Coursera Capstone Project

The Battle of Neighborhoods

Kenny Bai

June 2020

Report Content

1. **Introduction Section –** The business problems that need to be solved and the background of the problems

2. **Data Section –** Describe data requirements and sources needed to solve the problems

3. **Methodology Section –** Describe data science method we are using to solve the problems and statistical testing and machine learning technics that the report performed

4. **Result Section –** Discuss the result of the analysis

5. **Discussion Section –** Discuss the observations in the report and propose recommendations

6. **Conclusion Section –** Discuss the conclusion of the project

1. Introduction Section

1.1 Background

I am a finance analyst currently living and working in Australia. I live in Sydney with easy access to gym and wide range of restaurants. Recently, I've got an offer to be a CFO in one of the world largest FMCG companies, but the company require me to work and move to New York City and I would like to explore if I can find a place to live similar with one I live now. I have a passion in fitness and Japanese food; thus, a gym and a nice Japanese restaurant are something cannot to miss in my life. My life motto is "Eat More, Exercise More".

New York City, often called New York, has condensed population that it is easy to meet people professionally and personally. Museums, restaurants, bars and live entertainment offer great options for people love convenience. Living in the city affords one the opportunity to live a life time in a day. From a morning run in Central Park or gym, to some of the best food imaginable in the afternoon, to an evening watching comedy show in a club. Lives in New York really can be like the movies.

1.2 Problem to be solved

How to find an ideal neighborhood in New York City that meet the following conditions:

- The Borough it belongs to has good safety

- Nearby location has different good rating gym facilities
- Nearby location has different good rating Japanese restaurants
- List and visualize all the major parts of New York has good Japanese restaurant

2. Data Section

2.1 Data Requirements

- New York City Borough and Neighborhood information with Longitude and Latitude for visualization and mapping purpose.
- Restaurants detailed information (e.g. Japanese, Indian and etc) and gym in New York city and their locations to be located on Map
- Crime rate data by Borough level to indicate safety of each Borough
- Location of the future office (Time Square) for reference

2.1 Data Sources, Data Processing and Tools used

- New York City Borough and Neighborhood information with Longitude and Latitude for visualization and mapping purpose. **Data Sources:** https://cocl.us/new_york_dataset This data set has a total of 5 boroughs and 306 neighborhoods as well as the the latitude and longitude coordinates of each neighborhood.

- Restaurants detailed information (e.g. Japanese, Indian and etc) and gym in New York city and their locations to be located on Map. **Data Sources:** Foursquare API With API function in Foursquare app, we are able to locate gym and diverse restaurants and even the rating of them.

- Crime rate data indicate safety of each Borough. **Data Source:** https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD This data set shows list of every shooting incident that occurred in NYC going back to 2006 through the end of 2019, indicating the safety of NYC by Borough level.

- Location of the future office (near Time Square) for reference. **Data Source:** Google Map 1540 Broadway, New York, NY 10036, United States

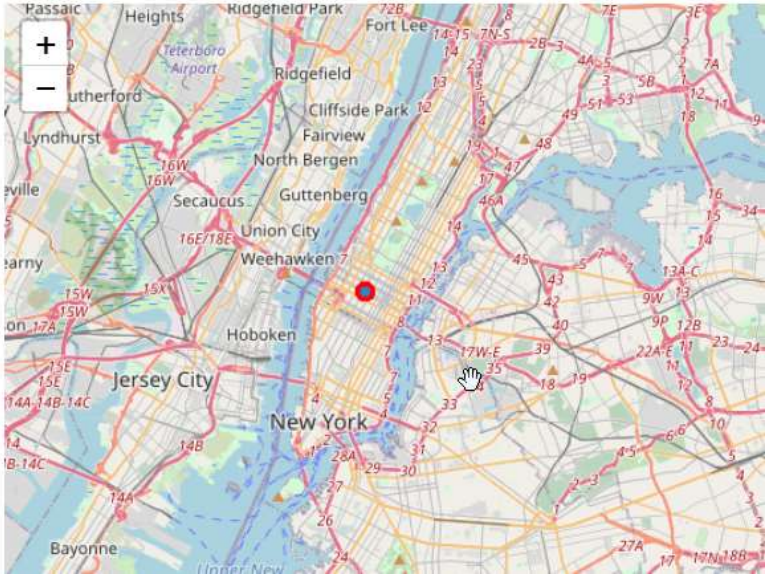
3. Methodology Section

3.1 General information of New York City Data

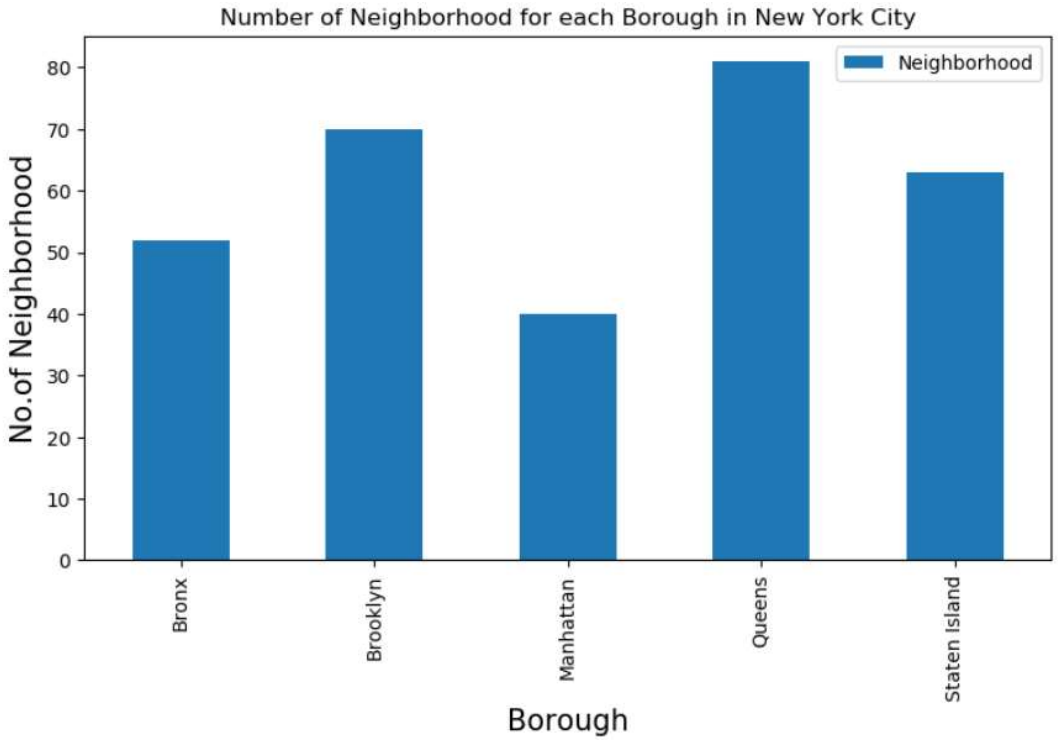
Firstly, we begin by collecting the New York City data from the above data source link and perform general analysis on the size of the data. We can conclude that New York City has today 5 Boroughs and 306 Neighborhoods. Also, we plot the future office location to the map to visualize the location.

```
In [50]: neighborhoods.shape
Out[50]: (306, 4)
```

Out[8]:



Then, we perform bar chart analysis on numbers of Neighborhoods.



3.2 Analysis on Safety in New York City

Firstly, we import the data from New York Police Department to show list of every shooting incident that occurred in NYC going back to 2006 through the end of 2019.

