Crime Rate Forecasting in Chicago

Agenda

01

Objective & Dataset

02

Crime Trend & Data Partitioning

03

Forecasting Models & Performances

04

Overall Performance Comparison

05

Recommendations & Conclusion

Objective & Dataset

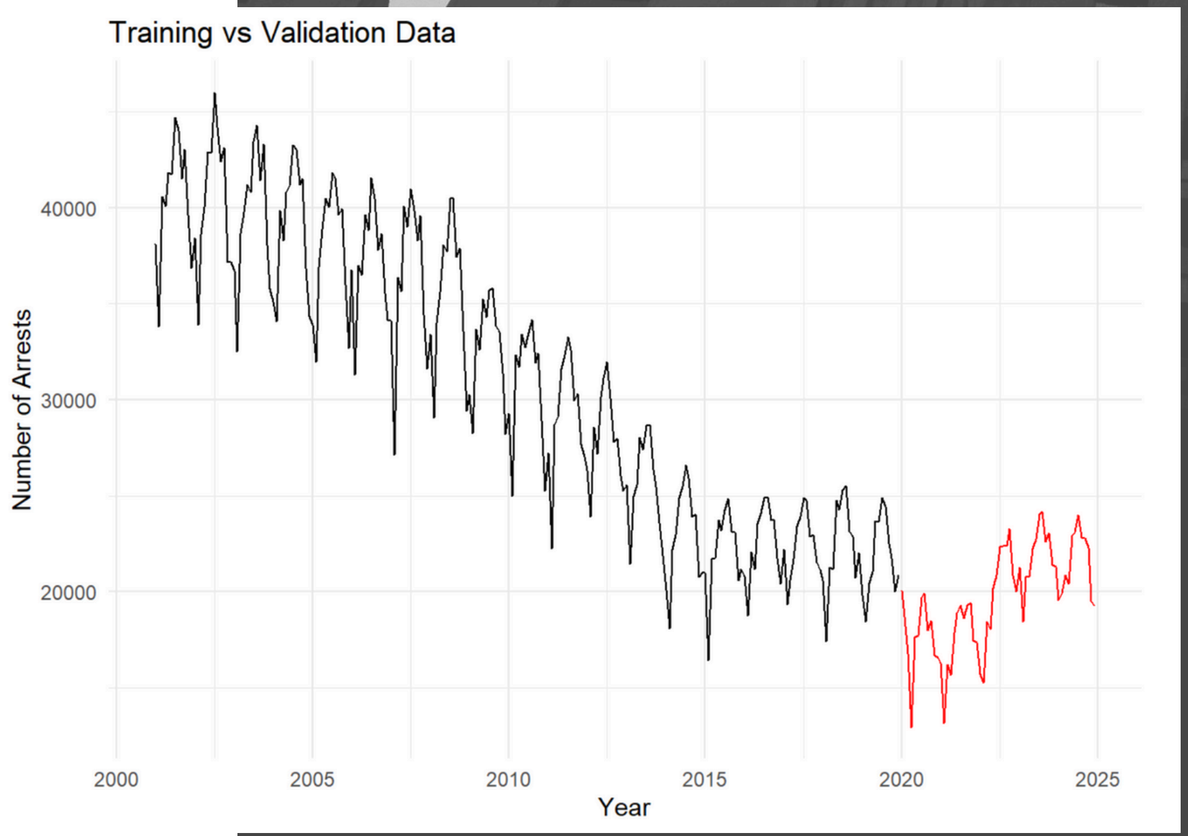
- Analyze and forecast crime rates using historical crime data for Chicago.
- Dataset: Chicago Crime Records (2001–Present)
  catalog.data.gov/dataset/crimes
- 7 Million+ crime records: Monthly aggregation & transformation
- Focused Crimes: THEFT and BATTERY

	Primary.Type	count
	<chr>	<int>
1	THEFT	1750314
2	BATTERY	1505260
3	CRIMINAL DAMAGE	939862
4	NARCOTICS	758763
5	ASSAULT	549217
6	OTHER OFFENSE	513850

Crime Trend & Data Partitioning

- Great fluctuation and about 50% rate decline over time
- Seasonal peaks observed during summer months
- Brutal winter's correlation to lower rates
- Training set from Jan 2001 to Dec 2019
- Validation set from Jan 2020 to Dec 2024.
- Covid Era

