



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
DE TRANSPORTES E GESTÃO TERRITORIAL

Felipe Pilleggi de Souza

**OBSERVATÓRIO DO MERCADO IMOBILIÁRIO COMO INSTRUMENTO
DE APOIO AO MONITORAMENTO DOS VALORES CADASTRAIS NO
ÂMBITO MUNICIPAL**

Florianópolis

2023

Felipe Pilleggi de Souza

**Observatório do Mercado Imobiliário como instrumento de apoio ao
monitoramento dos valores cadastrais no âmbito municipal**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Engenharia de Transportes e Gestão Territorial.

Orientador: Prof. Everton da Silva, Dr.

Florianópolis
2023

Ficha de identificação da obra elaborada pelo autor através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Pilleggi de Souza, Felipe
OBSERVATÓRIO DO MERCADO IMOBILIÁRIO COMO INSTRUMENTO DE APOIO AO MONITORAMENTO DOS VALORES CADASTRAIS NO ÂMBITO MUNICIPAL / Felipe Pilleggi de Souza ; orientador, Everton da Silva, 2023.
111 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial, Florianópolis, 2023.

Inclui referências.

1. Engenharia de Transportes e Gestão Territorial. 2. Observatório do Mercado Imobiliário. 3. Cadastro. 4. Avaliação em massa de imóveis. 5. valor do solo. I. da Silva, Everton. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial. III. Título.

Felipe Pilleggi de Souza

**Observatório do Mercado Imobiliário como instrumento de apoio ao
monitoramento dos valores cadastrais no âmbito municipal**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca
examinadora composta pelos seguintes membros:

Prof. Dr. Carlos Antônio Oliveira Vieira
Universidade Federal de Santa Catarina

Prof. Dr. Diego Alfonso Erba
RACAM

Profa. Dra. Ana Maria Benciveni Franzoni
Universidade Federal de Santa Catarina

Certificamos que esta é a versão original e final do trabalho de conclusão que foi
julgado adequado para obtenção do título de mestre em em Engenharia de Transportes e
Gestão Territorial.

Profa. Dra. Ana Maria Benciveni Franzoni, Coordenadora do Programa de Pós-
Graduação em Engenharia de Transportes e Gestão Territorial

Prof. Everton da Silva, Dr.

Orientador

Florianópolis

2023

AGRADECIMENTOS

Dedico este trabalho à minha esposa Juliana Ferrari Maistro, minha companheira de tantas lutas, que vivenciou comigo, por diferentes razões, o momento mais intenso de nossas vidas. Obrigado por tanto amor e carinho.

À minha filha Bella, que trouxe a luz e energia necessária para o avanço desta pesquisa.

À minha família, em especial aos meus pais, Suzani Cassiani e Carlos Eduardo Pilleggi de Souza, ambos professores universitários recém-aposentados, que sempre me ensinaram a ter um olhar crítico diante da sociedade por meio da perspectiva científica. São minhas grandes inspirações.

Ao meu orientador, Professor Everton, por todo o suporte durante o desenvolvimento da pesquisa e por ser uma constante fonte de inspiração como pesquisador e professor.

Ao pesquisador Diego Erba, grande visionário das políticas sociais e territoriais na América Latina.

Ao Grupo de Observação e Transformação do Território (GOTT/UFSC).

RESUMO

A pesquisa apresenta uma proposta de implementação e operacionalização do observatório do mercado imobiliário, de forma a fornecer um mecanismo de apoio ao monitoramento dos valores cadastrais de municípios. A abordagem adotada inclui a análise do cadastro imobiliário, modelagem, coleta e tratamento de dados de eventos de mercado. São avaliados os resultados da coleta automatizada de dados e o georreferenciamento por geocodificação. Adicionalmente, realizam-se procedimentos para o relacionamento entre eventos do mercado imobiliário e dados cadastrais. Por fim, esta pesquisa evidencia o potencial da análise de desempenho na supervisão dos valores cadastrais, além de propor a adoção de uma ferramenta de Business Intelligence para a visualização e interação com os dados.

Palavras-chave: Observatório do mercado imobiliário 1; Avaliação em massa de imóveis 2; Valores cadastrais 3.

ABSTRACT

The research presents a proposal for the implementation and operationalization of the real estate market observatory, in order to provide a support mechanism for monitoring the cadastral values of municipalities. The adopted approach includes the analysis of real estate registration, modeling, collection, and processing of market event data. The results of automated data collection and georeferencing through geocoding are evaluated. Additionally, procedures are conducted to establish a relationship between real estate market events and cadastral data. Finally, this research highlights the potential of performance analysis in supervising cadastral values, as well as proposing the adoption of a Business Intelligence tool for data visualization and interaction.

Keywords: Real estate observatory 1; Mass appraisal of properties 2; Cadastral values 3.

LISTA DE FIGURAS

Figura 1 – Observatório do Mercado Imobiliário (OMI) da Província de Córdoba	27
Figura 2 – Número de eventos de mercado levantados em 2021 pela IDECOR.	28
Figura 3 – Arquitetura dos servidores em nuvem do programa IDECOR (2019).....	29
Figura 4 – Fluxograma de procedimentos metodológicos da pesquisa.	37
Figura 5 – Fórmulas utilizadas como base para análise de desempenho.....	47
Figura 6 – Exemplo e requisição de API através do Software Postman.	52
Figura 7 – Fluxograma do processo de geocodificação.	57
Figura 8 – Exemplo de endereços de eventos de mercado coletados.	58
Figura 9 – Tela da aplicação Google Maps, exemplificando o padrão adotado de endereçamento pela tecnologia.....	58
Figura 10 – Exemplo de campos que são retornados pela API Google Geocoding.....	60
Figura 11 – Gráfico representando a porcentagem de eventos por cada tipo de precisão da API de geocodificação.	61
Figura 12 – Gráfico representando a porcentagem de eventos por cada tipo de precisão da API de geocodificação.	62
Figura 13 – Figura que apresenta a avaliação dos campos de eventos de mercado com relação aos dados cadastrais, aplicando-se convergência de denominação e outros ajustes tipológicos.	64
Figura 14 – Modelagem conceitual da proposta de um observatório do mercado imobiliário com os diferentes relacionamentos.	66
Figura 15 – Exemplo de Outliers avaliados, a partir de scripts sqls sobre a massa de dados.....	69
Figura 16 – Exemplo da tabela do banco de dados, representando os campos que foram utilizados para controle de georreferenciamento dos eventos de mercado.	70
Figura 17 – Exemplo de pontos geocodificados com precisão “ROOFTOP” (classe mais precisa a nível de porta), mas que não cruzavam com lotes cadastrais.	72
Figura 18 – Mosaico demonstrando situação em que um evento de mercado “rooftop” está posicionado sobre o alinhamento do muro físico. Esse detalhe é evidenciado tanto em vista aérea quanto em vista terrestre.	73
Figura 19 – Painel inicial de visualização do BI criado para o observatório do mercado imobiliário para o Município de Florianópolis.....	80
Figura 20 – Exemplo do Painel de BI sendo atualizados conforme os filtros cruzados, e a adequação dos gráficos para os mesmos.....	81
Figura 21 – Recorte do BI, que apresenta o comparativo entre o valor do metro quadrado (m ²) de eventos de mercado e o valor do m ² de valores venais.	82
Figura 22 – Gráficos do BI apresentando as diferenças entre o preço de mercado dos imóveis com relação ao valor venal, para cada imóvel do respectivo Distrito Administrativo.	84
Figura 23 - Gráfico do BI que apresenta a análise de desempenho de acordo com Nível de Avaliação, Mediana e Média Ponderada, além da linha de referência aceitável do IAAO (1991).	85

Figura 24 – Gráfico do BI que apresenta a análise de desempenho de acordo Uniformidade – Coeficiente de Dispersão.	85
Figura 25 – Gráfico do BI que apresenta a análise de desempenho de acordo Uniformidade – Diferencial Relacional ao Preço.....	86

LISTA DE QUADROS

Quadro 1 – Campos selecionados da tabela “cotr_imobiliario”.....	49
Quadro 2 – Campos selecionados da tabela lotes.....	50
Quadro 3 – Campos selecionados da tabela distritos administrativos.....	51
Quadro 4 – Campos selecionados a partir da coleta sistematizada de eventos do mercado.	53
Quadro 5 – Campos definidos para a tabela de observatório.	65
Quadro 6 – Scripts SQL para filtragem de possíveis outliers utilizando a análise de desvio padrão.....	65

LISTA DE TABELAS

Tabela 1 - Tabela dos quantitativos de anúncios imobiliários por tipologia.	56
Tabela 2 - Distribuição dos anúncios por tipo de precisão do endereço.	60
Tabela 3 - Distribuição dos anúncios por tipo de precisão do endereço para terrenos...	62
Tabela 4 - Vinculação dos eventos de mercado com a parcela cadastral para cada tipologia, a partir do preenchimento das inscrições imobiliárias pela equipe técnica....	75
Tabela 5 - Resultados dos cálculos obtidos na análise de desempenho para os Distritos Administrativos com no mínimo 5 casos.	78
Tabela 6 – Avaliação dos Valores em relação aos Limites Aceitáveis (IAAO e Brasil, 2009).....	79

LISTA DE ABREVIATURAS E SIGLAS

BI – Business Intelligence

CTM - Cadastro Territorial Multifinalitário

GOTT – Grupo de Observação e Transformação do Território

IBGE - Instituto Brasileiro de Geografia e Estatística

IDECOR - Instituto de Desenvolvimento e Planejamento de Cidades de Córdoba

IAAO - International Association of Assessing Officers

IPTU – Imposto Predial e Territorial Urbano

IPIUF – Instituto de Planejamento Urbano de Florianópolis

ITBI - Imposto sobre a Transmissão de Bens Imóveis

PVG – Planta de Valores Genéricos

SIG - Sistema de Informação Geográfica

SIT - Sistema de Informação Territorial

OMI – Observatório do Mercado Imobiliário

OUV – Observatório Urbano de Valores

SUMÁRIO

1 INTRODUÇÃO E RELEVÂNCIA DO TEMA.....	15
1.1.1 Objetivo Geral.....	19
1.1.2 Objetivos específicos	19
2 REVISÃO BIBLIOGRÁFICA	20
2. 1 AVALIAÇÃO DE IMÓVEIS.....	20
2.1.1 Avaliação em Massa de Imóveis	20
2.2 OBSERVATÓRIO DO MERCADO IMOBILIÁRIO - OMI	21
2.2.1 Projetos de observatórios na América Latina: Análise e Contribuições.....	24
2.2.1.1 - O caso de Fortaleza – CE.....	24
2.2.1.2 – O caso de Córdoba – Argentina.....	25
2.2.2 Web Crawlers e a coleta automatizada de eventos do mercado	30
2.2.3 Geocodificação	31
2.2.4 Sistema de Informação Geográfica (SIG) e Sistema de Informação Territorial (SIT).....	32
2.2.5 Banco de Dados	33
2.3 ANÁLISE DE DESEMPENHO	34
3 MATERIAIS E MÉTODOS.....	36
3.1 MATERIAIS.....	38
3.2 ANÁLISE DA ESTRUTURA DO CADASTRO IMOBILIÁRIO DE FLORIANÓPOLIS-SC.....	39
3.3 MODELAGEM CONCEITUAL E CONVERGÊNCIA DE DENOMINAÇÃO	40
3.4 COLETA SISTEMATIZADA DE EVENTOS DO MERCADO	41
3.4.1 Geocodificação dos endereços de anúncios imobiliários.....	41
3.5 MODELO FÍSICO E OPERAÇÕES PARA ARMAZENAMENTO DOS DADOS	43
3.6 MECANISMOS PARA DEPURAÇÃO DOS EVENTOS DE MERCADO	44
3.8 RELACIONAMENTO DOS DADOS DE EVENTOS DO MERCADO E CADASTRO IMOBILIÁRIO.....	45
3.9 ANÁLISE DE DESEMPENHO	46
3.10 - VISUALIZADOR DOS DADOS EM BI.....	47
4 RESULTADOS E DISCUSSÕES	48
4.1 ANÁLISE DA ESTRUTURA DO CADASTRO IMOBILIÁRIO DE FLORIANÓPOLIS – SC	48
4.2 - COLETA SISTEMATIZADA DE EVENTOS DE MERCADO.....	51
4.3 - GEOCODIFICAÇÃO DE ENDEREÇOS.....	56
4.4 - CONVERGÊNCIA DE DENOMINAÇÃO (DE - PARA) E MODELAGEM CONCEITUAL.....	63

4.5 OPERAÇÕES DE ARMAZENAMENTO E TRATAMENTO DOS DADOS ...	66
4.5.1 Mecanismos para depuração dos dados de eventos de mercado.....	67
4.5.1.2 - Análise do nível de georreferenciamento dos eventos do mercado.....	70
4.6 RELACIONAMENTO COM DADOS CADASTRAIS E PROCEDIMENTOS PARA ANÁLISE DE DESEMPENHO	76
4.7- VISUALIZADOR DE DADOS EM BI	79
4.7.1 Visualização de análise de desempenho	84
5 CONSIDERAÇÕES FINAIS.....	86
6 REFERÊNCIAS	89

1 INTRODUÇÃO E RELEVÂNCIA DO TEMA

A Lei 10.257, de 10 de julho de 2001, denominada Estatuto da Cidade, estabelece normas de ordem pública e interesse social que regulam o uso da propriedade urbana em prol do bem coletivo, da segurança e do bem-estar dos cidadãos, bem como do equilíbrio ambiental. Segundo Artigo 2º, inc. IV, a política urbana tem como uma de suas diretrizes o planejamento do desenvolvimento das cidades, da distribuição espacial da população e das atividades econômicas do Município e do território sob sua área de influência, de modo a evitar e corrigir as distorções do crescimento urbano e seus efeitos negativos sobre o meio ambiente (BRASIL, 2001).

O entendimento sobre a dinâmica imobiliária de uma cidade e do valor dos imóveis é importante para a gestão municipal, sob vários aspectos, como por exemplo: fiscal, planejamento e ordenamento territorial e social. Na estruturação de um Cadastro Territorial Multifinalitário (CTM), um conjunto de informações referentes ao valor dos imóveis é um dos principais cadastros temáticos, e contribuem para a definição de políticas públicas, com ênfase em políticas tributárias (DE CESARE ET AL, 2023). No âmbito da administração pública municipal, esse nível de informação subsidia a estimativa de cálculo dos tributos imobiliários, como é o caso do imposto predial e territorial urbano (IPTU), o imposto sobre a transmissão de bens imóveis (ITBI) e a contribuição de melhoria, bem como pode dar suporte a efetivação dos instrumentos de política urbana que dialogam com o mercado imobiliário, como a Outorga Onerosa do Direito de Construir (OODC), o Direito de Preempção, entre outros.

Dada à necessidade de avaliar o universo de imóveis localizados no município, visando estimar o valor cadastral dos imóveis, diferentes métodos e análises são empregados. Tais métodos e análises visam a avaliação em massa dos imóveis, que busca definir modelos que possam ser empregados para determinação do valor cadastral para todos os imóveis administrados pelo cadastro.

Para que os modelos de avaliação em massa produzam os melhores resultados possíveis e sejam também orientados por procedimentos normativos, é necessário um conjunto de dados do mercado imobiliário que seja estruturado, normalizado e representativo, tanto em escala territorial quanto em escala temporal. Isto é, informações

bem distribuídas na extensão do município, coletadas em intervalos de data relativos.

Com o avanço das tecnologias e geotecnologias nas últimas décadas, diferentes sistemas de informação surgem como soluções de apoio ao planejamento dos municípios da América Latina. Aquele que é recomendado para a finalidade da avaliação em massa de imóveis e monitoramento do mercado imobiliário é o Observatório do Mercado Imobiliário (OMI).

A Portaria 3242/2022 do Ministério do Desenvolvimento Regional trata do Observatório do Mercado imobiliário em um de seus artigos (Art. 20) e o define como um sistema de informação aberto, destinado à coleta e armazenamento contínuo de dados de mercado, incluindo: preços de transações imobiliárias, ofertas, aluguéis, custos de construção, avaliações prévias e indicadores relacionados ao setor imobiliário. Pode-se afirmar que pôs luz a um tema relevante, pois sua utilidade como ferramenta para análise do comportamento do mercado imobiliário é ampla, incipiente e para algumas áreas do conhecimento, ainda inexploradas.

No âmbito cadastral, principalmente nas jurisdições da América Latina, a utilização dos OMIs é crescente e bem estabelecida como instrumento de apoio a estudos voltados a políticas fiscais. Silva (2020) discorre que os OMIs são considerados por diferentes autores como sistemas destinados à coleta sistemática de dados, principalmente de preços dos imóveis os quais são registrados em uma base de dados e relacionados com o cadastro. Desta forma, possibilita-se a formação de conjunto de dados com os quais se realizam processo de análises exploratórias e modelagem do mercado imobiliário nos diferentes tipos de imóveis registrados na base cadastral.

A importância dos Observatórios e da avaliação em massa de imóveis no apoio a políticas públicas e planejamento das cidades vão além da justiça fiscal. De acordo com De Cesare (2005), iniquidades nas avaliações realizadas para tributação não apenas contrariam a justiça tributária. O Artigo 150 da Constituição Federal veda instituir tratamento desigual a contribuintes que se encontrem em situação equivalente. Outro ponto levantado pela autora é que o baixo desempenho nas avaliações realizadas para tributação está muitas vezes relacionado ao limitado acesso às informações de mercado, ao incorreto emprego de técnicas de estimação de valor, e a omissão de atributos fundamentais na estimação da base de cálculo.

Diante desse quadro de necessidades e da importância dos observatórios como ferramenta de monitoramento do mercado imobiliário, os municípios brasileiros tem a frente o importante desafio de implementar o OMI na administração pública municipal, a partir de um modelo de dados baseado na parcela cadastral e que permita a integração e interoperabilidade com outros sistemas.

Diante do exposto, coloca-se a seguinte questão norteadora para esta pesquisa: a estruturação de um Observatório do Mercado Imobiliário, destinado ao monitoramento do desempenho dos modelos de avaliação em massa de imóveis pode contribuir para as políticas públicas e à redução da iniquidade fiscal? Sabe-se, entretanto, que a estruturação de um observatório passa pelo desenho e implementação de uma base de dados e de procedimentos apoiados por soluções de software para sua operacionalização. Nesta lacuna é que a pesquisa proposta se ancora, propondo uma ferramenta de apoio aos municípios brasileiros.

Como resultado principal, proporcionado pelo cumprimento dos objetivos da pesquisa, apresenta-se um método de implementação e operacionalização de um OMI baseado em soluções de software e mecanismos de validação e qualificação dos dados. Resultado desenvolvido que teve como área de estudo o município de Florianópolis, SC, e que contou com apoio da Administração Municipal, a partir do Instituto de Planejamento Urbano de Florianópolis (IPUF) e do Grupo de Observação e Transformação do Território (GOTT) da Universidade Federal de Santa Catarina.

Destaca-se que ao longo da pesquisa pôde-se efetivamente dar início a estruturação do OMI no município de Florianópolis, que vem sendo mantido e alimentado com dados pelos técnicos e estagiários que acompanharam o desenvolvimento dos trabalhos e que contribuíram para o estabelecimento e aprimoramento dos procedimentos de coleta e validação dos dados levantados junto ao mercado imobiliário.

A estruturação da pesquisa iniciou-se com a "Introdução e Relevância do Tema", onde é apresentado o contexto e a importância da pesquisa. Os "Objetivos Geral e Específicos" delineiam os propósitos da investigação. Em seguida, é realizada a "Revisão Bibliográfica", que introduz os conceitos fundamentais, como a "Avaliação de Imóveis" e a "Avaliação em Massa", contextualizando a pesquisa no campo da avaliação imobiliária. A discussão sobre o "Observatório do Mercado Imobiliário (OMI)" e sua aplicação na América Latina ressalta a importância do monitoramento contínuo e da

coleta automatizada de dados. Além disso, tópicos como "Web Crawlers", "Geocodificação", "Sistemas de Informação Geográfica" e "Banco de Dados" são explorados para embasar a metodologia. A seção "Materiais e Método" detalha a abordagem adotada, incluindo a análise do cadastro imobiliário, modelagem, coleta e tratamento e visualização dos dados. A seção "Resultados e Discussões" apresenta os principais resultados da pesquisa, incluindo a análise do cadastro imobiliário, a coleta automatizada de eventos do mercado e os resultados da geocodificação de endereços de anúncios. Também é realizada a convergência de denominação, modelagem conceitual, juntamente com detalhes das operações de armazenamento e tratamento de dados, destacando a análise de georreferenciamento. Em seguida, investiga o relacionamento entre os dados de eventos e cadastrais, aplicando técnicas de avaliação de desempenho entre os mesmos. Por fim, elabora-se uma visualização dos dados e resultados a partir de uma ferramenta de Business Intelligence (BI).

1.1 OBJETIVO

1.1.1 Objetivo Geral

Propor um método para implementação e operacionalização de um observatório do mercado imobiliário, baseado em soluções de software livre e mecanismos de validação e qualificação dos dados.

1.1.2 Objetivos específicos

- a) Contribuir para um processo de monitoramento dos valores cadastrais quanto a sua relação com o comportamento dos preços do mercado imobiliário;
- b) Analisar e propor procedimentos baseados em software livre que possibilitem a coleta e validação sistemática de dados do mercado imobiliário;
- c) Oferecer um mecanismo de apoio à implementação ou a qualificação das políticas urbanas que se relacionam com o mercado imobiliário.

2 REVISÃO BIBLIOGRÁFICA

2.1 AVALIAÇÃO DE IMÓVEIS

A avaliação de imóveis é um processo técnico que tem por objetivo estimar o valor dos bens para uma determinada finalidade. Tem papel importante nas políticas públicas, sendo fundamental para a determinação dos cálculos de tributos municipais.

Erros e distorções na determinação da base de cálculo afetam diretamente a distribuição do valor do imposto. Portanto, devido unicamente à baixa qualidade das avaliações, ou seja, imperfeições de natureza essencialmente administrativa, a carga tributária pode ser indevidamente transferida entre diferentes grupos de renda (SABELLA, 1975).

Para Silva (2020), os estudos de avaliação de imóveis visam estimar o valor de mercado dos imóveis, sendo considerado pela norma brasileira como o valor mais provável de se negociar voluntariamente e conscientemente um ativo, dentro de um intervalo de data específico e dentro das condições atuais do mercado.

As avaliações de imóveis podem ser divididas em dois modelos: avaliação individual e avaliação em massa. A primeira é utilizada para diferentes propósitos no âmbito imobiliário, como levantamento de balanço patrimonial, financiamento de imóvel por bancos, garantias e no próprio processo de compra e venda. Já o segundo é utilizado principalmente para a tributação imobiliária (SILVA, 2020).

2.1.1 Avaliação em Massa de Imóveis

IAAO (1990) define a avaliação em massa de imóveis como o processo de estimar o valor de um grande número de unidades, dentro de um mesmo período de tempo, por meio de métodos padronizados e análises estatísticas. Para De Cesare (2003), esse tipo de avaliação consiste no desenvolvimento de um ou mais modelos genéricos, originários de uma amostra de dados, e os quais são aplicados às diferentes categorias de bens a serem avaliados.

Silva (2020) discorre que no âmbito da avaliação em massa de imóveis, aliado à utilização do observatório do mercado imobiliário, destaca-se o método comparativo de

dados do mercado. Este método estima o valor do bem pela comparação com dados do mercado.

É fundamental que haja um conjunto de dados que possa ser utilizado como amostra representativa do mercado, ou seja, que tenha uma distribuição equilibrada em seu domínio de modo que o conjunto de dados seja o mais representativo possível do universo que se pretende avaliar (SILVA, 2020).

Com relação a sua importância para a gestão territorial, e mais especificamente para as administrações municipais, Silva e Loch (2004) afirmam que:

A sistematização da avaliação em massa de imóveis vem preencher uma lacuna no desenvolvimento científico e tecnológico do cadastro técnico no que diz respeito à sustentabilidade fiscal dos municípios, crucial para equacionamento das contas públicas. Os três aspectos básicos (tripé) que fundamentam o cadastro técnico: físico, jurídico e econômico, tem neste último uma certa fragilidade quando o assunto é valor da propriedade, haja vista que se trata da base de cálculo de dois impostos de grande importância na arrecadação municipal (IPTU e ITBI) e que gera muito debate e inquietude na sociedade quando da proposição de alterações na carga tributária. (SILVA e LOCH, 2004, p.2).

Os observatórios, apoiando no processo de sistematização da avaliação em massa de imóveis, suprem uma lacuna no desenvolvimento científico e tecnológico do cadastro técnico, essencial para a sustentabilidade fiscal municipal, abordando o valor da propriedade, base de cálculo de impostos cruciais como IPTU e ITBI.

2.2 OBSERVATÓRIO DO MERCADO IMOBILIÁRIO - OMI

Para Eguino e Erba (2020), apesar do momento atual ser positivo com relação a interoperabilidade e georreferenciamento das informações no âmbito público e privado de uma forma geral, o mesmo ainda não ocorre no cadastro econômico. Existe a falta de padronização e de referência única em grande parte desses cadastros. Um dos motivos

destas inconsistências é que o valor dos imóveis é geralmente apurado no âmbito administrativo, caracterizados por processos não claros, subjetividade e manipulação de dados. O problema se agrava, pois, a determinação dos valores é aprovada nas instâncias de decisão política, e utilizada como base para definição dos diferentes tributos municipais e territoriais. Essa imprecisão gera iniquidade fiscal e afeta diretamente os contribuintes, o que desestimula o pagamento de impostos e contribuições no âmbito imobiliário.

Uma forma de resolver essa situação é estabelecendo políticas cadastrais que remetem todos os valores dos imóveis em um sistema único e inquestionável. Neste modelo, a referência dos valores deve ser a do próprio mercado imobiliário. Os anúncios de valor dos imóveis, sejam públicos ou de portais imobiliários, assumem o protagonismo contendo informações reduzidas e simplificadas, e que condicionam a demanda. Além disso, os valores estipulados nas transações imobiliárias já agregam o valor de intervenções públicas e privadas (EGUINO, ERBA, 2020).

Segundo De Cesare et al. (2023), é fundamental para a atividade de avaliação de imóveis o acesso a um conjunto de dados confiáveis sobre os preços praticados, que devem ser coletados de forma contínua e abranger todos os tipos de imóveis. Recomenda-se, no âmbito da administração municipal, a criação do Observatório do Mercado Imobiliário (OMI), um sistema de informação destinado à coleta e armazenamento contínuo de dados de mercado, que deve incluir preços de venda, aluguéis, custos de construção, avaliações prévias e indicadores relacionados ao setor imobiliário. Além disso, a formação da base de dados do OMI deve ser compartilhada com diferentes atores, incluindo parceiros potenciais como a Caixa Econômica Federal, a Secretaria do Patrimônio da União, os cartórios de registros de imóveis, entre outros (DE CESARE ET AL, 2023).

Seu desenvolvimento como conceito é incipiente e em constante transformação. A etimologia da palavra “observatório” é: de observar. Nos dicionários, seu principal significado ainda é o de estabelecimentos para observações astronômicas e meteorológicas. Porém, em uma nova perspectiva sendo amplamente utilizada pelas organizações públicas e privadas, tratam-se de instituições que se dedicam a observação, acompanhamento e divulgação de fenômenos ou informações. Nery (2014) discorre que o observatório é um sistema de informações criado por instituições públicas e privadas, para acompanhar a evolução de um fenômeno, de um domínio ou de um tema estratégico, no tempo e espaço. Podem ser utilizados para diferentes propósitos, podendo ter caráter

territorial, ambiental, social, legal ou de valores, podendo ser utilizados tanto para políticas públicas quanto para tributárias.

Eguino e Erba (2020) discorrem sobre o conceito mais amplo de observatório e as diferenças entre os observatórios territoriais e os observatórios do mercado imobiliário (OMI).

Um observatório é uma estrutura administrativa e tecnológica que monitora a cidade por meio de imagens e recenseamentos. Pode ser estruturado pela instituição que administra o cadastro territorial, por instituições privadas ou acadêmicas ou por alianças interinstitucionais que reúnam várias organizações com interesse comum em determinados espaços urbanos e em temáticas urbanas específicas. Enquanto os observatórios territoriais são criados com a finalidade de coletar e/ou publicar dados úteis para a definição de políticas públicas em geral, os OMIs realizam pesquisas com os dados para gerar mapas de valores, necessários para definir a política tributária sobre o imposto predial, a contribuição de melhorias e a recuperação de mais-valias (EGUINO, ERBA, 2020, p.26).

De acordo com De Cesare et al. (2023), no Capítulo 5 do livro "Cadastro territorial multifinalitário aplicado à gestão municipal", afirma-se que o OMI coleta não apenas preços relativos às transações imobiliárias, mas também dados sobre aluguéis, custos de construção, avaliações prévias e indicadores relacionados ao setor imobiliário. Esses dados são importantes para a realização de avaliações precisas e confiáveis dos valores dos imóveis. O sistema é idealmente aberto para permitir fácil acesso aos dados. Isso significa que os dados coletados pelo OMI devem estar disponíveis para consulta pública e serem acessíveis por diferentes atores interessados no mercado imobiliário.

A formação da base de dados do OMI deve ser compartilhada com diferentes atores, e sugere-se que a coordenação seja desempenhada pelo Poder Público municipal. Isso significa que diferentes instituições devem contribuir para a formação da base de dados do OMI, incluindo órgãos governamentais, empresas privadas e outras organizações. Parceiros potenciais incluem a Caixa Econômica Federal e outros agentes que financiam bens imóveis, a Secretaria do Patrimônio da União, as áreas de planejamento urbano e habitação do município, associações de corretores de imóveis,

entre outros. A colaboração desses parceiros é fundamental para garantir a qualidade e a abrangência dos dados coletados pelo OMI (DE CESARE ET AL, p.110).

Outro ponto importante mencionado pelos autores é com relação a importância do método comparativo direto dos dados de mercado na avaliação de imóveis, mas também dos desafios sobre a qualidade da informação. Embora seja ideal que a amostra de dados seja composta apenas por imóveis transacionados no período, há desafios relacionados à subdeclaração de preços para reduzir a carga tributária ou à declaração de preços superiores por razões fiscais relacionadas à lavagem de dinheiro. Portanto, é comum e desejável utilizar dados de ofertas, embora esses não representem evidências concretas do mercado imobiliário, uma vez que as transações não foram efetivadas. Nos modelos avaliatórios, é importante distinguir os preços de oferta dos preços de transação intermediados por variáveis que possam explicar esse comportamento.

Ainda, De Cesare et al. (2023, p.112) apontam que o uso de tecnologias e a disseminação de dados na internet ampliaram exponencialmente as informações disponíveis, resultando em amostras mais representativas para análises do mercado imobiliário. Algumas fontes fornecem dados georreferenciados e detalhes precisos dos imóveis, enquanto em outros casos é necessário identificar esses imóveis no cadastro territorial. A coleta de dados sobre a oferta de imóveis pode ser feita de forma individualizada, inclusive por meio de simulação de interesse na negociação. No entanto, é importante validar e associar os dados coletados automaticamente antes de integrá-los à base de dados do observatório. Devido à assimetria das informações no mercado imobiliário, é fundamental cruzar dados de diferentes fontes para complementação e verificação. Estabelecer parcerias ou relações de confiança com agências imobiliárias pode ajudar a reduzir o esforço na coleta de dados e confirmação dos preços de venda, levando em consideração a falta de transparência do mercado. (p. 112).

2.2.1 Projetos de observatórios na América Latina: Análise e Contribuições

2.2.1.1 - O caso de Fortaleza – CE

Fortaleza é um dos principais exemplos de cidades brasileiras que tiveram êxito na implementação de um observatório para apoio de políticas fiscais no município. Segundo o Instituto Brasileiro de Geografia e Estatística – IBGE (2021), a população estimada de 2021 do município é de 2,7 milhões, sendo um dos principais centros urbanos

do nordeste brasileiro. De acordo com Silva e Oliveira (2020), o município possui um cadastro com aproximadamente 780.000 imóveis, com uma grande variedade de edificações, e um mercado imobiliário altamente dinâmico, principalmente pela localização privilegiada na costa atlântica. Em 2009, Fortaleza iniciou seu processo de modernização do cadastro, migrando de um modelo cadastral ortodoxo, para um modelo de cadastro multifinalitário, através da implementação de um Sistema de Informações Territoriais (SIT). Este modelo de sistema abriu oportunidade para que outros setores da administração interagissem com os diferentes dados cadastrais e geográficos do município. O cadastro, que até então tinha uma função acessória, assumiu uma posição central na gestão territorial. A modelagem da base de dados passou a considerar o lote (ou a parcela) como o elemento central da gestão do cadastro (SILVA; OLIVEIRA,2020).

Ainda de acordo com os autores, com o avanço do CTM no município e fortalecimento do aspecto econômico, através da ampliação da arrecadação e da equidade, foi possível avançar na reestruturação do Observatório de Valores (OUV), que havia sido implementado inicialmente em 2005.

O OUV foi então estruturado como um módulo do SIT, recebendo melhorias estruturais, do corpo técnico, e tecnológicas, com o avanço das técnicas de raspagem de dados da internet, e posteriormente com as novas técnicas de *machine learning* (aprendizado de máquina), que permitiram gerar melhores estimativas na avaliação em massa de imóveis. A integração do OUV com o SIT permitiu a elaboração de mapas de valores imobiliários, aperfeiçoando os processos de avaliação e a cobrança do imposto predial. Outro importante avanço foi que o observatório permitiu a automatização do processo de cobrança do ITBI, importante fonte de renda do município (SILVA; OLIVEIRA, op. cit., pg. 38).

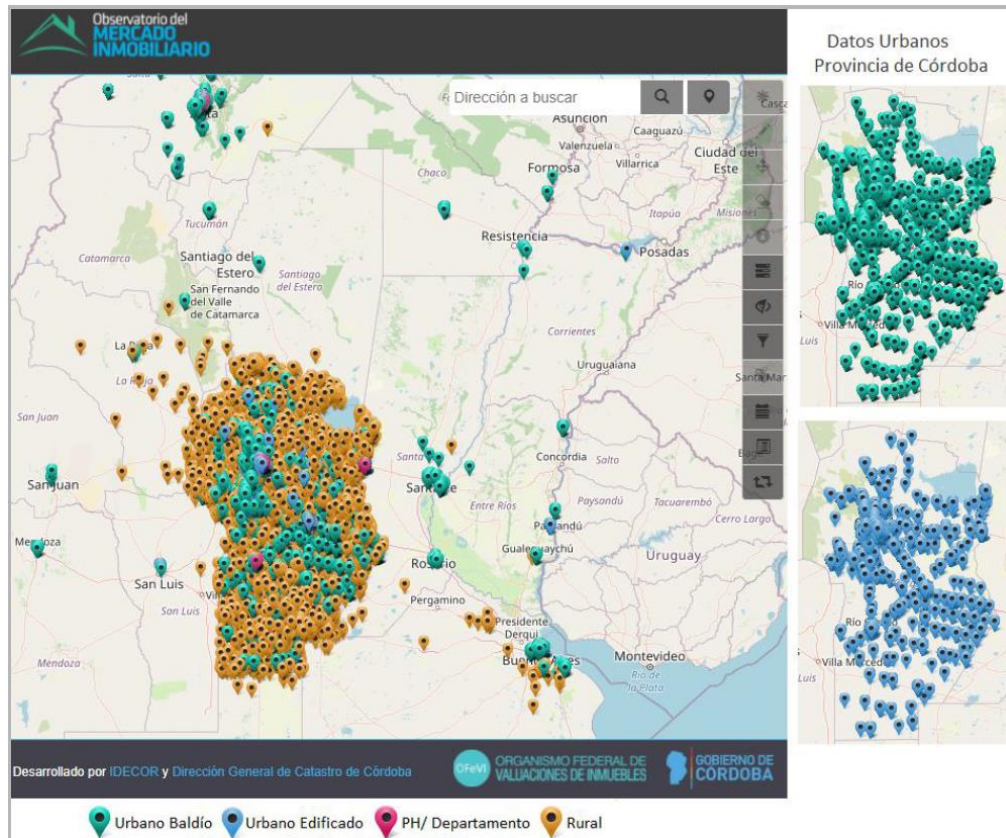
2.2.1.2 – O caso de Córdoba – Argentina

Na província de Córdoba – Argentina, o programa “Infraestructura de Datos Espaciales de la Provincia de Córdoba” (IDECOR) promove o desenvolvimento da Infraestrutura de Dados Espaciais (IDE) do território, estabelecendo uma integração entre os setores público, acadêmico e privado. Seu propósito consiste em fornecer uma ferramenta significativa para aprimorar a gestão das políticas públicas relacionadas ao território, bem como para estimular o desenvolvimento econômico e social (IDECOR, 2023).

Uma das finalidades do programa é determinar corretamente a valoração de imóveis e atualizar as técnicas e metodologias de valoração que são aplicadas pelo *Catastro Provincial*. Esses trabalhos são realizados no âmbito de um programa integral de reforma e inovação à valoração massiva de imóveis e fortalecimento do imposto imobiliário, impulsionado pelo Ministério de Finanças da Província de Córdoba. A equipe de trabalho é de caráter interinstitucional e multidisciplinar, composta por mais de 25 pessoas do IDECOR, com diferentes perfis (economistas, geógrafos, engenheiros, agrônomos, biólogos, agrimensores, arquitetos, profissionais de sistemas), e pesquisadores da Faculdade de Ciências Exatas, Físicas e Naturais (FCEfYN), da Universidade Nacional de Córdoba (UNC) (2021, p. 05).

Em 2021, através desta publicação “Estudo do mercado de solo urbano” (IDECOR, 2021), é descrito que a coleta de dados da instituição é realizada através do Observatório do Mercado Imobiliário (OMI) da Província de Córdoba, uma aplicação web desenvolvida pelo IDECOR que permite o registro e a sistematização de dados de mercado em uma base georreferenciada, com o objetivo de analisar a evolução e a dinâmica dos preços de imóveis urbanos e rurais. Os dados são categorizados de acordo com o tipo de imóvel, como terrenos vazios, construídos, unidades condominiais (PH) ou rurais (p.13). A Figura 1 exemplifica a plataforma do observatório, e pode-se notar os marcadores diferenciados de acordo com a tipologia do imóvel.

Figura 1 – Observatório do Mercado Imobiliário (OMI) da Província de Córdoba



Fonte: Imagem em IDECOR (2021, p.14).

Uma das tarefas do estudo foi o levantamento de mercado, que permitiu capturar e registrar no Observatório provincial (OMI) mais de 9.000 dados urbanos. Os cálculos dos valores unitários da terra para todas as propriedades urbanas (cerca de 1,7 milhão de propriedades) puderam ser concluídos de forma abrangente e a qualidade dos resultados melhorou em relação aos anos anteriores.

Na Figura 2, pode-se observar a quantidade e o tipo de dados coletados para os estudos urbanos, de acordo com o ano das pesquisas.

Figura 2 – Número de eventos de mercado levantados em 2021 pela IDECOR.

Tabla 1: Cantidad de datos urbanos según el año de relevamiento y tipo de inmueble

Cantidad de Datos	Año de relevamiento					Total general
	2017	2018	2019	2020	2021	
Tipo de Inmueble	2017	2018	2019	2020	2021	
Baldíos	1.369	7.195	6.096	8.147	5.256	28.063
Edificados	718	2.700	2.356	3.059	4.043	12.876
Total general	2.087	9.895	8.452	11.206	9.299	40.939

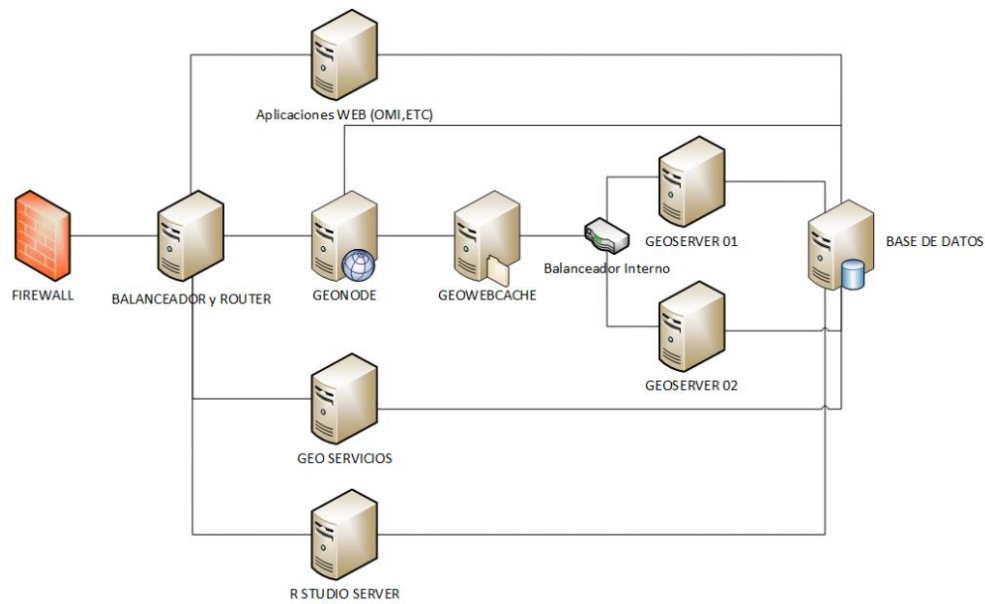
Fuente: Estudio Territorial Inmobiliario, 2021

Fonte: Imagem em IDECOR (2021, p.13).

O OMI de Córdoba é alimentado com informações fornecidas por diferentes fontes, como imobiliárias, corretores e proprietários. Para a coleta de dados em 2021, várias estratégias foram implementadas e posteriormente complementadas com outros recursos. Uma delas foi a colaboração de mais de 15 equipes municipais, compostas por funcionários e profissionais técnicos, que participaram do levantamento e continuaram contribuindo com dados para o observatório. A contribuição dos municípios representou, em média, cerca de 50% do total de dados coletados em 2021 nessas localidades. Esse grupo de usuários de 15 cidades contribuiu com um total de 3.400 dados para o Observatório (p.15).

Como um dos objetivos desta pesquisa é propor e operacionalizar um OMI, buscou-se compreender a infraestrutura do IDECOR. Na página oficial do programa, é apresentada a arquitetura de nuvem do programa, conforme Figura 3. De acordo com IDECOR (2019), o nó na nuvem utiliza o serviço "Open Cloud" da Telefónica, cuja infraestrutura é baseada em tecnologia fornecida pela "Huawei".

Figura 3 - Arquitetura dos servidores em nuvem do programa IDECOR (2019).



Fonte: IDECOR, 2019.

Conforme a publicação, inicialmente um servidor central coordena todos os acessos à infraestrutura, realizando tarefas de balanceamento e roteamento, e é protegido por um firewall para garantir a segurança dos dados.

Os serviços geoespaciais WMS, que respondem às consultas feitas ao portal GeoServer, são fornecidos por dois servidores conectados entre si. Esses servidores estão equipados com um balanceador de carga interno, que distribui as tarefas de forma eficiente para garantir o desempenho e a disponibilidade dos serviços. Para o armazenamento temporário de tiles, existe um servidor dedicado equipado com o GeoWebCache. Esse servidor é responsável por armazenar e disponibilizar as imagens de mapa em formato de quadros (tiles), permitindo um acesso rápido e eficiente aos dados geográficos. A fim de não impactar o desempenho dos usuários em geral, um terceiro servidor com GeoServer é disponibilizado especificamente para usuários e aplicativos externos que consomem geoserviços diretamente ou realizam downloads independentes do Visor de Mapas.

As aplicações web desenvolvidas pela IDECOR, como o portal Mapas Córdoba e o Observatório do Mercado Imobiliário (OMI), possuem um servidor independente para garantir um ambiente dedicado e otimizado para essas aplicações.

Quanto ao banco de dados geográfico, todas as aplicações acessam um único servidor unificado. Esse servidor utiliza o motor PostgreSQL + PostGIS, que oferece eficiência no gerenciamento de informações geográficas.

Além disso, a IDECOR possui um servidor dedicado para *data science* onde são executadas análises estatísticas avançadas e inteligência artificial. Esse servidor é equipado com o RStudio, uma plataforma popular para análise de dados e modelagem estatística (IDECOR, 2019).

2.2.2 Web Crawlers e a coleta automatizada de eventos do mercado

Web crawler – ou rastreador web - é um recurso para realizar buscas em páginas da internet, coletando e sistematizando seu conteúdo, sejam textos, imagens, vídeos, entre outros. Um bom exemplo de sua utilização são as ferramentas de busca, como o Google, que realiza a indexação de páginas para que estas apareçam nas próximas pesquisas realizadas pelos usuários e mantendo as mesmas atualizadas na base de dados (AHUJA; BAL; VARNICA, 2014).

A coleta automatizada de eventos do mercado em portais imobiliários pode ser implementada através da criação de scripts *web crawlers* capazes de realizar a extração de informações de páginas web de forma metódica e automatizada. Dependendo do propósito, é possível escrever um script em linguagem de programação, seja através de frameworks específicos de mapeamento das variáveis de HTML do *frontend*, como é o caso do *Scrapy* da linguagem *Python*, ou através do acesso e requisição da API REST dos portais imobiliários. Uma alternativa a elaboração de scripts de raspagem são os softwares desktop de interação visual, elaborados especificamente para facilitar e otimizar o processo de extração de dados, como é o caso do Octoparse e Parsehub.

A coleta automatizada de portais imobiliários, principalmente via web crawlers, gera um debate importante sobre sua legitimidade. Em recente publicação científica, intitulada de “Coleta de dados de imóveis de forma automatizada para fins de políticas públicas”, Silva, C. et al. (2020) apresentam a importância da coleta automatizada de eventos do mercado pelo Poder Público, e uma reflexão sobre sua legitimidade frente a instrumentos legais, como a LGPD (Lei Geral de Proteção de Dados Pessoais). Segundo os autores, A existência de legitimidade para o uso de web crawlers pela administração pública é importante e necessária, uma vez que se tem interesse em coletar apenas informações despersonalizadas dos imóveis, ou seja, informações que não permitam a associação entre o sujeito e os dados. Essa prática é respaldada pelo artigo 5º, incisos I, X e XI da LGPD. Além disso, é relevante destacar que o respeito ao Princípio da

Anonimidade, combinado com os parágrafos 3º e 4º do artigo 7º da LGPD, deve ser observado durante o tratamento de dados pessoais pela administração pública. Nesse sentido, a administração pública pode realizar o tratamento e uso compartilhado de dados necessários à execução de políticas públicas previstas em leis e regulamentos ou respaldadas em contratos, convênios ou instrumentos congêneres, conforme o inciso III, ou ainda para a realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais, conforme o inciso IV.

Vale ressaltar que o tratamento de dados pessoais cujo acesso é público deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização, conforme o parágrafo 3º, e que não é necessária a exigência de consentimento para os dados tornados manifestamente públicos pelo titular, resguardados os direitos do titular e os princípios previstos na LGPD, conforme o parágrafo 4º. Em relação aos agentes imobiliários, pode-se argumentar que a administração pública pode ser beneficiada pelo uso dos dados coletados com base nas políticas de solo (Silva, C. et al., 2020).

2.2.3 Geocodificação

Na área cadastral, é o georreferenciamento, conforme aponta Eguino e Erba (2020), que permite a estruturação da base de dados com localização, forma e dimensão de cada imóvel. Além disso, a adoção de um sistema único de referência garante a continuidade gráfica às informações territoriais e possibilita a interoperabilidade, ao integrar documentos cartográficos gerados por diferentes atores.

Literalmente, a palavra geocodificação, quando segmentada em seus componentes, tem como resultado o significado de “codificar as informações de referência geográfica da terra”.

Nas últimas décadas diferentes formas de georreferenciamento de dados foram desenvolvidas, dependendo da necessidade ou disponibilidade. Uma forma usual de georreferenciamento é a utilização dos sistemas de navegação por satélite, como o caso do GPS, determinando as coordenadas geográficas do local de interesse. Outro método de georreferenciamento é a geocodificação, que pode ser definida como processo de relacionamento de dados tabulares, que contêm informação de endereços textuais, com coordenadas do mundo real. (Skaba, 2009, pg. 57).

Para Skaba (2009), o conceito de geocodificação foi se transformando no tempo de acordo com as mudanças de disponibilidade de dados, mais precisamente geográficos. A crescente disponibilidade, confiabilidade e precisão (ou precisão) das séries de dados geográficos contribuíram com que o processo de geocodificação evoluísse continuamente, acompanhando o ritmo de desenvolvimento de produção destas bases de dados.

De acordo com Harries (1998), a história da geocodificação está ligada aos esforços do Departamento de Censo Americano (US Census Bureau) para encontrar maneiras de mapear os dados coletados nos Estados Unidos, endereço por endereço. No Censo de População e Habitação de 1960, os questionários foram enviados aos entrevistados e retirados de cada domicílio pelos entrevistadores. Em 1970, o plano foi utilizar o correio tanto para enviar quanto para retornar pesquisas. Isso exigiu capacidade de geocodificação e, posteriormente, o desenvolvimento de um guia de codificação de endereço (ACG). Inicialmente os esforços de geocodificação eram voltados apenas para endereços de ruas, mas a necessidade em estratificar por quadras e setores censitários logo foi adicionada. Para isso, as intersecções tiveram de ser numeradas e os intervalos de endereços reconciliados com as faces de quadras relativas.

2.2.4 Sistema de Informação Geográfica (SIG) e Sistema de Informação Territorial (SIT)

O Sistema de Informação Geográfica (SIG) é uma ferramenta importante na gestão municipal, e facilita a sistematização e o manejo dos dados e informações territoriais de forma eficiente, pois envolve o uso de diferentes componentes como hardware, software, redes de comunicação, recursos humanos, metodologia e processos (ERBA, 2013), capazes de gerenciar grandes volumes de dados georreferenciados.

Quando o SIG tem o objetivo de manipular os dados territoriais estes são denominados de Sistemas de Informação Territoriais (SIT). Dessa forma, o SIT é um sistema de informação destinado a administrar os dados territoriais, apoiando de forma eficiente a geração e produção de dados e informações pelas as entidades responsáveis (ERBA, 2007).

Para Mondal et al. (2016), O SIT é um sistema de gestão do território, composto por uma base de dados e informações alfanuméricas referenciadas espacialmente à parcela cadastral. Devido às atividades da administração territorial lidar com grandes quantidades de dados dinâmicos, é necessário um processo de manutenção contínuo na atualização do seu banco de dados. Estes processos devem prever procedimentos técnicos padronizados para coleta sistemática, atualização, processamento e distribuição de dados e informações.

Diferente do SIG, o SIT também registra e disponibiliza dados e informações que representam o relacionamento jurídico entre as pessoas e a terra. Trata-se de um sistema específico e detalhado para identificar partes do território e seus limites em grandes escalas cartográficas (parcelas), e é completo e geral de modo a cobrir toda a sua jurisdição de maneira sistemática (STUBKJAERS, 2006; MONDAL et al, 2016).

2.2.5 Banco de Dados

Os bancos de dados são uma tecnologia essencial em nossa sociedade atual e possuem papel crítico em praticamente todas as áreas do conhecimento em que computadores são utilizados, desde na área dos negócios, comércio eletrônico, engenharia, medicina, direito, educação e as ciências da informação (ELMASRI e NAVATHE, 2006).

Para a implementação de um sistema banco de dados, duas etapas são importantes: a modelagem conceitual e a criação do banco de dados. A primeira é um conjunto de conceitos usados para descrever a estrutura e operações do banco de dados. Nesta etapa, deve-se compreender e mapear os objetos e fenômenos que serão representados no ambiente computacional, assim como, a definição das classes e relacionamentos presentes no sistema. Já para a implementação física do banco de dados é indispensável um Sistema Gerenciador de Banco de Dados – SGBD onde são criadas as tabelas e relacionamentos previamente definidos na etapa de modelagem conceitual (PAZOLINI; SILVA; SOUZA; KOEHLER,2020).

O SGBD é a coleção de programas que permite ao usuário criar e manipular uma base de dados com facilidade. Dentre as principais funções estão a consulta, modificação,

atualização e geração de relatório dos dados armazenados (TAKAI; ITALIANO; FERREIRA, 2005).

Para Elmasri e Navathe (2006), trata-se de um sistema de software de propósito geral que facilita os processos de definição, construção, manipulação e compartilhamento de bancos de dados entre vários usuários e aplicações (ELMASRI e NAVATHE, 2006). Estes processos podem ser compreendidos como:

- Definição: Especificar os tipos de dados, as estruturas e as restrições para os dados a serem armazenados em um banco de dados;
- Construção: A construção de um banco de dados é o processo de armazenar os dados em alguma mídia apropriada controlada pelo SGBD;
- Manipulação: Funções de manipulação da base de dados, como pesquisas em banco de dados para recuperar um dado específico, atualização do banco, geração de relatórios dos dados;
- Compartilhamento: Permite aos diferentes usuários, programas e serviços acessar, de forma concorrente, o banco de dados. (p.3).

