

New York City and Neighborhood Crime Rate Data Analysis

Introduction

- ❑ New York is a busy city with multicultural presence from all over the world.
- ❑ City has lots of infrastructure, which keeps on expanding with all the growing business and people's needs.
- ❑ Data Analysis is very helpful in determining various aspects of the city for various reasons, for e.g.: which area is safest to expand, which area has low population density, which area is safest and could be successful for any new business, etc.
- ❑ The focus of this presentation is along the lines of finding the area with lowest crime rate in the New York metropolitan area.

The Data

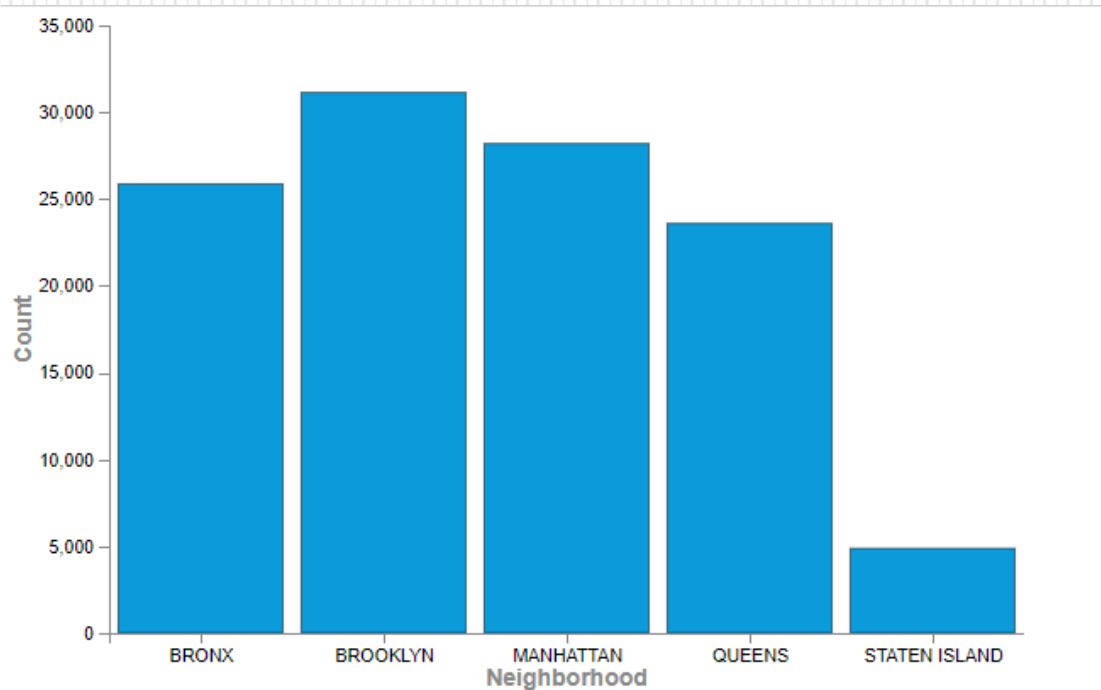
- Neighborhood with borough information data collected from https://cocl.us/new_york_dataset
- New York area and precinct data collected from the city of New York site open data <https://opendata.cityofnewyork.us/>
- Crime rate data collected from <https://data.cityofnewyork.us/Public-Safety/Citywide-Crime-Statistics/c5dk-m6ea>

The Methodology

- ❑ Load and Process the data
- ❑ Sorting and Parsing
- ❑ Visualize the data
- ❑ Another sub-level

