

Business Opportunities in Hyderabad

Manoj Ragala

2020/11/16

1. Introduction

Utilizing the Location based and venue related data for Hyderabad city, perform an analysis to find and report different business opportunities location wise, so that whenever someone wants to open a restaurant or a shopping mall, they can make use of this analysis to get insights and ideas about where to start the business within Hyderabad. This will help the people to make a better decision about their business faster and also its very insightful as well.

We will study about the existing businesses, where they are located and how they were distributed across the city using the information obtained through Foursquare API and Wikipedia data.

2. Data Collection and Curation

We will get the data using the following sources

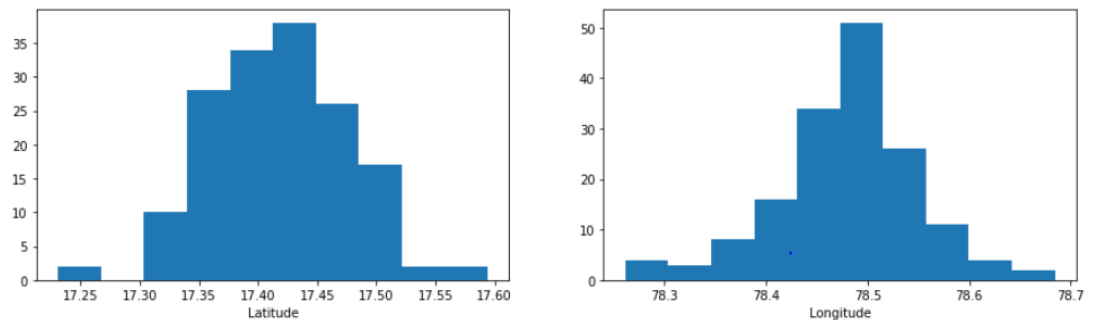
1. https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Hyderabad
2. <https://geopy.readthedocs.io/en/stable/>
3. Foursquare API

The first one has the data related to different locations and neighbourhoods in Hyderabad, we will copy or scrape the data to create a dataset in a meaningful form having 2 columns Neighbourhood and section of the city, then we will load the data into a python session and use the geopy library methods to fetch the latitudes and longitudes of different neighbourhoods in the city. Then we can use the Foursquare API to get the different venue details around each geographical location. For each API call a JSON file will be returned, it's better to understand the json data using a json viewer to better understand the structure of the file which makes it easy when we create our final dataset. Utilizing the API we make calls with latitude and longitude to get venue details for each neighbourhood. Then we pre-process the data to extract relevant features like id, name, latitude, longitude, category id, category name from the JSON file for each of venue. The following is a sample table that showcases the venue related information.

venue_name	venue_lat	venue_lon	venue_cat_id	venue_cat_name
Barista (Prasads Imax)	17.413033	78.465625	4bf58dd8d48988d1e0931735	Coffee Shop
Wonton	17.441918	78.357974	4bf58dd8d48988d145941735	Chinese Restaurant
Subway	17.425549	78.457191	4bf58dd8d48988d1c5941735	Sandwich Place
The Spicy Venue	17.435141	78.411046	4bf58dd8d48988d10f941735	Indian Restaurant
78 Degrees East	17.417746	78.543179	4bf58dd8d48988d10f941735	Indian Restaurant
Subway	17.401094	78.486590	4bf58dd8d48988d1c5941735	Sandwich Place
M.E.S colony park	17.497960	78.507471	4bf58dd8d48988d163941735	Park
Vivanta by Taj	17.443498	78.460556	4bf58dd8d48988d1fa931735	Hotel
Courtyard Hyderabad	17.424329	78.486719	4bf58dd8d48988d1fa931735	Hotel
blue sea tea & Snacks	17.436877	78.505008	4bf58dd8d48988d1dc931735	Tea Room

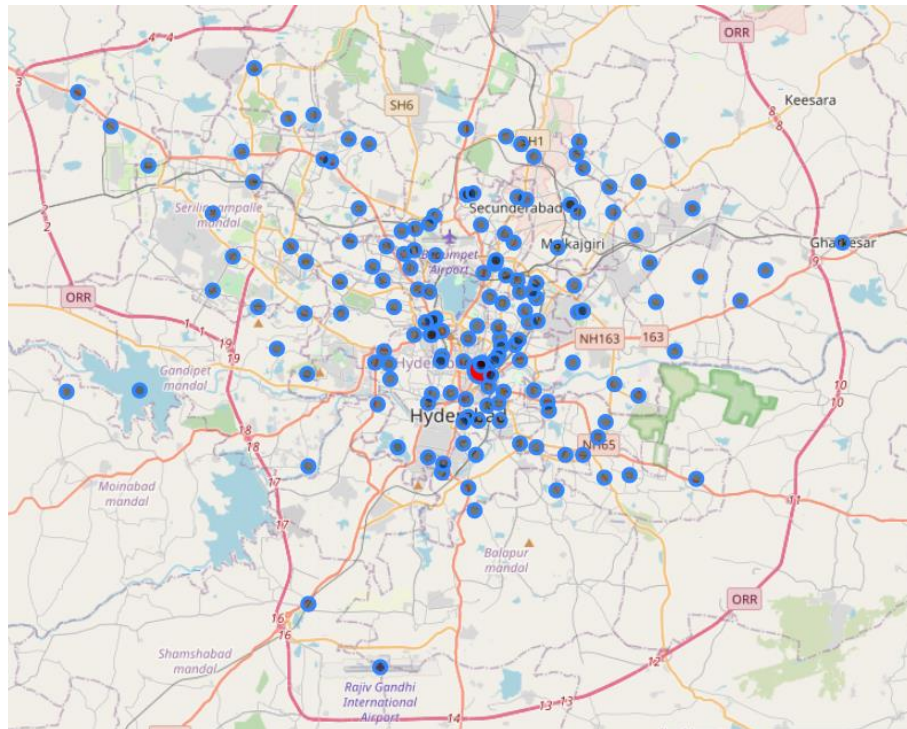
3. Exploratory Data Analysis

Let's have a look at the Latitude and Longitude distributions of neighbourhoods in Hyderabad.

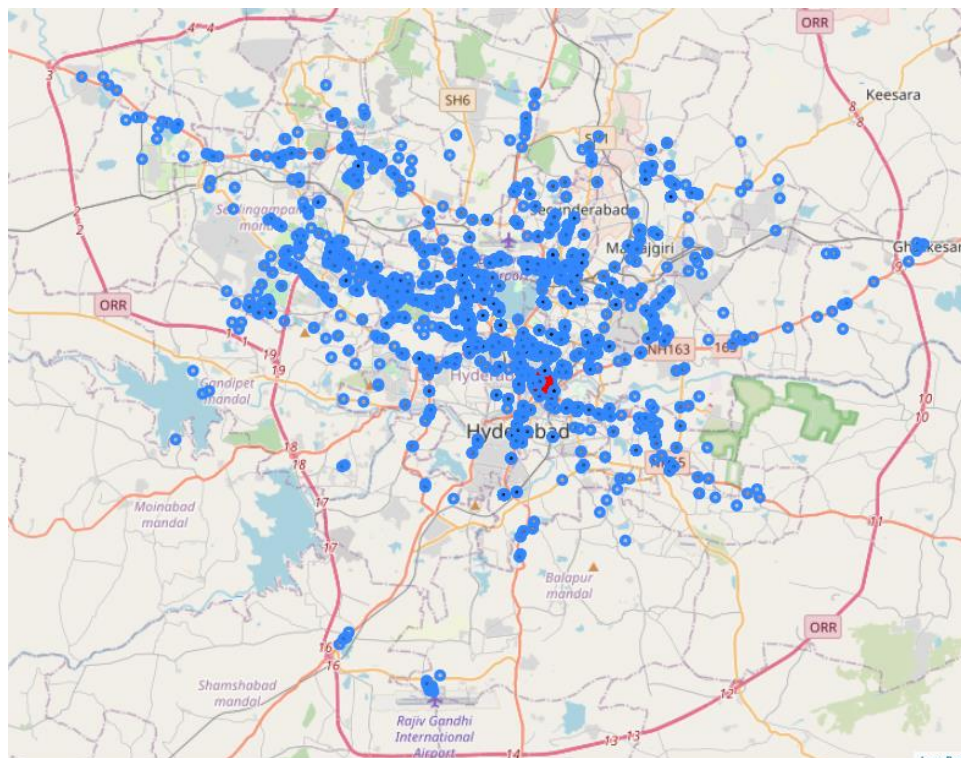


Hyderabad is 1,566 kilometres south of Delhi, 699 kilometres southeast of Mumbai, and 570 kilometres north of Bangalore by road. Situated in the southern part of Telangana in south eastern India, along the banks of the Musi River, located on the Deccan Plateau in the northern part of South India. Greater Hyderabad covers 625 km², making it one of the largest metropolitan areas in India. With an average altitude of 542 metres, Hyderabad lies on predominantly sloping terrain of grey and pink granite, dotted with small hills, the highest being Banjara Hills at 672 metres. The city has numerous lakes referred to as sagar, meaning "sea". Examples include artificial lakes created by dams on the Musi, such as Hussain Sagar, Osman Sagar and Himayat Sagar.

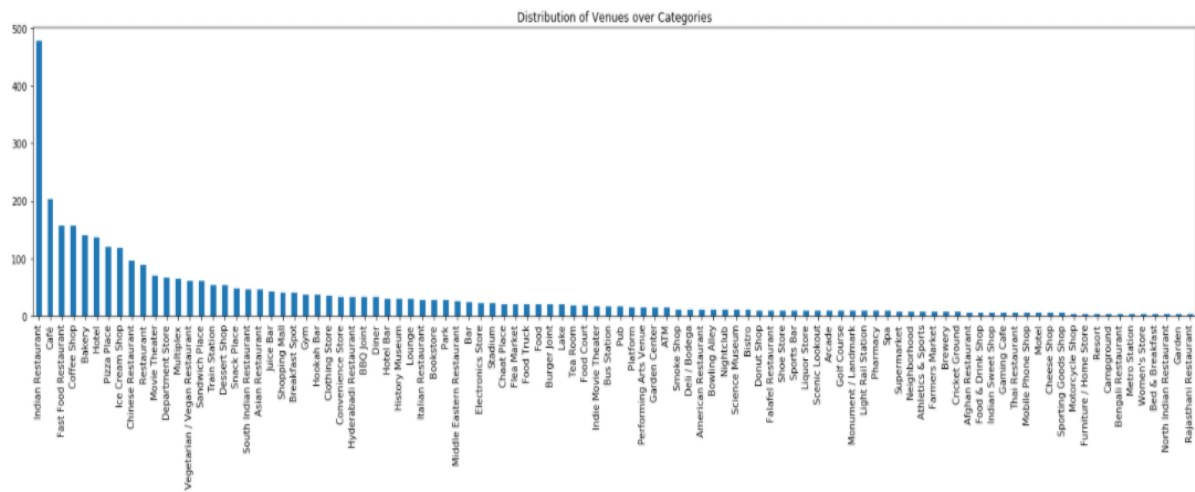
Hyderabad has vast number of neighbourhoods spread across the city, let's have a look at how these neighbourhoods are distributed



Lets have a look at how different venues were spread across the city.



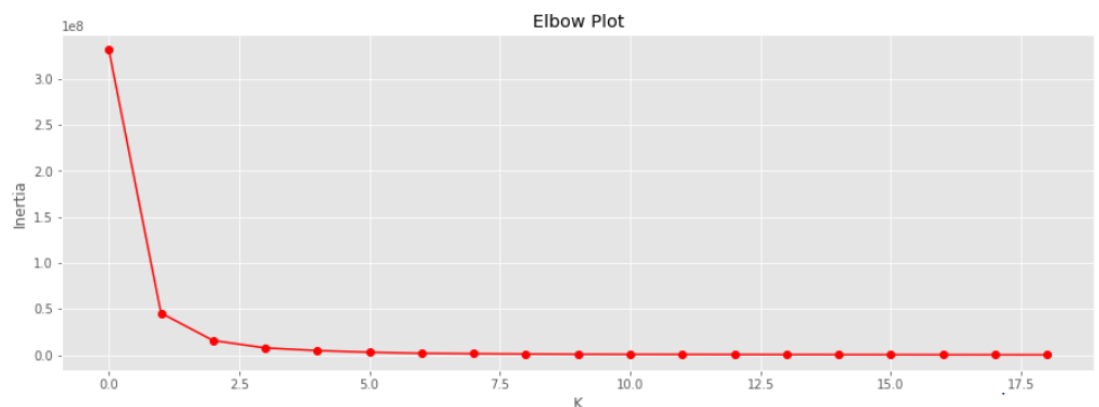
There are around 180 different venue categories, Lets look at how they are spread across the city.



Hyderabad City is divided into several sections. We have Central, Eastern, Western, Northern, North Eastern, North Western, South Eastern, South Western and Old City.

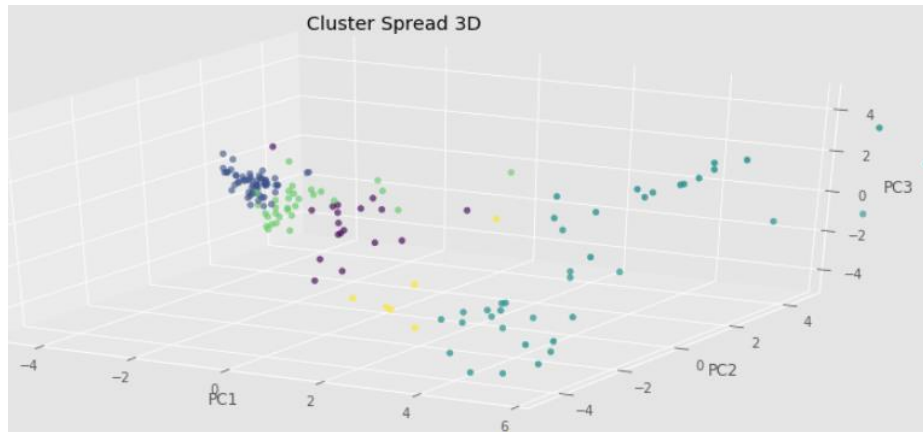
4. Clustering

Using the pandas library create the dummy venue category variables of the dataset, applying group by on neighbourhoods and summation on these dummy variables will give us a new dataset with all numeric variables except for the neighbourhoods, after standardizing the dataset lets create an elbow plot to get an idea about number of clusters we can create.



From the above plot we can confidently select 4 or 5 clusters, lets select 5.

Applying Kmeans Clustering on the dataset will create 5 clusters of the neighbourhoods. Let's look at these clusters, using PCA algorithm we can reduce the dimensionality of the dataset without losing much of the variance lets create 3 principal components and visualize these clusters.



Also let's look at how the neighbourhoods within each cluster and major venue contributors per cluster.

Cluster 1 :

Cluster	venue_cat_name	neighborhood
0	Indian Restaurant	66
1	Café	22
2	Coffee Shop	22
3	Bakery	21
4	Fast Food Restaurant	17
5	Restaurant	15
6	Chinese Restaurant	13
7	Movie Theater	13
8	Ice Cream Shop	12
9	Pizza Place	12

```
['adikmet' 'begum bazaar' 'chilkalguda' 'dar-ul-shifa' 'erragadda'
'gautham nagar' 'golnaka' 'gudimalkapur' 'habsiguda' 'hafeezpet'
'kakatiya nagar' 'musheerabad' 'padmanabha nagar colony' 'padmarao nagar'
'parsigutta' 'sikh village' 'sitaphalmandi' 'tarnaka']
```

Cluster 2 :

Cluster	venue_cat_name	neighborhood
80	Indian Restaurant	38
81	Train Station	21
82	Café	18
83	Bakery	14
84	Department Store	14
85	Pizza Place	14
86	ATM	13
87	Breakfast Spot	12
88	Coffee Shop	11
89	Fast Food Restaurant	11

```
['alwal' 'ammuguda' 'attapur' 'bachupally' 'badangpet' 'bairamalguda'
'bandlaguda' 'barkas' 'bhel township' 'boduppal' 'bowenpally' 'champapet'
'chanchalguda' 'chandrayan gutta' 'cherlapally' 'dammaiguda' 'falaknuma'
'ferozguda' 'gandipet' 'ghatkesar' 'hastinapuram' 'hayathnagar'
'jagadgirigutta' 'kandlakoya' 'karmanghat' 'kismatpur' 'kushaiguda'
'lal darwaza' 'langar houz' 'lothkunta' 'madannapet' 'madina, hyderabad'
'malkajgiri' 'mallapur' 'meerpet' 'mir alam tank' 'moosapet' 'mylargadda'
'nagole' 'narapally' 'nawab saheb kunta' 'old alwal' 'old bowenpally'
'patancheru' 'peerzadiguda' 'pocharam' 'pragathi nagar' 'puppalguda'
'purana pul' 'ramanthapur' 'rc puram' 'santoshnagar' 'serilingampally'
'shamirpet' 'umdanagar' 'uppuguda' 'vanasthalipuram' 'yapral']
```

Cluster 3 :

	Cluster	venue_cat_name	neighborhood
173	2	Indian Restaurant	290
174	2	Café	113
175	2	Hotel	110
176	2	Fast Food Restaurant	94
177	2	Coffee Shop	90
178	2	Bakery	88
179	2	Ice Cream Shop	86
180	2	Chinese Restaurant	75
181	2	Restaurant	57
182	2	Multiplex	49

['a.c. guards' 'abids' 'aghapura' 'ameerpet' 'badichowdi'
 'bagh lingampally' 'balkampet' 'banjara hills' 'barkatpura' 'bazarghat'
 'begumpet' 'chikkadpally' 'chintal basti' 'domalguda' 'film nagar'
 'gachibowli' 'hitec city' 'hyderguda' 'jubilee hills' 'kachiguda'
 'khairatabad' 'kphb' 'kukatpally' 'lakdikapool' 'madhapur' 'marredpally'
 'masab tank' 'mehdipatnam' 'mg road (james street)' 'nampally' 'patny'
 'punjagutta' 'rp road' 'rtc x roads' 'saifabad' 'secunderabad'
 'somajiguda' 'sr nagar' 'srinagar colony' 'sultan bazar' 'vikrampuri'
 'yousufguda']

Cluster 4 :

	Cluster	venue_cat_name	neighborhood
290	3	Indian Restaurant	54
291	3	Pizza Place	39
292	3	Café	35
293	3	Fast Food Restaurant	29
294	3	Department Store	22
295	3	Coffee Shop	16
296	3	Bakery	14
297	3	History Museum	13
298	3	Movie Theater	12
299	3	Train Station	11

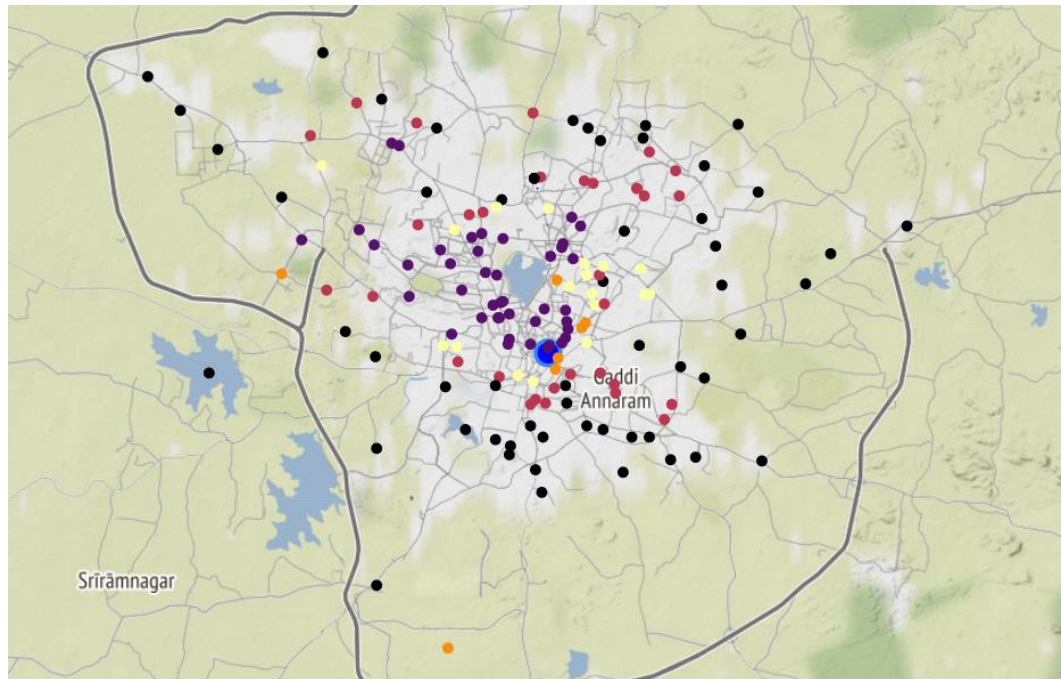
['a. s. rao nagar' 'alijah kotla' 'allwyn colony' 'borabanda' 'dabirpura'
 'dhoolpet' 'dilsukhnagar' 'fateh nagar' 'gaddiannaram' 'hasmathpet'
 'jamia osmania' 'jeedimetla' 'karkhana' 'karwan' 'khajaguda' 'lb nagar'
 'malakpet' 'miyapur' 'moghalpura' 'moosarambagh' 'moula-ali'
 'namalagundu' 'neredmet' 'nizampet' 'old neredmet' 'raidurg' 'safilguda'
 'sainikpuri' 'sanathnagar' 'saroornagar' 'trimulgherry' 'yakutpura']

Cluster 5 :

	Cluster	venue_cat_name	neighborhood
376	4	Indian Restaurant	29
377	4	Coffee Shop	17
378	4	Café	15
379	4	Hotel	12
380	4	Movie Theater	11
381	4	Ice Cream Shop	10
382	4	Pizza Place	7
383	4	Platform	6
384	4	Sandwich Place	6
385	4	Snack Place	6

['azampura' 'chaderghat' 'kavadiguda' 'nallakunta' 'nanakramguda'
 'rajiv gandhi international airport' 'shanker mutt']

Let's look at how neighbourhoods were clustered geographically.



5. Conclusions

Few of the observations after clustering are

- 1) There is a pattern in how the clusters were grouped, the clusters have spread from center to outside of the city in a layered fashion.
- 2) It's interesting to see that neighbourhoods of Old city away from City central were grouped into a cluster also neighbourhoods away from the center were grouped into a cluster.
- 3) Though geographically near, neighbourhoods got clustered depending on the kind of venues they have.

After doing the analysis its evident that Hyderabad has high density of food related venues, they are everywhere. For someone who wants to start a food related business it would be great for them to open their restaurant in above mentioned cluster 3 neighbourhoods like Gachibowli, Hitech City, Jubilee hills, Film nagar. For others who are quite not sure, they can start a Hotel in Cluster 3 or Cluster 4 neighbourhoods, or they can even start restaurants in places where they are scarce like in cluster 5 neighbourhoods. Since the food related venues are ubiquitous, it's better to think about other businesses like super markets, theatres, small scale markets etc.

6. Future Directions

We can extend this analysis by getting the population data per neighbourhood and do further analysis to understand the customer base for each kind of business. Also, we can combine most trending venues data with the population data to understand what makes a successful business.