

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

**Indicador alternativo de potencial econômico baseado no consumo de energia elétrica:**  
abordagem a partir do comportamento financeiro dos beneficiários do Cadastro  
Único para Programas Sociais

FERNANDA BARSANTE NICOLELA

SÃO PAULO

2020

FERNANDA BARSANTE NICOLELA

**Indicador alternativo de potencial econômico baseado no consumo de energia elétrica: abordagem a partir do comportamento financeiro dos beneficiários do Cadastro Único para Programas Sociais**

**Relatório final apresentado à Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas como requisito para o Programa Institucional de Bolsas de Iniciação Científica (CNPq).**

**Área de conhecimento:** Estatística e Administração Pública.

**Orientador:** Prof. Dr. Eduardo de Rezende Francisco.

SÃO PAULO

2020

## ÍNDICE

1. Introdução .....	6
2. Revisão Bibliográfica.....	7
2.1. Renda e Energia .....	7
2.1.1. Indicador Econômico Tradicional e o Critério de Classificação Econômica Brasil .....	7
2.1.2. Associação entre Renda e Consumo de Energia Elétrica.....	9
2.1.3. Consumidor de Energia Elétrica de Baixa Renda .....	12
2.2. Cadastro Único e Tarifa Social de Energia.....	12
2.2.1. O Cadastro Único para Programas Sociais .....	12
2.2.2. Tarifa Social de Energia Elétrica .....	13
3. Métodos.....	14
3.1. Estatística Espacial.....	15
4. Resultados .....	18
4.1. Seleção de Variáveis .....	18
4.2. Análise exploratória dos dados .....	19
4.3. Regressões Espaciais .....	26
4.3.1. Spacial Autoregression (SAR).....	26
4.3.2. Geographically Weighted Regression (GWR).....	26
5. Conclusões .....	27
Referências.....	29
Anexos .....	32
Anexo A: OLS da Renda a partir do Consumo de Energia Elétrica e IDEB .....	32
Anexo B: SAR Matriz Adjacência 1ª ordem da Renda a partir do Consumo de Energia e IDEB.....	32
Anexo C: GWR com <i>kernel bi-square</i> da Renda a partir do Consumo de Energia e do IDEB .....	33

## ÍNDICE DE GRÁFICOS, FIGURAS, TABELAS E QUADROS

<b>Tabela 1:</b> Estimativas de renda domiciliar mensal para os estratos socioeconômicos 2019	8
<b>Tabela 2:</b> Distribuição (%) da População Brasileira por Classe Econômica por Região Brasileira (2019)	9
<b>Tabela 3:</b> Descontos da Tarifa Social de Energia Elétrica por faixa de consumo (em kWh)	14
<b>Tabela 4:</b> Variáveis selecionadas para os modelos de regressão linear e espacial do estudo	15
<b>Tabela 5:</b> Resultados encontrados	28
<b>Figura 1:</b> Mapas (croquis) das 456 áreas de ponderação do município de São Paulo, representados; (esq.) Renda Mensal Domiciliar Média e (dir.) Consumo Médio Mensal Residencial de Energia Elétrica	11
<b>Figura 2:</b> Função distância de cada ponto espacial observado (i) – modelo GWR	17
<b>Figura 3:</b> Mapa de Renda Domiciliar (Cadastro Único) nos municípios paulistas	20
<b>Figura 4:</b> Mapa de Consumo de Energia Elétrica médio dos beneficiários da TSEE nos municípios paulistas	21
<b>Figura 5:</b> Mapa do Índice de Desenvolvimento da Educação Básica (IDEB) dos municípios paulistas	21
<b>Figura 6:</b> Box Map (hinge 1,5) da variável Renda Domiciliar (Cadastro Único) nos municípios paulistas	22
<b>Figura 7:</b> Box Map (hinge 1,5) da variável Consumo de Energia Residencial médio (TSEE)	22
<b>Figura 8:</b> Box Map (hinge 1,5) da variável IDEB dos municípios paulistas	23
<b>Figura 9:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Cluster Map</i> ) da variável Renda Domiciliar da amostra populacional do Cadastro Único dos municípios paulistas	25
<b>Figura 10:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Significance Map</i> ) da variável Renda Domiciliar da amostra populacional do Cadastro Único dos municípios paulistas	25
<b>Figura 11:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Cluster Map</i> ) da variável Consumo Médio de Energia Elétrica Residencial dos beneficiários da TSEE dos municípios paulistas	25
<b>Figura 12:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Significance Map</i> ) da Consumo Médio de Energia Elétrica Residencial dos beneficiários da TSEE dos municípios paulistas	25
<b>Figura 13:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Cluster Map</i> ) do IDEB dos municípios paulistas	25
<b>Figura 14:</b> Mapa do indicador de auto correlação espacial local ( <i>LISA Significance Map</i> ) do IDEB dos municípios paulistas	25
<b>Figura 15:</b> Mapa do Sistema Elétrico do Estado de São Paulo	28
<b>Gráfico 1:</b> Evolução do Consumo de Energia Elétrica e do PIB Real do Brasil de 1952 a 2018	10
<b>Gráfico 2:</b> Evolução anual do Consumo de Energia Elétrica Residencial e da Renda Familiar PIB Real do Brasil de 1990 a 2014	10
<b>Gráfico 3:</b> Histogramas de: 1) Renda Média Mensal Domiciliar (2018), 2) Consumo de Energia Elétrica Médio (2019), 3) IDEB (2018)	19
<b>Gráfico 4:</b> Gráfico de dispersão de Moran para a variável 1) Renda Domiciliar Média do Cadastro Único (2017) dos municípios do Estado de São Paulo, 2) Consumo Médio de Energia Elétrica Residencial (TSEE) e 3) IDEB dos municípios paulistas	24
<b>Quadro 1:</b> Resultados do Modelo SAR com matriz de adjacência Queen de 1ª ordem da Renda a partir do Consumo de Energia Elétrica e do IDEB no Estado de São Paulo	26
<b>Quadro 2:</b> Resultados da Regressão GWR com kernel bi-square da Renda a partir do Consumo de Energia Elétrica e do IDEB no Estado de São Paulo	27

## LISTA DE ABREVIATURAS E SIGLAS

ABEP	Associação Brasileira de Empresas de Pesquisa
ANEEL	Agência Nacional de Energia Elétrica
ANEP	Associação Nacional das Empresas de Pesquisa
BPC	Benefício de Prestação Continuada da Assistência Social
CADÚNICO	Cadastro Único para Programas Sociais
CCEB	Critério de Classificação Econômica Brasil
GWR	<i>Geographically Weighted Regression</i> (Regressão Ponderada Geograficamente)
IBGE	Instituto Brasileiro de Geografia e Estatística
IDEB	Índice de Desenvolvimento da Educação Básica
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
IPEA	Instituto de Pesquisa Econômica Aplicada
IPRS	Índice Paulista de Responsabilidade Social
IPVS	Índice Paulista de Vulnerabilidade Social
LISA	<i>Local Indicators of Spatial Association</i>
OLS	<i>Ordinary Least Squares</i> (Regressão por Mínimos Quadrados)
PERC	<i>Policy and Economic Research Council</i>
PIB	Produto Interno Bruto
PNAD	Pesquisa Nacional por Amostra de Domicílios
PNUD	Programa das Nações Unidas para o Desenvolvimento
POF	Pesquisa de Orçamentos Familiares
SAR	<i>Spatial Auto-Regressive model</i> (Modelo Espacial Auto-Regressivo)
SEADE	Fundação Sistema Estadual de Análise de Dados
TSEE	Tarifa Social de Energia Elétrica
WLM	<i>Weighted Linear Model</i>

## **Indicador alternativo de potencial econômico baseado no consumo de energia elétrica: abordagem a partir do comportamento financeiro dos beneficiários do Cadastro Único para Programas Sociais**

**Palavras-chave:** *Renda domiciliar, Consumo de Energia Elétrica, Cadastro Único, Tarifa Social de Energia, Big Data Analytics, Estatística Espacial.*

### **RESUMO**

O presente trabalho busca investigar o comportamento financeiro das famílias inscritas no Cadastro Único para Programas Sociais com vistas a construção de um indicador alternativo de potencial econômico, baseado em variáveis oriundas das empresas de serviços públicos, como o consumo de energia elétrica. Isto porque, uma vez que o tradicional indicador utilizado para obter informações sobre a situação econômica das famílias brasileiras não é significativamente preciso, um novo indicador se faz necessário no sentido de propiciar informações atuais e atualizáveis sobre a renda das famílias e, assim, auxiliar o direcionamento e ajuste de políticas públicas assistenciais. Para isso, utilizou-se a análise de informações alternativas no âmbito de técnicas do Big Data Analytics e Estatística Espacial e encontrou-se que o consumo de energia elétrica dos beneficiários da Tarifa Social de Energia Elétrica é estatisticamente significativo para responder a renda domiciliar dos inscritos no Cadastro Único para Programas Sociais.

## 1. Introdução

A utilização de resultados e informações estatísticas contribui significativamente para as tomadas de decisões na atualidade. Setores públicos e privados utilizam da produção de dados a partir das ferramentas estatísticas para definir suas metas e estratégias e priorizar e direcionar suas ações para alcançar seus objetivos de maneira eficiente: seja para formular, avaliar e implementar políticas públicas, seja para entender o comportamento dos consumidores e usar esse conhecimento no desenvolvimento de estratégias de marketing e comunicação. Logo, faz-se necessário coletar informações fidedignas a respeito da população, seus habitantes, suas famílias e suas características, em termos de cobertura, granularidade e suficiência de conceitos e dimensões em mensuração (FRANCISCO, 2010).

No âmbito da Administração Pública, as tomadas de decisões voltadas a resolver problemas sociais exigem embasamento para sua execução e eficácia. Para isso, o gestor público deve entender a dinâmica do território o qual suas ações se estabelecem para melhor direcionar e ajustar as políticas públicas, bem como incluir pautas na agenda governamental. Nesse sentido, e sob a perspectiva dos problemas relacionados às desigualdades socioeconômicas, destaca-se o Cadastro Único para Programas Sociais do Governo Federal (CadÚnico) como ferramenta de gestão para identificar e caracterizar as condições socioeconômica das famílias brasileiras de baixa renda e, assim, direcionar determinados benefícios.

O preenchimento do CadÚnico exige a declaração de renda da família – variável que, tradicionalmente, é utilizada para obter informações sobre a situação econômica das famílias brasileiras. A partir destas informações são elaborados estudos sobre condição de vida e pobreza (BUSSAB; FERREIRA, 1999) e, fundamentados nelas, elaboram-se ações com vistas a impactar positivamente a vida de milhares de famílias. No entanto, cabe salientar a dificuldade de se obter informações precisas, atuais e atualizáveis sobre esta variável, uma vez que está sujeita a super declaração, sub declaração, recusa e sazonalidade da fonte do rendimento das famílias (FRANCISCO, 2010; BUSSAB; FERREIRA, 1999). Logo, compromete-se a precisão sobre a condição financeira das famílias brasileiras e, conseqüentemente, o direcionamento de programas e incentivos.

Nessa perspectiva, estudos da *Policy and Economic Research Council* (PERC) evidenciam que indicadores alternativos provenientes das empresas de serviços públicos são cada vez mais necessários quando se almeja obter informações mais precisas sobre a condição financeira das pessoas, principalmente em mercados de crédito em que a lacuna de informação sobre a capacidade de adimplência dos indivíduos é uma significativa preocupação (TURNER et al., 2006). Nesse âmbito, destaca-se o serviço público de distribuição de energia elétrica como serviço essencial de maior capilaridade nacional, o qual é mensurado mensalmente, com informações de localização, histórico e sazonalidade.

Com base na forte relação identificada entre renda e consumo de energia no município de São Paulo (FRANCISCO, 2006, 2010), o presente trabalho objetiva estabelecer um indicador alternativo de potencial econômico que, por sua vez, será discutido considerando o parâmetro das famílias que integram políticas assistenciais nacionais. Enquanto premissa que justifica o presente estudo, tem-se que o desenvolvimento deste indicador alternativo pode contribuir para uma classificação econômica mais precisa do público que se serve desta categorização para o acesso a programas sociais de beneficiamento e superação das

desigualdades, assim como pode fornecer suporte à gestão e proposição de políticas públicas e programas assistenciais.

