

# Classifying Changes in County Incarceration Rates

---

CLASSIFICATION MODELING OF INCREASES  
OR DECREASES TO COUNTY-LEVEL JAIL  
INCARCERATION RATES

# Problem Statement

Using county level data on population dynamics, income and employment, and previous years' incarceration data, can we **predict if the jail incarceration rate will increase or decrease/stay the same?**



Help local stakeholders identify and address factors that contribute to increased jail rates.



Anticipating needs for over-capacity jails at county level.



Identify areas to target for decreasing incarceration trends.

# Policy Value Add

# Methodology & Data Sources

Incarceration Data from Vera Institute

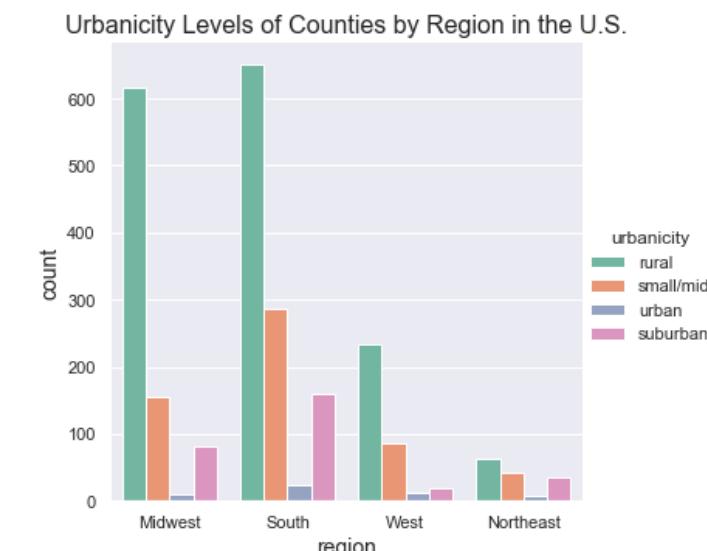
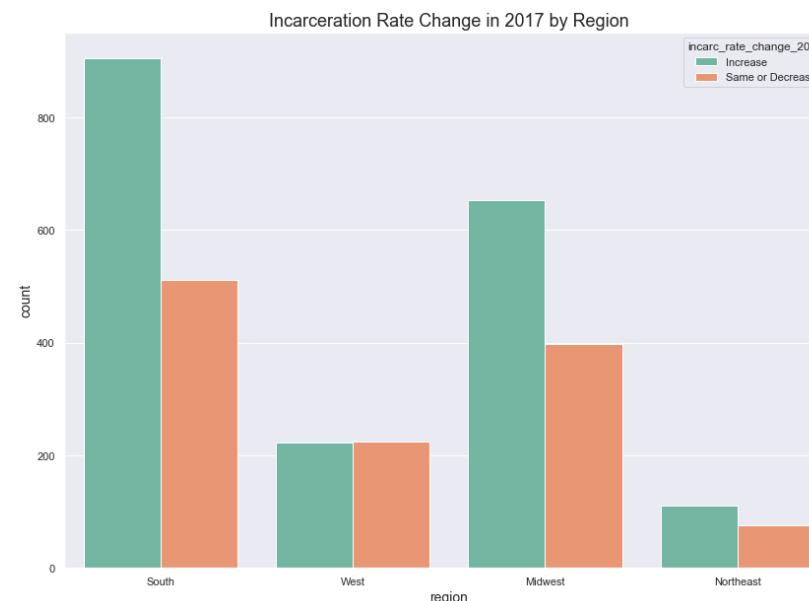
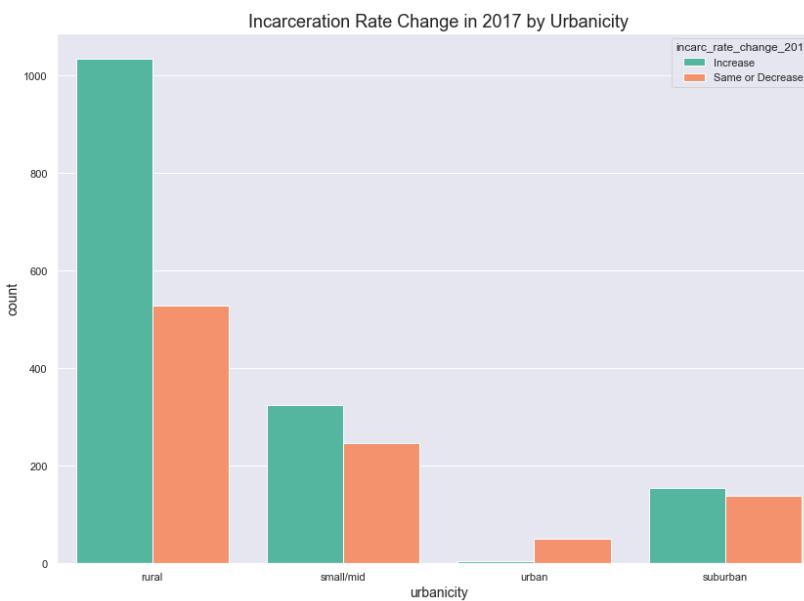
County level data on  
population/demographic dynamics,  
income, jobs, veterans from USDA ERS

