Analyzing the Impact of Economic Recessions on U.S Crime Rates and Financial Markets

Kiet Nguyen Shy-Lee Ben Ezer Kunj Patel

November 26, 2024

Abstract

This analysis explores the relationship between economic recessions and crime rates in the U.S., focusing on three distinct periods: the DotCom Bubble (2000), the 2008 Financial Crisis, and the COVID-19 Pandemic. Leveraging financial and crime datasets, the analysis reveals patterns of crime fluctuation during economic downturns, underscoring socioeconomic impacts. Key findings suggest that economic stress influences certain crime types, with complex variations observed across recession periods. The study aims to provide insights for policymakers and future research.

Keywords: Economic Recession, Crime Trends, Financial Markets, Socioeconomic Analysis, Economic Policy, Stock Market Volatility, Criminology, COVID-19 Pandemic, DotCom Bubble, 2008 Financial Crisis

Contents

1 Introduction

According to investigative reporter Steve Adams [2], crime and economics have a consistent relationship. However, he adds that crime rates have been continuously increasing for decades. Nonetheless, he reports a relative decrease in crime rates during the 2008 recession but an increase in crime rates in specific crime types during that time. He suggests this is most likely due to socioeconomic changes rather than the economy.

We decided to investigate the changes in crime rates during those economic periods and see if socioeconomic changes affected them. For example, after George Floyd's death in 2020, the Black Lives Matter movement was established. While the crime rate does not directly correlate with the COVID-19 pandemic, there was a rise in property-related crimes during that period.

The relationship between economic recessions and crime rates is a very popular branch of research within criminology and economics because it reflects the performance of 500 of the largest publicly traded companies in the U.S., across various sectors, including technology, healthcare, and financials. As the S&P 500 serves as a benchmark for overall market health, it allows us to gauge the stock market's response to financial shocks, making it a critical tool for understanding broader economic conditions.

We focus on three key eras—the DotCom Bubble (2000), the 2008-2009 Financial Crisis, and the COVID-19 Pandemic—because each represents a distinct type of economic shock, including market speculation, systemic financial collapse respectively. Analyzing these eras provides insights into how different types of financial crises affect both markets and society, offering lessons for future economic policy and resilience planning.

A number of studies that examined crime trends during the COVID-19 pandemic, such as Yang et al. [8] and Campedelli et al. [8], discovered variations in crime patterns as a result of normal activity modifications, containment measures, and economic strain. These results are consistent with Ashby's [4] study that used SARIMA models to estimate what crime rates may have been in the event that there had not been a pandemic. The study found that whereas some crimes (like robberies) dropped, other crimes experienced less change. Quednau [16] strengthens the idea that financial stress encourages criminal conduct by highlighting the connection between poverty and violent crime.

2 Exploratory Data Analysis

2.1 Financial Dataset

For this project, financial datasets are sourced from multiple sources, including Kaggle [5, 15], the Federal Reserve Economic Data [12], and Yahoo Finance [19]. These sources provided historical data on GDP, unemployment rates, and other economic indicators crucial to our analysis.

- **SP500:** Tracks the S&P 500, including daily open, close, highs, lows, and volume for 500 major U.S. companies. Updated daily.
- **Interest Rates (DGS10):** 10-Year Treasury rate, reflecting long-term interest rates and market sentiment on the economic outlook. Updated daily.
- **Inflation Rates (CPIAUCSL):** Consumer Price Index for urban consumers, measuring inflation via price changes in goods and services. Updated monthly.
- **GDP Growth Rates (A191RL1Q225SBEA):** Quarterly Real GDP growth, adjusted for inflation, showing economic expansion or contraction. Updated quarterly.
- **Unemployment Rates (UNRATE):** U.S. unemployment rate, indicating labor market health by the percentage of job-seeking unemployed individuals. Updated monthly.
- **Consumer Sentiment (UMCSENT):** Measures consumer confidence, which can signal future consumer spending and economic optimism or pessimism. Updated monthly.

- **VIX:** The "fear gauge," indicating 30-day market volatility expectations; higher values suggest greater market uncertainty. Updated daily.

We will look at the pre-, mid-, post-recession trend of each recession:

2.1.1 The Dot Com Recession (2001)

Figure 1: Financial Data Visualization - Part 1.

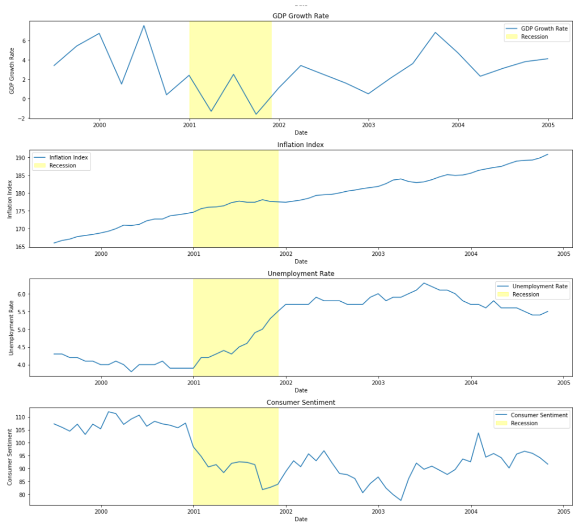

Figure 2: Financial Data Visualization - Part 2.

SP500 and VIX: Even before the recession officially started, we can see that the decreasing trend started to build up in the middle of 2000. Suggested by Langdon, McMenamin, Krolik, such downtrend matches service-producing sector downtrend, which led to the 2001 recession. The daily price difference within a day also builds up during mid 2006. During the recession, we can see that these trends became more clear, regarding an increase in fluctuation of price difference, trading volume and close price. Post-recession, not until mid 2003 did we achieve a more stable market and positive changes. From early 2002 to 2003, the effect of the economic recession worsened. VIX shows an even worse trend than the recession period.

Interest Rate: Before the recession happened, the interest rate was pretty high, forecasting a positive market. In mid-2000 the interest started a dropping trend. During the recession, the fluctuations were still stable, with a margin of change being less than 1%, even lower than the pre-recession process. Post-recession, within the 1-year span of 2002, the downtrend was much worse. This may be the government's effort to boost the economy, which led to a recovery at the end of 2003 / early 2004.

GDP Growth Rate: Before the recession, the growth rate fluctuated in 2000. The greatest decrease was around 7%. During the recession, it is clear that the recession is associated with a negative growth rate. Post-recession, the positive growth rate indicated that the U.S. had stepped out of the recession period.

Unemployment Rate: There was not a clear indicator to help forecast the unemployment. However, according to Langdon, McMenamin, Krolik, "in a short 10-month span—between March 10, 2000 and the end of 2000—the NASDAQ plunged more than 50 percent" to finance Internet startups. We can see a steep increase in the unemployment rate. The hype of Internet investment became "symptomatic of broader economic ills, which hurt the overall securities industry's revenues." In the personnel supply services, 556,000 jobs got their payroll cuts; employment in security and commodity brokerages shrank by 23,000. Post-recession, the rate of unemployment started to slow down, yet still gradually increased until mid-2003 when there were some positive decreasing trends.

Consumer Sentiment: Sharp decrease during the recession and only recovered post-recession in mid-2003. It seems like this recession has shown some forecastable signs during the mid of 2000 and achieved a recovery during mid-2003, given the patterns of changes.

