### **Name: Mihir Kudale**

### **Roll no: 2012076**

**Subject: MLDL**

### [**Practical 7 Clustering**](https://imcc.sppulms.in/course/view.php?id=113#section-7)

# Import Libraries

import pandas as pd

from sklearn.cluster import KMeans

from sklearn.preprocessing import MinMaxScaler

from matplotlib import pyplot as plt

%matplotlib inline

import warnings

warnings.filterwarnings('ignore')

# Load The Data

# df = pd.read\_csv('Countries\_exercise.csv')

# df.head()

# 

# df.shape

# 

# Visualization of DataSet

# plt.scatter(df.Longitude,df['Latitude'])

# plt.xlabel('Longitude')

# plt.ylabel('Latitude')

# 

# km = KMeans(n\_clusters=3)

# y\_predicted = km.fit\_predict(df[['Longitude','Latitude']])

# y\_predicted

# 

# df['cluster'] = y\_predicted

# df.head()

# 

# km.cluster\_centers\_

# 

# df1 = df[df.cluster==0]

# df2 = df[df.cluster==1]

# df3 = df[df.cluster==2]

# plt.scatter(df1.Longitude,df1['Latitude'],color='orange')

# plt.scatter(df2.Longitude,df2['Latitude'],color='red')

# plt.scatter(df3.Longitude,df3['Latitude'],color='blue')

# plt.scatter(km.cluster\_centers\_[:,0],km.cluster\_centers\_[:,1],color='purple',marker='\*',label='centroid')

# plt.xlabel('Longitude')

# plt.ylabel('Latitude')

# plt.legend()

# 

# Elbow Plot

# k\_rng = range(1,10)

# k\_rng

# 

# sse = []

# k\_rng = range(1,10)

# for k in k\_rng:

# km = KMeans(n\_clusters=k)

# km.fit(df[['Longitude','Latitude']])

# sse.append(km.inertia\_)

# plt.xlabel('K')

# plt.ylabel('Sum of squared error')

# plt.plot(k\_rng,sse)

# 