PyCrimes and Misdemeanors

UC Berkeley Data Science Bootcamp
- Steve Bonillas
- Laura Hickman
- Kento Nakajima
- Michelle Verger
- Christian Corona

Agenda

01. Executive Summary

- Goals of the project
- The questions we were looking to answer
- Preview of our initial findings

02. Tactics

- Deciding on a Topic
- Data Sources
- How the sources were used

03. Challenges

- What gave us trouble?
- How were we heroes?

04. Findings

- Factoids and Trivia (Fun!)
- Charts!
- Maps!
- Numbers!

05. Next Steps

- Blind Spots
- Additional Research Avenues
- Job Security

Executive Summary

Data Sources

- FBI Crime Data Explorer API
  - Property crime arrests by state
  - Number of law enforcement employees
  - Number of agencies reporting data
  - https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/docApi
- US Census Bureau
  - 2020 and 2010 Census results used for per capita calculations
  - https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html

Questions

- Between 2010 and 2020, what was the overall trend when it came to property crimes?
- Is there a correlation between population and arrest rates?
  - If those rates are adjusted per capita, does the picture change?
- Is there a correlation between the number of law enforcement employees and arrest rates?
- Historiography: Can we find any changes in the way the data was reported that might affect any trends that are seen?

Initial Findings

- Spoiler Alert
- Property crime steadily declined between 2010 and 2020
- Found correlation between population and raw arrest numbers.
  - When normalized for population, the arrest rates even out.
- Found correlation between number of police and number of arrests to a point, but after a certain level, diminishing returns.
- Reporting agencies and law enforcement numbers remained stable during the subject period.

Tactics

01. Setting our Goals

- Chose a project based on a team member's domain.
- Remain objective.
- Let the data tell the story.
- Work towards a thesis, not from one.

02. Divide the Labor

- Team members pursued areas that interested them. APIs, visualization, analysis, etc.
- Used regular communication to fill gaps as needed.

03. Find the Data

- Arrest Data was procured from the from the FBI Crime Data Explorer using APIs.
- Population data was pulled from the Census bureau and imported to the project via CSV.

04. Clean the Data

- API data was imported to Panda data frames and reviewed.
- Identified and corrected gaps.
- Clean data frames exported to CSVs to shorten later workflow and reduce API calls.

05. Analyze and Polish

- Comparison and Extrapolation
- Found common denominators in population and law enforcement rosters.
- Adjusting values per capita provided more context.
- Visualizations were generated in Python, with line, bar, and scatter graphs. Hvplot maps were utilized as well.

Challenges

Things we heroically overcame

- Michelle's broken arm.
- Pandemic data was unreliable.
  - Example: Florida went from reporting tens of thousands of property crime in 2020 to under 200 in 2021
  - So, we switched our focus.
- HvPlot had some formatting issues, which didn't allow for suppressing scientific notation without sacrificing other, more important formatting settings.

Blind Spot

- Laws and definitions vary from state to state. This could leave some of our data open to interpretation.

Some Quick Factoids

For the years 2010 to 2020, here are the states with the most arrests for...

Arson
(Setting Fires)

California

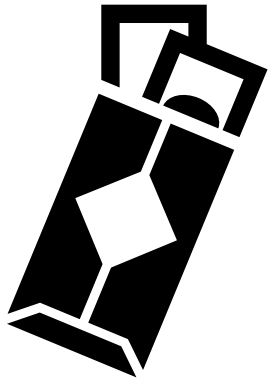
2,051 Arrests

Larceny
(Stealing Property)

Texas

11,5252 Arrests

Embezzlement
(Robbing Companies and Charities)



Virginia

1,676 Arrests

Fraud
(Stealing with Words)

North Carolina

18,242 Arrests

Some More Trivia

For the years 2010 to 2020, what states had the…

Highest Rate of Arrest (Per Capita)

Delaware

Averaged 1031 arrests per capita.

Delaware also had the single year highest arrest per capita with 1214 in 2012.

Tennessee was second with 832.

Lowest Rate of Arrest (Per Capita)

Illinois

Averaged 125 arrests per capita.

Washington DC was in our data with with an average of 30 per capita, but they aren't a state yet.

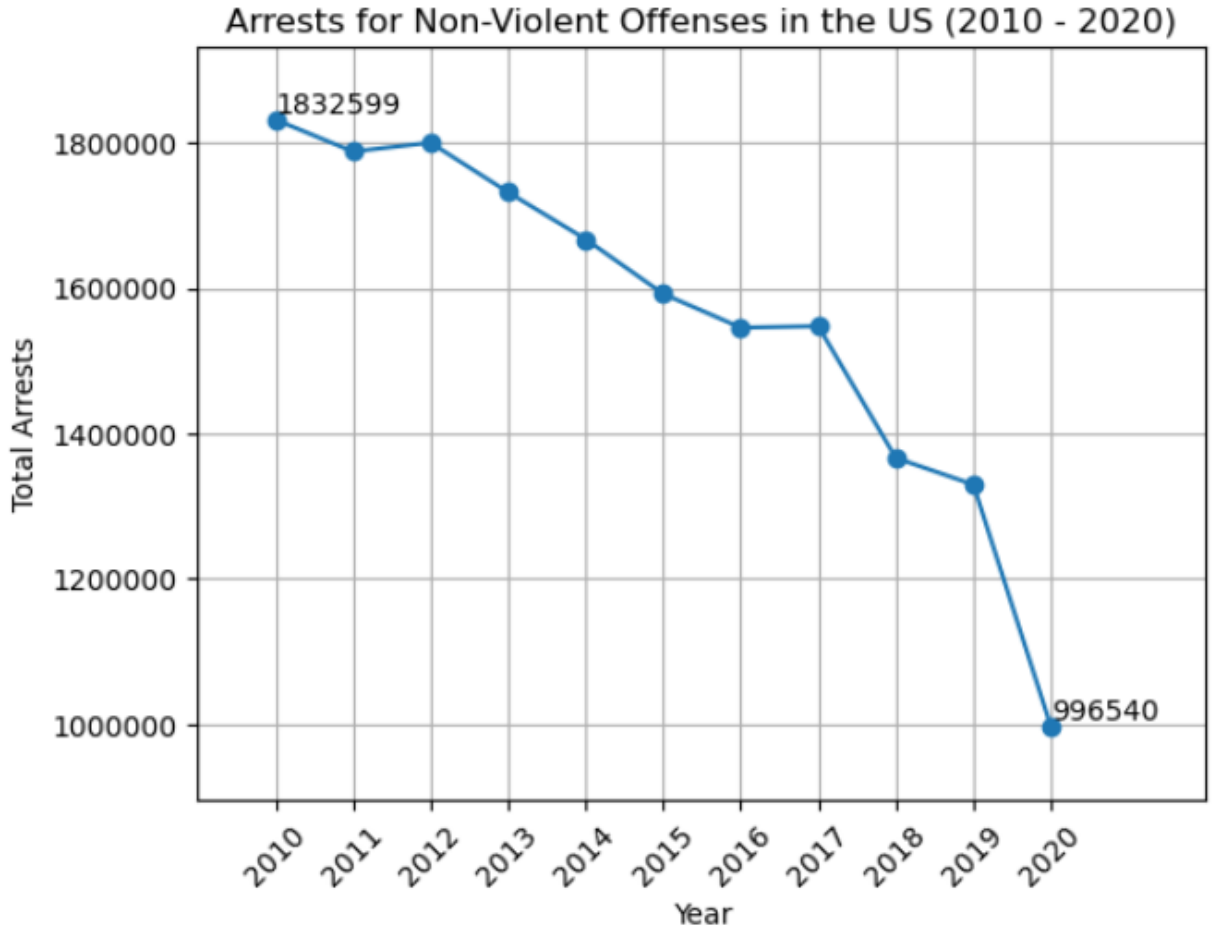
Single lowest year was Pennsylvania with 3.5 arrests per capita in 2020.

