Capestone Project

Introduction/Business Problem:

If a Businessman has a successful restaurant at a place in a city, He would tend to open his next restaurant at place similar to the successful restaurant's place in another city.

Let's assume I have a successful restaurant in Madhapur , Hyderabad. I am willing to open a restaurant in Central Bangalore neighborhood which is like Madhapur.

Aim: The goal is to find which neighborhood of Central Bangalore is like Madhapur , Hyderbad.

Data:

First, I will be needing latitude and longitude of Madhapur. I will use geopy library to get the latitude and longitude. Then I will be querying nearby places data from foursquare API.

Then I will use beautifulsoup library to get neighborhood places of Central Bangalore. For each place I would using geopy library to get latitude and longitude and foursquare API to get nearby places and concate them into single pandas data frame.

Methodology:

1. First I will collect the nearby data of Madhapur using Foursquare API.
2. Then I will get the region in Central Banglore neighbourhood using BeautifulSoup
3. Then I will get nearby data of every region in Central Bangalore and append them to a data frame.
4. Then I will append Madhapur data.
5. Then Preprocess the data to perform K-Means Clustering
6. Perform K-Means Clustering and find which region is more like Madhapur

Results:

After performing the Clustering, I found that Vansanth Nagar in Banglore is more likely to Madhapur Hyderabad.

K_Means Clustering

In [116]:
```
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=4).fit(train_x)
```

In [117]:
```
labels = kmeans.labels_
print(labels)
```

```
[1 2 2 1 1 0 0 1 3 3]
```

In [118]:
```
for x,y in zip(le.classes_,labels):
    print(x," : ",y)
```

```
Cantonment area  :  1
Domlur  :  2
Indiranagar  :  2
Malleswaram  :  1
Sadashivanagar  :  1
Seshadripuram  :  0
Shivajinagar  :  0
Ulsoor  :  1
Vasanth Nagar  :  3
madhapur  :  3
```

Discussion:

Vasanth Nagar should be our first choice. But Incase building a restaurant in Vasanth Nagar is not possible then we need to build at Seshadripuram or Shivajinagar (use k=3 in clustering, you will find these in same cluster as Madhapur).

K_Means Clustering

In [119]:
```
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=2).fit(train_x)
```

In [120]:
```
labels = kmeans.labels_
print(labels)
```

```
[0 1 1 0 0 1 1 0 1 1]
```

In [121]:
```
for x,y in zip(le.classes_,labels):
    print(x," : ",y)
```

```
Cantonment area  :  0
Domlur  :  1
Indiranagar  :  1
Malleswaram  :  0
Sadashivanagar  :  0
Seshadripuram  :  1
Shivajinagar  :  1
Ulsoor  :  0
Vasanth Nagar  :  1
madhapur  :  1
```

Conclusion:

The new restaurant should be built at Vasanth Nagar. If not possible then Seshadripuram or Shivajinagar should be our second choice.