

Samples of Anonymised Records (SARs)

Quick guide to SARs

1. What are the SARs? The SARs are a family of datasets drawn from the 1991 and 2001 UK Census. The SARs contain a separate record for each individual, but identifying information has been removed to protect confidentiality. The SARs datasets are similar to data from a survey, albeit with a much larger sample size thus permitting analysis of small sub-groups and small geographic levels. The SARs cover the full range of Census topics including, housing, education, health, transport, employment, ethnicity and religion.

2. Why use the SARs? The SARs allow flexible, multivariate analysis at the individual level. They are a unique data source for the investigation of a range of social issues including household composition, ethnic differences, education and employment. The SARs differ from traditional Census outputs in that they are not aggregated into pre-determined tables.

3. What datasets are available? The SARs family consists of the following datasets:

- The Individual SAR is a sample of individual census returns.

The 1991 Individual SAR (**I-SAR**) is a 2 per cent sample of over 1.1 million records. It contains a full range of census topics on individuals and summary information about households. In total 278 geographical areas are identified and include local authorities with a population of at least 120,000. Smaller local authorities have been grouped together.

The 2001 Licensed Individual SAR (**LI-SAR**) is a 3 per cent sample and contains over 1.75 million records. It contains a full range of census topics on individuals and summary information about households and the new information collected on qualifications, caring and religion. Geographical information is given down to Government Office Region.

- The Household SAR is a hierarchical sample of households and people within those households.

The 1991 Household SAR (**H-SAR**) is a 1 per cent sample of the population and contains 216,000 households and 500,000 people within these households. It is a hierarchical file that allows linkage between household and family members. The geographical base is the Registrar General's Standard Regions in 1991. The Household SAR contains about 40 variables though many derived variables have also been added.

The 2001: Special Licensed Household Sample of Anonymised Records (**SL-HSAR**) is a 1 per cent sample of the population and contains over 200,000 households and 500,000 people within households. Again it allows linkage between household and family members. It is limited to England and Wales and contains no geography.

- The 2001 Small Area Microdata (**SAM**)

The 2001 SAM is a sample of 5 per cent of the population of the UK, representing about 3 million individuals. The lowest geography is Local Authority (Unitary Authority in Scotland and Parliamentary Constituency in Northern Ireland), with some areas merged. The median sample size for an authority is 5.6 thousand records and nearly [eighty authorities have more than 10,000 records](#). The amount of individual detail in the SAM is less than in the Individual SAR because of the greater geographical detail in the SAM.

- The 2001 Controlled Access Microdata Samples (**CAMS**)

These are more detailed versions of the 2001 LI-SAR and SL-HSAR available for analysis in a safe setting within ONS. Further information can be found on the ONS web site at www.statistics.gov.uk/census2001/sar_cams.asp. Both files provide geography at Local Authority level and hold data for the entire UK. Codebooks for each file can be downloaded from the SARs web site.

4. Exploring the data online. Simple exploration is possible with [Nesstar](#).

Nesstar is a data exploration and analysis tool that allows you to look at the data without needing to register for it. This software makes it easy to specify and make your own tables or graphs from the SARs data. This is an excellent way of finding out whether the SARs are likely to be suitable for your purposes. There is a user guide in the top right hand corner of the Nesstar screen. After registration Nesstar can be used to its full potential, undertaking more complex analyses and downloading subsets of data in a wide variety of formats.

5. Registration and access. Once you have decided to use the SARs for your research you will need to register to obtain the data. There are three types of licenses corresponding to different levels of security.

- The [End User Licence agreement](#) (EUL) provides access to the following files:
 - 1991 Individual and Household files for Great Britain and Northern Ireland
 - 2001 Individual Licensed SAR
 - 2001 Small Area Microdata

Academic Users can register through Census.ac.uk and they receive the data free of charge from the SARs datasets page as a SPSS, STATA or tab de-limited file.

Non-Academic Users should follow the instructions on the web site at www.ccsr.ac.uk/sars/access/nonacademic/. CCSR will supply the 2001 SARs to public sector bodies free of charge and to the business sector for £1000 per file. Once registered, the data will be made available on CD with NSDSTAT.

Overseas users: can obtain information about use and access of EUL SARs on the [SARs web site](#).

- The **2001 Special Licence Household Sample of Anonymised Records** (SL-HSAR) requires a separate application procedure, which is more secure than the End User license agreement. Applicants need to complete an ONS Special Licence via the [UK Data Archive](#).
- Both 2001 [Controlled Access Microdata Samples \(CAMS\)](#) are available only through the Office for National Statistics (ONS) in a safe setting and require a much higher level of data stewardship. More information can be found on the [ONS web site](#).

6. Data Analysis. You can undertake straightforward analysis, including linear regression, on the whole data set or a subset, using [Nesstar](#). However, it is not possible to produce derived variables or do more advanced analyses using Nesstar. To do this you will need to use a package like SPSS or STATA, for which guides are available via the SARs [web site](#). CCSR offers one-day courses details of which can be obtained through the [CCSR Short Course Pages](#) as well as training [documentation](#). It will not be possible to analyse the SARs without reference to the documentation. You can access the User Guide and codebooks on the web before you register to use the data.

NSDstat is more user friendly statistics software and academics can download it for free of charge. For non-academics, purchase of the SARs includes a copy and licence for NSDstat for up to ten users. More information can be found at www.ccsr.ac.uk/sars/access/modes.html.

7. A teaching resource. Learning and teaching materials based on the SARs are freely available on the SARs web site. The materials include both methods (from tables to logistic regression) and substantive topics (for sociology, demography, social policy and other disciplines).

8. SARs support. CCSR provides a range of [support](#) for those who use or are interested in using the SARs. A helpdesk is staffed from 9.00 - 5.00, Monday to Friday and can be contacted at:

E: sars-helpdesk@manchester.ac.uk

T: 0161 275 4735

There is a regular [SARs newsletter](#) and a [SARs email list](#) for updates and discussion.

9. Publishing. We maintain a record of all publications which use the SARs. For a list of publications from the SARs or to add your publication to the list go to www.ccsr.ac.uk/sars/publications/.