Capstone Project - The Battle of the Neighborhoods

Applied Data Science Capstone by IBM/Coursera

Table of contents

Introduction: Business Problem

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted for people that are new to Hyderabad, it can be daunting to figure out what restaurants are worth going to and where they are and for the stakeholders interested in opening a restaurant in Hyderabad, India.

Since there are lots of restaurants in Hyderabad we will try to detect locations that are that are with good likes. We would also prefer locations as close to city center as possible.

We will use our data science powers to generate a few most promissing neighborhoods based on this criteria.

Data

For this assignment, I will be utilizing the Foursquare API to pull the following location data on restaurants in Hyderabad, IND.

1. Venue Name
2. Venue ID
3. Venue Location
4. Venue Category
5. Count of Likes

To acquire the data mentioned above, I will need to do the following:

Get geolocator lat and long coordinates for Hyderabad,IND. Use Foursquare API to get a list of all venues in Hyderabad

- Get venue name,
- venue ID
- location
- category
- likes

Methodology

Get the location coordinates using geocoder

```
from geopy.geocoders import Nominatim
# Get latitude and longitude
address = 'Hyderabad'
geolocator = Nominatim(user_agent="foursquare_agent")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print("Latitude is {} and Longitude is {}".format(latitude,longitude))
```

Use Foursquare API to fetch location data

```
search_query = 'restaurant'
radius = 10000
url =
'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&quer
y={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, latitude, longitude, VERSION,
search_query, radius, LIMIT)
results = requests.get(url).json()
#print(results)
venues=results['response']['venues']
df = json_normalize(venues)
df.head()
```

Perform Data wrangling and make it to required format

	name	categories	address	city	country	crossStreet	distance	formattedAddress	lat	lng	postalCode
0	Akbar Fast Food Restaurant	Indian Restaurant	Adj. Prince Hotel	Hyderabad	India	Meraj Hotel	2076	[Adj. Prince Hotel (Meraj Hotel), Hyderabad 50...	17.394486	78.442452	500028
1	Shah Ghouse Cafe and Restaurant	Indian Restaurant	Toli Chowli	Hyderabad	India	M	3911	[Toli Chowli (M), Hyderabad 500058, Telangana,...	17.396277	78.425094	500058
2	Al Rabea Al Arabi Restaurant	Middle Eastern Restaurant	NaN	NaN	India	NaN	3109	[India]	17.395362	78.432614	NaN
3	Voulga Restaurant	Restaurant	Darus Salaam	Hyderabad	India	NaN	853	[Darus Salaam, Hyderabad, Telangana, India]	17.382614	78.465828	NaN
4	Nimrah Restaurant And Bakery	Restaurant	NaN	Hyderabad	India	NaN	866	[Hyderabad 500004, TG, India]	17.392616	78.468165	500004

Fetch likes for each restaurant using foursquare API

	id	name	lat	lng	categories	total likes
0	4f9433d3e4b0949d21694780	Akbar Fast Food Restaurant	17.3945	78.4425	Indian Restaurant	1
1	4c81163ad92ea093375a3d72	Shah Ghouse Cafe and Restaurant	17.3963	78.4251	Indian Restaurant	75
2	517bee0be4b055745975c517	Al Rabea Al Arabi Restaurant	17.3954	78.4326	Middle Eastern Restaurant	3
3	50d1b77ee4b0d5316b4e0bff	Voulga Restaurant	17.3826	78.4658	Restaurant	0
4	4bcee9d1cc8cd13a5efdc4cf	Paradise Restaurant	17.4418	78.4874	Hyderabadi Restaurant	295

Analysis

Group the restaurants depending on type

```
: # let's star the process for re-categorizing the categories

  hyd_venues['categories'].unique()
```

```
: array(['Indian Restaurant', 'Middle Eastern Restaurant', 'Restaurant',
         'Hyderabadi Restaurant', 'Financial or Legal Service',
         'Andhra Restaurant', 'Vegetarian / Vegan Restaurant',
         'Fast Food Restaurant', 'Albaik Restaurant', 'Asian Restaurant',
         'Afghan Restaurant', 'Kebab Restaurant', 'Vashishta Restaurant',
         'Bar', 'Food', 'Snack Place', 'Bakery', 'Café',
         'Chinese Restaurant', 'Eagle bar & restaurant', 'Swad restaurant',
         'Multicuisine Indian Restaurant'], dtype=object)
```

```
: bars = ['pub', 'Eagle bar & restaurant', 'Bar', 'Dive Bar', 'Sports Bar']
  other = ['Financial or Legal Service','Food']
  indian_food = ['Restaurant','Indian Restaurant','Hyderabadi Restaurant',
  Cafe = ['Snack Place', 'Bakery', 'Café',]
  american_food = ['Burger Joint','Fast Food Restaurant','American Restaura
  chinese_food= ['Chinese Restaurant','Asian Restaurant']
  eastern_food=['Middle Eastern Restaurant','Albaik Restaurant','Afghan Res
```

