

Coursera Applied Data Science Capstone Project

Paul J. Walker

The Battle of the Neighborhoods – Report (Week 1)

1. Introduction Section

A. Business Problem

I. Background Information

The City of New York, usually called either New York City (NYC) or simply New York (NY), is the most populous city in the United States. NYC has an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles. A global power city, New York City has been described as the cultural, financial, and media capital of the world, and exerts a significant impact upon commerce, entertainment, research, technology, education, politics, tourism, art, fashion, and sports.

New York is a central hub for business across all markets and thus every industry or market in the area is highly competitive. In addition, the cost of conducting business as well as the associated risk is also some of the highest in the world. Because of this, new business ventures must analyze their markets carefully and efficiently. Various analyses of the area and its markets will provide a good base understanding of a specific business environment which in turn will help business strategically tackle their respective markets. This insight will help in the reduction of risk as well as maximize potential markets.

II. Problem Description

Restaurants are one of the most common new business types to emerge in a populated city and so there are thousands if not millions of competitors for restaurant owners. New York is one of the most known cities in the world for its excellent cuisines. The various cultures of NY has contributed to one of the most diverse and competitive restaurant industries in the world. Its food is inspired by an array of influential cultures such as;

- Central and Eastern European immigrants - bagels, cheesecake, hot dogs, knishes, and delicatessens
- Italian immigrants - New York-style pizza and Italian cuisine
- Jewish and Irish immigrants - pastrami and corned beef
- Chinese and other Asian immigrants 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.

- The city is home to "nearly one thousand of the finest and most diverse haute cuisine restaurants in the world", according to Michelin.

Clearly in order to not only thrive, but survive in this market, it is essential for business owners to strategically plan using analytics.

Various factors must be considered about the surrounding area of a location before it can be selected as a potential candidate. Some of these factors are;

- New York 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?
- Cuisine served / Menu of the competitors
- Segmentation of the Borough
- Untapped markets, saturated markets, etc.

and numerous others. Although a restaurant can have sufficient funding, it is imperative for them to choose a prime first location to begin its venture. The goal of a successful first location is to promote and grow the business in order to replicate the same success in numerous other locations.

B. Target Audience

As a data scientist I am to develop a method to select a successful first location. My objective for this task is to locate and recommend to the management of some Company X which neighborhood of New York City will be the best choice to start a new restaurant. Management expects to understand the rationale, methodology, results, and associated recommendations made. This will be useful for any who wish to open a new restaurant in the NYC area.

Company X wishes to open a new healthy, organic, and nutrient filled restaurant. Their menu will feature fresh made smoothies, soups, sandwiches, wraps, protein shakes, salads, and entrees.

C. Success Criteria

The success criteria for this project is simple – a recommendation of a borough/neighborhood to Company X based on lack of similar restaurants in that location, the proximity to the nearest suppliers of necessary ingredients.

2. Data Section

A. Required Data

I. Relevant & Content-specific data

Neighborhood Dataset: NYC has a total of 5 boroughs and 306 neighborhoods. In order for us to segment these neighborhoods and explore them we require a dataset that contains the 5 boroughs, the 306 neighborhoods in the boroughs, and the latitudinal and longitudinal coordinates of each of these neighborhoods.

Luckily this [dataset](#) exists on the internet and is easily accessible.

[15]:	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Farmers Markets Dataset: A farmers market is a public site used by two or more local producers for the direct sale and delivery of farm fresh products to consumers. Farmers markets can sell fresh fruits and vegetables, dairy products, fish, meat, baked goods, and other minimally processed foods.

We will be using the [NYC Department of Health and Mental Hygiene's farmers markets and food boxes](#) datasets to examine these markets.

Borough	Market Name	Street Address	Latitude	Longitude	Days of Operation	Hours of...
Bronx	170 Farm Stand	E 170th St & Townsend Ave	40.839882	-73.916783	Wednesday	2:30 - 6:30 p.m.
Bronx	Bissel Gardens Farmers Market	Baychester Ave & E 241st St	40.90152	-73.846937	Wednesday & Saturday	9 a.m. - 5 p.m.
Bronx	Bronx Borough Hall Greenmarket	Grand Concourse bet E 161st & 162nd Sts	40.826835	-73.922621	Tuesday	8 a.m. - 4 p.m.
Bronx	Bronx Family Center Go!Healthy Farm Sta...	1515 Southern Blvd	40.833734	-73.889934	Thursday	9:30 a.m. - 3:...
Bronx	BronxWorks Farm Stand	1130 Grand Concourse	40.832331	-73.91959	Thursday	10 a.m. - 4 p.m.
Bronx	CS211 Go!Healthy Farm Stand	1919 Prospect Ave	40.843546	-73.889755	Wednesday	2:30 - 5:30 p.m.
Bronx	Fordham Plaza Greenmarket	Park Ave bet E 189th St & E Fordham Rd	40.86153	-73.891063	Friday	8 a.m. - 4 p.m.
Bronx	Friends of Van Cortlandt Park Amalgamat...	Orloff Ave & Gale Pl	40.884532	-73.892172	Wednesday	2 - 7 p.m.
Bronx	Harvest Home Co-op City Farm Stand	Asch Loop & Aldrich St	40.870249	-73.830963	Wednesday	8 a.m. - 4 p.m.
Bronx	Harvest Home Hunts Point Farm Stand	E 163rd St & Hunts Point Ave	40.820986	-73.891613	Wednesday	8 a.m. - 4 p.m.
Bronx	Harvest Home Jacobi Hospital Farmers M...	1400 Pelham Pkwy S, at Eastchester Rd	40.857427	-73.847079	Tuesday & Friday	8 a.m. - 4 p.m.
Bronx	Harvest Home Mt. Eden Farmers Market	Mt Eden & Morris Aves	40.842924	-73.909558	Tuesday & Thursday	8 a.m. - 4 p.m.

II. Supporting data

NYC Demographics & Cuisines: For our analysis we will be using data from Wikipedia;

[New York Population](#)

[New York City Demographics](#)

[New York City Demographics \(Cont.\)](#)

[Cuisine of New York City](#)

NYC Geographical Coordinates: NYC geographical coordinates will be utilized as input for the Foursquare API which will be used to provide information for all surrounding venues in each neighborhood. We will then use the Foursquare API to explore the neighborhoods of NYC.

[29] :

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
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	Dunkin'	40.877136	-73.906666	Donut Shop
4	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop