

New York Police Department Complaints Visual Data Analysis



Aran Coll
N. V.

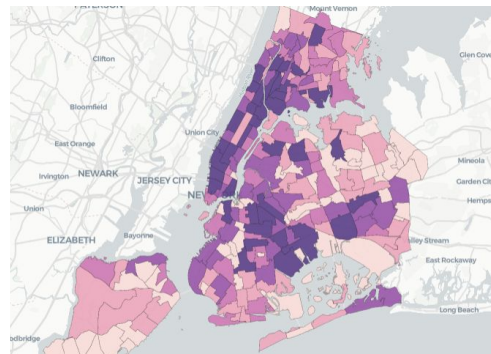
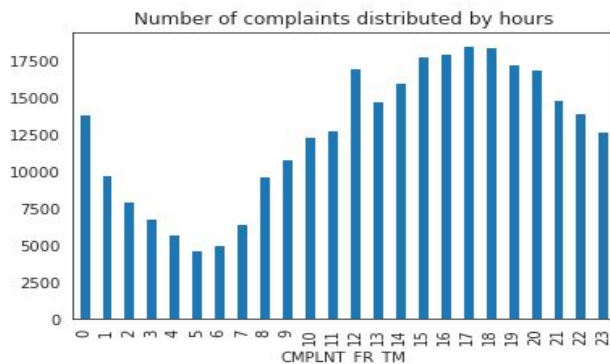
1. Objective of the project

GUARANTEE CITIZENS SECURITY

DATASET

<https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Current-Year-To-Date-/5uac-w243>

- New York City Police Department complaints data -> useful facts about crimes in NY



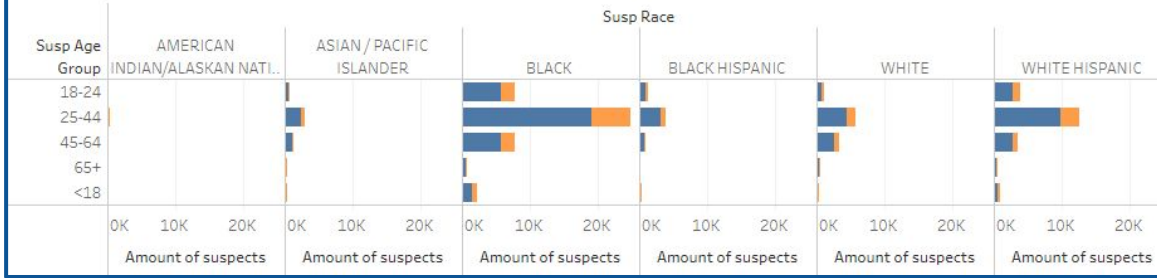
- Which groups should be more guarded (potential suspects) or protected (potential victims)?
- Which crimes are the most committed ones?
- Which areas are more dangerous? Where should be placed more police?

2. Description of the project

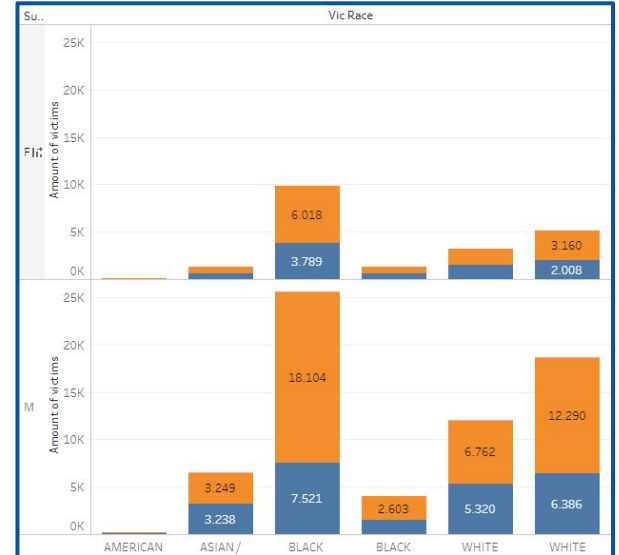
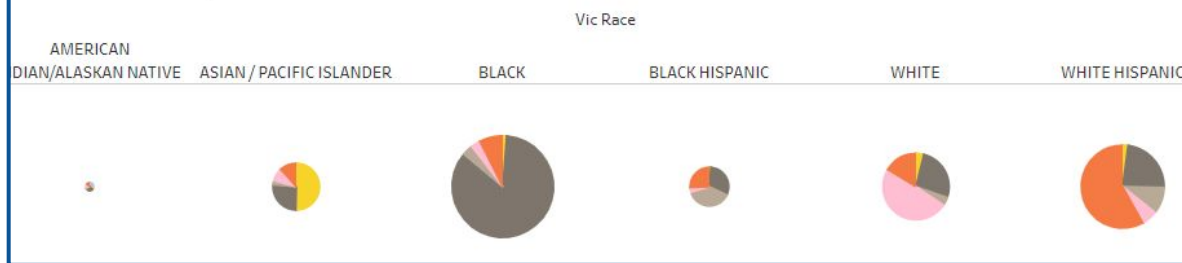
1. **Dataset Exploratory Data Analysis:** main characteristics of the variables, prepare the dataset.
2. **Distributions** to look for interesting information:
 - a. The number of complaints are very similar independent of the weekday.
 - b. The two more dangerous days are the first of June and the first day of the year.
 - c. The types of offenses that occur during the day and night are very similar, with some small differences
 - d. The hours with less reported reports are from 2 to 7 am
 - e. Violations tend to be committed by men to women.
3. **Clustering the profiles** of victims and suspects (age, race and gender)
4. **Predicting the race** of the non-identified suspects with GradientBoostingClassifier.
5. **Spatial analysis** and reordering of data grouping by neighborhoods.
6. Several **maps and visualizations** of the offenses by areas

