# CCD NONFISCAL UNIVERSE DATA FILE CONTENT AND USAGE NOTES

Contents

[CCD NONFISCAL UNIVERSE DATA FILE CONTENT AND USAGE NOTES 1](#_Toc65649448)

[File Structure and Naming Conventions 1](#_Toc65649449)

[Accessing CCD Data 1](#_Toc65649450)

[Introduction to Long Files 1](#_Toc65649451)

[Data File Naming Convention 2](#_Toc65649452)

[File Content 3](#_Toc65649453)

[Unit Identifiers 4](#_Toc65649454)

[Data Usage Notes 5](#_Toc65649455)

[Guidelines for Aggregating CCD Data 5](#_Toc65649456)

[Data Value Exceptions 5](#_Toc65649457)

[Location vs. Jurisdiction 6](#_Toc65649458)

[New York City Public School District 6](#_Toc65649459)

[Bureau of Indian Education Schools 6](#_Toc65649460)

[File Content 7](#_Toc65649461)

[Membership 7](#_Toc65649462)

[Staff 8](#_Toc65649463)

[National School Lunch Program (NSLP) eligibility – Free/Reduced-Price Lunch and Direct Certification 9](#_Toc65649464)

# CCD NONFISCAL UNIVERSE DATA FILE CONTENT AND USAGE NOTES

## File Structure and Naming Conventions

### Accessing CCD Data

Starting school year (SY) 2016-17, users interested in using CCD nonfiscal universe data files have the following options:

* **CSV and SAS files** – Prior to SY 2016-17, CCD data files were published as “wide” files in which each record represented a single entity (i.e., a state, a local education agency, or a school). This changed starting with the SY 2016-17 collection. The Directory and CCD School files are still published as wide files. The metric files (those with staff or student counts), are now published as “long” files in which each record contains a single data point and its unique set of qualifiers. The long files are the formal release; each data file has a corresponding companion file documenting the record layout and summary statistics for the variables in the data file.
* **Other file support**
  + SPSS – NCES provides a read-in utility for SPSS users. This utility is available for all CCD files and includes syntax to create fully labeled and formatted SPSS files from .csv source data.
  + Legacy file format– NCES also provides SAS programs that will convert the long files to wide files. The SAS programs can be found in the CCD Reference Library: <https://nces.ed.gov/ccd/reference_library.asp>. The wide files follow the format of CCD releases prior to SY 2014-15 which had a single file at each level (state, LEA and school). Data from all components (directory, membership, staff, etc.) are combined into a single file with a single record for each entity. The provided SAS programs are based on the 2016–17 CCD data files and they should be used as a baseline for following years’ data files. The programs should be adjusted as changes occur to CCD data files.
* **Table Generator** – The table generator function in CCD’s online Elementary/Secondary Information System (ElSi - <https://nces.ed.gov/ccd/elsi>/) enables users to select data from across all the years, levels and components of CCD without having to manipulate files or do any programming. The selected data can be downloaded as CSV or MS Excel files. Refer to the NCES blog post [Accessing the Common Core of Data (CCD)](../../../../Forms/AllItems.aspx?RootFolder=%2Fed%2FEDMITS%2FShared%20Documents%2F09%20%2D%20Change%20Management%2F01%20ECRB%2F01%20In%20Process&FolderCTID=0x012000425CAF913917DF43A883566EFE403D7C&View=%7BA8502C75%2D55D9%2D40F7%2DB672%2D7499FFD326CF%7D) for tips on using this tool.

### Introduction to Long Files

CCD nonfiscal universe data files for reporting years through SY 2015-16 were formatted as “wide” files where each record corresponded to a single entity (a state, an LEA or a school) and all data elements for that entity were included within that record. From SY 2016-17 forward, the CCD data files are released as long files. The “long” file format includes multiple records for each entity, each providing a separate data value, each with a set of qualifiers. Each of these individual records may be a detail count or a subtotal. For example, in the long file format, the detail SEA-level record for the number of third-grade Hispanic girls in Oregon would comprise the student count plus the required qualifiers (e.g., state ID, grade, race/ethnicity and sex). The membership file also includes subtotals by race/ethnicity and sex and totals for the reporting unit.

