The battle of the neighborhoods

By Niels Rijneveld

Coursera Capstone Project

Introduction

- Focus on New York City for XYZ Company
- Optimum location for a new restaurant business

Business problem

- Choice of first neighborhood to start a restaurant business
- Easy to replicate
- Low competition
- High demand
- Choice of menu

Success criteria

- Find the best neighborhood that meets the criteria

Facts of New York City

- Most populous city in the United States
- It is diverse and multicultural
- Provides lot of business opportunities
- It has a business-friendly environment and attracted many different players
- Global hub of business and commerce.
- The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theatre, fashion, and the arts in the USA.

Business environment

- Market is highly competitive
- Highly developed city so cost of doing business is also one of the highest
- New business venture or expansion needs to be analysed carefully
- One should strategically target the market
  - This will help in reduction of risk.
  - The Return on Investment will be reasonable.

Cuisine of New York City

- **Central and Eastern European immigrants**, especially Jewish immigrants: bagels, cheesecake, hot dogs, knishes, and delicatessens
- **Italian immigrants -** New York-style pizza and Italian cuisine
- **Jewish immigrants and Irish immigrants -** pastrami and corned beef
- **Chinese and other Asian restaurants -** sandwich joints, trattorias, diners, and coffee houses are ubiquitous throughout the city
- **Mobile food vendors -** Some 4,000 licensed by the city
- **Middle Eastern food** such as falafel and kebabs examples of modern New York street food
- Famous for fine dining **Michelin star 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 evident that to survive in such competitive market it is very important to strategically plan.

Factors to consider for location

- New York Population & Demographics
- Are there any Farmers Markets, Wholesale markets near by?
- Are there any venues like gyms, entertainment zones, parks near by?
- Who are the competitors in that location?
- Cuisine served by competitors
- Segmentation of the Borough
- Untapped markets
- Saturated markets

Data description

Data 1:

- Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.
- This dataset exists for free on the web. Link to the dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572

	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

Data description

Data 2: Farmers Markets data

- https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2
- A farmers' market is often defined as a public site used by two or more local or regional producers for the direct sale of farm products to consumers. In addition to fresh fruits and vegetables, markets may sell dairy products, fish, meat, baked goods, and other minimally processed foods.

	Borough	Market Name	Service Type	Street Address	Latitude	Longitude	Days of Operation	Hours of Operations	Season Dates	Accepts EBT	Open Year-Round	Stellar Cooking Demonstrations	Food Activities for Kids	Location Point
0	Brooklyn	Woodhull Hospital Youthmarket	Farmers Markets	Broadway & Flushing Ave	40.700726	-73.941932	Wednesday	9 a.m. - 2 p.m.	07/10/2019-11/27/2019	Yes	No	No	No	(40.700726, -73.941932)
1	Manhattan	Mount Sinai Hospital Greenmarket	Farmers Markets	E 99th St bet Madison & Park Aves	40.789169	-73.952743	Wednesday	8 a.m. - 5 p.m.	06/12/19-11/27/19	Yes	No	No	No	(40.789169, -73.952743)
2	Bronx	170 Farm Stand	Farmers Markets	E 170th St & Townsend Ave	40.839882	-73.916783	Wednesday	2:30 - 6:30 p.m.	07/10/2019-11/27/2019	Yes	No	No	Yes	(40.839882, -73.916783)
3	Manhattan	Greenmarket at Oculus Plaza	Farmers Markets	Church & Fulton Sts, on Oculus Plaza	40.711535	-74.010464	Tuesday	7 a.m. - 7 p.m.	07/09/2019-11/30/19	Yes	Yes	No	No	(40.711535, -74.010464)
4	Queens	Ditmars Park Youthmarket	Farmers Markets	Steinway St bet Ditmars Blvd & 23rd Ave, at Di...	40.772854	-73.906061	Saturday	8 a.m. - 3 p.m.	07/13/2019-11/23/2019	Yes	No	No	No	(40.772854, -73.906061)

Data description

Data 3: Data from Wikipedia pages

- New York Population
- New York City Demographics
- 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

	Racialcomposition	2010	1990	1970	1940
0	None	None	None	None	None
1	White	44.0%	52.3%	76.6%	93.6%
2	—Non-Hispanic	33.3%	43.2%	62.9%	92.0%
3	Black or African American	25.5%	28.7%	21.1%	6.1%
4	Hispanic or Latino (of any race)	28.6%	24.4%	16.2%	1.6%
5	Asian	12.7%	7.0%	1.2%	–

Data description

Data 4: foursquare

- New York city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City. The below is image of the Foursquare API data.

