Introduction: Business Problem

New York City's demographics show that it is a large and ethnically diverse metropolis. It is the largest city in the United States with a long history of international immigration. It was termed as a “melting pot”. to describe densely populated immigrant neighborhoods on the Lower East Side. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. Asian Americans flourish in New York City. With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, and French etc.

As New York is highly developed city, the cost of doing business is high, barriers to entry is high and the competition is high as well. Thus, any new business venture or expansion needs to be analyzed carefully. The sights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk and achieve higher return on investment.

Target Audience

F&M entrepreneurs and businesses who wants to open Indian restaurant in New York. They would want preliminary insights into the feasibility and potential of opening Indian restaurants in New York before conducting more in-depth street survey.

The objective is to find the most suitable location for the entrepreneur to open a new Indian Restaurant in New York City. The analysis would provide with entrepreneurs and businesses potential choices of Borough and Neighborhood that are good to start an Indian restaurant.

Criteria for Assessing Location Attractiveness for Indian Restaurant

Criteria 1: The parts of New York City with great Indian restaurants would be favorable to open an Indian restaurant.

- Hence, where are all the major parts of New York City that has great Indian restaurants?
- Do these major parts have potential for Indian Restaurant Market?

Criteria 2: The parts of New York City that lack Indian Restaurants would have opportunity for new Indian restaurant.

- Which areas lack Indian Restaurants?

Criteria 3: The place that are densely populated would be advantageous to have more restaurants particularly restaurants of authentic Asian flavors such as Indian cuisine.

- Hence, which are the places to that are densely populated or have expanding population?

Data

1. List of neighborhoods in New York City
    a. Data source : https://cocl.us/new_york_dataset
        i. This data set contains the required information of neighbourhoods of New York City that can be explored.
2. Latitude and Longitude of these neighborhoods

a. Data source : https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm
    i. Geo space data can be used to get the boundaries that will help to visualize choropleth map

	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

3. Population information of New York
    a. Data source; data.cityofnewyork.us
        i. Contains information of the population in 2000 and 2010

	borough	fips_county_code	nta_code	nta_name	population	year
0	Bronx	005	BX01	Claremont-Bathgate	28149	2000
1	Bronx	005	BX03	Eastchester-Edenwald-Baychester	35422	2000
2	Bronx	005	BX05	Bedford Park-Fordham North	55329	2000
3	Bronx	005	BX06	Belmont	25967	2000
4	Bronx	005	BX07	Bronxdale	34309	2000

4. Locations or venues of Indian restaurants.
    a. Foursquare API
        i. Contains all the venues in each neighbourhood that can be filtered to get only Indian restaurants.

	Neighbourhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Ripe Kitchen & Bar	40.898152	-73.838875	Caribbean Restaurant
2	Wakefield	40.894705	-73.847201	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Jackie's West Indian Bakery	40.889283	-73.843310	Caribbean Restaurant
4	Wakefield	40.894705	-73.847201	Rite Aid	40.889062	-73.842993	Pharmacy

Description

There is a total of 5 Boroughs and 306 Neighborhoods in New York. And there are about 147 indian restaurants across New York based on Foursquare classification.

Methodology

The process starts from extracting the various dataset through API calls to Foursquare and New York City population database. Next, the data will be processed through various data cleaning, wrangling and exploration techniques. After which, visualizations and mapping of the locations help to derive better insights into locations that are favorable to Indian restaurant. Then, clustering technique will

model the similarity across locations and propose centralized group of locations that business owners and entrepreneurs could look forward to exploring them.

Data wrangling and analytics involves getting aggregate sum and averages of information such as population size on Borough and Neighborhood levels. The following information will be obtained after the data processing stage.

1. Getting the count of population by Neighborhood and Borough in thousands

	Borough	Neighborhood	year	population
0	Bronx	Allerton	2000	28510
1	Bronx	Allerton	2010	28903
2	Bronx	Bedford Park	2000	55329
3	Bronx	Bedford Park	2010	54415
4	Bronx	Belmont	2000	25967

2. Getting the count of restaurants by Borough

	Borough	num_rest
0	Bronx	7
1	Brooklyn	25
2	Manhattan	33
3	Queens	75
4	Staten Island	7