The qualifiers of a detail data record constitute the “category set.” Each category has a limited number of “permitted values” or “category codes”; male and female for the category SEX, for example. Multiple category sets may be required for a single ED*Facts* file. The category sets within an ED*Facts* file are designated by a letter (e.g., “Category Set A”). The long files have a column for each group of permitted values that make up a category set. The membership files, for example, have one column each for grade, race/ethnicity, and sex. Refer to the companion files to see the category sets and allowable values for specific CCD nonfiscal data files.

The column TOTAL\_INDICATOR distinguishes between detail records and the subtotals or totals on the file. This column may specify a category set (e.g., “Category Set A – By Race/Ethnicity; Sex; Grade”); a reported subtotal (e.g. “Subtotal 4 – By Grade”); a derived subtotal (e.g., “Derived - Subtotal by Race/Ethnicity and Sex minus Adult Education Count”); or the total for the reporting unit (“Education Unit Total”).

Please note, the education unit total in the membership file includes counts for Adult Education. For the count of elementary/secondary students excluding Adult Education, use the record where TOTAL\_INDICATOR = ‘Derived - Education Unit Total minus Adult Education Count”; this is count comparable to the value MEMBER reported in previous years of the CCD.

In subtotal and total records, the value “No Category Codes” will appear in the column for the category not specified in the subtotal/total. The subtotals by race/ethnicity and sex, for example, show “No Category Codes” for “Grade”.

A permitted value of “Not Specified” is also possible for categorical variables. This identifies an unexplained discrepancy between a subtotal and total. When present, the numeric value in these records is a derived value: the difference between the sum of reported values within a category set and the reported total for a category set. For example, the ED*Facts* category set SEX has the permitted values of male and female. A state that does not count all students in those two categories may report a total for the category set that is greater than the sum of the two categories. The difference is calculated and reported as “Not Specified.”

Negative values are not used in the long file format to indicate exception values (missing, not applicable, or suppressed). Instead, these values appear as nulls.

### Data File Naming Convention

The CCD nonfiscal universe data files are available on the CCD web site: https://nces.ed.gov/ccd/ccddata.asp. The file naming convention is as follows:

**CCD\_SCH\_###\_xxyy\_f\_vv\_mmddyyyy**

* CCD is constant
* SCH indicates reporting level:
  + SCH – school
  + LEA – local education agency
  + SEA – state education agency
* ### is the three or six digit ED*Facts* file specification number:
  + 029 – Directory
  + 129 – CCD School (School only)
  + 052 – Membership
  + 059 – Staff
  + 033 – Free and Reduced Price Lunch (School only)
  + 141 – English Learners count (LEA only)[[1]](#footnote-1)
  + 002089 – Children with Disabilities (LEA only)
* xxyy is the school year (e.g. 1617 for the SY 2016-17)
* ‘f’ indicates the file format (‘w’ for ‘wide’; ‘l’ for ‘long’).
* vv indicates the version.
* mmddyyyy indicates the date file was extracted from DMS.

## File Content

The CCD nonfiscal universe data files are organized by topic area and released at three levels: school, LEA, and state. The table below provides an overview of the topic areas, the level at which the file is released, and a brief summary of the file content and records included. Finally, the table provides users with the ED*Facts* file numbers under which states submit the data.

