The Battle of Neighborhoods

Coursera Capstone Project

Khalid A.

Report Content

Introduction Section :

1.1 Discussion of the "backgroung situation" leading to the problem at hand:
1.2 Problem to be resolved
1.3 Audience for this project.

Data Section:

2.1 Data of Current Situation (current residence place)
2.2 Data required to resolve the problem

Methodology section :

3.1 Process steps and strategy to resolve the problem
3.2 Data Science Methods, machine learing, mapping tools and exploratory data analysis.

Results section :

Discussion of the results and how they help to take a decision.

Discussion section :

Elaboration and discussion on any observations and/or recommendations for improvement.

Conclusion section :

Decision taken and Report Conclusion.

Introduction Section

Scenario and Background

Let's assume I got client residing in Downtown Singapore. he currently live in Downtown Telok Ayer MRT metro station. he have been offered a job in Manhattan, NY. The key question is : How can he find a convenient and enjoyable place similar to where he live now in Singapore? In order to make a comparison and evaluation of the rental options in Manhattan NY, I must set some basis, therefore the apartment in Manhattan must meet the following demands: apartment must be 2 or 3 bedrooms desired location is near a metro station in the Manhattan area and within 1.0 mile (1.6 km) radius price of rent not exceed $7,000 per month etc.

Problem to be resolved:

How to find an apartment in Manhattan with the following conditions:

- Apartment with min 2 bedrooms.
- Monthly rent not to exceed US$7000/month.
- Located within walking distance (<=1.0 mile, 1.6 km) from a subway metro station in Manhattan.
- Venues and amenities as in his current residence.

Interested Audience

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 or Asia. Europe, US or Asia, Likewise, it can be helpful approach to explore the opening of a new business. The use of FourSquare data and mapping techniques combined with data analysis will help resolve the key questions arisen.

Data Section

Data Requirements

- Geodata for his current residence in Singapore with venues established using Foursquare.
- List of Manhattan (MH) neighborhoods with clustered venues established via Foursquare (as in Course Lab). 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)
- 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=new-
- Place to work in Manhattan (Park Avenue and 53rd St) for reference.

Data Sources, Data Processing and Tools used

- Singapore 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.

Methodology section

The Strategy to find the answer

The strategy is based on mapping the described data in section 2.0, in order to facilitate the choice of 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 pop ups on the map items will display rent price, location and cluster of venues applicable.

