

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Marco Antonio da Silva Rodrigues

**Armazenamento e Manipulação de Dados  
Espaciais no PostgreSQL/PostGIS**

**Uberlândia, Brasil**

**2018**

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Marco Antonio da Silva Rodrigues

**Armazenamento e Manipulação de Dados Espaciais no  
PostgreSQL/PostGIS**

Trabalho de conclusão de curso apresentado  
à Faculdade de Computação da Universidade  
Federal de Uberlândia, Minas Gerais, como  
requisito exigido parcial à obtenção do grau  
de Bacharel em Sistemas de Informação.

Orientador: Humberto Razente

Universidade Federal de Uberlândia – UFU  
Faculdade de Ciência da Computação  
Bacharelado em Sistemas de Informação

Uberlândia, Brasil

2018

Marco Antonio da Silva Rodrigues

## **Armazenamento e Manipulação de Dados Espaciais no PostgreSQL/PostGIS**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Sistemas de Informação.

Trabalho aprovado. Uberlândia, Brasil, 21 de dezembro de 2018:

---

**Humberto Razente**  
Orientador

---

**Professor**

---

**Professor**

Uberlândia, Brasil  
2018

# Agradecimentos

A minha mãe Adriana, por ter me ensinado os valores de honestidade, respeito e trabalho, que por mais que seja difícil não podemos desistir.

A minha amada esposa Sirlene que me incentivou a voltar à estudar, me apoiou nos momentos mais difíceis deste curso e desde que estamos juntos só me faz acreditar de que sou capaz.

A esta universidade, seu corpo docente pelo bom trabalho que realizam e principalmente ao meu orientador Humberto Razente, que me deu a oportunidade de realizar este trabalho e me ajudou com muitas correções e conselhos.

A todos os meus colegas de turma que passaram esses últimos anos comigo, mas principalmente ao meu amigo Tadeu Rodrigues, que me auxiliou em todas dificuldades que tive durante o curso.

# Resumo

Nas últimas décadas surgiram necessidades tais como; de automação e sistemas para o melhoramento no trânsito de grandes cidades, melhoria de infraestrutura, planejamento urbano, controle de queimadas, desmatamento, reflorestamento, agricultura, turismo, mercado imobiliário entre outros. São temas de grande importância e de diversos estudos, que com a utilização de base de dados geográficas podem ser mais bem detalhados e analisados, culminando em resultados mais precisos. Um banco de dados com extensão geográfica auxilia especialistas de diversas áreas, pois, não é necessário um grande conhecimento em banco de dados para operar um SGBD ou SIG, que oferecem ferramentas de apoio a decisões. As ferramentas além de não serem demasiadamente complexas tem um custo relativamente baixo e às vezes são gratuitas para fins acadêmicos. Um bom exemplo de uso de dados geográficos é o da COPEL (Companhia Paranaense de Energia), que com os aprimoramentos no setor elétrico nas últimas décadas precisou avaliar as metodologias de planejamento dos sistemas de distribuição e transmissão de energia. A partir de 1996 passou a utilizar um SIG com um aplicativo de cadastro que está integrado a um banco de dados com os elementos da rede de distribuição referenciados geograficamente para analisar o desempenho dos principais componentes da sua rede de distribuição ([DEV-MEDIA, 2008](#)). Esta monografia apresenta o suporte a dados geográficos em sistemas de gerenciamento de bancos de dados relacionais com destaque à linguagem SQL do PostGIS no PostgreSQL, bem como o desenvolvimento de um estudo de caso.

**Palavras-chave:** Georreferenciamento, Índices, Dados Espaciais, Schemas, Banco.

# Listas de ilustrações

Figura 1 – Mapa da cólera em 1832 de Charles Picquet.	16
Figura 2 – Mapa geopolítico de Minas Gerais.	19
Figura 3 – Tipos de geo-objetos.	20
Figura 4 – Exemplo de geo-campo da cidade de Uberlândia.	21
Figura 5 – Subclasses de geo-campo.	21
Figura 6 – Tipos de dados espaciais.	23
Figura 7 – Diagrama OMT-G do sistema de controle de ocorrências.	27
Figura 8 – Modelo Raster e Vetorial.	29
Figura 9 – CREATE INDEX	31
Figura 10 – CREATE INDEX CONCURRENTLY	31
Figura 11 – CREATE INDEX BRIN	32
Figura 12 – ST_AREA.	34
Figura 13 – ST_COVEREDBY	34
Figura 14 – CREATE EXTENSION.	39
Figura 15 – CREATE DATABASE.	40
Figura 16 – CREATE TABLE.	40
Figura 17 – CREATE pontos.	42
Figura 18 – CREATE polígonos	43
Figura 19 – Polígonos e Pontos	43
Figura 20 – Projeção Ferrovias Brasileiras	44
Figura 21 – Projeção Capitais Brasileiras	45
Figura 22 – Consulta áreas de uso sustentáveis	46
Figura 23 – Áreas de uso sustentáveis	46
Figura 24 – Cidades até 100 km de Uberlândia	46
Figura 25 – ST_ASEWKT	47
Figura 26 – Cidades a menos de 100 km de Uberlândia	47
Figura 27 – As 10 cidades mais perto de Uberlândia	47
Figura 28 – As 10 cidades mais perto de Uberlândia	48
Figura 29 – Áreas sustentáveis dentro do município de São Paulo	48
Figura 30 – Exibição das áreas sustentáveis dentro do município de São Paulo	49
Figura 31 – Proxy	56
Figura 32 – Site PostgreSQL	57
Figura 33 – Site EnterpriseDB	58
Figura 34 – 1º passo.	58
Figura 35 – 2º passo.	59
Figura 36 – 3º passo.	59

Figura 37 – 4º passo.	60
Figura 38 – 5º passo.	60
Figura 39 – 6º passo.	61
Figura 40 – 7º passo.	61
Figura 41 – 8º passo.	62
Figura 42 – 9º passo.	62
Figura 43 – 10º passo.	63
Figura 44 – Menu Iniciar - pgAmin4.	63
Figura 45 – Server	64
Figura 46 – Connection	64
Figura 47 – Extension	65
Figura 48 – Schema public	65
Figura 49 – Site IBGE	66
Figura 50 – Importando o banco	66
Figura 51 – Site QGIS	67
Figura 52 – Menu Iniciar - QGIS	68
Figura 53 – Interface QGIS	68
Figura 54 – Camada	69
Figura 55 – Conexão	69
Figura 56 – Adicionar Tabelas	70
Figura 57 – Visualização das tabelas	70

# **Lista de abreviaturas e siglas**

SGBD	Sistema de Gerenciamento de Banco de Dados
SGBDG	Sistema de Gerenciamento de Banco de Dados Geográfico
SIG	Sistema de Informação Geográfica
ESRI	Environmental Systems Research Institute
SRS	Sistema de Referência Espacial
SRID	Sistema de Identificação de Referência Espacial
OGC	Consórcio Geoespacial Aberto

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
1.1	Objetivos	11
1.2	Método	11
1.3	Resultados	12
1.4	Organização do Trabalho	13
<b>2</b>	<b>HISTÓRIA</b>	<b>14</b>
2.1	Começo da Análise Espacial	15
2.2	Primeiro SIG	15
<b>3</b>	<b>BANCOS DE DADOS GEOGRÁFICOS</b>	<b>18</b>
3.1	Modelos de representações	18
3.2	Tipos de dados espaciais	22
3.3	Operações para dados espaciais	23
3.4	Modelagem de Dados Espaciais	25
3.5	Modelos para representações de dados espaciais	27
3.5.1	Sistemas de referências espaciais	28
3.5.2	Modelos de análise	29
3.6	<b>Indexação Geográfica</b>	<b>30</b>
3.6.1	Índices GIST	30
3.6.2	Índices BRIN	31
3.6.3	Índices SP-GIST	32
3.6.4	Usando Índices	32
3.6.5	Consultas Espaciais	33
3.7	<b>Ferramentas</b>	<b>35</b>
3.7.1	PostGIS	35
3.7.2	Oracle Spatial	36
3.7.3	IBM DB2 Spatial Extender	37
3.7.4	SQL Server (SPATIAL DATA)	37
<b>4</b>	<b>ESTUDO DE CASO</b>	<b>39</b>
4.1	O Banco	39
4.2	O SIG	41
<b>5</b>	<b>CONCLUSÃO</b>	<b>50</b>
	<b>Conclusão</b>	<b>50</b>

<b>5.1</b>	<b>Trabalhos Futuros</b>	<b>50</b>
	<b>REFERÊNCIAS</b>	<b>51</b>
	<b>ANEXOS</b>	<b>54</b>
	<b>ANEXO A – INSTALAÇÃO DO POSTGRESQL</b>	<b>56</b>
	<b>ANEXO B – INSTALAÇÃO DO QGIS</b>	<b>67</b>

# 1 Introdução

A utilização de mapas é uma prática muito antiga com registros desde a Grécia antiga, Império Romano, entre outras civilizações. Eram utilizados com a mesma finalidade para qual utilizam hoje, ferramenta de orientação e localização, ou seja, registrar onde estão e o que têm a sua volta. Os primeiros mapas eram feitos em madeiras, esculpidos ou pintados. Foram muito utilizados para estratégias militares e de desbravamento, ([MARTINELLI, 2003](#)).

Com o passar dos séculos surgiu a cartografia que na verdade nunca foi apenas uma ciência que estuda a representação de mapas, pois por trás de todo mapa têm interesses políticos, econômicos e pessoais, onde é possível verificar o que o seu “vizinho” têm a mais que você. Hoje, com o avanço de novas tecnologias, a cartografia não poderia ficar de fora. Os mapas começaram a ser utilizados com a computação para fins militares e hoje são utilizados em diversas aplicações, desde simples aplicativos de buscas até redes sociais.

No Brasil é obrigatório o georreferenciamento em casos de alterações no registro imobiliário, a exemplo de compra e venda, desmembramento, remembramento, sucessão, partilha ou mudança de titularidade em propriedades rurais com mais de 100 hectares de acordo com a lei 10.267 (regulamentada pelo Decreto 4.449 e, em decorrência, entrou efetivamente em vigor na data de sua publicação em 30/10/02.), ([LEI..., 2001](#)), a Lei exige que o polígono resultante do georreferenciamento não se sobreponha a nenhum outro já certificado pelo Incra (Instituto Nacional de Colonização e Reforma Agrária). Para realizar uma georreferenciação é preciso conhecer as coordenadas geográficas do terreno. Esses dados são obtidos a partir de ferramentas que realizam análises das áreas que serão registradas no Incra. Tais ferramentas são conhecidas como SIG (sistemas de informação geográfica). Os dados geográficos também podem ser utilizados por exemplo, para realização de estudos ambientais, demográficos, climáticos, hidrográficos, etc. Um dado geográfico representa um objeto ou fenômeno para o qual a localização geográfica é uma característica inerente à informação e indispensável para analisá-la.

Os dados obtidos por meio de georreferenciamento podem ser armazenados como pontos, linhas ou polígonos. Essas informações são conhecidas como dados geográficos ou também dados espaciais, porque o que diferencia esse tipo de dado para os dados convencionais são os componentes espaciais neles contidos. Os dados espaciais são divididos em duas classes, vetorial e matricial ([LIU; ÖZSU, 2009](#)). No modelo vetorial os dados são armazenados por meio de coordenadas geográficas e no modelo matricial o armazenamento é por meio de linhas e colunas onde cada célula tem um valor correspondente ao atributo analisado e pode ser localizado pelos cruzamentos. Além de armazenar as formas geomé-

tricas simples é possível utilizar recursos mais avançados como regras de relacionamento, topologia, comportamento, anotações, imagens, objetos 3D, entre outros. Um banco de dados convencional consegue armazenar informações de endereços, que também são informações geográficas utilizadas como partes de um cadastro para futuras consultas, porém em um banco de dados espacial utiliza-se a informação de endereço para transformar em uma localização no espaço e relaciona-se com outras informações no espaço. Os principais sistemas gerenciadores de banco de dados (SGBD) possuem extensões para utilização de dados geográficos. Para utilizar as informações contidas no SGBDG há opções de SIGs onde é possível analisar as informações em representações geométricas (GOODCHILD, 2009).

## 1.1 Objetivos

O objetivo deste trabalho é o estudo sobre as formas de armazenamento, representações, classes, operações, funções, modelagem e consultas que são possíveis serem realizadas com dados espaciais. São apresentados os fenômenos geográficos que são Rios, Cidades, Vegetação, Solo; os tipos de dados espaciais que representam os fenômenos, um rio por exemplo, pode ser representado por um tipo *Line* e uma cidade pode ser representada por um *Point* ou por *Polygon*, essa escolha dependerá de qual é o intuito do projeto que está sendo desenvolvido; a maneira que foi encontrada para utilizar os dados espaciais em um banco de dados e integrá-lo à um SIG; o que um SIG bem estruturado é capaz de proporcionar à um usuário após estar devidamente configurado a um ou mais bancos de dados geográficos; quais são os tipos de dados mais utilizados e para quais objetos do mundo real são melhores indicados; os SGDBs que possuem extensões para utilização de dados espaciais, mas principalmente que é possível realizar com as funções que recebem como argumento tipos espaciais; como utilizar os SRSs (Sistemas de Referências Espaciais) no armazenamento de dados espaciais para que depois possam ser utilizados com escalas reais. Um breve entendimento sobre a história dos SIGs juntamente a informatização da cartografia.

Foi realizado um estudo de caso no qual é apresentado como criar um banco de dados com extensão espacial, como utilizá-lo e manipulá-lo, também como integrar o banco criado com um SIG. Ao final deste estão disponíveis dois anexos para que seja possível segui-lo e ao final dos procedimentos utilizar um banco de dados geográfico PostGIS e o SIG QGIS.

## 1.2 Método

Para realização deste trabalho foram utilizados o Portal de Periódicos da CAPES e as bibliotecas digitais da ACM (<http://dl.acm.org>) e IEEE Xplore Digital Library.

Durante as pesquisas foi verificado qual seria a melhor opção de Sistema de Gerenciamento de Banco de Dados e também qual Sistema de Informação Geográfica para utilizar no estudo de caso. As escolhas foram pelo banco de dados PostgreSQL e como SIG pelo QGIS por serem softwares livres de código aberto. Para efeito de demonstração de um banco de dados real, com dados que facilitarão o entendimento de quem avaliar ou utilizar este como estudo, optou-se pelo uso do banco de dados do IBGE (Instituto Brasileiro de Geografia e Estatística), que está disponível em seu site, com dados referentes a geografia do país.

### 1.3 Resultados

O resultado deste trabalho foi a realização do estudo de caso no capítulo 4, dividido nas Seções 4.1 e 4.2 onde foram empregados os conceitos estudados nos capítulos anteriores. Para realização do estudo de caso foi criado um ambiente composto de um gerenciador de banco de dados com extensão espacial (PostGis) que posteriormente foi utilizado para restaurar uma cópia do banco de dados bc250\_2017-11-08.tar, disponível no site do IBGE (IBGE, 2017), um SIG (QGIS). A cópia restaurada contém informações das principais agências nacionais de infraestrutura, sessenta e uma tabelas populadas, mais de mil funções do PostGis que recebem como entrada dados espaciais, índices espaciais, com isso, os exemplos citados no estudo de caso puderam ficar mais realistas (NETO; COSTA, 2017). Com o SIG e o SGBDG instalados e devidamente configurados foi realizada a conexão entre as duas ferramentas para que os dados presentes no banco pudessem ser visualizados em representações cartográficas que foram utilizadas com os dados restaurados, mas também com dados criados de maneira explicativa para futuros estudos em cima deste. Dentro do estudo de caso foram exploradas algumas funcionalidades do QGIS, como a possibilidade de criar camadas de visualização a partir de consultas realizadas em um SGBD interno da ferramenta (Figura 22). No estudo de caso estão citados exemplos de criação de tabelas (Figura 16, 17 e 18), inserção de dados (Figura 17 e 18), utilização de funções, representações de camadas (Figuras 19, 20, 21 e 23), criação do banco de dados (Figura 15) e extensão espacial (Figura 14), como o PostGis utiliza o formato WKT para realizar a conversão das coordenadas geográficas inseridas no banco em valores hexadecimais e com quais funções é possível realizar o processo reverso. São apresentadas quais as variações mais utilizadas e como os índices espaciais podem melhorar o tempo de resposta da consultas.

Com a conclusão deste trabalho foi verificado o quanto abrangente é o estudo sobre dados geográficos por mais que o intuito deste trabalho seja voltado para o que um banco de dados geográfico é capaz prover à quem o for utilizá-lo. Não pode ser deixado de lado o que ele ajuda socialmente e academicamente os cidadãos que o utilizam, principalmente por acadêmicos de outras áreas do conhecimento, pois a grande maioria de estudos utili-

zam um recorte de alguma localidade e posteriormente o comparam com outro recorte de algum trabalho já realizado. Para ambos os dados geográficos são de grande importância e sem eles os resultados ou a aplicação não seriam tão eficientes.

## 1.4 Organização do Trabalho

Este trabalho está dividido em cinco capítulos e dois anexos. O Capítulo 1 apresenta a utilização de mapas na história da humanidade, o surgimento da cartografia, tratando de temas que explicam de forma sucinta como podem ser aproveitados os SGBDGs pela sociedade, os objetivos do trabalho, métodos que foram usados para o desenvolvimento do mesmo. O capítulo 2 trata da história dos dados geográficos, como eram utilizados e com quais propósitos, também é descrito o desenvolvimento dos primeiros sistemas de informação geográficas. O capítulo 3 detalha de forma conceitual e com algumas ilustrações o que são dados geográficos, suas representações, operações, modelagens, como funcionam os índices geográficos e quais ferramentas são mais utilizadas para manipulá-los. No capítulo 4 é apresentado através de um estudo de caso como criar um banco de dados geográfico, suas principais características de armazenamento e criação de dados, também como integrá-lo a um SIG. A conclusão do trabalho é o capítulo 5 onde é apresentado o que foi aprendido com o trabalho e quais caminhos ainda podem ser explorados com o estudo de SGBDGs. No final há dois anexos para que seja possível criar exemplos semelhantes do estudo de caso, contido nos manuais A e B.

## 2 História

Entre as aplicações que mais utilizam bancos de dados espaciais, estão os SIGs, que são extensões modernas da cartografia já utilizadas há séculos, um SIG têm duas grandes diferenças e uma semelhança com os sistemas cartográficos antigos. A semelhança está que tanto um documento cartográfico quanto uma exibição de camadas em um SIG são representações de mapas simples onde é possível a adição de novas informações. Já as diferenças principais se dão que em um SIG não há limite para a adição de novas informações, visto que em uma mapa convencional há um limite físico para inserir novas informações, um SIG também utiliza análise e estatística para apresentar dados em apoio a argumentos particulares que um mapa tradicional de papel não pode realizar.

O desenvolvimento dos SIGs foi dividido em quatro grandes fases. A primeira fase se deu no início da década de 1960 até meados da década de 1970, que observou a criação de uma nova disciplina voltada para moldar a direção de pesquisa e desenvolvimento futuro de informações espaciais. A segunda fase, desde a década de 1970 até o início dos anos de 1980, viu a adoção de tecnologias por agências nacionais que levaram à um foco no desenvolvimento de melhores práticas. A fase três, entre 1982 até o final dos anos de 1980, viu o desenvolvimento e a exploração do mercado comercial em torno do SIG, enquanto a última fase foi desde o final dos anos 1980 e concentra-se em maneiras de melhorar a usabilidade da tecnologia, tornando as instalações mais centradas no usuário ([DEMPSEY, 2012](#)).

As Informações sobre a captação e desenvolvimento de SIG, particularmente por departamentos nacionais, são fragmentadas. Atualmente, nenhuma pesquisa disseminada nessa área foi realizada. Também parece provável que os primeiros estágios de desenvolvimento de SIGs no século XX foram caracterizadas por indivíduos que estavam perseguindo objetivos distintos em matéria de SIG e não havia prosseguimento com os dados de uma pesquisa para outra, não se levava em conta o que já tinha sido desenvolvido. Naquela época, o Laboratório de Computação Gráfica de Harvard, o Sistema de Informações Geográficas do Canadá, o Instituto de Pesquisa de Sistemas Ambientais e a Unidade de Cartografia Experimental do Reino Unido foram as principais influências no campo. Uma única direção não apareceu até que o assunto se tornasse o foco de intensa atividade comercial, uma vez que a tecnologia de imagem por satélite significava que aplicações em massa poderiam ser criadas para uso empresarial e privado, à época a empresa *Environmental Systems Research Institute* (ESRI) se tornou a organização dominante no campo, pois tornou-se a pioneira na produção de soluções para a área de informações geográficas com ênfase em estudos de uso do solo ([DEMPSEY, 2012](#)).

