COURSERA CAPSTONE PROJECT

The Battle of the Neighborhood
(Week - 2)

CONTENT:

- Introduction
- Data
- Methodology
- Results (Screenshots)
- Conclusion

INTRODUCTION

1.1 Scenario and Background

I am currently living in Singapore, within walking distance to Downtown "Telok Ayer MRT metro station" . I also enjoy great venues and attractions, such as international cuisine, entertainment and shopping. I have an offer to move to work to Manhattan NY and I would like to move if I can find a place to live similar with similar venues.

1.2 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 my current residence.

1.3 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. Lastly, this project is a good practical case for a person developing Data Science skills.

DATA SECTION

2.1 Data Requirements

- Geodata for 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 Nominator.
- 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

2.2 Data Sources, Data Processing and Tools used

- Singapore data and map is to be created with use of Nominator , Foursquare and Folium mapping.
- Manhattan neighborhoods were obtained from Wikipedia and organized by Neighborhoods with geodata via Nominator 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 Nominator.
- 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

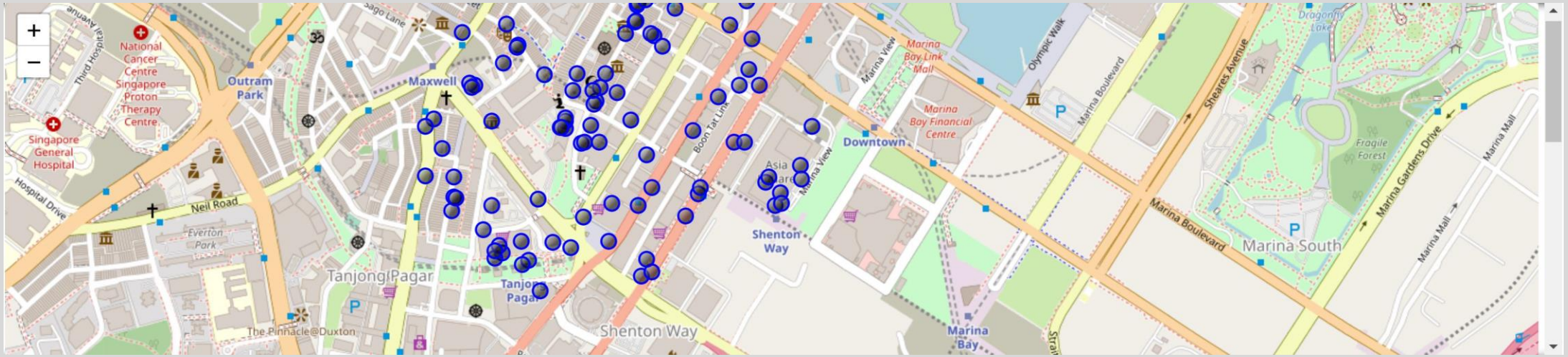
METHODOLOGY

The Strategy to find the answer:

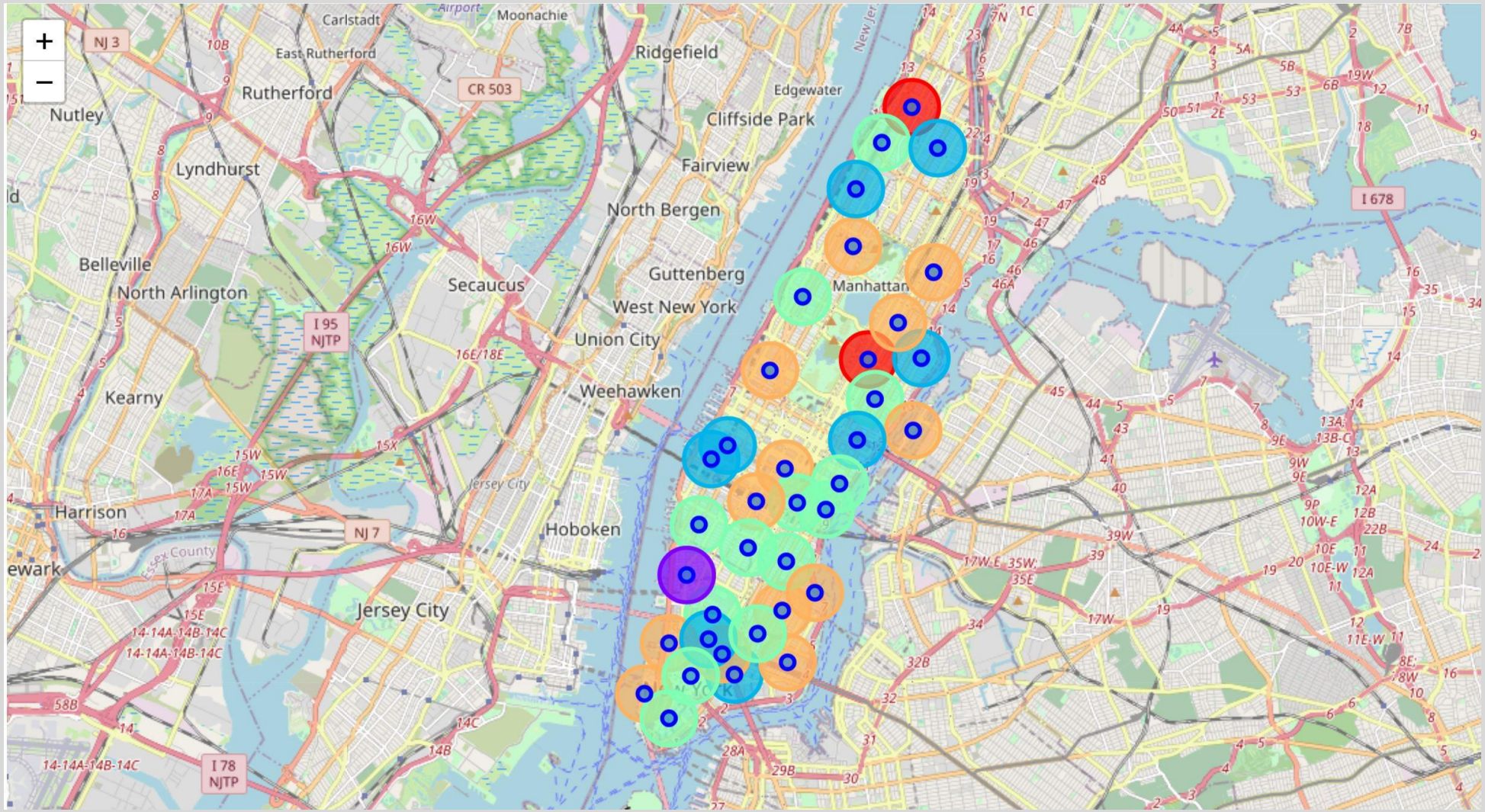
- The strategy is based on mapping the described data in section 2.0, in order to facilitate the choice of at least two candidate places for rent.
- The information will be consolidated in ONE MAP where one can see the details of the apartment, the cluster of venues in the neighborhood and the relative location from a subway station and from work place.
- A measurement tool icon will also be provided.
- The 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 Nomination to get latitude and longitude of subway stations and also for each of (144 units) the apartments for rent listed.
- Geopy_distance and Nomination were used to establish relative distances. Seaborn graphic was used for general statistics on rental data.
- Maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy.

RESULTS (SCREENSHOTS)

