Report for eCommerce Transactions Dataset

Introduction :

The dataset consists of three files: Customers.csv, Products.csv, and Transactions.csv, which provide valuable insights into customer behavior, product information, and transaction details. The primary goal of this analysis was to perform exploratory data analysis (EDA) to identify key patterns, trends, and actionable business insights. The dataset was preprocessed to ensure there were no missing values, and appropriate corrections were made to the data types for accurate analysis. Afterward, the three files were merged to create a comprehensive dataset, enabling a detailed exploration of customer purchasing patterns and product performance.

Data Cleaning and Preprocessing :

The dataset was first inspected for missing values, but no missing data was found, eliminating the need for data imputation or removal. The next step involved correcting the data types for some columns to ensure consistency:

- SignupDate and TransactionDate columns were converted to datetime format.
- Certain columns, such as Price, were verified and retained as float64 data types.

Once the data was cleaned, the three individual datasets (Customers.csv, Products.csv, Transactions.csv) were merged using the common columns CustomerID and ProductID. This provided a unified dataset, which allowed for a more detailed analysis of transaction details, customer behavior, and product performance.

Exploratory Data Analysis (EDA) and Insights :

After merging the datasets, several important insights were derived through the analysis:

- **High Sales Products**:
  By aggregating total sales per product, we identified the products with the highest revenue. The top-selling products were those with both high sales volume and high price points. This insight can guide future inventory management and marketing strategies, focusing on promoting these high-demand products.

Top 10 High Sales Products
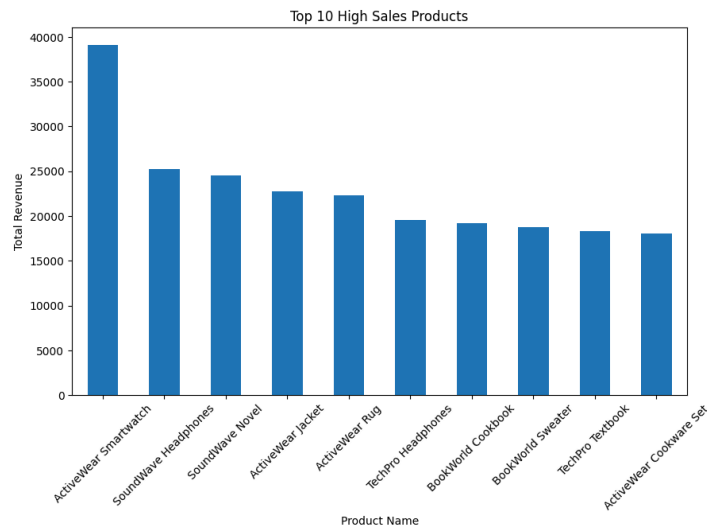
- **Sales Distribution Over Regions:**
  The sales distribution across regions highlighted significant differences in performance. Some regions contributed more to the total sales, indicating either higher customer demand or more effective marketing efforts in those areas. Conversely, regions with lower sales suggest the need for targeted promotions or localized strategies. Additionally, regions with high-value sales but low quantities sold could benefit from strategies to increase both the volume and revenue, such as offering targeted deals or promoting higher-ticket products more effectively.


Sales Distribution by Region
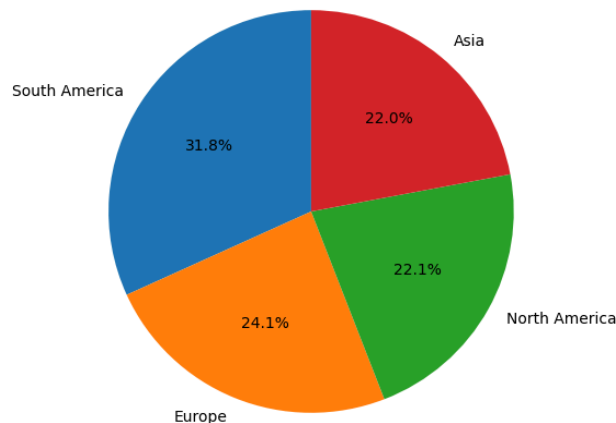
- **Top 10 Customers' Monthly Purchases**:
  An analysis of the top 10 customers showed distinct patterns in their monthly purchasing behavior. These customers contributed significantly to the total sales, especially during specific months. Targeted promotional efforts could help further engage these top customers and boost their purchasing activity in the future.

