**The Battle of the Neighborhoods – Report**

1. **Introduction & Business Problem:**

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, health, theater, fashion, and the arts in the United States. 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.

**Business Problem**

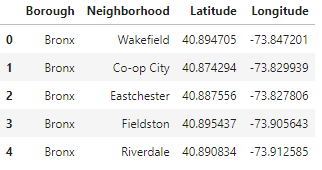
The City of New York is famous for its health. Its fitness culture includes an array of international techniques influenced by the city's immigrant history. Fitness centers have become so popular in the United States now it seems that there is one on every corner, not only in major cities but also in smaller cities. Starting a fitness center can be a great business opportunity, but you need to distinguish yourself from others to enjoy long-term success.

If you plan exclusive techniques and one-on-one personalized instructing, then it would suffice charging high prices, attracting sophisticated high end fitness clients, if you plan to teach a lot of people the same techniques then charging affordable prices to the mases would be adequate, with more high traffic area like residential area or nearby office space.

My client wants to open his business in Manhattan area, so I focus on that borough during my analysis. We define potential neighborhood based on the number of fitness centers which are operating right in each neighborhood. Manhattan has full potential but also is a very challenging district to open a business because of high competition. New fitness center should be open in an area that inadequate neighborhood in this way the center can attract more customers. Therefore, this analysis is necessary to ensure that we have enough customers and that we are not so close to other fitness places.

1. **Data**

**Data 1:** Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. This dataset exists for free on the web. Link to the dataset is: https://geo.nyu.edu/catalog/nyu\_2451\_34572

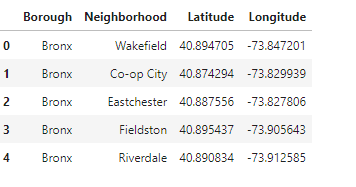


**Data2:** Network city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City. The below is image of the Foursquare API data.

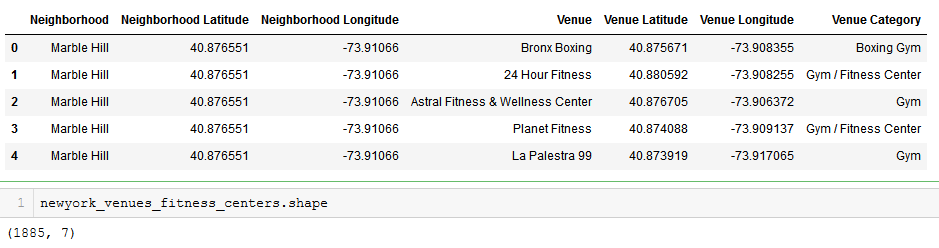
In addition, Fitness category Id 4bf58dd8d48988d175941735 is used for retrieving data from Foursquare API.

