Assignment No-6

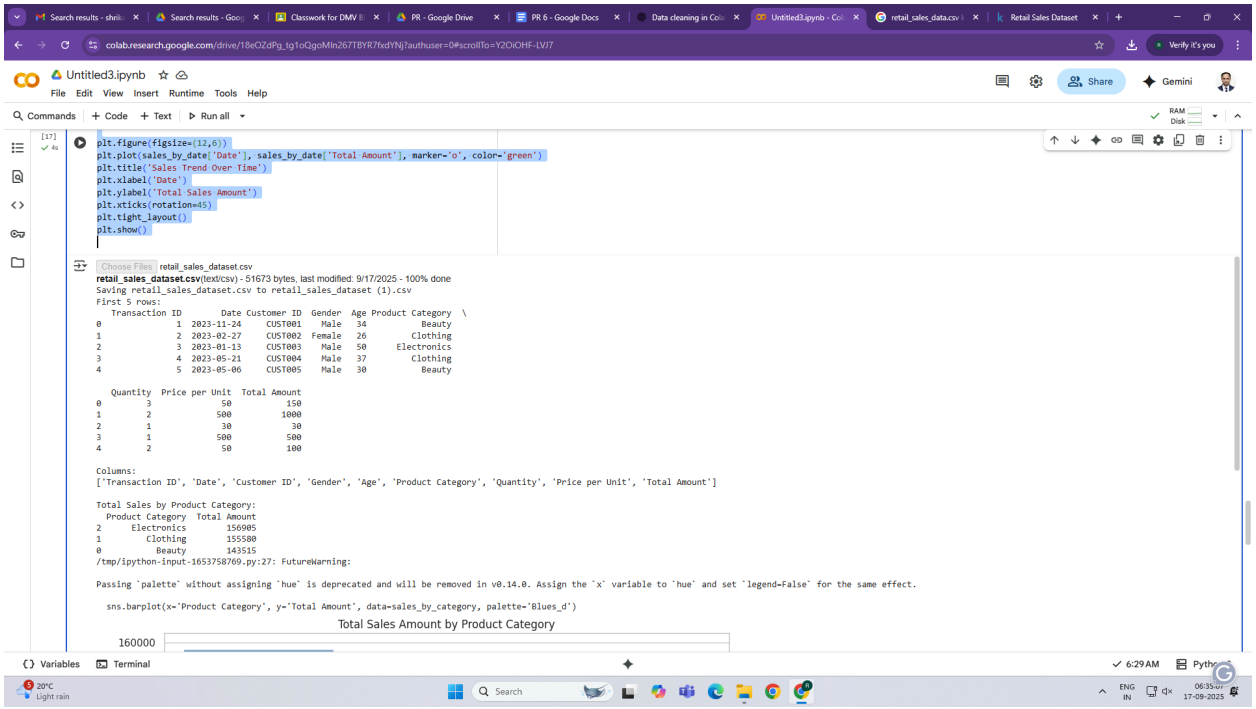
CODE:-

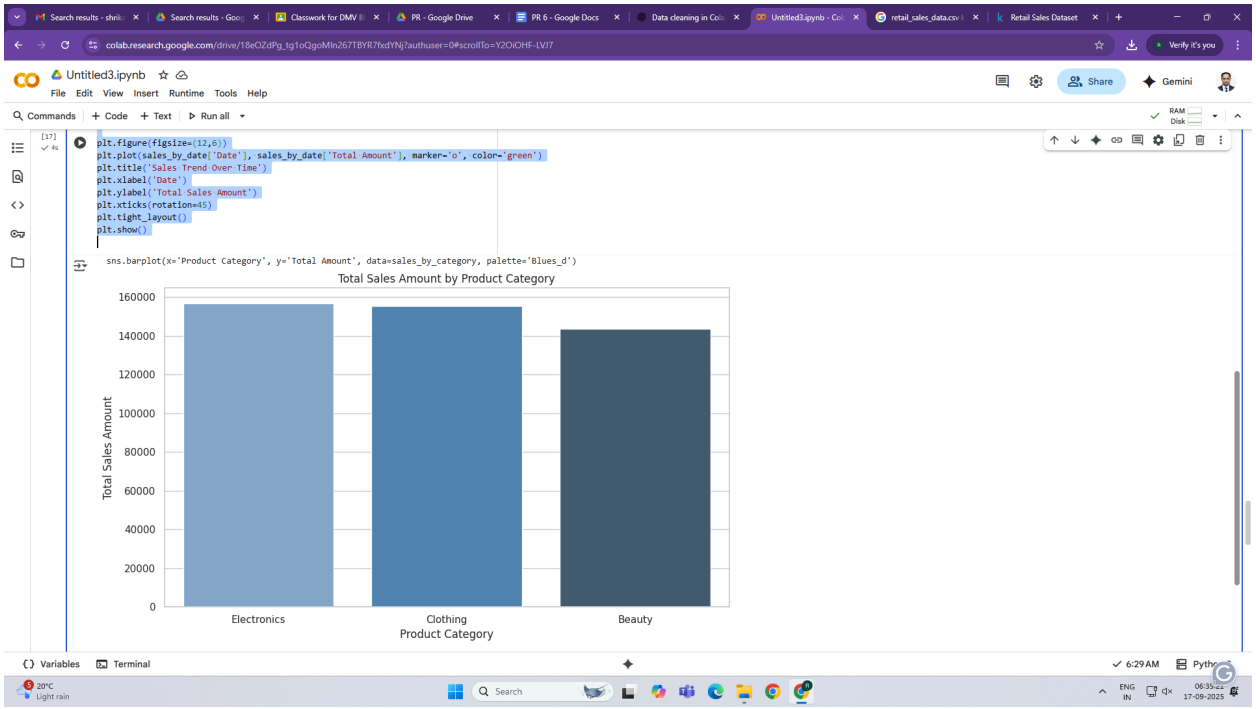
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from google.colab import files

sns.set(style="whitegrid")

# Upload and import dataset
uploaded = files.upload()
filename = next(iter(uploaded))
df = pd.read_csv(filename)

# Quick exploration
print("First 5 rows:")
print(df.head())

print("\nColumns:")
print(df.columns.tolist())

# 1. Aggregate total sales by Product Category
sales_by_category = df.groupby('Product Category')['Total
Amount'].sum().reset_index().sort_values(by='Total Amount',
