

ARBOVIS: Uma aplicação de visualização de dados sobre Arboviroses

ARBOVIS: A data visualization application on Arboviruses

Talitha Alessandra Medeiros Ribeiro¹, Larissa de Sousa Moreira Gusmão¹, Damires Yluska de Souza Fernandes¹

(1) Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB), Av. Primeiro de Maio, João Pessoa - PB, medeirostalitha@gmail.com, larissamgusmao@gmail.com, damires@ifpb.edu.br.

Resumo:

O volume de informações em saúde, em especial no cenário das arboviroses, cresceu juntamente com os desafios para trabalhar com os dados de forma que análises úteis possam ser feitas e usadas para apoiar a tomada de decisões. Este artigo apresenta a Arbovis, uma aplicação para análise visual de dados sobre arboviroses no estado da Paraíba. A arbovis disponibiliza visões interativas e customizadas sobre os referidos dados, baseadas em filtros que podem ser aplicados para gerar gráficos. Os dados apresentados foram obtidos através da Secretaria de Saúde do Estado da Paraíba e do IBGE. Como resultados iniciais, foram desenvolvidos gráficos de barras, barras horizontais e de linhas. Através da visão interativa, é possível identificar quais municípios possuem maior incidência de casos das doenças, por meio do cruzamento entre quantidade de habitantes e de casos, e também em quais houve maior quantidade de casos registrados. Dessa maneira, a Arbovis apresenta impacto positivo na gestão e na prevenção de doenças e epidemias, contribuindo para estudos e análises de pesquisadores, profissionais da saúde, gestores e da população em geral.

Palavras-chave: arboviroses; análise de dados; visualização de dados; aplicação web.

Abstract:

The volume of health information, particularly in the arboviruses scenario, has grown along with the challenges of working with data in such a way that useful analyzes can be produced and used to support decision making. In this light, this paper presents Arbovis, an application for visual analysis of data on arboviruses. Arbovis provides interactive and customized views of the data, based on filters that can be applied to generate graphs. The data presented were obtained from the Department of Health of the State of Paraíba and also from the IBGE. As initial results, bar graphs, horizontal bars and lines were developed. Through an interactive view, it is possible to identify which cities have a higher incidence of arboviruses cases. This is obtained by crossing and integrating data from some data sources with registered cases. Thereby, Arbovis provides a positive impact on the management and prevention of diseases and epidemics, thus contributing to studies and analyzes of researchers, health professionals, managers and the population in general.

Keywords: arboviruses; data analytics; data visualization; web application.

1 Introdução

A demanda por informações em saúde cresce a cada dia juntamente com os desafios para que sua utilização possa trazer resultados úteis à população. Em especial, o cenário associado às doenças transmitidas pelo mosquito *Aedes aegypti* (arboviroses) tem causado necessidades específicas de estudos e definições de políticas públicas. Diante disso, diversos órgãos institucionais, além de pesquisadores e profissionais de saúde, trabalham gerando dados sobre arboviroses.

À medida que os dados se tornam mais volumosos, com estruturas heterogêneas e com alta velocidade de

produção, como é o caso de dados em massa, ou *big data* (GANDOMI; HAIDER, 2015), a complexidade nas tarefas de análise de dados aumenta. Isso se deve a fatores como o próprio volume de dados, a quantidade de atributos a serem gerenciados e interpretados (variedade de dados) e também a aspectos da qualidade dos dados (THABET; SOOMRO, 2015). Nesse sentido, estratégias de ETL (Extração, Transformação e Carga) e de visualização da informação podem ajudar (BANSAL; KAGEMANN, 2015).

Estratégias de ETL referem-se a processos que permitem extrair dados de diferentes fontes, transformá-los para atender

as necessidades de padronização e uso, e carregá-los em repositórios de destino. Com os dados extraídos, o próximo passo é visualizá-los. A visualização da informação pode ser definida como o uso de representações visuais interativas de dados abstratos para amplificar a cognição (SHNEIDERMAN; PLAISANT; HESSE, 2013). Está associada à transformação de dados abstratos em, por exemplo, gráficos e imagens, denominados visualizações (VIEIRA; CORREA, 2011).

Nesse contexto, surge o termo “análise visual de dados” que busca integrar técnicas de visualização da informação com as de análise de dados. Mais precisamente, o termo *Visual Analytics* ou Análise Visual foi introduzido por Thomas e Cook (2005) quando se percebeu que a representação visual utilizando apenas técnicas de visualização de informação não era suficiente para viabilizar a análise de dados em massa (CARMO, 2015). Técnicas de interação e manipulação direta foram associadas à atividade de análise de modo que o usuário (que tem o papel de analista de dados) pudesse ser o centro e o coordenador da atividade. Estratégias que incluem a aplicação de filtros nos dados, operações de seleção de objetos visuais, navegação, exploração por zoom, construção de representações visuais podem ampliar as possibilidades das tarefas analíticas (FEKETE, 2013). Tais estratégias podem facilitar a personalização das tarefas de análises de dados e torná-las mais intuitivas.

Nessa perspectiva, este artigo propõe a aplicação Arbovis¹, que visa apresentar dados sobre arboviroses, no estado da Paraíba, por meio de visualizações interativas. A versão apresentada neste trabalho é uma extensão da aplicação de Santiago et al. (2018), que não incluía opções de interatividade e análises visuais construídas dinamicamente. Como resultado, o ambiente de análise de dados se torna mais rico, interativo e as informações geradas são mais compreensíveis.

¹ <http://arboviz.siderg.com.br/>

2 Objetivos

O objetivo da Arbovis é mostrar visões significativas sobre arboviroses, distribuindo filtros de apresentação de forma que o usuário possa montar visualizações específicas dos dados de acordo com suas necessidades. Desse modo, as visualizações pretendem contribuir para análise visual de dados sobre arboviroses por meio de recursos de interatividade como filtros nos dados, operações de seleção de objetos, navegação e exploração em representações gráficas diferentes.

3 Trabalhos Relacionados

Alguns exemplos de aplicações serviram como base para compor a abordagem proposta por este trabalho. A VisPublica (RIBEIRO et al., 2012) apresenta técnicas de visualização da informação (InfoVis) que são aplicadas para facilitar a transparência de dados públicos. Esta aplicação permite que os usuários utilizem as técnicas e tecnologias mostradas no catálogo para gerar as visualizações dos seus próprios dados.

A Vis-Saúde (KAIESKI, 2014) é um sistema de código aberto, no qual os registros de saúde pública são usados para fornecer análises. A aplicação propõe uma metodologia de análise e visualização da correlação entre as variáveis associadas à ocorrência de uma enfermidade em uma região.

A DataViva² é uma plataforma de pesquisa com dados abertos que tem como objetivo fornecer dados estatísticos sobre saúde, transporte público, economia, educação e outros domínios de dados associados a várias cidades brasileiras. Recentemente, a DataViva incluiu dados fornecidos pelo DATASUS, o órgão público da saúde do governo brasileiro.

4 Procedimentos Metodológicos

Para viabilização do desenvolvimento da aplicação, foram realizadas reuniões com

² <http://dataviva.info/pt/>

gestores da Secretaria de Saúde do Estado da Paraíba. O levantamento dos requisitos foi feito por meio de entrevistas, em que foram identificados os aspectos que os gestores consideram relevantes para a análise dos dados e que gostariam de visualizar na aplicação.

Por meio dessas reuniões, foram obtidas planilhas em formato CSV com os dados sobre as arboviroses. Os municípios e os dados sobre o quantitativo de habitantes foram obtidos através do IBGE³.

Para embasamento técnico, foram realizados testes com tipos de gráficos diferentes utilizando as linguagens de programação Python⁴ e JavaScript⁵, o objetivo era identificar qual apresentação e tecnologia seriam mais eficazes para o desenvolvimento da análise interativa visual.

A Figura 1 mostra o processo empregado para o desenvolvimento da aplicação.

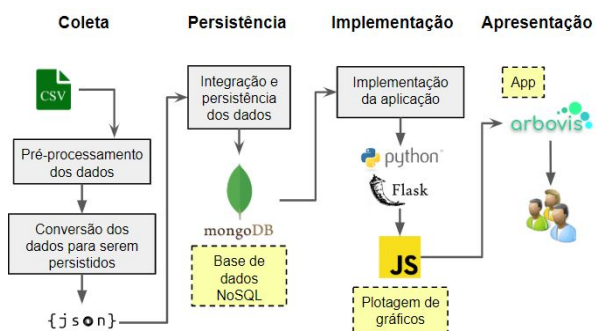


Figura 1: Visão geral do processo para implementação. Fonte: Os autores.

Como mostra a Figura 1, para a implementação da aplicação, primeiramente foi feito o pré-processamento dos dados obtidos. Após isso, os dados foram convertidos para o formato JSON⁶, foram integrados e persistidos na base de dados MongoDB⁷. Em seguida, a aplicação foi

³

<https://www.ibge.gov.br/geociencias-novoportal/organizacao-do-territorio/estrutura-territorial/15761-areas-dos-municipios.html?=&t=downloads>

⁴ <https://www.python.org/>

⁵ <https://www.javascript.com/>

⁶ <http://www.json.org/>

⁷ <https://www.mongodb.com/>

implementada utilizando Python juntamente com o microframework Flask⁸. A integração entre o Python e o MongoDB foi feita através da distribuição PyMongo⁹. Por fim, foi iniciada a plotagem dos gráficos com a biblioteca JavaScript Chart.js¹⁰.

5 Arbovis: abordagem e implementação

A Arbovis foi desenvolvida com foco em apresentar informações sobre as arboviroses de maneira clara e interativa para gestores, profissionais de saúde e para a população em geral.

5.1 Abordagem e funcionalidades

As principais funcionalidades da aplicação são i) filtrar dados, ii) selecionar gráfico e iii) selecionar visualização. De acordo com a pesquisa montada pelo usuário, a consulta é realizada no banco de dados e o resultado é retornado para o usuário. Uma visão geral da abordagem da Arbovis é apresentada na Figura 2.

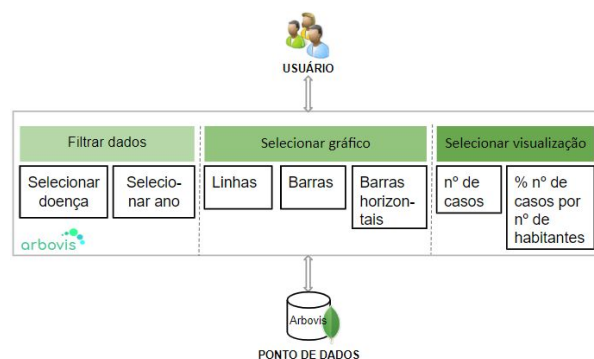


Figura 2: Abordagem da Arbovis. Fonte: Os autores.

5.2 Implementação e cenário de uso

Com o objetivo de facilitar o acesso as visualizações, a seção de indicadores foi evidenciada no painel da página principal, na barra de navegação e no rodapé. A Figura 3 mostra a tela principal da aplicação, com destaque para a opção de montar o gráfico de acordo com os filtros disponibilizados.

⁸ <http://flask.pocoo.org/>

⁹ <https://api.mongodb.com/python/current/>

¹⁰ <https://www.chartjs.org/>



Figura 3: Página principal da Arbovis. Fonte: Arbovis.

Além das funcionalidades principais, o usuário também é informado sobre a sintomatologia das arboviroses, há também uma breve descrição sobre as doenças (Figura 4).



