



EXPLORING VENUES FOR APARTMENT IN MANHATTAN

Capstone Project - The Battle of Neighborhoods

Author

Prashant Joshi

Contents

1. Introduction
 - a. Background
 - b. Problem
 - c. Interest
2. Data acquisition and cleaning
 - a. Data Sources
 - b. Data Cleaning
3. Methodology
 - a. Exploratory Data Analysis
 - b. Modelling
4. Results
5. Discussion
6. Conclusion

1 Introduction

a. Background

In this era of globalization, many people tend to relocate, either to a city within the same country, or to a completely different country. Going to another country is always a whole new experience for anyone. People might not know about the locality they are going to, and might not rely on the locals. Getting a proper is a tedious task. Due to this, the people, especially the students and working-class people suffer a lot in finding a better place to rent which is affordable and nearer to several amenities. These things eventually becomes turning factors in searching the places for rent.

b. Problem

To overcome this, one of the several possible ways is to analyze the current location and target location, keeping in mind that the target location must have possible amenities within the proximity. To achieve this, considering Manhattan as the target location, I want to find the place to be rented, with following characteristics –

- Apartment having at least 2 bedrooms
- Monthly rent must not exceed UD\$7000/month
- Locality having similar amenities and venues as compared to current location
- Apartment having walking distance from a subway metro station

c. Interest

The person who may have desire to relocate to overseas cites like Manhattan might get benefit from this project. International students and the working people being shifted to Manhattan may find this project relevant and help themselves with finding a better apartment.

2. Data Acquisition and Cleaning

a. Data Sources

Given below is the list of data required in this project:

- Along with geodata, list of boroughs and neighborhoods of Manhattan
- Dataset of subway metro stations with their respective locations is scraped from [Wikipedia](#).
- Dataset of Apartments for rent containing address and price respectively were gathered from websites of [rantmanhattan.com](#), [nestpick.com](#) and [realtor.com](#).
- List of all neighborhoods of Manhattan to be used to cluster

b. Data Cleaning

A csv file was compiled with the rental place that indicated: areas of Manhattan, address, number of beds, area and monthly rental price. The algorithm was used to create all the geodata using Nominatim. With the use of geolocator = Nominatim(), it was possible to determine the latitude and longitude for the subway metro locations as well as for the geodata for each place listed. “Great Circle” function from geolocator was used to calculate distance between two points, as in the case to calculate average rent price for units around each subway station and at 1.6 km radius. Foursquare is used to find the avenues at Manhattan neighborhoods in general and a cluster is created to be able to search for the venues depending on the location shown.

Some illustrations of Datasets used to collect the required data.

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Fig: List of Manhattan Neighborhoods

	Address		Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
0	West 105th Street	Upper West Side		2.94	5	3400	10000	NaN	NaN
1	East 97th Street	Upper East Side		3.57	3	2100	7500	NaN	NaN
2	West 105th Street	Upper West Side		1.89	4	2800	5300	NaN	NaN
3	CARMINE ST.	West Village		3.03	2	1650	5000	NaN	NaN
4	171 W 23RD ST.	Chelsea		3.45	2	1450	5000	NaN	NaN

Fig: Manhattan Flats Price

	sub_station	sub_address	lat	long
71	50 St Br	New York, NY 10019, United States	2.624409	-1.333834
72	57 Street Station	New York, NY 10019, United States	-1.013137	0.013050
73	23 Street Station	New York, NY 10010, United States	-0.759999	-0.245240
74	14 Street / 8 Av	New York, NY 10014, United States	-0.477321	0.185143
75	MTA New York City	525 11th Ave, New York, NY 10018, USA	-0.422317	0.747684

Fig: Manhattan Subway Station list

3 Methodology

This section contains the main component of the report where the data is gathered and prepared for analysis. The tools described are used here and the Notebook cells indicates the execution of steps.

a. Exploratory Data Analysis

The strategy is based on mapping the above described data in section 2.0, in order to facilitate the choice of at least two candidate places for rent. Choice is made based on the demands imposed: location near a subway, rental price and similar venues to Singapore. This visual approach and maps with popups labels allow quick identification of location, price, and feature, thus making the selection very easy.