ascending=False)
print("\nTotal Sales by Product Category:")
print(sales_by_category)

# 2. Bar plot of sales by Product Category
plt.figure(figsize=(10,6))
sns.barplot(x='Product Category', y='Total Amount',
data=sales_by_category, palette='Blues_d')
plt.title('Total Sales Amount by Product Category')
plt.xlabel('Product Category')
plt.ylabel('Total Sales Amount')
plt.tight_layout()
plt.show()
```

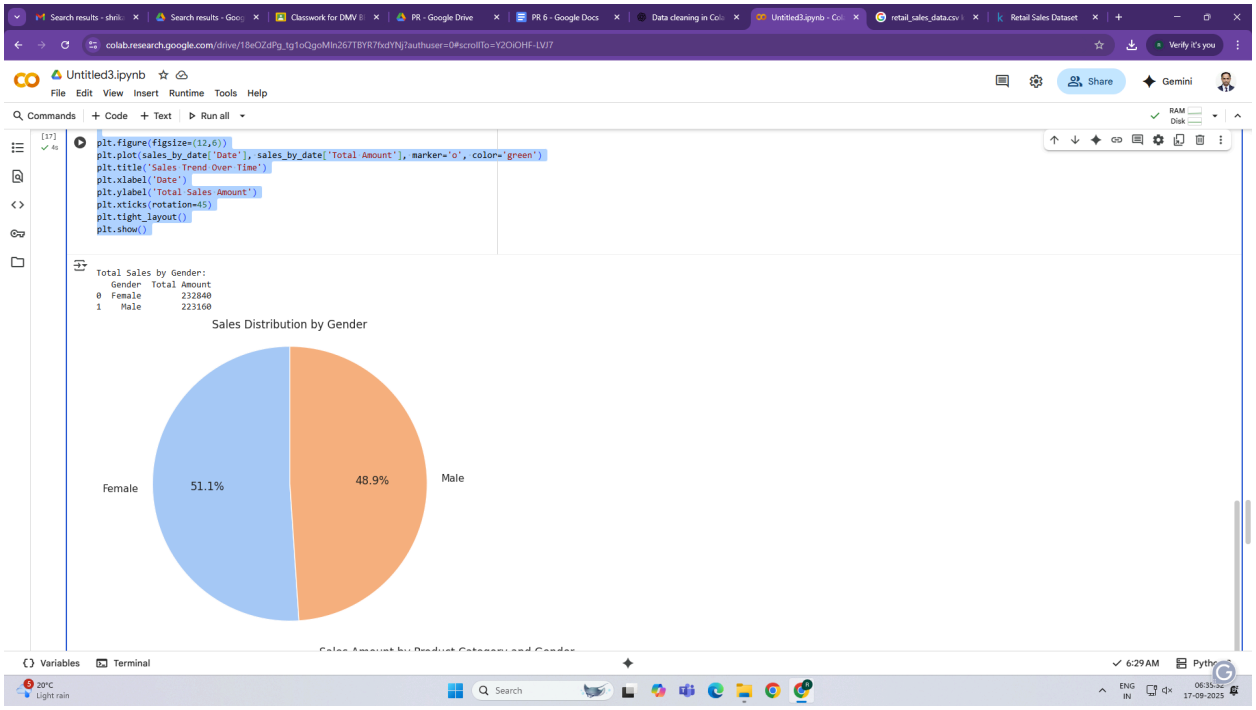
```
# 3. Aggregate total sales by Gender
sales_by_gender = df.groupby('Gender')['Total Amount'].sum().reset_index()
print("\nTotal Sales by Gender:")
print(sales_by_gender)

