

A guide to using Scottish and UK datasets

Growing up in Scotland (GUS)

GUS is a longitudinal dataset, looking at the lives of children and their families in Scotland. New information about children and families is added every year. The survey has representative sample of families from every local authority in Scotland.

The interviews at each sweep of Cohort 1 collected information from the child's main carer about a range of issues related to young children and their parents including:

- Characteristics and circumstances of children and their families in Scotland
- Housing, neighbourhood and community
- Food and eating
- Activities
- Child health and development
- health service contact
- Parenting styles and responsibilities
- Childcare and work-life balance
- Parental physical and emotional health
- Child, parent and family social networks
- Parental employment, income and education

The GUS survey contains several data collection methods, including face to face interviews with carers and children, self-completion questionnaires, cognitive assessment, measurements of height and weight done by interviewer, and information from the child's health records. For more on designing and developing a questionnaire for children in GUS, [see here](#).

Survey years, sweeps and initial sample sizes

GUS began in 2005. The survey design consisted of recruiting an initial total of 8,000 parents in 2005, comprising two sub-cohorts of children. Subsequently a new birth cohort was introduced in 2011:

- Child cohort (CC): c. 3000 children born between June 2002 and May 2003 – 4 'sweeps' of data collected from families when child aged just under 3 years to just under 6 years. There are no current plans to collect further data from this group.
- Birth cohort 1 (BC1): c. 5000 children born between June 2004 and May 2005 – data collected annually from families when children aged between 10 months and just under 6

1 See <http://discover.ukdataservice.ac.uk/catalogue/?sn=5760&type=Data%20catalogue> for an account of the edition history of the survey and <http://growingupinScotland.org.uk/about-gus/where-are-we-now/> for an overview of recent and current

activities, including data collection rounds and the publication of research findings.

years then periodically until children in Primary 6. 7 sweeps have been collected, the last sweep is expected to be collected during the 2014/15 academic year.

- Birth cohort 2 (BC2): c. 6000 children born between March 2010 and February 2011 – data collected from families when children aged 10 months, just under 3 years and just under 5 years. Two sweeps have been collected. The third sweep is expected in 2015.

GUS ‘Ages and st ages’

Child’s age	Cohort/Year of data collection (Sweep)		
	Child cohort	Birth cohort 1	Birth cohort 2
10 months		2005/06 (S1)	2010/11 (S1)
Age 2		2006/07 (S2)	
Age 3	2005/06 (S1)	2007/08 (S3)	2013 (S2)
Age 4	2006/07 (S2)	2008/09 (S4)	
Age 5	2007/08 (S3)	2009/10 (S5)	2015 (S3)
Age 6	2008/09 (S4)	2010/11 (S6)	
Age 8		2012/13 (S7)	
Primary 6 (Age 10)		2014/15 (S8)	

There are a number of small but important differences in the design used for Birth Cohort 2 and the earlier cohorts. This is explained in page 5 of the [user guide](#) for BC 2, Sweep 1.

Geographical variables

The regular (non secure access) file had area urban-rural classification, a Scottish index of multiple deprivation summary variable, and a derived variable identifying some of the Scottish Health Board areas. The secure access data contains postcode information.

Useful websites

All documentation, including paper copies of the Questionnaire, is available from the UK Data Service and from the project website. GUS data can be accessed from the UK Data Service.

Project website

The [“Using GUS data”](#) tab has information on the types of data collected and the collection mode, [Data documentation](#) including a topic guide, user guide for each sweep, the questionnaires and variables lists. The last section under the “Using GUS data” tab provides a list of current and completed projects that have used GUS data.

Workshops to promote and support the use of GUS data have been held every 2 years. [This webpage](#) has links to the materials that were provided at previous workshops on survey design, topics and themes, documentation and how to access data. There are practical exercises and tips to use SPSS which are useful to understand how to work with the variables from GUS.

The Publications tab lists reports published by the Scottish Government and journal articles published in peer reviewed journals by researchers using data from GUS.

This [guide](#) made by the GUS project has some information on the data, and a range of clear and useful tools you could use for data manipulations and data management (such as how to reduce the number of variables, how to recode and compute, and how to use weights and stratification variables).

[UK Data Service](#)

The UKDS has separate pages for each wave of GUS, which can be accessed through the search box. The “Documentation” tab contains for the selected cohort and sweep(s) the data documentation (includes carer questionnaire), interviewer and coder instructions and user guide. The “Variable” tab contains a list of all the variables. One can select a variable together with a category, upon which the unweighted frequency and percentage is provided.

[The Scottish Government website](#)

Has news releases of results from GUS, research reports and other associated documents. The best way to find them is through the search bar on the top right corner of the webpage. Examples include:

- <http://www.scotland.gov.uk/News/Releases/2012/05/growingup10052012>
- <http://www.scotland.gov.uk/Publications/2013/02/3280>. Growing Up in Scotland: Birth Cohort 2. Results from the First Year. This GUS report provides a detailed insight into the first set of data collected from the study’s second birth cohort around the time they were all aged 10 months old.
- <http://www.scotland.gov.uk/Publications/2010/04/21131609/0> – Growing Up in Scotland - 2010 - Topic Research Findings: Research Findings No.1/2010: Growing Up in Scotland: The Circumstances of Persistently Poor Children. Summary Report. This is one of a series that summarises key findings from the fourth sweep of GUS.

Missing values

All missing or similar categories were coded with a negative integer:

- Questions that did not apply to the child or carer were coded as -1 (“not applicable”, for instance, no partner in household).
- Missing values were also coded as -1 and system missing values -1 or -3.

- The code for 'No information' provided or 'not available' is -3.
- 'Don't know' is coded as -8.
- When 'Refused' to answer, the code -9 was given.

Weights

In GUS there are two weights for each cohort on all datasets plus a separate weight for analysis of the partner interview data.

Selection of the correct weight is dependent on the data you are using and analysis being undertaken:

- Cross-sectional weight
 - Use for any cross-sectional analysis of single sweep data only. For Sweeps 3 to 6 there are two components to this weight:
 - Individuals who had responded at all previous sweeps.
 - Individuals who had responded at Sweep 1 but had missed one or more interviews in sweeps 2-5.
- These two sample populations are to be treated separately during the weighting. This is because the second group respondents are likely to have different response behaviour to respondents in the first group (this is suggested by their much lower response rates).
- Longitudinal weight
 - Use for analysis of more than one sweep of data
 - Weight used should be from the latest sweep (i.e. if analysing sweep 3 and sweep 5 data, use sweep 5 longitudinal weight).
 - The final sweep weight is the product of the sweep non-response weight and the sweep interview weight. For each cohort the final weights were scaled to the responding sweep sample size, this makes the weighted sample size match the unweighted sample size.
- Sweep 2 Partner interview weight
 - Use for any analysis of Partner interview data
- See [User Guides](#) for more details.