| **CCD Nonfiscal Files** | **File Level(s)** | | | **File Content Overview** | **Records included** | **ED*Facts* File** |
| --- | --- | --- | --- | --- | --- | --- |
| SEA | LEA | School |
| Directory | X | X | X | * Unique identifiers * Contact information * Operational status * Type and Charter status * Information on grades offered | All operational and non-operational entities | FS029  FS039 |
| School Characteristics |  |  | X | Data on school programs including Title I, Magnet and National School Lunch Program (NSLP) | Operational entities only | FS129 |
| Membership | X | X | X | Data disaggregated by grade, sex, and race/ethnicity | Operational entities with grades that expect membership only | FS052 |
| Staff | X | X | X | SEA- and LEA-level files include all FTE staff (teachers and staff by professional category and level)  School files include FTE teachers only | Operational entities only | FS059 |
| Free and reduced-price lunch |  |  | X | Counts of students eligible for   * Free and reduced-price lunches * Directly certified as eligible for free lunches under NSLP | Operational entities only | FS033 |
| English Learners[[2]](#footnote-2) |  | X |  | Counts of English learners | Operational entities only | FS141 |
| IDEA |  | X |  | Data for school-age and early childhood special education students | Operational entities only | FS002  FS089 |

To facilitate correct use of the data files, CCD data users should refer to the following additional materials for each data file and year combination:

* **Release notes**. Each CCD public file release includes a brief release notes document. This document highlights any data changes from the prior year and shows directory frequencies for the current year alongside similar frequencies for several prior years. For 1a releases, the release notes also include the US and Outlying row from the summary tables released with the files.
* **Companion files**. All CCD public file releases include a set of detailed companion files. Each companion file includes the file layout, permitted values for categorical variables, and data frequencies. The links to the companion files are co-located with the CCD public release files.
* **State data notes**. These files are also part of each public release and provide additional contextual information including state responses to questions about data quality. The links to the state data notes are co-located with the CCD public release files.
* **ED*Facts* file specifications**. These are the documents ED*Facts* provides to states as guidance as they prepare their ED*Facts* files for submission. The definitions and guidance in these files help inform use of the resulting CCD public data files. File specifications for all ED*Facts* files and years are available here: <https://www2.ed.gov/about/inits/ed/edfacts/file-specifications.html>

## Unit Identifiers

CCD data users commonly merge multiple CCD data files, pulling together data across years, topic area, and/or file level. The description of the unique unit identifiers will help users complete these data merges successfully.

There are two types of unit identifiers in the school and LEA data files: 1) unit identifiers assigned by the U.S. Department of Education, and 2) unit identifiers assigned by the reporting SEAs.

* **Identifiers Assigned by the U.S. Department of Education**. Refer to the NCES blog post [Accessing the Common Core of Data (CCD)](https://nces.ed.gov/blogs/nces/2021/01/04/default) for a detailed explanation of this identifier
  + Note that NCES LEAIDs and school IDs do not normally change from year to year. There have been exceptional instances when changes to ID fields were made. These changes were due to a processing anomaly or to correct cases where SCHID was not unique within a state and effected a limited number of schools. The CCD reference library ([https://nces.ed.gov/ccd/reference\_library.asp](https://nces.ed.gov/ccd/reference_library.asp%20)) includes crosswalks of the ID changes so that these schools can be tracked longitudinally.
* **Identifiers Assigned by the State Education Agencies (SEAs)**. SEAs assign unit identifiers to their schools and LEAs for their own data tracking and management systems that are different than the unit identifiers assigned to their schools and LEAs by the U.S. Department of Education. The CCD nonfiscal universe data file publishes these state-assigned unit identifiers on the School Universe and LEA Universe files as the variables ST\_SCHID and ST\_LEAID. Data users should note that these unit identifiers may not be unique across SEAs and that the rules that govern the assignment, usage, and formatting of these identifiers may vary from SEA to SEA and even from LEA to LEA within an SEA.

## Data Usage Notes

### Guidelines for Aggregating CCD Data

For many analyses, it is useful to aggregate data from lower levels to higher levels. However, higher level aggregations can be misleading when data at lower levels are missing. NCES strongly encourages researchers to consider the possible effect of missing or suppressed data on their analyses.

