Using the National Survey of Family Growth

www.cdc.gov/nchs/nsfg.htm
nsfg@cdc.gov

Kimberly Daniels, Ph.D.; Gladys Martinez, Ph.D.; Casey Copen, MPH, Ph.D.
National Center for Health Statistics

Workshop presented at the
2015 NCHS National Conference on Health Statistics
North Bethesda, MD
August 24, 2015
Outline of NSFG Workshop

- Background of the NSFG
- Survey Design and Sample
- Questionnaire Content
- User Tools – How to access data, documentation and support
- Hands-on practice with User Tools and Data
- Questions and discussion
Purposes of the National Survey of Family Growth (NSFG):

• **Addressing Sec 306 of the Public Health Service Act:**
  “NCHS shall collect statistics on...family formation, growth, & dissolution.”

• **Explaining variations in birth rates:** contraception, sexual activity, infertility.

• **Monitoring risk of HIV and sexually transmitted diseases:** # of sexual partners, condom use, drug use, same-sex contact.

• **Describing relationships & families:** Marriage, divorce, cohabitation; Roles of fathers in raising children.

• **Describing attitudes** about sex, marriage, parenthood.
Proximate Determinants of Fertility
Intermediate Variables

Background Characteristics

- Race/ethnicity
- Religion
- Labor force participation
- Education
- Income
- Access to health care
- Family background
- Community environment (economic, social, etc)

Intercourse variables:
- Timing of first intercourse
- Percent of women who ever had intercourse
- Time spent in marriage or cohabitation (separation, divorce)
- Frequency of intercourse

Conception variables:
- Contraceptive use
- Sterilization
- Infertility

Pregnancy outcome (gestational) variables:
- Miscarriage and stillbirth
- Induced abortion

Fertility

Live births
Percent of pregnancies by outcome and race and Hispanic origin: United States, 2009

Non-Hispanic white

- Live birth: 70%
- Induced abortion: 18%
- Fetal loss: 12%

Non-Hispanic black

- Live birth: 49%
- Induced abortion: 15%
- Fetal loss: 36%

Hispanic

- Live birth: 68%
- Induced abortion: 15%
- Fetal loss: 17%

<table>
<thead>
<tr>
<th>Year</th>
<th>Scope (15-44)</th>
<th>N</th>
<th>Over-Samples</th>
<th>Average Length</th>
<th>Contractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973</td>
<td>Ever-Married Women</td>
<td>9,797</td>
<td>Black women</td>
<td>60 Min</td>
<td>NORC</td>
</tr>
<tr>
<td>1976</td>
<td>Ever-Married Women</td>
<td>8,611</td>
<td>Black women</td>
<td>60 Min</td>
<td>Westat</td>
</tr>
<tr>
<td>1982</td>
<td>All Women</td>
<td>7,969</td>
<td>Black women &amp; teens</td>
<td>60 Min</td>
<td>Westat</td>
</tr>
<tr>
<td>1988</td>
<td>All Women</td>
<td>8,450</td>
<td>Black women</td>
<td>70 Min</td>
<td>Westat</td>
</tr>
<tr>
<td>1995</td>
<td>All Women</td>
<td>10,847</td>
<td>Blacks &amp; Hispanics</td>
<td>100 Min</td>
<td>RTI</td>
</tr>
<tr>
<td>2002</td>
<td>All Women &amp; Men</td>
<td>12,571</td>
<td>Blacks, Hispanics, &amp;</td>
<td>W= 85 min</td>
<td>U of Michigan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teens</td>
<td>M= 60 min</td>
<td></td>
</tr>
<tr>
<td>2006-2010</td>
<td>All Women &amp; Men</td>
<td>22,682</td>
<td>Blacks, Hispanics, &amp;</td>
<td>W=~80 min</td>
<td>U of Michigan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teens</td>
<td>M=~60 min</td>
<td></td>
</tr>
<tr>
<td>2011-2013</td>
<td>All Women &amp; Men</td>
<td>10,416</td>
<td>Blacks, Hispanics, &amp;</td>
<td>W=~80 min</td>
<td>U of Michigan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teens</td>
<td>M=~60 min</td>
<td></td>
</tr>
</tbody>
</table>
NSFG Funders

- National Center for Health Statistics (NCHS)
- Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
- Office of Population Affairs (OPA)

- Children’s Bureau (Admin for Children, Youth and Families)
- ACF’s Office of Planning, Research, & Evaluation (OPRE)
- Office of the Assistant Secretary for Planning and Evaluation (OASPE)

- CDC’s Division of HIV/AIDS Prevention
- CDC’s Division of STD Prevention
- CDC’s Division of Reproductive Health
- CDC’s Division Cancer Prevention & Control
- CDC’s Division of Birth Defects and Developmental Disabilities
Survey Design and Sample

For more detailed information on the methodology and design of the NSFG, see:


2011+ NSFG Currently in the Field

• Interviewing began in September 2011 and is now in the 4th year
• Interviews are conducted in-person in the homes of respondents, using laptop computers
• Average interview length of 80 min for females and 60 min for males, including 15-20 min of ACASI content
• Highly skilled female interviewers, trained specifically on NSFG for one week
• $40 Incentive offered for the interview
• Teens 15-17 must have signed parental consent and also provide assent themselves.
• Interviews conducted in both English & Spanish (since the 1973 NSFG)
• Age range expansion from 15-44 to 15-49 this fall.
### 2011-2013 National Survey of Family Growth

<table>
<thead>
<tr>
<th>Category</th>
<th>Sample Size</th>
<th>Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>All 15-44</td>
<td>10,416</td>
<td>73%</td>
</tr>
<tr>
<td>Females 15-44</td>
<td>5,601</td>
<td>73%</td>
</tr>
<tr>
<td>Males 15-44</td>
<td>4,815</td>
<td>72%</td>
</tr>
<tr>
<td>Teens 15-19</td>
<td>2,125</td>
<td>75%</td>
</tr>
</tbody>
</table>
Content of the NSFG Questionnaires
Pregnancy history for women: 2011-2013

- Outcome and gestational length
- Timing of learning she was pregnant#
- Timing of first prenatal care visit#
- Smoking during pregnancy#
- Intendedness at time of conception

- For live births:
  - Birth weight
  - Payment for delivery
  - Breastfeeding
    - Initiation
    - Duration of any breastfeeding
    - Duration of exclusive breastfeeding

# = Not asked for abortions
Fathering information for men: 2011-2013

• Men are asked about biological children fathered within the context of specific relationships, as well as non-biological children he has ever lived with

