

Introduction: Business Problem

Mumbai is the financial, commercial and entertainment capital of India and contributes to 6.16% of India's GDP. According to United Nations, Mumbai is the second-most populous city in India after Delhi. Investors and entrepreneurs all over the world want to invest in Mumbai due to its business opportunities and large labor-force.

Due to its high population, the food and beverage industry can prosper with a reasonably moderate investment and large profit margins. Thus, this project (report) can provide a perspective to entrepreneurs and investors from all over the world looking to invest in beverage and food industry in Mumbai. The analysis will provide a perspective on various areas (clusters) of Mumbai and will be able to tell where opening a beverage shop or a restaurant can prove profitable to the investor.

Data

The data for the list of neighborhoods in Mumbai is obtained from https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

The different venues and its queries is obtained from the FourSquare APIs. The dataframe will contain top 10 venues in each cluster of areas and will provide adequate information to make a decision about where to open beverage shops or restaurants. K-means clustering algorithm of machine learning will be used to cluster the areas (districts) of Mumbai with high accuracy and will be used to find pattern between the different districts and clusters.

Thus, the data will be used to meet the following objectives:

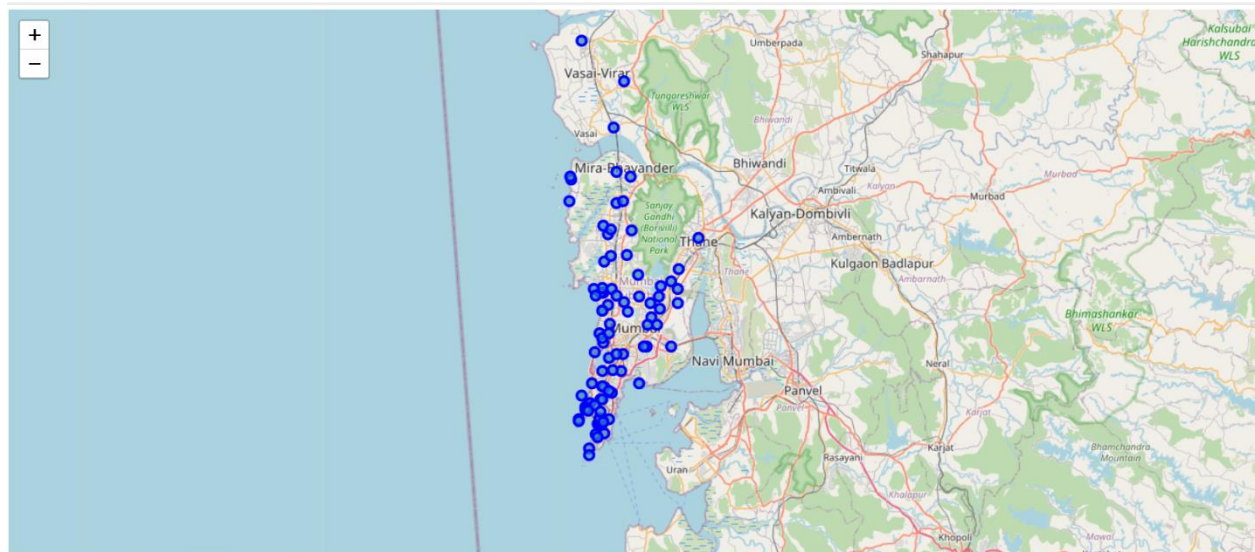
- Neighborhoods and Areas of Mumbai
- Trending Venues of the Areas
- Categorizing the venues
- Cluster different Areas of Mumbai using K-means algorithm

Then we gather the data from the Wikipedia page and extract the table or the Dataframe from the site using BeautifulSoup package and append it to our Dataframe.

	Neighborhood	Borough	Latitude	Longitude
0	Amboli	Andheri,Western Suburbs	19.1293	72.8434
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833
2	D.N. Nagar	Andheri,Western Suburbs	19.124085	72.831373
3	Four Bungalows	Andheri,Western Suburbs	19.124714	72.82721
4	Lokhandwala	Andheri,Western Suburbs	19.130815	72.82927
5	Marol	Andheri,Western Suburbs	19.119219	72.882743
6	Sahar	Andheri,Western Suburbs	19.098889	72.867222
7	Seven Bungalows	Andheri,Western Suburbs	19.129052	72.817018
8	Versova	Andheri,Western Suburbs	19.12	72.82
9	Mira Road	Mira-Bhayandar,Western Suburbs	19.284167	72.871111

Note: (snapshot of the Dataframe, not the entire Dataframe)

We use the Folium library and geopy library to get the map of Mumbai and all the neighborhoods marked in it.



