

# Data Documentation

---

## About

The College Scorecard project is designed to increase transparency, putting the power in the hands of students and families to compare colleges and see how well schools are preparing their students to be successful. This project provides more data than ever before to help students and families compare college costs and outcomes as they weigh the tradeoffs of different colleges, accounting for their own needs and educational goals.

These data are provided through federal reporting from institutions, data on federal financial aid, and tax information. These data provide insights into the performance of schools that receive federal financial aid dollars, and the outcomes of the students of those schools.

## Accuracy and Privacy

Many elements are available only for Title IV recipients, or students who receive federal grants and loans. These data are reported at the individual level to the National Student Loan Data System (NSLDS), which is used to distribute federal aid, and published at the aggregate institutional level. While some schools report these data at the campus level (8-digit OPE ID), data produced for this site are rolled up to the institution level (6-digit OPE ID). In these cases, IPEDS institutions sharing a common 6-digit OPEID are all assigned the (student-weighted) average outcome or median outcome for students across all branches of the institution for NSLDS or tax-data derived measures.

All (NSLDS) and Treasury elements are protected for privacy purposes; those data are shown as PrivacySuppressed. Note that for many elements, we have also taken additional steps to ensure data are stable from year to year and representative of a certain number of students. For most elements, data are pooled across two years of data to reduce year-over-year variability in figures (i.e. repayment rate, debt figures, earnings). Moreover, for elements that we expect to highlight on the consumer-facing College Scorecard, a separate version of the element is available that suppresses the data for schools with fewer than 30 students in the denominator to ensure the data are as representative as possible.

## Root

These select items refer to the most basic information in the data set.

<b>Dev-category</b>	root
<b>ID</b>	Integer
	Data files are provided at the unit ID level, which is the unique identification number assigned to postsecondary institutions as

surveyed through the Integrated Postsecondary Education Data System (IPEDS) of analysis reported in IPEDS.

OPEID is the identification number used by the U.S. Department of Education's Office of Postsecondary Education (OPE) and Federal Student Aid Office (FSA) to identify schools that have Program Participation Agreements (PPA) so that its students are eligible to participate in Federal Student Financial Assistance programs under Title IV regulations. This is a 6-digit number and is also provided as an 8-digit number with a 2-digit suffix to identify branches, additional locations, and other entities that are part of the eligible institution.

**OPEID and Unit ID Crosswalk** As a part of this project, the National Center for Education Statistics and the Federal Student Aid office collaborated to help create a crosswalk between IPEDS IDs (UnitID) and FSA IDs (OPEID). This is a critical, and previously unavailable, map between the two systems' differing definitions of an institution. It includes the history of institutions from 2000 to 2014, including new locations, changes in affiliation, and closures. The OPEID/Unit ID map across years represents this crosswalk.

### **Currently Operating**

Boolean

Schools that are noted as "currently operating" by the Office of Federal Student Aid are noted with a 1; schools that are not currently operating (i.e. have closed or merged, or are no longer eligible for Title IV aid) are noted with a 0.

## **About the School**

A number of the elements produced in this API provide basic descriptive information about the school in question. These include: identifiers, location, degree type and profile, programs offered, and the academic profile of students enrolled. Most of these elements are already available through the IPEDS.

### **Dev-category**

**school**

### **Name**

String

The institution's name (INSTNM), as reported in IPEDS.

### **Location**

String

The institution's location, as reported in IPEDS (CITY, STABBR, ZIP). This is also reported using latitude and longitude (integer; LATITUDE and LONGITUDE).

### **URLs**

String

Each institution reports the URL of its homepage (INSTURL) and the URL of its net price calculator (NPCURL). Both URLs are included on the College Scorecard. These URLs are included only in the 2013 data file.

**Main Campus/Branch**

Boolean and Integer

The main campus column (main) identifies whether the school's Unit ID represents the main campus of the school or not, where 1 is a main campus and 0 is not. The branch campus column (NUMBRANCH) identifies the number of branch campuses at that school.

**Accrediting Agency**

String

The institutional accreditor for each school is identified (AccredAgency) to facilitate further analysis. Those data are maintained by the Department of Education's Office of Postsecondary Education, and are included only in the 2013 file, but represent data from Summer 2015.

**Degree Type**

Integer

There are several elements that identify the degree profile of the institution. Highest degree (HIGHDEG) identifies the highest award level offered at the school. Predominant degree (PREDEG) identifies the type of degree that the school primarily awards; for instance, a school that awards 40 percent bachelor's degrees, 30 percent associate degrees, and 30 percent certificate programs would be classified as predominantly bachelor's degree awarding. School degree (sch\_deg) uses the predominant degree calculations, but recodes missing elements (0=N/A) and predominantly graduate schools (4=Graduate) according to the program length reported to NSLDS. School degree is not available after 2012, because it was used to determine the NSLDS completion and transfer rates, which are not available after that year.

**Public/Private Nonprofit/****Private For-Profit**

String

This element (CONTROL) identifies whether the institution's governance structure is public, private nonprofit, or private for-profit.

**Carnegie Classifications**

String

The Carnegie Foundation classifies institutions in several ways. These data include the Basic classification (CCBASIC), the Undergraduate Profile (CCUGPROF), and the size and setting classification (CCSIZSET).

**Special Mission or  
Religious Affiliation**

String

Institutions that are identified as minority-serving institutions, including (for this purpose) Historically Black Colleges and Universities and Tribal Colleges and Universities, are flagged with an indicator (HBCU=Historically Black Colleges and Universities; PBI=Predominantly Black Institutions; ANNHI=Alaska Native-/Native Hawaiian-serving Institutions; TRIBAL=Tribal Colleges and Universities; AANAPII=Asian American-/Native American-Pacific Islander-serving Institutions; HSI=Hispanic-serving Institutions; NANTI=Native American Non-Tribal Institutions). These are all recent grantees of the Department's minority-serving institution programs or designated in statute.

Institutions are also identified by their religious affiliation (RELAFFIL), including not reported or no affiliation. These data are reported by institutions to IPEDS.

Data on special missions are provided only for the most recent year, in the 2013 data file, but represent data from Summer 2015.

**Distance-Only**

Boolean

Schools that are identified as distance education-only (DISTANCEONLY) programs are identified with a 1. Users may find specific programs of study that are offered as distance education-only programs through the CIP variables (i.e. CIP01ASSOC is an online-only program if it is identified as a 2). These data are not reported prior to 2011.

**Revenue/Cost of the School**

Integer

The data files include several elements related to institutional finance. The net tuition revenue per full-time equivalent student (TUITFTE) uses tuition revenue minus discounts and allowances, and divides that by the number of FTE undergraduate and graduate students. Instructional expenditures per FTE student (INEXPFTE) uses instructional expenditures divided by the number of FTE students. The average faculty salary (AVGFACSAL) produces the average faculty salary per month, by dividing the total salary outlays by the number of months worked for all full-time, nonmedical instructional staff. Prior to the 2011-12 academic year, when months worked were reported in groups, the value for 9-10 months is estimated as 9.5 months and the value for 11-12 months is estimated as 11.5 months. Values prior to the 2003-04

academic year are limited to degree-granting institutions for consistency with values in subsequent years.

## Heightened Cash Monitoring 2 Boolean

The U.S. Department of Education places institutions on a Heightened Cash Monitoring (HCM) payment method to provide additional oversight of cash management. HCM2 is the type of HCM that indicates more serious financial or federal compliance issues. These data are maintained by the Department of Education's Office of Federal Student Aid. Users should look to the 2013 data file for the most recent flag, but these data were reported as of Summer 2015.

## Academics

This information describes the types of academic offerings available at each school.

**Dev-category** academics

**Programs Offered by Type** Integer

The Classification of Instructional Programs (CIP) provides a structure in which to track and report in fields of study. Two types of program data are included in these data. The first set (PCIP[01-54]) provide the percentage of degrees awarded in each two-digit CIP code field of study. The second set (CIP[01-54][CERT1/CERT2/ASSOC/BACHL/CERT4]) identifies whether the institution offers the program, at what level, and whether the institution offers the program and level through an exclusively distance education program. It is calculated from counts of awards made in each CIP, as reported in IPEDS.

## Admissions

Several elements, including some that haven't been published before, identify demographic and other

**Dev-category** admissions



**Admission Rate**

Float

Colleges report to IPEDS their admissions rate, defined as the number of admitted undergraduates divided by the number of undergraduates who applied. For institutions with multiple branches, ADM\_RATE includes the admissions rate at each campus, while ADM\_RATE\_ALL represents the admissions rate across all campuses, defined as the total number of admitted undergraduates across all branches divided by the total number of undergraduates who applied across all branches.

**SAT and ACT Scores**

Float

Test scores of enrolled students are not reported for all institutions, but may help students to find a school that is a good academic match. Our files include the 25<sup>th</sup> and 75<sup>th</sup> percentiles of SAT reading (SATVR\* for \_25 and \_75), writing (SATWR\* for \_25 and \_75), math (SATMT\* for \_25 and \_75); and of ACT English (ACTEN\* for 25 and 75), writing (ACTWR\* for 25 and 75), math (ACTMT\* for 25 and 75), and cumulative (ACTCM\* for 25 and 75) scores. We also have derived midpoints of the ACT (ACT\*MID for CM, EN, MT, and WR) and SAT (SAT\*MID for VR, MT, WR, but not cumulative) scores; however, note that institutions do not report those midpoints to IPEDS. SAT and ACT data are available from 2001 on; however, SAT writing scores are available only from 2006 on, and ACT writing data are available only from 2008 on.

