Capstone Project:

The Battle of Neighborhoods - Week 2

Finding a suitable house for rent in Manhattan New York

Melonica Mohapatra

Sept. 24th , 2019

Report Content

1. Introduction Section :

- The “business problem” to be solved by this project and who may be interested

2. Data Section:

- Describe Data requirements and Sources needed to solve the problem

3. Methodology section:

- Main component of the report - Execute data processing, describe/discuss any exploratory data analysis and/or inferential statistical testing performed, and/or machine learnings used.

4. Results section:

- Discussion of the results and finding of answer

5. Discussion section:

- Discussion of observations noted and any recommendations

6. Conclusion section:

- Answer chosen and conclusions.

1.0 Introduction

1.1 Scenario and Background

I am an engineer and currently living in Gajuwaka, Visakhapatnam, India. I am staying in a two bedroom flat. Restaurant and bus terminal is nearby my home and I enjoy movies and international cuisine available nearby. I have applied to universities for doing MS in US. New York is one of the most likely place to get.

1.2 Problem to be resolved:

Manhattan being chosen to be the place to stay, this project revolves around finding a two bedroom flat around Manhattan not exceeding monthly rent $7000. As a student I will have to

travel to university routinely. Public transport like metro and other facilities would be required for daily commuting. Trending venues to relax with international cuisine and entertainment also would be a criteria besides a cap of $7000 a month as rent.

1.3 Interested Audience

This project would be used to analyze data and find most suitable apartment out of available apartments around Manhattan. The methodology and tools are extremely useful to do analysis on data by using foursquare APIs. Lat Long based information helps to zero in the target. The similar approach may help to find out a suitable location to start a business too. I believe the methodology, tools and strategy used in this project

is relevant for a person or entity considering moving to a major city in US, Europe. The use of FourSquare APIs for geo data mapping techniques will help resolve the key questions.

2.0 Data Section

Gajuwaka is a major suburb of Visakhapatnam City, India. The neighbourhood of Visakhapatnam is considered the biggest shopping district in Andhra Pradesh by revenue. Though conceived as a residential locality, it is now one of the principal shopping districts of the city. Gajuwaka area of Visakhapatnam have Highest per capita income in Andhra Pradesh. Gajuwaka is one of the busiest shopping districts of Andhra Pradesh. The neighbourhood is considered to be the biggest shopping district in Andhra Pradesh by revenue. There are a number of apparel, jewellery, theaters, shopping malls based in Gajuwaka. Several hotels and restaurants with international cuisine are available. Public transport is conveniently available to go to work.

The New York City (NYC) or simply New York (NY), is the most populous city in the United States. New York City comprises 5 boroughs sitting where the Hudson River meets the Atlantic Ocean. At its core is Manhattan, a densely populated borough that's among the world's major commercial, financial and cultural centers. The following activities are performed to gather data for the project.

2.1 Data Requirements

- Geodata for current residence in Gajuwaka, Visakhapatnam with venues established by using Foursquare APIs.
- List of Manhattan (MH) neighborhoods with clustered venues established via Foursquare APIs

https://en.wikipedia.org/wiki/List_of_Manhattan_neighborhoods#Midtown_neighborhoods

- List of subway metro stations in Manhattan with addresses and geo data (lat,long):
https://en.wikipedia.org/wiki/List_of_New_York_City_Subway_stations_in_Manhattan
https://www.google.com/maps/search/manhattan+subway+metro+stations/@40.7837297,-74.1033043,11z/data=!3m1!4b1
https://www.google.com/maps/search/manhattan+subway+metro+stations/@40.7835603,-74.1033047,11z/data=!3m1!4b1

- List of apartments for rent in Manhattan area with information on neighborhood location, address, number of beds, area size, monthly rent price and complemented with geo data via Nominatim.

http://www.rentmanhattan.com/index.cfm?page=search&state=results
https://www.nestpick.com/search?city=newyork&page=1&order=relevance

&display=grid&currency=USD
- Place to work in Manhattan (Park Avenue and 53rd St) for reference

2.2 Data Sources, Data Processing and Tools used

- Gajuwaka data and map is to be created with use of Nominatim , Foursquare and Folium mapping
- Manhattan neighborhoods were obtained from Wikipedia and organized by Neighborhoods with geodata
via Nominatim for mapping with Folium.
- List of Subway stations was obtained via Wikipedia, NY Transit web site and Google map,
- List of apartments for rent was consolidated from web-scraping real estate sites for MH. The geolocation (lat,long) data was found with algorithm coding and using Nominatim.
- Folium map was the basis of mapping with various features to consolidate all data in ONE map where one can visualize all details needed to make a selection of apartment

3.0 Methodology

The Strategy to find the answer:

The strategy is based on mapping the collected data as described in section 2.0, in order to zero in at least two candidate places for rent. The information will be consolidated in ONE MAP where one can see the details of the apartment, the cluster of venues in the neighborhood and the relative location from a subway station and from work place. A measurement tool icon will also be provided. The popups on

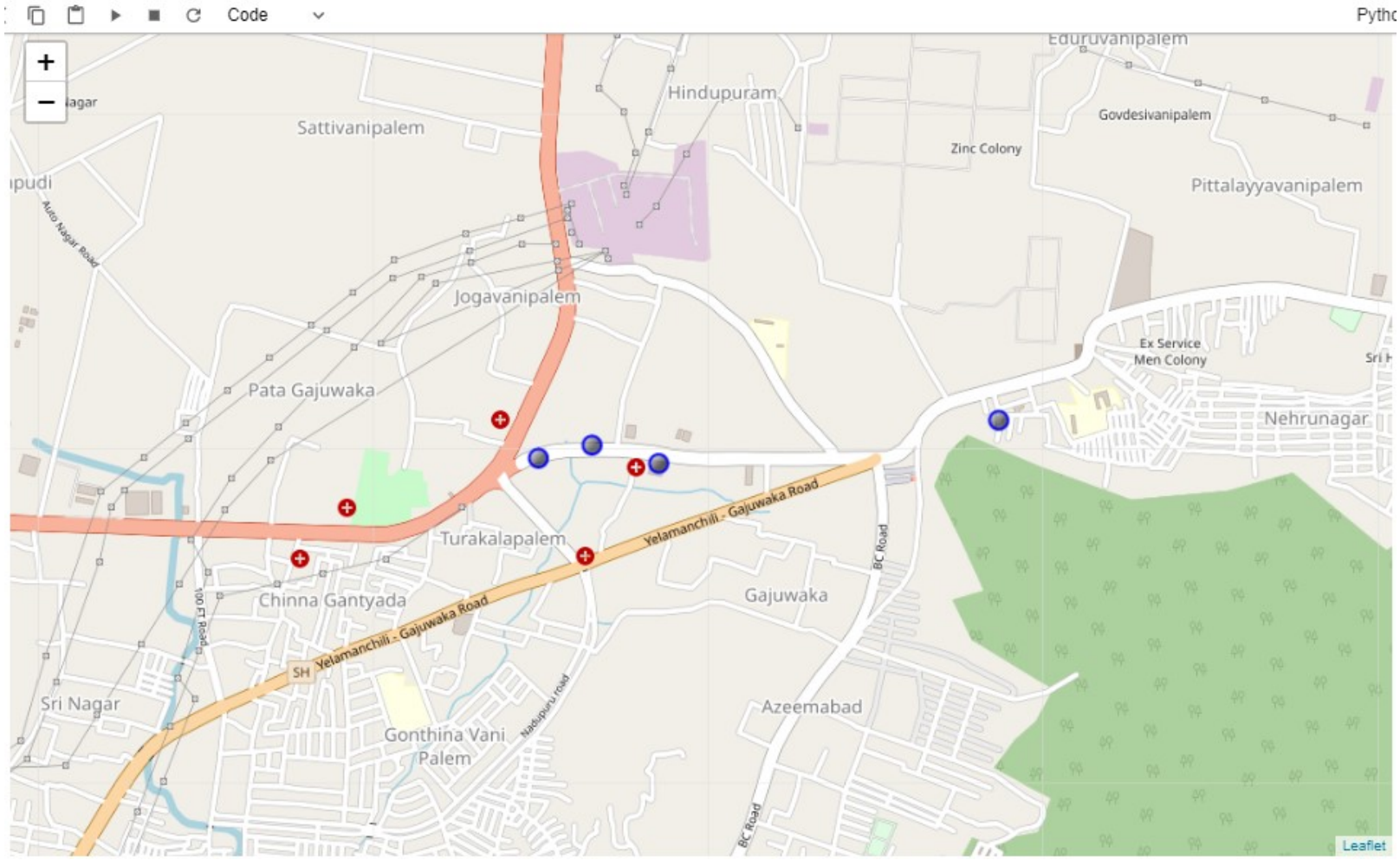
the map items will display rent price, location and cluster of venues applicable.

The Tools:

Web-scraping of sites is used to consolidate data-frame information which was saved as .csv files for convenience and to simply the report. Geodata was obtained by coding a program to use Nominatim to get latitude and longitude of subway stations and also for each of (144 units) the apartments for rent listed. Geopy_distance and Nominatim were used to find out relative distances. Seaborn graphic was used for general statistics on rental data. Maps with popups labels facilitate quick identification of location, price and feature.

