Predicting the Apt Place for Italian Restaurant startup in Hyderabad

Suneetha Vanjivakam

Apr 29, 2020

1. Introduction

1.1 Background

This document is targeted to provide accurate predictions for stakeholders interested in starting an Italian restaurant in Hyderabad, India. We will use our data science powers to generate a few most promising neighbourhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders. Some information about Hyderabad history and prominence in IT development in India is as below.

Hyderabad is the capital of the Indian state of Telangana. It is a historic city noted for its many monuments, temples, mosques and bazaars. The city of **Hyderabad** was founded by the Qutab Shahi sultan Muhammad Quli Qutb Shah in 1591 CE. It was built around the Charminar, which formed the centre piece of the city. Hyderabad's central location between the Deccan Plateau and the Western Ghats, and industrialisation throughout the 20th century attracted major Indian research, manufacturing, educational and financial institutions. Since the 1990s, the city has emerged as an Indian hub of pharmaceuticals and biotechnology. The formation of special economic zones and HITEC City dedicated to information technology has encouraged leading multinationals to set up operations in Hyderabad.

1.2 problem

Data that might contribute to predicting the apt location might include latitude , longitude, cuisines of existing restaurants in all areas around 60 kms of the city , Italian restaurant density in existing locations nearer to City Centre , Italian restaurant density in existing locations nearer to City IT Hub , which can attract more International Foodies.

1.3 Interest

Obviously , stakeholders who are interested to start an Italian restaurant in Hyderabad would be interested in accurate prediction of the location.

2. Data acquisition and cleaning

2.1 Data Sources

Most Data related to restaurants locations , existing cuisines , latitudes and longitudes are fetched using Zomato api which is one of the leading online food delivery app in Hyderabad. Four square api is used to fetch the Hyderabad latitudes and longitudes in the project and re-use them to draw the folium maps.

2.2 Data Cleaning

Data Downloaded using Zomato api were combined into one table. There were multiple entries for each location in city, which were identified using latitudes and longitudes. Three parameters were considered for each restaurant such as latitude, longitude, cuisines.

2.3 Feature Selection

After Data Cleaning, Sample Dataset looks like as below, which shows the locality, that is borough name, latitude, longitude, cuisines available in that restaurant which is used to identify an Italian restaurant.

	latitude	longitude	locality	cuisines
0	17.437113	78.399917	Jubilee Hills	Cafe, Bakery, Italian, Desserts
1	17.416618	78.410692	Film Nagar	Desserts, Cafe, Italian
2	17.435744	78.457869	Begumpet	North Indian, Chinese, South Indian, Bakery
3	17.423434	78.431609	Banjara Hills	Cafe, American, Continental, Italian
4	17.433380	78.405089	Jubilee Hills	American, Healthy Food, Mexican, Desserts
...	...	...	...	...
95	17.443548	78.368384	Gachibowli	Bakery
96	17.404941	78.476671	Basheer Bagh	Fast Food, Street Food, Ice Cream, Desserts, B...
97	17.448024	78.378759	Hitech City	Cafe, Desserts
98	17.423808	78.412692	Jubilee Hills	Fast Food, Salad
99	17.436496	78.399494	Jubilee Hills	Italian

100 rows × 4 columns

3 Exploratory Data Analysis

3.3 Predicting the target Location

Based on definition of our problem, factors that will influence our decision are:

- Number of Italian restaurants in each locality
- Distance of Italian restaurants to City Centre or City IT Hub

Based on the above two factors , we will provide a conclusion on location prediction to start Italian restaurant in Hyderabad.

3.4 Relationship between the Restaurant Density in the Locality and Predicted Location

The Restaurant density in the locality is very important criteria to predict the desired location of Italian restaurant start-up . Here is the snapshot of the locality wise Italian restaurant count .

```
Italian_lat_long_df['locality'].value_counts()
```

```
Jubilee Hills                                  24
Gachibowli                                     13
Madhapur                                       11
Banjara Hills                                  11
Hitech City                                     9
Film Nagar                                      7
Begumpet                                        4
Himayath Nagar                                  2
Somajiguda                                      2
Sheraton Hyderabad Hotel, Gachibowli            2
Taj Falaknuma, Falaknuma                        1
Kompally                                        1
Falaknuma                                       1
Miyapur                                         1
Inorbit Mall, Hitech City                       1
Basheer Bagh                                    1
Holiday Inn Express & Suites                    1
Hyderabad Marriott Hotel & Convention Centre    1
Kondapur                                        1
Aalankrita Resort, Shamirpet                    1
Necklace Road                                   1
The Park, Somajiguda                            1
Tolichowki                                      1
ITC Kohenur                                     1
The Golkonda Resorts & Spa, Gandipet            1
Name: locality, dtype: int64
```

