



BY ORDER OF



IN PARTNERSHIP



CHICAGO POLICE
DEPARTMENT

PRESENTED BY



PRECRIME
TASK FORCE

CRIME PREDICTIONS & FORECASTING

PRECRIME TASK FORCE

INTRODUCTION

Chicago Police Department recently participated in the annual State Defense and Security Summit 2021 and have identified the need for a change in their policies on a city level. To increase the prevention effort of high crimes in the city and equipping proper skillsets and resources for efficiency moving forward.

As the first phase of revolutionizing policing work, they decided to set up a new department called PreCrime Task Force (PCTF) as representative of the force, informing and geared frontline officers with intel towards potential crimes trends.

'Prevent, deter and detect crime'





“Our endgame is arrests for the precursors to crimes”

“So, every day we’re going to be clearing the corners .. to protect people of Chicago”

David Brown
Superintendent of the Chicago Police Department

BRIEF BACKGROUND

Chicago City crime rate far exceeds the average of its country, United States and state, Illinois on every aspect from violent crime such as murder, rape to simple crime such as theft and assault.

Notably, in the past decade, Chicago's crime rates have seen a decreased in crime trend but the reasons why are not known to the government nor the state police. The total number of crime rates are still high.

While Chicago is known for series of violent and property crimes mentioned in the media, the percentage difference was measure based on individual crime. Other types of crime not mention may be greater in numbers and magnitude.

TASK

Assembly of Forensic Data Scientist to serve as the brains of PreCrime Task Force to predict and forecast crimes in Chicago.

SCOPE OF WORK

- To understand trends and patterns of crimes in Chicago
- Develop a model to churn out accurate predictions and forecasting for deployment to assist the frontline officers in their line of duty
- First phase of the plan: Targeting high crime rates



DATA SCIENCE PROCESS



Data
Acquisition



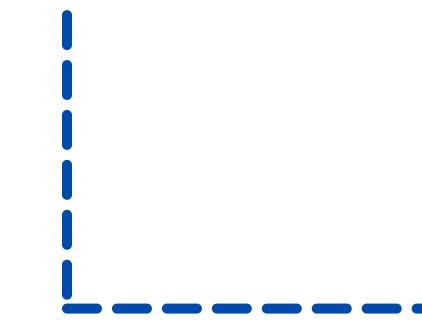
Exploratory Data
Analysis



Modeling &
Evaluation



Data
Cleaning



Data
Pre-processing



DATA ACQUISITION & CLEANING

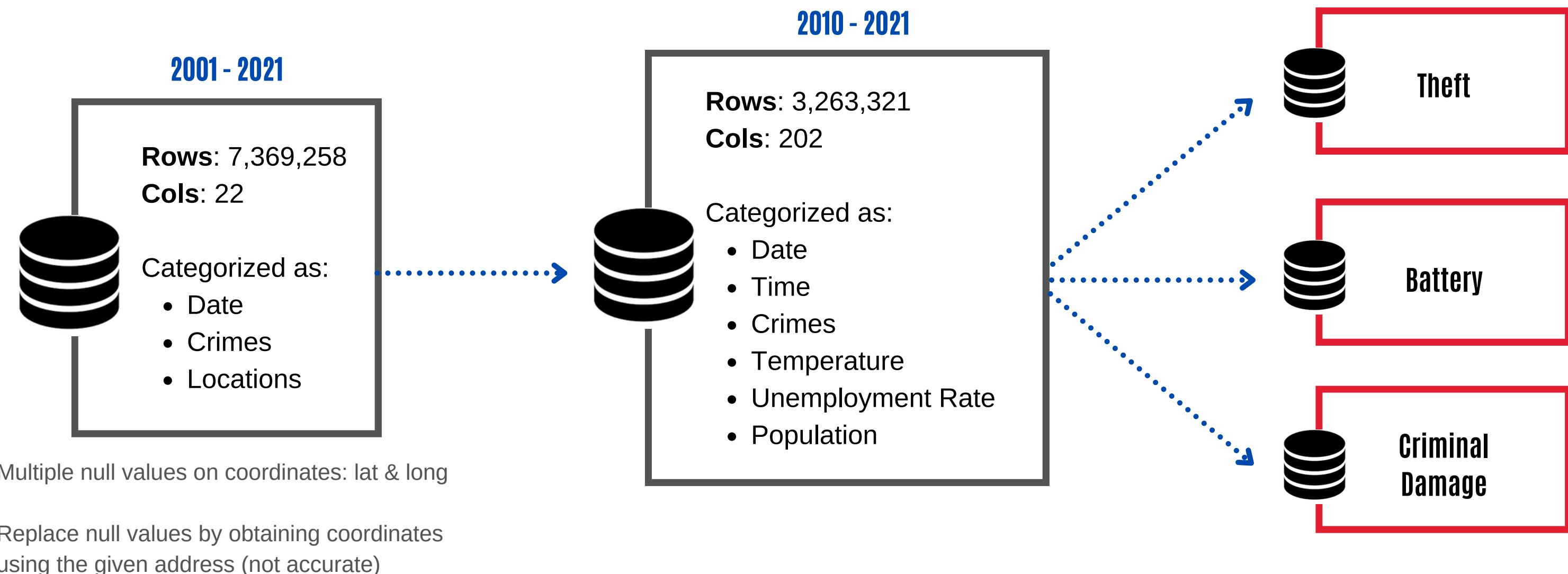
PRECRIME TASK FORCE

CRIME DATASET



CHICAGO
DATA PORTAL

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to 2021 and extracted from the Chicago Police Department (CLEAR)

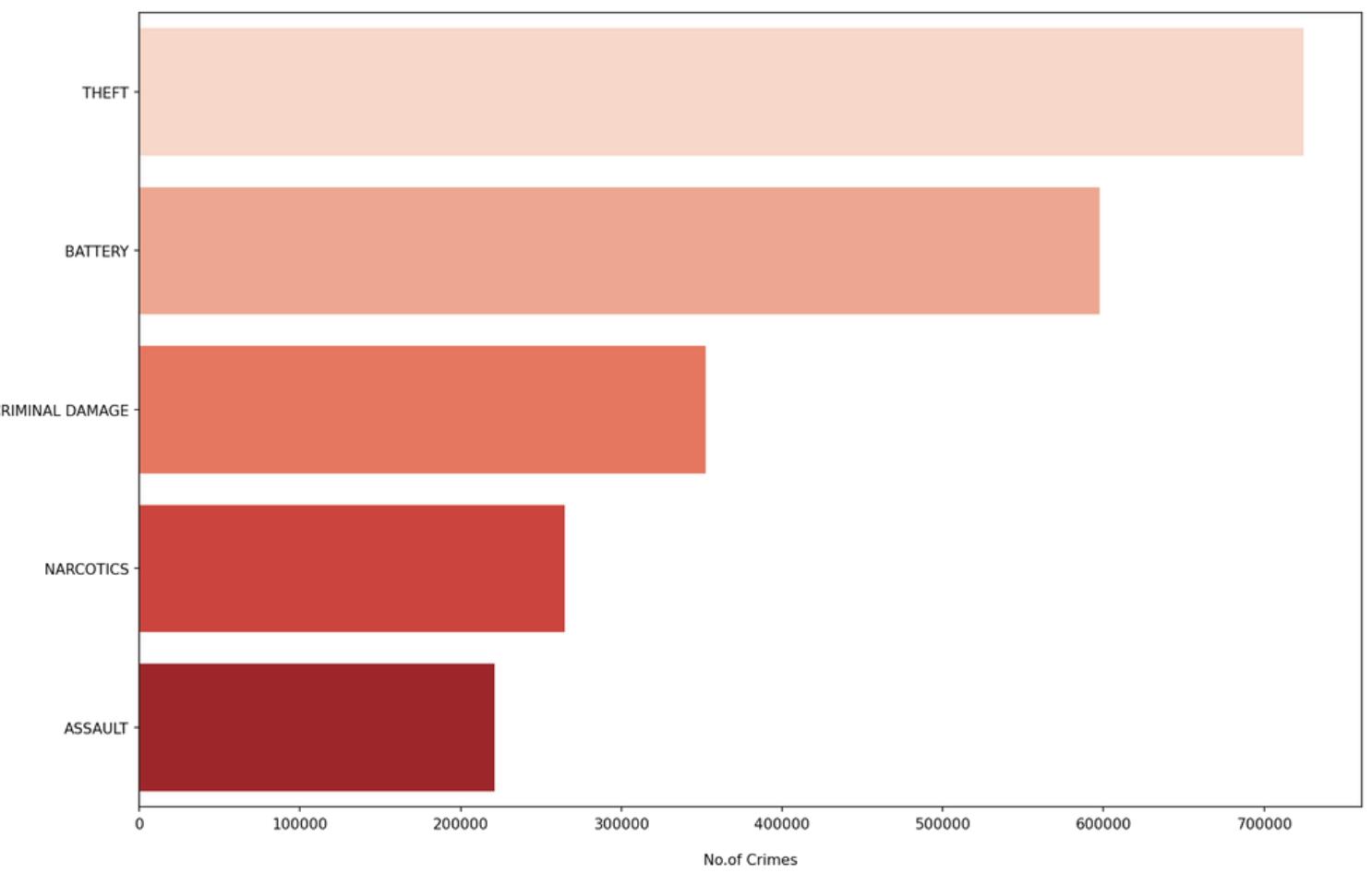


A woman with dark hair and a headband is looking upwards and to the left. She is positioned in front of a complex, semi-transparent digital interface. This interface features a large globe of the Earth, various data charts, graphs, and circular dashboards. One dashboard in the foreground is labeled "ROBOT". The overall theme is a blend of technology and data analysis.

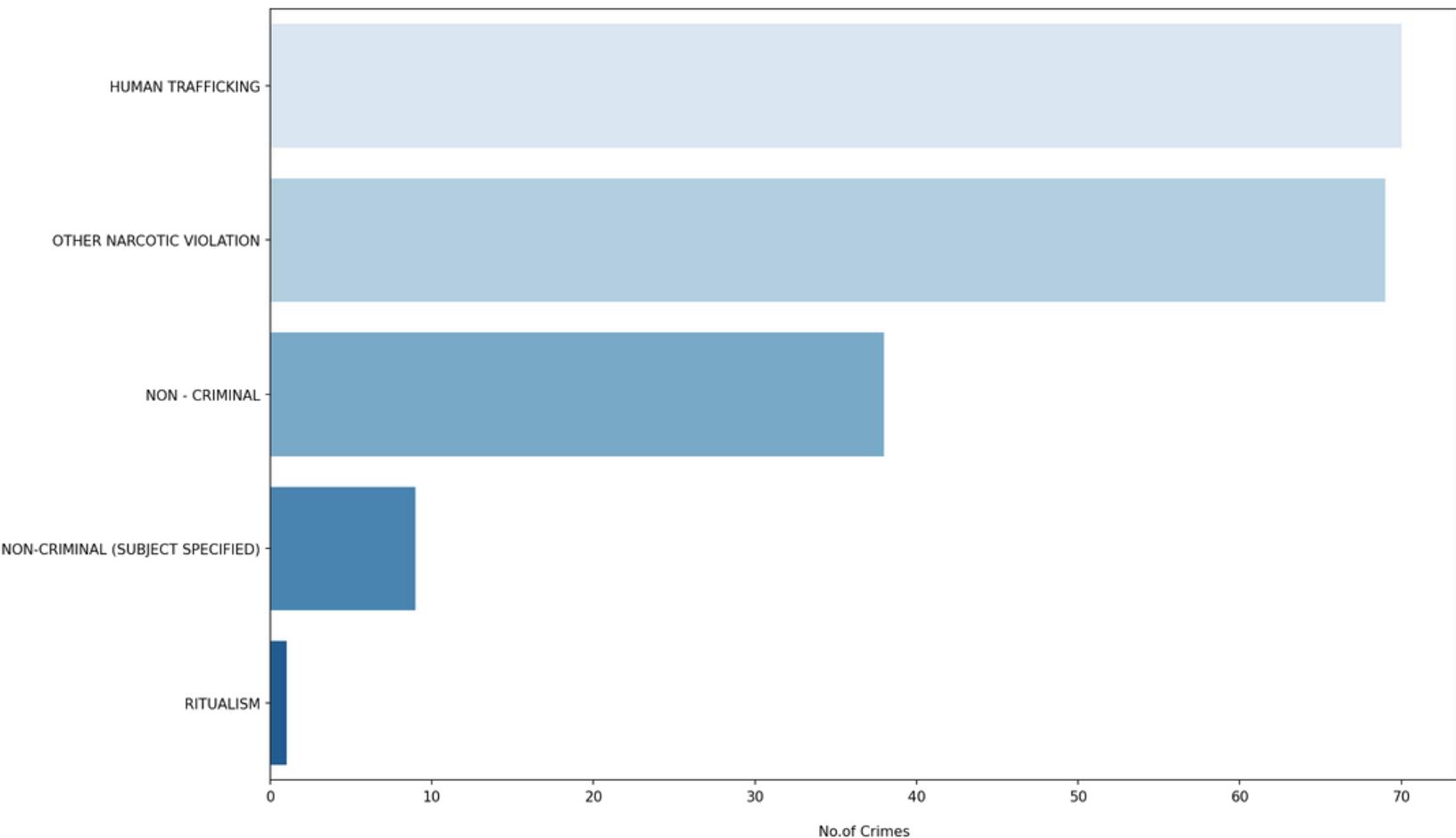
EXPLORATORY DATA ANALYSIS

PRECRIME TASK FORCE

Top 5 - Crimes Committed



Bottom 5 - Crimes Committed



Top 5

- Theft - 724,602
- Battery - 597,640
- Criminal Damage - 352,512
- Narcotics - 264,973
- Assault - 221,014

