

Data Documentation
Preliminary Loan Debt Data by Field of Study
version: 5/21/19

Introduction:

This document describes College Scorecard's *Preliminary Loan Debt Data by Field of Study* data file. Given these data are preliminary, the U.S. Department of Education (the Department) does not recommend using these data for informing enrollment decisions. College Scorecard will provide a final version appropriate for a wider audience (i.e. prospective students) in the fall of 2019 after institutions have had a chance to revise the underlying data from which these metrics are derived.

Data methodology:

The *Preliminary Loan Debt Data by Field of Study* data file provides the borrower count, median cumulative loan debt, and mean cumulative loan debt by field of study for institutions that participate in the Federal student financial assistance programs administered by the Department under Title IV of the Higher Education Act of 1965 as amended. The unit of analysis for these calculations is the unique combination of an institution (designated by a 6-digit Office of Postsecondary Education Identification Number (OPEID)), a field of study within an institution classified using a 4-digit Classification of Instructional Program (CIP) code, and a credential level code. Programs include those administered by institutions located in the United States, other jurisdictions of the United States (e.g., Puerto Rico), or in foreign countries. In addition, the data file includes the number of awards conferred by field of study as reported by title IV administering institutions to the Integrated Postsecondary Education Data System (IPEDS).

College Scorecard calculated the borrower count, median cumulative loan debt, and mean cumulative loan debt using data in the National Student Loan Data System (NSLDS) by taking the following actions:

- Define and calculate *observations* as the unique combinations of a graduation event, a student, an OPEID, a 4-digit CIP code, and a credential level.
 - De-duplicate observations. If there are multiple graduation events for the same student, OPEID, 4-digit CIP code, and credential level, use only the most recent graduation event and remove all other observations. If multiple graduation events share the same date, and that is the most recent event within the category, use only one of the observations (doesn't matter which one) and remove others. See Appendix I for an example.
- Assign a dollar amount to each observation using the following methodology:
 - For the student and OPEID being evaluated, sum all the loan disbursements (original loan amount less cancellations and refunds) from the William D. Ford Federal Direct Loan Program (Direct Loans), the Federal Family Education Loan Program (FFEL Program), and Graduate PLUS Loan Program with the exception of:
 - loans disbursed after the graduation event¹

¹ The disbursement date refers to the "first disbursement" date to capture all loans used to obtain credits prior to and credits for graduation at the evaluated program

- loans disbursed at a different academic level (i.e. undergraduate versus graduate) based on the following coding scheme:
 - Undergraduate credentials:
 - Undergraduate Certificates or Diplomas
 - Associate's Degrees
 - Bachelor's Degrees
 - Post-baccalaureate Certificates
 - Graduate credentials
 - Master's Degrees
 - Doctoral Degrees
 - First Professional Degrees
 - Graduate / Professional Certificates
- Loans with a Current Loan Status of 'CA' (Cancelled)
 - Exclude observations that are assigned a zero dollar amount.²
- Group observations into combined two-award-year³ cohorts and calculate the count, mean, and median cumulative debt for all observations in each OPEID, 4-digit CIP code, credential combination.

College Scorecard calculated the IPEDS count by taking the following actions:

- Download the appropriate IPEDS Completions survey component data files and associated institutional directory information
- Recode IPEDS award levels to correspond to the credential levels from NSLDS:

IPEDS Award Level	NSLDS Credential Level
Less-than-1-year	Undergraduate Certificates or Diplomas
At least 1- but less-than-2-year	
At least 2- but less-than-4-year certificates	
Associate's degrees	Associates Degrees
Bachelor's degrees	Bachelor's Degrees
Post-baccalaureate certificates	Post-baccalaureate Certificates
Master's degrees	Master's Degrees
Doctor's degrees-research/scholarship	Doctoral Degrees
Doctor's degrees-other	
Doctor's degrees - professional practice	First Professional Degrees
Post-master's certificates	Graduate/Professional Certificates

- Create subtotal records for each IPEDS institution that summarize the IPEDS award level 6-digit CIP code records at the 4-digit CIP code NSLDS credential level by summing the first-major and second-major award counts across records.

