suposições são necessárias para a utilização de forma correta do modelo de regressão linear clássico:

- Os erros têm distribuição normal, média zero e variância constante;
- Os erros são não correlacionados dois a dois;
- As variáveis explicativas  $X_1, X_2, \dots, X_k$  assumem valores fixos.

Em notação matricial, podemos escrever o modelo de regressão linear da seguinte forma:

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{e}, \quad (3.3)$$

onde:

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \underline{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

### 3.2.1 Estimadores dos Parâmetros do Modelo

Os estimadores dos parâmetros da regressão linear podem ser obtidos através do método de mínimos quadrados ordinários, encontrando os vetores de estimadores  $\hat{\underline{\beta}}$ , com componentes  $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n)$  que minimizam a soma dos quadrados dos resíduos (SQE). Note que podemos reescrever o SQE da seguinte forma matricial:

$$SQE = \sum_{i=1}^n e_i^2 = \underline{e}'\underline{e} \quad (3.4)$$

$$= (\underline{Y} - \underline{X}\underline{\beta})'(\underline{Y} - \underline{X}\underline{\beta}) \quad (3.5)$$

$$= \underline{Y}'\underline{Y} - \underline{Y}'\underline{X}\underline{\beta} - \underline{\beta}'\underline{X}'\underline{Y} + \underline{\beta}'\underline{X}'\underline{X}\underline{\beta} \quad (3.6)$$

$$SQE = \underline{Y}'\underline{Y} - 2\underline{\beta}'\underline{X}'\underline{Y} + \underline{\beta}'\underline{X}'\underline{X}\underline{\beta} \quad (3.7)$$

Então podemos derivar e igualar a zero o SQE para obter o ponto de mínimo para os

valores de  $\beta$ ,

$$\begin{aligned} \frac{\delta SSE}{\delta \underline{\beta}} &= -2\underline{X}'\underline{Y} + 2\underline{X}'\underline{X}\underline{\beta} = 0 \\ \iff (\underline{X}'\underline{X})\underline{\beta} &= \underline{X}'\underline{Y} \end{aligned} \quad (3.8)$$

Tomando a matriz inversa de  $\underline{X}'\underline{X}$ , temos

$$\iff \underline{\beta} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y}$$

e finalmente chegamos ao modelo de regressão linear ajustado

$$\hat{Y} = \underline{X}\hat{\beta}$$

tendo os vetores dos resíduos

$$\hat{e} = \underline{Y} - \hat{Y} \quad (3.9)$$

## 3.2.2 Inferência do Modelo

### 3.2.2.1 Significância individual dos parâmetros do modelo

A significância dos coeficientes estimados de regressão parcial é avaliada através do Teste t de Student, aonde verifica-se a importância individual da variável que está sendo incluída no modelo de regressão com  $k$  variáveis independentes e testa as seguintes hipóteses:

$$\begin{cases} h_0 : \beta_k = 0 \\ h_1 : \beta_k \neq 0 \end{cases}$$

Desta forma, a estatística do teste é dada por

$$t_0 = \frac{\hat{\beta}_k}{\sqrt{\hat{\sigma}^2 C_{kk}}} \sim t(n - p - 1)$$

onde onde  $C_{kk} = (X'X)^{-1}$ .

### 3.2.2.2 Significância global do modelo

Já o teste F consiste em testar a hipótese nula que todos os coeficientes de inclinação da equação de regressão múltipla são simultaneamente iguais a zero contra a hipótese que nem todos eles são, tendo  $H_0$ :

$$F_0 = \frac{\frac{SQR}{p}}{\frac{SQE}{(n-p-1)}} \sim F(P; n - p - 1)$$

### 3.2.3 Análise de Resíduos

A análise de resíduos serve para verificar a adequabilidade do modelo de regressão. Para isso, é fundamental analisar se os resíduos atendem aos pressupostos básicos feitos anteriormente, como independência, homoscedasticidade e distribuição normal. Ademais, quando se trata de regressão linear múltipla, também é necessário verificar se existe multicolinearidade entre as variáveis explicativas.

Os resíduos são definidos como a diferença entre os valores da variável resposta observada e a os valores da variável resposta prevista, isto é,

$$e_i = y_i - \hat{y}_i, \quad i = 1, 2, \dots, n.$$

#### 3.2.3.1 Diagnóstico de Normalidade

O Q-Q Plot é um gráfico muito útil para auxiliar na inspeção da normalidade dos resíduos através dos seus quantis. Esse método consiste em comparar os valores observados, com a curva teórica da distribuição normal representada por uma diagonal. Há evidências de que a distribuição dos erros segue uma distribuição normal, quando ela apresenta um bom ajuste a curva. Junto a essa análise gráfica, vamos utilizar de forma complementar o Teste de Kolmogorov-Smirnov para testar as hipóteses:

$$\begin{cases} h_0 : \text{Os dados seguem uma distribuição normal} \\ h_1 : \text{Os dados não seguem uma distribuição normal} \end{cases}$$

Se baseando na estatística de teste:

$$d = \max\{\max(|F(x_i) - F_i(x_i)|); \max(|(x_i - 1) - F_i(x_i)|)\} \quad (3.10)$$

O teste observa a máxima diferença absoluta entre a função de distribuição teórica acumulada, assumindo que ela seja normal e sua frequência dos valores amostrais. A ideia é que para cada valor de  $x$ ,  $F_0(x)$  esteja próximo de  $F(x)$ , isto é, sob  $H_0$ , espera-se que as diferenças entre  $F_0(x)$  e  $F(x)$  sejam pequenas.

### 3.2.3.2 Diagnóstico de Homoscedasticidade

Uma das hipóteses mais importante do modelo de regressão linear é a de que os erros devem ter variância constante, ou em outras palavras, que seja homoscedásticos. Uma forma de verificar essa condição é através da análise do gráfico dos resíduos versus valores ajustados, onde esperamos observar os pontos distribuídos de maneira aleatória em torno da reta  $y = 0$ , indicando assim que a variância dos resíduos é constante.

### 3.2.3.3 Diagnóstico de Independência

O teste de Durbin-Watson é usualmente utilizado para verificar a independência dos resíduos, no qual identifica a presença de autocorrelação entre os resíduos (RAWLINGS[9]). Ele se baseia nos valores da estatística de teste

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (3.11)$$

onde os  $e_i$  são os desvios da regressão ajustada pelo método de mínimos quadrados ordinários e  $d$  toma valores entre zero e quatro.

Ele testa as seguintes hipóteses:

$$\begin{cases} h_0 : \rho = 0 \\ h_1 : \rho \neq 0 \end{cases}$$

Regra de decisão:

- $h_1 : \rho > 0$

$$d < d_l \Rightarrow \text{Rejeitar } H_0$$

$d > d_U \Rightarrow$  Não rejeitar  $H_0$

$d_L \leq d \leq d_U \Rightarrow$  Teste é inconclusivo

- $h_1 : \rho < 0$

$4 - d < d_l \Rightarrow$  Rejeitar  $H_0$

$4 - d > d_U \Rightarrow$  Não rejeitar  $H_0$

$d_L \leq 4 - d \leq d_U \Rightarrow$  Teste é inconclusivo

Quando os resíduos apresentam autocorrelação, podemos aplicar a transformação de Prais-Winstem, para resolver esse problema, transformando a primeira observação como se segue:

$$Y_1\sqrt{1 - \rho^2} \text{ e } X_1\sqrt{1 - \rho^2}$$

### 3.2.3.4 Análise de Multicolinearidade

O termo multicolinearidade é utilizado para designar quando há duas variáveis independentes fortemente correlacionadas, o que pode levar a uma estimação imprecisa. Uma maneira de investigar a presença de Multicolinearidade é por meio do VIF (Variance Inflation Factor). Ele é calculado com base nos elementos da diagonal principal de  $(X'X)$ , onde podemos definir o  $j$ -ésimo elemento como:

$$VIF_{jj} = \frac{1}{1 - R_j^2} \quad (3.12)$$

em que  $R_j^2$  é o coeficiente de determinação da regressão de  $X_j$  sobre as outras variáveis explicativas.

Quando maior a correlação entre as variáveis, mais alto será o valor do VIF. Na literatura indica problema de multicolinearidade quando esse valor é maior que 10 (GUJARATI[8]).

## 3.2.4 Medidas de qualidade do ajuste

### 3.2.4.1 Coeficiente de Determinação ( $R^2$ )

O coeficiente de determinação é usado para medir a qualidade do ajuste de uma regressão e indica o percentual da variação total de  $Y$  explicada pelo modelo de re-

gressão(GUJARATI[8]), sendo calculado da seguinte forma:

$$R^2 = \frac{SQE}{SQT} * 100 \quad (3.13)$$

### 3.2.4.2 Coeficiente de Determinação Ajustado ( $R_{aj}^2$ )

Para comparar dois modelos de regressão com a mesma variável resposta é necessário levar em consideração o número de variáveis explicativas do modelo. Podemos fazer isso considerando uma forma alternativa do coeficiente de determinação, que tem como equação:

$$R_{aj}^2 = 1 - \frac{\frac{SQR}{(n-k)}}{\frac{SQT}{(n-1)}}$$

### 3.2.4.3 Critério de Informação de Akaike (AIC)

O AIC impõe uma medida corretiva pelo acréscimo de regressores, sendo mais dura do que o  $R^2$  e sua equação é a seguinte:

$$\ln AIC = \frac{2k}{n} + \ln \frac{SQR}{n}, \quad (3.14)$$

onde  $\ln AIC$  é o logaritmo natural de AIC e  $\frac{2k}{n}$  é o fator de correção.

Na comparação entre dois modelos ou mais, a escolha vai levar em consideração aquele que apresentar o menor valor do AIC.

### 3.2.5 P-valor

O p-valor é o nível mais baixo de significância no qual a hipótese nula pode ser rejeitada, sendo assim, quando ela for menor que o nível de significância adotado nesse trabalho, será rejeitada a hipótese nula.

### 3.2.6 Modelos de Regressão Polinomial

Nem sempre a regressão linear produz um ajuste adequado ao modelo estudado. Em algumas situações, a curva desse modelo pode apresentar um comportamento polinomial. Nesses casos, a melhor opção é utilizar o método de regressão polinomial, que é uma das



classes do modelo de regressão múltipla. A equação que descreve esse modelo é a seguinte:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}^2 + \dots + \beta_k X_{ki}^m + e_i \quad (3.15)$$

Os estimadores desse modelo também podem ser obtidos através do método dos mínimos quadrados visto anteriormente, onde vamos estimar  $m + 1$  coeficientes e tomar as derivadas parciais em função das  $m + 1$  estimativas e por fim, resolver um sistema de  $m+1$  equações lineares concomitantes (CHAPRA[10]).

Neste trabalho, os modelos de regressão polinomiais serão usados para estimar a tendência do número de matrículas no período de 2007 a 2015. Assim, X será a variável que representa o tempo. Em geral, a variável que representa o tempo (o ano da matrícula) é centrada no ponto médio da série ( $t - t_0$ , onde  $t_0 = [(t_{\max} + t_{\min})/2]$ ), para "se evitar a correlação entre os termos da equação de regressão" (LATORRE[11]).

### 3.2.7 Modelo Teórico

O município tem que dispor de meios e condições para atender a demanda, e entre alguns fatores que podem interferir na taxa de atendimento, segundo a literatura, são: o número de estabelecimentos de ensino que cada município possui, renda per capita, índice de Gini. Calcula-se que quanto maior o IDH do município, maior é a taxa de atendimento, e quanto menor a taxa de urbanização, menor seja a taxa de atendimento. A figura 1 ilustra o modelo teórico e, portanto, apresenta todas as variáveis que poderão ser consideradas no modelo de regressão linear.

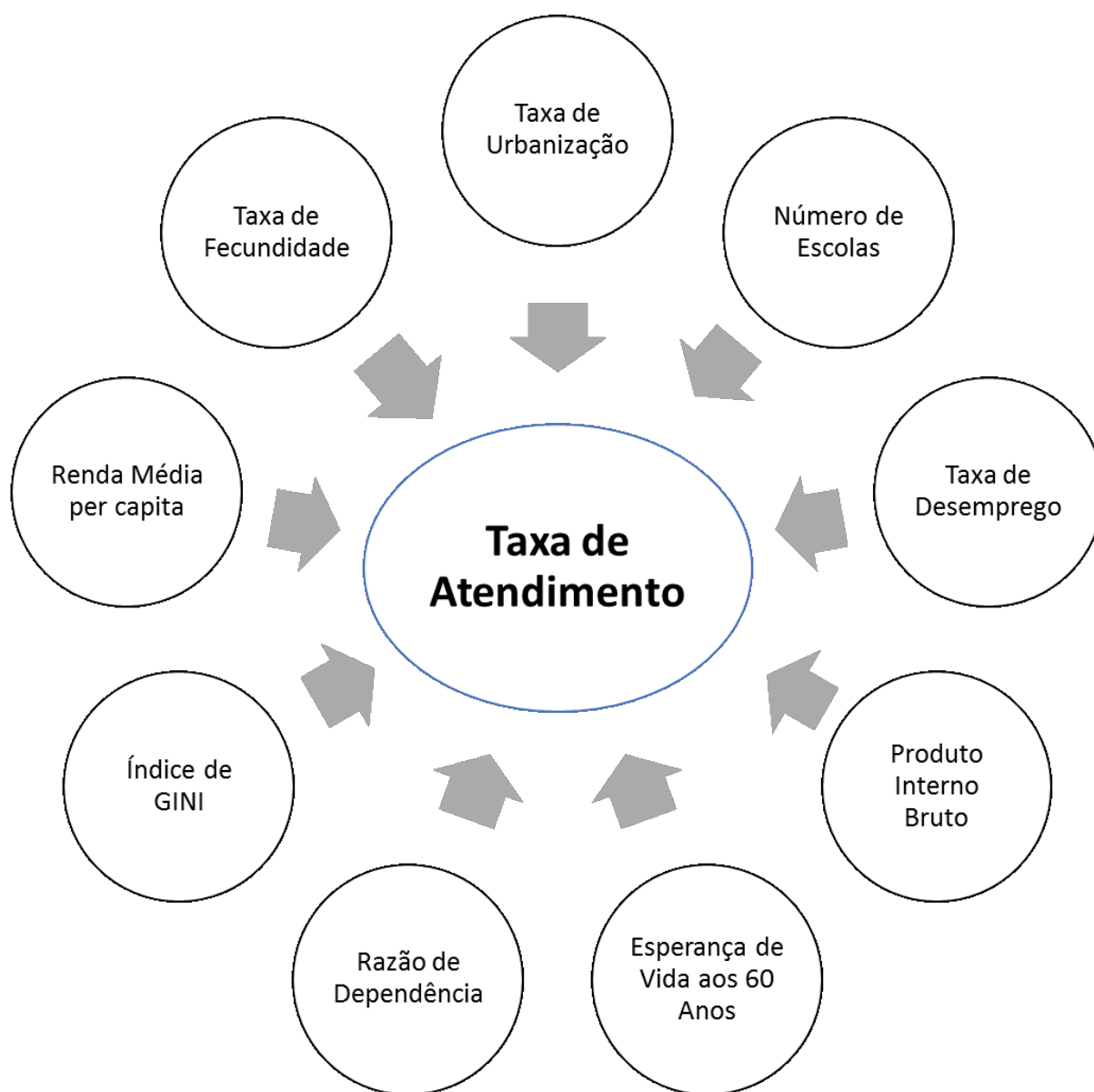


Figura 1: Diagrama com as variáveis que influenciam na taxa de atendimento

### 3.2.8 Método de Seleção das Variáveis

Nesse trabalho foi utilizado o método da eliminação retroativa. Inicialmente, começamos com todas as variáveis incorporadas no modelo e determinamos o p-valor do teste t para todas elas, caso todos os p-valores sejam menores que  $\alpha$ , não eliminamos nenhuma variável e encerramos o processo. Caso contrário, eliminamos aquela que apresenta o maior p-valor e voltamos a ajustar o modelo sem a variável eliminada. Esse processo é repetido até que todas as variáveis apresentem um p-valores menor que  $\alpha$  (HAIR[12]).

## 3.3 Análise Espacial

### 3.3.1 Análise Exploratória

A análise espacial serve como uma ferramenta para compreender as relações e características de um fenômeno, tendo como base a sua localização espacial. Os dados espaciais podem ser tradicionalmente classificados em três diferentes tipos: dados de padrão de pontos, dados espacialmente contínuos e dados de área.

Neste trabalho foi realizada a análise de dados espaciais de área. Dados dessa natureza têm a sua localização associadas a áreas que são delimitadas por polígonos, geralmente municípios, bairros ou setores censitários. Nelas não temos o conhecimento exato da localização do evento, e sim um valor agregado por área(CÂMARA [13]). O objetivo é investigar a existência de um padrão aleatório nos dados observados além de produzir mapas com melhores estimativas através de estruturas ocultas.

#### 3.3.1.1 Matriz de Proximidade Espacial

A maneira como os dados estão dispostos no espaço se constitui em um importante indicador da interação espacial, obtido através de medidas de associação ou autocorrelação espacial, permitindo assim, estimar a magnitude da dependência espacial. Há várias formas de se medir a autocorrelação a partir de diversos indicadores, que em comum, utilizam o conceito da matriz de proximidade espacial, também conhecida como matriz de vizinhança. Ela estima a variabilidade de dados de áreas, descrevendo o arranjo espacial dos objetos. Na matriz  $W(n \times n)$ , cada elemento  $w_{ij}$  representa uma medida de proximidade entre as cidades  $i$  e a cidade  $j$ . Ela pode ser calculada utilizando três critérios diferentes: em função do centróide das áreas, do comprimento das fronteiras ou da existência de um lado comum (MARQUES [14]).

### 3.3.1.2 Índice Global de Moran

O Índice global de Moran (I) é um dos indicadores de autocorrelação espacial mais empregados para verificar a dependência espacial e sua expressão considerando apenas o primeiro vizinho é a seguinte:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (Z_i - \bar{Z})(Z_j - \bar{Z})}{\sum_{i=1}^n (Z_i - \bar{Z})^2}, \quad (3.16)$$

onde:  $n$  representa o número de observações;  $Z_i$  o valor da Taxa de Matrícula na área  $i$ ;  $\bar{Z}$  o valor médio da Taxa de Matrícula na região de estudo e  $W_{ij}$  os elementos da matriz normalizada de proximidade espacial.

O Índice de Moran varia entre -1 e 1. Valores próximos de 0 indicam independência espacial, próximo de 1 indicam correlação espacial positiva, isto é, o valor da Taxa de Matrícula de uma cidade tende a ser semelhante aos valores vizinhos. Já valores próximos de -1 indicam uma correlação negativa.

Feito o cálculo, o próximo passo é estabelecer sua validade estatística, ou seja, verificar se de fato os valores medidos representam uma correlação espacial significativa. Para isso, associasse a estatística de teste à uma distribuição de probabilidade estatística, sendo mais frequente a distribuição normal. Uma alternativa, sem a necessidade de pressupostos em relação à distribuição é o teste de pseudo-significância. Nele são geradas distintas permutações dos valores associados às regiões: de maneira que em cada permutação é produzido um novo arranjo espacial, onde os valores estão redistribuídos entre as áreas. Se o valor observado inicialmente estiver localizado em umas das extremidades da distribuição simulada podemos concluir que o valor obtido é significamente estatístico [14].

### 3.3.1.3 Índice Local de Associação Espacial

O Índice Global de Moran apresenta um único valor como medida de associação espacial para todo o conjunto de municípios. Já o Índice Local de Associação Espacial (LISA – Local Indicators of Spatial Association), produz um valor específico para cada município, permitindo a identificação de agrupamentos de municípios com valores semelhantes e "outliers". O Índice Local de Associação Espacial Pode ser calculado da seguinte forma:

$$I_i = Z_i \frac{W z_i}{\sigma^2}, \quad (3.17)$$

onde:  $Z_i$  representa o valor do desvio do objeto  $i$ ;  $Wz_i$  o valor médio dos objetos vizinhos de  $i$  e  $\sigma^2$  é a variância da distribuição dos valores dos desvios.