2.3 ANÁLISE DE DESEMPENHO

Segundo De Cesare (2003), O desempenho das avaliações de propriedades para fins tributários é importante para garantir a equidade e justiça fiscal. As avaliações devem ser compatíveis com a realidade do mercado, consistentes, explicáveis e justas para os contribuintes. Iniquidades nas avaliações são comuns e o controle de desempenho é essencial para garantir a qualidade do trabalho. O desempenho de avaliação é medido pela mensuração do nível e uniformidade de avaliação. A equidade da avaliação é alcançada quando as avaliações mantêm uma relação consistente com o valor comercial das propriedades. Iniquidades de avaliação ocorrem quando algumas classes de propriedades são avaliadas em níveis significativamente diferentes das outras classes. O controle de desempenho de avaliação é realizado por meio da comparação entre o valor de avaliação e o preço do imóvel. É importante ajustar os preços de venda para uma data fixa e utilizar uma amostra independente de dados. O controle de desempenho deve ser realizado para eliminar ou reduzir iniquidades nas avaliações.

A análise do desempenho da avaliação objetiva é a medição de dois aspectos: Nível e uniformidade da avaliação (IAAO 1999). Segundo De Cesare (2003), o nível de valorização refere-se à porcentagem ou proporção pela qual os imóveis são avaliados em relação ao valor de mercado; enquanto a uniformidade está relacionada ao tratamento tributário igual em relação às propriedades individuais.

De Cesare (2009) apresenta as principais medidas empregadas para verificar o nível e a uniformidade das avaliações, assim como as indicações dos limites considerados aceitáveis pelo IAAO (International Association of Assessing Officers) para a análise das medidas. O Instituto recomenda o uso da mediana do quociente entre o valor da avaliação e o preço de venda para identificar o nível pelo qual um grupo de imóveis foi avaliado com relação aos valores praticados no mercado imobiliário, demonstrando quão próximos os imóveis foram avaliados do nível legal ou desejado. A medida mais empregada para verificar a uniformidade das avaliações é o Coeficiente de Dispersão.

A análise de desempenho da avaliação consiste em comparar o valor da avaliação com o valor comercial do imóvel. O valor efetivamente empregado na tributação pode eventualmente ser diferente do valor da avaliação. Todavia, é importante que o comportamento do mercado seja mantido (SILVA, 2021).

Ainda de acordo com Silva (2021), o modelo tem as vantagens de ser ajustável a um grande número de situações, podendo ajustar relações lineares e não lineares entre preços de venda e valores avaliados, mesmo tendo a flexibilidade de modelar relações do tipo “S”, em que propriedades de preços médios são precificadas corretamente, as propriedades de baixo valor são supervalorizadas e as propriedades de alto valor são subestimadas.

3 MATERIAIS E MÉTODOS

Os materiais e métodos aplicados no desenvolvimento da pesquisa buscaram contemplar os objetivos estabelecidos, apresentando os procedimentos voltados para coleta e análise dos dados, bem como a definição das técnicas empregadas para validação e discussão dos resultados. Vale salientar que tais procedimentos foram balizados no referencial teórico relacionado ao tema da dissertação, que orientou a definição dos modelos conceitual e físico da base de dados do observatório do mercado imobiliário e das técnicas empregadas nas análises.

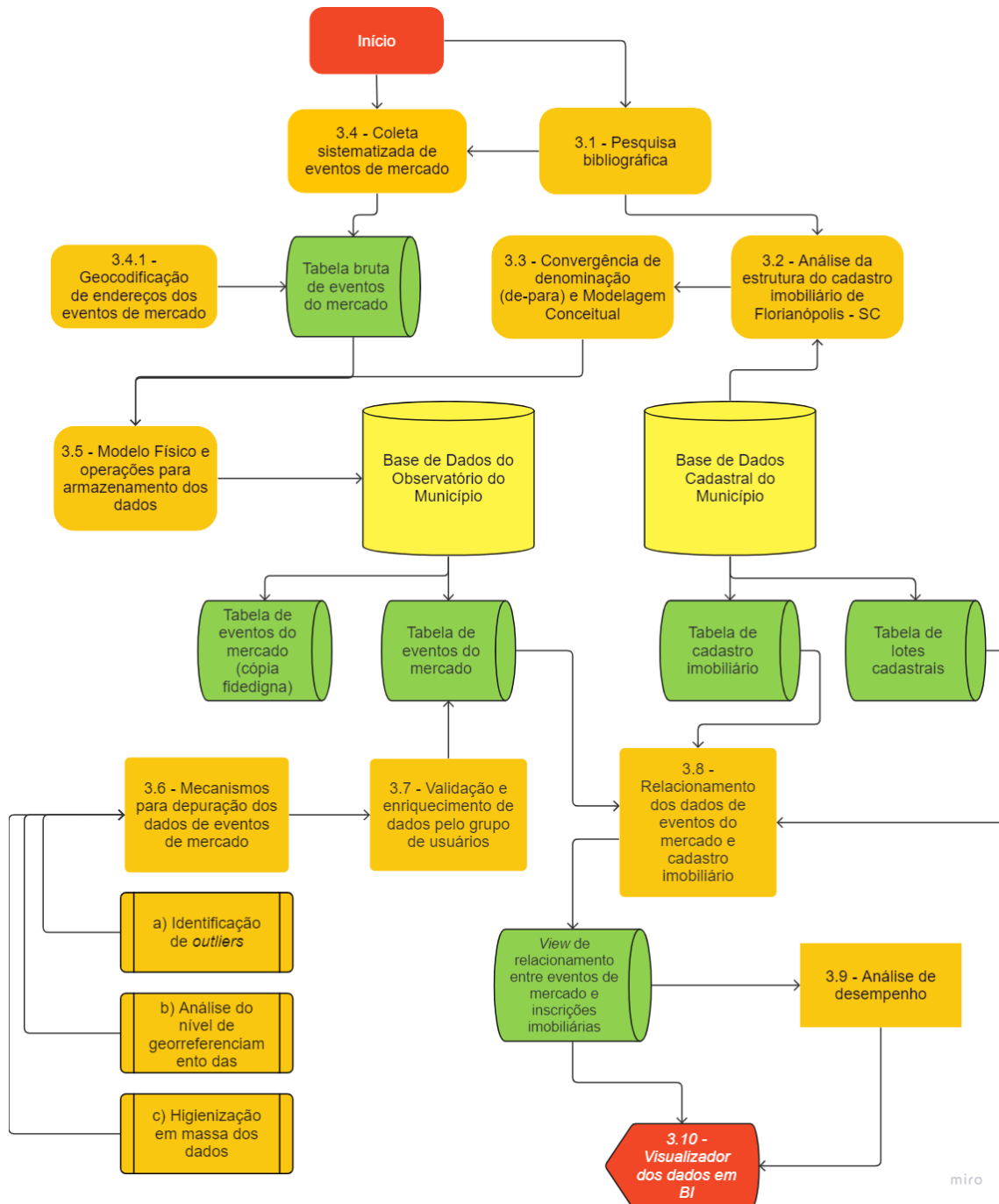
Um ponto importante na realização do trabalho, por ter o município de Florianópolis como área de estudo, foi a realização da análise e entendimento da estrutura cadastral e tributária definida em suas bases de dados, de modo a compreender suas particularidades.

Concomitantemente ao entendimento das bases de dados em uso pelo município, foi realizada a coleta de eventos do mercado imobiliário, sendo sistematizada com o emprego de um banco de dados geográfico e então normalizadas de acordo com particularidades do município. Nesta etapa, diferentes processos foram executados para depuração dos dados. Entre eles, a colaboração do grupo técnico do IPUF, que contribuiu para a validação, correção e cruzamento de dados.

Realizada a revisão dos dados levantados junto ao mercado imobiliário, procedeu-se a vinculação dos eventos ao cadastro imobiliário, tornando possível o emprego de ferramentas de análises e de monitoramento dos valores cadastrais.

A Figura 4 apresenta o fluxo de procedimentos metodológicos da pesquisa, além dos produtos relacionados aos mesmos.

Figura 4 – Fluxograma de procedimentos metodológicos da pesquisa.



Fonte: Elaborado pelo autor.

3.1 MATERIAIS

A definição dos materiais a serem utilizados na pesquisa contou com a busca no referencial teórico sobre observatório do mercado imobiliário, com foco nas experiências nacionais e latino-americanas. Pôde-se observar neste levantamento as principais ferramentas que embasaram a implementação de observatórios e assim decidir o rumo da pesquisa. Certamente a leitura dos textos possibilitou ir além da própria definição dos materiais a utilizar, pois os conceitos, funções e modelos propostos foram úteis às demais etapas do trabalho.

Ainda com o intuito de identificar os materiais de apoio à pesquisa, foi realizado o estudo do sistema cadastral do município, que compreendeu os distintos bancos de dados atrelados à gestão e uso dos dados cadastrais, sendo eles: banco de dados do IPUF e o Banco de dados do Geoprocessamento Cooperativo. Neste último, inclui-se ainda a análise das tabelas do banco de dados do Sistema de Tributos Municipais, disponíveis e integradas ao banco de dados em questão.

O banco de dados do IPUF é um conjunto de informações geográficas e cadastrais utilizadas para o planejamento e a gestão urbana da cidade de Florianópolis, em Santa Catarina, Brasil. Esses dados são coletados e mantidos pelo instituto, com o objetivo de apoiar a tomada de decisões em relação à gestão territorial, ambiental e urbana do município. Este banco de dados possui integração com as bases de dados das diferentes secretarias da Prefeitura, como por exemplo o banco de dados do Geoprocessamento Cooperativo de Florianópolis, e o banco de dados do Sistema de Tributos Municipais.

O banco de dados do Geoprocessamento Cooperativo de Florianópolis é um conjunto de informações geográficas e cadastrais utilizado pela Prefeitura de Florianópolis, em Santa Catarina, Brasil, para a gestão e planejamento do município. É o banco de dados específico para o Sistema de Geoprocessamento Cooperativo, que se trata de uma aplicação web para visualização e edição dos dados cartográficos e cadastrais, além de mapas temáticos das diferentes Secretarias Municipais. Este banco possui integração com o banco de dados do Sistema de Tributos Municipais, e as tabelas provenientes dessa integração foram também analisadas para compreensão da estrutura da informação e dos mecanismos de atualização cadastral.

Dado que um dos propósitos da pesquisa foi a viabilidade de aplicação dos resultados no âmbito municipal, de modo a trazer benefícios diretos à sociedade e o poder público, e principalmente levando em consideração as diferentes realidades orçamentárias dos diferentes municípios, optou-se por realizar a pesquisa com a utilização de ferramentas de código aberto e gratuitas, ou de baixo custo. Tais ferramentas foram empregadas para coletas sistematizadas de eventos do mercado, para visualização e interação com os dados, além de mecanismos para realizar a análise de desempenho dos modelos de avaliação em massa dos imóveis. A seguir são apresentados os softwares utilizados nesta pesquisa:

- a) Pacote LibreOffice Versão: 4.2
- b) OMT-G Designer (Web - Acesso em 2021-2022).
- c) PostgreSQL Versão: 12.
- d) PostGIS Versão: 3.1.
- e) Dbeaver Versão: 22.1.0.
- f) QGIS Versão: 3.20.0.
- g) Miro. (Web - Acesso em 2021-2022).
- i) Google Looker Studio: 2021.1.
- j) Notepad++ Versão: v8.4.

3.2 ANÁLISE DA ESTRUTURA DO CADASTRO IMOBILIÁRIO DE FLORIANÓPOLIS-SC.

A compreensão do modelo cadastral visou a análise da estrutura geométrica, econômica e jurídica do sistema cadastral do município, direcionando as análises na infraestrutura do banco de dados, na estrutura dos dados cadastrais, no modelo de cálculo do valor dos imóveis e no parcelamento do solo. Estes elementos foram considerados fundamentais para a modelagem do observatório do mercado imobiliário.

O primeiro passo foi o acesso ao banco de dados cadastral do município fornecido para a presente pesquisa, a partir da cooperação entre a Universidade Federal de Santa Catarina (UFSC), por meio do Grupo de Observação e Transformação do Território (GOTT), e a Prefeitura de Florianópolis, envolvendo o IPUF. O município

disponibilizou parte de sua base de dados para diferentes estudos acadêmicos, visando apoiar os avanços científicos que possam contribuir para a elaboração de seu próprio observatório.

Com o fim de respeitar a Lei Geral de Proteção dos Dados (LGPD), assim como a segurança da rede interna da Prefeitura, foram utilizados instrumentos para garantir o sigilo das informações municipais e a segurança de acesso. Para pesquisadores externos que não fazem parte do corpo técnico da prefeitura, como o autor desta pesquisa, foram utilizados os seguintes instrumentos administrativos: assinatura de termos de uso dos dados da intranet do IPUF, utilização de VPN com usuário e senha para acesso a rede, além de usuário e senha para acesso ao banco de dados.

3.3 MODELAGEM CONCEITUAL E CONVERGÊNCIA DE DENOMINAÇÃO

Esta etapa visou compreender a estrutura da base cadastral do município, para então desenvolver uma modelagem do banco de dados geográfico do observatório do mercado imobiliário contemplando as particularidades da administração territorial, e assim integrar este instrumento no sistema cadastral do município.

A partir dos avanços realizados pelo Grupo de Observação e Transformação do Território (GOTT), especificamente por meio do modelo conceitual proposto na dissertação de mestrado de PAZOLINI (2018), intitulada “Observatório de Valores Imobiliários: Modelagem Conceitual Baseada em OMT-G”, foram realizadas análises e adaptações desta modelagem no sentido de atender as particularidades presentes no modelo proposto nesta pesquisa.

Aliado à modelagem conceitual, está o processo de convergência de denominação, popularmente conhecido como De-Para. Nesta etapa, foram levantados os campos da estrutura de dados cadastrais dos imóveis do município, e equiparados com os campos dos dados dos eventos do mercado imobiliário. Foi criado um dicionário de relacionamento apresentando os campos equivalentes de cada uma das tabelas. Por exemplo, é possível que eventos de anúncios imobiliários façam a distinção tipológica entre “sala comercial” e “escritório”. Caso o cadastro imobiliário trate ambos como uma única variável denominada “sala”, o adequado é tratar as diferentes tipologias como tal. Cabe ressaltar, que apesar da normalização e agrupamento de variáveis em casos específicos, as originais foram preservadas na tabela bruta de coleta de dados do mercado imobiliário, garantindo a informação inicial para possível consulta.

3.4 COLETA SISTEMATIZADA DE EVENTOS DO MERCADO

Os dados do mercado imobiliário podem ser obtidos de diferentes fontes de informação. Atualmente a fonte que disponibiliza maior quantidade de eventos corresponde aos sites de imobiliárias, que por sua vez, em bom número, tem suas informações concentradas em portais especializados, como por exemplo: Vivareal, ZAP Imóveis e OLX.

Para as informações disponíveis em websites, há possibilidades de busca por acesso visual a cada um dos eventos de mercado registrados nas bases de dados dos portais, onde as informações são transcritas para uma planilha; ou por meio de algoritmos de busca automatizada (API REST), utilizando-se *Web Crawlers*, que captura os dados disponíveis e armazena diretamente em tabela de banco de dados geográfico.

Nesta etapa da pesquisa optou-se por utilizar mecanismos de busca automatizada de dados de portais imobiliários. Para a extração das informações diretamente dos portais web, avaliou-se que há duas principais formas: i) mapeando as variáveis de interesse do código fonte da página, que geralmente estão na composição HTML, e efetuando o *parse*, compreendido pelo processo de coletar as informações segmentando-as nos diferentes campos pré-determinados. Ou então ii) o acesso direto a *API REST* no portal imobiliário. Neste último, o dado apresenta-se estruturado, facilitando o processo de armazenamento e de atualização incremental. Utilizou-se a segunda alternativa como base para o levantamento dos eventos de mercado. Os dados coletados foram salvos em uma planilha CSV para aplicação de geocodificação dos endereços.

3.4.1 - Geocodificação dos endereços de anúncios imobiliários

Após a busca automatizada de anúncios de portais imobiliários web, a etapa seguinte consistiu em georreferenciar os endereços dos imóveis, utilizando a API Google Geocoding como geocodificador de endereços. A opção pela utilização desta API se deu pela vasta base de dados de endereços e informações de localização que o Google mantém em constante atualização.

Para as análises que se desenvolvem sobre o comportamento do mercado imobiliário, o conhecimento da localização dos eventos é fundamental para embasar melhor as estimativas e conclusões sobre a relação dos preços com o espaço. Neste

sentido, um dos importantes desafios da atualização dos dados em um observatório do mercado imobiliário é a definição da localização de cada evento, principalmente os que são anunciados em portais web.

No processo de busca automatizada de anúncios, e especificamente no processo de capturar informações de localização dos portais, estes dados são disponibilizados de duas formas: i) o endereço do imóvel de forma descritiva declarado pelo anunciante, e ii) as coordenadas geográficas alfanuméricas disponibilizadas por parte dos portais imobiliários.

De modo geral, baseando-se em experiências na busca automatizada de anúncios, é possível afirmar que os portais imobiliários adotam uma mesma estratégia para geração dessas coordenadas geográficas: utilizam o próprio endereço do imóvel (descritivo) e aplicam as APIs de geocodificação sobre ele, com o propósito de georreferenciar os imóveis e apresentá-los em mapas online nas páginas dos anúncios.

Dado o exposto, e de modo a evitar o retrabalho de executar as APIs de geocodificação em endereços disponíveis, o mais adequado seria a utilização das coordenadas já processadas pelas APIs de geocodificação utilizadas pelos portais imobiliários. O processo em si é eficaz para a captura de coordenadas, porém, a imprecisão gerada no georreferenciamento não é controlada, pois não se tem acesso aos metadados da geocodificação.

Nesse sentido, ao estudar o algoritmo da API de geocodificação Google Geocoding, conclui-se que esta apresenta, como resposta de requisição, uma variável fundamental para compreender a qualidade posicional dos dados: o campo *location_type*, que apresenta a precisão de geocodificação de acordo com a documentação da própria API do Google¹.

Como alguns portais imobiliários apenas executam o algoritmo de geocodificação a fim de georreferenciar seus imóveis, mas sem disponibilizar o campo *location_type* para o público em geral, optou-se por aplicar a API do Google Geocoding em todos os endereços disponíveis na pesquisa web de anúncios imobiliários. Desta forma, além das coordenadas, o propósito foi armazenar a qualidade posicional dos eventos, e assim propiciar uma análise mais objetiva por parte dos usuários envolvidos,

¹ Google Geocoding API - Solicitação e resposta de geocodificação: <https://developers.google.com/maps/documentation/geocoding/requests-geocoding?hl=pt-br>. Acesso em 2022.

visando a vinculação dos eventos de mercado com o correspondente imóvel da base cadastral.

A precisão do geocodificador possui 4 classes, são elas: *rooftop*, *range interpolated*, *geometric center* e *approximate*, de acordo com o manual da API Google Geocoding. De modo geral, o *rooftop* indica que o resultado da geocodificação reflete uma posição geográfica precisa, para a qual existe informação de localização até o nível do endereço do imóvel. O *range interpolated* indica que o resultado da geocodificação representa uma aproximação entre dois pontos conhecidos no endereço, ou seja, não está no endereço do imóvel, porém posicionado pela distância métrica entre dois endereços já conhecidos. O *geometric center* indica que o resultado da geocodificação é o centro geométrico de um resultado, como por exemplo, o centróide de uma linha (logradouro) ou polígono (bairro). Por fim, a classe *approximate* apresenta a estimativa de um lugar aproximado, em muitos casos indica que o processo de geocodificação falhou.

3.5 MODELO FÍSICO E OPERAÇÕES PARA ARMAZENAMENTO DOS DADOS

Esta etapa visou implementar a infraestrutura necessária para um observatório do mercado imobiliário em uma base de dados pré-existente no município. Para tanto, fez-se a configuração e estruturação do banco de dados geográfico mediante a utilização do Sistema Gerenciador de Banco de Dados (SGBD) PostgreSQL Server. Esta operação compreendeu a definição dos grupos de usuários, usuários, databases, tabelas, relacionamentos, permissões e funções no banco de dados.

O modelo físico foi implementado usando o banco de dados PostgreSQL, com a extensão espacial PostGIS. Complementarmente foi utilizado o programa Dbeaver, uma aplicação gráfica de acesso ao SGBD capaz de reproduzir as seguintes etapas: conexão com servidor; criação da base de dados; criação do esquema; criação dos grupos de usuários e usuários; criação da extensão espacial; criação das tabelas do modelo; criação de funções e gatilhos (*triggers*).

Após a implementação do modelo físico do observatório no SGBD do município, foram desenvolvidos os *scripts* de criação das tabelas do observatório em linguagem SQL, contemplando os ajustes do subitem 4.3 e 4.7, respeitando as particularidades da base de dados do município. Neste processo ocorreu a transformação do tipo de dado e propriedade de campo das tabelas de origem, considerando a estrutura das tabelas de destino (modelo físico implementado). Na sequência, foi avaliado o tamanho máximo de

um valor de campo, tipo de campo (textual, numérico, geométrico, entre outros), se se tratavam de campos indexados e a necessidade de uma chave primária.

Outra importante operação foi a adequação dos dados georreferenciados no subitem 4.6. As coordenadas alfanuméricas resultantes da geocodificação foram transformadas em *geometry*, tipo de dado geométrico da extensão espacial PostGIS, respeitando os sistemas de referência e de coordenadas da base de dados municipal.

Essas operações, embora tenham ocorrido em uma etapa avançada da pesquisa, foram estudadas antes mesmo da modelagem conceitual e serviram para orientar o desenvolvimento do trabalho.

3.6 MECANISMOS PARA DEPURAÇÃO DOS EVENTOS DE MERCADO

Esta etapa consistiu em definir procedimentos para correção e validação dos dados de anúncios imobiliários, tanto a partir de *scripts* SQLs de correção em massa, quanto a partir de correção específica com o uso do software de geoprocessamento QGIS. Três principais procedimentos foram realizados:

a) Análise de casos atípicos (outliers)

Nesta etapa foram criados scripts para avaliação de dados com comportamento fora do padrão, comumente denominados de *outliers* em estudos estatísticos. Barnett e Lewis (Barnett e Lewis, 1994) indicam que uma observação de ponto fora da curva, ou *outlier*, é aquela que parece desviar-se significativamente de outros membros da amostra em que ocorre. Johnson (1992) define um outlier como uma observação em um dado conjunto que parece ser inconsistente com o restante desse conjunto de dados.

O primeiro passo foi definir quais campos da tabela eram considerados cruciais para análise de outliers: preço e área. Após esta etapa, definiu-se os filtros para isolamento da amostra a ser avaliada. Por exemplo, amostras de eventos de mercado da tipologia “terreno” com o tipo de negócio “venda”. Foram então criados *scripts* SQLs, utilizando algoritmos de análise estatística, para selecionar estes eventos de mercado fora da curva.

b) Análise do nível de georreferenciamento dos eventos de mercado

Esta etapa consistiu em analisar o georreferenciamento dos eventos de mercado. Foram utilizadas duas principais frentes para análise: Scripts SQLs espaciais para validar

a qualidade posicional das coordenadas geográficas; e a resposta de precisão correspondente aos retornos da API Google Geocoding para tais coordenadas.

c) Higienização em massa dos dados

Esta etapa consistiu em analisar as informações alfanuméricas da pesquisa de mercado, buscando corrigir ou eliminar dados inconsistentes nos diferentes campos da tabela. Foram avaliados erros comuns gerados na digitação de anúncios imobiliários por anunciantes, como por exemplo a utilização do ponto no lugar da vírgula, principalmente no separador milhar e/ou separador decimal.

3.7 VALIDAÇÃO E QUALIFICAÇÃO DOS DADOS PELO GRUPO TÉCNICO DO IPUF

Nesta etapa foram avaliados os meios de disponibilização dos dados do banco de dados para o grupo técnico do IPUF para análise e correção dos dados de mercado. Foi preparado um ambiente de consulta de dados utilizado o sistema de informação geográfica QGIS.

O grupo técnico teve acesso direto ao banco de dados para análise e correção dos dados de mercado. Neste processo de validação e qualificação, foram realizadas correções de inconsistências e erros identificados pelos usuários, além da inserção de novas informações relevantes. Dessa forma, a parceria com o IPUF foi fundamental para a validação e qualificação dos dados de mercado, garantindo a confiabilidade e a qualidade dos dados utilizados na presente pesquisa.

3.8 RELACIONAMENTO DOS DADOS DE EVENTOS DO MERCADO E CADASTRO IMOBILIÁRIO

Segundo Silva (2006), a criação de uma base de dados do mercado imobiliário é um passo importante para avançar na qualificação da tributação imobiliária. Isso ocorre porque o valor de mercado das propriedades varia devido a diferentes atributos, e quanto maior o espaço urbano, maior a heterogeneidade desses valores. Quando essa base de dados é vinculada ou relacionada ao cadastro técnico, ela permite agregar uma grande variedade de informações, capazes de fornecer os atributos necessários para a definição de modelos que levem a valores prováveis de mercado de maneira eficiente.

Neste contexto, esta etapa teve como objetivo o vínculo dos registros dos eventos de mercado, após devidamente georreferenciadas, com as parcelas do cadastro territorial, utilizando-se a inscrição imobiliária como elemento de ligação.

3.9 ANÁLISE DE DESEMPENHO

O objetivo desta etapa foi aplicar as medidas de desempenho (nível e uniformidade), de modo a analisar e monitorar os valores cadastrais calculados pelos modelos de avaliação em massa em uso nos municípios. Para o estudo, utilizou-se dados sobre preços dos imóveis disponíveis no observatório do mercado imobiliário de Florianópolis relacionados com o cadastro imobiliário. As medidas foram calculadas para os valores cadastrais dos terrenos baldios do município.

Utilizou-se, como referência para elaboração dos cálculos e para aplicação das medidas para análise de desempenho, o quadro apresentado na Figura 5, proveniente da pesquisa realizada por Machado et al. (2020), e que já agrupam em um quadro estruturado as principais fórmulas utilizadas pelo IAAO e pelas Diretrizes Nacionais na implantação do Cadastro Técnico Multifinalitário (BRASIL, 2009).

Figura 5 – Fórmulas utilizadas como base para análise de desempenho.

