

Local authority/census abstracts

Strand organiser: Eileen Howes, Greater London Authority

Using local administrative data sources to improve census coverage and for post census validation.

Piers Elias, Tees Valley Unlimited

This paper considers the administrative data sources available to Local Authorities to help in the build-up to the Census to identify hard-to-count groups and post census to use as a validation tool. Local Authorities have access to a variety of data sources which hold information potentially useful for this purpose and this presentation considers some of the data sources available at a variety of geographical levels (individual or household level, Census Output Area, Polling District). These could be used as background evidence to assist the Census Area Managers to target resources towards the hardest-to-count areas. The aim is to provide additional information above that provided by, or available from, ONS. Tees Valley Unlimited, working on behalf of five Unitary Authorities in the NE, is developing a points-based system to identify and map these areas. However, whether we will be able to gain access to the data available and obtain the appropriate permissions in time to be able to use the data fully may not be known until this very conference.

Email: piers.elias@teesvalleyunlimited.gov.uk

Census Quality Assurance Studies.

Jason Thomas, Jonathan Smith, Office for National Statistics

The quality assurance (QA) process for the 2011 census is essential for the timely identification of systematic and respondent error and ultimately in the production of robust census estimates. A central part of the census QA strategy is to work closely with local authorities (LAs) in developing the QA approach. This paper reports on the pilot study carried out from February to April 2010 and the emerging QA studies from May to September 2010. The purpose of the studies is to source and analyse LA data and 'soft' intelligence that may inform QA checks. The LAs for the pilot studies were chosen based on total gross migration flow and 2011 census non-response modelling to identify those LAs anticipated to be most difficult to enumerate in 2011. A wide geographic spread across England and Wales was also considered and most of the 10 Government Office Regions (GORs) were represented. This approach will be adopted for the QA studies which will include LAs from all GORs. Initial findings from the pilot include an improved understanding of the data held by LAs, legal implications around data access, data transfer options and storage capacities. ONS have maintained a dialogue with LAs. This has allowed the opportunity for LAs to provide feedback on the process and additional information on the data they provided. For census QA the pilot study has provided early insights into data sources and checks ONS has planned to undertake. The QA studies will provide further intelligence on analysis approaches and data sources.

Email: jason.thomas@ons.gsi.gov.uk

Census 2011 Quality Assurance Checks

David Fallows, Paula Guy, Office for National Statistics

The 2011 Census will arguably pose the biggest enumeration challenge ever in the UK. The scale of the operation required to enumerate the entire UK population in order to produce robust population estimates is huge and the job of making sure this process has been accurate is equally challenging. There are limitless factors which could impact on the accuracy of Census estimates. In order to identify and quantify these factors, a series of checks for all Census questions, comparing the characteristics and plausibility of the distributions of responses to Census questions against comparator data will be carried out. Checks comparing Census estimates against administrative and survey data sources will range from looking at demographic indicators such as sex ratios and fertility rates to comparing the age distributions of subpopulations such as migrants or students. Key checks will be incorporated into an automated system that for each check will display diagnostics of any significant discrepancies between Census estimates and expected values. The automation of checks provides the functionality to drilldown into both variables and geography to isolate the source of error. For example if the Census estimates of males for an LA are lower than expected values, the user can easily identify if the undercount is concentrated in a particular age group ie missing 20 to 24 yr old males or if the difference is concentrated in a specific MSOA or population subgroup. The automated system also includes the functionality to view automated check diagnostics at above LA levels of geography for all processed LAs, providing an emerging picture of Census variable distributions against comparator distributions at higher levels of geography throughout processing. This allows the user to identify trends or errors developing either nationally at the GOR level or for bespoke groupings of LAs such as LAs with large student populations or alternatively rural LAs. A supporting ad-hoc system will provide a secure environment where additional checks will be run and anomalies identified through the automated system can be further investigated. This paper will outline the Census QA checks being carried out for every LA. It will clarify the differences and relationship between automated and ad-hoc systems as well as the content, functionality and visualisation of the checks.

