

The Quality Management Strategy (QMS) in the New Zealand 2001 Census of Population and Dwellings and How this Shapes the Future.

Nancy McBeth, Catherine McGill, Tracy Moore, Denise McGregor, Statistics New Zealand

E-mail: Nancy_McBeth@stats.govt.nz

Introduction

New Zealand has a long tradition of conducting five-yearly population censuses. Unlike many other countries, New Zealand still uses the enumerator drop-off and collect method, with every one of its four million people required to answer the same set of questions. In both 1996 and 2001, Statistics New Zealand (SNZ) provided bilingual forms in English and Māori (the language of New Zealand's indigenous population), which collected the same information as the English-only forms.

New Zealand census forms consist of individual forms completed by or on behalf of everyone in New Zealand on census night, and dwelling forms. The latter provides information on the attributes and characteristics of dwellings, as well as information used to produce family and household statistics.

Following the data collection phase, 2001 Census data was processed using scanning and image-based processing, and then evaluated and macro-edited. This process involved assessing the data against previously developed fitness for use specifications, to determine whether the data was of an acceptable quality. Following this process, data of a fit standard was published. Unlike previous censuses, data was available within 12 months of census day, which met user expectations. One of the reasons that the 2001 Census was successful in meeting this goal was the approach to quality management.

For the 1996 Census, SNZ aimed to provide good quality data within budget, and to have the output database ready for use a year after census day. However, this was not achieved, which demonstrated the difficulties in delivering a high level of quality given a fixed budget and timeframe. For the 2001 Census, it was clear that a different approach would be needed, and this became known as the 2001 Census Quality Management Strategy (QMS). The strategy sought to provide direction for the 2001 Census in terms of quality and ensuring that agreed output quality levels were met. The QMS marked a significant change in philosophy, with a move to an overarching strategy which would be followed through the developmental and operational phases of the census (ie development, enumeration, processing, evaluation and output).

The first section of this paper outlines the 2001 Census QMS, focusing on the output quality goals that were set

and the strategies employed to achieve them. The second section looks at the quality of census variables through case studies. The paper concludes by looking at the review of the 2001 QMS and the resulting recommendations for the 2006 Census.

2001 Census Quality Management Strategy

The 2001 Census quality philosophy reflected major shifts in Statistic New Zealand's wider strategic direction. The overall strategic direction for the 2001 Census was set by the Government Statistician, and took a clear output focus. At a similar time, wider corporate thinking led to the development of a more strategic approach to quality across the organisation, in the form of the draft Statistics New Zealand Quality Assurance Framework. This provided a framework for managing and improving quality in surveys.

Output Quality Goals for 2001

Central to the strategic approaches was clearly setting out the quality goals, not just in terms of a process, but rather in terms of the outputs. This approach required a clear identification of the key output quality priorities, which could then be used as the basis for coordinated decision-making across the phases. In terms of the 2001 Census, the key output goals were to:

- maximise the coverage of dwellings and people
- produce relevant and useful outputs
- produce outputs on time
- produce outputs consistent with previous censuses
- produce transparent and well-documented methodology underlying the census outputs
- conduct the census within budget
- control and reduce respondent load, and
- evaluate the data and methodologies used.

Strategies to Achieve Output Quality Goals

With such a wide ranging set of goals, it was critical to identify a set of strategies that could be consistently applied across the census phases. A key feature of the strategies was that they were able to be operationalised across the census cycle, but retain the output focus.

Five key strategies were identified in the QMS to achieve the output quality goals, and in doing so, producing data of sufficient quality:

1. differentiate between levels of output quality

2. agree on quality standards for outputs in 2001
3. identify areas of high risk to quality throughout the census process
4. manage and reduce this risk
5. monitor and measure quality and provide feedback to users.

Each of the strategies is outlined below.

Strategy One: Differentiate Between Levels of Quality

Learning from the experiences of 1996, it was decided in planning for the 2001 Census to prioritise levels of output quality. This decision reflected that SNZ did not have the resources or time to produce all outputs at a high level of quality. Three different levels of data quality were applied to outputs throughout the census development and operational phases. Topic-based census variables were listed under their relevant quality level according to their importance and use (foremost, defining, supplementary). The distinction between these three levels and the implications of differing levels of quality were then to be considered in all aspects of the development of the 2001 Census. Despite being a relatively small agency, SNZ has a wide range of subject-matter expertise, connected to different parts of the census data user community. Categorising the full range of census variables into the three levels was carried out in conjunction with these subject-matter areas.

