

# Customer Clustering Report

## 1. Introduction:

The task is on customer segmentation/ clustering to use customer as well as transaction information and the goal is to group customers into clusters based on similarities in their profiles and transactions.

## 2. Clustering Algorithm:

**K-Means** algorithm for clustering is used and the customers are divided into **4 clusters** based on the elbow method.

Number of Clusters: 4

## 3. Clustering Metrics:

- **DB Index:** The DB index helps in evaluating how well-separated and compact the clusters are. A lower value indicates better clustering. **1.23** is the metric which is indicating better clustering.
- **Silhouette Score:** The silhouette score is of **0.47** which indicates clusters are not well defined but moderately defined closer to indicating a good clustering.

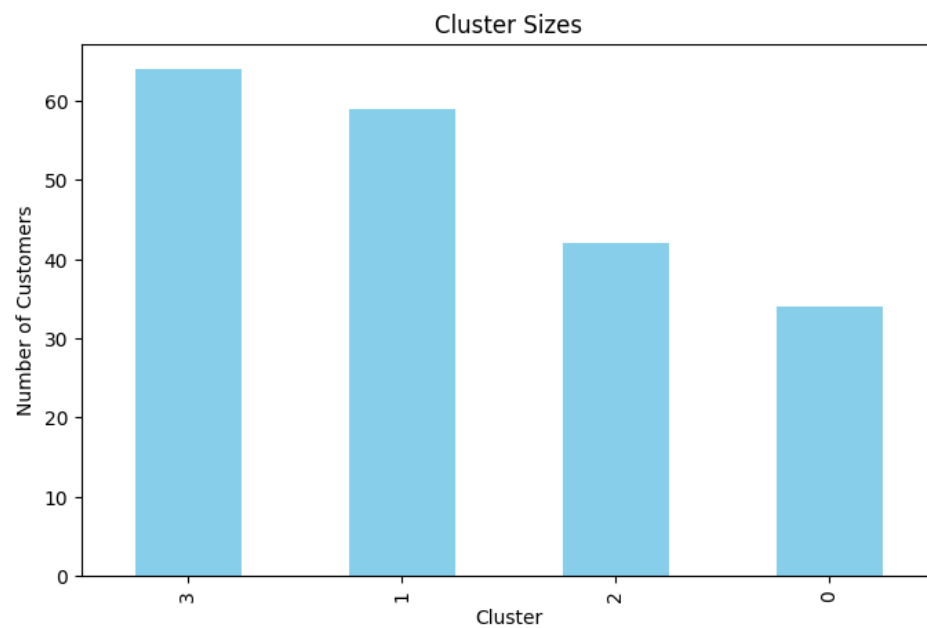
## 4. Cluster Visualizations:

- **PCA Scatter Plot:**



The plot shows that the clusters are separated with a little bit of overlapping.

- Bar chart:



The barchart indicates that the cluster sizes are varied and distinct.

---

## 5. Conclusion:

- Successfully divided the customers into **4 clusters** using K-Means.
- The **Davies-Bouldin Index** (1.23) and **Silhouette Score** (0.47) show that the clusters are fairly formed.
- The **K-Means algorithm** performed reasonably well for segmenting customers. These clusters can now be used for marketing, personalized recommendations, or further analysis of customer behaviour.