4.0 Execution and Results

Current residence Neighborhood in Gajuwaka



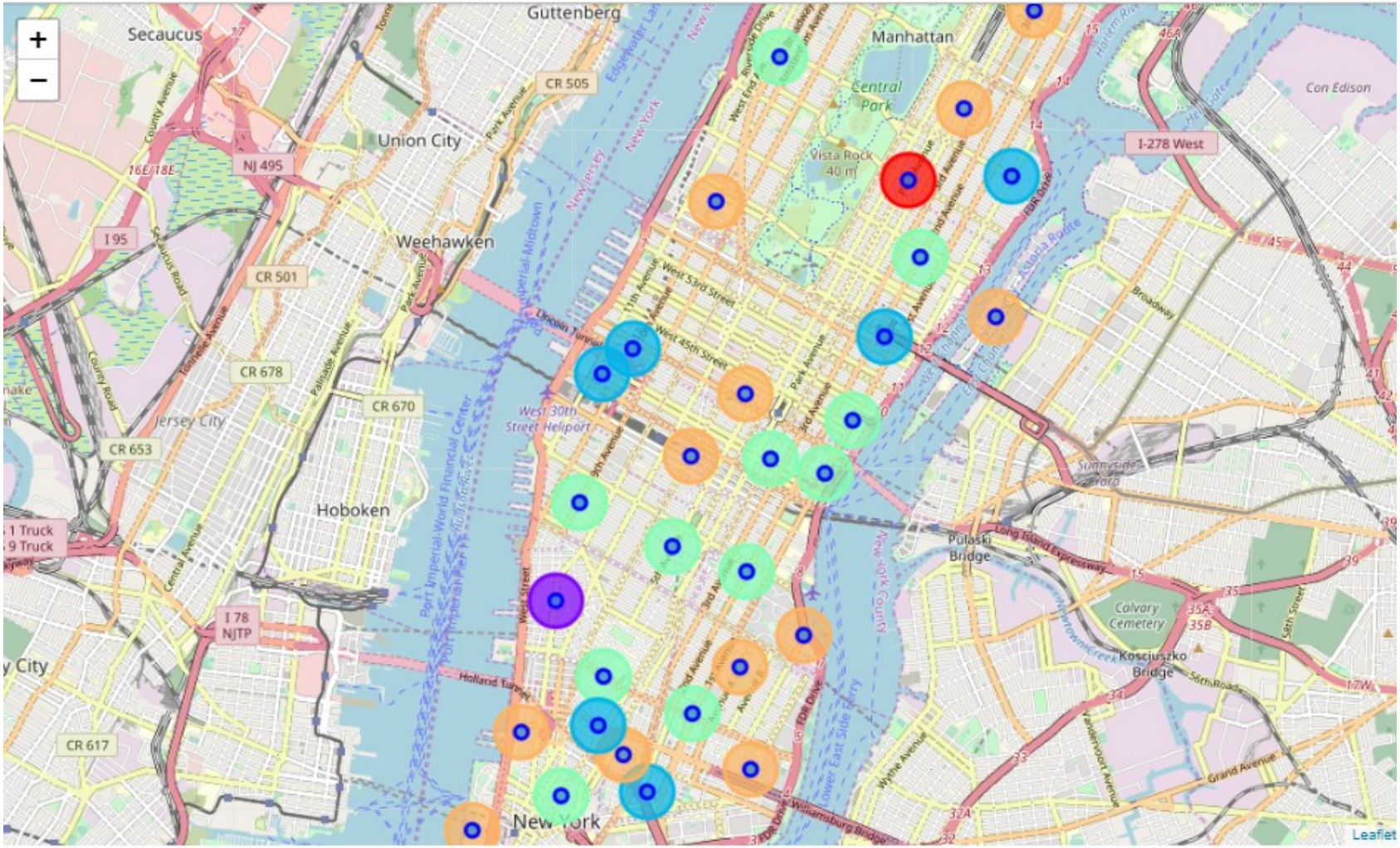
Venues around Neighborhood in

```
SGnearby_venues.head(10)
```

	name	categories	lat	lng
0	Best western - Ramachandra	Hotel	17.686031	83.206629
1	Mohini Theatre	Movie Theater	17.685447	83.208834
2	Amul Scooping Parlour	Ice Cream Shop	17.686811	83.219998
3	Four Seasons	Hotel	17.685594	83.204890

Map of Gajuwaka residence place with venues in Neighborhood - for reference

Manhattan Map - Neighborhoods and Cluster of Venues



GeoData Manhattan apartments for rent

```
# csv files with rental places with basic data but still wihtout geodata ( latitude and longitude)
mh_rent=pd.read_csv('MH_flats_price.csv')
mh_rent.head()
```

	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
0	West 105th Street	Upper West Side	2.94	5.0	3400	10000	NaN	NaN
1	East 97th Street	Upper East Side	3.57	3.0	2100	7500	NaN	NaN
2	West 105th Street	Upper West Side	1.89	4.0	2800	5300	NaN	NaN
3	CARMINE ST.	West Village	3.03	2.0	1650	5000	NaN	NaN
4	171 W 23RD ST.	Chelsea	3.45	2.0	1450	5000	NaN	NaN

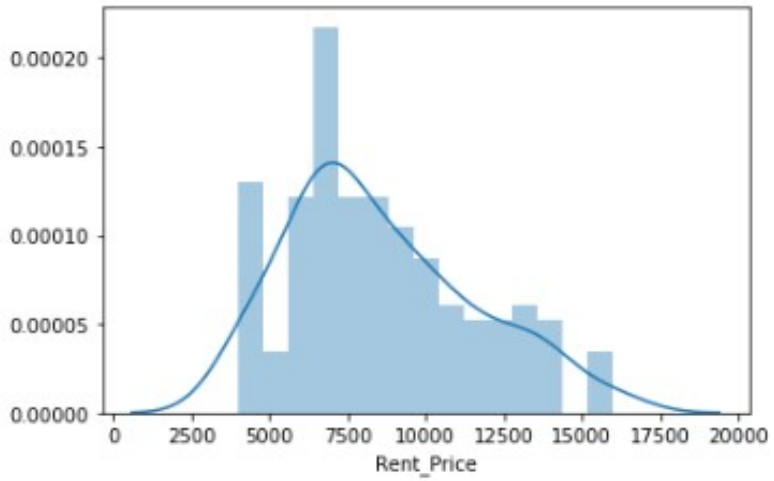
```
mh_rent.tail()
```

	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
139	200 East 72nd Street	Rental in Lenox Hill	5.15	3.0	1700	8750	NaN	NaN
140	50 Murray Street	No fee rental in Tribeca	7.11	2.0	1223	8700	NaN	NaN
141	300 East 56th Street	No fee rental in Midtown East	3.87	3.0	2100	8118	NaN	NaN
142	1930 Broadway	No fee rental in Central Park West	5.06	2.0	1600	8095	NaN	NaN
143	33 West 9th Street	Rental in Greenwich Village	6.67	2.0	1500	10000	NaN	NaN

