How Did the COVID-19 Pandemic Affect Crime In Chicago? A Data Analysis Approach to Crime Data Correlations

AZAL ALHADIDI, HOLLAND HYSMITH, and REKHA BHUPATIRAJU

1 INTRODUCTION

The COVID-19 pandemic called for global shutdown in the year 2020, heightening a sense of isolation, fear, and anger among billions of people. With the ongoing friction of shutting down businesses and mask mandates, less people were able to carry out their normal lives. Interestingly enough, crime decreased at a rate of 37% worldwide after stay-at-home issues were ordered [1]. According to the American Law Institute in 1962, crime has been defined as an "offense by the Model Penal Code or any other statute of this State, for which a sentence of imprisonment is authorized" [2]. However, cities like Chicago that accounted for half of the U.S. homicides for 2016, create speculation if the pandemic had an effect on their crime rate as well [3]. Understanding the connections of densely populated cities with crime behavior is imperative to help identify potential motives and prevent further crime from occurring in the future. As machine learning models and data analysis methods mature in the computational landscape, applications in criminal studies have utilized their techniques. A recent study was conducted to analyze crime in South Africa following a linear regression model. This model was designed to help predict crime across all nine provinces dependent on the population, number of police stations, and types of crimes committed [Figure 1]. A separate study observed 2013 crime statistics in the state of Mississippi by utilizing linear regression, additive regression, and decision stump algorithms. The goal of this work was to categorize which cities of Mississippi were more likely to have certain crimes occur and how many crimes happened based upon population [Figure 2]. Therefore, for this work, an array of data analysis techniques will be constructed to correlate crime trends in Chicago during the pandemic. Understanding connections such as locations where crimes were committed, types of crimes, and times in which they occurred can be a helpful resource for understanding how global phenomenon can impact large cities. This information can be used for future events to help mitigate prevention of crime occurring in the city of Chicago.

2 DATA

For this project, the analytical study leverages the Chicago crime data set. Data is obtained from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. This data set is about the reported incidents of different crime types (except murdering) across Chicago state from 2001 to the present, with weekly updates. The data set records may be updated by the authority based on further crimes' investigation further on. This data set

Azal Alhadidi, Holland Hysmith, and Rekha Bhupatiraju

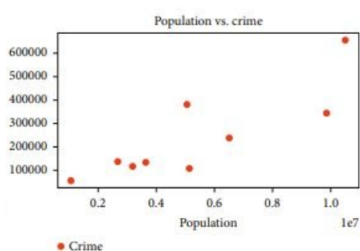Fig. 1. Linear regression approach to crime based on population.

Fig. 2. Results for violent crime per 100k of population.

Algorithm	Correlation Coefficient	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Root Relative Squared Error
Linear Regression Model	1	0.004	0.006	0.0009%	0.001%
Additive Regression Model	0.97	116	168	26%	27%
Decision Stump Model	0.78	276	379	62%	62%

can reflect enormous trends due to impacting factors, including the Covid19 pandemic. Hence, this report will focus on exploring the Chicago crime data during the pandemic and the effect of some applied precautions by the authority (e.g., lock downs and isolation) on the crime rate and type.

The data set consists of several necessary fields related to the reported crimes, including reporting code, type of the crime, demographic location, crime description, and other crime and police-related fields. The data set contains many crime features (about 30 features) and a huge number of records (more than 6 million) as it has accumulated records since 2001. This would make the analytical process difficult and time and computational resources-consuming. Since we intend to explore the effect of Covid19 on the Chicago crime incidents, and since the pandemic spread in the US occurs around December 2019, we will utilize part of the data set related to the crime data set of the last two-three years. Therefore, the primary data set will be diminished by writing a script for shortening it and putting the trimmed data in a new shape to serve our analytical needs

The data set will be analyzed and preprocessed to figure out the exciting patterns using Python Jupyter Notebook with Pandas, Numpy libraries. The Matplotlib library is also utilized for presenting the analytical results. The Chicago crime dataset is available in CSV format on the Kaggle website[6].It is freely available for everybody to use under the following terms provided by the data set source[7].

3 METHODS

Is this an in-depth data analysis or a machine learning problem? What type of ML or data analyses techniques might you utilize for this study? For ML problems, is there a simple baseline that you can use as a starting point? What metric will you use to measure success? Is there a definition of "success" (especially important for unsupervised/data analysis problems)?

4 CONCLUSION

This study will provide a quantitative view of the effect of the pandemic's impact on crime rates and response. If we find no statistically significant patterns, this could be indicative of a lack of an effect that could still tell us something about the population's behavior. Any result could be informative for governments and other organizations as it could inform their response to future events.

biblatex test.bib

Let's cite! Einstein's journal paper [?] and Dirac's book [?] are physics-related items.

How Did the COVID-19 Pandemic Affect Crime In Chicago? A Data Analysis Approach to Crime Data Correlations

REFERENCES

105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156