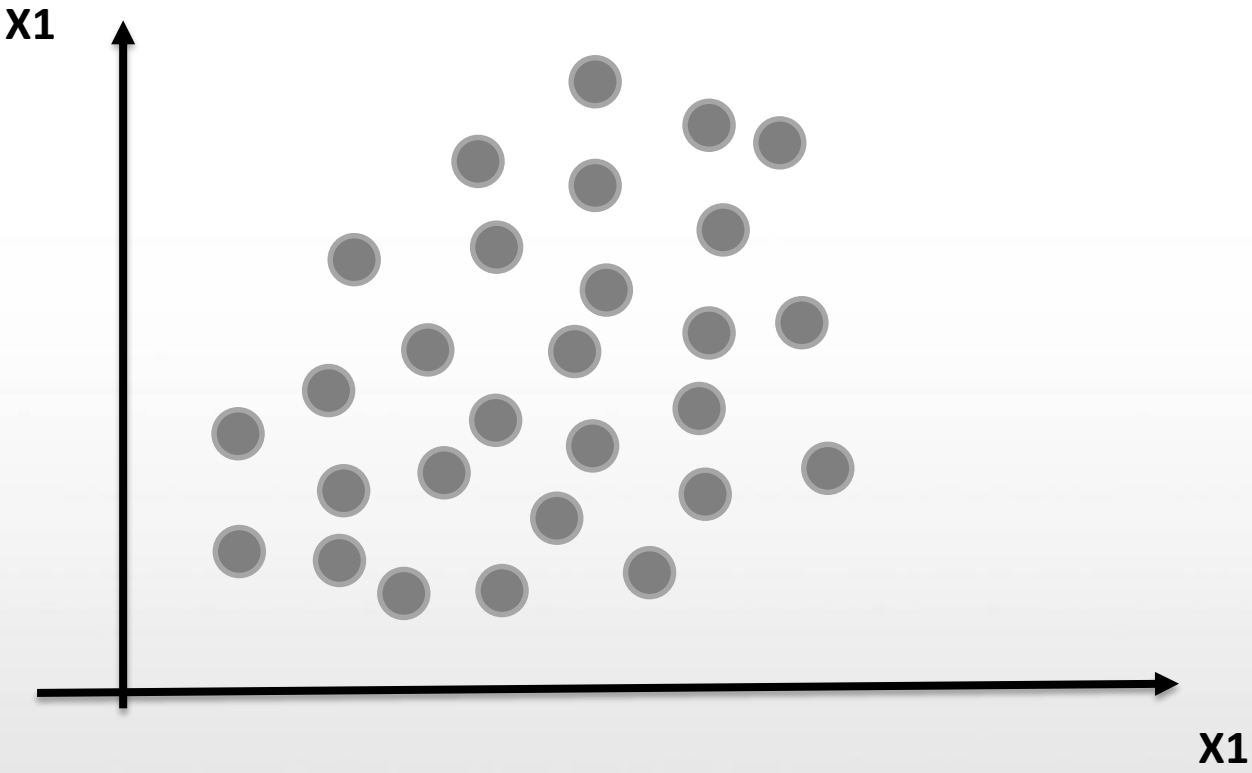
Unsupervised Learning

K-Means Clustering

K Means Clustering

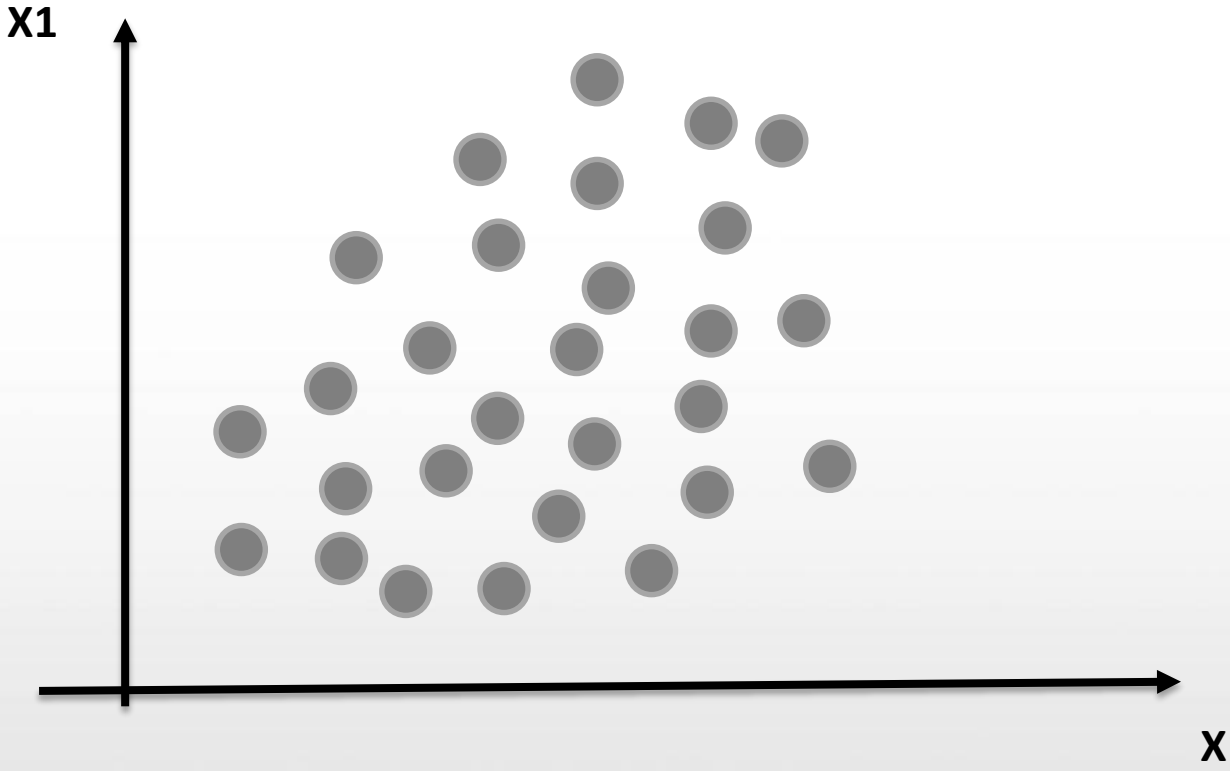


Steps in K -Means

- Step 1 : Choose the number K of clusters
- Step 2: Select at random K points, the centroids
- Step 3: Assign each data point to closest centroid
- Step 4: Computer and place the new centroid of each clusters
- Step 5: Reassign each data point to the closest centroid. Go to step 4

