Stats NZ
Tatauranga Aotearoa

Data sources and methodology for iwi affiliation in the 2023 Census

2023 Census | Tatauranga 2023

Te Kāwanatanga o Aotearoa
New Zealand Government

Citation
Stats NZ (2024). *Data sources and methodology for iwi affiliation in the 2023 Census*. Retrieved from www.stats.govt.nz.

ISBN 978-1-991307-17-0

Stats NZ accessed the data in the IDI for use in the 2023 Census in accordance with security and confidentiality provisions of the Data and Statistics Act 2022. Only people authorised by the Data and Statistics Act 2022 are allowed to see data about a particular person, household, business, or organisation and the results in this paper have been confidentialised to protect these groups from identification.

Published in September 2024 by
Stats NZ Tatauranga Aotearoa
Wellington, New Zealand

Contact
Stats NZ Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4600

www.stats.govt.nz

Contents

List of tables and figures

List of tables

List of figures

Purpose and summary

Purpose

Data sources and methodology for iwi affiliation in the 2023 Census outlines the data sources and methodology used to determine values for the iwi affiliation variable in the 2023 Census, including use of alternative data sources to mitigate missing census responses.

Editing, data sources, and imputation in the 2023 Census provides more information on the processes used for other census variables.

Summary of key points

By design, the 2023 Census combines census response data and alternative data sources.

The iwi affiliation variable is one of the highest-priority variables released in the 2023 Census. It is used to provide iwi, iwi-related groups, and Māori with information to support their aspirations, to inform decisions that affect iwi and Māori, and to meet the Crown's partnership obligations under Te Tiriti o Waitangi.

However, non-response in the 2023 Census means that some data is missing. Without alternative data sources, iwi affiliation would not be available for many people of Māori descent who are included in census data. This missingness would also differ across different subpopulations.

The methodology for iwi affiliation in the 2023 Census was co-developed and co-tested by the Data Iwi Leaders Group (Data ILG) and Stats NZ. Working with Te Kāhui Raraunga Charitable Trust (the mahi arm of Data ILG), we have ensured that iwi Māori data expertise was prioritised in the design of the methodology. The methodology uses alternative data sources to fill missing and residual values for the iwi affiliation variable. This builds on mahi undertaken by Data ILG and Stats NZ to produce 2018 Census iwi estimates. While alternative data sources were used for some other variables in the 2018 Census, this is the first time alternative data sources have been used for iwi affiliation in official census data.

The alternative data sources in this paper include historical census data (from the 2013 and 2018 Censuses), as well as administrative (admin) data from the Ministry of Education (MoE) and the Ministry of Social Development (MSD). The methodology also includes the use of parents', grandparents', or great-grandparents' iwi affiliation where parental relationships are found in Department of Internal Affairs (DIA) data.

The use of alternative data sources increases the percentage of the 2023 Māori descent census usually resident population that have a value for iwi affiliation from 716,208 people (73.2 percent) to 954,042 people (97.5 percent).

Mana Ōrite Relationship Agreement provides more information on the relationship between Stats NZ and the Data Iwi Leaders Group.

Te Kāhui Raraunga provides more information on the Trust.

Background information

Stats NZ published 2023 Census: High Level Design in October 2021. The combined model relied on full field enumeration as the primary source of census information, with alternative data sources used to fill missing responses. The alternative data sources used in this model include historical census data, admin data, and statistical imputation.

As part of the joint work programme of the Mana Ōrite Relationship Agreement, Stats NZ had ongoing discussions about the 2023 Census iwi affiliation variable with Data ILG. This led to governance decisions, informed by the advocacy and leadership of Data ILG, that improved the quality of the variable. These decisions included making iwi affiliation a priority 1 variable[1] and agreement on the key performance indicators[2] for the response rate and coverage rate of the Māori descent population. In addition, there was exploration into the benefits of reusing some of the 2018 Census mitigation methodology and the use of alternative data sources.

The order and use of alternative data sources in this process, within the 2023 Census design and processes, have been worked on together by Data ILG technicians and Stats NZ under the Mana Ōrite Relationship Agreement. The final decisions align with recommendations made by Data ILG and Stats NZ.

The interim collection response rate for people of Māori descent in the 2023 Census is 77.0 percent, compared with the national rate of 88.3 percent (Stats NZ, 2024a) See Interim coverage rates, collection response rates, and data sources for the 2023 Census for more details.

---

[1] Priority 1 topics are those that make up the core reason for conducting a census and have the highest output quality need. For more details, please refer to Appendix in 2023 Census: Final content report.

[2] The key performance indicator for national response rate for Māori descent was set for greater than or equal to 90 percent and the national net coverage rate for people of Māori descent was set for greater than or equal to 98 percent. For more details, please refer to Measuring the success of the 2023 Census.

The interim coverage rate for people of Māori descent is 97.2 percent, compared with the national rate of 99.1 percent. Coverage rates are higher than collection response rates, as some people who did not respond to the census are added through the use of admin data (also called admin enumeration).

The first section of this paper covers the definition, use, and classification of iwi affiliation. The next section details the iwi affiliation question in the 2023 Census, and the subject population. The third section has an overview of the 2023 Census combined model. Following on there is a summary of alternative data sources and details of the methodology. The final sections include results, quality assessment of the variable, and strengths and limitations of the methodology.

Background to the iwi affiliation concept

This section describes Stats NZ's definition of iwi affiliation and lists uses of the data. It also describes the classification, how responses are mapped to the classification, and the changes to the classification.

Definition used by Stats NZ

An iwi (a Māori tribe) is one of the largest kinship groupings and is generally made up of several hapū that are all descended from a common ancestor. Hapū are clusters of whānau (families) where the whānau is usually an extended family grouping consisting of children, parents, often grandparents, and other closely related kin.

For statistical purposes, Stats NZ defines iwi as a whakapapa-based kinship grouping that generally has several hapū and one or more active marae, and a recognised structure that represents the interests of the iwi, such as a rōpū whakahaere, committee, or board.

Uses of iwi affiliation data

Iwi affiliation data is published by Stats NZ and on Te Whata – a data platform tailored specifically by iwi for iwi. Te Whata was designed by Te Kāhui Raraunga Charitable Trust to support iwi to be designers and users of data. Te Kāhui Raraunga Charitable Trust was established in 2019 to undertake mahi on behalf of Data ILG.

Iwi affiliation data is used to:

- monitor the performance of Treaty of Waitangi obligations by the Crown and iwi
- help the allocation of resources and funds to and within iwi

- help with Waitangi Tribunal decisions on, for example, land ownership and fishing rights
- help iwi organisations, local and central government agencies, and non-governmental organisations to evaluate, plan, and provide services to iwi in areas such as housing, health, and education.

Iwi affiliation classification

Iwi affiliation uses a two-level hierarchical classification. Census iwi and iwi-related groups V2.1.0 has more detail for examining the classification.

Level 1 of the classification lists 18 broad categories. These include:

- 12 categories that group iwi into geographical regions for statistical purposes
- a category for confederations and waka, where the iwi is not named
- a category for when iwi are named but the region is not known
- a category for when hapū are affiliated to more than one iwi
- a category for when the region is known but the iwi is not named
- a category for people who ‘Don’t know’ their iwi
- a category for ‘Not elsewhere included’ categories.

Level 2 of the classification lists iwi and iwi-related categories:

- Within the 12 geographical regions, there are 139 ‘Fully coded iwi’ categories.
- Within the four level 1 categories for ‘Confederations and waka’, ‘Iwi named, region not known’, ‘Hapū affiliated to more than one iwi’, and ‘Region known, iwi not named’, there are a total of 45 ‘Partially coded iwi’ categories.
- There is a category for ‘Don’t know’.
- The ‘Not elsewhere included’ category contains the residual categories.

The categories ‘Refused to answer’, ’Response unidentifiable’, ‘Response outside scope’ and ‘Not stated’ are called residual categories. A valid response for the iwi affiliation variable is any response except responses that fall into these residual categories. The response ‘Don’t know’ is considered a valid response.

Multiple responses could be provided to the iwi affiliation question. People reporting more than one iwi affiliation are counted in each category they stated. Therefore, the total number of responses in a table will be greater than the number of respondents.

Data is only published at level 2 of the classification. The geographic regions provide structure to the long list of iwi and are not intended for reporting purposes.

The separate classification Census Iwi grouping recode V2.1.0 is used to publish data with iwi grouped together. These groupings are based on geographic, historical, and tribal groupings of iwi and/or hapū.

Mapping iwi affiliation responses to the classification

When processing census responses to the iwi affiliation question, text responses are coded into the various categories in the iwi classification.

First, a series of rules are used to try to automatically match the responses with a 'synonym'. Synonyms are a list of names (sometimes including hapū), or names with different spellings, that correspond to each iwi in the classification category. The list of synonyms for each version of the classification are available on Ariā. Responses that cannot be automatically matched are sent to the Manual Intervention team who had specific capability and knowledge in manually coding the iwi affiliation responses. If a response could still not be classified, it is coded to one of 'Refused to answer', 'Response unidentifiable', or 'Response outside scope'.

