**Capstone Report - The Battle of the Neighborhoods (Week 2)**

**Applied Data Science Capstone by IBM/Coursera**

**Problem Background:**

In this module we will evaluate the competitive market and opportunity cost associated with doing business in New York City, NY. Since the establishment of the US, New York City has been the leader in driving revenue for the nation in every facet. From being the financial capital of the US coupled with its attractiveness to international business New York City has attracted many diverse cultures, heritages, and ethnicities. With that comes the ability to create a niche business or compete in the open market while minimizing downside risk. The results from this module will enhance insight while translating the data in efforts to discover a strategic approach in local direct marketing.

**Problem Description**

With the emergence of the food culture and growth from self-proclaimed foodies New York City has become one of the primary cuisine hubs of the world. There are no favorites which allows for ample opportunities for a restaurant investor who may be new to the area.

1. Chinese, Japanese, Korean cuisines
2. Italian Cuisine
3. Brazilian Cuisine
4. Cajun Cuisine
5. African, Jamaican, Caribbean
6. Greek Cuisine
7. Indian Cuisine
8. Jewish Cuisine
9. Mexican Cuisine
10. Thai Cuisine
11. Mediterranean Cuisine
12. American Cuisine
13. French Cuisine

With all the cuisine choices you can understand exactly how competitive and overwhelming a simple food choice can be. Add in food trucks, bodegas, and fusion restaurants just to name a few niches. Somethings to consider before opening a new restaurant location in New York City are:

1. Current population and tourism in New York City.
2. Diversity of demographics, heritages, ethnicities of New York City.
3. Opportunity cost to provide organic and vegan products from local farmers markets.
4. Distance from entertainment and multipurpose areas and food selection.
5. Statistics on Untapped and Saturated areas
6. Top competitors in those markets
7. Cuisine Options
8. Segmentation of the Boroughs

This research was designed based on the company’s goal to obtain multiple locations over the next five years. Along with franchising the importance of current market data will strategically expedite revenue and return on investment.

**Target Audience**

I was brought on to this project to lead the team of data scientist due to my vast knowledge of the area and industry. My team and I will discover opportunities and make recommendations for future locations across New York City. We understand how vital this initial location is to the growth of the overall portfolio and pride ourselves on following through for years to come.

**Success Criteria**

The success of this project will be based off of the immediate ability to generate profits based on the marketing efforts from research. If this initial location can maintain profitability for two years Lipsey Enterprise will invest in another location in emerging markets.

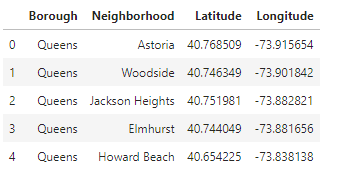
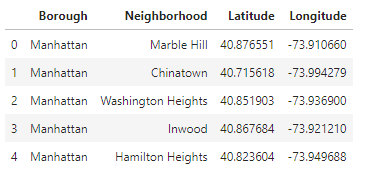
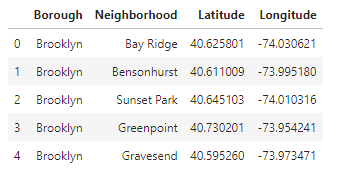
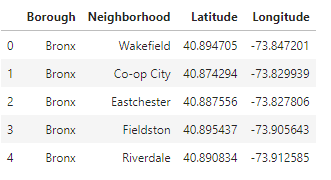
**Data Collected**

**Primary location: New York City**

The following data will be utilized to evaluate New York City as the primary market for potential restaurant locations.

**Data 1:**

The dataset that we will be utilizing to provide a proper assessment can be reviewed here:  
<https://geo.nyu.edu/catalog/nyu_2451_34572>



The data is segmented into a total of 5 boroughs consisting of 306 neighborhoods that we will be exploring based on longitude and latitude of each neighboring community.

**Data 2:**

The 2nd dataset will provide us with data of local eateries in New York City consisting of mobile food trucks, food carts, snack bars and restaurants.

<https://data.cityofnewyork.us/Recreation/Directory-of-Eateries/8792-ebcp>

This dataset also includes a list of urban agricultural sites that can be utilize for city gardens and urban farms.  
<https://data.cityofnewyork.us/Environment/City-owned-sites-that-are-available-and-potentiall/qchy-end3>

Having access to organic and vegan products in a major city like New York is vital to operating a restaurant.

**Data 3:**

Wikipedia is our main resource for our 3rd dataset.

