The Battle of the Neighborhoods – Report

1. Introduction & Business Problem.

New York City comprises 5 boroughs sitting where the Hudson River meets the Atlantic Ocean. At its core is Manhattan, a densely populated borough that's among the world's major commercial, financial and cultural centers.

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 analyzed 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.

A well established restaurant chain for Bulgarian traditional food is looking for an opportunity to expand its business in North America. As an employee it is my job to analyze big cities like New York, San Francisco, Toronto etc. and propose a suitable location for the restaurant. In order to accomplish the task I need to gather the following information:

1. Gather information about New York City boroughs and neighborhoods.
2. List the nearby venues like restaurants, coffee shops, shopping centers business centers.
3. Determine most preferred food.
4. Pinpoint the most suitable place to open a restaurant in Manhattan and Brooklyn.

If a business wants to stay competitive many other thigs must be takin into consideration like population, demographics, markets and access to ingredients, competition etc.

The developed can be used to recommend a suitable location for opening a restaurant in any city around the world. If a business wants to stay competitive many other thigs must be takin into consideration like population, demographics, markets and access to ingredients, competition etc. Once the analysis is done it can be presented to the management for the purpose of making the decision whether it is suitable to expand the business.

2. Data

New York City comprises 5 boroughs sitting where the Hudson River meets the Atlantic Ocean. At its core is Manhattan, a densely populated borough that's among the world's major commercial, financial and cultural centers. The data for the 5 boroughs and 306 neighborhoods is available at https://geo.nyu.edu/catalog/nyu_2451_34572.

A data for the local market is also needed in order to analyze the availability of needed food and ingredients. In the below link Farmer market data is given:

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.

For further analysis we need to gather information about Now York population, demographics and cuisine. That data can be scraped from various websitelike wikipedia:

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

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

2.1. The data about New York neighborhoods is put into a dataframe as follows:

	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

The data is available at https://cocl.us/new_york_dataset.

2.2. Farmer markets data is available at https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2
The information is used to create a dataframe and a map of the available markets

2.3. Demographics data: https://en.wikipedia.org/wiki/Demographics_of_New_York_City. Using that link to scrape a table and create a dataframe.

2.4. Cuisine data: https://en.wikipedia.org/wiki/Cuisine_of_New_York_City

3. Methodology

In the first step we collected data about New York City:

1. Demographics data.
2. Boroughs, neighborhoods, and preffered food.
3. Available markets for supplies.
4. determinied that the preferred food in the city is Italian.

Second step is to analyze the collected data and define:

1. Use foursquare API to fetch the number of restaurants in Manhattan and Brooklyn boroughs.
2. The market saturation for restaurant business for Manhattan and Brooklyn boroughs.
3. Confirm that the preferred food is Italian (since it is for the whole city).
4. Use Kmeans to cluster the restaurants.
5. Pinpoint the most appropriate neighborhood or area to open a traditional Balkan food restaurant.

Data analysis

New york city Geographical Coordinates Data.

