Customer Segmentation Using Clustering

1. Methods Implemented

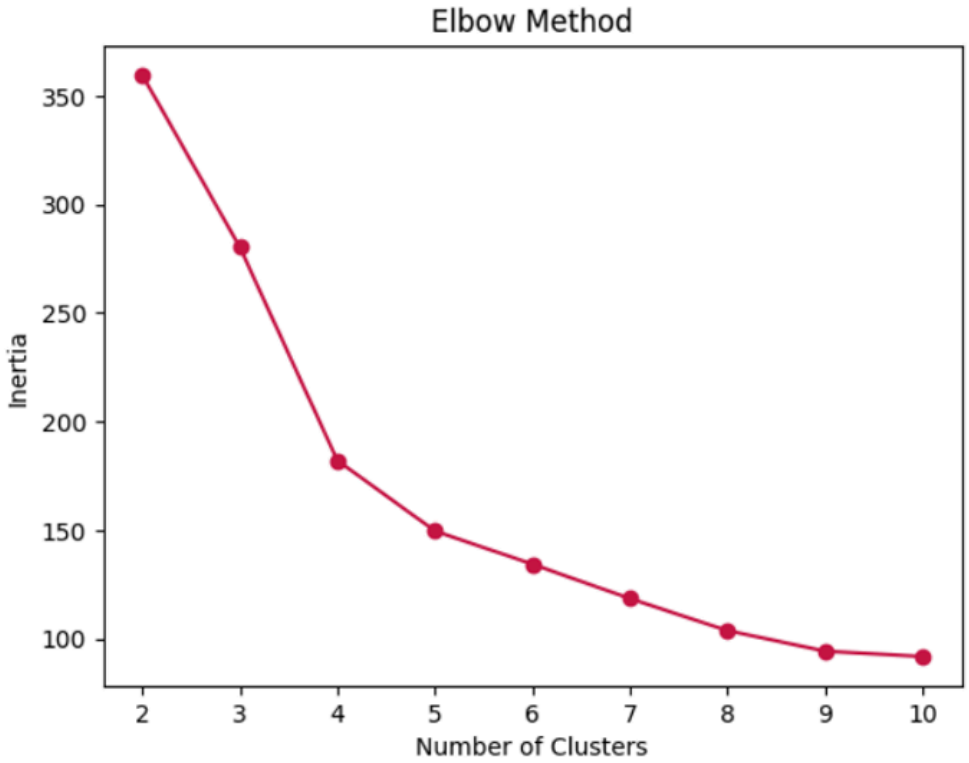
The following clustering techniques were applied:

- KMeans Clustering
- Agglomerative Clustering
- Hierarchical Clustering

2. Methods for Determining Optimal K

Two standard approaches were used to identify the ideal number of clusters:

- **Elbow Method**: Measures variance reduction to determine the point of diminishing returns.



- **Silhouette Method:** Evaluates cluster compactness and separation

3. Feature Engineering

Customer segmentation was performed using RFM Recency, Frequency, Monetary) features, defined as:

- **Recency:** Number of days since the customer's last transaction.
- **Frequency:** Total number of unique transactions made by the customer.
- **Monetary:** Value Sum of all transaction values for each customer.

