



Bureau of Justice Statistics

Using Record Linkage to Assess Data Quality in Demographic Information Collected Across Federal Justice Agencies

Mark Motivans, Ph.D.

In workshop: Using Record Linkage of Administrative Records to Improve Federal Statistics

Federal Committee on Statistical Methodology

College Park, Maryland

October 23rd, 2024



Bureau of Justice Statistics

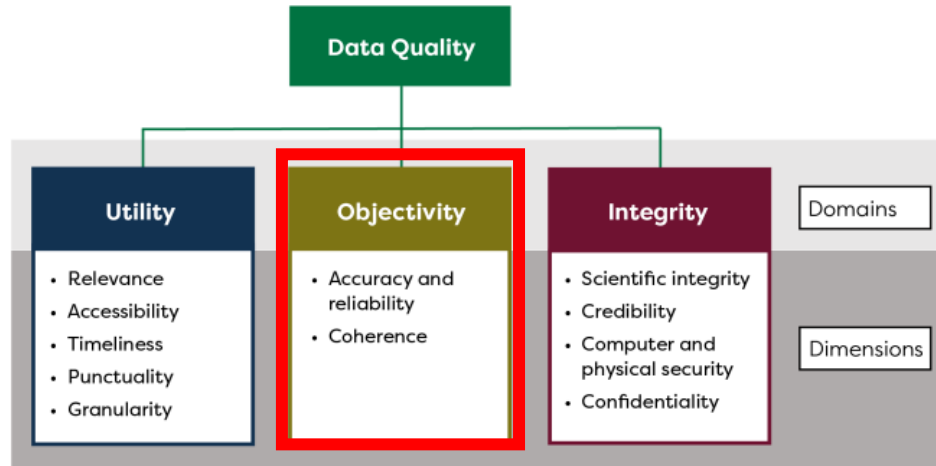
Presentation outline

- Problem statement
- Data and Methods
- Findings
- Discussion



- FCSM *Data Quality Analysis Working Group* defined data quality as:
 - “the degree to which data capture the desired information using appropriate methodology in a manner that sustains public trust”

Figure ES 1. The FCSM Data Quality Framework





Objectivity domain refers to whether information is accurate, reliable, and unbiased...

- Benefits of data quality assessment
 - Provides evidence to support data use for purposes beyond why they were collected
 - Promotes a more informed use of information
 - Prompts agency insights into areas in need of improvement
 - Data quality assessments generally less common for integrated and secondary-use data



We use linked administrative data from the BJS Federal Justice Statistics Program

- We focus on the source and comparability of data across two agencies (Federal Bureau of Prisons and the U.S. Sentencing Commission)
- The variables we investigated included:
 - Sex
 - Race
 - Ethnicity
 - U.S. citizenship status
- For linked records, demographic information is present in both datasets
 - To what extent is demographic information similar between agencies?



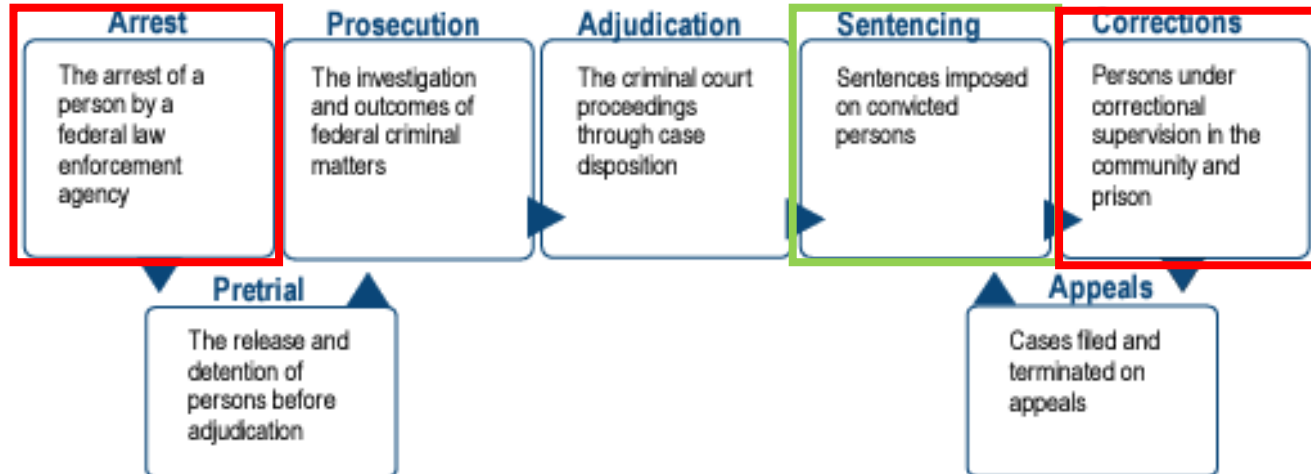
BJS

Bureau of Justice Statistics

Data and Methods

Sources of demographic information on persons processed

Exhibit 1. Federal Criminal Case Processing



Sources of demographic information: FBI Fingerprint Card

APPLICANT FD-258 (Rev. 11-1-20) 1110-0046		LEAVE BLANK		TYPE OR PRINT ALL INFORMATION IN BLACK		FBI		LEAVE BLANK	
LAST NAME		FIRST NAME		MIDDLE NAME					
SIGNATURE OF PERSON FINGERPRINTED		ALIASES AKA		O R I					
RESIDENCE OF PERSON FINGERPRINTED		DATE OF BIRTH		DOB		Month		Year	
DATE		SIGNATURE OF OFFICIAL TAKING FINGERPRINTS		CITIZENSHIP		RACE		HAIR	
EMPLOYER AND ADDRESS		YOUR NO.		UNIVERSAL CONTROL NO.		LHCN		LEAVE BLANK	
REASON FINGERPRINTED		ARMED FORCES NO.		MNU		CLASS			
		SOCIAL SECURITY NO.		SOC		REF.			
		MISCELLANEOUS NO.		MNU					