Stats NZ notes that, as part of this process, synonyms are regularly updated and the methodology used should be continually improved.

Changes to the classification

The iwi classification was first developed in 1991, followed by the statistical standard in 1994. Minor updates to the iwi classification have generally taken place before each Census of Population and Dwellings but it was not until 2016 with the Recommendations from the 2016 review of the statistical standard for iwi that the standard was first reviewed in its entirety.

Some key recommendations from this research and consultation (Stats NZ, 2016) were to:

- highlight that the standard is for statistical purposes (Recommendation 1)
- update the iwi standard and classification (Recommendation 2)
- change the process for including groups in the classification (Recommendation 3).

When the classification is updated, some iwi that were classified as a single iwi in the old classification are classified as multiple iwi in the new classification. In this situation, the synonyms identified with the iwi category in the historical classification are reallocated to the appropriate iwi category in the new classification, as well as having new synonyms added. There are cases where some iwi that are newly added to the classification for the 2023 Census do not have any past synonyms. That is, if an iwi A is split in the classification into iwi A (Region 1) and iwi A (Region 2), and

the synonyms for iwi A did not include region information, then historical census responses for iwi A will be coded to iwi A, region not known.

The admin data sources use a range of versions of the classification and are provided to Stats NZ already coded to a version of the iwi classification by the agency that collected the data. The actual response or synonyms used are not provided to Stats NZ, and so we are not able to code these to the current classification. For iwi in historical classifications that have been split into multiple iwi in the 2023 classification, the iwi codes are recoded to the iwi with the largest population. That is, if an iwi A in the old classification is split into iwi A1 and iwi A2 in the new classification, then responses to A will be recoded to the larger of A1 and A2.

There are several different versions of the iwi classification. The 2023 Census uses the Census iwi and iwi-related groups V2.1.0 classification, which is based on the Iwi and iwi-related groups statistical classification V2.1.0.

The 2018 Census used the Iwi and iwi-related groups statistical classification V1.0.0. The 2013 Census used the Iwi - New Zealand Standard Classification V6.0.0.

The current data from MoE uses the Iwi and iwi-related groups statistical classification V1.0.0. Current data from MSD uses the Iwi - New Zealand Standard Classification V5.0.0.

See Appendix 1 for a summary of changes to the iwi classification over time.

Iwi affiliation in the 2023 Census

This section describes the subject population and the iwi affiliation question in the 2023 Census.

In the 2023 Census, Stats NZ uses the Census iwi and iwi-related groups V2.1.0 classification. There is more information about the iwi affiliation variable and its quality at Iwi affiliation – 2023 Census: Information by concept.

Subject population for iwi affiliation

The subject population for iwi affiliation is the Māori descent census usually resident population.

The population we apply the iwi affiliation methodology to is the census usually resident population who are of Māori descent, or who don’t know if they are of Māori descent. This is made up of people who responded to the 2023 Census with a response of ‘Yes’ or ‘Don’t know’ to the Māori descent question, as well as those

who did not answer that question, or did not respond to the 2023 Census, and for whom an alternative data source was used to assign a value of ‘Yes’ or ‘Don’t know’. People who don’t know if they are of Māori descent were asked the iwi affiliation question, as it can be used to determine their Māori descent value for the Māori descent – electoral variable. Data sources and imputation for Māori descent in the 2023 Census will be updated on 23 October 2024 to include more information about the Māori descent – electoral variable.

For people missing Māori descent information in the 2023 Census, deterministic derivation can determine if they are of Māori descent. If someone is missing Māori descent information in the 2023 Census and all of the alternative data sources, but they affiliate to an iwi in the 2023 Census or any alternative data source (including from their parents/grandparents/great-grandparents), then they are assigned as being of Māori descent. This means that a person having iwi affiliation (in the 2023 Census or an alternative data source) can lead to their inclusion in the subject population for iwi affiliation.

Data sources and imputation for Māori descent in the 2023 Census has more information on the methodology for Māori descent in the 2023 Census.

Iwi affiliation question in the 2023 Census

People who report ‘Yes’ or ‘Don’t know’ to the Māori descent question in the 2023 Census were asked if they know the name(s) of their iwi.

Stats NZ worked with Te Kāhui Raraunga Charitable Trust to redesign the iwi affiliation question in the online form. The online form presented a map of Aotearoa where respondents were provided three ways to respond:

- Use the map to select a region, then select their iwi from a list of iwi in that region. This is shown in figure 1.
- Bypass the map and select an iwi from the list of iwi (that could also be filtered by region). This is shown in figure 2.
- Manually type the iwi, where an as-you-type function would produce options from the list of iwi. Respondents could also manually type their iwi without using the as-you-type functionality.

The list of iwi displayed to respondents included the 139 fully-coded iwi. Up to 16 iwi could be reported by a person using the online form. Design of forms for the 2023 Census has further information on the map view and why these changes were made.

On the 2023 Census paper forms, five fields were available for respondents to list iwi. With the paper forms, respondents were provided with the list of 139 iwi names in the Guide Notes.

Respondents could choose between English and te reo Māori on the online form using a toggle. Meanwhile, the paper forms were available in three formats: English, bilingual (English and te reo Māori), and large print. Figure 3 shows the iwi affiliation question in the paper form of the 2023 Census.

Figure 1

Iwi affiliation question in the 2023 Census online form – map view

Figure 2

Iwi affiliation question in the 2023 Census online form – list view

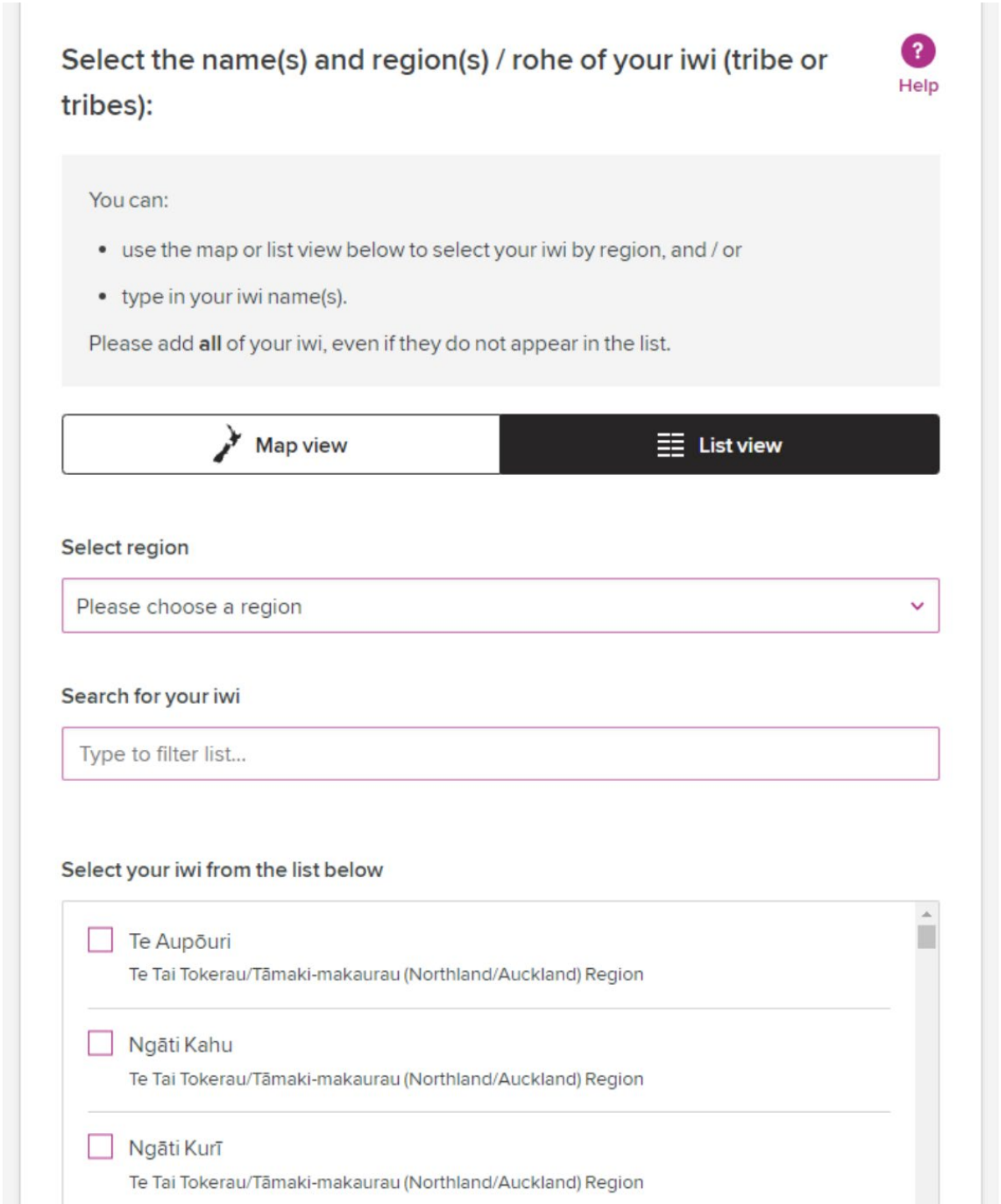
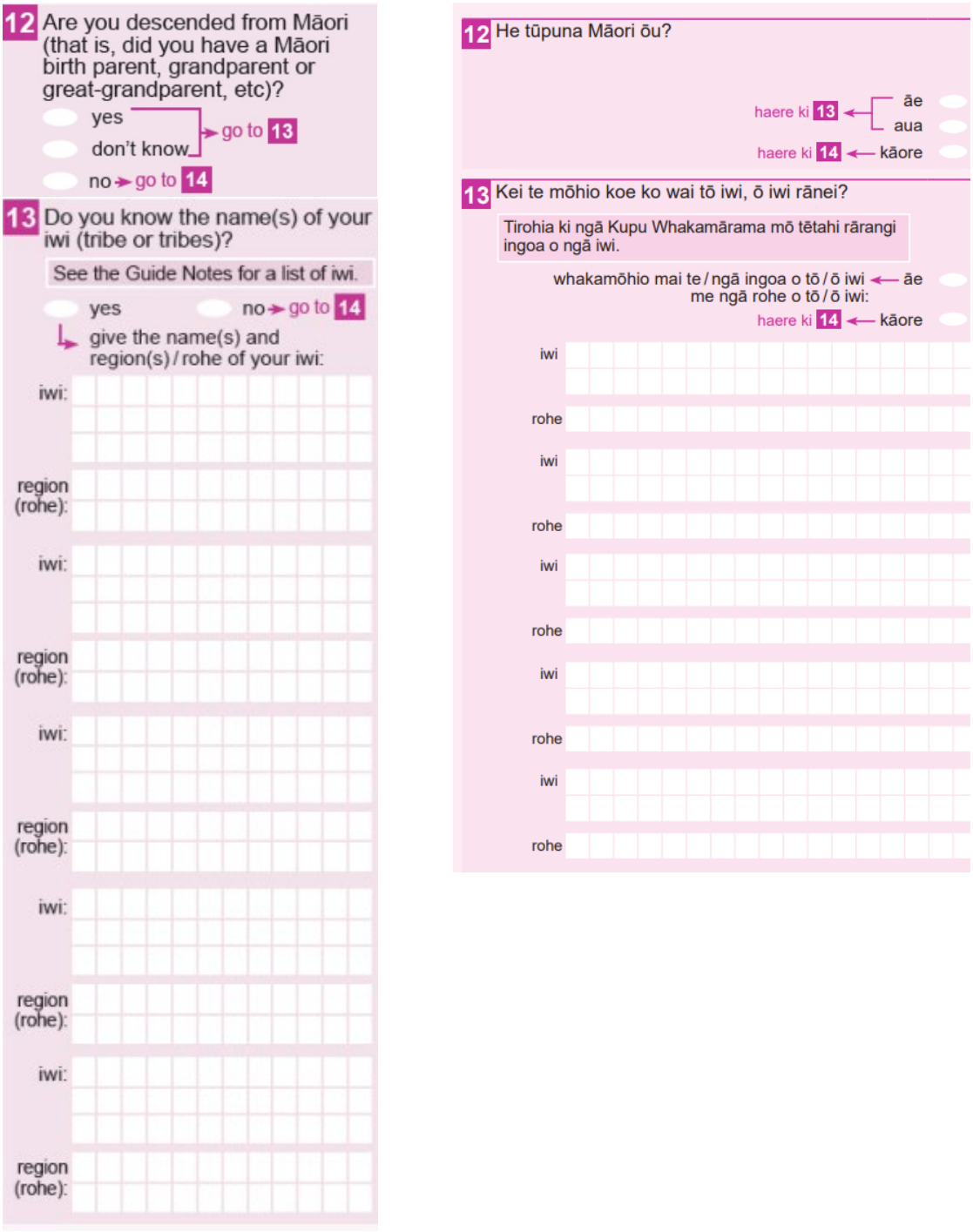


