

**U.S. Department of Education Office for Civil Rights** 

# 2021–22 Civil Rights Data Collection Data File User's Manual

Release 1.0

**January 17, 2025** 

This user's guide was produced under U.S. Department of Education Contract No. 47QTCA18D00JV Task Order No. 91990019F0392. Rosa Olmeda serves as the contracting officer's representative.

## **U.S.** Department of Education

Miguel Cardona Secretary of Education

January 17, 2025

This user's guide is in the public domain. Authorization to reproduce it in whole or in part is granted. Permission to reprint this publication is not necessary. On request, this publication is available in alternate formats, such as Braille, large print, or CD Rom. For more information, please contact the Department's Alternate Format Center at (202) 260–0852 or (202) 260–0818.

# **Document Control**

## **DOCUMENT INFORMATION**

Title:	2021–22 Civil Rights Data Collection Data File User's Manual					
Issue Date:	January 17, 2025					
Security Level:	Unclassified – For Official Use Only					

## **DOCUMENT HISTORY**

Version Number	Date	Summary of Change
1.0	1/17/2025	This is the user manual for the 2021–22 CRDC data file.

# Contents

Content	ts	4
1. Pu	rpose	5
2. Ab	out CRDC Data	5
2.1	CRDC Respondents	6
2.2	Survey Forms and Reporting Categories	6
2.3	New Data Elements for the 2021–22 CRDC	7
2.4	EDFacts Data and CRDC	7
3. Ov	verview of the 2021–22 CRDC Submission Process	7
3.1	Online Data Submission and Response Rate	7
3.2	CRDC Partner Support Center	8
3.3	Addressing Errors	8
3.4	Certification	8
4. Po	st-Collection Overview	9
4.1	Data Quality Outreach and Corrections	9
4.2	Data Quality Corrections and Explanations	10
4.3	Data Quality Assessment and Findings	10
5. Ov	verview of CRDC Data Files and Data Elements	10
5.1	Restricted-Use Data File	10
5.2	Public-Use Data File	11
5.3	Data File Types and Structure	11
5.4	Reserve Codes and System Errors	12
5.5	Data Types	

## 1. Purpose

This manual provides data users with details regarding the 2021–22 Civil Rights Data Collection (CRDC). It includes general information about the CRDC and describes different aspects of the 2021–22 collection, including response rates, data quality concerns, data file details, collection procedures, and privacy protection methodology. This manual also includes an overview of post-submission data corrections, outreach efforts, and subsequent updates to the data files. It is intended to be used in conjunction with both the public-use and restricted-use data files.

## 2. About CRDC Data

The CRDC is a mandatory survey of public local educational agencies (LEAs) and schools in the 50 states, Washington, D.C., and the Commonwealth of Puerto Rico. The CRDC collects information about student enrollment, access to courses, programs and school staff, and school climate factors, such as harassment or bullying and student discipline. Most data collected by the CRDC are disaggregated by race/ethnicity, sex, disability, and English Learner (EL) status.

The collection is an essential part of the <u>U.S. Department of Education (Department) Office for Civil Rights' (OCR)</u> strategy for administering and enforcing civil rights laws in the nation's public LEAs and schools. OCR may use CRDC data as it investigates complaints alleging discrimination, determines whether the federal civil rights laws it enforces have been violated, initiates proactive compliance reviews to focus on particularly acute or nationwide civil rights compliance problems, and provides policy guidance and technical assistance to educational institutions, parents, students, and others.

Section 203(c)(1) of the 1979 Department of Education Organization Act conveys to the Assistant Secretary for Civil Rights the authority to "collect or coordinate the collection of data necessary to ensure compliance with civil rights laws within the jurisdiction of the Office for Civil Rights" (20 U.S.C. § 3413(c)(1)). The civil rights laws enforced by OCR for which the CRDC is used to collect data include:

- Title VI of the Civil Rights Act of 1964, which prohibits discrimination based on race, color, and national origin;
- Title IX of the Education Amendments of 1972, which prohibits discrimination based on sex; and
- Section 504 of the Rehabilitation Act of 1973, which prohibits discrimination on the basis of disability.

