

Criminal Cases in State Courts (CCSC)

**United States Department of Justice
Bureau of Justice Statistics**

Electronic Data Extraction Guide

January 2020

Prepared by:

The Urban Institute/National Center for State Courts

Contents

What is the study about?	2
Who do I contact if I have questions?.....	2
Will the data be secure and kept confidential?	3
How do I submit my file(s)?	4
When is the submission due?	4
How do I prepare the data for my jurisdiction?	4
File Structure	4
File Format	5
Supporting Documentation	5
How do I identify eligible cases?.....	5
What data elements should I include?	6
What if I am unable to provide all the requested data?.....	7
What happens after we submit the data?	7
Appendix A. Requested Data Elements, Definitions, and Formats.....	8

Criminal Cases in State Courts (CCSC) Electronic Data Extraction Guide

What is the study about?

The U.S. Department of Justice (DOJ), Bureau of Justice Statistics (BJS) has initiated a data collection of case-level data on felony and misdemeanor cases in state criminal courts across the United States with the goal of developing nationally representative state court statistics. BJS previously maintained two reporting programs—the National Judicial Reporting Program (NJRP), last completed in 2006, and the State Court Processing Statistics (SCPS), last completed in 2009—to collect and analyze case-level data from criminal courts. The Criminal Cases in State Courts (CCSC) leverages current advances in court data systems and standards to update the reporting on state court activity. The CCSC sample will be drawn from all felony and misdemeanor cases disposed in state criminal courts in 2019. This information is critical to understanding the number and types of cases processed in state criminal courts nationwide, along with other characteristics of these cases and the associated defendants.

BJS anticipates that the CCSC will inform 1) the field of court administration of current practice and trends in criminal case processing and 2) the public of how criminal cases are handled in the United States. Findings from the CCSC will be reported in aggregate form, precluding the possibility of revealing the identity of individual defendants or jurisdictions, and any identifying information will be removed or anonymized before archiving the data for research purposes. The CCSC can also serve as a benchmark to state court administrators and policymakers who wish to compare their own data to these national estimates. In other words, states and counties could better understand their own level of criminal justice activity in a broad, national context.

The Urban Institute (Urban) and the National Center for State Courts (NCSC) will collect these data on behalf of BJS. The ability to collect data electronically from automated information systems depends on each jurisdiction's capacity to extract the necessary data from their data management system. BJS, Urban and NCSC recognize this variation in state court systems and have developed this data extraction guide to support all modes of data collection. This data extraction guide will (1) describe the secure electronic data submission process and (2) provide a set of reporting instructions and standards for jurisdictions that have the capacity to extract data in electronic format. Appendix A includes a list of the data elements we seek to collect along with their definitions, preferred formats, and example values.

Who do I contact if I have questions?

An CCSC research team member from Urban or NCSC will contact your jurisdiction by email to initiate and facilitate a data request. For questions about extracting or submitting data, please reach out to this original Urban or NCSC point of contact or:

CCSC Help Desk

Toll Free: 855-279-3515

Email: CCSC@urban.org

Mail: 500 L'Enfant Plaza SW, Washington, D.C. 20024

For general questions about the CCSC study:

Suzanne Strong

Statistician and CCSC Project Manager

Bureau of Justice Statistics

810 Seventh St., NW

Washington, DC 20531

202-616-3666

Will the data be secure and kept confidential?

Consistent with its statutory obligations (34 U.S.C. § 10134), BJS only uses information collected under its authority for statistical or research purposes. Further, BJS is required by law to protect the confidentiality of all personally identifiable information it collects or acquires in conjunction with BJS-funded projects (34 U.S.C. § 10231), and must maintain the appropriate administrative, physical, and technical safeguards to protect the identifiable information against improper use or unauthorized disclosure. BJS will not use or reveal data identifiable to a private person, except as authorized under 28 CFR § 22.21 and § 22.22. The BJS Data Protection Guidelines summarize the federal laws, regulations, and other authorities that govern information acquired under BJS's authority, and are published on the BJS website: https://www.bjs.gov/content/pub/pdf/BJS_Data_Protection_Guidelines.pdf. Urban and NCSC are required to adhere to these same requirements as a condition of funding. Any data collected as part of the CCSC cannot and will not be used in any enforcement actions or administrative procedures to limit access to a program or activity that receives Federal funds or other Federal financial assistance (Title VI of the Civil Rights Act).

BJS only publishes de-identified data at the aggregate level in its project findings, reports, data files, and other statistical products. BJS archives its published data and related data documentation (e.g., user guides) at the National Archive of Criminal Justice Data (NACJD), located at the University of Michigan. To the extent practical, BJS removes, masks, or collapses direct and indirect identifiers prior to sending data to NACJD to protect confidentiality. NACJD takes additional precautions to protect confidentiality, including conducting a comprehensive disclosure risk review to determine the appropriate level of security that should be applied to the data. For more information on data requiring additional security protections, please see: <https://www.icpsr.umich.edu/icpsrweb/content/NACJD/restricted.html>.

