

|  |
| --- |
| Capstone Project Report |
|  |
| February 8  Authored by: Sneha Devadas |

# Identifying an Ideal Location for a Supplement Shop in Manhattan

Introduction

Manhattan, known regionally as **the City** and the urban core of the New York metropolitan area, is the most densely populated of the five boroughs of New York City. Manhattan has been described as the cultural, financial, media, and entertainment capital of the world, and the borough hosts the United Nations Headquarters. Anchored by Wall Street in the Financial District of Lower Manhattan, New York City has been called both the most economically powerful city and the leading financial center of the world. Manhattan real estate has since become among the most expensive in the world, with the value of Manhattan Island, including real estate, estimated to exceed US$3 trillion in 2013; median residential property sale prices in Manhattan approximated US$1,600 per square foot ($17,000/m2) as of 2018, with Fifth Avenue in Midtown Manhattan commanding the highest retail rents in the world, at US$3,000 per square foot ($32,000/m2) per year in 2017.

To setup and run a business entity in such location requires a great deal of consideration and deliberation. Among important factors such as financial, operational and legal considerations, the location of the proposed business venture plays a very important role. It is so powerful a factor that it can make or break the business. We need to identify such a location that it effectively addresses the gap in demand. In other words, we need to identify an area which has a lot of market potential waiting to be tapped. One way to identify this is by analyzing the presence and popularity of such business ventures in the proposed area versus the demand.

**Business Problem/Problem Statement**

*Identify a location for setting up a new Supplement Store in one of the neighborhoods in Manhattan, which may help to maximize the sales.*

**Stakeholder**

An individual or a partnership interested to setup a Supplement Shop in the neighborhoods of Manhattan.

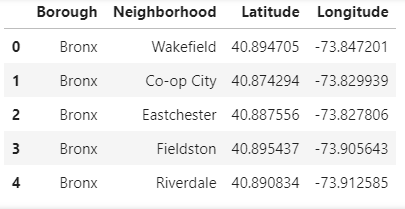
**Data**

To address the business problem defined in the previous section, we need the following data:

* Name of all the neighborhoods in Manhattan along with the geographical coordinates (latitude and longitude).
* List of popular venues in each of these neighborhoods.

“newyork\_data.json” file available in the following link (<https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json>) contains the a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough in New York , where Manhattan is one borough, as well as the latitude and longitude coordinates of each neighborhood.

For example, below is snapshot of the data extracted from “newyork\_data.json” file.



This dataset is cleaned and processed to retrieve data of the neighborhoods in Manhattan alone.

List of popular venues in each of the neighborhood is obtained with the help of Foursquare API. The data is obtained in the form of a json file, which is then processed to retrieve the desired information i.e., the venue name, category to which it belongs, the latitude and longitude of the venue.

For example, the information extracted from the response.json file for the neighborhood Stuyvesant Town in Manhattan is as follows.



These datasets are then used for segmenting and clustering the neighborhoods in Manhattan. Further analysis of these clusters will help us to identify the neighborhood or list of neighborhoods which are promising.

**Methodology**

In order to choose an ideal neighborhood to set up a Supplement Shop it is necessary to understand the characteristics of each neighborhood. The neighborhood chosen should be such that there is a considerable demand for the service we are offering and there are few market players or competitors. In this project, the approach used was to analyze the popular venues in a particular neighborhood, to identify potential demand as well as possible competitors. In simple words, look for a neighborhood which has no or few supplement shops among the popular venues. At the same time, it should also host venues related to health and fitness such as gym, playgrounds, hospital etc.

After extracting and preprocessing the required data, the neighborhoods were visualized on a map for better understanding.

