Task 1: Exploratory Data Analysis (EDA) and Business Insights

Divya Aggarwal
IIT Roorkee
Email id : d_aggarwal@hs.iitr.ac.in
Contact Number: +91-8307114763

The datasets provides a comprehensive look into customer behavior, product performance, and transaction trends. Below is a summary of key insights derived from the data:

Customer Insights

1. **Customer Distribution by Region:**
    - South America leads with 59 customers, followed by Europe (50), North America (46), and Asia (45).
2. **Signup Trends:**
    - Majority of signups occurred in 2024 (79), followed by 2022 (64), and 2023 (57).
3. **Top Customers (by Transaction Count):**
    - Customers **C0109** and **C0156** made the most transactions (11 each).

Product Insights

1. **Top 5 Best-Selling Products (by Quantity):**
    - **P059** and **P054** sold the most (46 units each), followed by **P029 (45)**, **P079 (43)**, and **P061 (43)**.
2. **Revenue Contribution by Category:**
    - **Books** generated the highest revenue at $192,147.47.
    - Categories **Electronics**, **Clothing**, and **Home Decor** followed, contributing $180,783.50, $166,170.66, and $150,893.93 respectively.
3. **Descriptive Statistics for Product Prices:**
    - Average Price: $267.55
    - Maximum Price: $497.76 (likely premium items in Electronics or Home Decor).
    - Minimum Price: $16.08 (likely from the Books category).

Transaction Insights

1. **Monthly Transaction Trends (2024):**
   - January had the highest number of transactions, followed by months like September and July. November had the lowest.
2. **Transaction Distribution:**
   - A vast majority of transactions (996) occurred in 2024, while only 4 were in late 2023.
3. **Descriptive Statistics for Transactions:**
   - Average Quantity per transaction: **2.54 items**.
   - Maximum Total Value: $1991.04 (likely a high-ticket product or bulk purchase).
4. **Correlational insights:**
   - **Quantity vs. TotalValue (0.609972)**: There's a moderate positive correlation between the quantity of items purchased and the total transaction value, meaning that higher quantities typically lead to a higher total value of the transaction.
   - **TotalValue vs. Price (0.722714)**: There is a strong positive correlation between the total value of a transaction and the price of the products. This indicates that higher-priced products contribute more to the total value of the transaction.