

2018 Census External Data Quality Panel: Minutes of Meeting on 12 April 2019

Date and time	12 April 2019, 9am to 3:30pm
Location	Stats NZ BNZ Centre, 120 Hereford Street Christchurch
Present - panel members	Richard Bedford – co-chair Alison Reid – co chair Barry Milne Donna Cormack Ian Cope – via video conference Len Cook – via video audio conference Tahu Kukutai – via audio conference Thomas Lumley
Present – Stats NZ	Adele Quinn, Manager Census Analytics Carol Slappendel, Deputy Government Statistician Christine Bycroft, Principal Statistician Gareth Meech, Senior Manager Census Data Quality/ panel secretariat Kathy Connolly, General Manager Census Liz MacPherson, Government Statistician Steph Prosser, Senior Analyst Census Vince Galvin, Chief Methodologist
<p><u>Meeting minutes</u></p> <p><u>Review previous minutes and action points</u> Richard Bedford chaired the meeting. This is likely to be the last full panel meeting that he will chair. The full panel meeting in May will be chaired by newly appointed co-Chair Alison Reid.</p> <p>The panel confirmed that the minutes contained a fair summary of discussions at the March meeting.</p> <p>The Chair agreed to work through the action points at the end of the meeting to ensure most time was devoted to methods. At that time, the Chair and secretariat worked through the action point list from the March meeting, removing those that can be closed.</p> <p><u>Updated Terms of Reference</u> The Chair noted that the panel had an opportunity to review the updated version and that both Chairs were happy to endorse version 2.0 of the External Data Quality Panel Terms of Reference. The substantive changes focus on the extension of the panel sitting period and the recognition of a co-Chair arrangement. This version will be published on the website on, or soon after, 29 April 2019.</p> <p><u>Report back from 1 April in-committee meeting</u> At their panel only meeting on 1 April, members focused on the internal papers they were preparing on specific issues, and discussed whether they had enough information from Stats NZ to give more formal advice relating to methods. The following papers have been sent to Stats NZ. These discuss issues or concerns where the panel needs additional clarification:</p> <ul style="list-style-type: none"> - “Iwi data in 2018 census” received by Stats NZ on 22 March 2019 - “Census 2018 - issues to consider in assessing data quality and inter-censal comparability” received 5 April 2019 - “Use of administrative data in Census 2018” received 4 April 2019 - “What is the IDI” received 8 April 2019 	

Stats NZ thanked the panel for producing the documents and reiterated that they found them useful. Kathy noted that the presentation that had been prepared for this meeting would try to address some of the significant issues raised. Stats NZ will produce a written response back to the panel for each paper in due course.

The panel was not able to further its assessment of the overall methodology that Stats NZ was employing to improve coverage of the census, due to the absence of documentation. Once this documentation is available, the panel will need time to fully reflect on the methodology that had been employed.

The Chair also noted that there had been a lot of negative media articles about the 2018 census in the previous few weeks, and that members are not able to publicly comment on issues being raised. This was a source of frustration for some members of the panel.

The panel are supportive of Stats NZ being as transparent as possible in its announcements on 29 April.

Methods and data quality session

Christine presented slides that provided a refresher on the completed methods work to date. As the slides were only loaded to the shared workspace on the morning of the meeting, panel members had not had a chance to see them prior to the meeting.

At several stages of the presentation, panel members noted discomfort in Stats NZ using the term “admin enumeration” as it promotes a sense of enumeration, when none existed before census collection. Stats NZ is actually carefully selecting additional people from lists. The panel noted the risks around inventing language ‘after the fact’. Stats NZ acknowledged that their language around this is developing. Some international sources use the same term, but the method utilised here is not exactly the same as overseas. Some of the methodological approaches taken by Stats NZ have not been done in other countries.

It was also suggested by a panel member that Stats NZ needs to employ statistical terminology and language, rather than terms that imply value judgements.

The panel suggested changing the language ‘Census undercount’ into something like ‘indicative coverage gap’. Undercount suggests it has been measured by the PES.

The panel also suggested that the term ‘signs of life’ should be amended to read something like ‘signs of activity’.

The panel noted that Stats NZ needs to be very clear that they are using real data from admin sources and that it is not imputing data for individuals. It also needs to be very clear that reference is being made to individuals and not households.

When discussing the IDI spine, the panel noted that it is important for StatsNZ to be clear if the spine has names attached.

