

# Opening a Restaurant in Delhi, India

Aarushi Singh

- Introduction

Venturing into a business in hospitality/food industry can be tricky. The list of things to keep in mind before owning a business in the sector is quite long. One among them is to choose the locality in such a manner that it helps in the business's long-term profitability and it builds a strong foothold in the area of choice. While being a resident of a city may help with a surface level knowledge of the hip localities initially, there is lots that might get missed when relying on personal experiences alone. This is where data comes in.

We can leverage information found on the Foursquare API to segment the neighbourhoods of Delhi based on the restaurants in the area and accordingly give recommendations backed by data to a new business owner looking to open a restaurant in the city. The information can also be helpful to customers for finding popular neighbourhoods for different kinds of food joint.

**Business Problem:**

A new business owner is looking to open a restaurant in Delhi. Where would you recommend that they open it?

- Data

All data related to restaurants and venues in Delhi are procured by Foursquare API which is a location application that recommends nearby venues and local businesses in the user's area. All data is retrieved in JSON format and then converted to a dataframe in Pandas. The information on the neighbourhoods of Delhi are scraped from the Wikipedia page of the same name. There are total 139 neighbourhoods in Delhi. To call a query from the Foursquare API, we need the latitude and longitude coordinates of the neighbourhoods. To get the coordinates data, we use the library `geopy.geocoders`.

Wikipedia link of Neighbourhoods of Delhi:

[https://en.wikipedia.org/wiki/Category:Neighbourhoods\\_in\\_Delhi](https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Delhi)

Final look of the dataframe (first 10 rows):

	Neighbourhood	Latitude	Longitude
0	Ashok Nagar (Delhi)	28.636	77.1018
1	Ashok Vihar	28.6995	77.1848
2	Ashram Chowk	28.6831	77.1901
3	Babarpur	28.6874	77.2798
4	Badarpur, Delhi	28.4932	77.303
5	Bali Nagar	28.6541	77.1282
6	Bawana	28.7997	77.0329
7	Ber Sarai	28.5484	77.1815
8	Bhajanpura	28.701	77.261
9	Chanakyapuri	28.5947	77.1885

- Methodology

After web scraping the page from Wikipedia and getting the location coordinates and any required data pre-processing, the data is used to make calls to the Foursquare API.

There are different types of calls we can make to the Foursquare API which procures different venue information to us. For our use case, we gauge the attractiveness of each neighbourhood of Delhi by exploring the different kind of venues in each area. For this reason, we make a regular API call to explore each neighbourhood within a 500 metres radius, and it lists us the desired venue recommendations (with a limit of 100 venues).

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Ashok Nagar (Delhi)	28.636021	77.101822	McDonald's	28.636260	77.097104	Fast Food Restaurant
1	Ashok Nagar (Delhi)	28.636021	77.101822	Giani Ice Cream	28.638913	77.101122	Ice Cream Shop
2	Ashok Nagar (Delhi)	28.636021	77.101822	Haldiram's tilak nagar	28.638989	77.101133	North Indian Restaurant
3	Ashok Nagar (Delhi)	28.636021	77.101822	CCD tilak nagar	28.636264	77.097048	Café
4	Ashok Nagar (Delhi)	28.636021	77.101822	Grillz Jail Road	28.635471	77.097031	Fast Food Restaurant
5	Ashok Vihar	28.699453	77.184826	ELIXIR FERTILITY CENTRE	28.700034	77.188093	Health & Beauty Service
6	Ashok Vihar	28.699453	77.184826	Bakers Stop	28.700495	77.188716	Bakery
7	Ashok Vihar	28.699453	77.184826	Invitation Banquet	28.696018	77.185953	Diner
8	Ashok Vihar	28.699453	77.184826	Gola Northend	28.701242	77.189288	Indian Restaurant
9	Ashram Chowk	28.683134	77.190083	S Square	28.683455	77.187857	Chinese Restaurant

Ultimately, we want to analyse each neighbourhood by the kind of activity/place that they are most known for, hence it's important to list each venue's category in the dataframe.