OCR's implementation of the regulations for each of these statutes requires recipients of the Department's federal financial assistance to submit to OCR "complete and accurate compliance reports at such times, and in such form and containing such information" as OCR "may determine to be necessary to enable [OCR] to ascertain whether the recipient has complied or is complying" with these laws and implementing regulations (34 C.F.R. § 100.6(b), 34 C.F.R. § 106.81, and 34 C.F.R. § 104.61). In addition, pursuant to a delegation by the Attorney General of the United States, OCR shares in the enforcement of Title II of the Americans with Disabilities Act of 1990, which prohibits discrimination based on disability. Any data collection that OCR has determined to be necessary to ascertain or ensure compliance with these laws is mandatory.

Since the 2011–12 school year, OCR has collected data from all public LEAs and their schools in the 50 states and Washington, D.C. Over time, the CRDC's collection universe has grown to include long-term secure justice facilities, charter schools, alternative schools, and schools that focus primarily on serving students with disabilities. OCR added the Commonwealth of Puerto Rico to the CRDC, beginning with the 2017–18 CRDC. From 1968 to 2010, civil rights data were collected from a sample of public LEAs and their schools, except for the 1976 and 2000 collections, which included data from a universe of LEAs. Prior to the 2004 CRDC, the survey was administered as the Elementary and Secondary School Civil Rights Survey. To access historical civil rights data from 1968 through 1998, visit the CRDC Archival Tool.

## 2.1 CRDC Respondents

For the CRDC, the LEA serves as the primary responding administrative unit.<sup>1</sup> LEAs are responsible for reporting data for themselves and their schools. Some state educational agencies (SEAs) submit some or all data on behalf of their LEAs, but LEAs are required to certify the data. Florida is the only SEA that certifies data for all of its LEAs. Seventeen SEAs provided all or partial data for the 2021–22 CRDC on behalf of their LEAs.<sup>2</sup>

## 2.2 Survey Forms and Reporting Categories

The CRDC is comprised of LEA- and school-level data. Prior to administration of the survey and pursuant to the Paperwork Reduction Act of 1995, OCR's proposed CRDC must undergo two public notice and comment periods—an initial 60-day period followed by a 30-day period—and receive the U.S. Office of Management and Budget's (OMB) approval. To help OCR prepare the initially proposed CRDC that undergoes the 60-day public comment period, OCR gathers information about emerging civil rights issues from program offices, stakeholders, and the public. After the 60-day comment period, OCR reviews the comments received, prepares responses to the comments, and depending on the comments, may or may not make changes to the proposed CRDC. The responses and the proposed CRDC then undergo a 30-day public comment period. After that comment period, OCR considers the comments received, crafts responses to the comments, and revises the proposed CRDC, if deemed appropriate. OCR then submits the responses and the proposed CRDC to OMB, for review and approval. After OMB approves the CRDC, OCR develops the LEA-level and school-level CRDC survey forms.

The LEA-level and school-level 2021–22 CRDC survey forms are available at https://civilrightsdata.ed.gov/data.

In general, for reporting of race/ethnicity data to the Department, recipients of Department financial assistance are required to follow the Department's "Final Guidance on Maintaining, Collecting, and Reporting Racial and Ethnic Data to the U.S. Department of Education" dated October 2007 (72 Fed. Reg. 59,266). The Department's guidance requires recipients to report race data by seven major racial/ethnic

\_\_\_

<sup>&</sup>lt;sup>1</sup> An LEA is "a public board of education or other public authority legally constituted within a State for either administrative control or direction of, or to perform a service function for, public elementary schools or secondary schools in a city, county, township, school district, or other political subdivision of a State, or for a combination of school districts or counties as are recognized in a State as an administrative agency for its public elementary schools or secondary schools." See <u>34 CFR § 303.23</u>.