Figure 3

Māori descent and iwi affiliation questions in the 2023 Census paper form



2023 combined census model

This section describes how 2023 Census responses are combined with alternative data sources to reduce the amount of missing data in the census dataset.

How the 2023 combined census model works

The 2023 Census has implemented a combined model by design, with alternative data sources used to supplement census responses. When an individual has not responded to the census and we are confident in the quality of the admin record, we use admin data to add records (referred to as admin enumerations).

Methodology for using admin data to count people in the 2023 Census provides more information about the admin enumeration process.

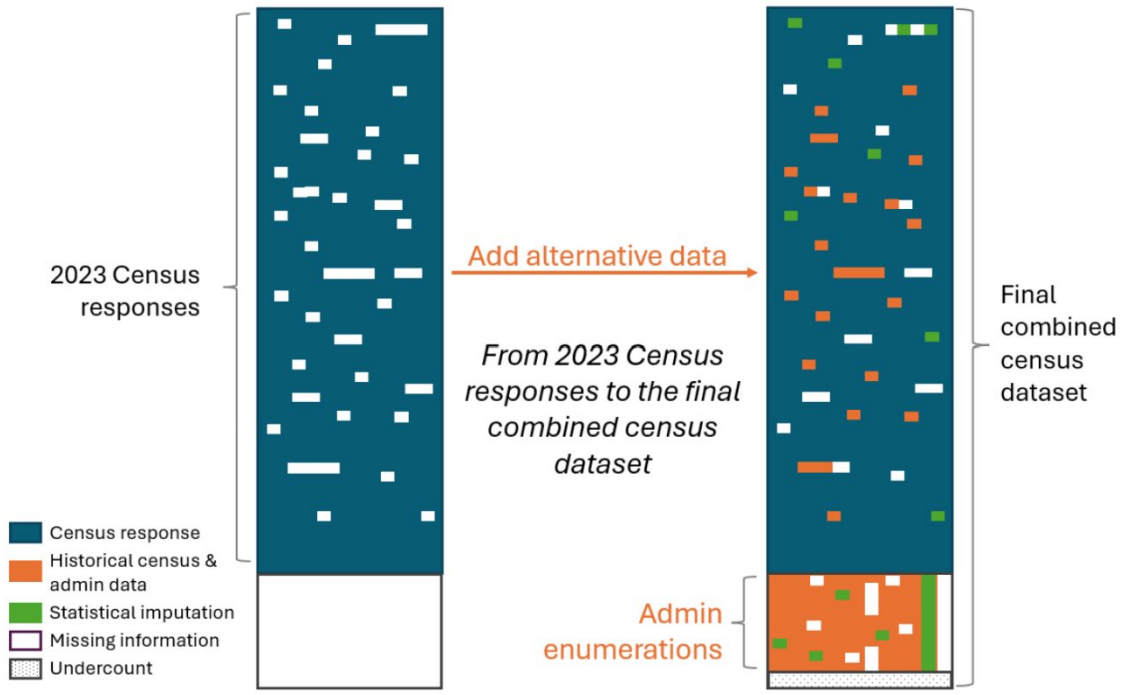
When we are missing responses within a census record, we use historical census data, admin data, and statistical imputation to fill in the gaps.

Editing, data sources, and imputation in the 2023 Census provides more information.

Figure 4 shows the difference between filling in variables and admin enumeration.

Figure 4

Admin enumeration of records and data sourcing of variables



Note: Statistical imputation was not used for the iwi affiliation variable.

Integrated Data Infrastructure (IDI)

Integrated Data Infrastructure (IDI) is a large research database, which holds microdata (unit-level data, such as at the individual or business level) about people and households. Data is gathered from a range of government agencies and non-

government organisations (admin data) and from Stats NZ surveys (including the 2013 and 2018 Censuses).

Integrated data provides more information about how data is linked together, or integrated, to form the IDI.

Unlike most access to the IDI, which uses Stats NZ's Data Lab, development of the methodology for filling in missing values in census variables was conducted in a separate secure environment. This environment followed security protocols similar to those used by IDI and followed frameworks like the Five Safes framework and Ngā Tikanga Paihere.

Privacy impact assessment for the use of admin data in the 2023 Census provides more information.

Linking census respondents to the IDI

2023 Census respondents are linked to the IDI spine using a probabilistic approach, which confirms a link when there is a high enough likelihood that sets of records match to the same individual.

Linking 2023 Census respondents to the Integrated Data Infrastructure describes this process in more detail. Overall, 97.9 percent of all census responses were linked (Stats NZ, 2024b). For the census usually resident population, 98.2 percent of responses are linked, including 97.3 percent of responses identified as being of Māori descent.

These links enable us to identify where we are missing census responses for an individual, including for the iwi affiliation question. We can then use admin and historical census data about that individual to fill in these gaps.

Table 1 shows how the combined model looks for people of Māori descent (the subject population for iwi affiliation) and their iwi affiliation values. This shows the breakdown of respondents and admin enumerations, as well as whether their iwi affiliation value is sourced from a 2023 Census response, an alternative data source, or is missing. In the Results section we discuss in more detail how the different alternative data sources contribute, and how that interacts with the different sources of Māori descent information.