2.1.2 The Housing Bubble (2007-2009)

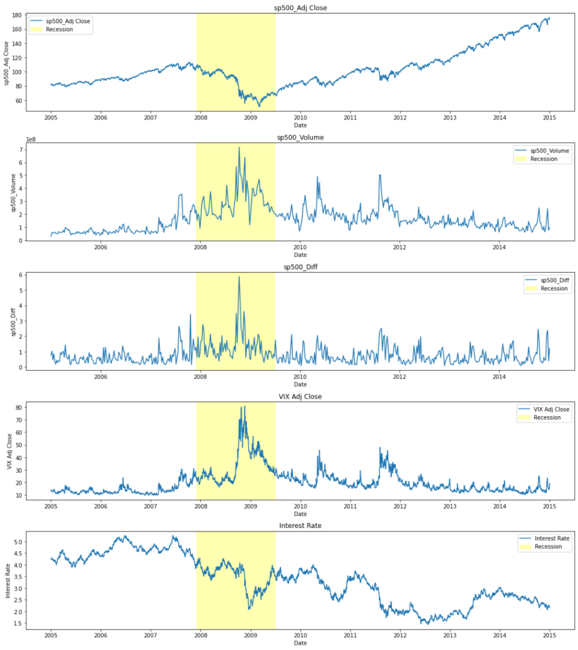

Figure 3: Housing Bubble Visualization - Part 1.

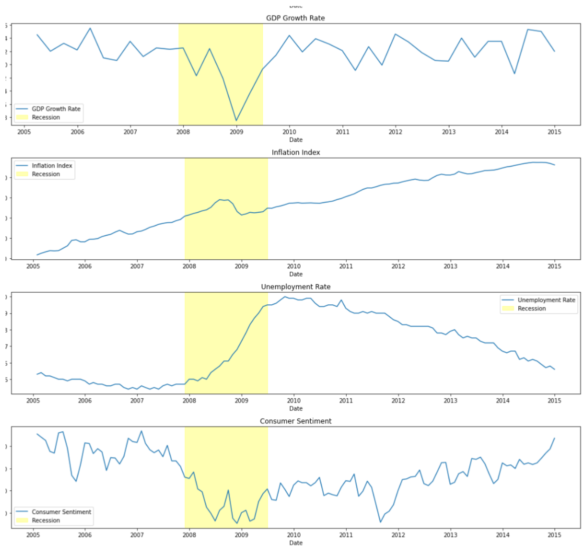

Figure 4: Housing Bubble Visualization - Part 2.

S&P 500 and VIX: Before the recession, the price of S&P 500, trading volume, and price fluctuation built up faster since the middle of 2006. VIX showed some signs of increasing but was not clear until the beginning of 2007. During the recession, the peaks of all trends were during the middle of 2008. Several reasons were prompted to explain this. Firstly, Lehman Brothers declaring bankruptcy triggered panic throughout the U.S. and global financial market. Secondly, the securitization system that generated mortgage-backed securities (MBS) from mortgages became opaque and unclear. Two major financial institutions, Fannie Mae and Freddie Mac, made loans to a lot of low-credit applicants. Together, they held 1.5 trillion dollars in bonds and thousands of billion dollars in issued MBS. These explain the sharp increase in VIX and drop in S&P 500 performance. Post-recession, both S&P 500 and VIX performance suggested that the financial market only recovered in early 2012, where the fluctuation became more stable.

Interest Rate: Before the recession, the interest rate was pretty stable and only showed some suspicious signs toward the end of 2007. It seems like the financial market was confident in that current growing state of the market. During the recession, the interest rate started to drop fast. The previous period with high interest rates finally came back in this period, especially in the housing market as families could not keep up with the payment of their mortgage. Interest rate drops indicated the pessimistic view of the financial market. Post-recession, the interest rate started out stable and became lower over the year until mid-2013. Interest rate increase in mid-2013 indicated the market somewhat recovered.

GDP Growth Rate and Inflation Index: Before the recession, both had a very expectable and stable pattern. During the recession, a sharp decrease peaked in early 2009. This reflected the effect of the above crisis. Right after the recession, both seem to have become fairly stable.

Unemployment Rate: Before the recession, the unemployment rate was on the way to decrease and showed no sign of the market recession. This shows how great the expectation of the market performance during that time. During the recession, we can see a steep increase, almost doubled just in the span of two years. According to the BLS, ”in late 2009, more than 15 million people were unemployed. Total employment, as measured by the Current Population Survey (CPS), dropped by 8.6 million, or almost 6 percent.” Post-recession, the unemployment stayed relatively high showing the market inability to adjust. Only until 2011 did the trend cool down.

Consumer Sentiment: Drop dramatically during the recession and only became more stable toward the end of 2014. This recession was more visible in the financial market and showed some sign during mid-2006. Even though it was announced to end in mid-2009, the recovery seems to take until 2014, 5 years later.

2.1.3 The Covid-19 Recession

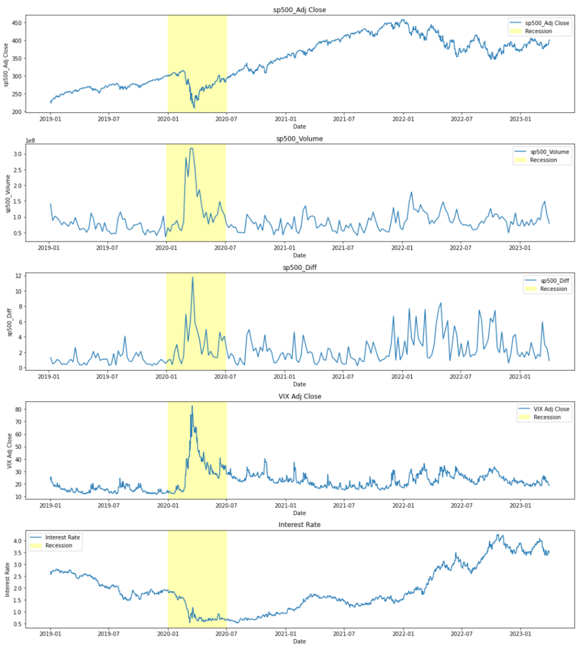

Figure 5: Covid-19 Data Visualization - Part 1.

Figure 6: Covid-19 Data Visualization - Part 2.

S&P 500 and VIX: During the 2020 recession triggered by the COVID-19 pandemic, the S&P 500 experienced a rapid, steep decline in early March, while the VIX spiked to record levels as investor uncertainty surged. The pandemic brought unprecedented economic shutdowns and disruptions across industries, leading to massive sell-offs. Post-recession, thanks to the Federal Reserve’s swift response with aggressive monetary easing, coupled with government stimulus packages, helped stabilize the market, leading to a sharp recovery by late 2020.

Interest Rate: During the recession, interest rates fell to near-zero levels in 2020 as the Federal Reserve moved quickly to counter the economic impact of the pandemic. This drastic drop reflected the Fed’s attempt to stimulate borrowing and spending, providing liquidity to the economy. Post-recession, rates remained low, with gradual rate increases.

GDP Growth Rate and Inflation Index: During the recession, the GDP growth rate saw one of its sharpest contractions in modern history during Q2 2020, as lockdowns and restrictions heavily curtailed economic activity. Post-recession, both GDP growth and inflation showed signs of recovery by late 2020. However, it seems like the inflation grows quite fast ever since.

