TASK - 2

-BY
D shrisharang

1) Upload Superstore.csv document
2) I used jupyther and tableau both

1)
Jupyther

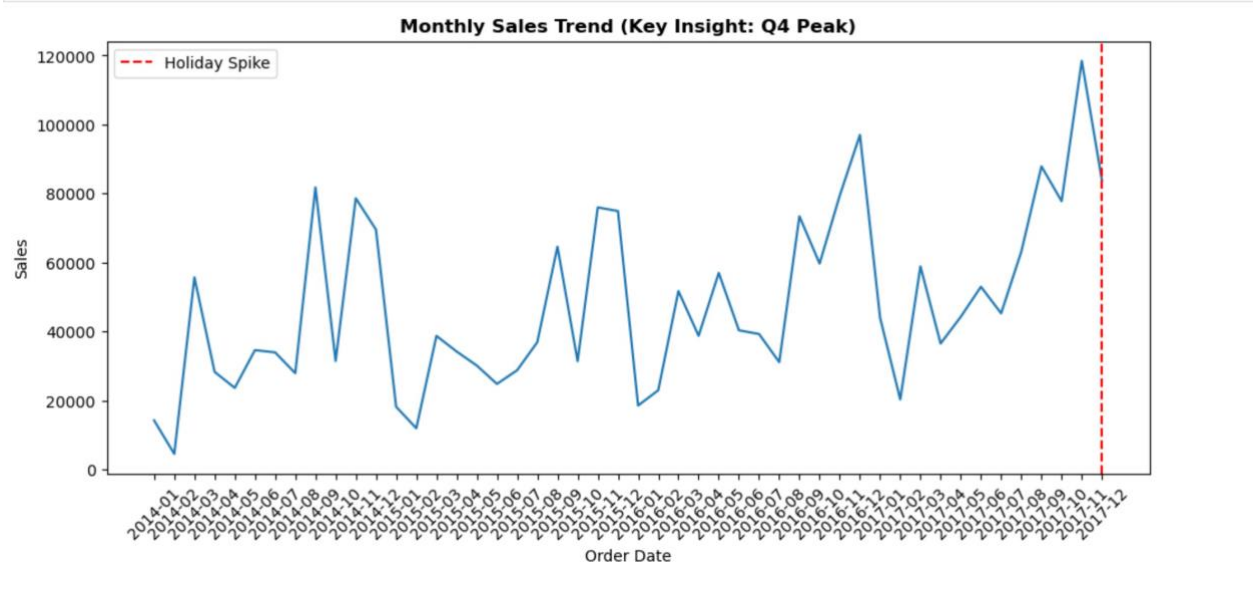
Code:-

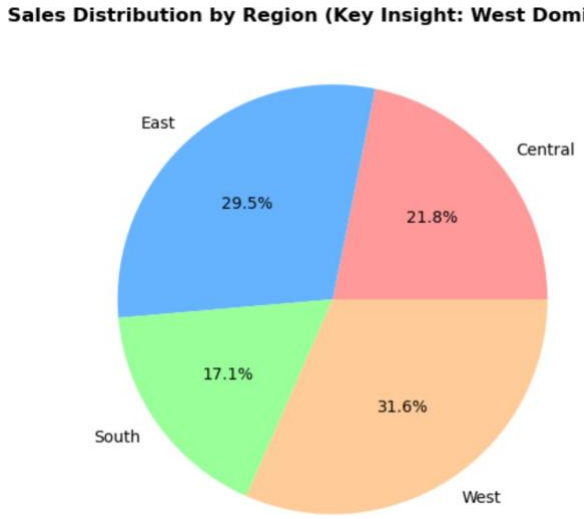
```
import pandas as pd

encodings = ['latin1', 'ISO-8859-1', 'cp1252', 'utf-8']

for encoding in encodings:
    try:
        df = pd.read_csv("Desktop/Superstore.csv",
encoding=encoding)
        print(f"Success with encoding: {encoding}")
        break
    except UnicodeDecodeError:
        continue

df['Order Date'] = pd.to_datetime(df['Order Date'])
```