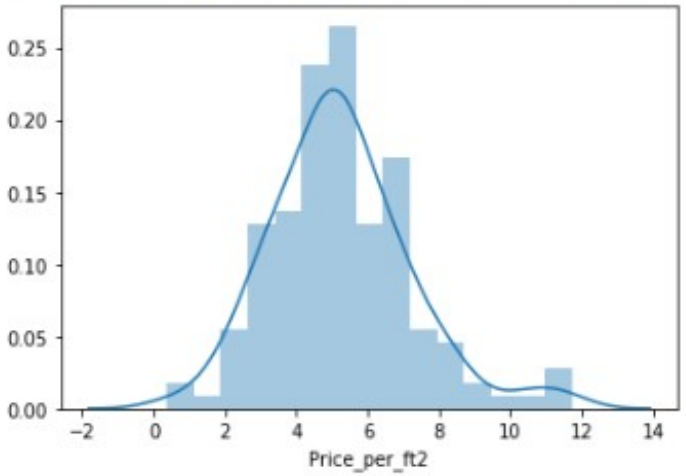
Rental Price Statistics Manhattan Apartments

Budget US$7000/month is around the mean

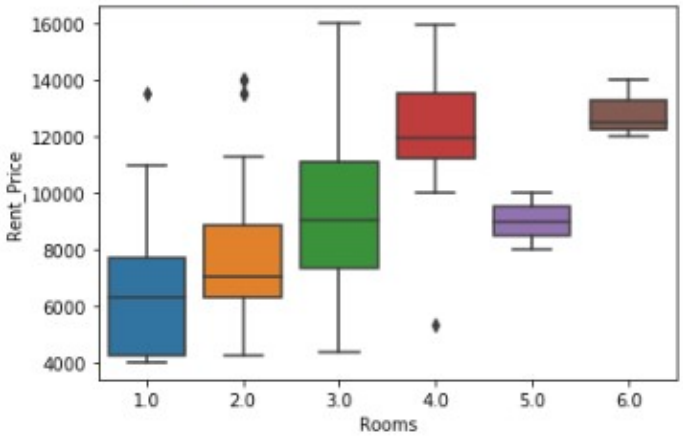
<matplotlib.axes._subplots.AxesSubplot at 0x7fccdb690a90>



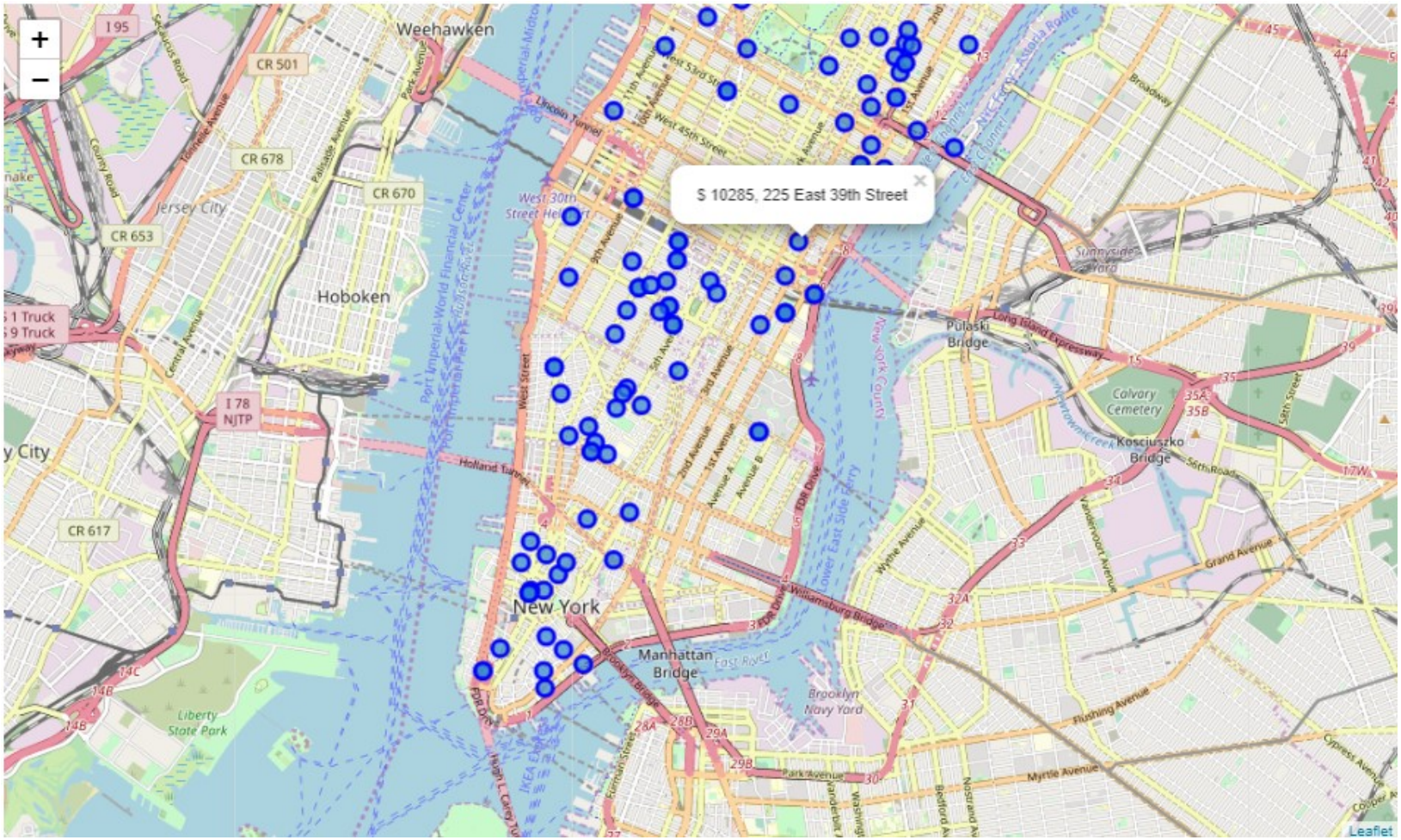
<matplotlib.axes._subplots.AxesSubplot at 0x7fccd9b25438>