Desse modo, como hipóteses da pesquisa, busca-se responder, sob a perspectiva da análise de dados das famílias residentes nos municípios do Estado de São Paulo, se existe relação positiva entre Consumo de Energia Elétrica Residencial dos beneficiários da Tarifa Social de Energia Elétrica (TSEE) e a Renda Familiar dos inscritos no Cadastro Único - quanto maior o consumo, maior a renda e se existe influência geográfica na explicação dessas variáveis. Para solucioná-las utilizou-se regressão linear multivariada da variável renda (dependente) na presença de outras variáveis (independentes) definidas no sentido de corroborar com características determinantes da renda. Em sequência aplicou-se métodos da Estatística Espacial para analisar a distribuição das variáveis no espaço geográfico.

Como principais achados do trabalho destaca-se que a Renda Domiciliar Média das famílias cadastradas no Cadastro Único tem relação com o Consumo de Energia Elétrica Domiciliar daqueles que são beneficiários da Tarifa Social de Energia (TSEE) e com o Índice de Desenvolvimento de Educação Básica (IDEB) das respectivas cidades – uma das variáveis selecionadas que corroboram com a renda. Isso será melhor detalhado nos Capítulos 4 e 5.

A pesquisa divide-se em cinco partes. A primeira apresentará uma revisão bibliográfica e algumas análises estatísticas iniciais a respeito da relação entre energia e renda. A segunda aborda questões relativas à metodologia utilizada, com destaque para o uso de ferramentas de Estatística Espacial. A terceira e quarta partes, por sua vez, tratam das análises propriamente ditas, bem como do processo para obtenção dos resultados e do estabelecimento do indicador. Por fim a última parte versa sobre os principais achados da pesquisa e tece as considerações finais acerca dos trabalhos realizados.

## **2. Revisão Bibliográfica**

### **2.1. Renda e Energia**

#### **2.1.1. Indicador Econômico Tradicional e o Critério de Classificação Econômica Brasil**

Renda é, tradicionalmente, o indicador utilizado para obter informações sobre a situação econômica das famílias brasileiras. Como dito anteriormente, é um desafio a obtenção de dados precisos e atualizáveis a seu respeito - uma vez que está sujeita a super declaração, sub declaração, recusa e sazonalidade da fonte do rendimento das famílias (FRANCISCO, 2010; BUSSAB; FERREIRA, 1999). Com isso em vista, Institutos de Pesquisa utilizam indicadores referente a Classe Econômica ou o Poder de Consumo baseados, entre outros, na posse de bens duráveis da família.

Nesse ambiente destaca-se o indicador Critério de Classificação Econômica Brasil (CCEB) - definido pela Associação Nacional de Empresas de Pesquisa (ANEP) em 1996 e mantido, desde 2004, pela Associação Brasileira de Empresas de Pesquisa (ABEP) - que classifica a população brasileira em segmentos econômicos. Para este critério, são consideradas determinadas características dos domicílios brasileiros, tais como a existência e a quantidade de bens materiais, o acesso a serviços públicos e a escolaridade do chefe de família e, a partir delas as famílias são classificadas em um critério de pontos, os quais são usados como ferramenta de divisão de classes socioeconômicas.



Tal critério é atualizado a cada ano para melhor se adequar a realidade econômica dos domicílios brasileiros. Destaca-se a versão de 2015, intitulada Novo Critério Brasil, que alterou a metodologia adotada para o cálculo do critério: passou a ser considerado, em detrimento da declaração de renda, variáveis indicadoras de renda permanente fundamentadas nos dados da Pesquisa de Orçamentos Familiares (POF) do Instituto Brasileiro de Geografia e Estatística (IBGE). Dentre as principais alterações ocorridas, destaca-se a inclusão de variáveis de infraestrutura urbana - pavimentação de ruas e água encanada - além do ajuste de itens domiciliares sujeitos a evolução da tecnologia e sua incorporação na vida cotidiana, como inclusão de microcomputador, micro-ondas e motocicleta.

O critério vigente, de junho de 2019, estima, a partir das respostas ao questionário referente às categorias de produtos e serviços utilizados pela família e à informações da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) de 2018, a renda domiciliar mensal da população brasileira, dividindo-a em seis estratos socioeconômicos denominados A, B1, B2, C1, C2 e DE, como se vê na Tabela 1:

**Tabela 1:** Estimativas de renda domiciliar mensal para os estratos socioeconômicos 2019

Estrato Sócio Econômico	Renda média domiciliar
A	25.554,33
B1	11.279,14
B2	5.641,64
C1	3.085,48
C2	1.748,59
DE	719,81
<b>TOTAL</b>	<b>3.014,01</b>

Fonte: ABEP - Associação Brasileira de Empresas de Pesquisa (2019)

Bem como observado em estudos realizados pela ABEP em 2019, constata-se que os valores de rendas estimados possuem variância elevada, com sobreposições importantes nas rendas entre as diferentes classes. Isso significa que explorar apenas a renda não é eficiente para a determinação do nível socioeconômico e, por isso, segundo o estudo, este indicador não substitui ou complementa o questionário sobre as variáveis indicadoras de renda permanente e de serviços utilizados pela família.

Ademais, o Novo Critério Brasil ressalta importantes dissemelhanças regionais (Tabela 3), além de não ser adequado para a caracterização de famílias alocadas nas extremidades da distribuição de renda (FRANCISCO, 2019; MATTAR, 1996; SILVA, 2002) - especialmente, para o consumidor de rendimentos mais baixos. O parâmetro fornece uma segmentação das grandes massas e concilia-se bem com os estudos de abrangência nacional. Para as regiões ou parcelas específicas, há a necessidade de especialização ou adaptações do Critério Brasil para a caracterização inicial da população analisada que, muitas vezes, abrange variáveis que melhor caracterizam a idoneidade para o consumo. O uso de indicadores de consumo que dispõem de cobertura e utilidade mais amplas, mostram-se úteis nesse processo de caracterização dos consumidores.

**Tabela 2:** Distribuição (%) da População Brasileira por Classe Econômica por Região Brasileira (2019)

Classe	Brasil	Sudeste	Sul	Nordeste	Centro Oeste	Norte
1 - A	2,5%	3,1%	3,0%	1,3%	4,0%	1,3%
2 - B1	4,4%	5,6%	5,3%	2,3%	5,3%	2,8%
3 - B2	16,5%	20,5%	21,3%	9,1%	18,8%	8,6%
4 - C1	21,5%	25,1%	26,6%	14,8%	22,4%	13,8%
5 - C2	26,8%	26,3%	28,1%	26,4%	28,0%	26,7%
6 - D - E	28,3%	19,4%	15,7%	46,1%	21,5%	46,8%
total	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: ABEP - Associação Brasileira de Empresas de Pesquisa (2019)

### 2.1.2. Associação entre Renda e Consumo de Energia Elétrica

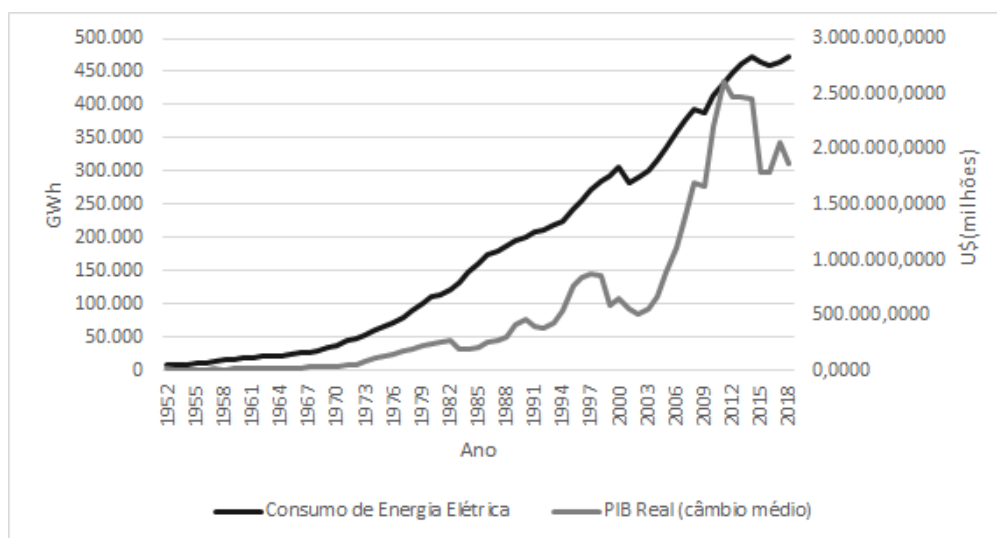
Frente os impasses decorrentes da utilização do indicador econômico tradicional, estudos da *Policy and Economic Research Council* (PERC) evidenciam que indicadores alternativos provenientes das empresas de serviços públicos são cada vez mais necessários quando se almeja obter informações mais precisas sobre a condição financeira das pessoas. Nesse âmbito, destaca-se o serviço público de distribuição de energia elétrica: de acordo com PNADC 2019, 99,8% dos domicílios particulares permanentes tinham acesso à eletricidade, configurando o serviço de maior capilaridade nacional. Assim, a variável “Consumo de Energia Elétrica” propicia análises estatísticas atuais e atualizáveis - visto que é um serviço coletado mensalmente, com informações de localização, histórico e sazonalidade.

Numa perspectiva histórica, a energia elétrica esteve intimidante relacionada a desenvolvimento – compreendido como sinônimo de industrialização, urbanização e consumo de bens duráveis, os quais possuem a eletricidade como insumo básico. Com o aprimoramento da utilização deste insumo, este passou a estar intimamente ligado a melhoria do bem-estar dos indivíduos no sentido de proporcionar, por exemplo, a conservação de alimentos, a facilitação da comunicação, o desenvolvimento tecnológico, além de significativos avanços nas áreas de saúde, educação e economia (CORREIA-SILVA *et al.*, 2007). Outrossim, o consumo de energia elétrica fornece parâmetros sobre a facilitação do desenvolvimento de atividades geradoras de renda exercidas no próprio domicílio.

Desse modo, a variável “Consumo de Energia Elétrica” diz muito a respeito das condições sociais e econômicas das famílias e, por conseguinte, de determinada sociedade. Esta variável é capaz de ajudar a inferir uma melhor classificação socioeconômica a partir de faixas de consumo e contemplar uma melhor definição do consumidor de baixa renda, especialmente nas regiões de menor acesso e maior dificuldade de levantamento de bens de consumo.

Para além dos pontos positivos a respeito da utilização Consumo de Energia Elétrica como base para a construção do indicador econômico alternativo, cabe salientar a significativa relação estatística entre esta variável e o PIB Real – tradicional indicador utilizado para retratar a riqueza dos países. No período de 1952 a 2018, apresentado no Gráfico, a correlação entre as duas variáveis foi de 0,92, ou seja, PIB e Consumo de Energia Elétrica apresenta uma relação forte e positiva. Nessa lógica, a partir de uma análise visual, é possível perceber que ambas as linhas possuem tendências similares. Entretanto, destaca-se que a partir de 2012 o PIB Real

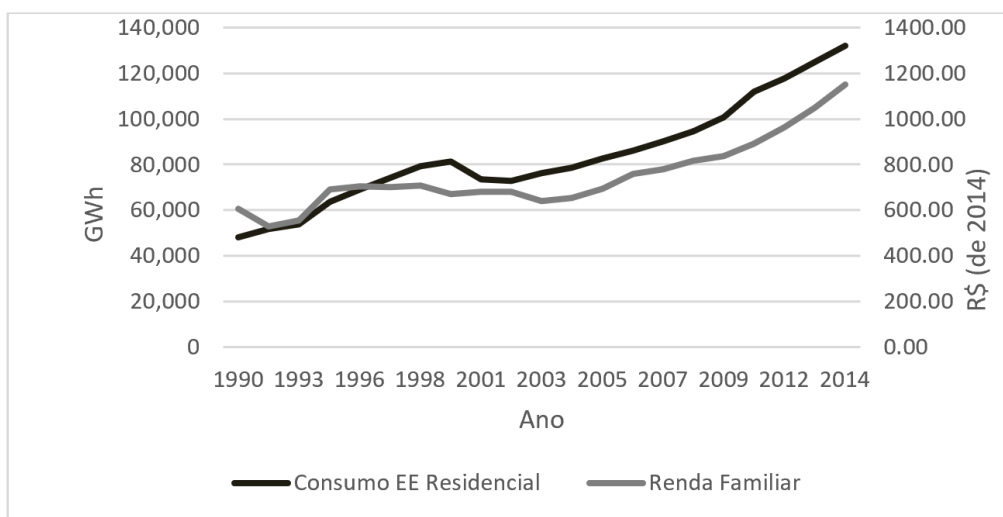
brasileiro entrou em principalmente devido, principalmente, segundo o governo e o IBGE, às condições climáticas que afetaram o desempenho da agropecuária<sup>1</sup>.