The processing of these DATA and its mapping will allow to answer the key questions to make a decision:

- What is the cost of available rental places that meet the demands?
- What is the cost of rent around a mile radius from each subway metro station?
- What is the area of Manhattan with best rental pricing that meets criteria established?
- What is the distance from work place (Park Ave and 53rd St) and the tentative future rental home?
- What are the venues of the two best places to live? How the prices compare?
- How venues distribute among Manhattan neighborhoods and around metro stations?
- Are there trade-offs between size and price and location?
- Any other interesting statistical data findings of the real estate and overall data.

b. Modelling

Manhattan Neighborhoods – Data and Mapping

Cluster neighborhood data was produced with Foursquare during course lab work. A csv file was produced containing the neighborhoods around the 40 Boroughs. Now, the csv file is just read for convenience and consolidation of report.

Examine a particular Cluster - print venues. After examining several cluster data, I concluded that cluster # 2 resembles closer the Singapore place, therefore providing guidance as to where to look for the future apartment. Assign a value to 'kk' to explore a given cluster.

Places for rent in Manhattan

Several Manhattan real estate webs were web scrapped to collect rental data, as mentioned in section 2.0. The result was summarized in a csv file for direct reading, in order to consolidate the process.

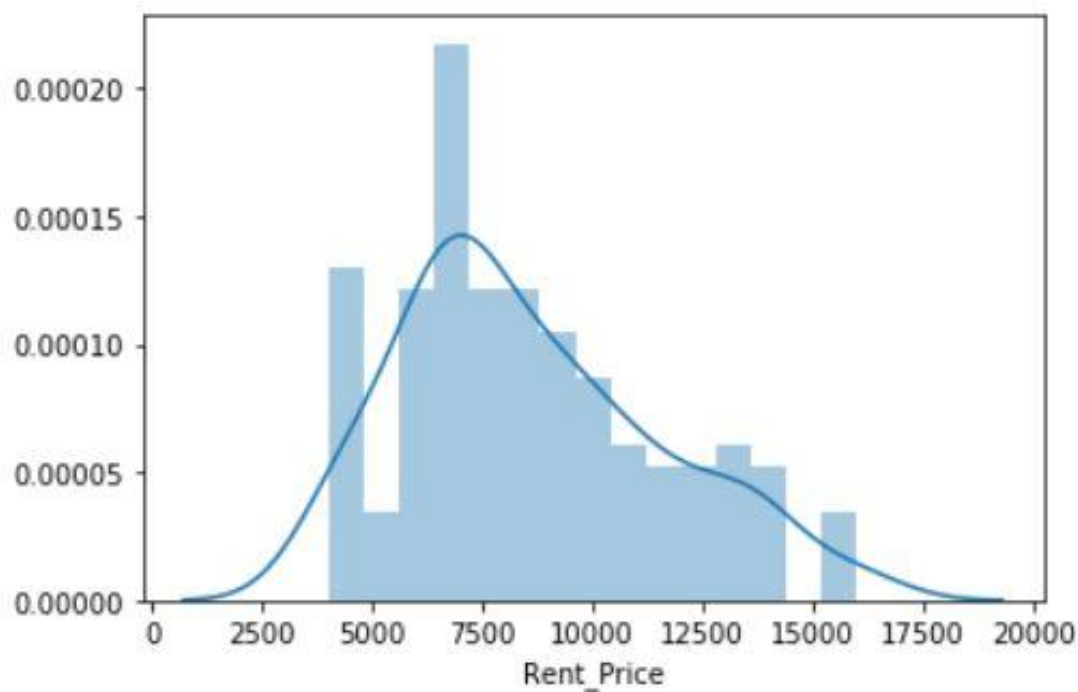
The initial data for 144 apartments did not have the latitude and longitude data (NaN) but the information was established in the following cell using an algorithm and Nominatim.

Obtain geodata (Lat, long) for each rental place in Manhattan with Nominatim Data was stored in a csv file for simplification report purposes and saving code processing time in future.