Figura 4: Página principal da Arbovis. Fonte: Arbovis.

A página de indicadores contém a visualização dos dados sobre as doenças. Esta seção é composta por filtros que o usuário pode selecionar de acordo com o objetivo da sua pesquisa, as opções selecionadas geram o gráfico dinamicamente (Apêndice A). Os filtros apresentados são i) Selecionar doença; ii) Selecionar ano; iii) Selecionar tipo de gráfico; iv) Selecionar tipo de visualização.

Considerando um cenário no qual no filtro de doença é selecionada a chikungunya, no filtro de ano de ocorrência é selecionado o ano de 2017, o gráfico apresentado é o de linhas, que por padrão já é selecionado, e número de casos absoluto (Apêndice B), é possível perceber que o município mais representativo, com 4556 casos registrados, é João Pessoa. Entretanto, se for aplicada a visualização por porcentagem do número de casos por número de habitantes, como mostrado no Apêndice C, o município de São Bentinho se destaca entre os outros, com 5.35% de casos por número de habitantes.

Outro cenário pode ser visto no Apêndice D, em que são selecionados a doença dengue e o ano de 2016 com gráfico de barras e número de casos absolutos registrados por município. Neste contexto, o município com mais registros é João Pessoa, com 7701 casos. Já no cenário de porcentagens do número de casos pelo número de habitantes, como mostrado no Apêndice E, o município mais significativo é Juripiranga, com 27.58% de casos por número de habitantes.

Através das visões apresentadas, é possível identificar quais são os municípios que necessitam de maior atenção do governo, por possuírem maior incidência de casos das doenças. Dessa forma, as visualizações disponibilizadas pela Arbovis possibilitam ações mais assertivas por parte dos gestores.

5.3 Avaliação da Arbovis

Com o objetivo de analisar aspectos da Arbovis como a usabilidade na utilização e apresentação dos indicadores, foi realizada uma avaliação inicial com um gestor da Secretaria de Saúde da Paraíba.

No aspecto da utilização da Arbovis, o gestor considerou a aplicação fácil de usar e também indicou que ela o ajudou a encontrar informações importantes sobre as arboviroses.

Outro aspecto apontado é a apresentação dos indicadores, em que o usuário considerou os filtros adequados para

a realização das consultas e que tornaram a aplicação interativa. Além disso, informou que a aplicação apresenta praticidade nos resultados apresentados. Como sugestão de melhoria, foi sugerido acrescentar índices de infestação predial por município junto às taxas de incidência dos agravos, estes são dados que serão futuramente integrados à base existente.

6 Considerações Finais

A Arbovis é pautada em um problema social e possibilita o uso de dados reais em um cenário de demandas concretas a esses dados. Os resultados obtidos com esta aplicação possuem impacto especial na gestão em relação a prevenção de doenças e epidemias causadas pelo *Aedes aegypti*.

Pesquisadores, professores, gestores ou quaisquer interessados poderão consultar os dados de forma integrada e analisá-los de acordo com suas necessidades. Para os gestores, a aplicação é significativa pois os resultados obtidos através das visões customizadas podem contribuir para a tomada de decisão. Entretanto, para expandir as possibilidades de análise dos gestores esperamos incluir em trabalhos futuros o uso de mapas e averiguar a eficácia dessa visualização na gestão.

Também podem ser identificadas necessidades específicas de diferentes perfis, para contribuir em maior profundidade em seus estudos e análises. Um perfil que se destaca para ser explorado é o profissional de saúde.

7 Referências

BANSAL, S.; KAGEMANN, S. Integrating Big Data: A Semantic Extract-Transform-Load Framework. **Computer**, vol.48, n.3, p. 42-50, 2015.

CARMO, A.. **Exploração e análise de dados coletados pelo Sistema Integrado de Monitoramento Ambiental utilizando técnicas de Visual Analytics**. 2015. Dissertação de Mestrado. Faculdade de Ciências e Tecnologia (FCT), Presidente

Prudente, 2015. Disponível em: <<https://repositorio.unesp.br/handle/11449/138391>>. Acesso em: 04 ago. 2018.