Table 1

Iwi affiliation counts by unit source in the census file, and source of the iwi affiliation variable, at a high level				
Population of Māori descent	Unit source		Iwi affiliation variable source	
978,246	Census individual form	734,745	2023 Census	716,208
			Alternative data sources	15,618
			No information	2,922
	Individuals on the household listing only	40,668	2023 Census	0
			Alternative data sources	30,138
			No information	10,530
	Admin enumeration	202,833	2023 Census	0
			Alternative data sources	192,084
			No information	10,749
Source: Stats NZ				

Limitations of alternative data sources

The alternative data sources we use to fill in missing responses may not reflect how an individual would have responded to the 2023 Census. This could be due to changes in an individual's knowledge or perception over time, differences in version of the iwi classification, differences in collection method based on the government agency, or for other unknown reasons. This can lead to inconsistencies between data sources.

We provide an indication of the quality of each alternative data source through metric 1 of the quality rating scale.

Data quality assurance in the 2023 Census provides more information on quality ratings.

Alternative data sources for iwi affiliation in the 2023 Census

This section describes the alternative data sources used to replace missing and residual values in the iwi affiliation variable for the 2023 Census. In the following subsections, we discuss in detail how iwi affiliation is collected in each source and whether it aligns with the classification for the 2023 Census variable.

Historical census data

Throughout this paper, we use 'historical census data' to refer to data from the 2013 and 2018 Censuses.

The iwi classifications used in historical censuses have differences from the classification used in the 2023 Census, as additional iwi have been added. Historical census data may not represent iwi who have been added to the classification used in the 2023 Census (Census iwi and iwi-related groups V2.1.0) since the 2013 Census or the 2018 Census or that have had changes made to their classification (for example, iwi split by region). For this reason, Stats NZ uses synonyms to enable a wide range of historical census responses to be matched to the most suitable iwi in the most recent classification.

Iwi affiliation in the 2013 Census

People who responded ‘Yes’ to the Māori descent question in the 2013 Census were asked if they knew the name(s) of their iwi. However, anyone who responded on paper could provide an iwi irrespective of their response to the Māori descent question.

The 2013 Census used paper and online forms and the majority of responses were received on paper. On the paper form, respondents could write up to five iwi, and the list of iwi names was provided in the Guide Notes. The 2013 Census online form displayed five sets of boxes for respondents to list their iwi without the need for a respondent to click a button to add more iwi.

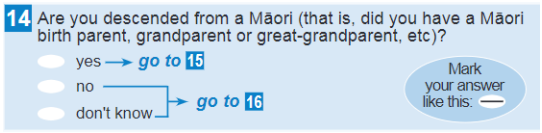
Table 2 includes a summary of the metadata for the iwi affiliation variable in the 2013 Census.

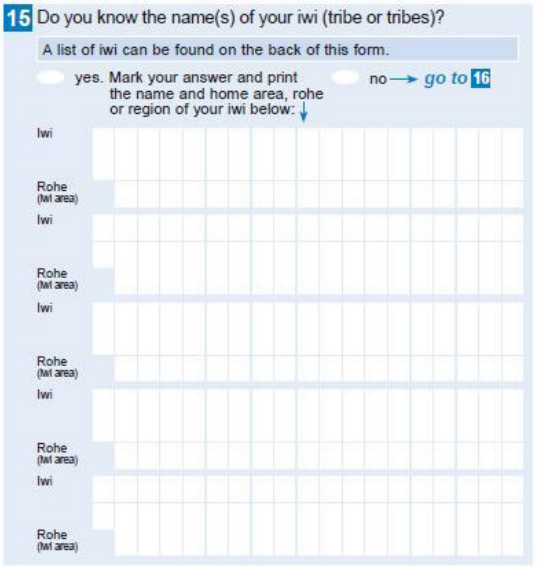
There is more information about the iwi affiliation variable in the 2013 Census at Iwi 2013 Census (information by variable and its quality).

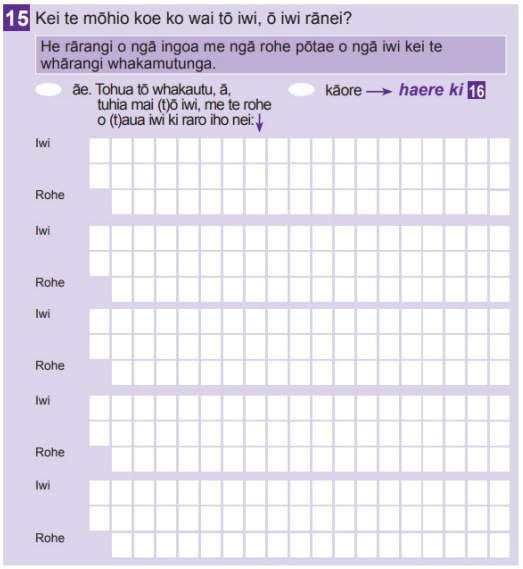
Figure 5

Māori descent and iwi affiliation questions in the 2013 Census paper and online forms

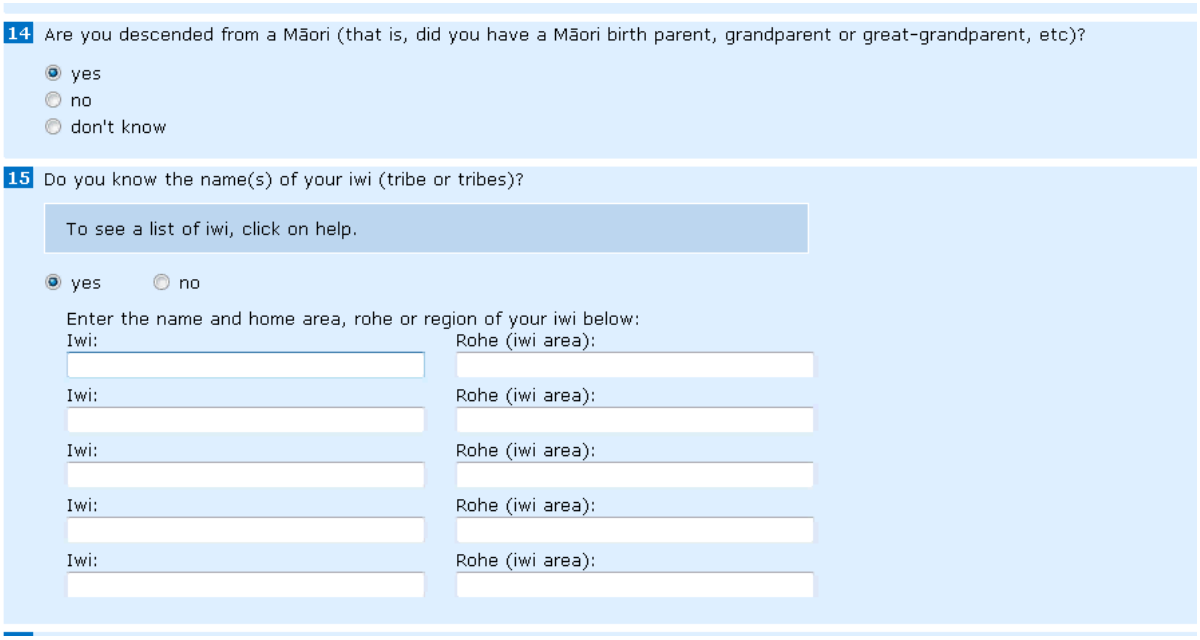
2013 Census paper forms







2013 Census online form



Iwi affiliation in the 2018 Census

People who reported ‘Yes’ or ‘Don’t know’ to the Māori descent question in the 2018 Census were asked if they knew the name(s) of their iwi. However, anyone who responded on paper could provide an iwi irrespective of their response to the Māori descent question.

The 2018 Census used paper and online forms and most responses were received online. In the 2018 Census online form, respondents could enter up to five iwi. Respondents were presented with a single iwi field with an ‘add’ button for adding more iwi. When part of an iwi name was typed in the ‘iwi’ field, an as-you-type list appeared that an iwi could be selected from, or respondents could type their iwi in full. The as-you-type list in the 2018 Census online form included 299 entries. These included common synonyms and alternative spellings for both fully and partially coded iwi in addition to the official descriptors.

Following this, a region name could be typed in the ‘Region’ field. A list of iwi was provided if the respondent clicked on a link at the top of the fields (see figure 6 below).

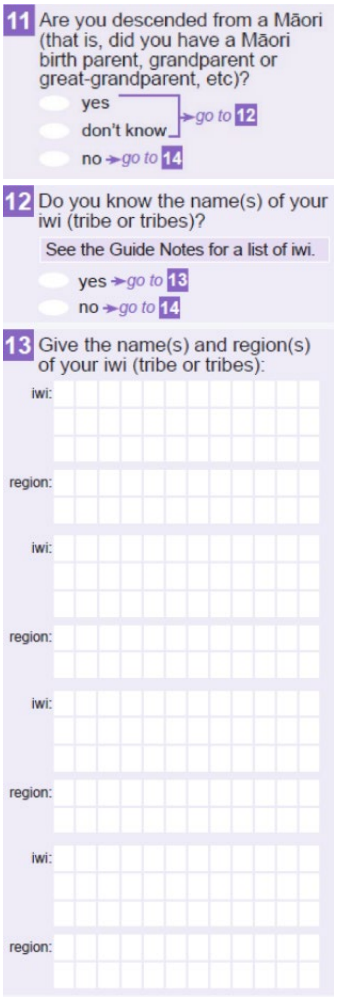
With the paper form, respondents could write in up to four iwi and were provided with the list of iwi names in the Guide Notes.