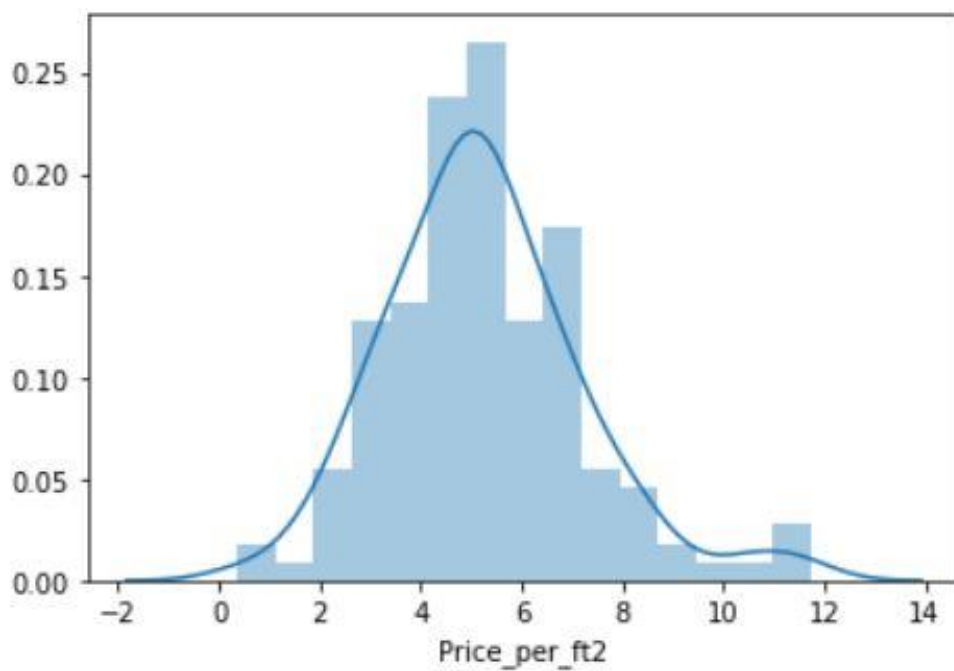
Manhattan apartment rent price statistics

A US 7000 Dollar per month rent is actually around the mean value - similar to Singapore! wow!


```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc769824ac8>
```

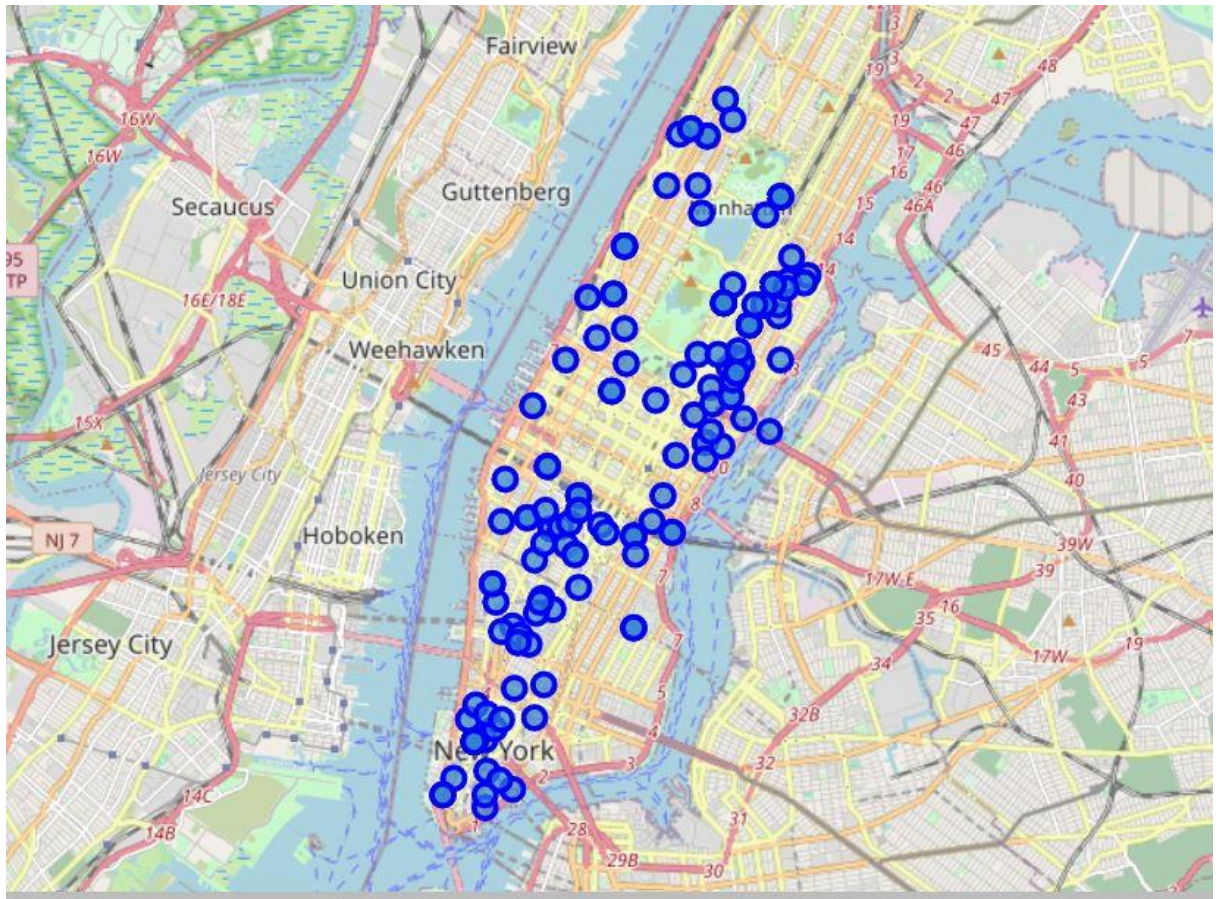


```
<matplotlib.axes._subplots.AxesSubplot at 0x7fc7697a6160>
```

