Q4 — Clustering (12 points)

Tie-breaking rule: In case of ties, assign to cluster with lowest centroid norm.

Part 1 (4 points) — Hierarchical Clustering

Points: {2, 5, 10, 14, 17, 26}. Repeatedly merge the two nearest clusters (single linkage).

Step	Clusters before merge	Min distance	Merge
1	{2}, {5}, {10}, {14}, {17}, {26}	d(2,5) = 3 and d(14,17) = 3 (tie)	Merge {2} and {5} → {2,5}
2	{2,5}, {10}, {14}, {17}, {26}	d(14,17) = 3	Merge {14} and {17} → {14,17}
3	{2,5}, {10}, {14,17}, {26}	d(10,{14,17}) = 4	Merge {10} and {14,17} → {10,14,17}
4	{2,5}, {10,14,17}, {26}	d({2,5},{10,14,17}) = 5	Merge → {2,5,10,14,17}

$$\boxed{\text{Final clusters: } \{2, 5, 10, 14, 17\} \text{ and } \{26\}}$$

Part 2 (4 points) — K-Means (initial centroids: 5, 17)

Iteration 1 — Assign

Point	d(·, 5)	d(·, 17)	Cluster
2	3	15	C_1
5	0	12	C_1
10	5	7	C_1
14	9	3	C_2
17	12	0	C_2
26	21	9	C_2

C_1 = {2, 5, 10}, C_2 = {14, 17, 26}

Iteration 1 — Update centroids

- $\mu_1 = (2 + 5 + 10)/3 = 17/3 \approx 5.67$
- $\mu_2 = (14 + 17 + 26)/3 = 57/3 = 19$

Iteration 2 — Assign

Point	d(·, 17/3)	d(·, 19)	Cluster
2	$11/3 \approx 3.67$	17	C_1
5	$2/3 \approx 0.67$	14	C_1
10	$13/3 \approx 4.33$	9	C_1
14	$25/3 \approx 8.33$	5	C_2
17	$34/3 \approx 11.33$	2	C_2
26	$61/3 \approx 20.33$	7	C_2

Assignments unchanged → **converged.**

$$\boxed{\text{Final clusters: } \{2, 5, 10\} \text{ and } \{14, 17, 26\}}$$

Note: Hierarchical and K-means produce different results on the same data!

Part 3 (4 points) — Algorithm Selection for Figure 1

(a) Left dataset — Two concentric rings

Answer: Hierarchical clustering

K-means assigns points to the nearest centroid, producing convex (Voronoi) partitions. It cannot separate two concentric rings — both rings share approximately the same center, so K-means would cut through both with a linear boundary. Hierarchical clustering with single linkage follows chain-like connectivity within each ring and correctly identifies the two ring-shaped clusters.

(b) Right dataset — One large cluster + three small clusters

Answer: Hierarchical clustering (None also accepted)

There are 4 visually distinct groups but we need K=2 clusters. K-means with K=2 would place two centroids and split the space with a Voronoi boundary, likely cutting through or merging groups in unnatural ways — it cannot discover that the natural grouping has 4 clusters when forced to use 2. Hierarchical clustering (single linkage) can at least produce a dendrogram reflecting the true cluster structure, and cutting at the right level could yield a reasonable 2-way split (e.g., the large cluster vs. the three small clusters grouped together). However, since neither algorithm is ideal for forcing 4 natural groups into 2 clusters, "None" is also accepted.