3.1 Visualizations

Suspects race - age group



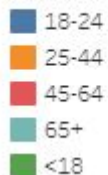
Victim race - Suspect race



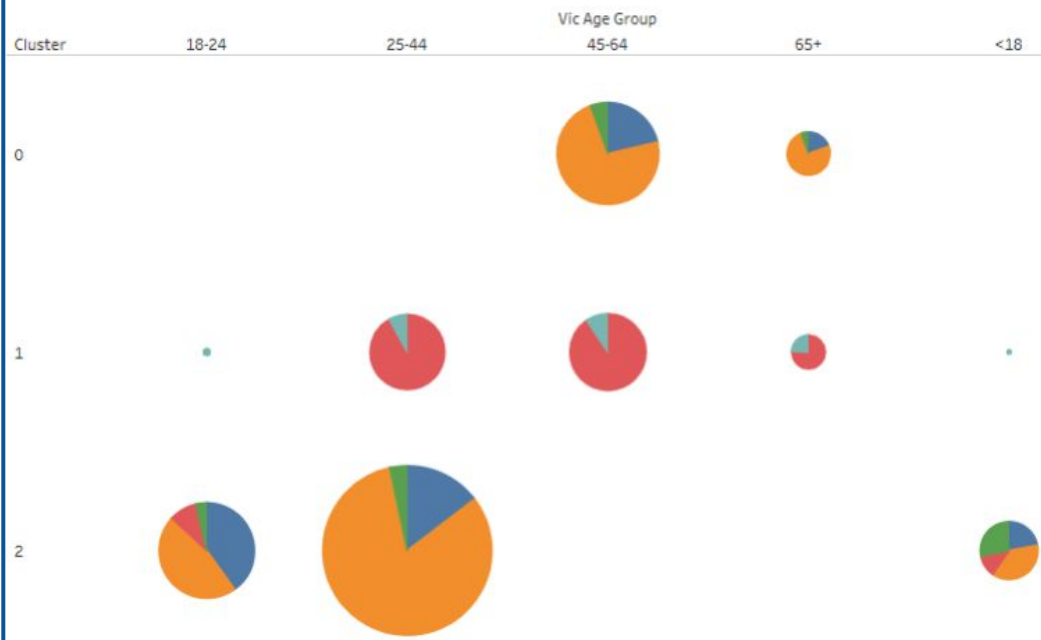
Analysis of the profiles

- **25 -44**
- Mainly **black and white hispanic**
- Tend to attack people from the same race
- More male suspects and female victims
- **Cluster 0:** victims >56, suspects 25-44
- **Cluster 1:** victims all ages, suspects >45
- **Cluster 2:** young victims <44, suspects 18-44
- Distribution of offense category proportional in all clusters.

