

CAPSTONE PROJECT
ON
HOUSING SALES PRICES &
VENUE DATA ANALYSIS OF
NEIGHBORHOODS OF PUNE CITY, INDIA.



SUBMITTED BY
SWAPNIL ASHOK UDGIR

INDEX

Sr. No.	Particulars	Page No.
1	Introduction	3
2	Data Description	4
3	Methodology	5
4	Results	6
5	Discussion	12
6	Conclusion	12
7	References	13

HOUSING SALES PRICES & VENUE DATA ANALYSIS OF NEIGHBORHOODS OF PUNE CITY, INDIA.

1. INTRODUCTION

Pune is a bustling metropolis in the Indian State of Maharashtra and the eighth most populous city in India, with an estimated population of about 7.2 million as of 2019. It has been ranked as one of the most liveable cities in India several times [1].

Along with its extended city limits Pimpri Chinchwad and the three cantonment towns of Pune, Khadki and Dehu Road, Pune forms the urban core of the eponymous Pune Metropolitan Region (PMR). The Pune Metropolitan Region (PMR) has grown to 7,256 km² made up of the ten talukas of the Pune district. The twin cities of Pune and Pimpri-Chinchwad along with the three cantonment areas of Pune, Khadki and Dehu Road form the urban core of the PMR, which also includes seven municipal councils and 842 villages [1].

Situated 560 metres (1,837 feet) above sea level on the Deccan plateau on the right bank of the Mutha river, Pune is also the administrative headquarters of its namesake district [1].

The city is considered to be the IT capital of India. It is also known as the "Oxford of the East" due to the presence of several well-known educational institutions. The city has emerged as a major educational hub in recent decades, with nearly half of the total international students in the country studying in Pune. Research institutes of information technology, education, management and training attract students and professionals from India and overseas [1].

As a resident of this city, I decided to use Pune city in my project. The city is divided into 41 boroughs and 100 neighborhoods in total. However, the fact that the boroughs are squeezed into an area of approximately 331.26 square kilometres causes the city to have a very intertwined and mixed structure. The scope of this project is limited to Pune city only [1].

Rapid industrialisation since the 1960s has led to large influx of people to the city. Housing supply has not kept pace with demand, causing the number of slum dwellings to increase [1].

Pune is a city with a high population and population density. Being such a crowded city leads the owners of shops and social sharing places in the city where the population is dense. When we think of it by the investor, we expect from them to prefer the boroughs and neighborhoods where there is a lower real estate cost and the type of business they want to install is less intense. If we think of the city residents, they may want to choose the regions where real estate values are lower, too. At the same time, they may want to choose the boroughs and neighborhoods according to the social places density. However, it is difficult to obtain information that will guide investors in this direction, nowadays.

When we consider all these problems, we can create a map and information chart where the real estate index is placed on Pune city and each neighborhood is clustered according to the venue density.

2. DATA DESCRIPTION

To consider the problem the data can be listed as below:

I found the list of 41 Boroughs and 100 Neighborhoods of Pune city from the Pune Municipal Corporation (PMC) official website. I cleaned the data and reduced it to 84 Neighborhoods by removing the Neighborhoods whose areas are partially included in the Pune city [2]. I could not find the geojson file of administrative divisions' data of Pune city from any of the available repositories. So, I used Google Map, 'Search Nearby' option to get the centre coordinates of the each Borough and Neighborhood. I prepared a csv file for each Borough and Neighborhood along with their location coordinates such as Latitude and Longitude [5].

I used Foursquare API to get the most common venues of given Borough and Neighborhood of Pune city [3].

There is not many public data related to demographic and social parameters for the Pune city. Therefore I had to set-up my own data tables in most cases. In this case, I collected latest per square feet Housing Sales Price (HSP) Averages for each Borough and Neighborhood of Pune city from housing retail website [4].

I cleaned the data and reduced it to Pune city where I used it to create bar map of Housing Sales Price Index of Pune city.

3. METHODOLOGY

I prepared the master data in csv format and stored it in my Git hub repository. The main components of the data were Boroughs, Neighborhoods, Latitude, Longitude and Average Housing Prices of Pune city.

I got the names of the 41 Boroughs and 100 Neighborhoods from Pune Municipal Corporation (PMC) official website. I reduced the scope of the study to 84 Neighborhoods by removing the Neighborhoods whose areas are partially included in the Pune city [2].