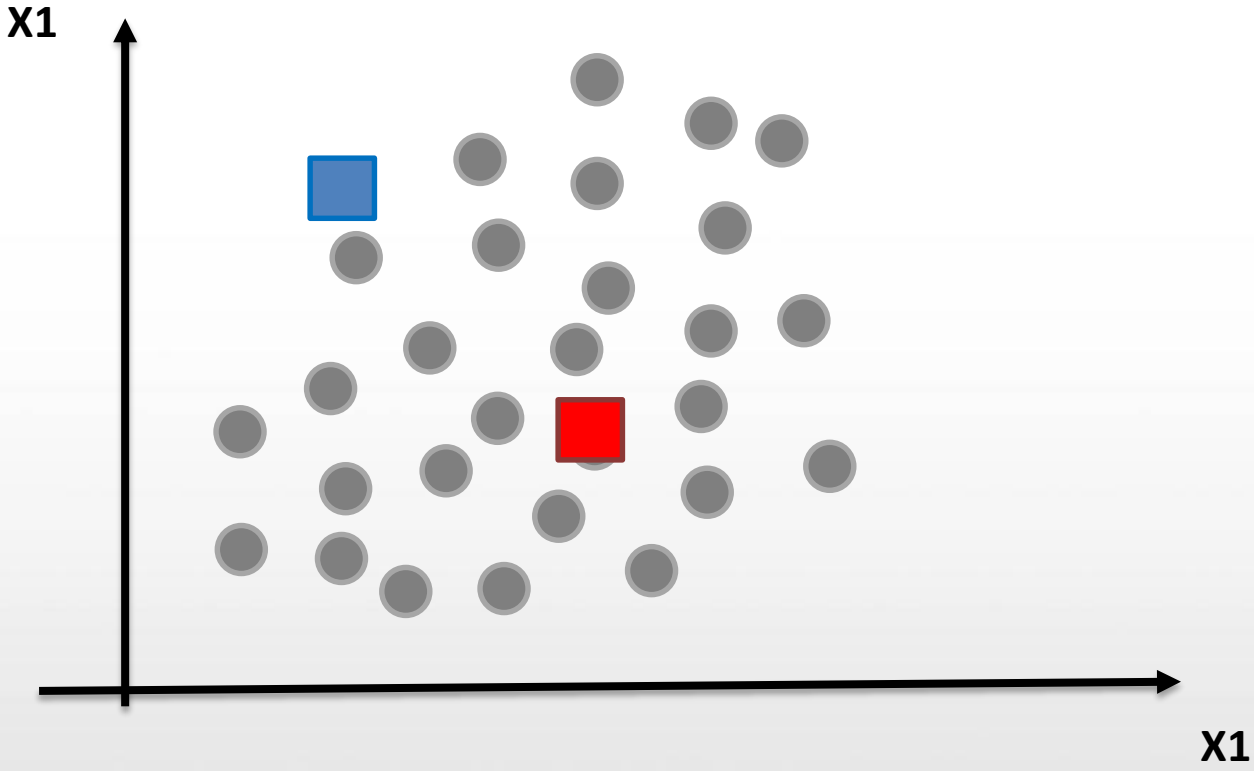
K Means Clustering

Step 1: choose number K of clusters : K = 2



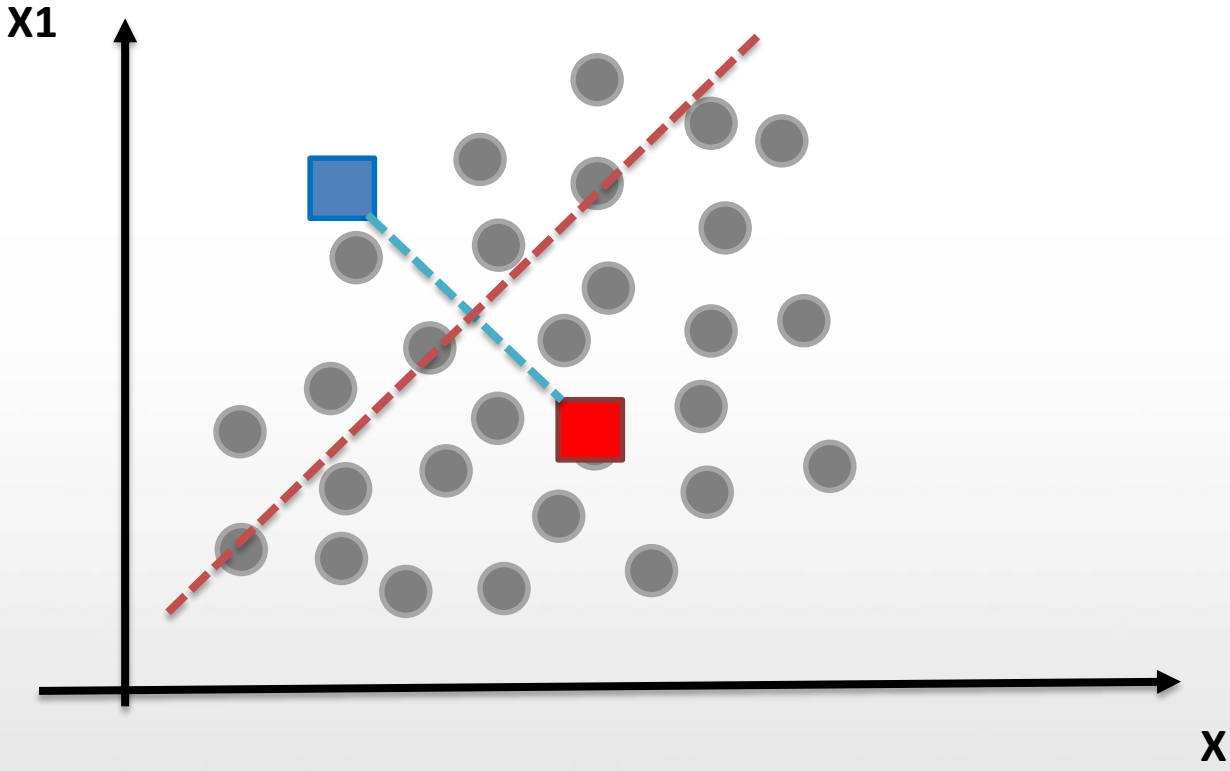
K Means Clustering

Step 2: Select random K points, centroid