Level 1: Foremost Variables

The decision was made to categorise those variables which are core to the census as foremost variables. These variables had the highest priority in terms of quality, time and resources: age, sex, ethnicity, and location. In addition, the output of final counts of population and dwellings was also classified as a foremost variable. Variables in this category are prescribed by legislation, but also in the New Zealand context are the fundamental demographic reason for conducting a census programme.

To ensure these variables were of the highest quality, a set of instructions was provided for the various census projects to follow. These instructions included having all questions relating to foremost outputs on the front page (of the English Individual Form) to maximise responses and allow less intrusive enumerator checking during collection. In the case of extreme non-response, collection of only the foremost variables was an option of last resort. During the processing phase, the respondent's actual response was to be taken unless very strong evidence suggested change to the response was required. Micro editing was to be used to actively look for introduced and respondent error. Most of these variables (age, sex, usual residence meshblock¹ and census night meshblock) were to be imputed for item non-response. Macro editing was used during the

evaluation phase if problems were found with the data or the processing system that needed to be addressed.

Level 2: Defining Variables

Those variables which define key subject populations for which the census provides measures that are important for policy development, evaluation or monitoring were categorised as defining variables. Defining variables are used frequently in cross tabulations with foremost variables. They represent key sub-populations and measures that are of high public interest, for example, families and households, labour force status, country of birth, and income. These variables are closely linked to the main purpose of a census, and in the New Zealand context may only be available in detail, for example at subnational level, from the population census.

These variables had second priority in terms of resources across all phases of the census. For example, enumerators were only instructed to check foremost variable questions and, therefore, if respondents did not answer some defining variable questions, this would not be detected during enumeration. Less editing and operator intervention was to be carried out during the processing and evaluation phases. Consequently, some inconsistencies, from respondent error and introduced error, could be left in the data. One variable, labour force status, was imputed.

Level 3: Supplementary Variables

In the absence of a wider integrated social statistics strategy in New Zealand, the New Zealand population census contains a number of variables which do not directly fit in with the main purpose of a census, but are still of importance to certain groups. These variables are categorised as supplementary variables and had third priority in terms of effort and resources. However, there were minimum quality standards that had to be met in order to make the output data suitable for use. Examples of supplementary variables include: occupation, language, religion, workplace address, and rent. As with the defining variables, supplementary variables had no checking in the enumeration phase. There was to be even less editing and operator intervention during the processing and evaluation stages. It was recognised that some inconsistencies (from respondent error and introduced error) would be left in the data. The instructions also made it clear that these variables may not be output at the finest level without a set of caveats. No imputation was to be applied to these outputs. The option of sampling was available for these variables.

Strategy Two: Agree on Quality Standards for Outputs in 2001

In addition to differentiated levels of output quality, SNZ aimed to provide census data that met at least a minimum quality standard required by users so that the

data was relevant and useful. This minimum quality standard for variables was identified in the form of 'Fitness for Use Specifications', which reflected the three quality levels and user requirements. These specifications were to contain all of the information that the census development teams would require to develop systems and processes that would meet the quality standards. The specifications were 'living documents' that could be revised until the Census Dress Rehearsal, after which time a formal agreement was required to make changes.

Strategy Three: Identify Areas of High Risk to Quality Throughout the Census Process

An important part of ensuring quality in a census is the identification of risks. Early in the development of the 2001 Census, risks for the major phases of census were identified. These in turn were prioritised so that key risks to census had the highest priority in terms of resources for risk management and monitoring. These risks and the risk mitigation strategies were monitored by the managers of the various phases, and risks across the project were monitored by the project governance team.

Strategy Four: Manage and Reduce this Risk

Two main tools were used to manage risk in 2001: Statistics Project Management and a Testing Strategy.

Statistics Project Management

An SNZ in-house project management tool was used to manage the 2001 Census, which ensured that all projects were planned, tested, well consulted on and would not adversely affect quality of outputs. An important part of this tool was quality control, which was carried out through a review process for each project deliverable. Each project had a quality plan that included the deliverables and the quality standards against which the deliverables were to be assessed. Quality checking was carried out by nominated quality reviewers who, either individually or collectively, reviewed the deliverables to check they were correct, complete and in accordance with any applicable standards. This review process was to help provide a higher level of quality for deliverables and thus for the census process as a whole.

