

## Report on the Customer Segmentation using Clustering

### 1. Methods Tried –

- KMeans Clustering
- Agglomerative Clustering
- Hierarchical Clustering

### 2. Methods for finding optimal K

- Elbow Method
- Silhouette Method

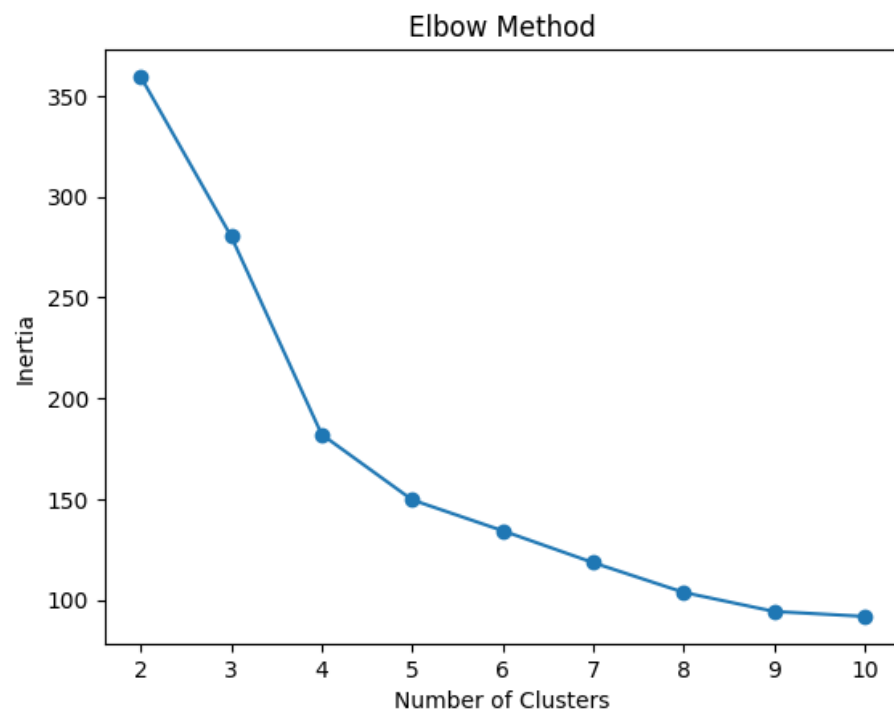
### 3. Feature Engineering

#### Recency, Frequency and Monetary (RFM) based features

1. Recency: This feature indicates the number of days since a customer's most recent transaction.
2. Frequency: This represents the total number of unique transactions a customer has made.
3. Monetary: This measures the total monetary value a customer has spent, calculated by summing the 'TotalValue' of all transactions for each customer.

