Description of Data

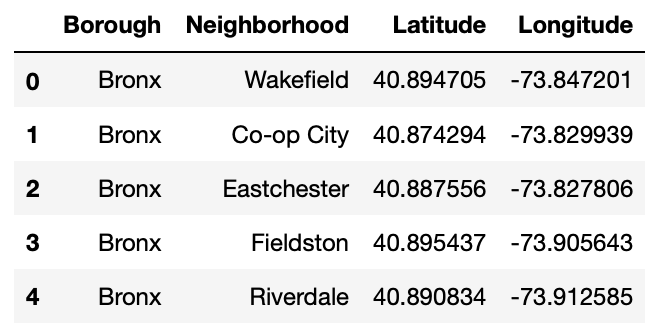
For this project, we will use New York Neighborhood data that provide previously in Coursera course, and we narrow down to neighborhoods within Manhattan, which includes 40 major neighborhoods. To help our target audience to make decision on which is a better neighborhood for them to open a new restaurant, I spent more time in obtaining different data that can exposed neighborhoods in different perspective.

Therefore, I have found some demographic data for each neighborhood in Manhattan such as, Median Household Income, Average Household Income, Per-capita Income and Population of each neighborhood. These demographic data will be useful since it is possible that neighborhood with higher average household income or with larger population may good for business. We will explore about the demographic data and get some summary statistics as well.

Finally, we could scrape data by using Foursquare API to analysis what kind of characteristics the surrounding venue have within each neighborhood in Manhattan. This can help investor to get better idea what type of competitor(restaurant) they may have if they open their restaurant in specific neighborhood.

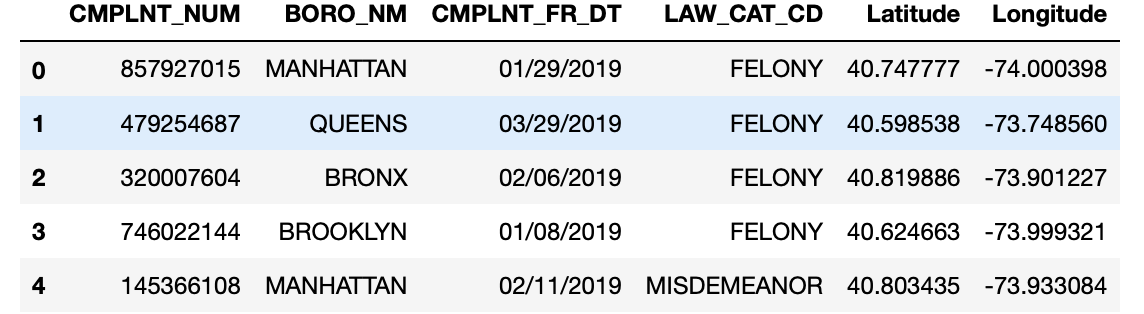
Data Source from:

1. New York Neighborhood data: <https://geo.nyu.edu/catalog/nyu_2451_34572>



Latitude and Longitude can help we locate each neighborhood in Manhattan. Most importantly, we can utilize this location information to help in Foursquare API when we try to scrap all restaurant within each neighborhood latter.

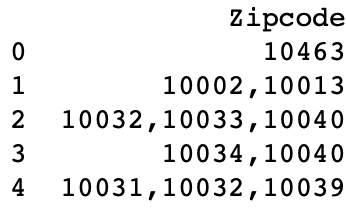
2. New York crime data: <https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Map-Year-to-Date-/2fra-mtpn>



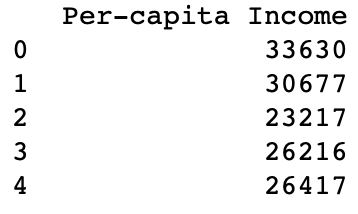
This dataset includes different types of crimes reported to the New York City Police Department (NYPD) from 2016 till this year 2019.It includes 24 features and 222398 rows of data. From this dataset, we will focus on 6 features as above screenshot shown. We can use these data to explore about crime in each neighborhood.

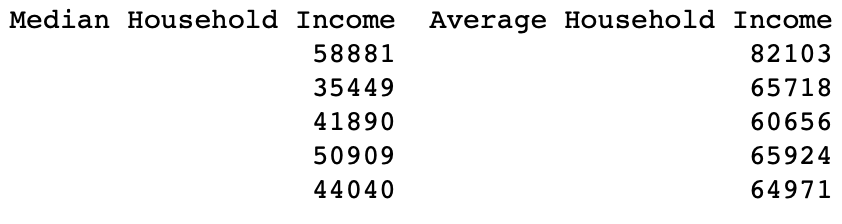
3.'zipcode' data was from wikipidia that I search one by one according to the Neighborhood

Use to locate each specific neighborhood and helps in grabs other demographic data online.



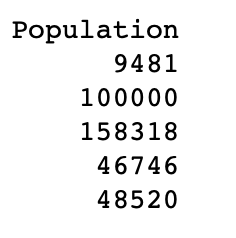
4.Income data was searched one by one from website 'https://www.incomebyzipcode.com/' by passing the zipcode from above.





5.Part of the Population data are from https://www.worldatlas.com/articles/manhattan-neighborhoods-by-population.html, and part of them are searched from wikipidia.

The following is a head of the data, and will show approximately how many residents will be in each neighborhood.



The whole dataset will look like this after putting everything together.