It is hard to find an icon that conveys "Being let go with a warning."

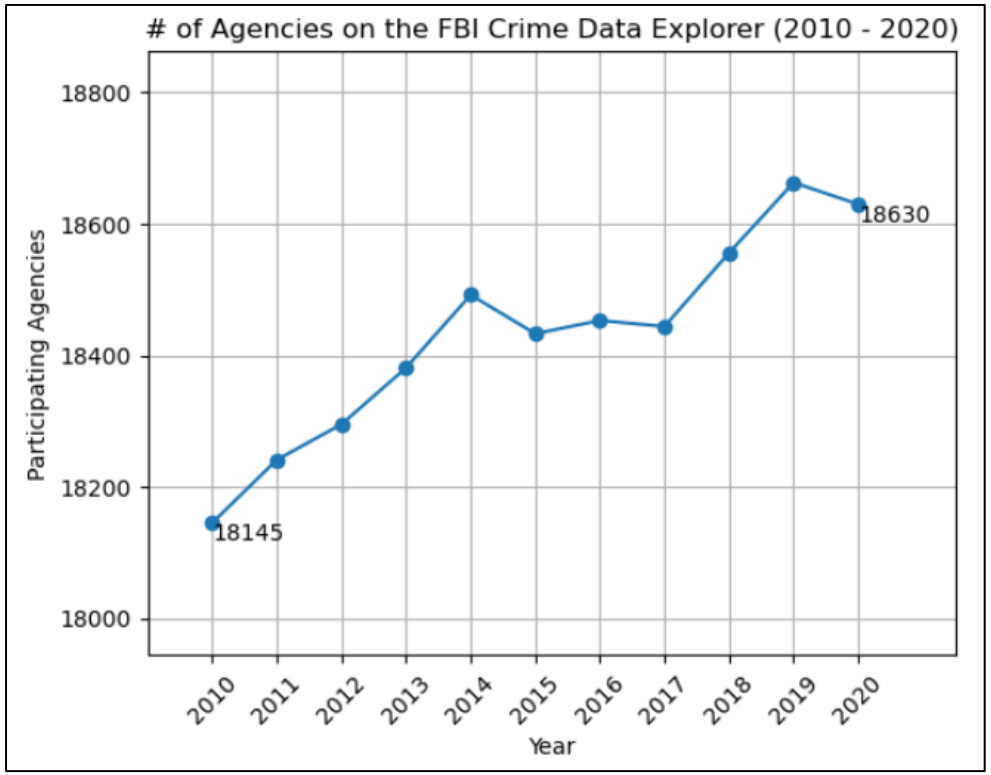
Total Arrests in the US (2010 – 2020)

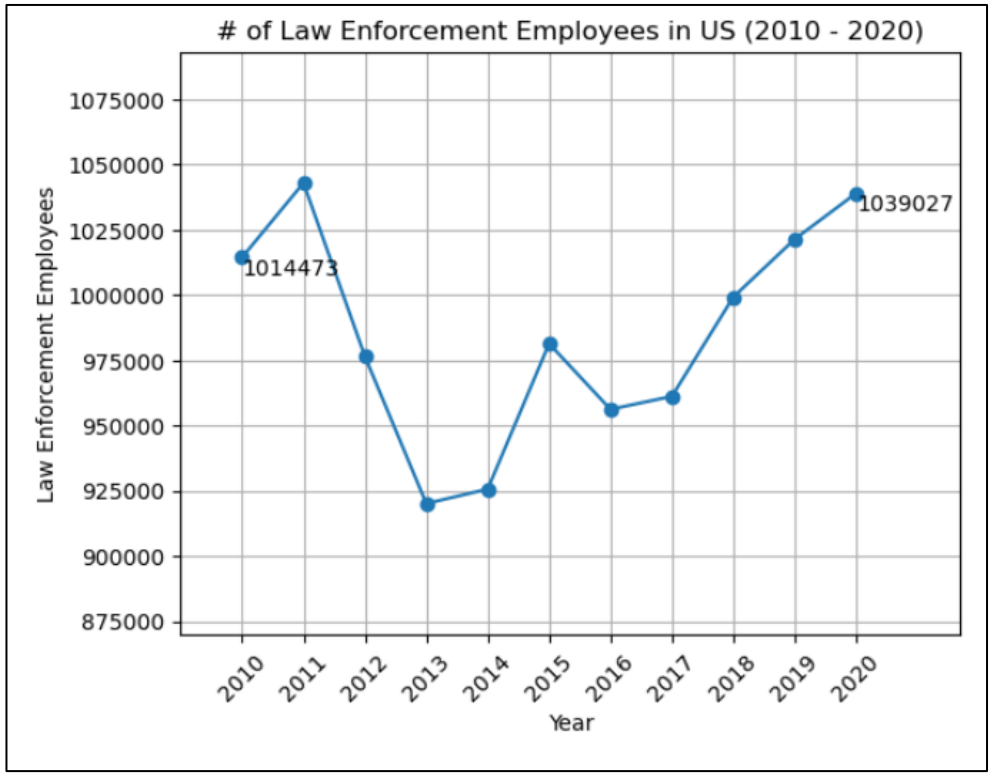
- One of our first findings was a significant and steady downward trend in reported property crime arrests.

