

## DESENVOLVIMENTO DE SITE PARA GEOVISUALIZAR CASOS DE MICROCEFALIA REGISTRADOS NO BRASIL, NO PERÍODO DE 2015 A 2024 COM USO DO PYTHON

Maria Clara Leite de Castro<sup>1</sup>, Marilée Patta<sup>2</sup>

<sup>1</sup>Aluna do Curso de Sistemas de Informação da Unimontes, Montes Claros, Brasil  
(mclaraleiteuni@gmail.com)

<sup>2</sup>Professora do Curso de Sistemas de Informação da Unimontes, Montes Claros, Brasil

**Resumo:** Desenvolveu-se em Python a geovisualização dos casos de microcefalia no Brasil (2015 - 2024). Com dados do RESP-Microcefalia e do SINASC (DATASUS), e com as bibliotecas Pandas, Geobr, Folium e Plotly, além de HTML, CSS e Javascript, criou-se interface com mapas coropléticos e gráficos dinâmicos reativos a tempo, espaço, gênero e etiologia, apresentando-se a taxa de microcefalia por 10 mil nascidos vivos em escala municipal. Constatou-se que o pico epidemiológico foi em 2015-2016, com forte concentração espacial nos estados do Nordeste e equilíbrio entre sexos.

**Palavras-chave:** Mapeamento da Microcefalia com Python; Geovisualização da Microcefalia no Brasil;

### INTRODUÇÃO

A epidemia de microcefalia surgiu no país em 2015. Associada ao vírus Zika, representou um grande desafio para a saúde pública brasileira. O aumento súbito de casos evidenciou não apenas a gravidade da situação, mas também a necessidade urgente de ferramentas que permitissem uma compreensão mais clara da distribuição temporal e espacial desses casos. Diante desse problema, o Ministério da Saúde instituiu o sistema de Registro de Eventos em Saúde Pública – Microcefalia (RESP-Microcefalia) em novembro de 2015. Criou uma base de dados para monitorar e investigar casos suspeitos e confirmados. No entanto, a análise desses dados por meio de métodos tradicionais apresenta limitações para a compreensão integral do fenômeno. Sendo assim, a geovisualização de dados é uma ferramenta fundamental para transformar informações complexas em conhecimento acessível e compreensível. Dessa forma, por meio de técnicas de visualização espacial, torna-se possível identificar padrões, desigualdades regionais e tendências temporais que podem auxiliar tanto a pesquisa acadêmica quanto a formulação de políticas públicas de saúde.

Neste contexto, refletiu-se sobre os impactos significativos que a epidemia gerou, especialmente em regiões socioeconomicamente vulneráveis. Nesses cenários, pesquisadores e gestores de saúde ainda encontram dificuldades para explorar os dados oficiais disponíveis de forma visual e intuitiva. Isso justifica a necessidade de transformar tais registros

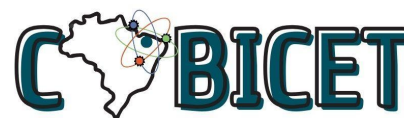
públicos em visualizações acessíveis por meio de uma aplicação web voltada à análise exploratória dos casos ocorridos entre 2015 e 2024.

A geovisualização com uso do Python busca responder às seguintes questões centrais: onde se concentraram os maiores registros de microcefalia no Brasil entre 2015 e 2024, e quais indicadores evidenciam os locais de maior acometimento, orientando assim tanto o desenvolvimento técnico da solução quanto a seleção das variáveis e representações visuais mais adequadas para a análise.

Assim, desenvolveu-se um site para geovisualizar os casos de microcefalia registrados no Brasil no período de 2015 a 2024 com uso do Python, o que se desdobra nos seguintes objetivos específicos: construir visualizações interativas para análise temporal dos casos; desenvolver gráficos dinâmicos para análise por gênero, região e classificação; implementar um mapa geográfico para visualização da distribuição espacial dos registros; e permitir a exploração interativa dos dados por meio de filtros na interface web.

A microcefalia é uma condição neurológica rara caracterizada por um tamanho reduzido da cabeça, geralmente decorrente de um desenvolvimento incompleto do cérebro. Essa anomalia pode estar associada a várias causas, incluindo infecções durante a gestação, como é o caso da infecção pelo vírus Zika (TEIXEIRA et al., 2016).

Em 2015, observou-se um aumento incomum de nascimentos com microcefalia, especialmente na



região Nordeste do Brasil, o que chamou a atenção das autoridades de saúde (TEIXEIRA et al., 2016). Antes disso, o vírus Zika era considerado de baixa gravidade (TEIXEIRA et al., 2016), mas a correlação com o aumento de casos de malformações levou especialistas a suspeitarem de sua associação com a microcefalia (LOWE et al., 2018).

A partir disso, o Ministério da Saúde decretou estado de emergência em saúde pública em novembro de 2015, seguido por declaração semelhante da Organização Mundial da Saúde em fevereiro de 2016.

A microcefalia, geralmente causada por alterações no desenvolvimento cerebral, pode trazer consequências variáveis para a saúde, desde atrasos leves no desenvolvimento até deficiências motoras e intelectuais graves, como paralisia cerebral. Essas condições estão associadas a anomalias como calcificações cerebrais, ventriculomegalia e distúrbios de migração neuronal (SCHULER-FACCINI et al., 2016).

Além de afetar o sistema nervoso, a microcefalia pode provocar dificuldades de alimentação, crises convulsivas e problemas sensoriais. Sua origem pode estar relacionada a fatores genéticos, ambientais ou infecções congênitas, como as causadas pelo vírus zika, que se destacaram em acometimentos de surtos (SCHULER FACCINI et al., 2016).

O aumento dos casos de microcefalia no Brasil exigiu uma resposta urgente das autoridades. Assim, o Ministério da Saúde instituiu o Registro de Eventos em Saúde Pública (RESP) em novembro de 2015, para monitoramento e notificação dos casos de microcefalia e outras alterações congênitas relacionadas ao vírus Zika. Diante da crise, foram adotadas medidas emergenciais, como a notificação obrigatória dos casos suspeitos, o fortalecimento das ações de combate ao mosquito transmissor e a ampliação da rede de cuidados para crianças com necessidades especiais.

As ações de prevenção passaram a incluir o uso de repelentes, roupas protetoras, mosquiteiros, além de campanhas de conscientização sobre o controle do mosquito e a importância do saneamento básico.

O Departamento de Informática do Sistema Único de Saúde (DATASUS), criado em 1991, é responsável por prover soluções tecnológicas e sistemas de informação que apoiam a gestão do SUS. Assim, ele assegura a produção, integração e disseminação de dados essenciais ao planejamento e controle das ações em saúde. Além disso, ele reúne diversos bancos de dados e sistemas de informação, como o Sistema de Informações sobre Nascidos Vivos (SINASC), o Sistema de Informações Hospitalares (SIH), o Sistema de Informação de Agravos de

Notificação (SINAN), entre outros. Essas bases são amplamente utilizadas para análises estatísticas, estudos epidemiológicos e planejamento de ações de saúde pública.

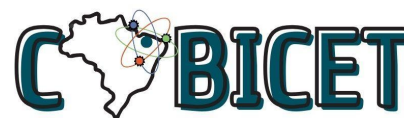
Com a criação do sistema de Registro de Eventos em Saúde Pública – Microcefalia permitiu o registro detalhado de dados clínicos, laboratoriais e epidemiológicos, contribuindo para a compreensão da magnitude do problema e para a formulação de políticas públicas direcionadas. O Sistema de Informações sobre Nascidos Vivos (SINASC) foi implantado oficialmente no Brasil a partir de 1990, com o objetivo de coletar, processar e disponibilizar dados sobre todos os nascimentos ocorridos em território nacional.

O principal instrumento de coleta do sistema, denominado Declaração de Nascido Vivo (DN), é utilizado para o registro oficial do nascimento. Esse documento é preenchido por profissionais de saúde ou parteiras tradicionais em casos de partos assistidos, sendo posteriormente digitado e consolidado pelas Secretarias Municipais de Saúde.

Em relação à questão tecnológica, discorre-se sobre geovisualização, Python, Pandas, Geobr, Folium, Plotly, HTML, CSS e Javascript.

A geovisualização pode ser definida como a visualização de dados que possuem um componente geográfico, frequentemente usada para dados com estruturas complexas que envolvem espaço, tempo e vários atributos temáticos (CADENAS, 2016). Dessa forma, a análise espacial é de grande importância para a área da saúde e a epidemiologia, sendo utilizada para diversos fatores, como: 1) Exploração de dados de saúde: Ferramentas de geovisualização permitem a exploração e interação com dados espaciais de forma dinâmica e em tempo real (ROBINSON, 2005). Isso é crucial para pesquisadores em saúde explorarem dados multivariados espaço-temporais; 2) Identificação de desigualdades e padrões: A geovisualização, por meio da análise espacial de dados, pode revelar padrões na distribuição de indicadores de saúde e evidenciar desigualdades regionais (PIMENTEL et al., 2012). Por exemplo, um estudo sobre desigualdades em saúde no Brasil utilizou geovisualização para demonstrar a existência de disparidades regionais no acesso e utilização de serviços de saúde, favoráveis às regiões Sul e Sudeste (PIMENTEL et al., 2012); 3) Apoio à tomada de decisão e planejamento: Mapas temáticos podem fornecer assistência no raciocínio visual e são amplamente utilizados para facilitar a tomada de decisão em diversas áreas. Na saúde, isso inclui o planejamento urbano e a gestão de desastres.

O Python é uma linguagem de programação de alto nível, versátil e amplamente adotada em aplicações



de ciência, tratamento e geovisualização de dados (VAN ROSSUM; DRAKE, 2009). No contexto deste trabalho, foi utilizada uma cadeia de bibliotecas específicas para realizar desde a conversão de arquivos proprietários de saúde até a renderização cartográfica final.

Pandas é a principal biblioteca para análise e manipulação de estruturas de dados em Python. Através do uso de *DataFrames*, o Pandas permite a leitura massiva de arquivos, limpeza de inconsistências, tratamento de dados faltantes, agrupamentos espaciais e temporais, além da junção (merge) entre as bases do RESP-Microcefalia e do SINASC para o cálculo preciso da taxa de incidência (MCKINNEY, 2010).

O Geobr é um pacote computacional desenvolvido pelo Instituto de Pesquisa Econômica Aplicada (IPEA) que permite o acesso direto às malhas geográficas oficiais do Brasil, disponibilizadas pelo IBGE em formato digital.

O Folium é uma biblioteca baseada na tecnologia *JavaScript Leaflet.js* que permite a criação de mapas interativos utilizando o ambiente Python. O Folium foi aplicado para construir o mapa coroplético, viabilizando a renderização de camadas informativas, controle de *zoom*, escalas de cores integradas e recursos de interatividade web, como as caixas de texto informativas acionadas pela ação do usuário.

O Plotly é uma biblioteca gráfica de código aberto que permite a construção de gráficos interativos e dinâmicos com alta qualidade visual. Pyreaddbc e Dbfread são ferramentas técnicas essenciais para a etapa de engenharia de dados.

O *HyperText Markup Language - HTML* é uma linguagem de marcação que define a estrutura de uma página na web. Ele funciona como a base, organizando os elementos que compõem o site, como textos, imagens, vídeos, links, botões e tabelas. É através do HTML que se estabelece quais informações serão exibidas e como elas estão hierarquicamente organizadas na página (MELO, 2021).

O *Cascading Style Sheets - CSS* é responsável por definir a aparência do site, sendo possível personalizar cores, fontes, tamanhos, alinhamentos, espaçamentos e até criar animações. O CSS separa o conteúdo (feito em HTML) da parte visual, facilitando a manutenção, tornando o desenvolvimento mais organizado (MELO, 2021).

O *JavaScript* é a linguagem que adiciona comportamento e interatividade ao site. Por meio dele, é possível fazer com que a página responda a ações do usuário, como cliques, preenchimento de formulários, rolagem e seleção de elementos. Além disso, ele permite criar funcionalidades como mapas

interativos, gráficos dinâmicos, filtros, carregamento de dados sem recarregar a página, entre outros (MDN Web Docs, 2025).

