

Exploratory Data Analysis (EDA)

Data Overview

Customers Dataset (Top 5 Rows):

CustomerID	CustomerName	Region	SignupDate
C0001	Lawrence Carroll	South America	10-07-2022
C0002	Elizabeth Lutz	Asia	13-02-2022
C0003	Michael Rivera	South America	07-03-2024
C0004	Kathleen Rodriguez	South America	09-10-2022
C0005	Laura Weber	Asia	15-08-2022

Products Dataset (Top 5 Rows):

ProductID	ProductName	Category	Price
P001	ActiveWear Biography	Books	169.3
P002	ActiveWear Smartwatch	Electronics	346.3
P003	ComfortLiving Biography	Books	44.12
P004	BookWorld Rug	Home Decor	95.69
P005	TechPro T-Shirt	Clothing	429.31

Transactions Dataset (Top 5 Rows):

TransactionID	CustomerID	ProductID	TransactionDate	Quantity	TotalValue	Price
T00001	C0199	P067	25-08-2024 12:38	1	300.68	300.68
T00112	C0146	P067	27-05-2024 22:23	1	300.68	300.68
T00166	C0127	P067	25-04-2024 07:38	1	300.68	300.68
T00272	C0087	P067	26-03-2024 22:55	2	601.36	300.68

Summary Statistics

Customers:

- Total Rows: 200
- Unique Regions: 4
- Most Frequent Region: South America

Products:

- Total Rows: 100
- Unique Categories: 4
- Highest Price: \$497.76

Transactions:

- Total Rows: 1,000
- Average Total Value: \$689.99

Visualizations

Customer Distribution by Region

Observation Points:

- The bar chart shows the distribution of customers across different regions.
- South America has the highest number of customers, making it the largest market.
- Europe comes second, showing a strong customer base.
- Asia and North America have comparable customer counts, with North America slightly ahead of Asia.
- Africa has the lowest number of customers, indicating potential for growth in this region.



Price Distribution of Products

Observation Points:

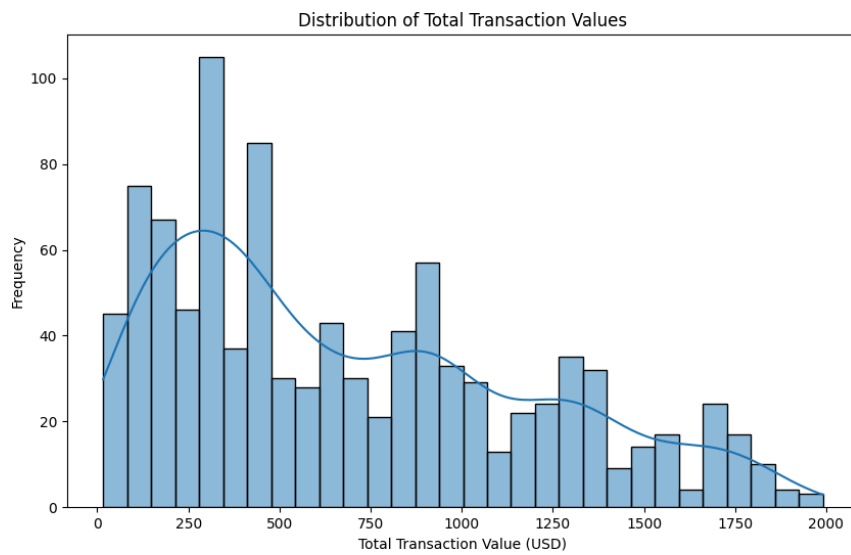
- The price distribution is skewed towards products priced between \$100 and \$400.
- There is a noticeable peak in the \$200-\$300 range, indicating that most products fall within this price bracket.
- A smaller number of products are priced below \$100 or above \$400.
- The distribution appears normal with a slight right skew, which may indicate that higher-priced products are less common.