#### 3.3.1.4 Diagrama de Espalhamento de Moran

Outra forma de visualizar a dependência espacial é através do gráfico de espalhamento de Moran. Para tal, devemos normalizar os valores da TA, subtraindo os valores dele por sua média e dividindo pelo desvio padrão, podendo assim analisar o comportamento da variabilidade espacial. O próximo passo é a construção de um gráfico bidimensional dos valores normalizados pela média dos vizinhos, sendo dividido em quatro quadrantes, com as seguintes interpretações:

- Q1 (valores positivos, médias positivas) e Q2 (valores negativos, médias negativas): indicam pontos de associação espacial positiva, no sentido que uma localização possui vizinhos com valores semelhantes.
- Q3 (valores positivos, médias negativas) e Q4 (valores negativos, médias positivas): indicam pontos de associação espacial negativa, no sentido que uma localização possui vizinhos com valores distintos.

#### 3.3.1.5 Box Map, Lisa Map e Moran Map

No Box Map, os municípios são classificados e coloridos levando em consideração a sua posição em relação aos quadrantes do gráfico de espalhamento. Para geração do Lisa Map, analisamos a significância dos valores do índice de Moran Local obtido para cada município, se baseando hipótese de não existência de autocorrelação espacial. Os municípios são classificados em quatro diferentes grupos: não significantes ( $p > 0,05$ ); com significância entre 0,05 e 0,01; com significância entre 0,01 e 0,001; e menor que 0,001. O Moran Map segue a mesma ideia do Lisa map, porém apenas os municípios que obtiveram valores significantes no LISA são destacados, aparecendo classificados em quadro grupos, de acordo com sua localização no quadrante do gráfico de espalhamento e o restante dos municípios são classificados como sem significância.

## 4 Análise dos Resultados

Neste trabalho todas as análises foram feitas no programa R version-3.3.2 e RStudio Version-1.0.44. Os gráficos foram gerados utilizando o pacote "ggplot2" do R. Foi considerado o nível de significância de 5 % em todos os testes realizados.

### 4.1 Análise de Tendência Utilizando Modelos de Regressão Polinomial

Foi realizado a análise de tendência do período compreendido entre 2007 e 2015 utilizando modelos de regressão polinomial para séries históricas, sendo considerado o número de matrículas totais da EJA no ensino médio como variável dependente ( $Y$ ), e os anos como variável independente ( $X$ ). A fim de evitar a autocorrelação entre os termos da equação de regressão, foi aplicada a transformação da variável "ano" em ano-centralizado (ano menos o ponto médio do período de estudo), sendo o ano de 2011 utilizado como ponto médio.

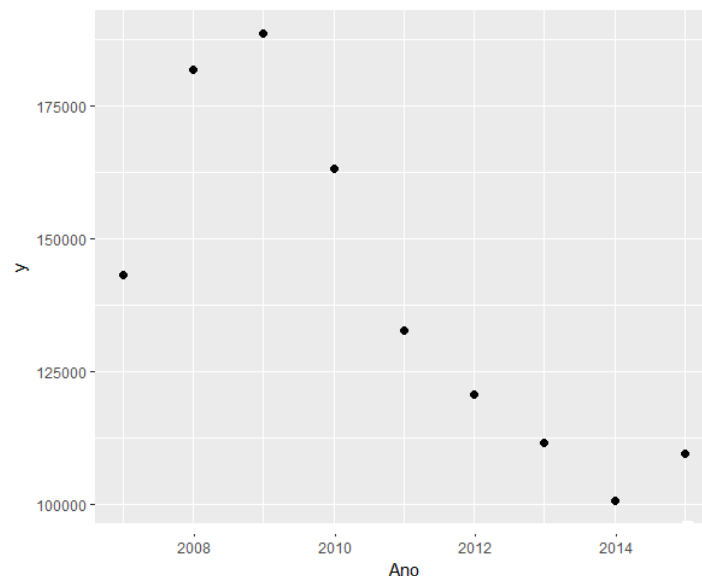


Figura 2: Número de matrículas em cada ano centralizado.

Na figura 2 observa-se o diagrama de dispersão que expressa a relação entre o número de matrículas totais e os anos de estudo, auxiliando visualmente a escolha dos candidatos a melhor curva para representar essa relação. Nele nota-se um aparente comportamento polinomial com uma tendência declinante iniciada a partir do ano de 2010. Para chegar na função que descreve a relação entre o número de matrículas totais e os anos, foi realizada a modelagem pela técnica de regressão, considerando o modelo linear simples até o modelo quártico. Os resultados obtidos podem ser vistos na tabela 1.

Tabela 1: Estimativas dos parâmetros dos modelos polinomiais

Modelo	Linear		Quadrático		Cúbico		Quártico	
Estimador	Estimativa	p-valor	Estimativa	p-valor	Estimativa	p-valor	Estimativa	p-valor
Intercepto	138987	0,0000	144950,1	0,0000	144950,1	0,0000	138028,25	0,0000
$\hat{\beta}_1$	-9581	0,0079	-9581,1	0,0120	-24304,0	0,0005	-24257,09	0,0002
$\hat{\beta}_2$	-	-	-894,4	0,4810	-894,4	0,1386	2743,99	0,0817
$\hat{\beta}_3$	-	-	-	-	1247,7	0,0032	1244,35	0,0009
$\hat{\beta}_4$	-	-	-	-	-	-	-220,97	0,0338

Analisando e comparando os quatros diferentes modelos, tem-se que apenas o modelo linear apresentou um p-valor significativo para todos os seus estimadores. Já o modelo quártico apresentou valores significativos para três dos seus estimadores e uma valor bem próximo para o estimador  $\hat{\beta}_2$ . A tabela 2 apresenta os resultados para o Teste F, Teste de Durbin-Watson (DW), o qual foi feito utilizando o pacote "lmtest" do R, além do coeficiente de determinação ajustado ( $R_{aj}^2$ ).

Tabela 2: P-valor dos testes de qualidade de ajuste dos modelos estimados

Modelo	DW (p-valor)	F (p-valor)	$R_{aj}^2$
Linear	0,0411	0,0079	0,6091
Quadrático	0,0049	0,0305	0,5832
Cúbico	0,0004	0,0009	0,9240
Quártico	0,1865	0,0005	0,9730

O modelo que apresentou melhor ajuste foi o quártico, porém tendo um estimador que não é significativo. Nos demais modelos, o pressuposto de autocorrelação não foi atendido, tendo o teste de Durbin-Watson apresentado um valor não significativo. Na literatura [8], recomendasse a correção de Prais-Winsten, quando temos um número pequeno de amostras e tratamento do problema envolvendo a autocorrelação. A tabela 3, apresenta

os modelos polinomiais com a correção de Prais-Winsten utilizando o pacote "prais" do R.

Tabela 3: Estimativas dos parâmetros dos modelos com a correção de Prais-Winsten

Modelo	Linear		Quadrático		Cúbico		Quártico	
	Estimativa	p-valor	Estimativa	p-valor	Estimativa	p-valor	Estimativa	p-valor
Intercepto	138142	0,0000	146024	0,0000	145449,6	0,0000	136789,81	0,0000
$\hat{\beta}_1$	-8588	0,0251	-8014	0,0557	-24395,8	0,0007	-23909,24	0,0000
$\hat{\beta}_2$	-	-	-1097	0,4447	-993,8	0,1275	3602,89	0,0123
$\hat{\beta}_3$	-	-	-	-	1256,6	0,0035	1206,02	0,0002
$\hat{\beta}_4$	-	-	-	-	-	-	-283,52	0,0068

Depois de aplicada a correção de Prais-Winsten, observou-se que os modelos linear e quártico obtiveram um p-valor significativo. Na tabela 4, observa-se que todos os modelos foram significativos para o Teste F e o modelo quártico foi o que apresentou o melhor ajuste, com o coeficiente de determinação,  $R^2 = 99,98\%$ , e o coeficiente de determinação ajustado,  $R_{aj}^2 = 99,97\%$ , indicando que uma parte significativa da variabilidade da variável "Número de Matrículas" foi explicada pelo modelo.

Tabela 4: P-valor dos testes de qualidade de ajuste dos modelos com correção de Prais-Winsten estimados

Modelo	F (p-valor)	$R_{aj}^2$
Linear	0,0000	0,9688
Quadrático	0,0000	0,9570
Cúbico	0,0000	0,9945
Quártico	0,0000	0,9997

A equação que representa o modelo escolhido é a seguinte:

$$\hat{Y}_i = 136789,81 - 23909,24X + 3602,89X^2 + 1206,02X^3 - 283,52X^4 \quad (4.1)$$

E a curva de tendência do modelo final, pode ser visto na figura 3, onde observa-se que os valores estimados para os números de matrícula, estão bem próximos dos valores reais.



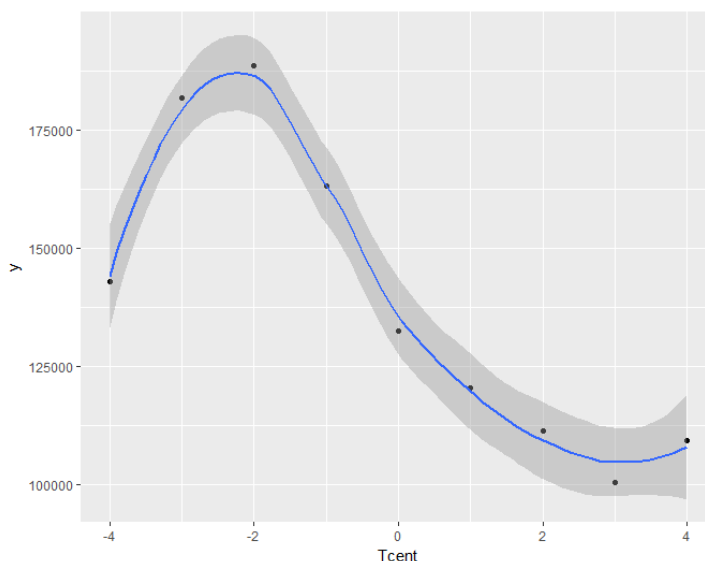


Figura 3: Número de matrículas em cada ano centralizado, ajustado pelo modelo quártico com correção de Prais-Winsten

## 4.2 Análise Exploratória de Dados Espaciais

A distribuição do número de matrículas nos municípios do estado do Rio de Janeiro pode apresentar algum tipo de associação entre municípios vizinhos, em diferentes escalas espaciais. Através da análise espacial exploratória, foi analisada se há a existência da autocorrelação espacial, onde os valores das características das localidades geograficamente próximas com valores parecidos indicariam uma autocorrelação positiva e, analogamente, os menos similares, uma autocorrelação negativa, para pares aleatórios de observações. Foram utilizados os pacotes "sp", "maptools", "spdep", "rgdal" e "ggmap" do R para a análise espacial e o pacote "spacemaker" para o AIC.