Variável em estudo: quociente entre o valor avaliado e o preço de venda.		Limites aceitáveis
Nível de avaliação	Média aritmética: soma da variável estudada dividida pelo número de observações.	0,90 – 1,10 do nível desejado (IAAO, 1999) 0,70 – 1,00 (BRASIL, 2009)
	Mediana: classificação das observações de acordo com o valor da variável estudada. A mediana é a observação central no caso de um número ímpar de observações; ou a média das duas observações centrais no caso de um número par de observações.	
	Média ponderada: soma do valor avaliado dividido pela soma do preço de venda para todo o conjunto de imóveis em análise.	
Uniformidade	Coeficiente de Dispersão (%) (CD): 1. Cálculo do desvio de cada observação (quociente) em relação à mediana, em módulo (valores absolutos); 2. Soma desses desvios no módulo; 3. Divisão pelo número de observações (obtenção do desvio absoluto médio); 4. Divisão pela mediana; 5. Multiplicação por 100 (resultado em porcentagens).	i. = < 10 – 15% residencial ii. = < 15 – 20%, demais segmentos. (IAAO, 1999) < = 30% (BRASIL, 2009)
	Diferencial relacionado ao preço (DRP): é a média dividida pela média ponderada, com um indicativo de regressividade se o resultado for > 1; ou progressividade se o resultado for < 1.	i. Se < 0.98 indica progressividade. ii. Se > 1.03 indica regressividade. (IAAO, 1999)

Fonte: adaptado de De Cesare (2012) e Silva *et al.* (2015).

Fonte: Machado et al. (2020).

Além das fórmulas, realizou-se uma sequência de procedimentos para avaliar e utilizar apenas os dados consistentes, tanto para o lado dos dados de eventos de mercado, quanto aqueles do cadastro imobiliário. As amostras foram separadas pelos diferentes Distritos de Florianópolis, e selecionados aqueles que continham um número mínimo de eventos aceitável para a análise. Foram eles: Sede Insular, Lagoa da Conceição, São João do Rio Vermelho, Cachoeira do Bom Jesus, Sede Continental, Campeche, Pântano do Sul, Ribeirão da Ilha, Canasvieiras, Ratoles, Santo Antônio de Lisboa, Ingleses do Rio Vermelho, Barra da Lagoa.

3.10 - VISUALIZADOR DOS DADOS EM BI

Com o objetivo de tornar públicas as informações do Observatório e possibilitar avaliação interna do grupo técnico com relação ao comportamento do mercado

imobiliário, foi desenvolvido um Business Intelligence (BI) gratuito, utilizando a plataforma Looker Studio (anteriormente conhecida como Google Data Studio). Por meio dessa plataforma, foram elaborados gráficos de visualização e implementados filtros dinâmicos, que permitem uma análise e visualização mais precisas dos eventos de mercado e dos dados cadastrais imobiliários.

4 RESULTADOS E DISCUSSÕES

4.1 ANÁLISE DA ESTRUTURA DO CADASTRO IMOBILIÁRIO DE FLORIANÓPOLIS – SC

Esta etapa do estudo visou a avaliação da estrutura cadastral do município, com a finalidade de estabelecer possíveis relacionamentos entre suas informações e as das tabelas de eventos de mercado. O objetivo foi o de padronizar os campos existentes nas referidas tabelas e, assim, desenvolver um observatório plenamente funcional.

Para realizar a análise da estrutura cadastral foram obtidas informações sobre as características dos imóveis cadastrados no município, tais como sua localização, área total do lote, número de unidades, valor venal e outros atributos considerados importantes. Neste sentido, a primeira etapa consistiu na análise da tabela COTR_IMOBILIARIO, a qual representa a principal tabela da base cadastral do município. Os dados desta tabela são administrados Sistema de Tributos Municipais, sendo a referida tabela espelhada e integrada ao banco de dados do IPUF e o banco de dados do Sistema de Geoprocessamento.

O número de registros na tabela do cadastro imobiliário, quando da análise realizada, era de 472574 unidades, que representam o total de unidades autônomas do Município. Dentro de cada parcela, também denominada “lote” pelos sistemas tributário e cadastral, é possível identificar a presença de uma ou mais unidades autônomas. Os casos de múltiplas unidades em um mesmo lote devem ser avaliados cuidadosamente nos processos de vinculação com os eventos de mercado, de modo a evitar-se relacionamentos equivocados.

A tabela é composta por um total de 109 campos. Avaliou-se quais deles seriam pertinentes à estruturação do observatório do mercado imobiliário, focando principalmente em informações que pudessem ser relacionadas com dados de eventos de mercado. O Quadro 1 apresenta quais campos selecionados para apoiar a pesquisa:

Quadro 1 – Campos selecionados da tabela “cotr_imobiliario”.

Campo	Definição
nu_insc_imbl	Inscrição Imobiliária
tp_cobr	Cobertura
tp_estr	Estrutura
tp_ocpc	Ocupação
tp_sitc	Situação
tp_utlz	Utilização
tp_patr	Patrimônio
tp_ocpc_lote	Ocupação do Lote
tp_pedl	Pedologia
tp_topg	Topografia
tp_sitc_qudr	Situação Quadra
qt_area_cons	Área Construída
qt_area_trbt_terr	Área Tributável Terreno
vl_venl_prdl	Valor Venal Predial
vl_m2_prdl	Valor M2 Predial
vl_venl_terr	Valor Venal Terreno
vl_m2_terr	Valor M2 Terreno
aa_cons	Ano de Construção
qt_area_lote	Área Lote
qt_area_undd	Área da Unidade
nu_imvl	Número de Porta
cd_logr	Código do Logradouro
cd_lote	Código do Lote
cd_quadra	Código da Quadra

Fonte: Elaborado pelo autor.

A tabela “cotr_imobiliario” não inclui os dados geográficos, como a geometria dos lotes, por exemplo. Com o objetivo de estabelecer o relacionamento entre os dados de “cotr_imobiliario” e os registros da base geoespacial (geometrias dos lotes), foi realizada uma análise na tabela “lotes”, fornecida pelo IPUF. O quadro 2 apresenta a tabela “lotes” e seus campos, juntamente com suas respectivas traduções:

Quadro 2 – Campos selecionados da tabela lotes.

Campo	Definição
gid	Identificador Geométrico Único
inscricao	Inscrição Imobiliária
numero	Número do Lote
geom	Campo Geométrico (Polígono - SRID: 31982 SIRGAS2000 UTM22S

Fonte: Elaborado pelo autor.

A tabela de lotes é composta por 120744 registros, que por tratar-se de uma tabela de dados geográficos, este número pode ser considerado como o total de polígonos na mesma. A tabela possui inscrições imobiliárias até o nível de código de lote. Isso significa que, ao se relacionar com a tabela “cotr_imobiliario”, é possível que um lote da tabela “lotes” seja vinculado com mais de um registro da primeira.

Outra tabela empregada foi a dos Distritos Administrativos de Florianópolis, sendo utilizada especialmente para a agregação dos eventos de mercado por localidade, por meio do cruzamento espacial entre os eventos e a geometria de cada distrito. Essa escolha se deveu ao fato de que os bairros indicados nos anúncios imobiliários não coincidem com os limites oficiais do Município, sendo comum a adoção de bairros informais ou localidades.

Os Distritos de Florianópolis da base municipal são: Barra da Lagoa, Cacheira do Bom Jesus, Campeche, Canasvieiras, Ingleses do Rio Vermelho, Lagoa da Conceição, Pantano do Sul, Ratonas, Ribeirão da Ilha, Sede Continental, Sede Insular, Santo Antônio de Lisboa, São João do Rio Vermelho. Foi analisada a estrutura da tabela e definiu-se como sendo de interesse ao trabalho os campos apresentados no Quadro 3.

Quadro 3 – Campos selecionados da tabela distritos administrativos.

Campo	Definição
gid	Identificador Geométrico Único
cd_dist_ad	Código do Distrito Administrativo
mm_dist_ad	Nome do Distrito Administrativo
geom	Campo Geométrico (Polígono - SRID: 31982 SIRGAS2000 UTM22S)

Fonte: Elaborado pelo autor.

4.2 - COLETA SISTEMATIZADA DE EVENTOS DE MERCADO

O processo de coleta sistematizada de eventos do mercado consistiu inicialmente em obter acesso as APIs REST dos portais imobiliários. Para realizar a coleta das informações utilizou-se os filtros disponíveis na plataforma para a coleta dos dados de interesse. Nesse sentido, optou-se por utilizar o filtro cidade: “Florianópolis”, tipo de negócio: “aluguel” e “venda”, tipo de imóvel: “todos os imóveis”.

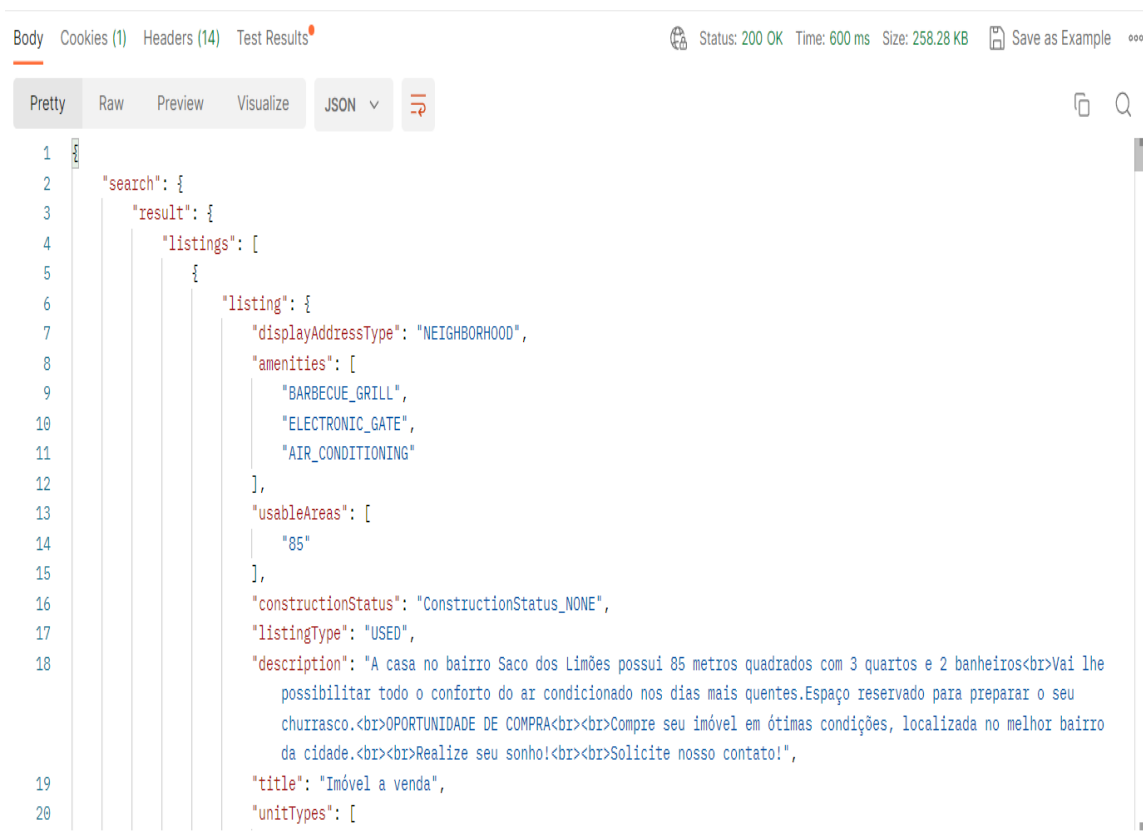
Na busca por compreender a estrutura dos dados provenientes do retorno da API, utilizou-se inicialmente o software Postman, em sua versão gratuita. Trata-se de um software que serve principalmente para solicitação de APIs e que possui recursos que facilitam a visualização dos dados. Utilizou-se o comando GET² para solicitação das informações. Na Figura 6, pode-se observar o retorno das informações, que são apresentadas em formato JSON³, estruturado e indentado.

² Os comandos de requisição de API mais comuns são:

GET: usado para recuperar informações de um recurso especificado. Este comando é padronizado pelo protocolo HTTP. Disponível em: <https://tools.ietf.org/html/rfc7231>.

³ JSON: JavaScript Object Notation é um formato de intercâmbio de dados leve, que é fácil de ler e escrever. Ele é baseado em um subconjunto da linguagem de programação JavaScript e é independente de plataforma. O JSON é comumente utilizado para transmitir dados entre um servidor e uma aplicação web, como uma alternativa ao XML. Disponível em: <https://www.json.org/json-pt.html>

Figura 6 – Exemplo e requisição de API através do Software Postman.



```
1 {}
2 "search": {
3   "result": {
4     "listings": [
5       {
6         "listing": {
7           "displayAddressType": "NEIGHBORHOOD",
8           "amenities": [
9             "BARBECUE_GRILL",
10            "ELECTRONIC_GATE",
11            "AIR_CONDITIONING"
12          ],
13          "usableAreas": [
14            "85"
15          ],
16          "constructionStatus": "ConstructionStatus_NONE",
17          "listingType": "USED",
18          "description": "A casa no bairro Saco dos Limões possui 85 metros quadrados com 3 quartos e 2 banheiros<br>Vai lhe possibilitar todo o conforto do ar condicionado nos dias mais quentes.Espaço reservado para preparar o seu churrasco.<br>OPORTUNIDADE DE COMPRA<br><br>Compre seu imóvel em ótimas condições, localizada no melhor bairro da cidade.<br><br>Realize seu sonho!<br><br>Solicite nosso contato!",
19          "title": "Imóvel a venda",
20          "unitTypes": [
```

Fonte: Elaborado pelo autor.

Com a requisição exemplificada na Figura 6, analisou-se quais campos seriam coletados para armazenamento. Neste ponto, a pesquisa bibliográfica sob os diferentes observatórios existentes na América Latina, aliado à análise realizada no item 3.2 (Análise da estrutura de dados de Florianópolis), serviram como base para a definição dos campos de interesse. Toda informação útil com respeito a características físicas, características geográficas, características quantitativas, dados sobre a origem do anúncio, e informações adicionais que contribuíssem para qualificação de um imóvel foram armazenadas. O Quadro 4 apresenta os campos de interesse que foram selecionados e a definição de cada um deles.

Quadro 4 – Campos selecionados a partir da coleta sistematizada de eventos do mercado.

Campo	Definição
id	Identificador Único
created_at	Data de criação
updated_at	Data de atualização
title	Título
description	Descrição
street	Rua
street_number	Número do endereço
zip_code	Código postal
neighborhood	Bairro
city	Cidade
state	Estado
country	País
area_size	Área
price	Preço
iptu_price	Preço do IPTU
condominium_price	Preço do condomínio
link	Link
advertiser	Anunciante
phone	Telefone
bedrooms	Quantidade de quartos
bathrooms	Quantidade de banheiros
suite	Quantidade de suítes
parking	Quantidade de vagas de garagem
unit_type	Tipo de unidade
list_type	Tipo de lista
business_type	Tipo de negócio
construction_status	Estado de construção
external_id	ID externo
origin_id	ID de origem

Fonte: Elaborado pelo autor.

Cada um dos campos apresentados desempenha um papel importante na obtenção e organização das informações do OMI. A seguir são apresentados a função de cada um deles:

- id (Identificador Único): Serve como um identificador exclusivo para cada registro, facilitando a referência e o acesso aos dados específicos de um anúncio.

- `created_at` (Data de criação) e `updated_at` (Data de atualização): Permitem rastrear quando o registro foi criado e atualizado pela última vez, fornecendo informações sobre a cronologia dos dados coletados.
- `title` (Título): Contém o título ou nome do anúncio, que geralmente fornece uma breve descrição do imóvel ou propriedade anunciada.
- `description` (Descrição): Contém informações detalhadas sobre o imóvel, incluindo características, comodidades, localização e outras informações relevantes fornecidas pelo anunciante.
- `street` (Rua), `street_number` (Número do endereço), `zip_code` (Código postal), `neighborhood` (Bairro), `city` (Cidade), `state` (Estado) e `country` (País): Esses campos fornecem informações sobre a localização do imóvel, o que é crucial para a análise do mercado imobiliário em diferentes regiões e áreas geográficas.
- `area_size` (Área): Indica a área do imóvel em metros quadrados ou outra unidade de medida padrão, permitindo a comparação e avaliação de imóveis com base em seu tamanho.
- `price` (Preço), `iptu_price` (Preço do IPTU) e `condominium_price` (Preço do condomínio): Esses campos fornecem informações sobre os preços associados ao imóvel, incluindo o preço de venda, custo do IPTU e valor do condomínio, que são fatores importantes na análise do mercado imobiliário.
- `link` (Link): Armazena o link ou URL do anúncio original, permitindo o acesso direto ao anúncio na plataforma de origem.
- `advertiser` (Anunciante) e `phone` (Telefone): Esses campos registram informações sobre o anunciante, como o nome do anunciante e o número de telefone de contato, o que pode ser útil para fins de referência e acompanhamento.

- bedrooms (Quantidade de quartos), bathrooms (Quantidade de banheiros), suite (Quantidade de suítes) e parking (Quantidade de vagas de garagem): Esses campos fornecem informações sobre as características do imóvel em termos de quartos, banheiros, suítes e vagas de garagem, auxiliando na análise das preferências e tendências do mercado imobiliário.
- unit_type (Tipo de unidade), list_type (Tipo de lista), business_type (Tipo de negócio) e construction_status (Estado de construção): Esses campos classificam o tipo de unidade, o tipo de lista (venda, aluguel, etc.), o tipo de negócio (residencial, comercial, etc.) e o estado de construção do imóvel, fornecendo informações adicionais para análise e categorização dos dados.
- external_id (ID externo) e origin_id (ID de origem): Esses campos podem ser usados para registrar IDs adicionais associados ao anúncio, como IDs de referência da plataforma.

O próximo passo foi a criação de um *script Python* para a coleta sistemática de dados. Utilizou-se então a função GET, por meio da biblioteca *requests*. Isso possibilitou a obtenção dos dados desejados na plataforma em formato JSON, contendo as informações dos imóveis disponíveis na plataforma.

Obtidos os dados, foi realizada a etapa de tratamento, utilizando as funções de manipulação de dados do Python para selecionar apenas as informações relevantes, descartar as informações desnecessárias e aplicar a formatação adequada para garantir a consistência dos dados.

Em seguida foi criada uma tabela no banco de dados PostgreSQL para armazenar os dados coletados, com o emprego do comando CREATE TABLE, projetando a tabela intitulada de “properties” (propriedades) com as colunas correspondentes aos dados obtidos na API e definindo os tipos de dados de cada coluna. Utilizou-se a biblioteca *psycopg2* do Python para estabelecer uma conexão com o banco de dados PostgreSQL e inserir os dados obtidos da API na tabela. Essa etapa foi realizada com um loop que percorreu todos os registros da resposta JSON, inserindo cada registro na tabela.

Através de uma abordagem de coleta de dados planejada, foi possível reunir informações de maneira mensal no Município de Florianópolis. Essa prática visou estabelecer uma base de dados dinâmica que serve e servirá como alicerça para o

observatório atualmente em fase de implementação. Cumpre ressaltar, entretanto, que um marco temporal foi delimitado para nortear o desenvolvimento deste estudo, englobando o período de Setembro de 2020 a Março de 2021.

Neste período foram levantados 61.365 eventos de mercado para o município de Florianópolis. Destes, um total de 11.425 com o “tipo de transação” = “aluguel” e outros 49.933 como “tipo de transação” = “venda”. Com relação à tipologia dos imóveis, foram coletados:

Tabela 1 – Tabela dos quantitativos de anúncios imobiliários por tipologia.

Tipologia	Número de Eventos	Percentual
apartamento	37.348	60.88%
galpão	89	0.15%
terreno	4.149	6.76%
especial	26	0.04%
sala/loja	6.159	10.04%
casa	13.594	22.14%
Total	61365	100%

Fonte: Elaborada pelo autor.

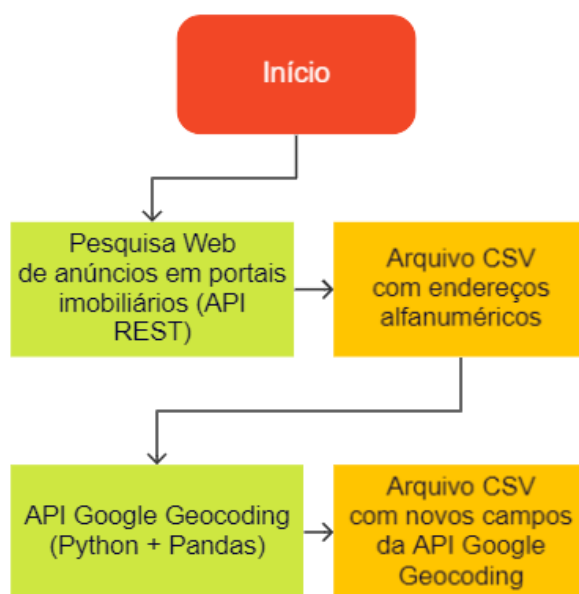
Observa-se pelos números apresentados na Tabela 1 que o seguimento de apartamentos se destaca em número de anúncios (60,86%), demonstrando o crescimento da demanda por este tipo de imóvel no município. O segmento de casas tem um número expressivo (22,15%), porém não tanto quando o de apartamentos. No caso do segmento de terrenos, o percentual é importante (6,76%), mas demonstra também a relação com o número de terrenos vazios existentes na cidade. Possivelmente alguns dos anúncios de terrenos correspondem a imóveis construídos em que a renovação da ocupação seja mais vantajosa aos que procuram empreender no ramo imobiliário.

4.3 - GEOCODIFICAÇÃO DE ENDEREÇOS

Para análise da precisão do georreferenciamento dos dados levantados no mercado imobiliário, adotou-se três procedimentos, conforme segue: i) coleta dos endereços: foi realizada a utilização dos endereços alfanuméricos dos imóveis em oferta, sendo os dados coletados organizados em uma tabela no formato CSV; ii) aplicação do algoritmo de geocoding: neste procedimento foram realizadas às requisições da API

Google Geocoding via linguagem Python e biblioteca Pandas para a leitura do endereço dos imóveis na planilha em formato CSV; e iii) com os resultados da qualidade de precisão vinculados aos endereços de imóveis, a última etapa consistiu em unificar estas informações com a tabela bruta de anúncios imobiliários, preenchendo os campos “point_precision” com os valores de precisão, e os campos “latitude” e “longitude” com as coordenadas geográficas. A Figura 7 apresenta o fluxograma deste processo.

Figura 7 – Fluxograma do processo de geocodificação.



Fonte: Elaborado pelo autor.

Visando exemplificar as etapas apresentadas na Figura 7, a Figura 8 exibe um recorte dos resultados de endereços alfanuméricos obtidos na pesquisa de eventos do mercado, já salvo em planilha de extensão CSV. A coleta sistemática foi realizada campo a campo, então nota-se que cada elemento que compõe um endereço encontra-se em um campo distinto, sendo eles: "street", "street_num", "zip_code", "neighborhood", "city", "state" e "country".

Figura 8 – Exemplo de endereços de eventos de mercado coletados.

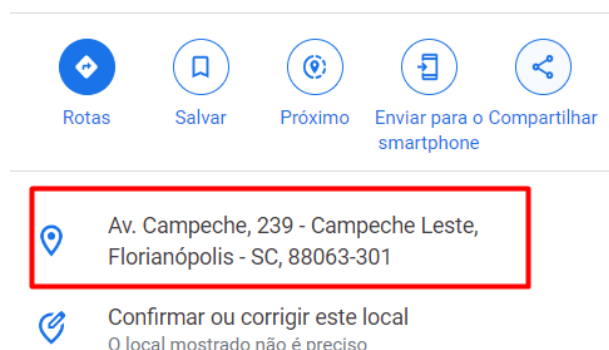
	F	G	H	I	J	K	L
1	street	street_num	zip_code	neighborho	city	state	country
2	Avenida Pequeno Príncipe	2	88063000	Campeche	Florianópolis	SC	BR
3	Rua Marinas do Campeche	406	88065185	Campeche	Florianópolis	SC	BR
4	Rua Gilmar Darli Vieira	400	88063650	Campeche	Florianópolis	SC	BR
5	Rua Silyjo Lopes Araújo	555	88048391	Rio Tavares	Florianópolis	SC	BR
6	Servidão São Francisco de Assis	33	88063360	Campeche	Florianópolis	SC	BR
7	Avenida Pequeno Príncipe	1000	88063000	Campeche	Florianópolis	SC	BR
8	Rua Jerônimo Venâncio das Chagas	200	88063660	Campeche	Florianópolis	SC	BR
9	Avenida Campeche	805	88063300	Campeche	Florianópolis	SC	BR
10	Servidão Jaborandi	1154	88065035	Campeche	Florianópolis	SC	BR
11	Rua Gilmar Darli Vieira	312	88063650	Campeche	Florianópolis	SC	BR
12	Rua Nicolau João de Abreu	224	88063600	Campeche	Florianópolis	SC	BR
13	Avenida Pequeno Príncipe	1000	88063000	Campeche	Florianópolis	SC	BR
14	Avenida Campeche	1670	88063300	Campeche	Florianópolis	SC	BR
15	Servidão Canto das Pérolas	24	88063076	Campeche	Florianópolis	SC	BR
16	Avenida Campeche	805	88063300	Campeche	Florianópolis	SC	BR
17	Rua Heloisa Rojo Machado	131	88066066	Campeche	Florianópolis	SC	BR
18	Rua Silyjo Lopes Araújo	321	88048391	Rio Tavares	Florianópolis	SC	BR
19	Rua Ipê da Praia	700	88064086	Ribeirão Da Ilha	Florianópolis	SC	BR
20	Rua Gilmar Darli Vieira	312	88063650	Campeche	Florianópolis	SC	BR
21	Rua Nivaldo Dias	102	88063670	Campeche	Florianópolis	SC	BR
22	Rua Pau de Canela	606	88063505	Campeche	Florianópolis	SC	BR
23	Rua Otávio Cruz	138	88063620	Campeche	Florianópolis	SC	BR
24	Rua Nivaldo Dias	52	88063670	Campeche	Florianópolis	SC	BR
25	Rua Reverendo Gelson dos Santos Castro	676	88048340	Rio Tavares	Florianópolis	SC	BR

Fonte: Elaborado pelo autor.

Para enviar os endereços para a API de Geocodificação, foi necessário agrupar os campos em um campo único de endereço, considerando ainda a correta ordenação dos mesmos e os respectivos símbolos de separação dos campos, de forma a seguir os padrões da API, garantindo assim uma maior precisão na geocodificação.

Esta ordenação padrão foi possível a partir da observação de como são apresentados os endereços na plataforma Google Maps, conforme Figura 9:

Figura 9 – Tela da aplicação Google Maps, exemplificando o padrão adotado de endereçamento pela tecnologia.



Fonte: Elaborado pelo autor.