Monthly Purchases Heatmap of Top 10 Customers
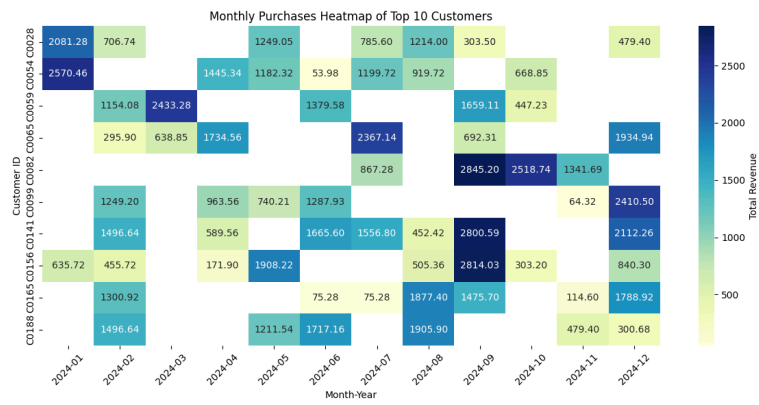
● **Category-Based Purchases**:
By analyzing sales by product category, it became clear that certain categories, such as electronics, generated higher sales in terms of value, while categories like clothing saw higher volume in terms of units sold. This insight helps in understanding consumer preferences and allows the business to adjust product marketing strategies accordingly.


Revenue by Category

Units Purchased by Category
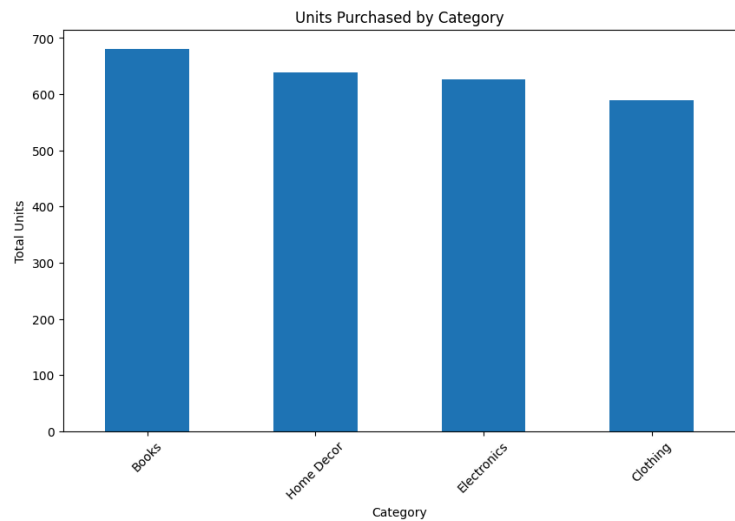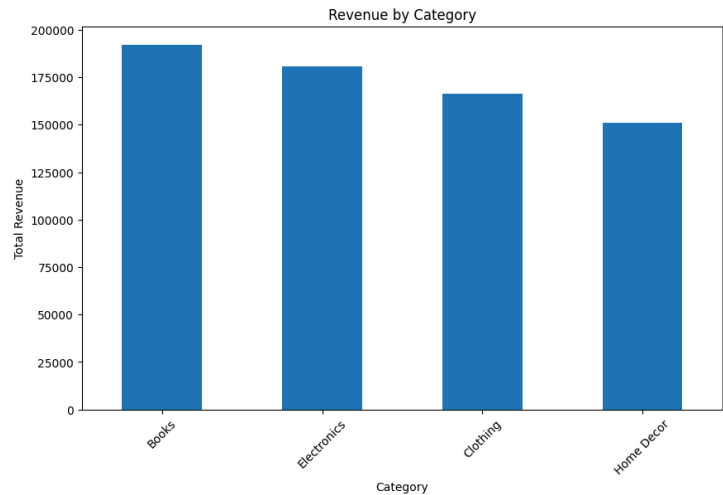
- **Purchasing Trends by Region**:
  Regional analysis revealed that certain products were more popular in specific regions. For example, certain electronic products were more frequently purchased in urban regions, suggesting that regional marketing campaigns could be tailored to meet these specific demands.

