

National Data Archive

National Data Enclave Access Request Form

Objectives

The National Data Enclave (NDE) was established by the National Data Archive to allow researchers meeting certain qualifications, and under strict supervision, to access confidential statistical micro data files. NDE provides a mechanism whereby researchers can access detailed data files in a secure environment, without jeopardizing the confidentiality of respondents.

Location

The NDE is located at ...

[provide physical location, plus tel, fax, email and website]

NDE Operations

Researchers can access the data on-site, where they are provided with the data, computer equipment, software, office space, and NDE staff supervision.

Data

- NDE staff constructs the necessary data files before the guest researcher arrives, and ensures that no restricted data leave the facility.
- Researchers proposing multiple analyses that employ multiple data sets will have access to only one dataset at a time. Under no circumstance will researchers be permitted any opportunity to merge datasets on their own.
- NDE allows researchers to supply their own anonymous data to be linked with NDE data sets to create merged data sets that will be stored at NDE. The researcher-supplied data may consist of proprietary data collected and "owned" by the researcher or other publicly available data legally obtained by the researcher. Researchers MUST provide NDE staff with complete documentation of any data proposed to be merged with NDE data. Researchers expecting to use merged files are responsible for interacting with NDE staff to ensure that their data can be merged with the data. NDE will accept user data files in SAS, SPSS or Stata format.
- NDE periodically creates and maintains backup copies of all computer files. Backup files are stored in a secure storage area accessible by NDE staff only, although they may be made available to researchers who need to return for additional analyses. These backup files will contain user-supplied data as well as the merged files. These backup files will be destroyed upon the written request of the user.

Computer equipment

- NDE facility has [N] user workstations and a black and white laser printer located in a secure room. The NDE computers have no electronic link to the Internet, and have been configured so that removable media such as CD-ROM or DVD writers, floppy disks or USB ports are inaccessible to users.

- NDE workstations consist of [Pentium X XXX MHz] computers running [Windows NT / other?].

Software

- CPro, EPI-Info, SAS, SPSS, and Stata are installed on the workstations, in addition of MS-Office applications. Additional programming/analytic languages can be added as needed. For more information on the version of the software available at NDE, please contact us.
- Researchers must have sufficient expertise in data analysis, and be able to conduct their analyses with one of the software provided. NDE does not provide technical support in data analysis.

Office space

- Researchers must work under the supervision of NDE staff and only during normal working hours (Monday-Friday, 8:30 a.m. - 5:00 p.m.).
- Admittance to NDE will be limited to the researchers whose names are included in the Research Proposal. Researchers will be required to show photo identification before admittance.
- A maximum of 3 collaborating researchers can sit at a computer station.
- Scheduling time at NDE is on a first-come, first-served basis.

NDE staff supervision (for disclosure review)

- External researchers are not allowed to bring documents, manuals, books, etc., that may enable them to identify and disclose confidential information they access at NDE. Neither are they allowed to bring into the NDE cell phones, pagers, or other devices which would enable them to communicate with persons outside of the NDE.
- Researchers may not save output, files, or programs to transportable electronic media. NDE staff can copy output or programs to transportable media, if requested.
- Researchers may take the results of their analyses off-site only after disclosure review by NDE staff. Disclosure review consists of looking for tabular cells less than five, tables with geographic variables in any dimension, models with geographic variables (or variables tantamount to geographic variables) as outcome variables, or case listings.
- All logs will be printed or electronically archived and will be kept by NDE, which will retain only the programs and procedures run by external researchers. The logs will not include results from their research.
- All computer output generated by statistical programs and all hand-written notes based on such computer output are subject to disclosure review by NDE staff before removal from NDE. Output is restricted to summary tables. In no case may any table contain cells with fewer than 5 observations. If found, these small cells are suppressed, generally by obliterating the cell. To assure that small cells cannot be calculated from the other cells in the same row or column, staff make illegible the totals for the rows and columns corresponding to the small cell. Once disclosure review is completed, researchers receive a photocopy of the final tabulations. NDE staff will use best practices in determining whether tabular data are identifiable and will be conservative in their decisions. NDE decisions are final and not subject to negotiation by researchers.

Admission costs

Researchers using NDE will be charged for space and equipment rental, and for staff time necessary for supervision, disclosure limitation review, maintenance of computer facilities (including both hardware and software), and the creation and maintenance of data files required by the researcher. The cost of accessing NDE is given below:

Affiliation of the primary data investigator	Setup charge and file creation (fix cost)	Use of facilities (per day, per computing station)
<i>National users</i>		
Staff from NDE member agency	Free	Free
Other public agency	[in national currency]	[in national currency]
University / research center	[in national currency]	[in national currency]
NGO	[in national currency]	[in national currency]
<i>International users</i>		
Research in partnership with staff from NDE member agency	Free	Free
International organization	[in US\$]	[in US\$]
University / research center	[in US\$]	[in US\$]
NGO	[in US\$]	[in US\$]

An additional amount may be charged as needed for special handling, such as the merging of additional data or creating custom file formats, or acquisition and installation of specific non-standard software. The amount will be determined by discussion between the researcher and NDE staff.

