Introduction

This project deals with the exploring the neighborhoods of Mumbai. This project is designed for those who are looking to start a restaurant business in Mumbai. Mumbai is a financial capital of India and houses all types of cultures. One can find a restaurant of any cuisine in Mumbai. This project is for those who are looking to open a new hotel or restaurant.

The Foursquare API is used to get the details of venues in Mumbai. These venues are then clustered based on data received from the API and then analyzed using the K-Mean Clustering and Silhouette Score.

The target audience of this project are the small-scale hotel owners and restaurant owners who are planning to open branches in Mumbai and its neighborhoods. The project aims to answer the following questions:

1. What is the best location to open a new Hotel in Mumbai?
2. Which place is most suitable for starting a Mall in Mumbai?

Data Requirement

The data required for this projects is taken from two sources:

1. Wikipedia Page for Neighborhoods in Mumbai -
   http://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mumbai
2. Foursquare APIS

Mumbai has 93 neighborhoods and this is seen in the Wikipedia page.

	Area	Location	Latitude	Longitude
0	Ambol	Andheri,Western Suburb	19.129	72.843
1	Chakala, Andher	Western Suburb	19.11138	72.86083
2	D.N. Naga	Andheri,Western Suburb	19.12408	72.83137
3	Four Bungalow	Andheri,Western Suburb	19.12471	72.8272
4	Lokhandwal	Andheri,Western Suburb	19.13081	72.8292

The Foursquare API will return a list of venues when we provide it with the latitudes and longitudes.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ambol	19.129	72.843	Cafe Arfa	19.128930	72.847140	Indian Restaurant
1	Ambol	19.129	72.843	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
2	Ambol	19.129	72.843	Subway	19.127860	72.844461	Sandwich Place
3	Ambol	19.129	72.843	Cafe Coffee Day	19.127748	72.844663	Coffee Shop
4	Ambol	19.129	72.843	Apple Service Centre	19.128036	72.842718	IT Services

