



Secure Microdata Files – User Guide

1. Introduction

Census Microdata

The 2011 Census Microdata products provide an alternative source of data to the standard 2011 Census releases which consist of tables containing counts of how many people in one area have a certain attribute or attributes.

In contrast, microdata contain information from the individual respondents and households, this means that census microdata contain a rich source of data for researchers to use in their own analyses.

These data have been treated to protect the confidentiality of the respondents. Rather than data being delivered as tables of counts, data are stored in the same way as if the data had been collected by conducting an anonymous sample survey, with a sample of cases being made available and individual records containing information on a limited number of topics.

Secure Microdata Files

A variety of microdata products are available, balancing the level of detail in the data, the sample size and how to access the data.

The Secure Microdata Files represent the products with the highest level of detail and the largest sample size. As a result, they are also protected by the highest level of access limitation and are only available to Approved Researchers via the Virtual Microdata Laboratory (VML). Details of other microdata products can be found via the [ONS Census Microdata webpage](#).

The Secure Microdata Files consist of random samples of 10 per cent of people or households in the 2011 Census output database for England and Wales. These include people classed as both usual residents and short-term residents. There are two Secure Microdata Files:

The Individual file, which includes person-level data on 257 variables for over 5 million individuals. The lowest level of geography available is local authority. Records include people resident in households and communal establishments.

The Household file, which includes person and household level data on 245 variables for over 5 million individuals within more than 2.4 million households. This is a sample of 10 per

cent of all households and includes individual person-level records for each person living within those households. Where households contained no enumerated persons, records in the sample include only household-level data. The Household file allows linkage between individuals in the same family and the same household. Local authority is the lowest level of geography available.

The household and individual files do not overlap and cannot be combined.

2. About the Secure Microdata Files

2.1 2011 Census

Census data are collected every 10 years in England and Wales, as governed by the 1920 Census Act. Census day for the 2011 Census was the 27th of March 2011. Full information on how the 2011 Census in England and Wales was conducted is available on the 2011 Census page on the ONS website <http://www.ons.gov.uk/ons/guide-method/census/2011/index.html>.

The 2011 Census is the most complete available source of information on the population. However, despite efforts to reach everyone and obtain the most accurate information possible, no census is perfect and some people are inevitably missed. Further information on how the 2011 Census was conducted, and the treatment of missing data, is available in the [Quality and Methods](#) section of the [2011 Census User Guide](#).

2.2 UK Censuses and Microdata Products

The [Office for National Statistics](#) (ONS) is responsible for carrying out the census in England and Wales. Simultaneous but separate censuses took place in Scotland and Northern Ireland. These were run by the [National Records of Scotland](#) (NRS) and the [Northern Ireland Statistics & Research Agency](#) (NISRA) respectively.

The 2011 Census Microdata Secure Files released by ONS cover the countries of England and Wales only. Similar products for Northern Ireland are available from [NISRA](#). NRS have plans to produce similar products for Scotland.

2.3 Population Base

The Secure Files include data from the total England and Wales population, which includes usual residents, short-term residents and students living away from home during term-time (categorised in the 'population base' variable).

The main population base for published statistical tables from the 2011 Census is the usual resident population as at census day, 27 March 2011. For 2011 Census purposes, a usual resident of the UK is anyone who, on census day, was in the UK and had stayed or intended to stay in the UK for a period of 12 months or more, or had a permanent UK address and was outside the UK and intended to be outside the UK for less than 12 months.

If the data from the Secure Files are not filtered to include only usual residents, results may differ from published statistical tables. Many of the variables included in the dataset, however, contain values only for usual residents.

Students and schoolchildren in full-time education studying away from the family home were counted as usually resident at their term-time address. Basic demographic information only (name, sex, age, marital status and relationship) was collected at their non-term time ('home' or vacation) address. In the Secure Files, information collected at their non-term time address is available by filtering the Population Base variable to include only 'students living away from home during term-time'. Datasets including both 'students living away from home during term-time' and 'usual residents' could thus include duplicate information on basic demographic variables for some individuals. Therefore careful consideration should be given when choosing the appropriate population base for data exploration, and when comparing findings from the Secure Files with published reports.

Residents in communal establishments are included in the Individual Secure Microdata File but not the Household Secure Microdata File.

2.4 Geography

The lowest level of geography for the Individual and Household Secure Files is local authority.

2.5 Variables

Separate code books are available for the secure individual file and the secure household file.

Further details on how these variables were derived from the 2011 Census questionnaire, including definitions and derivations, can be found in the [2011 Census variable and classification information guides](#).

