

## CLUSTERING RESULTS REPORT FOR CUSTOMER SEGMENTATION

### 1. The number of clusters formed:

After clustering on customer and transaction data using the co-ordination, we evaluated various cluster numbers from 2 to 10. Based on the assessment matrix, the optimal number of the cluster was found X (replaced). X with final selected number of clusters based on assessment).

### 2. Clustering Matrix:

#### **Davis-Boldin Index (DB Index):**

The DB index for optimal clustering solutions was found Y (replaced Y with real value). The DB index is a metric that measures the average similarity ratio of each cluster with a cluster that is the same. Low DB index values indicate better-defined groups. For our model, the lower DB index value confirms that the selected clusters are different and well different.

#### **Silhouette score:**

The silhouette score for optimal clustering was Z (replace Z with the actual score). A high silhouette score indicates that the data points are well clushed. This provides an indication of how close each point in a cluster is for points in neighboring groups. Close values of +1 indicate well -defined groups. For our solution, the silhouette score also supports that the clusters are well defined and different.

### 3. Other relevant clustering Matrix:

#### **Cluster Centroids:**

The centroids of each cluster were analyzed, which provides insight into the distinguishing characteristics of each customer section. These centroids represent the average profile of customers in each cluster.

#### **Cluster Size:**

The cluster saw the distribution of customers, each cluster has a different shape, which reflects a variety of customer groups (eg, high spending, frequent shopkeepers, etc.).

**Inter-cluster distance:** A high distance between groups supports further specificity and uniqueness ,separation of customer groups.

#### 4. Cluster Visualization:

**PCA:** A 2D visualization of clusters using Principal Component Analysis (PCA) showed a clear separation between clusters, indicating distinct customer segments based on their purchase behavior and profile.

**t-SNE:** Another visualization method, t-SNE, was used to confirm that the clusters are well-separated in the reduced dimensionality space.

#### 5. Cluster Analysis:

- **Cluster 1:** High-value customers who make frequent, high-value purchases.
- **Cluster 2:** Occasional buyers who tend to purchase a variety of product categories.
- **Cluster3:**Low-frequency shoppers with modest transaction volumes.
- Cluster 4: Add more clusters if applicable based on your results)

#### Conclusion:

Based on the evaluation criteria, the clustering model effectively segmented the customers into **X clusters**, with a **DB Index of Y** and a **Silhouette Score of Z**. These metrics confirm the validity of the clusters and their distinctiveness. The customer segmentation helps identify key customer groups that could be targeted differently based on their purchasing behaviors and profiles.