Distribution of Total Transaction Values

Observation Points:

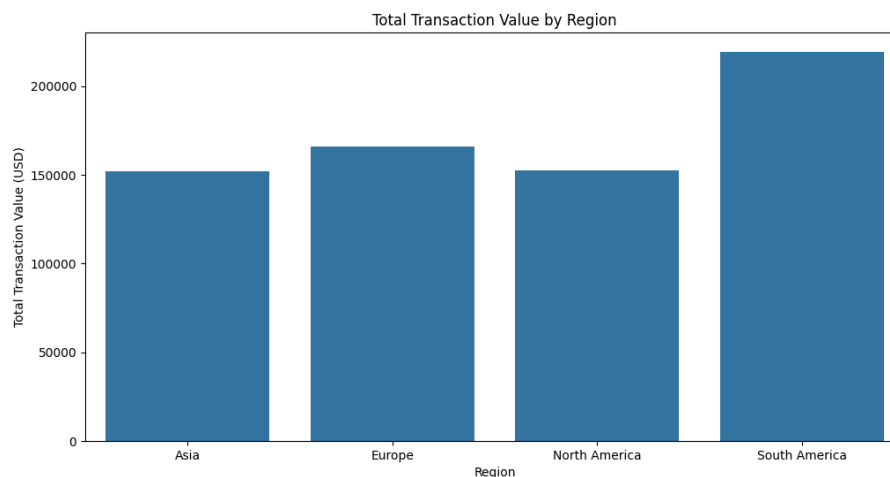
- The transaction values have a wide distribution, with the majority of transactions falling within the low-to-medium value range.
- A few high-value transactions contribute to a long right tail in the distribution, suggesting that there are some significant outliers in transaction value.
- The kde plot supports the fact that most of the transactions are on the lower end, but a few are significantly higher.



Total Transaction Value by Region

Observation Points:

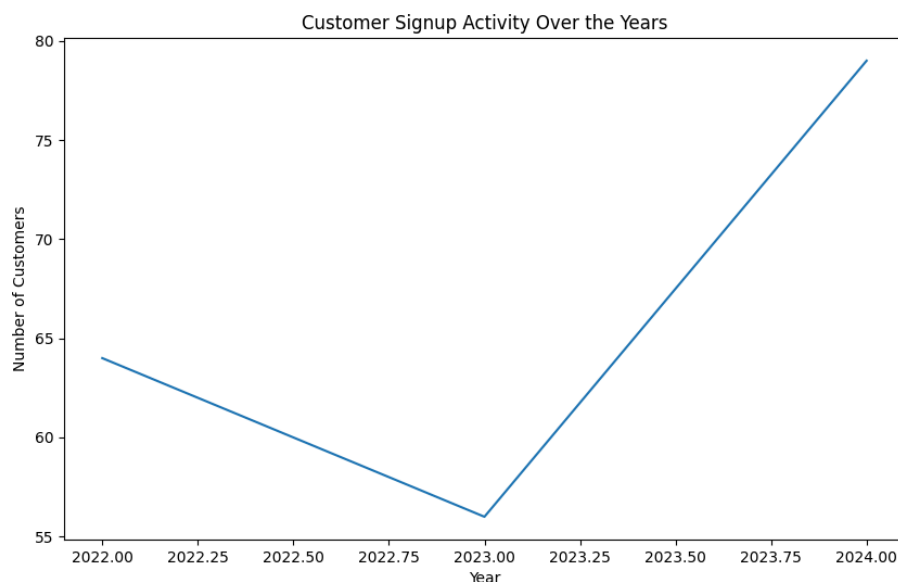
- South America surpasses other regions in transaction value, highlighting its strong performance in this dataset.
- Europe leads in total transaction value, closely followed by North America and Asia, indicating these regions are significant contributors to overall sales.
- Africa remains the lowest in total transaction value, consistent with its smaller customer base.
- The revenue distribution aligns with the customer spread, though South America's higher transaction values suggest a strong consumer presence in that region.



Customer Signup Activity Over the Years

Observation Points:

- The trend indicates a decline in customer signups from 2022 to 2023, suggesting a potential slowdown in customer acquisition during this period.
- Following this decline, there is a noticeable rebound in signups from 2023 to 2024, indicating renewed growth and successful customer engagement strategies.
- This pattern highlights fluctuations in customer signup activity, which could be influenced by external factors or shifts in marketing approaches.
- The chart reflects the platform's ability to regain momentum in expanding its customer base after a period of decline.



Average Transaction Value by Product Category

Observation Points:

- Clothing has a slightly higher average transaction value, indicating a marginally higher price range for these items.
- Books and Electronics have similar average transaction values, reflecting moderately priced items in these categories.
- Home Decorations have the lowest average transaction value, suggesting more affordable products in this category.
- This data highlights varying price points across product categories, with Clothing standing out for its relatively higher average transaction value.

