

```
In [ ]: # Name : Shubham Sapkal
# Roll No. : 2012118
# subject: ML DL
# practical no. : 6

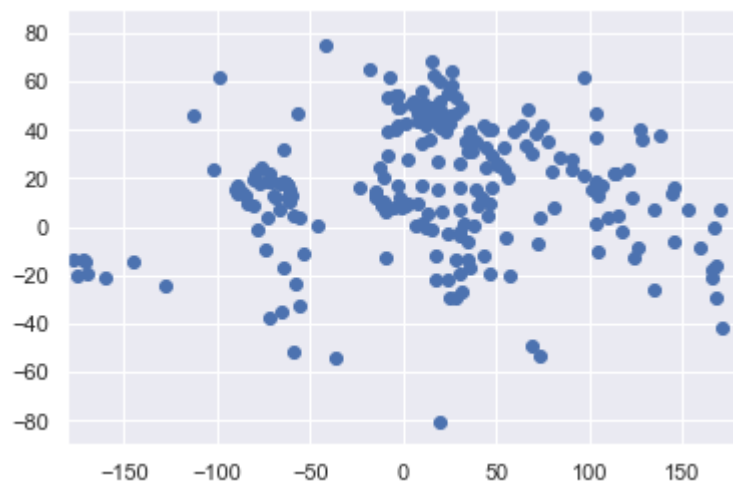
#K means sample code

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
from sklearn.cluster import KMeans

raw_data = pd.read_csv('Countries_exercise.csv')

#Remove the duplicate index column from the dataset.
data = raw_data.copy()

plt.scatter(data['Longitude'], data['Latitude'])
plt.xlim(-180,180)
plt.ylim(-90, 90)
plt.show()
```



```
In [ ]: #Create a copy of that data and remove all parameters apart from Longitude and Latitude.
x = data.iloc[:,1:3]

#Clustering
kmeans = KMeans(3)
kmeans.fit(x)

#Clustering Results
identified_clusters = kmeans.fit_predict(x)
identified_clusters

data_with_clusters = data.copy()
data_with_clusters['Cluster'] = identified_clusters
data_with_clusters

plt.scatter(data['Longitude'], data['Latitude'], c=data_with_clusters['Cluster'], cmap = 'rainbow')
plt.xlim(-180,180)
plt.ylim(-90, 90)
plt.show()
```