	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	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Jimbo's	40.891740	-73.858226	Burger Joint
4	Wakefield	40.894705	-73.847201	Kingston Tropical Bakery	40.888568	-73.859885	Bakery

Analytical approach

- New York City neighborhood has a total of 5 boroughs and 306 neighborhoods

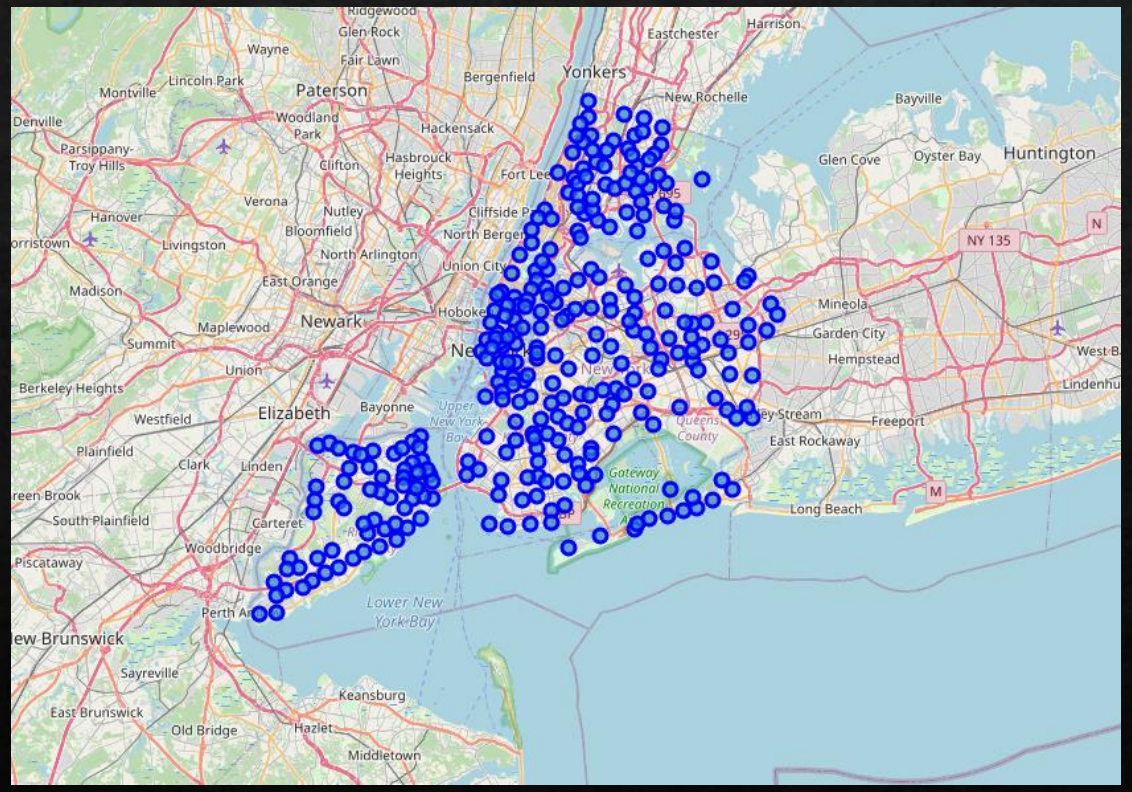
- PART 1 - Clustering of Manhattan and Brooklyn
- PART 2 - Clustering of Bronx, Queens and Staten Island.

- Only restaurant data is filtered from foursquare.com venues data and utilized for this project.
- This is done using the following exploratory data analysis.

Methodology

Data 1- New York city Geographical Coordinates Data.