Unemployment Rate: During the recession, unemployment skyrocketed in early 2020, reaching levels unseen since the Great Depression, with over 20 million jobs lost within weeks. Post-recession, the rate declined as the economy began to reopen and government support programs were implemented, but it remained elevated into early 2021.

Consumer Sentiment: Consumer sentiment plunged during the early months of the pandemic due to widespread uncertainty, job losses, and health concerns. Sentiment improved gradually as government support programs took effect and vaccine rollouts began, with more stability returning by late 2021.

2.2 Crime Dataset

For this part of the analysis, we obtained crime data from multiple sources, including the NYPD arrest records and detailed crime records from Chicago and Los Angeles, each dataset offering a unique perspective on crime during economic downturns. [6, 18]

Here are the sources: NYPD Arrest Data, Chicago Crime Data, LA Crime Data.

- **ID:** Unique identifier for each crime record.
- **Case Number:** The Chicago Police Department RD Number (Records Division Number), unique to each incident.
- **Date:** The date when the incident occurred, sometimes estimated.
- **Block:** The partially redacted address where the incident occurred, provided to the nearest block for privacy.
- **IUCR:** Illinois Uniform Crime Reporting code that is linked to the primary type and description of the crime.
- **Primary Type:** The primary description of the crime based on the IUCR code.
- **Description:** A more detailed description or subcategory of the crime.
- **Location Description:** A description of the location where the incident occurred.
- **Arrest:** Indicates whether an arrest was made.
- **Domestic:** Indicates whether the incident was related to domestic violence as defined by the Illinois Domestic Violence Act.
- **Beat:** The smallest police geographic area where the incident occurred.
- **District:** The police district where the incident occurred.
- **Ward:** The ward (City Council district) where the incident occurred.
- **Community Area:** The community area where the incident took place.
- **FBI Code:** Crime classification as per the FBI's National Incident-Based Reporting System (NIBRS).
- **X Coordinate:** The x coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection, adjusted for privacy.
- **Y Coordinate:** The y coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection, adjusted for privacy.
- **Year:** The year when the incident occurred.
- **Updated On:** The date and time when the crime record was last updated.
- **Latitude:** Geographic latitude of the incident, adjusted for privacy.

- **Longitude:** Geographic longitude of the incident, adjusted for privacy.
- **Location:** The location where the incident occurred in a format that supports mapping and geographic operations within data portals, adjusted for privacy.

The analysis of Chicago's crime data from 2001 to 2024 examined the effects of economic conditions on society. After being cleaned and classified by kind of crime, the data was examined to find patterns over time.

Crime Rates per Year in Chicago: There is a declining trend with notable jumps during the 2008 housing bubble crisis and the COVID-19 outbreak, pointing to a likely link between economic stress and crime rates. Seasonal trends and other changes were also noticeable, emphasizing the need for more detailed time-series study.

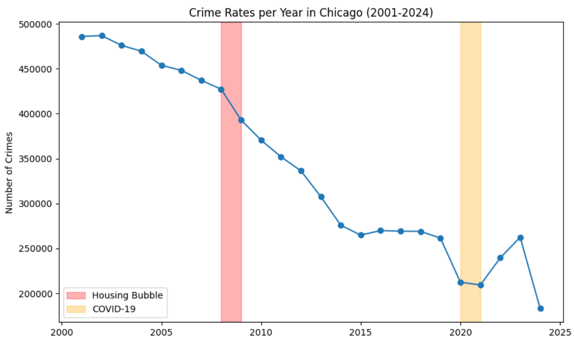

Figure 7: Crime Rates per Year in Chicago (2001-2024).

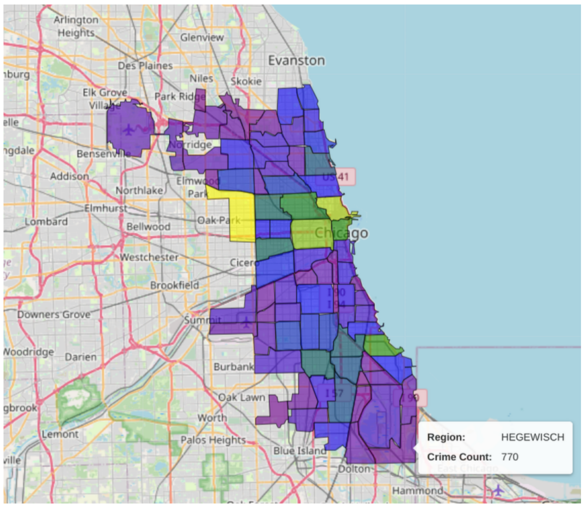

Figure 8: Crime Hotspots by Area in Chicago (2023).

Crime Hotspots by Area in Chicago - 2023: This map depicts crime distribution in Chicago in 2023. The color gradient, which runs from purple to yellow, depicts locations with variable crime densities, making it easier to identify the zones with the greatest and lowest crime rates.

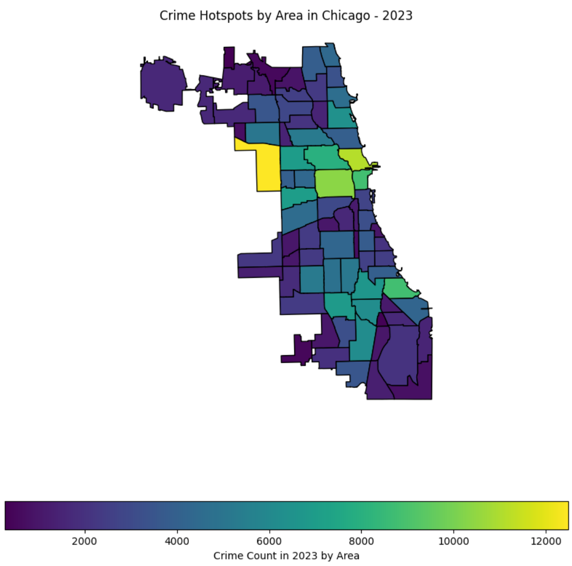

Figure 9: Crime Hotspots by Area in Chicago - 2023

NYPD Arrests by Crime Type Over Time: This line chart depicts the number of arrests for three categories of crimes: homicide, burglary, and robbery, as reported by the NYPD. The trend indicates a relatively consistent number of homicide arrests, depicted by the blue line, which is substantially lower than burglary and robbery rates. The green line represents burglaries and shows a continuous reduction over time, but the red line for robberies shows variation with a noteworthy uptick towards recent dates, illustrating changes in crime patterns and law enforcement activities over time.

2.3 Crime and Economy

2.3.1 Housing Bubble:

During this period, a surprising trend emerged: a significant decrease in crime rates, counter to expectations that economic stress might fuel criminal activity. This decline, observable across 6 out of 7 categories, stemmed from complex interactions of economic, social, and demographic factors. "For 2009, the FBI reported an 8 percent drop in the nationwide robbery rate and a 17 percent reduction in the auto-theft rate from the previous year" [10].

One key explanation is that fewer opportunities for crime emerged during this period. Economic disruptions from the housing bubble kept many people at home, reducing the likelihood of property crimes such as theft and burglary. Home security measures also improved, making homes harder targets for criminals. Additionally, technological advancements like mobile phones and security cameras acted as strong deterrents [17].

