Retail Sales Data Analysis Report

1. Executive Summary

The objective of this project was to analyze retail sales data to uncover trends, customer behavior, and business insights. The dataset contained transactional details such as Transaction_ID, Date, Customer_ID, Gender, Age, Product Category, Quantity, Price per Unit, and Total Amount.

The project was executed using Excel, SQL, Python (Pandas, Matplotlib, Seaborn), and Power BI to demonstrate a complete data analytics workflow. This report summarizes data cleaning, transformation, exploratory analysis, and visualization tasks, followed by business insights and recommendations.

2. Dataset Description

The dataset used for analysis contains 1000+ records with the following columns:

- Transaction_ID
- Date
- Customer_ID
- Gender
- Age
- Product_Category
- Quantity
- Price_per_Unit
- Total_Amount
- Derived Columns: Age_Group, Sales_Classification, Year, Quarter, Month, Day_of_Week

3. Data Cleaning & Preparation

Key data cleaning and preparation tasks included :

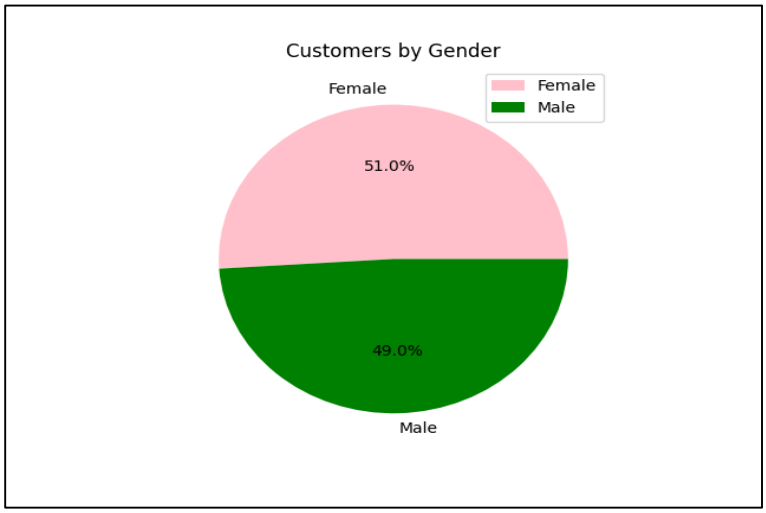
- Removed duplicate Transaction_IDs.
- Checked and handled NULL values.
- Converted Date column from text to proper Date format.
- Added calculated columns:
  - Age_Group (Young, Adult, Senior).
  - Sales_Classification (Low, Medium, High).
- Ensured proper data types using SQL DESCRIBE and Python dtypes().

4. Exploratory Data Analysis (Python)

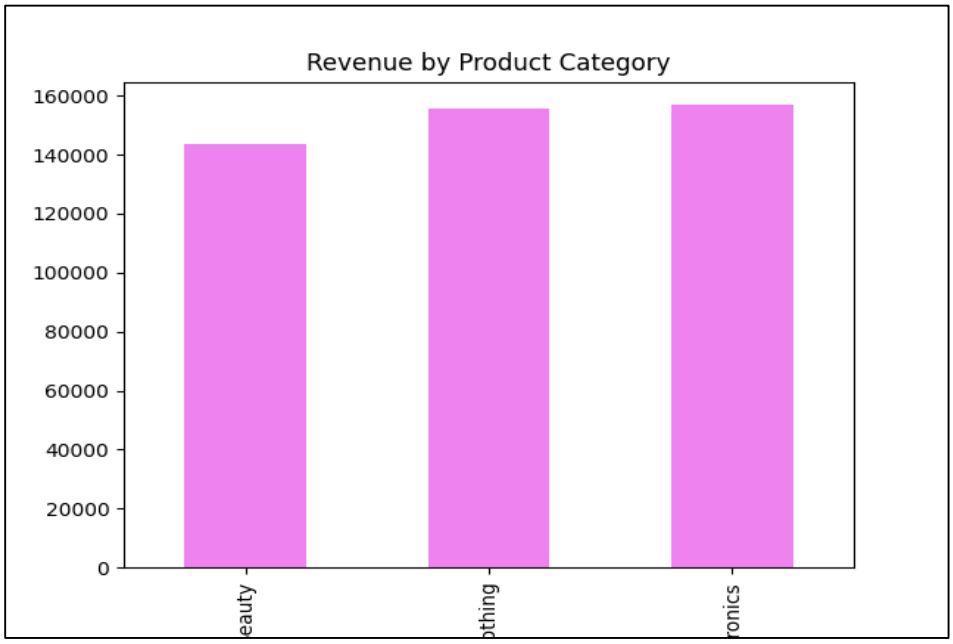
Exploratory data analysis was performed using Pandas, Matplotlib, and Seaborn. Charts and graphs were generated to identify sales trends, customer demographics, and product performance.

