

# Estudo Comparativo entre dengue, chikungunya e zikavírus na cidade de Recife (PE), no ano 2021, usando a linguagem Python

Angélica Félix de Castro  
Departamento de Computação  
Universidade Federal Rural do Semi-Árido  
Mossoró, Brasil  
[angelica@ufersa.edu.br](mailto:angelica@ufersa.edu.br)

Amanda Gondim de Oliveira  
Departamento de Computação  
Universidade Federal Rural do Semi-Árido  
Mossoró, Brasil  
[amandagondim@ufersa.edu.br](mailto:amandagondim@ufersa.edu.br)

George Felipe Fernandes Vieira  
Departamento de Computação – Bolsista da  
UFERSA  
Universidade Federal Rural do Semi-Árido  
Mossoró, Brasil  
[georgeffv1@gmail.com](mailto:georgeffv1@gmail.com)

**Resumo** - Arboviroses são as doenças causadas pelos chamados arbovírus, que incluem o vírus da dengue, zika vírus, febre chikungunya e febre amarela. São um problema de saúde pública no Mundo principalmente pelo potencial de dispersão, pela capacidade de adaptação a novos ambientes e hospedeiros, pela possibilidade de causar epidemias extensas e pela susceptibilidade universal. O presente trabalho consiste em um estudo dos casos de arboviroses em Recife, capital do estado de Pernambuco, Nordeste do Brasil, no ano de 2021. Primeiramente, foram extraídos dados de fontes oficiais da Prefeitura Municipal da cidade do Recife sobre casos de dengue, chikungunya e zika vírus no ano de 2021. Após essa extração, comandos Python foram aplicados para entender como ocorreu o surto dessas doenças específicas no Recife e para gerar mapas de calor dos bairros mais afetados pelas arboviroses. Foi possível concluir que as arboviroses não são motivos preocupantes de morte e não causam alta hospitalização dos pacientes. Quanto aos mapas de calor, os bairros de Varzea, Ibura, Iputinga, Imbiribeira, Vasco da Gama e Dois Unidos foram os mais afetados pela presença das três doenças. Em termos comparativos, dengue e chikungunya apresentam sintomas semelhantes (febre e cefaleia).

**Palavras-chave:** Arboviroses, Python, Mapas de Calor

## I. INTRODUÇÃO

É notório que o Brasil vem sendo afligido por doenças com graves repercussões para a saúde de sua população conforme indicadores operacionais e epidemiológicos, entre elas as Arboviroses (ou Arbovírus), conjunto de patologias formado principalmente pela Dengue, Chikungunya e Zikavírus [1]. Esses três agravos foram sinalizados pelo Ministério da Saúde - entidade superior de orientação e formulação de políticas a nível federal que devem ser seguidas pelos profissionais de saúde das esferas Estaduais e Municipais - como uma doença negligenciada, que merece receber atenção especial dos profissionais de saúde que atuam na atenção e na vigilância e dos gestores das esferas federal, estadual e municipal.

A área da saúde tem sido uma das áreas mais beneficiadas pela tecnologia, por ser considerada detentora de problemas clássicos, possuidores de todas as peculiaridades necessárias, para serem instrumentalizados por tais sistemas [2]. Neste contexto, a avaliação dos classificadores e sua implementação em um sistema que auxilie no diagnóstico de doenças como a dengue, chikungunya e zika, que hoje afetam uma grande parte da população mundial, é totalmente possível [3].

A área de ciência e análise de dados também tem sido amplamente utilizada no estudo e proliferação dessas arboviroses. Em se tratando de computação aplicada à saúde, desde os anos 2000 até hoje, são desenvolvidos estudos utilizando redes neurais, algoritmos de aprendizado de máquina, até mesmo modelos estatísticos, que tem como intuito contribuir ao combate do dengue e reverter esse quadro alarmante que a doença paira à nação [4].