The percentage of cases (schools or LEAs) reporting valid values can be used to evaluate the reliability of the aggregated number. The denominator in this percentage is the number of cases in the aggregation excluding those cases not expected to report the item (i.e., cases with a value flagged ‘Not Applicable’); the numerator is the number of those schools providing valid data (values greater than or equal to zero). The differences between them are cases of missing or suppressed values.

NCES suggests the following guidelines for interpreting this percentage.

* If the percentage is greater than 95 percent, the aggregation can be considered reliable.
* If the percentage is less than 95 percent, but more than 85 percent, the aggregation should be interpreted with caution.
* If the percentage is less than 85 percent, the aggregation should be considered unreliable.

### Data Value Exceptions

**Zero Counts:** Zero counts are valid data values; they do not indicate missing, not-applicable, or suppressed values. A zero count shows that a data element was expected and reported, and there were no instances of this data element. For example, a school that currently offers education services to students in grade 12 but currently does not enroll any grade 12 students would report a grade 12 enrollment/membership count equal to zero (0).

**Missing Data**: Missing data occur when a value for a data element was expected but not reported. For example, a school that enrolls students in grade 12 but cannot or did not report the number of 12th-graders in the school has missing data for grade 12 enrollment/membership. In CCD data files, missing numeric data appear as null with DMS\_FLAG set to “Missing.” Missing categorical data are coded as “Missing.”

**Not Applicable Data:** A data element is “not applicable” when a value for a data element was not expected, and the state did not report a value for that element. For example, an elementary school with a high grade of 5 would not have applicable grade 12 enrollment/membership. In CCD data files, not-applicable numeric data appear as null with DMS\_FLAG set to “Not applicable.” Not applicable categorical data are coded as “Not applicable.”

**Suppressed Data:** LEA- and school-level data values that have been suppressed appear as null with DMS\_FLAG set to “Suppressed.” Prior to SY 2016-17, suppressed values were set to “-9.”

### Location vs. Jurisdiction

* In the CCD nonfiscal universe data files, schools and LEAs are included under the jurisdiction responsible for their administration, which is not necessarily the state where the LEA or school is located. Some examples of this situation include
* SEAs that operate schools in a neighboring state;
* charter school operators/administrators based in one state that operate schools in one or more states other than the state in which they are based. These entities are often used as the LEA for the charter school they operate or administer; and
* schools administered by the Bureau of Indian Education (BIE) and the Department of Defense Education Activity (DoDEA) that are located in various states across the country.

In these cases, the state abbreviations used in the physical location or the mailing address will not necessarily reflect the SEA responsible for administering the school or LEA. The first two digits of the LEA and school ID numbers are the state code (FIPST) for the responsible state-level jurisdiction. FIPST (or STATENAME) should be used to group schools and LEAs by their administrative state. BIE and DoDEA have been assigned their own codes in CCD for this purpose.

### New York City Public School District

Prior to SY 2005–06, New York City (NYC) schools were reported as part of a single, regular LEA. In SY 2005–06, New York began reporting NYC schools as part of a supervisory union with 33 member LEAs. NCES continues to report NYC schools as a single LEA. However, on the CCD nonfiscal universe data files, they appear in the supervisory union configuration. NCES aggregates data from the 33 component LEAs with the single supervisory union to provide estimates for the NYC Public Schools. This aggregation is done by selecting schools or LEAs with a FIPST code = 36 (for the state of New York) and a supervisory union identification number (UNION) = 300.

### Bureau of Indian Education Schools

The BIE supports elementary and secondary schools and dormitories located on over 50 reservations in numerous states, including off-reservation boarding schools and schools directly controlled by tribes and tribal school boards under contracts or grants with the BIE. The bureau also funds residential programs for students at boarding schools and at dormitories housing those attending nearby tribal or public schools.