Testing Strategy

In the 1996 Census, there was a short timeframe for the development and testing of the systems. The testing methods were time consuming and there was insufficient rigorous testing, resulting in data quality problems. A more structured testing programme was clearly required for the 2001 Census.

In the 2001 Census, testing was planned to minimise the risk of any parts of the census process not working to specification and affecting quality. A full testing

programme was developed to test each discrete module and its interaction with other modules. The final census test, the Census Dress Rehearsal, was a full test of all the systems and their inter-relationships, with only minor changes allowed before the actual census. It was expected that a greater emphasis on testing during the development stage would reduce the number of problems occurring during the census itself.

Strategy Five: Monitor and Measure Quality and Provide Feedback to Users

Monitoring Quality in the 2001 Census

The overall aim of quality monitoring for the 2001 Census was to ensure data quality met the requirements of the quality standards, that the processes established to minimise error were working effectively, and that unanticipated error did not affect quality. The QMS specified that quality monitoring was to be focused on the high risk areas, and in terms of variables, focused on the foremost and defining variables. Information from monitoring had to be provided in time to influence the quality of the results (through rectification).

The evaluation phases contained a number of distinct processes which enabled data quality to be monitored. This included macro evaluation, where data was looked at from a macro viewpoint to identify expected data trends and comparisons with other data sources. Systems evaluation monitoring included checks of automatic system routines and derivations. Both areas of the evaluation phase occurred during processing to allow major data and system problems to be identified and fixed where necessary. In addition to the evaluation phase, a management information system was developed to monitor results and determine whether processes were meeting targets of cost, time and budget. This system was also used to provide early warnings of deviations from agreed targets so that action could be taken. Contracts with external companies were also monitored and information provided on the quality of the data.

Measuring Quality in the 2001 Census

The quality standards outlined the data quality required for key uses of the data. While it is not possible to provide an overall measure of quality achieved, there are various indicators or measures for contributors to quality. In general, the key indicators of quality in 2001 were to be:

- response rates; both unit (person or dwelling) and item (individual question)
- coverage (mis-enumeration)
- error from recognition, operators, or other system error
- amount of imputation required (at unit and item level)

- degree of editing required, and degree to which respondents' answers have been adjusted, and
- ability to compare 2001 data with data from previous censuses.

The MIS and quality monitoring (including evaluation) as outlined above were intended to provide information about the data quality at enumeration, processing and output. In addition, a post-enumeration survey (PES) was conducted to provide an independent check of coverage (undercount and overcount).

Provide Feedback to Users

As well as monitoring and measuring quality, it is important to inform users of the quality of data they are receiving. SNZ has provided summary metadata on the quality of census data for external users. In addition, more detailed information on data quality was available to SNZ users of census data through the official census output tool and evaluation documents.

The strategies therefore combined to ensure a strong, but targeted approach to quality in the 2001 Census. They enabled future decisions about quality to be seen as a core, but controlled, element of the various census processes. With the QMS placing strong emphasis on forward planning, SNZ was confident that, in conducting the 2001 Census, it would be able to better manage the quality, timeliness and budget trade-offs. However, the success of any theory is shown in the practice, and the next section outlines SNZ's experiences in applying the QMS.

Variable Case Studies

This section of the paper examines the final quality of variables from the differing data quality levels, to provide examples of how the QMS was applied to different phases of the 2001 Census and to analyse how well this worked. This was achieved by assessing whether the stated quality goals and strategies were met and followed, and whether the actions to be taken for variables of each quality level were carried out. A variable from each quality level has been selected: age (level one: foremost variable), personal income (level two: defining variable) and occupation (level three: supplementary variable).

Foremost Variable: Age

Consistent with its status as a foremost variable, the age question was on the front page of both the English and bilingual individual forms. Considerable time was spent checking the quality of this data during processing and evaluation. Edits were used to look for respondent error and introduced error, and problems with age data that were identified during evaluation were generally fixed.

Sometimes it was not possible to implement the QMS as intended because certain aspects of the census system were developed that did not allow it. For

example, as specified in the QMS, enumerators checked on the doorstep that respondents had answered the questions related to the foremost variables on the front page of the questionnaire, but it was not possible to do this check when forms were put in privacy envelopes or mailed back.

