

Project Proposal

For our project, we would like to explore the Chicago Police Data Project data provided by The Invisible Institute (<https://invisible.institute/>). As part of this exploration, we would like to combine datasets that are publicly available data from <https://data.cityofchicago.org/> (such as crime data, arrest data, settlement data, etc.) to glean specific insights about the data. Through the use of the CPDP dataset and external datasets such as the Chicago PD Settlement Data, Chicago Crime Data, Chicago Arrest Data, and Socioeconomic Indicators Data, we wish to analyze the level and severity of police intervention in different areas around Chicago and its effects on the community. As part of efforts from an ongoing project in a different class, basic insights about the data will already be available to our group. Examples of questions we will already have answers to include: in general, how does the level of police intervention vary across communities with different socioeconomic statuses? To what degree, if at all, are reports of police brutality handled differently across the varying socioeconomic areas of Chicago?

In our project for this class, we would like to expand the scope of these retrospective questions to make predictions about how crime will be addressed in certain parts of Chicago in the future. More specifically, we would like to build models to clearly answer the following questions:

- Given the socioeconomic status of a neighborhood, predict whether or not the neighborhood suffers from high police brutality rates.
- Given the beat, block, officer rank, subject gender, subject race, predict whether the trr (tactical response report, described further in detail here: <http://how.cdpd.works/glossary/what-are-tactical-response-reports-trr>) report is a case of police brutality.
- Given the crime rate(s) of an area, predict the prevalence of police violence/ brutality.