

Census 2001 Review and Evaluation

Census Geography: Executive Summary

ONS is carrying out a review and evaluation of the 2001 Census in England and Wales which will culminate in a Data Quality report being published, followed by a General Report.

Plans for individual reports on specific aspects of the Census operation and a timetable for release have been published.

Each report is written in isolation and is subject to amendments as processing progresses and further information comes to light.

Reports will be released on the ONS website in the form of a high level Executive Summary and a more detailed Evaluation Report.

November 2003

Content	Page
Project Objectives.....	2
Background.....	2
Methodology.....	2
Assessment and Lessons Learnt.....	3
Conclusions.....	4

Census 2001 Review and Evaluation

Project Objective

To provide and maintain an integrated, accurate and timely geographical service to all aspects of the 2001 Census, including geographical information, advice and support.

To minimise data redundancy and duplication of effort by establishing and maintaining a central database of census geography information and ensure, as far as is cost-effective and practical, that strategies for meeting census geography requirements are implemented in an integrated manner.

Background

Although only responsible for England and Wales geography the project took on a coordination role with the General Register Office for Scotland (GROS) and the Northern Ireland Statistical Research Agency (NISRA) to ensure that systems which make use of geography data took account of any differences.

The aims of Census Geography were to ensure that:

- enumeration areas were created which facilitated the efficient and accurate distribution and collection of census forms by enumerators, whilst attempting to equalise the workloads;
- a central geography database and systems were in place to assist in the processing of census forms by resolving queries and ensuring that all records had an accurate and valid postcode and grid reference assigned to them; and
- a separate geography for output, based on grouping postcodes, was created and made available to users.

Methodology

The main change for the 2001 Census was the introduction of major automation in the design of enumeration areas and subsequently in the creation of Output Areas (OAs). Advances in affordable computer power, Geography Information System (GIS) software and digital data, from the national mapping agency Ordnance Survey (OS) made this possible.

The use of GIS at the front end allowed the enumeration areas to be split from the areas used for the release of census output. This was a crucial and innovative advance. Enumeration areas and output areas serve two very different purposes and previously the same building block had to be used for both purposes, which gave a sub-optimal solution for either enumeration or output.

Enumeration Area Planning

An automated Geography Area Planning System (GAPS) developed jointly with a private sector company ESRI (UK) (the GIS software supplier) was used to plan the enumeration areas. The supplier was selected after a competitive European Union tendering exercise and contained the novel requirement that the supplier used two Office for National Statistics (ONS) staff as part of the development team. GAPS effectively automated the majority of the tasks previously done manually and allowed the operation to be completed in a shorter timescale with a tenth of the staff.

The requirement for a smaller field force for the 2001 Census, due to the use of post back for census forms, meant that GAPS could be used to plan enumeration areas that were larger than those used during the 1991 Census. It also meant that more flexibility could be introduced in order to enable enumerators to manage more than one area or for two or more enumerators to cover a number of districts jointly.

GAPS made extensive use of OS digital maps, boundaries and the ADDRESS-POINT gazetteer which were supplied at an attractive price after a Service Level Agreement (SLA) was signed between the two offices. Outputs from the system meant that enumerators were provided with 'customised' maps and lists of addresses preprinted into their enumeration record books (ERBs).

Processing of census forms

A geography database set up as part of the enumeration area planning exercise was used extensively during the processing of census forms to check the validity of postcodes and allow grid references to be added. Postcode and address queries raised during this process were passed to the geography team for resolution and various automated and manual systems (collectively

Census 2001 Review and Evaluation

known as the LOckheed Martin Address System (LOMAS)) were used to ensure the final census database contained accurate and valid postcodes and associated grid references.

Output Geography

Another first for 2001, and what many consider to be a world leading operation, was the separation of the collection geography from the output geography, again facilitated by the combined use of a GIS, digital boundaries and the fully grid referenced census database. An automated zoning system for output purposes (Output Area Production System (OAPS)) was developed jointly with Professor David Martin, from Southampton University, whose expertise was invaluable in developing theory into a working system.

OAs were created, as far as possible, by grouping together postcodes, thus allowing better integration between geographical information referenced by census and postcode geographies. All OAs were created above confidentiality thresholds, with population sizes standardised, internal social homogeneity maximised and irregular geographical shapes minimised.

Assessment and Lessons Learnt

Geography is crucial to the census operation, being present in all stages. The successful development and implementation of the 2001 geographical system is all the more impressive given that it was a key innovation. Innovation always carries risk; in this case the risks were well managed and contained. The detailed lessons learnt during that process are described below.

Although in live running the geography systems worked well, overall, the geographical aspects of each stage of the census process were not always considered sufficiently at the design and development stage. Geography therefore needs to assume a higher profile and become more integrated in the overall conduct and management of a future census. Geography 'experts' need to be involved, from the beginning, in the design of any systems which will make use of geography data.

