INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DE GOIÁS
CAMPUS GOIÂNIA
DEPARTAMENTO ACADÊMICO DA ÁREA III
COORDENAÇÃO DE ENGENHARIA CARTOGRÁFICA E DE AGRIMENSURA

POLLIANNA VIEIRA DA SILVA

TRABALHO DE CONCLUSÃO DE CURSO

ANÁLISE ESPACIAL DO ÍNDICE DE DESENVOLVIMENTO DA EDUCAÇÃO BÁSICA (IDEB) DAS ESCOLAS MUNICIPAIS DE APARECIDA DE GOIÂNIA-GO

GOIÂNIA, GO
2019

POLLIANNA VIEIRA DA SILVA

ANÁLISE ESPACIAL DO ÍNDICE DE DESENVOLVIMENTO DA EDUCAÇÃO BÁSICA (IDEB) DAS ESCOLAS MUNICIPAIS DE APARECIDA DE GOIÂNIA-GO

Trabalho de Conclusão de Curso (TCC) apresentado à Coordenação de Geomática, do Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Campus Goiânia, como parte dos requisitos necessários para a obtenção do título de Engenharia Cartográfica e de Agrimensura.

Orientador: Prof. Dr. Giovanni de Araujo Boggione

GOIÂNIA, GO
2019

Si381a Silva, Pollianna Vieira da.
Análise espacial do índice de desenvolvimento da Educação Básica (IDEB) das escolas municipais de Aparecida de Goiânia-GO / Pollianna Vieira da Silva. – Goiânia: Instituto Federal de Educação, Ciência e Tecnologia de Goiás, 2020.
71 f.: il.

Orientador: Dr. Giovanni de Araujo Boggione.

TCC (Trabalho de Conclusão de Curso) – Curso Superior de Engenharia Cartográfica e de Agrimensura, Departamento de Áreas Acadêmicas III, Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

1. Sistema de Informação Geográfica. 2. Análise Espacial. 3. Processamento Digital de Imagens. I. Boggione, Giovanni de Araujo (orientador). II. Instituto Federal de Educação, Ciência e Tecnologia de Goiás. III. Título.

CDD 526.982

Ficha catalográfica elaborada pela Bibliotecária Karol Almeida da Silva Abreu CRB1/ 2.740
Biblioteca Professor Jorge Félix de Souza,
Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia.

INSTITUTO FEDERAL
Goiás

MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
SISTEMA INTEGRADO DE BIBLIOTECAS

TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO NO REPOSITÓRIO DIGITAL DO IFG - ReDi IFG

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia de Goiás, a disponibilizar gratuitamente o documento no Repositório Digital (ReDi IFG), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IFG.

Identificação da Produção Técnico-Científica

[ ] Tese	[ ] Artigo Científico
[ ] Dissertação	[ ] Capítulo de Livro
[ ] Monografia – Especialização	[ ] Livro
[X] TCC - Graduação	[ ] Trabalho Apresentado em Evento
[ ] Produto Técnico e Educacional - Tipo: _____	

Nome Completo do Autor: Pollianna Vieira da Silva
Matrícula: 2015103113900 60
Título do Trabalho: Análise Espacial do Índice de Desenvolvimento da Educação Básica (IDEB) dos Escolas municipais de Aparecida de Goiânia - Go.

Autorização - Marque uma das opções

1. (X) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG;
2. ( ) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG somente após a data ___/___/_____;
3. ( ) Não autorizo disponibilizar meu trabalho no Repositório Digital do IFG.

Ao indicar a opção **2 ou 3**, marque a justificativa:

( ) O documento está sujeito a registro de patente.
( ) O documento pode vir a ser publicado como livro, capítulo de livro ou artigo.
( ) Outra justificativa: _____

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O/A referido/a autor/a declara que:

i. o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
ii. obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
iii. cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

Goiânia, 12/11/20.
Local Data

Pollianna Vieira da Silva
Assinatura do Autor e/ou Detentor dos Direitos Autorais

TERMO DE APROVAÇÃO

POLLIANNA VIEIRA DA SILVA

ANÁLISE ESPACIAL DO ÍNDICE DE DESENVOLVIMENTO DA EDUCAÇÃO BÁSICA (IDEB) DAS ESCOLAS MUNICIPAIS DE APARECIDA DE GOIÂNIA - GO

Trabalho de Conclusão de Curso apresentado à Coordenação de Geomática, do Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia, como parte dos requisitos necessários para a obtenção do título de Bacharelado em Engenharia Cartográfica e de Agrimensura.

Aprovado em, 19 de dezembro de 2019.

Prof. Dr. Giovanni Araújo Boggione (IFG)
(Orientador)

Prof. Dr. Fabio Campos Macedo (IFG)
(Membro da Banca Examinadora)

Prof. Dr. Nilton Ricetti Xavier de Nazareno (IFG)
(Membro da Banca Examinadora)

GOIÂNIA, GO
2019

DEDICATÓRIA

A minha família pelo incentivo de sempre e ao meu Primo Bruno (in memorian) por ter me presenteado com um motivo para seguir em frente.

AGRADECIMENTOS

A Deus por toda fé e força nesse ano difícil.

Aos meus pais e irmão pelo incentivo nos meus estudos.

A minha família no geral por acreditar em mim.

Ao meu orientador Profº Dr. Giovanni de Araujo Boggione pela dedicação, paciência, boa vontade e parceria.

A todos os meus amigos (as) da turma, pelo companheirismo no decorrer desses anos.

A Reizane e a Gabriella por não me deixarem desistir, e por sermos as meninas superpoderosas do time Boggione.

A professora e pesquisadora Drª Flávia Feitosa, pela disponibilidade e colaboração nas análises.

Ao meu amigo Leandro pelo suporte técnico diário.

A Secretaria de Planejamento e Regulação Urbana de Aparecida de Goiânia pela ajuda com os dados, especialmente a Janaina, Carolina, Cristina, Lorena, Raquel, Amanda e Lya.

Ao meu afilhado Bruno Filho que me motiva todos os dias a ver o lado bom da vida.

E ao IFG pela oportunidade de crescimento e conhecimento.

EPÍGRAFE

Recria tua vida, sempre, sempre.
Remove pedras e planta roseiras e faz doces. Recomeça”.
Cora Coralina

RESUMO

ANÁLISE ESPACIAL DO ÍNDICE DE DESENVOLVIMENTO DA EDUCAÇÃO BÁSICA (IDEB) DAS ESCOLAS MUNICIPAIS DE APARECIDA DE GOIÂNIA-G0.

AUTORA: POLLIANNA VIEIRA DA SILVA
ORIENTADOR: DR. GIOVANNI DE ARAUJO BOGGIONE

Aparecida de Goiânia-GO é um munícipio do Estado de Goiás que vem sendo reconhecido pelo seu crescimento industrial e populacional. O censo do ano de 2010 estimou uma população total de 455.657 pessoas e a estimativa para o ano de 2018 foi de 565.957 habitantes, um aumento significativo de 24%. Devido a isso, há demandas crescentes na área de serviços destinados à população, necessitando, portanto, de melhorias nos setores da infraestrutura da cidade. Dentro da demanda de infraestrutura educacional as escolas têm papel inquestionável. Todos os municípios têm uma grande demanda na área da educação e é um desafio atender a todos com um ensino de qualidade. Em 2007 foi criado o IDEB que é o índice de desenvolvimento da educação básica e tem a função de mensurar a qualidade da educação oferecida. É intenção desta pesquisa mostrar se existem relações espaciais entre o IDEB do Município de Aparecida de Goiânia e varáveis sócio econômicas e ambientais, utilizando ferramentas de Banco de Dados Geográficos, Sistema de Informação Geográfica, Mineração de Dados e Processamento Digital de Imagens. Foram criados: um banco de dados para consultas espaciais dentro de um ambiente SIG; um mapa de conforto ambiental utilizando técnicas de processamento digital de imagens; relações não espaciais utilizando mineração de dados; e análises exploratórias utilizando técnicas de análise de correlação espacial. O minerador APRIORI encontrou relações entre o IDEB, à renda e esgoto. As análises espaciais mostraram que as famílias com a maior renda estão agrupadas na região norte da cidade onde faz limite com a cidade de Goiânia. A análise exploratória para a abordagem utilizando os dados apenas dos Setores onde existem escolas, a Regressão Clássica foi a mais confiável. Já para a abordagem utilizando todos os Setores a regressão mais confiável foi a Spatial Lag.

Palavras chave: IDEB, Censo 2010, Educação, Análise Espacial, Análise Exploratória.

ABSTRACT

SPATIAL ANALYSIS OF THE BASIC EDUCATION DEVELOPMENT INDEX (IDEB) OF THE APARECIDA DE GOIÂNIA-GO SCHOOLS.

AUTHOR: POLLIANNA VIEIRA DA SILVA
ADVISOR: DR. GIOVANNI DE ARAUJO BOGGIONE

Aparecida de Goiânia-GO is a municipality in the state of Goiás that has been recognized for its industrial and population growth. The 2010 census estimated a total population of 455,657 people and the estimate for 2018 was 565,957 inhabitants, a significant increase of 24%. Because of this, there are growing demands in the area of services for the population, therefore requiring improvements in the city's infrastructure sectors. Within the demand for educational infrastructure, schools have an unquestionable role. All municipalities have a great demand in the area of education and it is a challenge to serve everyone with quality education. In 2007, IDEB was created, which is the development index of basic education and has the function of measuring the quality of education offered. The intention of this research is to show if there are spatial relationships between IDEB in the Municipality of Aparecida de Goiânia and socio-economic and environmental variables, using tools of Geographic Database, Geographic Information System, Data Mining, Digital Image Processing and Geostatistics. Information from the 2010 Census, School Census 2018, risk data from the municipality and satellite images were used. A database was created for spatial queries within a GIS environment; an environmental comfort map using digital image processing techniques; non-spatial relationships using data mining; and exploratory analyzes using spatial correlation analysis techniques. The miner APRIORI found relationships between IDEB, income and sewage. Spatial analyzes showed that families with the highest income are grouped in the northern region of the city where it borders on the city of Goiânia. The exploratory analysis for the approach using data only from the sectors where schools exist, Classical Regression was the most reliable. For the approach using all Sectors, the most reliable regression was Spatial Lag.

Keywords:IDEB, 2010 Census, Education, Spatial Analysis, Exploratory Analysis.

LISTA DE FIGURAS

Figura 1 – Mapa de Localização. 28
Figura 2 – Fluxograma Geral 29
Figura 3 – Fluxograma Metodológico. 30
Figura 4 – Atributos do Censo 2010 31
Figura 5 – Janela de Visualização do Qgis. 34
Figura 6 – Sistema Gerenciador de Banco de Dados Geográfico 34
Figura 7 – Mapa de Temperatura. 36
Figura 8 – Índice de Vegetação da Diferença Normalizada (NDVI) 37
Figura 9 – Mapa de Declividade. 38
Figura 10 – Mapa de dados de risco 39
Figura 11 – Mapa de Conforto Urbano Ambiental. 40
Figura 12 – Setores Censitários que possuem escolas 42
Figura 13 – Pontos centrais dos Setores Censitários. 43
Figura 14 – Diagrama de Voronoi 44
Figura 15 – Exemplo de Consulta Espacial. 45
Figura 16 – Consulta de renda média baixa 46
Figura 17 – Consulta de renda média alta 46
Figura 18 – Consulta de coleta de esgoto maior que 80% 47
Figura 19 – Consulta de coleta de esgoto menor que 20%. 48
Figura 20 – Cruzamento de dados de Renda, IDEB e Esgoto 48
Figura 21 – Cruzamento de altos valores de Renda, IDEB e Esgoto. 49
Figura 22 – Seleção de Variáveis 51
Figura 23 – Quadrantes do Gráfico de Moran. 52
Figura 24 – Índice de Moran Local para onde existem escolas 53
Figura 25 – Mapa Cluster para onde existem escolas. 54
Figura 26 – Mapa de Significância para onde existem escolas 54
Figura 27 – Regressão Clássica e 9 variáveis para onde existem escolas. 55
Figura 28 – Regressão Spatial Lag e 9 variáveis para onde existem escolas 57
Figura 29 – Regressão Clássica e 4 variáveis para onde existem escolas. 58
Figura 30 – Regressão Spatial Lag e 4 variáveis para onde existem escolas 59
Figura 31 – Índice de Moran Local de todos os Setores Censitários. 61

Figura 32 – Mapa Cluster de todos os Setores Censitários 61
Figura 33 – Mapa de Significância de todos os Setores Censitários. 62
Figura 34 – Regressão Clássica e 9 variáveis para todos os Setores Censitários 63
Figura 35 – Regressão Spatial Lag e 9 variáveis para todos os Setores Censitários .. 64
Figura 36 – Regressão Clássica e 4 variáveis para todos os Setores Censitários 65
Figura 37 – Regressão Spatial Lag e 4 variáveis para todos os Setores Censitários. ... 66

LISTA DE QUADROS

Quadro 1 – Dados utilizados referentes ao censo escolar. 32
Quadro 2 – Atributos utilizados referentes ao censo 2010. 32
Quadro 3 – Pesos aplicados na geração do Mapa de Conforto Ambiental. 36
Quadro 4 – Classes associadas ás variáveis utilizadas. 50
Quadro 5 – Comparação das regressões para onde existem escolas. 59
Quadro 6 – Comparação dos R^2 para onde existem escolas. 60
Quadro 7 – Comparação das regressões de todos os Setores Censitários. 66
Quadro 8 – Comparação dos R^2 de todos os Setores Censitários. 67

LISTA DE EQUAÇÃO

Equação 1 – Cálculo para obtenção das notas do IDEB. 22

LISTA DE ABREVIATURAS E SIGLAS

IDEB	Índice de Desenvolvimento da Educação Básica
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
SIG	Sistema de Informação Geográfica
SGBD	Sistema Gerenciador de Banco de Dados
SRTM	Shuttle Radar Topography Mission
NDVI	Índice de Vegetação com Diferença Normalizada
SAEB	Sistema Nacional de Avaliação da Educação Básica
MDE	Modelo Digital de Elevação
WEKA	Coleção de Algoritmos de Machine Learning e Data Mining
APRIORI	Algoritmo para geração de regras de associação

SUMÁRIO

1 INTRODUÇÃO 17

1.1TEMA 18

1.2 JUSTFICATIVA 18

2 OBJETIVOS 20

2.1 OBJETIVO GERAL 20

2.2 OBJETIVOS ESPECÍFICOS 20

3. REVISÃO DE LITERATURA 21

3.1. O IDEB 21

3.2 SISTEMAS DE INFORMAÇÕES GEOGRáFICAS E BANCO DE DADOS GEOGRÁFICOS 23

3.3 DESCOBERTA DE CONHECIMENTO E MINERAÇÃO DE DADOS 25

3.4 ANÁLISE EXPLORATÓRIA E ANÁLISE ESPACIAL 26

4. MATERIAIS E METODOLOGIA 27

4.1 ÁREA DE ESTUDO 27

4.2 MATERIAIS 28

4.3 PROCEDIMENTOS METODOLÓGICOS 29

4.3.1 BANCO DE DADOS GEOGRÁFICO 31

4.3.2 MAPA DE CONFORTO AMBIENTAL 35

4.3.3 MINERAÇÃO DE DADOS 40

4.3.4 ANÁLISE ESPACIAL 41

5. RESULTADOS 42

5.1 BANCO DE DADOS GEOGRÁFICOS E A INTEGRAÇÃO COM O AMBIENTE SIG 44

5.2 MINERAÇÃO DE DADOS 49

5.3 ANÁLISE EXPLORATÓRIA 52

6. CONCLUSÃO 68

7. REFERÊNCIAS..... 69

1 INTRODUÇÃO

Aparecida de Goiânia é um Munícipio de destaque no Estado de Goiás pelo seu crescimento Industrial, o que trouxe visibilidade e necessidade de melhorias na infraestrutura da cidade. O Censo do ano de 2010 estimou uma população total de 455.657 pessoas e a estimativa para 2018 de acordo com o IBGE foi de 565.957 habitantes. Aparecida é hoje um pólo industrial em diversos segmentos, e toda essa industrialização trouxe a necessidade de um desenvolvimento da infraestrutura do Município para poder atender os munícipes no que diz respeito aos serviços básicos, como a educação. A educação é um direito básico de acordo com o artigo 205 da Constituição Federal. A rede de Educação Municipal Básica de Aparecida de Goiânia no ano de 2018 contava com 59 unidades de Ensino Municipal Básico com um total de 23.974 alunos.

O objeto de estudo nesse trabalho são as Escolas Municipais, sua distribuição no perímetro do Município, seus desempenhos tendo como referência o IDEB (Índice de Desenvolvimento da Educação Básica) e suas relações com os indicadores sociais disponibilizados pelo Censo 2010, dados de segurança e indicadores ambientais. Foi realizada uma análise em três etapas, a primeira foi a criação de um banco de dados geográfico para consultas espaciais e visualização em um ambiente SIG, a segunda foi a etapa de mineração de dados e por fim a etapa de análise espacial e elaboração de mapas temáticos.

No Brasil o INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) divulga microdados da educação, que é considerado o Censo da educação. Nesse Censo se encontram todos dos dados relacionados à estrutura física da escola, informações a cerca de saneamento básico e até formação dos professores. O IDEB, objeto de estudo deste trabalho, foi criado em 2007 pelo INEP. TOSTA (2015) descreve o IDEB como o principal instrumento de avaliação da Educação Básica Nacional, mas seu resultado não agrada a todos os segmentos, por isso existem muitos questionamentos sobre a fidedignidade dos dados disponibilizados. A nota técnica do IDEB pontua a importância dele nos quesitos de identificar quais escolas estão com problemas no rendimento e proficiência, e assim, monitorar a evolução dos alunos.

Serão utilizados os dados oficiais do INEP e IDEB, além dos dados fornecidos pela Diretoria de Urbanismo da Prefeitura Municipal de Aparecida de Goiânia, dados de

Segurança Pública disponibilizados pela ouvidoria do Estado de Goiás e o mapa de conforto ambiental do município.

A educação é um pilar básico de direito a todo cidadão brasileiro para um melhor desenvolvimento da sociedade, mas na prática isso não significa que esse ensino oferecido atenda as expectativas em amplos sentidos. Na atualidade uma grande maioria tem acesso, mas também existe uma grande taxa de desistentes, de repetentes, e quando são submetidos a alguma avaliação de peso a probabilidade é de que a maioria não tenha um desempenho satisfatório. Dentro dessas possibilidades podemos entender que o sistema educacional pode ser falho deixando assim as notas do IDEB muitas vezes abaixo das metas pré-estabelecidas.

As notas do IDEB são importantes porque permitem o monitoramento de como anda o ensino no país, podendo detectar as escolas onde os alunos estejam com baixo desempenho e também monitorar o desempenho a longo prazo, INEP (2019).

A utilização de análises espaciais utilizando as técnicas de Geoestatística e Mineração de Dados é uma alternativa interessante para trazer transparência para informações a respeito da educação além de mostrar se existem ou não padrões dentro das variáveis utilizadas, e tentarão responder às questões como: Há relação espacial entre o IDEB e a situação socioeconômica das famílias dos alunos matriculados? Há alguma anomalia no cruzamento dos dados a partir do senso comum? Além de tudo isso, esse cruzamento de informações pode ajudar a traçar um perfil da cada escola, que podem ser disponibilizadas em um ambiente SIG.

1.1TEMA

Análise espacial de indicadores das escolas municipais de Aparecida de Goiânia/GO e suas relações com dados socioeconômicos e ambientais.

1.2 JUSTFICATIVA

Os municípios têm uma grande demanda na área da educação e um desafio é conseguir atender à todos, implantando escolas onde uma maior quantidade de pessoas terá acesso. E nessa parte entra a infraestrutura básica, como água, energia, coleta de lixo, transporte público, conforto ambiental, segurança pública, entre outros.

Esse trabalho tem o intuito de investigar variáveis que possam apresentar um cenário do comportamento espacial do IDEB, indo além de respostas esperadas, utilizando SIG, Banco de Dados, Mineração de Dados e Processamento Digital de Imagens.

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral do trabalho é mostrar se existem relações espaciais entre o IDEB das escolas municipais de Aparecida de Goiânia e varáveis sócio econômicas e ambientais.

2.2 OBJETIVOS ESPECÍFICOS

- Construção de cenários onde se possa avaliar a distribuição espacial das escolas;
- Analisar variáveis que possam ou não interferir nas notas do IDEB, como segurança pública, conforto ambiental e demais informações contidas no Censo 2010;
- Analisar de forma exploratória a dependência espacial das variáveis;
- Verificar a autocorrelação espacial de indicadores socioambientais, utilizando índice global de Moran e avaliando suas significâncias estatísticas;
- Elaboração de mapas oriundos de análises espaciais.

3. REVISÃO DE LITERATURA

3.1. O IDEB

O IDEB é o Índice de Desenvolvimento da Educação Básica que foi criado em 2007 como forma de avaliar a qualidade do ensino no Brasil. Para se fazer os cálculos leva-se em consideração os dados do Censo Escolar, as médias do SAEB (Sistema de Avaliação da Educação Básica) para as unidades da Federação e para o País e a Prova Brasil que é aplicada para os Municípios. As notas do IDEB variam de 0 a 10 (INEP,2019).

Na nota técnica disponibilizada pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), o IDEB é descrito como um indicador da qualidade educacional onde combina as informações da Prova Brasil ou do SAEB. Ele foi desenvolvido com a finalidade de sintetizar as informações de desempenho dos exames padronizados com as informações a cerca do rendimento escolar. A importância do IDEB nos termos de diagnóstico para melhoramento da politica publica na rede de educação é detectar em quais escolas os alunos possuem uma baixa proficiência e monitorar o desempenho dos alunos (Nota Técnica, 2019).

As notas utilizadas nesse trabalho foram as do IDEB de 2017 para anos iniciais, porque a maioria das Escolas Municipais de Aparecida de Goiânia não possuem nota de IDEB para anos finais em 2017. O IDEB é calculado de 2 em 2 anos e por isso foram utilizadas as informações de 2017. A Nota Informativa do IDEB 2017 diz que os resultados foram calculados a partir do desempenho no SAEB 2017 e das taxas de aprovação, que foram calculadas de acordo com o Censo Escolar 2017 (IDEB, Nota Informativa 2019).

Atualmente o INEP é o órgão responsável pelas estatísticas escolares. Ele foi criado em 1937 por Lei, e foi chamado inicialmente de Instituto Nacional de Pedagogia. Sua nomenclatura foi mudando com o passar dos anos e já em 1938 passou a ser chamado de Instituto Nacional de Estudos Pedagógicos. O ano de 1944 foi um marco, porque nesse ano foi lançada a Revista Brasileira de Estudos Pedagógicos, onde todos os estudos do instituto foram publicados. Já em 1952 Anísio Teixeira assumiu a direção do Instituto sendo ele quem deu ênfase a importância das pesquisas educacionais. Em 1972 veio uma nova denominação, Instituto Nacional de Estudos e Pesquisas Educacionais. Em 1976 a sede do INEP foi fixada em Brasília. Em 1985 com a nova república no poder, houve um novo redesenho institucional. Na década de 90 o órgão passou por uma nova reestruturação, e em 1997 foi

transformado em uma autarquia federal. Nos dias atuais é responsável pelos levantamentos estatísticos e avaliações no âmbito educacional. (INEP, 2019)

O IDEB fornece dois dados importantes: o fluxo escolar e as médias e desempenho nas avaliações. Pode-se dizer que ele é um medidor da qualidade da educação no País. (INEP, 2019)

Para se calcular o IDEB, algumas variáveis são utilizadas tais como a nota da Prova Brasil, o SAEB e os resultados do Censo Escolar. O INEP descreve a função do IDEB como agregação ao enfoque pedagógico dos resultados das avaliações, permitindo traçar metas para melhorar a qualidade educacional. (INEP, 2019)

O calculo para as notas do IDEB de forma geral são feitas de acordo com a Equação 1:

$$IDEB_{ji} = N_{ji} P_{ji}; \qquad 0 \leq N_j \leq 10; 0 \leq P_j \leq 1 \text{ e } 0 \leq IDEB_j \leq 10$$

Em que i = ano do exame (SAEB e Prova Brasil) e do Censo Escolar;

N_{ji} = média da proficiência em Lingua Portuguesa e Matemática, padronizada para um indicador entre 0 e 10, dos alunos da unidade j, obtida em determinada edição do exame realizado ao final da etapa de ensino;

P_{ji} = indicador de rendimento baseado na taxa de aprovação da etapa de ensino dos alunos da unidade j , Nota técnica IDEB (2019).

TOSTA (2015) diz que o IDEB é o principal instrumento para se avaliar a educação básica brasileira, e por isso é o responsável por mudanças que possam melhorar a politica educacional do País.

Em 2017 o Município de Aparecida de Goiânia alcançou as seguintes médias do IDEB: Para 4ª e 5ª séries a média foi de 5,5 e para a 8ª série e 9º ano a nota foi de 4,3. A nota projetada para 2019 para a 4ª e 5ª séries é de 5,9 e para a 8ª série e 9º ano, 5,0. (IDEB, 2019). Essas informações são de acesso livre, e isso é positivo porque permitem aos pais terem conhecimento das notas das escolas que pretendem matricular seus filhos.

Soares, J.F e Xavier, F.P (2013) discutem como o IDEB mudou a forma de avaliação do ensino e que a partir dele se leva em conta o aprendizado e o desenvolvimento do aluno e não apenas os processos de aprendizado e a gestão em si. No trabalho, eles apresentaram a forma como o IDEB é avaliado mostrando que não é um calculo simples como parece ser o senso comum. As notas são avaliadas entre 0 e 10, mas no atual cenário a nota 6,0 não é considerada ruim. Essa aplicação de notas também pode ter um peso negativo futuramente, porque existe uma tendência de as escolas quererem alcançar a nota 10 e para isso utilizarem

de alguns métodos que não são indicados pedagogicamente. Citam um ponto importante da avaliação: só é avaliado o aluno que estiver presente no dia da aplicação da prova Brasil e isso pode explicar porque algumas escolas de Aparecida de Goiânia não foram avaliadas. Os alunos por vários motivos podem ter preferido não fazer a prova, e se a quantidade for menor que a porcentagem exigida, a nota não é computada. E isso também abre a oportunidade de a escola "escolher" os alunos que farão a prova e assim a nota pode não refletir a realidade de todos os alunos.

Soligo, V. (2015), aborda a importância de o poder público ter acesso e investigar as variáveis que possam ter relações com o IDEB tendo os indicadores sociais como variáveis de pesquisa. As variáveis usadas no trabalho foram: o índice de Gini, o indicador de renda per capita, taxa de analfabetismo e IDEB.

Santos K.M; Santos M. J e Carniello M.F (2019), buscaram uma associação entre as escolas que estavam localizadas em regiões intraurbanas e a qualidade de ensino. O estudo mostrou que há uma independência espacial e este resultado não corroborou com as respostas esperadas, já que os autores tinham uma ideia de que se a escola fosse mais afastada das áreas centrais a qualidade do ensino seria pior.

3.2 SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS E BANCO DE DADOS GEOGRÁFICOS

Um SIG (Sistema de Informação Geográfica) tem por definição uma estrutura geral definida. Seus componentes são: Interface com usuário; Entrada e integração de dados; Funções de consulta e análise espacial; Visualização e plotagem; Armazenamento e recuperação de dados. Existe uma hierarquia entre esses componentes, Câmara e Queiroz (2001) explicam que o nível mais próximo do usuário é a interface homem-máquina, o nível intermediário é a parte de entrada, edição, análise, visualização e saída e no nível mais interno do sistema que é um sistema de gerência de banco de dados geográficos e é nesse nível que tem o armazenamento e a recuperação dos dados e dos atributos.

Câmara, G. e Queiroz, G.R. (2001) afirmam que "o termo SIG é aplicado para sistemas que realizam o tratamento computacional de dados geográficos e recuperam informações não apenas com base em suas características alfanuméricas, mas também através de sua localização espacial...".

Thomé (1999) listas funcionalidades de um SIG (Sistema de Informação Geográfica) como as mudanças de projeções e as diversas análises espaciais feitas a partir de um banco de dados. Essa funcionalidade será de fundamental importância para a visualização dos cenários propostos nesse trabalho.

Câmara, G. e Queiroz, G.R. (2001) também listam as principais características de um SIG, que são: Inserir e integrar, numa única base de dados, informações espaciais provenientes de dados cartográficos, dados censitários e cadastro urbano e rural, imagens de satélite, redes e modelos númericos de terreno; oferecer mecanismos para combinar as várias informações, através de algoritmos de manipulação e análise, bem como para consultar, recuperar, visualizar e plotar o conteúdo da base de dados georreferenciados.

Um SIG tem uma estrutura pré definida de forma hierárquica: existe a interface com o usuario que é o primeiro degrau, em seguida o usuário entra com os dados e os integra e a partir disso já se pode fazer consultas espaciais, visualizar e plotar esse cenário. Por último, armazenar essas informações em um banco de dados geográfico. (CÂMARA , 2001)

O sistema de bancos de dados possui três níveis de arquitetura, são eles: Nível de armazenamento ou interno, Nível do usuário ou externo e Nível lógico ou conceitual. O nível interno é aquele que mais se aproxima de um meio de armazenamento físico e se ocupa do modo como os dados são armazenados dentro do sistema; O nível externo fica mais próximo ao usuário e se ocupada do modo como são vistos por usuários individuais; e o nível conceitual é um nível indireto entre os outros dois. (DATE, 2004)

Com o passar dos tempos e o aumento da circulação de informações, se fez necessário estudar métodos onde fosse mais fácil lidar com o grande fluxo de dados. Com o surgimento dos computadores o compartilhamento, a edição e o filtro desses dados fluem de forma mais rápida e eficaz.

Alves (2004) descreve que um Banco de Dados é um conjunto de dados que possuem significado, que não são dados aleatórios e tem a capacidade de se relacionarem. Já o Sistema de Gerenciamento de Banco de Dados (SGBD) é o conjunto de softwares que gerenciam um banco de dados.

Carvalho, J.P e Delgado, J.P.M (2019) inseriram os dados do EDUCACENSO em um banco de dados geográficos para futuras consultas espaciais em relação ao transporte escolar municipal, além da geração de um mapa de origem-destino. Eles descrevem que o objetivo principal do trabalho é identificar onde são as maiores concentrações de deslocamentos e relacionar isso com os alunos da rede municipal que utilizam o transporte municipal que é

fretado pelo município, ou seja, utilizaram as informações geográficas para gerar os mapas de origem/destino a partir dos dados do censo educacional.

3.3 DESCOBERTA DE CONHECIMENTO E MINERAÇÃO DE DADOS

A descoberta de conhecimento em bancos de dados (KDD) é o processo não trivial de identificar em dados, padrões que sejam válidos, novos (previamente desconhecidos), potencialmente úteis e compreensíveis, visando melhorar o entendimento de um problema ou um procedimento de tomada de decisão. É uma tecnologia que trabalha em diversas áreas: estatística, inteligência computacional, reconhecimento de padrões e banco de dados. (GOLDSCHMIDT, 2005).

Um dos objetivos da descoberta de conhecimento em bancos de dados é tornar padrões compreensíveis para humanos, visando promover uma melhor compreensão dos próprios dados. Embora seja um tanto subjetivo medir compreensibilidade, um dos fatores frequentes é a medida de simplicidade. O fator de compreensão dos dados está relacionado à intuitividade da representação destes, bem como da granularidade alta o suficiente para que estes sejam compreendidos.

A Mineração de Dados é uma das etapas envolvidas no KDD. Ela tem o potencial de trabalhar com um grande volume de informações, encontrando padrões entre elas. (Fayyad, U.; Shapiro, G.P.; Smyth, P. ,1996)

Segundo Calixto, K.E.A.; Gusmão, R.P.de; Segundo, C.V.N. (2017), a mineração de dados é um processo de análise projetado para exploração de dados de grande volume. Por meio de algoritmos, a mineração de dados busca descobrir padrões de comportamento, prever resultados e criar novas informações.

Nesse trabalho os algoritmo a serem utilizados pelo minerador serão o APRIORI, que é um algoritmo que utiliza regras de associação e que pode descobrir os conjuntos de itens que satisfaçam um limite inicial estipulado, independente da quantidade de itens; a seleção de atributos, que apresenta as variáveis mais correlacionadas dado o conjunto de dados e a clusterização que busca encontrar agrupamentos com alguma similaridade.

Fonseca, S.O e Namen, A.A (2016) utilizaram a mineração de dados para analisar os dados do INEP a fim de identificar o perfil dos professores de matemática do 9º do ensino fundamental utilizando o minerador para encontrar padrões. Utilizaram o minerador de dados devido à grande quantidade de informação disponibilizada. A mineração indicou algums

pontos favoráveis para o aprendizado de matemática como a assiduidade dos professores, quantidade de conteúdo previsto, entre outros.

3.4 ANÁLISE EXPLORATÓRIA E ANÁLISE ESPACIAL

A análise espacial diz respeito a localização geográfica do objeto de estudo. No caso desse trabalho a análise em questão é a geolocalização das escolas municipais que possuem nota do IDEB. A análise exploratória é uma ferramenta muito útil para o entendimento das relações entre as variáveis baseadas na dependência espacial.

A dependência espacial é uma característica inerente à representação de dados através de divisões territoriais. Pode ser entendida como a tendência que uma variável tem de ser explicada por amostras dos seus vizinhos mais próximos do que o restante da localização.

Essa dependência espacial pode ser medida de diversas formas. Uma delas é o Índice de Moran que mede a autocorrelação espacial, indicando um grau de associação espacial presente no conjunto de dados.

Existem dois tipos de Índice de Moran: o global e o local. No global é feita a autocorrelação levando em consideração o primeiro vizinho na amostra associando um valor para o conjunto todo, já no local é feita uma autocorrelação mais detalhada para cada área. (Luzardo, A.J.R.; Castaneda Filho, R.M; Rubim, I.B., 2017). Nesta pesquisa utilizaremos o Índice de Moran local.

Souza, H.G; Tabosa, F.J.S (2016) trabalharam com análise espacial de todos os municípios do Ceará, utilizando os dados do INEP e do Instituto de Pesquisa e Estratégia econômica do Ceará. Estipularam um valor para a variável dependente que é o IDEB dos anos iniciais e finais e as variáveis explicativas são as disponibilizadas pelo Instituto de Pesquisa do Ceará. Chegaram a conclusão de que as séries finais não conseguiram atingir a meta, e que as séries iniciais, em sua maioria, conseguiram.

4. MATERIAIS E METODOLOGIA

Para a execução desta pesquisa, foram realizadas duas abordagens para o problema. O município de Aparecida de Goiânia possui 677 setores censitários, mas apenas 52 destes setores possuem escolas com nota de IDEB. Embora o dado oficial contemple 59 escolas, 7 delas não possuíam nota e constavam como nota 0 (zero). Para não contar como outlier, já que os valores das notas estavam entre 4,9 e 6,4, essas escolas foram descartadas das análises.

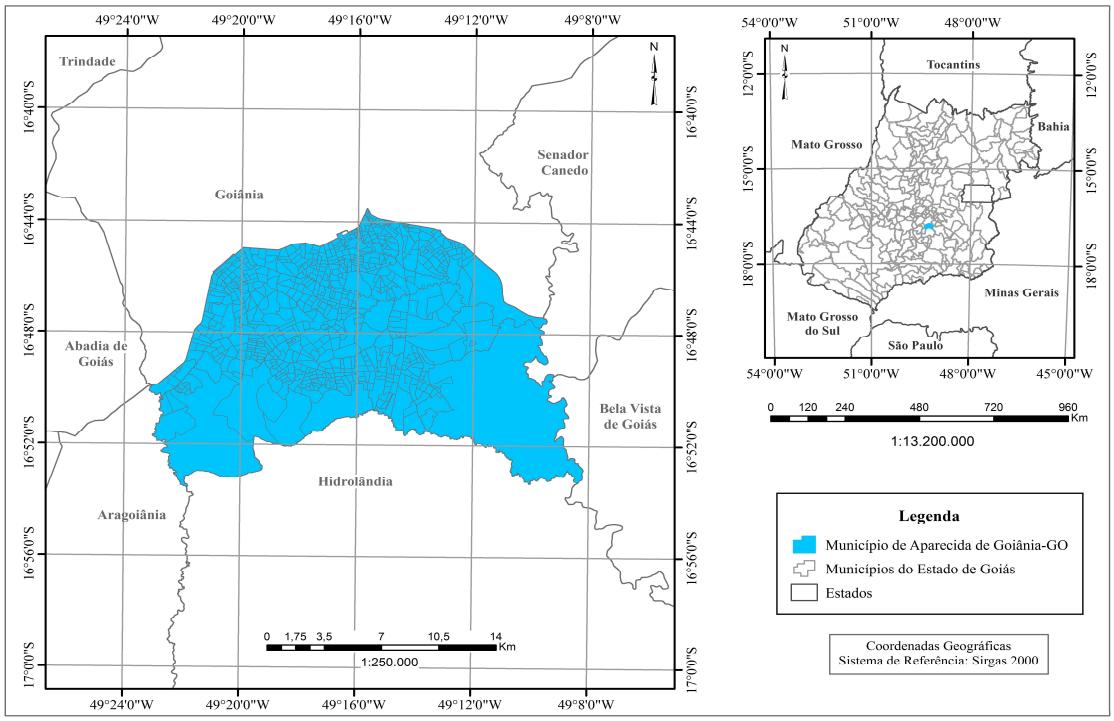
Na primeira abordagem os valores das 52 escolas foram "distribuídos" para as escolas vizinhas usando um critério de proximidade espacial e o interpolador de vizinho mais próximo. Desta forma, os 679 distritos passaram a ter um valor de IDEB associado à sua área.

Na segunda abordagem foram utilizados apenas os distritos censitários que possuíam valores na planilha do IDEB fornecidos pelo INEP, ou seja, apenas os 52 setores. Os procedimentos metodológicos são os mesmos para as duas abordagens, e a diferença dos resultados serão expostos no item de resultados.

4.1 ÁREA DE ESTUDO

A área de estudo é o Município de Aparecida de Goiânia-GO que faz parte da região metropolitana de Goiânia, sendo privilegiado por sua localização central no país. Possui uma população estimada para 2019 de 578.179 habitantes e uma área aproximada de 291,13 km². A Figura 1 mostra a localização do município dentro do estado de Goiás.

Figura 1 - Mapa de Localização.



Fonte: A autora, 2019.

4.2 MATERIAIS

Os materiais utilizados no desenvolvimento deste trabalho foram:

- Planilhas do Censo 2010 disponibilizadas pelo IBGE;

- Planilhas do Censo da Educação 2018 disponibilizadas pelo INEP;

- Mapa de conforto ambiental do município; que foi gerado a partir de uma imagem Landsat 8-OLI com resolução espacial nominal de 30 metros, a Banda 10 termal da Landsat 8-OLI (Operational Land Image) com resolução espacial nominal de 100 metros, NDVI (Normalized Difference Vegetation Index), Declividade que foi feita a partir de uma imagem SRTM (Shuttle Radar Topography Mission) com elevação espacial nominal de 30 metros e os dados de riscos do município;

- Arquivos vetoriais do Município disponibilizados pela Secretaria de Planejamento e Regulação Urbana de Aparecida de Goiânia;

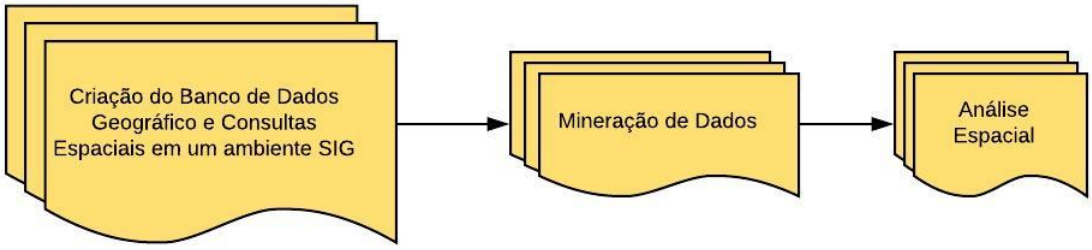
- Notas do IDEB 2017 disponibilizados pelo INEP;

- Dados de segurança por bairro de Aparecida de Goiânia-GO fornecido pela Secretaria de Segurança Pública do Estado.

4.3 PROCEDIMENTOS METODOLÓGICOS

A metodologia foi realizada em três etapas complementares. A primeira foi a criação de um Banco de Dados Geográfico e consultas espaciais em um ambiente SIG, a segunda foi a descoberta de conhecimento utilizando mineração de dados e a terceira foi a etapa de análise espacial.