Monthly Arrest Counts Over Time



Forecasting Models & Performances

01

Regression

02

Exponential Smoothing

03

Auto Arima

04

Neural Network

05

Snaive

06

Simple Average

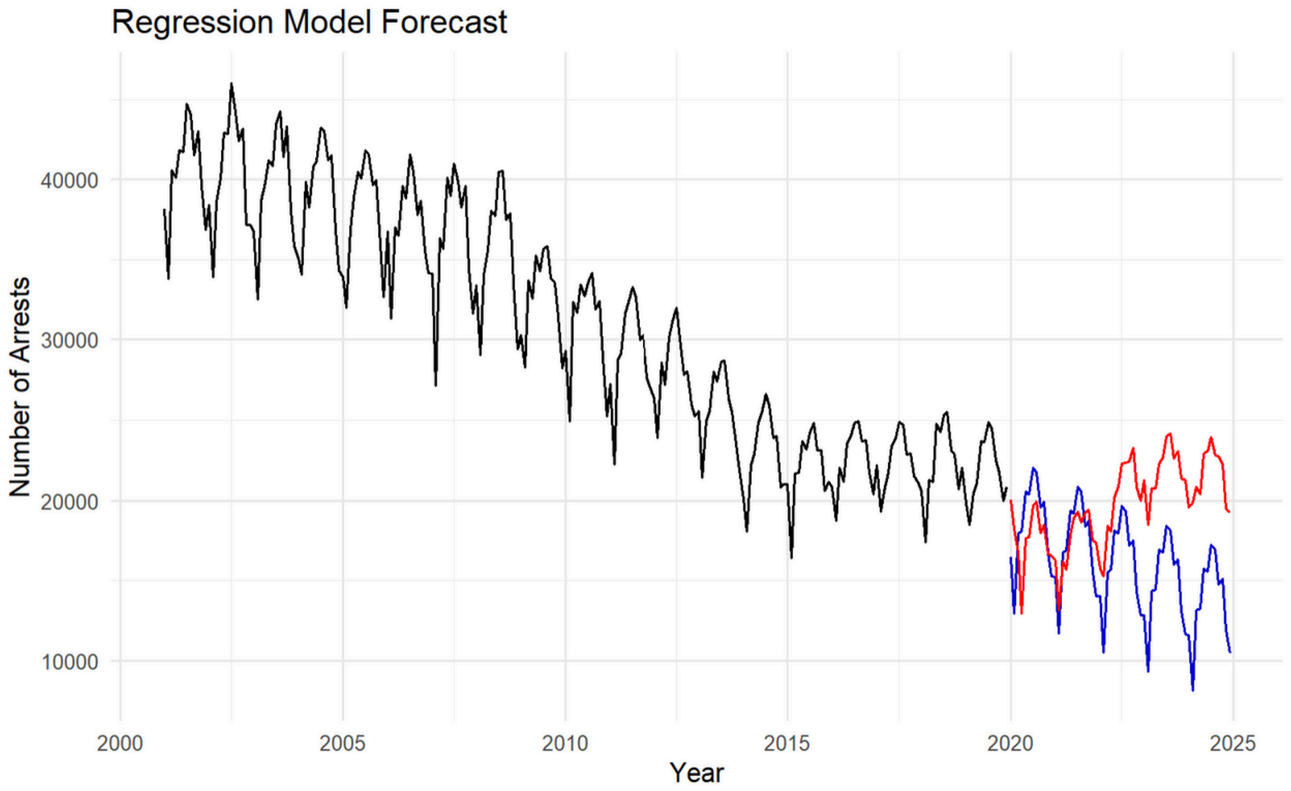
07

Trimmed Mean

Regression Model

- Captures long-term crime trends and seasonal spikes effectively but biggest covid hit
- On average, we are off by about 4516 arrests per month
- MAPE suggests forecast is off by 22.27%, on average