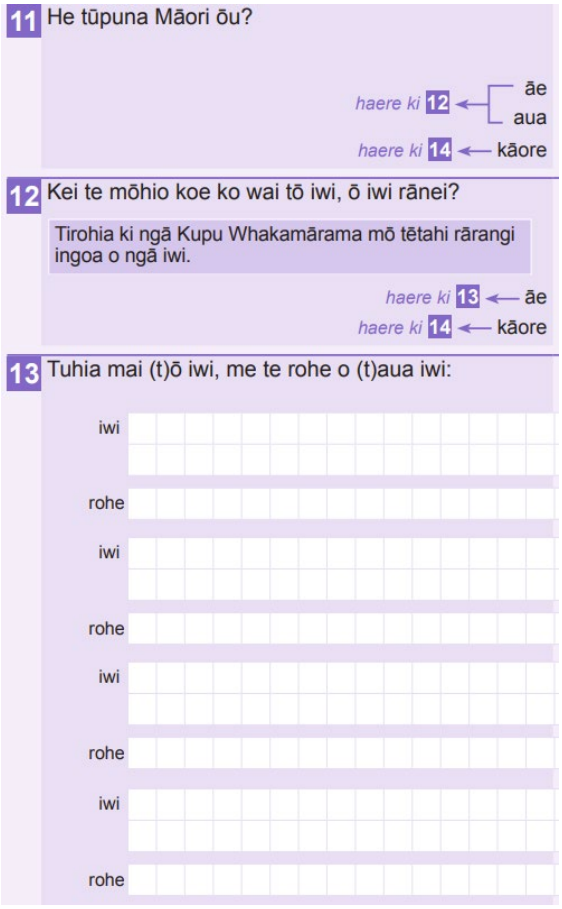
The number of iwi reported per person for the 2018 Census was lower for online responses compared with the 2018 Census paper forms and the 2013 Census, which is attributed to the form design. There is some discussion of this in Design of forms for the 2023 Census.

Figure 6

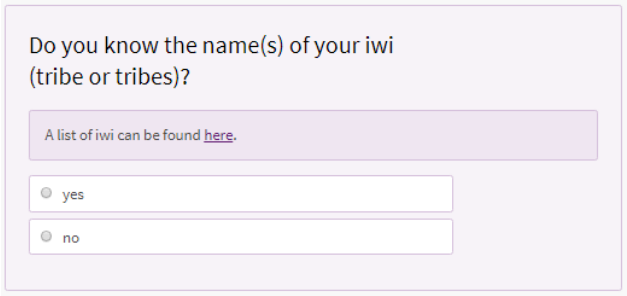
Māori descent and iwi affiliation questions in the 2018 Census paper and online forms

2018 Census paper form





Iwi affiliation questions in the 2018 Census online form



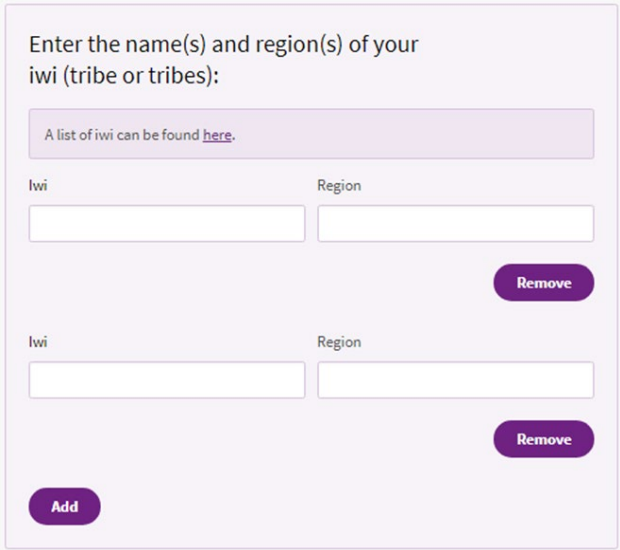


Table 2 includes a summary of the metadata for the iwi affiliation variable in the 2013 and 2018 Censuses.

Table 2

Summary of metadata for the 2013 and 2018 Censuses		
Feature	2013 Census	2018 Census
Number of iwi affiliations	Up to **five iwi** affiliations	Up to **five iwi** affiliations online, up to **four iwi** affiliations on paper
Reported by [1]	Self-reported	Self-reported
Classification [2]	Iwi - New Zealand Standard Classification V6.0.0 (2009)	Iwi and iwi-related groups statistical classification V1.0.0 (2017)
Online/Paper	Online/paper	Online/paper
Online form filling functionality	Manual entry of iwi and region	Manual entry of iwi and region including the use of as-you-type functionality with an "add" button to add additional iwi in the online form
Timestamps	Reference date	Reference date

[1] Census is in principle self-reported, but may be completed by someone else, particularly for children.

[2] The historical census data is re-coded to the 2023 classification by using the updated synonym list.

Source: Stats NZ

As stated earlier, a valid response to the iwi affiliation question is any response except those coded to 'Refused to answer', 'Response unidentifiable', 'Response outside scope', or 'Not stated'. In the 2013 Census, 96.7 percent of the Māori descent population had a valid response. Note that no alternative data sources were used for Māori descent in the 2013 Census, so this number is not directly comparable with results from the 2018 and 2023 Censuses.

In the 2018 Census, the level of collection response achieved was low, particularly for people of Māori descent (72.8 percent) (Stats NZ, 2024a). As a result, there was

an extremely high level of missing data for the iwi affiliation variable. Report of the Independent Review of New Zealand's 2018 Census examines some of the causes of the low collection response, including decisions around reduced access to paper forms, as well as a lack of meaningful engagement with iwi Māori leaders, and other communities, on how to increase engagement with census.

Even though the iwi affiliation counts were not officially published, a methodology for producing iwi affiliation estimates was jointly developed by Data ILG and Stats NZ. Methodology for the 2018 iwi affiliation estimated counts outlines its detail. Even after filling missing responses using data from the 2013 Census and using parents' responses for children under 15, there was only a valid iwi value for 84 percent of the Māori descent population (Stats NZ, 2021).

To enable production of estimated counts, extra steps were applied that reduced the impact of missing data. Weights were created to 'rate up' those with a known iwi affiliation. The purpose of this process was to increase total iwi affiliation counts to account for the missingness, and to reduce the potential of estimates being biased due to the people responding to the census having different characteristics than people who did not respond.

Data ILG and Stats NZ worked together to conclude that the approach used in the 2018 Census was not suitable for 2023 Census. It is important to note that the methodology for the 2023 Census only uses 2018 Census responses and excludes any mitigations.

There is more information about the iwi affiliation variable in the 2018 Census at Iwi (information about this variable and its quality).

Admin data

A person may provide their iwi affiliation data to different government agencies as part of interactions with them, possibly including multiple occasions for the same agency. Sometimes, a person may provide different iwi affiliation information to different agencies or at different times.

MoE Schools data, MoE Tertiary data, and MSD Student Loans and Allowances (SLA) data are the three admin data sources in the IDI that are used for the 2023 Census iwi affiliation methodology. Other data sources that collect iwi affiliation information were not considered as the context of the collection is not considered appropriate.

Table 3 shows the metadata for each admin source.

Table 3

Summary of metadata for MoE Schools, MoE Tertiary, and MSD SLA			
Feature	MoE Schools (Student Personal Information)	MoE Tertiary (enrolments)	MSD SLA
Population	Schooling ages – 5 and above	Tertiary enrolments	Eligible recipients of financial support for study
Number of iwi affiliations	Up to **three iwi** affiliations per event	Up to **three iwi** affiliations per event	Up to **five iwi** affiliations per event
Reported by	Parents/guardian	Self-reported	Self-reported
Current Classification	Iwi and iwi-related groups statistical classification V1.0.0 (2017)	Iwi and iwi-related groups statistical classification V1.0.0 (2017)	Iwi - New Zealand Standard Classification V5.0.0 (2005)
Online/Paper	Online/paper	Online/paper	Online/paper
Online Form filling functionality	Drop down list, no free text, displays iwi name that matches as you type	Can differ by institution	Drop down list, displays iwi name that matches as you type
Timestamps	Not available	Study start date, Year, Study end date	Study start date, Year, Study end date
Data available: (in October 2023 in the IDI)	From 2003 to August 2022	From 2002 to August 2022	From 1999 to July 2022

Note:
MoE Ministry of Education
MSD Ministry of Social Development
SLA Student Loans and Allowances
The current classification, online/paper and form filling functionality entries relate to the current collection system and will have been different in the past.
Source: Stats NZ

Ministry of Education Schools and Tertiary data

Out of the MoE tables available in the IDI, there are two data tables that include information about iwi affiliation. These tables are as below.

- MoE Schools: Student Personal Information table, that includes student iwi affiliation information from schools since 2003.
- MoE Tertiary: Tertiary enrolments table that includes information about tertiary students' iwi affiliation since 2002.