**Costs**

Information about the costs to students of a school can provide important context for students and families as they seek to evaluate the tradeoffs of access, affordability, and outcomes. While several of the elements in this category are commonly used elements from IPEDS, others are new and/or different elements derived from NSLDS.

**Dev-category****cost****Average Cost of Attendance,  
Tuition and Fees**

Integer

The average annual cost of attendance includes tuition and fees, books and supplies, and living expenses for all full-time, first-time, degree-/certificate-seeking undergraduates who receive Title IV aid. These data are available separately for academic year institutions (COSTT4\_A) and for program-year institutions (COSTT4\_P). For program-year institutions, the cost of attendance represents the program with the largest enrollment at the institution for the length of the full program, regardless of program length, and includes undergraduates who first enrolled at any time during the academic year. Data are not reported prior to 2009.

Additionally, the cost data include the tuition and required fees of the institution. They are provided for in-state students (TUITIONFEE\_IN), out-of-state students (TUITIONFEE\_OUT), and program-year institutions (TUITIONFEE\_PROG). For academic year institutions, tuition and fees are calculated for full-time and first-time students. For non-academic-year (program-year or continuous enrollment) institutions, tuition and fees

represent those charged to first-time, full-time undergraduates for the largest program at the institution, regardless of program length. Some institutions have different tuition and fees for in-district students that are not reflected in this metric. Data are not reported prior to 2000.

### Average Net Price

Integer

There are several elements in the dataset that describe the average net price, derived from the full cost of attendance (including tuition and fees, books and supplies, and living expenses) minus federal, state, and institutional aid, for undergraduate Title IV-receiving students. Average net price (NPT4\_\* for \_PUB [public colleges; for public institutions, this metric is limited to undergraduates who pay in-state tuition] and \_PRIV [private colleges]) includes a weighted average of all undergraduate, Title IV-receiving students; whereas average net price by income quintile (NPT41\_\*, NPT42\_\*, NPT43\_\*, NPT44\_\*, and NPT45\_\* for \_PUB and \_PRIV, as well as \_PROG [program—year reporters] and \_OTHER [other academic calendar schools]) provides the net price separately for each income bracket for those students. Income quintiles are: (1) \$0-\$30,000; (2) \$30,001-\$48,000; (3) \$48,001-\$75,000; (4) \$75,001-\$110,000; and (5) \$110,000+. We also included several combinations of net price quintiles, which rely on a weighted average of those quintiles (i.e. NPT4\_048\_\* is the net price for the first two income brackets, \$0-\$48,000). Net price data are not reported prior to 2009.

### Student Body

Several elements, including some that haven't been published before, identify demographic and other details about the student body of the school. Some of the elements are available through the Integrated Postsecondary Education Data System, and others were produced using the National Student Loan Data System.

#### Dev-Category

student

#### Number of Undergraduate Students

Integer

This element (UGDS) includes the number of degree/certificate-seeking undergraduates enrolled in the fall, as reported to IPEDS. In 2001, degree-/certificate-seeking status was not collected, so data for that year (UG) are reported as a separate element.

#### Undergraduate Student Body by Race:

Float

These data are reported by institutions to IPEDS; and rely on self-reported data. This includes the total enrollment of undergraduate, degree-seeking students, based on fall enrollment, who are: white (UGDS\_WHITE), black (UGDS\_BLACK), Hispanic (UGDS\_HISP), Asian (UGDS\_ASIAN), American Indian/Alaska Native (UGDS\_AIAN), Native Hawaiian/Pacific Islander (UGDS\_NHPI), two or more races (UGDS\_2MOR), non-resident aliens (UGDS\_NRA), and race unknown (UGDS\_UNKN). Note that prior to 2008, the categories were different, and are designated as separate elements. Additionally, in 2000-01, degree-/certificate-seeking status was not collected, so data for that year are also reported as a separate element.

**Undergraduate Students by  
Part-Time/Full-Time Status**

Float

This element includes the proportion of degree/certificate-seeking undergraduates enrolled part time in the fall term, as calculated from IPEDS data (PPTUG\_EF). Note that this metric does not apply during 2000-01, when degree-/certificate-seeking status was not collected; data for that year are reported separately (PPTUG\_EF2) and calculated from the number of part-time undergraduates divided by the total number of undergraduate. Both sets of data include non-first-time undergraduates.

