

Attachment A
NCSES Data Tools and Product User Needs Analysis
User needs contact script and survey instrument
5/217/2016

Initial contact by NSF

Dear <Name>,

We have identified you as having an interest in science and engineering (S&E) data either because you have contacted us or have published on S&E topics. In order to meet your current and future needs, we at the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation would like to better understand how you use our data products. NCSES is the principal source of analytical and statistical reports, data, and related information that describe and provide insight into the nation's S&E resources. The Center distributes a variety of data on S&E topics including research and development (R&D) funding, R&D performance, and the S&E workforce.

In the next few days, you will receive an email with a link to a short questionnaire from an independent third-party research organization, SRI International, on our behalf. The questions will ask about your work and your needs related to S&E data. Your identity will be kept confidential and ~~all your~~ answers will be recorded anonymously-will be kept confidential. Your feedback is a critical element to our ability to learn and improve.

Your participation is voluntary. This study is authorized by law (42 U.S.C. 1862 Section 3.a.6.). The OMB control number for this study is 3145-0174.

If you have questions regarding the questionnaire, please do not hesitate to contact me. We look forward to your participation.

Sincerely,
May Aydin
<Signature>

Follow-up contact from SRI

Dear <Name>

In the past few days, May Aydin at the National Science Foundation's National Center for Science and Engineering Statistics (NCSES) sent you an email about a questionnaire to gather information about the needs of data users. I have included the email below for your reference.

Please find the link to the questionnaire here:
<link>

The average completion time is 15 minutes.

If you have any questions about the project please contact May at maydin@nsf.gov. For technical questions, please contact Christina Freyman at christina.freyman@sri.com.

Sincerely,
Christina Freyman
SRI International
christina.freyman@sri.com

<paste email from above here>

Survey instrument

V132

5/178/2016

Intro

The National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF) is the principal source of statistical data and related analyses on the nation's science and engineering resources. NCSES statistical data is available to the public via a variety of online products ranging from microdata files to table-building tools to state-level and institution reports.

In an effort to improve these products, we are conducting a short survey of individuals with an interest in science and engineering data. Please help us by taking this 10–15 minute survey by X (date).

This survey is voluntary and your answers will be recorded anonymously. By continuing with this survey you are giving your consent for participation. This study is authorized by law (42 U.S.C. 1862 Section 3.a.6.). The OMB control number for this study is 3145-0174.

Please use the previous and next buttons to navigate the survey. At any time you may choose to save the survey by selecting the “Resume later” button. When you return, please select the “Load unfinished survey” button below.

If you have any questions about the survey please contact me at christina.freyman@sri.com.

A note on privacy

This survey is anonymous.

The record of your survey responses does not contain any identifying information about you, unless you identify yourself in a text box. The identifying token you used to access this survey will not be stored together with your responses. It is managed in a separate database and will only be updated to indicate whether you did (or did not) complete this survey. There is no way of matching identification tokens with survey responses.

[Items in square brackets will not be seen by respondent]
[Page 1 Section A: Respondent Classification]

[A1.] **Which of the following types of federal government data or data products have you used?**

Check any that apply

[Respondent will be given an error message if “None...” and another choice is selected.]

- Workforce data (e.g., Occupational Employment Statistics)
- Economic data (e.g., Gross Domestic Product, consumer spending)
- Education data (e.g., National Assessment of Educational Progress)
- Health care data (e.g., mortality data)
- Other type(s) not listed above (please specify in the box to the right)
- None, I have not used federal statistical data

[IF ANSWER TO Q1 IS OTHER THAN “NONE, I HAVE NOT USED FEDERAL STATISTICAL DATA” CONTINUE, IF ANSWER IS “NONE, I HAVE NOT USED FEDERAL STATISTICAL DATA” THEN SKIP TO SECTION D]

[A2.] **How would you classify your primary occupation?**

Choose one of the following answers

- Student (undergraduate or graduate) or postdoctoral fellow
- Academic researcher
- University administrator
- K-12 educator
- Nonprofit researcher or policy analyst
- For-profit employee
- Government employee
- Journalist/reporter
- Other occupation not listed (please specify in the box to the right)

[A3.] **In the past two years, from which federal government agencies have you accessed statistical data or data products?**

Check any that apply

[Mandatory for skip logic.]