3. Getting the Amenities within a Neighborhood via one-hot encoding

	Neighbourhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Amphitheater	Animal Shelter	Antique Shop
0	Allerton	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.016129	0.0	0.0	0.0
1	Annadale	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.117647	0.0	0.0	0.0
2	Arden Heights	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
3	Arlington	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.045455	0.0	0.0	0.0
4	Arrochar	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0

4. Getting the Top 10 venue for Neighborhood via frequency ranking

	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	Allerton	Pizza Place	Donut Shop	Fried Chicken Joint	Caribbean Restaurant	Fast Food Restaurant	Sandwich Place	Pharmacy	Convenience Store	Supermarket	Bus Station
1	Annadale	Restaurant	Trail	American Restaurant	Pizza Place	Pharmacy	Diner	Business Service	Sports Bar	Food	Train Station
2	Arden Heights	Park	Bank	Dog Run	Pizza Place	Sandwich Place	Liquor Store	Pharmacy	Spa	Coffee Shop	Trail
3	Arlington	Bus Stop	Boat or Ferry	Hardware Store	Nightlife Spot	Automotive Shop	Department Store	Discount Store	Donut Shop	Convenience Store	Comic Shop
4	Arrochar	Beach	Italian Restaurant	Baseball Field	Bus Stop	Polish Restaurant	Middle Eastern Restaurant	Supermarket	Mediterranean Restaurant	Bagel Shop	Historic Site

5. Getting the density by Neighborhood

	Borough	Neighborhood	population	num_rest	density
49	Bronx	Soundview	175.210	7	0.039952
16	Bronx	East Concourse	121.245	7	0.057734
61	Bronx	Williamsbridge	118.741	7	0.058952
2	Bronx	Bedford Park	109.744	7	0.063785
53	Bronx	University Heights	108.523	7	0.064502

6. Getting the population growth by Neighborhood

	Borough	Neighborhood	pop_growth
73	Brooklyn	Borough Park	38.2130
81	Brooklyn	Canarsie	35.1695
240	Queens	Jackson Heights	31.5265
276	Queens	South Ozone Park	27.2510
49	Bronx	Soundview	25.7275

Analysis would be done on two levels, first by Borough and next by Neighborhood. Potential Boroughs were selected to be good areas such as Manhattan, Brooklyn and Bronx out of the 5 Boroughs in New York. Next, neighborhoods within these Boroughs are selected and recommended by taking into considerations restaurant density, Indian restaurant popularity and population growth.

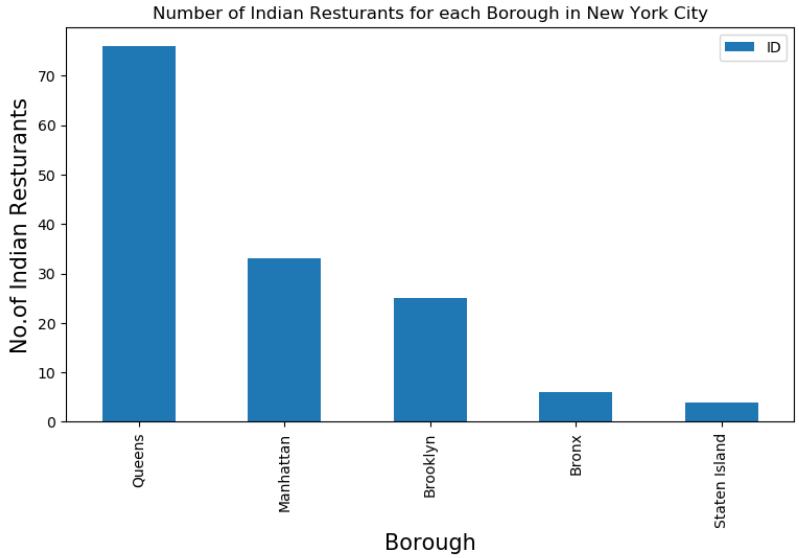
Data visualizations aid in picturing the exact locations to explore for Indian restaurants. It could also provide geospatial information on the surrounding especially amongst clusters of locations. The following image is all the Neighborhoods in New York using Geopy and Folium libraries to create a map of New York city with neighborhoods superimposed on top.

Clustering models could be explored for this unsupervised learning problem with dataset that is not labelled. The model results would aggregate locations that have similarities in the amenities within

the Neighborhood and hence would give a more accurate insight on which type of Neighborhood would best fit an Indian Restaurant. This is achieved by calculating the average overall rating for each cluster after the optimal clusters are decided by the model via the KMeans elbow method.

