

Capstone Project - The Battle of Neighborhoods (Week 2)

Where should I open a European-style Vegetarian
restaurant in India?

Business Problem

- A customer wants to open a new European style vegetarian restaurant in Delhi and Mumbai in India.
- Since India is a diverse country accommodating both foreign and local communities, my customer has asked me to help him in finding the best location for his first restaurant of his food chain in Delhi and Mumbai.
- Since Delhi is the capital of India and Mumbai being the financial capital, it attracts both foreigners and nationals to establish their business, or to study or for tourism.
- Delhi itself has 7 districts with 73 locations and Mumbai has 12 districts with 114 locations. We aim to find a location with good amount of people and low competition.

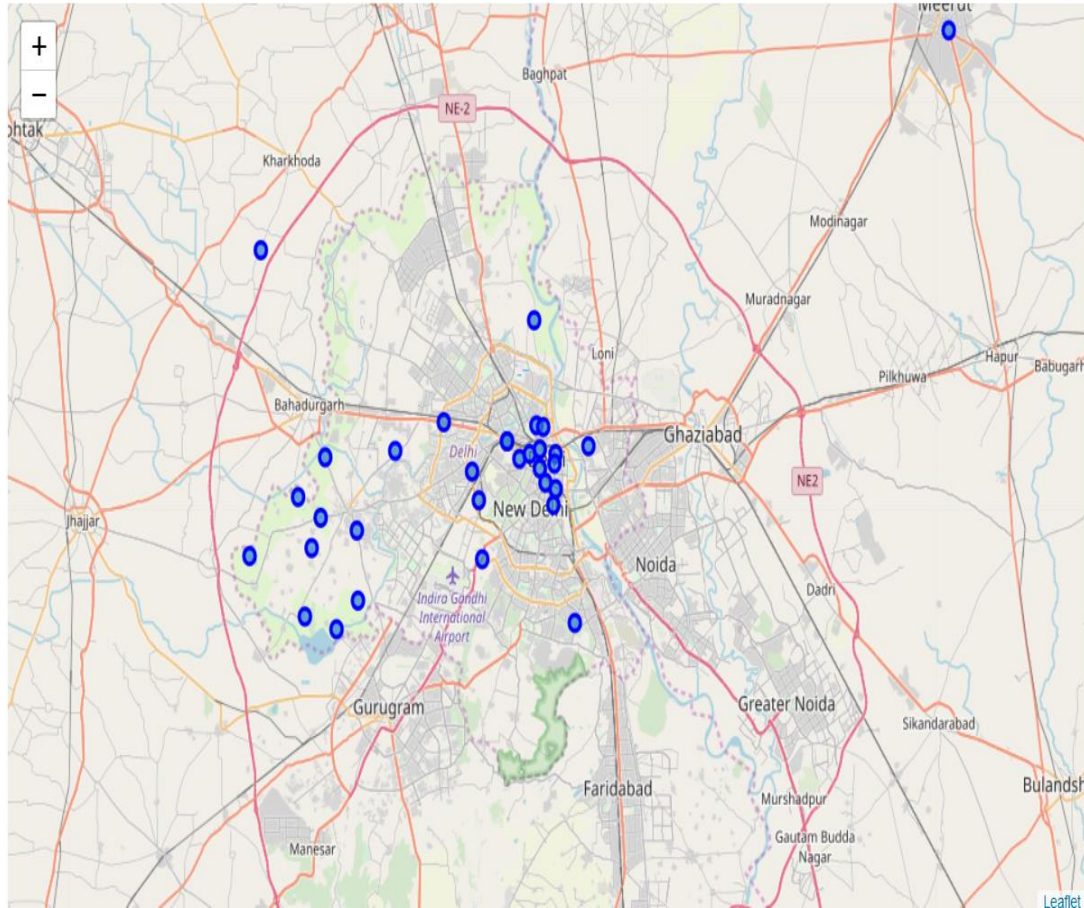
Data

- In order to find a good location we will need to analyse the neighbourhood of Delhi and Mumbai. The following csv file can be used for this:
 - <https://github.com/pareshprakash/capstone-project/Delhi-District.csv>
 - <https://github.com/pareshprakash/capstone-project/blob/master/Mumbai-District.csv>
- The Four Square API can be used to compare the neighborhood of Delhi and Mumbai and based on the comparison we can choose a location for the restaurant.
- We need to find out localities which are full of convenience stores, restaurants, train station, airport terminals.

