Date: 24th Jan,2021

Analysing & Visualizing rent houses data in Mumbai

Vishal Sudhir Sawant

1.1 Introduction

Mumbai is densely populated city on India's west coast. A financial center. It's India's largest city with over 603.4km area and a population of 1.84 crores. It serves as an economic hub of India, contributing 10% of factory employment, 25% of industrial output.

As I mentioned earlier that Mumbai serves as an economical hub of India, lots of peoples move to Mumbai for doing a job, finding a job, education, labour work, etc. as we know Mumbai is a densely populated city so, it's hard to find the house on rent(in the budget of person) in this type of city.

So, to solve this problem I will be going to analyse all the rents house data of Mumbai and find some reliable house for users.

1.2 Description of the data

After searching a couple of data provider sites, I have found perfect data set on **kaggle.com**, I'm going to use **"Flats for Rent in Mumbai"** data set from Kaggle. This data set contains data from 12/08/2019 to 14/01/2020 total of 34349 rows.

Followings are a list of columns with a small description of each:

area	Floor area of the property
bathroom_num	Number of bathrooms available
bedroom_num	Number of bedrooms available
city	City in which property is located
desc	Text description of the property
dev_name	Name of property developer
floor_count	Total number of floors in building
floor_num	Floor on which property is located
furnishing	Furnishing status
id	Unique ID
latitude	Latitude of location
id_string	Unique ID string used to scrape a particular HTML page element.
locality	Locality in which property is located
longitude	Longitude of location.
post_date	Date on which property was listed on website.
poster_name	Name of poster
price	Price of the property
project	Name of the residential complex
title	Title of the property ad on the website
trans	Type of property transaction

type	Type of residential complex
URL	URL of the individual property
user_type	Type of user who posted the ad

This data set contains 23 columns but, we will filter the data set and remove the un-relevant column and keep only those column which will useful to us. we will also be going to apply data cleaning on the filtered dataset.

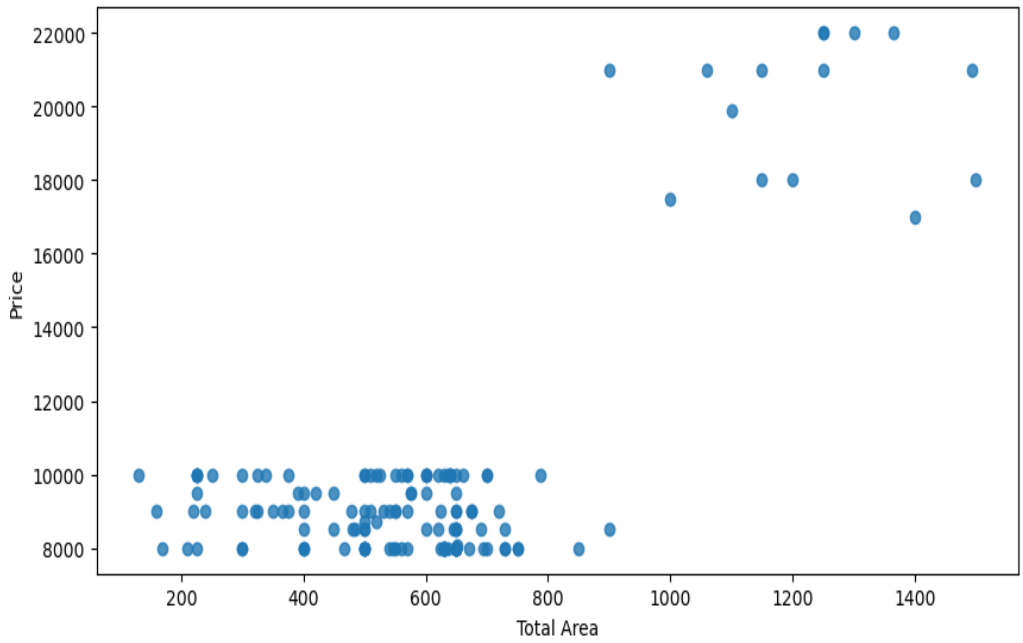
area	desc	latitude	locality	longitude	price	title
225	1 Bath,Unfurnished,1 floor,South facing Single...	19.233532	Borivali West	72.839195	11500	1 BHK Apartment for Rent in Borivali West
225	1 Bath,Unfurnished,Upper Basement floor 1 BHK,...	19.009829	Antop Hill	72.825241	16000	1 BHK Apartment for Rent in Antop Hill
225	1 Bath,Unfurnished,1 floor,West facing hi i ha...	72.874649	Azad Nagar	19.174179	18000	1 BHK Apartment for Rent in Azad Nagar
225	1 Bath,Semi-Furnished,4 floor,North - East fac...	19.134399	Andheri East	72.868393	10000	1 BHK Apartment for Rent in Andheri East
225	1 Bath,Semi-Furnished,1 floor One Room Kitchen...	19.115490	Andheri East	72.872696	15000	1 BHK Apartment for Rent in Andheri East