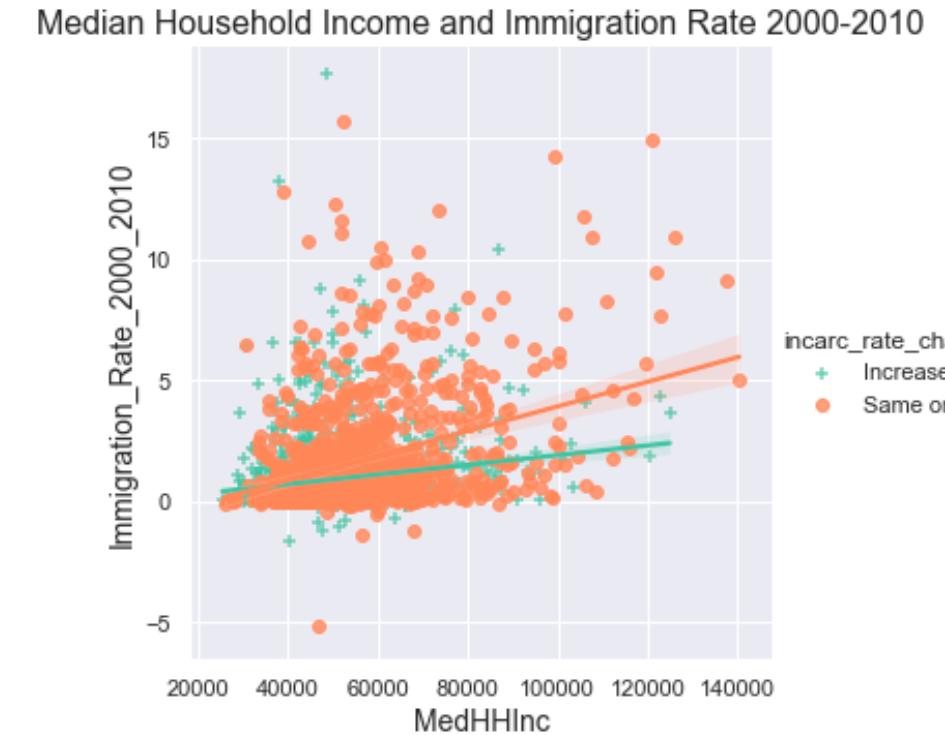
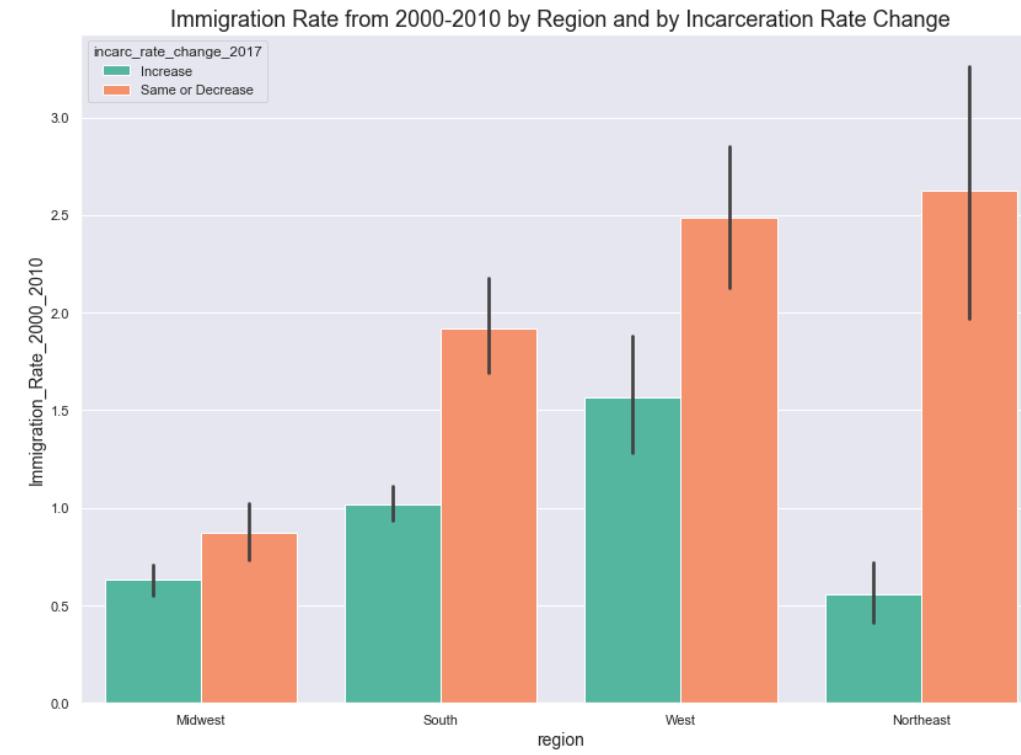
OSEMIn Framework