- The numbers shown here are an aggregate of arrests across all states, for all offenses classified by the FBI as ***Property Crimes***, which are:
  - Arson
  - Larceny – Theft
  - Vehicle Theft
  - Burglary
  - Embezzlement
  - Forgery & Counterfeiting
  - Fraud
  - Vandalism
  - Dealing in Stolen Property



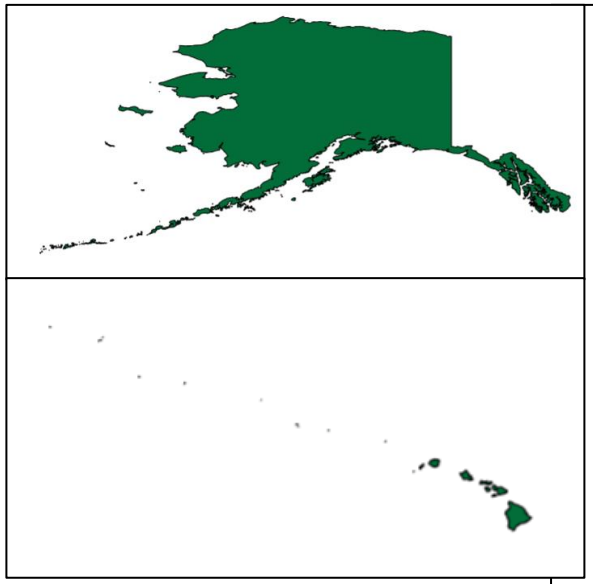
Double Checking Our Findings

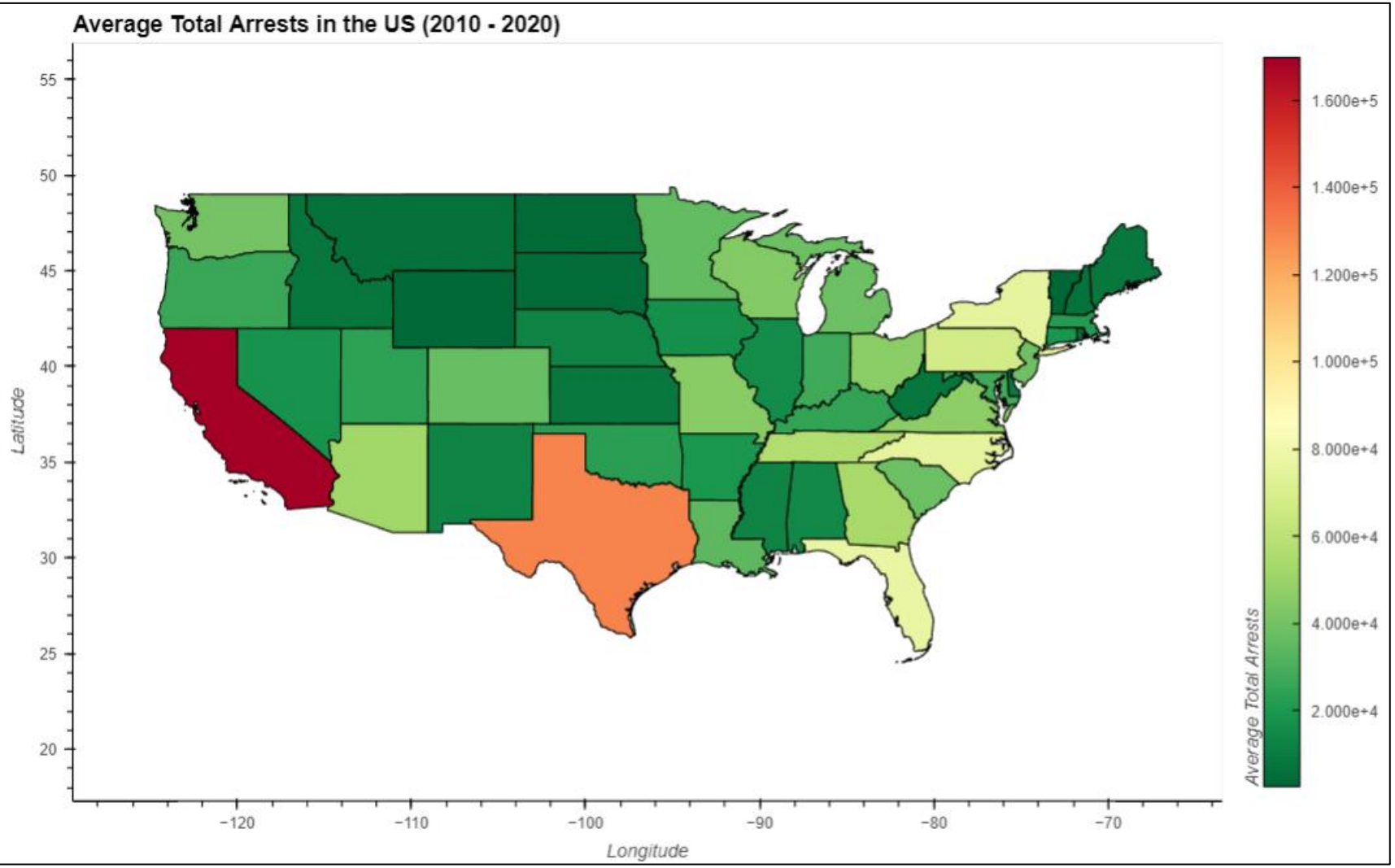




- These charts show that the number of agencies reporting to the Crime Data Explorer and the number of law enforcement personnel remained relatively stable throughout the subject period.
- This provides added context to our overall arrest data, which shows a steady downward trend. If these charts showed drops in reporting agencies or in law enforcement officers, those could be possible outside factors that influenced the drop in reported arrests, but that does not appear to be the case.