These data will be available within the VML in each of the following formats: CSV, SPSS (.sav), STATA (.dta) or SAS (.sas7bdat).

2.6 Continuity between 2001 and 2011

There are some differences between the 2001 and 2011 censuses, which may have implications for users. Full information on differences between the census questions and definitions can be found in the [comparability over time](#) section of the [2011 Census User Guide](#).

Information on the 2001 Census Samples of Anonymised Records (SARS), the equivalent products to the 2011 Census Microdata, are available from the [UK Data Service website](#).

3. How were the samples drawn?

3.1 Sample size

Sample sizes have been chosen to be consistent with statistical disclosure control considerations and user requirements. The number of records that are within the sample and unique within the census database was measured as a proportion of the number of records that are unique within the sample. The level for this proportion was set in order to determine sufficient uncertainty.

The sample composition chosen will remain confidential, as was the case for the 2001 Census microdata.

3.2 Stratification

Stratification enables the characteristics of a sample to be proportionally representative of the population by dividing the population into strata based on key characteristics. Random samples taken from each stratum are then pooled to form the final sample.

The Secure File samples are stratified by census [output area](#) within Local Authority. This method ensures good representation of data, and that the sample is more evenly spread and is consistent with the user requirement for a multipurpose product that can be used for a wide variety of analyses. It also controls against extreme sample selection, ensuring for instance that an entire output area is not selected at random.

4. Confidentiality & Protection of Personal Data

Preserving the confidentiality of personal information provided by the public on their census questionnaires remains a top priority for the census. (See more detail on [data confidentiality in the 2011 Census outputs](#)).

[Statistical disclosure control](#) applied to 2011 Census data, including targeted record swapping and restriction of detail, additionally protects the attributes of individuals and ensures that there is always uncertainty surrounding any apparent attribute disclosure.

The Microdata Secure Files contain samples of 10 per cent of people and no personal identifiers (name, address, date of birth) have been included in any 2011 Census Microdata products. However, due to the risk of disclosure, these products are only available to individuals who have been granted [Approved Researcher](#) status. Additionally, the data will only be available to access via the [Virtual Microdata Laboratory](#), a highly secure environment from which no data can be exported without specific approval.

5. How to gain access to the Secure Microdata Files

Access to the Secure Individual Microdata File and the Secure Household Microdata File will be available via the Virtual Microdata Laboratory (VML). To request an application form for accessing a secure microdata sample, contact VML.Service.Desk@ons.gsi.gov.uk

The VML is currently being renovated to provide a much improved service and will be in place from 2015. Therefore, in the meantime, access to the Secure Microdata Files will be available via an interim VML, a temporary version of the VML housed at the Office for National Statistics in Titchfield. Access will only be available from this site. Once the new VML is in place, the range of locations from which the data can be accessed will be expanded.

Access to the Secure Microdata Files within the VML will only be granted to individuals with 'Approved Researcher' status. This status will only be granted if the researcher is able to demonstrate that he/she:

- has the appropriate knowledge and experience necessary for handling potentially disclosive personal information;

- has provided satisfactory evidence supporting their application that illustrates their professionalism and technical competence to carry out the research proposal; and
- demonstrates a commitment to protecting and maintaining the confidentiality of the data during the creation of outputs and publications that arise during the proposal.

Additionally, the research must be able to demonstrate that their research serves one of the following public benefits:

- Supporting the formulation and development of public policy or public service delivery.
- Forms part of the programme of research covered by the National Data Strategy or otherwise supported directly or indirectly by the Economic and Social Research Council.
- Supports an obligation of public law (e.g. Local Development Plans).
- Explores new statistical methods that can be used to produce statistics that serve the public good.

Further information and the relevant application forms are available from the [approved researcher accreditation](#) page on the ONS website.

6. Other Census products and getting more information

Further information from the 2011 Census on output products already published or planned for publication in the future can be found on the [2011 Census data web pages](#).

For further information about census results please contact [Census Customer Services](#).

[Census ad hoc tables](#) contain combinations of data that are not available in standard publications. They have been requested by the media or other user. Once created, ad hoc tables are published for all to use. Ad hoc tables will be constrained by the data that are available at each stage, by the similarity to what ONS plans to release, and by statistical disclosure control. Tables can be pre-ordered from [Census Customer Services](#).

Census Customer Services

Email: census.customerservices@ons.gsi.gov.uk

tel: +44 (0)1329 444972

fax: +44 (0)1633 652981

minicom: +44 (0)1329 444517

Census Customer Services
ONS
Titchfield
Fareham
Hants
PO15 5RR