Report of the 2023 Census External Data Quality Panel

An assessment of Stats NZ's methodological approach to producing the 2023 Census file (updated)

Barry Milne

Thomas Lumley

Andrew Sporle

October 2024 (updated November 2024)

Contents

Executive Summary

The interim response rate for the 2023 Census was 88.3%, necessitating the use of alternative data sources to complete the census file – as had happened for the 2018 Census. This report of the 2023 Census External Data Quality Panel assesses Stats NZ's methods for completing the 2023 Census file. The 2023 Census used a 'combined model' by design, which combined census responses with alternative data sources. To a large extent, the methods used by Stats NZ for 2023 Census are the same as those used in the 2018 Census, though some methodological changes have been made. Where there have been changes, this report considers how these changes are likely to affect data quality. The assessment focuses on 'Priority 1' individual variables as well as the 'Activity limitations' (Priority 3) variable, which has been examined because of the importance of this variable to the disabled community in Aotearoa New Zealand. The variables assessed are:

- Usually resident population count (total)
- Usually resident population count (by territorial authority and Auckland local board)
- Māori descent (output)
- Māori descent (electoral)
- Level 1 ethnicity
- Age
- Gender
- Iwi affiliation
- Activity limitations.

We give ratings for variables using the same five-level scale (very high, high, moderate, poor, very poor) used by Stats NZ in their data quality assurance process.

Methods used in the 2023 Census

Stats NZ augmented the census file with admin enumerations – census records that do not correspond to a returned census form but do correspond to an actual person resident in New Zealand. This method is a repeat of the method used in 2018, and involved linkage to the Integrated Data Infrastructure (IDI) to determine which admin records to enumerate. Linkage quality was slightly better in 2023 compared with 2018, which may result in a slight improvement in 2023 data.

A new method used in 2023 was the use of the XGBoost prediction algorithm to identify usual residence address for admin enumerations. Testing indicates that this algorithm improves upon the approach used in 2018 so should result in better address determination.

Two new issues were identified regarding Stats NZ's methods for the 2023 Census. First, 2023 Census had a very high rate of duplicate census forms – mostly involving people being listed on more than one household/dwelling form. The process to remove duplicates was designed to have a low risk of removing genuinely distinct individuals, and while 538,000 duplicate records were removed, it is estimated that only about 3,000 duplicate forms remain in the census file. This suggests that duplicates are not a major issue for data quality.

Second, as for 2018 Census, some admin enumerations had insufficient evidence to place them into a household, but sufficient evidence to locate their meshblock of residence. This evidence might involve evidence of recent residency – that is, people who used to live in a meshblock but now live somewhere

else. In such a situation, the meshblock assignment would be incorrect, and would distort very small-scale population estimation.

Alternative data sources were used when data were missing for some variables (item non-response). For the variables assessed by the Panel, the approach taken for 2023 was largely similar to the approach taken for 2018, though there were several minor changes: (i) use of two historic censuses: 2018 and 2013; (ii) use of electoral roll data for the Māori descent variables; (iii) use of parental data for the Māori descent variables; (iv) 'Don't know' values were able to be imputed for the Māori descent (output) variable; and (v) use of alternative data sources, including historic census data, admin data, and parental data, for the 'Iwi affiliation' variable. In addition, because data on gender were collected for the first time in 2023 Census (only data on sex had been collected previously), a process was developed for item non-response for gender. This process involved sourcing data from the Ministry of Social Development (MSD) where this was available and using imputation otherwise.

Assessment of variables

- *Usually resident population count (total and by territorial authority and Auckland local board).* The count of the population is likely to be 'very high' quality, both for the overall population and for territorial authorities and Auckland local boards. The quality of methods for using admin data to count large geographies (such as territorial authorities) should ensure that the overall population counts for cyclone-affected areas are accurate.
- *Māori descent (output).* The Māori descent (output) variable can be considered 'very high' quality. The addition of new supplementary data sources has improved the coverage for the Māori descent variable, especially for those with Māori descent. However, the change in methods has resulted in yet another break in the Māori descent time series, with the most recent three censuses using different methods.
- *Māori descent (electoral).* The Māori descent (electoral) variable can also be considered 'very high' quality. This variable used the same supplementary data sources as the Māori descent (output) variable, and is likely to be of similar quality.
- *Level 1 ethnicity.* As was the case in 2018, the quality of Level 1 ethnicity data varies across ethnicity: the quality can be considered 'very high' for European, Asian, and MELAA groups, but 'high' for Māori and Pacific groups. The quality in the latter two groups is likely to be variable by age, with ethnicity data being of lesser quality in younger age groups.
- *Age*. As largely the same methods were used to source Age data in 2023 and 2018, with less reliance on alternative data sources in 2023, the Age variable in 2023 should again be rated 'very high' quality.
- *Gender*. Counts for males and females are closely in line with expectations and can be considered to be 'very high' quality. However, there are too many unknowns to judge whether the count for 'Another gender' is accurate.
- *Iwi affiliation.* Iwi affiliation should be considered 'moderate' quality. The methodology used by Stats NZ to complete the iwi affiliation variable greatly increased coverage, but coverage will be lower for older ages.
- *Activity limitations.* Data are unavailable for 15.5 to 15.6% of those aged 5 years and older across the six 'Activity limitations' questions. As such, the data should be considered 'poor' quality.

Summary and conclusions

Stats NZ took sensible and justifiable steps to implement methods to make the 2023 Census file more complete and of as high quality as it could be. The panel endorses the methods used by Stats NZ to fill gaps in the 2023 Census, which were largely the same as the methods undertaken to fill gaps in the 2018 Census. These methods are now well-established and work well.

While the methods implemented by Stats NZ improved the quality of 2023 Census data, the variables assessed by the panel were of varying quality, as described above. Moreover, these methods are not perfect substitutes for census responses. The use of alternative data sources works well for some purposes, such as counting the population, but less well for (i) populations with lower response rates – 2023 Census, like 2018 Census, achieved lower response rates for Māori and Pacific populations; (ii) populations with lower quality alternative data sources, such as 'Iwi affiliation'; and (iii) variables without alternative data sources, such as 'Activity limitations'.

As such, if a fully administrative census were to be implemented, it will likely be of lower quality than a full enumeration census. Well-designed surveys may address some of these issues but will struggle to capture small populations and small areas well. For this reason, the panel believes that Stats NZ needs to give priority to finding ways to collect data directly where alternative sources are not available or are clearly of lower quality. A full enumeration census would achieve this, but may be difficult to attain.

