Meanshift is a non parametric, feature space analysis technique for locating the maxima of a density function. It builds on the concept of Kernel density estimation.
It is being used extinsively in the filed of image processing and computer vision.

Given a set of data points, the algorithm iteratively assigns each data point towards
the closest cluster centroid and direction to the closest cluster centroid is determined
by where most of the points nearby are at. So each iteration each data point will move
closer to where the most points are at, which is or will lead to the cluster center.
When the algorithm stops, each point is assigned to a cluster.

The way it works is that it selects a data point and then creates a sort of area around it (circular area) whose radius/bandwidth is decided by the function bandwidth_estimator. Then all other data points falling under that area are considered and the center of that area(just like center of gravity) is calculated using their weigths.This shifts the kernel to a higher density region. Then another area is created with the same bandwidth and the process is repeated till the kernel/center doesnt shift.
This was explained for just one data point. The algorithm does it for all data points till convergence is achieved and the final centers that are calculated are the ones with the highest density around them.
In a way this algorithm is calculating the points where the density/mode of data points is highest. For this reason it is also called Mode seeking algorithm.

The basic idea of this code is to show the working and result of MeanShift algorithm. We have not taken dataset from any online source. Instead we will be using the make_blobs function to generate data for the algorithm.

Firstly we define our own centers. This helps in verifying the efficiency of our algorithm later on. When we generate the data points next, we use these predefined centers as centers for generated data. We also have to select a bandwidth, which can be a tricky task but thanks to estimate_bandwidth function, this can be taken care of.
Next we implement the meanshift and take as output, number of clusters and the predicted centers. We then compare manually the accuracy of each. In our case we can see that the predicted centers and number of clusters are highly accurate.

Unlike the popular K-Means cluster algorithm, mean-shift does not require specifying the number of clusters in advance. The number of clusters is determined
by the algorithm with respect to the data.

PROS --
Application independent
The shape of cluster doesnt matter
Predefined number of clusters not required
Robust to outliers

CONS --
Computationally expensive (O(n^2))
Output depends on window size

This was a very basic implementation of Meanshift.