Predicting Location for Establishing a Restaurant in Hyderabad City

Applied Data Science Capstone Project by

Sandeep Kumar Vengala

Problem Statement

- To find an optimal location for setting up a restaurant.
- Specifically, this project is targeted to stakeholders interested in setting up any food-joint/restaurant in Hyderabad, Telangana, India.
- We try to detect locations that are not already crowded with restaurants and are as close to the city centre as possible.

Data Acquisition

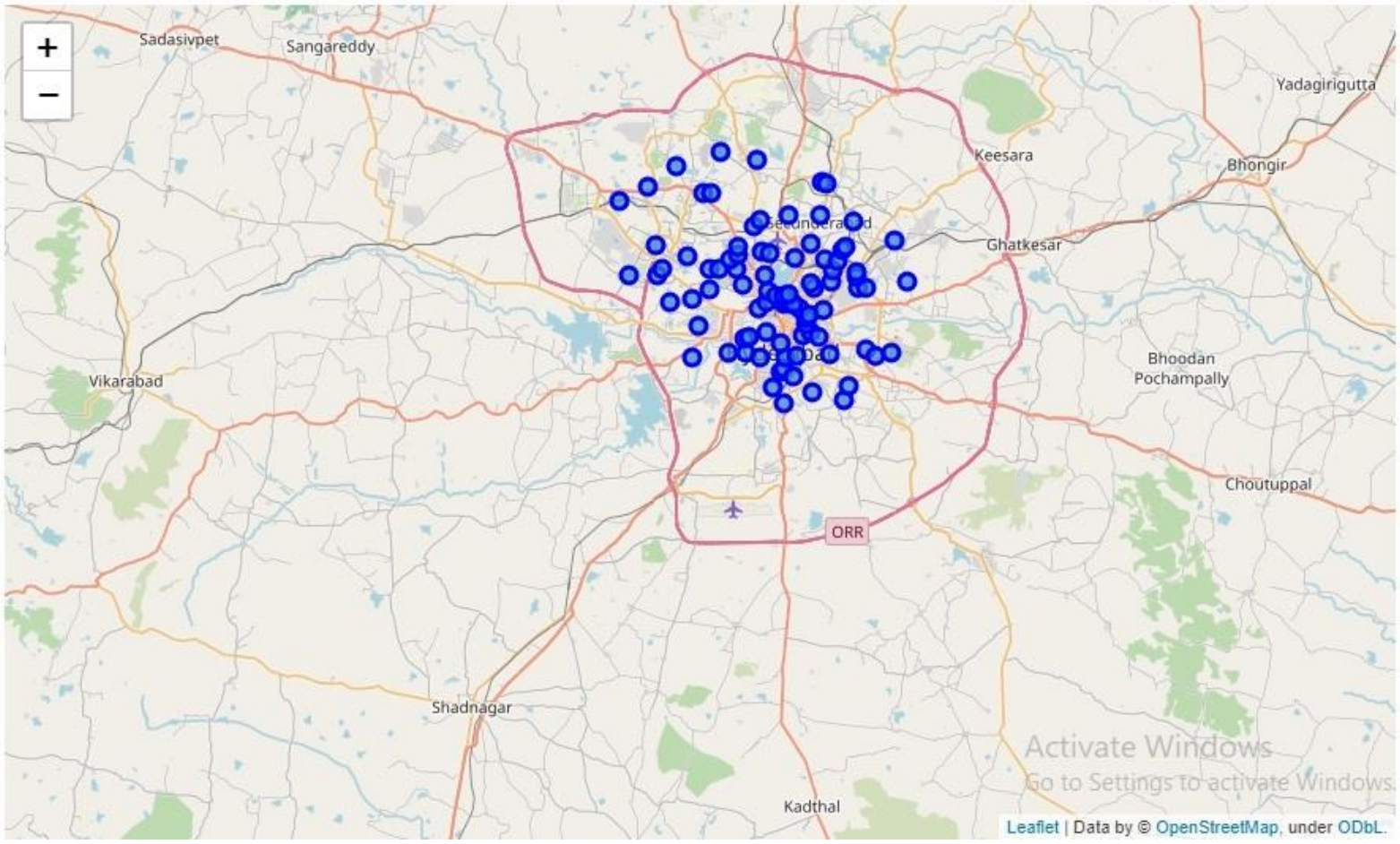
- **Following data sources will be needed to extract/generate the required information:**
- names of all neighbourhoods will be extracted from the website 'http://pincode.india-server.com/cities/hyderabad/' using Pandas package in python
- approximate addresses of centres of these neighbourhoods will be obtained using GeoPy Geocoder package in python
- number of restaurants and their type and location in every neighbourhood will be obtained using Foursquare API

Data Cleaning

The initial 234 neighbourhoods were filtered down to 41 neighbourhoods with a 3-stage filtering process:

- First, 234 neighbourhoods were filtered down to 128 due to lack of location data.
- Then, they were further filtered down to 89 neighbours with the criterion that only neighbourhoods within 25 Kms from the city centre are to be considered.
- Finally, these neighbourhoods were furthered filtered down to 41 due to lack of data for proper clustering of neighbourhoods.

