Neighbourhood Comparison

1. Problem Introduction:

1.1 Problem to be resolved

The goal of this project is to compare the neighbourhoods of two cities and determine how similar or dissimilar they are using the skills learned during the IBM Data Science course, and in particular using the tools to explore location data and geographical location.
More specifically, in this project I will be looking at how similar or dissimilar are Downtown Singapore and Manhattan, with the goal to find a rental apartment in Manhattan which meets the following demands:

- apartment must be 2 or 3 bedrooms
- desired location is near a metro station in the Manhattan area and within 1.5 mile radius
- rent price should not exceed 8.000$ per month
- top amenities in the selected neighbourhood
- desirable to have venues such as coffee shops, restaurants, wine stores, gym and food shops

1.3 Interested Audience

Anyone who considers moving to a major city, since the approach and methodologies used here are applicable in all cases.
Using Foursquare data an mapping techniques together with data analysis and data visualisation it will be possible to solve the problem.

2. Data Description

2.1. Required Data

Using Foursquare I can find the venues in a predefined area in Singapore.
In order to make a good choice for an apartment in Manhattan, the following data is required:

- List of neighbourhoods in Manhattan with their Geodata (latitude and longitude)
- List of the subway metro stations in Manhattan with geodata
- List of apartments for rent in Manhattan with descriptions (number of beds, price, location, address)
- Venues and amenities in the neighbourhoods

A list of Manhattan neighbourhoods is already available from previous LABs.
A list of the Manhattan's subways can be found on the following links:
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
For the list of apartments the following link can be used:
http://www.rentmanhattan.com/index.cfm?page=search&state=results & https://www.realtor.com/apartments/Manhattan_NY

A csv file can be compiled indicating the: area of Manhattan, address, number of beds, sq ft and monthly rental price.
Using the Geolocator (Nominatig) I can obtain the geolocation of the subways and of the apartments (with the latitude and the longitude)

This how the data of the Manhattan neighbourhood looks like:

	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

This is how the subway station data with geolocation looks like

	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

This is how the collected rental data will look like (latitude and longitude just need to be added)

	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

2.2. Data utilization

Using Foursquare and geopy data I can map the top 10 venues for all Manhattan neighbourhoods and cluster them in groups. Using foursquare and geopy I can map the location of subway metro stations , separately and on top of the above clustered map in order to be able to identify the venues and amenities near each metro station. I can also map the location of rental places linked to the subway locations. I could create a map that depicts, for instance, the average rental price per square ft, around a radius of 1.0 mile around each subway station. I will be able to quickly point to the popups to know the relative price per subway area. Addresses from rental locations will be converted to geodata(lat, long) using Geopy-distance and Nominatim.

