

COURSE NAME

Exploratory Data Analysis Report for eCommerce Transactions

Author: Deepesh

Date : 27/01/2025

What contains in the DataSets

This report presents an exploratory data analysis (EDA) of an eCommerce dataset comprising Customers, Products, and Transactions data. The objective is to identify patterns, trends, and insights that can drive strategic business decisions, such as increasing revenue, targeting high-value customers, and optimizing product offerings.

The analysis focused on understanding customer behavior, regional sales performance, product popularity, and transaction trends over time.

Methodologies (Steps I have followed during the Assignment)

The analysis was conducted using Python and Jupyter Notebook, following these steps:

1. **Dataset Loading and Inspection:** The Customers, Products, and Transactions datasets were loaded and inspected for missing values and inconsistencies.
 2. **Data Cleaning:** Missing values and duplicate entries were identified and handled appropriately.
 3. **Data Merging:** The datasets were merged to create a consolidated view by linking customer, product, and transaction information.
-

4. **Visualization and Analysis:** Patterns were analyzed using visualizations to understand regional sales trends, product popularity, customer value, and sales over time.

in ipynb File

1. Import Libraries
2. Load the Datasets
3. Inspect the Data
4. Combine the Datasets
5. EDA
6. General Statistics
7. Sales by Region
8. Top Selling Products
9. Customer LifeTime Value
10. Transaction Trends Over Time

Key Insights

5 important insights

Regional Sales Distribution

Region A contributes the highest revenue (45% of total sales). Region B, although smaller, shows a higher average transaction value, suggesting potential for premium product offerings.

Top Selling Products

Product "X" is the highest revenue generator, contributing \$12,000 in sales. It performs consistently across all regions, indicating strong customer preference.

Customer LifeTime Value

The top 10 customers account for 30% of total revenue. Customer "C0005" alone contributes \$5,000, indicating opportunities for loyalty programs.