Import Libraries required for the project

- Regex
- Json
- NumPy
- pandas
- Beautiful Soup
- Requests
- Scikit Learn
- Geo Py
- Folium
- Matplotlib
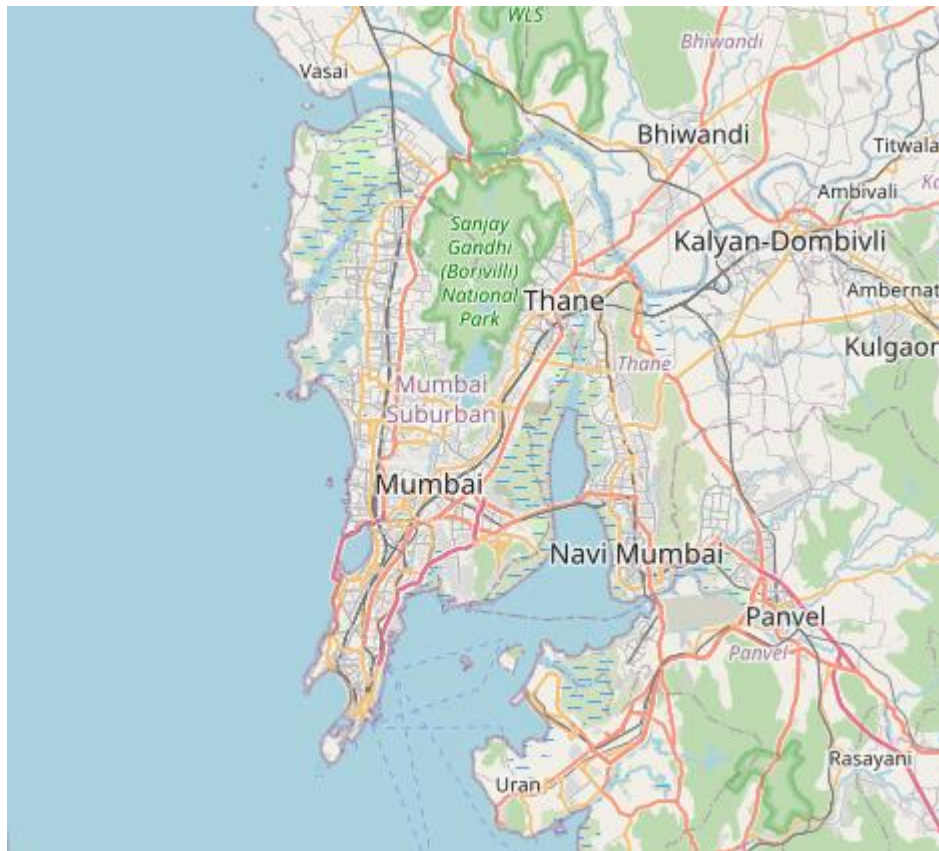
Map of Mumbai
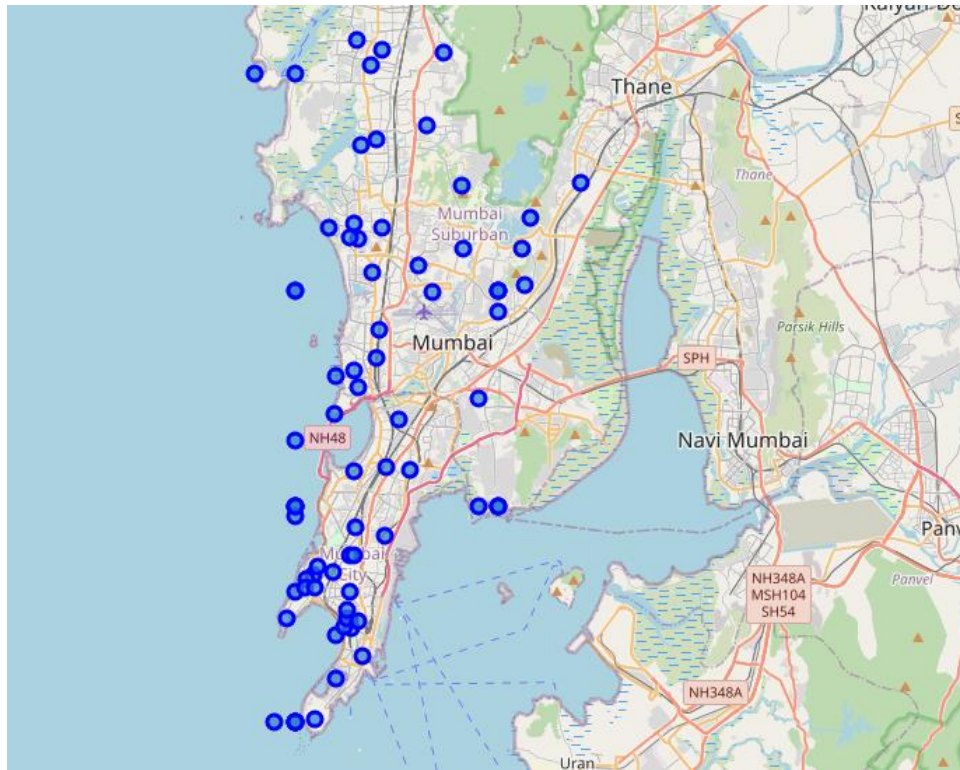
Map of Mumbai with neighborhoods superimposed.

Methodology

We have the data that has latitudes and longitudes of the neighborhoods in Mumbai and the venues that are nearby. We will only consider neighborhoods for which we have considerable number of venues available. We will have a look at the neighborhood that has the highest number of venues. We have 93 neighborhoods in Mumbai. From the foursquare Api we have received 858 venues. We will take a look at the venue categories and then look at how many unique types of venues were received. We use the KMean Clustering, One Head Encoding. These cluster are used to find the best place to open a restaurant.

