Mumbai & Delhi

Let the Data find the similitude!

The study Entails the Data analysis of various factors for two cities like food, venues to visit,neighbourhood etc.

BUSINESS PROBLEM

Mumbai and Delhi are the two major metro cities in India. There has been a war for supremacy in terms of quality of life, jobs, education, entertainment and recreational facilities that these cities have to offer to its residents. This data science

project attempts to analyze the neighborhoods in each of these two cities and tries 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 metro cities.

The deciding factor for most would be on how lively, supportive, vibrant and unique each of the cities can be when compared to each other.

The business problem in this study assumes that people who would be interested in this study are those who would like to create a projection of potential life and activities in these metro city neighborhoods if the subject moves to live in one of them. The decision to choose one over the other would depend on popular venues in the neighborhoods in each of these metro cities.

DATA

For any “data science project” data is of paramount importance. For this study, we needed data about neighborhoods in each of these metro cities. The neighborhoods with their respective districts and postal codes for all would serve us well for this study.

For this project we need the following data :

- New Delhi neighborhood data that contains list Locality, districts, pin codes.
- Data source : "https://www.movingsolutions.in/blog/2020/02/05/pin-codes -of-delhi-locations

- Mumbai neighborhood data that contains list Locality, districts, pin codes
- Data source: "https://www.mapsofindia.com/pincode/india/maharashtra/mumbai

Description : we will parse the required information from the above said website. And we will use this data set to explore various locality of new delhi city and mumbai

- Nearby places in each locality of new delhi city and mumbai
- Data source : Fousquare API
- Description : By using this api we will get all the venues in each neighborhood of both the cities.

Approach

- Collect the new delhi and Mumbai city data from above mentioned websites .
- Using FourSquare API we will find all venues for each neighborhood in both the cities
- Filter out all venues that are nearby by locality.
- Using aggregative rating for each coomon venue to find the best places and sort them accordingly.

For any “data science project” data is of paramount importance. For this study, we needed data about neighborhoods in each of these metro cities.

In this study, we will pass the websites containing the various pin codes of delhi and Mumbai along with their neighborhood, read it into a pandas Dataframe and curate it to remove the data related to all other cities, towns, and places which are not

Mumbai or Delhi, since we are only interested in comparing these two biggest metro cities in India.

We shall then clean up the unnecessary columns , which is not relevant or useful for our current study. Post office names (***office name***) will be used as the neighborhood names in each of the regions such as Mumbai or Delhi.

Neighborhood names with the same ***Pincode*** will be combined as a single row.

Foursquare API will be used to find the longitude and latitude of each of the neighborhoods in both Mumbai and Delhi. This will form the dataset we will use for this study.

The first few records of the dataset we now have after cleanup and curation appear as below.

:

	Neighborhood	Pincode	District
1	A I staff colony	400029	Mumbai
2	Aareymilk Colony	400065	Mumbai
3	Agripada	400011	Mumbai
4	Airport	400099	Mumbai
5	Ambewadi	400004	Mumbai

Same is being done for neighborhoods of Delhi as shown below.

	Neighborhood	Pincode	Borough
0	Anand Vihar	110092	East Delhi
1	Azad Nagar	110051	East Delhi
2	Babarpur	110032	East Delhi
3	Balbir Nagar	110032	East Delhi
4	Bhajan Pura	110053	East Delhi

