EDA Report

1. Introduction

This report presents the exploratory data analysis (EDA) of an eCommerce Transactions dataset comprising Customers, Products, and Transactions data. The objective is to understand the data structure, uncover trends, and derive actionable business insights. The analysis focuses on customer demographics, product sales, and transaction patterns.

---

2. Data Understanding

- **Dataset Overview:**

    o **Customers.csv: Contains 4 columns (CustomerID, CustomerName, Region, SignupDate) and 200 records.**

    o **Products.csv: Contains 4 columns (ProductID, ProductName, Category, Price) and 100 records.**

    o **Transactions.csv: Contains 7 columns (TransactionID, CustomerID, ProductID, TransactionDate, Quantity, TotalValue, Price) and 1000 records.**

- **Missing Values:**

    o **No missing values in Customers, Products, or Transactions datasets.**

- **Duplicate Records:**

    o **No duplicate records found in any dataset.**

- **Key Metrics:**

    o **Total Transactions: 1000**

    o **Unique Customers: 200**

    o **Unique Products: 100**

---

3. Exploratory Analysis

Customer Analysis:

- **Customers are distributed across regions, with South America generating the highest total revenue of $219,352.56.**

- **Customer signup trends show steady growth over time.**

Product Analysis:

- **Products span multiple categories, with Books and Electronics contributing significantly to sales.**

- **Product prices range from $16.08 to $497.76, with an average price of $272.55.**

Transaction Analysis:

- South America generates the highest revenue ($219,352.56), followed by Europe ($166,254.63).

- The top 10 products account for a significant portion of total quantity sold.

Visualization Highlights:

1. Region-wise Sales: A bar chart highlights regional contributions to revenue.

2. Category-wise Sales: A bar chart shows sales by product category.

3. Top Products: A bar chart identifies the most sold products.

4. Sales Trend Over Time: A line chart illustrates revenue trends by month.

5. Correlation Heatmap: Displays relationships between numerical features (e.g., TotalValue, Quantity).

---

4. Insights

1. **South America contributes over 40% of total revenue:** This highlights the importance of focusing marketing efforts in this region while exploring untapped areas for growth.

2. **Top 10 products drive a significant portion of sales:** Expanding the product portfolio and reducing reliance on a few products could mitigate risk.

3. **Books and Electronics dominate sales:** Increasing offerings in these categories could further boost revenue.

4. **Sales trends show steady growth:** Indicates effective customer acquisition and retention strategies.

5. **High correlation between Quantity and TotalValue:** Suggests bulk-purchase promotions could maximize revenue.

---

5. Conclusion

The analysis provides actionable insights such as targeting high-performing regions, optimizing inventory for top-selling products, and leveraging sales trends for strategic planning. These findings can guide the company in enhancing sales, customer acquisition, and overall business performance.