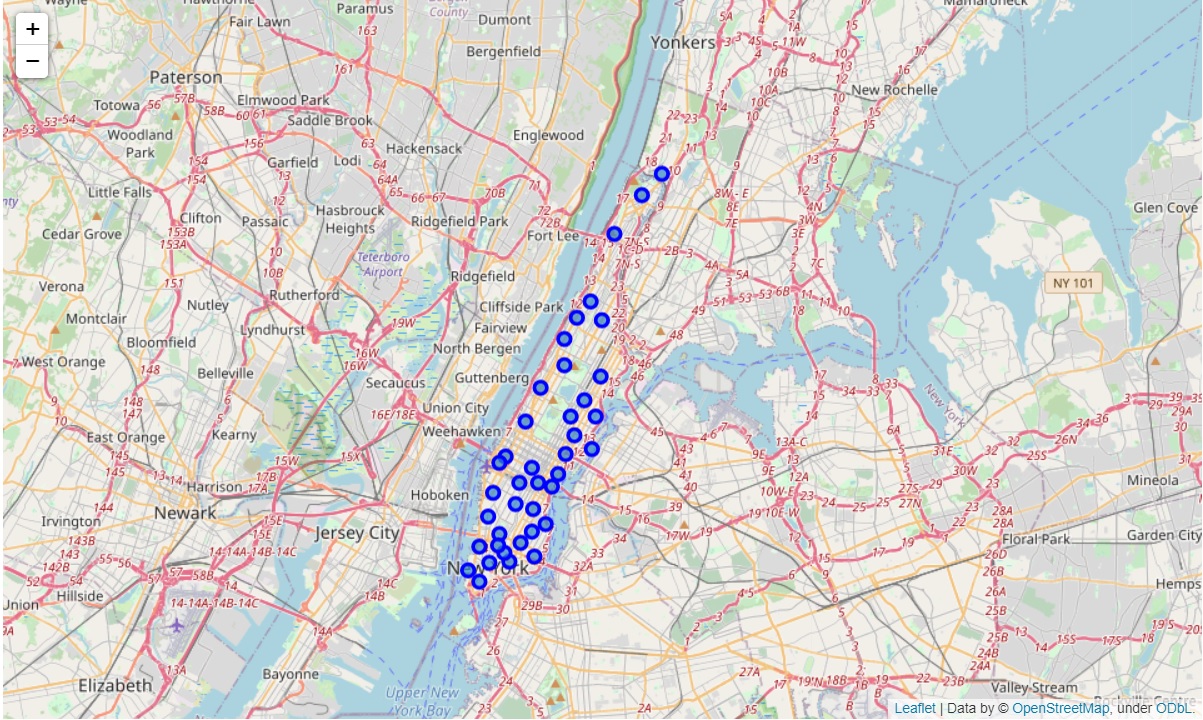
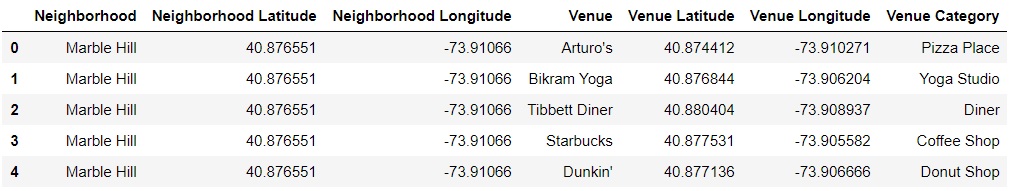


Figure Neighborhoods of Manhattan

Next step is to identify the popular venues in each neighborhood. We used Foursquare API to provide the necessary data. The response was processed to obtain a dataset as shown below.

Table Processed Foursquare API output



This data was then consolidated on the basis of neighborhood. Further, frequency of occurrence of each category in each neighborhood was calculated. Based on the frequency of occurrence, top 10 venues were selected for each neighborhood.

Table Top 10 venues of each neighborhood in Manhattan



The above dataset was used for segmenting and clustering the neighborhoods of Manhattan into three groups using k-means clustering algorithm. The value of k was determined by optimizing the value of Squared Error or Cost function as shown below.

