

Requesting access to data archived at the North Carolina Education Research Data Center

The North Carolina Education Research Data Center (NCERDC) facilitates problem-focused research on education by sharing data with qualified researchers.

The NCERDC data on students, teachers, and schools are available only for use in research directly related to education issues and policies. Eligibility, procedures for requesting access to the data, and data security requirements are described in the document [NCERDC Data Use Agreement](#) and briefly outlined below:

Summary Checklist

1. Data application cover sheet
2. Proposal including abstract, description of the project and intended use of NCERDC data
3. IRB approval from the Principal Investigator's institution
4. Signed confidentiality agreements from the PI and all research staff
5. Data Security Plan
6. Data Request Form
7. Signed Data Disclosure and Destruction Agreement
8. For faculty only - Compensation to the Data Center for access to the data and time spent on the project by Data Center personnel.
9. For students only - Faculty sponsorship of the proposal, indicated by a letter of support.

Guidelines and forms for these application components are available for download at <https://childandfamilypolicy.duke.edu/research/nc-education-data-center/application-forms/>

Please read through requirements below for important details regarding each of these application components.

Forms can be emailed to kara.bonneau@duke.edu.

Eligible Institutions and Researchers

Data can be released to: an institution of higher education or a non-profit research institution, or a government agency located within the United States. The institution must have established protocols for an Institutional Review Board for the protection of human subjects, to review proposals for research projects employing sensitive data. Researchers requesting NCERDC data must have their current primary affiliation with an eligible institution, or be a currently enrolled student in a doctoral program at an eligible institution.

I. Procedures for Obtaining Data

A. All Researchers:

As stated in the Data Use Agreement, all researchers who access data from the Data Center are required to:

- Complete and sign the Data Use Agreement.
- Sign an Investigator confidentiality agreement.
- Ensure that their research assistants and associates sign a confidentiality agreement and abide by the procedures for protecting confidentiality of the data.
- Obtain their institution's IRB approval of their project. This must be at the full or expedited review level. *A statement of IRB review exemption does not meet the NCERDC requirements.*
- Complete a [Data Request form](#). Each requested file should include a statement explaining why that particular dataset is relevant for the research project. This should be more detailed than “used as a control variable” and should be specific to each file, not copied and pasted across numerous data requests.
- Provide a data security plan that meets NCERDC standards. The NCERDC requires that the original de-identified data files, and all resulting temporary and derived data files, must be stored on a secure network server with protections and restrictions appropriate for sensitive data. *External devices (such as laptops or hard drives) and cloud storage options (such as Google Drive) do not meet the [Data Center security requirements](#).*
- Agree to send all manuscripts (including working papers and press releases) to DPI and the Data Center prior to submitting them for publication.
- Agree to destroy all copies of the data, original and derived, on the date stated in the [Disclosure and Data Destruction Agreement](#). Data destruction dates between January 1, 2022 and December 31, 2026 are approved contingent upon renewal of the NCERDC’s agreement with DPI. *Dates beyond 2026 cannot be accepted at this time.*
- Agree to provide the Data Center with an access fee to cover the administrative costs and personnel time spent on project review, setting up accounts, and providing data files and documentation. The standard fee is \$3480, equivalent to two days of Data Center services. This fee also covers limited email consultation to the investigator and/or research staff as to the origins, structure, and general content of the data files sent. This fee is waived for researchers who are doctoral candidates, and may occasionally be reduced for early career faculty who are not able to obtain external funding for data acquisition. We encourage these researchers to seek funding sources within their universities, and will provide them with any materials that they may need for this process.

Requests for data customization will involve additional fees to compensate the Data Center for time that personnel spend on the project. During the proposal review process, the Data Center will provide the researcher with an estimate of the number of days of Data Center services required for customization.

Doctoral Candidates: In addition to the requirements for all researchers, student applicants must have a faculty member sponsor the proposal by writing a letter of support, indicating that the faculty member will assume responsibility for data security as detailed in the data security plan. Due to resource constraints, students may have access only to existing data files. The Data Center will not waive the data customization fee for dissertation research.

II. Guidelines for Evaluating Proposals

Following the instructions in the Data Use Agreement, proposals should include a proposal abstract (less than 250 words) and should include the title of the project, names of investigators, and funding source (or proposed funding source). The abstract should also state the goals of the project and the Data Center's role in meeting those goals. The proposal also should include a project description of the following: goals, significance, specific research questions, analysis plan, and how the requested data files and measures will be used for addressing this project plan. If the project requires effort from the Data Center beyond providing existing files (e.g., creating customized data files), these activities must be described and justified in the proposal. Most proposals are 5-10 pages long.

Data Center personnel will check completed proposals for compliance with conditions established in the data use agreement, evaluate their feasibility, and ensure that different scholars do not duplicate each other's efforts. The NCERDC also will set priorities for accommodating different projects using the following guidelines.

A. Mission

The NCERDC will give higher priority to proposals that meet the North Carolina Department of Public Instruction priorities. Proposals with a different focus are still eligible for data, but if all other factors are equal, proposals fitting the Data Center's goals of supporting research that is useful for education policy will have a higher priority.

B. Geography

The NCERDC provides data access to researchers from many different universities. The North Carolina Department of Public Instruction has determined that eligible researchers must have their primary affiliation at an institution located within the United States, and that the data files released to eligible researchers may not be analyzed or stored at any co-investigator's institution outside of the United States.

C. Feasibility

Projects that require substantial time investment from Data Center personnel may receive lower priority, depending upon the Data Center's schedule of commitments at the time that the proposal is submitted.

III. Different Kinds of Requests Researchers May Make of the Data Center

A. Existing Data

The Data Center can accommodate requests for copies of data files that are currently held in the Data Center archive. When an application for using these data is approved, the NCERDC will provide access to the requested data through a secure server. Requests for additional years of data or additional data files can be made as addenda to an approved application. Requests for additional files resulting from a shift in research questions may require submission of a new proposal.

B. New Products

1. Customized Data:

A researcher might want the Data Center to merge, concatenate, or subset various existing files. Although this type of request requires some programming from the Data Center, it does not raise new confidentiality concerns. A fee for customization will apply.

2. Linking Data by Names and other Identifying Information

a. Linking to data collected by a research team

Researchers might conduct a survey of students in one school and want to match the survey data to Data Center files. This type of request raises additional confidentiality concerns because the researcher already knows the identity of the students. Such requests will only be considered if the researcher has written documentation of consent to access student academic records, and the consent process is approved by the NCERDC and the researcher's IRB. If this condition is met, and in consultation with the North Carolina Department of Public Instruction, the Data Center will use identifying information from the researcher's data to match that file to Data Center data and provide a dataset in which researchers are able only to identify their study participants.

b. Linking to administrative data files

Some researchers may wish to use only de-identified data from other administrative data files, or data records for a sub-group of students, such as students enrolled in a specific curriculum program. These data should be requested by the research team directly from the administrative data source, specifying the need for identifiers that would permit the Data Center staff to match those individual records to student records in the NCERDC data. The researchers would then receive an encrypted linked file from the NCERDC that would not permit identifying individual students. Requests of this type will be considered on a case by case basis, reviewed carefully for confidentiality concerns, and may be subject to separate review and approval by NCDPI.

All researchers will have signed the Researcher Confidentiality Agreement in which they promise not to try to identify any individual. Projects that involve matching data from external sources are the most labor-intensive for the Data Center to undertake, and will require the researcher to offset Data Center costs. An estimate of the customization fee will be provided to the applicant when a project description has been submitted to the Data Center.