Key analyses included:

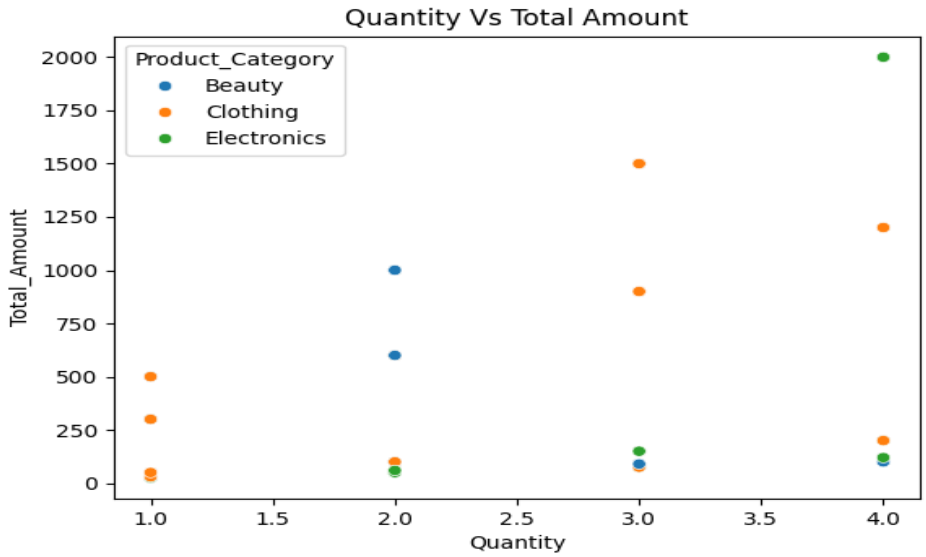
- Gender distribution of customers (51% female, 49% male).



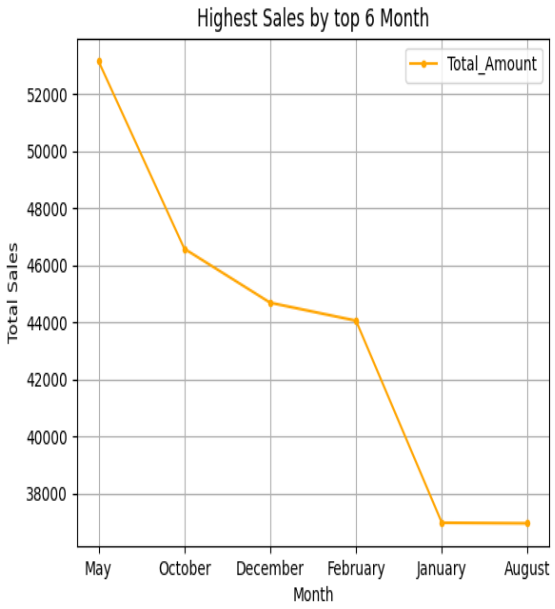
- Electronics category generated the highest revenue.

