INFORMATION ONLY – DO NOT USE TO REPORT

Guide to Public-Use Data Files

Survey of Graduate Students and Postdoctorates in Science and Engineering: Fall 2021

The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) collects count data for graduate students (separately for master's and doctoral students), postdoctoral appointees (postdocs), and doctorate-holding nonfaculty researchers (NFRs) in science, engineering, and selected health units.

Graduate student counts are collected by enrollment level, enrollment status, sex, citizenship, and race and ethnicity (2021 questions 2 and 3). Full-time graduate student counts are also collected by primary source and mechanism of financial support (2021 question 4).

The 2021 GSS public-use data files are organized so that the institution, school, and GSS code–level data are available for each year all within a single record.[1] The data also include the Integrated Postsecondary Education Data System (IPEDS) UNITID, which allows researchers the ability to link the GSS institutions to other institutional data sources.

Public-use data are available for download by year. Key variables are also available for trend data tabulations through the National Centers for Science and Engineering Statistics (NCSES) interactive tool (https://ncsesdata.nsf.gov/ids/).

DATA FILE FORMATS

The GSS public-use file is available in Excel and SAS formats.

The files available for download are as follows (YYYY stands for 1972–2021 GSS data collection years):

- gssYYYYc_xlsx.zip files – contain gssYYYY_code.xlsx Excel files
- gssYYYYc_sas.zip files – contain gssYYYY_code.sas7bdat SAS data sets

For compatibility issues with prior-round data, Excel files are divided by question number into three worksheets:

- Race (questions 2 and 3) – pt_tot_all_races_v … ft_frst_wmen_unknown_v
- Support (question 4) – ft_tot_all_srces_v ... ft_oth_mech_self_sup_v
- Postdoc (questions 5 and 6) – pd_tot_all_srces_v ... med_degr_oth_non_fcty_v

As a convenience, the institution, school, and GSS code–level data are replicated for each associated record rather than providing separate institution, school, and GSS code–level files that would need to be merged together. The data contain the institution, school, and GSS code–level data gathered in question 1, and the variables are placed as leading variables in the SAS files and columns in the Excel

[1] The GSS public-use data files are also available at the GSS organizational unit level by contacting the GSS project officer.

files. SAS and Excel files are available by data collection year, and these files include all variables (institution, school, GSS code, part-time race, full-time race, support, and postdoc and NFR data all on one record). Starting in 2017, data are also available separately for master's- and doctoral-level enrollment.

DATA FILE IDENTIFIERS

IDs are created to uniquely identify institutions and schools.

The ID structure is the following:

- year – identifies the year in which the data are collected
- institution_id – identifies the institution
- school_id – identifies a school that is unique across all institutions and not just a sequence within institution
- gss_code – identifies the type of program and is unique within a school

The combination of year, school_id, and gss_code forms a unique record across the GSS data files.

As indicated in the ID structure, institutions can have more than one school—thus, there is the need for the school_id variable. The school name is reported as the name of the school within the institution and is not unique across all schools. A unique name for the school containing the institution name and school name is available in the variable "full_school_name."

In 2014, the GSS collected data from 708 institutions, including 151 institutions that were added to the survey universe as the result of a comprehensive frame evaluation and 2 institutions that were dropped as the result of this frame evaluation. The final 2014 data file includes data from the 706 institutions eligible for the survey at the end of the 2014 cycle. Due to the frame expansion, these data are not comparable to prior-year data. To enable trend comparisons with 2013 and prior years, an additional "2014old" data file is available that contains the 2014 data from the 557 institutions eligible for the survey prior to the frame update. The "year" variable is coded to 2014old in the supplemental file and to 2014 in the main, final 2014 data file. More detailed information about the frame expansion can be found in a Special Report, *Assessing the Impact of Frame Changes on Trend Data from the Survey of Graduate Students and Postdoctorates in Science and Engineering* (https://www.nsf.gov/statistics/2016/nsf16314/).