FEKETE, J. Visual Analytics Infrastructures: From Data Management to Exploration. **Computer**, Institute of Electrical and Electronics Engineers, Visual Analytics: Seeking the Unknown, vol.46, n.7, p.22-29, 2013.

GANDOMI, A.; HAIDER, M. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, Ontario, v.35, p.137-144, 2015. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0268401214001066>>. Acesso em: 05 out. 2018.

KAIESKI, N. **Vis-Saúde - Uma metodologia para visualização e análise de dados de saúde pública**. 2014. Dissertação de Mestrado (UNISINOS), São Leopoldo, 2014. Disponível em: <<http://www.repositorio.jesuita.org.br/handle/UNISINOS/3619>>. Acesso em: 04 ago. 2018.

RIBEIRO, F. et al. VisPublica: uma proposta para aprimorar a transparência de dados públicos. In: **VIII Simpósio Brasileiro de Sistemas de Informação**, 2012, São Paulo. p. 480-491.

SANTIAGO, M., ALENCAR A., FERNANDES A., SOUZA, D. (2018). A Semantic-based Approach for Facilitating Arbovirus Data Usage. 20^o International Conference on Enterprise Information Systems (ICEIS), Funchal, 2018.

SHNEIDERMAN, B., PLAISANT, C., HESSE, B. Improving health and healthcare with interactive visualization methods. **Computer**, vol. 46, n. 5, p. 58-66, 2013.

THABET, N.; SOOMRO, T. Big Data Challenges. **Journal of Computer Engineering & Information Technology**.

Dubai, 2015. Disponível em:
<https://www.researchgate.net/publication/282281171_Big_Data_Challenges>. Acesso em: 04 out. 2018.

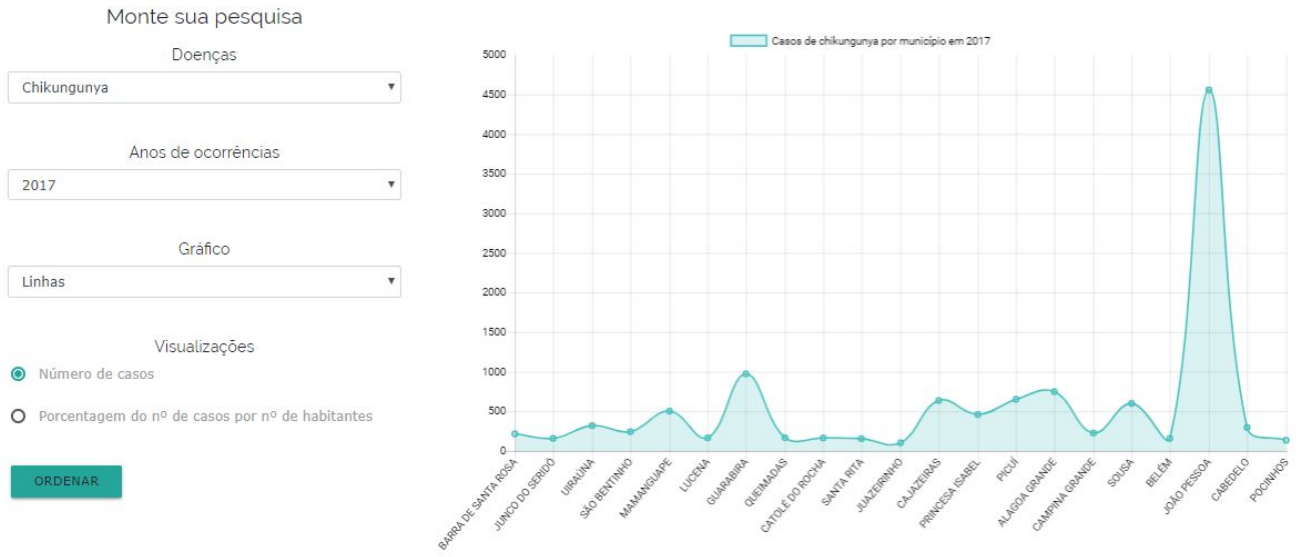
THOMAS, J. J.; COOK, K. A. (2005).
Illuminating the Path: The Research and
Development Agenda for Visual Analytics.
[S.l.]: National Visualization and Analytics Ctr.
Paperback. ISBN 0769523234.