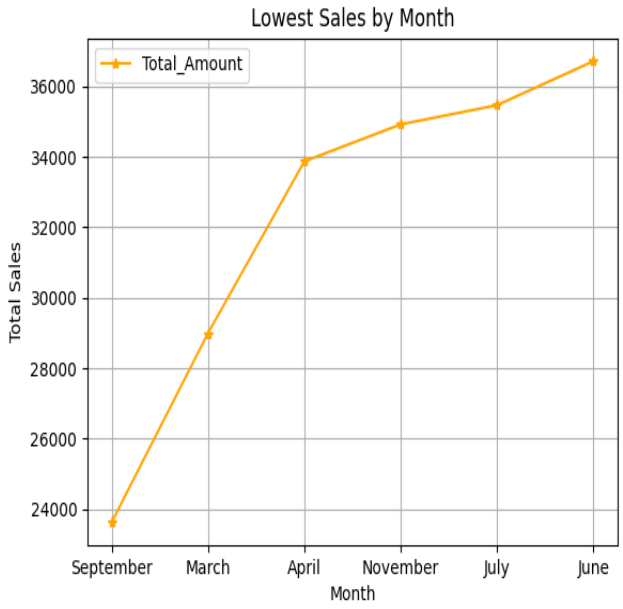


- Scatter plot of Quantity vs Total_Amount showed a positive correlation.



- Monthly sales analysis revealed seasonal peaks.

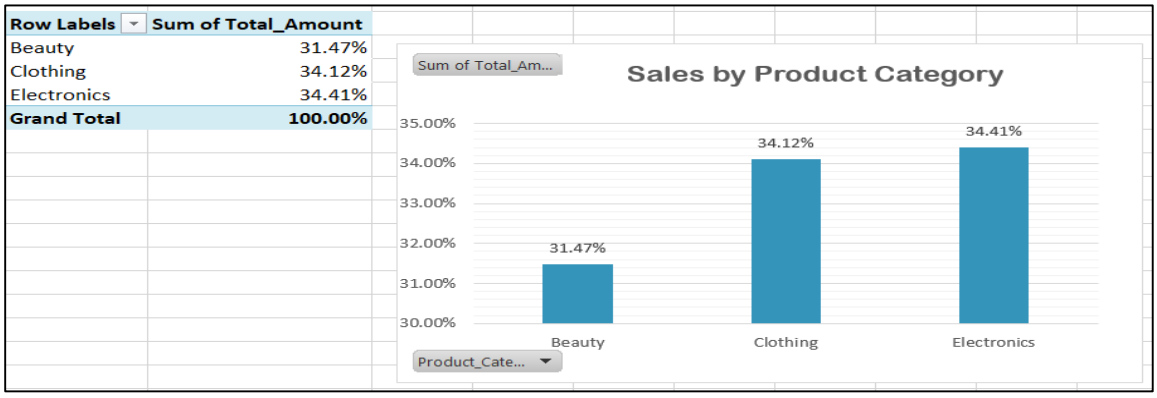




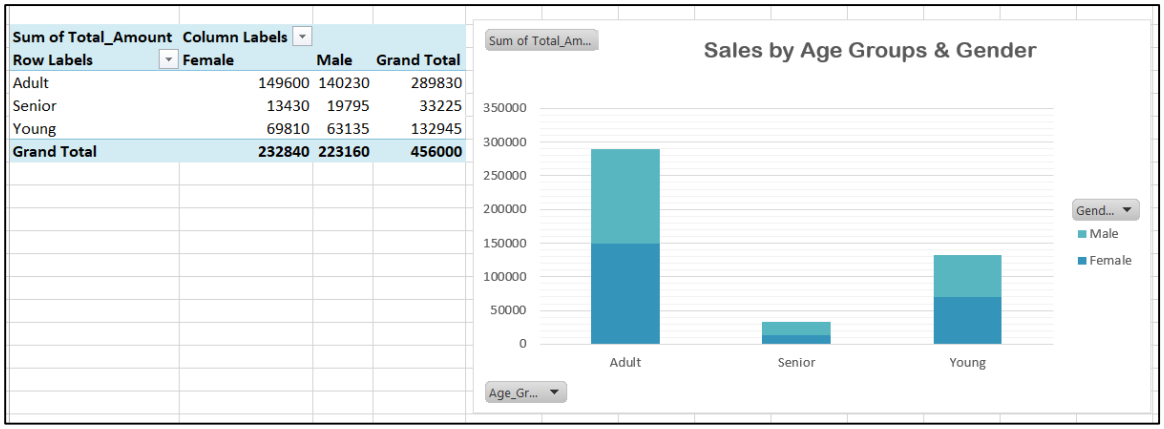
5. Excel Analysis

Excel was used for quick validation, transformations, and reporting:

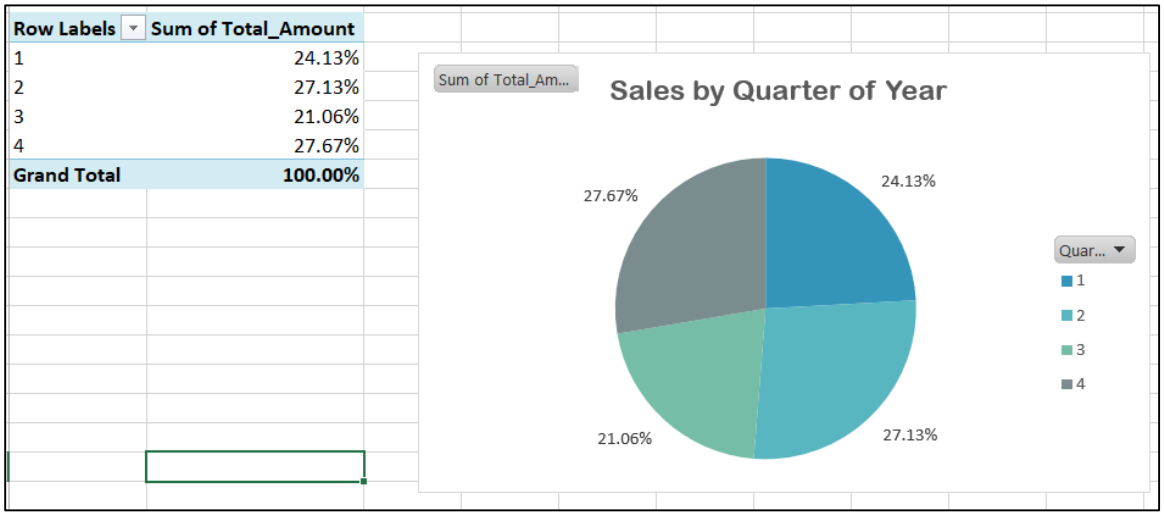
- VLOOKUP and INDEX-MATCH applied to map customer/product details.
- PivotTables created to analyze revenue by category and region.
- Conditional Formatting used to highlight low/negative sales.
- Charts generated to summarize sales trends.



Row Labels	Sum of Total_Amount
Beauty	31.47%
Clothing	34.12%
Electronics	34.41%
Grand Total	**100.00%**



Sum of Total_Amount	Column Labels		
Row Labels	Female	Male	Grand Total
Adult	149600	140230	289830
Senior	13430	19795	33225
Young	69810	63135	132945
Grand Total	**232840**	**223160**	**456000**

