

**Infotact Solutions**

**Sales Data Analysis and Dashboard (Excel/Power BI)**

**Project 1**

Data Analytics **Team 7**

Batch 5.1

**Team Members**

Sandhya S

Pavithra S

Manya

Adithya

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl no.** | **Topic** | **Page no** |
| **1.** | Introduction | **3** |
| **2.** | Problem Statement | **3** |
| **3.** | Objectives | **3** |
| **4.** | Tools and Technologies used | **4** |
| **5.** | Software and Hardware Requirements | **4** |
| **6.** | Dataset Used | **5** |
| **7.** | Dataset Overview | **5-6** |
| **8.** | Data Analysis And Findings | **6-7** |
| **9.** | Business Insights | **7** |
| **10.** | Recommendation | **7** |
| **11.** | Screenshots | **8-19** |
| **12.** | Conclusion | **19** |
| **13.** | Reference | **19** |

**Project Title: *Barista Coffee Sales Insights***

**1.Introduction:**

In a data-centric era, the ability to interpret and analyze data effectively is crucial for business growth. This project focuses on analyzing coffee sales data to uncover patterns, trends, and insights that can guide decision-making. Using a combination of Microsoft Excel, Power BI, and Python (via Jupyter Notebook), we have developed a full-scale analytical report and interactive dashboard to help stakeholders identify key performance metrics and areas for business improvement.

The project primarily deals with data from a coffee shop business and aims to assist in understanding customer behavior, sales trends, and the performance of products and regions.

**2. Problem Statement:**

Businesses often possess large volumes of raw sales data but lack the analytical tools and processes to derive meaningful insights. Key challenges include:

* **Data Complexity:** Raw datasets often contain inconsistencies, missing values, and lack structure.
* **Static Reports:** Traditional reporting lacks interactivity and real-time updates.
* **Decision Delay:** Insights are often delayed due to manual processes and lack of visualization.

Our solution addresses these problems by building an end-to-end data pipeline that cleans, analyzes, and visualizes data interactively and efficiently.

**3. Objectives:**

The goals of this project are:

* To clean and structure raw coffee sales data for analysis.
* To conduct detailed analysis on sales trends across various dimensions.
* To build dynamic, user-friendly dashboards that allow data exploration.
* To provide strategic recommendations based on data-driven findings.

**Inventory & Marketing Planning:**  
 - Stock up on Coffee and Cake during Q4 to meet high seasonal demand.  
 - Plan combo offers featuring Pizza and Merchandise.  
  
- **Promotions Strategy:**  
 - Increase Flat Discount and BOGO offers to boost low-performing times (Evening/Night).  
 - Continue targeted promos for Coffee, which shows high responsiveness.  
  
- **Customer Engagement:**  
 - Expand loyalty programs to further increase repeat purchases.  
 - Use morning promotions for Tea, targeting the most active time block.  
  
- **Regional Campaigns:** - Launch special campaigns in Downtown and Uptown stores to improve performance.  
 - Analyze underperforming SKUs (e.g., Cold Drinks in winter) and consider seasonal adjustments.

**4. Tools and Technologies Used:**

| **Tool/Technology** | **Purpose** |
| --- | --- |
| Microsoft Excel | Creating pivot table analysis, KPI cards & basic charts |
| Power BI | Data modeling, DAX calculations, dashboard creation |
| Python (Jupyter) | Automated data processing, advanced visualizations |
| Pandas, Matplotlib, Seaborn | Python libraries used for data manipulation and visualization |
| Source Data | Coffee Sales Dataset in Excel format, integrated into Power BI |

**5. Software and Hardware Requirements:**

**5.1 Software Requirements:**

| **Software** | **Version/Details** |
| --- | --- |
| Microsoft Excel | 2016 or later |
| Power BI Desktop | Latest version (recommended) |
| Python | Version 3.8 or above |
| Jupyter Notebook | Installed via Anaconda or pip |
| Operating System | Windows 10 or later (64-bit) |
| Additional Libraries | pandas, numpy, matplotlib, seaborn |

**5.2 Hardware Requirements:**

| **Component** | **Minimum Requirement** | **Recommended** |
| --- | --- | --- |
| Processor | Intel Core i5 or AMD Ryzen 5 | Intel i7 or higher |
| RAM | 8 GB | 16 GB or more |
| Storage | 1 GB free disk space | SSD with 10+ GB free |
| Display | 1366 x 768 resolution | Full HD (1920 x 1080) |
| Graphics | Integrated Graphics | Dedicated GPU (for large visuals) |

**6. Dataset Used:**

**Dataset Title: *Barista* *Coffee Sales Dataset***

**Source Files:**

* **Barista Coffee Sales.xlsx (Primary dataset used for analysis)**
* **Used in combination with baristaa.pbix and Barista.ipynb for visualization and coding**

**Origin:**

* **Sourced from Kaggle and internal mock sales data generation**
* **Simulates a multi-regional coffee shop chain’s historical transactions**

**Content Summary:**

| **Field/Column** | **Description** |
| --- | --- |
| **Date** | **Transaction date of the sale** |
| **Product** | **Name of the coffee product sold (e.g., Merchendise, Coffee, Cake)** |
| **Quantity** | **Number of units sold per transaction** |
| **Revenue** | **Total revenue from the sale** |
| **Location** | **Sales Location (e.g., Aiport, Downtown)** |
| **Customer Type** | **Loyalty or new customer** |

**Usage:**

* **Used in Excel for pivot-based breakdowns**
* **Loaded into Python for transformation and trend analysis**
* **Modeled in Power BI for dashboard development**

**Preprocessing Done:**

* **Missing value treatment**
* **Date conversion for time-series analysis**
* **Region and product standardization**
* **Removal of duplicates and normalization**

**7. Dataset Overview:**

**Source:**

* The dataset is titled Coffee Sales.xlsx and includes historical sales records.

**Attributes:**

* **Product Name**
* **Quantity Sold**
* **Revenue (Total Sale Amount)**
* **Location**
* **Customer Type (e.g., Loyalty/New)**
* **Date/Time**

**Preprocessing Steps:**

* Replaced missing values and standardized column headers.
* Removed duplicates and validated data types.
* Converted date fields for temporal analysis.

**8. Data Analysis & Findings:**

**A. Excel Analysis:**

* Created pivot tables to identify:
  + Top-selling products
  + Monthly sales breakdown
  + Regional performance variations

**Key Insights:**

* Products like **Merchendise**, **Pizza**, **Cold drinks** and **Coffee** had the highest sales volume.
* The **Downtown** and **Suburbs** loaction led in total revenue generation.

**B. Python (Jupyter Notebook) Analysis:**

Performed **Exploratory Data Analysis (EDA)** using pandas, matplotlib, and seaborn.

**Key Analysis Included:**

* **Trend Over Time:** Monthly and quarterly revenue analysis.
* **Customer Segmentation:** Behavior patterns between loyalty and one-time customers.
* **Correlation Heatmaps:** To identify relationships between sales variables.

Python CopyEdit

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

df = pd.read\_excel("Coffee Sales.xlsx")

sns.lineplot(x="Month", y="Revenue", data=df)

plt.title("Monthly Revenue Trends")

**Findings:**

* Sales consistently increased in **Q4**, indicating a strong **seasonal trend**.
* Loyalty customers spent significantly more per transaction.

**C. Power BI Dashboard:**

Power BI was used to build an interactive dashboard featuring:

* **KPIs:** Total Sales, Revenue Growth %, Units Sold
* **Visual Charts:** Bar charts, pie charts, line graphs, and maps
* **Interactivity:** Slicers by region, customer type, and product

**Dashboard Highlights:**

* **Dynamic Filtering:** Instantly compare sales by region or product category.
* **Heat Maps:** Show concentration of sales by geographic region.
* **Drillthrough Capability:** Users can drill down from product category to individual SKUs.

**9. Business Insights:**

**Business Insights:**  
  
- **Product Performance:**  
 - Merchandise, Pizza, and Coffee are consistent top performers.  
 - Coffee sales in 2023 show a dramatic increase, indicating growing demand.  
  
- **Store Locations:**  
 - Downtown and Suburbs locations generate the highest sales; Uptown and Airport underperform comparatively.  
  
- **Seasonal Trends:** - Q4 consistently shows revenue spikes across most product categories, especially Cake and Coffee.  
  
**- Promotion Effectiveness:**  
 - Promo usage (50–52%) is closely tied to increased sales, especially for Coffee.  
 - Flat Discounts and BOGO generate the most revenue in promotional campaigns.  
  
- **Customer Behavior:**  
 - Morning is the most active purchase time, especially for Tea.  
 - Loyalty/repeat customers spend more and engage frequently in promotions.

* **Top Products:** Merchendise , Pizza and Coffee consistently led sales across all periods.
* **Seasonal Patterns:** Q4 is a critical revenue-generating period.
* **Customer Loyalty:** Loyalty program members deliver higher per-sale revenue.
* **Regional Trends:** South region lags behind; needs targeted marketing.
* **Product Gaps:** Underperformance in cold beverages during winter months.

