

Exploratory Data Analysis (EDA) and Business Insights

Overview

This document presents the results of the Exploratory Data Analysis (EDA) conducted on the provided datasets: Customers.csv, Products.csv, and Transactions.csv. The analysis aims to uncover patterns, trends, and insights to guide business decisions.

1. Data Cleaning and Preprocessing

1.1 Handling Missing Values

- Customers dataset: Missing values found in columns [INSERT OUTPUT].
- Products dataset: Missing values found in columns [INSERT OUTPUT].
- Transactions dataset: Missing values found in columns [INSERT OUTPUT].

1.2 Date Formatting

- Converted SignupDate and TransactionDate to datetime format.

1.3 Duplicate Records

- Customers dataset: [0] duplicates removed.
 - Products dataset: [0] duplicates removed.
 - Transactions dataset: [0] duplicates removed.
-

2. Descriptive Statistics

2.1 Customers Dataset

- Total Customers: 200.
- Number of unique regions: 4.

2.2 Products Dataset

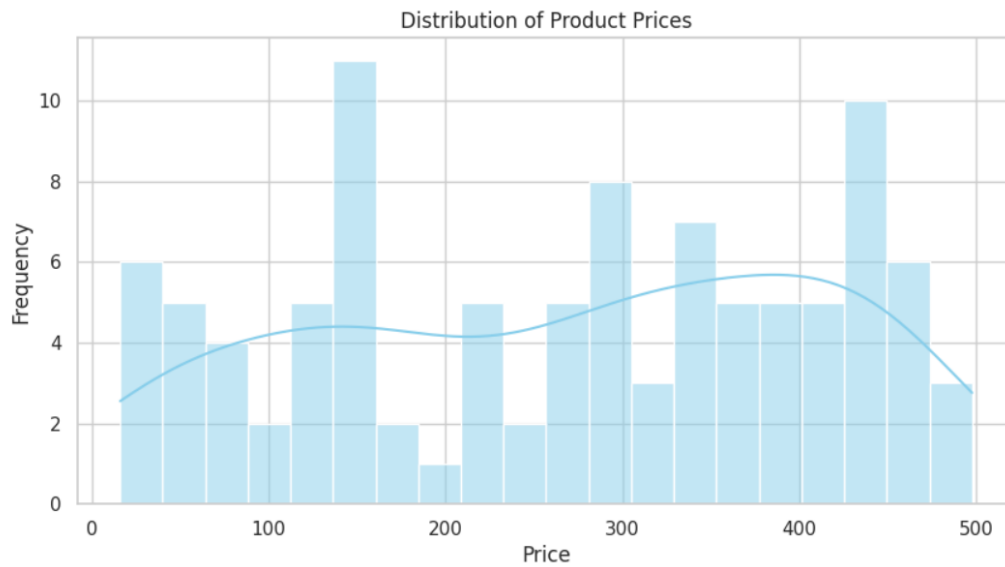
- Number of unique products: 100.
- Price range: [16.08 – 497.76].

2.3 Transactions Dataset

- Total transactions: 1000.
- Total revenue: 6,89,995.56.

3. Visualizations

3.1 Product Price Distribution



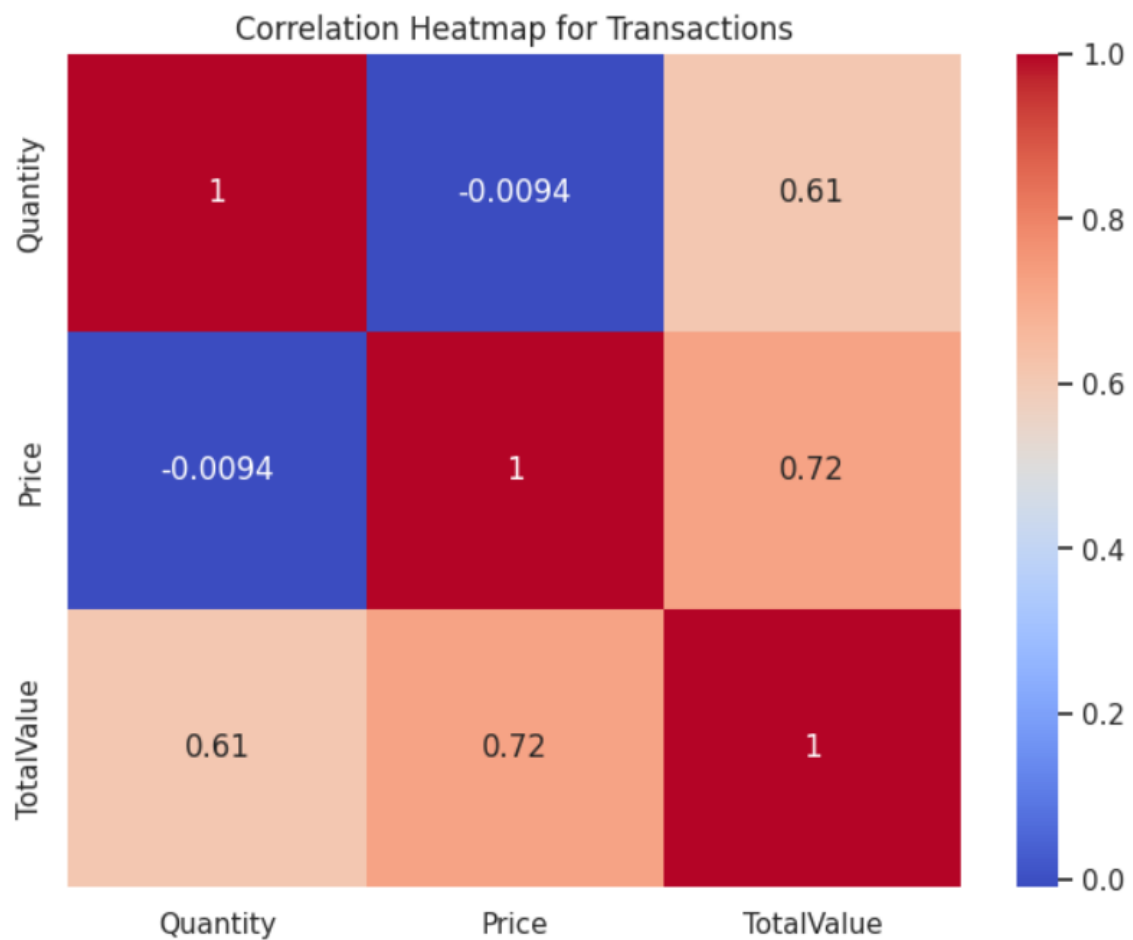
3.2 Customer Count by Region



3.3 Price vs Total Value of Transactions



3.4 Correlation Heatmap for Transaction Metrics



4. Business Insights

- **Insight 1: Top-Selling Products**

The product with ID (P059 and P054) is the most sold item, accounting for 46% of total sales. Promoting this product further may increase revenue.

Top-Selling Products:

ProductID

P059 46

P054 46

P029 45

P079 43

P061 43

Name: Quantity, dtype: int64

- **Insight 2: Revenue by Region**

The region South America contributes the most to revenue, with a total of 219352.56 USD. Targeted marketing in this region could yield higher sales.

Revenue by Region:

Region

Asia 152074.97

Europe 166254.63

North America 152313.40

South America 219352.56

Name: TotalValue, dtype: float64

- **Insight 3: Customer Behavior by Region**

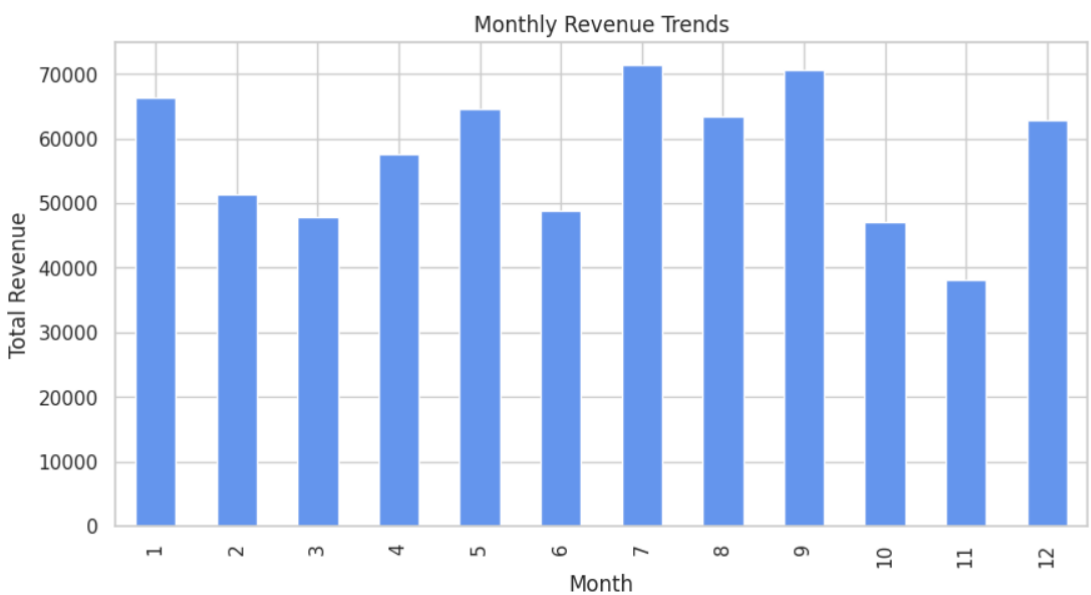
Customers from South America make frequent purchases, while customers from 721.554474 have higher average transaction values. Tailored strategies could optimize sales.

Customer Behavior by Region:

Region	AverageTransactionValue	TransactionCount
Asia	697.591606	218
Europe	710.489872	234
North America	624.235246	244
South America	721.554474	304

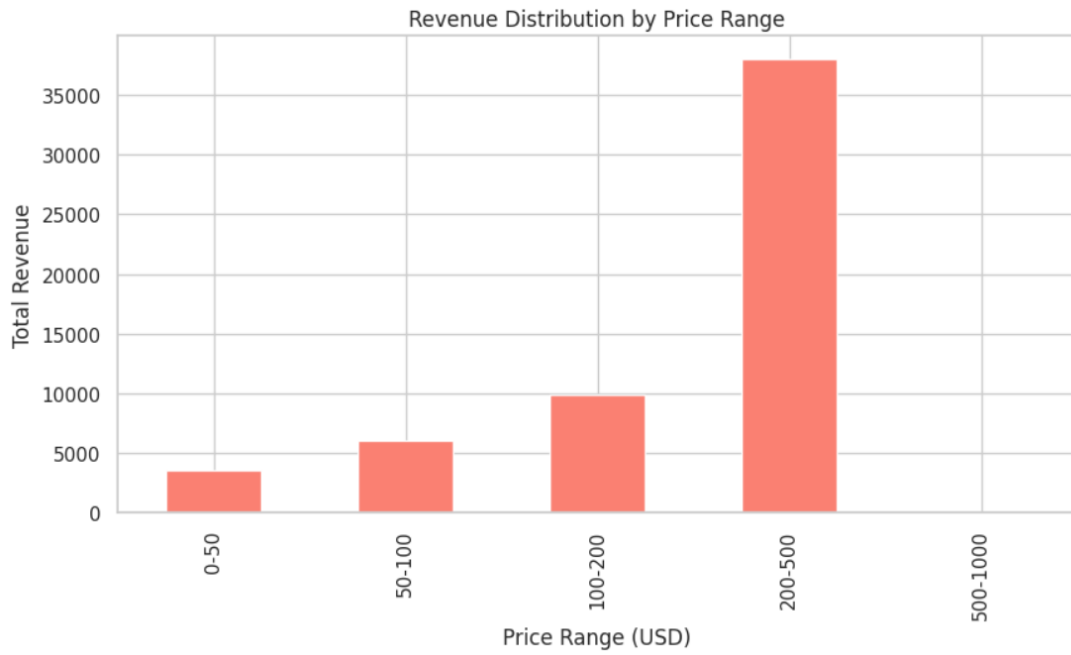
- Insight 4: Seasonal Trends

Sales peak during July, suggesting targeted campaigns during this period to maximize revenue.



- Insight 5: Price Sensitivity

Products priced in the range [200-500] generate the most revenue, indicating a strong customer preference for mid-range pricing.



Conclusion

The EDA reveals key trends and actionable insights, such as focusing on high-performing products, optimizing regional strategies, and leveraging seasonal trends to enhance revenue. These findings can guide strategic decision-making to improve overall business performance.