• For each child:
  – Age and sex
  – Living arrangements (where child lives now)
  – When did he learn about the pregnancy (before delivery or after)
  – Paternity establishment
  – Marital or cohabiting status with child’s mother at time of delivery
  – For each child < 19 years old:
    • Wantedness of pregnancy when he learned about the pregnancy
    • Happiness about pregnancy when he learned about the pregnancy

• Fathering behaviors and activities with children:
  – Items vary by age of child and co-residential status

• Other pregnancies he fathered that did not end in live birth (Numbers ending in miscarriage, stillbirth, abortion)
Measures of contraceptive use: 2011-2013

• **Females**
  – Ever-use of each specific method
  – Method use at first intercourse
  – First method ever used
  – Partner-specific use (up to 3 “most recent partners”)
  – Month-by-month method calendar spanning 3 - 4 years prior to interview (including month of interview)
  – Consistency of condom use in past 4 weeks & past 12 months

• **Males**
  – No method calendar, but comparable contraceptive use items asked within specific relationships
  – Method use at first & last intercourse with each of up to 3 “most recent partners”
  – Past 12 months: all methods used with partner, method used the most, frequency of using any method, & consistency of condom use
Marriage, Dissolution, and Cohabitation: 2011-2013

Marital histories
- Start/end dates, premarital cohabitation, # of children with each spouse, spouse’s prior marriage history & # of children

Cohabitation histories
- Start/end dates, outcomes, engagement status, expectations to marry for current cohabiters, children with each partner
- Men: limited to cohabs with women he ever married + 1st cohab partner ever

Selected spouse/partner characteristics
- Age, race/Hispanic origin, education
Infertility, reproductive health, and health services: 2011-2013

- Infertility services ever used
  - Help to get pregnant
  - Help to prevent miscarriage
  - Length of time tried to become pregnant before 1st visit for help

- Health conditions and behaviors related to reproduction
  - Diabetes, ovulatory problems, fibroids, etc.
  - Vaginal douching
  - Pelvic inflammatory disease (PID)

- Birth control and medical services used in past 12 months
Types of data collected in Audio CASI

• Pregnancies ever had or fathered
• Incarceration experience (males only)
• Alcohol, smoking, (females only), and drug use in last 12 months
• Vaginal, oral, and anal sex with opposite-sex partners
• Same-sex sexual experience (greater detail for males)
• Non-voluntary vaginal intercourse (18-44 only)
• HIV/STI risk behaviors; STI history
• Sexual attraction and sexual orientation/identity
• Income and public assistance received
NSFG User Tools and Data Resources
The National Survey of Family Growth (NSFG) gathers information on family life, marriage and divorce, pregnancy and infertility, use of contraception, and men's and women's health. The survey results are used by the U.S. Department of Health and Human Services and others to plan health services and health education programs, and to do statistical studies of families, fertility, and health. Links to some of those studies are included on this website, under "Publications and Information Products."

What's New

- 2011-2013 National Survey of Family Growth (NSFG) Public Use Data Files (12/2014)

Publications

- Sexual Activity, Contraceptive Use, and Childbearing of Teenagers Aged 15-19 in the United States (7/2015)
- Three Decades of Nonmarital First Births Among Fathers Aged 15-44 in the United States (6/2015)
- Interpregnancy Intervals in the United States: Data From the Birth Certificate and the National Survey of Family Growth (4/2015)
Key Statistics from the National Survey of Family Growth

B Listing

• Birth control - See Contraception
• Births (mothers)
  • Age at first birth
  • Marital or cohabiting status of mother at time of birth
  • Number of children ever born (parity)
  • Premarital/nonmarital status of births
  • Attitudes about childbearing - See Attitudes: Parenthood and childbearing
• Births (fathers)
  • Age at first birth fathered
  • Marital or cohabiting status of father at time of birth

Age at first birth fathered

| Mean age at the birth of their first biological child for men 15-44 years of age who ever fathered a biological child |
|--------------------------------------------------|----------------|----------------|
| 2002 | 2006-2010 | 2011-2013 |
| 25.1 | 25.1 | 25.4 |
Publications and Information Products

2011-2013 Reports

- Sexual Activity, Contraceptive Use, and Childbearing of Teenagers Aged 15–19 in the United States (7/2015)
- Three Decades of Nonmarital First Births Among Fathers Aged 15–44 in the United States (6/2015)

2006-2010 Reports

- Interpregnancy Intervals in the United States: Data From the Birth Certificate and the National Survey of Family Growth (4/2015)
- Recent Declines in Nonmarital Childbearing in the United States (8/2014)
2011–2013 NSFG Reports (more to come)
Bibliography for the National Survey of Family Growth

List of Reports and Articles Using NSFG Data

- 2006-2010 [PDF - 117 KB]
- Cycle 6 [PDF - 119 KB]
- Cycle 5 [PDF - 97 KB]
- Cycle 4 [PDF - 73 KB]
- Cycle 3 [PDF - 70 KB]
- Cycles 1 and 2 [PDF - 76 KB]
## Data files for 2011-2013 NSFG

<table>
<thead>
<tr>
<th>Data files</th>
<th>Availability</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Public-use data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-- Female respondent (5,601 records)</td>
<td>Public, free</td>
<td>Downloadable, NSFG website</td>
</tr>
<tr>
<td>-- Female pregnancy (9,543 records)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-- Male respondent (4,815 records)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACASI data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-- Female respondent</td>
<td>Public, free (previous years available by request)</td>
<td>Downloadable, NSFG website</td>
</tr>
<tr>
<td>-- Male respondent</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>REGION data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-- Female and male respondent file</td>
<td>Public, free, but available by request</td>
<td>NSFG team</td>
</tr>
<tr>
<td><strong>Contextual data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-- Time of interview, state and county</td>
<td>Restricted, fee based</td>
<td>Research Data Center</td>
</tr>
<tr>
<td>-- Time of 2010 census, state</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CASEID= linking variable across files
2011-2013 NSFG: Public Use Data Files, Codebooks, and Documentation

Persons with disabilities experiencing problems accessing this page should contact CDC-INFO at 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 interactive codebook

- User's Guide:
  - Main Text [PDF - 1.16 MB]
    - Part 1 General Information for Users
    - Part 2 Topic-Specific Notes
  - Appendix 1: File Indexes for 2011–2013 NSFG

On this Page

- Codebooks and Documentation
- Variance Estimation Examples
- Questionnaires
- Downloadable Data Files
- Program Statements
- Other Data Files
Downloadable Data Files