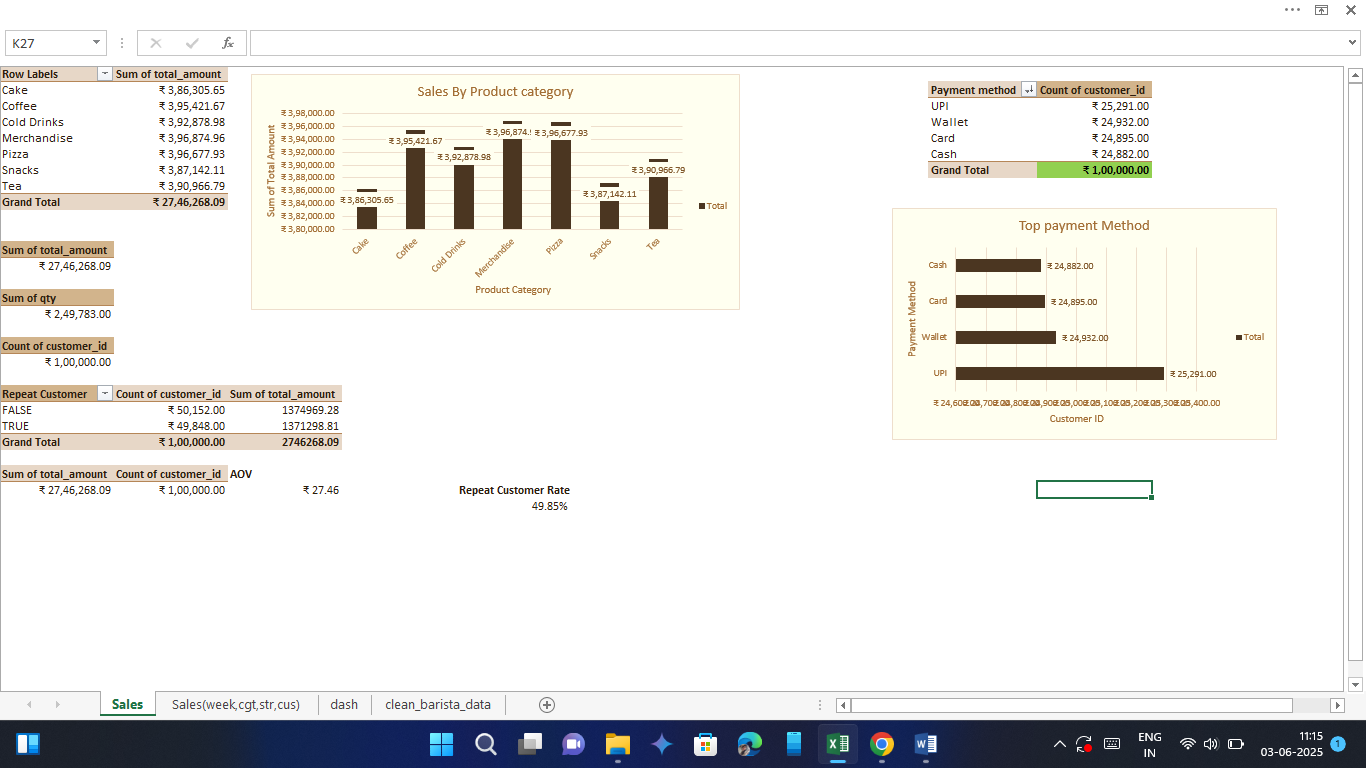
**10. Recommendations:**

* **Inventory Optimization:** Prepare high-selling SKUs in advance for Q4.
* **Geographic Expansion:** Focus on boosting sales in the South with offers and campaigns.
* **Marketing Strategy:** Use loyalty incentives to increase repeat purchases.
* **Product Mix Adjustments:** Phase out consistently underperforming products.
* **Dashboard Usage:** Implement Power BI dashboard at operational levels for real-time tracking.

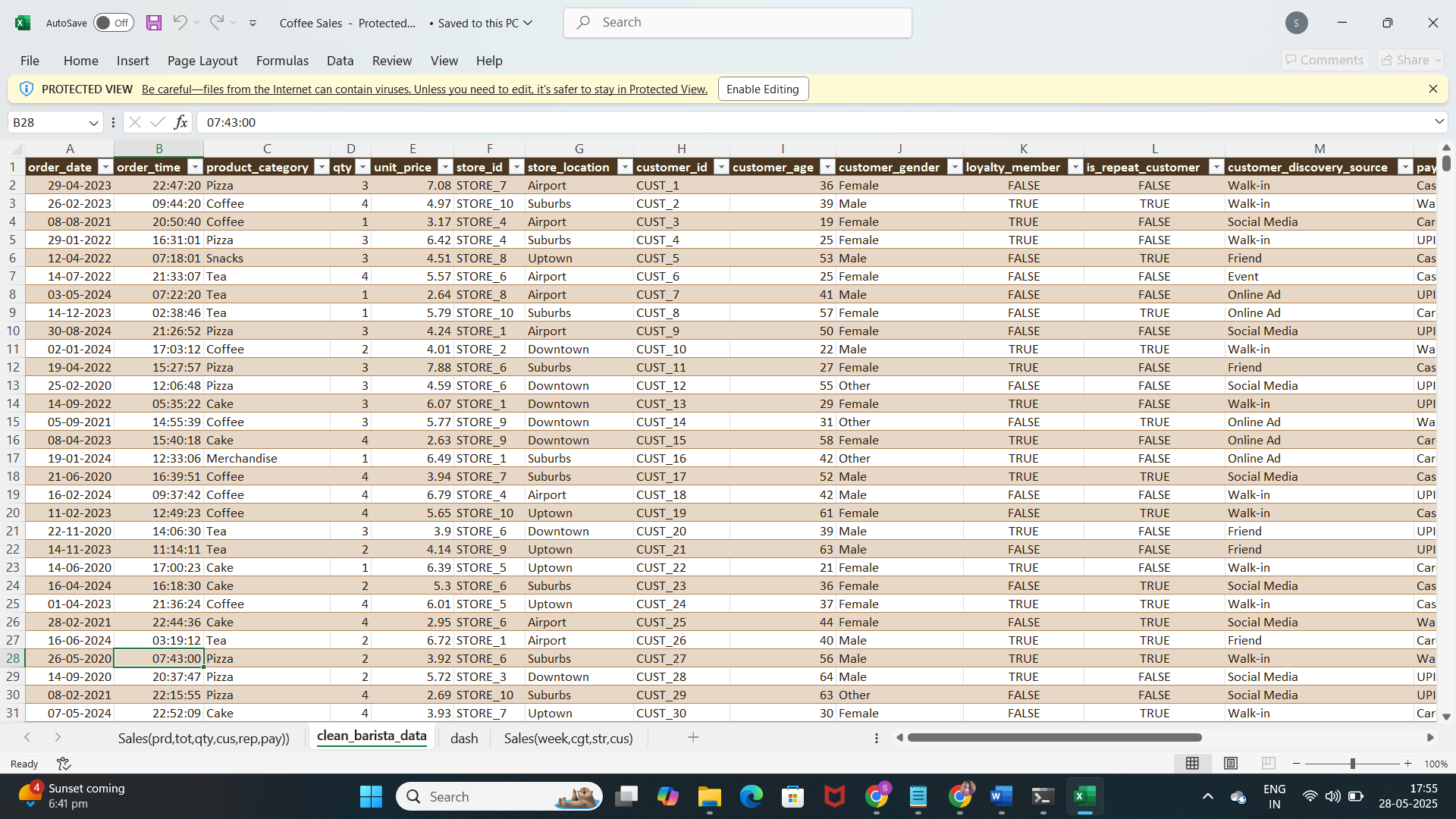
**11. Screenshots:**

**Excel Part:**

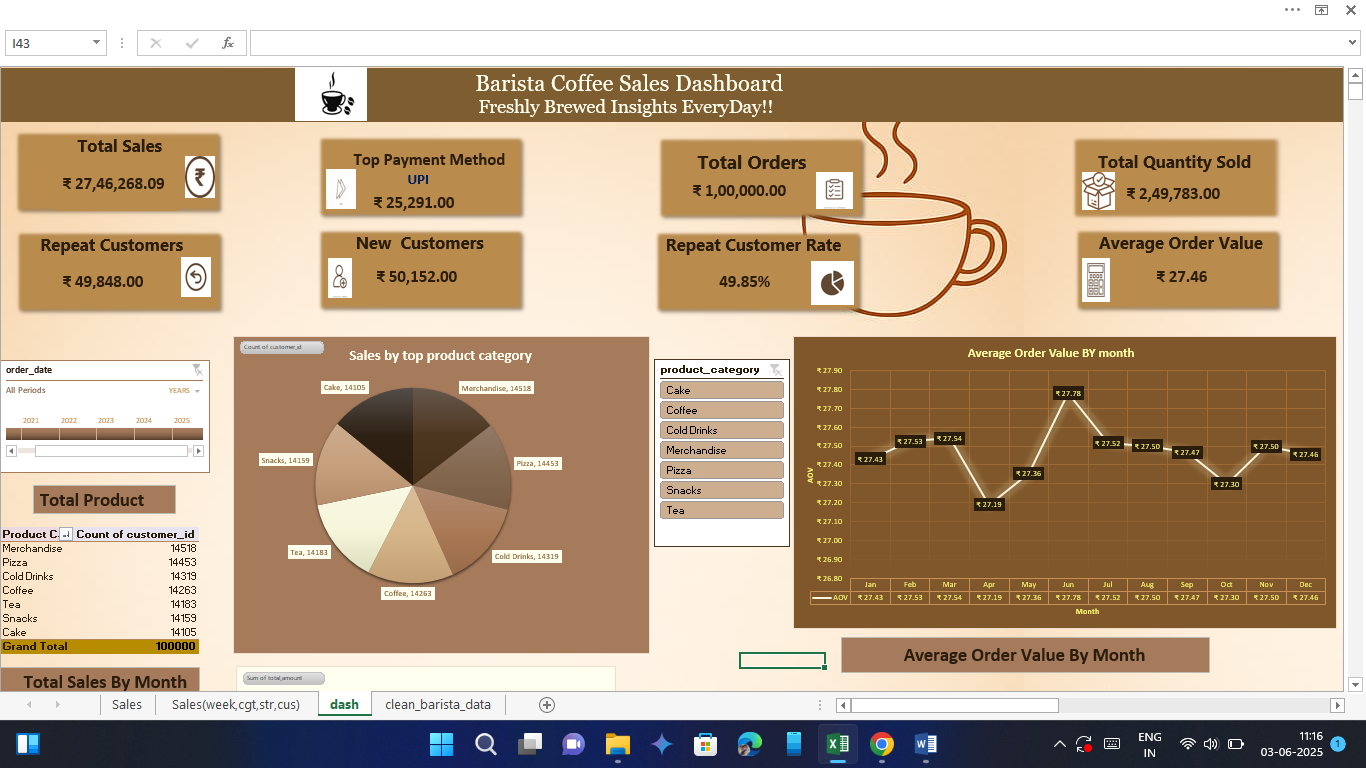
**Sales:**

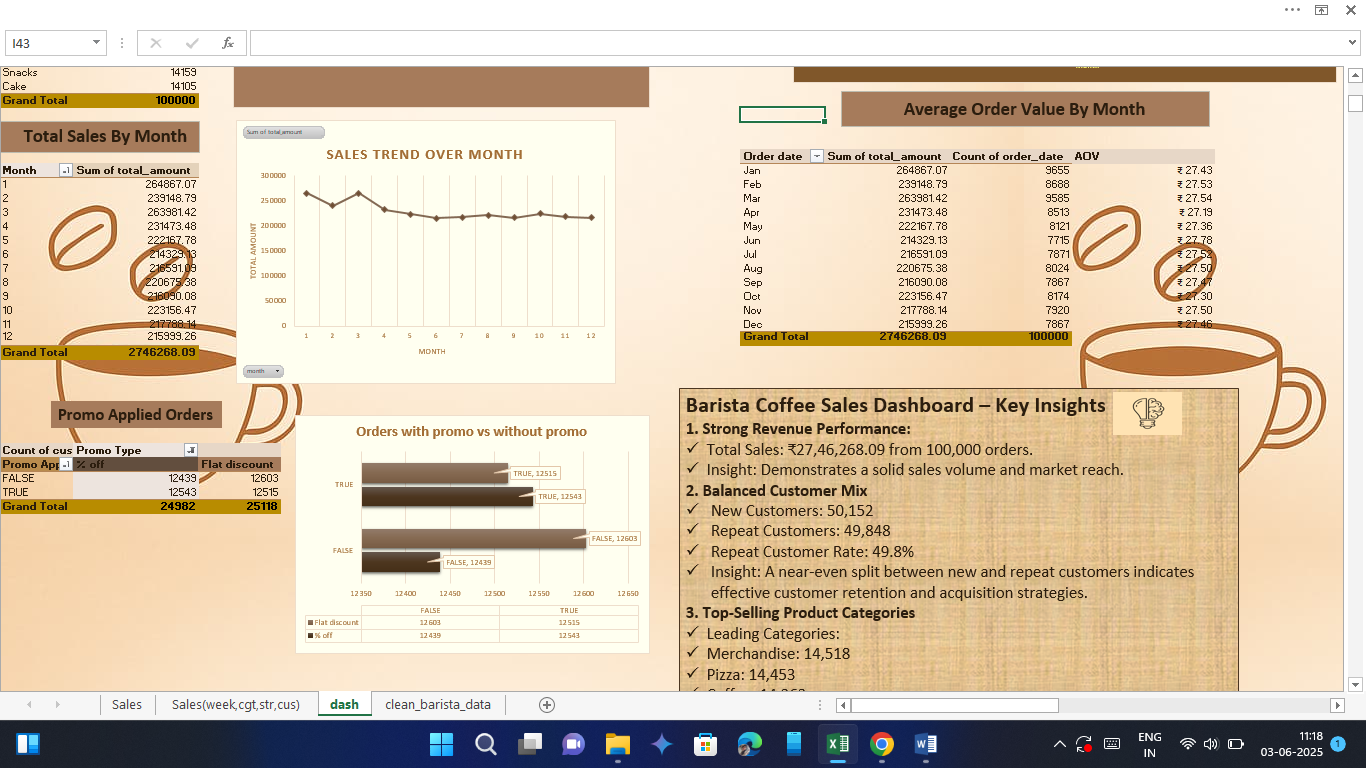