## 2.1 Começo da Análise Espacial

A primeira aplicação documentada do que poderia ser classificado como um SIG foi na França em 1832. O geógrafo francês Charles Picquet criou uma representação baseada em mapas da epidemiologia da cólera em Paris, representando os 48 distritos de Paris com diferentes tonalidades de cores de meio-ton, uma versão inicial de um mapa de calor. O mapa, publicado no relatório *Rapport sur la marche et les effets du choléra-morbus*, em Paris, é provavelmente o primeiro uso da análise espacial na epidemiologia. Na Figura 1 é possível verificar o mapa de Charles Picquet.

Outra situação semelhante levou John Snow a descrever as mortes por cólera em Londres usando pontos em um mapa em 1854. O mapa de Snow era importante porque não era apenas uma apresentação de dados, foi feita uma tentativa de apresentar um argumento desenvolvido a partir de uma análise espacial dos dados exibidos no mapa.

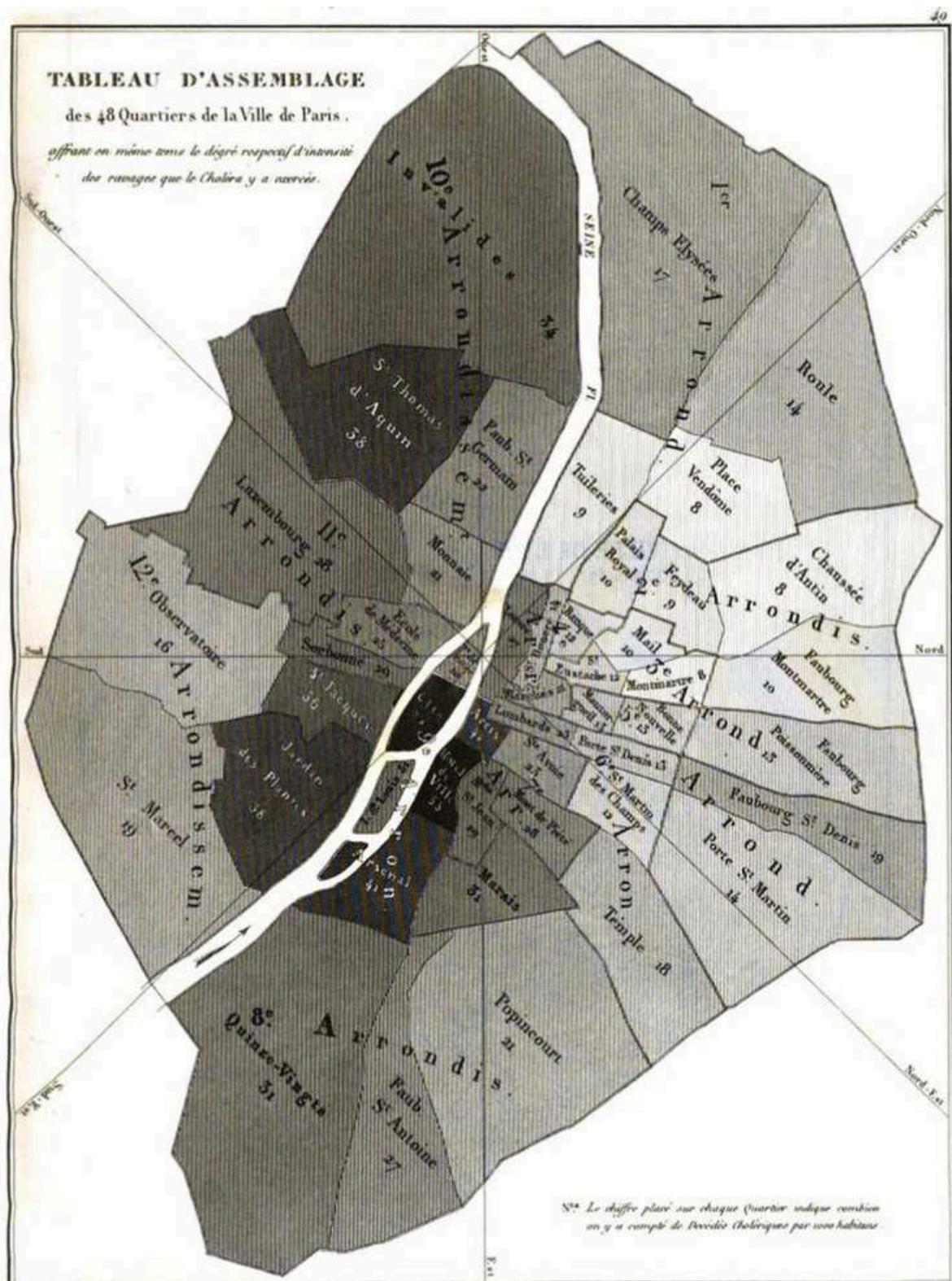
O próximo passo significativo no desenvolvimento de sistemas de informação geográfica modernos foi no início do século XX. Uma técnica de impressão conhecida como *Photozincography*, que foi um processo fotográfico desenvolvida no século XIX por Sir Henry James, a *Photozincography*, era uma maneira revolucionária de copiar negativos fotográficos em zinco para ser usada na criação de mapas, contornos, gravuras e reprodução de imagens. *Photozincography* é, em certo sentido, a primeira máquina copiadora produzida e foi usada para separar camadas de um mapa ([HANNAVY, 2008](#)). Embora pareça ser um SIG, isso não representa um SIG completo, pois não há oportunidade de fornecer uma análise dos dados mapeados ([DEMPSEY, 2012](#)).

## 2.2 Primeiro SIG

Com o avanço das armas nucleares foi necessário que aplicações de hardware e mapeamento fossem desenvolvidas e com essas tecnologias o primeiro SIG operacional foi lançado em Ottawa, Canadá. Esse SIG inicial foi desenvolvido para armazenar, comparar e analisar dados sobre o uso da terra no Canadá. O sistema foi aprimorado ao longo dos anos de 1970 e 1980 até meados dos anos de 1990, altura em que foi impulsionado por hardware de textitMainframe e continha conjuntos de dados de toda a massa terrestre canadense. Logo começou a ser utilizado para fins militares.

De acordo com ([DEMPSEY, 2012](#)) durante os anos de 1970 e 1980 desenvolvimentos na consciência espacial e em como lidar com dados espaciais estavam sendo feitos em centros acadêmicos importantes, como Harvard e Esri. Na década de 1990, a Esri, uma das maiores empresas de software SIG, lançou o ArcView, que era uma solução de desktop para produzir sistemas de mapeamento através de uma interface baseada no Windows. O padrão ArcView foi logo adotado por muitas organizações governamentais, empresariais,

Figura 1 – Mapa da cólera em 1832 de Charles Picquet.



Mapa sombreado mostrando mortes de cólera por mil habitantes para cada um dos 48 distritos de paris de 1832 por Charles Picquet ([DEMPSEY, 2012](#)).

de defesa e não-governamentais devido à sua interface gráfica e facilidade de uso, ainda hoje o padrão do ArcView é utilizado, um exemplo é o QGIS que é muito semelhante.

Durante a próxima década, a Internet viu a adoção de tecnologias SIGs em níveis mais baixos, a medida que os custos caíam e a tecnologia chegava ao alcance das autoridades locais. Ao mesmo tempo, a disseminação da Internet forneceu um meio simples para acessar mapas. Como a Esri encorajou as organizações a adicionar conjuntos de dados aos conjuntos de mapas que já estavam sendo disponibilizados pela Internet. Setores-chave como o governo, órgãos não-governamentais e empresas de serviços públicos pareceram estar desenvolvendo uma abordagem para compartilhar dados e há evidências de compartilhamento significativo em tais plataformas, como o de conjuntos de dados em plataformas compartilhadas, até chegar aos dias atuais onde qualquer pessoa com o mínimo de conhecimento de computadores e banco de dados por utilizar os SGIs existentes.

# 3 Bancos de Dados Geográficos

Este capítulo apresenta uma breve compilação do Capítulo *Spatial Data Warehouses* do livro *Data Warehouse Systems* ([VAISMAN; ZIMÁNYI, 2014](#)) e do Tutorial sobre Bancos de Dados Geográficos ([QUEIROZ; FERREIRA, 2006](#)).

O banco de dados geográfico é um dos componentes mais importantes para quem trabalha com SIGs. Um erro muito comum para quem o utiliza com um SIG é que inicialmente não se preocupam com a organização, armazenamento e normatização dos dados, um banco de dados espacial bem estruturado pode solucionar esses erros. Outra importante pergunta para quem está criando um banco de dados espacial é para que ele está sendo criado, pois sem essa resposta podem-se criar um banco com muita informação desnecessária para a aplicação que utilizará o banco, um exemplo, o banco de dados de um estudo médico que quer determinar onde há mais incidência de determinada doença não precisa conter tantas informações quanto o banco de dados do Google Maps, com informações de restaurantes, rodovias e etc. As informações que são armazenadas em um banco de dados são divididas duas formas, dados Alfanuméricos, que são textos e números, esses tipos de dados são utilizados em um banco de dados relacional comum, também existe a forma de dados espaciais que podem representar informações sobre o local físico e a forma de objetos geométricos. Esses objetos podem ser locais como cidades, edifícios, representados por pontos ou objetos mais complexos como países, estradas ou lagos, representados por múltiplos polígonos, conjuntos de linhas ou polígonos simples. Cada tipo de dado espacial, assim como dados comuns podem ser utilizados em operações como exemplo, medir distância, perímetro, área e etc. Nas seções a seguir, será apresentado os modelos de representações, os tipos de dados espaciais, operações possíveis com eles, as formas de modelagens, os índices mais utilizados e as ferramentas que utilizam.

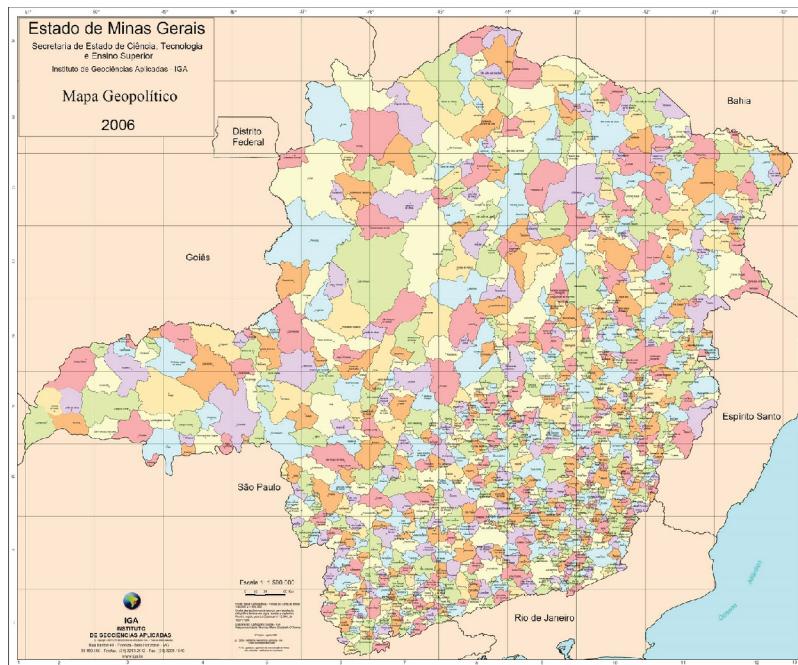
## 3.1 Modelos de representações

Os dados espaciais podem ser classificados em dois grandes modelos formais de representações para trabalharmos em nível conceitual da modelagem de dados, geo-objetos e geo-campos;

**Geo-objetos:** No modelo geo-objetos são representadas coleções de entidades distintas ou semelhantes com delimitações bem especificadas, ou seja, são individualizáveis no mundo real e cada componente tem sua identidade e seu atributo. Divisões políticas e terrenos em um loteamento são alguns exemplos de geo-objetos. Na Figura 2 está representado o mapa político de Minas Gerais, onde é possível identificar todos os geo-objetos que representam as cidades do Estado. O uso de coleções de geo-objetos é muito frequente em

bancos de dados espaciais, pois é muito proveitoso tratar geo-objetos parecidos de forma consistente. Por exemplo, os municípios do Estado de Minas Gerais, os estados do Brasil, e os países da Europa. A ideia de coleções de geo-objetos ainda é aproveitada para propor um modelo orientado a objetos para dados espaciais.

Figura 2 – Mapa geopolítico de Minas Gerais.



Fonte: Instituto de Geociências Aplicadas — IGA (2006)

No modelo OMT-G os geo-objetos são divididos em duas classes, geo-objeto com geometria e geo-objeto com geometria e topologia. Na classe de geo-objeto com geometria são representados apenas objetos que têm atributos geométricos, como, linhas, pontos e polígonos, é possível representar edificações, ruas, árvores, etc. Já na classe de geo-objetos com geometria e topologia, representa além de objetos que possuem características geométricas, também representam propriedades topológicas de conectividade, podendo ser representados por nós e segmentos orientados. É especializada em classes do tipo Nô, linha unidirecional e bidirecional. Alguns exemplos desta classe são estradas, redes de esgotos, rios, etc.

As duas classes dos geo-objetos são divididas em subclasses que representam:

- Polígono - objetos de uma área, podem aparecer conectados, como cidades dentro de um estado ou isolados, como os continentes no planeta terra.
- Ponto - objetos que possuem um único par de coordenadas ( $x, y$ ), como por exemplo na representação de árvores, ponto de ônibus, hidrantes, etc.

- Linha - objetos lineares sem exigência de conectividade, como divisões de terrenos, limites de municípios, etc.
- Nó - objetos aos quais linhas encerram-se ou cruzam-se com um grafo. Possui a propriedade de conectividade, garantindo a conexão com a linha. Alguns exemplos de nós podem ser vistos nos cruzamentos de ruas, redes de esgotos, etc.
- Linha Unidirecional - objetos lineares que tem o começo e fim da linha em um nó e possuem uma direção específica. Como exemplo existem trechos de uma rede de esgoto, que indicam a direção do fluxo da rede.
- Linha Bidirecional - objetos lineares que começam e terminam em um nó e que são bidirecionais. Cada linha bidirecional deve estar conectada a dois nós ou a outra linha bidirecional. Como exemplo há trechos de uma rede de água, onde a direção do fluxo pode em dois sentidos.

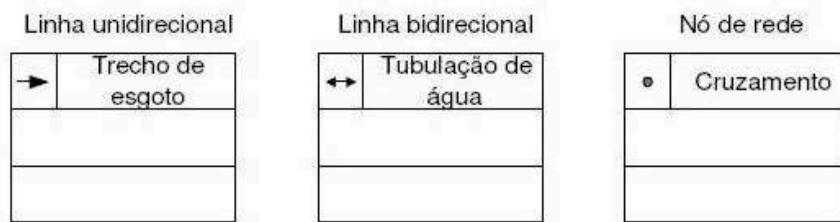
Na Figura 3 está apresentado como os tipos de geo-Objetos são representados em um diagrama.

Figura 3 – Tipos de geo-objetos.

#### Geo-objetos com geometria



#### Geo-objetos com geometria e topologia



Fonte: (QUEIROZ; FERREIRA, 2006)

**Geo-campo:** Neste modelo o espaço observado é contínuo sem delimitações, onde dentro do geo-campo é possível existir inúmeros geo-objetos ou mesmo todo o geo-campo estar dentro de um único geo-objeto. São fenômenos de variação contínua no espaço e tempo. Tipo de vegetação, solo e relevo, são exemplos de geo-campos. Na Figura 4 é possível verificar que tanto dentro do retângulo vermelho ou fora, as propriedades da imagem são as mesmas sem nenhuma distinção prévia. As delimitações de um geo-campo são de uma forma arbitrária, na imagem da cidade de Uberlândia-MG, os limites correspondem apenas aos limites que decidiu-se enquadrar ou em outros casos as limitações da ferramenta que foi utilizada para registrar o geo-campo. Assim, os geo-campos podem ser divididos em partes e ainda assim manter sua propriedade essencial.

Figura 4 – Exemplo de geo-campo da cidade de Uberlândia.



Fonte: Google Maps (2018)

O modelo OMT-G define a classe geo-campo em cinco subclasses, Isolinhas, Subdivisão Planar, Tesselação, Amostras e Rede Triangular Irregular, como exemplificado na Figura 5. Cada uma das subclasses possui um padrão simbólico de representação:

Figura 5 – Subclasses de geo-campo.

Rede triangular irregular	Isolinhas	Subdivisão planar	Tesselação	Amostras
Temperatura	Curvas de nível	Pedologia	Imagem LANDSAT	Pontos cotados

Fonte: (DELBONI; BORGES; LAENDER, 2005)

- Rede triangular Irregular - representa o conjunto de grades triangulares de pontos

que cobrem todo o domínio espacial. Um exemplo de rede triangular irregular é visto em modelagem de terreno.

- Isolinhas - são coleções de linhas fechadas que não se cruzam ou se tocam. Cada círculo representa uma valor associado. Exemplo, curvas de nível, curvas de temperatura e curvas de ruído, etc.
- Subdivisão Planar - são conjuntos de subdivisões com independência de uma para outra, elas não se sobrepõem. Exemplo, tipos de solo, divisão de bairros, divisões administrativas e temáticas.
- Tesselação - são conjuntos de subdivisões com células regulares que não se sobrepõem e que cobrem completamente este domínio. As células são individuais entre elas. Exemplo, imagem de satélite.
- Amostras - é a forma de representar pontos regulares ou não, distribuídos pelo espaço observado. Exemplo, Cidades, distritos, estações meteorológica, etc.

## 3.2 Tipos de dados espaciais

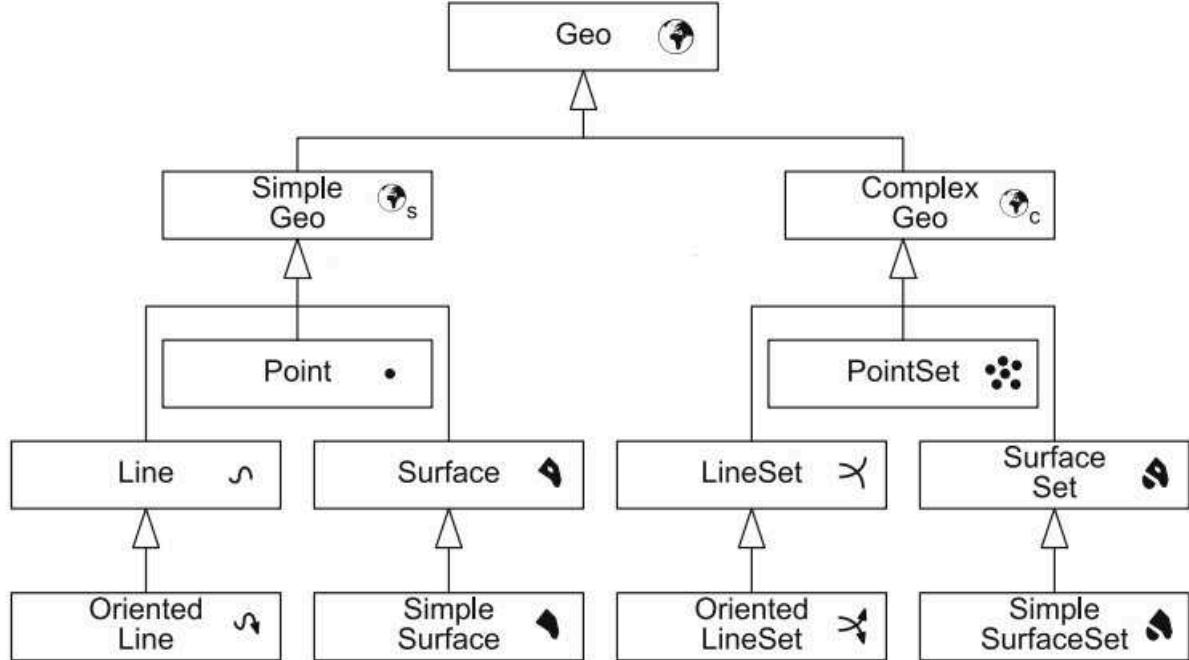
Nesta seção serão apresentados os tipos de dados equivalentes aos modelos de representações descritos na seção anterior. As representações espaciais correspondem a um objeto do mundo real a qual uma aplicação precisa armazenar suas características espaciais. Para tal, um tipo de dado espacial precisa ter o componente descritivo para especificar as características gerais dos objetos espaciais e são representados por tipos de dados tradicionais (Strings e numéricos), também é necessário um componente espacial que define a extensão do objeto no espaço de interesse.

Para representar as extensões espaciais existem vários tipos de dados, que seguem uma hierarquia.

O tipo de dado espacial Geo é o mais genérico de todos, os outros tipos de dados espaciais são abstrações do tipo Geo, como apresentado na Figura 6. Na hierarquia os próximos tipos de dados são o SimpleGeo e ComplexGeo.

