

GreenTech Project

Introduction

The world is currently experiencing one of the biggest crises in history that was the result of just a nano-meter assembly of nucleic acid and protein. The SARS-COV2 virus (aka COVID-19) reminds us of the fragility of humanity - its economy, the ecosystem and how important it is to take care of our surrounding environment. The COVID-19 crisis has completely thrown the economy into an almost unprecedented chaos. But in these dark times there is still a lot of hope. Decisions made during the “days after” will be extremely important for the future of our economy, planet/environment and mankind. Each of us have the choice to either think about how we can create a better world with our skills and knowledge, or if we will just fall back on our old ways. The first hope is that people will change their way of thinking, and work in new ways while making environmentally friendly decisions. The second hope is that the crash of the economy creates a unique opportunity for young entrepreneurs to come out from under the shadow of big companies. Perhaps the aftermath will force people and businesses to rely more on skills and ideas rather than on money. In this context I decided to act at my level to contribute to the global effort by using high technology for the environment. This project outlines the opening of an electronic store in Manhattan dedicated to the environment.

The business will need to be close to a university in order to attract students, scientists, or people who are passionate about creating their own projects for the environment. The neighborhood should have the following:

- A low crime rate
- A high population density
- A low average rent price

Once the best neighborhood for the project is found, the study will focus on this neighborhood and identify the exact, best location to open the store. The store must be located in a commercial zone within a 100m radius of a science & technology laboratory or teaching building.

Data

The crime rate, population density and rent price data were taken from the city of New York's official data provider website NYC Open Data. These data are combined with the neighborhood geometry found on the same website in order to make choropleth maps. Then, university venues were found using the Foursquare API. A clustering by using k-means were performed in order to see how university venues are grouped.

Manhattan, New York City, NY, USA

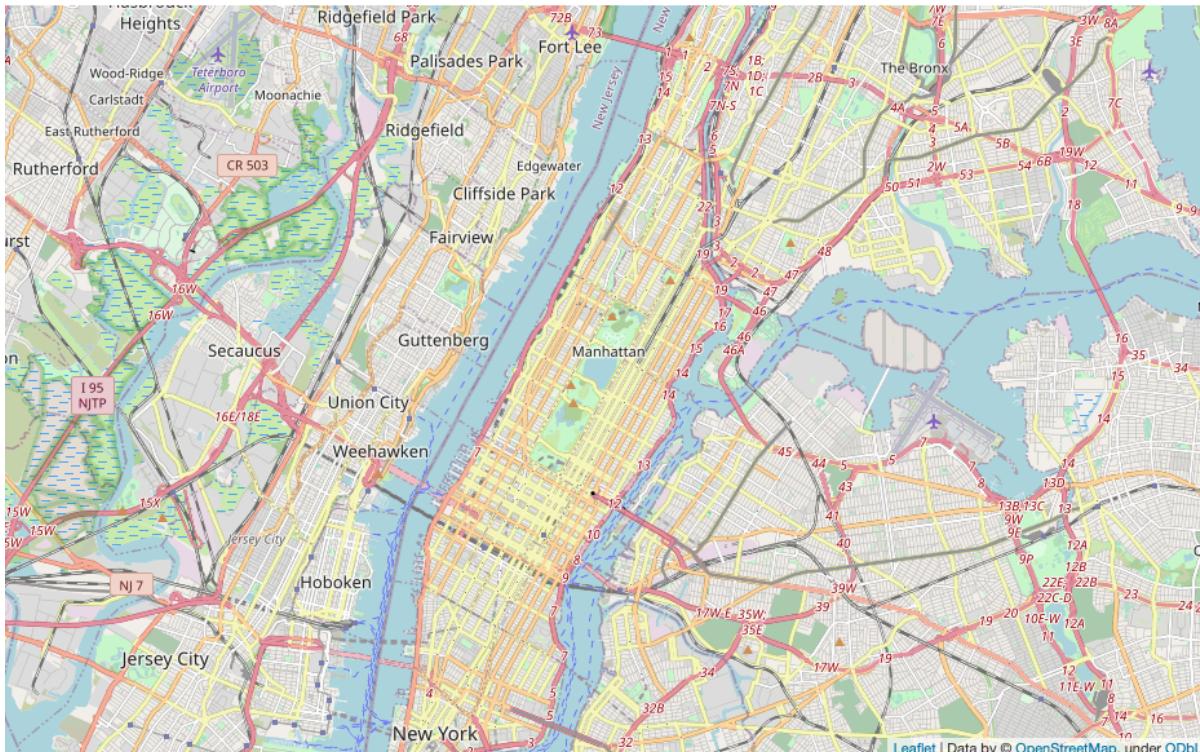


Figure1: Map of Manhattan.

