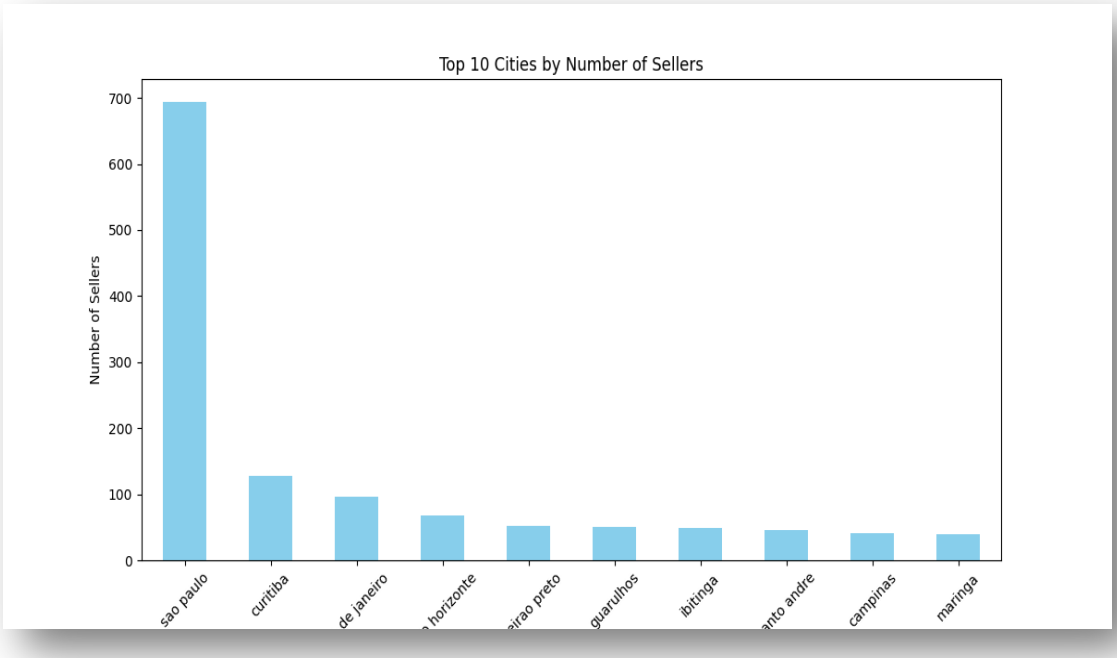
Market Basket Analysis

Purpose of the Code

The code is designed to analyze data on e-commerce sellers across Brazilian cities and states. The objective is to create visualizations that show the distribution of sellers by city and by state, helping to identify key areas with high seller concentration.

Code Description

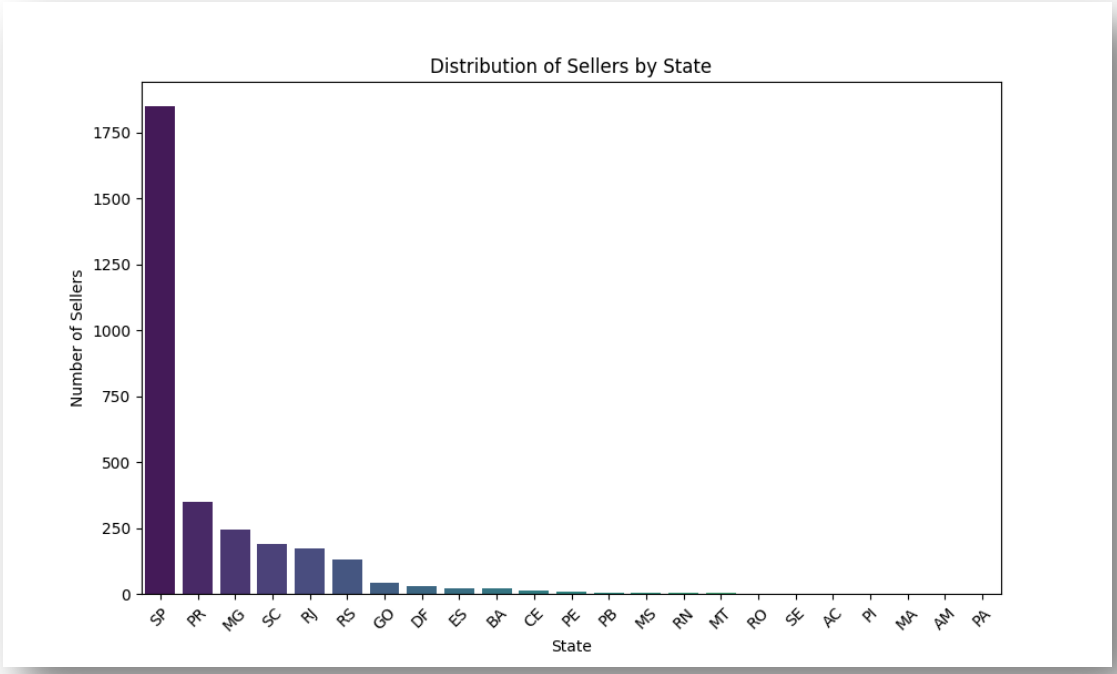
```
#Importing necessary libraries
import matplotlib.pyplot as plt
import pandas as pd


#Loading the dataset
data = pd.read_csv('brazilian_ecommerce_data.csv')


#Grouping data by city and state to calculate the number of sellers
city_data = data['city'].value_counts().head(١٠)
state data = data['state'].value counts()
3. Output Visualizations
Plot 1: Top 10 Cities by Number of Sellers
#Plotting top 10 cities
plt.figure(figsize=(10,6))
city_data.plot(kind='bar', color='skyblue')
plt.title('Top 10 Cities by Number of Sellers')
plt.xlabel('City')
plt.ylabel('Number of Sellers')
plt.xticks(rotation=45)
plt.show ()
```



The first plot visualizes the top 10 cities with the highest number of sellers. The city **São Paulo** leads by a large margin, with over 700 sellers, while other cities like **Curitiba** and **Rio de Janeiro** have considerably fewer sellers.

Plot 2: Distribution of Sellers by State

```
#Plotting seller distribution by state
plt.figure(figsize=(12,8))
state_data.plot(kind='bar', colormap='viridis')
plt.title('Distribution of Sellers by State')
plt.xlabel('State')
plt.ylabel('Number of Sellers')
plt.xticks(rotation=45)
plt.show()
```



The second plot shows the distribution of sellers by state. **São Paulo (SP)** again has the highest concentration, with over 1750 sellers, followed by **Paraná (PR)**, **Minas Gerais (MG)**, and **Santa Catarina (SC)**.

The remaining states have significantly fewer sellers.

4. Conclusions

The analysis reveals a strong concentration of e-commerce sellers in São Paulo, both by city and state, indicating its economic importance within Brazil's e-commerce landscape.