

Optimal Gym Locations in Brooklyn

Alexander John Newman

February 22, 2021

1. Introduction

1.1 Background

Amidst the Covid-19 pandemic, lockdowns and restrictions have been imposed worldwide. Countries all over the world have responded differently to this virus. Some have banned international travel while others have maintained the “business as usual” status quo. Nonetheless, major cities all over the world, mainly dense ones (regarding population), unsurprisingly have seen the highest case numbers. This often times leads to more stringent measures such as shutting down bars, restaurants, and gyms. These venues previously mentioned serve as a staple in certain countries and major cities, New York City being one of those places. With the obesity epidemic that has run rampant throughout the United States for a number of years, gyms became an integral part of the culture as people worked to improve upon facets of their strengths and their immune systems. With gyms being shut down in many different parts of the country, this causes many people to sink into much more sedentary lifestyles they may not have been experiencing before. With this comes expected weight gain and weakened immune systems, the two things you don’t want when Covid-19 is a concern.

1.2 Problem

A client of mine would like to know where the most saturated and unsaturated markets are in the borough of Brooklyn. Using this data, we can then make an appropriate decision as to where to open this health/wellness center.

1.3 Interest

While this is both a useful business in terms of helping others with their health goals, it’s also a potentially lucrative business. On the tail end of the current coronavirus pandemic, people will start becoming more health-conscious. In a society of people trying to be as healthy as possible, a gym has the client base to prosper.

2. Data Acquisition/Cleaning

2.1 Data Sources

The data was obtained from NYU's [Spatial Data Repository](#). This Repository had an associated .json file containing all of the information of the five boroughs that we would need to pull our data from. Our project focuses specifically on the borough of Brooklyn which we sort through to find optimal locations for a new gym.

2.2 Data Cleaning

The data cleaning portion of this project wasn't too labor-intensive. All that needed done was to select the borough in which we wanted to investigate. Every location was listed by its respective latitude/longitude coordinates. After selecting the borough of Brooklyn, I then was able to use the Foursquare API to search for Venues related to gym/fitness centers and display them on a Google Map. This allowed for a viewer friendly interpretation of the dataframes displayed prior to this point in the Jupyter notebook.

Any missing values from the provided dataset were ignored. The final merged dataframe before I began clustering had a row count of 1792. This means that within the borough of Brooklyn, there are 1792 venues (as of NYU's 2014 data) that either are or are related to gym/fitness centers. This list includes all sorts of different styles of gyms (martial arts, climbing, etc) and health centers (chiropractors, weight loss centers, etc). To obtain the consumer information we want, we needed to sort the clusters by popularity of said venues.

2.3 Data Selection

The goal of this project is to find the optimal location in Brooklyn to open a gym. In order to gain a general idea, we want to group the neighborhoods of the borough into clusters. Following this clustering, we want to create dataframes that sort the venues by popularity per cluster.

By sorting the clusters by popularity of a particular venue, we're able to cut out all of the venues that provide mostly null values for our dataset. This highlights the importance of places like gyms, gyms/fitness centers, yoga studios, and martial arts centers (these appear to be the hotspots in all clustered neighborhoods).

Lastly, I make a couple personal decisions, such as disregarding the data concerning the 4th-nth most popular venues as well as assuming that a 'gym' is the same thing as a 'gym/fitness center'.

3. Data Analysis

3.1 Using k-means Clustering

Of the 70 selected neighborhoods within Brooklyn, five were given further consideration as a possible location to open this gym. These five were selected as the best options from their respective clusters.

Within each cluster, the neighborhoods are listed by order of their indices with the following columns being listed in order of popularity of each venue. To start, we'll analyze cluster 0 first:

Cluster 0

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Bay Ridge	Gym / Fitness Center	Gym	Yoga Studio	Martial Arts School
1	Bensonhurst	Gym / Fitness Center	Martial Arts School	Gym	Yoga Studio
3	Greenpoint	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
14	Brownsville	Gym / Fitness Center	Gym	Baseball Field	Boxing Gym
21	Red Hook	Gym / Fitness Center	Track	Martial Arts School	Baseball Field
22	Gowanus	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
23	Fort Greene	Gym / Fitness Center	Yoga Studio	Gym	Martial Arts School
27	Starrett City	Gym / Fitness Center	Gym	Martial Arts School	Gym Pool
35	Dyker Heights	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
40	Downtown	Gym / Fitness Center	Gym	Yoga Studio	Gym Pool
41	Boerum Hill	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
43	Ocean Hill	Gym / Fitness Center	Martial Arts School	Boxing Gym	Gym
53	Fort Hamilton	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
56	Rugby	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
64	Broadway Junction	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
67	Highland Park	Gym / Fitness Center	Gym	Martial Arts School	Baseball Field
69	Erasmus	Gym / Fitness Center	Gym	Yoga Studio	Residential Building (Apartment / Condo)