1. First factor: Crime rate

In order to visualize the crime rate, an API (published by NYC open data on [Socrata](#)) was called. Because of the data size (all crimes in NYC since 2000) it was necessary to call the API with specific parameters in order to reduce the size downloaded.

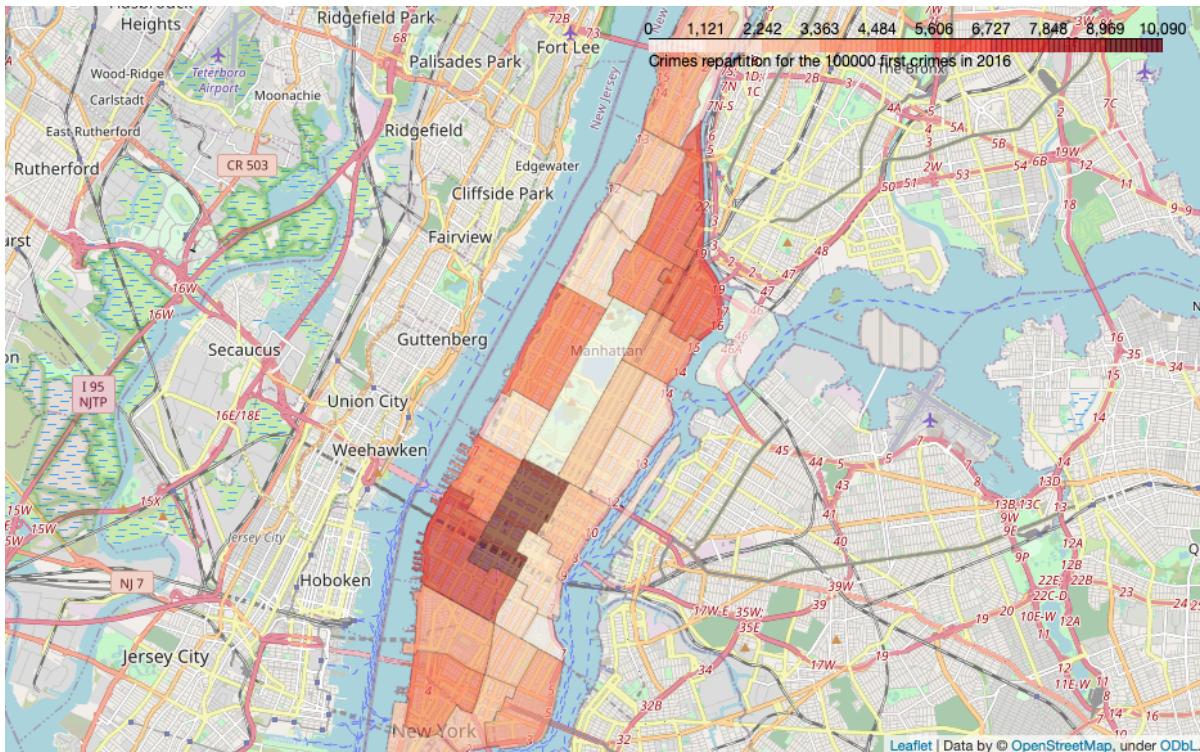


Figure 2: Crime rate in each neighborhood. The data were the first 100,000 crimes between 2016 and 2017.

2. Second factor: Population density

The population density by NTA was also downloaded on Open Data NYC website. Data comes from the 2000 and 2010 censuses for all boroughs. In order to be the closest possible to today's population density, the data was taken for 2010 and only in Manhattan.

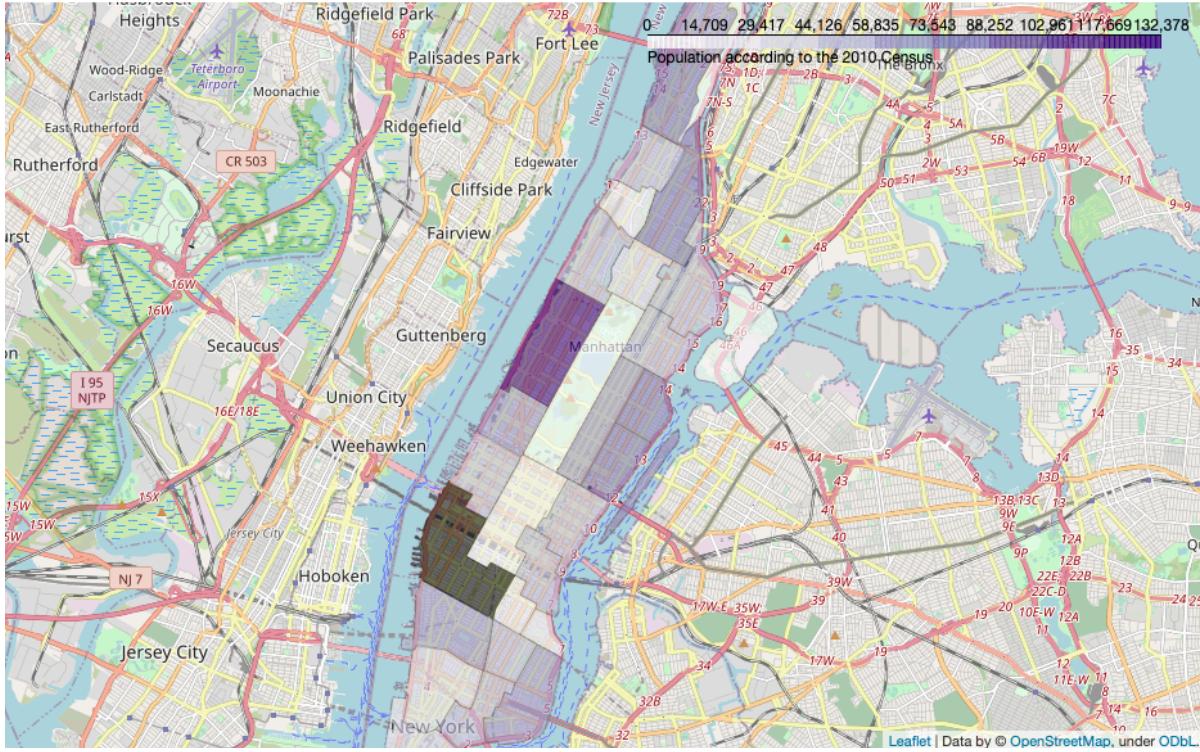


Figure 3: Population number in each neighborhood according to the 2010 census.

3. Third factor: the proximity to a university

The business will mostly interest students so it's important to be close to a university and/or near all venues related to universities. To do so, a Foursquare request was performed in order to see all venues related to universities in Manhattan. Then all the venues were put in a DataFrame.

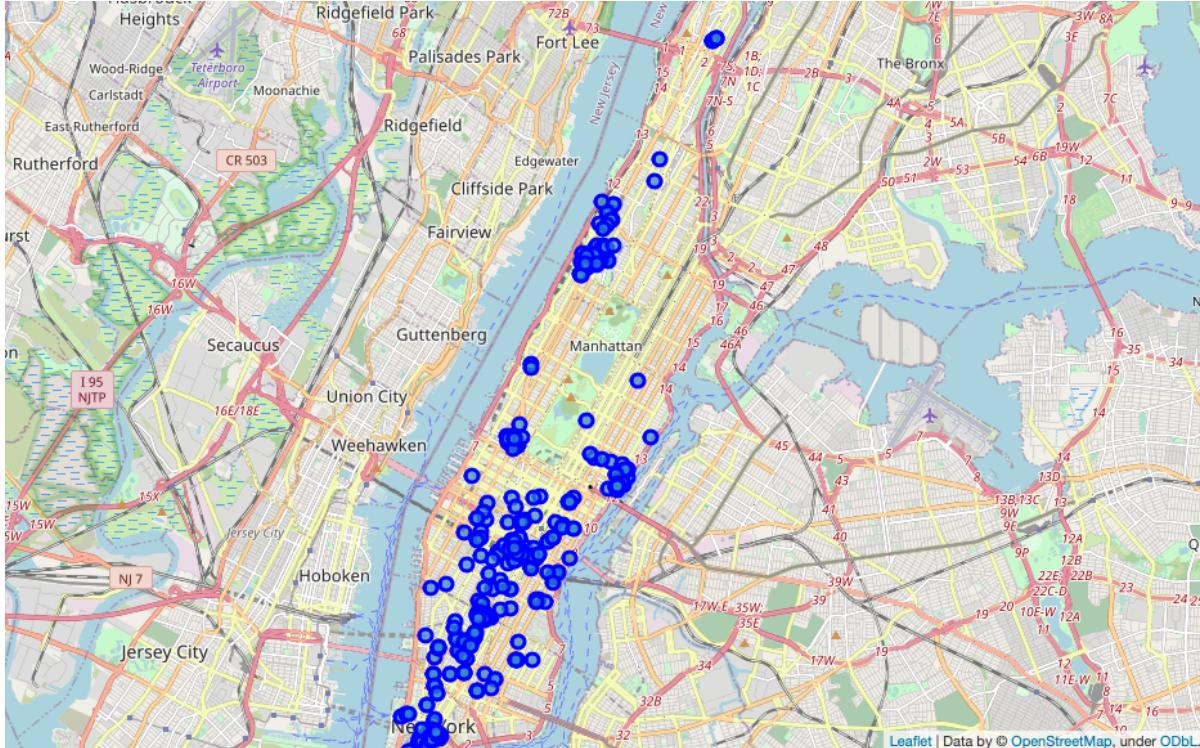


Figure 4: Visualization of all venues related to universities in Manhattan.

Clustering

Let's see what is the spatial distribution and where is the highest concentration of venues related to university. To do so, a k-mean clustering was performed.

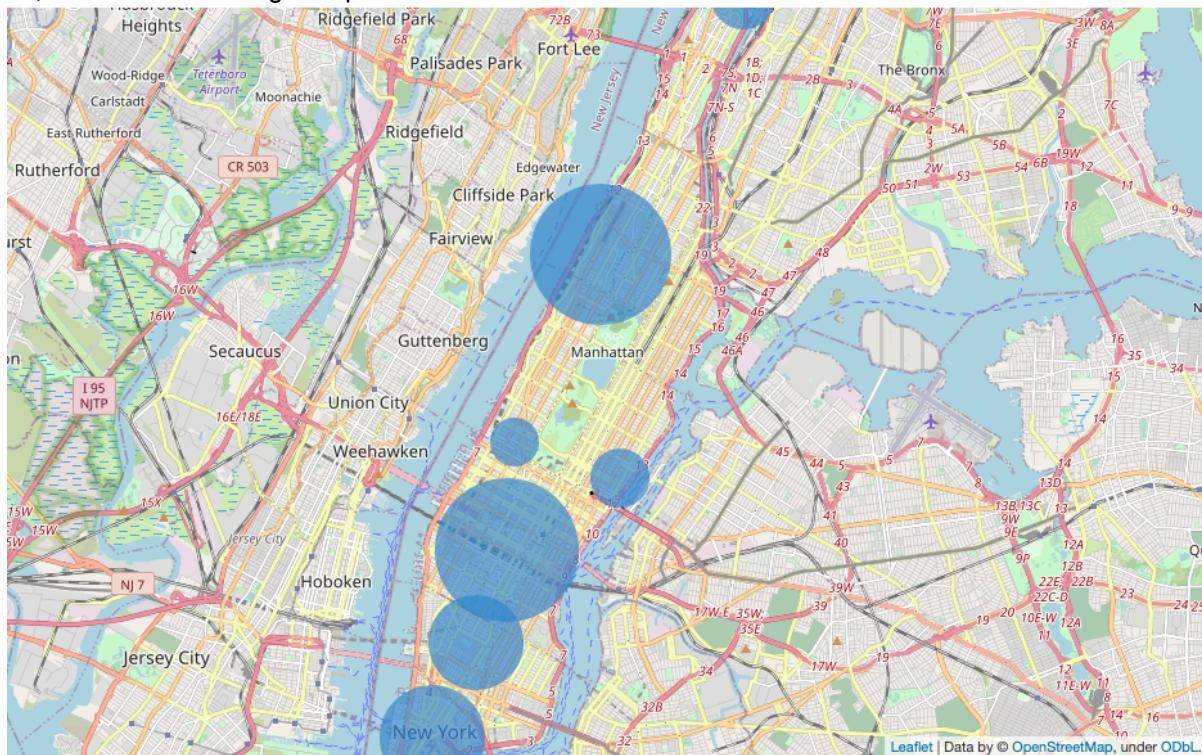


Figure 5: Spatial distribution of venues related to universities after performing a k-mean clustering

Neighborhoods selection

We can observe clusters are located in 7 different neighborhoods:

- Marble Hill-Inwood
- Morningside Heights
- Lincoln Square
- Midtown-Midtown South
- Lenox Hill-Roosevelt Island
- West Village
- Battery Park City-Lower Manhattan

So let's take a closer look at these neighborhoods

4. Fourth factor: the gross rent price in the neighborhoods

In order to reduce costs before the business starts to take-off, it's beneficial to select a location where the rent price is not too high. From the CSV file downloaded we can see that the 2 most expensive neighborhoods are:

- Battery Park City-Lower Manhattan
- Midtown-Midtown South

And the 2 most affordable neighborhoods are:

- Marble Hill-Inwood
- Morningside Heights

Table1: Median of the gross rent price in each neighborhood of interest

Neighborhood	Price (\$)
Marble Hill-Inwood	1279
Morningside Heights	1575
Lincoln Square	2391
Midtown-Midtown South	2951
Lenox Hill-Roosevelt Island	2289
Battery Park City-Lower Manhattan	3228
West Village	2337