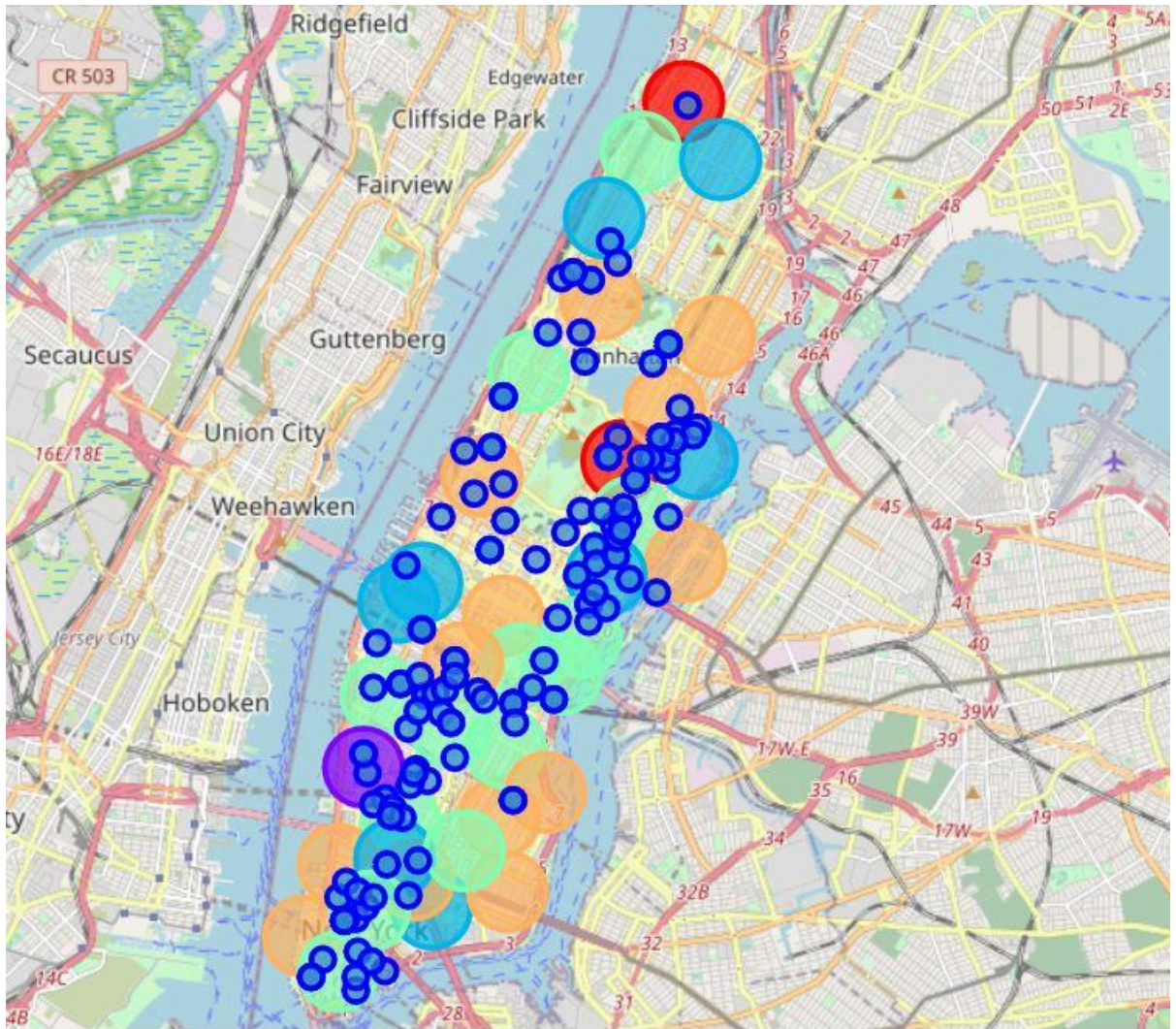


Map – Manhattan apartments for rent

The popups will indicate the address and the monthly price for rent thus making it convenient to select the target apartment with the price condition stipulated (max US7000).



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



Now one can explore a particular rental place and its venues in detail

In the map above, examination of apartments with rental place below 7000/month is straightforward while knowing the venues around it.

We could find an apartment with at the right price and in a location with desirable venues. The next step is to see if it is located near a subway metro station, in next cells work.