MoE provides guidelines on how iwi affiliation data is to be collected, along with a standard enrolment form as a guideline, but does not follow-up with individual institutions. MoE collects up to three iwi affiliations per enrolment, using a classification currently aligned to the Iwi and iwi-related groups statistical classification V1.0.0 (2017). However, previously collected MoE data has not been updated to that classification. For more information, please refer to Iwi Data: Collection and Use.

The MoE Schools table records compulsory school enrolments for each student (as filled by their parents/guardian) every time they enrol in a school (which generally starts at age 5). Hence, one person can have multiple records if they enrol more than once. The MoE Schools table does not contain a timestamp variable. This means we cannot determine how old the responses are.

In the MoE Tertiary table, iwi affiliation is collected every time a student enrols in a tertiary institution (students enrol every year of study at minimum). Hence, a student can have multiple records over the years if they enrol more than once over the duration of their certification. The data has a timestamp that helps in identifying the age of records.

Overall, the MoE collection of iwi affiliation is a well-established system that has guidelines available and that uses Stats NZ classification. However, there might be some variation in the actual practice of data collection, especially when it comes to the updates of the iwi classification.

Ministry of Social Development Student Loans and Allowances data

The Student Loans and Allowances dataset from MSD originally comprised student allowances data since 1999, and then expanded in 2000 to include student loans. Scholarship data was added subsequently. The dataset in the IDI contains information provided by applicants, education provider(s), the Tertiary Education Commission (approved study programmes or qualifications), the New Zealand Qualifications Authority (Top Scholar Awards), and MoE (TeachNZ Scholarships).

Each record contains up to five iwi affiliation entries aligned to the Iwi - New Zealand Standard Classification V5.0.0 (2005). We assume that the form is most likely filled by the students themselves, although there could be situations where a parent or whānau member fills the form. There is also the possibility of multiple records if the student is eligible for more than one monetary support item (for example, allowance, loan, or scholarship).

This data source allows for the student to retrospectively change or update their iwi affiliation information if needed. However, these changes are only provided to the IDI if made within five years of the data originally being provided. This means that if a person changes their iwi affiliation information relating to a date more than five years ago, the IDI will not have a record of the updated response. This data also has a timestamp available for understanding the recency of the data.

Overall, the MSD Student Loans and Allowances dataset has a system for collecting iwi affiliation data which allows for retrospective updates. However, the classification

currently used is not up to date with the more recent iwi affiliation classification used by Stats NZ.

Quality of alternative data sources

This section examines the quality of the alternative data sources being used. The analysis presented gives context to the decision to use these sources for the 2023 Census iwi affiliation methodology (described in the next section). As with the methodology, some of the sources are combined for this analysis. The 2013 Census and 2018 Census are combined into a historical census dataset, and the MoE Tertiary and MSD Student Loans and Allowances datasets are combined into one dataset.

To quantitatively assess the alternative data sources, we used two quality measures – coverage and consistency.

Coverage

Coverage describes the relationship between the population of interest and the people present in the dataset. Coverage of iwi affiliation for each alternative data source was calculated by dividing the number of respondents in the 2023 Census with valid responses in alternative data sources by the number of people in the 2023 Māori descent census usually resident population.

Table 4 shows coverage of alternative data sources: the historical census dataset, the combined MoE Tertiary and MSD Student Loans and Allowances dataset, and the MoE Schools data.

Table 4

Coverage of alternative data sources

Population of interest	Alternative data source	Coverage (percent)
2023 Māori descent census usually resident population	Historical census data set	70.3
	Combined MoE Tertiary and MSD SLA data set	37.4
	MoE Schools	28.7

Note:
MoE Ministry of Education
MSD Ministry of Social Development
SLA Student Loans and Allowances
Source: Stats NZ

Of those who responded that they are of Māori descent in the 2023 Census, 37.4 percent have valid iwi affiliation values (including ‘Don’t know’) in the combined MoE Tertiary and MSD Student Loans and Allowances dataset and 28.7 percent have valid iwi affiliation values (including ‘Don’t know’) in MoE Schools data. Overall, 61.6

percent of the 2023 Māori descent census usually resident population have an iwi affiliation in one of these three admin datasets. There is an analysis of the overlap of the three admin data sources in Appendix 3.

Consistency

The second measure that we used to measure the quality of alternative data sources is consistency. We looked at consistency in two ways:

- Consistency by individual – compares the iwi affiliation of a person between two sources.
- Consistency by iwi – compares the people who affiliate with a given iwi between two sources.

When calculating consistency by individual, we consider the agreement between iwi affiliations found in 2023 Census responses and alternative data sources. We do this for people in the 2023 Māori descent census usually resident population who gave a valid response to the iwi affiliation question, and who have an iwi affiliation in alternative data source. Agreement is calculated as the proportion of valid responses from the 2023 Census that are also found in alternative data sources for each individual. This methodology of measuring consistency captures how well alternative data sources could replace entire census responses and includes all valid responses (including ‘Don’t know’). Table 5 gives details of how this measure is calculated using all possible scenarios.

The consistency by individual for different sources compared with the 2023 Census is shown in table 6. The sources shown here are the combined historical census dataset, the combined MoE Tertiary and MSD Student Loans and Allowances dataset, and MoE Schools data.

The numbers presented here differ from the numbers shown in the section Assessing quality that are used to calculate the metric 1 weights, as the populations are different. The population in this section is people who both responded to the iwi affiliation question in the 2023 Census and are present in each alternative data source. The population in the calculation of the metric 1 weights only includes those people where that source is their preferred alternative data source.

Table 5

Example of the consistency by individual calculation
Calculating consistency by individual between the 2023 Census and an alternative data source

Person	2023 Census iwi affiliation responses	Alternative data source iwi affiliations	Matches	Number of 2023 Census iwi affiliations	Match rate per person
1	A	A	1	1	1
2	A, B	A, B	2	2	1
3	A	B	0	1	0
4	A, B, C, D, E	A	1	5	0.2
5	A	A, B, C, D, E	1	1	1
6	Don't know	Don't know	1	1	1
Sum of match rate per person					4.2
Number of people					6
Overall consistency score (sum of match rate per person / number of people)					0.7

Source: Stats NZ

Table 6

Consistency by individual for alternative data source, compared with the 2023 Census

Data source	Consistency by individual (percent)
Combined historical census data set	66.7
Combined MoE Tertiary and MSD SLA data set	62.6
MoE Schools	51.2

Note:
MoE Ministry of Education
MSD Ministry of Social Development
SLA Student Loans and Allowances
Source: Stats NZ

The numbers presented here differ from the numbers shown in the section Assessing quality that are used to calculate the metric 1 weights, as the populations are different. The population in this section is people who both responded to the iwi affiliation question in the 2023 Census and are present in each alternative data source. The population in the calculation of the metric 1 weights only includes those people where that source is their preferred alternative data source.

To measure consistency by iwi for an alternative data source, the population for each iwi is people in the 2023 Māori descent census usually resident population who responded to the 2023 Census iwi affiliation question with that iwi, and who are

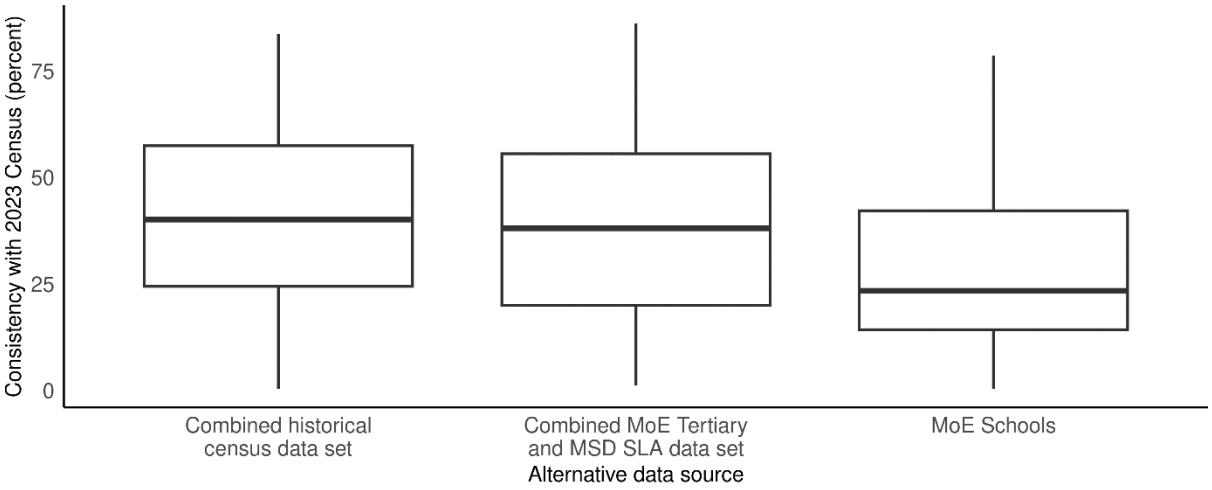
present in that alternative data source. The consistency by iwi is the proportion of that population that gave that iwi in the alternative data source. An example of this calculation is shown in table 7. Summary information showing the range of consistency scores by iwi is shown in figure 7.

