

Scotland's Census:

2011 census Secure Microdata – user guide

**Version 0.1
15 February 2016**

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 Virtual Microdata Laboratory (VML). Details of other Microdata products can be found on the [Scotland Census](#) website.

The secure Microdata files consist of random samples of 10 per cent of people or households in the 2011 Census output database for Scotland, These include people classed as both usual residents and schoolchildren and full-time students living away from home during term time. There are two secure Microdata files:

The Individual file, which includes person-level data on 145 variables for over half a million individuals. The lowest level of geography is Council Area. Records include people resident in households and communal establishments.

The Household file, which includes person and household level data on 123 variables for over half a million individuals within more than 200,000 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. Council Area 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 Scotland, as governed by the 1920 Census Act. Census data for the 2011 Census was the 27th of March 2011. Full information on how the 2011 Census in Scotland was conducted is available on the [Scotland Census website](#).

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 census methodology section of the [Scotland Census website](#).

2.2 UK Censuses and Microdata products

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

The 2011 Census Microdata Secure samples released by NRS cover Scotland only. Similar products for [England & Wales](#) and [Northern Ireland](#) are available.

2.3 Population Base

The Secure Files include data from the total Scottish population, which includes Schoolchildren and full-time students living away from home during term time (categorised in the 'term-time address indicator' variable).

The main population base for published statistical tables from the 2011 Census is the usual resident population as at census data, 27th March 2011. For 2011 Census purposes, a usual resident of the UK is anyone who, on census data, 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.

Schoolchildren and full-time students 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 address. In the Secure Files, information collected their non-term time address is available by filtering the term time address indicator variable to include only 'At another address'. 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 no in the Household Secure Microdata File.

2.4 Geography

The lowest level of geography for the Individual and Household Secure Files is Council Area.

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 other associated metadata, can be found in the [Scotland Census variables and classifications](#) page.

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

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 are unique within the census database were 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 enable 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 Council Area. This method ensure 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, the 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 (VML), 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

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 the 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 Scotland Census](#) website.

For further information on census results, please contact [Census Customer services](#).

Census Customer Services

Email: customer@gro-scotland.gsi.gov.uk

Telephone: 0131 314 4299

Census Customer Services
National Records of Scotland,
Room 1/2/9
Ladywell House
Ladywell Road
Edinburgh
EH12 7TF