The decline in crime during this period also reflected demographic shifts. The aging population reduced the proportion of young males, the demographic most associated with criminal activity. The Stanford report highlights that crime rates often decline with the maturation of a population, as older individuals are statistically less likely to commit crimes [11].

While the bubble's prosperity contributed to declining crime, it also created dynamics that shifted criminal activity patterns. For example, during economic booms, drug-related and violent crimes often decline as illicit markets contract. During the housing bubble, urban areas experienced gentrification, reducing crime-prone environments by displacing vulnerable communities to less crime-ridden suburbs [14].

Punishment and incarceration trends also played a significant role in the crime decline. During the recession, the U.S. correctional population grew dramatically, peaking at over 7.3 million under correctional supervision, compared to 1.8 million in 1980 [11]. This effect came at a significant societal cost, disproportionately impacting marginalized communities and sustaining stark racial disparities in incarceration. While punishment strategies have been a factor, criminologists argue they cannot fully account for the broader, sustained crime decline [1].

This highlights that crime does not consistently align with economic cycles. Studies show that during the recession, people spent more time at home due to high unemployment, reducing opportunities for certain crimes. Despite financial struggles, a combination of social cohesion, technological deterrents, and high incarceration rates helped maintain low crime levels [16].

2.3.2 COVID Recession:

During the COVID-19 pandemic, crime was vastly impacted, leading to notable shifts in various types of criminal activity [8]. The most significant increase in crime type was homicide rates, which increased by nearly 30% during the pandemic. Experts correlate this attribute to several factors, such as social and economic stress, disruption of community programs, and reduced police presence during the pandemic. However, this increase is not the same for other crime types.

Despite the increase in violent crimes such as homicide, the overall violent crime rate dropped by about 22%. That includes sexual assault, rape, and robbery, which is most likely due to the lockdown (i.e., limited interactions in public spaces) [4].

There was a rise of about 9% in auto theft crimes since people had limited activity to monitor their vehicles during the lockdown (as it was unnecessary to routinely check your car). Hence, there was more opportunity for criminals to steal cars, which again, could be due to lack of police presence [17].

There was a significant rise in hate crimes against Asian-Americans because of harmful rhetoric linking COVID-19 to China. That caused a rise in bias-motivated violence against said ethnic group. Furthermore, there was a relative increase in hate crimes against Black Americans, highlighting how the pandemic's tension attributed to social tensions [13].

However, while economic hardship could have been a source for the changes during the economic crisis of COVID-19, experts suggest that the unique circumstances of the pandemic could challenge the correlation between the recession and crime rates [2].

3 Methodology

3.1 Research Question 1

How did major financial events (such as the 2008-2009 Crisis, COVID-19 pandemic, and DotCom Bubble) affect GDP and the S&P 500, and what patterns of recovery can be observed?

To address this question, we utilize algorithms tailored to the goal of binary classification to analyze the impact of these financial events on economic indicators. The following modeling techniques are employed:

- **Logistic Regression:** This method provides coefficients for each indicator, allowing us to learn how each variable impacts the economy's state by estimating the probability of a recession.
- **Random Forest Classifier:** This model offers a neutral perspective on the importance of each indicator due to its robustness against overfitting. It is particularly effective for understanding non-linear relationships in the data. Feature importance scores can further highlight which indicators are most helpful in predicting economic states.
- **Gradient Boosting Classifier:** This model excels at capturing non-linear relationships between indicators and provides insights into the relative importance of features.
- **Long Short-Term Memory (LSTM):** As a black-box model designed for time-series data, LSTM leverages a three-gate architecture (Forget Gate, Input Gate, and Output Gate) to maintain long-term and short-term dependencies effectively. This structure mitigates issues like vanishing or exploding gradients often seen in traditional recurrent neural networks.

Each model is evaluated for its performance in identifying recessionary periods and patterns of economic recovery based on the behavior of GDP and S&P 500.

3.2 Research Question 2

How have economic recessions influenced crime rates across the United States, and what patterns of change occurred during and after these periods?

To analyze this question, we used crime data from Chicago spanning 2001 to 2024. The dataset focuses on monthly crime counts, enabling us to observe trends over time, particularly during periods of economic downturn and subsequent recovery.

SARIMA Model: The Seasonal Autoregressive Integrated Moving Average (SARIMA) model was trained on historical crime data to capture seasonality and trends. Forecasts were generated for the period between January 2001 and December 2025. The model produced:

- A significant spike in crime around 2001.
- A general downward trend in crime rates post-2009, corresponding with the recovery from the 2008 financial crisis.

BSTS Model: The Bayesian Structural Time Series (BSTS) model was employed to investigate the relationship between economic variables (e.g., GDP growth) and specific crime types such as theft. This model incorporates prior beliefs and quantifies uncertainty in its forecasts, making it a powerful tool for analyzing historical data and predicting trends.

The analysis using these models revealed the following:

- Crime rates, particularly theft, are influenced by broader economic conditions.
- During recessions, GDP growth and unemployment rates significantly affect theft rates.

The SARIMA model successfully captures seasonality and patterns in crime data across time, whereas the BSTS model uses probabilistic reasoning to investigate how economic variables such as GDP growth and unemployment affect crime rates, particularly stealing. Combined, these techniques provide a solid framework for analyzing crime trends during economic downturns, with room for improvement through further data and model adjustment.

3.3 Research Question 3

How do crime rates correlate across different states during economic downturns, and can we develop a predictive model to forecast crime trends based on these factors?

In the literature regarding crime rates during inflation times, we learned that there is a decrease in the general crime rate; however, according to a study by Rosenfeld, during the 2008-2009 inflation, there was an increase in three crime types: robbery, burglary, and homicide. Furthermore, according to Pequero et al., during the COVID-19 pandemic, there was a rise in domestic violence cases as a direct response to the mandated lockdown. Thus, we attempt to record those specific crime types and total crime in general according to the associated periods.

In our modeling techniques, we use two types of questions. One is polynomial regression, and the other is logistic regression. The polynomial regression was significantly more successful in predicting our crime rates over time than the logistic regression, as follows:

- **Polynomial Regression:**
  - **NYPD:** Mean Squared Error: 0.00039963219251845167 R^2 Score: 0.999999998914167
  - **Chicago:** Mean Squared Error: 0.000037502136016184204 R^2 Score: 0.9999999999773476
  - **LA:** Mean Squared Error: 0.000000036254297794709 64 R^2 Score: 0.999999999997635

- **Random Forest:**
  - **NYPD:** Mean Squared Error: 50899.85471231627 R^2 Score: 0.6089480333684092
  - **Chicago:** Mean Squared Error: 34.84000570912033 R^2 Score: 0.0412860786984286
  - **LA:** Mean Squared Error: 2220.1402847619047 R^2 Score: 0.773303035657138

As shown in the results above, the polynomial regression model consistently outperforms the random forest model, particularly for the Chicago and LA datasets, where the R^2 score approaches nearly perfect prediction accuracy. In contrast, the random forest model demonstrates weaker performance, especially for the NYPD and Chicago datasets.

