Housing Sales Prices & Venues Data Analysis of Mumbai

Sumit Singh

30th April 2020

Introduction

▶ Background

- Mumbai is one of the largest metropolises in the world where over 19.98 million people live and it has a population density of 21000 people per square kilometer. As a resident of this city, I decided to use Mumbai in my project. The city is divided into 4 suburban zone in total. However, the fact that the four suburban zones are squeezed into an area of approximately 603 square kilometers causes the city to have a very intertwined and mixed structure [1].

▶ Problem Descriptions

- As you can see from the figures, Mumbai 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 suburban zone 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 region 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 Mumbai and each neighborhoods is clustered according to the venue density.

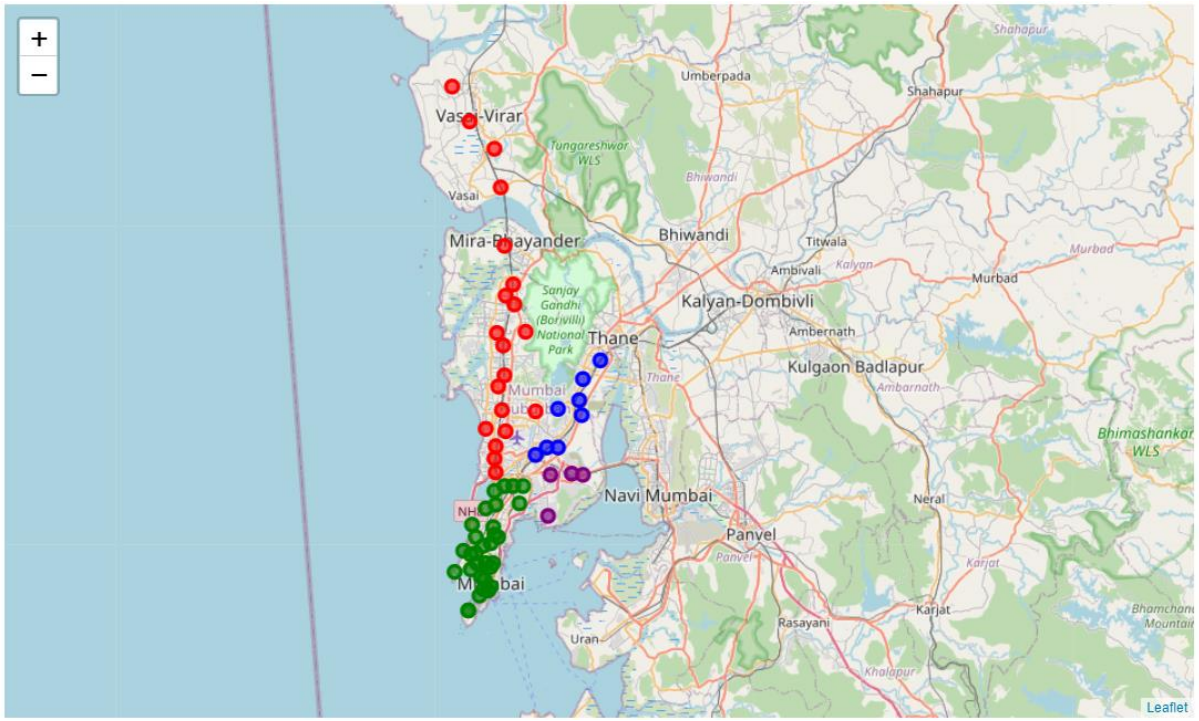
Data Description

- I found the neighborhoods of Mumbai from Wikipedia. The coordinate of neighborhoods was obtained from google map. https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai
- I cleaned the data and reduced it to city of Istanbul where I used it to create choropleth map of Housing Sales Price for each neighborhoods was obtained from 99acres property dealing site.
- https://www.99acres.com/property-rates-and-price-trends-in-mumbai
- I used Forsquare API to get the most common venues of given Neighborhoods of Mumbai. https://developer.foursquare.com/
- I used Google Map, 'Search Nearby' option to get the center coordinates of the each Neighborhood. https://www.google.co.in/maps/
- All compiled data was also stored in Github repository. https://github.com/Sumsys/Applied_Data_Science_Capstone_Project/tree/master/Week5