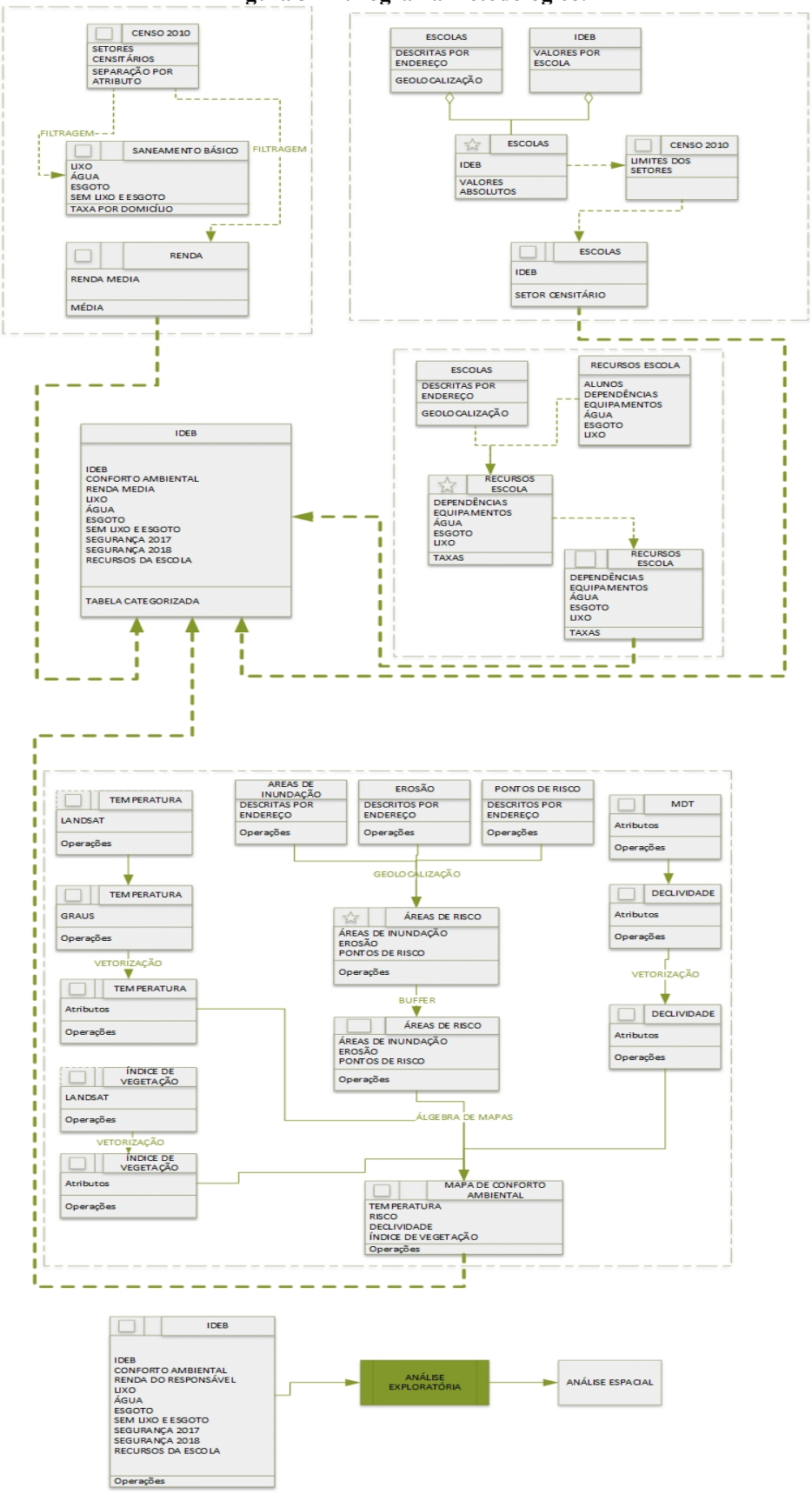
Figura 2 - Fluxograma Geral.



Fonte: A autora, 2019.

A Figura-3 apresenta o fluxograma metodológico utilizando diagrama de transformação. As etapas são descritas em ordem de execução.

Figura 3- Fluxograma metodológico.



Fonte: A autora, 2019.

4.3.1 BANCO DE DADOS GEOGRÁFICO

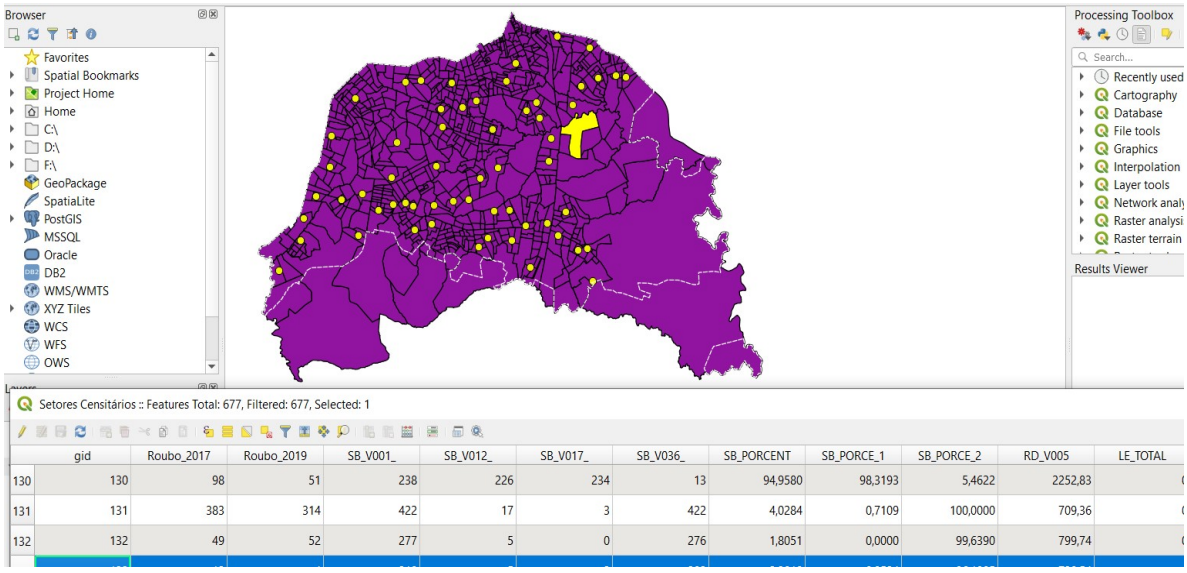
Os dados do censo 2010 foram filtrados deixando somente os atributos de interesse, tais como os de saneamento básico (lixo, água, esgoto, e as residências sem coleta de lixo e esgoto) e o atributo de renda média por domicílio. Esses dados foram inseridos no arquivo vetorial de setores censitários.

Os dados das escolas estavam em forma de endereços e os dados do IDEB podem ser consultados no site do INEP individualmente. Essas duas informações foram geolocalizadas em forma de ponto no centroide da área do setor censitário.

Os mesmos dados das escolas e mais os dados disponibilizados pelo censo da educação que contém informações referentes a escola, tais como: se possui coleta de água, como são as dependências da escola, os equipamentos disponíveis, esgoto e lixo, também foram espacializados como pontos no centroide da área do setor censitário.

Para elaboração do Banco de Dados às informações do Censo 2010, das Escolas e de segurança publica foram inseridas ao arquivo vetorial de Setores Censitários disponibilizados pelo IBGE como mostra a Figura 4.

Figura 4 – Atributos Censo 2010.



Fonte: A autora, 2019.

Os atributos utilizados relacionados às escolas são apresentados no Quadro 1:

Quadro 1 - Dados utilizados referentes ao Censo escolar.

Dados retirados do Censo Escolar	
E_IN_AGUA_	Água consumida pelos alunos na escola passa por um processo de filtragem;
E_IN_AGU_1	Abastecimento de água - Rede pública;
E_IN_AGU_2	Abastecimento de água - Poço artesiano
E_IN_AGU_3	Abastecimento de água - Cacimba/Cisterna/Poço
E_IN_ESGOT	Esgoto sanitário - Rede pública;
E_IN_ESG_1	Esgoto sanitário – Fossa
E_IN_LIX_1	Destinação do lixo – Queima
E_IN_LIX_3	Destinação do lixo – Recicla
E_IN_LABOR	Dependências existentes na escola - Laboratório de informática
E_IN_LAB_1	Dependências existentes na escola - Laboratório de ciências
E_IN_QUADR	Dependências existentes na escola - Quadra de esportes Coberta
E_IN_QUA_1	Dependências existentes na escola - Quadra de esportes Descoberta
E_IN_QUA_2	Dependências existentes na escola - Quadra de esportes coberta ou descoberta
E_IN_EQUIP	Equipamentos existentes na escola - Aparelho de televisão

Fonte: Censo,2010.

Os atributos utilizados relacionados aos Setores Censitários são apresentados no Quadro 2.

Quadro 2 - Atributos utilizados referentes ao Censo 2010.

Atributos do Censo 2010	
Dados retirados da Tabela Domicilio01 – Referentes à Saneamento Básico (SB)	
SB V001	Domicílios particulares e domicílios coletivos
SB V012	Domicílios particulares permanentes com abastecimento de água da rede geral
SB V017	Domicílios particulares permanentes com banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via rede geral de esgoto ou pluvial
SB V036	Domicílios particulares permanentes com lixo coletado por serviço de

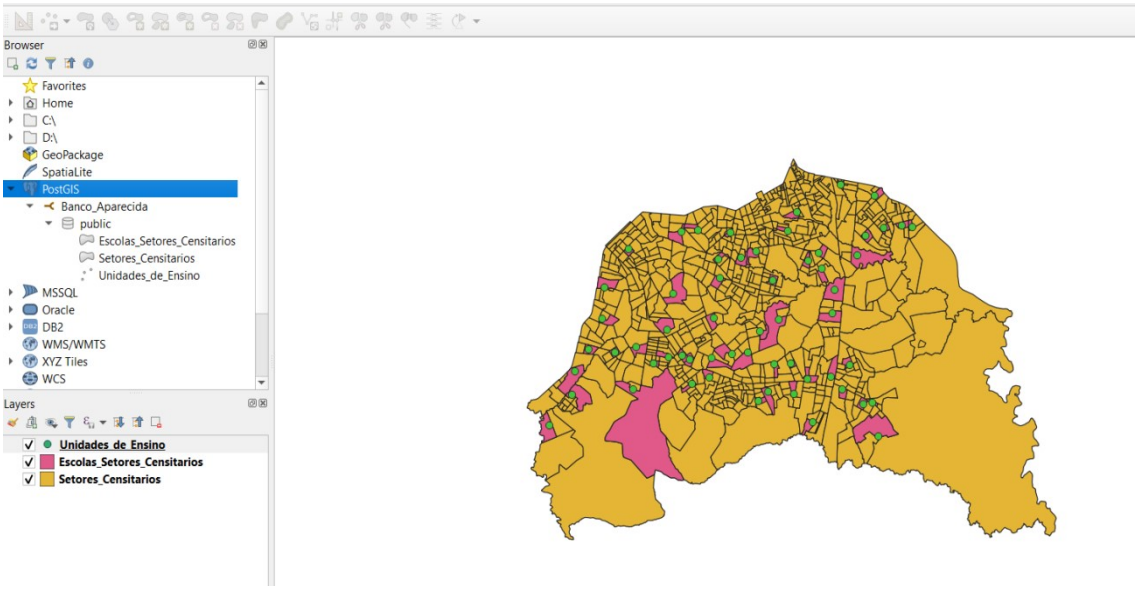
	limpeza;
Dados retirados da Tabela Domicilio01-Referentes à Lixo e Esgoto	
LE V001	Domicílios particulares e domicílios coletivos
LE V236	Domicílios particulares permanentes com outra forma de destino do lixo, banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via fossa séptica
LE V237	Domicílios particulares permanentes com outra forma de destino do lixo, banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via fossa rudimentar
LE V238	Domicílios particulares permanentes com outra forma de destino do lixo, banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via vala
LE V239	Domicílios particulares permanentes com outra forma de destino do lixo, banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via rio, lago ou mar
LE V240	Domicílios particulares permanentes com outra forma de destino do lixo, banheiro de uso exclusivo dos moradores ou sanitário e esgotamento sanitário via outro escoadouro
Dados retirados da Tabela Basico_GO – Referente a Renda Média	
RD V005	Valor do rendimento nominal médio mensal das pessoas responsáveis por domicílios particulares permanentes (com e sem rendimento);

Fonte: Censo, 2010.

Depois de unir as informações das tabelas aos arquivos vetoriais de interesse, inserimos os dados através do SGBD (Sistema Gerenciador de Banco de Dados), utilizando o software QGis. Assim quando efetuamos alguma alteração no arquivo vetorial no ambiente SIG a mesma já será feita dentro do Banco de Dados Geográfico. O Banco de dados utilizado foi o PostgreSQL com extensão espacial PostGis.

A Figura 5 mostra a janela de visualização do SIG com os arquivos vetoriais dentro do Sistema Gerenciador de Banco de Dados.

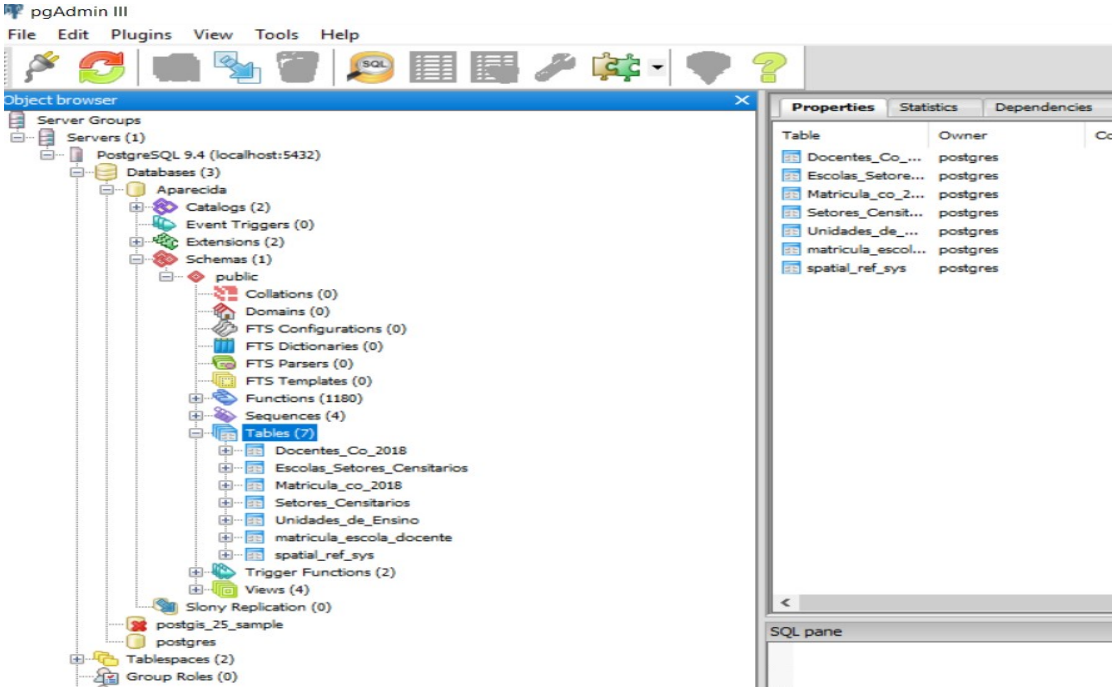
Figura 5- Janela de visualização do QGIS.



Fonte: A autora, 2019.

Na Figura 6 são mostradas todas as informações espacializadas e as não espacializadas, dentro do pgAdmimIII. Essas informações são as tabelas do Censo 2010 e Censo da Educação.

Figura 6 - Sistema Gerenciador de Banco de Dados Geográfico.



Fonte: A autora, 2019.

O município de Aparecida de Goiânia contava com um total de 23.974 alunos da rede Municipal (não entram as escolas com convênios) no ano de 2018. Desse total foram retiradas as sete escolas que não possuem nota no IDEB 2017. A nota informativa do IDEB pontua os motivos para que algumas escolas não possuam notas de IDEB no ano de 2017. As que se encaixam no contexto de escolas públicas municipais são:

- Escolas Públicas que oferecem ensino fundamental e/ou ensino médio que não realizaram o SAEB por terem menos de 10 alunos matriculados nas etapas avaliadas (5º ano/4ªsérie, 9ºano/8ªsérie e 3ª ou 4ªsérie do ensino médio), conforme declaração prestada ao Censo Escolar 2017;

- Escolas pertencentes às redes municipais que não aderiram ao SAEB 2017;

- Escolas que realizaram o SAEB 2017, mas não prestaram informações ao Censo Escolar sobre os alunos aprovados e por isso não tiveram a taxa de aprovação calculada;

- Escolas em que o número de alunos participantes do SAEB 2017 não alcançaram 80% dos alunos matriculados na etapa avaliada (Portaria nº 447, de 24 de maio de 2017). Nota informativa IDEB, (2019). No ano de 2017, 52 escolas municipais foram avaliadas e possuem notas de IDEB no Município de Aparecida de Goiânia-Go.

4.3.2 MAPA DE CONFORTO AMBIENTAL

O Mapa de Conforto Ambiental tem o objetivo de mostrar as áreas que oferecem uma melhor qualidade de vida baseado em questões que abordam conforto de acordo com o senso comum. A dificuldade é escolher os melhores indicadores e nesse trabalho foram utilizados como variáveis os dados de Declividade, Temperatura, Índice de Vegetação (NDVI) e os dados sobre os riscos existentes no município.

O mapa de conforto ambiental do município foi gerado a partir de uma adaptação do trabalho de Carneiro (2009). Para tal, foi utilizado um conjunto de imagens Landsat 8-OLI do dia 07/09/2019, de onde foram extraídos os dados de temperatura e índice de vegetação; SRTM de onde se extraiu a declividade, e os dados de risco fornecidos pela Defesa Civil. Para se chegar ao mapa final foi feita uma álgebra de mapas inserindo pesos nas variáveis como é apresentado no Quadro 3.

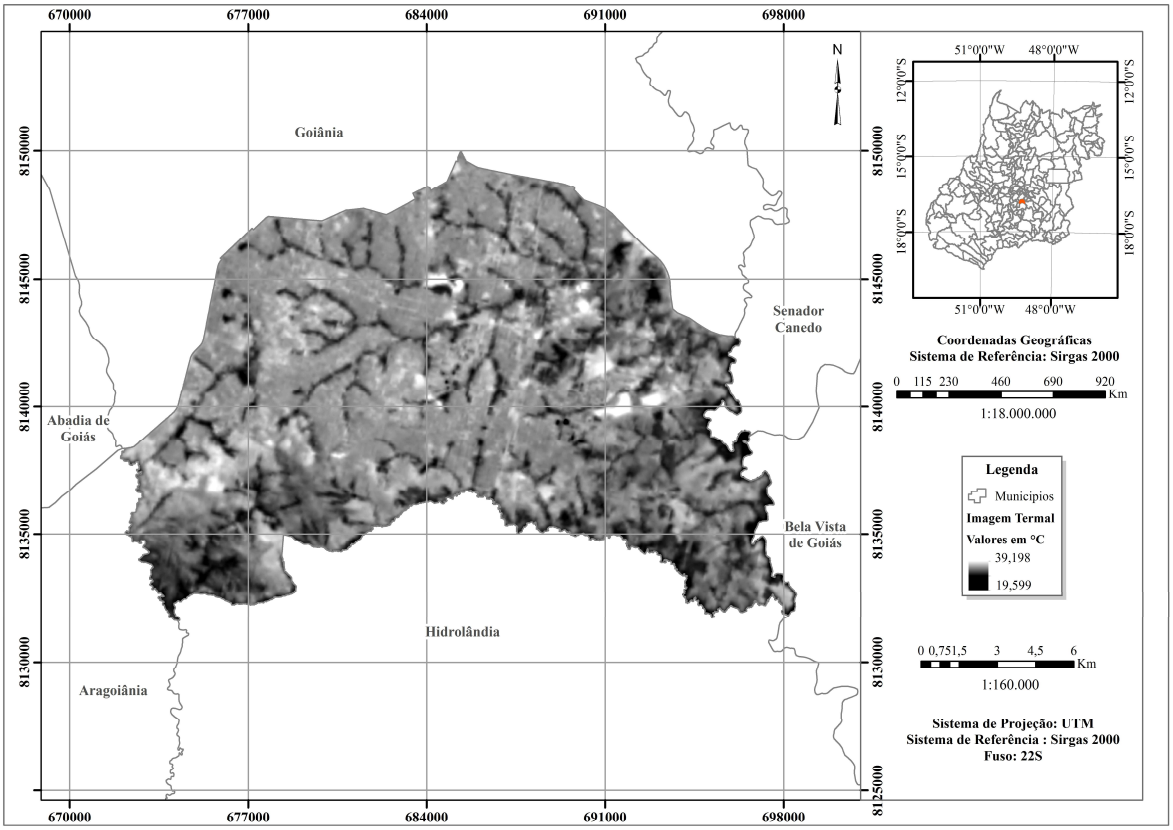
Quadro 3 - Pesos aplicados na geração do mapa de conforto ambiental.