1. R. THUMB		2. R. INDEX		3. R. MIDDLE		4. R. RING		5. R. LITTLE	
6. L. THUMB		7. L. INDEX		8. L. MIDDLE		9. L. RING		10. L. LITTLE	
LEFT FOUR FINGERS TAKEN SIMULTANEOUSLY				L. THUMB		R. THUMB		RIGHT FOUR FINGERS TAKEN SIMULTANEOUSLY	

FEDERAL BUREAU OF INVESTIGATION UNITED STATES DEPARTMENT OF JUSTICE CJIS DIVISION/CLARKSBURG, WV 26306

1110-0046

APPLICANT

THIS CARD FOR USE BY:

- LAW ENFORCEMENT AGENCIES IN FINGERPRINTING APPLICANTS FOR LAW ENFORCEMENT POSITIONS,*
- OFFICIALS OF STATE AND LOCAL GOVERNMENTS FOR PURPOSES OF EMPLOYMENT, LICENSING, AND PERMITS, AS AUTHORIZED BY STATE STATUTES AND APPROVED BY THE ATTORNEY GENERAL OF THE UNITED STATES, LOCAL AND COUNTY ORDINANCES, UNLESS SPECIFICALLY BASED ON APPLICABLE STATE STATUTES TO NOT SATISFY THIS REQUIREMENT**
- U.S. GOVERNMENT AGENCIES AND OTHER ENTITIES REQUIRED BY FEDERAL LAW**
- OFFICIALS OF FEDERALLY CHARTERED OR INSURED BANKING INSTITUTIONS TO PROMOTE OR MAINTAIN THE SECURITY OF THOSE INSTITUTIONS.

Please review the helpful information to aid in the successful processing of hard copy civil fingerprint submissions in order to prevent delays or rejections. Hard copy fingerprint submissions must meet specific criteria for processing by the Federal Bureau of Investigation.

Ensure all information is typed or legibly printed using blue or black ink.
Enter data within the boundaries of the designated field or block.
Complete all required fields. (If a required field is left blank, the fingerprint card may be immediately rejected without further processing.)

* The required fields for hard copy civil fingerprint cards are: ORI, Date of Birth, Place of Birth, NAM, Ser, Date Fingerprinted, Reason Fingerprinted, and proper completion of fingerprint impression boxes.

Do not use highlighters on fingerprint cards.
Do not enter data or labels within "Leave Blank" areas.
Ensure fingerprint impressions are rolled completely from nail to nail.
Ensure fingerprint impressions are in the correct sequence.
Ensure notations are made for any missing fingerprint impression (i.e. amputation).
Do not use more than two white per fingerprint impression blocks.
Ensure no stray marks are within the fingerprint impression blocks.

Training aids can be ordered online via the Internet by accessing the FBI's website at: fbi.gov, click on "Fingerprint", then click on "Ordering Fingerprint Cards & Training Aids". Direct questions to the Biometric Services Section's Customer Service Group at (204) 625-5590 or by e-mail at <identity@fbi.gov>.

Social Security Account Number (SSAN): Pursuant to the Privacy Act of 1974, any Federal, state, or local government agency that requests an individual to disclose his or her SSAN, is responsible for informing the person whether disclosure is mandatory or voluntary, by what authority or other authority the SSAN is solicited, and what uses will be made of it. In this instance, the SSAN is solicited pursuant to 28 U.S.C. 534 and will be used as a unique identifier to confirm your identity, because many people have the same name and date of birth. Disclosure of your SSAN is voluntary; however, failure to disclose your SSAN may affect completion or approval of your application.

PRIVACY ACT STATEMENT

Authority: The FBI's acquisition, preservation, and exchange of fingerprints and associated information is generally authorized under 28 U.S.C. 534. Depending on the nature of your application, supplemental authorities include Federal statutes, State statutes pursuant to P.L. 92-544, Presidential Executive Orders, and Federal Regulations. Providing your fingerprints and associated information is voluntary; however, failure to do so may affect completion or approval of your application.

Principal Purpose: Certain determinations, such as employment, licensing, and security clearances, may be predicated on fingerprint-based background checks. Your fingerprints and associated information/biometrics may be provided to the employing, investigating, or otherwise responsible agency, and the FBI for the purpose of comparing your fingerprints to other fingerprints in the FBI's Next Generation Identification (NGI) system or its successor systems (including civil, criminal, and talent repositories) or other available records of the employing, investigating, or otherwise responsible agency. The FBI may retain your fingerprints and associated information/biometrics in NGI after the completion of this application and, while retained, your fingerprints may continue to be compared against other fingerprints submitted to or retained by NGI.

Routine Uses: During the processing of this application and for as long thereafter as your fingerprints and associated information/biometrics are retained in NGI, your information may be disclosed pursuant to your consent, and may be disclosed without your consent as permitted by the Privacy Act of 1974 and all applicable Routine Uses as may be published at any time in the Federal Register, including the Routine Uses for the NGI system and the FBI's Biometric Routine Uses. Routine uses include, but are not limited to, disclosures to: employing, governmental or authorized non-governmental agencies responsible for employment, contracting, licensing, security clearances, and other suitability determinations; local, state, tribal, or federal law enforcement agencies; criminal justice agencies; and agencies responsible for national security or public safety.

PAPERWORK REDUCTION ACT NOTICE

According to the Paperwork Reduction Act of 1995, no persons are required to provide the information requested unless a valid OMB control number is displayed. The valid OMB control number for this information collected is 1110-0046. The time required to complete this information collected is estimated to be 10 minutes, including time reviewing instructions, gathering, compiling, reviewing and submitting the information collection. If you have any comments concerning the accuracy of this time estimate or suggestions for reducing the burden, please send to: Department Clearance Officer, United States Department of Justice, Justice Management Division, Policy and Planning Staff, Washington, DC 20535.

INSTRUCTIONS:

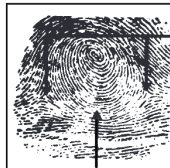
- PRINTS MUST GENERALLY BE CHECKED THROUGH THE APPROPRIATE STATE IDENTIFICATION BUREAU AND ONLY THOSE FINGERPRINTS FOR WHICH NO DISQUALIFYING RECORD HAS BEEN FOUND LOCALLY SHOULD BE SUBMITTED FOR FBI SEARCH.
 - IDENTITY OF PRIVATE CONTRACTORS SHOULD BE SHOWN IN SPACE "EMPLOYER AND ADDRESS". THE CONTRIBUTOR IS THE NAME OF THE AGENCY SUBMITTING THE FINGERPRINT CARD TO THE FBI.
- UNIVERSAL CONTROL NUMBER, IF KNOWN, SHOULD ALWAYS BE FURNISHED IN THE APPROPRIATE SPACE.

