Assignment 3 (Hands on Experience with Clustering)

1. Generate 4 clusters using the following

```
import numpy as np

N = 80

mean1 = [5, 15]
mean2 = [5, 11]

mean3 = [15,15]
mean4 = [15,11]

cov = [[3, 0], [0, 3]]

np.random.seed(50)
X = np.random.multivariate_normal( mean1, cov, int(N/4))
X = np.concatenate ((X, np.random.multivariate_normal ( mean2, cov, int(N/4))))
X = np.concatenate ((X, np.random.multivariate_normal ( mean3, cov, int(N/4))))
X = np.concatenate ((X, np.random.multivariate_normal ( mean4, cov, int(N/4))))
```

2. Implement using Python K-means clustering with $K = 4$. Do not use existing functions. Set maximum number of iterations to 50. Comment your program. Run your program more than one time. Also evaluate the clustering result using SSE.

   Comment on the performance. Identify situations when the method does well and when it does not. (40%)

3. Repeat step 2 but this time implement K-means++. (40%)

4. Comment on the relative performance of K-means and K-means++ qualitatively and quantitatively. (20%)

What you should hand in

One Jupyter notebook containing your code, output, and answers.