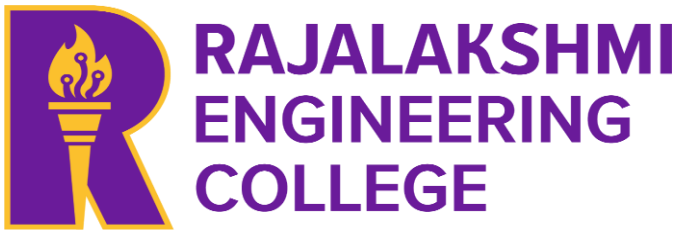
RAJALAKSHMI ENGINEERING COLLEGE
(Autonomous)

RAJALAKSHMI NAGAR, THANDALAM, CHENNAI-602105



AD23632 - FRAMEWORK FOR DATA AND VISUAL ANALYTICS

LABORATORY RECORD NOTEBOOK

Name : T.R SKANDHA

Year / Branch / Section : III- AIML

Register No. : 2116231501156

Semester : V

Academic Year : 2025-2026

INDEX

PYTHON IMPLEMENTATION :-

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

jsmart = pd.read_csv("jsmart.csv")

print("Dataset preview:")
print(jsmart.head())
print("\nAvailable columns:", list(jsmart.columns))

plt.figure(figsize=(10,6))
jsmart['Country'].value_counts().head(10).plot(kind='bar', color='skyblue',
edgecolor='black')
plt.title("Top 10 Countries by Orders")
plt.xlabel("Country")
plt.ylabel("Number of Orders")
plt.xticks(rotation=45)
plt.show()

plt.figure(figsize=(6,6))
jsmart['Segment'].value_counts().plot(
    kind='pie',
    autopct='%1.1f%%',
    startangle=90,
    cmap='Set3'
)
plt.title("Distribution of Market Segments")
plt.ylabel("")  # remove extra label
plt.show()

plt.figure(figsize=(8,5))
plt.hist(jsmart['lon'], bins=20, alpha=0.7, label="Longitude", color="orange",
edgecolor="black")
plt.hist(jsmart['lat'], bins=20, alpha=0.7, label="Latitude", color="green",
```

```
edgecolor="black")
plt.title("Geographic Distribution (Lon/Lat)")
plt.xlabel("Values")
plt.ylabel("Frequency")
plt.legend()
plt.show()

plt.figure(figsize=(8,5))
sns.countplot(x="Ship Mode", data=jsmart, palette="Set2", edgecolor="black")
plt.title("Orders by Shipping Mode")
plt.show()

plt.figure(figsize=(10,6))
sns.countplot(x="Region", hue="Segment", data=jsmart, palette="Set3",
edgecolor="black")
plt.title("Orders by Region and Segment")
plt.xticks(rotation=45)
plt.show()

plt.figure(figsize=(8,5))
sns.boxplot(x="Ship Mode", y="lon", data=jsmart, palette="Set2")
plt.title("Longitude Distribution by Shipping Mode")
plt.show()