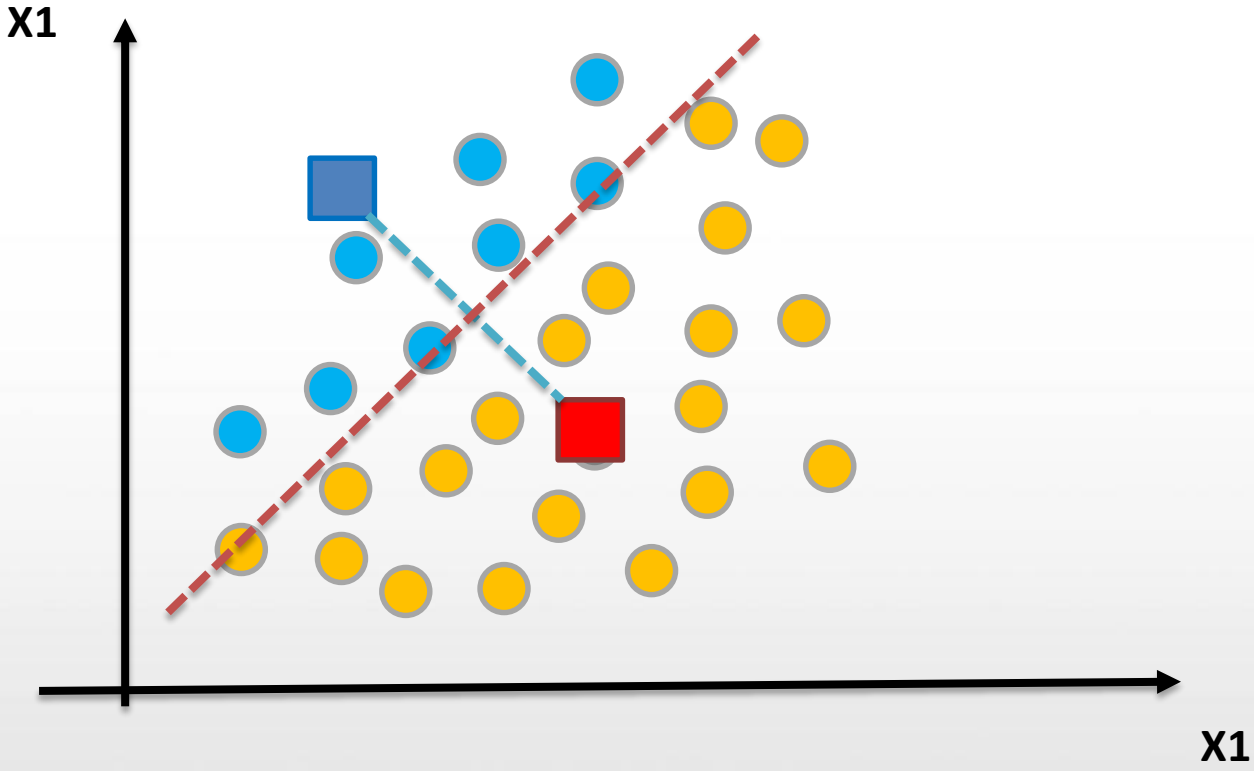
K Means Clustering

Step 2: Assign each data point to the closest centroid → that forms K clusters

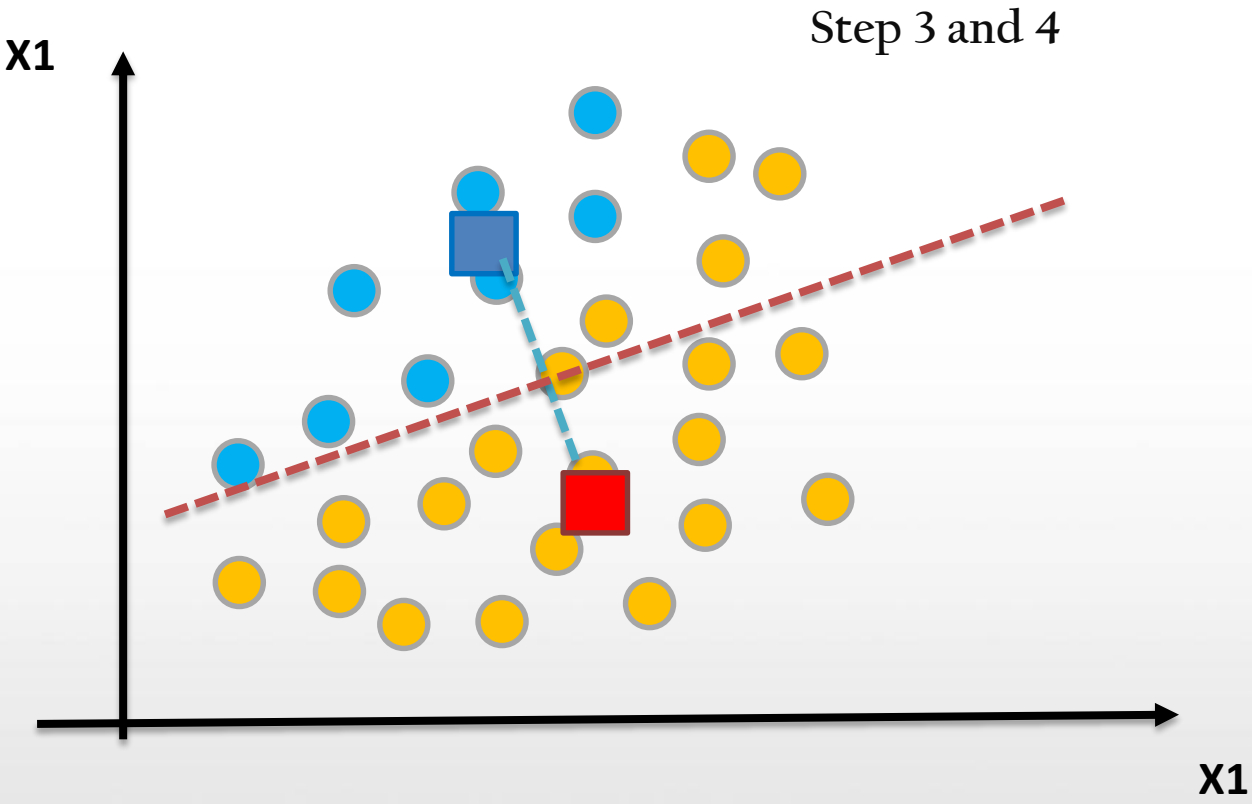


K Means Clustering

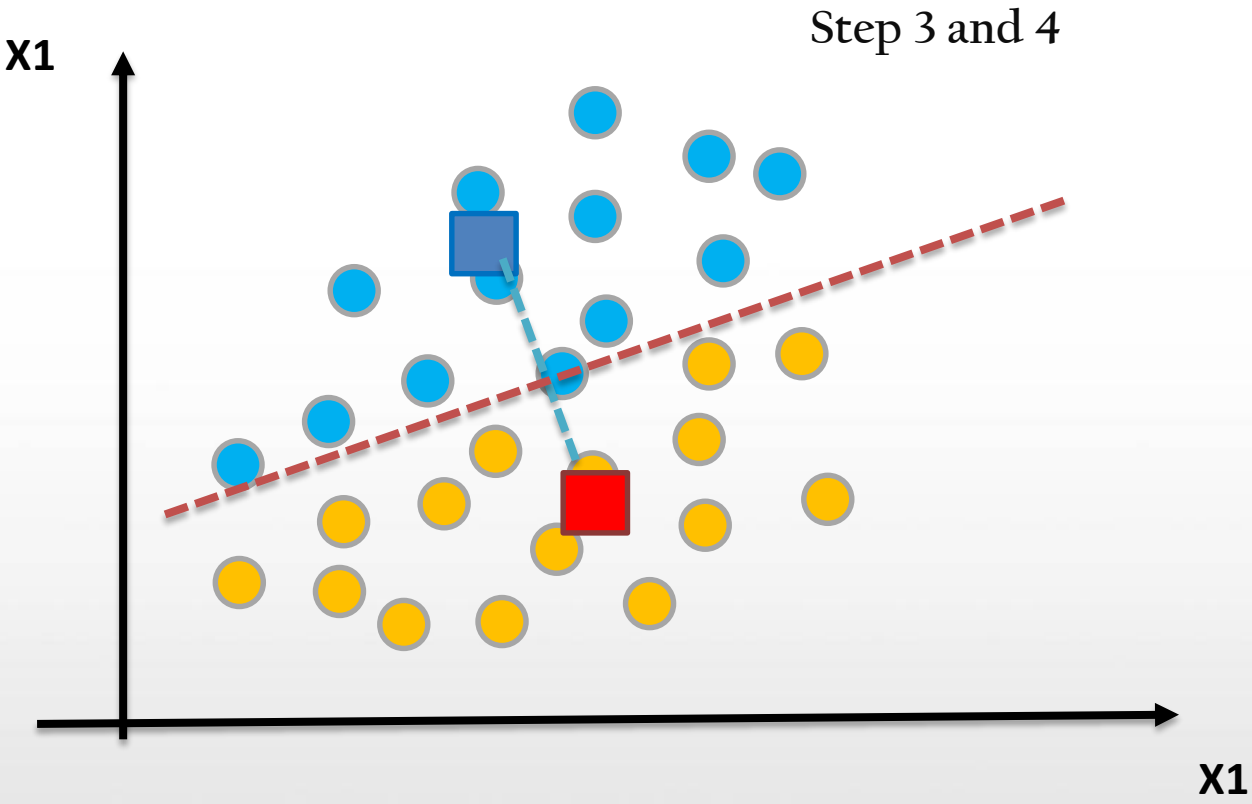
Step 3 and 4



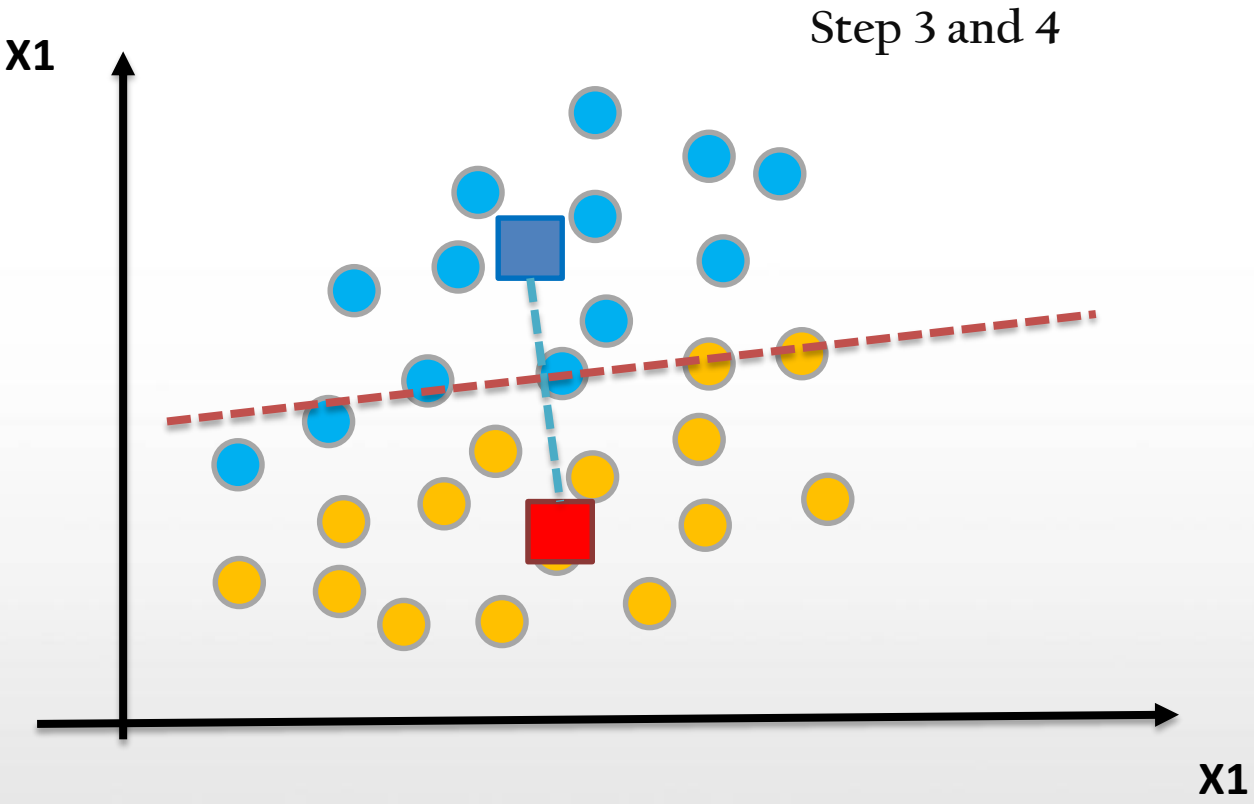
K Means Clustering



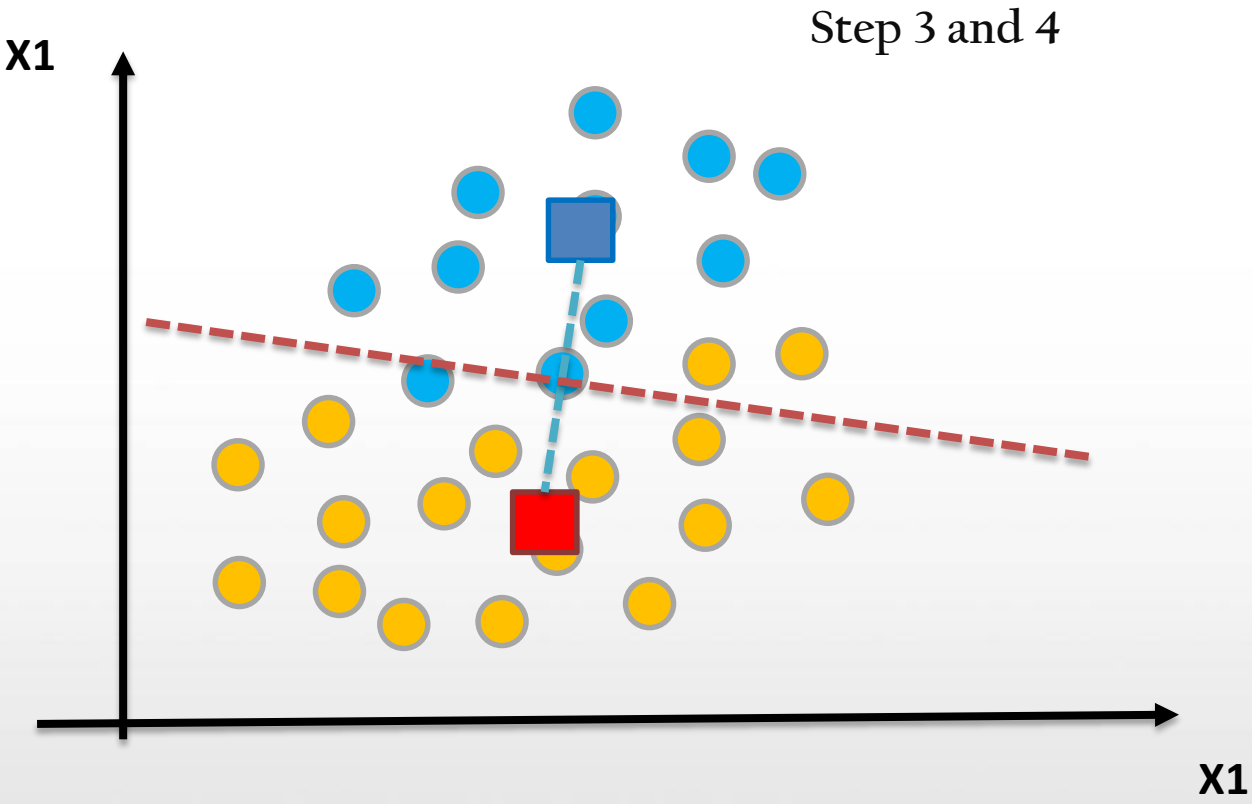
K Means Clustering



