

Guidance on the use of public use data sets:

This guidance represents the Northwestern University IRBs practice regarding investigators' use of specified "public use" data sets. Secondary data analysis of publicly available data is a common research method. Increasingly federal agencies supporting research require investigators to make the data they collect publicly available. Additionally, many professional organizations and journals require that research data sets of published works be made accessible to encourage scholarly consideration and possible replication of the research.

That said, publicly available data does not constitute research with human subjects (as defined under 45CFR46.102) as it does not involve access to identifiable private information about people or the people are already identifiable in the public domain. The primary mission of the IRB is protection of participants in research. If people are already identified in the public or if there are no identifiers, there is no one for the IRB to protect so secondary analysis of publicly available data does not require NU IRB review.

Note: Researchers *can* submit research involving public use data for documentation of NU IRB acknowledgement of "not human research" but are not required to do so. If you want to submit to the IRB use the [HRP-503 Human Research Determination form](#).

Additional considerations:

1. If part of a research project involves secondary analysis of publicly available data in the context of additional procedures that do meet the criteria for human research, the whole project needs to be submitted for IRB review.
2. This guidance also does not apply to restricted use files of secondary data that may be provided under special agreements by the external agencies, repositories or similar organizations. Such secondary datasets typically include potentially identifying information, unusually sensitive and/or private data, etc. The distinction between public use and restricted use data is generally determined by the primary source, and use of restricted data requires IRB review and typically a confidentiality agreement. Projects that use restricted data must be approved by the NU IRB using the Data Analysis Protocol.
3. Additionally, if you think you have a publicly available dataset that is not listed below, you can submit the HRP-503 determination form to assist in determining if the dataset meets the requirements for a public use dataset. A link to the dataset website must be provided.
4. Researchers who wish to propose an additional agency or data set be added to the public data sources list should contact NU IRB by sending an e-mail request with a subject heading of "Request for Addition to Public Use Data List".
5. Secondary analysis of data derived from already completed research projects may not qualify for a "not human research" determination. Investigators planning to make use of such data must submit an application using the Data Analysis Review protocol to the NU IRB for review and approval.

Sources of Public Use Data (that do not require NU IRB review):

The following are known repositories or publicly available data. Note that some of the repositories and other data sources also have restricted data. Secondary analysis of restricted data must be submitted to the NU IRB for consideration as human research.

[American National Election Studies \(ANES\)](#)

[Better Access to Data for Global Interdisciplinary Research \(BADGIR\)](#)

[CDC: National Center for Health Statistics](#)

[Health and Retirement Study-University of Michigan \(HRS\)](#)

[Inter-University Consortium for Political and Social Research](#)

[Integrated Public Use Microdata Samples \(IPUMS\)](#)

[NCES: National Center for Education Statistics](#)

[National Child Development Study](#)

[Panel Study of Economic Dynamics \(PSID\)](#) – University of Michigan

[Pew Research Center](#)

[QDR: Qualitative Data Repository](#)

[Roper Center for Public Opinion Research](#)

[Survey of Consumer Attitudes \(SCA\)](#)

[University of Wisconsin-Madison Data and Information Services Center \(DISC\)](#)

[U.S. Bureau of Census](#)

[U.S. Bureau of Labor Statistics](#)

See also the Northwestern University Library Data Resources Guide: [Social Science Data Resources](#) (<http://libguides.northwestern.edu/data>). The Social Science Data Librarian Ann Zald (anne.zald@northwestern.edu, 847-467-6441) can be contacted for assistance in accessing data files for NU faculty, staff and students. Some of these data resources are publicly available, some are restricted and some are resources purchased by the NUL for use by NU researchers and may involve a data use agreement or other data sharing mechanism.