

TECNOLOGIA EM SISTEMAS PARA INTERNET

**Rosana da Silva Soares
Kennedy Viana Aguiar
Aurélio Vinícius França dos Santos**

**RELATÓRIO 2 DE PRÁTICA INTEGRADA
DE
CIÊNCIA DE DADOS E INTELIGÊNCIA ARTIFICIAL**

Brasília - DF

15 de Março de 2021

Sumário

1. Objetivos	3
2. Descrição do problema	4
3. Desenvolvimento	5
3.1 Código implementado	5
4. Considerações Finais	6
Referências	7

1. Objetivos

Este segundo relatório tem como objetivo apresentar a exploração de dados do site The National Ufo Reporting, com gráficos e mapas dos quatro estados que possuem maiores relatos de OVNIS. Além também criar um mapa do País inteiro dos EUA a fim de demonstrar as ocorrências de cada estado, e um mapa apenas para o estado da Califórnia.

2. Descrição do problema

Na primeira parte do projeto de prática integrada conhecemos o site The National Ufo Reporting, onde foram analisados os dados sobre avistamento de ovnis, transformando esses dados em tabelas. Para melhor visualização dos mesmos, foi proposto desenvolver uma nova exploração desses dados em formatos de gráficos e mapas, com o intuito de, construir uma representação visual da quantidade de ocorrências por cidade e estado.

3. Desenvolvimento

O desenvolvimento do problema proposto se deu da seguinte forma:

Passo 1: trabalhar em cima dos dados coletados;

Passo 2: Coletar informações dos tipos de ovnis, cidades, e quantidades de aparições em cada local.

Passo 3: gerar mapas com as informações coletadas.

Passo 4: Dar um merge nas informações coletadas com as coordenadas das cidades

Passo 5: Gerar mapa dos EUA e pontilhar os relatos respectivamente em seus devidos locais

Onde na Califórnia está localizada a maior quantidade de visualizações de objetos voadores não identificados? Investigue e descreva a possível razão para esse local ter a maior quantidade de visualizações..

A cidade de los angeles concentra o maior quantidade de ufo reports, os possíveis motivos são: maior concentração populacional vinculado com a cultura da cidade como filmes, eventos e convenções sobre aliens. Além disso a proximidade com o estado de nevada, onde está localizada a área 51 contribui para tal cultura. Não somente a isso o estado da califórnia é estado mais cosmicamente inclinado , o que contribui para a visualização do espaço.

3.1 Código implementado

```
# IMPORTACOES DE MODULOS E BIBLIOTECAS
```

```
pip install zipcodes
from folium.plugins import HeatMap
import folium
import zipcodes
from bs4 import BeautifulSoup
from urllib.request import urlopen
from urllib.error import HTTPError
import csv
import matplotlib
import matplotlib.pyplot as plt
import numpy as np
```

```
import pandas as pd
from collections import Counter
```

```
# Criando o dataframe
data = pd.read_csv('OVNIS.csv')

# Os 4 estados com maiores aparições
x = data.estado.value_counts()
x = dict(list(x.items())[0:4])

# Coletando todos os registros de formatos de cada estado
formatos = {}
for i, row in data.iterrows():
    if(row.estado in x):
        if(row.estado not in formatos):
            formatos[row.estado] = []
        else:
            formatos[row.estado].append(row.formato)
```

```
# Montando a estrutura para plotagem
dados = ([], [], [], [])
index = 0

for i in formatos.values():
    c = Counter(i)
    dados[index].append(c['Light'])
    dados[index].append(c['Circle'])
    dados[index].append(c['Fireball'])
    dados[index].append(c['Triangle'])
    index += 1
```



```
# Plotagem do gráfico 1

dim = len(dados[0])
w = 0.75
dimw = w / dim

fig, ax = plt.subplots()
x = np.arange(len(dados))
for i in range(len(dados[0])):
    y = [d[i] for d in dados]
    b = ax.bar(x + i * dimw, y, dimw, bottom=0.001)

ax.set_xticks(x + dimw)
ax.set_xticklabels(["CA", "WA", "FL", "TX"])

ax.set_title('Gráfico dos estados com mais ocorrências')
ax.set_xlabel('Estados')
ax.set_ylabel('Número de ocorrências')

plt.show()
```

```
#gerando o grafico de barras gráfico 2
Light = np.array((1167, 442, 623, 555))
Circle = np.array((602, 290, 463, 252))
Fireball = np.array((562, 267, 486, 183))
Triangle = np.array((397, 267, 225, 173))

states = ['CA', 'WA', 'FL', 'TX']

plt.figure(figsize=(7,5))

plt.bar(states, Light, color = 'green')
plt.bar(states, Circle, color = 'yellow')
plt.bar(states, Fireball, color = 'blue')
plt.bar(states, Triangle, color = 'red')

plt.xlabel('estados')
plt.ylabel('visualizações')
plt.title('grafico de barras')
plt.legend(('Light', 'Circle', 'Fireball', 'Triangle'))

plt.show()
```

```

# GERANDO CSV COM AS COORDENADAS DE CADA CIDADE

csv_columns = ['ID', 'cidade', 'estado', 'lat', 'long']
csvfile = 'coords.csv'
coords = []

for i in ovnis:
    try:
        lat = zipcodes.filter_by(city=ovnis[i]['cidade'],
state=ovnis[i]['estado'])[0]['lat']
        lng = zipcodes.filter_by(city=ovnis[i]['cidade'],
state=ovnis[i]['estado'])[0]['long']
    except IndexError:
        lat = 32.7199
        lng = -117.1805

    coord = {
        'ID': i,
        'cidade': ovnis[i]['cidade'],
        'estado': ovnis[i]['estado'],
        'lat': lat,
        'long': lng
    }
    coords.append(coord)
    print('COORDS: ', len(coords))

    if i == '1000':
        break

try:
    with open(csvfile, 'w') as csvfile:
        writer = csv.DictWriter(csvfile, fieldnames=csv_columns)
        writer.writeheader()
        for crd in coords:
            writer.writerow(crd)
except IOError:
    print("I/O error")