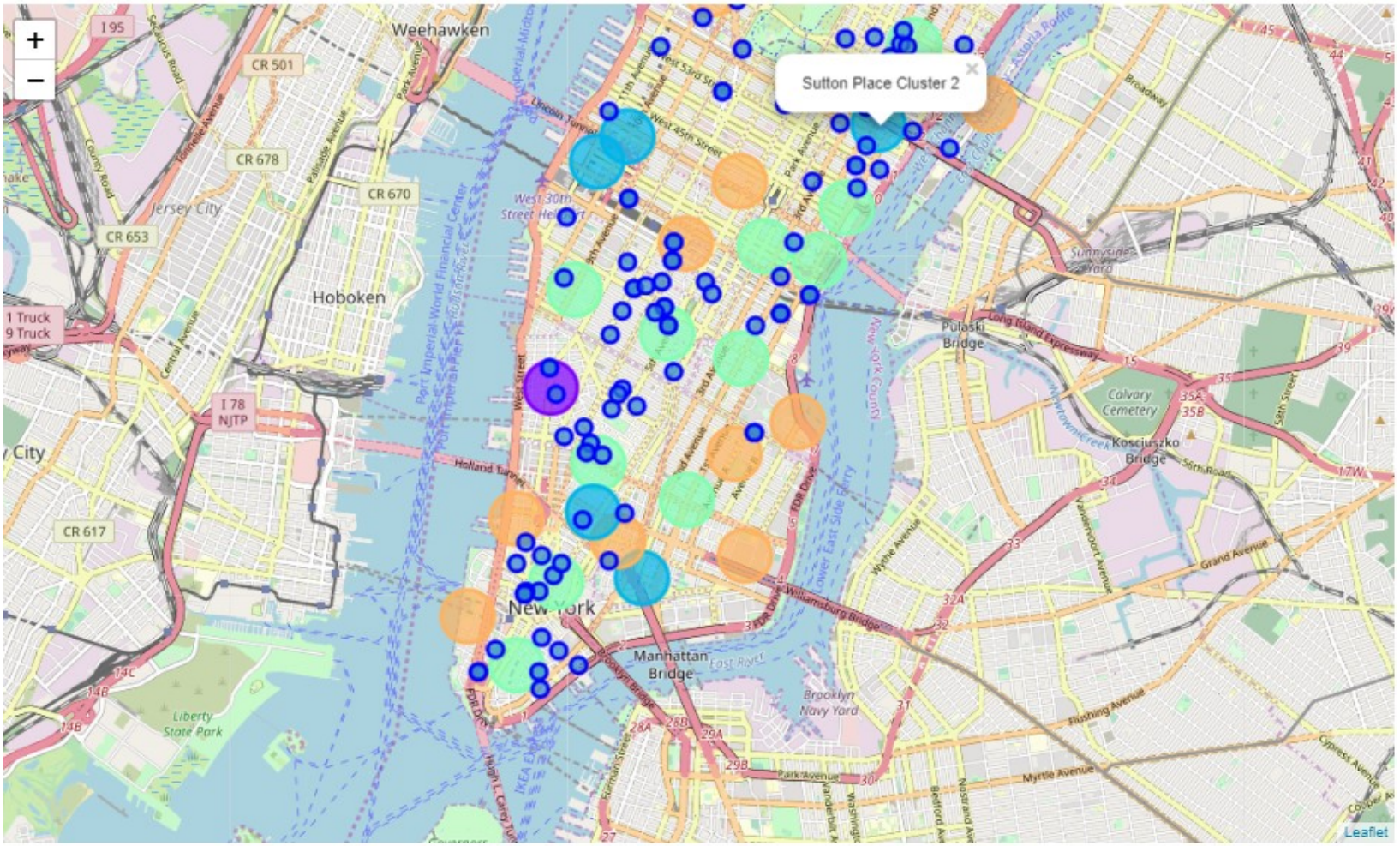
<matplotlib.axes._subplots.AxesSubplot at 0x7f20558f6208>