**Gráfico 1:** Evolução do Consumo de Energia Elétrica e do PIB Real do Brasil de 1952 a 2018

**Fonte:** Elaboração própria a partir de dados do Banco Central do Brasil e da Eletrobrás (IPEADATA)

Assim como o PIB Real, a variável Renda Familiar do brasileiro também guarda uma significativa relação com a variável Consumo de Energia Elétrica Residencial. A partir da elaboração do gráfico abaixo, ficou visivelmente clara a existência de uma correlação positiva entre essas variáveis no período de 1990 a 2014. Matematicamente o valor dessa correlação é de 0,83, o que implica uma relação forte e positiva entre as variáveis.



**Gráfico 2:** Evolução anual do Consumo de Energia Elétrica Residencial e da Renda Familiar PIB Real do Brasil de 1990 a 2014

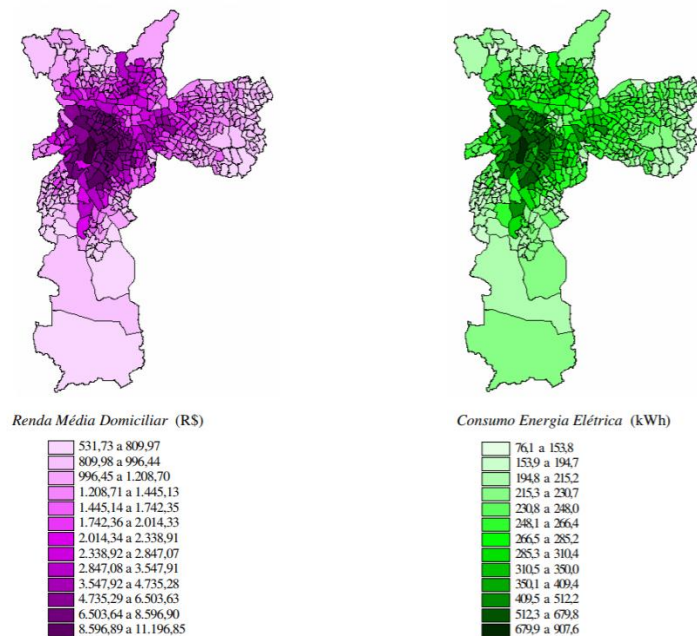
**Fonte:** Elaboração própria a partir de dados IPEADATA.

<sup>1</sup> <http://g1.globo.com/economia/agronegocios/noticia/2013/03/clima-e-barreiras-sanitarias-prejudicaram-pib-da-agropecuaria-em-2012.html#:~:text=Rio%2C%2001%20%2D%20Os%20problemas%20clim%C3%A1ticos,2%2C%25%20no%20ano.>

Concomitantemente a esta visão, constata-se que as distribuições dos fenômenos Renda e Consumo de Energia Elétrica no espaço geográfico não se dão aleatoriamente. Isto porque a dinâmica urbana, essencialmente relacionada a motivações históricas, políticas e sociais, determina a localização das indústrias e empresas de comércio e serviços de forma a favorecer sua proximidade à matéria prima (ou aos insumos) e ao mercado consumidor. A população, portanto, se estabelece dentro desse contexto espaço-temporal.

Ademais, usando como base a “primeira lei da geografia” (TOBLER, 1970), as “coisas” mais próximas tendem a ser mais parecidas entre si, ou seja, pessoas que vivem em uma mesma vizinhança são mais parecidas, por familiaridade cultural, razões econômicas ou histórias. Fato este que ajuda a explicar bairros com perfis distintos. Dessa forma, não se pode deixar de explorar a influência que a geografia tem na explicação da Renda.

Nesse sentido, estudos anteriores evidenciaram – por meio de modelos de regressão linear múltipla - que o Consumo de Energia Elétrica contribui para a explicação da variabilidade da Renda em áreas censitárias do IBGE (FRANCISCO, 2006). A Figura 1 apresenta as 456 áreas de ponderação do Censo Demográfico do IBGE para o município de São Paulo de acordo com sua Renda Mensal Domiciliar Média e seu Consumo Médio Mensal Residencial de Energia Elétrica. Assim, a partir de uma análise visual, é possível perceber a tendência da concentração das duas variáveis ser de natureza geográfica - à medida que as áreas da ponderação se afastam do centro, o valor das medidas diminuem (padrão Centro-Periferia). Desse modo, faz-se necessário, na construção do indicador alternativo de potencial econômico, a utilização de métodos da Estatística Espacial.



**Figura 1:** Mapas (croquis) das 456 áreas de ponderação do município de São Paulo, representados; (esq.) Renda Mensal Domiciliar Média e (dir.) Consumo Médio Mensal Residencial de Energia Elétrica. **Fonte:** FRANCISCO, 2010.

### **2.1.3. Consumidor de Energia Elétrica de Baixa Renda**

A distribuição de energia elétrica configura um serviço público essencial. Assegurado normativamente pela Lei nº 7.783/1989<sup>2</sup>, tais serviços são aqueles prestados pela Administração Pública - direta ou indireta, cuja interrupção coloca em risco iminente a sobrevivência, a saúde ou a segurança da população (BRASIL, 1989). Desse modo, a comercialização da distribuição de energia é submetida a regras que visam a proteção do consumidor<sup>3</sup>, principalmente o de baixa renda.

Este papel é atribuído a Agência Nacional de Energia Elétrica, agência reguladora encarregada de proteger e promover os direitos dos consumidores, através da prestação do serviço de energia amparada em critérios de equidade, ética, defesa da concorrência e garantia de direitos individuais e coletivos (BORGER et al., 2015). Dentre as ações para promoção dos serviços públicos relacionados ao fornecimento de energia, vale citar o estabelecimento, por parte da administração pública brasileira, da Tarifa Social de Energia Elétrica (TSEE) - que será tratada em seguida.

Por outro lado, o país ainda sofre com um alto grau de desigualdade no fornecimento de infraestrutura para uma parcela relevante da população, cujo acesso aos serviços essenciais, tal qual a energia, não são garantidos de maneira plena e/ou suficiente. Em decorrência da falha da administração pública em prover esses serviços de forma universal, e tendo em vista outros fatores (culturais, sociais e econômicos), uma parcela dos consumidores de baixa renda é pressionada a praticar ligações por conta própria, o que acaba por gerar instalações impróprias e tem por uma das consequências a fraude no pagamento da energia elétrica. Alguns dos fatores que impulsionam esse cenário são: desemprego, elevado nível de impostos e encargos na tarifa da energia elétrica, o elevado nível de perdas (técnicas e comerciais), os custos crescentes de produção (FRANCISCO, 2006, 2010) e a desigualdade no fornecimento de infraestrutura pública e acesso aos serviços essenciais.

No que diz respeito a parcela de consumidores de baixa renda, cabe salientar que, de acordo com a Pesquisa Nacional de Amostras por Domicílio (PNAD) Contínua de 2019, cerca de 13,8 milhões de brasileiros vivem na extrema pobreza – 6,7% da população nacional. Sob a perspectiva de Prahalad e Hart (2002), esta parcela da população – base da pirâmide, configura um mercado imenso e inexplorado pelas organizações que, em sua maioria, orientam seus produtos/serviços para o topo da pirâmide. Sem deixar a rentabilidade de lado, esses autores defendem e estimulam a adoção de estratégias de direcionamento de investimentos e inovações para melhorar a qualidade de vida da população de baixa renda – que necessitam de instrumentos de gestão capazes de identificá-las com êxito.

## **2.2. Cadastro Único e Tarifa Social de Energia**

### **2.2.1. O Cadastro Único para Programas Sociais**

---

<sup>2</sup> A Lei nº 7.783/1989, conhecida como Lei da Greve, regulamenta o art. 9º, § 1º, da Constituição Federal de 1988 que discorre sobre a existência de serviços ou atividades essenciais da comunidade. Para além da distribuição de energia elétrica, são exemplos de serviços públicos essenciais o abastecimento de água, os serviços de telecomunicação, de saneamento básico, transporte público etc.

<sup>3</sup> Vide Código de Defesa do Consumidor – Lei nº 8.078/1990.

O Cadastro Único para Programas Sociais do Governo Federal (Cadastro Único) é uma ferramenta de gestão que permite identificar e caracterizar a condição socioeconômica das famílias brasileiras de baixa renda. O Cadastro Único configura uma porta de entrada à participação de políticas e programas sociais de enfrentamento à pobreza e às desigualdades sociais.

Criado em 2001, mediante Decreto nº 3.877, o Cadastro Único surgiu para centralizar o cadastramento de programas sociais que estavam sendo implementados na época - como os Programas Bolsa Escola, o Auxílio-Gás e o Bolsa Alimentação, criados no segundo mandato do Presidente Fernando Henrique Cardoso. O intuito era integrar as informações dos beneficiários para assim facilitar a coordenação das ações e aumentar a eficiência da gestão dos governos federal, estadual e municipal.

Este instrumento consolidou-se e fortaleceu-se com a criação do Programa Bolsa Família (PBF), em 2003 - programa de transferência direta de renda. Com o intuito de proporcionar alívio imediato à situação de pobreza e extrema pobreza enfrentada por milhões de famílias brasileiras, assim como induzir a inclusão social e o acesso a direitos básicos, o PBF foi elaborado durante o primeiro mandato do governo Lula como decorrência da expansão, consolidação e unificação dos programas sociais citados anteriormente. Desse modo, em 2004, o PBF foi sancionado pela Lei nº. 10.836 como um programa destinado às ações de transferência de renda com condicionalidades.

O programa é fundamentado em três eixos: (i) complemento de renda, para aliviar imediatamente a pobreza; (ii) acesso a direitos, por meio das condicionalidades de saúde, educação e assistência social, para que as futuras gerações da família quebrem o ciclo da pobreza, e (iii) articulação com outras ações sociais para estimular o desenvolvimento das famílias.

Entende-se por famílias em situação de extrema pobreza aquelas que possuem renda mensal de até R\$ 89,00 por pessoa, ao passo que as em situação de pobreza, entre R\$ 89,01 e R\$ 178,00 por pessoa - desde que seja composta por crianças ou adolescentes de 0 a 17 anos. Segundo o Governo Federal, o Cadastro Único reúne cerca de 73 milhões de cadastros, sendo, aproximadamente, 14 milhões os beneficiários do PBF.

Hodiernamente, o Cadastro Único é utilizado para selecionar os beneficiários de diversos outros programas sociais além do PBF, tais como: Tarifa Social de Energia Elétrica (TSEE), Programa Minha Casa Minha Vida, Benefício de Prestação Continuada (BPC), Programa Cisternas, Água para Todos, Telefone Popular, Programa Brasil Alfabetizado, entre outros.

### **2.2.2. Tarifa Social de Energia Elétrica**

A Tarifa Social de Energia Elétrica (TSEE) é um benefício concedido pelo Governo Federal, coordenado e regulamentado pela Agência Nacional de Energia Elétrica (ANEEL) - agência reguladora na forma de autarquia federal, vinculada ao Ministério das Minas e Energia que regula o setor elétrico brasileiro. A TSEE consiste na concessão de descontos na conta de energia elétrica para os consumidores enquadrados na subclasse residencial baixa renda.

Criada pela Lei nº 10.438 de 2002 e alterada e regulamentada pela Lei nº 12.212 de 2010, a TSEE é assegurada às residências habitadas por: (i) famílias inscritas no Cadastro Único, com renda familiar mensal per capita menor ou igual a meio salário mínimo nacional; (ii) idosos com 65 anos ou mais ou pessoas com deficiência, que recebam o Benefício de Prestação Continuada da Assistência Social (BPC); (iii) e, excepcionalmente, famílias inscritas no Cadastro Único com renda mensal de até 3 salários mínimos, que tenham entre seus membros pessoa em tratamento de saúde domiciliar que demande uso contínuo de equipamentos hospitalares que consumam energia elétrica (BRASIL, 2010).

