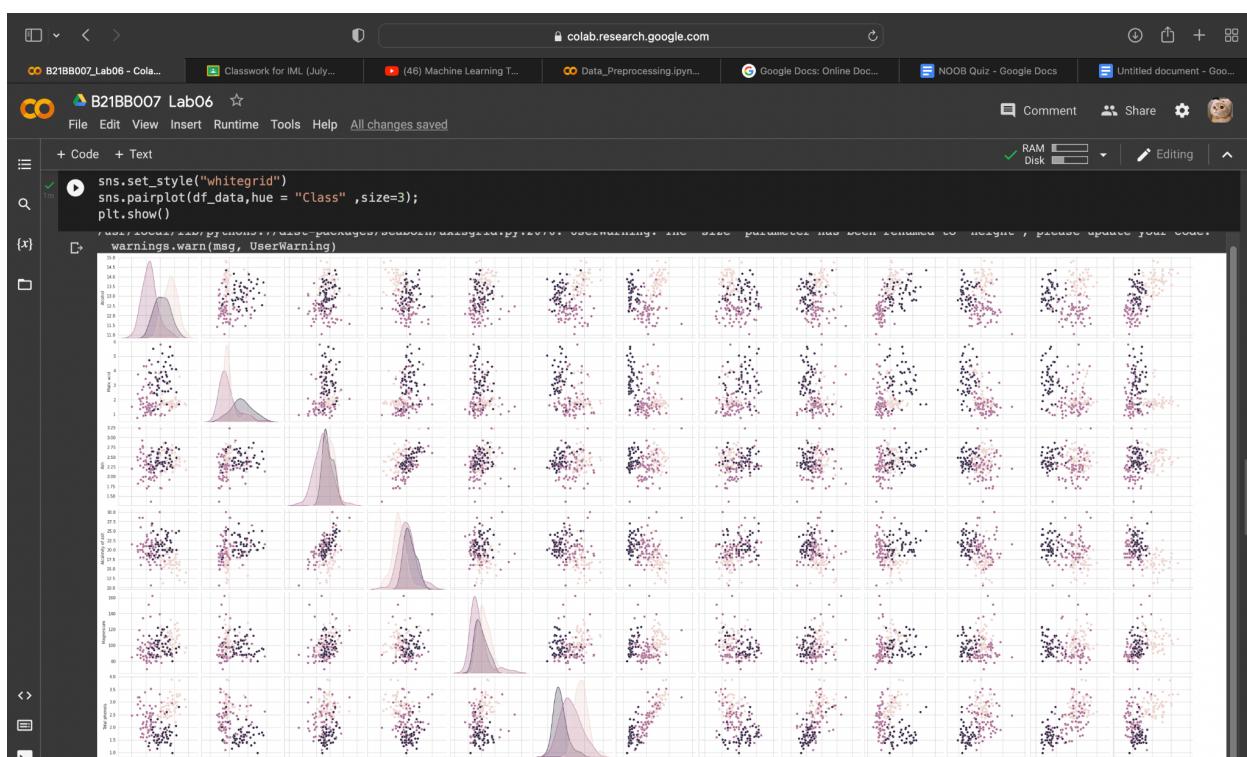
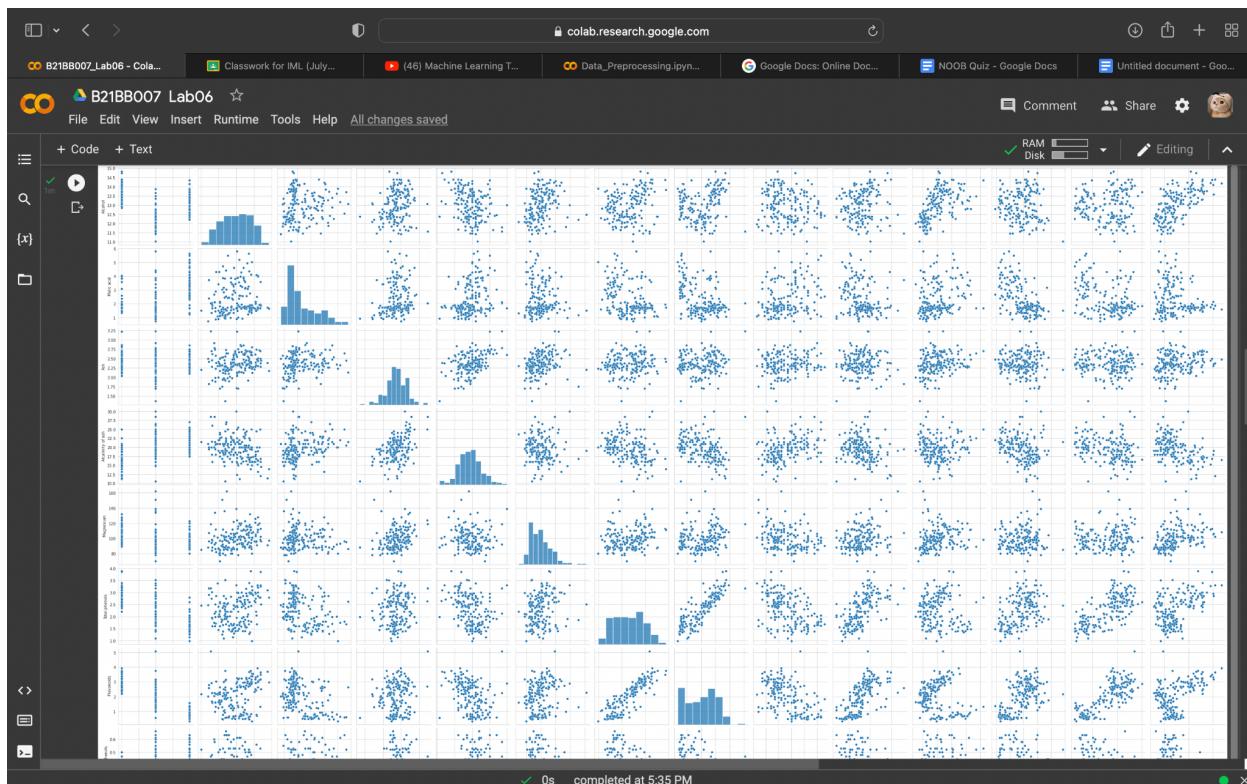


Lab06
Smruti dawale
B21BB007

Que 1)

Here I have plotted the graphs using the seaborn library.
I even did preprocessing , removed the nan values and scaled the data using MinMax Scaler .



Que 3)

The screenshot shows a Google Colab notebook titled "B21BB007_Lab06". The code cell contains the following Python script:

```
[19]     true_labels_list.append(2)

[20] from sklearn.cluster import KMeans
      kmeans = KMeans(n_clusters = 3, init = 'k-means++', max_iter = 600, n_init = 140, random_state = 0)
      y_kmeans = kmeans.fit_predict(x)

      print("True Class Labels :-")
      print(true_labels_list)

      print("Predicted Class labels :-")
      print(kmeans.labels_)
```

The output of the code shows two lists of labels. The first list, "True Class Labels", is a long sequence of ones and twos. The second list, "Predicted Class labels", is a sequence of ones, twos, and threes. Below the code cell, there is a scatter plot with three clusters of points (red, orange, green) and their corresponding centroids (black dots). A legend at the bottom identifies the clusters by color.

As you can see the predicted array using the k means matches almost with the original values , indicating that our model was trained accurately .The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K.I have chosen k = 3.

The plot of clusters with centroids is given below .