Circled is the neighborhood Ocean Hill. Ocean Hill was the only neighborhood with a 'gym' listed as the '4th Most Common Venue' according to the Foursquare API. This is the lowest rank of popularity for a 'gym' within this cluster. However, 'Gym / Fitness Center's appear to be the most common venue for cluster 0.

Cluster 1

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
6	Sheepshead Bay	Yoga Studio	Gym	Gym / Fitness Center	Martial Arts School
7	Manhattan Terrace	Gym	Gym / Fitness Center	Pilates Studio	Weight Loss Center
13	Prospect Heights	Gym	Yoga Studio	Gym / Fitness Center	Cycle Studio
15	Williamsburg	Gym	Yoga Studio	Gym / Fitness Center	Pilates Studio
16	Bushwick	Gym	Yoga Studio	Gym / Fitness Center	Martial Arts School
17	Bedford Stuyvesant	Gym	Yoga Studio	Gym / Fitness Center	Athletics & Sports
18	Brooklyn Heights	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
19	Cobble Hill	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
20	Carroll Gardens	Gym / Fitness Center	Gym	Pilates Studio	Yoga Studio
24	Park Slope	Gym / Fitness Center	Gym	Yoga Studio	Martial Arts School
26	East New York	Gym	Yoga Studio	Gym / Fitness Center	Track
32	Coney Island	Yoga Studio	Gym	Gym / Fitness Center	Residential Building (Apartment / Condo)
33	Bath Beach	Gym / Fitness Center	Gym	Yoga Studio	Martial Arts School
37	Marine Park	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio
38	Clinton Hill	Gym	Yoga Studio	Gym / Fitness Center	Pilates Studio
42	Prospect Lefferts Gardens	Gym	Gym / Fitness Center	Yoga Studio	Gym Pool
44	City Line	Yoga Studio	Martial Arts School	Gym	Gym / Fitness Center
47	Prospect Park South	Gym	Gym / Fitness Center	Yoga Studio	Martial Arts School
49	East Williamsburg	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
50	North Side	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
51	South Side	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
52	Ocean Parkway	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio
54	Ditmas Park	Gym	Gym / Fitness Center	Yoga Studio	Pilates Studio
63	Weeksville	Yoga Studio	Gym / Fitness Center	Gym	Pilates Studio
66	Homecrest	Gym / Fitness Center	Yoga Studio	Gym	Pilates Studio
68	Madison	Gym / Fitness Center	Gym	Martial Arts School	Yoga Studio

Circled is the City Line neighborhood. Here, we see that both 'Gym' and 'Gym / Fitness Center' are listed as the 3rd and 4th most common venues, respectively. Being that City Line is the only neighborhood in this cluster where this is the case, it was an easy decision to select it as an optimal location for a new gym to open up.

Cluster 2

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
5	Brighton Beach	Gym	Gym / Fitness Center	Yoga Studio	Athletics & Sports
8	Flatbush	Gym	Gym / Fitness Center	Martial Arts School	Yoga Studio
9	Crown Heights	Gym	Gym / Fitness Center	Gym Pool	Boxing Gym
10	East Flatbush	Gym	Yoga Studio	Gym / Fitness Center	Residential Building (Apartment / Condo)
28	Canarsie	Gym	Gym / Fitness Center	Martial Arts School	Weight Loss Center
29	Flatlands	Gym	Yoga Studio	Baseball Field	Boxing Gym
46	Midwood	Gym	Gym / Fitness Center	Martial Arts School	Dance Studio
48	Georgetown	Gym	Martial Arts School	Gym / Fitness Center	Baseball Field
55	Wingate	Gym	Gym / Fitness Center	Pilates Studio	Gym Pool
58	New Lots	Gym	Track	Gym / Fitness Center	Martial Arts School
59	Paerdegat Basin	Gym	Gym / Fitness Center	Weight Loss Center	Baseball Field
61	Fulton Ferry	Gym	Gym / Fitness Center	Yoga Studio	Boxing Gym
62	Vinegar Hill	Gym	Gym / Fitness Center	Gym Pool	Yoga Studio
65	Dumbo	Gym	Yoga Studio	Gym / Fitness Center	Pilates Studio

Circled is the Flatlands neighborhood. In this cluster, Gyms are clearly the most popular type of venue of people to attend. However, Flatlands is the only neighborhood where ‘Gym / Fitness Center’ doesn’t appear in the top 4 most common/popular places for locals to attend/visit. Thus, it appears that the Flatlands neighborhood has the smallest amount of competition as far as this gym market is concerned.

