

# Who is policing the police?

Team #3 case study : 12/12/2020

# Problem

- Increased visibility of police misconduct in media
- Lack of data on police crime



Source: Chad Davis

# Why should we care?

- Allow comparisons of departmental crime rates
- Allow comparisons of types of police crime
- Help create accountability by understanding agency's response to police crime
- Enhance police-community relationship through data-driven policy solution

# 650

crimes committed by the police  
officers attracted media attention



# Agenda



## Media Attention

Analysis of newspaper headlines



## Geospatial Analysis

Crime data with respect to a county



## Time Series

Crimes and trends over time



## Crimes

What were the most committed crimes?



## Offenders

Possible connection between officers' ranks and conviction times



## Careers

How many of them kept their jobs after committing crimes?

# 01

## Media Attention

NLP analysis of the  
newspapers  
headlines

# Predominant Topics Discovered

01

## Drugs & Alcohol

Drug-related offenses  
and DUI, disorderly  
conduct

02

## Domestic Abuse

Spouse battery  
and injury  
infliction

03

## Sexual Assault

Child pornography and  
underage assault,  
stalking and attempting  
lewd acts, rape

04

## Fraud

False report filing,  
embezzlement,  
theft





## Overall sentiment of the headlines

73%

Negative

25%

Neutral

1.4%

Positive

# 02

## Geospatial Analysis

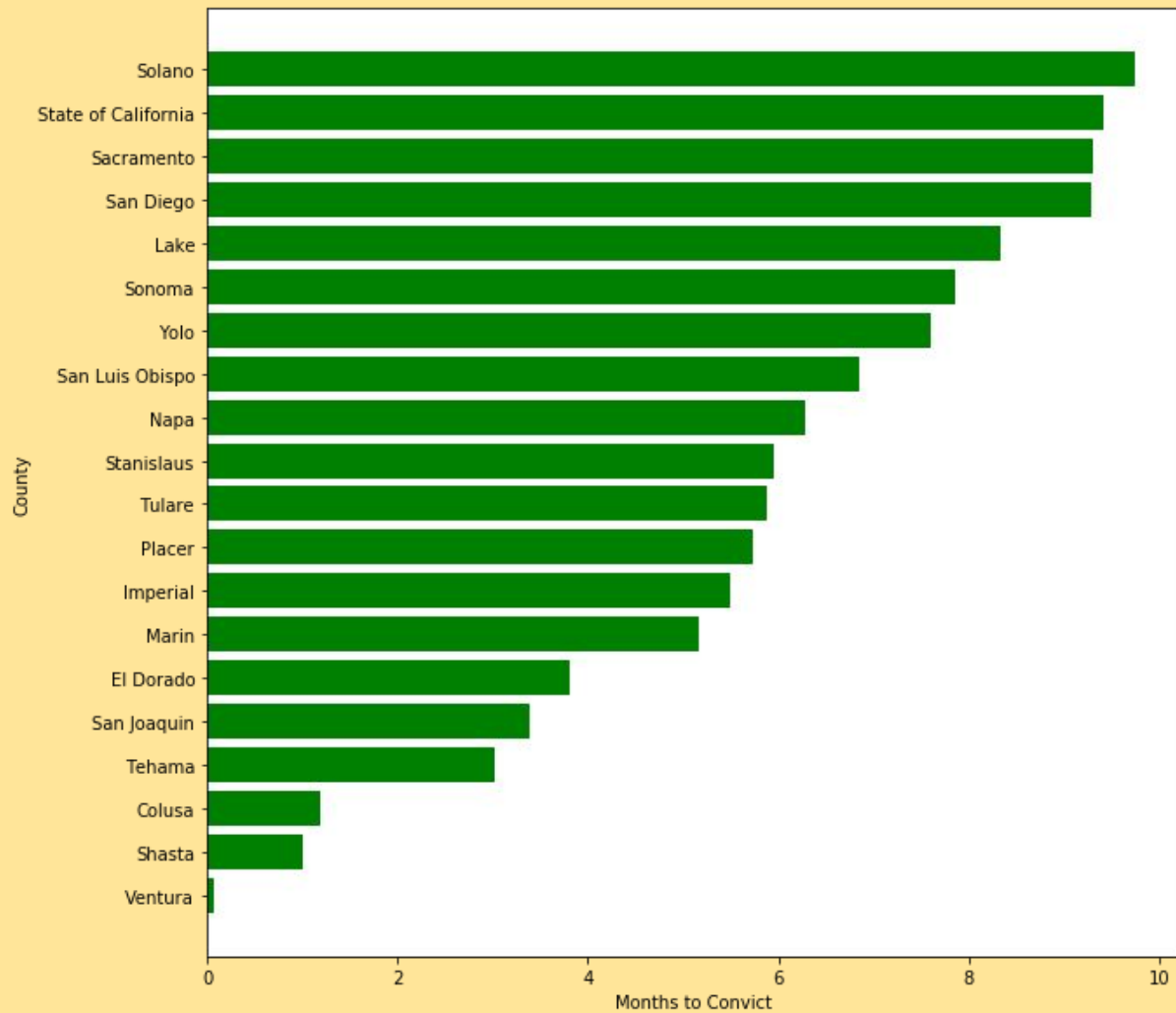
Discoveries per county

## Average Values for Counties, sorted by months to convict

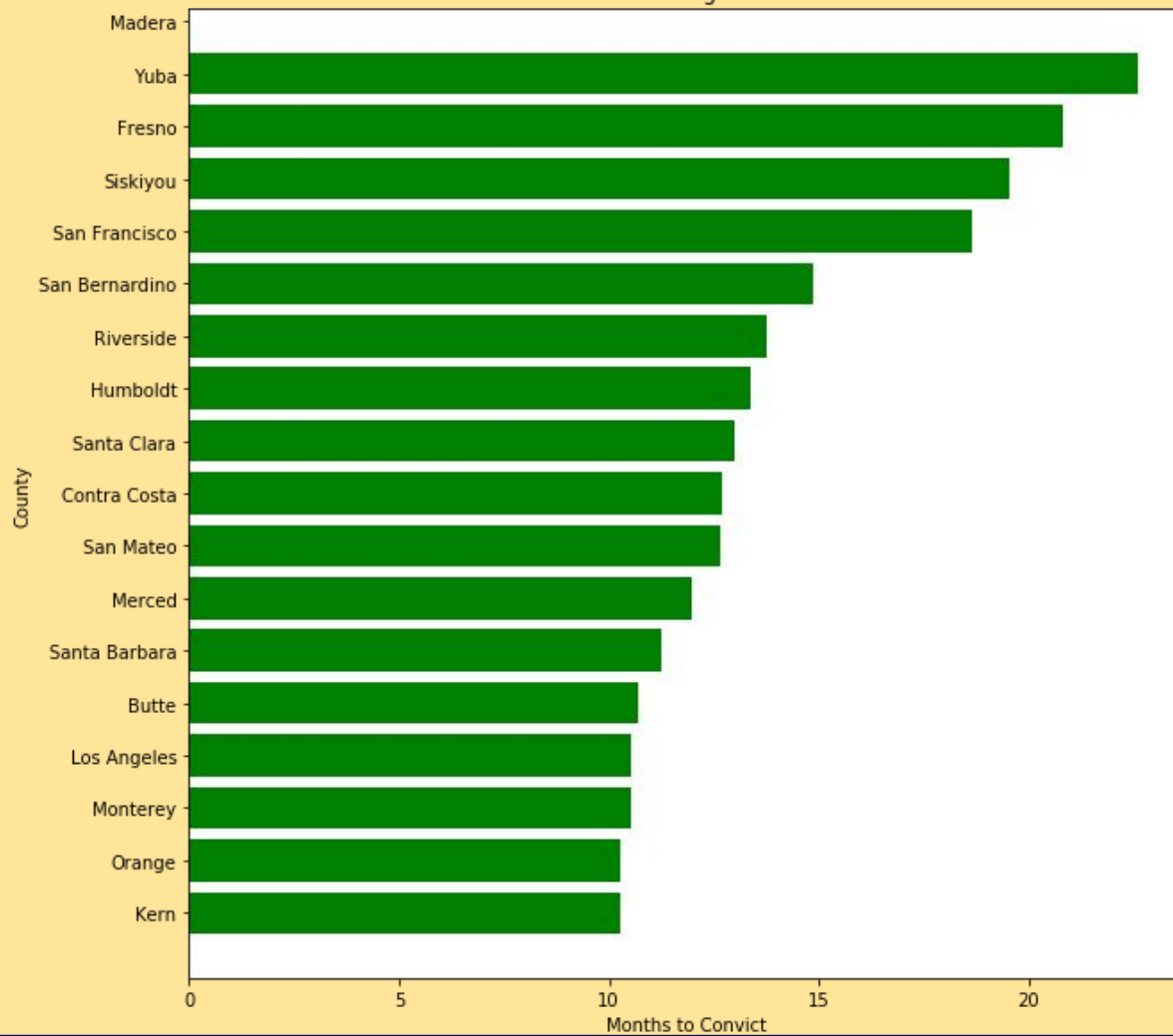
County	Figure
Ventura	0.0
Shasta	1.0
Sacramento	9.3
Solano	9.7

**Average for State of California: 9.4**

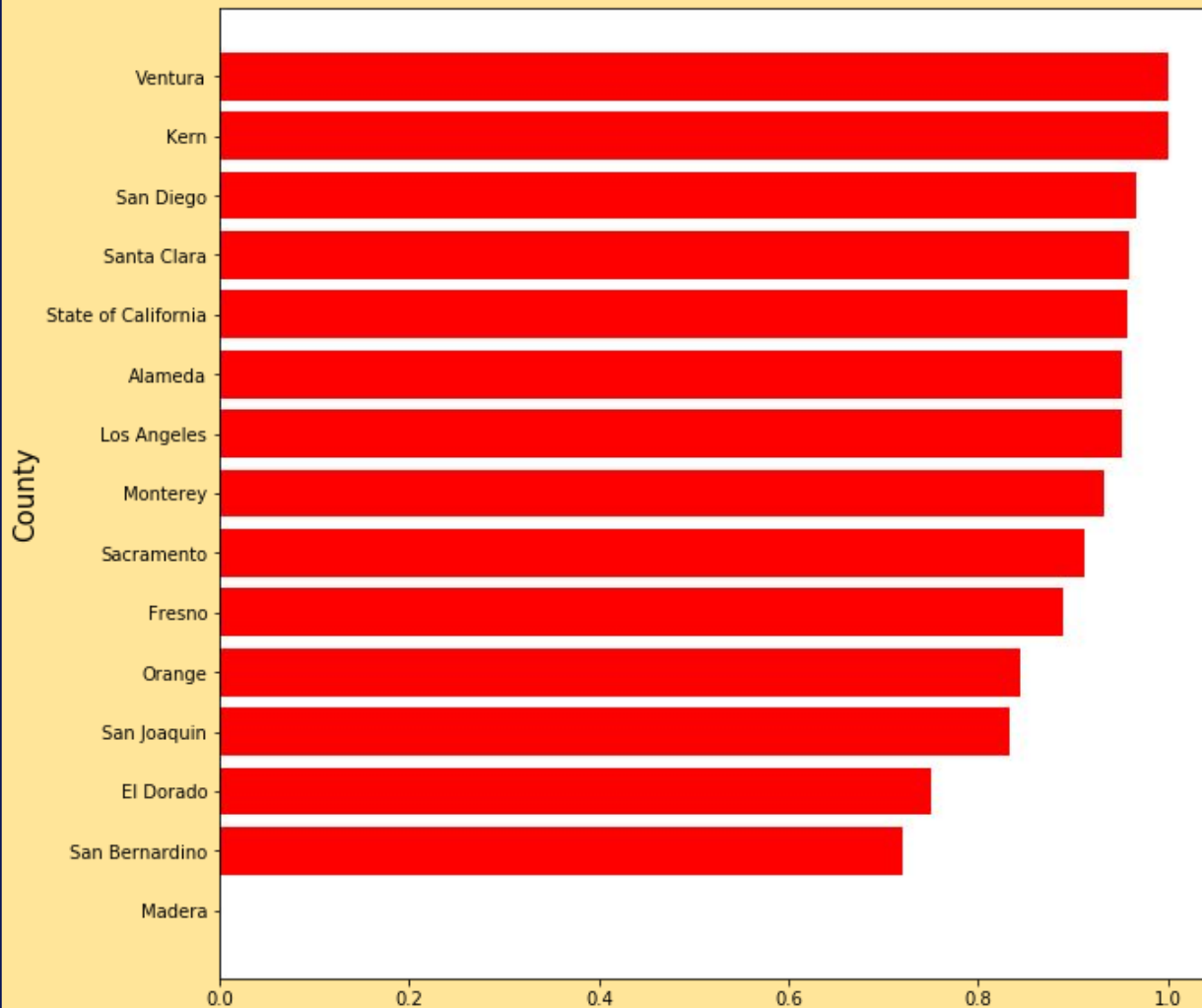
20 Counties with the Shortest Conviction Times



20 Counties with the Longest Conviction Times







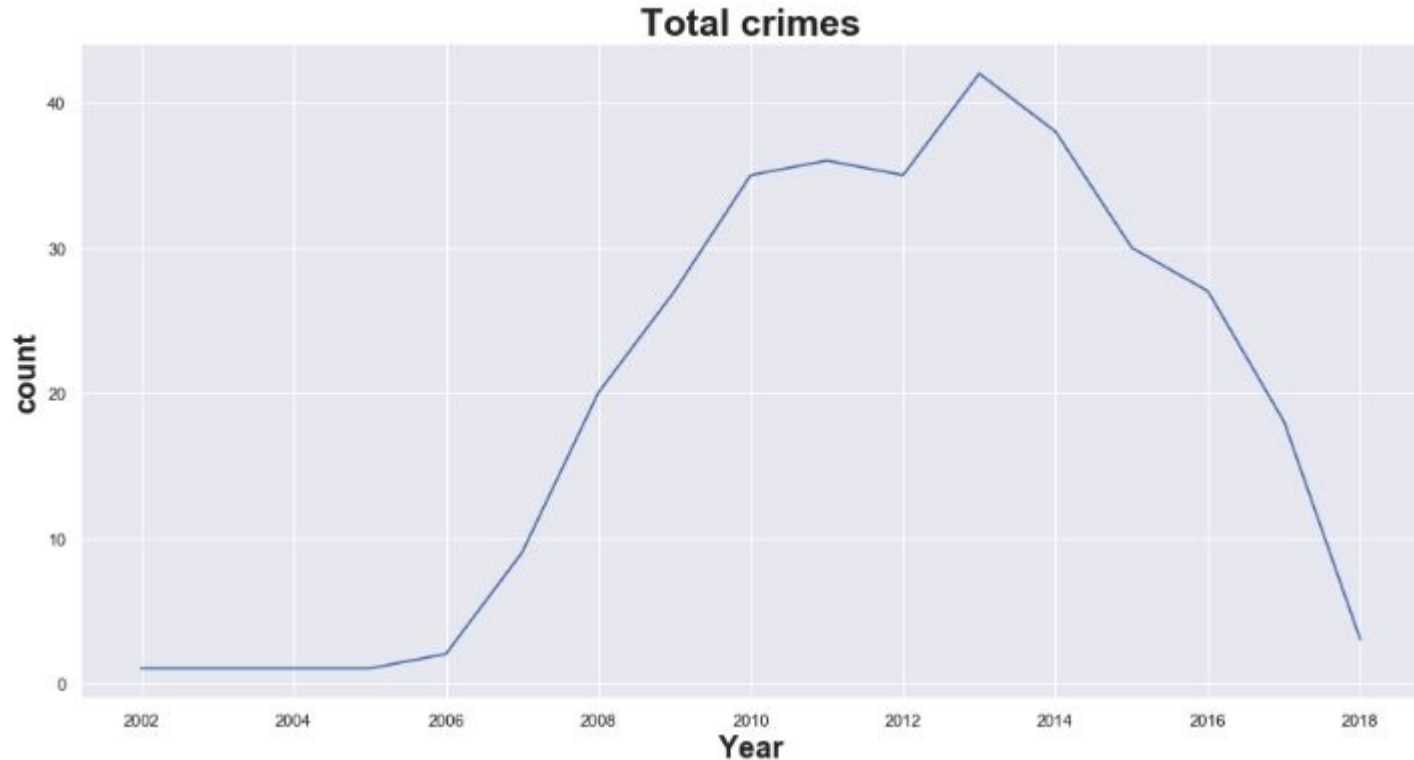
15 Smallest Percent Average Display Publicly Counties

# 03

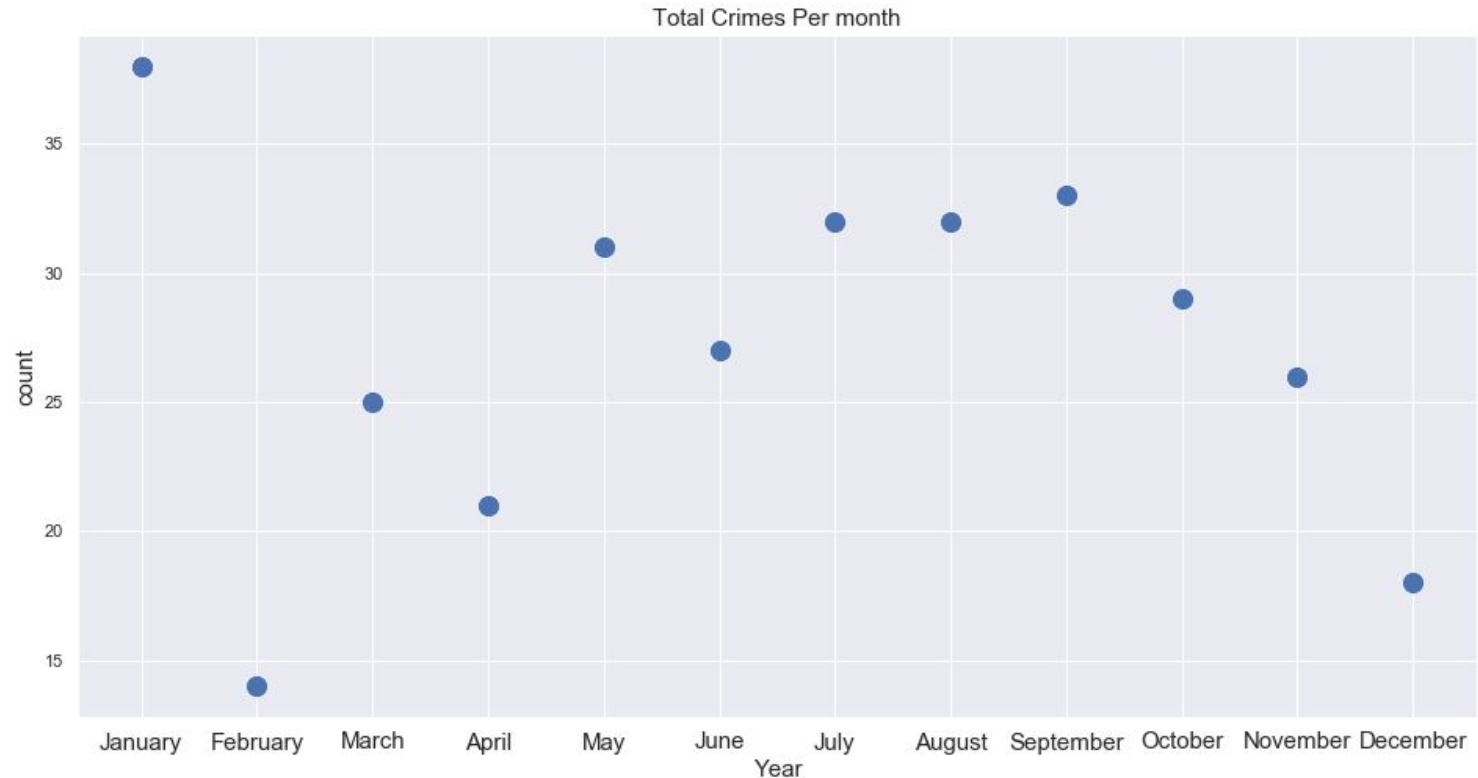
## Time Series

Trends over time

# Total crimes reported per year for convicted officers



# Total crimes reported per month for convicted officers

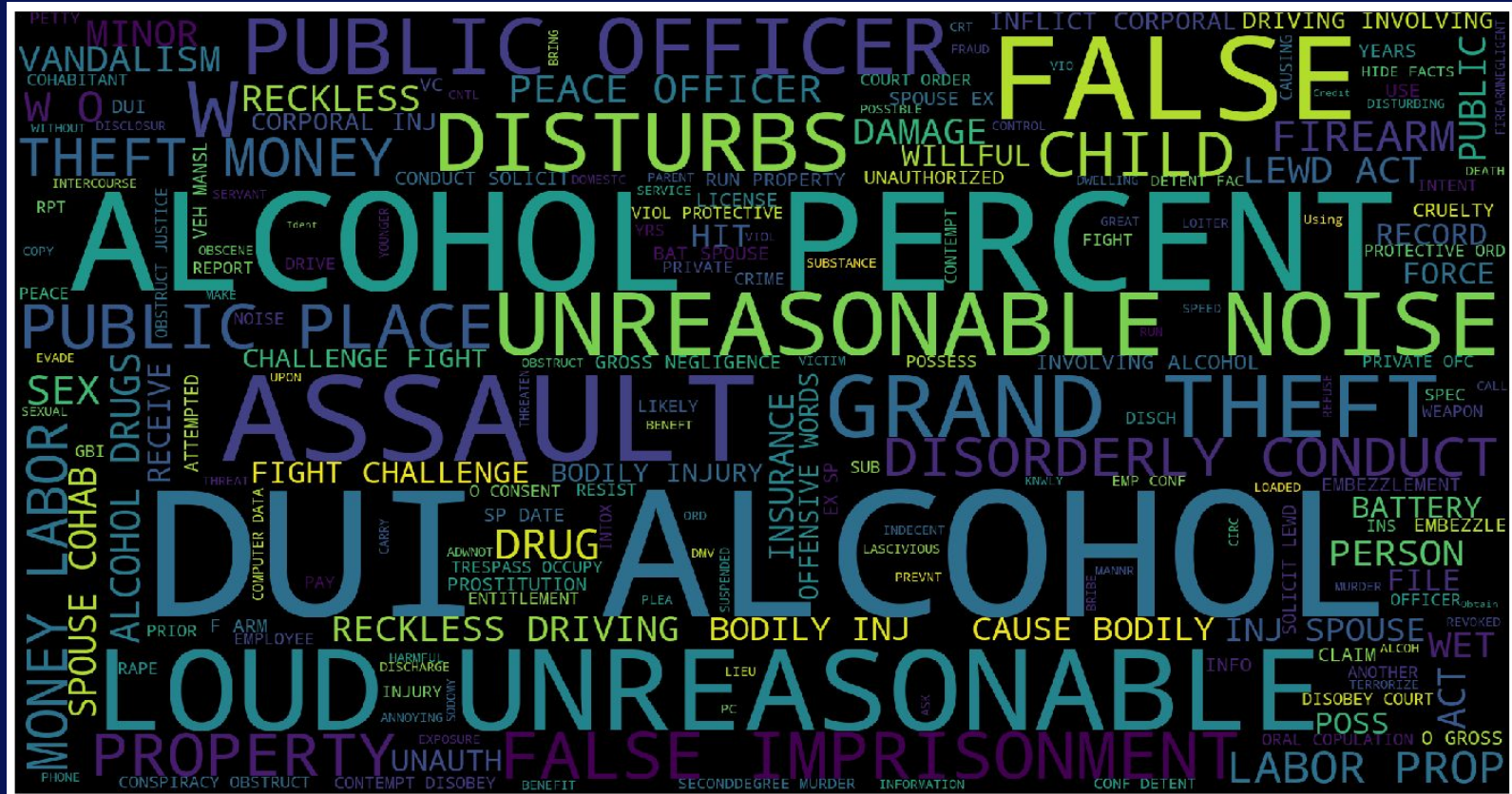


# 04

## Crimes and offenders

Occurrences, consequences, and outcomes

# Frequent Terms Convictions





# Top Charges

DUI

Domestic Violence

Assault (non-domestic)