**Dataset:**



**Dashboard:**

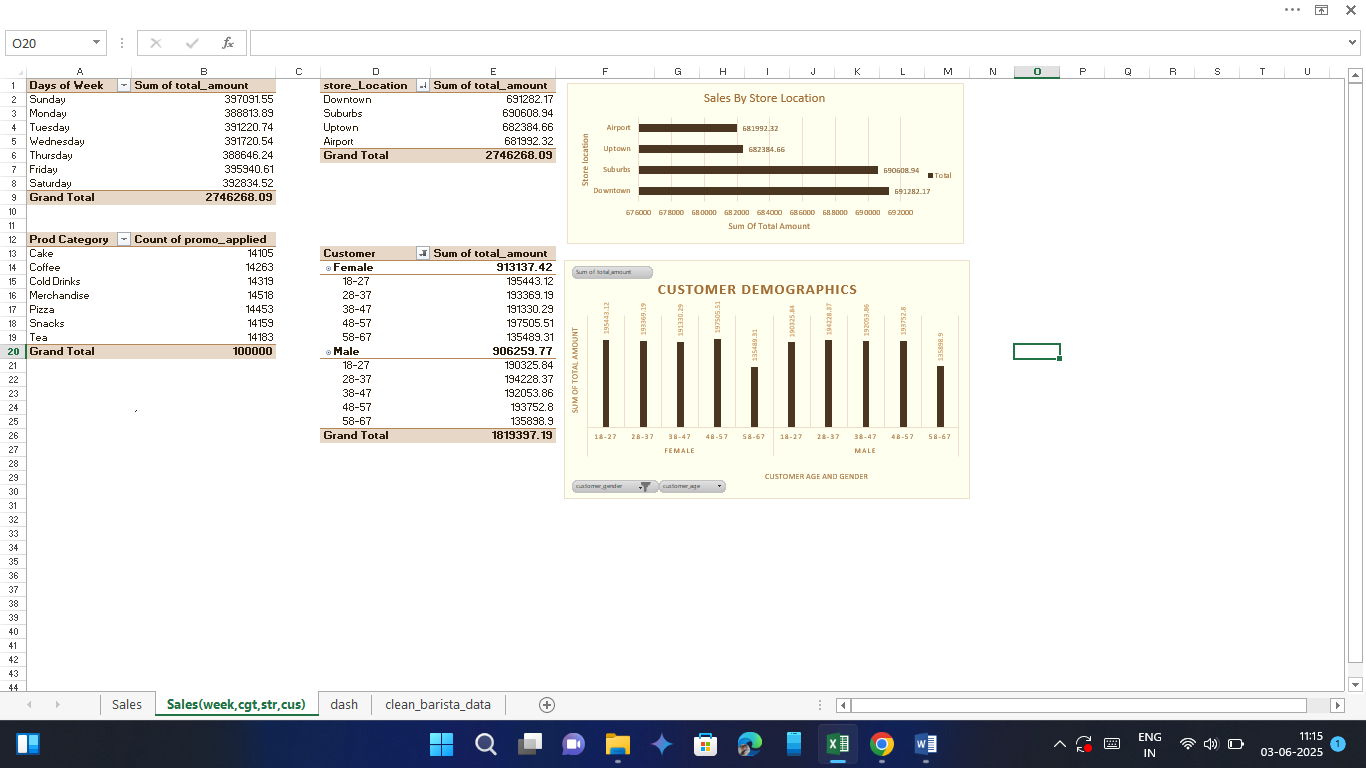




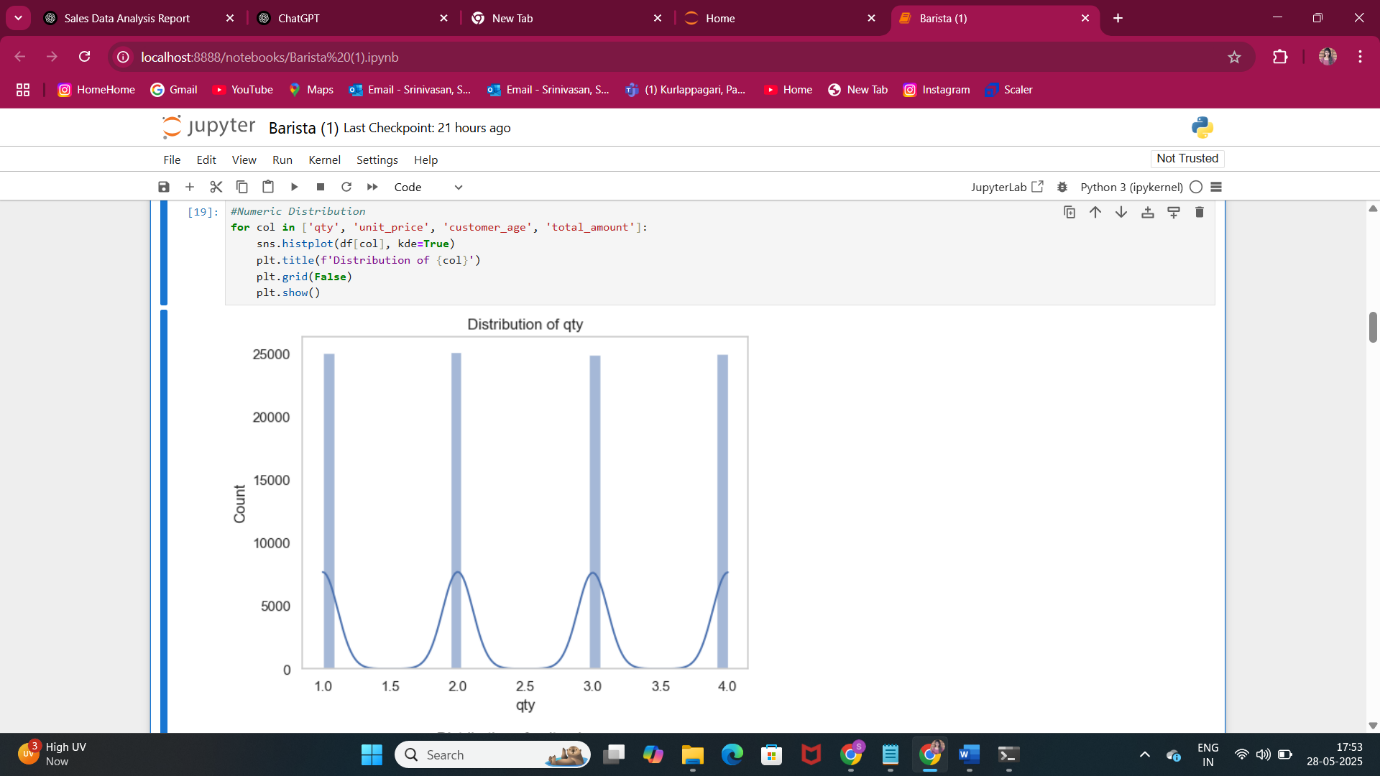
**After Applying Timeline and Product Slicer (2023, Coffee):**