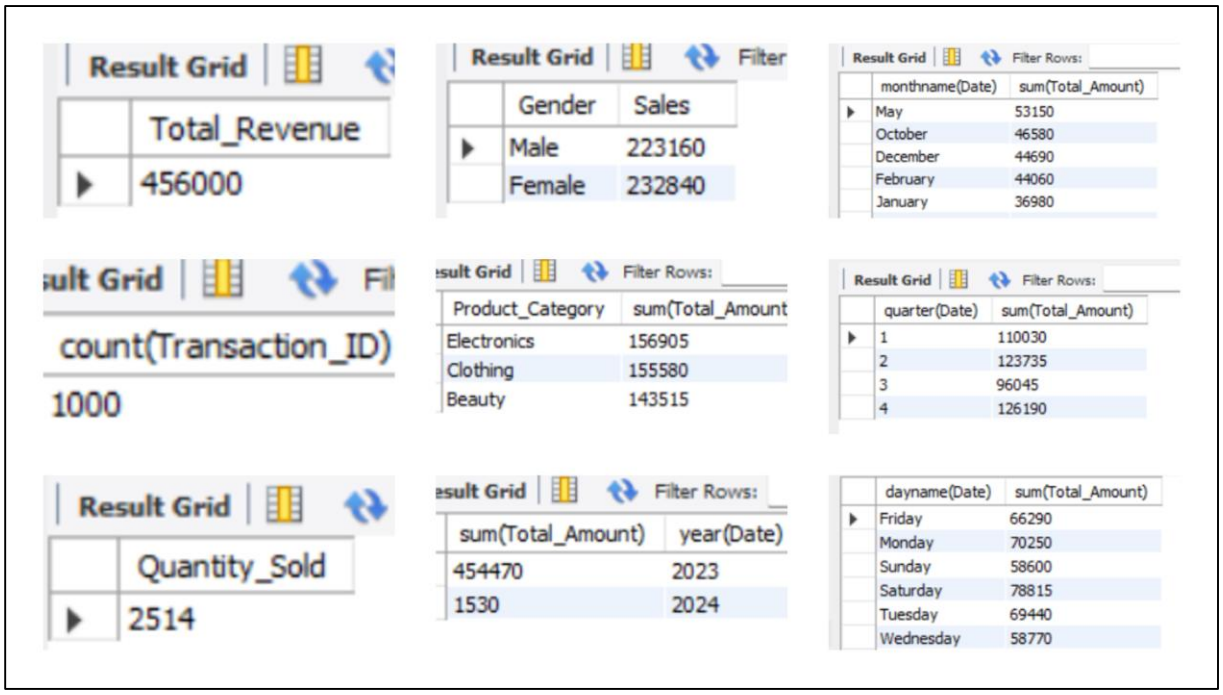


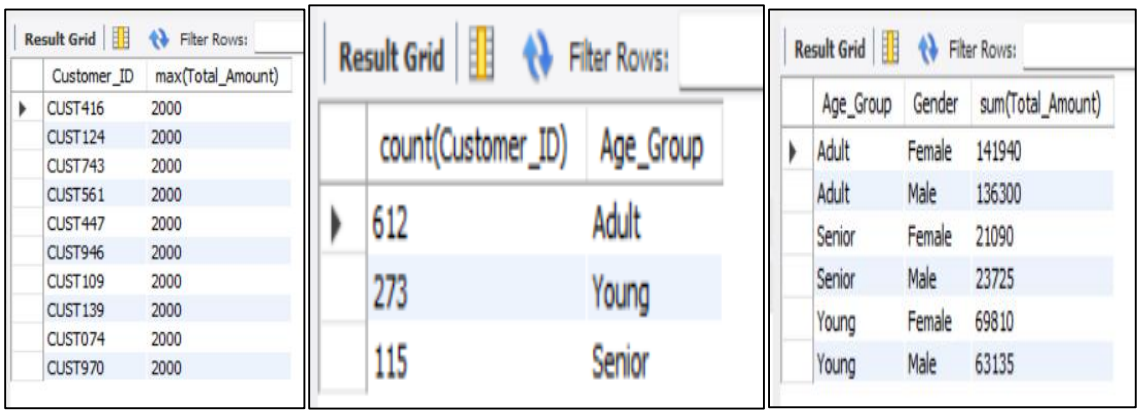
Row Labels	Sum of Total_Amount
1	24.13%
2	27.13%
3	21.06%
4	27.67%
Grand Total	**100.00%**

6. SQL Analysis

SQL queries were used for structured analysis and business insights:

- Total Revenue, Transactions, and Quantity Sold.
- Gender-based sales analysis.
- Sales performance by Product_Category.
- Time-based insights (Year, Month, Quarter, Day of Week).
- Top 10 customers by revenue.
- Customer segmentation by Age_Group and Gender.
- Festival season sales patterns (October-November).
- Business Insights Queries.