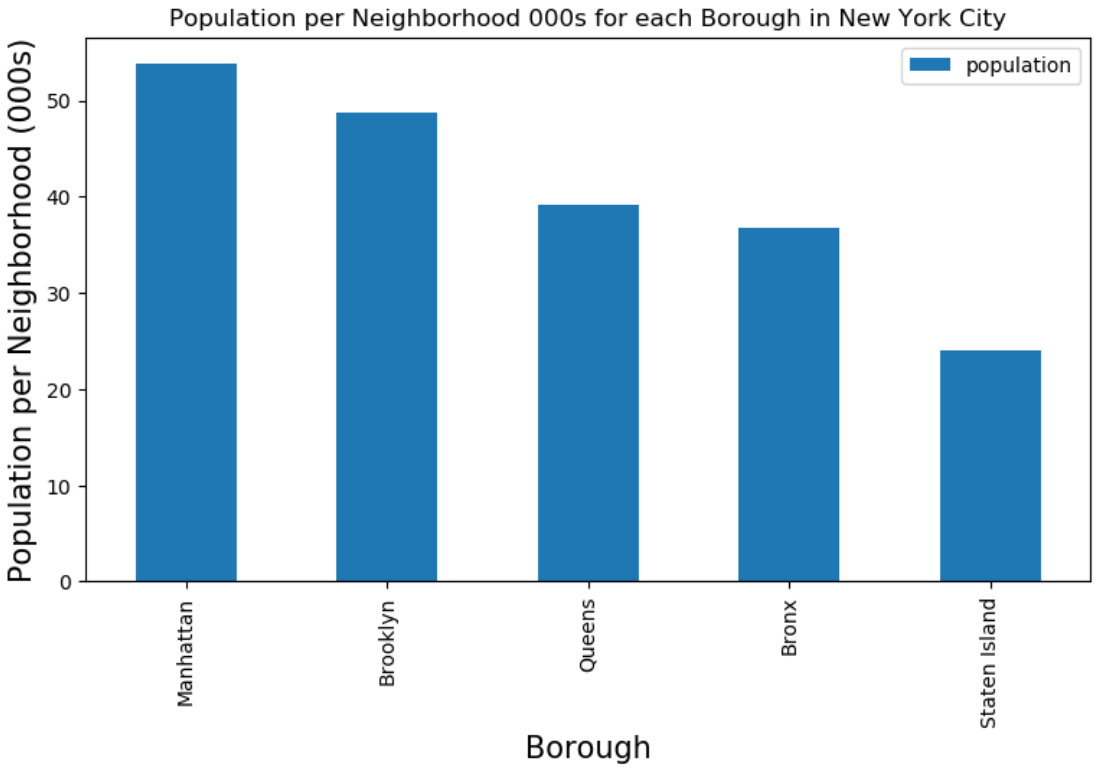
Results

The major parts of New York that has great Indian restaurants are Queens, Manhattan, Brooklyn.



Neighborhood	pop_growth	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	res
City Island	NaN	Harbor / Marina	Seafood Restaurant	Italian Restaurant	Deli / Bodega	Bar	American Restaurant	French Restaurant	Bank	Baseball Field	Thrift / Vintage Store	0.4
Edgewater Park	NaN	Italian Restaurant	Pizza Place	Deli / Bodega	Bar	Donut Shop	Spa	Diner	Latin American Restaurant	Sushi Restaurant	Asian Restaurant	0.4
Manhattan Beach	NaN	Italian Restaurant	Turkish Restaurant	Café	Eastern European Restaurant	Dessert Shop	Seafood Restaurant	Sandwich Place	Bar	Russian Restaurant	Bus Stop	0.5
Ditmas Park	NaN	Caribbean Restaurant	Mexican Restaurant	Pizza Place	Coffee Shop	Grocery Store	Bar	Thai Restaurant	Indian Restaurant	Mobile Phone Shop	Latin American Restaurant	0.5
Schuylerville	8.792	Pizza Place	Italian Restaurant	Diner	American Restaurant	Donut Shop	Pharmacy	Supermarket	Mexican Restaurant	Sandwich Place	Latin American Restaurant	0.4

To analyse further on the potential of Borough and Neighbourhood, restaurant density should be taken into account. Criteria 1 was further met with a deeper analysis into the restaurant density of the Neighbourhood based on how frequent Restaurants are 1km away from the centre which suggests that these Neighbourhoods are popular to have Restaurants.

