Exploratory Data Analysis (EDA) and Business Insights

Insights on Customer and Product Trends

1. Customer Lifetime Value (CLV) Analysis

Revenue and Transaction Frequency: We analyzed total revenue per customer along with transaction frequency.

Customer Lifecycle: The difference between the first and the most recent transaction was computed to understand customer behavior over time.

Segmented CLV Analysis: The dataset is evenly distributed across three segments: Low, Medium, and High CLV.

Average CLV by segment shows a clear trend-

High CLV: Highest average value.

Low CLV: Lowest average value.

Top 10 most valuable and least valuable customer were also determined- The most valuable customer is from south africa (CO153) with approx CLV of 14200 and least valuable is customer C0032 with only 224 CLV.

2. Product Category Sales Trends Over Time

By merging the products and transactions datasets, we grouped the data by Year-Month, Category, and calculated the Sum of Total Value for each group.

Key insights:

Books: Sold the most in May and the least in November.

Clothing: Sales peaked in September, while March had the lowest.

Electronics: Highest sales in July, with the lowest in February.

Home Décor: Sold the most in December and the least in November.

3. Region-Based Segmentation for Marketing and Inventory

Combining transactions and customer datasets revealed sales by region-

Total Spend: Highest in South America and lowest in North America.

Quantity of Goods: Most goods sold in South America but least in Asia.

Key Insight:

While North Americans spend slightly less, they purchase a greater quantity of items, emphasizing value-for-money preferences.

4. Multiple Purchases by a Single Customer

Purchase Frequency:

Over 40 customers bought 5 products.

3 customers made 11 purchases each.

13 customers made only 1 purchase.

Key Insight:

Total purchase is low when to number of customers is small but eventually grows in middle and again there is a steep as purchases number increases.

5. Sign-Up Trends

Sign-up patterns show seasonal fluctuations:
Sign-ups peaked in early 2022. Decline during summer and winter, with a recovery just before autumn each year. This pattern suggests customers are more likely to sign up in anticipation of colder months, potentially driven by holiday shopping needs.

6. Top 1% and Bottom 1% Customer Demographics

Top 1% Customers:
Average Spending: $272.74
Most Frequently Purchased Category: Books
Month with Highest Spending: January
Region with Most Spending: South America

Bottom 1% Customers:
Average Spending: $188.33
Most Frequently Purchased Category: Clothing
Month with Highest Spending: November
Region with Most Spending: South America

Key Insight:
Both the highest and lowest spenders are from South America, reflecting a broad spectrum of spending habits within the region.