Eram Manasia
11/3/2019

- **A brief section addressing feedback from the project proposal**

  N/A

- **A 200+ word summary of findings from the data exploration**

From my data exploration, I have decided to make the following changes to the project:

1. The collection of data has new data sources:
    a. https://www1.nyc.gov/site/nypd/stats/crime-statistics/borough-and-precinct-crime-stats.page#bronx : This is for crime statistics for each borough.
    b. https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm : List of NYC zip codes. This is to add all zip codes to its respective borough.
    c. https://www.psc.isr.umich.edu/dis/census/Features/tract2zip/ : Excel sheet for median income, average income, and population by zip code.
2. Instead of distinguishing data by zip code, it will now be done by borough. So, there will be an analysis of borough income vs crime rate. The reason for this is because it is difficult to find data for crime rate distinguished by zip code.

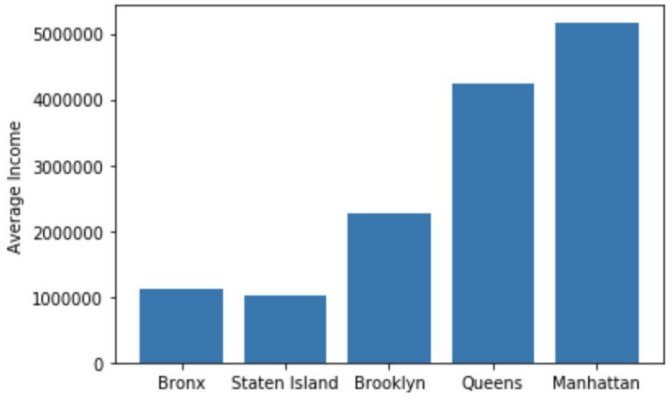
So far, data exploration has only been done on the income aspect of the project. According to the bar graphs I have generated, the list below have the boroughs in order from greatest to least for the criteria defined:

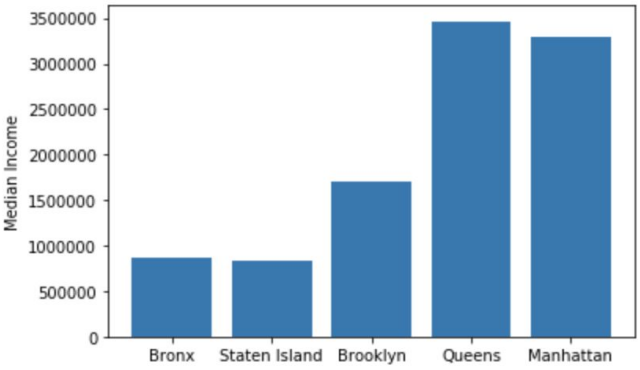
1. Average Income : Manhattan, Queens, Brooklyn, Bronx, Staten Island
2. Median Income : Queens, Manhattan, Brooklyn, Bronx, Staten Island
3. Population : Brooklyn, Queens, Manhattan, Bronx, Staten Island
4. Zip code Count: Queens, Manhattan, Brooklyn, Bronx, Staten Island

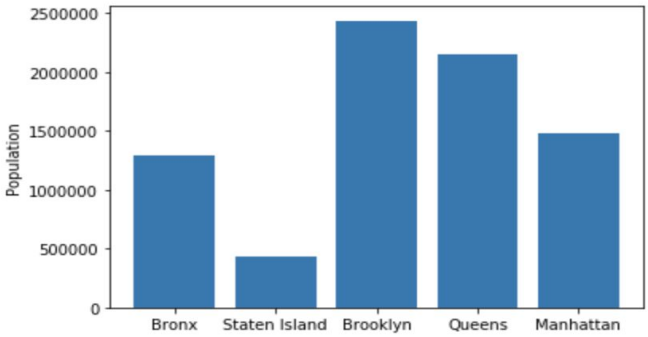
For the crime rate aspect of the project, the analysis will be by borough. There will need to be edits made to the excel sheet itself which will take about a week to do so. I will also need to look for more data sources so that I can finalize a year for analysis.

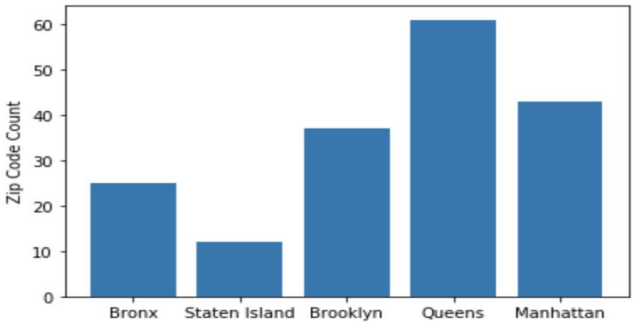
- **At least two visuals from exploratory data analysis**

Eram Manasia
11/3/2019









Eram Manasia
11/3/2019

- **Any concerns you are having about the project at-large**

I am having difficulty finding data on crime rate statistics. I will be doing research in the following weeks for the data.