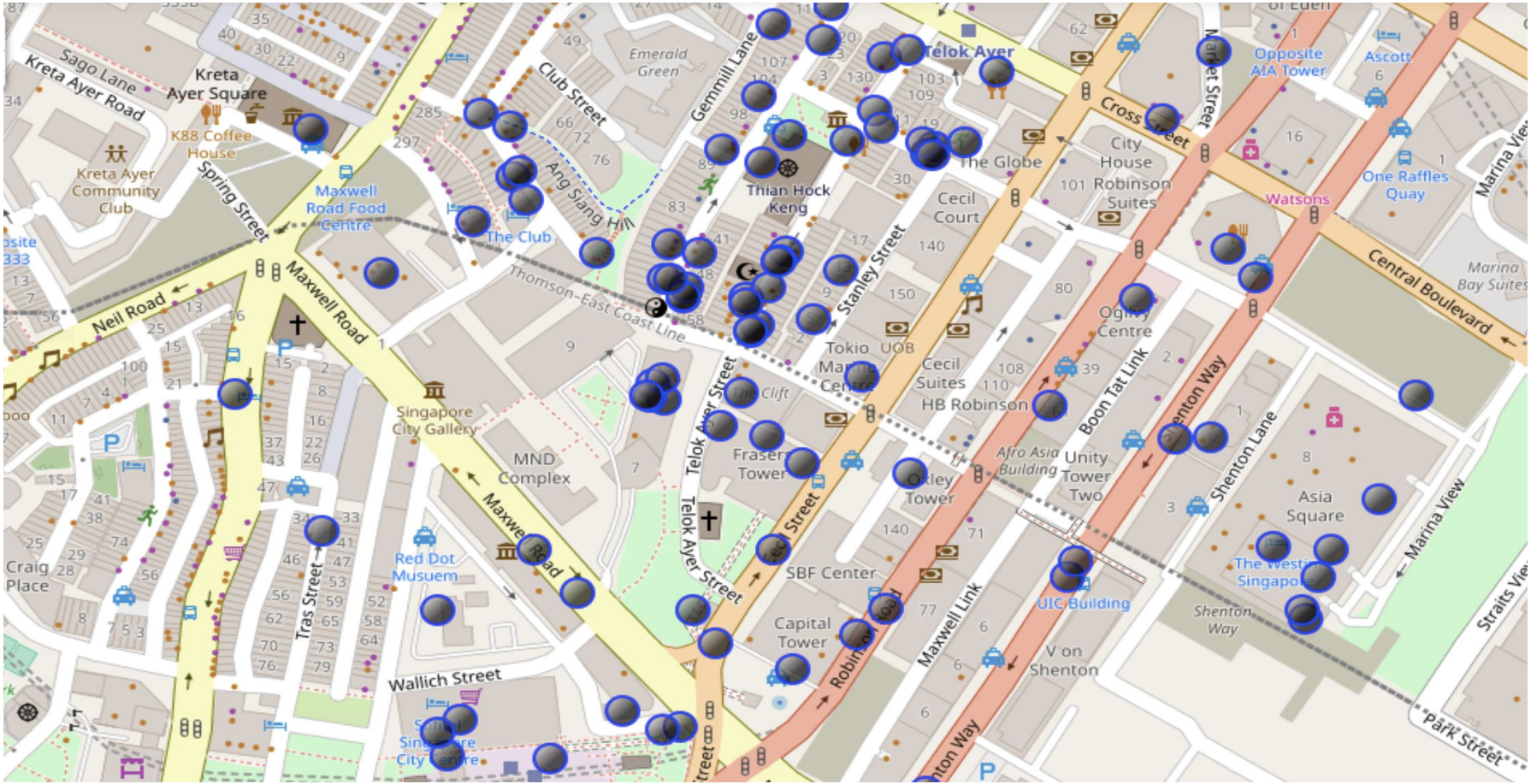
The Tools

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 establish relative distances. Seaborn graphic was used for general statistics on rental data.

Maps with pop ups labels allow quick identification of location, price and feature, thus making the selection very easy

