



# National Survey of Sexual Attitudes and Lifestyles, 2010-2012: Teaching Dataset

## User Guide

Author: Alle Bloom and Jennifer Buckley

Version: 1.0

Date: October 2020

## Contents

Introduction .....	3
Background to Natsal.....	3
Survey design .....	3
Response rate.....	4
Overview of the Natsal teaching dataset .....	4
Topics.....	4
Additional variables.....	4
Sexual conservatism .....	4
Age at birth of first child .....	5
Missing Values .....	5
Don't know and refusal .....	5
Not applicable.....	5
Survey methodology variables .....	6
Survey weight .....	6
Survey design variables .....	6
Variables in the dataset .....	7
Variable list .....	7
Code for creating the dataset (SPSS Syntax) .....	12

## Introduction

The [British National Surveys of Sexual Attitudes and Lifestyles \(Natsal\)](#) are some of the largest and most detailed studies of sexual attitudes and behaviour in the world.

This teaching dataset contains data from Natsal-3, which interviewed 15,162 adults aged 16-74 in 2010-2012. The data all comes from the [original study accessible via the UK Data Service](#) but with changes to make it more accessible for teaching including reducing the number of variables.

This user guide provides a background to Natsal including details of the survey design along with details of the teaching dataset including a variable list, syntax used to create the dataset and full codebook.

## Background to Natsal

The National Surveys of Sexual Attitudes and Lifestyles ran first in 1990-1991. Natsal originated in the mid-1980s as a response to the HIV epidemic and the need for population-based data to understand transmission. Further studies have taken place every 10 years since with Natsal-2 in 1999-2001 and Natsal-3 in 2010-2012. Natsal-4 was under development at the time of writing.

Data from these studies have improved our understanding of sexual behaviour and provided evidence to inform policy and practice. [Initial findings from Natsal-3 are available to read online via the Lancet.](#)

## Survey design

Natsal-3 involved face-to-face interviews, using Computer-Assisted Personal Interviewing (CAPI), and a self-completion part, using Computer-Assisted Self Interviewing (CASI).

The face-to-face interview included questions on general health, family when growing up, learning about sex, first sexual experience, contraception, periods and menopause, attitudes to different kinds of relationships and sociodemographic details including cohabitation history.

More sensitive questions came in the self-completion part to reassure respondents of confidentiality and to avoid embarrassment. Topics covered in this part include questions about sexual histories and recreational drug use.

Not all respondents received the self-completion questions, it depended on how they answered previous questions (survey routing). Those with no sexual experience did not get the CASI questionnaire, while those with some sexual experiences got a shortened version. This design feature means that some cases are recorded as 'not applicable' for variables coming from the CASI part of the survey. Overall, 1.9% of respondents were not eligible for this module at all, and of the 98.1% that were eligible for any form of the CASI, 3.9% were given the shortened version and 2.2% refused to complete any of the CASI.

## Response rate

The sample design for Natsal-3 required that one randomly selected adult aged 16-74 (16-34 in the boost samples) be interviewed in each eligible sampled address. In total 59,412 addresses were sampled, 24,924 in the core sample and 18,537 and 15,951 in the boost samples.

Once removing those who were ineligible (based on building type or being out of the age range), those whose eligibility was unknown (through lack of contact or refusing information) and the estimated ineligibles, 26,274 estimated eligible addresses remained.

In total, interviews were completed with 15,162 respondents.

Using AAPOR's formula for calculating RR3 (Response rate 3) NATSAL-3 achieved an overall response rate of 57.7%.

## Overview of the Natsal teaching dataset

### Topics

The teaching dataset contains data on a wide of topics including:

- family background
- sources of sex education
- methods and sources of contraception
- sexual attraction and sex history
- fertility
- alcohol, smoking and other drug use
- relationships including relationship status and happiness
- attitudes towards sexual lifestyles and behaviours such as adultery, same sex relationships and sex in the media
- and many demographic variables

A list of variables by topic is can be found below.

### Additional variables

This teaching dataset contains additional variables designed to support teaching and learning.

### Sexual conservatism

The variable *attconservative* measures sexual attitudes.