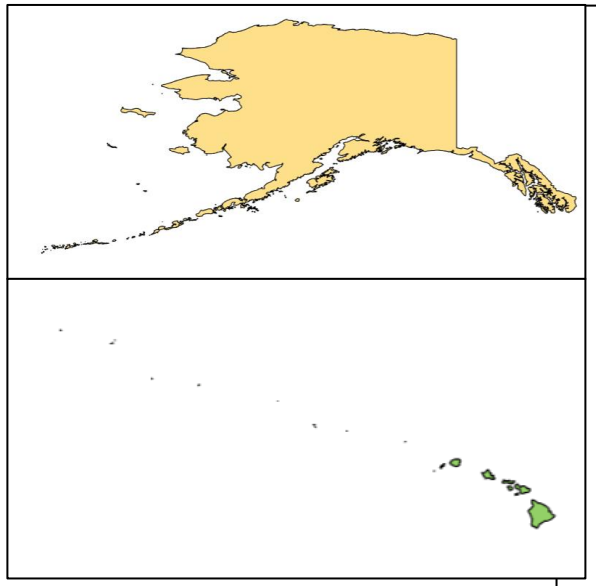
Average Total Arrests in the US

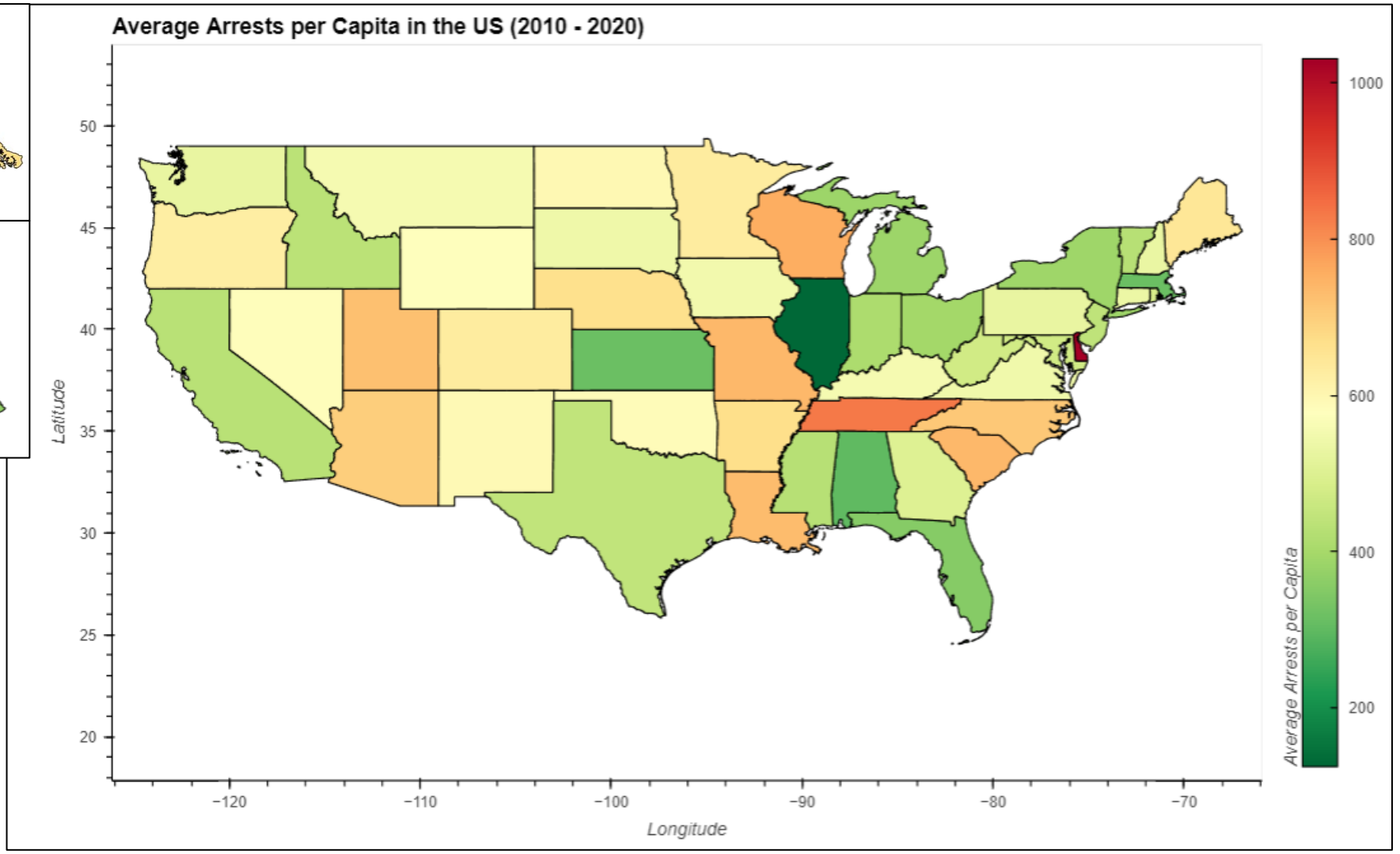




- This is our heatmap of total arrests in the US between 2010 and 2020.
- We took the average of each state's arrests and used hvplot to give a color gauge to our map to illustrate the areas with the greatest concentration of arrets.