Execution and Results

Current residence neighborhood in Singapore

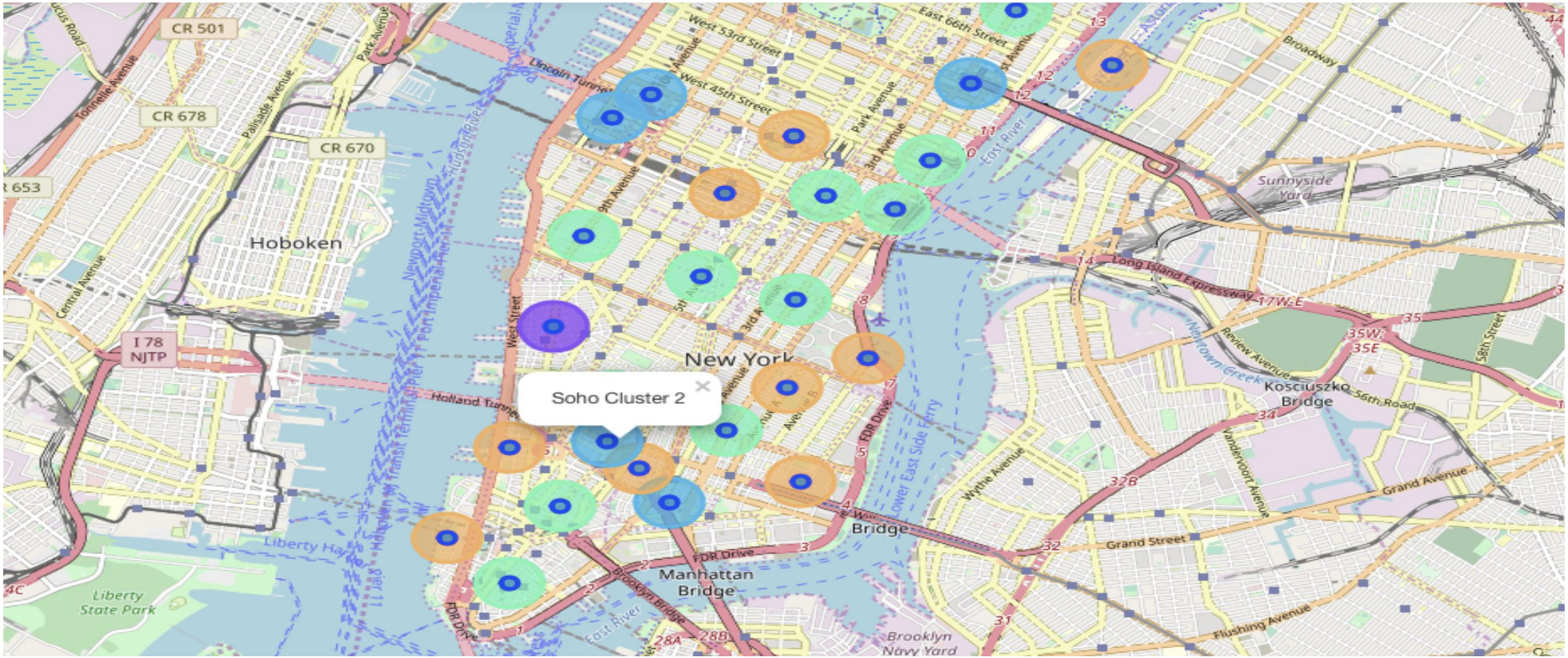


venues around current residence in Singapore

```
# Venues near current Singapore residence place
SGnearby_venues.head(10)
```

	name	categories	lat	lng
0	Napoleon Food & Wine Bar	Wine Bar	1.279925	103.847333
1	Park Bench Deli	Deli / Bodega	1.279872	103.847287
2	Native	Cocktail Bar	1.280135	103.846844
3	Muchachos	Burrito Place	1.279175	103.847082
4	Matt's | The Chocolate Shop	Dessert Shop	1.280462	103.846950
5	Freehouse	Beer Garden	1.281254	103.848513
6	PS.Cafe	Café	1.280468	103.846264
7	왕대박 Wang Dae Bak Korean BBQ Restaurant	Korean Restaurant	1.281345	103.847551
8	Ancient Therapy	Massage Studio	1.280413	103.847481
9	Oven & Fried Chicken	Korean Restaurant	1.280479	103.847522

Map of Manhattan neighborhoods



Geodata Manhattan apts for rent

```
mh_rent=pd.read_csv('MH_rent_latlong.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	40.799771	-73.966213
1	East 97th Street	Upper East Side	3.57	3.0	2100	7500	40.788585	-73.955277
2	West 105th Street	Upper West Side	1.89	4.0	2800	5300	40.799771	-73.966213
3	CARMINE ST.	West Village	3.03	2.0	1650	5000	40.730523	-74.001873
4	171 W 23RD ST.	Chelsea	3.45	2.0	1450	5000	40.744118	-73.995299