4 Data Visualization and Analysis

4.1 Research Question 1

The first thing I noticed about this question was that the dataset is highly skewed and biased toward the non-recession label (class 0). With 11,523 data points, only 1,146 of them are labeled as recession (class 1). This means that if the model has to guess randomly, the chance that it labels a period as recession will only be: [12, 14, 19]

$$\frac{1,146}{11,523} = 9.95\%.$$

This will serve as our lower bound to track our performance.

In deciding which performance metric should be prioritized, I believe recall (how many recessions are predicted) is a better bet. The reason is that, when predicting a recession, we want to be able to classify as many recessions as possible. This will help us have more time and allocate more resources to mitigate the effects of the recession. Precision (how many predicted recessions are actual recessions), on the other hand, may not be as useful, as the cost of misclassifying a recession is not as severe as missing a recession.

For such an imbalanced dataset where the cost of a false positive (a predicted recession is not an actual recession) and false negative (a recession is misclassified to non-recession) is asymmetric, PR-AUC is a better indicator than ROC-AUC as it looks at the precision-recall curve. At this point, we use 0.5 as the threshold for our model when evaluating metrics such as precision, recall, etc.

Fine-tuning process: We applied feature engineering to mitigate the noise by removing unnecessary columns such as 'sp500_Open', 'sp500_High', 'sp500_Close', 'sp500_Low', 'VIX High', 'VIX Close', 'VIX Low', and 'VIX Open'. At the same time, we added columns such as 'VIX Diff' and 'sp500_Diff', which indicate the daily change of these two indices. We also applied 'class_weight' in favor of the minority class. For hyperparameter tuning, we focused on finding the hyperparameters that maximize the recall of the model.

Post improvement, here is the performance of each model:

- **Logistic Regression:**

```
PR-AUC: 0.6673431892516889
              precision    recall  f1-score   support

           0       1.00      0.67      0.80      4432
           1       0.11      1.00      0.20       182

    accuracy                           0.69      4614
   macro avg       0.56      0.84      0.50      4614
weighted avg       0.96      0.69      0.78      4614

Best Hyperparameters: {'C': 0.009, 'class_weight': {0: 1, 1: 9}, 'max_iter': 7000, 'solver': 'liblinear'}
```

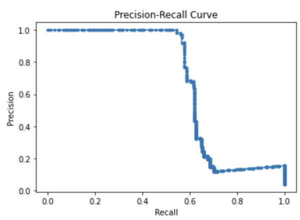

Figure 10: Logistic Regression: PR-AUC and Classification Metrics.

- **Random Forest Classifier:**

PR-AUC: 0.8656540088266116

	precision	recall	f1-score	support
0	1.00	0.94	0.97	4432
1	0.43	1.00	0.60	182
accuracy			0.95	4614
macro avg	0.71	0.97	0.78	4614
weighted avg	0.98	0.95	0.96	4614

Best Hyperparameters: {'max_depth': 15, 'max_features': 'log2', 'min_samples_leaf': 1, 'min_samples_split': 10, 'n_estimators': 75}

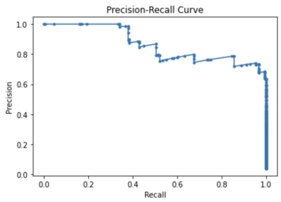

Figure 11: Random Forest Classifier: PR-AUC and Classification Metrics.

- **Gradient Boosting Classifier:**

PR-AUC: 0.8807365969683628

	precision	recall	f1-score	support
0	1.00	0.96	0.98	4432
1	0.51	1.00	0.67	182
accuracy			0.96	4614
macro avg	0.75	0.98	0.83	4614
weighted avg	0.98	0.96	0.97	4614

Best Hyperparameters: {'learning_rate': 0.05, 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 4, 'min_samples_split': 10, 'n_estimators': 200}

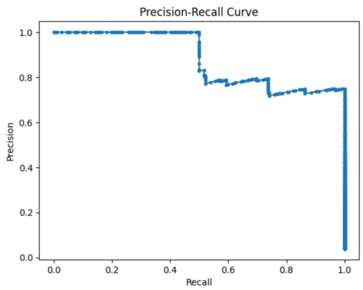

Figure 12: Gradient Boosting Classifier: PR-AUC and Classification Metrics.

- **Long Short-Term Memory (LSTM):** Works best with the RMSprop optimizer.

PR-AUC: 0.9979689366786141

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1511
1	1.00	0.98	0.99	135
accuracy			1.00	1646
macro avg	1.00	0.99	0.99	1646
weighted avg	1.00	1.00	1.00	1646

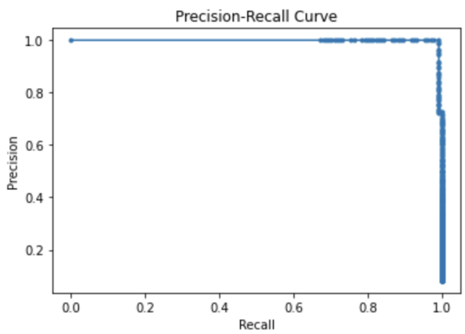

Figure 13: Long Short-Term Memory (LSTM): PR-AUC and Metrics.

4.2 Research Question 2

For this section of the research, we used two separate models to look at the influence of economic downturns on crime trends. These models help us understand how crime patterns change over time, as well as how recessions like the 2008 Financial Crisis and the COVID-19 epidemic affect crime rates. [7, 18]

SARIMA model: Employed to forecast the crime rates over time, taking into account the seasonal nature of crime. The model was fit using historical crime data from Chicago between 2001 and 2024. The plot of the SARIMA forecast provides an overview of the observed and predicted crime rates during the period under analysis.

The plot's colored zones show these recessionary times, with crime dropping significantly during the 2008 Financial Crisis and again during the COVID-19 epidemic. This shows a possible link between economic downturns and lower crime rates; however, certain crime categories may have responded differently.

This time series analysis strengthens our findings by showing crime rates on a monthly basis. The vertical bands in this picture represent the 2008 financial crisis and the COVID-19 epidemic. The pattern demonstrates an instantaneous decline in crime following the start of the Financial Crisis, and the same tendency continues during the COVID-19 epidemic. This shows a substantial short-term criminal response to economic hardship, which was most likely influenced by a variety of social and economic variables throughout these times.

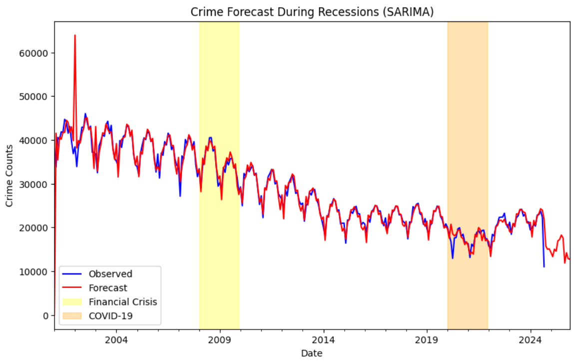

Figure 14: Crime Forecast During Recessions (SARIMA).

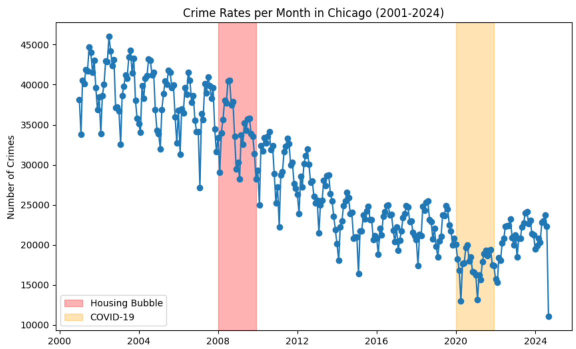

Figure 15: Crime Rates per Month in Chicago (2001-2024).