<sup>&</sup>lt;sup>2</sup> SEAs providing all or partial data for the 2021–22 CRDC include: Arkansas, Colorado, Delaware, Florida, Georgia, Hawaii, Iowa, Kansas, Kentucky, Maine, Massachusetts, North Carolina, Oklahoma, Puerto Rico, Virginia, West Virginia, and Wisconsin.

categories (i.e., American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino of any race, Native Hawaiian or Other Pacific Islander, Two or more races, and White). Under current Department guidance, recipients may choose to disaggregate these categories further to address their own needs at the state level but must categorize race data according to these seven categories for the Department's data collections, such as the CRDC and EDFacts collections. The best way for users to understand the specifics of disaggregation is to review the LEA-level and school-level survey forms and Appendix Sheet B. CRDC Data Dictionary in the accompanying 2021–22 Appendix Workbook.

#### 2.3 New Data Elements for the 2021–22 CRDC

Most of the data elements collected for the 2020-21 CRDC continued to be collected in the 2021-22 CRDC. However, OCR did introduce new data elements for the 2021–22 CRDC and retired others. New data elements introduced for the 2021–22 CRDC were optional for LEAs to report and are only available in the CRDC restricted-use data file.<sup>3</sup> To view the data elements for the 2021–22 CRDC, see <u>Appendix Sheet B. CRDC Data Dictionary</u> in the accompanying 2021–22 Appendix Workbook.

#### 2.4 EDFacts Data and CRDC

EDFacts is a centralized data collection through which SEAs submit pre-K through grade 12 data to the Department to enable better data analysis and data usage in policy development, planning, and management. OCR co-stewards the collection of some EDFacts data files. Theses EDFacts data files co-stewarded by OCR for the 2021-22 CRDC are forthcoming.

## 3. Overview of the 2021–22 CRDC Submission Process

#### 3.1 Online Data Submission and Response Rate

The CRDC's online data submission system opened on December 11, 2023, and closed on February 26, 2024. LEAs submit data into the submission system by manually keying the data directly into the system, uploading data files into the system, or both. LEAs could opt to use either method or a combination of the two methods. All data reported were stored in the same data repository and were visible in the online data entry screens and all system-generated reports. For more details on manual data entry in the submission system and flat file submission see the <u>CRDC Submission System User Manual</u>.

The CRDC online data submission system included a resources page accessible from every screen to provide data submitters with guidance on survey items and tips on avoiding data submission mistakes. These resources were also available outside of the online system via the <u>CRDC Resource Center</u>.

For the 2021–22 CRDC, 6,409 LEAs (36.2%) used the online data entry method only, while 11,295 LEAs (63.8%) used the flat file submission method for at least part of their data entry. OCR achieved 99.9% reporting and certification from required data submitters for the 2021–22 CRDC, which included 17,704 LEAs and 98,010 schools.

<sup>&</sup>lt;sup>3</sup> For more information about the CRDC restricted-use data file, see Section 5.1 of this document.

## 3.2 CRDC Partner Support Center

For the 2021–22 CRDC, the CRDC Partner Support Center (PSC) provided ongoing support to data submitters. During the data submission period, the PSC followed up with uncertified LEAs on a weekly basis to offer assistance and conducted intensive outreach efforts to unresponsive LEAs. After the data submission period ended, the PSC also conducted outreach to LEAs to request data corrections for some data quality errors, when necessary. The PSC responded to telephone calls and e-mails from data submitters who needed technical and, in many cases, substantive assistance. Examples of technical assistance provided include logging into the online submission system, creating additional accounts, explaining the different navigation methods in the online submission system, helping to enter the data manually and via the file upload tool, resolving errors, and helping with certifying data. Examples of substantive assistance provided include explaining survey questions, definitions, and other topical content related to the CRDC. In addition to fielding data submitter questions, the PSC conducted regular outreach, including recurring e-mails to communicate important dates and helpful tips about the collection process.

## 3.3 Addressing Errors

OCR implements real-time, automatic data quality checks in the submission system to help data submitters correct potential errors prior to completing the data submission. For the CRDC, these checks are commonly referred to as "business rules" and are categorized as either errors or warnings. Errors require data submitters to correct or, in some cases, submit a justification or "reason code" to certify their data. Reason codes allow submitters to explain why the value submitted was valid. Warnings remind data submitters to review their data submissions but do not require data corrections. In the submission system, data submitters could view errors and warnings on a summary page that displayed all data errors, warnings, and pending data quality checks. Data submitters had the opportunity to correct errors in the submission system prior to certifying the data. For a complete list of business rules for the 2021-22 collection, see Appendix Sheet G. Submission System BRs in the accompanying 2021-22 Appendix Workbook. If an error could not be corrected, then the data submitter was instructed to enter a reason and/or comment, or in the case of missing values, provide an action plan for reporting the required data in the next collection.<sup>4</sup>

## 3.4 Certification

To complete the submission of data for the CRDC, either the LEA superintendent or another person designated by the LEA superintendent is required to certify the data. Certification attests that the data are "true and correct" to the best of the certifier's knowledge.