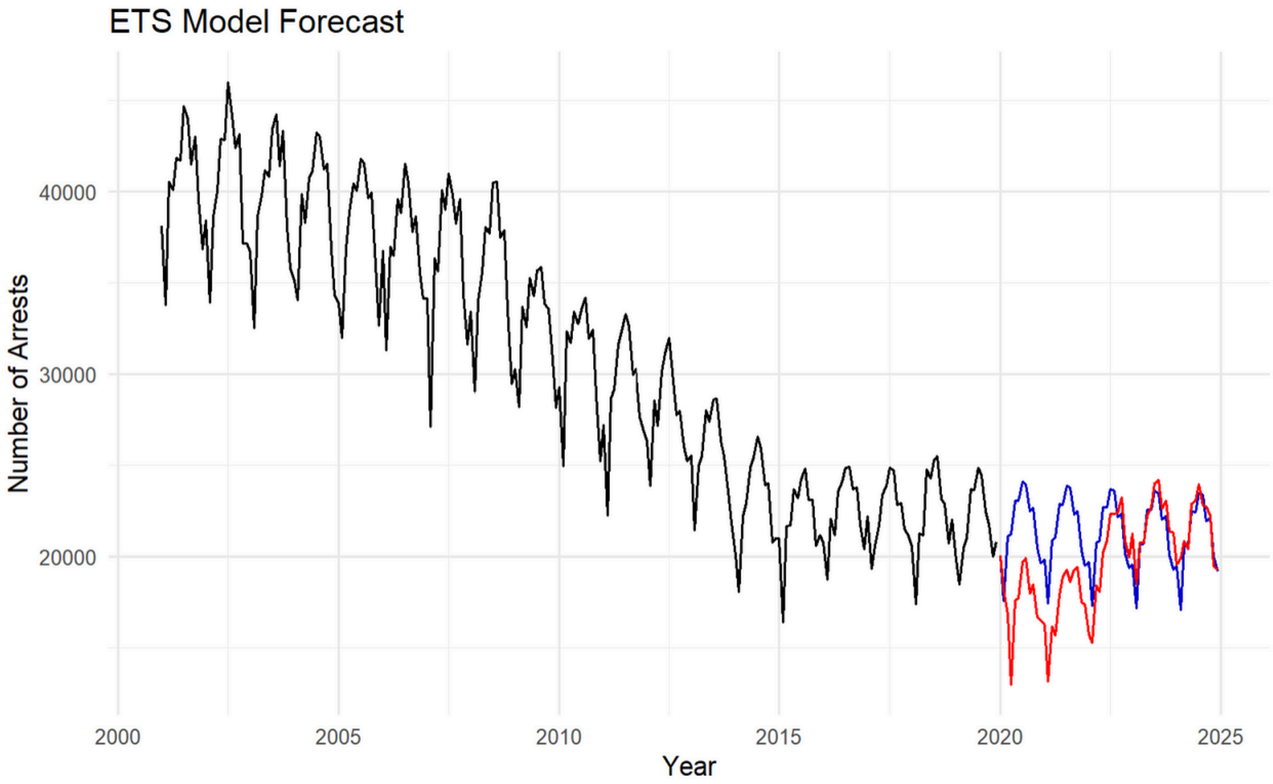
MAE	MAPE
4515.954409	22.271051



Exponential Smoothing Model

- Smooths seasonal patterns but over-smooths crime spikes.
- On average, we are off by about 2221 arrests per month
- MAPE suggests forecast is off by 12.59%, on average

