'''

DS 5230

Summer 2022

HW2_Problem_1: K-Means Theory

Hongyan Yang

'''

A) prove that E step update on membership (\pi) achieves the minimum objective given the current centroids (\mu)

   For any observation xi that changed membership at E step. Its contribution to the objective function changed from (xi – mu) ^ 2 to (xi – mu*) ^ 2 where mu* is the new centroid that xi changed membership to. Because mu* is the centroid closest to xi rather than mu or other centroids, the objective function is guaranteed to decrease at E step. Thus, update on membership achieves the minimum objective.

B) prove that M step update on centroids (\mu) achieves the minimum objective given the current memberships (\pi)

   For any random selected "centroid" mu in each cluster k, its contribution to the objective function equals (x1 - mu) ^ 2 + (x2 – mu) ^ 2 + … + (xn – mu) ^ 2. After expanding this function of mu and taking mu's first derivative, we find that the objective function exists a minimum value when mu equals the mean of observations in cluster k, which exactly equals the updated centroid at M step. Thus, update on centroid achieves the minimum objective.

C) Explain why K-Means has to stop (converge), but not necessarily to the global minimum objective value.

   Given an initial cluster assignment of each observation, the objective function keeps decreasing with the iterations of E steps and M steps, which have been proved to achieve the minimum objective at each iteration. So, the clustering obtained will continually improve until the result no longer changes. However, because the objective function of K-Means is not convex. The result obtained will depend on the initial (random) cluster assignment of each observation and K-Means will converge to a local optimum but not necessarily to the global optimum.