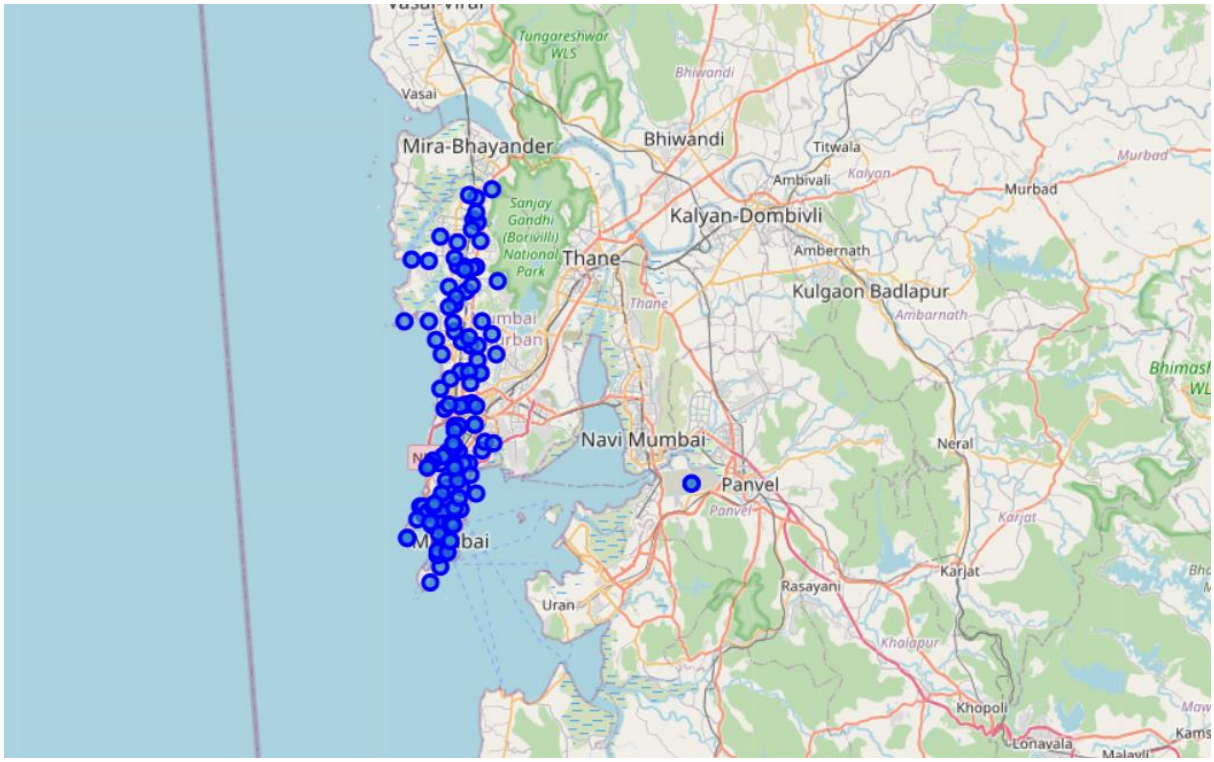
The next step is to enhance the dataset with the required information. We would need the longitude and latitude values for the neighborhoods. We will use the *Nominatim* library from *geocoders.geopy* package to find the longitude and latitude for each of the neighborhoods and would eventually create a dataset having all the necessary columns for our analysis.

