



National Center for Health Statistics












## National Survey of Family Growth

# NSFG Cycle 6 (2002): Public Use Data Files, Codebooks, and Documentation

Persons with disabilities experiencing problems accessing this page should contact CDC-INFO at [CDC-INFO@cdc.gov](mailto:CDC-INFO@cdc.gov), 800-232-4636 or the TTY number at (888) 232-6348 and ask for a 508 Accommodation PR#9342. If emailing please type 508 Accommodation PR#9342 without quotes in the subject line of the email.

## Codebooks and Documentation










- Codebooks:
  - Webdoc, the NSFG's interactive online codebook, was deactivated as of 12/31/20. Public-use file indexes (Appendix 1a, 1b, and 1c linked below) can be searched to find variable names or identify relevant variables based on key words in variable labels. The section letters noted in the file indexes correspond to the sections of the codebook, so the file indexes indicate which codebook PDF will contain your variables of interest.
  - [Female codebook](#)  [PDF – 7.7 MB]
  - [Male codebook](#)  [PDF – 4 MB]
  - [Pregnancy codebook](#)  [PDF – 850 KB]
- User's Guide
  - [Main Text](#)  [PDF – 209 KB]
  - [Appendix 1 \(File indexes\)](#)  [PDF – 432 KB]
  - [Appendix 2 \(Recode specs\)](#)  [PDF – 374 KB]
- Notes for Users: new and revised variables
  - [Revised Health Insurance Variable](#)  [PDF – 12 KB]
  - [METRO Residence Coding](#)  [PDF – 75 KB]
  - [Household Variables, User's Guide & Codebook](#)  [PDF – 3 KB]

- [Household Variable – parent type](#)  [PDF – 32 KB]

## Variance Estimation Examples

Each of the PDF files below provides programs and output in SAS, SUDAAN, STATA, and WesVar for nine examples of variance estimation. These examples are intended to cover many of the common types of variance estimation you may do in your analyses. If you have any questions about any of these examples, please e-mail the NSFG staff at [nsfg@cdc.gov](mailto:nsfg@cdc.gov).

### Variance Estimation:

- [Example 1. Percent of women using the oral contraceptive pill by age](#)  [PDF – 753 KB]
- [Example 2. Total number of women using the oral contraceptive pill by age](#)  [PDF – 346 KB]
- [Example 3. Mean numbers of children ever born \(PARITY\), by race and Hispanic origin for women 20-44 years of age](#)  [PDF – 332 KB]
- [Example 4. Linear regression of PARITY on age, race and Hispanic origin, and education for women 20-44 years of age](#) [PDF – 419 KB] 
- [Example 5. Percent distribution of pregnancies by wantedness status, by race and Hispanic origin, and age at pregnancy outcome](#)  [PDF – 544 KB]
- [Example 6. Percent of male and female teenagers 15-19 years of age who have ever had sex, by sex and race and Hispanic origin \(combined male/female files\)](#)  [PDF – 227 KB]
- [Example 7. Percent of males 20-44 years of age who have ever fathered a child, by race and Hispanic origin](#)  [PDF – 247 KB]
- [Example 8. Percent of males and females who strongly agree that “a young couple should not live together unless they are married” by sex and age \(combined male/female files\)](#)  [PDF – 290 KB]
- [Example 9. Logistic regression for the probability of strongly agreeing that “a young couple should not live together unless they are married” by age, sex, race and Hispanic origin, and education \(combined male/female files\)](#)  [PDF – 445 KB]

## Questionnaires

- [Description of the two questionnaire formats](#)
- [Female Questionnaire](#)
- [Male Questionnaire](#)

## Downloadable Data Files

- [Female Respondent Data File \(2002FemResp.dat\)](#)
- [Female Pregnancy Data File \(2002FemPreg.dat\)](#)

- Male Respondent Data File (2002Male.dat)
- Household Variables, ASCII Data File (2002 HHvars.dat)
- Household Variable – parent type, ASCII Data File (HHPARTYPNEWASC.DAT)
- Household Variable – parent type, SAS Data File (hhpartypnew.sas7bdat)
- Current Insurance Variable, ASCII Data File (2002curr\_ins.dat)
- Current Insurance Variable, SAS Data File (c6\_curr\_ins.sas7bdat)

## Program Statements

- SAS Program Statements
  - Female respondent files
    - 2002FemRespInput.sas
    - 2002FemRespValueLabel.sas
    - 2002FemRespVarLabel.sas
  - Male respondent files
    - 2002MaleInput.sas
    - 2002MaleValueLabel.sas
    - 2002MaleVarLabel.sas
  - Pregnancy files
    - 2002FemPregInput.sas
    - 2002FemPregValueLabel.sas
    - 2002FemPregVarLabel.sas
  - Household variables
    - 2002HHvars.sas
- SPSS Program Statements
  - Female respondent file (2002FemRespSetup.sps)
  - Male respondent file (2002MaleRespSetup.sps)
  - Pregnancy file (2002PregSetup.sps)
- STATA Program Statements
  - Female respondent files
    - 2002FemResp.dct
    - 2002FemResp.do
  - Male respondent files
    - 2002Male.dct
    - 2002Male.do
  - Pregnancy files
    - 2002FemPreg.dct
    - 2002FemPreg.do
  - Household Variables
    - 2002HHvars.sas