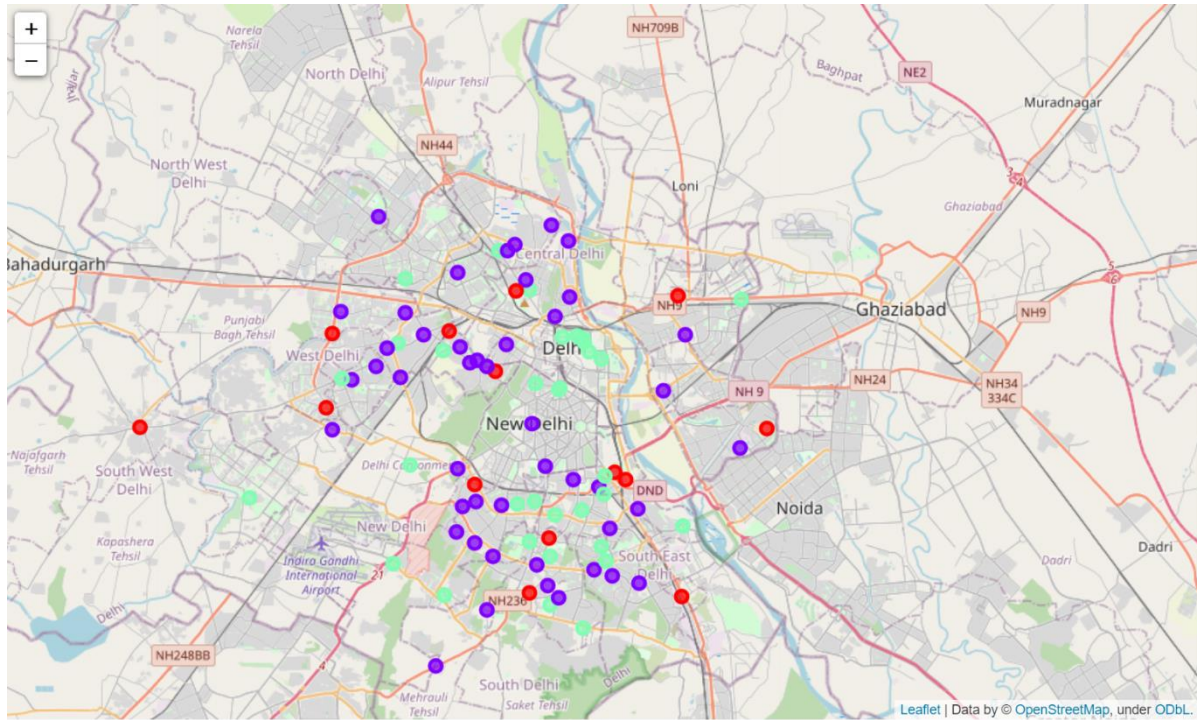
Next, we count the number of different categories in each neighbourhood. We notice that the number of categories returned in each neighbourhood is quite skewed. Since foursquare API returns local businesses and attractions, we may assume that neighbourhoods with categories less than 4 are areas with little commercial avenues. Hence, any owner of a potential restaurant would not be interested in these areas. Therefore, we continue our analysis of Delhi by removing these areas from our consideration. This will also ensure that our clustering results are not skewed.

The next thing to do is to transform our categorical features (i.e venue categories) to integers. For this purpose, we use one hot encoding on VenueCategory. After which, we group the neighbourhoods by the mean of the frequency of occurrence each category. This helps us to determine the top 4 most common venues of each neighbourhood.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Ashok Nagar (Delhi)	Fast Food Restaurant	Ice Cream Shop	North Indian Restaurant	Café
1	Ashok Vihar	Indian Restaurant	Health & Beauty Service	Bakery	Diner
2	Bali Nagar	Indian Restaurant	Light Rail Station	Chinese Restaurant	Ice Cream Shop
3	Ber Sarai	Shopping Mall	Hotel	Airport Food Court	History Museum
4	Chandni Chowk	Indian Restaurant	Snack Place	Market	Jewelry Store
5	Chittaranjan Park	Market	Pizza Place	Fast Food Restaurant	Women's Store
6	Civil Lines, Delhi	Indian Restaurant	Smoke Shop	Italian Restaurant	Light Rail Station
7	Connaught Place, New Delhi	Indian Restaurant	Café	Chinese Restaurant	Bar
8	Dabri, New Delhi	Movie Theater	Department Store	Grocery Store	Food & Drink Shop
9	Dariba Kalan	Indian Restaurant	Market	Snack Place	Hotel
10	Daryaganj	Indian Restaurant	Restaurant	Hotel	Road

Now that our data is prepared with the features required for clustering. Since, our aim is to segment the neighbourhoods on the basis of its top venues, we need a clustering algorithm. We use the K-means clustering which is a partitioning method to segment each neighbourhood according to its top 4 venues. We initialise the algorithm with  $k = 3$ .