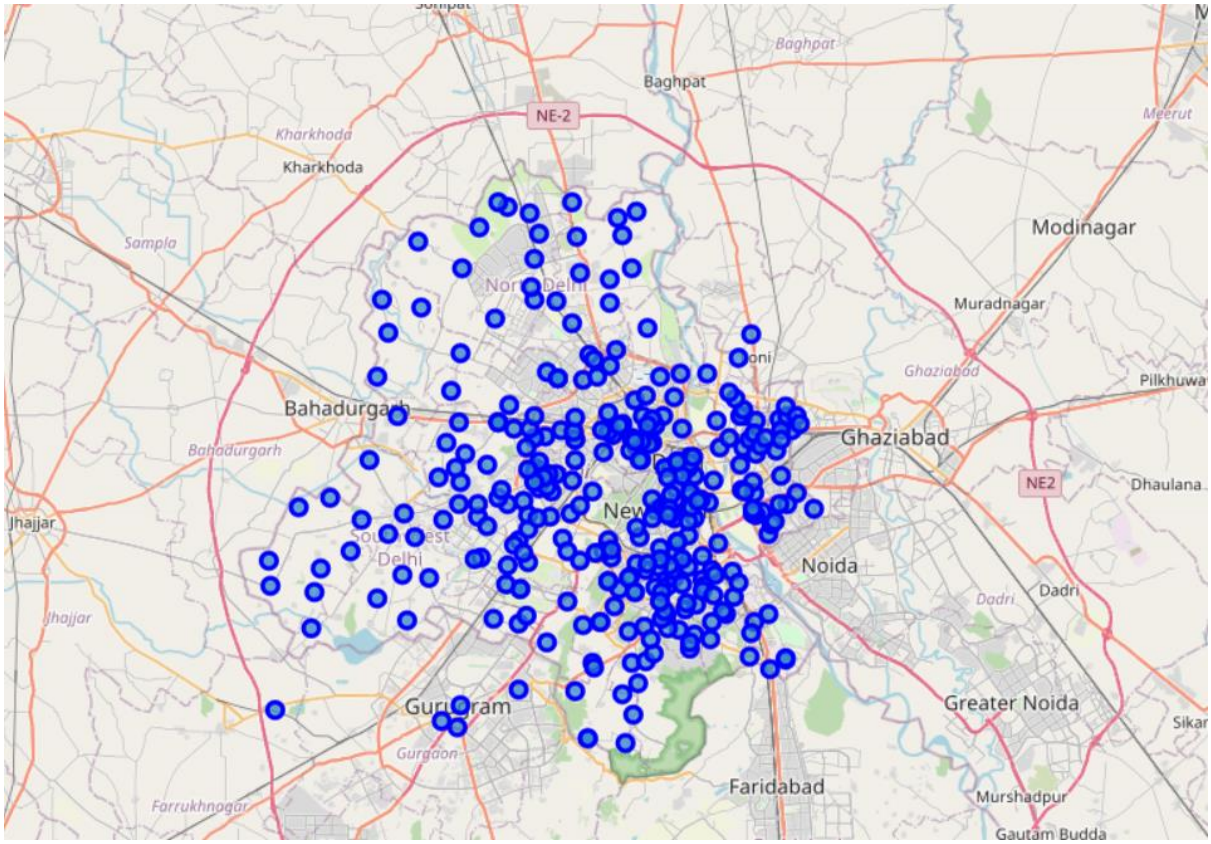
	Neighborhood	Pincode	District	Latitude	Longitude
3	Agripada	400011	Mumbai	18.975302	72.824898
4	Airport	400099	Mumbai	19.090201	72.863808
5	Ambewadi	400004	Mumbai	19.186776	72.859313
6	Andheri	400053	Mumbai	19.119698	72.846420
7	Andheri East	400069	Mumbai	19.115883	72.854202

Curating the same for the neighborhoods in Delhi

	Neighborhood	Pincode	Borough	Latitude	Longitude
0	Anand Vihar	110092	East Delhi	28.641115	77.312502
1	Azad Nagar	110051	East Delhi	28.662682	77.279515
2	Babarpur	110032	East Delhi	28.687431	77.279755
3	Balbir Nagar	110032	East Delhi	28.683790	77.290754
5	Bhola Nath nagar	110032	East Delhi	28.669127	77.285241
7	Chilla	110091	East Delhi	28.595984	77.300936
9	Dilshad Garden	110095	East Delhi	28.675826	77.321516
11	G.T.B. Hospital	110095	East Delhi	28.684560	77.311077
12	Gandhi Nagar	110031	East Delhi	28.453175	77.015329
13	Gandhi Nagar bazar	110031	East Delhi	28.656097	77.235436
14	Geeta Colony	110031	East Delhi	28.650101	77.275921
15	Ghazipur	110096	East Delhi	28.628183	77.325442

We now will join the neighborhoods of both the cities and have the necessary information to visualize the neighborhoods for both the cities on a *folium* map.





Analyzing the neighborhoods

We will use the Foursquare API to find the top venues in the neighbourhoods of Mumbai and Delhi. This will help us understand the nature of life the neighborhoods have to offer in both of the Metropolitan city.