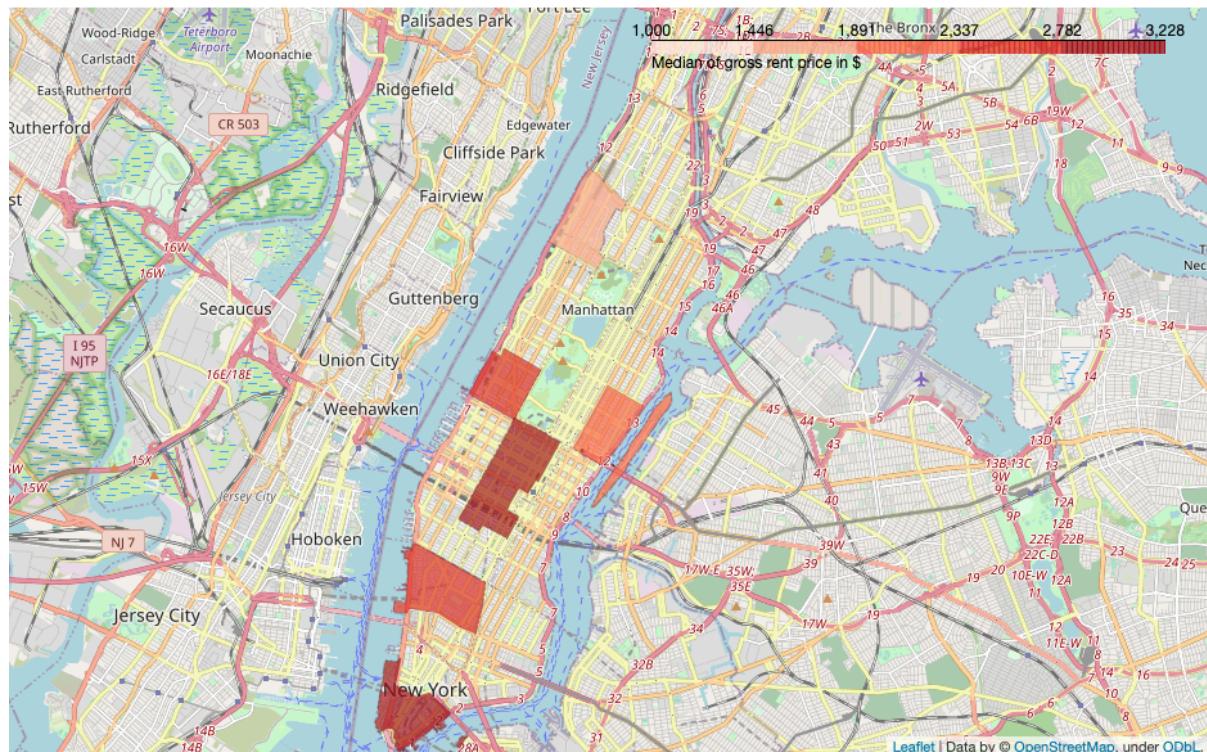


Figure 6: Visualization of the median gross rent price spatial distribution for the neighborhood of interest.

5. Attractivity

Neighborhoods have to be compared against one another in order to see which one is the best according to the 4 factors (criminality rate, population density, proximity to a university and rent price). From the 4 factors, an "attractivity score" was calculated which is:

$$A = \frac{PD}{CR + RP} + NV$$

Where PD is the population density, CR the crime rate, RP the rent price and NV the number of venues related to universities (in the cluster).

Then, the values were normalized in order to compare the 7 neighborhoods.

$$A_{Normalized} = \frac{A_{Max} - A}{A_{Max} - A_{Min}}$$

Where A is the attractivity for the current neighborhood, A_{Max} the maximum value of the attractivity score of all neighborhoods and A_{Min} the minimum value. The most attractive neighborhood will have an $A_{Normalized}$ Index = 1 and the least attractive one will have a $A_{Normalized}$ Index = 0.