Os descontos são cumulativos e decrescentes, por faixa de consumo, incidente na conta de energia elétrica, conforme tabela abaixo:

**Tabela 3:** Descontos da Tarifa Social de Energia Elétrica por faixa de consumo (em kWh)

<b>Consumo Mensal</b>	<b>Desconto</b>
de 0 a 30 kWh	65%
de 31 kWh a 100 kWh	40%
de 101 kWh a 220 kWh	10%
a partir de 221 kWh	0%

Fonte: ANEEL, 2020.

### 3. Métodos

Para obter os resultados acerca do indicador alternativo de potencial econômico, a pesquisa concentrar-se-á em análise e estudos de técnicas de Big Data e de métodos de inferência estatística - regressão linear, múltipla e espacial. Aplicou-se os modelos de *Ordinary Least Squares (OLS)*, *Spatial Autoregression (SAR)*, *Geographically Weighted Regression (GWR)* com vistas a identificar a dependência entre as variáveis e como ela está distribuída no espaço.

Inicialmente, para embasar a hipótese do presente trabalho, realizou-se estudos bibliográficos a respeito da relação entre Consumo de Energia Elétrica e Renda. Para além da perspectiva histórica dessa relação, publicações de órgãos públicos e reportagens também foram analisadas para compreender a real implementação e impacto do serviço público essencial de energia e dos programas sociais existentes para amenizar lacunas de desigualdade socioeconômicas.

Em sequência, fez-se a busca e seleção das variáveis para a realização dos modelos estatísticos. Nas regressões as variáveis são classificadas em dependente ou independentes – isto para encontrar a dependência que uma variável possui (dependente) em função de outra (s) (independente(s)) e, assim, construir o indicador. A previsão da variável dependente, neste estudo, refere-se a Renda Domiciliar Média das famílias do Cadastro Único - presente nos microdados amostrais desidentificados do CadÚnico, disponíveis pela plataforma do governo

federal da Secretaria de Avaliação e Gestão da Informação (SAGI), e as independentes foram definidas no sentido de corroborar com características determinantes desta variável.

Entre elas destaca-se a variável Consumo de Energia Elétrica dos Beneficiários da Tarifa Social de Energia Elétrica (TSEE). Sua obtenção foi possível devido o contato direto com gestora ANEEL, que disponibilizou, sob a Lei de Acesso à Informação (LAI), o consumo de energia elétrica faturado (kWh) de cada uma das famílias beneficiárias do programa do Estado de São Paulo.

As demais variáveis independentes selecionadas são:

**Tabela 4:** Variáveis selecionadas para os modelos de regressão linear e espacial do estudo

Variáveis	Ano	Fonte
Consumo de Energia Elétrica dos beneficiários do TSEE	2019	ANEEL
Quantidade de famílias cadastradas	2018	Microdados do Cadastro Único para Programas Sociais
Quantidade de pessoas por família		
PIB do município p (Em mil reais correntes)	2017	Fundação SEADE
PIB per capita do município (Em reais correntes) Índice Paulista de Responsabilidade Social (IPRS) - Dimensão Riqueza		
Participação do município no PIB do Estado (em%)		
Índice de Desenvolvimento da Educação Básica (IDEB) - Ensino Fundamental - Anos Iniciais - Rede Pública		
Índice Paulista de Responsabilidade Social (IPRS) - Dimensão Riqueza	2018	
Empregos Formais		
Índice Paulista de Vulnerabilidade Social	2010	

### 3.1. Estatística Espacial

Para poder observar as distribuições das variáveis em mapas e descobrir influências geográficas faz-se necessário o uso de ferramentas da estatística espacial. Esta análise permite



ao gestor público entender a dinâmica do território o qual suas ações se estabelecem. É através de uma visão holística do território que se faz possível compreender padrões existentes no espaço geográfico, assim como mensurar propriedades e relacionamentos no que diz respeito a localização espacial do fenômeno em estudo (CÂMARA *et al.*, 2003).

Dessa forma, um dos modos de incorporar o espaço geográfico nos modelos estatísticos é integrar termos autorregressivos espaciais baseados em proximidade ou adjacência. Uma das possibilidades é utilizar as Matrizes de Vizinhança, definidas por contiguidade ou distância – ferramenta que demonstra se o comportamento de uma variável é influenciado pelo comportamento dos seus vizinhos. Recomenda-se que seja feita uma análise da distribuição espacial das variáveis estudadas e uma formulação teórica prévia para ajudar a determinar o critério que será utilizado na construção da matriz (FRANCISCO, 2010).

Uma das maneiras de estudar a dependência espacial da variável dependente é através do Modelo Autorregressivo (SAR) – no qual a matriz de vizinhança (W) é incorporada como uma variável dependente espacialmente defasada (Wy) multiplicado pelo coeficiente espacial autoregressivo ( $\rho$ ). A hipótese nula nessa regressão, estipulada por  $\rho = 0$ , é que não há autocorrelação espacial.

$$\mathbf{SAR}: y = \rho W y + X \beta + \varepsilon$$

Uma medida de autocorrelação espacial é o índice  $I$  de Moran - considerado um indicador de correlação entre o valor de uma variável observada em uma região e os valores da mesma variável observados em suas regiões vizinhas (GRIFFITH, 1987). Calcula-se da seguinte forma:

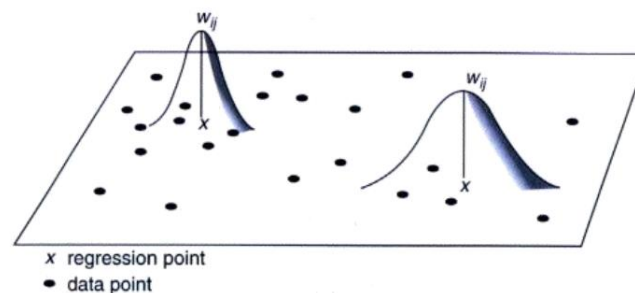
$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

em que  $y$  é a variável de interesse e  $w_{ij}$  é uma célula da matriz de vizinhança. Para que o cálculo do índice de Moran represente um indicador de autocorrelação espacial, os pesos das vizinhanças de cada observação devem ser ponderados para que todas as linhas da matriz de

vizinhança tenham soma igual a 1. Desta forma, temos que  $\sum_{i=1}^n \sum_{j=1}^n w_{ij} = n$ . O indicador varia de -1 (perfeita dispersão) a 1 (perfeita associação), passando pelo valor de padrão espacial totalmente aleatório (0) (FRANCISCO, 2010).

Cabe salientar que cada observação  $i$  possui uma contribuição local para compor o indicador global de dependência espacial. O conjunto de contribuições locais, ou componentes individuais, denomina-se *Local Indicators of Spatial Association (LISA)*, o qual pode ser visualizado espacialmente e testado para significância estatística para prover uma indicação de padrões de concentração espacial da variável em estudo na área (FRANCISCO, 2010).

Em paralelo ao LISA destaca-se o método *Geographically Weighted Regression* (GWR)<sup>4</sup>, utilizado com vistas a aumentar a confiabilidades das medidas globais. O modelo GWR elabora modelos locais para cada ponto espacial observado (i), ponderando todas as observações, baseadas em k vizinhos, em função da distância (ou de qualquer medida de vizinhança) deste ponto.



**Figura 2:** Função distância de cada ponto espacial observado (i) – modelo GWR.  
**Fonte:** FOTHERINGHAM; BRUNSDON, CHARLTON, 2002.

A expressão numérica do modelo é dada a seguir:

$$\mathbf{GWR:} \quad y(g) = \beta_0(g) + \beta_1(g)x_1 + \beta_2(g)x_2 + \dots + \beta_p(g)x_p + \varepsilon$$

em que  $g$  é um vetor dos  $n$  pontos que configuram o espaço e cada um deles possui um parâmetro  $\beta$  distinto. Assim, teremos  $n$  regressões – uma a cada ponto  $g_i$  – o que torna o modelo espacial mais preciso.

Cabe salientar que a escolha da função de proximidade (ou ponderação) espacial é muito importante para o modelo, pois ela determina a influência espacial das observações entre si. Os diferentes esquemas de ponderação são denominados *kernels*, os quais podem ser classificados como: *bi-square* fixo, *gaussiano* fixo, *bi-square* adaptável e *gaussiano* adaptável – que variam em relação ao espaçamento e a densidade de observações. Essa variação pode ser aplicada da seguinte forma:

- **Kernel Gaussiano:**  $w_{ij} = e^{-0,5(d_{ij}/h)^2}$  , se  $d_{ij} < h$  (largura de banda)  
 (fixo)  
 $w_{ij} = 0$  , caso contrário;

---

- **Kernel Bi-square:**  $w_{ij} = \left(1 - (d_{ij}^2 / h^2)\right)^2$  , se  $d_{ij} < h$  (largura de banda)  
 (fixo)  
 $w_{ij} = 0$  , caso contrário;

---

- **Kernel Gaussiano:**  $w_{ij} = e^{-0,5(d_{ij}/h)^2}$  , se  $g_j$  é um dos  $k$  vizinhos mais próximos de  $g_i$  e  $h$  (largura de banda) é a distância do  $k$ -ésimo vizinho mais próximo.  
 (adaptável)  
 $w_{ij} = 0$  , caso contrário;

<sup>4</sup> Ou Regressão Ponderada Geograficamente - desenvolvido por Fotheringham, Charlton e Brunson (1997, 2002).

- **Kernel *Bi-square*:** (adaptável)  $w_{ij} = \left(1 - (d_{ij}^2 / h^2)\right)^2$ , se  $g_j$  é um dos  $k$  vizinhos mais próximos de  $g_i$  e  $h$  é a distância do  $k$ -ésimo vizinho mais próximo.  
 $w_{ij} = 0$ , caso contrário.

## 4. Resultados

### 4.1. Seleção de Variáveis

Diante diversas variáveis independentes para análise é interessante utilizar o método *stepwise* para excluir variáveis insignificantes para o modelo pretendido. Tal método selecionou as seguintes variáveis independentes como significantes (nível de significância 10%) para responder a renda domiciliar nos inscritos no Cadastro Único: consumo de energia elétrica dos beneficiários do TSEE e o IDEB dos municípios - índice municipal calculado com base nas taxas de aprovação escolares e no desempenho dos alunos em avaliações nacionais aplicados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), o qual, para além de auferir a qualidade da educação básica brasileira, serve de parâmetro para os repasses federais a estados e municípios na área da educação.

O primeiro modelo elaborado foi a partir da Regressão Linear Multivariada (OLS) não espacial, o qual evidenciou que o consumo de energia elétrica e IDEB são significantes para explicar a renda em análise tendo em vista seus valores-P de, respectivamente, 0,004 e 0,031 – menores que o nível de significância adotado - vide anexo A. O  $R^2$  foi baixo, 2,02%, e a equação encontrada foi:

$$\mathbf{Renda = 763 - 1,301 Energia + 34,1 IDEB}$$

o que significa que: para o consumo de energia e o IDEB igual a zero, a renda domiciliar é R\$ 763; para cada R\$ 1 a mais na renda domiciliar, o consumo de energia elétrica cai 1,3 kWh; para cada R\$ 1 a mais na renda domiciliar, o IDEB do município aumenta 34,1 unidades de medida.

Cabe salientar que era esperado que o parâmetro de Consumo de Energia Elétrica fosse positivo – razões que serão discutidas no Capítulo 5. Já o parâmetro positivo do IDEB já era esperado, pois é inegável que a renda impacta positivamente o desempenho escolar – logo, quanto maior a renda domiciliar dos inscritos no CadÚnico, maior o IDEB municipal<sup>5</sup>.

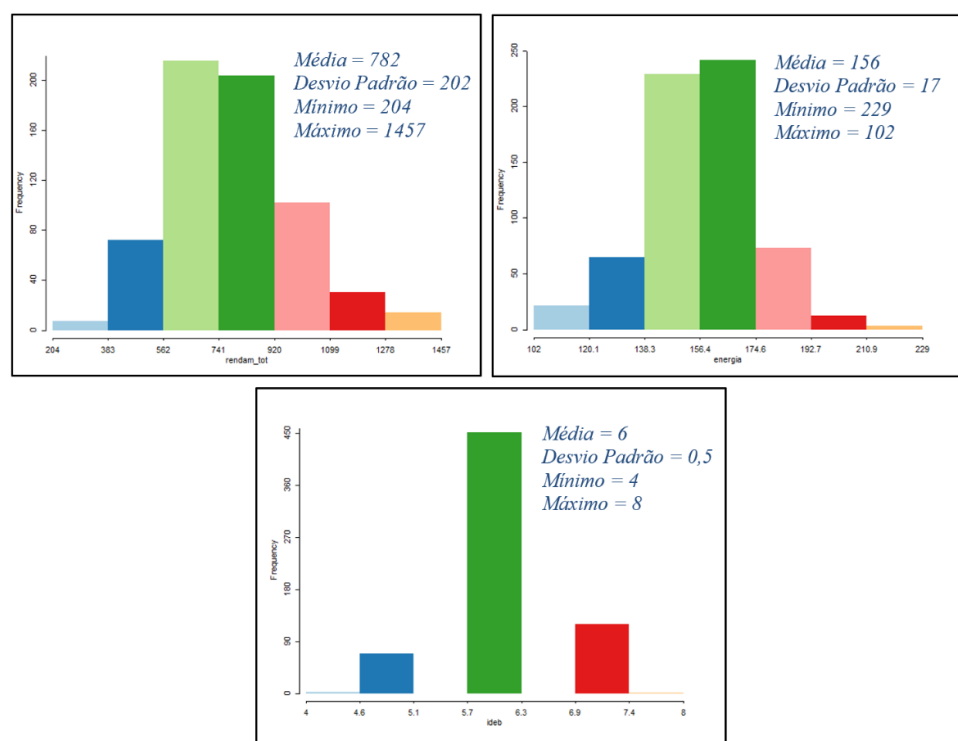
---

<sup>5</sup> Nesse ambiente cabe salientar estudos que evidenciam o impacto negativo da pobreza no IDEB. Por meio de regressões estatísticas, provou-se que a presença de alunos em situação de pobreza possui efeito negativo significativo no IDEB da escola (DUARTE, 2013).

## 4.2. Análise exploratória geográfica dos dados

O primeiro passo após o método *stepwise*, que selecionou as variáveis energia e IDEB, foi a análise exploratória – técnicas estatísticas que possibilitam resumir as características principais dos dados, frequentemente a partir de métodos visuais. Elaborou-se um histograma para análise descritivas das variáveis e, no âmbito da Estatística Espacial, foram adotados os métodos: *Natural Breaks*, *Box Map*, *LISA Map* e o diagrama de dispersão de Moran para representar as variáveis no mapa, bem como padrões espaciais em suas distribuições.

Os histogramas do Gráfico 3 evidenciam as distribuições das variáveis supracitadas e suas principais descrições estatísticas. Com 645 observações (todos os municípios do Estado de São Paulo), as distribuições de Renda e Consumo de Energia Elétrica mostraram-se assimétricas à direita, ao passo que a distribuição do IDEB apresentou uma concentração muito grande de observações em torno da média da distribuição (453 observações). A renda domiciliar média dos inscritos no CadÚnico é R\$ 782 - em termos de comparação, a renda domiciliar per capita média do Estado de São Paulo, em 2019, foi de R\$ 1.945,73, segundo o IBGE. O consumo de energia elétrica médio de uma família beneficiária do TSEE é 156 kWh, correspondente à faixa de 10% de desconto na conta de luz – sendo que uma geladeira antiga, em média, consome 150 kWh de consumo por mês<sup>6</sup>. A média do IDEB paulista é 6 – igual ao valor da meta nacional estabelecida até 2022.

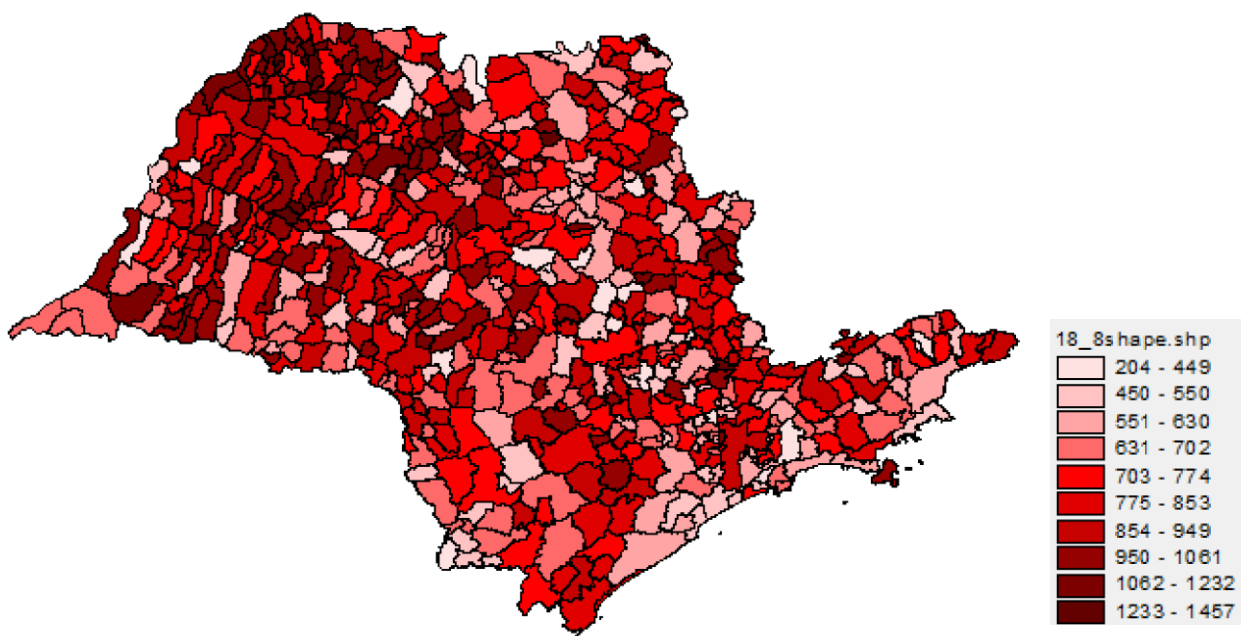


**Gráfico 3:** Histogramas de: 1) Renda Média Mensal Domiciliar (2018), 2) Consumo de Energia Elétrica Médio (2019), 3) IDEB (2018).  
**Fonte:** Elaboração própria com uso do software GeoDA  
**Nota:**  $n = 645$

<sup>6</sup> [http://www2.aneel.gov.br/arquivos/PDF/17-05\\_material\\_3.pdf](http://www2.aneel.gov.br/arquivos/PDF/17-05_material_3.pdf)

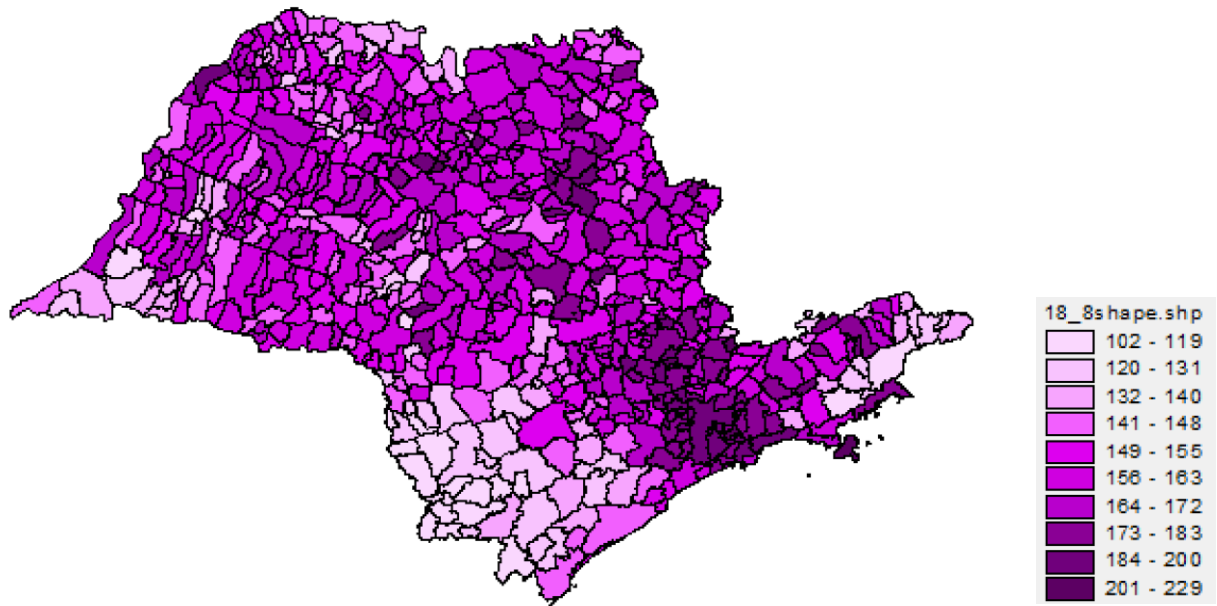
Na perspectiva espacial, o Natural Breaks é um mapa temático que utiliza o método do algoritmo de Jenks (Natural Breaks) (1967), método recomendado pela literatura para ser o primeiro utilizado para a representação de uma variável em nível espacial. Ele se fundamenta na técnica de maximizar a variância interclasses utilizando uma base matemática para determinar o valor de seus intervalos. Entre o maior e menor valor da distribuição, identificam-se as maiores diferenças entre um valor e o outro para assim determinar os intervalos. Para representar renda e consumo de energia as observações são divididas em 10 classes representadas numa escala de cores intuitivas (a cor clara significa menor renda/consumo de energia e a escura, maior renda/consumo de energia). A representação do IDEB, por sua vez, teve suas observações divididas em 5 classes. Naturalmente, dado o menor intervalo de dados sob análise, o número de classes deste também o foi.

Na Figura 1 é possível identificar que a Renda Média Domiciliar (Cadastro Único) por município apresenta uma concentração bem maior no norte do Estado, bem como os melhores resultados do IDEB – Figura 3. O Consumo de Energia Elétrica Residencial, por sua vez, é maior na região do entorno da capital de São Paulo e no litoral do Estado, como visto na Figura 2.