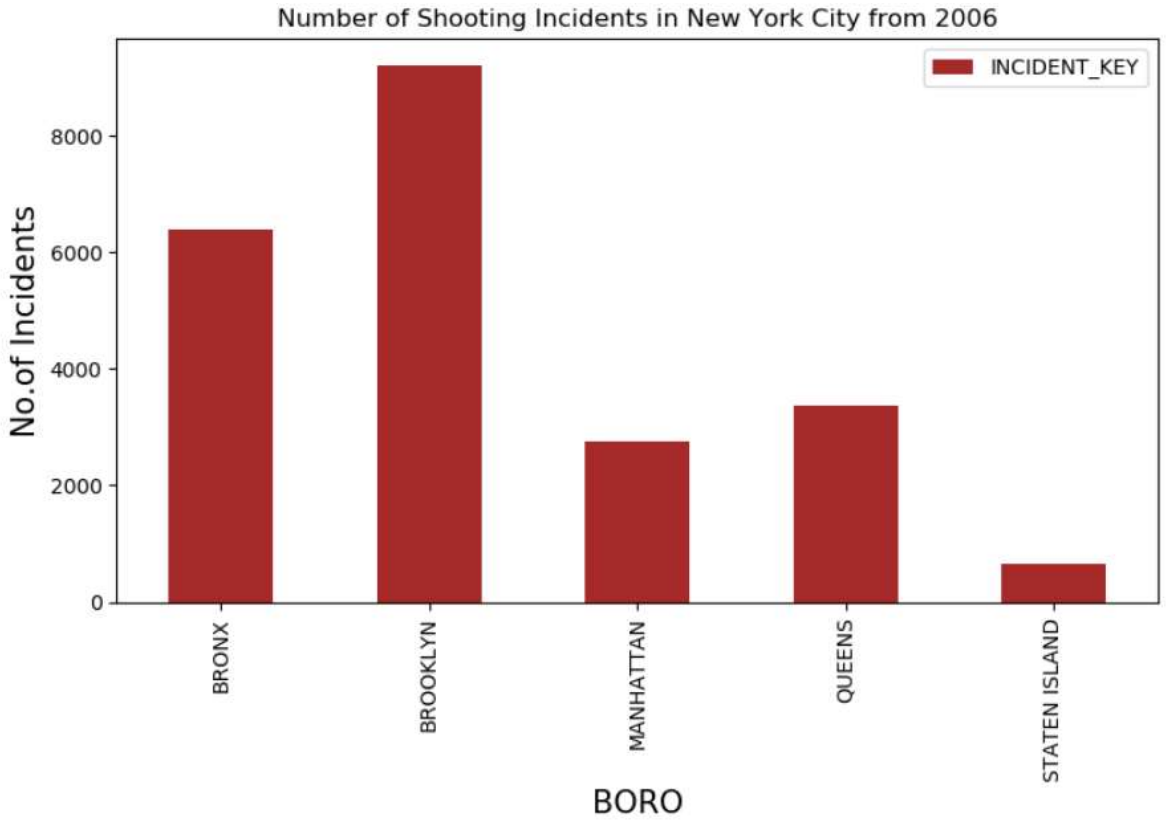
In [11]:
```
#Download crime data
import pandas as pd
import io
import requests
url="https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD"
s=requests.get(url).content
crimedf = pd.read_csv(io.StringIO(s.decode('utf-8')))

crimedf.head()
```

Out[11]:

	INCIDENT_KEY	OCCUR_DATE	OCCUR_TIME	BORO	PRECINCT	JURISDICTION_CODE	LOCATION_DESC	STATISTICAL_MURDER_FLA
0	169180023	09/09/2017	3:57:00	STATEN ISLAND	120	0.0	NaN	False
1	169180027	09/09/2017	22:15:00	BRONX	50	0.0	NaN	False
2	169180025	09/09/2017	18:35:00	BROOKLYN	79	2.0	MULTI DWELL - PUBLIC HOUS	False
3	169180024	09/09/2017	15:20:00	BROOKLYN	67	0.0	NaN	False
4	169180022	09/09/2017	0:20:00	BROOKLYN	61	2.0	MULTI DWELL - PUBLIC HOUS	False

After filtering the data, we can find that numbers of shooting incident in New York City that can indicate general safety level in each Borough. The safer the Borough, the better place for individual to live in. And the below results shows that Brooklyn has the highest shooting incidents in NYC while Staten Island and Manhattan shows relatively low numbers of incidents.



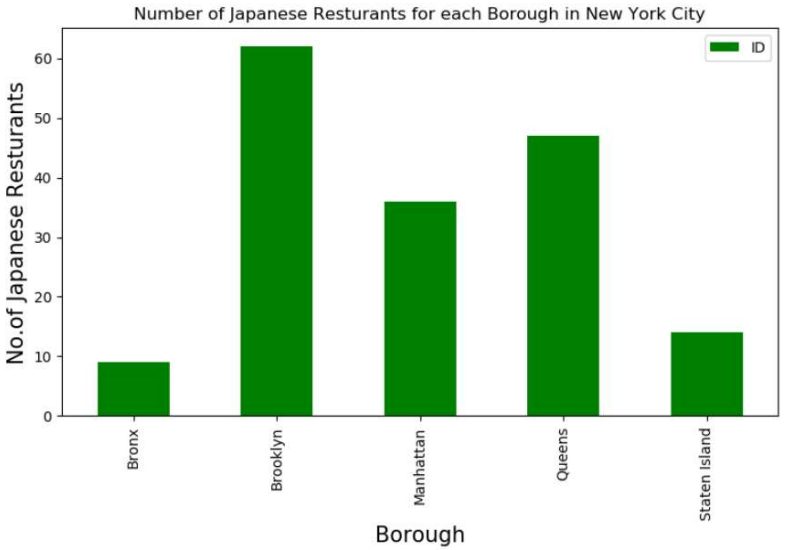
3.3 Analysis on Japanese Restaurant and Gym facility

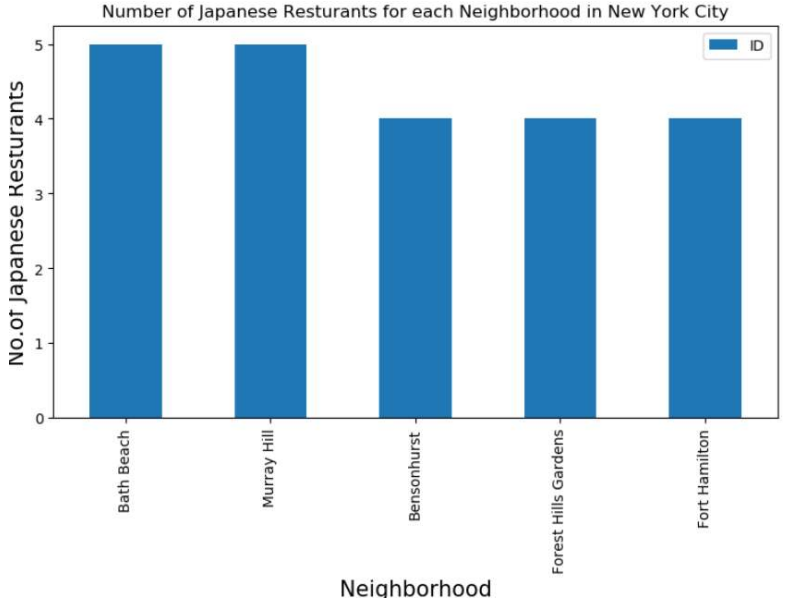
Next we use Four Square API function to help us to locate all the Japanese Restaurant in all the neighborhoods to find out which neighborhoods has the highest numbers of Japanese restaurant and Gym. Below is a view on how many Japanese restaurant in each neighborhood.

```
    print('(',count,'/',len(neighborhoods),')','Japanese restaurant in '+Neighb
    for resturant_detail in jap_rest.values.tolist():
        id, name , category = resturant_detail
        japrest = japrest.append({'Borough': Borough,
                                  'Neighborhood': Neighborhood,
                                  'ID': id,
                                  'Name' : name
                                 }, ignore_index=True)
    count+=1
```

```
( 1 / 306 ) Japanese restaurant in Wakefield, Bronx:0
( 2 / 306 ) Japanese restaurant in Co-op City, Bronx:0
( 3 / 306 ) Japanese restaurant in Eastchester, Bronx:0
( 4 / 306 ) Japanese restaurant in Fieldston, Bronx:0
( 5 / 306 ) Japanese restaurant in Riverdale, Bronx:2
( 6 / 306 ) Japanese restaurant in Kingsbridge, Bronx:2
( 7 / 306 ) Japanese restaurant in Marble Hill, Manhattan:0
( 8 / 306 ) Japanese restaurant in Woodlawn, Bronx:0
( 9 / 306 ) Japanese restaurant in Norwood, Bronx:0
( 10 / 306 ) Japanese restaurant in Williamsbridge, Bronx:0
( 11 / 306 ) Japanese restaurant in Baychester, Bronx:0
( 12 / 306 ) Japanese restaurant in Pelham Parkway, Bronx:0
( 13 / 306 ) Japanese restaurant in City Island, Bronx:1
( 14 / 306 ) Japanese restaurant in Bedford Park, Bronx:0
( 15 / 306 ) Japanese restaurant in University Heights, Bronx:0
( 16 / 306 ) Japanese restaurant in Morris Heights, Bronx:0
```

Below Bar chart 1 shows that Brooklyn, Queens, Manhattan have more Japanese restaurant than the other two boroughs. In terms of neighborhood, Bar chart 2 shows Bath Beach and Murray Hill have highest numbers of Japanese restaurants.