- Female Respondent Data File (2011-2013_FemResp.dat)
- Female Pregnancy Data File (2011-2013_FemPreg.dat)
- Male Respondent Data File (2011-2013_Male.dat)

Program Statements

- SAS Program Statements
  - Female respondent file (2011-2013_FemRespSetup.sas)
  - Pregnancy file (2011-2013_FemPregSetup.sas)
  - Male respondent file (2011-2013_MaleSetup.sas)
- SPSS Program Statements
  - Female respondent file (2011-2013_FemRespSetup.sps)
  - Male respondent file (2011-2013_MaleRespSetup.sps)
- STATA Program Statements
  - Female respondent file
    - 2011-2013_FemRespSetup.dct
    - 2011-2013_FemRespSetup.do
  - Pregnancy file
    - 2011-2013_FemPregSetup.dct
    - 2011-2013_FemPregSetup.do
  - Male respondent file
    - 2011-2013_MaleSetup.dct
    - 2011-2013_MaleSetup.do

SAS, SPSS, and STATA program statements


Other Data Files

The NSFG ACASI data for 2011-2013 are now included as part of the main NSFG public use files. In addition to the main 2011–2013 public use data files, there are other files available containing Region
Click on “I accept…” and then you’ll be directed to FTP site:
FTP directory /pub/Health_Statistics/NCHS/Datasets/NSFG/ at ftp.cdc.gov

FTP directory /pub/Health_Statistics/NCHS/Datasets/NSFG/ at ftp.cdc.gov

To view this FTP site in File Explorer: press Alt, click View, and then click Open FTP Site in File Explorer.

Up to higher level directory

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Size</th>
<th>Filename</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/04/1998</td>
<td>12:00AM</td>
<td>0</td>
<td>.cache</td>
</tr>
<tr>
<td>09/04/1998</td>
<td>12:00AM</td>
<td>0</td>
<td>.cache+</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>19,838,925</td>
<td>1973NSFGData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>31,716,306</td>
<td>1976NSFGData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>32,648,974</td>
<td>1982NSFGData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>30,022,850</td>
<td>1988FemRespData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>55,991,727</td>
<td>1988PregData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>141,726,902</td>
<td>1995FemRespData.dat</td>
</tr>
<tr>
<td>04/05/2012</td>
<td>12:00AM</td>
<td>8,916,776</td>
<td>1995PregData.dat</td>
</tr>
<tr>
<td>11/15/2010</td>
<td>12:00AM</td>
<td>150,852</td>
<td>2002curr_ins.dat</td>
</tr>
<tr>
<td>02/23/2005</td>
<td>12:00AM</td>
<td>6,089,664</td>
<td>2002FemPreg.dat</td>
</tr>
<tr>
<td>02/23/2005</td>
<td>12:00AM</td>
<td>37,664,704</td>
<td>2002FemResp.dat</td>
</tr>
<tr>
<td>11/29/2007</td>
<td>12:00AM</td>
<td>289,133</td>
<td>2002HHvars.dat</td>
</tr>
<tr>
<td>02/23/2005</td>
<td>12:00AM</td>
<td>14,719,936</td>
<td>2002Male.dat</td>
</tr>
<tr>
<td>10/12/2011</td>
<td>12:00AM</td>
<td>11,557,488</td>
<td>2006_2010_FemPreg.dat</td>
</tr>
<tr>
<td>10/12/2011</td>
<td>12:00AM</td>
<td>76,780,587</td>
<td>2006_2010_FemResp.dat</td>
</tr>
<tr>
<td>10/12/2011</td>
<td>12:00AM</td>
<td>47,281,635</td>
<td>2006_2010_Male.dat</td>
</tr>
<tr>
<td>12/11/2014</td>
<td>12:00AM</td>
<td>4,580,640</td>
<td>2011_2013_FemPregData.dat</td>
</tr>
<tr>
<td>12/11/2014</td>
<td>12:00AM</td>
<td>27,769,758</td>
<td>2011_2013_FemRespData.dat</td>
</tr>
<tr>
<td>12/11/2014</td>
<td>12:00AM</td>
<td>22,327,155</td>
<td>2011_2013_MaleData.dat</td>
</tr>
<tr>
<td>04/06/2010</td>
<td>12:00AM</td>
<td>15,960,000</td>
<td>C5FieldworkData.asc</td>
</tr>
<tr>
<td>11/15/2010</td>
<td>12:00AM</td>
<td>312,320</td>
<td>C6_curr_ins.sas7bdat</td>
</tr>
<tr>
<td>01/20/2015</td>
<td>04:26PM</td>
<td>214,016</td>
<td>cyc2psu.sas7bdat</td>
</tr>
<tr>
<td>04/02/2010</td>
<td>12:00AM</td>
<td>312,320</td>
<td>hhpartypnew.sas7bdat</td>
</tr>
<tr>
<td>04/02/2010</td>
<td>12:00AM</td>
<td>150,852</td>
<td>HHPARTYPNEWASC.DAT</td>
</tr>
<tr>
<td>06/18/2015</td>
<td>11:31AM</td>
<td>Directory</td>
<td>sas</td>
</tr>
<tr>
<td>12/10/2014</td>
<td>12:00AM</td>
<td>Directory</td>
<td>spss</td>
</tr>
<tr>
<td>12/10/2014</td>
<td>12:00AM</td>
<td>Directory</td>
<td>stata</td>
</tr>
</tbody>
</table>

NSFG Documentation for 2011-2013

• User’s Guide
  – Part 1: Overview
  – Part 2: Topic-specific notes for analysts
  – 7 Appendixes including:
    • File Indexes (App 1)
    • Syntax Guidance for Common File Manipulations (App 2)
    • Recode Specs & Recode “Cross-walks” (Apps 3 & 4)
    • Questionnaire Changes since 2006-2010 NSFG (App 5)
    • FAQ for the NSFG (App 6)
    • List of restricted use variables and modified variables (App 7)

• Questionnaires in 2 formats:
  – CAPI-Lite Questionnaires
  – CAPI Reference Questionnaires (CRQs)

• Codebooks (entries for every variable on the 3 data files)
  – Webdoc: interactive, online, searchable
  – PDFs can be created from Webdoc
User’s Guide – Part 2

• **Topic-specific information** such as:
  – Abortion under-reporting for women
  – Quality of birth reporting for women
  – Sex education topics
  – Religion variables

• **File-specific information** on particular variables or series that may help in your analyses
  – Data issues or problems
  – Guidance on how variables are mapped (e.g., family planning services, method use)
Each codebook entry anchors you to file, section, series, etc.

