

Clustering and K-means Clustering

What is clustering?

Clustering is grouping data points into groups where data points in one group are similar to each other.

What is clustering?

Machine Learning: Clustering



By color



By shape



By size



etc...

Methods of Clustering

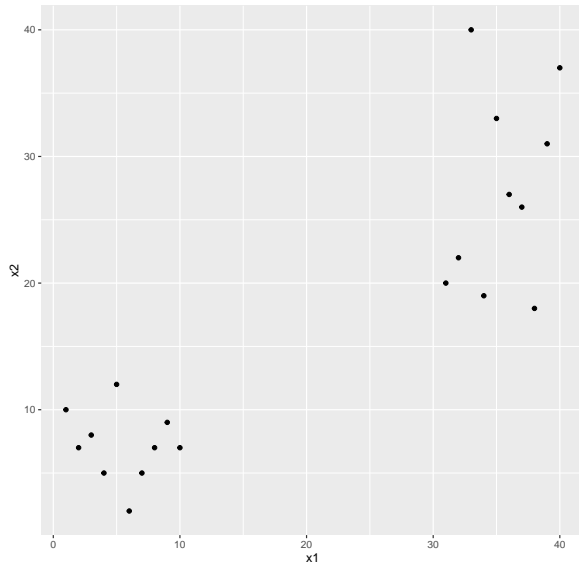
We will cover two clustering methods:

- ▶ K-means clustering and
- ▶ Hierarchical clustering

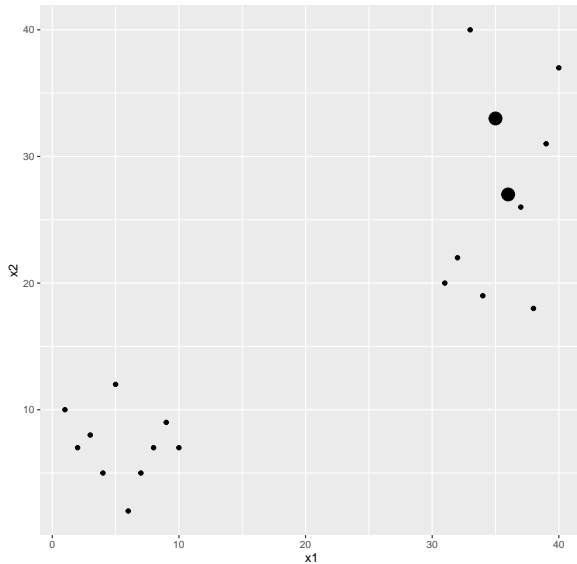
K-means Clustering

- ▶ Data
- ▶ Visualize Data
- ▶ Result of K-means clustering

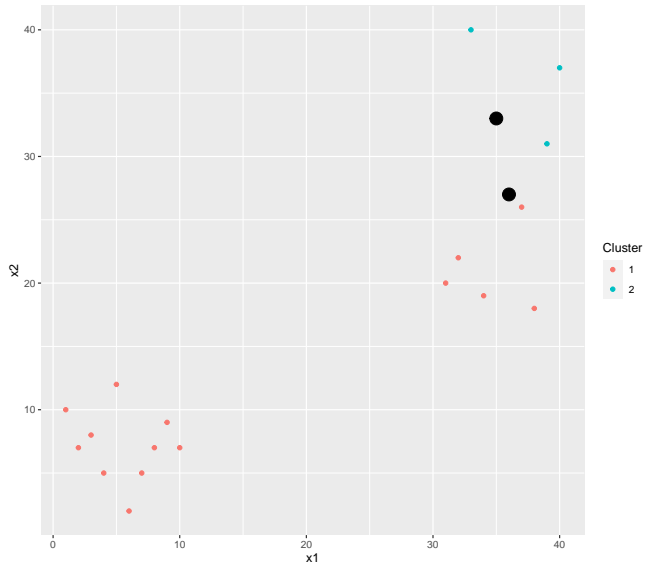
Step 1



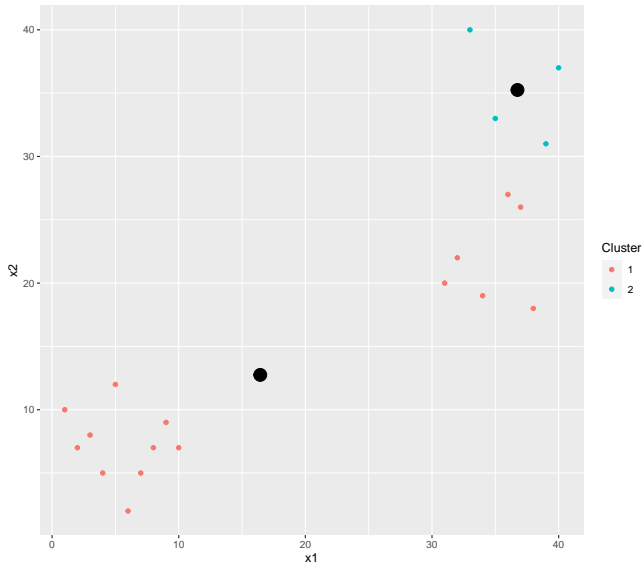
Step 1: Randomly select centroids



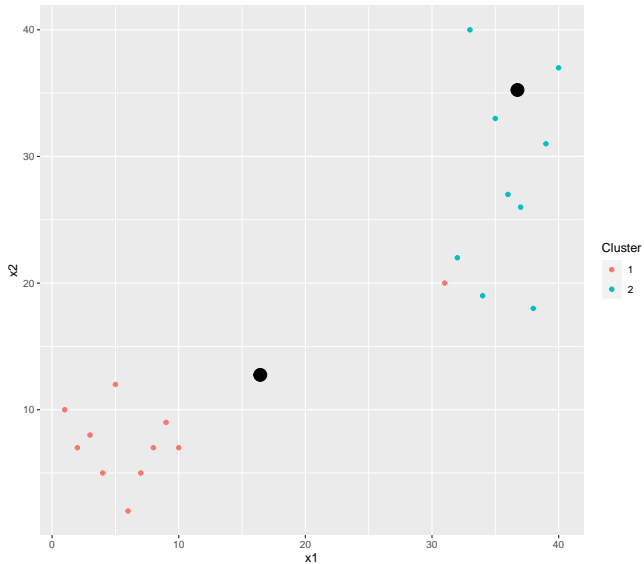
Step 1: Collect points for each clusters



