In [19]:
```
import pandas as pd
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [20]:
```
import openpyxl
data=pd.read_excel('/home/placement/Documents/country.xlsx',engine='openpyxl')
```

In [21]:
```
data
```

Out[21]:

	country	lat	long	language
0	austria	47.68	13.33	english
1	beljium	50.63	4.67	english
2	germany	51.15	10.43	german
3	india	22.00	78.00	hindi
4	sreelanka	7.66	80.63	simhala
5	southafrica	-20.33	28.00	english

In [22]:
```
x=data.drop(['country','language'],axis=1)
```

In [23]:
```
x
```

Out[23]:

	lat	long
0	47.68	13.33
1	50.63	4.67
2	51.15	10.43
3	22.00	78.00
4	7.66	80.63
5	-20.33	28.00

In [24]:
```
from sklearn.cluster import KMeans
Kmeans=KMeans(n_clusters=2)
ypred=Kmeans.fit_predict(x)
ypred=Kmeans.predict(x)
```

In [25]:
```
ypred
```

Out[25]:
```
array([1, 1, 1, 0, 0, 0], dtype=int32)
```

In [26]:
```
data['category']=ypred
```

In [27]:
```
data
```

Out[27]:

	country	lat	long	language	category
0	austria	47.68	13.33	english	1
1	beljium	50.63	4.67	english	1
2	germany	51.15	10.43	german	1
3	india	22.00	78.00	hindi	0
4	sreelanka	7.66	80.63	simhala	0
5	southafrica	-20.33	28.00	english	0

In [28]:
```
from sklearn.cluster import KMeans
Kmeans=KMeans(n_clusters=3)
ypred=Kmeans.fit_predict(x)
ypred=Kmeans.predict(x)
```

In [29]:
```
ypred
```

Out[29]:
```
array([1, 1, 1, 0, 0, 2], dtype=int32)
```

In [30]:
```
data['category']=ypred
```

In [31]:
```
ypred
```

Out[31]:
```
array([1, 1, 1, 0, 0, 2], dtype=int32)
```

In [32]:
```
data['category']=ypred
```

In [33]:
```
data
```

Out[33]:

	country	lat	long	language	category
0	austria	47.68	13.33	english	1
1	beljium	50.63	4.67	english	1
2	germany	51.15	10.43	german	1
3	india	22.00	78.00	hindi	0
4	sreelanka	7.66	80.63	simhala	0
5	southafrica	-20.33	28.00	english	2

In [ ]:
```

```