**Capstone Project – Battle of Neighborhoods**

**Introduction & Background**

A major food-chain firm, XYZ is looking at opening its first restaurant in New York. Assumption is that XYZ does not have much budget constraint and is looking at entering the market. The investment is majorly driven by the demand and customer preferences.

Thus, it is of essence to analyze the existing restaurants in New York to get a flavor of the type of cuisines, their locations and customer ratings.

XYZ is expecting to look at the most famous cuisines/type of restaurants and the neighborhood wherein there exist an opportunity to invest and open a new restaurant.

Although while deciding on the cuisine & location of restaurant there are a lot of parameters that can be taken into account like rent of location, connectivity, competitor analysis, proximity to offices/colleges/schools along with financial decisions based on cost benefit analysis, we will be focusing on 2 key parameters:

* The most famous cuisines/restaurant types in New York
* The location wherein an opportunity exist

Target Audience is XYZ leadership.

**Data**

We have used the following data for the project:

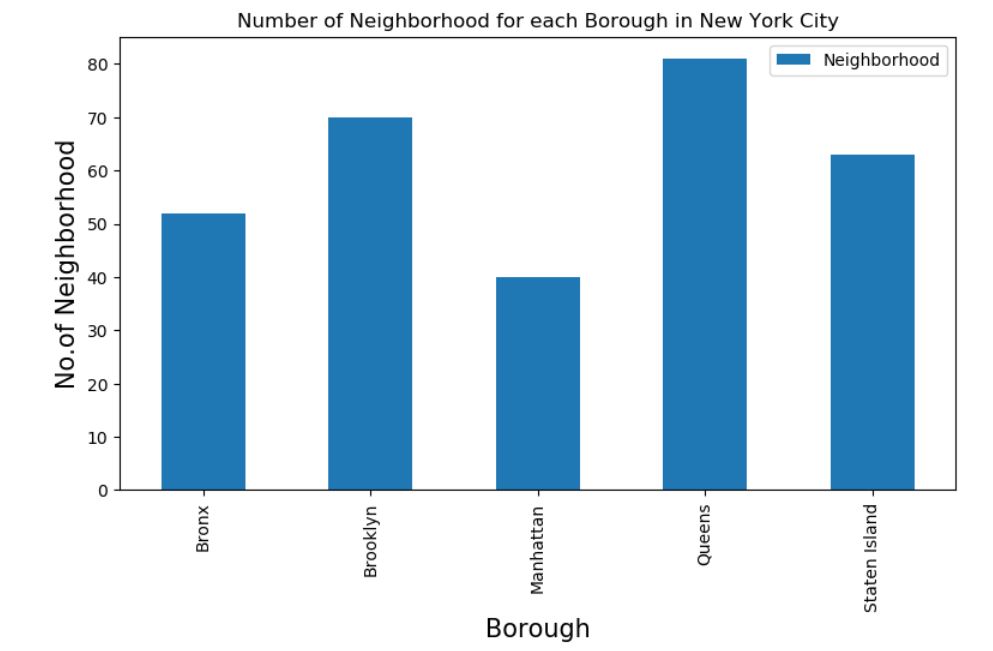
1. New York City data that contains list of Borough, Neighborhoods along with latitude & longitude information
   1. Source: <https://cocl.us/new_york_dataset>
2. Foursquare API data
   1. We will be using this data to get details on the venues in each neighborhood and exploring the different types of restaurants

**Methodology**

The following steps have been followed:

1. *Importing Libraries:* First, we import all the necessary libraries
2. *Importing Neighborhood Data:* Once we have all the relevant libraries in place we collect and import data:
   1. *New York Neighborhood Data:* Data consisting of the list of Borough and corresponding neighborhoods in New York. This data also contains the geographical coordinates for each neighborhood i.e. latitudinal and longitudinal coordinates. Data has 306 rows & 4 columns.

| **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| --- | --- | --- | --- |
| Bronx | Wakefield | 40.894705 | -73.847201 |
| Bronx | Co-op City | 40.874294 | -73.829939 |
| Bronx | Eastchester | 40.887556 | -73.827806 |
| Bronx | Fieldston | 40.895437 | -73.905643 |
| Bronx | Riverdale | 40.890834 | -73.912585 |