**Undergraduate Students  
by Family Income**

Float

Using data from NSLDS, these elements calculate the percentage of Title IV-receiving students who had family incomes from the FAFSA in each of five income quintiles (INC\_PCT\_LO = \$0-\$30,000; INC\_PCT\_M1 = \$30,001-\$48,000; INC\_PCT\_M2 = \$48,001-\$75,000; INC\_PCT\_H1 = \$75,001-\$110,000; and INC\_PCT\_H2 = \$110,001+). These data are also produced separately for dependent (DEP\_INC\_PCT\_\*\*) and independent (IND\_INC\_PCT\_\*\*) students. Data are produced for rolling two-year pooled cohorts; for instance, the 2013 data file includes data for the 2013 and 2014 cohorts of students, and the 2012 data file includes data for the 2012 and 2013 cohorts of students.

The data files also include the counts of students in each bracket used for the average net price by income quintile measurements (NUM[1-5]\_[PUB or PRIV]). This element includes the number of full-time, first-time, degree/certificate-seeking undergraduates who received Title IV



aid. Separate metrics are calculated for public institutions and private institutions. This metric is not available prior to the 2010 academic year.

### **Retention Rate**

Float

Available through IPEDS, retention rate identifies (separately) the share of full-time and part-time students in the prior year, at four-year (RET\_FT4 and RET\_PT4) and less-than-four-year institutions (RET\_FTL4 and RET\_PTL4), who return to the institution after the first year. Data are not available prior to 2004.

### **Undergraduate Student Body by Age**

Float

These data (UG25abv) are reported by institutions every other year to IPEDS. This element identifies the share of students at the institution who are ages 25 to 64, as reported by the institutions fall reporting date. The most recent data are available in the 2013 data file.

### **Share of First-Generation Students**

Float

This element identifies the share of federally aided students at the institution who report that they are first-generation on the FAFSA (PAR\_ED\_PCT\_1STGEN). The data are also available separately for the share of students who report that their parents' highest education level is middle school (PAR\_ED\_PCT\_MS), high school (PAR\_ED\_PCT\_HS), and some postsecondary education (PAR\_ED\_PCT\_PS). Data are produced for rolling two-year pooled cohorts; for instance, the 2013 data file includes data for the 2013 and 2014 cohorts of students, and the 2012 data file includes data for the 2012 and 2013 cohorts of students. For these variables, years refer to award years (e.g., award year 2013 begins on July 1, 2012, and ends June 30, 2013).

### **Number of Schools to Which Students Sent FAFSAs**

Float

When students complete their applications for federal student aid (FAFSA), they must list at least one college (and may elect to list several) to receive their information. This element depicts the percentage of students who send their FAFSA report to at least 2, 3, 4, or 5 institutions (APPL\_SCH\_PCT\_GE\* for 2, 3, 4, or 5). Data are produced for rolling two-year pooled cohorts; for instance, the 2013 data file includes data for the 2013 and 2014 cohorts of students, and the 2012 data file

includes data for the 2012 and 2013 cohorts of students. For these variables, years refer to award years (e.g., award year 2013 begins on July 1, 2012, and ends June 30, 2013)

Note that in the interest of consistency, the data were also produced separately by the Treasury Department for the cohorts used to generate earnings data. Those data (fsend\_\* for 1-5) measure the share of students who submitted FAFSAs to only one college; or at least 2, 3, 4, or 5 schools. Those data are available for the 2005-2006 pooled cohorts (see 2005 file); the 2003-2004 cohorts (see 2003 file); the 2001-2002 cohorts (see 2001 file); the 1999-2000 cohorts (see 1999 file); and the 1996-1997 cohorts (see 1997 file).

### **Undergraduate Student Demographics for Earnings Cohorts**

Float and Integer

In the interest of consistency and data completeness, the Treasury Department produced demographic elements for each of the cohorts for which it produced earnings measures. These include: the share of female students (female) and students over 23 at entry (agege24), using Social Security Administration data; the share of married students (married); the share of dependent students (dependent); the share of veteran students (veteran); and the share of first-generation students (first\_gen). It also includes the average and median family income of students (faminc and md\_faminc) and separately of independent students (faminc\_ind); the percentage of the population from students' home ZIP codes by race (pct\_\* for white, black, asian, hispanic), education level (pct\_\* for ba [bachelor's degree over the age of 25] and grad\_prof [professional degree over the age of 25]), nationality (pct\_born\_us), poverty (poverty\_rate), household income (median\_hh\_inc), and employment status (unemp\_rate) using Census data. Those data are available for the 2005-2006 cohorts (see 2005 file); the 2003-2004 cohorts (see 2003 file); the 2001-2002 cohorts (see 2001 file); the 1999-2000 cohorts (see 1999 file); and the 1996-1997 cohorts (see 1997 file). For these variables, years refer to award years (e.g., award year 2013 begins on July 1, 2012, and ends June 30, 2013).