Rate the restaurant food quality as poor, average and good depending on total number of likes

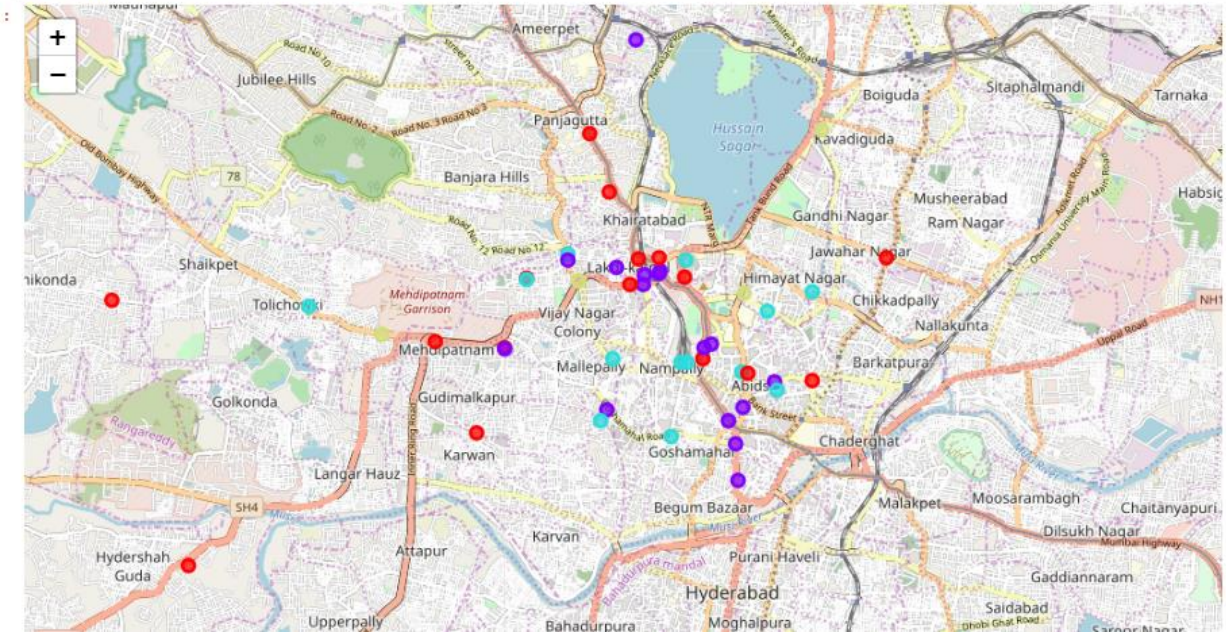
hyd_venues

	id	name	lat	lng	categories	total likes	total likes_cat	categories_new
0	4f9433d3e4b0949d21694780	Akbar Fast Food Restaurant	17.3945	78.4425	Indian Restaurant	1	avg avg	indian food
1	4c81163ad92ea093375a3d72	Shah Ghouse Cafe and Restaurant	17.3963	78.4251	Indian Restaurant	75	great	indian food
2	517bee0be4b055745975c517	Al Rabea Al Arabi Restaurant	17.3954	78.4326	Middle Eastern Restaurant	3	avg avg	eastern_food
3	50d1b77ee4b0d5316b4e0bff	Voulga Restaurant	17.3826	78.4658	Restaurant	0	poor	indian food
4	4bcee9d1cc8cd13a5efdc4cf	Paradise Restaurant	17.4418	78.4874	Hyderabadi Restaurant	295	great	indian food
5	5a21602775a6ea748fd77429	Nimrah Restaurant And Bakery	17.3926	78.4682	Restaurant	0	poor	indian food
6	5823557cda82023188157a2f	Grills restaurant	17.3861	78.4568	Financial or Legal Service	0	poor	other
7	5686bc36498ece24b0f97da8	Zoha Restaurant	17.3931	78.4576	Hyderabadi Restaurant	0	poor	indian food
8	5a1ade2ae96d0c5d8b3dc3da	Sri Anupama Family Restaurant	17.3848	78.4559	Andhra Restaurant	0	poor	indian food
9	4eb54b825c5c5a531ff0d6cd	Woodland Restaurant	17.4021	78.4858	Vegetarian / Vegan Restaurant	0	poor	indian food
10	50ba59a4e4b077f48d74046b	Azizia Restaurant	17.3926	78.4673	Restaurant	0	poor	indian food
11	4e4fb510b0fb088f3c1ece7f	Paradise Restaurant	17.4036	78.4528	Indian Restaurant	13	great	indian food
12	52f8e44c498ea024d4f989e2	Dhruva Bar and Restaurant	17.4009	78.3869	Restaurant	2	avg avg	indian food
13	4df9c2f6d4c064db03aa050d	Venue, vegetarian Restaurant	17.393	78.4704	Indian Restaurant	1	avg avg	indian food
14	51d53e5e498e120a89aa3aab	Dwaraka (Veg Restaurant)	17.4233	78.4543	Indian Restaurant	2	avg avg	indian food
15	5406d59f498e5c9379534db9	Venkey's Veg - A Multi Cuisine Restaurant	17.3946	78.4706	Fast Food Restaurant	0	poor	american food