1. In this we load the data and explore data from newyork_data.json file.

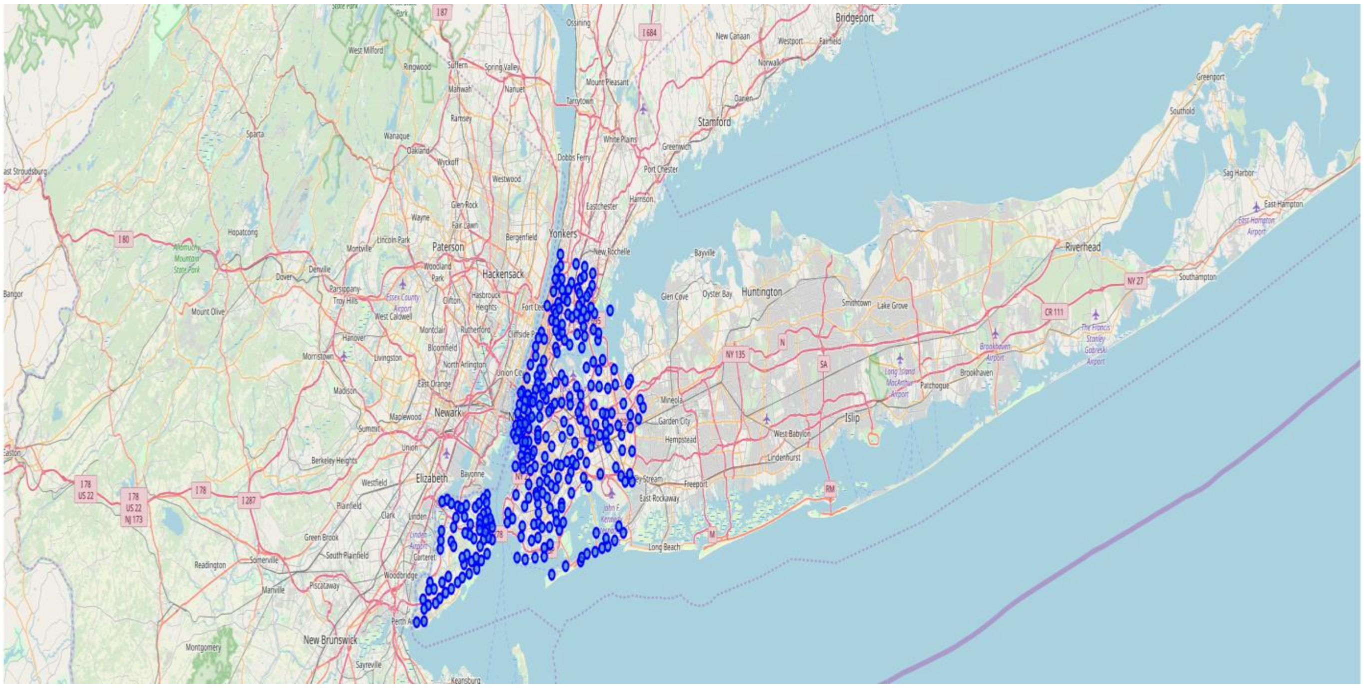
2. Transform the data of nested python dictionaries into a pandas dataframe.

3. This dataframe contains the geographical coordinates of New York city neighborhoods.

4. This data will used to get Venues data from Fouresquare.

5. We used geopy and folium libraries to create a map of New York city with neighborhoods superimposed on top.

New York neighbourhood visualization



The DOHMH Farmers Markets and Food Boxes dataset.

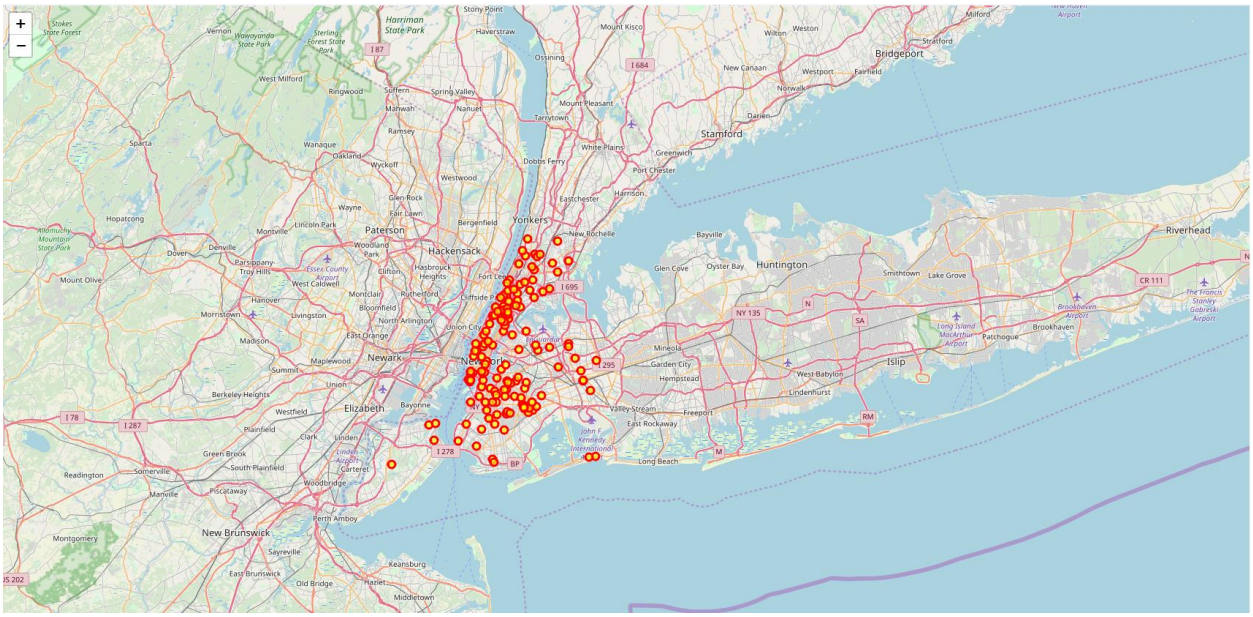
In this we will be using the data of Farmers Markets data.