Susp Age Group

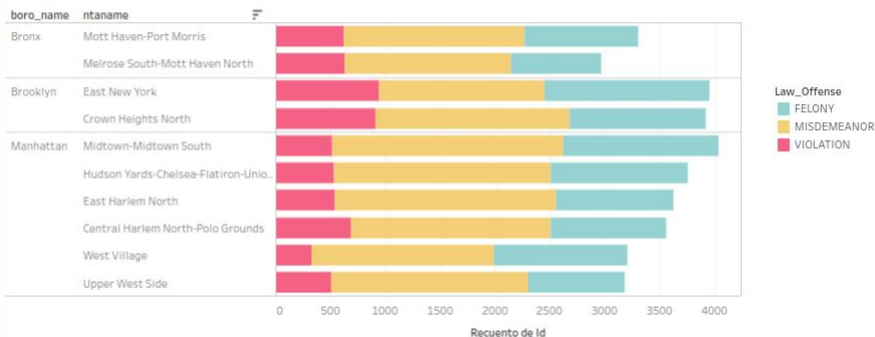


Clusters per ages

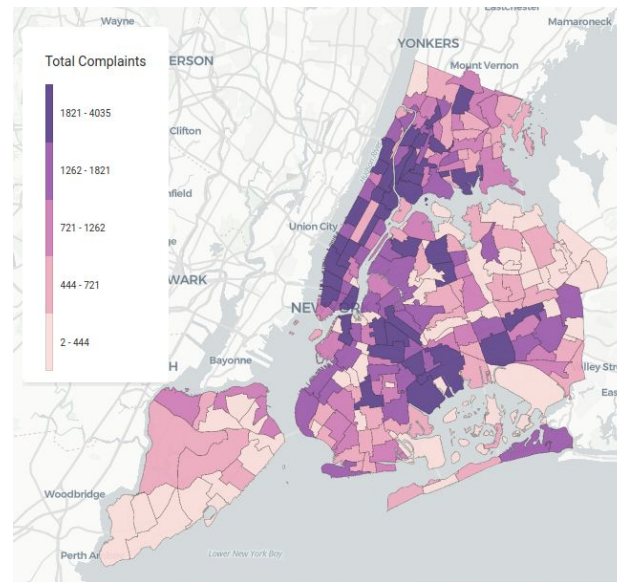
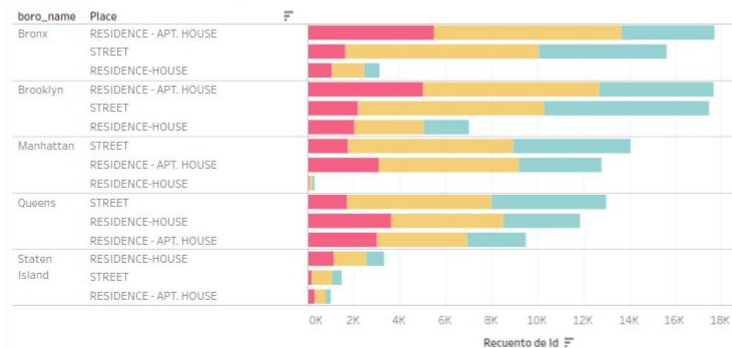


3.2 Visualizations

Neighborhood Tabulation Areas with more complaints

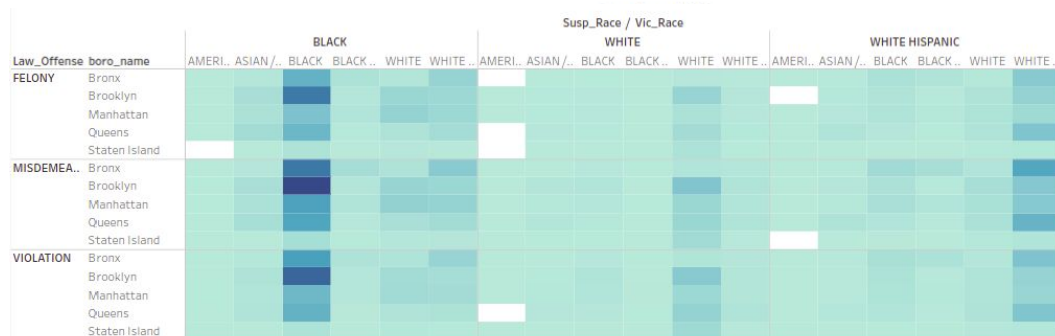


Common locations for Complaints

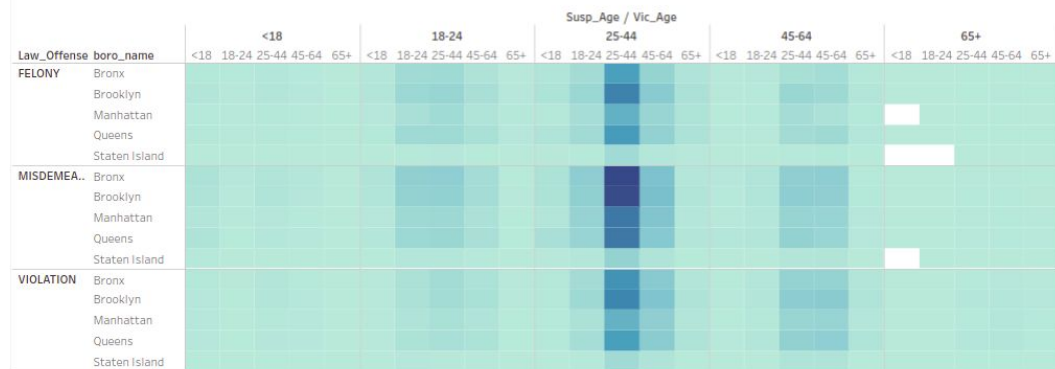


Analysis of the profiles by location

Suspect and Victim Race



Suspect and Victim Age Range



- No clear relation between profiles and borough.
- No clear relation between type of offense and borough either.
- Tend to attack people from the same race.
- Concentrated in the three middle age ranges for both suspects and victims.

3.3 Final conclusions

Suspects and victims profiles have very similar distributions for all the municipalities.



Manhattan,
Brooklyn and Bronx



MISDEMEANOR:
petit larceny,
harassment,
criminal mischief



25-44, black and
white, men



25-44, black and
white, women

4. Next steps

- Join the dataset with demographic and socioeconomic data of the neighbourhoods, so that crimes can be related with the number of inhabitants and the citizens status.
- This would complement and enrich our conclusions with the real situation of every county of New York.