I could not find the geojson file of administrative divisions' data of Pune city from any of the available repositories. So, I used Google Map, 'Search Nearby' option to get the centre coordinates of the each Borough and Neighborhood of Pune city [5]. I prepared a csv file for each Borough and Neighborhood of Pune city along with their location coordinates such as Latitude and Longitude.

	Borough	Neighborhood	Latitude	Longitude
0	Mundhwa - Magarpatta City	Amanora Park Town	18.521754	73.940587
1	Dhankawadi - Ambegaon Pathar	Ambegaon Pathar	18.458489	73.842559
2	Vadgaon Dhayari - Suncity	Anand Nagar	18.474473	73.822394
3	Pune Vidyapeeth - Wakdewadi	Ashok Nagar	18.550072	73.839188
4	Kharadi- Chandannagar	Ashoka Nagar	18.559982	73.942650
5	Aundh - Bopodi	Aundh	18.559488	73.803733
6	Mundhwa - Magarpatta City	B.T Kawade Road	18.518254	73.906662
7	Baner - Balewadi - Pashan	Balewadi	18.578653	73.770933
8	Baner - Balewadi - Pashan	Balewadi Gaon	18.576415	73.778786
9	Baner - Balewadi - Pashan	Balewadi Phata	18.569090	73.783903

Fig. 3.1

I used python folium library to visualize geographic details of Pune city, its boroughs and Neighborhoods. I created a map of Pune city with Neighborhoods superimposed on top. I used latitude and longitude values to get the visual as below:

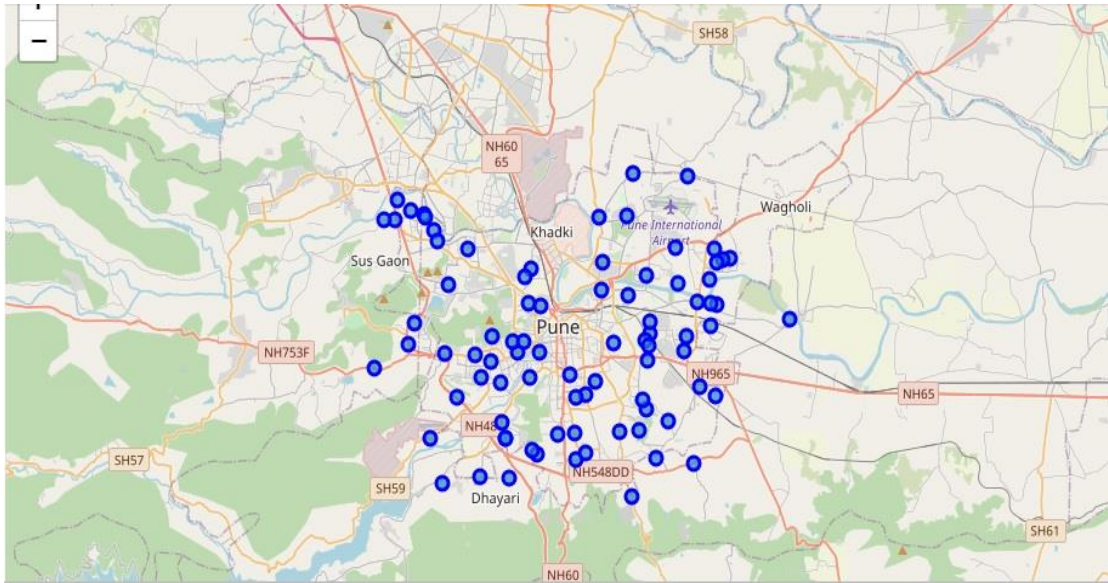


Fig. 3.2

I used the Foursquare API to explore the Boroughs and Neighborhoods of Pune city and segment them [3]. I set a limit of 100 venues per neighborhood with a radius of 1000 metres from the location coordinates of the neighborhood. Given below is the head of Venue name, Venue Category, Venue Latitude, Venue Longitude.

[19]:	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amanora Park Town	18.521754	73.940587	Crazy Noodles	18.518636	73.936016	Chinese Restaurant
1	Amanora Park Town	18.521754	73.940587	Amanora Town Centre	18.518747	73.935204	Shopping Mall
2	Amanora Park Town	18.521754	73.940587	INOX	18.518703	73.935990	Multiplex
3	Amanora Park Town	18.521754	73.940587	Spaghetti Kitchen	18.518721	73.934763	Italian Restaurant
4	Amanora Park Town	18.521754	73.940587	Dario's Espresso	18.518689	73.936042	Italian Restaurant

Fig. 3.3

4. RESULTS

I plotted a bar graph of Neighborhoods v/s Number of Venues returned by the Foursquare API which is shown below. From the bar graph, we can clearly see that Top 5 Neighborhoods having venues greater than 75 are Aundh, Prabhat Road, Model Colony, Kalyani Nagar and Koregaon Park. The Bottom 5 Neighborhoods having venues less than 10 are Shivane, Keshav Nagar, Yewalewadi, Handewadi and Kale Padal.