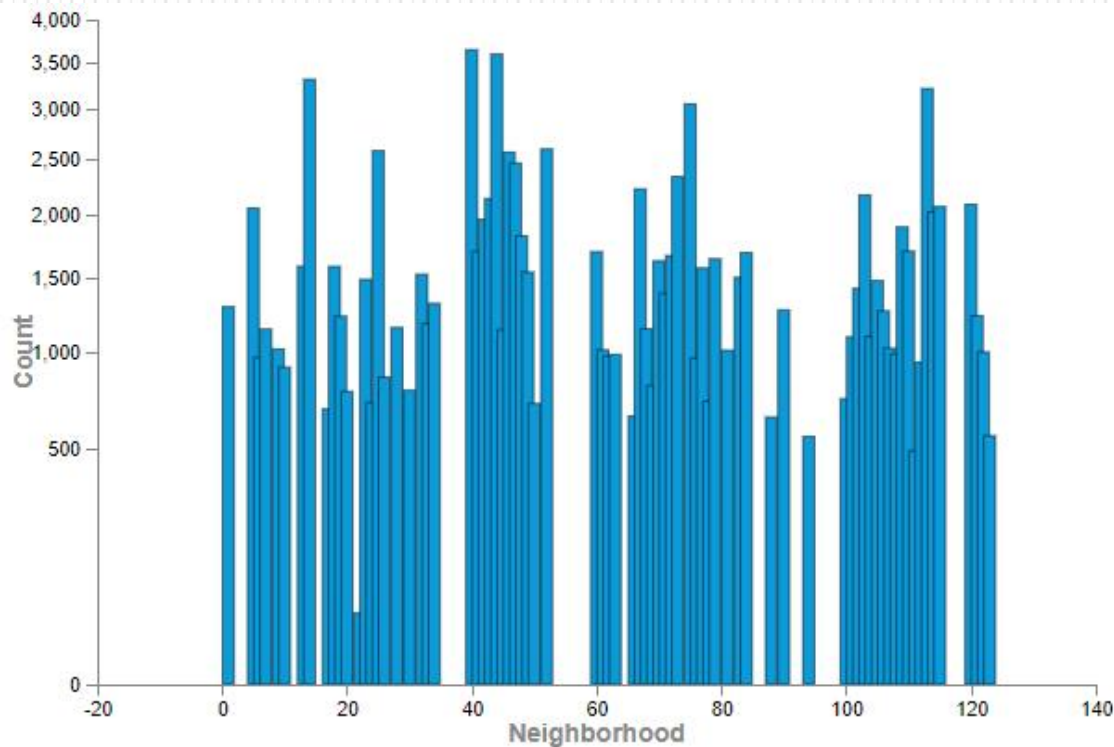
Crime Count by Neighborhood

(for current year)

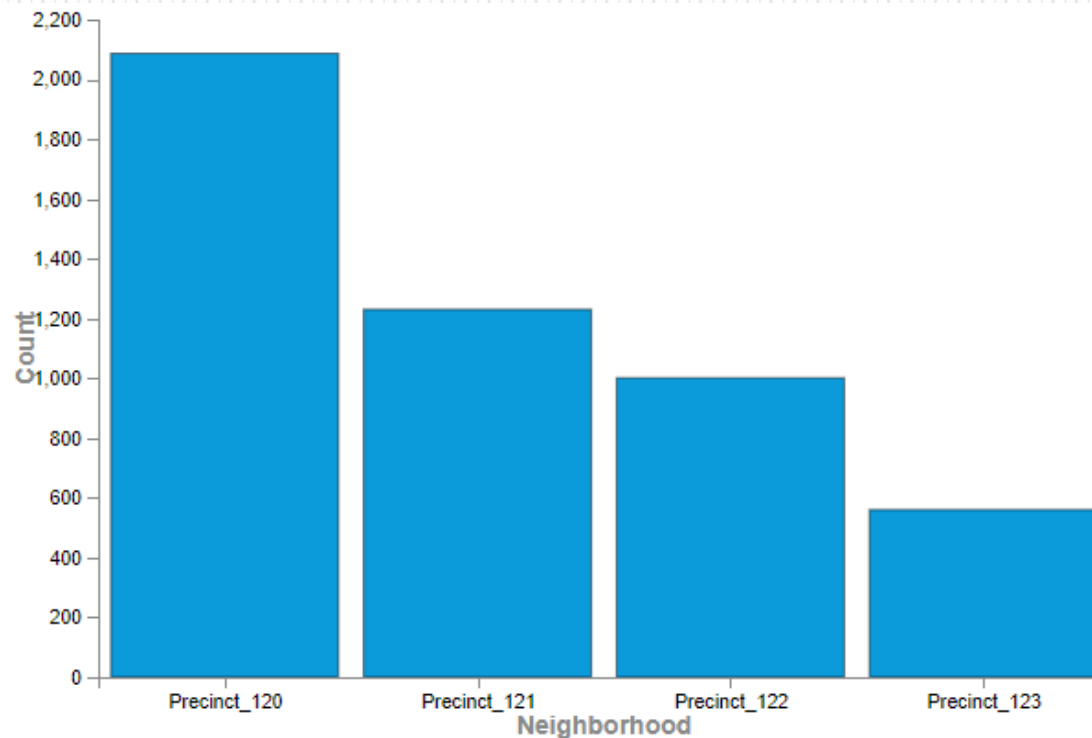


Crime Count by Precinct

(for current year)



Crime Count in Staten Island' Precinct



Discussions and Recommendations

The City of New York Open Data enables us to gain an understanding of the crime volume by type by area but not specific enough to understand the distribution properties.

There is value to the city to explore the detailed crime data using data science to predict frequency, location, timing and conditions to best allocated resources for the benefit of its citizens and it's police force. However, human behavior is complex requiring thick profile data by individual and the conditions surrounding the event(s). To be sufficient for reliable future prediction it would need to demonstrate validity, currency, reliability and sufficiency.

It would be interesting to further study the Census data and if this captures the population that is renting or more temporary/transient population, given the City is a multicultural population hub.

Conclusion

Using the datasets from the City of New York Open Data project we were able to analyze, discover and describe neighborhoods and crime.

While overall, the City of New York Open Data is interesting, it misses the details required for true valued quantitative analysis and predictive analytics which would be most valued by investors and developers to make appropriate investments and to minimize risk.

The Open Data project is a great start and empowers the need for a model to be developed where citizens of digital age can share their data as they wish for detailed analysis that enables the creation of valued services.