

Insights Derived from Visualizations and EDA

1. Distribution of Customers by Region

Insight:

The visualization shows that North America has the highest number of customers, followed by other regions like Europe and Asia. This indicates that the business is heavily concentrated in North America, contributing the most to the customer base. Other regions, although represented, have significantly fewer customers compared to North America.

This trend could imply a strong presence or effective marketing strategies in North America, better accessibility to products, or a higher demand for the offerings in that region. Conversely, the underperformance in other regions may indicate a need for increased promotional efforts, logistical improvements, or tailored products to fit local demand.

2. Distribution of Products by Category

Insight:

The visualization reveals that Books and Electronics are the most purchased categories. These two categories dominate product sales, suggesting they are the most popular or essential for customers. Other categories, while present, may lack the same appeal or importance.

This trend highlights the potential need to diversify inventory within these popular categories or better promote the less popular ones. It could also indicate seasonal or event-driven spikes in demand for these categories, such as holidays or sales campaigns.

3. Trends in Transaction Dates

Insight:

The monthly sales trend shows significant fluctuations in sales, with highest sales occurring in the 7th and 9th months, while the biggest drop is observed in the 11th month. This suggests potential seasonality in purchasing patterns, possibly tied to holidays, back-to-school shopping, or seasonal promotions. The sharp decline in the 11th month might indicate a post-seasonal slump or reduced consumer spending.

Businesses can leverage these insights to optimize inventory management, marketing campaigns, and resource allocation to align with peak demand periods. The decline in November calls for strategies to counteract this drop, such as targeted discounts or promotional offers.

4. Distribution of Quantity Purchased

Insight:

The quantity distribution reveals a wavy KDE pattern within a frequency range of 0–50, with quantities varying primarily between 1 and 4, and the highest frequency at 4. This implies most customers purchase small quantities, suggesting either a retail-focused model or individual consumption rather than bulk buying.

Businesses can interpret this as an opportunity to create promotional bundles or discounts for higher quantities, encouraging customers to purchase more at once.

5. Distribution of Total Transaction Value

Insight:

The total transaction value distribution shows that \$2,000 has the lowest frequency, while \$250 has the highest frequency with a peak of 60 transactions. This suggests that most transactions fall within a moderate spending range, with fewer high-value purchases.

This insight can guide pricing strategies, promotional offers, and segmentation of customers based on spending capacity. The lower frequency of high-value transactions may indicate a need for premium product bundles or targeted marketing toward high-income segments.