

# Applied Data Science Capstone Project Report

## The Battle of Neighborhoods – Report

### 1. Introduction

As suggested in the \*Instructions\* section of the assignment, this problem statement deals with the stakeholders who want to open a new restaurant in New York city. Since the data regarding the city is readily available and easy to find we have decided to go ahead with this city. In this project using the geographical location data from Foursquare and other location data providers, we aim to recommend places/neighborhoods where the business owners can open a restaurant in order to maximize their profits. This will be done by carrying out analysis on the frequency of use of other similar restaurants in the nearby places.

#### 1.1 Why New York City?

Further refining the idea, we focus only on opening *\*North Indian\** restaurants in New York City. Since the Indian population has increased substantially in the surrounding areas of New York City over the past few years, the demand for Indian cuisine is on the rise. North Indian cuisine is also one of the most famous Indian cuisines worldwide and has demand from the non-Indian community in New York. Also, New York city is one of the costliest cities to live in the USA. This is also true for opening a new business or expanding an already existing business. Since, there is a high cost associated with the business expansion in the city, it becomes very important that the owner does a thorough market research before setting up or adding a new restaurant to his/her business.

Applying the techniques learnt in the lab earlier in this course, we can get tabular as well as visual information about the various existing places in New York City as well as the frequency of their use. Since New York City is a large area to be analyzed and the owner is affluent enough, we focus our study on the area of ***Manhattan*** for this project.

#### 1.2 Business Problem

Indian cuisine is famous all over the world. New York City has observed a high immigration rate of Indian residents from the US as well as throughout the world. Thus, it is obvious that to support the tourism industry as well as the working class in NYC, more Indian restaurants must be opened. North Indian cuisine, being one of the most famous Indian cuisine, has a high demand. Thus it would be a great idea if we analyze the performance of the Indian restaurants in NYC, especially Manhattan, to see areas where we can potentially open new restaurants.

#### 1.3 Target Audience

Since we are dealing with the Manhattan area in NYC, we are obviously dealing with affluent businessmen who can afford to open a large-scale restaurant chain in such areas. Also, they need to be in areas where maximum people are satisfied. So, the targeted audience is the Indian tourists, workers also and anyone who loves North Indian cuisine. The objective is to recommend to the management a neighborhood where opening a restaurant would lead to the maximum profit.

#### 1.4 Location Preferences

We need to see that the restaurant opened has a substantial Indian population living or working around in the neighborhood or it is opened in tourist places which attract many Indians or people loving Indian food. This can be capture using previous location data from Foursquare and other location data providers.

However, various factors need to be studied in order to get the best location to open our new restaurant. This project is aimed at capturing all the relevant data and providing a suitable solution.

## 2. Data

As executed in the lab for \*Neighborhood Segmenting and Clustering\* in Week 3 lab, we can use the same dataset from the same location for our project. In order to perform the analysis on locations to open a North Indian restaurant in Manhattan, NYC, we need to have .json file and appropriate API data i.e. location data from data providers. This can be obtained from the following sources:

A) NYC has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. Luckily, this dataset exists for free on the web. We can get the .json file from this link: [https://geo.nyu.edu/catalog/nyu\\_2451\\_34572](https://geo.nyu.edu/catalog/nyu_2451_34572)

B) We also must use the Foursquare Venues list from where we can get the data regarding the previously opened restaurants. For this we can use the following links and codes:

--> From Foursquare Venues Categories - <https://developer.foursquare.com/docs/resources/categories>

--> North Indian Restaurant ID - 54135bf5e4b08f3d2429dfdd

Since the data for specifically the North Indian Restaurants is very low, we must use the data pertaining to Indian Restaurants in the region of Manhattan. For this we can use the following category id:

Indian Restaurant ID - 4bf58dd8d48988d10f941735

## 3. Methodology

In this project, I have used the same methodology as used in the Segmenting and Clustering Lab Session explained during the course in Week 3. To explain the project in detail, we can go stepwise as follows:

- Initially, we import all the necessary libraries from the python packages into our notebook. These libraries include libraries for vector calculations, data frame manipulations and for visualization tools such as maps and other features.
- We then download the required '.json' file from the link provided in the lab session which contains the geographical data about the city of New York with all its boroughs and neighborhoods.
- This data isn't in a pandas dataframe format and must be converted into one before utilizing this data for further manipulations. Thus, we change the original data into a pandas dataframe format using a 'for' loop as suggested in the lab session. We get the following dataframe afterwards:

```
neighborhoods.head()
```