Finally, the 93 neighbourhoods look like:



Where red signifies the 1<sup>st</sup> cluster, purple signifies the 2<sup>nd</sup> cluster, and green signifies the 3<sup>rd</sup> cluster.

## • Result

Among 93 neighborhoods of Delhi, we see a total of 14 neighborhoods in the 1st cluster (numbered 0), 46 neighborhoods in 2nd cluster (numbered 1) and 33 neighborhoods in 3rd cluster (numbered 2). The heads of the clusters look like:

Cluster #0:

	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
27	Geetanjali Enclave	28.5304	77.2021	0.0	Café	Metro Station	Coffee Shop	Women's Store
35	Gulmohar Park	28.5571	77.213	0.0	Park	Theater	Café	Music Venue
37	Janakpuri	28.6219	77.0875	0.0	Park	Gym / Fitness Center	Café	Bakery
62	Mayur Vihar Phase - 3	28.6115	77.3363	0.0	ATM	Music Venue	Martial Arts Dojo	Burger Joint
63	Meera Bagh	28.6588	77.0908	0.0	ATM	Convenience Store	Spa	Café
70	Najafgarh	28.6123	76.9824	0.0	ATM	Food & Drink Shop	Mobile Phone Shop	Hookah Bar
71	Nanakpura	28.5839	77.1714	0.0	Department Store	Asian Restaurant	Café	Bus Stop
75	Naveen Shahdara	28.6773	77.286	0.0	ATM	Men's Store	Gym / Fitness Center	Vegetarian / Vegan Restaurant
80	Nizamuddin East	28.5897	77.2504	0.0	Train Station	Park	Café	Food Court
89	Patel Nagar	28.6598	77.157	0.0	Women's Store	Café	ATM	Train Station
93	Rajendra Nagar, Delhi	28.6397	77.1834	0.0	Café	Food & Drink Shop	Sandwich Place	Pizza Place
104	Sarai Kale Khan	28.5866	77.2562	0.0	Train Station	Food Court	Bus Station	Moving Target
105	Sarita Vihar	28.5286	77.2883	0.0	Park	Market	Juice Bar	Train Station
110	Shakti Nagar, Delhi	28.6798	77.1949	0.0	Fried Chicken Joint	Café	Playground	Concert Hall

Cluster #1:

	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Ashok Nagar (Delhi)	28.636	77.1018	1.0	Fast Food Restaurant	Ice Cream Shop	North Indian Restaurant	Café
7	Ber Sarai	28.5484	77.1815	1.0	Shopping Mall	Hotel	Airport Food Court	History Museum
11	Chittaranjan Park	28.5388	77.2492	1.0	Market	Pizza Place	Fast Food Restaurant	Women's Store
12	Civil Lines, Delhi	28.6769	77.225	1.0	Indian Restaurant	Smoke Shop	Italian Restaurant	Light Rail Station
14	Dabri, New Delhi	28.6108	77.0911	1.0	Movie Theater	Department Store	Grocery Store	Food & Drink Shop
18	Dayanand Colony	28.5622	77.2476	1.0	Restaurant	Indie Movie Theater	Fast Food Restaurant	Market
21	Derawal Nagar	28.6998	77.1901	1.0	Shopping Mall	Bakery	Snack Place	Pizza Place
22	Dhaura Kuan	28.5919	77.1617	1.0	Burger Joint	Sake Bar	Golf Course	Bakery
25	East Patel Nagar	28.6456	77.1726	1.0	Pizza Place	Hotel	Donut Shop	Bed & Breakfast
28	Ghitorni	28.4938	77.1492	1.0	Indian Restaurant	Pizza Place	Sandwich Place	Sporting Goods Shop
31	Govindpuri	28.5352	77.2638	1.0	Farmers Market	Pharmacy	Burger Joint	Bar
32	Greater Kailash	28.5419	77.2385	1.0	Restaurant	Park	BBQ Joint	Liquor Store
36	Hauz Khas	28.5443	77.2067	1.0	Chinese Restaurant	Metro Station	Nightclub	Lounge
38	Jangpura	28.5825	77.2415	1.0	Sandwich Place	Fast Food Restaurant	Hotel	Chinese Restaurant
43	Karol Bagh	28.653	77.189	1.0	Snack Place	Asian Restaurant	Dessert Shop	Hookah Bar
44	Keshav Puram	28.6889	77.1617	1.0	Moving Target	Boutique	Basketball Court	Bakery
50	Krishna Nagar, Delhi	28.6578	77.2902	1.0	Bakery	Pizza Place	Men's Store	Gym