- National Center for Science and Engineering Statistics at the National Science Foundation
- Bureau of Economic Analysis
- Bureau of Justice Statistics
- Bureau of Labor Statistics

- Bureau of Transportation Statistics
- U.S. Census Bureau
- Economic Research Service at the U.S. Department of Agriculture
- Energy Information Administration
- National Agricultural Statistics Service
- National Center for Health Statistics
- National Center for Education Statistics at the Department of Education
- Office of Research, Evaluation, and Statistics at the Social Security Administration
- Statistics of Income Division at the Internal Revenue Service
- Other agency not listed above (please specify in the box to the right)

[A4.] Which of the following topics have you searched for or used federal government statistical data for in the past two years?

Check any that apply

- STEM (science, technology, engineering & mathematics) education
- Science and engineering workforce
- Federal or state government research and development (R&D) spending and activities
- Business R&D spending and activities
- University R&D spending and activities
- Other U.S. science and engineering topic(s) not listed above (please specify in the box to the right)
- None, I have not searched for this kind of data

[IF ANSWER TO A4 IS OTHER THAN “None, I have not searched for this kind of data”]

[A5a.] How easy or difficult is it to locate science and engineering data for your needs?

Choose one of the following answers

- Very difficult
- Somewhat difficult
- Somewhat easy
- Very easy

[IF ANSWER TO A4 IS OTHER THAN “None, I have not searched for this kind of data”]

[A5b.] How easy or difficult is it to access science and engineering data for your needs?

Choose one of the following answers

- Very difficult
- Somewhat difficult
- Somewhat easy
- Very easy

[IF ANSWER TO A4 IS OTHER THAN “None, I have not searched for this kind of data”]

[A5b.] **Please describe the problems you encountered, if any.:**

[Text box – will code the themes]

[A6.] **Over the past two years, how often have you used government data or data products?**

Choose one of the following answers.

[Mandatory for skip logic.]

- Less than once per year
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[Page Break]

[Section B: Power and Intermittent users

A6= EVERY 2–3 MONTHS OR MONTHLY OR EVERY 2 WEEKS OR WEEKLY OR DAILY

QUESTIONS MARKED “NCSES DATA=NO” WILL BE SHOWN ONLY TO RESPONDENTS WHO DID NOT CHOOSE NCSES IN A3 ABOVE.

QUESTIONS MARKED “NCSES DATA=YES” WILL BE SHOWN ONLY TO RESPONDENTS WHO DID CHOOSE NCSES IN A3 ABOVE

[NCSES data=NO

B0.] **Why have you not used NCSES data tools or data products?**

Check any that apply

- I was not aware of NCSES data.
- I was aware of NCSES data, but the available datasets do not apply to my questions.
- NCSES has datasets relevant to my questions, but I was unable to access them.
- NCSES data applies to my questions, but available datasets do not include the variables I want.
- NCSES data applies to my questions, but does not provide the level of granularity I want.
- Other reason(s) not listed above (please specify in the box to the right)

[ALL

B1.] **In what ways have you used federal government statistical data?**

Check any that apply

- I have downloaded and analyzed microdata.
- I have downloaded pre-formatted tabular data.
- I have generated table files online and downloaded them as formatted data files.
- I have generated visualizations online.
- I have copied and pasted data from online tables into a spreadsheet or other document.
- I have downloaded reports and imported formatted tables and figures into another document.
- I have accessed short reports (such as InfoBriefs) for informational purposes without extracting tables or figures.
- Other way(s) not listed above (please specify in the box to the right)

[IF B1= YOU DOWNLOADED AND ANALYZED MICRODATA.]

[B1A-1]. **Have you downloaded and imported federal statistical microdata files into a statistical package (e.g., SPSS, SAS, STATA, R)?**

- Yes
- No
- Uncertain

[B1A-2]. **Have you downloaded and imported federal statistical microdata files into a software package to generate visualizations (e.g., Tableau, D3)?**

- Yes
- No
- Uncertain

[B1A-3]. **Have you downloaded and imported federal statistical microdata into a database or spreadsheet (e.g., Microsoft Access, Microsoft Excel)?**

- Yes
- No
- Uncertain

[B1B.] Are there other ways in which you've downloaded and imported federal statistical microdata files not included above?