VIEIRA, J., CORREA, R. Visualização da
Informação na construção de interfaces

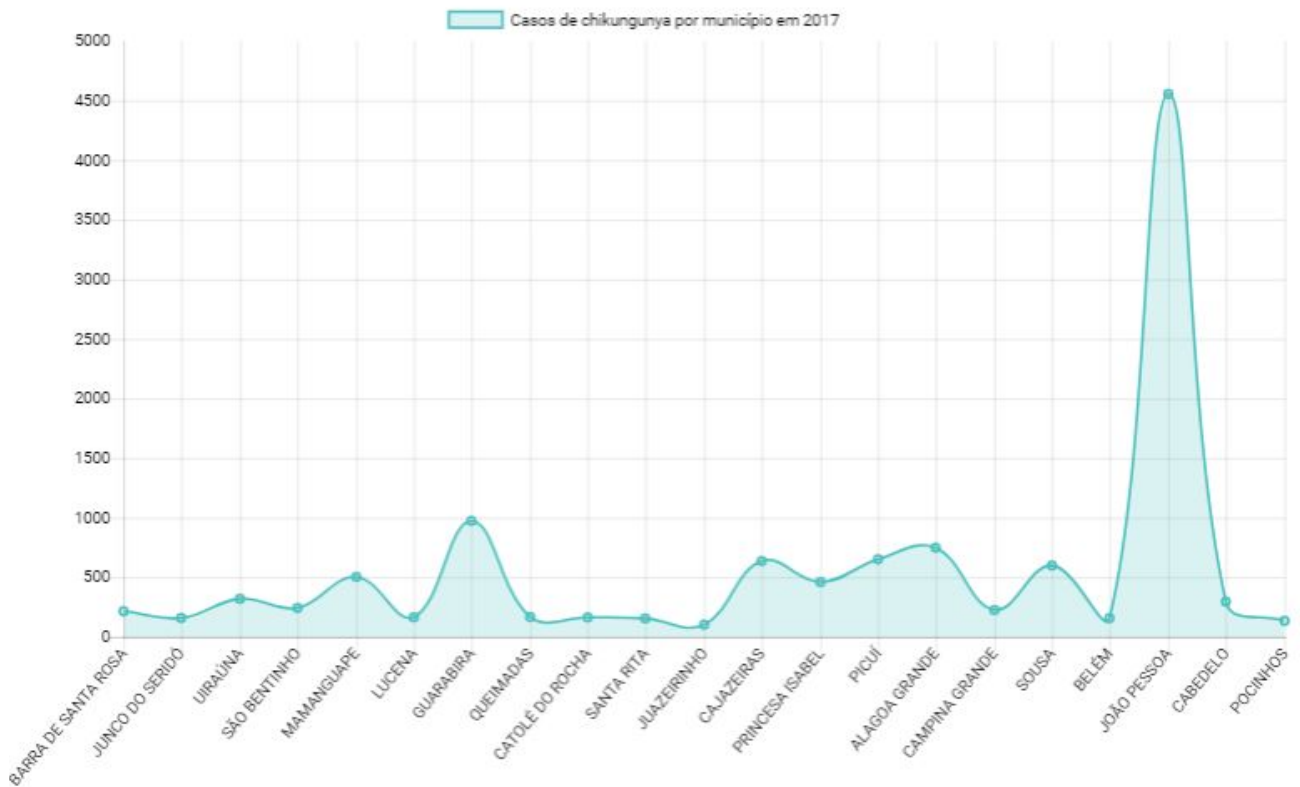
amigáveis para Sistemas de Recuperação de
Informação. **Encontros Bibli: revista
eletrônica de biblioteconomia e ciência da
informação**, Florianópolis, v. 16, n. 32, p.
73-93, out. 2011. ISSN 1518-2924.

