

Analysis of Libraries/Bookstores in Hyderabad



Introduction

Hyderabad, the capital of southern India's Telangana state is a major center for the technology industry, it's home to many upscale restaurants and shops. There are a lot of people here who love to read.

Book lovers in India are always on the lookout for a nearby library where they could go, relax and read. Libraries provide us with a plethora of opportunities to explore and tickle our creativity. They should be located at places where they can be easily accessible to the public as well as provide a conducive environment for reading. There are a few public libraries in Hyderabad(India), but as a business decision for opening a public library needs a lot of consideration. The location specially needs to be considered so that it is accessible not only to the school and college going students but also to their parents and the elderly.

Business Problem

The business problem that this capstone project will be looking at is to find an appropriate location for a book lover/investor to open a library/book store for the public. Using Data Science methodologies this project aims to provide an answer to the question : What is the recommended place to open a public library/book store in Hyderabad?

Target Audience

This project is useful to investors/book lovers who are interested in opening a public library/book store in the city of Hyderabad, India.

Data

In order to get this problem solved we will need the following data:

- The list of neighbourhoods of Hyderabad. This can be obtained from scraping the wikipedia page.
- The latitudes and longitudes of those neighbourhoods. This is required to plot the map. This can be done using the geocoder.
- To perform the clustering of neighbourhoods we will also need the data related to libraries. In order to obtain this we will need to use the Foursquare API([Foursquare Venues & Places Database](#)) and retrieve the library/book store category from it.

Methodology

The wikipedia page used for the data wrangling was

https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Hyderabad,_India

Beautifulsoup was used to parse the data and the neighbourhood places were put in a dataframe.

	Neighborhood
0	A. S. Rao Nagar
1	A.C. Guards
2	Abhyudaya Nagar
3	Abids
4	Adikmet

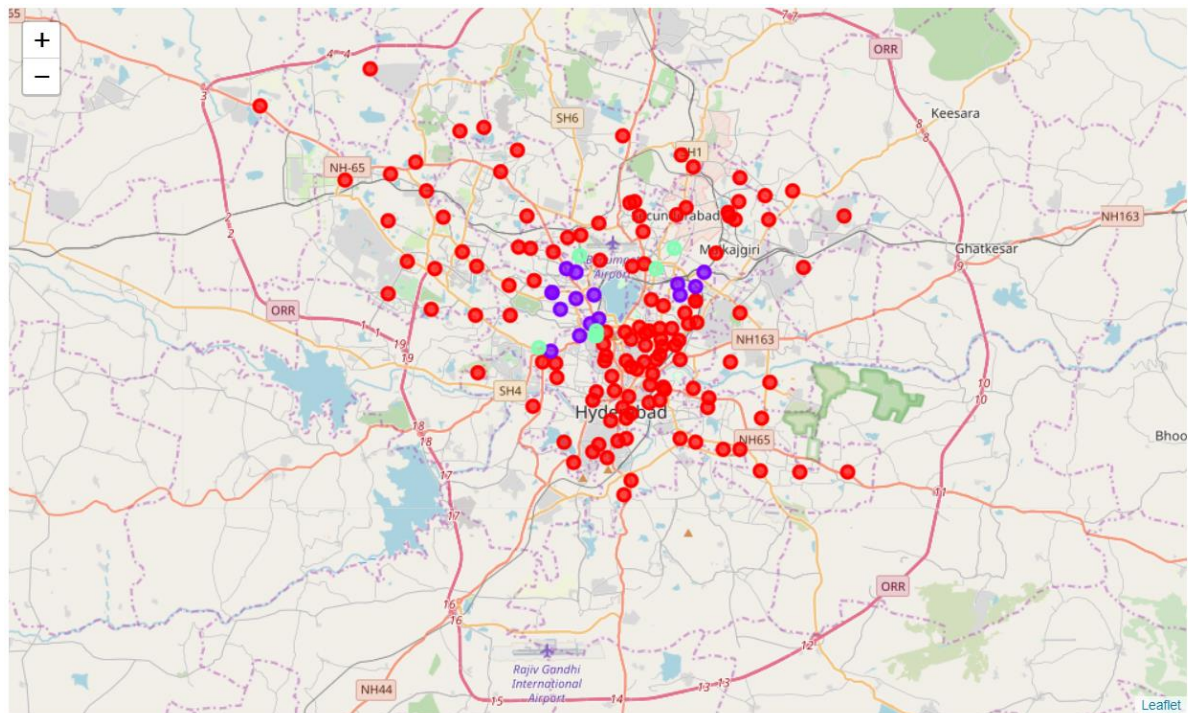
Using geocode the latitude and longitude of each were retrieved.

