

The dataset that I will be using is Civilian Complaints Against New York City Police Officers from ProPublica. This dataset is a list of data of about 33,000 civilian complaints that were previously filed against New York City police officers. Although this dataset is constantly being updated, I will be using the latest update that is out currently, which released July 2020. The data within this dataset includes information about the officer that received a complaint such as their name, id, rank, ethnicity, gender, age, and command they report to. There is also data about the complaint itself such as the date the complaint was received, closed, allegation, and outcome. Finally, there is information about the complainant such as their ethnicity, gender, and age. The results of the complaints are categorized into three different groups. The first being substantial evidence, which means that the alleged accusation was found to be true and it violated the rules of the police department. The second category is exonerated, which means the accusation was true, however it did not violate any New York Police Department rules, usually leading to no action being taken against the officer. The last category is unsubstantiated, meaning an investigation of the incident did occur, however nothing could be fully confirmed. A category that was not included in this dataset was complaints that were filed and found to be untrue, meaning this dataset only contains complaints that were found to be true, or have the potential to be true and could not be confirmed.

Some questions that I can ask using this data are as follows. Does the amount of complaints that an officer receives increase due to any factors relating to the officer themselves such as their age, rank, gender, ethnicity, or command they report to? Do the complaints received overall tend to lead to action (substantiated), or no action (exonerated and unsubstantiated)? Are there any specific complaint that is received more than others? Do the complaints that are received more often lead to a specific board disposition more often? Are there any factors from the complainant (age, gender, or ethnicity) that appear more often in certain complaints over others? Does the length of time it takes between the date the complaint was received and the date the complaint was closed affect the outcome?

## Reference List

Derek Willis, Eric Umansky and Moiz Syed. The NYPD Files. Search Thousands of Civilian Complaints Against New York City Police Officers. 2020. Retrieved October 7, 2020 from <https://projects.propublica.org/nypd-ccrb/>