

The Battle of the Neighborhoods – Report

1. Introduction and Business Problem

- **Problem Background**

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a 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, theater, fashion, and the arts in the United States.

This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analysed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

- **Problem Description**

A restaurant is a business which prepares and serves food and drink to customers in return for money, either paid before the meal, after the meal, or with an open account. The City of New York is famous for its excellent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history.

- 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 coffeehouses are ubiquitous throughout the city
- mobile food vendors - Some 4,000 licensed by the city
- Middle Eastern foods such as falafel and kebabs examples of modern New York street food

- 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 evident that to survive in such competitive market it is very important to strategically plan. Various factors need to be studied in order to decide on the Location such as :

- New York Population
- New York City Demographics
- Are there any Farmers Markets, Wholesale markets nearby so that the ingredients can be purchased fresh to maintain quality and cost?
- Are there any venues like Gyms, Entertainment zones, Parks nearby where floating population is high
- Who are the competitors in that location?

- **Target Audience:**

To recommend the correct location, XYZ Company Ltd has appointed me to lead of the Data Science team. The objective is to locate and recommend to the management which neighborhood of Newyork city will be best choice to start a restaurant. The Management also expects to understand the rationale of the recommendations made.

This would interest anyone who wants to start a new restaurant in New York city.

- **Success Criteria:**

The success criteria of the project will be a good recommendation of borough/Neighborhood choice to XYZ Company Ltd based on Lack of such restaurants in that location and nearest suppliers of ingredients.

2. Data

One city will be analyzed in this project: New York City.

We will be using the below datasets for analyzing New York city

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 2:

Second data which will be used is the 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>

Website-<https://www.grownyc.org/greenmarketco/foodbox> GrowNYC's Fresh Food Box Program is a food access initiative that enables under-served communities to purchase fresh, healthy, and primarily regionally grown produce well below traditional retail prices.

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.

	FacilityName	Service Category	Service_Type	Address	Address 2	Borough	ZipCode	Latitude	Longitude	AdditionalInfo	StartDate	EndDate	Monday	Tuesday	Wednesday	Thursday
0	Inwood Park Greenmarket	Farmers Markets and Food Boxes	Farmers Markets	Isham St bet Seaman & Cooper	NaN	Manhattan	10034	40.869009	-73.920320	Open year-round	NaN	NaN	NaN	NaN	NaN	NaN
1	82nd Street Greenmarket	Farmers Markets and Food Boxes	Farmers Markets	82nd St bet 1st & York Aves	NaN	Manhattan	10028	40.773448	-73.948954	Open year-round	NaN	NaN	NaN	NaN	NaN	NaN
3	125th Street Farmers Market	Farmers Markets and Food Boxes	Farmers Markets	125th St & Adam Clayton Powell Jr Blvd	NaN	Manhattan	10027	40.808981	-73.948327	Market open dates: 6/13/2017 to 11/21/2017	06/13/2017	11/21/2017	NaN	10am-7pm	NaN	NaN
4	170 Farm Stand	Farmers Markets and Food Boxes	Farmers Markets	170th St & Townsend Ave	NaN	Bronx	10452	40.840095	-73.916827	Market open dates: 7/5/2017 to 11/22/2017	07/05/2017	11/22/2017	NaN	NaN	2:30pm-6:30pm	NaN
5	175th Street Greenmarket	Farmers Markets and Food Boxes	Farmers Markets	175th St bet Wadsworth Ave & Broadway	NaN	Manhattan	10033	40.845956	-73.937813	Market open dates: 6/29/2017 to 11/30/2017	06/29/2017	11/30/2017	NaN	NaN	NaN	8am-5pm

Data 3:

For the below analysis we will get data from Wikipedia as given below :

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

Data 4:

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	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	Rite Aid	40.896521	-73.844680	Pharmacy
2	Wakefield	40.894705	-73.847201	Cooler Runnings Jamaican Restaurant Inc	40.898283	-73.850478	Caribbean Restaurant
3	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
4	Wakefield	40.894705	-73.847201	Dunkin Donuts	40.890631	-73.849027	Donut Shop

3. Methodology

- **Business Understanding**

Our main goal is to get optimum location for new restaurant business in New York City for the company

- **Analytic Approach**

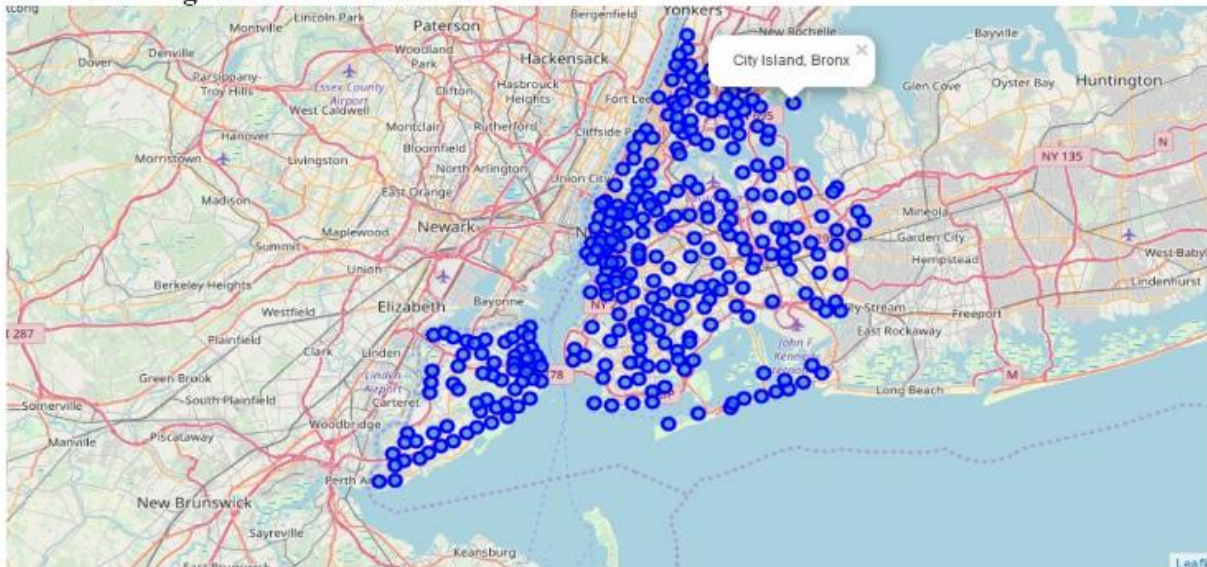
New York city neighborhood has a total of 5 boroughs and 306 neighborhoods. The first part of this project is the clustering of Manhattan and Brooklyn. The second part is the clustering of Bronx, Queens and Staten Island.

