



FINDING A NIGHT CLUB LOCATION IN NEW YORK

In the time of COVID using Data Science Techniques

MOTIVATION

- Finding a location for a New Night club in New York City is the goal
- An optimal location would be one where other highly rated night clubs are located
- The assumption is club patrons will club hop switching from one club to another. (Even when there is a cover charge.)
- If a neighborhood is a hot spot for clubbing it will have numerous night clubs. Such a neighborhood is already attracting patrons so locating a club there would be beneficial.
- A new night spot which is in proximity will attract the curious. Walking distance and short taxi rides



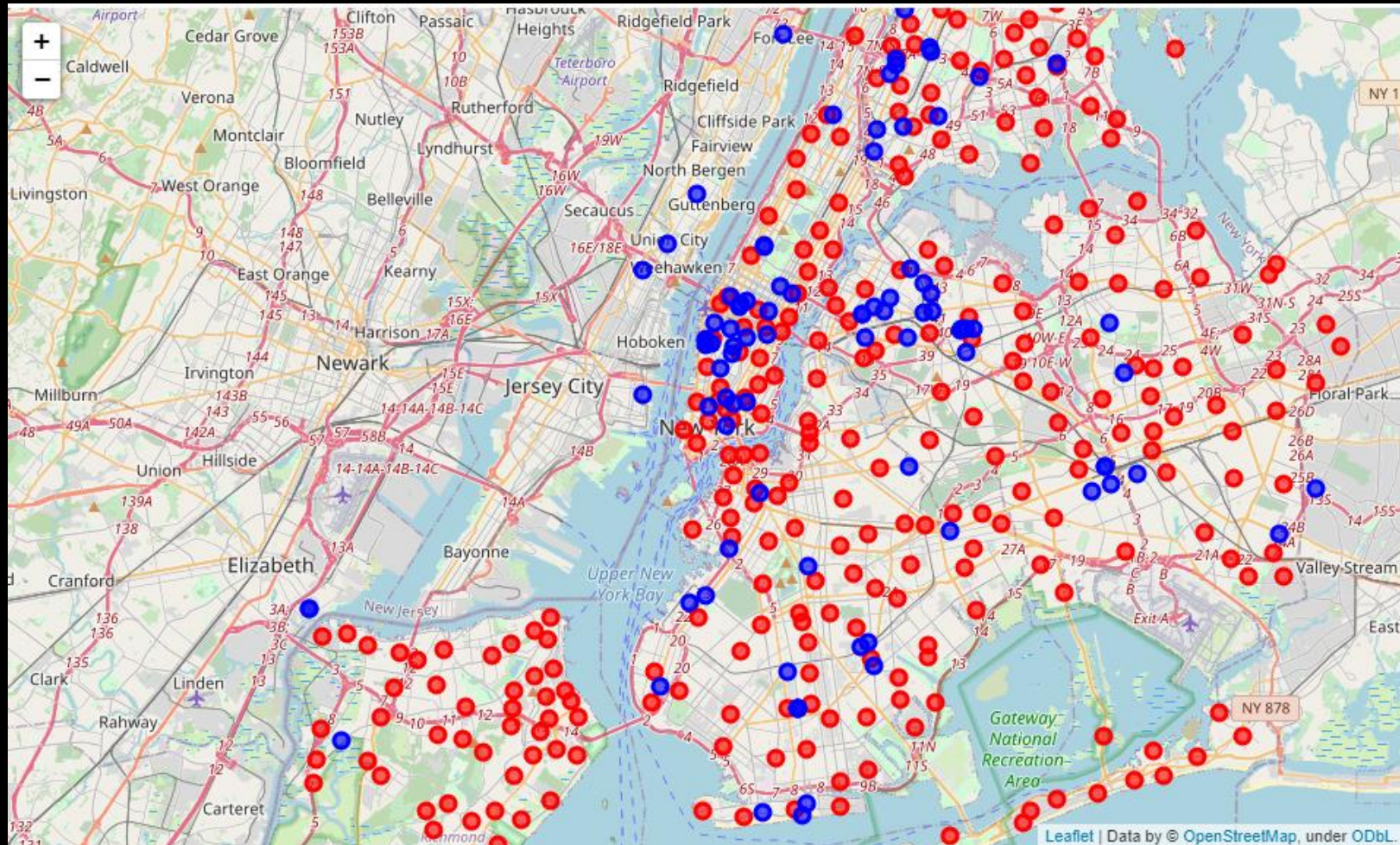
CONSIDERATIONS

- Many clubs are closed in New York City because of COVID-19
- Data is limited through the FourSquare API
- There is less traffic and fewer patrons
- However, it is even more critical to find a location with clubs still open that are still rated highly.