### **Financial Aid**

Federal financial aid, including Pell Grants and federal student loans, helps many students, particularly low-income students, access and afford a higher education. Data on the amount of debt that students can expect to borrow and the loan performance of former students may help students find the best option in their price range.

<b>Dev-Category</b>	<b>student</b>
<b>Percent of Undergraduates Receiving Federal Loans</b>	<p>Float</p> <p>This element (PCTFLOAN), as reported in IPEDS, shows the share of undergraduate students who received federal loans in a given year. It can provide important context to figures related to debt, repayment, and non-repayment. This figure may be influenced by the eligibility for federal loans and the extent to which students apply for federal loans, as well as by the cost of the programs. In particular, many community colleges are sufficiently low-cost to have low federal loan borrowing rates; and it may be difficult to compare borrowing behaviors for those institutions. These data are not available prior to 2009.</p>
<b>Percentage of Pell Students</b>	<p>Float</p> <p>This element (PCTPELL), pulled from IPEDS, shows the share of undergraduate students who received Pell Grants in a given year. This is an important measure of the access a school provides to low-income students. However, it may not capture all low-income students. Students who are undocumented immigrants or foreign nationals are not eligible to receive Pell Grants, and some low-income students may not have completed the FAFSA to receive federal aid, but those students may have similar financial circumstances to Pell recipients, or may be just on the other side of Pell eligibility, creating a cliff effect. Additionally, in some states (such as California), state financial aid may be sufficient to cover costs at community colleges, in particular; so those students may not seek or receive a Pell Grant. These data are not available prior to 2008.</p>
<b>Cumulative Median Debt</b>	<p>Integer</p> <p>This is the median loan debt accumulated at the institution by all student borrowers of federal loans who separate (i.e., either graduate or withdraw) in a given fiscal year, measured at the point of separation (DEBT_MDN). More specifically, the measure represents the sum of all originated federal loans over students' college education—e.g., if a student receives a federal loan for \$2,000 for each of eight semesters, their cumulative debt is recorded as \$16,000.</p> <p>These data are available for all borrowers at the institution, as well as disaggregated by completion status (GRAD_DEBT_MDN for students who completed and WDRAW_DEBT_MDN for students who withdrew</p>

without completing); by FAFSA family income (LO\_INC\_DEBT\_MDN = \$0-\$30,000; MD\_INC\_DEBT\_MDN = \$30,001-\$75,000; and HI\_INC\_DEBT\_MDN = \$75,001+); by dependent (DEP\_DEBT\_MDN) and independent (IND\_DEBT\_MDN) status; by Pell status (PELL\_DEBT\_MDN for students who ever received a Pell Grant and NOPELL\_DEBT\_MDN for students who never received a Pell Grant); by gender (FEMALE\_DEBT\_MDN and MALE\_DEBT\_MDN); and by first-generation status (FIRSTGEN\_DEBT\_MDN and NOTFIRSTGEN\_DEBT\_MDN).

At schools where large numbers of students withdraw before completion, a lower median debt level could simply reflect the lack of time that a typical student spends at the institution. Therefore, the Department of Education recommends using the typical debt level for students who complete (GRAD\_DEBT\_MDN\_SUPP or GRAD\_DEBT\_MDN10YR\_SUPP for the debt level expressed in monthly payments). Additionally, this measure can be placed in context by looking at the borrowing rate of students at the institution (PCTFLOAN; see above); at schools where few students borrow, the numbers may represent outliers.

Data are produced for rolling two-year pooled cohorts; for instance, the 2013 data file includes data for the 2013 and 2014 cohorts of students, and the 2012 data file includes data for the 2012 and 2013 cohorts of students. For these variables, years refer to award years (e.g., award year 2013 begins on July 1, 2012, and ends June 30, 2013).

## Completion

College completion is associated with other positive outcomes, like finding a job and successfully repaying student loans, and is an important metric for evaluating the experiences of students at the institution. However, both existing and new methods of measuring completion have limitations.

### Dev-category

### completion

Completion rates for first-time, full-time students

### 150 Percent IPEDS

### Completion Rate

Float

Currently, institutions report to the Integrated Postsecondary Education Data System on the completion rates for first-time, full-time students who begin school in the fall semester and complete within 150 percent of the expected time to completion (C150\_4 for four-year schools and C150\_L4 for less-than-four-year schools); those rates are also available disaggregated by race (C150\_4\_\* and C150\_L4\_\*, for \_WHITE [white],

\_BLACK [black], \_HISP [Hispanic], \_ASIAN [Asian], \_AIAN [American Indian/Alaska Native], \_NHPI [Native Hawaiian/Pacific Islander], \_2MOR [two or more races], \_NRA [non-resident alien], and \_UNKN [race unknown]).