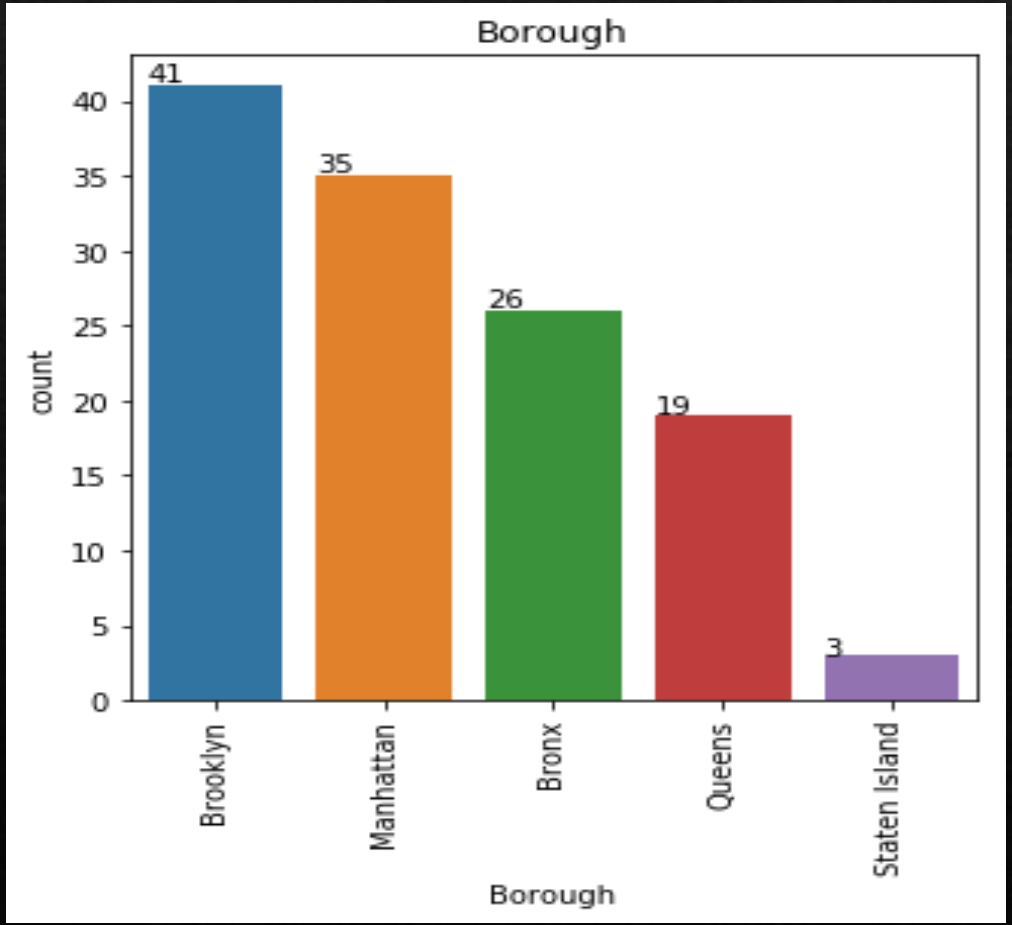
- Load the data and explore data - newyork_data.json
- Transform the data of nested python dictionaries into a pandas data frame.
- The data frame contains the geographical coordinates of New York city neighborhoods.
- Data will used to get Venues data from Foursquare.
- Geopy and folium libraries used to create a map of New York City with neighborhoods superimposed on top.



Methodology

Data 2 - DOHMH Farmers Markets and Food Boxes dataset.

- Here we will be using the data of Farmers Markets
- There are totally 124 Farmers Markets in New York city.
- Highest number are in Manhattan and Brooklyn; lowest in Queens, Bronx and Staten Island.



Methodology

Farmers Market visualization - New York City

Methodology

Data 3: To analyze New York City Population, Demographics and

- We used BeautifulSoup python library.
  - Beautiful Soup is a Python package for parsing HTML and XML documents (including having malformed markup, i.e. non-closed tags, so named after tag)
  - It creates a parse tree for parsed pages that can be used to
  - extract data from HTML, which is useful for web scraping

Methodology – New York Population

- Manhattan (New York County) is the geographically smallest and most densely populated borough.
- Brooklyn (Kings County), on the western tip of Long Island, is the city's most populous borough.
- Queens (Queens County), on Long Island north and east of Brooklyn, is geographically the largest borough.

	Borough	County	Estimate_2017	GrossDomesticProduct	square_miles	square_km	persons_sq_mi
0	The Bronx	Bronx	1,418,207	42.695\n	30,100	42.10	109.04
1	Brooklyn	Kings	2,559,903	91.559\n	35,800	70.82	183.42
2	Manhattan	New York	1,628,706	600.244\n	368,500	22.83	59.13
3	Queens	Queens	2,253,858	93.310\n	41,400	108.53	281.09
4	Staten Island	Richmond	476,143	14.514\n	30,500	58.37	151.18
5		City of New York	8,336,817	101,000	842.343	302.64	783.83
6		State of New York	19,453,561	89,000	1,731.910	47,214	122,284

Methodology – New York City Demographics

New York City is the most populous city in the United States, with an estimated record high of 8,336,817 residents as of 2017, incorporating more immigration into the city than outmigration since the 2010 United States Census.

	Racialcomposition	2010	1990	1970	1940
0	None	None	None	None	None
1	White	44.0%	52.3%	76.6%	93.6%
2	—Non-Hispanic	33.3%	43.2%	62.9%	92.0%
3	Black or African American	25.5%	28.7%	21.1%	6.1%
4	Hispanic or Latino (of any race)	28.6%	24.4%	16.2%	1.6%
5	Asian	12.7%	7.0%	1.2%	–

Methodology – Cuisines of New York City

- This data has been manually prepared. Data is taken from Wikipedia page - https://en.wikipedia.org/wiki/Cuisine_of_New_York_City
- Data used to create word cloud.

NEW YORK CITY CUISINE - Most Preferred Food in New York City – Italian, Indian, Mexican.