Table 7

Example of the consistency by iwi calculation			
Iwi	**Iwi population in 2023 Census (excluding alternative data sources)**	**Iwi population in alternative data source who also gave that iwi in 2023 Census**	**Consistency by iwi (percent)**
A	100	30	30
B	100	0	0
C	200	200	100
Source: Stats NZ			

Figure 7

Summary of consistency scores by iwi for different alternative data sources



Source: Stats NZ

Note: This shows the upper and lower quartiles, the median, and the whiskers extend to the furthest point no more (less) than 1.5 times the inter-quartile range from the upper (lower) quartile. This excludes 'Don't know' responses.

Figure 7 shows that the distribution of consistency by iwi scores has significant overlap across all three alternative data sources, with the distribution being more similar between the historical census dataset and the combined MoE Tertiary and MSD Student Loans and Allowances dataset.

This assessment of quality of alternative data sources was used to guide the decision on the use of these sources.

2023 Census iwi affiliation methodology

In this section we outline how the decisions on the use of alternative data sources were made. We then give an overview of the methodology for iwi affiliation and describe in more detail how iwi affiliation data from the above sources is used in the methodology.

Decisions on use of alternative data sources

Data ILG technicians and Stats NZ worked together under the Mana Ōrite Relationship Agreement to develop the methodology for iwi affiliation, including the use of alternative data sources. The final decisions align with recommendations made by Data ILG and Stats NZ.

Stats NZ summarised the factors contributing to these decisions, listed below:

- How closely the classification used in alternative data sources aligns with the classification used in the 2023 Census.
- How closely the collection methodology in alternative data sources aligns with the collection methodology used in the 2023 Census.
- The recency of the collection of the alternative data sources.
- The quality (coverage and consistency) of the iwi affiliation data from the alternative data sources.

We aim to reflect the response an individual may have given if they responded to the iwi affiliation question in the 2023 Census.

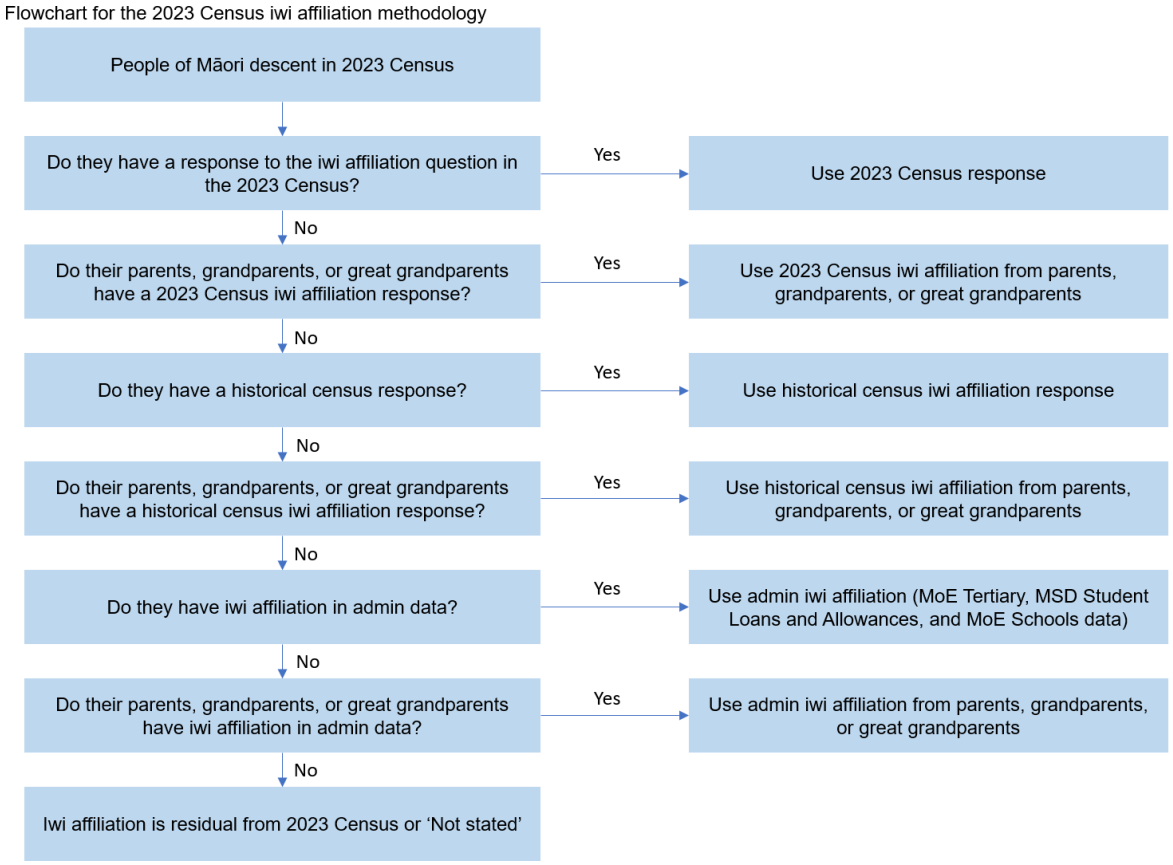
Overview of the methodology

To improve the quality of iwi counts in the 2023 Census, we used the following alternative data sources:

- Iwi affiliation data sourced from parents, grandparents, and great-grandparents. These relationships are identified using DIA birth records.
- Iwi affiliation data from the 2013 and 2018 Censuses.
- Iwi affiliation data from the last 20 years from MoE Tertiary data, MSD Student Loans and Allowances, and from MoE Schools data.

The methodology for iwi affiliation used in 2023 Census follows the flowchart shown in figure 8.

Figure 8



Note: 'Response' means any valid response, including 'Don't know'.

The data sources are ordered, and, for a given person, a valid value for iwi affiliation is searched for sequentially across the sources according to their place in the order, until one is found. If, after all the data sources have been checked, there is no valid value for iwi affiliation, then the person will either be assigned a residual value from their 2023 Census form (where applicable), or will be assigned as 'Not stated'. The ordering of the sources considered the factors listed earlier, relating to classification, collection context, recency of data, and the quality of data from the different sources. These factors relate both to how an iwi affiliation value is determined for each record, and the impact on the counts that different iwi will experience.

The methodology is:

1. Start with respondent's response to the iwi affiliation question in the 2023 Census.
2. If no valid response, use:
    a. parents'
    b. grandparents'
    c. great grandparents'

    valid iwi response from the 2023 Census.
3. If still no valid value, use combined iwi affiliation response from historical censuses.

4. If still no valid value, use:
    a. parents'
    b. grandparents'
    c. great grandparents'

    valid iwi response from historical censuses.
5. If still no valid value, use iwi affiliation value from admin data:
    a. combined MoE Tertiary and MSD Student Loans and Allowances data
    b. MoE Schools data.
6. If still no valid value, use:
    a. parents'
    b. grandparents'
    c. great grandparents'

    valid iwi affiliation value from the admin data.
7. If still no valid value, keep residual iwi affiliation response from 2023 Census or set iwi as 'Not stated'.

A 'valid' iwi affiliation value is any response except 'Refused to answer', 'Response unidentifiable', 'Response outside scope', or 'Not stated'. For the iwi affiliation variable, 'Don't know' is a valid value. Parents, grandparents, and great-grandparents are identified using DIA birth records.

The Māori descent value of parents/grandparents/great-grandparents, and the Māori descent value in a historical census, are not considered when using alternative data sources for iwi affiliation. Note that the admin sources used here do not record Māori descent.

Sourcing iwi affiliation from alternative data

Historical census data

Historical census data used in this methodology only includes responses directly from census forms. The latest synonym list is applied to all historical census responses. This means that while the collection of the data was under an older version of the iwi classification, we are able to recode the data to the latest (2023) version of the iwi classification. However, there are cases where some iwi that are newly added to the classification for the 2023 Census do not have any past synonyms as mentioned in the section Changes to the classification.

We combine iwi responses from the 2013 and 2018 Censuses in this methodology. If a 'Don't know' response is found with other iwi, then only the iwi are chosen as the final value. In all other cases, distinct iwi in the combination of historical census responses are chosen as the final value. Note that the combination of responses does not exceed 16 iwi as historical censuses allowed a maximum of five responses

each, and so the most responses an individual can have from those two sources is 10.

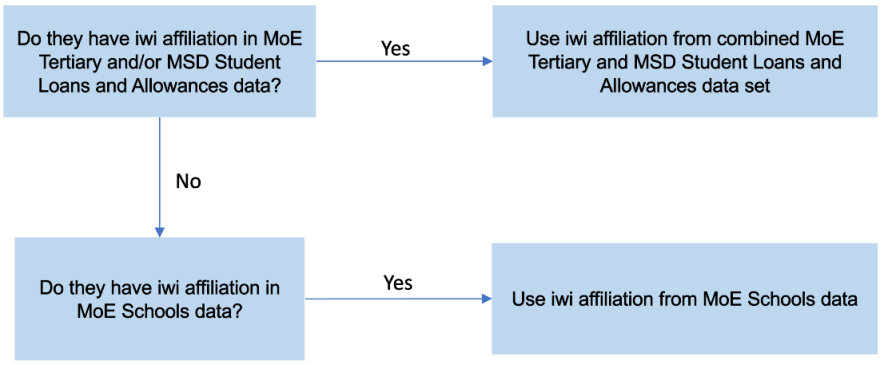
Admin data

When using admin data, we first look for a value in the MoE Tertiary data and the MSD Student Loans and Allowances data. If there is a valid value in either or both, then the combination of the two sources is used. If there is no valid value, then we look in the MoE Schools data. If there is a valid value we use that, or else we go on to the next source.