The result doesn't mean that inquiry run all possible results in neighborhoods. The results depend upon the Latitude and Longitude information and here we just run single Latitude and Longitude information for each neighborhood. We can increase the possibilities with more Latitude and longitude information of neighborhood.

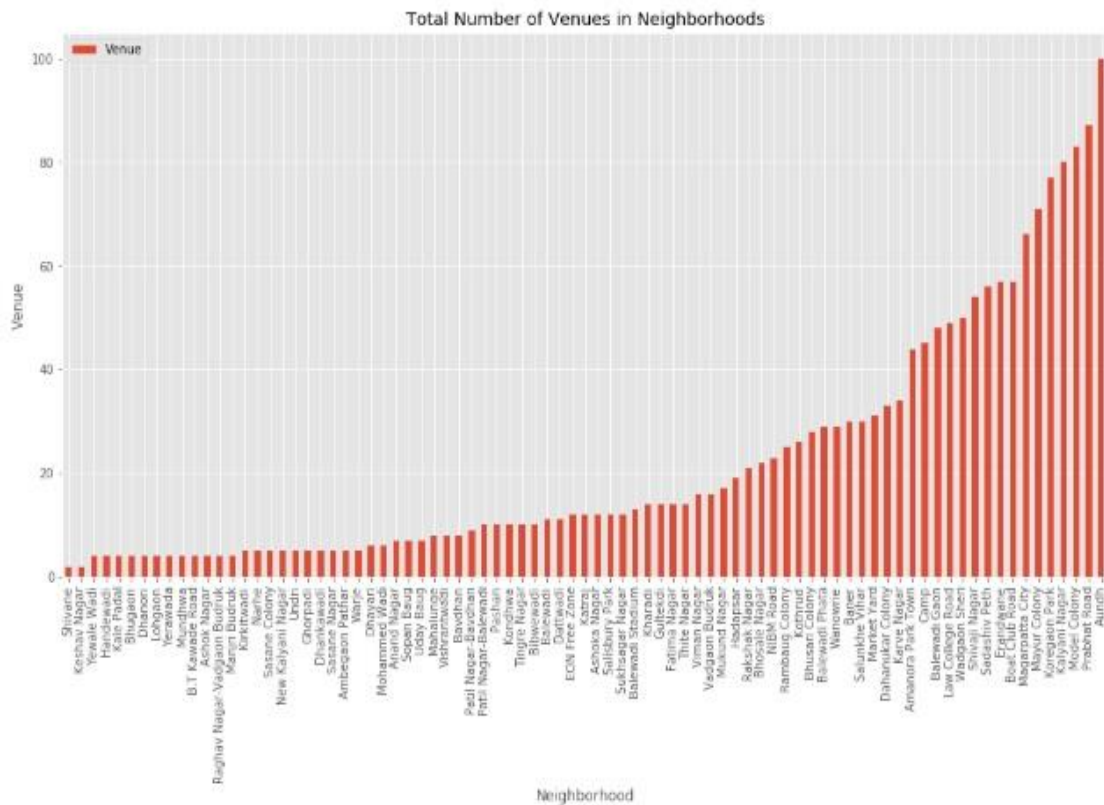


Fig. 4.1

In summary of this graph 175 unique categories were returned by Foursquare, then I created a table which shows list of top 10 venue category for each neighborhood in below table.

[24]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Amanora Park Town	Italian Restaurant	Bar	Donut Shop	Indian Restaurant	Fast Food Restaurant	Chinese Restaurant	Coffee Shop	Department Store	Burger Joint	Bowling Alley
1	Ambegaon Pathar	Café	Convenience Store	Coffee Shop	Auto Workshop	Eastern European Restaurant	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Exhibit
2	Anand Nagar	Indian Restaurant	Pizza Place	Café	Fast Food Restaurant	Snack Place	Ice Cream Shop	Deli / Bodega	Farmers Market	Farm	Falafel Restaurant
3	Ashok Nagar	Playground	Pub	Café	Diner	Zoo	Dumpling Restaurant	Farm	Falafel Restaurant	Exhibit	English Restaurant
4	Ashoka Nagar	Restaurant	Pizza Place	Breakfast Spot	Hotel	Gym	Ice Cream Shop	Multicuisine Indian Restaurant	North Indian Restaurant	Go Kart Track	Tea Room

Fig. 4.2

As we can see we have some common venue categories in boroughs and neighborhoods. I used unsupervised learning K-means algorithm to cluster the boroughs. K-Means algorithm is one of the most common cluster method of unsupervised learning. I ran the K-Means algorithm to cluster the neighborhoods into **5** clusters.