Weapon laws

Larceny/theft

Sex offenses with minor

Fraud

Sex offenses forcible

Drug/narcotic

Traffic Violations

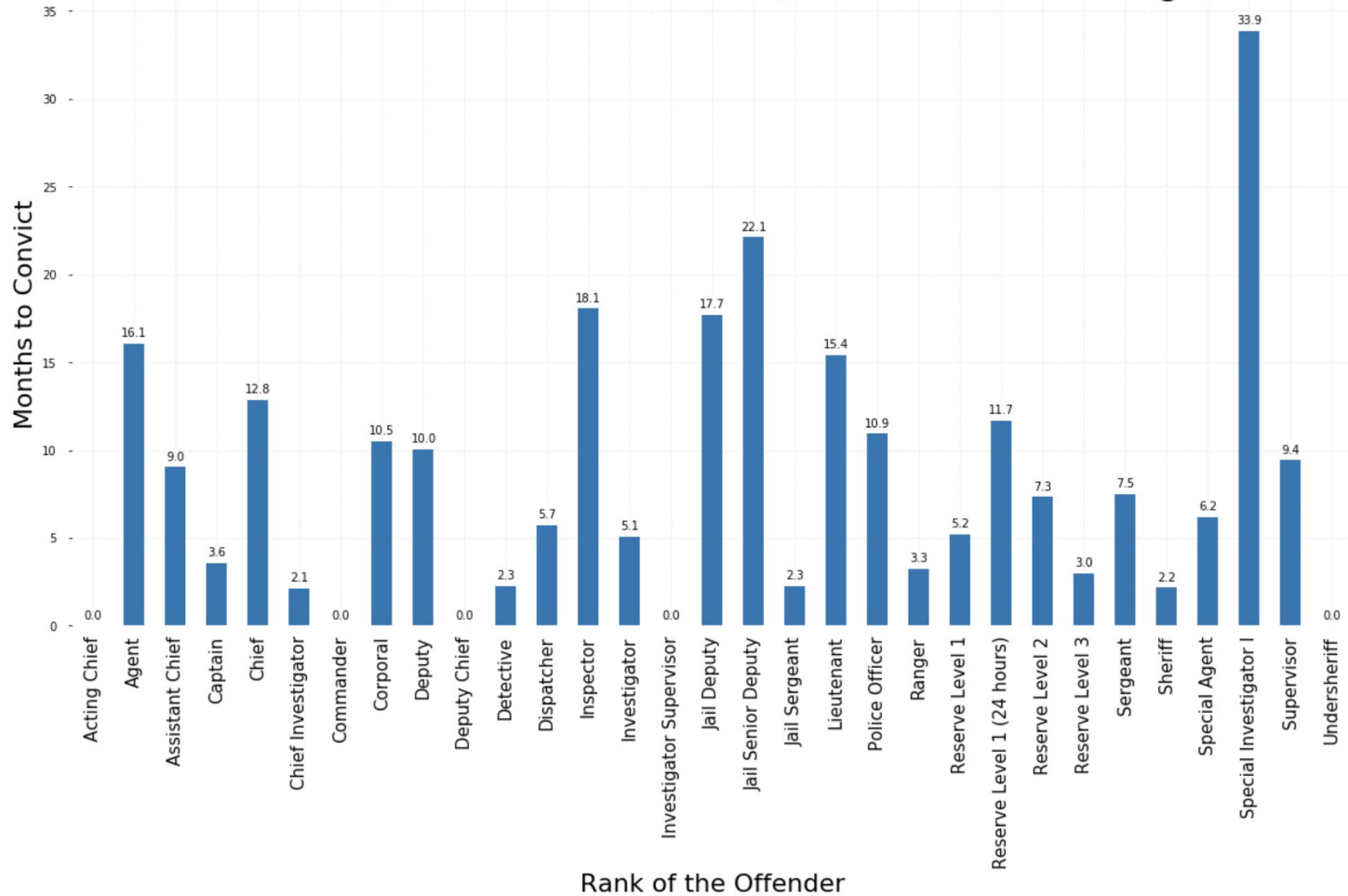
Embezzlement

Child abuse

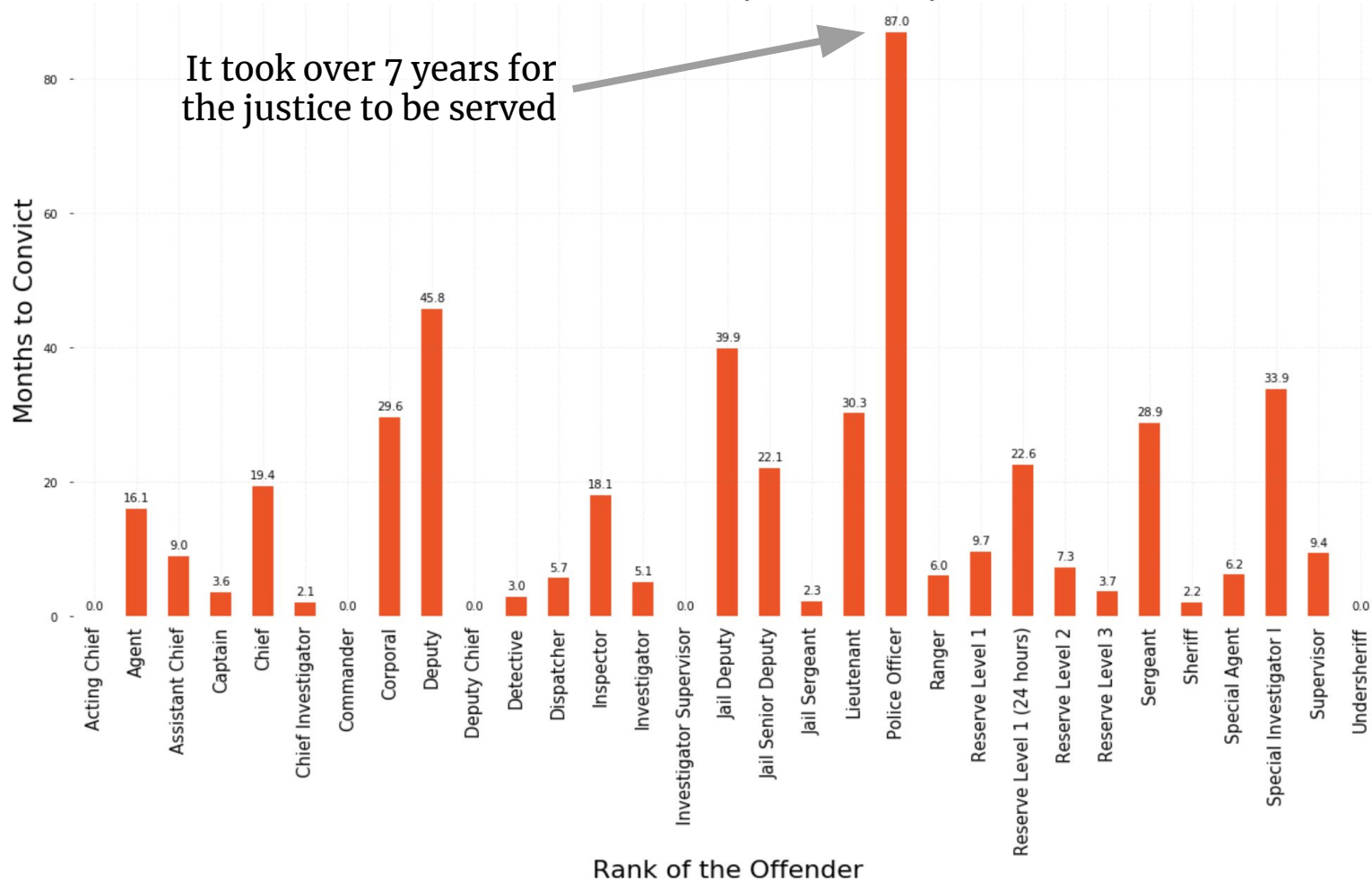
Disorderly conduct

Sex offenses (non-forcible)