Some state governments provide some support to BIE schools within their borders. To account for this funding, the state governments may include a BIE school in their ED*Facts*/CCD reporting. In a limited number of cases, the state also reports the membership for these schools, creating the possibility of double-counting these students in CCD nonfiscal universe data. There are a small number of these schools and LEAs (less than 20 schools and 10 LEAs) in three states. In prior years of CCD, flags were included on the school and LEA files to identify these schools. In order to make the DMS file generation process more efficient, these flags were dropped with the 2016-17 release. A list of these schools and LEAs is available in the CCD reference library ([https://nces.ed.gov/ccd/reference\_library.asp](https://nces.ed.gov/ccd/reference_library.asp%20)s).

## File Content

### Membership

**Head Counts**: All student membership counts in CCD are head counts, not FTEs. A student may attend more than one school, but each student is counted only once, in the school where he/she spends most of the school day, sometimes called the “home school” or “school of record.” A student is counted in the membership of the SEA responsible for the student and in the LEA and school that he/she actually attends. Some LEAs ‘tuition-out’ students to schools in other LEAs or even private schools. These students are still counted in the membership of the LEA responsible for them, but they are not reported in membership of any of that LEA’s schools. For this reason, school membership aggregated to the LEA level does not always equal the membership reported by the LEA.

**Ungraded:** “Ungraded” is included as a grade-level designation in CCD; however, not all states use this designation. For these states, counts of ungraded students and ungraded teachers are reported as “not applicable.”

**Adult Education:** Some public elementary/secondary schools also provide adult education, which may engage the elementary/secondary teachers in those schools. Generally, adult education (AE) is outside the scope of the CCD. In order to provide data users with all the data relevant to public schools’ functions, information about AE is included in the CCD. This information includes a grades-offered flag for AE and counts of AE enrollment. AE is not considered in assigning GSHI and GSLO values or a school’s LEVEL value. In the long files (SY 2016-17 and later), use the field “Total Indicator Description” to use the subtotal most appropriate for the research approach. In SY 2015-16 and prior AE enrollment is not included in the membership totals by race/ethnicity or in the MEMBER variable, which is the total enrollment for elementary/secondary grades only. A separate enrollment variable (TOTAL) includes AE membership.

**Shared-Time Reporting**: A “shared-time” school is a school with students who attend for only part of the school day. Because students are counted only in the membership of the school where they spend most of the school day, they may not be counted in the membership of the shared-time school. For example, a student may attend a regular high school for most of the day and a career/technical (CTE) high school part time. That student is typically counted in the membership of the regular high school, not the CTE high school. The shared-time flag on the CCD school file identifies schools with part-time students. The membership of shared-time school understates the number of students the school actually serves.

**Grades offered vs. enrollment by grade:** The grades-offered flags and the low-grade and high-grade variables derived from them reflect the grades that characterize the LEA or school. A school or LEA may offer a grade but not have any students enrolled in that grade in a given year, reporting a zero count for that grade. Conversely, a school or LEA may report a small number of students enrolled in a grade it does not offer; these are typically small or isolated schools and LEAs that are accommodating students for whom it would be a hardship to attend a school or LEA that formally offers the grade.

The lowest grade offered flag is only set to “prekindergarten” when a school or LEA also offers kindergarten and grade 1, or when prekindergarten and/or kindergarten are the only grades offered. Otherwise, the lowest grade offered is the next lowest grade above kindergarten. This is to avoid mischaracterizing, for example, a high school that offers prekindergarten but no other elementary grades.

**Differences in Cross-Level Totals:** A student is counted in the membership of the SEA responsible for the student and in the LEA and school that he/she actually attends. An LEA or a state may “tuition-out” a student to another LEA, state, or even a private school. For this reason, the school membership aggregated to the LEA level may not equal LEA membership, and school membership and LEA membership aggregated to the state level may not equal state-reported membership.

**Changes in race/ethnicity reporting:** Prior to SY 2008-09, NCES asked states to report student enrollment counts by five racial/ethnic categories. Starting in SY 2008-09, two additional race/ethnic categories were added (Hawaiian/Pacific Islander and “two or more races”). For reporting in SY 2008-09 and SY 2009-10, SEAs had the option to use either the 5-category reporting or the 7-category reporting. A flag was added to the data file to indicate which category set was used. Since SY 2010-11, all reporting SEAs have reported enrollment counts using all seven categories.