1 – Introduction

In the 2018 Census, the unexpectedly low response rate led Stats NZ to develop methods to augment the census data with administrative data and information from the 2013 Census ('alternative data sources'). These approaches were examined in great detail by the 2018 Census External Data Quality Panel, who produced reports on the methodology and on the resulting data quality of most of the Census variables (Census External Data Quality Panel, 2019a, 2019b, 2019c, 2020). The 2023 Census had an interim response rate of 88.3% (Jack & Bowlby, 2024). This time there were plans in place to use alternative data sources to complete the census file. Stats NZ learned from the lessons from the 2018 Census and implemented a 'combined model' by design, which combined census responses with alternative data sources (Using a combined census model for the 2023 Census).

The 2023 Census External Data Quality Panel was established to assess Stats NZ's methods for completing the 2023 Census file. The panel (two of whom were members of the 2018 External Data Quality Panel) has been able to rely heavily on the 2018 reports. To a large extent, the methods used by Stats NZ in this census are the same as those used in the 2018 Census, though some methodological changes have been made. Where there have been changes in methods we have considered how these changes are likely to affect data quality. We have also considered whether there were issues that the 2018 panel missed, but we have not sought to redo the work of the 2018 panel from the beginning.

This report is also different from the 2018 Census data quality report in that our terms of reference are restricted to the Priority 1 individual variables including the two versions of the Māori descent variable, as well as one Priority 3 individual variable, 'Activity limitations' (Stats NZ, 2021a). The panel specifically requested to examine the quality of 'Activity limitations' because of the importance of this variable to the disabled community in Aotearoa New Zealand. We have not reviewed other Priority 2 or Priority 3 variables systematically, though we will sometimes comment on them in passing. The variables we assess are:

- Usually resident population count (total)
- Usually resident population count (by territorial authority and Auckland local board)
- Māori descent (output)
- Māori descent (electoral)
- Level 1 ethnicity
- Age
- Gender
- Iwi affiliation
- Activity limitations.

In section 2 we describe the methods by Stats NZ for completing the 2023 Census file, with a particular focus on methods that deviate from or have been added to the methods employed for the 2018 Census.

In section 3, we assess the variables listed above. We give ratings for variables using the same five-level scale (very high, high, moderate, poor, very poor) used by Stats NZ in their data quality assurance process (Stats NZ, 2024a). We take account of a number of factors when assessing quality, including Stats NZ's three assessment metrics: data sources and coverage, consistency and coherence, and accuracy of responses (Stats NZ, 2024a); the (very similar) criteria used by the 2018 Census External

Data Quality Panel: coverage, consistency, comparability and contemporaneity (Census External Data Quality Panel, 2019a); and whatever other factors we consider relevant.

Section 4 summarises our findings and offers some concluding remarks.

2 – Methods used in the 2023 Census

2.1. Overview

The Integrated Data Infrastructure (IDI) combines government data on individuals (e.g., use of the health system, receipt of a benefit, school attendance) into a linked database (Stats NZ, 2022a). Everyone who has entered New Zealand, who was born in New Zealand, and who has had taxes paid or withheld will be included in the IDI, forming an 'ever resident' population (Milne et al., 2019; Statistics New Zealand, 2019). The IDI includes essentially everyone who currently lives in New Zealand, plus many others. Filtering out those who are known to have left the country or died and those who have had no recent contact with the system gives an estimated resident population (IDI-ERP). The IDI-ERP provides an upper bound on who lives in New Zealand, and the census provides a lower bound.

In constructing the census file, Stats NZ uses IDI-ERP records to create admin enumerations, to narrow the gap between the IDI-ERP and the census. Admin enumerations are census records that do not correspond to a returned census form but do, with sufficiently high confidence, correspond to an actual person resident in New Zealand. The Stats NZ report, *Methodology for using admin data to count people in the 2023 Census* (Stats NZ, 2024f) describes the guiding principles and the steps that are taken to add records, and notes where the steps have changed for the 2023 Census. Briefly, the first step is to add people to households for which there are census forms if there is strong evidence that they belong there. Second, households are created for non-responding addresses from the people who are predicted to live at that address, given some checks that the set of people forms a plausible household. Finally, some people are assigned only a meshblock – a small geographical area containing 30-60 households – rather than a specific address, if they have a high predicted probability of living in that meshblock but cannot be reliably assigned to a household and address in the meshblock.

2.2. Linking to the IDI

Linking census records to the IDI serves two purposes. First, those not linked form the pool of records that potentially can be used for admin enumerations – that is, records that were missed by the census but should be included in the usually resident population (11% of records in 2023, (Stats NZ, 2024f)). Second, it enables incomplete census records – those with missing data for some variables – to be completed using historical census data or administrative data available in the IDI. Linking census records to the IDI in 2023 followed largely the same methods as those used in 2018, and the statistical appropriateness of these methods (if not the social and cultural license to undertake them) was endorsed by the 2018 Census External Data Quality Panel (Census External Data Quality Panel, 2019c).

Linking is undertaken in 6 'passes' which use different blocking (stratifying) and linking variables, with less strict matching requirements in later passes, to identify likely 'true' matches (Stats NZ, 2024e). There were minor changes in the linking methodology used in 2023, including:

- *Linking of absentees.* An absentee is a person usually living at an address for which a census household or dwelling form exists but was not present at that address on census night. Absentees were not linked to admin records in 2018 but were in 2023. Absentees were first linked to individual forms to see if they have responded elsewhere. If they had not, and if the absentee record was successfully linked to the IDI and determined to be in New Zealand on census night, they were considered admin enumerations. This new methodology should result in

slightly better data because admin enumeration absentees will be placed in a household or dwelling, which may not have happened if linking was not undertaken for absentees.