Email: david.fallows@ons.gsi.gov.uk

Straightening out the Censuses: Adventures in open standards, dimensalisation, description, comparison and dissemination of UK Census datasets.

Justin Hayes, University of Manchester

The aggregate outputs from UK censuses are unparalleled resources of information on the demographic and socio-economic characteristics of the UK population. Unfortunately, the size and complexity of the census datasets, combined with a lack of clear and consistent definitional structures and poor linkage of data and metadata (numbers and meaning) have traditionally presented considerable barriers to easy and informed use of census information. Developing open standards technologies and methods offer potentials for creating and enforcing structures, linking data and metadata, and describing and making census information available in comprehensive, comparable and flexible forms which make it easier to obtain, understand and make use of, and encourage the development of improved and innovative dissemination applications. The Census Dissemination Unit (CDU), based within Mimas at The University of Manchester, is funded by the ESRC Census Programme to disseminate the aggregate census to the UK academic community. This paper will describe the experiences and outcomes of the CDU's application of these methods and technologies to the 2001 Census.

Email: justin.hayes@manchester.ac.uk

Challenges and approaches in designing flexible delivery mechanisms for Census 2011.

Joe Traynor, Office for National Statistics

Introduction & Aims The possibilities offered by flexible online data delivery mechanisms represent a major step forward in terms of the ability of statistical and other agencies to meet user needs in the 21st century. Transforming policy into practice requires detailed analysis and consideration. The aims of this paper are twofold: - To describe the challenges involved in potentially delivering 2011 Census data via new online delivery mechanisms; and - To outline the ONS approach to exploring the ways in which flexible output delivery might be achieved, given complex mutually-dependent constraints.

Methods Detailed specifications based on 2001 Census outputs have been developed and users have been asked to submit their comments and requirements on the basis of this template. In the meantime, analysis of 2001 Census outputs by geography, age, table population and variable has started, with a view to identifying possible multi-dimensional datasets that would meet users' specified needs, as well as offering increased flexibility in outputs, while balancing the disclosure issues created by generating outputs on the basis of 19 geographies, multiple non-nested age bands and approximately 400 variables.

Results & Conclusions This is a work-in-progress: this section will describe and interpret in detail the complex relationships and dependencies that must be considered in the design of flexible outputs for 2011 Census data, as well as the research to date on how ONS might produce multi-dimensionalised datasets that would: - Maintain the level of detail that was available in 2001; - Produce consistent outputs (inter- and intra-tabular consistency); - Increase flexibility of outputs; - Consider data visualisation options; and - Protect individuals from attribute disclosure-risk.

Email: joe.traynor@ons.gsi.gov.uk

London's industrial structure 1981 to 2001.

Richard Cameron, Greater London Authority

The release of 2001 Census workplace data for small areas in London has opened up the opportunity for detailed analyses on the location of industry in London. The dataset comprises numbers working in each London ward by industry, occupation and ethnic group, and also the numbers travelling to work in each ward from each local authority in England and Wales. The data will allow us to present a picture of the detailed structure and commuting patterns of London's industries in 2001. It is also intended to look at the changes in industrial structure across censuses from 1981 to 2001, however this is complicated by differences in classifications, geography and coding over time. We hope to present ways to mitigate these differences.

Email: richard.cameron@london.gov.uk

Enabling local authority users to produce small area population projections.

Harvey Snowling, General Register Office for Scotland

The General Register Office for Scotland (GROS) publishes population projections for Scotland and its local authority areas every two years. However, there is some demand for population projections from planners who need data at more local level. It is not feasible for central government to produce projections for every level of geography within the local authority areas of Scotland. So, GROS will be releasing small area data that can be aggregated up to enable local authority users to produce population projections at different geographies. GROS has been engaged in a project with the Centre for Census & Survey Research and Fife Council, using POPGROUP, to

identify the data required to enable users to produce good quality population projections for small areas. As a result of this work, GROS has produced a sensitivity report, guidance for creating projections using POPGROUP, and has identified the data required to produce good quality projections. This paper looks at the results of the sensitivity analysis that examined the effects of running scenarios based on fertility, mortality and migration assumptions that were varied according to the level of detail that was provided. For example, in the case of fertility assumptions the investigation found that, for most purposes, the use of total fertility rates estimated from births, together with the national age-specific pattern, should be sufficient. But local age-specific fertility rates can add some accuracy to the projections, so the provision of 'age of mother' data at some level of detail is beneficial.