## MATERIAL E MÉTODOS

A unidade de análise da pesquisa são os casos de microcefalia registrados nos Municípios do Brasil no período de 2015 a 2024.

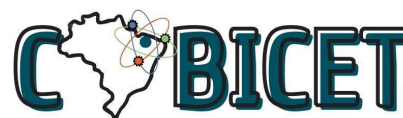
Os dados extraídos do RESP-Microcefalia constituem a base principal para a análise espacial e temporal da distribuição dos casos no Brasil entre os anos de 2015 e 2024. O SINASC foi utilizado como fonte complementar de dados para o cálculo do número de nascidos vivos por município, servindo como denominador para a construção da taxa de incidência de microcefalia. Dessa forma, o cruzamento entre os dados do SINASC e do sistema RESP-Microcefalia permitiu uma análise mais precisa da distribuição dos casos, considerando o tamanho populacional de cada localidade.

A técnica utilizada é a documental, utilizando dados secundários e públicos de duas bases de dados do Departamento de Informática do Sistema Único de Saúde (DATASUS): RESP-Microcefalia (base de dados que contém as notificações de casos suspeitos e confirmados de microcefalia no período de 2015 a 2024) e SINASC (dados coletados a partir das Declarações de Nascido Vivo (DN), utilizados para quantificar o volume total de nascimentos por município e servir como o denominador estatístico).

Uma vez que o DATASUS distribui os microdados do SINASC compactados no formato proprietário *Data Base Compressed - DBC*, a biblioteca *pyreaddbc* foi utilizada para descompactá-los em formato DBF, enquanto a *dbfread* permitiu a leitura e conversão final dessas estruturas para arquivos CSV manipuláveis.

Do banco de dados de Registro de Eventos em Saúde Pública, foram selecionadas e tratadas as seguintes variáveis categóricas e numéricas:

- **CODMUNRES:** Código de 6 dígitos que identifica o município onde reside a mãe do recém-nascido. Esta variável foi utilizada como chave primária para o merge com a malha do IBGE;
- **CLASSIFIN:** Atributo qualitativo nominal utilizado na composição do mapa de geovisualização, através da realização de uma filtragem de todos registros para considerar exclusivamente aqueles rotulados como "Confirmado". Para os gráficos dinâmicos do painel, mantiveram-se as demais classificações ("Provável", "Em investigação", "Inconclusivo", "Descartado" e "Excluído");
- **SEXO:** Variável categórica mapeada para os valores "Masculino", "Feminino" e "Ignorado",



utilizada para avaliar a distribuição biológica dos casos de microcefalia;

- REGIAORES e REGIAONOT: Indicadores das regiões de residência e notificação (Norte, Nordeste, Sudeste, Sul e Centro-Oeste), utilizados para a análise espacial da evolução da epidemia;
- ANO\_NOT: Atributo temporal que registra o ano em que o evento suspeito foi inserido no sistema, sendo usado para a análise temporal da evolução dos casos de microcefalia.

Do Sistema de Informações sobre Nascidos Vivos, a extração concentrou-se na obtenção do denominador demográfico populacional de recém-nascidos:

- CODMUNRES: Utilizado para agregar e contabilizar o volume total de partos bem-sucedidos em cada localidade;
- ANO\_NASC (Ano de Nascimento): Variável temporal extraída a partir do tratamento dos metadados dos arquivos anuais via expressões regulares (Regex), garantindo o pareamento anual correto com os registros do RESP.

O indicador desenvolvido para a geovisualização é a Taxa de Incidência de Microcefalia, calculada no nível de granularidade municipal. Na literatura epidemiológica aplicada a anomalias congênitas, adota-se o total de nascidos vivos como a população sob risco real de manifestação do evento, evitando distorções que ocorreriam caso fosse utilizada a população geral de habitantes. A métrica foi padronizada para uma base de 10.000 nascidos vivos, expressa matematicamente pela equação:

$$\text{Taxa de Incidência (município)} = \frac{\text{Casos Confirmados (RESP)}}{\text{Total de nascidos vivos (SINASC)}} \times 10.000 \quad (1)$$

A biblioteca Geobr foi fundamental para carregar diretamente no código os polígonos geométricos de todos os municípios brasileiros, servindo de base para o mapeamento espacial (PEREIRA et al., 2019).

A taxa de incidência foi indexada aos polígonos geoespaciais obtidos via biblioteca Geobr, permitindo que a categorização visual pelo algoritmo do Folium dos municípios brasileiros através de um gradiente cromático (colormap), dividindo a severidade do acometimento em níveis visualmente identificáveis na interface do site.

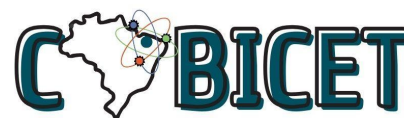
Com os dados em formato acessível, a etapa de higienização concentrou-se em garantir a consistência das variáveis e mitigar o impacto de dados faltantes ou corrompidos. Utilizando a biblioteca Pandas, os seguintes procedimentos foram executados: Padronização de Datas: Conversão de strings e inteiros truncados para o formato padrão de

data (datetime), tratando inconsistências em colunas críticas como data de notificação e data de nascimento; Mapeamento de Dicionários: Substituição de códigos numéricos nominais do DATASUS por suas respectivas descrições textuais baseadas nos dicionários oficiais (ex: conversão das categorias da variável CLASSIFIN para termos como "Confirmado", "Descartado" ou "Em investigação"); Tratamento de Missing Values: Agrupamento de registros nulos, em branco ou inconsistentes sob o rótulo unificado de "Ignorado".

A última etapa consistiu na modelagem espacial e construção da interface de visualização para o usuário final: Obtenção de Malhas Geográficas: Utilização da biblioteca Geobr para extração direta das geometrias político-administrativas dos municípios brasileiros mapeadas pelo IBGE; Compatibilização e Otimização Espacial: Alinhamento dos códigos de município para uma chave comum de 6 dígitos (utilizando preenchimento de zeros à esquerda com zfill) e aplicação de algoritmos de simplificação de geometria para reduzir o peso computacional do arquivo final; Construção do Mapa Coroplético: Geração do mapa interativo com a biblioteca Folium, configurando escalas de cores gradientes dinâmicas, inserção de legendas em HTML/CSS e estilização de caixas de informação ao passar o mouse (tooltips); Painel de Gráficos e Filtros: Integração do mapa a uma página web estruturada em HTML5, estilizada com CSS3 e dinamizada com JavaScript. Os gráficos complementares (temporal, por gênero, por região e classificação) foram renderizados de forma reativa a filtros de seleção utilizando a biblioteca Plotly.js.

A abordagem analítica fundamentou-se nos preceitos da Estatística Descritiva e da Análise Exploratória de Dados (AED), processadas inteiramente de forma computacional. O tratamento dos dados quantitativos e a geração dos indicadores foram conduzidos sob dois eixos principais de investigação: 1) Eixo Temporal: Avaliação do comportamento sequencial das notificações ao longo do recorte cronológico delimitado (2015 a 2024). Utilizou-se o cálculo de frequências absolutas e relativas anuais para identificar períodos de pico epidêmico, estabilização ou declínio no registro de casos; 2) Eixo Espacial: Mapeamento da distribuição geográfica das ocorrências no território nacional, com granularidade estendida desde o nível macrorregional até o nível municipal.

O tratamento estatístico evitou distorções demográficas por meio do cálculo da Taxa de Incidência por 10.000 nascidos vivos, permitindo a comparação equitativa do risco epidemiológico entre municípios de diferentes portes populacionais. Complementarmente, foram aplicadas segmentações por categorias nominais (como sexo biológico e



classificação final do caso) para estratificar o perfil demográfico dos agravos.

Por fim, a difusão e a apresentação dos resultados foram materializadas na interface de uma aplicação web interativa, desenvolvida em arquitetura baseada em arquivos de dados abertos. Visando otimizar o consumo de memória e dispensar a necessidade de uma infraestrutura complexa de servidores de bancos de dados relacionais (como SGDBs tradicionais), o sistema consome os dados de saúde previamente tratados e estruturados em formato CSV.

A renderização das informações ocorre diretamente no lado do cliente. A linguagem HTML5 estrutura os componentes, as folhas de estilo em CSS3 ditam a identidade visual e o posicionamento fluido dos painéis, e os algoritmos em JavaScript (em conjunto com as bibliotecas Plotly.js e Folium, através do mapa embutido via iframe) realizam a leitura dinâmica das tabelas. Esse arranjo viabiliza que filtros interativos de tempo, gênero e região atualizem os gráficos, fornecendo um ambiente intuitivo para a exploração de cenários por gestores e pesquisadores.

## RESULTADOS E DISCUSSÃO

Necessitou-se realizar a transformação do formato dos dados do DATASUS. Os arquivos do SINASC e do RESP são disponibilizados publicamente na extensão .dbc (*Data Base Compressed*), sendo incompatíveis com as principais bibliotecas de análise de dados em Python. Assim, foi desenvolvido um script que automatiza o processo de conversão. Inicialmente, utilizou-se a função `dbc2dbf()` da biblioteca `Pyreaddbc`, responsável por descompactar cada arquivo .dbc de uma pasta e gerar arquivos no formato .dbf. Em seguida, a biblioteca `dbfread` foi utilizada para ler e interpretar os registros contidos nesses arquivos. Posteriormente, a biblioteca `Pandas` converteu essas informações para o formato *DataFrame*, realizando a exportação dos dados em arquivos *Comma-Separated Values* - .csv. O processo foi automatizado pela função `glob()` da biblioteca `Glob`, responsável por identificar todos os arquivos .dbc presentes na pasta de origem. Além disso, as pastas destinadas aos arquivos .dbf e .csv foram criadas automaticamente utilizando a função `makedirs()` da biblioteca `Os`.

Com os dados do SINASC e RESP convertidos para o formato acessível, fez-se necessário fazer o tratamento e padronização da base de pesquisa. Depois da conversão, os dados do RESP-Microcefalia foram distribuídos em arquivos CSV separados por ano. Dessa forma, cada um dos arquivos foi lido pela função `glob` da biblioteca `Glob`, e todos foram consolidados em único *Dataframe* por meio da função `concat()` da biblioteca `Pandas`. Após a concatenação, foram eliminados os registros

completamente vazios e valores ausentes foram preenchidos com strings vazias, preparando o conjunto de dados para a etapa de mapeamento.

Nos dados do DATASUS há campos *datas* como inteiros de 8 dígitos no formato DDMMAAAA, sem separadores, o que inviabiliza sua interpretação como tipo *datetime* pelo `Pandas`. Dessa forma, foi desenvolvida uma função que recebe esse valor e o converte para *string*, aplica preenchimento de zeros à esquerda para garantir que *datas* com dígitos suprimidos sejam recuperadas corretamente, e por fim utiliza-se da função `to_datetime()` da biblioteca `Pandas`, transformando todas colunas de data da base no formato `%d%m%Y`.

Como os dados do DATASUS estão codificados em formato de números inteiros, foi realizada a rotulagem dos mesmos com base no dicionário de dados oficial do sistema. Dessa forma, foi implementado um script em Python utilizando estruturas de dicionário para realizar o mapeamento de todos os códigos numéricos aos seus respectivos significados textuais, por uso do método `.map()` do `Pandas`.

Para garantir coerência, foi implementada uma função que unifica sob o rótulo "Ignorado" todos os registros nulos, em branco, com o código 9 (número usado pelo DATASUS para ignorados), espaços em branco ou valores *None*. Essa função foi aplicada a todas as colunas da base, incluindo os campos de identificação geográfica, que tiveram seus códigos municipais padronizados em 6 dígitos após conversão para inteiro e *string*. O resultado final do processamento foi exportado para um arquivo de extensão .csv, que passou a servir como a fonte de dados central para todas as visualizações do projeto.

Com as bases devidamente tratadas, a etapa seguinte consistiu no cálculo do indicador epidemiológico central do projeto: a taxa de incidência de microcefalia por 10.000 nascidos vivos por município.