AP7-1 Need to be clear what the IDI spine is, and produce a document collating date ranges for each spine data source used in the 2018 Census.

The panel questioned how Stats NZ removed duplicates in 2013 Census. The response was that there was no process for this in place in 2013. Methods are much more rigorous in that area in 2018.

During the discussion about admin households, it was noted that the methods quality assurance paper provided to the panel prior to the meeting contains reference to making an arbitrary judgement about where the line was drawn when to including or excluding people in the model. Stats NZ responded that there is a method behind this and it can be better explained.

During the presentation on ethnic distribution by source group, the panel asked for similar analysis of NZ European and Asian level 1 ethnicity alongside the Māori and Pacific peoples analysis provided.

AP7-2 More analysis of ethnic distribution for NZ European and Asian ethnic groups and the percentages of these that contribute to the Māori descent population.

Advice was given by the panel to avoid saying that the 2018 Census data is 'better than' 2013 without evidence – Stats NZ needs to be cautious about such comparisons in the upcoming 29 April media release.

With reference to the slide about minimum accuracy targets, the panel raised the question around what Stats NZ's customers will be more interested in - levels of change or actual counts. Users expect consistency – but a balance needs to be struck. It was noted that the denominator for many significant rates will have changed.

Reference was made to the variable warrant of fitness document for some key variables that the panel wants to see at the May meeting.

Stats NZ presented summaries of sensitivity analysis work completed to date.

The panel suggested that Stats NZ either run a technical seminar for DHBs or ensure they have been invited to the current seminar series.

The panel was interested in reading the Dot Loves Data report about electoral count and boundary sensitivity.

AP7-3 StatsNZ to load DLD report to shared workspace.

StatsNZ presented proposed changes to iwi count methodology to create more accurate historic census counts on the 2018 classification. The panel noted that if the method is refuted, it could undermine the 2013 counts and that new iwi will not have been fairly represented given the list effect in the questionnaire.

There was not clear support expressed by the panel for the changes.

There was some discussion about the use of data from prisons. Tahu and Donna expressed serious concerns about the ethics of obtaining identifiable prisoner data from Corrections without prisoner consent and strongly advised Stats NZ against doing so.

The panel asked whether Māori had been involved in the decision making. They also expressed the view that Stats NZ needs to be unequivocal in its April 29 media release that the 2018 Census will not produce high quality iwi data.

Programme critical path and April key messages

The panel would like more definite timeframes about the drafting of method papers, so that they can take this into account when finalising their meeting schedule and work plan through to September.

AP7-4 Stats to update method paper timeline spreadsheet.

Regarding the key messages for the 29 April announcement, the panel provided the following advice:

- Use terms such as 'indicative coverage estimate' rather than 'undercount'.
- Be clear that Stats NZ are not talking about the official PES measure, but an 'indicative coverage measure'
- Tell the public that some of the more detailed results might not be of sufficient quality to be termed "official statistics"

- Say that a lot of the methods work is brand-new and that while it had not been anticipated that some of the changes in approach would have to be used for the 2018 census, there are some advantages in implementing these approaches.
- There will not be a census in 2021
- 'Fit for purpose' can be problematic – perhaps talk about 'known quality'
- Note that StatsNZ has done more work to assess the quality of census data than ever before
- StatsNZ is in a much better place now to start development of the next census
- Can we be up-front about names and addresses and how we do the linkage?
- Make sure StatsNZ documents why we think we can use admin data – what are the facts about this?
- Need to be clear about what anonymisation is in this context.
- If talking about Māori and iwi data it would be best to have a te reo version of the media release so it can be used quickly by Māori language media.
- Whether StatsNZ should publish census night counts at the same time as usual residence.
- Suggested StatsNZ should be confident about what the quality of the data is and is not.
- Shouldn't underplay problems with variables such as families and households.
- Acknowledge that because this methodological work was not planned, there was not an opportunity to have an open and honest conversation with the public and Māori stakeholders. Need to be more upfront about what this means for Māori data governance going forward

Legal basis and ethics session

Carol shared StatsNZ interactions with the Office of the Privacy Commissioner. They have reviewed all three PIAs and provided feedback. The OPC is comfortable with Stats NZ use of administrative data in the census from a privacy perspective.

The Privacy Act allows for data to be re-purposed for statistical and research purposes as long as the outputs are confidentialised, which StatsNZ always ensures.