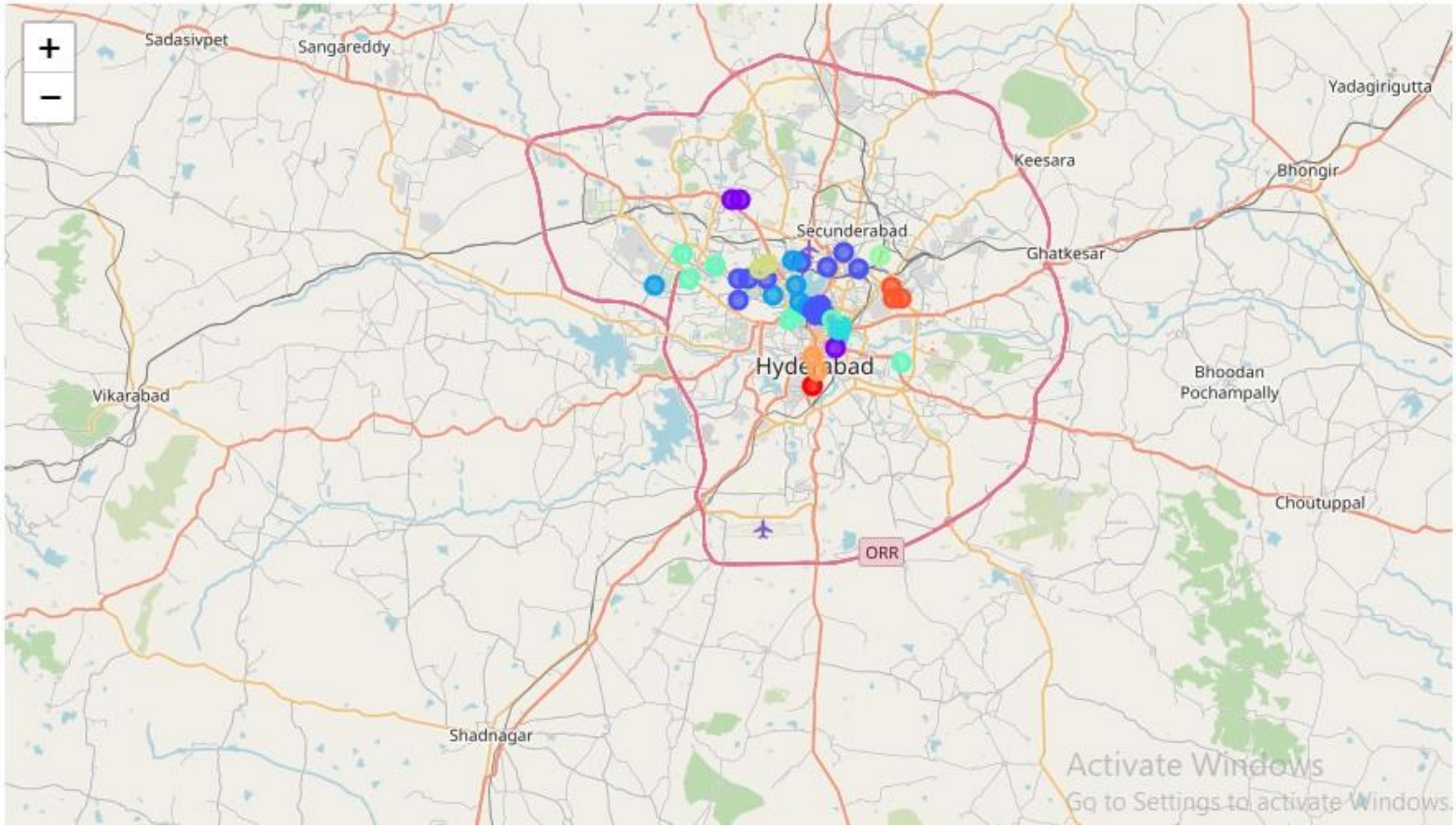
Map of Hyderabad with relevant neighbourhoods