Then we use the FourSquare API calls to get the top 100 nearby and popular venues for all the neighborhoods in the boroughs of Mumbai.

Methodology

Our objective is to find the best suitable areas in Mumbai to open a restaurant or a beverage shop. We will use the K-means clustering algorithm to achieve our objective.

We will use the one-hot encoding method on the Venues Dataframe and then group it by Neighborhoods. The one-hot encoding will return the venue categories as column per neighborhood, and then it will be grouped together to provide the weighting of venue type occurrence on each neighborhood.

K-means clustering algorithm will help us cluster the neighborhoods based on the top venues in the encoded dataframe and provide cluster labels for similar neighborhoods. We will then observe the clusters one by one to determine it's content and then provide the appropriate recommendation.

Data Analysis

We use the one-hot encoding approach to analyze each neighborhood and then group it by neighborhoods and get the 10 most popular and common venues around each neighborhood in the boroughs of Mumbai.

	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	Gym	Coffee Shop	Indian Restaurant	Bakery	Dim Sum Restaurant	Farmers Market	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
1	Altamount Road	Café	Pizza Place	Sandwich Place	Bakery	Indian Restaurant	Theater	Coffee Shop	Dance Studio	Creperie	Cosmetics Shop
2	Amboli	Park	Fast Food Restaurant	Gym	Coffee Shop	Sandwich Place	Chinese Restaurant	Indian Restaurant	Dessert Shop	Electronics Store	Donut Shop
3	Amrut Nagar	Indian Restaurant	Café	Electronics Store	Fast Food Restaurant	Restaurant	Asian Restaurant	Falafel Restaurant	Bookstore	Bowling Alley	Brewery
4	Asalfa	Park	Men's Store	Hostel	Playground	Food Truck	Dhaba	Event Space	Electronics Store	Donut Shop	Dog Run
5	Ballard Estate	Harbor / Marina	Convenience Store	Hotel	Indian Restaurant	Grocery Store	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
6	Bandstand Promenade	Scenic Lookout	Gym	Indian Restaurant	Fast Food Restaurant	Lounge	Beach	Café	Food Truck	Chinese Restaurant	Italian Restaurant

Machine Learning algorithm

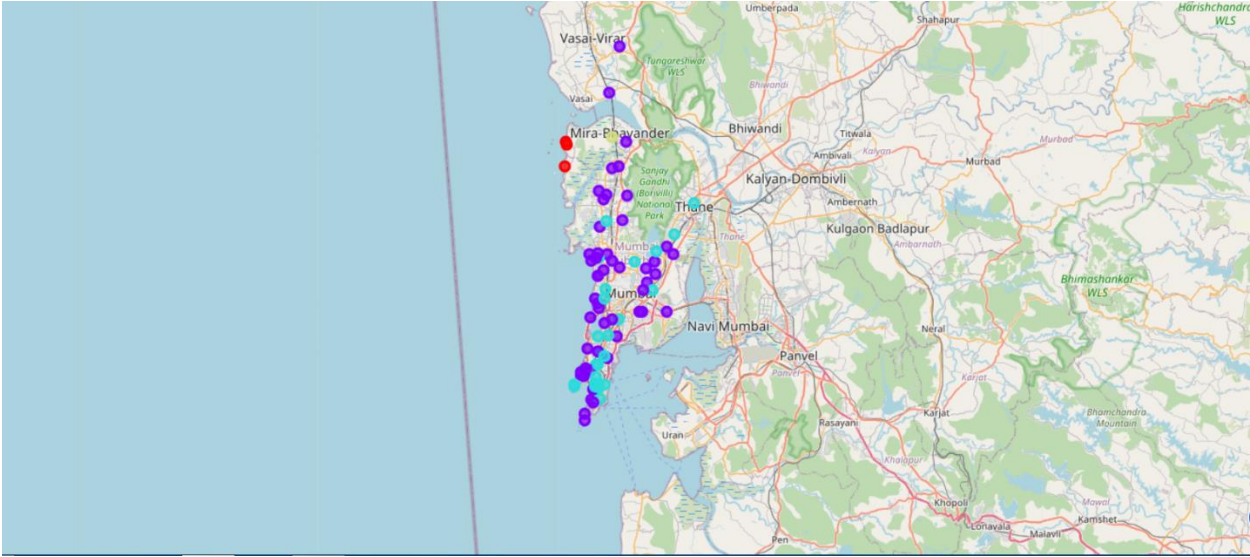
K-means Clustering

We use K-means clustering method to cluster the neighborhoods. We assign the number of clusters to be 4 (at elbow point). We get a new Dataframe with cluster labels assigned to each neighborhood.