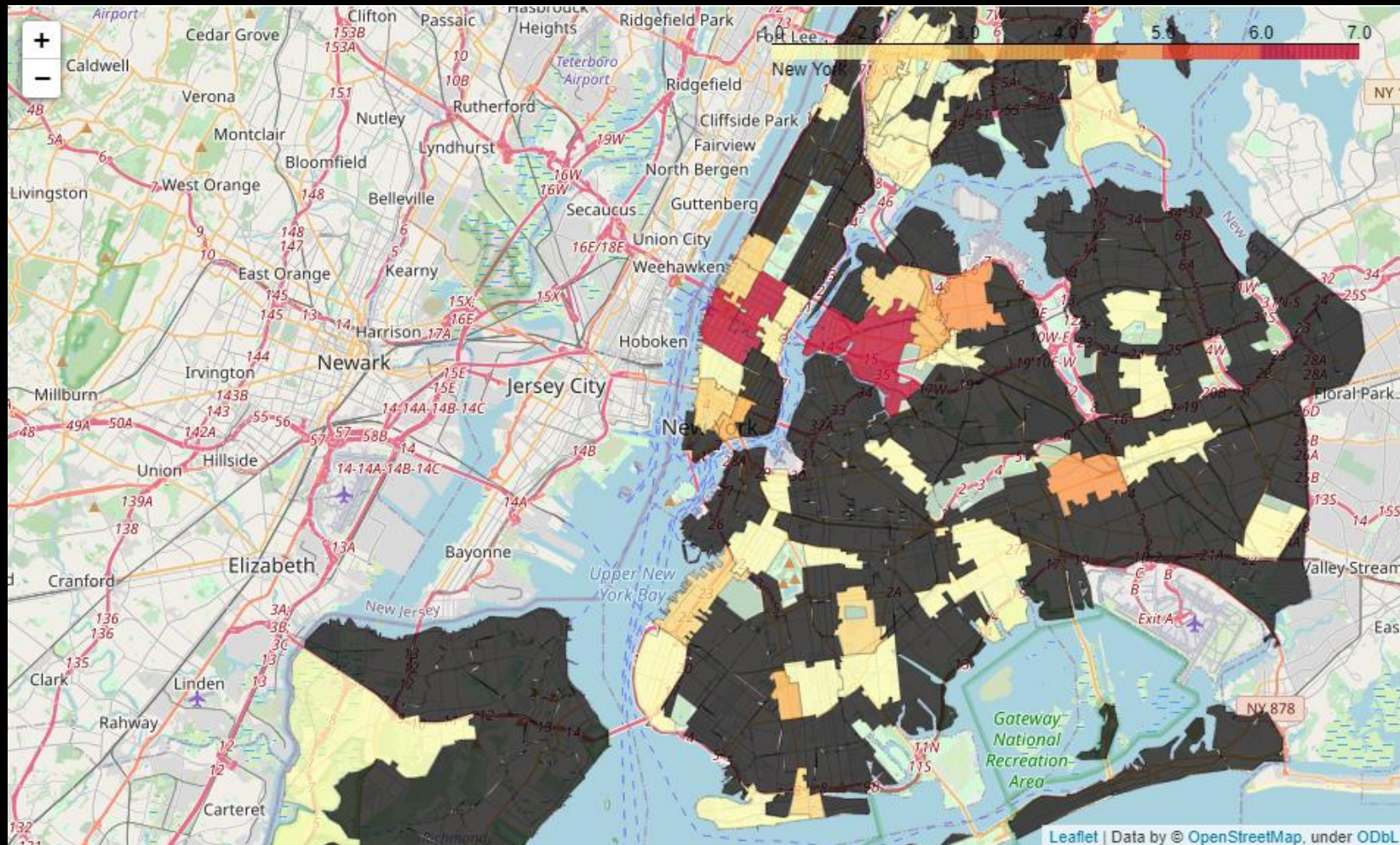
DATA ACQUISITION AND CLEANING

- New York neighborhood data from Coursera assignment:
 - https://geo.nyu.edu/catalog/nyu_2451_34572
 - 306 entries having latitude and longitude of geographical centers of neighborhoods
- New York boundary data from geo json file:
 - <https://github.com/veltman/snd3/blob/master/data/nyc-neighborhoods.geo.json>
 - 188 entries having the geographical boundaries of neighborhoods as polygons
 - Some neighborhoods were combined into single polygons
- Foursquare API used to get Night Clubs based on distance from geographical centers of neighborhoods