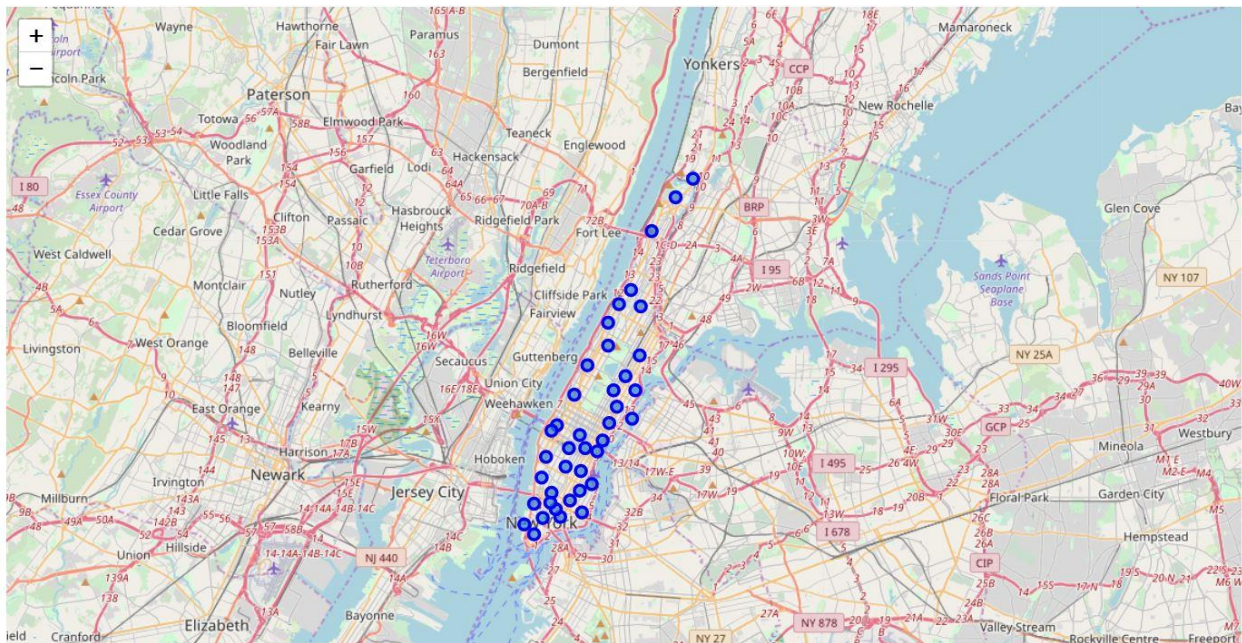
	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

