DATA

[Capstone Project - The Battle of Neighborhoods (Week 1)]

Data Section

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 Nominatim. 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 Data Sources, Data Processing and Tools used

- Singapore data and map is to be created with use of Nominatim , Foursquare and Folium mapping
- Manhattan neighborhoods were obtained from Wikipedia and organized by Neighborhoods with geodata via Nominatim 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 Nominatim.
- 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 Nominatim to get latitude and longitude of subway stations and also for each of (144 units) the apartments for rent listed.

Geopy_distance and Nominatim 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.

-Achal Shakywar