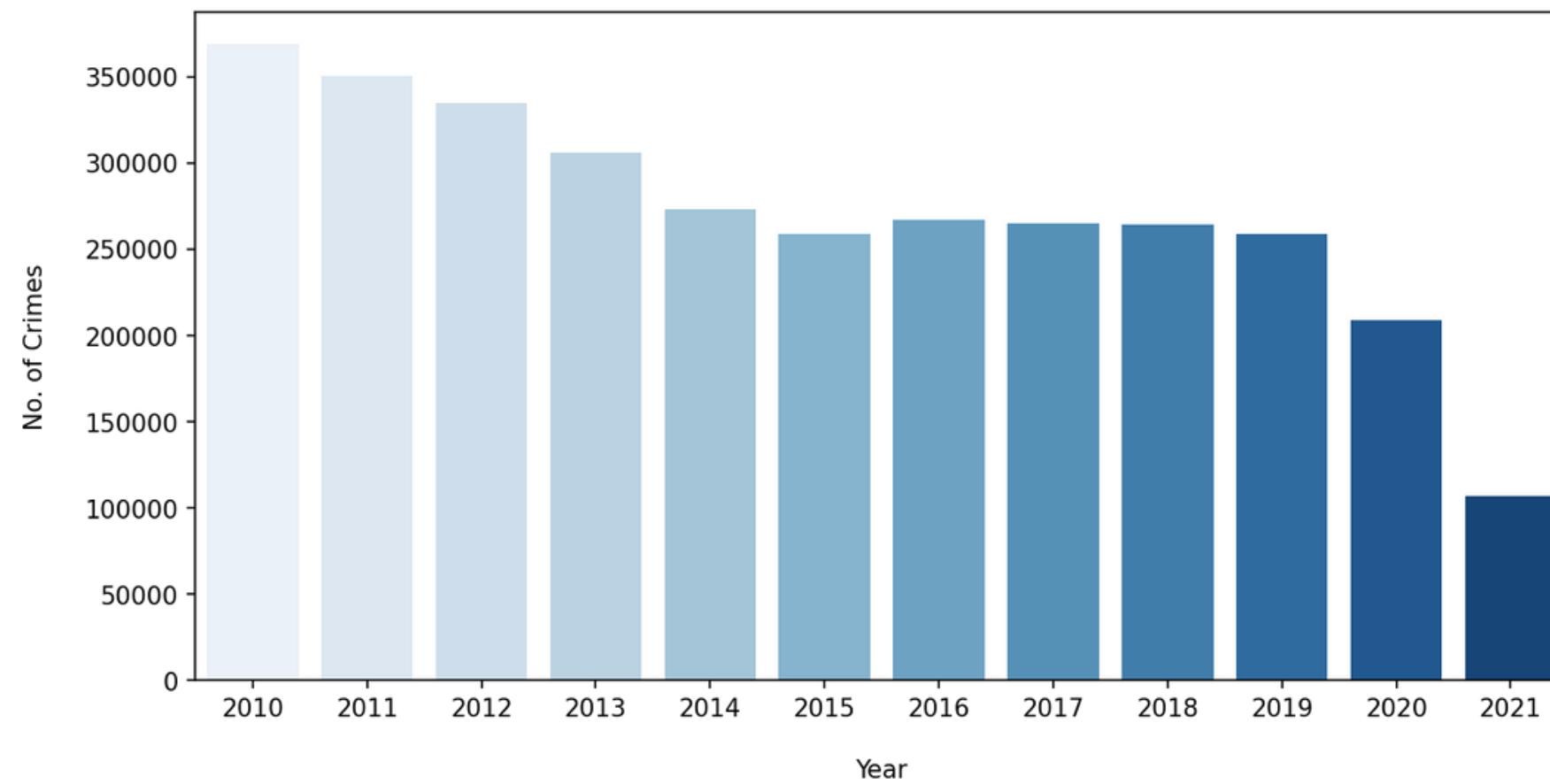
Bottom 5

- Ritualism - 1
- Non-Criminal - 9
- Non-Criminal (Subj. Specified) - 38
- Other Narcotics Violation - 69
- Human Trafficking - 70

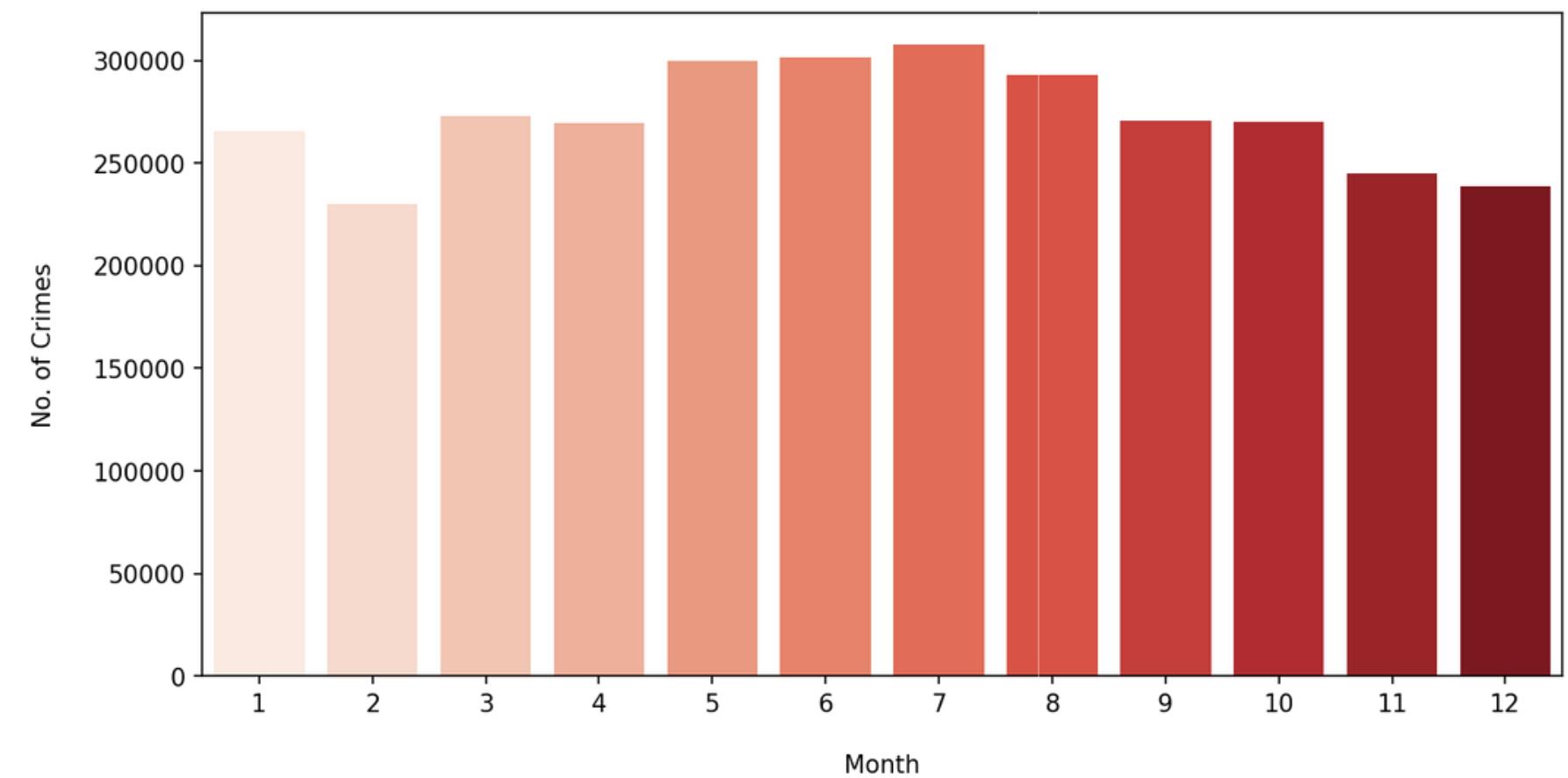
TOP & BOTTOM 5 CRIMES



Overall Crime Rates (Year)



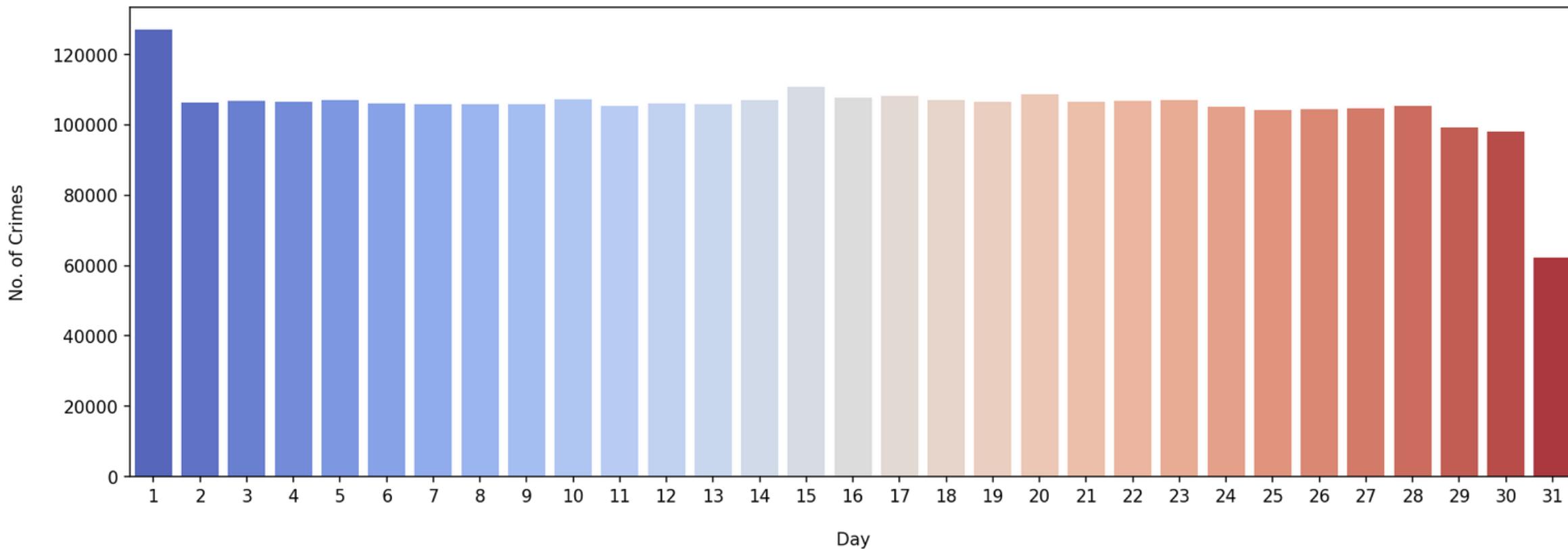
Overall Crime Rates (Month)



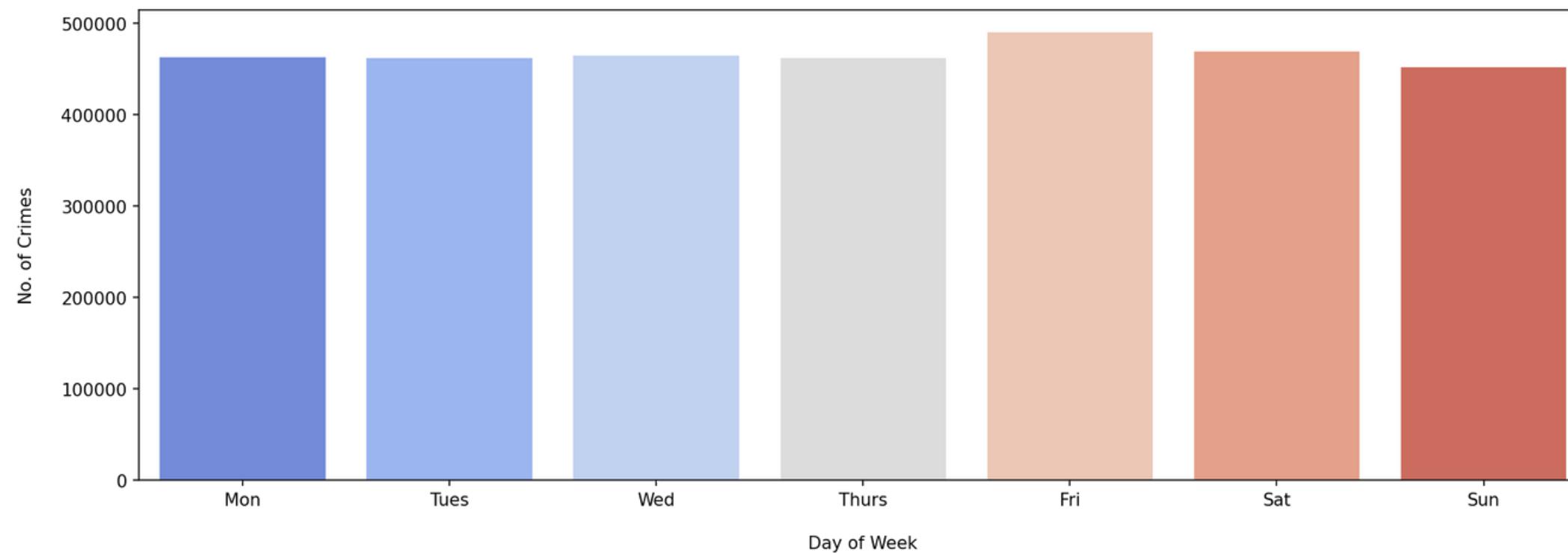
OVERALL CRIMES RATES (YEAR & MONTH)



Overall Crime Rates (Day)

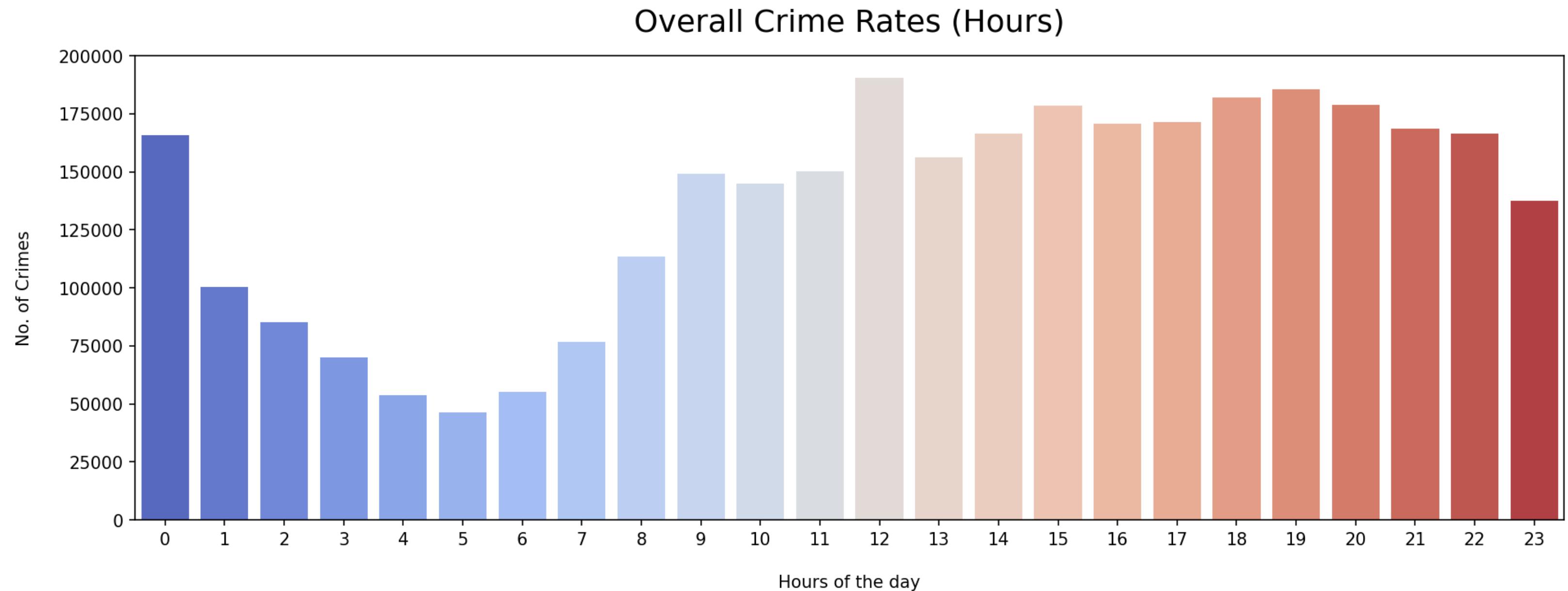


Crime Rates (Day of Week)



OVERALL CRIMES RATES (DAY & DAY OF THE WEEK)



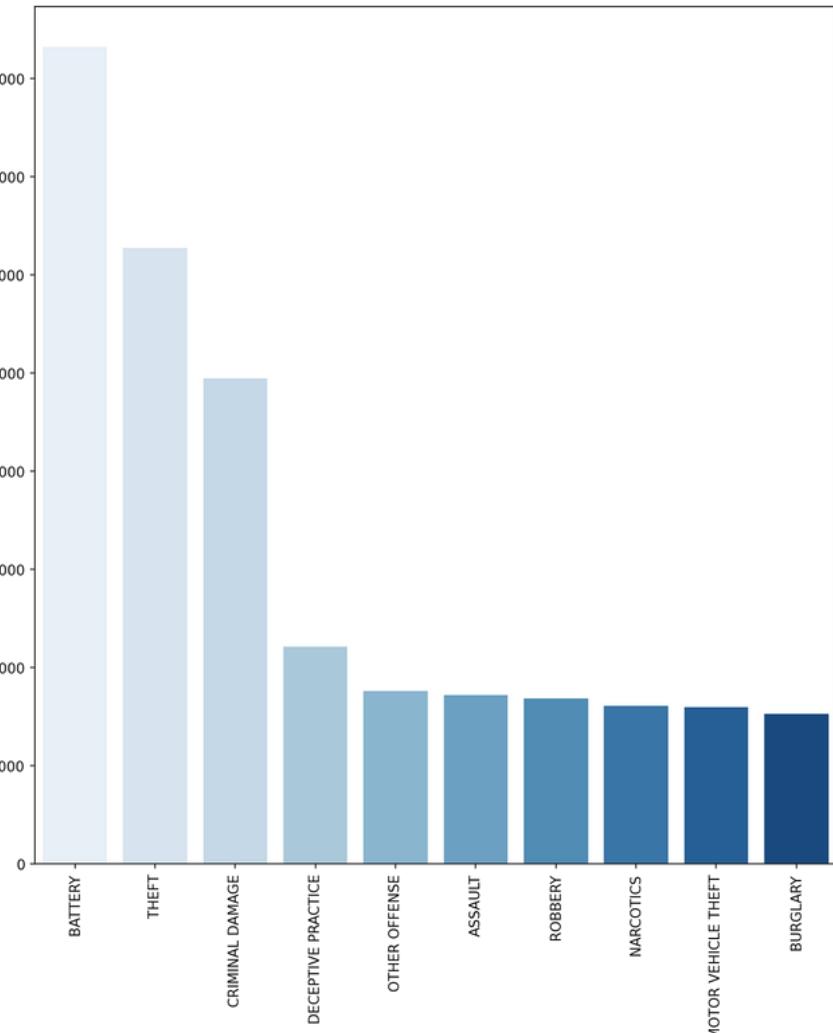


OVERALL CRIMES RATES (HOURS)

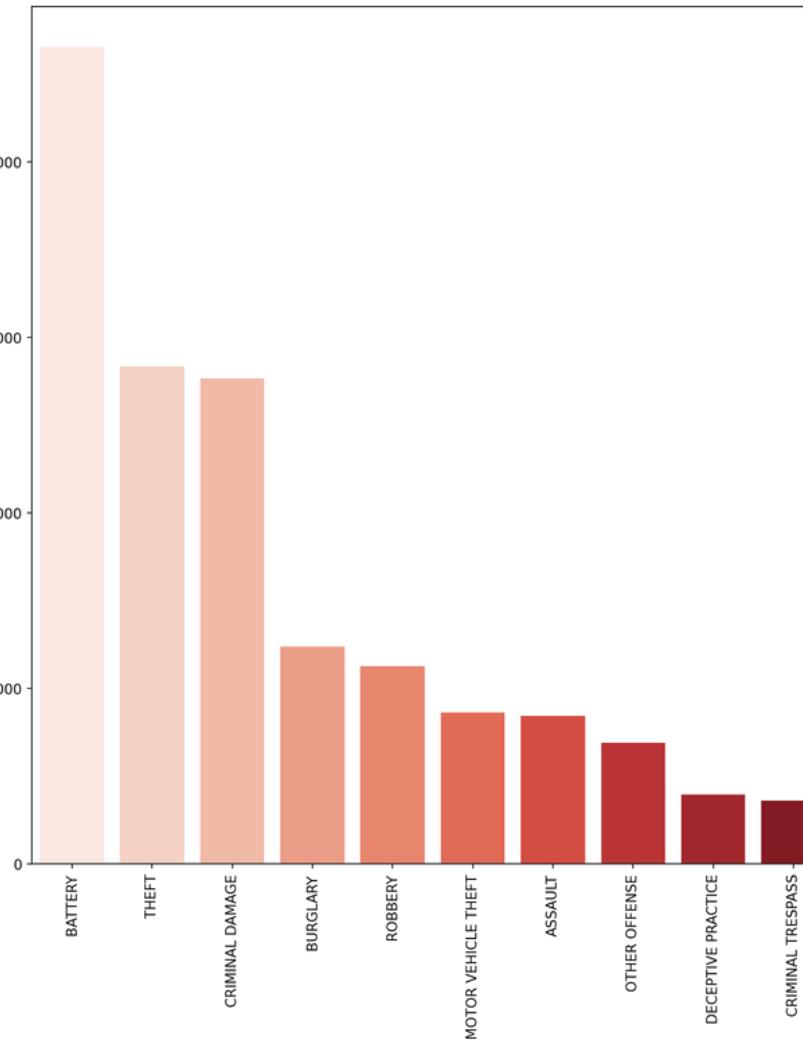


LATE NIGHT TO MORNING

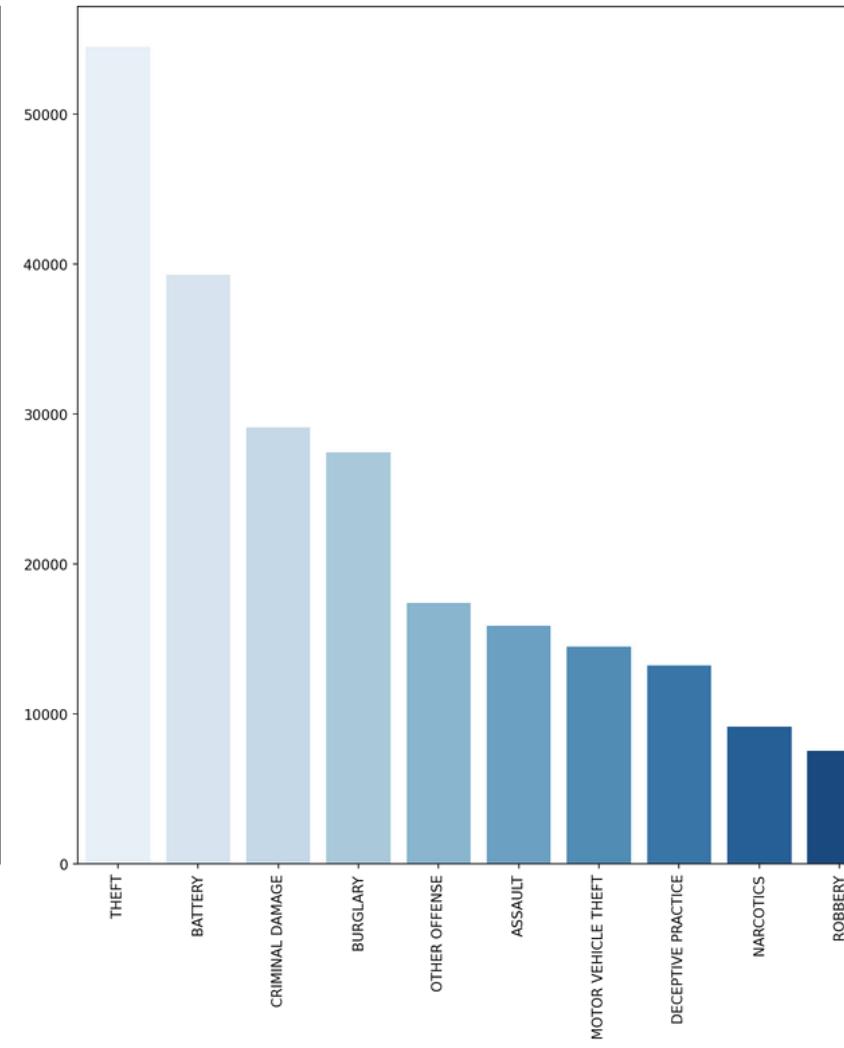
Late Night (12am to 3am)



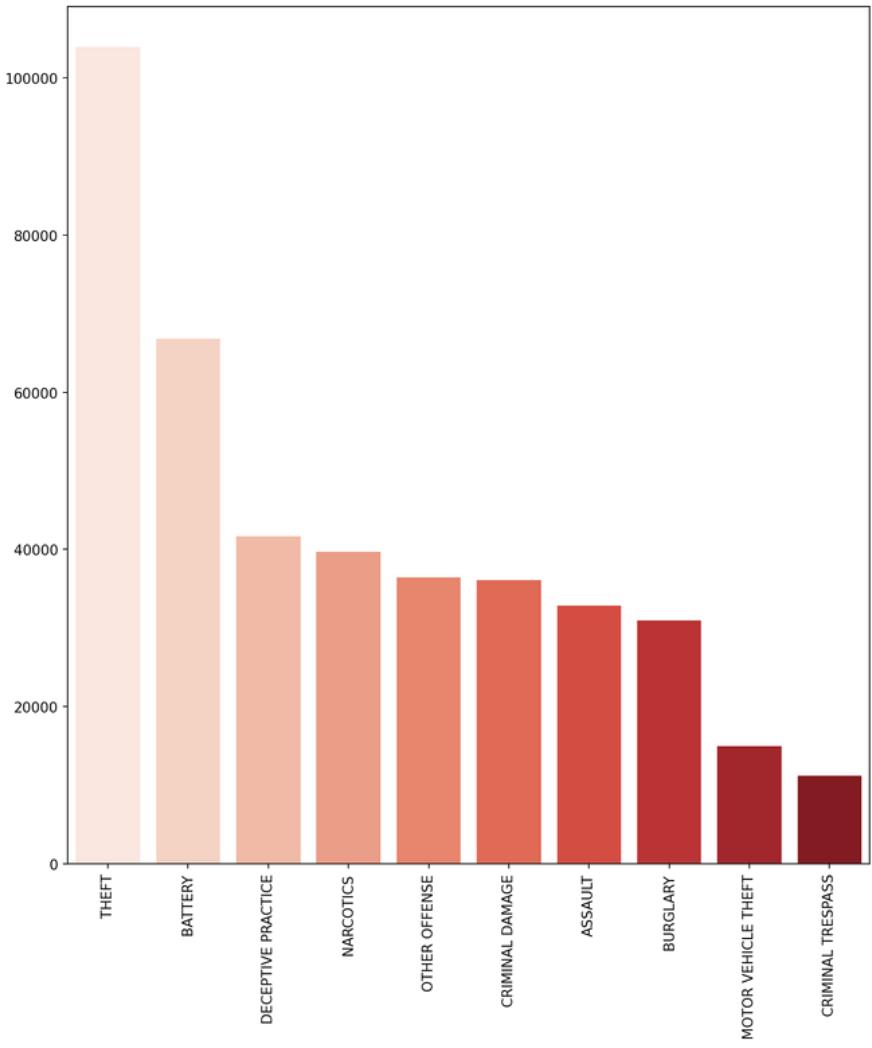
Towards Morning (3am to 6am)



Early Morning (6am to 9am)



Late Morning (9am to 12pm)

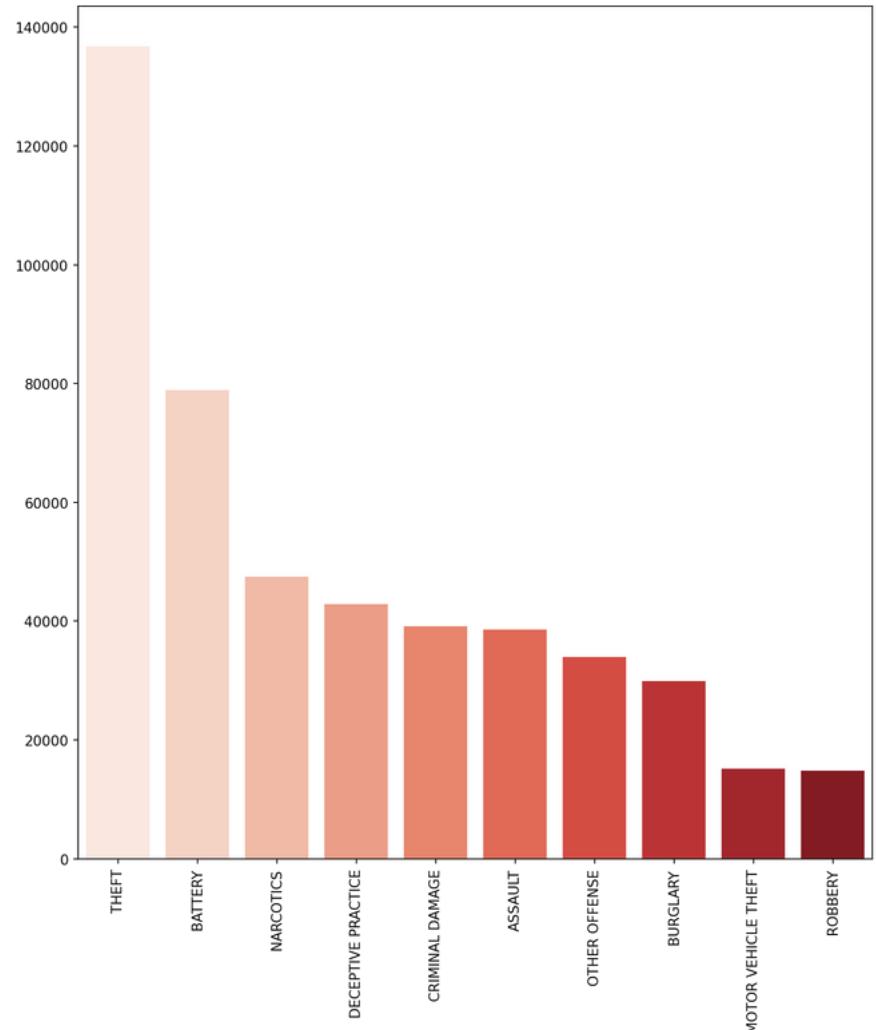


CRIMES VS HOURS

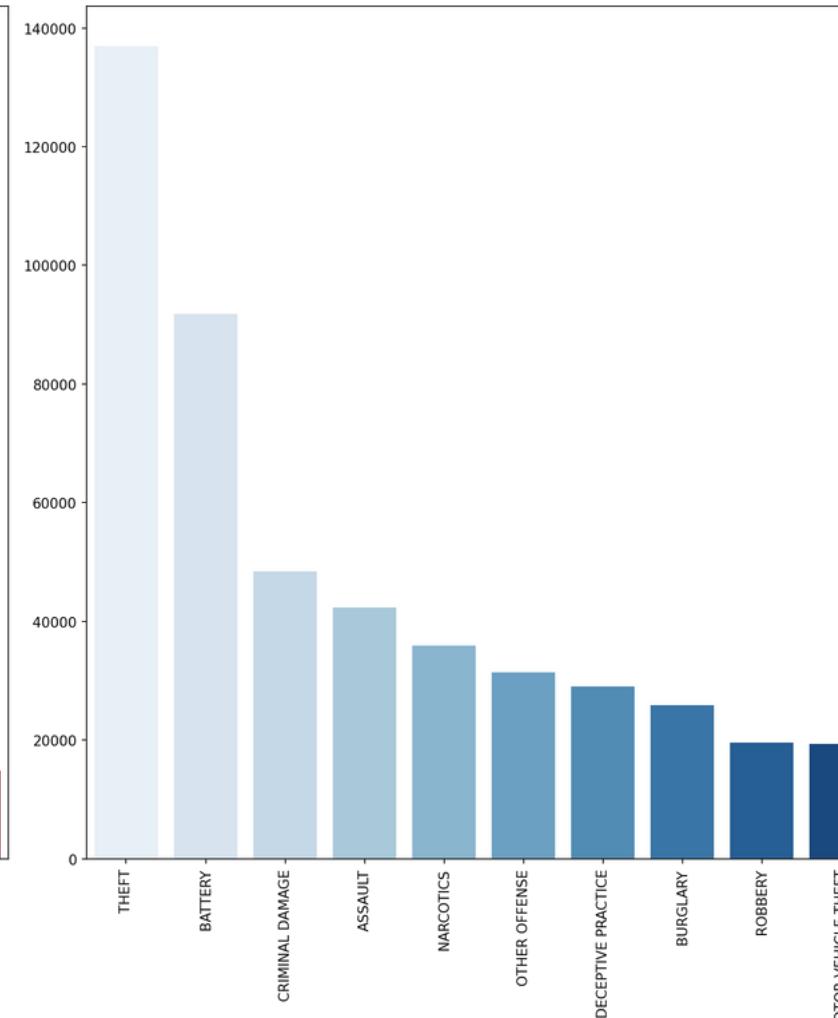


AFTERNOON TO EVENING

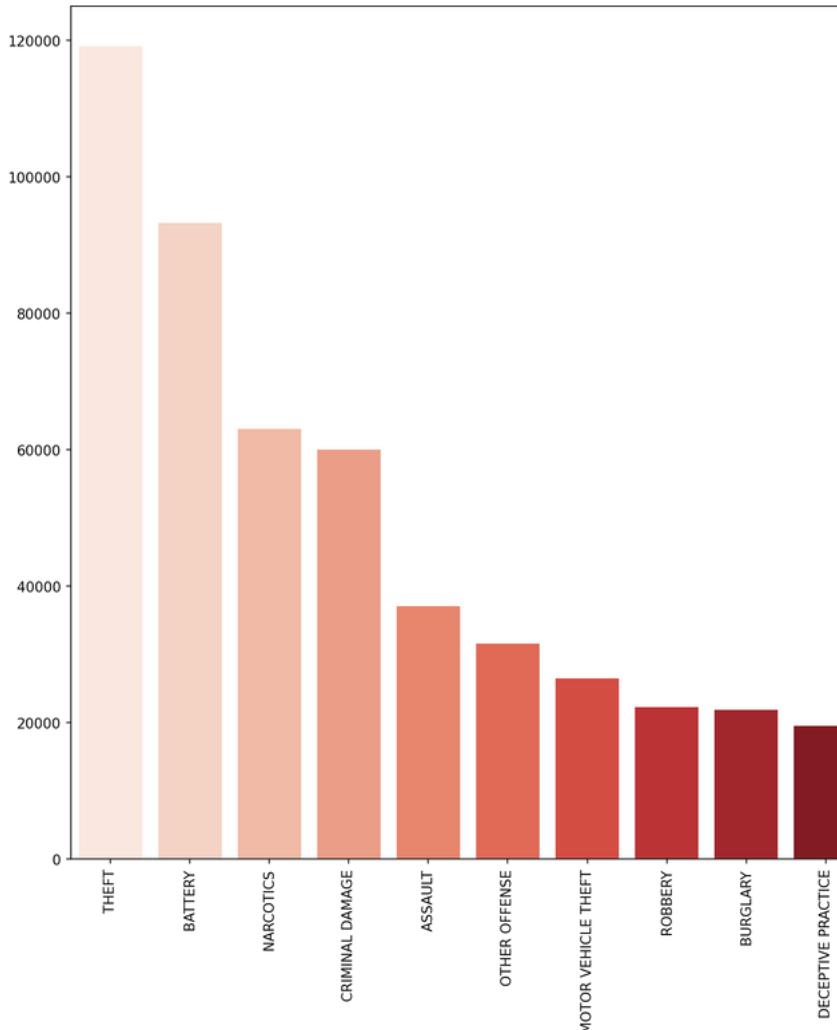
Early Afternoon (12pm to 3pm)



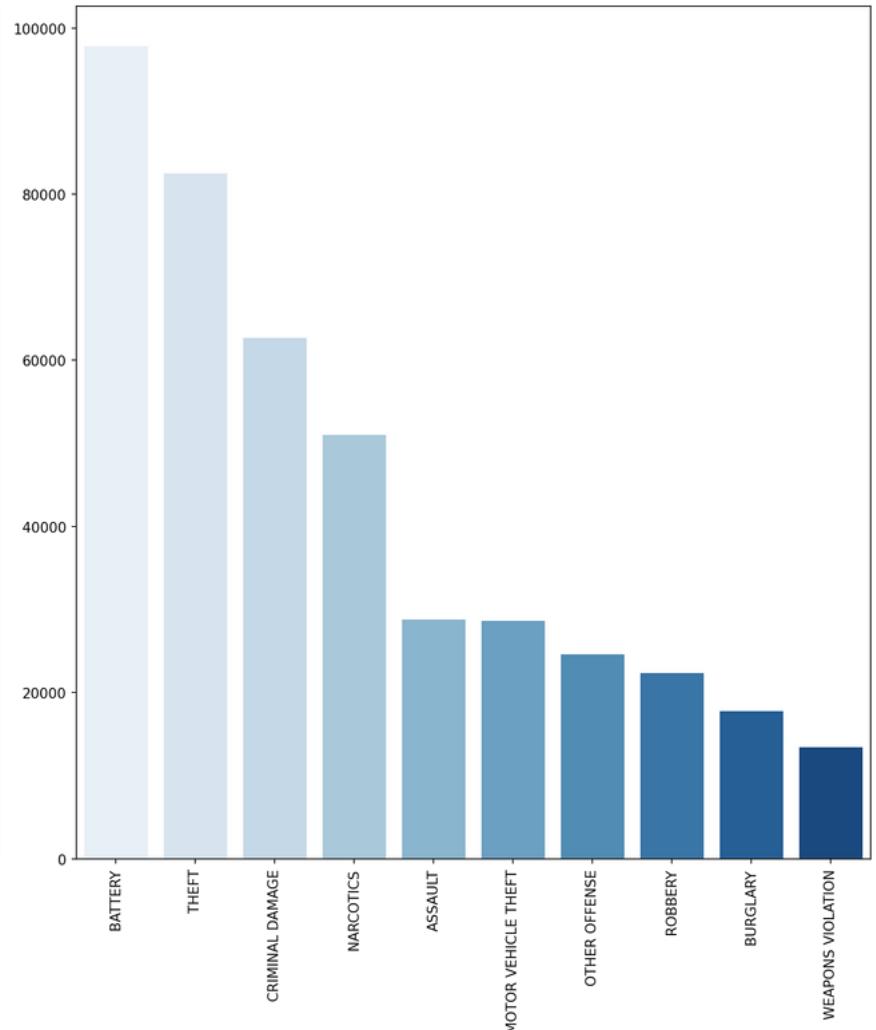
Late Afternoon (3pm to 6pm)



Evening (6pm to 9pm)



Late Evening (9pm to 12am)



CRIMES VS HOURS

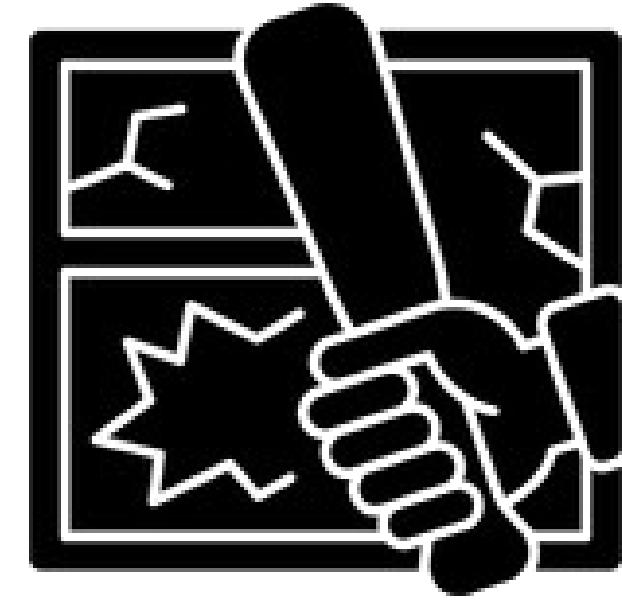




THEFT



BATTERY

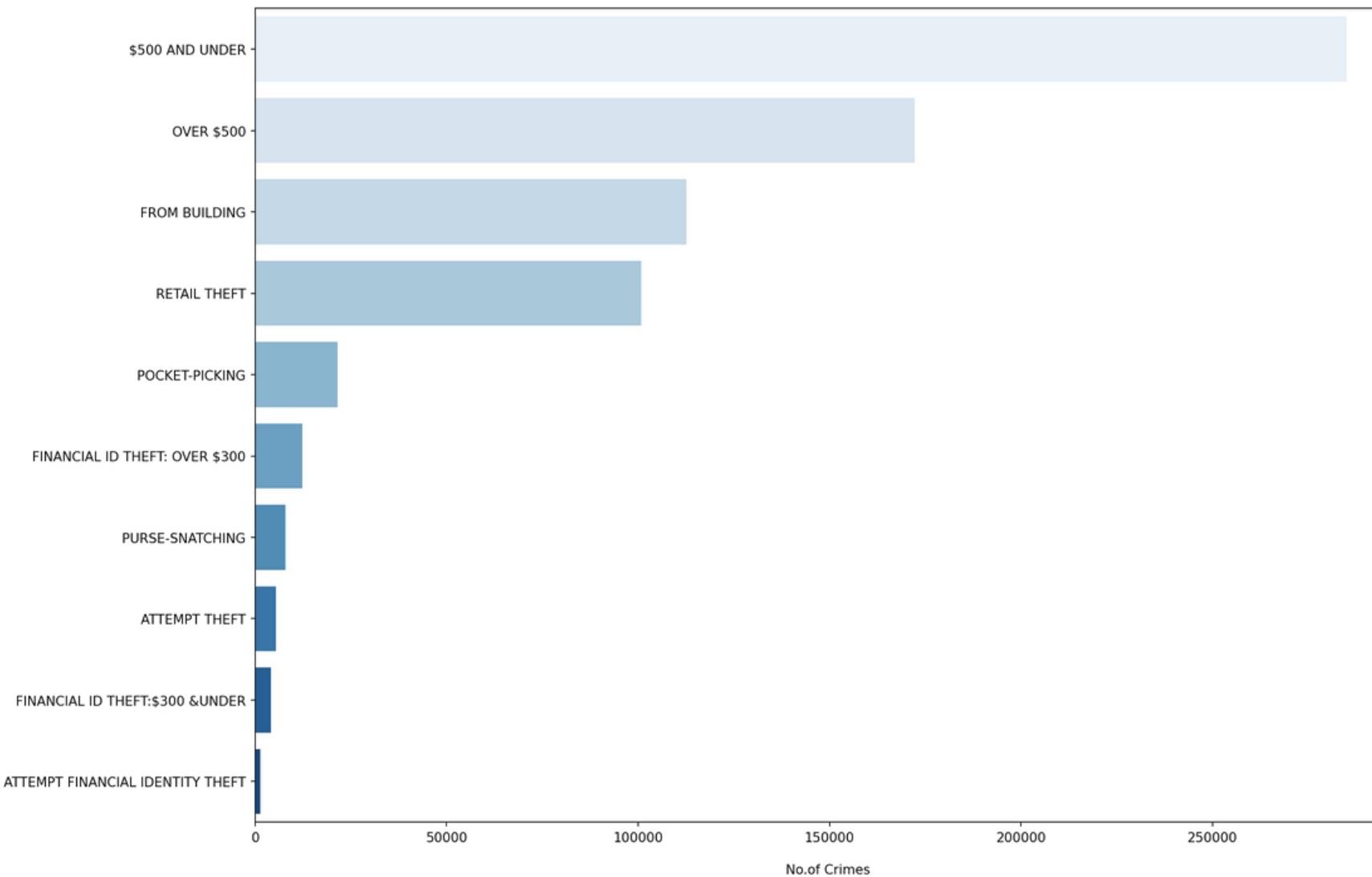


CRIMINAL DAMAGE

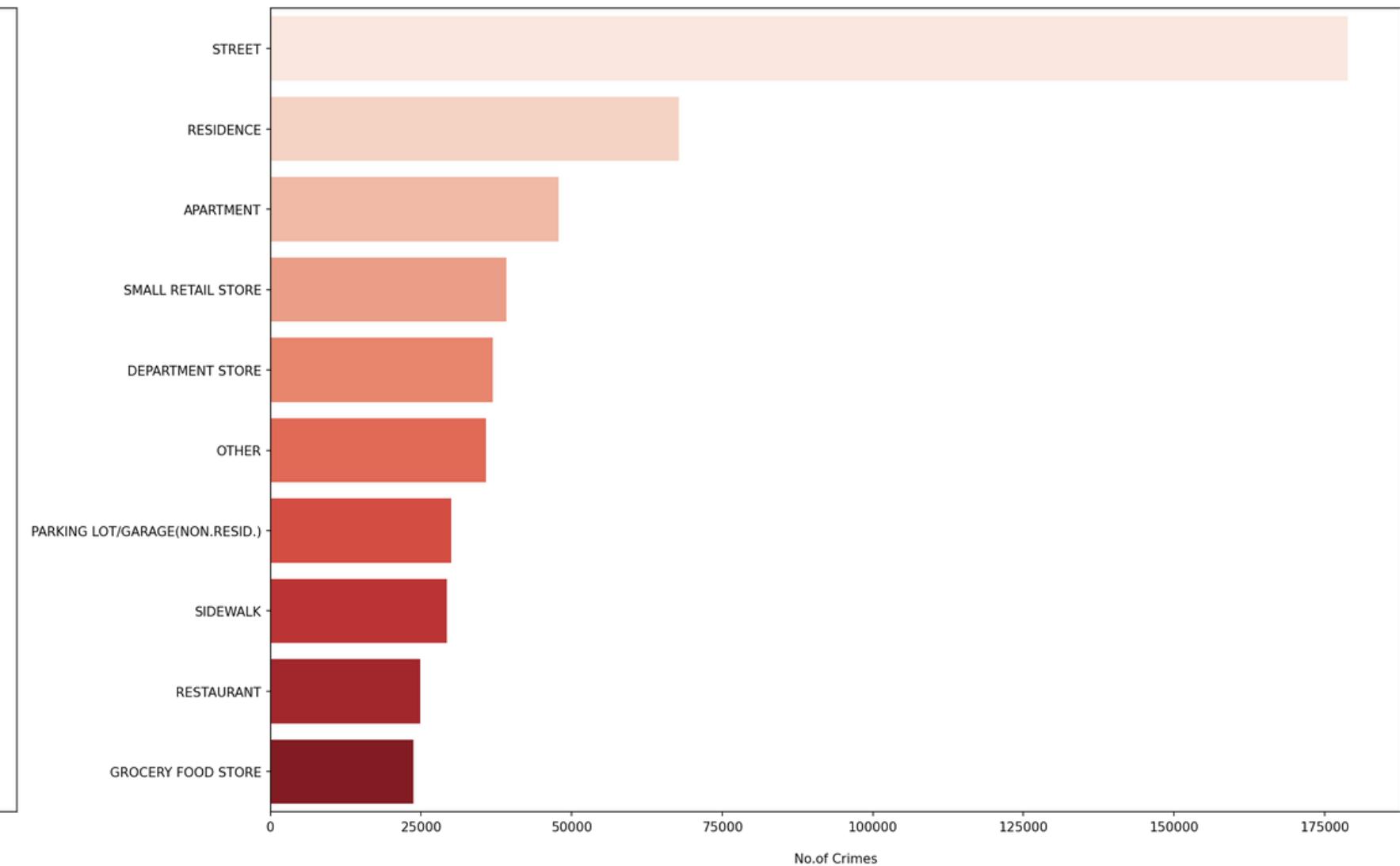
FIRST PHASE: IN-DEPTH OF CRIME



Causes of Theft



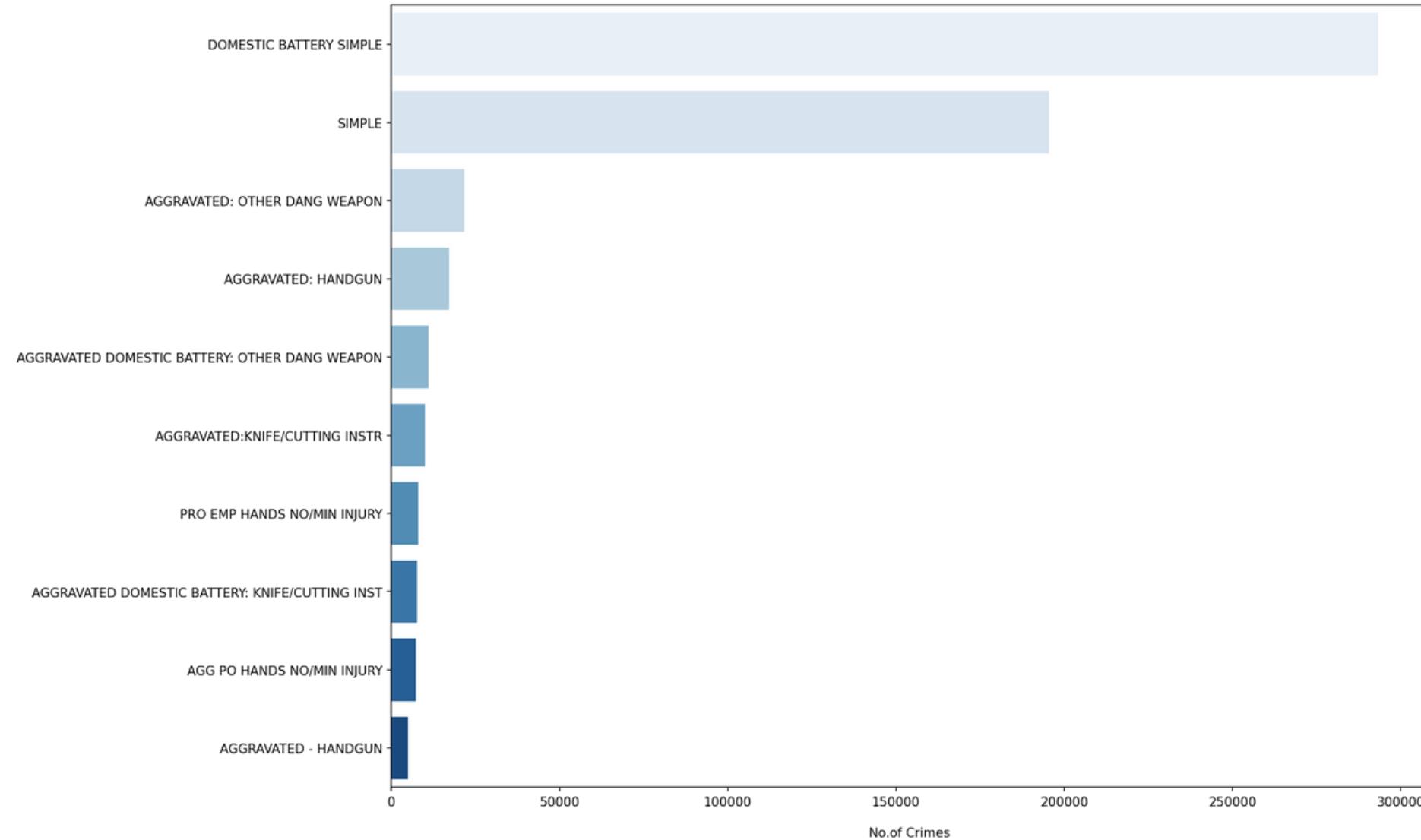
Locations of Theft



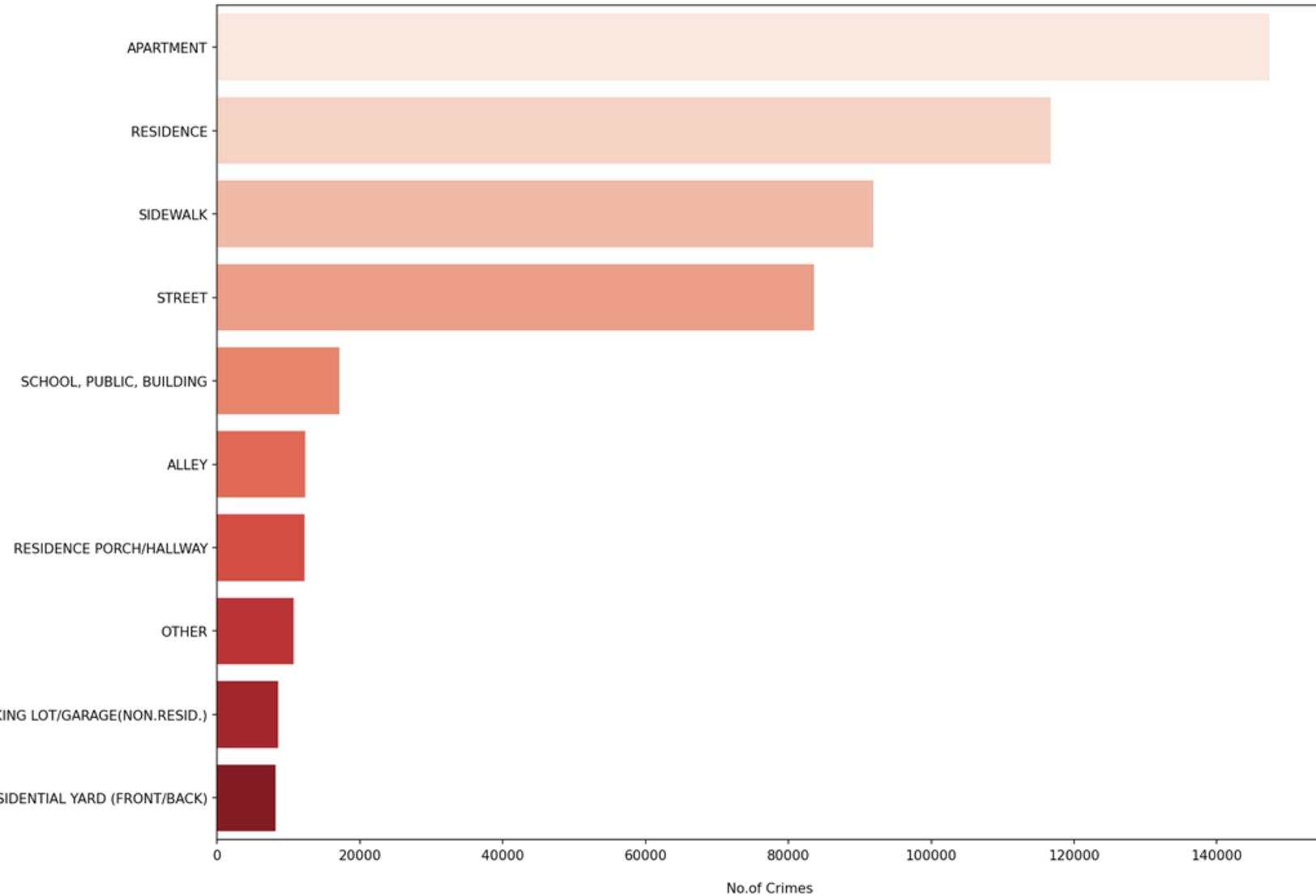
IN-DEPTH OF CRIME: THEFT



Causes of Battery



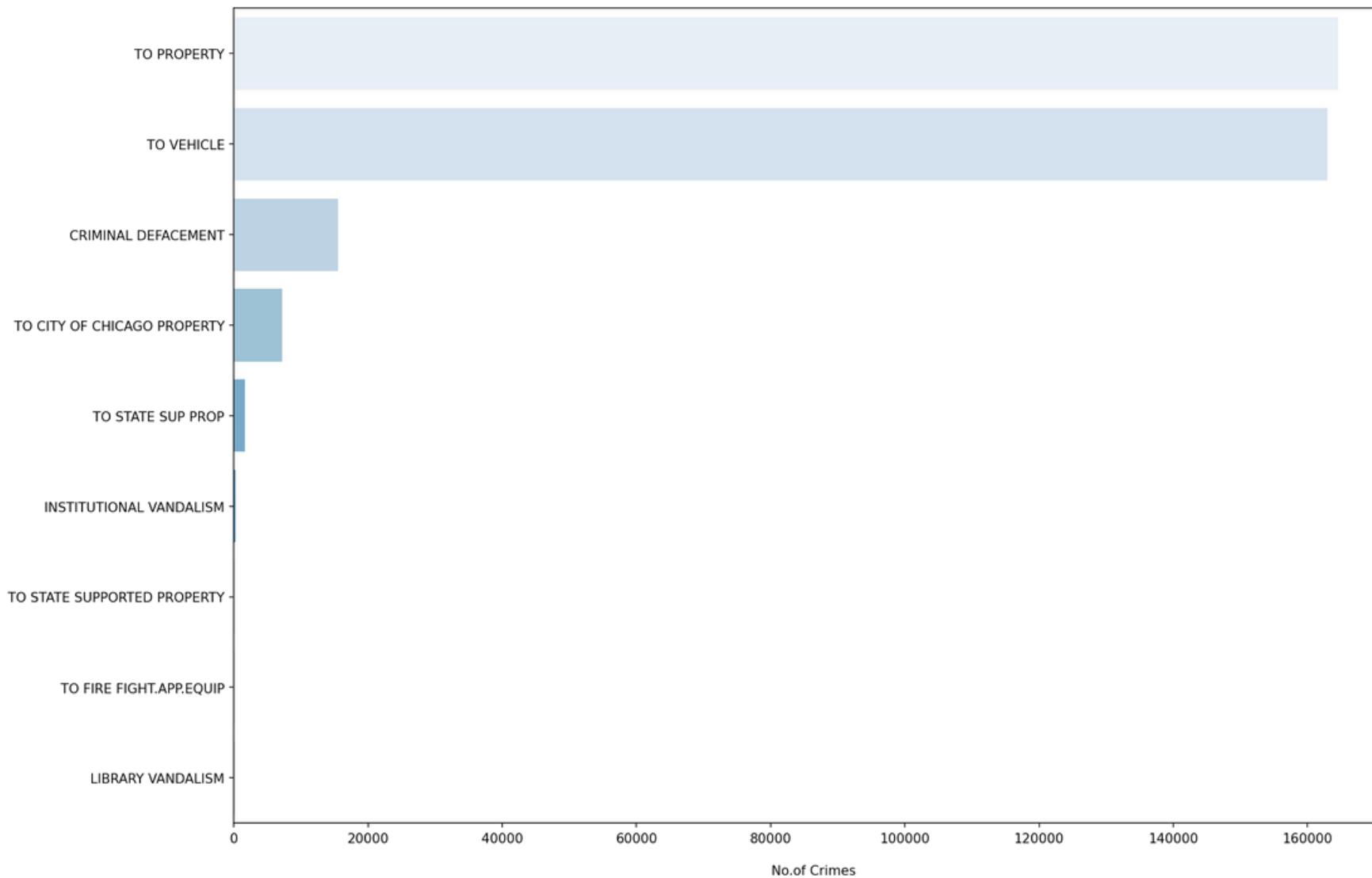
Locations of Battery



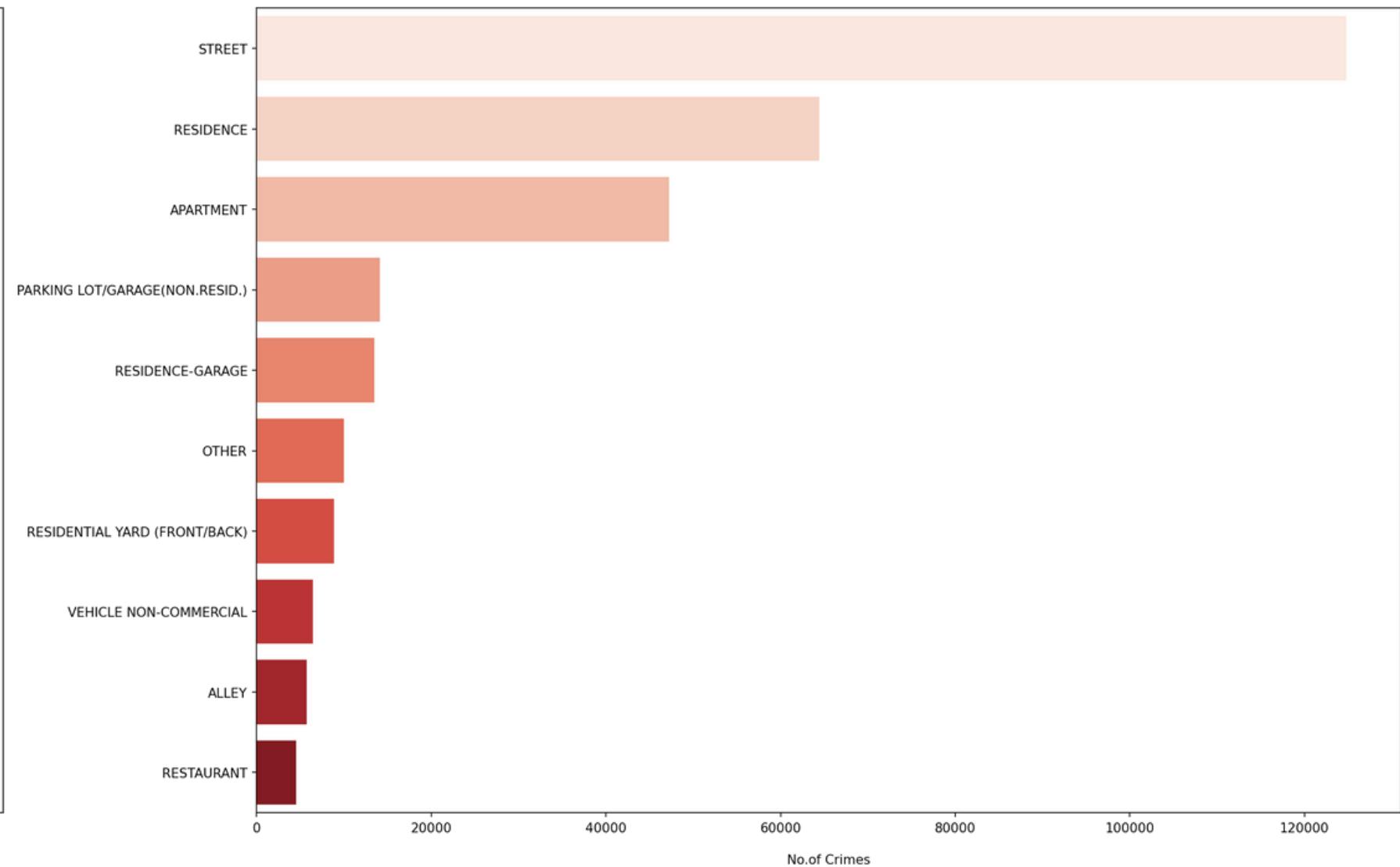
IN-DEPTH OF CRIME: BATTERY



Causes of Criminal Damage



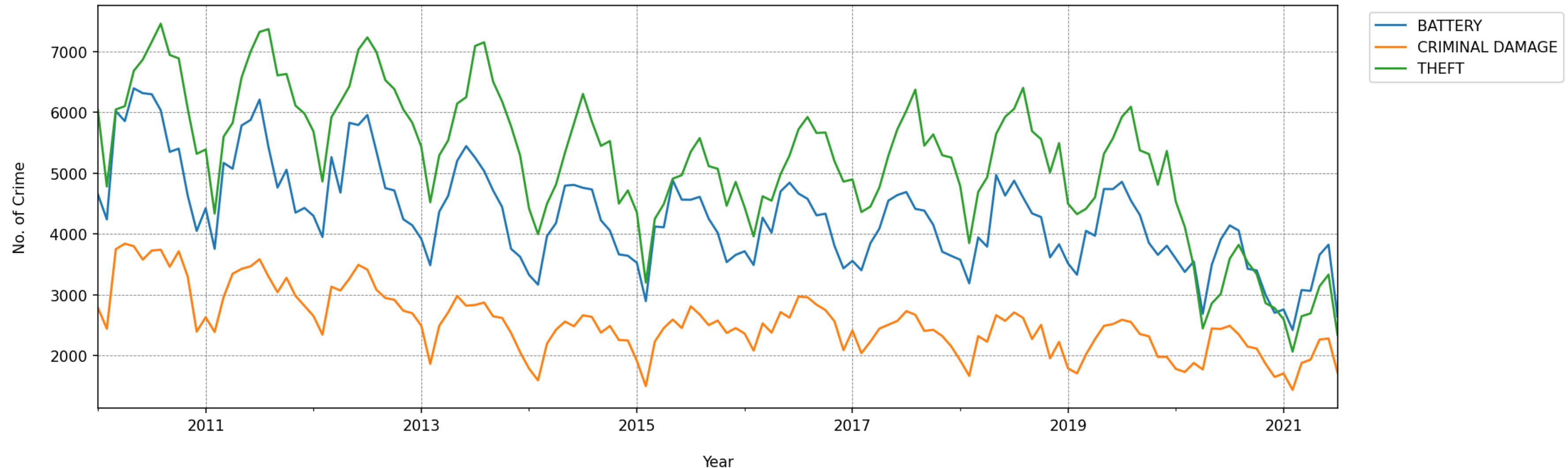
Locations of Criminal Damage



IN-DEPTH OF CRIME: CRIMINAL DAMAGE



Different Crime Types - Crime Rates



CRIMES RATES TRENDS





DATA PRE-PROCESSING

PRECRIME TASK FORCE

One hot encoding of categorical features

- **Description**
 - Theft - \$500 & above, Retail, Pocket-Picking ..
 - Battery - Simple, Aggravated: Dangerous Weapon ..
 - Criminal Damage - To Property, To Vehicle ..
- **Location Description**
 - Apartment, Residence, Street ..
- **Arrest**
 - Arrestable, Non-Arrestable
- **Domestic**
 - Domestic, Non-Domestic

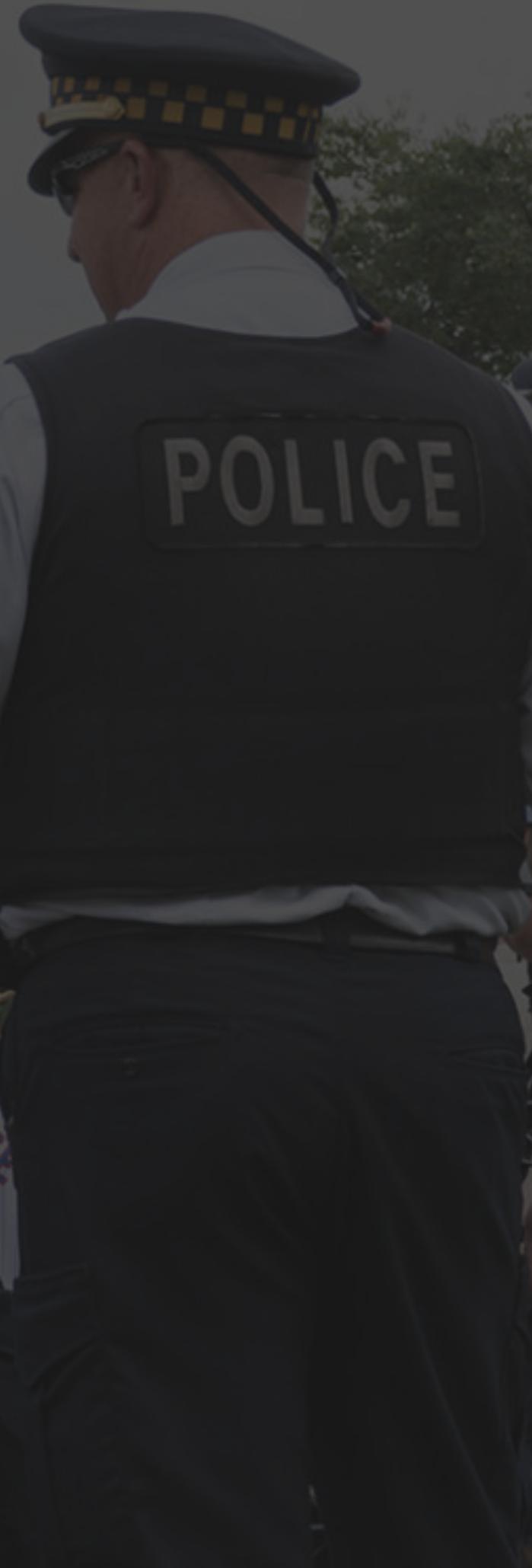
Extract features from date & time

- **Date**
 - Year, Month, Day, Day of Week
- **Seasons**
 - Spring, Summer, Autumn, Winter
- **Time of the Day**
 - Night, Morning, Afternoon, Evening
