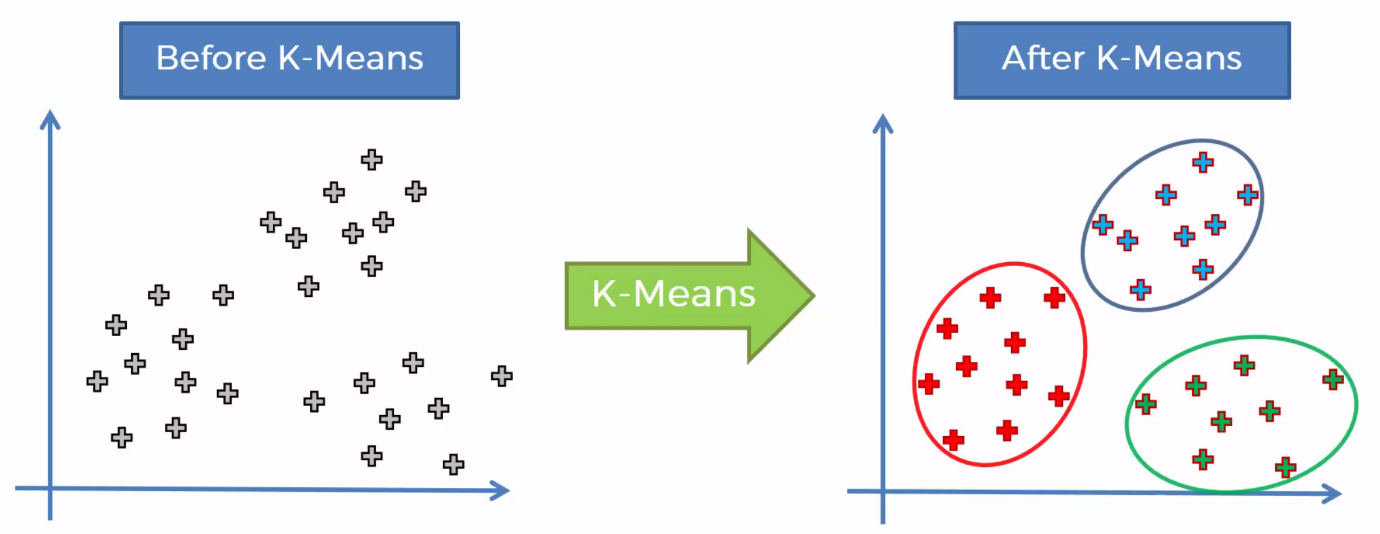
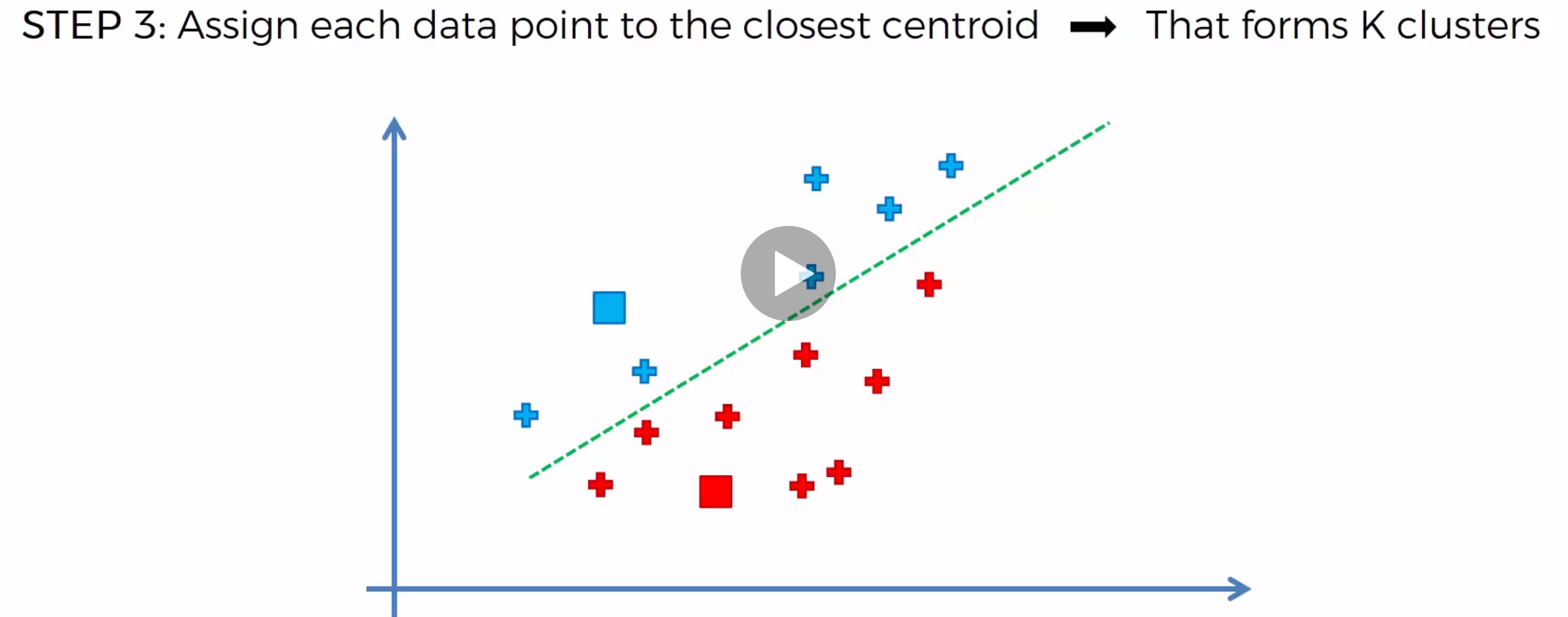
K-Means Intuition: Understanding K-Means