- *Using both sex at birth and gender for linking.* For the first time, 2023 Census collected data both on 'gender' and 'sex at birth' on the form completed by individuals. Previously, only data on 'sex' had been collected, and this variable was used for linking. However, *partial responses* in 2023 – those listed on a household or dwelling form but for whom there was no accompanying individual form – only had data recorded for gender. While sex at birth tends to be a more reliable linking variable (because data on sex at birth are more commonly recorded on administrative records), for linking census data to the IDI in 2023 both sex at birth and gender were used, though in different passes ((Stats NZ, 2024e), p 16-17). This enabled linkage for partial responses that otherwise would not have occurred.
- *Initials and last two letters of first names used as linking variables.* To avoid errors with short names (e.g., "Lin") being linked to longer names of which they are a subset (e.g., "Linda"), initials of the first names and last two letters of the first name were used as additional linking variables in 2023. The impact of this change is unclear (Stats NZ, 2024g, p 19), though it likely affects a small subset of potential links.
- *IDI person-address records that linked in one pass included in subsequent passes.* Census-IDI linking is one-to-one at the person-level but makes use of IDI address records which are one-to-many (i.e. one person may have several address records – on average each person has five). In 2018, if a person-address link was made, this record was removed from subsequent passes, whereas in 2023, the person-address that was linked was also available for linkage in subsequent passes. This allows for more linkage records and potentially more linkage possibilities – ultimately, the link with the highest weight or earliest pass is kept. This resulted in improvements in "approximately a few thousand records" (Stats NZ, 2024g, p 20).

Overall the new methodologies are likely to have resulted in small improvements in linkage quality compared with 2018. Linkage was achieved for 97.9% of records (98.2% excluding absentees), compared to 97.7% (excluding absentees) in 2018, and linkage improved from 2018 to 2023 for males and females across all ages and for all ethnic groups (Statistics New Zealand, 2019; Stats NZ, 2024e).

2.3. Address determination

A census record comes with an address, even though this address is removed before the data are made available in the IDI. There are many people in the IDI-ERP database who have more than one listed address; they may have moved, or they may have given a relative's address for some official data collection, or the data may simply be wrong. On top of this, the addresses for people in the same household may not all be equally well represented in the administrative data; different data sources collect addresses at different times and with different reliability. The admin enumeration process must pick a single best address from those available and must try to pick the same address for all the people in a household.

In 2018, the choice was based on an explicit set of rules. In 2023 the initial choice is based on a machine-learning model and it is then checked for consistency with other records assigned the same address. The machine-learning model is described in *Predicting usual residence address from admin data in the 2023 Census* (Stats NZ, 2024g). It uses the XGBoost algorithm to estimate a probability of being correct for each address, based on information such as the source of the address, the recency and consistency of

the report, the number of addresses reported for that person, the person's age, and whether the address actually contains a dwelling. XGBoost is a highly-regarded prediction tool, and one that has performed well in competitive evaluations of prediction accuracy in many applications. The algorithm is trained on addresses from census forms and tested on both census forms and addresses from the Household Labour Force Survey. In the available testing data, where the truth is known, most people have a single most-likely address, this address has a very high predicted probability, and the address is correct. In records from people who did not return a Census form there are (unsurprisingly) somewhat more people who do not have a high-confidence address; this is especially true for young adults. However, the testing shows that the model is more accurate than the previous approach, and that this accuracy improvement holds for all ages and for all Level 1 ethnicities (Stats NZ, 2024g). A similar XGBoost model was used to evaluate whether a proposed household generated from administrative data was likely to be a genuine household living at the same address.

The changes in address determination should improve the admin enumerations and increase the accuracy of all variables compared to the 2018 approach.

2.4. New issue: duplicate census forms

The 2023 Census had an unusually high rate of duplicate census forms, partly because of the efforts to increase participation and perhaps partly because of people relocating after flood and cyclone damage to their homes. In total, 538,000 duplicate records were removed. This includes scenarios where multiple individual forms were completed for a person, as well as scenarios where an individual was listed on multiple dwelling forms. The de-duplication process was designed to have a low risk of removing genuinely distinct forms, but it had to be performed for more records than expected, in a constrained time period. By manually checking a stratified random sample of records, Stats NZ estimates that about 3,000 duplicate forms remain in the census file (information provided to the Panel by Stats NZ). This estimate of remaining duplicates is likely to be qualitatively reliable and suggests that duplicates are not a major issue for data quality.

2.5. New issue: what counts as accurate meshblock enumeration?

Admin enumerations into a meshblock rather than an address occur when there is a highly predicted probability that an individual lives in a particular meshblock, but there is not a sufficiently high probability that they are part of a household at any specific address in that meshblock. One explanation for this is that the individual does live in that meshblock, but their address data are of lower quality or consistency than average; they have genuinely lived at more than one address in the meshblock, or they have recorded more than one address in the meshblock because of data errors. If this explanation is true, meshblock enumeration is the best solution; the individual really lives in the meshblock, but their actual address and household membership is not discernable.

Another possibility is that some meshblock enumerations are people who used to live in that meshblock and now live somewhere else. If their previous address is now occupied by someone else with reliable data, the address will be unavailable to assign to them, and they may not have lived in their new address long enough to generate reliable evidence for the new address. Under this interpretation the assigned meshblock would be incorrect *for census night* but would have been correct fairly recently. The meshblock assignment would be incorrect, and would distort very small-scale population estimation, but would correctly indicate the characteristics of the area.

We have no way to determine the extent to which one or other of these interpretations (or some other explanation) is true for meshblock enumerations, but they do suggest caution with population estimation at very small geographical scales.

2.6. Alternative data sources

Alternative data sources were used when data were missing for some variables (item non-response). The approach taken for 2023 was largely similar to the approach taken for 2018: item non-response from 2023 Census was filled in, first, from data for the same individual from either historical (previous) census data or administrative data, and, second, imputation was used only when there was no suitable source from either historical census data or administrative data (Stats NZ, 2024d). Changes from 2018 included:

- Two historical censuses were available: 2018 and 2013. Use of 2013 Census data was only considered when data from 2018 Census was unavailable.
- The use of electoral roll data for the Māori descent variables. Linking to the electoral roll was undertaken independently of linking to administrative records in the IDI, and data were made available through a data sharing agreement with the Electoral Commission. Electoral roll data were fourth in the priority rankings for the Māori descent variables.
- The use of parental data for the Māori descent variables, where Department of Internal Affairs (DIA) birth registrations are used to identify parents (and grandparents). Māori descent values from parents' 2023 Census, historical census, or admin data can then be used for their children. Parental data were fifth in the priority ranking for Māori descent variables. Grandparent data were sixth in the priority ranking.
- 'Don't know' values were able to be imputed for the Māori descent (output) variable.
- The use of individual data from alternative data sources, including historical census data, admin data, and parental data for the 'Iwi' variable. As with Māori descent, DIA birth records are used to identify parents (and grandparents and great-grandparents). The iwi affiliation from parents' 2023 Census, historical census, or admin data can then be used for their children.
- Gender, measured for the first time in 2023, sourced data from the Ministry of Social. Development (MSD) when it was not available from 2023 Census data, and used statistical imputation when neither census nor MSD data were available (Stats NZ, 2022b, 2024d).