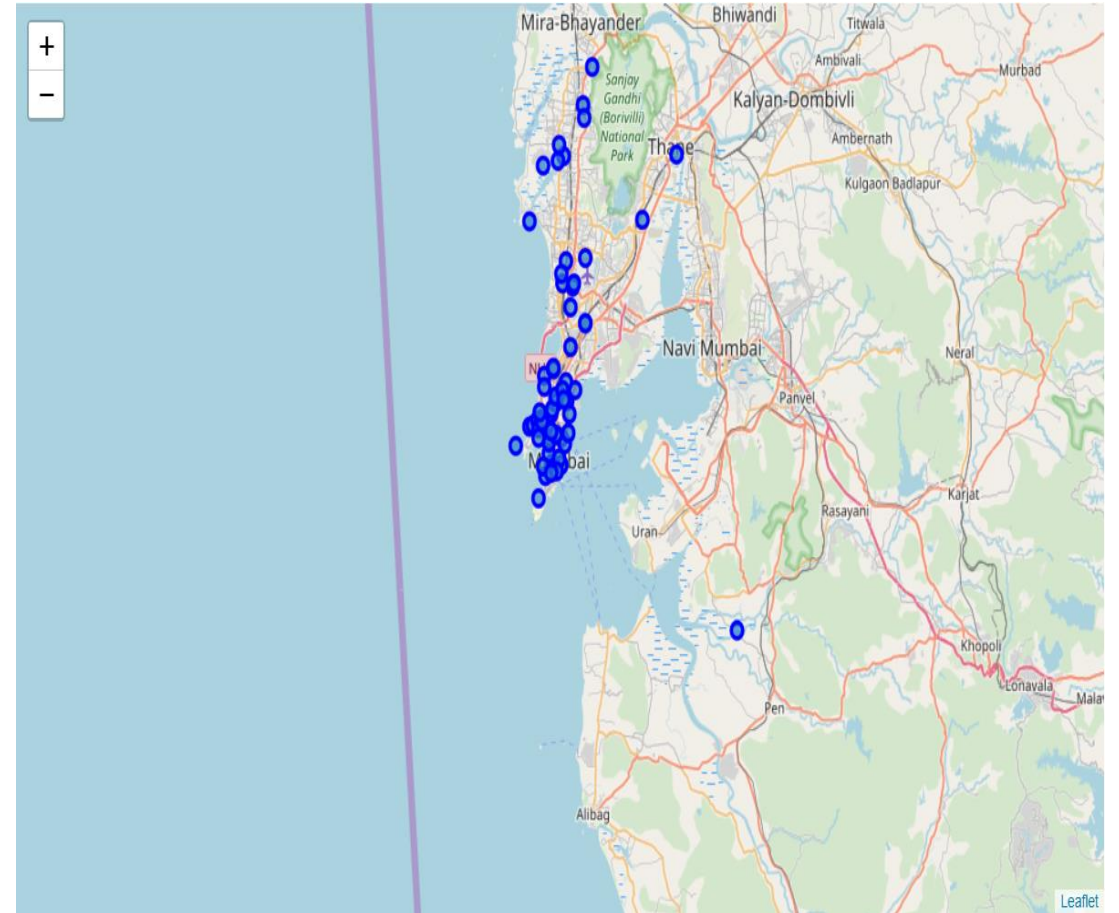
Methodology

- For each location in Delhi and Mumbai, all convenience stores, restaurants, train station data was found out using the Four Square API.
- The number or ratio of each of them was then computed.
- The data was then arranged to find out the most common venue in each of the districts arranged in descending order.
- K-means clustering was used to analyse the neighbourhood and find out the location with the highest number of the above locations.
- Based on the clusters, the location with the maximum number of people can be identified and then that location can be analysed for convenience stores, restaurants, train station and airport.
- The analysis made can then be used to determine if that location is an ideal location for the customer or not.

