



Census Data Quality Advisory Working Group

Minutes of Census Data Quality Advisory Working Group 19 May 2011

Present

Peter Scrimgeour
 Laura Murison
 David Blue
 Esta Clark
 Kim Reimann (minutes)
 Glen Bramley
 Euan Smith
 Hazel McShane
 Lorraine Cook
 Alistair Harvey
 Allison Craig
 Andrew Ballinghall
 Cameron Thomas
 Heidi Goodship
 Jan Freeke
 Lesley Mann
 Tom Snowling

Organisation

Census Director, National Records of Scotland (NRS)
 Census Data Quality, NRS
 Census Data Quality, NRS
 Population and Migration Statistics, NRS
 Census Division, NRS
 Heriot-Watt University
 Scottish Government (Scottish Neighbourhood Statistics)
 Scottish Government (Local Government Finance)
 Convention of Scottish Local Authorities (COSLA)
 Edinburgh City Council
 North Ayrshire Council
 Fife Council
 Highland Council
 Scottish Borders Council
 Glasgow City Council
 North Lanarkshire Council
 Aberdeen City Council

Apologies:

Tony Jenkins	Dundee City Council
Paul Davison	Stirling Council/Local Area Research and Intelligence Association (LARIA)
Kirsty MacLachlan	Head of Demography Division, NRS

1. Welcome, minutes and actions from 12 January 2011

- 1.1 Peter welcomed everyone to the third meeting of the Census Data Quality Advisory Working Group.
- 1.2 The minutes from the previous meeting were accepted, and noted that Laura had actioned the point in paragraph 5.3 – templates were taken to the next Population and Migration Statistics (PAMS) meeting for comment.

2. Census Data Quality Assurance Strategy - Laura Murison

- 2.1 Laura presented an overview of the Census Data Quality Assurance Strategy paper which outlines the strategy for the next 18 months and is to be published online at the end of the month. Comments on all aspects of this paper are welcome.

- Phase 1: July – Dec 2011. Initial feel for the data, to check for possible errors. Data will be received in 10 blocks throughout this period.
 - Phase 2: Jan- April 2012. This might be considered to be the most challenging phase as it involves extensive quality assurance of the population estimates by local authority, age and sex.
 - Phase 3: Aug- Nov 2012. Once output areas are created, this phase will check to see if the geographies that are created will be satisfactory. Will check to see if editing and imputation processes are satisfactory at output area level.
- 2.2 Comparator sources: the rest of the paper addresses what adjustment methods might be used in order to address issues. Certain local authorities will be targeted for analysis against expected Scottish Longitudinal Study (SLS) members where comparator local authorities would not be appropriate. Comments especially welcome regarding comparator sources.
- 2.3 Placeholders: Will be used to adjust for bias within households and those missed. The Labour Force Survey (LFS) also helps inform the adjustment for household bias (i.e. whether whole households or only people within households are to be imputed). Around 96-97% of census questionnaires were returned.
- 2.4 There are two stages to the census coverage process. The first is to decide what the ratio estimates should be. The second creates the skeleton records for missed people
- 2.5 It was suggested that an outline of coverage estimation and adjustment processes should be included in the Census Data Quality Assurance Strategy paper and that the meaning of the ratio estimates should be better explained.
- 2.6 As many comparator sources as possible and as appropriate will be used.

Action Point 1: All asked to provide comments on the Census Data Quality Assurance strategy paper by the end of May.

Action Point 2: The UK Statistics Authority are in the process of writing their Phase 2 of the Special Assessment of the 2011 Population Census. The census team is to share this strategy paper with the assessors.

Action Point 3: Laura to circulate the slides from Neil Jackson's presentation from a previous meeting and to include an outline of coverage estimation and adjustment in the Census Data Quality Assurance Strategy.

Action Point 4: David to clarify method B (p12/13) of the strategy paper regarding how Child Benefit Claimants information can be used here.

3. Quality Assurance Checks and Tolerances - David Blue

- 3.1 David went through the main points from the checks and tolerances and identified sub-groups table that was circulated. This included information on checks, stages, comparator sources and tolerance methods for geography.
- 3.2 The quality assurance process begins as soon as the data is received.

- 3.3 Issues were discussed regarding registration of children – counting those that are not registered for school and potential alternative sources of acquiring this data.
- 3.4 The group advised that further education students data should be included in the student check.
- 3.5 Council tax data: Household Estimates and Projections branch may get this with student exemptions which could be helpful for analysis.
- 3.6 Glen wondered if using a +/-5% tolerance for error for property checks was too large. David explained that tolerance checks are automated at the beginning to help identify potential problem areas that might need special consideration. They will be revised after this first automated process and so this level may change. Glen emphasised the importance of the property count check as there should be a reconciliation of property counts against questionnaires received in the census.
- 3.7 Euan wondered if using the same tolerances for all data zone levels might identify new builds only. David explained that Household Estimates and Projections branch should have work on this for areas undergoing lots of growth. Laura also put forth that the assessor's portal will also be used and this should help identify new build information. Property counts will be used, along with mid-year population estimates. It was suggested that changes in small area population estimates from 2001 to 2009 could be examined in advance.
- 3.8 The Royal Mail's postcode address file comes from a few different sources. This address file is what the census uses initially to help identify where to send out census forms.
- 3.9 David is currently doing tests of the Data Quality Management System (DQMS) to see how well it works and to test the tolerance levels.
- 3.10 Jan asked Esta whether population estimates will be produced for use in census Quality Assurance (QA) and Esta confirmed that estimates for census day and 2011 rolled forward from 2010 will be produced initially at local authority level and then at data zone level a few months later.
- 3.11 Cameron outlined the interest that there is in migrant workers which are a hard to reach group. The group agreed that there are no data sources that reliably record when migrant workers leave the country again. The migrant workers scan will be used in QA but it does not record exits.

Action Point 5: Laura to check what data on Council Tax exemptions Household Estimates and Projections branch receive and how this might inform QA processes.

Action Point 6: David to present results of his tests of the Data Quality Management System when complete.

4. **Local Authority Templates (these are included in Appendices 3a – 3c in the Census Data Quality Assurance Strategy attached at item 2) – Update on completion and data sources - Laura Murison**

4.1 Laura explained that work will begin incorporating feedback in January 2012. Shetland was given as an example of a response that had come in. Shetland highlighted topics they thought would help in analysis such as their migrant worker movements, students, and the closure of an RAF base. Laura welcomed further comments for aggregate levels and went through some of the existing comments received.

4.2 Laura went through the list of local authority data sources that have been identified and the group worked through which should be suggested to local authorities to give more guidance. The list of suggested sources is:

- Detailed breakdowns of council tax exemptions (e.g. occupied exemptions categorised by students, armed forces etc.). Counts of each type of exemption by local authority (and data zone if possible).
- House constructions/demolitions close to Census Day.
- Large communal establishment changes since the last census (e.g. closure of a prison or armed forces base).
- Households in multiple occupancy. Counts by local authority (and data zone if possible).
- Private landlords. Counts of privately renting households by local authority (and data zone if possible).

Action Point 7: Laura to write again to Population and Migration Statistics (PAMS) representatives who received the local authority templates initially to reiterate that aggregate data is requested for quality assurance purposes and to circulate the list of suggested data sources.

Action Point 8: Local authorities hold some level of data regarding landlord register. Laura, Alistair and Tom to consider how to action private rented accommodation by data zone data. However, this data might realistically only be held at the local authority level.

Action Point 9: Laura to pursue information on Council Tax exemptions and landlord register and to pass back findings to the working group.

5. **Update on Local Authority Summaries - David Blue**

5.1 David presented that local authority summaries will be published in-house, around the end of June. There will be one for every local authority and one for Scotland and will be updated as and when new data is received.

Action Point 10: David to circulate local authority summaries for each local authority to the relevant PAMS member along with the Scotland level summaries.

6. Live running – frequency of meetings of the advisory group and approach - Laura Murison

- 6.1 It was agreed that there should be a couple of meetings during Phase 1 to July. The 1st meeting in Sept/Oct, when half of the processing units are complete, and the 2nd in December 2011 when final processing units data is received.
- 6.2 Phase 2, demographics and QA coverage process, checks on comparator sources: for the third and fourth meetings for February & April 2012.
- 6.3 For Phase 3, information on prior analysis undertaken will be distributed before meeting so that members can be briefed on findings and findings can be discussed. One meeting in October 2012 and one wrap-up/ evaluation meeting shortly after this.
- 6.4 Local authorities were asked to pass in their completed Data Confidentiality agreement forms at the end of the meeting.

Action Point 11: It will be necessary to sign another data confidentiality form from the NHS Central Register. Laura to circulate this form by email.

Action Point 12: Members of the working group to sign the three confidentiality agreements and post to Laura.

7. Local Authority Aggregate Data – Geography - Laura Murison

- 7.1 Some issues with aggregate-level geographies were discussed. Cameron mentioned that there is an ongoing problem with GIS shape files. Laura stated that Tricia Couper could be contacted at National Records of Scotland (NRS) if members are unsure if they have the correct updated postcodes data and geographies. These are updated regularly.
- 7.2 Euan mentioned that data zone geographies are still in the process of changing, as Scottish Government (SG) are still to respond to the consultation. 2011 census output areas will grow to incorporate other areas. Those output areas that have shrunk enough will merge with other output areas.

8. Any Other Business and Date of Next Meeting

- 8.1 Access to the 2001 Census Data: Supertable freely available for download from Space Time Research. (for more information contact Ed Turnbull at NRS).
- 8.2 Once Scotland's Census Results On Line (SCROL) 2 is released, it will also contain SCROL 1 information from the 2001 census. There are no plans at the moment to add in earlier data from previous censuses (before 2011). SCROL 2 will also adopt a new name. Further information on this will be at the next PAMS meeting.
- 8.3 Date of Next Meeting: September/October 2011