# K-means clustering Algorithm

## Test Datasets

scikit-learn is a Python library for machine learning that provides functions for generating a suite of test problems.

They are also useful for better understanding the behavior of algorithms in response to changes in hyperparameters.

## Classification test problems:

we will look at three classification problems: blobs, moons and circles.

### Blobs Classification Problem

The [make\_blobs()](http://scikit-learn.org/stable/modules/generated/sklearn.datasets.make_blobs.html) function can be used to generate blobs of points with a Gaussian distribution.

https://machinelearningmastery.com/generate-test-datasets-python-scikit-learn/

K- means clustering using IRIS Dataset:

Dataset is having 150 records consisting of Setosa, versicolor and virginica

Based on sepal length sepal width petal length petal width we need to classify the flower as setosa or versicolor or virginica.

K-mean clustering: