



Data Strategy and Evaluation Seminar Series

DATA STRATEGY, PART 3

Mia Bird &
Ryan Souza

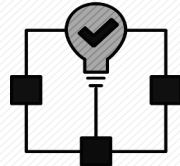
HOUSEKEEPING/ZOOM



STRATEGIC FRAMEWORK FOR DATA-DRIVEN AND EVIDENCE BASED PRACTICE



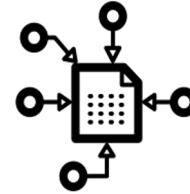
**Program
Inventory &
Literature Review**



**Logic Models &
Process Maps**



**Contracting
Practices**



**Data Strategy
& Dashboards**



**Program
Evaluation
& Cost Benefit
Analysis**

Seminar Series 1: Programming

Seminar Series 2: Data Strategy and Evaluation

Data Strategy, Part 1

- Accessing County Data (Nov 2020)

Data Strategy, Part 2

- Accessing State Data (Jan 2021)

Data Strategy, Part 3

- Integrating County and State Data (Mar 2021)

Data Dashboards

- Visualizing Data (May 2021)

Outcomes Evaluation

- Evaluating Program Outcomes (July 2021)

DATA & EVALUATION SERIES CORE COMPONENTS

DATA STRATEGY, PART 3 OVERVIEW

Integrating County & State Data

- County and DOJ data
 - Use Case: Linking Jail data with DOJ data for pretrial analysis
- County & CDCR data
 - Use Case: Linking Probation data with CDCR PRCS release data

LINKING
JAIL DATA
WITH DOJ DATA



PREPARING JAIL DATA FOR LINKING

1

Select & clean
key elements

2

Consolidate
booking &
release types

3

Clean offense
codes & create
hierarchy

4

Construct
indicators

5

Collapse to
booking event
level

SELECT & CLEAN KEY ELEMENTS

Jail data systems include more information than needed

- It is helpful to narrow the dataset to the key elements needed for your analysis before undertaking data linking

Key elements often include:

- CII number, local identifier, date of birth, gender, race/ethnicity
- Booking number, date, type, offense code, and charge level
- Release date and type

It's helpful to pre-clean the data by resolving any mis-matching information across individuals (such as inconsistencies in DOB)

CONSOLIDATE BOOKING & RELEASE TYPES

Coding schemes for jail data vary across systems, and it may be helpful to consolidate booking and release types into a common set of types

Booking Type

- New/fresh arrest
- Warrant
- Probation violation
- Parole violation
- Commitment
- Hold, transfer, other

Release Type

- Time served
- Sentenced capacity
- Pretrial (capacity, bail, program, OR, CR)
- Dismissed, acquitted
- Released to probation, parole, program
- Detention only, transferred, other
- Decreased

CLEAN OFFENSE CODES & CREATE OFFENSE HIERARCHY

Booking offenses are represented by penal codes in jail data

It's helpful to clean booking offenses using [law enforcement code tables](#)

The cleaning process involves connecting offense codes to the:

- Detailed offense description
- DOJ seriousness level

When jail data codes do not match to standard law enforcement codes, errors or instances of local codes may need to be resolved manually

CONSTRUCT INDICATORS

Once the jail data are clean, it's helpful to construct key indicators at the booking level, such as:

- Indicators for booking and release type
- Indicators for any person, property, drug, or other offense type
- Indicators for serious, violent, sexual, or DV offense
- Total number of charges, felony charges, or misdemeanor charges

These indicators allow you to characterize bookings in important ways and to retain this information once you collapse jail data

COLLAPSE TO BOOKING EVENT LEVEL

It's helpful to collapse each jail booking event into one observation (or row)

- Sometimes we call this “flattening the file”
- The flat file includes only one row for each booking event, but there may be multiple rows/booking events for an individual in the data

Data can be collapsed in different ways depending on your goal

- In this use case, we sort the data based on the DOJ seriousness level
- We then collapse the data to retain only the most serious booking charge, along with the indicators we created related to this booking event

QUESTIONS
OR COMMENTS



PREPARING DOJ CRIMINAL HISTORY DATA FOR LINKING

1

Select arrest events

2

Select conviction dispositions

3

Clean offense codes & create hierarchy

4

Construct indicators

5

Collapse to cycle level

SELECT ARREST EVENTS

DOJ Criminal History data are organized by cycles that begin with events

- Each cycle has an ID that connects all subsequent steps within that cycle

Cycles may begin with an arrest event or with other types of events

- For the purpose of building criminal history, we typically focus on events that begin with arrests
- Therefore, a common first step in setting up DOJ data is to select cycles that begin with an arrest event

Arrests provides us with one key component of criminal history

SELECT CONVICTION DISPOSITIONS

Once cycles beginning with an arrest event are selected, then we can look to subsequent court actions to identify convictions

Court action steps will include a set of possible dispositions

- For the purpose of building criminal history, we typically focus on conviction dispositions
- Therefore, a common second step in setting up DOJ data is to identify and select conviction dispositions
- Through this process, you reduce the complexity of the dataset to focus on the information relevant to your purpose

CLEAN OFFENSE CODES & CREATE HIERARCHY

Penal codes and offense descriptions are attached to arrest and conviction offenses in DOJ data

- As was the case with the jail booking data, it may be helpful to clean these offense codes using [law enforcement code tables](#)

The cleaning process allows you to match to the DOJ seriousness level and, therefore, to identify the most serious arrest and conviction charges

- We commonly characterize the arrest or conviction based on the most serious charges (e.g. felony arrest, felony person arrest, etc.)

