

IBM Data Science Professional Specialization Capstone
Final Report

**The Battle of the Neighborhood:
Launching a new business in an already emerged
market space**

Sean Gill - June 2019

The intention of this report is to attempt to solve a problem or question through applying data science methodology on location data received from Foursquare API.

1. Introduction & Business Queries

Introduction:

New York City is the most populous city in the United States of America as reported by the United States Census Bureau in 2017 at approximately 8.623 million with an upwards trend. While being quite populated, New York City has been the continuous homes to many immigrants and cultures which have flourished for decades. This trend can be seen in New York City's MOIA Annual Report of 2018. (We will break down the demographics of these immigrants in the data breakdown section.) Being such the culture hub, it is no wonder that New York City is an international hotbed of opportunity. Any business would need to strategize carefully to maximize their reach and exploit the diversity present in New York City's many neighborhoods. This will ultimately reduce the possibility of risk for any new business venture.

With such a diverse population and culture, New York City has through a wide range of cuisines and restaurants that can cater to varying price points. In this project, our business is a restaurant that creates homemade Mexican food that welcomes influences from various Latino cultures.

The clients for this project are a small family wants to open their own homemade family-owned dine-in restaurant in a New York City neighborhood with a quick

connection to the downtown area. As this is their very first attempt at a business venture, they have requested insight on several issues. This process would also be beneficial for anyone who wants to start a new business in New York City or any other major city. This brings us to questions of logistics .

Business Queries:

When delving into the logistics, the future restauranteurs had three main areas of concern which they hoped data analytics could shed some light on.

Here are those concerns:

- ▶ Where would a brand new Mexican/ Latino restaurant find the space to distinguish themselves from the numerous competition and maximize income?
- ▶ Which neighborhood might have a supportive demographic (presumably Latino/ South American as it would be similar to the restaurant's culinary influence) to help sustain business?
- ▶ How much competition might they face regarding similar restaurants? (Regarding proximity to NYC's downtown area)

2. Data Acquisition and Preparation

The neighborhoods of New York City have a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, I have created a dataset that contains the 5 boroughs and the neighborhoods that exist in

each borough as well as the the latitude and longitude coordinates of each neighborhood (The link to such datasets: https://geo.nyu.edu/catalog/nyu_2451_34572

- https://cocl.us/new_york_dataset) in order to pin-point possible locations.

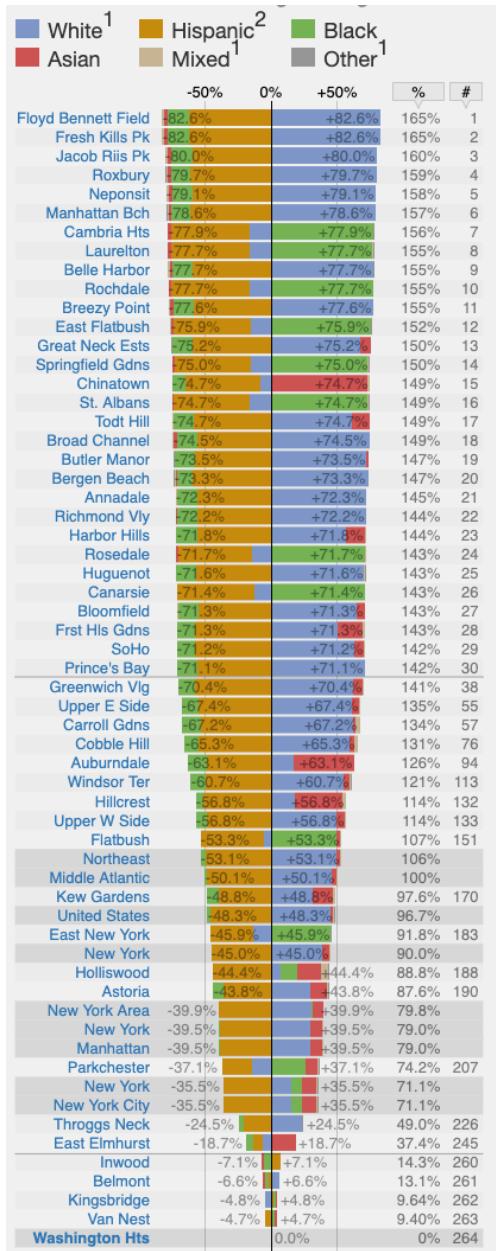


Figure 1

New York City's MOIA Annual Report 2018 was used to examine possible target demographics and extrapolate which backgrounds that might show the most interest. As for any business, having a clientele lays the foundation for progressing towards a bountiful economic future. Through the MOIA's report, we can find the top ten countries of origin for foreign born city residents (Table 1) and use them to base several assertions on. However from this report it is important to remember that we cannot