2.7. Within-household donor information.

For ethnicity, if data were not available from 2023 Census, historical censuses, or admin data, ethnicity was copied from the person in the household closest in age to the census individual. The within-household donor process was used for ethnicity for the 2018 Census.

2.8. Statistical imputation.

Stats NZ used the CANadian Census Edit and Imputation System (CANCEIS) system for the 2023 Census:

"The CANCEIS system searches records that are near neighbours on a master list to find potential donors who are good matches on a set of specified matching variables. Once the good matches are found, the system … choose[s] the closest match." (Census External Data Quality Panel, 2019c, p 21).

CANCEIS has been used previously for a number of variables for 2018 Census and is a well-established imputation approach.

3 – Assessment of variables

3.1. Usually resident population count (total and by territorial authority and Auckland local board)

The count of the population is likely to be accurate and 'very high' quality, both for the overall population and for subpopulations including territorial authorities and Auckland local boards. Eighty-nine percent of the usually resident population count were derived from census responses and 11% from admin enumerations (Stats NZ, 2024f). The panel is of the opinion that the methods used to count individuals in the usually resident population from admin data are likely to be sound – as was concluded when similar methods were used for the 2018 Census (Census External Data Quality Panel, 2019c). We note that census data collection was heavily impacted by tropical Cyclone Gabrielle for some territorial authorities, and these areas relied disproportionately on admin enumerations (Stats NZ, 2024f). However, the quality of methods for using admin data to count large geographies (such as territorial authorities) should ensure that the overall population counts for cyclone-affected areas are accurate.

As noted above in section 2.5, the possible inaccuracy of meshblock enumerations leads us to caution against estimating population at very small geographical scales, such as meshblocks, but population estimation of territorial authority and local board areas are likely to be accurate.

3.2. Māori descent (output)

The Māori descent question in the 2023 Census is used to determine two variables: Māori descent (output) – discussed in this section; and Māori descent (electoral) – discussed in section 3.3. The Māori descent (output) variable has three valid responses: 'Māori descent', 'No Māori descent' and 'Don't know'. This variable is used to determine the Māori descent census usually resident population, and "to determine the subject population for the 'Iwi affiliation' variable" (Stats NZ, 2024b).

The 2023 Census included a change in the extent of alternative data sources used for Māori descent (output). On the one hand this has resulted in improved coverage, but on the other hand has resulted in another break in the time series for Māori descent data.

In total, 85.3% of Māori descent (output) data were obtained from 2023 Census responses (Stats NZ, 2024b). When Māori descent was not reported on a 2023 Census form, data for Māori descent (output) were obtained in the following priority order (Stats NZ, 2024b):

1. 2018 Census (6.5%)
2. 2013 Census (2.6%)
3. DIA birth registrations (2.9%)
4. Electoral roll data (0.6%)
5. Parental data from census and admin data sources (0.1%)
6. Grandparent data from census and admin data sources (<0.1%)
7. 'Deterministic derivation' through an iwi affiliation record in some data source (0.1%)
8. Statistical imputation (2.0%).

Of these eight sources, 'Don't know' responses were able to be obtained from sources 1–3, 5, 6, and 8. Stats NZ had planned to source data from great-grandparents (between step 7 and 8), but no data from great-grandparents were available among the small fraction (2.1%) whose Māori-descent (output) had not already been determined by step 7.

The use of additional data sources has improved the coverage for the Māori descent (output) variable, resulting in a decreased use of imputation to replace missing data (only 2%). The availability of another historical census file (2018 Census) is useful. Two point six (2.6) percent of data for the Māori descent (output) variable was derived from the 2013 Census, which is now a decade old. Older data sources cannot account for increases in genealogical knowledge or willingness to report Māori descent during the intervening time, so are likely to be a source of potential undercount of those reporting Māori descent. However, it is worth noting that 2023 Census and 2013 Census responses matched for 93% of individuals who reported Māori descent at both censuses (Stats NZ, 2024b).

The introduction of new data sources of reported descent improves coverage and quality as the sources used for completion of the Māori descent (output) variable are now all sources where descent is reported by the individual or by a parent or grandparent. Self-reported descent from the electoral roll and iwi affiliation responses constitutes information reported at a comparable time period to the census so will not be impacted by any increase in genealogical awareness over time. The use of DIA records for completing adult descent information involves using information that is likely to be older than census-sourced information. The use of birth records to link children's Māori descent information with the self-reported source from a parent will increase completeness but it is likely to undercount those of Māori descent in those age groups. This is due to potential differential availability in information on both parents on the birth record, which may be lower for male parents and for Māori parents.

The use of admin data will also result in improving the information for the younger age groups in particular, but does create age-related differences in the timeliness of the data source and whether it was self-reported. The removal of the within household donor process for determining Māori descent will result in improved quality of the data as information is no longer dependent on an assumption of shared descent which cannot be verified.

The increased use of alternative data sources has markedly reduced the need for imputation; only 2% of all responses have been produced through imputation in the 2023 Census dataset with only 1.5% of those in the 'Māori descent' category being produced in this way. For the 'no Māori descent' category the percentage imputed is slightly higher at 2.1% (Stats NZ, 2024b). The use of imputation to include 'don't know' values produces a dataset where all possible values of the Māori descent variable have been subject to the data sourcing processes.

In summary, the addition of new alternative data sources has improved the coverage for the Māori descent (output) variable, especially for those with Māori descent. However, the change in methods has resulted in yet another break in the Māori descent (output) time series, with the most recent three censuses using different methods (census reports only in 2013, use of a limited set of alterative data sources in 2018, and use of an extended and modified set of alternative data sources in 2023).

The combination of high quality data sources, assessed accuracy of responses and data source consistency support considering the overall quality of the Māori descent (output) data as very high – in agreement with Stats NZ ratings. However the relatively higher reliance on older data sources for those of Māori descent will result in the data being of lower quality than for those with no Māori descent, with a potential impact of undercounting the Māori descent population. As such, the quality of Māori descent information for the Māori descent population should be considered slightly lower than the quality of the no Māori descent population.

3.3. Māori descent (electoral)

The Māori descent (electoral) variable has two valid responses: 'Māori descent' and 'No Māori descent'. This variable is used to determine the census Māori electoral population which in turn helps determine the number and boundaries of both Māori electorates and general electorates as required by the Electoral Act 1993. The Māori descent (electoral) variable is also used to determine the estimated resident population of Māori descent (Stats NZ, 2024b).