Recode specs excerpt from Appendix 3 of User’s Guide

Search feature for key words or question numbers; can check by file
SECTION A: Calendar Instructions; Demographic Characteristics; Household Roster; Childhood Background

(AA) Age and date of birth
(AB) Marital/cohabiting status
(AC) Hispanic origin and race
(AD) Household roster
(AF) Regular school and GED
(AG) Childhood background

Create codebook for SECTION A: Calendar Instructions; Demographic Characteristics; Household Roster; Childhood Background
OUTCOME (315-315)

Variable Type: pregnancy recode

Description: Pregnancy outcome (recode)

<table>
<thead>
<tr>
<th>value</th>
<th>label</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LIVE BIRTH</td>
<td>6670</td>
</tr>
<tr>
<td>2</td>
<td>INDUCED ABORTION</td>
<td>993</td>
</tr>
<tr>
<td>3</td>
<td>STILLBIRTH</td>
<td>71</td>
</tr>
<tr>
<td>4</td>
<td>MISCARRIAGE</td>
<td>1451</td>
</tr>
<tr>
<td>5</td>
<td>ECTOPIC PREGNANCY</td>
<td>120</td>
</tr>
<tr>
<td>6</td>
<td>CURRENT PREGNANCY</td>
<td>238</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>9543</td>
</tr>
</tbody>
</table>

Universe: Applicable for all pregnancies

Rencode specs: pdf

SAS logic:

```sas
IF outcome=1 then OUTCOME=1; * impute this case */
ELSE if outcome=2 then OUTCOME=1;
ELSE if outcome=3 then OUTCOME=6;
ELSE if outcome=4 then OUTCOME=6;
ELSE if outcome=5 then OUTCOME=6;
ELSE if outcome=6 then OUTCOME=6;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=2;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=3;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=4;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=5;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=6;
ELSE if PREGEND1=1 or PREGEND2=1 or PREGEND3=1 then OUTCOME=6;
ELSE if BC-10 HOWIND1=3 (live birth) then OUTCOME=1;
ELSE if HOWEND in (8,9) then OUTCOME=1; /* impute this case */
END;
```

User note: If interested in all outcomes for a multiple-pregnancy pregnancy (i.e., 2 or more fetuses), use raw variables BC-1 PREGENDn. To determine if a pregnancy resulted in a multiple birth (e.g., twins, triplets), see computed variable nbirth_n.
2011-2013 Contextual Data Files

- Use remotely or in an RDC (NCHS, CDC/Atl, or Census Data Centers)
  ✦ application required (see: http://www.cdc.gov/rdc/)
  ✦ charges apply
- Smaller file for the 2011-2013 data
  - State FIPS code at time of interview
  - County FIPS code at time of interview
  - State FIPS code for residence at 2010 Census
Cycles 1-6, 2006-2010, and 2011-2013 of the NSFG

- Public use files from all previous cycles can be downloaded
- Questionnaires, User’s Guides, and Codebooks are on the NSFG web site

- ACASI/Omitted Items files (Cycle 4, 5, 6, & 2006-2010) are available free, upon request

- Contextual files for Cycle 5, 6, and 2006-2010 data available through RDC
NSFG - sample weights & design variables

The sample weights adjust for:

• Oversampling
• Screener and Interview non-response
• Adjusted to Census Bureau population estimates (age, race, sex)

• SAS, STATA, SPSS all have software that will use weights and design variables to compute correct sampling errors.
• Specific syntax examples are on the NSFG web site based on 2002 data, and can be adapted for 2011-2013 data.

It is important to use the weights & design variables to make valid national estimates using the NSFG.
# Weight variables for 2011-2013 NSFG

<table>
<thead>
<tr>
<th>Description</th>
<th>Variable name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample weight</td>
<td>Wgt2011_2013</td>
</tr>
<tr>
<td>Stratum</td>
<td>SEST</td>
</tr>
<tr>
<td>Cluster</td>
<td>SECU</td>
</tr>
</tbody>
</table>
Pooling data across NSFG cycles (female data) (can pool additional data files)

<table>
<thead>
<tr>
<th>Design Variable</th>
<th>2002</th>
<th>2006-2010</th>
<th>2011-2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stratum variable</td>
<td>SEST</td>
<td>SEST</td>
<td>SEST</td>
</tr>
<tr>
<td>Cluster/Panel Variable</td>
<td>SECU_R – fem resp</td>
<td>SECU</td>
<td>SECU</td>
</tr>
<tr>
<td></td>
<td>SECU_P – fem preg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final post-stratified, fully adjusted case weight</td>
<td>FINALWGT</td>
<td>WGTQ1Q16</td>
<td>WGT2011_2013</td>
</tr>
</tbody>
</table>

To pool data, define 4 new variables for each file, and then “set” or append the new datasets.

**2002 NSFG:**
- stratvar=sest
- panelvar=secu_r
- weightvar=finalwgt
- survey=2002

**2006-10 NSFG:**
- stratvar=sest
- panelvar=secu
- weightvar=wgtq1q16
- survey=2008

**2011-13 NSFG:**
- stratvar=sest
- panelvar=secu
- weightvar=wgt2011_2013
- survey=2012
Combining male & female data from 2011-2013

• Make sure you are using comparable variables
  – Consult recode specs and male/female recode crosswalk
  – This point also important when pooling data across years

• See example in Appendix 2 of 2011-2013 User’s Guide

DATA FEMDATA;
SET FEMRESP (KEEP=CASEID [other variables you wish to include]);
R_SEX=1; ** female;
RUN;

DATA MALEDATA;
SET MALERESP (KEEP=CASEID [other variables you wish to include]);
R_SEX=2; ** male;
RUN;

data MF_POOLED;
set femdata maledata;
RUN;
To learn more about the NSFG:

- Visit the NSFG webpage
  www.cdc.gov/nchs/nsfg.htm

- Join the NSFG Announcements ListServ

- Contact the NSFG team
  Email: nsfg@cdc.gov
  Phone: 301-458-4222

**NSFG team:** Joyce Abma, Anjani Chandra (team leader), Casey Copen, Kim Daniels, Jill Daugherty, Isaedmarie Febo-Vazquez, Gladys Martinez

Amy Branum, Chief, Reproductive Statistics Branch
Delton Atkinson, Director, Division of Vital Statistics
National Survey of Family Growth Examples with SAS and Stata code

1. Percent of never-married females aged 15-19 who have ever had sexual intercourse

2. Percent of sexually experienced females who have ever used the Pill

3. Percent of births in the 5 years before the interview that were intended at conception, by age and marital status at birth
Hands-On Example #1: Percent of never-married teens 15-19 who have ever had sexual intercourse: United States, 1988-2013

HADSEX by survey year, limited to teens (AGER 15-19) with FMARITAL=5

Use recode HADSEX instead of raw variable rhadsex.
HADSEX ( 4431-4431 )

Variable Type : recode

Description : Whether R has ever had sexual intercourse with a male (RECODE)

<table>
<thead>
<tr>
<th>value</th>
<th>label</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>YES, R EVER HAD INTERCOURSE</td>
<td>4858</td>
</tr>
<tr>
<td>2</td>
<td>NO, R NEVER HAD INTERCOURSE</td>
<td>743</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>5601</td>
</tr>
</tbody>
</table>

Universe : Applicable for all respondents

Recode specs: pdf
SAS Code: Never married female teens who had ever had sex

*create a library name to point to;
libname nsfg 'c:\wherever the data is stored';

options nocenter nofmterr;
*create value labels;
proc format;
value yesno 1="yes" 2="no";
run;
*open dataset and keep specified variables;
data FEMALE;
set library.dataset /* (replace with your PUF fem resp file filename) */
(keep=caseid hadsex fmarital ager sest secu wgt2011_2013);
*divide weight by 1000 to get numbers in thousands – optional;
wgt1000=wgt2011_2013/1000;
* Create subpopulation variable for all never married teens;
  nmteen=2;
  if ager lt 20 and fmarital=5 then nmteen=1;
run;
proc sort data=female out=FSORTED;
  by SEST SECU;
run;

*weighted frequency SE and 95% CI;
proc surveyfreq data=Fsorted;
cluster SECU;
stratum SEST;
title "Percentage of never married female teens aged 15-19 who ever had sex: 2011-2013";
table nmteen*hadsex/ NOCELLPERCENT NOTOTAL NOFREQ NOWT CL row NOSPARSE;
weight wgt1000;
format hadsex yesno.;
run;
SAS Output: Never married female teens who had ever had sex

Percentage of never married female teens aged 15-19 who ever had sex: 2011-2013

The SURVEYFREQ Procedure

Data Summary

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Strata</td>
<td>18</td>
</tr>
<tr>
<td>Number of Clusters</td>
<td>72</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>5601</td>
</tr>
<tr>
<td>Sum of Weights</td>
<td>60887.363</td>
</tr>
</tbody>
</table>

Table of nmteen by HADSEX

<table>
<thead>
<tr>
<th>nmteen</th>
<th>HADSEX</th>
<th>95% Confidence Limits for Percent</th>
<th>Row Percent</th>
<th>Std Err of Row Percent</th>
<th>95% Confidence Limits for Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>5.8065 - 7.8342</td>
<td><strong>44.0588</strong></td>
<td>2.8203</td>
<td>38.4045 - 49.7131</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>7.2945 - 10.0251</td>
<td>55.9412</td>
<td>2.8203</td>
<td>50.2869 - 61.5955</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>78.9827 - 82.7258</td>
<td>95.6631</td>
<td>0.4692</td>
<td>94.7224 - 96.6037</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>2.8828 - 4.4484</td>
<td>4.3369</td>
<td>0.4692</td>
<td>3.3963 - 5.2776</td>
</tr>
</tbody>
</table>
Stata Code: Never married female teens who ever had sex

*open dataset and keep specified variables
set more off
use "\cdec\project\femresp_2011_2013.dta", clear /* (replace with your PUF fem resp file filename) */
keep CASEID HADSEX FMARITAL AGER SEST SECU WGT2011_2013

*divide weight by 1000 to get numbers in thousands – optional
gen wgt1000=WGT2011_2013/1000

*create and assign value labels
label define yesno 1 "yes" 2 "no"
label values HADSEX yesno

*create teen never married subpop variable
gen teennm=.
replace teennm= 1 if AGER<=19 & FMARITAL==5

*set survey info
svyset [pweight=wgt1000], strata(SEST) psu(SECU)

*run survey tabulation of hadsex
tab HADSEX , subpop(teennm) /*unweighted frequencies*/
svy: tab HADSEX , subpop(teennm) count format(%13.2fc) /*weighted counts*/
svy: tab HADSEX , subpop(teennm) se percent ci /*weighted frequencies, se, 95%CI*/
Stata Output: Never married female teens who ever had sex

```
. svy: tab HADSEX, subpop(teenrm) se percent ci /*weighted frequencies, se, 95%CI*/
(running tabulate on estimation sample)

Number of strata = 18
Number of PSUs  = 72
Number of obs   = 1025
Population size = 9425.4695
Subpop. no. of obs = 1025
Subpop. size   = 9425.4695
Design df      = 54

<table>
<thead>
<tr>
<th>Whether R has ever had sexual intercourse with a male (RECODE)</th>
<th>percentages</th>
<th>se</th>
<th>lb</th>
<th>ub</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>44.06</td>
<td>2.82</td>
<td>38.5</td>
<td>49.77</td>
</tr>
<tr>
<td>no</td>
<td>55.94</td>
<td>2.82</td>
<td>50.23</td>
<td>61.5</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: percentages = cell percentages
    se      = linearized standard errors of cell percentages
    lb      = lower 95% confidence bounds for cell percentages
    ub      = upper 95% confidence bounds for cell percentages

. end of do-file
```
Hands-On Example #2: Percent of sexually experienced females 15-44 who have ever used the Pill: United States, 1982-2013

PILLR by survey year, limited to sexually experienced women (HADSEX=1)

Use recode PILLR instead of raw variable PILL.

Raw variable, pill, for all women

http://www.icpsr.umich.edu/icpsradmin/nsfg/variable/614733?studyNumber=9998
Link to pdf with recode specs. The recode variable corrects for later reports of pill use in Section E after question EA-1 (shown on previous slide) was asked.

http://www.icpsr.umich.edu/icpsradmin/nsfg/variable/616467?vg=7220&studyNumber=9998
SAS Code: Ever use of the pill among sexually experienced women

options nocenter nofmterr;
*create value labels;
proc format;
value yesno 1="yes" 2="no";
run;

*open dataset and keep specified variables;
data FEMALE;
set library.dataset /* (replace with your female respondent file filename) */
(keep=caseid wgt2011_2013 sest secu pillr hadsex);

*divide weight by 1000 to get numbers in thousands – optional;
wgt1000=wgt2011_2013/1000;
run;

proc sort data=female out=FSORTED;
by SEST SECU;
run;