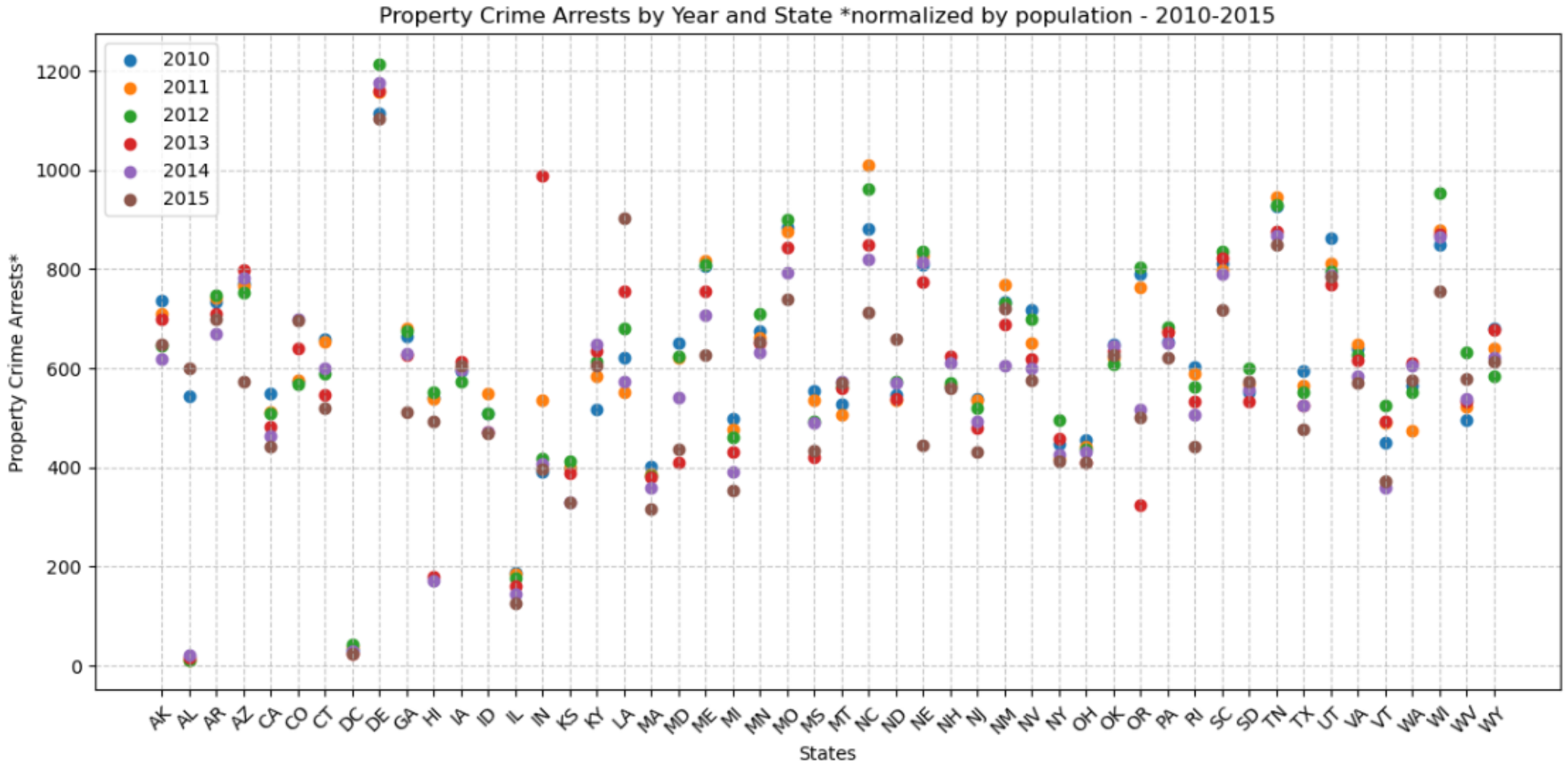
Average Arrests Per Capita



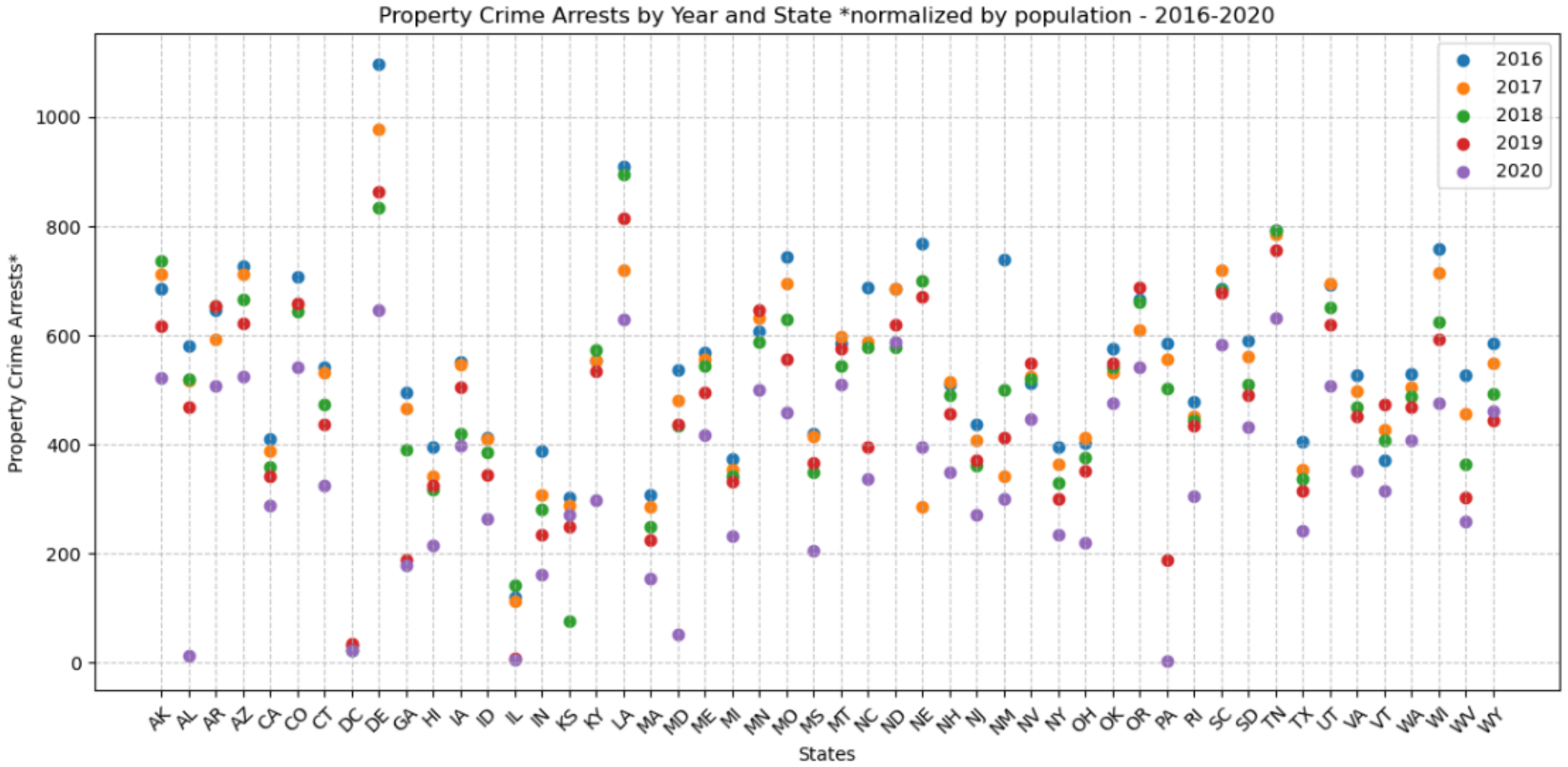


- This map tells a different story.
- We adjusted the arrest numbers per capita, to visualize the Arrest Rate.
- What we see is arrest rates are more normalized and that the hot spots are no longer just the population centers.

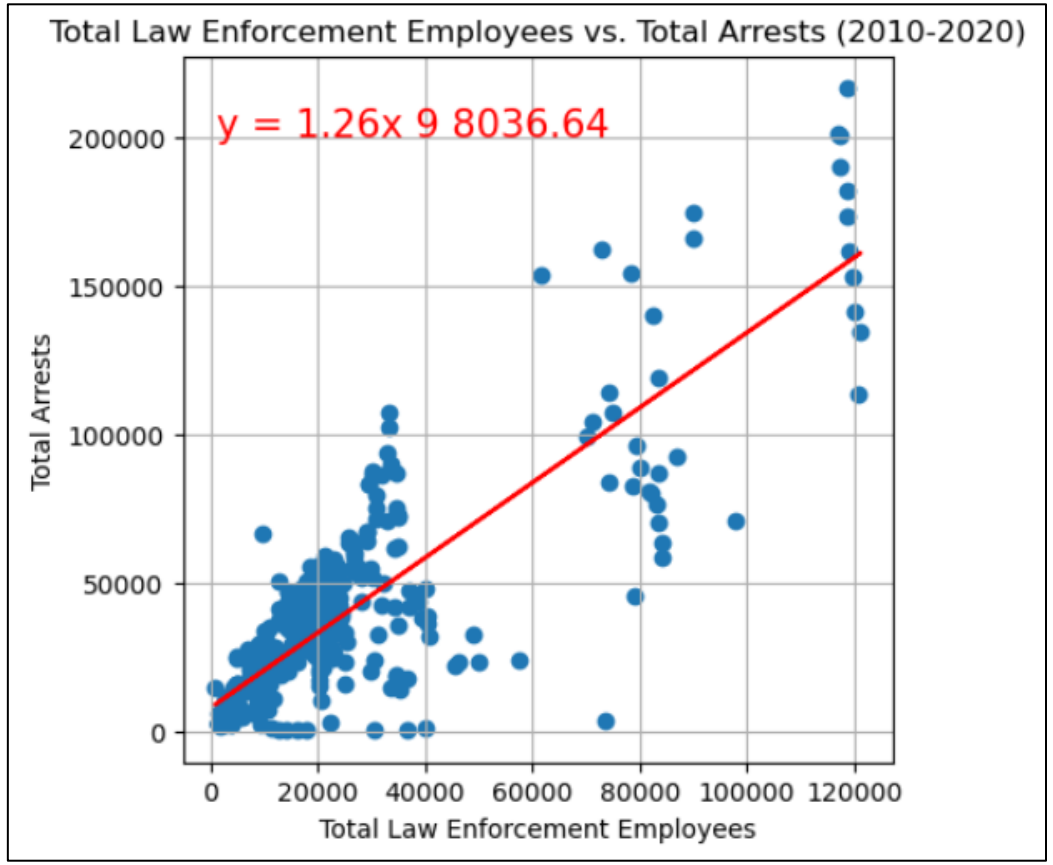
Arrests by Year & State (2010 - 2015)

