

Task 1 - Exploratory Data Analysis

An exploratory analysis of an e-commerce dataset containing **customer**, **product**, and **transaction** information.

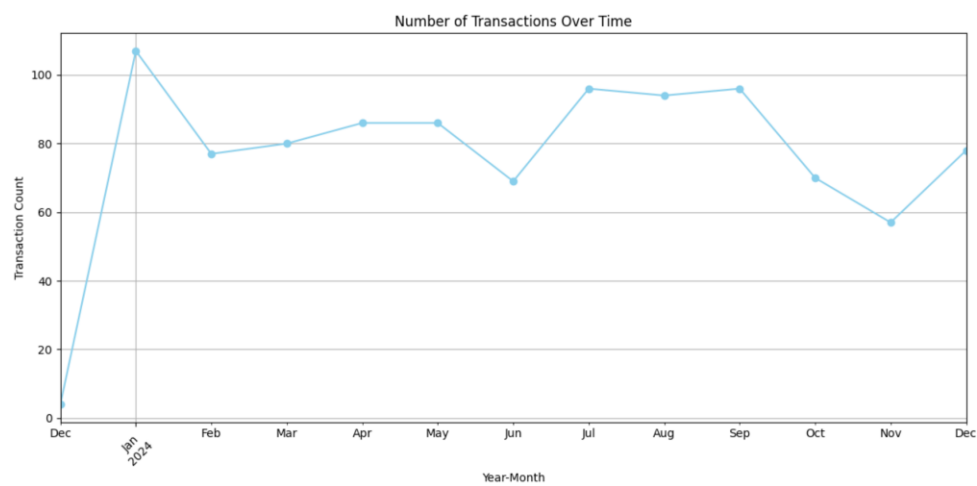
Dataset Overview

- 1. **Customers**
 - 200 unique customers across four global regions.
 - Includes sign-up dates ranging from early 2022 to late 2024.
- 2. **Products**
 - 100 distinct products categorized into **Books**, **Electronics**, **Clothing**, and **Home Decor**.
 - Prices range from \$16.08 to \$497.76, with a mean of \$267.55.
- 3. **Transactions**
 - 1,000 transaction records, each indicating the product purchased, quantity, transaction date, and total value.
 - Average transaction value is \$689.99, suggesting moderate- to high-value purchases.

Key Insights:

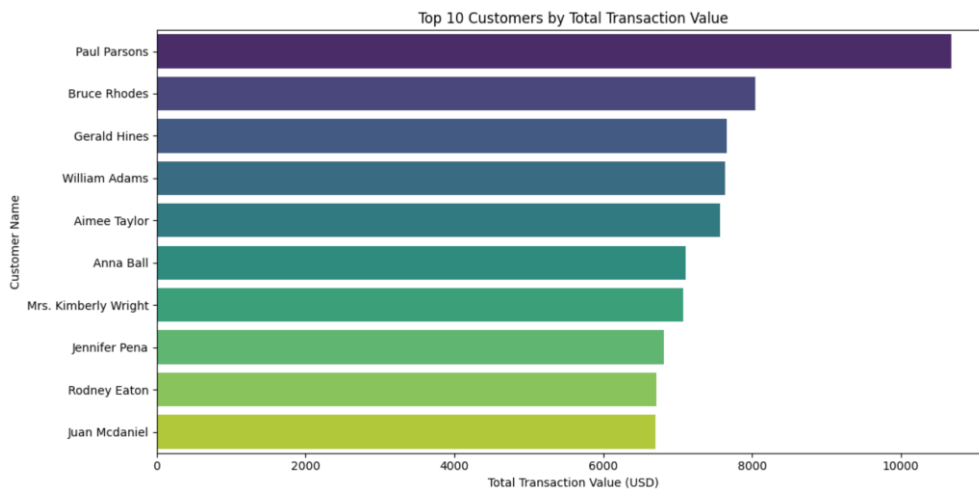
1. Transaction Surge in Early 2024:

Tracked the monthly count of transactions from December 2023 to December 2024, revealing a significant jump in early 2024. This likely reflects a post-holiday or new-year promotional effect.