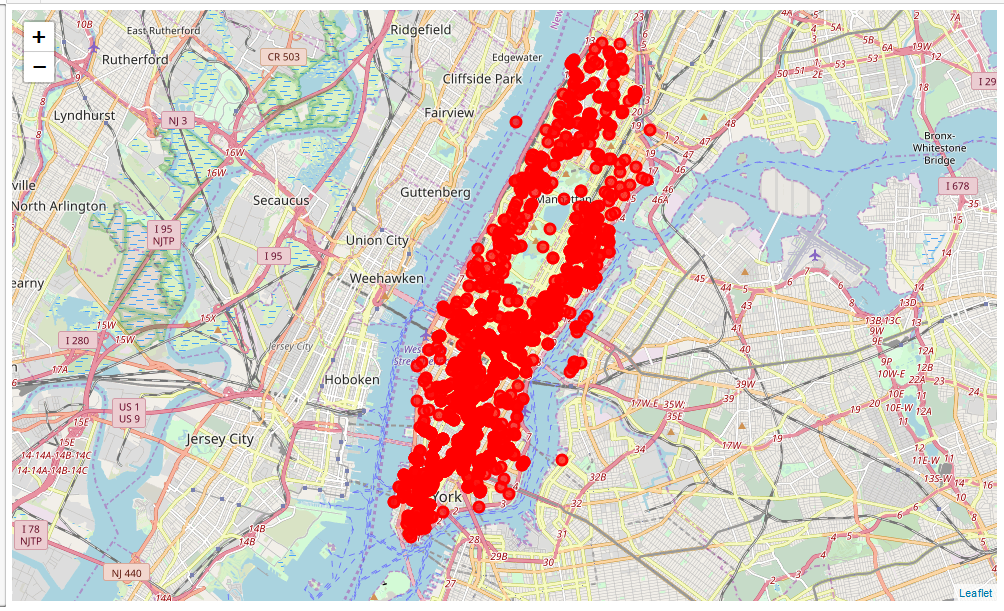
1. **Methodology**

In this project, I will use the basic methodology as taught in Week 3 lab.

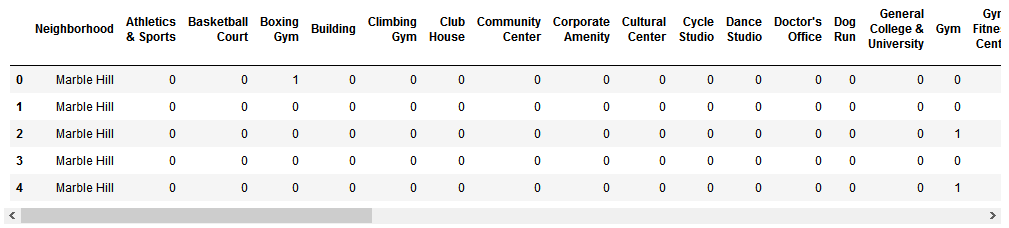


Above, I have done convert addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighborhoods in Manhattan, New York. After that, explore function to get fitness center categories in each neighborhood.

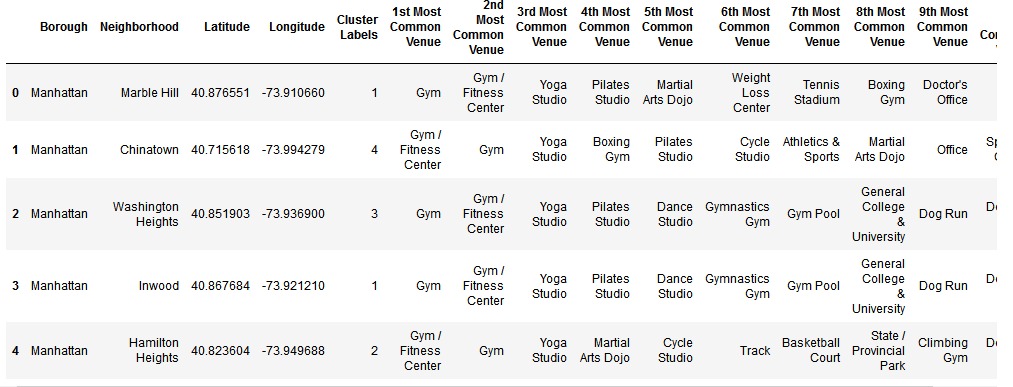


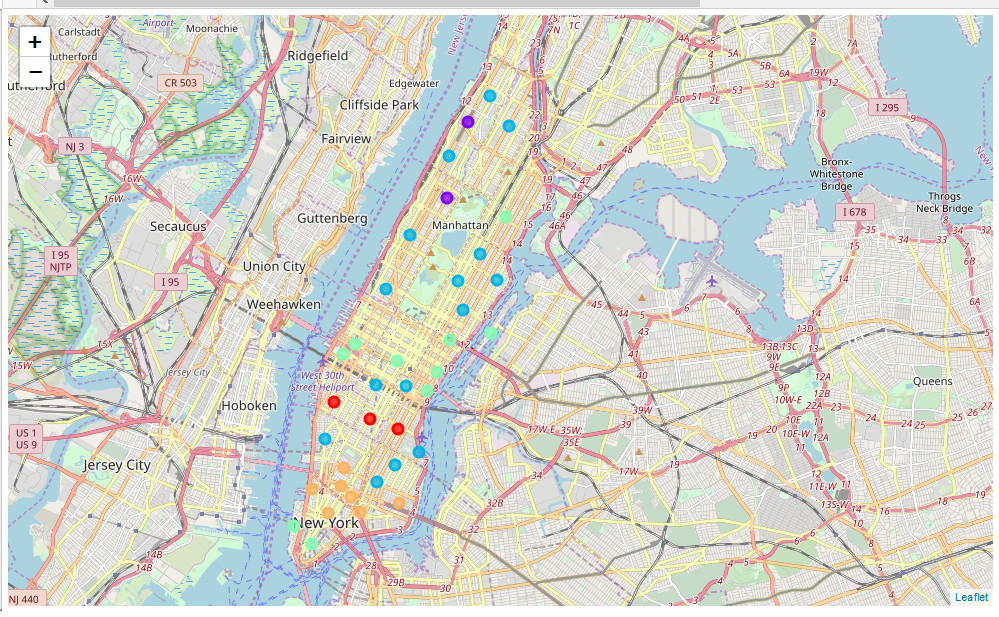


Fitness centers in Manhattan



Then use this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.





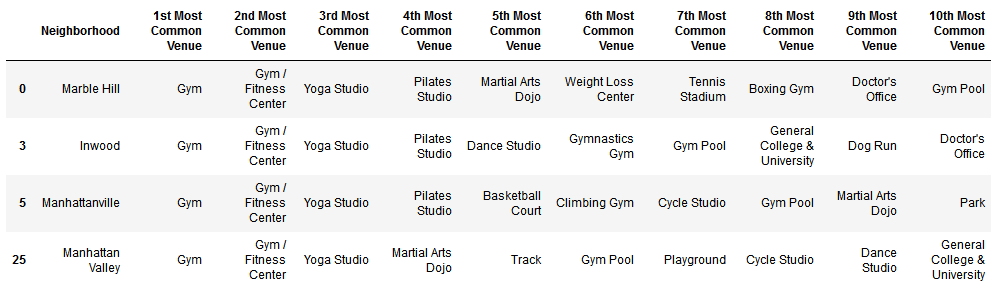
1. **Results**

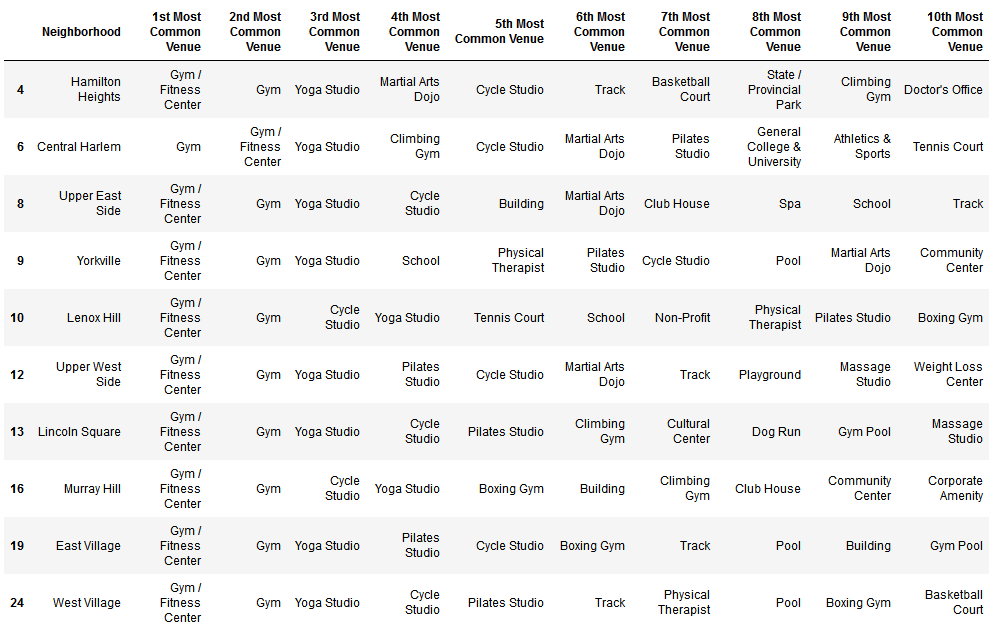
**K-mean Cluster** Using K-mean to clustering data area with less number of fitness centers