Below is the merged table with cluster labels for each neighborhood.

[27]:

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
	Mundhwa - Magarpatta City	Amanora Park Town	18.521754	73.940587	3	Italian Restaurant	Bar	Donut Shop	Indian Restaurant	Fast Food Restaurant	Chinese Restaurant	Coffee Shop	Department Store	Burger Joint	Bowling Alley
	Dhankawadi - Ambegaon Pathar	Ambegaon Pathar	18.458489	73.842559	4	Café	Convenience Store	Coffee Shop	Auto Workshop	Eastern European Restaurant	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Exhibit
	Vadgaon Dhayari - Suncity	Anand Nagar	18.474473	73.822394	3	Indian Restaurant	Pizza Place	Café	Fast Food Restaurant	Snack Place	Ice Cream Shop	Deli / Bodega	Farmers Market	Farm	Falafel Restaurant
	Pune Vidyapeeth - Wakdevadi	Ashok Nagar	18.550072	73.839188	3	Playground	Pub	Café	Diner	Zoo	Dumpling Restaurant	Farm	Falafel Restaurant	Exhibit	English Restaurant
	Kharadi - Chandannagar	Ashoka Nagar	18.559982	73.942650	3	Restaurant	Pizza Place	Breakfast Spot	Hotel	Gym	Ice Cream Shop	Multicuisine Indian Restaurant	North Indian Restaurant	Go Kart Track	Tea Room

Fig. 4.3

I superimposed the clusters of neighborhoods on the map of Pune city which is shown below.

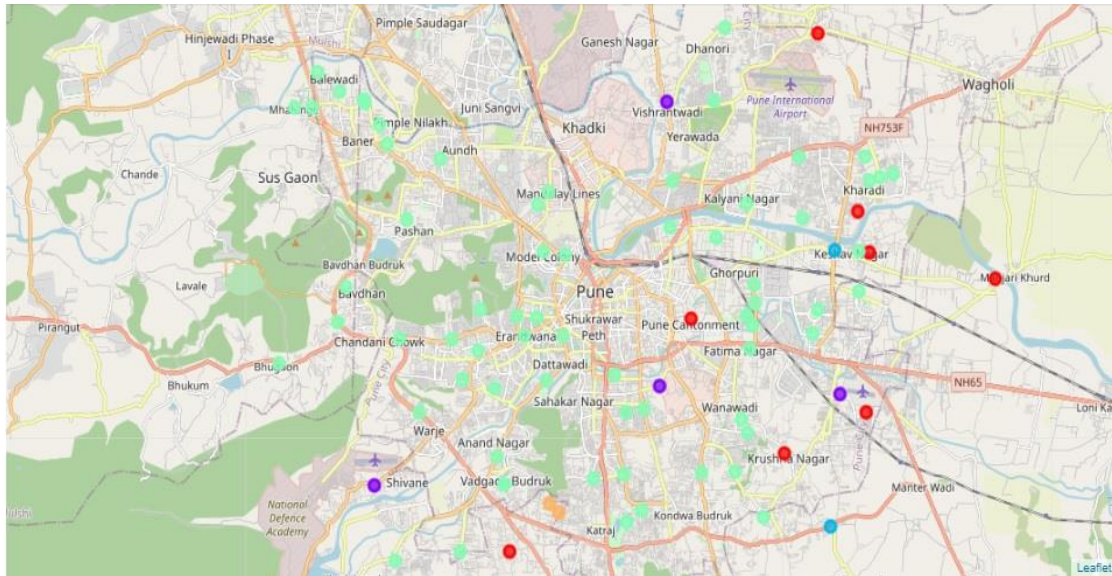
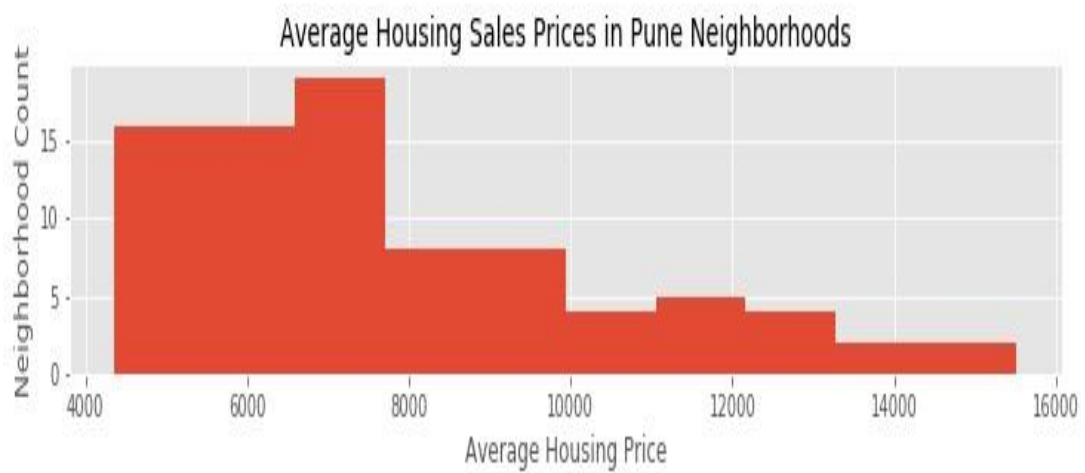


