

Best Neighborhood to Live !!!

Manoj Raghorta

Introduction

It is imperative for every individual to use one's time, energy and money wisely.

we propose an application which has the capability to minimize every individual's effort while planning to live in a city or planning to stay on a holiday at particular location.

every individual would want to stay at a place from where one could have easy access to their location of interests.

We propose a framework which takes input from user about the places they want to explore, and find the best neighbourhood which has shortest access to most number of interested venues.

DataSets

Mumbai Neighborhood Wikipedia

https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

Apart from Wikipedia data, Geocoder API and Foursquare API are used to get the latitude and longitude data sets.

Web Extraction

I have extracted the neighborhoods of Mumbai Metropolitan city. We can generally do this task by using 'Beautifulsoup' API and giving the wikipedia page containing interested city's neighborhood information.

The extracted information may be encoded in xml or pure HTML format which will need clean up to get the names of the neighborhoods.

Unwanted information is purified only to access the neighbourhood names and then converted into a dataframe

Get Neighborhood Locations

The data frame of neighborhoods we formed needs to be appended with their respective latitude and longitude information.

We used the 'Geolocator API' to get these latitudes and longitudes.

we can also use Foursquare API yet using Geolocator is convenient as we only want location information and nothing else.

The resultant data frame consists of neighborhood names and their respective latitude and longitude values.

Venue Search Request Using Foursquare

Based on the user's preference, one provides the query string to search for in venues.

Here we use Foursquare API request by providing query string, client credentials, search radius and center point to begin search from.

As I am working on Mumbai Neighborhoods, I have set a place named 'Kurla' as center and provided it latitude and longitude to begin mapping the radius and begin our search.

Results we get from query request is in json format which needs to be parsed to get the relevant information. We extract names of venues, its latitudes and longitudes from the json data.

Maps Visualization

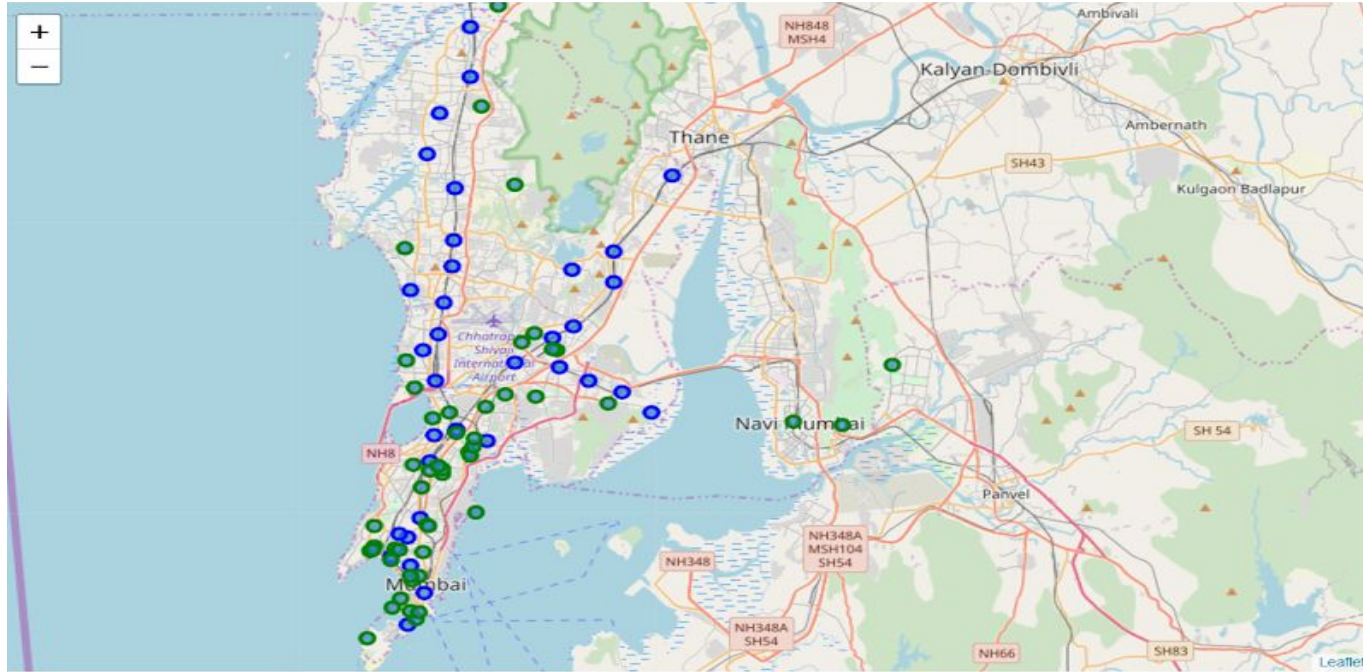
The user would definitely want the application to be more of an intuitive and graphical rather than just statistics oriented.

First, we create a map of all neighborhoods in the city using Folium API which uses the neighborhoods latitude and longitude.

Then we create another map of venues we generated using user's search query.

Then we superimpose the two maps together which results in getting a pictorial view of how neighborhoods and venues generated are situated on the citi's map.

Superimposed Map of Results



Blue Neighborhoods

Green Venues

Statistical Analysis

Though we provide the user a pictorial presentation of how neighborhoods and venues one requested are situated, there is not a statistical proof in numbers which extends the proof we discovered through maps.

We calculate the distance from each neighborhood to each venue generated and do an analysis on how many venues are situated in each neighborhood's near proximity area and what is their distance.

We use the 'Haversine Formula' to calculate the distance between each venue to each neighborhood.

After further analysis, we calculate how many venues fall in the near proximity of each neighborhood.

The near proximity or limit is assumed as 2.5 kilometers in this project which can be variable

Statistical Analysis Results

Andheri	0
Bandra	4
Borivali	1
Dahisar	1
Goregaon	0
Jogeshwari	0
Juhu	0
Kandivali west	1
Khar	2
Malad	0
Santacruz	1
Vile Parle	0
Ghatkopar	3
Kanjurmarg	0
Kurla	6
Mulund	0
Powai	0
Vidyavihar	4
Vikhroli	0

Chembur	3
Govandi	1
Mankhurd	1
Trombay	1
Antop Hill	6
Byculla	7
Colaba	6
<u>Dadar</u>	<u>12</u>
Fort	8
<u>Girgaon</u>	<u>11</u>
<u>Kalbadevi</u>	<u>11</u>
<u>Kamathipura</u>	<u>11</u>
Matunga	8
Parel	10
Tardeo	10

In above results, you can see neighbourhood named 'Dadar ' has the highest number of research places so the most suitable place to live if you are a research person will be Dadar. Yet the neighbourhoods Girgaon, Kalbadevi, Kamathipura are also very near the research places so you can also live over there too

Conclusion

We extracted the neighborhoods of Mumbai city and generated their latitudes and longitudes using Geocoder API.

Then we extracted the search query results given by the user for one's interest of venues using Foursquare API and purified the data and converted the data in pandas data frames.

Using Folium API we create separate maps of neighborhoods and venues and then both maps are superimposed on each other.

Finally we perform statistical analysis to build a distance matrix which provides how many venues fall in the near proximity of every neighborhood.

Based on our result, user can identify which neighborhood is best suitable to live and most accessible to one's suggested interested venues.