Locations in DELHI



Locations in MUMBAI

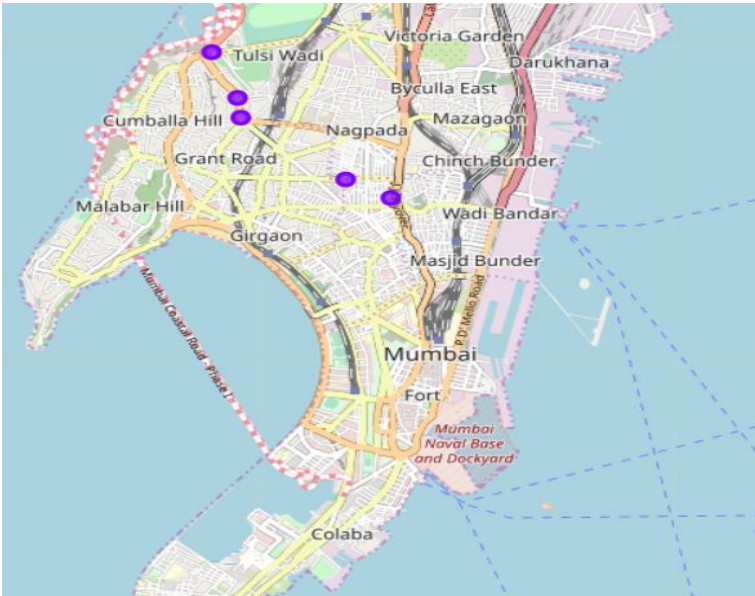
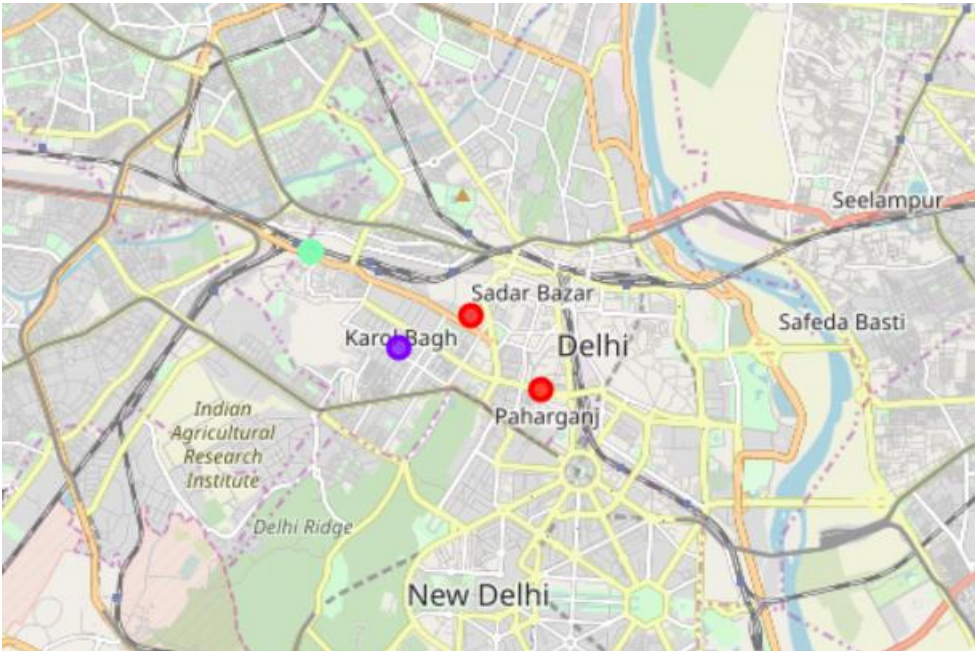


- According to the maps we can find out the most dense localities in Delhi and Mumbai and can make our progress from those locations.
- Chosen locations are South Delhi, Delhi and Mumbai Central, Mumbai

After the k-means clustering was done, the resulting clusters with the maximum density and the most common venues in Delhi and Mumbai was then identified.