```
mh_rent.tail()
```

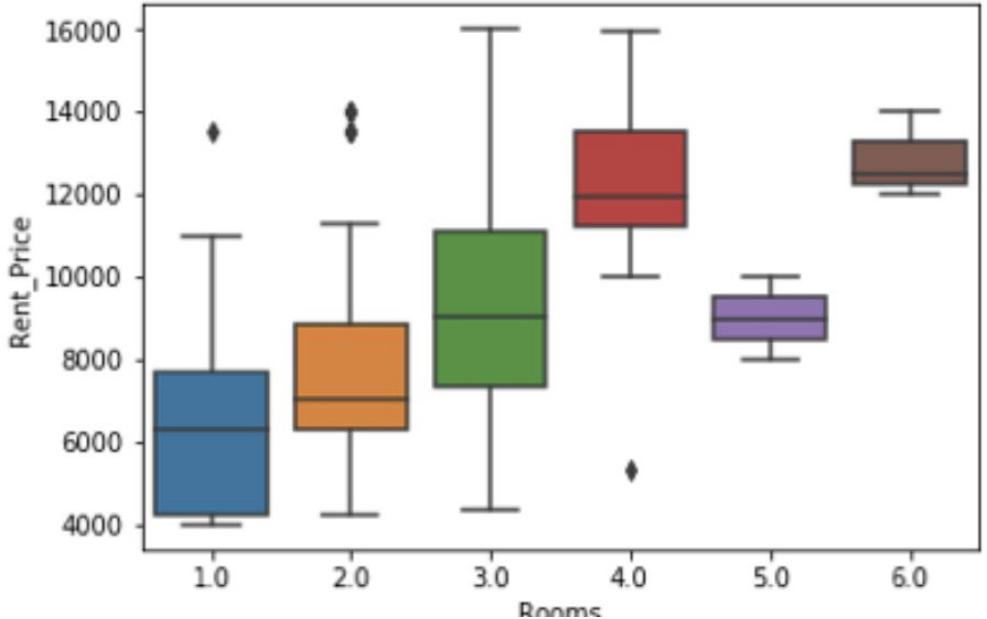
	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
139	200 East 72nd Street	Rental in Lenox Hill	5.15	3.0	1700	8750	40.769465	-73.960339
140	50 Murray Street	No fee rental in Tribeca	7.11	2.0	1223	8700	40.714051	-74.009608
141	300 East 56th Street	No fee rental in Midtown East	3.87	3.0	2100	8118	40.758216	-73.965190
142	1930 Broadway	No fee rental in Central Park West	5.06	2.0	1600	8095	40.772474	-73.981901
143	33 West 9th Street	Rental in Greenwich Village	6.67	2.0	1500	10000	40.733691	-73.997323

Manhattan apartment rent price statistics

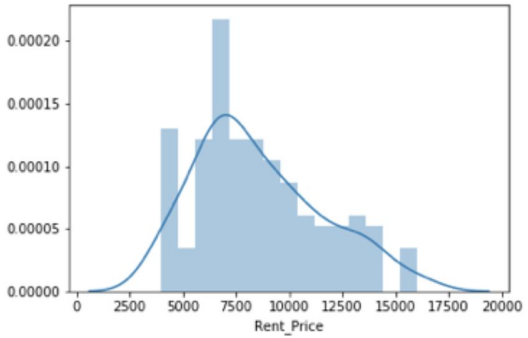
A US 7000 Dollar per month rent is actually around the mean value