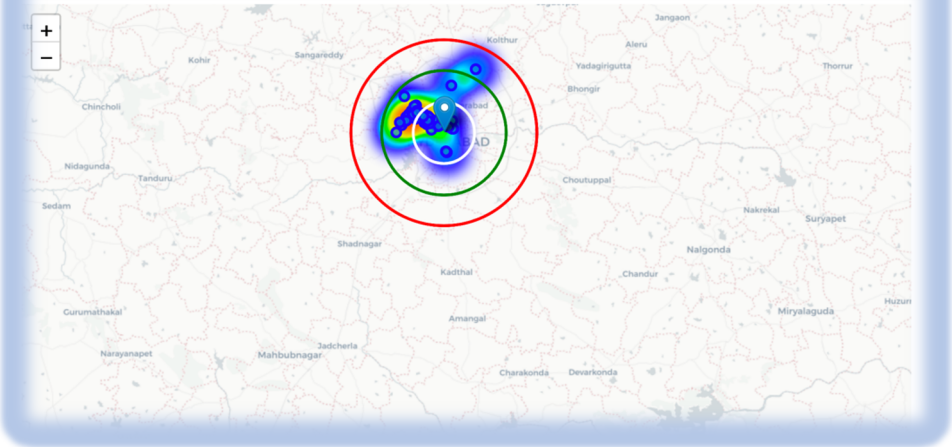
4. K- Means Clustering with HeatMaps

Step 1:

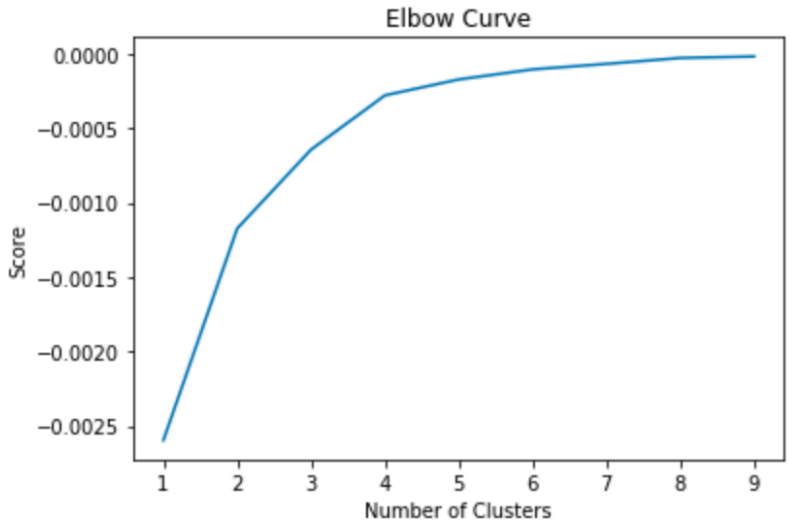
First , we took folium map to show three different regions to plot the latitudes and longitudes of all Italian restaurants with in 10km , 20km , 30km from City Centre as below.

Below Folium map shows that the core cluster of restaurants are located within 10 km from the Centre of the Hyderabad city.

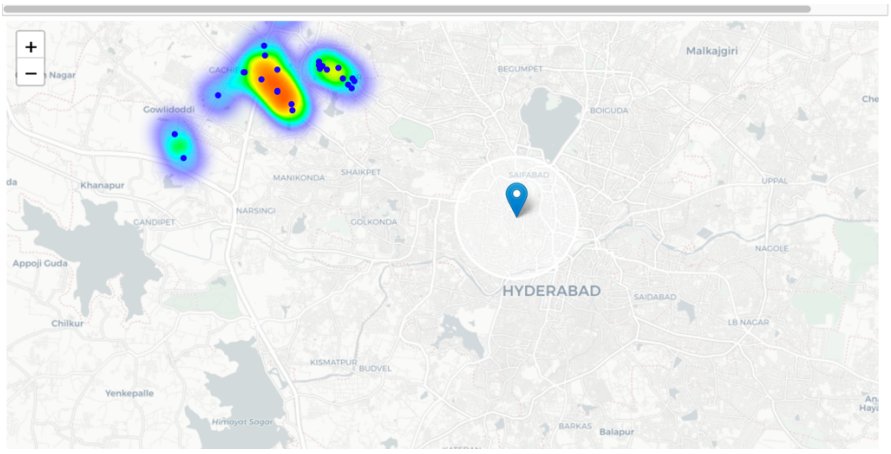
```
map_berlin = folium.Map(location=hyderabad_center, zoom_start=10)
folium.TileLayer('cartodbpositron').add_to(map_Hyderabad) #cartodbpositron cartodbdark_matter
HeatMap(heat_data).add_to(map_Hyderabad)
folium.Marker(hyderabad_center).add_to(map_Hyderabad)
folium.Circle(hyderabad_center, radius=10000, fill=False, color='white').add_to(map_Hyderabad)
folium.Circle(hyderabad_center, radius=20000, fill=False, color='green').add_to(map_Hyderabad)
folium.Circle(hyderabad_center, radius=30000, fill=False, color='red').add_to(map_Hyderabad)
map_Hyderabad
```



Scatter plot from K-Means:



Step 2: Applied Clustering algorithm and we can see that there is one cluster of locations that has less density of Italian restaurants nearer to IT Hub of the city, which can be best considered for starting Italian restaurant in Hyderabad. As there are multinationals working in IT Hub, there are high chances that the restaurant can attract many foodies around the tech park.



5. Conclusion

Hence , we explored the data from Zomato api , to get the list of Italian restaurants by their location co-ordinates , framed the data into a pandas data frame , analyzed the data , plotted the folium maps to show the core density of restaurants in the Hyderabad city and showed that stake holders can consider near IT Hub , rather than city Centre which is already cluttered with Italian restaurants.