# 4. Pie chart for sales by Gender
plt.figure(figsize=(6,6))
plt.pie(sales_by_gender['Total Amount'], labels=sales_by_gender['Gender'],
autopct='%1.1f%%', startangle=90, colors=sns.color_palette('pastel'))
plt.title('Sales Distribution by Gender')
plt.axis('equal')
plt.show()

# 5. Aggregate sales by Product Category and Gender
sales_by_cat_gender = df.groupby(['Product Category', 'Gender'])['Total
Amount'].sum().reset_index()

# Pivot for stacked bar plot
pivot_table = sales_by_cat_gender.pivot(index='Product Category',
columns='Gender', values='Total Amount').fillna(0)

# 6. Stacked bar plot: Sales by Product Category and Gender
pivot_table.plot(kind='bar', stacked=True, figsize=(10,7),
colormap='Paired')
plt.title('Sales Amount by Product Category and Gender')
plt.xlabel('Product Category')
plt.ylabel('Total Sales Amount')
plt.xticks(rotation=45)
plt.legend(title='Gender')
plt.tight_layout()
plt.show()

# 7. Optional: Aggregate sales by Date to analyze sales trend over time
df['Date'] = pd.to_datetime(df['Date'])
sales_by_date = df.groupby('Date')['Total Amount'].sum().reset_index()

