The Battle of the Neighborhoods

Opening a New Restaurant in NYC

1- Introduction & Business Problem:

- **Problem Background:**

The City of New York, is the most populated city in the United States. It is multicultural. It provides the best place to start a business. It has encouraged many different players into the market.

The city is a major center for many financial activities including restaurant businesses but the market is highly competitive. Thus, any new business project needs to be thoroughly analyzed. The understandings derived from the business analysis will have a great impact in influencing any business decision. This will definitely help in the reduction of the risk of any kind.

- **Problem Description:**

A restaurant is a business is planning to launch in the center in New York city. It is planned to prepare and serve food.. The City of New York is famous for its various cuisines. It includes international cuisines influenced by the city's multinational environment:

1. Central and Eastern European immigrants, especially Jewish immigrants - bagels, cheesecake, hot dogs, knishes, and delicatessens
2. Italian immigrants - New York-style pizza and Italian cuisine
3. Jewish immigrants and Irish immigrants - pastrami and corned beef
4. Chinese and other Asian restaurants, sandwich joints, diners, and coffeehouses are ubiquitous throughout the city
5. mobile food vendors - Some 4,000 licensed by the city
6. Middle Eastern foods such as falafel and kebabs examples of modern New York street food
7. It is famous for not just Pizzerias, Cafe's but also for fine dining Michelin starred restaurants. The city is home to "nearly one thousand of the finest and most diverse haute cuisine restaurants in the world", according to Michelin.

So it is obvious that to start a business in a such competitive market it is very important to do your homework before making this huge and risky step. Many factors need to be analyzed and studied in order to decide on the Location such as :

1. New York Population
2. New York City Demographics
3. Are there any Farmers Markets, Wholesale markets etc nearby so that the ingredients can be purchased fresh to maintain quality and cost?

4. Are there any venues like Gyms, Entertainment zones, Parks etc nearby where floating population is high etc
5. Who are the competitors in that location?
6. Cuisine served / Menu of the competitors
7. Segmentation of the Borough
8. Untapped markets
9. Saturated markets etc
   The list can go on...

- **Target Audience:**

To recommend the correct location, the Company has appointed me to lead of the Data Science team. The objective is to **locate and recommend** to the stakeholders which neighborhood of New York City will be best choice to start a restaurant. The Management also expects to understand the rationale of the recommendations made.

This would also be an interest of anyone who plans to start a new restaurant in New York city.

- **Success Criteria:**

The success criteria of the project will be a good recommendation of neighborhood choice to the stakeholders based on lack of such restaurants in that location and nearest suppliers of ingredients.

2- Data

The Target City : **New York, USA.**

The following data will be used to analyze the city:

- **Data 1 :** Neighborhoods in NYC

This dataset exists for free on the web. Link to the dataset is : https://geo.nyu.edu/catalog/nyu_2451_34572

- **Data 2 :** DOHMH Farmers Markets and Food Boxes dataset. In this we will be using the data of Farmers Markets.

https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2
.

- **Data 3 :** For the below analysis we will get data from wikipedia as given below :

1. New York Population
2. New York City Demographics
3. Cuisine of New York city

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
https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City

- **Data 4 :** Foursquare API to explore neighborhoods in New York City.

3- Analytic Approach:

New York City has 5 boroughs and 306 neighborhoods. The first part of the project will initiate a clustering of Manhattan and Brooklyn and the second part will handle clustering of Bronx, Queens and Staten Island.

Exploratory Data Analysis:

Data 1- New york city Geographical Coordinates Data

1. Loading and exploring data from newyork_data.json file.

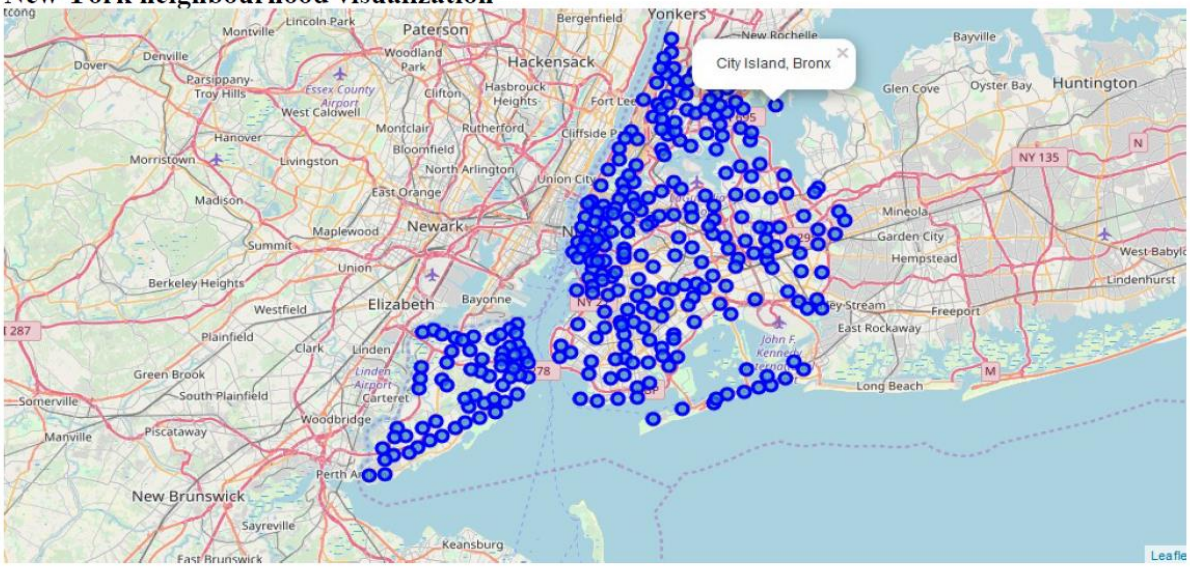
2. Transforming the data that is nested python dictionaries into a pandas data frame.

3. The data frame contains the coordinates of New York city neighborhoods.

4. The data will be used to get Venues data from Fouresquare API.

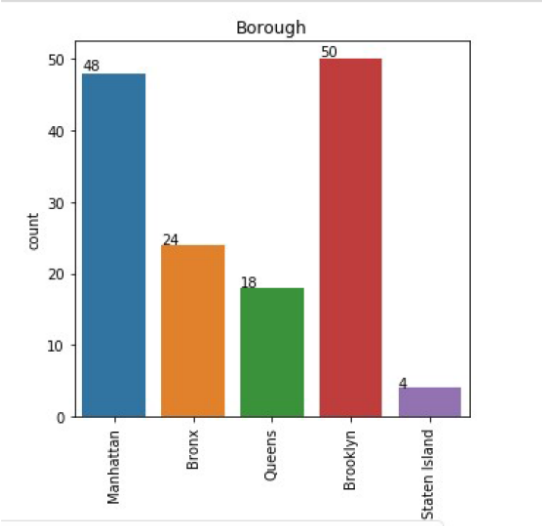
5. By using geopy and folium libraries to create a map of New York city with neighborhoods.