1. LOOP



THE LINES BETWEEN CENTER OF LOOP AND DELTA MUST SHOW

2. WHORL



THESE LINES RUNNING BETWEEN DELTAS MUST BE CLEAR

3. ARCH



ARCHES HAVE NO DELTAS

FD-258 (Rev. 11-1-20)



Bureau of Justice Statistics

Data and Methods

Sources of demographic information: Presentence Investigation Report

WORKSHEET FOR PRESENTENCE REPORT

(See Publication 107 for Instructions)

1. FACESHEET DATA		
Defendant's Court Name:		
Defendant's True Name:		
Docket No.:	District:	
Judge/Magistrate:	Sentencing Date:	
USPO:	Arrest Date:	
Assistant U.S. Attorney (Name, address, telephone)		Defense Counsel (Name, address, telephone)
DEFENDANT'S IDENTIFICATION		
Defendant's Names: (List every name the defendant has used, e.g., name given at birth, name given at adoption, nickname, alias, names used as a result of marriage, etc.)		
Date of Birth:	Age:	Place of Birth:
Race: <input type="checkbox"/> White <input type="checkbox"/> Black <input type="checkbox"/> American Indian/Alaskan Native		Hispanic Origin:
<input type="checkbox"/> Asian or Pacific Islander <input type="checkbox"/> Unknown		<input type="checkbox"/> Hispanic <input type="checkbox"/> Not Hispanic <input type="checkbox"/> Unknown
Sex:	Country of Citizenship:	Immigration Status:
No. of Dependents:	Education:	SSN:
FBI No.:	U.S. Marshal's No.:	Other ID No.:
Defendant's Legal Address:		
(Number and Street)		(Apartment)
(City)	(State)	(Zip)
Defendant's Current Address:		
(Number and Street)		(Apartment)
(City)	(State)	(Zip)

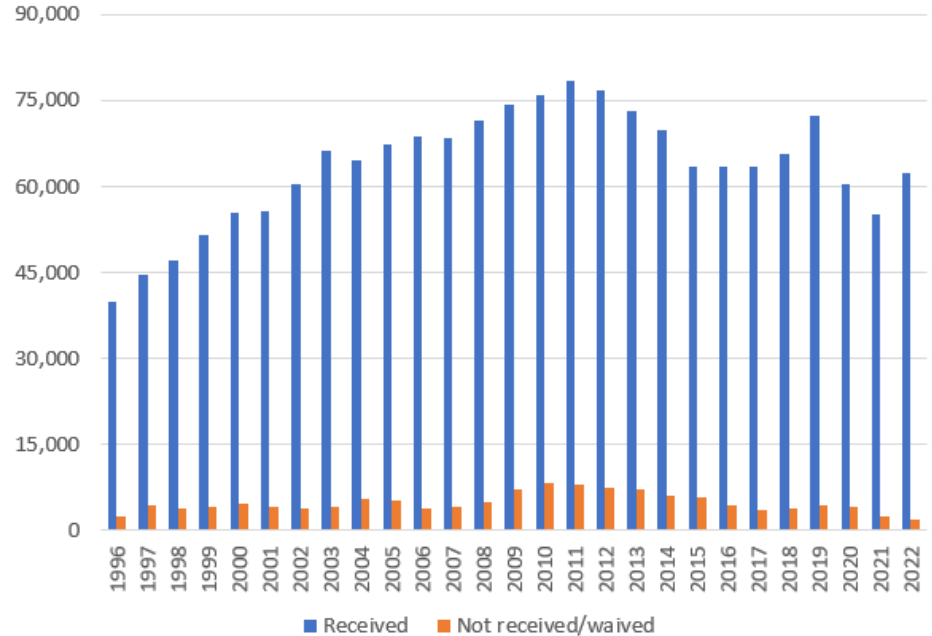


Bureau of Justice Statistics

Data and Methods

The U.S. Sentencing Commission reports whether they receive the PSR in their data collection

U.S. Sentencing Commission's Presentence Investigation Report Status, FY 1996-2022



Source: U.S. Sentencing Commission. Sourcebook of Federal Sentencing Statistics (annual).



Bureau of Justice Statistics

Data and Methods



Expectations

- Data are keyed from paper forms or captured during interview
 - Source of information is not always recorded (e.g., self-report, transcribed from official record, agent decision, etc.)
- The Presentence Investigation Report (PSR) is recorded as the source for information in the U.S. Sentencing Commission data and whether they received it
- Will match rates improve with PSR availability or worsen when PSR is not available?

Data Used: Screening Criteria

- Screening criteria are used to limit records of persons admitted to prison from FY 2016-2022 to those within scope of USSC data
- Criteria exclude the following:
 - persons violating supervision
 - cases from DC Superior Court
 - military prisoners transferred to BOP
 - international treaty transfers
 - persons boarded for a state offense
- About 80% of total records were included

Fiscal year	Persons admitted to federal prison		Met screening criteria	
	Number	Percent	Number	Percent
Total	374,011	100%	298,689	80%
2016	62,329	100%	51,666	83%
2017	60,208	100%	48,951	81%
2018	59,285	100%	47,662	80%
2019	58,929	100%	45,477	77%
2020	36,951	100%	28,786	78%
2021	45,642	100%	35,884	79%
2022	50,667	100%	40,263	79%

Source: Bureau of Justice Statistics, based on data from the Federal Bureau of Prisons, SENTRY database, fiscal years 2016-2022.