*weighted frequency SE and 95% CI;
proc surveyfreq data=FSORTED;
cluster SECU;
stratum SEST;
title "Ever use of pill among sexually experienced women 2011-2013";
table hadsex*pillr/NOCELLPERCENT NOTOTAL NOFREQ NOWT CL row NOSPARSE;
weight wgt1000;
format pillr yesno. hadsex yesno.;
run;
**SAS Output: Ever use of the pill among sexually experienced women**

### Ever use of pill among sexually experienced women 2011-2013

**The SURVEYFREQ Procedure**

**Data Summary**

<table>
<thead>
<tr>
<th>Number of Strata</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Clusters</td>
<td>72</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>5601</td>
</tr>
<tr>
<td>Sum of Weights</td>
<td>60887.363</td>
</tr>
</tbody>
</table>

### Table of HADSEX by PILLR

<table>
<thead>
<tr>
<th>HADSEX</th>
<th>PILLR</th>
<th>95% Confidence Limits for Percent</th>
<th>Row Percent</th>
<th>Std Err of Row Percent</th>
<th>95% Confidence Limits for Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>66.0542</td>
<td>72.0541</td>
<td><strong>78.7618</strong></td>
<td>1.1706</td>
</tr>
<tr>
<td>no</td>
<td></td>
<td>16.7342</td>
<td>20.5068</td>
<td>21.2382</td>
<td>1.1706</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>1.9274</td>
<td>3.3976</td>
<td>21.6017</td>
<td>2.7360</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>8.1122</td>
<td>11.2136</td>
<td>78.3983</td>
<td>2.7360</td>
</tr>
</tbody>
</table>

Note: The highlighted row indicates the focus on the frequency and confidence intervals for the use of the pill among sexually experienced women.
STATA Code: Ever use of the pill among sexually experienced women

*open dataset and keep specified variables
set more off
use "\cdc\femresp_2011_2013.dta", clear /* (replace with your PUF fem resp file filename) */
keep CASEID WGT2011_2013 SEST SECU PILLR HADSEX

*divide weight by 1000 to get numbers in thousands – optional
gen wgt1000=WGT2011_2013/1000

*create and assign value labels
label define yesno 1 "yes" 2 "no"
label values HADSEX PILLR yesno

*set survey info
svyset [pweight=wgt1000], strata(SEST) psu(SECU)

*run cross tabulation of hadsex and pillr
tab HADSEX PILLR, row /*unweighted frequencies*/
svy: tab HADSEX PILLR, count format(%13.2fc) /*weighted counts*/
svy: tab HADSEX PILLR, se row percent ci /* weighted frequencies, se, 95%CI*/
### STATA Output: Ever use of the pill among sexually experienced women

```
.svyr: tab HADSEX PILLR, se row percent ci /* weighted frequencies, se, 95%CI*/
(running tabulate on estimation sample)

Number of strata = 18
Number of PSUs = 72
Number of obs = 5601
Population size = 60887.363
Design df = 54

<table>
<thead>
<tr>
<th>Whether R has ever had sexual intercourse with a male (RECODE)</th>
<th>Ever used the pill for any reason (RECODE)</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>78.76</td>
<td>21.24</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.171)</td>
<td>(1.171)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[76.32, 81.01]</td>
<td>[18.99, 23.68]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>21.6</td>
<td>78.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(2.736)</td>
<td>(2.736)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[16.62, 27.59]</td>
<td>[72.41, 83.38]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>yes</td>
<td>71.72</td>
<td>28.28</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.458)</td>
<td>(1.458)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[68.7, 74.55]</td>
<td>[25.45, 31.3]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: row percentages (linearized standard errors of row percentages) [95% confidence intervals for row percentages]

Pearson:
Uncorrected chi2(1) = 974.9328
Design-based F(1, 54) = 366.9956 P = 0.0000

. end of do-file
```
Hands-On Example #3: Percent of births in the 5 years before the interview that were intended at conception, by mother’s age and marital status at delivery: United States, 2011-2013

Source: NCHS special tabulation. For prior data releases see Table 2 and Figure 2 in Mosher WD, Jones J, Abma J. Intended and Unintended Births in the United States: 1982-2010. NHHSR #55. Hyattsville, MD: National Center for Health Statistics. 2012.
WANTRESP (350-350)

Variable Type: pregnancy recode

Description: Wantedness of pregnancy - respondent - Cycle 5, 6 version (recode)

<table>
<thead>
<tr>
<th>Value</th>
<th>Label</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LATER, OVERDUE</td>
<td>566</td>
</tr>
<tr>
<td>2</td>
<td>RIGHT TIME</td>
<td>4012</td>
</tr>
<tr>
<td>3</td>
<td>TOO SOON, MISTIMED</td>
<td>2913</td>
</tr>
<tr>
<td>4</td>
<td>DIDN'T CARE, INDIFFERENT</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>UNWANTED</td>
<td>1924</td>
</tr>
<tr>
<td>6</td>
<td>DON'T KNOW, NOT SURE</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>9543</td>
</tr>
</tbody>
</table>

Universe: Applicable for all pregnancies

Recode specs: pdf

http://www.icpsr.umich.edu/icpsradmin/nsfg/variable/613604?studyNumber=9998
SAS Code (1): Percent of births that were intended at conception, by age and marital status at delivery

options nocenter nofmterr;

data preg201113;
  set library.dataset /* (replace with your PUF preg file filename) */
   (keep = AGER WANTRESP OUTCOME DATEND RMAROUT6 AGEPREG cmintvw WGT2011_2013 SECU SEST);

  *divide weight by 1000 to get numbers in thousands – optional;
  WGT1000 = WGT2011_2013/1000;

  * Create a new variable based on the wantedness recode (WANTRESP) to combine two categories to result in "intended" births
  * and assign 'didn't care' and 'don't know' to the category "intended" (individual analyst's decision)
  * all other codes will remain as they are on the recode.
  WANTREC=WANTRESP;
  if WANTRESP in (1,2) then WANTREC=1;  * if pregnancy occurred at the right time or later than R wanted, we classify it as INTENDED ;
  if WANTRESP in (4,6) then WANTREC=1;  * if respondent didn't care/was indifferent, or didn't know, we coded as INTENDED ;

  * Create a new variable based on the "marital status at outcome" recode to combine the 4 "not currently married" categories into 1
  MARCOHAB = .;
  IF RMAROUT6 = 1    THEN MARCOHAB = 1; ELSE   *MARRIED AT TIME OF OUTCOME; 
  IF RMAROUT6 = 5    THEN MARCOHAB = 2; ELSE   *COHABITING AT TIME OF OUTCOME;
  IF RMAROUT6 in (2 3 4 6) THEN MARCOHAB = 3;  *NOT MARRIED/NOT COHABITING AT TIME OF OUTCOME;