The exclusion of part-time students, transfer students, and students who do not start during the fall from IPEDS completion rates makes the rates less relevant for those populations of students. First-time full-time students make up fewer than half of all college students, or even less in some sectors of institutions (e.g. community colleges). Furthermore, although schools have the option to report transfer outcomes for first-time full-time students, many choose not to. In light of these limitations, IPEDS will soon collect graduation rates for four cohorts of students based on their full-time and part-time, and first-time and non-first-time, status.

To reduce variability from year to year, the data are also available as pooled completion rates across two years on a rolling basis (C150\_4\_POOLED); among institutions with fewer than 30 students in the combined cohorts, the measure was created based on a four-year cohort instead (see poolyrs for the number of years used in the rolling average). The data are also suppressed for schools with fewer than 30 students, where the outcome of a single student could dramatically change the rate (C150\_4\_POOLED\_SUPP). These data may be most appropriate for consumer-facing tools.

Pooled data are available only in the 2013 data file (2007-2008 cohorts as collected in 2012 and 2013 for four-year institutions, or 2005-2008 cohorts as collected in 2010, 2011, 2012, and 2013 in the event of very small institutional cohorts). One-year (non-pooled) measures are included for every year (C150\_4 and C150\_L4).

## **200 Percent IPEDS Completion Rate**

Float

Currently, institutions also report to the Integrated Postsecondary Education Data System on the completion rates for first-time, full-time students who begin school in the fall semester and completion within 200 percent of the expected time to completion (C200\_4 for four-year schools and C200\_L4 for less-than-four-year schools). Although these rates face the same limitations as the 150 percent completion rate above, we used the 200 percent completion rates for less-than-four-year institutions, with similar pooling over years and suppression for

small schools (C200\_L4\_POOLED\_SUPP). This is primarily because at many two-year institutions, even students who begin as first-time, full-time tend to experience more bouts of part-time status, extending the time required to complete.

Pooled data are available only in the 2013 data file (2009-2010 cohorts as collected in 2012 and 2013 for less-than-four-year institutions, or 2005-2010 cohorts in the event of very small institutional cohorts as collected in 2010, 2011, 2012, and 2013).

## Outcomes for Title IV Students

### NSLDS Completion and Transfer Rates

Float

Using data from the National Student Loan Data System, the Department also produced completion rates that track institutional outcomes for students who receive federal financial aid. Each school has all possible outcomes reported: share of students who died (DEATH\_YR\*\_RT), completed at the original institution (COMP\_ORIG\_YR\*\_RT), transferred and completed at a four-year institution (COMP\_4YR\_TRANS\_YR\*\_RT), transferred and completed at a two-year institution (COMP\_2YR\_TRANS\_YR\*\_RT), withdrew from the original institution (WDRAW\_ORIG\_YR\*\_RT), transferred and withdrew from a four-year institution (WDRAW\_4YR\_TRANS\_YR\*\_RT), transferred and withdrew from a two-year institution (WDRAW\_2YR\_TRANS\_YR\*\_RT), still enrolled at the original institution (ENRL\_ORIG\_YR\*\_RT), transferred and is still enrolled at a four-year institution (ENRL\_4YR\_TRANS\_YR\*\_RT), transferred and is still enrolled at a two-year institution (ENRL\_2YR\_TRANS\_YR\*\_RT), status unknown at the original institution (UNKN\_ORIG\_YR\*\_RT), transferred to a four-year institution and status is unknown (UNKN\_4YR\_TRANS\_YR\*\_RT), and transferred to a two-year institution and status is unknown (UNKN\_2YR\_TRANS\_YR\*\_RT).

Each of those rates is available at two (\_YR2\_RT), three (\_YR3\_RT), four (\_YR4\_RT), six (\_YR6\_RT), and eight (\_YR8\_RT) years after entering the school. This can be translated into percentages of expected time to completion; for two-year schools, it's 100%, 150%, 200%, 300%, and 400% of expected time to completion, and for four-year schools, it's 50%, 75%, 100%, 150%, and 200% of expected time to completion.

Additionally, each rate is disaggregated for low-income students (LO\_INC\_\* = \$0-\$30,000 in FAFSA family income), middle-income students (MD\_INC\_\* = \$30,001-\$75,000 in FAFSA family income), and high-income students (HI\_INC\_\* = \$75,001+ in FAFSA family income); dependent (DEP\_\*) and independent (IND\_\*) students; male (MALE\_\*) and female (FEMALE\_\*) students; students who ever received a Pell grant (PELL\_\*) and students who never received a Pell grant (NOPELL\_\*); students who ever received a federal loan at the measured school (LOAN\_\*) and students who never received a loan there (NOLOAN\_\*); and for first-generation students (FIRSTGEN\_\*) and not-first-generation students (NOT1STGEN\_\*).