Imagens	Pesos aplicados
Risco	0.15
NDVI	0.45
Declividade	0.15
Temperatura	0.25

Fonte: A autora, 2019.

Para se chegar aos dados de temperatura, foi feita uma conversão do valor do pixel para graus célsius. Os dados de temperatura foram extraídos da Banda 10 termal da Landsat 8-OLI do dia 07 de setembro de 2019. Os valores da temperatura variaram entre 19,59°C e 39,20º. A Figura 7 mostra que as áreas mais claras são aquelas onde a temperatura é alta, e as partes mais escuras são onde a temperatura é mais amena.

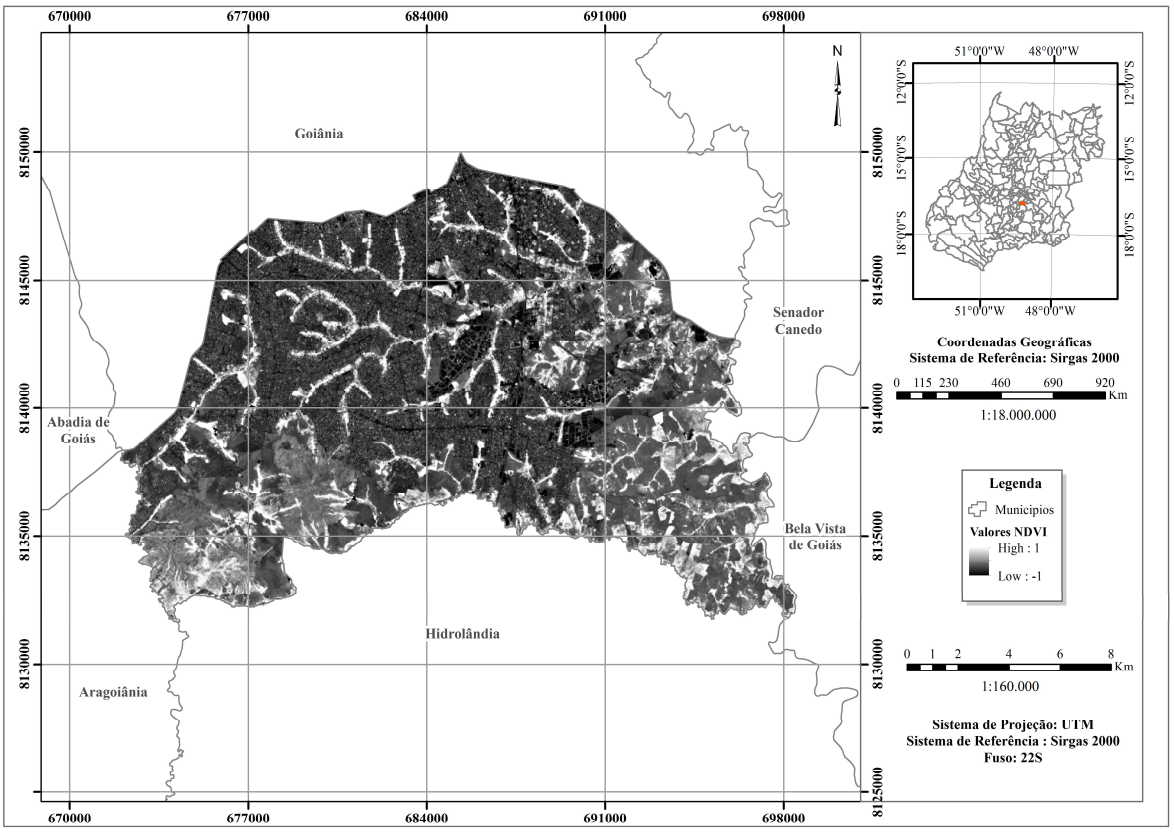
Figura 7- Mapa de Temperatura.



Fonte: A autora, 2019.

O intuito de um Índice de Vegetação é utilizar duas bandas espectrais para realçar as características da vegetação, no caso do Landsat 8 OLI foram utilizadas as bandas 4 e 5, sendo a 4 do Vermelho e a 5 do Infra-Vermelho. O NDVI é um Índice de Vegetação Normalizado, seus valores variam entre -1 e 1, onde 1 significa uma alta quantidade de vegetação e -1 indica a ausência de vegetação. Na Figura 8 é possível observar que as áreas que estão mais próximas a algum curso d'água está perto do valor 1. Indicando que essas regiões possuem algum tipo de vegetação.

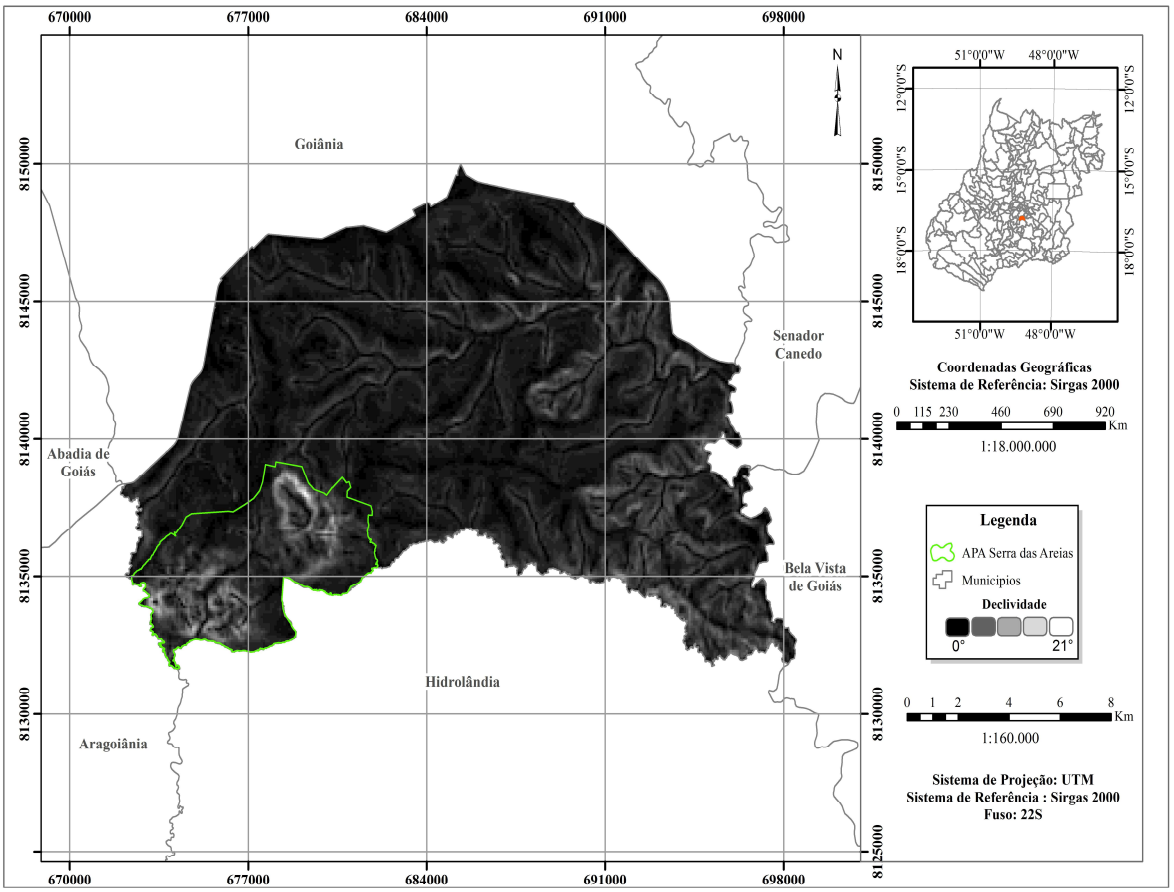
Figura 8- Índice de Vegetação da Diferença Normalizada (NDVI).



Fonte: A autora, 2019.

Os dados de declividade foram extraídos automaticamente da imagem SRTM que é um MDE (Modelo Digital de Elevação). Os valores encontrados variam entre 0° e 21°. As áreas mais claras são referentes aos valores mais altos de declividade, e os mais escuros se referem aos valores mais baixos. As áreas onde se encontram as maiores declividades estão concentradas dentro do perímetro da Área de Preservação Ambiental Serra das Areias como mostra a Figura 9.

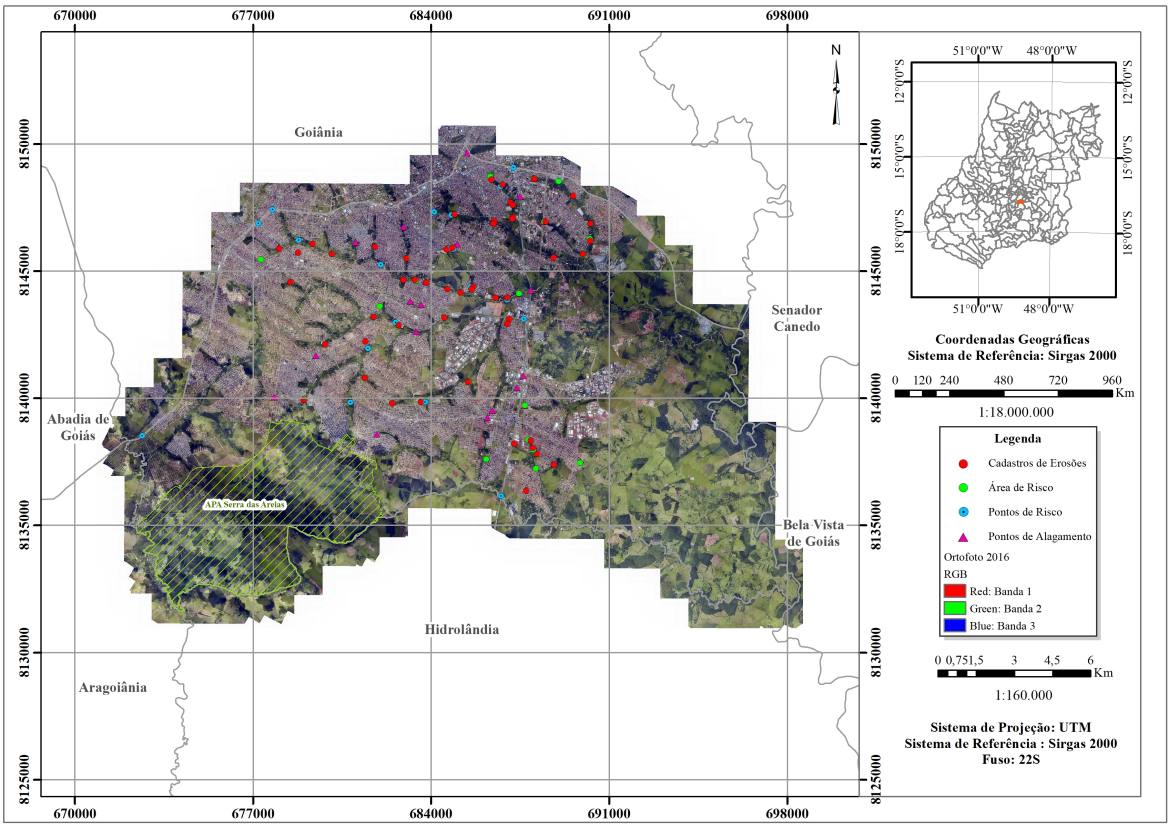
Figura 9 - Mapa de Declividade.



Fonte: A autora, 2019.

Os dados de riscos utilizados foram: Cadastro de erosões, Áreas de risco, Pontos de risco e Pontos de alagamentos, que foram disponibilizados pela Defesa Civil á Secretaria de Planejamento e Regulação Urbana. Os dados foram utilizados para criação de pontos com a localização fornecida (geolocalização). O resultado é mostrado na Figura 10. Feita a geolocalização, foi gerado um buffer de 300m que para determinar as áreas com maior probabilidade de risco, e outro buffer de 500m para uma área com risco moderado.

Figura10-Mapas de dados de risco.

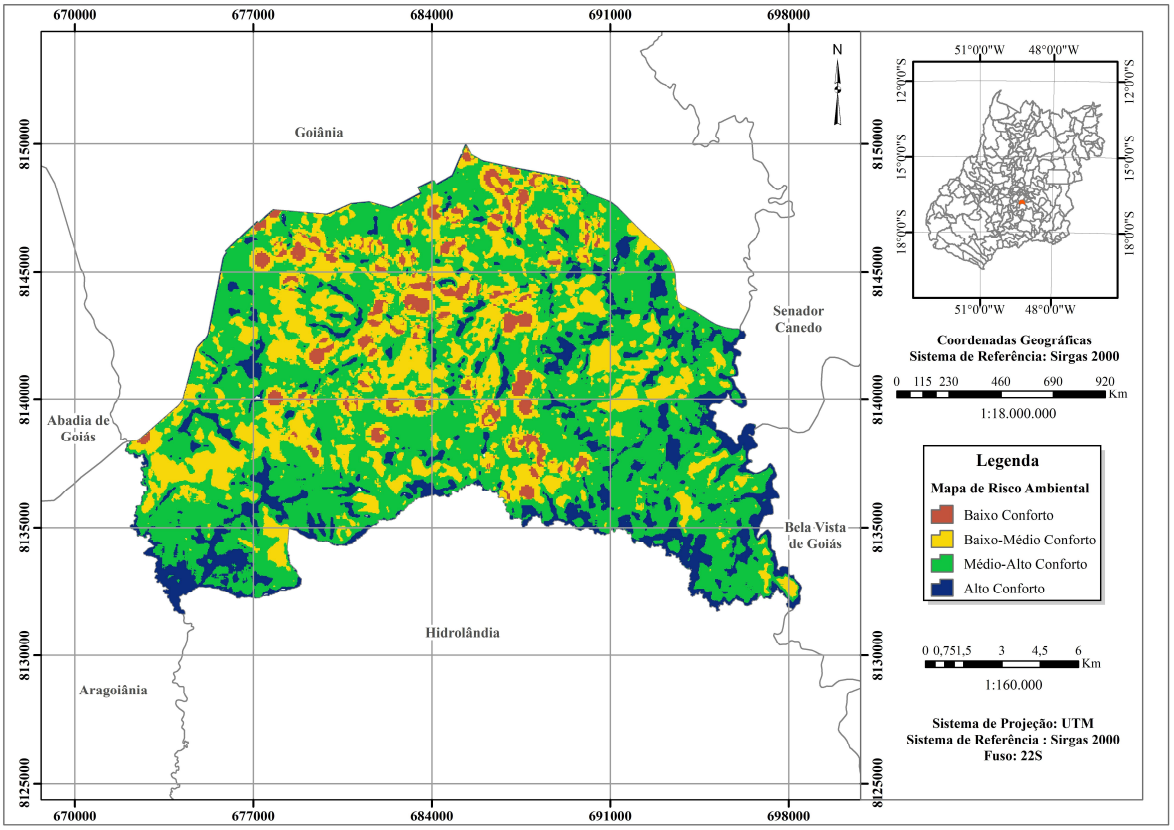


Fonte: A autora, 2019.

De posse dos dados normalizados, foi realizada a álgebra de mapas para apontar as localidades com melhor conforto ambiental. Nos dados, apenas a informação de NDVI não é inversamente proporcional aos demais. Como critério, considerou-se que áreas com mais vegetação possuem maior conforto, enquanto que temperatura, risco e declividade, quanto maiores, pior o cenário.

Após a álgebra de mapas foi gerado o Mapa de Conforto como mostra a Figura 11. Ele foi classificado em quatro intervalos: Baixo conforto, Baixo-Médio conforto, Médio-Alto conforto e Alto conforto. É possível observar que as áreas de alto conforto estão afastadas da área urbana, se localizando nas extremidades do município. As maiores áreas estão entre o baixo-médio conforto e o médio-alto conforto. O baixo conforto são algumas áreas pontuais distribuídas por todo o município.

Figura11- Mapa de Conforto Urbano Ambiental.



Fonte: A autora, 2019.

Este dado gerado também foi adicionado ao Banco de Dados para a análise com o minerador de dados e a análise espacial.

4.3.3 MINERAÇÃO DE DADOS

Nesta etapa do trabalho, os dados foram categorizados em intervalos arbitrários descartando o atributo espacial. Este procedimento é necessário para a utilização do programa WEKA, o minerador de dados. Dentre os algoritmos úteis para o problema desta pesquisa se destacam o seletor de variáveis, que encontra as variáveis mais correlacionadas com a variável dependente e o algoritmo de regras de associação. Para encontrar possíveis associações entre a variável dependente (IDEB) e demais variáveis foi utilizado o algoritmo APRIORI.

O WEKA é um software livre, produzido pela Universidade de Waikato da Nova Zelândia. Dentro dele existem vários algoritmos para serem utilizados na Mineração de Dados.

Foi utilizado o APRIORI que é o algoritmo proposto em AGRAWAL; SIKRANT para a mineração de conjuntos de itens frequentes. Usando este algoritmo, podemos descobrir todos os conjuntos de itens que satisfaçam um limite mínimo de frequência, independente do seu tamanho. Esse algoritmo se baseia na seguinte propriedade: “Todo subconjunto não vazio de um conjunto de itens frequentes também dever ser frequente.” Isto significa que dentro de um conjunto de itens frequente de tamanho k, todos os seus subconjuntos de tamanho m ($0 < m < k$) também são frequentes. O que é lógico; pois um conjunto não tem como ser mais frequente que um de seus subconjuntos.

4.3.4 ANÁLISE ESPACIAL

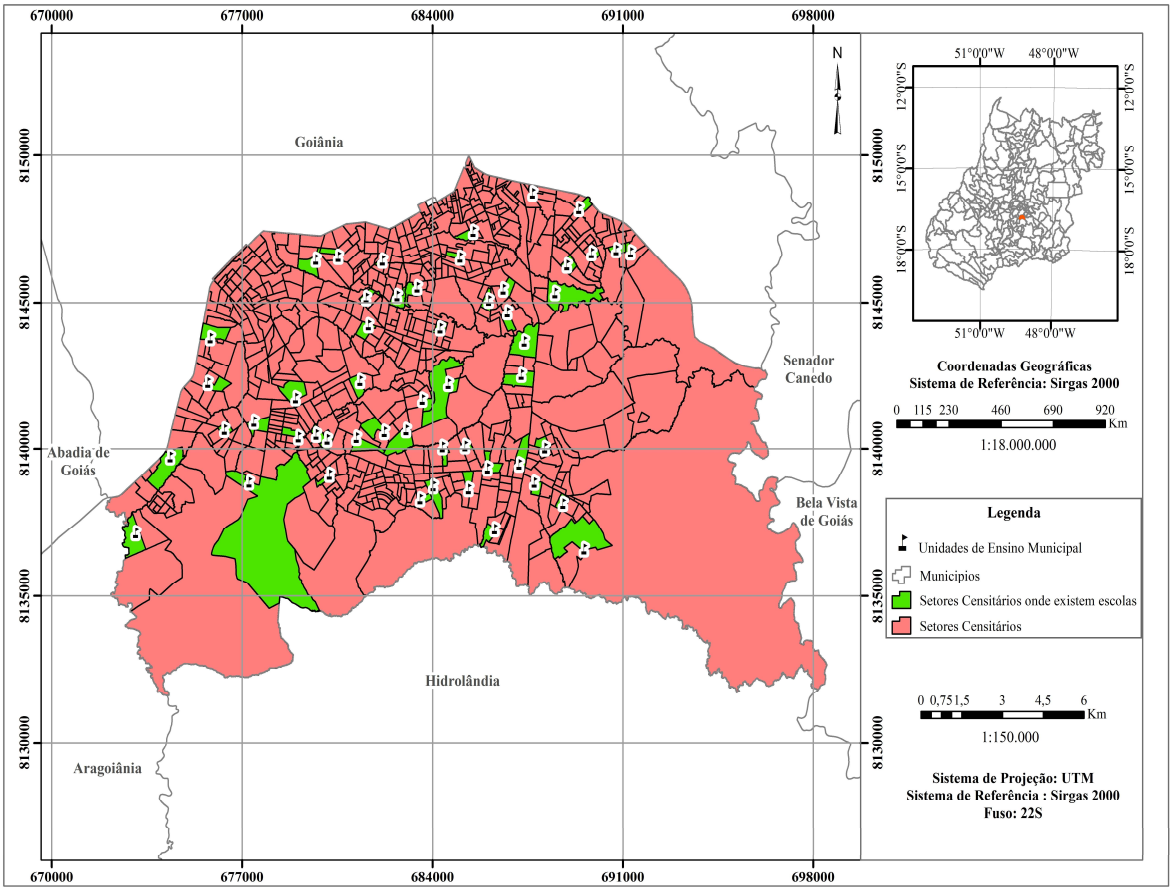
Na última etapa do trabalho foi realizada a análise exploratória e a análise espacial do conjunto de dados catalogados e gerados para esta pesquisa.