Methodology

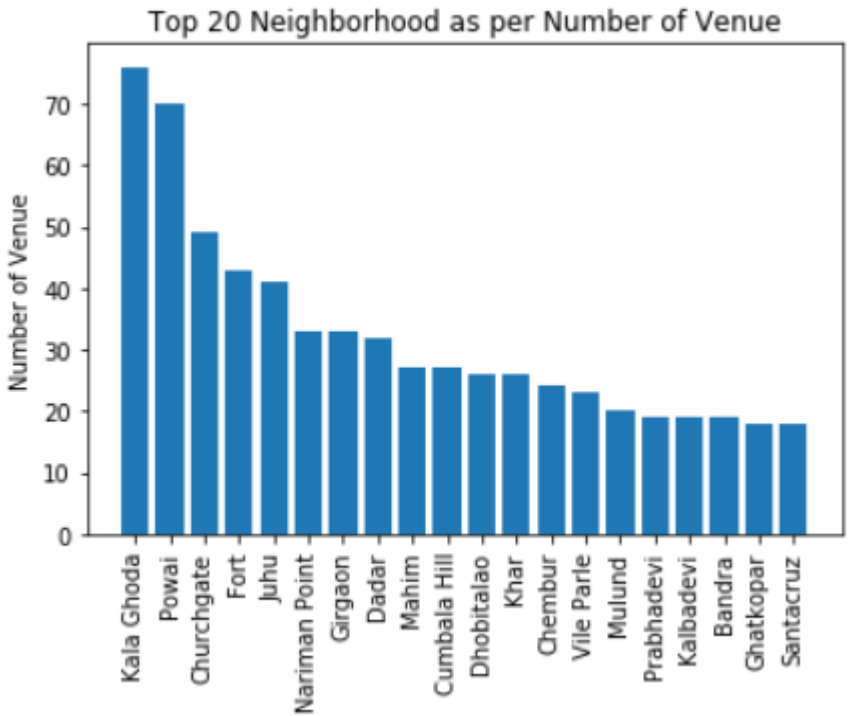
▸ Neighborhood Data

	Neighborhood	Suburbs_Zone	Latitude	Longitude	Avarage House Price	grid
0	Andheri	Western Suburbs	19.119000	72.847000	18000	19
1	Marol	Western Suburbs	19.117300	72.884000	16000	19
2	Bhayandar	Western Suburbs	19.290000	72.850000	7500	ot
3	Bandra	Western Suburbs	19.054444	72.840556	32000	23
4	Borivali (West)	Western Suburbs	19.238100	72.852300	16000	11



▸ Venue Data from Foursquare API

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Andheri	19.119	72.847	Merwans Cake shop	19.119300	72.845418	Bakery
1	Andheri	19.119	72.847	Narayan Sandwich	19.121398	72.850270	Sandwich Place
2	Andheri	19.119	72.847	McDonald's	19.119691	72.846102	Fast Food Restaurant
3	Andheri	19.119	72.847	Cafe Alfa	19.119667	72.843560	Indian Restaurant
4	Andheri	19.119	72.847	Vaibhav Restaurant	19.118235	72.847991	Indian Restaurant



Methodology (Contd.)

