

Visualização de Casos de Doenças Distribuídos por Bairros na Cidade de Santa Maria: Um Módulo de Processamento e Geoprocessamento para o Programa PPSUS UFN

Felipe dos Santos Martinez
Curso de Sistemas de Informação
Universidade Franciscana
CEP 97010-032 – Santa Maria, RS, Brasil
f.martinez@ufn.edu.br

Sylvio André Garcia Vieira
Curso de Sistemas de Informação
Universidade Franciscana
CEP 97010-032 – Santa Maria, RS, Brasil
sylvio@ufn.edu.br

Resumo—Neste artigo, descreve-se o desenvolvimento de um módulo de monitoramento para o Programa Pesquisa para o Sistema Único de Saúde da Universidade Franciscana (PPSUS UFN). O intuito deste módulo é identificar pontos de ocorrência de diferentes doenças com base nos dados fornecidos pelo sistema da Prefeitura da cidade de Santa Maria no Rio Grande do Sul.

Keywords: Python, Geoprocessamento, Saúde pública, Processamento de dados, PPSUS.

I. INTRODUÇÃO

Surtos e epidemias por doenças infecciosas têm se tornado um grande desafio de saúde pública em todo o mundo. Tais eventos afetam a qualidade de vida das pessoas, sobrecarregam os sistemas de saúde e geram graves consequências econômicas e sociais. Por exemplo, conforme a Organização Mundial de Saúde (OMS)[1] estima-se que 600 milhões de pessoas – quase 1 em cada 10 pessoas no mundo – adoecem e 420.000 morrem todos os anos devido a Doenças de Transmissão Hídrica e Alimentar (DTHA), ou seja, ingestão de água e/ou alimentos contaminados, capazes de causar doenças infecciosas, o que resulta na perda de 33 milhões de anos de vida saudáveis. Já as crianças menores de 5 anos carregam 40 por cento desta carga de DTHA, causando 125.000 mortes a cada ano.

Esses fatos demonstram a necessidade e urgência em desenvolver abordagens inovadoras e eficazes para detectar, monitorar e responder com antecedência a possíveis crises de saúde pública, garantindo que o poder público seja capaz de anteceder e supervisionar as possíveis mazelas e enfermidades que afetam as sociedades, poupando vidas e garantindo o bem-estar público.

Atualmente, há na cidade de Santa Maria no estado do Rio Grande do Sul, um projeto na Universidade Franciscana (UFN) que desenvolve um sistema piloto nomeado como Pesquisa para o sistema único de saúde da Universidade Franciscana (PPSUS UFN), que foi inicialmente concebido pelo Laboratório de Práticas dos cursos de computação da UFN (LAPINF UNF), junto ao Mestrado Profissional em

Saúde Materno Infantil. Este sistema começou em um projeto da Fundação de amparo à pesquisa do estado do Rio Grande do Sul (FAPERGS) apoiando o mestrado da estudante Márian O. Pires [2].

Segundo o Ministério da saúde [3] PPSUS é uma iniciativa inovadora por adotar um modelo de gestão descentralizado e participativo, envolvendo gestores, profissionais de saúde, pesquisadores e representantes da sociedade civil. Esse Programa constitui uma ferramenta potencialmente indutora para que os principais problemas de saúde da população figurem entre as linhas prioritárias de investigação dos pesquisadores brasileiros, tendo a relevância sócio-sanitária como critério norteador para a definição dos temas prioritários de pesquisa. Sendo assim, observou-se a necessidade de criar um módulo capaz de gerar relatórios de monitoramento de surtos e epidemias, permitindo que os órgãos públicos competentes possam realizar ações embasadas nos dados apresentados.

De acordo com Vasconcelos, V. V. et al. [4], a utilização do geoprocessamento na gestão de riscos desempenha um papel crucial ao permitir a análise de dados espaciais, possibilitando uma compreensão aprofundada do território e fornecendo subsídios para a tomada de decisões estratégicas no planejamento e implementação de medidas de redução de riscos e prevenção de desastres.

Como parte do cotidiano, Unidades Básicas de Saúde (UBS) produzem uma vasta quantidade de dados relacionados a prontuários de pacientes e que passando pelos processos necessários podem gerar informações relevantes para a prevenção e soluções de possíveis surtos e epidemias.

Assim, no presente trabalho, teve-se como objetivo geral criar um módulo de monitoramento para o sistema PPSUS UFN capaz de agir como ferramenta de vigilância contínua e adaptável, capaz de melhorar monitoramento e visualização diário de doenças através de geoprocessamento e permitindo ações rápidas do poder público com base em dados de casos como, por exemplo: dengue, coronavírus, DTHA, toxoplasmose, etc.

Já como objetivos específicos, visou-se assegurar a padronização e a consistência dos dados por meio de uma correção manual dos nomes de bairros, no intuito de garantir a precisão das informações processadas, além de inserir esses dados com geoprocessamento para gerar resultados confiáveis. Também, buscou-se a visualização e análise dos dados dos casos de doenças, desenvolvendo relatórios gráficos. Por fim, criou-se mapas detalhados da distribuição espacial de diferentes doenças.

II. REFERENCIAL TEÓRICO

No intuito de melhor compreensão, nesta seção descrevem-se os conceitos necessários e relevantes para o desenvolvimento deste trabalho apresentado. Cada subseção abordará aspectos cruciais para a compreensão e contextualização do tema, fornecendo uma base teórica sólida.