Methodology – Cuisines of New York City

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

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

Methodology – Cuisines of New York City

- **QUEENS CUISINE** - Most Preferred Food in Queens is – Italian, Indian, Pakistani

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

Methodology - Foursquare

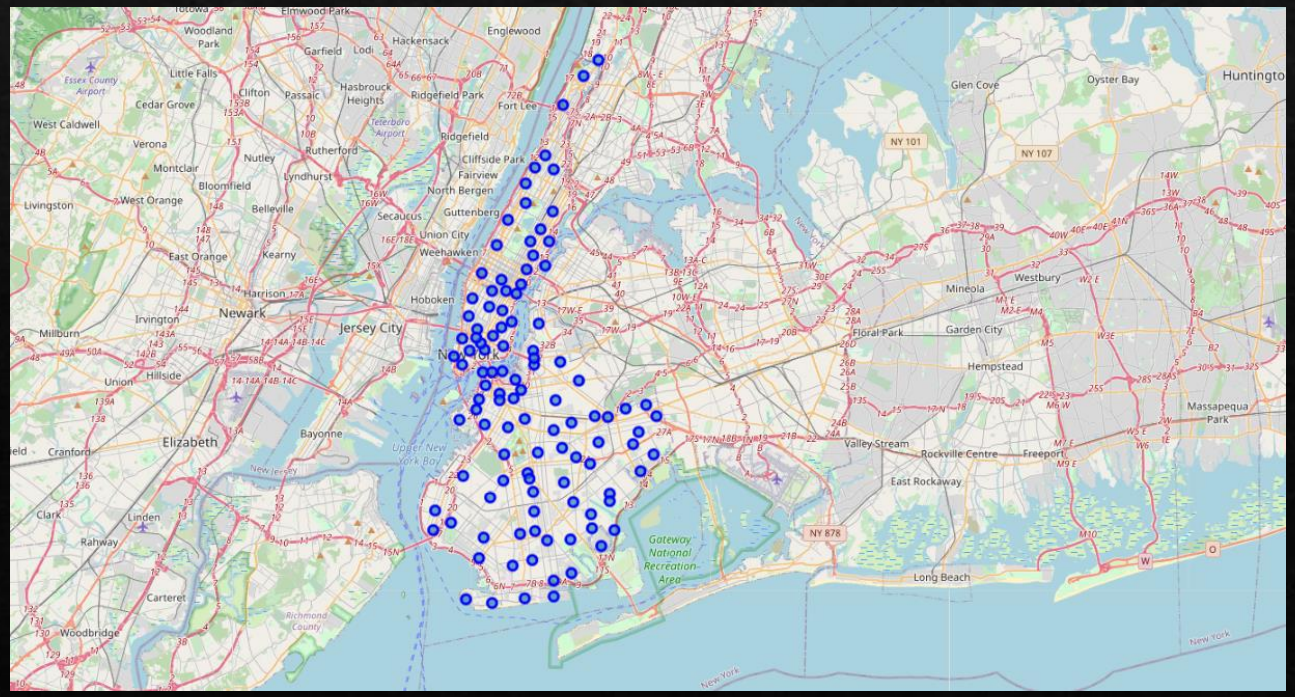
- NewYork city geographical coordinates data has be utilized as input for the Foursquare API, that has been leveraged to provision venues information for each neighborhood.
- We used the Foursquare API data to explore neighborhoods in New York City.
- Using the geographical coordinates of each neighbourhood foursquare API calls are made to get top 200 venues in a radius of 1000 meters