The data sources and priority order used for the Māori descent (electoral) variable for the 2023 Census are the same as those used for the Māori descent (output) variable described in section 3.2 above. However, the methods used to determine this variable differ slightly, primarily because the Māori descent (electoral) variable has just two valid options whereas the Māori descent (output) variable also has a 'Don't know' option (see section 3.2). Essentially what this means is that at each data source stage, those without either of the responses 'Māori descent' or 'No Māori descent' are passed on to the next data source to determine if a valid response can be obtained. For example, whereas 85.3% of Māori descent (output) data were obtained from 2023 Census responses, only 83.0% of Māori descent (electoral) data were obtained from 2023 Census responses (Stats NZ, 2024b). The 2.3% difference represents n=114,357 who responded 'Don't know' to the Māori descent question at the 2023 Census.

When the response to the Māori descent question at the 2023 Census was either 'Don't know' or was not provided, data for Māori descent (electoral) were obtained from the following data sources, in this priority order (Stats NZ, 2024b):

1. 2018 Census (7.6%)
2. 2013 Census (3.0%)
3. DIA birth registrations (3.2%)
4. Electoral roll data (0.8%)
5. Parental data from census and admin data sources (0.2%)
6. Grandparent data from census and admin data sources (<0.1%)
7. 'Deterministic derivation' through an iwi affiliation record in some data source (0.1%)
8. Statistical imputation (2.2%).

The same comments about data quality made in section 3.2 apply here: the 2023 Census has used additional data sources which have increased the coverage for the Māori descent electoral variable, but this creates a change in the time series which is important for both the electoral population determination as well as the calculation of the Māori descent estimated resident population. The continued reliance on previous census records for over 10 percent of the responses means that some (3%) of the key information for the next general election will be over 13 years old.

The overall quality of the Māori descent (electoral) variable should be rated as very high – though slightly lower for the Māori descent population due to the differences in reliance on less timely and less self-reported data sources. The use of additional data sources has reduced the requirement of imputation for this variable with 97.8% of sources being either self or kin reported. However, the reliance on previous census data is likely to result in a slight undercount of the Māori descent population due to changes in individual knowledge and reporting of their Māori descent since the 2013 or 2018 Censuses. Interestingly, an evaluation of the 'consistency' of alternative data sources (the extent to which alternative data sources agree with census responses for individuals with valid data for both), and

the overall contribution of different data sources to the final variable, indicates that the Māori descent (electoral) variable is likely to be of higher quality than the Māori descent (output) variable, because of the low consistency of 'Don't know' responses (Stats NZ, 2024b).

3.4. Level 1 ethnicity

The use of administrative data and other alternative data sources to 'fill gaps' in Level 1 ethnicity data improves the counts of Level 1 ethnic groups. However, while overall counts of Level 1 ethnicity are improved by the use of alternative data sources, the accuracy is likely to be worse than if all data were derived from census responses. There are four reasons for this, as noted by the 2018 External Data Quality Panel. Specifically, alternative data sources: (i) are less likely to be derived from self-report; (ii) may not adhere to the Ethnicity Standard (2005 V2.0) used in the census; (iii) have the potential to be reported a considerable time before census day, noting that reports of ethnicity do change over time and with context; and (iv) may be less amenable to the reporting of multiple ethnicities (Census External Data Quality Panel, 2019c, p 52).

We also note that data for 27 % of Level 1 Māori and 25% of Level 1 Pacific ethnicity were sourced from alternative data sources, compared to 11% for European, 11% for Asian, 14% for Middle Eastern, Latin American and African (MELAA), and 15% for Other (Ethnicity – 2023 Census: Information by concept). Digging deeper, for Māori, 16% were sourced from historical census data, 10% from administrative data and 1% from imputation; and for Pacific, 14% were sourced from historical census data, 10% from administrative data and 1% from imputation. Thus, as was the case in 2018, the quality of Level 1 ethnicity data varies across ethnicity; the quality can be considered 'very high' for European, Asian and MELAA, but 'high' for Māori and Pacific. We agree with Stats NZ's ratings in this regard (Ethnicity – 2023 Census: Information by concept).

3.5. Age

The 2023 Census Age variable is 'very high' quality. Eighty-nine point seven (89.7) percent of data for Age were sourced from 2023 Census responses, 10.2% from admin data, and 0.1% from all other methods (Age – 2023 Census: Information by concept). Largely the same methods were used to source Age data in 2023 and 2018, with less reliance on alternative data sources in 2023. When these methods were used previously in 2018 the Age variable was rated by the 2018 Census External Data Quality Panel as 'very high' quality (Census External Data Quality Panel, 2019c), so the Age variable in 2023 should similarly be rated 'very high' quality.

3.6. Gender

Data on gender was collected for the first time at the 2023 Census. Previous censuses have collected data on sex only, whereas data on both 'sex at birth' and 'gender' were collected (in separate questions) at the 2023 Census. The Stats NZ standard on gender defines gender as "a person's social and personal identity as male, female, or another gender or genders that may be non-binary" (Stats NZ, 2021b). Accurate gender information is important not just for communities represented by the information but also because Stats NZ has adopted a 'gender by default' policy for data outputs, replacing the sex variable (Stats NZ, 2021b).

The 2023 Census question offered the three options of 'Male', 'Female', or 'Another gender', with an optional text box to elaborate if 'Another gender' was selected. Where a census response was not provided alternative data sources were used, including admin data (from MSD only as this is the only

administrative source that consistently collects data on 'Another gender'), and statistical imputation using the 'CANCEIS' approach described above.

Overall, 89.7% of data for gender were sourced from census responses, 2.6% from administrative data (MSD), and 7.6% from imputation. Results were:

Male	49.27%
Female	50.32%
Another gender	0.41%

The following discussion focuses on the 'Another gender' category. It should be noted that none of the panel are experts in gender diversity or its measurement, so the discussion should be interpreted with this caveat in mind.

Slightly fewer of those classified as 'Another gender' had data from alternative sources: only 1.7% of data were sourced from administrative data, and 7.1% from imputation (see above for use of alternative data sources for the gender variable overall). Assuming those in the 'Another gender' category were as likely to complete a census form as the overall population, this might indicate that 'Another gender' may be undercounted from alternative data sources. However, the overall impact of this is small – around 0.005 percentage points, assuming census-level proportions of gender categories are obtained from alternative data sources. Moreover, it is unknown whether those in the 'Another gender' category are as likely, more likely, or less likely to complete a census form. The reporting of another gender is strongly patterned by age with over half of those reporting another gender being in the 15–29 years age groups.

