**Market Segmentation Report**

**Data Import and Preprocessing**

* Imported necessary libraries: pandas, numpy, matplotlib, seaborn, StandardScaler from sklearn.preprocessing, PCA from sklearn.decomposition, and various clustering and classification algorithms from sklearn.cluster and sklearn.tree.
* Suppressed warnings for cleaner output.
* Loaded the dataset from "/content/sample\_data/Customer Data.csv".
* Checked the shape, information, and summary statistics of the dataset.

**Exploratory Data Analysis (EDA)**

* Visualized the distribution of numerical features using kernel density estimation (KDE) plots.
* Displayed the distribution of each numerical feature using histograms.
* Plotted the correlation matrix heatmap to understand the pairwise correlation between features.

**Clustering**

* Scaled the data using StandardScaler.
* Performed Principal Component Analysis (PCA) and reduced the dimensionality to 2 for visualization purposes.
* Employed the Elbow Method to determine the optimal number of clusters for K-Means.
* Applied K-Means clustering with 4 clusters.
* Visualized the clusters in the PCA space.
* Examined the cluster centers and assigned cluster labels to the original dataset.
* Saved the K-Means model using joblib and exported the clustered dataset to a CSV file.
* Classification
* Prepared the data for classification by separating features and labels and splitting into training and testing sets.
* Trained a Decision Tree Classifier on the training data.
* Evaluated the model's performance using confusion matrix and classification report.
* Saved the trained model using pickle.

**Conclusion**

The analysis involved preprocessing, EDA, clustering, and classification.

A comprehensive pipeline was built for understanding the dataset, clustering customers, and predicting customer clusters using a decision tree classifier.