



Methodological Notes

What's New and Improved

1. Data quality enhancements: The previous completions report had breaks in the trend line mainly impacting data after 2013 due to three significant data reporting and processing changes:
 - Beginning in 2011, two changes improved the definition of first-time beginning college students:
 - (1) current dual enrolled students were excluded, and
 - (2) data enhancements enabled reliable exclusion of students with prior enrollments over a longer historical timeframe (impacting mainly older students)
 - The third change relates to improved student-level data matching in the Clearinghouse database that caused unusually large rate changes from 2012 to 2013.

To provide a trend line with improved validity at all time points, all three data quality enhancements were made across all cohorts analyzed in this report.

2. Prior dual enrollment: Former dual enrollment students is now a filter available on the dashboard.
3. Neighborhood income: For the first time, this report includes students' neighborhood income starting with the 2010 cohort.
4. Race/ethnicity: Completion rates were previously analyzed by race/ethnicity but for the first time yearly success components were analyzed by race/ethnicity starting with the 2011 cohort.

Data Coverage

National Student Clearinghouse data track enrollments nationally and are not limited by institutional and state boundaries. As of fall 2022, institutions actively submitting enrollment data to the Clearinghouse account for 97 percent of the nation's postsecondary enrollments in degree-granting institutions. The enrollment data used in this report provide unduplicated headcounts for fall starters from 2007 to 2018.

Cohort Identification, Data Cut, and Definitions

This report examines completion and progress towards completion, up to eight years from first enrollment, for cohorts of first time-in-college degree-seeking students who started their postsecondary studies at U.S. colleges and universities in the fall semesters from 2007 to 2018. To limit the cohorts to first-time, undergraduate students only, the study uses data from the

Clearinghouse's enrollment reporting and DegreeVerify services to confirm that students included in the study fulfilled the following conditions:

1. Enrolled in a Title IV degree-granting institution in fall semester of corresponding cohort year, excluding territories outside the U.S. (e.g., Guam, Puerto Rico, or the U.S. Virgin Islands).
2. Did not have a previous enrollment record, as shown in the Clearinghouse data, prior to the first day of enrollment in the fall of the corresponding cohort year, unless the previous enrollment record was before the student turned 18 years old (qualified as former dual enrollment students).
3. Did not receive any credential from a postsecondary institution prior to the first day of enrollment in the fall of the corresponding cohort year, according to Clearinghouse data, unless the award date was before the student turned 18 years old (former dual enrollment).
4. Had at least one legitimate enrollment status throughout the study period; that is, enrolled for at least one term with full-time, part-time (i.e., halftime or less than half-time), or withdrawal status.
5. Showed intent to seek a degree or certificate. That is:
 - For students who started at four-year institutions, enrolled at least one term with an intensity of halftime or higher.
 - For students who started at two-year institutions, either:
 - o Enrolled full time for at least one term before August 10, of the respective cohort year or
 - o Enrolled three-quarter time for at least one term or half time for any two terms before December 31, of the respective cohort year

We excluded 226,339 students from the 2018 cohort who began at two-year institutions as nondegree-seeking students as a result.

To account for students with postsecondary enrollments due to dual enrollment, students were excluded if they met one or both of these criteria:

1. Earned a high school diploma after the first day of enrollment in the fall of the corresponding cohort year.
2. Were younger than 18 at the time of enrollment.

Race and Ethnicity

Race and ethnicity data are available starting with the fall 2011 cohort. The race and ethnicity categories include Asian, Black, Hispanic, White, Native American, Native Hawaiian or other Pacific Islander, Two or More Races, or Other. The Other category includes International and Unknown/Missing. Approximately 15.2% percent of the data are unknown or missing race/ethnicity over the eight cohorts for which these data are available.

Former Dual Enrollment Students

Our definition of first-time students excludes current dual enrollment students, impacting largely two-year institutions. The cohorts do include *former* dual enrollment students: first-time college students who had previously taken dual enrollment courses. In this report, these are the students who enrolled in college courses while still in high school prior to the fall semester for the corresponding year. Students were identified as former dual enrollment students if they had an enrollment or credential record prior to their diploma data or an enrollment or credential record prior to the corresponding fall semester before they turned 18 years old. Former dual enrollment students represent an average of 22.5% of all the students included in this report. Reference the table below for the average percentage of prior dual enrollment students by sector across all cohorts.

Starting Sector	Percentage of Students with Former Dual Enrollment
Public 4-Year	25.1%
Private Nonprofit 4-Year	19.5%
Private For-Profit 4-Year	6.1%
Public 2-Year	16.6%
Public PAB	19.0%

Starting Enrollment Intensity

In this report, starting enrollment intensity is the student’s enrollment intensity at the time of postsecondary enrollment. Enrollment intensity is reported as full-time, part-time, or less than part-time. Part-time and less than part-time students are grouped together and identified as part-time.

Completion and Concurrent Completion

For this report, we examined completion by first-time students. For overall completion, we defined completion as having obtained a credential at any institution within the eight-year study period. In these cases, completion is defined as obtaining the credential sought or a higher credential. Clearinghouse data provide a unique headcount of U.S. college enrollments during each term, which allows for the tracking of individuals including those with concurrent completion. Some students show more than one completion awarded at multiple institutions on the same day. In these instances, the highest credential earned was selected and if the credential type is the same at both institutions, the completion at the starting institution is prioritized.

Definition of Progress and Success

Students are counted in only one outcome category per year. Where the student is counted is based on a hierarchical order of outcomes listed below. Academic years run from July 1st through June 30th.

Graduated from Starting Institution: Students received a credential from the same institution in which they started college. Once a student completes a credential at the starting institution, the student will be counted in this category across all subsequent years, unless they earn a higher-level credential in a later year. In that case, the outcome category for that year and following years will reflect graduation but in relation to the institution type for the higher-level credential.

Transferred & Graduated from Other Institution: Students did not receive a credential from the starting institution but did receive a credential from a different institution. Students will be counted in this category in all subsequent years unless they earn a higher-level credential in a later year. In that case, the outcome category for that year and following years will reflect graduation but in relation to the institution type for the higher-level credential.

Enrolled at Starting Institution: Students have not yet received a credential but are still enrolled at the starting institution.

Transferred and Enrolled at Other Institution: Students have not yet received a credential but are still enrolled at an institution other than the starting institution.

Not Enrolled: Students had no enrollments during the academic year in question. This is not a terminal outcome. In other words, a student can be counted as “not enrolled” in one academic year but then counted as “enrolled” in a subsequent year if the student returns to college.

Transferred and No Longer Enrolled: Students had no enrollments for the academic year in question, but in a previous academic year had transferred to an institution other than the starting institution.

Never Transferred and Not Enrolled: Student had no enrollments during the academic year in question and had never transferred to an institution other than the starting institution.

Imputation of Values for Gender

The Clearinghouse’s coverage of student gender has increased dramatically for enrollments occurring in recent years. However, imputation of gender for the majority of enrollment records is still necessary in order to use the data for research studies using older cohorts. To meet this need, the Research Center developed an imputation process based on first names. Previously submitted name gender pairs throughout the Clearinghouse database are used to determine the probability of any first name being associated with either gender. To increase the accuracy of the imputation process, the Research Center also draws on name-gender data from the Social Security Administration (SSA) and the U.S. Census Bureau. Because the Clearinghouse collects transactional data, its data contain many more unique first names than other sources. The imputation used only those pairs in which the name occurred in at least two instances and was associated with a single

gender in at least 95 percent of the instances. The SSA and Census data sets were used to ensure that name-gender pairs were consistent across every data set in which they occurred and to enhance the imputation process by contributing name-gender pairs that did not occur in the Clearinghouse data. The imputation process that yielded additional gender codes produced an average gender coverage rate of 95.3% percent across all cohorts.

Neighborhood Income Measure

The neighborhood income measure provides information about the relative socioeconomic level of students' pre-college neighborhoods for students originating from the 50 U.S. states and Washington, D.C. Neighborhood income is highly correlated with other indicators of neighborhood socioeconomic status (SES) such as home ownership, educational attainment, employment, and poverty. Research suggests that [students hailing from higher SES neighborhoods have better outcomes](#) in terms of college attendance and lifetime earnings, likely due to factors such as [access to high-quality schools, high-achieving peer groups, healthier natural environments, and limited exposure to violence and the criminal justice system](#). It is important to note that this is not a measure of family or individual income. Not all students who come from high-income neighborhoods come from high-income families and the same is true of students from low-income neighborhoods.

