

Crime in Boston

Problem statement

While we all aspire to minimize crime, it is necessary to first understand the severity of the issue and contributing factors. Pinpointing the exact causes of crime is impossible, as it is a highly nuanced and complex issue. However, factors such as gross income, economic disparity, and government infrastructure/support are often strong indicators. For example, it is widely believed that there's an increase in crime in regions that have a large variance in citizens' income levels.

You will work with datasets drawn from the social world, which are thus influenced by both personal- and institutional-level biases and discrimination. Consider how the processes contributing to the creation of these datasets (calling 911, believing survivors of crime, documenting events, etc.) might make some "crimes" more visible and others invisible.

You should note that academic studies have found evidence that some predictive policing algorithms perpetuate systemic racism in the United States. One criticism of these algorithms is that, instead of predicting criminality, they predict likelihood of arrest - a function of the police presence already in a neighborhood and individual officers' discriminatory policing practices.

You are tasked with inspecting factors that may be correlated and/or contributing to occurrences of Boston crimes and the types thereof (e.g., wire fraud, tax evasion, robbery, vandalism, etc). You will need to carefully consider the ethical implications of your analysis, including your selection of data and methods as well as the possibility of unintended consequences when predictive policing tools are applied in the real world.

Primary project goal: to analyze and model crime in the greater Boston area using the Boston Crime Incidence Reports (along with at least one other geographic location: Cambridge, New York City, etc.) and other data resources helpful to predict crime, understand crime's effect on society, or understand the effects of predictive policing on society.

Data Resources

1. **Crime Incidence Reports (feel free to use other locations as well)**
 - [Boston](#)
 - [Cambridge](#)
 - [New York City](#)
2. **A few selected Boston OpenData resources**
 - [Boston Street Light Locations](#)
 - [Boston Property Assessments](#)
 - [Boston Neighborhood Demographics](#)

High-level project goals

1. Explore, Investigate, and Visualize various factors of the Crime data: minimally, geographic differences and time trends.
2. Describe and incorporate additional data sources that you will use to help you understand or use the Crime data.
3. Train and evaluate models using the Crime data (either as predictors or as the primary response in some fashion) and your own data sources.
4. Use and interpret your models to discuss the causes and correlate of or effects due to crime in the Boston/Cambridge area.
5. Discuss the social and ethical implications of your methods and findings.

References

1. ["The Geography of Violence, Alcohol Outlets, and Drug Arrests in Boston"](#)
2. ["Exploring the Social Construction of Crime by Neighborhood: News Coverage of Crime in Boston"](#)
3. [Do crime hot spots affect housing prices?](#)
4. ["Crime and Property Values in the 1990s"](#)
5. [Can Street Lighting Reduce Crime?](#)
6. ["Predictive policing algorithms are racist. They need to be dismantled."](#)
7. [Race after Technology: Abolitionist Tools for the New Jim Code](#)
8. [UK Government: Report commissioned by CDEI calls for measures to address bias in police use of data analytics](#)