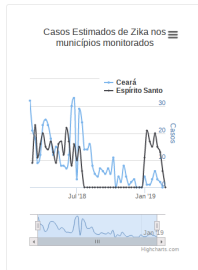
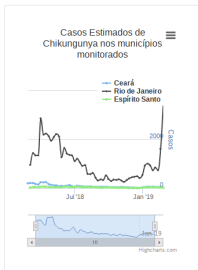
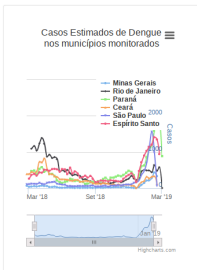


O projeto InfoDengue: Python no dia-a-dia da pesquisa

Flávio Codeço Coelho

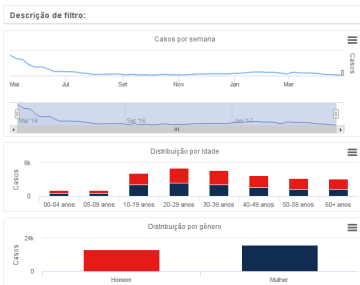
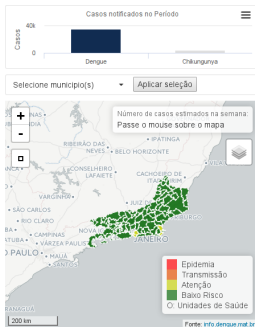
March 20, 2019

Situação Epidemiológica - 790 Municípios monitorados



Rio de Janeiro

Última atualização: 4 de Maio de 2017



Infraestrutura de Software

Open-source: 18 repositórios: Linguagem Python R.

The screenshot shows the GitHub repository page for 'Alerta Dengue project'. At the top, there is a profile picture of a blue map of Brazil and the repository name. Below this, a description states: 'Alerta Dengue project is the maintainer of the infodengue website which provides real-time epidemiological analyses to more than 500 cities.' It also lists the location as 'Rio de Janeiro, Brazil', the website 'http://info.dengue.mat.br', and the email 'alerta_dengue@focruz.br'. Navigation tabs include 'Repositories 18', 'People 18', 'Teams 0', 'Projects 0', and 'Settings'. A search bar is present with filters for 'Type: All' and 'Language: All', along with 'Customize pins' and a 'New' button. The main content area lists several repositories:

- AlertaDengue**: Portal de dados do Projeto Alerta Dengue. Includes tags for science, surveillance, epidemiology, and dengue-fever. It is a Jupyter Notebook repository with 13 stars, 8 forks, GPL-3.0 license, and 1 issue needs help. Updated 2 days ago.
- AlertaDengueAnalise**: Repository for R with 2 stars and 6 forks. Updated 2 days ago.
- PySUS**: funções e classes para auxiliar no tratamento de dados do SUS. Includes tags for data-science, geospacial, and health. It has 19 issues, 12 forks, and 3 issues need help. Updated 6 days ago.
- vigibot**: Telegram bot to interact with epidemiologists, providing insights about disease surveillance.

On the right side, there are two sidebars. The 'Top languages' sidebar shows Jupyter Notebook (orange), Python (blue), R (light blue), HTML (red), and CSS (purple). The 'People' sidebar shows a grid of 18 profile pictures and an 'Invite someone' button.

Captura automática de dados

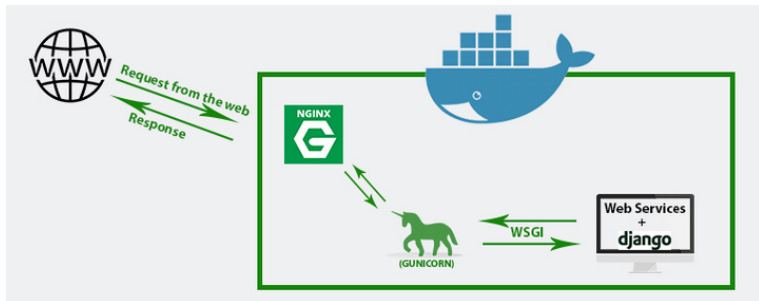
- ▶ Dados de Estações meteorológicas (METAR)
- ▶ Dados de imagens de satélites (Google Earth Engine, NASA)
- ▶ Dados do SUS
- ▶ Dados do Twitter
- ▶ Execução assíncrona da captura com Celery



<https://github.com/AlertaDengue/AlertaDengueCaptura>

Aplicação Web

Deploy com Docker Compose + Ansible



Análise de Dados e Visualização

- ▶ Dados armazenados em banco PostgreSQL
- ▶ Consultas, Manipulação e análise com Pandas
- ▶ Visualização: Javascript, Plotly, leaflet, etc.

Leaflet 

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



plotly



PostgreSQL 

Modelagem Matemática

- ▶ EDOs
- ▶ Processos Estocásticos



Modelo SIR

Modelo epidemiológico, com taxa de transmissão variável.