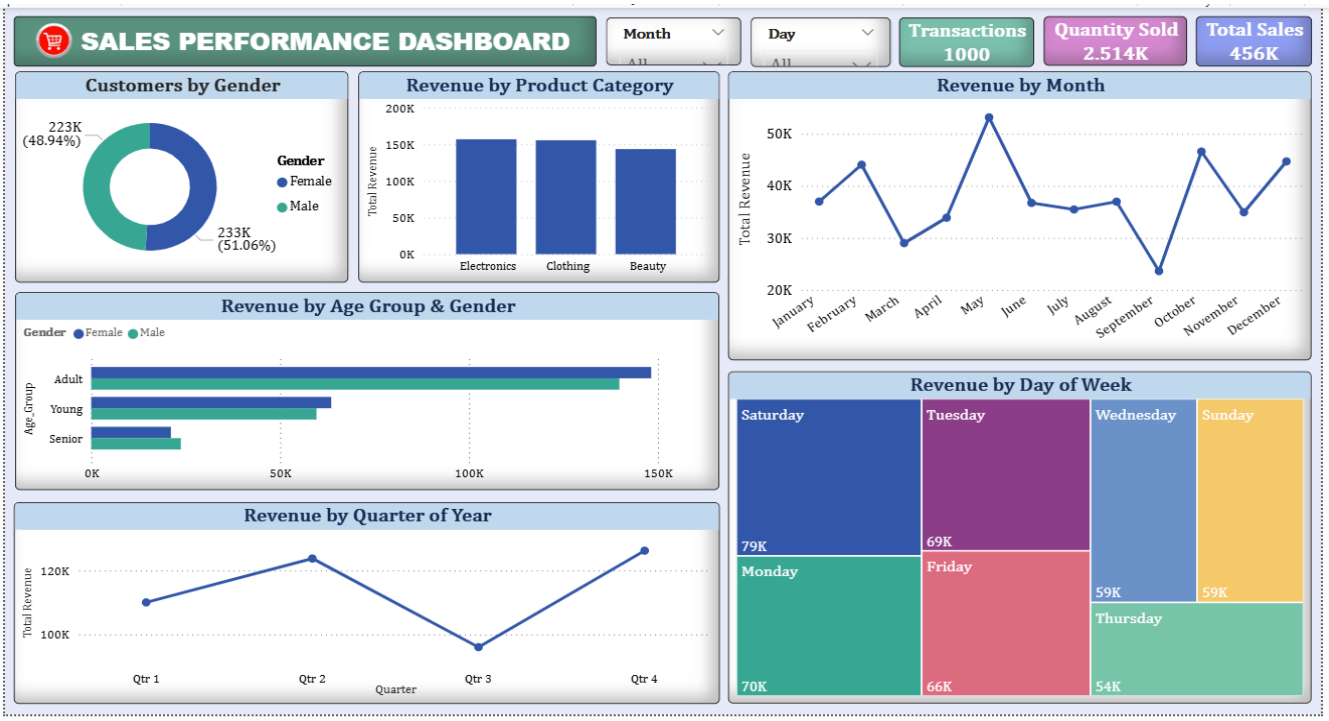




7. Power BI Dashboard

An interactive Power BI dashboard was created to present insights visually.

- The dashboard included:
- Sales by Region and Category.
- Revenue, Quantity Sold, Transactions KPIs.
- Customer segmentation visuals.
- Sales by Month & Day of Week.
- Sales by Quarter of Year.
- Customers by Gender.



8. Key Insights

- Majority of customers are Female (51%).
- Electronics category contributes the highest revenue.
- Customers buying more quantity tend to spend more overall.
- Seasonal peak sales observed during Q4.
- Adults are contributing approx. 55% of Revenue.
- Senior age group spends significantly on Electronics.
- Highest revenue generated during festival months for Clothing (October, November).
- Highest Revenue in all months is May showing significant milestone of Highest Sales.

9. Conclusion

This project showcased end-to-end data analytics skills across multiple tools (Excel, SQL, Python, Power BI).

By cleaning, transforming, and analyzing sales data, key business insights were uncovered, such as customer demographics, product performance, and seasonal trends.

The findings can support business decision-making in marketing, inventory planning, and customer targeting.