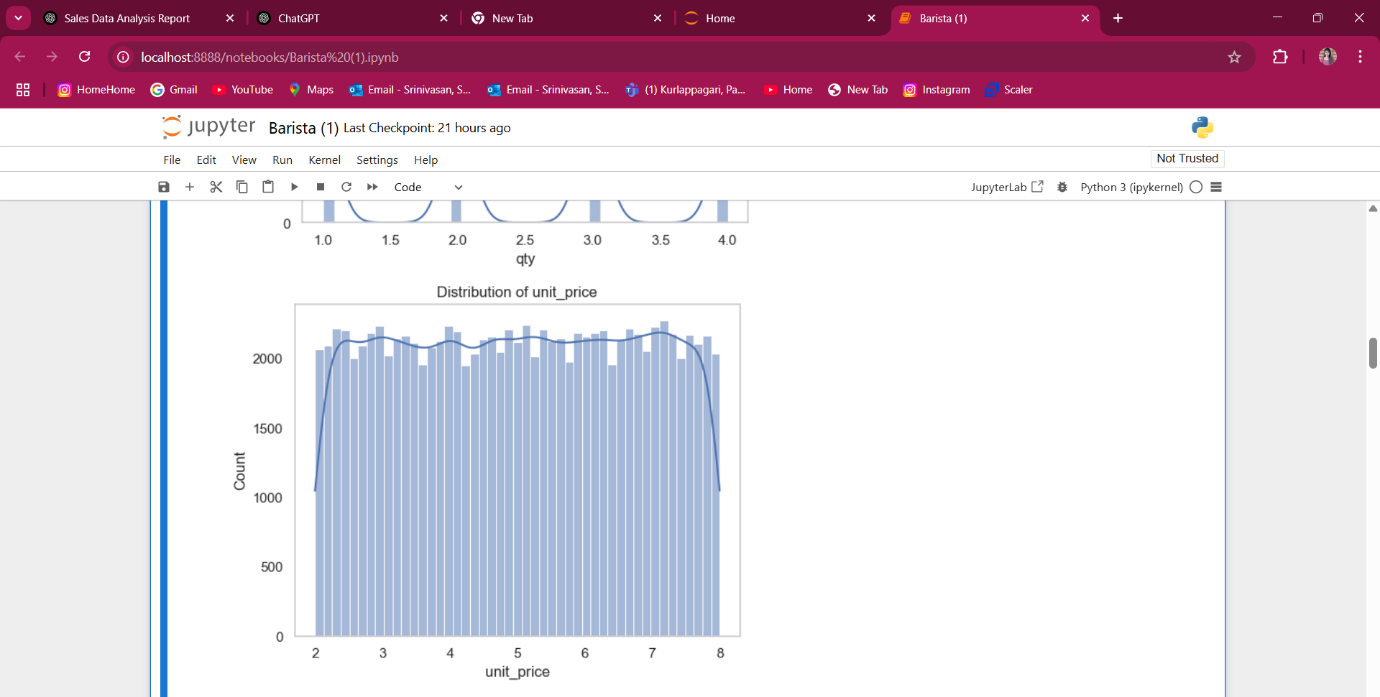


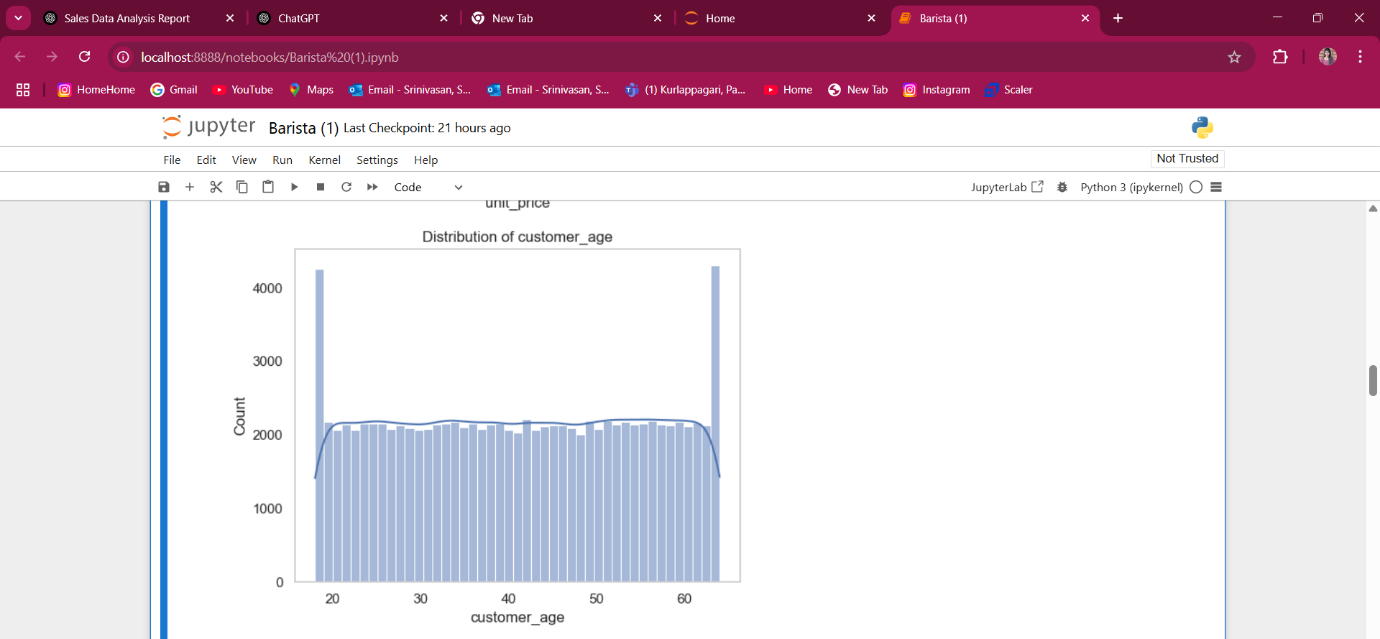
**Sales 2:**

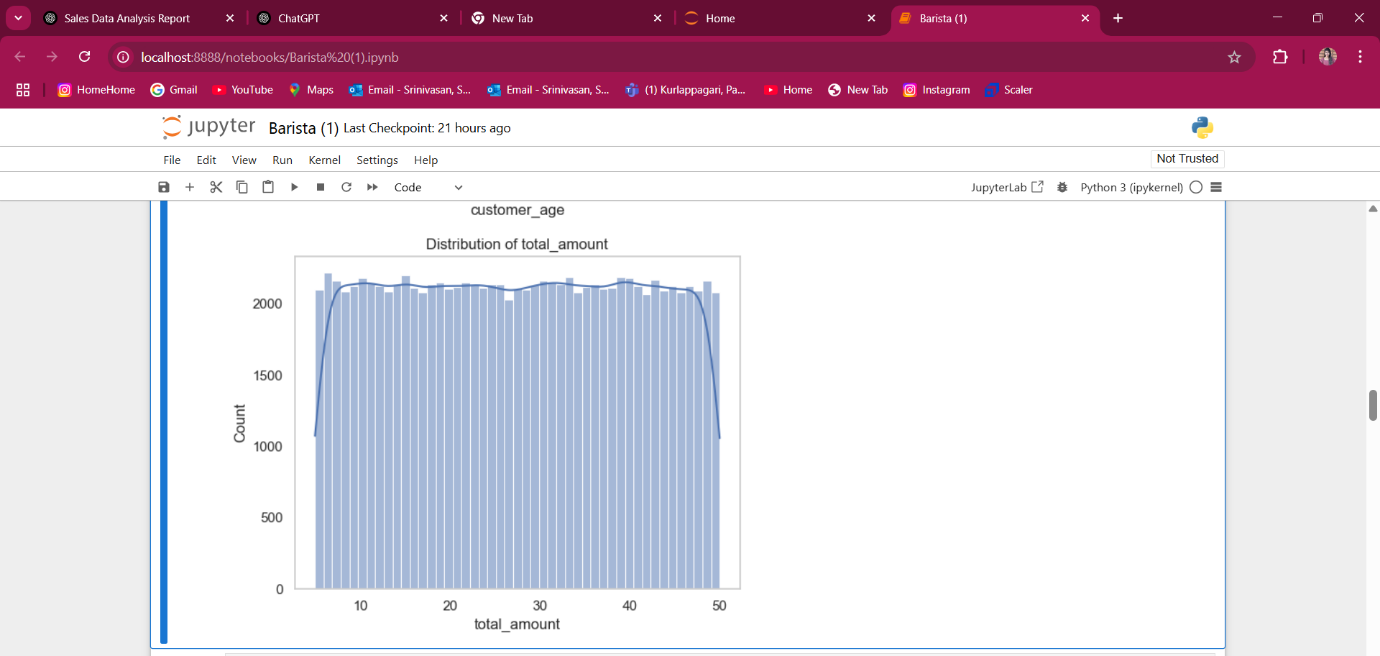


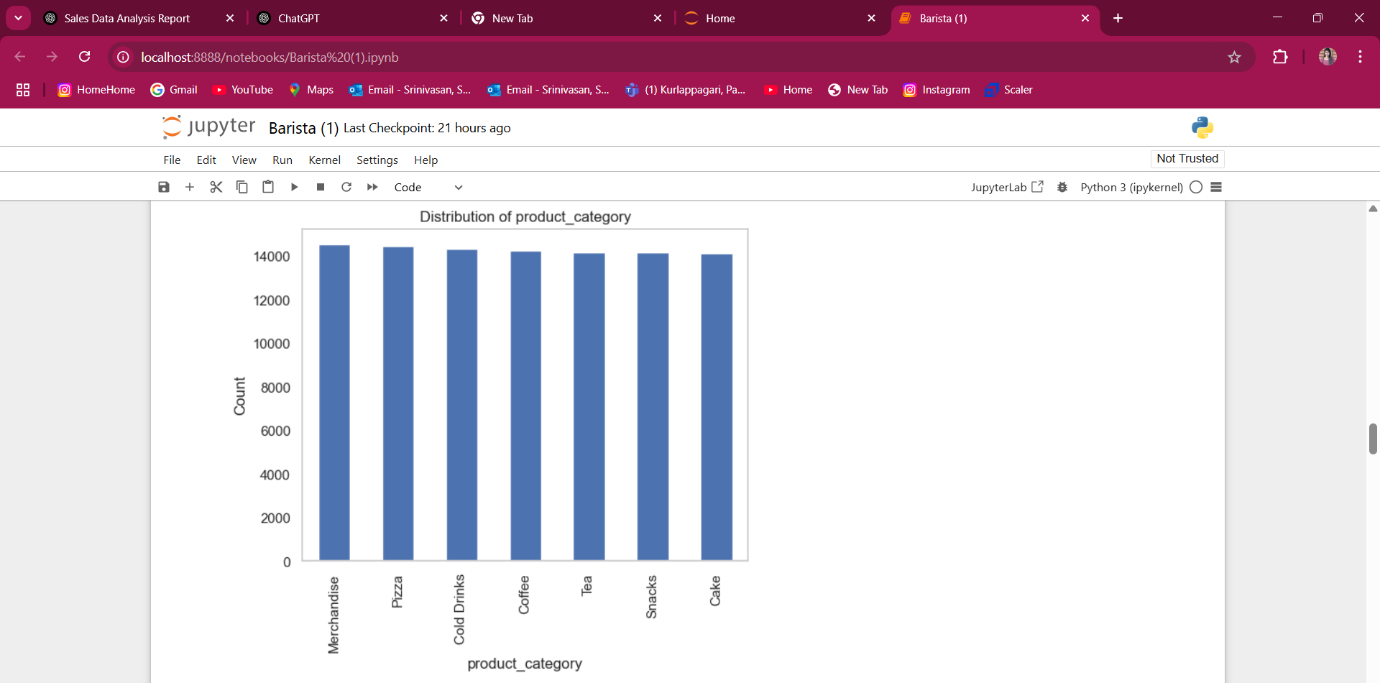
**PYTHON (EDA):**

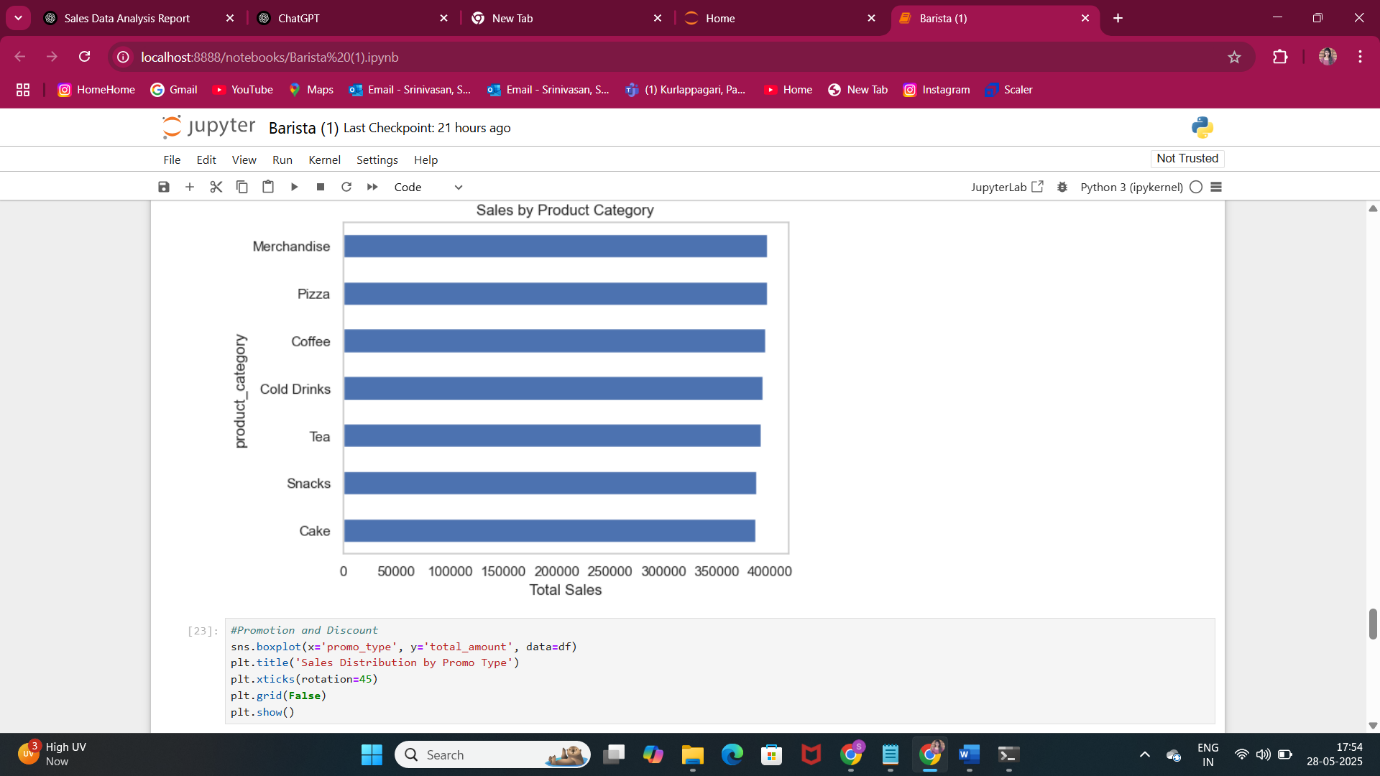


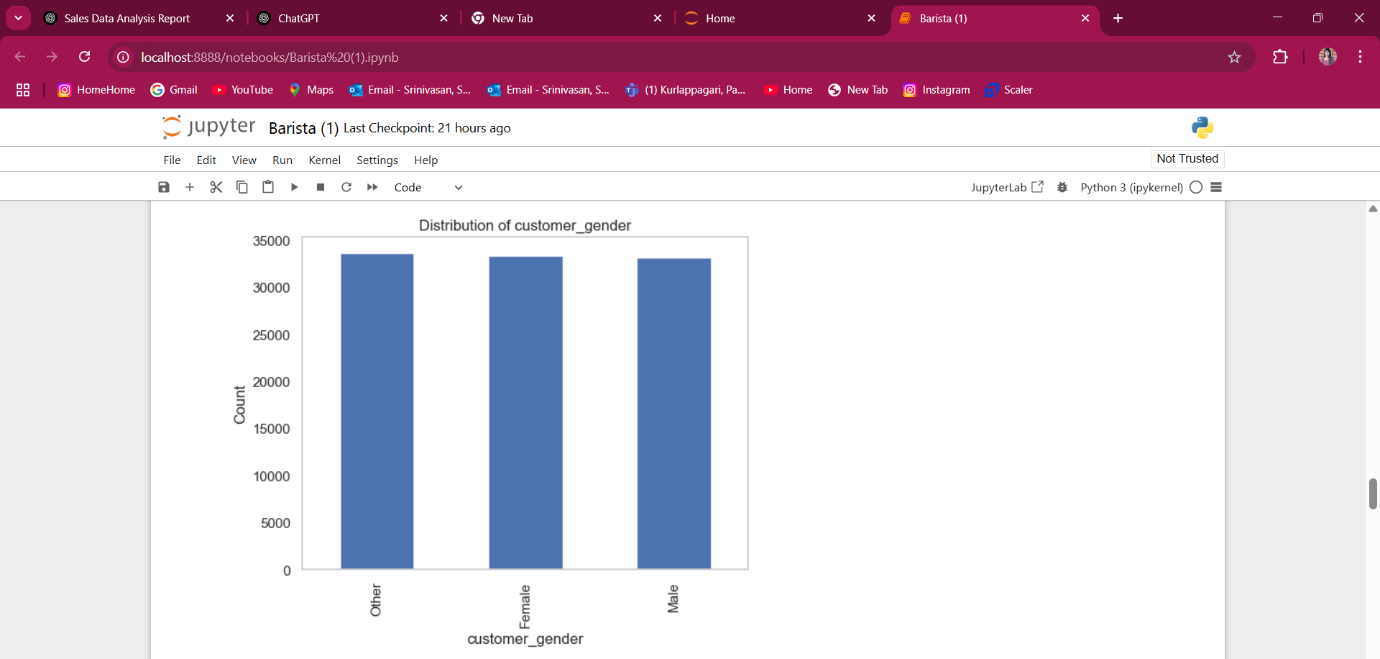


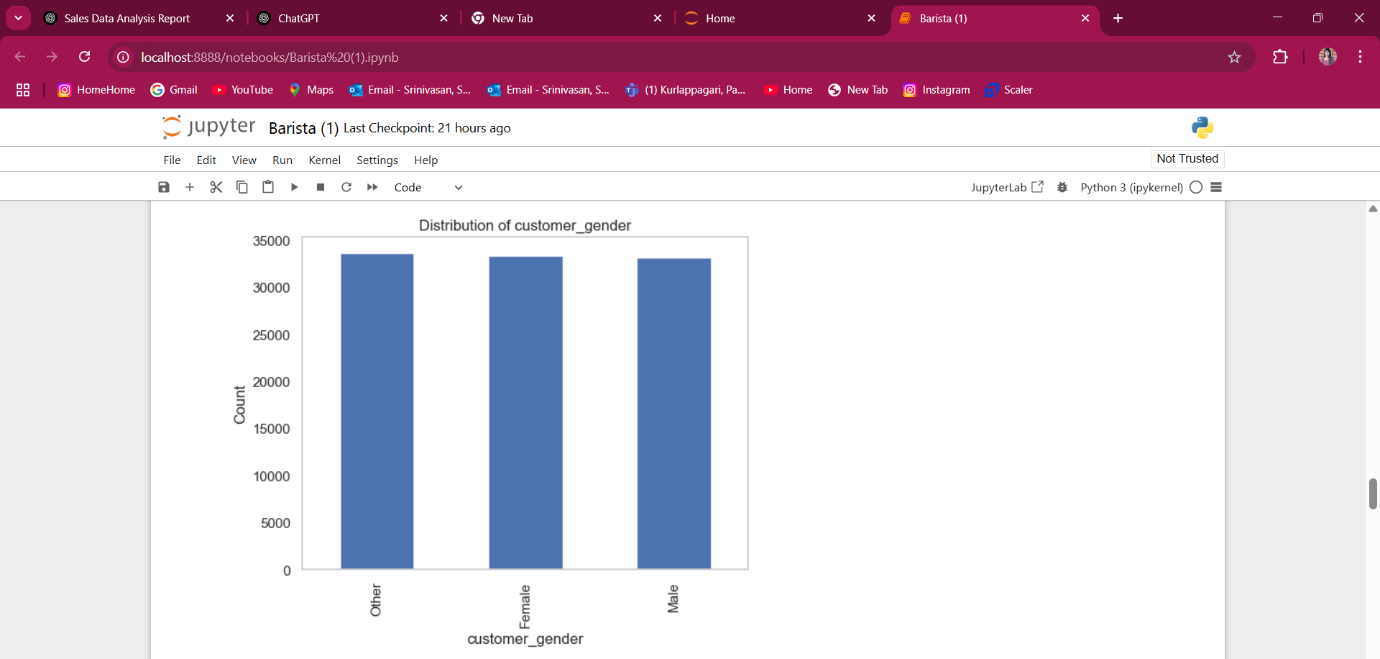