Month:-

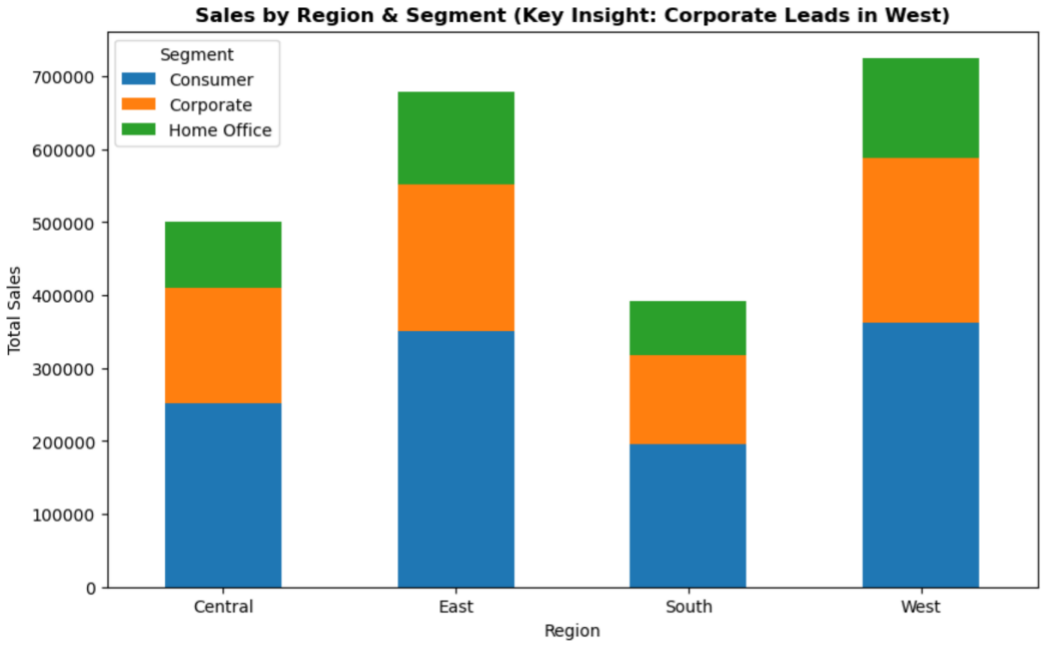
```
# Extract month-year and aggregate sales
monthly_sales = df.groupby(df['Order Date'].dt.to_period('M'))
['Sales'].sum().reset_index()
monthly_sales['Order Date'] = monthly_sales['Order
Date'].astype(str)

plt.figure(figsize=(12, 5))
sns.lineplot(data=monthly_sales, x='Order Date', y='Sales',
color='#1f77b4')
plt.title("Monthly Sales Trend (Key Insight: Q4 Peak)",
fontweight='bold')
plt.xticks(rotation=45)
```