Disponível em:
<<https://periodicos.ufsc.br/index.php/eb/articloe/view/1518-2924.2011v16n32p73>>. Acesso em: 04 ago. 2018.

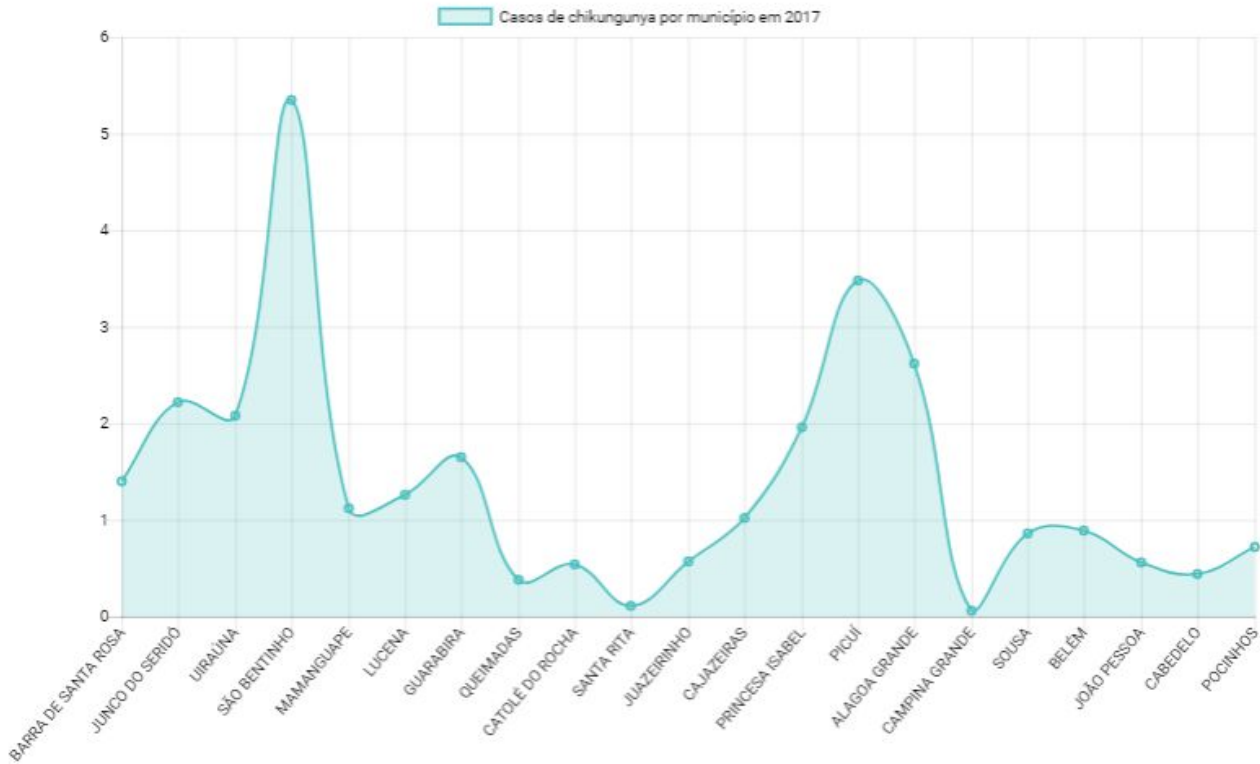
Apêndice A - Filtros selecionados para visualização customizada. Fonte: Arbovis.