PART – 1 Brooklyn and Manhattan

PART – 2 Bronx, Queens and Staten Island

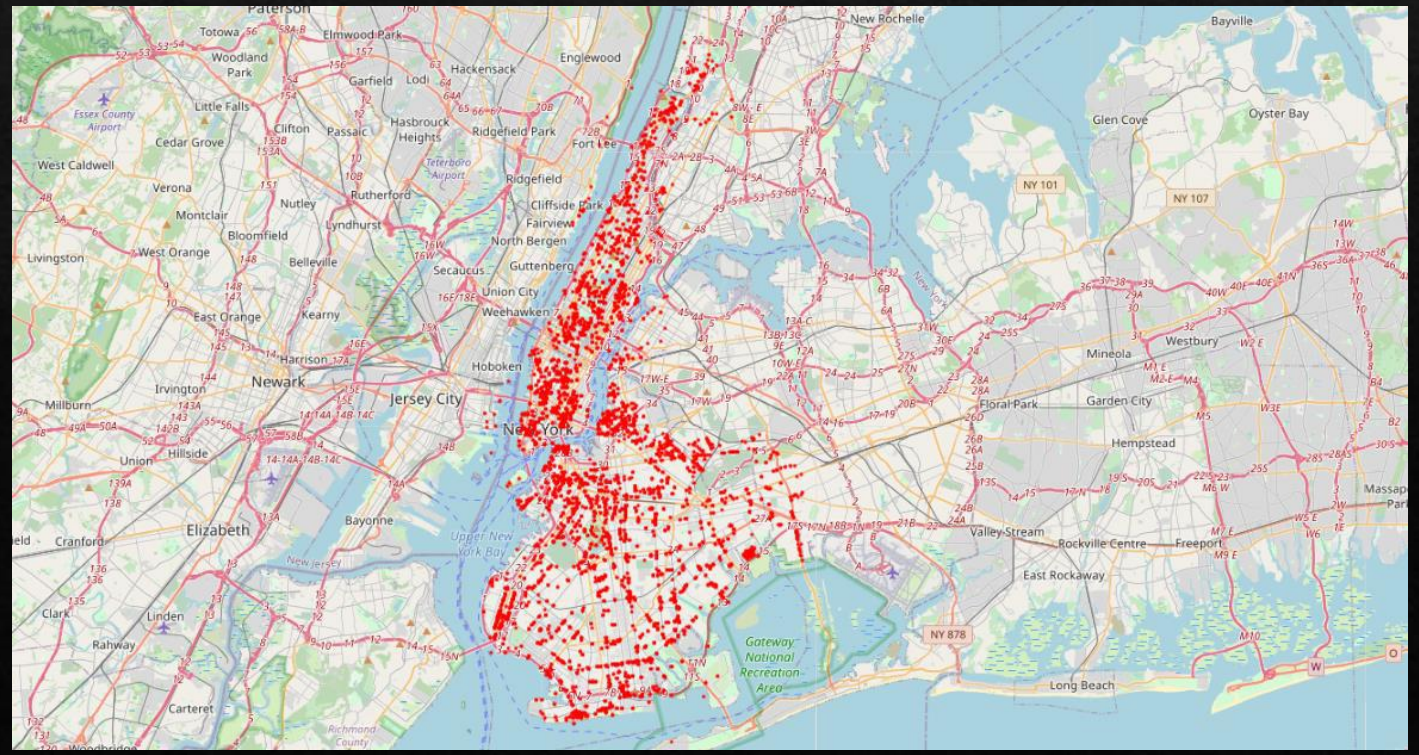
Methodology Part 1 – Brooklyn & Manhattan

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
1	Marble Hill	40.876551	-73.91066	Sam's Pizza	40.879435	-73.905859	Pizza Place
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Garden Gourmet Market	40.881350	-73.903389	Gourmet Shop
4	Marble Hill	40.876551	-73.91066	Indian Road Café	40.872922	-73.918459	Café



Methodology Part 1 – Brooklyn & Manhattan

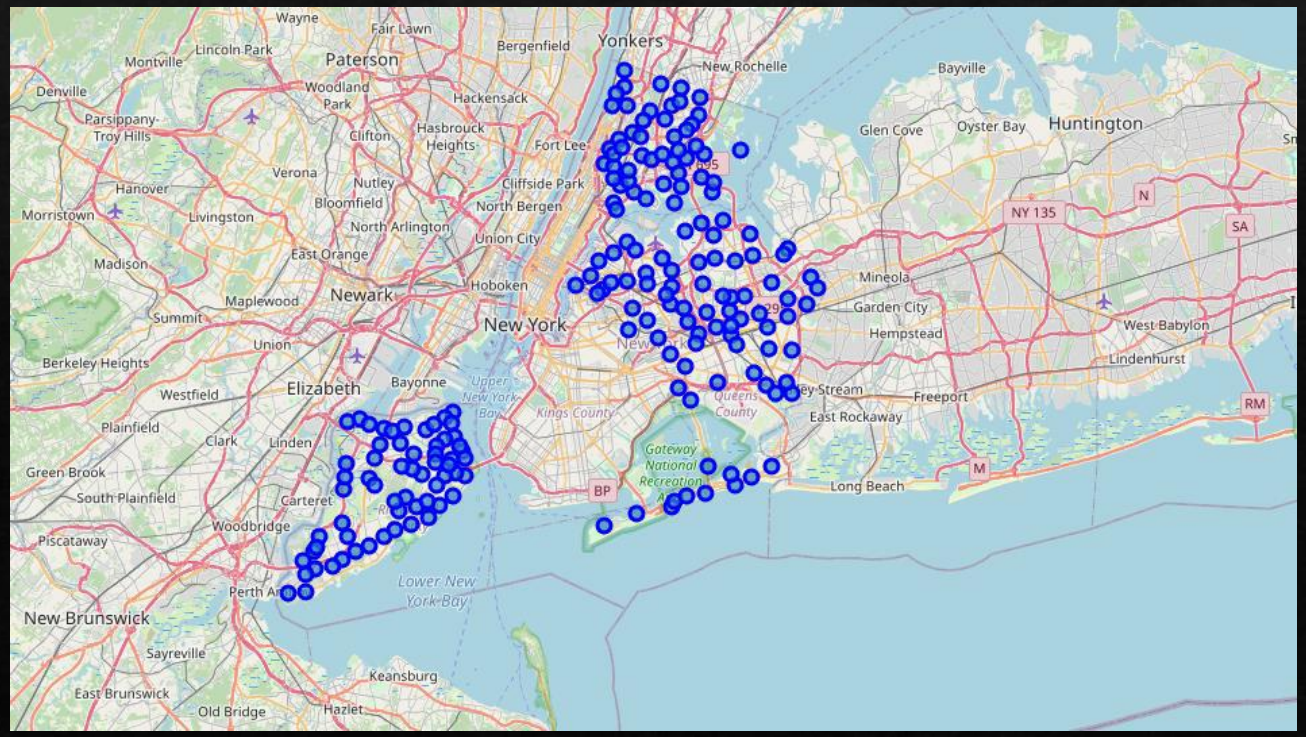
Brooklyn and Manhattan Venues Visualization: Generated the below Brooklyn and Manhattan Venues Visualization. The "BM_venues" dataframe has 10997 venues and 300 unique venue types.