Bureau of Justice Statistics

Link Rate

- 94% of BOP admission records from FY 2016-2022 linked with USSC sentencing records from FY 1994-2022
- About 75% of all admissions were matched (mostly due to screening rules)

Fiscal year	Persons admitted to federal prison	With linked records	
	Number	Number	Percent
Total	298,689	280,058	94%
2016	51,666	45,879	89%
2017	48,951	44,182	90%
2018	47,662	43,813	92%
2019	45,477	42,866	94%
2020	28,786	27,957	97%
2021	35,884	35,529	99%
2022	40,263	39,832	99%

from the Federal Bureau of Prisons, SENTRY database, fiscal years 2016-2022 linked with data from the U.S. Sentencing Commission, Monitoring File, fiscal years 1994-2022.

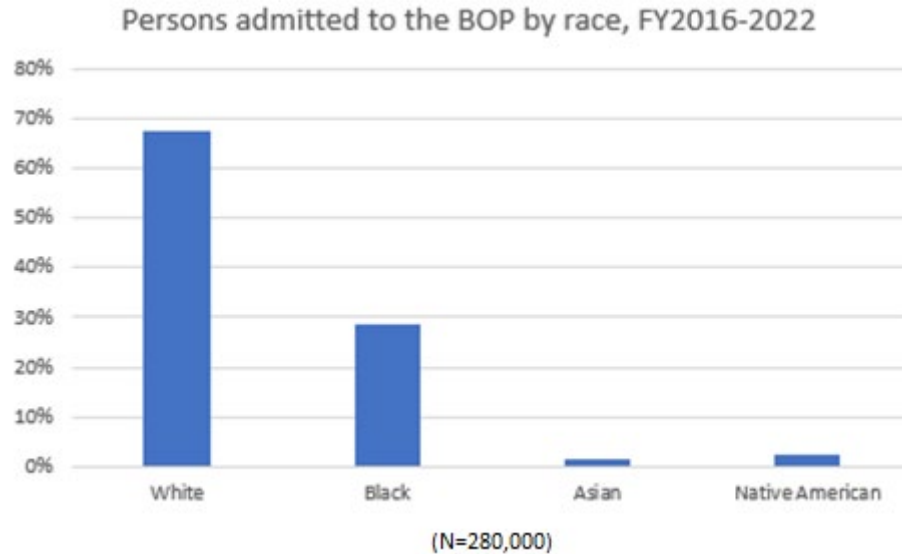




Race classification used by USSC and BOP

Agency	Stage	Values
U.S. Sentencing Commission (USSC)	Sentencing	<ul style="list-style-type: none">• White/Caucasian• Black/African American• American Indian/Alaska Native• Asian or Pacific Islander• Multi-racial• Other• Not Available in Docs Non-US American Indian
Federal Bureau of Prisons (BOP)	Persons admitted to federal prison	<ul style="list-style-type: none">• White• Black/African American• American Indian• Asian/Native Hawaiian/Other Pacific Islander• Unknown

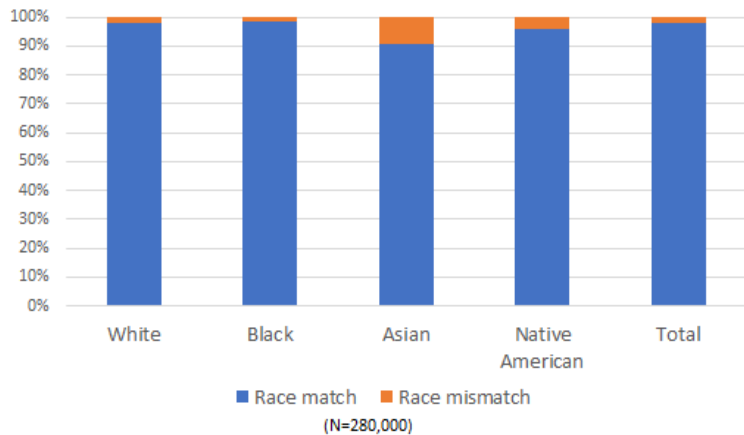
Distribution of sample by race



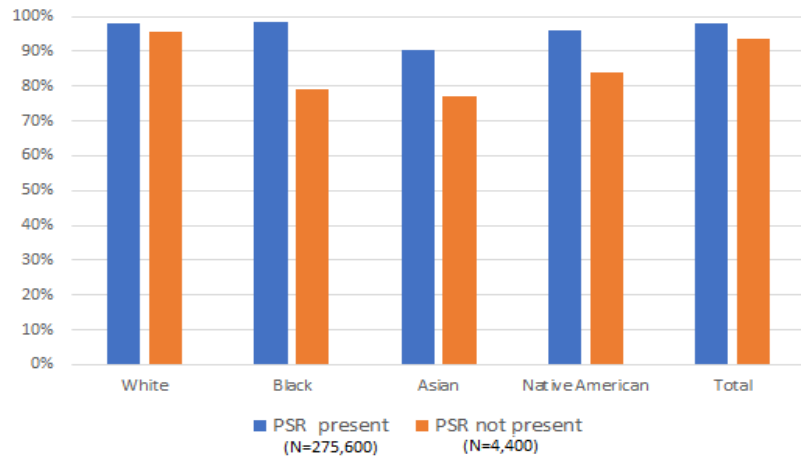
Bureau of Justice Statistics

Match rate for race

Percent of linked records where race matched, FY 2016-2022



Percent of linked records where race matched by whether PSR was present, FY 2016-2022