Analysis

We first look at the distribution of venues in the dataset that we have made.
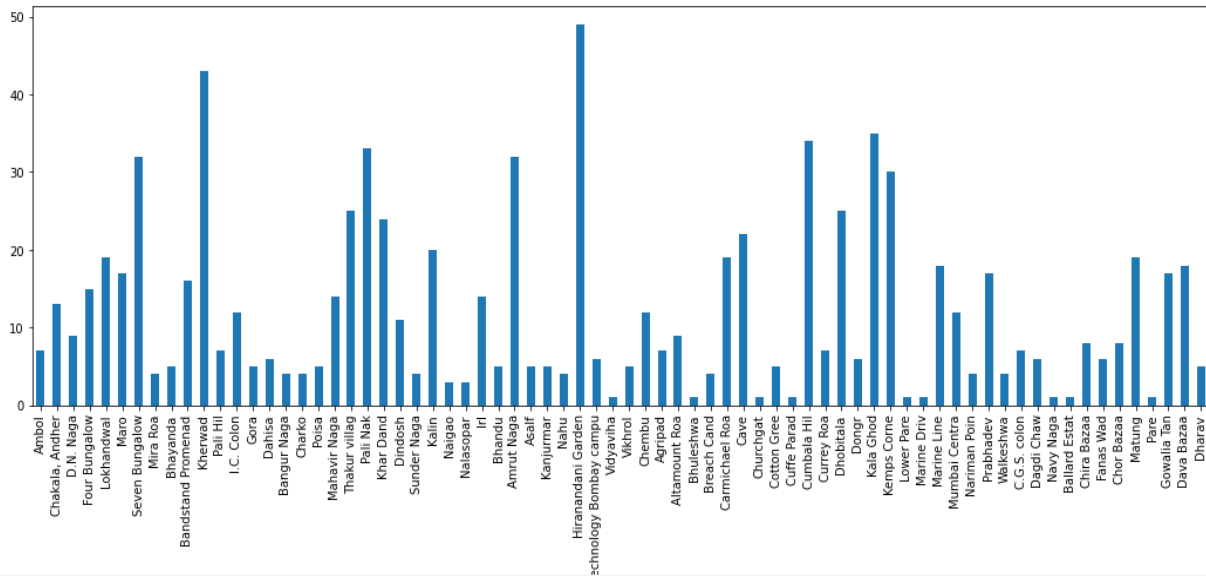
We can see that some neighborhoods have very little venues data available. These areas are not required in the study. For filtering we remove all the samples with less than 10 venues.
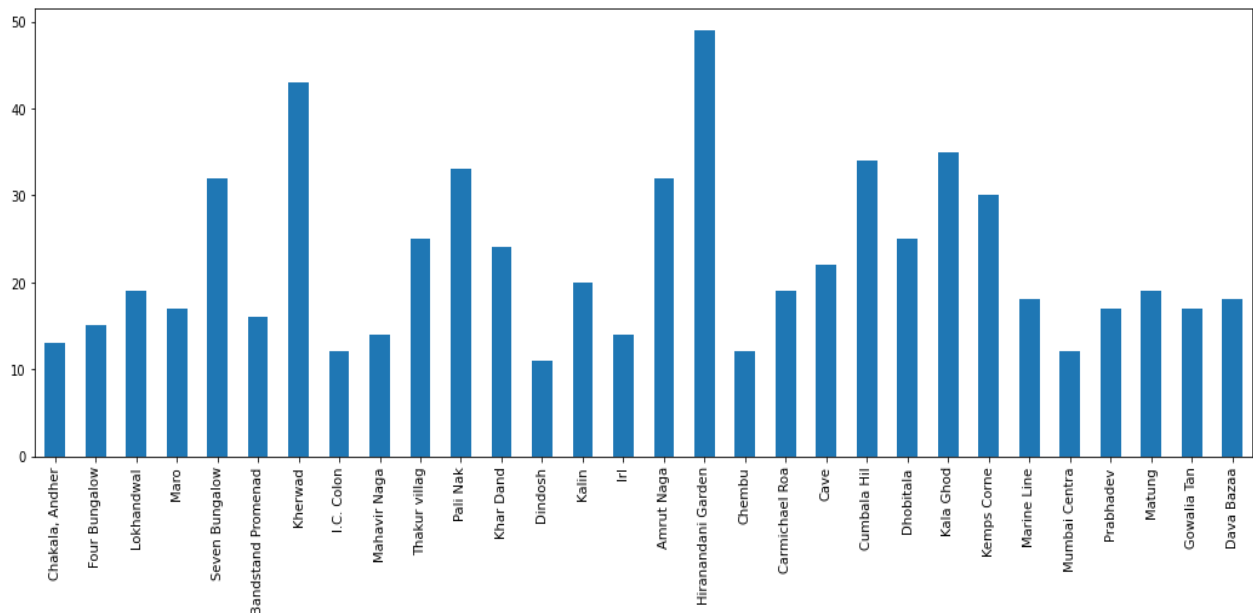
The above figure shows only neighborhoods that are greater than 10 venues.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
7	Chakala, Andher	19.11138	72.86083	Cafe Coffee Day	19.112272	72.861106	Café
8	Chakala, Andher	19.11138	72.86083	Shree	19.112256	72.861113	Restaurant
9	Chakala, Andher	19.11138	72.86083	Courtyard Mumbai International Airport	19.114167	72.864131	Hotel
10	Chakala, Andher	19.11138	72.86083	Sai Palace Hotel	19.115373	72.860571	Hotel
11	Chakala, Andher	19.11138	72.86083	NH1 Kitchen and Bar	19.111335	72.858639	Cocktail Bar

We will use One Hot Encoding to get the different venues and their categories in the different neighborhoods.

	Neighborhood	Accessories Store	Afghan Restaurant	American Restaurant	Amphitheater	Arcade	Art Gallery	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Bagel Shop	Bakery	Bar	Beach	Beer Bar	Beer Garden	Bengali Restaurant
7	Chakala, Andher	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Chakala, Andher	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	Chakala, Andher	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Chakala, Andher	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Chakala, Andher	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Find the mean by grouping by neighborhoods.