Com a correta padronização, criou-se um campo intitulado "address" no arquivo CSV. Utilizando a fórmula “=CONCAT” do LibreOffice, agrupou-se todos os campos nesse único campo de endereço. A partir daí iniciou-se o processo de elaboração do *script* em linguagem *Python* de requisição dos endereços.

A elaboração do *script Python* de requisição da API de Geocodificação foi precedida da criação de um usuário Google para geração de uma chave de API (*API KEY*), que foi utilizada como um dos parâmetros de requisição. Outro requisito para implementação do *script* foi a importação de bibliotecas *Python*, sendo elas: *Pandas*, *Requests*, *Logging* e *Time*. Estas bibliotecas foram responsáveis por realizar a leitura e exportação de arquivos em CSV, interpretação dos arquivos “*json*” gerados pela API Google Geocoding, permissões de acesso ao usuário da API e definição de tempo para cada processo do *script* de requisição.

O *script Python* foi desenvolvido para realizar as seguintes operações:

- Ler o arquivo CSV de eventos do mercado;
- Rodar a API de Geocodificação sobre o campo “address”;
- Armazenar os novos campos criados pela API no mesmo CSV de origem.

A API retorna, como resposta a requisição de geocodificação, o campo *formatted_address*, que é responsável por apresentar o endereço alfanumérico completo do retorno da requisição, o campo *location_type*, que apresenta os tipos de precisão de geolocalização alcançados pelo geocodificador, o campo *lat* apresentando as coordenadas de latitude e o campo *long* com as coordenadas de longitude, ambos no sistema de referência WGS84 (Sistema Geodésico Mundial 1984, de coordenadas geográficas) (Figura 10). Para casos em que não houve qualquer retorno da API, ou seja, em que os campos vieram nulos, foi utilizado a variável “*PRECISION_NONE*”. Estes campos adicionais foram salvos e acrescentados ao arquivo CSV de origem.

Figura 10 – Exemplo de campos que são retornados pela API Google Geocoding.

	B	C	D	E
1	formatted_address	latitude	longitude	location_type
2	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
3	Biguaçu River, State of Santa Catarina, Brazil	-27.5118086	-48.7640716	APPROXIMATE
4	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
5	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
6	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
7	Centro, Biguaçu - SC, Brazil	-27.5023677	-48.6464068	APPROXIMATE
8	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
9	R. Cónego Rodolfo Machado, 1587 - Rio Caveiras, Biguaçu - SC, 88160-000, Brazil	-27.5087614	-48.647761	ROOFTOP
10	Biguaçu, State of Santa Catarina, Brazil	-27.4955667	-48.65536600000001	APPROXIMATE
11	Rod. SC 407, nº2546 - Bairro Deltaville, Biguaçu - SC, 88164-183, Brazil	-27.4896628	-48.6663754	ROOFTOP
12	R. Sebastião Laurentino da Silva, 1002 - Córrego Grande, Florianópolis - SC, 88037-400, Brazil	-27.6080096	-48.504618	ROOFTOP
13	Servidão da Figueira Velha, 241 - Córrego Grande, Florianópolis - SC, 88037, Brazil	-27.6068693	-48.4975804	RANGE_INTERPOLATED
14	R. Marinas do Campeche, 11 - Campeche Central, Florianópolis - SC, 88065-185, Brazil	-27.680711	-48.5094613	RANGE_INTERPOLATED
15	Servidão Vila Harmonia, 35 - Campeche, Florianópolis - SC, 88063-500, Brazil	-27.6749882	-48.4847167	ROOFTOP
16	Servidão Olindina Maria Lopes, 538 - Morro das Pedras, Florianópolis - SC, 88066-028, Brazil	-27.701706	-48.503727	ROOFTOP
17	Servidão do Cravo Branco, 1 - Rio Tavares Central, Florianópolis - SC, 88063-522, Brazil	-27.6663135	-48.484505	RANGE_INTERPOLATED
18	Rod. Aparício Ramos Cordeiro, 111 - Autódromo, Florianópolis - SC, 88065-500, Brazil	-27.6824812	-48.5113251	RANGE_INTERPOLATED
19	Rod. SC 407, nº2546 - Bairro Deltaville, Biguaçu - SC, 88164-183, Brazil	-27.4896628	-48.6663754	ROOFTOP
20	R. das Gardenias, 20, Palhoça - SC, Brazil	-27.6175919	-48.6840815	ROOFTOP
21	R. das Gardenias, 20, Palhoça - SC, Brazil	-27.6175919	-48.6840815	ROOFTOP
22	R. José Bonifácio, 48 - Estreito, Florianópolis - SC, 88070-770, Brazil	-27.5891313	-48.5858161	ROOFTOP
23	Av. Afonso Delambert Neto, 740 - 03 - Lagoa da Conceição, Florianópolis - SC, 88062-000, Brazil	-27.6057794	-48.4680414	ROOFTOP
24	Rod. Balçicero Filomeno, 3633 - Ribeirão da Ilha, Florianópolis - SC, 88064-001, Brazil	-27.7006244	-48.5413806	ROOFTOP
25	R. do Calafate, 12 - Pantanal, Florianópolis - SC, 88040-008, Brazil	-27.6115854	-48.5245604	ROOFTOP
26	R. Oláyo Bilac, 251 - Jardim Eldorado, Palhoça - SC, 88133-350, Brazil	-27.630291	-48.6639839	RANGE_INTERPOLATED
27	R. Oláyo Bilac, 251 - Jardim Eldorado, Palhoça - SC, 88133-350, Brazil	-27.630291	-48.6639839	RANGE_INTERPOLATED
28	Av. Deltaville, 100 - Bairro Deltaville, Biguaçu - SC, 88160-000, Brazil	-27.4868021	-48.6691497	RANGE_INTERPOLATED
29	Av. Wilson Castelo Branco, 800 - Bairro Deltaville, Biguaçu - SC, 88160-000, Brazil	-27.4878582	-48.6731493	RANGE_INTERPOLATED
30	R. da Escova de Garrafa, 571 - Cachoeira do Bom Jesus, Florianópolis - SC, 88056-363, Brazil	-27.4202274	-48.4249933	RANGE_INTERPOLATED
31	Av. Wilson Castelo Branco, 800 - Bairro Deltaville, Biguaçu - SC, 88160-000, Brazil	-27.4878582	-48.6731493	RANGE_INTERPOLATED
32	Jureré, Florianópolis - State of Santa Catarina, Brazil	-27.4535435	-48.5112213	APPROXIMATE
33	Av. Red Park, 230, Florianópolis - SC, 88060-223, Brazil	-27.4752863	-48.4089274	ROOFTOP
34	R. Maestro Vila Lobos, 11 - João Paulo, Florianópolis - SC, 88030, Brazil	-27.5676133	-48.5093897	RANGE_INTERPOLATED

Fonte: Elaborado pelo autor.

Os campos retornados pela API são importantes informações para validação da precisão de georreferenciamento dos eventos, e deste modo fizeram parte do processo de revisão de eventos do mercado, realizado com apoio dos técnicos do IPUF envolvidos nas atividades do observatório do mercado imobiliário.

A coleta automatizada de eventos do mercado imobiliário totalizou 61.365 registros de anúncios, correspondentes aos diferentes segmentos do mercado. A Tabela 2 apresenta quantidades e porcentagens obtidas para cada tipo de classe de precisão da API de Geocodificação:

Tabela 2 – Distribuição dos anúncios por tipo de precisão do endereço.

Tipo de precisão	Quantidade de eventos	Porcentagem
ROOFTOP	24.430	39,83%
RANGE_INTERPOLATED	10.393	16,93%
GEOMETRIC_CENTER	22.898	37,32%
APPROXIMATE	3.260	5,32%
PRECISION_NONE	384	0,62%
Total	61.365	100%

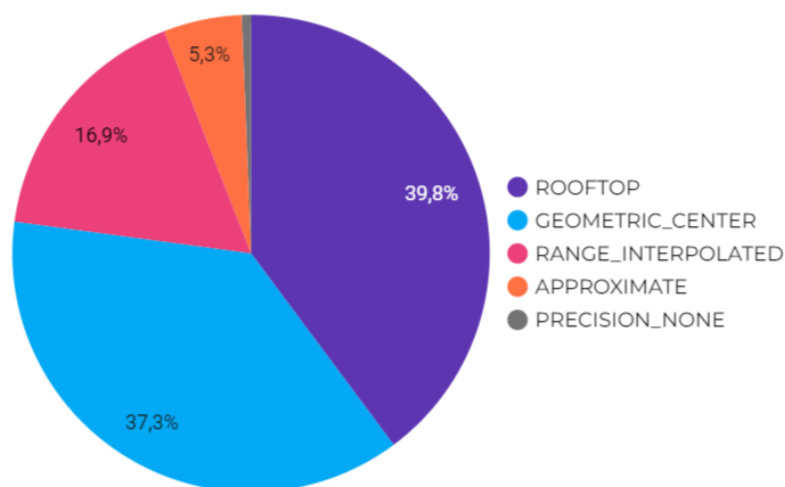
Fonte: Elaborado pelo autor.

Observa-se pelos números apresentados na Tabela 2 que os tipos de precisão que tiveram maior resultado foram ROOFTOP e GEOMETRIC_CENTER, respectivamente. Esta é uma informação importante para a análise da qualidade posicional dos dados de

anúncios imobiliários, assim como a qualidade da base de endereçamento do Google. Como o ROOFTOP é uma precisão que depende obrigatoriamente de um número de porta reconhecido pela inteligência de reconhecimento visual do Google Street View, e que este posiciona o ponto justamente sobre a localização do número de porta, é possível afirmar que 39,83% dos eventos de mercado tiveram um ponto georreferenciado com uma precisão ao nível de imóvel. Ainda assim, os técnicos municipais foram orientados a analisar cada registro para validar se a informação estava realmente precisa. De modo geral, pode-se afirmar que os tipos de precisão tiveram uma distribuição relativamente equilibrada nas diferentes classes, conforme demonstra a Figura 11.

De modo geral, pode-se afirmar que os tipos de precisão tiveram uma distribuição relativamente equilibrada nas diferentes classes, conforme demonstra a Figura 11.

Figura 11 – Gráfico representando a porcentagem de eventos por cada tipo de precisão da API de geocodificação.



Fonte: Elaborado pelo autor.

Cabe ressaltar que o trabalho desenvolvido estabeleceu como foco das análises o segmento de terrenos (vazios). Com isso, fez-se a análise da distribuição das precisões especificamente para os eventos deste segmento. Nota-se pela Figura 12 que a distribuição das precisões não é semelhante ao comportamento de todo o conjunto de dados. Há maior imprecisão na precisão das localizações. Pode-se afirmar que este

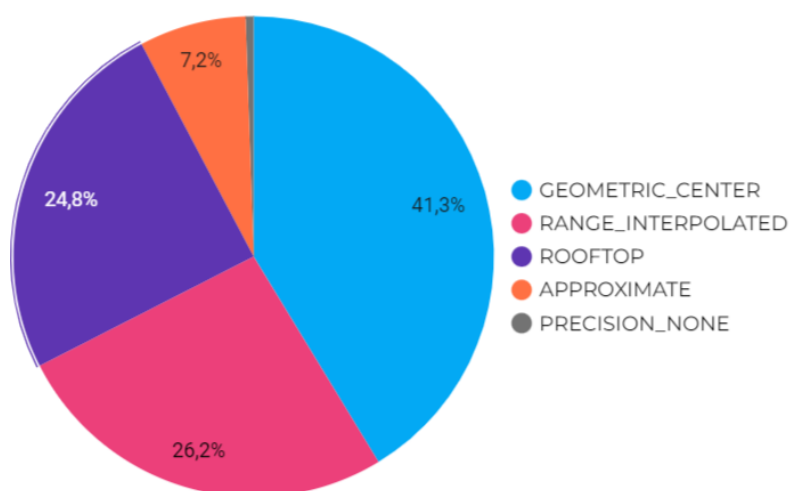
comportamento se deve ao fato de que os terrenos vazios não possuem endereçamento completo, ou seja, não possuem um número de porta. A Tabela 3 apresenta esta distribuição por tipo de precisão para os imóveis de tipologia “terreno”.

Tabela 3 – Distribuição dos anúncios por tipo de precisão do endereço para terrenos.

Tipo de precisão	Quantidade de eventos	Porcentagem
ROOFTOP	1012	24.75%
RANGE_INTERPOLATED	1070	26.25%
GEOMETRIC_CENTER	1687	41.28%
APPROXIMATE	294	7.19%
Precision_NONE	22	0.54%
Total	4085	100.00%

Fonte: Elaborado pelo autor.

Figura 12 – Gráfico representando a porcentagem de eventos por cada tipo de precisão da API de geocodificação.



Fonte: Elaborado pelo autor.

A revisão do conjunto de dados levantado possibilitou constatar que a precisão da localização de terrenos é prejudicada pela falta de números de porta físicos em suas fachadas, o que dificulta a codificação desses endereços. Como o Google utiliza tecnologia de mapeamento terrestre em 360 graus como um dos principais recursos para a construção de sua base de endereçamento, o georreferenciamento através do número de

porta em terrenos torna-se um desafio para os terrenos baldios. Esta menor precisão pode ser evidenciada na classe “ROOFTOP”, que reduziu de 39,83% (conjunto total) para 24.75% (conjunto dos terrenos).

4.4 - CONVERGÊNCIA DE DENOMINAÇÃO (DE - PARA) E MODELAGEM CONCEITUAL

Nesta etapa do estudo o objetivo foi comparar os dados coletados de eventos de mercado com o cadastro imobiliário do município. Para isso, foi avaliada tanto a tabela de eventos de mercado quanto a tabela do cadastro imobiliário, com relação a seus campos e valores, dando-se assim o início ao processo de convergência de denominação entre as tabelas.

Optou-se por utilizar as nomenclaturas dos campos e valores do cadastro como o padrão a ser seguido, uma vez que o objetivo do observatório é justamente dar apoio ao cadastro municipal. Dessa forma, foram realizados ajustes na tabela de eventos de mercado para adequá-los à base de dados do cadastro.

Por meio de reuniões realizadas com o grupo técnico do IPUF no mês de agosto de 2021, foi elaborada uma planilha Excel com a finalidade de mapear os diversos campos presentes na tabela de eventos de mercado, bem como descrevê-los, definir o tipo de campo correspondente (numérico, lista suspensa, entre outros) e adequá-los de acordo com os campos e variáveis do sistema cadastral em vigor.

A Figura 13 exemplifica como este processo foi realizado.

Figura 13 – Figura que apresenta a avaliação dos campos de eventos de mercado com relação aos dados cadastrais, aplicando-se convergência de denominação e outros ajustes tipológicos.

A	B	C	D	E	F	G	H	I	J
Create table	nome	descrição	preenchimento	correlação	nome no bd	tipo	obrigatôr	lista suspensa	TEM NO
mmlink SERIAL	mmlink		automático		mmlink	SERIAL			
inscricao VARCHAR (17)	inscricao imobiliária	incricão completa,	manual		inscricao	VARCHAR (17)			sim
tp_fonte TEXT	tipo de fonte	qual o tipo da fonte	manual lista suspensa		tp_fonte	TEXT	mobiliariacretor autônomo	cartórioT	-
fonte TEXT	nome da fonte		automático	advertiser	fonte	TEXT			
phone TEXT	telefone da fonte	telefone da fonte	automático	phone	phone	TEXT			
link TEXT	link da fonte	link da fonte	automático	link	link	TEXT			
external_id TEXT	código do anúncio		automático	external_id	external_id	TEXT			
origin_id TEXT	anunciante	código do anúncio na		origin_id	origin_id	TEXT			
situacao TEXT	situação na quadra	situação do imóvel na	manual lista suspensa		situacao	TEXT			sim ->
description TEXT	descrição	é a descrição do	automático	description	description	TEXT			não
title TEXT	título	título do anúncio do	automático	title	title	TEXT			
street TEXT	logradouro	logradouro principal	automático	street	street	TEXT			sim
street_number NUMERIC	número	número de porta	automático	street_number	street_number	NUMERIC			
zip_code NUMERIC	CEP	CEP	automático	zip_code	zip_code	NUMERIC			
bairro TEXT	bairro	bairro em que está	automático		bairro	TEXT			sim
cidade TEXT	cidade	cidade em que está			cidade	TEXT			
testada NUMERIC	Testada principal	testada principal do	manual		testada	NUMERIC			sim
usage_type TEXT	tipo de uso		automático	usage_type	usage_type	TEXT			
tipo TEXT	tipo imóvel	tipo do imóvel, se é	manual lista suspensa	type	tipo	TEXT		predialterritorial	é construído ->
caract TEXT	característica	se é único,	manual lista suspensa	caract	caract	TEXT	únicocond. horizontalcond. vertical		
tipologia TEXT	tipologia	se o imóvel é casa,	automático com relacionament	unit_type	tipologia	TEXT	ventosalagalpãotelheirogaragemiscir		sim - C
paredes TEXT	paredes	tipo de paredes do	manual lista suspensa		paredes	TEXT	alvenariamadeiraconcretovidro		sim - S
padrao TEXT	padrão	qual o padrão do	manual lista suspensa		padrao	TEXT	atomédico altomédio baixo		sim - L
conserv TEXT	conservação	estado de	manual lista suspensa		conserv	TEXT	ótombonregulapéssimoruf		não
n_pavim NUMERIC	Número de pavimento	quantidade de	manual		n_pavim	NUMERIC			
pavimento NUMERIC	Pavimento no imóvel	em que pavimento o	manual		pavimento	NUMERIC			sim
area_size NUMERIC	área do anúncio	área do anúncio	automático		area_size	NUMERIC			
area_terreno NUMERIC	Área do terreno	área do terreno	manual		area_terreno	NUMERIC			sim
area_constr NUMERIC	Área total construída	área total construída	manual		area_constr	NUMERIC			sim
t_uni_lote NUMERIC	Totál unidades lote	numero total de			t_uni_lote	NUMERIC			sim
area_unidade NUMERIC	Área da unidade (privativa)	área da unidade (área	manual		area_unidade	NUMERIC			
area_tot NUMERIC	Área total (comum + privativa)	área total (área	manual		area_tot	NUMERIC			
condominio_preco NUMERIC	valor do condomínio	valor do condomínio			condominio_preco	NUMERIC			
point_precision TEXT	localidade da localização do anúncio	localidade da localização do a	automático	point_precision	condominio_preco	TEXT	im localizaçãobairrodistrictoloteamentologradourolote		
qualidade TEXT	Qualidade da localização		manual lista suspensa		qualidade	TEXT			
validacao TEXT	Validação		manual lista suspensa		validacao	TEXT		nãosimdescartado	
obs TEXT	Observação da validação		manual		obs	TEXT			
alug_venda TEXT	Aluguel ou venda	valor está relacionado	automático	business_type		TEXT			
price NUMERIC	valor	valor total do imóvel	automático	price	price	NUMERIC			
iptu_price NUMERIC	valor do IPTU	valor do IPTU do	automático	iptu_price	iptu_price	NUMERIC			
valor_m2 NUMERIC	valor do metro quadrado	valor por metro	automático		valor_m2	NUMERIC			
bedrooms NUMERIC	quartos	número de quartos do	automático	bedrooms	bedrooms	NUMERIC			
bathrooms NUMERIC	banheiros	número de banheiros	automático	bathrooms	bathrooms	NUMERIC			
suite NUMERIC	suites	número de suites do	automático	suite	suite	NUMERIC			
parking NUMERIC	vagas de garagem	número de vagas de	automático	parking	parking	NUMERIC			
data_evento DATE	data do evento		automático		data_evento	DATE			
data_atualizacao DATE	data da última atualização	registra através de	automático trigger	-	data_atualizacao	DATE			
usuario TEXT	usuário que editou a última atualiz	registra através de	automático trigger	-	usuario	TEXT			
geom GEOMETRY (POINT, 31982)	geometria	geometria ponto	-	-	geom	GEOMETRY (POINT, 31982)			

Fonte: Elaborado pelo autor.

A partir da definição junto a equipe técnica, chegou-se no consenso de nomenclatura e definição dos campos, conforme Quadro 5:

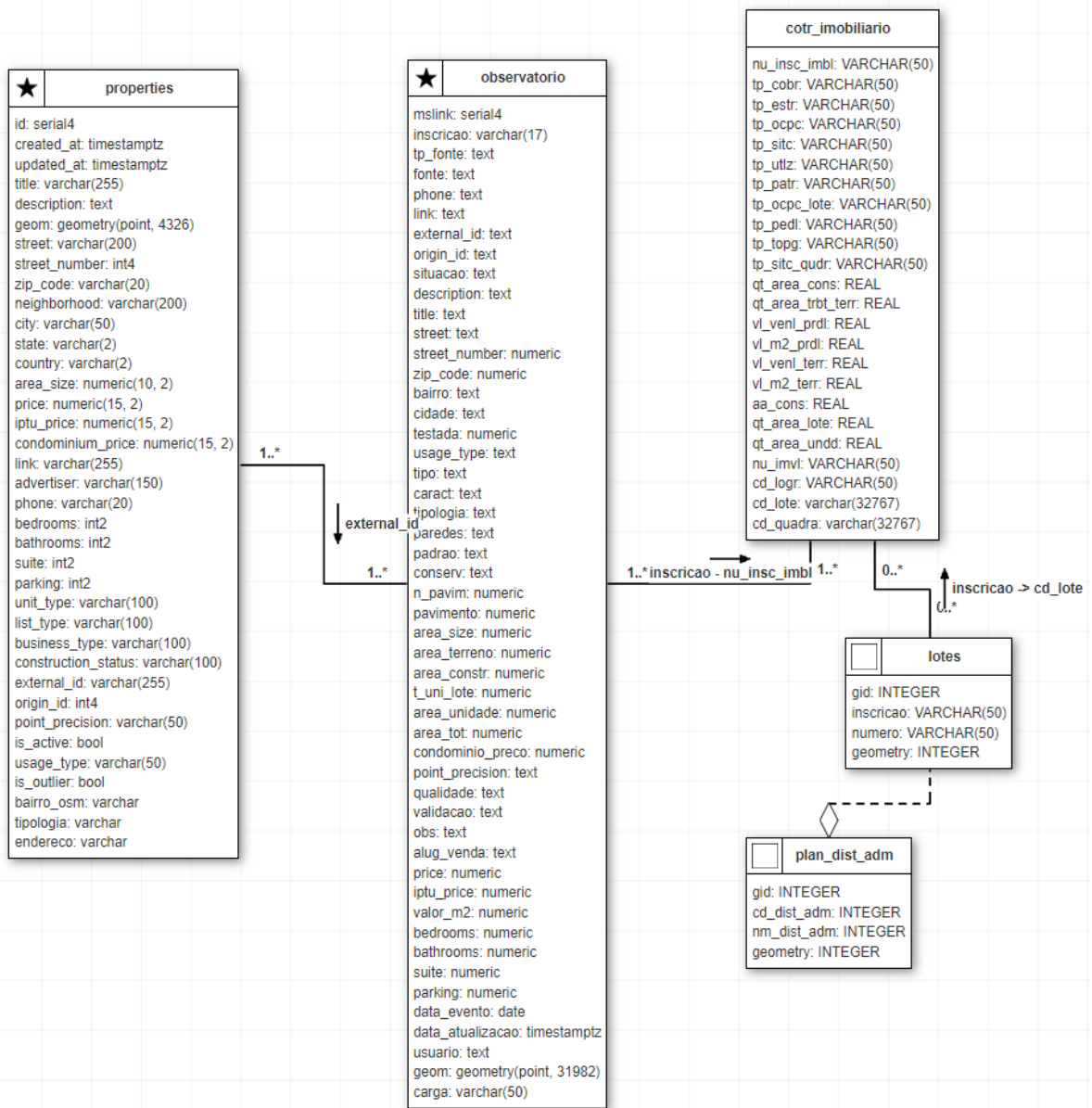
Quadro 5 – Campos definidos para a tabela de observatório.

Campo	Definição
mmlink	Número serial único para cada registro na tabela
inscricao	Número de identificação imobiliária com 17 dígitos
tp_fonte	Tipo de fonte usada para obter as informações do imóvel
fonte	Nome da fonte onde a informação do imóvel foi obtida
phone	Número de telefone da fonte
link	Link da fonte onde a informação do imóvel foi obtida
external_id	Código do anúncio na fonte
origin_id	Código do anunciante na origem
situacao	Situação do imóvel na quadra
title	Título do anúncio do imóvel
street	Logradouro principal do imóvel
street_number	Número da porta do imóvel
zip_code	CEP do imóvel
bairro	Bairro onde o imóvel está localizado
cidade	Cidade onde o imóvel está localizado
testada	Comprimento da testada principal do imóvel
usage_type	Tipo de uso do imóvel
tipo	Tipo do imóvel (predial ou territorial)
caract	Característica do imóvel
tipologia	Tipologia do imóvel
paredes	Tipo de parede do imóvel
padrao	Padrão do imóvel
conserv	Estado de conservação do imóvel
n_pavim	Número de pavimentos do imóvel
mmlink	Número serial único para cada registro na tabela
inscricao	Número de identificação imobiliária com 17 dígitos
tp_fonte	Tipo de fonte usada para obter as informações do imóvel
fonte	Nome da fonte onde a informação do imóvel foi obtida
phone	Número de telefone da fonte

Fonte: Elaborado pelo autor.

Após a definição dos campos e valores a serem utilizadas, desenvolveu-se a modelagem conceitual da base de dados utilizando o método OMT-G, de forma a planejar e avaliar o correto relacionamento entre os dados correspondentes. A Figura 14 apresenta a respectiva modelagem.

Figura 14 – Modelagem conceitual da proposta de um observatório do mercado imobiliário com os diferentes relacionamentos.



Fonte: Elaborado pelo autor.

4.5 OPERAÇÕES DE ARMAZENAMENTO E TRATAMENTO DOS DADOS

Nesta implementação, criou-se a tabela com anúncios originais intitulada “properties”, que contemplou todas as informações brutas da coleta de dados, sem as devidas adequações e correções, preservando a origem da informação. Outra tabela foi criada com as mesmas informações intitulada “observatório”, mas para a devida edição e correção dos eventos do mercado. Esta tabela segue as definições realizadas no processo de modelagem conceitual e convergência de denominação (subitem 4.3), sendo a

nomenclatura dos campos e atributos padronizada de acordo com a estrutura da base cadastral municipal.