A. Doenças e Epidemias

Doença [5] é uma enfermidade ou estado clínico, independentemente de origem ou fonte, que represente ou possa representar um dano significativo para os seres humanos. Epidemias [5] é denominação utilizada em situações em que a doença envolve grande número de pessoas e atinge uma larga área geográfica.

B. Unidades Básicas de Saúde (UBS)

As Unidades Básicas de Saúde [6], que são as principais estruturas físicas da Atenção Básica, são instaladas próximas da vida dos usuários, desempenhando um papel central na garantia de acesso a uma saúde de qualidade. As unidades oferecem uma diversidade de serviços realizados pelo SUS, incluindo: acolhimento com classificação de risco, consultas de enfermagem, médicas e de saúde bucal, distribuição e administração de medicamentos, vacinas, curativos, visitas domiciliares, atividade em grupo nas escolas, educação em saúde, entre outras.

C. Processamento de Dados

Segundo Louzada Neto [7], o processamento de dados envolve a manipulação e análise de grandes volumes de informações para extrair insights significativos. No contexto da saúde pública, isso é fundamental para entender e controlar doenças, permitindo a identificação de padrões de infecção e a monitorização da propagação de doenças, como destaca Lima [8]. Ferramentas de análise de dados processam informações de diversas fontes, como registros médicos e dados epidemiológicos.

D. Geoprocessamento

Monteiro [9] define o geoprocessamento como a aplicação de técnicas de análise de dados geoespaciais para visualizar e analisar dados geográficos. Isso é particularmente útil em situações de epidemias [16], onde a doença envolve um grande número de pessoas e se espalha por uma ampla área geográfica. Utilizando Sistemas de Informação Geográfica (SIG), profissionais de saúde pública podem mapear a distribuição de doenças, identificar áreas de risco e monitorar a disseminação

de epidemias. Câmara et al. [10] destacam que o geoprocessamento permite a análise de fatores ambientais que contribuem para a propagação de doenças, facilitando a implementação de medidas preventivas mais eficazes e o desenvolvimento de estratégias de longo prazo.

E. Python

Python [11] é uma linguagem de programação interpretada, interativa e orientada a objetos. Ela oferece suporte a vários paradigmas de programação, como programação procedimental, funcional e orientada a objetos. Python é conhecida por possuir uma sintaxe clara e poderosa, combinando recursos como módulos, exceções, digitação dinâmica, tipos de dados de alto nível e classes.

Além disso, possui interfaces para chamadas de sistema, bibliotecas e sistemas de janelas, sendo extensível em C ou C++. Python é uma linguagem portátil, sendo executada em diversas variantes do Unix, incluindo Linux e macOS, além do Windows. Sua versatilidade e ampla gama de recursos fazem dela uma escolha popular entre os desenvolvedores em diferentes áreas, que disponibiliza conjuntos de pacotes e bibliotecas voltadas a ciência de dados. Dentre as bibliotecas e pacotes que foram utilizados neste trabalho, pode-se citar:

- Pandas [12]: é uma biblioteca *open source* que fornece estruturas de dados de alto desempenho e ferramentas para análise e manipulação de dados.
- NumPy [13]: pacote voltado para computação científica com uma coleção de funções matemáticas para trabalhar com estrutura de dados.
- Matplotlib [14]: é uma biblioteca de plotagem 2D do Python que produz números de qualidade de publicação em vários formatos de cópia impressa e ambientes interativos entre plataformas.
- GeoPandas [15]: é um projeto de código aberto para facilitar o trabalho com dados geoespaciais em python. GeoPandas estende os tipos de dados usados pelos pandas para permitir operações espaciais em tipos geométricos.
- Django [16] é um *framework* desenvolvido para cumprir prazos rápidos e ajudar os desenvolvedores experientes, cuidando da parte web para que o programador possa se concentrar apenas no seu aplicativo.

F. HTML

(Linguagem de Marcação de HiperTexto) [17] é o bloco de construção mais básico da web. Define o significado e a estrutura do conteúdo da web.

G. Sqlite3

SQLite [18] é uma biblioteca C que fornece um banco de dados leve baseado em disco que não requer um processo de servidor separado e permite acessar o banco de dados usando uma variante não padrão da linguagem de consulta SQL. Alguns aplicativos podem usar SQLite para armazenamento interno de dados. Também é possível prototipar um aplicativo usando SQLite e depois portar o código para um banco de dados maior, como PostgreSQL ou Oracle.

H. Data Frame

Um data frame [19] é uma estrutura de dados semelhante a uma matriz, porém as suas colunas têm nomes e podem conter dados de tipagem diferente. Um data frame pode ser visto como uma tabela de uma base de dados, em que cada linha corresponde a um registro (linha) da tabela. Cada coluna corresponde às propriedades (campos) a serem armazenadas para cada registro da tabela.

I. Shapefile

Um shapefile [20] é um formato de armazenamento de dados de vetor da Esri para armazenar a posição, a forma e os atributos de feições geográficas. É armazenado como um conjunto de arquivos relacionados e contém uma classe de feição. Os shapefiles normalmente contêm grandes feições com muitos dados associados e foi historicamente utilizado em aplicativos de desktop GIS.

III. TRABALHOS CORRELATOS

Nesta seção, será apresentados trabalhos relacionados ao tema, além de ferramentas e tecnologias semelhantes que podem ser utilizadas no desenvolvimento deste projeto.