It has been derived using a statistical method called principal component analysis (PCA), which can be used to analyse patterns of responses to multiple questions. One benefit of methods such as PCA is that they can help measure complex concepts that can be hard to measure with a single survey question.

In this case, we used 7 questions measuring opinions relating to sexual behaviours and lifestyles, these were:

- What is your opinion about a person having one night stands? Right/wrong
- The rest of the questions asked respondents to say whether they agreed strongly, agreed, neither agreed nor disagreed, disagreed, disagreed strongly with the following statements:
  - It's OK to have sex with someone without being in love with them
  - People are under a lot of pressure to have sex nowadays
  - It is natural for people to want sex less as they get older
  - Men have a naturally higher sex drive than women
  - There's too much sex in the media these days
  - Young people today start having sex too early

Using PCA, we derived a variable where a higher response indicates a more conservative attitude towards sexual behaviour in general. The new variable is a continuous variable (with a distribution close to normal) and can be examined using techniques such as correlation and regression.

The SPSS syntax used to create the variable is in the syntax section below.

Age at birth of first child

The variable Dage1ch derives from age1ch but with top and bottom coded values removed. Variables that are top-coded group all observations with values above a set upper bound into one category. In age1ch, this upper bound is anyone who had their first child over the age of 40. Bottom-coded variables groups all observations with values below a lower limit. In age1ch the lower bound is anyone who had their first child under the age of 15.

Creating a new version without the top and bottom coding is designed to simplify analysis of the variable using methods such as OLS linear regression. The original variable age1ch remains in the dataset.

Missing Values

Don't know and refusal

When asked a question, survey participants may respond 'do not know' or refuse to answer. Such responses are recorded using specific codes, often something distinctive compared to other values such 99 or 999.

To help with analysis of this dataset, most 'not answered' and 'don't know' values are pre-set as missing values in the SPSS and Stata versions. The main exception is the sexual attitudes variables where the category labelled 'Depends/Don't Know' is kept as a valid value.

Not applicable

Many variables in the dataset have a category of not applicable, generally coded as -1. The not applicable category applies when a specific case is not given a value, usually because they were not asked the question. For instance, the variable age1ch records the age of respondent at the birth of their 1<sup>st</sup> child; naturally, this variable is only applicable to those who have had children.

Additionally, not applicable categories apply for questions asked within the more sensitive self-completion part of the survey, which not all participants received (see section on [survey design](#) above for more details).

## Survey methodology variables

### Survey weight

Many survey datasets contain variables called survey weights. These variables are made by the data collectors for you to apply when analysing data. We use weights to make sample data better represent the population it's designed to reflect by adjusting for over and under-represented cases. Under and over-representation can result from the complex sampling methods used in large scale surveys and other issues including non-response.

In this teaching dataset, the survey weight variable **Total\_wt** is for use when analysing the total sample (including the boosts of younger people). The weight accounts for unequal selection probabilities from the sample design (selection weight) and for non-response by adjusting the distribution of age, sex and regional profiles to match the general population (poststratification weight).

For more information about the weight variable, see the [Natsal-3 documentation here](#).

### Survey design variables

Natsal-3 uses a stratified, clustered sample design with postcode sectors as the Primary Sampling Units (PSUs). The sampling frame is the (small users) Postcode Address File (PAF).

The teaching datasets contains a weight variable and variables for strata and PSU (`total_wt`, `strata` and `psu_scrm`), which can be used to account for clustering, weights and post-stratification.

In addition to strata, the dataset contains three additional grouped strata variables. These grouped variables can be used when stratas contain too few PSUs to use 'strata'.

## Variables in the dataset

The teaching dataset contains 130 of the original 1224 variables in the dataset.

The survey questions can be found in the [questionnaire accessible with the main study data](#).

The documentation that accompanies this teaching dataset includes a codebook for this teaching dataset.

## Variable list

### Survey Methodology

Variable Name	Variable Label
sin2	Scrambled serial number
dateyoi	Year of Interview
total_wt	Total weight for Natsal 2010(core and boost samples)
psu_scrm	Scrambled psu identifier
strata	Strata
stratagr1	Grouped strata, combining adjacent strata
stratagr2	Grouped strata, combining 3 adjacent strata
stratagr3	Grouped strata, combining 34 adjacent strata

### Demographic

Variable Name	Variable Label
dage	Respondent's age at interview,years
rdoby	Respondent's year of birth
rsex	Respondent's sex
ethnicgrp	Ethnic group that respondent considers themselves belonging to, grouped
sexid	Sexual identity

### Health and Disability Variables

Variable Name	Variable Label
health	Respondents opinion of own health
disabil2	Longstanding illness, disability or infirmity
bodymass	BMI

### Body measurement

Variable Name	Variable Label
bodymass	BMI

### Alcohol and smoking

Variable Name	Variable Label
drink	Ever drink alcohol nowadays
alcohol2	Average alcoholic consumption per week (grouped into: less or more than recommended)
smoking	Current smoking status (light, heavy, non, ex)

### Drugs

Variable Name	Variable Label
drcannabis	Ever taken: cannabis (non-injected)
drampheta	Ever taken: amphetamines (non-injected)
drcoaine	Ever taken: cocaine (non-injected)
drcrack	Ever taken: crack (non-injected)

drecstasy	Ever taken: ecstasy (non-injected)
drnonihero	Ever taken: non-injected heroin (non-injected)
dracidlsd	Ever taken: acid/LSD (non-injected)
drcrysmeth	Ever taken: crystal meth (non-injected)
dramylnit	Ever taken: amyl nitrates (non-injected)
drothnonpre	Ever taken: other non-prescribed drugs (non-injected)
inject2	Ever injected non-prescribed drugs/other substances
drugsyr2	drug use in last year

#### Depression

Variable Name	Variable Label
mscore	Total score on PHQ-2
depscr	Screen positive for current depression(based on PHQ-2, mood)

#### Sexual attraction and Number of partners

Variable Name	Variable Label
attscale	Sexual Attraction
expscale	Heterosexual/homosexual sexual experience
hetlife	No. of het. sex partners, life
samlife	No. of same sex partners, life

#### Marital and Relationship status

Variable Name	Variable Label
marstat	Marital status at present
relstat	Relationship status at the time of interview
relstatgp	Relationship status, grouped
relstatgp2	Relationship status (grouped)
relsat	How happy in relationship with partner

#### Fertility

Variable Name	Variable Label
nochall	Number of natural children (excl stillborn),All respondents
dage1ch	Age of respondent at birth of 1st child (no top or bottom coding)
age1ch	Age of respondent at birth of 1st child
child18	Respondent had their first child before the age of 18 years
nochb435	Did not have any natural children by the age of 35 (those aged 35+ only)
adopchld	Has adopted children
stepchld	Has Step-children

#### Contraception

Variable Name	Variable Label
fpsorcg	Source of contraceptive supplies, grouped
fppref2	Preferred source of contraception
usfsteril	Usually use: female sterilisation
usmsteril	Usually use: male sterilisation
usconpill	Usually use: contraceptive pill
uscondom	Usually use:male condom
usfemidom	Usually use: femidom
usemcpill	Usually use: morning after pill
usemcyiud	Usually use: emergency IUD
uscoil	Usually use: IUD
usmirena	Usually use: mirena



uscap	Usually use:cap
usinjectn	Usually use: contraceptive injections
usgels	Usually use: foams/gels/sprays/pessaries
usrhythm	Usually use: rhythm method
uswithdrw	Usually use: withdrawal
usimplant	Usually use: contraceptive implants
usother	Usually use: other contraception
usconptch	Usually use: Contraceptive patch
usnotused	Not had a usual contraceptive method in the last year

#### Family structure

Variable Name	Variable Label
parents4	Family structure when aged 14, (5 groups)
sibpos2	Position in family when growing up
siblive	live with siblings (natural, half, adopted or step) at any time when growing up

#### Sex Education

Variable Name	Variable Label
talkmapa	Ease with which discussed sex with parents around age 14
sexeducoc	Source of sex education - doctor/nurse/clinic
sexeduintad	Source of sex education - Internet-sexual advice websites
sexedusch	Source of sex education – lessons at school
sexeduinto	Source of sex education – Internet-other
sexedubmn	Source of sex education - books/magazines/newspapers
sexedufrn	Source of sex education - friends of about own age
sexedutvid	Source of sex education - TV/radio/DVDs/videos
sexedumum	Source of sex education - mother (incl step/adoptive)
sexedadad	This is missing from the dataset when using the syntax below
sexedufsp	Source of sex education - 1st girlfriend/boyfriend/sexual partner
sexeduintpn	Source of sex education - Internet - pornographic websites
sexedupor	Source of sex education - Pornographic magazines/films
sexedubsis	Source of sex education - brother/sister (incl half/step/adopted)
sexeduoth	Source of sex education – Other
sexeduani	Source of sex education - animals/growing up on a farm
sexedurel	Source of sex education - other relative(s)
sexeduofri	Source of sex education - older friends
sexeduofrn	Source of sex education - friends(age not specified)
sexeduexp	Source of sex education - through (own) experience
sexedufil	Source of sex education - cinemas/film (other than TV or video)
sexeduedc	Source of sex education - lessons/courses(i.e. formal teaching)

#### Opinions and attitudes

Variable Name	Variable Label
rwadult	Opinion about adultery when married : right/wrong
rwcasual	Opinion about one night stands? right/wrong.
rwsamm	Opinion about sexual relations between two adult men? right/wrong
rwsamf	Opinion about sexual relations between two adult women? right/wrong

snnolov	Sex without love OK
snpres	People are under pressure to have sex
snold	Natural for people to want sex less as they get older
snsexdrv	Men have a naturally higher sex drive than women
snmedia	Too much sex in the media
snearly	Young people today start having sex too early
sngayadp	Gay men should be able to adopt children
snlesadp	Lesbians should be able to adopt children
snsexed	Sex Education encourages young people to have sex
uagesex	Respondents perception of the proportion of young people having sex before 16
attconservative	Sexually conservative attitudes scale, higher score = more conservative attitudes

#### Religion

Variable Name	Variable Label
religimp	Importance of religion and religious beliefs now
whirel	Which religion belongs to
oftrelig	Frequency of attendance to religious services/meetings

#### Socio-economic

Variable Name	Variable Label
ractivhi3	Respondent's Highest occupation code
employed	Employment status at interview
rempstat	Employee/self-employed?
rnssecgp_6	Respondent's NSSEC code (8 groups)
rnssecgp_4	Respondent's NSSEC code (5 groups)
income	Household income per year (incl benefits, pensions etc)
parsc3	Parent's social class grouped (orig: par1occ, par2occ, par3occ)
educ3	Highest educational qualification (academic/non-academic), grouped
netacc	Does your household have access to the Internet from home?

#### Deprivation

Variable Name	Variable Label
adj_imd_quintile	Quintile of adjusted IMD (For Great Britain)
qimd	England: IMD score (2010) -Quintiles
qwimd	Wales: IMD score (2011) -Quintiles
qsimd	Scotland: IMD score (2009) -Quintiles

#### Accommodation/settlement variables

Variable Name	Variable Label
tenure	Own / rent accommodation
livehere	Always lived in village / town / city

#### Geographic variables

Variable Name	Variable Label
gor_l	Government Office Region(inner & outer London specified)
urindew	Urban/Rural Indicator (England & Wales)
urindsc	Urban/Rural Indicator (Scotland)