A figura 4 mostra o número de matrículas na EJA no ensino médio por cada mil habitantes classificado segundo seus quantis, entre os anos de 2007 e 2015, divididos em 3 diferentes triênios. No 1º triênio podemos observar que a região metropolitana possuía uma das menores taxas de matrículas do estado do Rio de Janeiro e grande parte dos municípios do Noroeste Fluminense estavam acima da mediana. Já entre os anos de 2010 e 2012, percebemos uma drástica diminuição de matrículas nos municípios que possuíam uma taxa superior a 19 matrículas por mil habitantes. E no último triênio, essa redução foi generalizada, com quase todos municípios apresentando uma taxa inferior a 12 matrículas por mil habitantes.

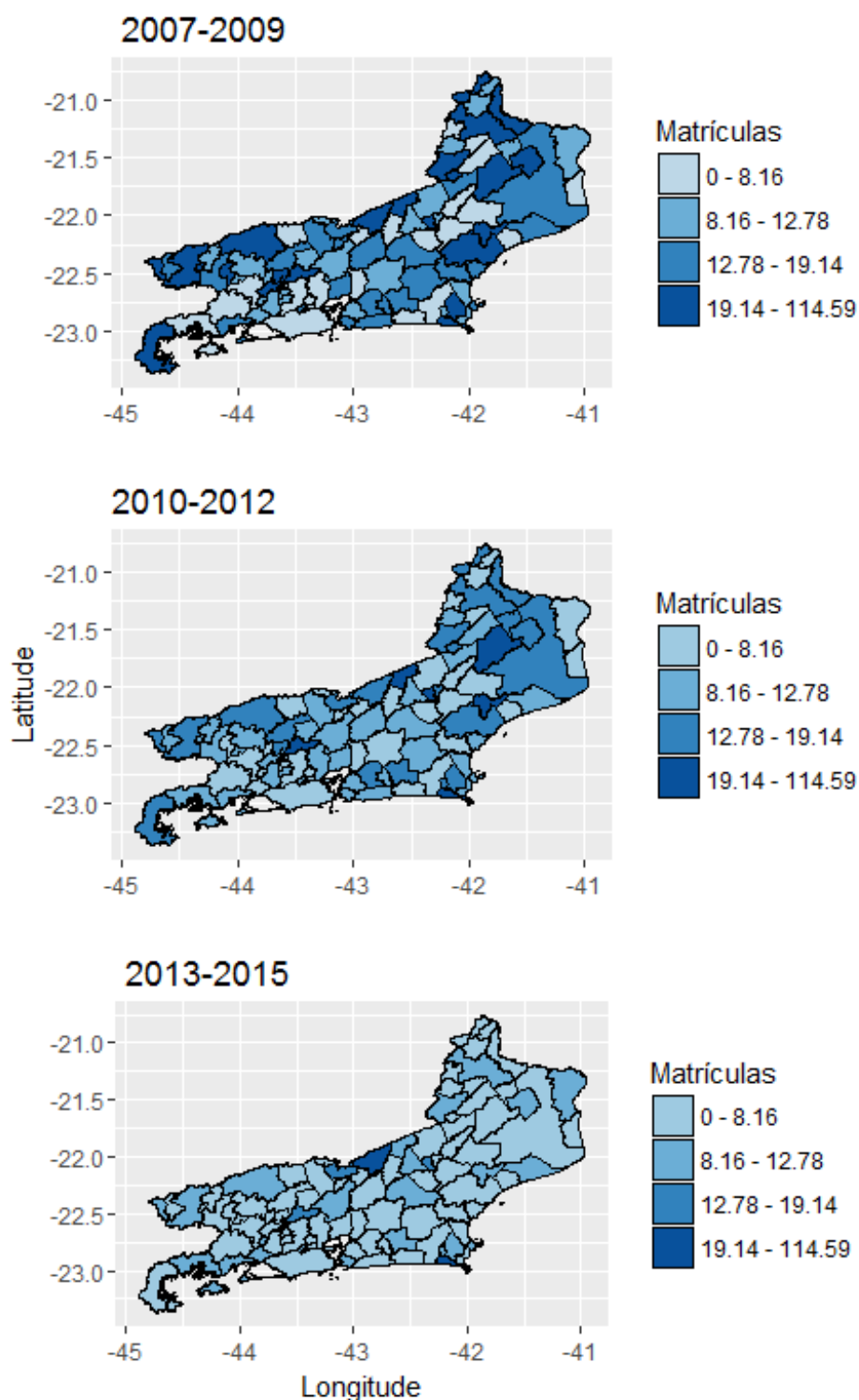


Figura 4: Mapa do Estado do Rio de Janeiro segundo o número de matrículas (por 1000 hab.) da EJA no período de 2007 a 2009.

A partir dos dados apresentados nos mapas da figura 4, foram geradas matrizes de vizinhanças (W): contígua e de distância geográficas variando o número de vizinhos, indo de um até seis. A matriz de vizinhança para 5 vizinhos foi a que apresentou o menor valor

de AIC, sendo assim a escolhida. A figura 5 mostra as conexões entre os municípios do estado do Rio de Janeiro levando em consideração os 5 vizinhos mais próximos.

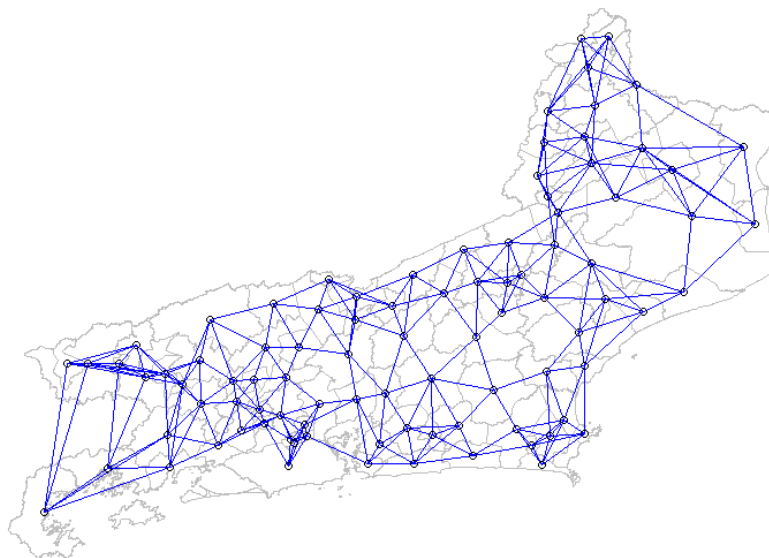


Figura 5: Mapa do Rio de Janeiro com as conexões entre os municípios considerando os 5 vizinhos mais próximos

Na tabela 5, é apresentado o Índice Global de Moran para os 3 períodos, e o nível de significância, segundo o critério de 5 vizinhos mais próximos. Nela observa-se que os valores do Índice Global de Moran foram próximos a zero em todos os casos, com o p-valor não significativo. Conclui-se que não existe uma autocorrelação espacial, seja qualquer um o período observado.

Tabela 5: Índice Global de Moran e o nível descritivo

Período	Índice de Moran	p-valor
1º triênio	-0,012631166	0,5126
2º triênio	-0,038212038	0,6630
3º triênio	-0,050012176	0,1787

O Índice de Associação Local (LISA) auxiliou na visualização de áreas de "clusters"

e de zonas de transição, através dos valores do Gráfico de Espalhamento de Moran, representados no Box Map. Na figura 6, nota-se que a região metropolitana no período de 2007 a 2009 concentra a maior parte dos seus municípios no quadrante Q2, indicando uma correlação positiva entre os municípios vizinhos. Os municípios localizados no quadrante Q3, que são aqueles que possuem valores positivos com médias negativas, e no quadrante Q4, que possuem valores negativos com médias positivas, podem ser interpretados como aqueles que se encontram mais afastados da reta de regressão linear e indica que esses municípios não seguem o mesmo processo de dependência espacial dos outros.

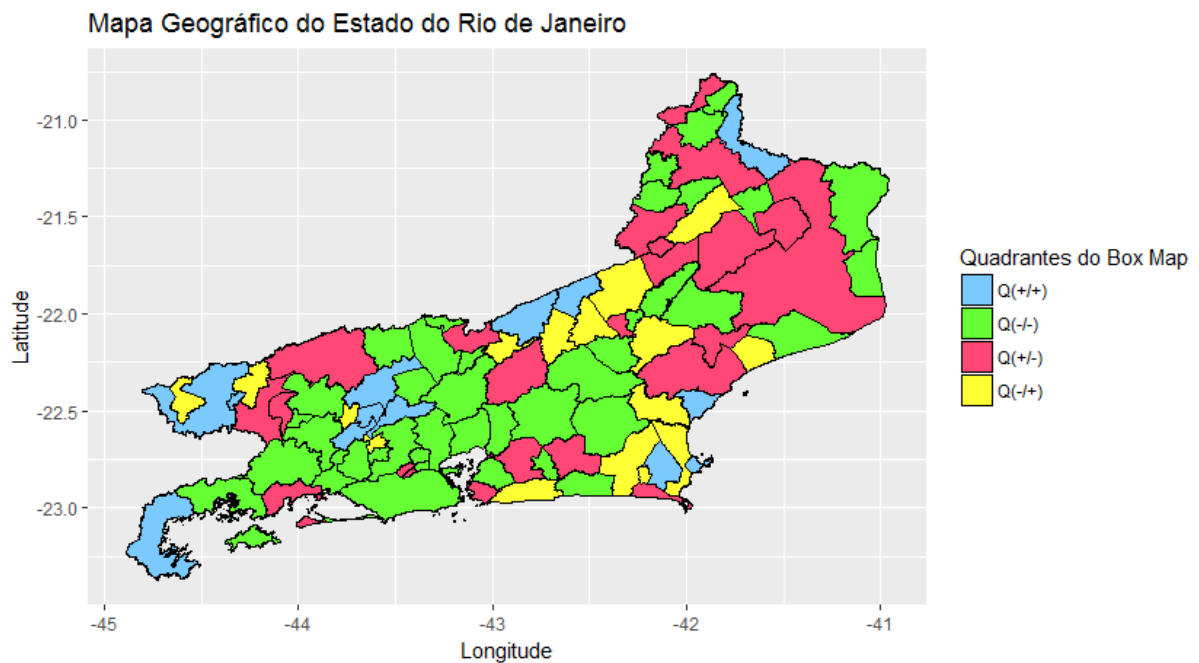


Figura 6: Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 1º Triênio, 2007-2009

Na figura 7, observa-se uma concentração de municípios com alta correlação espacial nas regiões Norte e Noroeste Fluminense. O município do Rio de Janeiro continua com uma correlação positiva com seus vizinhos.

