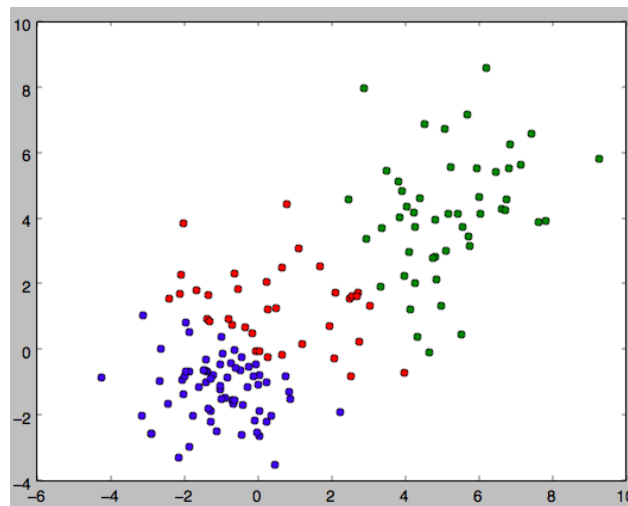


# K-means

The python implementation of K-means can be described as follows. Use python function `random.sample` to randomly choose several different points as initial cluster centroids. In the `kmeans` function, we firstly use function `cluster`, which counts the distances between different data points and three centroids. According to the minimal distance between each point and centroid, we divide several points into three initial clusters. Then in the function `calmeans`, we calculate new centroids based on new divided data points. These two functions are continuously executed in function `kmeans` until the positions of centroids are unchanged.

A scatter plot of K-means on givens dataset goes like:



```
centroids = (0.4971103635555555, 1.2669637546111112)
            (-1.0393701035692309, -1.2380392655538464)
            (5.172903915591836, 4.135913677061226)
```