DataFrame fed to K-Means Clustering Algorithm

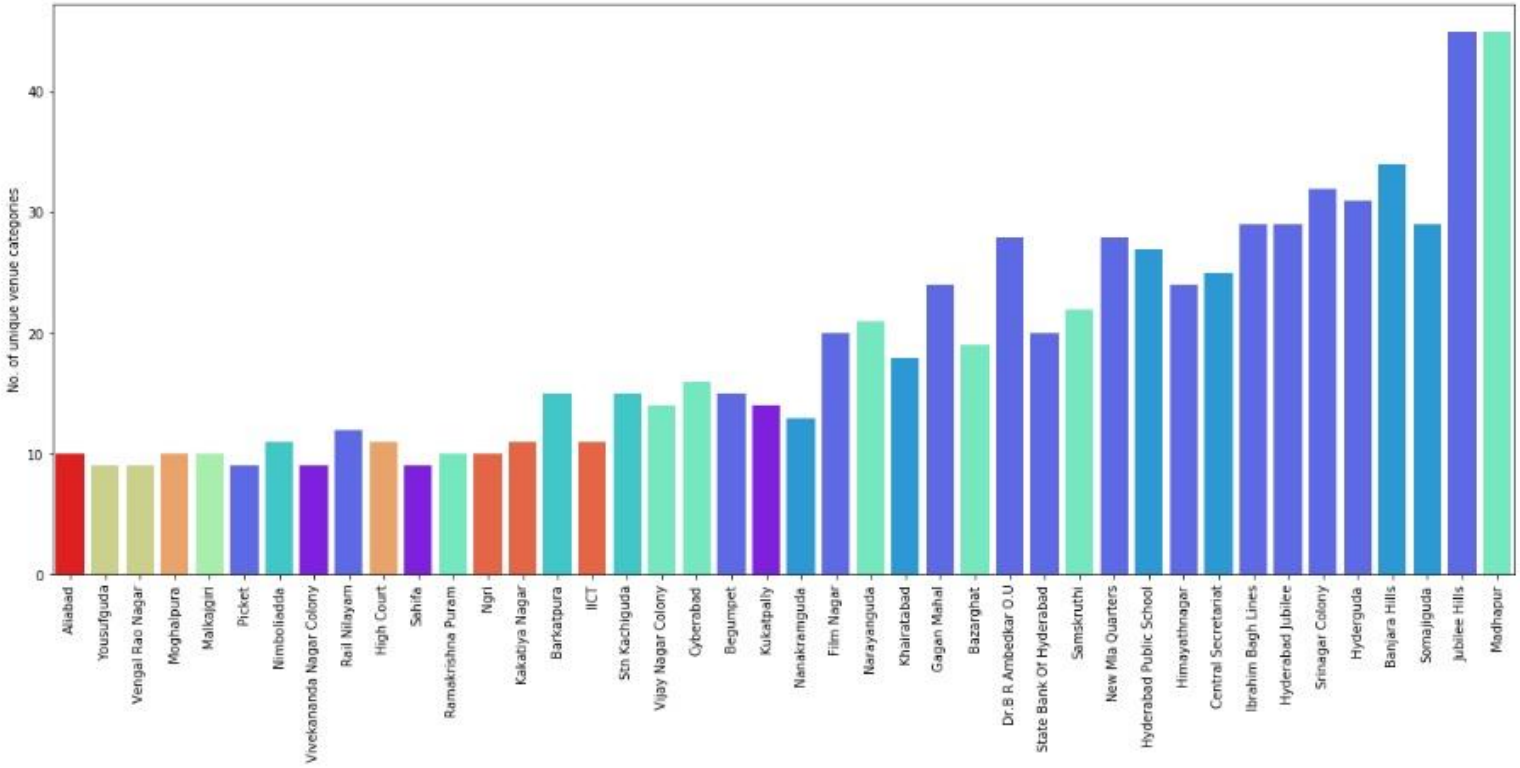
	Neighborhood_Name	Accessories Store	Afghan Restaurant	Airport	American Restaurant	Andhra Restaurant	Arcade	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	...	Sporting Goods Shop	Sports Bar
0	Aliabad	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.100000	0.000000	...	0.000000	0.000000
1	Banjara Hills	0.000000	0.000000	0.00	0.017241	0.000000	0.000000	0.000000	0.017241	0.000000	...	0.000000	0.000000
2	Barkatpura	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000
3	Bazarghat	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000
4	Begumpet	0.000000	0.000000	0.05	0.000000	0.000000	0.050000	0.000000	0.050000	0.000000	...	0.000000	0.000000
5	Central Secretariat	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.025000	0.000000	0.000000	...	0.000000	0.000000
6	Cyberabad	0.000000	0.000000	0.00	0.000000	0.050000	0.000000	0.000000	0.100000	0.000000	...	0.000000	0.000000
7	Dr.B R Ambedkar O.U	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.027778	0.000000	0.000000	...	0.027778	0.000000
8	Film Nagar	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.043478	0.000000	...	0.000000	0.000000
9	Gagan Mahal	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.031250	0.000000
10	High Court	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000
11	Himayathnagar	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.025000	0.000000
12	Hyderabad Jubilee	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000
13	Hyderabad Public School	0.000000	0.000000	0.00	0.000000	0.000000	0.025000	0.000000	0.000000	0.000000	...	0.025000	0.000000
14	Hyderguda	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.020000	0.000000
15	IICT	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000

Map of Hyderabad with neighborhoods colour-coded according to their clusters.