New York neighbourhood visualization

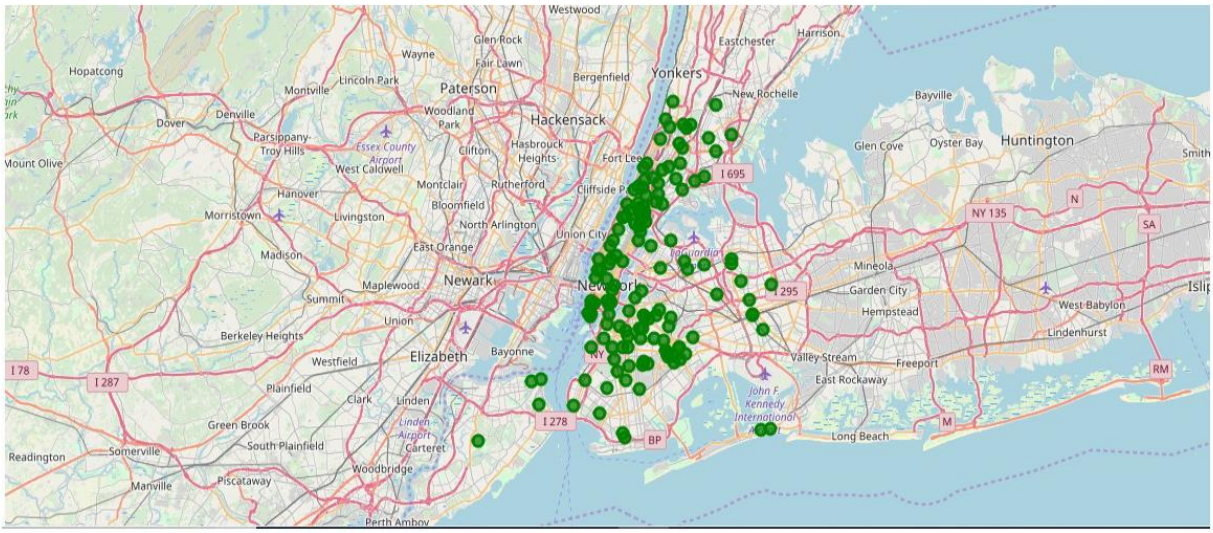


Data 2- DOHMH Farmers Markets and Food Boxes dataset

Using Farmers Markets data. There are totally 144 Farmers Markets in New York city. We notice that the highest numbers are in Manhattan and Brooklyn, whereas the lowest numbers are in Queens, Bronx and Staten Island.



Farmers Market visualisation-New York City



Data 3 : Analyzing New York city Population, Demographics and Cuisine

Scrapping data from Wikipedia and using BeautifulSoup library.

1.New York Population: Insights from the data

1- Manhattan is most densely populated borough.

2- Brooklyn is the city's most populous borough.

3- Queens is geographically the largest borough.

	Borough	County	Estimate_2017	square_miles	square_km	persons_sq_mi	persons_sq_km
0	Manhattan	New York	1,664,727	22.83	59.13	72,033	27,826
1	The Bronx	Bronx	1,471,160	42.10	109.04	34,653	13,231
2	Brooklyn	Kings	2,648,771	70.82	183.42	37,137	14,649
3	Queens	Queens	2,358,582	108.53	281.09	21,460	8,354
4	Staten Island	Richmond	479,458	58.37	151.18	8,112	3,132
5		City of New York	8,622,698	302.64	783.83	28,188	10,947
6		State of New York	19,849,399	47,214	122,284	416.4	159

2.New York City Demographics

New York City is the most populous city in the United States with an estimated record high of 8,622,698 residents as of 2017. The racial composition is as given below. This is the reason New York city has restaurants serving cuisine from many countries such as Indian, African, Japan etc. This also increases the scope for restaurants business in New York City.

	Racialcomposition	2010	1990	1970	1940
0	White	44.0%	52.3%	76.6%	93.6%
1	—Non-Hispanic	33.3%	43.2%	62.9%	92.0%
2	Black or African American	25.5%	28.7%	21.1%	6.1%
3	Hispanic or Latino (of any race)	28.6%	24.4%	16.2%	1.6%
4	Asian	12.7%	7.0%	1.2%	–

3.Cuisine of New York city

Data is derived from Wikipedia page https://en.wikipedia.org/wiki/Cuisine_of_New_York_City
Using a word cloud, the most Preferred Food in New York City –Italian, Purto Rican, Mexican, Jewish, Indian, Pakistani & Dominican.

BROOKLYN CUISINE -Most Preferred Food in Brooklyn is –Italian, Purto Rican & Mexican

MANHATTAN CUISINE - Most Preferred Food in Manhattan is – Italian, American, Puerto Rican and Indian.

QUEENS CUISINE - Most Preferred Food in Queens is – Indian, Irish, Pakistani and Mexican

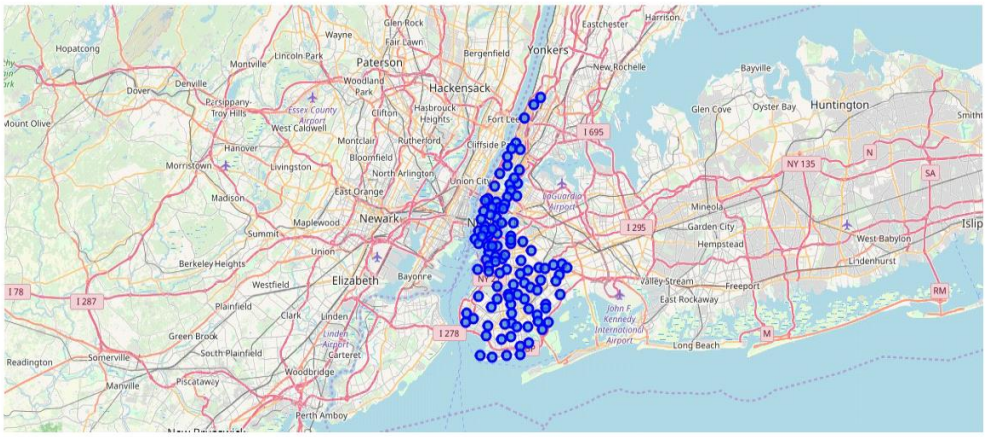
THE BRONX CUISINE - Most Preferred Food in The Bronx is – Italian,Puerto Rican, Albanian and Dominican.