- **Exploratory Data Analysis**

Data 1: New York City Geographical Coordinates Data

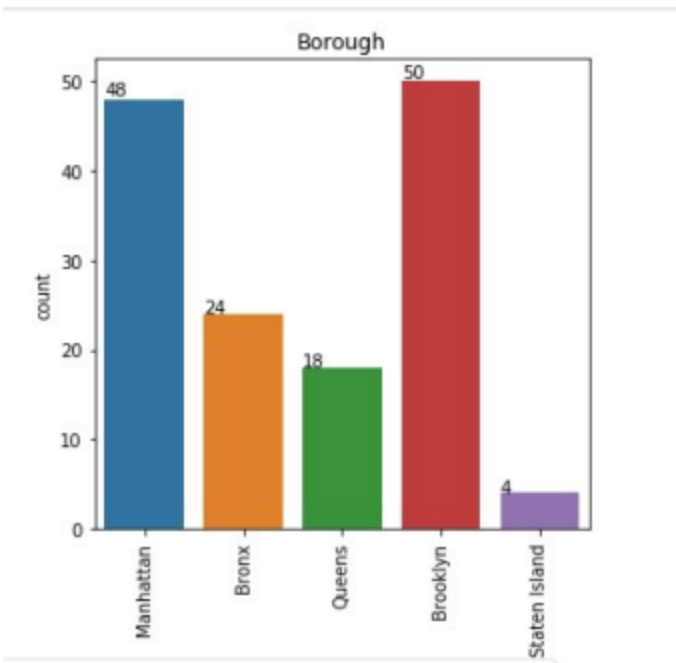
- Load the data and explore data from newyork_data.json file.
- Transform the data of nested python dictionaries into pandas dataframe.
- The dataframe contains the geographical coordinates of New York City neighbors
- Get Venues data from Foursquare.
- Used geopy and folium libraries to create a map of New York city with neighborhoods

New York neighbourhood visualization

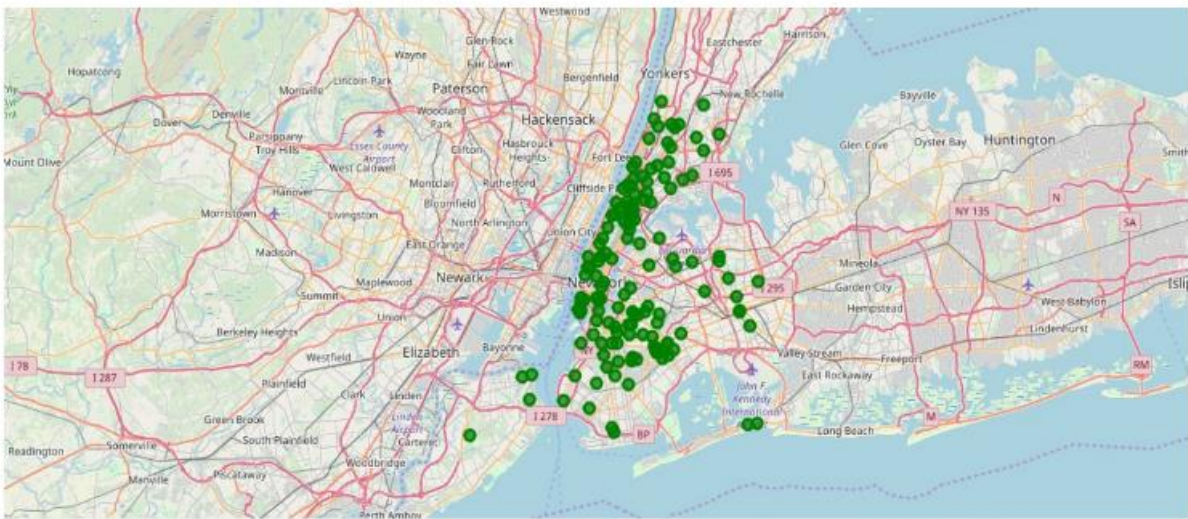


Data 2: the second data which is used is the DOHMH Farmers Markets and Food Boxes dataset.

There are totally 144 Farmers Markets in the city. The highest number are in Manhattan and Brooklyn. The picture below shows the evidence



Farmers Market visualisation-New York City



4. Results:

Brooklyn and Manhattan:

In the Map Visualization, we see the different types of clusters created by using K-Means