Na construção do denominador demográfico, os arquivos convertidos do SINASC na primeira etapa, foram lidos em blocos de memória para viabilizar o processamento de arquivos de grande volume sem esgotar os recursos de memória disponíveis. Para cada bloco, o ano de referência foi extraído do nome do arquivo, com tratamento para nomes de arquivos em formato de 2 ou 4 dígitos. Além disso, exclui-se automaticamente arquivos nacionais com prefixo DNBR, evitando duplicidade de contagens. Os registros de cada bloco foram agrupados pela coluna de código de município de residência e ano de nascimento, contabilizando o total de nascidos vivos por município e ano. Por fim, todos os agrupamentos foram concatenados e somados para gerar a taxa de

incidência total por município para todo o período analisado.

Para a consolidação dos Casos Confirmados, o arquivo tratado do RESP foi carregado e filtrado para reter exclusivamente os registros com a classificação final confirmada de microcefalia, e agrupados por código de município de residência, resultando no total de casos confirmados por município ao longo de todo o período.

No cálculo e exportação do indicador, as duas bases foram unidas por meio de um merge com a chave do código de município de residência. A taxa de incidência foi então calculada pela equação (1). A escolha de nascidos vivos como denominador, em detrimento da população geral, está alinhada às recomendações da literatura epidemiológica aplicada a anomalias congênitas, pois representa a população efetivamente sob risco de manifestação do evento (SCHULER-FACCINI et al., 2016). O resultado foi exportado para um arquivo usado para gerar mapas.

No mapa coroplético interativo foi realizado um script que implementa a etapa de geovisualização, transformando o indicador epidemiológico calculado em uma representação cartográfica interativa. O script carrega o arquivo da taxa de incidência e garante que os códigos municipais estejam padronizados com 6 dígitos (ver Figura 1).

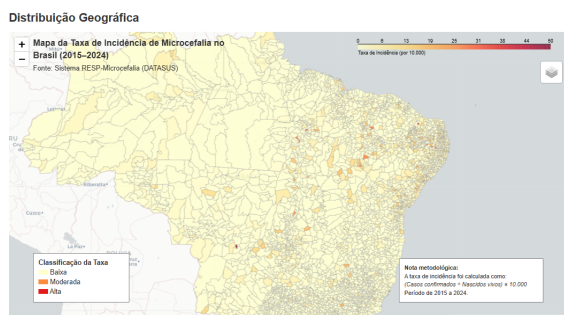


Figura 1. Mapa Coroplético Interativo

Na obtenção e compatibilização da malha geográfica, as geometrias dos municípios brasileiros foram obtidas diretamente via biblioteca Geobr, por meio da função `read_municipality()`, que retorna os polígonos geoespaciais oficiais do IBGE para o referido ano de referência. No entanto, os códigos municipais retornados pelo IBGE contêm 7 dígitos, enquanto o DATASUS adota a convenção de 6 dígitos. Para alinhar as duas bases, aplicou-se uma operação que trunca o último dígito verificador do código do IBGE e preenche com zeros à esquerda, garantindo a correspondência exata nas operações de junção.

Para reduzir o tamanho do arquivo GeoJSON gerado, foi aplicado o método `.simplify()` sobre os polígonos geométricos, que simplifica as fronteiras municipais sem descaracterizar a topologia das divisas. O mapa

coroplético foi construído com a biblioteca Folium, configurado sobre o provedor de *tiles* cartodbpositron para um visual limpo e legível. A escala de cores gradiente foi gerada pela biblioteca Branca, variando do amarelo-claro (baixa incidência) ao vermelho-intenso (alta incidência), proporcionando uma leitura visual intuitiva da severidade epidemiológica. A função *style function* mapeia cada polígono municipal à cor correspondente na escala conforme o valor da taxa de incidência. Cada município recebeu um *tooltip* interativo que exibe o nome do município, Unidade Federativa e a taxa de incidência ao passar o cursor. Esse processo é relativo à Otimização Geoespacial e Renderização do Mapa.

Assim, os elementos de interface foram incorporados ao Mapa. Na camada HTML do mapa via folium, foram adicionados um título posicionado no canto superior esquerdo identificando o mapa e sua fonte de dados; uma legenda textual no canto inferior esquerdo categorizando visualmente os níveis "Baixa", "Moderada" e "Alta" incidência com amostras de cor correspondentes; e uma nota metodológica no canto inferior direito explicando a fórmula de cálculo e o período de abrangência dos dados. O arquivo resultante é incorporado na interface principal do dashboard por meio de uma tag `<iframe>`.

A aplicação web foi concebida sob uma arquitetura leve e descentralizada de processamento no lado do cliente. Esta decisão de projeto eliminou a necessidade de servidores de aplicação robustos e sistemas gerenciadores de bancos de dados relacionais em tempo de execução, garantindo alta portabilidade e velocidade de resposta. O ecossistema é mantido pela integração de arquivos CSV, folhas de estilo customizadas e uma interface em HTML5.

O painel de controle foi desenvolvido para atuar como o núcleo de interatividade da aplicação. A seção de filtros foi disposta em uma estrutura de grade responsiva via CSS *Grid*, permitindo seleções múltiplas e simultâneas por meio de elementos nativos do HTML. Os filtros disponíveis abrangem cinco dimensões analíticas (ver Figura 2): 1) Ano: Isola anos específicos ou recortes sequenciais do período avaliado; 2) Gênero: Filtra registros classificados como "Masculino", "Feminino" e "Ignorado"; 3) Classificação: Filtra por desfechos da investigação clínica ("Confirmado", "Provável", "Em investigação", "Descartado", "Inconclusivo" e "Excluído"); 4) Região: Isola as notificações pelas cinco grandes regiões epidemiológicas do país; 5) Etiologia: Filtra por provável causa dos casos de microcefalia.



