



# Scotland's Census 2011

## Microdata Teaching File

### User Guide

June 2014



## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. Confidentiality and protection of personal data</b>	<b>3</b>
<b>3. 2011 Census Microdata Products</b>	<b>3</b>
<b>4. Purpose of the 2011 Census Microdata Teaching File</b>	<b>4</b>
<b>5. Sampling</b>	<b>4</b>
<b>6. How the Teaching File should be cited</b>	<b>5</b>
<b>7. Appendix 1: Contents of the Teaching File</b>	<b>6</b>

### 1. Introduction

1.1 [Scotland's census](#) is the largest and most complex statistical exercise undertaken in Scotland. The census has collected information about the population every 10 years since 1801 (except in 1941 when no census was taken due to the Second World War). For the 2011 Census, National Records of Scotland (NRS) asked people to fill in a census questionnaire to reflect their circumstances on 27 March 2011. The census [questionnaire](#) included 13 questions about each household (eg type of accommodation, number of rooms etc) and up to 35 questions for each person in the household (eg age, sex, marital status, ethnic group, health, qualifications).

1.2 Key users of census information include both central and local government, academia, organisations undertaking research, the private, business and voluntary sectors and the general public. The census is widely acknowledged as playing a fundamental and unique role in the provision of comprehensive and robust population statistics. Census information is needed to help government develop policies and initiatives, and for local authorities to plan services and to make effective use of resources that benefit the people of Scotland. This information underpins the allocation of billions of pounds of public money each year to provide services like education, transport and health. Decisions are taken every day using census statistics. For example, the numbers of school spaces, houses, care homes, hospitals and fire services are all influenced by the census.

1.3 Census statistics are used to help understand the increasingly diverse nature of Scotland's population by capturing the similarities and differences in the populations' characteristics locally and nationally. The census provides a rich source of data to learn about our population and are used for research and teaching purposes.

1.4 NRS is publishing a number of [standard output](#) tables containing summary statistics that describe the characteristics of Scotland's population. However, these tables do not always provide the flexibility needed for specific research questions to be answered. In order to make census data more accessible for research purposes, NRS is producing a number of microdata products based on a sample of anonymised census data at an individual person and household level.

## 2. Confidentiality and protection of personal data

2.1 By law we keep your personal census details confidential for 100 years. Until then, we will not share or publish information that would reveal personal information about you or the place you live. We ensure that measures are taken so that the microdata products remain anonymised and confidential in line with other census products.

2.2 Statistical disclosure techniques are employed to ensure that the risk of inadvertent disclosure in statistical outputs is minimised. The [ONS website](#) includes a description of the principles of the Statistical Disclosure Control (SDC) methodology applied by all the UK census offices for the 2011 Census.

2.3 In addition to SDC the individual level samples included in all census microdata products will have all personal information (name, address, date of birth) removed to help ensure confidentiality is maintained. If required, variables will either be removed or their output categories collapsed to high levels to maintain confidentiality.

## 3. 2011 Census Microdata Products<sup>1</sup>

3.1 The first microdata product to be published is a person level Teaching File. This comprises a sample of 63,388 records (approximately 1 per cent) relating to people living (usually resident) in Scotland at the time of the 2011 Census. For each individual, information is available for 15 separate characteristics (e.g. age, sex marital status) to varying degrees of detail. Further information on the categories available in the Teaching File for each characteristic is provided in Appendix 1, and additional background on definitions is available on our [website](#).

3.2 The Teaching File has been published as a comma separated format text file (.csv) on the [NRS Census 2011 website](#) and is available to everyone subject to Crown Copyright protection (see section 6).

3.3 Equivalent Teaching Files for [England and Wales](#) and for [Northern Ireland](#) have been published by the Office for National Statistics (ONS) and the Northern Ireland Statistics and Research Agency (NISRA) respectively.

3.4 In addition to the Teaching File, NRS are considering two further 2011 Census microdata products that will only be available under controlled access arrangements, as detailed below:

- Secure products: An anonymised 10 per cent sample of individuals and an anonymised 10 per cent sample of households. These samples will not overlap and will provide the most detail for the characteristics included. They will only be available in a secure setting equivalent to the Virtual Microdata Laboratory (VML) managed by ONS. In order to gain access to the data an application would need to be submitted and approved by the data providers and signed contracts and licence agreements are required.

---

<sup>1</sup> Sometimes referred to as “Sample of Anonymised Records”, or “SARs”.

- Safeguarded products: An anonymised 5 per cent sample of individuals will be available to approved researchers via specific end-user license arrangements. The sample will be drawn from the 10 per cent secure individual sample mentioned above and will therefore overlap with it. However, there will be a reduced level of detail available in the safeguarded product, reflecting the associated access arrangements and the paramount need to preserve the confidentiality of personal census information on individuals.

3.5 Further proposals on the content and access arrangements for the 2011 Census secure and safeguarded microdata products for Scotland will be provided later in 2014.

#### **4. Purpose of the 2011 Census Microdata Teaching File**

4.1 The primary purpose of the Teaching File is as an educational tool aimed at:

- assisting with the teaching of statistics and geography in schools and education establishments;
- providing an introduction to the detail, available in microdata products in order to give users the skills and information necessary to make use of the more detailed products to come;
- encouraging wider use of census data by providing a way of examining census data beyond the standard tables.

4.2 Due to the relatively small sample size and limited detail available for each of the characteristics included, the Teaching File may not always be the most appropriate data set for analysis. For example:

- Using a [published table](#) that includes the entire population is always more appropriate when considering the population of Scotland as a whole.
- It may be more appropriate to use more detailed classifications available from a standard census table or (once available), from the secure/safeguarded census microdata products. For example, if the research question is looking at differences between those aged 18 and over and those aged under 18 then it would be better to use data that has single year of age rather than 16-24 age grouping available in the Teaching File.

4.3 If detailed multivariate analysis is required, then the small sample size of the Teaching File may result in reduced reliability for any conclusions drawn. If detailed multivariate analysis is required then consideration should be given to further collapsing categories or using the secure/safeguarded microdata products with their larger sample sizes.

#### **5. Sampling**

5.1 A random 1 per cent sample of individual person records was selected from the complete 2011 Census database. The distribution of each variable was compared to the distribution of the complete 2011 Census to make sure that the Teaching File is representative of the population as a whole. As this is a sample, it is subject to

sampling error, which arises whenever variable estimates are based on a sample rather than a full census of the population. Care should be taken when generalising findings from any sample to a wider population, with due consideration to the reliability of results.

## **6. How the Teaching File should be cited**

6.1 All material on the NRS website is covered by Crown Copyright protection unless otherwise stated.

6.2 Information in the Teaching File (not including logos) can be reproduced or adapted freely without a specific application for a licence under the terms of the [Open Government Licence](#) as long as the source is acknowledged. If you reproduce the information then you have to include the following source accreditation: National Records of Scotland licensed under the Open Government Licence. If you adapt the information you have to include the following source accreditation: Adapted from data from National Records of Scotland licensed under the Open Government Licence.

## 7. Appendix 1: Contents of the Teaching File

Variable Name	Number of Categories	Variable Values
Record_Number	Unique reference ID	Unique reference ID
Region	1	S92000003
Residence_Type	2	C. Resident in a Communal Establishment P. Not resident in a Communal Establishment
Family_Composition	6 (also X)	0. Not in a family 1. Married/same-sex civil partnership couple family 2. Cohabiting couple family 3. Lone parent family (male head) 4. Lone parent family (female lead) 5. Other related family X. No code required (residents of a communal establishment)
Sex	2	1. Male 2. Female
Age	8	1. 0 to 15 2. 16 to 24 3. 25 to 34 4. 35 to 44 5. 45 to 54 6. 55 to 64 7. 65 to 74 8. 75 and over
Marital_Status	5	1. Single (Never married or never registered a same-sex civil partnership) 2. Married or in a same sex-civil partnership 3. Separated, but still legally married or still legally in a same-sex civil partnership 4. Divorced or formerly in a same-sex civil partnership which is now legally dissolved 5. Widowed or surviving partner from a same-sex civil partnership
Student (Schoolchild or full-time student)	2	1. Yes 2. No
Country_Of_Birth	2	1. UK 2. Non UK
Health (General health)	5	1. Very good health 2. Good health 3. Fair health

		<ul style="list-style-type: none"> <li>4. Bad health</li> <li>5. Very bad health</li> </ul>
Ethnic_Group	5	<ul style="list-style-type: none"> <li>1. White</li> <li>2. Mixed or multiple ethnic group</li> <li>3. Asian</li> <li>4. African</li> <li>5. Caribbean or black</li> <li>6. Other ethnic group</li> </ul>
Religion	9	<ul style="list-style-type: none"> <li>1. No religion</li> <li>2. Christian</li> <li>3. Buddhist</li> <li>4. Hindu</li> <li>5. Jewish</li> <li>6. Muslim</li> <li>7. Sikh</li> <li>8. Other religion</li> <li>9. Not stated</li> </ul>
Economic_Activity	9 (also X)	<ul style="list-style-type: none"> <li>1. Economically active: Employed</li> <li>2. Economically active: Self-Employed</li> <li>3. Economically active: Unemployed</li> <li>4. Economically active: Full-time student</li> <li>5. Economically inactive: Retired</li> <li>6. Economically inactive: Student</li> <li>7. Economically inactive: Looking after home or family</li> <li>8. Economically inactive: Long-term sick or disabled</li> <li>9. Economically inactive: Other</li> <li>X. No code required (Aged under 16)</li> </ul>
Occupation	9 (also X)	<ul style="list-style-type: none"> <li>1. Managers, Directors and Senior Officials</li> <li>2. Professional Occupations</li> <li>3. Associate Professional and Technical Occupations</li> <li>4. Administrative and Secretarial Occupations</li> <li>5. Skilled Trades Occupations</li> <li>6. Caring, Leisure and Other Service Occupations</li> <li>7. Sales and Customer Service Occupations</li> <li>8. Process, Plant and Machine Operatives</li> <li>9. Elementary Occupations</li> <li>X. No code required (People aged under 16 and people who have never worked)</li> </ul>
Industry	13 (also X)	<ul style="list-style-type: none"> <li>1. Agriculture, forestry and fishing</li> <li>2. Mining and quarrying; Manufacturing;</li> </ul>

		<p>Electricity, gas, steam and air conditioning system; Water supply</p> <p>3. Construction</p> <p>4. Wholesale and retail trade; Repair of motor vehicles and motorcycles</p> <p>5. Accommodation and food service activities</p> <p>6. Transport and storage; Information and communication</p> <p>7. Financial and insurance activities</p> <p>8. Real estate activities; Professional scientific and technical activities; Administrative and support service activities</p> <p>9. Public administration and defence</p> <p>10. Education</p> <p>11. Human health and social work activities</p> <p>12. Arts; entertainment and recreation</p> <p>13. Other</p> <p>X. No code required (People aged under 16 and people who have never worked)</p>
Hours_Worked_Per_Week	4 (also X)	<p>1. Part-time: 15 or less hours worked</p> <p>2. Part-time: 16 to 30 hours worked</p> <p>3. Full-time: 31 to 48 hours worked</p> <p>4. Full-time 49 or more hours worked</p> <p>X. No code required (People aged under 16 and people not working)</p>
Approximate_Social_Grade	4 (also X)	<p>1. A,B</p> <p>2. C1</p> <p>3. C2</p> <p>4. D,E</p> <p>X. No code required ( People aged under 16 and people resident in communal establishments)</p>