```
sns.boxplot(x='Rooms', y= 'Rent_Price', data=mh_rent)
```

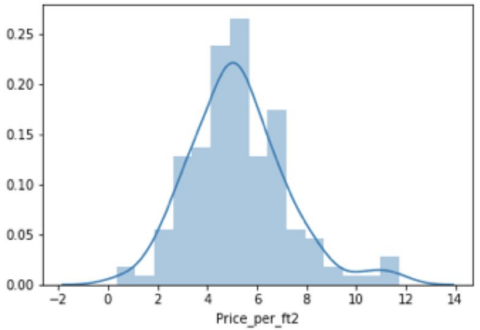
```
<matplotlib.axes._subplots.AxesSubplot at 0x1a25f2a2b0>
```



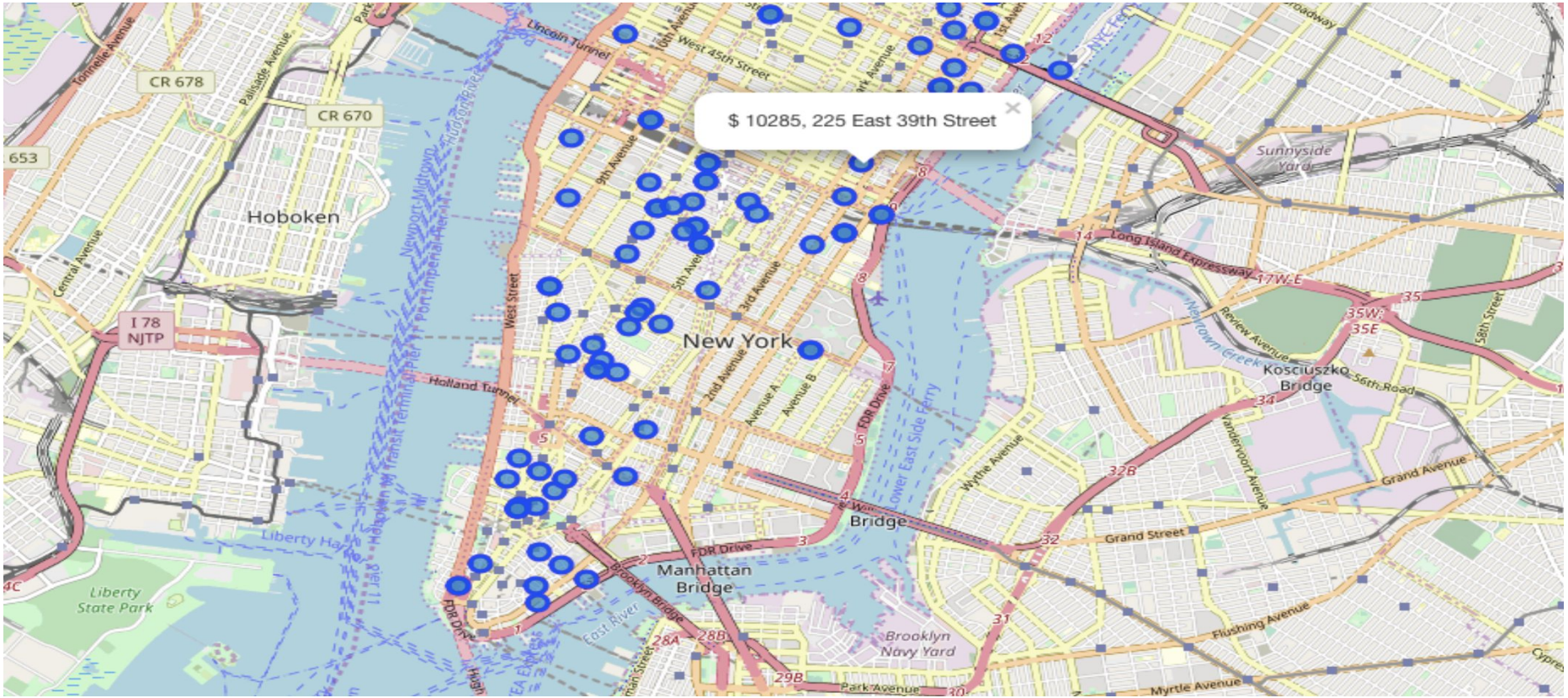

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a25dd8400>
```