plt.figure(figsize=(8,6))
sns.heatmap(jsmart.corr(numeric_only=True), annot=True, cmap="coolwarm",
linewidths=0.5)
plt.title("Correlation Heatmap")
plt.show()
```

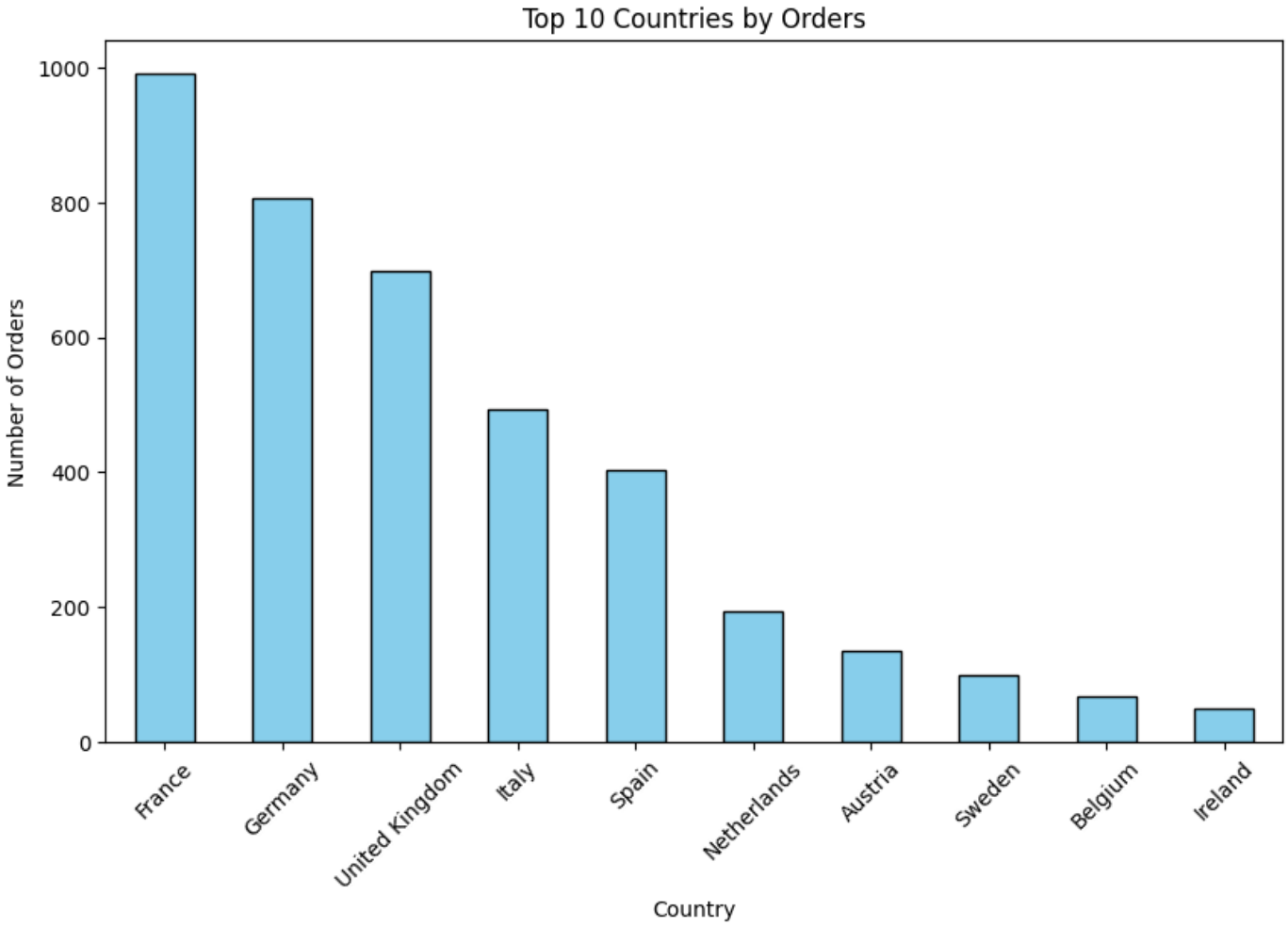
OUTPUT :-

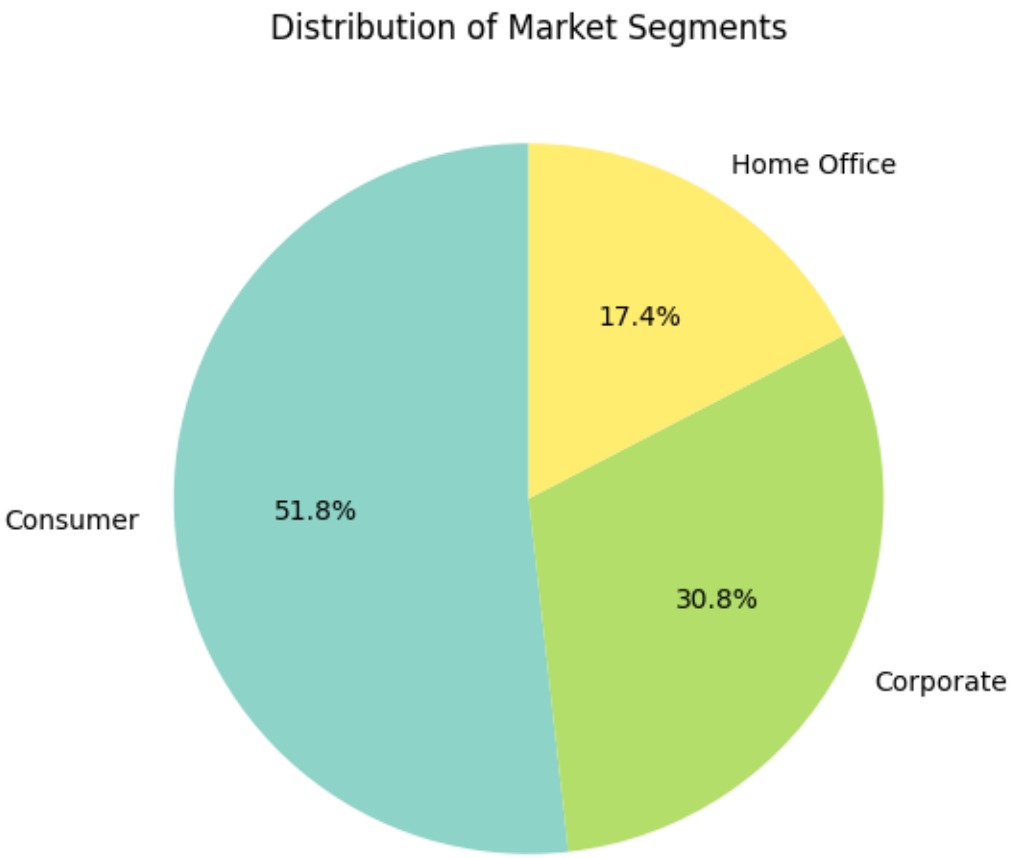
Dataset preview:

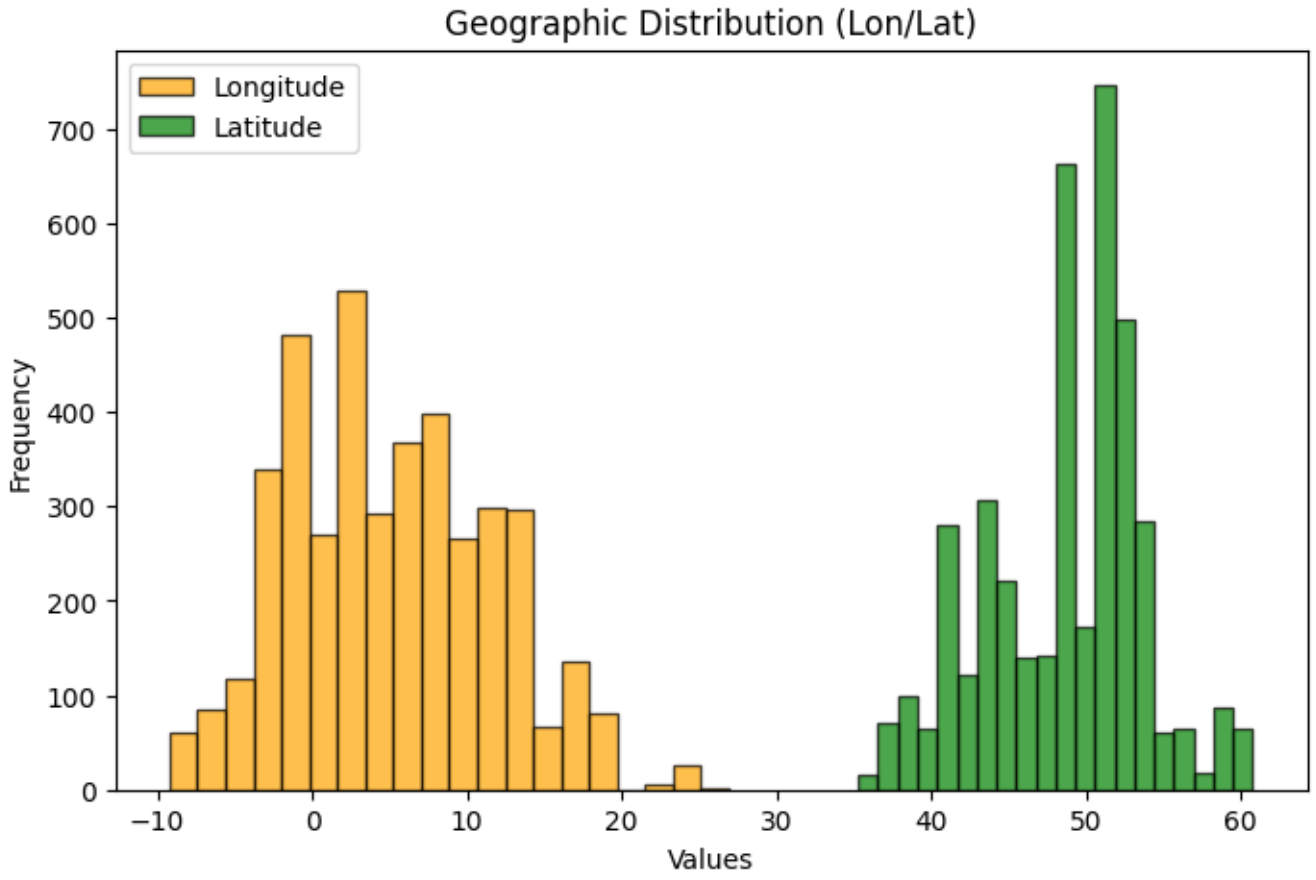
```
        Order ID Order Date     Customer Name        City        Country  \
0  BN-2011-7407039   1/1/2017       Ruby Patel   Stockholm         Sweden
1  AZ-2011-9050313   1/3/2017    Summer Hayward   Southport  United Kingdom
2  AZ-2011-6674300   1/4/2017  Devin Huddleston     Valence         France
3  BN-2011-2819714   1/4/2017       Mary Parker  Birmingham  United Kingdom
4   AZ-2011-617423   1/5/2017      Daniel Burke  Echirolles         France

