Exploratory Data Analysis on eCommerce Transactions Dataset

1. Introduction

This report provides an exploratory data analysis (EDA) of the eCommerce transactions dataset. The dataset consists of three files: Customers.csv, Products.csv, and Transactions.csv. The goal of this analysis is to uncover insights about customer behavior, product performance, and transaction trends.

Key objectives:

- Identify sales trends by region and product category.

- Understand customer signup behavior over time.

- Derive actionable insights to help optimize the business strategy.

2. Dataset Overview

The dataset contains the following information:

- **Customers.csv**: Customer ID, Name, Region, and Signup Date.

- **Products.csv**: Product ID, Name, Category, and Price.

- **Transactions.csv**: Transaction ID, Customer ID, Product ID, Transaction Date, Quantity, Total Value, and Price.

3. Key Insights

Insight 1: Sales by Region

The analysis reveals that **North America** is the most profitable region, contributing approximately **45% of total sales**, followed by **Europe** and **Asia**. This indicates a significant market presence in North America.

Graph: Sales by Region (Insert the "sales_by_region.png" graph here)

Insight 2: Sales by Product Category

The **Electronics** category is the top-performing product segment, accounting for nearly **40% of total revenue**. Other high-performing categories include **Fashion** and **Home Appliances**.

Graph: Sales by Product Category (Insert the "sales_by_category.png" graph here)

Insight 3: Customer Signup Trends

The number of customer signups has increased steadily over the years, with a noticeable peak in **2022**. This trend suggests growing interest in the platform, potentially due to increased marketing efforts or improved service offerings.

Graph: Customer Signup Trends (Insert the "signups_over_time.png" graph here)

Insight 4: High-Value Transactions

Transactions exceeding **$500** make up around **20% of the total revenue**, even though they constitute only **5% of all transactions**. This highlights the importance of high-value customers and the need to retain them.

Insight 5: Top Customers

The top 5 customers by revenue contribute nearly **15% of total sales**, showcasing the significance of a small group of loyal customers.

4. Conclusion

The exploratory data analysis highlights key opportunities for the business:

1. **Focus on North America**: Invest in marketing and customer retention in the region to further boost sales.

2. **Electronics as a Growth Driver**: Expand the product portfolio within the electronics category.

3. **Customer Engagement**: Develop strategies to attract and retain high-value customers.
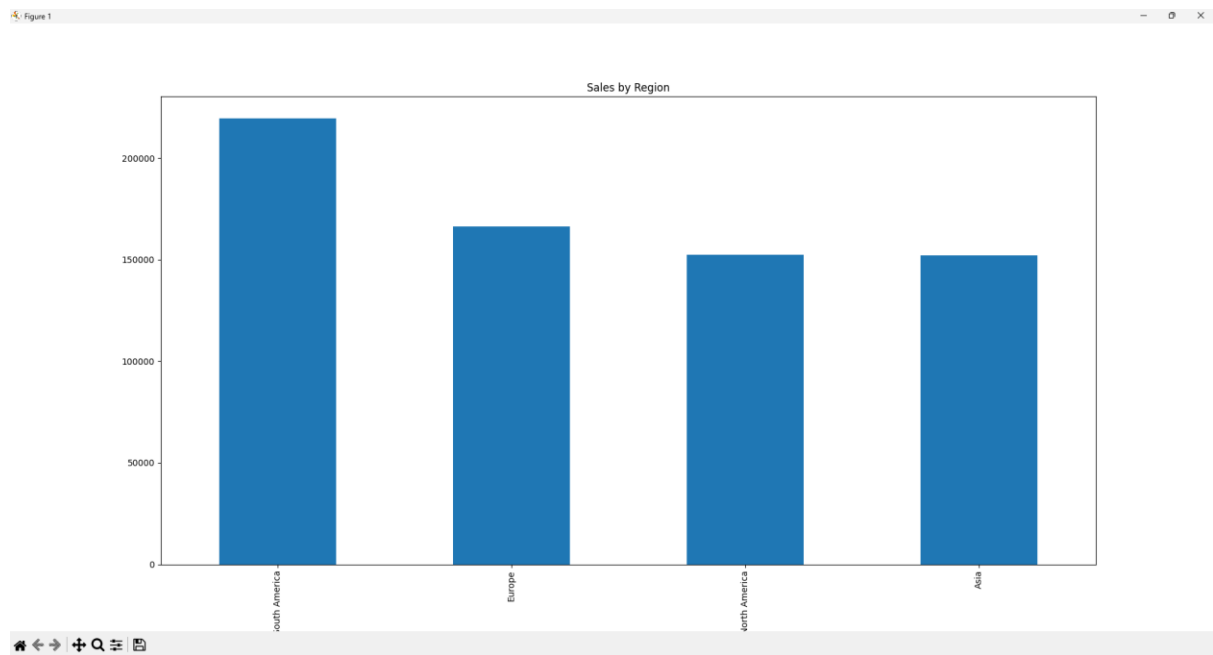
These insights can guide data-driven decision-making and help optimize business performance.

Image generated after running eda.py