Binary Classification Modeling

# Incarceration Rates per Urbanicity & Region

HIGHER SHARE OF COUNTIES WITH INCREASED INCARCERATION RATES ARE RURAL, IN SOUTH AND MIDWEST REGIONS



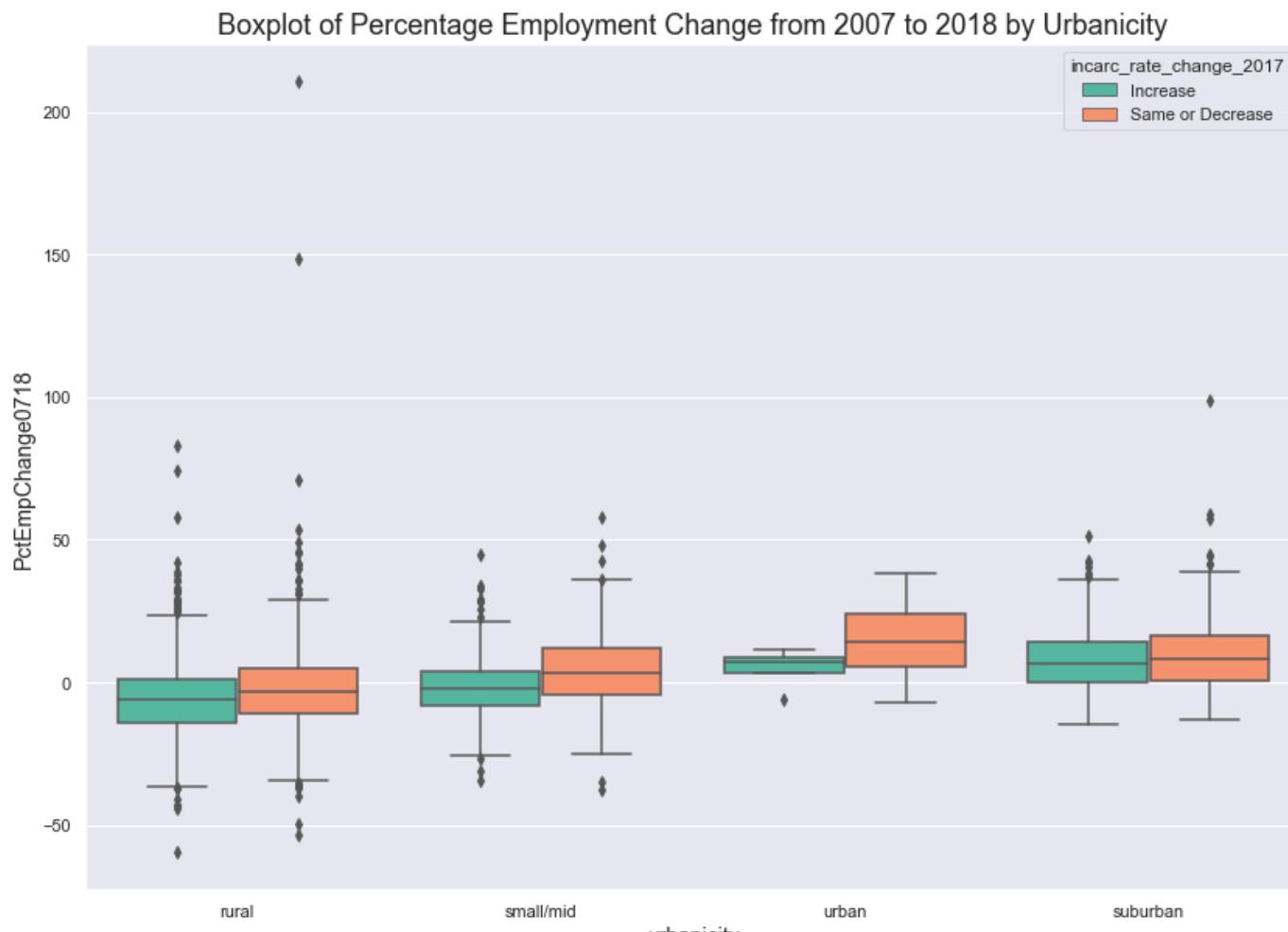


# Immigration Rates

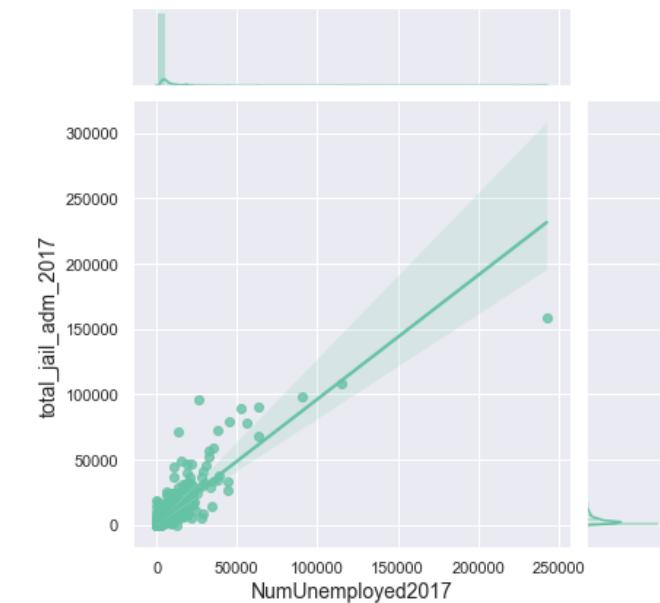
---

# Employment Rate Changes

---



Employment rate percentage changes are slightly lower for counties with increased incarceration rates.



# Findings

Using incarceration-related data + other county-level data:

- Best Guess: 60/40 for Increased versus Decreased/Steady rate
- Baseline Model Score: 78%
- Improved Model Score: 87%

Using only other county-level data (no incarceration-related data included):