Ciência de Dados pode ser definida como os processos, modelos e tecnologias que estudam os dados durante todo o seu ciclo de vida: produção, armazenamento, transformação, análise e descarte [5]. Algumas operações podem acontecer na etapa de análise de dados: é possível recorrer a gráficos, tabelas e outros recursos visuais que permitam uma visualização mais clara dos dados coletados. Uma tabela dinâmica, por exemplo, pode ajudar a classificar e filtrar os dados a partir de diferentes variáveis. Durante esta etapa de análise de dados, a utilização da linguagem Python em uma base de dados pode ser extremamente benéfica, por possuir bibliotecas com modelos matemáticos, fórmulas, testes estatísticos e aprendizagem de máquina [5].

O presente trabalho possui como objetivo principal analisar os dados de arboviroses na cidade de Recife, capital de Pernambuco, Nordeste do Brasil, no ano de 2021, através da obtenção de bases de dados governamentais do estado. A ideia é aplicar conceitos de análise e ciência de dados, usando a linguagem Python, para compreender como essas doenças se comportaram na capital recifense.

## II. REFERENCIAL TEÓRICO

### A. Arboviroses

Em grandes cidades, a proliferação de doenças infecciosas ocorre com maior intensidade e o controle epidêmico é mais desafiador devido ao aumento da população e da grande interação social. Não só o Brasil, mas também vários países do mundo, principalmente com clima tropical, apresentam epidemias [6] de doenças transmitidas pelo mosquito *Aedes aegypti*, popularmente conhecido como mosquito da Dengue. Além de Dengue, o mosquito também transmite doenças como Zika e Chikungunya, sendo todas elas integrantes de um grupo maior de doenças denominadas de arboviroses.

A dengue pode se manifestar de diferentes formas. Ela pode ser assintomática ou pode evoluir até quadros graves, como a dengue hemorrágica. A chikungunya e a zika são doenças causadas por vírus, chamados CHIKV e KIKV respectivamente, com um conjunto de sintomas que é bastante similar ao da dengue. A transmissão do zika vírus também pode acontecer por meio do contato sexual com uma pessoa que tenha sido infectada pela doença [7] e [8].

### B. Análise de Dados

A Análise Exploratória de Dados (AED) teve seu início com J.W. Tukey, em 1977, e visa aumentar o conhecimento do pesquisador sobre uma população a partir de uma amostra. Dessa forma, é possível descrever a AED como um conjunto de métodos adequados para a coleta, a exploração e descrição e interpretação de conjuntos de dados numéricos. Estes métodos permitem a exploração dos dados com intuito de identificar padrões de interesse e a representação dos dados caracterizados por estes padrões. Embora as técnicas da AED sejam simples, geralmente são métodos robustos (válidos para uma grande gama de situações e modelos) e resistentes (insensíveis aos erros grosseiros ou dados estranhos) [9].

#### B.1) Uso de Python em Análise Exploratória de Dados

Python é uma linguagem de programação muito utilizada em atividades de análise de dados e ao contrário de linguagens como R, Python é uma linguagem de programação de propósito geral, ou seja, não é exclusiva para atividades de análise de dados [9]. As características mais relevantes do Python são a sua capacidade de integração com outras linguagens e seu sistema de bibliotecas maduro. Algumas bibliotecas do Python são fortemente analíticas e oferecem ferramentas de ciência de dados composta de funções altamente otimizadas para trabalhar, configuração ideal de memória, pronta para realizar operações de script com desempenho ideal [10].

### III. METODOLOGIA APLICADA

O presente trabalho se caracteriza como um estudo exploratório que avaliou os dados de notificação das arboviroses, no ano de 2021 na área metropolitana do Recife no estado de Pernambuco. A pesquisa foi realizada com base nas informações disponíveis no portal oficial<sup>1</sup> da Prefeitura do Recife que registrou 9217 casos de dengue, 15035 casos de chikungunya e 533 de zikavírus, no ano de 2021.

Para esse trabalho, os dados foram estudados e analisados por meio de técnicas e ferramentas da área de análise de dados com o objetivo de fornecer uma melhor visualização e interpretação dos dados brutos disponíveis pela Prefeitura da cidade.

Do ponto de vista prático, a implementação desse trabalho foi realizada com o uso da linguagem de programação Python devido à sua capacidade de integração com outras linguagens e seu sistema de bibliotecas maduro, as quais oferecem ferramentas de ciência de dados compostas de funções altamente otimizadas. As bibliotecas do Python utilizadas no presente trabalho foram: Pandas, Matplotlib, Numpy e PyLab.