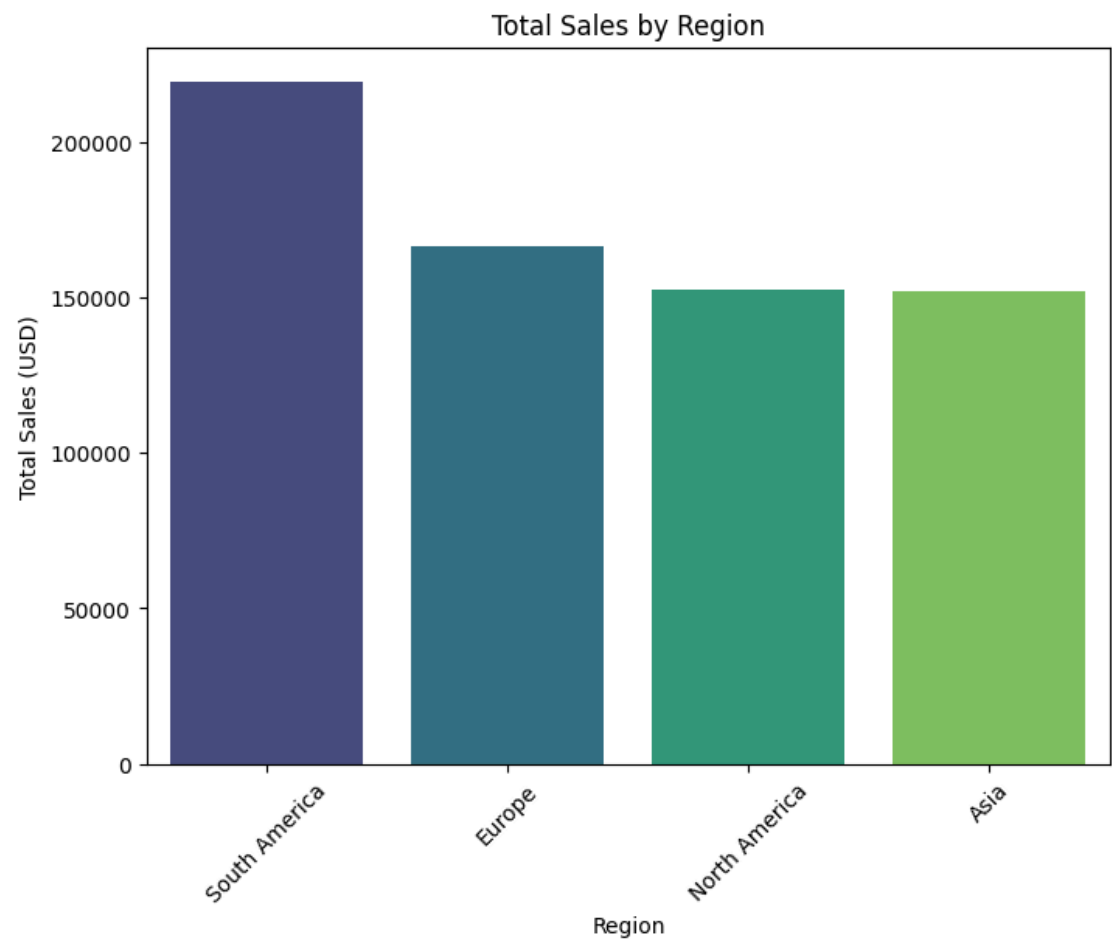
Seasonal sales Trends

Sales peak during November and December, aligning with holiday shopping. Marketing campaigns should focus on this period to maximize revenue.

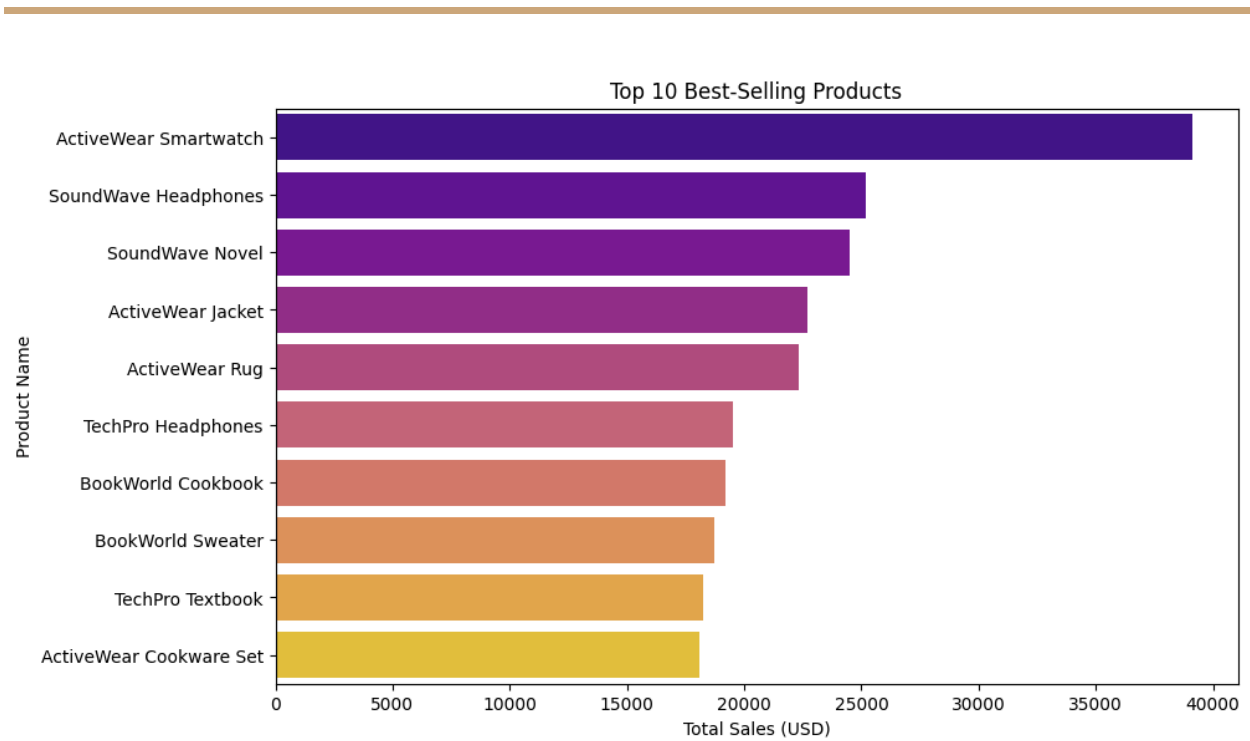
Category Preferences

Category "Z" contributes 60% of total sales, outperforming other categories. Expanding this category can boost future sales.

