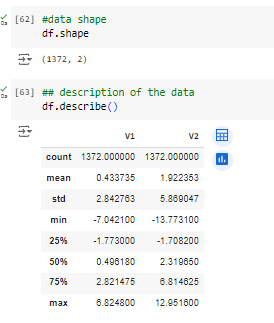
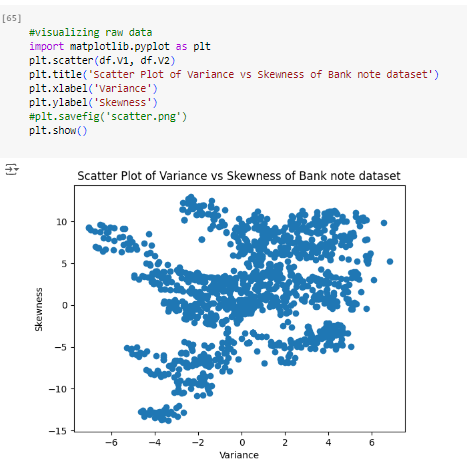
Clustering project 5\_2

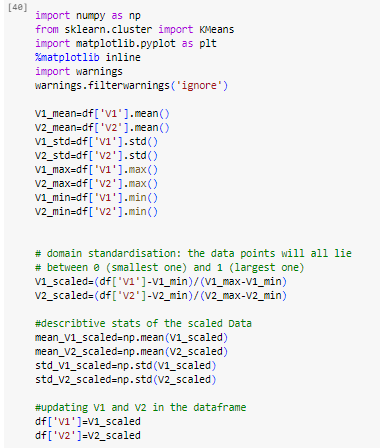
1-I loaded the Banknote authentication dataset. I used pandas library to load csv file. and I Calculated Statistical Measures to describe the data and visualized raw data using scatter plot



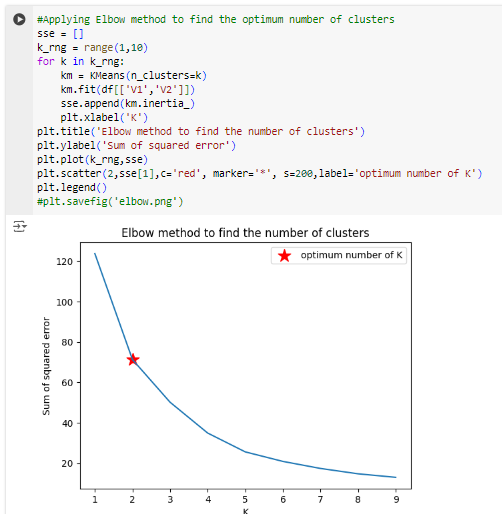




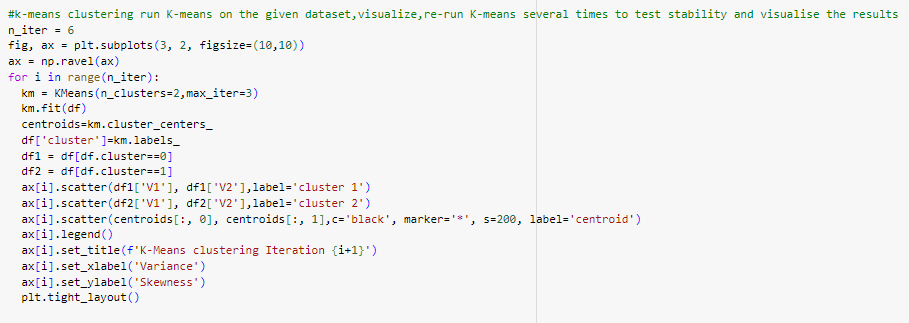
The Banknote Authentication dataset contains 1372 instances with 2 features: variance, skewness of wavelet-transformed images of banknotes. statistical measures, e.g. mean and standard deviation were calculated. These two features got standardized from 0 to 1

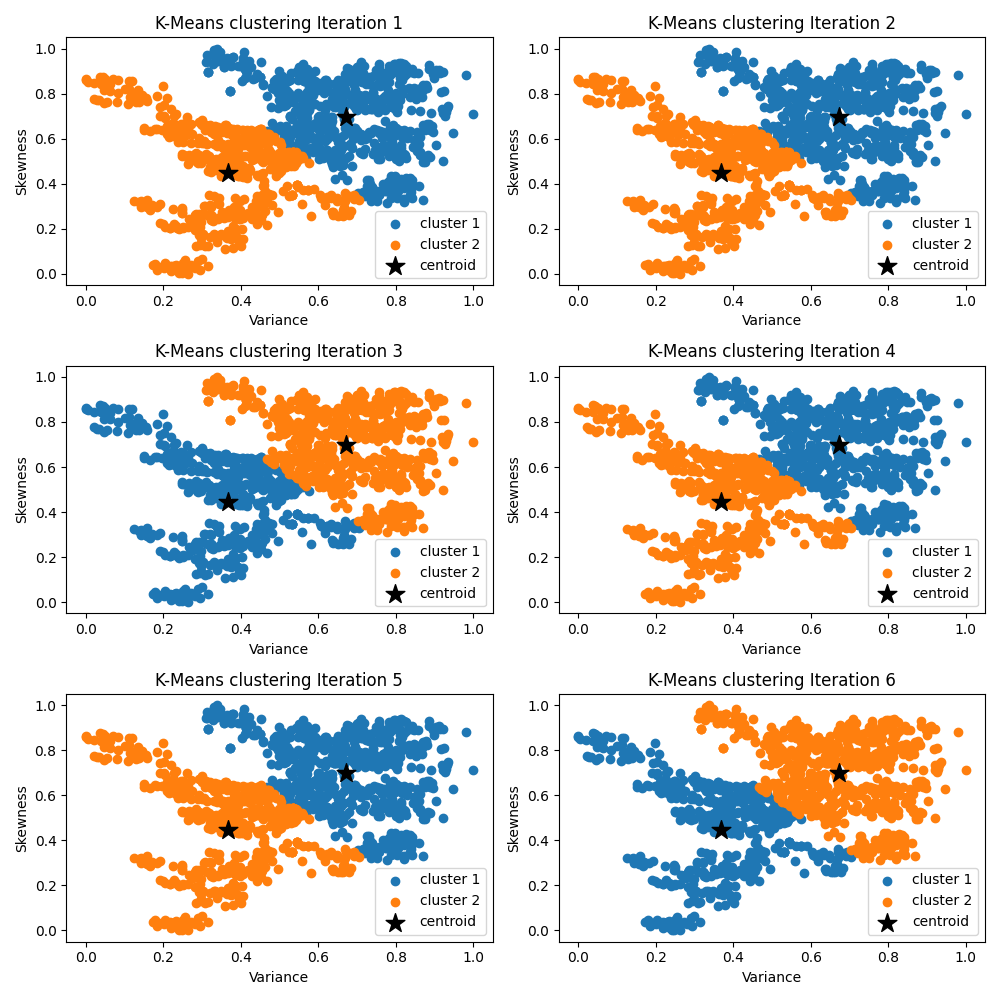


2- I used scatter plot to graph the normalized data of two features Variance Vs Skewness of Bank note. I added the centroids resulted from the k\_mean clustering.I choosed number of clusters=2 after applying Elbow method to determine the optimum number of clusters.



3- I visualize the clustering results using a scatter plot and To check the stability of the K-means algorithm, I re-run it multiple times with different random states and visualize the results using a scatter plot





5- Compare the visualizations from different runs. the clusters are consistent across different random states. the algorithm is stable

6-Brief describtion

“In this assignment, I used the Banknote authentication dataset to train a K-means clustering model.I chosed number of clusters to be 2 after Applying the Elbow method to find the optimum K number. I ran the K-means algorithm multiple times with different random states and visualized the results. The clusters were generally consistent across different runs, indicating that the K-means algorithm is relatively stable for this dataset. The visualizations showed clear separation between the clusters, suggesting that the features used are effective for distinguishing between genuine and fake banknotes.”