1. Population of New York City.
2. Demographic of New York City.
3. List of Cuisines of New York City.  
   Links to Wikipedia sources:  
   <https://en.wikipedia.org/wiki/New_York_City>  
   <https://en.wikipedia.org/wiki/Economy_of_New_York_City>  
   <https://en.wikipedia.org/wiki/Portal:New_York_City>  
   <https://en.wikipedia.org/wiki/Cuisine_of_New_York_City>

**Data 4:**

Foursquare API will provide the content needed in order to discover venues for each New York City neighborhood in our initial dataset. (5 Borough and 306 Neighborhoods)

**Methodology**

In this project we will direct our efforts on detecting areas of New York City that have low restaurant density. We will limit our analysis to area ~6km around city center.

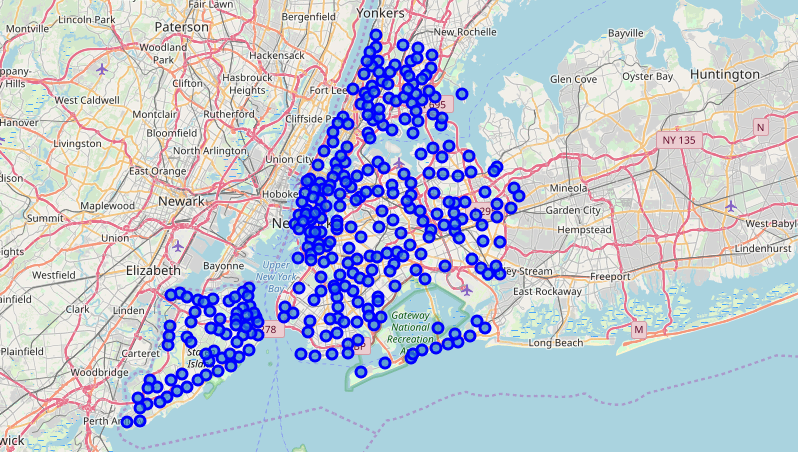
In first step we have collected the required data: location and type (category) of every restaurant within 6km from New York City center. We have also identified African restaurants (according to Foursquare categorization).

**Analytic Approach:**

According to the feedback from our sources New York City is comprised of 5 main boroughs which consist of 306 neighborhoods. The main areas we targeted based on demographic and popularity are Manhattan, Queens, Brooklyn and the Bronx.

Data 1: Geographic Coordinates:

* Our initial data was pulled from the file newyork\_data.json
* We then transition the data into pandas dataframe
* Next, we pinpoint our coordinates for 306 neighborhoods in New York City
* Our Foursquare API provided us with the appropriate data for venues
* We imported Geopy and Folium to create our visualizations for New York City
* The neighborhoods are indicated by blue marker.



Data 2: Venues and Eateries: The 2nd dataset will provide us with data of local eateries in New York City consisting of mobile food trucks, food carts, snack bars and restaurants. <https://data.cityofnewyork.us/Recreation/Directory-of-Eateries/8792-ebcp>

This dataset also includes a list of urban agricultural sites that can be utilize for city gardens and urban farms. <https://data.cityofnewyork.us/Environment/City-owned-sites-that-are-available-and-potentiall/qchy-end3> Having access to organic and vegan products in a major city like New York is vital to operating a restaurant.

Data 3: Wikipedia is our main resource for our 3rd dataset. We scraped these Wikipedia pages in order to acquire our data. Beautiful Soup was the tool that we imported into python. This allowed us to extract data and convert html into a useful format.

Population of New York City. Demographic of New York City. List of Cuisines of New York City. Links to Wikipedia sources: <https://en.wikipedia.org/wiki/New_York_City> <https://en.wikipedia.org/wiki/Economy_of_New_York_City> <https://en.wikipedia.org/wiki/Portal:New_York_City>

Results: Manhattan has a population of 72,033 persons per square mile. (Updated:2015) Making Manhattan the most populated county in the US. Central Harlem has the population and several existing African restaurants.

