

# **Annex to the Registrars General's Agreement**

## **2011 CENSUS OUTPUTS: AIMS**

### **1 INTRODUCTION**

The UK Registrars General have collectively committed to joint working on the 2011 Census with the aim of maximising harmonisation across the separate Census operations. The objectives of the Registrars General are set out in the Registrars General's Agreement

([http://www.statistics.gov.uk/about/census/census2011/2011census\\_prog.asp](http://www.statistics.gov.uk/about/census/census2011/2011census_prog.asp)).

This document is an annex to that agreement and sets out key high level aims for ensuring, so far as practicable, that the outputs from the 2011 Census meet user needs.

### **2 UK CONSISTENCY**

The Registrars General of England and Wales, Scotland and Northern Ireland have, subject to the need for approval where appropriate by the relevant legislature, agreed to aim for harmonisation on a number of key aspects of the 2011 Censuses. These include:

- a common population base;
- common questions to be agreed wherever possible, diverging only in response to clear user needs, with such divergences minimised and the scope for cross-comparison maximised;
- common methodologies on disclosure control and estimation;
- consistent, coherent and accessible statistics for the UK and for each component country, with a joint database (and/or a common data schema) seen as a desirable way of facilitating that outcome; and
- a common approach to output specifications, data quality, data formats and timing of releases.

### **3 DATA ACCESS**

#### **3.1 Aim: Outputs free at point of delivery**

A key aim underpinning user access to 2011 Census outputs is that all standard output will be free to users at the point of delivery.

Through the Census Access project (England and Wales), SCROL (Scotland) and NICA (Northern Ireland) the three Census Offices delivered all standard 2011 output free to users at the point of delivery. The funding mechanism for these projects differed across the Census Offices; for example the bulk of the England and Wales project was funded through the Invest to Save Budget (ISB) initiative with further funding from a consortium of stakeholders – Department of Health, Local Government Association, Economic and Social Research Council, Office of the Deputy Prime Minister (ODPM).

Further work is required to determine which outputs will be laid before the Parliaments, and to develop the funding model for standard outputs from the 2011 Census that are not laid before Parliament.

### **3.2 Aim: Dissemination methods and media to keep up to date with technological innovation**

An overarching aim for the dissemination of 2011 Census outputs is that methods and media keep up to date with technological innovation. However a balance will need to be struck between taking account of any future technological innovations and developing the necessary dissemination tools and systems in sufficient time. Research will also be needed to establish the extent to which Census requirements may be met by corporate dissemination solutions and tools.

The current assumption is that standard pre-planned outputs will be disseminated via a number of media - paper reports, on-line, and writable media (CD, DVD) but with dissemination likely to be web dominated and paper reports kept to the minimum to meet the legal requirements for reports laid before parliament under section 4.1 of the two Census Acts. The user requirement for outputs on writable media will need to be established.

### **3.3 Aim: Flexible table generation on-line**

In addition to pre-planned outputs, the aim is to provide the facility for flexible table generation on-line. This would be at two levels;

- User-defined extractions from, and tallying of, standard tables whereby users can specify sub-groups and/or geographical areas from within published tables to construct tables of interest. (Facilities of this type are already available for local 2001 Census statistical output currently disseminated on the NeSS/SCROL/NICA websites);
- The facility for user-defined tables utilising standard variables but used in a bespoke fashion. Such tables could be generated via hyper-cubes or from protected micro-data. Software that enables this functionality has already been developed and is available as part of the 'Space Time Research' package of tools, which includes the Super-cross tabulation package used for 2001 Census outputs. The Australian Bureau of Statistics (ABS) is currently testing this with the aim of utilising it in its 2006 Census outputs to increase data utility. Statistical Disclosure Control software would also need to be developed to support this facility to apply adequate disclosure protection to the user generated tables. (The ABS is currently developing disclosure control methodology to support this facility).

The facility of flexible table generation online has the potential to reduce the user requirement for standard output and commissioned output from the 2011 Census. Approximately 80 per cent of commissioned output from the 2001 Census involved standard variables used in a bespoke fashion.

Clearly, this aim is subject to the development of sufficiently robust statistical disclosure control methodology and security systems.

### **3.4 Aim: On-line facility to produce graphs and maps of standard and user-defined output**

The facility for on-line mapping and graphing of 2011 Census output disseminated on the NeSS/SCROL/NICA websites is already available. There are also interactive maps and charts on the NS on-line web pages. If the facility for on-line user defined output is developed there is the potential for joining up all this functionality to provide on-line charting and mapping facilities for both standard and user defined 2011 Census output.