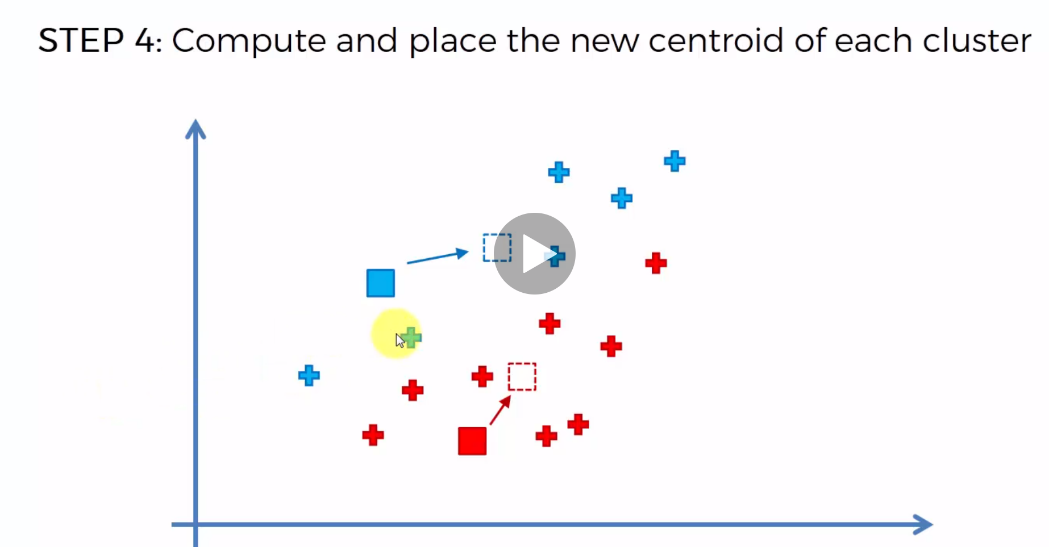
Work with multi-dimensional

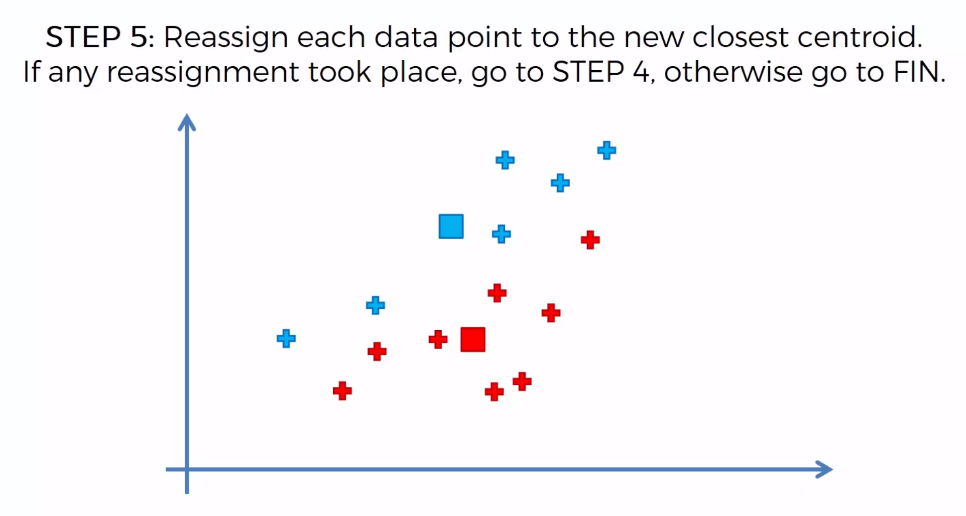
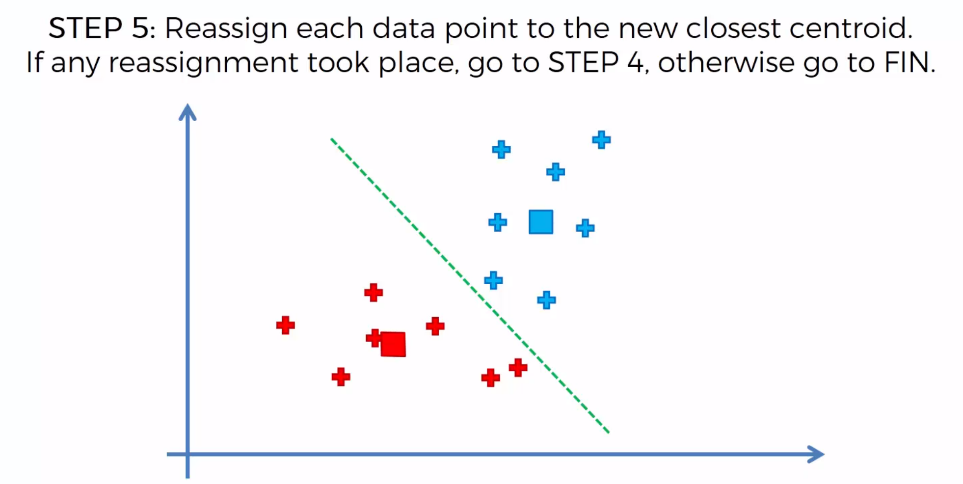
Step 1: Choose the number K of clusters

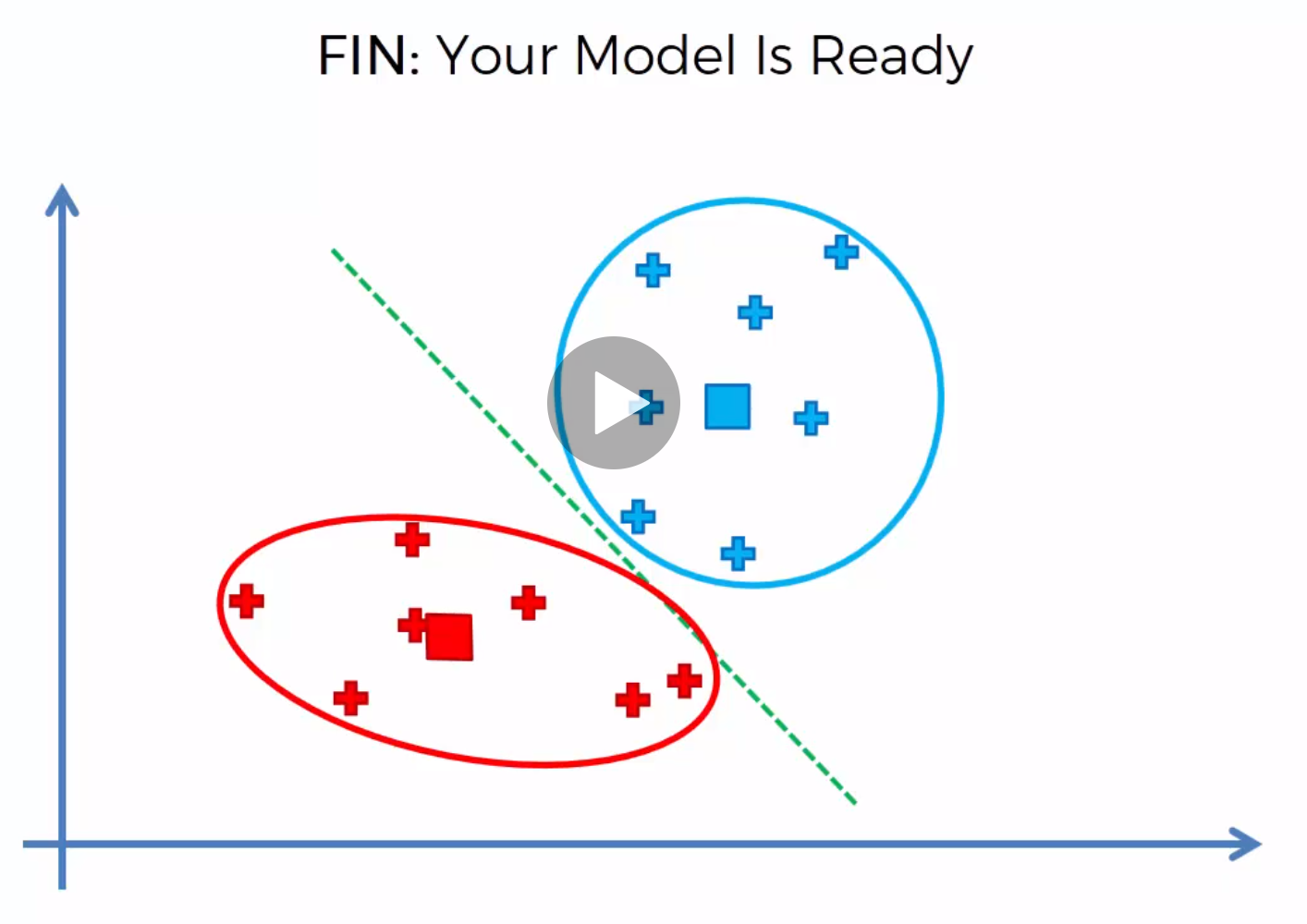
Step 2: Select at random K points, the centroids (not necessarily from your dataset)

Step 3: Assign each datapoint to the closest centroid -> Forms K clusters



Step 4: Compute and place the new centroid of each cluster

Step 5: Reassign each data point to the new closest Centroid. If any assignment took place, go to Step 4, otherwise go to Finish



K-Means Intuition: Random Initialization Trap.

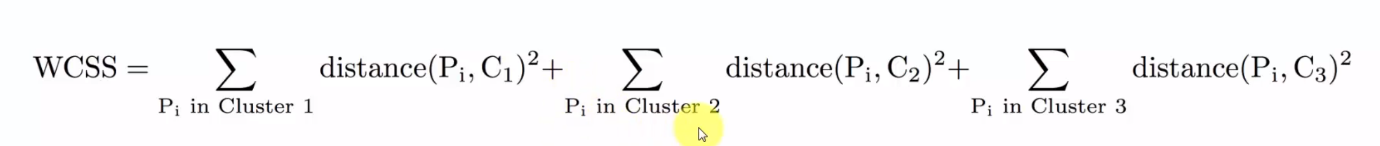
Easily spotted Clusters can be a trap.

K-Means++ algorithm. All happens in the background. Tools that you use can be implementing will be selecting should be specific.

K-Mean Intuition. Selecting the correct number of clusters. There is an algorithm for it.

Metric to be imposed – WCSS Within Cluster sum of squares.

The less the WCSS the better and goodness of fit the model. However,



WCSS vs Number of Clusters (Look for the elbow)