Email: Harvey.Snowling@gro-scotland.gsi.gov.uk

The UK Statistics Authority's Assessment Programme: Population Statistics.

Mark Pont, UK Statistics Authority

As part of its role to promote and safeguard the production of official statistics, the UK Statistics Authority has a statutory responsibility to assess official statistics against the Code of Practice. Statistics that comply with the Code are designated as National Statistics. Since the introduction of the new Code in January 2009, the Statistics Authority has published 34 assessment reports, covering the assessments of around 160 official statistical outputs. A programme is in place to re-assess all existing National Statistics, alongside assessments of other official statistics by the end of 2012. Some of ONS's Migration Statistics were assessed as part of the initial programme during 2009. At the same time, the Statistics Authority carried out a separate comprehensive review of the steps being taken across government to improve migration statistics. In addition, the first phase of a special assessment of the 2011 Censuses in the UK was published in spring 2010. The Statistics Authority plans to assess the remaining official population, migration and demographic statistics, published for the UK and its constituent countries during autumn 2010. This talk will describe the assessment process, some of its findings and successes to date, and explain the important role of users of statistics in contributing views about the statistics in question.

Email: mark.pont@statistics.gsi.gov.uk

Languages, ethnicity, and education in London.

*Michelle vonAhn, London Borough of Newham; Ruth Lupton, London School of Economics;
Charley Greenwood, Institute of Education; Dick Wiggins, Institute of Education*

The increasing linguistic diversity of the UK attracts much interest and debate among public service providers, educationalists and the public. Using 2008 PLASC data for London's state pupils reveals that over 300 languages are spoken. Around 60% of London pupils are English speakers however, there are over 40 languages spoken by more than 1,000 pupils. Bengali, Urdu and Somali are the top three languages spoken in London, other than language. This paper will illustrate the geographical distribution of languages spoken by in London pupils and demonstrate the implications for service provision. We will also illustrate how some of the ethnic categories that are widely used in analysis of Census data hide substantial linguistic diversity, particularly 'Black African' and 'White Other' and provide evidence to show that Language spoken provides a means to better understand the relationship between ethnicity and educational performance.

Email: michelle.von.ahn@newham.gov.uk

Measuring and managing the impact of immigration.

Pip Tyler, Yorkshire & Humber Regional Migration Partnership; Peter Boden, University of Leeds & Edge Analytics Ltd.

International migration continues to exercise the minds of the public, policy makers and politicians alike. There is a clear need and appetite to understand the way in which movements of people into and out of the UK's regions impact and shape future policy interventions. But the debate around immigration is often built around anecdote and perceived change rather than evidence led. Too often in the past, the poor quality of individual data sources has created a sense that immigration is essentially 'unknowable', and difficult. However, recent work by the University of Leeds and the Yorkshire and Humber Regional Migration Partnership (YHRMP) has demonstrated the way in which bringing various strands of data together at a local level can help tell a coherent narrative about how local authority areas are evolving and changing through migration. This paper reports on how the YHRMP has used its Regional Migrant Databank to create Local Migration Profiles, to initially assist partner agencies in the development of Joint Strategic Needs Assessments. The paper will illustrate the potential for the Databank and its Profiles to examine how asylum seekers are housed across local areas, in Childcare Sufficiency Assessments to contribute to knowledge about those most in need of child care support, and in giving an overall picture of the local community to the emerging UK Border Agency Local Immigration Teams.

Email: p.boden@leeds.ac.uk