	Neighborhood	Accessories Store	Afghan Restaurant	American Restaurant	Amphitheater	Arcade	Art Gallery	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Bagel Shop	Bakery	Bar	Beach	Beer Bar	Beer Garden	Be Resta
0	Amrut Naga	0.0	0.03125	0.03125	0.0	0.0	0.0	0.0	0.0	0.031250	0.0	0.0	0.000000	0.000000	0.0000	0.0	0.0	0.00
1	Bandstand Promenad	0.0	0.00000	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0625	0.0	0.0	0.00
2	Carmichael Roa	0.0	0.00000	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.105263	0.000000	0.0000	0.0	0.0	0.05
3	Cave	0.0	0.00000	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.090909	0.045455	0.0000	0.0	0.0	0.00
4	Chakala, Andher	0.0	0.00000	0.00000	0.0	0.0	0.0	0.0	0.0	0.076923	0.0	0.0	0.000000	0.000000	0.0000	0.0	0.0	0.00

We get the top venues in the above dataset. We then use K Mean Clustering algorithm on this dataset. The K values that provide the best score is 8.

	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	Amrut Naga	Café	Indian Restaurant	Fast Food Restaurant	Clothing Store	Bookstore	Mediterranean Restaurant	Sandwich Place	Bowling Alley	Electronics Store	Chinese Restaurant
1	Bandstand Promenad	Scenic Lookout	Cocktail Bar	Fast Food Restaurant	Boutique	Boat or Ferry	Lounge	Café	Beach	Food Truck	Gym
2	Carmichael Roa	Bakery	Fast Food Restaurant	Pizza Place	Hotel	Italian Restaurant	Food Truck	Electronics Store	Deli / Bodega	Park	Chinese Restaurant
3	Cave	Indian Restaurant	Chinese Restaurant	Café	Bakery	Jewelry Store	Bridal Shop	Movie Theater	Flea Market	Cheese Shop	Bar
4	Chakala, Andher	Restaurant	Hotel	Pizza Place	Café	Fast Food Restaurant	Salon / Barbershop	Multiplex	Punjabi Restaurant	Asian Restaurant	Cocktail Bar

Results
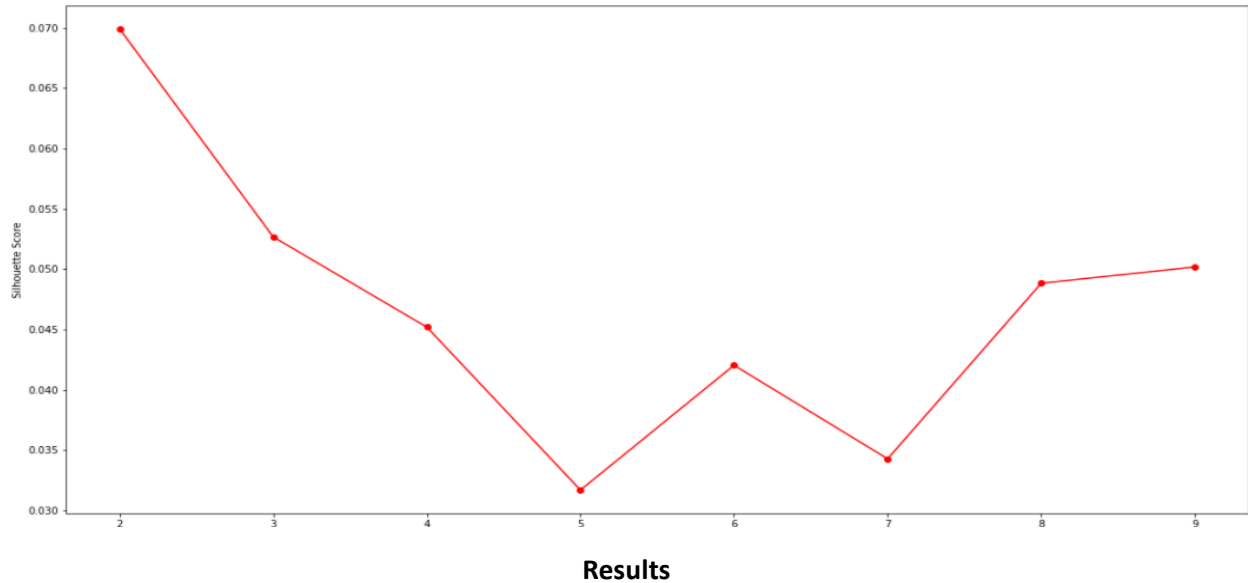
We create 8 clusters and analyze them.