Table 1 **Country of origin for NYC immigrants**

| | | Number | % of all foreign-born |
|---|--------------------|---------|-----------------------|
| 0 | Dominican Republic | 422,244 | 13.5% |
| 1 | China | 318,581 | 10.2% |
| 2 | Mexico | 183,197 | 5.8% |
| 3 | Jamaica | 176,718 | 5.6% |
| 4 | Guyana | 136,760 | 4.4% |
| 5 | Ecuador | 129,693 | 4.1% |
| 6 | Haiti | 88,383 | 2.8% |
| 7 | Trinidad & Tobago | 83,222 | 2.7% |
| 8 | Bangladesh | 80,448 | 2.6% |
| 9 | India | 76,424 | 2.4% |

precisely confirm the distribution of these residents and their subsequent families in exact neighborhoods. The closest data found supporting any form of distribution of ethnicity in local neighborhoods from www.statisticalatlas.com, which collects its data from the U.S. Census, which show some form of local distribution through a point of deviation as compared to Washington Heights, a New York City neighborhood as seen in *Figure 1*. In lieu of pure unbiased data, the precise location of the demographics was

deemed ultimately unnecessary. As long as the demographic exists in a localized area, the possibility of interaction with a business or product would be approximately the same.

Foursquare's services were used to explore and examine New York City's numerous neighborhoods. Foursquare data was also deployed to parse out Venues, User tips, Feedback, location and Ratings of Neighboring restaurants that would be eventual competition. As seen below in Figure 2:

Figure 2

| | name | categories | address | cc | city | country | crossStreet | distance | formattedAddress |
|---|---------------------------------|--------------------|----------------------|-----------|-------------|----------------|-------------------------------|-----------------|---|
| 0 | Chipotle Mexican Grill | Mexican Restaurant | 111 Fulton St | US | New York | United States | at William St | 333 | [111 Fulton St (at William St), New York, NY 1... |
| 1 | Mad Dog & Beans Mexican Cantina | Mexican Restaurant | 83 Pearl St | US | New York | United States | at Stone St | 1010 | [83 Pearl St (at Stone St), New York, NY 10004... |
| 2 | Chipotle Mexican Grill | Mexican Restaurant | 2 Broadway Frnt 4 | US | New York | United States | btwn Beaver & Stone St | 1045 | [2 Broadway Frnt 4 (btwn Beaver & Stone St), N... |
| 3 | Chipotle Mexican Grill | Burrito Place | 625 Broadway | US | New York | United States | btwn Bleeker & Houston St | 1688 | [625 Broadway (btwn Bleeker & Houston St), Ne... |
| 4 | Chipotle Mexican Grill | Mexican Restaurant | 200 Varick St Frnt B | US | New York | United States | Varick between Houston & King | 1732 | [200 Varick St Frnt B (Varick between Houston ... |