StatsNZ asked the panel about their degree of confidence in the method. After discussion, the Chair said they would discuss at the in-committee meeting on 15 April. There may not be a general statement from the panel in April.

The next meetings dates are:

- in-committee only 15 April, Auckland
- in-committee only 8 May, Auckland
- full meeting 23 May in Auckland
- full meeting 24 June in Wellington (to fit in with PANZ conference on 20 and 21 June).
- July TBC in May

Action log

Ref	Date raised	Description	Owner	Date required	Progress	Status
Meeting 2 actions						
AP 2-4	23/10/18	More information requested about what investigations we are doing into 2013 counts and response rates – are we assuming that 2013 is correct?	Panel	By Dec meeting	9/11 Information presented that include 2013 response rates and counts. Gareth to check with Richard whether this should be closed by 23 Nov. 6/12 Gareth to discuss with Richard	Closed

					18/2 Stats may need a specific 2013 methods paper to document any inconsistencies in approach and will be captured in methodology papers and metadata documentation. Gareth to upload proposed method papers (completed 18 Feb). Panel to review paper and offer comments by 6 March. 11/4 Summary spreadsheet of proposed papers to be published uploaded on 10 April. Close	
AP 2-6	23/10/18	Need more detail about the totality and components of the imputation methods initially planned and now being implemented.	Gareth	By 9 Nov	Sessions at 30 Oct meeting and priority at full 9 Nov meeting. 9/11 Information presented at meeting. Gareth to check with Richard whether this should be closed by 23 Nov. 6/12 Gareth to discuss with Richard 18/2 Spreadsheet detailing variables and imputation processes to be sent to panel prior to 6 March meeting. 6/3 Methods paper and variable source spreadsheet sent to panel. Confirm with Richard whether this item can be closed before 12 April in committee meeting. 11/4 Panel still requires fuller methods paper which is planned to be written. Close	Closed
AP 2-16	23/10/18	Re electoral implications, more information is required about the representativeness of	Kathy and Gareth	By 9 Nov	Discussion planned for 9 Nov meeting. 9/11 No additional electoral data was presented (as not	Closed

		the data and impacts of imputations. A range of scenarios, including constitutional impacts for Māori, would be useful			complete yet). Updated draft electoral calculations to be presented to panel in Dec meeting. 6/12 Electoral data has not yet been run – targeting Feb meeting update. 18/2 Updated electoral data to be presented to 6 March meeting 6/3 Data is not ready for presentation to the panel. To be provided to the panel before or at the 12 April meeting. Note that Dot loves Data has been contracted to look at the sensitivity of the detailed electoral calculations. 11/4 Slides included in April meeting presentation. Close.	
Meeting 3 actions						
AP 3-10	9/11/18	Clarification is needed about what data is used in the census administration enumeration (referred to as IDI data in the presentation), and a list of the current variables being imputed, where it comes from, whether the data suppliers are aware of its use and whether it uses identifiable or de-identified person data.	Panel members	By 7 Dec	6/12 Leave open, not addressed yet 14/2 PIA includes detail, but hard to follow. Suggest adding a flowchart to aid communication. Panel to review the PIA and give feedback. 6/3 One panel member had provided feedback. Carol stated that the PIAs on the website require further revision work has started on this. Has worked with legal counsel on creating a generic statement about the legal basis of the use of admin data. A statement will go into an appendix before the (likely end) April announcement. Leave open.	Closed

					11/4 Legal basis document added to shared space. Close.	
AP 3-11	9/11/18	Present criteria about the thresholds being used to make decisions about methods.	Gareth and Christine	By 7 Dec	6/12 Some information in Dec meeting present, more info to come after meeting. Leave open. 14/2 Panel will be using Canadian quality framework. Request more information about quality plans at March meeting. 6/3 An update about the quality measures was presented at 6 March meeting. Variable quality rating scale paper to be sent to panel before 12 April meeting. 11/4 Variable rating scale paper loaded to workspace 10 April. Close and add to May meeting agenda.	Closed
Meeting 4 actions						
AP 4-6	7/12/18	Document 'What is the legal basis that it is OK to do what we are doing with admin data and what are the ethical implications we have considered?' 'What commitment has Stats NZ given to respondents and data suppliers about what we are doing with admin data? What have they said? What are the key data sources?'	Carol	By 14 Feb meeting	14/2 Although Liz discussed the issues with the panel at meeting on 5 Feb, panel still wants a statement from Stats about the legal basis of admin sources. Without, it is a risk for the panel to endorse. High ethical standard set by the Stats NZ at the 2017 data summit. 6/3 This matter continues to be a source of concern to panel and Stats is addressing some of these concerns in the variable list (see AP 3-10) 11/4 On agenda for April meeting. Some of the panel does not agree that the Stats Act is not	Closed