O tipo de dado SimpleGeo representa os tipos *Point*, *Line* e *Surface*. O SimpleGeo é um tipo abstrato, isto é, nunca é instanciado, após a criação de um valor SimpleGeo, é necessário especificar qual dos seus subtipos caracteriza o novo elemento. Um valor SimpleGeo pode ser usado, por exemplo, para representar genericamente cidades, enquanto que uma pequena cidade pode ser representada por um *Point* apenas em forma de identificar que em determinado local existe uma cidade; já uma grande cidade pode-se utilizar o tipo *SimpleSurface*, que representa o polígono referente a cidade.

Figura 6 – Tipos de dados espaciais.



Fonte: ([VAISMAN; ZIMÁNYI, 2014](#))

O tipo ComplexGeo representa qualquer conjunto heterogêneo de geometrias que podem incluir *PointSet*, *LineSet* e *SurfaceSet*. O tipo ComplexGeo pode ser usado para representar um sistema de águas consistindo de rios (*OrientedLine*), lagos (*Surface*) e reservatórios (*Point*).

### 3.3 Operações para dados espaciais

Assim como os dados comuns têm uma grande quantidade de operações que os suportam, como operações lógicas, aritméticas, conversão, etc. Os tipos de dados espaciais podem ser utilizados em vários tipos de operações que são divididas em cinco classes, sendo elas:

**Topológicas:** Operações topológicas especificam dois valores espaciais relacionados um com o outro. Elas são muito usadas em aplicações espaciais por exemplo, se dois municípios tem uma fronteira comum, uma rodovia cruza um país ou uma cidade está localizada em um estado. As definições de relações topológicas são baseadas na definição de fronteiras, interior e exterior dos valores espaciais. Intuitivamente, o exterior de um valor espacial é composto de todos os pontos dos espaços adjacentes que não pertence ao valor espacial. O interior de um valor espacial é composto de todos seus pontos que não pertencem à fronteira. A fronteira é definida para os diferentes tipos de dados espaciais da seguinte forma, um ponto tem uma fronteira vazia e seu interior é o próprio

ponto. A fronteira de uma linha é dada por seus pontos extremos, previsto que eles podem ser distinguidos. Dentro da classe topológica é possível realizar operações que retornam resultados booleanos dos seguintes tipos:

- Intersects/Disjoint: São relações inversas, quando uma aparece a outra não. Duas geometrias são disjuntas se o interior e a fronteira de um objeto se intersectam apenas no exterior de outro.
- Equals: Verifica se uma geometria é igual a outra se elas compartilharem os mesmos conjuntos de pontos.
- Overlaps: Duas geometrias são sobrepostas se o interior de cada uma intersecta o interior e o exterior de ambas.
- Contains/Within: São relacionamentos simétricos, **A contains B** se e apenas se **B within A**. Uma geometria contém outra se o interior de um objeto é localizado no interior de outro objeto e as fronteiras dos dois objetos não fazem intersecção.
- Touches: Duas geometrias se encontram se elas cruzam-se, mas seus interiores não.
- Covers/CoveredBy: São relacionamentos simétricos. **A cobre B** se e apenas se **B** for coberto por **A**. Uma geometria cobre a outra se incluir todos os pontos da outra, dentro da geometria. Isto significa que a primeira geometria contém o interior de uma outra, como definido anteriormente, mas sem a restrição que as fronteiras das geometrias cruzam-se.
- Crosses: Esta operação identifica se uma geometria cruza outra e a dimensão dessa interseção é menor que a maior dimensão das geometrias.

**Predicates:** São operações que recebem dois valores espaciais como entrada e também retornam valores booleanos como respostas:

- IsEmpty: Determina se uma geometria está vazia ou não.
- OnBorder e InInterior: Determinam, respectivamente, se um ponto pertence a fronteira ou ao interior de outra geometria.

**Unárias:** Operações unárias recebem como argumento um valor espacial e retornam um novo valor espacial:

- Boundary: Retorna o anel (Borda) exterior de uma superfície.
- Buffer: Retorna uma geometria contendo todos os pontos cuja distância até a geometria passada como argumento é menor ou igual a uma determinada distância.

- Centroid: Retorna o ponto central de uma geometria.
- ConvexHull: Retorna a geometria de convexo mínimo que encerra uma geometria.

**Binárias:** Operações Binárias recebem como argumento dois (ou mais) valores espaciais e retornam um novo valor espacial. Intersecção, união, *Difference* e *SymDifference*, operam como na teoria dos conjuntos habituais e retornar uma geometria obtida pela aplicação da operação correspondente às geometrias fornecidas como argumento:

- Intersection: Retorna o uma geometria resultado da interseção de duas ou mais geometrias passadas como argumentos.
- Union: Retorna uma geometria resultado da união de duas ou mais geometrias passadas como argumentos.
- Difference: Retorna uma geometria (Em alguns casos a geometria resultante possui mais de um polígono) resultado das diferenças entre os argumentos passados como argumento.
- SymDifference: Retorna uma geometria com os pontos que não se cruzam entre as geometrias passadas como argumentos.

**Numéricas:** Operações Numéricas recebem uma ou duas geometrias como argumento e retornam um valor numérico correspondente a operação utilizada:

- NoComponents: Retornam o número máximo de subconjuntos disjuntos de uma geometria.
- Length: Retorna o comprimento de uma geometria
- Area: Retorna a área de uma geometria.
- Perimeter: Retorna o tamanho da Boundary ou Surface.
- Distance: Retorna a distância mínima entre todos os pares de pontos de duas geometrias.
- Direction:Retorna o ângulo em graus da linha entre dois pontos passados como argumentos.

### 3.4 Modelagem de Dados Espaciais

A modelagem de dados espaciais é mais trabalhosa do que a de dados tradicionais, pois, envolve características espaciais, por exemplo localização, o tempo de observação,

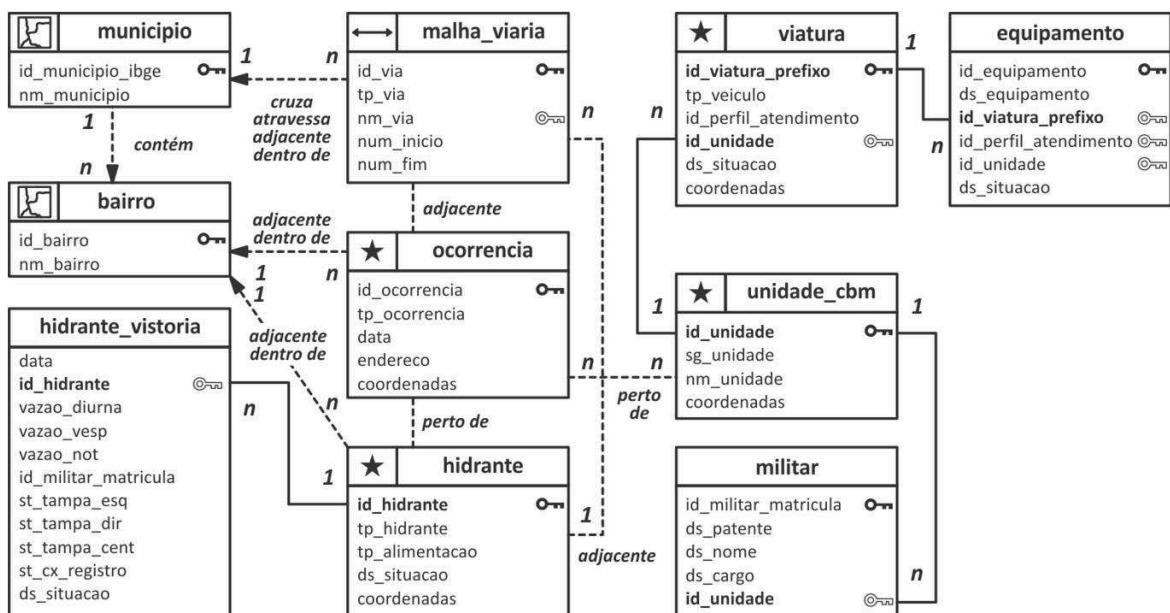
a precisão de obtenção/representação e existência de relações espaciais. Existem diversos modelos que podem ser utilizados para realizar a modelagem de dados espaciais, alguns são apresentados a seguir:

- GeoOOA (Object-oriented analysis for geographic informationsystems): As extensões incluídas pelo modelo GeoOOA permitem a representação de contextos semânticos que diferem entidades com ou sem representação espacial, utiliza-se de pictogramas para identificar a relação, entidade e classe espacial, diferencia relacionamentos espaciais dos básicos e suporta abstrações de classes espaciais.
- GMOD (Geographic Model): Nesse modelo os sistemas espaciais têm dois tipos de classes, uma que simboliza entidades associadas a alguma localização espacial (geo-Classe) e outra que não possui referência espacial, mas está relacionada com as primeiras (Classes Convencionais).
- GeoIFO (Geographic Is-a relationships, Functional relationships, complex Objects): Os objetos do mundo real são ligados ao espaço através de sua localização. A localização de um objeto inclui sua posição, geometria, tamanho e orientação. Em modelos semânticos a posição de um objeto pode ser modelada como um tipo especial de atributo, assim, posição seria uma função que retornaria para cada objeto espacial uma parte do espaço.
- GISER (Geographic Information Systems Entity Relational model): É baseado em quatro conceitos básicos, espaço-tempo, feições, coberturas e objeto espacial. As feições representam entidades geográficas que são divididas em elementos denominados coberturas. Uma cobertura consiste de um conjunto de objetos espaciais, distribuídos no tempo e no espaço.
- GeoFrame: Fornece um diagrama de classes básicas para auxiliar o projetista nos primeiros passos da modelagem conceitual de dados de uma nova aplicação. O GeoFrame inclui, por exemplo, a modelagem dos aspectos espaciais da informação geográfica e a diferenciação entre objetos convencionais e objetos/campos geográficos.
- GeoProfile: Caracteriza-se por ser uma proposta para integrar as particularidades dos principais modelos conceituais de dados espaciais. Assim, o GeoProfile não é um novo modelo, mas sim uma compilação e integração dos construtores de SIG.
- OMT-G (Object Modeling Technique for Geographic Applications): É o modelo mais conhecido, pois, é o que têm pesquisas mais recentes, ele baseado em três conceitos principais, classes, relacionamentos e restrições de integridade espaciais. As classes e relacionamentos representam as primitivas básicas para a criação de

um esquema estático de uma aplicação. A identificação de restrições de integridade espacial consiste em identificar as condições para garantir a integridade do banco de dados (BORGES; JR.; LAENDER, 2001).

Segundo (LIZARDO; JR., 2014), "OMT-G é um modelo de dados dotado de recursos para o projeto de bancos de dados e aplicações geográficas. O OMT-G parte das primitivas definidas para o diagrama de classes da Unified Modeling Language (UML), introduzindo primitivas geográficas com o objetivo de aumentar a capacidade de representação semântica daquele modelo e, portanto reduzindo a distância entre o modelo mental do espaço a ser modelado e o modelo de representação usual".

Figura 7 – Diagrama OMT-G do sistema de controle de ocorrências.



Fonte: (SANTOS, 2013)

Na Figura 7 é possível verificar um diagrama utilizando o modelo OMT-G sobre as ocorrências do Corpo de Bombeiros do Estado do Pará. É possível observar que as classes espaciais tem um pictograma identificando qual representação espacial ela identifica.

### 3.5 Modelos para representações de dados espaciais

Para representar os dados espaciais em um mapa de forma coerente com o mundo-real, salvaguarda a escala da representação, se faz necessário a utilização da referência espacial (SRID) que será utilizado pelo SIG para colocar o objeto no seu devido lugar.

### 3.5.1 Sistemas de referências espaciais

Para representar o formato da Terra existem duas referências principais que são usadas, o geoide e o elipsoide. O geoide é um modelo de referência para a superfície da Terra que se equipara com o nível médio do mar e sua extensão imaginária através dos continentes. É utilizado em geodésia (Geodésia é a ciência que estuda as dimensões, forma e o campo de gravidade da Terra, permitindo analisar, medir e representar o espaço geográfico do planeta com precisão.) para medir elevações de superfície precisas. Contudo, o geoide não é muito prático para produzir mapas. O elipsoide é uma superfície matematicamente definida que se aproxima do geoide. O elipsoide é usado para medir a localização dos pontos de interesse usando latitude e longitude. Estas são medidas dos ângulos (em graus) a partir do centro da Terra para um ponto na superfície da Terra. A Latitude mede ângulos nas direções Norte-Sul, enquanto a Longitude mede ângulos nas direções Leste-Oeste. Enquanto um elipsoide aproxima a forma da Terra, um DATUM (modelo matemático teórico da representação da superfície da Terra ao nível do mar) define onde na Terra um elipsoide será ancorado.

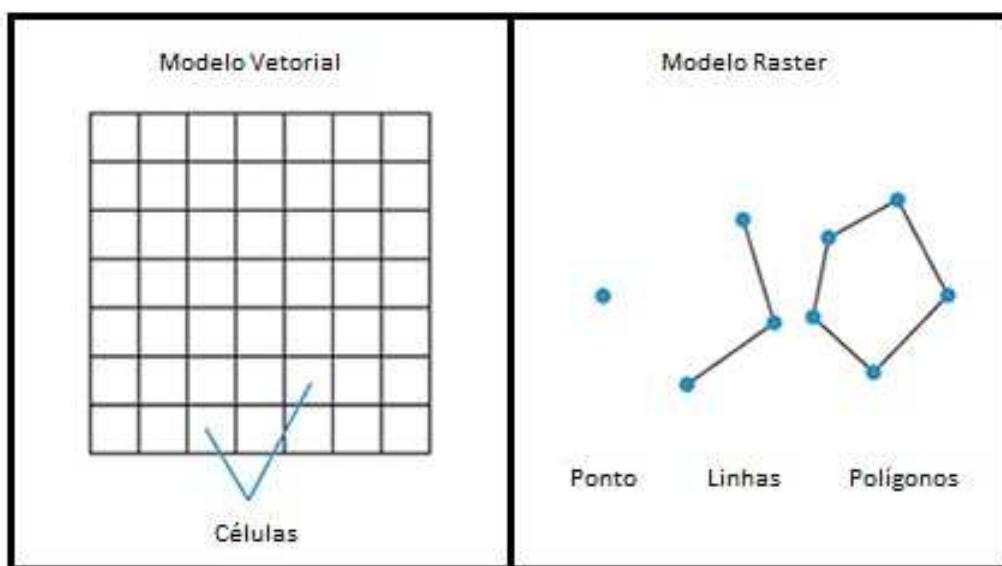
Para produzir um mapa, a superfície de referência da curvatura da Terra (geoide), aproxima-se de um elipsoide, que deve ser transformado no plano do mapa por meio de uma projeção. Portanto, um ponto na superfície de referência da Terra com coordenadas geográficas expressadas por latitude e longitude é transformado em coordenadas (x,y) cartesianas (ou mapa) representando posições no mapa plano, contudo, como uma projeção de mapa necessariamente causa deformações, diferentes projeções são usadas para diferentes propósitos, dependendo qual informação é preservada, nomeadamente, formas e ângulos, áreas, distância ou direções. Essas características estão em conflito (não é possível preservar as formas e ângulos, bem como a área), e portanto, a importância atribuída a cada uma dessas características dita a escolha de uma projeção específica. Um sistema de referência espacial (SRS) atribui coordenadas em um espaço matemático para uma localização no espaço do mundo real. Um SRS define pelo menos as unidades de medidas do sistema de coordenadas subjacentes (tal como graus ou metros), as coordenadas máximas e mínimas (também referido como os limites), a unidade linear de medida padrão, se os dados são planares ou esferoides (Um esferoide é uma superfície quádrica em três dimensões obtida através da rotação de uma elipse ao redor de um de seus eixos principais) e informação de projeção para transformar o dado em outro SRSs. SRSs são em geral bons para apenas uma região específica do globo. Se duas geometrias estão no mesmo SRS, elas podem ser sobrepostas sem distorção, se não for o caso, elas devem ser transformadas. Como há milhares de SRSs, cada um é identificado por um sistema de identificação de referência espacial (SRID).

### 3.5.2 Modelos de análise

A maioria dos elementos que existem na natureza podem ser representados por formas geométricas (pontos, linhas ou polígonos, isto é, vetores) ou por células com informações (raster), conforme exemplificado na Figura 8. São formas intuitivas e versáteis de ilustrar o espaço, que ajudam a entender melhor os elementos em estudo de acordo com sua natureza.

Dependendo da maneira de representar o espaço que eles usam, podemos classificar os SIGs em dois grandes modelos ou formatos:

Figura 8 – Modelo Raster e Vetorial.



Fonte: Autoria própria(2018)

- **Modelo Vetorial.** A escolha de um modelo ou outro dependerá se as propriedades topológicas são importantes para a análise. Se esse for o caso, o modelo de dados vetoriais é a melhor opção, mas sua estrutura de dados, embora muito precisa, é muito mais complexa e isso pode retardar o processo. Também é mais fácil optar por uma estrutura de dados vetoriais quando é necessário demonstrar mais de um atributo no mesmo espaço. O modelo vetorial fornece uma coleção de tipos de dados para representar objetos espaciais no computador. Portanto, por exemplo, enquanto no nível abstrato um objeto linear é definido como uma coleção infinita de pontos, no nível de implementação cada linha deve ser aproximada usando pontos, linhas e curvas como primitivas.
- **Modelo Raster.** Se a análise que interessa não requer o recurso a propriedades topológicas, o uso do formato raster é mais rápido, simples e eficiente. No nível lógico,

campos contínuos devem ser representados de maneira discreta. Por isto, é preciso primeiro dividir o domínio espacial em um número finito de elementos, chamado de Tesselação. Então, será necessário atribuir um valor do campo a um ponto representativo para cada elemento da partição.

## 3.6 Indexação Geográfica

Os índices espaciais são muitos importantes para realização de consultas em uma grande base de dados, porque sem sua utilização seria necessário realizar uma busca sequencial por todos os registros das tabelas consultadas, o que demandaria grande esforço computacional dependendo do tamanho da base de dados. A forma utilizada pela indexação para diminuir o tempo de uma consulta é organizando os dados em uma hierarquia de árvore de pesquisa. O PostgreSQL da suporte para três tipos de árvores de pesquisas, B-Trees, R-Trees e Índices GIST. Nesta seção o índice GIST é o mais comentado, pois, é o que está sendo utilizado no estudo de caso no Capítulo 4, além do que as árvores B-Trees e R-Trees não são tão eficientes para consultas de dados espaciais como os índices GIST. Existem índices que têm como principais características melhorar o desempenho da leitura e gravação de dados em tabelas muito extensas, o exemplo abordado nas próximas subseções é o índice BRIN. Os índices SP-GIST são recomendados para estruturas de dados não balanceadas, pesquisas que são compatíveis com a regra de particionamento podem ser muito rápidas, pois, é em partições que os espaços de busca são divididos.

### 3.6.1 Índices GIST

De acordo com o trabalho *What is the GIST?: Understanding the Use of Public GISTS on GitHub* ([WANG et al., 2015](#)) um Índice GIST (Generalized Search Tree), em português Árvore de Pesquisa Generalizada, é uma API que permite a desenvolvedores da linguagem C utilizem a indexação do PostgreSQL. Isso acontece porque o GIST adiciona uma API ao sistema de indexação do PostgreSQL que pode ser implementado para dados espaciais. Os índices GIST utilizam uma estrutura semelhante a um Índice B-Tree, que utiliza uma cópia ordenada dos dados da coluna indexada, realiza uma busca binária do valor da consulta, divide em índices menores e quando a tabela pesquisa é muito grande o índice salva os valores em uma árvore binária ou balanceada para que o PostgreSQL gerencie a memória adequadamente. Mas na utilização do Índice GIST o PostgreSQL não cria a estrutura de árvore, mas sim, trabalha com a implementação de funções, como *Union*, *Penalty* e *PickSplit*. Quando uma consulta é executada o PostgreSQL usa a função *Consistent* que recebe uma parte da árvore e uma consulta e retorna *NO* se os dois argumentos não retornam um dado item, caso contrário *MAYBE*.

O grande diferencial do índice GIST é que quem implementa a API pode decidir

que tipo de dados indexar e como organizar os valores na Árvore GIST, o PostgreSQL simplesmente chama a função *Consistent* para encontrar os valores de destino sem se importar com os valores dados no argumento ou com a aparência da árvore.

Na Figura 9 está um exemplo de criação de um índice GIST na base de dados utilizada no Capítulo 4, neste exemplo de índice a coluna utilizada é a geom que é do tipo *geometry*.

Figura 9 – CREATE INDEX

```

1  CREATE INDEX i_lim_pais_a_geom
2  ON bc250.lim_pais_a USING gist
3  (geom)
```

Fonte: Autoria própria(2018)

A Construção de um índice espacial é computacionalmente muito cara e demorada, então durante a criação não é permitido que seja gravado novos dados na tabela, para que isso não atrapalhe um ambiente de produção é possível utilizar o modo *CONCURRENTLY*. Quando esta opção é usada, o PostgreSQL irá construir o índice sem realizar nenhum bloqueio que impeça inserções, atualizações ou exclusões concorrentes na tabela, na Figura 10.