NSLDS records the first time students received aid, rather than the first time they enroll. Therefore, students were placed in cohorts based on students' responses to a question on the FAFSA about their grade level; students indicating they were first-year students were placed in the completion cohort for that award year; second-year placed in the previous cohort; and third- and fourth-year students were placed two cohorts prior to reduce the risk of misreporting on the FAFSA.

Based on those cohorts, the Department used the enrollment records reported by institutions to determine the percentage of students completed within a given period. Where schools had not provided any information, including a completion status, students were recorded as non-completers. Additionally, students who transferred to another school (i.e., separated from the original institution and subsequently received Title IV aid at another institution) were reported separately.

The Department identified several limitations that led us to exclude these rates from our consumer tool until schools have another opportunity to improve their reporting. Some schools seem to report completion and withdrawal interchangeably, since both indicate the student is entering repayment. Similarly, because NSLDS is used for administering financial aid and for those purposes separating from the school only mattered for students with loans who then entered repayment, schools were not required to report completion status for students who receive only grants prior to 2012, so historical rates reflect exceptionally low completion rates for Pell-only students. Reporting of part-time and full-time status is exceptionally spotty, so it is not possible from the data to separate out those students; and, of course, the rates do not include non-Title-IV students, and may therefore be unrepresentative of the outcomes of some colleges.

Data files include data for different cohorts measured at different times. Data in the XXXX merged file corresponds to the AAAA cohort year based on the measurement period designated as YRZ in the variable name, such that XXXX=AAAA-1+Z. For example, the variable names with \*YR4\* in the 2013 data file refer to the 2010 cohort measured in 2014. Those same variables in the 2012 data file refer to the 2009 cohort measured in 2013. For these variables, years refer to award years (e.g., award year 2013 begins on July 1, 2012, and ends June 30, 2013)

## Earnings

One of the most common reasons students cite in choosing to go to college is the expansion of employment opportunities. To that end, data on the earnings and employment prospects of former students can provide key information. To measure the labor market outcomes of individuals attending institutions of higher education, data on cohorts of federally aided students were linked with earnings data from de-identified tax records and reported back at the aggregate, institutional level.

There are two notable limitations that researchers should keep in mind for all of these metrics. First, the data are not yet available to produce program-level earnings data. Research suggests that the variation across programs within a school may be even greater than aggregate earnings across schools; for instance, STEM and health majors frequently earn more than students who study in other fields. Second, the data include only Title IV-receiving students, so figures may not be representative of schools with a low proportion of Title IV-eligible students. Additionally, the data are restricted to students who are not enrolled (enrolled means having an in-school deferment status for at least 30 days of the measurement year), so students who are currently enrolled in graduate school at the time of measurement are excluded.

### Dev-category

### earnings

### Mean and Median Earnings

Integer

Mean (mn\_earn\_wne\_p\*) and median (md\_earn\_wne\_p\*) earnings are for the institutional aggregate of all federally aided students who enroll in an institution each year and who are employed but not enrolled. Earnings are defined as the sum of wages and deferred compensation from all non-duplicate W-2 forms received for each individual, plus positive self-employment earnings from Schedule SE. Data are available for each year starting six years after a student enrolls in college, up to 10 years after the student enrolls; enrollment dates are estimated based on FAFSA self-reporting, as with the completion rate cohort construction described above.

Mean earnings are also available in years 6 and 10 after entry to the institution (but not between those years) disaggregated by



FAFSA family income (mn\_earn\_wne\_inc1\_p\*, mn\_earn\_wne\_inc2\_p\*, and mn\_earn\_wne\_inc3\_p\*); by dependent status (for dependents, mn\_earn\_wne\_indep0\_p\*; and for independents, mn\_earn\_wne\_indep1\_p\*); and by gender (for female, mn\_earn\_wne\_male0\_p\*; and for male, mn\_earn\_wne\_male1\_p\*).

Earnings are based on measurement periods with years after cohort entry labeled in the variable names with a \*pZ at the end. For a given pooled cohort AAAA-BBBB, where BBBB is AAAA+1, data are included on merged file XXXX based on XXXX=AAAA+Z. For example, ten-year earnings (\*p10) in the 2011 data file refer to the 2001-2002 pooled cohorts measured in the 2011 and 2012 calendar years. While earnings are measured based on calendar years, cohort years for earnings variables are based on award years.

Those data are available for the 2005-2006 cohorts; the 2003-2004 cohorts; the 2001-2002 cohorts; the 1999-2000 cohorts; and the 1996-1997 cohorts. Disaggregated earnings data are available only at 6 and 10 years after entry for the relevant cohorts.