- Baseline Model Score: 55%
- Improved Model Score: 76%

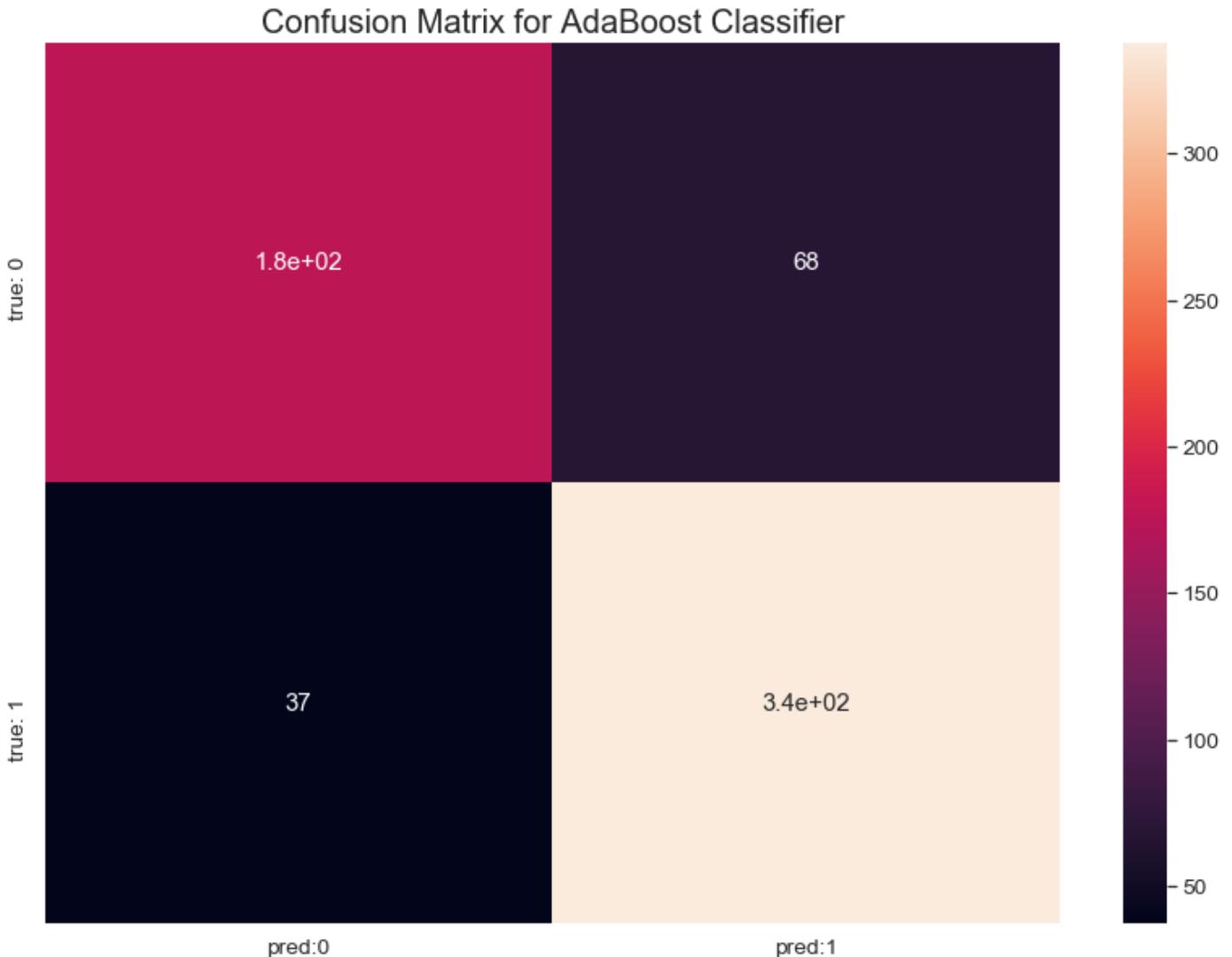
Difficulty with accurately classifying counties with decreased/steady rates

# Findings – with incarceration data

---

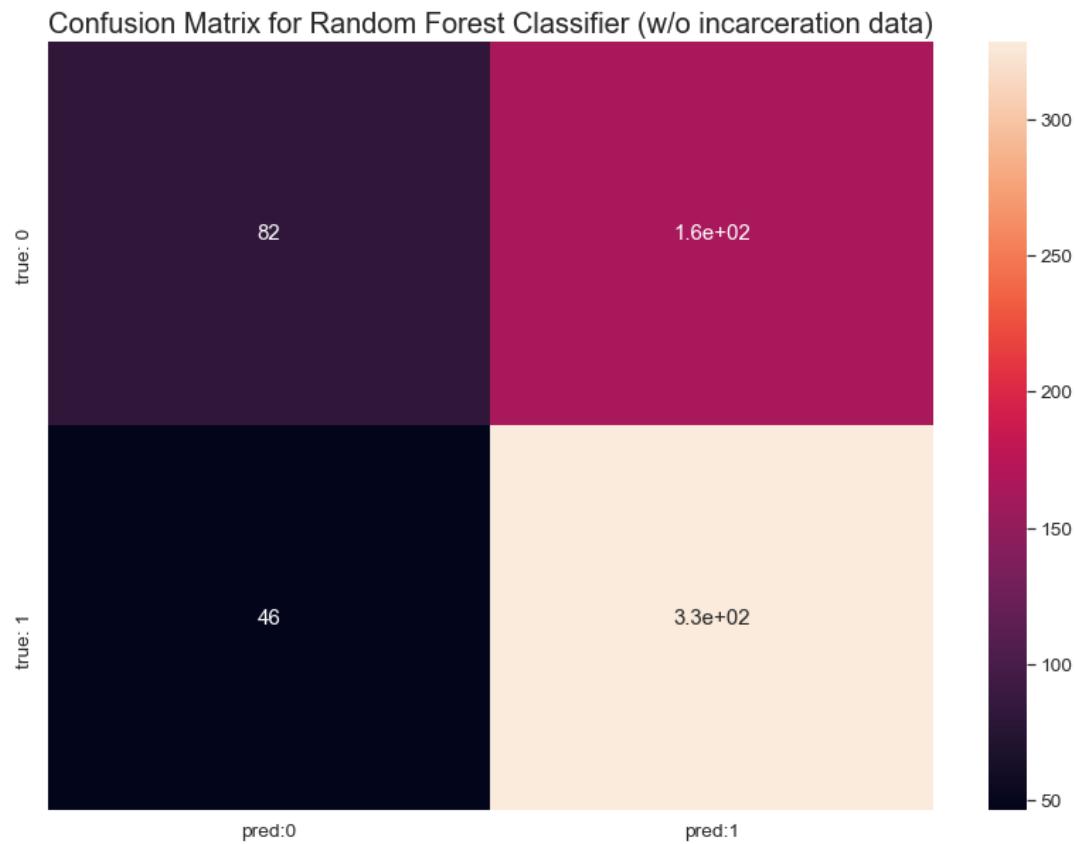
Previous years' incarceration rates were among the most important, along with population change rates for 2010 to 2018 and from 2017 to 2018.

~520 counties accurately classified



# Findings – without incarceration data

---



Key features with most predictive power:

- percentage of Non-Veterans who are on Disability;
- population change & natural change rates (2000-2010, and 2010-2018);
- median household income;
- percentage of the population aged 65 or older in 2010;
- and percentage employment change 2017-2018.

## Future Considerations

Incorporate additional county-level data, including election data and political party distributions.

Feature engineering to improve model performance and possibly reduce the number of total features.

Classification models on top  $n$  features to see if limited features could be used for continuous monitoring and predicting.

Regression analysis to predict the actual rates per county.



# Thank you

---

## Data Sources:

Vera Institute of Justice database of county and jurisdiction level incarceration rates from 1970 to 2017: [https://github.com/vera-institute/incarceration\\_trends](https://github.com/vera-institute/incarceration_trends)

The Database is comprised of the following Bureau of Justice Statistics (BJS) data collections: the Census of Jails (COJ), which covers all jails and is conducted every five to eight years since 1970, and the Annual Survey of Jails (ASJ), which covers about one-third of jails-and includes nearly all of the largest jails-that has been conducted in non-census years since 1982, and the BJS National Corrections Reporting Program (NCRP) data collection.

*"Incarceration Trends is supported by Google.org, the John D. and Catherine T. MacArthur Foundation Safety and Justice Challenge, and the Robert W. Wilson Charitable Trust."* <http://trends.vera.org/incarceration-rates?data=pretrial>

The United States Department of Agriculture's (USDA) Economic Research Service (ERS): <https://catalog.data.gov/dataset/county-level-data-sets>