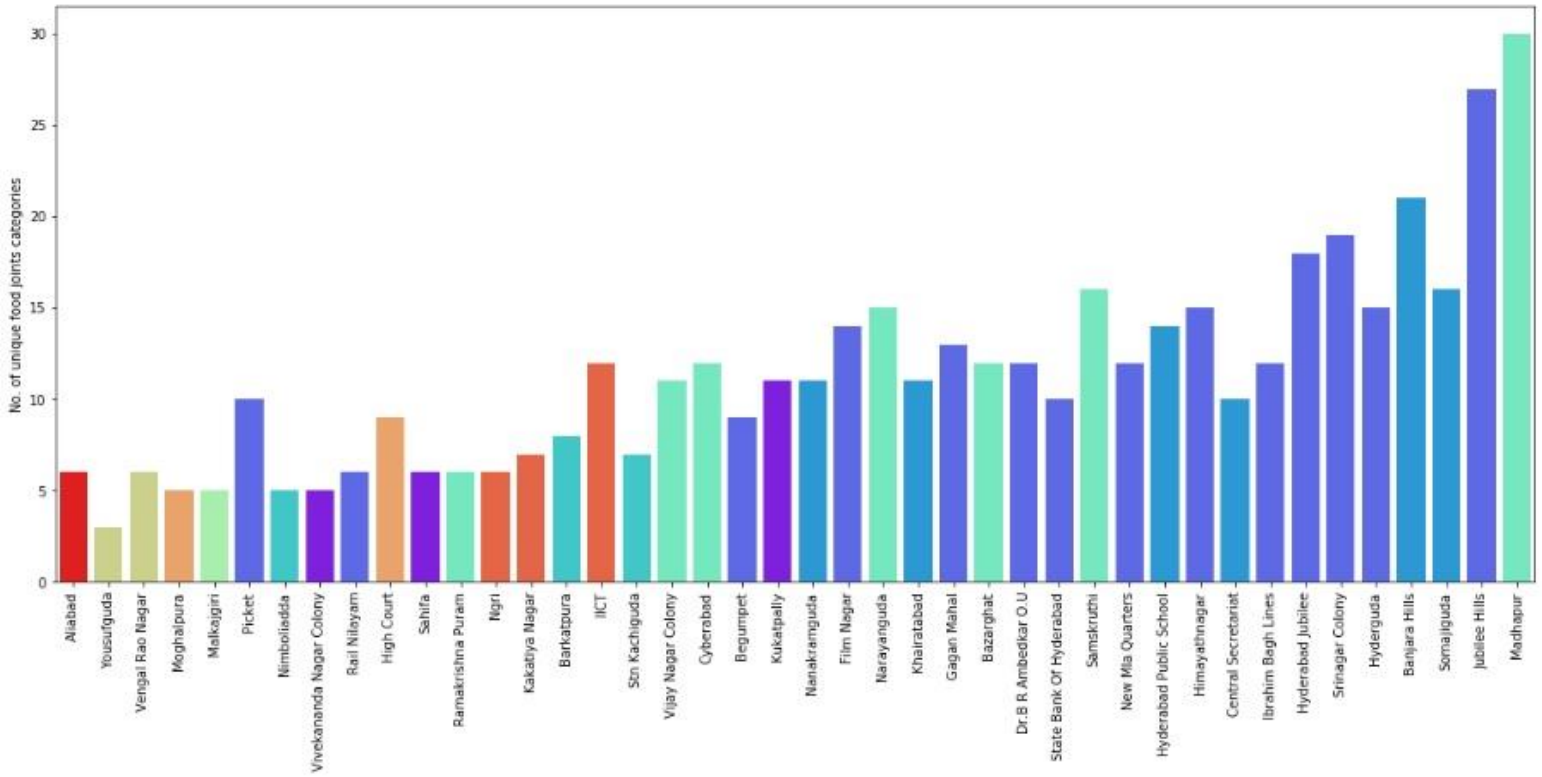
Bar graph

- **The bar graph below shows the diversity of amenities along Y axis and corresponding neighborhood names along X axis. The bars are color coded according to their cluster number.**



Bar graph

- **The bar graph below shows the number of food joints along Y axis and corresponding neighborhood names along X axis. The bars are color coded according to their cluster number.**



Conclusion

- **Optimal locations were identified as those neighbourhoods which are part of the best clusters (having neighbourhoods with most diverse amenities) yet have the least number of food joints. The optimal locations obtained are:**
  - Rail Nilayam
  - RamaKrishna Puram
  - PicketB
  - Begumpet
  - Cyberabad

- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighbourhoods and their locations, 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.