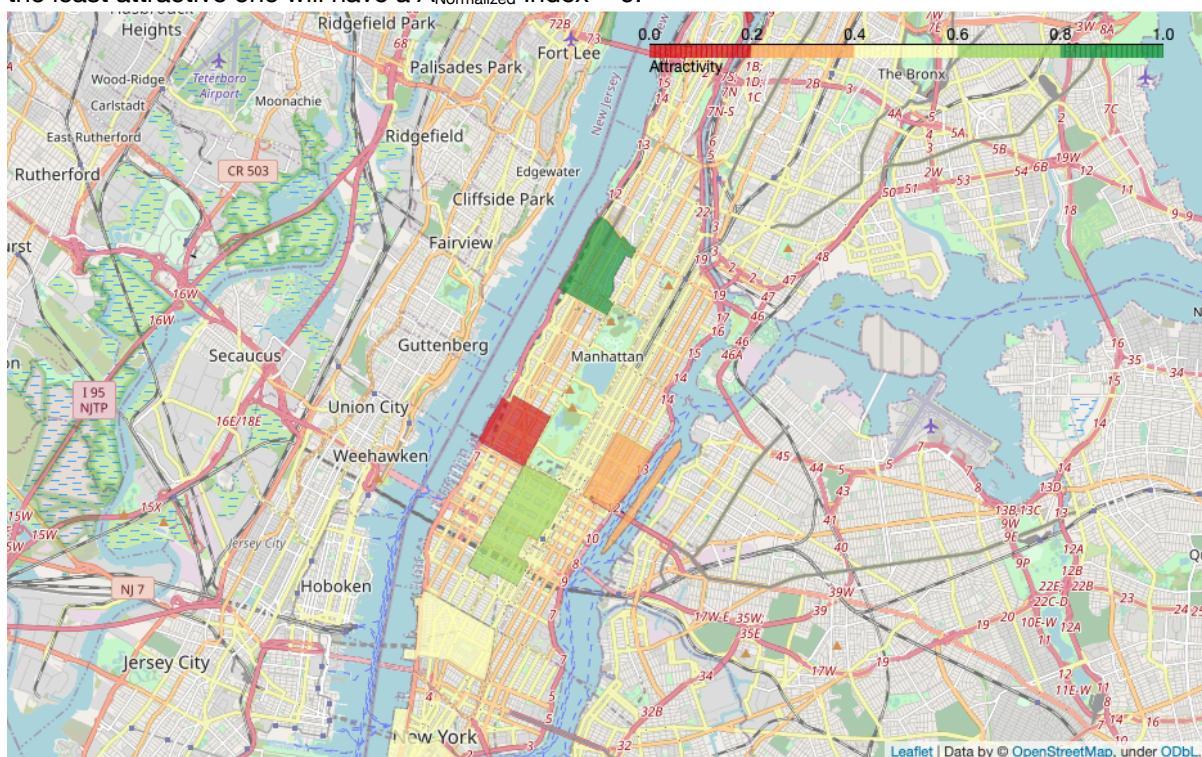


Figure 7: Visualization on map of the $A_{Normalized}$ index which a quantification of how a neighborhood is attractive for the project according to the study of the four factors. In green the most attractive neighborhood, in red the less attractive neighborhood.

6. Location in Morningside Heights

The combination of these different factors makes Morningside Heights the best place to start the GreenTech business project. But the question now is: "Where exactly in Morningside Heights?" Commercial/business places can't open everywhere because the city of New York delimits specific zones according to the activity (ie: household, industry, commerce, offices). So let's explore this neighborhood where it's possible to rent an office/commercial space.

Data acquisition and reorganization

On the NYC Open Data website there are the official commercial/office zones. So the data were downloaded and reorganized in order to see these zones only in the neighborhood of interest: Morningside Heights. To do so, each commercial/office space coordinates were checked to determine if they are within the neighborhood area. Then a new geojson file was created.

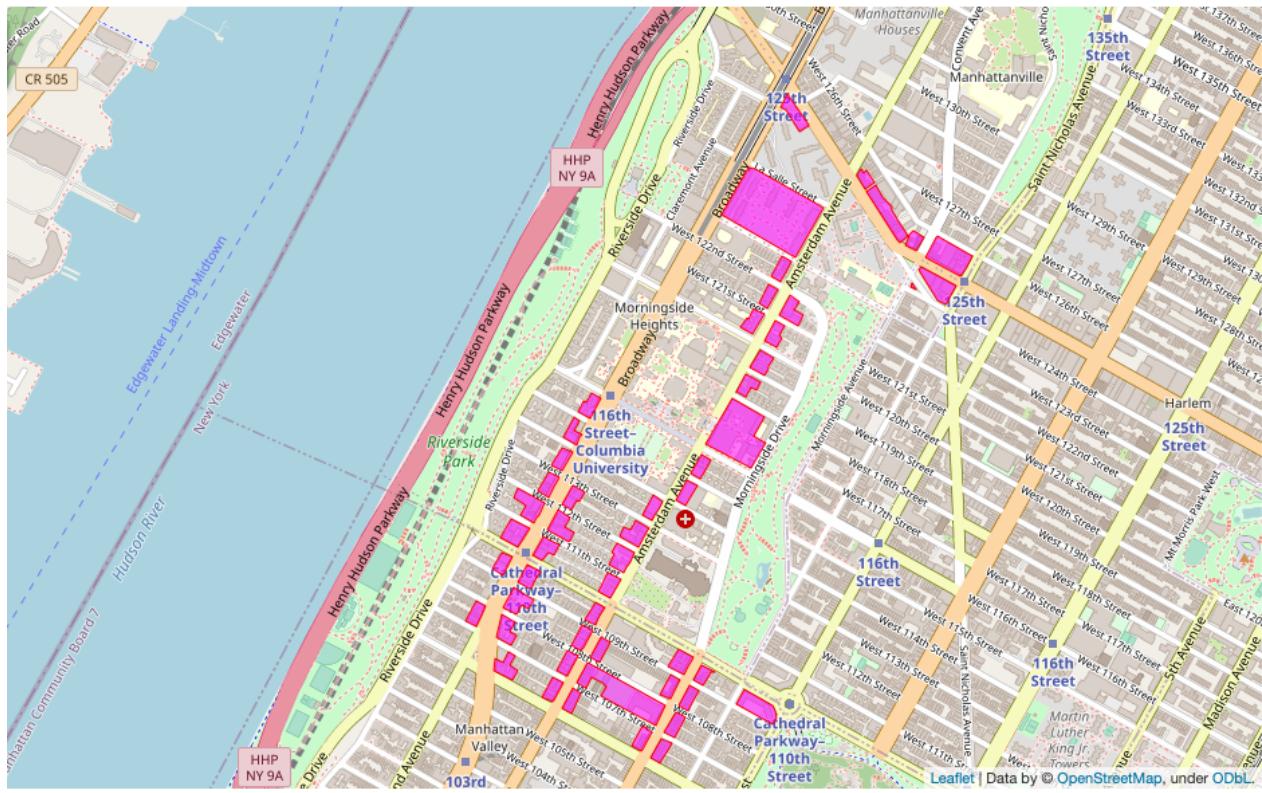


Figure 8: Visualization of the reserved commercial zones where is possible to open a store (in purple) which was determined by the city of New York.

Looking for the best commercial zone

Now that all commercial zones are defined, it's important to define which ones are the best for the GreenTech project. One of the first purposes is to attract people who are interested in technology, electronics etc. So the best spaces would be close to electronic/computer/engineering/technology/environment/science buildings. To find those kinds of venues, a Foursquare query was performed. Then, if a commercial zone is within a radius of 100m from a venue it becomes a candidate zone. The addresses of each candidate were found thanks to a reverse geocoder.

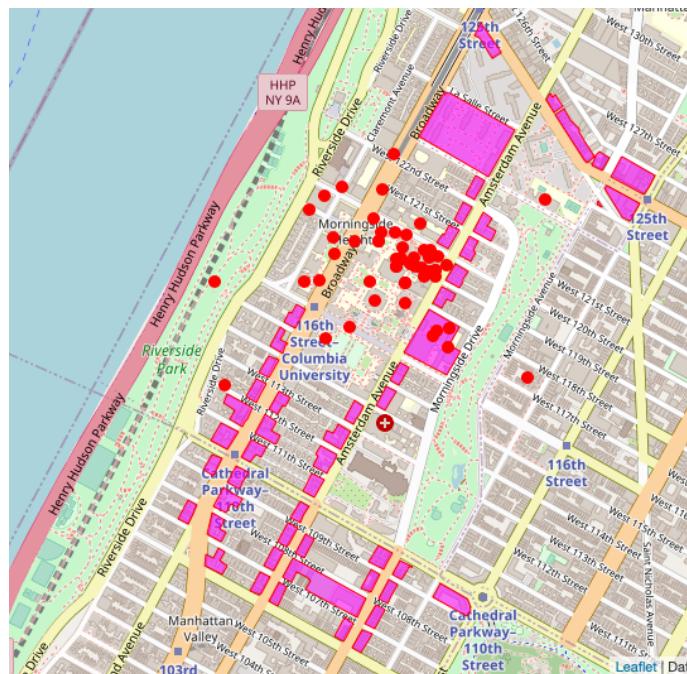


Figure 9: Visualization of the reserved commercial zones (in purple) and venues related to science, technology, environment, electronic, computer, engineering Lab or teaching building (red dots)

Study of the venues' distribution and concentration

A heat map is useful to see the spatial distribution of the venues and to see, among the potential candidates, which one is the closest to the 'cluster'



Figure 10: Visualization of the reserved commercial zones (in purple), venues related to science, technology, environment, electronic, computer, engineering Lab or teaching building (red dots), and their spatial distribution/concentration (color gradient from blue (the lowest density), to red (the highest density))

Candidates identification.

The best places are marked in blue which are the commercial zones within a 100m radius from a venue.