In 2017, the GSS made significant changes to how the data were collected: it revised the taxonomy used to classify disciplines and added separate variables for master's and doctoral students. Due to these changes, the 2017 data are not comparable to prior years. A set of bridge estimates was created to permit comparisons to previous years and for trend analyses. These estimates are labeled 2017old and are available at the broad-field level for all combined graduate student variables and for postdoc variables. Since 2017old codes are at the broad-field level, fewer records are in 2017old than in 2016 or 2017new. For more information about the updated GSS taxonomy and the changes to the 2017 GSS and 2017old, including how to use 2017old, see appendix C, "Taxonomy Change Documentation," in the GSS 2018 Methodology Report.

The following summary table is provided to help users confirm that they have imported the data files properly for a given year. It enumerates the number of institutions, schools, and GSS code–level records that were included in the GSS for the given year.

Year	Number of unique entities		
	Institutions	Schools	Records
1972	259	328	4,079
1973	262	340	5,688
1974	284	375	6,384
1975	584	682	7,741
1976	594	693	7,815
1977	601	704	8,007
1978	599	708	8,088
1979	629	745	8,252
1980	626	742	8,306
1981	622	736	8,240
1982	609	724	8,134
1983	609	723	8,015
1984	412	530	7,388
1985	412	525	7,437
1986	412	527	7,486
1987	416	533	7,575
1988	606	723	8,416
1989	609	726	8,515
1990	610	727	8,597
1991	609	726	8,719
1992	608	725	8,851
1993	606	723	8,969
1994	605	722	9,092
1995	603	720	9,184
1996	603	720	9,140
1997	601	722	9,162
1998	601	721	9,198
1999	599	719	9,267
2000	596	716	9,260
2001	601	720	9,302
2002	596	715	9,345
2003	593	712	9,377
2004	591	710	9,302
2005	588	702	9,225

Year	Number of unique entities		
	Institutions	Schools	Records
2006	588	707	9,271
2007	582	700	9,563
2008	579	708	9,782
2009	575	703	9,909
2010	574	692	10,047
2011	565	686	10,041
2012	565	684	10,115
2013	564	680	10,143
2014old	557	671	10,233
2014	706	821	10,666
2015	711	824	10,830
2016	714	828	10,892
2017old	713	827	4,923
2017	703	814	11,303
2018	715	817	11,604
2019	714	809	11,804
2020	712	806	12,832
2021	699	787	12,905

Users can analyze data across GSS data collection years by concatenating the GSS data across multiple years to create a longitudinal data set. The "year" variable, which indicates the GSS data collection year, will need to be used as a key variable in the ID structure. The following summary table is provided to help users confirm that they have concatenated data properly across years. It enumerates the number of institutions, schools, and GSS code–level records that were ever included in the GSS.

Years	Number of unique entities		
	Institutions	Schools	Records
1972–2021	890	1,079	453,263

INTEGRATED POSTSECONDARY EDUCATION DATA SYSTEM (IPEDS) UNITID

One feature that should help facilitate analysts' use of the data is the inclusion of the IPEDS UNITID. The IPEDS UNITID is linked to the School ID. The 2020 version of IPEDS is the latest version that was used to link to the schools. For convenience, UNITID is replicated in the file across years for the same school. If schools are not reported in the 2020 IPEDS file, the UNITID field is filled with a reserve code value of "999999."

DATA ITEMS

The data variables collected over the years have consisted primarily of full- and part-time graduate students and postdocs, with detailed information on full-time student source (federal, nonfederal, other) and mechanism (e.g., fellowships, traineeships, research assistantships, or teaching assistantships) of primary support, sex, field of study, and citizenship. Starting with 2017 data, all graduate student data are available separately for master's and doctoral graduate students.

The collection of full- and part-time graduate student race and ethnicity data was introduced as an optional item in the 1979 survey and became a standard survey item in 1980. Data on the sex of part-time graduate students have been collected since 1977. Race and ethnicity data separated by sex have been collected for full- and part-time graduate students since 1993 and for first-time, full-time students since 1999.

Prior to 2017, graduate student data were collected as combined master's and doctoral data. Starting in 2017, master's and doctoral student data were separated and are available to data users as distinct variables for master's and doctoral students as well as combined for all graduate students.

Postdoc and NFR data were collected as combined counts until 1977, collected separately starting in 1979, and expanded substantially in 2010. See appendix A for more details.