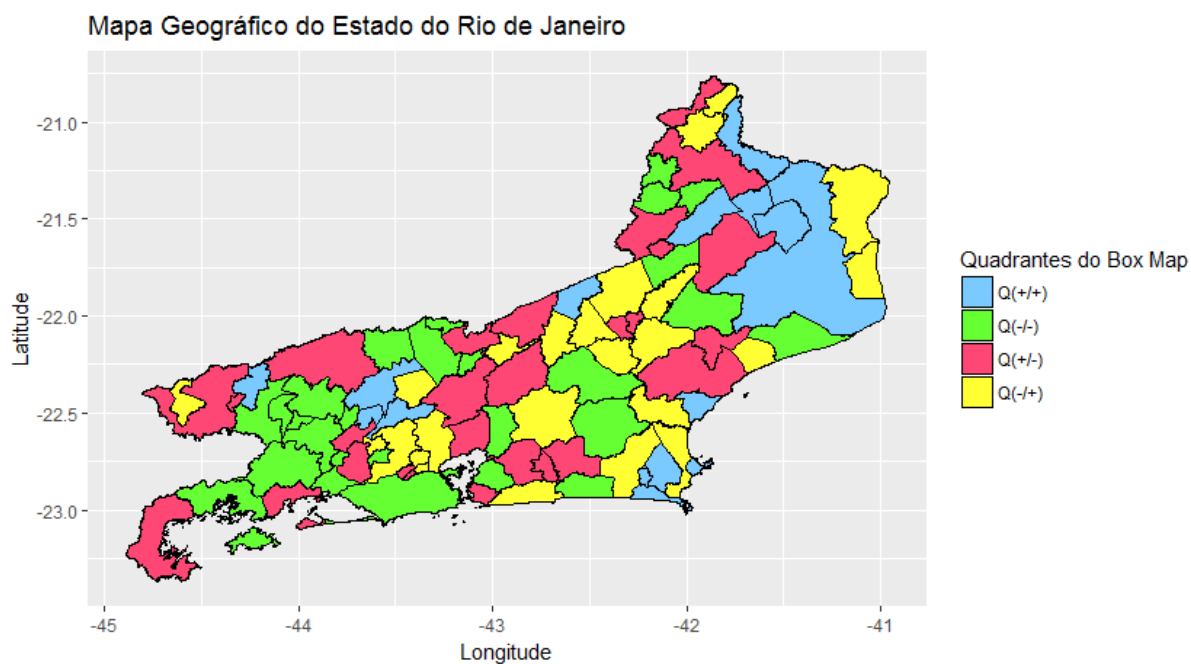


Figura 7: Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 2º Triênio, 2010-2012

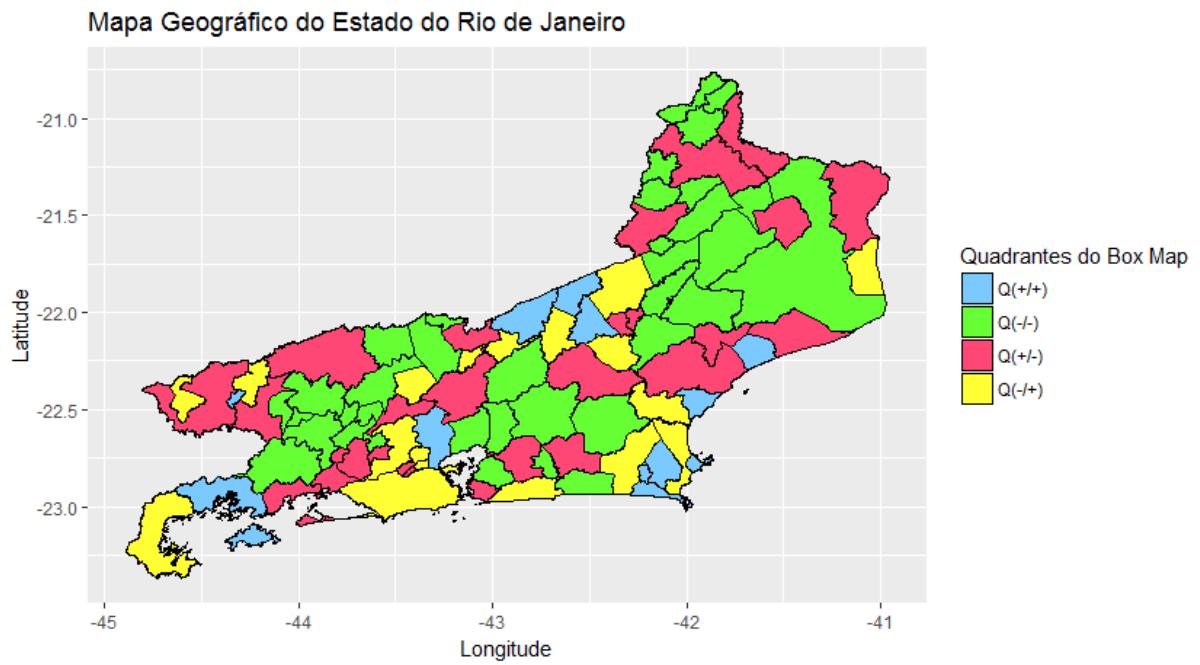


Figura 8: Box-map dos municípios, classificados de acordo com sua posição no gráfico de espalhamento de Moran no 3º Triênio, 2013-2015

Na figura 9 encontram-se os mapas "Moran Map", onde cada município foi colorido de acordo com sua posição no Gráfico de Espalhamento de Moran juntamente com seu nível de significância. No primeiro triênio verifica-se que os municípios de Carmo e Iguaba Grande estavam localizados no quadrante Q4, voltando a aparecer nos mesmo quadrante no 3º triênio. No 2º triênio, nenhum dos municípios apresentou um valor significativo e no 3º triênio, além dos municípios já citados, temos a cidade de Arraial do Cabo, localizada no quadrante Q4 e o município de São Sebastião do Alto no quadrante Q1, indicando uma correlação positiva.



Figura 9: Moran Map

### 4.3 Indicador de Taxa de atendimento para EJA por Município

A figura 10 mostra as taxas de atendimento para EJA no estado do Rio de Janeiro. De forma geral, as tonalidades de cor mais escura correspondem as melhores taxas de cobertura e, as mais claras, as taxas não satisfatórias de atendimento à EJA de ensino médio. A taxa de atendimento encontra-se superior a 30% nos municípios de Arraial do Cabo e Conceição do Macabu. No primeiro, ela corresponde a 92% e no segundo a 34%. Por outro lado, os municípios de Carapebus, Comendador Levy Gasparian, Laje do Muriaé, Rio Claro, Rio das Flores, São José de Ubá, São João da Barra, São Sebastião do Alto e Trajano de Moraes apresentam taxas de atendimento igual a zero para população com Ensino Fundamental completo e Ensino Médio incompleto. Um total de 40 municípios apresenta taxa de atendimento de até 10%. Desses, 11 estão no patamar igual ou inferior a 5%, com 4 municípios na Região Serrana, 2 na Região Noroeste Fluminense, 2 na Região Norte Fluminense e um na Região das Baixadas Litorâneas.

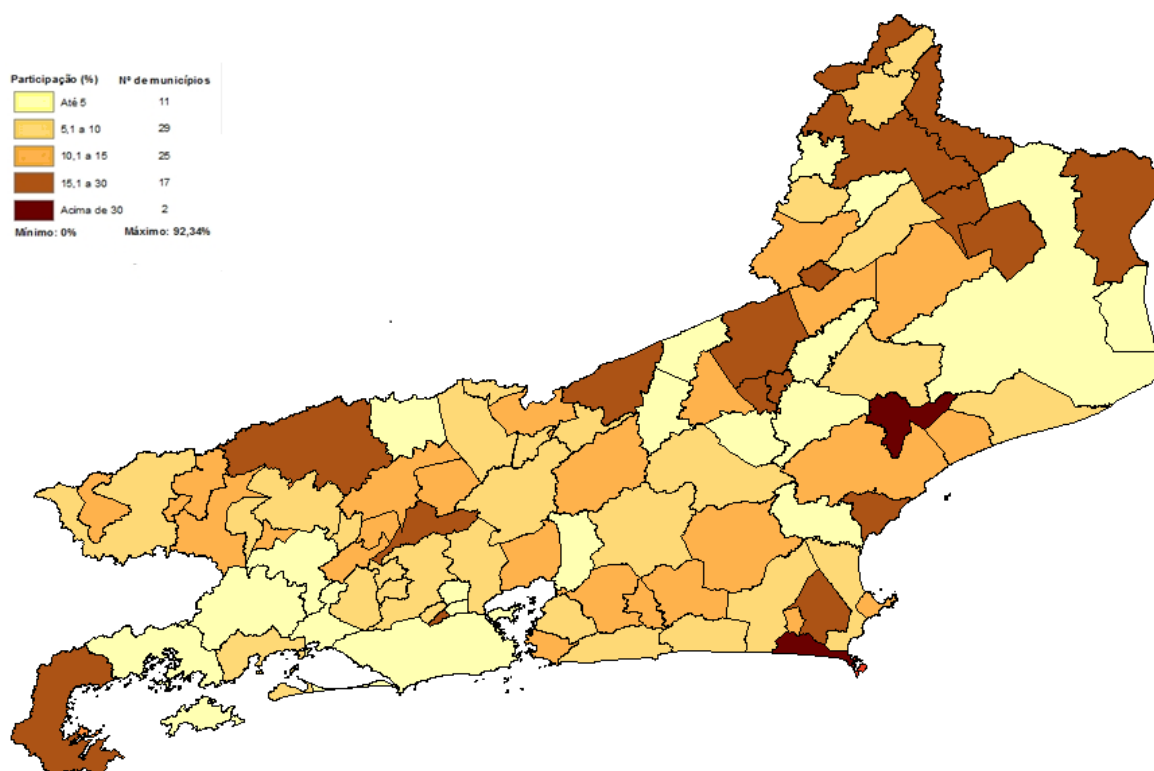


Figura 10: Taxa de Atendimento nos Municípios do Estado do Rio de Janeiro em 2010