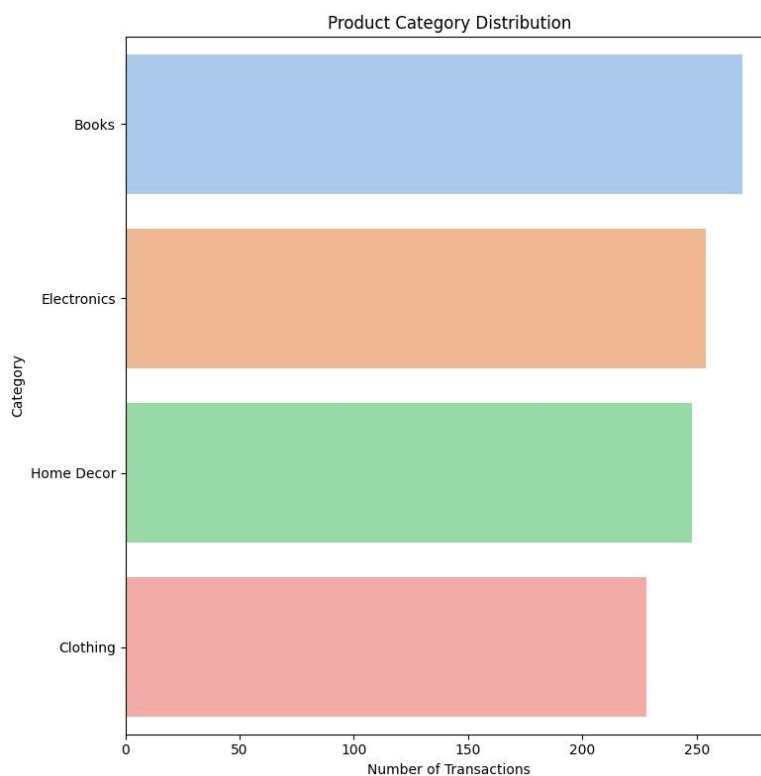
2. Revenue Concentration Among Top Buyers:

A small subset of customers (top 10) drives a substantial portion of overall sales. Personalizing offers and loyalty programs for these high-value customers could increase retention.



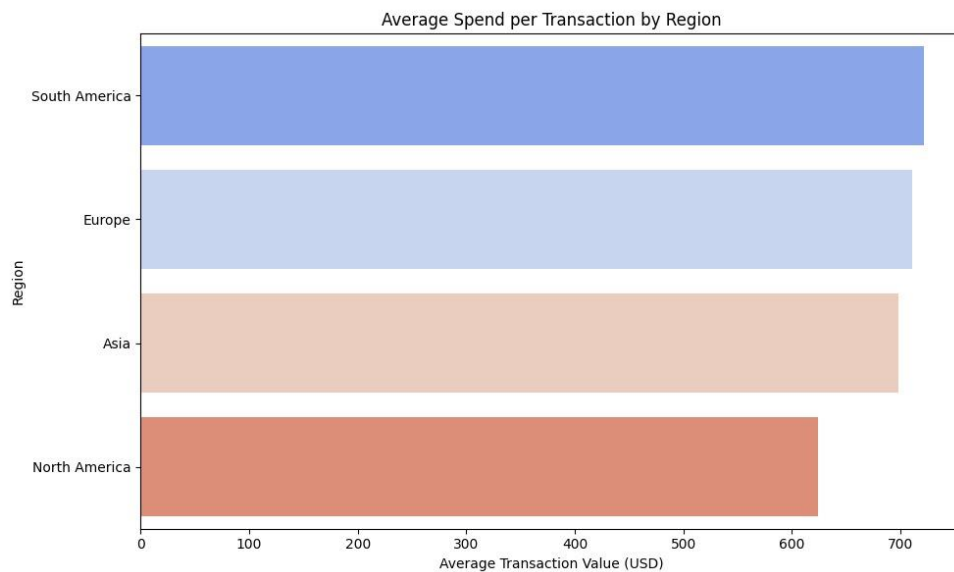
3. Dominance of Books and Electronics:

Books appear most frequently purchased and also top total revenue, with Electronics closely following and These categories may benefit from increased stock and focused promotions.



