

Data Science Capstone Project

The Battle of Neighbourhoods – Mumbai vs. New York

By: Akash Kadam

August 2020



Introduction

Mumbai and New York, two cities that never sleep. Both are the financial hub of their respective nations. Each of them attracts immigrants (who come to follow their dreams).

Mumbai, formerly known as **Bombay**, capital of **Maharashtra** state, **India**. It is the country's financial and commercial centre and its principal port on the Arabian Sea. Located on Maharashtra's coast, Mumbai is India's most-populous city and it is one of the largest and most densely populated urban areas in the world. Mumbai is home for 1.84 Cr people. The total area of Mumbai is 603.4 km². Average temperature of Mumbai is 26.9°C.

New York City, officially **the City of New York**, city and port located at the mouth of Hudson River, north-eastern U.S. It is the largest and most influential American metropolis. New York is the most populous and most international city in the United States of America. New York City is in reality a collection of many neighbourhoods.

City's population is 1.9 Cr. Area is 790 km². Average temperature of New York is 17.1°C.

Business Problem

The Objective of this project is to make an attempt to analyse the neighbourhoods in each of these two cities and try to understand what is popular in them and what they have to offer to someone who is contemplating to make a choice on seeking a life in either of the cities. The decision to choose one over another would depend on popular venues in the neighbourhoods in each of these cities.

Target Audience

People who would be interested in this study are those who would like to create a projection of potential life and activities in these metropolitan city neighbourhoods if the subject moves to live in one of them.

Data

For any “Data Science Project” data is of paramount importance. To solve the problem, we will need the following data:

- List of neighbourhoods in Mumbai and New York.
- Latitude and Longitude coordinates of those neighbourhoods. This is required in order to plot the data on map and also to get the venue data.

- Venue data, particularly the top venues in 500 meter radius of the neighbourhoods. We will use this data to perform clustering on the neighbourhoods.

Sources of data and methods to extract them:

- **Mumbai**

This [Wikipedia page](#) contains a table named “Mumbai neighbourhood coordinates”. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup packages.

- **New York**

We will download the New York data using “wget” command in Python

“wget -q -O 'newyork_data.json'

https://cocl.us/new_york_dataset”. Note that this is a .json file and we need to import json library in Python to handle this file.

- **Venues Data**

Now, we will use Foursquare API to get the venue data for the neighbourhoods. Foursquare API will provide many categories of the venue data.

Methodology

On a high level, we do the following. After tidying up and exploring the data, we will apply the K-means machine learning technique for creating clusters of districts. We will use the silhouette score for choosing the optimal number of clusters.

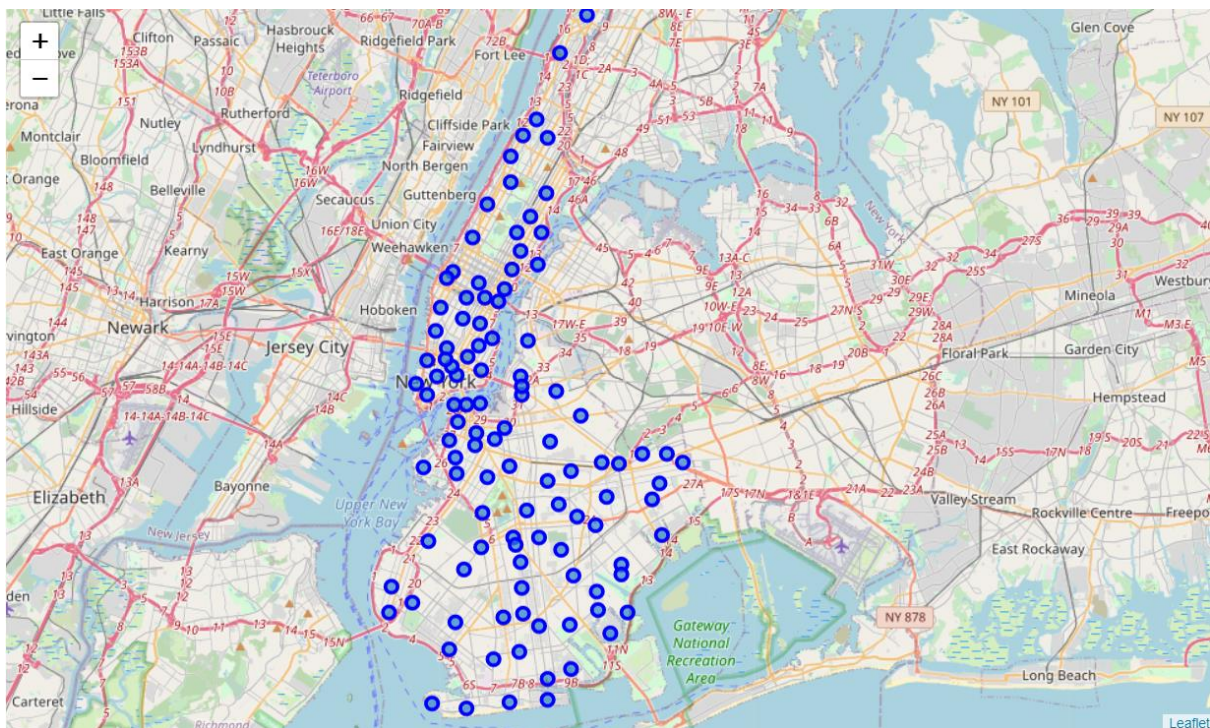
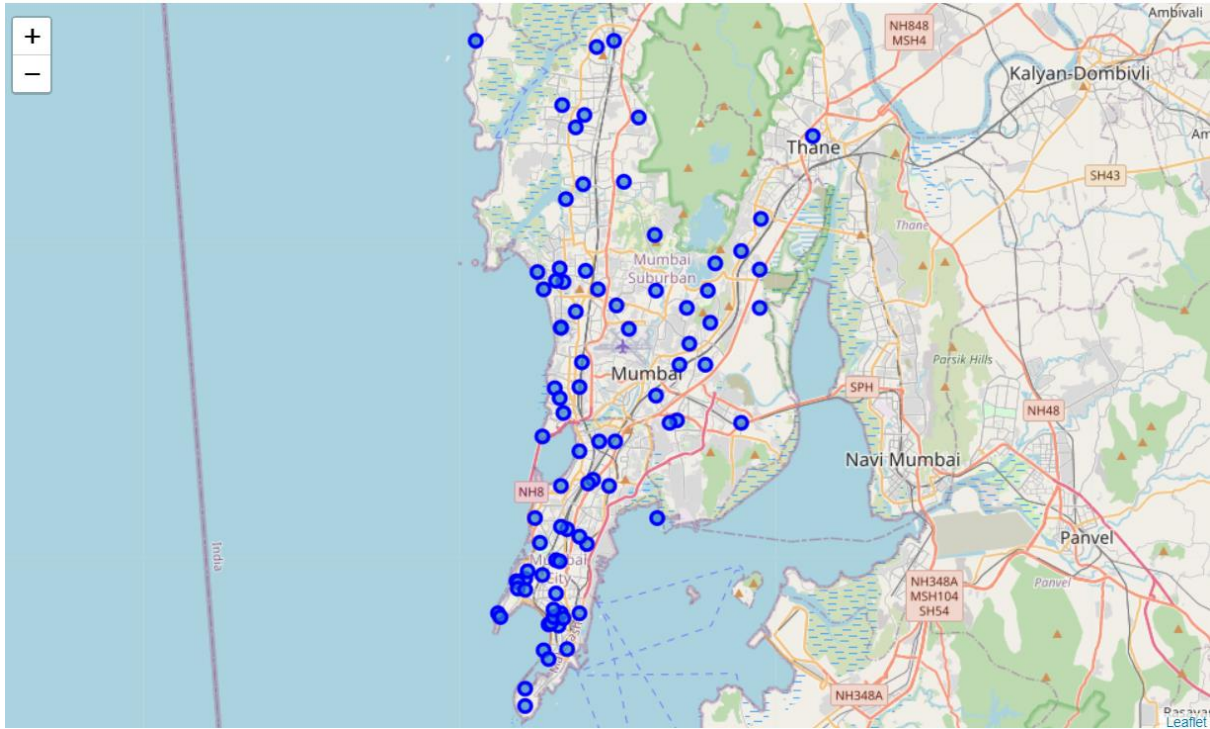
First, let's create a Dataframe which contains the list of Neighbourhoods in Mumbai with the respective geo coordinates. When done, it will look like this. Use `Dataframe.head()` function to view first 5 rows in the Dataframe.

	Neighbourhood	Location	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

Do the same for New York data.

	Borough	Neighbourhood	Latitude	Longitude
6	Manhattan	Marble Hill	40.876551	-73.910660
100	Manhattan	Chinatown	40.715618	-73.994279
101	Manhattan	Washington Heights	40.851903	-73.936900
102	Manhattan	Inwood	40.867684	-73.921210
103	Manhattan	Hamilton Heights	40.823604	-73.949688

Then, we will create a map of Mumbai and New York with Neighbourhoods superimposed on top. We will do this using **folium** library in Python. **folium** makes it easy to visualize data that's been manipulated in **Python** on an interactive leaflet map.



In the next step, we collect venues for each Neighbourhood and see which venues are the most common. In this step, the venue data was collected from Foursquare via API. After collecting the data and organising into a pandas dataframe, we have a Dataframe that looks like this (this is only a portion of the whole Dataframe).

	Neighbourhood	Location	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Amboli	Andheri, Western Suburbs	19.1293	72.8434	Cafe Arfa	19.128930	72.847140	Indian Restaurant
1	Amboli	Andheri, Western Suburbs	19.1293	72.8434	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
2	Amboli	Andheri, Western Suburbs	19.1293	72.8434	Subway	19.127860	72.844461	Sandwich Place
3	Amboli	Andheri, Western Suburbs	19.1293	72.8434	Cafe Coffee Day	19.127748	72.844663	Coffee Shop
4	Amboli	Andheri, Western Suburbs	19.1293	72.8434	Bhardawadi Ground	19.126143	72.843548	Park

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop
4	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop

Then, We will apply analytical methods to find the unique venue categories in the Mumbai and New York neighbourhoods and we will create a Dataframe that contains each of the neighbourhoods with the frequency of occurrence for each of the venue category.

From our analysis, we see that there are **176** unique venue categories in Mumbai neighbourhoods. Indian, Chinese, Thai, American, Spanish, Mediterranean, Deli restaurants, Burger joints, Tea shops, Cafes, Concert halls, theatres, Boutiques, Bowling Alleys, Bars, Flea markets, Harbors, Gourmet shops, Night Clubs, Pubs, Bagel shops, Pharmacies and Spas being some of them.

There are **381** unique venue categories in New York neighbourhoods. Coffee Shop, Pizza Place, Restaurants, Café, Cocktail Bar, Bakery being some of them.

Then we will get the top 10 venues in each of the neighbourhoods of Mumbai and New York.

	Neighbourhood	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	Aarey Milk Colony	Snack Place	Yoga Studio	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner	Dim Sum Restaurant	Dhaba
1	Agripada	Bakery	Coffee Shop	Indian Restaurant	Restaurant	Gym	Amphitheater	Dhaba	Fast Food Restaurant	Farmers Market	Falafel Restaurant
2	Altamount Road	Café	Coffee Shop	Sandwich Place	Pizza Place	Theater	Indian Restaurant	Bakery	Yoga Studio	Department Store	Electronics Store
3	Amboli	Coffee Shop	Indian Restaurant	Chinese Restaurant	Sandwich Place	Fast Food Restaurant	Park	German Restaurant	Donut Shop	Diner	Gourmet Shop
4	Amrut Nagar	Indian Restaurant	Café	Restaurant	Fast Food Restaurant	Asian Restaurant	Electronics Store	Park	Pizza Place	Bowling Alley	Falafel Restaurant

	Neighbourhood	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	Bath Beach	Chinese Restaurant	Pizza Place	Dessert Shop	Donut Shop	Bubble Tea Shop	Fast Food Restaurant	Italian Restaurant	Asian Restaurant	Gas Station	Pharmacy
1	Battery Park City	Park	Hotel	Gym	Coffee Shop	Shopping Mall	Memorial Site	Food Court	Sandwich Place	BBQ Joint	Plaza
2	Bay Ridge	Italian Restaurant	Greek Restaurant	Pizza Place	Spa	Bar	American Restaurant	Pharmacy	Hookah Bar	Playground	Chinese Restaurant
3	Bedford Stuyvesant	Coffee Shop	Café	Bar	Pizza Place	Deli / Bodega	Grocery Store	Tiki Bar	Thrift / Vintage Store	Cocktail Bar	Gift Shop
4	Bensonhurst	Chinese Restaurant	Italian Restaurant	Ice Cream Shop	Pizza Place	Sushi Restaurant	Donut Shop	Sporting Goods Shop	Pet Store	Grocery Store	Coffee Shop

Cluster the neighbourhoods in Mumbai and New York based on the similarity of top common venues.

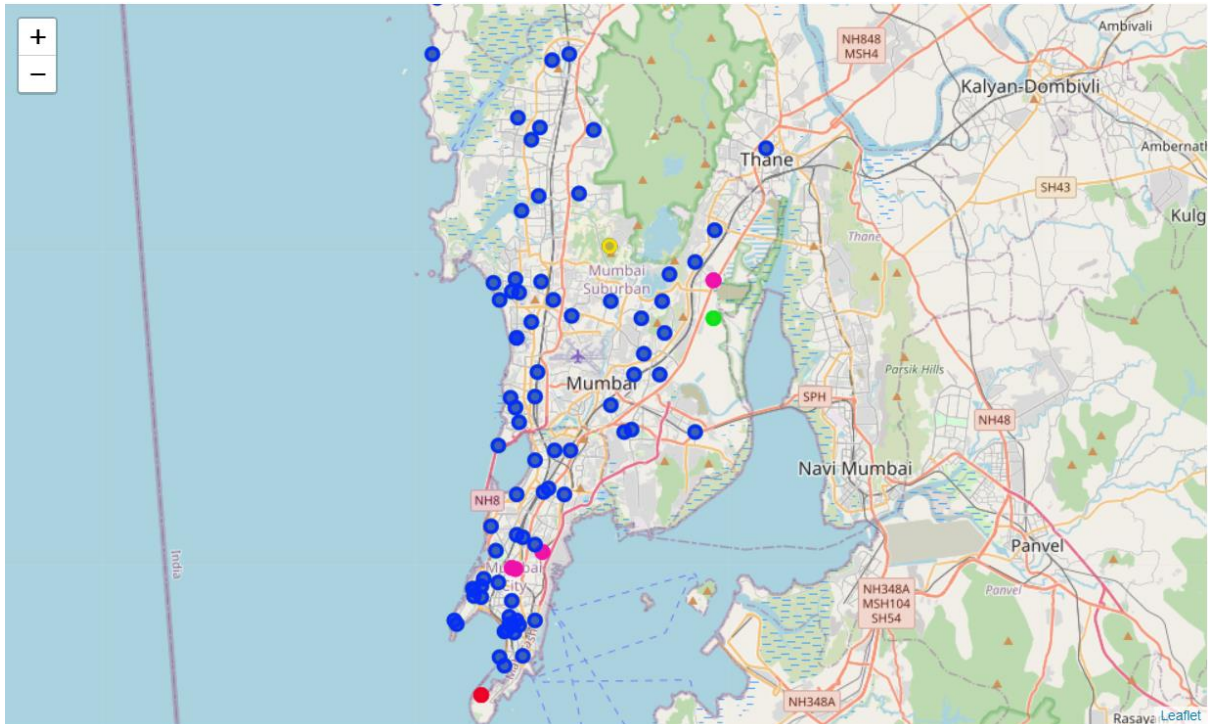
Let us now apply a clustering algorithm to group the neighbourhoods based on the similarity in types of venues they have. By clustering, we also provide information to users on a common type of neighbourhood. We will use the k-Means clustering approach to cluster the neighbourhoods. k will be selected as 6.

This means that we will group the neighbourhoods into 6 clusters. Each of the neighbourhoods gets a Cluster Label assigned.

	Neighbourhood	Location	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
0	Amboli	Andheri, Western Suburbs	19.129300	72.843400	0.0	Coffee Shop	Indian Restaurant	Chinese Restaurant	Sandwich Place	Fast Food Restaurant	Park	German Restaurant
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833	0.0	Hotel	Multiplex	Fast Food Restaurant	Restaurant	Wine Shop	Gym / Fitness Center	Seafood Restaurant
2	D.N. Nagar	Andheri, Western Suburbs	19.124085	72.831373	0.0	Gym / Fitness Center	Cocktail Bar	Pizza Place	Chinese Restaurant	Snack Place	Indian Restaurant	Light Rail Station
3	Four Bungalows	Andheri, Western Suburbs	19.124714	72.827210	0.0	Electronics Store	Chinese Restaurant	Pizza Place	Restaurant	Residential Building (Apartment / Condo)	Juice Bar	Fish Market

	Borough	Neighbourhood	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
0	Manhattan	Marble Hill	40.876551	-73.910660	2	Coffee Shop	Discount Store	Gym	Yoga Studio	Tennis Stadium	Pizza Place	Diner
1	Manhattan	Chinatown	40.715618	-73.994279	0	Chinese Restaurant	Bakery	Cocktail Bar	Vietnamese Restaurant	Optical Shop	Salon / Barbershop	American Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	0	Café	Bakery	Bank	Mobile Phone Shop	Grocery Store	Supermarket	Sandwich Place
3	Manhattan	Inwood	40.867684	-73.921210	4	Mexican Restaurant	Lounge	Café	Restaurant	Park	Frozen Yogurt Shop	Bakery
4	Manhattan	Hamilton Heights	40.823604	-73.949688	2	Pizza Place	Deli / Bodega	Coffee Shop	Café	Mexican Restaurant	Yoga Studio	Bakery

We will then use the Mumbai dataset with cluster labels assigned to visualize the cluster in **folium** map.



A piece of important information this map provides is that many neighbourhoods in Mumbai are of similar nature concerning the venues they have around, indicated by the cluster marked in blue.

Let us now dig a little deeper into how the neighbourhoods are clustered and what is the characteristic of the cluster that is very common across most neighbourhoods in Mumbai.

Cluster Label 0

	Neighbourhood	Location	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	
0	Ambohi	Andheri, Western Suburbs	19.129300	72.843400	0.0	Coffee Shop	Indian Restaurant	Chinese Restaurant	Sandwich Place	Fast Food Restaurant	Park	German Restaurant	[
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833	0.0	Hotel	Multiplex	Fast Food Restaurant	Restaurant	Wine Shop	Gym / Fitness Center	Seafood Restaurant	C
2	D.N. Nagar	Andheri, Western Suburbs	19.124085	72.831373	0.0	Gym / Fitness Center	Cocktail Bar	Pizza Place	Chinese Restaurant	Snack Place	Indian Restaurant	Light Rail Station	
3	Four Bungalows	Andheri, Western Suburbs	19.124714	72.827210	0.0	Electronics Store	Chinese Restaurant	Pizza Place	Restaurant	Residential Building (Apartment / Condo)	Juice Bar	Fish Market	S

This is the biggest cluster containing 81 Neighbourhoods. These Neighbourhoods are the Blue coloured dots in the map. The Neighbourhoods belonging to this cluster are popular for having Indian Restaurant, Chinese Restaurant, Pub, Fast Food Restaurant, Bar, Recreation Centre, Market, Multiplex, Train/Bus Station. These Neighbourhoods may interest people having diverse food choices starting from Indian, Chinese, Seafood, Falafel, Dumpling and Mediterranean. We see that this neighbourhood would be something that would be interesting to those who depend more on the public commute since these neighbourhoods are closer to train and bus stations. There are many Pubs, Nightclubs, Cafes, Bars which may attract people who like to Party a lot. Cluster 1, 2, 3 are having very few neighbourhoods.

Cluster Label 1

Neighbourhood	Location	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
Worli	Goregaon, Western Suburbs	19.148493	72.881756	1.0	Snack Place	Yoga Studio	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner	Dim Sum Restaurant	Dhaba

Very few neighbourhoods belong to this cluster, making this unique. They are the Yellow dots in the map. As you can see they are far away from the main city. They are popular for Event Space, Snack Place, Yoga Studio, Diner. This neighbourhood is suitable for middle class families with low budget for buying/renting a house.

