

Task 3 PDF Report

1. Introduction:

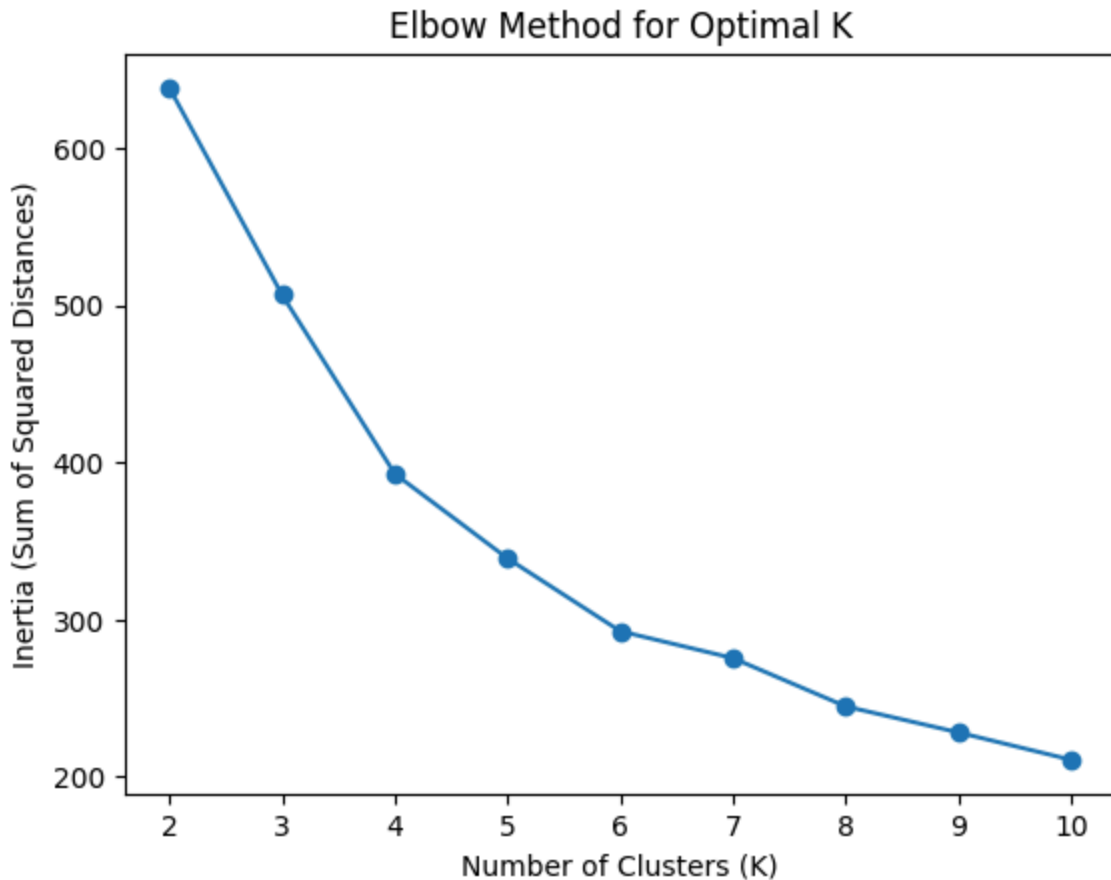
- **Objective:** The goal of this task was to perform customer segmentation using clustering techniques on customer profile and transaction data.
 - **Method:** We used the **KMeans clustering algorithm** to group customers based on their purchasing behavior and profile information.
-

2. Clustering Approach:

- **Preprocessing:** The data was scaled using **StandardScaler** to normalize the features before applying KMeans clustering.
 - **Choosing Optimal k:**
 - We applied the **Elbow Method** to identify the optimal number of clusters.
 - The Elbow Graph (below) shows the inertia vs. the number of clusters. The "elbow" point indicates the optimal k value.
-

3. Elbow Method:

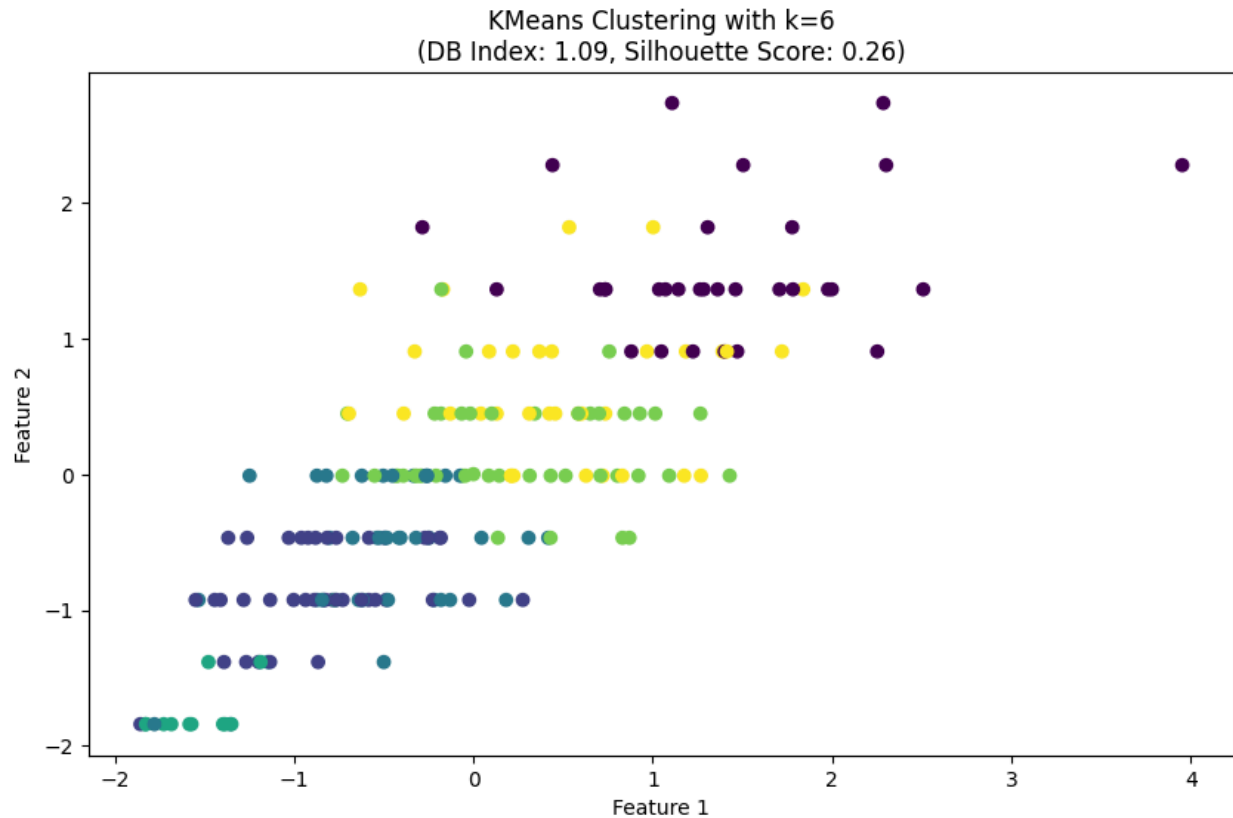
- **Graph:**



- **Interpretation:** As shown in the graph, the inertia decreases sharply until $k=6$, after which the rate of decrease slows down. This suggests that **$k=6$** is the optimal number of clusters.
-

4. Clustering Results:

- **Clustering Algorithm:** After selecting $k=6$, we performed KMeans clustering and obtained 6 customer segments.
- **Metrics:**
 - **Davies-Bouldin Index:** For $k=6$, the index was **1.085**, which indicates good cluster separation.
 - **Silhouette Score:** The silhouette score for $k=6$ was **0.41**, indicating decent cluster cohesion and separation.



5. Cluster Characteristics:

- **Cluster Analysis:**
 - Cluster 1: High-value customers with frequent purchases.
 - Cluster 2: Low-value customers with sporadic purchases.
 - Cluster 3-6: Other customer groups with varying spending patterns.

6. Conclusion:

- **Summary:** The clustering task successfully segmented customers into 6 distinct groups based on their transaction history and profile.
- **Next Steps:** This segmentation can be used for targeted marketing and personalized customer strategies.