A. O Impacto do Uso da Biblioteca Pandas do Python como Ferramenta de Análise de Dados Referente aos Casos Graves de COVID-19 Notificados pela Secretaria Estadual de Saúde do Recife entre Julho de 2021 a Junho de 2022

O artigo de Wagner Vidal Xavier da Silva [21], visa investigar o impacto do uso da biblioteca Pandas do Python como ferramenta de análise de dados em relação aos casos graves de COVID-19 notificados pela Secretaria Estadual de Saúde do Recife.

O estudo utilizou dados de notificações de casos graves de COVID-19 no período de julho de 2021 a junho de 2022, disponibilizados pela Secretaria Estadual de Saúde do Recife. Utilizando a biblioteca Pandas do Python, foram realizadas análises exploratórias e descritivas dos dados, como a identificação de padrões de distribuição dos casos graves de COVID-19 ao longo do tempo, bem como a correlação entre a idade dos pacientes e a gravidade dos casos.

A partir da análise da base de dados dos casos graves de Covid-19, compreendeu o estudo, por exemplo, a possibilidade de ser traçado um perfil dos distritos sanitários que mais notificaram casos da doença, o que pode ser interessante para traçar rotas de mobilidade, implementação de centros médicos e suporte hospitalar conforme a localização. Há diversas possibilidades de interpretação, dependendo da necessidade de quem precisa da informação.

Outro resultado factível de análise foi a questão do quantitativo de pessoas a partir do gênero, é observável como a doença se manifesta em homens e mulheres, se existe algum sintoma diverso, o tempo de recuperação assim como a porcentagem de óbito entre o gênero, etc. Verificam-se possibilidades de informação.

Estes resultados indicaram que o uso da biblioteca Pandas do Python permitiu uma análise mais detalhada e eficiente dos

dados, com a possibilidade de manipulação, organização e visualização dos dados de forma mais fácil e rápida. Além disso, o estudo identificou padrões na distribuição dos casos graves de COVID-19 ao longo do tempo, bem como a correlação entre a idade dos pacientes e a gravidade dos casos.

B. Sistema computacional automático para geração de relatórios epidemiológicos a partir de dados do datasus

O trabalho de Eliseu et. Al. [22] tem por objetivos desenvolver um programa computacional online que unificasse a metodologia de acesso, busca e interpretação de dados do DATASUS.

Foi estruturado na programação PHP 7 e Java Script, a partir dos dados do DATASUS e do IBGE. Utilizando as variáveis localização geográfica, ano, sexo e Classificação Internacional de Doenças (CID-10). Após a construção o programa, foi selecionado um dos CID disponíveis, C25 – Neoplasia Maligna do Pâncreas (C25) para que os resultados do programa fossem comparados com o método tradicional (manual) de criação de relatórios epidemiológicos.

Como resultados, comparado com o método tradicional, ambos tiveram mesmos resultados de óbitos e populações. A mortalidade em 105 habitantes teve diferença de 0,04 devido aos arredondamentos do método tradicional. Contudo, as significâncias dessas diferenças forma acima de 0,99 obtendo nenhuma diferença significativa entre os métodos. O programa cria mapas geográficos atualizados de mortalidade por unidades federativas. Os autores afirmam que o programa desenvolvido é um eficiente método de acesso, busca, interpretação e visualização dos relatórios epidemiológicos.

C. Utilização da linguagem Python na análise de dados de dengue em Recife, capital do estado de Pernambuco, Nordeste do Brasil e definição de linha de tendência usando Machine Learning

O objetivo deste trabalho de Angélica Félix et al [23], foi analisar os dados da dengue em Recife no ano de 2020 e definir uma linha de tendência para os casos de dengue em Pernambuco utilizando técnicas de Machine Learning. Para a análise dos dados da dengue em Recife em 2020, foram utilizadas bases de dados oficiais do estado de Pernambuco. Foram consultadas informações relacionadas à evolução dos casos de dengue, como sintomas, necessidade de internação, presença de doenças autoimunes e coinfeção com chikungunya. A análise foi realizada por meio de gráficos e mapas de calor.

Os resultados da análise dos dados da dengue em Recife em 2020 mostraram que a maioria dos casos apresentava sintomas de febre e cefaleia, com poucos pacientes necessitando de internação. A ocorrência de conjuntivite, doenças autoimunes e coinfeção com chikungunya foi baixa. Quanto à linha de tendência nos casos de dengue em Pernambuco, a aplicação da Regressão Linear indicou uma tendência decrescente ao longo dos anos. A previsão sugere que os casos de dengue em Recife tendem a diminuir no futuro, desde que não ocorram eventos extraordinários que afetem essa tendência. O trabalho também

menciona uma redução significativa nos casos confirmados de arboviroses em Recife em 2022 em comparação com o ano anterior.

Os trabalhos correlatos apresentados serão de grande utilidade no atual trabalho de conclusão de curso, pois tem estreita relação com o uso de dados publicos, uso de geoprocessamento e a utilização da Linguagem Python com a Biblioteca Pandas, que igualmente serão utilizadas neste trabalho.