                 State   Region      Segment Ship Date     Ship Mode  \
0            Stockholm    North  Home Office  1/5/2017  Economy Plus
1              England    North     Consumer  1/7/2017       Economy
2  Auvergne-Rhône-Alpes  Central     Consumer  1/8/2017       Economy
3              England    North    Corporate  1/9/2017       Economy
4  Auvergne-Rhône-Alpes  Central  Home Office  1/7/2017      Priority

         lon        lat
0  18.068581  59.329324
1  -3.010113  53.645708
2   4.892360  44.933393
3  -1.890401  52.486243
4   5.718034  45.142151
```

Available columns: ['Order ID', 'Order Date', 'Customer Name', 'City', 'Country', 'State', 'Region', 'Segment', 'Ship Date', 'Ship Mode', 'lon', 'lat']

``


Top 10 Countries by Orders
Number of Orders
0
200
400
600
800
1000
France
Germany
United Kingdom
Italy
Spain
Netherlands
Austria
Sweden
Belgium
Ireland
Country


Distribution of Market Segments
Home Office
17.4%
Consumer
51.8%
30.8%
Corporate

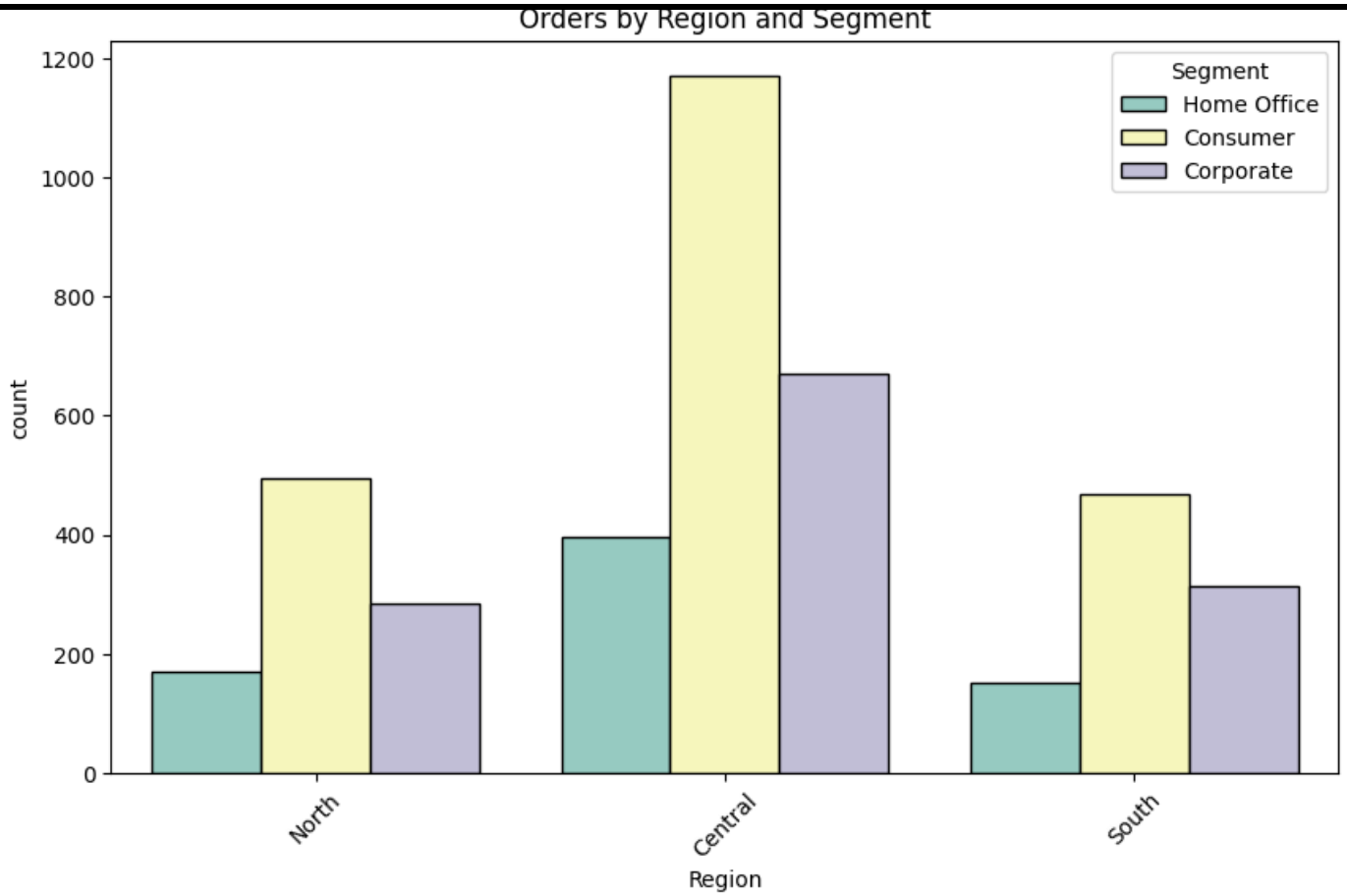


/tmp/ipython-input-3560960110.py:57: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.countplot(x="Ship Mode", data=jsmart, palette="Set2", edgecolor="black")
```