plt.figure(figsize=(12,6))
plt.plot(sales_by_date['Date'], sales_by_date['Total Amount'], marker='o',
color='green')
plt.title('Sales Trend Over Time')
```

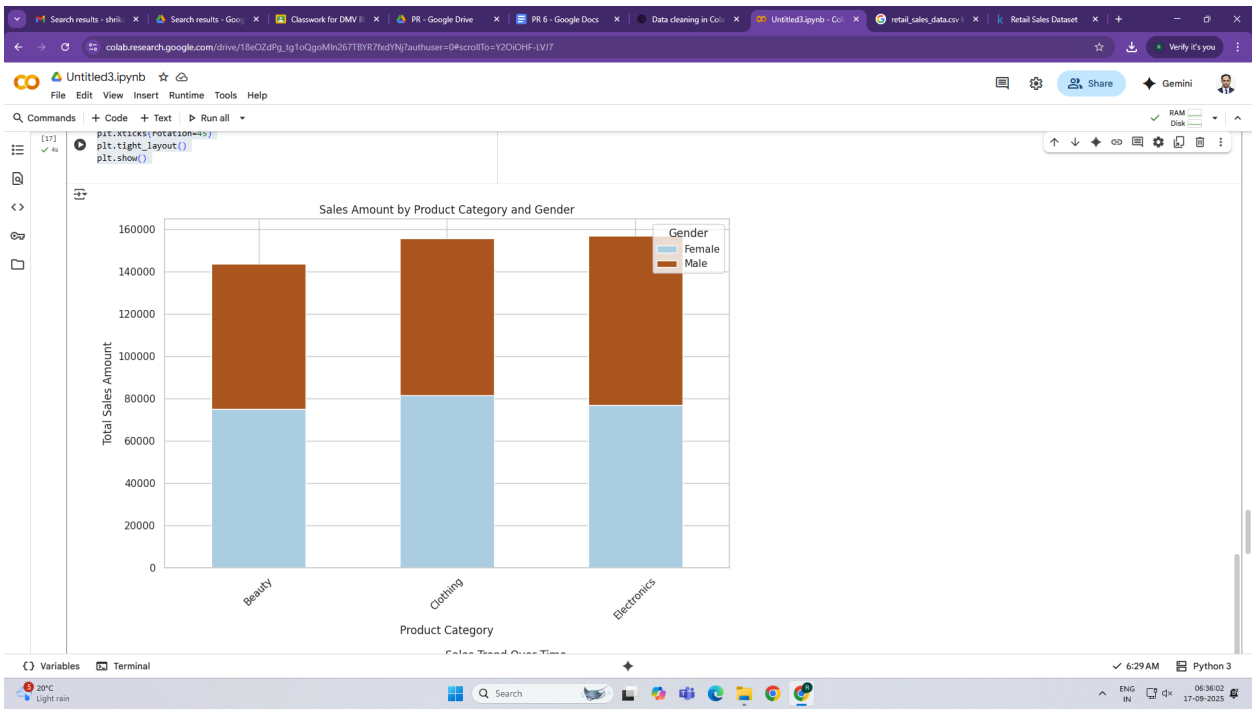
```
plt.xlabel('Date')
plt.ylabel('Total Sales Amount')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