	Neighborhood	Latitude	Longitude
0	A. S. Rao Nagar	17.479950	78.556834
1	A.C. Guards	17.402804	78.459487
2	Abhyudaya Nagar	18.990477	72.844057
3	Abids	17.389478	78.477182
4	Adikmet	17.409550	78.513094
5	Afzal Gunj	17.389178	78.465273
6	Aghapura	17.345630	78.472680
7	Aliabad, Hyderabad	17.360545	78.480102
8	Alijah Kotla	17.504362	78.414985
9	Allwyn Colony	17.502229	78.508858
10	Alwal	17.777683	78.570507
11	Amberpet	17.437501	78.448251

Using Foursquare api the venue names and categories were obtained

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	A. S. Rao Nagar	17.47995	78.556834	The Coffee Cup	17.483180	78.552104	Café
1	A. S. Rao Nagar	17.47995	78.556834	Cafe Coffee Day	17.481262	78.555077	Café
2	A. S. Rao Nagar	17.47995	78.556834	Domino's Pizza	17.475035	78.553141	Pizza Place
3	A. S. Rao Nagar	17.47995	78.556834	Fifth Avenue Bakers	17.487673	78.542793	Bakery
4	A. S. Rao Nagar	17.47995	78.556834	Woodland Restaurant	17.476646	78.566404	Snack Place

Used KMeans clustering created 3 clusters. Using Folium created a map of all the clusters



Cluster 0

	Neighborhood	Bookstore	Cluster Labels	Latitude	Longitude
0	A. S. Rao Nagar	0.000000	0	17.479950	78.556834
105	Malkajgiri	0.000000	0	17.344964	78.441656
106	Malkajgiri mandal	0.000000	0	17.498161	78.356763
107	Mallapur	0.000000	0	17.357528	78.477084
108	Mallepally	0.000000	0	17.468531	78.420670
109	Manikonda	0.000000	0	17.374353	78.516083
110	Manovikas Nagar	0.000000	0	17.450807	78.422242
111	Marredpally	0.000000	0	17.466816	78.559216
112	Masab Tank	0.000000	0	17.419142	78.498573
113	Mehboob ki Mehendi, Hyderabad	0.000000	0	17.422291	78.516970
114	Mehdipatnam	0.000000	0	17.823182	78.536153

Cluster 1

	Neighborhood	Bookstore	Cluster Labels	Latitude	Longitude
156	Radha Krishna Nagar	0.030000	1	17.439446	78.442856
10	Amberpet	0.030000	1	17.437501	78.448251
38	Chatta Bazaar	0.030000	1	17.409678	78.457052
21	Balkampet	0.030000	1	17.417746	78.439901
80	Kavadiguda	0.030000	1	17.426352	78.434940
166	Safilguda	0.037037	1	17.429733	78.517426
89	Krishna Nagar, Hyderabad	0.047619	1	17.437722	78.521976
103	Maharajgunj	0.030000	1	17.402962	78.450754
104	Mahatma Gandhi Road (Secunderabad)	0.024390	1	17.394263	78.434251
76	Kamala Nagar, Hyderabad	0.030000	1	17.412426	78.461823
167	Saidabad	0.030000	1	17.425505	78.458520
129	Nagole	0.045455	1	17.425601	78.508515
137	Nawab Saheb Kunta	0.030000	1	17.423647	78.448406
143	Padmarao Nagar	0.030000	1	17.426352	78.434940
37	Chandrayan Gutta	0.031250	1	17.431532	78.507162

Cluster 2

	Neighborhood	Bookstore	Cluster Labels	Latitude	Longitude
20	Bairamalguda	0.016129	2	17.446923	78.450451
133	Nampally, Hyderabad	0.013514	2	17.439702	78.494785
26	Basheerbagh	0.010000	2	17.402787	78.460449
102	Madinaguda	0.019608	2	17.450794	78.504897
86	Kondapur, Ranga Reddy district	0.010000	2	17.405368	78.459949
141	Noorkhan Bazar	0.010000	2	12.981796	77.594961
1	A.C. Guards	0.010000	2	17.402804	78.459487
121	Moosapet	0.017544	2	17.396251	78.427061

Results/Observations/Conclusion

Most of the bookstores are in the cluster 1 and the least in cluster 0 and moderate in cluster 2. Hence this shows that a person keen on opening a library/bookstore would benefit by opening it somewhere in the cluster 0, preferably near Hitech city. Hitech city is the IT hub of Hyderabad, with many schools as well in this area. So having a library/bookstore in this area would benefit everyone. It would be a win win deal for both the customers as well as the business person. With moderate competition and unique selling ideas even cluster 2 is good to open a bookstore/library. It would be ideal to avoid cluster 1 for opening a bookstore/library as there seems to be many in this cluster and the business might have to suffer because of major competition.