Final Project – Shubh Ganatra

Coursera Capstone - REPORT CONTENT

- Introduction Section: Discussion of the business problem and the interested audience in this project.
- Data Section: Description of the data that will be used to solve the problem and the sources.
- 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.
- Results section: Discussion of the results.
- Discussion section: Elaboration and discussion on any observations noted and any recommendations suggested based on the results.
- Conclusion section: Report Conclusion.

Introduction

I am an aspiring entrepreneur looking for a suitable place to open his new restaurant in New York City, for this I require various types of data including all the localities and neighborhoods, places already existing in those localities, the footfall in those places, what type of places are doing really good i.e. are trending in those specific localities. There are many platforms available via Google or social media platforms to get such information, but it is rewarding to do it myself with self-learned tools.

I believe this is a relevant project for a person or entity considering starting a venture in a major city in Europe, US or Asia, since the approach and methodologies used here are applicable in all cases. The use of Foursquare data and mapping techniques combined with data analysis will help resolve the key questions arisen. Lastly, this project is a good practical case toward the development of Data Science skills.

Data

Description of Data used

I am currently a new comer in NYC hence do not have any information about the localities and neighborhoods, hence,

In order to make a good choice of a location for a new restaurant in NY, the following data is required:

- List/Information on Boroughs with geodata.
- List/Information about the neighborhoods in each borough with geodata.
- Listed places of interest in specified areas with descriptions (type of establishment, footfall, rating)
- Ranking Of venues and amenities in the neighborhoods (e.g. top 10)

A json file was created which will be read in order to create a data frame and its mapping. The csv file ‘newyork_data.json’ has the following below data structure. The file will be directly read to the Jupiter Notebook for convenience and space savings. The clustering of neighborhoods and mapping will be shown however.

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

- List of Boroughs and Neighborhoods of New York with their geodata (latitude and longitude)
- List of places in every neighborhood
- Descriptions of the places, including geodata, type of establishment, footfall
- The top trending venues in the neighborhood
- The different types of venues existing in each neighborhood
- The number of venues in each category of a neighborhood

The data will be used as follows:

Use Foursquare and goopy 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 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.

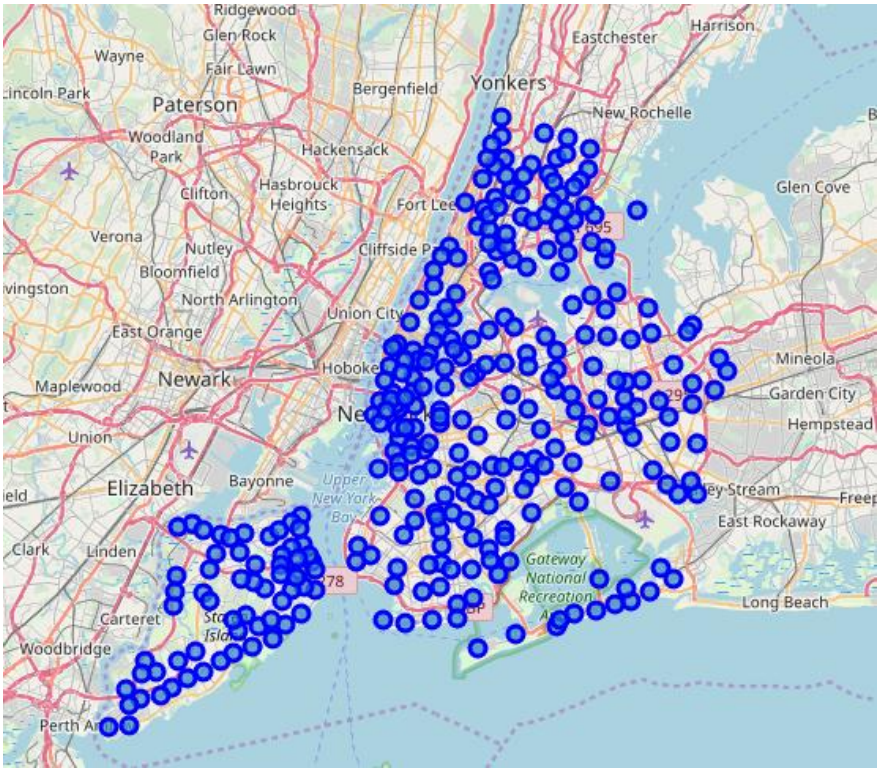
The processing of these DATA will allow to answer the key questions to make a decision