Urban and NCSC project staff will receive data sets from contributing agencies through a Secure File Transfer Protocol (SFTP) set up by Urban's Technology and Data Science team. Personally-identifiable information is encrypted while in transit, and once a dataset is received it will be transferred to a secure encrypted drive. Access to the data will be limited to those employees who have a need for such data and have signed a confidentiality pledge. The confidentiality pledge includes an agreement to comply with all data security and human subjects' protection requirements. Any data set(s) electronically transmitted to BJS will be over the DOJ's Office of Justice Programs (OJP) secure transfer site.

We understand that your agency may have preexisting policies in place around data sharing, and we will work with your agency to meet any data transfer or agreement requirements you may have. While each of the identified data elements were selected to fully understand the processing of cases disposed in state criminal courts, we recognize that not all data elements may be collected or readily available electronically for public use. Please notify us if you limit the amount or type of data you can release.

How do I submit my file(s)?

You can use Urban's SFTP to transfer your agency's data. Urban will provide you with a username and password to log in to the project's website and deposit the electronic data. You may deposit as many files as necessary and in any format. For example, your jurisdiction may want to deposit several files that must be linked to get complete case data, as well as a Word document that provides the record layout for the files and linking instructions. Additional files, as necessary, may be submitted by your jurisdiction and separately tracked.

An Urban research team member will contact your jurisdiction to initiate and facilitate the data submission process. If needed, the Urban research team will provide guidance and assistance with regard to uploading data via the SFTP. If your jurisdiction requires an alternative means of submission, Urban will work with you to find the most convenient secure method for you to submit your data.

Please do not transfer data over non-secure mechanisms, such as unencrypted file transfer over the internet or unencrypted email.

When is the submission due?

The target date is **XXXX**, but we understand that may not be possible for every agency. Please contact us and we will work with you to set a realistic target date.

How do I prepare the data for my jurisdiction?

This section provides guidance on reporting specifications, including file structure, supporting documentation, case identification and selection, and variable and value definitions.

There is no required format for the data you submit; use whatever is most convenient for you. The suggested coding classifications and value labels – developed to be consistent with NCSC's draft National Open Court Data Standards (NODS) – are provided in Appendix A to assist with data submission. Your project point of contact will also provide technical assistance, including guidance in interpreting the data request. We recognize that the participating courts are unique and have different data systems with different levels of detail and data sharing capabilities. Rather than placing the burden on the data contributor to submit data in this uniform format, we will accept any data that each jurisdiction can provide and in turn will process it into a standard format that can be used to produce the national estimates.

To do this effectively, NCSC and Urban Institute will study and compare the submitted data to the extract guideline specifications to determine if they require any restructuring. For each submission, NCSC and Urban Institute will communicate with contributing jurisdictions to clarify any questions raised during the review. For files that require substantial restructuring, Urban Institute and NCSC staff will consult with jurisdiction personnel to obtain file documentation that will enhance understanding of the data. During this process, NCSC and Urban Institute will be in close communication with the jurisdiction to ensure that the data are processed and interpreted correctly.

File Structure

Both case- and event-level data is requested. You can provide data in any format that is convenient for you; **please provide supporting documentation to ensure that we interpret the data properly.**

One way to structure data is to provide one file at the case level and multiple event-level files with clearly established relationships between the appropriate fields and records of source tables and the targeted output layout. Using this format, the case level file would include one record for each disposed case. If there are multiple charges, they would all appear in the same record. Event-level files could include more than one record for each event in the case (e.g. separate records for each attorney assignment or court hearing). Case or docket numbers (with defendant identifiers, if applicable) should be included in each file to allow for record linking.

File Format

The preferred electronic file formats include:

- Text files (fixed width, delimited);
- Excel (or .csv) files;
- Stata data files;
- SAS data files;
- R data files;
- SPSS data files;
- Access database; or
- SQL server database.

Supporting Documentation

Please provide some tracking information and relevant documentation for your data submission, including:

- Date of submission
- Primary point of contact: name, organization, position, address, telephone, and e-mail address
- File format and format version (e.g., SQL server DB, Access DB, text files (fixed width, delimited), Excel, etc.)
- Known data quality problems or limitations
 - Missing data:
 - System-missing (data element not available in system)
 - Unit-missing (large % of missing values for a particular data element)
 - Common data errors, system-wide or with specific elements (e.g. misspellings, redundancy, duplication, contradictory values)
- Data formatting information:
 - Diagrams, data dictionaries, and/or field descriptions where available
 - Documentation of variables that diverge from the definitions provided in Appendix A
 - Any lookup table(s) that helps to translate data values
 - Graphical representation of the physical data source to show underlying database architecture that would inform the restructuring of data files
 - Information on the level at which data is provided for each table (e.g., case-level, event-level)

How do I identify eligible cases?