Payment is expected in advance of the use of NDE.

Payments should be made to: [Provide instructions on mode of payment]

Submission of Research Proposals

Researchers must submit proposals using the form below. Prospective researchers are encouraged to check with NDE staff prior to writing their proposals to ensure that the data of interest can be made available to them. Researchers should develop their proposals in a way that facilitates the ability of NDE staff to create the analytic files required for the project. Proposals should be explicit regarding the variables needed as well as any case selection required. Only those data items required to conduct the proposed analyses will be included in the analytic data file and the proposals should address why the requested data are needed for the proposed study. Overly large and complex projects or poorly defined projects will require extensive communication between NDE staff and the researchers proposing the project, and this can cause the process to move slowly. Work to prepare data files can be accomplished most expeditiously if large, complex projects are subdivided into manageable parts and requested data are clearly defined.

Researchers wishing to link NDE data with external data should provide the external data to NDE staff in advance of their entry to and use of NDE.

Upon receipt, the Research Proposal will be evaluated by a review committee convened for that purpose. The following criteria apply to proposal review:

- o Scientific and technical feasibility of the project.
- o Availability of resources at NDE.
- o Risk of disclosure of restricted information.

Researchers should note that approval of their application does not constitute endorsement by NDE of the substantive, methodological, theoretical, or policy relevance or merit of the proposed research. NDE approval only constitutes a judgment that this research, as described in the application, is not an illegal use of the requested data file and that there is high probability that the project can be successfully done in NDE.

National Data Archive

Application for Access to Confidential Data in the National Data Enclave

Title and reference number of the dataset(s) you are requesting (use the exact title and reference number as listed in our survey catalog):

Instructions

This form is to be mailed or faxed to the National Data Archive, with a cover letter printed on the sponsoring agency letterhead.

Mail to: [address]
 [address]
 [address]

Fax to: [fax number, with country and area code]

Access to confidential datasets is only granted when there is a legally registered sponsoring agency (university, company, research centre, national or international organization, etc.).

Information you provide in this form will not be shared with others, unless a breach to the legal agreement is confirmed in which case the National Data Archive may inform partner data archives.

Terms

In this agreement,

- 1 *'Primary Data Investigator'* refers to the investigator who serves as the main point of contact for all communications involving this agreement. The Primary Data Investigator assumes all responsibility for compliance with all terms of this Data Access Agreement by employees of the receiving organization.
- 2 *'Other Investigators'* refers to individuals other than the *Principal Investigator*, including research assistants, who will have access to the restricted data.
- 3 *'Receiving Organization'* refers to the organization / university/ establishment which employs the *Primary Data Investigator*.
- 4 *'Representative of the Receiving Organization'* refers to an individual who has the authority to represent the *Receiving Organization* in agreements of this sort.

Section A. *Primary Data Investigator*

First name	
Last name	
Title	Prof/ Dr/ Mr/ Mrs/ Ms
Organization	
Position in organization	
Postal address	
Telephone (with country code)	
Fax (with country code)	
E-mail	

Section B. *Other Investigators*

Provide names, titles, and affiliations of any other members of the research team who will have access to the restricted data or to output derived from these data.		
Name (last / first)	Position	Affiliation

A current resume or Curriculum Vitae for each person who will participate in the research activity must be provided with this request. Resumes or CVs must specify nationality.

Section C. *Receiving Organization*

Organization name	
Type of organization	<input type="checkbox"/> Line ministry / public administration <input type="checkbox"/> University <input type="checkbox"/> Research centre <input type="checkbox"/> Company <input type="checkbox"/> International organization <input type="checkbox"/> Non-governmental agency (national) <input type="checkbox"/> Non-governmental agency (international) <input type="checkbox"/> other (specify) _____
Organization website (URL)	
Postal address	

Section D. *Representative of the Receiving Organization*

First name	
Last name	
Title	Prof/ Dr/ Mr/ Mrs/ Ms
Position in organization	
Postal Address	
Telephone (with country code)	
Fax (with country code)	
E-mail	

Section E. *Description of intended use of the data*

<p><i>Please provide a description of your research project (project question, objectives, methods, expected outputs, partners). <u>Explain why publicly available datasets are not sufficient for the research purposes.</u> If information is insufficient, your request may be rejected or additional information will be asked. <u>This information may be provided in an attached appendix to this request.</u></i></p>
<p>List of expected output(s) and dissemination policy</p>
<p>Will you need to merge the dataset with other data? <input type="checkbox"/> YES <input type="checkbox"/> NO</p> <p>If YES specify all other datasets that will need to be merged.</p> <hr/> <hr/> <hr/> <hr/> <hr/>

Section F. *Identification of data files and variables needed*

The National Data Enclave provides detailed metadata on its website, including a description of data files and variables for each dataset. Researchers are requested to indicate which subset of variables or cases they are interested in, in order to allow the NDE to prepare the data files.