4.1 Current residence Neighborhood in Singapore



4.2 Manhattan Map - Neighborhoods and Cluster of Venues



4.3 GeoData Manhattan apts for rent

```
mh_rent=pd.read_csv('MH_rent_latlong.csv')
mh_rent.head()
```

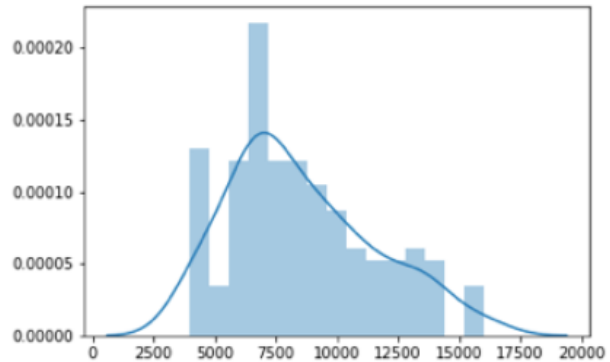
	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
0	West 105th Street	Upper West Side	2.94	5.0	3400	10000	40.799771	-73.966213
1	East 97th Street	Upper East Side	3.57	3.0	2100	7500	40.788585	-73.955277
2	West 105th Street	Upper West Side	1.89	4.0	2800	5300	40.799771	-73.966213
3	CARMINE ST.	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()
```

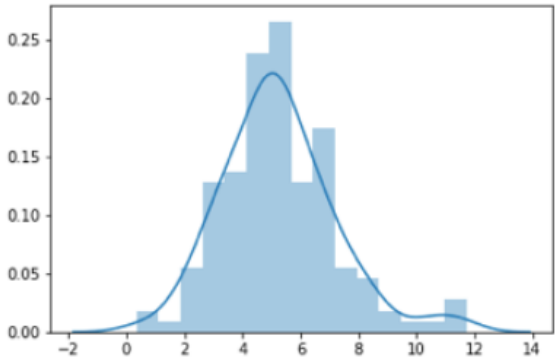
	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
139	200 East 72nd Street	Rental in Lenox Hill	5.15	3.0	1700	8750	40.769465	-73.960339
140	50 Murray Street	No fee rental in Tribeca	7.11	2.0	1223	8700	40.714051	-74.009608
141	300 East 56th Street	No fee rental in Midtown East	3.87	3.0	2100	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

4.4 Rental Price Statistics MH Apartments

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

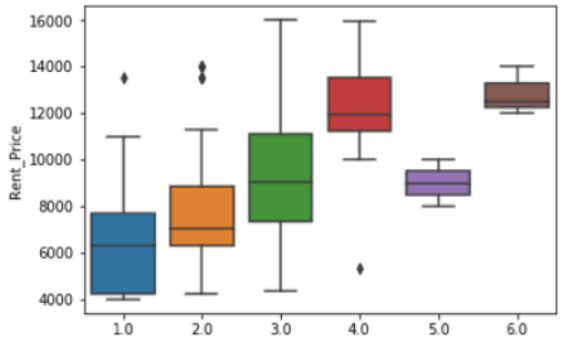



```
<matplotlib.axes._subplots.AxesSubplot at 0x1a2415fc18>
```

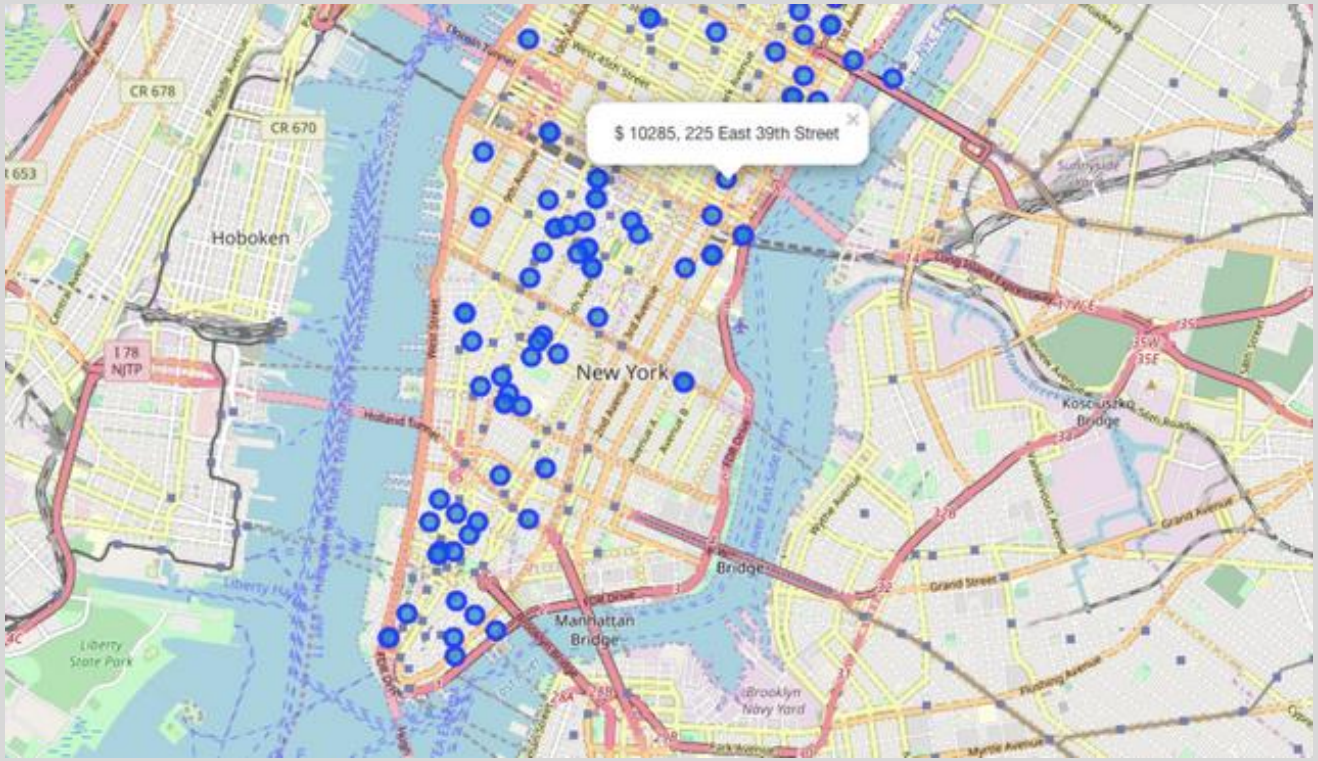