- What are total the number of neighborhoods in the user specified borough?
- How many establishments already exist in the selected neighborhood from the above list?
- What are the different categories in which the establishments can be divided into?
- What ratings have the currently existing venues received and what is the footfall?
- How venues distribute among neighborhoods?

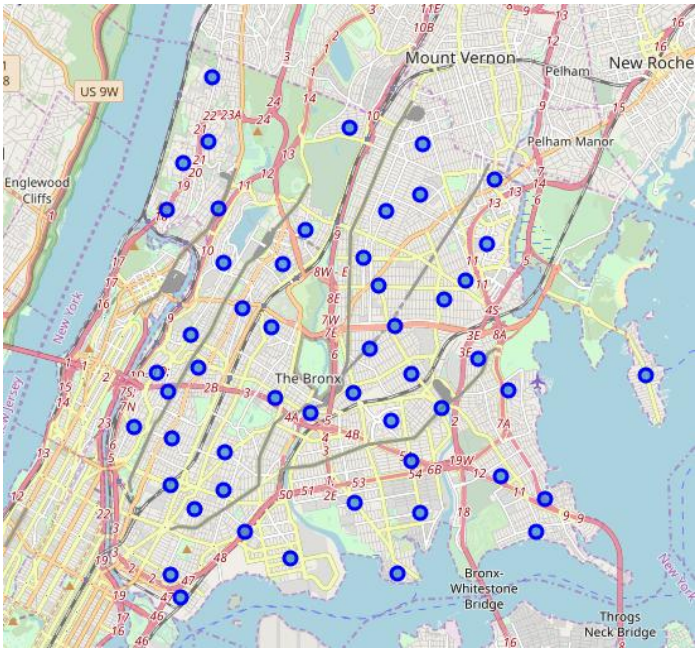
Methodology

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 indicate the execution of steps.

The analysis and the strategy: The strategy is based on mapping the above described data in section 2.0, in order to facilitate the choice of a neighborhood for the establishment of a new venture. The choice is made based on neighborhood characteristics: Number of already existing establishments, their description, footfall and rating, also other trending places in the neighborhood. This visual approach and maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy



Map Of New York With All Neighborhoods Marked



Map Of A Selected Borough (Bronx) With All Neighborhoods Marked

The processing of these DATA and its mapping will allow to answer the key questions to make a decision:

- The total number of establishments in a neighborhood
- The top 5 establishments in every neighborhood
- The top 5 establishments with rating and footfall in every neighborhood
- Different types of venues in every neighborhood
- Number of venues in each category

Results

Displayed all the above criterias of information required i.e. details regarding the neighborhoods and various establishments of those neighborhoods in New York City

	name	lat	lng
categories			
American Restaurant	1	1	1
Arcade	1	1	1
Bank	1	1	1
Baseball Field	1	1	1
Bus Station	1	1	1
Convenience Store	1	1	1
Discount Store	2	2	2
Donut Shop	2	2	2
Electronics Store	1	1	1
Fast Food Restaurant	1	1	1
Fried Chicken Joint	1	1	1
Gym / Fitness Center	1	1	1
Mattress Store	1	1	1
Men's Store	1	1	1
Mexican Restaurant	1	1	1
Pet Store	1	1	1
Pizza Place	1	1	1
Playground	1	1	1
Sandwich Place	1	1	1
Spanish Restaurant	1	1	1
Supermarket	1	1	1

```
----Baychester----
            venue  freq
0  Discount Store  0.09
1      Donut Shop  0.09
2      Playground  0.04
3  Mattress Store  0.04
4     Bus Station  0.04
```