## Other Data Files

When the National Center for Health Statistics (NCHS) collected the data from respondents to the National Survey of Family Growth (NSFG), those respondents were promised that the information they provided would be held “strictly confidential.” The NCHS is legally required to keep that promise. In order to do so, some variables could not be included on Public Use Files, either because they pose additional risk of disclosure or they contain data that would increase the consequences of disclosure. These items are made available to the research community in the following data files:

The three other Cycle 6 NSFG data files are:

- [ACASI Data](#)
- [Interview Variables](#)
- [Contextual \(Geographic\) Data](#)

## ACASI Data

The NSFG Cycle 6 questionnaires contained a number of items designed to provide a comprehensive description of current and past behavior related to the risk of acquiring sexually transmitted infections (STI), including the Human Immunodeficiency Virus, or HIV, the virus that causes AIDS. These questions were asked via Audio Computer-Assisted Self-Interviewing, or ACASI, in which the respondent hears the question through headphones or reads it from the laptop screen and enters the answer directly into the computer. The object of ACASI was to give respondents a more private opportunity to report this sensitive information.

The ACASI files include most of the items from the ACASI portion of the NSFG Cycle 6 interview (female section J and male section K). The series on income and sources of income were collected in ACASI, but they are included on the Cycle 6 Public Use Files.

- The questions included in ACASI were largely the same for male and female respondents.
- Comparable items were asked about drug use, risk behaviors for sexually transmitted infections (STI, including HIV), and experience with STI.
- Both male and female respondents were given an opportunity to re-report their experience with pregnancies or fathering pregnancies that were previously reported directly to the interviewer.
- All adult respondents (18-44) were asked about non-voluntary sexual intercourse and types of force they may have experienced, if they reported non-voluntary intercourse.
- While the main interviewer-administered portion of the NSFG interview was limited to heterosexual vaginal intercourse, in ACASI all respondents were asked about other types of sexual activity, including oral and anal sex and same-sex partners.

Due to a change in NCHS policy, made effective in March 2020, the ACASI Data Files for 2002 are no longer accessible via special data use agreement. These files can only be accessed through the NCHS Research Data Center (RDC). For more information about using the RDC, including access and associated charges, visit the [RDC website](#).

For additional information or questions about these files, researchers may contact the NSFG staff [nsfg@cdc.gov](mailto:nsfg@cdc.gov)

## Interviewer Variables




Interviewers' observations of the interview process and circumstances are collected in the NSFG. These serve as measures of the respondent's environment and help assess factors that may affect data quality. In Cycle 6 of the NSFG, this paper-and-pencil questionnaire was expanded to include a more complete assessment of these factors. The data and documentation are now available for this file, which contains the responses of over 260 female interviewers, for the 12,571 respondents in the 2002 NSFG. These data are available through NCHS' Research Data Center (RDC). Please contact NSFG staff at [nsfg@cdc.gov](mailto:nsfg@cdc.gov) or staff of the RDC at [rdca@cdc.gov](mailto:rdca@cdc.gov). You can find further information at the [RDC Website](#).


## Contextual (Geographic) Data

Contextual or geographic data provide information on the context, or community, in which respondents live. Geographic data may include information for the region, state, county, census tract, or block group in which the respondent lives. A contextual data file for Cycle 6 is available now for use by the research community. These data are only available for use through the NCHS Research Data Center, or RDC, described below. The contextual data for Cycle 6 are drawn from 4 major sources:

- The 2000 Census Summary Files are the source of some contextual data, and many of these variables are available at the County, Census Tract, and Block Group levels.
- The second source is the County and City Data Book. This information is available only at the county level.
- The third source is a file of variables relating to family planning services and the need for services, all at the county level.
- The fourth source is a file of rates of selected sexually transmitted diseases (STD's), also at the county level.

Most of these variables are available for the respondent's residence at 2 points in time – the date of interview in 2002, and at the date of the census, on April 1, 2000.


The Cycle 6 [Contextual Data File Codebook](#)  [PDF – 3.3 MB] and a list of variables contained on the NSFG [Cycle 6 Contextual Data Files](#) , are available for viewing or downloading. Variables with large numbers of missing values are not included in the contextual data files. We also have available for viewing a [complete list of variables with missing values for 6,000 or more cases](#)  [PDF – 25 KB].

Researchers can see the Cycle 5 contextual variable list in a previous NSFG report published by NCHS in April 2003 (see Vital and Health Statistics [Series 23, Number 23](#) , by Mosher, Deang, and Bramlett, particularly Appendix II).

Researchers may also request that other variables be added to existing NSFG files in the NCHS Research Data Center. For example, a researcher might add state-level variables indicating provisions of a welfare law or a law relating to health care coverage. (Please contact the NSFG staff for specific instructions). There are charges for the use of the RDC, which are explained at the [RDC Website](#).

Researchers may also find useful information for working with NSFG data through the RDC in the Series 23, Number 23 report mentioned above or contact [nsfg@cdc.gov](mailto:nsfg@cdc.gov). For more information about using the RDC, visit the [RDC Website](#).

Last Reviewed: September 27, 2021

 Was this page helpful?

Yes    Partly    No