```

```
# Gera o mapa dos Estados Unidos com suas respectivas ocorrências
(1997-2017)
coords = pd.read_csv('coords.csv')

lat = coords.lat.tolist()
lng = coords['long'].tolist()

m = folium.Map(location=[37.8427887, -98.3807258],
tiles='cartodbdark_matter', zoom_start=4.25)
HeatMap(list(zip(lat, lng)), overlay=False).add_to(m)
m
```

```
# Gera as coordenadas das cidades dos Estados Unidos e salva em csv
csv_columns = ['ID', 'cidade', 'estado', 'lat', 'long']
csvfile = 'coords_ca.csv'
coords = []

for i in ovnis:
    try:
        if ovnis[i]['estado'] == 'CA':
            lat = zipcodes.filter_by(city=ovnis[i]['cidade'],
state=ovnis[i]['estado'])[0]['lat']
            lng = zipcodes.filter_by(city=ovnis[i]['cidade'],
state=ovnis[i]['estado'])[0]['long']
        else:
            continue
    except IndexError:
        continue

    coord = {
        'ID': i,
        'cidade': ovnis[i]['cidade'],
        'estado': ovnis[i]['estado'],
        'lat': lat,
        'long': lng
    }
    coords.append(coord)

coords = pd.DataFrame(coords)
coords.to_csv('coords.csv')
```

```
# Gera as coordenadas das cidades da Califórnia e salva em csv
coords = pd.read_csv('coords.csv')
coords_ca = []
for i, row in coords.iterrows():
    if row.estado == 'CA':
        try:
            lat = row.lat
            lng = row['long']
        except IndexError:
            continue

        coords_ca.append({
            'ID': i,
            'cidade': row.cidade,
            'estado': row.estado,
            'lat': lat,
            'long': lng
        })
    else:
        continue

coords_ca = pd.DataFrame(coords_ca)
coords_ca.to_csv('coords_ca.csv')
```

```
# Gera o mapa da Califórnia com suas respectivas ocorrências
cca = pd.read_csv('coords_ca.csv')

lati = cca.lat.tolist()
longi = cca['long'].tolist()

mapa = folium.Map(location=[37.2454668, -120.7021918],
tiles='cartodbdark_matter', zoom_start=6)
HeatMap(list(zip(lati, longi)), overlay=False).add_to(mapa)
mapa
```

4. Considerações Finais

A sprint foi muito desafiadora. Conhecemos novas bibliotecas e novos métodos, o que está garantindo uma grande gama de conhecimento.

Referências

The National UFO Reporting Center. Nuforc, 2021. Disponível em: <<http://www.nuforc.org/>>. Acesso em: 14 de Março. 2021.

Python package index. Python, 2021. Disponível em: <<https://pypi.org/project/zipcodes/>>. Acesso em 14 de Março.2021.

Matplotlib Basemap Toolkit. Matplotlib, 2016. Disponível em:<<https://matplotlib.org/basemap/>>. Acesso em: 14 de Março. 2021.