CS 5565, ECE5590CI, CS465R Clustering practice

1. (k-means clustering)
   Given the following data points with initial class assignments of either Red or Blue, use the k-means algorithm to assign points to clusters using *squared Euclidean distance* as the measure.

B	(1,1)
R	(2,2)
R	(3,3)
B	(4,4)
R	(1,6)
R	(3,5)
B	(4,7)

2. (Dendrogram (single linkage)) Suppose you have the following data points:

	X	Y
P1	0.40	0.53
P2	0.22	0.38
P3	0.35	0.32
P4	0.26	0.19
P5	0.08	0.41
P6	0.45	0.30

   Construct a dendrogram using single linkage using:

   (a) Euclidean distance.

   (b) Manhattan distance.

   Construct a dendrogram using complete linkage using:

   (a) Euclidean distance.

   (b) Manhattan distance.

3. Note: On the exam you will be given the distance matrix for the dendrograms, so you will not need to compute distances. This is not an assignment, is not to be handed in, and will not be graded. You only need to complete enough to understand how the algorithms work.