- **Weekend**
 - Weekday, Weekend



VARIABLES THAT AFFECT CRIME RATES

- Population density and degree of urbanization.
- ***Variations in the composition of the population, particularly youth concentration***
- Stability of the population with respect to residents' mobility, commuting patterns, and transient factors
- ***Economic conditions, including median income, poverty level, and job availability***
- ***Different types of holidays***
- Cultural factors and educational, recreational, and religious characteristics.
- Family conditions with respect to divorce and family cohesiveness.
- ***Weather climate and temperature***
- Citizens' attitudes toward crimes
- Crime reporting practices of the citizenry



Acquisition of data for new features

- United States Holidays
- Temperature
 - The average temperature (daily)
- Unemployment Rate
- Population Age
 - 5 to 7
 - 18 to 24
 - 25 to 34
 - 35 to 44
 - 45 to 54
 - 65 to 74
 - 75 & beyond



DATA PRE-PROCESSING

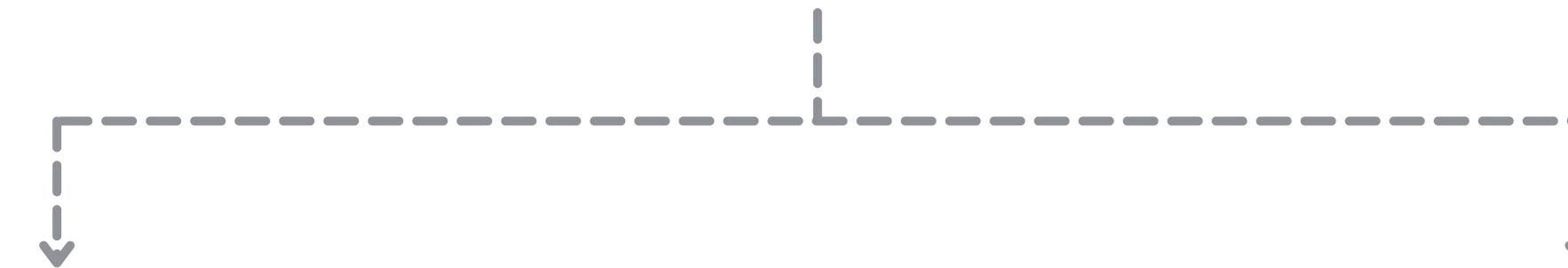


MODELLING

PRECRIME TASK FORCE



TIME SERIES MODELS



TRADITIONAL TIME SERIES MODELS

Univariate Models

- Holts Winter Exponential Smoothing
- ARIMA
- SARIMAX
- Prophet

MACHINE LEARNING MODELS

Regressors

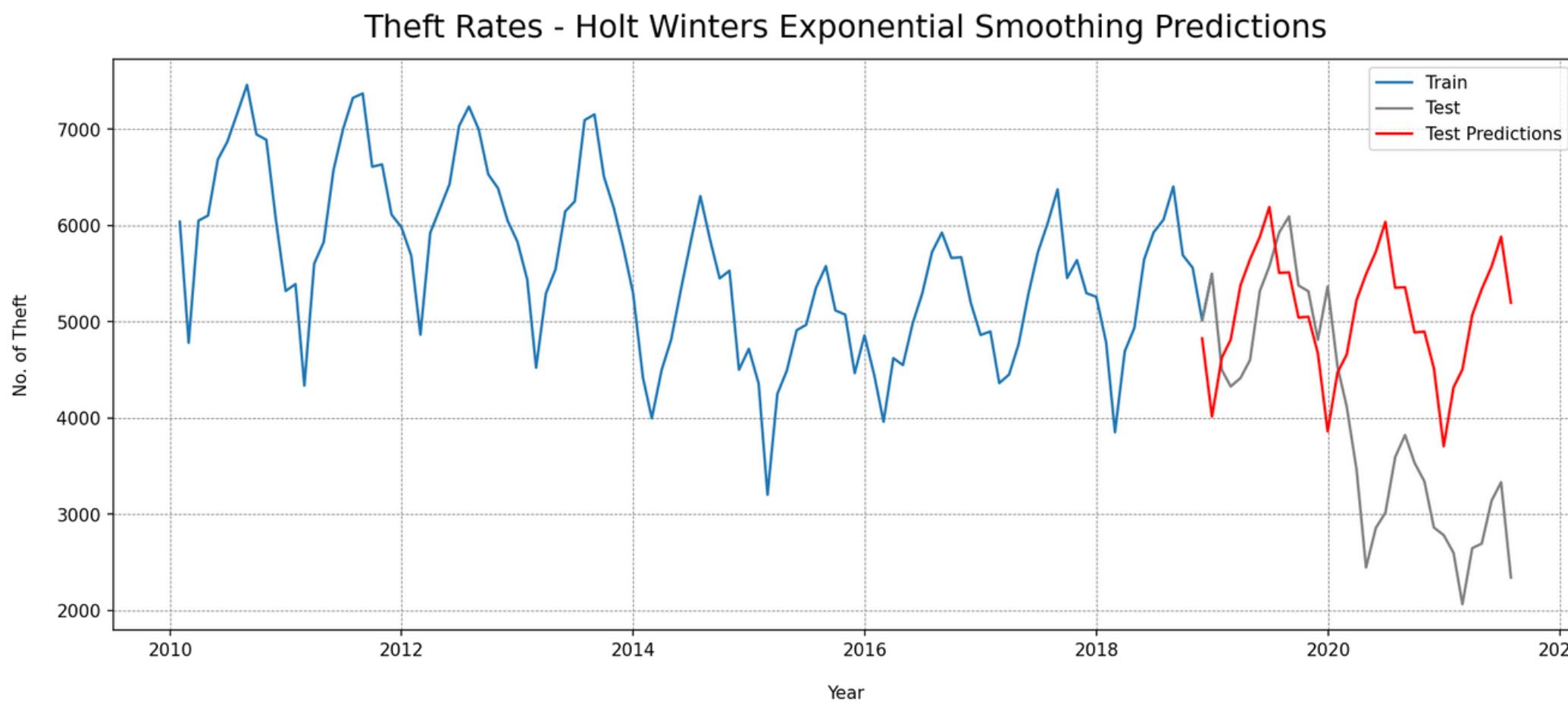
- Gradient Boosting
- Adaptive Boost
- XG Boost





THEFT

Traditional Time Series Models



HOLT WINTERS
ES

1685.73

RMSE

44.2%

MAPE

| MODEL | RMSE | MAPE |
|---------|---------|------|
| SARIMA | 2122.66 | 0.57 |
| SARIMAX | 2972.70 | 0.72 |
| PROPHET | 2136.79 | 0.56 |

HIGHLIGHTS

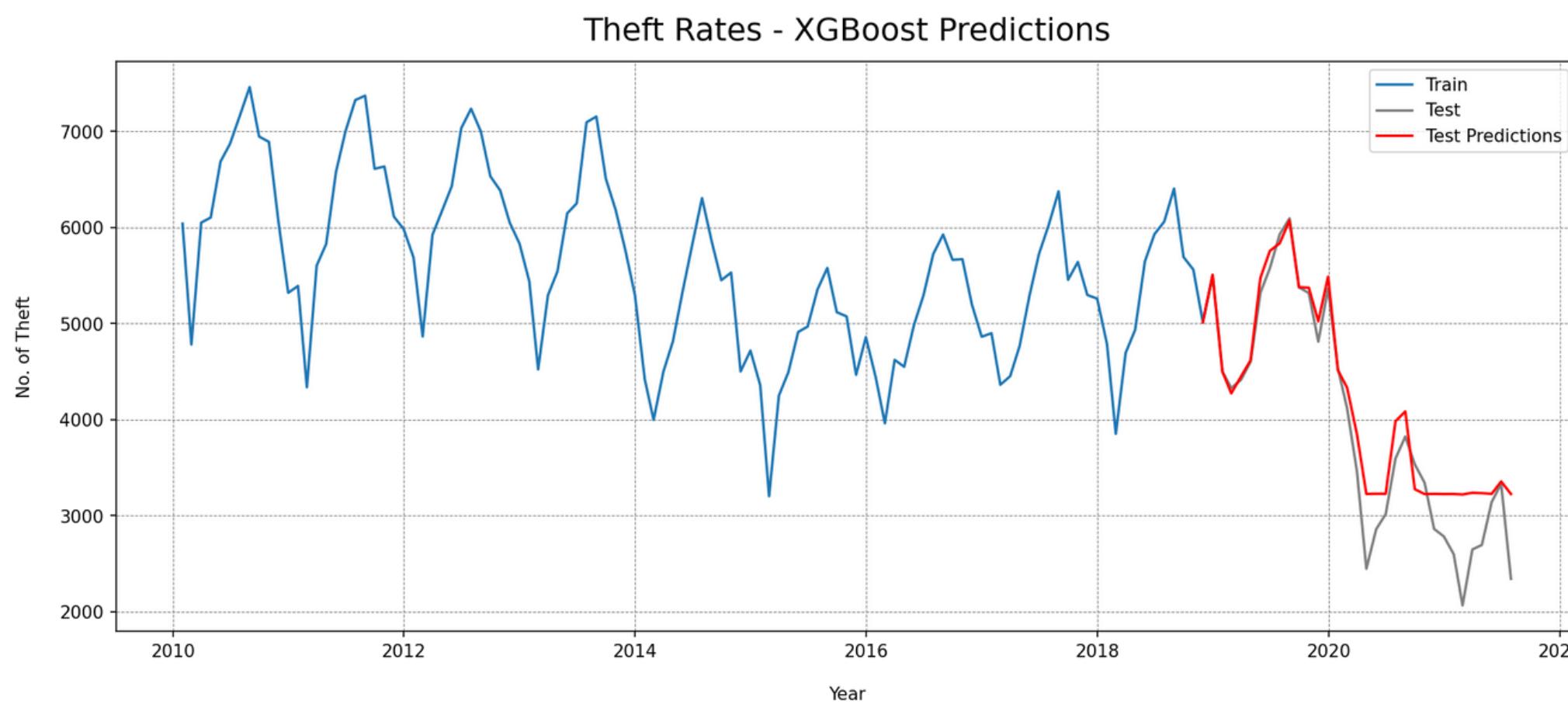
Model did not capture the exponential drop from 2020 to 2021.

Further investigations shows that due to the pandemic, COVID-19 and regulations in placed, people are lockdown in their own homes hence the setback



THEFT

Machine Learning Models



XG BOOST

384.89

RMSE

9.3%

MAPE

HIGHLIGHTS

| MODEL | RMSE | MAPE |
|----------------|--------|-------|
| GRADIENT BOOST | 635.09 | 0.107 |
| ADA BOOST | 384.89 | 0.165 |

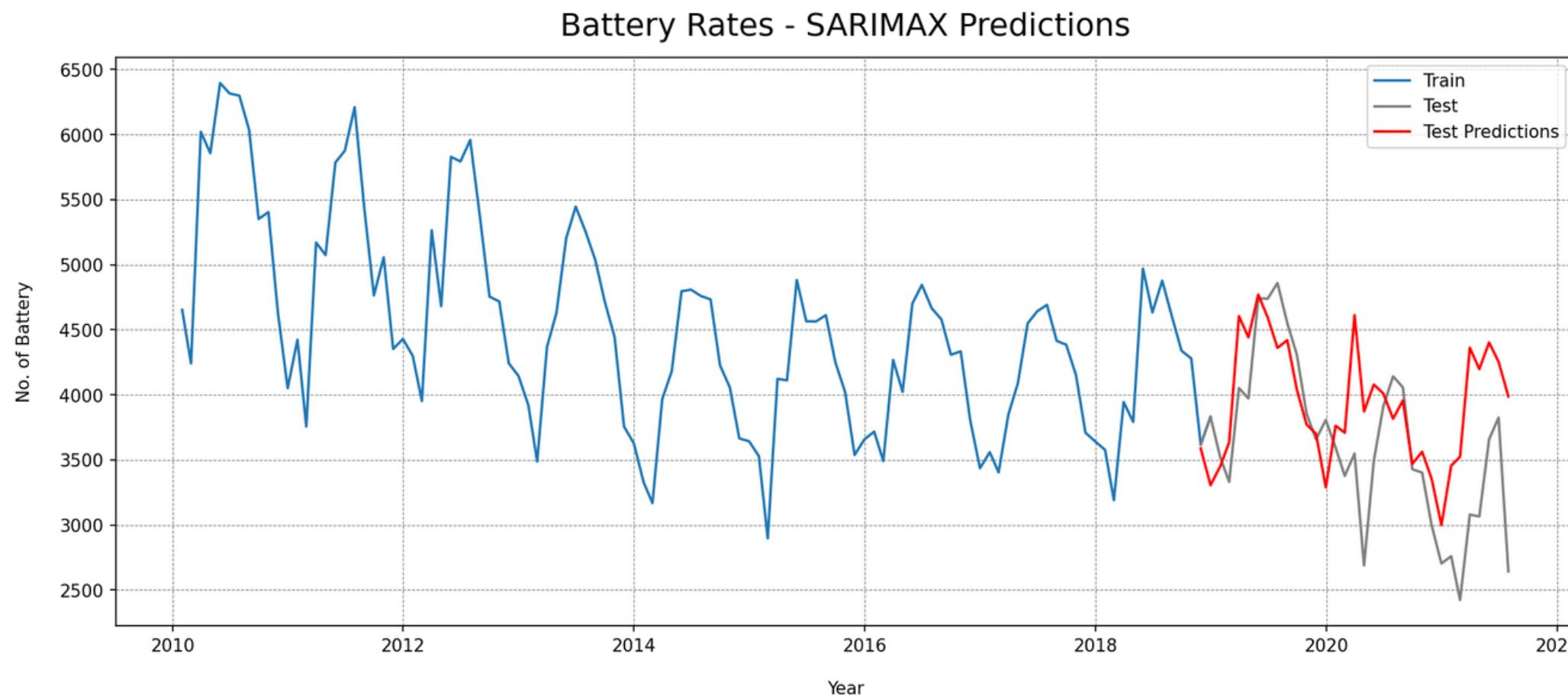
FEATURE IMPORTANCE

- district_2
- residential yard (front/back)
- night
- morning
- evening



BATTERY

Traditional Time Series Models



SARIMAX

605.21

RMSE

14.2%

MAPE

HIGHLIGHTS

Model did not capture the exponential drop from 2020 to 2021.

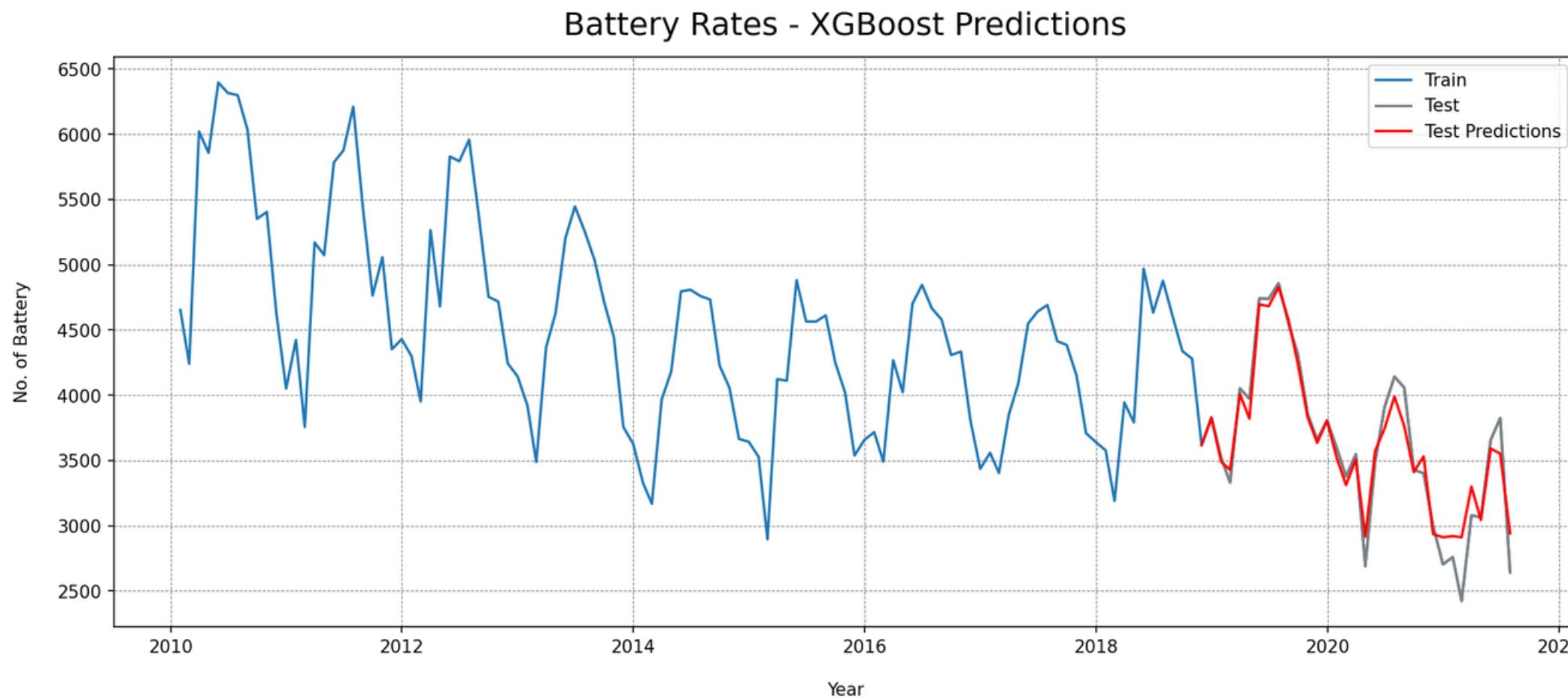
Further investigations shows that due to the pandemic, COVID-19 and regulations in placed, people are lockdown in their own homes hence the setback

| MODEL | RMSE | MAPE |
|-----------------|--------|------|
| HOLT WINTERS ES | 957.50 | 0.24 |
| SARIMA | 908.92 | 0.22 |
| PROPHET | 762.83 | 0.17 |



BATTERY

Machine Learning Models



XG BOOST

155.62

RMSE

3.4%

MAPE

| MODEL | RMSE | MAPE |
|----------------|--------|-------|
| GRADIENT BOOST | 201.10 | 0.044 |
| ADA BOOST | 230.35 | 0.053 |