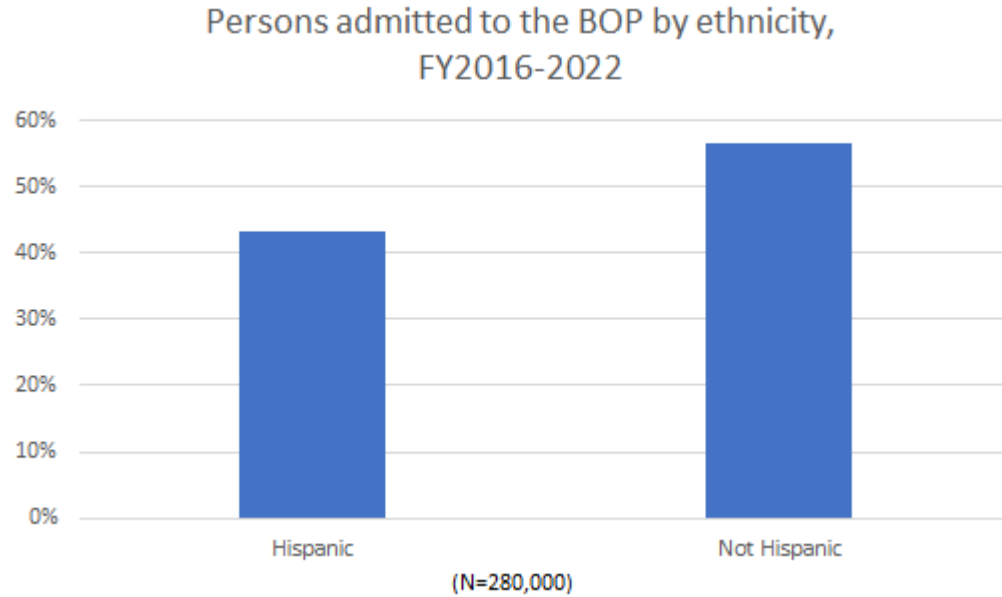
Bureau of Justice Statistics



Hispanic origin (ethnicity) classification used by USSC and BOP

Agency	Stage	Values
U.S. Sentencing Commission (USSC)	Sentencing	<ul style="list-style-type: none">• Information on Hispanic Origin Not Available• Non-Hispanic• Hispanic
Federal Bureau of Prisons (BOP)	Persons admitted to federal prison	<ul style="list-style-type: none">• Not known• Hispanic• Non-Hispanic

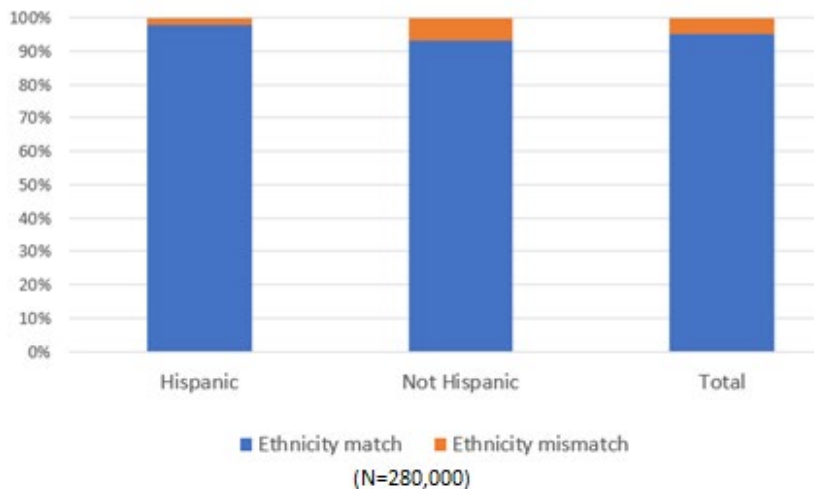
Distribution of sample by ethnicity



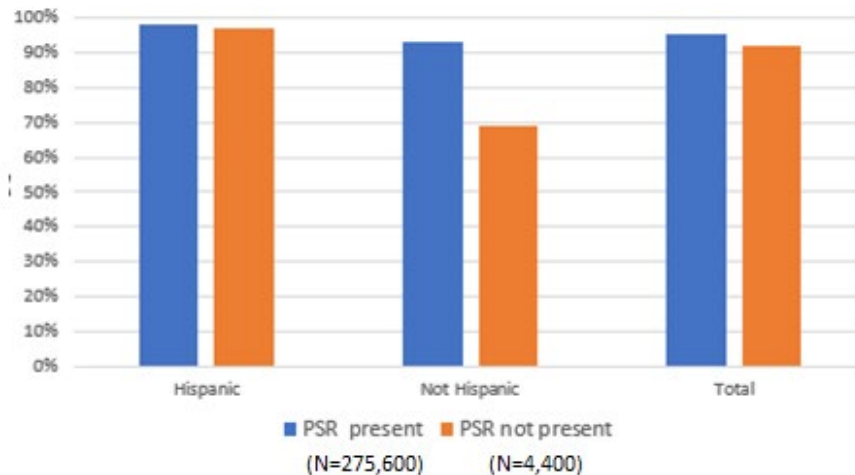
Bureau of Justice Statistics

Match rate for Hispanic origin (ethnicity) of person

Percent of linked records where ethnicity matched, FY 2016-2022



Percent of linked records where ethnicity matched by whether PSR was present, FY 2016-2022



Bureau of Justice Statistics

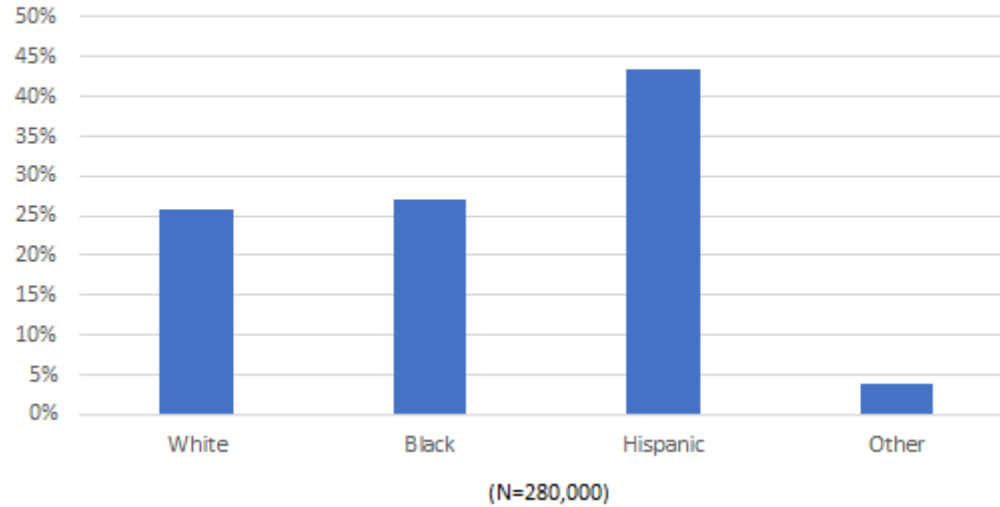


Race/Hispanic classification used by USSC and BOP

Agency	Stage	Values	Comment
U.S. Sentencing Commission (USSC)	Sentencing	<ul style="list-style-type: none">• White• Black• Hispanic• Other• Missing, Indeterminable, or Inapplicable	<ul style="list-style-type: none">• Constructed by USSC
Federal Bureau of Prisons (BOP)	Persons admitted to federal prison	<ul style="list-style-type: none">• White• Black• Hispanic• Other	<ul style="list-style-type: none">• Constructed by author using BOP race and Hispanic origin variables