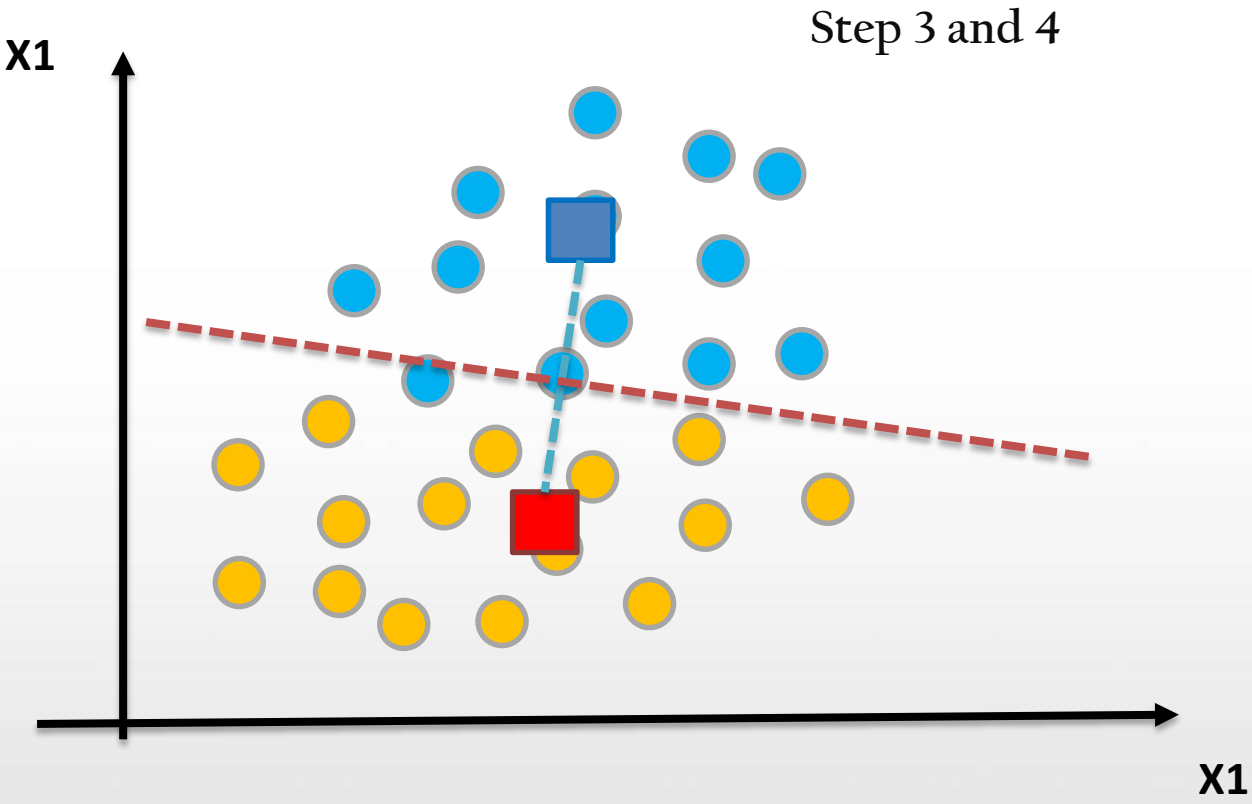
K Means Clustering



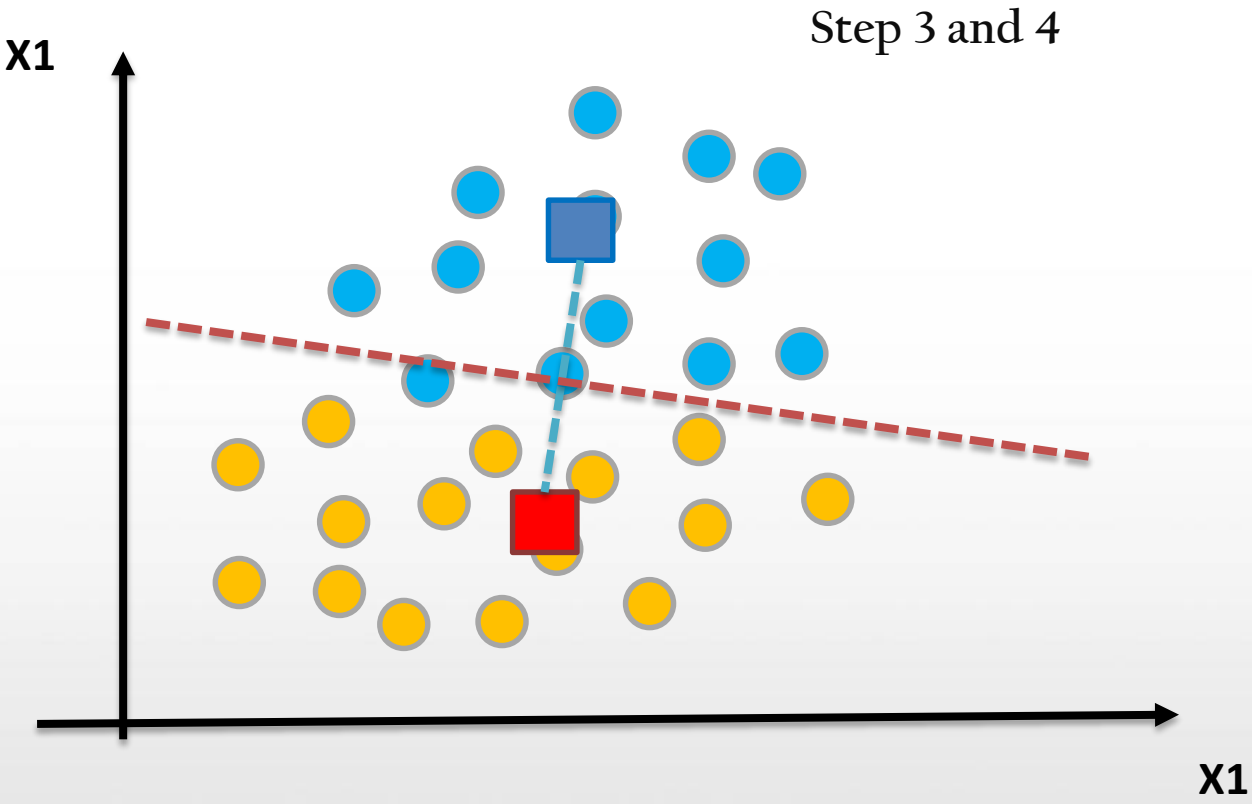
K Means Clustering



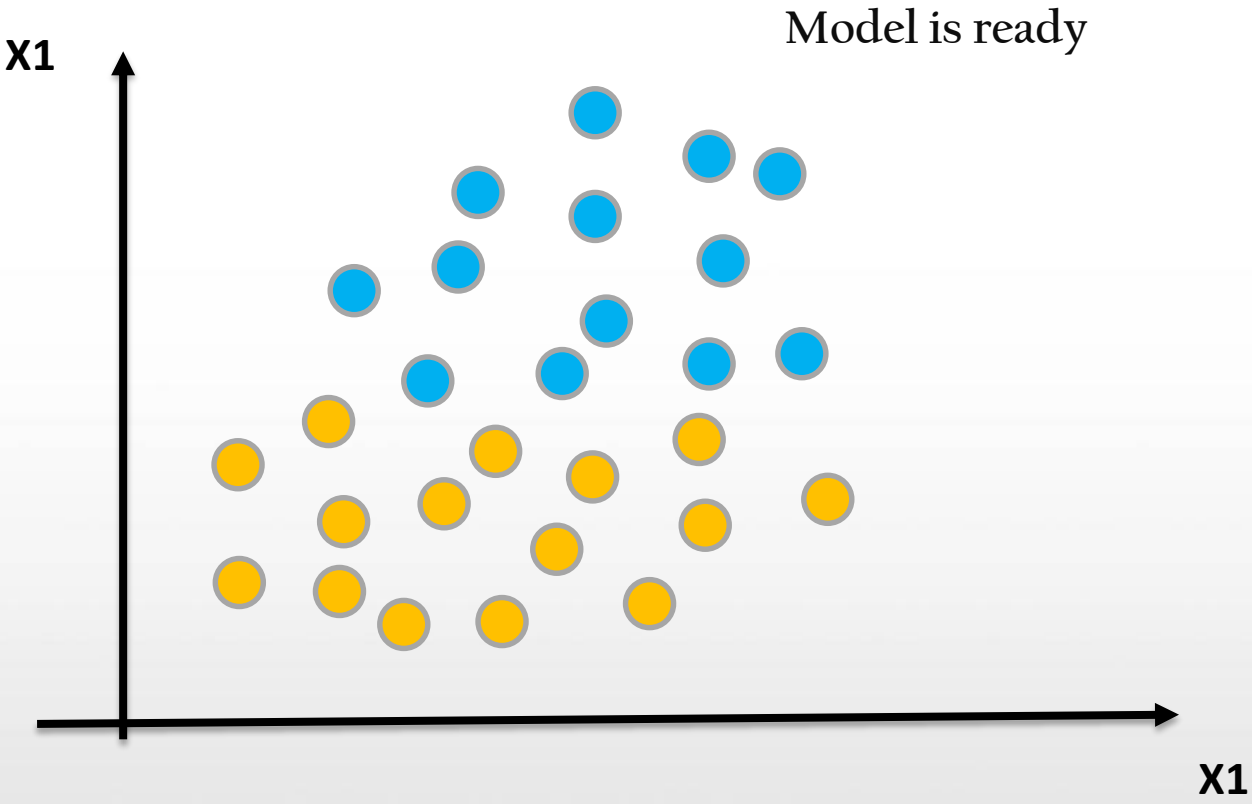
K Means Clustering



K Means Clustering

