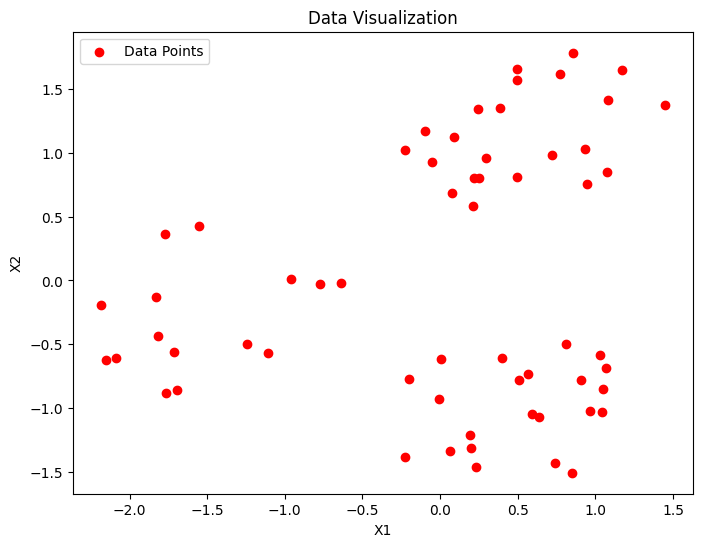
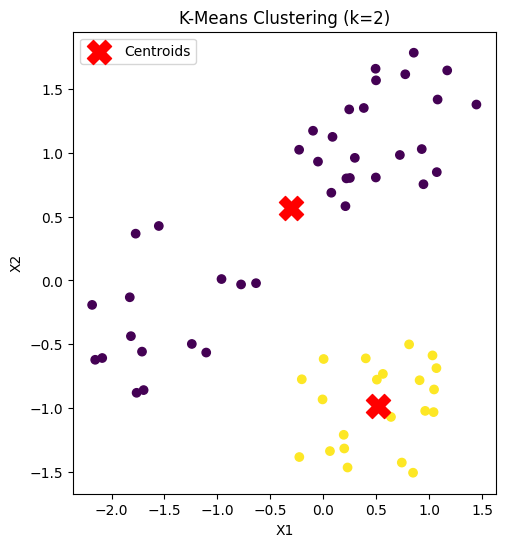
Name: Shreyanshu Yadav

Roll: 21052793

ML Assignment 3

The notebook performs k-means clustering on a dataset loaded from a CSV file kmeans.csv.

* The dataset is first normalized using the normalize function, which calculates the mean and standard deviation of the dataset and normalizes the data by subtracting the mean and dividing by the standard deviation.
* The kmeans function implements the k-means clustering algorithm. It takes the dataset, number of clusters k, and maximum number of iterations as input.
* The initialize\_centroids function initializes the centroids randomly by selecting k indices from the dataset without replacement.
* The assign\_clusters function assigns each data point to the closest centroid based on Euclidean distance.
* The update\_centroids function updates the centroids by calculating the mean of all data points assigned to each centroid. If no points are assigned to a centroid, it is randomly reinitialized.
* The normalized dataset is visualized using a scatter plot with x1 and x2 as the x and y axes, respectively.
* 
* The k-means clustering algorithm is run for k=2 and k=3, and the resulting clusters and centroids are visualized using scatter plots.
* The centroids and clusters for k=2 and k=3 are stored in centroids\_2, clusters\_2, centroids\_3, and clusters\_3, respectively.
* 
* 