Methodology Part 2 – Bronx, Queens & Staten Island

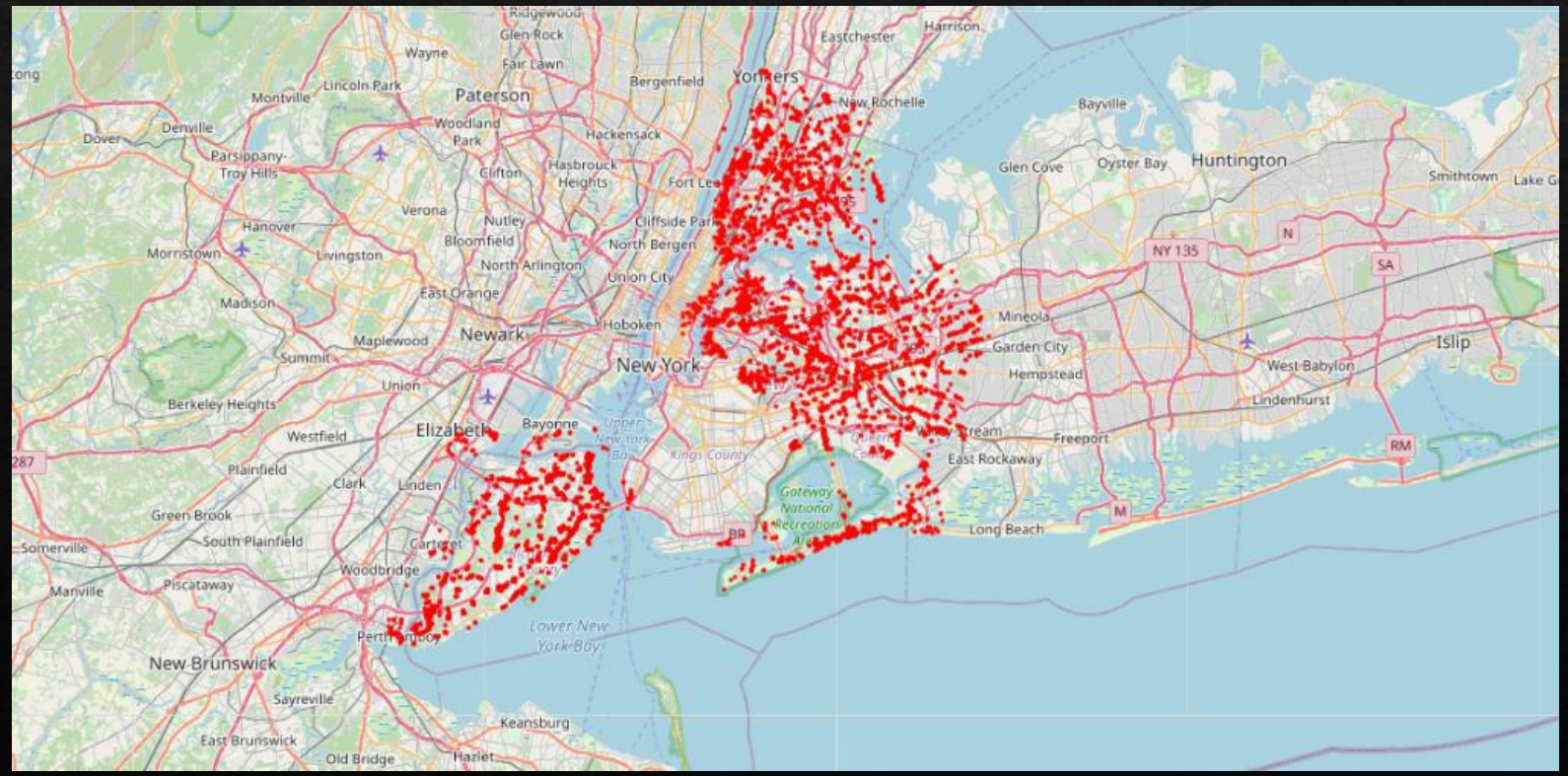
[42]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
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	Ali's Roti Shop	40.894036	-73.856935	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Jimbo's	40.891740	-73.858226	Burger Joint
4	Wakefield	40.894705	-73.847201	Kingston Tropical Bakery	40.888568	-73.859885	Bakery



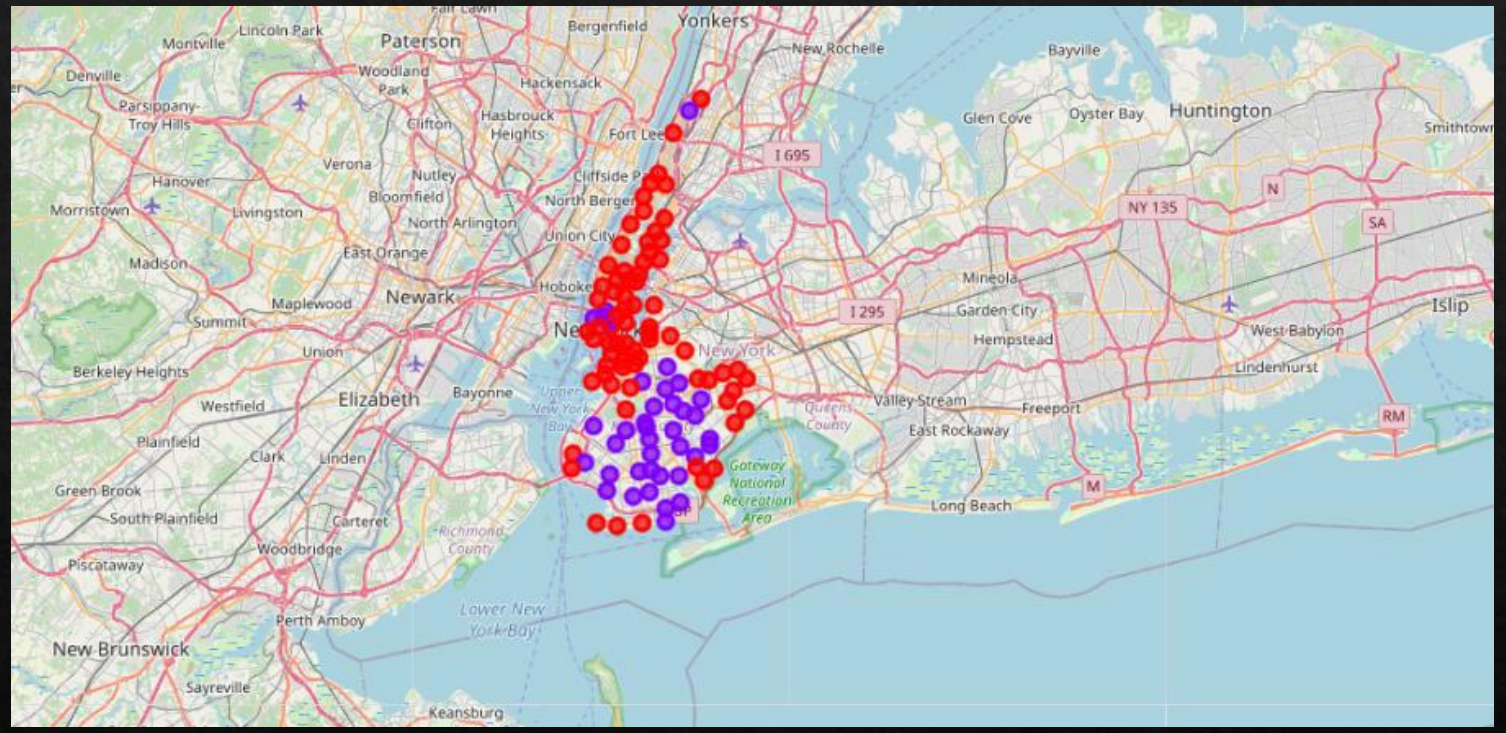
Methodology Part 2 – Bronx, Queens & Staten Island

Bronx, Queens and Staten Island Venues Visualization : The "BQS_venues" dataframe has 18173 venues and 345 unique venue types.