IV. METODOLOGIA

O módulo desenvolvido de processamento e geoprocessamentos de dados do sistema de saúde da Prefeitura de Santa Maria está sendo integrado no sistema PPSUSUFN. Esse sistema, inicialmente, foi desenvolvido como produto tecnológico por uma egressa do Mestrado Profissional em Saúde Materno Infantil da Universidade Franciscana (UFN) [2]. Em um primeiro momento, seriam utilizados dados oriundos do Departamento de Informática do Sistema Único de Saúde (DATASUS), que disponibiliza informações que poderiam servir para subsidiar análises objetivas da situação sanitária, servindo como base tomadas de decisão alicerçadas em evidências e elaboração de programas de ações de saúde, utilizando dados de saúde fornecidos pelos municípios e estado. No entanto, a principal limitação dessa fonte de dados é o tempo em que é disponibilizado os dados para consulta, o que pode levar vários meses.

Devido à necessidade de criar um módulo capaz de gerar relatórios que identifiquem inícios de epidemias em estágios embrionários, optou-se por utilizar dados fornecidos pela Prefeitura municipal de Santa Maria (PMSM), acessados por meio de credenciais obtidas com a Secretaria Municipal de Saúde (SMS) pela pessoa do Secretário Municipal.

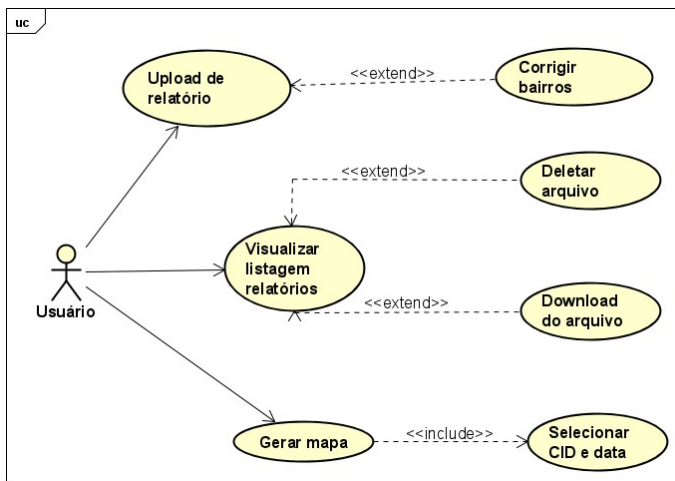


Figura 1: Diagrama de caso de uso. [24]

Como pode ser observado no Diagrama de caso de uso (Figura 1), o usuário interage com três páginas: Upload de relatório, relatórios, Gerar mapa. A seguir, será comentada as funcionalidades e características das páginas Upload de relatório e relatórios.

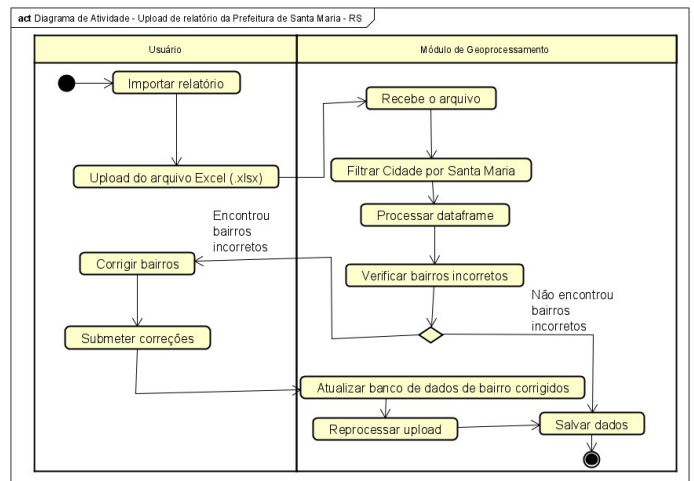


Figura 2: Diagrama de atividade da funcionalidade Upload. [24]

A. Upload de relatório

Os dados utilizados no sistema são exportados do sistema de saúde PMSM por meio de um arquivo nomeado como relatório sintético por localidade, no formato .xlsx (Excel). Esse documento contém informações sobre a quantidade de casos de CID por bairro.

Upload de Relatório



Figura 3: Seleção de datas para renomear arquivo upado e vincular data ao banco de dados junto com o conjunto de dados do arquivo. [24]

Conforme ilustrado no Diagrama de Atividades (Figura 2), ao fazer o upload do relatório, o usuário define a data do relatório exemplificado na figura 3, o sistema recebe o arquivo e o transforma em um dataframe. Em seguida, processa o dataframe, realizando a filtragem específica para a cidade de Santa Maria, ajusta os nomes das colunas e organiza o conteúdo para evitar possíveis erros.

O processo de tratamento dos dados inclui a padronização dos nomes das colunas para facilitar a manipulação e a filtragem por localidade, garantindo que apenas informações

de Santa Maria sejam processadas. Além disso, é realizada a verificação e correção dos nomes dos bairros. Nomes de bairros incorretos ou não padronizados são identificados e corrigidos com base em uma tabela de referência.

Para tanto, é utilizado o arquivo de Malha de Setores Censitários do IBGE, manipulado para utilizar os bairros e distritos de Santa Maria, chamado “Bairrosedistritos.json”. Neste arquivo contém a coluna com nome de bairros e distritos corretos. Caso sejam encontrados bairros incorretos, o sistema encaminha para que o usuário faça as correções necessárias conforme mostrado na figura 4 e reprocessa o *upload* para garantir a integridade e precisão dos dados armazenados.

Após o processamento e correção dos dados, estes são salvos em um banco de dados. O sistema utiliza um modelo de banco de dados relacional onde os dados do relatório, incluindo as correções dos bairros, são armazenados em tabelas apropriadas. O modelo de dados inclui as classes Relatório, Casos e CorrecaoBairro, definidas conforme a figura 5.