```
sns.boxplot(x='Rooms', y= 'Rent_Price', data=mh_rent)
```

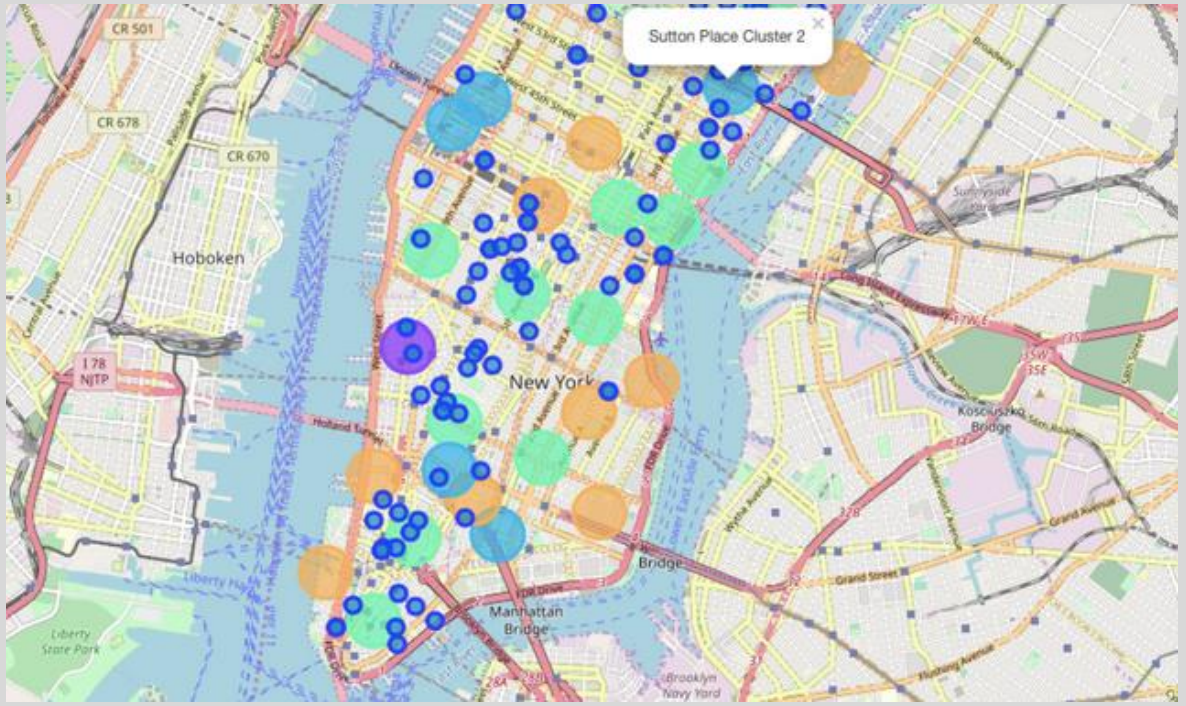
```
<matplotlib.axes._subplots.AxesSubplot at 0x1a25f2a2b0>
```



4.5 Apartments for Rent in MH



4.6 MH apt for rent with venue clusters



4.7 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

4.8 Manhattan subway stations 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

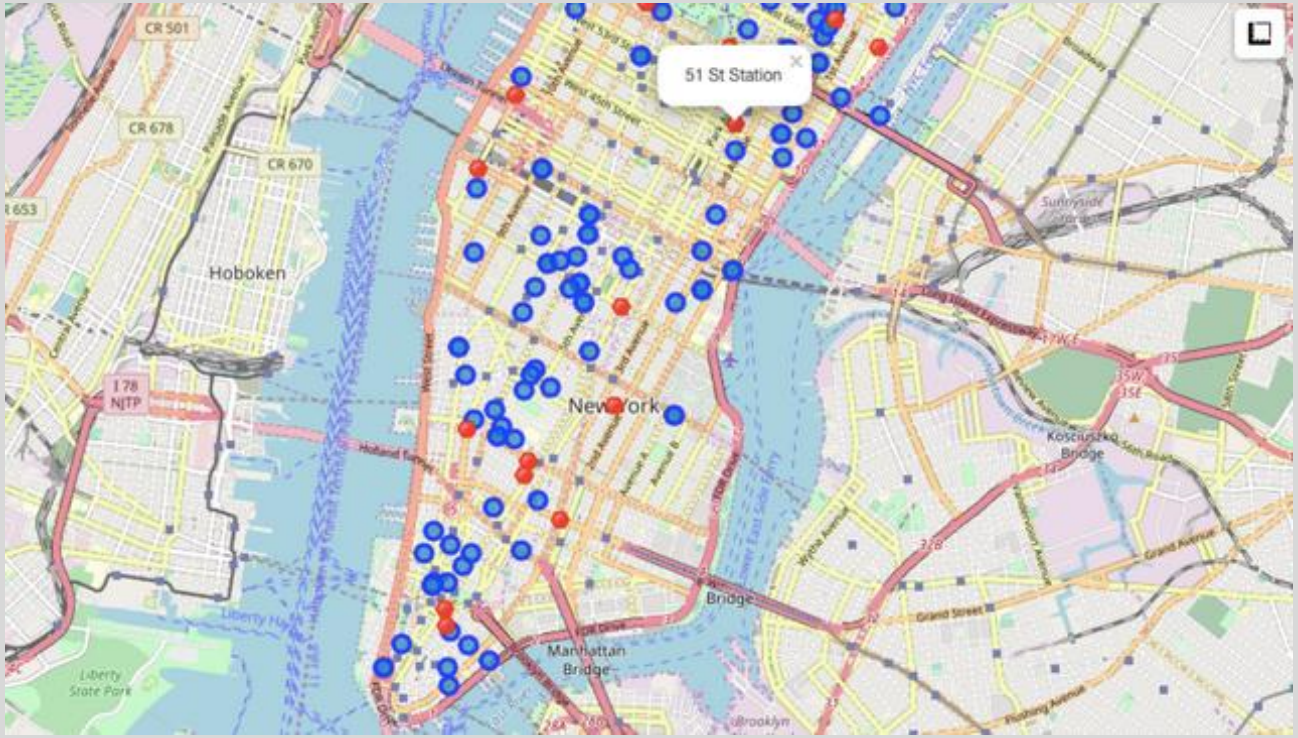
```
# removing duplicate rows and creating new set mhsub1
mhsub1=mh.drop_duplicates(subset=['lat','long'], keep="last").reset_index(drop=True)
mhsub1.shape
```