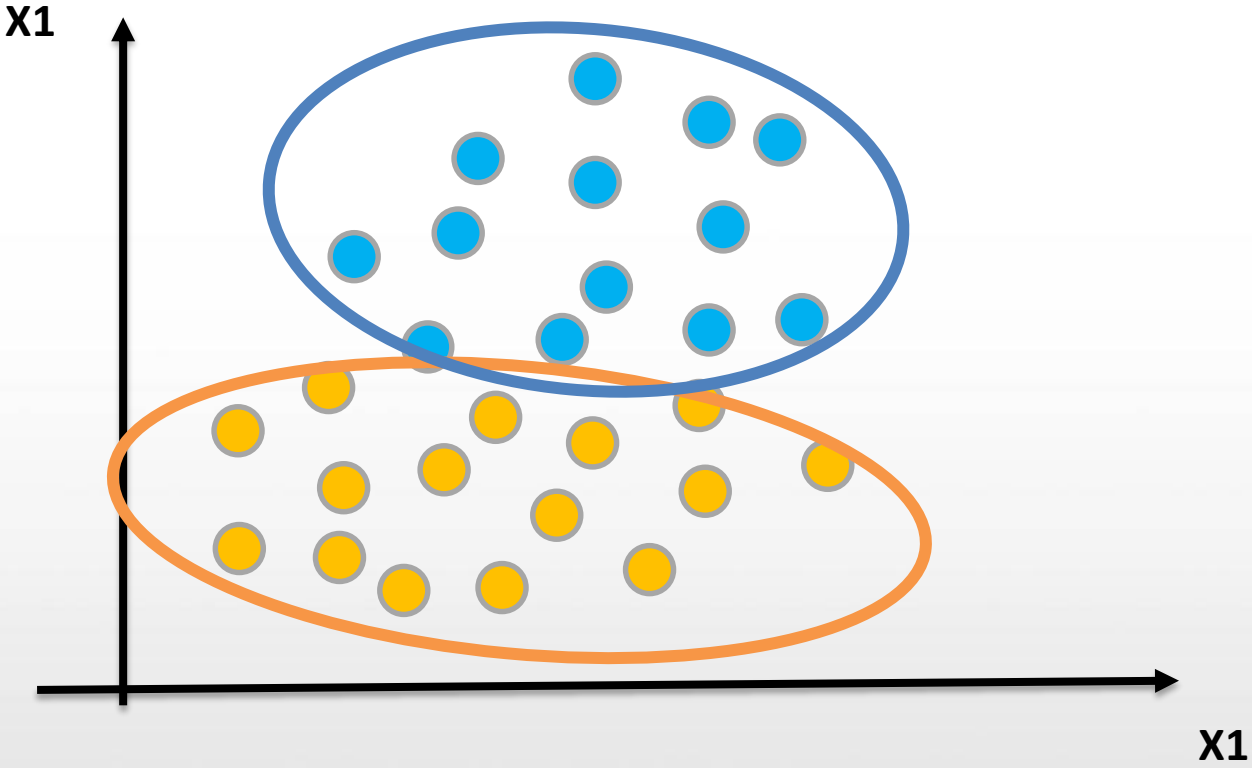


K Means Clustering

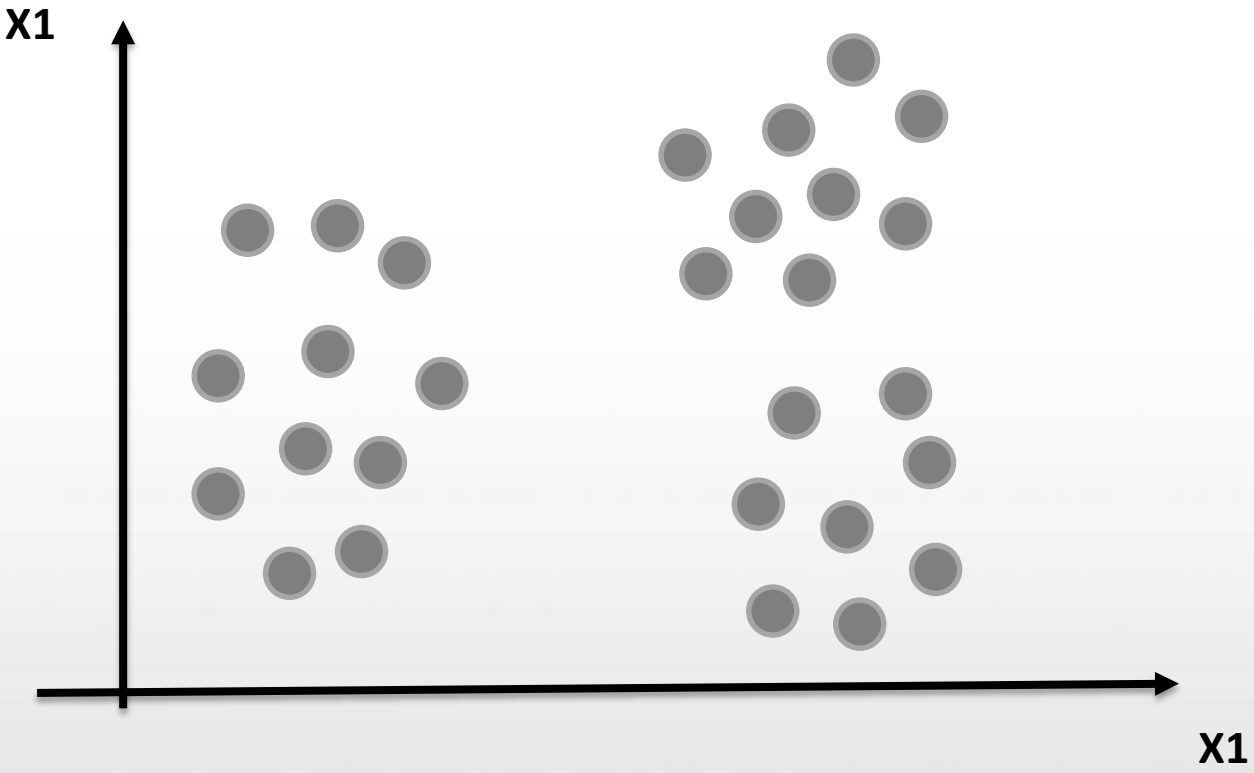


K Means Clustering

Model is ready



How many cluster should I choose