The data provided should include all felony and misdemeanor cases disposed in state criminal courts of general and limited jurisdiction, excluding municipal courts, in 2019. Case dispositions may be entered

as nolle prosequi, guilty finding or verdict, dismissal, or acquittal, among others (e.g. directed verdict, incompetent to stand trial). Cases disposed in 2019 but not sentenced should be included. Please exclude from the data file all cases still pending a final decision from a judicial officer (e.g., those in which there is not a disposition for the original charge(s)). For appeals, include all cases appealed from a limited jurisdiction court to a general jurisdiction court if there is a final disposition in the general jurisdiction court. Do not include any cases that are remanded, unless the remand includes a final disposition in either a limited or general jurisdiction court. Exclude any violations of probation or parole hearings or cases.

The data file should include the following case types:

- Felonies;
- Misdemeanors;
- Criminal traffic;
- Appeals or transfers from limited jurisdiction courts to general jurisdiction courts;
- Transfers from other jurisdictions; and
- Juvenile court cases transferred to adult court.

The data file should exclude the following types of cases if possible. If it would be too difficult to remove these cases, please note this in your submission.

- Cases disposed in municipal courts, unless it would be too difficult to remove these cases;
- Civil traffic;
- Appellate court cases;
- Juvenile court cases (except if a juvenile is tried by the adult court);
- Cases in which a defendant was diverted but has not yet completed the diversion program; and
- Violations of probation or parole.

BJS defines a criminal case as a set of all charges against a single defendant arising out of one incident. An incident with a continuing course of conduct (e.g., sexual assaults against one victim over time; thefts committed by one defendant multiple times) meets this definition. If your definition of a case differs from this (e.g., single citation per case, multiple defendants per case), please call your project point of contact or the CCSC Help Line at 855-279-3515 and note this in your submission. A case should include all charges under the same case number; if multiple charges were disposed at different points in time, then select cases where the most serious charge was disposed in 2019.

What data elements should I include?

Urban and NCSC are requesting data elements pertaining to multiple points throughout the life of a case. Appendix A lists the data elements we are requesting and their definitions, which are consistent with the draft NCSC NODS. Preferred formats and example values for each data element are also provided in Appendix A. Please note that preferred formats and example values are optional; **please provide data in the format most convenient for you along with appropriate documentation of data structure and values.**

We plan to collect data elements from the following categories:

- General case information;
- Defendant information;
- Charge information;
- Sentencing information;
- Hearing information;
- Attorney information;
- Pretrial release information;
- Diversion and problem solving court information; and
- Warrant information.

What if I am unable to provide all the requested data?

If your agency does not collect one or more data elements, or providing them would be too burdensome or is prohibited under agency policy, then exclude those data elements from your data submission. If you are unable to provide many of the elements, please reach out to your point of contact or the CCSC Help Desk at 855-279-3515.

What happens after we submit the data?

Urban and NCSC will verify the contents of the data file. We may reach out to clarify or confirm any questions about variable formats or values. Typically, this will be within 2-3 weeks of your submission. Urban or NCSC will contact you to review any questions. Having a thorough understanding of your data is vital to constructing a valid CCSC dataset and ensuring that we are properly interpreting and using your data.

Appendix A. Requested Data Elements, Definitions, and Formats

Category	Data Element	Definition	Format	Example Values
General Case Information	Jurisdiction identifier	County name; district name; circuit name	String	County name; district name; circuit name
	Court level	Level of court.	String	District; circuit
	Court case identifier	Anonymized series of characters that identify the court case that follows a case through court levels. This identifier should be assigned to each charge within the same case.	Numeric or string	
	Initiating action	Type of action that initiated the case.	Numeric or String	1=Summons; 2=Arrest; 3=Warrant; 4=Warrantless arrest; 5=Indictment
Defendant Information	Defendant identifier	Anonymized series of characters that identify the same individual within the court system, across cases and courts within the state.	Numeric or string	
	Age at arrest	Individual's age as of arrest/offense date. If age is not available, please provide date of birth.	Numeric	
	Gender or sex	Individual's gender.	Numeric or String	1 = Male; 2 = Female
	Ethnicity	Individual's ethnicity determines whether a person is of Hispanic origin or not.	Numeric or String	1=Spanish/Hispanic/Latino; 0=Not Spanish/Hispanic/Latino
	Race	Individual's self-identification with one or more racial groups. NOTE: Self-identification is much preferred, but in its absence, the perception of race by the criminal justice officer(s) who had the first contact with the defendant can serve as a proxy.	Numeric or String	1=White; 2=African American; 3=American Indian or Alaskan; 4=Asian 5=Pacific Islander or Native Hawaiian; 5=Other 6=Two or more races
	Defendant indigency status	Indicator of whether individual was determined to be indigent at any point during the case.	Numeric or String	1=Yes; 0=No
	Attorney type	Type of counsel at disposition/sentencing.	Numeric or String	1 = private/retained; 2=public defender; 3=assigned/appointed; 4=no counsel/pro se