Figura 10 – CREATE INDEX CONCURRENTLY

```

1  CREATE INDEX CONCURRENTLY i_lim_pais_a_geom
2  ON bc250.lim_pais_a USING gist
3  (geom)
```

Fonte: Autoria própria(2018)

### 3.6.2 Índices BRIN

Os Índices BRIN (*Block Range Index*), em português índice de intervalo de blocos, concentram-se em casos de usos semelhantes aos dos índices SP-GIST, porque têm resultados melhores quando existe um ordenamento dos dados e são muito grandes. O índice BRIN ajuda em consultas grandes quando os dados são de certa maneira agrupados por si só, como códigos de endereços postais (CEP), que são semelhantes em uma mesma cidade, o BRIN garante que os CEPs estejam próximos uns dos outros no disco rígido. O índice BRIN permite que grandes quantidades de dados semelhantes sejam podados (KERSTIENS, 2017).

De acordo com o manual ([BASTON, 2018](#)) a ideia de um índice BRIN para dados espaciais é armazenar apenas a caixa de *bouding* englobando todas as geometrias contidas em todas as linhas em um conjunto de blocos de tabelas, chamado de intervalo. Obviamente, este método de indexação só será eficiente se os dados forem fisicamente ordenados de uma maneira em que as caixas de projeção resultantes para intervalos de blocos serão mutuamente exclusivas. O índice resultante será realmente pequeno, mas em muitos casos será menos eficiente que um índice GIST.

Na Figura 11 está a sintaxe para construção de um índice BRIN.

Figura 11 – CREATE INDEX BRIN

```

1   CREATE INDEX i_rel_pico_p_geom
2   ON bc250.rel_pico_p USING brin
3   (geom)
```

Fonte: Autoria própria(2018)

### 3.6.3 Índices SP-GiST

Os índices SP-GIST ([AREF; ILYAS, 2001](#)) são mais úteis quando seus dados possuem um elemento de agrupamento natural e também não são uma árvore igualmente equilibrada. Um ótimo exemplo disso são números de telefone (pelo menos dos EUA). Eles seguem um formato de: 3 dígitos para o código de área, 3 dígitos para prefixo (historicamente relacionado ao switch de uma operadora de telefonia) e 4 dígitos para o número da linha. Isso significa que você tem um agrupamento natural em torno do primeiro conjunto de 3 dígitos, em torno do segundo conjunto de 3 dígitos, e os números podem se espalhar em uma distribuição mais uniforme. Mas, com números de telefone alguns códigos de área têm uma saturação muito maior do que outros. O resultado pode ser que a árvore esteja muito desequilibrada. Por causa desse agrupamento natural na frente e a distribuição desigual de informações, dados como números de telefone poderiam ser um bom argumento para o SP-GIST.

### 3.6.4 Usando Índices

Ao serem criados os índices aceleraram o acesso aos dados sem que o usuário perceba, quem decide quando o índice será utilizado é o *query planner*, mas não é sempre que ele utiliza o índice, fazendo com que às vezes as consultas sejam realizadas de maneira sequencial percorrendo toda a tabela, dependendo da seletividade.

Existem maneiras para descobrir se os índices espaciais que foram criados estão sendo utilizados, como especificado pelo manual de desenvolvimento ([BASTON, 2018](#)), destacado nos seguintes itens:

- Em primeiro lugar, lê-se o plano de consulta e verifique se a sua consulta realmente tenta calcular o que necessita. Uma condição inoperante do *JOIN* esquecida ou a tabela errada, pode inesperadamente trazer *EXPLAIN* na frente da consulta.
- Segundo, certifique-se de que as estatísticas sejam reunidas sobre o número e as distribuições de valores em uma tabela, para fornecer ao planejador de consulta melhores informações para tomar decisões sobre o uso do índice. *VACUUM ANALYZE* irá calcular ambos.
- Deve-se aspirar regularmente seus bancos de dados de qualquer forma. Muitos DBAs do PostgreSQL têm o *VACUUM* sendo executado como um trabalho *cron off-peak* em uma base regular.
- Se a limpeza não ajudar, force temporariamente o planejador à usar as informações de índice usando o conjunto *enable\_seqscan to off; command*. Dessa forma, pode-se verificar se o planejador é capaz de gerar um plano de consulta acelerado por índice para sua consulta. Deve-se usar este comando apenas para depuração, em geral, o planejador sabe melhor do que você sobre quando usar índices. Depois de executar sua consulta, não se esqueça de voltar *ENABLE\_SEQSCAN*, para que outras consultas utilizem o planejador normalmente.
- Se definido, *enable\_seqscan off*, ajuda sua consulta a ser executada, seu Postgres provavelmente não está ajustado para seu hardware. Se você achar que o planejador está errado sobre o custo das varreduras sequenciais versus índice, tente reduzir o valor de *random\_page\_cost* no *postgresql.conf* ou usando o conjunto *random\_page\_cost para 1.1*. O valor padrão para o parâmetro é 4, tente defini-lo como 1 (em SSD) ou 2 (em discos magnéticos rápidos). Diminuir o valor torna o planejador mais inclinado a usar as varreduras de índice.
- Se definido, *enable\_seqscan off*, não ajudar a sua consulta, pode acontecer que você use uma construção Postgres ainda não é capaz de desvendar. Uma sub-consulta com seleção *inline* é um exemplo, você precisa reescrevê-la para que o planejador de formulários possa otimizar, digamos, uma *LATERAL JOIN*.

### 3.6.5 Consultas Espaciais

A partir da criação dos índices é possível realizar consultas utilizando-os com funções espaciais com melhores desempenhos. Um exemplo de função na qual é utilizado

o índice GIST é a função *ST\_AREA*, que calcula a área de um polígono e retorna o valor em graus de circunferências. Na Figura 12 é possível observar uma consulta que utiliza-se da função *ST\_AREA* recebendo como argumento a coluna geom da tabela *bc250.lim\_municipio\_a* (Essa tabela está disponível no banco de dados importado para realização do estudo de caso no cap.4) e multiplica o retorno pela constante 12321 para descobrir quais municípios do Brasil têm área maior do 100000 quilômetros quadrados. O resultado da consulta retornou quatro cidades.

Figura 12 – ST\_AREA.

```

1   SELECT nome
2     FROM bc250.lim_municipio_a
3    WHERE (st_area(geom)*12321) > 100000

```

Fonte: Autoria própria(2018)

Na Figura 13, o exemplo dado é da função *ST\_COVEREDBY* que recebe duas geometrias como argumento verifica se a primeira geometria está totalmente dentro da segunda, retornando um valor booleano. No exemplo verificamos se uma cidade (60235 = Uberlândia) pertence a um estado (59522 = Minas Gerais), para compor a consulta foi utilizado a função *ST\_ASEWKT* para converter a coluna geom da tabela *bc250.loc\_cidade\_p* (Essa tabela está disponível no banco de dados importado para realização do estudo de caso no cap.4) do modelo WKT utilizado para uma representação onde está o SRID e as coordenadas geográficas do ponto onde é localizado a cidade. O resultado da consulta é *true*.

Figura 13 – ST\_COVEREDBY.

```

1   SELECT St_coveredby((SELECT St_asewkt(Cast((SELECT geom
2     FROM bc250.loc_cidade_p
3     WHERE id_objeto = 60235) AS TEXT)))
4   , Cast(
5     (SELECT geom
6       FROM bc250.lim_unidade_federacao_a
7       WHERE id_objeto = 59522) AS TEXT))

```

Fonte: Autoria própria(2018)

## 3.7 Ferramentas

Para armazenar, manipular e consultar dados espaciais é necessário o acesso a um *Geodatabase*, que é uma base de dados relacionais que suporta tais tipos de dados. Os principais SGBDs podem armazenar *Geodatabases* com grandes volumes de dados, além de manipular formas geométricas, como linhas, pontos, polígonos e usar recursos mais avançados, como topologia, anotações, comportamento, objetos 3D, entre outros.

Manter um banco de dados espacial é uma tarefa que requer muita responsabilidade, tempo e dinheiro, pois, os dados podem constituir uma parte vital para o funcionamento de uma empresa, por exemplo, uma companhia de fertilizantes não pode de forma alguma perder as suas informações sobre as localizações que está utilizando determinado tipo de defensivo agrícola ou mesmo ter as informações incorretas em sua base de dados. Por esses motivos a gestão dos dados espaciais é uma das mais importantes funções desempenhadas por um SGBDG e um SIG. A definição do que é SIG sempre gera muita discussão. Alguns autores incluem a fase de aquisição dos dados, outros não, alguns excluem os produtos (mapas) como parte do GIS, outros incluem os profissionais como peças de um GIS.

Uma definição da University of Edinburgh que parece bem apropriada e prática de que SIG é um sistema informatizado para captura, armazenamento, verificação, integração, manipulação, análise e visualização de dados relacionados a posições na superfície terrestre. O SIG nada mais é que a ferramenta utilizada para manipular, armazenar e consultar os dados geográficos/espaciais, seja em forma de dados alfanuméricos ou com a utilização direta por mapas, essa forma é quando o usuário do SIG realiza uma consulta e é retornado um mapa ou uma imagem com a posição consultada ([MELO, 2016](#)).

O principais SGBDs possuem extensões para dados espaciais que podem ser conectados a SIGs e assim visualizar os dados em forma de representações espaciais. Segue alguns exemplos dos principais SGBDGs disponíveis.

### 3.7.1 PostGIS

O PostGIS é a ferramenta para dados espaciais dos bancos de dados relacionais da PostgreSQL. No PostGIS há suporte para objetos geográficos, onde as consultas de localizações sejam realizadas através de comandos SQL. A licença desse software é pública, o PostGIS é distribuído gratuitamente de acordo com os termos da Licença GNU/GPL (General Public License). Outras vantagens em utilizar o PostGIS é sua compatibilidade com os padrões OGC (Open Geospatial Consortium, organização internacional que regulamenta as características que os dados geográficos devem ser representados), é possível utilizar índices espaciais e centenas de recursos para análise e tratamento espacial, possui facilidade para exportar/importar dados, em sua estrutura estão dispostos diversas ferra-

mentas de conversão nativas do sistema, fácil integração com aplicações SIG e WebGIS, o PostGIS é capaz de comunicar-se com a maioria deles, com mais facilidade se utilizarem os padrões OGC.

O PostGIS é uma excelente alternativa para uso comercial, pois além de ser gratuito, é a ferramenta para construção de banco de dados geoespaciais *open source* mais utilizada no mundo, pois tem a mesma estabilidade e rapidez que os softwares proprietários oferecem.

### 3.7.2 Oracle Spatial

O Oracle Spatial 11g, é a versão para dados espaciais do Banco de Dados *Oracle Enterprise Edition*, sua licença é do tipo proprietário, o Oracle utilizou pela primeira vez o recurso de dados espaciais com uma modificação no Oracle 4 feita por pesquisadores que trabalham com o CHS, *Canadian Hydrographic Service*, empresa do governo canadense que é responsável por coletar, gerenciar, transformar e disseminar dados e informações batimétricas, hidrográficas e náuticas em papel e cartas náuticas eletrônicas, bem como publicações e serviços hidro-espaciais, incluindo serviços de atualização de avisos para envio e/ou avisos aos navegantes. Uma equipe de desenvolvimento conjunto da equipe da CHS e da Oracle redesenhou o kernel do Oracle, resultando na *Spatial Data Option* ou "SDO" para Oracle 7.

É uma das ferramentas com mais funcionalidades disponíveis no mercado, apresenta recursos para utilização de aplicações que se fazem necessário recursos para dados geoespaciais baseados em localização e sistemas empresariais de informações espaciais. O Oracle Spatial tem um número maior de recursos de localização básicos em comparação aos banco de dados convencionais da Oracle, com ele é possível gerar *buffers* agregados espaciais, calcular área, distância e referência linear, inclui também o tipo de dados Geo-Raster para armazenar e gerenciar dados e metadados de raster de imagens e em grade, modelos de dados de rede e topologia, mecanismos de geocodificação e criação de rotas, APIs para implantação rápida e fácil de serviços de mapeamento, geocodificação e criação de rotas, funções de análise espacial e exploração, plataforma completa de gerenciamento de dados para os requisitos avançados de qualquer aplicativo geoespacial ou empresarial habilitado para localização, suporte para dados tridimensionais, e novos tipos de dados suportam armazenamento e gerenciamento de nuvens de pontos e modelos de terreno, encontrados em domínios como planejamento urbano, segurança nacional ou produção de mapas baseada em Lidar (*Light Detection And Ranging*, é uma tecnologia óptica de detecção remota que mede propriedades da luz refletida de modo a obter a distância e/ou outra informação a respeito um determinado objecto distante) ([IHM; LOPEZ; RAVADA, 2007](#)).

### 3.7.3 IBM DB2 Spatial Extender

O DB2 *Spatial Extender* disponibiliza recursos de análise e gerenciamento a dados espaciais para bancos de dados DB2, sua licença é do tipo proprietário, ele incorpora um SIG ao DB2 e implementa a especificação OGC. Com o DB2 *Spatial Extender* é possível implementar colunas em uma tabela capazes de armazenar dados espaciais, como a localização de um marco, uma rua ou "pedaço" de terra. Para que um banco DB2 comum possa utilizar as funções e dados espaciais faz-se necessário que o *Spatial Extender* esteja ativado. Após ativado será possível criar tabelas com colunas espaciais que o próprio DB2 *Spatial Extender* converte dados espaciais em formato de armazenamento a partir de formatos externos como WKT, (Representação de texto), WKB (Representação binária), GML (Representação em linguagem de marcação geográfica) e representação do formato Esri (formato Esri é suportado por diversos programas de Geoprocessamento é formato vetorial mais amplamente utilizado no mundo inteiro.). O suporte ao Formato Esri se dá porque foi justamente com a colaboração da *Environmental Systems Research Institute* que o DB2 *Spatial Extender* foi desenvolvido. A Esri que é líder de mercado em GISs e todas as ferramentas e aplicativos ESRI, bem como todos os aplicativos que usam a API ESRE SDE, funcionarão no estado em que se encontram com o DB2 *Spatial Extender*.

O DB2 *Spatial Extender* permite que os usuários visualizem e analisem dados operacionais e de negócios espaciais. Informações de localização espacial, como latitude e longitude, assim como muitos outros sistemas de coordenadas, podem ser derivadas automaticamente de dados existentes, armazenadas em um formato espacial padrão da indústria e recuperadas para análise com o auxílio de índices espaciais ([IBM KNOWLEDGE CENTER, 2013](#)).

### 3.7.4 SQL Server (SPATIAL DATA)

A Microsoft adicionou suporte no SQL Server 2008, SPATIAL DATA, com a introdução de tipos de dados espaciais nativos para representar objetos espaciais. Ao mesmo tempo, a Microsoft adicionou a funcionalidade necessária para acessar e indexar dados espaciais, fornecer otimizações baseadas em custo e suportar operações como a interseção de dois objetos espaciais. Os tipos de dados espaciais do SQL Server, que são baseados em coordenadas de latitude e longitude, permitem armazenar esses objetos e disponibilizá-los para os aplicativos de chamada. O SQL Server suporta dois tipos de dados espaciais, *Geometry* e *Geography*, cada um dos tipos de dados espaciais tem seu próprio uso. Por exemplo, o tipo *Geography* geralmente é usado para armazenar os dados de GPS de um aplicativo, enquanto o tipo *Geometry* é frequentemente usado para mapear um objeto tridimensional, como um prédio. Juntos, os dois tipos de dados espaciais suportam 16 tipos de objetos de dados. Dos 16, 11 são instanciáveis (podem ser representados) em um banco de dados. Eles derivam determinadas propriedades de seu tipo de dados pai e

são essas propriedades que os distinguem como objetos, pontos ou sequências de linhas ([KATIBAH; STOJIC, 2011](#)).

## 4 Estudo de Caso

Este trabalho apresenta o estado da arte da manipulação de dados geográficos por meio de Sistemas de Gerenciamento de Bancos de Dados Relacionais por meio da linguagem SQL. Para tanto, criou-se um experimento prático para demonstração das técnicas e ferramentas abordadas.

Foi utilizado para criar o banco de dados a ferramenta pgAdmin4, no banco foi criada uma extensão PostGIS. A partir do site do IBGE ([IBGE, 2017](#)), foi obtido o arquivo bc250\_2017-11-08.tar que depois de restaurado no banco, é criado um schema chamado bc250, contendo informações geográficas de todo território brasileiro.

Figura 14 – CREATE EXTENSION.

```

1 CREATE EXTENSION postgis
2   SCHEMA public
3   VERSION "2.5.0";

```

Fonte: Autoria própria (2018)

### 4.1 O Banco

O banco criado recebeu o nome de IBGE, com a criação do schema bc250 no banco, foram importadas 61 tabelas que têm informações de Hidrografia, Sistema de Transportes, Energia e Comunicações, Abastecimento de Água e Saneamento Básico, Estrutura Econômica, Localidades, Limites e Relevo de todo o território nacional. Os dados contidos nesse banco são referentes a estudos e pesquisas das mais diversas instituições reguladoras, desde a ANA (Agência Nacional de Águas) à SNUC (Sistema Nacional de Unidades de Conservação), segundo o próprio IBGE o bc250 é "Um conjunto de dados geoespaciais de referência, estruturados em bases de dados digitais, permitindo uma visão integrada do território nacional, sendo de grande importância para projetos de planejamento regional, de cunho ambiental e de gestão do território."([NETO; COSTA, 2017](#)).

No comando para criação do banco de dados, foi utilizado como padrão de grafia 'Portuguese\_Brazil.1252' em LC\_COLLATE e LC\_CTYPE, ordem de classificação de cadeia e classificação de caracteres, respectivamente, essas categorias afetam a ordem de classificação dos índices, portanto, devem ser mantidas fixas ou os índices nas colunas

de texto ficarão corrompidos. O limite de conexões ficou como -1, que é o padrão do PostgreSQL. A codificação utilizada foi UTF-8.

Figura 15 – CREATE DATABASE.

```

1   CREATE DATABASE "Ibge"
2       WITH
3           OWNER = postgres
4           ENCODING = 'UTF8'
5           LC_COLLATE = 'Portuguese_Brazil.1252'
6           LC_CTYPE = 'Portuguese_Brazil.1252'
7           TABLESPACE = pg_default
8           CONNECTION LIMIT = -1;

```

Fonte: Autoria própria(2018)

Dentre as 61 tabelas que foram importadas no schema bc250, existem registros dos mais diversos tipos, como, cidades, capitais, rios, ferrovias, cemitérios, barragens, etc. Na grande maioria das tabelas foram criados variáveis com tipos de dados espaciais que são utilizados posteriormente em um SIG para representar de maneira ilustrativa suas localizações em um mapa.

Um exemplo é a tabela loc\_capital\_p, que têm armazenada os registros das vinte e sete capitais dos estados do Brasil.

Figura 16 – CREATE TABLE.

```

1 CREATE TABLE bc250.loc_capital_p
2 (
3     id_objeto integer NOT NULL,
4     nome character varying(100) COLLATE pg_catalog."default",
5     nomeabrev character varying(50) COLLATE pg_catalog."default",
6     geometriaaproximada character varying(3) COLLATE pg_catalog."
7         default",
8     tipocapital character varying(20) COLLATE pg_catalog."default
9         ",
10    geom geometry(Point,4674),
11    id_produtor integer,
12    id_elementoprodutor integer,
13    CONSTRAINT loc_capital_p_pk PRIMARY KEY (id_objeto)
14 )

```

Fonte: Autoria própria (2018)

Nesta tabela existem variáveis com tipos de dados padrões como sua chave primária, id\_objeto, que é representada por um tipo inteiro ou a variável nome que é do tipo string. Por outro lado existe a variável geom que é do tipo geometry, onde a forma de representação é Point e o sistema de referência espacial é o 4674 que representa o geocêntrico da América do Sul, é possível verificar todos SRID na tabela **spatial\_ref\_sys** que está integrada a extensão PostGIS, o campo **srtext** desta tabela contém informações de qual local no planeta faz referência o SRID.

O tipo de dado espacial mais básico usado no PostGIS é o tipo **geometry**, esse tipo de dados utiliza um formato de codificação chamado WKT (Well Known Text), que foi definido pelo Open Geospatial Consortium (OGC) na Norma Internacional ISO 19125-1: 2004, “Informações geográficas - Acesso a recursos simples - Parte 1: Arquitetura comum”, ([LOTT, 2013](#)). No PostgreSQL a função que codifica a string de um Insert que contém as coordenadas e o tipo de dados especificado para armazenar no banco de dados é a função **st\_geometryfromtext(text,integer)** que gera um valor hexadecimal e posteriormente é utilizado pelo SIG para representar de em forma de figuras geométricas em um mapa de acordo com as coordenadas geográficas e o sistema de referência espacial.

Na Figura 17 segue um exemplo simples da criação de duas tabelas e de suas representações, também há uma imagem com a representação das duas tabelas em forma de camadas no QGIS. Uma tabela com a representação de seis pontos utilizando o tipo **Point** e outra com a representação de dois polígonos cobrindo cinco dos seis pontos utilizando o tipo **Polygon**.

## 4.2 O SIG

O SIG que escolhido para utilizar com os dados do banco em diferentes projeções de mapas foi o QGIS, que é um software livre, multiplataforma. No QGIS é possível criar, editar e visualizar mapas utilizando os dados já preenchidos em um banco de dados espacial, é necessário a criação de uma conexão entre o QGIS e o banco de dados. As tabelas do banco de dados conectadas ao QGIS são demonstradas pelo SIG através de camadas, onde uma pode se sobrepor a outra ou completando-a.

Na Figura 20 é apresentado a projeção dos limites dos países da América do Sul e dos trechos ferroviários, através de duas tabelas do banco, a tabela lim\_pais\_a e tra\_trecho\_ferroviario\_l, onde cada uma das tabelas tem uma variável espacial, MultiPolygon e MultiLineString, respectivamente.

Em outro caso há a projeção das unidades federativas com suas respectivas capitais, nessa projeção foram utilizadas duas tabelas, lim\_unidade\_federacao\_a e loc\_capital\_p, cada qual com uma variável espacial com os tipos MultiPolygon e Point, respectivamente.

Figura 17 – CREATE pontos.

```
1      CREATE TABLE PUBLIC.pontos
2      (
3          id_objeto INTEGER NOT NULL,
4          nome CHARACTER varying(100) collate pg_catalog."default",
5          geom geometry(point ,4674),
6          CONSTRAINT pontos_pk PRIMARY KEY (id_objeto)
7      );
8
9      INSERT INTO PUBLIC.pontos
10         VALUES      (1,
11                      'ponto 1',
12                      St_geometryfromtext('Point(-18.882468 -48.315629)', 4674)
13                      );
14
15      INSERT INTO PUBLIC.pontos
16         VALUES      (2,
17                      'ponto 2',
18                      St_geometryfromtext('Point(-18.648696 -48.192846)', 4674)
19                      );
20
21      INSERT INTO PUBLIC.pontos
22         VALUES      (3,
23                      'ponto 3',
24                      St_geometryfromtext('Point(-24.957886 -53.459450)', 4674)
25                      );
26
27      INSERT INTO PUBLIC.pontos
28         VALUES      (4,
29                      'ponto 4',
30                      St_geometryfromtext('Point(-26.898875 -49.077590)', 4674)
31                      );
32
33      INSERT INTO PUBLIC.pontos
34         VALUES      (5,
35                      'ponto 5',
36                      St_geometryfromtext('Point(-20.153760 -44.908527)', 4674)
37                      );
38
```

Fonte: Autoria própria (2018)

Figura 18 – CREATE poligonos

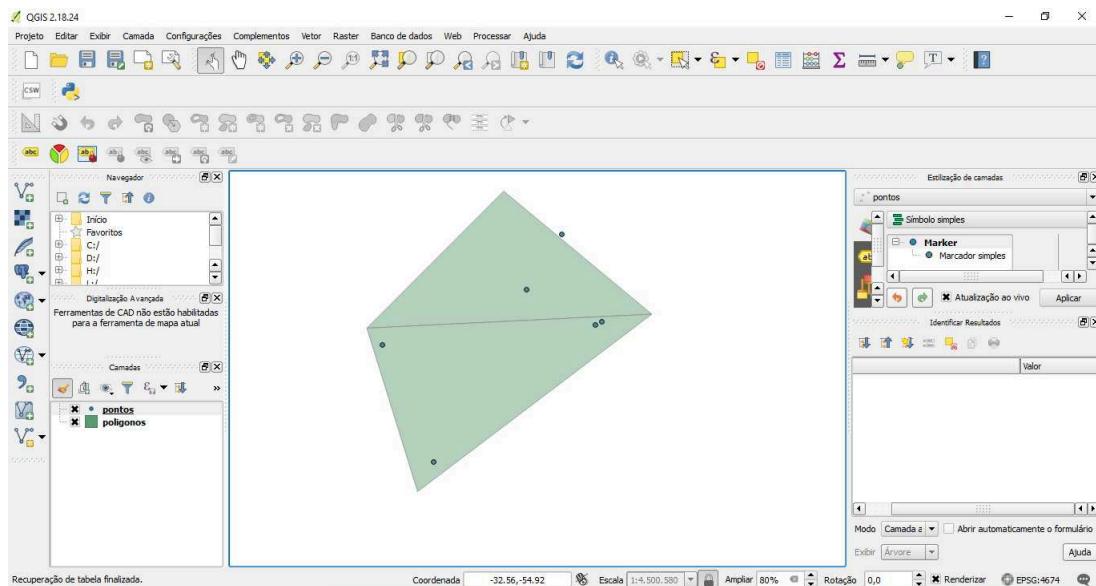
```

1   INSERT INTO PUBLIC.poligonos
2     VALUES      (1,
3       'Meu territorio',
4       St_geometryfromtext('Polygon((-16.775688 -47.911559,
5         -25.567584 -54.570176, -27.474901 -48.436373, -16.775688
6         -47.911559))', 4674));
7
8   INSERT INTO PUBLIC.poligonos
9     VALUES      (2,
10      'territorio 2',
11      St_geometryfromtext('Polygon((-16.775688 -47.911559,
12        -22.336031 -43.295736, -27.474901 -48.436373, -16.775688
13        -47.911559))', 4674));

```

Fonte: Autoria própria (2018)

Figura 19 – Polígonos e Pontos



Obtido por meio da visualização dos dados da aba Camadas na ferramenta QGIS.

Uma funcionalidade do QGIS é realizar consultas no banco de dados conectado através do gerenciador interno e os resultados, caso tenham como retorno colunas com valores espaciais, o usuário pode criar camadas temporárias e exibi-las com as camadas oriundas do banco de dados. Na Figura 22 é possível observar uma consulta na tabela `bc250.lim_unidade_uso_sustentavel_a` usando como sub-consulta a tabela `bc250.lim_municipio_a` para descobrir quais áreas de uso sustentáveis são maiores que o município de São Paulo. Para descobrir o tamanho das áreas e da cidade de São

Figura 20 – Projeção Ferrovias Brasileiras



Obtido por meio da visualização dos dados da aba Camadas na ferramenta QGIS.

Paulo foi utilizado a função **ST\_AREA** que utiliza a coluna espaciais geom e devolve como argumento o tamanho em grau de georreferência.

Na Figura 23 a consulta é projetada em forma de camada. Também é possível criar visões para posteriormente projetá-las. Na Figura 24 mostra o exemplo de uma consulta onde são obtidas as cidades com distâncias menores ou iguais a 100 km de Uberlândia, o número '1' na comparação representa 100 quilômetros de acordo com a escala especificado no manual do banco (NETO; COSTA, 2017). Para calcular a distância a função **ST\_DISTANCE**, que recebe como argumento duas coordenadas. Para obter a coordenada da cidade de Uberlândia no formato que o Postgis aceito foi utilizado a função **St\_ASEWKT** que recebe o hexadecimal armazenado na variável **geom** e converte-o, como é verificável na Figura 25, na projeção de camadas do QGIS exposta na Figura 26,

Figura 21 – Projeção Capitais Brasileiras



Obtido por meio da visualização dos dados da aba Camadas na ferramenta QGIS.

está exibida os pontos na cor verde de cidades que estão a menos de 100 km de Uberlândia, o ponto vermelho representa a própria cidade.

Na figura 27 é apresentado uma consulta que obtém as 10 cidades mais próximas de Uberlândia e na 28 a consulta é projetada. Outro exemplo da Figura 29 onde a consulta verifica dentre todas as áreas de uso sustentáveis do Brasil quais estão dentro do município de São Paulo, nessa consulta foi utilizada a função **ST\_COVEREDBY** para verificar quais a geometrias da tabela **lim\_unidade\_uso\_sustentavel** estão dentro da geometria referente a São Paulo. Na figura 30 está a apresentação das camadas referentes a consulta da 29 junto com a camada da tabela **lim\_municipio\_a**.

Figura 22 – Consulta áreas de uso sustentáveis

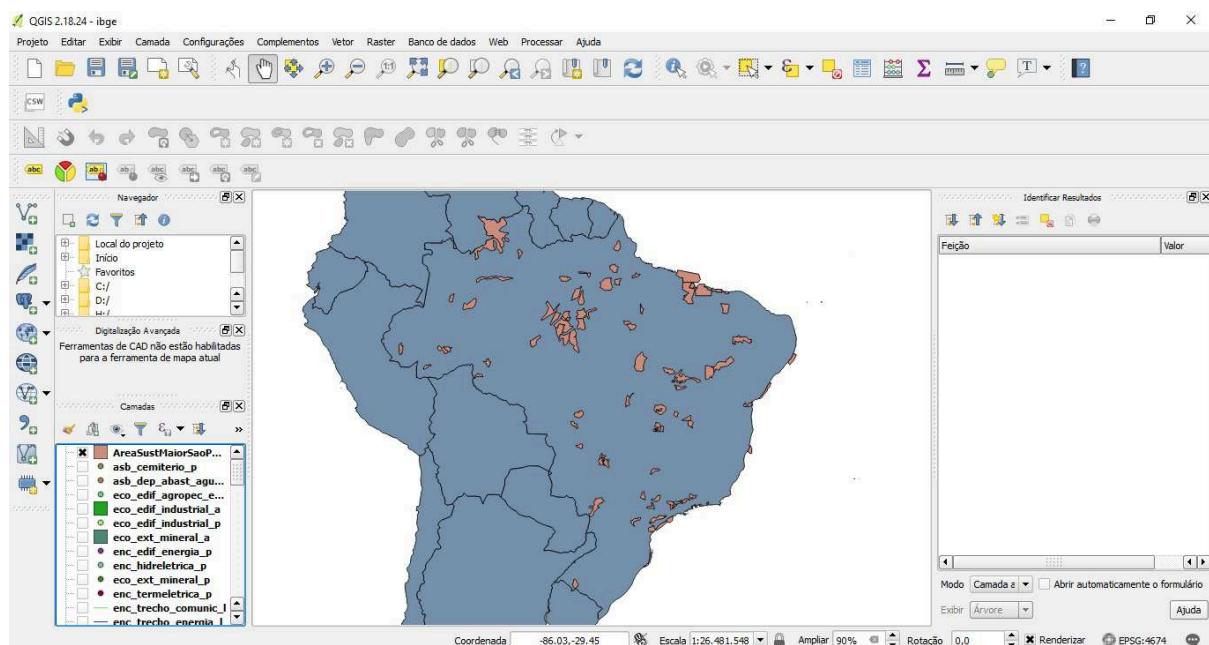
```

1   SELECT *
2   FROM   bc250.lim_unidade_uso_sustentavel_a
3   WHERE  St_area(geom) > (SELECT St_area(geom)
4                           FROM   bc250.lim_municipio_a
5                           WHERE  id_objeto = 58933)

```

Fonte: Autoria própria (2018)

Figura 23 – Áreas de uso sustentáveis



Obtido por meio do projetor de camadas disponível no QGIS.

Figura 24 – Cidades até 100 km de Uberlândia

```

1   SELECT *
2   FROM   bc250.loc_cidade_p
3   WHERE  St_distance(geom, 'SRID=4674;POINT (-48.2770000149999
4   -18.9189999959999)' )
5   <= 1
      AND id_objeto != 60235

```

Fonte: Autoria própria (2018)

Figura 25 – ST\_ASEWKT

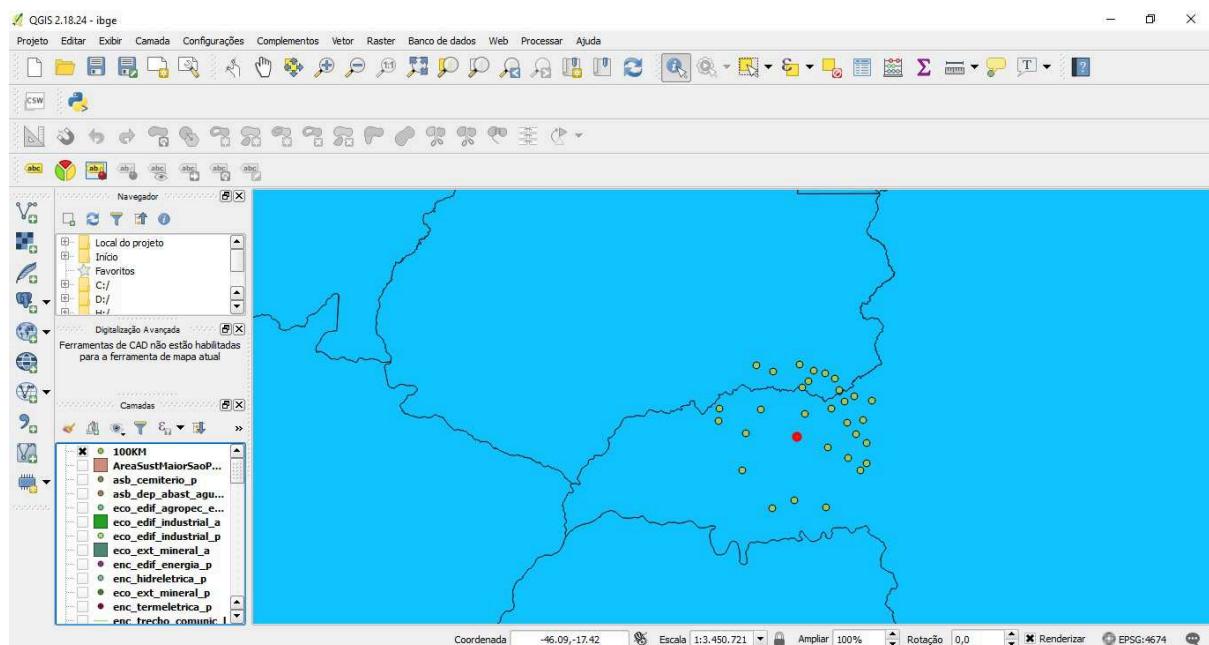
```

1   SELECT St_asewkt((SELECT geom
2     FROM   bc250.loc_cidade_p
3     WHERE  id_objeto = 60235))
4
5   >>>
6   SRID=4674; POINT (-48.2770000149999 -18.9189999959999)

```

Fonte: Autoria própria(2018)

Figura 26 – Cidades a menos de 100 km de Uberlândia



Obtido por meio do projetor de camadas disponível no QGIS.

Figura 27 – As 10 cidades mais perto de Uberlândia

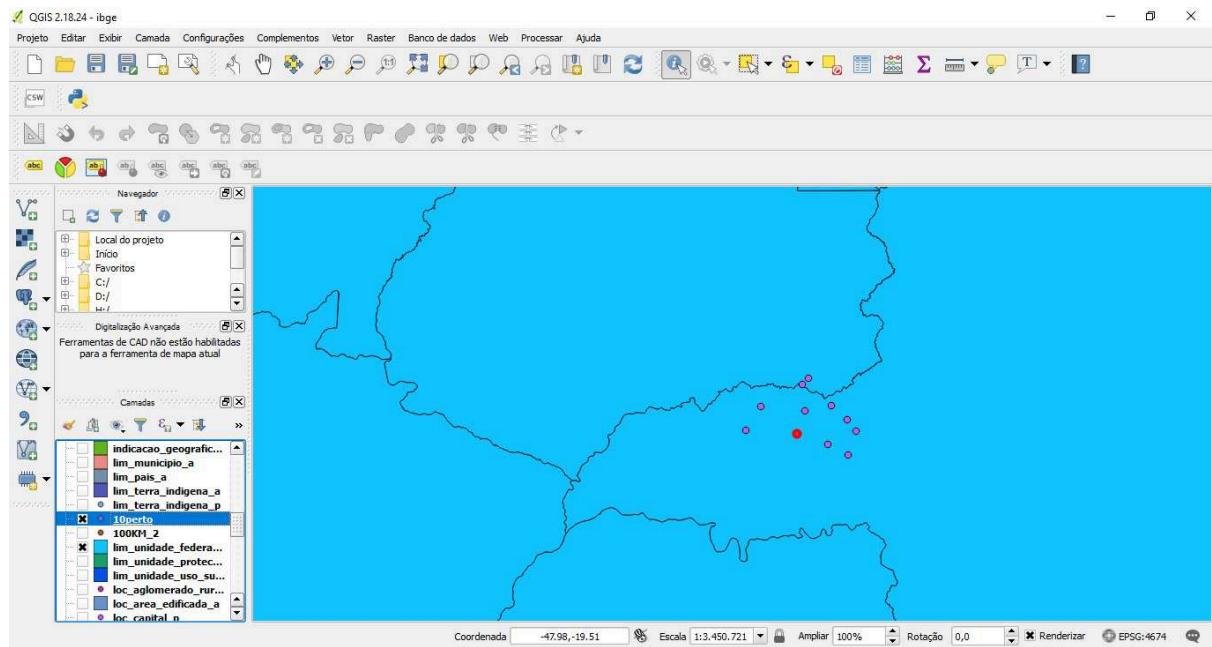
```

1   SELECT St_distance(geom , 'SRID=4674;POINT (-48.2770000149999
2   -18.9189999959999)')
3   AS
4   distancia ,
5   *
6   FROM   bc250.loc_cidade_p
7   ORDER  BY distancia ASC
8   LIMIT  11

```

Fonte: Autoria própria(2018)

Figura 28 – As 10 cidades mais perto de Uberlândia



Obtido por meio do projetor de camadas disponível no QGIS.

Figura 29 – Áreas sustentáveis dentro do município de São Paulo

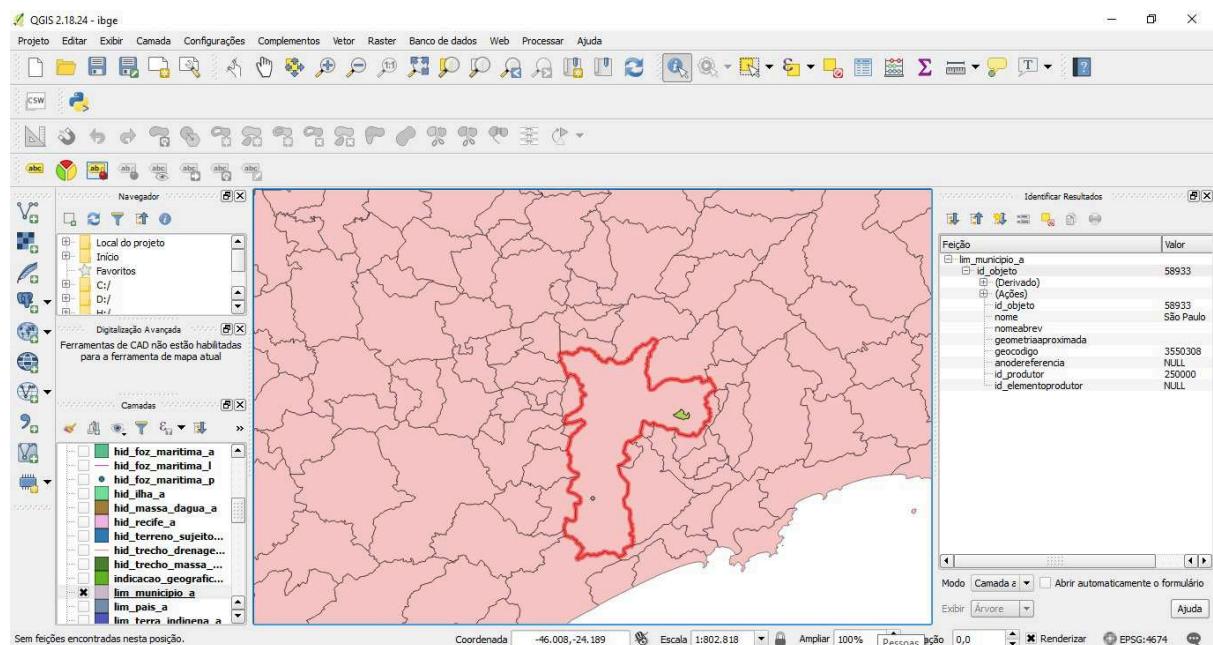
```

1  SELECT distinct(sus.geom) AS area, sus.id_objeto
2  FROM bc250.lim_unidade_uso_sustentavel_a sus
3      inner JOIN bc250.lim_municipio_a
4  ON (SELECT St_coveredBy((SELECT St_asewkt(Cast(( sus.geom )
5          AS
TEXT))), ,
6          Cast((
7              SELECT geom
8          FROM
9              bc250.lim_municipio_a
10             WHERE
11                 id_objeto = 58933)AS TEXT)))

```

Fonte: Autoria própria(2018)

Figura 30 – Exibição das áreas sustentáveis dentro do município de São Paulo



Obtido por meio do projetor de camadas disponível no QGIS.

