Customer Segmentation / Clustering

-Done By: Rama Krishna Chaithanya Seela

- **Clustering Model used:** KMeans

  - I used Elbow Method to choose the optimal number of clusters. Because In K-Means clustering, we start by randomly initializing k clusters and iteratively adjusting these clusters until they stabilize at an equilibrium point. However, before we can do this, we need to decide how many clusters (k) we should use.

- **The number of clusters formed:** 8

  - The optimal number of clusters is identified by plotting the inertia (sum of squared distances within clusters) against the number of clusters and finding the "elbow point," where adding more clusters results in diminishing returns. The analysis likely showed a significant decrease in inertia at 8 clusters.

- **Davies-Bouldin Index: 0.4644838757580652**

  - Lower DB Index values generally signify better clusters, as they suggest greater intra-cluster similarity and inter-cluster separation. The result of 0.464 suggests that the clustering model has performed well, producing meaningful groupings with minimal overlap. This makes the clusters suitable for applications like customer segmentation, where well-defined groups are crucial for targeted strategies.

- **Silhouette Coefficient: 0.532039360980534**

  - The Silhouette Coefficient evaluates how well each data point fits within its assigned cluster compared to other clusters. This score reflects moderately strong cluster cohesion and separation, suggesting that most points are well-assigned to their respective clusters.