Finding The Best Apartment to Rent in Manhattan N.Y

Gaby Bou Tayeh

March 2020

1 Introduction

1.1 Motivation and background

Backgroud: I am a third-year Ph.D. student currently living in Paris. Paris is a living city with a lot of diverse venues, in addition, the public transportation system in Paris is very efficient and well managed, It is very easy to go from point A to point B regardless of the location of these two points. However, renting an apartment in Paris is very expensive, and it varies highly according to the boroughs.

Motivation: I have been offered a great opportunity to work in Manhattan, NY. Although I am very excited about it, I want to make sure to find a fairly similar neighborhood to the one I am currently living at. I am interested in venues, metro station proximity, and apartment prices. On a personal level, I believe that this project is a good practice towards the development of my Data Science skills.

2 Objective:

My objective is to to find an apartment to rent in Manhattan NY under certain conditions:

- The apartment rental price should be within my 7000-8000 US $ Budget.

- The Neighborhood where the apartment is located should have roughly similar venues to the one I am currently living in.

- The apartment should not be too far from the nearest metro station (within 500m).

3   Targeted audience:

I believe that this is a relevant project for a person or an entity considering moving to a major city in Europe, the US or Asia. The approach and the used methodology could be adapted and applied to cities other than Manhattan. The use of FourSquare data and mapping techniques combined with data analysis will help us decide which apartment to rent.

4   Description of the used data

In order to make a final decision on which apartment to rent in Manhattan NY, the following data is required:

- Information on Manhattan's neighborhoods alongside the coordinates of each neighborhood(latitude, longitude).

- Information on the subway metro stations in Manhattan (location, coordinates).

- a List of apartments that are published for rent in the Manhattan area with descriptions (price, location, address, etc..)

- List of venues in the Manhattan neighborhoods (e.g. top 10)

```
mh_neigh_data.tail():

        Borough    Neighborhood    Latitude    Longitude

    35  Manhattan    Turtle Bay    40.752042    -73.967708

    36  Manhattan    Tudor City    40.746917    -73.971219

    37  Manhattan    Stuyvesant Town    40.731000    -73.974052

    38  Manhattan    Flatiron    40.739673    -73.990947

    39  Manhattan    Hudson Yards    40.756658    -74.000111
```

Figure 1: mh_neigh_data.csv

4.1   Data sources

The list of Manhattan neighborhoods is worked out during Lab exercise of the applied data science course (a CSV file containing the final results was saved locally which will be retrieved later and read into a Pandas Dataframe). a part of the csv file 'mh_neigh_data.csv' content is illustrated in Figure 1.

A list of Manhattan subway metro stations was prepared and it was completed by scraping data from this Wikipedia page [1]. Additional information

was also scrapped from NY Transit authority and Google maps [2]. The geolocation of the station was obtained using "Nominatim". Details will be shown later in the Notebook. The subway CSV file is "MH_subway.csv"', Figure 2 below illustrates a part of it:

mhsub.tail(): sub_station sub_address lat long

	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

Figure 2: MH_subway.csv

A list of apartments for rent was collected by web-browsing real estate companies in Manhattan [3, 4, 5]. A CSV file ("MH_flats_prices.csv") was compiled indicating: location of the apartment (Area), address, number of beds, and monthly rental price. Nominatim is used again to retrieve the coordinates (lat, long) of each apartment. Figure 3 shows a snapshot of the data related to the apartments.

	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

Figure 3: MH_flats_prices.csv

4.2 How the data can serve our objective?

The Nominatim (from geopy) and the Foursquare API will be used to retrieve the coordinates of the clustered neighborhoods, apartments, metro stations and display them on an interactive map using Folium. Information about the name of each neighborhood and borough, the price of each apartment, and the distance between apartments and their nearest metro station will be easily accessible using the interactive map. All this information will help us decide which apartment to rent.

5 Methodology

In this section, we will present an exploratory analysis of the previously described data, we will briefly explain the data collection phase, pre-processing, and visualization. For more detailed information the interested readers can refer to the final Notebook.

5.1 Exploring Manhattan's neighborhoods

A CSV file of the clustered neighborhoods from the previous Notebooks (Part3) was saved locally. Now, the CSV file is just read for convenience and consolidation of the report. A portion of the content of the file is illustrated in Figure 4.

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

Figure 4: Manhattan's neighborhoods DataFrame

Also from the last lab exercise, I have stored the clustered neighborhoods alongside the top 10 venues in each one of them in a CVS file. Figure 5 shows a portion of its content.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	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 Com V
0	Manhattan	Marble Hill	40.876551	-73.910660	2	Coffee Shop	Discount Store	Yoga Studio	Steakhouse	Supplement Shop	Tennis Stadium	Shoe Store	Gym	
1	Manhattan	Chinatown	40.715618	-73.994279	2	Chinese Restaurant	Cocktail Bar	Dim Sum Restaurant	American Restaurant	Vietnamese Restaurant	Salon / Barbershop	Noodle House	Bakery	B Tea
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Bakery	Mobile Phone Shop	Pizza Place	Sandwich Place	Park	Gym	Latin American Restaurant	Resta
3	Manhattan	Inwood	40.867684	-73.921210	3	Mexican Restaurant	Lounge	Pizza Place	Café	Wine Bar	Bakery	American Restaurant	Park	F Y
4	Manhattan	Hamilton Heights	40.823604	-73.949688	0	Mexican Restaurant	Coffee Shop	Café	Deli / Bodega	Pizza Place	Liquor Store	Indian Restaurant	Sushi Restaurant	San

Figure 5: Apartments DataFrame

5.2 Apartment rental prices and geo-location

Now after we have all the necessary information for each neighborhood, let's load the apartment prices data. Figure 3 shows a portion of the content of the loaded CSV containing information about apartments that are currently posted for rental. As we can see in the figure, we have the address of each apartment.