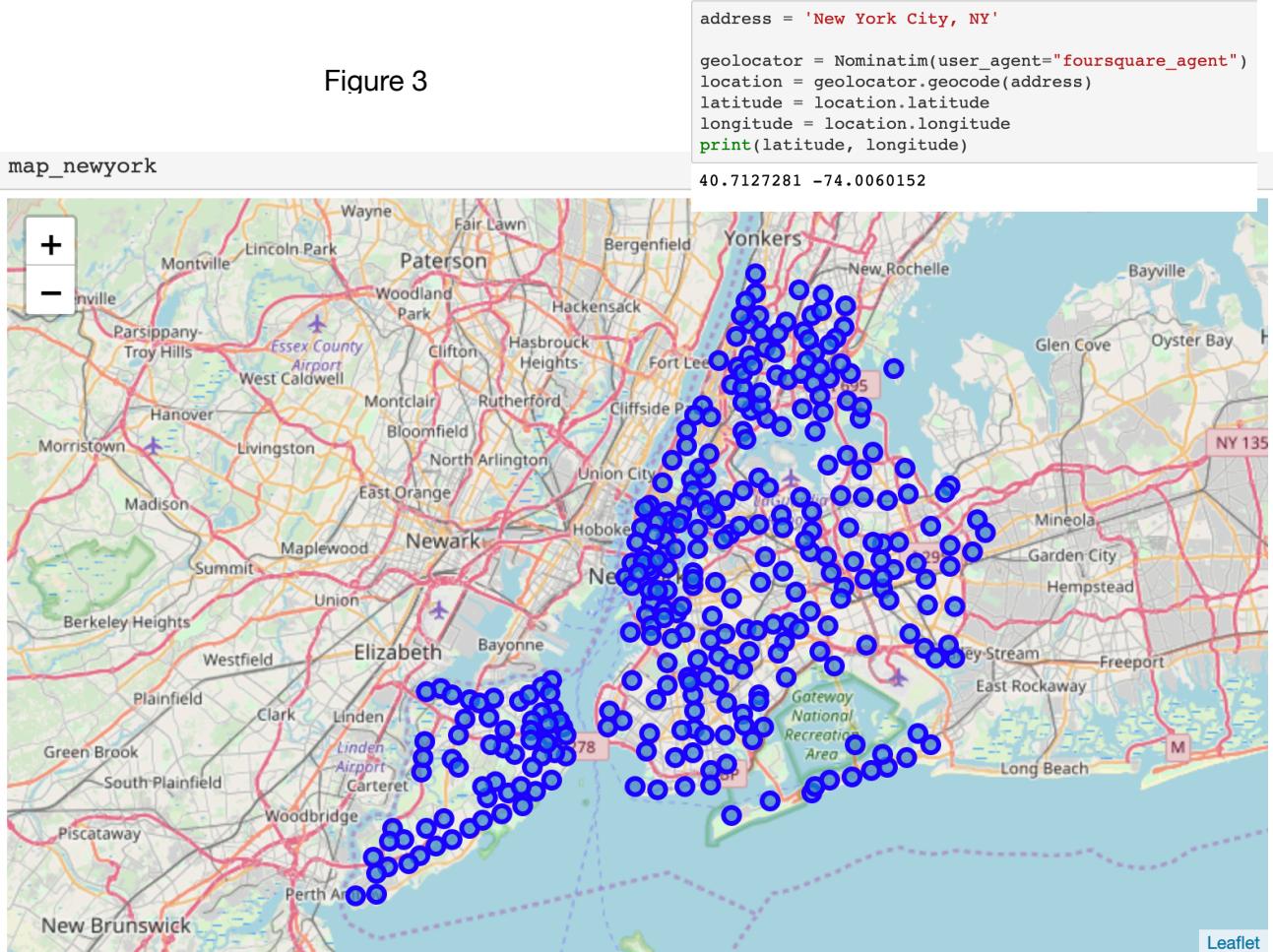
3. Methodology

Visualization:

Find Coordinates for New York City's downtown area.

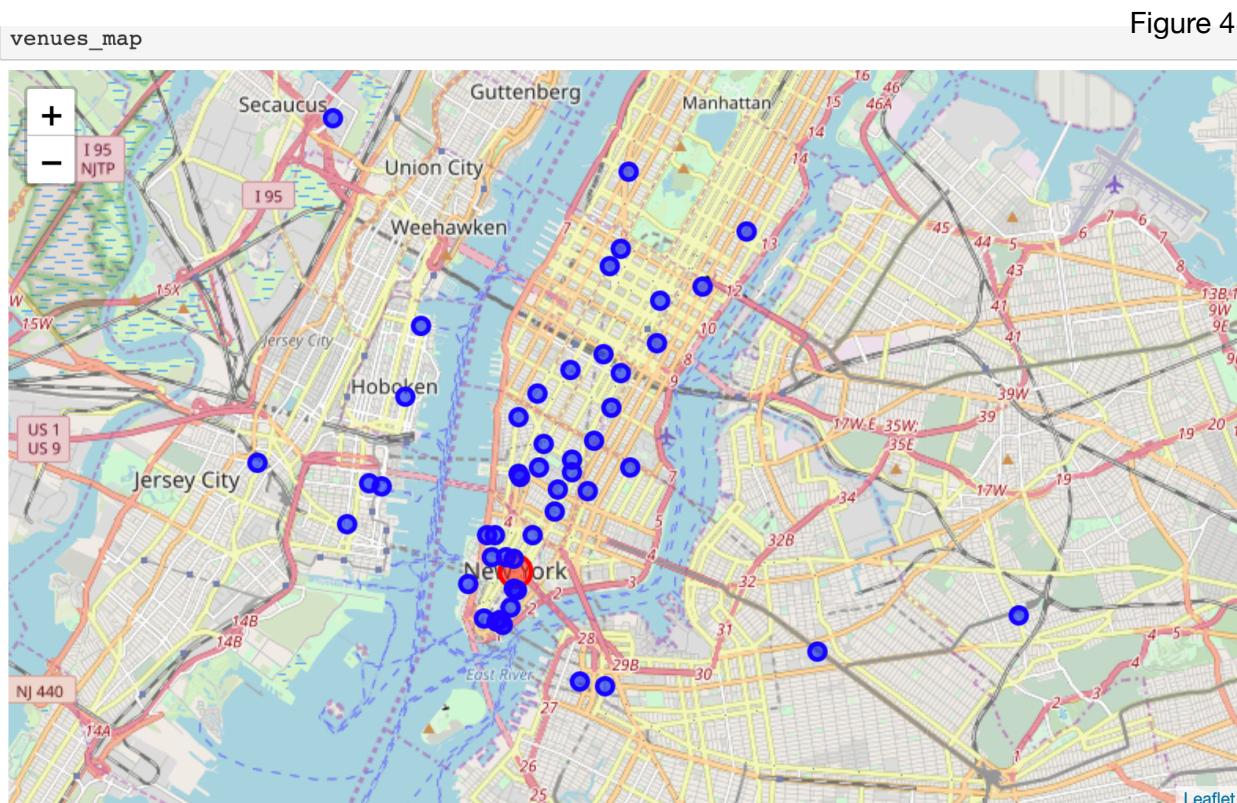
From these coordinates we can define our prospective neighborhoods of New York City using latitude and longitude with our JSON file, as seen in *Figure 3*.

