CENSUS ADVISORY GROUP

AG (14) 01 - 2011 Census Highlight Report (Summer 2014)

This highlight report describes progress for the following topics, not covered on the agenda items:

1. Recent and upcoming publications

2. UKSA assessment

3. NRS update

4. NISRA update

5. ONS Geography update

6. Eurostat

7. VML

8. Successful projects in government award

1.   Recent and upcoming publications

1.1 General Report

As is traditional after a census, ONS will publish a General Report.   The 2011 Census General Report for England and Wales will be published in the summer of 2014 and will review the entire 2011 Census operation from the early consultation and planning stages through to the production and dissemination of outputs, evaluation, and beyond. It will provide a wealth of detail about how the census was carried out and what we have learned to take forward in our plans for any future census. The report is aimed at both the experienced and occasional user of census data, but it is hoped the wider public may also find the report useful and informative.

The previous General Report, from 2001, is available from the link: The Census 2001 General report for England and Wales

1.2 2011 Census data catalogue

The 2011 Census data catalogue was published on 13 May 2014. The old release-numbered navigation was replaced with intuitive channels to the data.

Three ways to search:
- by table type,
- by topic, and
- by release date.

Each table type has its own page:
- Population and household estimates
- Key and Quick statistics
- Detailed Characteristics
- Local Characteristics
- Alternative population statistics
- Statistics for the UK
- (Small population groups will follow when the armed forces tables are released).

Each page lists all appropriate releases; each release includes:
- links to tables,
- links to layouts,
- link to release page, and any other pertinent content, such as statistical bulletin, interactive maps etc.,
- summary of contents,
- publication date and original release number.

1.3 2011 Area Classification for Output Areas

The release of a new *2011 Area Classification for Output Areas (2011 OAC)*, produced by University College London on behalf of ONS, is scheduled for release on Friday 18 July. This will include for all UK Output Areas, code and name information for the three hierarchies of the 2011 OAC - Supergroups, Groups and Subgroups, and supporting material including maps, a methodology note and pen portraits.

If you have any questions relating to the 2011 Area Classification, please contact ONS using the mailbox: 2011OAC@ons.gov.uk

1.4 Microdata Teaching File

ONS published the Microdata Teaching File on 23 January 2014 as a csv file available to download from the ONS website. This anonymised, random, one per cent sample of census records provides a useful resource for users of census outputs to analyse census data in a way that is not possible using standard census outputs.

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

- encouraging wider use of census data by providing a way of examining census data beyond the standard tables;
- providing an introduction to the detail, metadata and data formats included in microdata products in order to give users the skills and information necessary to make use of the more detailed products to come; and,
- assisting with the teaching of statistics and geography at GCSE and higher levels

Preserving the confidentiality of personal information provided by the public on their census questionnaire remains a priority for the census. The Microdata Teaching File is a 1 per cent sample of people and a small number of their characteristics. Additional measures have been taken to ensure that individuals cannot be identified:

- No personal identifiers (name, address, date of birth) have been included in the Microdata Teaching File
- Potentially disclosive output variables have been either completely removed or have been aggregated to reduce the level of detail available for each record. In particular, geographic information is limited to region (e.g. North West or London).

1.5 Key and Quick Statistics for Output Areas in the UK

On 13 June 2014 ONS published 26 Key Statistics and 47 Quick Statistics tables at Output Area level for the United Kingdom. These combine information previously published separately by the census offices responsible for the conduct of the census in each constituent country of the UK. The tables are provided for all output areas in the UK, in a simple csv format suitable for additional processing by users.

1.6 2011 Census Quality Survey (CQS) report

The 2011 Census Quality Survey report was published on 14 February 2014. The report describes the CQS design and presents its findings in terms of agreement rates between the 2011 Census and CQS for each question asked in the survey. It suggests possible reasons for differences based on analysis of the differences.

While there is no single overall measure of the accuracy of a census, there are several indicators of its quality. One of these is the level of respondent error, and the principal aim of the Census Quality Survey (CQS) was to estimate this**.** The CQS was a voluntary survey carried out across England and Wales after the 2011 Census. It aimed to measure the accuracy of answers given to census questions by asking a sample of households the census questions again in a face-to-face interview. By comparing the responses given in the CQS to those given in the census, agreement rates were calculated which provide an indication of how accurately the 2011 Census questionnaire had been completed by the general public.
.

2.   UKSA assessment

The UK Statistics Authority (UKSA) has initiated their special assessment of the 2011 Census (phase3). The UKSA assessment team is gathering evidence to inform their assessment and will be contacting some users and stakeholders to gain feedback on the 2011 Census. The assessment covers the three censuses in the UK and an assessment report is expected late in 2014 or early 2015.

This is an important part of the process that ensures that we at ONS are in compliance with the Code of Practice for Official Statistics whilst carrying out the census. More information on the previous assessments can be found at: http://www.ons.gov.uk/ons/guide-method/census/2011/how-our-census-works/how-we-planned-the-2011-census/independent-assessments/special-assessment-of-the-2011-censuses-in-the-uk/index.html

3.   NRS update

2011 Census statistics for Scotland

Release 2

The three outstanding Quick Statistics tables (on household deprivation, central heating, and long-term health conditions) were published on 9 April 2014 as Release 2D.

Release 3

Release 3 of the 2011 Census results for Scotland comprises publication of Local Characteristics (LC) and Detailed Characteristics (DC) tables for Scotland. These are being published by broad topic throughout 2014 in the following order:

- Ethnicity, Identity, Language and Religion
- Population and Households
- Health
- Labour Market & Education
- Housing and Accommodation
- Transport

The first batch of DC and LC tables was published (as Release 3A) on 27 February 2014. Four further batches of tables (Releases 3B, 3C, 3D and 3E) have been published since then, with the next batch (Release 3F) due for publication on 25 June.

Census Data Explorer

The key dissemination route for the statistical outputs from Scotland's 2011 Census is the Census Data Explorer - http://www.scotlandscensus.gov.uk/ods-web/home.html. This includes all the standard output tables that have been published to date, area profiles of census snapshot data for areas selected by users, a range of interactive maps and charts, a data warehouse (primarily for users of bulk data) and metadata and other supporting information.

Microdata teaching File

A 2011 Census microdata Teaching File for Scotland is due to be published on 25 June. Publication of safeguarded and secure microdata files of Scottish census data are planned for publication later in the year; NRS will provide users with proposals relating to the content and access arrangements for these additional microdata products.

Origin-destination statistics

Origin-destination statistics for Scotland will be included as part of the UK-wide data due to be published in the summer by ONS. NRS intends to consult users separately about proposals on additional origin-destination statistics for Scotland, eg on travel to places of study.

Further information in relation to statistical outputs from Scotland's 2011 Census can found in the Outputs Prospectus on the NRS census website - http://www.scotlandscensus.gov.uk/ - which is updated on a regular basis.

4.   NISRA update

4.1 Standard Releases to date – (all now complete)

First Release: Population and Household Estimates
Mostly July / September 2012, completed February 2013

Second Release: Key and Quick Statistics
Main public release December 2012, completed February 2013

Third Release: Detailed Characteristics
Started May 2013, completed November 2013

Fourth Release: Local Characteristics
Single release, completed March 2014

4.2 Other NI Releases to date

Northern Ireland's new local council areas (elections 22 May 2014)

- September 2013: Population Estimates for the 11 new LGDs
- January 2014: Key Statistics 're-issued'
- Statistical profiles used by BBC on Election Day

Report of the Registrar General 2012

- November 2013: included an overview chapter on the 2011 NI Census

4.3 UK and Specialist

UK releases:
Co-ordinated by ONS, as NRS data become available

Eurostat:
Co-ordinated and delivered (for UK) - by ONS

Flow-data (Origin-Destination) + Micro-data:
Co-ordinated and delivered (at UK level) by ONS
NISRA working with ONS and NRS
In principle, NISRA plans to mirror ONS release timetable,
(eg Micro-data Teaching file released January 2014)
Final decisions to be taken on the full range of products.

4.4 Other planned releases

- NISRA plans to review settlement classification and urban / rural classification during 2014/2015

- Single hard-copy 'summary report' planned for late 2014, based on Key Statistics

- General Report scheduled for late 2014

- Northern Ireland Longitudinal Study. 2011 Census link 'soft launch' in December 2013. Full launch event on 5 June 2014.

- Joint publication with CSO Ireland on 2011 Census in Ireland and Northern Ireland, scheduled for June 2014

4.5 Benefits Realisation

- Project underway

- Plan to conduct user survey in June 2014

Further details on all outputs available on the NI Census website at www.nisra.gov.uk/census

5.   ONS Geography update

5.1 Built-up areas and sub-divisions

2011 Census data for built-up areas in England and Wales provide information on the villages, towns and cities where people live, and allows comparisons between people living in built-up areas and those living elsewhere. The data was used to underpin a revised version of the rural-urban classification in England and Wales.

Census data for built-up areas (previously called urban areas) have been produced every 10 years since 1981. For 2011 Census outputs the areas have been renamed built-up areas. This renaming avoid previous inconsistency in terminology with the rural-urban classification, and provide a name that more accurately describes the geography.

A new methodology to capture the areas has been used in the 2011 version of the dataset, but it follows the rules used previously so that results are broadly comparable. As before, the definition follows a 'bricks and mortar' approach, with areas defined as built-up land with a minimum area of 20 hectares (200,000 m^2), while settlements within 200 metres of each other are linked. Built-up area sub-divisions are also identified to provide greater detail in the data, especially in the larger conurbations.

Datasets available for built-up areas include digital boundaries, lookups from OA to built-up area and from built-up area to a number of output geographies.

5.2 Rural/urban classification

A rural/urban classification for 2011 OA, LSOA and MSOA was published in August 2013. A rural-urban classification for higher level geographies is ongoing, with the local authority level classification expected to be published in June 2014.

The classifications are revised versions of those produced after the 2001 Census, but with additional detail in the urban domain. They were created by the University of Sheffield on behalf of a government working group.

5.3 Workplace Zones (WZs)

This new 2011 Census geography was derived by splitting or merging the 2011 OAs until they contain consistent numbers of workers. WZs are intended as a more suitable statistical and mapping base for business statistics than the residence-based 2011 Output Areas (OAs).

There are 53,578 WZs, with a mean worker population of 456.

Digital boundaries, with worker population weighted centroids, and lookup files to a number of geographies are available to complement the first census release on workplace zones on 23rd May 2014.

5.4 Best-fit percentage indicator

For each lookup from OA, LSOA and MSOA to a higher geography, a best-fit percentage indicator is now provided. This shows (where the population of the best-fitted OA/LSOA/MSOA is not wholly contained within the higher geography) the percentage of the OA/LSOA/SOA population that actually falls within the boundary of the higher geography it is best-fitted to.

2011 Census enumeration postcodes linked to 2011 OA, LSOA, MSOA, wards, parishes and local authorities

A lookup file of 2011 Census enumeration postcodes in England and Wales linked to the 2011 parishes, wards and local authorities in which they were located. 2011 Census enumeration postcodes are those postcodes that were recorded during the 2011 Census as containing one or more usual residents.

The enumeration postcodes have been assigned to the parishes, wards and local authority districts by plotting each postcode's centroid (the grid reference location of the average address in the postcode) directly in to the area boundaries.

More information about the lookup file can be found here. The lookup file can be downloaded from the Open Geography portal at https://geoportal.statistics.gov.uk/geoportal.

