



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

Author: Alle Bloom and Jennifer Buckley

Version: 1.0

Date: October 2020

## Instructor notes

This short guide is designed to help those planning teaching and other learning activities using the Natsal teaching dataset. It discusses the contents of the dataset and ways it could be used for teaching quantitative research methods.

### Research topics in the dataset

The dataset contains variables across a range of research topics. Example questions and topics to explore with the data include:

- The role of family background (including household type, position in the family and ease of talking to parents around sex) on behaviour and attitudes
- What determines age at birth of 1<sup>st</sup> child?
- How do attitudes towards sexual lifestyles vary?
- How do we learn about sex? Does this vary across the generations and depending on family or socioeconomic background?
- Who takes what drugs?
- What are the most popular forms of contraception?
- How does religion determine sexual attitudes and behaviours?

### Interval and continuous variables

Below is a list of interval or continuous variables in the dataset, with some relevant features of these variables noted.

- Age (*dage*)
- BMI (*bodymass*)
- Sexual attitudes: Factor scores with normal distribution (*attconservative*)
- Number of heterosexual partners (*hetlife*) (Right-skew)
- Number of same sex partners (*samlife*)(Right-skew)

- Age of respondent at birth of 1st child - The source variable (*age1ch*) includes top and bottom coded values; however, we have added a recoded version (*dage1ch*) with these values removed. The values then range between 15 and 40.
- Depression Score (*mscore*)

## Ideas for teaching key quantitative methods topics

### Factors relating to age at which individuals have their first child – Linear Regression

A good option for teaching linear regression. A first step could be to look at differences by *sex* (*rsex*) (we would expect men to be older than women) and *education* (*educ3*) (*those with no qualifications are younger at the birth of their first child*). Students can also generate their own theories and consider appropriate model building strategies. Some further suggestions include:

- Deprivation - a model with *qmid* and *dage1ch* shows deprivation is related to having a first child earlier.
- Experience of sex education - difficulty discussing sex with parents is related to having a first child later.
- Religion
- Sexual attitudes

### Factors affecting depression score – Logistic Regression

The binary variable *depscr* – *whether respondent screens positive for current depression* is a good option for teaching logistic regression. Once again, a good first step may be to explore differences in *sex* (*rsex*) and perhaps *age* (*dage*). Further analysis could include:

- Exploring measures of deprivation (*qmid*) and socio-economic inequality (*rnssecgp\_6 nssec variables, netacc*) - as deprivation increases so does the likelihood of screening positive for depression.
- Exploring drug use (*drX variables, drugsyr2*) - those who have used drugs in the last year are more likely to screen positive for depression than those who haven't.

### Religious importance and Sexual attitudes – Chi-Square Test

The religious beliefs variable and the sexual attitudes variables can be used to create interesting cross-tabulations and conduct Chi-square tests. A good starting point might be to cross-tabulate the importance of religious beliefs variable (*religimp*) and the opinion about one-night stands variable (*rwcasual*). Another interesting option is a cross-tabulation of *sex* (*rsex*) and the ease with which respondents discuss sex with their parents (*talkmapa*).

### Sexual attitudes and Drug use – Factor Analysis and Latent Variable Modelling

The sexual attitudes opinion variables (used to create the *attconservative* variable) are a good option for teaching factor analysis. The factor analysis used to create the variable can be replicated and the students asked to name the factor. Conducting a factor analysis to explore the *drX* variables also creates some interesting findings with different factors based on the type of drug – e.g. party drugs vs. hard drugs. The variables on drug and alcohol use can also be used for more complex modelling such as LCA and mediation modelling e.g. exploring pathways into drug use.

## Model building

In addition to analysis techniques, the dataset is well-suited for teaching thoughtful and considered model building. With the variety of the variables in the dataset, students can practice applying theoretical skills to build varied models based on literature.

## Weights and complex sample design

The teaching dataset can be used to teach the analysis of complex samples. 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 main weighting variable is the 'total\_wt' variable. This variable is the overall weight for the total sample (including the boost samples of younger people).

The teaching dataset contains variables for strata and PSU (strata and psu\_scrm). In addition to variable strata, the dataset contains three additional grouped strata variables. These grouped variables can be used when stratas contain too few PSUs to use strata.