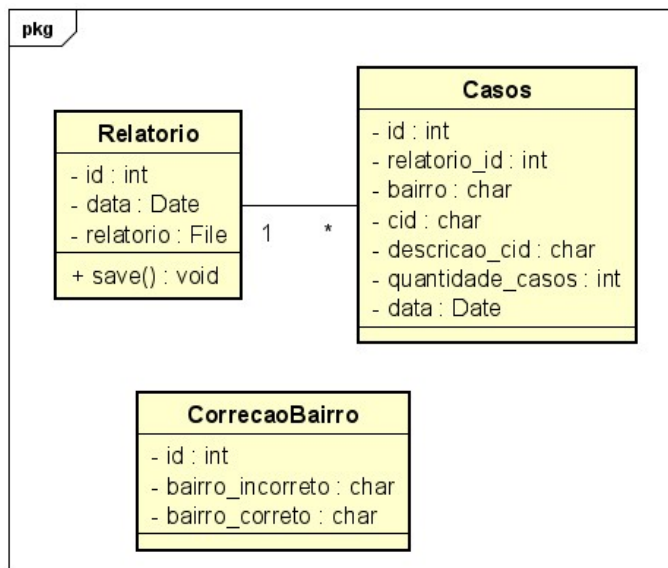


Figura 5: Diagrama de classes. [24]

Correção de nomes dos bairros

Para ajudar na correção, visite o site com a [Lei Complementar Nº 42/2006](#) e [Openstreetmap](#).

| | |
|---|---|
| ARCO IRES | QUADRA 107 |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| CACHACRA DAS FLORES | RENASCENCA-PATRONATO |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| COAB FERNANDO FERRARI | SAO VALENTIM |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| DISTRITO DE BOCA DOD MONTE | SIPRIANO DA ROCHA |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| KM 2 | STE DE DEZEMBRO |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| N S DE LURDES | VILA ALMEIDA |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| PAIMS | VILA CAUDRO |
| <input type="text" value="Selecione o bairro correto"/> | <input type="text" value="Selecione o bairro correto"/> |
| PINEIRO MACHADO | <input type="button" value="Salvar Mapeamentos"/> |
| <input type="text" value="Selecione o bairro correto"/> | |

Figura 4: Página com os nomes que devem ser corrigidos. [24]

O modelo “Relatorio” armazena os dados do relatório, incluindo o arquivo carregado e a data do relatório. O modelo “Casos” armazena as informações dos casos por bairro, CID e quantidade de casos, vinculados ao relatório correspondente. O modelo “CorrecaoBairro” armazena as correções dos nomes dos bairros, mapeando nomes incorretos e os adequando para os corretos. Assim, o sistema garante que os dados estejam corretamente estruturados e disponíveis para consultas e visualizações posteriores.

Para corrigir manualmente os nomes dos bairros que estavam incorretos, bem como para substituir nomes de vilas e ruas pelos nomes adequados dos bairros, utilizou-se o documento Lei Complementar Nº 42, de 29 de dezembro de 2006 [25]. Essa lei fornece a lista de vilas e loteamentos pertencentes a cada bairro. Além disso, foram realizadas pesquisas em ferramentas de mapas, como o *OpenStreetMap*, e consultas no Google.

A Figura 6 ilustra a quantidade de registros de bairros corretos e incorretos encontrados nos relatórios de janeiro antes de aplicar a correção. Sem tal correção, antes de salvar os arquivos, haveria uma perda aproximada de 57,57% de casos que não seriam mapeados, o que representa um impacto significativo na qualidade e na integridade dos dados.

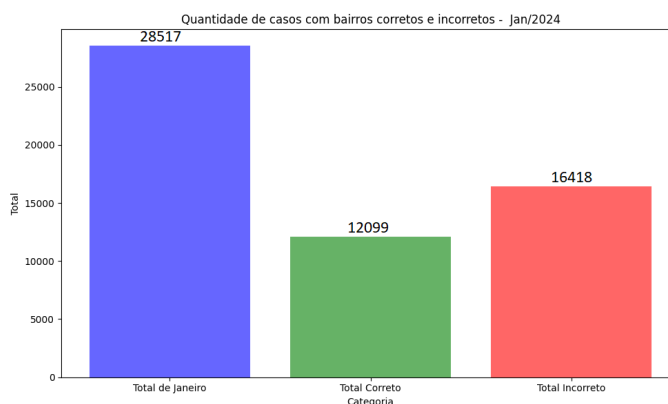


Figura 6: Gráfico com número de casos com bairros corretos e incorretos antes da correção manual. [24]

A Figura 7 ilustra a diferença substancial entre os dados inicialmente incorretos e os dados após a aplicação das correções, destacando a importância desse processo.

