Predicting Crimes in Philadelphia

Fiona Chin (flc38), Erik Enriquez (eee37), Yumeng Niu (yn227)

Problem

Our goal is to analyze Philadelphia crime data to predict future crimes and understand what are the common factors that influence different types of criminal actions. What are the key characteristics in predicting crime?

Dataset

We are primarily using the dataset on Kaggle "Philadelphia Crime Data" (https://www.kaggle.com/mchirico/philadelphiacrimedata). From this dataset, we are trying to examine the historical crime data in Philadelphia from 2006 to 2016 to predict number of criminal incidents in the future, categorized by districts and crime types. The primary factors we want to use includes district, time, latitude, longitude, number of incidents in the past, and type.

We are also planning to incorporate other Philadelphia related datasets into our project (on OpenDataPhilly https://www.opendataphilly.org) to find any significant relationships between crime data and factors such as weather, income, and location. These factors can be useful in predicting crime in specific districts of Philadelphia as well as what characteristics are strongly correlated to the types of crime happening.

Value

The problem is important because predicting crime can help determine the required size and budget of Philly's police department. Predicting crime rates can potentially lower the death toll associated with crime and make Philly a safer place to live.

We believe we are likely to succeed because we are passionate about solving Philly's crime crisis. In addition, we have a fairly large dataset for this problem, which gives us space to fully utilize various algorithms and methods learnt in this class to examine this problem from different angles. Moreover, with the sources of additional related datasets, we are allowed to investigate seemly unrelated parameters and their impact on crime rate.