3. Methodology

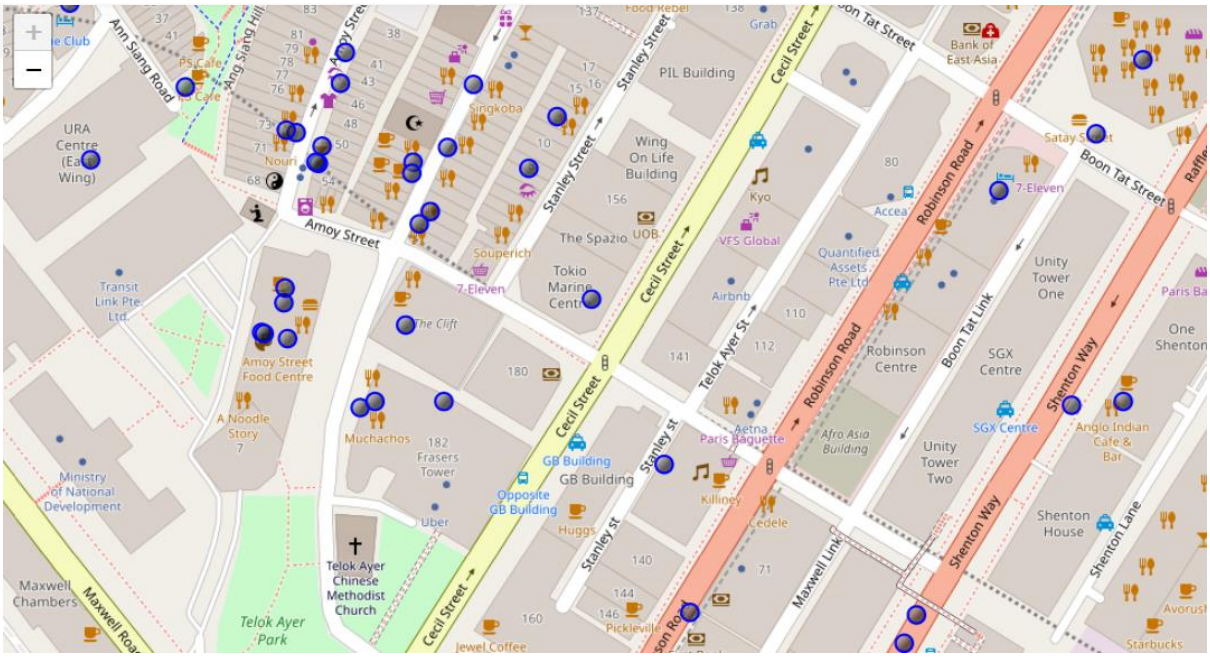
The methodology is based on the mapping of the data in section 2 in different ways, thus allowing throughout the use of a visual approach, i.e. by means of maps with popups labels, a quick identification of apartments according to the demands.

The processing of the DATA and its mapping, will allow to answer the key questions such as:

- 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 are the venues of the two best places to live? How do the prices compare?
- How venues distribute among Manhattan neighbourhoods and around metro stations?
- Any other interesting statistical data findings of the real estate and overall data.

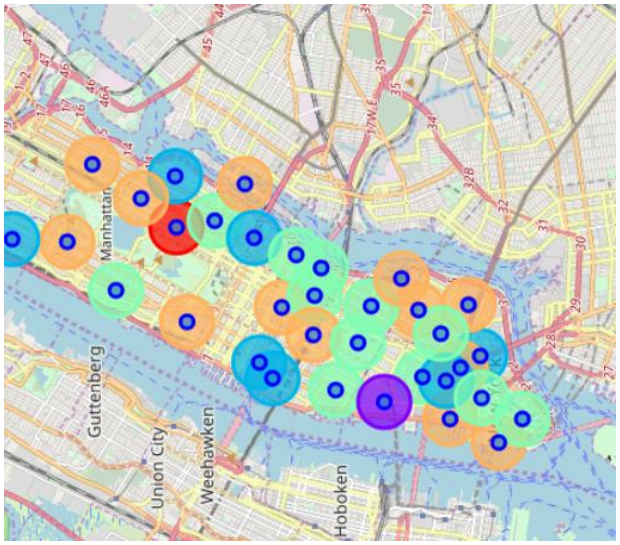
Let's first map the venues of the neighbourhood in Singapore in order to find similar neighbourhood in Manhattan.
The first 10 are showed below:

	name	categories	lat	lng
0	Napoleon Food & Wine Bar	Wine Bar	1.279925	103.847333
1	Pepper Bowl	Asian Restaurant	1.279371	103.846710
2	Native	Cocktail Bar	1.280135	103.846844
3	Park Bench Deli	Deli / Bodega	1.279872	103.847287
4	Mellower Coffee	Café	1.277814	103.848188
5	Muchachos	Burrito Place	1.279072	103.847026
6	Dumpling Darlings	Dumpling Restaurant	1.280483	103.846942
7	Sofitel So Singapore	Hotel	1.280017	103.849813
8				
9				



Below are showed the Manhattan neighbourhood with the top 10 clustered venues

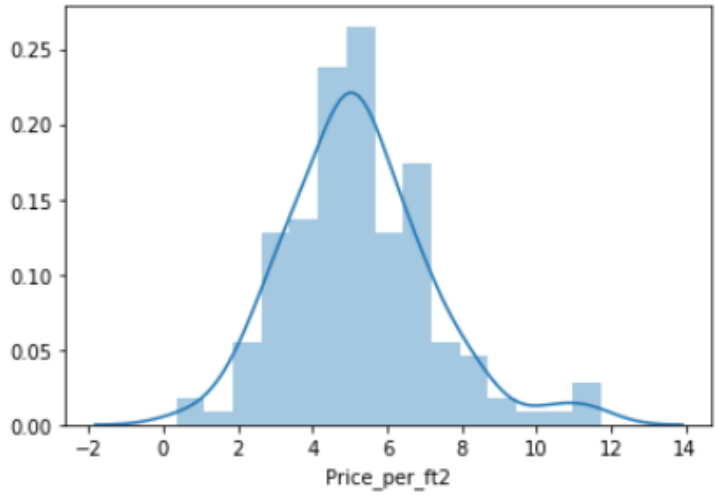
	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	2	Coffee Shop	Discount Store	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
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Bakery	Mobile Phone Shop	Pizza Place	Sandwich Place	Park	Gym	Latin American Restaurant
3	Manhattan	Inwood	40.867684	-73.921210	3	Mexican Restaurant	Lounge	Pizza Place	Café	Wine Bar	Bakery	American Restaurant	Park
4	Manhattan	Hamilton Heights	40.823604	-73.949688	0	Mexican Restaurant	Coffee Shop	Café	Deli / Bodega	Pizza Place	Liquor Store	Indian Restaurant	Sushi Restaurant