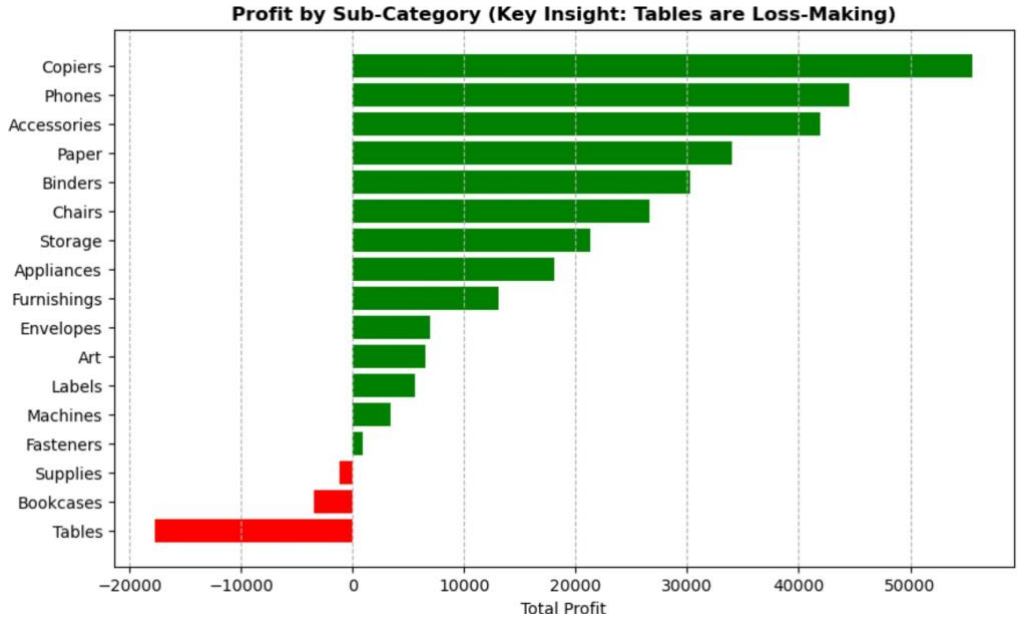
```
plt.axvline(x='2017-12', color='red', linestyle='--',
label='Holiday Spike')  # Highlight Dec
plt.legend()
plt.show()
profit_by_subcat = df.groupby('Sub-Category')
['Profit'].sum().sort_values()

plt.figure(figsize=(10, 6))
bars = plt.barh(profit_by_subcat.index, profit_by_subcat.values,
                color=['red' if x < 0 else 'green' for x in
profit_by_subcat.values])
plt.title("Profit by Sub-Category (Key Insight: Tables are Loss-
Making)", fontweight='bold')
plt.xlabel("Total Profit")
plt.grid(axis='x', linestyle='--')
plt.show()


region_sales = df.groupby('Region')['Sales'].sum()

plt.figure(figsize=(6, 6))
plt.pie(region_sales, labels=region_sales.index, autopct='%1.1f%
%',
        colors=['#ff9999','#66b3ff','#99ff99','#ffcc99'])
plt.title("Sales Distribution by Region (Key Insight: West
Dominates)", fontweight='bold')
plt.show()


segment_region = pd.crosstab(df['Region'], df['Segment'],
values=df['Sales'], aggfunc='sum')
segment_region.plot(kind='bar', stacked=True, figsize=(10, 6),
                    color=['#1f77b4', '#ff7f0e', '#2ca02c'])
plt.title("Sales by Region & Segment (Key Insight: Corporate Leads
in West)", fontweight='bold')
plt.ylabel("Total Sales")
plt.xticks(rotation=0)
```