Data 4 : Using Foursquare API

The data has been leveraged to provide venues information for each neighborhood using the Foursquare API data to explore neighborhoods.

1- Brooklyn and Manhattan

Brooklyn and Manhattan Visualization :

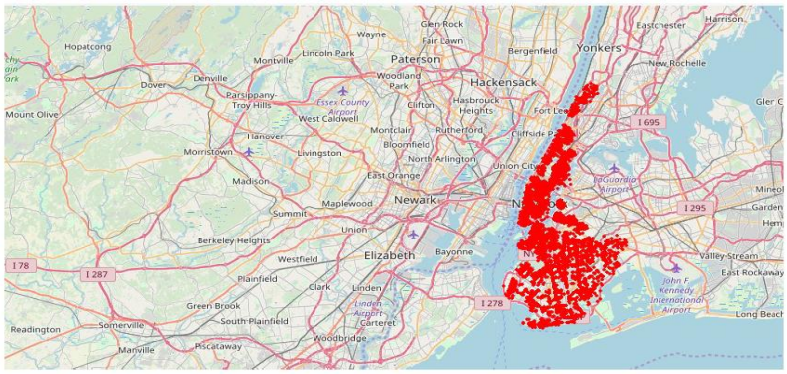


By using the geographical coordinates of each neighborhood, foursquare calls were made to get top 200 venues in a radius of 1000 meters. The venues are given below:

Brooklyn and Manhattan Venues :

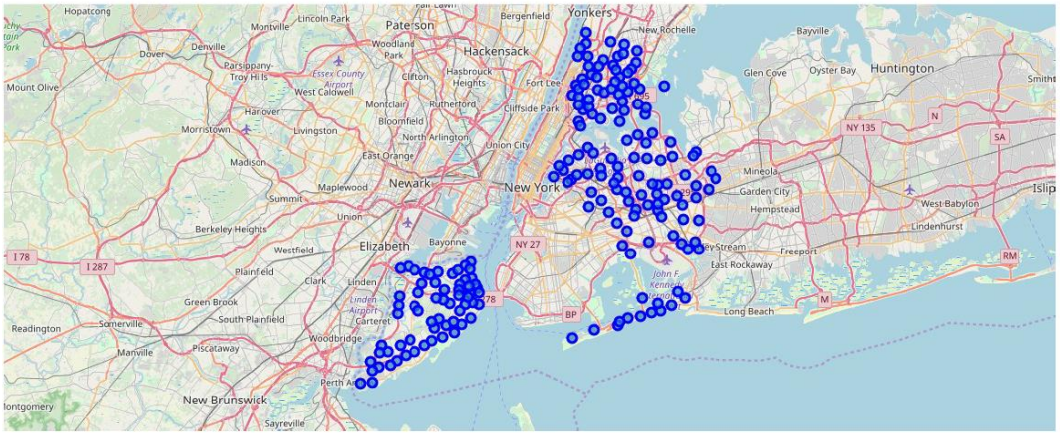
	Neighborhood	NeighborhoodLatitude	NeighborhoodLongitude	Venue	VenueLatitude	VenueLongitude	VenueCategory
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Sam's Pizza	40.879435	-73.905859	Pizza Place
4	Marble Hill	40.876551	-73.91066	Loeser's Delicatessen	40.879242	-73.905471	Sandwich Place

Brooklyn and Manhattan Venues Visualization :



2- Bronx, Queens and Staten Island :