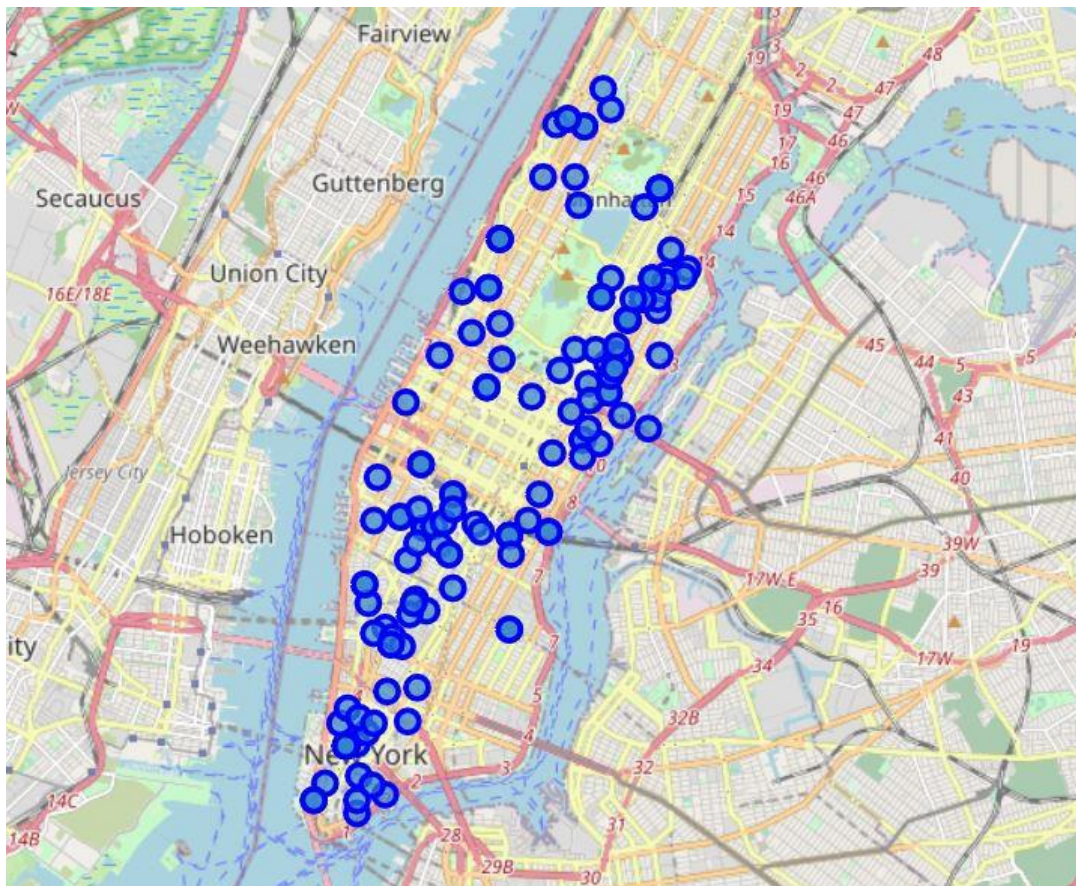
Mapping Manhattan Subway locations

Manhattan subway metro locations (address) was obtained from web scrapping sites such as Wikipedia, Google and NY Metro Transit. For simplification, a csv file was produced from the 'numbers' (Apple excel) so that the reading of this file is the starting point here.

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 display rental address and monthly rental price and the subway station name.

