TASK 3: Clustering and Segmentation

The purpose of this analysis was to segment the customers into distinct groups based on their purchasing behavior, using clustering techniques. These clusters can be leveraged to develop personalized marketing strategies and improve overall business operations.

Methodology

Data Preprocessing: The dataset was scaled using StandardScaler() to standardize the features for better clustering performance.

Clustering Algorithm: Kmeans Clustering algorithm is used and Elbow method is used to find optimum no of clusters.

Results: Optimal Number of Clusters: Based on the Elbow Method, the optimal number of clusters was determined to be 4. Each cluster represents a unique group of customers with similar purchasing patterns.

Evaluation metrics:

DB Index: evaluates the compactness and separation of clusters; lower values are better.

Silhouette Score: measures the consistency within clusters; ranges from -1 to 1.

Metrics value:

DB Index: 1.2249965469608148

Silhouette Score: : 0.33230070099758946