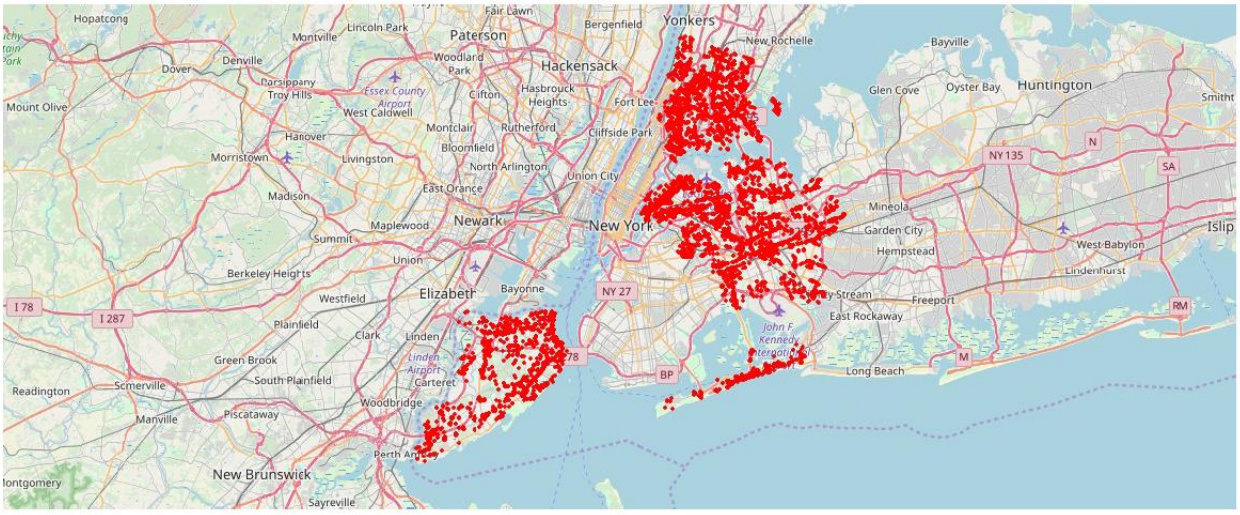
Bronx, Queens and Staten Island Neighborhoods Visualization :



Bronx, Queens and Staten Island Venues Visualization :

	Neighborhood	NeighborhoodLatitude	NeighborhoodLongitude	Venue	VenueLatitude	VenueLongitude	VenueCategory
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Ripe Kitchen & Bar	40.898152	-73.838875	Caribbean Restaurant
2	Wakefield	40.894705	-73.847201	Jackie's West Indian Bakery	40.889283	-73.843310	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
4	Wakefield	40.894705	-73.847201	Rite Aid	40.896521	-73.844680	Pharmacy

Bronx, Queens and Staten Island Venues Map Visualization :



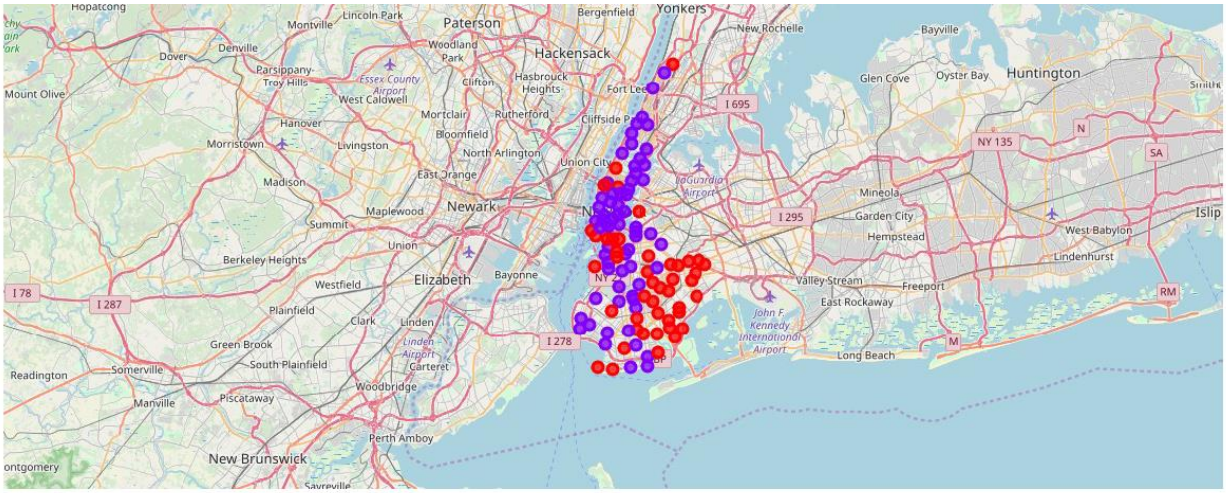
4- RESULTS :

The restaurant data were used for Brooklyn & Manhattan clustering and Bronx, Queens and Staten Island clustering.