Notice that the icon in the top-right corner is a "ruler" that allows to measure the distance from a rental place to a specific subway station

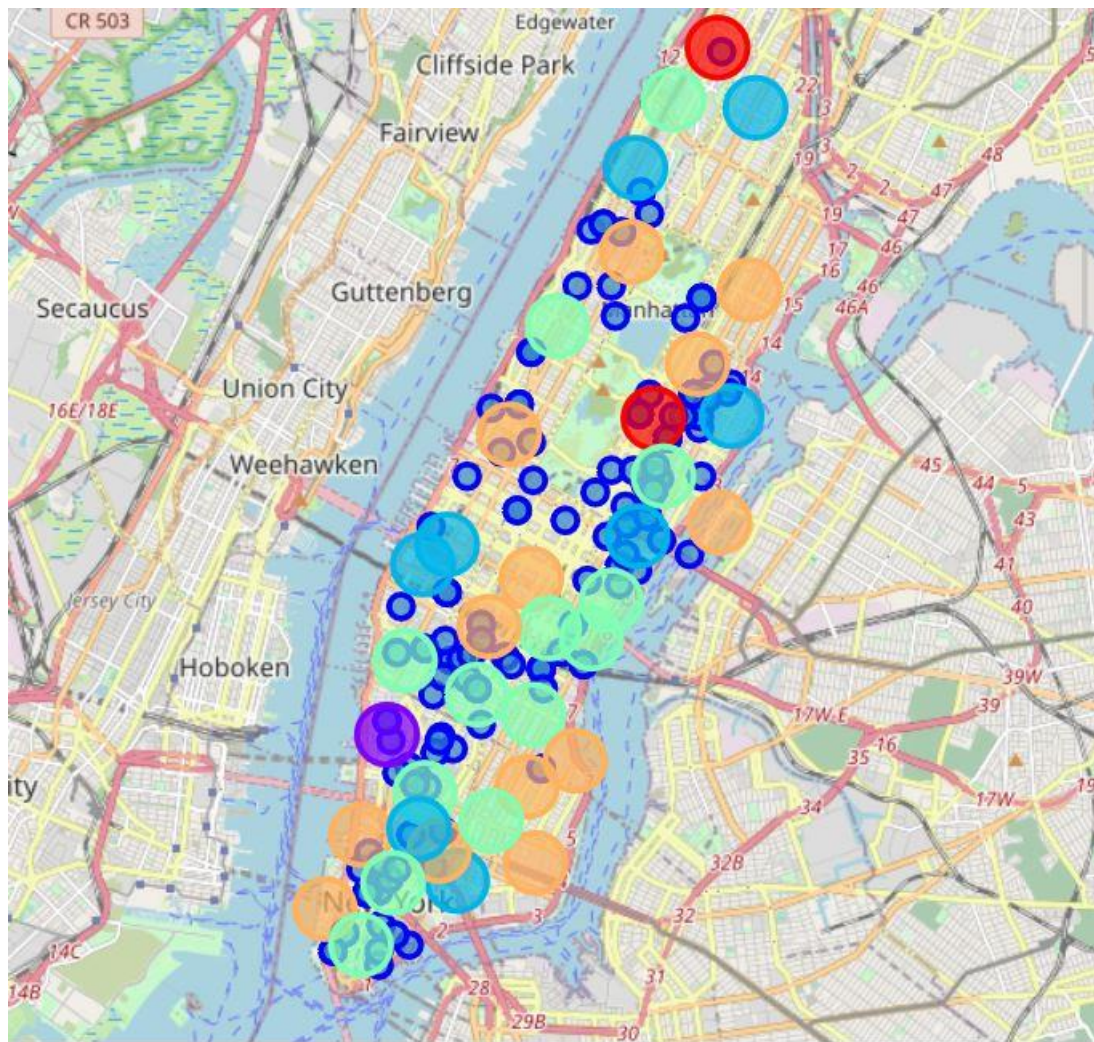


4 Results

Single Consolidated Map

Let's consolidate all the required information to make the apartment selection in one map

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.



