

Report For Customer Segmentation/ Clustering

- ◆ **Library Imports:** Common libraries like pandas, numpy, matplotlib, seaborn and scikit-learn are used for data manipulation and clustering.
- ◆ **Data Loading and Cleaning:** Used the data from Customers.csv and transactions.csv. There were no missing values.
- ◆ **Clustering Methods:** KMeans Clustering was used, the data was segmented with respect to the total money that each customer had spent. Using elbow method we found that 3 is the optimal number of clusters to be formed.
- ◆ **Total number of clusters formed:** 3
- ◆ **Visualization:** Use of scatterplots.
- ◆ **Metrics:** Evaluating clusters using **silhouette score:** 0.5387817110061265 and **Davies-Bloudin Index:** 0.5546790774972087