² An example of a student who is not in the borrower-only cohort is a student who received a Pell Grant, but not a federal loan.

³ An award year begins July 1 and ends June 30 the following year.

- Combine IPEDS institution records associated with the same 6-digit OPEID, 4-digit CIP code, and credential level into a single record by summing across IPEDS institutions with the same 6-digit OPEID.
- Aggregate IPEDS single award year files into two-award-year groupings by summing records with the same OPEID, 4-digit CIP code, and credential level across award years.

After calculation of both the NSLDS and IPEDS data, College Scorecard match-merged the records by combined award years, OPEID, 4-digit CIP code, and credential level.

Note that the NSLDS and IPEDS counts may differ. Reasons for this difference include, but are not limited to:

- IPEDS including completions that are not included in the NSLDS cohort definition (e.g., completions from non-borrower Title IV students or non-Title IV students)
- Institutions categorizing awards differently (e.g., CIP code selection) for the two different data systems.
- Imperfect alignment of credential level definitions.
- Deduplication of programs during the NSLDS data generating process – some programs, especially focused, short-term certificate programs, may be classified under the same CIP code due to limitations of the CIP while truly being distinct programs.

Data presented in the combined NSLDS and IPEDS data files have been treated with cell suppression methods to reduce the risk of disclosure. To further reduce disclosure risk, information about the specific suppression rules is not available to the public.

Data metrics:

The data metrics for each field of study include

- **COHORT:** The two award years during which the borrower completion cohort was established. The two years provided in this variable represent the end of the award years covered (e.g., 2015_2016 indicates the July 1, 2014 to June 30, 2015 and July 1, 2015 to June 30, 2016 combined award years).
- **OPEID:** The FSA-assigned institution identification number (truncated, six-digit version) is contained in this variable.
- **NAME:** The institution name associated with the OPEID.
- **TYPE:** The type of institution (public, private, proprietary, or foreign).
- **CIPCODE:** The 4-digit Classification of Instructional Programs (CIP) code identifying the program's field of study. For more information about the CIP, see: <https://nces.ed.gov/ipeds/cipcode/>
- **CIPDESC:** The text description of the field of study associated with the CIP code
- **CREDLEV:** The level the credential awarded to the student completing the program of study. Credentials are categorized into the following levels:
 - 1: Undergraduate Certificate or Diploma
 - 2: Associate's Degree
 - 3: Bachelor's Degree
 - 4: Post-baccalaureate Certificate

5: Master's Degree

6: Doctoral Degree

7: First Professional Degree

8: Graduate/Professional Certificate

- **CREDESC:** The text description of the credential level
- **COUNT:** The number of individuals in the cohort.
- **DEBTMEDIAN:** The calculated median debt amount for members of the cohort
- **DEBTMEAN:** The calculated mean debt amount for members of the cohort
- **IPEDSCOUNT:** The number of awards conferred, as captured by the IPEDS Completions component.

Appendix I Graduation Event History for a Hypothetical Student

Award date	OPEID	CIP code	Credentia level	Cumulative debt @ OPEID=111111	Cumulative debt @ OPEID=22222	De-duplication Action
July 1	111111	111101	AA	\$1000		Exclude, not the most recent graduation event in the category. ⁴
August 1	111111	111101	AA	\$1500 ⁵		Exclude, not the most recent graduation event in the category.
September 1	111111	111102	AA	\$3000		Include, the most recent graduation event in the category.
October 1	111111	111102	BA	\$4000		Include, the most recent graduation event in the category.
November 1	111111	222201	BA	\$5000		Since these are simultaneous awards in the same four-digit CIP code category, include just one of the simultaneous awards in the calculation. ⁶
November 1	111111	222202	BA	\$5000		
December 1	222222	111101			\$1000	Include, the most recent graduation event in the category.

⁴ Category refers to unique combinations of six-digit OPEID, four-digit CIP code, and credential level

⁵ Indicates the student took out \$500 of loans at OPEID=111111 after receiving an award in July 2016 and before receiving an award in August 2016. More specifically, the cumulative amount in July was \$1000 and the cumulative amount in August was \$1500, so the loan amount disbursed between the two events is \$500.

⁶ Simultaneous awards should have the same cumulative debt amount