Apartments for Rent in Manhattan



Manhattan apartments for rent with venue clusters



Venues of cluster 3

```
## kk is the cluster number to explore
kk = 3
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == kk, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Inwood	Mexican Restaurant	Lounge	Pizza Place	Café	Wine Bar	Bakery	American Restaurant	Park	Frozen Yogurt Shop	Spanish Restaurant
5	Manhattanville	Deli / Bodega	Italian Restaurant	Seafood Restaurant	Mexican Restaurant	Sushi Restaurant	Beer Garden	Coffee Shop	Falafel Restaurant	Bike Trail	Other Nightlife
10	Lenox Hill	Sushi Restaurant	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Pizza Place	Burger Joint	Deli / Bodega	Gym	Sporting Goods Shop	Thai Restaurant
12	Upper West Side	Italian Restaurant	Bar	Bakery	Vegetarian / Vegan Restaurant	Indian Restaurant	Coffee Shop	Cosmetics Shop	Wine Bar	Mexican Restaurant	Sushi Restaurant
16	Murray Hill	Sandwich Place	Hotel	Japanese Restaurant	Gym / Fitness Center	Coffee Shop	Salon / Barbershop	Burger Joint	French Restaurant	Bar	Italian Restaurant
17	Chelsea	Coffee Shop	Italian Restaurant	Ice Cream Shop	Bakery	Nightclub	Theater	Art Gallery	Seafood Restaurant	American Restaurant	Hotel
18	Greenwich Village	Italian Restaurant	Sushi Restaurant	French Restaurant	Clothing Store	Chinese Restaurant	Café	Indian Restaurant	Bakery	Seafood Restaurant	Electronics Store
27	Gramercy	Italian Restaurant	Restaurant	Thrift / Vintage Store	Cocktail Bar	Bagel Shop	Coffee Shop	Pizza Place	Mexican Restaurant	Grocery Store	Wine Shop
29	Financial District	Coffee Shop	Hotel	Gym	Wine Shop	Steakhouse	Bar	Italian Restaurant	Pizza Place	Park	Gym / Fitness Center
31	Noho	Italian Restaurant	French Restaurant	Cocktail Bar	Gift Shop	Bookstore	Grocery Store	Mexican Restaurant	Hotel	Sushi Restaurant	Coffee Shop
32	Civic Center	Gym / Fitness Center	Bakery	Italian Restaurant	Cocktail Bar	French Restaurant	Sandwich Place	Coffee Shop	Gym	Yoga Studio	Park
35	Turtle Bay	Italian	Coffee Shop	Steakhouse	Wine Bar	Sushi	Hotel	Noodle House	Indian	Japanese	French
36	Tudor City	Café	Park	Pizza Place	Mexican Restaurant	Greek Restaurant	Sushi Restaurant	Hotel	Deli / Bodega	Diner	Dog Run
38	Flatiron	Italian Restaurant	American Restaurant	Gym	Gym / Fitness Center	Yoga Studio	Vegetarian / Vegan Restaurant	Bakery	Clothing Store	Cosmetics Shop	Cycle Studio

Manhattan subway stations geo data

	sub_station	sub_address	lat	long
0	Dyckman Street Subway Station	170 Nagle Ave, New York, NY 10034, USA	40.861857	-73.924509
1	57 Street Subway Station	New York, NY 10106, USA	40.764250	-73.954525
2	Broad St	New York, NY 10005, USA	40.730862	-73.987156
3	175 Street Station	807 W 177th St, New York, NY 10033, USA	40.847991	-73.939785
4	5 Av and 53 St	New York, NY 10022, USA	40.764250	-73.954525

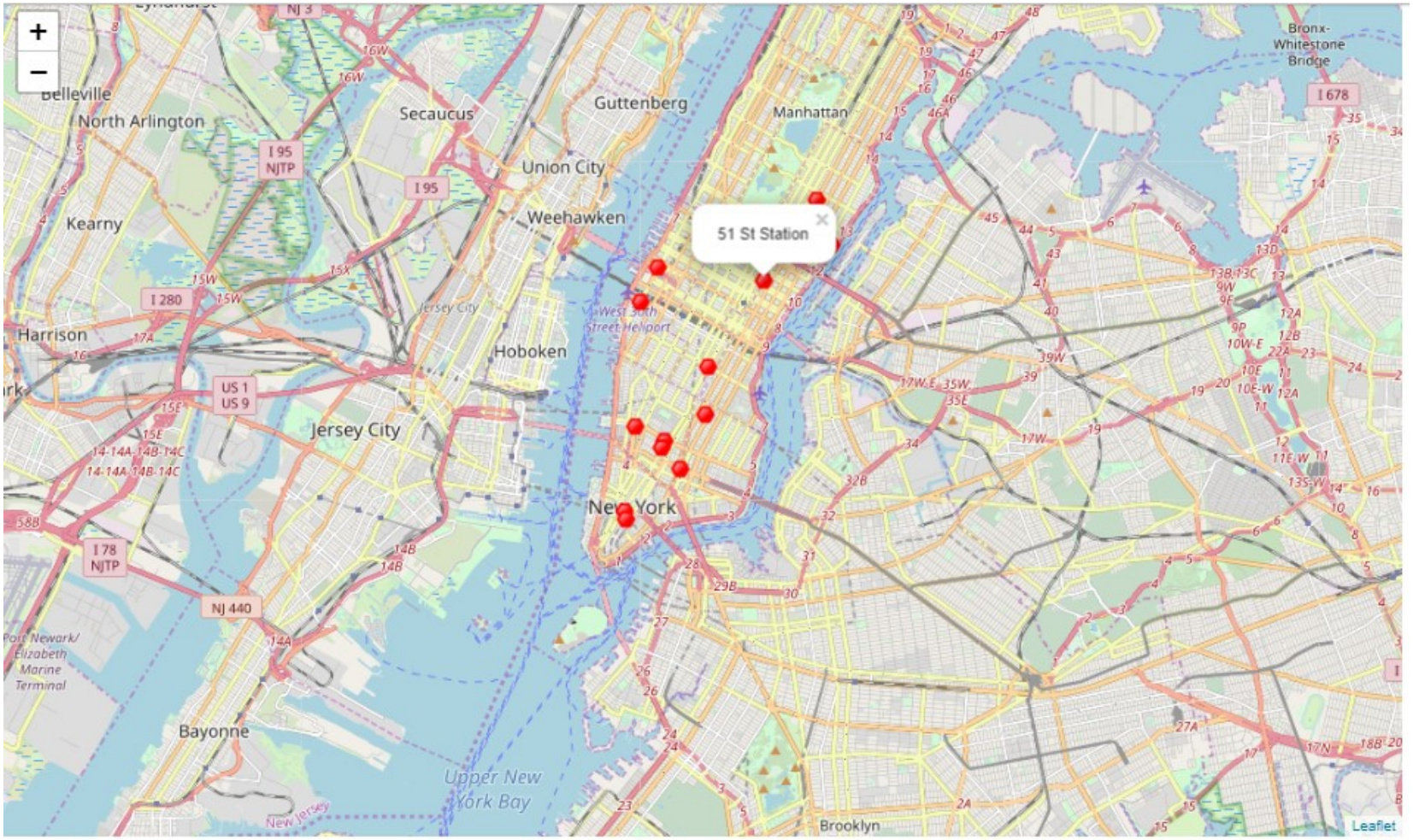
```
# removing duplicate rows and creating new set mhsub1
mhsub1=mh.drop_duplicates(subset=['lat','long'], keep="last").reset_index(drop=True)
mhsub1.shape
```

