ECommerce Transactions Analysis: Exploratory Data Analysis (EDA)

Table of Contents

Overview

This task involves performing EDA to understand customer behavior, product sales, and transactional trends. By analyzing the data, the project identifies patterns and generates insights that can help businesses improve customer engagement, optimize product offerings, and increase revenue.

Dataset Description

The analysis uses three datasets:

1. **Customers.csv**
    - CustomerID: Unique identifier for each customer.
    - CustomerName: Name of the customer.
    - Region: Continent where the customer resides.
    - SignupDate: Date of customer signup.
2. **Products.csv**
    - ProductID: Unique identifier for each product.
    - ProductName: Name of the product.

- Category: Product category.
- Price: Price of the product (in USD).

3. **Transactions.csv**
    - TransactionID: Unique identifier for each transaction.
    - CustomerID: ID of the customer who made the transaction.
    - ProductID: ID of the product sold.
    - TransactionDate: Date of the transaction.
    - Quantity: Quantity of the product purchased.
    - TotalValue: Total value of the transaction (in USD).

Steps Performed

1. **Data Cleaning and Preparation**:
    - Handled missing and duplicate data.
    - Parsed date columns and extracted relevant time-based features (e.g., month-year for trends).
    - Merged datasets for richer analysis (e.g., joining customer, product, and transaction data).
2. **Visualization and Analysis**:
    - **Customer Signup Trends**:
        - Analyzed signups over time to identify patterns.
        - Created a line plot of signups by month-year.
    - **Total Sales by Region**:
        - Aggregated transaction data by region.
        - Visualized total sales using bar charts.
    - **Top-Selling Products**:
        - Ranked products based on the total quantity sold.
        - Visualized the top 10 products using bar charts.

3. **Deriving Business Insights**:
    - Extracted key insights from EDA to inform business decisions.

Key Insights

1. **Signup Trends**:
    - Customer signups show significant seasonal variation, with spikes in certain months (e.g., holiday seasons or promotional periods).
2. **Regional Sales Performance**:
    - South America leads in total sales, followed by Europe, indicating strong market presence in these regions.
3. **Top Products**:
    - Products like "ActiveWear Smartwatch" and "SoundWave Headphones" are the best-sellers, emphasizing customer interest in tech and wearable products.
4. **Customer Retention Opportunities**:
    - Customers purchasing high-value items often return for repeat transactions, indicating loyalty opportunities in specific segments.
5. **Cross-Selling Insights**:
    - Customers buying electronics often purchase complementary products, suggesting cross-selling strategies.

Technologies Used

- **Programming Language**: Python
- **Libraries**: pandas, numpy, matplotlib, seaborn
- **Development Environment**: Jupyter Notebook

Conclusion

The exploratory data analysis (EDA) of the eCommerce Transactions dataset provided critical insights, summarized as follows:

- **Seasonal Trends**: Customer signups show seasonal spikes, indicating opportunities for targeted marketing during peak periods.
- **Regional Performance**: South America leads in total sales, highlighting regional disparities and opportunities for expansion elsewhere.
- **Product Popularity**: Top-performing products, such as "ActiveWear Smartwatch" and "SoundWave Headphones," emphasize the value of tech-focused offerings.
- **High-Value Customers**: Loyal, repeat customers present significant opportunities for targeted loyalty programs.
- **Cross-Selling Potential**: Transaction patterns reveal opportunities to bundle and cross-sell complementary products.

These insights establish a foundation for further analysis in predictive modeling and customer segmentation.