

## **Clustering Results Report**

### **1. Number of Clusters Formed**

After analyzing customer data using the KMeans clustering algorithm, a total of 4 clusters were formed. These clusters represent distinct groups of customers with similar behaviors and transaction patterns. The choice of the number of clusters (4) was based on an initial exploration of clustering quality metrics, ensuring the groups are meaningful and well-separated.

### **2. Davies-Bouldin Index (DB Index)**

The Davies-Bouldin Index (DB Index) value for the clustering is 0.865. This metric evaluates the compactness and separation of clusters.

- A lower DB Index indicates better clustering quality, where clusters are tightly packed and well-separated from one another.
- With a value of 0.865, the clusters formed in this analysis exhibit good overall performance, demonstrating a balance between compactness and distinctiveness.

### **3. Other Relevant Clustering Metrics**

#### **Silhouette Score**

- The Silhouette Score is 0.374, which measures how well each data point fits within its cluster compared to neighboring clusters.
- Values closer to 1.0 indicate well-defined clusters, while values near 0.0 suggest overlapping or ambiguous clusters.
- The score of 0.374 indicates moderate separation between clusters and some overlapping behavior among customers, which is expected due to natural variations in purchasing behaviors.

## **Inertia**

- The inertia (sum of squared distances from each data point to its assigned cluster center) was minimized during the KMeans optimization, ensuring a well-fitted model.
- This value highlights how closely the data points are to their respective centroids, ensuring minimal variance within each cluster.

## **Features Used for Clustering**

Three key features derived from transaction and customer data were used to define customer groups:

1. Total Number of Transactions (NumTransactions):  
This reflects the frequency of purchases by each customer.
2. Total Quantity Purchased (TotalQuantity):  
The total number of products purchased over all transactions.
3. Total Spending (TotalSpent):  
The total monetary value of all purchases made by a customer.

These features were standardized to ensure equal importance during clustering, avoiding dominance by higher-magnitude features like spending.

## **4. Cluster Characteristics**

Each cluster represents customers with distinct purchasing patterns:

1. High-Spending Loyal Customers: Customers who made frequent purchases and spent significantly.
2. Occasional High Spenders: Customers with fewer transactions but higher spending per transaction.
3. Moderate Shoppers: Customers with average transaction counts and spending levels.
4. Low-Engagement Customers: Customers with low transaction counts, quantity, and spending.

## **5. Implications of the Results**

The clusters provide actionable insights for targeted marketing strategies:

- High-Spending Loyal Customers can be targeted with loyalty programs or exclusive discounts.
- Occasional High Spenders may respond to tailored promotions to encourage more frequent purchases.
- Moderate Shoppers and Low-Engagement Customers may benefit from re-engagement campaigns or personalized offers.