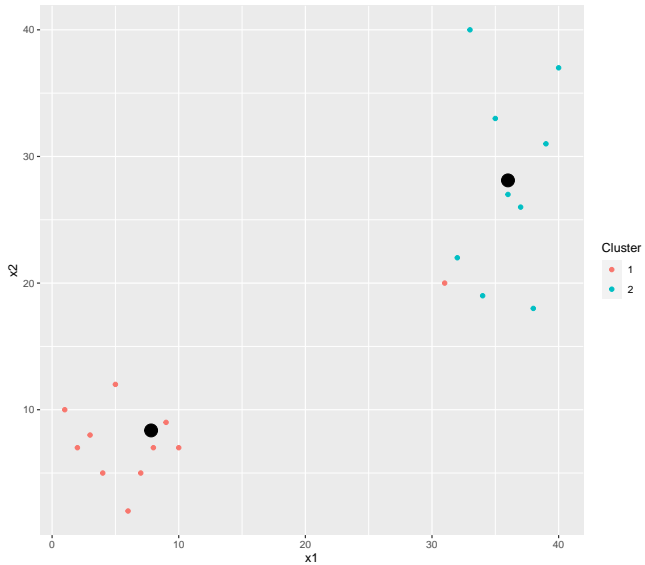
Step 2: Relocate centroids



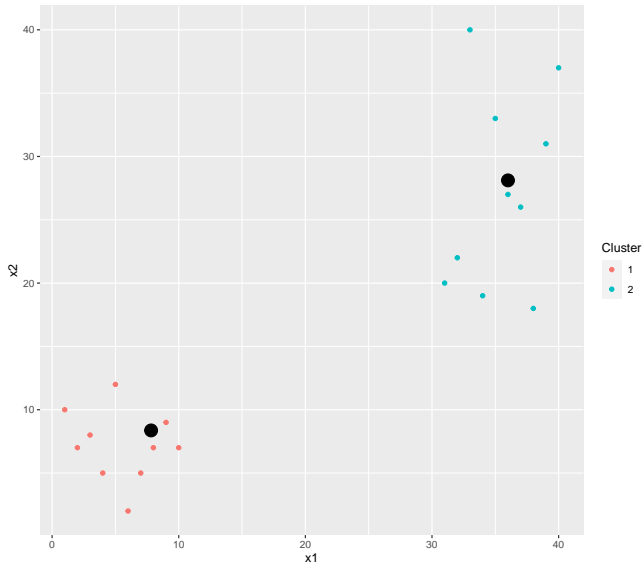
Step 2: Collect points for each clusters



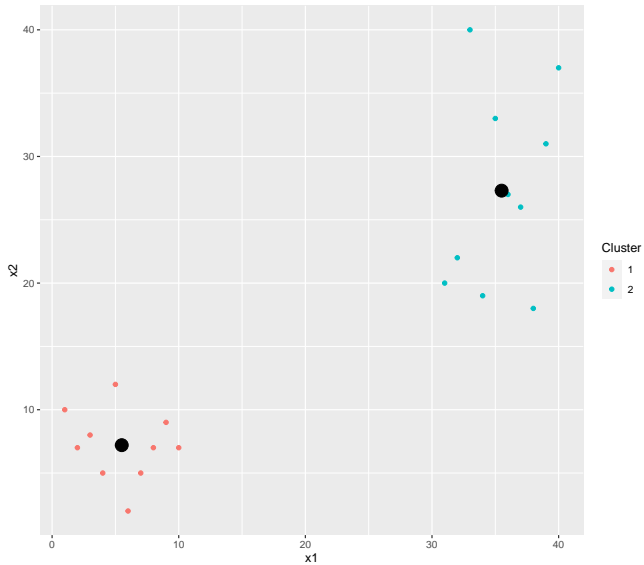
Step 2: Relocate centroids



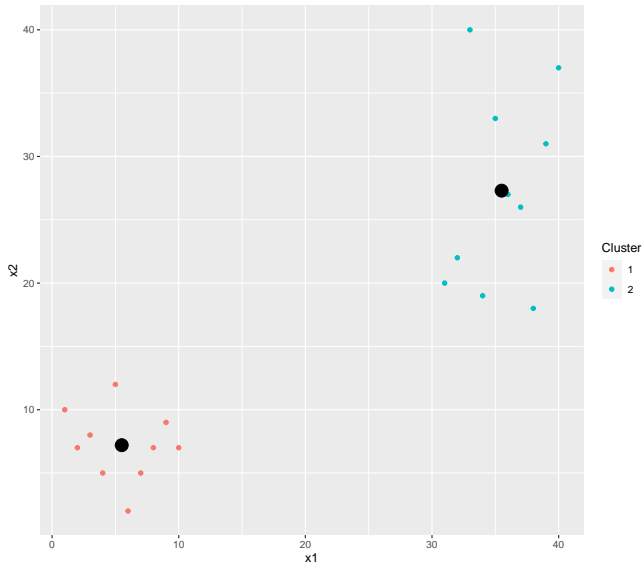
Step 2: Collect points for each clusters



