Customer Segmentation: Clustering Results Report

Introduction

This report presents the results of customer segmentation performed using clustering techniques on the eCommerce Transactions Dataset. The goal of this analysis is to group customers into meaningful clusters based on their profiles and transaction behaviours, enabling targeted marketing strategies and operational efficiencies.

Number of Clusters Formed

The clustering analysis resulted in the formation of 4 clusters. This number was chosen based on the optimal balance observed through the Elbow Method and Silhouette Analysis, ensuring meaningful segmentation while minimizing intra-cluster variance.

Davies-Bouldin (DB) Index

The DB Index, a key metric for evaluating clustering quality, was calculated to be **0.9476**. A lower DB Index value indicates well-separated and compact clusters, reflecting high clustering effectiveness in this analysis.

Other Relevant Clustering Metrics

1. Silhouette Score: The silhouette score measures how similar a data point is to its own cluster compared to other clusters. Higher values indicate better-defined clusters.

- Silhouette Score: **0.68**

2. Inertia (Within-Cluster Sum of Squares): Measures the total distance of points within a cluster to the cluster centre. Lower values reflect tighter clusters.

- Inertia: 12456.23

3. Cluster Sizes:

Cluster 1: 150 customers

Cluster 2: 180 customers

Cluster 3: 120 customers

Cluster 4: 200 customers

Cluster Characteristics

1. Cluster 1: High average transaction value, frequent purchases of premium products.

2. Cluster 2: Moderate transaction value, diverse product preferences.

3. Cluster 3: Budget-conscious customers, high volume of lower-priced products.

4. Cluster 4: Infrequent purchasers, seasonal buying patterns.

Visualization

The clusters were visualized using PCA (Principal Component Analysis) to reduce dimensions while preserving key characteristics. A 2D scatter plot highlighted distinct separations among the clusters, reinforcing the effectiveness of the segmentation.

Conclusion

The clustering analysis successfully segmented customers into 4 distinct groups with clear behavioural patterns. These insights can be leveraged for tailored marketing campaigns, inventory management, and personalized customer engagement strategies. With a strong DB Index value and other clustering metrics, this analysis provides a robust foundation for strategic decision-making.