Scotland's Census 2021
Statistical Disclosure Control and Outputs
Stakeholder Event Summary

September 2017

Table of Contents

Objective ID :

1.   Introduction

National Records of Scotland (NRS) held a stakeholder event on 28 June 2017 on Statistical Disclosure Control(SDC) and Outputs. NRS held two sessions, a morning and afternoon session, and invited stakeholders to give their view on the SDC methods and outputs proposed for Scotland's Census 2021.

NRS invited users who had previously indicated they would be interested in consulting on Statistical Disclosure Control (SDC) and/or outputs. The event was also promoted through the Scotland's Census newsletter and on the Scotland's Census website.

The event had a presentation on the proposed SDC method for 2021 and the outputs dissemination system. The presentation and an example flexible table builder were made available online after the event, along with an online survey to users who had been unable to attend the event to give feedback. This survey closed Friday 1st September 2017.

In total 33 stakeholders attended over the two sessions, representing a range of different areas including; central government, local government, academics, special interest groups and members of the public.

2.   Statistical Disclosure Control

Statistical Disclosure Control is used to prevent the release of any data that could lead to identification or disclosure of information about individuals or households. In 2011, NRS used targeted record swapping and table redesign to protect against disclosure. This was a time consuming process as every table had to be manually checked for disclosure. This delayed the release of outputs and meant that users did not always receive the table they wanted.

For the 2021 Census we are planning to continue using targeted record swapping, but will combine this with cell key perturbation rather than table redesign. A similar method is currently used by the Australian Bureau of Statistics on their census data. This method makes small changes to some of the cells in a table to add more uncertainty to protect against disclosure, particularly disclosure which occurs when one table is compared to other tables with similar data. Therefore, this method will allow users to create their own tables and would speed up the release of the data.

Stakeholders were asked what they thought of the proposed SDC methods.

Users liked the fact that the method would speed up the release of the data and that the method applied uncertainty in a repeatable way so that tables reproduced at a different time or by a different person would always stay the same. Users also liked the idea that some higher level geographies may be unperturbed.

Users asked whether univariate data would also be available unperturbed. This is more complicated than leaving high level geographies unperturbed, but it is being considered and will require more research.

Objective ID :

Overall most stakeholders appeared to be happy with the proposed SDC methodology, as it would speed up the release of data and allow users to create their own tables. A few stakeholders questioned whether cell perturbation was necessary or whether the combination of record swapping and imputation provided enough protection. Over-imputation was also suggested as an SDC method, but as this requires discarding some census records to impute false records this is not seen as an acceptable method.

Some stakeholders were interested in what impact the cell key perturbation would have on linked data and data research. This will be considered in the future but there is potential for a form of the cell key perturbation code to be supplied so that similar measures can applied to outputs produced in other environments.

3.   Outputs

Stakeholders were shown mockups for potential 2021 Census webpages with suggested features and functionality. Users were then given an example flexible table builder to use, to test the functionality and their experience using the website.

3.1   2021 webpage mockups

The mockup of the 2021 Census website homepage appeared to be popular with stakeholders. Users liked the latest news and publications section at the side of the page. Stakeholders liked the idea of using symbols to represent different categories of data as they were easy for users to recognise, but there were lots of comments around the example icons chosen as users did not feel these were representative of the population and felt they were too restrictive. There were some suggestions on how these icons could be improved which will be considered.

NRS is proposing to use more visualisations to represent the data in 2021. Stakeholders liked the visualisations and were keen to have lots of functionality around these. Users would like to be able to compare two areas and different variables, as well as being able to add these visualisations to their own documents or webpages. Stakeholders were particularly interested in being able to see time series data that compare census data over time. It was asked whether you could hover on visualisations and graphs to get the corresponding number, this had been considered but hovering may result in accessibility issues so this will need further consideration.

NRS is also proposing a mobile friendly site or app to allow users to easily access high level data on a range of devices. Stakeholders thought that this could be useful, but had no preference for an app over a website.

For 2021, we intend to redesign the table search to allow users to search by variable name or geography rather than the table number. We also plan to enable users to select the area from the map first so that the search lists all tables available for that geography. Stakeholders liked the move away from searching by table number.

Objective ID :

Users were interested in whether they would be able to use the map search for lower level geographies and to aggregate geographies.

NRS also had mockups of what a proposed flexible table builder could look like, including messages on SDC and what might happen a table didn't pass SDC checks. Stakeholders thought that these messages would be useful, and wanted to know what level of detail would be available for each variable and what groupings would be available. Our current idea is that we would make the most detailed breakdown possible available, subject to passing SDC, and users would be able to aggregate their own categories. Stakeholders also wanted to be able to send the link to a table that they created to others, so they would get exactly the same results, this is one of the requirements for the system.

3.2   Flexible Table Builder

NRS had an example flexible table builder and a worksheet with suggested exercises for stakeholders to try so that they could explore the full functionality of a flexible table builder. Stakeholders were invited to give their feedback on sticky notes and NRS staff were available to help use the website and answer any questions.

Stakeholders likes the ability to view the data as a graph or in a map, but would like more options to customise these. Users also liked that data could be downloaded in different formats.

Attendees felt that the flexible table builder could be more user friendly and easy to use. Some of the functionality was not obvious and stakeholders thought this needed to be explained more clearly, and the user guide needed to be modified and expanded on. It was suggested that training on how to use the builder would be helpful. Users also found that the site did not work as well on tablet devices.

4.   Online Feedback Survey

After the event the material was put online with an online survey for interested parties to complete. This survey was open for 7 weeks. Users were asked to look at the presentation before completing the survey.

NRS received 18 responses to the survey, but 3 were excluded from analysis as they did not contain any information.

NRS asked respondents how satisfied they were the SDC proposals for the 2021 census. 8 users were satisfied or very satisfied, 3 were neither satisfied nor dissatisfied, and 4 were dissatisfied. This is a much higher proportion of dissatisfied stakeholders than at the event where we were able to discuss the SDC methodology with users and answer any questions, and there were no comments on why these 4 were dissatisfied with the SDC. Out of these 4, 2 stated that they thought that allowing users to make their own tables was a good idea, this can only be done with the SDC methodology proposed.

Objective ID :

Respondents were asked for their comments on the proposal to allow users to make their own tables. 79 per cent of respondents thought this was a good or very good idea.

NRS asked how users found the use of the flexible table builder and for any comments on the system. 9 users thought it was easy or very easy to use and 2 thought it was difficult or very difficult. However some users who found the system easy to use expressed concern that it may be more difficult for users with less knowledge of the census or using data. It was also suggested that some training may be needed. Some users would also like the functionality to create derived variables based on multiple existing variables.

Respondents were also asked about whether a mobile friendly site would be useful. Users thought that this may be useful for frequent users and could be useful when out or looking for data on a specific area.

5.   Next Steps

NRS will consider all of the feedback received and will use this to inform the continued development of SDC and Outputs for Scotland's Census 2021. In particular, we will consider how we incorporate users suggestions, particularly around functionality and accessibility in the design of the flexible table builder and the web pages. We will also continue to engage with stakeholders during this process as our plans develop further to ensure that the 2021 outputs meet user needs.

More information and details about upcoming events can be found on the Scotland's Census website, by subscribing to the Scotland's Census newsletter and following us on Twitter @NatRecordsScot

Objective ID :