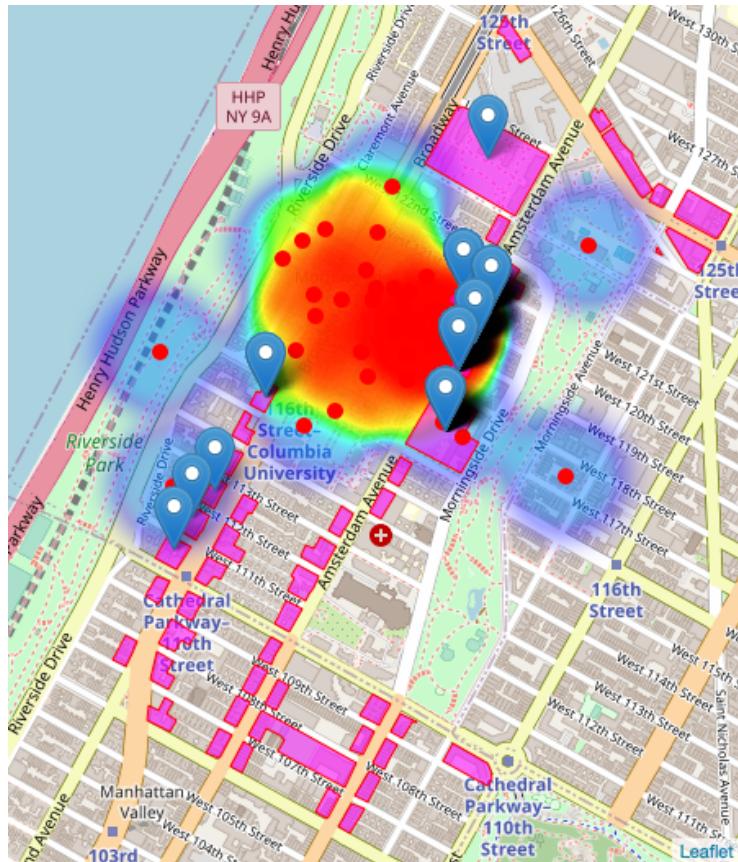


Figure 11: Visualization of the reserved commercial zones (in purple), venues related to science, technology, environment, electronic, computer, engineering Lab or teaching building (red dots), and their spatial distribution/concentration (color gradient from blue (the lowest density), to red (the highest density)). The blue markers are the commercial zones within a 100m radius from a venue.

Final Result

For a more readable map, only the candidates' marker was kept. This map represents the final results which are the best places (with their postal addresses) to start the GreenTech project which is located:

- In the neighborhood with the best ratio between university venues, population density, crime rate, rent price
- In a reserved commercial zone
- Within a radius of 100m from a Science/Computer/Technology/Electronic/Environment lab/teaching building

Addresses
Five Guys 2847 Broadway
Bank Street College of Education 610 West 11...
2881 Broadway Morningside Heights
Jerome Greene Hall West 116th Street Morning...
Regnor Court 601 West 115th Street
424 West 119th Street Morningside Heights
The Poinsettia West 120th Street Morningside...
1225 Amsterdam Avenue Morningside Heights
Teachers College West 122nd Street Morningsi...
90 La Salle Street Manhattan

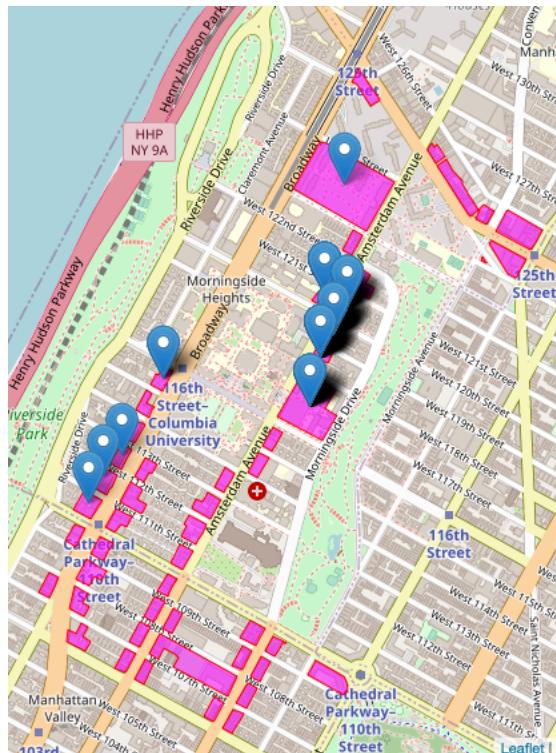


Figure 10: The blue markers are the commercial zones within a 100m radius of a venue related to science, technology, environment, electronic, computer, engineering Lab or teaching building.

Conclusion and discussion

Now we know the addresses of the best places for the GreenTech project. The study could have been more specific with, for example, real estate data in order to know the available places to buy or rent. Moreover, the access to “foot traffic data” from Foursquare could have been really interesting in order to see the most frequented places, which would determine the most profitable location.