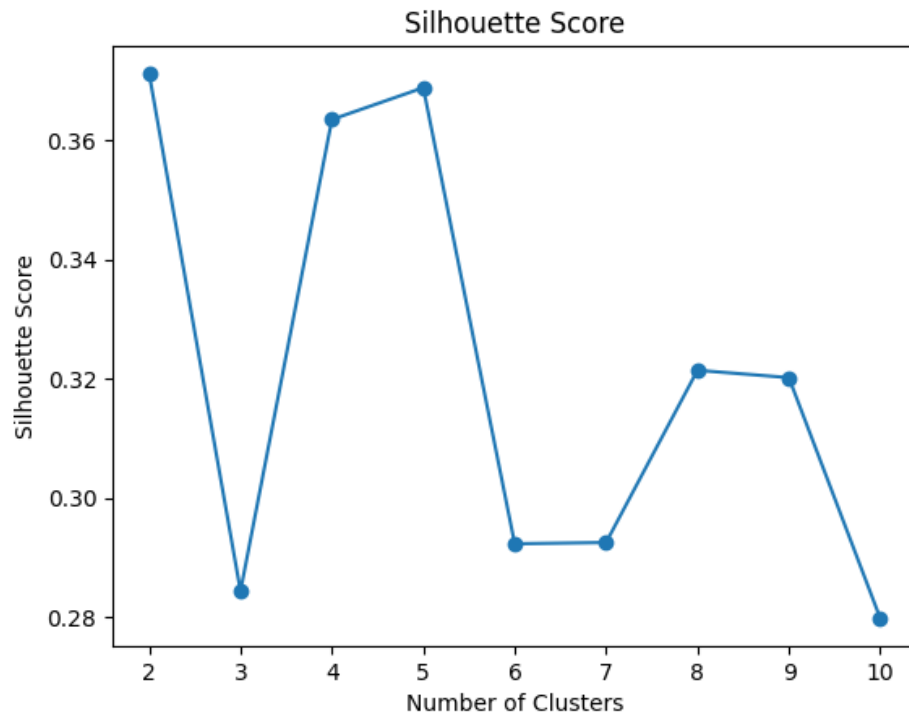
### 4. Plots for optimal K

- Elbow Method



Optimal K = 4

- **Silhouette Method**



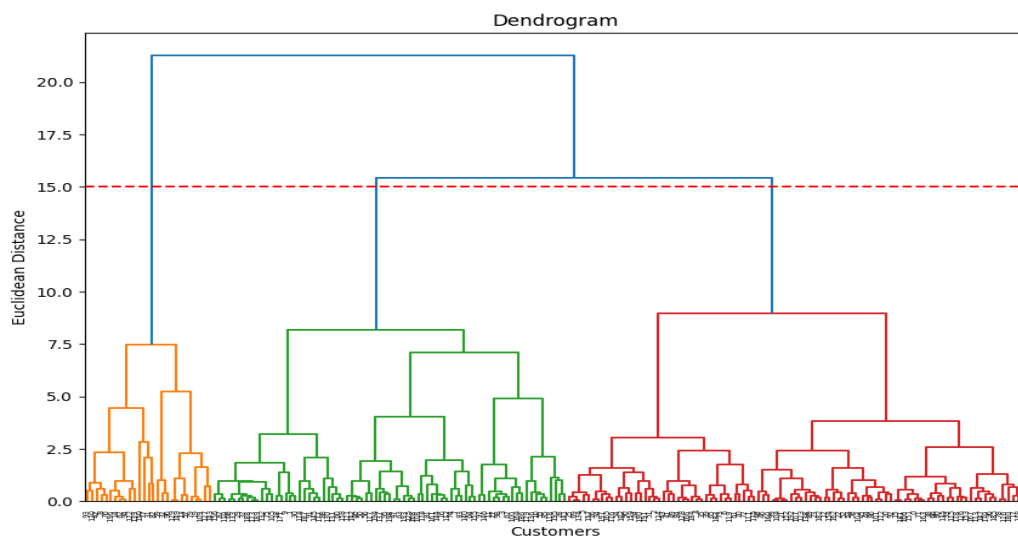
Optimal K = 5

## 5. Results –

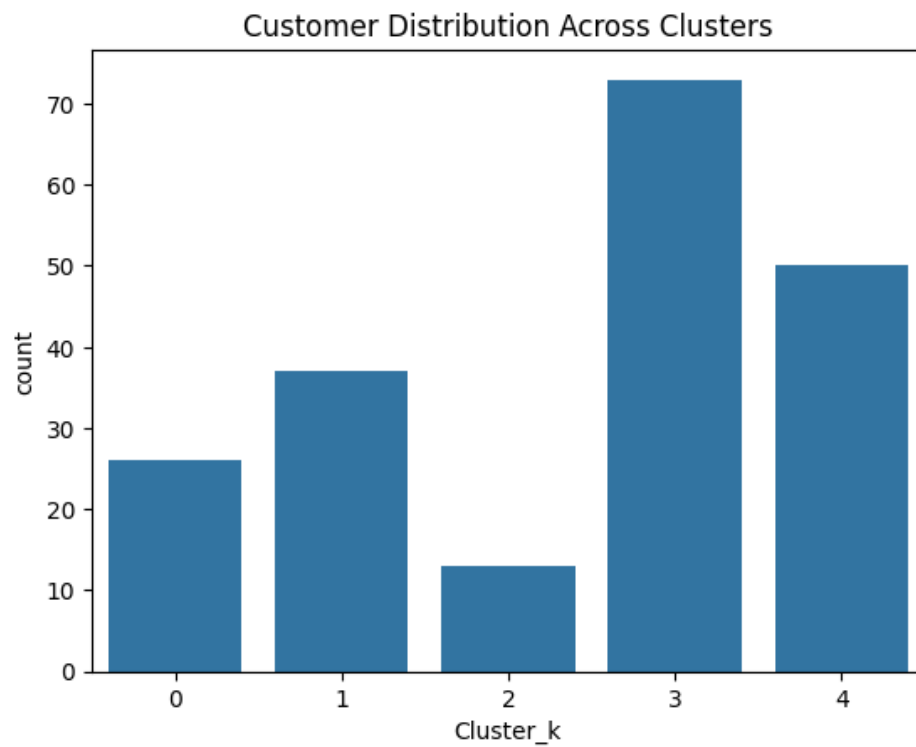
Methods	DB Index	Silhouette Score	Number of clusters
KMeans	0.9860	0.267	5
Agglomerative	0.9860	0.2675	4
Hierarchical	0.9860	0.2675	4

## 6. Plots

### 1. Dendrogram Plot from Agglomerative clustering



## 2. Distribution across clusters in KMeans Clustering



## 3. Customer Segments (Recency Vs Monetary Value)

