## Applied Data Science Capstone by IBM/Coursera

**Table of contents**

* [Introduction: Business Problem](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#introduction)
* [Data](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#data)
* [Methodology](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#methodology)
* [Analysis](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#analysis)
* [Results and Discussion](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#results)
* [Conclusion](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#conclusion)

## 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={}&query={}&radius={}&limit={}'.format(CLIENT\_ID, CLIENT\_SECRET, latitude, longitude, VERSION, search\_query, radius, LIMIT)

results = requests.get(url).json()

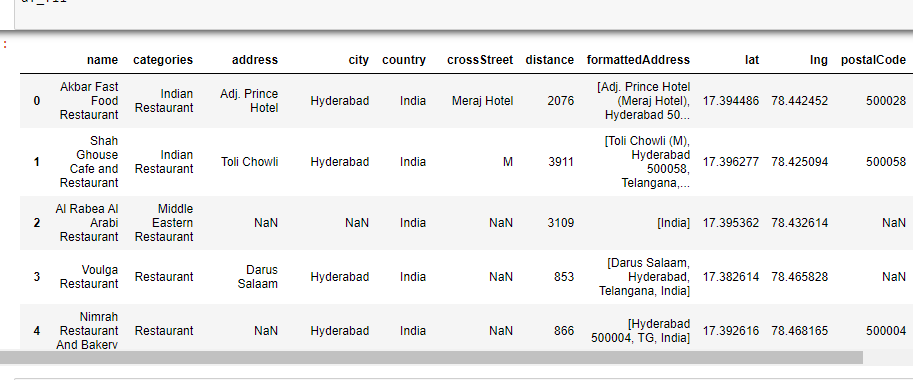
#print(results)

venues=results['response']['venues']

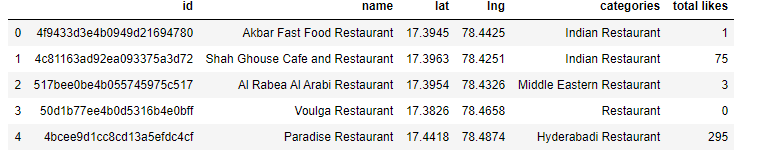
df = json\_normalize(venues)

df.head()

### Perform Data wrangling and make it to required format



### Fetch likes for each restaurant using foursquare API



## [Analysis](http://localhost:8888/notebooks/Desktop/Learn/DS_Capstone/Coursera_Capstone/Cstone_Final.ipynb#analysis)

**Group the restaurants depending on type**

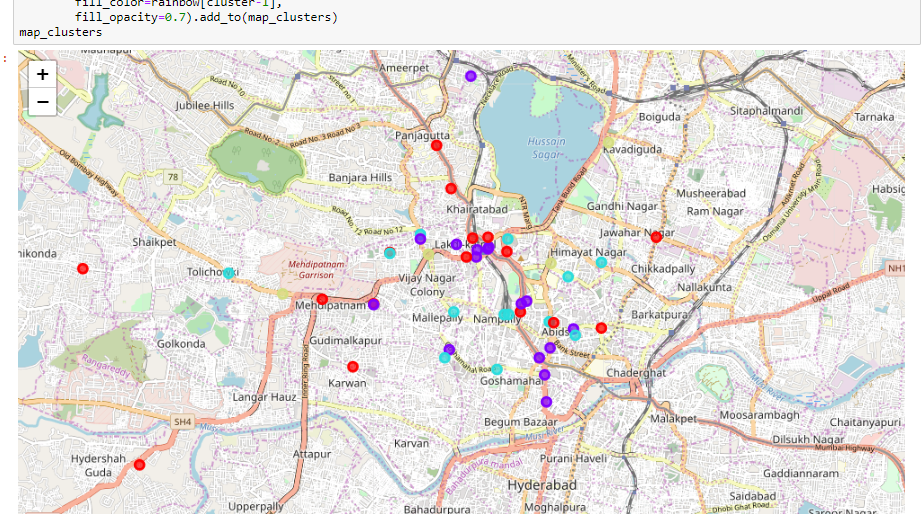


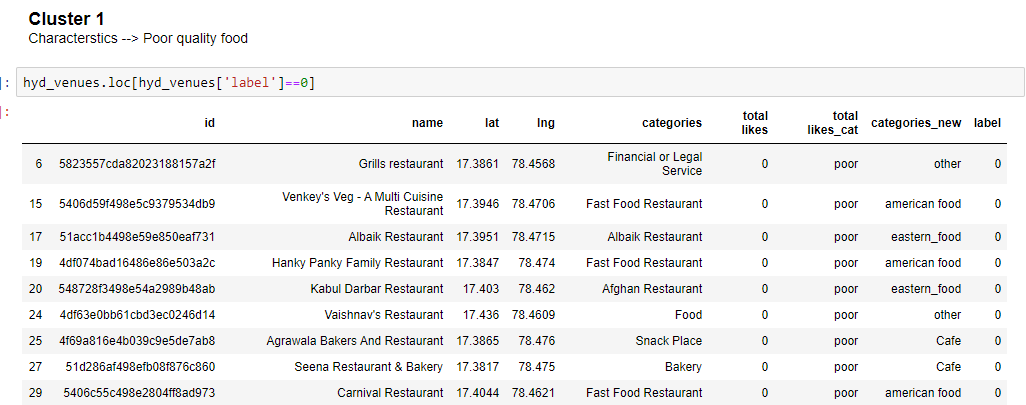
**Rate the restaurant food quality as poor, average and good depending on total number of likes**



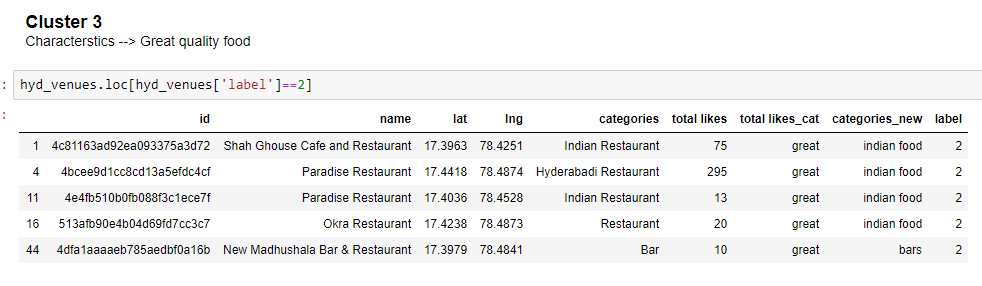
## Results and Discussion

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









## 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.