```
(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

4.9 Apts for rent (blue) and subway stations (red)



CONCLUSION

• I feel rewarded with the efforts, time and cash spent. I believe this course with all the topics lined is well worthy of appreciation.

• This project has shown Pine Tree State an employment to resolve a real scenario that has impacting personal and money impact mistreatment knowledge Science tools.

• The mapping with geological formation could be a terribly powerful technique to consolidate data and create the analysis and call thoroughly and confidently. I'd suggest for use in similar things.

• One should keep up with recent tools for DS that continue to appear for application in many business fields.


thank you
gracias
merci
danke
謝謝
teşekkür ederim
ngiyabonga
obrigado
bedankt
dziękuję
спасибо
go raibh maith agat
sukriya
kop khun krap
terima kasih
감사합니다
mochchakkeram
tapadh leat
dank je
grazie
arigatô
takk
dakujem
мерси
ευχαριστώ
ありがとう
рахмат
Баярлалаа
mersi
kia ora
barka
welalin
tack
spas
faafetai lava
vinaka
blagodaram
misaotra
matondo
paldies
grazzi
mahalo
хвала
asante
manana
obrigada
tenki
murakoze
chokrane
mamnun
дякую
djiere dieuf
tau
sulpáy
gràcies
gratias ago
chnorakaloutioun
taiku
akun
dankon
köszönöm
mauruuru
hvala
gràcie
bayarlalaa
dhanyavad
kiitos
dankie
nandri
nanni
спасибі
enkosi
sobodi
dekuji
mèsi
sagolun
najis tuke
kam sah hamnida
rahmat
তোমাকে ধন্যবাদ
tanemirt
rahmet
diolch
xièxie
dhanyavadagalu
shukriya
merce
trugarez