

COURSERA CAPSTONE

IBM APPLIED DATA SCIENCE

CAPSTONE

HYDERABAD CITY : LOCALITY RECOMMENDATION SYSTEM

ARJUN GIDWANI

JUNE 2020

PROBLEM DESCRIPTION

- The objective of this project is to analyse and select the best possible localities in the city of Hyderabad for a contract working executive to live in/relocate to.
- A challenge faced by Sindhi working executives relocating to Hyderabad for 1-2 year projects from Mumbai, Ahmedabad and Pune living with family on rent in Hyderabad, is finding the right locality to live in based on the population density, the rental rates and the availability of specific conveniences in the locality. The committee members of the Sindhi Community Centre of Hyderabad approached me to find a solution for this problem.
- **BUSINESS QUESTION:** If a Sindhi working executive is planning to relocate to Hyderabad, based on his specific requirements , in which locality would you recommend he shift to?

DATA

DATA REQUIRED:

- - A list of the Zones, Circles and Wards which come under the Greater Hyderabad Municipal Corporation.
- The Coordinates (The Latitudes and Longitudes) of each ward (equivalent of a neighbourhood). This is required particularly to plot the map and to get the venue data using the Foursquare API.
- Venue Data
- Ward-wise average rent data for 2BHK
- Population Data for each Ward
- These are important for clustering and setting the parameters of the recommender model.

DATA

SOURCES OF DATA:

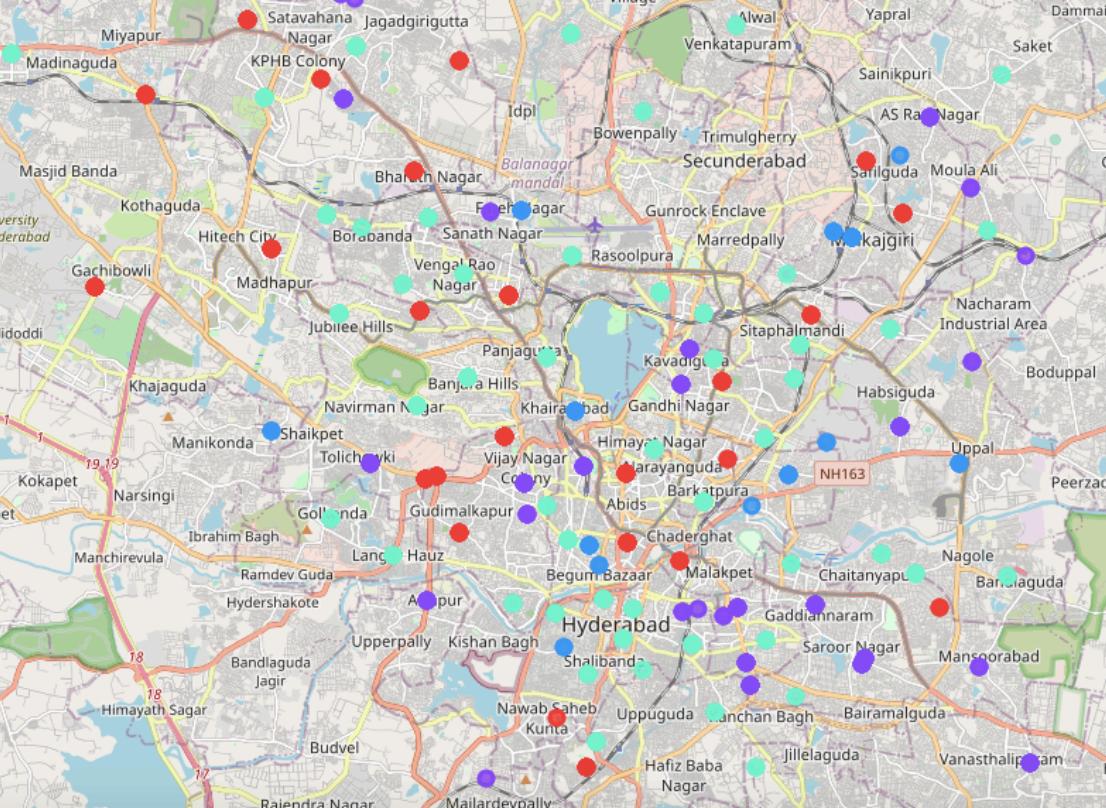
- https://www.ghmc.gov.in/Documents/GHMC_Circles.pdf
- <https://www.ghmc.gov.in/Documents/Wards.pdf>
- <https://www.makaan.com/price-trends/property-rates-for-rent-in-hyderabad>
- <https://indikosh.com/city/708723/greater-hyderabad>

METHODOLOGY

- Create the main data sheet using the data collected and read it into the program.
- Identify and Explore the Venues using Foursquare API
- Perform Clustering using K-Means Clustering
- Create the Recommender System

RESULTS

- Hyderabad City Map with the four clusters



RESULTS

- The Result List produced gives us the names of ideal localities with similarity to Ameerpet (in house rental rates, population and in the number of vegetarian restaurants) and the top 3 venues in those localities. The ranking is also given in order to assist the user in making a better informed choice.

[84] :					
	Ward	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Ranking
0	Ahmed Nagar	Venue Category_Indian Restaurant	Venue Category_Pizza Place	Venue Category_Hotel	[0.23475159972514792]
1	Ameerpet	Venue Category_Indian Restaurant	Venue Category_Fast Food Restaurant	Venue Category_Women's Store	[0.5303522750453613]
2	Azampura	Venue Category_Hotel	Venue Category_Tourist Information Center	Venue Category_Indian Restaurant	[0.41677435957312037]

DISCUSSION

- An alternate cluster number (k number) could significantly alter the results. Based on a high or low k number the results could be either overfitted or underfitted. It, therefore, is very important to conduct an analysis of the number of clusters prior to clustering. In this program we have used the elbow method to arrive at k (the number of clusters).
- To be able to factor in vegetarian restaurant options along with giving greater priority to the rental rates and the population density, we need to create a Non-vegetarian restaurant category which is equal to 1- the total number of Vegetarian restaurants for each ward. In the final algorithmic equation, we give this parameter the least weightage so as to give a higher ranking to wards with Vegetarian Restaurants.

CONCLUSION

- The Hyderabad Locality Recommendation system relies on the important factors of Avg.Rent, Population ,the special parameter (which in this case is the Vegetarian Restaurant requirement) and the Foursquare API to analyse the venues in each Ward/Locality.

THANK YOU