4. Optimal K Values

- **Elbow Method Suggested K 4.**
- **Silhouette Method Recommended K 5.**

5. Results

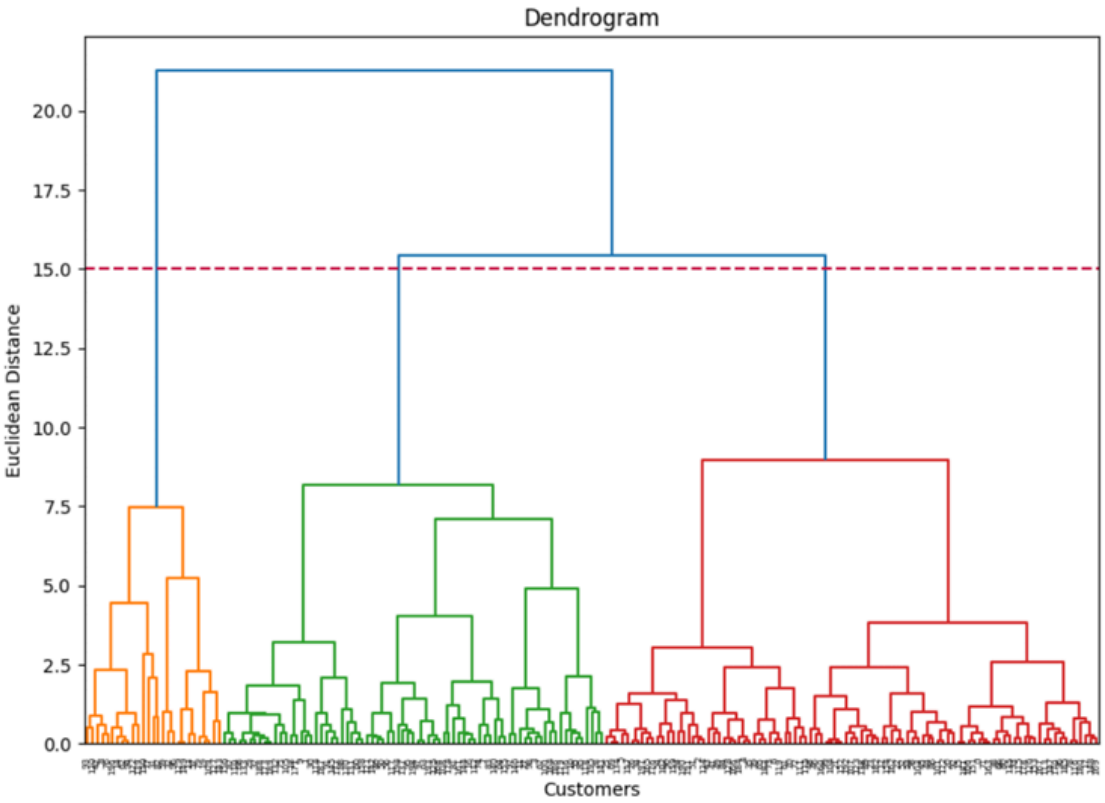
The clustering techniques were evaluated based on the Davies-Bouldin Index DB Index), Silhouette Score, and the number of clusters. The results are summarized below:

Method	DB Index	Silhouette Score	Number of Clusters
KMeans	0.9860	0.2670	5
Agglomerative	0.9860	0.2675	4
Hierarchical	0.9860	0.2675	4

