



CAPSTONE PROJECT

Eduardo

Introduction

- In this project we will try to find an optimal location for a Gym. Specifically, this report will be targeted to stakeholders interested in opening an business in New York City, USA.
- Since there are lots of Gym's in New York we will try to detect locations that are not already crowded with business. We are also particularly interested in areas with no business in vicinity. We would also prefer locations as close to city center as possible, assuming that first two conditions are met.

Data

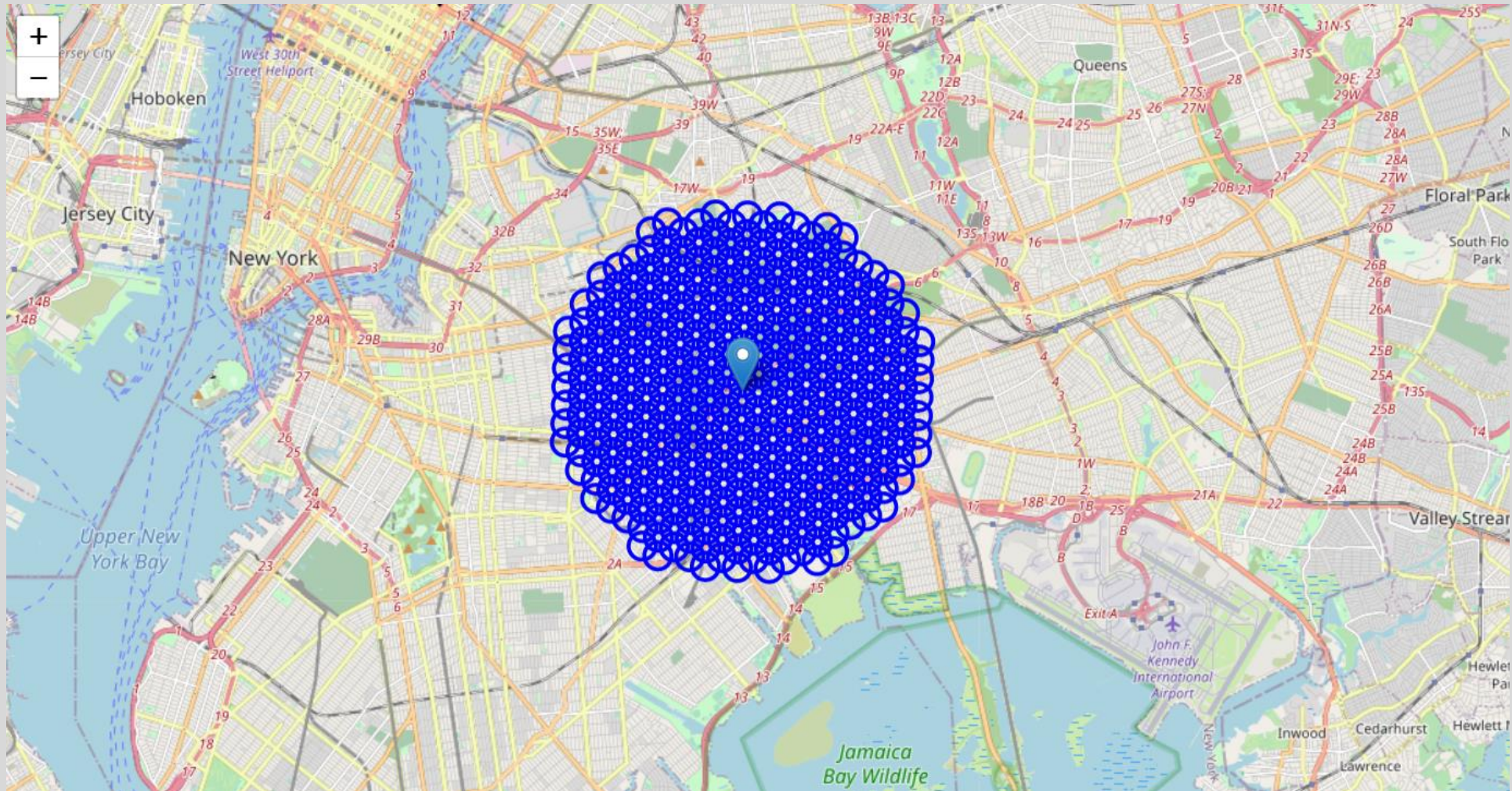
Based on definition of our problem, factors that will influence our decision are:

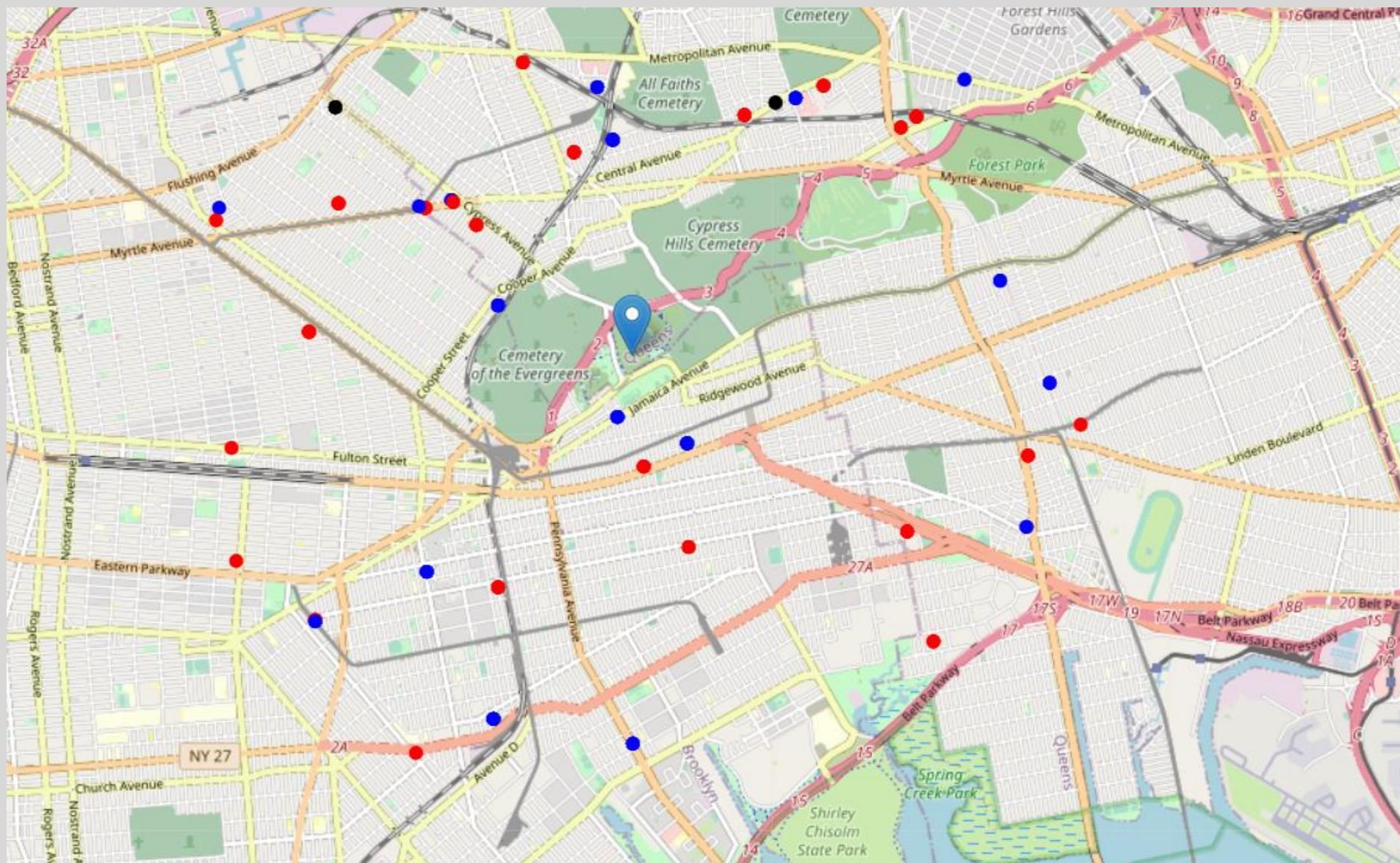
- number of existing business in the neighborhood (any type of business)
- number of and distance to business in the neighborhood, if any
- distance of neighborhood from city center

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

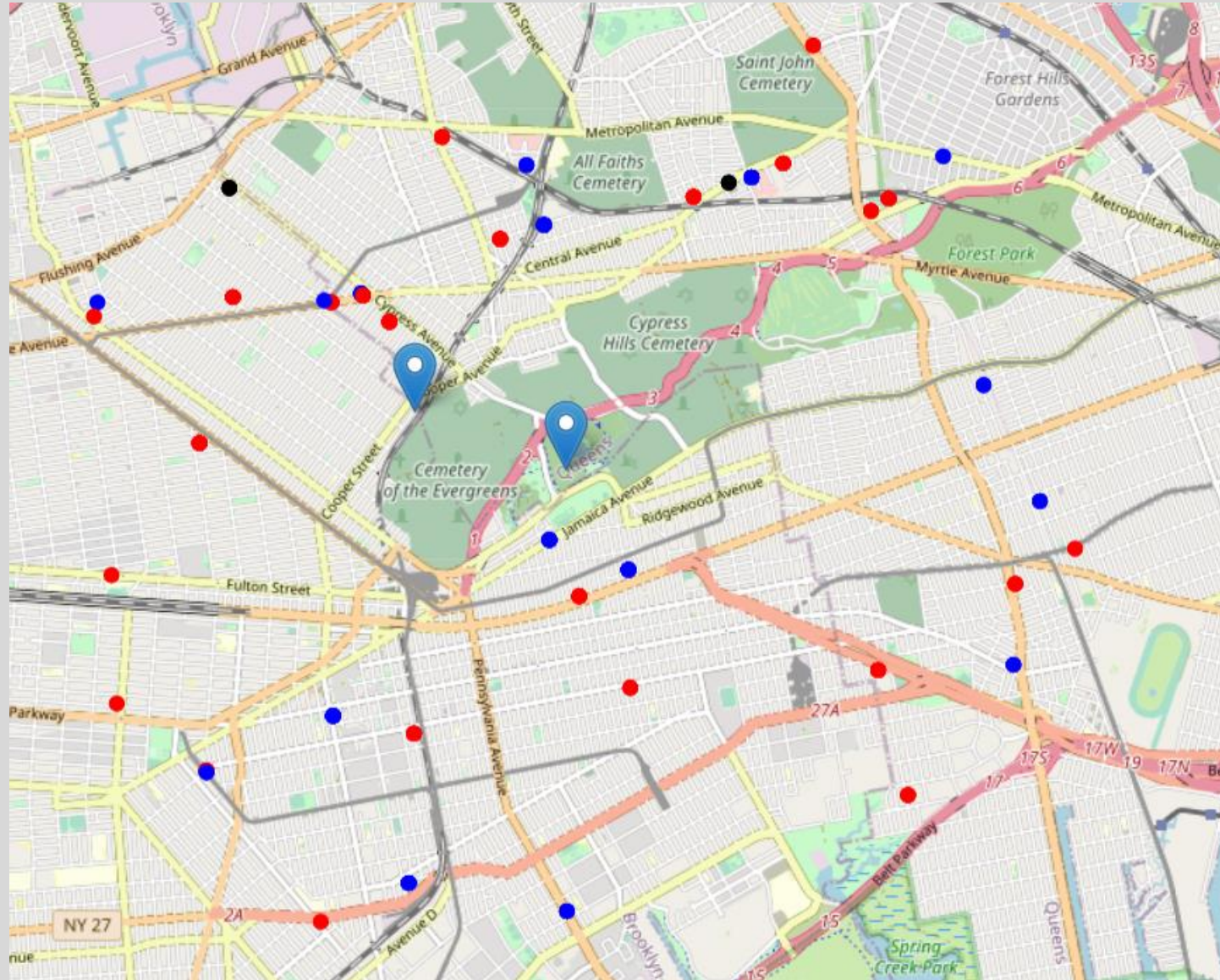
Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be search and chose and approximate addresses of centers of those areas will be obtained using Foursquare API
- number of Gym's and their type and location in every neighborhood will be obtained using Foursquare API
- coordinate of New York center will be obtained using Foursquare API of well known New York location





01/08/2021



Results and Discussion

Our analysis shows that although there is a great number of gym in New York center (~2000 in our initial area of interest which was 12x12km around Highland Park), there are pockets of low gym density fairly close to city center. Highest concentration of gym was detected north and south from Highland Park, so we focused our attention to areas east and west, corresponding to boroughs Ocean Hill and Knews Gardens. Another borough was identified as potentially interesting (Richmond Hill and Woodhaven, west from Highland Park), but our attention was focused on Ocean Hill and Knews Gardens which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low gym density.

After directing our attention to this more narrow area of interest (covering approx. 5x5km south-east from Highland Park) we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two business in radius of 250m and those with an Gym closer than 400m were removed. Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Conclusion

Purpose of this project was to identify New York city areas close to center with low number of business (particularly Gym) in order to aid stakeholders in narrowing down the search for optimal location for a new Gym. By calculating gym density distribution from Foursquare data we have first identified general boroughs that justify further analysis (Ocean Hill and Knews Gardens), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby business. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal Gym location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.