Cluster 1:

The top 3 venues in this cluster are bakery, juice bar, paper and office supply store.

	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
7	I.C. Colon	Bakery	Juice Bar	Paper / Office Supplies Store	Arcade	Indian Restaurant	Tennis Court	Tea Room	Chinese Restaurant	Fast Food Restaurant	Food Truck

Cluster 2:

The top spots are coffee shop, Indian restaurants, ice cream shops and fast-food restaurants.

	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
2	Lokhandwal	Coffee Shop	Lounge	Accessories Store	Juice Bar	Pizza Place	Clothing Store	Pub	Recreation Center	Residential Building (Apartment / Condo)	Restaurant
8	Mahavir Naga	Indian Restaurant	Pizza Place	Market	Gym	Park	Coffee Shop	Café	Dessert Shop	Donut Shop	Snack Place
9	Thakur villag	Ice Cream Shop	Breakfast Spot	Fast Food Restaurant	Pizza Place	Indian Restaurant	Chinese Restaurant	Dessert Shop	Department Store	Residential Building (Apartment / Condo)	Snack Place
25	Mumbai Centra	Fast Food Restaurant	Train Station	Ice Cream Shop	Chinese Restaurant	Department Store	Movie Theater	Snack Place	Indian Restaurant	Pizza Place	Vegetarian / Vegan Restaurant

Cluster 3:

The top restaurants can be seen in the below figure.

	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
1	Four Bungalow	Accessories Store	Sports Club	Women's Store	Fish Market	Juice Bar	Market	Park	Pizza Place	Playground	Residential Building (Apartment / Condo)
3	Maro	Snack Place	Indian Restaurant	Bakery	Gym	Asian Restaurant	Convenience Store	Flea Market	Restaurant	Department Store	Diner
4	Seven Bungalow	Café	Chinese Restaurant	Ice Cream Shop	Pub	Bar	Seafood Restaurant	Dim Sum Restaurant	Electronics Store	Bistro	South Indian Restaurant
5	Bandstand Promenad	Scenic Lookout	Cocktail Bar	Fast Food Restaurant	Boutique	Boat or Ferry	Lounge	Café	Beach	Food Truck	Gym
6	Kherwad	Bakery	Café	Indian Restaurant	Gourmet Shop	Clothing Store	Chinese Restaurant	Pizza Place	Snack Place	Fast Food Restaurant	Bar
10	Pali Nak	Indian Restaurant	Bakery	Bar	Fast Food Restaurant	Seafood Restaurant	Café	Asian Restaurant	Salad Place	Coffee Shop	Deli / Bodega
13	Kalin	Indian Restaurant	Clothing Store	Women's Store	Chinese Restaurant	Moving Target	Bakery	Light Rail Station	Sandwich Place	Steakhouse	Yoga Studio
15	Amrut Naga	Café	Indian Restaurant	Fast Food Restaurant	Clothing Store	Bookstore	Mediterranean Restaurant	Sandwich Place	Bowling Alley	Electronics Store	Chinese Restaurant
16	Hiranandani Garden	Indian Restaurant	Restaurant	Café	Chinese Restaurant	Bar	Bakery	Department Store	Italian Restaurant	Coffee Shop	Arcade
18	Carmichael Roa	Bakery	Fast Food Restaurant	Pizza Place	Hotel	Italian Restaurant	Food Truck	Electronics Store	Deli / Bodega	Park	Chinese Restaurant

Conclusion

Cluster 1 — No. of venues: Tea Room, Paper / Office Supplies Store, Food Truck, Indian Restaurant, Chinese Restaurant

Cluster 2 — No. of venues: Ice Cream Shop, Dessert Shop, Indian Restaurant, Snack Place, Pizza Place

Cluster 3 — No. of venues: Chinese Restaurant, Bakery, Fast Food Restaurant, Indian Restaurant, Café

Cluster 4 — No. of venues: Flea Market, Bar, Indian Restaurant, Café, Fast Food Restaurant

Cluster 5 — No. of venues: Restaurant, Spa, Fast Food Restaurant, Ice Cream Shop, Shopping Mall

Cluster 6 — No. of venues: Gastropub, Chinese Restaurant, Fast Food Restaurant, Electronics Store, Indian Restaurant

Cluster 7 — No. of venues: Restaurant, Salon / Barbershop, Fast Food Restaurant, Café, Punjabi Restaurant

Cluster 8 — No. of venues: Bridal Shop, Chinese Restaurant, Bakery, Café, Indian Restaurant

I.C. Colon Cluster 1

Leaflet | Data by © OpenStreetMap, under ODbL