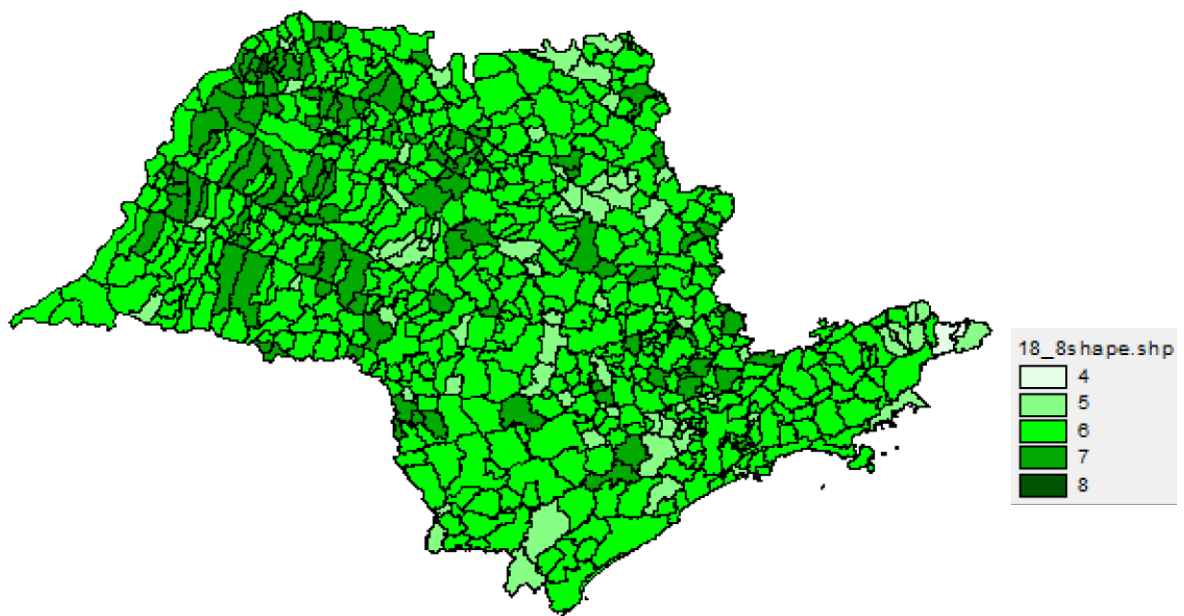


**Figura 3:** Mapa de Renda Domiciliar (Cadastro Único) nos municípios paulistas – 10 classes (natural breaks)

**Fonte:** Elaboração própria com uso do software ArcView.

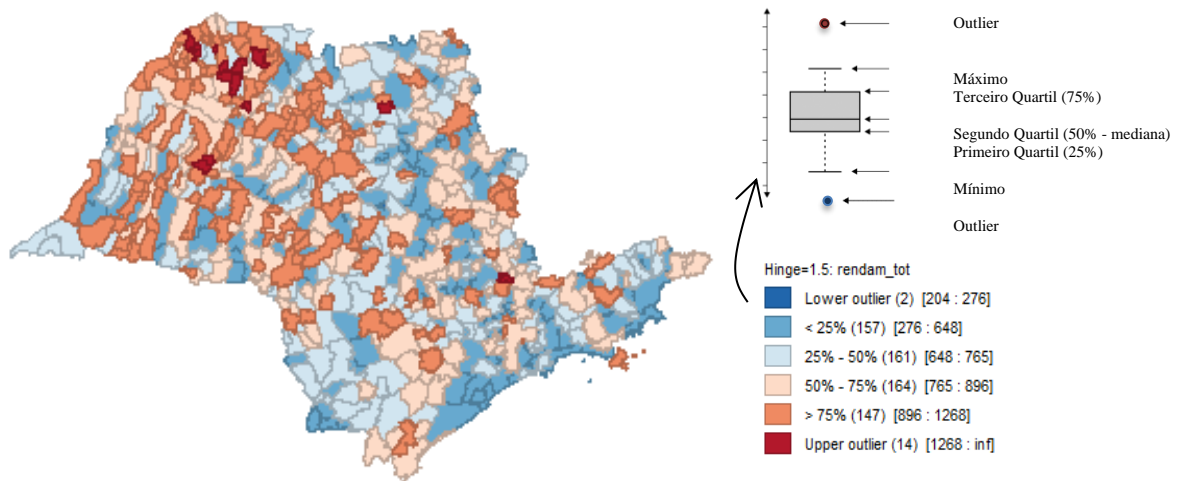


**Figura 4:** Mapa de Consumo de Energia Elétrica médio dos beneficiários da TSEE nos municípios paulistas – 10 classes (natural breaks).  
**Fonte:** Elaboração própria com uso do software Arcview.



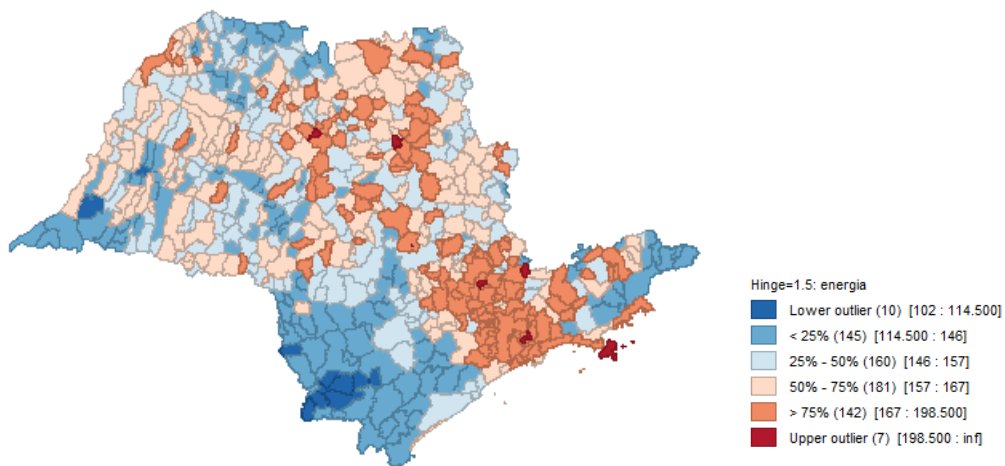
**Figura 5:** Mapa do Índice de Desenvolvimento da Educação Básica (IDEB) dos municípios paulistas – 10 classes (natural breaks).  
**Fonte:** Elaboração própria com uso do software Arcview.

Com vistas a identificar outliers e padrões espaciais amplos no conjunto de dados, utiliza-se o *Box Map* – mapa das variáveis baseado nos limites do *boxplot* para identificar concentração espacial de municípios de alta ou baixa renda/consumo/IDEB. A Figura 5 apresenta-se dividida, predominantemente, em dois lados: o esquerdo com valores maiores (tons de vermelho) e o direito, menores (tons de azul). A Figura 5 apresenta concentração maior de outliers a sudoeste do Estado em tons de azul) e ao leste, centro-sul e litoral em tons de vermelho.



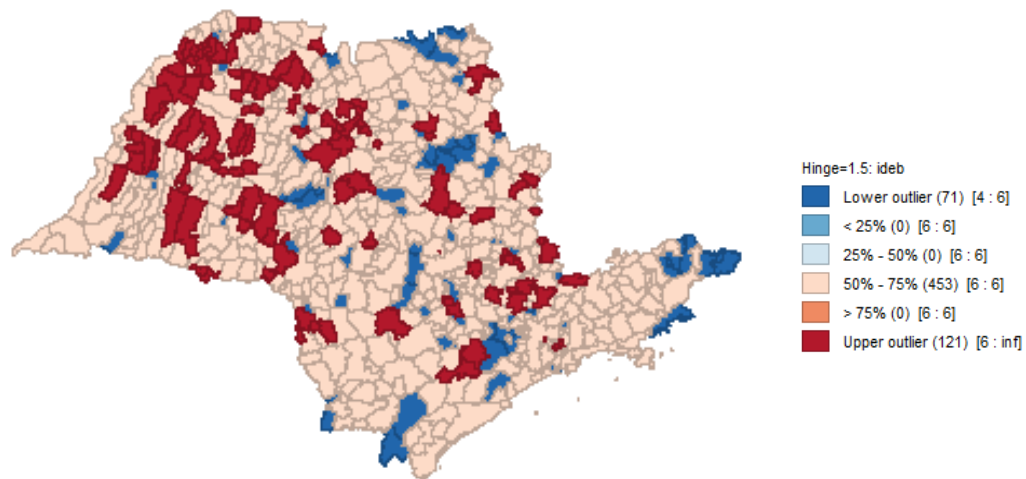
**Figura 6:** Box Map (hinge 1,5) da variável Renda Domiciliar (Cadastro Único) nos municípios paulistas

**Fonte:** Elaboração própria com uso do software GeoDa.



**Figura 7:** Box Map (hinge 1,5) da variável Consumo de Energia Residencial médio (TSEE).

**Fonte:** Elaboração própria com uso do software GeoDa.



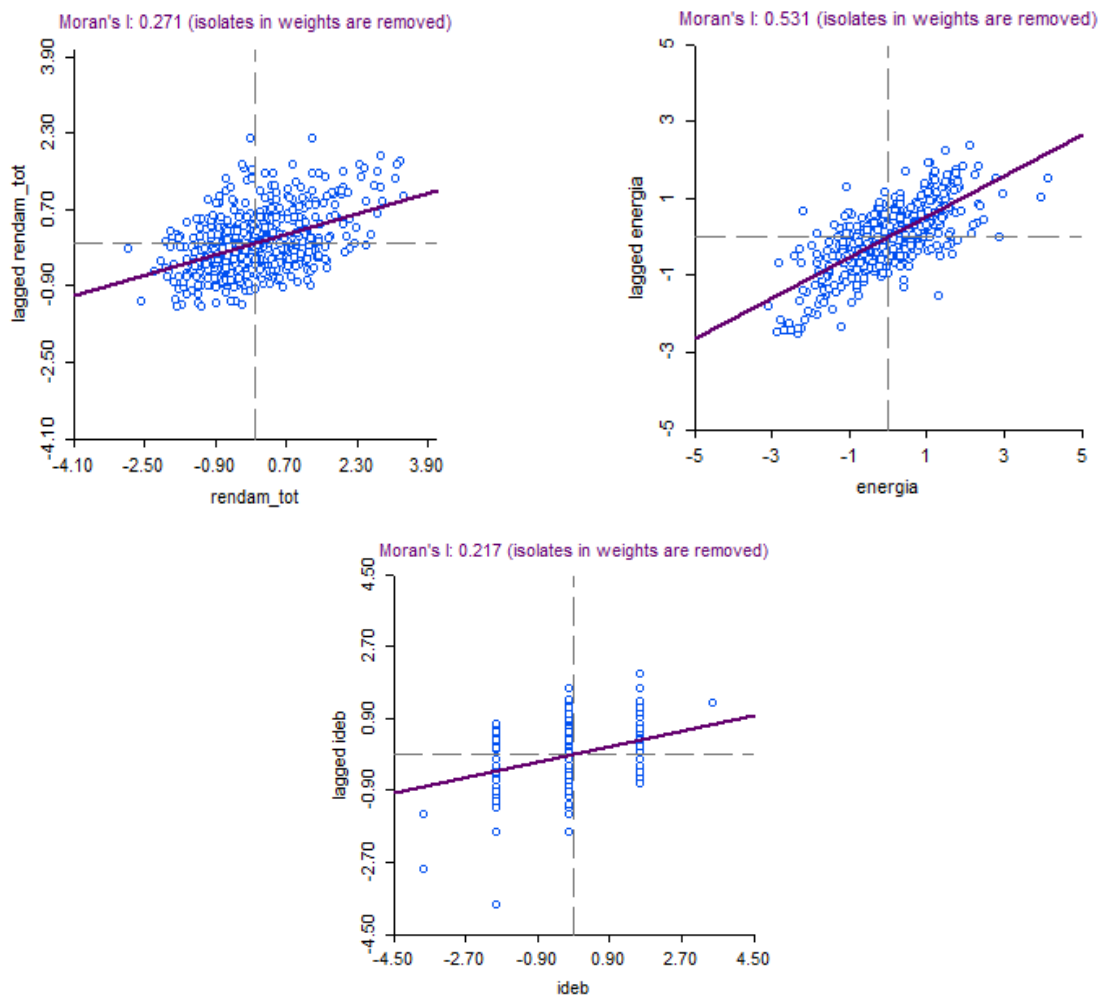
**Figura 8:** Box Map (hinge 1,5) da variável IDEB dos municípios paulistas

**Fonte:** Elaboração própria com uso do software GeoDa.

Para aprofundar a dependência espacial das variáveis, foi utilizado o indicador de auto correlação espacial I de Moran para avaliar as propriedades de vizinhanças presentes. Busca-se investigar o grau de dependência entre os municípios analisados, ou seja, se a mesma vizinhança apresenta características parecidas. Dentre outros tipos, foi escolhido realizar a matriz de vizinhança, por contiguidade, do tipo Queen, de ordem 1 para a pesquisa em questão – sendo considerado vizinho aqueles que são fisicamente adjacentes, aonde todas as áreas que possuem inteseccção não-nula e direta com a área em análise são consideradas vizinhas. Encontrou-se que, em média, cada município possui 5,7 vizinhos (sendo o máximo 23 e o mínimo 0).

O Índice I de Moran, no contexto deste trabalho, é a correlação entre a variável renda (eixo x) e a variável renda média dos vizinhos (eixo y) – idem para energia e IDEB. O Gráfico 3 e 5 mostram a dependência espacial da variável Renda e IDEB, sendo seu I de Moran, respectivamente, igual 0,27 e 0,22 – considerado uma relação baixa (mas não desprezível) entre a variável x e y. Além disso, os valores expressos no gráfico não se assemelham a uma reta, característica de uma correlação e dependência espacial alta. A variável consumo de energia, por sua vez, apresentou o maior I de Moran, de 0,53, considerado uma relação forte entre variável consumo de energia (eixo x) e a variável consumo de energia médio dos vizinhos (eixo y).