Figura 2. Filtros e cartões de métricas

A reatividade da interface é gerenciada com o uso do *JavaScript*, utilizando a biblioteca *PapaParse* para realizar o carregamento rápido e o parsing do arquivo tratado dos registros do RESP diretamente na memória do navegador. Quando o usuário aciona os botões de controle ("Aplicar Filtros" ou "Limpar Filtros"), a rotina atualiza três cartões de métricas em destaque: 1) o Total de Casos capturados no filtro corrente; 2) a quantidade de Casos Confirmados com seu percentual representativo sobre o total de registros selecionados; e 3) o número absoluto de Municípios Afetados em comparação com a totalidade das municipalidades brasileiras.

Os gráficos estatísticos foram implementados com a biblioteca *Plotly.js*, que permite a construção de visualizações dinâmicas e interativas renderizadas inteiramente no navegador. Foram desenvolvidos seis gráficos complementares: 3 gráficos temporais de linhas (série histórica anual por caso, por sexo e por etiologia) e 3 gráficos de distribuição (gráfico de setores para a exibição de casos por sexo, e gráficos de barras para a região e etiologia dos casos).

Os resultados obtidos na dimensão temporal revelam o perfil epidêmico clássico observado na literatura. Os registros apontam um pico epidemiológico expressivo entre os anos de 2015 e 2016, período correspondente à introdução e rápida dispersão do vírus Zika em território nacional, o que motivou a criação emergencial do próprio sistema RESP-Microcefalia pelo Ministério da Saúde em novembro de 2015 ( ver Figura 3). Nos anos subsequentes (a partir de 2017), os dados evidenciam uma tendência acentuada de declínio no número de novas notificações, seguida de estabilização em patamares baixos e controlados, que se estende até o final do período analisado em 2024. Esse comportamento de cauda longa reflete o sucesso relativo das medidas sazonais de controle vetorial do mosquito *Aedes aegypti* e o desenvolvimento de imunidade na população exposta.

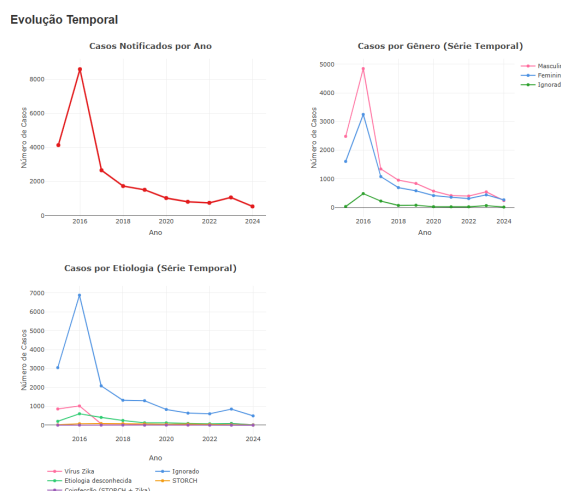


FIGURA 3. Gráficos de evolução temporal

Foi realizada análise espacial e identificação das Zonas de maior incidência. Assim, a distribuição espacial dos dados revela um padrão de desigualdade regional marcante durante todo o período. Ao utilizar a taxa de incidência por nascidos vivos, evita-se que cidades grandes distorçam os resultados apenas por terem mais habitantes, permitindo identificar onde a crise foi realmente mais grave. A análise dos resultados aponta forte concentração geoespacial das maiores taxas de incidência nos municípios da região Nordeste do Brasil. Essa distribuição concentrada corrobora amplamente a literatura da saúde coletiva, que associa a maior severidade do surto inicial no Nordeste a fatores socioambientais favoráveis à proliferação do vetor: climas predominantemente quentes e úmidos, vulnerabilidades históricas de saneamento básico e o desabastecimento crônico de água. Essas condições impulsionaram o armazenamento precário de água em recipientes desprotegidos, multiplicando os criadouros do mosquito e intensificando a taxa de transmissibilidade do vírus Zika em gestantes no início da série histórica (TEIXEIRA et al., 2016; LOWE et al., 2018). Assim, apresenta-se a geovisualização das regiões geográficas do Brasil (ver Figura 4,5,6,7 e 8)