▸ K-Means Cluster

- Top 10 Venues

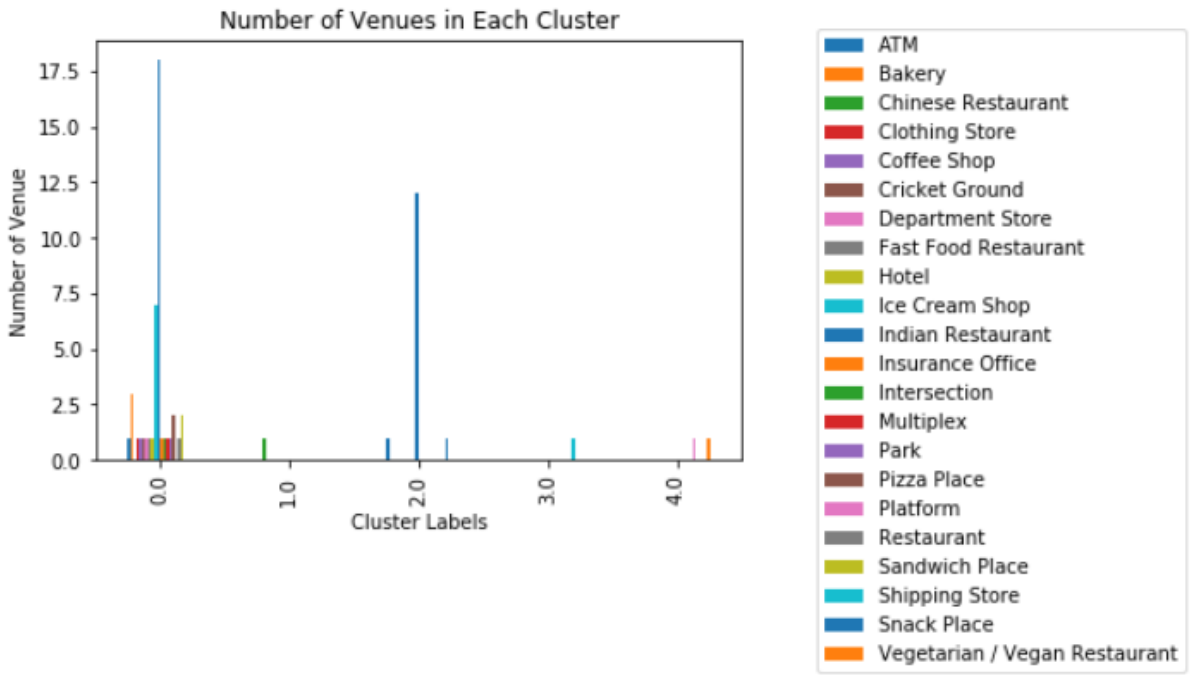
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agripada	Bakery	Indian Restaurant	Coffee Shop	Gym	Diner	Flower Shop	Flea Market	Field	Fast Food Restaurant	Electronics Store
1	Andheri	Indian Restaurant	Fast Food Restaurant	Bakery	Platform	Restaurant	Sandwich Place	Burger Joint	Food Truck	Food Court	Gym
2	Antop Hill	Indian Restaurant	Grocery Store	Train Station	Diner	Flower Shop	Flea Market	Field	Fast Food Restaurant	Electronics Store	Dumpling Restaurant
3	Bandra	Indian Restaurant	Café	Paper / Office Supplies Store	Italian Restaurant	Brewery	Breakfast Spot	Restaurant	Furniture / Home Store	Lake	Platform
4	Bhandup	Indian Restaurant	Bar	Donut Shop	Food	Flower Shop	Flea Market	Field	Fast Food Restaurant	Electronics Store	Dumpling Restaurant

