The Battle of Neighbourhood

(Capstone Project Final Report)

Introduction

- New York is vibrant and fast-growing city in the world with different ethnic population from world over.
- With the different ethnicity there is wide difference in culture and eating habits.
- New York City is composed of five boroughs, each of which is a county of the State of New York.
- The five boroughs—Brooklyn, Queens, Manhattan, the Bronx, and Staten Island.
- There are many bakers in New York City, with wide variety baked item.

Problem

- The problem is to identify and provide better insight to the bakers who are in baking business in New York and its neighbourhood.
- Identify best suitable locations for bakers in New York
- Identify locations with potential Bakers
- Identify locations where there less Bakers.

Data Collection

1. New York City data that contains list Boroughs, Neighbourhoods along with their latitude and longitude.
    - Data source : https://cocl.us/new_york_dataset
    - Description: This data set contains the required information. And we will use this data set to explore various neighbourhoods of New York City.
2. The details of Bakers in each neighbourhood of New York city
    - Data source : Foursquare API
    - Description: Foursquare API is used to get all the venues in each neighbourhood of New York. And filter is applied to select Bakery as venues.

Methodology

1. We begin by collecting the New York city data from the following link "https://cocl.us/new_york_dataset“

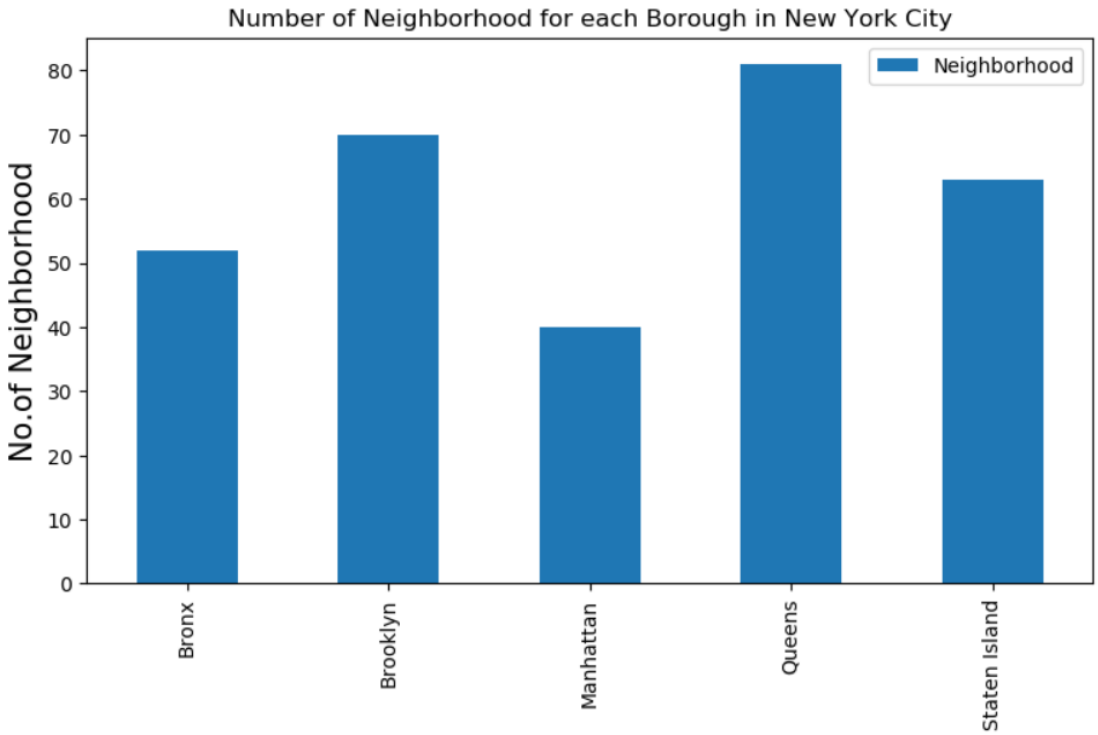
```
In [20]: new_york_data.head()
```

Out[20]:

	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

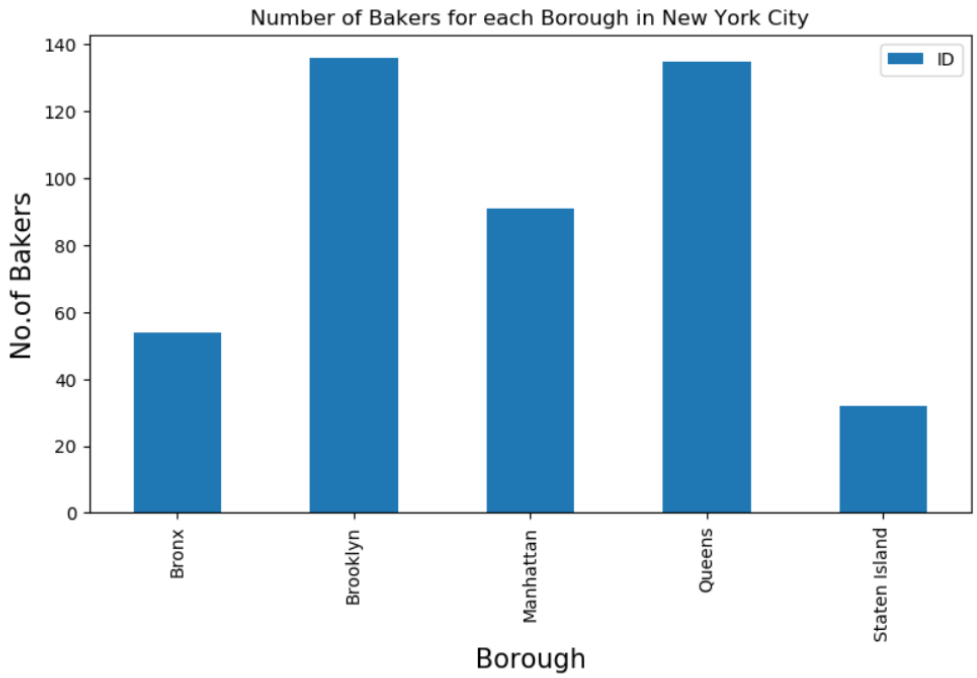
```
In [10]: new_york_data.shape
Out[10]: (306, 4)
```

2. The number of neighbourhoods for each borough is visualised



3. We then filter the bakery venues from data frame.

```
( 1 / 306 ) Bakery in Wakefield, Bronx:2
( 2 / 306 ) Bakery in Co-op City, Bronx:2
( 3 / 306 ) Bakery in Eastchester, Bronx:1
( 4 / 306 ) Bakery in Fieldston, Bronx:1
( 5 / 306 ) Bakery in Riverdale, Bronx:0
( 6 / 306 ) Bakery in Kingsbridge, Bronx:2
( 7 / 306 ) Bakery in Marble Hill, Manhattan:2
( 8 / 306 ) Bakery in Woodlawn, Bronx:2
( 9 / 306 ) Bakery in Norwood, Bronx:0
( 10 / 306 ) Bakery in Williamsbridge, Bronx:3
( 11 / 306 ) Bakery in Baychester, Bronx:2
( 12 / 306 ) Bakery in Pelham Parkway, Bronx:3
( 13 / 306 ) Bakery in City Island, Bronx:0
( 14 / 306 ) Bakery in Bedford Park, Bronx:0
( 15 / 306 ) Bakery in University Heights, Bronx:1
( 16 / 306 ) Bakery in Morris Heights, Bronx:0
( 17 / 306 ) Bakery in Fordham, Bronx:2
( 18 / 306 ) Bakery in East Tremont, Bronx:0
( 19 / 306 ) Bakery in West Farms, Bronx:0
```



4. Using FourSquare API the likes,tips, rating of the bakery are accessed.

```
In [170]: bakery_stats_ny_csv.head()
```

Out[170]:

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Wakefield	4cd3edd25366548131a129ed	E&L Bakery	7	6.5	4
1	Bronx	Wakefield	0	0	0	0.0	0
2	Bronx	Co-op City	4a720b8ef964a52036da1fe3	Panera Bread	118	8.3	38
3	Bronx	Co-op City	4c433c8fce54e21ef8c50d1a	The Cookie Factory	7	8.1	4
4	Bronx	Co-op City	0	0	0	0.0	0