The chart below shows how crime growth rates (year-over-year) differed between recessionary and non-recessionary eras. The crime growth rate is noticeably negative during recessions, as seen by the red bars, particularly in 2009 and 2020. This lends credence to the theory that economic downturns result in decreased crime rates in some situations. In contrast, non-recession eras (blue bars) show greater positive growth, demonstrating that crime rates have recovered following recessions. The steep decreases observed during the COVID-19 epidemic suggest that lockdowns and restricted social interaction had a significant impact on crime rates.

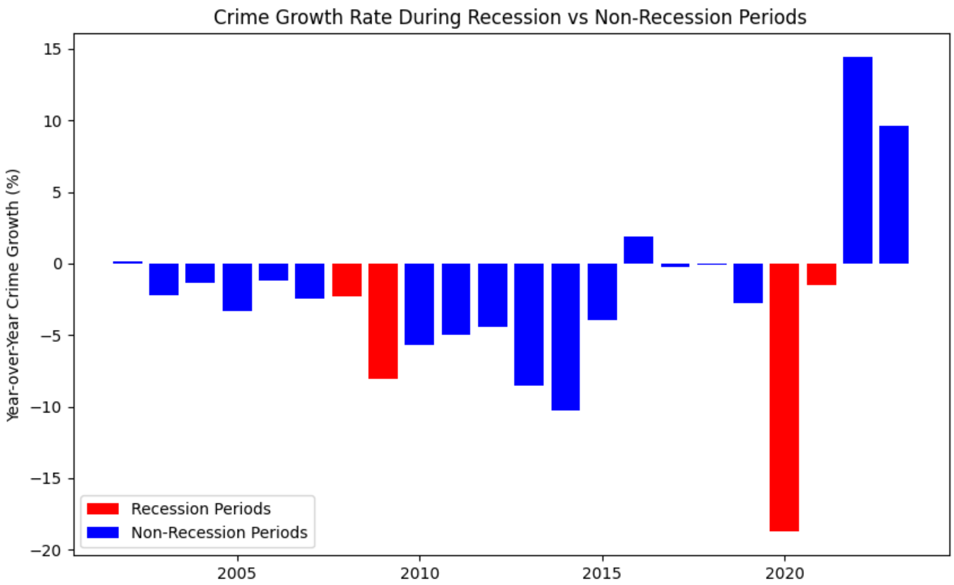

Figure 16: Crime Growth Rate During Recession vs Non-Recession Periods.

BSTS model: Implemented to better understand the link between theft offenses, GDP growth, and unemployment rates. This model contributed to capturing the dynamic impacts of economic variables on crime, allowing for posterior inference on the underlying economic determinants.

In the BSTS model, we examined the impact of GDP growth and unemployment rates on stealing offenses. The posterior distributions for `beta_gdp` and `beta_unemployment` show that both economic factors have a substantial influence on theft rates, with negative coefficients for both, implying that theft rates are negatively connected to GDP growth and unemployment rates. The model's credible intervals demonstrate the strength of these associations.

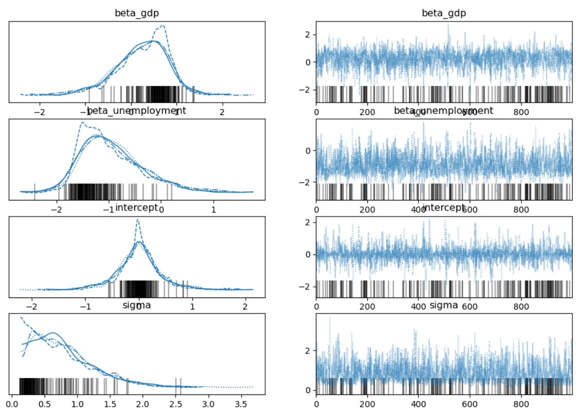

Figure 17: Trace plots and posterior distributions for the BSTS model parameters.

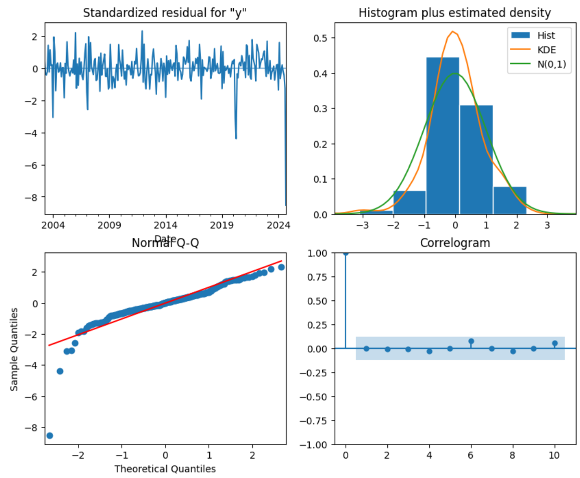

Figure 18: Diagnostics of residuals for the SARIMA model

Our data reveals a consistent link between economic downturns and crime rates. Crime rates typically fall during recessions, as seen in the 2008 Financial Crisis and the COVID-19 epidemic. The SARIMA model was excellent in forecasting these trends, but the BSTS model helped us understand the underlying economic causes of these crime patterns. These findings can help policymakers and law enforcement organizations plan for future economic downturns and their possible influence on crime rates.

4.3 Research Question 3

For this part of our research, we used polynomial regression and random forest to see if we can predict crime over time. The random forest method failed to predict anything, but the polynomial regression of degree 5 successfully predicted the crime with deficient mean squared error and R^2 scores.

While the polynomial regression was successful in predicting the data, it was based only on the sum of crimes per day. Hence, we cannot produce a graph to show the success of the model on our data. We will showcase here several examples of how different crime types changed over time in New York, Chicago, and Los Angeles. In New York and Chicago, we looked at the data quarterly, and in Los Angeles, we looked at the data monthly.

NYPD: In the NYPD dataset, we can see a rise in all inflation periods (except for one), which corresponds with the literature.

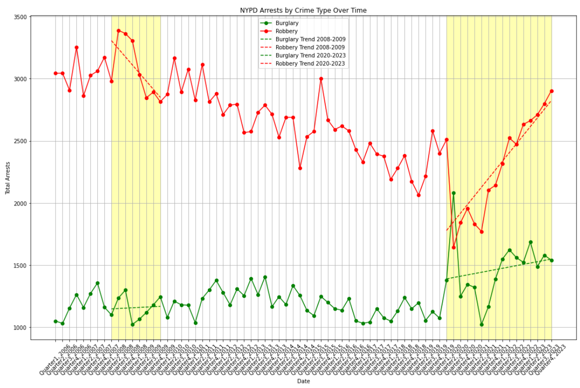

Figure 19: Burglary and Robbery Trends During Recession Periods.

