Assignment 4

1. Manually perform K Means clustering on Manual_Data.xlsx. There are 10 data points given and you have to separate them into 2 clusters.

2. Carbon and nitrogen emission tests of 2 different types of vehicles were done. Test results are provided in Excel sheets. Your task is to identify the two groups of vehicles from the data.

a) Plot the data (Data.xlsx) to get an idea of the data distribution. Plot Result 1 on x-axis and Result 2 on y-axis. Report your visual observations.
b) Apply K-Means clustering on the data to find out the 2 clusters. Make appropriate plots.
c) Plot the data (Data_GMM.xlsx) to get an idea of the data distribution. Plot Result 1 on x-axis and Result 2 on y-axis. Report your visual observations.
d) Apply Gaussian Mixture Model on the Data_GMM.xlsx to find out the 2 clusters. Make appropriate plots.
e) Compare the two methods used.

3. Manually(pen-paper) perform the PCA analysis of the data vectors Y1, Y2 presented in thefollowing table. Calculate the eigen vectors corresponding to two principle directions and transform the data into the new coordinate space.

Y1	Y2
2	1
3	4
5	0
7	6
9	2

Submission Details

1. Submit a zip file on moodle named “EntryNumber.zip” with all the code files and a **report with all the graphs and analysis**. Only Matlab & python are allowed.
2. **Use of GMM, PCA and K-means libraries are strictly not allowed.**