

Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Restaurant in Creativity City of Gastronomy
(Hyderabad), India



Introduction

Hyderabad, India is not only known as the 'City of pearls' but recently it has also been awarded by UNESCO as the 'Creative city of Gastronomy'. It is one of India's most visited tourist places with lots of tourist attraction spots like Charminar, Safdarjung Museum, Golconda Fort etc. Also, people visiting this place never forget to miss out on its worldwide famous cuisine 'Hyderabadi Biryani'. Restaurants are like a non-stop destination for all the food lovers. The central location and the large crowd gatherings near shopping malls or Inox (Cine theatre) provide great advantage and distribution channels for restaurants. As a result, there are many chains of restaurants in the city of Hyderabad and many more are coming up. Opening a restaurant allows property developers to earn consistent rental income. Of course, as with any business decision, opening a new restaurant requires serious consideration and lots of planning. Particularly, the location of the restaurants is one of the most important decisions that will determine whether the restaurant will be a success or a failure.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Hyderabad, India to open a new Indian restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Hyderabad, India, if a business person is looking to open a new restaurant, where would you recommend that he open it? If in case there are already two Indian restaurants which have good ratings, will it be risky to open new one near these restaurants? What all factors will help him to run his business above average? (Out of scope for this project: Budget for Kitchen tools, restaurant furniture and decorations, hire new chef's and waiters, budget everybody's salary. Decide on Menu details, restaurant advertisement, publish discount coupons, restaurant website.)

Target Audience of this project

This project is particularly useful to property developers, investors or business men or a freelancer looking to open or invest in new Indian restaurants in the capital city of Telengana i.e. Hyderabad. This project is timely as the city has been recently awarded by UNESCO as creative city of Gastronomy

Data

To solve the problem we need the following data:

- List of neighbourhoods in Hyderabad. This defines the scope of this project which is confined to the city of Hyderabad(excluding Secunderabad), the capital city of the Telengana in South India.

Neighborhood	
0	A. S. Rao Nagar
1	A.C. Guards
2	Abhyudaya Nagar
3	Abids
4	Adikmet
5	Afzal Gunj
6	Aghapura
7	Aliabad, Hyderabad
8	Alijah Kotla
9	Allwyn Colony

- Latitudes and Longitudes coordinates of those neighbourhoods. This is required in order to plot the map and also get the venue data.

	Neighborhood	Latitude	Longitude
0	A. S. Rao Nagar	17.455470	78.543890
1	A.C. Guards	17.392977	78.456867
2	Abhyudaya Nagar	17.337590	78.562350
3	Abids	17.389800	78.476580
4	Adikmet	17.410600	78.515150
5	Afzal Gunj	17.377510	78.480050
6	Aghapura	17.387349	78.466993
7	Aliabad, Hyderabad	17.346330	78.472550
8	Alijah Kotla	17.360690	78.479960

- Venue data, particularly data related to cafeteria, restaurants. We will use this data to perform clustering on the neighbourhoods.

Sources of data and methods to extract them

- Wikipedia page for neighbourhoods
[https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Hyderabad, India](https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Hyderabad,_India)
- Geocoder package for latitude and longitude coordinates.
- Foursquare API for Venue Data.