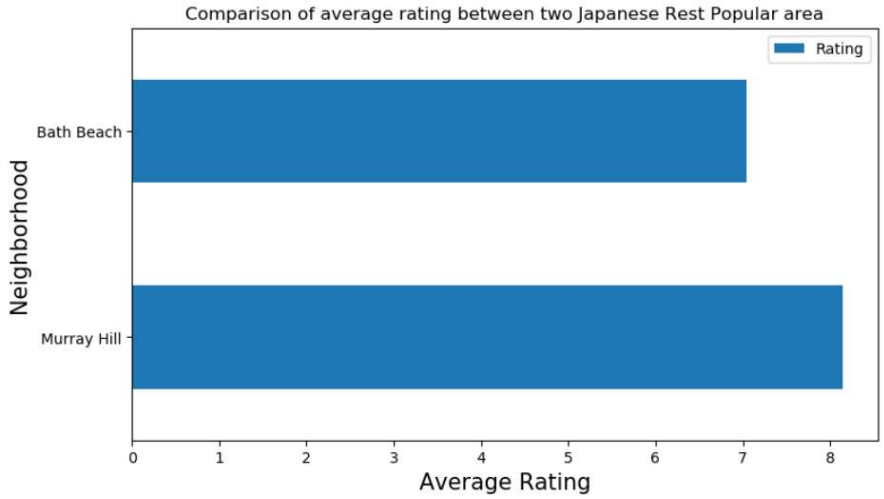

Number of Japanese Resturants for each Borough in New York City
ID
No.of Japanese Resturants
0
10
20
30
40
50
60
Bronx
Brooklyn
Manhattan
Queens
Staten Island
Borough


Number of Japanese Resturants for each Neighborhood in New York City
ID
No.of Japanese Resturants
0
1
2
3
4
5
Bath Beach
Murray Hill
Bensonhurst
Forest Hills Gardens
Fort Hamilton
Neighborhood

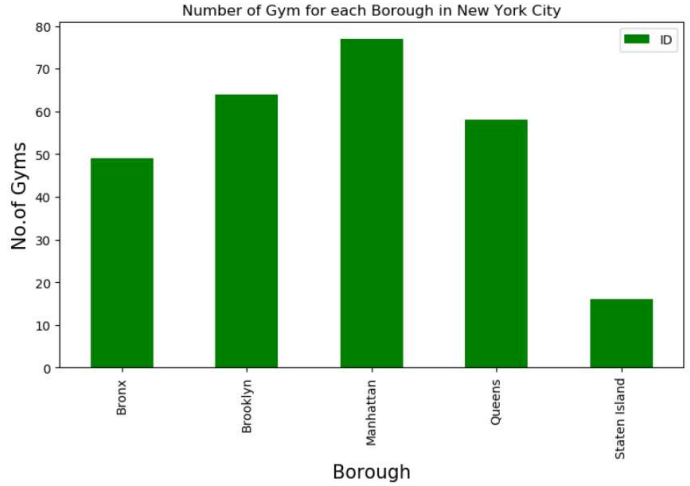
Next we pick 2 of these neighborhoods to look at the rating of the venues by using four square api.

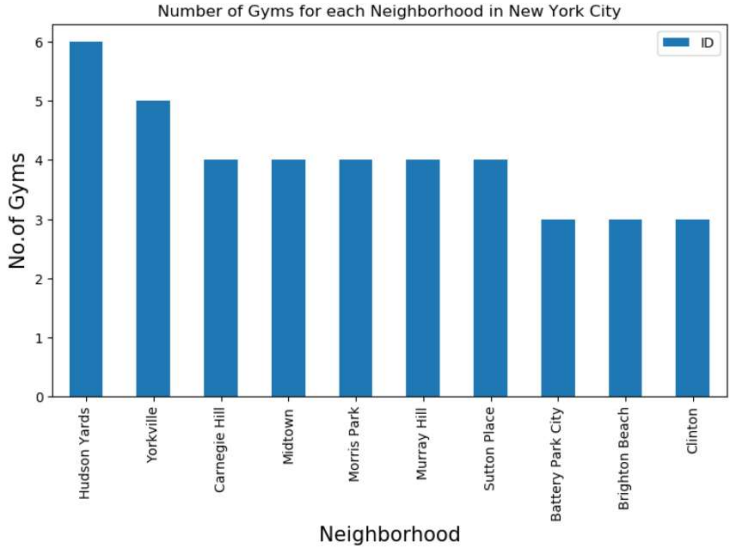
```
                         ID          Name  Likes  Rating  Tips
0  540dc46d498e86f0e5059dd2  Zuma New York    454     8.4    84
( 1 / 10 ) processed
                         ID            Name  Likes  Rating  Tips
0  591caee89deb7d0f69be77a2  Omusubi Gonbei     14     8.3     1
( 2 / 10 ) processed
                         ID              Name  Likes  Rating  Tips
0  49db8b67f964a520d85e1fe3  Aburiya Kinnosuke    279     8.8    91
( 3 / 10 ) processed
                         ID            Name  Likes  Rating  Tips
0  4bc8eee83740b713fcbe5d65  Northern Sushi      6     7.9     2
( 4 / 10 ) processed
                         ID        Name  Likes  Rating  Tips
0  5cf07afa66f3cd002c8b3ae6  Izakaya Mew      7     7.3     1
```