LEAs must submit all required data and resolve errors before certification. Submissions that do not meet these criteria cannot be certified. Data submitters that had trouble resolving errors received help from the PSC.

<sup>&</sup>lt;sup>4</sup> Select LEAs had the opportunity to correct their data after the close of the submission system during post-collection data quality outreach. See <u>Section 4.1 Data Quality Outreach and Corrections</u>.

## 3.4.1 Force Certification

For the 2021–22 CRDC, some data submissions required force certification. Force certification occurs when a data submitter is unable to resolve a submission system business rule error. When this occurs, the PSC manually overrides the error(s), allowing the LEA to certify. For a complete list of force certified LEAs, see Appendix Sheet F. Force Certification in the accompanying 2021–22 Appendix Workbook.

#### 3.4.2 Certification with an Action Plan

In cases where an LEA is unable to report data, OCR requires the LEA to submit an action plan. In an action plan for the CRDC, LEAs are required to describe the specific steps and timeline that they will follow to ensure that the data are collected for future CRDCs. To submit an action plan, LEAs had to first contact the PSC, and then once the plan was approved by OCR, the LEA could certify its data. For the 2021–22 CRDC, 95 LEAs submitted action plans. See <u>Appendix Sheet D. Action Plans</u> in the accompanying 2021-22 Appendix Workbook for a complete list of LEAs that submitted action plans for the 2021–22 collection.

#### 3.4.3 Certification with a Quick Plan

Quick plans are shorter versions of CRDC action plans and allow LEAs to provide an explanation for missing data and an abbreviated plan of action for collecting and submitting the data for the next CRDC. For the 2021–22 CRDC, quick plans were limited to the following topic areas: allegations of harassment or bullying, mechanical restraint, physical restraint, seclusion, referrals to law enforcement, and school-related arrests. LEAs were not required to contact the PSC for assistance prior to creating and submitting a quick plan, and OCR did not review the quick plan prior to certification. See <a href="Appendix Sheet E. Quick Plans">Appendix Sheet E. Quick Plans</a> in the accompanying 2021–21 Appendix Workbook for a complete list of LEAs that submitted quick plans for the 2021–22 collection.

## 4. Post-Collection Overview

## 4.1 Data Quality Outreach and Corrections

After the close of the 2021–22 CRDC, OCR performed a series of post-submission data quality checks to identify issues in data submissions. These checks focused on issues of internal consistency and overall data reasonableness (e.g., duplicate data, summation to totals, zeros, and comparison with other data sources). A total of 297 post-submission data quality checks were performed to identify these issues. These checks comprise 119 post-submission data quality checks (or "ad hoc" checks) and 178 data validation checks. OCR selected 20 of the ad hoc data quality checks to include in the post-submission data quality outreach and corrections period. Additionally, LEAs flagged with these data quality checks were contacted regarding any data validation checks they were identified with. In total, the corrections process comprised 195 data quality checks. For a complete list of the data quality checks included in the post-submission data quality outreach and corrections period, see Appendix Sheet H. Data Quality Checks in Outreach in the accompanying 2021–22 Appendix Workbook. From June 12, 2024, to July 8, 2024, LEAs

\_\_\_

<sup>&</sup>lt;sup>5</sup> Data validation checks were introduced for the 2021-22 CRDC. While similar to the general data quality checks, data validation checks monitor for broader data quality concerns across all modules, including 1) Zero Values, 2) Low Values, 3) High Values, 4) Snapshot Enrollment Comparisons, 5) Related Elements, and 6) Duplicate Reports. In contrast, ad hoc data quality checks examine narrow and topical data quality issues.