## 4.4 Análise de Regressão

Nesse estudo, buscamos saber quais foram os fatores que contribuíram para a variabilidade da taxa de atendimento nos municípios que pertencem ao estado do Rio de Janeiro, para isso a variável resposta foi definida sendo a "Taxa de Atendimento" do  $i$ -ésimo município. Na tabela 6 têm-se as variáveis resposta e explicativas do modelo teórico proposto, juntamente com suas medidas descritivas. Na tabela 7, tem-se a legenda para as respectivas variáveis.

Tabela 6: Estatísticas descritivas das variáveis resposta e explicativas-RJ-2010

	<b>V1</b>	<b>V2</b>	<b>V3</b>	<b>V4</b>	<b>V5</b>	<b>V6</b>	<b>V7</b>	<b>V8</b>	<b>V9</b>	<b>V10</b>
Mínimo	0,000	0,423	1,490	354,000	43,900	36,500	8.732,00	0,651	77,590	1,300
1° <i>quartil</i>	0,058	0,478	6,910	543,700	51,450	75,650	10.602,00	11,339	79,610	1,680
Mediana	0,094	0,507	8,400	604,400	53,100	87,750	12.985,00	16,663	80,380	1,820
Média	0,113	0,506	8,360	644,200	53,430	84,340	24.449,00	17,805	80,650	1,850
Desvio Padrão	0,111	0,040	2,451	210,531	3,324	15,240	35.530,35	0,816	1,601	0,249
3° <i>quartil</i>	0,135	0,531	10,210	701,600	55,750	97,030	23.306,00	21,926	81,800	2,020
Máximo	0,923	0,639	13,480	1951,100	61,000	100,000	293.750,00	40,581	84,390	2,500

Tabela 7: Legenda das variáveis resposta e explicativas

<b>V1</b>	Taxa de Atendimento
<b>V2</b>	Índice de GINI
<b>V3</b>	Taxa de Desemprego
<b>V4</b>	Renda Média per capita
<b>V5</b>	Razão de Dependência
<b>V6</b>	Taxa de Urbanização
<b>V7</b>	PIB per capita
<b>V8</b>	Número de Escolas
<b>V9</b>	Esperança de Vida (>60 anos)
<b>V10</b>	Taxa de Fecundidade

A análise de Multicolinearidade feita através dos fatores de inflacionamento das variâncias (VIF), no qual não foram detectados valores acima de quatro, indicando que as variáveis explicativas do modelo proposto não estão fortemente correlacionadas. Foi utilizado o método da eliminação retroativa, e a partir deste método, foram excluídas as variáveis não significativas "Índice de GINI", "Taxa de Desemprego", "Renda Média per capita", "Razão de Dependência", "PIB per capita", "Esperança de Vida aos 60 anos" e "Taxa de

Fecundidade” do modelo final, além do intercepto que apresentou-se como não significativo. Na tabela 8 pode-se visualizar as estimativas do modelo sem intercepto.

Tabela 8: Estimativas dos parâmetros do modelo sem intercepto

Estimador	Estimativa	p-valor
$\hat{\beta}_0$	-	-
$\hat{\beta}_1$	0,0008392	0,00274
$\hat{\beta}_2$	0,0245955	0,04207

O próximo passo foi a validação do modelo através da verificação dos pressupostos necessários apresentados no capítulo de metodologia. Na figura 11, temos o histograma dos resíduos do modelo, onde podemos perceber que a curva teórica da distribuição normal não se encaixa perfeitamente sobre o histograma gerado, dando indícios de que os resíduos não possuem normalidade, sendo realizado posteriormente o teste de Kolmogorov no qual rejeitou-se a hipótese nula.

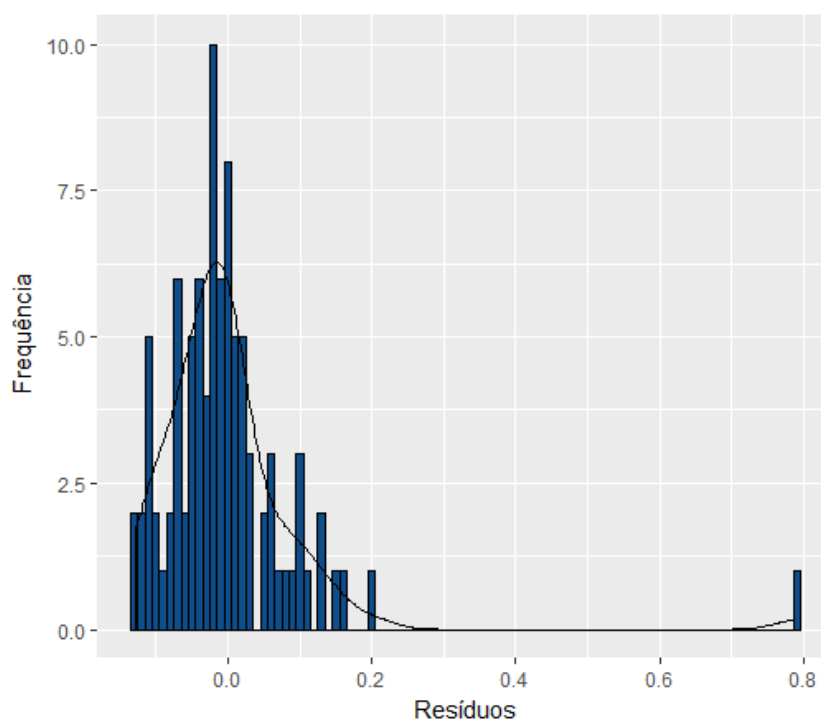


Figura 11: Histograma dos resíduos (modelo com "outliers")

Na figura 12, temos o boxplot dos resíduos, onde podemos notar um valor extremo ("outliers") dentre as observações, o qual poderia estar influenciando os estimadores do nosso modelo e igualmente trazendo a violação de normalidade dos resíduos.

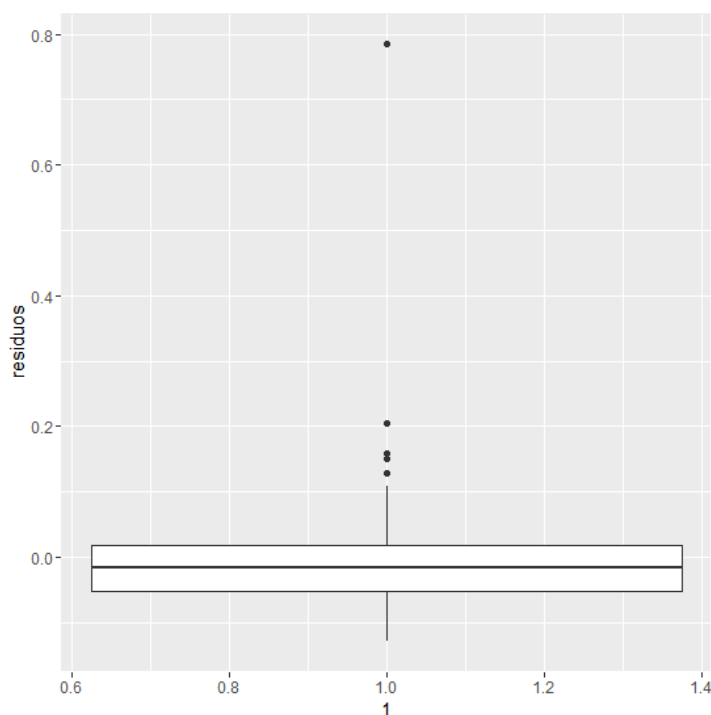


Figura 12: Boxplot dos resíduos (modelo com "outliers")

Então, o modelo escolhido, sem o intercepto seria:

$$\hat{Y}_i = 0,0008392V6 + 0,0245955V8 \quad (4.2)$$

A escolha do modelo sem intercepto faz sentido, uma vez que, se a "Taxa de Urbanização"(V6) e o "Número de Escolas"(V8) forem nulos, o valor esperado da "Taxa de Atendimento" será igual a 0. Mantendo a variável "Número de Escolas" constante, o efeito parcial da "Taxa de Urbanização" no valor médio estimado da "Taxa de Atendimento" será igual a 0,0008392. Ao passo que mantendo a variável "Taxa de Urbanização" constante, o efeito parcial do "Número de Escolas" no valor médio estimado da "Taxa de Atendimento" será igual a 0,0245955.

Com a preocupação de não retirar uma observação importante do banco de dados, foi minuciosamente analisado o valor obtido dessa observação e chegou-se a conclusão que seria muito improvável um município chegar a 92% da taxa de atendimento, sendo, portanto, retirado do banco de dados o município de Arraial do Cabo.

Novamente foi feito a seleção das variáveis explicativas do modelo pelo método da eliminação retroativa, apresentando as mesmas variáveis do modelo anterior como significativa e o intercepto não significativo. O teste F avalia a hipótese de que nenhuma das

variáveis explicativas tem qualquer relação com a variável resposta. Sob esta hipótese, observou-se um p-valor inferior a 0,001, onde podemos concluir que pelo menos uma das variáveis explicativas está relacionada com a variável resposta. Na tabela 9 podemos visualizar as estimativas do modelo sem intercepto .

