Capstone Project: Battle of the Neighborhoods

Chintan Patel

14th June 2020

1. Introduction

1.1 Background:

Indian ethnic food is "the next big thing" in the USA food industry. According to Mr. Anshu Dua, founder and chief executive officer of The Chaat Co., New York _"Indian cuisine has been underrepresented in the U.S. market for quite some time. Most Americans experience Indian food in poorly branded, Taj Mahal-themed restaurants with all-you-can-eat buffets featuring mysterious curries, rice, and naan. While delicious, that experience has not evolved in most cities for at least the past 30 years.[1] "As per Washington Post report, there are, more than 40,000 Chinese restaurants around the country, and roughly the same number of Mexican restaurants, but only about 5,000 Indian restaurants. According to Krishnendu Ray, a professor at New York University who has been studying the cuisine's rise for more than a decade the Indian cuisine is the next "ethnic food trend." There are about 2.7 million Indians in the U.S. -- about half a million in New York City.[2]

One of the main advantages of opening new Indian cuisine restaurant in New York is steadily increasing Indian origin people immigrating to this cosmopolitan, and a substantial number of them are professionals. Though, of course, there are also Indians in less professional fields but most of them belong to higher up in income bracket, they have more than enough purchasing power if they have suitable option for original ethnic Indian Food. About 30 percent to 50 percent of Indian immigrants have substantial cultural capital, and many are Anglophones. The prestige of Indian immigrants gets linked to prestige of Indian food.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] *US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.*

1.2 Business Problem:

Although there may be enough Indian Cuisine Restaurants in New York Area to cater the needs of Indian Origin people for Indian food, due to constant influx of new Immigration there is still scope for a new restaurant which can provide authentic taste of India. This project aims to identify the suitable location for a new Indian Restaurant at the borough level of detail using Foursquare API.

In this project, I want to try to find the best borough in New York City to open an Indian Cuisine Restaurant. The challenge is to find a suitable neighborhood which is close enough to some amenities and venues, has fewer competitors, and of course is affordable to rent. An ideal location should be one where there is highest concentration of Indian Americans and should also be near a place where many people visit such as near a cinema, park, garden, playground, and theater.

For this project, I am going to create a simple guide on where to eat based on Foursquare likes, restaurant category and geographic location data for Indian restaurants in New York city. Then I will find Indian restaurants per borough and then highest number of Indian Americans per borough. Then I will try to extract the information on which borough has lowest number of restaurants per population, so that my new venture would have minimum competition and maximum market potential.

1.3 Problem Statement

1. Which is the best rated Indian cuisine restaurant in New York City?
2. In what Neighborhood or borough should I open an Indian cuisine restaurant to have the best chance of being successful?

1.4 Target Audience

- Business personnel who wants to invest or open an Indian restaurant in NY. This analysis will be a guide to start Indian restaurants targeting the Indian Americans.
- Freelancer who loves to have their own restaurant as a side business.
- Indian origin person who wants to move in the neighborhoods with lots of option for Indian restaurants in NY.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

- Business Analyst or Data Scientists, who wish to analyze the neighborhoods of NY using Exploratory Data Analysis

2. Data

   2.1 To answer the above questions, data on New York City neighborhoods, boroughs to include boundaries, latitude, longitude, restaurants, and restaurant ratings and tips are required. The elaborative New York City data containing the neighborhoods and boroughs, latitudes, and longitudes has been obtained from the data source:

   https://cocl.us/new_york_dataset.

   2.2 We also need enough data of Indian Americans living in New York city as per The American Community Survey (ACS) which is an ongoing survey by the U.S. Census Bureau. It regularly gathers information previously contained only in the long form of the decennial census, such as ancestry, citizenship, educational attainment, income, language proficiency, migration, disability, employment, and housing characteristics. These data are used by many public-sector, private-sector, and not-for-profit stakeholders to allocate funding, track shifting demographics, plan for emergencies, and learn about local communities. Sent to approximately 295,000 addresses monthly (or 3.5 million per year), it is the largest household survey that the Census Bureau administers.[3] The required data specially a table is scraped from the following Wikipedia page:

   https://en.wikipedia.org/wiki/Indians_in_the_New_York_City_metropolitan_region

   2.3 All data related to locations and quality of Indian restaurants will be obtained via the Foursquare API utilized via the Request library in Python.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] *US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.*

3. Methodology

- The required data related to New York borough and neighborhood has been collected from https://cocl.us/new_york_dataset and then it was cleaned and processed into a data frame.