Figure 3



4. Results

With this parameter set, Foursquare's API was used to overlay similar restaurants with similar qualities of the perspective restaurant. Mexican restaurants with a relative close proximity to the downtown area. While using Foursquare, related Mexican restaurants were gathered in a 16000 meter radius (~10 miles) The red circle denotes the downtown area, while blue circles denotes Mexican restaurants in Figure 4. This segmentation was set with a limit of 60 however only 50 were returned in the request



as in Figure 5. In terms of the boroughs, the radius affected 3 boroughs; Manhattan, Brooklyn and Queens. As per Figure 4, Manhattan holds the most values, then Brooklyn and one outlier in Queens. It is important to remember that these restaurants were selected through the radius and restaurants committed to Foursquare's API. There could be more related restaurants outside the scope of these parameters.

```
In [49]: dataframe_filtered.name
```

```
Out[49]: 0           Chipotle Mexican Grill
1           Mad Dog & Beans Mexican Cantina
2           Chipotle Mexican Grill
3           Chipotle Mexican Grill
4           Chipotle Mexican Grill
5           Dahlia's Mexican Restaurant
6           Chipotle Mexican Grill
7           Chipotle Mexican Grill
8           Órale! Mexican Kitchen
9           Chipotle Mexican Grill
10          Chipotle Mexican Grill
11          Chipotle Mexican Grill
12          Chipotle Mexican Grill
13          Chipotle Mexican Grill
14          Chipotle Mexican Grill
15          Chipotle Mexican Grill
16          Sinigual Contemporary Mexican Cuisine
17          QDOBA Mexican Eats
18          Chipotle Mexican Grill
19          Chipotle Mexican Grill
20          Chipotle Mexican Grill
21          Chipotle Mexican Grill
22          QDOBA Mexican Eats
23          Chipotle Mexican Grill
24          Zaragoza Mexican Deli-Grocery
25          Chipotle Mexican Grill
26          Panchito's Mexican Restaurant
27          Chipotle Mexican Grill
28          Chipotle Mexican Grill
29          La Sirena Mexican Folk Art
30          Chipotle Mexican Grill
..
```

Figure 5

```
dataframe_filtered.shape
```

```
: (50, 16)
```

From the reported data, we can see that the first snapshot of data is dominated by restaurant chains. Namely, Chipotle Mexican Grill and QDOBA Mexican Eats. This could prove a special business angle for our clients, Homemade food vs Fast-food.

From the previous figure, (Figure 4) we can see that Manhattan is completely occupied by Mexican style restaurants. Our clients want to stay close to the city center while being offered an opportunity where their restaurant can emerge among the local competition. From this assertion, West Brooklyn satisfies these requirements as it provides direct links back to the downtown area via bridges.

5. Conclusion

For our new clients, West Brooklyn appears to be the best starting location. It is close to the city center with direct access to the downtown area via two bridges. The only substantial competition was with Guadalupana Mexican Bakery with a rating of 8.3/10.

```
result = requests.get(url).json()
try:
    print(result['response']['venue']['rating'])
except:
    print('This venue has not been rated yet.')
```

8.3

However the bakery not a close direct competitor in terms of distance, but also because of their specialty. A bakery and not a restaurant. Lastly, a base demographic of Latinos were found in New York City. 4 out of 10 top immigrants, mainly from Guyana, Mexico, Dominion Republic, and Ecuador as displayed in Table 1. This evidence might give a supportive clientele but that would be purely speculation. However it may still comfort the client.

6.Discussion

- Brooklyn is an open area without direct competition

- Accuracies of models are subjected to change depending on opening/closures of other restaurants.
- Plus the considerations of the surrounding businesses/restaurants were only gathered through Foursquare. There might be unofficial or unverified businesses to consider.
- There is high competition in Manhattan, so it would be considered very risky to setup there.
- A more detailed analysis might be required to finalize the decision making situation for the clients. Including, but not limited to factors such as rent, human traffic or a comprehensive review of past restaurants in the area.