Distribution of sample by race/ethnicity

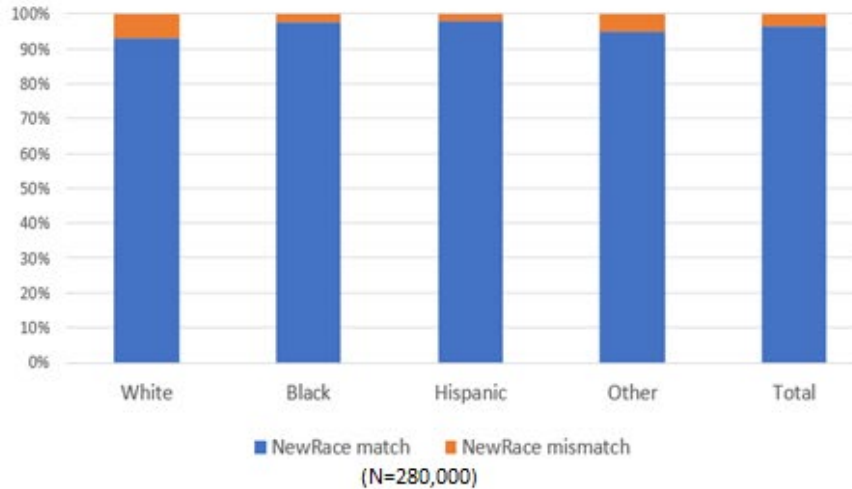
Persons admitted to the BOP by race/ethnicity,
FY2016-2022



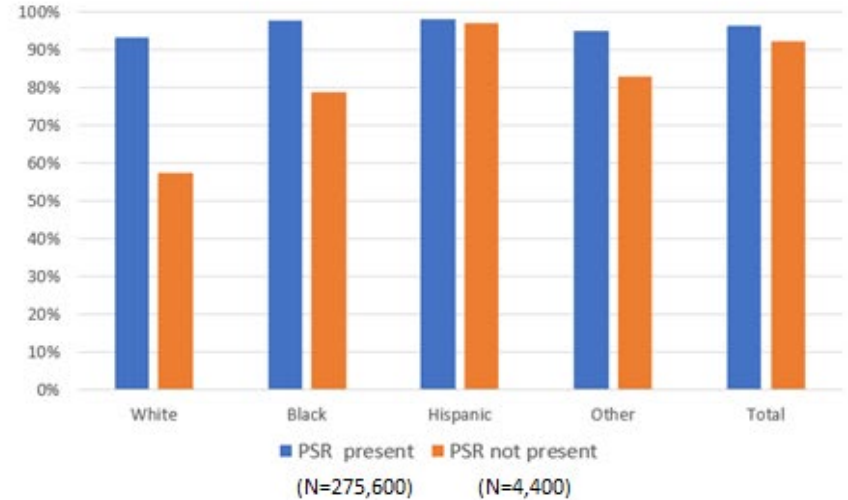
Bureau of Justice Statistics

Match rate for race/ethnicity

Percent of linked records where race/ethnicity (newrace) matched, FY 2016-2022



Percent of linked records where race/ethnicity (newrace) matched by whether PSR was present, FY 2016-2022



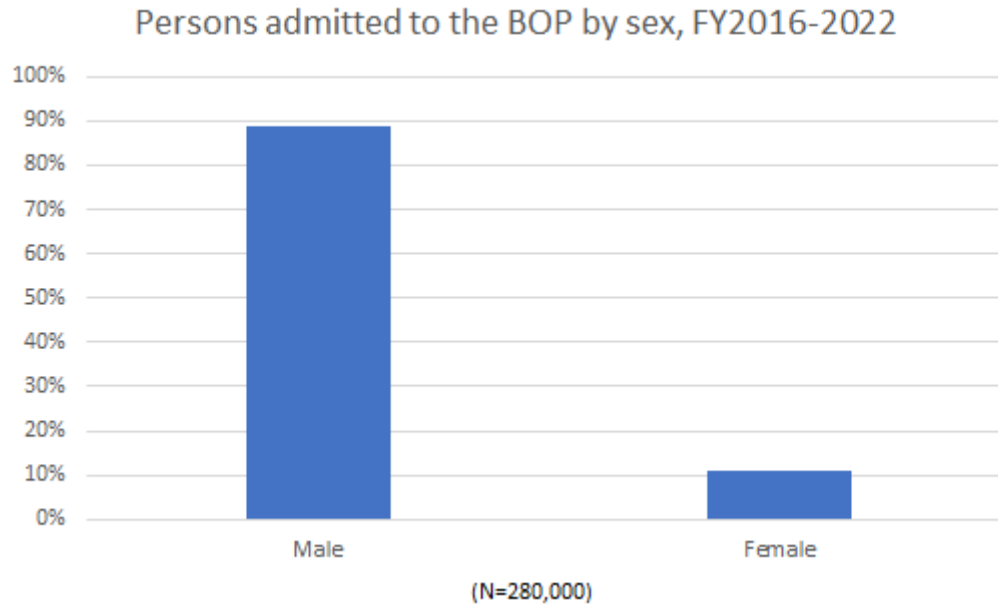
Bureau of Justice Statistics



Sex classification used by USSC and BOP

Agency	Stage	Values
U.S. Sentencing Commission (USSC)	Sentencing	<ul style="list-style-type: none">• Male• Female• Missing, Indeterminable, or Inapplicable
Federal Bureau of Prisons (BOP)	Persons admitted to federal prison	<ul style="list-style-type: none">• Male• Female• Not known

