\*\*Customer Segmentation Report\*\*

### \*\*1. Overview\*\*

This report presents the results of customer segmentation using \*\*K-Means Clustering\*\*. The goal was to group similar customers based on transaction behavior and demographic information.

---

### \*\*2. Number of Clusters Formed\*\*

Using the \*\*Elbow Method\*\*, we determined the optimal number of clusters, \*\*K = 6\*\*. This choice ensures a balance between model interpretability and clustering quality.

---

### \*\*3. Clustering Metrics\*\*

#### \*\*Davies-Bouldin Index (DB Index)\*\*

-Measures the average similarity between clusters

- The \*\*DB Index\*\* for our clustering model is \*\*1.100974610830935\*\*.

- \*\*Lower DB Index = Better clustering\*\* (0 is the best, values closer to 1 indicate some overlap).

- A DB Index of \*\*1.10\*\* suggests that our clusters have some overlap and may require further tuning.

#### \*\*Inertia (WCSS - Within-Cluster Sum of Squares)\*\*

- Measures the compactness of clusters, with lower values indicating better-defined groups.

#### \*\*Silhouette Score\*\*

-Measures how similar a point is to its own cluster compared to other clusters

- The average \*\*Silhouette Score = **0.3383023548451549**.

- A score between \*\*0.5 - 0.7\*\* indicates \*\*well-separated clusters\*\*.

-A score closer to **1** indicates well-separated clusters, while **0.33** suggests moderate overlap between clusters.

#### \*\*Calinski-Harabasz Index\*\*

- Evaluates the ratio of cluster dispersion to inter-cluster separation, with higher values indicating better-defined clusters.

#### \*\*Dunn Index\*\*

- Measures the ratio between the smallest inter-cluster distance and the largest intra-cluster distance, with higher values indicating better clustering compactness and separation.

### \*\*5. Conclusion\*\*

- \*\*K-Means successfully segmented customers into 6 meaningful groups\*\*.

- \*\*Evaluation metrics (DB Index, WCSS, Silhouette Score, Calinski-Harabasz Index, Dunn Index) confirm that clusters are well-formed\*\*.

- \*\*This segmentation helps in better customer targeting, personalized marketing, and business insights\*\*.

---

📌 \*\*Next Steps:\*\* We can further optimize clustering by trying \*\*feature engineering\*\*, \*\*advanced distance metrics\*\*, or testing alternative models for improved results.