



Quem se prepara, não para.

Business Intelligence

4º período

Professora: Michelle Hanne

Tutorial ETL + indicadores no PowerBI

Executar o script de coleta de dados sobre as epidemias: dengue, Chikungunya, zika em Belo Horizonte de 2022 até 2023

Baixar o arquivo

ColetaDadosEpidemia_Tutorial.ipynb

Executar o arquivo. Será gerado no seu drive um arquivo com o nome

dadosepidemiologicos.csv

```
# 3106200 = Belo Horizonte
url = "https://info.dengue.mat.br/api/alertcity"
geocode = 3106200
```

```
format = "csv"
ew_start = 1
ew_end = 53
ey_start = 2022
ey_end = 2023
```

```
disease = ['dengue', 'chikungunya', 'zika']
for n in disease:
    print(n)
    params = (
        "&disease="
        + f"{n}"
        + "&geocode="
        + f"{geocode}"
        + "&disease="
        + f"{n}"
        + "&format="
        + f"{format}"
        + "&ew_start="
        + f"{ew_start}"
        + "&ew_end="
        + f"{ew_end}"
        + "&ey_start="
        + f"{ey_start}"
        + "&ey_end="
        + f"{ey_end}"
    )
```

```
url_resp = "?".join([url, params])
print(url_resp)
```

```
dados = pd.read_csv(url_resp, index_col='SE')
dados['disease'] = n
if (n=='dengue'):
    dadosdengue=dados
elif (n=='zika'):
    dadoszika=dados
elif (n=='chikungunya'):
    dadoschiku=dados
```

O Script faz um loop nos tipos de epidemias, coleta os dados e salva cada um em um dataframe

```
final_df = pd.concat([dadosdengue, dadoszika,  
dadoschiku], ignore_index=True)
```

Concatena os 3 dataframes em apenas um único “final_df”

```
final_df.to_csv("/content/drive/MyDrive/Colab  
Notebooks/Exercicio_Epidemiologia/dadosepidemiologicos.csv", index=True)
```

```
final_df.head()
```

Gera o arquivo final de saída

ETL – separando as tabelas Fato e Dimensão

Tabela Fato

data_iniSE
casos_est
casos_est_min
casos_est_max
casos
tweet
umidmed
tempmin
tempmax
disease

Tabelas Dimensões

```
✓ [36] df_nivel.head()
```

	nivel	descricao
0	4	red
1	3	orange
2	2	yellow
3	1	green

```
✓ df_doenca.head()
```

	disease	indice
0	dengue	1
1	zika	2
2	chikungunya	3

```
✓ df_receptivo.head()
```

	receptivo	descricao
0	1	favorable
1	0	unfavorable

ETL – separando as tabelas Fato e Dimensão

Executar o script
de para criar as
tabelas
dimensões e fato

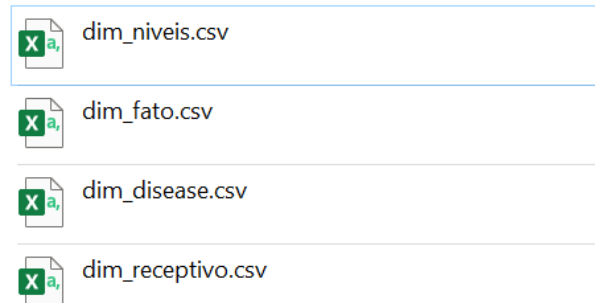
Baixar o arquivo

Manipulando_Dados_Epidemias.ipynb

Executar o arquivo. Será gerado no seu drive as tabelas fato e dimensões.

PowerBI

- 1- Acessar o PowerBI
- 2- Inserir os arquivos CSV criados no python:
- 3- Fazer a separação da coluna Data em outra tabela dimensão



PowerBI - Dashboard

- 1- Mostrar a comparação do número de casos de cada epidemia ao longo do tempo
- 2- Fazer um filtro por tipo de doença que deverá ser aplicado em todo o relatório
- 3- Fazer um gráfico comparando as variáveis temperatura mínima vs casos, aplicar o filtro por seleção do tipo de doença.
- 4- Fazer um gráfico comparando as variáveis temperatura mínima e máxima, aplicar o filtro por seleção do tipo de doença.
- 5- Fazer um gráfico comparando o número de casos com tweets.

Fazer um filtro por tipo de doença que deverá ser aplicado em todo o relatório