

Slide 1 of 17

Title slide: Acquiring Micro-level NCES Data

Slide 2 of 17

This module introduces users to the NCES micro-level datasets and explains how to access them. It explains the differences between public-use and restricted-use data and describes the processes for acquiring public-use datasets and their supporting materials. Additionally, this module explains what constitutes restricted-use data, describes the reasons for NCES having restricted-use data, and provides information related to acquiring restricted-use datasets and their supporting materials.

Throughout this module, underlined blue screen text indicates a link to additional resources.

Slide 3 of 17

Micro-level data are detailed data about each individual survey respondent. In a microdata file, each respondent's record includes his or her responses to each survey item and any additional constructed variables and/or data added from external sources. An example of data from an external source would be data about the school in which the student is enrolled that are taken from the NCES school universe data.

It is important to note that microdata files can include data about either individuals or institutions. In studies there is often one set of microdata for students, one set for teachers of the students in the study, and one set for the school the student attends.

Slide 4 of 17

NCES is required by law to protect the confidentiality of respondents to NCES surveys who provide individually identifiable data. Whenever possible, NCES applies disclosure protections to micro-level data to allow for the release of public-use data. However, there are times when researchers need to use the more detailed restricted-use data to conduct a specific analysis. There are also instances where the amount of data perturbations and protections that would be required to produce a public-use file are deemed to make the data unusable for research and analytic purposes, so only a restricted-use file is produced. NCES established a licensing system to provide qualified researchers access to restricted-use files, which will be discussed later.

The first step in protecting the confidentiality of individual respondents is to remove all direct identifiers, such as names or addresses from micro-level data files. NCES then uses perturbation techniques to introduce a controlled amount of noise into each data file. These perturbations are used to guard against a data user matching a micro-level data file against an external micro-level data file and identifying individual respondents with certainty through such matches. These perturbations and protections are included in both restricted-use and public-use data files and do not affect the overall data quality.

To produce a public-use data file, NCES staff conducts disclosure risk analyses to identify data elements that, if released as reported, could disclose the identity of an

Acquiring Micro-level NCES Data

individual survey respondent. Data elements that are identified as disclosure risks are masked or coarsened by recoding using top coding, bottom coding, or categorical recoding throughout. In some cases, a data element is suppressed from the public-use file completely.

Slide 5 of 17

NCES produces codebooks, ASCII data files, PDF versions of the survey User Manual, and PDF versions of the data collection instruments, with the exception of material that is copyright-protected.

There are different ways to acquire public-use datasets.

Survey information and documentation can be accessed and downloaded from the appropriate NCES study website. Data products and survey publications can be located and downloaded using the electronic catalog at <http://nces.ed.gov/pubsearch/>.

Some NCES data collections provide an Electronic Codebook, or ECB, software on a CD-ROM or DVD to facilitate the use of NCES data.

Other NCES data collections make the data available for download in various formats and software programs. Some NCES data collections make datasets available both ways such as through eDAT. Specific instructions for acquiring individual datasets are included in the DLDT survey-specific modules.

Slide 6 of 17

Restricted data includes information that could enable the identification of specific schools, individual administrators, teachers, or students; therefore, it is masked or suppressed in a public-use file. Such information is confidential and protected by laws that provide for the protection of individually identifiable data. Failure to comply with these laws is punishable as a Class E felony.

Restricted-use data files contain more detailed information than the public-use files. Thus restricted-use data files may be considered sensitive or unique. Researchers who are interested in accessing restricted-use data must apply for, and be granted, a restricted-use data license from NCES.

NCES is required by law to protect the individually identifiable data provided by NCES survey respondents under a pledge of confidentiality. Under the applicable confidentiality laws, individuals granted access to restricted-use data are not permitted to publish any data that could be used to identify an individual respondent. They are also prohibited from using the data for a purpose other than the statistical purpose for which the data were supplied. Individuals granted access to restricted-use data must provide a sworn statement agreeing to observe the limitations of the law. Anyone granted access who knowingly publishes or communicates any individually identifiable information will be subject to fines of up to \$250,000 and/or up to 5 years in prison.

Slide 7 of 17

The restricted-use license is a contract through which IES/NCES loans a qualified organization the data under certain terms and conditions. The requestor must state a justification for the data use, which should demonstrate both that the requestor's proposed use is statistical and the research question can be answered with data available in the restricted-use file being requested and that the research question cannot be answered adequately with data available in a public-use file. NCES does not evaluate the merits of the proposed line of research, just that the line of research can only be addressed with restricted-use data.

Additionally, the requestor must submit the required documents, agree to keep the data secure from unauthorized disclosure at all times, agree to unscheduled inspections of compliance by the IES data security personnel to ensure compliance, agree to not publish or release unrounded, unweighted counts, and agree to submit all work products to NCES for disclosure review prior to sharing with anyone not authorized on a license.

While this training has referred only to a license up to this point, there are actually two types of permissions for acquiring restricted-use data: a restricted-use license, which is granted to universities and other non-governmental organizations, and a memorandum of understanding, which is an agreement entered into with other federal agencies.

Slide 8 of 17

Documentation for a License Application starts with a formal request which is completed online via the web. To request a new license, go to the following URL <https://nces.ed.gov/statprog/licenseapp/licenseinfo>. go. Read the information and click on "continue". Be sure to read the Restricted-Use Data Procedures Manual and then check the box indicating you have done so.

Slide 9 of 17

Keep in mind that this form is for submitting a new formal request for a new restricted-use data license. You do not need to submit this formal request if: a) You have an active License, and are trying to get access to additional restricted-use data. You should login to your current existing License and submit an "add-data" request, or b) You have already submitted a formal request and the IES Data Security Office has notified you via email of any problems that need to be addressed before your application can be approved. You must use the login email that was sent to the PPO the day the original formal request was submitted. That login email will allow you to edit your already submitted formal request. If you are unsure whether you should be submitting a new formal request or not, please send an email to IESData.Security@ed.gov. Applicants must read the Restricted-Use Data Procedures Manual prior to submitting the application. This manual outlines required procedures and details the security requirements. Once you have read the manual, check the box indicating you have done so and click "request new license".

Slide 10 of 17

There are five steps to obtain an NCES restricted-use data license. As you can see, the applicant must provide information regarding the project officers, data files, project details, users, and then confirm the application.

Slide 11 of 17

The first step requires that the researcher determine several parties to the license. The Principal Project Officer, or PPO, is responsible for day-to-day operations involving the requested data. The PPO must be at the level of a Postdoctoral researcher or higher or he or she must be an experienced researcher in an independent research organization. The Senior Official, or SO, is a person in the organization with the legal authority to sign a contract for the organization. The System Security Officer, or SSO, oversees security of the data. A license may also include up to seven authorized users who are granted permission to access the data through the license.

Researchers below the post-doctoral level can access restricted-use data by being listed as an authorized user on someone else's license. Each authorized user must sign a notarized affidavit of nondisclosure that names the data sets that user has received approval to access. The Principal Project Officer and Senior Official must understand and acknowledge that they are responsible for overseeing the use of the data by anyone listed as an authorized user on their license, and that they will be held accountable for any misuse. Also, each user listed on the License, including the PPO, SSO, and all authorized users are required to complete a short online training course that covers the data security procedures and disclosure prevention measures required under the terms of the License. Users may log into the training through the NCES website, using the following web address:

<https://nces.ed.gov/statprog/licenseapp/CertificationLogin.asp>

The applicant must attest that all personnel have completed the license user training in this section. If users have not completed the training, they will not be able to move on to the next step in the process.

Slide 12 of 17

The second step in the process requires the applicants to select the data of interest. Start by selecting a survey. If you don't see the survey for which you are looking, look under "Individual Surveys". Next, click data files from the second column to add them to your license request. To de-select a data file, click a checked data file from either the second or third columns.

Slide 13 of 17

The third step in the process is to select the project details. Applicants must provide project information. USERS

Acquiring Micro-level NCES Data

MUST ANSWER SEVEN QUESTIONS ABOUT THEIR RESEARCH OBJECTIVES. PLEASE NOTE IF ALL SEVEN QUESTIONS ARE NOT ANSWERED, YOUR APPLICATION WILL BE REJECTED. The seven questions are as follows:

1. What is your Project Title
2. Describe the final research objective.
3. Articulate your specific research questions
4. Identify the specific variables or sets of variables to be used
5. Describe the statistical techniques to be used
6. Education Community Served
7. Explain why the restricted-use data are necessary for your research, as opposed to data that are already publicly available (e.g. on a released public use data file or in one of the data tools using restricted-use data in the NCES DataLab).

Slide 14 of 17

In addition to the principle project officer (PPO), senior official (SO), and the system security officer (SSO) you may designate up to six professional/technical and support staff who will access the restricted-use data. Access to the secure project office and data is limited to only those persons who are listed below. Each person listed is required to have a signed and notarized Affidavit of Nondisclosure on file with IES. Each user listed on the License, including the PPO, SSO, and all authorized users are required to complete a short online training course that covers the data security procedures and disclosure prevention measures required under the terms of the License. Users may log into the training through the NCES website, using the following web address:

<https://nces.ed.gov/statprog/licenseapp/CertificationLogin.asp>

Slide 15 of 17

The last step is to confirm your application and submit.

Slide 16 of 17

Once you have obtained a restricted-use data license, you may also use the online licensing system to request an amendment to an existing license. Amendments include requests to add or remove researchers from your license or requests to add additional data to your license. You will need to provide a description of your research and the justification for needing additional restricted-use data. You will also need to enter the license number and Principal Project Officer's email address.

Slide 17 of 17

This module has provided you with information needed to assist you in accessing NCES micro-level datasets. The module's objectives as well as important resources that have been provided throughout the module are summarized here

Users may elect to assess their learning by answering multiple choice questions pertaining to this module through the Check Your Knowledge option located at the bottom of the screen. This is an optional tool that users may use at any time to review key concepts. You may now Check Your Knowledge, proceed to the next module in the series, or click the exit button to return to the landing page.