Cluster 3

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
29	Sea Gate	Yoga Studio	Gym / Fitness Center	Track	Football Stadium

Cluster 3 calls for a small gated community in the southeast peninsula of Brooklyn. Sea Gate is a gated residential area, with what I assume is not much space for commercial properties and the likes. We’ll talk about the viability of Sea Gate more in the discussion section.

Cluster 4

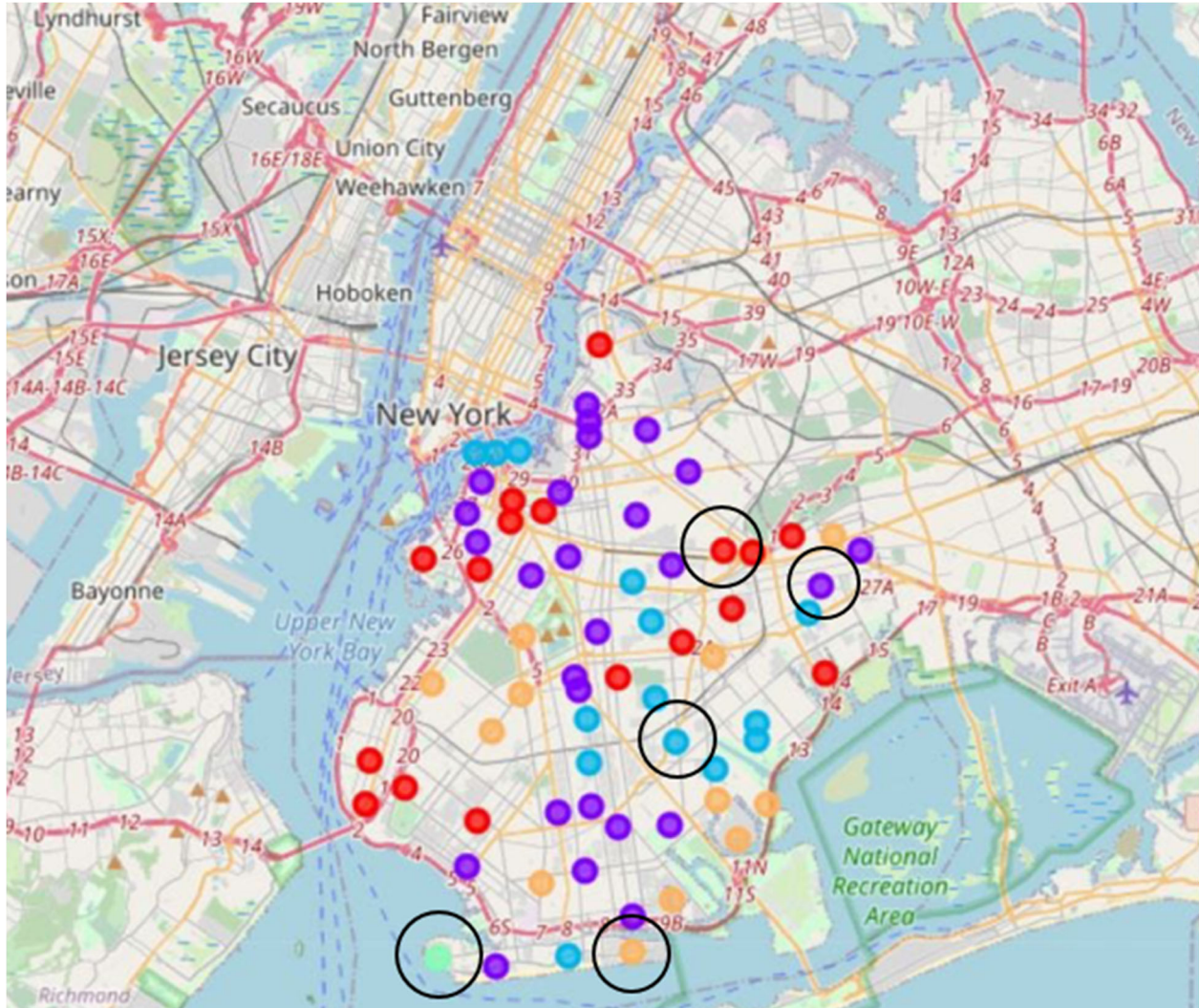
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
2	Sunset Park	Gym	Yoga Studio	Martial Arts School	Cycle Studio
4	Gravesend	Gym	Martial Arts School	Gym / Fitness Center	Yoga Studio
11	Kensington	Gym	Martial Arts School	Yoga Studio	Gym / Fitness Center
12	Windsor Terrace	Gym	Martial Arts School	Yoga Studio	Pilates Studio
25	Cypress Hills	Gym / Fitness Center	Gym	Martial Arts School	Baseball Field
30	Mill Island	Gym	Martial Arts School	Gym Pool	Yoga Studio
31	Manhattan Beach	Martial Arts School	Gym	Yoga Studio	Pilates Studio
34	Borough Park	Gym	Yoga Studio	Martial Arts School	Baseball Field
36	Gerritsen Beach	Gym	Gym / Fitness Center	Martial Arts School	Medical Center
45	Bergen Beach	Martial Arts School	Gym Pool	Gym	Yoga Studio
57	Remsen Village	Gym	Martial Arts School	Gym / Fitness Center	Baseball Field
60	Mill Basin	Gym	Martial Arts School	Gym / Fitness Center	Gym Pool

