**Report Content**

1.Introduction Section : ⁃ The “business problem” to be solved by this project and who may be interested

2.Data Section: ⁃ Describe Data requirements and Sources needed to solve the problem

3.Methodology section: ⁃ Main component of the report - Execute data processing, describe/discuss any exploratory data analysis and/or inferential statistical testing performed, and/or machine learnings used.

4.Results section: ⁃ Discussion of the results and ﬁnding of answer

5.Discussion section: ⁃ Discussion of observations noted and any recommendations

6.Conclusion section: ⁃ Answer chosen and conclusions.

**1. Introduction Section:**

**Discussion of the business problem and the audience who would be interested in this project.**

**Description of the Problem and Background:**

**Scenario:**

I am a data scientist from India currently working in XYZ Company in Singaore. I currently live within walking distance to bus stop and I enjoy many ammenities and venues in the area, such as various restaurants, cafes, food shops and entertainment. I have been offered a great opportunity to work for a leader firm in ABC Company. I am very excited and I want to use this opportunity to practice my learnings in Coursera in order to answer relevant questions arisen. For that I have to move some other place. For example I have to move from Signapore to New York The key question is : How can I find a convenient and enjoyable place similar to mine 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:

1. apartment must be 2 or 3 bedrooms
2. desired location is near a metro station in the Manhattan area and within 1.0 mile (1.6 km) radius
3. price of rent not exceed $7,000 per month
4. top ammenities in the selected neighborhood shall be similar to current residence
5. desirable to have venues such as coffee shops, restaurants Asian Thai, wine stores, gym and food shops
6. as a reference, I have included a map of venues near current residence in Singapore.

**Business Problem:**

The challenge is to find a suitable apartment for rent in Manhattan NY that complies with the demands on location, price and venues.

**2. Data Section:**

**Description of the data and its sources that will be used to solve the problem**

**Description of the Data:**

The following data is required to answer the issues of the problem:

1. List of Boroughs and neighborhoods of Manhattan with their geodata (latitud and longitud)
2. List of Subway metro stations in Manhattan with their address location
3. List of apartments for rent in Manhattan area with their addresses and price
4. Preferably, a list of apartment for rent with additional information, such as price, address, area, # of beds, etc
5. Venues for each Manhattan neighborhood ( than can be clustered)
6. Venues for subway metro stations, as needed

**How the data will be used to solve the problem**

The data will be used as follows:

1. Use Foursquare and geopy data to map top 10 venues for all Manhattan neighborhoods and clustered in groups ( as per Course LAB)
2. Use foursquare and geopy data to 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 ammenities near each metro station, or explore each subway location separately
3. Use Foursquare and geopy data to map the location of rental places, in some form, linked to the subway locations.
4. create a map that depicts, for instance, the average rental price per square ft, around a radious of 1.0 mile (1.6 km) around each subway station - or a similar metrics. I will be able to quickly point to the popups to know the relative price per subway area.
5. Addresses from rental locations will be converted to geodata( lat, long) using Geopy-distance and Nominatim.
6. Data will be searched in open data sources if available, from real estate sites if open to reading, libraries or other government agencies such as Metro New York MTA, etc.

The procesing of these DATA will allow to answer the key questions to make a decision:

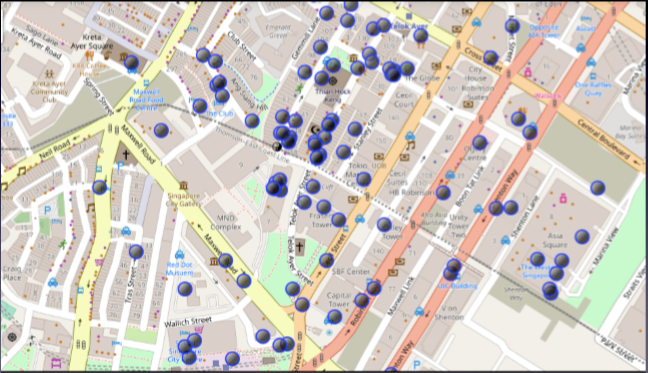
1. what is the cost of rent (per square ft) around a mile radius from each subway metro station?
2. what is the area of Manhattan with best rental pricing that meets criteria established?
3. What is the distance from work place ( Park Ave and 53 rd St) and the tentative future home?
4. What are the venues of the two best places to live? How the prices compare?
5. How venues distribute among Manhattan neighborhoods and around metro stations?
6. Are there tradeoffs between size and price and location?
7. Any other interesting statistical data findings of the real estate and overall data.

