Question 1

$\sum_{i,i'_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = 2 \cdot \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_k j)^2$
the number of observations within each cluster is donated by$|C_K|$, and the value of mean is
$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{C_k}$
we can substitute the mean with the summation and then
we have the value for the left hand side of the equation except for the constant 2
$\frac{1}{|C_k|} \sum_{i,i'_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = \frac{2}{|C_k|} \cdot \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x'_i j)^2$

K-means algorithm reduces the objective function since its an iterative process first it assigns random means to the clusters and then group points according to how close they are to the specific cluster and then it will find a new cluster means that reduces the objective function within the cluster, after that the k-means re-assign the point to the cluster based on how close the points are to the new-mean, and the algorithm only stops when there objective function doesnot change - no more improvment - or reached a specific threshold

Question 2

(a) no, we cannot use k-means clustering with this dataset since k-means assumes the classes to have spherical distribution, and then try to find centers with neat spheres around them, and given this situation, it will definitely fail to do so. It will be able to minimize the within-cluster sum of squares but will mix the different groups together.

(b) we recommend using the average linkage method, for non-spherical data, the average linkage method is considered to be the most robust linkage method, as it is less sensitive to noise and outliers, and it often produces compact and well-separated clusters.