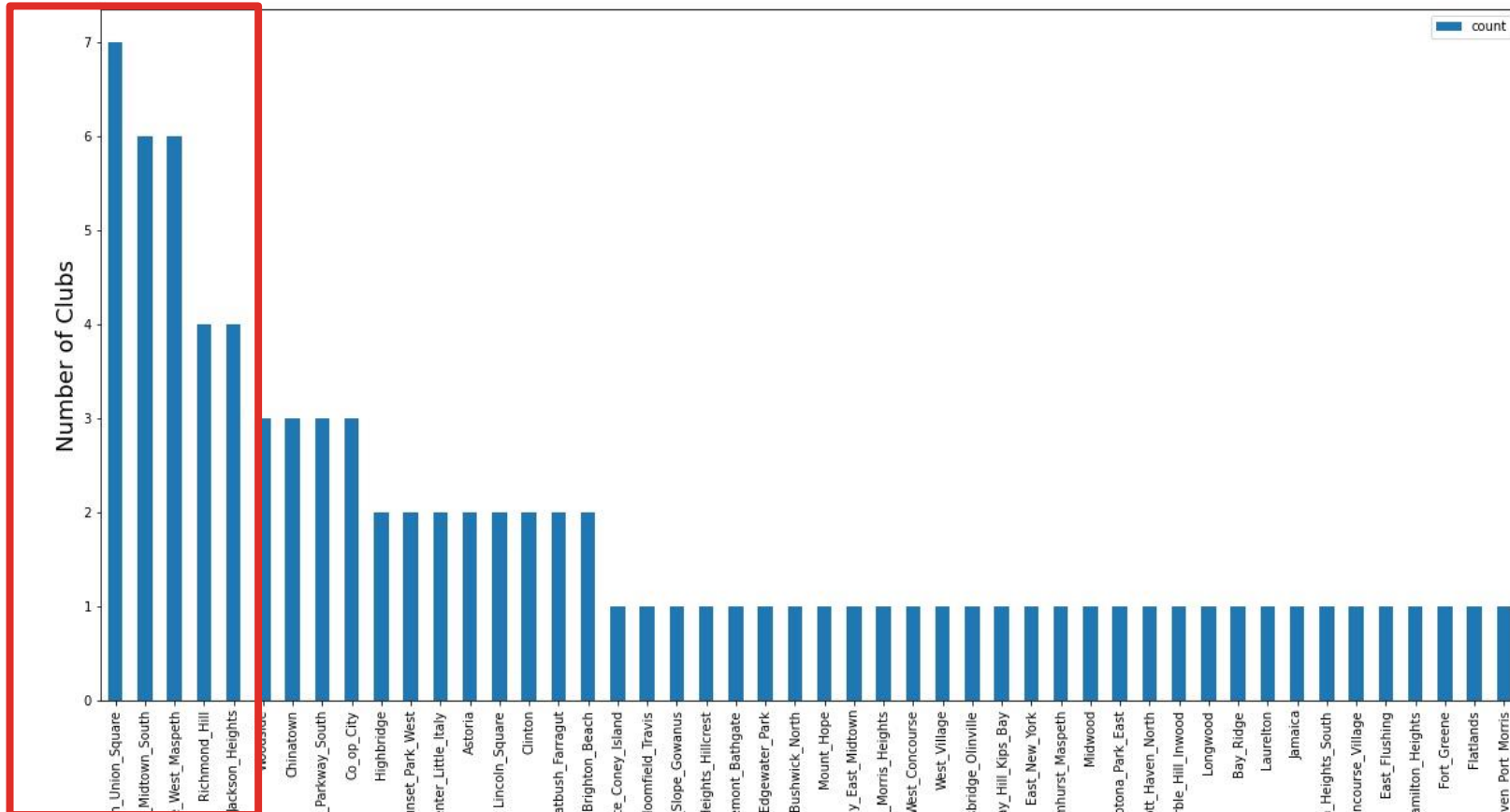
CLUB LOCATIONS AND NEIGHBORHOOD CENTERS



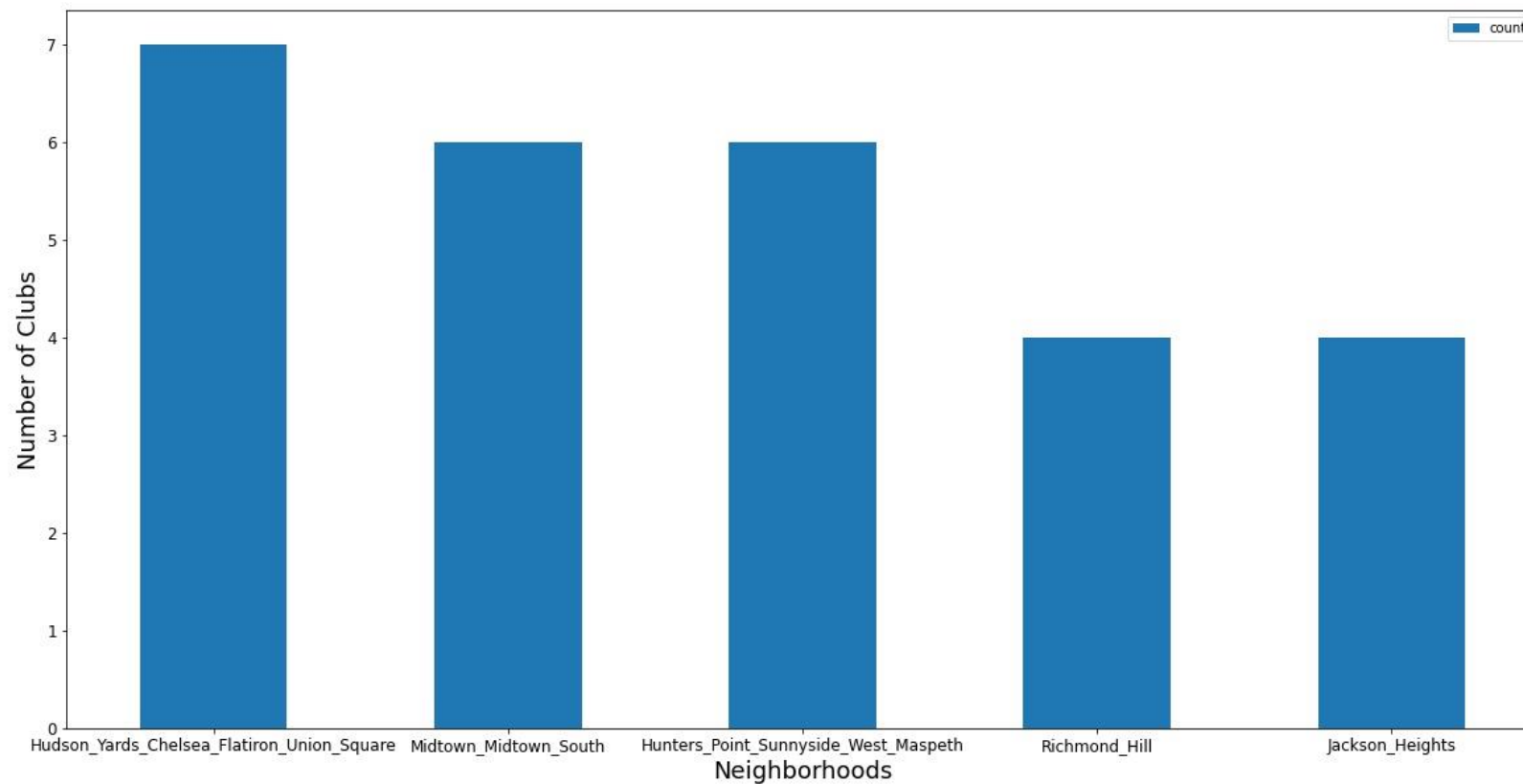
BINNING CLUBS GEOGRAPHICALLY



RANKING THE TOP NEIGHBORHOODS



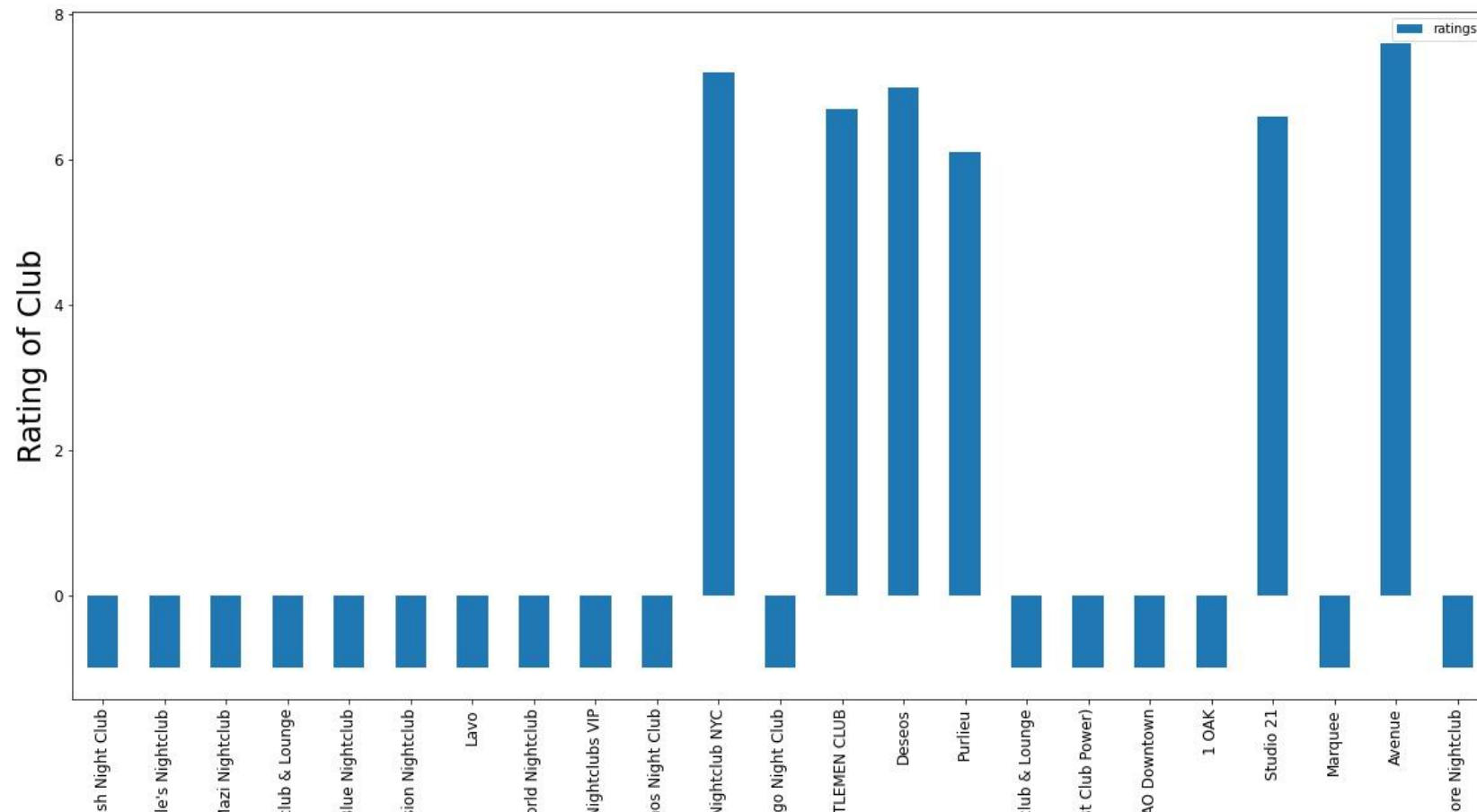