Prospective data users should note that data items and types of institutions surveyed have varied over the years of the survey. For example, prior to 1992, permanent residents and temporary visa holders were collected as combined counts; from 1992 through the present, permanent residents are counted with U.S. citizens.

Major barriers for trend analysis using the GSS data are that the GSS institutions changed substantially over time and that improvements in data collection methodology can cause trend breaks. See "Historical Changes in the GSS" section for more details.

The variable list is grouped similarly to the questions presented in the current GSS survey worksheet.

- Variables 1–22 – institution, school, and GSS code–level data
- Variables 23–52 – part-time graduate student data (combined master's + doctoral student counts)
- Variables 53–82 – full-time graduate student data (combined master's + doctoral student counts)
- Variables 83–112 – first-time, full-time graduate student data (combined master's + doctoral student counts)
- Variables 113–226 – source and mechanism of primary support for full-time graduate student data (combined master's + doctoral student counts)
- Variables 227–56 – part-time master's student data
- Variables 257–86 – full-time master's student data
- Variables 287–316 – first-time, full-time master's student data
- Variables 317–430 – source and mechanism of primary support for full-time master's student data
- Variables 431–60 – part-time doctoral student data
- Variables 461–90 – full-time doctoral student data

- Variables 491–520 – first-time, full-time doctoral student data
- Variables 521–634 – source and mechanism of primary support for full-time doctoral student data (doctoral student counts)
- Variables 635–787 – postdoctoral appointee data
- Variables 788–802 – NFR data

Prospective data users should note that graduate student data were collected for part-time students (question 2) and then for full-time students (question 3). To obtain counts for graduate students (regardless of full-time or part-time), users must sum the pt_* and ft_* variable fields.

The data collected on the number of postdocs and NFRs in GSS-eligible units were expanded in 2010, and significant effort was made to ensure that appropriate personnel were providing these data. Overall counts for both groups also increased substantially in 2010. It is unclear how much of these increases represent actual growth in postdocs and NFRs and how much results from improved data collection. More information on the improved data collection and changes in postdoc data can be found in an InfoBrief, *Counts of Postdoctoral Appointees in Science, Engineering, and Health Rise with Reporting Improvements*, available at https://www.nsf.gov/statistics/infbrief/nsf13334/ and in NFR data in a Working Paper, *Examining the Reporting of Nonfaculty Doctorate Researchers in the Survey of Graduate Students and Postdoctorates in Science and Engineering,* at https://www.nsf.gov/statistics/2015/ncses15201/. Additionally, due to the changes in 2017 data collection, many institutions reported more NFRs due to underreporting in the past. Therefore, the trend data for both postdocs and NFRs should be used with caution.

The table in appendix B lists the data items, labels, and data type available in the GSS files.

Historical values for the following institution-level data items—hdg_inst, toc_code, hbcu_flag, land_grant_flag, and Carnegie_code—were overwritten and replaced with the most recent value for the institution through 2006. That is, the variables will be the 2006 value for all years 1972–2006. Beginning in 2007, these items have been updated if there is a change for that year, and historical values will not be overwritten.

The following code labels are provided for discrete items indicated in the data item table above.

hdg_inst — Institution highest degree
1 = Doctorate-granting (at least one science and engineering [S&E] doctoral organizational unit)
2 = Master's-granting

toc_code — Institution type of control
1 = Public
2 = Private

institution_state — Institution state code (a two-character state abbreviation)

AL = Alabama	MT = Montana
AK = Alaska	NE = Nebraska
AZ = Arizona	NV = Nevada
AR = Arkansas	NH = New Hampshire
CA = California	NJ = New Jersey
CO = Colorado	NM = New Mexico