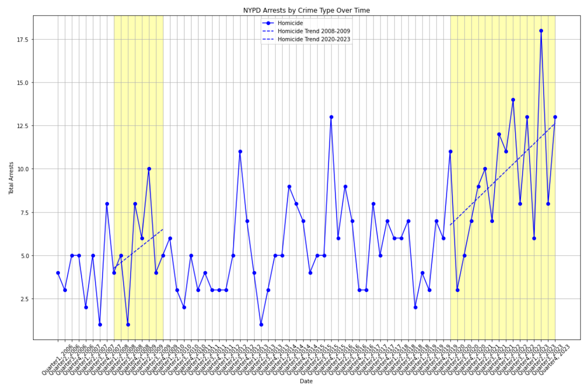

Figure 20: Homicide Trends During Recession Periods.

Chicago: The Chicago dataset did not have the same offense specification of "homicide" arrest like the other datasets; hence, we did not research homicide offenses in Chicago. Nevertheless, during the inflation period, we see a decrease in burglary and robbery crimes, which is the opposite of the literature.

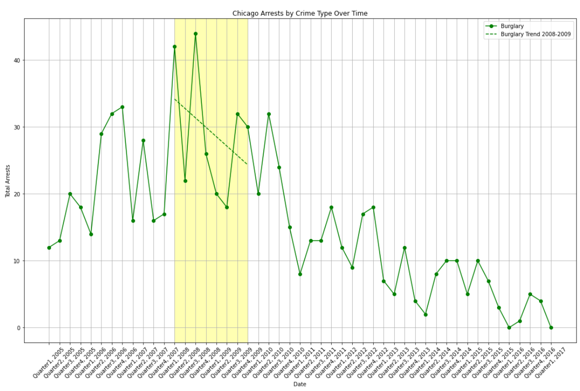

Figure 21: Burglary Trends During the 2008 Financial Crisis.

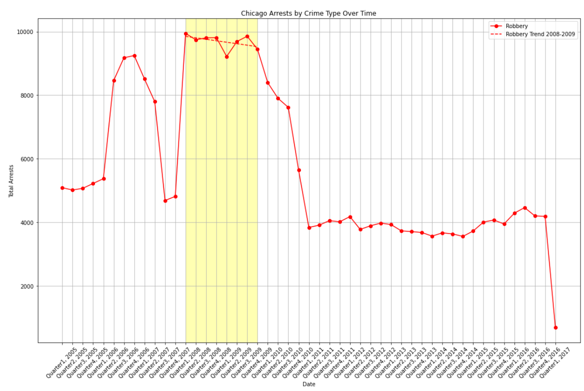

Figure 22: Robbery Trends During the 2008 Financial Crisis.

Los Angeles: Despite the information in the literature, the LA dataset failed to show a rise in robbery and burglary crimes during the lockdown; however, we can see a rise in homicide cases, which corresponds with the literature.

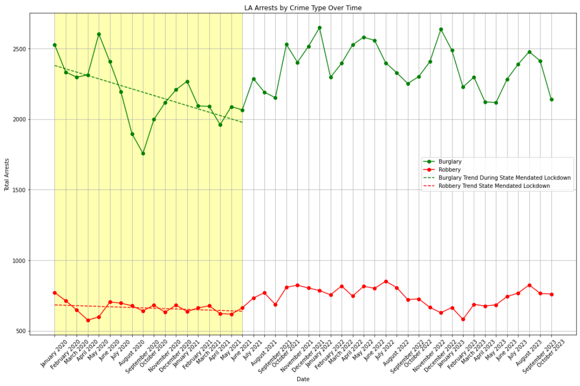

Figure 23: Burglary and Robbery Trends During State-Mandated Lockdown.

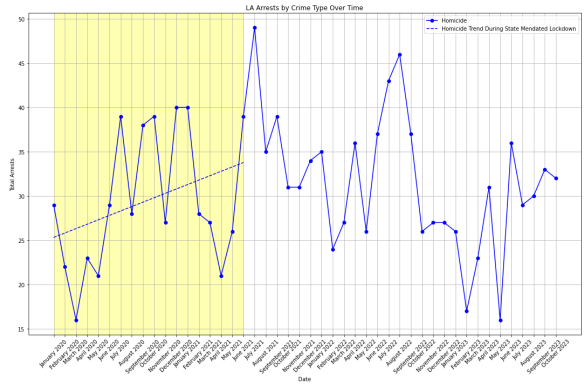

Figure 24: Homicide Trends During State-Mandated Lockdown.

during

According to the literature, there is supposed to be a rise in domestic violence arrests during the COVID-19 lockdown. However, from the graph, we can see there was a decrease in domestic violence arrests during that period. We suspect that this is due to the lack of the "domestic" crime specification, as we filtered out specific offenses' descriptions and residential regions to create hypothetical domestic violence crimes only from the dataset.

Now, we will show the word cloud for each city to showcase the most common crime types in each region.

Figure 25: Word Cloud for NYPD.

Figure 26: Word Cloud for Chicago.

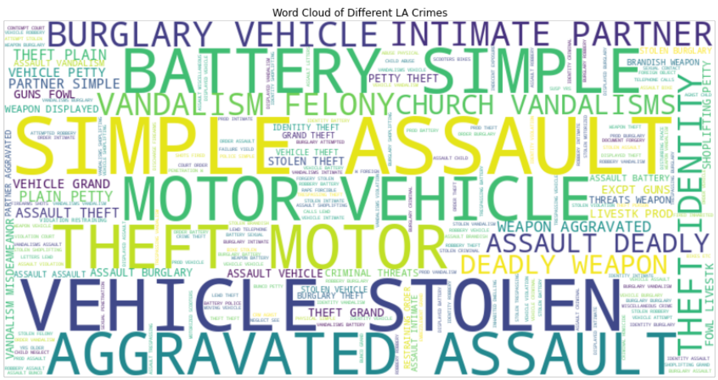

Figure 27: Word Cloud for LA.

As seen in the word clouds, there is a significant difference between the most common offenses in each dataset. Furthermore, we speculate that each state has different descriptions for the same offense. That can explain why we were unable to produce a graph for homicide cases in Chicago, as well as why we had to count several descriptions to produce the robbery offenses in said dataset.

5 Challenges

5.1 Class Imbalance in Financial Recession Prediction:

Models struggled with the 90-10 imbalance in recession vs. non-recession data, favoring non-recession periods. To address this, PR-AUC and class weighting techniques were employed to improve the models' sensitivity to recession periods.

5.2 Modeling Crime Data:

Reporting standards varied significantly among states, requiring extensive normalization and feature engineering. Models such as SARIMA and BSTS demanded careful adjustment to align with economic events and produce accurate results.

5.3 Ethical Concerns:

The possibility of bias in predictive policing raised ethical concerns, particularly the risk of perpetuating stereotypes in neighborhoods. To minimize stigmatization, the focus was placed on reporting results and prioritizing economic support over punitive measures.

5.4 Domestic Violence Data Limitations:

Since the datasets did not include explicit domestic violence criteria, particularly in the Los Angeles dataset, we had to hypothesize probable indicators for crimes to classify them as domestic violence. This limitation constrained our ability to investigate domestic violence trends comprehensively during the pandemic.

5.5 Inconsistent Crime Descriptions Across Datasets:

For the datasets used to investigate crime trends in New York, Chicago, and Los Angeles, there were significant differences in the descriptions of the same crimes. As a result, we were unable to cohesively analyze trends in robberies, burglaries, and homicides since the categorizations were not consistent across datasets.