During the lab exercises, they provided us with a method that enables us to extract the latitude and longitude of a specific address. This method illustrated below is used to extract the geo-location of each apartment. Figure 6 show the final Dataframe after adding the lat and the long.

$geolocator = Nominatim()$

$location = geolocator.geocode(address)$

	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

Figure 6: apartments information

5.3 Manhattan's subway stations

For the subway station, we had their names and addresses, and we needed to get their latitude and longitude. Using the same method used previously for apartments, we were able to get the latitude and longitude of every subway station. Figure 7 shows the final DataFrame.

	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

Figure 7: Subway Stations DataFrame

5.4 Visualisation on a map

Now after having all the information that we needed, let's visualize them on a Map. Figure 8 show the neighborhood clusters (big red-colored circles), the metro stations (small red dots), and the apartments (small blue dots). A ruler tool is available at the top right corner, the user can use this tool to measure distances on the map.

Selecting the best apartment to rent: The consolidate map will be used to explore our options. Let's say I am interested in living in neighborhoods belonging to Cluster 2 (close to the one I currently live in Paris). And my budget is limited to 7000-8000 US Dollars a month.
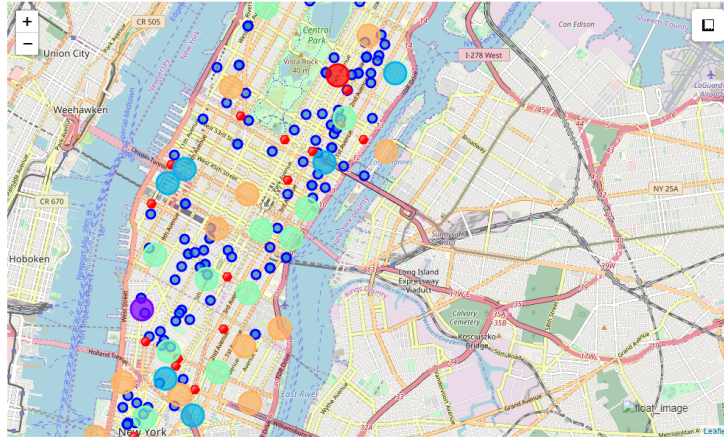
5

Figure 8: Final Map

Here are 3 options (there are obviously more, I will just give 3 examples)

1. Neighborhood: Soho, apartment address: 93 Grand Street, Price: 6900, distance to nearest subway station: 461 m.

2. Neighborhood: Sutton Place, apartment address: 221 East Str60th Street, Price: 8000, distance to nearest subway station: 106 m.

3. Neighborhood: Morning side Heights, apartment address: 1080 Amsterdam Avenue, Price: 7792, distance to nearest subway station: 2297 m.

Option 3 is not to be considered! too far away from a subway station (time is money). Am I willing to pay 1100 US dollars more to live in an apartment that is closer to a subway station (option 2)? I don't think it worth it, walking a 300m hundred meters (roughly 4 minutes) to save 1100 dollars is a good deal. Therefore I will stick to option 1

6   Discussion

In order to improve our decision about which apartment to rent, it is possible to explore other information. here's a list of a few features that could be taken into consideration

- Population density: What if the person looking for an apartment prefers a calm neighborhood on a noisy dense one?

- Distance to work location: The metro station location is important, but what if we found two similar apartments, one is 100m away from the nearest metro station, and the other is 500m away. However, from the former, it takes 40min to arrive at work and from the later, it only takes a

few minutes. The apartment that is slightly farther from its nearest metro station would be a better choice.

- In the previous lab, in this specialization, we analyzed the crime rate in Manhattan N.Y, we could also use this to alert newcomers to the city of certain neighborhoods that are not very safe.

7   Results

In this report, I present a comprehensible easy to use tool that enables a person moving to Manhattan to find similar neighborhoods to the one he is currently living in with the right rental price. For future work, in order to improve this tool, more features could be added, such as population density, crime rates in each neighborhood, and distance to work form each apartment.

References

[1] "Wikipedia - Subway Stations." https://en.wikipedia.org/wiki/List_of_New_York_City_Subway _stations_in$_M$anhattan. $Accessed : 2010 - 01 - 03.$

[2] "Google Maps - Subway Stations." https://www.google.com/maps/search/manhattan+subway +metro+stations/@40.7837297,-74.1033043,11z/data=!3m1!4b1. Accessed: 2020-01-03.

[3] "Rentmanhattan." http://www.rentmanhattan.com/index.cfm?page=search& state=results. Accessed: 2020-01-03.

[4] "Realtor." https://www.realtor.com/apartments/Manhattan_NY. Accessed: 2020-01-03.

[5] "Nestpick." https://www.nestpick.com/search?city=new-york&page=1& order=relevance&district=manhattan&gcid=CjwKCAiAjNjgBRAgEiwAGLlf2hkP3A- cPxjZYkURqQEswQK2jKQEpv_ MvKcrIhRWRzNkc_r- fGi0lxoCA7cQAvD_BwE&type=apartment&display=list. Accessed: 2020-01-03.