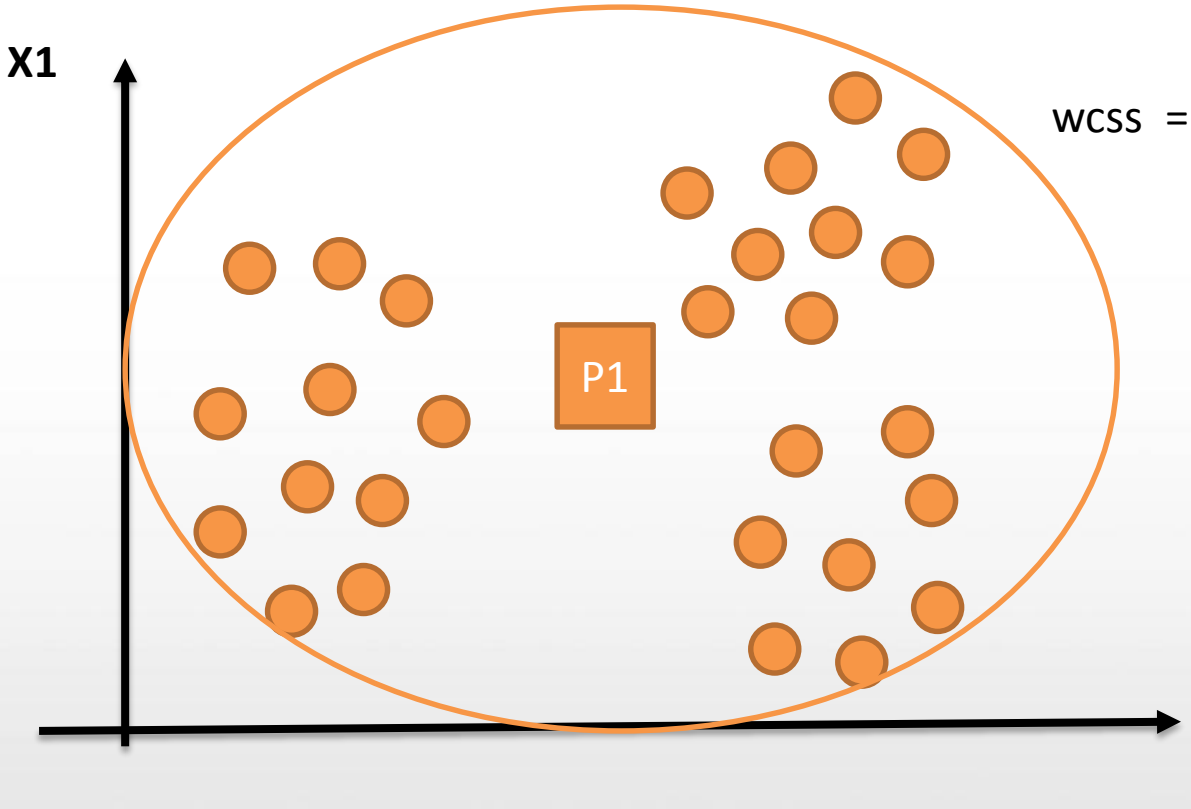


Choosing right number of clusters using **within-cluster sums of squares (WCSS)**

$$\text{WCSS} = \sum distance\,(P_i, C_1)^2 + \sum distance\,(P_i, C_2)^2 + \ldots.$$

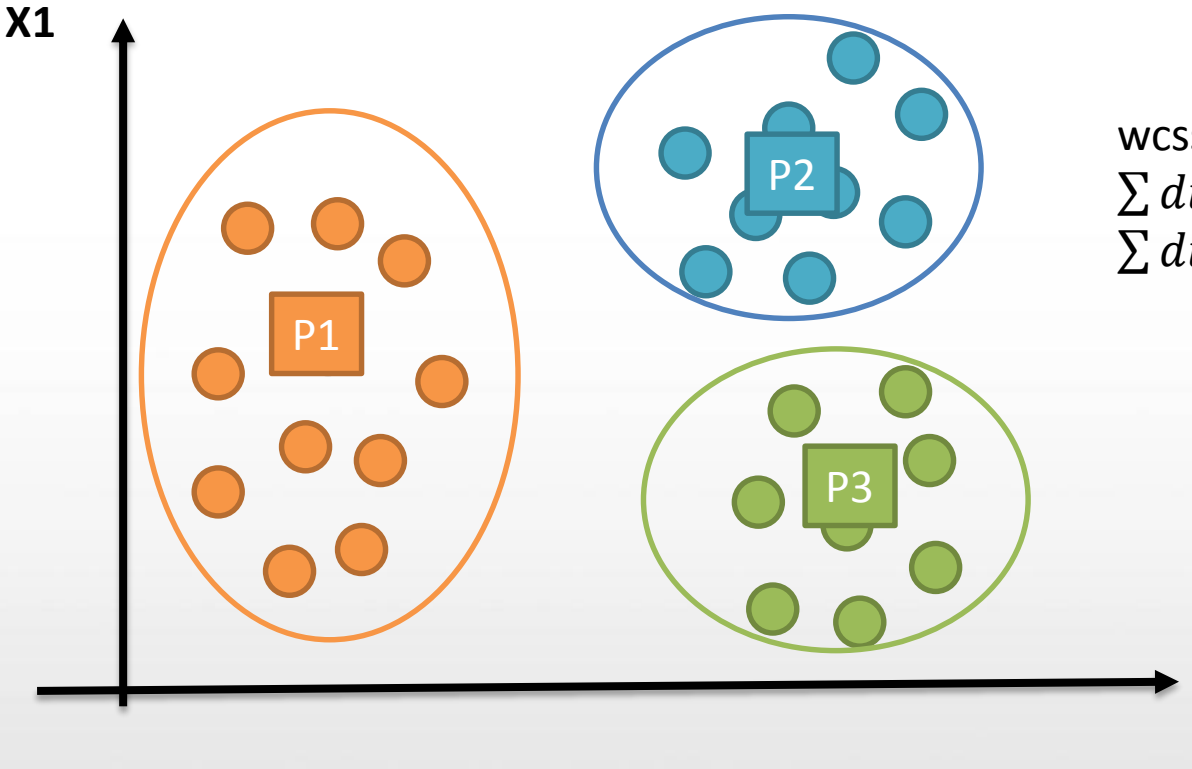
Choose wcss should be minimum

How many cluster should I choose



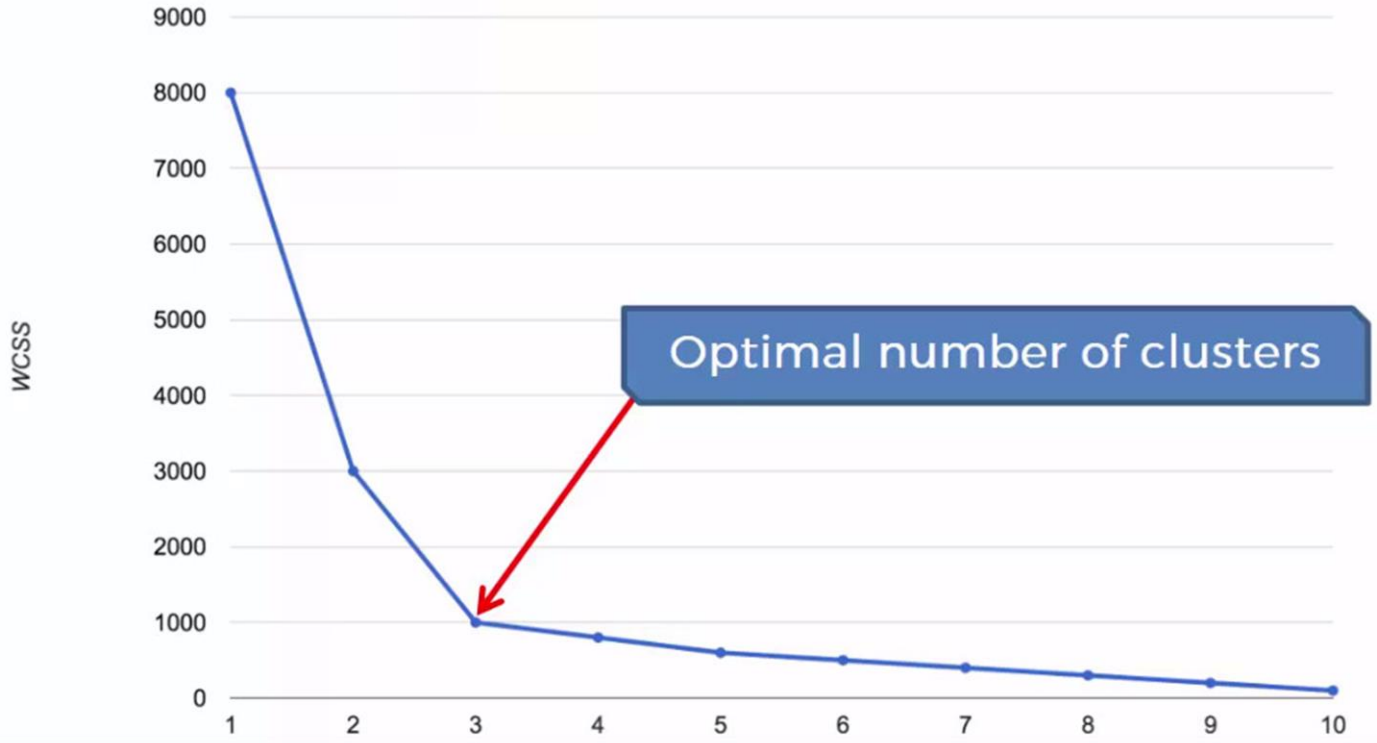
$$\text{wcss} = \sum distance\ (P_i, C_1)^2$$

How many cluster should I choose



wcss = $\sum distance\ (P_i, C_1)^2 +$
$\sum distance\ (P_i, C_2)^2 +$
$\sum distance\ (P_i, 3)^2$

The Elbow Method



Hands on Clustering

IOTGYAN