### **3.5 Aim: Comparisons between 2001 and 2011 Census outputs**

The aim is to produce comparisons between 2001 and 2011 Census outputs following research into reliability of measures of change taking into account issues such as changes to:

- definitions or questions;
- the geography for Census outputs; and
- statistical disclosure control methodology

Areas of poor coverage in either or both Censuses may mean that comparisons over time at OA level would not be robust for some areas.

Further research will be needed to determine what time series are meaningful and at what geographical level and additionally whether it will be possible to enable on-line user defined comparisons between 2001 and 2011 Census outputs.

### **3.6 Aim: To maximise data utility**

Different levels of access for different Census outputs are being considered in order to maximise data utility. Census tables for small populations, for example some ethnic groups particularly at low geographical levels, would be very sparse. Statistical disclosure control can compromise data utility because of the need to suppress detailed breakdowns. The aim is to make all tabular output 'fit for purpose' within disclosure control constraints. Data likely to be compromised could be subject to lower levels of statistical disclosure control and made available to users only under special licence or accessed in data labs with outputs checked prior to removal. An example could be the Origin/Destination tables which are very sparse. It may be appropriate to make these matrices available publicly at Super Output Area/Data Zone level, and available under licence at Output Area level, and allow users interested in specific journeys access to detailed micro-data in a safe setting.

Clear and equitable criteria for access would govern this arrangement.

### **3.7 Aim: Output prospectus and timetable to be published pre-release**

A pre-release outputs prospectus and timetable for statistical outputs and metadata will be published. Every endeavour will be made to keep to the published timetable but quality will not be compromised. Any unforeseen problems likely to result in an inability to meet the published timetable will be explained and communicated to users as early as possible, with revised publication dates.

### **3.8 Aim: Concurrent first release of statistics across the UK**

This aim is in accordance with the Registrar Generals' agreement.

### **3.9 Aim: Statistics will be released concurrently for all areas within England and Wales/Scotland/Northern Ireland**

It is probably inevitable that publication of statistics will be staged, for example in 2001 there were separate releases for population estimates, Key Statistics, Census Area Statistics, and Standard Tables. As in 2001, at each stage statistics will be released concurrently for all areas.

### **3.10 Aim: Concurrent publication of appropriate metadata with associated statistical outputs**

Not all quality measures will be available at the time of statistical release, however the intention is to publish quality measures such as response rates and imputation rates concurrently with associated statistical outputs. Other more specialised evaluation will be published later in accordance with the published timetable.

### **3.11 Aim: Joined-up and comparable UK outputs**

There will inevitably be some differences in questions across the UK countries to reflect local data needs and the decisions of devolved administrations. However, where differences do occur the aim will be to map results to a common framework to enable UK comparability at some level.

### **3.12 Aim: Joining-up Census outputs with other statistical outputs**

A high level aim is to join-up 2011 Census outputs with other statistical outputs. However, this could range from simply a link from census tabular output to other data on same topic through to the production of topic-specific reports that draw on Census and other data. One example would be building on the 2001 'Focus On' Reports.

Further discussion is required to determine which sources should contribute to topic output in the future and the extent to which Census outputs should be joined with other statistical outputs. If the Census Offices were to put significant resources into this it is likely that topic experts would need to spend less time on other activities and hence other outputs in the two year period around the Census may not be produced.

## **4 STATISTICAL DISCLOSURE CONTROL**

### **4.1 Aim: Common UK statistical disclosure control methodology for Census 2011 outputs that minimises disclosure risk whilst maximising data utility**

The Registrars General's statement of agreement for 2011 UK Census includes aiming for a common Statistical Disclosure Control (SDC) methodology.

2001 Census outputs were subject to differing SDC methodologies across the UK which led to much discontentment amongst users and impeded UK compatibility.

The goal for 2011 Census is for a common SDC methodology to be adopted across the UK which protects against the risk of disclosure whilst maximising data utility (taking into account user output requirements) and which complies fully to the commitment to confidentiality on the Census form and to the legislatures.

In 2001 there were different disclosure control methodologies across the UK because of different views about the level of acceptable risk and the perception of what constitutes disclosure. High level discussions across the UK offices on these SDC policy issues are planned with the intention of coming to an agreed view on these risks and perception issues as the base for all further methodological research for 2011 Census outputs.

Further research will cover SDC methodological issues for all Census outputs including pre-planned tables, flexible outputs, commissioned output and micro-data and for different levels of access (public, licensed and safe setting)

Common SDC methodologies will be fully evaluated in terms of a utility/risk continuum and in the context of the agreed SDC policy position.

It is intended that this work will be completed before consultation on detailed output requirements commence.