identified as failing one or more of these data quality checks were given the opportunity to review and correct their data in the online submission system. In total, 7,128 LEAs were selected for this data quality outreach. Of these LEAs, 5,941 (83%) responded to the corrections process.<sup>6</sup>

## 4.2 Data Quality Corrections and Explanations

LEAs that participated in data quality outreach submitted data corrections, submitted a reason code or explanation about their reported data, or both. About 51.1% (3,639/7,128) of responding LEAs corrected a portion of their submitted data during the outreach period, and 69.5% (4,954/7,128) of responding LEAs submitted at least one reason code or explanation<sup>7</sup> for issues included in outreach. LEAs submitting reason codes or explanations indicated that they: corrected their data; were unable to correct their data; or, believed their data were correct as originally reported.

If an LEA indicated that they were unable to correct their data, they were prompted to provide an additional written explanation. These written explanations were categorized by OCR using the following categories: 1) LEA is unable to verify, 2) LEA did not understand analysis or process, 3) school or LEA type is justification for data, or 4) other.

Of the LEAs that provided a reason code or explanation, the majority indicated their data were correct as reported. Approximately 32% (2,266/7,128) of LEAs submitting reason codes or explanations noted that their data were correct as reported. Some of their data were incorrect and corrected the issue. Less than 10% (663/7,128) of LEAs submitting reason codes or explanations indicated that they could not correct at least some of their data. In these cases, LEAs were allowed to submit written explanations. For more details on reason codes and written explanations submitted by LEAs during the data quality outreach and corrections period, see <u>Appendix Sheet I. Explanations from LEAs</u> in the accompanying 2021–22 Appendix Workbook.

## 4.3 Data Quality Assessment and Findings

After the data quality corrections and outreach period closed, data were reviewed for quality. Results of the data quality assessment for the 2021-22 CRDC are forthcoming.

## 5. Overview of CRDC Data Files and Data Elements

For the 2021–22 CRDC, OCR released two data files: a restricted-use data file and a public-use data file.

## 5.1 Restricted-Use Data File

The restricted-use data file is the primary analysis and investigative file primarily used internally by OCR. This file is the basis for enforcement research, as well as for internal program and policy analysis. In addition, the restricted-use data do not undergo any data quality suppression and privacy protections are

<sup>&</sup>lt;sup>6</sup> Responding LEAs submitted data corrections, submitted an explanation about their reported data, or both.

<sup>&</sup>lt;sup>7</sup>A reason code is an LEA-provided explanation during the corrections period, in response to an identified data quality issue. Since reason codes and explanations are self-reported, the response indicated by the LEA may not reflect the status of their data. For example, if an LEA uses a reason code to state they corrected data, this does not necessarily mean that the data have been corrected.

not applied. The data in this file are minimally edited to preserve the original, certified data as submitted. All CRDC data are included in the file, including both required and optional data. <u>Appendix Sheet B. CRDC Data Dictionary</u> in the accompanying 2021-22 Appendix Workbook lists all data elements, both required and optional, for the 2021–22 CRDC.

The Department's Institute for Education Sciences is responsible for granting access to the restricted-use data file. For information on how to request access to the restricted-use data file, see <a href="Restricted Use Data">Restricted Use Data</a> Licenses.

#### 5.2 Public-Use Data File

The public-use data file is intended for public-use. It is released to maximize the use of statistical information while protecting the disclosure of student information. The data file only includes required data elements for the 2021–22 CRDC. Some of the data have been perturbed to protect the identities of the students for whom these data were reported. For more information on data included in the public-use data file, see <a href="Appendix Sheet B. CRDC Data Dictionary">Appendix Sheet B. CRDC Data Dictionary</a> in the accompanying 2021-22 Appendix Workbook.

#### 5.2.1 Privacy Protection

To prevent the disclosure of identifying information, student counts in the public-use data file were privacy protected by making small, random adjustments to the data. This process, also called perturbing, used a low-frequency perturbation routine. The methodology added or subtracted one case to blur the data and used random data swapping. The routine protected true zeros, except for outcome data (e.g., Algebra passing), for which zeros were included in the perturbation routine. The routine was applied to all student count data elements. It is important to note that the perturbation routine aggregates to the specified change thresholds at the state and national levels, but there is no threshold for aggregation at the LEA-level. Caution should be taken when making comparisons between LEA- and school-level form information as these may not align due to perturbation methods.

