

1. You have four medicines, each having two attributes X (weight index) and Y (pH). The data is given below:

Object	Attribute 1 (X): weight index	Attribute 2 (Y): pH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

You know that there are two clusters in the dataset. You can assume the initial value of the cluster centers to be $c_1 = (1, 1)$ and $c_2 = (2, 1)$.

Run the k-means clustering algorithm using

- Manhattan distance
- Euclidean distance

In how many steps does the algorithm converge and find the final cluster assignments.

2. You are given 8 data points and their 2 attributes as:

$A1=(2,10)$, $A2=(2,5)$, $A3=(8,4)$, $A4=(5,8)$, $A5=(7,5)$, $A6=(6,4)$, $A7=(1,2)$, $A8=(4,9)$.

To help you, the Euclidean distance between each data point is given below:

	A1	A2	A3	A4	A5	A6	A7	A8
A1	0	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$
A2		0	$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$
A3			0	$\sqrt{25}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{53}$	$\sqrt{41}$
A4				0	$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$
A5					0	$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$
A6						0	$\sqrt{29}$	$\sqrt{29}$
A7							0	$\sqrt{58}$
A8								0

a. Assuming that $k=3$ and the following points as initial centers: A1, A4, and A7. Run the k-means algorithm and output the final centers and the final cluster assignments.

b. Does it make a difference in the result if you assume the following initial centers: A2, A5, and A8.

3. Perform agglomerative clustering on the following dataset using single and complete link approaches. Output the resulting dendograms.

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

4. Perform agglomerative clustering on the following data points using single link, complete link, and average link approaches.

$A1=(2,10)$, $A2=(2,5)$, $A3=(8,4)$, $A4=(5,8)$, $A5=(7,5)$, $A6=(6,4)$, $A7=(1,2)$, $A8=(4,9)$.