Total Sales Trends Per Region
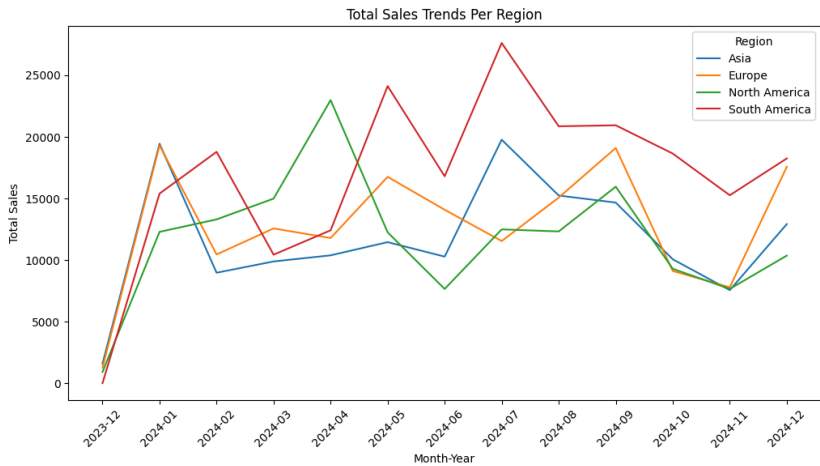
- **Average Customer Spend**:
  The average purchase value per customer was calculated, revealing that a small group of customers made up the majority of the total sales. This insight is valuable for customer segmentation and can inform strategies to retain high-value customers while nurturing lower-spending customers.

Average Purchase per Customer per Month
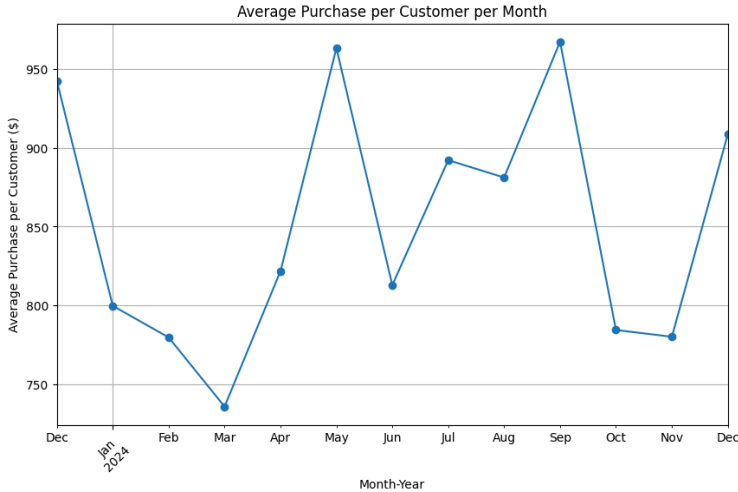
Conclusion :

The EDA provided several valuable insights into customer purchasing behavior, product sales, and regional preferences. By focusing on high-performing products, understanding purchasing patterns, and tailoring marketing campaigns to specific regions and customer segments, the company can improve sales performance and customer engagement. These insights are essential for making data-driven business decisions and enhancing the company's overall strategy.