There are totally 144 Farmers Markets in New York city. Highest number are in Manhattan and Brooklyn.

And lowest in Queens, Bronx and Staten Island.

The proof of this is as given below.

Farmers Market visualisation-New York City



To analyze New York city Population, Demographics and Cuisine , scrapped the data from Wikipedia pages given above in the data section. 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 soup). It creates a parse tree for parsed pages that can be used to extract data from HTML, which is useful for web scraping .

Insights from the data :

1. Manhattan (New York County) is the geographically smallest and most densely populated borough.
2. Manhattan's (New York County's) population density of 72,033 people per square mile (27,812/km²) in 2015 makes it the highest of any county in the United States and higher than the density of any individual American city.
3. Brooklyn (Kings County), on the western tip of Long Island, is the city's most populous borough.
4. Queens (Queens County), on Long Island north and east of Brooklyn, is geographically the largest borough.

	Borough	County	Population	GDP($)	GDP(per capita)	square miles	square km	persons/sq. miles	person/sq. km.
0	The Bronx	Bronx	1,471,160	28.787	19,570	42.10	109.04	34,653	13,231
1	Brooklyn	Kings	2,648,771	63.303	23,900	70.82	183.42	37,137	14,649
2	Manhattan	New York	1,664,727	629.682	378,250	22.83	59.13	72,033	27,826
3	Queens	Queens	2,358,582	73.842	31,310	108.53	281.09	21,460	8,354
4	Staten Island	Richmond	479,458	11.249	23,460	58.37	151.18	8,112	3,132
5		City of New York	8,622,698	806.863	93,574	302.64	783.83	28,188	10,947
6		State of New York	19,849,399	1,547.116	78,354	47,214	122,284	416.4	159

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,incorporating more immigration into the city than outmigration since the 2010 United States Census. 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.

	Racial composition	2010	1990	1970	1940
0	White	44.0%	52.3%	76.6%	93.6%
1	—Non-Hispanic	33.3%	43.2%	62.9%[249	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%[249	1.6%
4	Asian	12.7%	7.0%	1.2%	–

Cuisine 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 . Using this data we did word cloud.

NEW YORK CITY CUISINE:

Most Preferred Food in New York City –Italian, Purto Rican, Mexican, Jewish, Indian, Pakistani & Dominican.