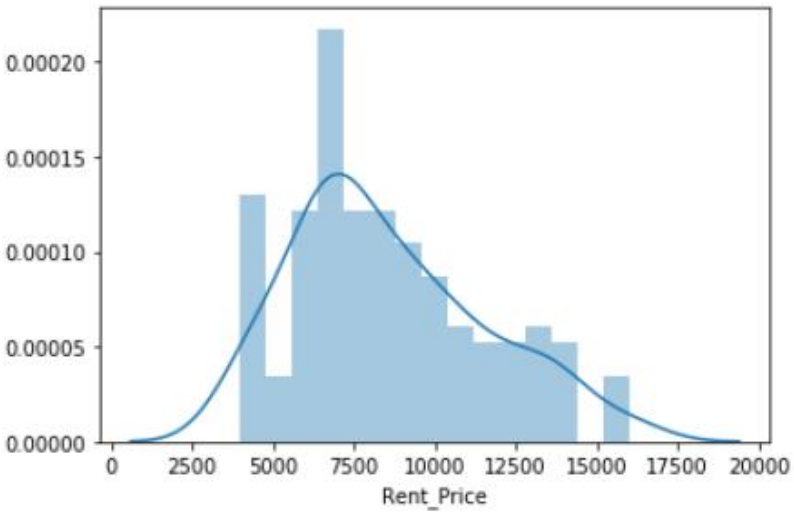


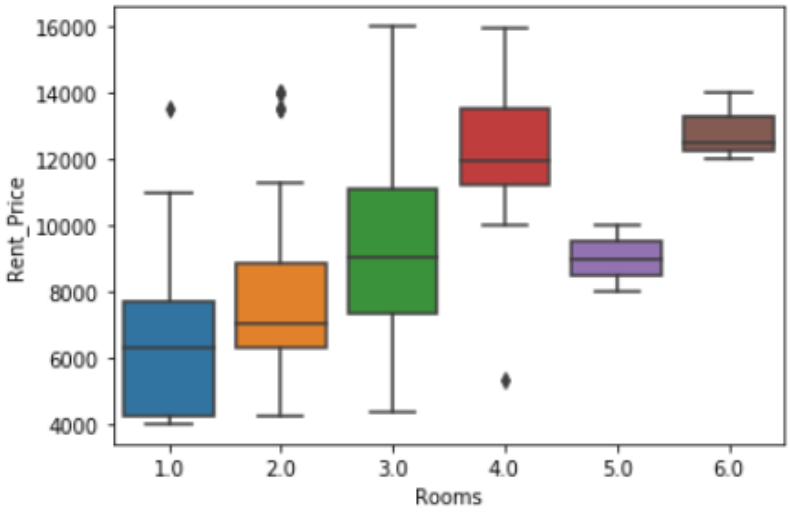
After examining the cluster data, it turns out the cluster 2 reassembles the most the place in Singapore (table below).

	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

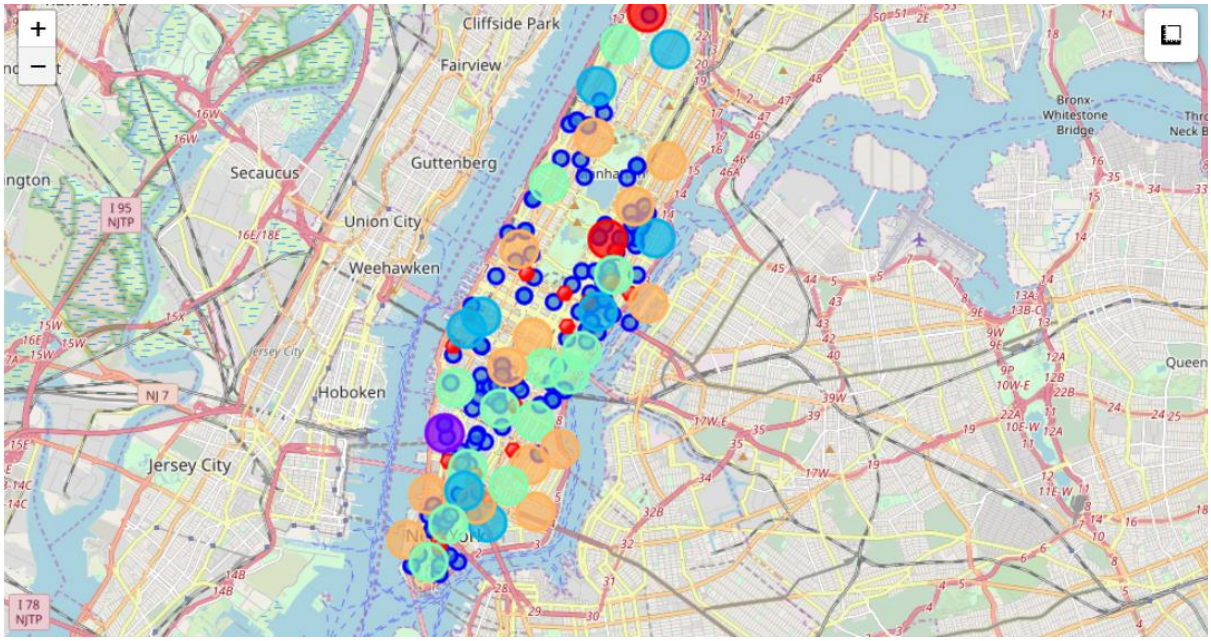
Once we have mapped the rent places in Manhattan we can analyse the price, and I have obtained the following result:







We have everything we need to map rental places with price, subway locations and cluster of venues.
This allows us to have a quick and clear overview on apartments that are in line with my demands.



Results and Conclusion

From the map above I have chosen the apartment on 305 East 63rd Street in the Sutton Place Neighbourhood and near 'subway 59th Street' station, Cluster # 2 Monthly rent : 7500 Dollars

The methodology use allows for a quick and easy selection of apartments while using an interactive map.
Also, it allowed me to compare the similarity of neighbourhoods in different cities.