A tabela disponibilizada para qualificação e validação dos dados possibilitou que os usuários realizassem a edição, usando, principalmente, software QGIS para corrigir as inconsistências alfanuméricas e geométricas. Também foi possível adicionar novos dados de mercado. Para analisar e monitorar estas edições de tabelas pelos diferentes usuários, foram criados mecanismos de auditoria. O primeiro utilizado compôs-se de funções (*functions*) para atualizar os campos de data de alteração e usuário de edição, a partir de gatilhos (*triggers*) do banco de dados. Esta função é acionada pelo gatilho após a alteração de qualquer um dos valores das colunas de uma tabela, incluindo sua componente espacial.

Outro mecanismo utilizado foi a criação de uma tabela de auditoria, a partir de uma função do banco de dados - também acionada por um gatilho - que monitora todas as ações realizadas na tabela alvo, que neste caso é a tabela de edição de anúncios imobiliários. Diferente do primeiro mecanismo, que apenas atualiza informações diretamente na tabela de edição, esse armazena em uma nova tabela os dados manipulados. Para toda alteração, uma nova linha foi inserida nesta tabela de auditoria com uma cópia dos dados da tabela alvo, e um comentário sobre qual ação foi realizada. Por exemplo, caso um anúncio seja removido da tabela de edição de anúncios imobiliários por um usuário, uma nova linha será inserida na tabela de auditoria contemplando todos os dados daquele anúncio, além de um comentário de que ele foi removido.

O script de ambos mecanismos pode ser consultado no **APÊNDICE I** desta pesquisa.

4.5.1 Mecanismos para depuração dos dados de eventos de mercado

A depuração de dados de eventos de mercado envolveu a utilização de uma variedade de técnicas e ferramentas com o objetivo de alcançar maior consistência, precisão e completude dos dados, antecedendo uma análise e validação visual a ser realizada pelos profissionais envolvidos, a ser apresentada no subitem 4.5.2.

Um dos procedimentos realizados foi a identificação de *outliers* (casos atípicos). Foram criados *scripts sql* para filtragem desses casos, e posteriormente a correção individual de forma visual. O procedimento envolveu a filtragem de casos em que os valores observados estavam muito acima ou abaixo dos valores médios, geralmente

resultado de erros de digitação dos anunciantes dos portais imobiliários e/ou erros na importação da base de dados de imobiliárias para os portais. Para esta atividade, utilizou-se o número de desvios padrão, que indica o quanto os valores de um conjunto de dados estão dispersos em relação à média. Definiu-se três principais campos para essa análise: *price* (preço), *area_size* (área) e o *valor_m2* (valor de metro quadrado).

Para criação da seleção de desvio padrão utilizou-se a função nativa do PostgreSQL, denominada STDDEV. Aplicou-se o valor de 2 desvios padrões acima ou abaixo da média para avaliar possíveis outliers, conforme convenção de análise estatística. Definiu-se grupos de amostras para avaliação, filtrando o tipo de transação e o tipo de imóvel.

No Quadro 6 é apresentado o script utilizado para filtrar os dados de eventos do mercado de acordo com a tipologia de "terreno" e o tipo de transação de "venda". Além disso, foi incluído o campo "link" para permitir a avaliação visual direta da fonte (portal imobiliário), com o objetivo de constatar que o possível erro está na origem do dado, e não em alguma falha no procedimento de coleta automatizada dos dados. Outro ponto favorável da observação dos anúncios diretamente no portal foi poder-se analisar a descrição do anúncio, que por vezes é complementada por alguma informação específica de área, que pode revelar o motivo das discrepâncias de valores. Por exemplo, é possível que um anunciante declare mais de um terreno dentro de um anúncio, e declare o preço de apenas um deles, gerando uma inconsistência na relação preço x área.

Quadro 6 – Scripts SQL para filtragem de possíveis outliers utilizando a análise de desvio padrão.

```
-- DESVIO PADRÃO DOS VALORES NEGOCIADOS -
SELECT price, area_size, valor_m2, link FROM
observatorio.observatorio
WHERE price > (SELECT AVG(price) + (2 * STDDEV(price)) FROM
observatorio.observatorio
where tipologia = 'terreno' and alug_venda = 'venda'););
-- DESVIO PADRÃO DOS VALORES DE M2 NEGOCIADOS -
SELECT price, area_size, valor_m2, link FROM
observatorio.observatorio
WHERE valor_m2 > (SELECT AVG(valor_m2) + (2 * STDDEV(valor_m2)) FROM
observatorio.observatorio
where tipologia = 'terreno' and alug_venda = 'venda' );
-- DESVIO PADRÃO DAS ÁREAS NEGOCIADAS -
SELECT price, area_size, valor_m2, link FROM
observatorio.observatorio
WHERE area_size > (SELECT AVG(area_size) + (2 * STDDEV(area_size))
FROM observatorio.observatorio
where tipologia = 'terreno' and alug_venda = 'venda' ););
```

Fonte: Elaborado pelo autor.

A Figura 15 exemplifica os possíveis outliers no campo “valor_m2”. Os casos apresentados foram filtrados pela análise de desvio padrão, e evidenciam valores discrepantes da realidade do município. Um dos exemplos apresenta um terreno sendo vendido no município por 980 milhões de reais, tendo uma superfície de apenas 151 m². Mesmo com a especulação imobiliária presente no município de Florianópolis, esse dado é completamente fora da realidade do mercado.

Figura 15 – Exemplo de Outliers avaliados, a partir de scripts sqls sobre a massa de dados.

Grid	price	area_size	valor_m2	link
1	15,500,000	0	[NULL]	https://www.vi
2	40,000,000	[NULL]	[NULL]	https://www.vi
3	75,000,000	0	[NULL]	https://www.vi
4	935,307,000	140	6,680,764.29	https://www.vi
5	980,000,000	151	6,490,066.23	https://www.vi
6	111,111,111	52	2,136,752.13	https://www.vi
7	416,000,000	520	800,000	https://www.vi
8	39,900,001	106	376,415.1	https://www.vi
9	44,999,990	609	73,891.61	https://www.vi
10	18,000,000	350	51,428.57	https://www.vi
11	16,000,000	414	38,647.34	https://www.vi
12	14,421,000	385	37,457.14	https://www.vi

Fonte: Elaborada pelo autor.

Outro procedimento realizado para garantir maior consistência dos dados foi a higienização e tratamento dos dados, utilizando scripts SQL para avaliação em massa. Entre os scripts aplicados sobre a massa de dados, estão:

- Scripts de exclusão de valores nulos nos campos de interesse;
- Script de expressão regular para alterar campos numéricos de vírgula para ponto nas casas decimais. Utilizou-se a expressão **replace** (“campo de interesse”, ‘,’, ‘.’) de forma a normalizar todas as variáveis de um campo, mesmo que existissem alguns eventos com padrão de casas decimais distintos;
- Script de expressão regular para remover caracteres especiais. Utilizou-se a expressão **regexp_replace**(“campo de interesse”, ‘[^0-9]’, ‘’, ‘g’) para

remover qualquer caractere especial diferente de números em campos teoricamente numéricos;

- Script para alterar o tipo de formatação de campo, quando necessário. Campos do tipo texto foram alterados para numéricos quando a informação de origem tinha somente dados numéricos.

4.5.1.2 - Análise do nível de georreferenciamento dos eventos do mercado

A partir resultados obtidos na geocodificação de endereços, iniciou-se o trabalho de avaliação do nível de georreferenciamento dos eventos de mercado, e consequentemente a correção destes. Para esse processo, houve apoio do grupo técnico do IPUF na validação do posicionamento dos pontos, com a utilização do Software QGIS.

Foi elaborado um mecanismo de validação dos dados, a partir da interpretação do nível de georreferenciamento gerado pelo processo de geocodificação. A avaliação foi realizada utilizando a ordem dos mais precisos para os menos precisos. Iniciou-se com os registros de eventos classificados como “ROOFTOP” e na sequência o “RANGE INTERPOLATED”.

Conforme demonstra a Figura 16, o campo “qualidade” foi criado para definição da precisão de georreferenciamento do evento de mercado a partir da análise manual do grupo técnico. O campo “validação” foi criado para controle de pontos já revisados, o campo “obs” foi criado para observações referentes a revisão manual dos eventos, sejam aquelas relacionadas ao controle de georreferenciamento, ou dos outros campos preenchidos.

Figura 16 – Exemplo da tabela do banco de dados, representando os campos que foram utilizados para controle de georreferenciamento dos eventos de mercado.

point_precision	qualidade	validacao	obs
ROOFTOP	1	1	aérea do geo é diferente da do anúncio, área do geo 419
ROOFTOP	1	1	[NULL]
ROOFTOP	1	1	aérea do geo é diferente da do anúncio, área do geo 372.41
ROOFTOP	1	1	A aérea do geo é diferente da do anúncio, área do geo 16200
ROOFTOP	3	1	
ROOFTOP	4	1	[NULL]
ROOFTOP	1	1	[NULL]
ROOFTOP	1	1	A aérea do geo é diferente da do anúncio, área do geo 20000
ROOFTOP	2	1	Area no cadastro de 750m ²
ROOFTOP	3	1	Área que consta no cadastro é em conjunto com a área comum do condomínio sendo
ROOFTOP	3	1	Área que consta no cadastro é em conjunto com a área comum do condomínio sendo

Fonte: Elaborado pelo autor.

Foram definidas as seguintes classes de “qualidade” posicional após validação:

1. sem localização
2. bairro/distrito
3. loteamento
4. logradouro
5. lote

Na Figura 16 pode-se constatar que no exemplo apresentado estão os eventos classificados na situação `point_precision = “ROOFTOP”`, ou seja, a melhor classe de precisão de geocodificação. Contudo, nota-se que não necessariamente isso reflete em um georreferenciamento preciso ao nível de lote, já que durante a revisão dos técnicos foi constatado que nem todos os registros se encontravam em situação de “qualidade” = 5 (lote), ou seja, no nível de precisão do lote.

As principais causas para a inconsistência da qualidade posicional da geocodificação “ROOFTOP” são:

1. **Imprecisões entre as bases cartográficas:** estas imprecisões ocorrem principalmente com relação a escala dos produtos e posicionamento. A base de lotes cadastrais do município possui diferentes produtos cartográficos, que são atualizados com o tempo à medida que ocorrem os parcelamentos do solo no município, assim como as atualizações cartográficas cadastrais em massa. Essa base pode conter deslocamentos e imprecisões geométricas nas diferentes áreas do território, que quando relacionadas com a base de geocodificação do Google podem não realizar o cruzamento espacial de intersecção. A mesma imprecisão cartográfica ocorre na base cartográfica da geocodificação, que da mesma forma não favorece o cruzamento espacial, mesmo que em princípio os dados estejam apontando para o mesmo imóvel. Na Figura 17 são apresentados dois exemplos de eventos de mercado com nível de precisão “ROOFTOP”, mas que pelas diferenças cartográficas não se relacionaram espacialmente.

Figura 17 – Exemplo de pontos geocodificados com precisão “ROOFTOP” (classe mais precisa a nível de porta), mas que não cruzavam com lotes cadastrais.



Fonte: Elaborado pelo autor.

2. **Posicionamento do ponto “ROOFTOP”:** O georreferenciamento dos pontos "ROOFTOP" é realizado, em parte, utilizando técnicas de reconhecimento de imagens terrestres do Google Street View. Nesse processo, ocorre a triangulação a partir de diferentes imagens terrestres, para posicionar o ponto exatamente no número de porta físico do imóvel. Esse procedimento garante alta precisão no georreferenciamento do imóvel, mas geralmente o ponto é posicionado na frente do imóvel, onde normalmente está o número de porta físico. Esse fato pode prejudicar o relacionamento espacial do imóvel com a base cadastral de lotes, uma vez que, por vezes, os limites cartográficos dos lotes, mesmo que com alto nível de precisão, não coincidem com a situação de fato. Na Figura 18 exemplifica-se esta situação, demonstrando o alinhamento do muro físico e a divergência com a extremidade do lote cadastral.

Figura 18 – Mosaico demonstrando situação em que um evento de mercado “rooftop” está posicionado sobre o alinhamento do muro físico. Esse detalhe é evidenciado tanto em vista aérea quanto em vista terrestre.



Fonte: Elaborado pelo autor.

A análise dessas divergências evidenciou a importância da validação humana sobre os pontos georreferenciados a partir de mecanismos de geocodificação, mesmo que sejam considerados de alto nível de precisão pela API. Existem alternativas para viabilizar o cruzamento espacial de forma automática nestes casos citados, como a criação de um script espacial por distância ao lote mais próximo do ponto de evento de mercado, mas ainda sim pode acarretar em relacionamento incorreto.

4.5.2 Validação e qualificação de dados pela equipe técnica municipal

Esta etapa da pesquisa teve grande importância, tanto para o avanço da presente pesquisa quanto para o início do desenvolvimento do observatório no Município de Florianópolis. O envolvimento da equipe técnica municipal é recomendado para a

manutenção de um observatório. De acordo com De Cesare *et al.* (2023), "a manutenção do valor dos imóveis atualizados no cadastro territorial dos municípios requer formação de uma equipe permanente de trabalho, composta por servidores habilitados e capacitados para a função" (p. 127). Devem ser utilizadas as alternativas legais para a contratação de serviços, e que a capacitação do corpo técnico municipal esteja inclusa na contratação dos serviços, bem como o acompanhamento da equipe de técnicos do município nas principais etapas do trabalho.

O processo de validação e qualificação dos dados, por parte da equipe técnica, sobre as informações levantadas pela coleta automatizada de eventos do mercado foi realizada nos seguintes campos:

inscricao: o campo de inscrição, ou seja, a correta associação entre eventos de mercado e as respectivas parcelas cadastrais foi realizada exclusivamente pela equipe técnica, tomando como base para avaliação o georreferenciamento por geocodificação; o texto descritivo dos anúncios; as áreas do evento do mercado com relação a área cadastral e as fotos associadas aos anúncios dos portais imobiliários;

situacao: a situação na quadra, que foi preenchida com os seguintes campos:

meio da quadra;

esquina/mais de uma frente

vila

cond. horizontal

encravado

gleba

aglomerado;

testada: Testada Principal – Informações foram preenchidas com base no cadastro, ou por fotointerpretação das imagens Google Satélite;

area: priorizou-se considerar a área do anúncio, mesmo que divergente ao cadastro. Mas foi avaliado casos atípicos e editados quando eram claramente erros de digitação;

qualidade: a qualidade posicional do evento de mercado, após a validação da equipe técnica. São eles:

1 - preciso, no lote

2 - no logradouro

3 - no loteamento / condomínio

4 - no bairro;

validação: o campo validação foi um importante campo de controle de eventos que já haviam sido avaliados, mesmo que não corrigidos ou enriquecidos com informação. Suas valores são:

01 – validado;

02 – validado mas não encontrado.

Nulo – não validado;

observacao: campo de observação, para armazenar qualquer descrição adicional sobre uma revisão de um evento de mercado.

A atividade visando avaliar o posicionamento dos eventos de mercado e associá-los as parcelas cadastrais, por meio da inserção da inscrição imobiliária no campo “**inscrição**”, foi uma etapa importante do trabalho e que possibilitou o desenvolvimento das outras etapas da pesquisa: análise de desempenho e do monitoramento de valores cadastrais. O enfoque da avaliação, tanto para fins de pesquisa quanto para os estágios iniciais de estabelecimento do observatório, concentrou-se nos eventos de mercado de tipologia “terrenos”. Os resultados obtidos nesse relacionamento são apresentados na Tabela 4.

Tabela 4 – Vinculação dos eventos de mercado com a parcela cadastral para cada tipologia, a partir do preenchimento das inscrições imobiliárias pela equipe técnica.

Tipologia	Quantidade de eventos vinculados	Total de eventos	Porcentagem relativa de vinculados
Galpão	1	83	0.14%
sala/loja	4	6159	0.06%
casa	108	13594	0.06%
terreno	463	4149	11.16%
apartamento	512	37348	1.37%
Total	1088	61359	1.77%

Fonte: Elaborado pelo autor.

Entre os diferentes grupos de dados apresentados na Tabela 3, categorizados de acordo com suas respectivas tipologias, destaca-se o grupo "terreno", tomado como segmento de mercado para as análises neste trabalho. Esse grupo recebeu prioridade no processo de validação, o qual foi realizado de forma completa até o momento do corte temporal realizado para a pesquisa. As outras tipologias não foram revisadas por completo, razão pela qual limitou-se as análises ao seguimento de terrenos. Essa

priorização mencionada deve-se ao fato de que os dados relacionados a terrenos possibilitam a elaboração da planta de valores genéricos, que é fundamental para a avaliação em massa de imóveis.

Nesta amostra específica de “terreno” pode-se observar que apenas 463 (11.16%) dos 4149 eventos de mercado do seguimento de terrenos foram devidamente vinculados a parcela cadastral pelos técnicos municipais. Ao resgatar os resultados obtidos no subitem 4.31, em que no processo de geocodificação de terrenos foram obtidos 1012 (24.75%) terrenos com precisão de número de porta (“ROOFTOP”), evidencia-se que, mesmo em casos em que a precisão seja máxima e o ponto georreferenciado posicionado sobre um imóvel, a análise de outras informações demonstrou a importância da revisão caso a caso do posicionamento dos eventos.

Entre os diferentes motivos da incorreta vinculação, um dos que mais se destaca são aqueles em que a área total declarada pelo anunciante é consideravelmente divergente da área total do terreno na base cadastral. Esses casos são muito comuns em áreas urbanas com forte dinâmica no uso e ocupação do solo, em que lotes são desmembrados em um curto espaço de tempo, e a atualização cadastral não acompanha tais processos, ou mesmo casos de venda de terrenos de posse, sem a devida declaração no sistema cadastral. Estes motivos acabam trazendo à tona a importância do Cadastro Territorial Multifinalitário (CTM), em que a parcela é o elemento base para o funcionamento de uma coerente administração do território.

4.6 RELACIONAMENTO COM DADOS CADASTRAIS E PROCEDIMENTOS PARA ANÁLISE DE DESEMPENHO

Para o relacionamento entre a tabela de eventos de mercado e a base cadastral do Município, foi criada uma *view* materializada no banco de dados, relacionando três tabelas. Para o relacionamento entre a tabela de eventos de mercado e a base cadastral do Município, foi criada uma *view* materializada no banco de dados relacionando três tabelas: tabela “observatorio”, “cotr_imobiliario” e “lotes”. Ao utilizar uma estrutura de *view* materializada em vez de uma tabela, é possível atualizar os dados conforme necessário, pois ela funciona como uma consulta SQL "materializada" em formato de tabela. Isso simplifica o processo de atualização dos eventos de mercado e dos dados cadastrais, evitando a necessidade da utilização de linguagens de programação de alto nível, como Python, Java, entre outras. Por fim, essas informações foram unidas

novamente com as camadas espaciais de distritos administrativo de Florianópolis, sendo armazenados na tabela "observatorio.plan_distrito_adm". Esta união permitiu a associação dos lotes de terra com seus respectivos distritos administrativos, considerando que os Distritos Administrativos são as divisões territoriais mais consolidadas na administração municipal.

Esta *view* foi elaborada para duas principais finalidades:

- a) publicar as informações do observatório que foram associadas à base cadastral do Município a partir do BI;
- b) utilizá-la como base para a filtragem de amostras específicas para a análise de desempenho.

Foram levantados os critérios necessários para uma avaliação de desempenho consistente, tendo como principal objetivo propor mecanismos para filtragem e validação dos dados a serem utilizados. Nesta etapa foram filtrados apenas imóveis territoriais (terrenos), que continham: dados de valor venal na tabela cadastral, dados de preço e área na tabela de eventos de mercado, e que não tivessem dados nulos. Outro filtro definido para a análise de desempenho foi o de no mínimo 5 casos por Distrito. Os Distritos de Ratoles, Pântano do Sul, Ribeirão da Ilha, Barra da Lagoa, São João do Rio Vermelho e Sede Continental foram suprimidos da análise por não se enquadrarem nestes critérios.

No sistema cadastral, um mesmo lote possui múltiplas unidades autônomas, assim como múltiplas unidades de avaliação. Portanto, ao avaliar um imóvel relativo a um anúncio do mercado imobiliário e compará-lo com um imóvel da base cadastral, também se certificou a correspondência. Para isso, uma estratégia adotada foi utilizar um fator de tolerância de variação de área. Esse fator consistiu na divisão da área do terreno cadastrado pela área do terreno do evento de mercado. Foi estabelecido que esse fator não poderia apresentar uma variação superior a 20% em relação à área do terreno de referência, nem para mais nem para menos.

Após os procedimentos de relacionamento, filtragem e validação dos eventos de mercado para análise de desempenho, a próxima etapa consistiu na criação das fórmulas, conforme apresentadas na Figura 5 desta pesquisa. Os scripts SQL, disponibilizados no **APÊNDICE II**, foram empregados para os seguintes cálculos:

- a) Cálculo da média aritmética por distrito;

- b) Cálculo da mediana por distrito;
- c) Cálculo da média ponderada por distrito;
- d) Cálculo do coeficiente de dispersão por distrito;
- e) Verificação da uniformidade por distrito;
- f) Verificação de regressividade ou progressividade por distrito (diferencial relacionado ao preço).

Os resultados obtidos para cada Distrito Administrativo são apresentados na Tabela 5:

Tabela 5 – Resultados dos cálculos obtidos na análise de desempenho para os Distritos Administrativos com no mínimo 5 casos.

Distrito	Número de eventos	Média Aritmética	Mediana	Média Ponderada	Coefficiente de Dispersão	Diferencial Relacional Preço
Cachoeira Do Bom Jesus	25	0.77	0.53	0.39	86.24	1.96
Campeche	10	1.00	0.48	0.42	148.29	2.36
Canasvieiras	11	0.99	0.86	1.09	52.74	0.91
Inglese Do Rio Vermelho	10	1.05	1.12	0.35	26.68	3.05
Lagoa Da Conceição	7	0.89	0.88	0.78	43.86	1.14
Santo Antônio De Lisboa	10	0.54	0.14	0.51	352.92	1.07
Sede Insular	66	0.51	0.51	0.43	54.86	1.16

Fonte: Elaborado pelo autor.

Os resultados obtidos na análise de desempenho foram comparados aos limites aceitáveis e recomendados pela IAAO. Destaca-se que a nova portaria das diretrizes nacionais do CTM (Portaria nº 3.242/2022) não define tais limites, diferentemente da portaria anterior (nº 509/2009). Tal mudança se deveu ao fato de que os resultados da análise de desempenho devem orientar com o tempo as melhorias nos modelos de avaliação em massa, sem necessariamente classificá-los como aceitáveis ou não. Para efeitos de exemplo, empregou-se os limites estabelecidos na Portaria nº 509/2009.

Para o nível de avaliação, que leva em conta as três primeiras valores (Média Aritmética, Mediana e Média Ponderada), os limites aceitáveis são:

- a) IAAO: 0,90 – 1,10;

b) BRASIL: 0,70 – 1,00.

Para a uniformidade, que avalia o Coeficiente de Dispersão (CD) e o Diferencial Relacionado ao Preço (DRP), os limites aceitáveis são:

a) CD:

- IAAO: < 10-15% residencial, ou < 15-20% demais;
- BRASIL: <= 30%.

b) DRP:

- IAAO: < 0,98 = progressividade, e > 1,03 = regressividade.

A Tabela 6 apresenta esta avaliação aplicada para cada conjunto de dados, fazendo referência aos dados da Tabela 5. Valores dentro do limite aceitável estão representados como Sim, enquanto valores fora do limite aceitável como Não.

Tabela 6 – Avaliação dos Valores em relação aos Limites Aceitáveis (IAAO e Brasil, 2009).

Distrito	Média Aritmética	Mediana	Média Ponderada	Coeficiente de Dispersão	Diferencial Relacional Preço
Cachoeira Do Bom Jesus	Não (IAAO) Sim (BRASIL)	Sim (IAAO) Sim (BRASIL)	Não (IAAO) Sim (BRASIL)	Não	regressividade
Campeche	Sim (IAAO) Sim (BRASIL)	Não	Não	Não	regressividade
Canasvieiras	Sim (IAAO) Sim (BRASIL)	Não (IAAO) Sim (BRASIL)	Sim (IAAO) Não (BRASIL)	Não	progressividade
Inglêses Do Rio Vermelho	Sim (IAAO) Não (BRASIL)	Não	Não	Não	regressividade
Lagoa Da Conceição	Não (IAAO) Sim (BRASIL)	Não (IAAO) Sim (BRASIL)	Não (IAAO) Sim (BRASIL)	Não	regressividade
Santo Antônio De Lisboa	Não	Não	Não	Não	regressividade
Sede Insular	Não	Não	Não	Não	regressividade

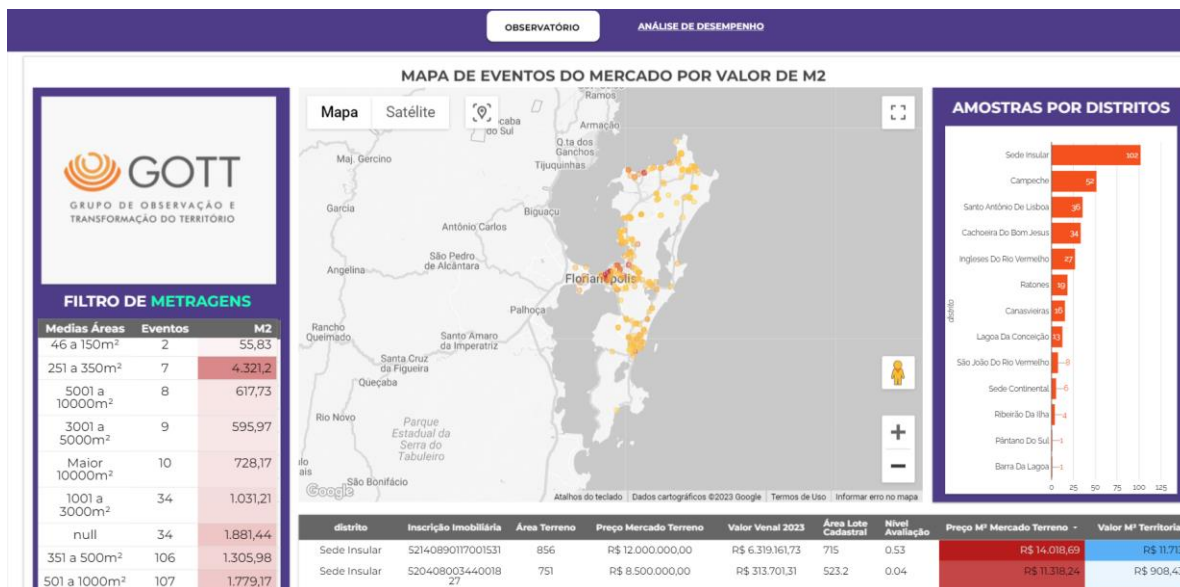
Fonte: Elaborado pelo autor.