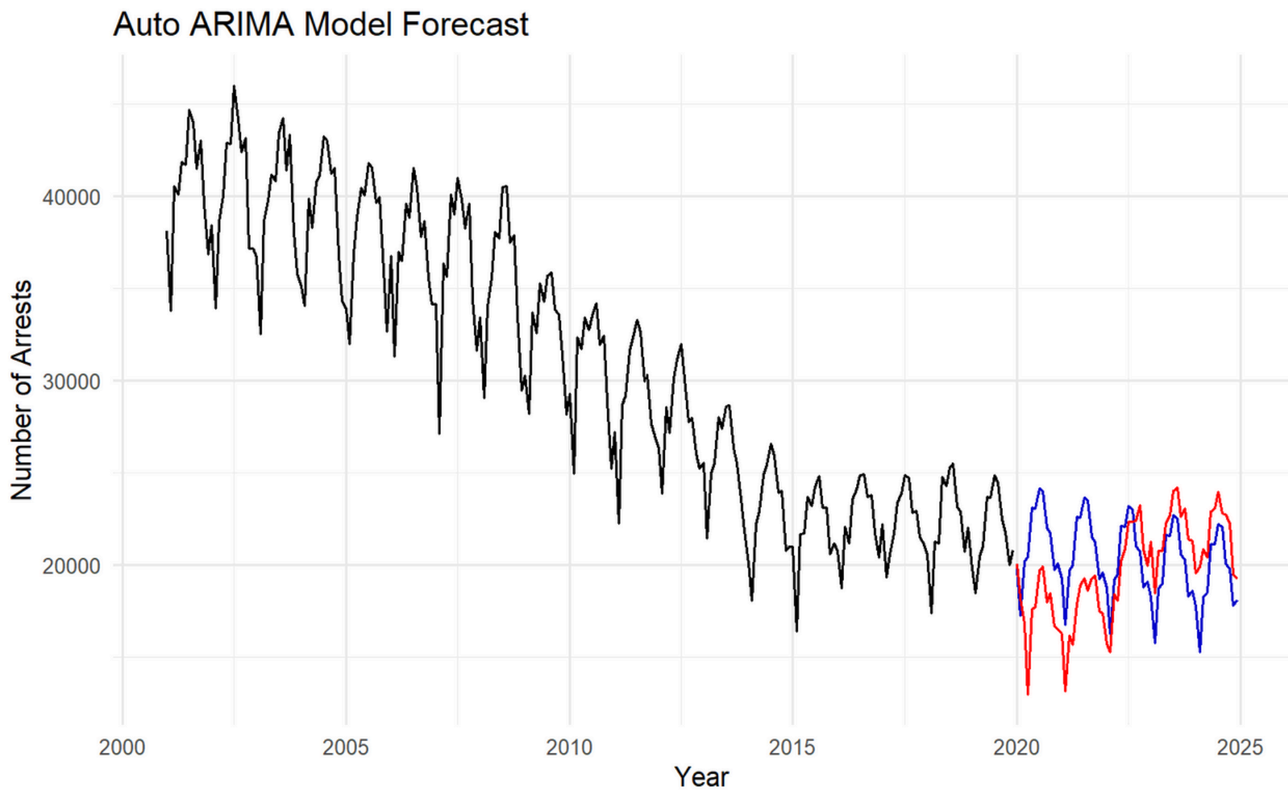
MAE	MAPE
2221.3531085	12.5922999



Auto ARIMA Model

- Handles short-term trends but struggles with seasonality.
- On average, we are off by about 2562 arrests per month
- MAPE suggests forecast is off by 13.85%, on average

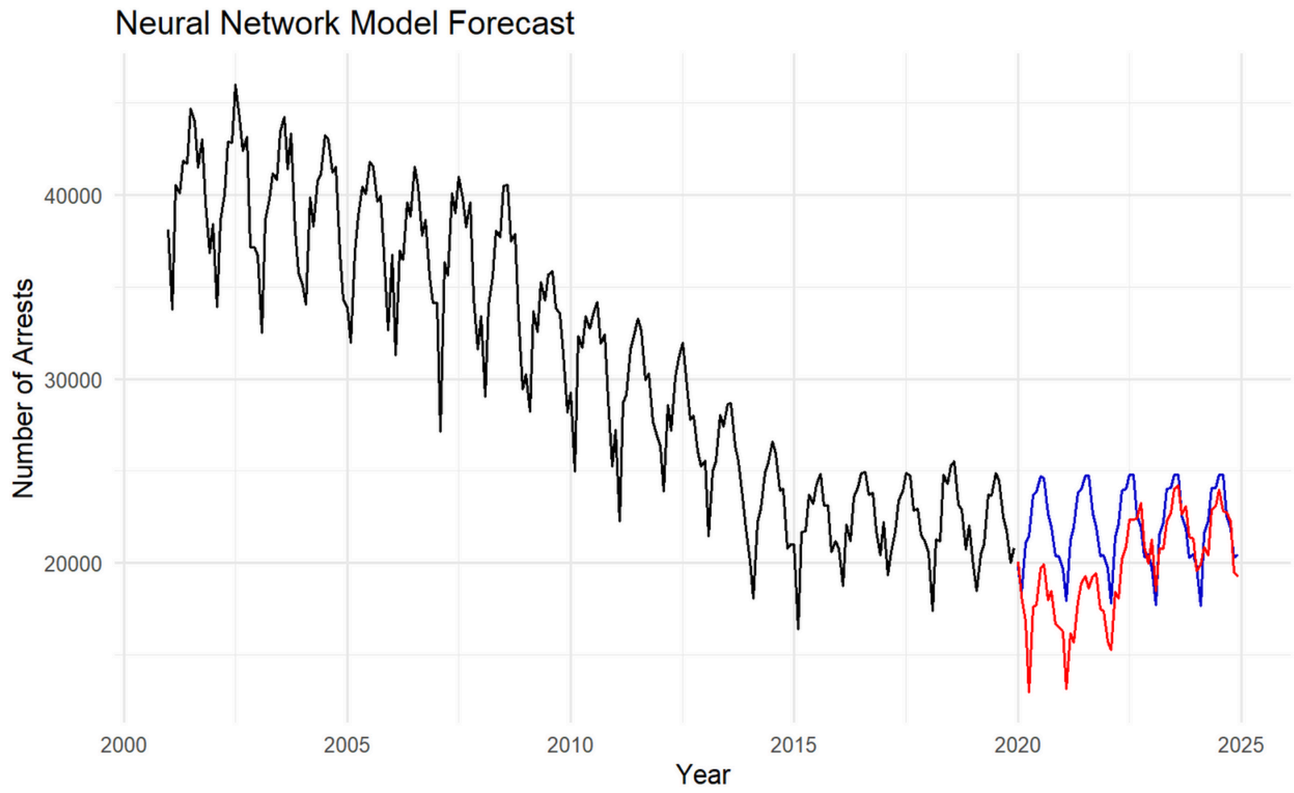
MAE	MAPE
2562.3604198	13.8464051



Neural Network Model

- Captures complex non-linear trends but prone to overfitting.
- On average, we are off by about 2624 arrests per month
- MAPE suggests forecast is off by 14.66%, on average

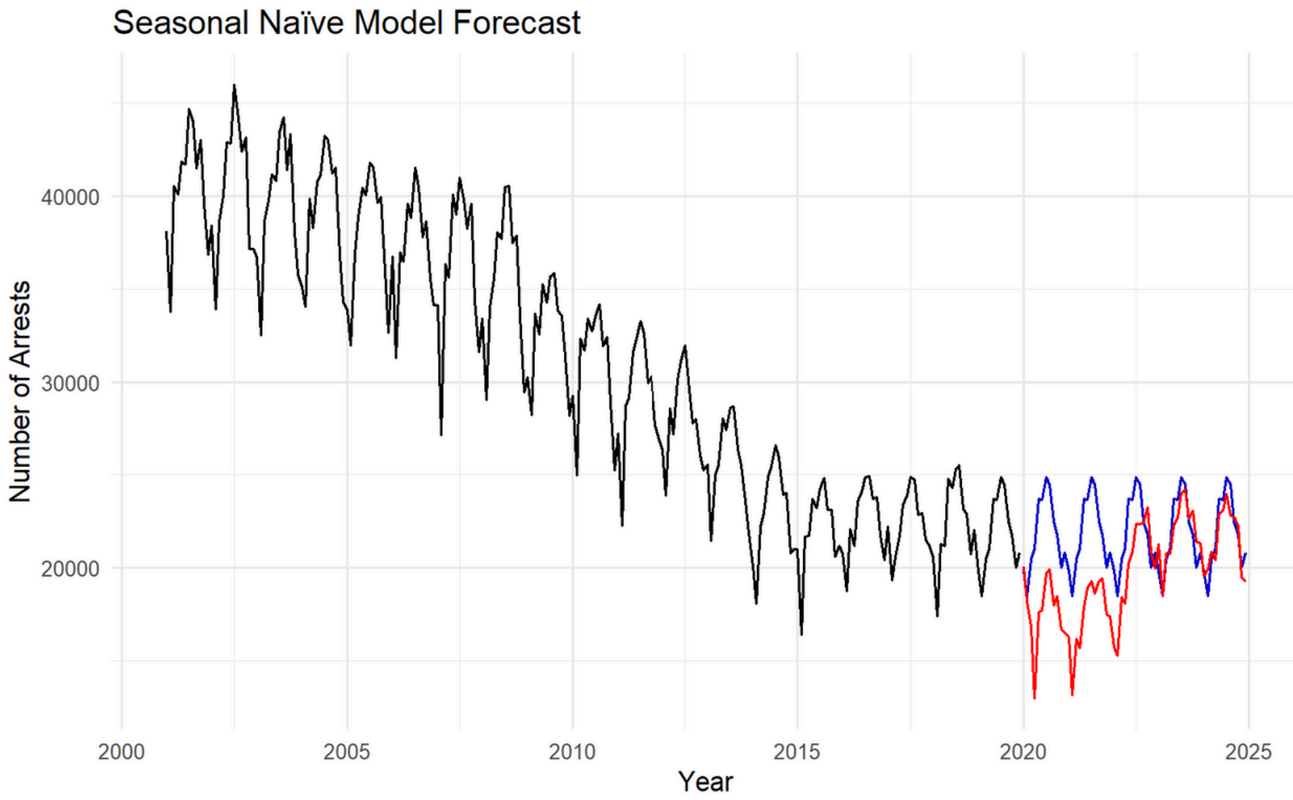
MAE	MAPE
2624.3099135	14.6617387



Seasonal Naïve Model

- Uses past seasonality to predict future crimes, serving as a baseline
- On average, we are off by about 2444 arrests per month
- MAPE suggests forecast is off by 13.74%, on average

