K = 2

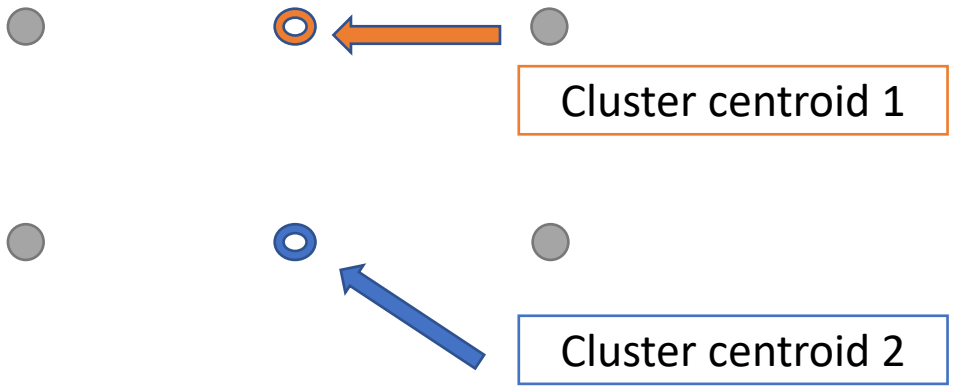
Let’s pretend we have only two four points we are trying to cluster, with k = 2.

K = 2

We can see that there are two clusters, and the most logical grouping are the two points on the left, and the two points on the right.

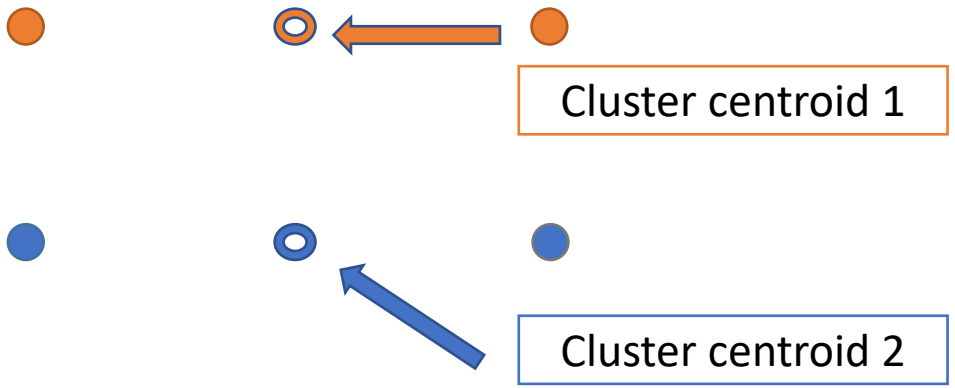
But is k-means guaranteed to find this solution?

K = 2



What happens if the cluster centroids are initialized at these points?

K = 2



These become our final state clusters! K-means is a stochastic algorithm, so our results will depend on how the cluster centroids are initialized. There is always the danger of falling into a local minimum.