**Anomaly Detection using K-Means Clustering Algorithm**

**Name:**

**Roll No:**

The data file is uploaded in the assignment folder. Data instances that fall outside of defined clusters could potentially be marked as anomalies. Write the program to do the following task. You have got the clusters using Lab 12 program. Fill the spaces below to each task with your code and output (if required).

1. What are the packages imported to answer the other questions?
2. Determine the value of K needs to be chosen for K-Means algorithm (use the Elbow Method)
3. Apply K-means clustering to the data ( As in Lab 12 Program).
4. Display the number of objects in each cluster.
5. Write a function that will return Series of distance between each point of the data and its distance with the closest centroid
6. Using the above defined function get the distance between each point of the data and its nearest centroid.
7. The largest distances are considered as Outlier. To get the number of outliers we will consider a fraction of the distance series. Display the number of outliers (k) using the outliers fraction value 0.1.
8. Set the threshold value to detect a point as outlier or not as the minimum of the largest k distance
9. Add another column in the dataframe to show the particular object is Outlier or not (1 or 0)
10. Plot a scatter plot using two different colours to indicate their class labels (Outlier or not)