# 5 Conclusão

Com a realização deste trabalho é possível observar o quanto abrangente é o estudo sobre dados espaciais, são diversas utilidades muito além de apenas formas de representar objetos do mundo real em mapas físicos ou digitais. Os dados espaciais auxiliam diversos outros setores da sociedade a partir de estudos realizados por técnicos e acadêmicos das mais variadas áreas do conhecimento.

Este trabalho tentou contribuir, compilando diversos outros estudos realizados referentes aos conceitos dos SGBDGs, suas utilizações, demonstrar que não é complexo trabalhar com tipos de dados espaciais, qualquer pessoa com o mínimo domínio em banco de dados pode realizar aventuras ou mesmo grandes inovações utilizando banco de dados espaciais.

## 5.1 Trabalhos Futuros

Entre os trabalhos futuros pode-se destacar a validação experimental dos índices espaciais para execução de consultas topológicas, aos vizinhos mais próximos (k-nearest neighbor queries) e por abrangência (range queries). Apoiando-se nos avanços nas tecnologias de reconhecimento de imagens, juntamente integrados aos SGBDGs pode-se realizar estudos mais detalhados sobre arborização em cidades e quantidade de áreas onde há drenagem de água para evitar alagamentos em dias de fortes chuvas. Esses são apenas alguns exemplos de uma vasta gama deles que podem existir.

## Referências

- AREF, W. G.; ILYAS, I. F. Sp-gist: An extensible database index for supporting space partitioning trees. *Journal of Intelligent Information Systems*, v. 17, n. 2, p. 215–240, Dec 2001. ISSN 1573-7675. Disponível em: <<https://doi.org/10.1023/A:1012809914301>>. Citado na página 32.
- BASTON, J. A. A. *PostGIS 2.5.2dev Manual*. Berkeley, CA, 2018. Disponível em: <<https://postgis.net/docs/>>. Acesso em: 11 nov. 2018. Citado 2 vezes nas páginas 32 e 33.
- BORGES, K. A. V.; JR., C. A. D.; LAENDER, A. H. F. OMT-G: an object-oriented data model for geographic applications. *GeoInformatica*, v. 5, n. 3, p. 221–260, 2001. Disponível em: <<https://doi.org/10.1023/A:1011482030093>>. Citado na página 27.
- DELBONI, T. M.; BORGES, K. A. V.; LAENDER, A. H. F. Geographic web search based on positioning expressions. In: *Proceedings of the 2005 Workshop on Geographic Information Retrieval*. New York, NY, USA: ACM, 2005. (GIR '05), p. 61–64. ISBN 1-59593-165-1. Disponível em: <<http://doi.acm.org/10.1145/1096985.1097000>>. Citado na página 21.
- DEMPSEY, C. *Open Source GIS and Freeware GIS Applications*. 2012. Disponível em: <<https://www.gislounge.com/history-of-gis/>>. Citado 3 vezes nas páginas 14, 15 e 16.
- DEVMEDIA, E. *Sistemas de Informações Geográficas: aplicações e utilidades - Parte 01*. 2008. Disponível em: <<https://www.devmedia.com.br/sistemas-de-informacoes-geograficas-aplicacoes-e-utilidades-parte-01/7782>>. Citado na página 4.
- GOODCHILD, M. F. Geographic information system. In: *Encyclopedia of Database Systems*. [s.n.], 2009. p. 1231–1236. Disponível em: <<https://link.springer.com/referencework/10.1007%2F978-0-387-39940-9>>. Citado na página 11.
- HANNAVY, J. *Encyclopedia of Nineteenth Century Photography*. New York: Routledge, 2008. Citado na página 15.
- IBGE. *bc250*. [S.l.], 2017. Disponível em: <<https://www.ibge.gov.br/geociencias-novoportal/cartas-e-mapas/bases-cartograficas-continuas/15759-brasil.html?=&t=downloads>>. Acesso em: 17 out. 2018. Citado 2 vezes nas páginas 12 e 39.
- IBM KNOWLEDGE CENTER. *O DB2 Spatial Extender*. [S.l.], 2013. Disponível em: <[https://www.ibm.com/support/knowledgecenter/es/SSEPGG\\_10.1.0/com.ibm.db2.luw.spatial.topics.doc/doc/csbp1001.html](https://www.ibm.com/support/knowledgecenter/es/SSEPGG_10.1.0/com.ibm.db2.luw.spatial.topics.doc/doc/csbp1001.html)>. Citado na página 37.
- IHM, J.; LOPEZ, X.; RAVADA, S. *Oracle Spatial 11g: gerenciamento avançado de dados espaciais para aplicativos empresariais*. [S.l.], 2007. Disponível em: <<https://www.oracle.com/technetwork/pt/database/spatial-11g-whitepaper1.pdf>>. Citado na página 36.

- KATIBAH, E.; STOJIC, M. *Dados espaciais (SQL Server)*. [S.I.], 2011. Disponível em: <<https://docs.microsoft.com/pt-br/sql relational-databases/spatial-spatial-data-sql-server?view=sql-server-2017>>. Citado na página 38.
- KERSTIENS, C. *A tour of Postgres Index Types*. São Francisco, CA, 2017. Disponível em: <<https://www.citusdata.com/blog/2017/10/17/tour-of-postgres-index-types/>>. Acesso em: 05 dez. 2018. Citado na página 31.
- LEI 10267. *Presidência da República - Casa Civil - Subchefia para Assuntos Jurídicos*, 2001. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/LEIS\\_2001/L10267.htm](http://www.planalto.gov.br/ccivil_03/leis/LEIS_2001/L10267.htm)>. Citado na página 10.
- LIU, L.; ÖZSU, M. T. (Ed.). *Encyclopedia of Database Systems*. Springer US, 2009. ISBN 978-0-387-35544-3. Disponível em: <<https://doi.org/10.1007/978-0-387-39940-9>>. Citado na página 10.
- LIZARDO, L. E. O.; JR., C. A. D. OMT-G designer: A web tool for modeling geographic databases in OMT-G. In: *Advances in Conceptual Modeling - ER 2014 Workshops, ENMO, MoBiD, MReBA, QMMQ, SeCoGIS, WISM, and ER Demos, Atlanta, GA, USA, October 27-29, 2014. Proceedings*. [s.n.], 2014. p. 228–233. Disponível em: <<https://link.springer.com/book/10.1007%2F978-3-319-12256-4>>. Citado na página 27.
- LOTT, R. *Geographic information - Well-known text representation of coordinate reference systems*. [S.I.], 2013. Disponível em: <<http://docs.opengeospatial.org/is/12-063r5/12-063r5.html>>. Citado na página 41.
- MARTINELLI, M. *Mapas da geografia e cartografia temática*. [S.I.]: Editora Contexto, 2003. Citado na página 10.
- MELO, E. *GIS*. 2016. Disponível em: <<http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm>>. Citado na página 35.
- NETO, W. J. S.; COSTA, P. D. A. V. *BASE CARTOGRÁFICA CONTINUA DO BRASIL, ESCALA 1:250.000 - BC250*. [S.I.], 2017. Disponível em: <[ftp://geoftp.ibge.gov.br/cartas\\_e\\_mapas/bases\\_cartograficas\\_continuas/bc250/versao2017/informacoes\\_tecnicas/Documentacao\\_bc250\\_v2017.pdf](ftp://geoftp.ibge.gov.br/cartas_e_mapas/bases_cartograficas_continuas/bc250/versao2017/informacoes_tecnicas/Documentacao_bc250_v2017.pdf)>. Citado 3 vezes nas páginas 12, 39 e 44.
- QUEIROZ, G. R.; FERREIRA, K. R. *Tutorial sobre Bancos de Dados Geográficos*. [S.I.], 2006. Disponível em: <[http://www.dpi.inpe.br/TutorialBdGeo\\_GeoBrasil2006.pdf](http://www.dpi.inpe.br/TutorialBdGeo_GeoBrasil2006.pdf)>. Citado 2 vezes nas páginas 18 e 20.
- SANTOS, L. S. dos. Banco de dados geográfico aplicado a gestao de informacao do corpo de bombeiros militar do estado do para. *geopara.blogspot*, 2013. Disponível em: <<http://geopara.blogspot.com/2013/03/banco-de-dados-geografico-aplicado.html>>. Citado na página 27.
- VAISMAN, A. A.; ZIMÁNYI, E. *Data Warehouse Systems - Design and Implementation*. Springer, 2014. (Data-Centric Systems and Applications). ISBN 978-3-642-54654-9. Disponível em: <<https://doi.org/10.1007/978-3-642-54655-6>>. Citado 2 vezes nas páginas 18 e 23.

WANG, W. et al. What is the gist?: Understanding the use of public gists on github. In: *Proceedings of the 12th Working Conference on Mining Software Repositories*. Piscataway, NJ, USA: IEEE Press, 2015. (MSR '15), p. 314–323. ISBN 978-0-7695-5594-2. Disponível em: <<http://dl.acm.org/citation.cfm?id=2820518.2820556>>. Citado na página 30.

## Anexos

Para poder trabalhar com a manipulação de dados espaciais primeiramente é necessário que seja instalado em uma máquina um gerenciador de banco de dados, um sistema de informação geográfica, criar, popular o banco e conectar o banco e o SIG. Nos anexos abaixo está detalhado quais são os passos que devem ser seguidos para realizar a instalação de todos os componentes para operar dados espaciais em um ambiente igual ao do estudo de caso.

# ANEXO A – Instalação do PostgreSQL

Foi optado instalar o PostgreSQL por ser um dos SGBDs de código aberto mais avançado que existe, para instalá-lo acessem o site do PostgreSQL <<https://www.postgresql.org/download/windows/>>, escolha o sistema operacional e clique no link **Download the installer**, os usuários serão direcionados para o site da EnterpriseDB, que é uma empresa privada que fornece software e serviços baseados no banco de dados de código aberto PostgreSQL, <<https://www.enterprisedb.com/downloads/postgres-postgresql-downloads>>, onde será possível escolher qual a versão do PostgreSQL, opte pela versão mais recente.

Execute o arquivo **postgresql-11.0-2-windows-x64** como administrador e prosseguir com os próximos passos.

- 1º passo: Clique na opção **next** para começar a instalação;
- 2º passo: Escolha o diretório onde será instalado o PostgreSQL;
- 3º passo: Escolha uma senha para acessar o PostgreSQL;
- 4º passo: Identifique um número de porta que o servidor escutará, por padrão a porta é 5432;
- 5º passo: Escolha a aplicação mais recente para instalar;
- 6º passo: Selecione o PostGIS na versão mais atual para instalação;

**Obs.:** Para a instalação do PostGIS a partir da rede da UFU por meio do proxy, após o término da instalação do PostgreSQL, deve-se fechar a aplicação StackBuilder, abrir um prompt de comandos, execute as seguintes instruções:

Figura 31 – Proxy

```

1 cd C:\Program Files\PostgreSQL\11\bin
2 set https_proxy=http://proxy.ufu.br:3128
3 set HTTPS_PROXY=http://proxy.ufu.br:3128
4 C:\Program Files\PostgreSQL\11\bin>stackbuilder.exe

```

- 7º passo: Escolha o local onde o Banco de dados será executado;
- 8º passo: Clique em **next** para iniciar a instalação;
- 9º passo: Será exibido a janela de progressão da instalação;
- 10º passo: A instalação foi realizada, clique em **Finish**.

Depois da instalação do PostgreSQL concluída, vá ao Menu Iniciar do Windows e digite **pgAdmin4**, execute como administrador, a aplicação será aberta no seu browser, a primeira coisa à fazer é criar uma conexão, para isso clique com o botão direito do mouse sobre a opção **Servers->Create->Server**, de um nome para o servidor do PostgreSQL na aba **General**, em seguida na aba **Connection** coloque a porta e a senha que você escolheu na instalação, também preencha o endereço onde o servidor ficará hospedado.

Após a criação do banco de dados, é necessário criar a extensão espacial, para tal clique com o botão direito do mouse no item **Extensions**, na aba **General** escolha a opção **postgis** e na aba **Definition** escolha o schema public.

Para utilizar o banco de dados que o IBGE disponibiliza, você precisa acessar o site <[https://downloads.ibge.gov.br/downloads\\_geociencias.htm](https://downloads.ibge.gov.br/downloads_geociencias.htm)> **geociências->cartas\_e\_mapas->bases\_cartograficas\_continuas->bc250->versao2017->postgis** e baixar o arquivo **bc250\_2017-11-08.tar**, em seguida vá ao pgAdmin4, clique com o botão direito do mouse sobre o banco de dados que você já criou e escolha a opção **Restore**, no formato escolha **Custom or tar**, procure o arquivo, escolha a opção postgres na opção **Role name**, clique em restore. Será criado um schema com o nome **bc250**, contendo as tabelas preenchidas com as variáveis espaciais.

Figura 32 – Site PostgreSQL

The screenshot shows the PostgreSQL website's "Windows installers" page. At the top, there's a navigation bar with links for Home, About, Download, Documentation, Community, Developers, Support, Donate, and Your account. A search bar is also present. Below the navigation, a banner states "18th October 2018: PostgreSQL 11 Released!". On the left, there's a "Quick Links" sidebar with options like Downloads (Binary, Source), Software Catalogue, and File Browser. The main content area is titled "Windows installers" and features a "Interactive installer by EnterpriseDB". It describes the installer as including the PostgreSQL server, pgAdmin, StackBuilder, and a package manager. It mentions graphical or silent install modes and compatibility with PostgreSQL 11. Below this, there's a "Platform support" section with a table showing PostgreSQL Version 11 running on 64 Bit Windows Platforms (2016, 2012 R2) and 32 Bit Windows Platforms.

PostgreSQL Version	64 Bit Windows Platforms	32 Bit Windows Platforms
11	2016, 2012 R2	

Figura 33 – Site EnterpriseDB

The screenshot shows the EnterpriseDB (EDB) website. At the top, there is a navigation bar with links for Transferências (Downloads), BLOG, PORTAIS (Portals), and LÍNGUA (Language). Below the navigation bar, there are dropdown menus for Produtos, Nuvem, clientes, Serviços, Treinamento, and Recursos. On the right side of the header, there are search, login, and contact links. The main content area has a dark blue header with the text "Download do PostgreSQL". Below this, there is a table listing PostgreSQL versions and their availability for different operating systems:

Versão do PostgreSQL	Linux x86-64	Linux x86-32	Mac OS X	Windows x86-64	Windows x86-32
11.0	N / D	N / D	Baixar	Baixar	N / D
10.5	Baixar	Baixar	Baixar	Baixar	Baixar
9.6.10	Baixar	Baixar	Baixar	Baixar	Baixar
9.5.14	Baixar	Baixar	Baixar	Baixar	Baixar
9.4.19	Baixar	Baixar	Baixar	Baixar	Baixar
9.3.24	Baixar	Baixar	Baixar	Baixar	Baixar

Figura 34 – 1º passo.



Figura 35 – 2º passo.

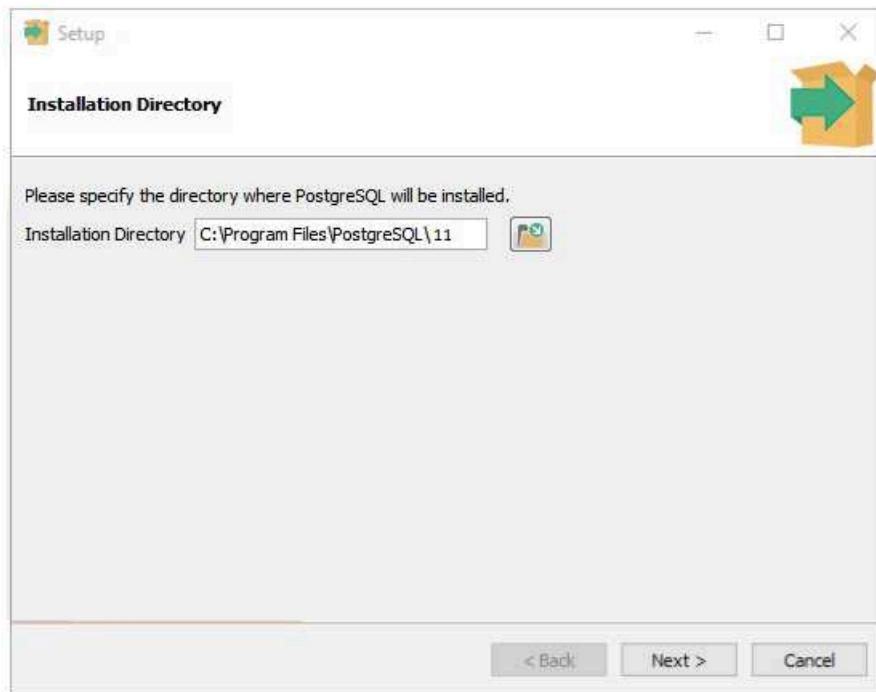


Figura 36 – 3º passo.

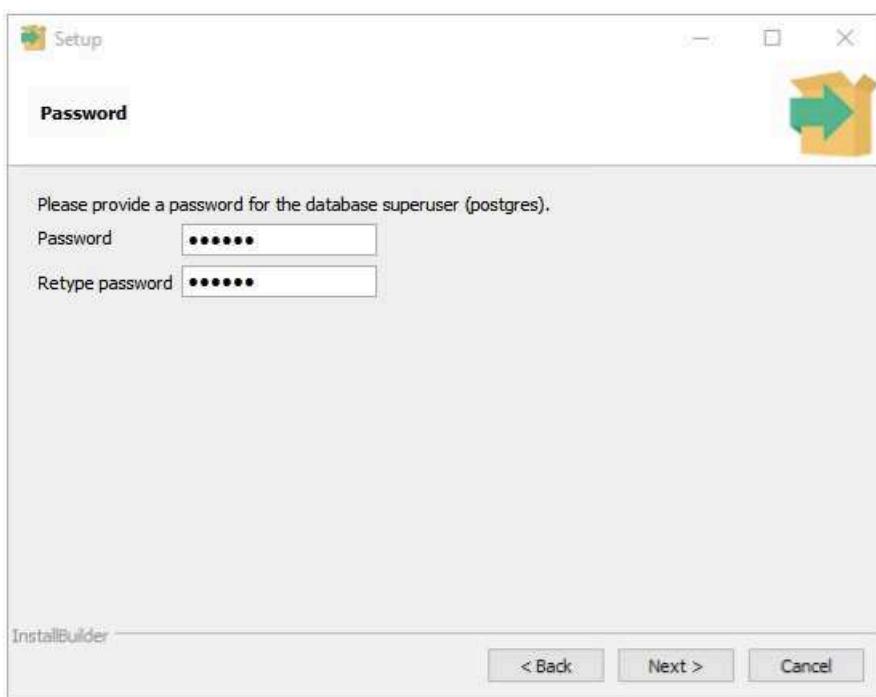


Figura 37 – 4º passo.

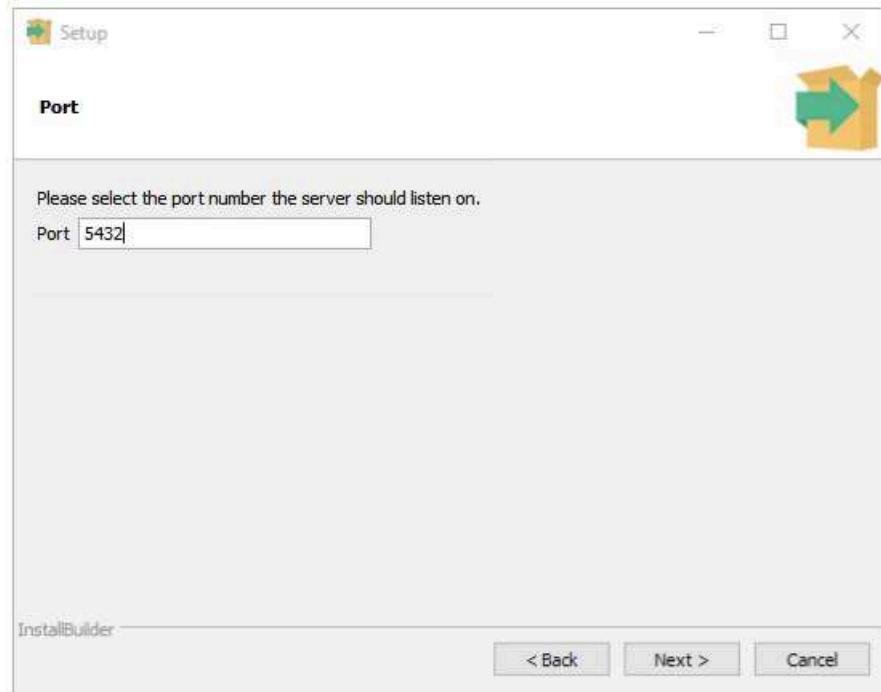


Figura 38 – 5º passo.