[ALL

B2.] **On average, how often do you visit the websites of federal government statistical agencies?**

Choose one of the following answers.

- Less than once per year

- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily
- Other (please specify in the box to the right)

[NCSES data=YES

B3.] In the past two years, about how frequently have you visited the NCSES website (<http://nsf.gov/statistics/>)?

Choose one of the following answers.

- Less than once per year
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily
- Other (please specify in the box to the right)

[ALL

B4.] What kind of data access and features do you feel that federal government statistical agencies should provide to meet the needs of users like yourself?

Check any that apply

- Access to detailed microdata
- Access to standard errors associated with data (where applicable)
- Time series data
- Customizable data tables
- Customizable data visualizations
- Detailed methodological reports
- Metadata such as data dictionaries, etc.
- Downloadable data in comma-separated values (.csv) or tabular separated values (.tsv) formats
- Downloadable data in a specific statistical software package format (e.g., SPSS, SAS, STATA)
- Analytical reports exploring policy topics using the agency's own data
- Access to relevant data from other government agencies in addition to the agency's own data
- Access to relevant data from non-government sources in addition to the agency's own data
- Mapping of variables to external standards (such as NAICS, IPEDS, SOC, etc.)
- Other item(s) not listed above (please specify in the box to the right)

[NCSES data=YES

B5.] **What are some examples of topics or issues you have researched using NCSES data?**

[text box encoding by theme]

[NCSES data=YES

B6.] **Do you hold an NCSES restricted data license?**

- Yes
- No

[NCSES data=YES

B7.] **In the past two years, which of the following data tools and/or data products have you used?**

Check any that apply

- **WebCASPAR.** The WebCASPAR Data System is a custom NCSES data tool that provides access to the Survey of Earned Doctorates/Doctorate Records File; the Survey of Federal Funds for Research and Development; Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions; Survey of Research and Development Expenditures at Universities and Colleges/Higher Education Research and Development Survey; Survey of Science and Engineering Research Facilities; and the Survey of Graduate Students and Postdoctorates in Science and Engineering.
- **SESTAT Data Tool.** The SESTAT Data Tool contains data from the National Survey of College Graduates (NSCG), the National Survey of Recent College Graduates (NSRCG), and the Survey of Doctorate Recipients (SDR).
- **SESTAT Metadata Explorer.** The SESTAT Metadata Explorer contains information on the variables included the National Survey of College Graduates (NSCG), the National Survey of Recent College Graduates (NSRCG) and the Survey of Doctorate Recipients (SDR).
- **SED Tabulation Engine.** The Survey of Earned Doctorates (SED) Tabulation Engine is a custom NCSES data tool that provides alternate access to a subset of the data available in WebCASPAR.
- **Public Use Microdata Files.** The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS), Higher Education Research and Development Survey (HERD), and the Federally Funded Research and Development Centers Research and Development Survey (FFRDC) Public Use Files provide access to NCSES data as raw data that may be opened in Excel or other statistical tools.
- **Academic Institutional Profiles.** Academic Institutional Profiles show selected NCSES information for U.S. academic institutions.
- **State Profiles.** Science and Engineering State Profiles is an interactive tool that presents S&E workforce and R&D data for U.S. states.
- **Data Tables.** Detailed statistical data provided in tabular formats.

[NCSES data=YES

B8.] The following questions ask about your average use of NCSES data tools or data products.

“Use” includes using locally stored data. It may include such actions such as viewing, transforming, analyzing, or citing in a report.

Using the following scale, please select your average use in the past two years for each of the listed data tools or data products.

[IF B7= WebCASPAR=CHECKED]

[B8A] In the past 2 years, how often have you used WebCASPAR?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= SESTAT Data Tool=CHECKED]

[B8B] In the past 2 years, how often have you used SESTAT Data Tool?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= SESTAT Metadata Explorer=CHECKED]

[B8C] In the past 2 years, how often have you used SESTAT Metadata Explorer?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= SED Tabulation Engine=CHECKED]

[B8D] In the past 2 years, how often have you used SED Tabulation Engine?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks

- Weekly or daily

[IF B7= Public Use Microdata Files=CHECKED]

[B8E] In the past 2 years, how often have you used public use microdata files?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= Academic Institutional Profiles=CHECKED]

[B8F] In the past 2 years, how often have you used Academic Institutional Profiles?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= State Profiles=CHECKED]

[B8G] In the past 2 years, how often have you used State Profiles?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[IF B7= Data Tables=CHECKED.]