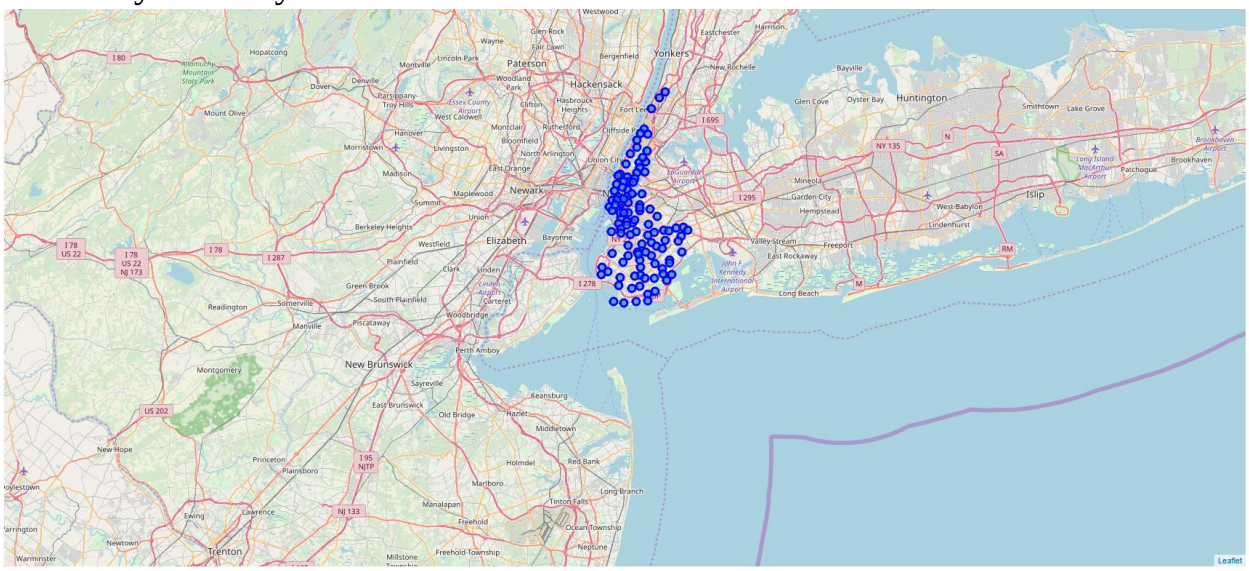
MANHATTAN CUISINE

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

BROOKLYN CUISINE

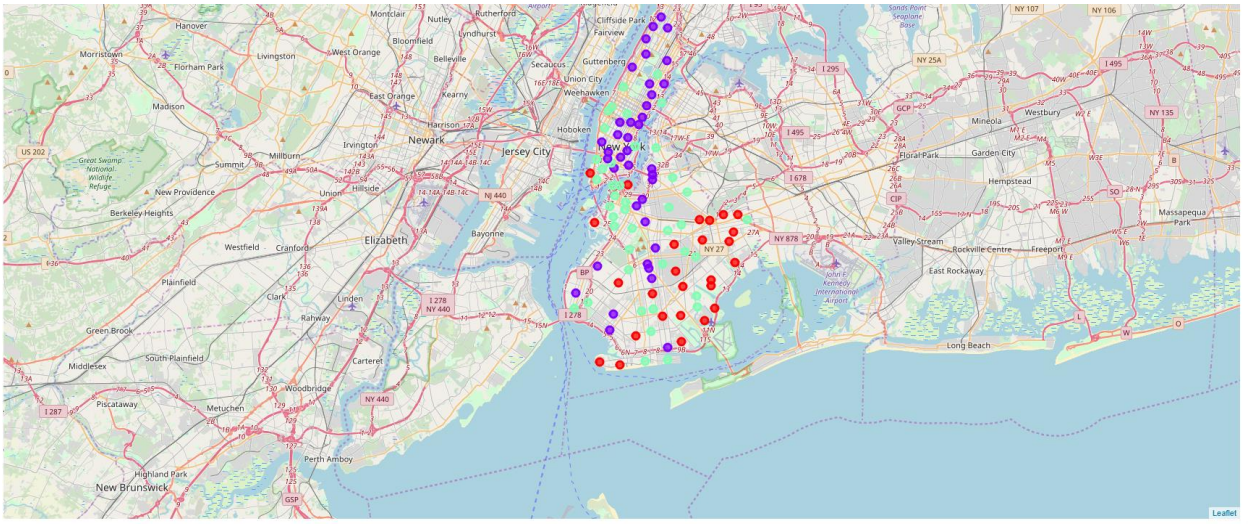
Most Preferred Food in Brooklyn is –Italian, Purto Rican & Mexican.

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. Brooklyn and Manhattan :



5. Results

From this venues data we filtered and used only the restaurant data for Brooklyn & Manhattan clustering and Bronx, Queens and Staten Island clustering. As we focussed only on restaurants business. Neighborhood K-Means clustering based on mean occurrence of venue category. In the below Map Visualization, we can see the different types of clusters created by using K-Means for Brooklyn & Manhattan.



Cluster0 : The Total and Total Sum of cluster0 has smallest value. It shows that the market is least saturated.

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

Cluster2: The Total and Total Sum of cluster2 has high value. It shows that the markets are saturated. Number of restaurants is high.

6. Discussion.

The analysis shows that the market is fairly saturated. There is not much room for opening new restaurant unless it has something
special to it. Offering a traditional Balkan food restaurant might have success, since there aren't many restaurants like that in Manhattan and Brooklyn boroughs.3

7. Conclusion

Starting a new business by opening a new Balkan food restaurant either in Manhattan or Brooklyn is a big risk, considering those boroughs are highly saturated with restaurants and most people prefer Italian food. If a decision is made to open a restaurant in those areas it should be in neighborhoods that belong to cluster2.

Further analysis of other New York City boroughs should be done in order to find the best place to open a new restaurant in New York.