- Top 10 Venues merge with Neighborhood data

	Neighborhood	Suburbs_Zone	Latitude	Longitude	Avarage House Price	grid	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 Com V
0	Andheri	Western Suburbs	19.119000	72.847000	18000	19	0.0	Indian Restaurant	Fast Food Restaurant	Bakery	Platform	Restaurant	Sandwich Place	B
1	Marol	Western Suburbs	19.117300	72.884000	16000	19	2.0	Snack Place	Indian Restaurant	Hotel	Gym	Diner	Hotel Pool	D
2	Bhayandar	Western Suburbs	19.290000	72.850000	7500	ot	3.0	Shipping Store	Zoo	Donut Shop	Food	Flower Shop	Flea Market	
3	Bandra	Western Suburbs	19.054444	72.840556	32000	23	2.0	Indian Restaurant	Café	Paper / Office Supplies Store	Italian Restaurant	Brewery	Breakfast Spot	Resta
4	Borivali (West)	Western Suburbs	19.238100	72.852300	16000	11	0.0	Park	Intersection	Pizza Place	Shop & Service	Soccer Field	Italian Restaurant	

▸ Cluster Classification based on K-Means

- Cluster 0 : “Multiple Social Venues”
- Cluster 1 : “Chinese Restaurant Venues”
- Cluster 2 : “Indian Restaurents Venues”
- Cluster 3 : “Shipping Store Venues”
- Cluster 4 : “Vegetarian Restaurant and Platform Venues”



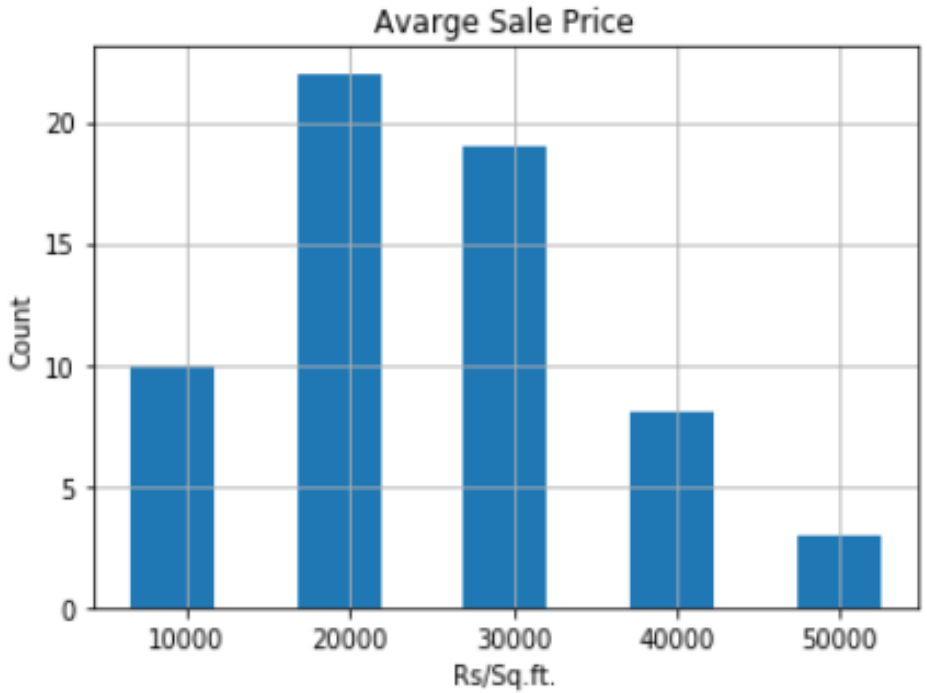
Methodology (Contd.)

▶ Most common Venue in each cluster

Cluster Labels	1st Most Common Venue	Count
0.0	ATM	1
	Bakery	3
	Clothing Store	1
	Coffee Shop	1
	Cricket Ground	1
	Department Store	1
	Fast Food Restaurant	1
	Hotel	1
	Ice Cream Shop	7
	Indian Restaurant	18
	Insurance Office	1
	Intersection	1
	Multiplex	1
	Park	1
	Pizza Place	2
	Restaurant	1
	Sandwich Place	2
1.0	Chinese Restaurant	1
2.0	ATM	1
	Indian Restaurant	12
	Snack Place	1
3.0	Shipping Store	1
4.0	Platform	1
	Vegetarian / Vegan Restaurant	1