CT = Connecticut
DE = Delaware
DC = District of Columbia
FL = Florida
GA = Georgia
GU = Guam
HI = Hawaii
ID = Idaho
IL = Illinois
IN = Indiana
IA = Iowa
KS = Kansas
KY = Kentucky
LA = Louisiana
ME = Maine
MD = Maryland
MA = Massachusetts
MI = Michigan
MN = Minnesota
MS = Mississippi
MO = Missouri
NY = New York
NC = North Carolina
ND = North Dakota
OH = Ohio
OK = Oklahoma
OR = Oregon
PA = Pennsylvania
PR = Puerto Rico
RI = Rhode Island
SC = South Carolina
SD = South Dakota
TN = Tennessee
TX = Texas
UT = Utah
VT = Vermont
VI = Virgin Islands
VA = Virginia
WA = Washington
WV = West Virginia
WI = Wisconsin
WY = Wyoming

hbcu_flag — Historically black college or university flag
0 = no
1 = yes

land_grant_flag — Land grant flag
0 = no
1 = yes

carnegie_code_1994 — Carnegie Code 1994 Classification
R1—Research Universities I
R2—Research Universities II
D1—Doctoral Universities I
D2—Doctoral Universities II
C1—Master's (Comprehensive) Universities and Colleges I
C2—Master's (Comprehensive) Universities and Colleges II
LA1—Baccalaureate (Liberal Arts) Colleges I
LA2—Baccalaureate (Liberal Arts) Colleges II
2YR—Associate of Arts Colleges
ART—Schools of Art, Music, and Design
BUS—Schools of Business and Management
ENG—Schools of Engineering and Technology
HLT—Other Separate Health Profession Schools
LAW—Schools of Law
MED—Medical Schools and Medical Centers

REL—Theological Seminaries, Bible Colleges, and Other Institutions Offering Degrees in Religion
TEA—Teachers Colleges
TRI—Tribal Colleges and Universities
OTH—Other Specialized Institutions
N/A—Not Classified

carnegie_code_2005 Carnegie Code 2005 Classification
carnegie_code_2010 Carnegie Code 2010 Classification

-1	(Not applicable)
0	(Not classified)
1	Assoc/Pub-R-S: Associate's—Public rural-serving small
2	Assoc/Pub-R-M: Associate's—Public rural-serving medium
3	Assoc/Pub-R-L: Associate's—Public Rural-serving Large
4	Assoc/Pub-S-SC: Associate's—Public Suburban-serving Single Campus
5	Assoc/Pub-S-MC: Associate's—Public Suburban-serving Multicampus
6	Assoc/Pub-U-SC: Associate's—Public Urban-serving Single Campus
7	Assoc/Pub-U-MC: Associate's—Public Urban-serving Multicampus
8	Assoc/Pub-Spec: Associate's—Public Special Use
9	Assoc/PrivNFP: Associate's—Private Not-for-profit
10	Assoc/PrivFP: Associate's—Private For-profit
11	Assoc/Pub2in4: Associate's—Public 2-year colleges under 4-year universities
12	Assoc/Pub4: Associate's—Public 4-year Primarily Associate's
13	Assoc/PrivNFP4: Associate's—Private Not-for-profit 4-year Primarily Associate's
14	Assoc/PrivFP4: Associate's—Private For-profit 4-year Primarily Associate's
15	RU/VH: Research Universities (very high research activity)
16	RU/H: Research Universities (high research activity)
17	DRU: Doctoral/Research Universities
18	Master's L: Master's Colleges and Universities (large programs)
19	Master's M: Master's Colleges and Universities (medium programs)
20	Master's S: Master's Colleges and Universities (small programs)
21	Bac/A&S: Baccalaureate Colleges—Arts & Sciences
22	Bac/Diverse: Baccalaureate Colleges—Diverse Fields
23	Bac/Assoc: Baccalaureate/Associate's Colleges
24	Spec/Faith: Special Focus Institutions—Theological seminaries, Bible colleges, and other faith-related institutions
25	Spec/Med: Special Focus Institutions—Medical schools and medical centers
26	Spec/Health: Special Focus Institutions—Other health professions schools
27	Spec/Eng: Special Focus Institutions—Schools of engineering
28	Spec/Tech: Special Focus Institutions—Other technology-related schools
29	Spec/Bus: Special Focus Institutions—Schools of business and management
30	Spec/Arts: Special Focus Institutions—Schools of art, music, and design
31	Spec/Law: Special Focus Institutions—Schools of law
32	Spec/Other: Special Focus Institutions—Other special-focus institutions
33	Tribal: Tribal Colleges

carnegie_code_2015 Carnegie Code 2015 Classification

-2	Not applicable, not in Carnegie universe (not accredited or non-degree-granting)
1	Associate's Colleges: High Transfer-High Traditional
2	Associate's Colleges: High Transfer-Mixed Traditional/Nontraditional
3	Associate's Colleges: High Transfer-High Nontraditional
4	Associate's Colleges: Mixed Transfer/Career & Technical-High Traditional
5	Associate's Colleges: Mixed Transfer/Career & Technical-Mixed Traditional/Nontraditional
6	Associate's Colleges: Mixed Transfer/Career & Technical-High Nontraditional