Tabela 9: Estimativas dos parâmetros do modelo sem "outliers" e sem intercepto

Estimador	Estimativa	p-valor
$\hat{\beta}_0$	-	-
$\hat{\beta}_1$	0,0007531	0,0000
$\hat{\beta}_2$	0,0229912	0,00333

Outra vez, seguimos com as verificações necessárias para a validação do modelo. Primeiramente foi verificado o histograma dos resíduos através da figura 13, o qual a curva teórica da distribuição normal já se mostrou melhor ajustada do que o modelo anterior, indicando que com a retirada do "outlier", a distribuição dos resíduos já se aproximava de uma curva normal. Foi feito então o teste de Kolmogorov, o que obteve um p-valor de 0,22, não rejeitando a hipótese nula, portanto os resíduos são normalmente distribuídos, tendo o primeiro pressuposto atendido.

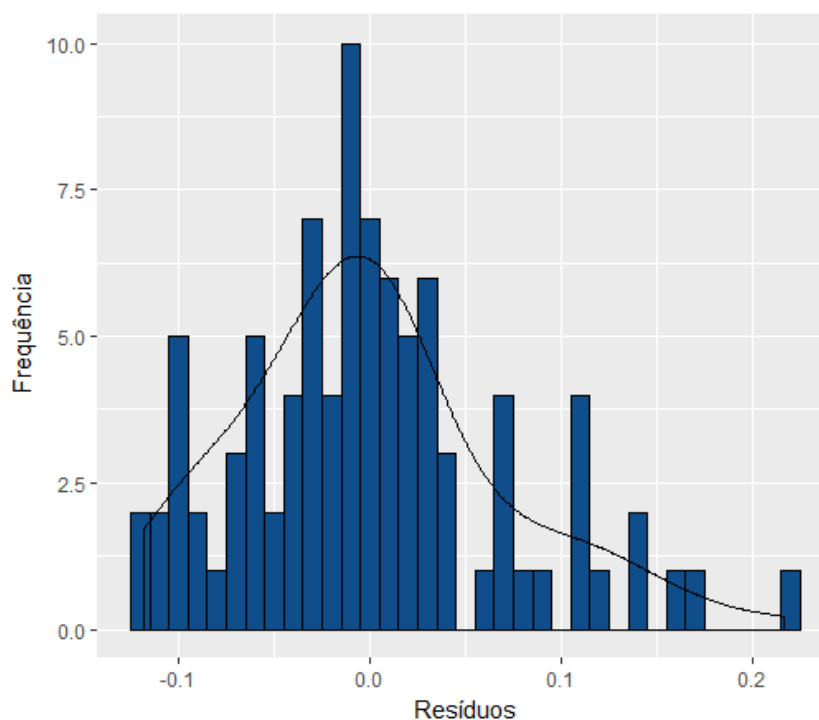


Figura 13: Histograma dos resíduos (modelo sem "outliers")

O próximo pressuposto analisado foi o de não correlação entre os resíduos do modelo,

ou seja, a não existência de autocorrelação do modelo. Para isso utilizou-se o teste de Durbin-Watson (DW), o qual apresentou um p-valor de 0,68, não rejeitando a hipótese nula, logo o pressuposto de não autocorrelação também foi atendido nesse modelo. O último pressuposto verificado foi em relação à homocedasticidade dos resíduos. Podemos ver através do gráfico de "resíduos versus valores ajustados" presente na figura 14 que os pontos estão distribuídos de forma aleatória, dando indícios de que os resíduos não apresentam heterocedasticidade.

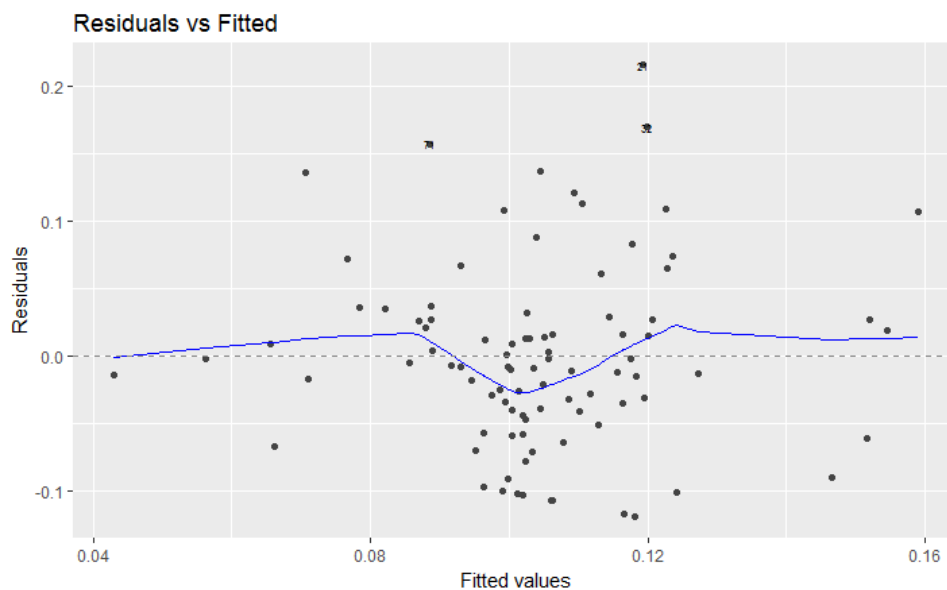


Figura 14: Gráfico resíduos versus valores ajustados (modelo sem "outliers")

Com isso, todos os pressupostos do modelo foram atendidos e temos a equação do modelo final estimado sendo a seguinte:

$$\hat{Y}_i = 0,0007531V6 + 0,0229912V8 \quad (4.3)$$

Interpretando os coeficientes do modelo de regressão linear múltipla, temos que, mantendo a variável "Número de Escolas" constante, então o efeito parcial da "Taxa de Urbanização" (V6) no valor médio estimado da "Taxa de Atendimento" será igual a 0,0007531. E mantendo a variável "Taxa de Urbanização" constante, então o efeito parcial do "Número de Escolas" (V8) no valor médio estimado da "Taxa de Atendimento" será igual a 0,0229912.

Os modelos selecionados mostram que a "Taxa de Atendimento" e a "Taxa de Urbanização" e o "Número de Escolas" estão relacionados de forma positiva.

## 5 Conclusão

Com mais de dois milhões de jovens e adultos no estado do Rio de Janeiro que não tiveram oportunidade de concluir seu ensino médio na idade recomendada, as políticas públicas voltadas para EJA são de fundamental importância. Esse trabalho permitiu avaliar o alcance que essa modalidade de ensino vem tendo nos últimos anos no estado e seus rumos, sinalizando para o poder público e sociedade civil o quadro preocupante que a EJA atravessa.

Através da análise de tendência, utilizando as ferramentas de regressão polinomial, conclui-se que ao longo da série histórica, as matrículas na EJA vieram apresentando uma tendência de queda, iniciada a partir do ano de 2010 e com um aparente indicativo que ela poderia ser interrompida em 2015. Porém, cabe destacar que nesse ano houve uma metodologia diferente para contabilizar o número de matrículas, no qual foram contabilizadas as matrículas da Educação à distância (EAD), tendo uma forte influência no aumento de matrículas em 2015.

O indicador da taxa de atendimento mostrou que os municípios atenderam muito aquém da demanda potencial, com muitos deles chegando a absurdos 0% de atendimento dentro da modalidade no ensino médio. Uma taxa de atendimento alta poderia contribuir na queda do número de matrículas nos últimos anos, porém, como nota-se, provavelmente não foi essa a causa da vertiginosa queda nos números de matrículas acompanhada através da análise de tendência.

Quanto à análise espacial, o município de Angra dos Reis iniciou 2007 e 2008 tendo em torno de 950 matrículas, já no ano seguinte ela deu um salto para 1853 matrículas e voltou a perder 1/3 das suas matrículas em 2010 e logo nos anos seguintes retornou ao seu patamar de 1500 matrículas. Caso como esse aonde tem uma queda acentuada e uma recuperação nos anos seguintes também podem ser observadas nos municípios de Barra Mansa em 2013; Bom Jardim em 2010; Itaguaí em 2009; Mesquita em 2013 e 2014; entre outros. Podemos notar esse movimento também no sentido contrário, onde em um ano

tem-se uma repentina alta de matrículas e logo volta a cair, como no município de Miguel Pereira, que mais do que dobrou o número de matrículas em 2011 e logo no outro ano retornou ao mesmo patamar. Outros municípios apresentam o mesmo comportamento, com um crescente aumento do número de matrículas e logo a seguir, uma drástica queda.

Esses fatos levam a alguns questionamentos: a falta de critérios técnicos para a decisão de quantas turmas da EJA abrir em um determinado ano, ficando a caráter do gestor a decisão, tomada muitas das vezes com base unicamente nas suas opiniões e intenções, sem nenhum tipo de argumentação concreta ou justificativa e de forma aleatória. Outro ponto é a qualidade das informações coletadas e repassadas para o INEP, colocando em dúvida se houve algum erro nesse processo. Esses fatos se refletem nos dados e mapas analisados, mostrando algumas disparidades e justificando a ausência de correlação espacial.

Quanto à análise de regressão, de antemão, sabemos que indivíduos que vivem em zonas rurais possuem uma maior dificuldade de encontrar um estabelecimento escolar que ofereça a EJA, então quanto mais urbanizada é uma cidade, maior é a taxa de matrícula dela. Outra conclusão que chegamos é que ter uma escola que ofereça o ensino médio regular, aumenta consideravelmente o número de matrículas na EJA daquela localidade. Muito desse movimento tem uma explicação na descaracterização do seu objetivo, que é de atender aquele aluno que está fora da sala de aula e com idade não considerada escolar, para servir como uma maneira de absorver alunos repetentes do ensino regular que estão com idade acima de 18 anos.

Esse trabalho contribuiu para apontar o grave cenário da EJA no estado do Rio de Janeiro, apresentando uma tendência de forte queda nos números de matrículas e a falta de alinhamento entre os municípios do estado e critérios de políticas educacionais visando atender essa demanda existente. Além da necessidade de direcionar o olhar sobre as regiões rurais, onde a taxa de atendimento é menor.

