K-Means clustering:

K-means is a clustering algorithm whose main goal is to group similar elements or data points into a cluster.

Note: 'k' in K-means represents the no. of clusters.

Eg: Suppose that the data mining task is to cluster points into three clusters.

Dataset: $A_1=(2,10)$; $A_2=(2,5)$, $A_3=(8,4)$
$B_1=(5,8)$; $B_2=(7,5)$; $B_3=(6,4)$
$C_1=(1,2)$; $C_2=(4,9)$

Initially A_1, B_1, C_1 points are centers of each cluster. Use Euclidian distance formula. Find distances from each centroid.

Initial centroids:
A_1: (2,10)
B_1: (5,8)
C_1: (1,2)

Datapoints			Distance to				Cluster	new clust	
			(2,10)	(5,8)	(1,2)		I_1	I_2	I_3
A_1	2	10	0.00	3.61	8.06		1	1	1
A_2	2	5	5	4.24	3.16		3	3	3
A_3	8	4	8.49	5.00	7.28		2	2	2
B_1	5	8	3.61	0.00	7.21		2	2	1
B_2	7	5	7.07	3.61	6.71		2	2	2
B_3	6	4	7.21	4.12	5.39		2	2	2
C_1	1	2	8.06	7.21	0.00		3	3	3
C_2	4	9	2.24	1.41	7.62		(2) → (1)		1

↓ final cluster

New centroids are
A_1: (2,10)
B_1: (6,6)
C_1: (1.5,3.5)

Current centroids:

New centroids:
A_1: (3,9.5)
B_1: (6.5,5.25)
C_1: (1.5,3.5)

Cluster 1: A_1, B_1, C_2
Cluster 2: A_3, B_2, B_3
Cluster 3: A_2, C_1