**3. 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

**4.Execution and Results:**

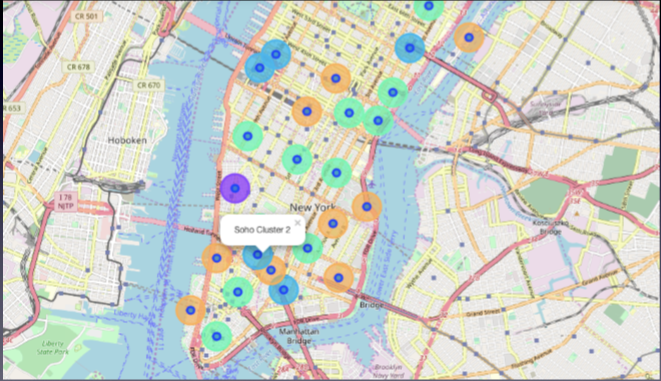
**Current residence Neighborhood in Singapore**



**Venues around Neighborhood in**



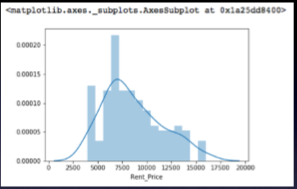
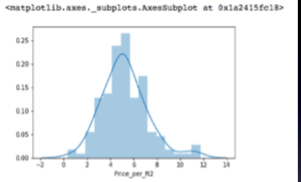
**Manhattan Map - Neighborhoods and Cluster of Venues**