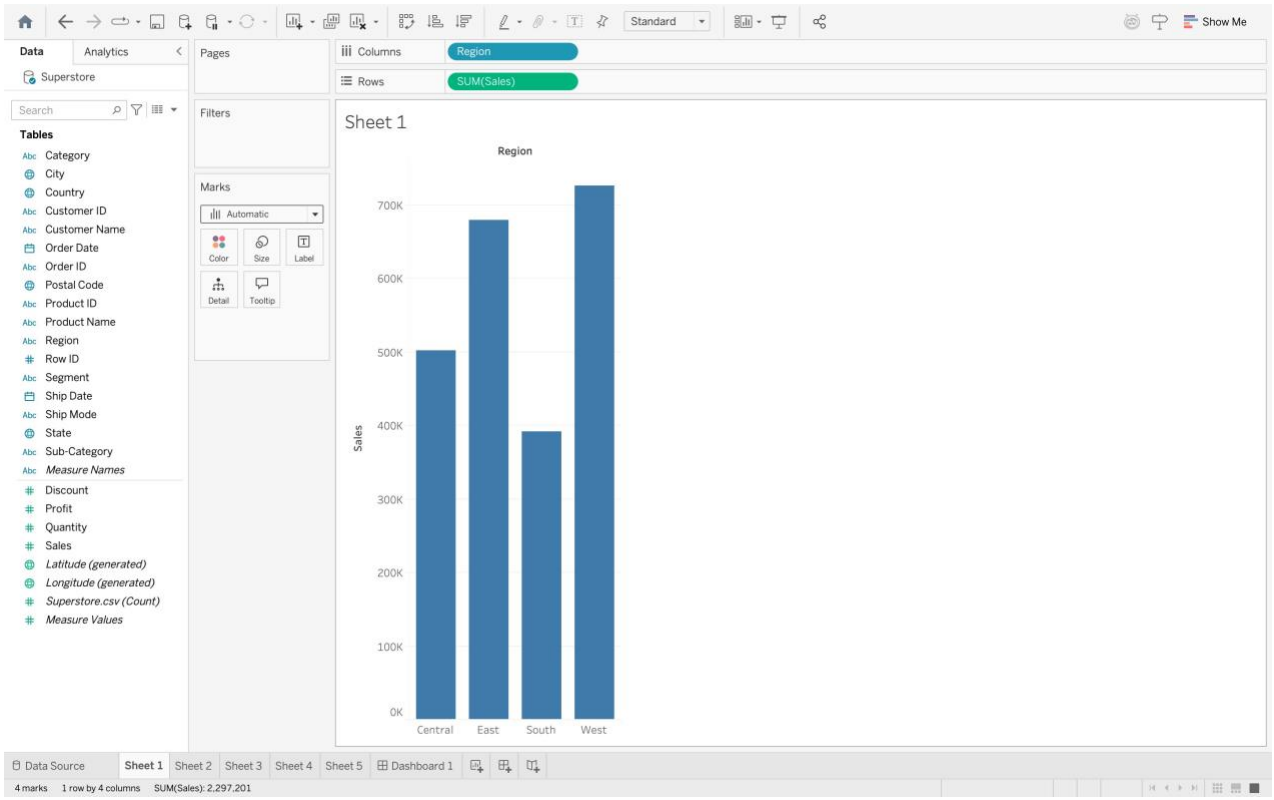

Sales Distribution by Region (Key Insight: West Dominates)
East
Central
29.5%
21.8%
17.1%
31.6%
South
West

