Clustering Analysis Report

Objective

This report evaluates the clustering results by comparing different numbers of clusters using metrics such as the Davies-Bouldin Index (DB Index) and the Silhouette Score. The goal is to identify the optimal number of clusters for the dataset.

Clustering Results Summary

Number of Clusters	Davies-Bouldin Index	Silhouette Score
2	1.041	0.369
3	1.133	0.292
4	0.946	0.327
5	1.020	0.308
6	0.996	0.313
7	1.022	0.290
8	1.051	0.276
9	1.079	0.264
10	1.085	0.270

Key Observations

1. **Davies-Bouldin Index**

   o The lowest DB Index is observed for **4 clusters** (0.946), indicating better-defined and more compact clusters compared to other configurations.

   o As the number of clusters increases beyond 4, the DB Index begins to rise, suggesting that adding more clusters leads to less distinct separation between them.

2. **Silhouette Score**

   o The highest Silhouette Score is observed for **2 clusters** (0.369), which indicates the best separation and cohesion for this configuration.

o   However, the Silhouette Score decreases significantly for 3 clusters and remains relatively lower for larger numbers of clusters.

3. **Trade-Off Between Metrics**

   o   While the Silhouette Score favors 2 clusters, the DB Index and business interpretability point toward **4 clusters** as a reasonable balance between compactness and meaningful segmentation.

Optimal Number of Clusters

Based on the evaluation of both metrics:

- **4 clusters** appear to be the most optimal choice, as they achieve the best DB Index and a relatively stable Silhouette Score.

Cluster Insights and Business Implications

The **4-cluster model** provides the following customer segments:

1. **Cluster 0**: **High-Value Loyal Customers**

   o   **Characteristics**:

       ▪   Highest average spending (TotalValue mean = 6075.93).

       ▪   Strong engagement with frequent purchases (Quantity mean = 21.55).

       ▪   Significant customer tenure (NumYears mean = 1.81).

   o   **Business Insights**:

       ▪   These customers are the backbone of revenue generation.

       ▪   Focus on retention strategies such as loyalty programs, personalized offers, and VIP experiences to maintain and grow their value.

2. **Cluster 1**: **Moderate-Spending Loyal Customers**

   o   **Characteristics**:

       ▪   Moderate average spending (3053.88).

       ▪   Decent purchase frequency (Quantity mean = 11.34).

       ▪   Long customer tenure (NumYears mean = 2.44).

   o   **Business Insights**:

- This segment has growth potential through upselling or cross-selling.

- Offer educational content, bundled deals, or tiered loyalty rewards to increase their spending.

3. **Cluster 2**: **New or Sporadic Buyers**

   o **Characteristics**:

   - Moderate average spending (3658.75).

   - Lower purchase frequency (Quantity mean = 13.75).

   - Short customer tenure (NumYears mean = 0.69).

   o **Business Insights**:

   - This group likely consists of new customers or those with inconsistent purchasing patterns.

   - Implement onboarding programs, welcome discounts, or retention campaigns to convert them into long-term customers.

4. **Cluster 3**: **Low-Value Customers**

   o **Characteristics**:

   - Lowest average spending (1498.85).

   - Infrequent purchases (Quantity mean = 5.68).

   - Short customer tenure (NumYears mean = 1.23).

   o **Business Insights**:

   - These customers contribute the least to revenue.

   - Target them with cost-effective reactivation strategies, such as seasonal promotions or limited-time offers, to maximize their engagement without significant investment.

Recommendations

1. **Retention and Growth**

   o Prioritize Clusters 0 and 1 by providing them with exclusive benefits to ensure long-term loyalty and increased spending.

2. **Retention for Short-Term Buyers**

- o Focus on Cluster 2 to improve retention rates through targeted campaigns, subscription offers, or reminders for repeat purchases.

3. **Cost-Effective Marketing for Low-Value Segments**

   - o Use Cluster 3 as a testing ground for low-cost marketing initiatives. Monitor their behavior to identify opportunities for reactivation or further segmentation.

4. **Strategic Investments**

   - o Allocate resources toward high-value segments (Clusters 0 and 1) while experimenting with scalable strategies for new or low-value customers.

Conclusion

The clustering analysis suggests that **4 clusters** provide the most meaningful segmentation. These insights highlight the diversity in customer behaviors and spending, enabling tailored strategies to maximize revenue and engagement. The business should focus on nurturing high-value customers while addressing the needs of new or sporadic buyers to ensure sustainable growth.