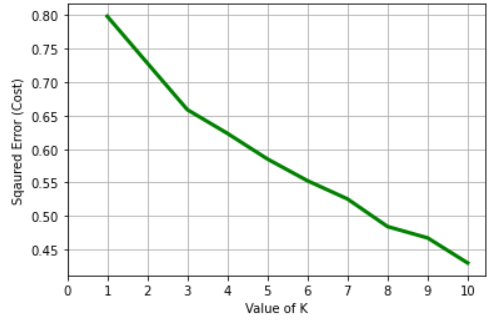
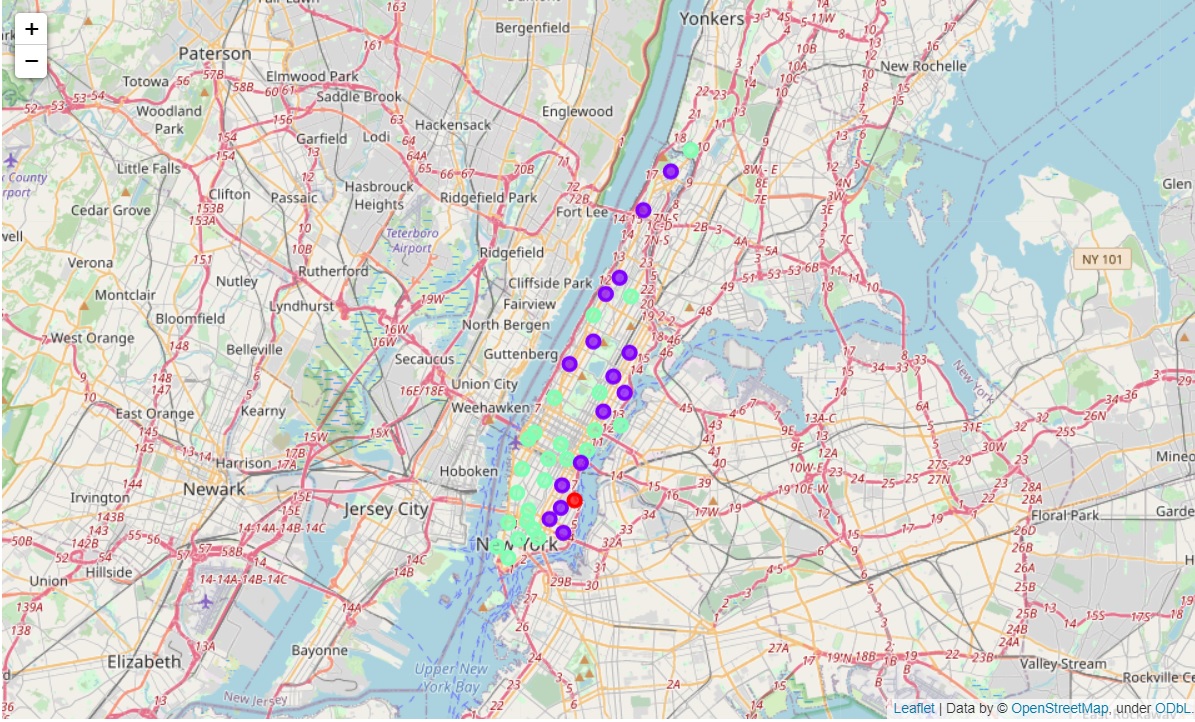


Figure Squared Error (Cost) for each value of k

The clusters obtained were then analyzed to identify defining characteristics.

**Results and Discussion**

Upon running k-means clustering, with k = 3 on the neighborhood data, following clusters were obtained.



Legend

Cluster 1

Cluster 2

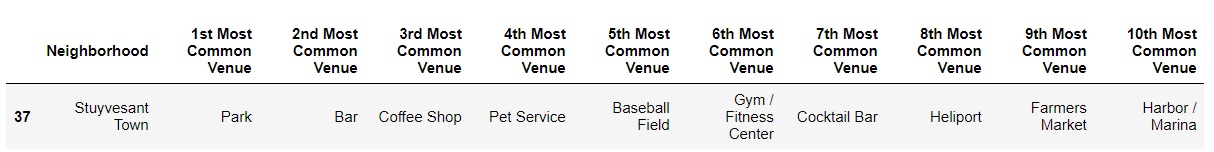
Cluster 3

Figure Clusters of Neighborhoods

Cluster 1

Neighborhoods and the top 10 popular venues of cluster 1 are given below. As can be seen clearly, it contains only one neighborhood namely Stuyvesant Town. Also, it has a lot of outdoor recreational and fitness related venues. This could possibly be a good fit for our supplement shop. Nonetheless let’s analyze the remaining clusters as well.

Table Neighborhood and popular venues of cluster 1



Cluster 2

Neighborhoods and the top 10 popular venues of cluster 2 are given below. Let’s get the frequency distribution of this data to identify the most popular venues of this cluster.