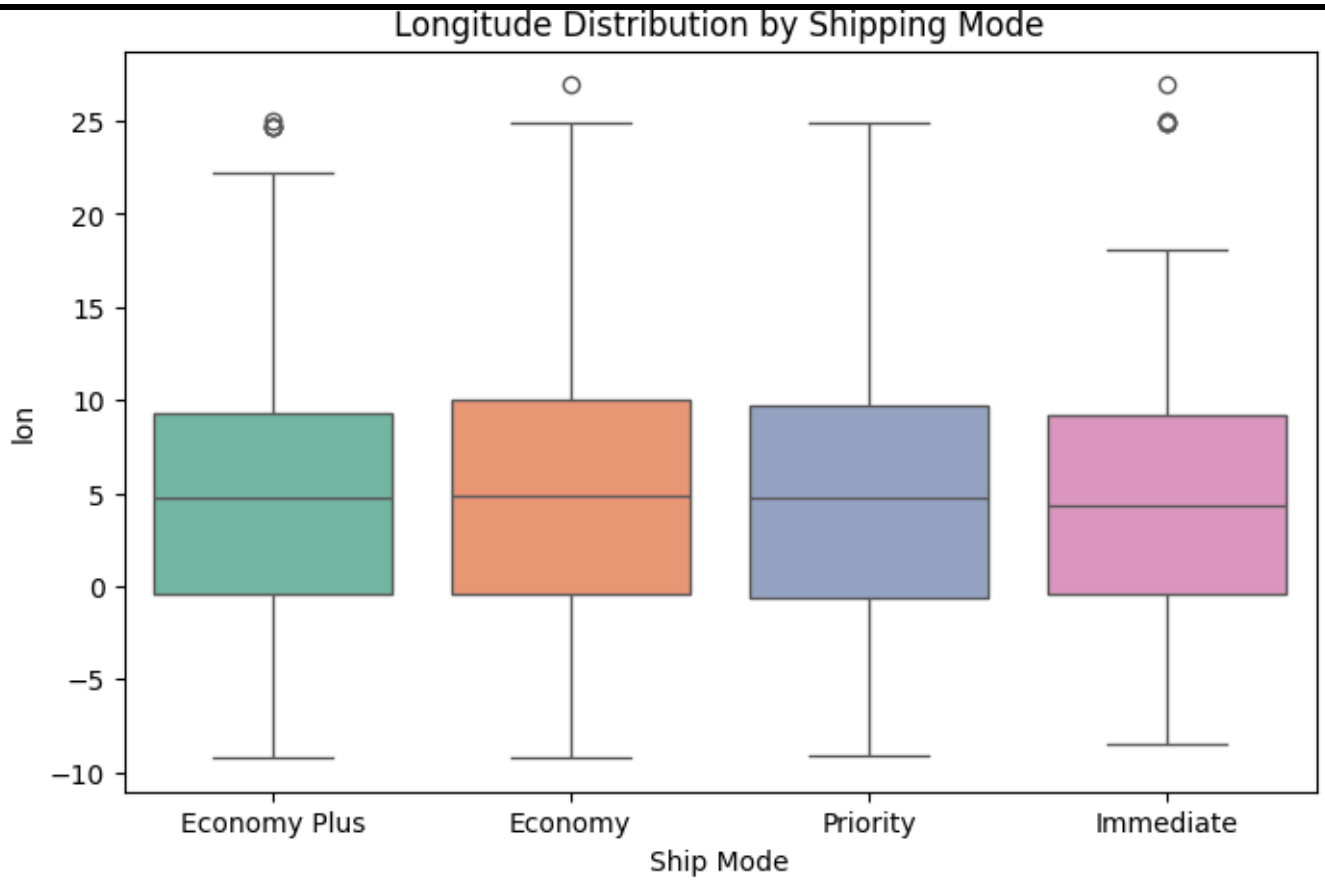

Orders by Shipping Mode
2500
2000
1500
1000
500
0
count
Economy Plus
Economy
Priority
Immediate
Ship Mode

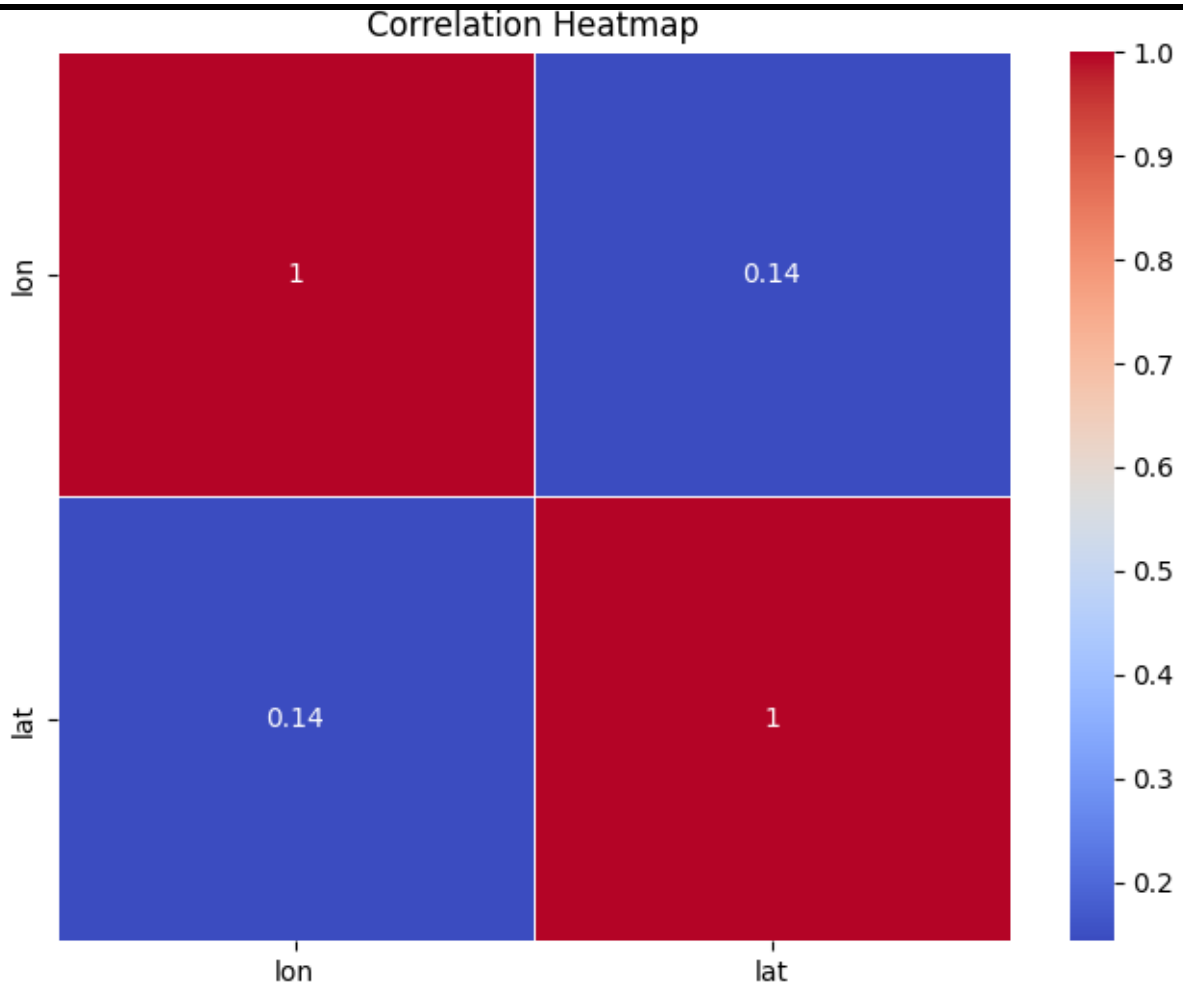


/tmp/ipython-input-3560960110.py:74: FutureWarning:

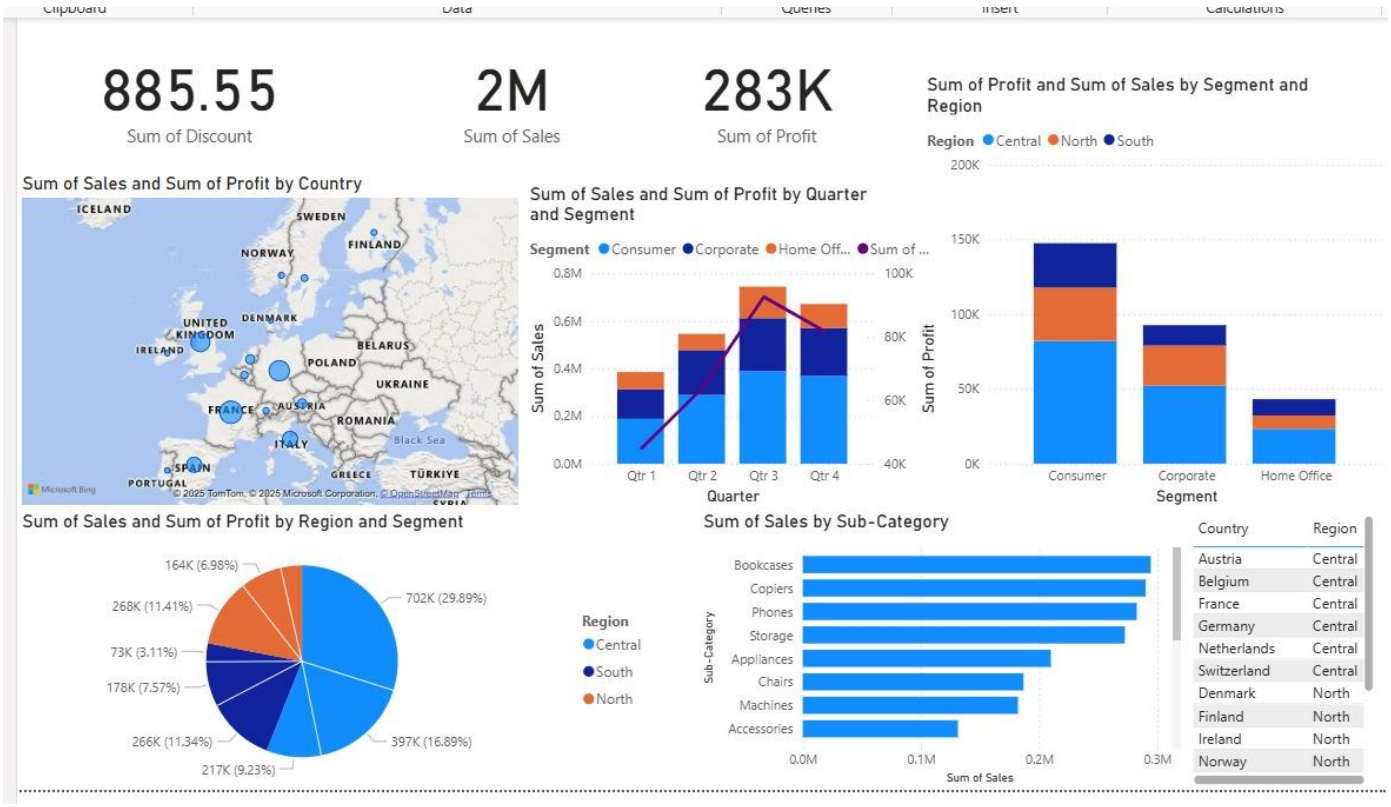
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

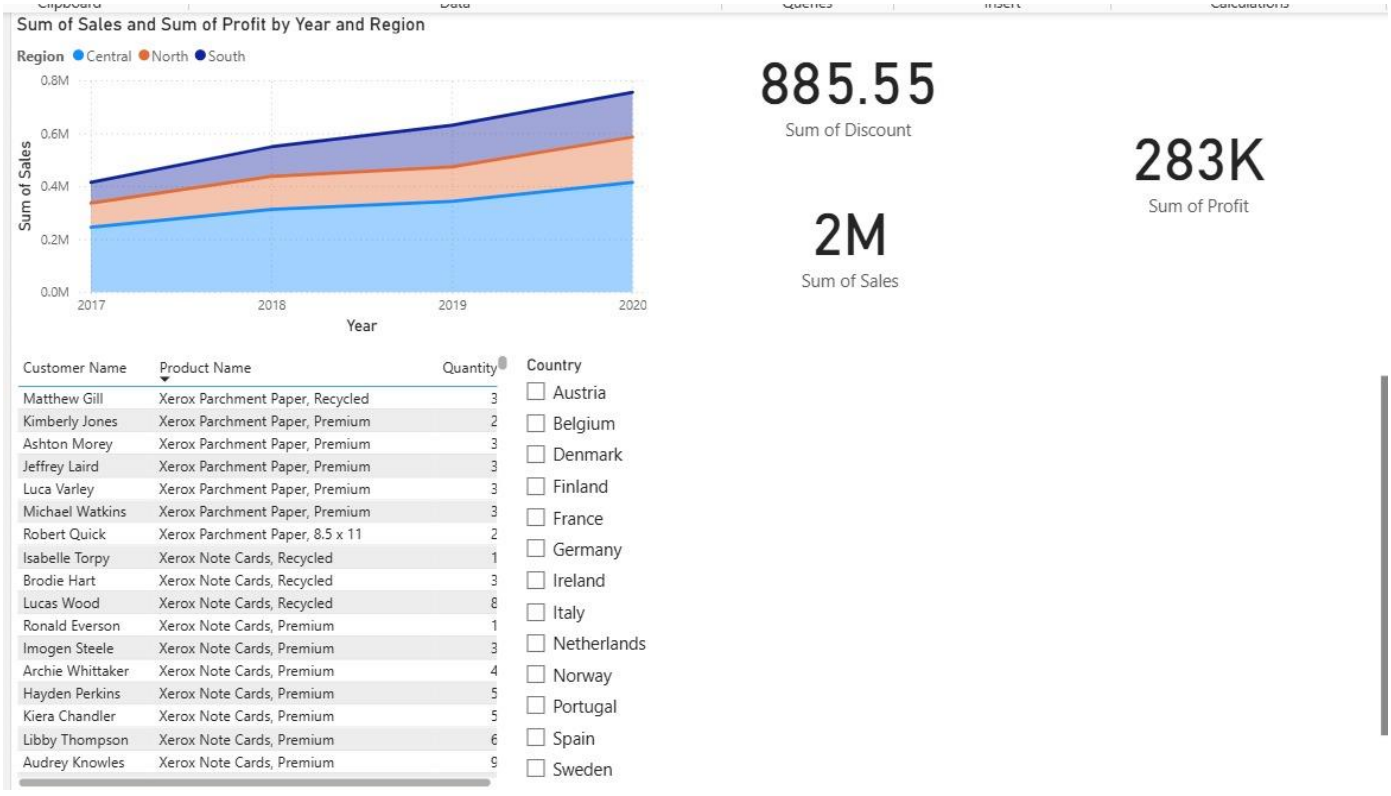
```
sns.boxplot(x="Ship Mode", y="lon", data=jsmart, palette="Set2")
```