For foremost variables such as age, respondents' actual responses were to be taken unless there was strong evidence to suggest they should be changed. Again, this action was partly carried out, but decisions were made to not fully implement it during the processing phase. It was decided not to allow changes to age data at certain stages, especially at the later stage of processing. The concern was that late changes may reverse earlier edits. It was also considered more useful for the evaluation stage to highlight processing issues and proposed changes through reprocessing.

Specifying quality standards for age proved to be a difficult task. The age data did not meet the quality criteria and it was felt that the standards may have been unrealistic. An example of the impractical standards was that there should be no more than 1 percent imputation for age. This is especially unreasonable if imputation at unit level is included. The result for 2001 was 3.7 percent imputation (including unit level) in comparison with 3.9 percent in 1996. The difference between the expectation and result reflects the difficulties in defining criteria for quality, in defining 'fitness for use' and the errors that make data 'unfit for use'. More work needs to be carried out for the 2006 Census on how to define and measure quality, and on what makes data unfit for use, including developing a better understanding of how the data is used.

The quality of the final age data was described in the evaluation report as 'acceptable'. This suggests that, although the age data was fit for use, its quality was somewhat lower than was desirable, particularly given that age was a foremost variable and was to be of the "highest quality possible". There was general agreement that more edits for age were needed and that the processing system needed to be designed so that operators could change age data when it was obviously wrong. But care needs to be taken that the integrity of data is preserved at all stages.

In summary, the QMS was applied for the age variable, although this example highlights the need to consider the balance between quality standards that are acceptable for users of census data and the practicalities of a census operation.

Defining Variable: Personal Income

Personal income was a second level (defining) variable. Where possible and relevant, the QMS was applied to this variable. The quality goal of having outputs that are consistent with previous censuses was met, although the removal in 2001 of edits that were used in 1996 has

affected comparisons with 1996 data. As for the age variable, it was felt that more edits on personal income would have been useful. The quality goal of having a transparent and well-documented methodology underlying the personal income output was well met. The actions to be taken for defining variables were followed for personal income.

The level of non-response (11.1 percent) for this variable, however, meant that an important quality goal – to have relevant and useful outputs – was not adequately met because the personal income data is not considered to be fit for some uses. These outputs include, for example, detailed regional analysis, population-based funding models where personal income is used as a variable, producing distributional statistics such as averages and medians, and looking at disparities at a regional and ethnic group level.

Overall, the evaluation document states that the data met most of the specified quality standards, but it is also noted that these quality statements were not very detailed. There were also conflicts between different elements of the QMS. For example, users specified personal income quality standards that contradicted the QMS instruction that some inconsistencies from respondent and introduced error could be left in the data for defining variables. This suggests that either these quality standards for personal income or some aspects of the QMS need to be reviewed or refined, so that quality standards are achievable under the QMS for future censuses.

Personal income measures traditionally have a high level of item non-response, and this was the case in 2001, where the level of item non-response still remained high. Non-response has proven to be a major issue that can result in data that is not fit for use. It would be beneficial for the QMS to place more emphasis on non-response and strategies for reducing it.

In general, the QMS was applied for the personal income variable, in that the actions to be taken for defining variables were followed. However, more attention is needed to address quality standards, in particular item non-response, in future censuses.

Supplementary Variable: Occupation

Occupation was a third level (supplementary) variable, and, in line with the QMS, it was treated as having a low priority for focussing resources and some inconsistencies were left in the data. The overall quality goals were met when examining the occupation variable, except for part of the goal relating to controlling and reducing respondent load. No action to reduce respondent load was taken for this variable, although it was controlled. It may be difficult to meet this goal for some variables as there may not always be much scope for reducing respondent load.

It was acknowledged in the QMS that responses for supplementary variables that were coded during the processing phase, such as occupation, would require more resources than a tick box or numeric response question. Occupation did receive sufficient attention as a supplementary variable throughout the processing and evaluation phases. However, it was recommended in the evaluation report that quality could be improved for the 2006 Census by including extra training on coding for operators and more testing of autocoding.

Despite this variable being assigned to the lowest quality category, the quality of the final data was still considered to be good. Therefore, application of the QMS as it applied to supplementary variables did not necessarily result in low quality data.