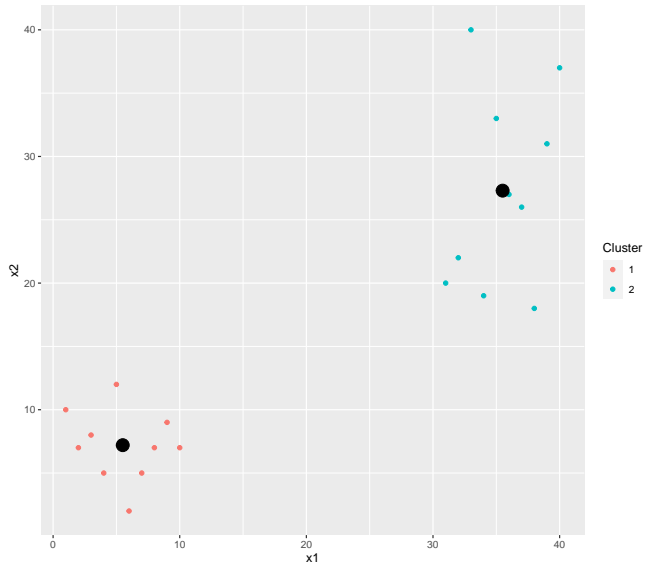
Step 2: Relocate centroids



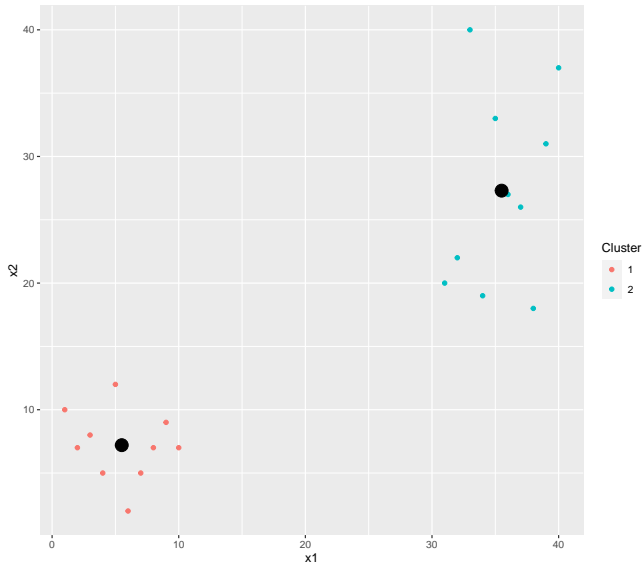
Step 2: Collect points for each clusters



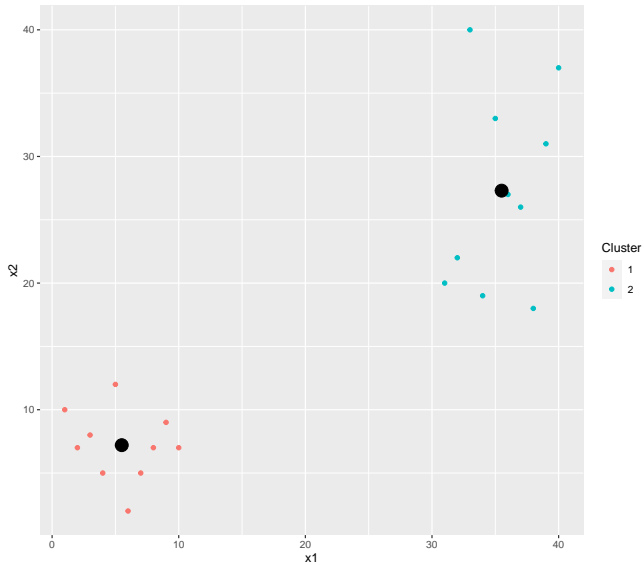
Step 2: Relocate centroids



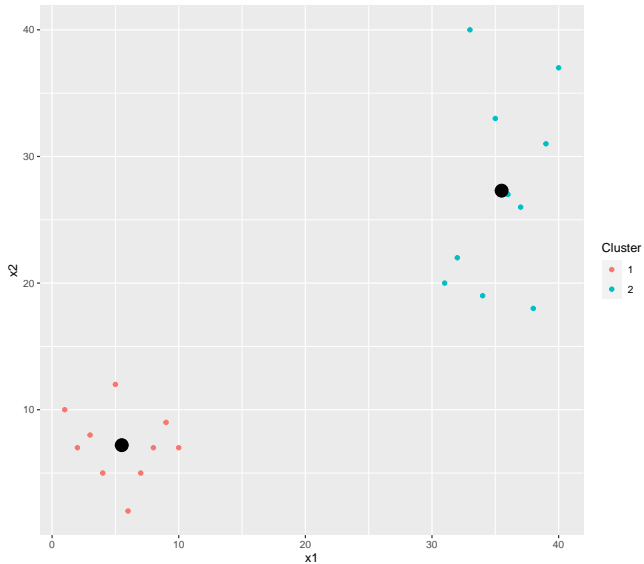
Step 2: Collect points for each clusters



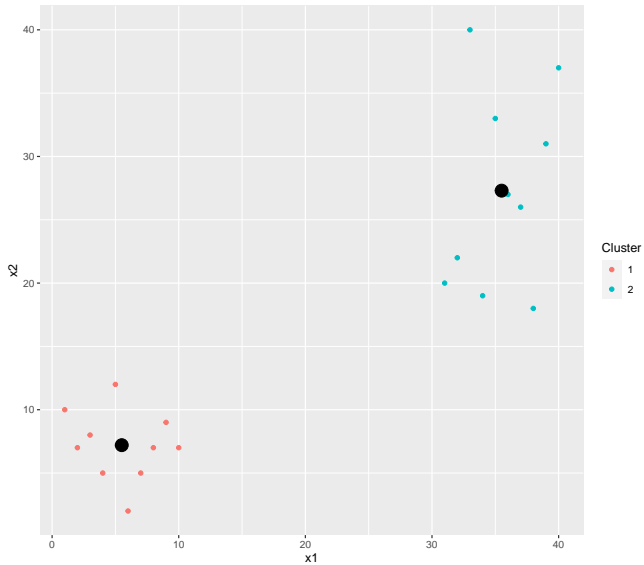
Step 2: Relocate centroids



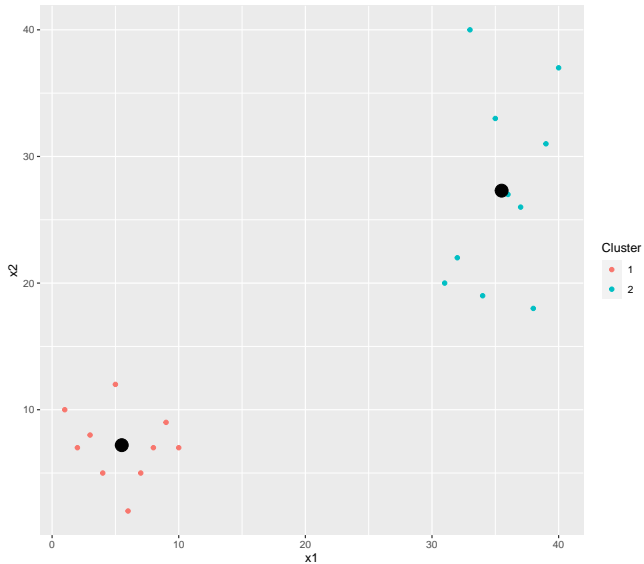
Step 2: Collect points for each clusters



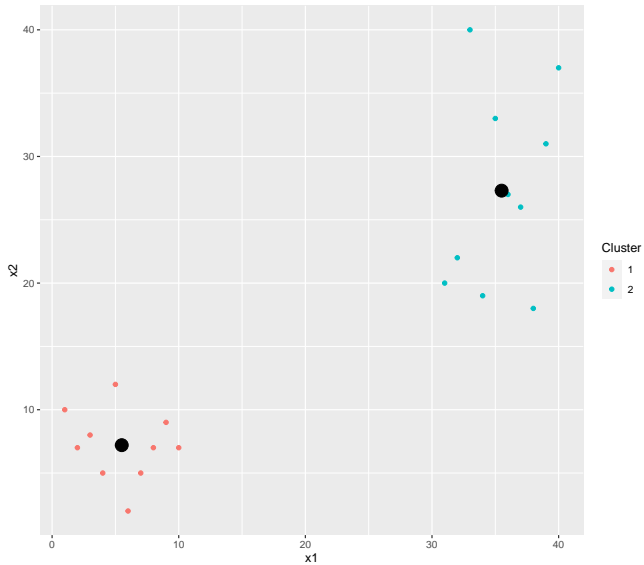
Step 2: Relocate centroids



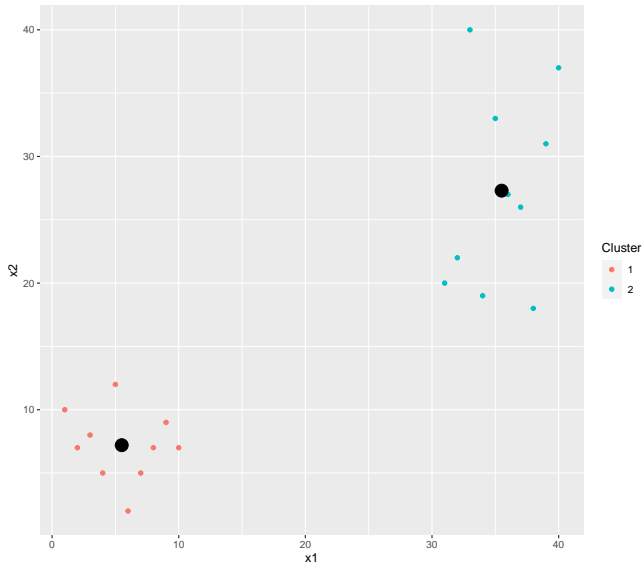
Step 2: Collect points for each clusters



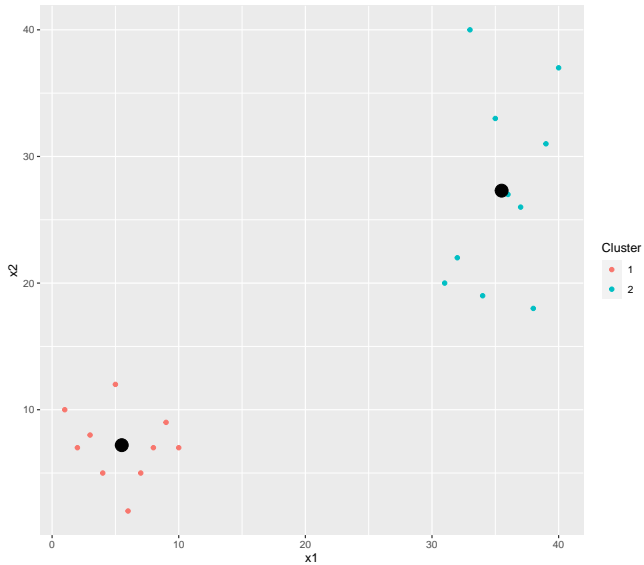
Step 2: Relocate centroids



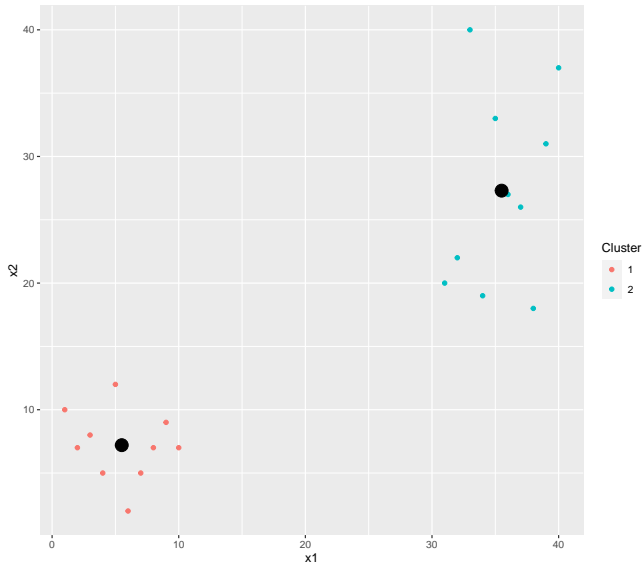
Step 2: Collect points for each clusters



Step 2: Relocate centroids



Step 2: Collect points for each clusters



Centroids

Cluster	x1	x2
1	5.5	7.2
2	35.5	27.3