## Code for creating the dataset (SPSS Syntax)

```
Get File 'eul_natsal_2010_for_archive.sav'.

*Create new variables
**PCA to create a factor for sexually conservative attitudes.
missing values snnolov snpres snold snsexdrv snmedia snearly snsexed
(8, 9).
FACTOR
  /VARIABLES rwcasual snnolov snpres snold
snsexdrv snmedia snearly
  /MISSING LISTWISE
  /ANALYSIS rwcasual snnolov snpres snold
snsexdrv snmedia snearly
  /PRINT INITIAL EXTRACTION FSCORE
  /PLOT EIGEN
  /CRITERIA MINEIGEN(1) ITERATE(25)
  /EXTRACTION PC
  /ROTATION NOROTATE
  /SAVE REG(ALL)
  /METHOD=CORRELATION.

*rename generated factor, only using the first one here (though
others interesting too).
COMPUTE attconservative = (FAC1_1)*-1.
VARIABLE LABELS attconservative 'Sexually conservative attitudes
scale, higher score = more conservative attitudes'.
Execute.
DELETE VARIABLES FAC2_1 FAC3_1.

*Create another version of the age of first child variable without
the top or bottom coding.
RECODE age1ch (14=-1) (41=-1) (ELSE=copy) into dagelch.
VARIABLE LABELS dagelch 'Age of respondent at birth of 1st child (no
top or bottom coding)'.
VALUE LABELS dagelch
-1 'Not applicable including top coded'
99 'Not answered'.

**Set missing values.
missing values ethnicgrp health disabil2 drink alcohol2 smoking
parents4 sibpos2 (9).
missing values sexedudoc sexeduintad sexedusch sexeduinto sexedubmn
sexedufrn fpsorcgp fppref2
sexedutvid sexedumum sexedudad sexedufsp sexeduintpn sexedupor
sexedubsis sexeduoth sexeduani sexedurel sexeduofri sexeduofrn
sexeduexp sexedufil sexeduedc
usfsteril usmsteril usconpill uscondom usfemidom usemcypill
usemcyiud uscoil usmirena uscap usinjectn usgels usrhythm uswithdrw
usimplant usother
  usconptch usnotused child18 nochb435 adopchld stepchld relsat
talkmapa drcannabis drampheta drcocaine drcrack drecstasy drnonihero
dracidlsd
```

```

drcrysmeth dramylnit drothnonpre inject2 drugsyr2 parsc3 oftrelig
netacc mscore depscr (-1, 9).
missing values rsex (-1, -8, -9).
missing values bodymass ractivhi3 rnssecgp_6 rnssecgp_4 income (99).
missing values agelch dagelch (-1, 99).
missing values attscale expscale (7,9).
missing values hetlife samlife (-1, 9995 THRU HIGHEST).
missing values nochall (999).
missing values rwadult rwcasual rwsamm rwsamf snnolov snpres snold
snsexdrv snmedia snearly sngayadp snlesadp snsexed uagesex relstat
relstatgp relstatgp2 tenure
livehere employed rempstat educ3 religimp (9).
missing values marstat (99).
missing values whirel (-1, 99).
missing values qimd qwimd qsimd (-1).

```

\*Adjust variable levels.

```

VARIABLE LEVEL ethnicgrp disabil2 parents4 fpsorcg child18 nochb435
drugsyr2 depscr relstat relstatgp relstatgp2
ractivhi3 employed rnssecgp_6 rnssecgp_4 parsc3 gor_1(NOMINAL).
VARIABLE LEVEL alcohol2 smoking educ3 (ORDINAL).

```

\*Save and order desired variables into a new file.

```

SAVE OUTFILE = 'natsal_3_teaching.sav'
/Keep = sin2 dateyoi total_wt psu_scrm strata stratagrp stratagrp2
stratagrp3 dage rdoby rsex ethnicgrp sexid health disabil2 bodymass
drink alcohol2 smoking drcannabis drampheta drcocaine drcrack
drecstasy drnonihero dracidlsd drcrysmeth dramylnit drothnonpre
inject2 drugsyr2 mscore depscr attscale expscale hetlife samlife
marstat relstat relstatgp relstatgp2 relsat nochall dagelch agelch
child18 nochb435 adopchld stepchld fpsorcg fppref2 usfsteril
usmsteril usconpill uscondom usfemidom usemcypill usemcyiud uscoil
usmirena uscap usinjectn usgels usrhythm uswithdrw
usimplant usother usconptch usnotused parents4 sibpos2 siblive
talkmapa sexeducoc sexeduintad sexedusch sexeduinto sexedubmn
sexedufrn sexedutvid
sexedumum sexedudad sexedufsp sexeduintpn sexedupor sexedubsis
sexeduoth sexeduani sexedurel sexeduofri sexeduofrn sexeduexp
sexedufil sexeduedc rwadult rwcasual rwsamm rwsamf snnolov snpres
snold snsexdrv snmedia snearly sngayadp snlesadp snsexed uagesex
attconservative religimp whirel oftrelig ractivhi3 employed rempstat
rnssecgp_6 rnssecgp_4 income parsc3 educ3 netacc adj_imd_quintile
qimd qwimd qsimd tenure livehere gor_1 urindew urindsc.

```