Neighborhood K-Means clustering based on mean occurrence of venue category:

K-Means clustering Algorithm was used to cluster the neighborhoods into two. k-means

Brooklyn & Manhattan:

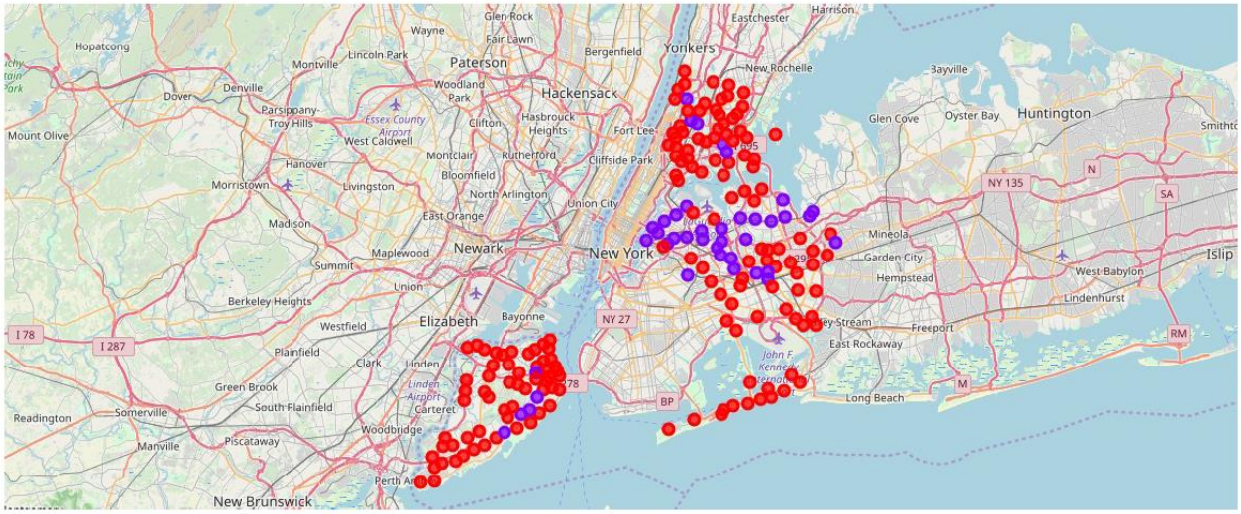


Cluster0 : The Total and Total Sum of cluster0 has smallest value. It shows that the market is not saturated.

Cluster1 : The Total and Total Sum of cluster1 has highest value. It shows that the markets are saturated. Number of restaurants are very high.

- As it is obvious from the map above there are no available neighborhoods in Brooklyn and Manhattan.

Bronx, Queens and Staten Island:



Cluster0 : The Total and Total Sum of cluster0 has smallest value. It shows that the market is not saturated. There are untapped neighborhoods. List is as given below.

	Borough	Neighborhood	Latitude	Longitude	Total	Cluster_Labels
0	Staten Island	Todt Hill	40.597069	-74.111329	0	0
1	Staten Island	Port Ivory	40.639683	-74.174645	0	0
2	Staten Island	Bloomfield	40.605779	-74.187256	0	0

Cluster1 : The Total and Total Sum of cluster1 has highest value. It shows that the markets are saturated. Number of restaurants are very high.

5- DISCUSSION:

1. There is an opportunity in growth for the Farmers markets in Bronx, Queens and Staten Island.

2. There is a possibility to explore multiple cuisine choices in Bronx, Queens and Staten Island.

3. Cuisines of many countries are already available In Manhattan and Brooklyn. It shows people love eating cuisines of various origins. So a risk can be taken along with great menu.

6- Conclusion and Recommendations:

After studying the various results a number of recommendations could be taken into consideration:

1- Bronx, Queens and Staten Island also has good number of restaurants but it can handle more. So an option of opening a new restaurant in on of these areas can be explored.
2- Brooklyn and Manhattan has high concentration of restaurant business. It is a high competitive market. But, you can compete as the demands for food and restaurants are increasing and a good menu can surely make the difference.
3- A restaurant with Indian flavor can be thoroughly checked as a good choice as it comes with lower risks and previously known competitors.
4- A deep data analysis with updated data can be done to determine the certainty of the findings.

.