# Predicting the location of new Italian restaurant in Banglore

**Md Zaid Hasib**

**5th February 2020**

## 1. Introduction : Bussiness Problem

India is country of diversity , it has several metro cities where people from different parts of India comes for jobs and better lives. Bangalore is one of them. Considered as IT hub of India, it has population of more than 10 million. Famous for different culture , food , and night life, Bangalore has all different types of places. You can also find different type of restaurants, from authentic south Indian to international cuisine. In this project we will analyse the number of Italian restaurants in the area and share the details with the shareholder, that which area would be best, if restaurant has to open.

## 2. Data Acquisition:

For the project, we are collecting data of Bangalore.

Steps to be followed for data acquisition:

1. With the help of geopy library, I got the longitude and latitude of Bangalore.

From the geopy library e are able to find the latitude and longitude of Banglore

```
[3]  1 print('Latitude = {}, Longitude = {}'.format(location.latitude, location.longitude))
     2 bang_lat,bang_long = location.latitude, location.longitude
```

```
Latitude = 12.9791198, Longitude = 77.5912997
```

2. As Bangalore is large city , so from the centre of Bangalore I am covering an area of 12x12 km. Further I have distributed the area into hexagonal

shape. Using python function convert the lat,lon in X-Y co-ordinates. We got the latitude and longitude of  all the 364 neighbourhood. From reverse geopy , by giving the lat,lon of the different cluster , we will able to find the area/address of that segment. By combining all the data's made one dataframe with address, lat, lon ,x,y and distance from centre as fields.

| | Address | Latitude | Longitude | X | Y | Distance from center |
|---|---|---|---|---|---|---|
| 0 | Fort High School, Krishna Rajendra Road, Maval... | 12.958821 | 77.573128 | 8.885556e+06 | 2.961840e+06 | 5992.495307 |
| 1 | Dharmaraya Swamy Temple Ward, South Zone, Beng... | 12.957730 | 77.575652 | 8.886156e+06 | 2.961840e+06 | 5840.376700 |
| 2 | Dharmaraya Swamy Temple Ward, South Zone, Beng... | 12.956640 | 77.578175 | 8.886756e+06 | 2.961840e+06 | 5747.173218 |
| 3 | Hindustan Petroleum, Proposed Hudson Cricle - ... | 12.955549 | 77.580699 | 8.887356e+06 | 2.961840e+06 | 5715.767665 |
| 4 | New car designer, Lalbagh Fort Road, Mavalli, ... | 12.954459 | 77.583222 | 8.887956e+06 | 2.961840e+06 | 5747.173218 |
| 5 | Lalbagh Botanical Gardens, ಪಾದಚಾರಿ ಮಾರ್ಗ, Mava... | 12.953369 | 77.585744 | 8.888556e+06 | 2.961840e+06 | 5840.376700 |
| 6 | Lalbagh Botanical Gardens, ಪಾದಚಾರಿ ಮಾರ್ಗ, Mava... | 12.952279 | 77.588267 | 8.889156e+06 | 2.961840e+06 | 5992.495307 |
| 7 | St. Joseph's School, Cottonpete Main Road, K R... | 12.962601 | 77.570305 | 8.884656e+06 | 2.962360e+06 | 5855.766389 |
| 8 | Bangalore Medical College Hospitals, ಸುಲ್ತಾನ್ ... | 12.961510 | 77.572829 | 8.885256e+06 | 2.962360e+06 | 5604.462508 |
| 9 | Srinidhi Sagar, 2, Krishna Rajendra Road, K R ... | 12.960419 | 77.575353 | 8.885856e+06 | 2.962360e+06 | 5408.326913 |

Fig1:Sample of data from dataframe

Now I have used four square API to get the restaurant present in the neighbour . From ('*https://developer.foursquare.com/docs/resources/categories'*).From t his I find out the list of the restaurants in the area and store the values in the v ariable.

Again, from API find out the list of Italian restaurant present in the neighbourhood. And made new data frame from all the extracted data.

| Address | Latitude | Longitude | X | Y | Distance from center | Restaurants in area | Distance to Italian restaurant |
|---|---|---|---|---|---|---|---|
| Fort High School, Krishna Rajendra Road, Maval... | 12.958821 | 77.573128 | 8.885556e+06 | 2.961840e+06 | 5992.495307 | 1 | 5225.267649 |
| Dharmaraya Swamy Temple Ward, South Zone, Beng... | 12.957730 | 77.575652 | 8.886156e+06 | 2.961840e+06 | 5840.376700 | 1 | 4813.516635 |
| Dharmaraya Swamy Temple Ward, South Zone, Beng... | 12.956640 | 77.578175 | 8.886756e+06 | 2.961840e+06 | 5747.173218 | 4 | 4444.824269 |
| Hindustan Petroleum, Proposed Hudson Cricle - ... | 12.955549 | 77.580699 | 8.887356e+06 | 2.961840e+06 | 5715.767665 | 4 | 4130.736395 |
| New car designer, Lalbagh Fort Road, Mavalli, ... | 12.954459 | 77.583222 | 8.887956e+06 | 2.961840e+06 | 5747.173218 | 1 | 3884.521019 |
| Lalbagh Botanical Gardens, ಪಾದಚಾರಿ ಮಾರ್ಗ, Mava... | 12.953369 | 77.585744 | 8.888556e+06 | 2.961840e+06 | 5840.376700 | 3 | 3719.680623 |

Fig1:Sample of data from dataframe along with the distance of the Italian restaurant

From the data from above data frames, I will use K-Means and distribute the map in segments. And again distribute the restaurants in cluster.

At last try to find out which cluster would be good for opening the Italian restaurant.

Data for the latitude and longitude of city we can get data also from kaggle: https://www.kaggle.com/rmenon1998/bangalore-neighborhoods/data#

## 3. Result

- Our analysis shows that although there is a great number of restaurants in Banglore (~2000 in our initial area of interest which was 12x12km around Banglore), there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north and west from Banglore, so we focused our attention to areas south, south-east and east, corresponding to and south-east corner of central area.
- Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

- Result of all this zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Italian restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Banglore center but not crowded with existing restaurants (particularly Italian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

# 4. Conclusion

- Purpose of this project was to identify Bangalore areas close to centre with low number of restaurants (particularly Italian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Italian restaurant. By calculating restaurant density distribution from foursquare data we have first identified general boroughs that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centres were created to be used as starting points for final exploration by stakeholders.
- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighbourhood etc.