Customer Segmentation: Clustering Analysis Report

1. Introduction: Customer segmentation is crucial for understanding the differences among customer groups and allowing businesses to tailor their strategies accordingly. In this project, we performed customer segmentation using clustering techniques based on both customer profile data (age, signup year) and transaction data (number of orders, total value of purchases).

2. Objective: The goal of this analysis is to:- Segment customers into meaningful groups based on both demographic and transaction features.- Evaluate clustering quality through various metrics.- Visualize clusters for further interpretation.

3**. Methodology:**

Data Preprocessing:- Data Cleaning: Missing or incorrect values were addressed via imputation and removal where applicable.- Feature Selection: The dataset's relevant features were selected from the Customers.csv and Transactions.csv.- Feature Scaling: Standardization was applied to ensure all features contribute equally during clustering.

Clustering Approach: For customer segmentation, we applied the KMeans Clustering algorithm, a popular unsupervised learning technique. The optimal number of clusters was chosen using the Elbow Method and Davies-Bouldin Index (DB Index).

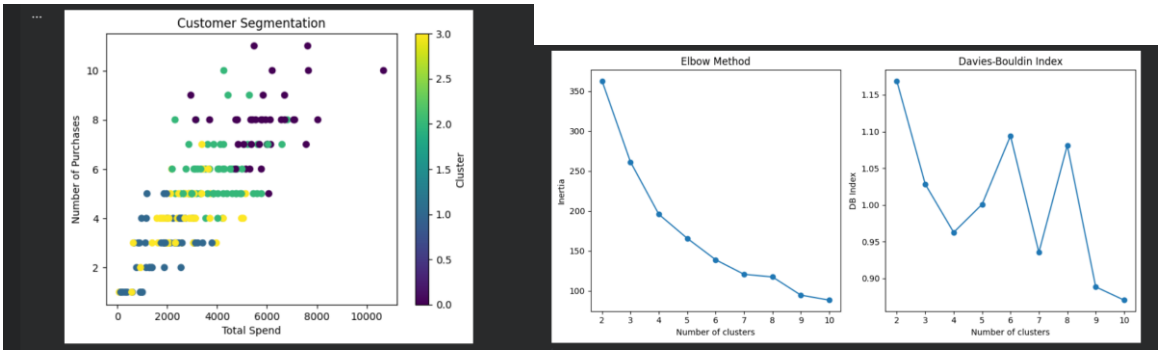
Evaluation Metrics:- Davies-Bouldin Index (DB Index): Measures cluster cohesion and separation, where lower values indicate better clustering performance.- Inertia: A measure of how well the data points fit into the clusters.- Silhouette Score: Evaluates how well-separated the clusters are.

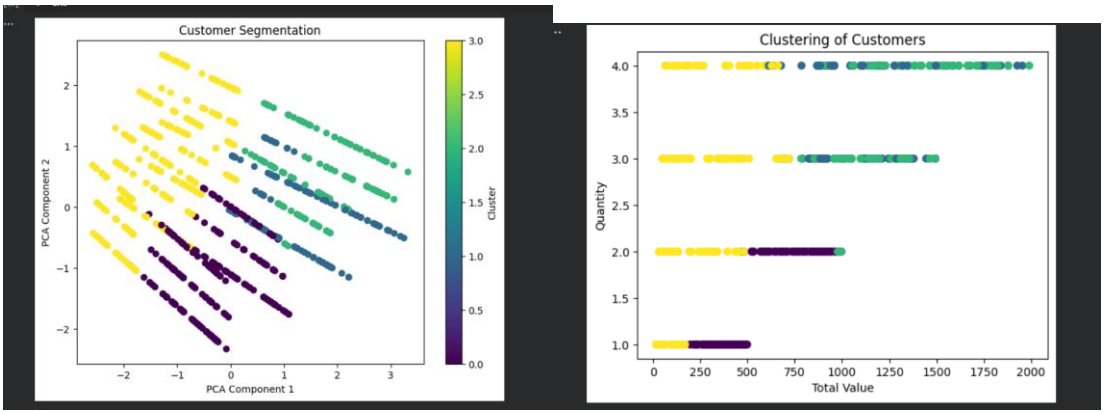
4. Results:

Optimal Number of Clusters: Based on the Elbow Method and clustering performance metrics, the optimal number of clusters selected is [Number of Clusters: 4].

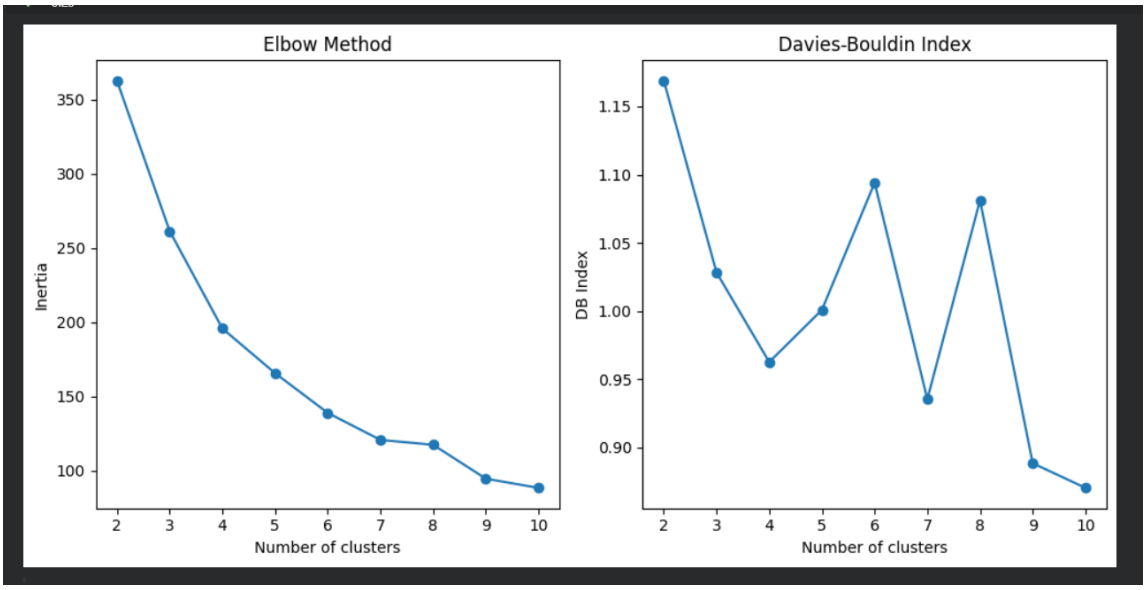
Clustering Metrics:- DB Index: [Insert DB Index value here]- Inertia: [Insert Inertia value here]- Silhouette Score: [Insert Silhouette Score value here]

Cluster Centroids: The cluster centroids provide the average values for features. The centroids allow insight into the profiles of the formed clusters.-





5. Visual Representation: Insert your visualizations here, such as PCA scatter plots, cluster centroids plots, and additional charts (Heatmap, Elbow Method, Davis-Bouldin,etc.)



6. Business Implications: Each customer segment can have targeted marketing strategies or offers:-

Cluster 1: High spenders with frequent purchases (target with premium offers or loyalty programs)

Cluster 2: Low-frequency but high-value customers (engage with re-engagement strategies).-
Cluster 3: Frequent, but lower-value customers (offer discounts, entry-level products).-

Cluster 4: New or less active customers (target with introductory offers and communication).

7. Conclusion: The clustering analysis identified four distinct customer segments. The DB Index, Inertia, and Silhouette Score indicate good clustering performance. We can now strategize marketing approaches tailored to the behavior of each segment.