- We can then either utilize a GeoCoder path or manually enter the co-ordinates of New York City in order to obtain the basic map of all the neighborhoods of New York City. This is our next step.
- Now, since we have obtained the location co-ordinates, we tend to focus our attention on the 'Manhattan' region in NYC. We then tabulate the neighborhoods of the region as shown below:

```
man_data = neighborhoods[neighborhoods['Borough'] == 'Manhattan'].reset_index(drop=True)
man_data.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

- Displaying the basic map of New York City and Neighborhoods of Manhattan we have the following:



Note: Since labeling these neighborhoods would not have looked aesthetically pleasing, I have decided not to label these neighborhoods. However, in the original notebook if you click on any of the blue dot you get the neighborhood name.

- The next step of this project is very important one as it requires us to give our Foursquare credentials and set up a client\_id and client\_secret code which will be used further ahead in the project. After doing this, we can use the second data set for our project, which is the 'category' data set from the Foursquare API where in we can compare and find out details of various restaurants already present in that area.



- Appropriately writing the code as written in lab for week 3 we see that we can list down all the Indian restaurants in the Manhattan region. There are lots of Indian restaurants which can be tabulated and displayed on the map as shown below:

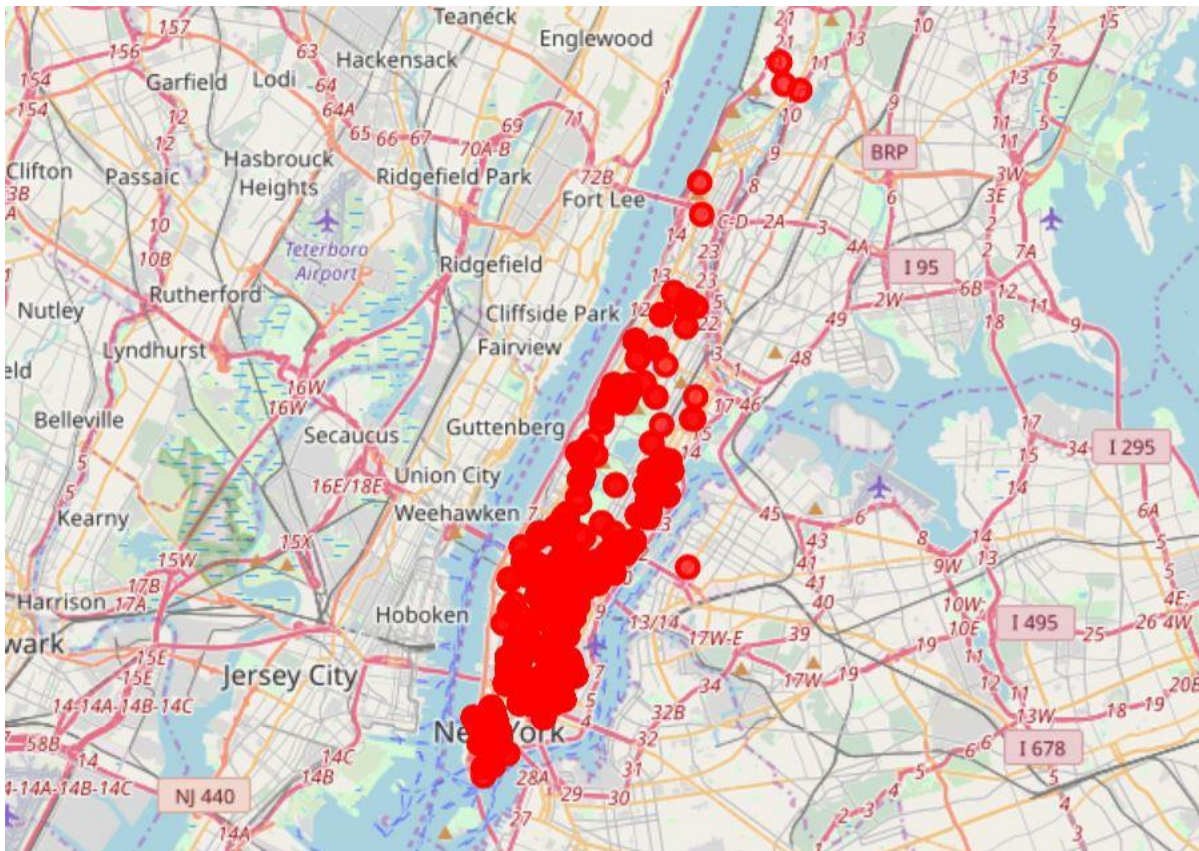
[11]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Riverdale Indian Cuisine	40.880886	-73.908800	Indian Restaurant
1	Marble Hill	40.876551	-73.910660	Tazmohol Indian Restaurant	40.879331	-73.903192	Indian Restaurant
2	Marble Hill	40.876551	-73.910660	Cumin Indian Cuisine	40.886459	-73.909816	Indian Restaurant
3	Chinatown	40.715618	-73.994279	Kabab Bites	40.720094	-73.995819	Indian Restaurant
4	Chinatown	40.715618	-73.994279	indi thai	40.719830	-73.990350	Indian Restaurant

[12]: `newyork_venues_ind.shape`

[12]: (1090, 7)

To be precise there are 1090 Indian restaurants in Manhattan region which are in the Foursquare database. Displaying these on the folium map, we get:



- In the next step we group all these restaurants by the 'Neighborhood' value and tabulate them for further analysis.

- We then perform one-hot encoding which is like the one performed in the lab session in week 3. On doing this we get similar results along with Indian restaurants, for a neighborhood.
- We later group these instances based on their Neighborhoods again so that we can get the mean value of the types of restaurants visited in this region. This will give us a better idea about the frequency of use of the restaurants.
- The next step involves the frequency analysis performed on each neighborhood based on the table received above. The two sample analysis results are as follows:

```

----Battery Park City----
              venue  freq
0  Indian Restaurant  0.84
1           Food Truck  0.08
2   Asian Restaurant  0.04
3   Tapas Restaurant  0.04
4     Burrito Place  0.00

```

```

----Carnegie Hill----
              venue  freq
0      Indian Restaurant  0.92
1  North Indian Restaurant  0.08
2      Asian Restaurant  0.00
3      Burrito Place  0.00
4      Chaat Place  0.00

```

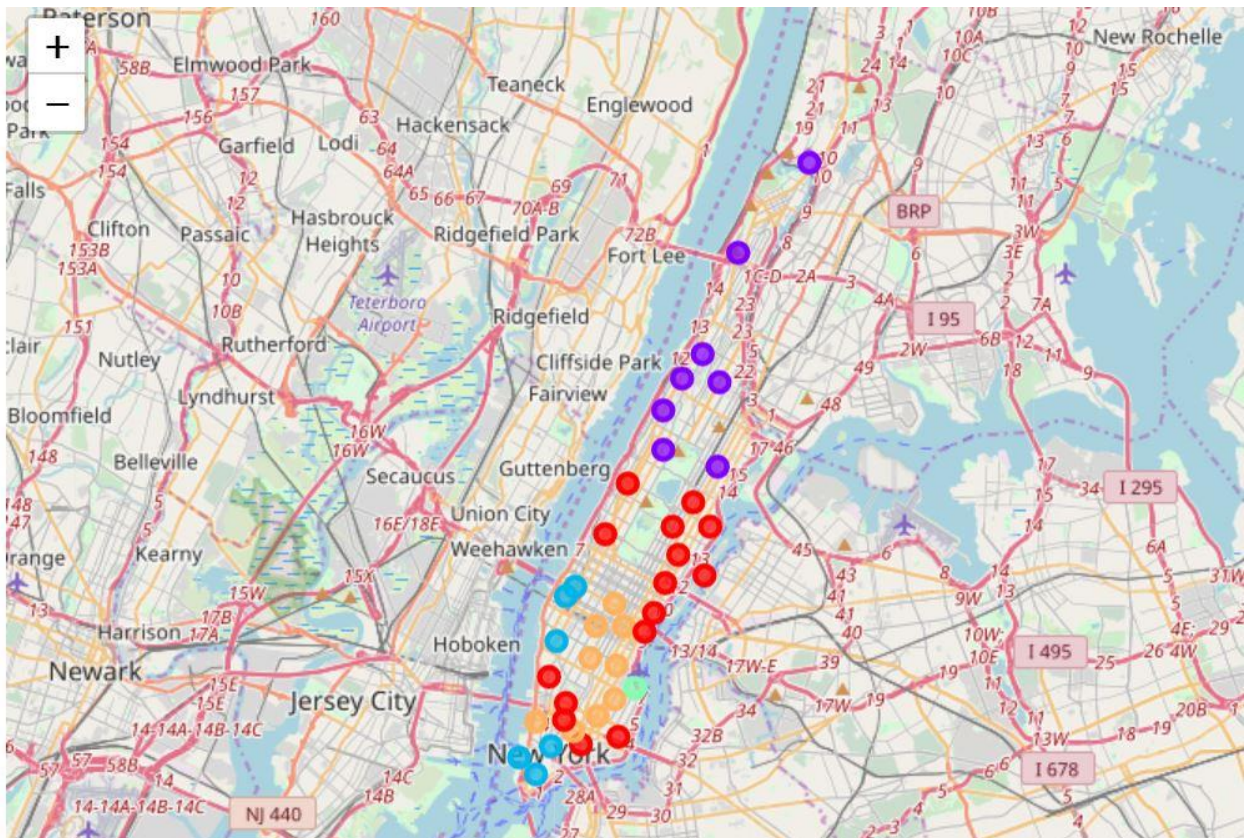
From the above figure we can easily find out which are the most common venues visited in that area and where can we think of opening our Indian restaurant or North Indian restaurant, to be specific.

- Such data can be organized into a tabular format, which can be used later to display up to the 10 most common venues in the neighborhood.
- The table is given onto the next page where we also see that we have added a column of 'Cluster Labels' in order to perform further analysis on the data.
- This clustering is carried by k-means clustering as denoted in the labs and we divide the entire region into 5 different clusters.