This lookup can be used with the 2011 Census estimates of the number of males, females and households for each enumeration postcode published in June 2013.

5.5 Dissemination of ONS geographic data–both portals now live

All geographic data released by ONS are freely available under the terms of the Open Government Licence .

5.6 Products

2011 Census Geography products can be accessed from the Open Geography portal. This provides open public access to view, query, download and visualise the spatial products that support national statistics, including all the 2011 Census geographic data products. The portal meets ONS strategic objectives to provide open, digital services and complies with the EU INSPIRE directive for harmonising the supply and access to spatial datasets across Europe

5.7 Linked data

ONS Geography core datasets can also now be accessed in machine-readable Linked Data format, at http://statistics.data.gov.uk.

6.   Eurostat

This release provides data for the UK as part of a European Union (EU) project to provide comparable statistics and transparency about the quality of census outputs across all EU member states. A total of 60 tables were provided ranging from UK geographical level down to LAU 2 on a wide range of census topics, these tables were made available as planned on 31st March 2014.

This release covers the whole of the UK and contains data already published as part of the 2011 Census for the UK as well as data on some new topics, including Country of Citizenship and Educational Attainment.

The data is freely available from the Census Hub, an EU tool for exploring the data. In addition to this, extracts of the selected data are also included in this release to download as reference tables. These extracts cover the topics of migration and workplace.

Further information about this release, including information on the variables used is also available as detailed metadata for the UK.

7. <u>VML</u>

7.1 Introduction

The Virtual Microdata Laboratory (VML) is ONS's flagship dissemination channel for researchers to access confidential microdata, which cannot be published openly, for approved research projects. The VML will be used to store the most detailed Microdata and Origin→Destination products from the 2011 Census.

This note gives Census Advisory Group members information about the VML and the application process to start a project, and an update on the project to replace the current infrastructure which will improve security, utility and the user experience.

7.2 Background

Access to the VML is provided for Approved Researchers working on defined and approved projects, through one of the secure rooms in London, Titchfield, Newport, Glasgow and Belfast.

Data accessed in this way, and any analyses undertaken, cannot be downloaded by researchers. Once analysis is complete, all results are checked to ensure that there is no risk of disclosure to specific individuals, households or businesses. These, non-disclosive, outputs are then emailed to the researcher by email and can be published and circulated as required.

The VML operates within a legal framework with penalties for researchers who break the rules they have signed up to.

7.2.1 Applying to use the VML

Access to the VML is only given once both the researcher, and the project to be completed, have been approved. The application and approval process is:

1.   Researcher contacts the VML team (vml@ons.gsi.gov.uk) and is asked to complete a Customer Request Form to describe the proposed project
1.1. Project purpose, data required and estimate of time to complete

2.   VML team and Data Owners agree whether project is appropriate use of data
2.1. For Census projects, this will be agreed through the Census Research Access Board (CRAB)

3.   Researcher applies for Approved Researcher status
3.1. Must demonstrate ability to complete proposed research project and commitment to maintaining confidentiality of data

4.   Researcher is given training in how to use the VML, the confidentiality requirements and the checks ONS will undertake to ensure outputs are non-disclosive

5.   Requested data are made available to researcher in a dedicated share within the VML

7.2.2 Data available within the VML

The VML holds a wide range of data from across ONS. These include:

- Business survey data (e.g. Business Structure, Annual Business Survey, Foreign Investment, Employment and Skills, Employee Relations)
- Earnings data (e.g. Annual Survey of Hours and Earnings)
- Social surveys (eg Labour Force Survey, General Lifestyle Survey)
- Longitudinal Study data

The VML will also hold the most detailed Microdata and Origin→Destination products from the 2011 Census.

7.3 VML Replacement Project

ONS is currently running a project to replace the VML, because the existing infrastructure will not be able to host and support all the data and analyses required in future.