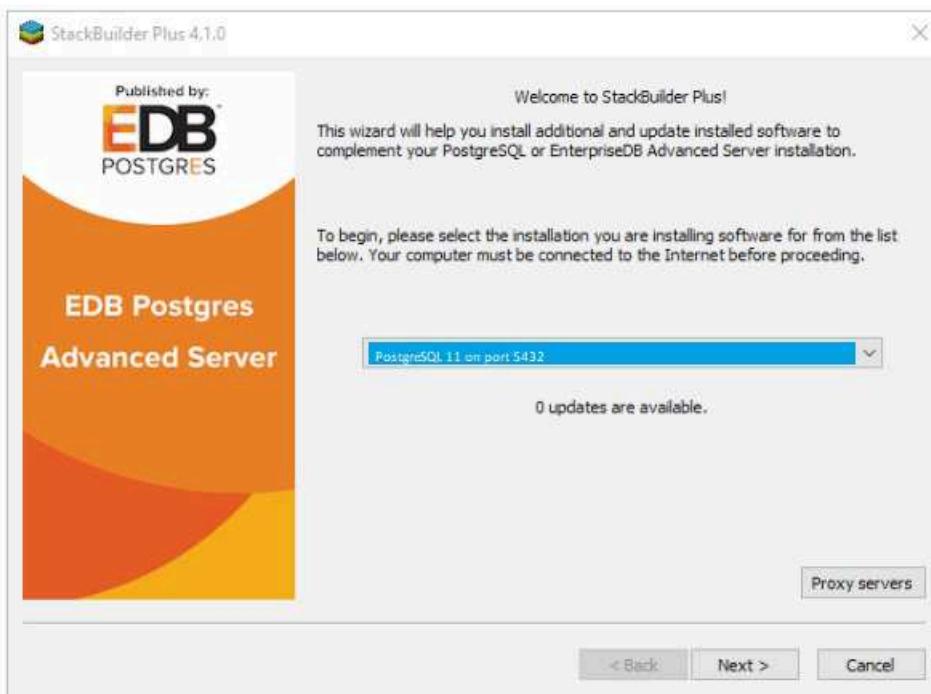


Figura 39 – 6º passo.

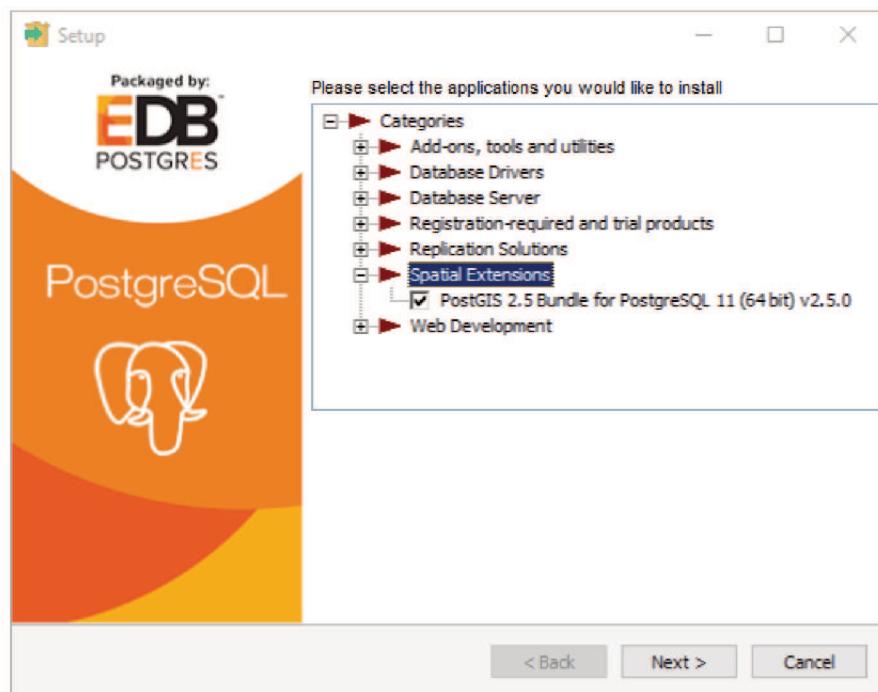


Figura 40 – 7º passo.

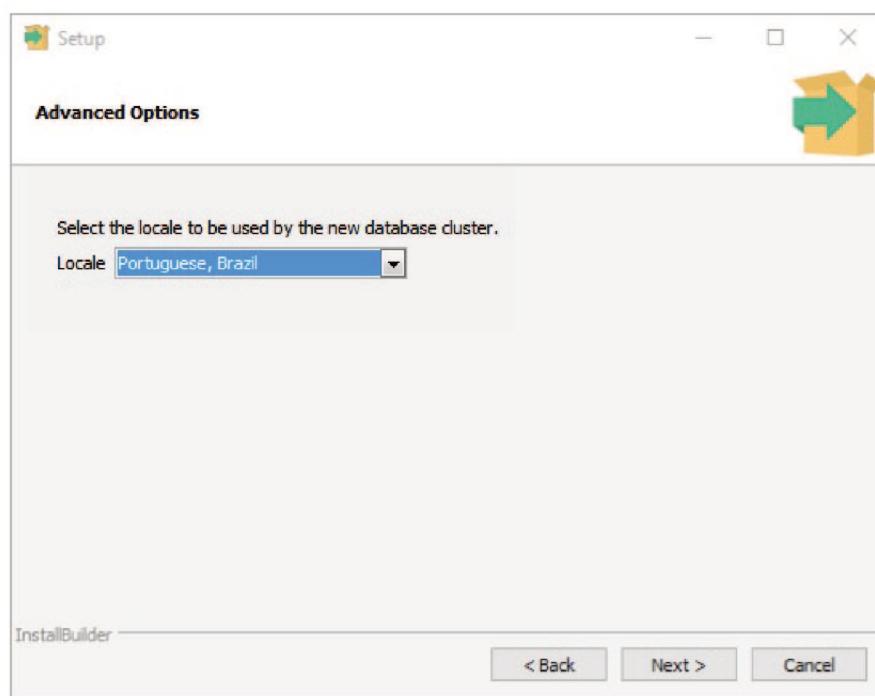


Figura 41 – 8º passo.

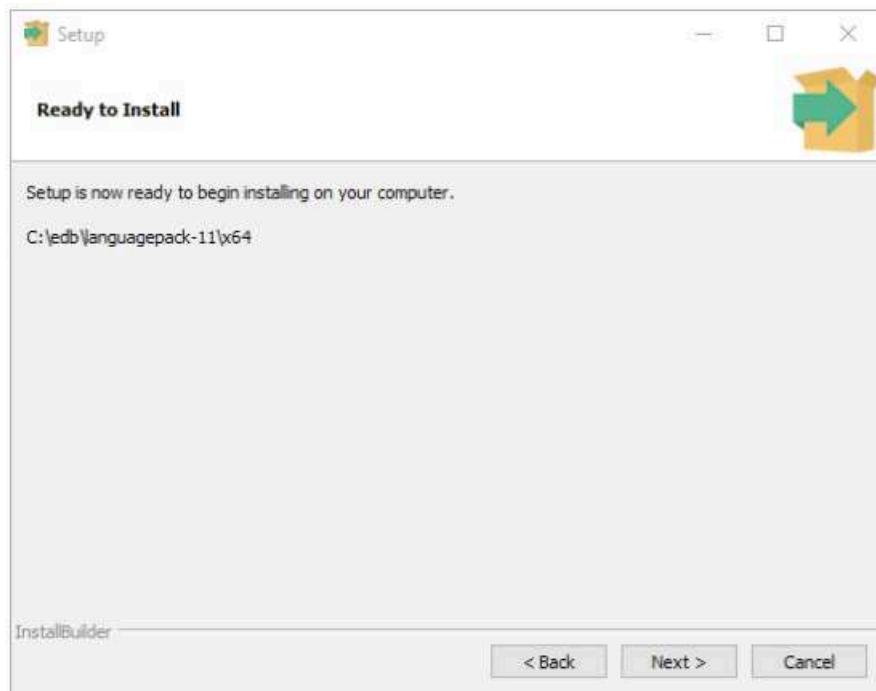


Figura 42 – 9º passo.

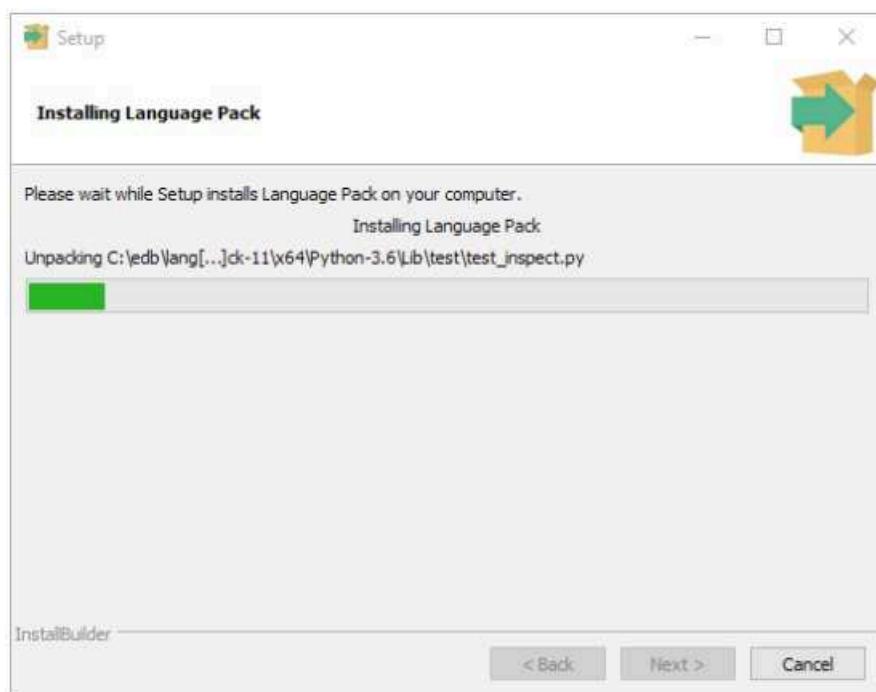


Figura 43 – 10º passo.



Figura 44 – Menu Iniciar - pgAdmin4.