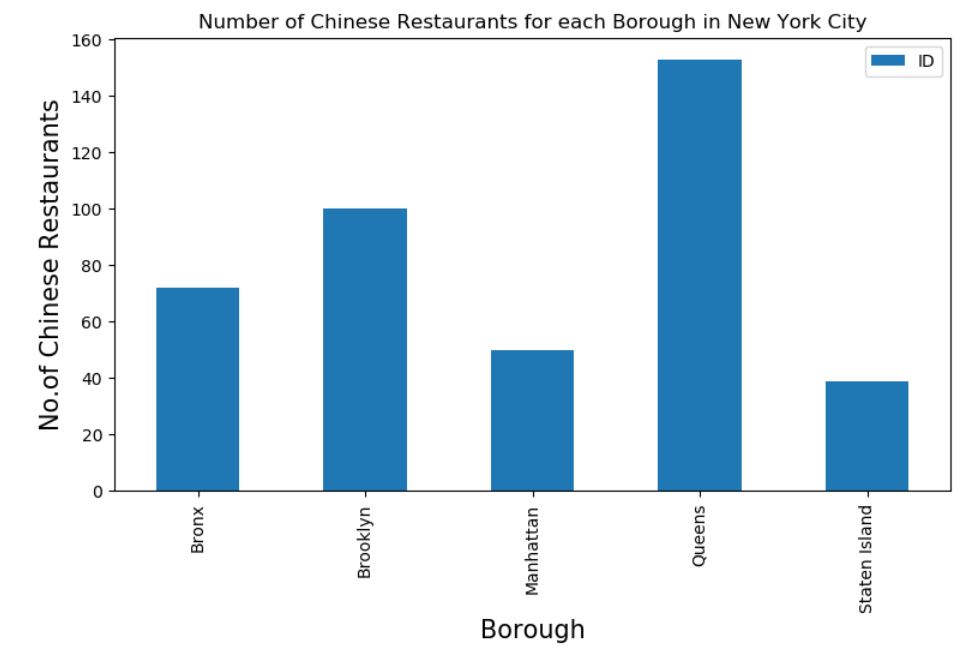
* 1. Access to Foursquare API data: This has been brought to use to explore the venues in the neighborhood in each borough of New York.

1. *Creating necessary functions to extract geographical coordinates, venue details:* We have created the following functions to help us extract call request for relevant information:
   * 1. *Geolocation(address):* This returns the latitude & longitude to the address
     2. *Get\_venues(lat,lng):* This sends a normal request to return a dataframe containing the details on list of venues i.e. ID, name & category corresponding to the latitude & longitude
     3. *Get\_venue\_details (venue\_id):* This returns a dataframe containing the details on the venue id. The details include ID, name, Likes, Rating, Tips
2. Once, we have the Neighborhood data, we explore the category of venues that are most common in these neighborhoods. We get the following list of venues with the count of venues in New York:

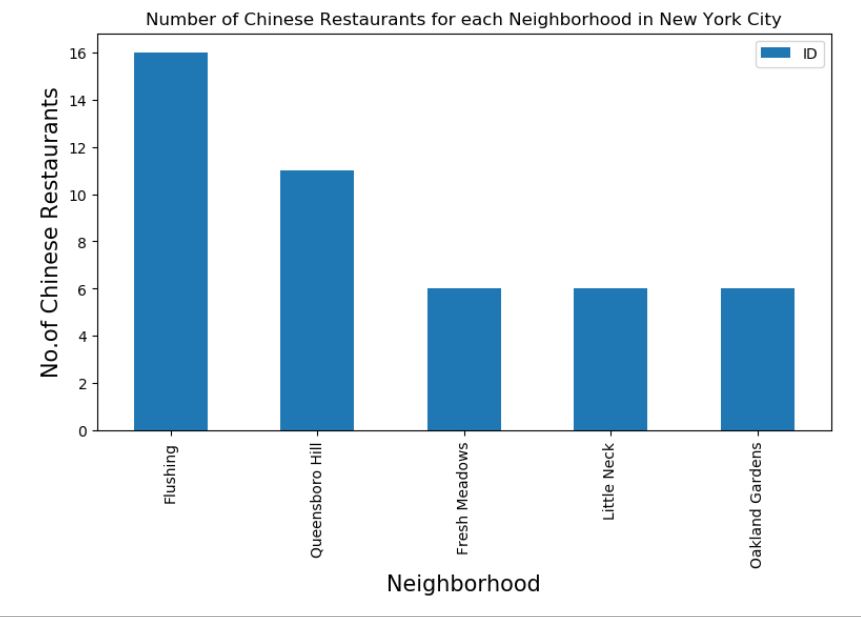
*[('Pizza Place', 1015), ('Italian Restaurant', 552), ('Coffee Shop', 535), ('Deli / Bodega', 474), ('Donut Shop', 460), ('Bakery', 447), ('Sandwich Place', 433), ('Chinese Restaurant', 421), ('Bar', 414), ('Park', 413), ('Pharmacy', 395), ('Grocery Store', 359), ('Mexican Restaurant', 329), ('Bank', 326), ('Ice Cream Shop', 325), ('Café', 322), ('American Restaurant', 319), ('Supermarket', 298), ('Fast Food Restaurant', 277), ('Gym / Fitness Center', 250)]*

As we can see, Pizza Places seem to be the most common type of restaurant, it might be very popular choice among the customers. Given that we have a limit of 500 premium calls within a day, we will analyze the Chinese restaurant which is also one of the most common cuisines in New York (8th in ranking here in terms of the number of restaurants across New York).

1. Let us assume that on analyzing other parameters, XYZ is looking at opening a Chinese restaurant in New York. Thus, a further analysis on the popularity and location of different Chinese restaurants can be leveraged.
2. Once we have finalized the cuisine as being Chinese, we explore the *Chinese restaurants by Borough*:



1. We can further look at the number of *Chinese Restaurants by Neighborhood:*



We observe that the neighborhood, Flushing in the borough Queens has a total of 16 Chinese restaurants in New York i.e. maximum among the neighborhoods

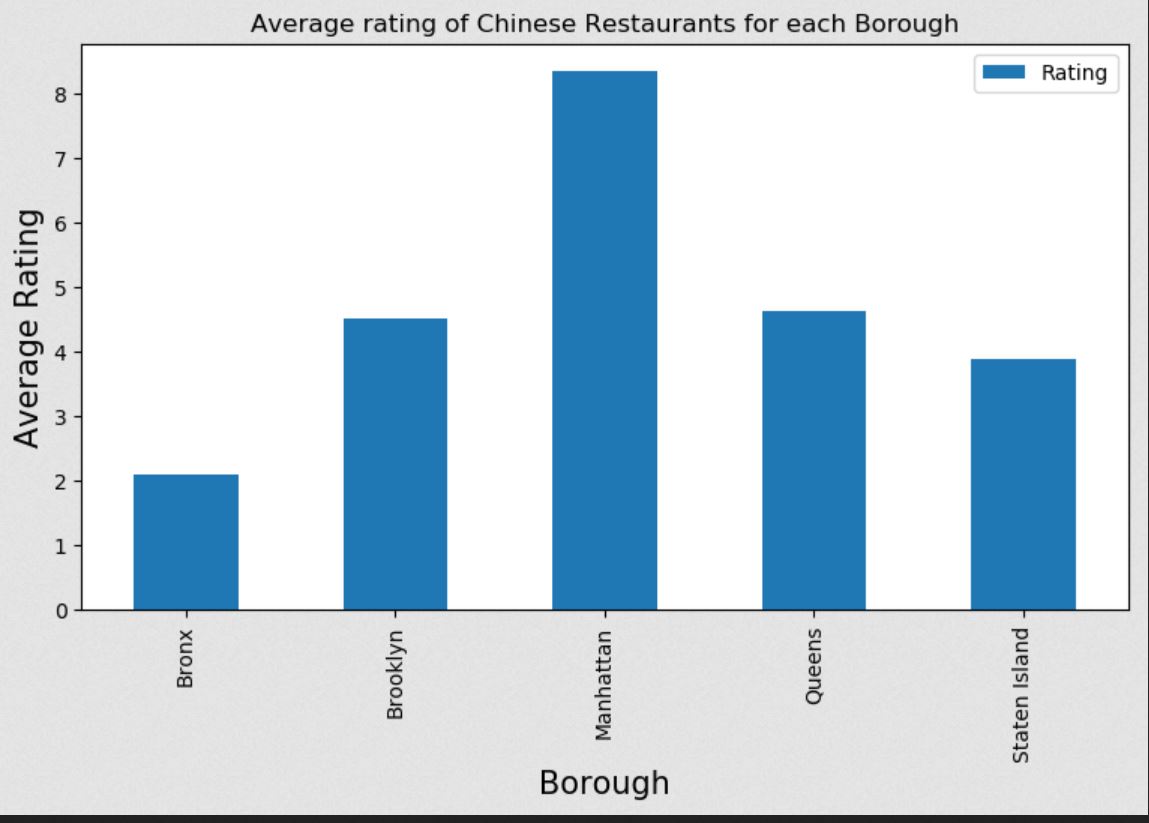
1. We can then look at the customer feedback for these restaurants: We use the get\_venue\_details to obtain the details on the restaurants from the Foursquare API.