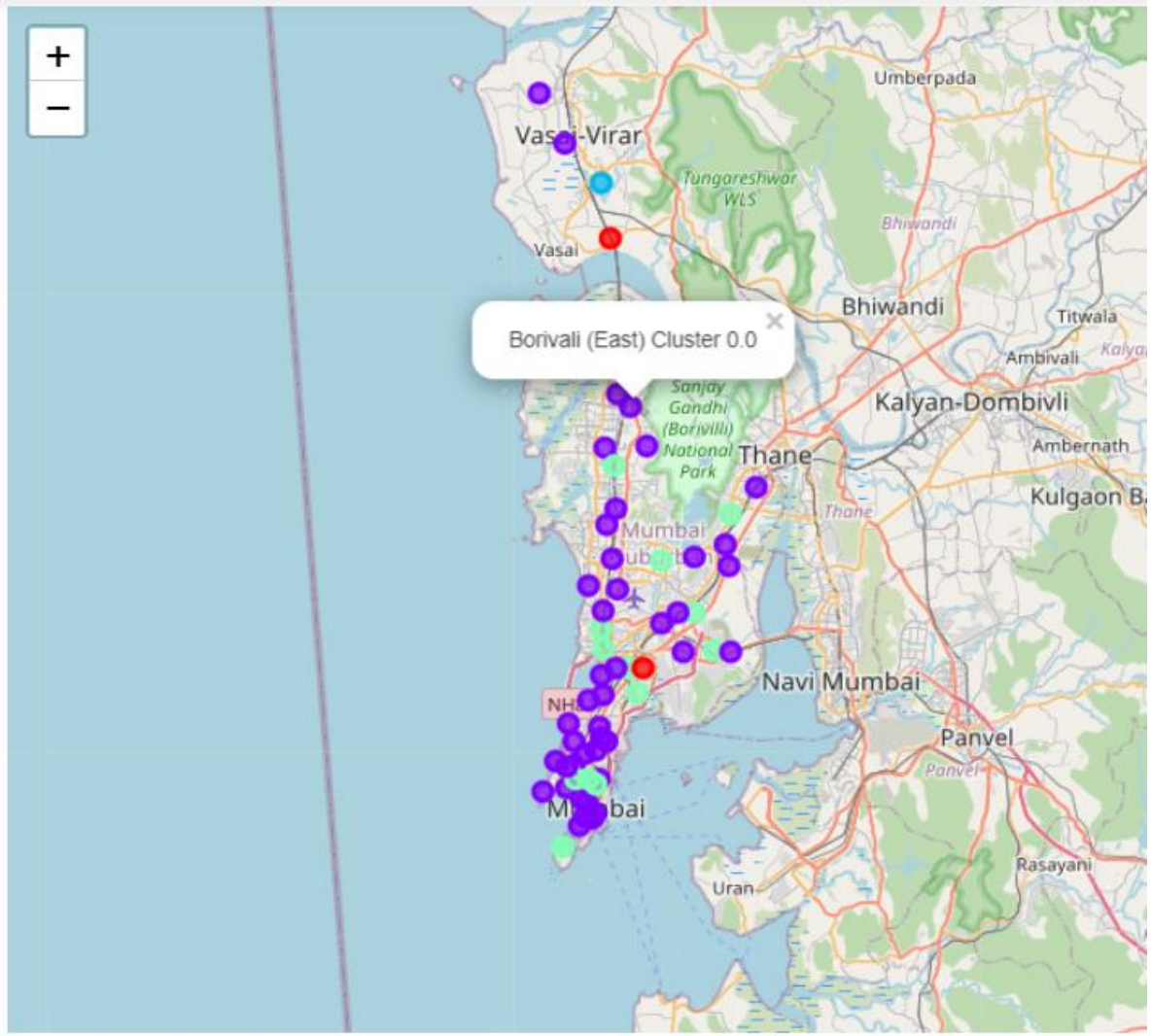
▶ Average Sales price

- 15000 ASP : “Low Level HSP”
- 15000-25000 ASP : “Mid-1 Level HSP”
- 25000-35000 ASP : “Mid-2 Level HSP”
- 35000-45000 ASP : “High-1 Level HSP”
- > 45000 AHP : “High-2 Level HSP”