- Wikipedia webpage containing data regarding the New York Demographics has been scraped using urllib and BeautifulSoup libraries.

- Foursquare API has been used to locate all venues and then filtered by Indian restaurants. Ratings, tips, and likes by users will be counted and added to the data frame.

- The restaurant Likes and Tips data has been sorted based on rankings

- Finally, the data be will be visually assessed using graphing from various Python libraries.

4. Analysis

After all the necessary libraries imported, to segment the neighborhoods of New York City, a dataset is required that contains the 5 boroughs and the neighborhoods, that exist in each borough, with respective latitude and longitude coordinates. This dataset is downloaded using the mentioned URL.

Once the .json file is downloaded, it is analyzed to understand the structure of the file. A python dictionary is returned by the URL and all the relevant data is found to be in the features key, which is basically a list of the Indian Restaurants in different neighborhoods. The dictionary is transformed, into a pandas data frame, by looping through the data and filling the data frame rows one at a time.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

As a result, a data frame is created with Borough, Neighborhood, Latitude and Longitude details of the New York City's neighborhood.

	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

Upon analysis, it is found that the data frame consists of 5 boroughs and 306 neighbourhoods.

Further, 'geopy' library is used to get the latitude and longitude values of New York City, which was returned to be Latitude: 40.71, Longitude: -74.01.

The Foursquare API is used to explore the neighborhoods for search query "Indian Restaurants". To access the API, 'CLIENT_ID', 'CLIENT_SECRET' and 'VERSION' is defined.

As a result, a data frame is created with all the Indian restaurants in New York City registered in Foursquare with their Foursquare Id, Name, Category, Latitude and Longitude details.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	location.lat	location.lng
0	5072fe2eb0ed0fcce2497f30	Indian Biryani Delights (Cart)	[{'id': '4bf58dd8d48988d1cb941735', 'name': 'F...	v-1592084761	False	Water St	btwn Maiden Lane & Pine St	40.713165	-74.010219
1	510aad1ae4b0bfb2583169b9	A Saffron Thread Fresh Indian	[{'id': '4bf58dd8d48988d10f941735', 'name': 'I...	v-1592084761	False	98 Chambers St	at Church St	40.714823	-74.007689
2	4a948b83f964a520f32120e3	Indian Express	[{'id': '4bf58dd8d48988d10f941735', 'name': 'I...	v-1592084761	False	18 Murray St	Church St.	40.713413	-74.008469
3	574eab2a498e3f5530cfdbca	Indian Beautiful Art (IBA) Crafts PVT Ltd	[{'id': '4bf58dd8d48988d103951735', 'name': 'C...	v-1592084761	False	100 Church St	NaN	40.712281	-74.010129
4	574e8ba2cd10da9c722374fb	Indian Beautiful Art (IBA) Crafts PVT Ltd	[{'id': '4bf58dd8d48988d1f4941735', 'name': 'D...	v-1592084761	False	100 Church St	NaN	40.716135	-74.006610

Fig. 1 List of all Indian Resturants from Foursquare Data with their location data

As, the aim is to segment the boroughs and neighborhoods of New York City with respect to the 'Indian Restaurants' in its vicinity, it is further required to fetch this data from all the 306 neighborhoods' venues. The Foursquare AP has been called to prepare neighborhood list that contains Indian restaurants for all 306 neighborhood venues.

```
( 297 / 306 ) Indian Resturants in Madison, Brooklyn:0
['Brooklyn', 'Madison', 40.60937770113766, -73.94841515328893]
( 298 / 306 ) Indian Resturants in Bronxdale, Bronx:0
['Bronx', 'Bronxdale', 40.85272297633017, -73.86172577555115]
( 299 / 306 ) Indian Resturants in Allerton, Bronx:0
['Bronx', 'Allerton', 40.86578787802982, -73.85931863221647]
( 300 / 306 ) Indian Resturants in Kingsbridge Heights, Bronx:0
['Bronx', 'Kingsbridge Heights', 40.8703923914147, -73.90152264513144]
( 301 / 306 ) Indian Resturants in Erasmus, Brooklyn:0
['Brooklyn', 'Erasmus', 40.64692606658579, -73.94817709920184]
( 302 / 306 ) Indian Resturants in Hudson Yards, Manhattan:0
['Manhattan', 'Hudson Yards', 40.75665808227519, -74.00011136202637]
( 303 / 306 ) Indian Resturants in Hammels, Queens:0
['Queens', 'Hammels', 40.58733774018741, -73.80553002968718]
( 304 / 306 ) Indian Resturants in Bayswater, Queens:0
['Queens', 'Bayswater', 40.611321691283834, -73.76596781445627]
( 305 / 306 ) Indian Resturants in Queensbridge, Queens:0
['Queens', 'Queensbridge', 40.756091297094706, -73.94563070334091]
( 306 / 306 ) Indian Resturants in Fox Hills, Staten Island:0
['Staten Island', 'Fox Hills', 40.61731079252983, -74.08173992211962]
```

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

After the list has been created I exported the list to.csv frmate for future use as we get only limited number of API calls in free account. The final list contains only 31 functioning restaurant with proper data.

```
In [60]: indian_rest_stats_ny.info()
         <class 'pandas.core.frame.DataFrame'>
         RangeIndex: 31 entries, 0 to 30
         Data columns (total 7 columns):
          #   Column        Non-Null Count  Dtype
         ---  ------        --------------  -----
          0   Borough       31 non-null     object
          1   Neighborhood  31 non-null     object
          2   ID            31 non-null     object
          3   Name          31 non-null     object
          4   Likes         31 non-null     object
          5   Rating        31 non-null     float64
          6   Tips          31 non-null     object
         dtypes: float64(1), object(6)
         memory usage: 1.8+ KB
```