Manhattan beach in Cluster 4 is our selected neighborhood of interest. I chose Manhattan Beach over Bergen Beach, as it wasn’t entirely clear what is meant by ‘Gym Pool’. A ‘Gym Pool’ could simply be describing a gym that happens to have a pool in it for members to use. With this assumption, Manhattan beach has less common/popular venues pertaining to ‘Gym’s and ‘Gym / Fitness Center’s.

3.2 Map Locations

Edited in post

Associated Map Locations



Circled neighborhoods (left to right): Sea Drive, Manhattan Beach, Flatlands, Ocean Hill, City Line

4. Results

4.1 Neighborhoods of Interest

In Brooklyn, the five neighborhoods with the highest probability of opening a successful gym are Sea Drive, Manhattan Beach, Flatlands, Ocean Hill, and City Line. Of the five, Flatlands and Ocean Hill appear to have the highest chance of being locations in which competition with other local gyms may be an issue. Sea Drive and Manhattan Beach showed to be the most optimal

places to open a new gym purely based off of the popularity of health/wellness related venues in the area. Lastly, City Line appears to have the potential for some competition to be present but is nothing close to being an oversaturated market. With Manhattan Beach and Sea Gate being secluded residential communities, I believe the best location for this gym would be in City Line.

5. Discussion

5.1 Other Variables

It's important here to consider the demographics of these neighborhoods. A couple outside factors play a big role in determining the demand for this product. Things like socioeconomic status, nationality, and age are all crucial factors in who is or is not among the demographic of clients using a fitness facility. Socioeconomic status will be a key indicator as to whether or not the community can afford your services. Without the income, citizens within the neighborhood will not be thinking about attending a gym. Although exercise is objectively good for your health, the habit of regularly working out is an integral part of more western cultures. Of course, being in America, and facing the obesity epidemic year after year, exercise has become a critical part of our society. However, in other nations, the people don't stress the need to regularly attend work out sessions. Being in a borough of NYC, a more diverse part of the United States, it's important to consider the ethnic background of potential clients within these neighborhoods. Lastly, and probably the most straight-forward limited factor not tested for, would be age. Children and senior citizens are not among the standard distribution of regular gym goers. To pick a neighborhood with an average age somewhere in the range of 25-45 would likely be the optimal scenario. Again, all these mentioned factors need testing for, in order to obtain a more accurate indication as to where the optimal gym location is.

Lastly, as previously mentioned, this data is taken from a 2014 dataset. It's entirely possible that the markets within these cities are entirely different and to consider these changes would be appropriate. Leveraging newer data on neighborhood venues as well as coupling that with demographic statistics per neighborhood would serve as the optimal data pool to derive deliverables from. Until then, City Line proves to be the most optimal spot to open a fitness facility or gym for the public.