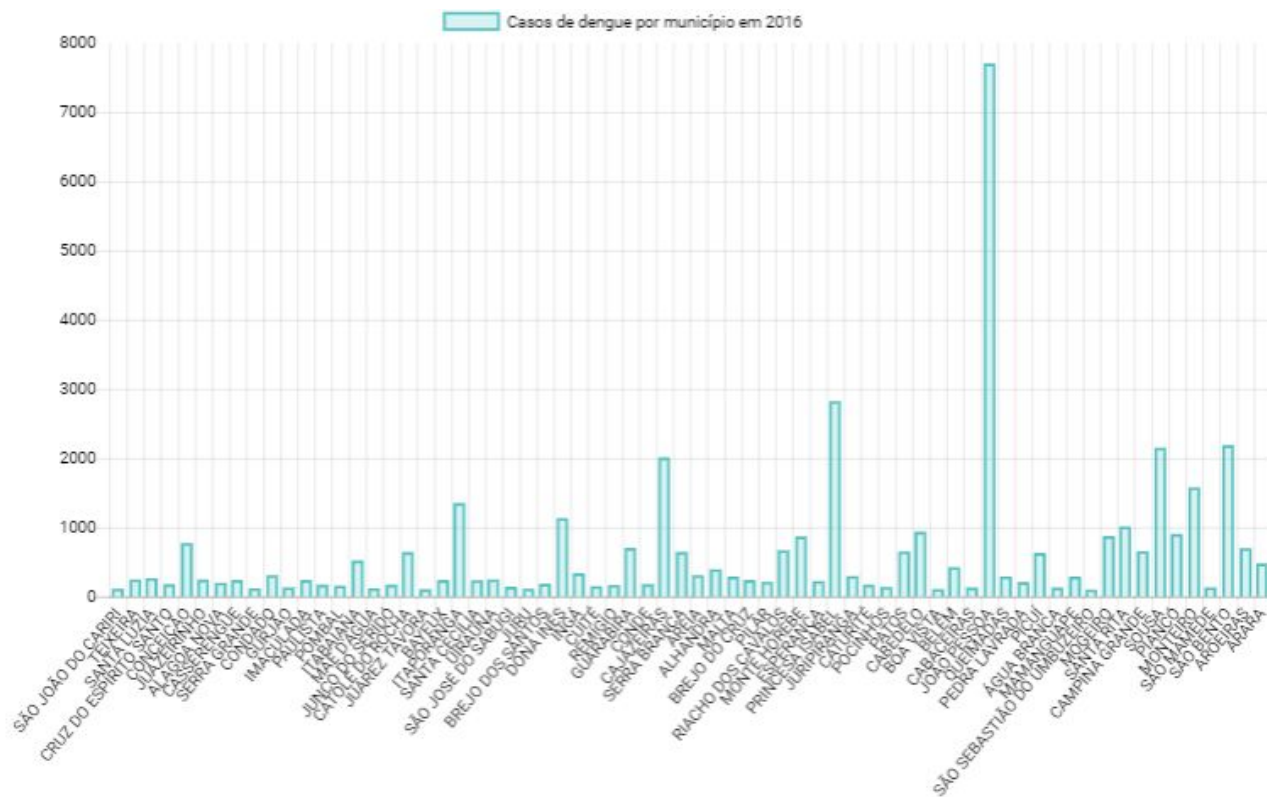
Apêndice B - Gráfico de linhas com casos de chikungunya por município em 2017 com número de casos total. Fonte: Arbovis.



Apêndice C - Gráfico de linhas com casos de chikungunya por município em 2017 com porcentagem de casos por número de habitantes. Fonte: Arbovis.



Apêndice D - Gráfico de barras com casos de dengue por município em 2016 com número de casos total. Fonte: Arbovis.



Apêndice E - Gráfico de barras com casos de dengue por município em 2016 com porcentagem de casos por número de habitantes. Fonte: Arbovis.