Longitude Distribution by Shipping Mode
lon
25
20
15
10
5
0
−5
−10
Economy Plus
Economy
Priority
Immediate
Ship Mode


Correlation Heatmap
lon
lat
1
0.14
0.14
1
lon
lat
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

POWER BI :-


885.55
Sum of Discount
2M
Sum of Sales
283K
Sum of Profit
Sum of Profit and Sum of Sales by Segment and Region
Region Central North South
Sum of Sales and Sum of Profit by Country
ICELAND
SWEDEN
FINLAND
NORWAY
UNITED KINGDOM
DENMARK
IRELAND
BELARUS
POLAND
UKRAINE
FRANCE
AUSTRIA
ROMANIA
ITALY
Black Sea
SPAIN
PORTUGAL
GREECE
TÜRKIYE
Sum of Sales and Sum of Profit by Quarter and Segment
Segment Consumer Corporate Home Off... Sum of ...
Sum of Sales
Sum of Profit
Qtr 1
Qtr 2
Qtr 3
Qtr 4
Quarter
Consumer
Corporate
Home Office
Segment
Sum of Sales and Sum of Profit by Region and Segment
164K (6.98%)
268K (11.41%)
73K (3.11%)
178K (7.57%)
266K (11.34%)
217K (9.23%)
397K (16.89%)
702K (29.89%)
Region
Central
South
North
Sum of Sales by Sub-Category
Bookcases
Copiers
Phones
Storage
Appliances
Chairs
Machines
Accessories
Sub-Category
Sum of Sales
Country
Region
Austria Central
Belgium Central
France Central
Germany Central
Netherlands Central
Switzerland Central
Denmark North
Finland North
Ireland North
Norway North


Sum of Sales and Sum of Profit by Year and Region
Region Central North South
Sum of Sales
Year
885.55
Sum of Discount
283K
Sum of Profit
2M
Sum of Sales
Customer Name
Product Name
Quantity
Matthew Gill Xerox Parchment Paper, Recycled 3
Kimberly Jones Xerox Parchment Paper, Premium 2
Ashton Morey Xerox Parchment Paper, Premium 3
Jeffrey Laird Xerox Parchment Paper, Premium 3
Luca Varley Xerox Parchment Paper, Premium 3
Michael Watkins Xerox Parchment Paper, Premium 3
Robert Quick Xerox Parchment Paper, 8.5 x 11 2
Isabelle Torpy Xerox Note Cards, Recycled 1
Brodie Hart Xerox Note Cards, Recycled 3
Lucas Wood Xerox Note Cards, Recycled 8
Ronald Everson Xerox Note Cards, Premium 1
Imogen Steele Xerox Note Cards, Premium 3
Archie Whittaker Xerox Note Cards, Premium 4
Hayden Perkins Xerox Note Cards, Premium 5
Kiera Chandler Xerox Note Cards, Premium 5
Libby Thompson Xerox Note Cards, Premium 6
Audrey Knowles Xerox Note Cards, Premium 9
Country
Austria
Belgium
Denmark
Finland
France
Germany
Ireland
Italy
Netherlands
Norway
Portugal
Spain
Sweden

EXPLAINTATION :-

1. Sum of Discount, Sales, and Profit

These key metrics show the overall performance indicators of the business.

- **Sum of Discount (885.55)** indicates total discounts given to customers.
- **Sum of Sales (2M)** represents the total revenue generated.
- **Sum of Profit (283K)** reflects the net earnings after expenses.

2. Sales and Profit by Country (Map Visual)

This map visual displays geographic sales and profit distribution across European countries.

- Larger circles denote higher sales volume.
- It helps identify top-performing markets like France and Germany.
- Useful for regional strategy and expansion planning.

3. Sales and Profit by Quarter and Segment

This clustered bar and line chart compares quarterly sales with profit trends.

- Segments include **Consumer**, **Corporate**, and **Home Office**.
- Q3 shows the highest sales and profit peak.
- It highlights seasonality and segment contribution.

4. Profit and Sales by Segment and Region

This stacked bar chart shows how each region contributes to total profit and sales.

- The **Central** region dominates overall performance.
- The **North** and **South** regions follow with smaller shares.
- Segment analysis aids in targeted marketing and budgeting.

5. Sales and Profit by Region and Segment (Pie Chart)

This pie chart illustrates proportional sales by region and segment.

- The **Central Region (29.89%)** leads in sales share.
- Each segment's contribution helps in balancing regional operations.

- Ideal for identifying underperforming areas.

6. Sales by Sub-Category

This horizontal bar chart breaks down sales by product sub-categories.

- **Bookcases**, **Copiers**, and **Phones** are top-selling items.

- It shows which product lines drive revenue growth.
- Useful for inventory and supply chain planning.

7. Sales and Profit by Year and Region

This area chart shows year-wise sales and profit trends across regions from 2017–2020.

- Sales increased steadily over time across all regions.
- The **Central Region** maintained the highest growth.
- Indicates consistent business expansion.

8. Customer and Product Details Table

Displays individual customer purchase data including product name and quantity.