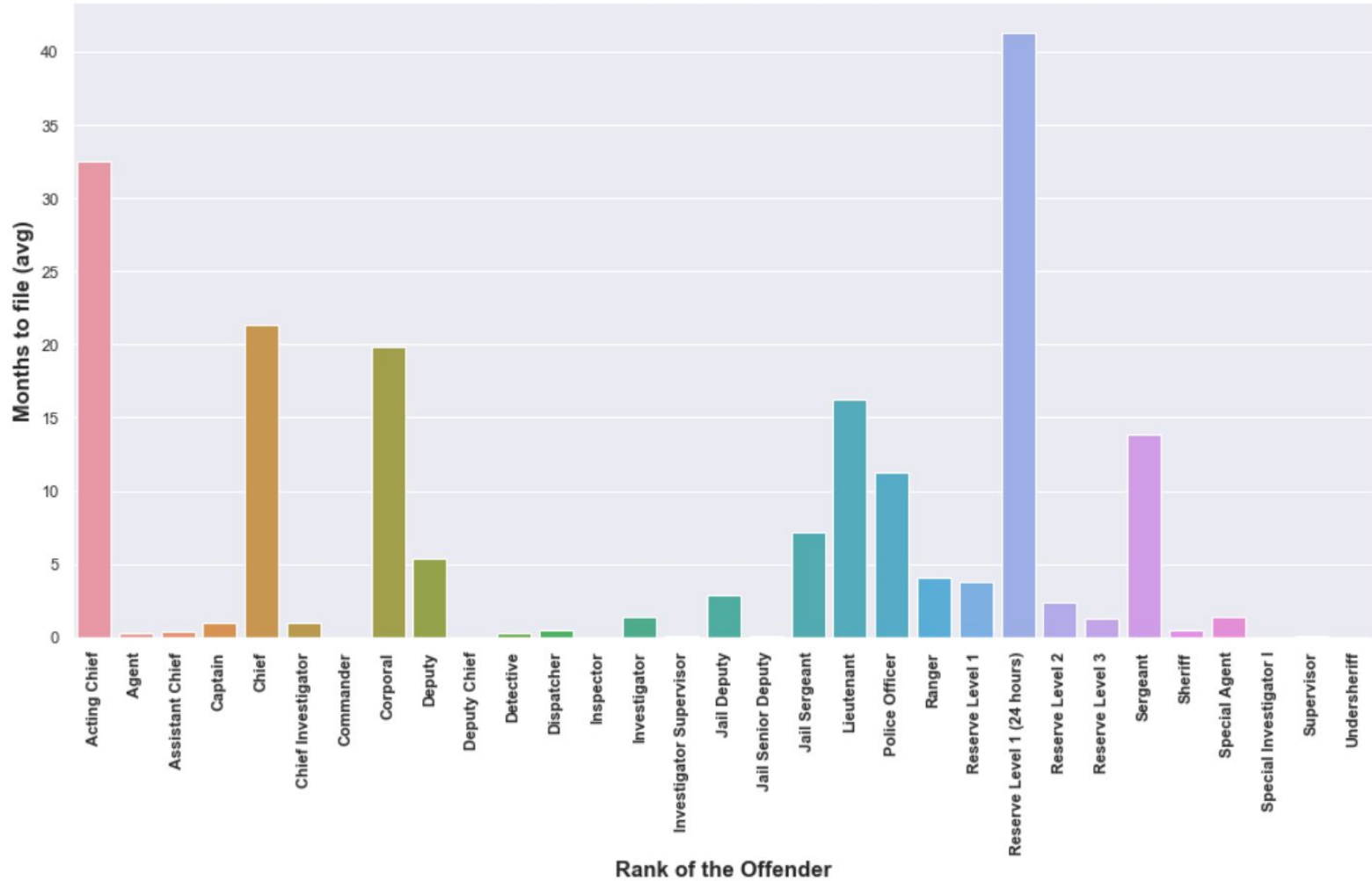
Rank vs. time it took to convict (in months) - on average