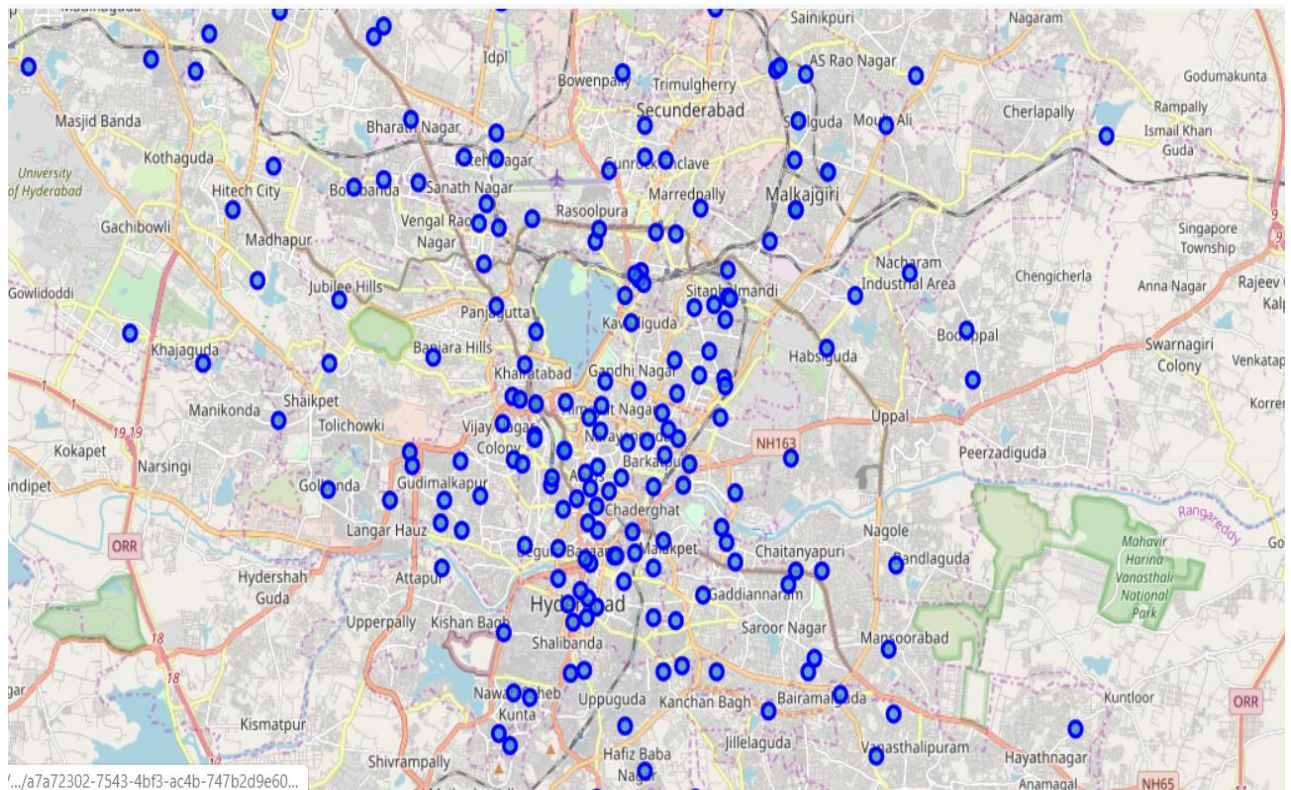
Methodology

- Web scraping Wikipedia page for neighbourhoods list.
- Get latitude and longitude coordinates using Geocoder.
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of the occurrence of each venue category.
- Filter venue category by Restaurant.
- Perform clustering on the data by using K-means clustering
- Visualise the map using Folium.

Results

Categorized the neighbours into 3 clusters

- **Cluster-1(Greater > 9)**- Neighbourhoods with high concentration of restaurant chains.
- **Cluster-2 (Less < 9 , greater >5)**- Neighbourhood with moderate number of restaurant chains.
- **Cluster-3(Less than<=5)**-Neighbourhood with less number of restaurant chains.



Discussion

- Most of the restaurants are concentrated in the central area of the city.
- Highest in cluster 1 and lowest in cluster 3
- Cluster 3 has low number of restaurants, hence may be a good place to open a restaurant.

Recommendations

- Open new restaurants in neighbourhoods in cluster 3 with little or no competition.
- Can open in neighbourhood 2 area with moderate competition if have unique cuisines and taste of food to stand out from the competition.
- Definitely avoid neighbourhood in cluster 1 as high number of restaurants already opened and very high competition.

Conclusion

- **Answer to business question:** The neighbourhoods in cluster 2 & 3 are the most preferred locations to open a new restaurant
- **Findings of this project** will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new restaurant.