Table Neighborhoods and popular venues of cluster 2



We notice that Mexican Restaurant, Coffee Shop, Café, Pizza Place, Italian Restaurant, Bar, Sushi Restaurant, Bakery, Deli / Bodega and Latin American Restaurant are the most popular venues. We also notice that fitness related or outdoor recreational venues are not that popular in this neighborhood. Different kinds of restaurants and coffee shops are predominant in this neighborhood.

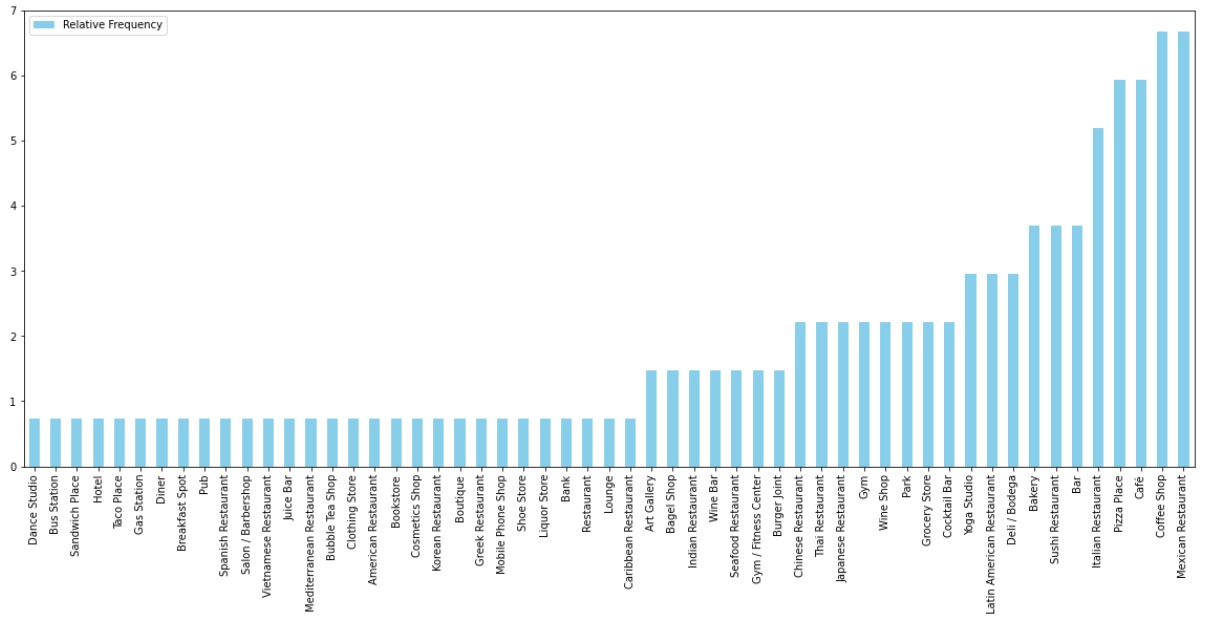


Figure Frequency Distribution of popular venues in cluster 2

Cluster 3

Neighborhoods and the top 10 popular venues of cluster 3 are given below. Let’s get the frequency distribution of this data to identify the most popular venues of this cluster.

Table Neighborhoods and popular venues of cluster 3



We notice that Coffee Shop, Italian Restaurant, American Restaurant, Gym/Fitness Center, Hotel, Park, Gym, Café, Cocktail Bar and Bakery are the most popular venues. We also notice that fitness related or outdoor recreational venues are quite popular in this neighborhood. This cluster is a balanced mix of restaurants, fitness and recreational amenities as well commercial establishments.

