In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [2]:
```
import openpyxl
#xls=pd.Excelfile('/home/placement/Downloads/country.xlsx',engine='openpyxl')
data=pd.read_excel('/home/placement/Documents/country.xlsx',engine='openpyxl')
```

In [3]:
```
#pip install  openpyxl
```

In [4]:
```
x=data.drop(['country','language'],axis=1)
```

In [5]:
```
x
```

Out[5]:

	lat	long
0	47.68	13.33
1	50.63	4.67
2	51.15	10.43
3	64.00	18.30
4	22.00	78.00
5	7.66	80.63

In [6]:
```
from sklearn.cluster import KMeans
Kmeans=KMeans(n_clusters=2)
ypred=Kmeans.fit_predict(x)
ypred=Kmeans.predict(x)
```

In [7]:
```
1 ypred
```

Out[7]:
```
array([0, 0, 0, 0, 1, 1], dtype=int32)
```

In [8]:
```
1 data['category']=ypred
```

In [9]:
```
1 data
```

Out[9]:

	country	lat	long	language	category
0	austria	47.68	13.33	english	0
1	belgium	50.63	4.67	english	0
2	germany	51.15	10.43	german	0
3	norway	64.00	18.30	english	0
4	india	22.00	78.00	hindi	1
5	srilanka	7.66	80.63	simhala	1

In [10]:
```
1 from sklearn.cluster import KMeans
2 Kmeans=KMeans(n_clusters=3)
3 ypred=Kmeans.fit_predict(x)
4 ypred=Kmeans.predict(x)
```

In [11]:
```
1 ypred
```

Out[11]:
```
array([2, 2, 2, 0, 1, 1], dtype=int32)
```

In [12]:
```
1 data['category']=ypred
```

In [13]:
```
1 data
```

Out[13]:

	country	lat	long	language	category
0	austria	47.68	13.33	english	2
1	belgium	50.63	4.67	english	2
2	germany	51.15	10.43	german	2
3	norway	64.00	18.30	english	0
4	india	22.00	78.00	hindi	1
5	srilanka	7.66	80.63	simhala	1

In [ ]:
```
1
```

In [ ]:
```
1
```

In [ ]:
```
1
```

In [ ]:
```
1
```