7	Associate's Colleges: High Career & Technical-High Traditional
8	Associate's Colleges: High Career & Technical-Mixed Traditional/Nontraditional
9	Associate's Colleges: High Career & Technical-High Nontraditional
10	Special Focus Two-Year: Health Professions
11	Special Focus Two-Year: Technical Professions
12	Special Focus Two-Year: Arts & Design
13	Special Focus Two-Year: Other Fields
14	Baccalaureate/Associate's Colleges: Associate's Dominant
15	Doctoral Universities: Highest Research Activity
16	Doctoral Universities: Higher Research Activity
17	Doctoral Universities: Moderate Research Activity
18	Master's Colleges & Universities: Large Programs
19	Master's Colleges & Universities: Medium Programs
20	Master's Colleges & Universities: Small Programs
21	Baccalaureate Colleges: Arts & Sciences Focus
22	Baccalaureate Colleges: Diverse Fields
23	Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's
24	Special Focus Four-Year: Faith-Related Institutions
25	Special Focus Four-Year: Medical Schools & Centers
26	Special Focus Four-Year: Other Health Professions Schools
27	Special Focus Four-Year: Engineering Schools
28	Special Focus Four-Year: Other Technology-Related Schools
29	Special Focus Four-Year: Business & Management Schools
30	Special Focus Four-Year: Arts, Music & Design Schools
31	Special Focus Four-Year: Law Schools
32	Special Focus Four-Year: Other Special Focus Institutions
33	Tribal Colleges

carnegie_code_2018 Carnegie Code 2018 Classification

-2	Not applicable, not in Carnegie universe (not accredited or non-degree-granting)
1	Associate's Colleges: High Transfer-High Traditional
2	Associate's Colleges: High Transfer-Mixed Traditional/Nontraditional
3	Associate's Colleges: High Transfer-High Nontraditional
4	Associate's Colleges: Mixed Transfer/Vocational & Technical-High Traditional
5	Associate's Colleges: Mixed Transfer/Vocational & Technical-Mixed Traditional/Nontraditional
6	Associate's Colleges: Mixed Transfer/Vocational & Technical-High Nontraditional
7	Associate's Colleges: High Vocational & Technical-High Traditional
8	Associate's Colleges: High Vocational & Technical-Mixed Traditional/Nontraditional
9	Associate's Colleges: High Vocational & Technical-High Nontraditional
10	Special Focus Two-Year: Health Professions
11	Special Focus Two-Year: Technical Professions
12	Special Focus Two-Year: Arts & Design
13	Special Focus Two-Year: Other Fields
14	Baccalaureate/Associate's Colleges: Associate's Dominant
15	Doctoral Universities: Very High Research Activity
16	Doctoral Universities: High Research Activity
17	Doctoral/Professional Universities
18	Master's Colleges & Universities: Large Programs
19	Master's Colleges & Universities: Medium Programs
20	Master's Colleges & Universities: Small Programs
21	Baccalaureate Colleges: Arts & Sciences Focus
22	Baccalaureate Colleges: Diverse Fields
23	Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's

24 Special Focus Four-Year: Faith-Related Institutions
25 Special Focus Four-Year: Medical Schools & Centers
26 Special Focus Four-Year: Other Health Professions Schools
27 Special Focus Four-Year: Engineering Schools
28 Special Focus Four-Year: Other Technology-Related Schools
29 Special Focus Four-Year: Business & Management Schools
30 Special Focus Four-Year: Arts, Music & Design Schools
31 Special Focus Four-Year: Law Schools
32 Special Focus Four-Year: Other Special Focus Institutions
33 Tribal Colleges

hhe_flag — High Hispanic Enrollment Flag

0 = no

1 = yes

The High Hispanic Enrollment Flag is based on IPEDS enrollment data and indicates that Hispanic students represent at least 25% of the undergraduate full-time equivalent (FTE) enrollment at one or more of the IPEDS reporting units within the GSS institution. IPEDS is conducted annually by the National Center for Education Statistics. FTE enrollment is calculated by equating three part-time students to one full-time student.