## **5 GEOGRAPHY**

### **5.1 Aim: An effective and flexible approach to output geographies**

ONS Census Division has been working with Neighbourhood Statistics and ONS Geography to establish the high level principles within which the 2011 Census output geography for **England and Wales** will be developed. Three key options have been identified

- (a) Define new geographies for 2011 Census outputs
- (b) Keep 2001 geographies largely unchanged (OAs, SOAs)
- (c) Keep 2001 geographies at some level (eg SOAs) and redraw OAs within this constraint

The current ONS recommendation for England and Wales is to maintain existing geographies for 2011 Census outputs. Some modification of the current OA geography will be required to take account of any significant population changes that will have occurred since the last census. (A paper on output geography for England and Wales is available at [http://www.statistic.gov.uk/census2001/cn\\_142.asp](http://www.statistic.gov.uk/census2001/cn_142.asp).)

In **Scotland** decisions have not yet been taken on high level geography principles, though it seems likely that they will be largely similar to those employed in recent Censuses.

In **Northern Ireland**, such decisions have yet to be taken and will need to be viewed in the context of the ongoing Review of Public Administration which is expected to result in the number of Local Government Districts being reduced from 26 to 7.

## **6 ADJUSTMENT AND QUALITY ASSURANCE**

### **6.1 Aim: Fully adjusted database**

The aim is for a database fully adjusted for missing responses to specific questions, (with the exception of any voluntary questions) and adjusted for estimated under and over coverage. It is also planned to undertake full (100 per cent) coding of questions. However, a cost/benefit analysis and confirmation of funding will be required before a final decision is reached.

### **6.2 Aim: No revisions**

Census output tables contain some 2 billion individual counts and take 2 years to produce. It would not be possible to update all of these counts, so the aim is to make no revisions to the 2011 Census outputs. (*Revisions* refer to changes to

headline census counts as a result of coverage error). This aim is underpinned by a commitment to put even greater effort into the quality assurance of results (see 6.4 below). Any necessary subsequent revisions will be made to mid-year population estimates if required.

While the timing of initial output release is likely to be similar to that of 2001 Census, the current plan is to capture and process data more quickly and complete coverage assessment processes faster than was the case in 2001 to allow more time for follow-up and quality assurance prior to publication.

### **6.3 Aim: Corrections policy in place in advance**

*Corrections* refer to changes to specific outputs as a result of coding or tabulation errors. Mechanisms will be put in place for users to report errors and for corrections to be prioritised. Users will be consulted on the most effective means of communicating corrections to the user population. Version control will be utilized to provide an audit trail and to enable users to access previous versions.

### **6.4 Aim: Joined-up and co-ordinated quality assurance policy to involve internal and external experts and bodies**

#### *Quality Assurance of Census-based population estimates*

The three Census Offices across the UK recognise the importance and potential benefits of a comprehensive and coordinated programme of quality assurance prior to the release of the 2011 results and are committed to identifying all available potential sources of useful information and relevant internal and external experts/bodies to assist with the work. For example, before the 2011 Census, ONS will liaise with Local Authorities to quality assure address lists for their areas, utilising alternative LA data sources. This will help inform the census field operation and will also provide additional information for internal ONS experts who will be responsible for quality assuring the estimates. GROS also intend to work with local authorities on quality assuring address lists, while NISRA will do likewise with Ordnance Survey

In any areas where a Census Office has particular concerns, we may discuss the estimates with the relevant LA or other government departments during the QA process, subject to strict confidentiality protocols and agreements.

#### *Quality Assurance of tables*

External bodies quality assured 2001 Census tables in an *ad hoc* manner with different users using different processes and identifying different errors after publication. If the facility for flexible tabular outputs from the 2011 Census is available there is the potential for a reduction in the volume of standard tabular output. This would make the internal QA process more manageable. In addition to internal QA of the tables produced, it is intended to plan early access to tables for selected users for quality assurance purposes.

## **7 SUMMARY OF CENSUS 2011 OUTPUT AIMS**

- Outputs free at point of delivery
- Dissemination methods and media to keep up to date with technological innovation
- On-line flexible table generation
- Improved data utility of some data outputs via differential access levels

- On-line mapping and graphing of standard and user-defined output
- 2001/2011 comparisons
- Pre-release output prospectus and timetable that is adhered to
- Concurrent first release of statistics across the UK
- Any given set of statistics will be released concurrently for all areas within a country of the UK
- Appropriate associated metadata published concurrently with statistical output
- Joined-up and comparable UK outputs
- Joining-up Census output with other statistical output
- Common UK SDC methodology for Census 2011 outputs that minimises disclosure risk whilst maximising data utility
- An effective and flexible approach to output geographies
- Fully adjusted database
- No revisions
- Corrections policy in place in advance
- Joined-up and co-ordinated quality assurance policy to involve internal and external experts and bodies.

## **ONS/GROS/NISRA**

**April 2006**