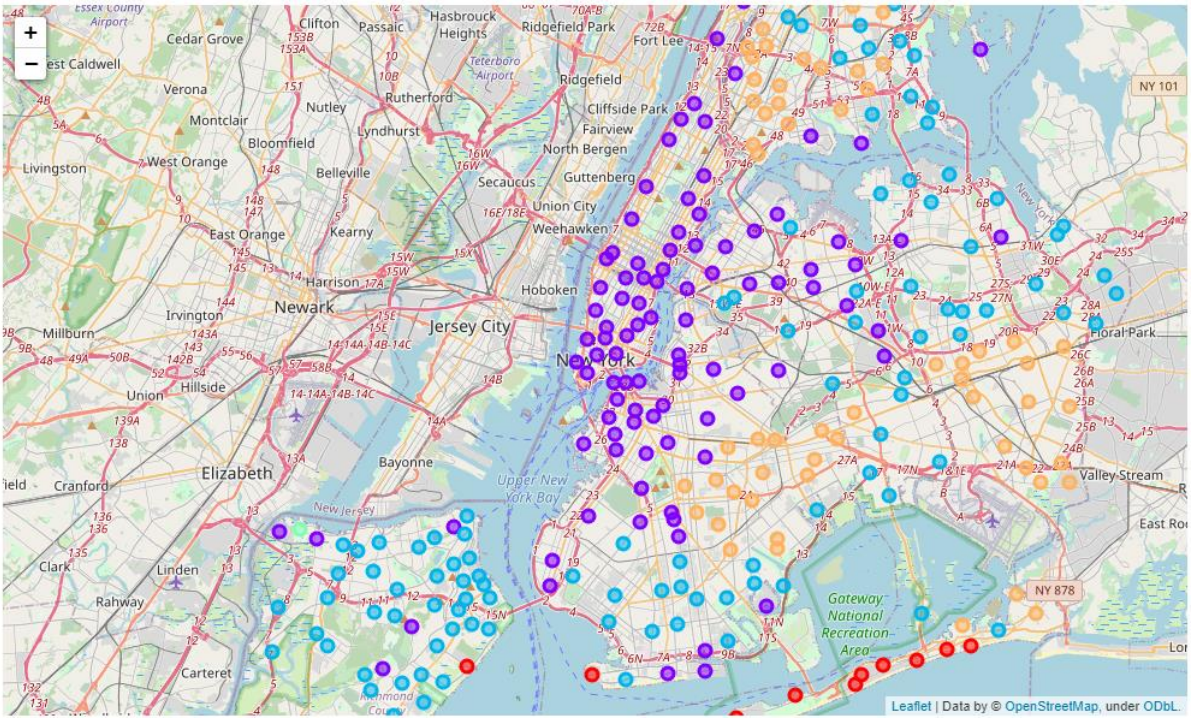


Here, we see that Manhattan and Brooklyn could have more Indian restaurants for their larger population size compared to Queens and Bronx. Currently, Indian restaurants in Manhattan and Brooklyn is about half that of Queens. This suggests that there can be new Indian Restaurants in these two Boroughs given that it is less competitive and there is greater potential for more customers, thereby meeting Criteria 2, that the parts of New York City with lack Indian Restaurants would be favourable to open an Indian restaurant.