[B8H] In the past 2 years, how often have you used Data Tables?

Choose one of the following answers.

- I do not use
- 1–2 times per year
- Every 2–3 months
- Monthly
- Every 2 weeks
- Weekly or daily

[PAGE BREAK]

[NCSES data=YES

B9.] Do you use NCSES data products that are composed for print, such as the Detailed Statistical Table Reports or NCSES InfoBriefs?

- Yes

- No

[If Yes, then show B9A and B9B.]

[B9A.] **How do you generally use NCSES data products that are composed for print?**

Check any that apply

- I print them.
- I view them on the website.
- I download them.
- Other method(s) (please specify in the box to the right)

[B9B.] **You indicated that you use NCSES data product that are composed for print. Of the choices presented below, please select the top three user options you would desire.**

Please select up to three options.

- View data products on website
- Print directly from website
- Print from PDF
- Print from Excel
- Download to PDF
- Download to Excel
- Other improvements (please specify in the box to the right)

[NCSES data=YES]

[B10.] **Are there any aspects to the NCSES datasets, tools, and products that you have found to be inconvenient or difficult to use?**

- Yes
- No

[IWhen yes is checked, B10A appears]

B10A. **Please describe the problem and, if resolved, how it was resolved in the box below.**

[text box encoding by theme]

[PAGE BREAK]

[NCSES data=YES]

[B11.] **Of the choices presented below, please select the top three improvements related to data availability and accessibility you would like to see to NCSES implement.**

Please select up to three options.

- Provide public use microdata files for all of its non-agency surveys
- Provide data in multiple file formats (e.g., .csv, .tsv)
- Provide data in formats compatible with specific software packages
- Provide more flexible modes for generating tables, charts and visualizations online
- Provide more sophisticated visualizations of NCSES data
- Provide webinars or other training on how to access NCSES data.
- Provide a central, searchable metadata tool that includes all available variables and their descriptions
- Provide access to data via an Application Programming Interface (API)
- Provide a single unified database that contains all NCSES data where one can search by keyword or theme
- Provide more accessible information on variables and methodological notes for each dataset
- Provide access to relevant data from other federal statistical agencies
- Provide access to relevant data from other third parties
- Other improvements (please specify in the box to the right)
- No improvements needed

NCSES data=YES

B12. **Of the choices presented below, please rank the top three improvements related to usability you would like to see implemented by NCSES.**

Please select up to three options.

- Harmonize data so that all time series data is available in one data file.
- Provide documentation on how survey questions have changed, with mappings of changes in variables for those questions.
- Provide a site where users can upload user-created code for cleaning and analysis.
- Provide a site where users can upload user-created analytical outputs or visualizations.
- Other improvements (please specify in the box to the right)
- No improvements needed

[IF B1= YOU DOWNLOADED AND ANALYZED MICRODATA.

B13.] **You indicated that you use microdata. Of the choices presented below, please select the top three file formats in which you want to use microdata.**

Please select up to three options.

- Comma separated values (.csv)
- Tabular separated values (.tsv)
- SPSS format (.spss)
- SAS format (.sas)
- Excel format (.xlsx)
- STATA format (.dta)
- Other file formats (please specify in the box to the right)

[IF B1= YOU DOWNLOADED PRE-FORMATTED TABULAR DATA.

B14.] You indicated that you use tabular data. Of the choices presented below, please select the top three file formats in which you want to use tabular data.

Please select up to three options.

- Comma-separated values (.csv)
- Tabular-separated values (.tsv)
- SPSS format (.spss)
- SAS format (.sas)
- Excel format (.xlsx)
- STATA format (.dta)
- PDF (.pdf)
- HTML (.html)
- Other file formats (please specify in the box to the right)

[ALL

B15.] Many statistical agencies provide online tools for users to build custom tables, visualizations, and/or perform statistical analyses. Of the choices presented below, please select the top three functionalities you would like to see on the NCSES website.

Please select up to three options.