This request if submitted to access:

- The whole dataset (all files, all cases)
- A subset of variables and/or cases as described below (note that variables such as the sample weighting coefficients and records identifiers will always be included in subsets):

Section G. *Software requirements*

The following software will be used by the researchers:

Software available at NDE:

- CPro
- EPI-Info
- SAS
- SPSS
- Stata

Other software: _____
(specify)

Notes:

- *NDE regularly upgrades its software. Contact us if you need more information on the version of each application available.*
- *Researchers who need software not provided as a standard by NDE will have to provide NDE with a valid license of the application, which will be installed by NDE staff for the duration of the research work (the license will remain the property of the researcher). Please contact NDE prior to finalizing this request to confirm the technical feasibility.*

Section H. *Data access agreement*

The Primary Data Investigator, the Other Investigators, and the Representative of the Receiving Organization agree to comply with the following:

- 1 Access to the confidential data will be limited to the Primary Data Investigator and Other Investigators listed in the application form, and who will sign the Affidavit of Confidentiality.
- 2 The data will only be processed for the stated statistical purpose. They will be used for solely for reporting of aggregated information, and not for investigation of specific individuals or organizations. Data will not in any way be used for any administrative, proprietary or law enforcement purposes.
- 3 The Primary Data Investigator undertakes that no attempt will be made to identify any individual person, family, business, enterprise or organization. If such a unique disclosure is made inadvertently, no use will be made of the identity of any person or establishment discovered and full details will be reported to the National Data Enclave. The identification will not be revealed to any other person not included in the data Access Agreement.
- 4 Any books, articles, conference papers, theses, dissertations, reports, or other publications that employ data obtained from the National Data Enclave will cite the source of data in accordance with the Citation Requirement provided with the dataset.
- 5 An electronic copy of all reports and publications based on the requested data will be sent to the National Data Enclave.

- 6 The original collector of the data, the National Data Enclave, and the relevant funding agencies bear no responsibility for use of the data or for interpretations or inferences based upon such uses.
- 7 Breaches of the agreement will be taken seriously and the National Data Enclave will take action against those responsible for the lapse if willful or accidental. Failure to comply with the directions of the National Data Enclave will be deemed to be a major breach of the agreement and may involve recourse to legal proceedings. The National Data Enclave will maintain and share with partner data archives a register of those individuals and organizations which are responsible for breaching the terms of the Data Access Agreement and will impose sanctions on release of future data to these parties.
- 8 The National Data Enclave reserves the right to terminate any project at any time that it deems that an investigator's actions will compromise confidentiality or ethical standards of behavior in a research environment.
- 9 No printouts, electronic files, documents, written notes, or media will be removed from the National Data Enclave until they have been scanned for disclosure risk by the National Data Enclave staff.
- 10 The Primary Data Investigator and Other Investigators may be barred from any future use of the National Data Enclave upon review and determination by the Director of the National Data Enclave that this is necessary to protect the integrity and confidentiality of the National Data Enclave.

Signatories

The following signatories have read and agree with the Data Access Agreement as presented in Section H above:

The Principal Data Investigator

Name _____ Signature _____ Date _____

The Representative of the Receiving Organization

Name _____ Signature _____ Date _____

National Data Enclave (NDE) - Affidavit of Confidentiality

NDE expects that all researchers will adhere to established standards and principles for carrying out statistical research and analyses. Researchers must conduct only those analyses which have received approval. Failure to comply will result in cancellation of the research activity and potential disbarment from future research activities in the NDE.

All researchers participating on an approved project will be asked to sign the following Affidavit of Confidentiality' and Agreement before accessing NDE.

I agree:

1. To make no copies of any files or portions of files to which I am granted access except those authorized by the National Data Enclave (NDE) staff. No confidential data or information viewed or otherwise obtained while I am a researcher in NDE will be removed from NDE.
2. To return to NDE staff all NDE restricted materials with which I may be provided during the conduct of my research at NDE and other materials as requested.
3. Not to use any technique in an attempt to learn the identity of any person, establishment, or sampling unit not identified on public use data files.
4. To hold in strictest confidence the identification of any establishment or individual that may be inadvertently revealed in any documents or discussion, or analysis. Such inadvertent identification revealed in my analysis will be immediately brought to the attention of NDE staff.
5. Not to remove any printouts, electronic files, documents, or media until they have been scanned for disclosure risk by NDE staff. I understand that NDE will perform a disclosure review and must provide approval to me before I remove any data from NDE, whether they are in electronic or paper form.
6. Not to remove from NDE any written notes pertaining to the identification of any establishment, individual, or geographic area that may be revealed in the conduct of my research at NDE.
7. To comport myself in a manner consistent with the principles and standards appropriate to a scientific research establishment.

I understand that deliberate violation of any of these conditions may result in cancellation of the data access agreement, and that I may be barred from any future use of NDE upon review and determination by the Director of NDE that this is necessary to protect the integrity and confidentiality of NDE.

Researcher's Name and Signature

Date

Witness's Name (NDE staff) and Signature

Date