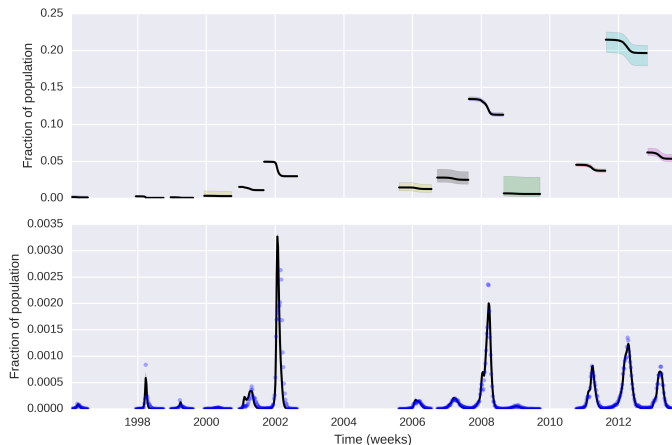
$$\frac{dS}{dt} = -\beta(t)SI$$

$$\frac{dI}{dt} = \beta(t)SI - \tau I$$

$$\frac{dR}{dt} = \tau I$$

Modelos + Dados

Ajustando modelo aos dados (Rio de Janeiro) para estimar $S(0)$ ¹.

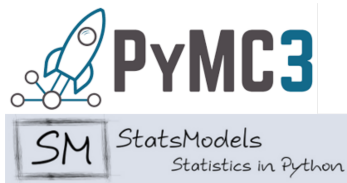


¹Estimating the Attack Ratio of Dengue Epidemics under Time-varying Force of Infection using Aggregated Notification Data FC Coelho, LM Carvalho Scientific Reports 5, 18455

Modelagem Estatística

No projeto Infodengue empregamos diversas técnicas de modelagem estatísticas.

- ▶ Inferência Bayesiana com PyMC3 e PyStan
- ▶ Modelagem estatística com R e StatsModels



Delimitando a estação Epidêmica

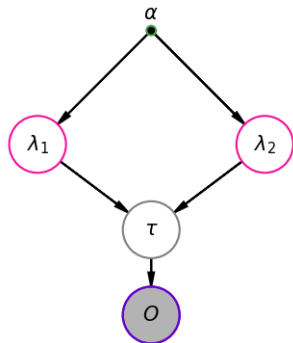
Sejam os casos observados de Zika um processo Poisson:

$$O_t \sim \text{Poisson}(\lambda_j).$$

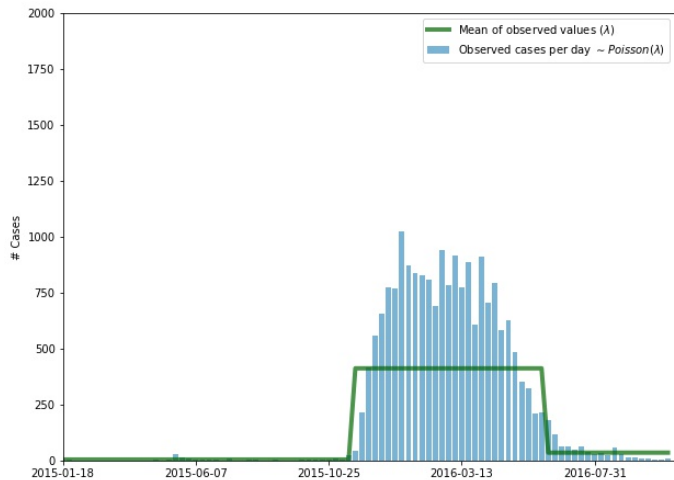
$$\lambda = \begin{cases} \lambda_1 & \text{if } t \leq \tau_1 \\ \lambda_2 & \text{if } \tau_1 \leq t \leq \tau_2 \\ \lambda_3 & \text{if } \tau_2 \leq t \end{cases}$$

Assuming these priors for the parameters: $\tau \sim \mathcal{U}\{a = 0, b = \hat{n}\}$ and $\frac{1}{\lambda_i} \sim \text{Exp}(\alpha = \lambda_i)$, we can estimate posterior distributions for them, given the observed data.

Bastos, M. M., & Coelho, F. C. (2018). Estimating under-observation and the full size of the 2016 Zika epidemic in Rio de Janeiro. PLOS ONE, 13(10), e0205001.



Delimitando a estação Epidêmica

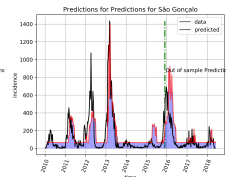
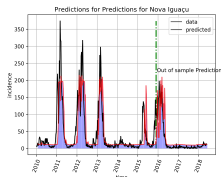
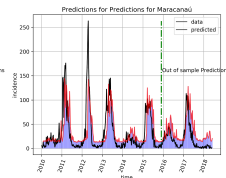
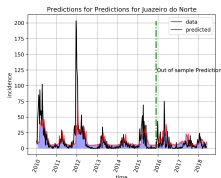
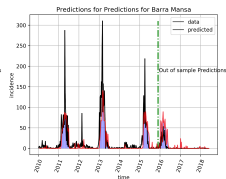
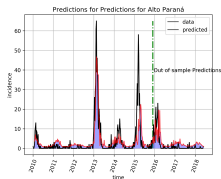


Aprendizagem de Máquinas

Previsão de séries temporais, não estacionárias, não ergódicas...

- ▶ Deep Learning: LSTM
- ▶ Tensorflow
- ▶ Keras
- ▶ Scikit-Learn

github.com/AlertaDengue/InfoDenguePredict



Interatividade

- ▶ VigiRobot: Robô epidemiologista capaz de interagir com profissionais de saúde fornecendo dados e análises sob demanda.
- ▶ Combinando NLU com Data-science



Equipe Infodengue

Não seria possível sem o esforço de muitos...



...