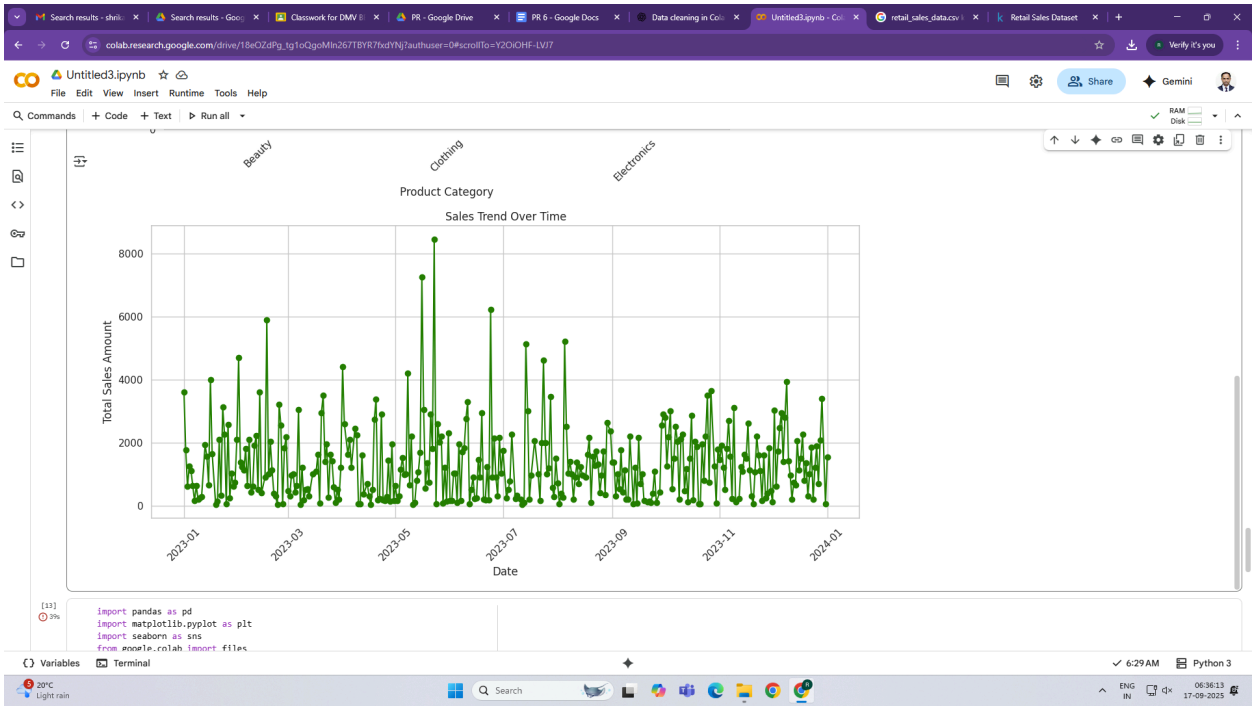
OUTPUT : –


Untitled3.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
plt.figure(figsize=(12,6))
plt.plot(sales_by_date['Date'], sales_by_date['Total Amount'], marker='o', color='green')
plt.title('Sales Trend Over Time')
plt.xlabel('Date')
plt.ylabel('Total Sales Amount')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
Choose Files retail_sales_dataset.csv
retail_sales_dataset.csv(text/csv) - 51673 bytes, last modified: 9/17/2025 - 100% done
Saving retail_sales_dataset.csv to retail_sales_dataset (1).csv
First 5 rows:
Transaction ID Date Customer ID Gender Age Product Category \
0 1 2023-11-24 CUST001 Male 34 Beauty
1 2 2023-02-27 CUST002 Female 26 Clothing
2 3 2023-01-13 CUST003 Male 50 Electronics
3 4 2023-05-21 CUST004 Male 37 Clothing
4 5 2023-05-06 CUST005 Male 30 Beauty
Quantity Price per Unit Total Amount
0 3 50 150
1 2 500 1000
2 1 30 30
3 1 500 500
4 2 50 100
Columns:
['Transaction ID', 'Date', 'Customer ID', 'Gender', 'Age', 'Product Category', 'Quantity', 'Price per Unit', 'Total Amount']
Total Sales by Product Category:
Product Category Total Amount
2 Electronics 156905
1 Clothing 155580
0 Beauty 143515
/tmp/ipython-input-1653758769.py:27: 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.barplot(x='Product Category', y='Total Amount', data=sales_by_category, palette='Blues_d')
Total Sales Amount by Product Category
160000
Variables Terminal
6:29 AM


Untitled3.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
plt.figure(figsize=(12,6))
plt.plot(sales_by_date['Date'], sales_by_date['Total Amount'], marker='o', color='green')
plt.title('Sales Trend Over Time')
plt.xlabel('Date')
plt.ylabel('Total Sales Amount')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
sns.barplot(x='Product Category', y='Total Amount', data=sales_by_category, palette='Blues_d')
Total Sales Amount by Product Category
Total Sales Amount
160000
140000
120000
100000
80000
60000
40000
20000
0
Electronics
Clothing
Beauty
Product Category
Variables Terminal
6:29 AM


Untitled3.ipynb
File Edit View Insert Runtime Tools Help
Share
Gemini
Commands + Code + Text Run all
plt.figure(figsize=(12,6))
plt.plot(sales_by_date['Date'], sales_by_date['Total Amount'], marker='o', color='green')
plt.title('Sales Trend Over Time')
plt.xlabel('Date')
plt.ylabel('Total Sales Amount')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
Total Sales by Gender:
Gender Total Amount
0 Female 232840
1 Male 223160
Sales Distribution by Gender
Female
51.1%
48.9%
Male
Variables
Terminal
6:29 AM


Untitled3.ipynb
File Edit View Insert Runtime Tools Help
Share
Gemini
Commands + Code + Text Run all
plt.tight_layout()
plt.show()
Sales Amount by Product Category and Gender
Gender
Female
Male
Total Sales Amount
160000
140000
120000
100000
80000
60000
40000
20000
0
Beauty
Clothing
Electronics
Product Category
Variables
Terminal
6:29 AM
Python 3




```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from google.colab import files
```