1.3 Methodology

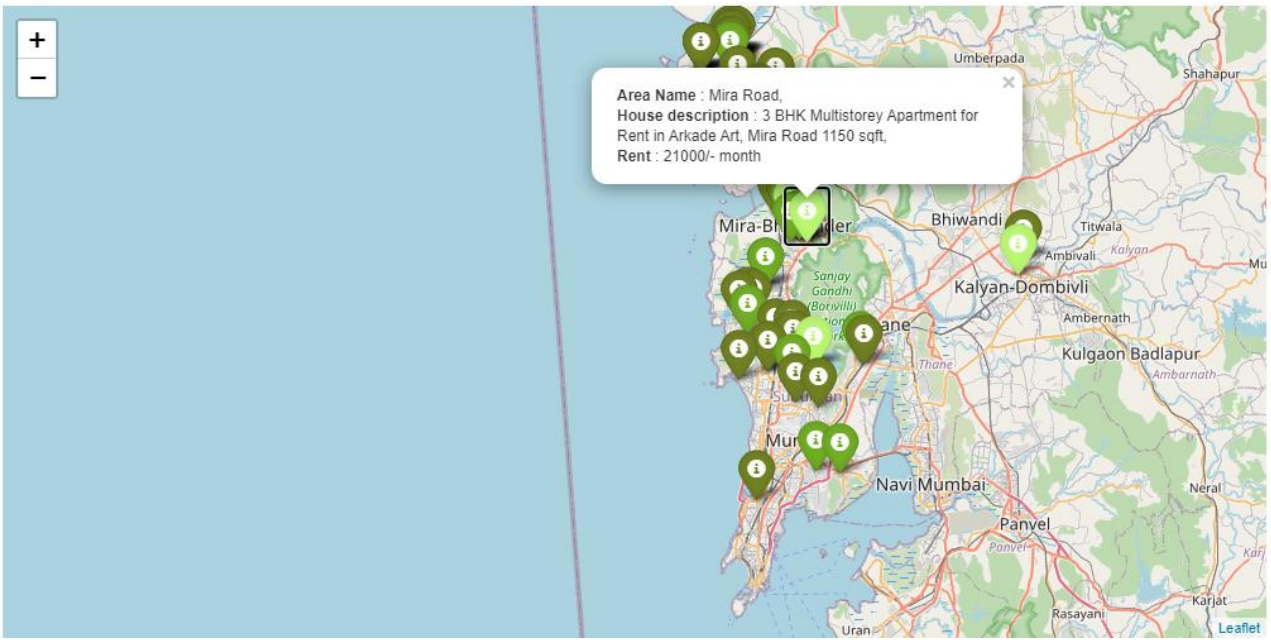
I used a folium library to visualize all the rent house data of Mumbai.

This shows relationship between total area of house and price of house.

```
: #lets plot the scatter plot to see the relation between Total area of house and price of house

figure(num=None, figsize=(10,5),dpi=100)

x = df.area
y = df.price


plt.scatter(x, y, alpha=0.8,cmap='viridis')
plt.xlabel('Total Area')
plt.ylabel('Price')
plt.show()
#dark green #097502
#light green #5CDD53
#light red #097502
#Dark light Red  #C08D8D
#Dark Red #C05252
```



And when you click on any marker you will able to see all primary information about that rent house.



The purpose behind the classification is to easily identify the house with low rent price to the highest rent price. I used foursquare APIs to explore the restaurants near all rent house locations and clusters them accordingly. Foursquare return all Venus near each house and then put all venues inside the data frame.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Naigaon East	19.360215	72.850167	Vaman Dhaba	19.356340	72.856704	Indian Restaurant
1	Vasai West	19.370682	72.812934	Western Inn	19.363244	72.813850	Restaurant
2	Virar	19.461619	72.800999	Domino's Pizza	19.457100	72.804253	Pizza Place
3	Virar	19.467032	72.801612	McDonald's	19.474018	72.804842	Fast Food Restaurant
4	Naigaon East	19.360215	72.850167	Vaman Dhaba	19.356340	72.856704	Indian Restaurant

I have also analysed the number of venues found by foursquare API for each area (locality)

I have also analysed the number of venues found by foursquare API for each area (locality)

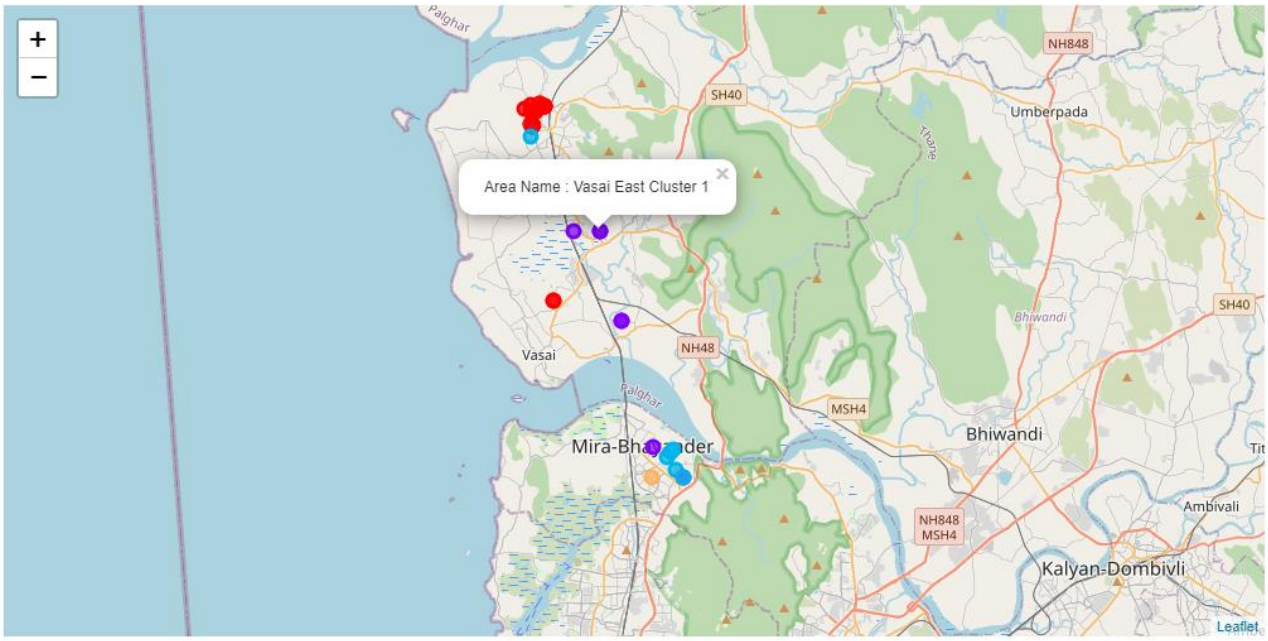
	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Bolinj	1	1	1	1	1	1
Goregaon East	1	1	1	1	1	1
Mira Road	4	4	4	4	4	4
Naigaon East	2	2	2	2	2	2
Poonam Gardens	1	1	1	1	1	1
Ramdev Park	1	1	1	1	1	1
Tirupati Nagar	1	1	1	1	1	1
Unique Garden	1	1	1	1	1	1
Vasai East	3	3	3	3	3	3
Vasai West	2	2	2	2	2	2
Vinay Nagar	1	1	1	1	1	1
Virar	11	11	11	11	11	11
Virar West	7	7	7	7	7	7

1.4 Result

As a result, I group Neighborhood('locality') by the number of top 5 common venues.
Then merged it with the dataframe.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Bolinj	Pizza Place	Restaurant	Mexican Restaurant	Indian Restaurant	Fast Food Restaurant
1	Goregaon East	Donut Shop	Restaurant	Pizza Place	Mexican Restaurant	Indian Restaurant
2	Mira Road	Pizza Place	Indian Restaurant	Restaurant	Mexican Restaurant	Fast Food Restaurant
3	Naigaon East	Indian Restaurant	Restaurant	Pizza Place	Mexican Restaurant	Fast Food Restaurant
4	Poonam Gardens	Mexican Restaurant	Restaurant	Pizza Place	Indian Restaurant	Fast Food Restaurant

After merging the venue data to the main data frame you can see the clusters map of restaurants near rent hour's locations Mumbai.



1.5 Discussion

know Mumbai is a densely populated city so, it's hard to find the house on rent(in the budget of person) in this type of city. So, to solve this problem I did the analysis of all the rent houses of Mumbai and find some reliable house by classifying them and also did exploring each nearby venue location.

1.6 Conclusion

By classifying, visualizing, and explore each nearby venue location, it is easier to make decisions for stakeholders like real-estate builders and brokers.