CONSTRUCT INDICATORS

Next, you can construct indicators for each cycle that provides information about arrests and convictions, such as:

- Indicators for arrests and convictions by the most serious charge level
- Indicators for arrests and convictions by the most serious offense category
- Indicators for serious serious, violent, sexual, or DV offense
- Total number of charges by charge level and offense category

These indicators allow you to characterize arrests and convictions in important ways and to retain this information once you collapse data

COLLAPSE TO CYCLE LEVEL

We can then collapse the data to the cycle level, meaning we will have one observation for each cycle that begin with an arrest event

Prior to collapsing, we can sort the data so that we retain the most serious arrest charge and most serious conviction charge in the cycle observation

Match DOJ data to Jail booking data based on individual CII number and arrest/booking date



Identify and select focal pretrial jail spells based on booking and release types for the period of interest, as well as DOJ information



Use DOJ arrest and conviction dates for these individuals prior to focal pretrial spell to construct criminal history

INTEGRATING DATA FOR A PRETRIAL ANALYSIS

QUESTIONS
OR COMMENTS



LINKING
PROBATION DATA
WITH CDCR DATA



PREPARING PROBATION DATA FOR INTEGRATION

1

Select & clean key
elements

2

Create supervision
type hierarchy

3

Construct indicators

4

Collapse to
supervision event
level

SELECT & CLEAN KEY ELEMENTS

Probation data systems may include more information than needed

- It is helpful to narrow the dataset to the key elements needed for your analysis before undertaking data linking

Key elements often include:

- CII number, local identifier, date of birth, gender, race/ethnicity
- Supervision start type & date
- Supervision end type & date
- Risk & need assessments
- Programs & services interventions
- Violations & sanctions

It's helpful to pre-clean the data by resolving any mis-matching information across individuals (such as inconsistencies in DOB)

CREATE SUPERVISION TYPE HIERARCHY

Some individuals starting probation supervision may have multiple overlapping supervision types

- For example, they may start a new mandatory supervision term, but may also continue a previously suspended PRCS term

Creating a probation type hierarchy involves determining how you want to characterize the probation supervision event

- For each probation start date, it is helpful to have a dominant supervision type
- For example, PRCS > mandatory supervision > felony probation

CONSTRUCT INDICATORS

It's helpful to construct key indicators at the probation event level, such as:

- Indicators for each supervision start type
- Indicators and/or scores for risk and needs assessments
- Indicators and/or counts for program or service participation
- Indicators and/or counts for violations and sanctions received
- Indicators for each supervision end type
- Indicators for additional/overlapping supervision terms

These indicators allow you to characterize probation terms in important ways and to retain this information once you collapse probation data

COLLAPSE TO PROBATION EVENT LEVEL

The flat file for probation would include only one observation per supervision event, but there be multiple rows/supervision events for an individual in the data

Data can be collapsed in different ways depending on your goal

- In the use case, we will focus on the PRCS population, so we sort the data such that PRCS is the dominant supervision type
- We then collapse the data to the most serious supervision type

PREPARING CDCR DATA FOR INTEGRATION

1

Select key elements

2

Select focal release

SELECT KEY ELEMENTS

CDCR typically provides release cohort data with summary-level indicators, including:

- Current offense category and type
- Prior violent and prior serious offenses
- Prior prison terms and returns to prison
- Prison length of stay
- Indicators of behavioral health needs
- California static risk assessment (CSRA) level

CDCR may also be able to provide data on COMPAS risk and needs assessments and in-custody programmatic interventions, which may require more work to construct summary indicators

SELECT FOCAL RELEASE

Individuals under PRCS supervision may have multiple CDCR stays in their prison history

To assess the characteristics of the recent PRCS population entering the county, we might choose to select the most recent release

- Doing so requires sorting the data by release date and collapsing the file to the most recent release

Alternatively, to look at the changes in characteristics over a long period we might include all PRCS releases for an individual to the county

Identify PRCS probation spells based on their start date



```
graph TD; A[Identify PRCS probation spells based on their start date] --> B[Match CDCR release data to these probation spells based on CII and the release/start date]; B --> C[The resulting data file will include a rich set of characteristics describing the PRCS population];
```

Match CDCR release data to these probation spells based on CII and the release/start date

The resulting data file will include a rich set of characteristics describing the PRCS population

INTEGRATING DATA FOR PRCS POPULATION COMPOSITION ANALYSIS

QUESTIONS AND DISCUSSION



MATERIALS, GUIDES, AND TEMPLATES

Materials available:

[DOJ Law Enforcement Code Tables](#)

[DOJ Offense Codes](#)

[DOJ Data Flow Charts](#)

[California Penal Code](#)

Data & Evaluation Series Info:

<https://www.counties.org/framework-seminar-series>

Support Hub Website:

<https://www.counties.org/csac-support-hub>



LOOKING AHEAD

PROGRAMMING SERIES, SEMINAR 1.4 (APRIL)

Cost-Benefit Analysis, Part 1:

This session will focus on the key concepts of applying economic analysis to justice systems and programs. We will explain the key drivers of cost in justice systems, as well as how to develop localized analysis of sentencing outcomes, recidivism, and resource use. Building on early seminars on program inventories, this will also address how to develop high quality cost estimates of programming interventions

DATA AND EVALUATION SERIES, SEMINAR 2.4 (MAY)

Data Dashboards:

This seminar provides an introduction to designing data dashboards, a key tool in making your data more accessible and flexible. We first discuss the process of designing a data dashboard that will allow you to track the key population characteristics, interventions, and outcomes that are important for your agency or county. We will then present and discuss several data dashboards concepts that can be adapted for your purposes.



Contact:

Ryan Souza, Program Director

RSouza@counties.org

(m) 916.628.1404