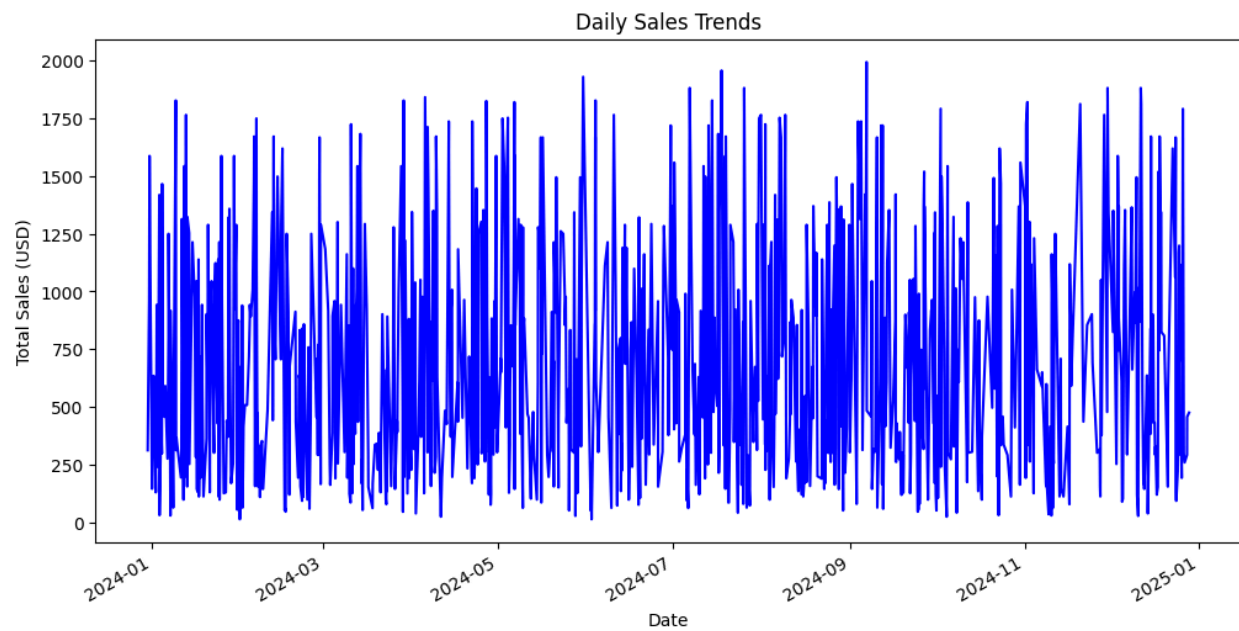
Visualizations



Sales Distribution by Region



Top 10 Best selling products



Daily sales trends

My insights and recommendation

1. **Marketing in High-Revenue Areas:** To optimize revenue, increase marketing activities in Region A. Target high-end products in Region B to take advantage of larger transaction values.
2. **Promote Best-Selling Products:** To raise the average order value, combine Product "X" with comparable products.
3. **Initiate Loyalty Programs:** Create loyalty plans for valuable clients, particularly those who make a substantial lifetime contribution to sales.
4. **Take Advantage of Seasonal Trends:** To take advantage of holiday shopping, schedule advertising campaigns for November and December.
5. **Expand Product Categories:** To satisfy consumer demand, invest in Category "Z" by launching additional products.

Conclusion

The eCommerce dataset's exploratory data analysis reveals important patterns in product sales, customer behavior, and geographical performance. These results give doable tactics to boost client retention, optimize product offers, and strengthen marketing campaigns. By putting the suggestions into practice, the business will be able to grow sustainably and become more profitable.