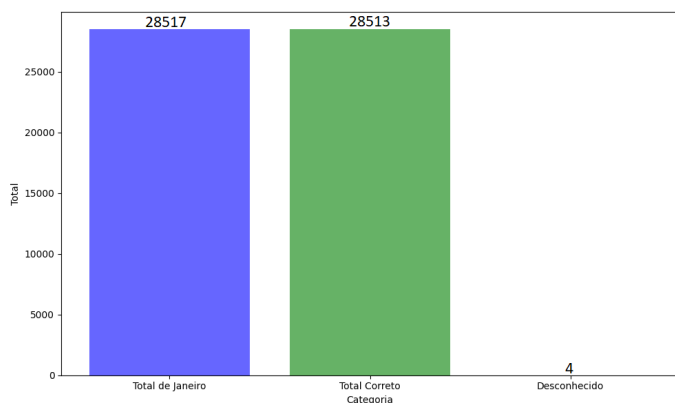


Figura 7: Gráfico com número de casos com bairros corretos e bairros que não foram possível identificar após correção manual. [24]

B. Listagem de relatórios

Na página de relatórios (Figura 8), constam todos os relatórios carregados pelo usuário e dois botões com as opções de *download* e *deletar*. A opção de *deletar* é utilizada para excluir o arquivo nos casos em que ocorra a adição de um arquivo na data errada. Ao selecionar o arquivo carregado que se deseja deletar, todos os casos vinculados a esse arquivo e o próprio arquivo são removidos do banco de dados e do sistema, deixando habilitada na página de “Upload de relatório” a opção de carregar o arquivo correto na data desejada.

Já a opção de *download* serve para obter o arquivo carregado no sistema. Isso permite ao usuário recuperar e acessar os dados originais sempre que necessário, facilitando a revisão ou a utilização dos dados fora do ambiente do sistema.

Lista de Relatórios

| Data | Arquivo | Ações |
|-----------------------|----------------------------|--|
| 1 de Janeiro de 2024 | relatorios/2024-01-01.xlsx | Download Deletar |
| 2 de Janeiro de 2024 | relatorios/2024-01-02.xlsx | Download Deletar |
| 3 de Janeiro de 2024 | relatorios/2024-01-03.xlsx | Download Deletar |
| 4 de Janeiro de 2024 | relatorios/2024-01-04.xlsx | Download Deletar |
| 5 de Janeiro de 2024 | relatorios/2024-01-05.xlsx | Download Deletar |
| 6 de Janeiro de 2024 | relatorios/2024-01-06.xlsx | Download Deletar |
| 7 de Janeiro de 2024 | relatorios/2024-01-07.xlsx | Download Deletar |
| 8 de Janeiro de 2024 | relatorios/2024-01-08.xlsx | Download Deletar |
| 9 de Janeiro de 2024 | relatorios/2024-01-09.xlsx | Download Deletar |
| 10 de Janeiro de 2024 | relatorios/2024-01-10.xlsx | Download Deletar |

Figura 8: Página que lista os relatórios que foram upados no sistema. [24]

V. RESULTADOS OBTIDOS

Após o upload e a correção manual da nomenclatura dos bairros, obteve-se uma base de dados com nomes de bairros padronizados, o que facilita futuros *uploads* e mantém a consistência dos dados. Essa base de dados também pode ser utilizada para outros tipos de relatórios que o sistema de saúde da Prefeitura de Santa Maria possua.

Casos de CID por Bairro

Selecione o CID:

A90 - DENGUE [DENGUE CLÁSSICO]

Data Inicial:

01/01/2024

Data Final:



Figura 9: Campo para realizar a filtragem de CID e data. [24]

No sistema, ao acessar a página “Gerar Mapa”, o usuário pode filtrar por CID e selecionar um intervalo de datas (Figura 9) para visualizar as ocorrências do CID escolhido. Com esses filtros, o sistema une os dados filtrado com o arquivo “Bairros-sm.json” através dos nomes dos bairros. Com essa união, é criando um dataframe, com as colunas: Nome do bairros, Código do cid, descrição do cid, quantidade de casos e dados geoespaciais, gerando um mapa (Figura 10) da cidade utilizando a biblioteca *matplotlib* que mostra a distribuição dos casos por bairro, facilitando a identificação de áreas com maior incidência da doença. Além disso, são gerados dois gráficos utilizando a biblioteca *matplotlib*: um gráfico de barras (Figura 11) que exibe a quantidade de casos por bairro e um gráfico de linhas (Figura 12) que mostra a ocorrência dos casos ao longo do tempo.

Essa funcionalidade é fundamental para visualizar de forma clara e intuitiva a distribuição espacial e temporal das ocorrências de doenças, permitindo uma análise mais eficiente e tomada de decisões baseada em dados. O gráfico de barras facilita a identificação dos bairros afetados, enquanto o gráfico de linha permite monitorar o progresso ou a regressão de casos ao longo do tempo na cidade.

VI. CONSIDERAÇÕES FINAIS

Este trabalho mostrou a importância da padronização e consistência dos dados, exemplificada pela correção manual dos nomes dos bairros, que assegurou a precisão e a confiabilidade das informações processadas. Com uma base de dados bem estruturada, foi possível gerar relatórios precisos e mapas detalhados que auxiliam na visualização espacial das ocorrências de doenças.