  * create a subpopulation variable for births in the last 5 years
  BIRTHFIV=.;
  if OUTCOME=1 and (DATEND >= (cmintvw - 60)) then BIRTHFIV=1;
run;
SAS Code (2): Percent of births that were intended at conception, by age and marital status at delivery

*assign value labels;
proc format;
value WANTF 1='Intended' 3='TOO SOON, MISTIMED' 5='UNWANTED';
value MARCOHF 1='MARRIED' 2='COHABITING' 3='NOT MARRIED OR COHABITING';
value AGEF 0-1999='UNDER 20 YEARS' 2000-2499='20-24 YEARS' 2500-4499='25-44 YEARS';
run;

proc sort data=preg201113 out=SORTED;
  by SEST SECU;
run;

*weighted frequency SE and 95% CI;
proc surveyfreq data=SORTED;
  weight WGT1000;
  cluster SECU;
  stratum SEST;
  tables BIRTHFIV*AGEPREG*WANTREC/ ROW CL NOCELLPERCENT NOTOTAL nowt NOSPARSE;
  tables BIRTHFIV*MARCOHAB*WANTREC/ ROW CL NOCELLPERCENT NOTOTAL nowt NOSPARSE;
  TITLE '2011-2013: Live births within the past 5 years - intendedness of birth at conception by age at birth and union status at birth';
  format AGEPREG agef. MARCOHAB marcohf. WANTREC wantf.;
run;
SAS Output: Percent of births that were intended at conception, by age at delivery

2011-2013: Live births within the past 5 years - intendedness of birth at conception by age at birth and union status at birth

The SURVEYFREQ Procedure

Data Summary

Number of Strata: 18
Number of Clusters: 72
Number of Observations: 9543
Sum of Weights: 105355.191

<table>
<thead>
<tr>
<th>AGEPREG</th>
<th>WANTREC</th>
<th>Frequency</th>
<th>95% Confidence Limits for Percent</th>
<th>Row Percent</th>
<th>Std Err of Row Percent</th>
<th>95% Confidence Limits for Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNDER 20 YEARS</td>
<td>Intended</td>
<td>69</td>
<td>1.9595 4.3538</td>
<td>31.4327</td>
<td>4.4238</td>
<td>22.5636 40.3018</td>
</tr>
<tr>
<td></td>
<td>TOO SOON, MISTIMED</td>
<td>143</td>
<td>4.1361 6.8114</td>
<td>54.5055</td>
<td>3.9174</td>
<td>46.6515 62.3595</td>
</tr>
<tr>
<td></td>
<td>UNWANTED</td>
<td>42</td>
<td>0.7466 2.0777</td>
<td>14.0618</td>
<td>2.9194</td>
<td>8.2087 19.9149</td>
</tr>
<tr>
<td>20-24 YEARS</td>
<td>Intended</td>
<td>303</td>
<td>10.8519 15.5311</td>
<td>53.0120</td>
<td>3.0025</td>
<td>46.9924 59.0317</td>
</tr>
<tr>
<td></td>
<td>UNWANTED</td>
<td>104</td>
<td>2.6193 4.7219</td>
<td>14.7509</td>
<td>2.1162</td>
<td>10.5082 18.9935</td>
</tr>
<tr>
<td>25-44 YEARS</td>
<td>Intended</td>
<td>883</td>
<td>44.9825 53.4571</td>
<td>75.6373</td>
<td>2.0329</td>
<td>71.5615 79.7131</td>
</tr>
</tbody>
</table>
SAS Output: Percent of births that were intended at conception, by marital status at delivery

2011-2013: Live births within the past 5 years - intendedness of birth at conception by age at birth and union status at birth

The SURVEYFREQ Procedure

Table of MARCOHAB by WANTREC
Controlling for BIRTHFIV=1

<table>
<thead>
<tr>
<th>MARCOHAB</th>
<th>WANTREC</th>
<th>Frequency</th>
<th>95% Confidence Limits for Percent</th>
<th>Row Percent</th>
<th>Std Err of Row Percent</th>
<th>95% Confidence Limits for Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARRIED</td>
<td>Intended</td>
<td>736</td>
<td>40.2688</td>
<td>49.9590</td>
<td>80.4069</td>
<td>1.7397</td>
</tr>
<tr>
<td></td>
<td>TOO SOON, MISTIMED</td>
<td>143</td>
<td>5.6247</td>
<td>8.9331</td>
<td>12.9733</td>
<td>1.3733</td>
</tr>
<tr>
<td></td>
<td>UNWANTED</td>
<td>71</td>
<td>2.3952</td>
<td>5.0331</td>
<td>6.6198</td>
<td>1.2180</td>
</tr>
<tr>
<td>COHABITING</td>
<td>Intended</td>
<td>326</td>
<td>11.8617</td>
<td>16.9464</td>
<td>55.5873</td>
<td>2.6630</td>
</tr>
<tr>
<td></td>
<td>TOO SOON, MISTIMED</td>
<td>189</td>
<td>6.2345</td>
<td>9.1781</td>
<td>29.7397</td>
<td>2.1226</td>
</tr>
<tr>
<td></td>
<td>UNWANTED</td>
<td>109</td>
<td>2.9336</td>
<td>4.6707</td>
<td>14.6730</td>
<td>1.7203</td>
</tr>
<tr>
<td>NOT MARRIED OR COHABITING</td>
<td>Intended</td>
<td>193</td>
<td>4.6275</td>
<td>7.4724</td>
<td>33.6474</td>
<td>3.2389</td>
</tr>
<tr>
<td></td>
<td>TOO SOON, MISTIMED</td>
<td>187</td>
<td>5.2544</td>
<td>8.1566</td>
<td>37.2932</td>
<td>3.3626</td>
</tr>
<tr>
<td></td>
<td>UNWANTED</td>
<td>127</td>
<td>3.6574</td>
<td>6.7926</td>
<td>29.0594</td>
<td>3.3913</td>
</tr>
</tbody>
</table>
**Stata Code (1): Percent of births that were intended at conception, by age and marital status at delivery**

*open dataset and keep specified variables*
set more off
use "\cdc\fempreg_2011_2013.dta", clear /* (replace with your PUF fem resp file filename) */
keep AGER WANTRESP OUTCOME DATEND RMAROUT6 AGEPREG CMINTVW WGT2011_2013 SECU SEST

*divide weight by 1000 to get numbers in thousands – optional*
gen wgt1000=WGT2011_2013 /1000