Cluster Label 2

Neighbourhood	Location	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
Vikhroli	Eastern Suburbs	19.11	72.94	2.0	ATM	Women's Store	Farmers Market	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner	Dim Sum Restaurant	Dhaba

The Neighbourhoods belonging to this Cluster are popular for having ATM, Event Space and Stores. These Neighbourhoods are represented by green coloured circles on the above map. This cluster is similar to cluster 1.

Cluster Label 3

Neighbourhood	Location	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
Cuffe Parade	South Mumbai	18.91	72.81	3.0	Garden	Yoga Studio	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner	Dim Sum Restaurant	Dhaba

Only one Neighbourhood in this cluster. It is represented by red dot in the map. It is unique because it is located in far south of the city. This area contains Gardens, Donut Shop, and Restaurants.

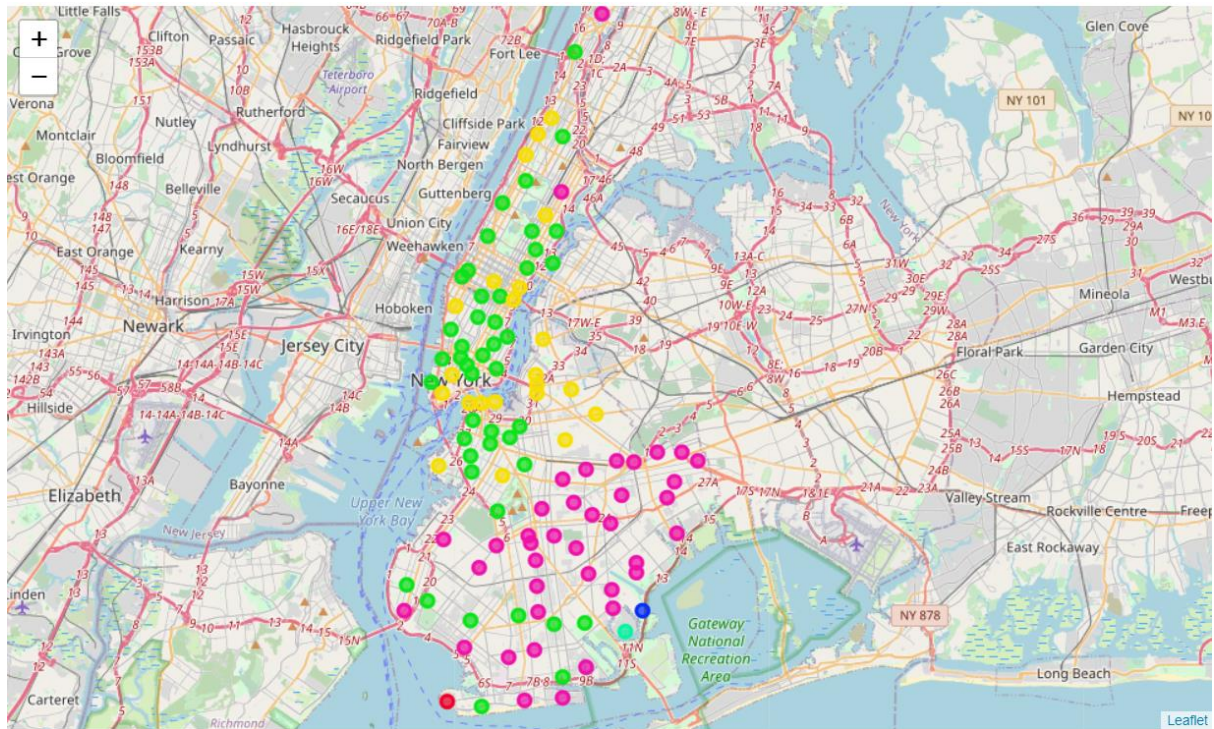
Cluster Label 4

	Neighbourhood	Location	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
40	Kanjurmarg	Eastern Suburbs	19.130000	72.940000	4.0	Multiplex	Bakery	Yoga Studio	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
52	Agripada	South Mumbai	18.977700	72.827300	4.0	Bakery	Coffee Shop	Indian Restaurant	Restaurant	Gym	Amphitheater	Dhaba	Fast Food Restaurant
59	Cotton Green	South Mumbai	18.986209	72.844076	4.0	Hookah Bar	Clothing Store	Plaza	Bakery	Yoga Studio	Department Store	Falafel Restaurant	Event Space
80	Dagdi Chawl	Byculla, South Mumbai	18.977129	72.829131	4.0	Indian Restaurant	Restaurant	Bakery	Yoga Studio	Dessert Shop	Falafel Restaurant	Event Space	Electronics Store