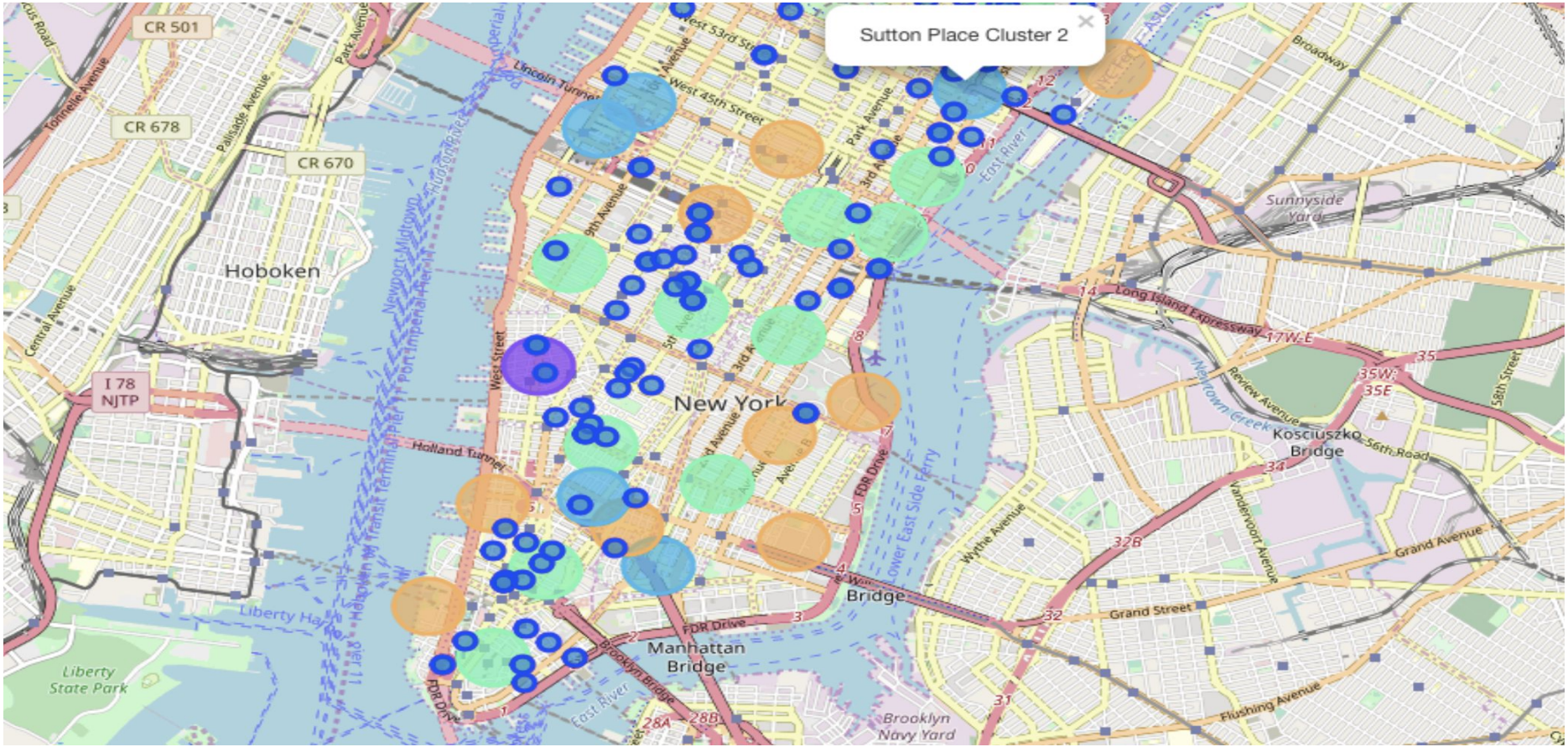

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a2415fc18>
```



Map of Manhattan apartments for rent



Map of Manhattan showing the places for rent and the cluster of venues



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_m
```