## Rank vs. time it took to convict (in months) - in extreme cases

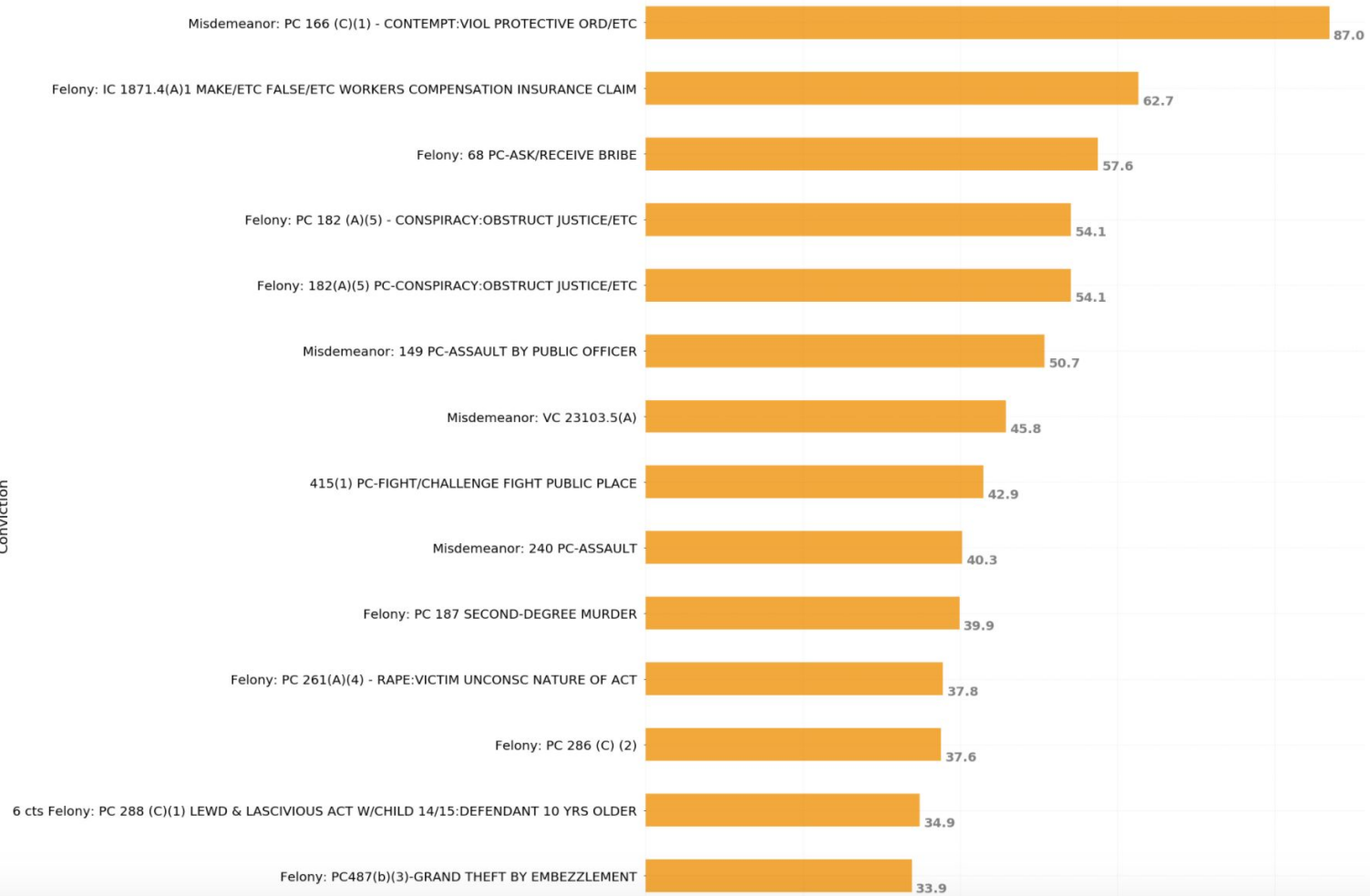


Rank vs time it took to file (months)



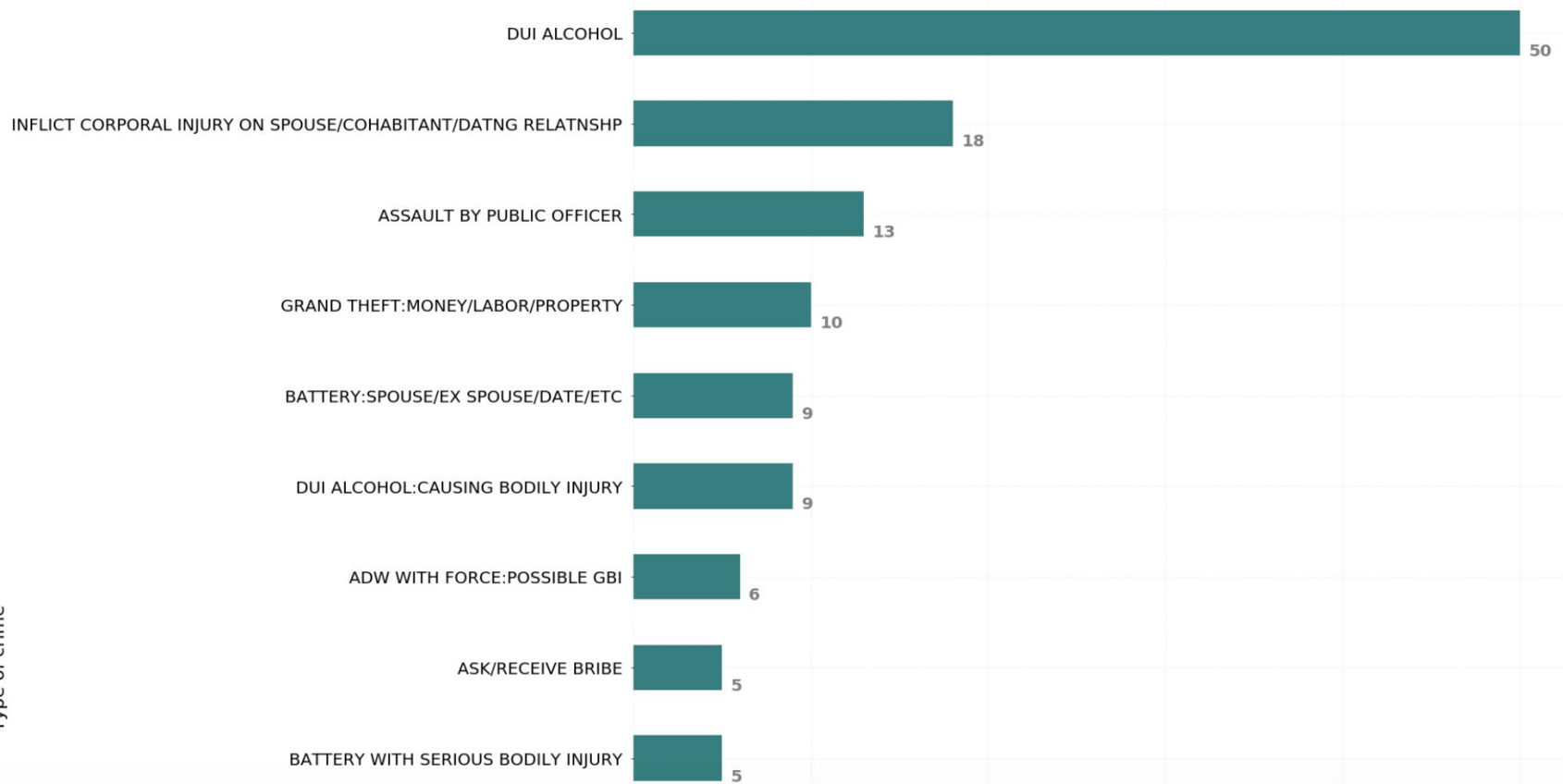
## Crime vs. time it took to convict (in months) - on average

Conviction



## Crime vs. number of times it occurred

Type of crime



\* Shows data for the first-time offenders



# Frequency of convictions as carried out

Outcome	Percent
CONVICTED	2.064
CONVICTED COMMITTED TO PRISON	6.651
CONVICTED SENTENCE NOT REPORTED	1.606
CONVICTED-FINE	3.67
CONVICTED-JAIL	3.44
CONVICTED-PROB/JAIL	51.376
CONVICTED-PROBATION	31.193

# Main Takeaways

## Frequency

Police misconduct  
is NOT an anomaly



DUI

Alcohol-related  
crimes were  
prevalent

## Media Coverage

Civil rights  
movements and their  
impact



Geographical  
Data

Why are there  
disparities in  
counties?

