P_1	2, 10	P_9	10, 12
P_2	2, 6 (Cluster 1)	P_{10}	7, 5
P_3	11, 11	P_{11}	9, 11
P_4	6, 9	P_{12}	4, 6
P_5	6, 4	P_{13}	3, 10
P_6	1, 2	P_{14}	3, 8
P_7	5, 10 (Cluster 2)	P_{15}	6, 11 (Cluster 3)
P_8	4, 9		

Ques: Perform K-Means Clustering on the given dataset. Assume $k=4$ and given initial cluster centres.

Ans

	D/C_1 (2,6)	D/C_2 (5,10)	C_3 (6,11)	C_4 (6,9)	Centroid
P(2,10)	4	3	4.123	4.123	2
(2,6)	0	5	6.403	5	1
(11,11)	10.29	6.082	5	5.385	3
(6,9)	5	1.414	2	0	4
(6,4)	4.472	6.082	7	5	1
(1,2)	4.123	8.944	10.29	8.602	1
(5,10)	5	0	1.414	1.414	2
(4,9)	3.605	1.414	2.828	2	2
(10,12)	10	5.385	4.123	5	3
(7,5)	5.099	5.385	6.08	4.123	4
(9,11)	8.602	4.123	3	3.605	3
(4,6)	2	4.123	5.38	3.605	1
(3,10)	4.123	2	3.162	3.162	2
(3,8)	2.23	2.82	4.24	3.162	1
(6,11)	6.403	1.41	0	2	3

$C_1 \rightarrow$ New Centroid $= \left(\frac{2+6+1+4+3}{5}, \frac{6+4+2+6+8}{5}\right)$

$= \left(\frac{16}{5}, \frac{26}{5}\right) = (3.2, 5.2)$

$C_2 \rightarrow \left(\frac{2+5+4+3}{4}, \frac{10+10+9+10}{4}\right) = \left(\frac{14}{4}, \frac{39}{4}\right) = (3.5, 9.75)$

$C_3 \rightarrow \left(\frac{11+10+9+6}{4}, \frac{11+12+11+11}{4}\right) = \left(\frac{36}{4}, \frac{45}{4}\right) = (9, 11.2)$

$C_4 \rightarrow \left(\frac{6+7}{2}, \frac{9+5}{2}\right) = \left(\frac{13}{2}, \frac{14}{2}\right) = (6.5, 7)$

ITERATION 2

	(3.2, 5.2)	(3.5, 9.75)	(9, 11.25)	(6.5, 7)	Centroid
(2, 10)	4.94	1.5206	7.11	5.408	2
(2, 6)	1.44	4.0388	8.75	4.609	1
(11, 11)	9.72	7.603	2.015	6.0207	3
(6, 9)	4.72	2.61	3.75	2.061	4
(6, 4)	3.046	6.269	7.846	3.041	4
(4, 2)	3.883	8.143	12.229	7.43	1
(5, 10)	5.126	1.5206	4.19	3.354	2
(4, 9)	3.883	0.901	5.4829	3.201	2
(10, 12)	9.6166	6.87	1.25	6.103	3
(7, 5)	3.805	5.9	6.56	2.06	4
(9, 11)	8.202	5.64	0.25	4.716	3
(4, 6)	1.131	3.78	7.25	2.69	1
(3, 10)	4.804	0.559	6.128	4.609	2
(3, 8)	2.807	1.82	6.823	3.64	2
(6, 11)	6.44	2.795	3.01	4.03	2

C1 → $\left(\frac{2+1+4}{3}, \frac{6+2+6}{3}\right) = \left(\frac{7}{3}, \frac{14}{3}\right) = (2.33, 4.66)$

C2 → $\left(\frac{2+5+4+3+3+6}{6}, \frac{10+10+9+10+8+11}{6}\right) = (3.833, 9.66)$

C3 → $\left(\frac{11+9}{2}, \frac{11+11}{2}\right) = (10, 11)$

C4 → $\left(\frac{6+6+7}{3}, \frac{9+4+5}{3}\right) = \left(\frac{19}{3}, \frac{18}{3}\right) = (6.33, 6)$

ITERATION 3

	(2.33, 4.66)	(3.833, 9.66)	(10, 11)	(6.33, 6)	CLUSTER
(2, 10)	5.35	1.864	8.06	5.894	2
(2, 6)	1.138	4.093	9.43	4.33	1
(11, 11)	10.74	7.291	1	6.841	3
(6, 9)	5.68	2.265	4.47	3.018	2
(6, 4)	3.72	6.06	8.062	2.027	4
(1, 2)	2.97	8.167	12.727	6.66	1
(5, 10)	5.97	1.215	5.099	4.21	2
(4, 9)	4.65	0.6808	6.324	3.798	2
(10, 12)	10.61	6.596	1	7.033	3
(7, 5)	4.68	5.634	6.708	1.203	4
(9, 11)	9.202	5.337	1	5.668	3
(4, 6)	2.141	3.66	7.810	2.33	1
(3, 10)	5.38	0.899	7.071	5.204	2
(3, 8)	3.406	1.857	7.615	3.884	2
(6, 11)	7.325	2.547	4	5.0108	2

C1 → $\left(\frac{2+1+4}{3}, \frac{6+2+6}{3}\right) = (2.33, 4.66)$

C2 → $\left(\frac{2+6+5+4+3+3+6}{7}, \frac{10+9+10+9+10+8+11}{7}\right) = (4.14, 9.57)$

C3 → (10, 11)

C4 → $\left(\frac{6+7}{2}, \frac{4+5}{2}\right) = (6.5, 4.5)$

ITERATION 4

	(2.33, 4.66)	(4.14, 9.57)	(10, 11)	(6.5, 4.5)	CLUSTER
(2,10)	5.35	2.18	8.06	7.106	2
(2,6)	1.38	4.16	9.43	4.74	1
(11,11)	10.74	7.007	1	7.905	3
(6,9)	5.68	1.945	4.47	4.527	2
(6,4)	3.72	5.87	8.062	0.707	4
(1,2)	2.97	8.195	12.727	6.041	1
(5,10)	5.97	0.961	5.099	5.7008	2
(4,9)	4.65	0.586	6.324	5.147	2
(10,12)	10.61	6.343	1	8.27	3
(7,5)	4.68	5.391	6.708	0.707	4
(9,11)	9.202	5.066	1	6.964	3
(4,16)	2.141	3.572	7.810	2.915	1
(3,10)	5.38	1.218	7.071	6.519	2
(3,8)	3.406	1.94	7.615	4.949	2
(6,11)	7.325	2.34	4	6.519	2

Since the new clusters are exactly the same as in iteration 3, hence the centroids also remain same. Thus we have reached the optimal clustering for k=4.

Hence, final centroids are :

C1 → (2.33, 4.66)

C2 → (4.14, 9.57)

C3 → (10, 11)

C4 → (6.5, 4.5)