| **Borough** | **Neighborhood** | **ID** | **Name** | **Likes** | **Rating** | **Tips** |
| --- | --- | --- | --- | --- | --- | --- |
| Bronx | Wakefield | 0 | 0 | 0 | 0 | 0 |
| Bronx | Wakefield | 0 | 0 | 0 | 0 | 0 |
| Bronx | Co-op City | 0 | 0 | 0 | 0 | 0 |
| Bronx | Co-op City | 4c66dcfaaebea593955a74d0 | Chinese Buffet | 7 | 4.9 | 10 |
| Bronx | Eastchester | 0 | 0 | 0 | 0 | 0 |

1. We looked at Restaurants with maximum Likes, Rating, Tips.
2. We then *sorted the neighborhoods & boroughs using Average Rating* as follows:

| **Neighborhood** | **Average Rating** |
| --- | --- |
| Chelsea | 9.20 |
| Flatiron | 9.20 |
| Greenpoint | 9.10 |
| Noho | 9.00 |
| North Side | 9.00 |
| Murray Hill | 9.00 |
| Midtown South | 9.00 |
| Gramercy | 9.00 |
| South Side | 9.00 |
| West Village | 8.95 |
| East Village | 8.95 |
| Upper West Side | 8.90 |
| Midtown | 8.90 |
| Little Italy | 8.80 |
| Tribeca | 8.80 |

| **Borough** | **Average Rating** |
| --- | --- |
| Manhattan | 8.356000 |
| Queens | 4.647712 |
| Brooklyn | 4.524000 |
| Staten Island | 3.884615 |
| Bronx | 2.104167 |



**Results:**

* Among boroughs, Queens is having the highest number of Chinese Restaurants.
* Within the neighborhoods, Flushing is having the highest number of Chinese Restaurants.
* The restaurant Budakkan in Chelsea, Manhattan has the highest rating, likes as well as tips.
* While Manhattan is having the highest average rating, Bronx is at the lowest.

**Discussion:**

Given the limits of data availability and to reduce the complexity we have taken into account the number of Chinese restaurants borough-wise, neighborhood-wise along with Average Rating of these restaurants as the key deciding factor to determine the potential location for the restaurant.

In a real-world problem, where if XYZ wants to open a restaurant in a location multiple parameters would be analysed which will cover not just the competitive landscape, but various other factors like the regulatory requirements in the region, XYZ’s capabilities, vision, cost-benefit analysis using future potential of market etc.

**Conclusion:**

* Chinese cuisine is one of the most famous cuisines in New York with potential to grow further.
* *Manhattan:* Although the number of Chinese restaurants in Manhattan is low as compared to other boroughs, the average rating of restaurants is highest. This may suggest high competition and lower demand. Thus, unless XYZ has plan in place to deal with the competition and XYZ is able to leverage its brand, it might not be the best option to go with.
* *Queens:* With Queens having the highest number of Chinese restaurants suggesting a very high demand which might be due to higher immigrant population from China in the neighborhood. But the competition will be stiff and it might not be much profitable.
* *Brooklyn:* The number of Chinese restaurants in Queens is much higher than Brooklyn but both have similar Average Ratings. Thus, Brooklyn appears to be having a potential for developing further in number.
* We can look at **Manhattan & Brooklyn as a potential borough** for opening the restaurant depending on the strategy in place for the competition along with other factors.
* **Depending on the competitive landscape, XYZ can look at neighborhood like Chelsea in Manhattan or Greenpoint in Brooklyn.**