“Injustice anywhere is a  
threat to justice  
everywhere.”

—Martin Luther King

# Meet the Team



Sasha  
Prokhorova

NLP analysis and  
data visualizations



Ben Chu

EDA and data  
visualizations



Nancy  
Henaku

Research



Karla Palos  
Castellanos

EDA & data  
visualizations



[Sasha's LinkedIn](#)

[Ben's LinkedIn](#)

[Nancy LinkedIn](#)

[Karla's LinkedIn](#)



[Sasha's Podcast](#)

[Ben's GitHub](#)

[Nancy's  
Publications](#)

[Karla's GitHub](#)



[Sasha's GitHub](#)

# Thanks!

Do you have any questions?

[prokhorova.sasha8@gmail.com](mailto:prokhorova.sasha8@gmail.com)

<https://medium.com/@prokhorova.sasha8>



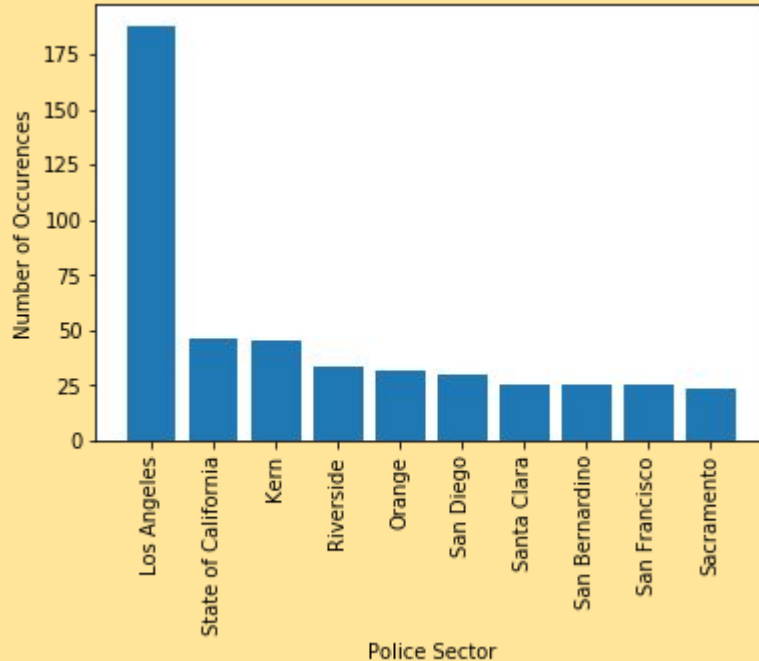
**CREDITS:** This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**. Please keep this slide for attribution.

# Appendix

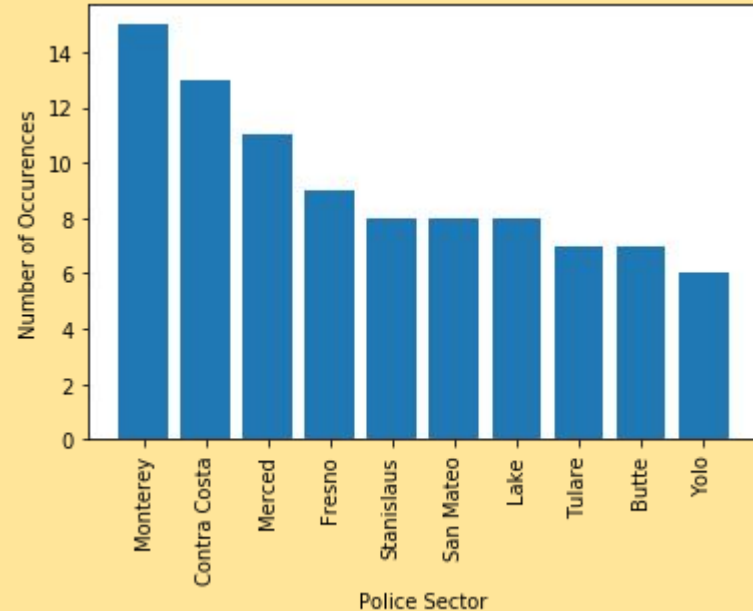


# Top Occurrences of Crimes

## Top 10 Counties by Number of Occurrences



## Top 11-20 Counties by Number of Occurrences



## Average Values for Counties, sorted by months to convict

Code:

```
df.groupby('county').mean().sort  
_values('months_to_convict')
```

	display_publicly	cii_number	months_to_convict
county			
Ventura	1.000000	1.032269e+09	0.065710
Shasta	1.000000	1.016808e+09	1.007550
Colusa	1.000000	1.022235e+09	1.182776
Tehama	1.000000	1.023671e+09	3.022649
San Joaquin	0.833333	8.266672e+08	3.384053
El Dorado	0.750000	1.024364e+09	3.811167
Marin	1.000000	1.020617e+09	5.158217
Imperial	1.000000	1.017602e+09	5.486766
Placer	1.000000	1.017551e+09	5.733177
Tulare	1.000000	1.021059e+09	5.881024
Stanislaus	1.000000	1.022624e+09	5.946734
Napa	1.000000	1.025040e+09	6.275283
San Luis Obispo	1.000000	1.020078e+09	6.850243
Yolo	1.000000	1.021962e+09	7.602620
Sonoma	1.000000	1.020150e+09	7.858888
Lake	1.000000	1.022703e+09	8.331060
San Diego	0.966667	1.137916e+09	9.283330
Sacramento	0.913043	1.192445e+09	9.304503
State of California	0.956522	9.882087e+08	9.411829
Solano	1.000000	1.015776e+09	9.741473