Para a análise exploratória foi calculado o Índice de Moran para avaliar a correlação espacial; e análise de regressão clássica e regressão Spatial Lag que é utilizada quando se deseja explicar a variável dependente a partir dela mesma e de outras variáveis explicativas.

5. RESULTADOS

Os resultados serão apresentados conforme as abordagens explicadas anteriormente. Uma abordagem trata os dados reais, ou seja, os setores censitários onde existem escolas (Figura 12), e a outra com esses dados estimados para o restante dos setores censitários (Figura 13). Uma ferramenta que calcula a menor distância entre os arquivos vetoriais foi utilizada, e em seguida foi feita uma junção desses dados, gerando assim outra tabela de atributos com valores extraídos das escolas mais próximas.

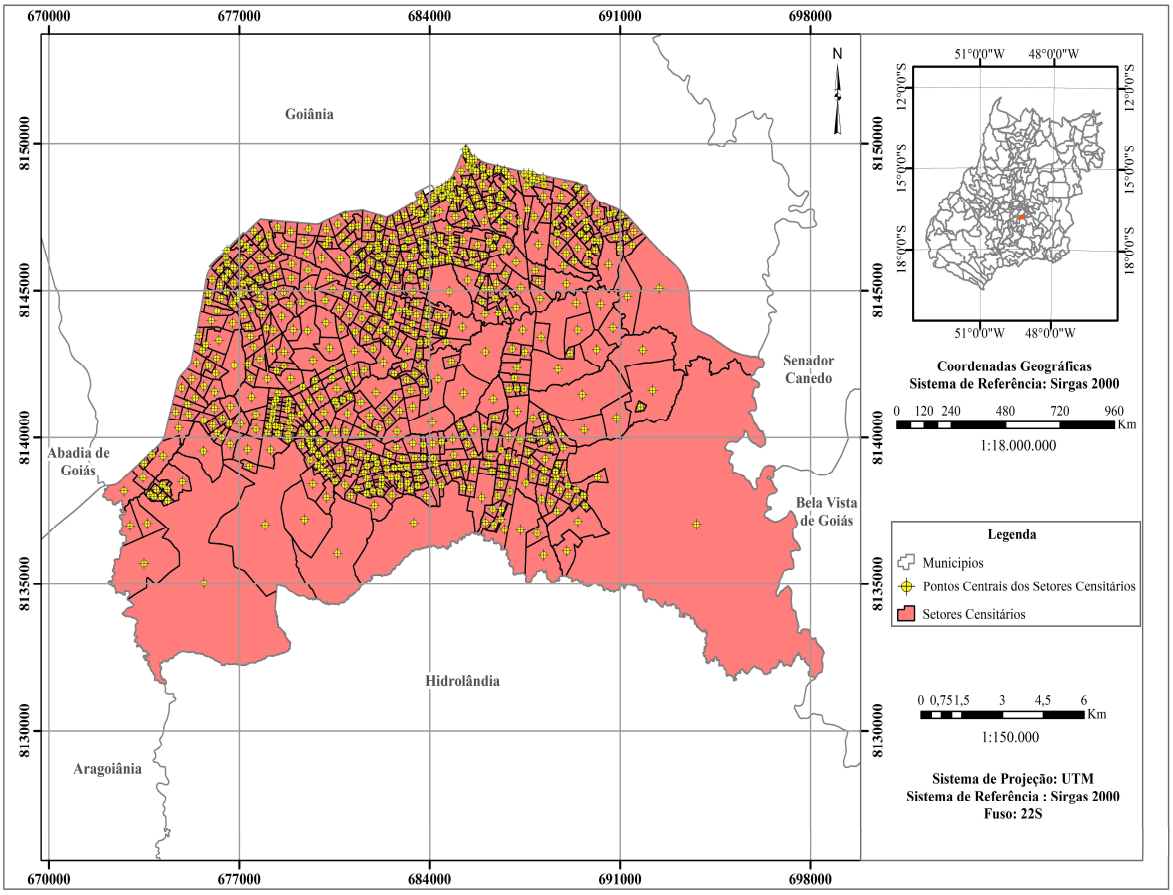
Figura12- Setores Censitários que possuem escolas.



Fonte: A autora, 2019.

O município de Aparecida de Goiânia é composto por 677 Setores Censitários, então foram gerados 677 pontos centrais como mostra a Figura 13.

Figura 13 - Pontos Centrais dos Setores Censitários.

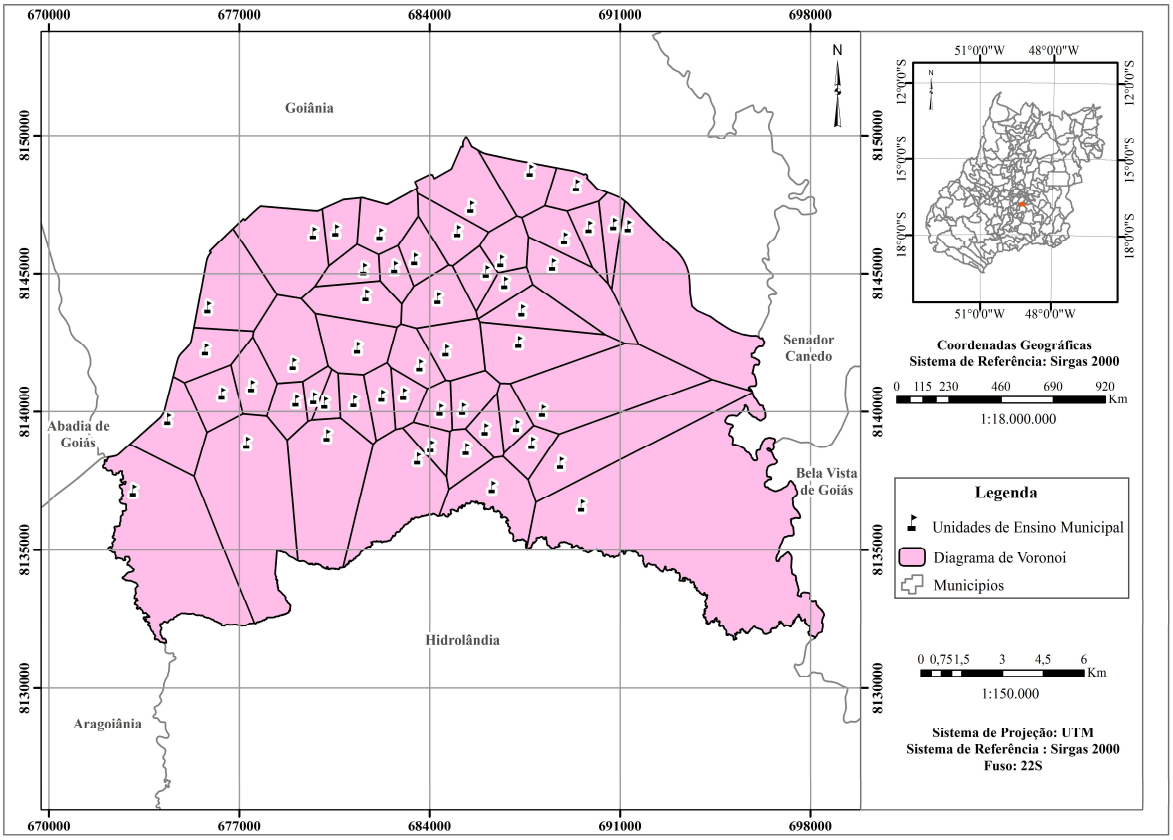


Fonte: A autora, 2019.

Esta abordagem de estimação dos pontos para os setores se mostrou adequada haja vista o teste que foi realizado utilizando diagramas de Voronoi, que considera um plano, contendo pontos que estão mais próximo de uma fonte do que de outra resultando em um polígono onde a distância entre a fonte e os pontos são as menores, para determinar área de abrangência de cada escola que possuía um dado real do IDEB.

A sobreposição dos valores estimados pelo diagrama de Voronoi e a utilização do vizinho mais próximo foi de 92 %, bem próximo de 100% o que indica a coerência da abordagem utilizada. A Figura 14 apresenta o diagrama calculado para o IDEB.

Figura 14 – Diagrama de Voronoi.

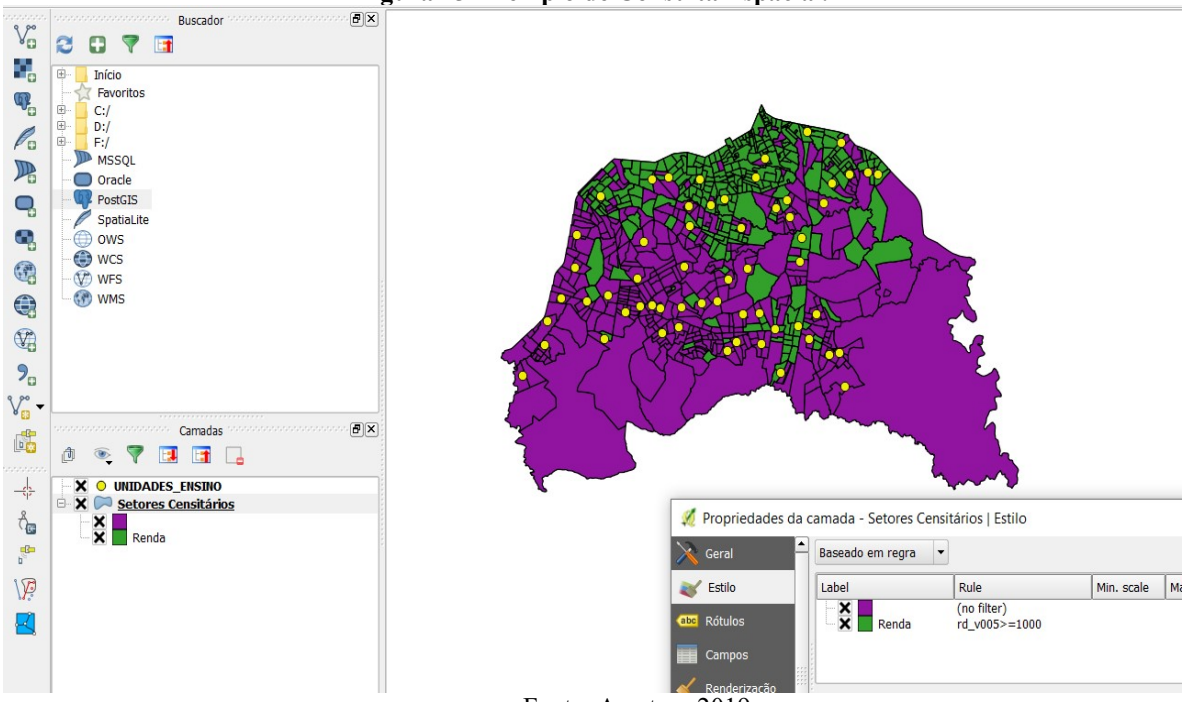


Fonte: A autora, 2019.

5.1 BANCO DE DADOS GEOGRÁFICOS E A INTEGRAÇÃO COM O AMBIENTE SIG

Foram feitas algumas consultas espaciais simples, com os dados que foram inseridos no Bando de Dados Geográfico para testagem do mesmo. Alguns exemplos são apresentados a seguir para mostrar visualmente os dados e sua distribuição pelos setores censitários.

Figura 15-Exemplo de Consulta Espacial.



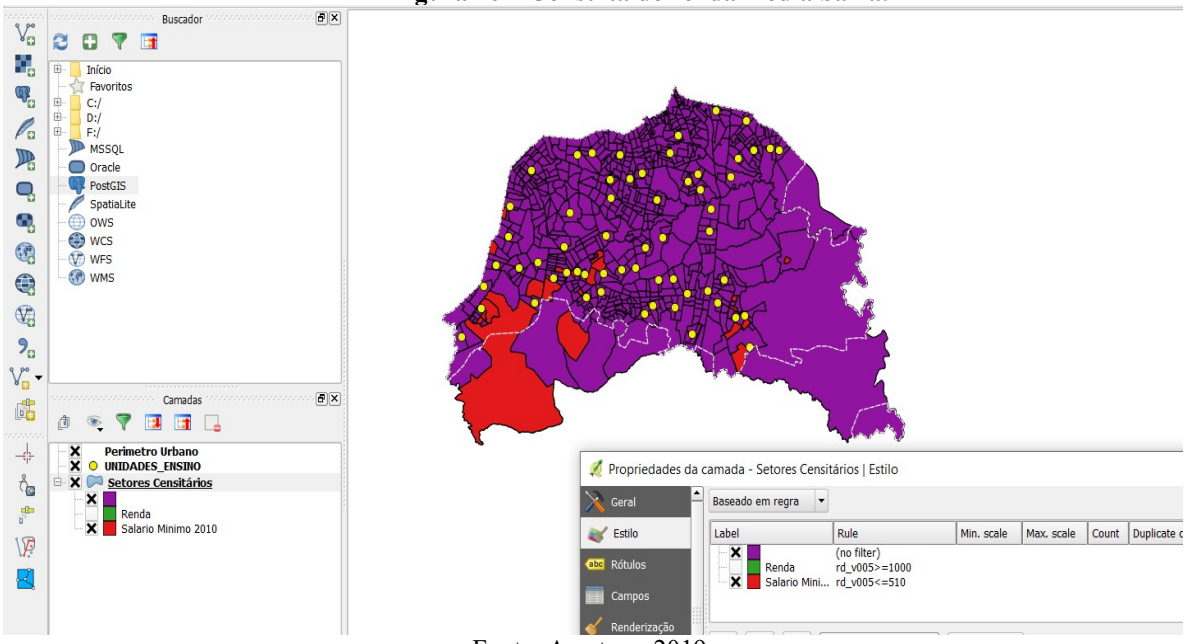
Fonte: A autora,2019.

A Figura 15 mostra um exemplo simples de consulta espacial em ambiente SIG, no caso desse trabalho no sofware QGIS. Por meio de uma regra foi solicitado que apresentasse na tela os Setores Censitários onde as rendas médias fossem maiores que R$1000,00. O Banco de Dados busca a informação e o SIG apresenta na tela. Pela região destacada, é perceptivel que a renda maior que R$1000,00 em 2010 ficou concentrada na parte norte da cidade que faz divisa com Goiânia, uma área bastante adensada e parte de trechos no decorrer da BR-153 .

Para a consulta seguinte foram selecionados os setores onde a renda média das familias eram menor ou igual a um salário minimo, que em 2010 era no valor de R$510,00, como mostra a Figura 16.

São poucos os setores censitários com esse salário, e os lugares onde foram apontados são em sua maioria áreas rurais e periféricas. A linha branca simboliza o Perímetro Urbano do Munícipio.

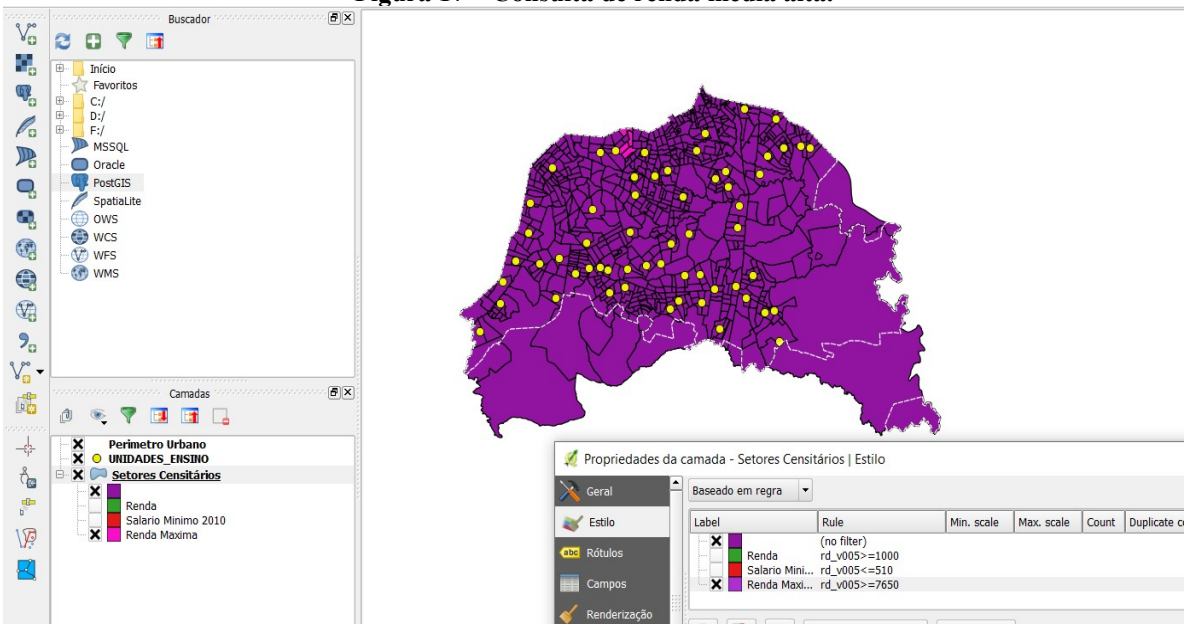
Figura 16 – Consulta de renda média baixa.



Fonte: A autora, 2019.

Para finalizar as análises sobre Renda média, foi criada uma regra para selecionar os Setores Censitários com rendimentos iguais ou superiores a quinze salários minimos, que na epoca daria um valor de R$7650,00. A Figura 17 mostra a seleção de cor rosa, são dois setores que ficam ao norte na divisa com Goiânia e que possuem condominios fechados, o que justifica o poder aquisitivo da média das familias que ali residiam.

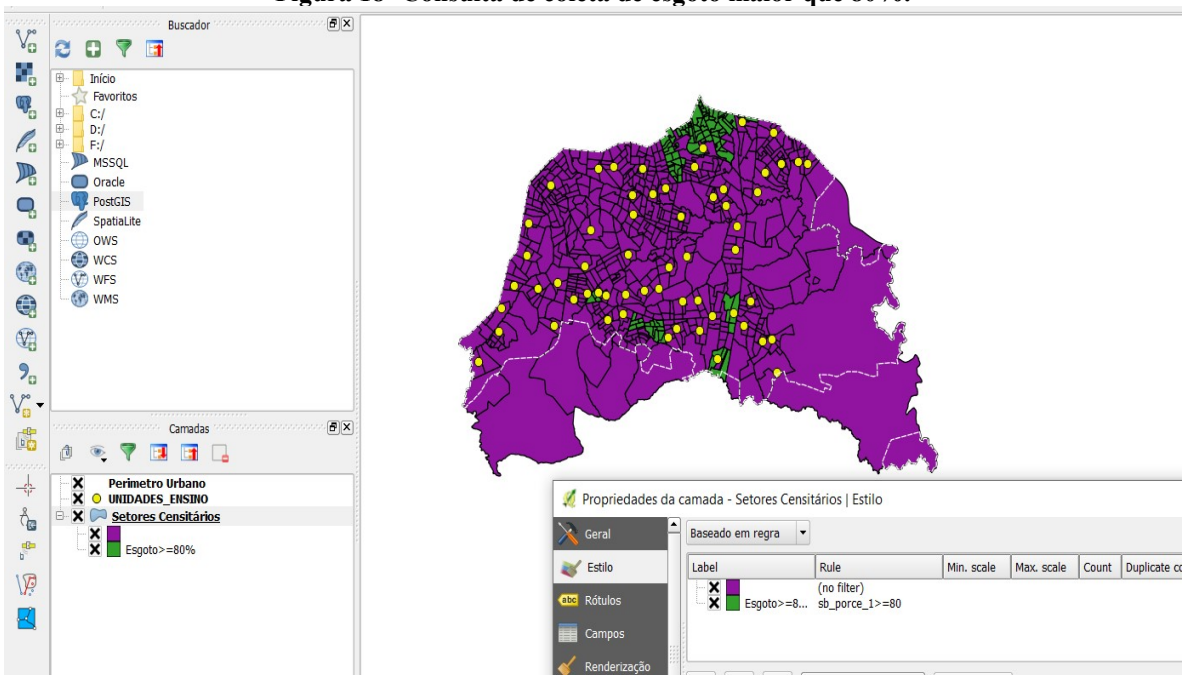
Figura 17 – Consulta de renda média alta.



