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()

**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()