```
manhattan_merged.head() # check the last columns!
```

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	1.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
1	Manhattan	Chinatown	40.715618	-73.994279	0.0	Indian Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	1.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
3	Manhattan	Inwood	40.867684	-73.921210	NaN	NaN	NaN	NaN	NaN	NaN
4	Manhattan	Hamilton Heights	40.823604	-73.949688	1.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant

- The last step is to perform a visual cluster analysis which reduce the different neighborhoods to a group of 5 clusters so that it is easier for us to identify regions which have relatively low density of Indian restaurants. We get the following results:





## 4. Results:

The main results of this project are obviously the clustered neighborhoods which provide some idea about the number of Indian restaurants in each neighborhood and the neighborhoods which can be in a similar region.

The k-means clustering provides the following results:

### a. Cluster 1 results:

manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]											
	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	Chinatown	Indian Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
7	Upper East Side	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
8	Yorkville	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
9	Lenox Hill	Indian Restaurant	North Indian Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant
10	Roosevelt Island	Indian Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant
11	Upper West Side	Indian Restaurant	South Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
12	Lincoln Square	Indian Restaurant	South Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant
17	Greenwich Village	Indian Restaurant	North Indian Restaurant	Food Truck	Dosa Place	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
19	Lower East Side	Indian Restaurant	Vegetarian / Vegan Restaurant	North Indian Restaurant	Deli / Bodega	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
22	Soho	Indian Restaurant	Dosa Place	North Indian Restaurant	Food Truck	Deli / Bodega	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
23	West Village	Indian Restaurant	Indian Chinese Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Hookah Bar	Himalayan Restaurant
29	Carnegie Hill	Indian Restaurant	North Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
33	Sutton Place	Indian Restaurant	Food Truck	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
34	Turtle Bay	Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant
35	Tudor City	Indian Restaurant	Food Truck	South Indian Restaurant	Chaat Place	Burrito Place	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant

### b. Cluster 2 results:

manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 1, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]											
	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	Marble Hill	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
2	Washington Heights	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
3	Hamilton Heights	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
4	Manhattanville	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
5	Central Harlem	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
6	East Harlem	Indian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Himalayan Restaurant	Food Truck
24	Manhattan Valley	Indian Restaurant	Himalayan Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Food Truck
25	Morningside Heights	Indian Restaurant	Himalayan Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar	Food Truck

c. Cluster 3 results:

<code>manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]</code>											
	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
13	Clinton	Indian Restaurant	Food Truck	South Indian Restaurant	Diner	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
16	Chelsea	Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Indian Chinese Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Hookah Bar	Himalayan Restaurant
27	Battery Park City	Indian Restaurant	Food Truck	Tapas Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
28	Financial District	Indian Restaurant	Food Truck	Tapas Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar
31	Civic Center	Indian Restaurant	Food Truck	Tapas Restaurant	Dosa Place	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant
38	Hudson Yards	Indian Restaurant	Food Truck	South Indian Restaurant	Diner	Vegetarian / Vegan Restaurant	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant	Hookah Bar

d. Cluster 4 results:

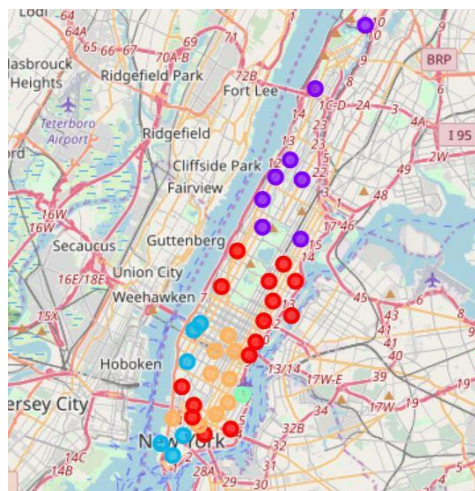
manhattan\_merged.loc[manhattan\_merged['Cluster\_Labels'] == 3, manhattan\_merged.columns[[1] + list(range(5, manhattan\_merged.shape[1]))]]

