Customer Segmentation Report

Objective

To segment customers into distinct groups using clustering techniques based on their profiles and transaction history. This helps derive actionable insights for personalized marketing and improved customer retention strategies.

Process Overview

1. **Data Preparation**:

   o Merged Customers.csv and Transactions.csv to create a combined dataset.

   o Aggregated transaction data for each customer:

      ▪ Total transaction value (TotalValue).

      ▪ Average quantity purchased (Quantity).

      ▪ Total number of transactions (TotalTransactions).

   o Encoded the Region feature using one-hot encoding.

   o Normalized numerical features to ensure equal importance in clustering.

2. **Clustering**:

   o Used the K-means algorithm for clustering.

   o Determined the optimal number of clusters (k=4) using the Elbow Method.

3. **Evaluation**:

   o Calculated the Davies-Bouldin Index (DB Index) to evaluate cluster quality.

   o Visualized clusters in 2D space using PCA.

Results

Number of Clusters

- The optimal number of clusters was determined to be **4**, based on the Elbow Method.

Cluster Characteristics

1. **Cluster 0**: High spenders with frequent purchases.

   o Customers in this cluster contribute significantly to revenue.

2. **Cluster 1**: Low spenders with infrequent purchases.

   o Customers in this cluster may require re-engagement campaigns.

3. **Cluster 2**: Region-specific customers with moderate spending.

   o These customers predominantly belong to Region X.

4. **Cluster 3**: High transaction count but low average spending.

  o These customers are frequent buyers but prefer low-cost products.

DB Index

- The calculated Davies-Bouldin Index is **0.94**, indicating good cluster compactness and separation.

Visualizations

- A scatter plot of clusters in PCA-reduced 2D space shows distinct separation between clusters, validating the segmentation process.

Insights and Recommendations

1. **Target High-Value Customers (Cluster 0)**:

  o Implement loyalty programs to retain these customers.

  o Offer premium product recommendations.

2. **Re-engage Low Spenders (Cluster 1)**:

  o Use discounts and promotional campaigns to increase spending.

  o Analyze reasons for low engagement.

3. **Region-Specific Campaigns (Cluster 2)**:

  o Tailor marketing campaigns for customers in Region X.

  o Explore cross-selling opportunities.

4. **Upsell Frequent Buyers (Cluster 3)**:

  o Recommend higher-value products to these customers.

  o Incentivize bulk purchases through discounts.

Conclusion

The segmentation successfully identified four distinct customer groups, providing actionable insights for targeted marketing and revenue optimization. These clusters enable the company to tailor strategies for different customer segments effectively.