In addition, data for some data elements were adjusted or suppressed<sup>8</sup> to minimize inconsistencies introduced by data perturbations. However, some inconsistencies may remain. Finally, the margin of difference between perturbed and unperturbed counts varied by the size of the data sample being perturbed.

#### 5.3 Data File Types and Structure

The CRDC collects data at two levels—the LEA-level and the school-level—using two different survey forms, referred to as the LEA Form and the School Form, respectively. Each form contains survey questions relevant to the reporting level. Copies of the 2021-22 CRDC survey forms are available at the following links:

LEA Form: <u>2021–22 Civil Rights Data Collection – -LEA Form (PDF) (ed.gov)</u>

<sup>&</sup>lt;sup>8</sup> Data suppressed for privacy protections are indicated with a "-12" reserve code. See 5.4 Reserve Codes and System Errors for more information.

School Form: <u>2021–22 Civil Rights Data Collection – School Form (PDF) (ed.gov)</u>

There are 34 topical data files for the 2021–22 CRDC—3 LEA-level files and 31 school-level files. These files correspond to the topical modules in which the 2021–22 CRDC data were reported.

For more information on how the restricted- and public-use data files are structured, see <u>Appendix Sheet</u> A. CRDC File Structure in the accompanying 2021-22 Appendix Workbook.

## 5.4 Reserve Codes and System Errors

In some instances, special or reserve codes were assigned to the data. Reserve codes provide information on variables that do not have reported values, or the reported value was removed by OCR. For the CRDC, reserve codes are standardized and do not represent an actual reported count or value.

The table below lists the reserve codes and their definitions for the 2021–22 CRDC.

Table 1. Reserve Code Values

Reserve Code Value	Definition				
-3	Skip Logic or Processing Failure				
-4	Missing Optional Data <sup>10</sup>				
-5	Action Plan/Quick Plans				
-6	Force Certified				
-9	Not Applicable/Skipped				
-12	Suppressed for Privacy Protections				
-13	Missing DIND skip logic				

For required data, missing values are represented with a -5 (action plan or quick plan) or -6 (force certified). Virtual schools or schools that offered virtual instruction only were skipped for several content areas, including restraint and seclusion and corporal punishment. For a complete list of content areas where this skip logic is applied, see the 2021-22 school and LEA forms.<sup>11</sup>

#### 5.4.1 System Submission Errors and Other Data Anomalies

There were instances when the CRDC data submission system did not function as expected. In cases where these issues could not be rectified due to system limitations, the data are assigned -3 in the CRDC data file. A select summary of those issues follows.

<sup>&</sup>lt;sup>9</sup> The restricted-use file structure is 36 topical data files, with 3 LEA-level files and 33 school-level files.

<sup>&</sup>lt;sup>10</sup> The "-4" reserve code is only in the restricted-use data file because it only applies to optional data elements.

<sup>&</sup>lt;sup>11</sup> The 2021-22 LEA Form is available at the following link: <u>2021–22 Civil Rights Data Collection – -LEA Form (PDF) (ed.gov)</u> and the 2021-22 School Form is available at the following link: <u>2021-22 Civil Rights Data Collection - School Form (PDF) (ed.gov)</u>

#### 5.4.1.1 Skip Logic Failure

There were instances where the skip logic failed to update related data fields. This functionality failure allowed data elements to require data entry when, according to their skip flag, those elements should have been skipped. The skip logic failure also allowed LEAs to skip data entry, even when their skip flag showed that those elements were required. The failure in the skip logic occurred due to intense system load, which may have caused the system to slow down and not update skip logic flags effectively. Values for data elements that had a skip logic failure were represented with a -3 reserve code.

#### 5.4.1.2 Autozero Functionality

There were instances where the autozero functionality in multiple modules did not function as expected for designated data fields. This functionality is meant to automatically apply zeros to data elements where missing values are not allowed and real numbers are required to be entered. In cases where the autozero function failed, zeros were incorrectly applied across data fields during data file processing. To address this, data were reverted to their original format and autozeroing was re-applied.

