In [111]:
```
import pandas as pd
```

In [112]:
```
import matplotlib.pyplot as plt
```

In [113]:
```
import openpyxl
xls=pd.ExcelFile("/home/placement/Desktop/divyasri/Lat Long details.xlsx")
```

In [114]:
```
data=pd.read_excel(xls)
```

In [115]:
```
data.describe()
```

Out[115]:

	Lat	Long
count	7.000000	7.000000
mean	30.955714	33.337143
std	31.821718	32.230456
min	-26.330000	4.670000
25%	14.780000	11.880000
50%	47.680000	18.300000
75%	50.890000	53.000000
max	64.000000	80.630000

In [116]:
```
X=data.drop(['Country','Language'],axis=1)
```

In [117]:
```
X
```

Out[117]:

	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	7.56	80.63
6	-26.33	28.00

In [118]:
```
import warnings
warnings.filterwarnings("ignore")
```

In [127]:
```
from sklearn.cluster import KMeans   #kmeans-clustering
kmeans=KMeans(n_clusters=2)
ypred=kmeans.fit_predict(X)
ypred=kmeans.predict(X)
```

In [128]:
```
ypred
```

Out[128]:
```
array([0, 0, 0, 0, 1, 1, 1], dtype=int32)
```

In [129]:
```
data['category']=ypred
```

In [130]:
```
data
```

Out[130]:

	Country	Lat	Long	Language	category
0	austria	47.68	13.33	english	0
1	belgum	50.63	4.67	english	0
2	germany	51.15	10.43	german	0
3	noway	64.00	18.30	english	0
4	India	22.00	78.00	hindi	1
5	srilanka	7.56	80.63	simhala	1
6	South africa	-26.33	28.00	english	1

In [131]:
```
from sklearn.cluster import KMeans   #kmeans-clustering
kmeans=KMeans(n_clusters=3)
ypred=kmeans.fit_predict(X)
ypred=kmeans.predict(X)
```

In [132]:
```
ypred
```

Out[132]:
```
array([0, 0, 0, 0, 1, 1, 2], dtype=int32)
```

In [133]:
```
data['category']=ypred
```

In [134]:
```
data
```

Out[134]:

	Country	Lat	Long	Language	category
0	austria	47.68	13.33	english	0
1	belgum	50.63	4.67	english	0
2	germany	51.15	10.43	german	0
3	noway	64.00	18.30	english	0
4	India	22.00	78.00	hindi	1
5	srilanka	7.56	80.63	simhala	1
6	South africa	-26.33	28.00	english	2

In [ ]:
```

```