System development was helped greatly by the use of integrated teams, with generalists and technical programming staff under the same management and this approach should be repeated. The more widespread use of the GIS and database software within ONS should make this easier to achieve.

Enumeration Area Planning

Overall this can be considered an overwhelming success with enumeration areas being planned in an extremely cost effective way. Enumerators were also provided with better, more up to date maps and the labour intensive task of writing addresses into their record books was largely removed. GAPS was also flexible enough to allow late changes, identified by Census District Managers (CDMs) during their check, to be incorporated and revised materials supplied.

The principle lessons learnt can be summarised as follows:

- the specialised nature of GIS systems meant that difficulties were encountered in finding sufficient, suitable qualified staff to develop the systems needed and this must be addressed in any future exercise;
- the task of ensuring that boundary data sets were consistent was very time consuming and sufficient resources should be made available in future at an early stage; and
- the printing of field staff maps and documentation was very time consuming and contracts with external printers should be set up well in advance to ensure that the materials are available when needed.

Processing of census forms

Although far more records than originally planned were passed to geography for query resolution, the flexible nature of the systems meant that they could be adapted to cope and ensure that crucial end dates were still met.

Census 2001 Review and Evaluation

The principle lessons learnt can be summarised as follows:

- experience gained this time will prove invaluable in designing systems for any future census and should enable an even quicker turn around time for query resolution; and
- careful consideration must be given to the balance between the need for timeliness and the need for total accuracy against a 'fit for purpose' solution especially in regard to grid reference allocation, as this has a significant impact on time and resource.

Output Geography

The creation of OAs was another resounding success which has created building blocks ideally suited to the presentation of census statistics as they are smaller than enumeration areas and, being constructed from post-codes allow other data to be easily referenced to them. Initial reaction from customers has been favourable, seeming to meet most of their needs.

A few users have commented that they would have liked more input into the physical design of the OAs but this was not possible due to time constraints and the need to retain uniformity over the whole of England and Wales. Users were consulted about the principles governing the creation of OA's and were provided with examples to illustrate the implementation of those principles. If development work was completed earlier in a future census, it would be possible to build on this aspect and allow a greater degree of involvement.

Another indicator of success is the fact that OAs have been adopted as the single small area building brick for National Statistics and are the basis for "Super Output Areas" to be used in Neighbourhood Statistics (NeSS).

The principle lessons learnt can be summarised as follows:

- OAs were created in a very cost effective manner over a short space of time to be fit for the purpose of disseminating 2001 Census Statistics; it may be possible to refine the boundaries to integrate them with the Ordnance Survey map database, and it will be necessary to consider the impact of changing topography on the boundaries in future.
- OAPS was basically developed within ONS by one person, despite strenuous efforts to find additional resources and this over-reliance on particular individuals must not be allowed to happen again; and
- changing to follow the new ONS Boundary Compliance Policy meant that OAs had to be created within wards in existence at the end of 2002, which caused major problems and delayed the production of OAs by 3-4 months. The timing of the introduction of new geography policies and the impact on outputs agreed with Census users should in future be considered carefully at an early stage in the policy formulation.

Conclusion

The moves from the mainly manual, labour intensive work carried out in the 1991 Census to the automated methods used in the 2001 Census were completed successfully. The development partnership with Professor David Martin resulted in an output geography system that is viewed as a world leader amongst countries with a complex geographical structure. However, technology moves on and the development work required for any future census should not be underestimated.

The importance of geography throughout the census operation needs to be recognised and sufficient and suitable resources put in place at an early stage. To help achieve this there is a need to ensure a strategy is in place to keep key specialist/technical staff throughout the life of any future project.

Census 2001 Review and Evaluation

Census Topics	Target Dates for Release
Legislation	Published
Non-Compliance (Executive Summary Only)	Published
Data Needs	Published
Geography	Published (Executive Summary)
Publicity	Published
Data Collection Development	Published
Data Collection Support	Published
Census Coverage Survey	Published
Processing	Published
Annex: Quality of Data Capture and Coding	Published
Downstream Processing	Published (Executive Summary)
Data Quality	
- Question non-response rates	Published
- Disclosure Control (Executive Summary only)	Published
- Data Validation (Executive Summary only)	Published
Edit & Imputation	Published
One Number Census	
- Quality Assurance	Published
- Lessons learnt (Executive Summary only)	Published
Output Policy	Published (Executive Summary)
Output Production	
- Part 1:Review of Output Released to date	Published (Executive Summary)
- Part 2:including Sample of Anonymised Records (SARs)/Origin Destination Matrices	Published
Census Access	Published
Programme Management	Published (Executive Summary)
Quality Report	Published
General Report	Published

Please note that the dates for release of individual evaluation reports noted above are target dates, and therefore subject to change. For the latest information please visit www.statistics.gov.uk/census2001/reviewevaluation.asp