

### Data Set Introduction

This data set contains information of complaints made to the NYPD by civilians as a result of their contact with police officers between the years of 1985 and 2020. The bulk of the records look like they come from 2000 and onwards. It is comprised of information on the officer, the complainant, the nature of the contact and the board disposition toward the complaint. Relevant information on the officer includes rank and command at time of incident, current rank and command, officer ethnicity, gender and age. Information on the complainant is comprised of ethnicity and age. Nature of contact includes, the initial reason for contact and the outcome, the allegation type and specific detail, the date complaint received and the date closed. Board disposition is the outcome of the complaint, which generally fall into three categories: exonerated, unsubstantiated, substantiated. The substantiated category includes in parenthesis the disciplinary action.

### Categorical Variable

The categorical variable of the data set is that of the Board Disposition column – however the three general categories contained within it pose a problem, particularly the data points labeled as ‘unsubstantiated’. That is, it is not simply a matter of every allegation resulting in either a substantiated allegation or exonerated. Initial counts of the 33,358 rows identify nearly half of the records fall into the unsubstantiated category. As it does not make sense to include them in the substantiated category, there are two options left to consider. First, combine the unsubstantiated and exonerated into a single ‘not substantiated’ category or completely discard those data points and work only with the binary exonerated/substantiated categories. While neither is ideal, for the sake of this analysis, it would seem to be better to completely disregard them, considering including them into the ‘exonerated’ category has the potential to misrepresent them. That is, while their status at the time of this report is unsubstantiated, the allegation could be completely true, just not proven, which does a disservice to those who reported the incident.

### Research Questions

The main goal of a model based on this dataset would be to identify particular subsets of officers, who could be described as ‘at risk’ which could be useful way of targeting preventative training. For example, can we predict the likelihood that a complaint will be of the substantiated category against a police officer of rank  $s$  in command group  $x$  in age range  $y$ . The goal of the selection of supporting sources would be used to compare if the general trends of the model are in agreement with established research.

- [1]Anthony A. Braga, Christopher Winship, Tom R. Tyler, Jeffrey Fagan, and Tracey L. Meares. 2014. The Saliense of Social Contextual Factors in Appraisals of Police Interactions with Citizens: A Randomized Factorial Experiment. *Journal of Quantitative Criminology* 30, 4 (2014), 599–627.
- [2] Louise Christian. 1989. Police Complaints Authority - Worse Than Useless? *Socialist Lawyer* 7 (1989), 15–17.
- [3] Amanda Geller and Jeffrey Fagan. 2019. Police Contact and the Legal Socialization of Urban Teens. *RSF: The Russell Sage Foundation Journal of the Social Sciences* 5, 1 (2019), 26–49. DOI:<https://doi.org/10.7758/rsf.2019.5.1.02>
- [4] Rory Kramer, Brianna Remster, and Camille z. Charles. 2017. BLACK LIVES AND POLICE TACTICS MATTER. *Contexts* 16, 3 (2017), 20–25. DOI:<https://doi.org/10.2307/26370533>
- [5] Stephen P. Savage. 2013. THINKING INDEPENDENCE: Calling the Police to Account through the Independent Investigation of Police Complaints. *The British Journal of Criminology* 53, 1 (2013), 94–112.
- [6] Graham Smith. 2001. Police Complaints and Criminal Prosecutions. *The Modern Law Review* 64, 3 (2001), 372–392.
- [7] Graham Smith. 2004. RETHINKING POLICE COMPLAINTS. *The British Journal of Criminology* 44, 1 (2004), 15–33.
- [8] Silvia Staubli. 2017. Discussion. In *Trusting the Police*. Transcript Verlag, 151–174. Retrieved October 8, 2020 from <http://www.jstor.org/stable/j.ctv1fxjrg.12>
- [9] Ian Waters and Katie Brown. 2000. POLICE COMPLAINTS AND THE COMPLAINANTS' EXPERIENCE. *The British Journal of Criminology* 40, 4 (2000), 617–638.
- [10] Robert E. Worden and Sarah J. McLean. 2017. Police Legitimacy. In *Mirage of Police Reform* (1st ed.). University of California Press, 42–68. Retrieved October 8, 2020 from <http://www.jstor.org/stable/10.1525/j.ctt1w8h1r1.7>
- [11] Robert E. Worden and Sarah J. McLean. 2017. Reflections on Police Reform. In *Mirage of Police Reform* (1st ed.). University of California Press, 178–196. Retrieved October 8, 2020 from <http://www.jstor.org/stable/10.1525/j.ctt1w8h1r1.14>