

Capstone Project 1 Proposal

By: Charlie Lee

Date: 1/17/2020

1. Problem

Crime data are used by law enforcement in ways where it provides predictions for resource allocation, budget formulation, planning, and other various purposes. The crime data benefits politicians, researchers, criminal justice professionals to comprehend crime and society. Also, chambers of commerce and tourism agencies review crime data to see how it impacts the particular geographic jurisdiction they serve at. The crime data are notorious by justice professionals to learn about nature, cause, and movement of crime over time.

With all that said, crime data are used in security and police work in an attempt to reduce criminal activities. Law enforcement can provide safer communities if they can foresee what type of specific crime are likely to occur at a certain time and location. Such crime data can bring us one step closer to prevent crime rather than reacting to them.

2. Client - Northern District of Illinois | Department of Justice

Client will be Northern District of Illinois - Department of Justice who is mainly concerned about the crimes in the city of Chicago.

Based on the data analysis provided, the client will be able to:

1. Budget more accurately
 - a. Accurately budget sufficient money to the proper locations and programs to save money and provide safer communities.
 - b. Help determine which programs in a certain community can receive criminal justice grant.
2. Efficient allocation of resources
 - a. Improve resource allocation during certain time at a given location.
 - b. Determine when and where more or less police officers will be required.
3. Predictive policing and initiative assessment
 - a. Help law enforcement to anticipate increased risk of a certain crime during specific months, therefore by being able to prevent from occurring.

- b. Insight of whether a certain crime increases or decreases during specific months and implement changes if needed.

3. Crime data

Dataset will be provided by the city of Chicago data portal

(<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data>).

Data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. This dataset reflects reported incidents of crime (with the exception of murders) that occurred in the city of Chicago from 2001 to present, minus the most recent seven day. The dataset gets updated on a daily basis.

4. Approach

Dataset will be sliced to analyze years from 2015 to 2018 (3 years).

2019 won't be included because per dataset, these crimes may be based upon preliminary information and may be changed at a later date based upon additional investigation. Instead, 2019 will be solely processed as a test set to see if the data shows trends as predicted. 2019 test set will be acquired at a later date to reduce any false data provided by the preliminary information.

Crime types will be categorized by the same way the city of Chicago reports in their dataset (called "Primary Type" in the dataset column) . These primary categories then will be separated by monthly to see if certain crime occurs more during certain months. Each crime category will be separated and be used as a train and test set. If possible, choropleth map will be used to show geographical information as well.

Data will be presented by:

1. Histogram - counts of certain crime vs. months (y vs. x) over the period of 3 years.
2. Line plot - counts of certain crime vs. months (y vs. x) over the period of 3 years.
Counts of certain crime vs. months (y vs. x) for 3 years combined.
3. Heat map
4. Box plot

5. Deliverable

The final report will be delivered in a PDF format report. The codes will be written in Jupyter notebooks, uploaded to GitHub.