HIGHLIGHTS

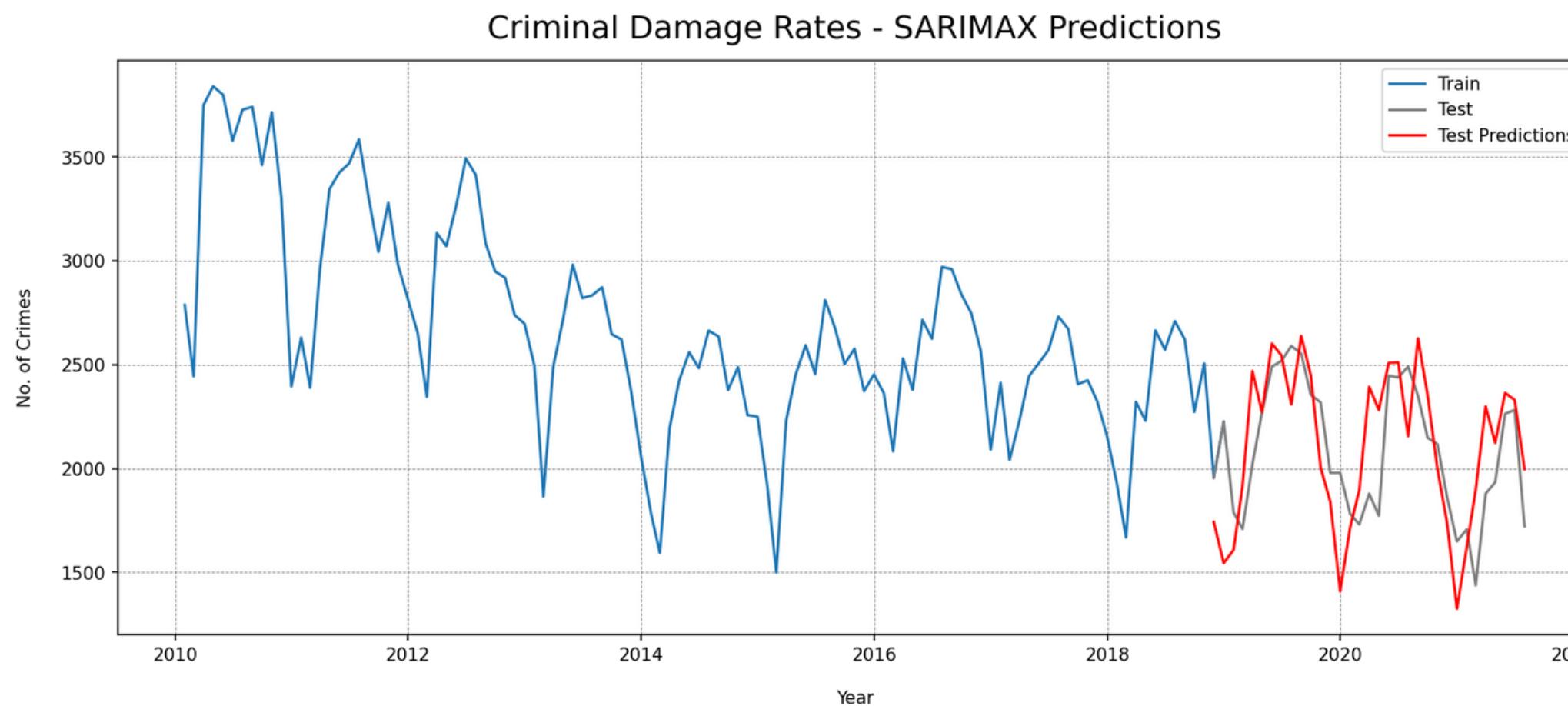
FEATURE IMPORTANCE

- district_7
- domestic_true
- domestic
- battery simple
- sidewalk
- evening
- district_3
- district_8
- street
- residence
- night



CRIMINAL DAMAGE

Traditional Time Series Models



SARIMAX

290.74

RMSE

11.8%

MAPE

HIGHLIGHTS

MODEL

HOLT WINTERS ES

SARIMA

PROPHET

RMSE

293.64

MAPE

0.12

0.15

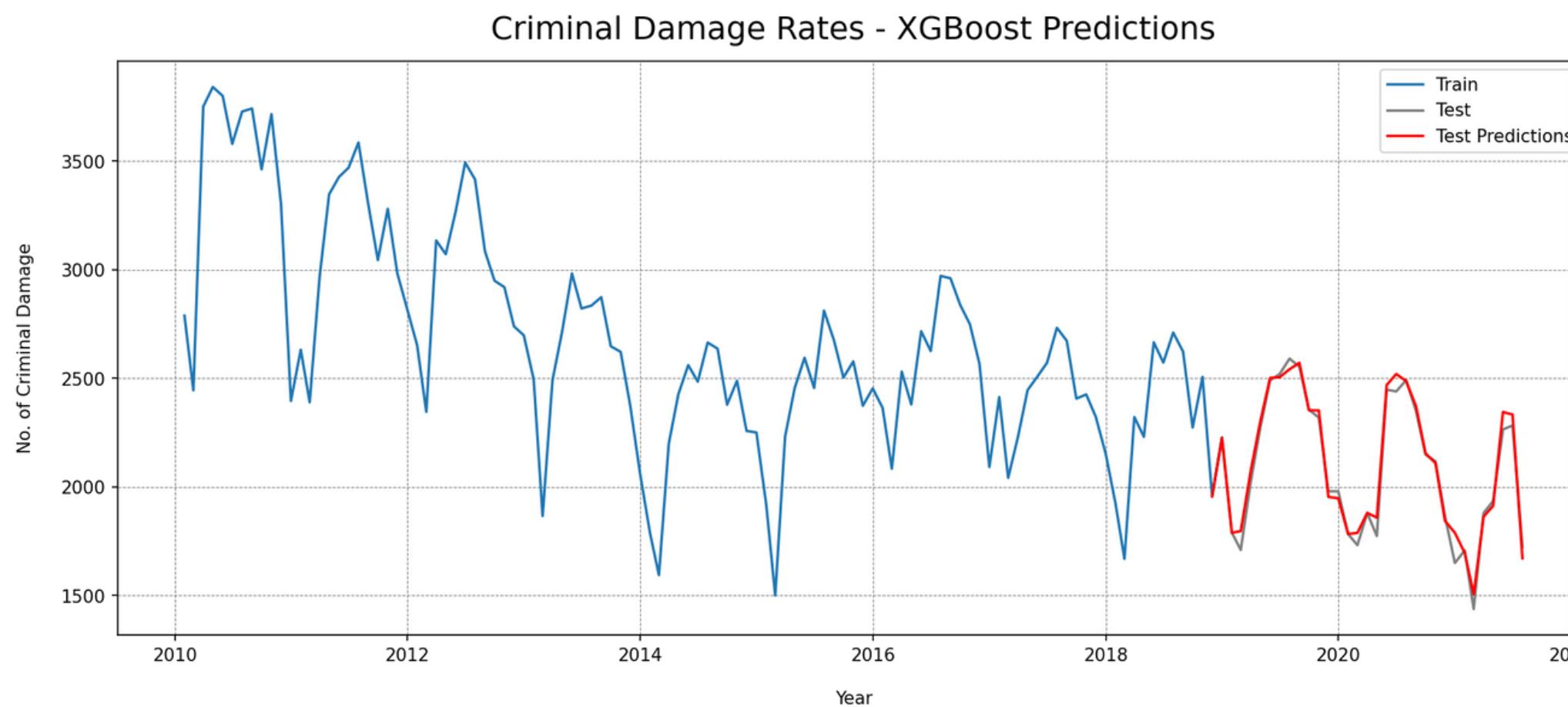
0.115

Criminal Damage rates generally followed the same trends as compared to previous 2 crime types.



CRIMINAL DAMAGE

Machine Learning Models



XG BOOST

46.64

RMSE

1.7%

MAPE

HIGHLIGHTS

FEATURE IMPORTANCE

- To property
- Street
- Residence
- To vehicle
- night
- district_4
- evening
- district_7
- district_25
- district_3

| MODEL | RMSE | MAPE |
|----------------|-------|-------|
| GRADIENT BOOST | 73.71 | 0.026 |
| ADA BOOST | 75.50 | 0.030 |



CONCLUSION & RECOMMENDATIONS

PRECRIME TASK FORCE

SUMMARY

THEFT

BATTERY

CRIMINAL DAMAGE

TTS

RMSE: 1685.73

MAE: 44.6%

METHOD: HOLTS WINTER ES

RMSE: 605.21

MAE: 14.2%

METHOD: SARIMAX

RMSE: 290.74

MAE: 11.8%

METHOD: SARIMAX

MLR

RMSE: 384.89

MAE: 9.3%

METHOD: XG BOOST

RMSE: 155.61

MAE: 3.4%

METHOD: XG BOOST

RMSE: 46.64

MAE: 1.7%

METHOD: XG BOOST

Regressor models have a better advantage due to features used to predict crimes





CONCLUSION

- Predicting crimes before they happen is simple to understand, but it takes a lot more than understanding the concept to make it a reality
- The first phase of deployment - Creation of UI for police to access historical values, predictions and forecast of crimes using XG Boost methods
- *Forecasting done using lag method of 1 month*

HIGHLIGHT 1

Based on limitations, gather more insightful and accurate data.

Also proper selection feature to better model the predictions, improving accuracy of prediction and minimized loss

HIGHLIGHT 2

Future Scope

Aside from using just a regression and times-series techniques, we can further look into the use of spatial data to predict crime as there many datas on locations

HIGHLIGHT 3

RECOMMENDATIONS

WORKING WITH DOMAIN EXPERTS

Relevant parties should be involved in this project such as representatives from the government, academic and private organization to provide insightful information and domain knowledge of the individual industry.

Tools and information are needed to predict a crime before it occurs. State officers themselves are humans and may have limitations digesting, processing and inferring big amount of data in order to predict crimes.

COMMUNITY INVOLVEMENT

Insufficient intel and equipments to aid in the understanding of where crime may take place. By engaging the members of public to educate them of potential crimes and reporting to the police when issues arises.

RESEARCH & DEVELOPMENT

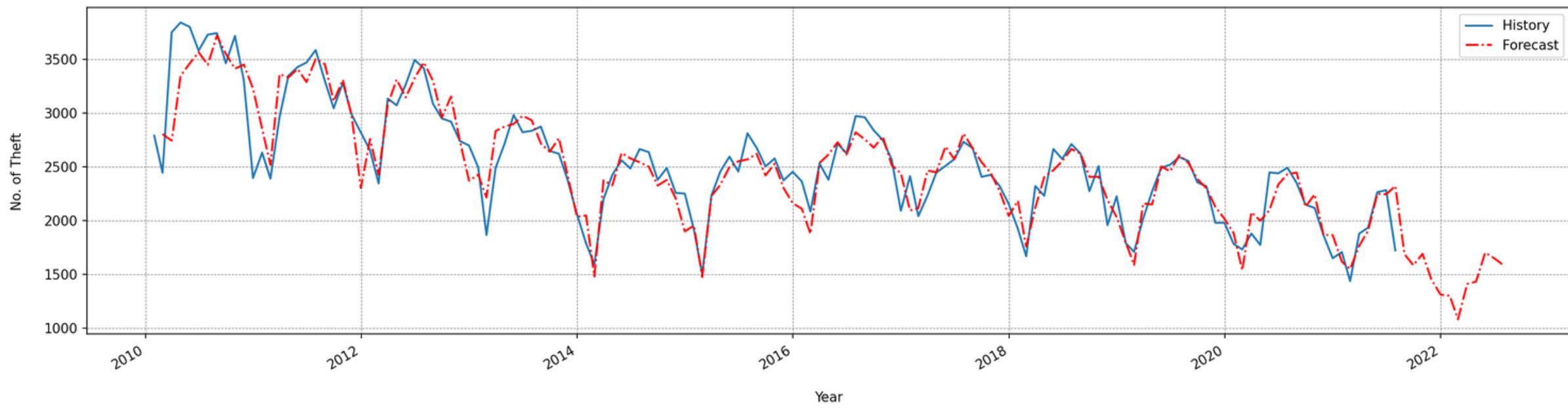
With the help of modern technology, we will be able to reinforce our intelligence towards different crimes in their city and well prepared to enforce law and ensure the safety of members of public.

CRIMINAL DAMAGE

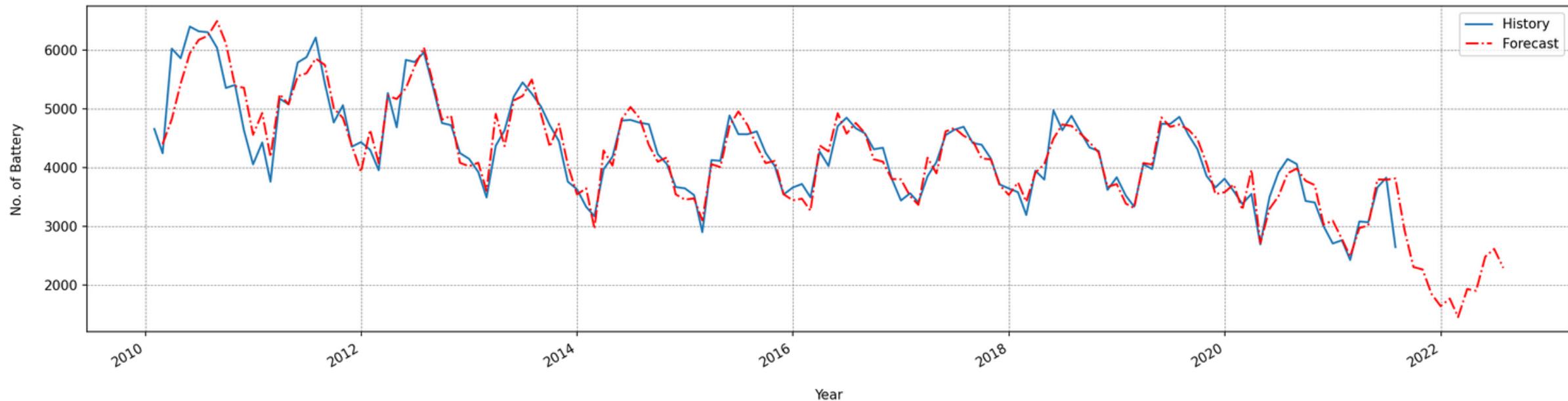
BATTERY

THEFT

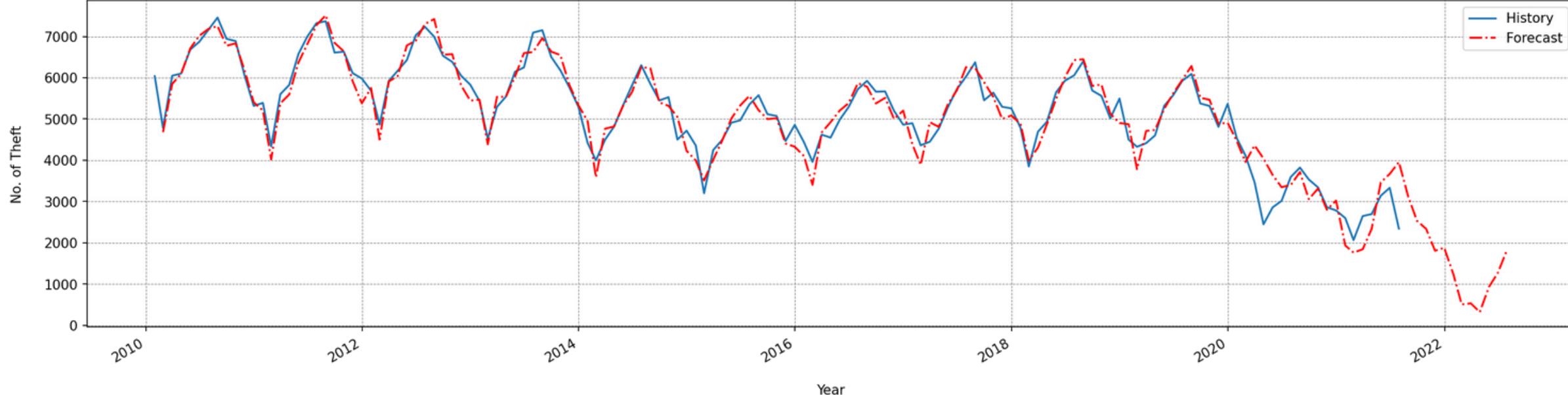
Criminal Damage Rates Forecast (Towards 2022) - SARIMAX Predictions



Battery Rates Forecast (Towards 2022) - SARIMAX Predictions



Theft Rates Forecast (Towards 2022) - Hotl Winter ES Predictions



FORECAST