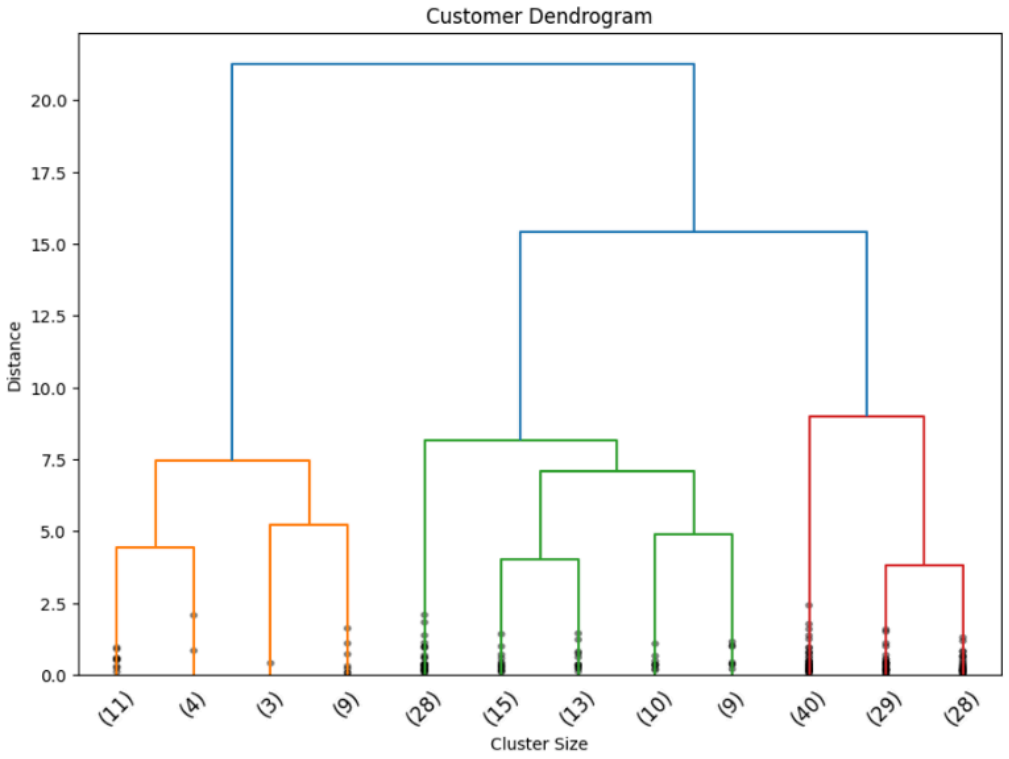
6. Visualizations

Key plots generated during the analysis include:

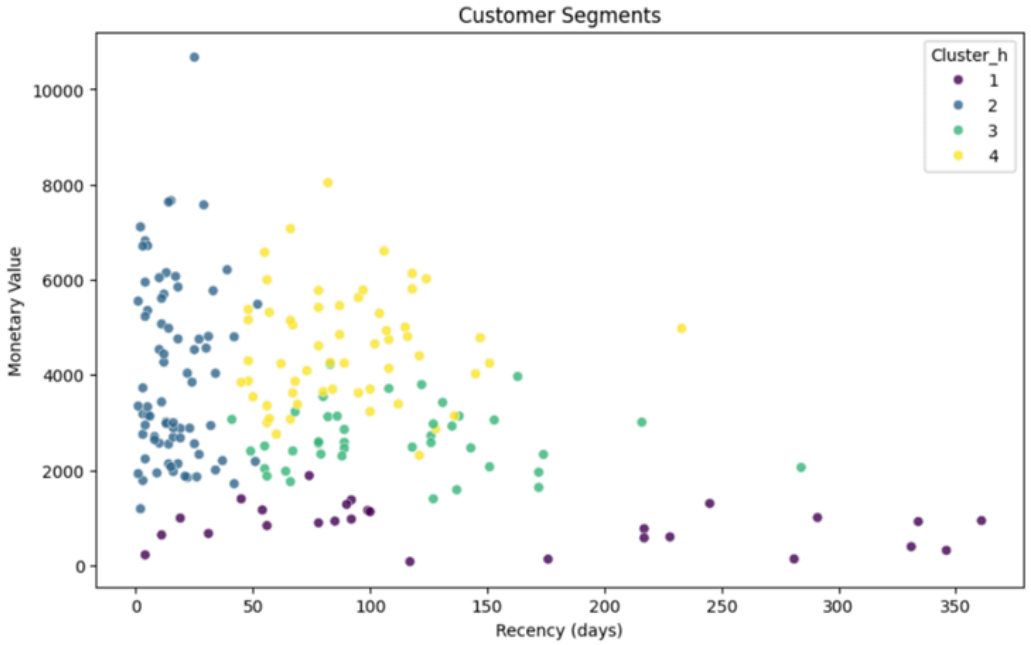
- **Dendrogram:** Visualized cluster hierarchies for Agglomerative Clustering.



- **Cluster Distribution:** Illustrated group sizes in KMeans Clustering.



- **Recency vs. Monetary:** Value Displayed customer segments based on spending patterns and transaction recency.



- Distribution across clusters in **KMeans Clustering**