## Cluster #2:

	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
1	Ashok Vihar	28.6995	77.1848	2.0	Indian Restaurant	Health & Beauty Service	Bakery	Diner
5	Bali Nagar	28.6541	77.1282	2.0	Indian Restaurant	Light Rail Station	Chinese Restaurant	Ice Cream Shop
10	Chandni Chowk	28.656	77.2322	2.0	Indian Restaurant	Snack Place	Market	Jewelry Store
13	Connaught Place, New Delhi	28.6314	77.2198	2.0	Indian Restaurant	Café	Chinese Restaurant	Bar
15	Dariba Kalan	28.6546	77.2334	2.0	Indian Restaurant	Market	Snack Place	Hotel
16	Daryaganj	28.6461	77.243	2.0	Indian Restaurant	Restaurant	Hotel	Road
19	Defence Colony	28.5712	77.2318	2.0	Indian Restaurant	Italian Restaurant	Bakery	Bar
20	Delhi Cantonment	28.5938	77.135	2.0	Indian Restaurant	Coffee Shop	Shopping Mall	Café
23	Dilshad Garden	28.6758	77.3215	2.0	Indian Restaurant	Light Rail Station	Diner	Pizza Place
24	Dwarka, Delhi	28.5772	77.0443	2.0	Indian Restaurant	Accessories Store	IT Services	Grocery Store
29	Gole Market	28.6337	77.2056	2.0	Indian Restaurant	Snack Place	Japanese Restaurant	Fast Food Restaurant
33	Green Park, Delhi	28.5555	77.2025	2.0	Indian Restaurant	Coffee Shop	Pizza Place	Donut Shop
41	Kailash Colony	28.5531	77.243	2.0	Bar	Indian Restaurant	Italian Restaurant	Flea Market
42	Kamla Nagar, New Delhi	28.6803	77.2021	2.0	Indian Restaurant	Breakfast Spot	Coffee Shop	Miscellaneous Shop
46	Khanpur, Delhi	28.5128	77.2324	2.0	Indian Restaurant	Chinese Restaurant	Plaza	Food & Drink Shop
47	Khari Baoli	28.6571	77.2211	2.0	Dessert Shop	Indian Restaurant	Market	Food & Drink Shop
51	Lajpat Nagar	28.5793	77.244	2.0	Sandwich Place	Indian Restaurant	BBQ Joint	Snack Place

## • Discussion

In the 1st cluster, among the popular venues are ATM, department/convenience stores, cafes and other utility businesses or parks. We can gauge that the cluster mostly consists of neighbourhoods with local market places having small or fast food joints among food venues.

The 2nd cluster, with the greatest number of venues offers very diverse categories. It consists of a vast number of food businesses like restaurants with different range of cuisines, cafes, pizzerias, ice cream shops, bakery etc. Coupled with food venues, are other hospitality businesses or tourist attractions (Eg. Hotels, Multiplex, Museums etc.)

The 3rd cluster contains neighbourhoods heavily popular for Indian restaurants. It also consists of a number of other restaurants and food joints, however the cuisines offered are very limited apart from Indian. Among the top 2 most common venues, food businesses are most common. It also has a number of lodging businesses like hotels and bed and breakfasts, but not much to offer in terms of other tourist attractions.

**Recommendations to business owner:** Out of the 3 clusters, the 2nd cluster (numbered 1) seems the most popular and conducive for any businessperson looking to get into restaurant business. The area is promising because the preferences are diverse in terms of customer's selection of the kind of food joint they wish to eat at (Restaurants, Cafes, Bakery etc.) and the cuisine preference is also diverse and so the owner may not be restricted in their business choices. The areas have strong tourist footfall owing to a number of attractions and lodging avenues offered, so there seems to be promising profitability for any new business to come up.

The 3rd cluster (numbered 2), although popular for restaurants has heavy inclination towards Indian cuisine. Any businessperson looking to open an Indian restaurant may consider the areas, however, there will be high competition and poaching customers from already established Indian restaurants may be tough. Businessperson coming with a franchise that has repute in similar localities in a different city may consider to open a restaurant here.

The 1st cluster (numbered 0) does not seem to be a cluster with a lot of businesses in the hospitality sector that may attract customers. Their food preferences are limited to fast food joints, that suggests that most customers don't visit the areas for a dine-in experience (like Restaurants, Cafes etc.). Other factors considered, if a businessperson is looking to take the risk and get the first mover advantage may invest in a restaurant in this cluster.

- Conclusion

The safest best for a businessperson looking to open a restaurant in the city are the neighbourhoods of 2nd cluster (numbered 1).