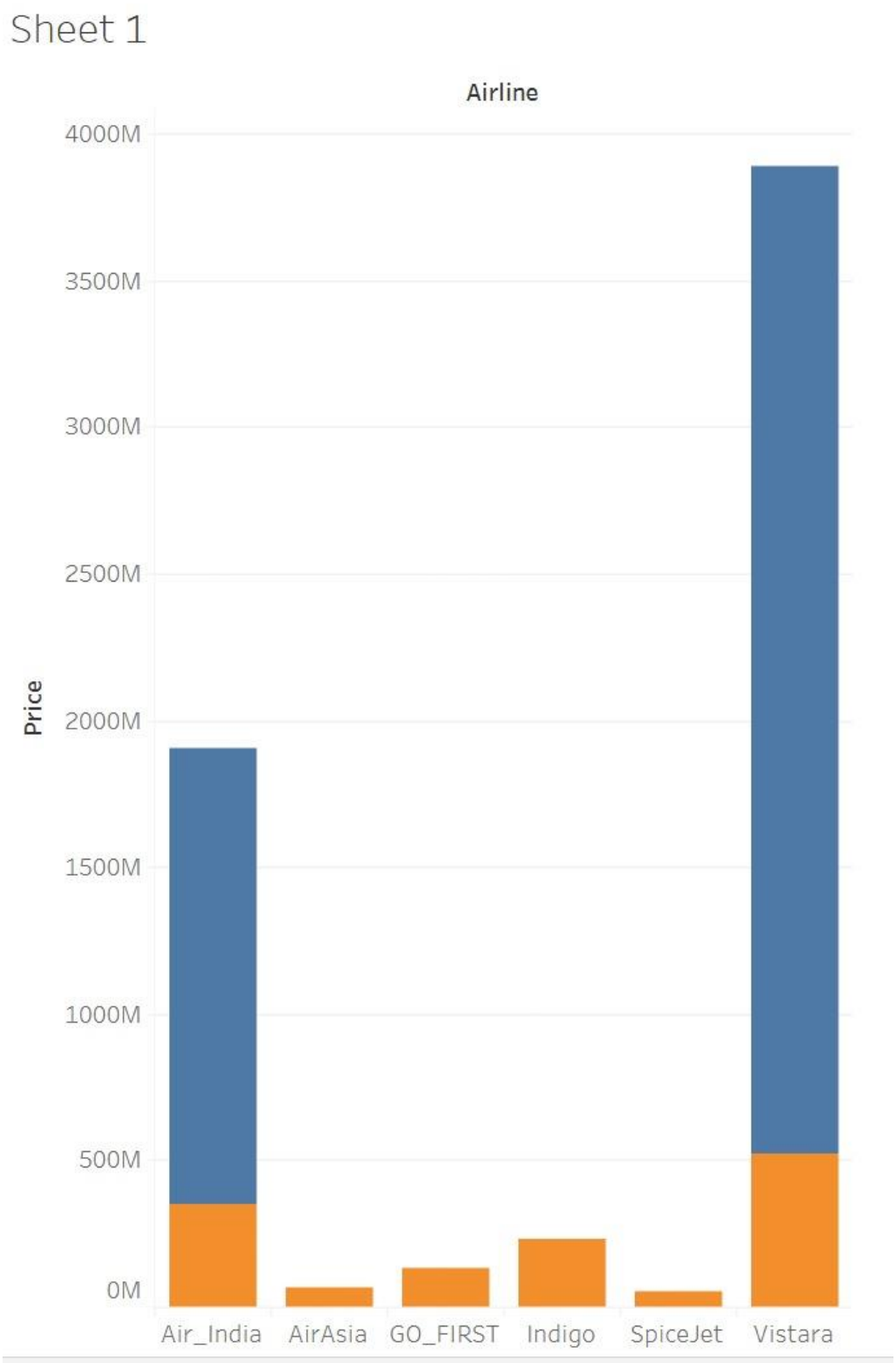
- Helps identify top customers and frequently purchased items.
- Useful for loyalty programs and customer segmentation.
- Enhances personalized sales strategies.