The current VML infrastructure is 9 years old and comprises an ad-hoc collection of hardware, using two operating systems that are no longer supported. As a result of these unsupported systems, the VML does not hold appropriate Business Information Level (BIL) accreditation (Level 3 would be appropriate for the current data held and data planned to be stored). There is a risk that any future software problems could result in Information Assurance closing the VML.

The VML is also now prone to technical failure, with three major incidents in 2013 which each took two weeks to resolve.

The VML is a 32 Bit environment, which limits the range of analytical software that can be used and means current, 64 Bit, versions of standard packages (e.g. MS Office, SAS, SPSS) cannot be made available. Data processing is also slow, which increases the time required for analysis. Some particularly large datasets, such as 2011 Census microdata samples, will be difficult to analyse without a 64 bit system and are likely to prove almost impossible to host within the current infrastructure.

Although there will be no immediate change to the application or access procedures, the new VML will bring significant benefits to ONS and its users. These benefits include:

- A BIL3 accredited environment, to host data and analysis securely
- A 64 bit environment, resulting in:
  - Use of the latest, and a wider range of, software tools which will greatly facilitate work done by current users
- Faster processing times, making analysis easier and less time-consuming
- Significantly lower risk of technical failure, and less time taken to resolve these
- Fewer complaints from key users, as major concerns (software limitations, processing time and technical instability) all addressed
- A dissemination mechanism within Government for the Administrative Data Research

For Census users, one of the key benefits is likely to be the increased flexibility in the range of software that can be used to analyse data. For example, it may be possible to make additional tools available within the new VML, to facilitate analysis of Origin→Destination data from the 2011 Census.

Once the new VML is implemented and fully functional, further benefits could be realised without additional cost to ONS through encouraging new uses and users of the VML as follows:

- Encourage take-up of remote VML access by Government departments
  - Following model currently used by BoE and BIS

  - Could possibly expand to Local Authorities to facilitate analysis
    - e.g. Analysis of 2011 Census local area O$\rightarrow$D data

- Increase the number of secure rooms, to maximise access

  - Currently limited geographically, which restricts and frustrates users from some parts of country
    - Labs in London, Titchfield, Newport, Belfast and Glasgow only

  - Can, in principle, be hosted in any building with PSN/GSI connection and appropriate physical security
    - Central Government and Local Authority buildings
    - Would require additional work to understand Property and Security considerations
      - Ownership of, and access to, terminals
      - Physical and Information security

7.4 Other Developments

At the same time as replacing the VML infrastructure, ONS are also working with the UK Data Service (UKDS) to review access procedures, training and output clearance processes. This will ensure consistency in all areas between the UKDS Secure Lab and the VML. As part of this work, ONS are reviewing the criteria for the approval of researchers and of VML projects with a view to, in particular, give users within commercial organisations greater clarity on how they can access and use these data.

In addition to ensuring consistency and giving greater clarity to users, this work will help ensure that all processes are appropriate, and will identify any opportunities to simplify these and reduce the application time.

8. Successful projects in government award

In March 2014, the first conference to celebrate successful projects in government was held. The 2011 Census was recognised alongside five other projects, including the 2012 Olympics and the new Firth of Forth road bridge (the Queensferry Crossing).

ONS was asked to identify four of the main factors which we believed were critical to ensuring success.

- Budget - having enough money allocated to meet the specified targets and cover unforeseen risks and issues.
- People - having enough people with the necessary skills and experience at the right time.
- Success criteria - ensuring these were clear, agreed and shared across the whole programme.
- Strong programme management - a 'relentless' focus on delivery.

This prestigious event was chaired by the director general of the Major Projects Authority, Dr Norma Wood and included contributions by Lord Brown of Maddingley, Rt Hon Margaret Hodge MP and Rt Hon Francis Maude MP.

At the conclusion of the conference, ONS were presented with an award to recognise the success of the 2011 Census and the hard work, commitment and skill brought to the programme by all those that worked on it.