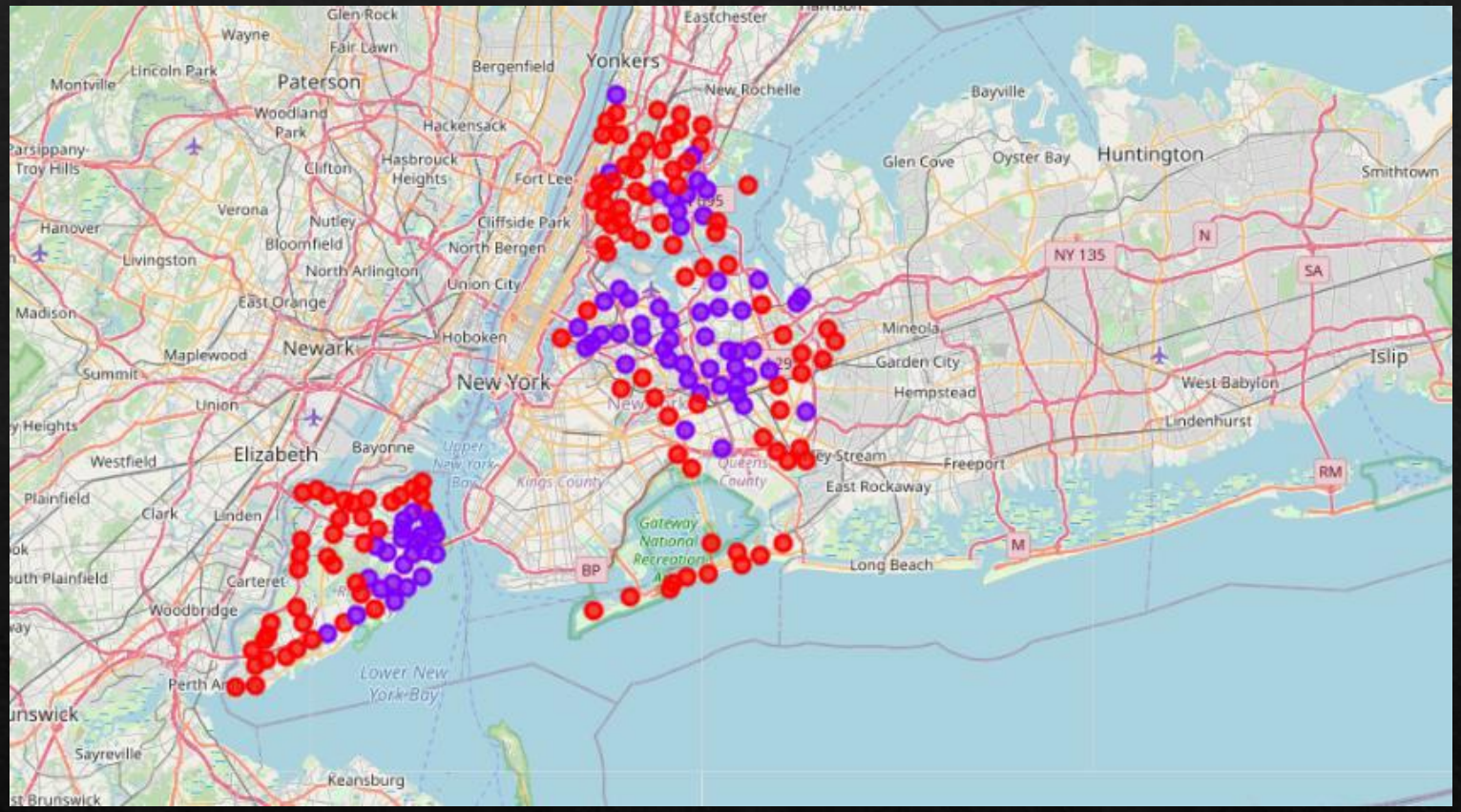
Results 1 – Brooklyn & Manhattan

- Segmenting and Clustering Neighborhoods
- 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 is very high. Lot of competition. Saturated neighborhoods.



Results 2 – Bronx, Queens & Staten Island

- Segmenting and Clustering Neighborhoods
- Cluster0: The Total and Total Sum of cluster0 has smallest value. It shows that the market is not saturated. There are untapped markets.
- Cluster1: The Total and Total Sum is very high. Lot of competition. Saturated neighborhoods.



	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

Conclusion & discussion

- There is scope to increase Farmers markets in Bronx, Queens and Staten Island.
- There is scope to explore cuisines of various countries in Bronx, Queens and Staten Island.
- In Manhattan and Brooklyn restaurants of cuisines of many countries are available. So if risk can be taken with great menu on board. It also shows people love eating cuisines of various countries.

Conclusion & discussion

- Analysis performed on limited data
- Re-run program with updated information
- Brooklyn and Manhattan has high concentration of restaurant business. Very competitive market.
- Bronx, Queens and Staten Island also have good number of restaurants but not as many as required. So this can be explored.
- As per the neighbourhood or restaurant type mentioned like Indian Restaurant, analysis can be checked. A venue with lowest risk and competition can be identified