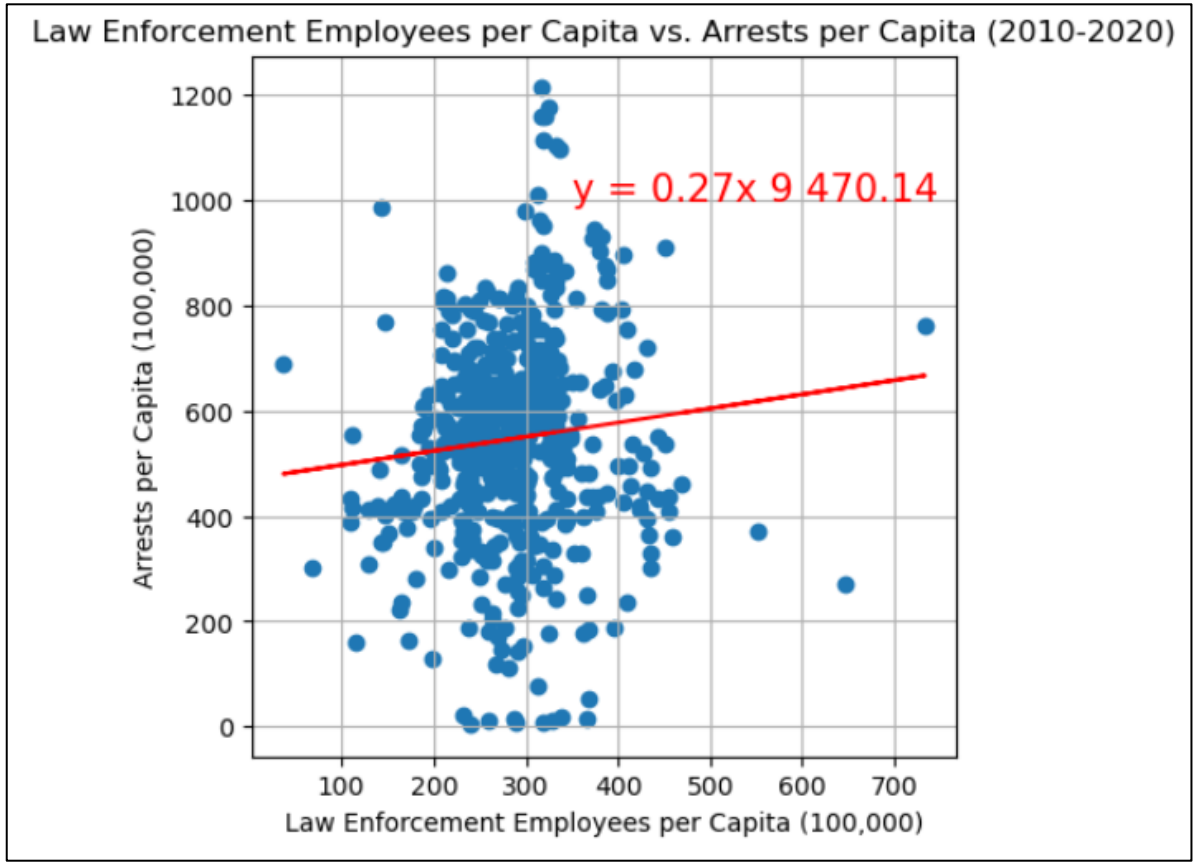


Arrests by Year & State (2016 - 2020)



