Project Overview:

In this part, you are required to implement the k-means algorithm and apply your implementation on the given dataset, which contains a set of 2-D points. You are required to implement two different strategies for choosing the initial cluster centers.

Strategy 1: randomly pick the initial centers from the given samples.

Strategy 2: pick the first center randomly; for the i-th center (i>1), choose a sample (among all possible samples) such that the average distance of this chosen one to all previous (i-1) centers is maximal.

You need to test your implementation on the given data, with the number k of clusters ranging from 2-10. Plot the objective function value vs. the number of clusters k. Under each strategy, plot the objective function twice, each start from a different initialization.

Algorithms:

k-Means Clustering

Resources:

A 2-D dataset to be downloaded from this link: [Dataset.](https://canvas.asu.edu/courses/45793/files/11859729/download?wrap=1)

Workspace:

Any Python programming environment.

Software:

Python environment.

Language(s):

Python. (MATLAB is equally fine, if you have access to it.)