Then as described in or problem statement I tried to find out the restaurnt with maximum likes and maximm tips by calling the Foursquare API and found out that Saffron Garden sittuated in S taten Island is the restaurant with maximum likes (i.e. 17) and also the same restaurant tops th e list of maximum Tips (i.e. 16) and surprisingly it is the restaurnt with maximum rating (i.e. 7.7/ 10) as well.

```
Borough                      Staten Island
Neighbourhood                    Fox Hills
ID             4c48da9f3013a59356c5f0e1
Name                        Saffron Garden
Likes                                   17
Rating                                 7.7
Tips                                    16
```
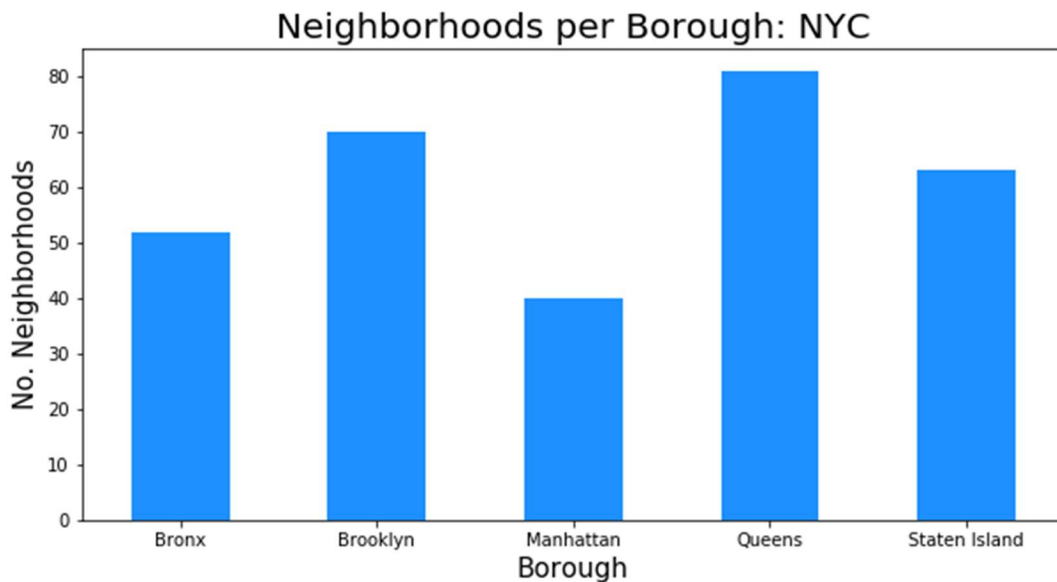
So, by analyzing the above data it was evident that we end up with rather smaller dataset with only 31 Indian Restaurant from Foursqure. So, instead of going for K means or any other clustering approach I decided to go with trditional approach of manual data crunching.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states
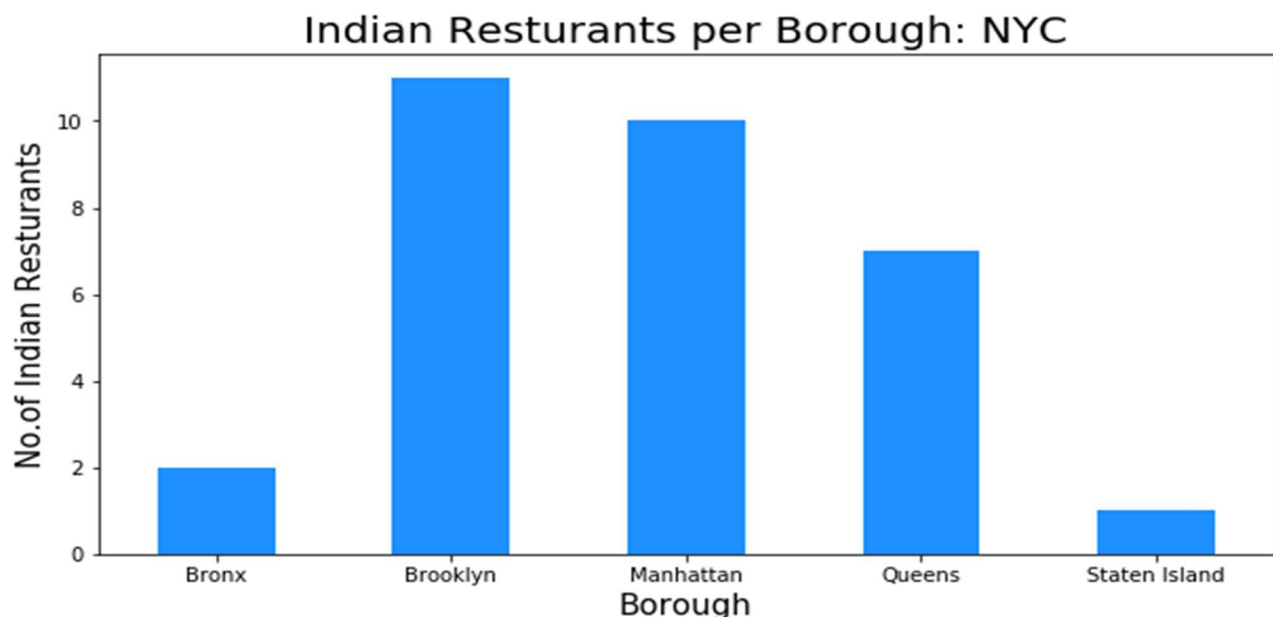
[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] *US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.*

I decided to find out the data of Indian restaurants per borough and neighbourhood. First I segmented the boroughs to decide how many neighbourhoods are there per borough. Using Matplotlib I plotted a bar graph that shows Neighbourhoods per boroughs in New York.

Neighborhoods per Borough: NYC

From the figuer we can gather that there are maximum number of neighbourhoods in Queens borough. And from the above analysis we knew that the best rated restaurant is sittuated in Staten Island and not in Queens. So, I decided to plot the restaurants per borough using matplotlib.