Results and Discussion

Cluster the restaurants using k-means and visualize the results on map using folium.

```
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)
map_clusters
```



Cluster 1
Characterstics --> Poor quality food

```
hyd_venues.loc[hyd_venues['label']==0]
```

	id	name	lat	lng	categories	total likes	total likes_cat	categories_new	label
6	5823557cda82023188157a2f	Grills restaurant	17.3861	78.4568	Financial or Legal Service	0	poor	other	0
15	5406d59f498e5c9379534db9	Venkey's Veg - A Multi Cuisine Restaurant	17.3946	78.4706	Fast Food Restaurant	0	poor	american food	0
17	51acc1b4498e59e850eaf731	Albaik Restaurant	17.3951	78.4715	Albaik Restaurant	0	poor	eastern_food	0
19	4df074bad16486e86e503a2c	Hanky Panky Family Restaurant	17.3847	78.474	Fast Food Restaurant	0	poor	american food	0
20	548728f3498e54a2989b48ab	Kabul Darbar Restaurant	17.403	78.462	Afghan Restaurant	0	poor	eastern_food	0
24	4df63e0bb61cbd3ec0246d14	Vaishnav's Restaurant	17.436	78.4609	Food	0	poor	other	0
25	4f69a816e4b039c9e5de7ab8	Agrawala Bakers And Restaurant	17.3865	78.476	Snack Place	0	poor	Cafe	0
27	51d286af498efb08f876c860	Seena Restaurant & Bakery	17.3817	78.475	Bakery	0	poor	Cafe	0
29	5406c55c498e2804ff8ad973	Carnival Restaurant	17.4044	78.4621	Fast Food Restaurant	0	poor	american food	0

cluster 2
Characterstics --> Average quality food

```
hyd_venues.loc[hyd_venues['label']==1]
```

	id	name	lat	lng	categories	total likes	total likes_cat	categories_new	label
0	4f9433d3e4b0949d21694780	Akbar Fast Food Restaurant	17.3945	78.4425	Indian Restaurant	1	avg avg	indian food	1
2	517bee0be4b055745975c517	Al Rabea Al Arabi Restaurant	17.3954	78.4326	Middle Eastern Restaurant	3	avg avg	eastern_food	1
12	52f8e44c498ea024d4f989e2	Dhruva Bar and Restaurant	17.4009	78.3869	Restaurant	2	avg avg	indian food	1
13	4df9c2f6d4c064db03aa050d	Venue, vegetarian Restaurant	17.393	78.4704	Indian Restaurant	1	avg avg	indian food	1
14	51d53e5e498e120a89aa3aab	Dwaraka (Veg Restaurant)	17.4233	78.4543	Indian Restaurant	2	avg avg	indian food	1
18	4ce8f32ae888f04d447c456b	Crystal Restaurant	17.4066	78.4963	Asian Restaurant	4	avg avg	chinese_food	1
21	575d8e7d498e5ed05b88906d	Bademiyan Kababs Restaurant	17.4031	78.46	Kebab Restaurant	3	avg avg	eastern_food	1
23	4df0f5b6d16486e86e5289e8	Narmada Bar & Restaurant	17.3911	78.4767	Bar	2	avg avg	bars	1
28	4e7f60e8be7b3ac644dbaa8b	Nasheman Restaurant	17.4039	78.4455	Indian Restaurant	2	avg avg	indian food	1

Cluster 3
Characterstics --> Great quality food

```
hyd_venues.loc[hyd_venues['label']==2]
```

	id	name	lat	lng	categories	total likes	total likes_cat	categories_new	label
1	4c81163ad92ea093375a3d72	Shah Ghouse Cafe and Restaurant	17.3963	78.4251	Indian Restaurant	75	great	indian food	2
4	4bcee9d1cc8cd13a5efdc4cf	Paradise Restaurant	17.4418	78.4874	Hyderabadi Restaurant	295	great	indian food	2
11	4e4fb510b0fb088f3c1ece7f	Paradise Restaurant	17.4036	78.4528	Indian Restaurant	13	great	indian food	2
16	513afb90e4b04d69fd7cc3c7	Okra Restaurant	17.4238	78.4873	Restaurant	20	great	indian food	2
44	4dfa1aaaaeb785aedbf0a16b	New Madhushala Bar & Restaurant	17.3979	78.4841	Bar	10	great	bars	2

Conclusion

The venues have been identified using Foursquare API, categorized, clustered and have been plotted on the map. The map reveals restaurants which are exceptionally good in Hyderabad Based on the visitor’s venue rating and price preferences. he/she can choose amongst the places.