Category	Data Element	Definition	Format	Example Values
	Failures to appear	Total number of failures to appear in the instant case	Numeric	
Charge Information	Arrest/offense date	Date when the defendant was arrested on this charge. If not available, please provide date on which the offense occurred.	Date	
	Charge identifier	Sequence number uniquely identifying each filing charge within a case (count number). This is not the statute number or case type code.	Numeric or string	
	Charge filing date	Date the charge was filed.	Date	
	Filing charge statute number	The statute number that defines the alleged behavior as a criminal offense.	Numeric or string	
	Filing charge description	Description of the offense at filing.	String	
	Filing charge degree	Charge class severity/degree at filing.	Numeric or String	1=Felony; 2=Gross misdemeanor; 3=Misdemeanor
	Filing charge degree detail	The specific class severity or degree provided by statutory laws for each charge against the defendant.	Numeric or String	1=Felony 1; 2=Felony 2; 3=Misdemeanor A; 4=Misdemeanor B; etc.
	Disposition charge statute number	The state statute number that defines the alleged behavior as a criminal offense for each disposed charge.	Numeric or string	
	Disposition charge description	Description of the offense at disposition.	String	
	Disposition charge degree	Charge class severity/degree at disposition.	Numeric or String	1=Felony; 2=Gross misdemeanor; 3=Misdemeanor
	Disposition charge degree detailed	The specific class severity or degree provided by statutory laws for each disposed charge.	Numeric or String	1=Felony 1; 2=Felony 2; 3=Misdemeanor A; 4=Misdemeanor B; etc.
	Charge disposition date	Date when the charge received a judgment or disposition.	Date	
	Charge sentencing date	Date when the charge received a sentence.	Date	
Charge manner of disposition	The manner in which the charge was disposed.	Numeric or String	1= Jury trial; 2=Bench/non-jury trial; 3=Guilty plea; 4=Entry into a drug court or other problem-solving court; 5=Diversion (not a	

Category	Data Element	Definition	Format	Example Values
				problem-solving court); 6=Dismissal/nolle prosequi; 7= Bindover/transfer; 8=Other
	Charge result of disposition	The judgment or disposition entered by the court for the charge.	Numeric or String	1=Conviction; 2=Acquittal; 3=Mistrial; 4=Entry into a problem-solving court docket; 5=Diversion (not to a problem-solving court docket); 6=Continued without a finding; 7=Dismissed/nolle prosequi; 8==Bindover/transfer; 9=Other
Sentencing Information	Sentence type	Type(s) of sentence(s) the defendant was ordered to serve.	Numeric or String	1=Death penalty; 2=Life in prison with no chance of release; 3=Life in prison with any chance of release; 4=State prison; 5=Jail; 6=Lifetime supervision; 7=Extended supervision/split sentence; 8=Probation; 9=Fine; 10=Restitution; 11=Community service; 12=Time served; 13=Other. This field should include all sentences imposed in the case.
	Sentence length	Length in days of minimum sentence imposed by the court. Includes indeterminate sentences. If sentences are recorded as minimum-maximum, please provide both and indicate the min and max.	Numeric	
	Sentence conditions	Additional conditions imposed at sentencing.	Numeric or String	1=Suspension or revocation of driver's license; 2=Sex offender registration; 3=Firearms eligibility; etc.
	Concurrent/consecutive sentence	Flag to indicate the sentence is to be served concurrently or consecutively.	Numeric	1=Yes; 0=No
	Time served credit length	Length of time (in days) spent in pretrial detention that was credited toward the sentence the defendant was mandated to serve.	Numeric	
	Fine amount ¹	Dollar amount of the fine the defendant was sentenced to pay.	Numeric	Dollar amount.
	Fees/costs amount	Dollar amount of the all fees/costs the defendant was mandated to pay.	Numeric	Dollar amount.

¹ If dollar amounts are not available for fine amount, fees/costs, or restitution, please indicate if any fines, fees, or restitution were ordered in the case.

Category	Data Element	Definition	Format	Example Values
	Restitution amount	Dollar amount of the restitution the defendant was sentenced to pay.	Numeric	Dollar amount.