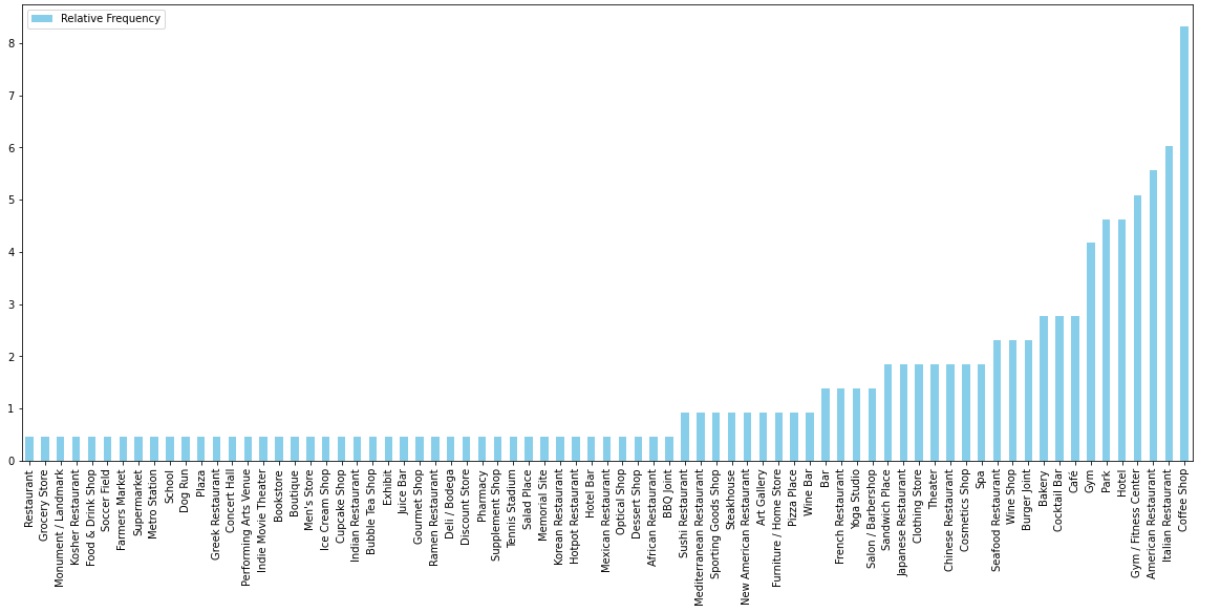


Figure Frequency Distribution of popular venues in cluster 3

Since fitness related venues are quite popular, we can consider the neighborhoods in this cluster as well as a promising location. Please note that the neighborhood of Marble Hill already has a Supplement Store. However, the remaining neighborhoods may prove to be promising, but further analysis is required to pin point the neighborhoods.

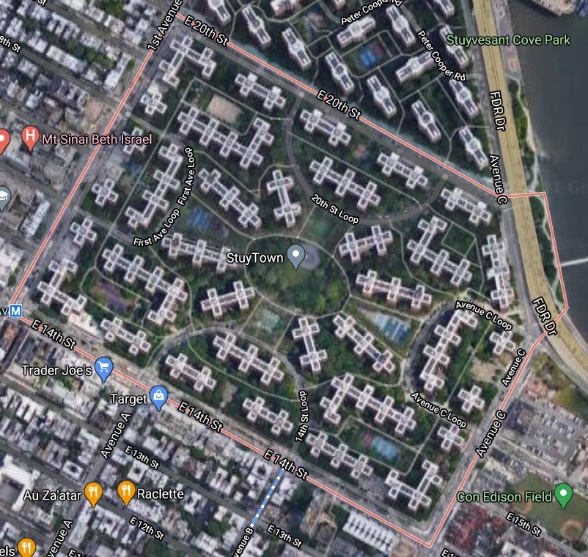
Focusing more on cluster 1, we find that Stuyvesant Town is a very unique neighborhood. **Stuyvesant Town–Peter Cooper Village** is a large, post-World War II private residential development on the east side of the New York City borough of Manhattan. The complex consists of 110 red brick apartment buildings on an 80-acre (32 ha) tract. This neighborhood has almost all sports grounds such as baseball field, tennis field etc. It also has a lot of recreational venues like parks, health and fitness related establishments such as gyms, hospitals etc. A list of 16 most popular venues of Stuyvesant Town is given below.

Table List of popular venues in Stuyvesant Town



Apart from Supplement Store, other promising business ventures in this neighborhood are Juice Bar, Dance Studio, Yoga Studio and Salad Place.

However, it should be noted that other factors like the operational cost, taxes and other legal requirements should also be analyzed for this neighborhood.

**Conclusion**

The neighborhoods of Manhattan were studied to identify a promising location for a Supplement Store. The defining feature data obtained from the Foursquare API was used for segmenting and clustering the neighborhoods into three groups using k-means clustering. Further analysis of the clusters suggests cluster 1 and cluster 3 as promising clusters. Among these neighborhoods Stuyvesant **Town** appears to be the best fit.