	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
36	Stuyvesant Town	Indian Restaurant	Vegetarian / Vegan Restaurant	North Indian Restaurant	Dosa Place	Deli / Bodega	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant	Indian Chinese Restaurant

e. Cluster 5 results:

manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 4, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]											
	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
14	Midtown	Indian Restaurant	Food Truck	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Diner	Vegetarian / Vegan Restaurant	Tapas Restaurant	Indian Chinese Restaurant	Hookah Bar
15	Murray Hill	Indian Restaurant	South Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	North Indian Restaurant	Diner	Tapas Restaurant	Pakistani Restaurant	Indian Chinese Restaurant	Hookah Bar
18	East Village	Indian Restaurant	Dosa Place	Deli / Bodega	Vegetarian / Vegan Restaurant	North Indian Restaurant	Food Truck	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
20	Tribeca	Indian Restaurant	Food Truck	Tapas Restaurant	Dosa Place	Asian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant
21	Little Italy	Indian Restaurant	Dosa Place	North Indian Restaurant	Food Truck	Deli / Bodega	Asian Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
26	Gramercy	Indian Restaurant	Vegetarian / Vegan Restaurant	South Indian Restaurant	North Indian Restaurant	Hookah Bar	Food Truck	Deli / Bodega	Chaat Place	Burrito Place	Tapas Restaurant
30	NoHo	Indian Restaurant	Dosa Place	Vegetarian / Vegan Restaurant	North Indian Restaurant	Food Truck	Deli / Bodega	Chaat Place	Tapas Restaurant	South Indian Restaurant	Pakistani Restaurant
32	Midtown South	Indian Restaurant	South Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Diner	Burrito Place	Tapas Restaurant	Pakistani Restaurant	North Indian Restaurant	Indian Chinese Restaurant
37	Flatiron	Indian Restaurant	South Indian Restaurant	Food Truck	Vegetarian / Vegan Restaurant	North Indian Restaurant	Hookah Bar	Diner	Deli / Bodega	Burrito Place	Tapas Restaurant

## 5. Discussions





As we can see from the map above, the Manhattan region has been divided into different clusters. Although there isn't a clear differentiating line for the 'red' cluster we can say that the visual results are satisfactory. Now reading the tabular results for the clusters from the 'Results' section we see that in cluster 1, the 2<sup>nd</sup> most common venues is indeed North Indian Restaurants, the topic of our interest. On further inspection we see that this cluster is represented by the 'red' color on the map. These are heavily congested regions and naturally the prices of these locations will be on the higher side. However, if we look at the Upper West Side Neighborhood, we see that the relative density of Indian restaurants is low, yet the demand for Indian food is higher than other regions. Thus, this can be one of the potential sites for a North Indian Restaurant if adequate funds are available.

The next region belongs to cluster 2 as there is relatively a smaller number of Indian restaurants as compared to the first cluster. Also, there is less congestion in this region which means that the land prices will be less as compared to cluster 1 neighborhoods. Thus, regions like Harlem, Upper Manhattan and Washington Heights are good neighborhoods to open a new Indian restaurant.

There are 3 more clusters to be considered. Although cluster 3 and 5 exhibit good demand for Indian food, it is slightly risky to open a North Indian restaurant as they appear in the 4<sup>th</sup> or 5<sup>th</sup> most common venues in the list. Also, since these neighborhoods lie in the Central Manhattan region, the prices would be the highest. Thus, it would be a bad idea to open a new restaurant here. Cluster 4 doesn't display enough data or statistical evidence from which we can conclude that we should/shouldn't open a new restaurant here as it has only one data point.

## 6. **Conclusion**

Although the goals of the project can be met on a higher level, like we have done in this project, this project can be refined further by providing more data to the algorithm like population data, ethnicity data, tourist information and eating preferences related to tourists as well as the working class travelling to Manhattan. With more such information, we will be able to pin-point our location as well as provide with alternate or back-up regions where we can potentially set-up a new Indian restaurant.

As far as this project is concerned, we were able to decide a rough or approximate region/neighborhood where we can open up a new North Indian restaurant by analytically studying the behavior and preferences of customers and qualitatively making decisions on the land price and competitions by looking and studying the cluster maps produced by our code.