

Nome: Daniel Alexandre Oliveira Modesto

RGM: 28595670

Curso: VISUALIZAÇÃO DA INFORMAÇÃO 20221

Link do Dataset:

<https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce>

Link do Mapa do Brasil utilizado:

https://github.com/filipegmedeiros/brazilian_choropleth_map

Link apresentação:

<https://www.loom.com/share/510c80ddd21f4f3ba45b52e59e15b187>

Plot 1

- Aqui se analisa o volume de compras feitas na plataforma Olist por período (mês) entre 2016 e 2018
- Cod fonte:

```
import pandas as pd;
import matplotlib.pyplot as plt;

series = pd.read_csv("olist_orders_dataset.csv", usecols =
["order_purchase_timestamp"], parse_dates=["order_purchase_timestamp"]);
series.groupby([(series.order_purchase_timestamp.dt.year), (series.order_purchase_timestamp.dt.month)]).count().plot(kind="bar")

plt.title("Volume de compras no site da Olist 2016-2018")
plt.ylabel("Volume de Compras")
plt.xlabel("Meses")
plt.show()
```

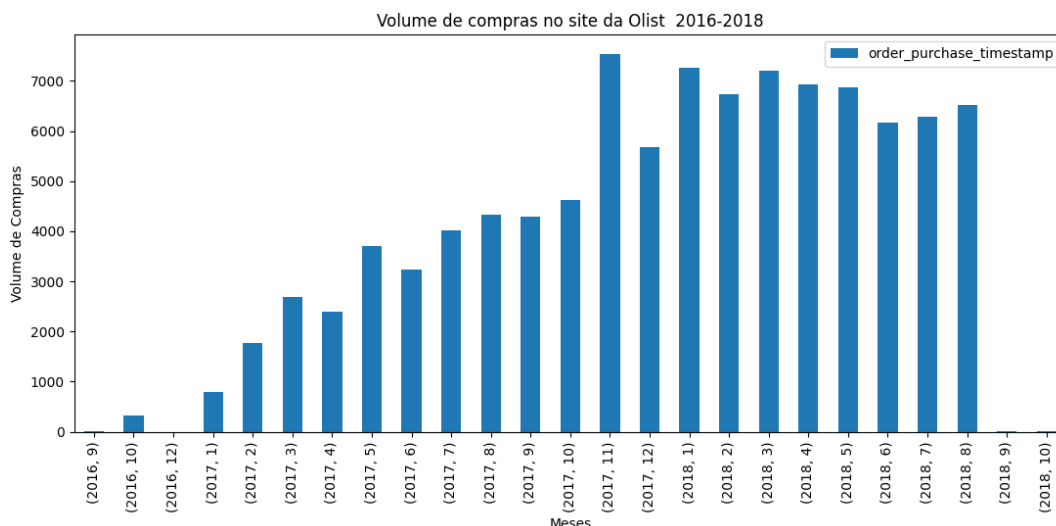


Tabela 1

Plot 2

- Análise da distribuição de vendas realizadas na plataforma Olist em relação ao preço do produto. Esta análise fornece informações da curva de tendência de preços dos itens de maior saída.
- Código fonte:

```

• import pandas as pd;
• import matplotlib.pyplot as plt;
•
• data = pd.read_csv("olist_order_items_dataset.csv", usecols=["price"])
• ndata = data.groupby(((data["price"]//5)*5).count())
• ndata.plot(style='k. ');
• plt.ylabel("volume de vendas")
• plt.xlabel("preço (R$)")
• plt.yscale("log")
• plt.xscale("log")
• plt.title("Distribuição de vendas por preço")
• plt.show();

