

Coursera Capstone Project Report

The Battle of the Neighborhoods

1. Introduction and Problem Description:

The New York City is the popular in the United States. It has diverse cultural and provides lot of business opportunities. The business market is highly competitive. The cost of living and doing business is high as it is highly developed city. Hence it is important to be properly and carefully analyze the environment when starting a new business venture or any expansion in the business market. 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.

1.1. Problem Statement

Doing business in restaurant that prepares variety of international cuisine has the big opportunity as the people living around the city are from different diversity and culture. The restaurant would provide the services such as 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.

The Central and Eastern European immigrants, especially Jewish immigrants prefer bagels, cheese cake, hot dogs, knishes, and delicatessens. The Italian immigrants prefer New York-style pizza and Italian cuisine. The Jewish immigrants and Irish immigrants choose pastrami and corned beef. The Chinese and other Asian restaurants, sandwich joints, trattorias, diners, and coffeehouses are ubiquitous throughout the city. Also the licensed mobile food vendors exist across the city. The Middle Eastern foods such as falafel, shawarma and kebabs examples of modern New York street food. The city is home to "nearly one thousand of the finest and most diverse 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.

1.2. Target Audience:

Various factors need to be studied in order to decide on the Location such as city Population, New York City Demographics. Are there any Farmers Markets, Wholesale markets etc nearby so that the ingredients can be purchased fresh to maintain quality and cost? Are there any venues like Gyms, Entertainment zones, Parks etc nearby where floating population is high etc. Who are the competitors in that location?

If a well-funded Company Ltd. need to choose the correct location to start its first venture. First move is very important, thereby choice of location is very important to be successful they can franchise the same in other locations To recommend the correct location, the Company Ltd has asked me to lead in 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 Newyork city. The success criteria of the project will be a good recommendation of Neighborhood choice to open the Company Ltd which based on location that is Lack of such restaurants and nearest suppliers of ingredients.

2. Data and references

2.1. Data Requirements

- The neighbor city for New York in the other side of Manhattan is Brooklyn has been taken to perform the neighborhood battle between two cities. Geodata for Brooklyn location with venues established using Foursquare
- List of Manhattan (MH) neighborhoods with clustered venues established via Foursquare (as in Course Lab).

https://en.wikipedia.org/wiki/List_of_Manhattan_neighborhoods#Midtown_neighborhoods

- List of categories shop in Manhattan with addresses and geo data (lat,long)

[10]:

	name	categories	lat	lng
0	Artist & Craftsman Supply	Arts & Crafts Store	40.674341	-73.984297
1	J.J. Byrne Playground	Park	40.672939	-73.984085
2	Whole Foods Market	Grocery Store	40.675262	-73.988863
3	Washington Park	Park	40.672582	-73.986570
4	Dirty Precious	Cocktail Bar	40.676317	-73.986753
5	Russo's Mozzarella & Pasta	Deli / Bodega	40.673132	-73.983030

2.2. Data Sources, Data Processing and Tools used

- Brooklyn data and map has been created with use of venue details from Foursquare and Folium mapping
- Manhattan neighborhoods were obtained from Wikipedia and organized by Neighborhoods with geodata via pid installation for mapping with Folium.
- List of Subway stations was obtained via Wikipedia, NY Transit web site and Google map.
- List of apartments for rent was consolidated from web-scraping real estate sites for MH. The geolocation (lat,long) data was found with algorithm coding and using Nominatim.
- Folium map was the basis of mapping with various features to consolidate all data in ONE map where one can visualize all details needed to make a selection of apartment.

The Tools: Web-scraping of sites is used to consolidate data-frame information which was saved as csv files for convenience and to simplify the report. Geodata was obtained by coding a program to use Nominatim to get latitude and longitude of subway stations and also for each of (144 units) the apartments for rent listed.

Geopy_distance and Nominatim were used to establish relative distances. Seaborn graphic was used for general statistics on rental data.