4.7- VISUALIZADOR DE DADOS EM BI

De forma a complementar os resultados da pesquisa, propôs-se a elaboração de um painel de visualização de dados do observatório em um sistema de *Business*

Intelligence (BI). Com este painel, objetivou-se disponibilizar um mecanismo aos municípios para realizar e acompanhar a análise de desempenho, bem como divulgar as informações coletadas e relacionadas a todos os atores envolvidos na implementação prática do observatório no município.

Figura 19 – Painel inicial de visualização do BI criado para o observatório do mercado imobiliário para o Município de Florianópolis.



Fonte: Elaborado pelo autor.

A tela inicial, correspondente a Figura 19, apresenta a visualização espacial dos eventos de mercado do segmento de terrenos baldios. São ao todo 463 imóveis, conforme apresentado na Tabela 3. Por esta visualização é possível perceber a distribuição espacial das ofertas de terrenos no município.

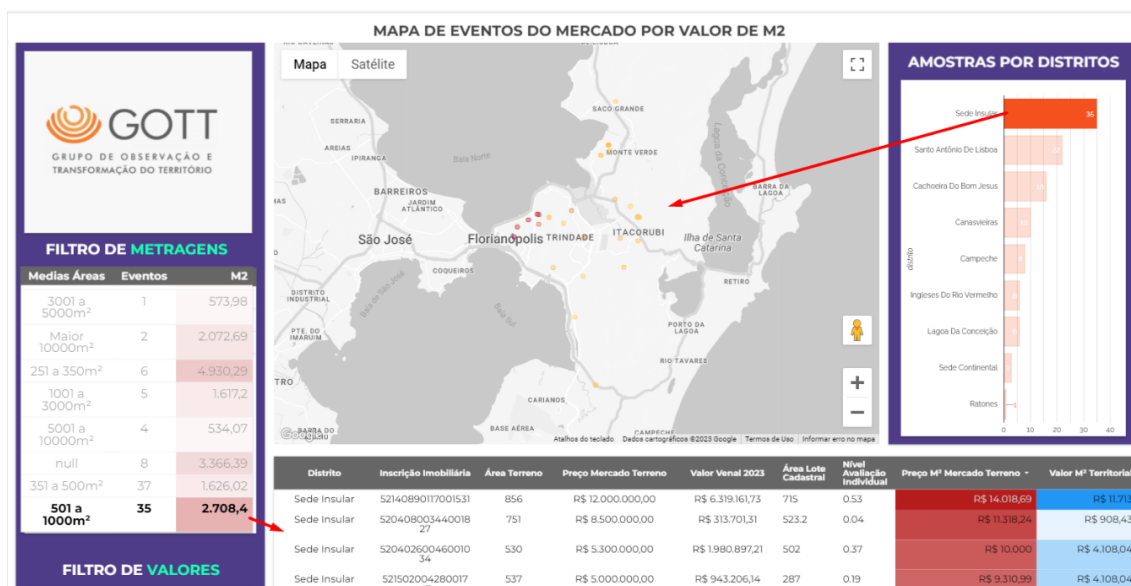
No lado esquerdo da tela inicial foi criado um quadro resumo que ilustra as médias das áreas dos eventos de mercado por intervalo de áreas. Cada linha do quadro representa um intervalo de classe, exibindo a quantidade de eventos e a média das áreas correspondente a esse grupo específico. Destaca-se que todos os gráficos do painel possuem filtros, que ao serem modificados refletem nas demais visualizações gráficas do painel. Abaixo deste gráfico, foi criado outro com a mesma estratégia de intervalos, porém com os preços praticados na venda do imóvel, os eventos dentro desses intervalos e a média do m².

No lado direito do painel pode-se visualizar o gráfico que apresenta o número de eventos de mercado por Distrito Administrativo. Esta informação é importante para monitorar o número de casos de cada Distrito, servindo como um indicador da representatividade da amostra, o que auxilia nas conclusões sobre os resultados da análise de desempenho, bem como na orientação de estratégias para ampliar o número de eventos de mercado em caso de baixa representatividade.

Abaixo da área de visualização do mapa são apresentados os registros dos eventos de mercado selecionados e seus respectivos campos de informações, sendo eles: Distrito, Inscrição Imobiliária, Área Terreno, Preço Mercado Terreno, Valor Venal 2023, Área Lote Cadastral, Nível de Avaliação Individual, Preço M² Mercado Terreno, Valor M² Territorial.

A Figura 20 exemplifica a aplicação dos filtros e os diferentes gráficos resultantes. No caso apresentado na Figura 20 foram filtrados os eventos do Distrito Administrativo Sede Insular, e apenas eventos de terrenos com superfície no intervalo de 501 a 1000 m².

Figura 20 – Exemplo do Painel de BI sendo atualizados conforme os filtros cruzados, e a adequação dos gráficos para os mesmos.



Fonte: Elaborado pelo autor.

Com isso, o mapa é atualizado para apresentar os resultados da seleção, e tais registros são apresentados de forma tabular abaixo da visualização espacial. É importante observar que o mapa está tematizado com base nos valores de metro quadrado (m²),

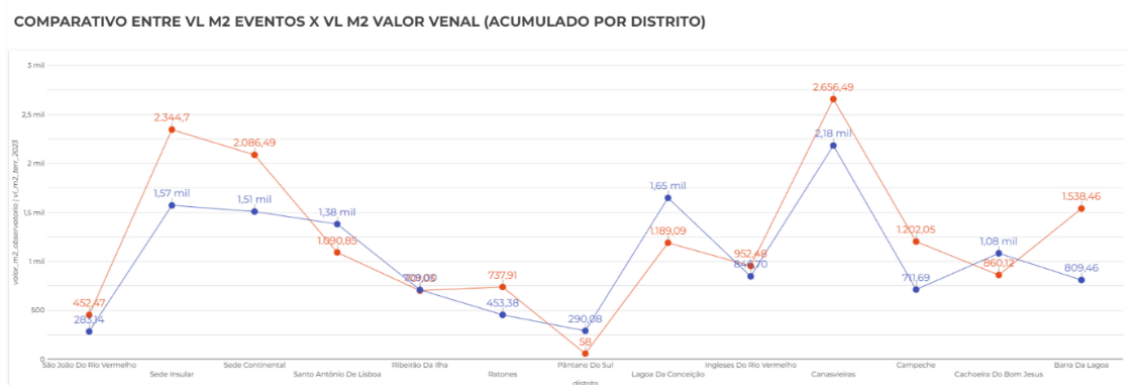
variando de tons amarelo claro para valores menores a tons de vermelho mais escuro para valores mais elevados. Essa representação visual possibilita a avaliação das áreas mais valorizadas e menos valorizadas do Distrito Sede Insular.

Na tabela apresentada na Figura 20 é possível observar as diferenças entre o valor do metro quadrado (m²) de eventos de mercado e o valor do m² de valores venais. A análise foi realizada por Distrito e considerou uma média ponderada dos valores. Essa informação é importante para entender as diferenças entre os preços praticados no mercado imobiliário e os valores cadastrais (venais), que são utilizados como referência para o cálculo do IPTU.

Observa-se ainda que a análise por Distrito permite identificar possíveis variações do preço de acordo com a localização do imóvel. Esses resultados são relevantes para subsidiar ações de políticas públicas e decisões no mercado imobiliário.

Na Figura 21 é apresentado o comparativo entre o preço médio do m² de eventos de mercado e o valor médio de m² dos valores cadastrais (venais), desta vez pelas amostras de cada Distrito, com suas respectivas inscrições imobiliárias.

Figura 21 – Recorte do BI, que apresenta o comparativo entre o valor do metro quadrado (m²) de eventos de mercado e o valor do m² de valores venais.



Fonte: Elaborado pelo autor

No gráfico da Figura 21 é possível observar, por Distrito, que as médias dos preços dos eventos do mercado podem ser significativamente superiores às médias dos valores cadastrais. No entanto, o oposto também ocorre. No Distrito Sede Insular, por exemplo, a média dos preços de terrenos nos eventos de mercado é de R\$ 2.344,7 por m², enquanto os valores cadastrais apresentam uma média de R\$ 1.570,00 por m², resultando

em uma diferença de aproximadamente 49,34% entre os valores praticados. Por outro lado, na Lagoa da Conceição, nota-se o contrário. A média dos preços dos eventos de mercado é de R\$ 1.189,09 por m², enquanto a média dos valores cadastrais é de R\$ 1.650,00 por m², gerando uma diferença de aproximadamente 27,9% entre os valores praticados. Quanto maior o número de eventos de mercado na base de dados do observatório do mercado imobiliário, mais representativo vai ser o resultado da análise estatística, gerando assim maior assertividade no monitoramento dos valores cadastrais.

A análise mencionada destaca, de maneira simples e objetiva, o potencial dos Observatórios do Mercado Imobiliário (OMIs) como ferramentas de apoio ao monitoramento dos valores cadastrais. A utilização de painéis de visualização, seja por meio do desenvolvimento de uma plataforma abrangente ou pelo aproveitamento de ferramentas disponíveis para visualização de dados, como os sistemas de Business Intelligence (BIs), possibilita uma interpretação mais aprofundada das informações provenientes dos observatórios. Isso viabiliza o acompanhamento visual e prático dos dados dos eventos de mercado e dos existentes no cadastro imobiliário.

Além do gráfico exemplificado na Figura 21 foram geradas outras visualizações para cada Distrito Administrativo, considerando a comparação individual de cada evento de mercado com os registros do cadastro imobiliário. Na Figura 22, a linha azul corresponde ao valor de metro quadrado cadastral e a vermelha o preço de metro quadrado praticado no mercado imobiliário. Esta visualização pode apoiar nas análises de avaliação individual dos dados, sendo útil para validar os dados e identificar possíveis casos atípicos (outliers) ou erros de armazenamento. É possível clicar em um dos eventos do mercado e visualizá-lo de forma independente no mapa superior.

Figura 22 – Gráficos do BI apresentando as diferenças entre o preço de mercado dos imóveis com relação ao valor venal, para cada imóvel do respectivo Distrito Administrativo.



Fonte: Elaborado pelo autor.

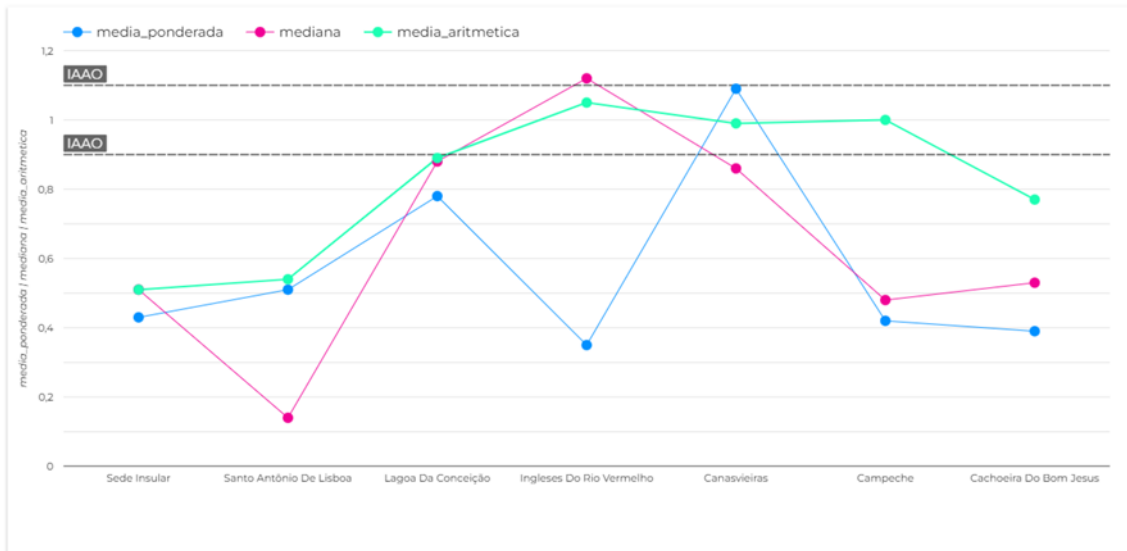
4.7.1 Visualização de análise de desempenho

Outro procedimento realizado para subsidiar a análise e visualização dos dados do OMI foi a criação de gráficos de análise de desempenho, utilizando os dados da Tabela 5. Com o objetivo de proporcionar o monitoramento contínuo dos valores de mercado e valores cadastrais, uma *view* específica foi desenvolvida para os scripts de desempenho, alimentando os gráficos do BI. Essa abordagem permite uma avaliação dinâmica, possibilitando a incorporação de novos eventos de mercado ao observatório, e o recálculo das fórmulas em tempo real.

Foram elaborados três gráficos para visualização das informações. A Figura 23 apresenta o gráfico correspondente ao nível de avaliação, com suas respectivas medidas e seus Distritos Administrativos. A Figura 24 e Figura 25 apresentam os gráficos de Uniformidade, com relação ao Coeficiente de Dispersão e Diferencial Relacional ao Preço, respectivamente.

Figura 23 - Gráfico do BI que apresenta a análise de desempenho de acordo com Nível de Avaliação, Mediana e Média Ponderada, além da linha de referência aceitável do IAAO (1991).

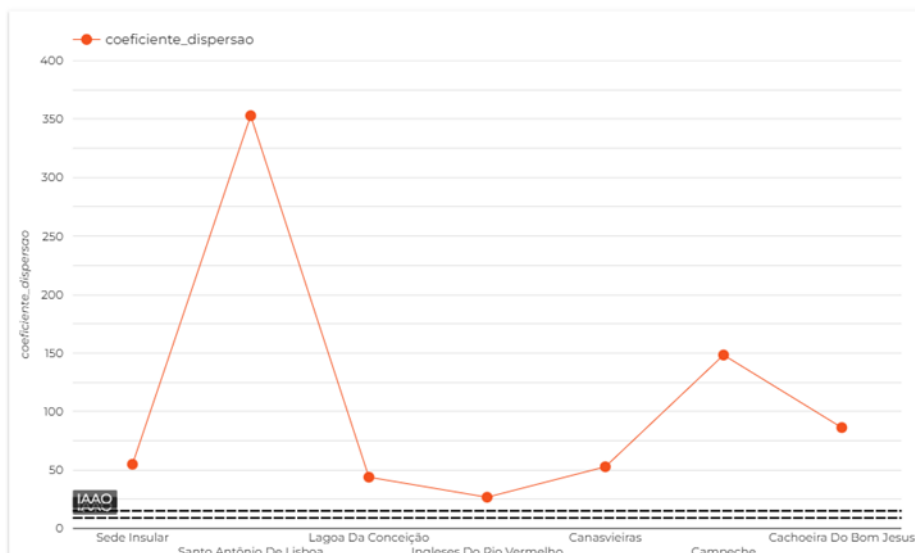
Nível de Avaliação - Média Aritmética, Mediana e Média Ponderada



Fonte: Elaborado pelo autor.

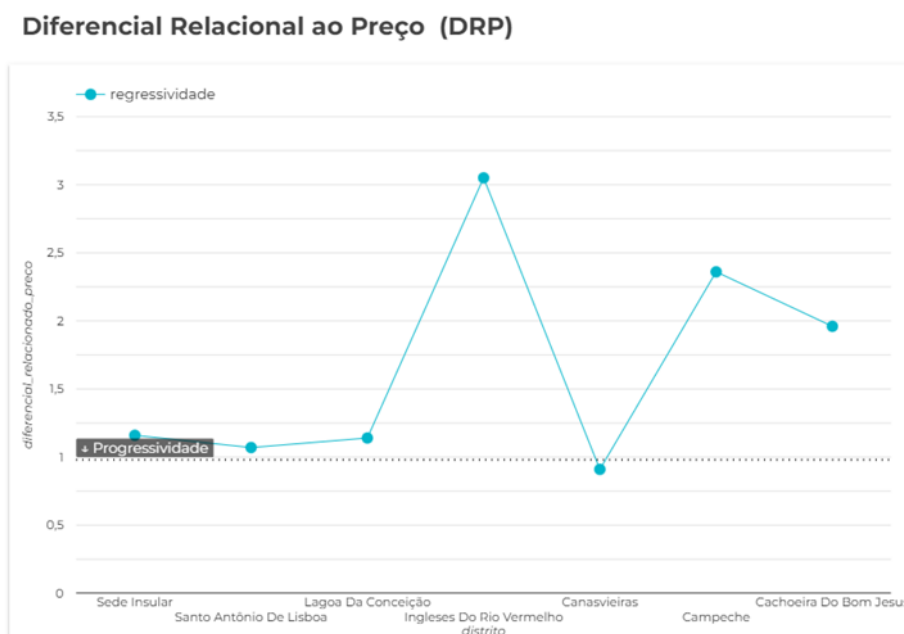
Figura 24 – Gráfico do BI que apresenta a análise de desempenho de acordo Uniformidade – Coeficiente de Dispersão.

Coeficiente de Dispersão (%) (CD):



Fonte: Elaborado pelo autor.

Figura 25 – Gráfico do BI que apresenta a análise de desempenho de acordo Uniformidade – Diferencial Relacional ao Preço.



Fonte: Elaborado pelo autor.

O projeto encontra-se disponível em: <https://lookerstudio.google.com/s/nUZ-gMKTHL8>

5 CONSIDERAÇÕES FINAIS

A implementação de um OMI envolve a criação de uma base de dados confiável e atualizada, bem como procedimentos apoiados por soluções de software para sua operacionalização, interoperabilidade e publicização de dados e indicadores. A presente pesquisa teve como objetivo propor um método para implementar e operacionalizar um OMI, a partir de softwares livres e gratuitos, de forma a fornecer um mecanismo de apoio ao monitoramento dos valores cadastrais dos municípios e contribuir para a formulação e qualificação de políticas urbanas.

O estabelecimento desse mecanismo desempenha um papel fundamental nos principais modelos de avaliação em massa de imóveis, uma vez que a precisão desses modelos é de extrema importância para assegurar uma distribuição mais justa e equitativa dos tributos sobre a propriedade entre os contribuintes.

No contexto da operacionalização dos OMIs, a adoção de estratégias e procedimentos que envolvem a coleta automatizada de eventos do mercado imobiliário

oferece possibilidades promissoras. Essa abordagem viabiliza o monitoramento contínuo do mercado imobiliário, proporcionando um processo mais sistemático, ágil e potencialmente mais eficiente em termos de quantidade de dados e velocidade de coleta. Além disso, ela pode ser uma alternativa menos onerosa em comparação aos modelos mais tradicionais, como as coletas realizadas em campo. Esta estratégia, por outro lado, evidencia a importância de sistemas de correção, higienização e validação dos dados, tanto a partir de técnicas computacionais, como pelo envolvimento de um grupo técnico especializado.

A geocodificação de endereços demonstrou ser um importante procedimento para validação da qualidade posicional de eventos coletados de forma automatizada em portais imobiliários. A utilização de mecanismos de geocodificação, que retornem além das coordenadas geográficas, o nível de precisão encontrado no georreferenciamento de sua base de endereçamento, é fundamental para o apoio na validação visual subsequente pelos autores envolvidos. Este mecanismo evidencia que a dificuldade no georreferenciamento de eventos de mercado territoriais é um desafio a ser superado. Durante o processo de geocodificação de endereços, avaliando toda a amostra de eventos, que inclui apartamentos, casas, lojas, entre outros, obteve-se aproximadamente 40% de retornos de precisão a nível de número de porta (ROOFTOP). No entanto, quando se trata de terrenos, essa precisão cai para 24,75%. É importante ressaltar que a baixa precisão é, principalmente, decorrente do fato de que anunciantes de terrenos em portais imobiliários declaram com menor frequência o endereço completo. Esse cenário pode ser explicado por dois principais motivos: a falta de acesso à informação cadastral a nível de número de porta por parte do contribuinte e/ou proprietário do imóvel e a omissão da informação de endereço completo por questões estratégicas de negociação, como por exemplo, indicar que o imóvel está em uma avenida principal do bairro, quando, na verdade, está localizado em uma rua próxima. É preciso buscar alternativas para superar essa limitação, a fim de aprimorar a qualidade dos dados utilizados em observatórios do mercado imobiliário e otimizar os resultados obtidos.

Apesar do momento atual mostrar-se promissor com relação a técnicas de automatização de processos, e o OMI ser uma tecnologia com potencial de ter cada vez mais ferramentas tecnológicas e processos automatizados em suas diferentes etapas, fica evidente a importância de se constituir um grupo técnico para sua gestão e manutenção. Os sistemas cadastrais utilizados em diferentes regiões da América Latina, incluindo outras partes do mundo, apresentam características e distinções específicas, muitas vezes

operando em nível municipal. Nesse contexto, é fundamental o conhecimento do sistema cadastral, das legislações referentes ao parcelamento do solo do município e dos modelos de cálculo de tributos associados ao mercado imobiliário para estabelecer uma conexão precisa entre os eventos de mercado e as parcelas cadastrais. Isso requer familiaridade com os procedimentos e normas adotados.

A utilização de soluções de software de código aberto e gratuito desempenha um papel fundamental na viabilização e operacionalização do Observatório do Mercado Imobiliário (OMI). Essas ferramentas, de fato, oferecem as funcionalidades necessárias para o adequado funcionamento eficiente do observatório, abrangendo desde a coleta até a análise e visualização de dados relevantes. Ao adotar esse modelo, a administração pública pode estabelecer seus próprios observatórios sem enfrentar altos custos de infraestrutura e licenças de software. Essa abordagem acessível e econômica democratiza o acesso às ferramentas necessárias para compreender e monitorar o mercado imobiliário, permitindo que mais instituições e órgãos públicos se beneficiem dessa capacidade analítica. Ao optar por softwares de código aberto e gratuito, pode-se aproveitar uma comunidade de desenvolvedores engajados e colaborativos, que contribuem constantemente com melhorias, atualizações e correções.

O OMI apresenta-se como um instrumento capaz de colaborar no monitoramento dos valores cadastrais no âmbito municipal, e a presente pesquisa contribuiu no sentido de demonstrar que, mesmo que em um estágio inicial da implementação do mecanismo no Município de Florianópolis, é possível realizar diferentes análises e validações sobre os dados cadastrais e seus respectivos valores venais. A partir dos resultados alcançados, abre-se um leque de oportunidades para aprimorar as análises de desempenho, corrigir inconsistências cadastrais, realizar estudos de avaliação em massa de imóveis, gerar PVGs e promover um planejamento territorial mais eficiente e sustentável.

6 REFERÊNCIAS

BARNETT V., LEWIS T., **Outliers in Statistical Data**. John Wiley, 1994.

BRASIL. MINISTÉRIO DAS CIDADES. Portaria nº 511, de 7 de dezembro de 2009. **Diretrizes para a criação, instituição e atualização do Cadastro Territorial Multifinalitário (CTM) nos municípios brasileiros**. Diário Oficial da União, Seção 1, nº 234, p. 75, 8 de dezembro de 2009. ISSN 1677-2339.

BRASIL. Lei nº 10.257, de 10 de julho de 2001. **Estatuto da Cidade**. Diário Oficial da União, Brasília, DF, 10 jul. 2001.

BRASIL. MINISTÉRIO DAS CIDADES. **Zeca Dastro e as Diretrizes para o Cadastro Territorial Multifinalitário**. Apoio: Lincoln Institute of Land Policy e Caixa Econômica Federal. Organizadores: Diego Alfonso Erba e Egláisa Micheline Pontes Cunha. 46p.

CARNEIRO, A. F. T.; ERBA, D. A.; AUGUSTO, E. A. A. **Cadastro multifinalitário 3d: conceitos e perspectivas de implantação no brasil**. Revista Brasileira de Cartografia, v. 64, n. 2, 11. 2012.

BRASIL. MINISTÉRIO DAS CIDADES. **Outorga Onerosa do Direito de Construir**. Programa Nacional de Capacitação das Cidades. Coleção Cadernos técnicos de regulamentação e implementação de instrumentos do Estatuto da Cidade. Volume 1, 2012.

BRASIL. MINISTÉRIO DAS CIDADES. **Parcelamento, Edificação ou Utilização Compulsórios (PEUC) e IPTU Progressivo no Tempo**. Programa Nacional de Capacitação das Cidades. Coleção Cadernos técnicos de regulamentação e implementação de instrumentos do Estatuto da Cidade. Volume 2. 2015, 80p.

DE CESARE, C. M., SILVA, E., SILVA, L. R. 2023. **Avaliação de Imóveis**. In: SILVA, E (org). Cadastro territorial multifinalitário aplicado à gestão municipal. 1. ed. Florianópolis: Universidade Federal de Santa Catarina, 2023. Cap. 5, p. 101-126.

DE CESARE, C. M. **Valuación de inmuebles para fines fiscales**. Lincoln Institute of Land Policy (2003).

DE CESARE, C. M. O Cadastro como Instrumento de Política Fiscal. In: **Cadastro multifinalitário como instrumento de política fiscal e urbana**. Orgs: Diego Alfonso Erba; Fabricio Leal de Oliveira; Pedro de Novais Lima Junior. Rio de Janeiro, 2005. pg 15-40

DE CESARE, C. M. CUNHA, E.M.P., 2012 - **Avaliação em massa de imóveis para fins fiscais: Discussão, análise e identificação de soluções para problemas e casos práticos** / Claudia M. De Cesare e Egláisa Micheline Pontes Cunha. Organização: Claudia M. De Cesare e Egláisa Micheline Pontes Cunha. Brasília: Ministério das Cidades, 2012. 116 p.: il. color.; 24cm. Programa Nacional de Capacitação das Cidades e Lincoln Institute of Land Policy.

ELMASRI, R.; NAVATHE, S. B. **Fundamentals of database systems**. Addison Wesley, 2005.

ENEMARK, STIG. **Land Administration Systems** - managing rights, restrictions and responsibilities in land. Map World Forum, Hyderabad, India, 2009.

ERBA, DIEGO ALFONSO; PIUMETTO, MARIO. **Para leer el suelo urbano**, catastros multifinalitarios para la planificación y el desarrollo de las ciudades de América Latina. Enfoque en políticas de suelo. Lincoln Institute of Land Policy. 2016 54p

ERBA, DIEGO ALFONSO (Org). **Sistemas de Información Geográfica aplicados a estudios urbanos**: experiencias latinoamericanas. Lincoln Institute of Land Policy, 2006.

ERBA, DIEGO ALFONSO. **O Cadastro Territorial**: Disposições gerais, benefícios e desafios. Seminário Regional Norte “Diretrizes Nacionais para o Cadastro Territorial Multifinalitário. Macapá, AP. Program on Latin America and the Caribbean. Lincoln Institute of Land Policy, Maio de 2014.