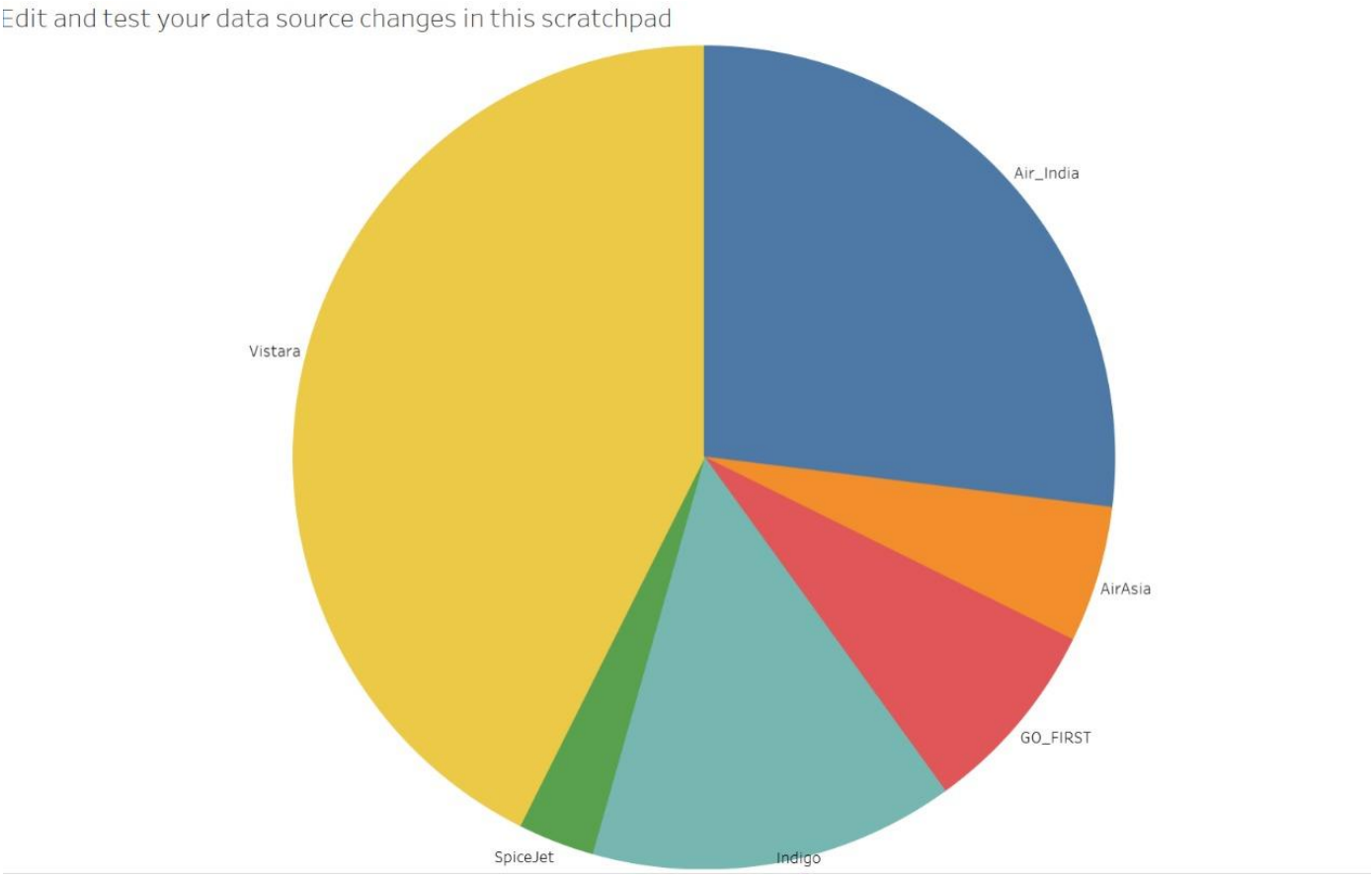
9. Country Filter Panel

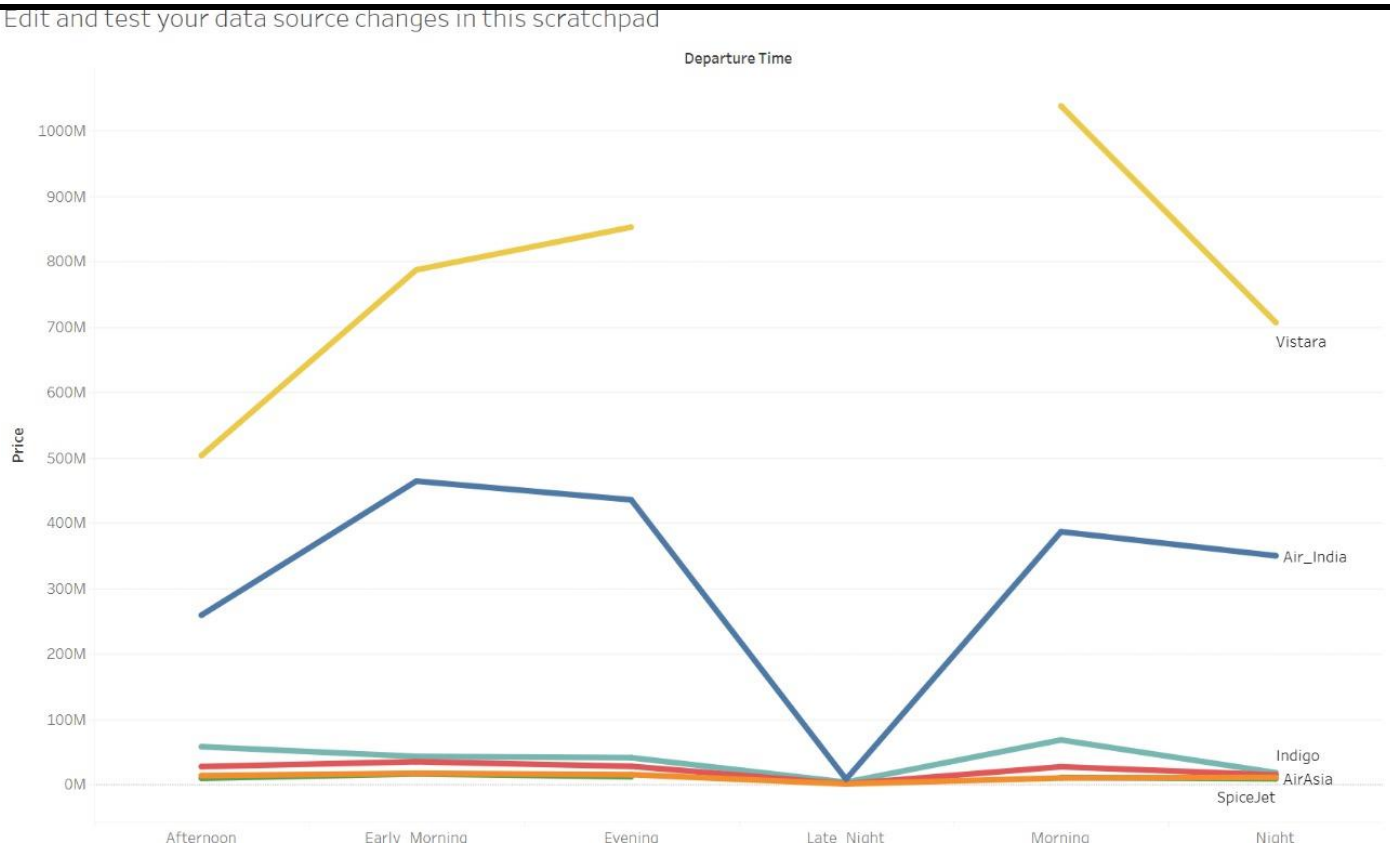
Allows filtering dashboard data by selected European countries.

- Dynamic filtering updates all visuals instantly.
- Helps compare performance between countries.
- Enables focused regional analysis.

TABLEAU :-


Sheet 1
Airline
Price
4000M
3500M
3000M
2500M
2000M
1500M
1000M
500M
0M
Air_India
AirAsia
GO_FIRST
Indigo
SpiceJet
Vistara


Edit and test your data source changes in this scratchpad
Air_India
AirAsia
GO_FIRST
Indigo
SpiceJet
Vistara



EXPLAINTATION :-

1. Airline vs Price (Bar Chart)

This bar chart compares the total flight prices across different airlines.

- **Vistara** and **Air India** show the highest total prices among all airlines.
- Low-cost carriers like **AirAsia** and **SpiceJet** have much smaller price totals.
- The chart helps identify premium vs budget airline pricing trends.

2. Airline Price Distribution (Pie Chart)

The pie chart illustrates each airline's share of the total flight price.

- **Vistara** holds the largest portion, indicating its dominance in pricing or bookings.
- **Air India** follows as the second major contributor.
- Airlines like **GO_FIRST**, **Indigo**, and **SpiceJet** occupy smaller shares, reflecting budget options.

3️⃣ Airline Price by Departure Time (Line Chart)
This line chart shows how flight prices vary by **departure time** for each airline.

- **Vistara** maintains consistently high prices across all time slots.
- **Air India** shows fluctuating prices, peaking in early morning and morning slots.
- Budget airlines remain relatively stable with lower prices across all times.

SUMMARY :-

The data visualization project effectively demonstrates how analytical dashboards can reveal key business insights using Power BI and Tableau.
In **Power BI**, sales and profit data were analyzed across regions, quarters, and product segments, highlighting trends like Central region dominance and Q3 profit peaks. The interactive visuals—such as maps, bar charts, and pie charts—helped identify top-performing countries, profitable customer segments, and product categories driving growth.

In **Tableau**, airline data was examined to understand pricing variations among major Indian airlines. Vistara and Air India emerged as premium carriers with the highest prices, while budget airlines like SpiceJet and AirAsia showed more economical trends. The departure-time analysis further revealed that flight prices fluctuate based on schedule and airline strategy.

Overall, both tools provided dynamic insights into performance, helping in **data-driven decision-making, comparative analysis, and strategic planning** through intuitive visual storytelling.