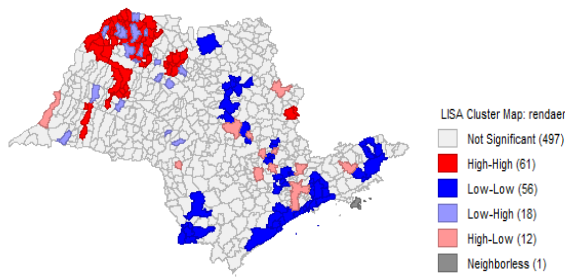


**Gráfico 4:** Gráfico de dispersão de Moran para a variável 1) Renda Domiciliar Média do Cadastro Único (2017) dos municípios do Estado de São Paulo, 2) Consumo Médio de Energia Elétrica Residencial (TSEE) e 3) IDEB dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.

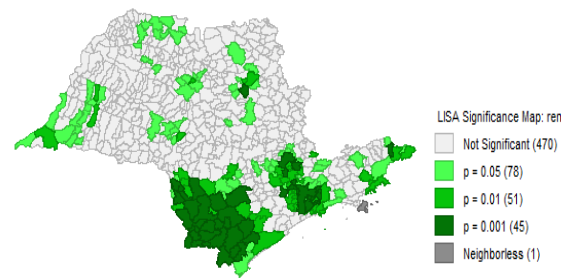
Em análise aos gráficos de dispersão acima, percebe-se que este é dividido em quatro quadrantes: renda média alta e vizinhos com renda média alta (*high-high*); renda média alta e vizinhos com renda média baixa (*high-low*); renda média baixa com vizinhos com renda média baixa (*low-low*) e renda média baixa com vizinhos com renda média alta (*low-high*) – idem para energia e IDEB. As Figuras 8, 10 e 12 – que configuram um mapa do indicador de auto correlação espacial local (LISA), apresentam os municípios e seus vizinhos a partir das cores vermelho escuro (*high-high*), vermelho claro (*high-low*), azul escuro (*low-low*), azul claro (*low-high*) e branco (municípios que não apresentaram significativa relação com sua vizinhança e que, portanto, podem facilmente mudar de quadrante). As Figura 8 mostra que a autocorrelação espacial da Renda Domiciliar Média do Cadastro Único é bem distribuída pelo estado de São Paulo; a Figura 10 apresenta um padrão de concentração maior das cores vermelho e azul, isto é, *high-high* e *low-low*, principalmente na região sudoeste, sul, leste e no entorno da capital; a Figura 12, por sua vez, apresenta, predominantemente, a concentração de *low-high*.

Os mapas apresentados em tons de verde, por sua vez, evidenciam a significância estatística de cada indicador local calculado e mapeado (Figuras 9, 11 e 13).



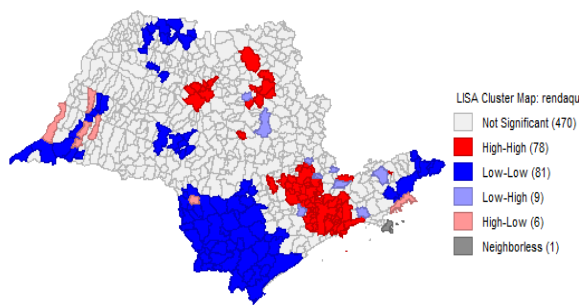
**Figura 9:** Mapa do indicador de auto correlação espacial local (*LISA Cluster Map*) da variável Renda Domiciliar da amostra populacional do Cadastro Único dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.



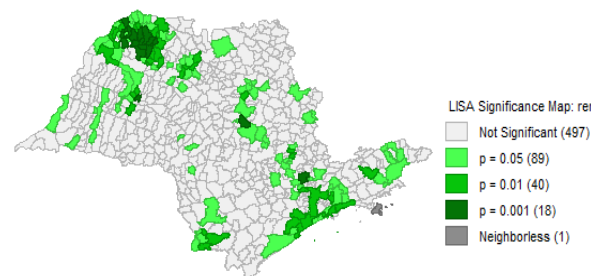
**Figura 10:** Mapa do indicador de auto correlação espacial local (*LISA Significance Map*) da variável Renda Domiciliar da amostra populacional do Cadastro Único dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.



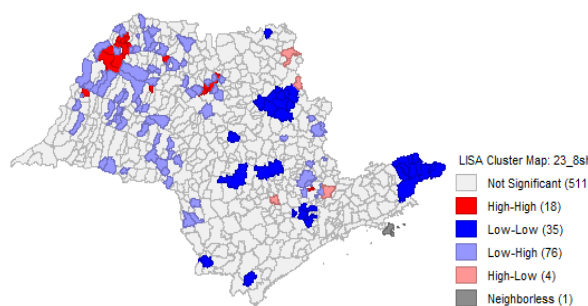
**Figura 11:** Mapa do indicador de auto correlação espacial local (*LISA Cluster Map*) da variável Consumo Médio de Energia Elétrica Residencial dos beneficiários da TSEE dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.



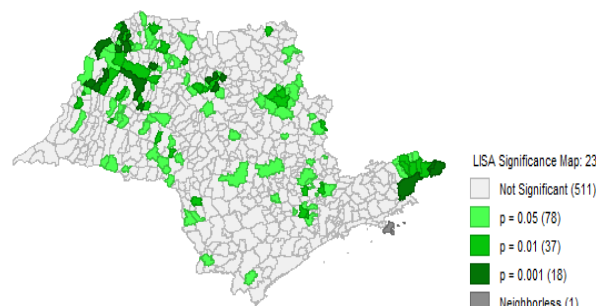
**Figura 12:** Mapa do indicador de auto correlação espacial local (*LISA Significance Map*) da variável Consumo Médio de Energia Elétrica Residencial dos beneficiários da TSEE dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.



**Figura 13:** Mapa do indicador de auto correlação espacial local (*LISA Cluster Map*) do IDEB dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.



**Figura 14:** Mapa do indicador de auto correlação espacial local (*LISA Significance Map*) do IDEB dos municípios paulistas.

**Fonte:** Elaboração própria com uso do software GeoDa.

### 4.3. Regressões Espaciais

#### 4.3.1. *Spatial Autoregression (SAR)*

*Spatial Autoregression (SAR)* é o modelo de regressão espacial global. As análises oriundas das expressões matemáticas indicadas nas metodologias do Capítulo 3 evidenciaram que o consumo de energia elétrica e o IDEB são capazes de explicar a renda familiar. Isto porque o que os valores-P encontrados, de 0,03 para energia e 0,00 para IDEB, são menores que 10% (valor do nível de significância adotado) - vide anexo B. Logo, não se rejeita a hipótese nula.

#### **SAR Matriz Adjacência 1ª ordem da Renda a partir do Consumo de Energia e IDEB**

Variável	Coefficiente	Erro Padrão	<i>z value</i>	Probabilidade
Intercepto	$\beta_0 = 242,50$	105,58	2,30	0,02
Energia	$\beta_1 = -0,88$	0,41	-2,15	0,03
IDEB	$\beta_2 = 58,04$	13,06	4,44	0,00

**Quadro 1:** Resultados do Modelo SAR com matriz de adjacência Queen de 1ª ordem da Renda a partir do Consumo de Energia Elétrica e do IDEB no Estado de São Paulo.

**Fonte:** Elaboração própria com uso do software GeoDa.

No que diz respeito a análise do coeficiente de determinação  $R^2$  (medida de qualidade de ajuste do modelo), o modelo espacial melhorou o modelo linear, tendo em vista que o coeficiente passou de 2,02% para 18,38% - sendo que quanto mais próximo de 1, melhor é o ajuste do modelo. A expressão da regressão SAR é:

$$\text{Renda} = 242 - 0,88 \text{ Energia} + 58 \text{ IDEB} + 0,417 \text{ WRenda}$$

Cabe salientar que, em comparação ao modelo não espacial, os parâmetros das variáveis independentes consumo de energia e IDEB perderam valor. Agora significam que: para o consumo de energia e o IDEB igual a zero, a renda domiciliar é R\$ 242; para cada R\$ 1 a mais na renda domiciliar, o consumo de energia elétrica cai 0,88 kWh; para cada R\$ 1 a mais na renda domiciliar, o IDEB do município aumenta 58 unidades de medida.

#### 4.3.2. *Geographically Weighted Regression (GWR)*

Como dito anteriormente, o método *Geographically Weighted Regression (GWR)*, diferentemente do SAR, elabora modelos locais para cada ponto espacial observado. Desse modo, não há apenas uma expressão para relacionar as variáveis, mas 645. Para calcular o GWR utilizou-se kernels adaptável do tipo *bi-square* - vide anexo C.

Os parâmetros de energia e IDEB podem assumir os seguintes valores:

## GWR com *kernel bi-square* da Renda a partir do Consumo de Energia e do IDEB

Variável	Mínimo	1º Quartil	Mediana	3º Quartil	Máximo
Intercepto ( $\beta_0$ )	-16,66	497,46	737,30	1036,30	2185,54
Energia ( $\beta_1$ )	-6,53	-2,68	-0,97	0,43	3,44
IDEB ( $\beta_2$ )	-78,59	8,17	38,96	52,27	121,46

**Quadro 2:** Resultados da Regressão GWR com *kernel bi-square* da Renda a partir do Consumo de Energia Elétrica e do IDEB no Estado de São Paulo.

**Fonte:** Elaboração própria com uso do software R Studio.

Ademais, o coeficiente de determinação  $R^2$  subiu para 30,14% – melhorando significativamente o modelo.

### 5. Conclusões

A presente pesquisa teve como principal propósito examinar a relação entre Consumo de Energia Elétrica dos beneficiários do TSEE e a Renda Familiar dos inscritos no Cadastro Único nos municípios do Estado de São Paulo. Investigou-se a adequação da variável consumo de energia da população baixa renda como fundamento para a construção de um indicador que possibilite a predição da Renda Domiciliar. Isto com vistas a contribuir para uma classificação econômica mais precisa do público que se serve desta categorização para o acesso a programas sociais de beneficiamento e superação das desigualdades, assim como fornecer suporte à gestão e proposição de políticas públicas e programas assistenciais.

Para isso, analisou-se os dados sob a perspectiva da estatística espacial com o intuito de encontrar padrões na distribuição geográfica das variáveis selecionadas, bem como melhorar o poder de predição e explicação dos resultados encontrados pela regressão linear (OLS). Foram utilizados os métodos *Spatial Auto-Regressive models* (SAR) e *Geographically Weighted Regression* (GWR).

Após os resultados obtidos, a pesquisa concluiu que:

- I. O modelo *stepwise* selecionou, dentre diversas variáveis selecionadas que corroboram com a variável dependente, o Consumo de Energia Elétrica dos beneficiários da TSEE e o IDEB.
- II. Dentre as variáveis analisadas, a que apresentou autocorrelação espacial mais significativa foi o Consumo de Energia Elétrica dos beneficiários da TSEE –  $I$  de Moran de cerca de 53,1% para Consumo de Energia, 27,1% para Consumo de Energia e 21,7% para IDEB.
- III. A Regressão Multivariada Linear (OLS) evidenciou que o Consumo de Energia Elétrica dos beneficiários da TSEE possui um poder de explicação baixo para a Renda Familiar dos cadastrados no Cadastro Único – o coeficiente de determinação foi de 2,02%.
- IV. O *Spatial Auto-Regressive models* (SAR) evidenciou que o Consumo de Energia Elétrica dos beneficiários da TSEE possui um poder de explicação maior para a Renda Familiar dos cadastrados no Cadastro Único – o coeficiente de determinação foi de 18,38%.