**Cluster 0**

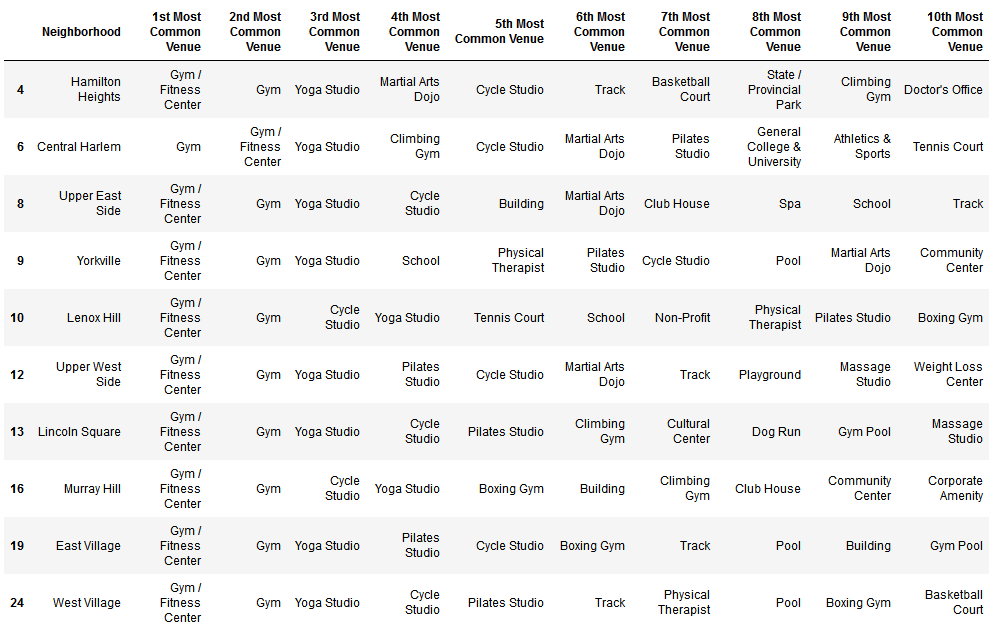


**Cluster 1**

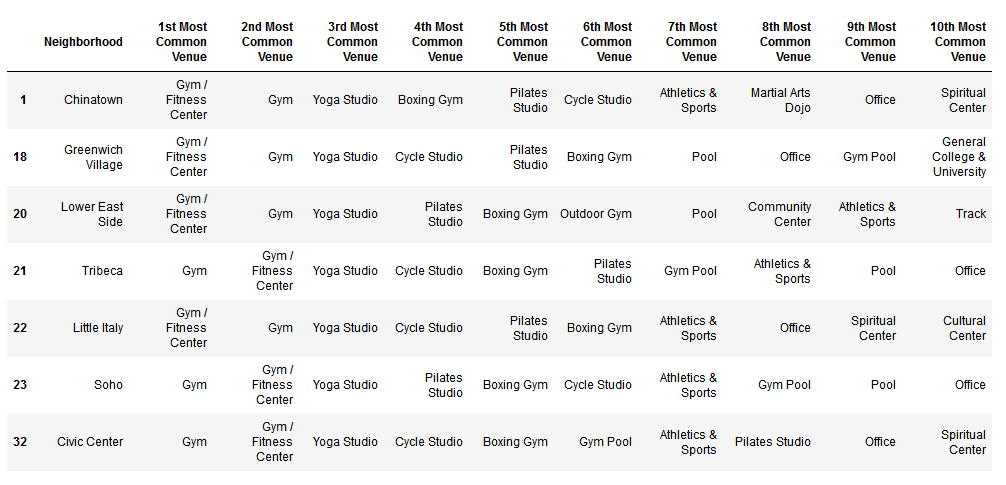


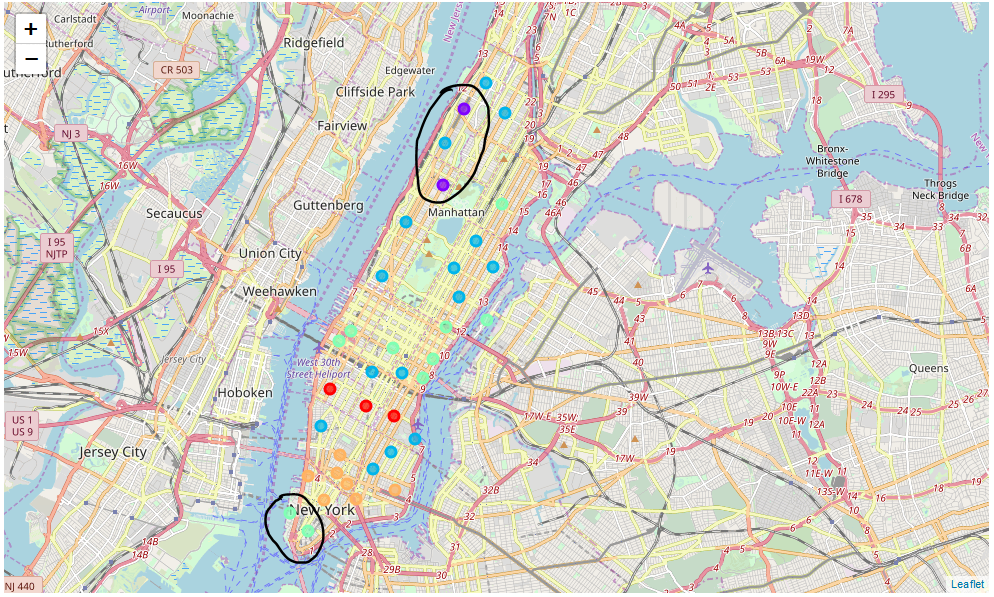
**Cluster 2**

**Cluster 3**



**Cluster 4**





Based on dataframe analysis above Cluster 0 (Chelsea) and Cluster 1 (Marble Hill) areas are the best places to open a new fitness center business.

1. **Discussion**

In this section, I would be discussing the observations I have noted and the recommendation that I can make based on the results.

This analysis is performed on limited data. This may be right or may be wrong. But if good amount of data is available there is scope to come up with better results.

* There is high competition in Midtown and Soho so it is very risky to open business in these areas.
* Chelsea has also potential where closes to Marble Hill area.
* It can be done more detailed analysis by adding other factors such as transportation, demographics of inhabitants.

Finally, Foursquare proved to be a good source of data but frustrating at times. Despite having a Developer account, I regularly exceeded my hourly limit locking me out for the day.

1. **Conclusion**

Although all of the goals of this project were met there is definitely room for further improvement and development as noted below. However, the goals of the project were met and, with some more work, could easily be developed into a fully pledged application that could support the opening a business idea in an unknown location.

As per the neighborhood or restaurant type mentioned like Fitness centers analysis can be checked. A venue with lowest risk and competition can be identified.