

Universidade Federal de Pernambuco - Centro de Informática

IF706 - Introdução a Ciência dos Dados

Projeto Prático.

Tema: Investigação de casos de Dengue, Zika e Chikungunya no Brasil

Equipe:

- Eduardo Luiz Silva <els6@cin.ufpe.br>
- Thiago Augusto dos Santos Martins <tasm2@cin.ufpe.br>
- Wellington Barbosa de Almeida <wba@cin.ufpe.br>

1. Proposta

Levando em consideração o contexto das doenças tropicais negligenciadas, em especial Dengue, Zika e Chikungunya, e seu impacto social e, segundo IANPHI (2016), econômico na América Latina, especialmente no Brasil, o projeto possui como proposta:

- I. Analisar a relação entre o perfil social dos indivíduos afetados e as incidências de doenças transmitidas pelo mosquito *Aedes Aegypti*;
- II. Investigar a influência da localidade física e climáticas em que se encontram tais indivíduos;
- III. Encontrar padrões nos casos relatados para assim sugerir abordagens de vigilância que possam ser implementados por governos.

O objetivo é analisar essas hipóteses e tentar validá-las ou não em relação a Recife, Pernambuco, mas se não for possível, devido a falta de dados, então o escopo será analisar globalmente.

2. Coleta de dados

As bases de dados a seguir foram o ponto de partida para a criação da proposta:

- **Casos de Dengue, Zika e Chikungunya**
 - Registro dos casos de Dengue, Zika e Chikungunya com registros nas unidades de saúde, públicas ou particulares.
 - Disponível em
<<http://dados.recife.pe.gov.br/dataset/casos-de-dengue-zika-e-chikungunya>>

- **Zika Virus Epidemic**
 - Disponível em <<https://www.kaggle.com/cdc/zika-virus-epidemic>>
- **Dengue, Temperatura e Chuvas em Campinas-SP**
 - Dataset contendo o número de casos mensais confirmados de dengue no município de Campinas/SP de 1998 a 2015. Além disso foram coletados dados de chuvas, temperatura média, mínima e máxima na cidade.
 - <https://www.kaggle.com/renangomes/dengue-temperatura-e-chuvas-em-campinas>

3. Metodologia de Desenvolvimento do Projeto

Para desenvolvimento do projeto será levado em consideração a criação de Sprints, que representarão as atividades a serem desenvolvidas no período curto de dias, para que seja possível desenvolver todo o escopo esperado até a apresentação do projeto .

As atividades foram desenvolvidas a se pensar no desenvolvimento completo do projeto, mas nada impede que possamos modificar prioridade de atividades, ou datas, já que pode-se finalizar uma sprint antes do tempo, ou por algum motivo, não conseguir terminar as atividades ao final do tempo estipulado.

| Atividades | Datas Importantes |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| Coleta de Dados Analisar viabilidade de utilização de dados locais ou globais. Se o escopo escolhido será nacional ou mundial. | 14/06 - Aula |
| Pré-processamento dos dados Analisar o que realmente será útil dos dataframes escolhidos. | 17/06 - Checkpoint |
| Análise Exploratória de Dados Verificar se inicialmente já podemos tirar conclusões sobre as hipóteses iniciais, ou se surgem outras hipóteses a analisar os dados. | 19/06 - Aula |
| Modelos de Aprendizado Analisar padrões nos dados e o que poderá ser utilizado para desenvolver um modelo que tira conclusões sobre novos dados, em relação às hipóteses escolhidas. | 21/06 - Aula |
| Visualização de Dados Análise discreta e concisa dos resultados, de forma gráfica. | 26/06 - Aula |
| Desenvolver slides para apresentação dos resultados | 28/06 - Aula |
| Entrega do Projeto | 01/07 |

Referências

IANPHI. **The Zika Virus Epidemic: Challenges and the Road Ahead**, 2016. Disponível em:<<http://ianphi.org/news/2016/zikavirusupdate.html>>.