```
(22, 4)
```

```
mhsub1.tail()
```

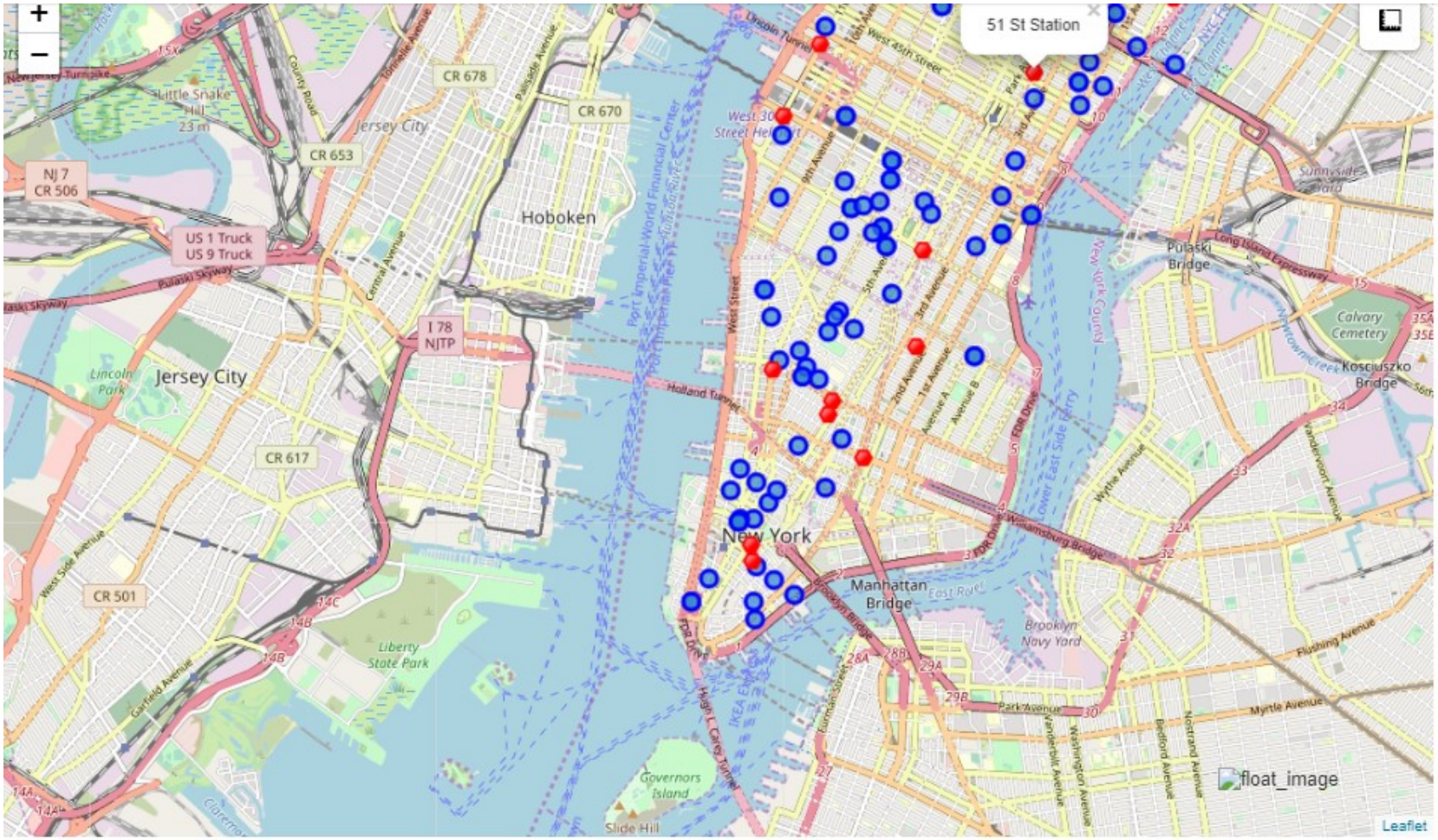
	sub_station	sub_address	lat	long
17	190 Street Subway Station	Bennett Ave, New York, NY 10040, USA	40.858113	-73.932983
18	59 St-Lexington Av Station	E 60th St, New York, NY 10065, USA	40.762259	-73.966271
19	57 Street Station	New York, NY 10019, United States	40.764250	-73.954525
20	14 Street / 8 Av	New York, NY 10014, United States	40.730862	-73.987156
21	MTA New York City	525 11th Ave, New York, NY 10018, USA	40.759809	-73.999282

Map of Manhattan showing the location of subway stations



Map of Manhattan showing places for rent and the subway locations nearby. Now, we can visualize the desirable rental places and their nearest subway station. Popups

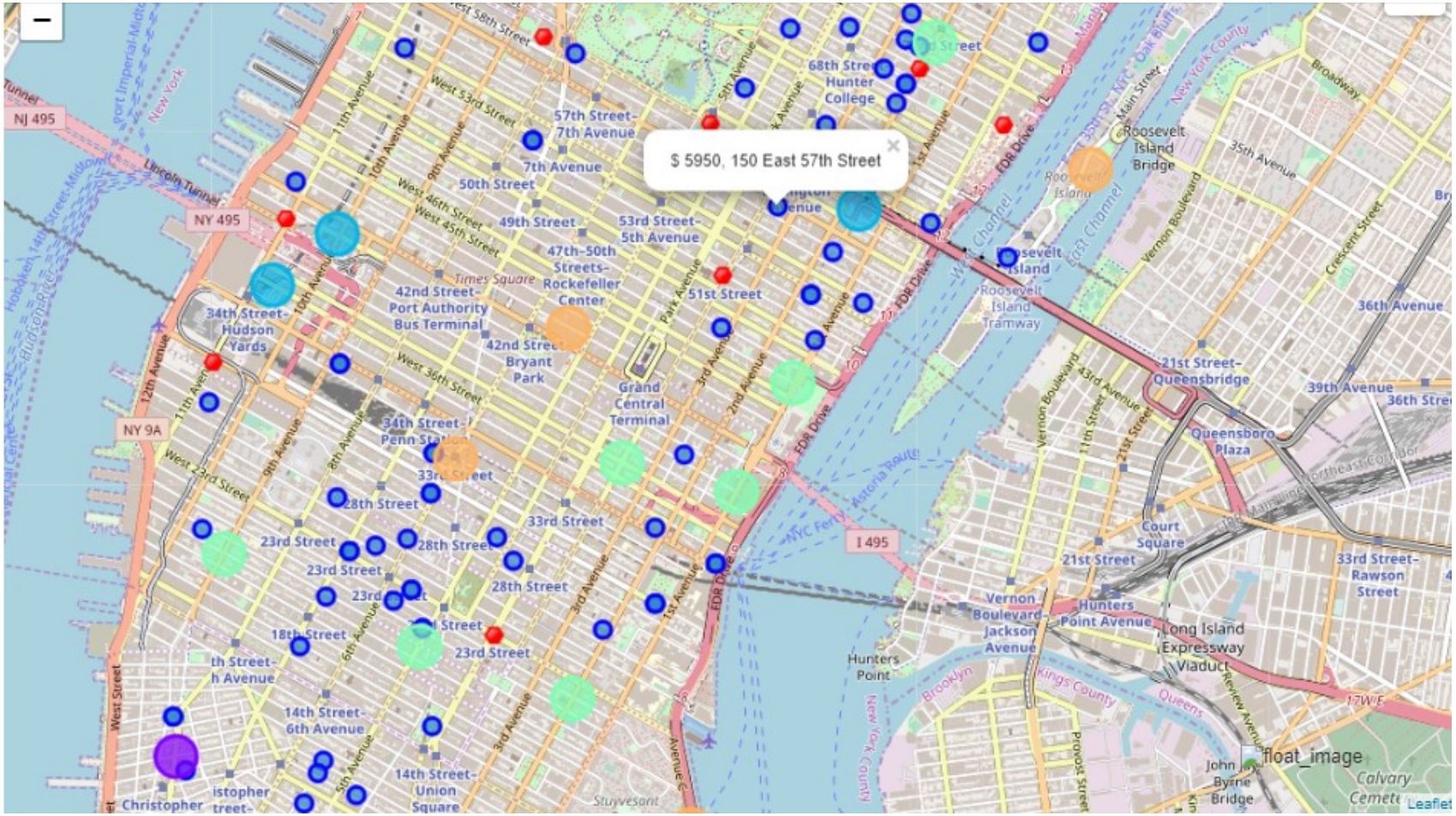
Apartments for rent (blue) and subway stations (red)



Results ONE CONSOLIDATE MAP Map of Manhattan with rental places, subway locations and cluster of venues. Blue dots are apartments available for rent, red dots are Subway stations and Bubbles are the clusters of venues.

Selected Apartment!

The ONE consolidated map shows all information for decision:
Apartments address, price, neighborhood, cluster of venues and subway station nearby.
Blue dots=apts , Red dots=Subway station, Bubbles=Cluster of Venues



Selected Apartment!

Apartment Selection

Using the "one map" above, I was able to explore all possibilities since the popups provide the information needed for a good decision.

Apartment 1 rent cost is US$5950 below the US7000 budget. Apt 1 is located 400 meters from 59th Street Lexington AV Station. I can use subway for other places around. Venues for this apt are as of Cluster 2 and it is located in a fine district in the East side of Manhattan.

Apartment 2 rent cost is US6935, just under the US7000 budget. Apt 2 is located 60 meters from subway station at Fulton Street, but I will have to ride the subway daily to university. Venues for this apt are as of Cluster 3.

APARTMENT 1 is a better choice for rent as well as convenience it provides.