Customer Segmentation and Clustering Report

Introduction

The objective of this analysis is to segment customers into distinct groups using clustering techniques.

The segmentation uses customer profile information and transaction history to identify patterns in customer behavior.

This report summarizes the clustering results, metrics, and key insights.

Number of Clusters Formed

After evaluating clustering performance across a range of 2 to 10 clusters,

the optimal number of clusters was determined to be X clusters.

This choice was guided by the Davies-Bouldin (DB) Index, a metric that evaluates cluster compactness and separation.

Clustering Metrics

- Optimal Number of Clusters (k): X
- Davies-Bouldin Index for Optimal Clusters: Y.Z (Lower values indicate better clustering performance)
- Average Intra-Cluster Distance: A measure of how similar the customers within each cluster are.
- Average Inter-Cluster Distance: A measure of how distinct each cluster is from the others.

Cluster Insights

Each cluster represents a unique customer segment with distinct characteristics. Below is a summary of the key attributes for each cluster:

Cluster	Total Customers	Avg Revenue (USD)	Avg Transaction Count	Avg Days Since Signup
Cluster 0	XX	$XXXX	XX	XXX
Cluster 1	XX	$XXXX	XX	XXX
Cluster 2	XX	$XXXX	XX	XXX
...	...	...	...	...

Visualizations

1. Davies-Bouldin Index vs. Number of Clusters:

- A line chart was plotted to show the DB Index across different values of k. The optimal k was identified as the point with the lowest DB Index.

2. PCA-Reduced Feature Visualization:

- A scatter plot using PCA (Principal Component Analysis) reduced dimensions provided a two-dimensional view of the clusters.

Each cluster is represented by a unique color, highlighting distinct separations among clusters.

3. Cluster Summary Bar Chart:

- Visual representations of average revenue, transaction count, and days since signup were created for each cluster to better understand customer behavior.

Key Takeaways

1. Distinct Customer Groups: The clustering algorithm identified X distinct customer groups, each with unique spending behavior and engagement patterns.
2. High Revenue Clusters: Clusters X and Y represent high-value customers with significantly higher average revenue and frequent transactions.
3. Low Engagement Groups: Clusters A and B represent low-engagement customers with fewer transactions and low lifetime value.
4. Opportunities for Targeted Marketing: By understanding cluster characteristics, tailored marketing strategies can be developed to improve engagement and retention.
5. Scalability: The clustering approach can be periodically re-applied to monitor shifts in customer behavior and update segmentation.

Conclusion

The clustering analysis successfully segmented customers into actionable groups.

The insights derived from these clusters can inform targeted marketing, resource allocation, and personalized customer engagement strategies.

Further analysis can incorporate additional data points or explore advanced clustering techniques to refine segmentation.