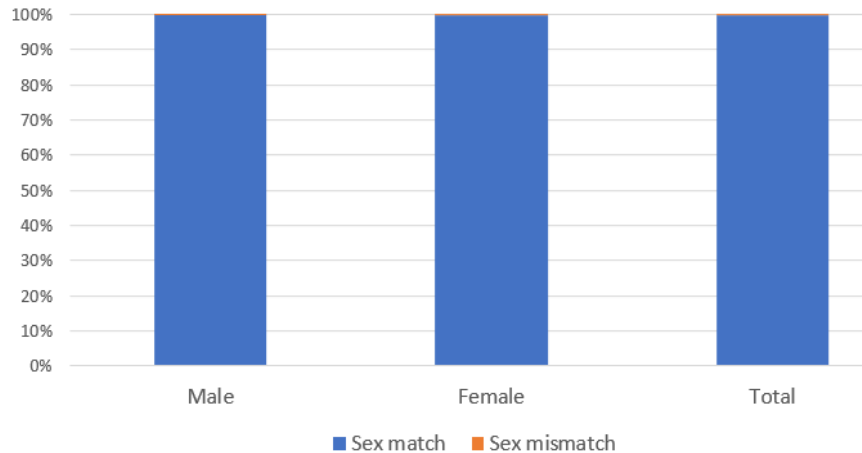
Distribution of sample by sex of person



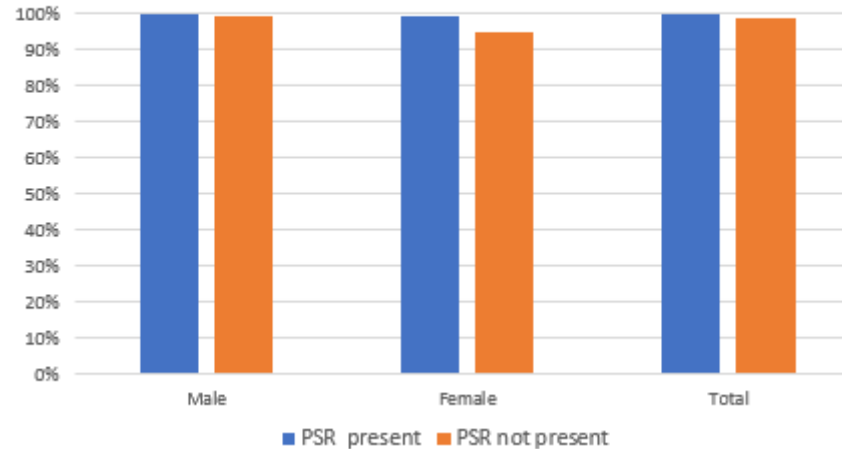
Bureau of Justice Statistics

Match rate for sex of person

Percent of linked records where sex matched,
FY 2016-2022



Percent of linked records where sex matched by whether
PSR was present, FY 2016-2022



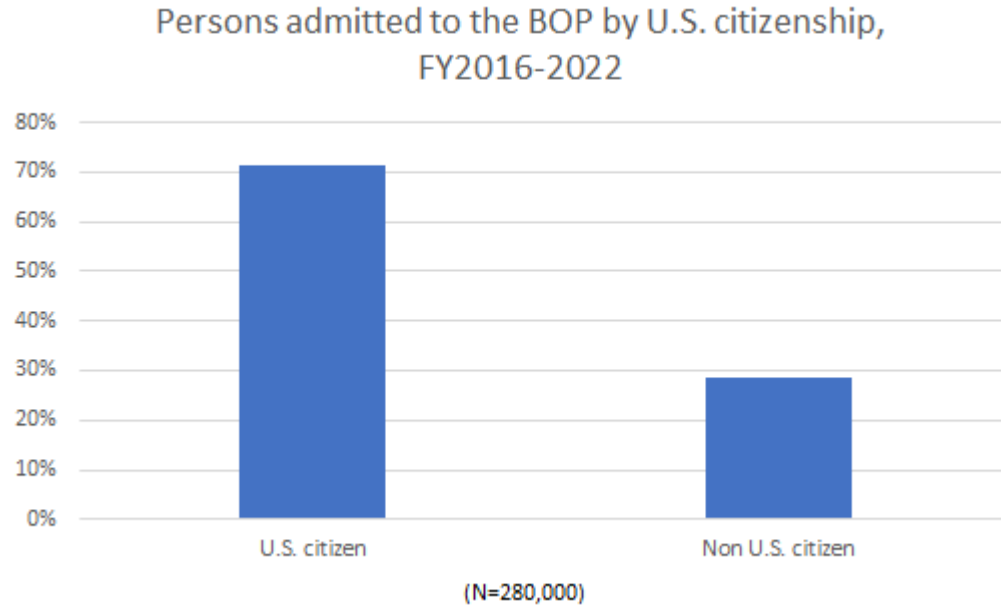
Bureau of Justice Statistics



U.S. citizenship status classification used by USSC and BOP

Agency	Stage	Values
U.S. Sentencing Commission (USSC)	Sentencing	<ul style="list-style-type: none">• U.S.• Non-U.S. (Includes Legal and Illegal Aliens)• Missing, Indeterminable, or Inapplicable
Federal Bureau of Prisons (BOP)	Persons admitted to federal prison	<ul style="list-style-type: none">• Missing• Non-US Citizen• US Citizen

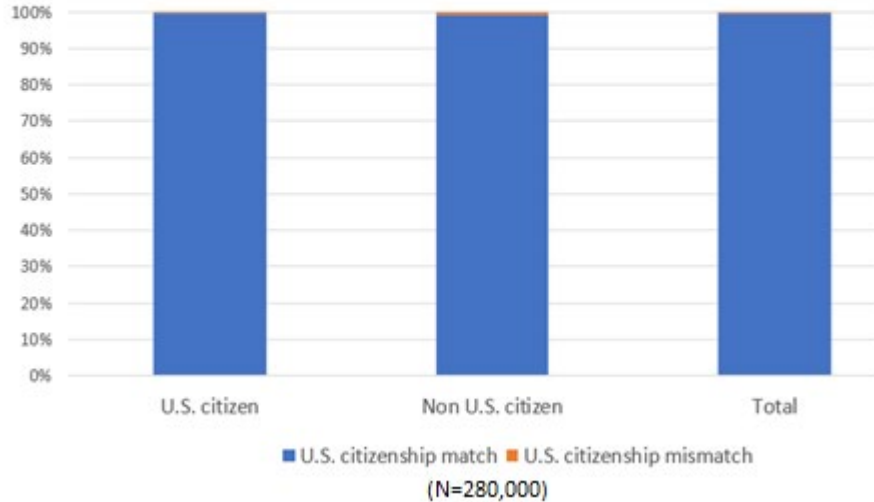
Distribution of sample by U.S. citizenship status