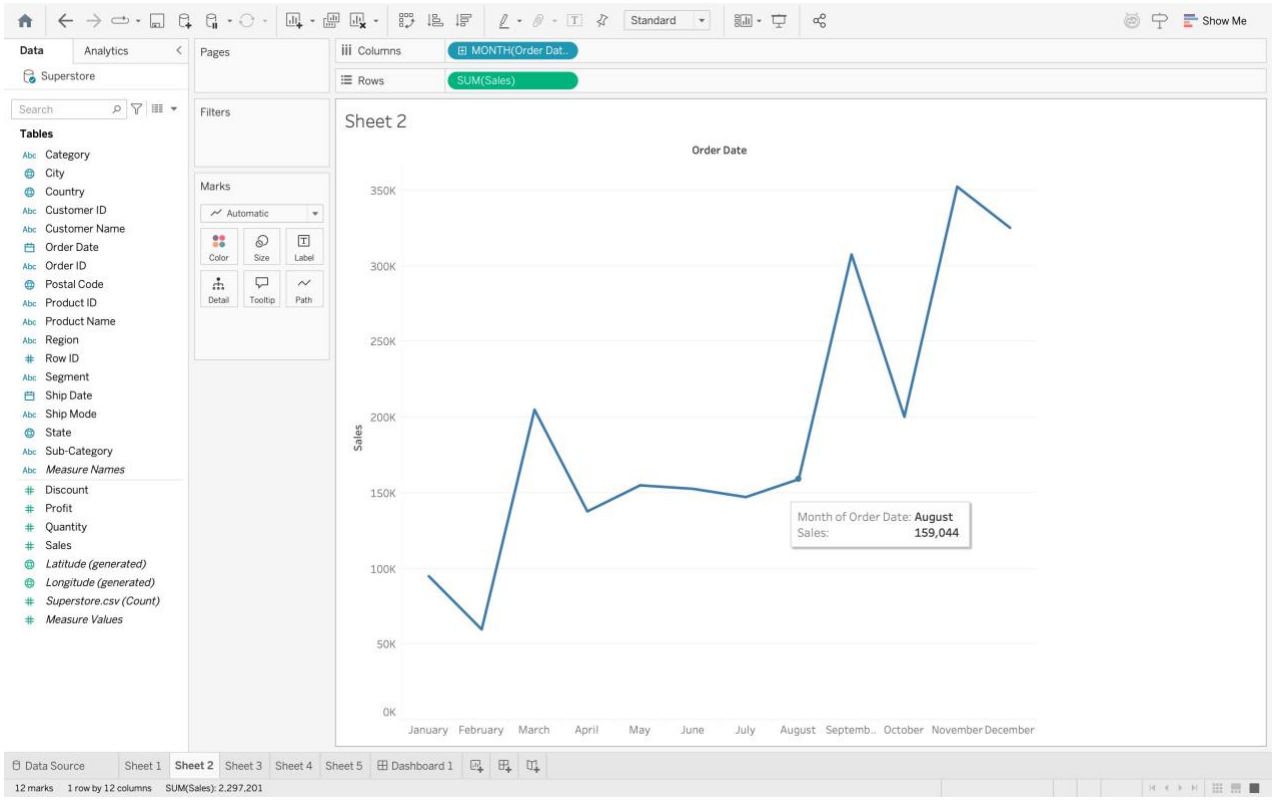
```
plt.show()
```

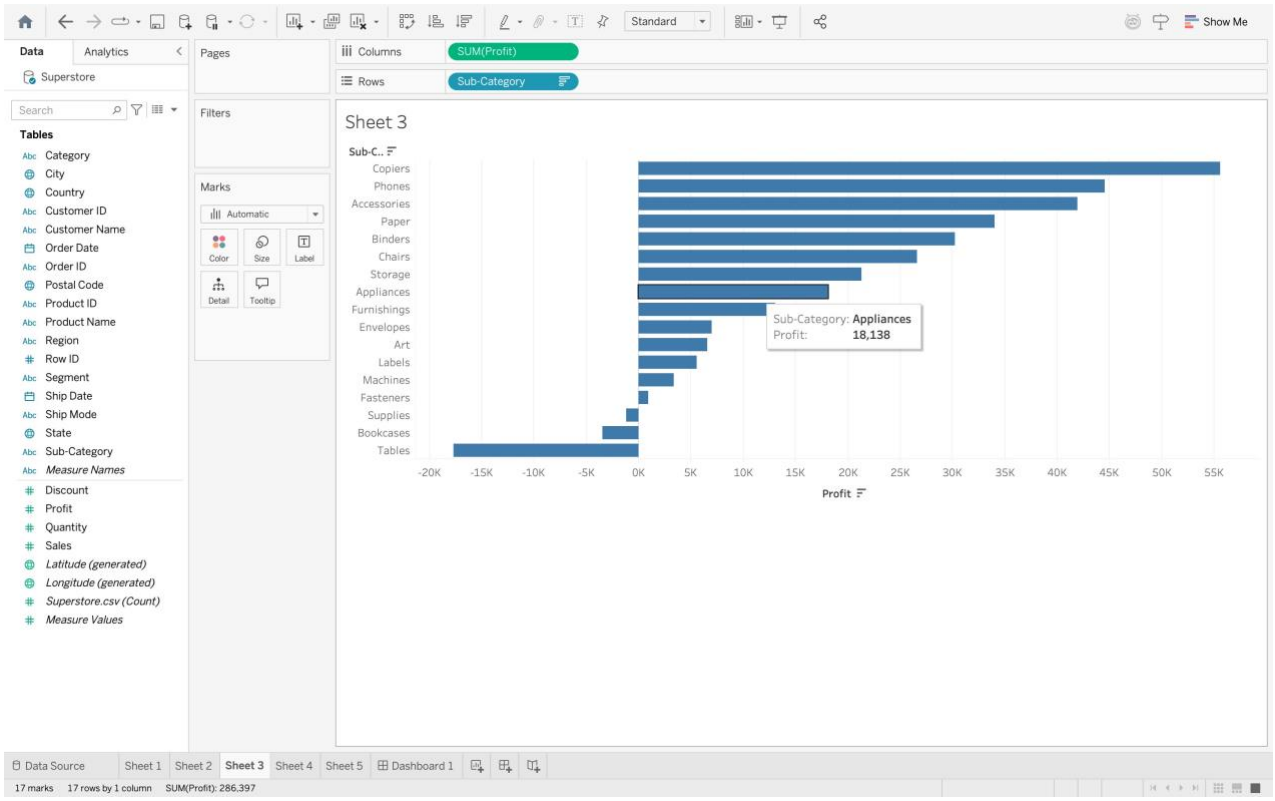

Sales by Region & Segment (Key Insight: Corporate Leads in West)
Segment
Consumer
Corporate
Home Office
Total Sales
700000
600000
500000
400000
300000
200000
100000
0
Central
East
South
West
Region