Após a extração e um pré-processamento dos dados, iniciou-se a fase de análise e mineração nos mesmos. Os resultados dessas análises serão apresentados a seguir.

#### Sintomas mais Comuns

De acordo com com [11], a dengue em Pernambuco não tem um índice alto de fatalidade, pois a maioria das pessoas conseguem a cura e o índice de internação é bem baixo (cerca de 4%). No estado de Pernambuco inteiro, no ano de 2021, houve 13.296 casos confirmados de dengue, 18.100 casos confirmados de chikungunya e outros 11 casos de zika. Desses números, foram identificadas 4 mortes [12]. É um índice de mortalidade baixo.

Ao analisar os dados da dengue e chikungunya da base de dados oficial de Pernambuco, foram identificados alguns sintomas presentes em ambos os casos. Na Figura 1, é possível de ver esses sintomas junto à presença da dengue.

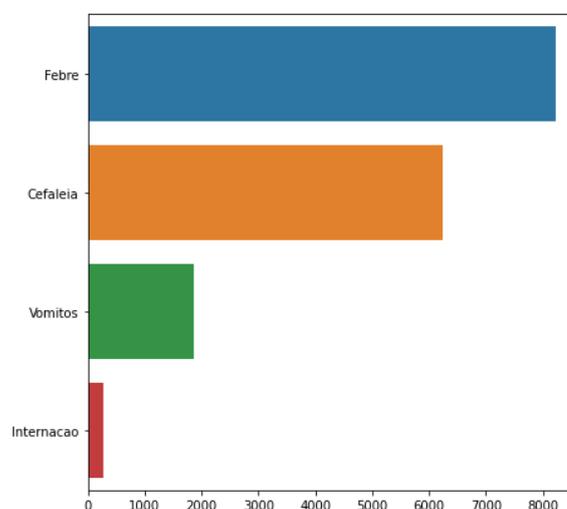


Fig. 1. Sintomas presentes na dengue e número de internações

Na Figura 2, notam-se os mesmos sintomas apresentados na Figura 1, mas dessa vez, em relação a chikungunya.

<sup>1</sup> <http://dados.recife.pe.gov.br/tr/dataset/casos-de-dengue-zika-e-chikungunya>

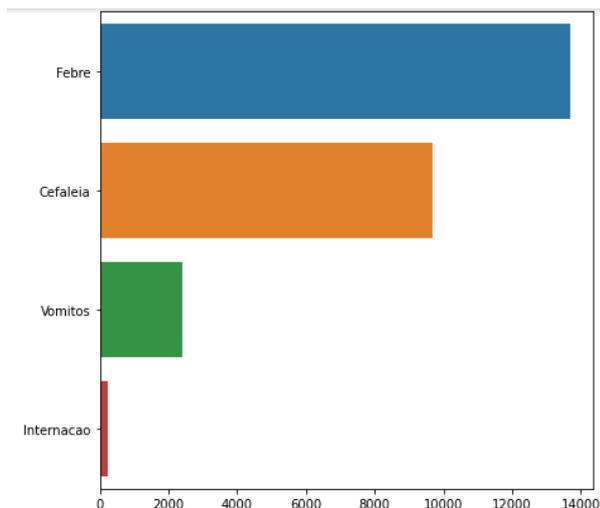


Fig. 2. Sintomas presentes nos casos de chikungunya e número de internações

Percebe-se que tanto na dengue quanto no chikungunya, o sintoma que mais aparece é febre, seguido de cefaleia. Também foi identificada uma quantidade razoável de pessoas que apresentaram vômitos ao longo da doença. Mesmo assim, os números de internações são baixos e a quantidade de mortes causadas por essas doença, é pequena. Na base de dados da zika vírus, não existem sintomas registrados.

### Regiões mais Afetadas na Cidade

Com o uso das bibliotecas Python, foi possível identificar os bairros de Recife que mais tiveram casos detectados de dengue, chikungunya e zikavírus no ano de 2021; bem como, gerar mapas de calor dessas detecções.

Na Tabela 1, aparecem os bairros, o tipo de arbovirose e a quantidade detectada. Nessa tabela, foram identificados apenas os 10 bairros que mais ocorreram dengue, os 10 bairros que mais ocorreram chikungunya e os 10 bairros que mais ocorreram zikavírus. Por isso, a existência de algumas células em branco.

TABELA 1. TABELA COM A RELAÇÃO DAS ARBOVIROSES NO ANO DE 2021 X QUANTIDADE DE CASOS DETECTADOS NOS BAIRROS

Bairro	Casos detectados no ano de 2021		
	Dengue (total de 9217 casos em 2021)	Chikungunya (total de 15035 casos em 2021)	Zikavírus (total de 533 casos em 2021)
Varzea	523	543	32
Ibura	520	855	17
Iputinga	415	456	20
Imbiribeira	402		
Vasco da Gama	304	496	22
Dois Unidos	271	707	18
Santo Amaro	271		
Boa Viagem	267		

Bairro	Casos detectados no ano de 2021		
	Dengue (total de 9217 casos em 2021)	Chikungunya (total de 15035 casos em 2021)	Zikavírus (total de 533 casos em 2021)
Nova Descoberta	247		21
Água Fria	245	757	
Campo Grande		584	
Cohab		493	49
Afogados		436	20
Linha do Tiro		425	
Cordeiro			28
Casa Amarela			16

Além disso, foi possível confeccionar os mapas de calor dos dados da Tabela 1. Nas Figuras 3, 4 e 5 são apresentados os mapas de calor com os bairros mais afetados pela dengue, chikungunya e zikavírus, respectivamente.

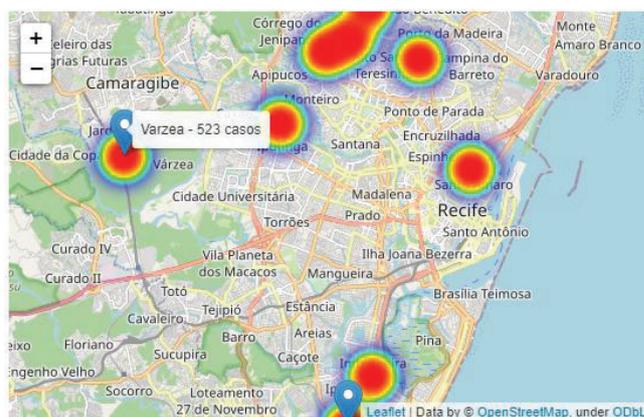


Fig. 3. Mapa de calor com os bairros que mais tiveram casos de dengue no ano de 2021, em Recife. O bairro Varzea foi o mais afetado, com 523 casos.



Fig. 4. Mapa de calor com os bairros que mais tiveram casos de chikungunya no ano de 2021, em Recife. O bairro Ibura foi o mais afetado, com 855 casos.

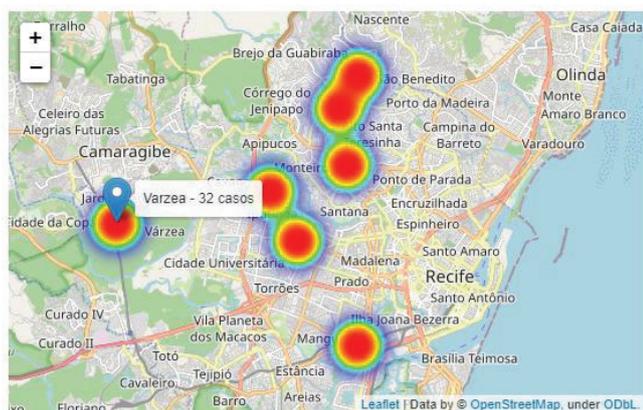


Fig. 5. Mapa de calor com os bairros que mais tiveram casos de zikavírus no ano de 2021, em Recife. O bairro Várzea foi o mais afetado, com 32 casos.

#### IV. RESULTADOS E CONCLUSÕES

A utilização de ciência e análise de dados na área da saúde tem sofrido uma crescente nos últimos anos. Conseguir extrair informação agregada a um dado tem sido de grande valia na tomada de apoio à decisão.