5. We will then sort Neighbourhoods and Borough the data keeping rating as the constraint.

```
In [178]: ny_neighborhood_stats_csv.sort_values(['Average Rating'],ascending=False).head(10)
```

Out[178]:

	Neighborhood	Average Rating
25	Woodlawn	8.800000
14	Melrose	8.800000
12	Longwood	8.800000
8	Fieldston	8.700000
0	Bay Ridge	8.500000
10	Hunts Point	8.100000
2	Belmont	8.083333
5	Country Club	8.033333
17	Pelham Bay	8.033333
23	Westchester Square	7.900000

6. Next we will consider all the neighbourhoods with average rating greater or equal 7.0 to plot on map

```
In [183]: ny_neighborhood_stats1=ny_neighborhood_stats_csv[ny_neighborhood_stats_csv['Average Rating']>=7.0]
```

7. We will join this dataset to original New York data to get longitude and latitude

Out[187]:

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Brooklyn	Bay Ridge	40.625801	-74.030621	8.500000
1	Bronx	Baychester	40.866858	-73.835798	7.550000
2	Bronx	Belmont	40.857277	-73.888452	8.083333
3	Bronx	Concourse	40.834284	-73.915589	7.100000
4	Bronx	Country Club	40.844246	-73.824099	8.033333
5	Bronx	Fieldston	40.895437	-73.905643	8.700000
6	Bronx	Fordham	40.860997	-73.896427	7.700000
7	Bronx	Hunts Point	40.809730	-73.883315	8.100000
8	Bronx	Kingsbridge	40.881687	-73.902818	7.550000
9	Bronx	Longwood	40.815099	-73.895788	8.800000
10	Manhattan	Marble Hill	40.876551	-73.910660	7.100000
11	Bronx	Melrose	40.819754	-73.909422	8.800000
12	Bronx	Pelham Bay	40.850641	-73.832074	8.033333
13	Bronx	Westchester Square	40.840619	-73.842194	7.900000
14	Bronx	Williamsbridge	40.881039	-73.857446	7.566667
15	Bronx	Woodlawn	40.898273	-73.867315	8.800000

8. Finally, we will visualize the Neighbourhoods and Borough based on average Rating using python's Folium library.

Result

So now we can answer the questions asked above in the Problem section of the notebook.

From our anaysis the answers to the above questions are:

A1) Baybridge(Brooklyn), Woodlawn(Bronx), Melrose(Bronx),(Bronx) are some of the best neighborhoods for bakers.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Brooklyn	Bay Ridge	40.625801	-74.030621	8.500000
1	Bronx	Baychester	40.866858	-73.835798	7.550000
2	Bronx	Belmont	40.857277	-73.888452	8.083333
3	Bronx	Concourse	40.834284	-73.915589	7.100000
4	Bronx	Country Club	40.844246	-73.824099	8.033333
5	Bronx	Fieldston	40.895437	-73.905643	8.700000
6	Bronx	Fordham	40.860997	-73.896427	7.700000
7	Bronx	Hunts Point	40.809730	-73.883315	8.100000
8	Bronx	Kingsbridge	40.881687	-73.902818	7.550000
9	Bronx	Longwood	40.815099	-73.895788	8.800000
10	Manhattan	Marble Hill	40.876551	-73.910660	7.100000
11	Bronx	Melrose	40.819754	-73.909422	8.800000
12	Bronx	Pelham Bay	40.850641	-73.832074	8.033333
13	Bronx	Westchester Square	40.840619	-73.842194	7.900000
14	Bronx	Williamsbridge	40.881039	-73.857446	7.566667
15	Bronx	Woodlawn	40.898273	-73.867315	8.800000

A2) Brooklyn have potential bakers.

A3) Staten Island ranks last in number of bakers.

A4) Bronx is the best place to stay if you prefer bakery.

Conclusion

As there is limit for access of data using Foursquare API, there is a wealth of information that can be accessed and analysed for productive purposes.