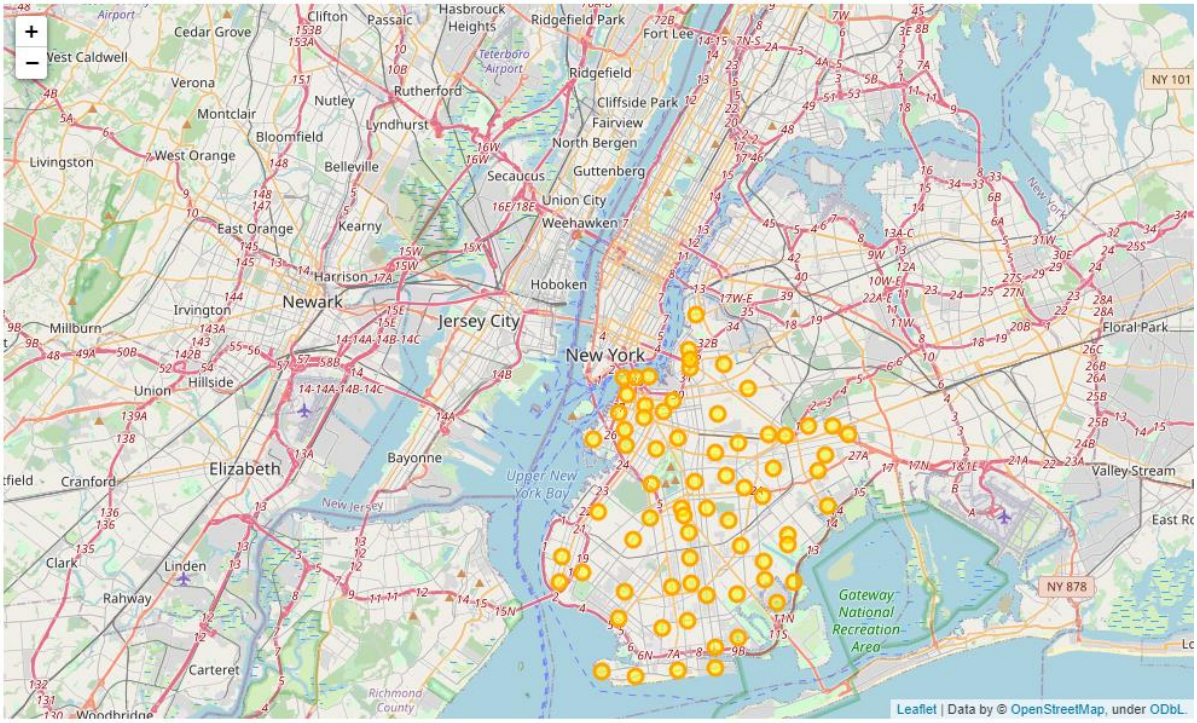
The place that are densely populated would be advantageous to have more restaurants particularly restaurants of authentic Asian flavors such as Indian cuisine. Hence, after analysis on the population changes, Boroughs such as Brooklyn and Manhattan have the most population whereas Boroughs such as Brooklyn and Bronx have the most population growth. There in all, Criteria 3 is met that Brooklyn, Manhattan and Bronx remains potentially attractive for new Indian Restaurants in that order. Specific neighbourhoods that could be good locations to set up would be Manhattan Beach, Brooklyn; Ditmas Park, Brooklyn.

The clustering analysis would give the same results that the favourable clusters 3, 4 and 1 are largely centred around Brooklyn, Manhattan, and Bronx.

```
Cluster Labels
3    0.571071
4    0.520260
1    0.505841
2    0.437170
0    0.255547
Name: overall_rating, dtype: float64
```



This matches with the below geospatial graph that Brooklyn and Manhattan have huge numbers of cluster 4 and 1. The top clusters 4 and 1 are in the area of Manhattan and Brooklyn and a bit further up would be Borough Bronx. This coincides with the conclusions derived from the previous visualizations and data analytics.



Discussion

The analysis helps any entrepreneurs or businesses that are keen to set up Indian Restaurants in New York. However, these insights are not able to cover all factors not excluding restaurants profiles,

residents and tourists' demographics and many other crucial factors that could impact a restaurant business. Further analysis such as street surveys are critical to obtain more consumer validation of the business potential. These insights derived from the analysis helps to narrow down the search and save costs for business owners to conduct further more specific market penetration tests. Moreover, it offers a bird-eye view to the drivers for restaurant business in New York.

The purpose of this analysis was to only provide info on areas in New York City that are potentially great for opening a new Indian Restaurant but not crowded with existing restaurants (particularly Indian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

The analysis could be more extended to include more granular details of demographic variety in each Neighbourhoods or the tourists that are visiting the neighbourhoods to complement the location-based analysis that is conducted during this research.

Conclusion

There is potential to start Indian Restaurant in Brooklyn and Manhattan as first priority. Good neighbourhoods. Our analysis shows that although there is a great number of restaurants in New York City, there are pockets of low restaurant density in Manhattan and Brooklyn. The good locations are mainly centred around the Upper New York Bay Area. Neighbourhoods within Brooklyn and Bronx has a high population growth while also having a low restaurant density. The recommendations are derived after achieving all the criteria that defines what would a good location be: high population and growth, low Indian restaurant density and high restaurant popularity. There is definitely a huge potential to open more Indian restaurant in these metropolitan Boroughs and Neighbourhoods.

Referrences

1. Foursquare API
2. New York City Data