As arboviroses são enfermidades tropicais endêmicas, que merecem receber atenção especial dos profissionais de saúde que atuam na atenção básica e na vigilância em saúde e dos gestores das esferas federal, estadual e municipal. Nesse sentido, o presente trabalho avaliou computacionalmente os dados de arboviroses de Recife, PE, no ano de 2021, a fim analisar esse quadro sanitário.

Após a aplicação de comandos e bibliotecas Python nas bases de dados oficiais da Prefeitura de Pernambuco, algumas conclusões foram obtidas:

- As arboviroses tem um índice de mortalidade baixo;
- Os sintomas mais comuns na dengue e Chikungunya são febre e cefaleia;
- O índice de internação de pessoas detectadas com arboviroses é baixo;
- Mapas de calor auxiliaram na visualização das arboviroses nos bairros de Recife;
- Os bairros Várzea, Ibura, Iputinga, Vasco da Gama e Dois Unidos merecem uma atenção especial pois houve confirmação das três arboviroses no ano de 2021

Como trabalho futuro, pretende-se aplicar essas técnicas e metodologias em outras cidades brasileiras, a fim de analisar a situação da presença de arboviroses nessas cidades.

#### AGRADECIMENTOS

Os autores agradecem à Universidade Federal Rural do Semi-Árido (UFERSA) pelo apoio financeiro e pela infraestrutura disponibilizada.

#### REFERÊNCIAS

- [1] P.M.S. Silva e S.A.Siebr, “A construção do repositório de dados da UFPPB: a experiência com o dataset de Arboviroses”, Métodos e Pesquisa em Administração, Seção: Pesquisas científicas. v. 4, n. 1, p. 48-54, 2019.
- [2] N.S. Nilson, Principles of Artificial Intelligence. Springer Verlag, Berlin, 1982.
- [3] W.C. Lima, “Estudo de classificadores e construção de dataset para diagnóstico de dengue, chikungunya e zika”, Trabalho de conclusão de curso, Departamento Acadêmico de Informática, Universidade Tecnológica Federal do Paraná, 2017.
- [4] P. I. Andrade e L.V. T. Brilhante, “Predizendo epidemias de Dengue, no Distrito Federal, utilizando algoritmos de Regressão”, Monografia de Graduação, Universidade de Brasília, 2018.
- [5] F. Amaral. 2016. Introdução à Ciência de Dados: Mineração de Dados e Big Data (1st. ed.). Vol. 1. Alta Books.
- [6] G. Cunha, “Surto, Epidemia, Endemia e Pandemia: Você sabe a diferença?”, 2017. Disponível em: <https://blog.enem.com.br/surto-epidemia-endemia-e-pandemia-voce-sabe-diferenca/>. Acesso em: 07 set 2022.
- [7] Rede D’or. Chikungunya. Disponível em: <https://www.rededorsaoluz.com.br/doencas/chikungunya>. Acesso em: 07 set 2022.
- [8] Rede D’or. Zika. Disponível em: <https://www.rededorsaoluz.com.br/doencas/zika>. Acesso em: 07 set 2022.
- [9] Tukey, John W. 1977. Exploratory Data Analysis. 1<sup>st</sup> Ed. Pearson. 688p.R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
- [10] G. R. Lopes, A. W. S. Almeida, A. C. B. Delbem and, C. F. M. Toledo. 2019. “Introdução à Análise Exploratória de Dados com Python”. In: Escola Regional de Computação Aplicada à Saúde (ERCAS 2019).
- [11] A. F. Castro e A. G. Oliveira, “Estudo sobre os casos de dengue utilizando técnicas de Análise de Dados: Um estudo de caso no estado de Pernambuco, Brasil”. In: 16th Iberian Conference on Information Systems and Technologies (CISTI), 2021, pp. 1-6, doi: 10.23919/CISTI52073.2021.9476432.
- [12] Portal Eletrônico G1 - PE. Disponível em: <https://g1.globo.com/pe/pernambuco/noticia/2021/11/16/recife-confirma-mais-duas-mortes-por-dengue-e-sobe-para-quatro-total-de-obitos-por-arboviroses-em-pe-em-2021.ghtml>. Acesso em: 08 set. 2022