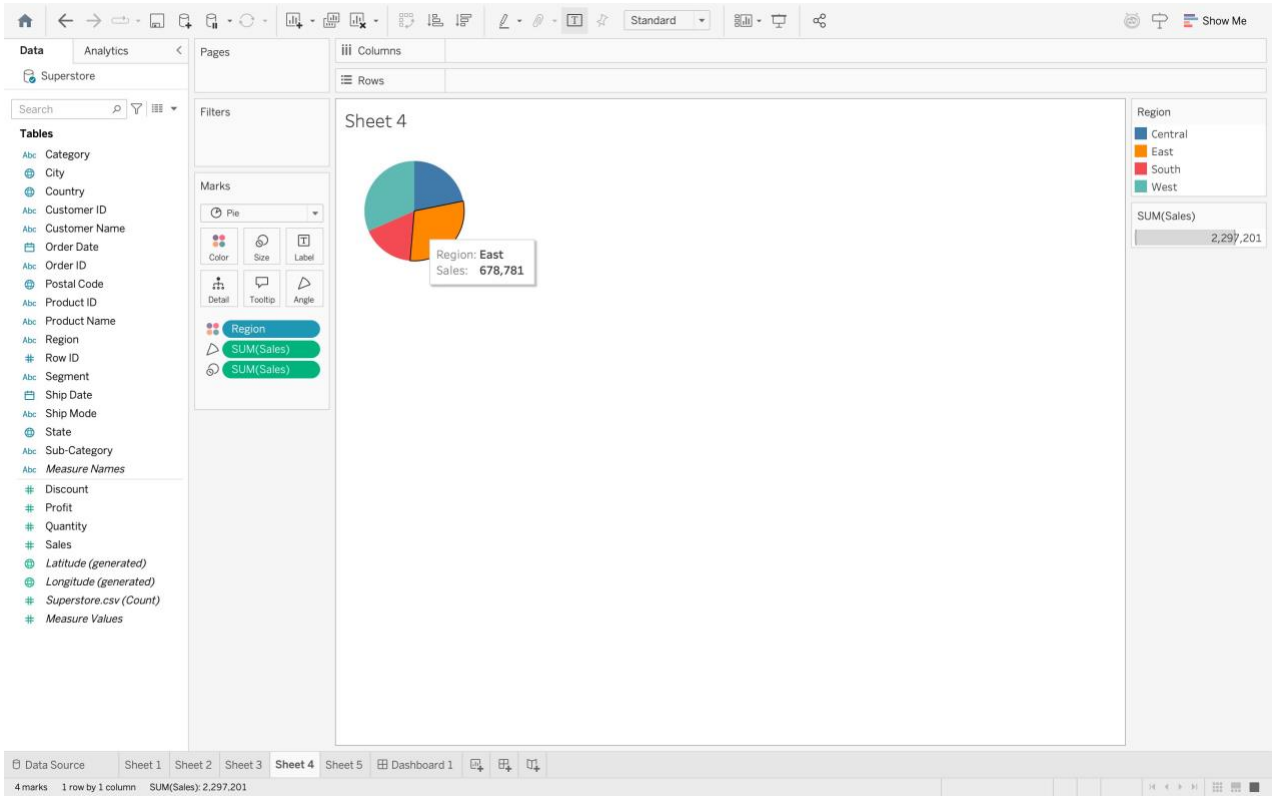

Profit by Sub-Category (Key Insight: Tables are Loss-Making)
Copiers
Phones
Accessories
Paper
Binders
Chairs
Storage
Appliances
Furnishings
Envelopes
Art
Labels
Machines
Fasteners
Supplies
Bookcases
Tables
−20000
−10000
0
10000
20000
30000
40000
50000
Total Profit

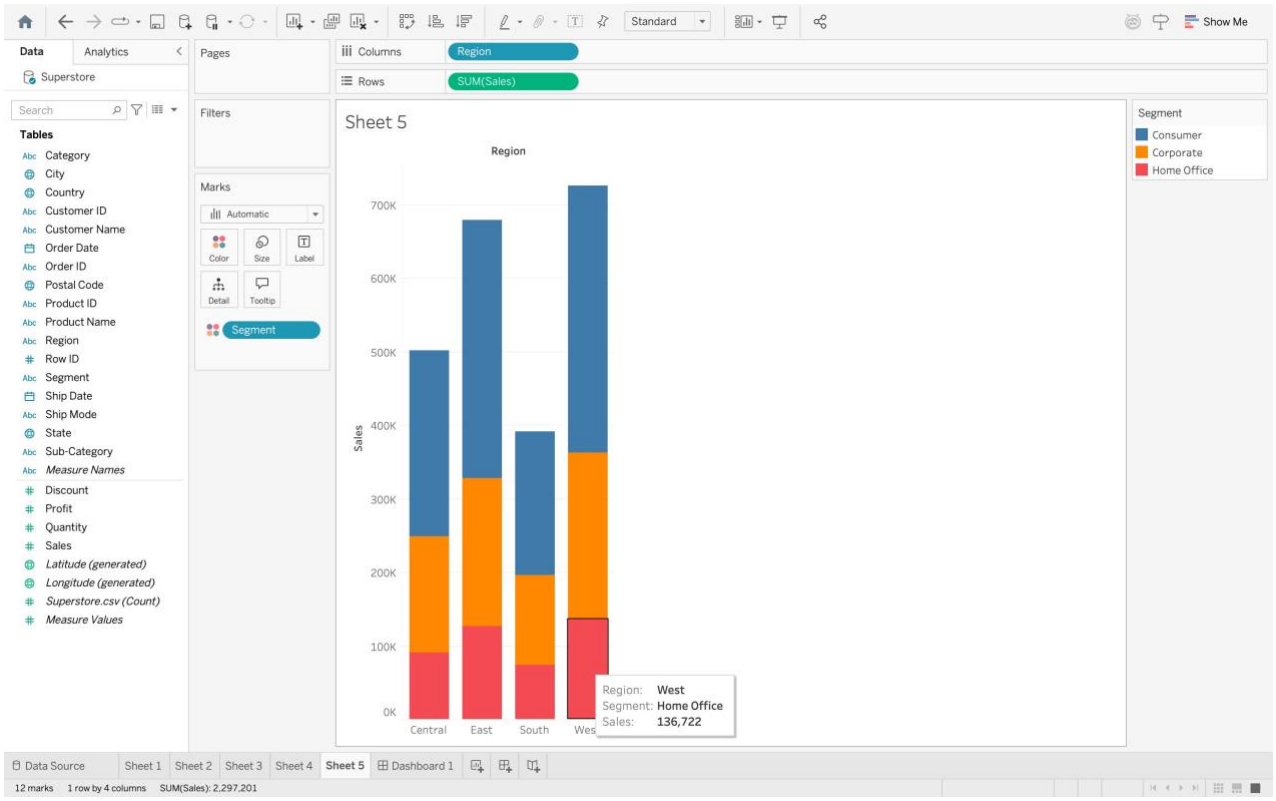
2)TABLEAU:-


Standard
Show Me
Data
Analytics
Superstore
Search
Tables
Category
City
Country
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Superstore.csv (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Columns
Region
Rows
SUM(Sales)
Sheet 1
Region
Sales
700K
600K
500K
400K
300K
200K
100K
0K
Central
East
South
West
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1
4 marks
1 row by 4 columns
SUM(Sales): 2,297,201