We will iteratively make Foursquare API calls for each of the neighborhoods in our dataset.

Next, we will employ statistically and analytical methods to find the unique venues/venue categories in the Mumbai and Delhi neighborhoods and we will build a Dataframe that calibrates each of the neighborhoods with the frequency of occurrence for each of the venue category.

From our analysis, we see that there are **218** unique venue categories in Mumbai and Delhi neighborhoods.

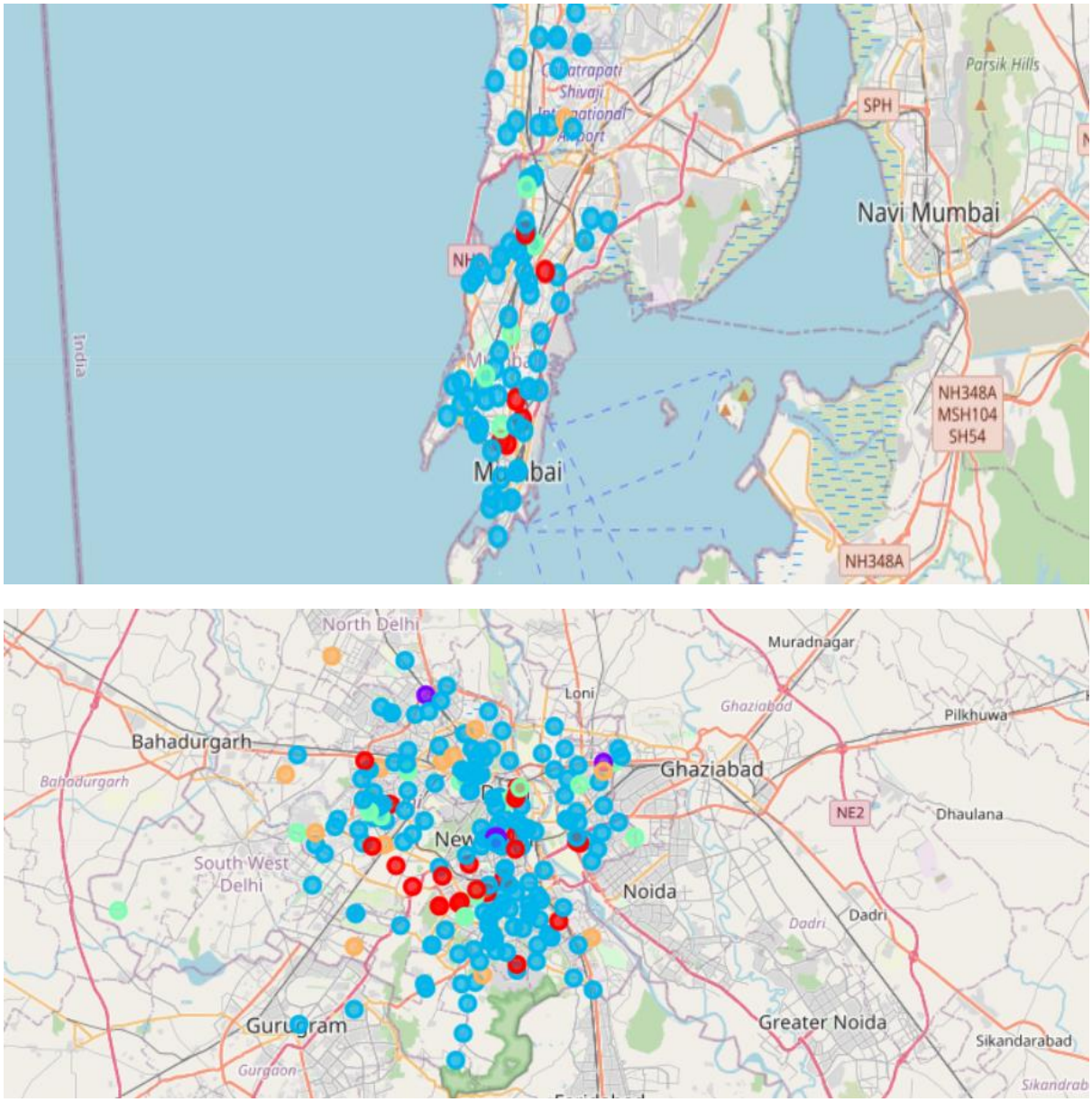
Atm , accessories store, Afgan restaurant, Airport,american restaurant , boutique shop, arcade, art gallery, Asian restaurant, auto workshop are some of them.

We then create a dataset that lists the top 5 common venues against each of the neighbourhoods in both of the cities . We get a representation such as below for all the neighbourhoods.

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

Given that we now have the required information regarding the top venues against each of the neighborhoods in Mumbai and Delhi , let us now apply a clustering algorithm to group the neighborhoods based on the similarity in types of venues they have. By clustering, we also provide information to users on a common type of neighbourhoods in both the cities. We will use the k-Means clustering approach to cluster the neighbourhoods. k will be selected as 5. This means that we will group the neighborhoods into 5 clusters. Each of the neighborhoods gets a Cluster Label assigned.

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



A piece of important information this map provides is that many neighborhoods in Mumbai and Delhi 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 neighborhoods are clustered and what is the characteristic of the cluster that is very common across most neighborhoods in both cities.

Cluster Label 0

	Longitude	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
5	72.859313	Women's Store	Indian Restaurant	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm
34	72.837053	Indian Restaurant	Flea Market	Pizza Place	Café	Discount Store	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm
65	72.830243	Indian Restaurant	Food Truck	Cheese Shop	Market	Yoga Studio	Duty-free Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant
72	72.834717	Indian Restaurant	Hookah Bar	Bakery	Yoga Studio	Donut Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market
78	72.827938	Indian Restaurant	Fast Food Restaurant	Yoga Studio	Discount Store	Flea Market	Fish & Chips Shop	Field	Farmers Market	Farm	Falafel Restaurant
102	72.856318	Indian Restaurant	Yoga Studio	Discount Store	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
122	72.847871	Hockey Arena	Indian Restaurant	Department Store	Yoga Studio	Donut Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market
139	72.838343	Indian Restaurant	Bakery	Chinese Restaurant	Playground	Café	Ice Cream Shop	Breakfast Spot	Restaurant	Arcade	Event Space

The neighborhoods belonging to this cluster is popular for having Indian restaurants, Cafes, markets,yoga studio and vegetarian joints. We see that this neighborhood would be something that a subsection of Indians would prefer if they want a scaled-down lifestyle with close to home vegetarian food.

Cluster Label 1

	Longitude	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
23	77.312395	Metro Station	Cafeteria	Yoga Studio	Donut Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm
245	77.213691	Metro Station	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
259	77.211989	Metro Station	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
266	77.212282	Metro Station	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
267	77.214397	Metro Station	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
296	77.149403	Metro Station	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant

The neighborhoods belonging to this cluster is popular for having Metro station, Donut shop,Flower shop, Flea market, and fast food joints. We see that this neighborhood would be something that would be interesting to those who depend more on the public commute since these neighborhoods are closer to metro station.

Cluster Label 2

	Longitude	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
3	72.824898	Coffee Shop	Bakery	Garden	Pool	Hostel	Airport Food Court	Airport Gate	Flower Shop	Flea Market	Fish & Chips Shop
4	72.863808	Airport	Airport Service	Jewelry Store	Yoga Studio	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant
7	72.854202	Indian Restaurant	Hotel	Shopping Mall	Light Rail Station	Electronics Store	Camera Store	Smoke Shop	Chinese Restaurant	Cricket Ground	Event Space
9	72.865256	Indian Restaurant	Trail	Gym / Fitness Center	Grocery Store	Yoga Studio	Discount Store	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market
11	72.840038	Park	Indian Restaurant	Fast Food Restaurant	Metro Station	Bar	Lake	Snack Place	Ice Cream Shop	Hardware Store	Discount Store
12	72.861599	Tea Room	Yoga Studio	Food	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
14	100.340014	Café	Indian Restaurant	Hostel	Halal Restaurant	Coffee Shop	Vegetarian / Vegan Restaurant	Bakery	Public Art	Hotel	Burger Joint
15	72.830267	Café	Indian Restaurant	Asian Restaurant	Smoke Shop	Dessert Shop	Arcade	Italian Restaurant	Bagel Shop	Chinese Restaurant	Coffee Shop