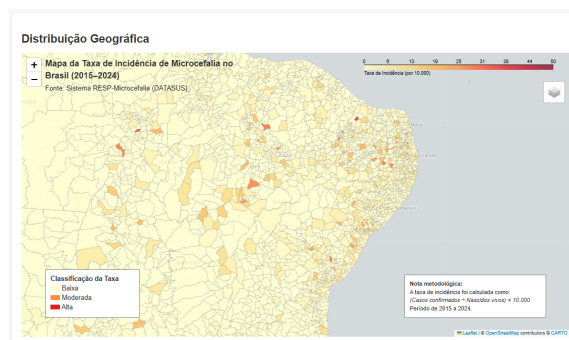


FIGURA 4. Mapa coroplético do Nordeste

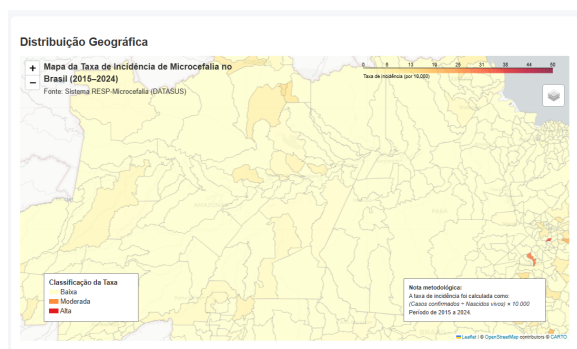


FIGURA 5. Mapa coroplético do Norte

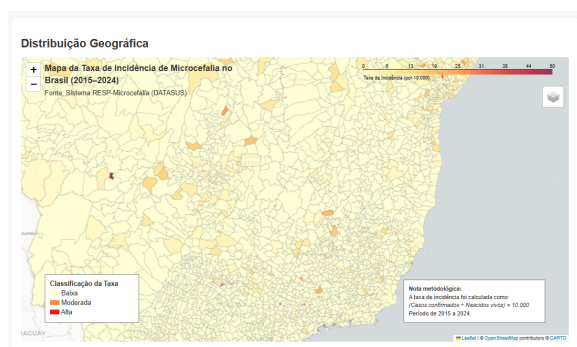


FIGURA 6. Mapa coroplético do Sudeste

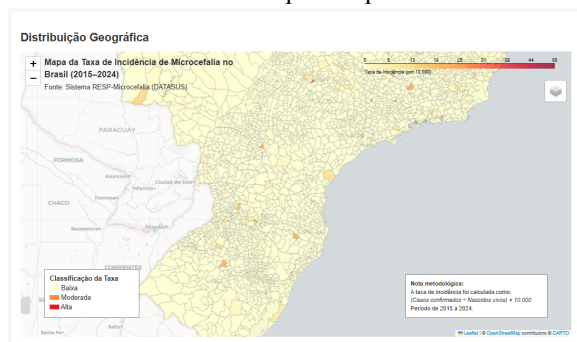


FIGURA 7. Mapa coroplético do Sul

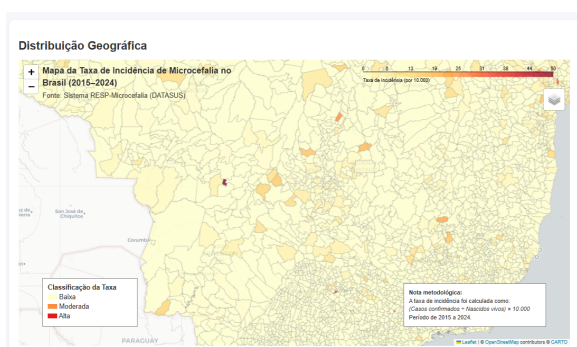


FIGURA 8. Mapa coroplético do Centro Oeste

Fez-se também a análise do perfil epidemiológico por Sexo e Classificação Final. O perfil demográfico das ocorrências, analisado por meio dos gráficos de distribuição implementados na interface, revelou um

equilíbrio entre casos de recém-nascidos do sexo masculino e feminino. Esse resultado é consistente com os fundamentos biológicos da condição: não há evidências clínicas ou epidemiológicas de predileção do vírus Zika por um sexo específico, confirmando que o agravo neurológico atinge de forma equitativa os fetos expostos à infecção materna intrauterina (SCHULER-FACCINI et al., 2016).

No que tange à classificação final, os resultados demonstraram que uma parcela expressiva dos casos suspeitos inseridos inicialmente no sistema ocorreu após investigação clínica e laboratorial aprofundada, categorizada como "Descartados" ou "Inconclusivos". Esse fenômeno está diretamente relacionado à sensibilidade inicial adotada pela vigilância epidemiológica nacional que, diante de uma emergência de saúde pública, optou por critérios de triagem amplos para garantir que nenhum caso potencial ficasse desassistido.

Por fim, a análise da variável etiologia valida a hipótese central da causa principal da epidemia: a prevalência majoritária do vírus Zika como agente primário do surto, sobressaindo-se em relação às outras infecções congênicas como do grupo STORCH (Sífilis, Toxoplasmose, Rubéola, Citomegalovírus e Herpes Simplex), que mantiveram suas taxas endêmicas basais históricas ao longo do período (ver Figura 9).

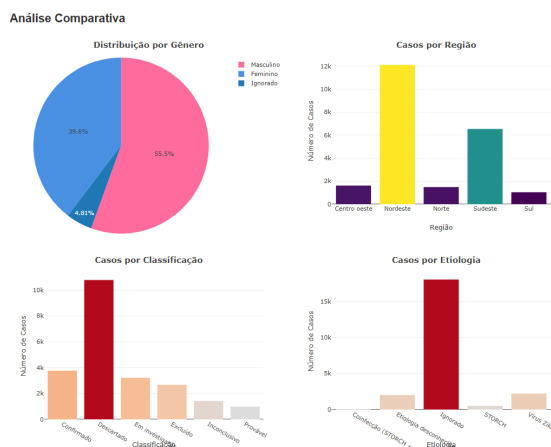
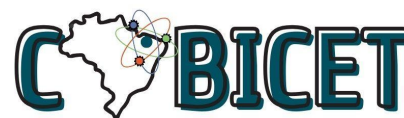


FIGURA 9. Gráficos de análise comparativa

## CONCLUSÃO

Desenvolveu-se uma aplicação web para geovisualização dos casos de microcefalia registrados no Brasil entre 2015 e 2024. A partir de dados públicos do RESP-Microcefalia e do SINASC (DATASUS), foi construído um fluxo completo em Python (abrangendo desde a conversão de arquivos no formato .dbc até a geração de um mapa coroplético interativo e painéis estatísticos dinâmicos), resultando em uma ferramenta capaz de



transformar registros epidemiológicos complexos em visualizações acessíveis e intuitivas.

Os resultados obtidos confirmaram os padrões descritos na literatura científica: o pico epidemiológico concentrado entre 2015 e 2016, associado à introdução do vírus Zika no território nacional; a forte concentração espacial das maiores taxas de incidência nos municípios do Nordeste, região marcada por condições socioambientais historicamente favoráveis à proliferação do vetor *Aedes aegypti*; e o equilíbrio proporcional entre os sexos dos recém-nascidos acometidos, coerente com a ausência de predileção biológica do vírus por sexo específico. A análise da variável etiologia reforçou a centralidade do vírus Zika como agente primário do surto, distinguindo-o das demais infecções congênitas do grupo STORCH, que permaneceram em seus níveis endêmicos basais.

Do ponto de vista técnico, a arquitetura adotada mostrou-se eficaz para a proposta do projeto, dispensando infraestrutura de servidores e bancos de dados relacionais, o que garante portabilidade e facilidade de distribuição da solução. A integração entre as bibliotecas Python (Pandas, Geobr, Folium e Plotly) e as tecnologias web (HTML5, CSS3 e JavaScript) demonstrou ser uma abordagem viável e reproduzível para projetos de saúde pública com dados abertos.

Como limitação, destaca-se que os dados do RESP-Microcefalia se encerram em 2024, não cobrindo integralmente o período de 2025 originalmente proposto, além de possíveis subnotificações inerentes aos sistemas de vigilância epidemiológica. Adicionalmente, a aplicação não foi disponibilizada publicamente até a conclusão deste trabalho, sendo sua publicação *online* e o compartilhamento do código-fonte em repositório aberto para etapas em um trabalho futuro. Nesse sentido, pretende-se também expandir a solução com a incorporação de dados do SINAN, a implementação de modelos preditivos de reemergência do vírus Zika e a adaptação da plataforma para análise de outras arboviroses de interesse epidemiológico, como dengue e chikungunya.