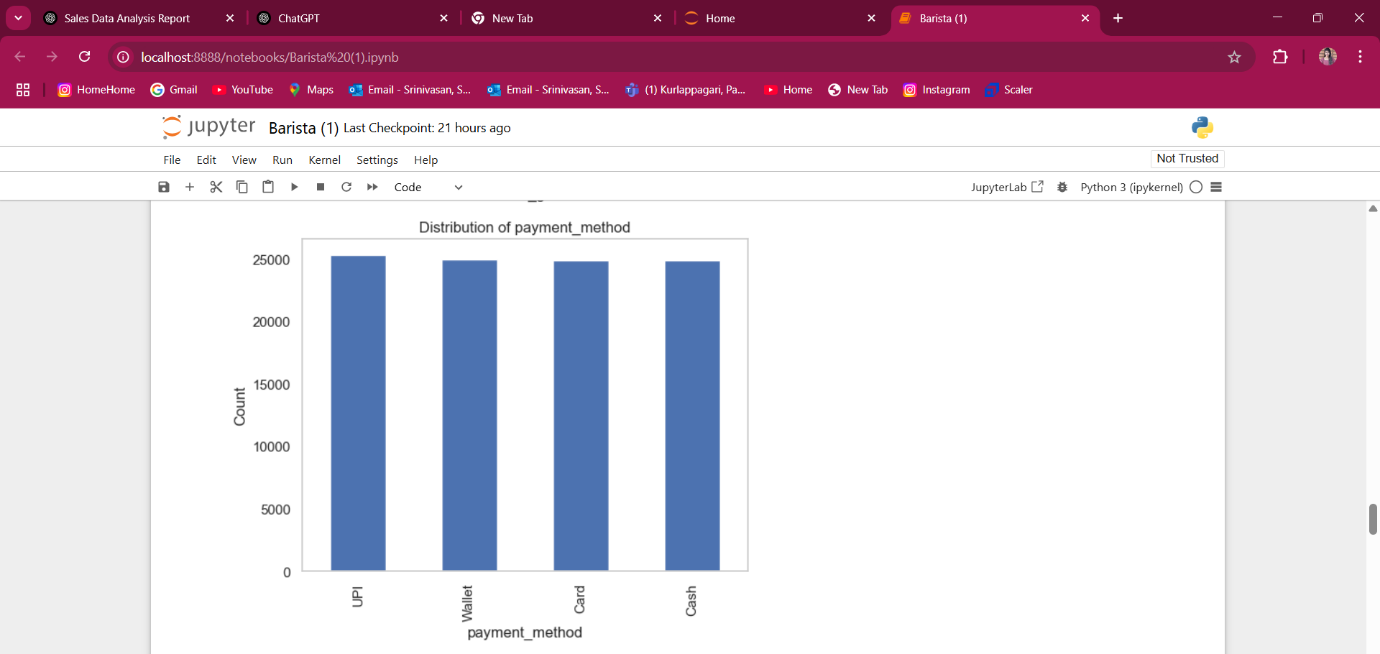


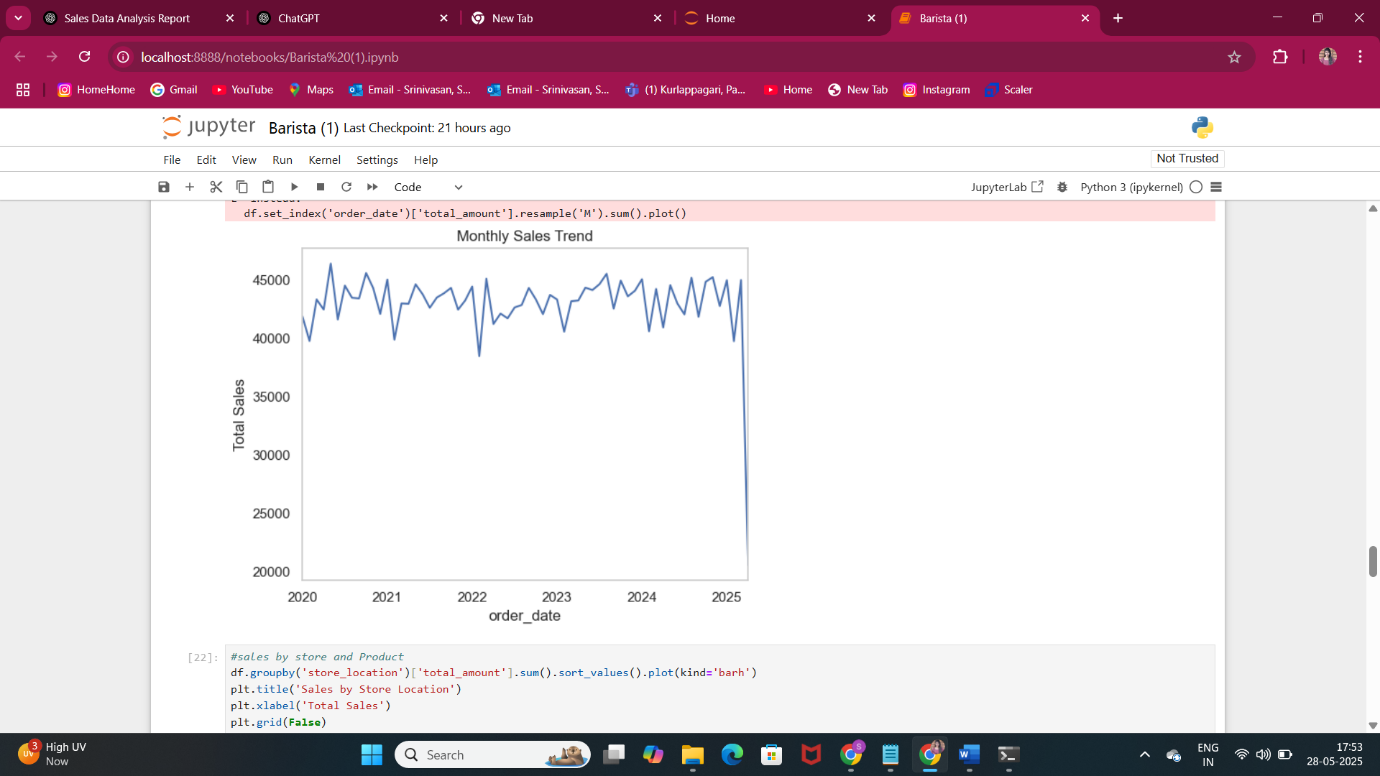


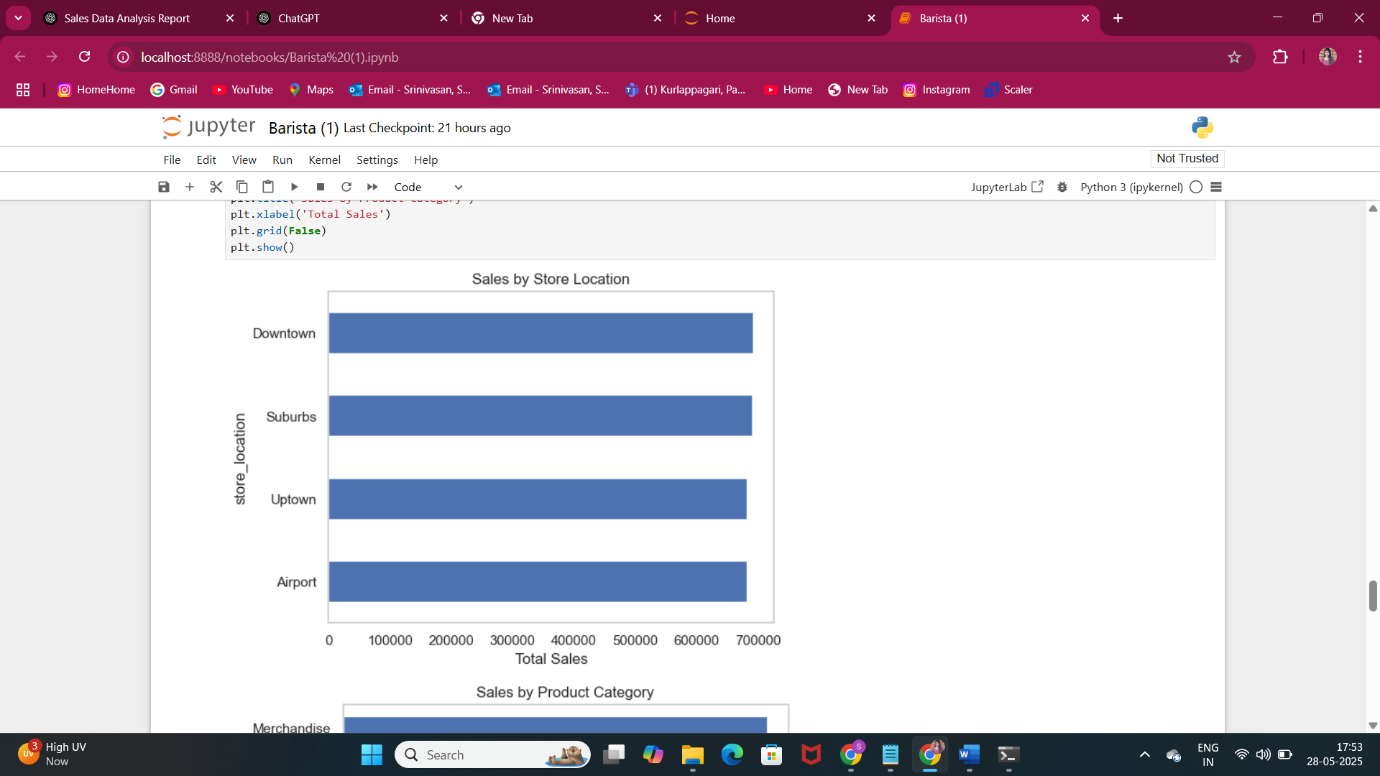


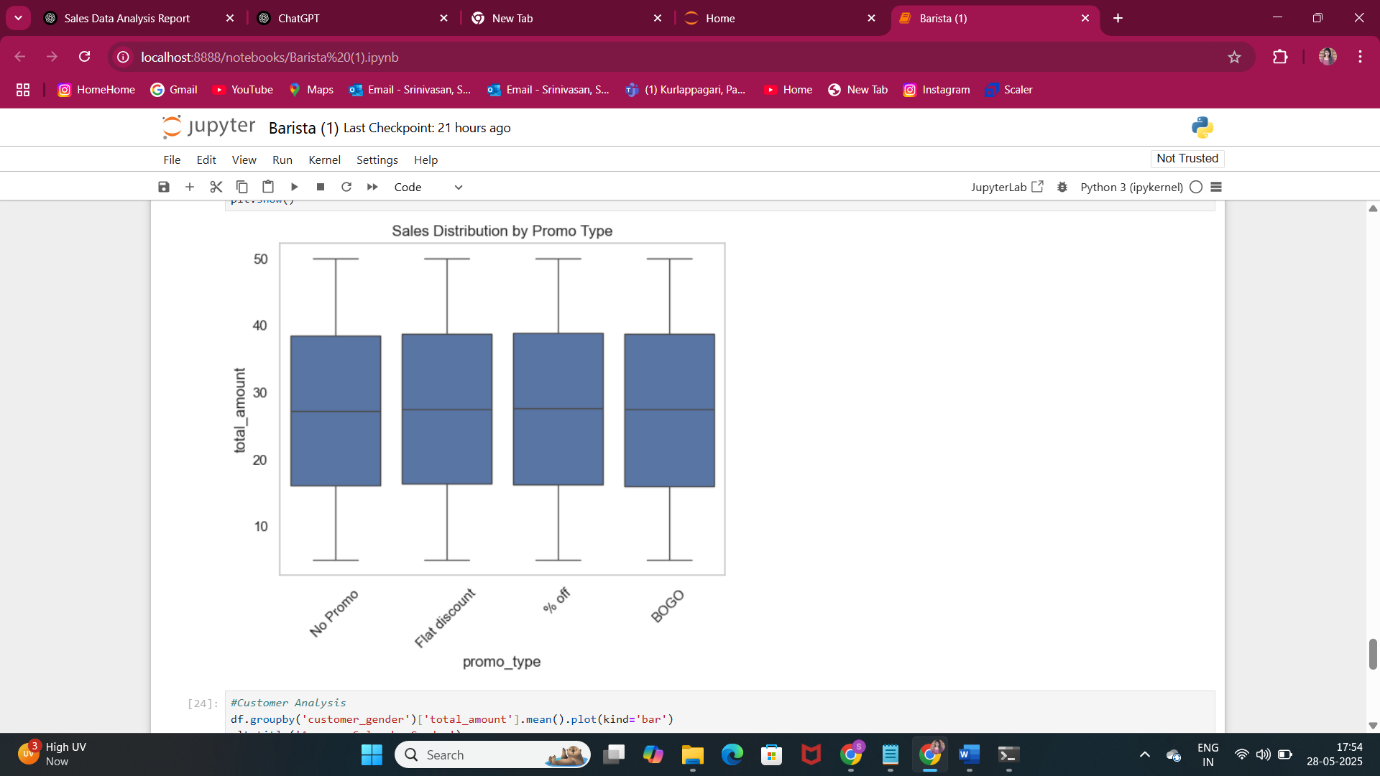








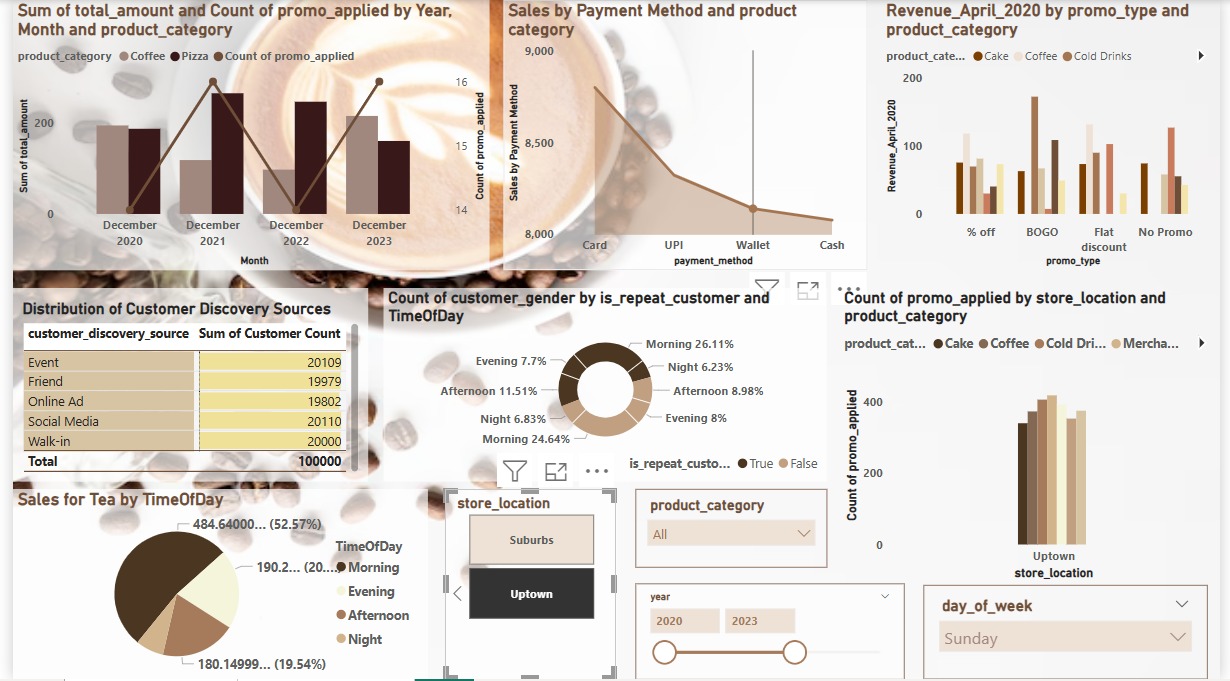


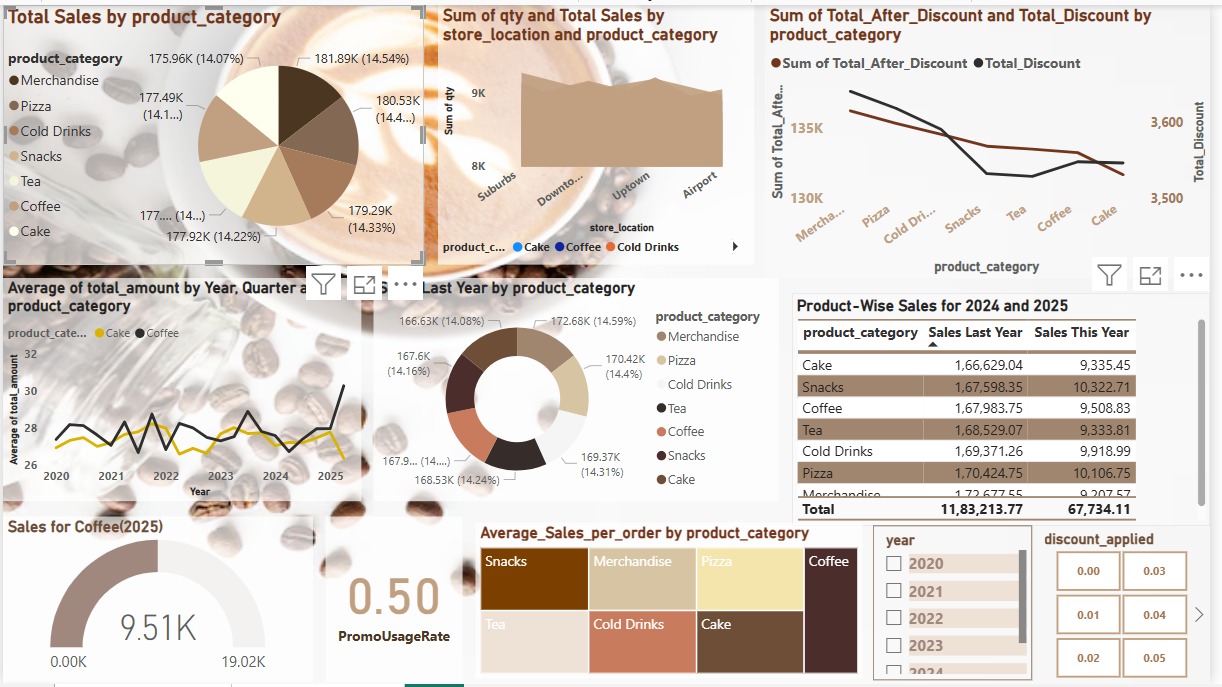


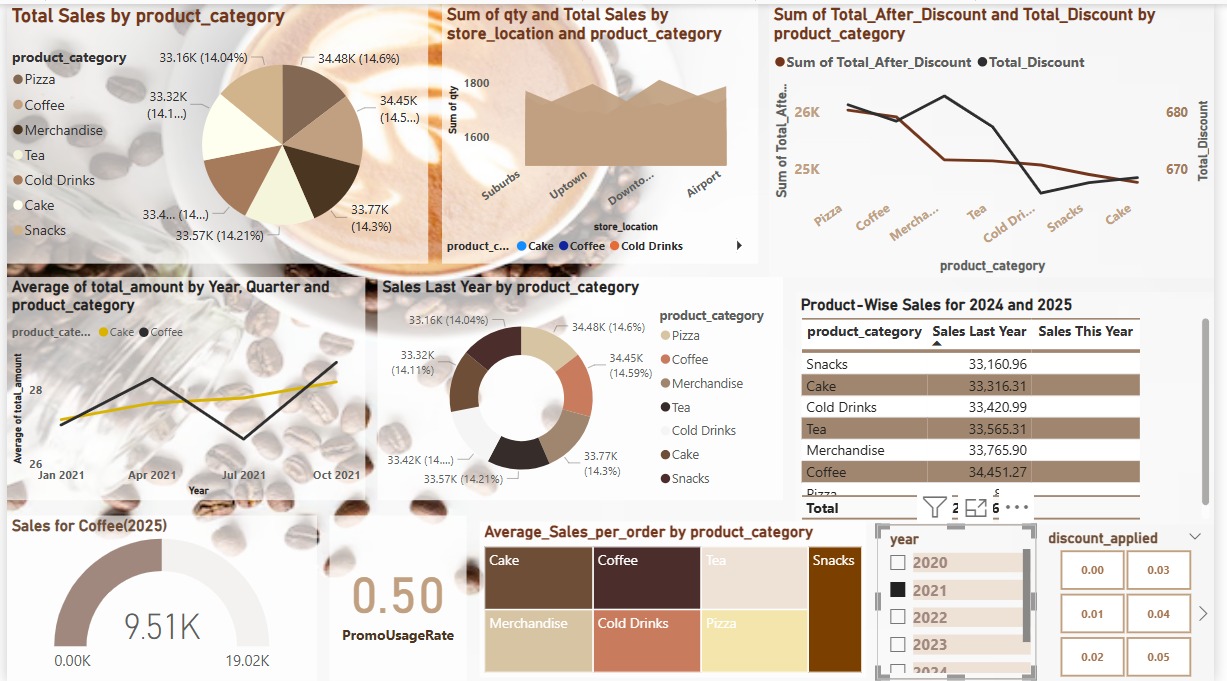