## Threshold Earnings

Float

This measure describes the fraction of former students earning over \$25,000 in 2014 constant dollars (gt\_25k\_p\*). It is available for each year from six to 10 years after entering the school. The threshold of \$25,000 was chosen because it corresponds approximately to the median wage of workers ages 25 to 34 with only a high-school degree. Establishing a baseline measure of success—is the typical student who attended this school in better financial circumstances than if he had begun working with only a high school diploma?—avoids many of the problems created by the lack of program-level data.

Earnings are based on measurement periods with years after cohort entry labeled in the variable names with a p\* at the end. For a given pooled cohort AAAA-BBBB, where BBBB is AAAA+1, data are included on merged file XXXX based on XXXX=AAAA-1+\*. For example, ten-year earnings (\*p10) in the 2011 data file refer to the 2001-2002 pooled cohort measured in the 2011 and

2012 calendar years. While earnings are measured based on calendar years, cohort years are based on award years.

Those data are available for the 2005-2006 cohorts; the 2003-2004 cohorts; the 2001-2002 cohorts; the 1999-2000 cohorts; and the 1996-1997 cohorts.

## Repayment

To provide a sense for the debt burden of attending college and the loan performance metrics for each school, we produced several elements using the National Student Loan Data System. These elements can provide useful information for students and families concerned about borrowing for college and interested in seeing borrowers' behavior after they leave the school.

### Dev-category

### repayment

#### Cohort Default Rate

Float

Cohort default rates are produced annually as an institutional accountability metric; schools with too-high default rates may lose access to federal financial aid. The three-year cohort default rate (CDR3) represents a snapshot in time; FY 2011 rates were calculated using the cohort of borrowers who entered repayment on their federal student loans between October 1, 2010 and September 30, 2011, and who defaulted before September 30, 2013. Three-year CDR data are not available prior to 2011. The three-year CDR replaced the two-year CDR, which is available prior to 2013.

The CDR is susceptible to gaming behavior that may push students toward forbearance and deferments, meaning they stay out of default but don't make progress on repaying their loans and may continue to accrue interest. The repayment rate resolves many of those issues.

The 2013 data file for three-year CDR variables refer to the 2011 fiscal year CDR cohort as measured in the 2013 fiscal year.

#### Repayment Rate on Federal Student Loans

Float

This element depicts the fraction of borrowers at an institution who have not defaulted on their federal loans and who are making progress in paying them down (i.e. have paid down at least \$1 in the principal balance on their loans) after leaving school (RPY\_\*YR\_RT). The rates are available for 1 (\_1YR\_RT), 3

(\_3YR\_RT), 5 (\_5YR\_RT), and 7 (\_7YR\_RT) years after leaving school. Repayment rates are generally considered superior to default rates, which measure only the worst-case scenario for repayment outcomes and which can be manipulated through the use of allowable nonrepayment options like deferments and forbearances.

Repayment rates are based on the set of federal loan borrowers who enter repayment in a given fiscal year, so the 2011 repayment cohort is based on students entering repayment from October 1, 2010 to September 30, 2011. Since students who graduate may not immediately enter repayment due to either their 6-month grace period, or being granted deferment because of hardship or upon entering graduate school, students are likely to enter repayment in a different year than when they exit (and are captured in the median cumulative debt metric). Students who have received either an in-school or military deferment are excluded from the calculation.

These data are available for all borrowers at the institution, as well as disaggregated by completion status (COMPL\_RPY\_\* for students who completed and NONCOM\_RPY\_\* for students who withdrew without completing); by FAFSA family income (LO\_INC\_RPY\_\* = \$0-\$30,000; MD\_INC\_RPY\_\* = \$30,001-\$75,000; and HI\_INC\_RPY\_\* = \$75,001+); by dependent (DEP\_RPY\_\*) and independent (IND\_RPY\_\*) status; by Pell status (PELL\_RPY\_\* for students who ever received a Pell Grant and NOPELL\_RPY\_\* for students who never received a Pell Grant); by gender (FEMALE\_RPY\_\* and MALE\_RPY\_\*); and by first-generation status (FIRSTGEN\_RPY\_\* and NOTFIRSTGEN\_RPY\_\*). The repayment rates are produced in rolling two-year averages to reduce variability from year to year; and the three-year repayment rate, which will be included on our consumer tool, is suppressed for schools with fewer than 30 borrowers in the two cohorts to produce more stable measures (RPY\_3YR\_RT\_SUPP).

Data included in the XXXX merged file based on pooled cohorts AAAA-BBBB, where BBBB is AAAA+1, and the measurement period after entering repayment designated in variable names with \*YRZ\* such that XXXX=AAAA+Z. For example, in the 2013 data file, the three-year repayment rate refers to the 2010-2011

pooled cohort as measured in 2013 and 2014. Data are produced for rolling two-year pooled cohorts for 2006-2007, 2007-2008, 2008-2009, 2009-2010, and 2010-2011. For these variables, years refer to fiscal years.