NAME - Himanshu.

Roll no. - 23/4017

Exam Roll no. - 23019582065.

SEM - 5th.

YEAR - 3rd.

Course - B.Sc. Computer Science.

Qs-1 Given

- Data points: {44, 28, 48, 26, 32, 14, 52, 50}.
- No. of clusters, K = 2
- Initial cluster centers: $C_1 = 5$, $C_2 = 38$

→ Iteration 1: Assign points to nearest ~~center~~ cluster.

Distance formula: $\boxed{d = |x - c|}$

→ clusters Centers

(i) $C_1 = 5$, $C_2 = 38$.

Point	Distance to 5.	Distance to 38	cluster.
44	39	6	2
28	23	10	2
48	43	10	2
26	21	12	2
32	27	6	2
14	9	24	1
52	47	14	2
50	45	12	2

Cluster assignments:

→ Cluster 1: {14}.

→ Cluster 2: {26, 28, 32, 44, 48, 50, 52}.

Update cluster Centers. C_1 = mean of cluster 1 = 14

C_2 = mean of cluster 2 = $\frac{26+28+32+44+48+50+52}{7} = \frac{280}{7} = 40$

Iteration 2: Assign points to nearest cluster (new centers)

clusters centers: (i) $C_1 = 14$, (ii) $C_2 = 40$.

Point	Distance to 14	Distance to 40	cluster.
44	30	4	2
28	14	12	2
48	34	8	2
26	12	14	1 (vs 2)

→ check 26:

· Distance to 14 = 12, · Distance to 40 = 14 → closest to 14. → cluster 1.

Point	Distance to 14	Distance to 40	cluster.
32	18	8	2
14	0	26	1
52	38	12	2
50	36	10	2.

→ cluster assignments:

(i) cluster 1: {14, 26}, (ii) cluster 2: {28, 32, 44, 48, 50, 52}.

update cluster centers.

$$C_1 = \frac{14+26}{2} = 20.$$

$$C_2 = \frac{28+32+44+48+50+52}{6} = \frac{254}{6} \approx 42.33.$$

Iteration 3: Assign points to nearest clusters (new centers).

cluster centers:

(i) $C_1 = 20$, (ii) $C_2 = 42.33$.

cluster assignments:-

· cluster 1 = {14, 26, 28}., cluster 2, {32, 44, 48, 50, 52}.

Compute SSE (Sum of Squared Errors).

$$SSE = \sum (x_i - \text{cluster center})^2$$

cluster 1: $(14-25)^2 + (26-25)^2 + (28-25)^2 + (32-25)^2 = (-11)^2 + 1^2 + 3^2 + 7^2 = 121 + 1 + 9 + 49 = 180.$

cluster 2: $(44-48.5)^2 + (48-48.5)^2 + (50-48.5)^2 + (52-48.5)^2$
$= (-4.5)^2 + (-0.5)^2 + 1.5^2 + 3.5^2 = 20.25 + 0.25 + 2.25 + 12.25 = 35.$

Total SSE = 180 + 35 + 215.

Q5-2 Iteration 1

Centroids: $C_1 = (185, 72)$, $C_2 = (170, 56)$.

Cluster assignment (nearest centroid by Euclidean distance).

- cluster 1: C1, C4, C5, C6.
- cluster 2: C2, C3.

Update centroids:

- C_1 = mean of cluster 1 = $((185+179+188)/4, (72+68+72+77)/4) = (183.5, 72.25)$.
- C_2 = mean of cluster 2 = $((170+168)/2, (56+60)/2) = (169, 58)$

SSE · cluster 1 ≈ 50.75, · cluster 2 ≈ 10.

Total SSE ≈ 60.75.

Iteration 2.

centroids: $C_1 = (183.5, 72.25)$, $C_2 = (169, 58)$.

cluster assignments: same as Iteration 1 → converged.

SSE: same as Iteration 1 → ≈ 60.75.