# Attachment #1: Data Requirements

**Date:** [please use MM/DD/YYYY format]

**The following files are available for licensing from the National Center for Science and Engineering Statistics (NCSES). Please select the restricted use data files that you would like to license**:

## Doctorate Records Files

Each year, the data from the Survey of Earned Doctorates (SED) are compiled into the Doctorate Records File (DRF). The DRF is a historical record of doctorate-degree graduates. Begun in 1920, the DRF contains annual information used to track the number of graduates in various files and other information related to the U.S. doctorate earning population.

[ ]  DRF (through 2020)

## Early Career Doctorate Survey

The Early Career Doctorates Survey (ECDS) gathers in-depth information about individuals who earned their first doctoral degree (PhD, MD, or equivalent) in the past 10 years and work at academic institutions and federally funded research and development centers. Unique in scope, the ECDS includes professional and research doctorate holders from all fields trained in the United States and abroad. Sponsored by NCSES and by the National Institutes of Health, the ECDS provides new data on the demographics, work experiences, and career paths of individuals in the first years following the completion of their doctoral studies. Data for the ECDS survey is available for the 2013 pilot study and the 2017 full survey.

[ ]  ECDS

## Scientists and Engineers Statistical Data System

SESTAT is the Scientists and Engineers Statistical Data System. This integrated system is a unique source of longitudinal information on the education and employment of the college-educated U.S. science and engineering workforce. The data are collected through biennial surveys:

* National Survey of College Graduates

The National Survey of College Graduates (NSCG) is a repeated cross-sectional biennial survey that provides data on the nation’s college graduates, with a focus on those in the science and engineering workforce. This survey is a unique source for examining the relationship of degree field and occupation in addition to other characteristics of college-educated individuals, including work activities, salary, and demographic information. NSCG data files are available for years 2013 and beyond.

Please note that the Secure Access Data Facility (SDAF) contains a reduced version of the NSCG restricted-use data. The SDAF version of the file is a publicly available dataset with a unique case identifier for each respondent (REFID) appended to the file. This allows approved researchers to merge different cycles of the NSCG survey and track respondents longitudinally. The full version of the NSCG restricted-use data may be obtainedthrough the Census Bureau’s [Federal Statistical Research Data Centers (FSRDC)](http://www.census.gov/fsrdc).

* National Survey of Recent College Graduates (discontinued after 2010)

The National Survey of Recent College Graduates (NSRCG) was a survey conducted from 1973 to 2010. It was a cross-sectional biennial survey that provided demographic and career information about individuals holding a bachelor’s or master’s degree in a science, engineering, or health field from a U.S. academic institution. Restricted use data for the NSRCG survey is available from 1993 until 2010 when the survey was discontinued.

* Survey of Doctorate Recipients

The Survey of Doctorate Recipients (SDR) is a longitudinal biennial survey conducted since 1973 that provides demographic career history information about individuals with a research doctoral degree in a science, engineering, or health field from a U.S. academic institution. The survey follows a sample of individuals throughout their careers from the year of their degree until age 76. The panel is refreshed each survey cycle with a sample of new doctorate earners. Restricted use data for the SDR is available starting in 1993.

The SDR uses the DRF (Doctorate Records File; referenced above) as its sampling frame. The SDR\_DRF datasets contain the DRF variables for the respondents selected into the SDR sample. Individuals can be linked across the two files to get a complete picture (i.e., demographic and career outcomes) of the respondent. The SDR\_DRF files are available starting in 1993. International Survey of Doctorate Recipients (ISDR) files are available for years 2010 and 2013. After the 2013 cycle, ISDR respondents are contained in the SDR file.

The ISDR file contains data for individuals who earned doctorate degrees in the United States but are now predicted to be living internationally.

* Integrated File for Scientists and Engineers Statistical Data System

The Scientists and Engineers Statistical Data System (SESTAT) is the result of the integration of the NSCG, NSRCG (through 2010 when the survey was discontinued), and SDR files. SESTAT provides a comprehensive picture of the number and characteristics of individuals in the United States with a bachelor's or higher-level degree, with a focus on those having science and engineering (S&E) degrees or working in S&E occupations.

**Please list 10–20 restricted-use variables by name from the files that you selected that you plan to use in your research project.**

# Attachment #2: Research Plan

**Date:** (please use MM/DD/YYYY format)

## Instructions

*For attachment #2, the following elements are requested:*

1. *A short description of your research.*

The detail required, for example, in a grant proposal is not necessary. Rather, we just need to get a summary of what you intend to do with the data so that we may evaluate the appropriateness of the data for your research. If you intend to use the licensed data for multiple projects, please describe each project separately. Once the license is issued, you are always at liberty to add additional data or projects.

1. *A project summary.*

The summary must tie to the data requirements you listed on the data requirements form. For SESTAT/NSCG/NSRCG/SDR, please reference which survey you are requesting restricted-use data for, as well as why the specific restricted-use variables are crucial to your research.

If you intend to match the data to any external datasets, please provide a short description of these datasets, what matching will be done, and the output.

1. Analytical approach.

*Please describe your analytical approach including any statistical techniques (e.g., crosstabulations, regression analysis, correlational, etc.) that you will use to analyze the restricted use data. Because of the costs associated with providing access to restricted use data, in some cases it is preferable for NCSES staff to carry out the analysis (under your direction) rather than provide access to the data. Knowing how you intend to analyze the data can help us make this assessment.*

1. *Justification for use.*

Include a section with a justification for why the publicly available data is not sufficient for your research. If you request multiple years of the SESTAT/NSCG/NSRCG/SDR, please include a justification for that.

1. *Names of* ***same-institution*** *collaborating researchers.*

Include a section naming all the collaborating researchers at your institution for your project(s). **Note that if you need to collaborate with researchers at other institutions, they will need a separate license.** The terms of collaboration will need to be addressed in the security plan.