Standard
Show Me
Data
Analytics
Superstore
Search
Tables
Category
City
Country
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Superstore.csv (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Path
Columns
MONTH(Order Dat..
Rows
SUM(Sales)
Sheet 2
Order Date
Sales
350K
300K
250K
200K
150K
100K
50K
0K
January
February
March
April
May
June
July
August
Septemb..
October
November
December
Month of Order Date: August
Sales: 159,044
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1
12 marks
1 row by 12 columns
SUM(Sales): 2,297,201


Data
Analytics
Superstore
Search
Tables
Category
City
Country
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Superstore.csv (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Columns
SUM(Profit)
Rows
Sub-Category
Standard
Show Me
Sheet 3
Sub-C..
Copiers
Phones
Accessories
Paper
Binders
Chairs
Storage
Appliances
Furnishings
Envelopes
Art
Labels
Machines
Fasteners
Supplies
Bookcases
Tables
Sub-Category: Appliances
Profit: 18,138
-20K
-15K
-10K
-5K
0K
5K
10K
15K
20K
25K
30K
35K
40K
45K
50K
55K
Profit
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1
17 marks
17 rows by 1 column
SUM(Profit): 286,397


Data
Analytics
Superstore
Search
Tables
Category
City
Country
Customer ID
Customer Name
Order Date
Order ID
Postal Code
Product ID
Product Name
Region
Row ID
Segment
Ship Date
Ship Mode
State
Sub-Category
Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Superstore.csv (Count)
Measure Values
Pages
Filters
Marks
Pie
Color
Size
Label
Detail
Tooltip
Angle
Region
SUM(Sales)
SUM(Sales)
Columns
Rows
Standard
Show Me
Sheet 4
Region: East
Sales: 678,781
Region
Central
East
South
West
SUM(Sales)
2,297,201
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1
4 marks
1 row by 1 column
SUM(Sales): 2,297,201


Data
Analytics
Superstore
Search
Tables
Abc Category
City
Country
Abc Customer ID
Abc Customer Name
Order Date
Abc Order ID
Postal Code
Abc Product ID
Abc Product Name
Abc Region
Row ID
Abc Segment
Ship Date
Abc Ship Mode
State
Abc Sub-Category
Abc Measure Names
Discount
Profit
Quantity
Sales
Latitude (generated)
Longitude (generated)
Superstore.csv (Count)
Measure Values
Pages
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Segment
Columns
Region
Rows
SUM(Sales)
Standard
Show Me
Sheet 5
Region
Sales
700K
600K
500K
400K
300K
200K
100K
0K
Central
East
South
West
Region: West
Segment: Home Office
Sales: 136,722
Segment
Consumer
Corporate
Home Office
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1
12 marks
1 row by 4 columns
SUM(Sales): 2,297,201


Sheet 1
Region
Sales
600K
400K
200K
0K
Central
East
South
West
Sheet 5
Region
Sales
600K
400K
200K
0K
Central
East
South
West
Se..
Re..
Sal..
Sheet 3
Sub-C..
Copiers
Phones
Accessories
Paper
Binders
Chairs
Storage
Appliances
Furnishings
Envelopes
Art
Labels
Machines
Fasteners
Supplies
-20K
-10K
0K
10K
20K
30K
40K
50K
Profit
Sheet 4
Sheet 2
Order Date
Sales
300K
200K
100K
0K
February
April
June
August
October
December
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 5
Dashboard 1