Fonte: A autora, 2019.

A proxima consulta foi relacionada a coleta de esgoto no Municipio. A regra usada foi: quais Setores Censitários possuem uma taxa igual ou maior de 80% de coleta de esgoto. A Figura 18 mostra as regiões selecionadas.

Figura 18- Consulta de coleta de esgoto maior que 80%.

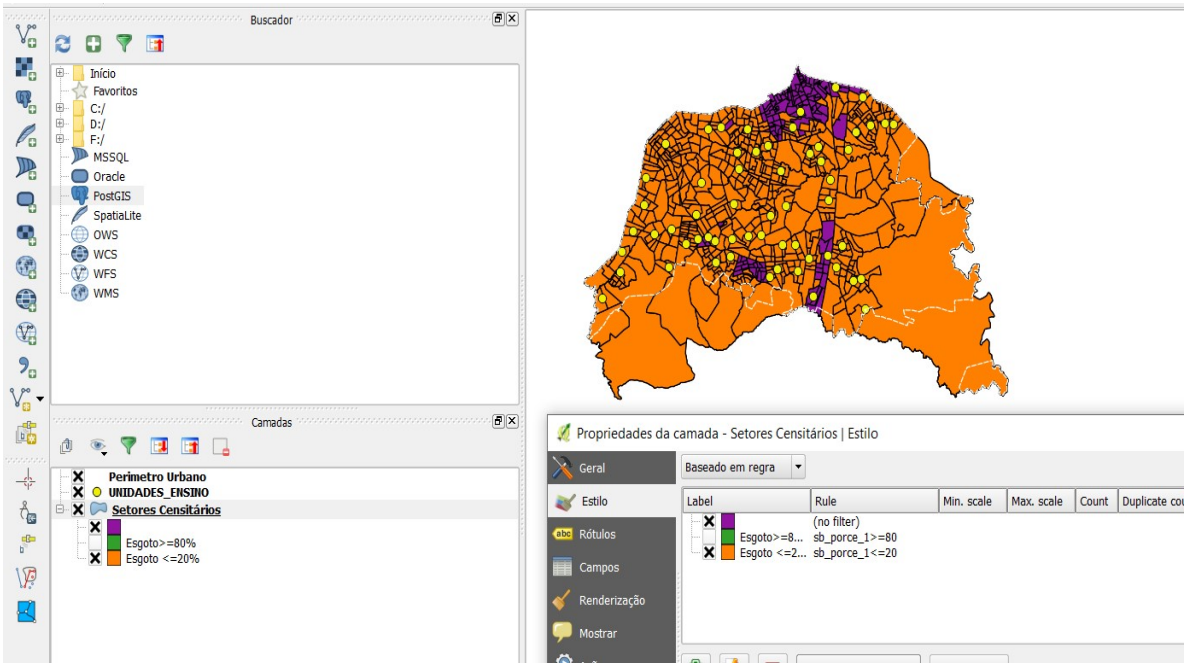


Fonte: A autora, 2019.

A parte norte da cidade e algumas regiões próximas ao centro do municipio eram beneficiadas, em 2010, com essa taxa de 80% de coleta de esgoto. São áreas bem adensadas e apenas quatro escolas possuiam esse beneficio (Escola Municipal Deraldo Lisboa dos Santos que possui IDEB de 5,5; Escola Municipal o Pequenino com IDEB de 6,0; Escola Municipal Jardim Bela Vista com IDEB 6,0 e Escola Municipal Guiomar Rosa de Oliveira com IDEB 5,9).

A Figura 19 mostra como o municipio no ano de 2010 era deficiênte na coleta de esgoto. A regra utilizada foi pedindo as regiões com a taxa igual ou menor que 20%. Toda a área em laranja entra nessa categoria, ou seja praticamente o municipio inteiro não possuía coleta de esgoto em 2010. A região norte próximo ao Buriti Shopping e a área dos Correios são uma das poucas que possuiam coleta de esgoto e uma pequena parte na região do Setor Central.

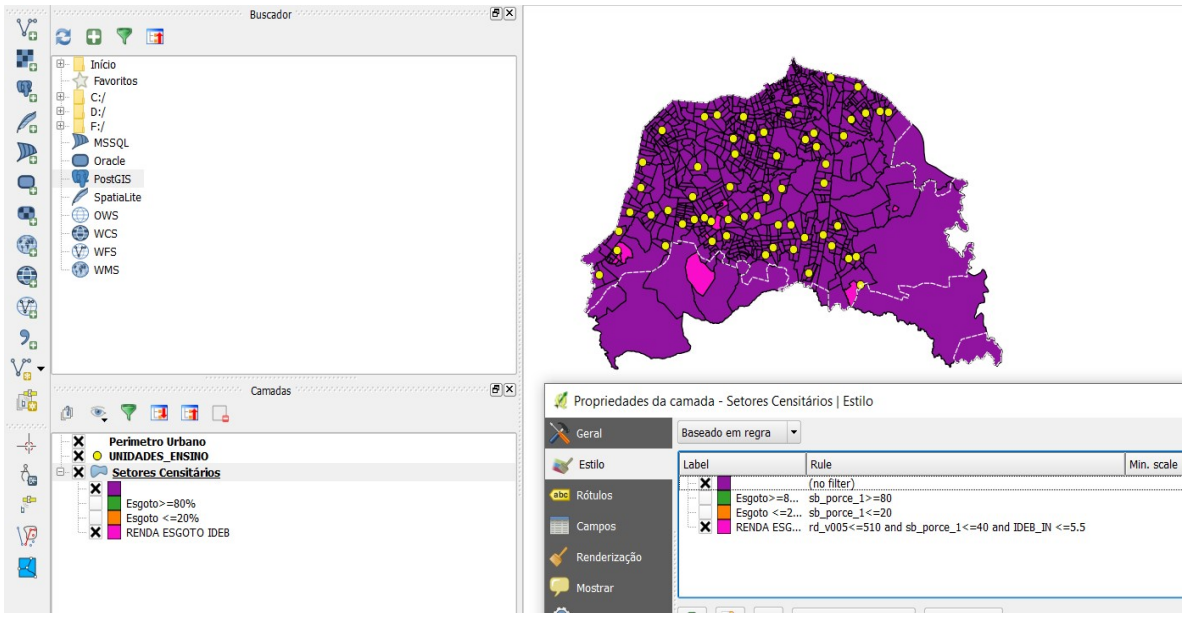
Figura 19 – Consulta de coleta de esgoto menor que 20%.



Fonte: A autora, 2019.

Para chegar no resultado na Figura 20, utilizamos os valores de renda menor ou igual R$510,00 (equivalente a um salário minimo em 2010), a porcentagem das residências com esgoto menor ou igual a 40% e as notas do IDEB menores ou iguais a 5,5.

Figura 20 – Cruzamento de dados de Renda, IDEB e Esgoto.

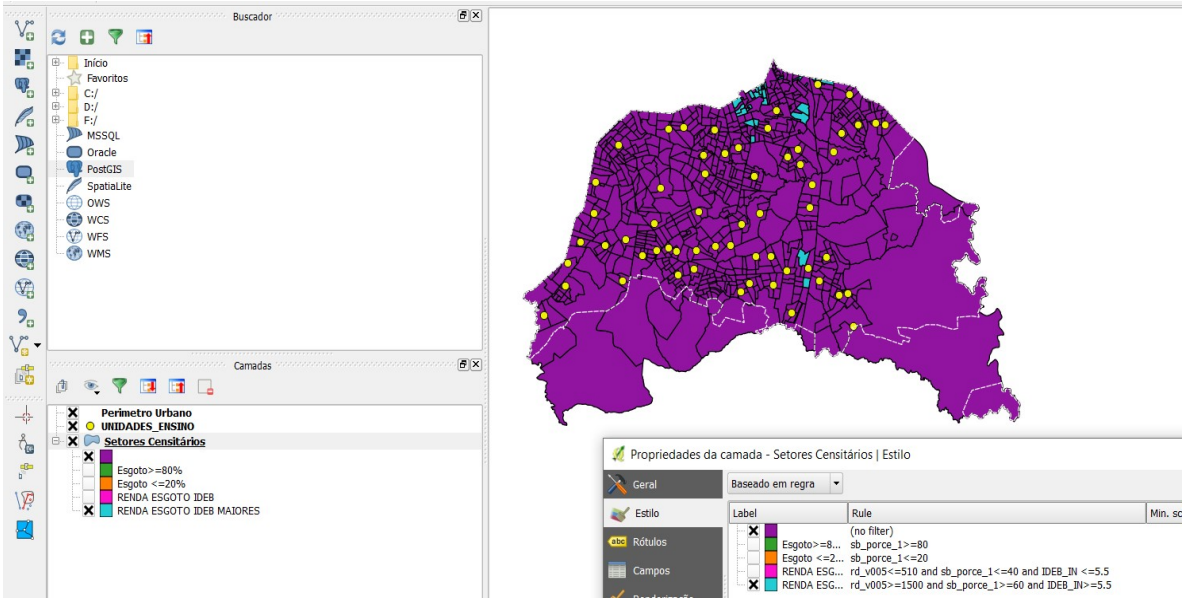


Fonte: A autora, 2019.

No total foram 10 Setores Censitários selecionados em áreas com baixo adensamento, sendo um setor localizado em área rural.

Na Figura 21 utilizamos os valores de renda maior ou igual a R$1500, a porcentagem de residências com esgoto maior ou igual a 60% e as notas do IDEB maiores ou iguais a 5,5.

Figura 21- Cruzamento de altos valores de renda, esgoto e IDEB.



Fonte : A autora, 2019.

Os Setores encontrados nessa regra estão em sua maioria ao norte do Município tendo apenas dois na região onde é o Setor Central, e nenhum em área rural.

5.2 MINERAÇÃO DE DADOS

Para esta etapa a proposta foi estabelecer possíveis associações entre as variáveis em discussão. Para tanto, um primeiro passo foi discretizar os valores de entrada para que o algoritmo de associação pudesse gerar as regras. O software utilizado foi o WEKA que permite utilizar diversas técnicas de mineração de dados. Entre elas, o algoritmo de associação APRIORI foi o escolhido. Pelo método 1, os dados socioeconômicos disponibilizados sob forma de valores numéricos foram discretizados conforme o Quadro 3.

Esta separação foi realizada em função da natureza dos dados. Os valores numéricos foram fatiados em intervalos que pudessem conter significativamente o conjunto de dados a ser representado.

Quadro 4 – Classes associadas às variáveis utilizadas.

RENDA MÉDIA				
MENOR OU IGUAL A 1 SALÁRIO	ENTRE 1 E 3 SALARIOS	ENTRE 3 E 5 SALÁRIOS	ENTRE 5 E 15 SALÁRIOS	MAIS QUE 15 SALÁRIOS
CLASSE E	CLASSE D	CLASSE C	CLASSE B	CLASSE A
LIXO E ESGOTO				
SE A %<= 20	SE A % >20 E <=40	SE A % >40 E <=60	SE A %>60 E <=80	SE A %>80
CLASSE E	CLASSE D	CLASSE C	CLASSE B	CLASSE A
ABASTECIMENTO DE ÁGUA				
SE A %<=20	SE A % >20 E <=40	SE A % >40 E <=60	SE A %>60 E <=80	SE A %>80
CLASSE E	CLASSE D	CLASSE C	CLASSE B	CLASSE A
COLETA DE ESGOTO				
SE A %<=20	SE A % >20 E <=40	SE A % >40 E <=60	SE A %>60 E <=80	SE A %>80
CLASSE E	CLASSE D	CLASSE C	CLASSE B	CLASSE A
COLETA DE LIXO				
SE A %<=20	SE A % >20 E <=40	SE A % >40 E <=60	SE A %>60 E <=80	SE A %>80
CLASSE E	CLASSE D	CLASSE C	CLASSE B	CLASSE A
MAPA DE CONFORTO AMBIENTAL				
BAIXO CONFORTO	BAIXO-MEDIO CONFORTO	MÉDIO-ALTO CONFORTO	ALTO CONFORTO	
SEGURANÇA PUBLICA				
SEM OCORRENCIA	BAIXA	MÉDIA	ALTA	
IDEB				
ENTRE 4,9 E 5,3	ENTRE 5,3 E 5,7	ENTRE 5,7 E 6,0	ENTRE 6,0 E 6,4	
CLASSIFICAÇÃO ESTRUTURA DAS ESCOLAS				
INSUFICIENTE	FRACO	MODERADO	BOM	

CLASSIFICAÇÃO MATERIAL DIDATICO DAS ESCOLAS			
INSUFICIENTE	FRACO	MODERADO	BOM

Fonte: A autora, 2019.

Alguns resultados obtidos com o algoritmo APRIORI são apresentados a seguir. Foram geradas 1000 (mil) regras de associação.

- IDEB=Entre_53_e_57:"COLETA=Sem_Coleta_Lixo_Esgoto_Classe_E, RENDA=Renda_Classe_D Classificacao_Conforto=Medio
- IDEB=Entre_57_e_60: "COLETA=Sem_Coleta_Lixo_Esgoto_Classe_E, RENDA=Renda_Classe_B
- IDEB=Entre_53_e_57: RENDA=Renda_Classe_C, ESGOTO=Esgoto_Classe_D,
- ESGOTO=Esgoto_Classe_E: IDEB=Entre_49_e_53,"RENDA=Renda_Classe_D"

Percebe-se que o APRIORI apresentou um padrão constante entre o IDEB, RENDA e ESGOTO o que mostra a relação com a qualidade socioeconômica da região. Nas áreas em que a renda é menor e a cobertura de esgoto não é suficiente o IDEB é baixo. O cenário oposto também é verdadeiro. Áreas com maior qualidade socioeconômica apresentam índices de IDEB maiores.

Outra análise com o minerador de dados foi a seleção de variáveis explicativas com relação à variável IDEB como mostra a Figura 22.

Figura 22- Seleção de Variáveis.

```
=== Attribute Selection on all input data ===

Search Method:
        Best first.
        Start set: no attributes
        Search direction: forward
        Stale search after 5 node expansions
        Total number of subsets evaluated: 45
        Merit of best subset found:    0.129

Attribute Subset Evaluator (supervised, Class (nominal):
        CFS Subset Evaluator
        Including locally predictive attributes

Selected attributes: 1,4,8,10 : 4
                     ESTRUTURA
```

Fonte: A autora, 2019.

Nesta análise utilizando o método Best Fisrt o minerador encontrou as variáveis: estrutura da escola, coleta de esgoto do setor censitário, segurança 2017 e classificação do conforto ambiental como as variáveis com maior poder explicativo das notas do IDEB.

5.3 ANÁLISE EXPLORATÓRIA

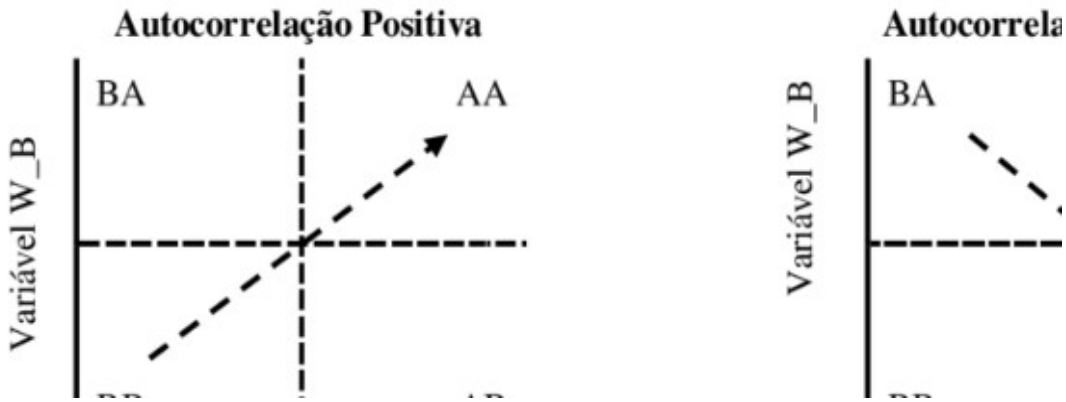
Na analise exploratória foram utilizadas as duas abordagens descritas anteriormente: uma com todos os setores censitários com as informações do IDEB adquiridas pelo método do vizinho mais próximo, e a outra abordagem onde foram utilizados apenas os setores censitários que possuem escolas.

Para se realizar as análises foi utilizado o Software livre GeoDa. Ele foi criado com a intenção de facilitar as análises de dados e os padrões espaciais. Dentro dele é possível fazer análises básicas de regressão espacial e autocorrelação espacial. (GeoDaCenter, 2019).

Câmara et all descreve que o Índice de Moran Global nos dá uma medida de associação espacial que existe dentro do conjunto de dados em questão, e que o Índice Local (LISA) mostra valores específicos para cada objeto, gerando assim os clusters que são os agrupamentos dos objetos que contém os valores dos atributos semelhantes.

Ramos, F.R (1999) diz que o Índice Global de Moran mostra a interdependência espacial de todos os conjuntos que estão em estudo, e já o Índice Local avalia a covariância entre um conjunto e uma vizinhança que definimos.O gráfico de do Índice de Moran é dividido em quatro quadrantes e por definição o primeiro quadrante é definido como baixo-alto, o segundo quadrante é alto-alto, o terceiro quadrante baixo-baixo e o quarto é alto-baixo como mostra a Figura 23.

Figura 23 – Quadrantes do Gráfico de Moran.



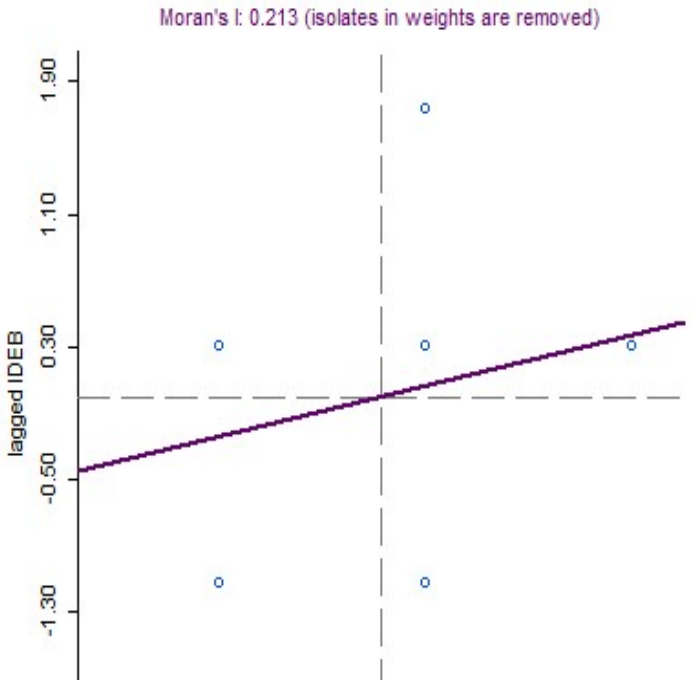
Fonte: Camara, M.R.G et all, 2013.

Seguem as analises exploratórias, primeiro dos setores censitários onde existem escolas.

5.3.1 ANÁLISE EXPLORATÓRIA DOS SETORES CENSITÁRIOS ONDE EXISTEM ESCOLAS.

O gráfico de Índice de Moran Local para os Setores Censitários onde existem escolas mostra que existe uma autocorrelação positiva como mostra a Figura 24. Seu valor foi de 0.213 que é um valor mais próximo à zero, podendo ser um indicativo de que não exista uma correlação espacial ou que essa correlação existe, mas é pequena.