4. South America Leads in Average Transaction Value:

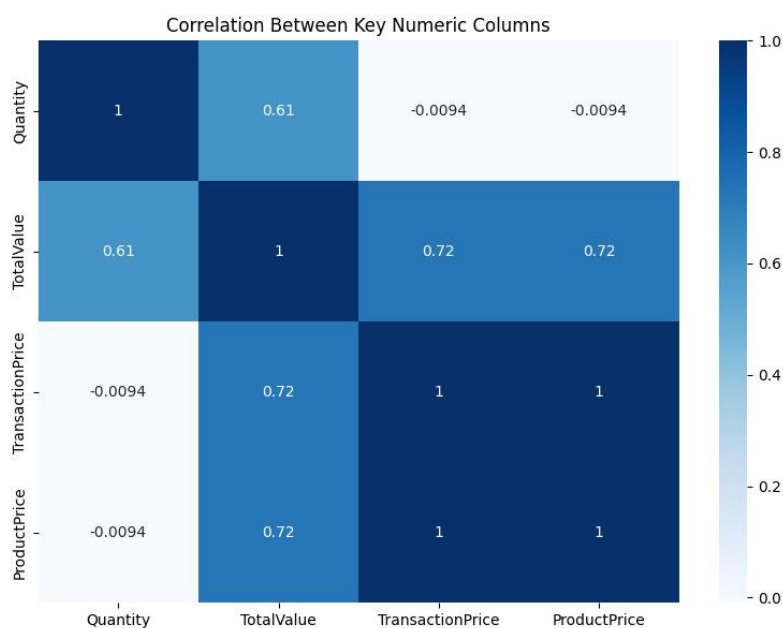
Compared average transaction values across countries. Despite a smaller customer base in some periods, the higher average spend makes it a premium market and Asia's larger transaction volumes come with lower average purchase size, suggesting a need for region-specific strategies.



5. Price as a Strong Revenue Driver:

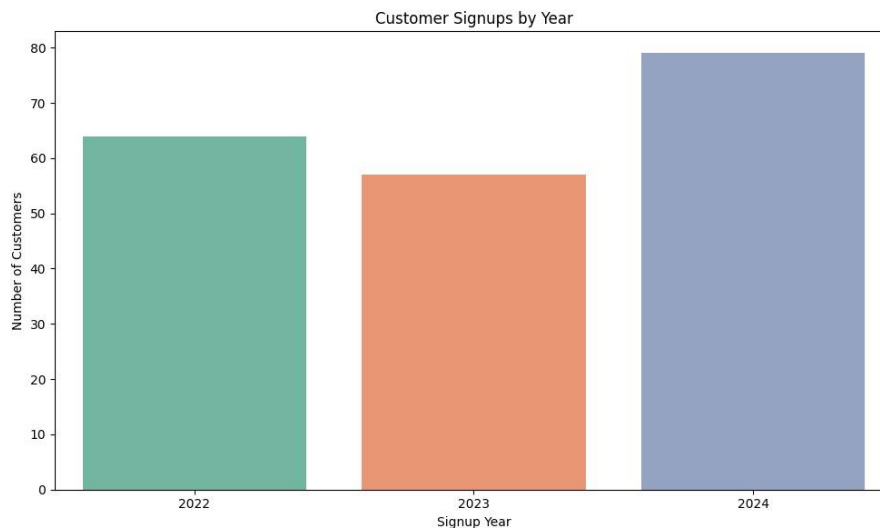
The correlation matrix reveals a **high (≈ 0.72)** correlation between product price and total value. Premium-priced items effectively boost revenue, though offering bulk deals may further leverage higher quantities

However, the correlation of 1.0 between Transaction Price and Product Price simply reflects their direct dependency, **not a revenue-driving insight**.



6. Customer Growth Trends:

The steady rise in customer sign-ups over the years indicates effective acquisition strategies. Continued focus on customer retention, especially for newer sign-ups, could sustain growth.



The Exploratory analysis highlights key opportunities for growth and optimization. A strong start in 2024 signals potential for leveraging seasonal promotions. Books and Electronics dominate sales, requiring strategic stock and marketing efforts. Top customers and premium-priced items drive significant revenue, emphasizing the need for personalization and pricing strategies. South America's high transaction value and Asia's volume-driven market suggest region-specific approaches. Finally, the steady growth in customer sign-ups reinforces the importance of retention strategies to sustain long-term success.

Code Link:

<https://colab.research.google.com/drive/1bqkkYTQwk393Pi3d4TKnIFYXjeoU7SO0?usp=sharing>