#### 5.4.1.3 ERRF215

The business rule ERRF215 monitors for when a submitter enters "NO" or all-zero counts for the Internet Access and Devices module. This rule was implemented as an error in the submission system for the 2021-22 collection but was reverted to a warning during the collection to ease burden.

#### 5.4.1.4 Referrals to Law Enforcement & School-related Arrests

In July 2024, the Government Accountability Office (GAO) released a report that reviewed school police and arrest data in the nation's K-12 schools. <sup>12</sup> It is important to note that the revisions made to the 2021-22 school form and associated definitions may have impacted reporting in the Referrals to Law Enforcement & School-related Arrests (ARRS) module. Therefore, data users should exercise caution when accessing these data, as they may not be directly comparable to previous collections. For more information on the revisions made to definitions for the 2021-22 collection, see the definitions crosswalk, LEA and School Crosswalk of Definitions and their Historical Changes.

#### 5.5 Data Types

There are three types of variables in the CRDC data files: reported data, calculated totals, and identifiers.

#### 5.5.1 Reported Data

The majority of data in the CRDC data files were reported by the LEA directly into the CRDC submission system. The CRDC data submission system auto-calculated values based on the reported data. All system-calculated data were visible to the LEA and could be changed.

#### 5.5.2 Calculated Totals

Calculated totals are values that were computed after the submission system closed and were subsequently added to the data files. Initially, totals were calculated by summing disaggregated data (e.g.,

<sup>&</sup>lt;sup>12</sup> The GAO report can be found here: <u>K-12 Education</u>: <u>Differences in Student Arrest Rates Widen when Race, Gender, and Disability Status Overlap.</u>

White female enrollment) greater than or equal to zero and by treating reserve code values as zeros. In instances where all disaggregated values had a reserve code, the calculated total was assigned the most negative reserve code reported in the group of disaggregated values. See the table below for an example of how calculated totals were computed in the CRDC data file.

Table 2. Calculated Totals Example

School	AM_F	AS_F	BL_F	HI_F	HP_F	TR_F	WH_F	LEP_F	IDEA_F	504_F	TOT_XX_F
Α	2	1	2	4	2	3	2	2	3	3	24
В	-5	-5	-6	-6	-5	-5	-5	-5	-5	-5	-6

LEAs are unable to view or change calculated totals. Some examples of calculated totals include: total overall enrollment by sex; total number of students who passed Algebra I in grade 7 or 8 by sex; and total number of single-sex interscholastic athletics teams. Calculated totals and variable names are shown in <u>Appendix Sheet C. Calculated Totals</u> in the accompanying 2021-22 Appendix Workbook.

#### 5.5.3 LEA and School Identifiers

The table below describes the three main identifiers (IDs) used in the data files.

Table 3. Data File Identifiers

Identifier	Definition
LEA ID	7-digit LEA identification code
SCH ID	5-digit school identification code
СОМВОКЕУ	Combined LEA ID and SCH ID, which often matches the Department's National Center for Education Statistics (NCES) ID

Generally, OCR utilizes the Department's National Center for Education Statistics (NCES) school and LEA IDs to identify schools and LEAs. However, there are cases where the CRDC and NCES IDs differ due to differences in definitions and collection procedures. For more information about these differences, and to link the CRDC to other Department data collections, read the NCES blog, "Building Bridges: Increasing the Power of the Civil Rights Data Collection (CRDC) Through Data Linking With an ID Crosswalk". LEAs that did not have an existing NCES ID were issued a new CRDC LEA ID, which consisted of the 2-digit state federal information processing series code followed by a 5-digit ID assigned by the PSC. Schools that did not have an existing NCES ID were issued a new ID, which consisted of the CRDC LEA ID and a temporary 5-digit school ID assigned by the PSC. A crosswalk to link the CRDC to other Department data collections is available in the accompanying Appendix Workbook under Appendix Sheet J. Crosswalk Information and Appendix Sheet K. CRDC and EDFacts Crosswalk.