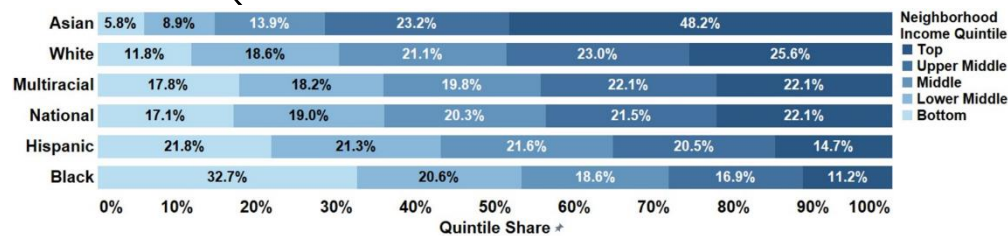
The measure utilizes street addresses reported to the Clearinghouse with each enrollment submission to locate students' homes in a particular census tract through geocoding. To best approximate socioeconomic origins, we use the first permanent address reported to the Clearinghouse for each student and apply the measure only to students who were 24 or younger at postsecondary entry (non-dual enrolled) and whose first higher education enrollment (including dual enrollment) occurred at age 19 or younger apply the measure only to undergraduate students 24 and younger whose first higher education enrollment (including dual enrollment) occurred at age 19 or younger. We apply the measure to cohorts starting in 2010 and on due to the availability of external data sources needed to construct the measure. After these three restrictions—location of the first address, age at first enrollment, and data year of the first address (after 2010)—are taken into account, we are able to include an average of 85.3 percent of the students across all cohorts in our panel for neighborhood income analyses.

Income data for each tract are sourced from the U.S. Census Bureau's American Community Survey (ACS) five-year estimates. These are adjusted using Regional Price Parity values from the Bureau of Economic Analysis to account for price level differences by state and metropolitan area. The quintiles referenced in this report are of tract median household income adjusted for household size. Quintiles are based on the national distribution of median household income, adjusted for household size, among all census tracts in the 50 states and D.C.

The vast majority (nearly 92%) of students included in the neighborhood income analyses in this report are successfully geocoded to a census tract. In this report, students with missing neighborhood income quintile values are those who have met the sample restrictions described above, but that (1) we were unable to geocode to a tract or (2) were geocoded to a tract for which ACS does not publicly publish income data. This missing rate was 8.3 percent across all cohorts. The inability to geocode (rather than geocoding to a tract without income data) accounts for nearly all of these missing cases. Geocoding non-matches typically arise from issues relating to the quality of address data such as PO Box addresses and incomplete street addresses. Address quality may be correlated with both neighborhood and household income.

For neighborhood income results by students' race/ethnicity, it is important to note that ethnoracial groups are not uniformly distributed across neighborhood income quintiles. To aid in interpreting these data, we provide a baseline estimate of the population of all 15–17-year-olds in the U.S. residing in tracts of each neighborhood income quintile for each of the race/ethnicity groups we report on, as well as for the Nation overall in figure M1.

Figure M1. National Distribution of 15–17-year-olds by Race/Ethnicity and Neighborhood Income Quintile



Source: U.S. Census Bureau American Community Survey, 2018-2022 5-year Estimates. NOTE: Age 15-17 selected to approximate high school-age and to avoid the undercounting of 18–24-year-olds living away from home while enrolled in postsecondary education. These individuals are not included in home census tract estimates.

Data Limitations

The data limitations in this report center mainly on the data coverage, the methods used for cohort identification, and the definition of key constructs (as outlined above).

Despite the challenges presented by low participation in the early years covered in this report, current Clearinghouse data nevertheless offer near-census national coverage, representing 97 percent of U.S. postsecondary enrollments. Keep in mind this report represents the completion rates for Clearinghouse institutions, not all institutions across all cohorts. The Data Coverage tables from our [Working with Our Data website](#) shows coverage by sector depending on the year in question; data should be interpreted accordingly.

Data limitations resulting from the cohort identification methods used in preparing this report should also be noted. Because the Clearinghouse data on designations for class year are incomplete, the researchers identified first-time undergraduate students via two indirect measures:

- No previous college enrollments recorded in the Clearinghouse data, and
- No previous degree awarded in the Clearinghouse’s historical DegreeVerify database.

Given these selection criteria, the sample for this report may include students who had enrollments or completions outside of Clearinghouse institutions, who had more than 30 Advanced Placement (AP) or International Baccalaureate (IB) credits. It is also possible that a small number of graduate students are included in the study cohort because of inconsistencies in the historical depth of DegreeVerify database records.

Finally, although Clearinghouse data contain some demographic information on students, historical coverage rates for the demographic data elements are uneven. Consequently, results on gender are based partially on imputed values, as described above and an average of 15.2 percent of students over the eight cohorts for which race/ethnicity data are available had the race/ethnicity value unknown or missing. No imputation was attempted for missing race/ethnicity data.

Suggested Citation

Lee, S., Berg, B., Gardner, A., Holsapple, M. & Shapiro, D. (December 2024), Yearly Progress and Completion. (Signature Report 23), Herndon, VA: National Student Clearinghouse Research Center.