This cluster is containing 4 Neighbourhood. Neighbourhoods in this cluster are popular for having Bakery, Clothing Store, Multiplex, Indian Restaurant and Event Space. This cluster is represented by pink coloured dots on the map of Mumbai.

Since, the objective of this study is to compare the neighbourhoods between the two cities of Mumbai and New York, and not really to compare neighbourhoods within Mumbai, we will put forth our conclusion from the study after doing a similar analysis on the neighbourhoods in New York.

Then use the New York dataset with cluster labels assigned to visualize the cluster in **folium** map.



Let us now examine the clusters.

Cluster Label 0

	Borough	Neighbourhood	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
1	Manhattan	Chinatown	40.715618	-73.994279	0	Chinese Restaurant	Bakery	Cocktail Bar	Vietnamese Restaurant	Optical Shop	Salon / Barbershop	American Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	0	Café	Bakery	Bank	Mobile Phone Shop	Grocery Store	Supermarket	Sandwich Place
6	Manhattan	Central Harlem	40.815976	-73.943211	0	African Restaurant	Cosmetics Shop	French Restaurant	American Restaurant	Bar	Chinese Restaurant	Seafood Restaurant
8	Manhattan	Upper East Side	40.775639	-73.960508	0	Italian Restaurant	Bakery	Coffee Shop	Gym / Fitness Center	Yoga Studio	Exhibit	Spa
9	Manhattan	Yorkville	40.775930	-73.947118	0	Italian Restaurant	Gym	Bar	Coffee Shop	Sushi Restaurant	Mexican Restaurant	Deli / Bodega
10	Manhattan	Lenox Hill	40.768113	-73.958880	0	Italian	Coffee	Pizza	Sushi	Cocktail Bar	Café	Burger Joint

This is the Largest Cluster containing 45 Neighbourhoods in it. It is represented by green dots in the map. These Neighbourhoods are popular for having Chinese Restaurant, Mexican Restaurant, Vietnamese Restaurant, American Restaurant, Italian Restaurant and many more types of Restaurants, Bakery, and Cafes. We can see that these Neighbourhoods are a great place for foodies.

Cluster Label 1

	Borough	Neighbourhood	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
85	Brooklyn	Bergen Beach	40.61515	-73.898556	1	Harbor / Marina	Baseball Field	Playground	Athletics & Sports	Fish Market	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

This Cluster contains only one neighbourhood. It is represented by blue dot. This neighbourhood is popular for Harbor, Playground, Baseball, Athletics & Sports. This place can interest people who like to play outdoor sports. Families with kids will be interested in this neighbourhood.

