



GENERAL LIFESTYLE SURVEY 2009: GUIDELINES FOR USING THE DATA

The General Lifestyle Survey as a module of the Integrated Household Survey

From 2008, the General Household Survey became a module of the Integrated Household Survey (IHS). In recognition, the survey was renamed the General LiFestyle Survey (GLF). Please refer to the IHS web page for further information

<http://www.statistics.gov.uk/statbase/Product.asp?vlnk=15381>

The sample design of the GLF is the same as the GHS and the questionnaire remains largely the same. The main change is that the GLF now includes the IHS core questions, which are asked across the IHS survey modules and cover employment, smoking prevalence, general health, ethnicity, citizenship and national identity. These questions are asked by proxy if an interview is not possible with the selected respondent (that is a member of the household can answer on behalf of other respondents in the household). This is a departure from the GHS which did not ask smoking prevalence and general health questions by proxy, whereas the GLF does from 2008. However, it should be stressed that the published smoking and health National Statistics from the GLF do not use proxy responses. Furthermore, the IHS smoking questions (and proxy responses to this) are only asked of those aged 18 and over. Nonetheless, the GLF continues to ask these questions (via self-completion) of those aged 16 and 17. The proxy responses to the IHS core smoking and health questions are included on the GLF micro dataset.

Changes to the drinking section

In recent years there have been a number of revisions to the methodology that is used to produce the alcohol consumption estimates. In 2006, the average number of units assigned to the different drink types and the assumption around the average size of a wine glass was updated. The revision to the methodology significantly increased the consumption estimates when compared to the original method. For further details please refer to the paper *Estimating alcohol consumption from survey data: updated method of converting volumes to units* National Statistics Methodology Series NSM 37 (ONS 2007), www.statistics.gov.uk/statbase/product.asp?vlnk=15067. The revised method has been used on the 2006, 2007 and 2008 micro datasets.

In addition to the revised method, a new question about wine glass size was included in the survey from 2008. Respondents are now asked whether they have consumed small (125 ml), standard (175 ml) or large (250 ml) glasses of wine. The data from this question are used when calculating the number of units of alcohol consumed by the respondent. It is now assumed that a small glass contains 1.5 units, a standard glass contains 2 units and a large glass contains 3 units. However, in 2006 and 2007 it was assumed that all respondents drank from a standard (175 ml) glass containing 2 units. The updated method has made little difference overall, but has slightly reduced the proportion of women exceeding 3 units on their heaviest drinking day in the week before interview. There are two reasons for this. Firstly, when glass size was analysed by sex and age, for most groups the average size was close to the average assumed under the previous method but for women aged 45-64 average size was lower and for men and women aged 65 and over it was much lower. Secondly, approximately 60 per cent of the units of alcohol consumed by women come from wine whereas only around 25 per cent of men's units do so. This means that any change to the calculation of units of alcohol coming from wine has a much greater effect on the total units for women than on the total for men.

The consumption variables that are affected by the updated method are included on the micro dataset. The published trend tables (*Smoking and drinking among adults, 2009*: <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=5756>) reflect the changes in drinking methodology by including three sections showing the original method up to 2006, then the revised method from 2006 to 2008 and finally the updated method from 2008 which uses the data about wine glass size.



1) GLF questionnaires

The 2009 questionnaires have been provided in Word format. The household questions are recorded first followed by the individual questions. This does not exactly reflect how the questionnaire is presented in the field, as in this case core household then individual questions from the Integrated Household Survey (IHS) are first asked, followed by the GLF specific topic questions. Furthermore, many IHS questions (smoking prevalence and general health) are asked by proxy where necessary, but the Word version of the questionnaire, does not currently reflect this.



2) GLF data files

The data files include households where a response from one or more individuals was missing. On person level files such individuals can be identified by the variable Schedtyp (where Schedtyp= 3). Two data files have been supplied; one at person level and the other at household level.

Person-level file

[glf09client_sl.sav](#) contains person-level data. There is one record for every individual, including children, living in each responding household. This file contains schedule variables (that is, variables which were on the original questionnaire) and the derived variables which have been programmed. All household-level variables have also been attached to all individuals in each household on the person level file.

The file includes two serial numbers which denote households and people. Case (Hserial anonymised) is a household identifier; each household has a unique value at Case, and this value is attached to everyone in the household. Persno is a person identifier; each person within a household has a separate person number.

This file can be used to run tables at person level or household level, but for ease of access we have also supplied a household level file.

Household file

[glf09clienhhld_sl.sav](#) contains household data. There is one record per responding household. The file contains schedule variables and derived variables. Some schedule variables have been deleted or grouped to maintain confidentiality.

Creating family file from the person level file

For families you use FUH=PERSNO (FUH is the person number of the family unit head).

Weight variable: Weight09

The data set is unweighted. Weight09 is the variable you should use to weight the data (See Appendix D Weighting in *General Lifestyle Survey 2009 Appendices.pdf* under <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=5756>). This weight applies to both household and individual level data.

3) Missing values

For each variable, there are three possible missing values. For the majority of variables, these are -6, -8, and -9. No values have been declared missing on the files supplied.

-6 means that the person was either a child (under 16 years of age) or not available for interview and is therefore not included in the relevant derived variable. For example, smoking questions are only asked of adults, so children are coded as -6 for smoking variables. Adults aged 16-59 are asked the family information questions, so children and adults aged 60 and over are coded as -6 for these variables.

-8 (NA) means 'No answer'. This is used where a respondent either does not know the answer or refuses to answer an individual question.

-9 (DNA) means 'Does not apply'. This is used when a household or respondent is eligible to be asked a question, but is routed past it by the flow of the questionnaire. For example, if a respondent is asked whether he or she is eligible to belong to their employer's pension scheme and the answer is 'no', he or she will be coded as -9 on the follow-up question which asked whether he or she belongs to the scheme.

From April 2007 respondents were no longer provided with the option to refuse the whole of the income section, as a consequence from 2008 dataset -7 values are not included, which were previously used to indicate such refusal.

-8 (NA) has the same meaning as for other variables.

-9 combines -6 and -9, and therefore includes people who are not eligible for the income section, and respondents who are routed past questions by the flow of the questionnaire. The exception to this is children and proxy respondents that are coded as -6.

Some words of warning

If you are using SPSS-NT, then it is very easy to save amended versions of SPSS files when you don't want to. When you try to quit SPSS, the dialogue box asks if you want to save the file. The intuitive reaction is to click on 'yes', but you should only do this if you want to save any amendments you have made. If you have used a 'keep' or 'select' statement, clicking 'yes' will save only the cases or variables you have selected. We recommend that you keep the CDROM or original file safe, just in case you inadvertently change the data files.

Both data files are large, and it takes time to run frequencies and tabulations on the full files. We recommend that you use a 'keep' statement to select relevant variables.

Social Survey Division at ONS is currently using version 12.0.1 of SPSS. You may find you have problems with missing values when trying to create derived variables. The solution we have used is to declare no missing values, then create the new variables, then declare missing values.

4) Excel spreadsheet (Variables – changes)

The list of variables on the Excel spreadsheet denotes changes made between the 2008 and the 2009 datasets including:

- Changes to variable names
- Variables (either derived or questionnaire) that have been dropped
- New variables (either derived or questionnaire) that have been introduced

All changes are listed on the Excel spreadsheet.



5) Specifications for derived variables

Also in the zipped file is a specification for derived variables on the data file.

GLF Team
Social Survey Division
Office for National Statistics
February 2011