

Lesson 19 (K-Means Clustering) The Jupyter notebook `K-Means.ipynb` contains code for generating random clusters of points and code that clusters these points using the K-Means clustering algorithm.

- (a) Experiment with generating random clusters. The matrix `STD` contains standard deviations of each attribute (e.g. `x` and `y` coordinates) of each cluster. Setting `rot = True` randomly rotates each cluster. Describe how to change:

increase cluster, smaller SSE

- the number of clusters `num_points length`
- the size of a cluster.
- the density of a cluster.
- the shape of a cluster.

- (b) Use the K-Means algorithm to cluster randomly generated points.

- Typically, the number of clusters is unknown in advance. Describe what happens when too few clusters are used in the K-Means algorithm. What happens when too many clusters are used?
 - How accurate is K-Means clustering when clusters have different sizes?
 - How accurate is K-Means clustering when clusters have different densities?
 - How accurate is K-Means clustering when clusters have non-globular shapes?
-