```

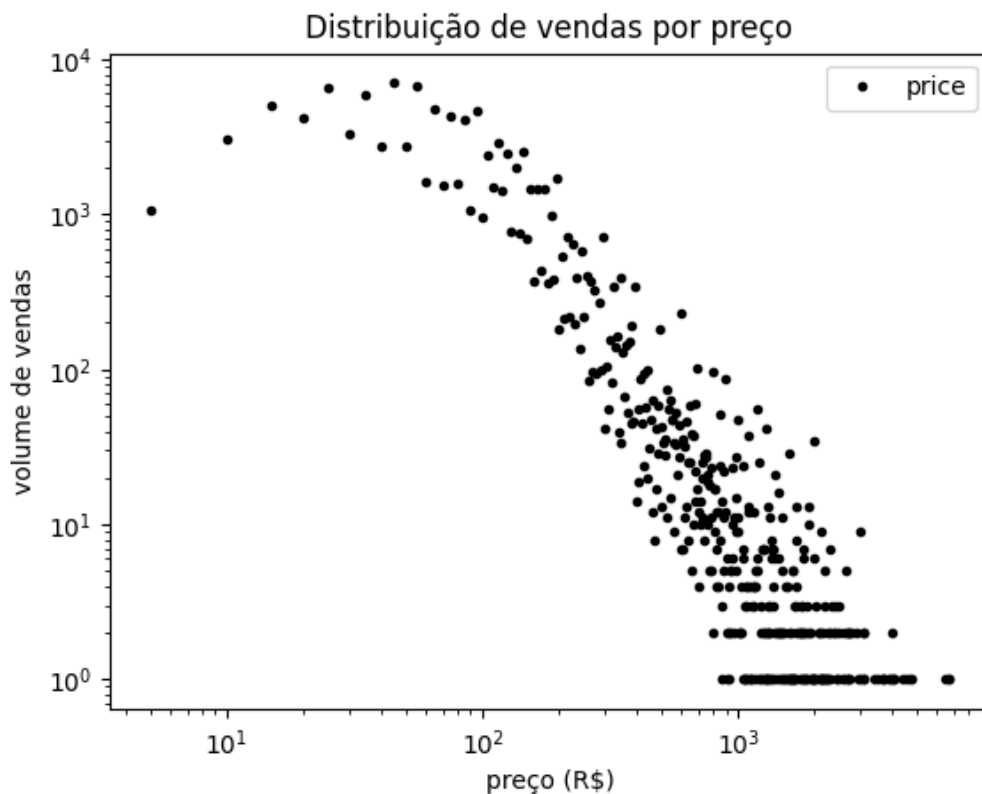


Tabela 2

Plot 3 e 4

- Análise de distribuição geográfica quantitativa do destino das vendas realizadas na plataforma Olist entre 2016 e 2018.
- Código fonte:

Mapa1:

```

import plotly.express as px
import pandas as pd
import json
data = pd.read_csv("olist_geolocation_dataset.csv", usecols=["geolocation_state"])
data.rename(columns={"geolocation_state": "estados"}, inplace = True)
ndata = data.value_counts().rename_axis("estados").reset_index(name="vendas");
geojson = json.load(open("brasil_estados.json"))
fig = px.scatter_geo(ndata, geojson=geojson, locations="estados", size = "vendas",
scope="south america")
fig.write_html('Mapa_vendas_por_estado_scatter.html', auto_open=True)

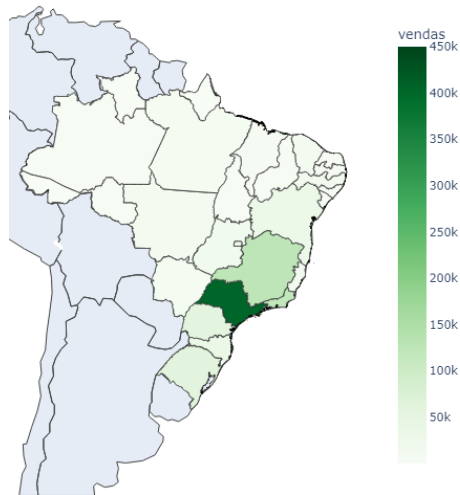
```

Mapa 2:

```
import plotly.express as px
import pandas as pd
import json
data = pd.read_csv("olist_geolocation_dataset.csv", usecols=["geolocation_state"])
data.rename(columns={"geolocation_state": "estados"}, inplace = True)
ndata = data.value_counts().rename_axis("estados").reset_index(name="vendas");
geojson = json.load(open("brasil_estados.json"))
fig = px.choropleth(ndata, geojson=geojson, locations="estados", color = "vendas",
color_continuous_scale="greens", range_color=(500,450000), scope="south america")
fig.write_html('Mapa_vendas_por_estado.html', auto_open=True)
```



mapa 1



mapa 2