Chicago Crime Analysis:

Crime Data From 2001 – 2023

By

Steven Delaney & William Akers

Description:

For our project we will analyze the Chicago reported crimes data set ranging from 2001 to 2023. With this data set we hope to answer

- Which areas in Chicago are hot spots for violent crime?
- Is there seasonality involved in the frequency of criminal activity?
- Which Districts and Wards have the best success rate at making arrests?
- Is there a correlation between type of crime and location?
- What trends are present in the types of crimes that occur, over the years?

Prior Work:

The University of Chicago Crime Lab has several research papers dedicated to Chicago Crime
https://crimelab.uchicago.edu/resources/2024-end-of-year-analysis-chicago-crime-trends/

The Chicago Police department uses summary statistics to measure success over time
https://www.chicagopolice.org/statistics-data/crime-statistics/

Journal of Data Analysis and Information Processing Vol.12 No.3, August 2024 uses the same data set as our project
https://www.scirp.org/journal/paperinformation?paperid=134329

Dataset:

https://www.kaggle.com/datasets/utkarshx27/crimes-2001-to-present

Our dataset was sourced from Kaggle and downloaded on Steven Delaney's machine. The data itself was collected by the Chicago Police Department and made available on the CLEAR (Citizen Law Enforcement Analysis and Reporting) System

Proposed Work:

Data Preprocessing & Cleaning:

- Outlier Analysis to determine if there are outliers in the data set that may skew results of the analysis.
- Cluster Analysis to determine if there are notable groupings that need to be categorized.
- Check for missing values and handle accordingly

Data Integration: At this stage we are only working with a single data set so integration may not be necessary.

List of Tools:

- **Visualization:** Tableau
- **Analysis:** PANDAS, Numpy, Matplotlib
- **Language:** Python
- **IDE:** VSCode, Jupyter
- **Collaboration:** GitHub, Outlook, Discord

Evaluation:

Cross Reference Prior Research

- Do our findings match those who have researched the same or similar data

Look for correlations that answer our research questions

- Our initial questions will guide our analysis and hopefully provide definitive answers to our initial set of questions

Predictive Modeling

- Using our dataset can we build a model that predicts the location of a crime based on what crime was committed and when.