Demographics: New York City has an estimated population of 8,622,698 (Updated:2017)

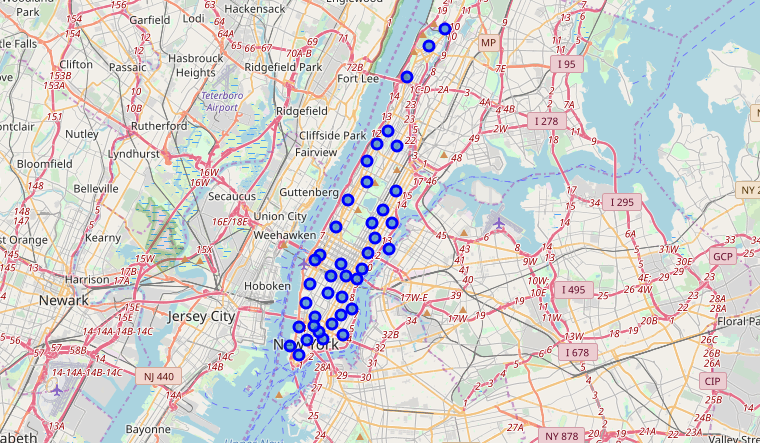
New York City Cuisines list was comprised using Beautiful Soup below: <https://en.wikipedia.org/wiki/Cuisine_of_New_York_City>

New York City Cuisine:

* Top Rated: Italian
* Puerto Rican
* Mexican
* Jewish
* African

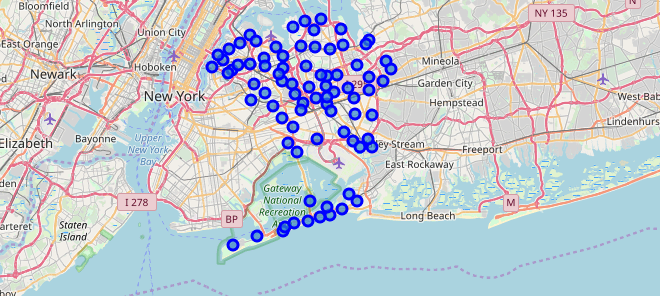
Manhattan Cuisine:

* Italian
* American
* Puerto Rican
* Indian



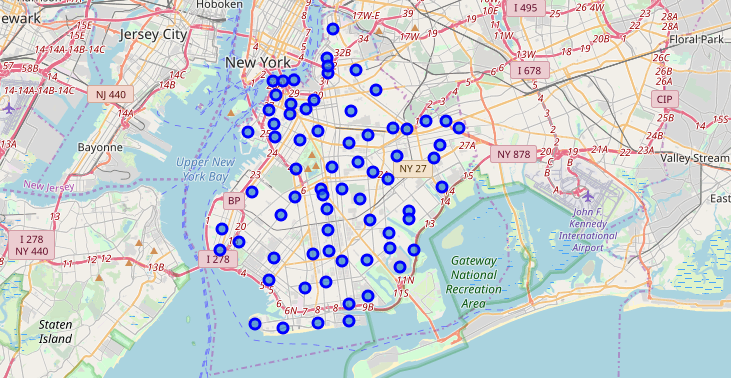
Queens Cuisine:

* Indian
* Irish
* Mexican
* Pakistan



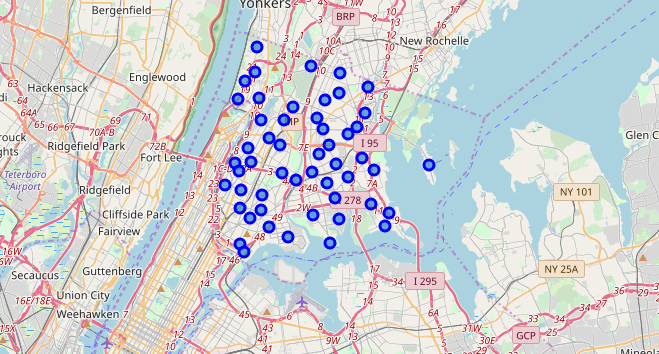
Brooklyn Cuisine:

* Italian
* Puerto Rican
* Mexican



Bronx Cuisine:

* Italian
* Puerto Rican
* Albanian
* Dominican



Data 4: Geographic Coordinates Foursquare API provided the data used to leverage the provisions of the venues listed within each neighborhood. Using the parameters to find the top 100 venues located within a 500KM radius.

**Conclusion**

After evaluating our results, we have concluded that Central Harlem would be the ideal spot to start the first Jamaican Fusion restaurant location. Based on the population and proximity we believe that there is enough traffic to generate sufficient business from marketing strategies. The data collected revealed that the second most sought out venue in Central Harlem was African Cuisine. There is no better market to begin investing in efforts to gain parts of the existing market share. Once the brand becomes established there will be opportunities to branch out due to the success of the first location.