	Pincode	Location	District	State	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	110005	Election Commission	SOUTH DELHI	DELHI	30.74108	76.77884	0	Indian Restaurant	Ice Cream Shop	Clothing Store	Miscellaneous Shop	Café	Fast Food Restaurant	Coffee Shop
1	110005	Anand Parbat Indl. Area	SOUTH DELHI	DELHI	28.66585	77.17347	2	Train Station	Airport Terminal	Convenience Store	Business Service	Fast Food Restaurant	Coffee Shop	Comic Shop
2	110005	Anand Parbat	SOUTH DELHI	DELHI	28.66585	77.17347	2	Train Station	Airport Terminal	Convenience Store	Business Service	Fast Food Restaurant	Coffee Shop	Comic Shop
3	110005	Bank Street	SOUTH DELHI	DELHI	43.65962	-70.25125	0	Seafood Restaurant	Hotel	Coffee Shop	Italian Restaurant	Bar	Ice Cream Shop	American Restaurant
4	110005	Desh Bandhu Gupta Road	SOUTH DELHI	DELHI	28.64519	77.21281	0	Hotel	Restaurant	Café	Pizza Place	Hostel	Indian Restaurant	Indian Chinese Restaurant
5	110005	Karol Bagh	SOUTH DELHI	DELHI	28.65156	77.18858	1	Fast Food Restaurant	Snack Place	Hotel	Indian Restaurant	Bakery	Coffee Shop	Comic Shop
6	110005	Master Prithvi Nath Marg	SOUTH DELHI	DELHI	28.65611	77.20108	0	Gift Shop	Coffee Shop	Asian Restaurant	Indian Restaurant	Fast Food Restaurant	Comic Shop	Convenience Store

	Pincode	Location	District	State	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	400007	Bharat Nagar	Mumbai Central	Maharashtra	8.53727	76.99409	0	ATM	Fireworks Store	Plaza	Arcade	Arts & Crafts Store	BBQ Joint	Restaurant
1	400007	Grant Road	Mumbai Central	Maharashtra	18.95929	72.83108	1	Indian Restaurant	Dessert Shop	Ice Cream Shop	Antique Shop	Arcade	Restaurant	Ice Cream Shop
2	400007	Swami Vivekanand Road	Mumbai Central	Maharashtra	28.24905	77.07279	0	ATM	Construction & Landscaping	Food Court	Fireworks Store	Fast Food Restaurant	Dessert Shop	Department Store
3	400007	Tardeo	Mumbai Central	Maharashtra	18.97243	72.81483	1	Chinese Restaurant	Fast Food Restaurant	Sandwich Place	Vegetarian / Vegan Restaurant	Pizza Place	Bengali Restaurant	Food Court
4	400008	Falkland Road	Mumbai Central	Maharashtra	21.67448	87.55792	2	Pharmacy	Vegetarian / Vegan Restaurant	Construction & Landscaping	Fireworks Store	Fast Food Restaurant	Dessert Shop	Department Store
5	400008	Kamathipura	Mumbai Central	Maharashtra	18.96172	72.82627	1	Indian Restaurant	Antique Shop	Arts & Crafts Store	BBQ Joint	Deli / Bodega	Food Court	Food Court
6	400008	Mumbai Central	Mumbai Central	Maharashtra	18.96972	72.81507	1	Chinese Restaurant	Fast Food Restaurant	Train Station	Vegetarian / Vegan Restaurant	Pizza Place	Bengali Restaurant	Food Court
7	400034	Hajiali	Mumbai Central	Maharashtra	18.97834	72.81214	1	Shopping Mall	Ice Cream Shop	Fast Food Restaurant	Department Store	Golf Course	Deli / Bodega	Restaurant
8	400034	Tulsiwadi	Mumbai Central	Maharashtra	22.31646	73.20885	1	Sandwich Place	Juice Bar	Pool Hall	Vegetarian / Vegan Restaurant	Coffee Shop	Fast Food Restaurant	Ice Cream Shop



Results

After analysing the cities we were able to find out the most common venues in both the cities

Delhi

- The most common venue found out to be restaurants. Thus this location can be chosen but it already possesses a lot of competition.

	District	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
0	SOUTH DELHI	76.77884	0	Indian Restaurant	Ice Cream Shop	Clothing Store	Miscellaneous Shop	Café	Fast Food Restaurant	Coffee Shop	Hotel
3	SOUTH DELHI	-70.25125	0	Seafood Restaurant	Hotel	Coffee Shop	Italian Restaurant	Bar	Ice Cream Shop	American Restaurant	Bakery
4	SOUTH DELHI	77.21281	0	Hotel	Restaurant	Café	Pizza Place	Hostel	Indian Restaurant	Indian Chinese Restaurant	Road
6	SOUTH DELHI	77.20108	0	Gift Shop	Coffee Shop	Asian Restaurant	Indian Restaurant	Fast Food Restaurant	Comic Shop	Convenience Store	Dessert Shop

Mumbai

- The most common venue has highest number of ATMs but the second most common location has a large number of restaurants. Thus this location is a good option but has a high competition.