We can see Murray Hill's Japanese restaurants have higher average rating.

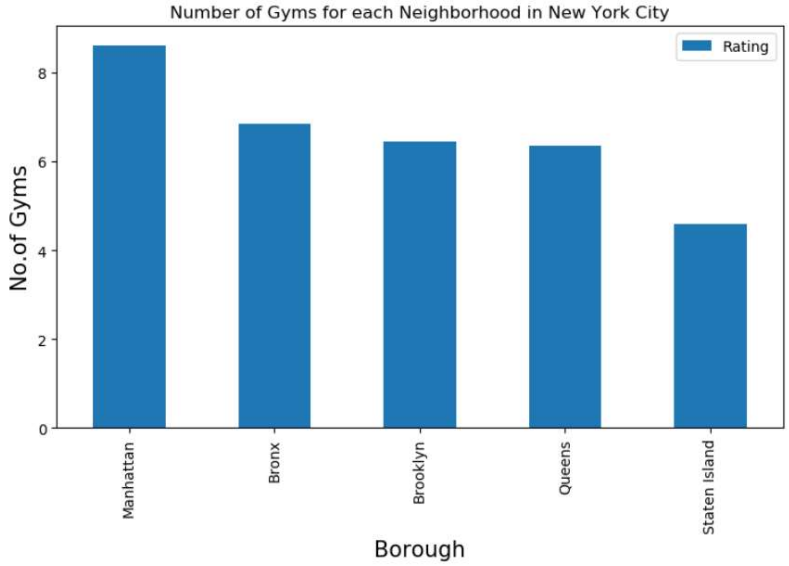


We perform similar analysis on Gym and found out that Manhattan has the highest numbers of gym facilities. Murray Hill as the second highest gym facility neighborhoods.





Summarizing the rating of gym in all boroughs, we can see the below. Manhattan's gym facility is good.



Murray Hill's Gyms' rating are above the average.

```
In [120]: MHgymrate= gymrate[gymrate['Neighborhood']=='Murray Hill']
          MHgymrate.head(20)
```

Out[120]:

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
116	Manhattan	Murray Hill	49c943d2f964a5204f581fe3	Equinox Park Avenue	364	8.5	60
117	Manhattan	Murray Hill	42af6f80f964a5205a251fe3	Equinox East 44th Street	259	9.1	38
118	Manhattan	Murray Hill	4c780d08566db60c5a4f400e	Manhattan Place Health Club	8	8.5	0
119	Manhattan	Murray Hill	4a047a18f964a520ff711fe3	Equinox East 43rd Street	120	8.6	18

4. Result Section

To summarize the results, we can see that from the safety point of view, Manhattan and Staten Island are two Borough with less shooting incidents. And in terms of Japanese restaurants, it turns out Brooklyn has highest numbers of Japanese Restaurant. And Bath Beach in Brooklyn and Murray Hill in Manhattan have the highest numbers of Japanese Restaurant. When it comes to rating, the ones in Murray Hill are better than those in Bath Beach. Regarding Gym facility, it is obvious that Manhattan excels. And when we look at Murray Hill's Gym facilities specifically, we can find that it has outstanding rating as well.

5. Discussion Section

From my personal point of view, combining all the observations of the results. The proposed neighborhood in NYC to live in is Murray Hill as it has great Japanese restaurants and gym facilities, plus it has good indication of safety in that Borough it belongs to.

6. Conclusion Section

This report has room for improvement and some of the assumptions are quite subjective. With more data point and time in the future, we can extend the topic to price of living to have more holistic views to the solution.