Cluster Label 2

Neighbourhood	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
Marble Hill	40.876551	-73.910660	2	Coffee Shop	Discount Store	Gym	Yoga Studio	Tennis Stadium	Pizza Place	Diner	Donut Shop	Pharmacy	Sandwich Place
Hamilton Heights	40.823604	-73.949688	2	Pizza Place	Deli / Bodega	Coffee Shop	Café	Mexican Restaurant	Yoga Studio	Bakery	Latin American Restaurant	Park	Cocktail Bar
Manhattanville	40.816934	-73.957385	2	Coffee Shop	Seafood Restaurant	Deli / Bodega	Bar	Italian Restaurant	Mexican Restaurant	Chinese Restaurant	Café	Boutique	Sushi Restaurant
Midtown	40.754691	-73.981669	2	Coffee Shop	Hotel	Clothing Store	Bakery	Theater	Sporting Goods Shop	Sandwich Place	Bookstore	Steakhouse	Gym
Chelsea	40.744035	-74.003116	2	Coffee Shop	Art Gallery	American Restaurant	Bakery	Seafood Restaurant	Hotel	Italian Restaurant	Nightclub	Bookstore	Market
Morningside Heights	40.808000	-73.963896	2	Park	American Restaurant	Bookstore	Coffee Shop	Deli / Bodega	Burger Joint	Sandwich Place	Ice Cream Shop	Farmers Market	Tennis Court
Financial District	40.707107	-74.010665	2	Coffee Shop	Pizza Place	Cocktail Bar	Gym / Fitness Center	Steakhouse	Italian Restaurant	Salad Place	Bar	American Restaurant	Café
Carnegie Hill	40.782683	-73.953256	2	Coffee Shop	Café	Yoga Studio	Gym / Fitness Center	Cosmetics Shop	Bookstore	Pizza Place	Bakery	French Restaurant	Gym
Civic Center	40.715229	-74.005415	2	Coffee Shop	Hotel	Cocktail Bar	Gym / Fitness Center	Yoga Studio	Spa	Park	French Restaurant	Italian Restaurant	Baker
Turtle Bay	40.752042	-73.967708	2	Sushi Restaurant	Coffee Shop	Café	Italian Restaurant	Park	Japanese Restaurant	French Restaurant	Seafood Restaurant	Deli / Bodega	Wine Bar

Neighbourhoods in this cluster are represented by yellow dots on the map. It contains 23 Neighbourhoods. This cluster is unique because it is having Park, Bar, Nightclub, Gym, and Yoga Studio. These Neighbourhoods may interest people who party a lot.

Cluster Label 3

Borough	Neighbourhood	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
Brooklyn	Sea Gate	40.576375	-74.007873	3	Bus Station	Beach	American Restaurant	Spa	Yoga Studio	Fishing Spot	Farm	Farmers Market	Fast Food Restaurant	F

This cluster is having only one neighbourhoods and the venues are very unique than other clusters. This Neighbourhood is popular for having Bus Station, Beach, Fishing Spot, and Markets.

This neighbourhoods is a good place for those who prefer Public Transport.

Cluster Label 4

	Neighbourhood	Location	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
40	Kanjurmarg	Eastern Suburbs	19.130000	72.940000	4.0	Multiplex	Bakery	Yoga Studio	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
52	Agripada	South Mumbai	18.977700	72.827300	4.0	Bakery	Coffee Shop	Indian Restaurant	Restaurant	Gym	Amphitheater	Dhaba	Fast Food Restaurant
59	Cotton Green	South Mumbai	18.986209	72.844076	4.0	Hookah Bar	Clothing Store	Plaza	Bakery	Yoga Studio	Department Store	Falafel Restaurant	Event Space
80	Dagdi Chawl	Byculla, South Mumbai	18.977129	72.829131	4.0	Indian Restaurant	Restaurant	Bakery	Yoga Studio	Dessert Shop	Falafel Restaurant	Event Space	Electronics Store

This cluster is represented by pink dots on the map. All the neighbourhoods in this cluster are in Brooklyn.

These Neighbourhoods are popular for having Bank, Pharmacy, Museum, Park, Grocery Store, Metro Station, Bus Station, Bookstore, Food, Supermarket.

This should be good for those who have children, since the venues close to these neighbourhoods are great to keep the children engaged.

Conclusion

In this Data Science Project, we have attempted to load the dataset for Mumbai (India) and New York (US). We have tried to analyse the neighbourhood regions in these metro cities based on the type of popular and top venues they have. We have clustered the neighbourhoods using K-means Clustering based on the most common top venues in each of the neighbourhood. Our intention with this project was to analyse and understand the difference in the type of life in these metro cities, which can offer decision points for anybody who is considering to settle in either of the metro cities and can get a peek into what type of experience and facilities he/she will be provided with.

Given our cluster information for both Mumbai and New York, we see that both Mumbai and New York neighbourhoods are a great place for a foodie. However, New York has more variety in food with immigrants making up the majority of the residents in some neighbourhoods.

Also due to the proximity of both the cities to the seashore, they offer for harbors, seafood, boat, and ferry rides.

Mumbai is a substantially more religious than New York. If you're a religious person (I'm not), then you've got a lot more avenues to indulge in relevant activities and mingle with like-minded people.

Also, with low cost of living Mumbai may interest people who are middle class and above.