**Coursera Capstone - REPORT**

**Battle of Neighborhoods**

**1. Introduction Section:**

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

**1.1 Scenario and Background**

I am a IT Consulting currently residing in Downtown Toronto. I currently live within walking distance to Downtown "Dundas Subway station" therefore I have access to good public transportation to work. Likewise, I enjoy many amenities in the neighborhood, such as international cuisine restaurants, cafes, food shops and entertainment.

I have been offered a great opportunity to work in Manhattan, NY. Even though I am looking forward to it, I am a bit nervous about the fact that i need to secure a comparable place to live in Manhattan.

Therefore, I decided to apply the learned skills during the Coursera course to explore ways to make sure my decision is factual and rewarding.

Of course, there are alternatives to achieve the answer using available Google and forums, but it rewarding doing it myself with learned tools.

**1.2 Problem to be resolved:**

The challenge to resolve is being able to find a rental apartment unit in Manhattan NY that offers similar characteristics and benefits to my current situation. Therefore, in order to set a basis for comparison, I want to find a rent a unit subject to the following conditions:

Apartment with min 2 bedrooms with monthly rent not to exceed US$5000/month

Unit located within walking distance from a subway metro station in Manhattan

Area with amenities and venues similar to the ones described for current location

**1.3 Interested Audience**

I believe this is a relevant project for a person or entity considering moving to a major city in anywhere, since the approach and methodologies used here are applicable in all cases.

**2. Data Section:**

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

**2.1 Data of Current Situation**

I Currently reside in the neighborhood of 'Bay Street' in Downtown Toronto. I use Foursquare to identify the venues around the area of residence which are then shown in the Singapore map shown in methodology and execution in section 3.0 . It serves as a reference for comparison with the desired future location in Manhattan NY

**2.2 Data Required to resolve the problem**

In order to make a good choice of a similar apartment in Manhattan NY, the following data is required: List/Information on neighborhoods form Manhattan with their Geodata ( latitud and longitud. List/Information about the subway metro stations in Manhattan with geodata. Listed apartments for rent in Manhattan area with descriptions ( how many beds, price, location, address) Venues and amenities in the Manhattan neighborhoods (e.g. top 10)

**2.3 sources and manipulation The list of Manhattan neighborhoods is worked out during LAb exercise during the course. A csv file was created which will be read in order to create a data frame and its mapping.**

The csv file 'Manhattan\_borough.csv' has the following below data structure. The file will be directly read to the Jupiter Notebook for convenience and space savings.

Borough Neighborhood Latitude Longitude Manhattan Marble Hill 40.87655078 -73.91065966 Manhattan Chinatown 40.71561842 -73.99427936 Manhattan Washington Heights 40.85190253 -73.93690028 Manhattan Inwood 40.86768396 -73.92121042 Manhattan Hamilton Heights 40.82360428 -73.94968792 Manhattan Manhattanville 40.81693443 -73.95738539

A list of Manhattan subway metro stops was complied in Numbers (Apple excel) and it was complemented with Wikipedia data ( <https://en.wikipedia.org/wiki/List_of_New_York_City_Subway_stations_in_Manhattan>) and information from NY Transit authority and Google maps (<https://www.google.com/maps/search/manhattan+subway+metro+stations/@40.7837297,-74.1033043,11z/data=!3m1!4b1>) for a final consolidated list of subway stops names and their address. The geolocation was obtained via an algorythm using Nominatim. Details will be shown in the execution of methodolody in section 3.0.

A list of places for rent was collected by web-browsing real estate companies in Manhattan : <http://www.rentmanhattan.com/index.cfm?page=search&state=results> <https://www.nestpick.com/search?city=new-york&page=1&order=relevance&district=manhattan&gclid=CjwKCAiAjNjgBRAgEiwAGLlf2hkP3A-cPxjZYkURqQEswQK2jKQEpv_MvKcrIhRWRzNkc_r-fGi0lxoCA7cQAvD_BwE&type=apartment&display=list><https://www.realtor.com/apartments/Manhattan_NY> A csv file was compiled with the rental place that indicated: areas of Manhattan, address, number of beds, area and monthly rental price. The csv file "nnnn.csv" had the following below structure. An algorithm was used to create all the geodata using Nominatim, as shown in section 3.0. The actual algorithm coding may be shown in 'markdown' mode because it takes time to run. With the use of geolocator = Nominatim() , it was possible to determine the latitude and longiude for the subway metro locations as well as for the geodata for each rental place listed. The loop algorythms used are shown in the execution of data in section 3.0 "Great circle" function from geolocator was used to calculate distances between two points , as in the case to calculate average rent price for units around each subway station and at 1.6 km radius. Foursquare is used to find the avenues at Manhattan neighborhoods in general and a cluster is created to later be able to search for the venues depending of the location shown.

**2.4 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 amenities 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 radius (1/1.5 miles) around each subway station - or a similar metric.

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.

**2.5 Mapping of Data**

The following maps were created to facilitate the analysis and the choice of the palace to live. Manhattan map of Neighborhoods Manhattan subway metro locations Manhattan map of places for rent Manhattan map of clustered venues and neighborhoods Combined maps of Manhattan rent places with subway locations Combined maps of Manhattan rent places with subway locations and venues clusters

**3. Methodology section:**

This section represents the main component of the report where the data is gathered, prepared for analysis. The tools described are used here and the Notebook cells indicates the execution of steps.

The analysis and the strategy:

1. The strategy is based on mapping the above described data in section 2.0, in order to facilitate the choice of at least two candidate places for rent.
2. The choice is made based on the demands imposed: location near a subway, rental price and similar venues to Toronto. This visual approach and maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy.
3. The processing of these DATA and its mapping will allow to answer the key questions to make a decision:
4. 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?