**K-Means**

1. How many starting configurations are there?
2. To initialize the 3-means algorithm, choose 3 distinct starting points from the given 6 points (a, b, c, d, e, f) in the 2D plane.

The number of possible starting configurations (i.e., distinct sets of 3 points from 6) is given by the combination formula:

6/3 =6! /3! (6-3)! = 20

There are 20 possible starting configurations when choosing 3 unique points from the 6 to initialize the KMeans algorithm.

Each configuration can lead to a different initial partitioning, and potentially, a different final clustering depending on the symmetry of the points and the behavior of the algorithm.

1. What are the stable 3-partitions?
2. A stable partition is a final cluster assignment where the algorithm no longer changes the groupings upon further iterations i.e., the cluster centers and memberships stay fixed.

A graph with numbers and symbols

AI-generated content may be incorrect.

Stable Partition (1): (a, d), (b, e), (c, f)

This partition follows a vertical alignment – left, right and center columns grouped.

A graph with black dots and orange dots

AI-generated content may be incorrect.

Stable Partition (2): Points (a, b), (c, e, f), (d)

This is an example of an imbalanced but stable configuration.

A graph with different colored dots

AI-generated content may be incorrect.

Stable Partition (3): Points (a, d), (b, c), (e, f)

This configuration aligns horizontally (Top and Bottom pairing)

* Discovered 7 unique stable 3-paritions.
* Indicating that different initial centroid choices can lead to different final clusterings, even for such a small, symmetrical dataset.

1. What is the number of starting configurations leading to each of the stable 3-partitions?
2. Based on the simulation of 20 possible initial centroid combinations, this is the count:

|  |  |  |
| --- | --- | --- |
| Partition | Cluster Grouping (Point Indices) | No. of Starting Configurations |
| 1 | (0, 3), (1, 4), (2, 5) | 8 |
| 2 | (0,), (1,2, 4, 5), (3,) | 2 |
| 3 | (0, 3), (1, 2), (4, 5) | 2 |
| 4 | (0, 1), (2, 5), (3, 4) | 4 |
| 5 | (0, 1, 3), (2,), (4, 5) | 1 |
| 6 | (0, 1, 3, 4), (2,), (5,) | 2 |
| 7 | (0, 3, 4), (1, 2), (5,) | 1 |

* The most frequent occurring stable configuration is Partition 1, appearing in 8 out of 20 runs.
* Several partitions are rare, occurring only once showing that the final clustering is highly dependent on the initial centroid selection.
* The variety in stable partitions even for a small, symmetrical dataset shows how initialization can strongly influence KMeans behavior.

1. What is the maximum number of iterations from any starting configuration to its stable 3-partition?
2. The maximum number of iterations required for KMeans to converge to a stable 3-partiton from any of the 20 initial centroid configurations is 2.

Every run (all 20 combinations) converged in exactly 2 iterations.

This quick convergence is expected due to:

This simplicity of the dataset (only 6 points)

Symmetrical layout in 2D Space.

The fact that centroids were initialized on actual data points.

This demonstrates how KMeans can converge rapidly on small, structured datasets and also highlights that convergence does not guarantee global optimality, only stability based in initialization.