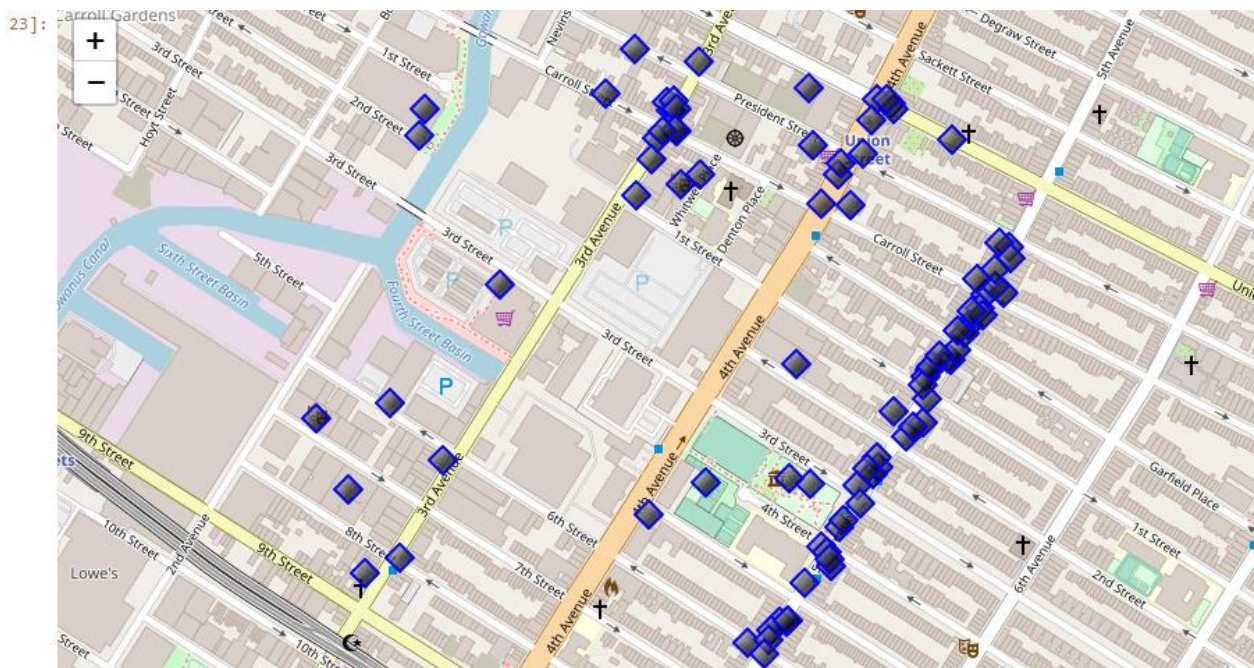
Maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy.

3. Methodology

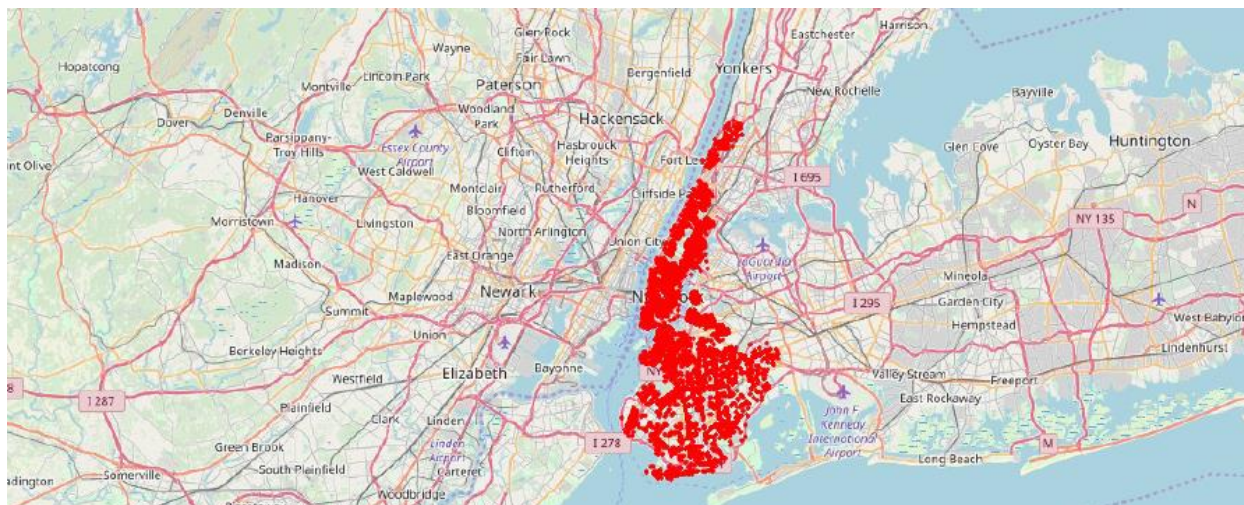
The methodology followed here is based on the mapping that described in data section. In order to facilitate the choice of at least two possibilities for opening the new restaurant, the information will be consolidated in ONE MAP where one can see the details of the neighborhood shops and the supplier details. The cluster of venues in the neighborhood and the relative location from a restaurant in Brooklyn. A measurement tool icon will also be provided. The popups on the map items will display living cost, location and cluster of venues applicable.

4. Execution and Results

4.1. Map of Brooklyn using latitude and longitude values



Brooklyn and Manhattan Venues Visualization has been generated to map the location and later used for trending venue findings.



Let's check how many venues were returned for each neighborhood.

`Brooklyn_venues.groupby('Neighborhood').count()`

Grouping the rows by neighborhood and by taking the mean of the frequency of occurrence of each category. Analyse with five common venues.