It is interesting that comparing the resulting data quality of this variable with that of personal income shows that it is possible for the final data quality of a supplementary variable to be higher than that of a defining variable. This is contrary to what would generally be expected under the QMS. It also demonstrates, in this instance at least, that the first requirement for quality data is that the question has been answered. If the question is not answered by a significant percentage of the population then the final data will not be of high quality and may not be fit for use, even if a lot of time and effort is put into editing, checking and fixing the data. This suggests that there should be more emphasis in the QMS on this aspect of data quality and on strategies to maintain or improve response rates or investigation into the need for imputation/modelling to estimate for non-response. In addition, the quality of the results also reflects the nature of the question, in this case, personal income is generally considered more sensitive to answer than occupation. The sensitivity of the topic needs to be taken into account when setting up quality standards.

Case Study Conclusions

Overall, the QMS was applied for the three variable case studies. The results indicated that while working within traditional census processes the application of the QMS did not necessarily lead to each variable meeting its assigned quality level. Quality standards needed to be more detailed and achievable under the QMS. Further work is required on how to define and measure quality, including determining fitness for use. There was general agreement that more edits were required. It would also be beneficial for the QMS to place a greater emphasis on strategies for maintaining or improving response rates and processes for handling missing data. There is a need to consider the balance between quality standards that are acceptable for users of census data and the practicalities of a census operation.

The QMS gave SNZ an indication of what the quality of the data was like. It helped us focus on quality and

attempt to define it, which was a big step forward for the census. The experiences of using a quality management strategy for the 2001 Census has laid a strong foundation for defining quality and fitness for use for the 2006 and future censuses.

2001 QMS Review

In addition to the case studies, an overall review of the QMS has been conducted. The findings so far indicate that the first two strategies of the QMS – differentiated levels of quality, and quality standards for outputs – were more familiar than the others. Indications have shown that the three levels of quality were easier to apply than the other strategies. The other three strategies represented good management practice and were generally implemented, but they were not always attributed back to the QMS.

Preliminary investigation as to the reasons for this suggest that a key element is the level of promotion and emphasis. In hindsight, the details of the QMS were not as well promoted as they could have been and did not provide clear enough direction on how quality should be managed in each phase and/or project within each phase.

Nevertheless, the concept of the QMS was a sound advance in methodology for 2001 that provided some useful guidelines. Quality was better understood, there were quality measures which could be discussed with users, and metadata relating to quality was provided for the first time. Further work is now required to refine and shape a quality management strategy for 2006.

Key Recommendations for the 2006 Census

A set of high level recommendations have been collated from our current review work that provide direction for managing quality in the 2006 Census. The key recommendations include using the 2001 QMS as a base to build on, and to continue with its general approach. In line with the 2006 Census strategic direction of being responsive to the needs of users, the QMS needs to involve users more, while at the same time being achievable and adaptable. The principle of prioritising variables should continue, with the need to investigate what quality levels are required and what is the most appropriate quality level for topics/variables/outputs. In addition, we need to use the experience from 2001 to improve the fitness for use specifications, including investigating how to better define and measure both quality and what makes data unfit for use. Appropriate quality standards need to be developed and as a consequence census processes may need to change to better achieve the quality standards.

Conclusion

It is proposed that the recommendations from the 2001 Census QMS review, in addition to investigations of

overseas and SNZ quality management strategies, will provide the basis for the development of a 2006 Census Quality Management Strategy. In 2001, we had a strategy that provided a starting point for managing quality throughout the census. The QMS was not fully implemented as originally envisaged, but it did provide a good guide to managing quality which had not been attempted for previous censuses. The 2001 Census met user expectations of data being available within a year of census day, and one reason for this success was the quality management approach. Risks were well managed and budgets were tightly controlled. Quality was better understood, there were quality measures which could be discussed with users, and metadata relating to quality was provided for the first time. For 2006, we aim to build on our experiences and develop a strategy that is more achievable taking into consideration appropriate quality standards and the processes to achieve them.

Acknowledgements

The authors wish to acknowledge the contribution of Len Cook, Dr Frank Nolan, Sarah Minson and Carolina Kol in the development of the 2001 Census Quality Management Strategy. The authors also wish to acknowledge the assistance of Angela Fabian in research for this paper, and the assistance of other SNZ colleagues in peer and quality reviewing the final paper.

References

- Statistics New Zealand (1998). Scope of Census, Census 2001: How Will We Shape It?.
- Statistics New Zealand (2001). Production and Presentation of Official Statistics: Strategies for Managing Quality.
- Statistics New Zealand (1999). High Level Quality Management Strategy.
- Statistics New Zealand (1999). Low Level Quality Management Strategy.

1 In this paper, 'meshblock' refers to the smallest geographic unit for which statistical data is collected and processed by SNZ.