	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

Manhattan Subway locations Geodata

click to scroll output; double click to hide

		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

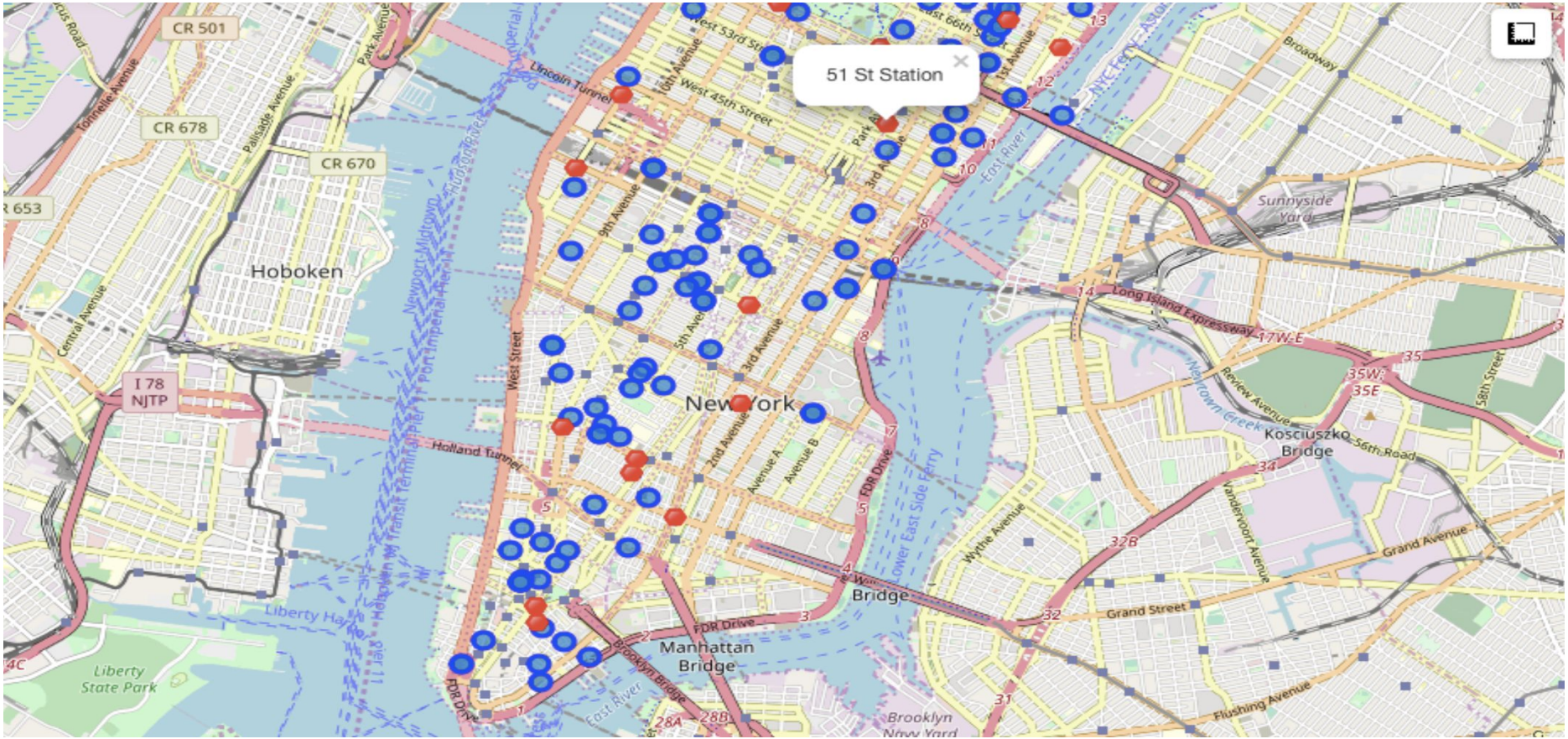
```
# removing duplicate rows and creating new set mhsub1
mhsub1=mh.drop_duplicates(subset=['lat','long'], keep="last").reset_index(drop=True)
mhsub1.shape
```

```
(22, 4)
```

```
mhsub1.tail()
```

