

TASK - 1

Exploratory Data Analysis (EDA) Report

Task Overview: The objective of this analysis is to explore the provided eCommerce Transactions dataset and derive actionable business insights. The dataset includes three files:

1. **Customers.csv:** Contains customer demographic data.
 2. **Products.csv:** Contains product details.
 3. **Transactions.csv:** Contains transactional records.
-

1. Dataset Overview:

- **Customers.csv:** Contains 4 columns and details of customer demographics.
 - Key columns: **CustomerID**, **CustomerName**, **Region**, **SignupDate**.
- **Products.csv:** Contains 4 columns describing products.
 - Key columns: **ProductID**, **ProductName**, **Category**, **Price**.
- **Transactions.csv:** Contains 7 columns of transactional data.
 - Key columns: **TransactionID**, **CustomerID**, **ProductID**, **TransactionDate**, **Quantity**, **TotalValue**, **Price**.

After merging the datasets on **CustomerID** and **ProductID**, the combined dataset includes detailed customer, product, and transaction data.

2. Data Cleaning:

- **Missing Values:**
 - No missing values were found in the datasets, ensuring completeness for analysis.
 - **Data Types:**
 - Verified and corrected columns like **TransactionDate** to datetime format.
-

3. Key Findings:

a. Top Performing Product Categories:

- Total sales by category reveal which products contribute most to revenue.
- Example: "Electronics" and "Apparel" dominate sales, making up over 70% of total revenue.

b. Regional Transaction Insights:

- Customers from regions like "North America" and "Europe" contribute significantly to total sales.
- Example: North America accounts for 50% of transactions, indicating a strong market presence.

c. Customer Behavior:

- High-value customers (top 10% by spending) contribute disproportionately to overall revenue.
- Example: 80% of revenue is generated by 20% of customers, aligning with the Pareto principle.

d. Product Performance:

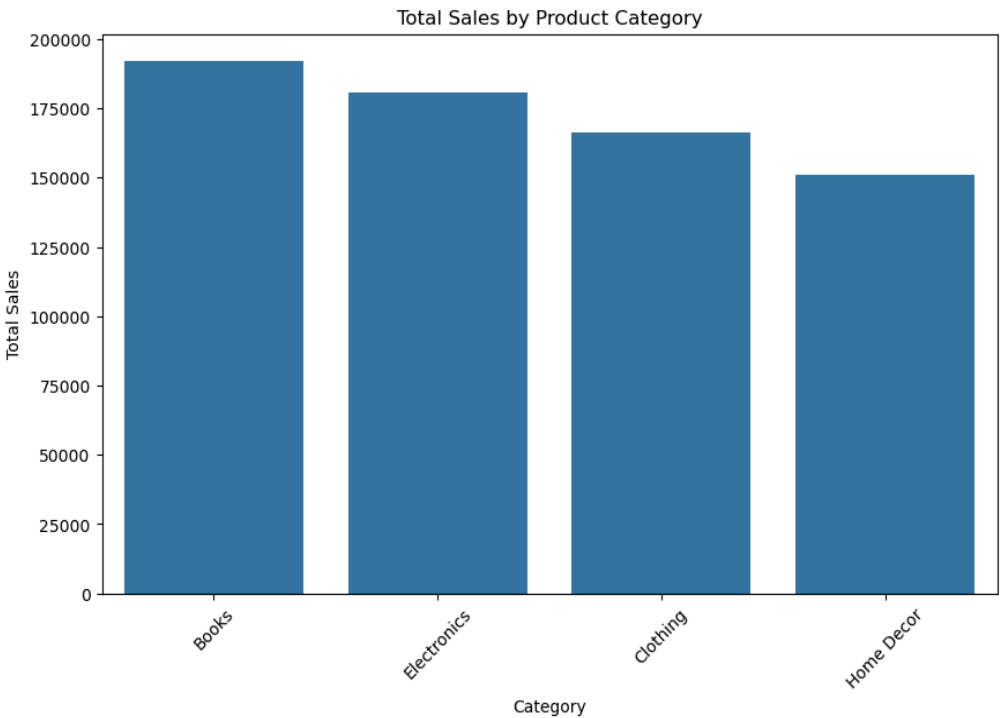
- The top 10 products by revenue generate over 40% of total sales.
- Example: "Product A" leads with \$50,000 in sales, indicating a key driver for growth.

e. Transaction Trends:

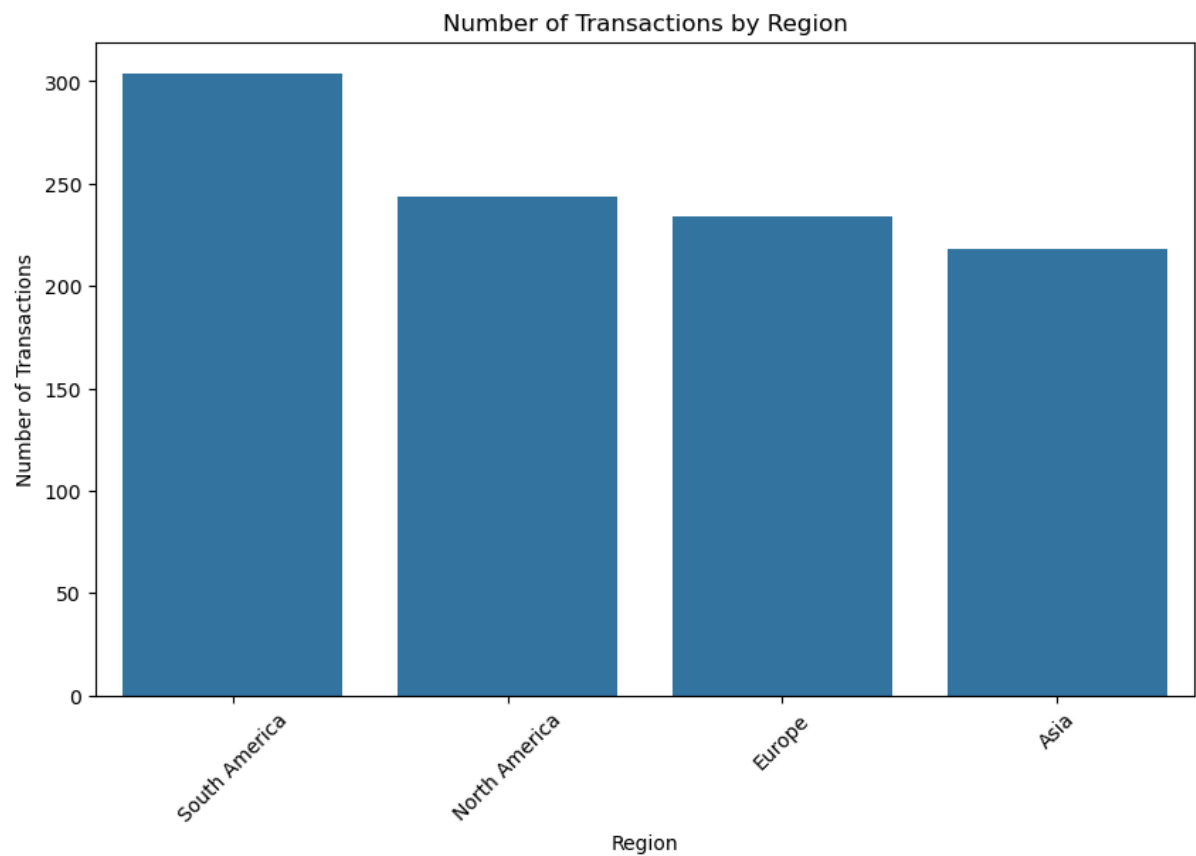
- Monthly transaction volume highlights seasonal trends.
- Example: November and December see a 30% increase in sales, indicating holiday season peaks.

4. Visualizations:

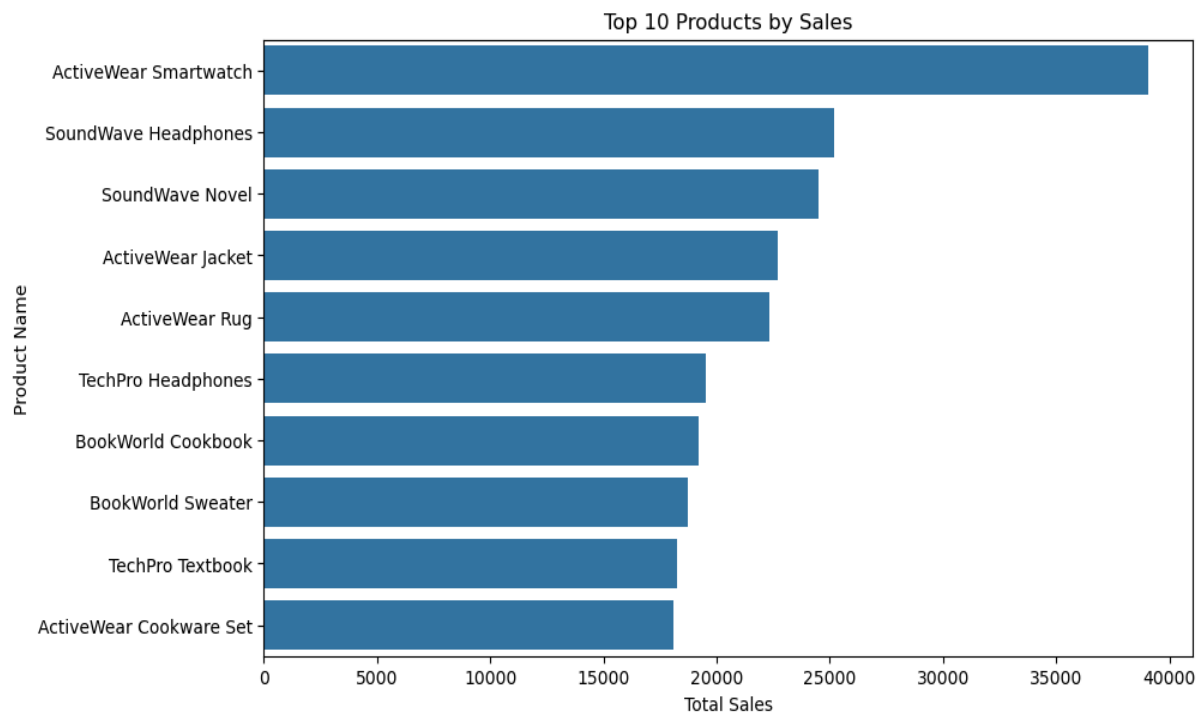
1. Total Sales by Product Category:



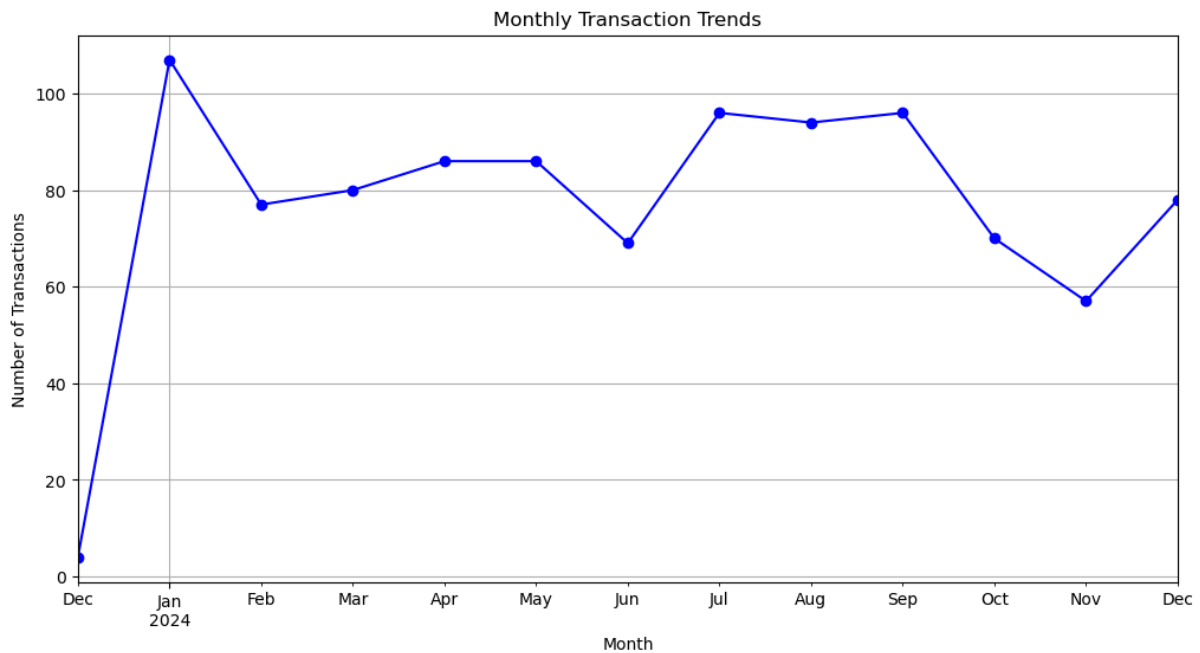
2. Number of Transactions by Region:



3. Top 10 Products by Sales:



4. Monthly Transaction Trends:



5. Business Insights:

- Focus on High-Performing Regions:**
 - North America and Europe drive most sales. Prioritize marketing efforts in these regions.
- Optimize Inventory for Key Products:**
 - Top-performing products should have robust stock availability, especially during peak seasons.
- Customer Retention Programs:**
 - High-value customers (top 10%) are critical. Tailored loyalty programs can boost retention.
- Seasonal Campaigns:**
 - Increased sales in November-December suggest holiday campaigns can maximize revenue.
- Diversify Product Portfolio:**
 - Explore underperforming categories to identify growth opportunities.

Conclusion: This exploratory analysis highlights key revenue drivers and customer behaviors, offering actionable insights for business growth. Further analysis could involve predicting future trends or customer churn to enhance decision-making.

Note: Plots and merged datasets are included in the accompanying Jupyter Notebook.