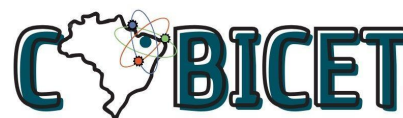
Por fim, aplicação desenvolvida representa uma contribuição concreta à intersecção entre ciência de dados, programação e saúde pública, demonstrando o potencial da geovisualização computacional como instrumento de apoio à pesquisa epidemiológica e à formulação de políticas públicas baseadas em evidências.

#### AGRADECIMENTOS

À Unimontes.

#### REFERÊNCIAS

- BR ADZAWLA, K. V. Analyzing geospatial data using folium and plotly. Disponível em: <<https://medium.com/@vevadzawla/analyzing-geospatial-data-using-folium-and-plotly-6c328e43fe8a>>. Acesso em: 14 jun. 2025.
- BRUNONI, D. et al. Microcefalia e outras manifestações relacionadas ao vírus Zika: impacto nas crianças, nas famílias e nas equipes de saúde. *Ciência & Saúde Coletiva*, v. 21, n. 10, p. 3297–3302, out. 2016.
- CADENAS, C. Geovisualization: Integration and visualization of multiple datasets using mapbox. 2014.
- DANICAT. GitHub - danicat/pysus: A library to open DATASUS files (\*.dbc) in Python. Disponível em: <<https://github.com/danicat/pysus>>. Acesso em: 14 jun. 2025.
- DATASUS. Disponível em: <<https://datasus.saude.gov.br/sobre-o-datasus/>>. Acesso em: 13 jun. 2025.
- ERICKSON, J. MySQL: Entendendo o que é e como é usado. Disponível em: <<https://www.oracle.com/br/mysql/what-is-mysql/>>. Acesso em: 12 jun. 2025.
- GEEKSFORGEES. Visualizing Geospatial Data using Folium in Python. Disponível em: <<https://www.geeksforgeeks.org/python/visualizing-geospatial-data-using-folium-in-python/>>. Acesso em: 14 jun. 2025.
- GOVERNO. Governo declara emergência por casos de microcefalia em Pernambuco. Disponível em: <<https://agenciabrasil.ebc.com.br/geral/noticia/2015-11/governo-declara-emergencia-em-saude-publica-por-casos-de-microcefalia-em-pe>>. Acesso em: 13 jun. 2025.
- HUMAN RIGHTS WATCH. Neglected and Unprotected. Disponível em: <<https://www.hrw.org/report/2017/07/13/neglected-and-unprotected/impact-zika-outbreak-women-and-girls-northeastern>>. Acesso em: 13 jun. 2025.
- IPEAGIT. GitHub - ipeaGIT/geobr: Easy access to official spatial data sets of Brazil in R and Python. Disponível em: <<https://github.com/ipeaGIT/geobr>>.
- LOWE, R. et al. The Zika Virus Epidemic in Brazil: From Discovery to Future Implications. *International Journal of Environmental Research and Public Health*, v. 15, n. 1, p. 96, 9 jan. 2018.



LUNA, J. C. GeoPandas Tutorial: An Introduction to Geospatial Analysis. Disponível em: <<https://www.datacamp.com/tutorial/geopandas-tutorial-geospatial-analysis>>. Acesso em: 14 jun. 2025.

MELO, D. O que é CSS [Cascading Style Sheets]? Disponível em: <<https://tecnoblog.net/responde/o-que-e-css-cascading-style-sheets/>>. Acesso em: 12 jun. 2025a.

MELO, D. O que é HTML? [Guia para iniciantes]. Disponível em: <<https://tecnoblog.net/responde/o-que-e-html-guia-para-iniciantes/>>. Acesso em: 12 jun. 2025b.

O que é JavaScript? - Aprendendo desenvolvimento web | MDN. Disponível em: <[https://developer.mozilla.org/pt-BR/docs/Learn\\_web\\_development/Core/Scripting/What\\_is\\_JavaScript](https://developer.mozilla.org/pt-BR/docs/Learn_web_development/Core/Scripting/What_is_JavaScript)>. Acesso em: 12 jun. 2025.

OLEMB. GitHub - olemb/dbfread: Read DBF Files with Python. Disponível em: <<https://github.com/olemb/dbfread>>. Acesso em: 14 jun. 2025.

PIMENTEL, E. C.; BRAGA, F. Desigualdades em Saúde no Brasil: Análise Comparativa através da Geovisualização de Indicadores. In: IV Congresso Internacional de Geografia da Saúde: Geografia, produção do conhecimento e práticas em saúde, Presidente Prudente.

RESP - Microcefalia. Disponível em: <<https://www.gov.br/saude/pt-br/composicao/svsa/sistemas-de-informacao/resp-microcefalia>>. Acesso em: 13 jun. 2025.

ROBINSON, A. Geovisualization and epidemiology : A general design framework. 2005.

ROBITEAILLE, F. Folium: Python Data, Leaflet.js Maps. Disponível em: <https://github.com/python-visualization/folium>. Acesso em: 14 jun. 2025.

SCHULER-FACCINI, L. et al. Possible association between Zika virus infection and microcephaly - Brazil, 2015. MMWR. Morbidity and mortality weekly report, v. 65, n. 3, p. 59–62, 2016.

TEIXEIRA, M. G. et al. The Epidemic of Zika Virus–Related Microcephaly in Brazil: Detection, Control, Etiology, and Future Scenarios. American Journal of Public Health, v. 106, n. 4, p. 601–605, abr. 2016.

WORLD. WHO Director-General summarizes the outcome of the Emergency Committee regarding clusters of microcephaly and Guillain-Barré syndrome. Disponível em: <<https://www.who.int/en/news-room/detail/01-02-2016-who-director>>

general-summarizes-the-outcome-of-the-emergency-committee-regarding-clusters-of-microcephaly-and-guillain-barré-syndrome>. Acesso em: 14 jun. 2025.