* Create a new variable based on the wantedness recode (WANTRESP) to combine two categories to result in "intended" births*
* and assign 'didn’t care’ and ‘don’t know’ to the category "intended" (individual analyst’s decision)*
* all other codes will remain as they are on the recode. ;*
gen wantrec=WANTRESP
replace wantrec=1 if WANTRESP==1 | WANTRESP==2 /* if pregnancy occurred at the right time or later than r wanted, we classify it as intended */
replace wantrec=1 if WANTRESP==4 | WANTRESP==6 /* if respondent didn't care/was indifferent, or didn't know, we coded as INTENDED */

/* Create a new variable based on the "marital status at outcome" recode to combine the 4 "not currently married" categories into 1 */
gen marcohab=.
replace marcohab=1 if RMAROUT6==1 /*married at the birth*/
replace marcohab=2 if RMAROUT6==5 /*cohabiting at the birth*/
replace marcohab=3 if RMAROUT6==2 | RMAROUT6==3 | RMAROUT6==4 | RMAROUT6==6 /*no union (not married, not cohabiting) at the birth*/

*create a subpopulation of births in last 5 years*
gen birthfiv=.
replace birthfiv=1 if (OUTCOME==1) & (DATEND>=CMINTVW - 60)

* create an age category variable*
gen agecat=.
replace agecat=1 if (AGEPREG>=0000) & (AGEPREG<=1999)
replace agecat=2 if (AGEPREG>=2000) & (AGEPREG<=2499)
replace agecat=3 if (AGEPREG>=2500) & (AGEPREG<=4499)
Stata Code (2): Percent of births that were intended at conception, by age and marital status at delivery

*create and assign value labels
label define wantf 1 "intended" 3 "too soon, mistimed" 4 "didnt care, indifferent" 5 "unwanted"
label define marcohf 1 "married" 2 "cohabiting" 3 "not married or cohabiting"
label define agef 1 "under 20 years" 2 "20-24 years" 3 "25-44 years"

label values agecat agef
label values marcoh marcohf
label values wantrec wantf

*set survey info
svyset [pweight=WGT2011_2013], strata(SEST) psu(SECU)

*run cross tabulation of intendedness and maternal age
tab agecat wantrec, subpop(birthfiv) row /* unweighted frequencies */
svy: tab agecat wantrec, subpop(birthfiv) count format(%13.2fc) /*weighted counts*/
svy: tab agecat wantrec, subpop(birthfiv) se row percent ci /* weighted frequencies, se, 95%CI */

*run cross tabulation of intendedness and marital status
tab marcoh marcohf wantrec, subpop(birthfiv) row /* unweighted frequencies */
svy: tab marcoh marcohf wantrec, subpop(birthfiv) count format(%13.2fc) /*weighted counts*/
svy: tab marcoh marcohf wantrec, subpop(birthfiv) se row percent ci /* weighted frequencies, se, 95%CI */
Stata Output: Percent of births that were intended at conception, by age at delivery

```
.svy: tab agecat wantrec, subpop(birthfiv) se row percent ci /* weighted frequencies, se, 95%CI */
```

<table>
<thead>
<tr>
<th></th>
<th>intended</th>
<th>too soon</th>
<th>unwanted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>under 20</strong></td>
<td>31.43</td>
<td>54.51</td>
<td>14.06</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(4.424)</td>
<td>(3.917)</td>
<td>(2.919)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[23.5,40.89]</td>
<td>[46.6,62.19]</td>
<td>[9.158,20.99]</td>
<td></td>
</tr>
<tr>
<td><strong>20-24 ye</strong></td>
<td>53.01</td>
<td>32.24</td>
<td>14.75</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(3.003)</td>
<td>(2.76)</td>
<td>(2.116)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[46.98,59.96]</td>
<td>[26.97,38]</td>
<td>[10.99,19.51]</td>
<td></td>
</tr>
<tr>
<td><strong>25-44 ye</strong></td>
<td>75.64</td>
<td>12.59</td>
<td>11.77</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(2.033)</td>
<td>(1.364)</td>
<td>(1.415)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[71.34,79.48]</td>
<td>[10.1,15.59]</td>
<td>[9.214,14.92]</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>65.57</td>
<td>21.69</td>
<td>12.74</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.676)</td>
<td>(1.262)</td>
<td>(1.175)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[62.13,68.85]</td>
<td>[19.27,24.33]</td>
<td>[10.57,15.29]</td>
<td></td>
</tr>
</tbody>
</table>

Key: row percentages
(linearized standard errors of row percentages)
[95% confidence intervals for row percentages]

Pearson:
Uncorrected chi2(4) = 255.4315
Design-based F(3.53, 190.87) = 36.1633  P = 0.0000
Stata Output: Percent of births that were intended at conception, by marital status at delivery

```
.svy: tab marcohab wantrec, subpop(birthfiv) se row percent ci /* weighted frequencies, se, 95%CI */
(running tabulate on estimation sample)

<table>
<thead>
<tr>
<th></th>
<th>wanted</th>
<th>too soon</th>
<th>unwanted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>marcohab</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>married</td>
<td>80.41</td>
<td>12.97</td>
<td>6.62</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.74)</td>
<td>(1.373)</td>
<td>(1.218)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[76.68,85.06]</td>
<td>[10.46,15.98]</td>
<td>[4.558,9.521]</td>
<td></td>
</tr>
<tr>
<td>cohabiti</td>
<td>55.59</td>
<td>29.74</td>
<td>14.67</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(2.663)</td>
<td>(2.123)</td>
<td>(1.72)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[50.2,60.84]</td>
<td>[25.67,34.16]</td>
<td>[11.55,18.47]</td>
<td></td>
</tr>
<tr>
<td>not marr</td>
<td>33.65</td>
<td>37.29</td>
<td>29.06</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(3.239)</td>
<td>(3.363)</td>
<td>(3.391)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[27.49,40.42]</td>
<td>[30.83,44.24]</td>
<td>[22.75,36.29]</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>65.57</td>
<td>21.69</td>
<td>12.74</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.676)</td>
<td>(1.262)</td>
<td>(1.175)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[62.13,68.85]</td>
<td>[19.27,24.33]</td>
<td>[10.57,15.29]</td>
<td></td>
</tr>
</tbody>
</table>

Key: row percentages
(Linearized standard errors of row percentages)
[95% confidence intervals for row percentages]

Pearson: Uncorrected chi2(4) = 318.6721
Design-based F(3.74, 201.82) = 45.8088  P = 0.0000

end of do-file
```