

Clustering:

1. Use single and complete link agglomerative clustering to group the data described by the following distance matrix. Show the dendrograms.

	A	B	C	D
A	0	1	4	5
B		0	2	6
C			0	3
D				0

2. Suppose that the task is to cluster the following eight points (with (x,y) representing location) into three clusters:

$A_1(2,10), A_2(2,5), A_3(8,4), A_4(5,8), A_5(7,5), A_6(6,4), A_7(1,2), A_8(4,9)$

The distance function is Euclidean distance. Suppose initially we assign  $A_1, A_4$  and  $A_7$  as the center of each cluster, respectively. Use the k-means algorithm to show only the three cluster centers after the first round execution.

3. Assume the following dataset is given: (2,2), (4,4), (5,5), (6,6), (9,9) (0,4), (4,0) . K-Means is run with  $k=3$  to cluster the dataset. Moreover, Manhattan distance is used as the distance function to compute distances between centroids and objects in the dataset. Moreover, K-Mean's initial clusters C1, C2, and C3 are as follows:

C1: {(2,2), (4,4), (6,6)}

C2: {(0,4), (4,0)}

C3: {(5,5), (9,9)}

Now K-means is run for a single iteration; what are the new clusters and what are their centroids?