Clustering Results Report

Overview
This report summarizes the results of customer segmentation performed using clustering techniques. The analysis combined customer demographic and transaction data to group customers into distinct clusters based on their behavior and profiles.

Results Summary

1. Number of Clusters Formed
- Optimal Number of Clusters: 4

  The optimal number of clusters was determined based on the Davies-Bouldin Index (DB Index) and visual inspection of cluster performance.

2. Davies-Bouldin Index (DB Index)
- DB Index Value: 0.65

  A lower DB Index value indicates better-defined clusters. This value reflects the separation and compactness of the clusters.

3. Other Clustering Metrics
- Silhouette Score: 0.45

  (Optional) The silhouette score measures how similar customers are within their clusters, with higher values indicating better clustering quality. A score of 0.45 suggests moderate cluster cohesion and separation.

- Cluster Sizes:

	CustomerID	Cluster
0	C0001	1
1	C0002	1
2	C0003	1
3	C0004	1
4	C0005	2

  The distribution of customers across clusters highlights their varying behavior and demographics.

Key Insights

1. Customer Grouping
Each cluster represents a distinct customer segment with shared characteristics, enabling tailored marketing strategies and product recommendations.

2. Cluster Characteristics

- Cluster 1: High-value customers with frequent transactions.
- Cluster 2: Occasional buyers with mid-range spending.
- Cluster 3: Price-sensitive customers purchasing in bulk.
- Cluster 4: New customers with limited transaction history.

3. Visual Representation

A PCA plot was used to reduce the data dimensions for visualization. It showed clear groupings of customers, reinforcing the distinctiveness of the clusters.

Conclusion

The clustering analysis successfully segmented customers into four distinct groups. The results provide actionable insights to improve customer targeting, enhance product offerings, and develop personalized marketing campaigns. The evaluation metrics, including the DB Index and silhouette score, confirm the quality of the clustering.

For further refinement, additional metrics such as customer lifetime value or churn probability can be integrated into future analyses.