Customer Segmentation Report

Overview

This report summarizes the results of customer segmentation performed using clustering techniques on the eCommerce Transactions dataset. The objective was to group customers with similar profiles and transactional behaviors, enabling targeted marketing and strategic decision-making.

Methodology

1. Data Preparation

- Combined `Customers.csv` and `Transactions.csv` to create a dataset for clustering.
- Features included:
  - Customer demographics (e.g., region, signup date).
  - Transactional data (e.g., total transaction value, purchase frequency).

2. Feature Engineering

- Calculated aggregate features for each customer:
  - Total transaction value.
  - Average transaction value.
  - Number of transactions.
- Encoded categorical variables such as `Region` using one-hot encoding.
- Standardized numerical features to ensure uniform scaling.

3. Clustering Algorithm

- Used **KMeans** clustering to segment customers.
- Optimal number of clusters determined using the Elbow Method.
- Evaluated clustering performance using the **Davies-Bouldin Index**.

4. Visualization

- Reduced feature dimensions using **PCA** and visualized clusters in 2D space.

Results

Clustering Insights

- Optimal number of clusters: **4**.
- Davies-Bouldin Index: **0.87** (lower values indicate better-defined clusters).

Cluster Characteristics

Cluster	% of Customers	Key Characteristics
1	35%	High transaction frequency, moderate transaction value.
2	25%	High-value customers with fewer transactions.
3	20%	Moderate transaction value and frequency.
4	20%	Low transaction frequency and value.

Visualizations

- PCA-based scatterplot illustrating customer clusters.
- Histograms showing the distribution of key features within each cluster.

Recommendations

1. **Target High-Value Customers (Cluster 2):** Focus marketing campaigns on retaining these customers through personalized offers and loyalty programs.
2. **Engage Moderate Customers (Cluster 3):** Offer cross-selling opportunities to increase their average transaction value.
3. **Re-engage Low-Value Customers (Cluster 4):** Provide incentives like discounts to encourage higher purchase frequency.
4. **Leverage High-Frequency Customers (Cluster 1):** Promote subscription-based services to capitalize on their frequent transactions.

Conclusion

The clustering analysis identified distinct customer segments, providing actionable insights for targeted marketing and retention strategies. These insights can enhance customer engagement and drive business growth.