school_type_code — School Type

G (Graduate) = Graduate Schools
M (Medical) = AAMC Member Medical Schools
N (Nursing) = Schools of Nursing
O (Osteopathic) = Schools of Osteopathic Medicine
P (Public) = Schools of Public Health
D (Dentistry) = Schools of Dentistry
V (Veterinary) = Schools of Veterinary Medicine
H (Hospital) = Hospital-Affiliated Medical Research Centers
A (Allied) = Schools of Professional or Allied Health
B (Both) = Schools with Programs from Both a Graduate School and an AAMC Member Medical School
U (Unique) = Other Health Sciences Schools

gss_code — GSS Degree Program (Field of Study) Code

Aerospace, Aeronautical, and Astronautical Engineering
101 Aerospace, Aeronautical, and Astronautical Engineering

Biological, Biomedical, and Biosystems Engineering
120 Biological, Biomedical, and Biosystems Engineering

Chemical, Petroleum, and Chemical-Related Engineering
104 Chemical Engineering
113 Petroleum Engineering

Civil, Environmental, Transportation and Related Engineering Fields
117 Architectural, Environmental, Construction and Surveying Engineering
105 Civil Engineering

Electrical, Electronics, Communications and Computer Engineering
118 Computer Engineering
106 Electrical, Electronics, and Communications Engineering

gss_code GSS Degree Program (Field of Study) Code

Industrial, Manufacturing, Systems Engineering and Operations Research
- 108 Industrial and Manufacturing Engineering
- 119 Systems Engineering and Operations Research

Mechanical Engineering
- 109 Mechanical Engineering

Metallurgical, Mining, Materials and Related Engineering Fields
- 121 Metallurgical, Mining, Materials and Related Engineering Fields

Other Engineering
- 102 Agricultural engineering
- 107 Engineering Mechanics, Physics, and Science
- 112 Nuclear Engineering
- 122 Engineering, other

Physical Sciences
- 201 Astronomy and Astrophysics
- 202 Chemistry
- 205 Materials Sciences
- 203 Physics
- 204 Physical Sciences, not elsewhere classified

Geosciences, Atmospheric Sciences, and Ocean Sciences
- 301 Atmospheric Sciences and Meteorology
- 302 Geological and Earth Sciences
- 303 Ocean and Marine Sciences
- 304 Geoscience, Atmospheric Sciences, and Ocean Sciences, not elsewhere classified

Mathematics and Statistics
- 404 Applied Mathematics
- 405 Mathematics
- 403 Statistics

Computer and Information Sciences
- 416 Artificial Intelligence, Informatics and CIS Topics
- 410 Computer Science
- 411 Computer and Information Sciences, general
- 413 Computer and Information Systems Security
- 415 Information Science and Studies
- 414 Information Technology
- 412 Computer and Information Sciences, not elsewhere classified

Agricultural and Veterinary Sciences
- 501 Agricultural Sciences
- 502 Veterinary Biomedical and Clinical Sciences

Natural Resources and Conservation
- 510 Environmental Science and Studies
- 511 Forestry, Natural Resources and Conservation