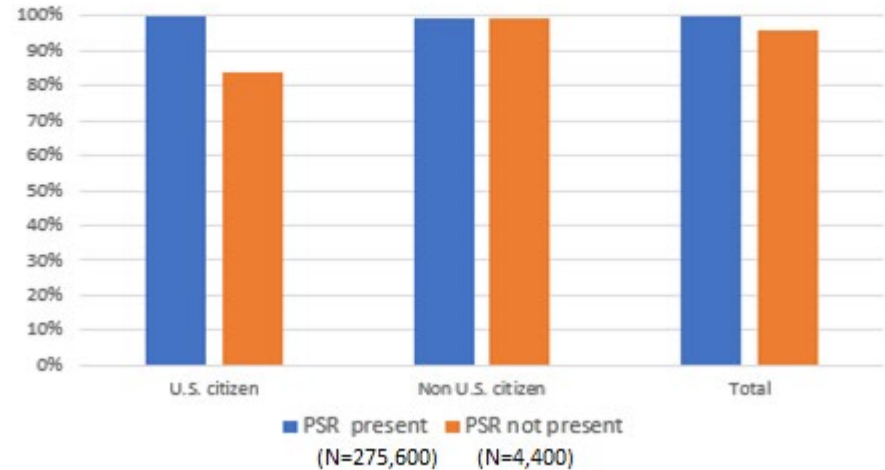
Bureau of Justice Statistics

Match rate for U.S. citizenship status

Percent of linked records where U.S. citizenship matched,
FY 2016-2022

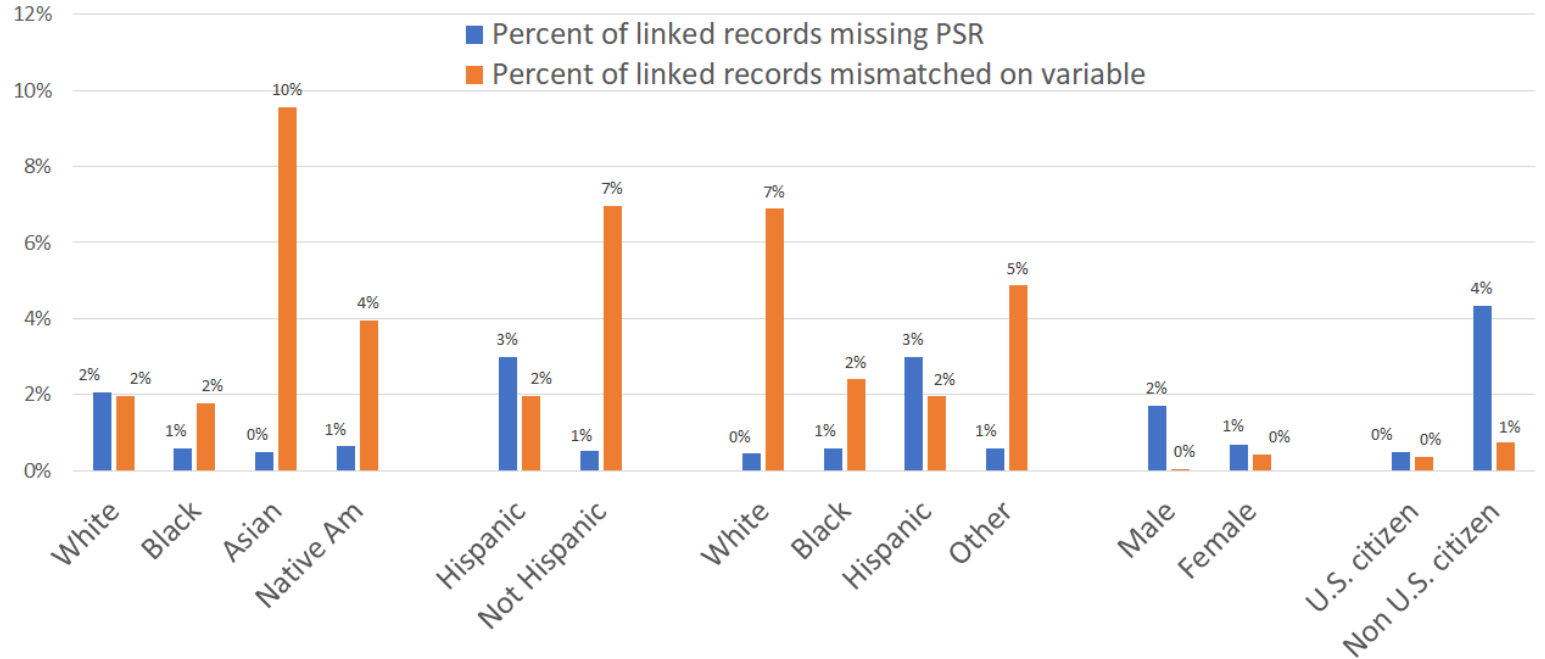


Percent of linked records where U.S. citizenship matched
by whether PSR was present, FY 2016-2022



Bureau of Justice Statistics

Summary of findings



Bureau of Justice Statistics



Findings/Discussion

- We found a very high level of consistency among demographic variables when comparing linked data
- Availability of PSR did not appear to substantially impact the match rates
- Future work includes:
 - More systematically collecting information about the source of information
 - More rigorously reporting on data quality issues with integrated data and how it may affect results
 - Additional work to understand the effects of using another agency's classification by modeling to show impact

Questions?

Askbjs@ojp.usdoj.gov

Data tool

<https://fccps.bjs.ojp.gov/>



Bureau of Justice Statistics

810 Seventh Street, NW, Washington, DC 20531

Phone: +1 (202) 307-0765

bjs.gov