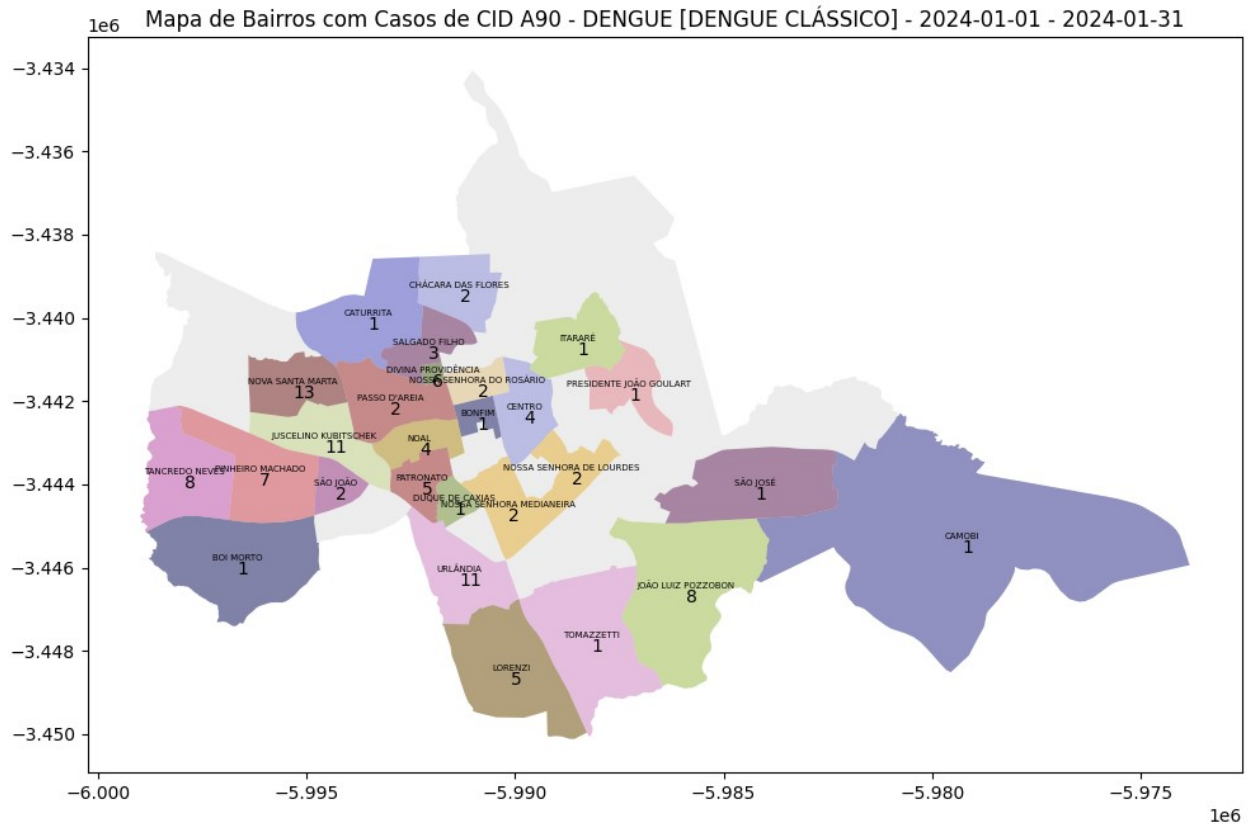


Figura 10: Mapa Gerado - CID A90 - Dengue - 01-01-2024 - 31-01-2024. [24]

Quantidade de Casos por Bairro

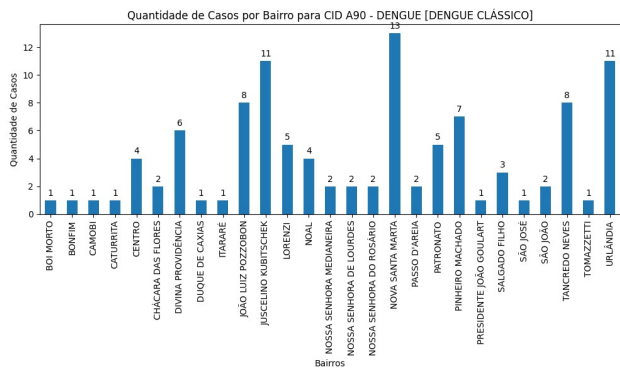


Figura 11: Gráfico com número de casos de CID A90 - Dengue por bairro - 01-01-2024 - 31-01-2024. [24]

Quantidade de Casos por Dia

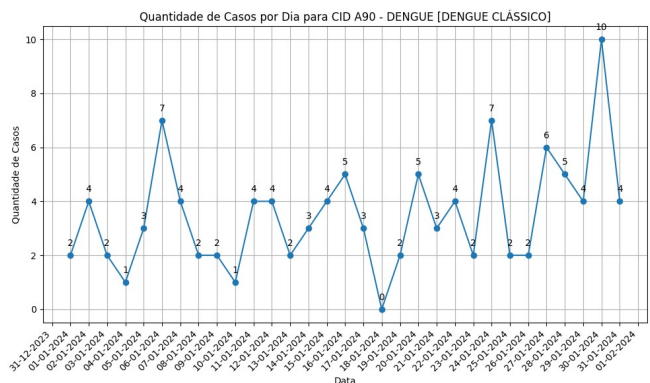


Figura 12: Gráfico com número de casos por dia de CID A90 - Dengue - 01-01-2024 - 31-01-2024. [24]

Ao integrar tecnologias de geoprocessamento com dados de saúde, conseguimos criar uma ferramenta para a visualização de doenças e epidemias. O uso de tecnologias de geoprocessamento não só aumentou a eficiência na visualização de surtos emergentes, mas também proporcionou uma vigilância contínua e dinâmica, permitindo adaptações rápidas conforme novos dados são inseridos no sistema. Isso reforça a relevância

do módulo como uma ferramenta precisa. Os resultados obtidos destacam a viabilidade e a eficácia de integrar dados de saúde e geoprocessamento, promovendo uma abordagem proativa na prevenção e controle de doenças.

A partir do que foi abordado neste trabalho, fica como sugestão para trabalhos futuros, a utilização de outras bibliotecas, como Plotly e Folium, para trazer interatividade aos gráficos

e mapas. Geração de mapas incluindo os distritos de Santa Maria. Além disso, a implementação de uma API que realize a integração entre o sistema de saúde da prefeitura e o módulo para automatizar a obtenção de dados, tornando o processo ágil. Também seria relevante automatizar a correção de dados utilizando a base de dados deste trabalho e técnicas de *machine learning*.

REFERÊNCIAS

- [1] “Doenças de Transmissão Hídrica e Alimentar (DTHA)”. Em: (2023). Disponível em <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/dtha>>.
- [2] Márian Oleques Pires. “SOFTWARE DE ASSISTÊNCIA AO PRÉ-NATAL: FERRAMENTA DE APOIO À GESTÃO DA INFORMÁTICA EM SAÚDE”. Em: (2022). Disponível em <http://www.tede.universidadefranciscana.edu.br:8080/bitstream/UFN-BDTD/1133/5/Dissertacao_MarianOlequesPires_ssemAssinaturas.pdf>.
- [3] Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia Brasil. Ministério da Saúde. Secretaria de Ciência. *Programa Pesquisa para o SUS: gestão compartilhada em saúde: Diretrizes Técnicas*. 5ª ed. 64 p. : il. Brasília: Ministério da Saúde, 2014.
- [4] V. v. et al. Vasconcelos. *CAPÍTULO 10: Geoprocessamento aplicado à gestão de riscos*. Editora UFABC, 2022.
- [5] Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Emergências em Saúde Pública. *Guia para investigações de surtos ou epidemias*. 2ª ed. 64 p. : il. Brasília: Ministério da Saúde, 2022.
- [6] “Atenção Básica ou Primária - Principal porta de entrada para o Sistema Único de Saúde (SUS)”. Em: (2023). Disponível em <<https://saude.rs.gov.br/atencao-basica-ou-primaria-principal-porta-de-entrada-para-o-sistema-unico-de-saude-sus>>.
- [7] F. Louzada Neto. *Introdução à Análise de Dados*. ed. Blucher, 2004.
- [8] F. Louzada Neto. *Técnicas de Processamento de Dados*. ed. LTC, 2018.
- [9] A. M. V. Monteiro. *Geoprocessamento para Análise Ambiental*. ed. Oficina de Textos, 2018.
- [10] A. M. V. Monteiro. *Fundamentos de Geoinformação*. ed. Embrapa, 2020.
- [11] Python Software Foundation. “FAQ geral sobre Python”. Em: (2023). Disponível em <<https://docs.python.org/3/faq/general.html#what-is-python>>.
- [12] “Pandas Documentation”. Em: (2023). Disponível em <<https://pandas.pydata.org/docs/index.html>>.
- [13] “Numpy”. Em: (2023). Disponível em <<https://numpy.org/>>.
- [14] “Matplotlib”. Em: (2023). Disponível em <<https://matplotlib.org/stable/index.html>>.
- [15] “Geopandas”. Em: (2023). Disponível em <<https://geopandas.org/en/stable/>>.
- [16] “Django”. Em: (2024). Disponível em <<https://www.djangoproject.com/start/overview/>>.
- [17] “HTML”. Em: (2024). Disponível em <<https://developer.mozilla.org/pt-BR/docs/Web/HTML/>>.
- [18] “Sqlite3”. Em: (2024). Disponível em <<https://www.sqlite.org/index.html>>.
- [19] Stephen Klosterman. *Projetos de Ciência de Dados com Python*. NOVATEC, 2019.
- [20] Esri. “Shapefiles”. Em: (2023). Disponível em <<https://enterprise.arcgis.com/pt-br/portal/latest/use/shapefiles.htm>>.
- [21] Wagner Vidal Xavier da Silva. “O Impacto do Uso da Biblioteca Pandas do Python como Ferramenta de Análise de Dados Referente aos Casos Graves de COVID-19 Notificados pela Secretaria Estadual de Saúde do Recife entre Julho de 2021 a Junho de 2022”. Em: (2022). Disponível em <https://www.researchgate.net/profile/Wellington-Dos-Santos/publication/369591716_Anaais_dov_Isimposio_de_Inovacao_e_MEngenharia_Biomedica_SABIO_2022/links/64239e2b92cfd54f84388297/Anaais-do-VI-Simposio-de-Inovacao-em-Engenharia-Biomedica-SABIO-2022.pdf> page = 23 >.
- [22] Eduardo Ottobelli Chiell Eliseu Luiz Kuiava Victor Antônio Kuiava e Daniel Navarini. “Sistema computacional automático para geração de relatórios epidemiológicos a partir de dados do datasus”. Em: (2020). Disponível em <<https://ojs.brazilianjournals.com.br/ojs/index.php/BJHR/article/view/20855/16650>>.
- [23] Angélica Félix et al DE CASTRO. “Utilização da linguagem Python na análise de dados de dengue em Recife - capital do estado de Pernambuco - Nordeste do Brasil e definição de linha de tendência usando Machine Learning.” Em: (2022). Disponível em <<https://tinyurl.com/4n3cdb8j>>.
- [24] Dos Autores. *Diagrama desenvolvido pelos autores do trabalho*. 2024.
- [25] RS Santa Maria. “LEI COMPLEMENTAR Nº 42, DE 29 DE DEZEMBRO DE 2006”. Em: (2006). Disponível em <<http://leismunicipa.is/tdben>>.