	District	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
0	Mumbai Central	76.99409	0	ATM	Fireworks Store	Plaza	Arcade	Arts & Crafts Store	BBQ Joint	Bengali Restaurant	Chinese Restaurant
2	Mumbai Central	77.07279	0	ATM	Construction & Landscaping	Food Court	Fireworks Store	Fast Food Restaurant	Dessert Shop	Department Store	Deli / Bodega

```
#Cluster 2 for Mumbai
mCentral_merged.loc[mCentral_merged['Cluster Labels'] == 1, mCentral_merged.columns[[2] + list(range(5, mCentral_merged.shape[1]))]]
```

	District	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
1	Mumbai Central	72.83108	1	Indian Restaurant	Dessert Shop	Ice Cream Shop	Antique Shop	Arcade	Restaurant	Deli / Bodega	Fireworks Store
3	Mumbai Central	72.81483	1	Chinese Restaurant	Fast Food Restaurant	Sandwich Place	Vegetarian / Vegan Restaurant	Pizza Place	Bengali Restaurant	Food Court	Italian Restaurant
5	Mumbai Central	72.82627	1	Indian Restaurant	Antique Shop	Arts & Crafts Store	BBQ Joint	Deli / Bodega	Food Court	Fireworks Store	Fast Food Restaurant
6	Mumbai Central	72.81507	1	Chinese Restaurant	Fast Food Restaurant	Train Station	Vegetarian / Vegan Restaurant	Pizza Place	Bengali Restaurant	Coffee Shop	Food Court
7	Mumbai Central	72.81214	1	Shopping Mall	Ice Cream Shop	Fast Food Restaurant	Department Store	Golf Course	Deli / Bodega	Indian Restaurant	Italian Restaurant
8	Mumbai Central	73.20885	1	Sandwich Place	Juice Bar	Pool Hall	Vegetarian / Vegan Restaurant	Coffee Shop	Fast Food Restaurant	Dessert Shop	Department Store

Discussion

- Based on cluster for each cities above, we believe that classification for each cluster can be done better with calculation of venues categories (most common) in each cities. Referring to each cluster, we can't determine clearly what represent in each cluster by using Foursquare - Most Common Venue data.
- However, we can make an assumption about each cluster as follow:
- **Cluster 1:** *Delhi:* Gift Shop
Cluster 2: *Delhi:* Restaurants
Cluster 3: *Delhi:* Train Station
Cluster 1: *Mumbai:* Indian Restaurants
Cluster 2: *Mumbai:* Mix Cuisine Restaurants
Cluster 3: *Mumbai:* Vegetarian Restaurants
- What is lacking at this point is a systematic, quantitative way to identify and distinguish different district and to describe the correlation most common venues as recorded in Foursquare. The reality is however more complex: similar cities might have or might not have similar common venues. A further step in this classification would be to find a method to extract these common venues and integrate the spatial correlations between different of areas or district.
- We believe that the classification we propose is an encouraging step towards a quantitative and systematic comparison of the different cities. Further studies are indeed needed in order to relate the data acquired, then observe it to more meaningful and objective results.

Recommendation

- The following analysis can be improved with following extensions:
 - Consider more categories. For example like "Night life" which is also a good source for customers. But also like "Restaurants", which even if not burger joints may be some concurrence if too many.
 - In the Locality itself, it can also be computed the distance between all the venues in order to find a place with the most number of potential customers.
 - Using smaller geographical areas like Neighborhoods could improve the accuracy for the scores.

Conclusion

- Using Foursquare API, we can capture data of common places all around the world. Using it, we refer back to our main objectives, which is to determine;
 - The similarities or dissimilarities in both the cities
 - Classification of area located inside the city whether it is restaurant, gift place, or others
- In conclusion, both cities Delhi and Mumbai are the center of attraction among India. However, to declare both cities are similar or dissimilar based on common venues visited is quite difficult. Both cities are similar in some venues also dissimilar in certain venues. And for classification based on common venues, again we must have more systematic or quantitative way to identify and declare this. Comparison can be made, but no such method or quantitative data to determine this. We hope in the future, a method to determine it can be established and explored for references.