**Power Bi:**

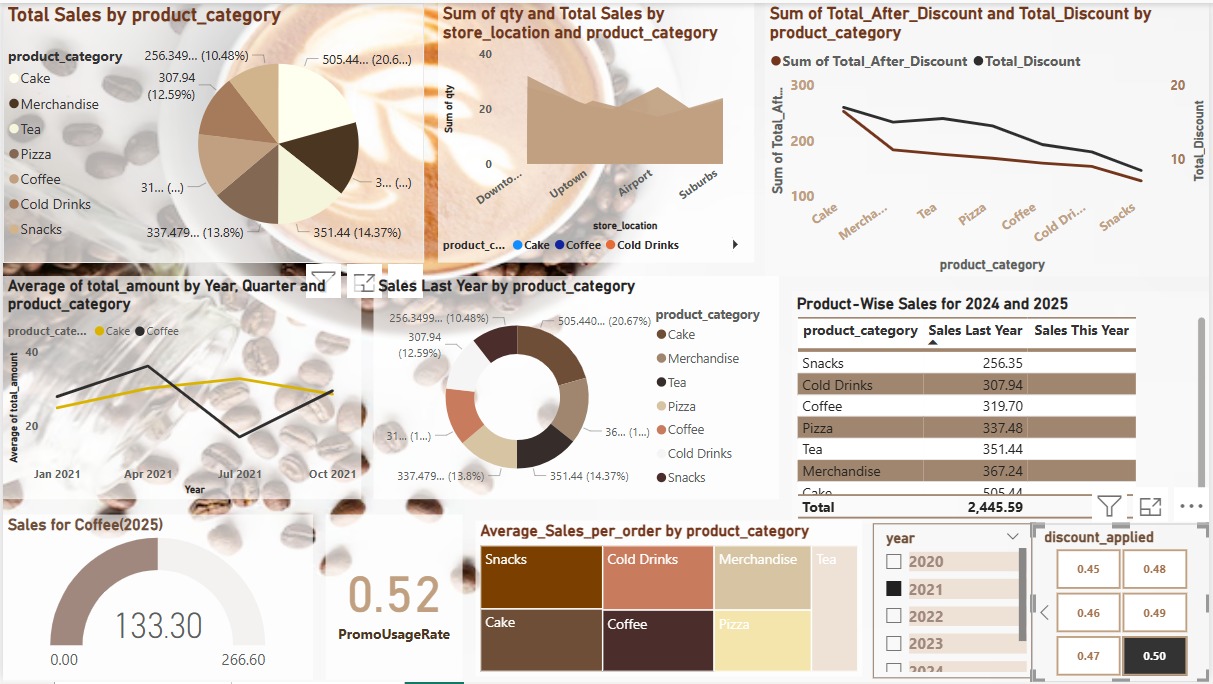












**12. Conclusion:**

This project demonstrated the power of data analytics in transforming simple sales logs into strategic business intelligence. Using a combination of tools, we were able to automate reporting, create interactive visualizations, and offer actionable insights. This solution can be scaled further for predictive analytics and integration with sales CRMs.

**13. References:**

* Kaggle Dataset: <https://www.kaggle.com/>
* Power BI Docs: <https://docs.microsoft.com/power-bi/>
* Python Pandas: https://pandas.pydata.org/
* Seaborn Visualization: https://seaborn.pydata.org/
* GitHub: <https://github.com/>