Against this, two sources of evidence suggest that the 'Another gender' may be overcounted in 2023 Census data. First, census estimates are higher than estimates from the Household Economic Survey (HES): 0.16% - 0.28%. A caveat is that the sample frame for HES excludes non-private dwellings, some of which (e.g. student hostels) may include a higher proportion of 'Another gender' – if so, HES may slightly underestimate 'Another gender', especially with the response being so strongly age-patterned. More generally, it is plausible that there is under-representation and/or under-reporting of 'Another gender' in surveys (Pho et al., 2023), so surveys may not necessarily be a good benchmark for comparison.

Second, the 2023 Census 'Another gender' count may be impacted by a level of facetious response. While facetious responses are hard to identify, higher counts of 'Another gender' in unexpected groups (e.g., 85+ years age group) might indicate facetious responses played some role. Note that higher counts of 'Another gender' were also found in expected groups – e.g., areas with a higher proportion of young people, particularly areas with tertiary institutions.

Taking all this into account, and with no gold-standard for comparison, it is not possible to judge whether the count for 'Another gender' is accurate – there are too many unknowns to make such an assessment. Counts for males and females are closely in line with expectations, however, and can be considered to be 'very high' quality.

3.7. Iwi affiliation

The quality of the 'iwi affiliation' variable was considered very poor for the 2018 Census, and "not suitable for providing an official iwi count" (Iwi (Information about this variable and its quality)). This was a function of the low response rate for the 2018 Census for those of Māori descent, and the lack of suitable alternative data sources.

For the 2023 Census, Stats NZ worked with technicians from the Data Iwi Leaders Group to develop a methodology to augment iwi affiliation responses from the 2023 Census with alternative data sources, in line with the 'combined census' model. The methodology involved using 2023 Census responses on iwi affiliation from parents, grandparents and great-grandparents – identified through DIA birth records – as the primary alternative data source. Historical census data were used next (from the individual first, then from parents, grandparents and great-grandparents), and finally admin data (again, from the individual first, then from parents, grandparents and great-grandparents) were used for those whose iwi affiliation could not be obtained from earlier steps (Stats NZ, 2024c). This is described in more detail below.

Three sources of admin data were used:

- Schools data from the Ministry of Education (MoE), which captures up to three iwi affiliations for all those with a school enrolment since 2003
- Tertiary data from the MoE, which captures up to three iwi affiliations for all those with a tertiary enrolment since 2002
- Student loans and allowances (SLA) data, which captures up to five iwi affiliations for those who have had a student loan since 1999, or have had a student allowance since 2000.

Data from MoE-tertiary and SLA were combined into a single 'tertiary' data source.

Iwi affiliation data were available from the 2023 Census form for 73.2% of those of Māori descent ("Māori descent" on the Māori descent (output) variable, n=978,246). Note, this includes those who indicated they did not know their iwi – overall, 61.2% of those of Māori descent answered the iwi affiliation question in the 2023 Census AND also reported at least one iwi.

When iwi affiliation data were not available from the 2023 Census form, the following priority order was used for data sources (Stats NZ, 2024c):

1. Parental iwi responses from the 2023 Census (8.3%)
2. Grandparent iwi responses from the 2023 Census (3.7%)
3. Great-grandparent iwi responses from the 2023 Census (0.1%)
4. Iwi responses from historical censuses (2018 and 2013) (8.5%)
5. Parental iwi responses from historical censuses (1.4%)
6. Grandparent iwi responses from historical censuses (0.2%)
7. Great-grandparent iwi responses from the historical censuses (<0.1%)
8. Iwi affiliation from 'tertiary' admin data (1.5%)
9. Iwi affiliation from 'schools' admin data (0.4%)
10. Parental iwi affiliation from admin data (0.2%)
11. Grandparent iwi affiliation from admin data (<0.1%)
12. Great-grandparent iwi affiliation from admin data (<0.1%).

After all of these steps were followed, 97.5% of those of Māori descent had iwi affiliation data (the remaining 2.5% had missing data for this variable). Eighty-two point five percent of those of Māori descent had iwi affiliation data AND reported at least one iwi.

Stats NZ rates the quality of 2023 Census iwi affiliation data as moderate (Iwi affiliation – 2023 Census: Information by concept). The panel agrees with this assessment. While coverage has been greatly

improved (from 73.2% to 97.5%), there are several quality issues with the data, including that (i) consistency between the 2023 Census and alternative data sources tends to be modest, and (ii) the alternative data sources typically do not cover the full population. For example:

- The extent to which iwi affiliations match between the 2023 Census and alternative data sources for individuals who reported both is only 66.7% for historical census, 62.6% for tertiary data, and 51.2% for MoE-schools data (Stats NZ, 2024c).
- Determining parental relationships from DIA birth records greatly affects the coverage of parental data. Because there is complete digitisation of birth records from 1985 onwards, and only patchy digitisation prior to 1985, those born before 1985 are less likely to have a parental link. In fact, very few aged 50+ had a parental link (Stats NZ, 2024c). Also, using DIA birth records will miss individuals not born in New Zealand, will tend to link to mothers more than fathers (Milne et al., 2020), and will not distinguish between biological parents and adoptive parents.
- Among those of Māori descent, 70.3% had an iwi recorded in a historical census, and only 37.4% had an iwi recorded in tertiary data and 28.7% in schools data (Stats NZ, 2024c). Those of Māori descent who have never attended tertiary institutions or attended prior to 2002 will be missed, as will those who enrolled in school prior to 2003.
- On the positive side, alternative data sources tended to include as many iwi affiliations (if not more), than reports from the 2023 Census (Stats NZ, 2024c).

In summary, while the methodology used by Stats NZ to complete the iwi affiliation variable is well-considered and appropriate given the low coverage of the 2023 Census, there are several quality and consistency issues to be aware of when using the data.

