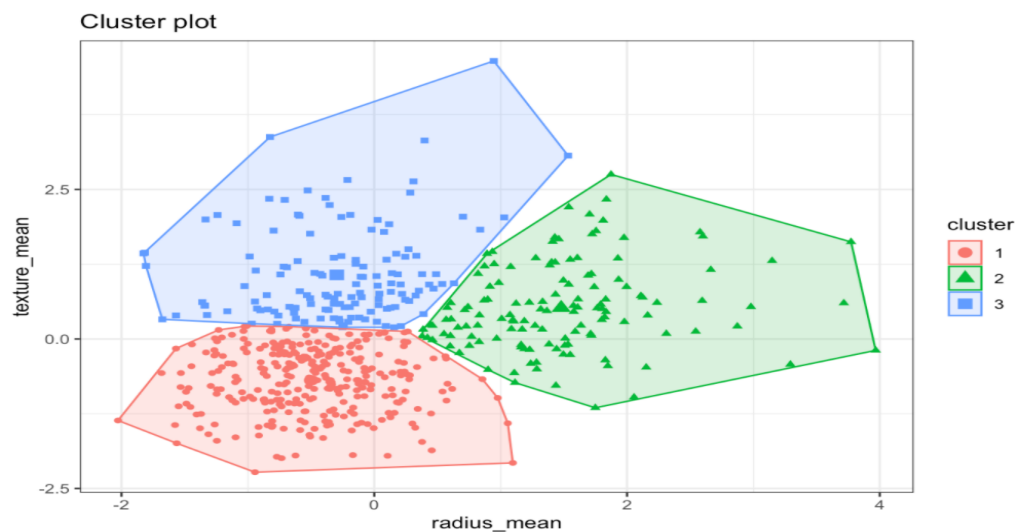# K-means Clustering

Clustering is the process of dividing the entire data into groups (also known as clusters) based on the patterns in the data. In clustering, we do not have a target to predict.

K-means clustering is an unsupervised machine learning algorithm. It makes inferences from datasets using only input vectors without referring to known, or labelled, outcomes.

For this, We need to define a target number k, which refers to the number of centroids we need in the dataset. The K-means algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible.

Useful Link -
- https://stanford.edu/~cpiech/cs221/handouts/kmeans.html
- https://bookdown.org/tpinto_home/Unsupervised-learning/k-means-clustering.html