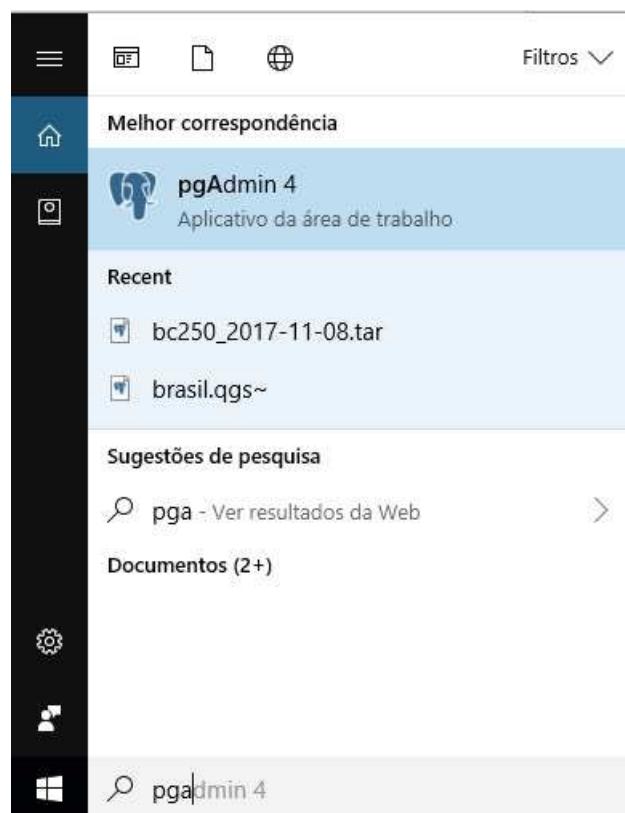


Figura 45 – Server

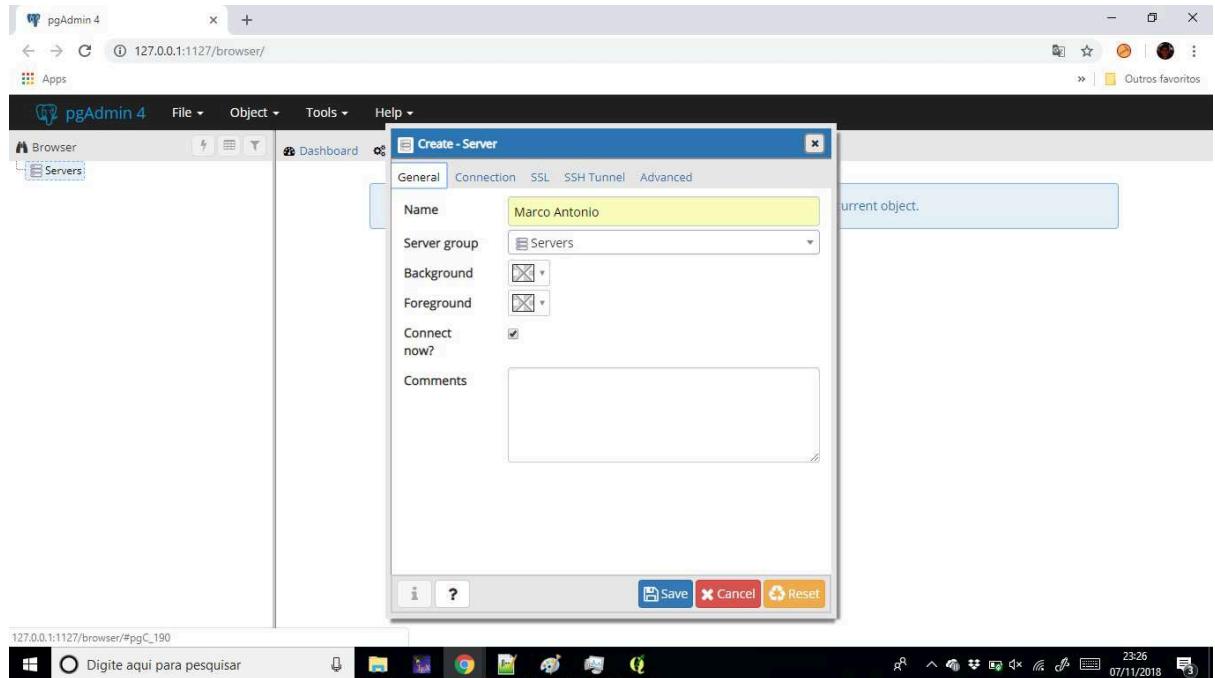


Figura 46 – Connection

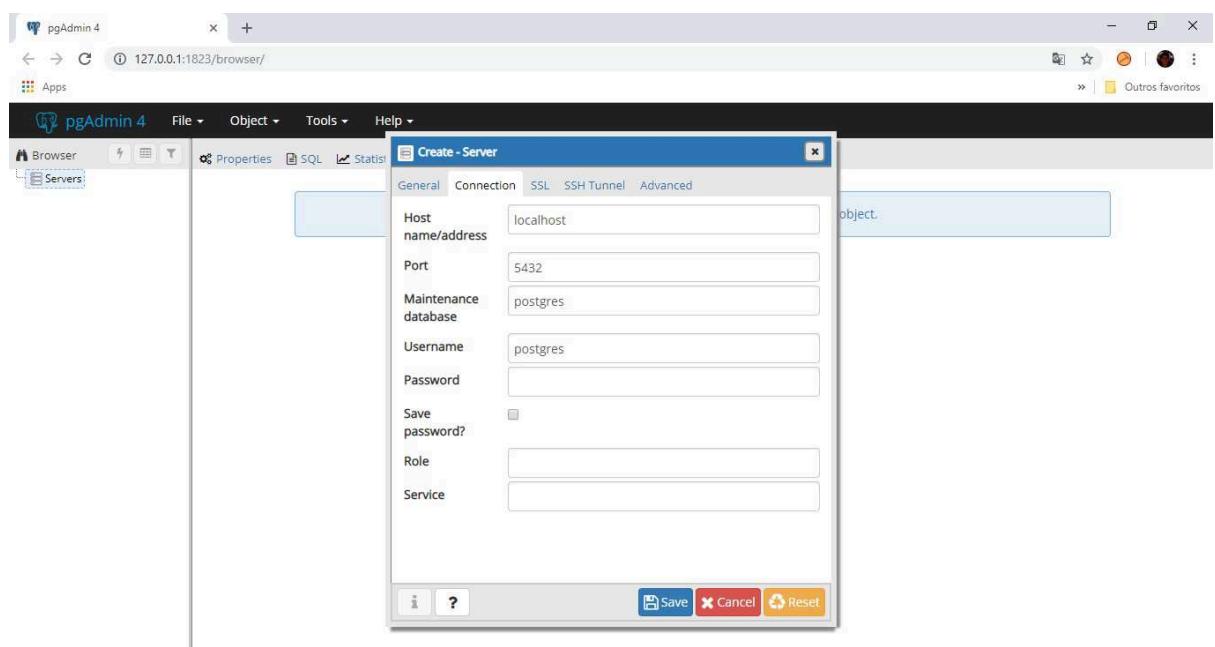


Figura 47 – Extension

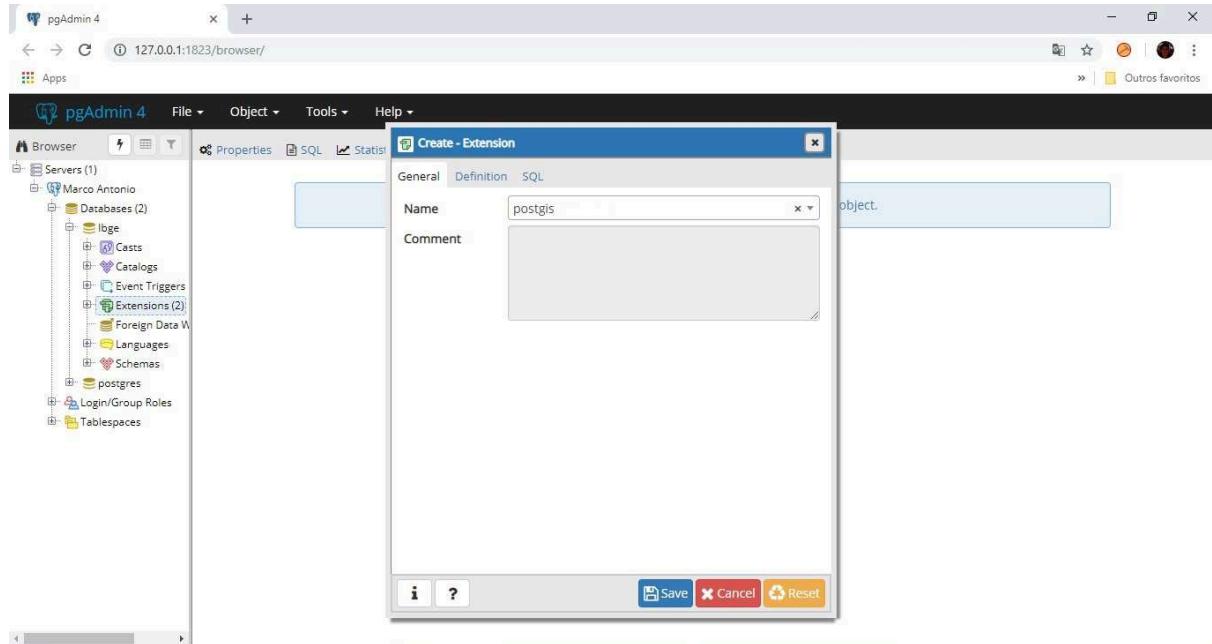


Figura 48 – Schema public

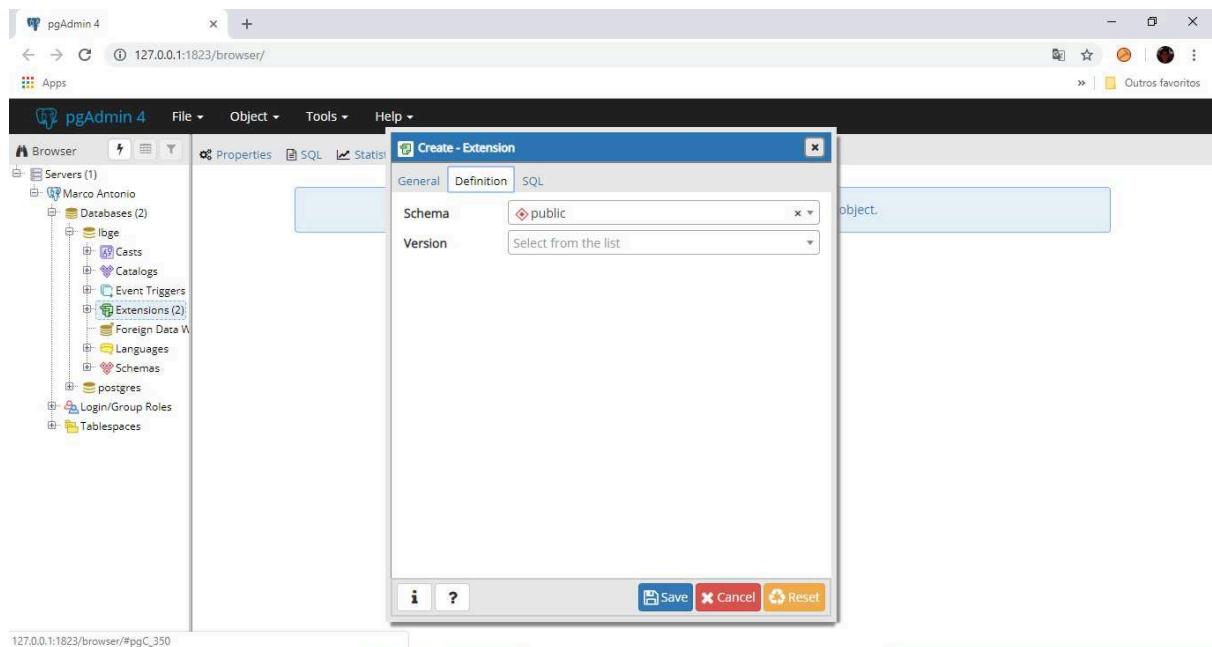


Figura 49 – Site IBGE

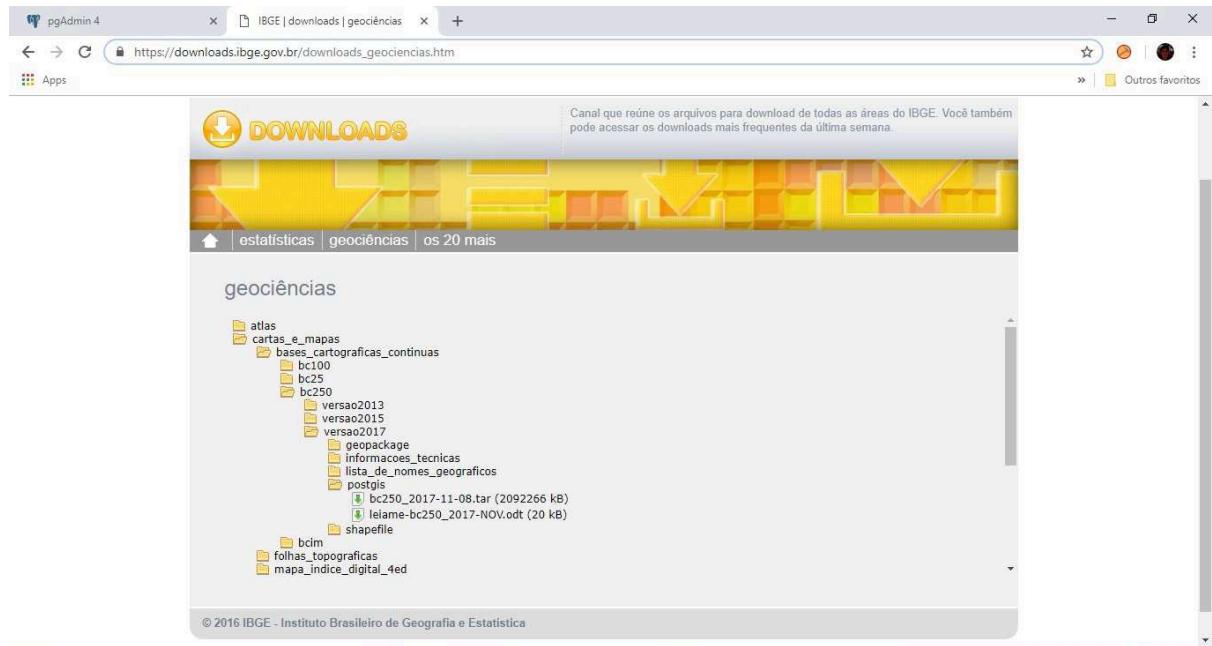
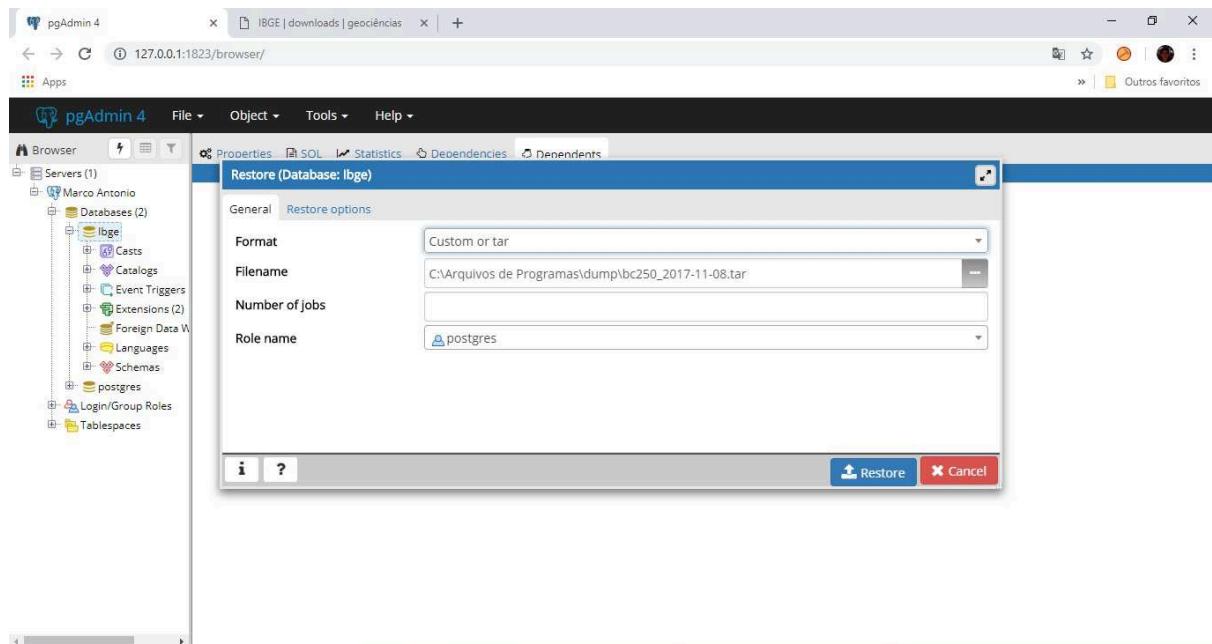


Figura 50 – Importando o banco



## ANEXO B – Instalação do QGIS

Para instalar o QGIS acesse o site <[https://www.qgis.org/pt\\_BR/site/forusers/download.html](https://www.qgis.org/pt_BR/site/forusers/download.html)>, escolha a versão do sistema operacional e baixe o arquivo, após isso execute o arquivo como administrador.

Para abrir o QGIS, dirija-se ao Menu Iniciar do Windows e digite QGIS e execute o arquivo **QGIS Desktop** como administrador.

Para poder utilizar os dados que estão no banco de dados criado no pgAdmin4 é preciso criar uma camada, vá em **Aba camada->Adicionar camada->PostGIS**, será aberto uma janela onde uma conexão precisará ser criada, clique na opção **novo** na janela **camada** e será aberta uma nova janela para você criar uma conexão entre o servidor e o QGIS.

Para exibir as informações do banco de em forma de representações espaciais é necessário ir novamente em **Aba camada->Adicionar camada->PostGIS**, clicar na opção **conectar**, digitar usuário e senha do PostgreSQL, selecionar o schema e adicionar as tabelas. Cada tabela adicionada será uma camada que se selecionada poderá ser visualizada conforme o tipo de representação espacial que está na variável espacial do banco de dados.

Figura 51 – Site QGIS

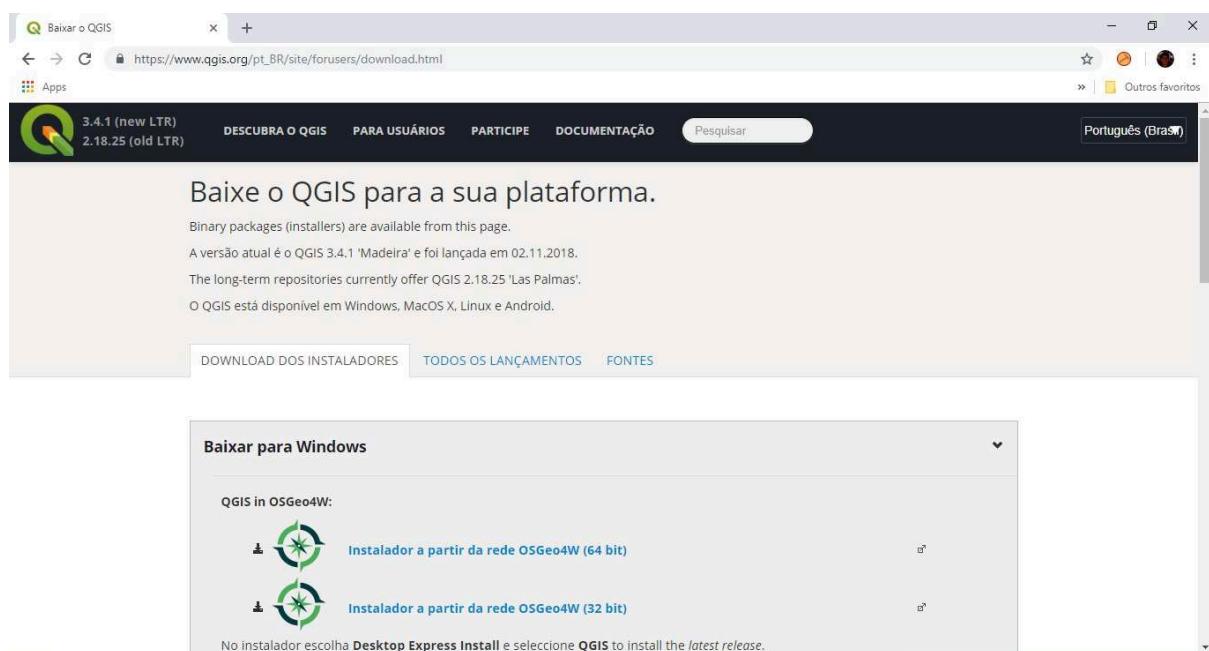


Figura 52 – Menu Iniciar - QGIS



Figura 53 – Interface QGIS

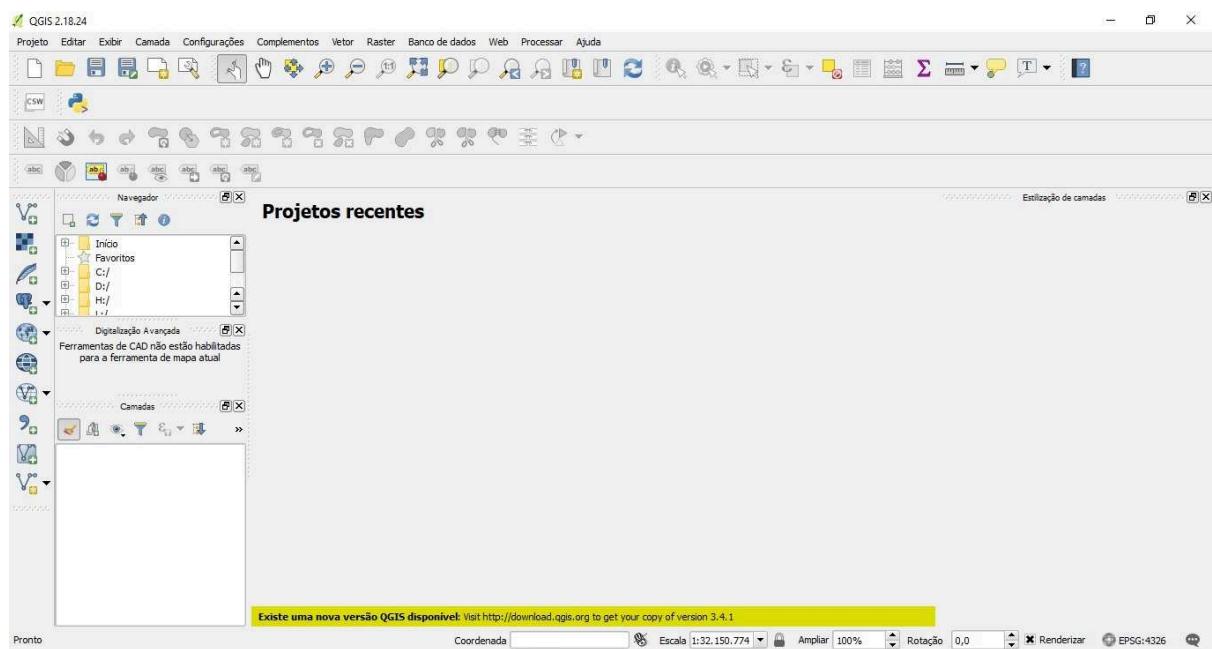


Figura 54 – Camada

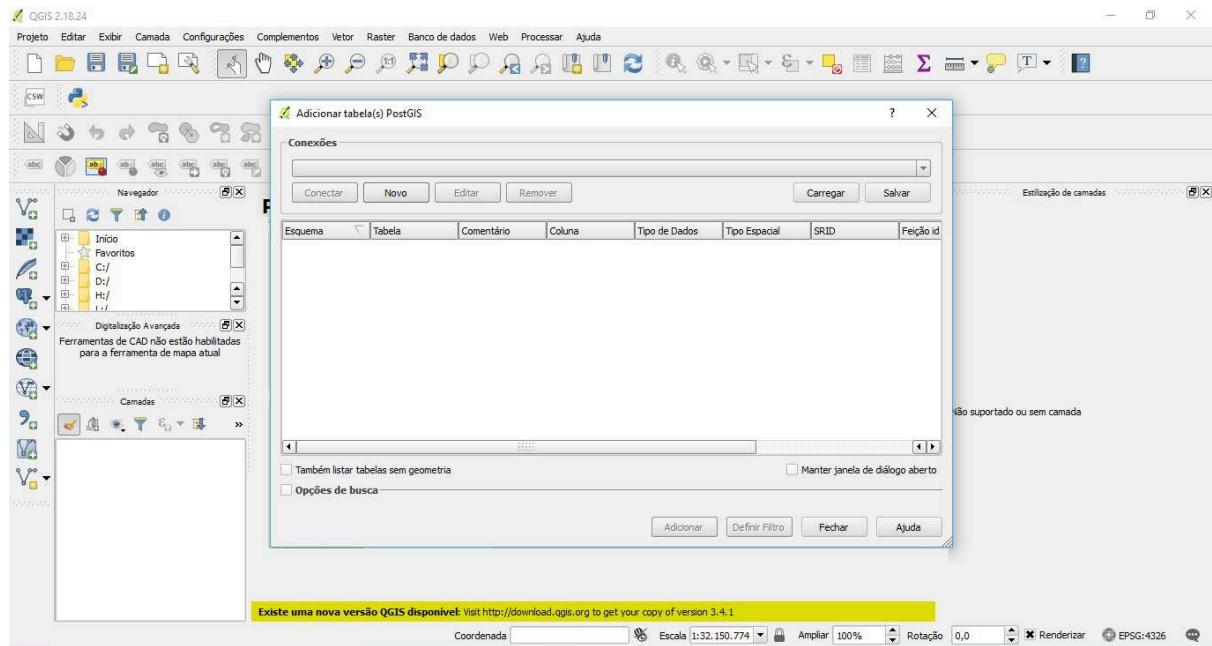


Figura 55 – Conexão

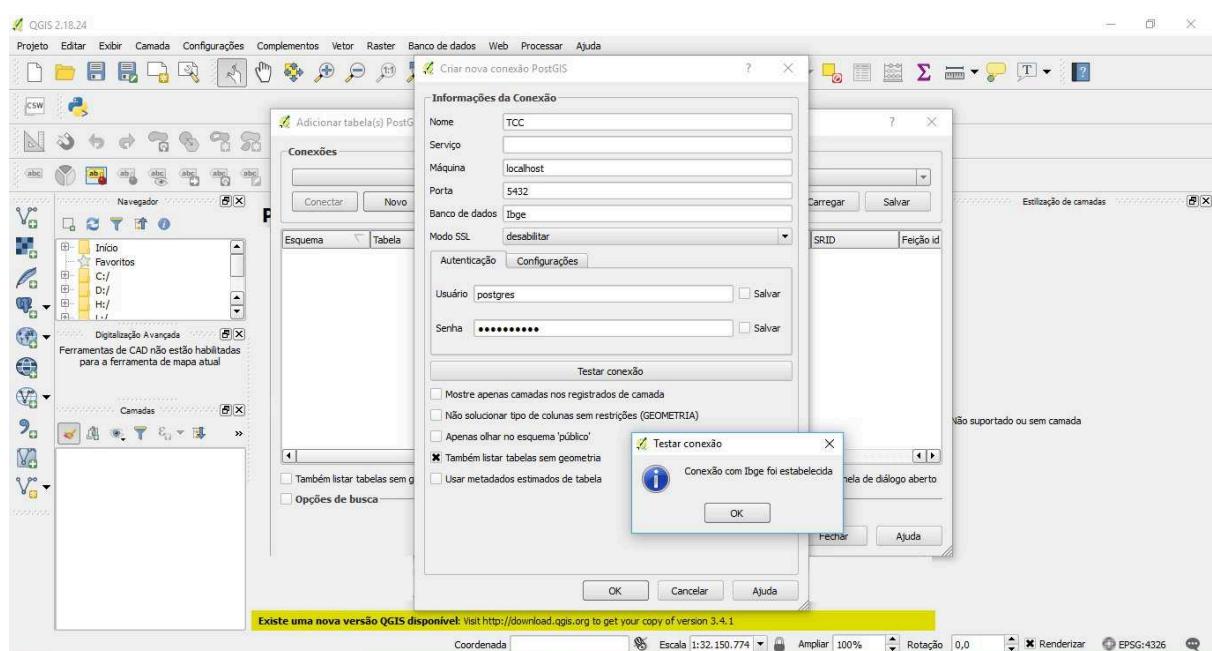


Figura 56 – Adicionar Tabelas

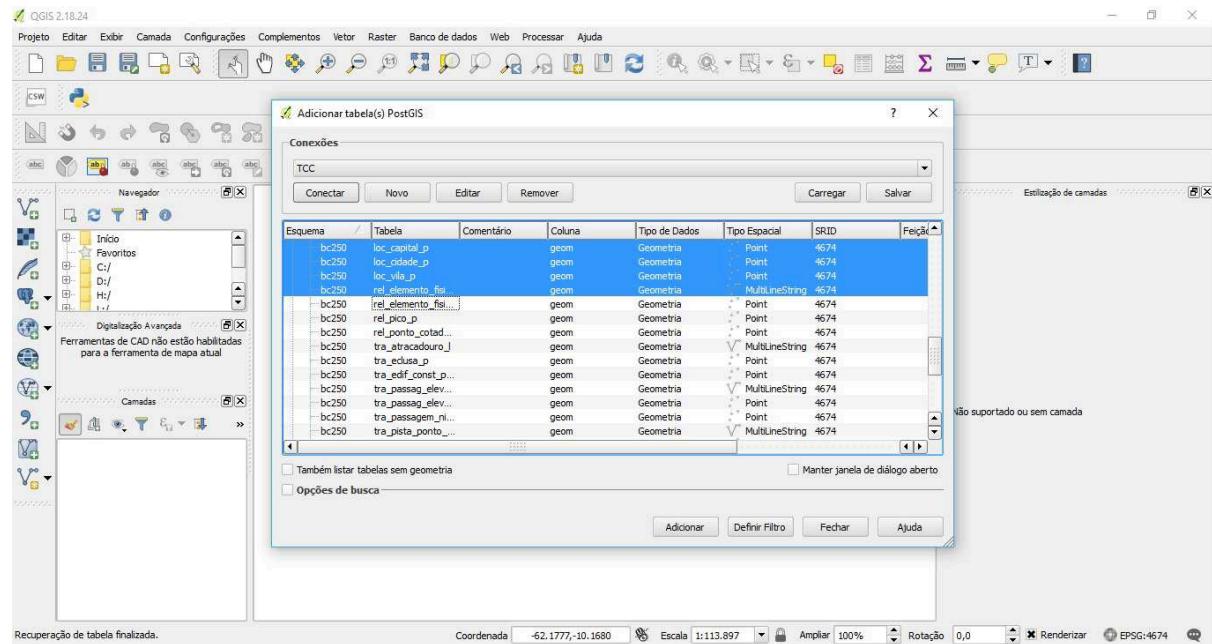


Figura 57 – Visualização das tabelas