The neighborhoods belonging to this cluster is popular for having a mix of Indian and Chinese restaurants, Cafe, Coffee shops etc.Also, these neighborhoods may interest people who have diverse food choices starting from Indian, Asian, Chinese, Italian to having Snacks, donutshops. These neighborhoods also provide for some recreational places such as Gyms, Parks, Bowling Alleys, Arcade, and Harbours.

Cluster Label 3

	Longitude	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
6	72.846420	Fast Food Restaurant	Food Court	Indian Restaurant	Bakery	Restaurant	Yoga Studio	Donut Shop	Fish & Chips Shop	Field	Farmers Market
8	72.846420	Fast Food Restaurant	Food Court	Indian Restaurant	Bakery	Restaurant	Yoga Studio	Donut Shop	Fish & Chips Shop	Field	Farmers Market
29	72.861808	Hotel	Fast Food Restaurant	Café	Discount Store	Flea Market	Fish & Chips Shop	Field	Farmers Market	Farm	Falafel Restaurant
35	72.832672	Multicuisine Indian Restaurant	Fast Food Restaurant	Platform	Train Station	Yoga Studio	Field	Farmers Market	Farm	Falafel Restaurant	Event Space
41	72.842876	Fast Food Restaurant	Flower Shop	Vegetarian / Vegan Restaurant	Breakfast Spot	Plaza	Indian Restaurant	Bar	Women's Store	Movie Theater	Farmers Market
74	72.864942	Gym / Fitness Center	Fast Food Restaurant	BBQ Joint	Yoga Studio	Donut Shop	Flea Market	Fish & Chips Shop	Field	Farmers Market	Farm
95	72.827395	Ice Cream Shop	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Dessert Shop	Yoga Studio	Donut Shop	Flea Market	Fish & Chips Shop	Field	Farmers Market

The main attraction in this neighborhood seems to be its proximity to Fast food restaurant, vegan restaurant, bakery,gym/fitness center, bar, yoga studio.

Cluster Label 4

	Longitude	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
20	72.856474	ATM	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
65	77.312267	ATM	Juice Bar	Discount Store	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
70	77.019182	ATM	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
71	77.097591	ATM	Juice Bar	Discount Store	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
87	77.105571	ATM	Pizza Place	Diner	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space
149	77.046691	ATM	Café	Discount Store	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
161	77.300518	ATM	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
193	77.197626	ATM	Ice Cream Shop	Discount Store	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
281	77.061882	ATM	Donut Shop	Flower Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
291	77.163475	ATM	Burger Joint	Donut Shop	Flea Market	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant

The main attraction in this neighborhood seems to be its proximity to ATM, fast food restaurant, Farmers Market, farm, Fish & chips shops, falafel restaurant.

Study findings & conclusion

In this project, we have attempted to load the dataset for two of India's prime metro cities and have tried to analyze the neighborhood regions in these metro cities based on the type of popular and top venues they have. We have clustered the neighborhoods based on the most common top venues in each of the neighborhood. Our intention with this project was to analyze and understand the difference in the type of life in these metros, 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 will be provided with.

Given our cluster information for both Mumbai and Delhi, we see that the clustering provides the group of venues that has more or less same kind of neighborhood in both of the cities. Also the analysis provides a insights for the audience who is

looking for same kind of neighborhood in Mumbai given its location in Delhi or vice-versa.