## Referências

- [1] VIEIRA, D. S. *Educação de Jovens e Adultos e Pluralidade Cultural: a realidade de um Colégio Supletivo de Ensino Médio*. 5. ed. [S.l.]: Pontifícia Universidade Católica do Rio de Janeiro, 2011.
- [2] RIBEIRO, V. M.; VÓVIO, C. L.; EDUCATIVA, A. *Educação de Jovens e Adultos: proposta curricular para o 1o. segmento do ensino fundamental*. [S.l.]: AC~ Ao Educativa, 1997.
- [3] RAUBER, A. M. d. T. R. *Concepções e Perspectivas de Educação: um estudo do Centro Estadual de Educação de Jovens e Adultos – CEEJA – Dourados/MS*. 1. ed. [S.l.]: UCDB, 2014.
- [4] BELUZO MAIRA FERREIRA ;TONIOSSO, J. P. O mobral e a alfabetização de adultos: considerações históricas. *Cadernos de Educação: Ensino e Sociedade*, n. 2, p. 196–209, 2015.
- [5] SOUZA, T. d. E. S. d. g. M. d. P. B. C. A. S. *O processo histórico de consolidação da Educação de Jovens e Adultos: as políticas públicas voltadas para EJA e a luta dos movimentos sociais para a efetivação do direito a educação*. Disponível em: <<http://www.anpae.org.br/simposio26/1comunicacoes/ThianadoEiradoSenadeSouza-ComunicacaoOral-int.pdf>>.
- [6] MACHADO, M. M. A trajetória da eja na década de 90: políticas públicas sendo substituídas por solidariedade. *REUNIÃO ANUAL DA ANPED*, ANPED São Paulo, v. 21, p. 59–73, 1998.
- [7] VENTURA, J. A oferta de educação de jovens e adultos de nível médio no estado do rio de janeiro: primeiras aproximações. *Revista Brasileira de Educação de Jovens e Adultos*, v. 4, n. 8, p. 09–35, 2016.
- [8] GUJARATI DAMODAR N.;PORTER, D. C. *Econometria Básica*. 5. ed. [S.l.]: AMGH, 2011.
- [9] RAWLINGS JOHN O; PANTULA, S. G. D. A. *Applied regression analysis: a research tool*. 2. ed. [S.l.]: Springer, 1998.
- [10] CHAPRA STEVEN C.; CANALE, R. P. *Métodos numéricos para engenharia*. 5. ed. [S.l.]: AMGH, 2011.
- [11] LATORRE, M. d. R. D. d. O. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Rev. Bras. Epidemiol*, n. 2, p. 145–152, 2001.



- [12] HAIR, J. F. et al. *Análise multivariada de dados*. 6. ed. [S.l.]: Bookman Editora, 2009.
- [13] CÂMARA, G. e. a. *Análise Espacial de Dados Geográficos*. 1. ed. [S.l.]: Livraria Virtual da EMBRAPA, 2004.
- [14] MARQUES, A. P. d. S. e. a. Análise exploratória de dados de área para índices de furto na mesorregião de presidente prudente-sp. *III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*, n. 2, p. 001–008, 7 2010. An optional note.

**ANEXO A - Número de Matrículas nos 92  
municípios entre os anos de  
2007 e 2015**

(Continua)

Município	Ano								
	2007	2008	2009	2010	2011	2012	2013	2014	2015
<b>Total</b>	143059	181749	188519	163074	132560	120513	111426	100582	109405
Angra Dos Reis	916	981	1853	633	1738	2278	1604	1539	1601
Aperibe	245	303	376	302	154	87	110	24	29
Araruama	739	961	889	1124	704	633	588	447	528
Areal	168	177	112	106	89	64	85	44	67
Armacao Dos Buzios	525	530	534	420	317	276	349	212	254
Arraial Do Cabo	2545	3096	3516	3591	738	343	1012	913	979
Barra Do Pirai	1204	1082	1049	957	949	715	561	408	521
Barra Mansa	1733	3178	3831	2821	1373	1563	930	1325	1530
Belford Roxo	2964	2528	2962	3003	2885	2015	3073	3346	1553
Bom Jardim	91	165	224	25	235	253	61	41	21
Bom Jesus Do Itabapoana	597	860	569	662	789	567	738	46	49
Cabo Frio	1383	1830	2292	2438	1710	1267	1213	1004	1038
Cachoeiras De Macacu	576	756	665	544	390	333	277	238	300
Cambuci	158	44	0	134	129	223	72	32	24
Campos Dos Goytacazes	8226	8502	8027	8298	5276	4627	3969	3543	3389
Cantagalo	171	246	196	134	117	105	68	43	95
Carapebus	0	0	0	0	0	0	25	33	42
Cardoso Moreira	244	311	214	255	173	74	73	57	54
Carmo	700	656	593	442	319	304	249	170	201
Casimiro De Abreu	617	451	303	115	168	299	542	442	283
Comendador Levy	56	148	82	0	0	0	41	51	38
Conceicao De Macabu	599	698	744	717	448	332	365	185	191
Cordeiro	868	803	689	474	367	416	410	247	428
Duas Barras	75	100	130	130	110	111	122	88	95
Duque De Caxias	4374	6619	9119	10351	7969	6982	6202	6598	6617
Engenheiro Paulo De Guapimirim	430	529	687	203	196	249	185	0	0
Iguaba Grande	168	337	300	394	305	207	191	226	224
Itaborai	3028	4821	3442	3693	3302	2873	1613	1492	1660
Itaguaí	1217	1229	357	680	462	662	860	952	1197
Italva	127	87	274	330	160	178	118	32	31
Itaocara	505	393	375	286	262	127	161	74	74
Itaperuna	1227	3315	3436	3380	1097	878	795	1162	1237
Itatiaia	529	543	374	443	215	173	178	116	130
Japeri	645	991	1047	852	760	463	530	432	603
Laje Do Muriae	0	0	0	0	0	0	10	38	67

(Continuação)

Município	Ano								
	2007	2008	2009	2010	2011	2012	2013	2014	2015
Macaé	4228	4079	4720	3213	3239	1855	1939	1756	1413
Macuco	0	63	112	101	84	73	62	20	38
Mage	2557	3344	3876	3211	2499	2263	2044	1253	1309
Mangaratiba	550	604	505	449	433	414	406	509	496
Marica	1709	1467	1463	1391	1037	865	852	546	498
Mendes	245	224	195	152	130	68	63	40	47
Mesquita	3359	3965	2869	2268	1947	2084	778	860	3097
Miguel Pereira	467	343	679	894	1935	512	507	281	236
Miracema	309	366	344	196	249	262	228	69	64
Natividade	158	199	142	149	72	95	138	26	24
Nilópolis	2078	3729	4606	4774	3654	2415	1786	1526	1385
Niteroi	9106	8868	8119	7157	5662	5729	4728	4931	4932
Nova Friburgo	2985	2850	2554	2050	1647	1582	1816	1285	1470
Nova Iguaçu	5400	5986	7121	7257	6986	5305	5017	4359	4637
Paracambi	997	1179	887	838	653	283	327	233	254
Paraíba Do Sul	539	678	607	512	357	281	271	215	164
Paraty	198	805	1210	1151	227	274	321	127	223
Paty Do Alferes	285	305	328	317	259	188	208	152	183
Petropolis	3076	3767	4357	3905	2742	3263	2124	2313	3949
Pinheiral	0	322	410	315	206	157	153	121	106
Pirai	106	175	137	114	124	57	67	77	63
Porciúncula	343	341	379	370	262	112	153	57	52
Porto Real	391	375	277	245	181	128	129	163	143
Quatis	133	144	201	161	143	123	123	83	60

(Conclusão)

Município	Ano								
	2007	2008	2009	2010	2011	2012	2013	2014	2015
Queimados	1390	1585	1521	1451	1066	917	1066	1058	1110
Quissama	263	350	271	133	184	262	337	125	89
Resende	6006	7048	2541	1574	2000	1992	1595	1302	1219
Rio Bonito	1047	1038	979	790	792	721	719	648	667
Rio Claro	0	0	0	0	0	0	29	53	47
Rio Das Flores	0	0	0	0	0	0	18	29	30
Rio Das Ostras	1371	2099	1493	2157	659	1067	805	823	1429
Rio De Janeiro	29426	43325	50299	38147	36830	38044	37274	34552	36250
Santa Maria Madalena	0	52	101	81	61	53	41	24	40
Santo Antonio De Padua	1171	1275	1019	687	530	423	323	456	610
Sao Fidelis	1116	999	1076	1105	718	697	429	392	417
Sao Francisco De	311	416	435	404	254	200	141	182	202
Sao Goncalo	8568	10674	10956	10586	9300	7255	6500	5443	6545
Sao Joao Da Barra	11	0	0	0	64	198	190	138	251
Sao Joao De Meriti	1576	3054	2842	3028	2242	1894	2162	1715	1863
Sao Jose De Uba	0	0	0	0	0	0	0	16	24
Sao Jose Do Vale Do Rio	158	171	148	120	72	64	85	74	65
Sao Pedro Da Aldeia	1796	2200	2350	2306	796	594	697	1090	1532
Sao Sebastiao Do Alto	0	0	0	0	0	0	1	0	0
Sapucaia	721	941	1104	424	307	273	385	314	360
Saquarema	560	549	518	549	389	231	327	289	319
Seropedica	413	897	966	990	852	1028	661	549	612
Silva Jardim	386	292	276	231	183	124	132	73	95
Sumidouro	46	32	46	38	39	24	32	0	0
Tangua	418	470	461	425	416	240	179	164	205
Teresopolis	2355	2796	2596	2348	1876	1607	1226	817	1151
Trajano De Moraes	0	0	0	0	0	0	14	27	30
Tres Rios	1347	1516	1477	1549	1021	960	821	1153	1388
Valenca	1323	1661	1689	1575	981	1230	1029	655	820
Varre-Sai	116	133	79	89	60	32	38	31	15
Vassouras	538	581	762	463	495	428	373	156	157
Volta Redonda	3757	6147	7036	2516	1578	1800	2019	1425	1662

Fonte: INEP - Censo Escolar.