Coursera Capstone Project
Coursera Capstone - REPORT CONTENT

1. Introduction Section: ▪Discussion of the business problem and the interested audience in this project.
2. Data Section:
   ▪Description of the data that will be used to solve the problem and the sources.
3. Methodology section ▪Discussion and description of exploratory data analysis carried out, any inferential statistical testing performed, and if any machine learnings were used establishing the strategy and purposes.
4. Results section ▪Discussion of the results.
5. Discussion section ▪Elaboration and discussion on any observations noted and any recommendations suggested based on the results.
6. Conclusion section ▪Report Conclusion.

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 currently live within walking distance to Downtown Telok Ayer MRT metro station and I enjoy many ammenities and venues in the area, such as various international cousine restaurants, cafes, food shops and entertainment. I have been offered a great opportunity to work for a leader firm in Manhattan, NY. I am very excited and I want to use this opportunity to practice my learnings in Coursera in order to answer relevant questions arisen. The key question is : How can I find a convenient and enjoyable place similar to mine now in Singapore? Certainly, I can use available real estate apps and Google but the idea is to use and apply myself the learned tools during the course. 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:

- apartment must be 2 or 3 bedrooms
- desired location is near a metro station in the Manhattan area and within 1.0 mile (1.6 km) radius
- price of rent not exceed $7,000 per month
- top ammenities in the selected neighborhood shall be similar to current residence
- desirable to have venues such as coffee shops, restaurants Asian Thai, wine stores, gym and food shops
- 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. The data required to resolve this challenge is described in the following section 2, below.

Interested Audience

I believe this is a relevant challenge with valid questions for anyone moving to other large city in US, EU or Asia. The same methodology can be applied in accordance to demands as applicable. This case is also applicable for anyone interested in exploring starting or locating a new business in any city. Lastly, it can also serve as a good practical exercise to develop Data Science skills.

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

How the data will be used to solve the problem

The data will be used as follows:
- Use Foursquare and geopy data to map top 10 venues for all Manhattan neighborhoods and clustered in groups (as per Course LAB)
- 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
- Use Foursquare and geopy data to map the location of rental places, in some form, linked to the subway locations.
- 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.
- Addresses from rental locations will be converted to geodata(lat, long) using Geopy-distance and Nominatim.
- 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:
- what is the cost of rent (per square ft) 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 53 rd St) and the tentative future 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 tradeoffs between size and price and location?
- Any other interesting statistical data findings of the real estate and overall data.