----Battery Park City----

	venue	freq
0	Park	0.08
1	Hotel	0.05
2	Coffee Shop	0.05
3	Gym	0.04
4	Memorial Site	0.03

----Carnegie Hill----

	venue	freq
0	Coffee Shop	0.06
1	Pizza Place	0.06
2	Café	0.04
3	Cosmetics Shop	0.04
4	Yoga Studio	0.03

----Central Harlem----

	venue	freq
0	African Restaurant	0.07
1	Cosmetics Shop	0.04
2	American Restaurant	0.04
3	Seafood Restaurant	0.04
4	French Restaurant	0.04

----Chelsea----

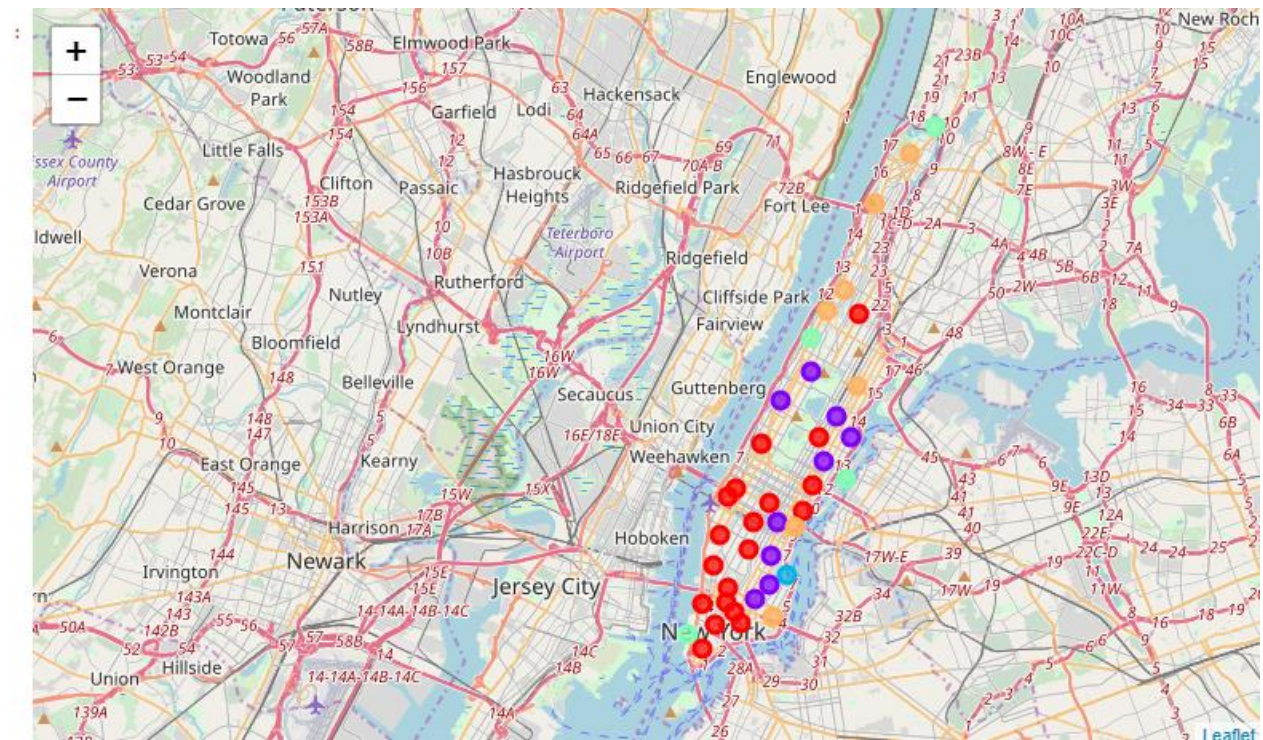
	venue	freq
0	Coffee Shop	0.06
1	Italian Restaurant	0.05
2	Bakery	0.05
3	Ice Cream Shop	0.04
4	Nightclub	0.04

1]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Battery Park City	Park	Hotel	Coffee Shop	Gym	Wine Shop
1	Carnegie Hill	Pizza Place	Coffee Shop	Cosmetics Shop	Café	French Restaurant
2	Central Harlem	African Restaurant	Cosmetics Shop	Bar	American Restaurant	Seafood Restaurant
3	Chelsea	Coffee Shop	Italian Restaurant	Bakery	Ice Cream Shop	Nightclub
4	Chinatown	Chinese Restaurant	Cocktail Bar	Bakery	American Restaurant	Vietnamese Restaurant

Run k-means to cluster the neighborhood into 5 clusters.

Examine the Clusters: each cluster and determine the discriminating venue categories that distinguish each cluster.



5. Conclusion:

The Manhattan and Brooklyn has the cuisine of many countries have been identified and selected to perform the k-means clustering to merge both the locations to analyze the battle of neighborhood in the determination of appropriate location to initiate the risk free business center to open the restaurant.

As per the neighborhood analysis of the restaurant in Brooklyn and Manhattan, the trending venues are not determined the selected the restaurant in the given location. A venue with lowest occurrence has been identified and fix as the feasible location to open the restaurant.