Indian Resturants per Borough: NYC

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

From the graph above we came to know that there are only 7 restaurants in Queens borough, I draw the list of those seven restaurants using foursquare API.

```
In [158]: indian_rest_ny[indian_rest_ny['Borough']=='Queens']
```

Out[158]:

	Unnamed: 0	Borough	Neighborhood	ID	Name
21	21	Queens	Astoria	5d8d9c5e0a2ffc0008f28855	The Gully
22	22	Queens	Kew Gardens	5782c9ce498edde587f5aa14	Tikka Indian Grill
23	23	Queens	Long Island City	50a287a7e4b0033f830f06db	Raj's Indian Kitchen
24	24	Queens	Rego Park	4b3a8aa5f964a520996925e3	Sajni 026
25	25	Queens	Bayside	4f1f4996e4b01ff351a7a50c	Ayna Agra Indian Restaurant
27	27	Queens	Jamaica Hills	4c434b2bd691c9b6ef8f8f0a	Sagar Restaurant
30	30	Queens	Sunnyside Gardens	4c48da9f3013a59356c5f0e1	Saffron Garden

Considering that Queens has the highest number of neighbourhoods and decent populaion, I thought just seven restaurants are not enough to cater the needs of so many people, but being a Data Scientist I cant just rely on intuition and this has to be supported by the strength of data.So I wanted to find number of Indian Americans living in the boroughs of New York and test my hypothesis.

To find more about demography of New York boroughs I scrapped a table from the wikipedia link already mentioned in the "Data" section. To scrape the table from the webpage I used 'urllib' and 'BeautifulSoup' libraries. I needed to extrct a partcular table from a webpage so I used .prettify function of beautiful soup librabry to view the HTML code to identify and separate the table I needed. Then the empty table was created and then after the required table was identified the empty table was populated using that data with only required columns.

	Rank	Borough	City	Indian_Americans	Density_per_square_mile	Percentage_in_municipality's_population
0	1\n	Queens	New York City	144,896	1,326.5	6.2
1	2\n	Brooklyn	New York City	25,270	357.9	1.0
2	3\n	Manhattan	New York City	24,359	1,060.9	1.5
3	4\n	The Bronx	New York City	16,748	398.6	1.2
4	5\n	Staten Island	New York City	6,646	113.6	1.4
5	\n	Total (2014)	New York City	227,994	753.4	2.7

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.

As it is evident from the table extracted from the census data of the New York City, there are total of 227,994 Indian Americans living in 5 boroughs of New York. Out of these more than half that is 144, 896 Indian Americans are living in Queens borough only! So, our hypotheiss stood the test of data.

5. Discussion:

This dataset only include Indian Americans meaning people of Indian Origin who are already citizens of USA. There is no proper data available of Indian professionals living in and around Queens borough on H1B visa and on Permanent Residence status, so they are not included in this dataset. But, if included it will only increase the number of people living in Queens borough who are financially well equipped to be our potential customers.

6. Result / Conclusion :

The crux of the report is, based on data available on Foursquare there are only 7 Indian restaurants catering the needs of 144, 896 people of Queens borough, effectively meaning *there is only one Indian Restaurant option available to approximately 21000 Indian Americans.* The JFK Airport which is used frequently by almost all Indian Americans living in New York area to travel to India and back, is added beebfit to the loction. Also as we have already examined out of seven restaurants situated in Queens, none of them is rated highest or liked by the customers, meaning there is huge potential and virtually unexplored market for Indian cuisine restaurants in Queens borough, New York.

[1] https://www.foodbusinessnews.net/articles/11822-the-time-is-right-for-authentic-indian-cuisine-in-the-united-states

[2] https://www.washingtonpost.com/news/wonk/wp/2015/03/04/why-delicious-indian-food-is-surprisingly-unpopular-in-the-u-s/

[3] *US Census Bureau. "ACS Information Guide". www.census.gov. p. 8. Retrieved July 31,2015.*