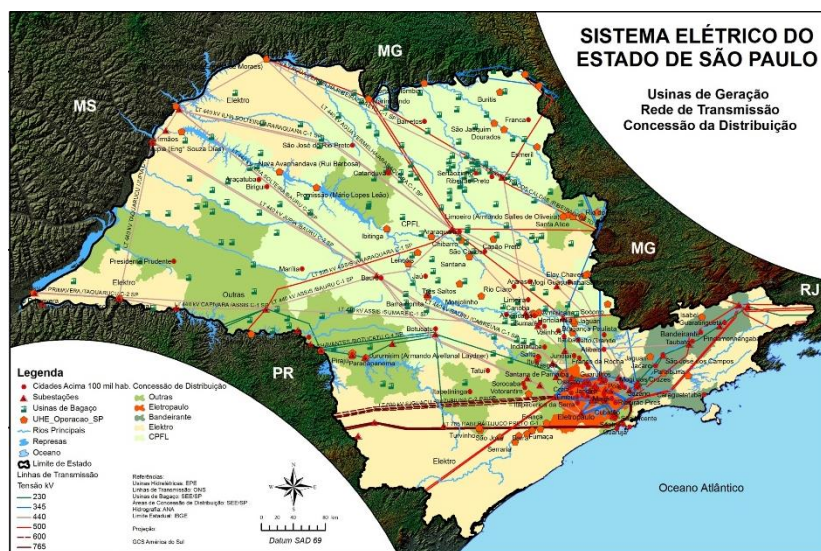
- V. Foi apresentada muitas vantagens com a adoção de modelos de dependência espacial na presente análise. O modelo local GWR melhorou significativamente o modelo global SAR. O coeficiente de determinação passou de 18,38% para 30,14%.
- VI. Os parâmetros do Consumo de Energia Elétrica dos beneficiários da TSEE na regressão linear (OLS) e no modelo SAR mostraram-se negativos – sendo iguais a - 1,30 e -0,88, respectivamente.

**Tabela 5:** Resultados encontrados

Modelo	OLS $y = X\beta + \varepsilon$	SAR $y(g) = \beta_0(g) + \beta_1(g)x + \varepsilon$	GWR - Kerney adaptável bi-square $y(g) = \beta_0(g) + \beta_1(g)x_1 + \beta_2(g)x_2 + \dots + \beta_p(g)x_p + \varepsilon$					
Parâmetros	Renda = 763 - 1,301 Energia + 34,1 IDEB	Renda = 242 - 0,88 Energia + 58 IDEB + 0,417 W Renda	Intercepto:	Min -16,66	1Q 497,46	MED 737,30	3Q 1036,30	Máx 2185,54
			Energia:	-6,53	-2,68	-0,97	0,43	3,44
			IDEB:	-78,59	8,17	38,96	52,27	121,46
R <sup>2</sup>	2,02%	18,38%	30,14%					

Inicialmente, esperava-se encontrar uma relação positiva entre renda domiciliar e o consumo de energia elétrica (quanto maior o consumo, maior a renda). Entretanto, o parâmetro da variável encontrado foi negativo. Este fato permite questionar o quanto a TSEE diz respeito a condição econômica da família que usufrui do benefício social e o quanto outros fatores, tal como a distribuição de infraestrutura e acesso a bens de consumo com maior eficiência energética, podem também impactar esta relação.

Nesse sentido, é interessante notar que o Mapa de Consumo de Energia Elétrica médio dos beneficiários da TSEE nos municípios paulistas (Figura 3) é visualmente similar ao Mapa do Sistema Elétrico do Estado de São Paulo (Figura 14). Isso reforça o entendimento mencionado acima, de que fatores como a existência e concentração de infraestrutura podem impactar os dados referentes ao consumo de energia elétrica. Tal fato pode ser representado, considerando, por exemplo, que as famílias beneficiárias que mais consomem energia elétrica residem na região paulista que mais possui infraestrutura de energia elétrica.



**Figura 15:** Mapa do Sistema Elétrico do Estado de São Paulo  
**Fonte:** Secretaria de Infraestrutura e Meio Ambiente do Estado de São Paulo.

Dentre os questionamentos mencionados acima, tal situação nos leva a refletir sobre a condição do uso da energia elétrica no que tange seu consumo inteligente e seguro por parte da população com menor poder aquisitivo. Cabe salientar que o consumo de energia elétrica médio de uma família beneficiária do TSEE é 156 kWh – sendo que uma geladeira antiga, em média, consome 150 kWh de consumo por mês. Assim, eletrodomésticos menos eficazes e mais antigos, oscilações nas redes de energia e instalações menos robustas interferem significativamente na conta de luz das famílias.

Sendo assim, tendo em vista a relação negativa entre consumo de energia elétrica e renda domiciliar da população baixa renda, o presente trabalho se posiciona no sentido de fornecer as bases para que estudos futuros possam avançar na compreensão da relação encontrada. Outra potencial informação encontrada para fundamentar estudos futuros é a similaridade encontrada entre a existência e concentração de infraestrutura de energia elétrica e o consumo por parte da população.

Desse modo, as informações encontradas – principalmente sob o ponto de vista espacial, podem servir como insumo para as empresas de distribuição de energia elétrica, bem como sua agência reguladora, fazerem melhor uso de suas informações para finalidades distintas das tradicionais aplicadas em sua gestão operacional. Isso sem contar que esta análise permite ao gestor público entender a dinâmica do território o qual suas ações se estabelecem e, assim, melhor orientar as políticas públicas.

## Referências

FRANCISCO, E. R. **Indicadores de Renda baseados em Consumo de Energia Elétrica: Abordagens Domiciliar e Regional na Perspectiva da Estatística Espacial**. Tese (Doutorado) - Escola de Administração de Empresas de São Paulo, 2010.

FRANCISCO, E. R. **Relação entre o Consumo de Energia Elétrica, a Renda e a Caracterização Econômica de Famílias de Baixa Renda no Município de São Paulo**. Tese (Mestrado) - Escola de Administração de Empresas de São Paulo, 2006.

BUSSAB, W. O.; FERREIRA, M. **Critério Brasil de Estratificação Socioeconômica: Aspectos Demográficos**. Anais do CLADEA, San Juan, Porto Rico, 34. 1999.

ABEP - Associação Brasileira de Empresas de Pesquisa. **Critério de Classificação Econômica - Base 2019**. Disponível em: <<http://www.abep.org/>>. Acesso em: 1 julho 2020.

KAMAKURA, Wagner; MAZZON, José Afonso. **Critérios de estratificação e comparação de classificadores socioeconômicos no Brasil**. RAE-Revista de Administração de Empresas, [S.l.], v. 56, n. 1, p. 55-70, jan. 2016. ISSN 2178-938X. Disponível em: <<http://bibliotecadigital.fgv.br/ojs/index.php/rae/article/view/59351>>. Acesso em: 10 Mai. 2019.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Nacional por Amostra de Domicílios Contínua: características gerais dos domicílios e dos moradores 2019**. Rio de Janeiro: IBGE, 2020.

MATTAR, Fauze N. **Porque os Métodos de Classificação Socioeconômicos utilizados no Brasil não funcionam**. In: ENANPAD, 20., 1996, Angra dos Reis.

SILVA, Ney L. **Critério Brasil: o mercado falando a mesma língua**. Pesquisa em foco – Informativo dos Associados da ANEP, dez. 2002.

CORREIA-SILVA, D.C.; SIMOES, J. E. M.; OLIVEIRA, C. D. C.. **Relação entre Desempenho Econômico e Consumo de Eletricidade no Brasil**. REFLEXÕES ECONÔMICAS, v. 2, p. 98-118-118, 2017.

BRASIL. SECRETARIA DE AVALIAÇÃO E GESTÃO DA INFORMAÇÃO. **Pesquisa de Qualidade do Cadastro Único para Programas Sociais do Governo Federal**. Disponível em: Acesso em: 15 dez. 2019.

LUPION, Bruno. **Mais 170 mil brasileiros entraram para a pobreza extrema em 2019**. 2020. Disponível em: <https://economia.uol.com.br/noticias/redacao/2020/05/18/mais-170-mil-brasileiros-entraram-para-a-pobreza-extrema-em-2019.htm>. Acesso em: 15 ago. 2020

**CLIMA e barreiras sanitárias prejudicaram PIB da agropecuária em 2012**. 2013. Disponível em: <http://g1.globo.com/economia/agronegocios/noticia/2013/03/clima-e-barreiras-sanitarias-prejudicaram-pib-da-agropecuaria-em->

2012.html#:~:text=Rio%2C%2001%20%2D%20Os%20problemas%20clim%C3%A1ticos,2%2C3%25%20no%20ano.. Acesso em: 15 ago. 2020.

Borger, Fernanda Gabriela, et al. **"Inovação social e sustentabilidade: consumo de energia elétrica em comunidades carentes no Brasil."** Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental 19.1 (2015): 71-79.

DUARTE, Natalia de Souza. **O impacto da pobreza no Ideb: um estudo multinível.** Revista Brasileira de Estudos Pedagógicos, Brasília, v. 94, n. 237, p. 343-363. Disponível em: <https://www.scielo.br/pdf/rbeped/v94n237/a02v94n237.pdf>. Acesso em: 15 agosto 2020.

PRAHALAD, C, K; HART, Stuart. **The Fortune at the Bottom of the Pyramid.** Strategy + Business, n. 26, First Quarter 2002.

JENKS, George F. 1967. **"The Data Model Concept in Statistical Mapping"**, International Yearbook of Cartography 7: 186–190.

CÂMARA, Gilberto et al. **Tutorial Análise Geográfica.** In: CONGRESO GEOBRASIL 2003, 2003, São Paulo

FOTHERINGHAM, A. Stewart; BRUNSDON, Chris; CHARLTON, Martin E. **Geographically Weighted Regression – the analysis of spatially varying relationships.** West Sussex, England: John Wiley & Sons, 2002.



## Anexos

### Anexo A: OLS da Renda a partir do Consumo de Energia Elétrica e IDEB

#### Regression Equation

rendam\_tot = 763 + 34,1 IDEB  
- 1,301 energia

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	763	125	6,12	0,000	
IDEB	34,1	15,8	2,16	0,031	1,00
energia	-1,301	0,444	-2,93	0,004	1,00

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
200,224	2,02%	1,72%	1,06%

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	531635	265817	6,63	0,001
IDEB	1	187606	187606	4,68	0,031
energia	1	343953	343953	8,58	0,004
Error	642	25737453	40089		
Lack-of-Fit	485	19818467	40863	1,08	0,276
Pure Error	157	5918986	37701		
Total	644	26269088			

### Anexo B: SAR Matriz Adjacência 1ª ordem da Renda a partir do Consumo de Energia e IDEB

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set      : 18_8shape
Spatial Weight : 23_8shape
Dependent Variable : rendam_tot   Number of Observations: 645
Mean dependent var : 782.099     Number of Variables : 4
S.D. dependent var : 201.81      Degrees of Freedom : 641
Lag coeff. (Rho) : 0.416766

R-squared      : 0.183765   Log likelihood : -4284.1
Sq. Correlation : -        Akaike info criterion : 8576.2
Sigma-square   : 33243     Schwarz criterion : 8594.08
S.E of regression : 182.327
```

Variable	Coefficient	Std. Error	z-value	Probability
W_rendam_tot	0.416766	0.0488753	8.52712	0.00000
CONSTANT	242.503	105.576	2.29695	0.02162
energia	-0.879231	0.408297	-2.15341	0.03129
ideb	58.0458	13.0641	4.44314	0.00001

#### REGRESSION DIAGNOSTICS

##### DIAGNOSTICS FOR HETEROSKEDASTICITY

##### RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	2	2.8434	0.24130

##### DIAGNOSTICS FOR SPATIAL DEPENDENCE

##### SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : 23\_8shape

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	66.9866	0.00000

===== END OF REPORT =====

## Anexo C: GWR com *kernel* bi-square da Renda a partir do Consumo de Energia e do IDEB

```
*****
*           Results of Geographically Weighted Regression           *
*****

*****Model calibration information*****
Kernel function: bisquare
Adaptive bandwidth: 116 (number of nearest neighbours)
Regression points: the same locations as observations are used.
Distance metric: Euclidean distance metric is used.

*****Summary of GWR coefficient estimates:*****
           Min.      1st Qu.      Median      3rd Qu.      Max.
Intercept  -16.65840  497.46633  737.30583 1036.14827 2185.5387
sp2$energia  -6.53379   -2.68167   -0.96591   0.43433   3.4391
sp2$IDEB    -78.59308   8.17265   38.95671   52.27297  121.4564
*****Diagnostic information*****
Number of data points: 645
Effective number of parameters (2trace(s) - trace(s's)): 58.02744
Effective degrees of freedom (n-2trace(s) + trace(s's)): 586.9726
AICC (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 8540.338
AIC (GWR book, Fotheringham, et al. 2002, GWR p. 96, eq. 4.22): 8488.606
Residual sum of squares: 18350692
R-square value: 0.301434
Adjusted R-square value: 0.2322567

*****
```