Crime Data Capstone Project Proposal

Phillip Spratling

Few things are as ubiquitous as crime. In any society, ever since there were laws there have been people who break them. Today, crime is such a large problem that billions of dollars are spent on police and crime analysis every year. Crime, if left unchecked, has the possibility to destroy a civilization, and thus it is the responsibility of a society's law enforcement to protect and ensure the safety of its citizens from each other.

Despite having a large police force, Atlanta has some of the highest crime rates in the country today. Forbes recently rated it the 6th most dangerous U.S. city, with a violent crime rate of 1,433 per 100,000 residents. Apart from the high crime rate, Atlanta is also home to some of the nation's best universities. In particular, the Georgia Institute of Technology is located right in the heart of Atlanta. Students come from across the world in pursuit of a higher education, and to do so they must learn to keep themselves safe. Students are advised to follow certain rules such as "Don't walk alone after dark!" or "Stay away from Home Park!". Even when following these rules, sometimes being the victim of a crime is inevitable. Knowing this, the Georgia Tech Police Department (GTPD) do all they can to keep their students safe.

In an effort to make crime around campus more transparent, the GTPD have since 2010 posted crime logs on their website available for public access. Every time an incident is reported, it is added to the log with various attributes such as the time of the incident, the type, and its location. My goal with this project is to analyze this data over the years and determine which factors play a role in predicting crime. I will then build a model to predict future crime. The analysis and predictions will enable the GTPD to better place their officers around campus, as well as help them decide which specific types of crimes to work towards reducing. It will also help Georgia Tech students know how to better keep themselves safe.

All of the data wrangling and analysis will be done with Python. The data will be gathered from the GTPD public website. As demographics of a city are a great predictor of crime, I will also merge demographics data for Atlanta throughout the years. I will begin by gathering and compiling the data, cleaning it, and preparing it for analysis. I will then explore into visualizing and answering the following questions, among others:

- How does crime vary with respect to time/ seasonality/ location etc.?
- How have overall crime rates/ specific types of crime changed over the years?
- How can patrol zones be improved?

After the data wrangling and exploratory data analysis, I will build a model to forecast crimes in the future.

Deliverables:

- All the code written throughout the project delivered in ipython notebooks
- Final Report detailing all my findings and methods
- A PowerPoint presentation