	Neighborhood	Borough	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
0	Amboli	Andheri,Western Suburbs	19.129300	72.843400	1.0	Park	Fast Food Restaurant	Gym	Coffee Shop	Sandwich Place	Chinese Restaurant	Indian Restaurant	Dessert Shop	Electronics Store	Donut Shop
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833	1.0	Café	Fast Food Restaurant	Restaurant	Hotel	Multiplex	Indian Restaurant	Falafel Restaurant	Salon / Barbershop	Diner	Asian Restaurant
2	D.N. Nagar	Andheri,Western Suburbs	19.124085	72.831373	2.0	Gym / Fitness Center	Indian Restaurant	Cocktail Bar	Pizza Place	Snack Place	Lounge	German Restaurant	Department Store	Electronics Store	Donut Shop
3	Four Bungalows	Andheri,Western Suburbs	19.124714	72.827210	1.0	Women's Store	Bar	Gym	Ice Cream Shop	Juice Bar	Fish Market	Market	Electronics Store	Pizza Place	Residential Building (Apartment / Condo)
4	Lokhandwala	Andheri,Western Suburbs	19.130815	72.829270	1.0	Pub	Lounge	Women's Store	Indian Restaurant	Coffee Shop	Cocktail Bar	Pizza Place	Department Store	Residential Building (Apartment / Condo)	Market

Data Visualization

We use the Folium library to visualize the clusters based on the cluster labels assigned to the neighborhoods.



Results Section

Here we see the results of our Data Analysis, Machine Learning algorithm (K-Means) to cluster the neighborhoods and Data Visualization

Cluster 1

Looking at Cluster 1 we can say that there is an abundance of restaurants of different cuisines (Indian, Dhaba, Seafood, Falafel (Greek), Donut Shop) but they are lacking in beverage shops like a coffee shops or cafe's.

	Borough	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
10	Mira-Bhayandar, Western Suburbs	Beach	Playground	Indian Restaurant	Resort	Bus Station	Women's Store	Dhaba	Event Space	Electronics Store	Donut Shop
15	Borivali (West), Western Suburbs	Resort	Seafood Restaurant	Aquarium	Indian Restaurant	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Dog Run
58	South Mumbai	Beach	Playground	Indian Restaurant	Resort	Bus Station	Women's Store	Dhaba	Event Space	Electronics Store	Donut Shop

This is the most extensive cluster spread across the entire map. While the Western Suburbs have plenty of restaurants of many cuisines and beverage shops of different kinds, the Eastern Suburbs and the Harbour Suburbs look to have a deficiency in good Chinese restaurants.



Cluster 3

While the South Mumbai areas have a lot of restaurants, bars and beverage shops, there is a deficiency in Andheri (Western Suburb area) in terms of coffee shops and cafe's.

