

Data Science Internship Assignment

Task 3: Customer Segmentation

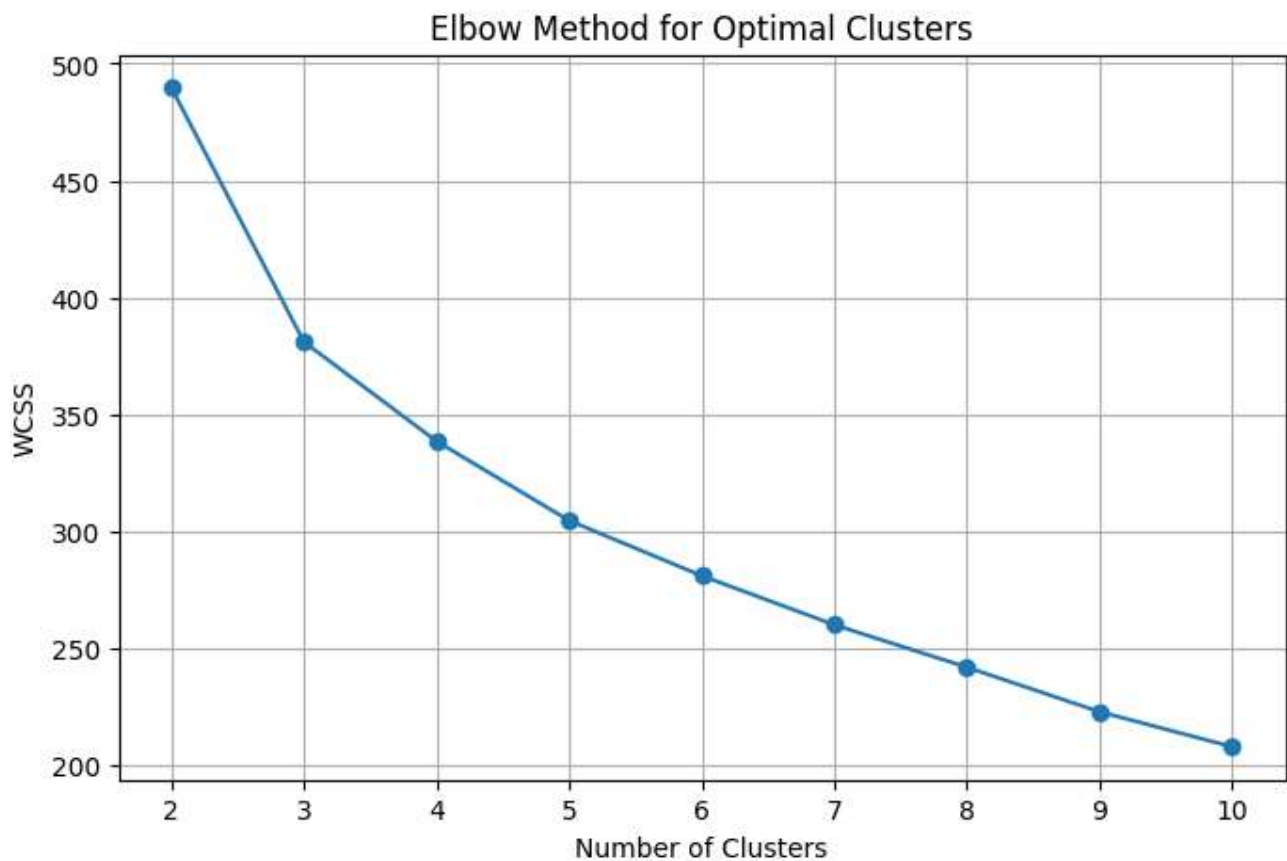
Rushikesh Borade

rishipb19@gmail.com

+91 7038235686

The number of clusters formed:

While performing this task, I used the Elbow method which uses WCSS (Within Cluster Sum of Squares) metric to determine the optimal number of clusters. Following graph shows the optimal point at which the rate of decrease in WCSS value becomes slowest:



Based on the graph, 5 clusters seems to be the most suitable option since the value of WCSS decreases at a slower rate at this point.

Hence, **number of clusters** was selected as **5**

Davies-Bouldin Index:

The DB – Index was calculated to be 1.37, given the number of clusters.

Clusters:

The following distribution of clusters is seen:

