

# Projeto Extensionista – Dengue no Brasil (2023–2025)

---

Aluno: Celso Muniz de Carvalho Neto  
Matrícula: 202203866681  
Professora: Simone Gama  
Curso: Ciência da Computação – UNESA  
Disciplina: Tópicos de Big Data – 2025.2

## 1. DIAGNÓSTICO E TEORIZAÇÃO

### Problemática e/ou problemas identificados

O Brasil enfrenta um aumento significativo nos casos de dengue, especialmente em 2024, com registros superiores a 6 milhões de notificações prováveis. A variação sazonal e o impacto de fatores climáticos contribuem para a dificuldade no controle e prevenção da doença. O problema central identificado é a alta incidência de casos, que gera sobrecarga no sistema de saúde e amplia os riscos à população.

### Justificativa

A escolha do tema se justifica pela relevância social e pela necessidade de utilizar dados reais para compreender e prever padrões de surtos. O estudo da dengue, por meio de análise de Big Data, possibilita desenvolver soluções baseadas em evidências que podem auxiliar autoridades públicas na tomada de decisão e conscientizar a sociedade.

### Objetivos/resultados/efeitos a serem alcançados

- Realizar análise exploratória dos casos de dengue de 2023 a 2025.
- Identificar os meses com maior concentração de casos.
- Construir gráficos que mostrem tendências anuais e mensais.
- Compreender a distribuição percentual dos casos e sua evolução.
- Contribuir para a conscientização sobre a gravidade da doença.

### Referencial teórico

De acordo com a Organização Mundial da Saúde (OMS) e o Ministério da Saúde, a dengue é uma das arboviroses mais graves da atualidade. Estudos de epidemiologia reforçam que análises de séries temporais permitem prever surtos, enquanto ferramentas de Big Data ampliam a capacidade de monitoramento. Referências em Ciência de Dados, como Han, Kamber & Pei (2012), destacam a importância da análise preditiva em saúde pública.

## 2. PLANEJAMENTO E DESENVOLVIMENTO DO PROJETO

### Plano de trabalho

- Coleta dos dados oficiais em formato CSV (SINAN/SVS-MS).
- Limpeza e unificação dos datasets.
- Organização das variáveis de interesse (ano, mês, número de casos).
- Criação de pipelines de análise no Google Colab (Python/Pandas).
- Construção de gráficos explicativos (evolução anual, distribuição por ano, evolução mensal e frequência mensal).
- Interpretação dos resultados e elaboração de relatório.

### Grupo de trabalho

Projeto realizado individualmente pelo aluno Celso Muniz de Carvalho Neto, que assumiu o papel de Full Stack Analytics, responsável por coleta, tratamento, análise e visualização dos dados.

### Detalhamento técnico do projeto

#### Explicação do tema escolhido

O tema dengue foi escolhido devido à sua relevância social, ao impacto nos serviços de saúde e à atualidade dos surtos recentes no Brasil. Analisar dados reais permite criar relatórios visuais que evidenciam a gravidade da doença e podem apoiar políticas públicas.

#### Explicação do código desenvolvido

O código foi implementado no Google Colab utilizando Python. Pandas para carregamento e tratamento dos dados, Matplotlib/Seaborn para criação de gráficos, pipeline de agrupamento por ano e por mês, cálculo de frequência percentual dos casos e ajuste de regressão linear para análise de tendência.

#### Explicação dos resultados alcançados

- Identificação de que 2024 foi o ano mais crítico, concentrando 67,9% dos casos registrados.
- Meses de março e abril apresentaram picos de infecção, com tendência de queda a partir de maio.
- Distribuição mensal mostrou que mais de 60% dos casos de 2025 ocorreram até abril.
- Gráficos explicativos possibilitaram visualizar padrões anuais e sazonais.

## 3. ENCERRAMENTO DO PROJETO

### Relatório coletivo

O projeto permitiu aplicar técnicas de Big Data em um problema real de saúde pública. Os resultados mostraram a gravidade da dengue no período analisado, especialmente em 2024, e apontaram a importância de políticas de prevenção voltadas para os meses mais críticos.

### **Relato de Experiência Individual**

A execução do trabalho no Google Colab proporcionou experiência prática na manipulação de dados reais e construção de dashboards em Python, consolidando o aprendizado em análise de dados e epidemiologia.

## **4. CONSIDERAÇÕES FINAIS**

O estudo demonstrou que a dengue continua sendo um dos maiores desafios de saúde pública no Brasil. A utilização de dados oficiais e ferramentas de Big Data mostrou-se eficiente para identificar padrões temporais e apoiar estratégias preventivas. Conclui-se que a integração entre ciência de dados e saúde pode trazer benefícios diretos para a sociedade.