For additional information on the Department of Education’s policies on reporting race and ethnicity, please see these web pages: [https://www2.ed.gov/policy/rschstat/guid/raceethnicity/index.html](https://www2.ed.gov/policy/rschstat/guid/raceethnicity/index.html%20) & <https://www2.ed.gov/policy/rschstat/guid/raceethnicity/questions.html>.

### Staff

The CCD nonfiscal universe data contain FTE staff counts of teachers and other staff. School-level data only include an FTE for teachers. The LEA- and SEA-level data include FTE counts of teachers by instructional level and other staff categories as follows:

| **TOTAL\_INDICATOR= Category Set A** | **Permitted Values for Submission** | **TOTAL\_INDICATOR= Derived - Major Staffing Category** |
| --- | --- | --- |
| Elementary School Counselors | ELMGUI | Guidance Counselors |
| Secondary School Counselors | SECGUI |
| School Counselors | GUI |
| Pre-kindergarten Teachers | PKTCH | Teachers |
| Kindergarten Teachers | KGTCH |
| Elementary Teachers | ELMTCH |
| Secondary Teachers | SECTCH |
| Ungraded Teachers | UGTCH |
| School Administrators | SCHADM | School Staff |
| School Administrative Support Staff | SCHSUP |
| Paraprofessionals/Instructional Aides | PARA |
| Librarians/media specialists | LIBSPE |
| Library/Media Support Staff | LIBSUP |
| Student Support Services Staff (w/o Psychology) | STUSUPWOPSYCH |
| School Psychologists | SCHPSYCH |
| LEA Administrators | LEAADM | LEA Staff |
| LEA Administrative Support Staff | LEASUP |
| Instructional Coordinators and Supervisors to the Staff | CORSUP |
| Other Support Services Staff | OTHSUP | Other Staff |

Note that prior to SY 2019-20, school psychologists were reported as part of the broader Student Support Services Staff category. Starting with SY 2019-20, School Psychologists were reported separately and the category name for the remaining support services staff was changed to Student Support Services Staff (w/o Psychology).

FTE is the amount of time required to perform an assignment stated as a proportion of a full-time position and is computed by dividing the amount of time employed by the amount of time normally required for a full-time position.

Not all states are able to report all of the staff categories collected by the CCD. Staff categories that are not reported on the SEA level are imputed. Unreported staff categories are not imputed on the school or LEA levels.

Aggregations of lower level staff data may not equal the FTE reported at a higher level. Typically, there are staff members working at higher levels that are not counted at the LEA or school level.

### National School Lunch Program (NSLP) eligibility – Free/Reduced-Price Lunch and Direct Certification

States submit two measures of eligibility for low- and reduced-cost meal programs – Free and Reduced-price Lunch and Direct certification (from SY 2016-17 forward). Refer to the NCES blog [Understanding School Lunch Eligibility in the Common Core of Data](../../../../Forms/AllItems.aspx?RootFolder=%2Fed%2FEDMITS%2FShared%20Documents%2F09%20%2D%20Change%20Management%2F01%20ECRB%2F01%20In%20Process&FolderCTID=0x012000425CAF913917DF43A883566EFE403D7C&View=%7BA8502C75%2D55D9%2D40F7%2DB672%2D7499FFD326CF%7D) for more information on reporting changes and guidance on how to use this CCD data element.

1. Starting with the SY 2019–20 reporting cycle, the English Learner data files will not be released with the CCD Universe Survey. Instead, these data will be released by the data steward through [ED Data Express](https://eddataexpress.ed.gov/). [↑](#footnote-ref-1)
2. Starting with the SY 2019–20 reporting cycle, the English Learner data files will not be released with the CCD Universe Survey. Instead, these data will be released by the data steward through [ED Data Express](https://eddataexpress.ed.gov/). [↑](#footnote-ref-2)