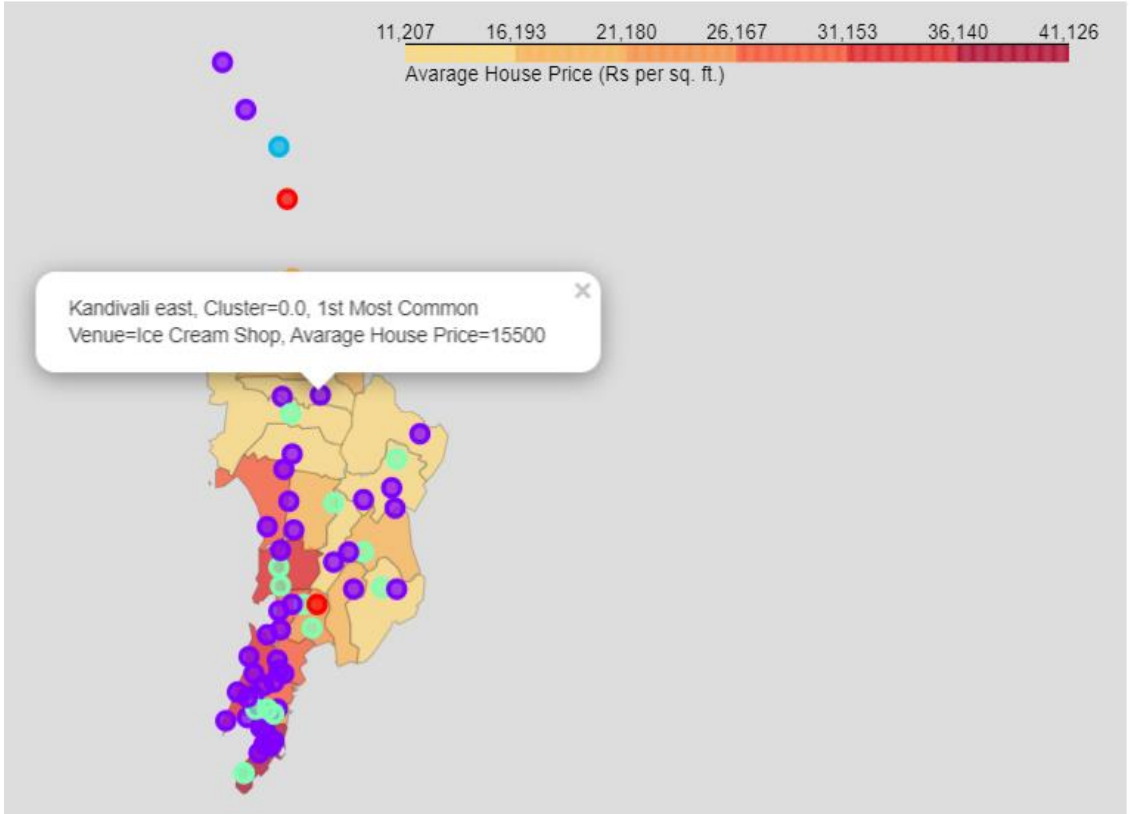
Results

▸ Neighborhood Cluster



▸ Average Sales price Choropleth Map

- Data provided for each Neighborhoos
  - Neighbourhood name,
  - Cluster,
  - 1st Most Common Venue,
  - Average Housing Price.



Conclusion and future directions

- As a result, people are turning to big cities 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.
- Not only for investors but also city managers can manage the city more regularly by using similar data analysis types or platforms.