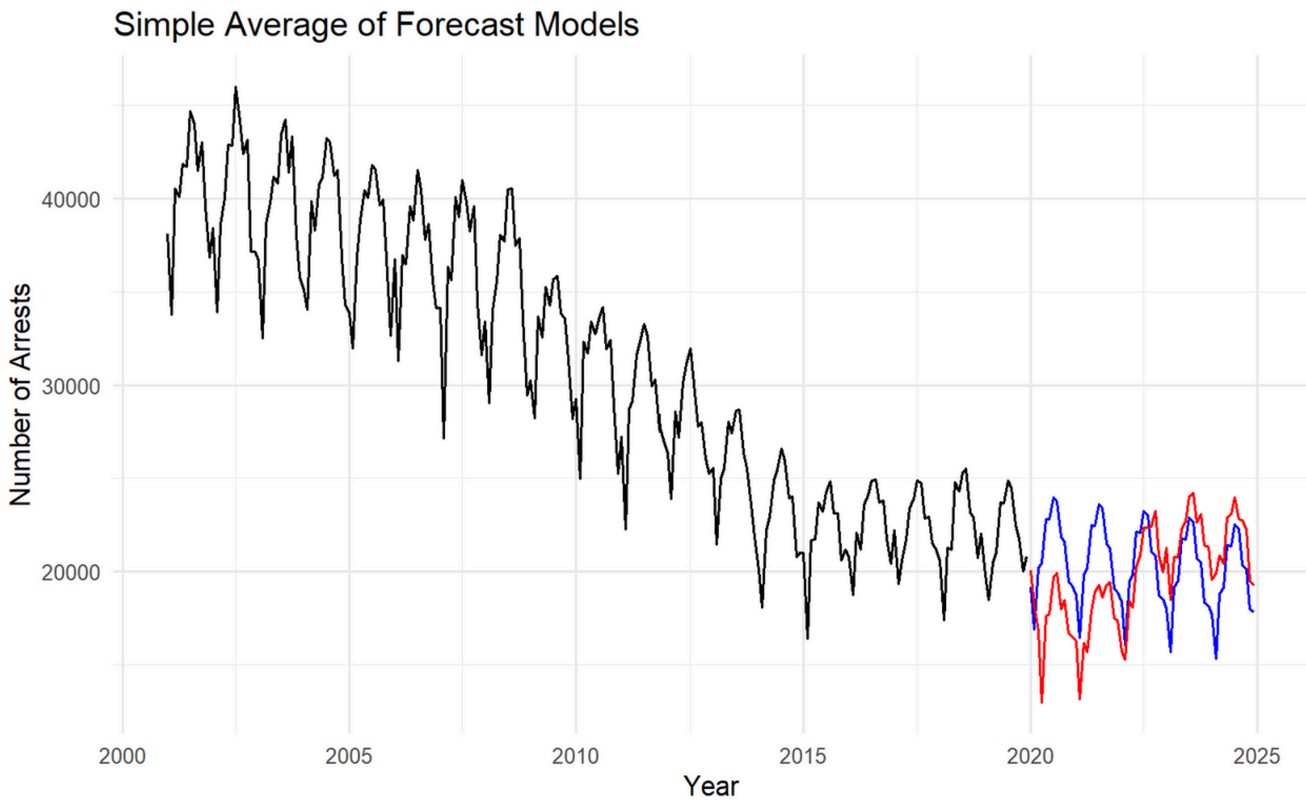
MAE	MAPE
2443.6500000	13.7401153



Simple Average Forecast

- Averages all models, smoothing extreme fluctuations.
- On average, we are off by about 2448 arrests per month
- MAPE suggests forecast is off by 13.25%, on average