ERBA, DIEGO ALFONSO (Org). **Catastro Multifinalitario**: aplicado a la definición de políticas de suelo urbano. Cambridge, MA. Lincoln Institute of Land Policy, 2007.

ERBA, Diego ALFONSO; PIUMETTO, Mario. **Catastro Territorial Multifinalitário**. Documento de trabajo del Lincoln Institute of Land Policy. Lincoln Institute of Land Policy, 2013. Disponível em: http://www.lincolninst.edu/sites/default/files/pubfiles/erba-wp14de1sp-full_0.pdf . Acesso em: 19/11/2021

ERBA, DIEGO ALFONSO; PIUMETTO, MARIO. **Mapa de valores del suelo de América Latina construido mediante crowdsourcing y aplicaciones SIG en la nube**. Informe Final, Lincoln Institute of Land Policy, 2016.

ERBA, DIEGO ALFONSO. O Cadastro Territorial: passado, presente e futuro. In: **Cadastro multifinalitário como instrumento de política fiscal e urbana**. Orgs: Diego Alfonso Erba; Fabricio Leal de Oliveira; Pedro de Novais Lima Junior. Rio de Janeiro, 2005. pg 15-40

GONZALEZ, MARCO. **Aplicação de técnicas de descobrimento de conhecimento em bases de dados e de inteligência artificial em avaliações de imóveis**. Porto Alegre, 2002. Tese de Doutorado, Programa de Pós-Graduação em Engenharia Civil da Universidade Federal do Rio Grande do Sul.

GONZALES, MARCO. **Plantas de Valores com dados de ITBI**. Alternativa Fiscal para os Municípios. Revista de Administração Municipal, Rio de Janeiro, v. 63, n.219, p. 92-99, 1996.

GUERRERO, MARTHA. **Diseño metodológico para crear Infraestructuras de Datos Espaciales a escala Ciudad-Región en Colombia**. Universidade Nacional de Colombia. Facultad de Ciencias Agrarias. Bogotá, Colombia. 2014.

HARRIES, KEITH. **Mapping Crime: Principle and Practice**. Chapter 4: Mapping Crime and Geographic Information Systems. 1999

Hawkis, D. M. 1980. **Identification of Outliers. Monographs on applied probability and statistics.** DOI: 10.1007/978-94-015-3994-4.

IAAO - **International Association of Assessing Officers. 1990. Property Appraisal and Assessment Administration.** Ed. By J.K. Eckert, IAAO: USA.

IDECOR (2019). **DECOR... en “la nube”!** In: IDECOR. Disponível em <https://www.idecor.gob.ar/idecor-en-la-nube/> Data: 02/03/2023.

IDECOR (2023). **Sobre o IDECOR.** In: IDECOR. Disponível em <https://www.idecor.gob.ar/que-es/>. Data: 02/01/2023.

IDECOR. (2021). **Estudio del Mercado de Suelo Urbano de la Provincia de Córdoba 2021.** Recuperado em 10 de setembro de 2022, de https://www.idecor.gob.ar/wp-content/uploads/2021/12/Informe-Final_Valores-Tierra-Urbana-2021_GobCba.pdf

Aurélio Marinho Jargas (2009). **Expressões Regulares - Uma Abordagem Divertida 3 ed.** [S.l.]: Novatec. 208 páginas. ISBN 978-85-7522-212-6.

JOHNSON R., **Applied Multivariate Statistical Analysis.** Prentice Hall, 1992.

KAUFMANN, J; STEUDLER, D; Etal. **Benchmarking Cadastral Systems.** FIG-Commission 7 - Cadastre and Land Management. Working Group 1998-2002 - Reforming the Cadastre. April, 2002.

MACHADO, J.V.C., Silva, E., Hochheim, N., Rosolem, G.P.N. (2020). **"Análise de Desempenho do Modelo de Avaliação de Imóveis para Fins Tributários de Joinville-SC: Estudo de Caso no Bairro Vila Nova"**. Artigo do 14 ° Cobrac. 2020.

MARIANI, GRACIELA. **Los Sistemas de Información Geográfica (SIG) - Caja de herramientas para un Nuevo Urbanismo.** Red Nuestras Ciudades. 2012.

MICHAELIS. **Dicionário Brasileiro da Língua Portuguesa.** 2021. Online. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/observatorio/>

PAZOLINI, T. **OBSERVATÓRIO DE VALORES IMOBILIÁRIOS: MODELAGEM CONCEITUAL BASEADA EM OMT-G, 2019.** Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Transportes e Gestão Territorial (PPGTG).

RESTREPO, LILIANA BUSTAMANTE; ERBA, DIEGO ALFONSO. **El Catastro Territorial en la República de Colombia.** in: El Catastro Territorial en América Latina y el Caribe. Org: Diego Alfonso Erba. Lincoln Institute of Land Policy. 2008. Pág. 125 - 141.

SABELLA, E. W. 1975. **An Empirical Procedure for Evaluating Local Assessment Compliance with State Assessment Law.** Journal of Economics and Business 28 (Fall): 55-59.

Secretaria de Coordenação e Governança do Patrimônio da União (SPU/ME).
Especificações Técnicas para Estruturação de Dados Geoespaciais Vetoriais - ET-EDGV 2.0: Especificação Técnica para Estruturação de Dados GeoEspaciais Vetoriais do Patrimônio Imobiliário Público Federal (ET-EDGV PIPB 2.0). 2020.

SKABA, DANIEL ALBERT. **Metodologias de geocodificação dos dados da saúde.** 2009. 155 f. Tese (Doutorado em Saúde Pública) - Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rio de Janeiro, 2009.

SMOLKA, MARTIN O; MULLAHY, LAURA. **Perspectivas urbanas:** temas críticos en políticas de suelo en América Latina. Lincoln Institute of Land Policy. 2010.

SILVA, C.B., LONGO, E.S., & SILVA, E. **Coleta de dados de imóveis de forma automatizada para fins de políticas públicas.** Artigo do 14 ° Cobrac. 2020.

SILVA, E. **Cadastro Técnico Multifinalitário: base fundamental para avaliação em massa de imóveis.** Florianópolis, 2006. XVIII, 201 f. Tese (Doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Produção.

SILVA, E. **Mercado de Suelo y Valuaciones Masivas.** LINCOLN INSTITUTE OF LAND POLICY. Red Académica de Catastro Multifinalitario. 2020

SILVA, EVERTON DA; VERDINELLI, MIGUEL ANGEL. **Avaliação em massa de terrenos em Blumenau (Santa Catarina - Brasil) usando análise fatorial de correspondência e regressão múltipla.** In : IX Congresso Nacional y IV Latinoamericano de Agrimensura, Cordoba, Rep. Argentina. Anais, Tomo III, 1997.

SILVA, EVERTON DA; VERDINELLI, MIGUEL ANGEL. **Proposta de avaliação coletiva de imóveis do tipo apartamento da cidade de Blumenau, SC.** In : 4º Congresso Brasileiro de Cadastro Técnico Multifinalitário, Florianópolis. Anais : CD, 2000.

STUBKJAERS, E. **The cadastre as a socio-technical system.** GIS Dev Asia Pac Mon Mag Geogr Info Sci 10(6):26–28,2006. Disponível em:

https://vbn.aau.dk/ws/portalfiles/portal/65320547/The_cadastre_as_a_socio_technical_system.pdf . Acessado em: 03 de Fevereiro de 2022.

TAKAI, OSVALDO KOTARO; ITALIANO, ISABEL CRISTINA; FERREIRA, JOÃO EDUARDO. **INTRODUÇÃO A BANCO DE DADOS.** São Paulo: Dcc-ime-usp, 2005.

APÊNDICE I

Criação das tabelas do observatório.

```
create schema public;

--CRIAR EXTENSÕES - Postgis

--DROP EXTENSION postgis;

SET search_path to public;

CREATE EXTENSION postgis schema public;

CREATE SCHEMA observatorio;

-- criar esquema do observatorio

CREATE SCHEMA auditoria;

-- criar esquema da auditoria

CREATE TABLE observatorio.base1 (
    id serial NOT NULL,
    codigo varchar NULL,
    datetime timestamptz NULL,
    url varchar NULL,
    siteorigem varchar NULL,
    "data" date NULL,
    tipologia varchar NULL,
    titulo varchar NULL,
```

preco float8 NULL,
condominio float8 NULL,
area float8 NULL,
anunciante varchar NULL,
telefone varchar NULL,
descricao varchar NULL,
caracteristica varchar NULL,
quarto int4 NULL,
suite int4 NULL,
banheiro int4 NULL,
vaga int4 NULL,
andar int4 NULL,
endereco varchar NULL,
bairro varchar NULL,
cep int4 NULL,
municipio varchar NULL,
uf varchar NULL,
urlmapa varchar NULL,
imagem varchar NULL,
imagemlista varchar NULL,
latitude float8 NULL,
longitude float8 NULL,
controle varchar NULL,
CONSTRAINT base1_pkey PRIMARY KEY (id)

)

```

WITH (
    OIDS=FALSE
);
ALTER TABLE observatorio.base1
    OWNER TO postgres;

-- permissões de grupos de usuarios
CREATE ROLE visualizador;
CREATE ROLE editor;
CREATE ROLE admin;

GRANT SELECT ON ALL TABLES IN SCHEMA observatorio TO
visualizador;

ALTER DEFAULT PRIVILEGES IN SCHEMA observatorio GRANT SELECT
ON TABLES TO visualizador;

ALTER DEFAULT PRIVILEGES IN SCHEMA observatorio GRANT
SELECT,INSERT,UPDATE ON TABLES TO editor;

GRANT ALL PRIVILEGES ON SCHEMA "observatorio" to admin;

-- criar usuarios e vincular aos seus grupos de permissao
CREATE USER usuario_visualizador WITH PASSWORD '12345'; --user e
senha (armazenada de forma criptografada)

GRANT visualizador to usuario_visualizador;

CREATE USER usuario_editor WITH PASSWORD '123456';--user e senha
(armazenada de forma criptografada)

GRANT editor to usuario_editor;

```

```
CREATE USER usuario_admin WITH PASSWORD '1234567';--user e senha  
(armazenada de forma criptografada)
```

```
GRANT admin to usuario_admin;
```

```
-- Create table
```

```
CREATE TABLE observatorio.observatorio (
```

```
    mmlink SERIAL,
```

```
    inscricao VARCHAR (17),
```

```
    tp_fonte TEXT,
```

```
    fonte TEXT,
```

```
    phone TEXT,
```

```
    link TEXT,
```

```
    external_id TEXT,
```

```
    origin_id TEXT,
```

```
    situacao TEXT,
```

```
    description TEXT,
```

```
    title TEXT,
```

```
    street TEXT,
```

```
    street_number NUMERIC,
```

```
    zip_code NUMERIC,
```

```
    bairro TEXT,
```

```
    cidade TEXT,
```

```
    testada NUMERIC,
```

```
    usage_type TEXT,
```


tipo TEXT,
caract TEXT,
tipologia TEXT,
paredes TEXT,
padrao TEXT,
conserv TEXT,
n_pavim NUMERIC,
pavimento NUMERIC,
area_size NUMERIC,
area_terreno NUMERIC,
area_constr NUMERIC,
t_uni_lote NUMERIC,
area_unidade NUMERIC,
area_tot NUMERIC,
condominio_preco NUMERIC,
point_precision TEXT,
qualidade TEXT,
validacao TEXT,
obs TEXT,
alug_venda TEXT,
price NUMERIC,
iptu_price NUMERIC,
valor_m2 NUMERIC,
bedrooms NUMERIC,
bathrooms NUMERIC,
suite NUMERIC,

```

parking NUMERIC,

data_evento DATE,

data_atualizacao TIMESTAMP WITH TIME ZONE,

usuario TEXT,

geom GEOMETRY(POINT,31982),

CONSTRAINT pk_observatorio_mslink PRIMARY KEY (mslink)

);

-- Create the spatial index on geom column of observatorio

CREATE INDEX SIDX_observatorio

ON observatorio.observatorio

USING GIST (geom);

-- Create função e trigger para atualizar a data de atualização a cada edição

CREATE OR REPLACE FUNCTION

observatorio.function_data_atualizacao_observatorio()

RETURNS TRIGGER AS $exemplo$

BEGIN

NEW.data_atualizacao = now();

RETURN NEW;

END;

$exemplo$ language 'plpgsql';

CREATE TRIGGER trigger_data_atualizacao_observatorio

```

```
BEFORE UPDATE OR INSERT ON observatorio.observatorio
FOR EACH ROW EXECUTE PROCEDURE
observatorio.function_data_atualizacao_observatorio();
```

-- Create função e trigger para atualizar o usuário a cada edição

```
CREATE OR REPLACE FUNCTION
observatorio.function_usuario_observatorio()
```

```
RETURNS TRIGGER AS $usuario$
```

```
BEGIN
```

```
NEW.usuario = user ;
```

```
RETURN NEW;
```

```
END;
```

```
$usuario$ language 'plpgsql';
```

```
CREATE TRIGGER trigger_usuario_observatorio
```

```
BEFORE UPDATE OR INSERT ON observatorio.observatorio
```

```
FOR EACH ROW EXECUTE PROCEDURE
observatorio.function_usuario_observatorio();
```

```
CREATE TABLE auditorobservatorio ia.observatorio (
```

```
mslink_audit SERIAL,
```

```
mslink INTEGER,
```

```
inscricao VARCHAR (17),
```

```
tp_fonte TEXT,
```

```
fonte TEXT,
```

phone TEXT,
link TEXT,
external_id TEXT,
origin_id TEXT,
situacao TEXT,
description TEXT,
title TEXT,
street TEXT,
street_number NUMERIC,
zip_code NUMERIC,
bairro TEXT,
cidade TEXT,
testada NUMERIC,
usage_type TEXT,
tipo TEXT,
caract TEXT,
tipologia TEXT,
paredes TEXT,
padrao TEXT,
conserv TEXT,
n_pavim NUMERIC,
pavimento NUMERIC,
area_size NUMERIC,
area_terreno NUMERIC,
area_constr NUMERIC,
t_uni_lote NUMERIC,

area_unidade NUMERIC,
area_tot NUMERIC,
condominio_preco NUMERIC,
point_precision TEXT,
qualidade TEXT,
validacao TEXT,
obs TEXT,
alug_venda TEXT,
price NUMERIC,
iptu_price NUMERIC,
valor_m2 NUMERIC,
bedrooms NUMERIC,
bathrooms NUMERIC,
suite NUMERIC,
parking NUMERIC,
data_evento DATE,
geom GEOMETRY(POINT,31982),
usuario TEXT,
data_alteracao TIMESTAMP WITH TIME ZONE,
operacao TEXT,

CONSTRAINT pk_observatorio_mslink_audit PRIMARY KEY
(mslink_audit)

);

-- Create the spatial index on geom column of observatorio

```

CREATE INDEX SIDX_observatorio

    ON auditoria.observatorio

    USING GIST (geom);

-- Create function para auditoria de delete, insert e update

CREATE OR REPLACE FUNCTION auditoria.function_audit_observatorio()

RETURNS TRIGGER AS $$

BEGIN

    IF ( TG_OP= 'DELETE') THEN

        INSERT INTO auditoria.observatorio (mslink, inscricao, tp_fonte, fonte,
phone, link, external_id, origin_id, situacao, description, title, street, street_number,
zip_code, bairro, cidade, testada, usage_type, tipo, caract, tipologia, paredes, padrao,
conserv, n_pavim, pavimento, area_size, area_terreno, area_constr, t_uni_lote,
area_unidade, area_tot, condominio_preco, point_precision, qualidade, validacao, obs,
alug_venda, price, iptu_price, valor_m2, bedrooms, bathrooms, suite,
parking,data_evento,geom, usuario, data_alteracao, operacao)

            VALUES (old.mslink, old.inscricao, old.tp_fonte, old.fonte, old.phone,
old.link, old.external_id, old.origin_id, old.situacao, old.description, old.title, old.street,
old.street_number, old.zip_code, old.bairro, old.cidade, old.testada, old.usage_type,
old.tipo, old.caract, old.tipologia, old.paredes, old.padrao, old.conserv, old.n_pavim,
old.pavimento, old.area_size, old.area_terreno, old.area_constr, old.t_uni_lote,
old.area_unidade, old.area_tot, old.condominio_preco, old.point_precision,
old.qualidade, old.validacao, old.obs, old.alug_venda, old.price, old.ipatu_price,
old.valor_m2, old.bedrooms, old.bathrooms, old.suite, old.parking, old.data_evento,
old.geom, user, now(), 'DELETE');

        RETURN OLD;

    ELSIF ( TG_OP= 'UPDATE') THEN

        INSERT INTO auditoria.observatorio (mslink, inscricao, tp_fonte, fonte,
phone, link, external_id, origin_id, situacao, description, title, street, street_number,
zip_code, bairro, cidade, testada, usage_type, tipo, caract, tipologia, paredes, padrao,

```

conserv, n_pavim, pavimento, area_size, area_terreno, area_constr, t_uni_lote, area_unidade, area_tot, condominio_preco, point_precision, qualidade, validacao, obs, alug_venda, price, iptu_price, valor_m2, bedrooms, bathrooms, suite, parking,data_evento,geom, usuario, data_alteracao, operacao)

VALUES (old.mmlink, old.inscricao, old.tp_fonte, old.fonte, old.phone, old.link, old.external_id, old.origin_id, old.situacao, old.description, old.title, old.street, old.street_number, old.zip_code, old.bairro, old.cidade, old.testada, old.usage_type, old.tipo, old.caract, old.tipologia, old.paredes, old.padrao, old.conserv, old.n_pavim, old.pavimento, old.area_size, old.area_terreno, old.area_constr, old.t_uni_lote, old.area_unidade, old.area_tot, old.condominio_preco, old.point_precision, old.qualidade, old.validacao, old.obs, old.alug_venda, old.price, old.iptu_price, old.valor_m2, old.bedrooms, old.bathrooms, old.suite, old.parking, old.data_evento, old.geom, user, now(), 'UPDATE'),

(new.mmlink, new.inscricao, new.tp_fonte, new.fonte, new.phone, new.link, new.external_id, new.origin_id, new.situacao, new.description, new.title, new.street, new.street_number, new.zip_code, new.bairro, new.cidade, new.testada, new.usage_type, new.tipo, new.caract, new.tipologia, new.paredes, new.padrao, new.conserv, new.n_pavim, new.pavimento, new.area_size, new.area_terreno, new.area_constr, new.t_uni_lote, new.area_unidade, new.area_tot, new.condominio_preco, new.point_precision, new.qualidade, new.validacao, new.obs, new.alug_venda, new.price, new.iptu_price, new.valor_m2, new.bedrooms, new.bathrooms, new.suite, new.parking, new.data_evento, new.geom, user, now(), 'UPDATE');

RETURN NEW;

ELSIF (TG_OP= 'INSERT') THEN

INSERT INTO auditoria.observatorio (mmlink, inscricao, tp_fonte, fonte, phone, link, external_id, origin_id, situacao, description, title, street, street_number, zip_code, bairro, cidade, testada, usage_type, tipo, caract, tipologia, paredes, padrao, conserv, n_pavim, pavimento, area_size, area_terreno, area_constr, t_uni_lote, area_unidade, area_tot, condominio_preco, point_precision, qualidade, validacao, obs, alug_venda, price, iptu_price, valor_m2, bedrooms, bathrooms, suite, parking,data_evento,geom, usuario, data_alteracao, operacao)

```
VALUES (new.mslink, new.inscricao, new.tp_fonte,
new.fonte, new.phone, new.link, new.external_id, new.origin_id, new.situacao,
new.description, new.title, new.street, new.street_number, new.zip_code, new.bairro,
new.cidade, new.testada, new.usage_type, new.tipo, new.caract, new.tipologia,
new.paredes, new.padrao, new.conserv, new.n_pavim, new.pavimento, new.area_size,
new.area_terreno, new.area_constr, new.t_uni_lote, new.area_unidade, new.area_tot,
new.condominio_preco, new.point_precision, new.qualidade, new.validacao, new.obs,
new.alug_venda, new.price, new.ipitu_price, new.valor_m2, new.bedrooms,
new.bathrooms, new.suite, new.parking, new.data_evento, new.geom, user, now(),
'INSERT');
```

```
RETURN NEW;
```

```
END IF;
```

```
END;
```

```
$$ language 'plpgsql';
```

```
-- Create trigger para auditoria
```

```
CREATE TRIGGER trigger_audit_observatorio
```

```
BEFORE UPDATE OR INSERT OR DELETE ON observatorio.observatorio
```

```
FOR EACH ROW EXECUTE PROCEDURE
```

```
auditoria.function_audit_observatorio();
```


APÊNDICE II

```
----- NIVEL E
UNIFORMIDADE

--SCRIPS INDEPENDENTES:
--Cálculo da média aritmética por distrito:

SELECT distrito, AVG(vl_venl_pgv_cheia2023 /
preco_observatorio) AS media_aritmetica
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20
and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20
) analise
GROUP BY distrito
having count (*) > 4

--Cálculo da mediana por distrito:

SELECT distrito, PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER
BY vl_venl_pgv_cheia2023 / preco_observatorio) AS mediana
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
analise
GROUP BY distrito
having count (*) > 4

--Cálculo da média ponderada por distrito:

SELECT distrito, SUM(vl_venl_pgv_cheia2023) /
SUM(preco_observatorio) AS media_ponderada
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
```

```

area_terreno_observatorio::numeric between 0.80 and 1.20 )
analise
GROUP BY distrito
having count (*) > 4

```

--Cálculo do coeficiente de dispersão por distrito:

```

SELECT
    distrito,
    ((SUM(ABS((vl_venl_pgv_cheia2023 / preco_observatorio) -
mediana)) / COUNT(*) / mediana)) * 100
    AS coeficiente_dispersao
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
t,
    (SELECT PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY
vl_venl_pgv_cheia2023 / preco_observatorio) AS mediana,
distrito as distrito_med
    FROM (select * from observatorio.analise_desempenho
where tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0
and valor_m2_observatorio is not null and vl_m2_terr_2023
is not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
    analise group by distrito) AS mediana_cte

    where distrito = distrito_med

GROUP BY distrito, mediana
having count (*) > 4

```

--Verificação da uniformidade (diferencial relacionado ao preço) por distrito:

```

SELECT analise.distrito, (AVG(vl_venl_pgv_cheia2023 /
preco_observatorio)) / media_ponderada AS
diferencial_relacionado_preco
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
analise,

```

```

(SELECT distrito, SUM(vl_venl_pgv_cheia2023) /
SUM(preco_observatorio) AS media_ponderada
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
analise
GROUP BY distrito
having count (*) > 4) ponderada
where analise.distrito = ponderada.distrito
group by analise.distrito, media_ponderada
having count (*) > 4

```

--Verificação de regressividade ou progressividade por distrito:

```

SELECT distrito, CASE
    WHEN AVG(vl_m2_terr_2023 / valor_m2_observatorio)
/ (SUM(vl_m2_terr_2023) / SUM(valor_m2_observatorio)) >
1.03 THEN 'regressividade'
    WHEN AVG(vl_m2_terr_2023 / valor_m2_observatorio)
/ (SUM(vl_m2_terr_2023) / SUM(valor_m2_observatorio)) <
0.98 THEN 'progressividade'
    ELSE 'equilibrado'
END AS tipo_avaliacao
FROM (select * from observatorio.analise_desempenho where
tipo = 'territorial' and vl_venl_pgv_cheia2023 != 0 and
valor_m2_observatorio is not null and vl_m2_terr_2023 is
not null and qt_area_lote::numeric /
area_terreno_observatorio::numeric between 0.80 and 1.20 )
analise
GROUP BY distrito
having count (*) > 9

```

-- atualizar

----- SCRIPT QUE JUNTA TODAS ELAS

--drop view

observatorio.niveis_analise_desempenho_distrito;

```

create view observatorio.niveis_analise_desempenho_distrito
as
select
    distrito,
    round(media_aritmetica::numeric,2) media_aritmetica,
    round(mediana::numeric,2) mediana,
    round(media_ponderada::numeric,2) media_ponderada,
    round(coeficiente_dispersao::numeric,2)
coeficiente_dispersao,
    round(diferencial_relacionado_preco::numeric,2)
diferencial_relacionado_preco,
    tipo_avaliacao

from (

WITH dados_filtrados AS (
SELECT *
FROM observatorio.analise_desempenho
WHERE tipo = 'territorial'
    AND vl_venl_pgv_cheia2023 != 0
    AND valor_m2_observatorio IS NOT NULL
    AND vl_m2_terr_2023 IS NOT NULL
    AND qt_area_lote::numeric /
area_terreno_observatorio::numeric BETWEEN 0.80 AND 1.20
),
dados_mediana AS (
SELECT distrito, PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER
BY vl_venl_pgv_cheia2023 / preco_observatorio) AS mediana
FROM dados_filtrados
GROUP BY distrito
HAVING COUNT(*) > 4
),
dados_media_ponderada AS (
SELECT distrito, SUM(vl_venl_pgv_cheia2023) /
SUM(preco_observatorio) AS media_ponderada
FROM dados_filtrados
GROUP BY distrito
HAVING COUNT(*) > 4
)
SELECT
    dados_filtrados.distrito,
    AVG(vl_venl_pgv_cheia2023 / preco_observatorio) AS
media_aritmetica,
    dados_mediana.mediana AS mediana,
    dados_media_ponderada.media_ponderada AS media_ponderada,
    ((SUM(ABS((vl_venl_pgv_cheia2023 / preco_observatorio) -
dados_mediana.mediana)) / COUNT(*) /
dados_mediana.mediana)) * 100 AS coeficiente_dispersao,
    AVG(vl_venl_pgv_cheia2023 / preco_observatorio) /
dados_media_ponderada.media_ponderada AS
diferencial_relacionado_preco,

```

```

CASE
  WHEN AVG(vl_m2_terr_2023 / valor_m2_observatorio) /
(SUM(vl_m2_terr_2023) / SUM(valor_m2_observatorio)) > 1.03
THEN 'regressividade'
  WHEN AVG(vl_m2_terr_2023 / valor_m2_observatorio) /
(SUM(vl_m2_terr_2023) / SUM(valor_m2_observatorio)) < 0.98
THEN 'progressividade'
  ELSE 'equilibrado'
END AS tipo_avaliacao
FROM dados_filtrados
JOIN dados_mediana ON dados_filtrados.districto =
dados_mediana.districto
JOIN dados_media_ponderada ON dados_filtrados.districto =
dados_media_ponderada.districto
GROUP BY dados_filtrados.districto, dados_mediana.mediana,
dados_media_ponderada.media_ponderada
HAVING COUNT(*) > 4
) yes

```