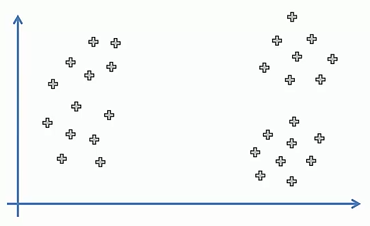
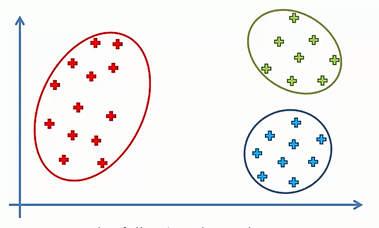
**RANDOM INITIALIZATION TRAP**

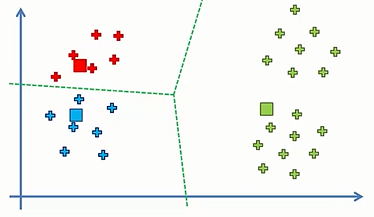
Here we have got a scattered plot.



If we choose k=3 clusters then this correct random initialization would lead us to the following three clusters.

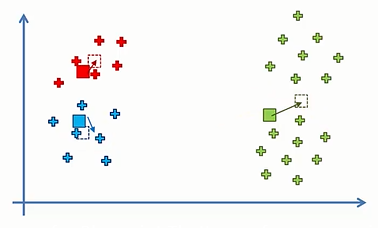


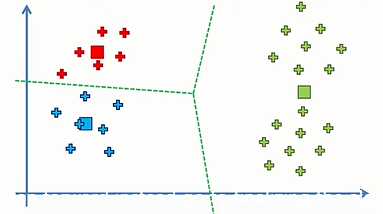
Now, if we select clusters in different locations we will get graph like this.



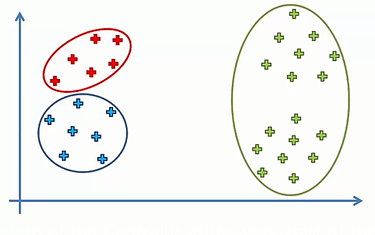
We are going to remove these clusters.

So, we are going to recalculate, compute and place the new centroid for each cluster





this time nothing changes and we get different clusters from those that we got the previous time.



=> Selection of the centroid at the beginning can potentially dictate the outcome of the algorithm.

But that’s not a good thing because centroid are selected at random.

Therefore modification in K – means algorithm is needed to overcome this problem.

Thus **K – Means ++ Algorithm** was introduced.