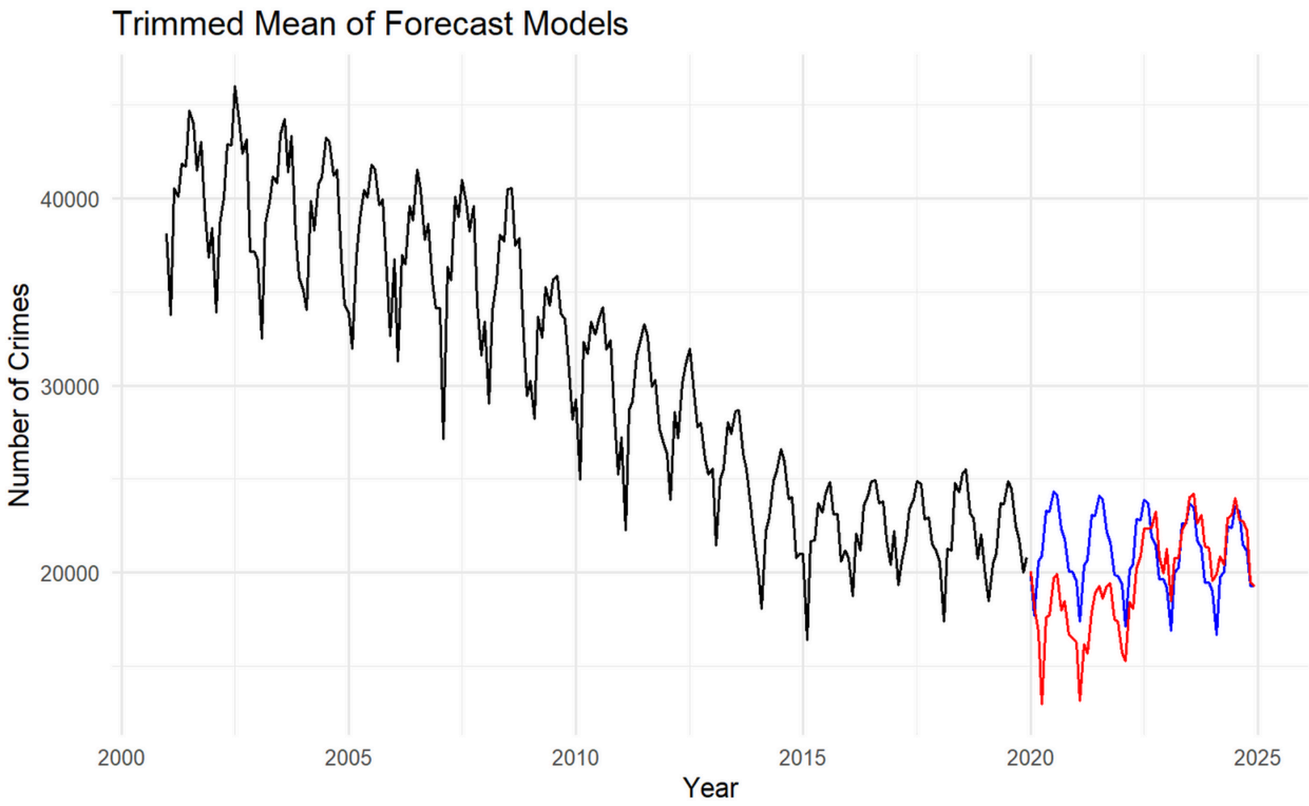
MAE	MAPE
2448.3788754	13.2477256



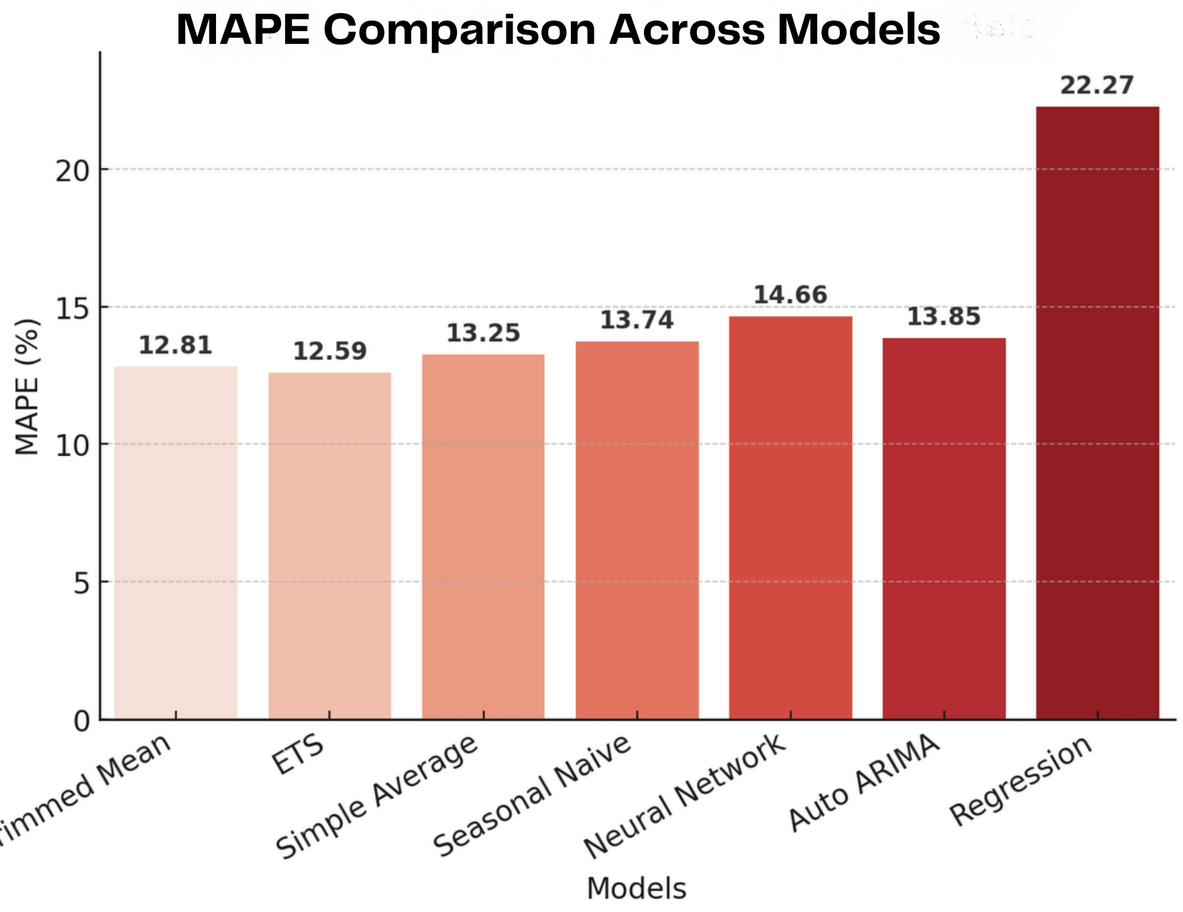
Trimmed Mean Forecast

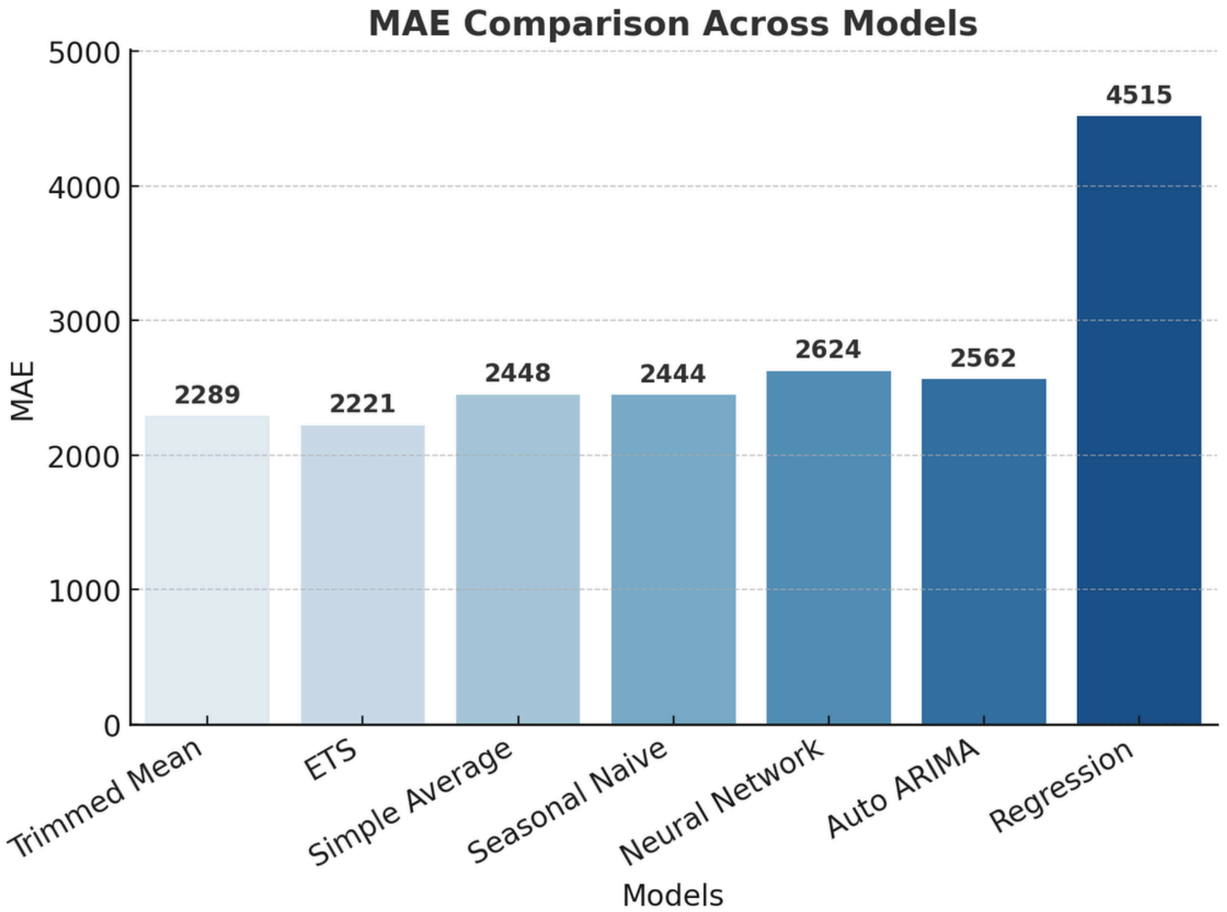
- Excludes outliers before averaging model forecasts.
- On average, we are off by about 2288 arrests per month
- MAPE suggests forecast is off by 12.81%, on average

MAE	MAPE
2288.915916	12.806158



Overall Performance Comparison





Conclusion

- The Trimmed Mean & ETS models demonstrated the best fit forecasters, providing reliable insights for future crime prevention efforts.
- Accurate time forecasts can empower law enforcement to transition from reactive responses to proactive strategies & optimize resources, reduce crime, and ensure public safety more effectively.

Recommendations

- Increase patrols by 25% during peak crime months (June–August).
- Continuously refine models by integrating socio-economic data.
- Identify high-crime areas via appropriate & extensive data