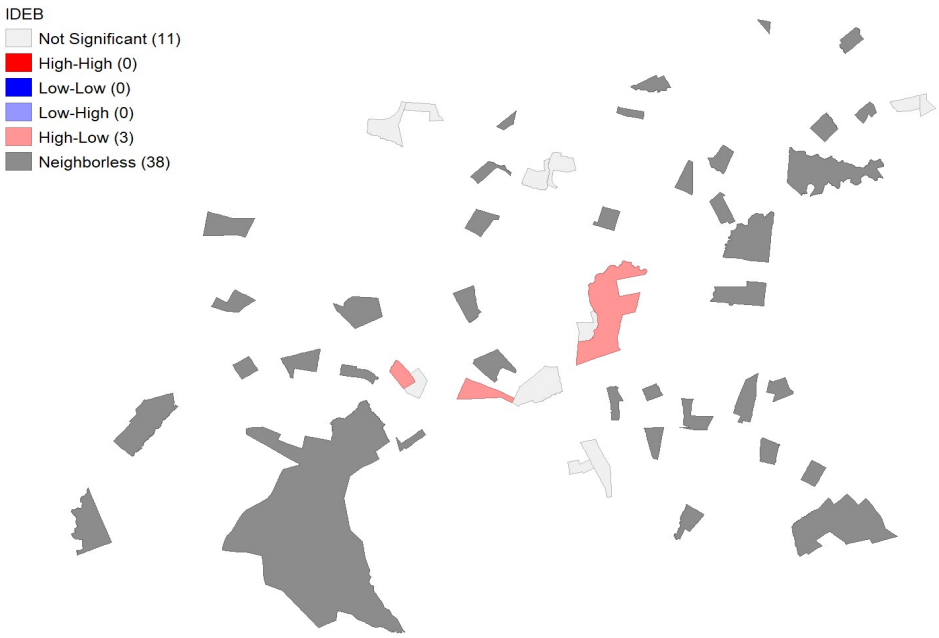
Figura 24 – Índice de Moran Local para onde existem escolas.



Fonte: A autora, 2019.

O Cluster apresentado na Figura 25 mostra que no caso dos setores censitários onde existem escolas a relação com a vizinhança é baixa e alguns são insignificantes porque não possuem uma proximidade, a maioria dos setores não tem vizinhos e estão espalhados pelo Município.

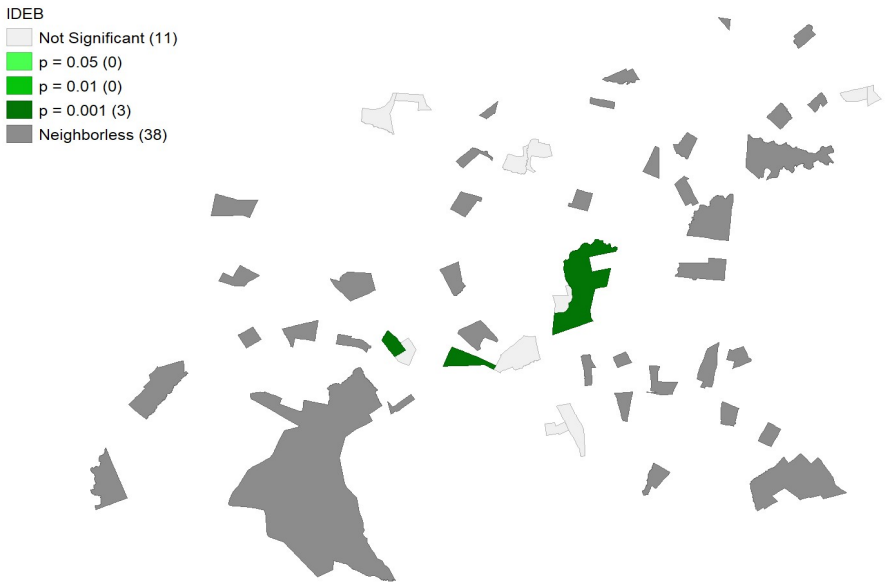
Figura 25– Mapa Cluster para onde existem escolas.



Fonte: A autora, 2019.

O mapa de significância ilustrado na Figura 26 tem como objetivo mostrar as áreas onde o indicador tem um nível de significância maior ou igual a 0.05. (Chagas, 2020). Nesse caso apenas três Setores obtiveram valor de p, que indica alguma significância das amostras.

Figura 26– Mapa de significância para onde existem escolas.



Fonte: A autora, 2019.

A Figura 27 mostra uma regressão clássica onde o IDEB foi utilizado como sendo a variável dependente e os dados de Roubo 2017, Roubo 2018, Abastecimento de Agua, Coleta de Esgoto, Coleta de Lixo, Renda Média, Dados de Conforto Ambiental e Residências sem Coleta de Lixo e Esgoto entraram como variáveis explicativas independentes.

Figura 27 – Regressão Clássica e 9 variáveis para onde existem escolas.

```
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  setores_censitarios_ESCOLAS_NUMERICO
Dependent Variable  :        IDEB  Number of Observations:    5
Mean dependent var  :       0.875  Number of Variables    :
S.D. dependent var  :   0.0675819  Degrees of Freedom     :   4

R-squared           :    0.192746  F-statistic            :
Adjusted R-squared  :    0.042559  Prob(F-statistic)      :
Sum squared residual:    0.191723  Log likelihood         :
Sigma-square        :  0.00445867  Akaike info criterion  :
S.E. of regression  :   0.0667733  Schwarz criterion      :
Sigma-square ML     :  0.00368698
S.E of regression ML:   0.0607205

------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic
------------------------------------------------------------------
          CONSTANT       0.687454       0.173138        3.97056
             R2017      0.0400229       0.261149       0.153257
             R2018       0.132322       0.301622       0.438703
          ABASAGUA      -0.024844      0.0282237      -0.880254
         COLESGOTO      0.0391677      0.0514932       0.760638
           COLLIXO      0.0917593       0.109982       0.834312
             RENDA       0.140415       0.556219       0.252446
         CONFORTO2      0.0986129       0.123796       0.796575
             SEMLE      -0.717991       0.833025      -0.861908
------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   53.823196
TEST ON NORMALITY OF ERRORS
TEST                     DF           VALUE             PROB
```

Fonte: A autora, 2019.

Como resultado tem se que: O Adjusted R-squared apresentou um valor de 0.042559 esse dado quer dizer o quanto da variação das variáveis dependentes é explicado pelas variáveis independentes. Ele varia entre 0 e 1, onde 0 quer dizer que as variáveis independentes não estão relacionadas com a variável dependente, e 1 significa que as variáveis independentes explicam a variação que acontece com a variável dependente. Então como o resultado está bem próximo à zero significa que essas variáveis não explicam o que ocorre com a variável dependente que é o IDEB.

O R-Squared (R- quadrado) é o resultado da regressão. Ele é a porcentagem da variação da variável dependente que é explicada por um modelo linear (Blog minitab, 2020). Nesse caso o modelo explica 19.27% da variável dependente.

A multicolinearidade apresentou o valor de 53.82. A professora e pesquisadora Flávia Feitosa da UFABC (Universidade Federal do ABC/SP) (via comunicação pessoal) justifica que os valores de multicolinearidade não devem ser maiores que 10, partindo desse pressuposto esse modelo possui uma multicolinearida alta. Este fato é um indicativo de que há variáveis com uma autocorrelação muito alta e isso acaba confundindo o modelo, é indicado que a regressão seja refeita com as variáveis que possuem menos autocorrelação.

O teste de não normalidade (Jarque Bera) apresentou o valor de p=0.52374. Na literatura o indicado é que esse valor seja menor ou igual a 0.05.

A heterocedasticidade realiza testes para ver se os resíduos da regressão não estão homogeneos. Se o valor de p for menor ou igual a 0.05 rejeita-se a hipotese nula porque os resíduos não estão homogeneos e isso indica algo de errado no modelo. Nessa regressão um dos valores foi de 0.13 que é maior que 0.05 e no segundo teste o valor foi de 0.017 que ficou a baixo de 0.05 então nesse teste temos evidencias contra a hipotese nula.

Nesse modelo apenas a constante que é o IDEB é significativa todas as outras não são, o valor de p delas está acima de 0.05.

Figura 28– Regressão Spatial Lag e 9 variáveis para onde existem escolas.

```
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMA
Data set            : setores_censitarios_ESCOLAS_NUMERICO
Spatial Weight      : setores_censitarios_ESCOLAS_NUMERICO
Dependent Variable  :        IDEB  Number of Observations:   52
Mean dependent var  :       0.875  Number of Variables   :   10
S.D. dependent var  :   0.0675819  Degrees of Freedom    :   42
Lag coeff.   (Rho)  :   0.0173656

R-squared           :    0.201211  Log likelihood        :    7
Sq. Correlation     : -            Akaike info criterion :   -1
Sigma-square        :  0.00364831  Schwarz criterion     :   -1
S.E of regression   :   0.0604013

------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value
------------------------------------------------------------------
           W_IDEB      0.0173656       0.023416       0.741611
         CONSTANT       0.679121       0.157274        4.31808
            R2017     -0.0178261         0.2473     -0.0720828
            R2018        0.20749       0.289432       0.716888
         ABASAGUA     -0.0238834       0.0255741      -0.933888
        COLESGOTO      0.0454973       0.0473313      0.961252
          COLLIXO      0.0977965        0.100004      0.977923
            RENDA       0.171667        0.505779      0.339412
        CONFORTO2      0.0905988        0.112478      0.805484
            SEMLE     -0.595165        0.773335      -0.769609
------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF      VALUE         PR
```

Fonte: A autora, 2019.

Na regressão espacial apresentada na Figura 28 o modelo explica apenas 20.12% da variação da variável dependente. A regressão Spatial Lag leva em conta a localização espacial das variáveis, além de atribuir uma autocorrelação espacial à variável dependente. Pode-se observar que o valor de Log Likelihood é baixo 72.16 o que indica que esse não é o melhor ajuste do modelo.Mais uma vez a única variável que é significativa é a constante, as outras não explicam a variável dependente.

A Figura 29 mostra uma Regressão Clássica que foi repetida, mas dessa vez com quatro variáveis: uma dependente e tres explicativas. A dependente é o IDEB e as independentes são: Roubo 2017, Coleta de Esgoto e Renda Media. Foi observado que o valor do R quadrado é de 13%, então modelo explica 13% da variavel dependente.

A multicolinearidade nesse caso ficou com um valor 8.46, um resultado melhor se comparado com a regressão onde foi utilizada nove variaveis. Miloca, S.A. e Conejo, P.D (2011) pontuam que uma das soluções para a multicolinearidade é a exclusão de algumas

variáveis independentes que sejam altamente correlacionadas, informação que corrobora com o que a pesquisadora Flávia Feitosa indicou. Isso mostra que nenhuma dessas variaveis está confundindo o modelo. Por isso foi feita uma regressão com nove variáveis e outra com quatro.

Figura 29– Regressão Clássica e 4 variáveis para onde existem escolas.

```
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  setores_censitarios_ESCOLAS_NUMERICO
Dependent Variable  :         IDEB  Number of Observations:    52
Mean dependent var  :        0.875  Number of Variables    :     4
S.D. dependent var  :    0.0675819  Degrees of Freedom     :    48

R-squared           :     0.130303  F-statistic            :      2
Adjusted R-squared  :     0.075947  Prob(F-statistic)      :    0.0
Sum squared residual:     0.206553  Log likelihood         :      6
Sigma-square        :   0.00430319  Akaike info criterion  :     -1
S.E. of regression  :    0.0655987  Schwarz criterion      :     -1
Sigma-square ML     :   0.00397218
S.E of regression ML:    0.0630252

------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic
------------------------------------------------------------------
          CONSTANT        0.86846      0.0314253        27.6357
             R2017       0.141225      0.0546148        2.58583
         COLESGOTO    0.00545578       0.041741       0.130706
             RENDA     -0.238075       0.437046      -0.544737
------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER    8.460071
TEST ON NORMALITY OF ERRORS
TEST                    DF            VALUE              PROB
Jarque-Bera              2              2.2047           0.33208

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                    DF            VALUE              PROB
Breusch-Pagan test       3              3.2786           0.35063
Koenker-Bassett test     3              6.5591           0.08736

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : setores_censitarios_ESCOLAS_NUMERICO
   (row-standardized weights)
```

Fonte: A autora, 2019.

A Figura 30 mostra uma nova regressão spatial lag e esse modelo consegue explicar a variação da variavel dependente em 13.81%. No campo de Probability observamos quais variáveis foram significativas nesse modelo, nesse caso Roubos 2017 e a constantes foram as variáveis significativas.

Figura 30– Regressão Spatial Lag e 4 variáveis para onde existem escolas.

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMA
Data set            : setores_censitarios_ESCOLAS_NUMERICO
Spatial Weight      : setores_censitarios_ESCOLAS_NUMERICO
Dependent Variable  :        IDEB  Number of Observations:   52
Mean dependent var  :       0.875  Number of Variables   :    5
S.D. dependent var  :   0.0675819  Degrees of Freedom    :   47
Lag coeff.   (Rho)  :   0.0154277

R-squared           :    0.138127  Log likelihood        :
Sq. Correlation     : -            Akaike info criterion :   -1
Sigma-square        :  0.00393644  Schwarz criterion     :    -
S.E of regression   :   0.0627411

------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value
------------------------------------------------------------
           W_IDEB     0.0154277      0.0225325       0.684688
         CONSTANT      0.864991      0.0305875        28.2793
            R2017      0.144473      0.0524734        2.75326
        COLESGOTO    0.00922489      0.0402552        0.22916
            RENDA     -0.252083       0.418195      -0.602789
------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF      VALUE        PR
```

Fonte: A autora, 2019.

Para avaliar qual é o melhor modelo é utilizado os resultados do Critério de Informação de Akaike, ele mostra qual a regressão é mais confiavel. Os valores obtidos nessa primeira abordagem onde foram utilizados os dados dos Setores Censitários onde existem escolas foram mostrados no Quadro 5.

Quadro 5 – Comparação das regressões para onde existem escolas.

Regressão Clássica utilizando 9 variaveis	Regressão Spatial Lag utilizando 9 variaveis
Akaike info criterion = -125.784	**Akaike info criterion = -124.328**
Regressão Clássica utilizando 4 variaveis	Regressão Spatial Lag utilizando 4 variaveis
Akaike info criterion = -131.909	**Akaike info criterion = -130.376**

Fonte: A autora, 2019.

Os resultados motram que o menor valor para regressão utilizando nove ou quatro variáveis são da regressão clássica, nesse caso ela é mais confiavel que a regressão spatial lag. Nesse caso a localização espacial não trouxe mais confiabilidade para o modelo.

No Quadro 6 foi apresentada uma comparação com os resultados do R-quadrado para se analisar qual dos modelos conseguiu ou não explicar a variavel IDEB.

Quadro 6 - Comparação dos R² para onde existem escolas.

R- quadrado da Regressão Clássica utilizando nove variáveis.	**19.27%**
R- quadrado da Regressão Spatial Lag utilizando nove variáveis.	**20.12%**
R- quadrado da Regressão Clássica utilizando quatro variáveis.	**13.03%**
R- quadrado da Regressão Spatial Lag utilizando quatro variáveis.	**13.81%**

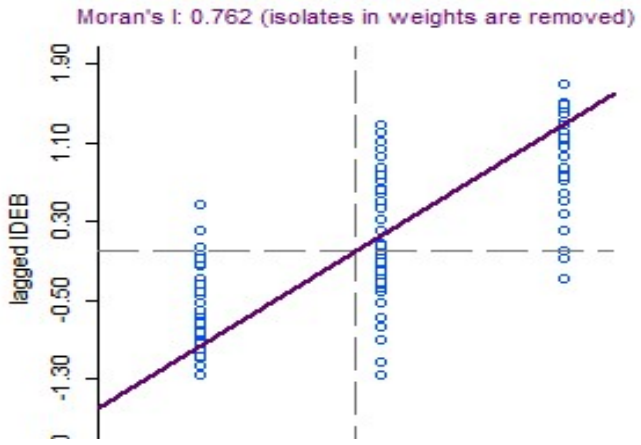
Fonte: A autora, 2019.

É observável que nenhum R-quadrado chegou proximo a 100%, indicando assim que esses modelos não explicam a variável dependente IDEB. A variável independente Roubo 2017 apareceu como sendo uma variável significante nos modelos onde foram utilizadas apenas quatro variáveis.

5.3.2 ANÁLISE EXPLORATÓRIA UTILIZANDO TODOS OS SETORES CENSITÁRIOS.

O índice de Moran local apresentou um valor de 0.762 mostrado na Figura 31, e isso indica que existe uma grande correlação espacial do IDEB e que possui uma autocorrelação positiva. Almeida (2012) argumenta que se a autocorrelação espacial é positiva significa que exista uma similaridade entre os valores dos atributos e a localização espacial. Este resultado já era esperado porque os valores foram “projetados” para os setores vizinhos mais próximos.

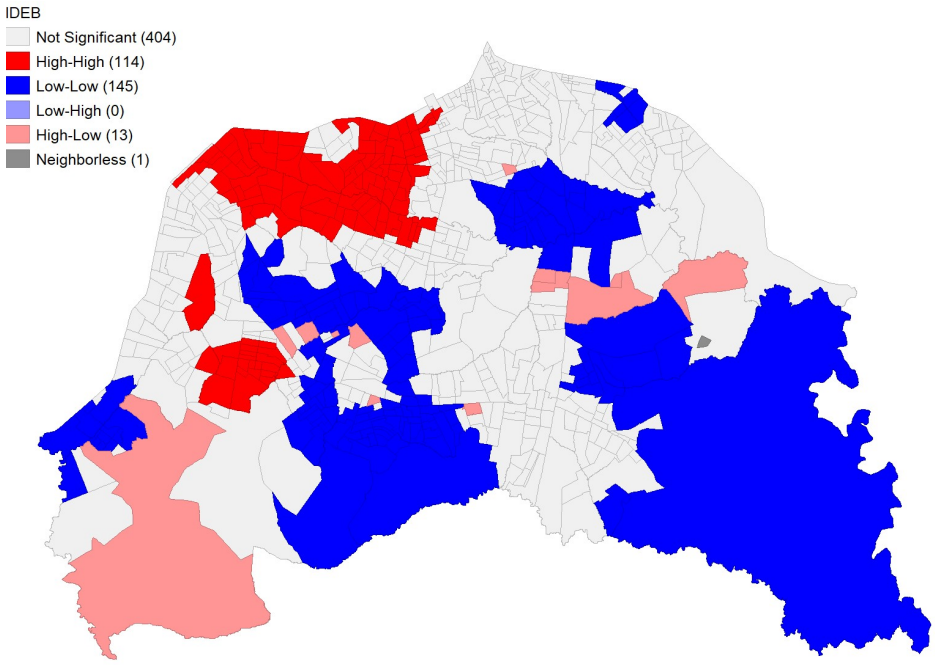
Figura 31– Índice de Moran Local de todos os Setores Censitários.



Fonte: A autora, 2019.

As áreas em vermelho mostradas no mapa de Cluster Figura 32 mostram onde o indicador e a vizinhança são mais altos. Olhando para a cidade percebemos que isso ocorre nas áreas mais ao norte que fazem divisa com Goiânia e que são áreas adensadas e os setores censitários são menores. Em azul são os clusters de baixos valores e estão em setores censitários maiores, menos adensados, e em glebas rurais ou áreas impróprias para ocupação além de serem mais periféricos. A maioria dos polígonos não possui significância.

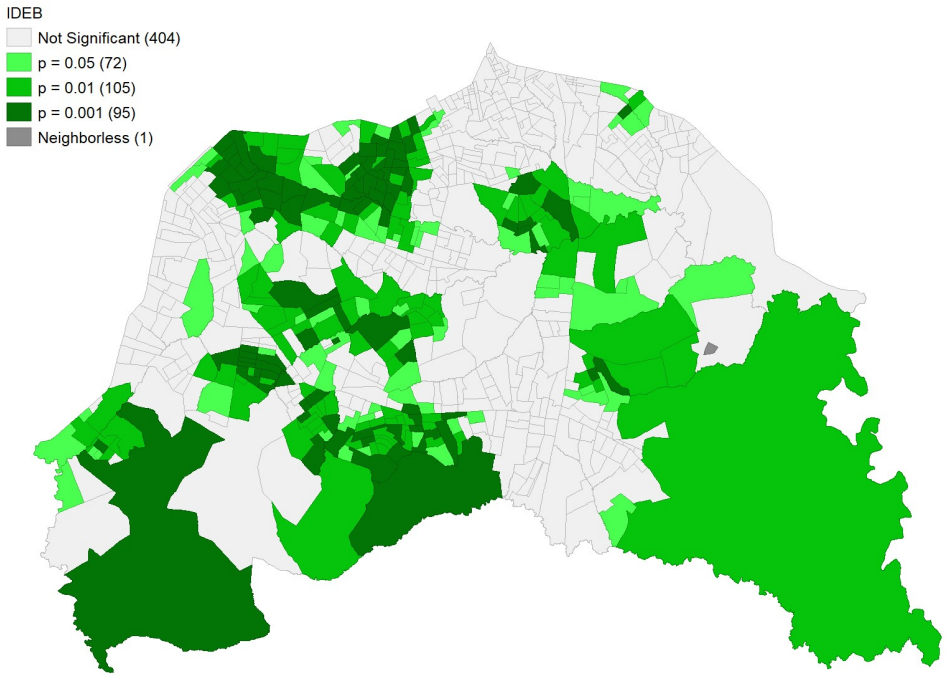
Figura 32– Mapa Cluster de todos os Setores Censitários.



Fonte: A autora, 2019.

Na abordagem de utilizar todos os Setores Censitários como apresenta a Figura 33, foram encontrados 72 polígonos com o nível de significância maior ou igual a 0.05. Os valores menores que 0.05 indicam a evidencia de que exista a hipótese nula. Nesse caso existiu uma maior significância do que na abordagem utilizando somente os setores onde existem escolas.

Figura 33 – Mapa de Significância de todos os Setores Censitários.



Fonte: A autora, 2019.

Na Regressão Clássica apresentada na Figura 34, utilizando nove variáveis e todos os Setores Censitários do Município é observado que o valor do Adjust R-squared e do R-squared continuam muito proximos de zero, indicando que mesmo utilizando todos os Setores as variaveis independentes não explicam a variavel dependente IDEB.

O valor do R-squared (R^2) explica o modelo em 5.74%, um valor muito pequeno. Como foram usadas muitas variaveis que podem ter confundido o modelo, a multicolinearidade ficou quase em 30. Aqui a variável renda média é significativa.

Figura 34– Regressão Clássica e 9 variáveis para todos os Setores Censitários.