3.8. Activity limitations

The Activity limitations items comprise the six questions from the Washington Group Short Set on Functioning, which assess difficulties in basic functioning in six areas: vision, hearing, mobility, cognition, self-care, and communicating (Washington Group on Disability Statistics, 2020). These questions are asked of census respondents aged 5-years and older. They are particularly vulnerable to low response rates as there are no alternative data sources that can be used to fill in missing responses. For the 2018 Census, data were unavailable for 16.4% - 16.8% of those aged 5 years and older across the six questions, which resulted in the data being rated of 'poor' quality (Activity limitations (information about this variable and its quality)). The data are only slightly better at the 2023 Census: data are unavailable for 15.5% - 15.6% of those aged 5-years and older across the six questions. As such, the data should be considered 'poor' quality again.

4 – Summary and conclusions

4.1. Summary

The 88.3% interim response rate of the 2023 Census required Stats NZ to implement methods to augment the census file. In response, Stats NZ took sensible and justifiable steps to implement methods to make the 2023 Census file more complete and of as high quality as it could be. The panel endorses the methods used by Stats NZ to fill gaps in the 2023 Census, which were largely the same as the methods undertaken to fill gaps in the 2018 Census, and which were endorsed by the 2018 Census External Data Quality Panel (Census External Data Quality Panel, 2019b, 2019c, 2019a, 2020).

While the methods implemented by Stats NZ improved the quality of 2023 Census data, the variables assessed by the panel – Priority 1 variables and the Priority 3 'Activity limitations' variable – were of varying quality. The Table below summarises the panel's rating of each variable using the same scale used by Stats NZ in their data quality assurance process: very high, high, moderate, poor, very poor (Stats NZ, 2024a).

Table. Quality ratings of the 2023 Census variables assessed by the panel.

Variables	**Panel rating**	**Comments**
1. Usually resident population count (total & TALB)	Very high	
2. Māori descent (output)	Very high	Slightly lower quality for those in the 'Māori descent' category compared to those in the 'no Māori descent' category.
3. Māori descent (electoral)	Very high	Slightly lower quality for those in the 'Māori descent' category compared to those in the 'no Māori descent' category.
4. Level 1 ethnicity	Very high/high	Very high: European, Asian, MELAA, Other High: Māori, Pacific
5. Age	Very high	
6. Gender	Very high/ Unable to judge	Male and female categories rated 'very high' Unable to judge the accuracy of the 'Another gender' category.
7. Iwi affiliation	Moderate	Coverage is lower for people of older ages.
8. Activity limitations	Poor	>15% missing data

Four variables can be considered 'very high' quality: *Usually resident population count*, both for the country as a whole and by territorial authority and local board; *Age*; *Māori descent (output); and Māori descent (electoral)*. The methods employed by Stats NZ – particularly the use of administrative data – are likely to produce very high quality estimates of the population count and the age distribution of the population. The methods employed are also likely to produce very high quality estimates of Māori

descent for output and electoral purposes, though may slightly underestimate the number of individuals of Māori descent.

Gender is also 'very high' quality for the 'Male' and 'Female' categories, but of unknown quality for the 'Another gender' category – there is no gold standard against which to compare this category, and too many unknowns about the data collection and quality of alternative data sources to definitively assess the quality of this category.

Level 1 ethnicity can be considered 'very high' quality for four of six ethnic groups: European, Asian, MELAA, and Other; but of only 'high' quality for the Māori and Pacific ethnic groups. Here, alternative data sources are likely to be of lower quality, so the use of such sources for more than a quarter of those included in the Māori and Pacific categories lowers the overall quality of data for these ethnic groups.

Iwi affiliation can be considered 'moderate' quality. While coverage has been greatly improved (from 73.2% to 97.5%), there are several quality issues with the data, including that (i) consistency between the 2023 Census and alternative data sources tended to be modest, and (ii) the alternative data sources typically did not cover the full population.

As was the case for 2018 Census, the low response rate for 2023 Census and the lack of alternative data sources to fill in missing responses means that *Activity limitations* should again be considered 'poor' quality. Without a suitable back-up data source (which seems unlikely to materialise), a near-complete census data collection is needed to ensure high quality data are available for *Activity limitations*. The needs of the disabled community are unlikely to be adequately served without such a data collection.

4.2. Conclusions

2023 Census was the second time Stats NZ had to augment the census file with alternative data sources – that is, responses not provided by a census respondent at census time. The first time this happened (in 2018) Stats NZ hastily developed methods to fill gaps in the census (Census External Data Quality Panel, 2019c). Second time around in 2023, Stats NZ used a combined model by design, which combined census responses with alternative data sources. As a result, Stats NZ were armed with well-developed processes to fill gaps in the 2023 Census, processes which have been refined and further developed since 2018. The panel is of the view that the methods employed by Stats NZ are now well-established and work well.

The methods employed are not perfect substitutes for census responses, however. The use of alternative data sources, particularly admin data, works well for some purposes, such as counting the population of the whole country, counting higher-level geographies (i.e., territorial authorities and Auckland local boards), and accurately capturing the age distribution of the population. The use of alternative data sources works less well for other purposes. For example, small area geographies (e.g., meshblocks) may be impacted by incorrect meshblock enumerations; Māori and Pacific ethnicity and level 2 and below ethnicities may be impacted by lower quality of ethnic data from administrative datasets; iwi affiliation is impacted by lower quality and inconsistent alternative data sources; and 'Activity limitations' is impacted by a lack of alternative data sources. Other census variables not covered in this report are also poorly served by alternative data sources, particularly family and household concepts that are reported at census time and for which administrative data sources are not readily available (Census External Data Quality Panel, 2019b; Stats NZ, 2024h), although there have been ongoing developments with the available data sources and methods for these concepts (Stats NZ, 2023).

For this reason, the panel believes that Stats NZ needs to give priority to finding ways to collect data directly where alternative sources are not available or are of clearly lower quality. A fully administrative census will be of lower quality, particularly for some ethnicities, for lower-level geographies, and for census topics not well served by administrative data. Well-designed surveys may address some of these issues but will struggle to capture small populations and small areas well. A full enumeration census should still be seen as the gold standard, while acknowledging that attaining high response rates across the population is becoming increasingly difficult. Tailoring initiatives to communities shows some promise, as demonstrated for the 2023 Census by Te Mana Whakatipu (Jack & Bowlby, 2024).

References

Census External Data Quality Panel. (2019a). *Assessment of Variables*. Statistics New Zealand. https://www.stats.govt.nz/reports/2018-census-external-data-quality-panel-assessment-of-variables

Census External Data Quality Panel. (2019b). *Final Report of the 2018 Census External Data Quality Panel*. Statistics New Zealand. https://www.stats.govt.nz/news/final-report-from-2018-census-external-data-quality-panel-now-available

Census External Data Quality Panel. (2019c). *Initial Report of the 2018 Census External Data Quality Panel*. Statistics New Zealand. https://www.stats.govt.nz/reports/initial-report-of-the-2018-census-external-data-quality-panel

Census External Data Quality Panel. (2020). *Data sources for key 2018 Census individual variables* (p. 1). Statistics New Zealand. https://www.stats.govt.nz/reports/2018-census-external-data-quality-panel-data-sources-for-key-2018-census-individual-variables/

Jack, M., & Bowlby, G. (2024). *Report of the Statutory Review of New Zealand's 2023 Census*. Statistics New Zealand. https://www.stats.govt.nz/reports/report-of-the-statutory-review-of-new-zealands-2023-census/

Milne, B. J., Atkinson, J., Blakely, T., Day, H., Douwes, J., Gibb, S., Nicolson, M., Shackleton, N., Sporle, A., & Teng, A. (2019). Data Resource Profile: The New Zealand Integrated Data Infrastructure (IDI). *International Journal of Epidemiology, 48*(3), 677–677e. https://doi.org/10.1093/ije/dyz014

Milne, B., Li, E., & Sporle, A. (2020). *Intergenerational Analyses Using the IDI: An update* (p. 19). COMPASS Research Centre, University of Auckland. https://cdn.auckland.ac.nz/assets/auckland/arts/our-research/research-institutes-centres-groups/compass/whole-population-data-analysis/Intergenerational-Links-IDI-Update.pdf

Pho, A. T., Bates, N., Snow, A., Zhang, A., Logan, R., Dastur, Z., Lubensky, M. E., Flentje, A., Lunn, M. R., & Obedin-Maliver, J. (2023). Asking sexual orientation and gender identity on health surveys: Findings from cognitive interviews in the United States across sexual orientations and genders. *SSM - Qualitative Research in Health, 4*, 100344. https://doi.org/10.1016/j.ssmqr.2023.100344

Statistics New Zealand. (2019). *Linking 2018 Census respondents to the Integrated Data Infrastructure*. Statistics New Zealand. https://www.stats.govt.nz/assets/Uploads/Methods/Linking-2018-census-respondents-to-the-integrated-data-infrastructure/Download-data/Linking-2018-census-respondents-to-the-integrated-data-infrastructure.pdf

Stats NZ. (2021a). *2023 Census: Final content report.* (p. 39). Stats NZ. https://www.stats.govt.nz/reports/2023-census-final-content-report#download-document

Stats NZ. (2021b). *Data standard for gender, sex, and variations of sex characteristics* (p. 36). Stats NZ. https://www.stats.govt.nz/methods/data-standard-for-gender-sex-and-variations-of-sex-characteristics/

Stats NZ. (2022a). *Integrated Data Infrastructure*. Stats NZ. https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/

Stats NZ. (2022b). *Methodologies for filling gaps in gender and sex at birth concepts for the 2023 Census* (p. 14). Stats NZ. https://www.stats.govt.nz/assets/Consultations/2023-Census-sex-and-gender-methodology-consultation/2023-Census-Methodology-for-filling-gaps-for-gender-and-sex-concepts-Consultation-technical-paper-December-2022.pdf

Stats NZ. (2023). *Experimental administrative population census (third iteration): Data sources, methods, and quality for household information* (p. 38). Statistics New Zealand. https://www.stats.govt.nz/research/experimental-administrative-population-census-third-iteration-data-sources-methods-and-quality-for-household-information/

Stats NZ. (2024a). *Data quality assurance in the 2023 Census* (p. 1). Stats NZ. https://www.stats.govt.nz/methods/data-quality-assurance-in-the-2023-census/#:~:text=When%20assessing%20the%20quality%20of,of%20data%20are%20evidence%2Dbased

Stats NZ. (2024b). *Data sources and imputation for Māori descent in the 2023 Census (updated)* (p. 37). Stats NZ. https://www.stats.govt.nz/methods/data-sources-and-imputation-for-maori-descent-in-the-2023-census

Stats NZ. (2024c). *Data sources and methodology for iwi affiliation in the 2023 Census* (p. 59). Statistics New Zealand. https://www.stats.govt.nz/methods/data-sources-and-methodology-for-iwi-affiliation-in-the-2023-census/

Stats NZ. (2024d). *Editing, data sources, and imputation in the 2023 Census*. (p. 46). Stats NZ. https://www.stats.govt.nz/assets/Methods/Editing-data-sources-and-imputation-in-the-2023-Census/editing-data-sources-and-imputation-in-the-2023-census.pdf#:~:text=Editing%2C%20data%20sources%2C%20and%20imputation%20in%20the%202023,is%20inconsistent%20or%20not%20usable%20for%20other%20reasons

Stats NZ. (2024e). *Linking 2023 Census responses to the Integrated Data Infrastructure 2023 Census | Tatauranga 2023* (p. 34). Statistics New Zealand. https://www.stats.govt.nz/methods/linking-2023-census-responses-to-the-integrated-data-infrastructure/

Stats NZ. (2024f). *Methodology for using admin data to count people in the 2023 Census*. (p. 47). Statistics New Zealand. https://www.stats.govt.nz/methods/methodology-for-using-admin-data-to-count-people-in-the-2023-census/

Stats NZ. (2024g). *Predicting usual residence address from admin data in the 2023 Census 2023 Census | Tatauranga 2023* (p. 48). Statistics New Zealand. https://www.stats.govt.nz/methods/predicting-usual-residence-address-from-admin-data-in-the-2023-census/

Stats NZ. (2024h). *Technical feasibility: Measuring population and dwellings using administrative data* (p. 53). Statistics New Zealand. https://www.stats.govt.nz/research/technical-feasibility-measuring-population-and-dwellings-using-administrative-data?utm_source=Statistics+NZ+Newsletters+%28Stats+News%2C+EDU%2C+Census+Advisory%29&utm_campaign=36c1678f60-EMAIL_CAMPAIGN_2019_12_16_07_06_COPY_01&utm_medium=email&utm_term=0_e8994da608-36c1678f60-506235361

Washington Group on Disability Statistics. (2020). *An introduction to the Washington Group on Disability Statistics question sets. 2020.* (p. 5). Washington Group on Disability Statistics. https://www.washingtongroup-disability.com/fileadmin/uploads/wg/Documents/An_Introduction_to_the_WG_Questions_Sets__2_June_2020_.pdf