

TECNOLOGIA EM SISTEMAS PARA INTERNET

**Emanuelle Ferreira da Silva
Lucas Matheus Silva
Neander Wendel Nobre Teixeira**

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

Brasília - DF

03/10/2020

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

O presente projeto se encontra na fase da exploração dos dados por meio de gráficos e mapas, tem como objetivo representar os dados que foram trabalhando na etapa anterior por meio de gráficos de barra agrupado e empilhado e por meio do mapa para uma melhor interpretação e visualização dos dados referentes as ocorrências registradas nos estados unidos.

2. Descrição do problema

O problema dessa etapa será dividido em duas etapas na construção dos gráficos de barra agrupada e empilhado e na criação do mapa onde os dados relacionado aos estados unidos serão explorados e mostrados nos resultados.

3. Desenvolvimento

Para o desenvolvimento dessa etapa foi necessário a utilização e algumas bibliotecas como o pandas para trabalhar, organizar, executar e salvar os dados, com o seaborn e matplotlib para a construção dos gráficos representados logo abaixo gerando um gráfico de barra empilhada e outro agrupada, foram utilizada a biblioteca pandasql por meio do sql.

3.1 Código implementado

1- Por meio da query é possível limitar os estados dos estados unidos para os quatros estados(TX, CA, WA, FL) com os seus respectivos shapes que será trabalhado por meio dos gráficos.

```
2 q = """
3
4 SELECT*from quantidade_estados_unidos where (State LIKE '%WA%' OR State LIKE '%CA%' OR state like '%TX%' or
5 State like '%FL%') and (Forma like '%Light%' OR Forma like '%Fireball%' OR Forma like '%Sphere%' OR Forma like '%Circle%') limit 16
6 """
7 quatro_estados_estados_unidos = pandasql.sqldf(q.lower(),locals())
8 quatro_estados_estados_unidos
```

Resultado do código acima

	State	quantidade de ocorrencia	forma
0	CA	2526	Light
1	CA	1216	Circle
2	WA	1152	Light
3	FL	1106	Light
4	CA	1012	Fireball
5	TX	870	Light
6	CA	815	Sphere
7	FL	658	Circle
8	FL	639	Fireball
9	WA	475	Circle
10	WA	465	Fireball
11	FL	402	Sphere
12	TX	394	Circle
13	WA	340	Sphere
14	TX	262	Fireball

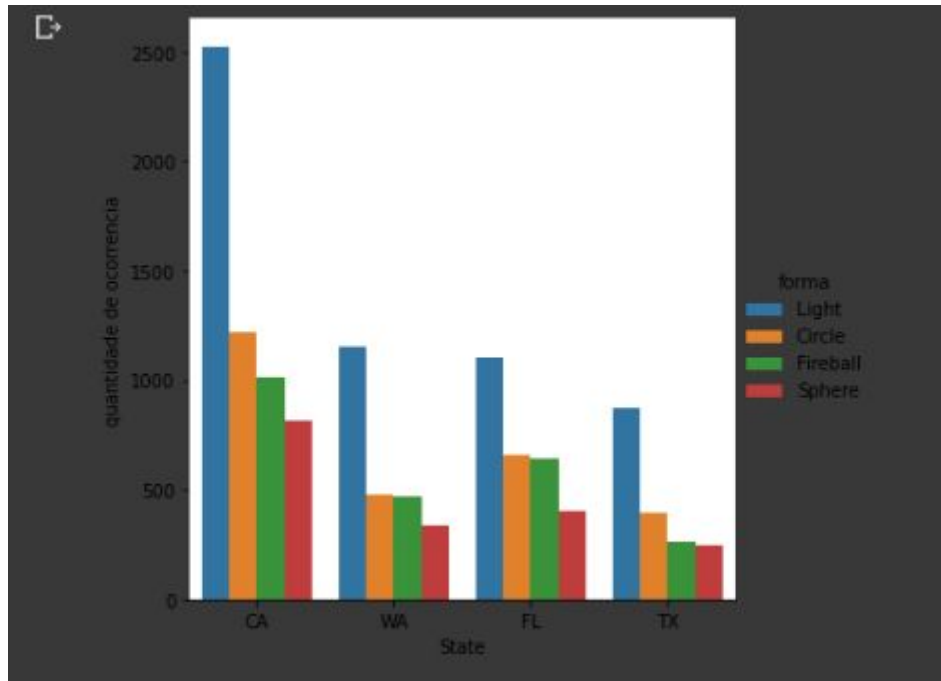
2- Importando as bibliotecas seaborn e matplotlib

```
2 import seaborn as sns
3 import matplotlib.pyplot as plt
```

3-Criando um novo data frame chamada grafic_bar, utilizando a biblioteca seaborn com a função cat plot com as propriedades para criar o gráfico dos quatros estados dos estados unidos.

```
2 grafic_bar = sns.catplot( x="State", y="quantidade de ocorrencia", kind="bar",hue="forma",data=quatro_estados_estados_unidos);
```

Resultado do código acima(gráfico gerado)



4-executando o df

```
1 quatro_estados_estados_unidos
```

Resultado da chamada do df

	State	quantidade de ocorrencia	forma
0	CA	2526	Light
1	CA	1216	Circle
2	WA	1152	Light
3	FL	1106	Light
4	CA	1012	Fireball
5	TX	870	Light
6	CA	815	Sphere
7	FL	658	Circle
8	FL	639	Fireball
9	WA	475	Circle
10	WA	465	Fireball
11	FL	402	Sphere
12	TX	394	Circle
13	WA	340	Sphere
14	TX	262	Fireball
15	TX	242	Sphere

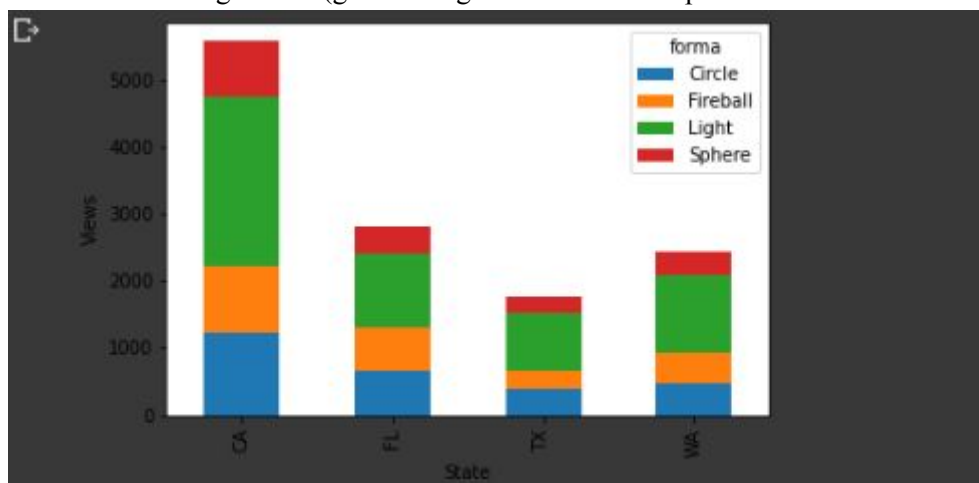
5-Baixando o gráfico por meio da função savefig

```
1 plt.savefig("grafic_bar.png")
```

6-Criando o gráfico de barra empilhado

```
5 grafic_bar_empilhado = quatro_estados_estados_unidos.groupby(["State", "forma"]).sum().unstack().plot(kind="bar", y="quantidade de
6 plt.xlabel('State')
7 plt.ylabel('Views')
8 plt.show();
9 grafic_bar_empilhado
```

Resultado do código acima(gerando o gráfico de barra empilhado



4. Considerações Finais

Podemos concluir que essa etapa foi fundamental para entendermos e trabalharmos na exploração dos dados por meio dos gráficos e por meio do mapa. Os resultados desse trabalho permite uma melhor análise, interpretação e visualização dos dados que foram apresentados e trabalhados no sprint 1. Em relação ao desenvolvimento do sprint 2 as tarefas foram divididas contudo os dois participantes não resolveram.

Referências

FIGUEIREDO, Vinicius. Seus Primeiros Passos com Data Scientist: Introdução ao Pandas. **Data Hackers**. São Paulo, 30 de maio de 2018. Disponível em:

<<https://medium.com/data-hackers/uma-introdu%C3%A7%C3%A3o-simples-ao-pandas-1e15ea37fa1>>. Acesso em: 08 de Set. de 2020

FAMETHEMES.Biblioteca seaborn com o matplotlib. **Vooo - Insights**.c2020

Disponível em:<<https://www.vooo.pro/insights/biblioteca-seaborn-com-o-matplotlib/>>. Acesso em:26 de set. de 2020

MATHEUS, Yuri. Matplotlib uma biblioteca Python para gerar gráficos interessantes. **alura. São Paulo**, 27 de Novembro de 2018. Disponível em:

<<https://www.alura.com.br/artigos/criando-graficos-no-python-com-a-matplotlib>>.Acesso em: 25 de set. de 2020.