Customer Segmentation Clustering Report

Overview

This report presents the results of customer segmentation using K-Means clustering. The analysis involved the following steps:

1. Data preprocessing and feature scaling.
2. Applying K-Means clustering.
3. Evaluating the clustering performance using Davies-Bouldin Index (DB Index).
4. Visualizing clusters using Principal Component Analysis (PCA).
5. Analyzing cluster centers for business insights.

Clustering Results

Number of Clusters Formed

The K-Means clustering algorithm was applied with 5 clusters, segmenting the customers into distinct groups based on their transaction behavior.

Davies-Bouldin Index (DB Index)

The Davies-Bouldin Index is used to measure the quality of clustering, where a lower value indicates better clustering performance. The calculated DB Index for the clustering results is:

DB Index: {db_index}

A DB Index value closer to zero suggests well-separated clusters with high intra-cluster similarity.

Cluster Centers

The cluster centers represent the average values of the standardized features within each cluster. These values help in understanding the characteristics of each segment.

Cluster	Total Spend	Total Products Bought	Recency
0	-1.478348	-1.520976	2.842258
1	-0.338375	-0.267860	-0.688415
2	-1.030844	-1.025428	0.116456
3	1.344399	1.362158	-0.481844

Cluster	Total Spend	Total Products Bought	Recency
4	0.206674	0.112256	0.493818

Interpretation of clusters:

- Cluster 0: Customers with low spending and high recency (likely inactive customers).
- Cluster 1: Average spenders with relatively low recency.
- Cluster 2: Below-average spenders with moderate recency.
- Cluster 3: High spenders with low recency (loyal customers).
- Cluster 4: Customers with moderate spending and higher recency.

PCA Visualization

To visualize the clusters, Principal Component Analysis (PCA) was applied to reduce the dimensionality of the data. The scatter plot below shows the clusters in a two-dimensional space:

The clustering visualization highlights the separation between customer segments, making it easier to interpret the segmentation patterns.

Customer Segmentation Summary Table

CustomerID	Total Spend	Total Products Bought	Last Purchase Date	Recency
C0001	-0.061701	-0.122033	2024-11-02	-0.258366
C0002	-0.877744	-0.448000	2024-12-03	-0.696600
C0003	-0.405857	0.203934	2024-08-24	0.731195
C0004	1.032547	1.670787	2024-12-23	-0.979331
C0005	-0.783929	-0.936951	2024-11-04	-0.286639

Conclusion

The customer segmentation analysis provides valuable insights into customer behaviors, enabling targeted marketing strategies. The clusters can help businesses identify loyal customers, inactive customers, and potential opportunities for customer re-engagement. Further analysis can include customer churn prediction and personalized marketing campaigns based on segmentation insights.