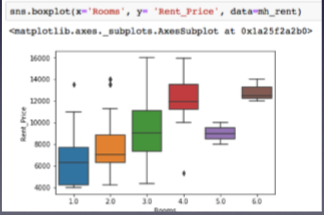


**GeoData Manhattan apts for rent**

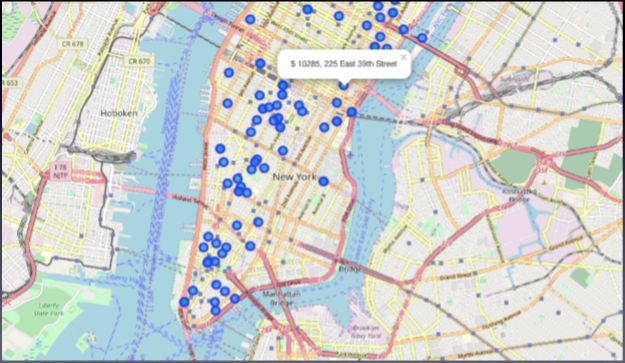


**Rental Price Statistics MH Apartments Budget US7000/month is around the mean**

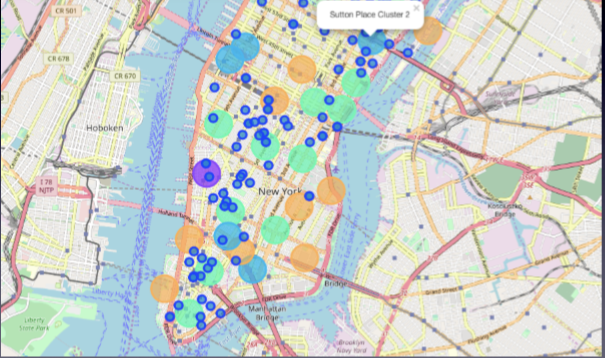




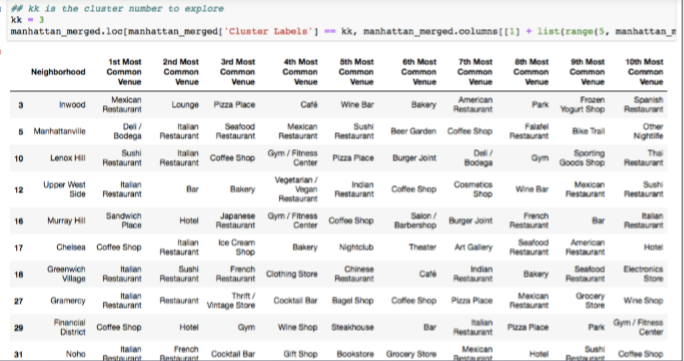
**Apartments for Rent in MH**



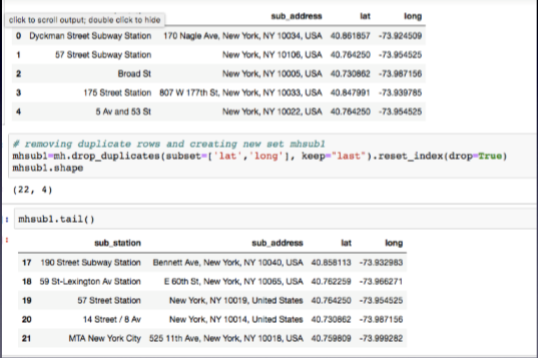
**MH apts for rent with venue clusters**



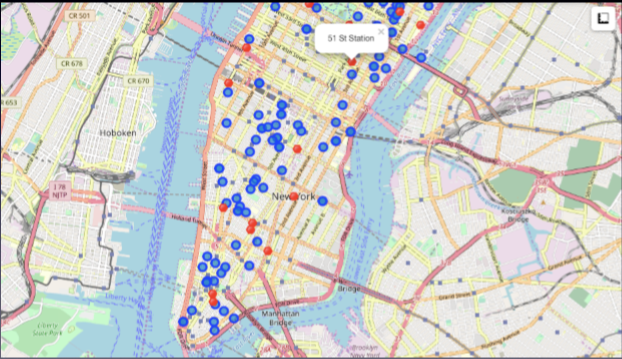
**Venues of cluster 3**



**Manhattan subway stations geodata**

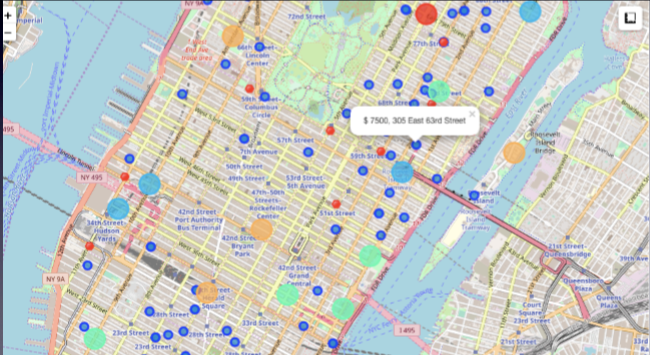


**Apts for rent (blue) and subway stations (red)**



**Selected Apartment!**

The ONE consolidated map shows all information for decision: Apartments address, price, neighborhood, cluster of venues and subway station nearby. Blue dots=apts , Red dots=Subway station, Bubbles=Cluster of Venues



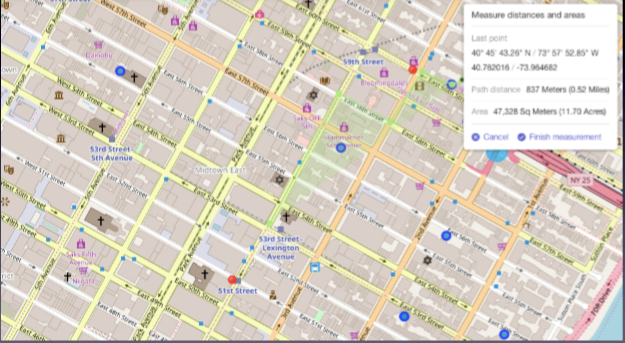
**Apartment Selection:**

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 US7500 slightly above the US7000 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 US6935, just under the US7000 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.

I will walk to work Walk from home to work is less than 1 km!



**Venus in Cluster 2 near future home**



# 5.0 DISCUSSION

### In general, I am positively impressed with the overall organization, content and lab works presented during the Coursera IBM Certification Course which helped me to learn a lot.

### 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.

### 6.0 CONCLUSIONS

### I feel rewarded with the efforts, time and money spent. I believe this course with all the topics covered is well worthy of appreciation.