6 Ethical Recommendations

This capstone project examines the relationship between economic recessions, crime rates, and market vs. crime dynamics. Each part of the project has ethical concerns that must be addressed to avoid unintended social consequences, misinterpretations, and prejudices. From forecasting recessions to analyzing crime trends during economic downturns, distinct ethical concerns arise that must be carefully addressed to avoid reinforcing stereotypes or perpetuating biases, particularly against economically vulnerable communities. This is especially relevant to African-American communities, which have historically been persecuted more harshly compared to other populations [3].

The prediction models utilized in this project, such as Logistic Regression, Random Forest, and Bayesian Structural Time Series (BSTS), each involve biases and assumptions that might affect outcomes, particularly in imbalanced datasets where one class dominates. For example, whether forecasting recessions or evaluating crime rates, an overemphasis on the dominant class might result in models that underperform on minority classes, distorting reality and potentially leading to policies that ignore substantial trends in crime or economic movements.

Regarding the predictive algorithms, even though grounded in quantitative methods, they cannot capture the full complexity of economic factors, especially with new trends (such as new technology usage) tied to human behavior and social impact. For example, the FBI has highlighted escalating AI-related threats, such as social engineering attacks and voice/video cloning scams, as emerging issues in 2024 [9]. Over-reliance on these models risks limiting economic diversity, leading to the convergence of business practices and sidelining innovation, diversity, and creativity within the economy.

Predictive modeling of crime during recessions may result in increased allocation toward punitive measures. Without careful consideration, such a preventive approach may divert funding away from community-based interventions, exposing society to the risk of intensifying policing. As the ACLU and other organizations note, predictive policing methods can reinforce bias and undermine public trust in law enforcement [3, 9].

Public policy decisions based on this approach risk oversimplifying complicated links between economic hardship and crime, potentially leading to targeted enforcement or economic policies that disproportionately affect low-income neighborhoods. Predictive crime research based on economic downturns, for example, may perpetuate negative assumptions about disadvantaged communities and worsen social disparities by boosting law enforcement activity in these areas while ignoring core socioeconomic causes. As research has demonstrated, socioeconomic assistance measures are frequently more successful in reducing crime than increased enforcement.

Our investigation did not reveal any direct correlation between recessions and rises in homicide, burglary, and robbery rates despite evidence from the literature. Furthermore, while prior research indicates a likely increase in domestic violence during pandemic lockdowns, our gathered data does not reflect this trend. Thus, we report significant differences between the literature and our findings.

6.1 Proposed Ethical Recommendations

We propose the following ethical recommendations [1]:

- **Standardized Crime Reporting:** There must be more standardized crime reporting and uniformity in crime descriptions across counties and states. Our findings were significantly diverse among the different datasets due to differences in descriptions and categorizations of the same crime types. By adopting more unified crime categorization, we should be able to report more accurate findings in our research and ensure that policy responses are well-informed and appropriate. Additionally, we recommend that investigative agencies work toward more comprehensive data collection methods, especially during recessions, so that it will be easier to identify crime trends during such periods [10, 15].

- **Sensitivity in Reporting Findings:** Researchers should report findings with sensitivity to avoid unintended stigma associated with specific communities or economic conditions. While certain crime types suggest an increase during recessions in the literature, findings that contradict this should be communicated carefully to avoid reinforcing harmful stereotypes [13].

- **Clear Description of Domestic Violence:** As the literature suggests a rise in domestic violence during lockdowns, there is an ethical responsibility to protect and ensure unbiased and respectful reporting of these vulnerable groups. Since in this specific research we had to improvise to identify potential domestic violence crimes, there should be a clear and consistent description of domestic violence assaults in future datasets. This ensures that hypothetical classifications are avoided, and vulnerable groups are accurately represented and protected.

References

[1] Research ethics in criminology. *Research Ethics Review*, 1:60–66, 2005. URL `https://journals.sagepub.com/doi/pdf/10.1177/174701610500100206`.

[2] Steve Adams. Crime trends during a recession, 2022. URL `https://www.skopenow.com/news/crime-trends-during-a-recession`.

[3] American Civil Liberties Union. Statement of concern about predictive policing by aclu and 16 civil rights privacy, racial justice, and technology organizations, 2024. URL `https://www.aclu.org/documents/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice`.

[4] M. P. J. Ashby. Initial evidence on the relationship between the coronavirus pandemic and crime in the united states. *Crime Science*, 9(6), 2020. `doi:10.1186/s40163-020-00117-6`.

[5] Benjibb. S&p 500 since 1950, 2021. URL `https://www.kaggle.com/datasets/benjibb/sp500-since-1950/data`. Accessed 6 Sept. 2024.

[6] Bureau of Labor Statistics. Labor market adjustments to recession, 2002. URL `https://www.bls.gov/opub/mlr/2002/02/art1full.pdf`.

[7] Bureau of Labor Statistics. The great recession and the great recovery, 2018. URL `https://www.bls.gov/opub/mlr/2018/article/great-recession-great-recovery.htm`.

[8] G. M. Campedelli, A. Aziani, and S. Favarin. Exploring the immediate effects of covid-19 containment policies on crime: An empirical analysis of the short-term aftermath in los angeles. *American Journal of Criminal Justice*, 45:704–727, 2020. `doi:10.1007/s12103-020-09578-6`.

[9] Federal Bureau of Investigation. Fbi warns of increasing threat of cyber criminals utilizing artificial intelligence, 2024. URL `https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence`.

[10] Federal Bureau of Investigation. U.s. big city crime data 2000-may 2024, 2024. URL `https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/explorer/crime/hate-crime`. Accessed 6 Sept. 2024.

[11] Federal Deposit Insurance Corporation. The financial crisis: An fdic study, 2008. URL `https://www.fdic.gov/bank/historical/crisis/chap1.pdf`.

[12] Federal Reserve Bank of St. Louis. Gross domestic product, 2023. URL `https://fred.stlouisfed.org/series/GDP/`. Accessed 6 Sept. 2024.

[13] Office of the High Commissioner for Human Rights. Racism and ai: "bias from the past leads to bias in the future", 2024. URL `https://www.ohchr.org/en/stories/2024/07/racism-and-ai-bias-past-leads-bias-future`.

[14] P. Paul and J. H. Pedtke. Historical patterns around financial crises, 2020. URL `https://www.frbsf.org/research-and-insights/publications/economic-letter/2020/05/historical-patterns-around-financial`.

[15] Aron Pollner. Fbi nibrs crime data: 2021 all states, 2021. URL `https://www.kaggle.com/datasets/aronpollner/us-fbi-nibrs-crime-data-2021-all-states`. Accessed 6 Sept. 2024.

[16] J. Quednau. How are violent crime rates in u.s. cities affected by poverty? *The Park Place Economist*, 28, 2021. URL `https://digitalcommons.iwu.edu/parkplace/vol28/iss1/8/`.

[17] Skopenow. Crime trends during a recession, 2024. URL `https://www.skopenow.com/news/crime-trends-during-a-recession`.

[18] World Economic Forum. How does the covid recession compare?, 2020. URL `https://www.weforum.org/agenda/2020/08/how-does-the-covid-recession-compare/`.

[19] Yahoo Finance. S&p 500 historical data, 2023. URL `https://finance.yahoo.com/quote/%5EGSPC/history/`. Accessed 6 Sept. 2024.