					prescriptive. Stats will provide a brief written response on the paper. Close.	
AP 4-8	7/12/18	Consult over when Stats NZ will return to the panel with a decision about public release of iwi data	Richard and Gareth	By 14 Feb meeting	14/2 Stats to cover in meeting on 14 Feb. Check with Richard if enough information to close. 6/3 iwi data leaders meeting key messages sent to panel. Decision made that Stats will not talk publicly about data quality before the April announcement. However, need to have conversation with iwi before then. Leave open until iwi data messaging is confirmed. Panel is writing up document to be sent to Stats about panel view about quality of iwi data. The panel noted that the 2018 Census webpage has a link to information about iwi that implies it will be available. Stats to review and update. 11/4 On agenda for April meeting. Close.	Closed
AP 4-9	7/12/18	Update panel once more detail has been completed on the quality framework including an ordering of decision making criteria.	Gareth	By 14 Feb meeting	14/2 Planned item on 6 March agenda. 6/3 Variable quality rating scale paper to be sent to panel before 12 April. 11/4 Paper loaded, set time in May agenda. Leave open.	Open
Meeting 5 actions						
AP 5-1	14/2/19	Need to clarify admin enumeration, admin imputation and other terminology. Thomas created a short paper of definitions to be reviewed by Stats NZ.	Adele, Christine	By 6 March	6/3 Thomas has written a paper for the panel about definitions of admin vs survey data. Kathy to send to SNZ staff for review. 11/4 Written response to be developed by	Closed

					Stats for each paper written by panel. Date to be confirmed. Close	
AP 5-6	14/2/19	Revise terms of reference mainly around the mechanics of the panel purpose, then for the panel to review and sign-off from Liz and Richard	Richard	By 6 March	6/3 Sent to Richard for initial review. Richard to review by 1 April. 11/4 Richard and Alison reviewed and spoke with Carol and Liz. To be confirmed at April meeting, effective from 29 April. Close.	Closed
AP5-8	14/2/19	Include 'what is a response' paper in the list of method papers that would assist the panel to write the reports	Christine	By 10 April	6/3 No update 11/4 Paper uploaded to shared space. Close.	Closed
Meeting 6 actions						
AP6-1	6/3/19	Provide the panel information on smoking question data as an example of a variable without other admin enumeration sources.	Adele	By 12 April	11/4 Unable to provide counts of lower priority variables for April meeting. Suggest smoking data and associated quality rating topic for May meeting.	Open
AP6-2	6/3/19	Present list of papers and approx. completion (or draft) timing so panel can base own reports off Stats documents.	Gareth	By 12 April	11/4 Updated list loaded to shared space however panel requires more specific dates quickly to know when to book upcoming meetings. Leave open.	Open
AP6-3	6/3/19	Add numbers on 'address not found' to the table at the end of the dwelling frame paper	Chris H	By 12 April	11/4 Additions were made, the new paper was uploaded. Close.	Closed
AP6-4	6/3/19	Upload variable quality rating scale paper to workspace before 12 April.	Gareth	By 12 April	11/4 Uploaded to space. Close.	Closed
AP6-5	6/3/19	Upload ethnic mobility paper to workspace	Gareth	By 12 April	11/4 Original uploaded on 29 March. Updated version 2 uploaded on 10 April. Close.	Closed
Meeting 7 actions						
AP7-1	12/4/19	Produce document about what the IDI spine is and is not. Include date ranges for	TBC	By 22 May		Open

		each spine data source used in the 2018 Census.				
AP7-2	12/4/19	Analysis of ethnic distribution for NZ European and Asian ethnic groups and the percentage of each source that contributes to Māori descent.	Adele	By 22 May		Open
AP7-3	12/4/19	Load Dot Loves Data electoral paper to shared workspace.	Gareth	By 29 April	12/4 Completed	Close
AP7-4	12/4/19	Update method paper timeline spreadsheet.	Gareth	By 29 April		Open

Released under the Official Information Act