FIVE NEIGHBORHOODS WITH THE MOST CLUBS



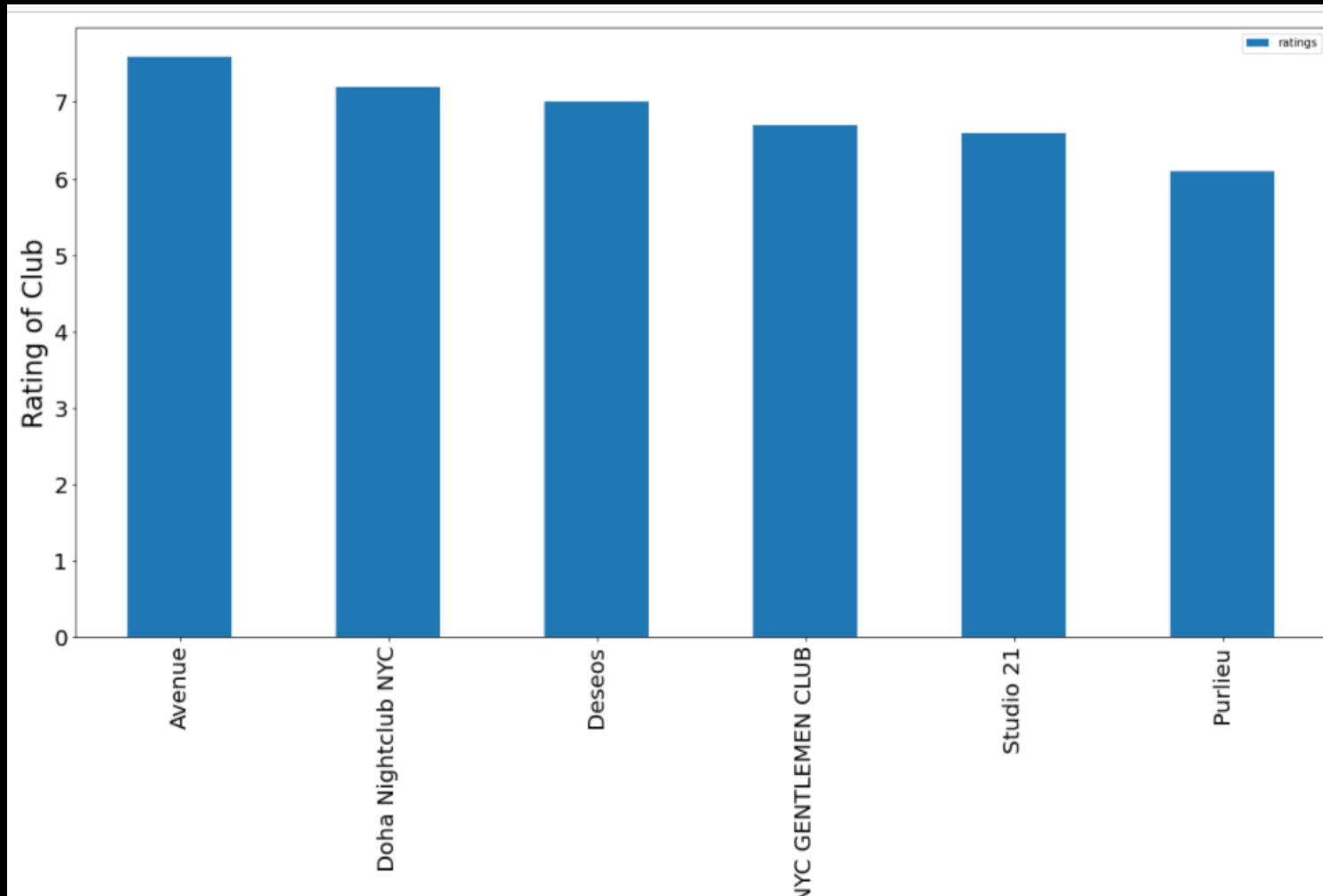
LIMITATIONS BECAUSE OF COVID-19

- Many clubs in New York at this time are closed
- Many clubs didn't have foursquare IDs to pull additional information
- Rating information was not available on all Night Clubs because of the closing
- Any club that did not have a rating was assigned a penalty of -1.
 - Any areas with many closed clubs would have a large penalty
- The assumption is it's better to open a club any club in an area with other open clubs instead of an area with no clubs.
- That is some night life is better than no night life

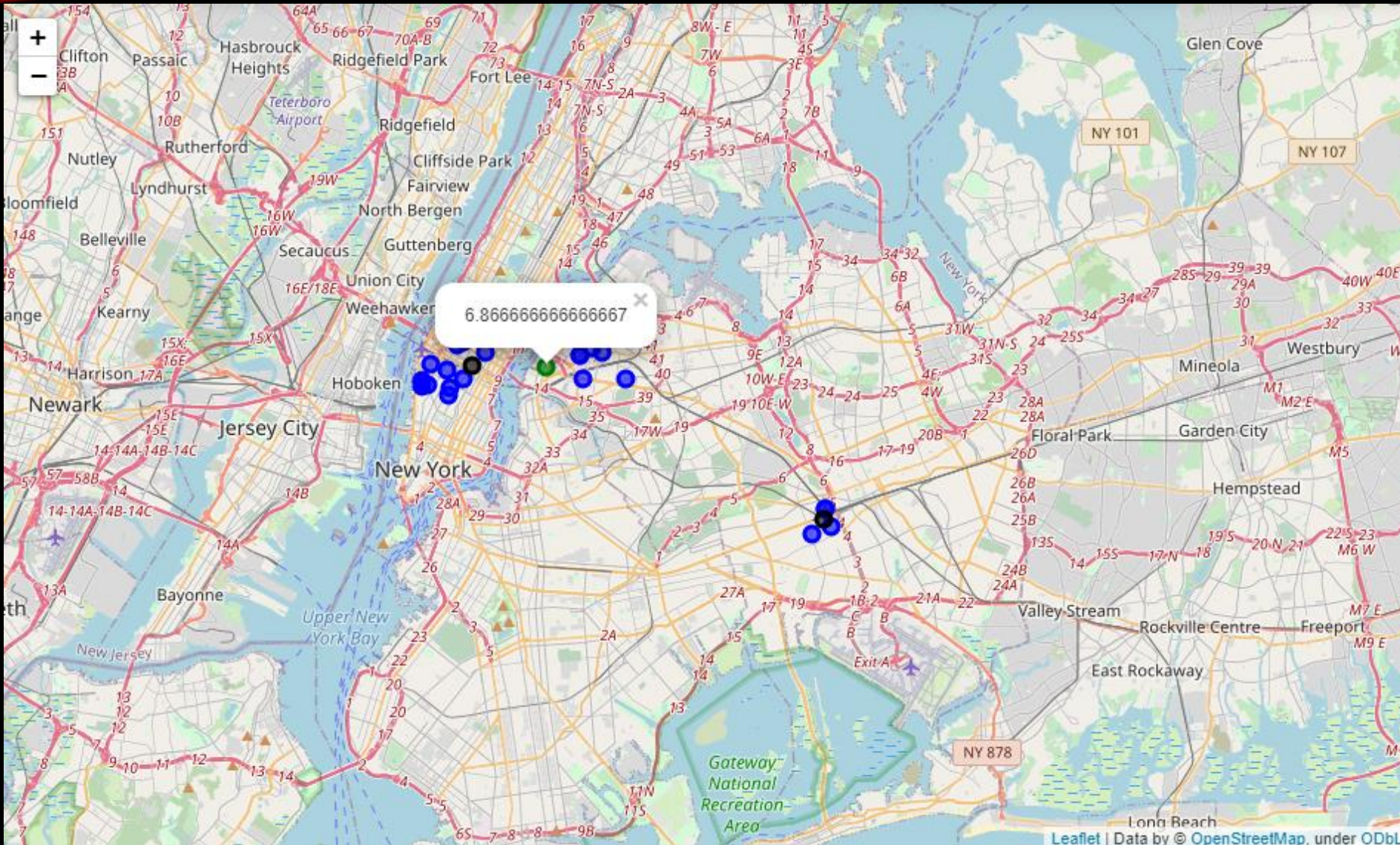
ADDING THE AVERAGE CLUB RATING



PLOTTING THE ONLY CLUBS THAT HAVE A RATING



CLUSTERING BY LOCATION AND RATING



- K-means Clustering
- 3 centroids

PICKING THE BEST NEIGHBORHOOD

- Using K-means clustering
- Three distinct clusters were used
- The average rating produced for each cluster was compared
- Only 1 cluster had a positive rating
- It also contained the most clubs with positive ratings
- The neighborhood seemed more resilient to a downturn in the economy due to covid. More open clubs mean that location is fairing better.

CONCLUSION AND FUTURE DIRECTIONS

- The centroid of the cluster with the highest average rating is located within the bounds of 'Hunters_Point_Sunnyside_West_Maspeth' boundary
- The closest Neighborhood center to the cluster centroid is "Hunters Point"
- These methods are in agreement so the neighborhood with clubs that have the highest average rating is "Hunters Point"
- The analysis could be run again after the covid crisis ends. Many venues are currently closed, but this neighborhood seems to have the highest resiliency with clubs still open.
- This location may lead the recovery of the night club scene.