

Homework- Kmeans-Clustering

Deadline: Thursday, (22-12-2022)

Consider we have the following points:

point	X	Y	Z
A	2	10	4
B	2	5	5
C	8	4	6
D	5	8	8
E	7	5	1
F	6	4	2
G	1	2	1
H	4	9	2

Using Manhattan distance

$$\sum_{i=1}^d |a_i - b_i|$$

Using K-Means clustering, cluster the given data into 3 clusters.

Initial centroids are points (2,10,4), (5,8,8), and (1,2,1)

- A- What will be the output clusters for the first, second, third iterations, show your calculations?
- B- What is the SSE for the clusters formed after the third iteration.