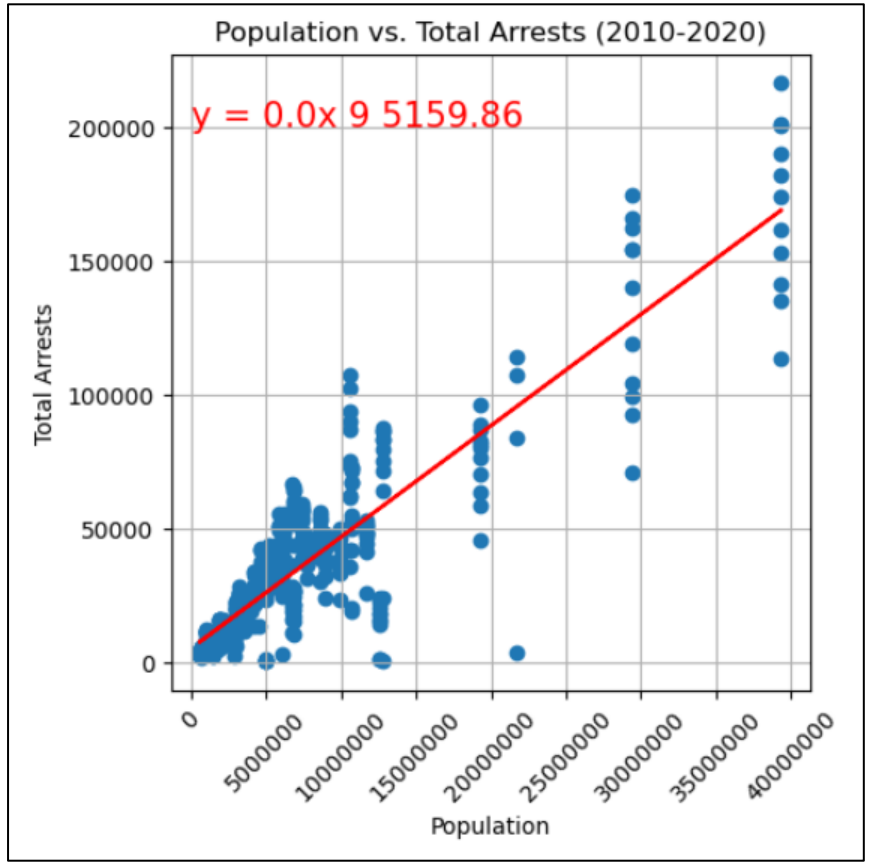
Law Enforcement and Arrests

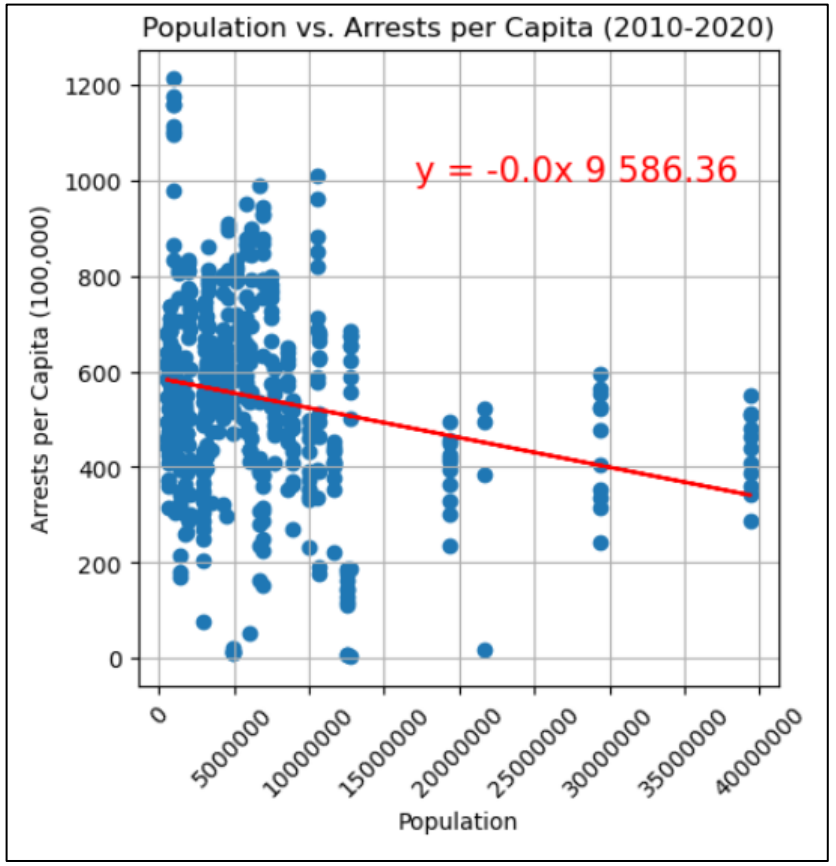




- The chart on the left illustrates the relationship between total number of law enforcement officers and total arrests in each state, each year. What it shows is that as the number of law enforcement officials grows, so do overall arrests. Since they are both linked to population, that make sense.
- On the right, we have both values adjusted per capita. What this shows us is that there is a region of efficiency between 200 and 400 law enforcement employees per capita where arrest rates are clustered. This likely gives us an idea of each state's budgeting considerations.

Population and Arrests

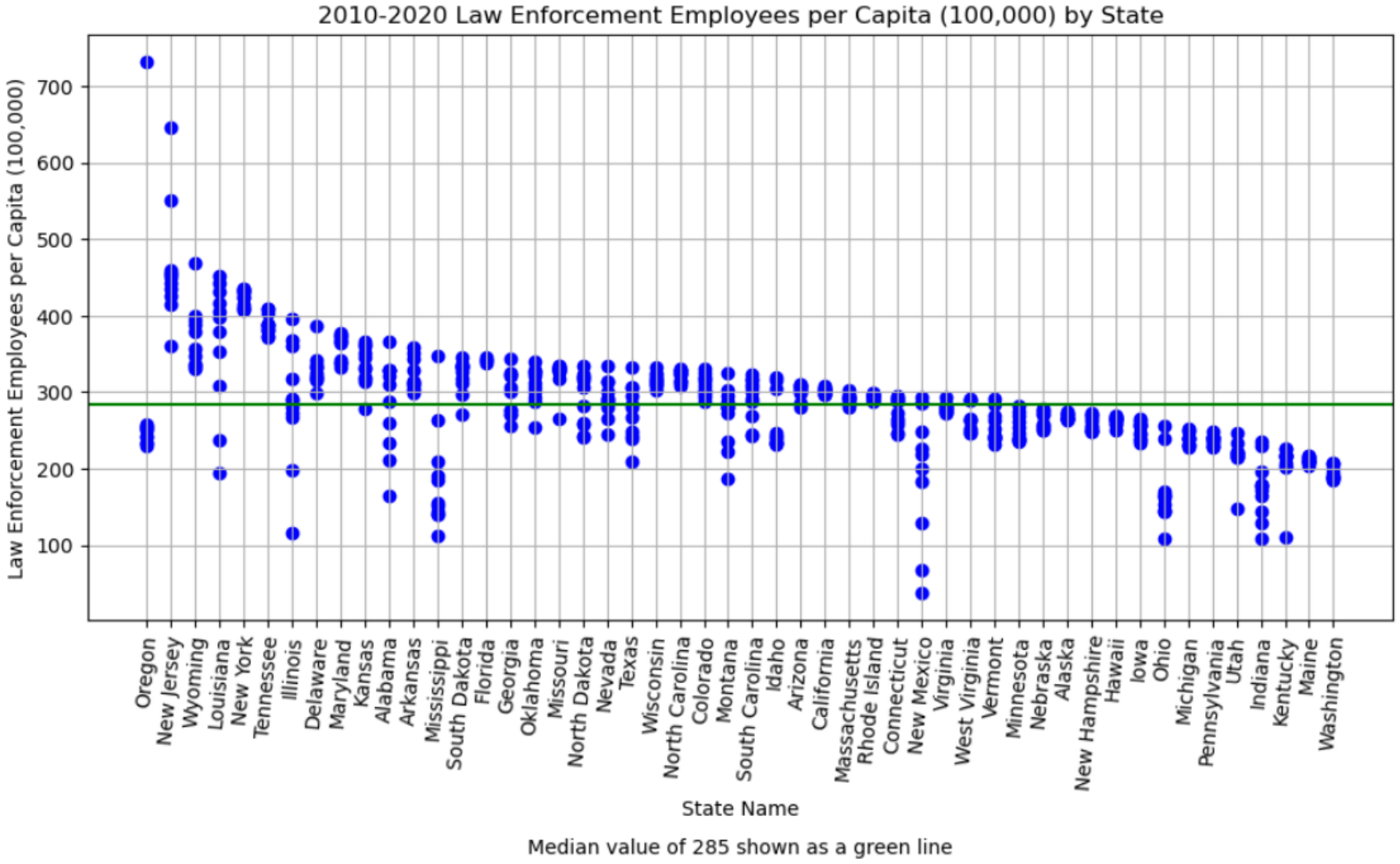




- The chart on the left shows population vs total arrests. The trendline shows a strong relation between population and overall arrest numbers.
- On the right, we see what happens when we adjust the population per capita. The trendline shows us that arrest rates shift downward slightly as population grows. This suggests an **Arrest Rate** which remains relatively steady, regardless of total population.

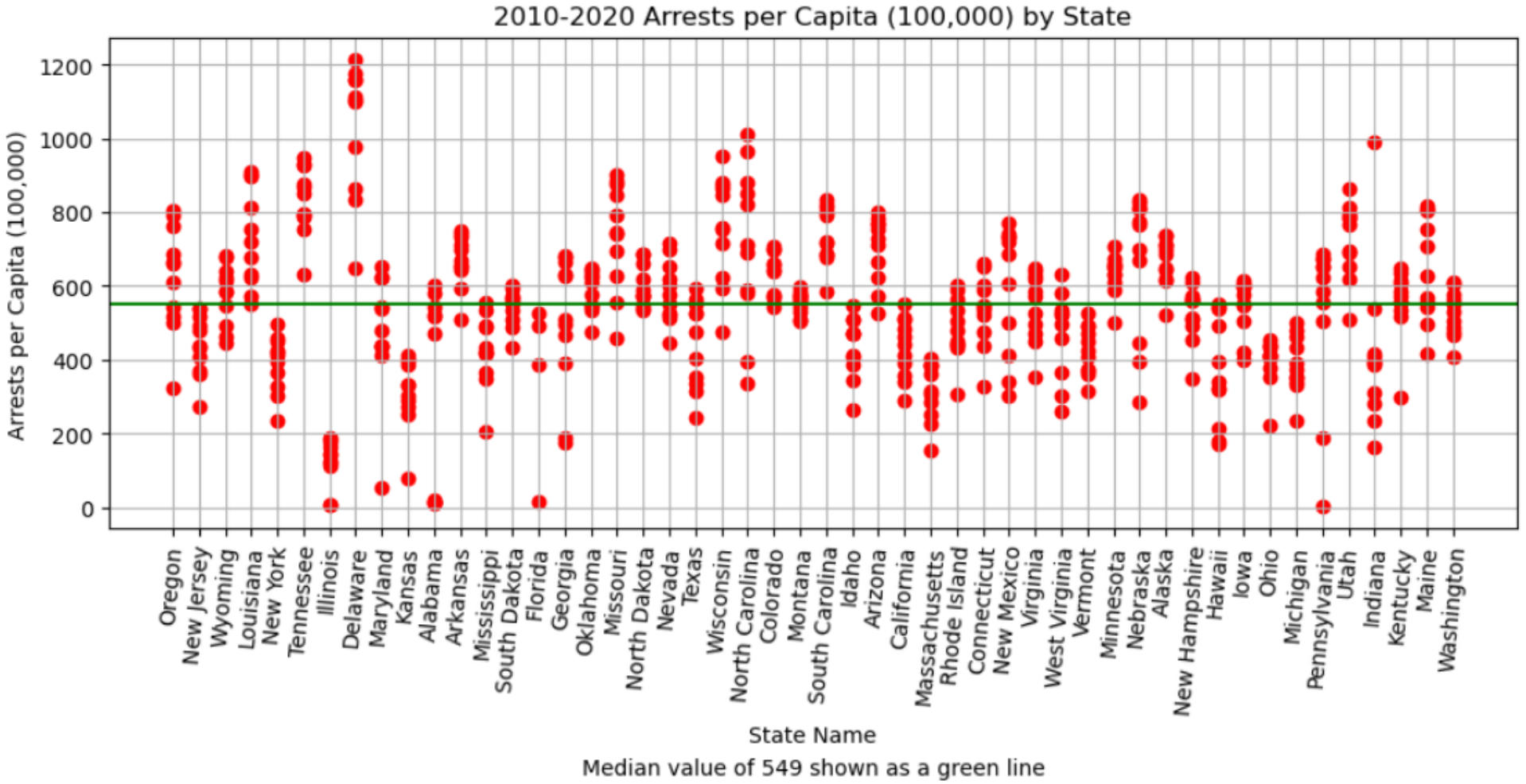
Police per Capita by State

Each dot on the vertical axis represents a year's worth of data for that specific state.



Arrests per Capita by State

Each dot on the vertical axis represents a year's worth of data for that specific state.



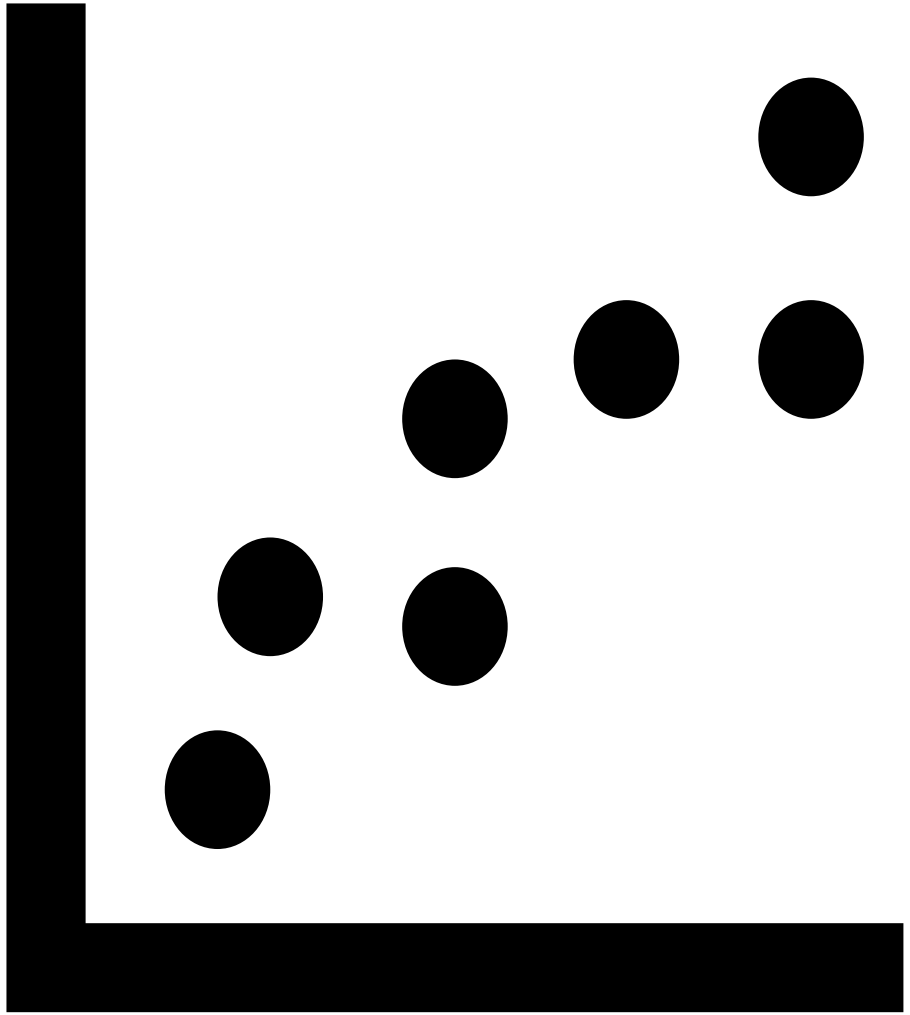
Next steps

Additional data sources to consider

Violent Crime

Education and Economic Data

Conviction and Incarceration Rates



Thank you

Steve Bonillas

Kento Nakajima

Michelle Verger

Laura Hickman

Christian Corona

PyCrimes and Misdemeanors