gss_code	GSS Degree Program (Field of Study) Code
	Biological and Biomedical Sciences
	602 Biochemistry
	603 Biology
	623 Biomedical Sciences
	605 Biophysics
	618 Biostatistics and Bioinformatics
	624 Biotechnology
	606 Botany and Plant Biology
	619 Cell, Cellular Biology and Anatomical Sciences
	620 Ecology and Population Biology
	621 Epidemiology
	610 Genetics
	611 Microbiological Sciences and Immunology
	622 Molecular Biology
	626 Neurobiology and Neuroscience
	612 Nutrition Science
	613 Pathology and Experimental Pathology
	614 Pharmacology and Toxicology
	615 Physiology
	616 Zoology and Animal Biology
	617 Biological and Biomedical Sciences, not elsewhere classified
	Health – Clinical Medicine
	701 Anesthesiology
	702 Cardiology and Cardiovascular Disease
	704 Endocrinology, Diabetes, and Metabolism
	705 Gastroenterology
	706 Hematology
	729 Medical Clinical Sciences and Clinical and Medical Laboratory Sciences
	707 Neurology and Neurosurgery
	708 Obstetrics and Gynecology
	703 Oncology and Cancer Research
	709 Ophthalmology
	710 Otorhinolaryngology
	711 Pediatrics
	713 Psychiatry
	712 Public Health
	714 Pulmonary Disease
	715 Radiological Sciences
	716 Surgery
	717 Clinical Medicine, not elsewhere classified
	Health – Other Health
	723 Communication Disorders Sciences
	718 Dental Sciences
	724 Kinesiology and Exercise Science
	719 Nursing
	720 Pharmaceutical Sciences
	722 Health-Related, not elsewhere classified

gss_code GSS Degree Program (Field of Study) Code

Psychology
- 803 Clinical Psychology
- 806 Counseling Psychology
- 804 Applied Psychology
- 815 Human Development
- 801 Psychology, General
- 805 Research and Experimental Psychology

Social Sciences
- 901 Agricultural Economics
- 902 Anthropology
- 916 Area, Ethnic, Culture, Gender, and Group Studies
- 911 Criminal Justice and Safety Studies
- 917 Criminology
- 903 Economics
- 904 Geography and Cartography
- 912 International Relations and National Security Studies
- 906 Linguistics
- 907 Political Science and Government
- 914 Public Policy Analysis
- 908 Sociology and Demography
- 918 Urban Studies and Affairs
- 919 Social Sciences, other

Multidisciplinary and Interdisciplinary Studies
- 982 Biological and Physical Sciences
- 981 Computational Science
- 984 Data Science and Data Analytics
- 983 International and Global Studies
- 980 Multidisciplinary and Interdisciplinary Studies

hdg_code Highest Degree Program Offered by GSS Code (may change from year to year)
- 1 = Doctorate-granting (at least one S&E doctoral organizational unit)
- 2 = Master's-granting
- 3 = No PhD or Master's Equivalent (only applicable to postdoc organizational unit)

HISTORICAL CHANGES IN THE GSS

Changes have been made to the coverage and content of the GSS to keep it relevant to the needs of data users. Such changes prevent precise maintenance of trend data; therefore, some data items are not available for all institutions in all years. Major changes in the data collected (with the year in which changes became effective) are available in the "Technical Notes" section of the *Graduate Students and Postdoctorates in S&E* Data Tables on the NCSES website at https://www.nsf.gov/statistics/gradpostdoc/.

DATA ELEMENT AVAILABILITY BY YEAR

Graduate student data are available annually for the years 1972–2021. Tables in appendix A show the types of data available in the GSS by survey year and the changes to the data item categories used to collect citizenship, race and ethnicity, source of support, and mechanism of support. Availability on the public-use data file is indicated by an "X." These tables do not show changes to the GSS coverage of institutions or field.

DATA IMPUTATION

The 2021 GSS collected 543 data items. Of the 543 data items collected in the GSS, the item nonresponse rates ranged from 1.68% to 6.43%. All missing data were imputed.

More information on the imputation is available in the "Technical Notes" section of the *Graduate Students and Postdoctorates in S&E* Data Tables on the NCSES website at https://www.nsf.gov/statistics/gradpostdoc/.

Imputation status codes are available on request in a supplemental data set that contains records at the organizational-unit level for each year of the study.

DATA CONFIDENTIALITY

Data from the GSS are not considered confidential because all data are based on aggregate counts for the unit.

SURVEY QUESTIONNAIRE

Copies of survey questionnaires for fall 1996–fall 2021 are available on the NCSES website at https://www.nsf.gov/statistics/questionnaires.cfm#EducationofScientistsandEngineers. Data users requiring information for earlier years can contact the National Science Foundation, GSS project officer.

In addition to reviewing the instructions and definitions provided in the questionnaire, data analysts are advised to use the most recent versions of the data for all years for trend data analyses.