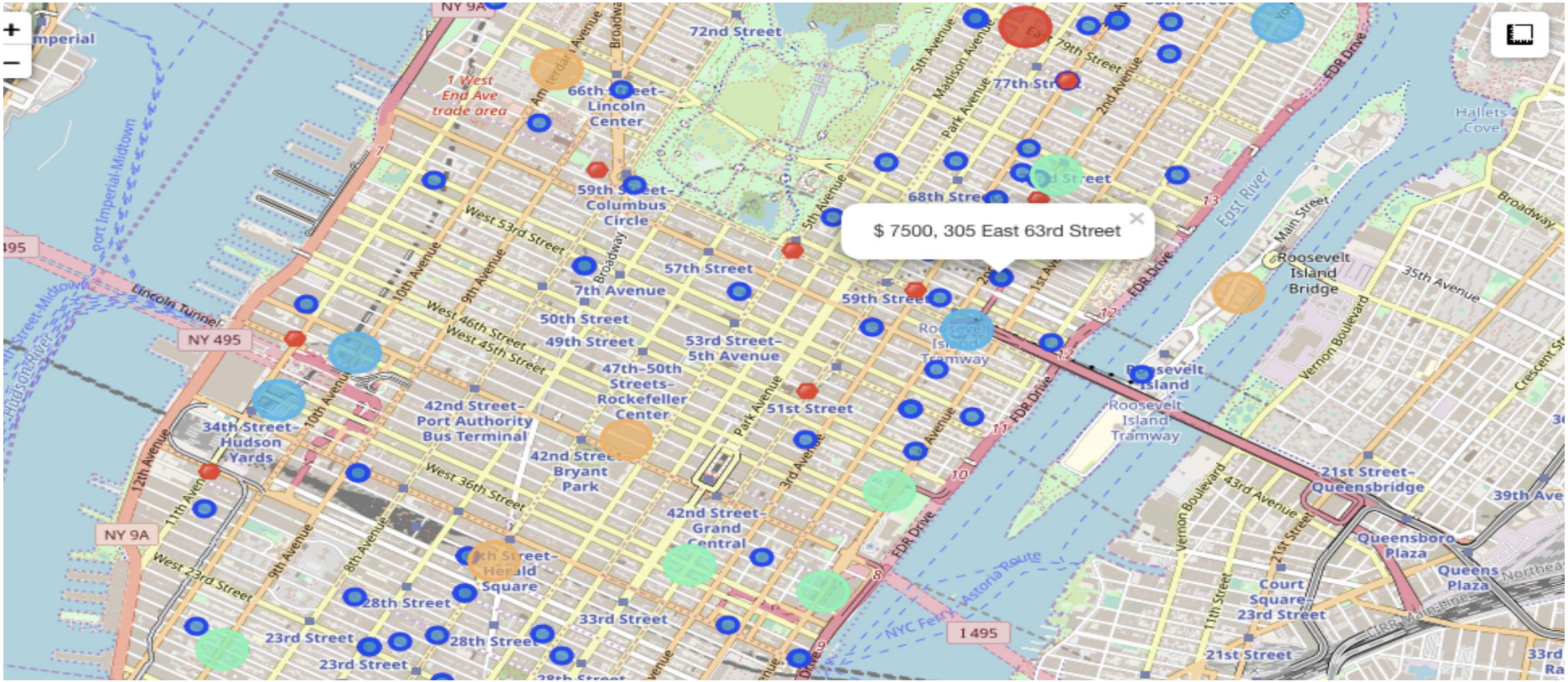
	sub_station	sub_address	lat	long
17	190 Street Subway Station	Bennett Ave, New York, NY 10040, USA	40.858113	-73.932983
18	59 St-Lexington Av Station	E 60th St, New York, NY 10065, USA	40.762259	-73.966271
19	57 Street Station	New York, NY 10019, United States	40.764250	-73.954525
20	14 Street / 8 Av	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 places for rent and the subway locations nearby



Map of Manhattan with rental places, subway locations and cluster of venues

Red dots are Subway stations, Blue dots are apartments available for rent, Bubbles are the clusters of venues



Apartment Selection

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

Apartment 1 rent cost is US7500 slightly above the US7000 budget. Apt 1 is located 400 meters from subway station at 59th Street and his work place (Park Ave and 53rd) is another 600 meters way. 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 he will have to ride the subway daily to work , possibly 40-60 min ride. Venues for this apt are as of Cluster 3.

5.0 DISCUSSION

I was able to explore all possibilities and find out 2 apartments:

- The apartment #1 rent cost is $7500 slightly above the $7000 budget. And apartment #2 rent cost is $6935, just under the $7000 budget.

- Apt #1 is located 400 meters from subway station at 59th Street, and Apt #2 is located 60 meters from subway station at Fulton Street.

- Apt #1 is located 600 meters from his workplace (Park Ave and 53rd). And Apt #2 far away from his workplace.

6.0 CONCLUSIONS

- Bouth Apt 1 and Apt 2 are wolking distance from the metro station.
- There's rent price difference between the Apt 1 and Apt 2 $565.
- The Apt 1 is walking distance from the client work and the Apt 2 is not.

If the client chose to rent Apt #1:

- he will spend ($565) monthly more than what he will pay for Apt 2.
- Walking distance from his work (he gonna save time and money spend on metro)

If the client chose to rent Apt #2:

- he will spend ($565) monthly less than what he will pay for Apt 1.
- Apt 2 far away from work (he gonna save time about (1h) and money about ($120) spend on metro).

End