Homework Set 11

Problem 1 (k-means)
Download the dataset "11HW1_KmeanData.mat". You will find variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ in the dataset. Use MATLAB function `kmeans` to cluster the 2 variables. Try $k = 4, 5,$ and 6. As part of your results, display all observations along with their labels in a plot for each variable and each k value.

Problem 2 (k-means)
In this problem, you will use k-means to compress images by reducing the number of colors it contains. Find the greyscale image "11HW2_PatchPanelsGrey.jpg" and the true-color image "11HW2_PatchPanels.jpg".

a) Display the images after data compression using k-means clustering for different values of $k = 2, 5, 10, 15,$ and 20.

b) What are the compression ratios for different values of k?

c) You will see that there is a trade-off between degree of compression and image quality. What would be a good value of k for each of the two images?