	Borough	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
2	Andheri, Western Suburbs	Gym / Fitness Center	Indian Restaurant	Cocktail Bar	Pizza Place	Snack Place	Lounge	German Restaurant	Department Store	Electronics Store	Donut Shop
5	Andheri, Western Suburbs	Indian Restaurant	Snack Place	Coffee Shop	Ice Cream Shop	Diner	Restaurant	Food	Bakery	Asian Restaurant	Convenience Store
25	Khar, Western Suburbs	Indian Restaurant	Bar	Lounge	Pub	Fast Food Restaurant	Dessert Shop	Beer Garden	Bengali Restaurant	Restaurant	Hotel
27	Malad, Western Suburbs	Indian Restaurant	Gym / Fitness Center	Coffee Shop	Chinese Restaurant	Hotel	Dessert Shop	Café	Bus Station	Diner	Lounge
28	Sanctacruz, Western Suburbs	Indian Restaurant	Women's Store	Dance Studio	Clothing Store	Chinese Restaurant	Moving Target	Platform	Middle Eastern Restaurant	Lounge	Sandwich Place
36	Ghatkopar, Eastern Suburbs	Indian Restaurant	Ice Cream Shop	Pizza Place	Snack Place	Bank	Bakery	Farmers Market	Multiplex	Arcade	Coffee Shop
38	Mulund, Eastern Suburbs	Indian Restaurant	Restaurant	Ice Cream Shop	Bus Station	Women's Store	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
41	Powai, Eastern Suburbs	Indian Restaurant	Concert Hall	Event Space	Coffee Shop	Diner	Bakery	Café	Dhaba	Falafel Restaurant	Electronics Store
46	South Mumbai	Gym	Coffee Shop	Indian Restaurant	Bakery	Dim Sum Restaurant	Farmers Market	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
48	South Mumbai	Indian Restaurant	Market	Fast Food Restaurant	American Restaurant	Restaurant	Food	Cheese Shop	Snack Place	Ice Cream Shop	Jewelry Store
51	South Mumbai	Indian Restaurant	Café	Fast Food Restaurant	Jewelry Store	Bakery	Portuguese Restaurant	Market	Cheese Shop	Chinese Restaurant	Bar
56	South Mumbai	Coffee Shop	Maharashtrian Restaurant	Plaza	Indian Restaurant	Women's Store	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
57	South Mumbai	Indian Restaurant	Bar	Bakery	Café	Music Store	Pizza Place	Bridal Shop	Portuguese Restaurant	Sandwich Place	Market
59	South Mumbai	Indian Restaurant	Dessert Shop	Coffee Shop	Café	Fast Food Restaurant	Seafood Restaurant	Chinese Restaurant	Juice Bar	Hotel	Boutique
63	South Mumbai	Convenience Store	Ice Cream Shop	Lighthouse	Indian Restaurant	Department Store	Coffee Shop	Cricket Ground	Diner	Falafel Restaurant	Event Space
68	South Mumbai	Indian Restaurant	Electronics Store	Café	Convenience Store	Pizza Place	Bakery	Dessert Shop	Smoke Shop	Snack Place	Theater
69	South Mumbai	Indian Restaurant	Vegetarian / Vegan Restaurant	Café	Tea Room	Sandwich Place	Snack Place	Movie Theater	Cosmetics Shop	Gym / Fitness Center	Gym
70	South Mumbai	Convenience Store	Ice Cream Shop	Lighthouse	Indian Restaurant	Department Store	Coffee Shop	Cricket Ground	Diner	Falafel Restaurant	Event Space
73	Byculla, South Mumbai	Indian Restaurant	Cupcake Shop	Art Gallery	Women's Store	Dim Sum Restaurant	Farmers Market	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
75	Fort South Mumbai	Harbor / Marina	Convenience Store	Hotel	Indian Restaurant	Grocery Store	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop

Cluster 4

This a single cluster and has many good restaurants nearby but not enough beverage shops.

	Borough	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
9	Mira-Bhayandar, Western Suburbs	Shipping Store	Women's Store	Dhaba	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Dog Run	Diner	Dim Sum Restaurant

Discussion Section

In Cluster 1 we observe that there is a certain deficiency in beverage shops. Looking at other clusters and seeing the variety of beverage shops like bars, pubs, coffee shops and cafe's being one of the most popular venues, it would be viable for an investor or an entrepreneur to open a beverage shop of their liking.

Cluster 2 is the most extensive cluster of all and covers a lot of the map of Mumbai especially central and southern Mumbai. While the western suburbs have a lot of restaurants and beverage shops, it would not be profitable to open either of them without facing stiff competition. The harbour suburbs lack in Chinese cuisine based restaurants and thus that could be a good investment for a lot less competition. The same can be said about eastern suburbs but due to less foot traffic, the profit margin might not be substantial and can face a little competition from a few chinese restaurants placed there.

Cluster 3 is concentrated in the south of Mumbai and traverses a little towards the western areas of Mumbai. This cluster shows that South Mumbai although may not be lacking in terms of restaurants and coffee shops, it is still a commercial side of Mumbai and thus foot traffic will be high and hence can provide to be a market for a good restaurant or beverage shop depending on intricate places where the investor wants to open the restaraunts or beverage shops. The Andheri area in the western suburb is a stand out in that cluster as the presence of coffee shops and cafe's is scarce and thus can provide a good, less competitve market for investors.

Cluster 4 contains only one borough, and thus looking at the popular venues we can safely suggest that the area is deficient in terms of beverage shops and options like a cafe, coffee shop, restobar, pub can all prove to be successful with less-hassle.

Conclusion

Investors and entrepreneurs may particularly find the South Mumbai region in cluster 3 and harbour suburbs in cluster 2 to be the most financially, and economically viable as the foot traffic in these areas is high with many commercial sites and offices situated in these areas and hence we can conclude this by saying that the investors will find it profitable to invest in a restaurant or a beverage shop in certain parts of Mumbai and be successful.

Note: Due to COVID-19 pandemic the FourSquare calls to get the most popular and nearby venues in all the neighborhoods of boroughs in Mumbai, may not be 100% accurate as the calls depend on the time when the call is made and also on the foot traffic which can prove to be less than the usual days because of the pandemic situation in Mumbai and also in the world. Although I still feel that the results we got will be accurate with a very low error rate.