Fig. 4.4

I scraped the Average Housing Prices data of Neighborhoods in Pune city from a popular the Real Estate Website and stored it in master data in csv format. Then, I visualized the frequency of average housing sales prices in neighborhoods of Pune city with the help of histogram:



Rs. / Sq. Ft.

Fig. 4.4

From the above histogram, we can define the ranges as below:

- < 6592 Rs. / Sq.Ft. AHP : “Low Level - 1 HSP”
- 6592 – 7707 Rs. / Sq.Ft. AHP : “Low Level - 2 HSP”
- 7707 – 9937 Rs. / Sq.Ft. AHP : “Mid-1 Level HSP”
- 9937 – 12167 Rs. / Sq.Ft. AHP : “Mid-2 Level HSP”
- 12167 – 13282 Rs. / Sq.Ft. AHP : “High-1 Level HSP”
- > 13282 Rs. / Sq.Ft. AHP : “High-2 Level HSP”

In summary section, one of my aim was also visualize the Average Housing Sale Prices for per square feet with bar graph. The Bar Graph is plotted as Neighborhoods of Pune city v/s Average Housing Sales Price which is given below. From the Bar Graph we can clearly visualise that the Top 5 Neighborhoods in Pune city having highest Average Housing Prices per Sq. Ft. are Law College road, Prabhat Road, Model Colony, Boat Club Road and Shivaji Nagar. The Bottom 5 Neighborhoods in Pune having Lowest Average Housing Prices per Sq. Ft. are Handewadi, Shivane, Kirkitwadi, Kale Padal and Ambegaon Pathar.

This data will help the people who want to buy right property or invest in right property in Pune city within their budget. The Bar Graph is plotted as Neighborhoods of Pune city v/s Average Housing Sales Price which is given below.

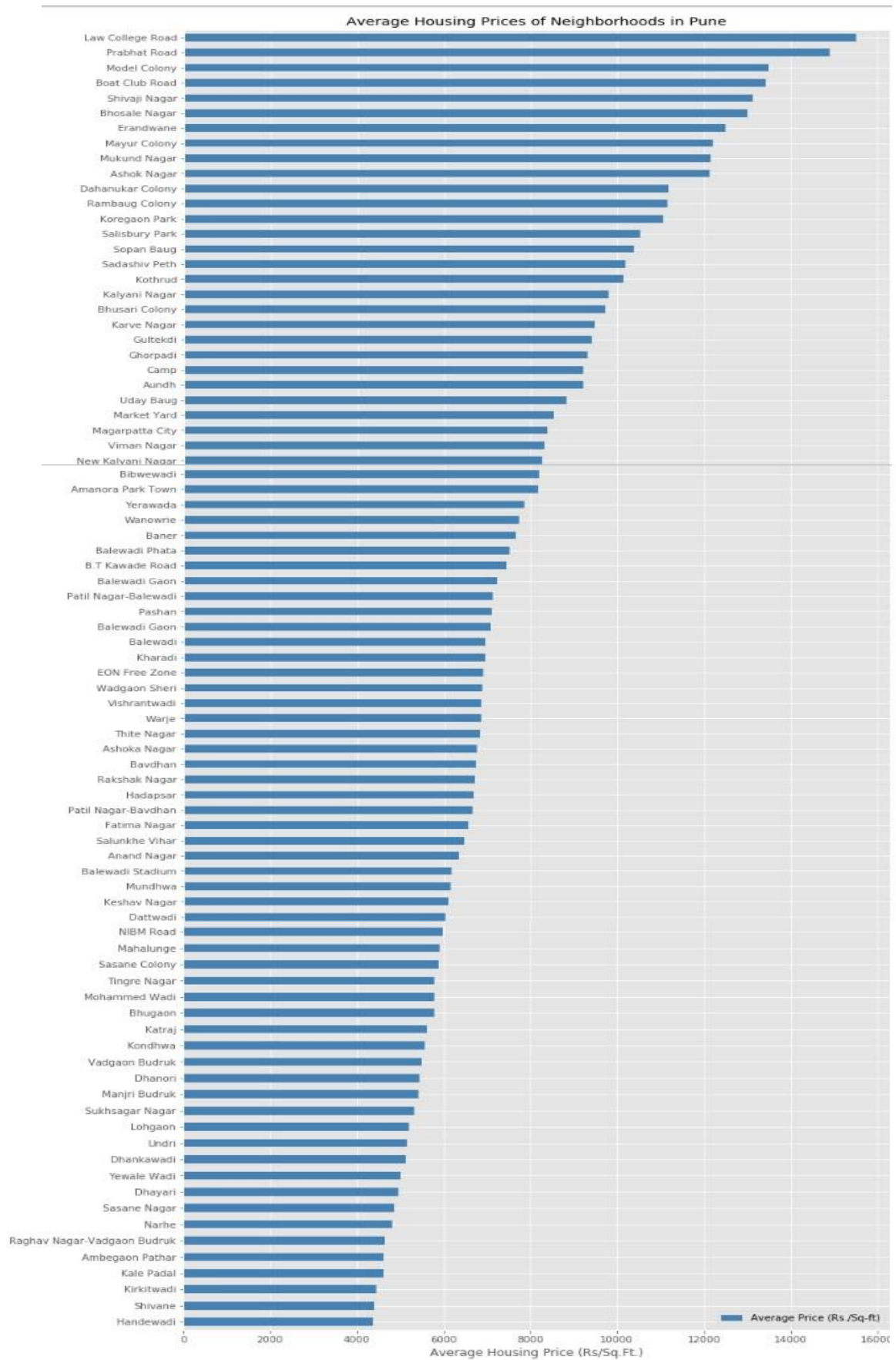


Fig. 4.5

5. DISCUSSION

Pune is a big city with a high population density. The total number of measurements and population densities of the 41 boroughs and 84 Neighborhoods in total can vary. As there is such a complexity, very different approaches can be tried in clustering and classification studies. Moreover, it is obvious that not every classification method can yield the same high quality results.

I used the K-means algorithm as part of this clustering study. I set the optimum k value to 5. However, only 84 neighborhood coordinates were used. For more detailed and accurate guidance, the data set can be expanded and the details of the neighborhood or street can also be explored.

I also found out the Top 10 venues in each of the neighborhood of Pune city using Foursquare API app.

I also performed data analysis through this information by adding the location coordinates of boroughs, neighborhoods and average home sales price as master data on GitHub. In future studies, these data can also be accessed dynamically from specific platforms or packages.

I ended the study by visualizing the data using Bar Graph and clustering information on the Pune city map. In future studies, web or telephone applications can be carried out to direct investors.

6. CONCLUSION

More and more people are turning to big cities such as Pune to start a business or work. For this reason, people can achieve better outcomes through their access to the platforms where such information is provided. This project will help people to choose the right neighborhood having easy access to required venues in Pune city within their budget.

This project will not only help investors to buy right properties but also help city managers to manage the city better by using similar data analysis types or platforms.

7. REFERENCES

- [1] Pune – Wikipedia <https://en.wikipedia.org/wiki/Pune>
- [2] Pune Municipal Corporation Website <https://pmc.gov.in>
- [3] Foursquare API <https://foursquare.com>
- [4] Housing Sales Prices of Pune Boroughs and Neighborhoods from <https://magicbricks.com>
- [5] Google Map – Pune.
- [6] <https://www.linkedin.com/pulse/housing-sales-prices-venues-data-analysis-ofistanbul-sercan-y%C4%B1ld%C4%B1z/>

-----*****-----