



National Center for Science and
Engineering Statistics

Survey

Survey of Doctorate Recipients: 2021 | 2021

The SDR provides data on the characteristics of science, engineering, and health research doctorate degree holders from U.S. academic institutions who are under the age of 76.

Survey Description

Survey Description

Survey Overview (2021 Survey Cycle)

Purpose

The Survey of Doctorate Recipients (SDR), conducted by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation, provides data on the characteristics of science, engineering, and health (SEH) doctorate degree holders. It samples individuals who have earned an SEH research doctoral degree from a U.S. academic institution and are less than 76 years of age. The SDR provides data useful in assessing the supply and characteristics of U.S.-trained SEH doctorates employed in educational institutions, private industry, professional organizations, and government in the United States, as well as in other countries worldwide.

Data collection authority

The information is solicited under the authority of the National Science Foundation Act of 1950, as amended, the America COMPETES Reauthorization Act of 2010, and the Confidential Information Protection and Statistical Efficiency Act of 2018. The Office of Management and Budget control number is 3145-0020.

Major changes to recent survey cycle

The 2021 SDR made two types of changes to the data collection instruments. First, for all modes of data collection, the survey included new questions to gauge the effects of the coronavirus pandemic on employment, specifically on labor force status, number of hours worked per week, salary, benefits, telecommuting options, and total earned income. The second change applied to the electronic instruments only. The Web and computer-assisted telephone interview (CATI) instruments included dependent interviewing (DI) methods for a targeted number of items within the employment question series to reduce respondent burden.¹

Key Survey Information

Frequency	Biennial.
Initial survey year	1973.
Reference period	The week of 1 February 2021.
Response unit	Individuals with an SEH research doctorate degree from a U.S. academic institution.
Sample or census	Sample.
Population size	Approximately 1,185,700 individuals.
Sample size	A total of 125,938 individuals.
Key variables	<ul style="list-style-type: none"> Demographics (e.g., age, race, sex, ethnicity, and citizenship) Educational history Employment status

Key variables

- Demographics (e.g., age, race, sex, ethnicity, and citizenship)
- Educational history
- Employment status

- Field of degree
- Occupation

Survey Design

Target population

The SDR target population includes individuals that meet the following criteria:

- Earned an SEH research doctoral degree from a U.S. academic institution prior to 1 July 2019
- Were not institutionalized or terminally ill on 1 February 2021
- Were less than 76 years of age as of 1 February 2021

Sampling frame

The Doctorate Records File (DRF) constructed from the annual Survey of Earned Doctorates, which is a census survey of all recipients of U.S. research doctoral degrees.

Sample design

The SDR uses a fixed panel design with a sample of new doctoral graduates added to the panel in each biennial survey cycle. For the 2021 SDR, all 2019 sample members who remained age eligible were retained for the 2021 cycle. As with prior survey cycles, a sample of 10,000 new graduates who had earned their degrees since the last SDR survey cycle, from 1 July 2017 to 30 June 2019, was added. The new graduates sample design followed the same sample design and sample stratification first introduced in 2019, defined by detailed fields of study, gender, and underrepresented minority status.

Data Collection and Data Processing

Data collection

The SDR uses a trimodal data collection approach: self-administered online survey, self-administered paper questionnaire (via mail), and CATI.

Data processing

The data collected in the SDR are subject to both editing and imputation procedures. The SDR uses both logical imputation and statistical (hot-deck) imputation as part of the data processing effort.

Estimation techniques

Because the SDR is based on a complex sampling design and subject to nonresponse bias, sampling weights are created for each respondent to support unbiased population estimates. The final analysis weights account for the following:

- Differential sampling rates
- Adjustments for unknown eligibility
- Adjustments for nonresponse
- Adjustments to align the sample distribution with the population distribution with respect to gender, race and ethnicity, location, degree year, and degree field data of the DRF.

Survey Quality Measures

Sampling error

Estimates of sampling errors associated with this survey were calculated using replicate weights.

Coverage error

Any missed doctoral graduates within the DRF derived from the SED would create undercoverage in the SDR. Reporting errors in the SED could lead to incorrect classification of doctorates as having or not having earned an SEH research doctorate, which could result in either overcoverage or undercoverage.

Nonresponse error

The weighted and unweighted response rates for the 2021 SDR were each 65%. Analyses of SDR nonresponse trends were used to develop nonresponse weighting adjustments to minimize the potential for nonresponse bias in the SDR estimates. A hot-deck imputation method was used to compensate for item nonresponse.

Measurement error

The SDR is subject to reporting errors from differences in interpretation of questions. Although three modes of response were offered (Web, mail, and CATI), 99% of sample members chose to respond via the Web instrument. As such, reporting error due to mode differences is significantly diminished.

Data Availability and Comparability

Data availability

Data from 1993 to present are available at the SDR website, <https://www.nsf.gov/statistics/srvydoctoratework/>.

Data comparability

Year-to-year comparisons can be made among the 1993 to 2021 survey cycles because many of the core questions remained the same. Small but notable differences exist across some survey years, such as the collection of occupation data based on more recent versions of the occupation taxonomy. Also, the SDR target population definition has changed over time as follows:

- Survey data prior to 2010 did not cover SEH doctorates residing outside of the United States.
- In 2010 and 2013, full coverage of SEH doctorates residing outside of the United States included only those having graduated since 2001. For graduates from earlier years, the coverage of those residing outside of the United States is partial.
- The 2015 SDR sample design improved population coverage in the 2015 and subsequent survey cycles to include all SEH doctorates awarded by U.S. institutions, regardless of the academic year of award or the recipient's post-graduation residency location.

Caution is recommended when interpreting or analyzing trends that span pre- and post-1991 surveys, pre- and post-2010 surveys, and pre- and post-2015 surveys given the noted changes in the survey design and target population.

- Overlap in sample cases across survey cycles support longitudinal analysis using SDR data. A longitudinal panel representing a cohort of SEH doctorate recipients awarded their degree prior to July 2013 and aged less than 66 years in 2015 was selected, and an initial longitudinal data file with imputation and weights accurately reflecting the longitudinal design was developed. The 2015–19 survey data for this panel is available at [Survey of Doctorate Recipients, Longitudinal Data: 2015–19](#).

Data Products

Publications

Data from the SDR are published in NCSSES InfoBriefs and data tables, available at <https://www.nsf.gov/statistics/srvydoctoratework/>. Information from this survey is also included in *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

Electronic access

The SDR public use data are available in the [SESTAT data tool](#) and in downloadable files through the [NCSSES data page](#). Access to restricted data for researchers interested in analyzing microdata can be arranged through a licensing agreement. For more information on licensing, see <https://nces.nsf.gov/about/licensing>.

Note

1 The Web and CATI instruments included DI methods for a targeted number of items within the employment question series. With DI, sample member responses from 2019 were preloaded into the 2021 SDR questionnaire and displayed for the respondent. For each of the DI questions, sample members first answered “yes” or “no” to indicate if the information displayed from their 2019 response still applied to the 2021 reference period. If not, the sample member provided updated information on the subsequent screen. Only sample members who participated in 2019 and reported working in both the 2019 and 2021 cycles were eligible for DI.