Problem Resolution - Select the apartment for rent

After examining, I have chosen two locations that meet the requirements which will assess to make a choice.

Apartment 1: 305 East 63rd Street in the Sutton Place Neighbourhood and near 'subway 59th Street' station, Cluster # 2 Monthly rent: 7500 Dollars

Apartment 2: 19 Dutch Street in the Financial District Neighbourhood and near 'Fulton Street Subway' station, Cluster # 3 Monthly rent: 6935 Dollars

Venues for Apartment 1 - Cluster 2

[39]:

	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
0	Marble Hill	Coffee Shop	Discount Store	Yoga Studio	Steakhouse	Supplement Shop	Tennis Stadium	Shoe Store	Gym	Bank	Seafood Restaurant
1	Chinatown	Chinese Restaurant	Cocktail Bar	Dim Sum Restaurant	American Restaurant	Vietnamese Restaurant	Salon / Barbershop	Noodle House	Bakery	Bubble Tea Shop	Ice Cream Shop
6	Central Harlem	African Restaurant	Seafood Restaurant	French Restaurant	American Restaurant	Cosmetics Shop	Chinese Restaurant	Event Space	Liquor Store	Beer Bar	Gym / Fitness Center
9	Yorkville	Coffee Shop	Gym	Bar	Italian Restaurant	Sushi Restaurant	Pizza Place	Mexican Restaurant	Deli / Bodega	Japanese Restaurant	Pub
14	Clinton	Theater	Italian Restaurant	Coffee Shop	American Restaurant	Gym / Fitness Center	Hotel	Wine Shop	Spa	Gym	Indie Theater
23	Soho	Clothing Store	Boutique	Women's Store	Shoe Store	Men's Store	Furniture / Home Store	Italian Restaurant	Mediterranean Restaurant	Art Gallery	Design Studio
26	Morningside Heights	Coffee Shop	American Restaurant	Park	Bookstore	Pizza Place	Sandwich Place	Burger Joint	Café	Deli / Bodega	Tennis Court
34	Sutton Place	Gym / Fitness Center	Italian Restaurant	Furniture / Home Store	Indian Restaurant	Dessert Shop	American Restaurant	Bakery	Juice Bar	Boutique	Sushi Restaurant
39	Hudson Yards	Coffee Shop	Italian Restaurant	Hotel	Theater	American Restaurant	Café	Gym / Fitness Center	Thai Restaurant	Restaurant	Gym

Venues for Apartment 2 - Cluster 3

	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

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\$7500 slightly above the US\$7000 budget. Apt 1 is located 400 meters from subway station at 59th Street and work place (Park Ave and 53rd) is another 600 meters way. I can walk to work place and 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 US\$6935, just under the US\$7000 budget. Apt 2 is located 60 meters from subway station at Fulton Street, but I will have to ride the subway daily to work, possibly 40-60 min ride. Venues for this apt are as of Cluster 3.

Based on current Singapore venues, I feel that Cluster 2 type of venues is a closer resemblance to my current place. That means that APARTMENT 1 is a

better choice since the extra monthly rent is worth the conveniences it provides.

5 Discussion

- In general, I am positively impressed with the overall organization, content and lab works presented during the Coursera IBM Certification Course
- I feel this Capstone project presented me a great opportunity to practice and apply the Data Science tools and methodologies learned.
- I have created a good project that I can present as an example to show my potential.
- I feel I have acquired a good starting point to become a professional Data Scientist and I will continue exploring to creating examples of practical cases.

6 Conclusion

- I feel rewarded with the efforts, time and money spent. I believe this course with all the topics covered is well worthy of appreciation.
- This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.
- The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision thoroughly and with confidence. I would recommend for use in similar situations.
- One must keep abreast of new tools for DS that continue to appear for application in several business fields.