- Build a custom table of counts with user-selected variables
- Build a custom table of results of data transformations on user-selected variables (such as subtotals)
- Produce custom graphical visualizations
- Include standard errors on any output
- Build a custom data set to export in your preferred file format
- Perform statistical significance testing on values
- Other functionalities (please specify in the box to the right)
- I would not use such functionalities

[ALL

B16.] Please list any online federal government data tools you have used at least once in the past two years.

[text box encoding by theme]

[ALL

B17.] **Are there any broader trends in data access and presentation that NCSES should look at?**

[text box encoding by theme]

[ALL

B18.] **Do you have other suggestions for ways that NCSES could improve accessibility and usability of NCSES data?**

[text box encoding by theme]

[Page Break

Section C: Occasional users of NCSES data

A6= LESS THAN ONCE PER YEAR OR 1-2 TIMES PER YEAR

QUESTIONS MARKED "NCSES DATA=NO" WILL BE SHOWN ONLY TO RESPONDENTS WHO DID NOT CHOOSE NCSES IN A3 ABOVE.

QUESTIONS MARKED "NCSES DATA=YES" WILL BE SHOWN ONLY TO RESPONDENTS WHO DID CHOOSE NCSES IN A3 ABOVE

NCSES data=NO

C0.] Why have you not used NCSES data tools or data products?

Check any that apply

- I was not aware of NCSES data.
- I was aware of NCSES data, but the available datasets do not apply to my question.
- NCSES has datasets relevant to my question, but I was unable to access them.
- NCSES data applies to my questions, but available datasets do not include the variables I want.
- NCSES data applies to my questions, but does not provide the level of granularity I want.
- Other reason(s) not listed above (please specify in the box to the right)

[ALL

C1.] In general, why have you looked for statistical data on science and engineering?

Check any that apply

- Class project
- Need a few statistics for a professional or government report
- Writing a story for a media publication
- Need a few statistics for a grant application
- Following up on a link or reference in an article
- Academic planning
- Other reason(s) not listed above (please specify in the box to the right)

[ALL

C2.] Which sources of information do you use to find out about where data on science and engineering is available?

Check any that apply

- Internet search
- Peer-reviewed journal articles
- Professor/advisor
- Presentation
- Media
- Other source(s) not listed above (please specify in the box to the right)

[ALL

C3.] In what ways have you used federal government statistical data?

Check any that apply

- I downloaded and imported microdata files.
- I downloaded pre-formatted tabular data.
- I used an online tool to generate visualizations, such as charts and graphs.
- I used an online tool to generate custom tables.
- I copied and pasted data from online tables into a spreadsheet or other document.
- I downloaded reports and imported formatted tables and figures into another document.
- I accessed short report (such as InfoBriefs) for informational purposes without extracting tables or figures.
- Other way(s) not listed above (please specify in the box to the right)

[ALL

C4.] What kind of data access and features do you feel that federal government statistical agencies should provide to meet the needs of users like yourself?

Check any that apply

- Information on survey methodology
- Estimates of standard error or uncertainty
- Time series data
- Mapping of variables to external standards (such as NAICS, IPEDS, SOC, etc.)
- Microdata
- Downloadable data in comma-separated values (.csv) or tabular-separated values (.tsv) formats
- Customizable data tables
- Customizable data visualizations
- Other item(s) not listed above (please specify in the box to the right)

ALL

[C5.] Many statistical agencies provide online tools for users to build custom tables, visualizations, and/or perform statistical analyses. Of the choices presented below, please select the top three functionalities you would like to see on the NCSES website.

Please select up to three options.

- Build a custom table of counts with user-selected variables
- Build a custom table of results of data transformations on user-selected variables (such as subtotals)
- Produce custom graphical visualizations
- Include standard errors on any output
- Build a custom data set to export in your preferred file format

- Perform statistical significance testing on values
- Other functionalities (please specify in the box to the right)
- I would not use such functionalities.

[ALL

C6.] **Do you have other suggestions for ways that NCSES could improve the accessibility and usability of NCSES data?**

[text box encoding by theme]

[Section D: Non-user of federal government statistical data

IF ANSWER TO Q1 IS "NONE, I HAVE NOT USED FEDERAL STATISTICAL DATA"]

D1. Thank you for your time and interest. This survey does not appear to be relevant to you.