```
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  setores_censitarios
Dependent Variable  :        IDEB  Number of Observations:  677
Mean dependent var  :    0.885377  Number of Variables   :    9
S.D. dependent var  :   0.0675071  Degrees of Freedom    :  668

R-squared           :    0.057453  F-statistic           :     5.
Adjusted R-squared  :    0.046165  Prob(F-statistic)     :3.64091
Sum squared residual:     2.90797  Log likelihood        :     88
Sigma-square        :  0.00435325  Akaike info criterion :    -17
S.E. of regression  :   0.0659792  Schwarz criterion     :    -17
Sigma-square ML     :  0.00429538
S.E of regression ML:   0.0655392

------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic
------------------------------------------------------------------------
          CONSTANT      0.869085      0.0235649        36.8806
             R2017      0.116831      0.0556083        2.10097
             R2018    -0.0784881      0.0622059       -1.26175
          ABASAGUA    0.00890114     0.00794379        1.12052
         COLESGOTO   -0.00812848     0.00879862      -0.923836
           COLLIXO     0.0054557      0.0132852        0.41066
             RENDA      0.109954      0.0372059        2.95529
         CONFORTO2    -0.0159141      0.0254187      -0.626079
             SEMLE    -0.0519945      0.0340441       -1.52727
------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   29.260675
TEST ON NORMALITY OF ERRORS
TEST                 DF           VALUE             PROB
```

Fonte: A autora, 2019.

Nessa regressão spatial lag apresentada na Figura 35 o modelo explica 64.09% a variação da variavel IDEB. Um valor alto de Log Likelihood indica que o modelo de ajuste de regressão pode estar adequado, nesse caso ele ficou com o valor 1168.49 que é considerado alto. Quanto menor for o Akaike info criterion mais confiavel é o modelo, o valor encontrado foi de 2316.98. Nesse modelo existe duas variáveis significativas que são o conforto ambiental e a coleta de lixo.

Figura 35– Regressão Spatial Lag e 9 variáveis para todos os Setores Censitários.

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMA
Data set            : setores_censitarios
Spatial Weight      : setores_censitarios
Dependent Variable  :        IDEB  Number of Observations:  677
Mean dependent var  :    0.885377  Number of Variables   :   10
S.D. dependent var  :   0.0675071  Degrees of Freedom    :  667
Lag coeff.   (Rho)  :    0.728837

R-squared           :    0.640890  Log likelihood        :    1
Sq. Correlation     : -            Akaike info criterion :   -2
Sigma-square        :  0.00163654  Schwarz criterion     :   -2
S.E of regression   :   0.0404541

------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value
------------------------------------------------------------------
           W_IDEB      0.728837      0.0258934        28.1476
         CONSTANT      0.317319      0.0261685         12.126
           R2017    0.00675775      0.0342723       0.197178
           R2018   -0.00411294      0.0382796      -0.107445
         ABASAGUA    0.00119516     0.00487584       0.245118
        COLESGOTO   -0.00483491     0.00539475      -0.896225
          COLLIXO    -0.0213021     0.00815758       -2.61132
            RENDA  -0.000987653      0.0228237     -0.0432732
        CONFORTO2    -0.0743324      0.0155928       -4.76709
            SEMLE    -0.0401173      0.0208789       -1.92142
------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF      VALUE        PR
```

Fonte: A autora, 2019.

A Figura 36 mostra que o resultado do modelo de regressão classica utilizando quatro variáveis foi de 4.51%, ou seja esse modelo não consegue explicar a variação da variavel dependente. É possível observar, nessas condições, que a multicolinearidade tem valor de 3.89 pressupondo que o modelo respondeu melhor com estas variáveis. Nesse modelo é observado que a variável independente roubo é significativo.

Figura 36 – Regressão Clássica e 4 variáveis para todos os Setores Censitários.

```
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  setores_censitarios
Dependent Variable  :        IDEB  Number of Observations:  677
Mean dependent var  :    0.885377  Number of Variables   :    4
S.D. dependent var  :   0.0675071  Degrees of Freedom    :  673

R-squared           :    0.045081  F-statistic           :
Adjusted R-squared  :    0.040825  Prob(F-statistic)     :8.224
Sum squared residual:     2.94614  Log likelihood        :
Sigma-square        :  0.00437763  Akaike info criterion :    -
S.E. of regression  :   0.0661636  Schwarz criterion     :    -
Sigma-square ML     :  0.00435176
S.E of regression ML:   0.0659679

------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic
------------------------------------------------------------
          CONSTANT      0.863249     0.00482196        179.024
             R2017     0.0530601      0.0115376        4.59889
         COLESGOTO   -0.00248162     0.00757105      -0.327777
             RENDA      0.123067      0.0363999        3.38097
------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   3.890481
TEST ON NORMALITY OF ERRORS
TEST                   DF           VALUE              PROB
Jarque-Bera             2            22.6445           0.00001

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                   DF           VALUE              PROB
Breusch-Pagan test      3            29.3950           0.00000
Koenker-Bassett test    3            49.5909           0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : setores_censitarios
```

Fonte: A autora, 2019.

A Figura 37 apresenta a regressão spatial lag levando em consideração quatro variaveis independentes, o valor do R-squared foi de 62.00%. Na heterocedasticidade o valor de p foi de 0.00000 então a hipotese nula foi rejeitada, indicando que possa existir um problema no modelo porque os residuos não estão homogeneos.

Figura 37– Regressão Spatial Lag e 4 variáveis para todos os Setores Censitários.

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMA
Data set            : setores_censitarios
Spatial Weight      : setores_censitarios
Dependent Variable  :        IDEB  Number of Observations:  677
Mean dependent var  :    0.885377  Number of Variables   :    5
S.D. dependent var  :   0.0675071  Degrees of Freedom    :  672
Lag coeff.   (Rho)  :    0.710798

R-squared           :    0.619940  Log likelihood        :    1
Sq. Correlation     : -            Akaike info criterion :   -2
Sigma-square        :  0.00173201  Schwarz criterion     :   -2
S.E of regression   :   0.0416174

------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value
------------------------------------------------------------------
           W_IDEB      0.710798      0.0274605        25.8844
         CONSTANT      0.254793      0.0243626        10.4584
           R2017    0.00872418     0.00727604        1.19903
        COLESGOTO   -0.00012053     0.00476245     -0.0253084
           RENDA     0.0024548      0.0229066       0.107166
------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF      VALUE        PF
Breusch-Pagan test                      3        51.5365     0.
```

Fonte: A autora, 2019.

O Quadro 7 apresenta os valores das regressões para a segunda abordagem, onde foram utilizados dados de todos os Setores Censitários do Municipio.

Quadro 7 – Comparação das regressões de todos os Setores Censitários.

Regressão Clássica utilizando 9 variaveis	Regressão Spatial Lag utilizando 9 variaveis
Akaike info criterion = -1750,55	**Akaike info criterion = -2316,98**
Regressão Clássica utilizando 4 variaveis	Regressão Spatial Lag utilizando 4 variaveis
Akaike info criterion = -1751,72	**Akaike info criterion = -2293,93**

Fonte: A autora, 2019.

Para a utilização de todos os Setores Censitários a regressão spatial lag aprensentou valores menores tanto para quando foi utilizado nove variáveis ou quando foram utilizadas quatro.

O Quadro 8 apresentou os valores de R-quadrado para as quatro Regressões feitas com os dados de todos os Setores Censitários.

Quadro 8 - Comparação dos R² de todos os Setores Censitários.

R- quadrado da Regressão Clássica utilizando nove variáveis.	**5.74%**
R- quadrado da Regressão Spatial Lag utilizando nove variáveis.	**64.09%**
R- quadrado da Regressão Clássica utilizando quatro variáveis.	**4.51%**
R- quadrado da Regressão Spatial Lag utilizando quatro variáveis.	**62.00%**

Fonte: A autora, 2019.

As Regressões Spatial Lag chegaram em uma porcentagem maior que 60%. Então elas explicam melhor a variabilidade das variáveis em torno da média. Em contrapartida os valores das Regressões Clássicas foram muito baixos, um apresentou o valor de 5.74% e o outro de 4.51%, então eles não conseguem explicar a relação com a variável dependente. Na Regressão Spatial Lag com as nove variáveis obteve-se o valor de 64.09% de R-quadrado. Apareceram duas variáveis independentes que foram significativas para o modelo, ou seja interferem no IDEB, foram elas o Conforto Ambiental e a Coleta de Lixo. Um lugar com um conforto ligado ao senso comum e que possui uma coleta de lixo regular interfere nas notas do IDEB.

6. CONCLUSÃO

Com todas as análises realizadas percebe-se que as consultas espaciais no ambiente SIG facilitam a compreensão de dados porque os associam visualmente. A utilização do Banco de Dados Geográfico é um avanço para a administração pública porque permite que vários órgãos trabalhem em cima dos mesmos dados ao mesmo tempo, isso traz uma agilidade interna nas atualizações e também auxilia o atendimento aos contribuintes.

A mineração de dados apresentou padrões, trazendo à tona associações interessantes dentro do conjunto de variáveis disponíveis. E confirmando essas associações da mineração, as análises exploratórias mostraram o quanto variáveis com autocorrelação podem atrapalhar o modelo de regressão.

Em relação a analise exploratoria dentro do GeoDa é possível concluir que para a primeira abordagem onde foram utilizados os dados apenas dos Setores Censitários onde existem escolas a Regressão Clássica foi a mais confiável tendo seus valores do Critério de Informação de Akaike menores. Seu R-quadrado não conseguiu explicar uma relação das variáveis independentes com a variável depentente IDEB.

Já na segunda abordagem quando foi utilizado os dados de todos os Setores Censitários do Município a Regressão mais confiável foi a Spatial Lag que leva em conta a posição no espaço dos dados, o R-quadrado delas conseguiu explicar 64.09% a relação da variável dependente IDEB com as variáveis independentes. Nessas Regressões Spatial Lag também foi observado que as variáveis independentes Conforto Ambiental e Coleta de Lixo tem uma relação com o IDEB.

O Índice de Moran Local apontou uma autocorrelação positiva do IDEB quando utilizamos todos os Setores Censitários e uma autocorrelação negativa quando foram utilizados apenas os Setores onde existem escolas.

As técnicas aqui apresentadas podem ser consideradas um “arcabouço” de procedimentos para avaliar um problema específico que envolva alguma variável no espaço. Embora as análises não apresentem um resultado único elas provocam a discussão do problema confrontando variáveis socioeconômicas, ambientais e educacionais.

7. REFERÊNCIAS

AGRAWAL, R., SRIKANT, R.: Fast Algorithms for Mining Association Rules. Proc, 20 th Int. Conf. Very Large Data Bases, VLDB, 1994

ALMEIDA, E. Econometria espacial aplicada. CAP.4. Campinas –SP. Editora Alinea,2012. Disponível em:< http://www.usp.br/nereus/wp-content/uploads/Almeida_cap_4.pdf> Acesso em: Dezembro, 2019.
Alves, W.P. Fundamentos de Banco de Dados. 1ª Ed. São Paulo: Érica, 2004. Pg:19 a 23.

BLOG MINITAB. Analise de Regressão. Disponível em: < https://blog.minitab.com/pt/analise-de-regressao-como-interpretar-o-r-quadrado-e-avaliar-a-qualidade-de-ajuste> Acesso em: Janeiro, 2020.

CALIXTO, K.E.A.; GUSMÃO, R.P.de; SEGUNDO, C.V.N. Mineração de dados aplicada a educação: um estudo comparativo acerca das características que influenciam a evasão escolar. VI Congresso Brasileiro de Informática na Educação-CBIE.DOI: 10.5753/cbie.sbie.2017.1447.

CÂMARA, G.; QUEIROZ, G.R. Arquitetura de Sistema de Informação Geográfica. São José dos Campos, INPE-8078-PRE/3893.

CÂMARA, G.; CAMARGO, E.C.G.; MONTEIRO, A.M.; NEVES, M.C.; RAMOS, F.R. Análise Exploratória Espacial de Dados Sócio-Econômicos de São Paulo. GISBRASIL 2000, Salvador, 2000. Disponível em :< http://www.dpi.inpe.br/gilberto/papers/marcos_gisbrasil2000.pdf> Acesso: Dezembro de 2019.

CARNEIRO, A.M.C. Uso do modelo de mistura MESMA e imagens ASTER para construção de um mapa de conforto urbano para Belo Horizonte - MG. Dissertação de Mestrado. Instituto de Geociências da UFMG. 2009.

CARVALHO, M.S.; SANTOS, R.S. Análise de dados espaciais em saúde pública: métodos, problemas, perspectivas. Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, 2005.

CHAGAS, A.L.S. Econometria Espacial Aula 4: GeoDA. Nereus (Núcleo de Economia Regional e Urbana da Universidade de São Paulo). Disponivel em: < https://edisciplinas.usp.br/pluginfile.php/4265581/mod_resource/content/0/Aula%204.pdf> Acesso: Janeiro, 2020.

DATE, C.J. Introdução a Sistemas de Banco de Dados. Rio de Janeiro. Elsevier, 2004.

ENNES, R.; IMAI, N.N.; TACHIBANA, V.M.; VICENTIN, D.M. Aplicação de Técnicas Utilizando Regressão Clássica e Espacial na Cidade de Presidente Prudente-SP. II Simpósio Brasileiro de Geomática. ISSN 1981-6251, p. 719-725. Julho, 2007.

FAYYAD, U.; SHAPIRO, G.P.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence,1996. 0738-4602-1996.

FONSECA, S.O; NAMEN, A.A. Mineração em bases de dados do INEP: uma análise exploratória para nortear melhorias no sistema educacional brasileiro. Educação em Revista|Belo Horizonte|v.32|n.01|p. 133-157 |janeiro-março 2016. Disponível em: < http://www.scielo.br/scielo.php?pid=s0102-46982016000100133&script=sci_abstract&tlng=pt> Acesso em: Novembro, 2019.

GEODACENTER. Apresentando o GeoDa. Disponível em :< http://geodacenter.github.io/index.html> Acesso em : Dezembro, 2019.

GOLDSCHMIDT, R.; PASSOS, E. Data Mining um guia prático. Conceitos, técnicas, ferramentas, orientações e aplicações. Rio de Janeiro. Elsevier, 2005.

IDEB. Nota Técnica. Disponível em: <http://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf> Acesso em 30/07/2019.

INEP. Nota Informativa. Disponível em:<http://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/nota_informativa_ideb2017_atualizado_Jun_2019.pdf> Acesso em : 16/09/2019.

INEP. Nota Técnica, Índice de Desenvolvimento da Educação Básica. Disponível em: <http://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf> Acesso em: 10 jun.2019.

INEP.Portal INEP. Disponível em:<http://porta.inep.gov.br/web/guest/inicio> Acesso em:10 jun.2019.

LANDIM, P.M.B.; YAMAMOTO, J.K. Geoestatística, conceito e aplicações. São Paulo. Oficina de Textos,2013.pg. 19 e 20.

LUZARDO, A.J.R.; CASTANEDA FILHO, R.M.; RUBIM, I.B. Análise espacial exploratória com o emprego do índice de moran.GEOgraphia, Niteroi, Universidade Federal Fluminense, Vol. 19, N.40, 2017.

MANZANO,J.A.N.G. PostgreSQL 8.3.0: Interativo: Guia de Orientação e Desenvolvimento. 1ª ed. São Paulo: Érica, 2008. Pg:15 a 19.

MEC. Ministério da Educação. Disponível em:<http://portal.mec.gov.br>Acesso em: 11 jun.2019.

MILOCA, S.A; CONEJO, P.D. Multicolinearidade em Modelos de Regressão. XXII Semana Academica da Matemática,2011. Disponível em:< https://www.estatistica.br/~yambar/MI404-Metodos%20Estatisticos/Aula%208-9%20Regress%E3o%20mult%20dim/inete%20adicional%20-

%20multicolinearidade%20em%20modelos%20de%20regressao.pdf>. Acesso em: Dezembro, 2019.

RAMOS, F.R. Indicadores de Autocorrelação Local Exemplo do Município de São Paulo. DPI SER 301Análise Espacial de Dados Geográficos,1999. Disponível em :< http://www.dpi.inpe.br/cursos/ser301/trabalhos/fred_lisa.pdf> Acesso em : Dezembro, 2019.

ROGERSON, P. Métodos estatísticos para geografia: um guia para o estudante. 3. Ed. Porto Alegre: Bookman, 2012, 348p.

SANTOS, KM; SANTOS, M.J; CARNIELLO, M.F. Analise Espacial da distribuição das escolas em relação a qualidade de ensino no município de porto velho-RO, Revista Brasileira de Gestão e Desenvolvimento Regional. G&DR. V. 15, N.6,Edição Especial, P.231 a 240, Nov.2019 Taubaté-SP. Disponível em:< https://www.rbgdr.com.br/revista/index.php/rbgdr/article/view/5197/855> Acesso : Dezembro, 2019.

SOARES, J.F; XAVIER, F.P. Pressupostos educacionais e estatísticos do IDEB. *Educ. Soc.* [online].2013, vol.34, n.124, pp.903-923. ISSN 0101-7330. http://dx.doi.org/10.1590/S0101-73302013000300013.< Disponível em: http://www.scielo.br/pdf/es/v34n124/13.pdf> Acesso em : Novembro, 2019.

SOLIGO, V. Relações entre condições sociais e o rendimento escolar aferido pelo IDEB na região sul. Disponível em:< pleiade.uniamerica.br/index.php/pleiade/article/view/274/459>. Acesso em: novembro, 2019.

SOUZA, F.C.M. Implementação de SIG e Mapas de Kernel visando Acessibilidade na Educação Superior. Recife, 2015.

SOUZA, H.G; TABOSA, F.J.S. Economia do Ceará em debate 2016. Análise Espacial do desempenho escolar da educação básica dos municípios do estado do ceará. P. 121. Disponível em:< https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2018/11/ENCONTRO_ECONOMIA_CEARA_EM_DEBATE_2016.pdf#page=121>. Acesso em: Dezembro, 2019.

THOMÉ, R. Interoperabilidade em Geoprocessamento: Conversão entre Modelos Conceituais de Sistemas de Informação Geográfica e Comparação com o Padrão Open Gis.

TOSTA, K.S. Análise da qualidade do IDEB como índice de desenvolvimento da educação básica: um estudo de caso do município de Campos dos Goytacazes/RJ. Dissertação (Mestrado em Políticas Sociais do Centro de Ciências do Homem) – Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, 2015.

VIEIRA, M.H.P.; PEREIRA, E.F.; VIEIRA, J.; LOUZADA, R.O.; SILVA, M.C.A.; FERREIRA, L.M. Mapas de Kernel como Alternativa ao Monitoramento Ambiental: Análise das industrias que realizam emissões nos municípios do estado de Mato Grosso do Sul, nos anos 2008 a 2018. In 7ºSimpósio de Geotecnologias no Pantanal, Jardim, MS, 20 a 24 de outubro de 2018. Embrapa Informática Agropecuária/INPE, p. 627-636.