Figure 9 gives details of how iwi affiliation is found using admin data sources.

Figure 9

Details of the admin data step



Where a person has different iwi affiliations reported at different times within an admin data source, all iwi affiliations they have reported are used. We believe it is appropriate to combine responses (rather than choose the most recent response) provided by a person due to the limitations in the number of iwi affiliations that can be stated at any one time, and that it isn't appropriate for Stats NZ to determine which of a person's responses within a data source should be used.

For each of the MoE Tertiary and MSD Student Loans and Allowances data, we include data for each person if they have a record from the past 20 years, or if we are unable to determine the age of their data, some of which is missing date information. This is discussed in Appendix 2. For MoE Schools data, there is no date information, so we are unable to filter by age of data, and all data is used.

Parental data

Based on the concept of whakapapa, we expect children to have the same iwi affiliation(s) as their parents. A parental step is used to fill in missing iwi affiliations

for a person by finding their parents and providing that person with their parents' response.

To identify parents in this methodology, we use DIA birth records. For each birth record, the person born and their parent(s) are separately linked to the IDI. As birth records from 1998 and later have been fully digitised (with most records since 1985 being fully digitised), they are much more likely to link to the IDI than older records, which are only partially digitised. This means for children born during or after 1985, we have a higher chance of linking them to their parents than for children born before 1985.

The parental step has been extended to also include grandparents and great-grandparents, as follows:

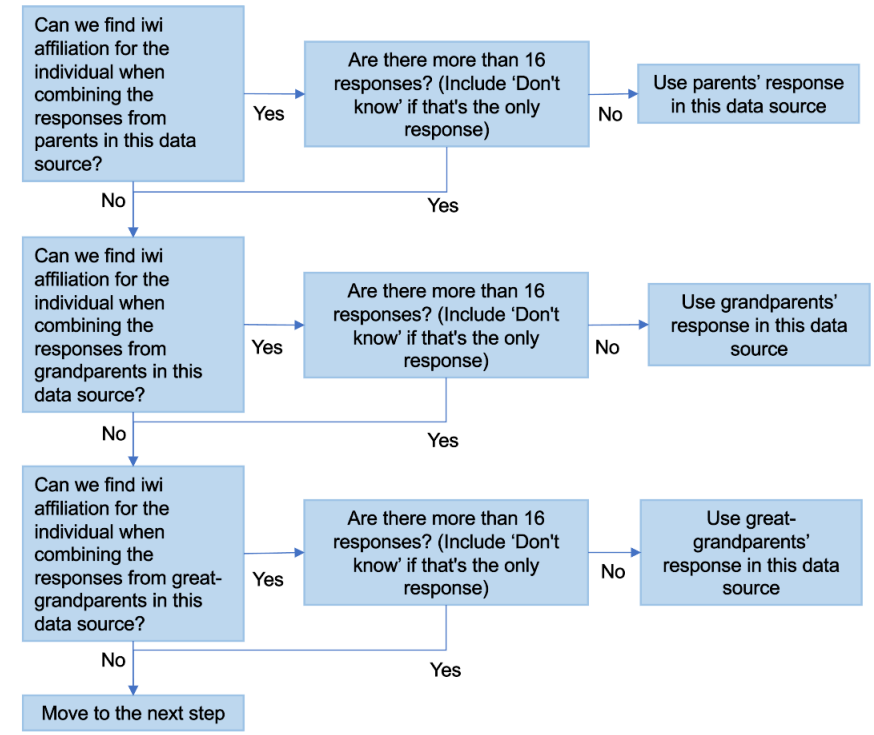
- Use DIA birth records to find the parents of the person with missing data and use their iwi affiliation data if valid.
- If the parents do not have valid data, find the grandparents (parents' parents), and use their iwi affiliation data if valid.
- If the grandparents do not have valid data, find the great-grandparents (grandparents' parents), and use their iwi affiliation data if valid.
- If the great-grandparents do not have valid data, continue with the next alternative data source in the methodology.

It is worth noting that we cannot identify cases where a child has been adopted in the DIA birth records, so there will be some instances where the parent found is not the birth parent. We do not consider the Māori descent value of the parents or child that may be present in the DIA birth records when determining parental relationships.

In each of the parental steps, we combine all valid responses (excluding residuals) across a generation. If there are valid iwi, then 'Don't know' responses are removed. This step does not require each of a person's parents, grandparents, or great-grandparents to be found in the DIA birth records or to have valid values for iwi affiliation. If any values are found for iwi affiliation, we only take values from that generation. If there are 16 or fewer distinct responses, then we use those responses. We restrict to 16 responses in line with the number of responses possible in the 2023 Census. As we are unable to determine which responses should be used if there are more than 16, we continue to the next generation (up to great-grandparents), and after that to the next source. The flowchart in figure 10 shows this.

Figure 10

Details of the parental steps



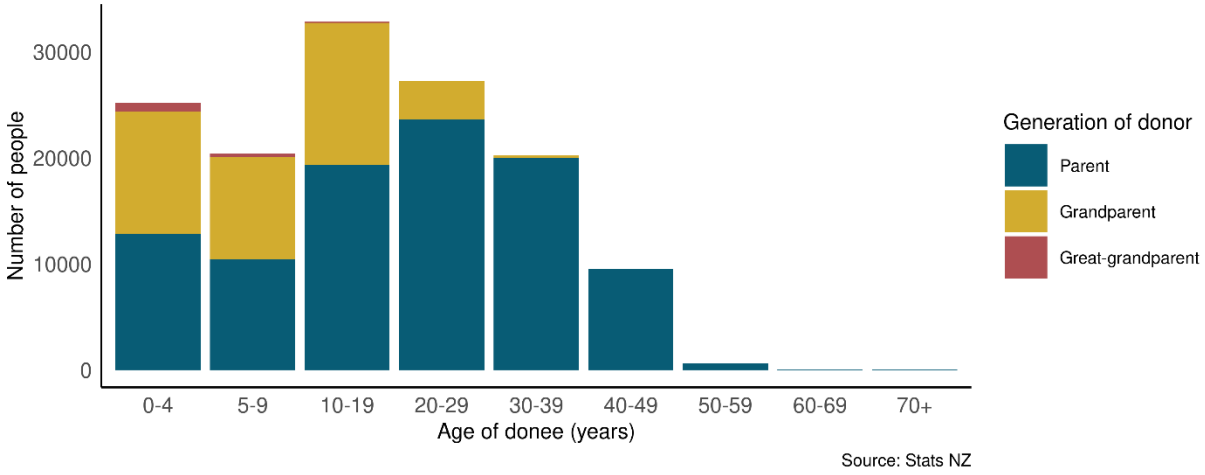
The approach of using parental data builds on the methodology used in the 2018 Census iwi estimates where a person missing iwi affiliation inherited iwi from any parents identified through Census Family and Household coding. This step was only applied to children younger than 15 years old and was limited to parents and children who lived in the same household. The 2023 Census methodology uses parental relationships from DIA birth records as they are more likely to reflect biological relationships, extend beyond the age of 14, and are not limited to people living within the same household.

While using parental steps does not cover people born overseas, it better acknowledges the parental lineage of the person and does not require that a person is living with their parent(s) to potentially benefit from their parent(s)' iwi affiliation information. The parental steps also remove the reliance on the Census Family and Household coding process, which has implications both for timeliness and accuracy.

Figure 11 shows the age distribution of people receiving iwi affiliation from the parental steps. It also shows the relationship of the donor(s) to the recipient, with most donors being parents, especially at older ages.

Figure 11

Age distribution of individuals who had their iwi affiliation filled in by the parental step, broken down by generation of donor(s)



We do not use iwi affiliation or Māori descent information from people who have passed away before 7 March 2023 (as identified using DIA death records) as that information is considered tapu and excluded from this process.

Results

This section summarises some of the results from applying the methodology. We present the proportion of data sources used in the 2023 Census, and the average number of iwi affiliations per person in records from the different data sources. We also present the distributions of the different data sources in iwi of different sizes, and for the different sources of the Māori descent variable. We also include results of the latter broken down by whether the person filled in an individual form, is on the household listing, or was admin enumerated.

Confidentiality

The 2023 Census confidentiality rules have been applied to the data in this paper. Counts are calculated using fixed random rounding to base 3 (FRR3) and suppression of ‘sensitive’ counts less than six, where tables report small populations. Individual figures may not always sum to stated totals.

Applying confidentiality rules to 2023 Census data and summary of changes since 2018 and 2013 Censuses contains more information.

Proportion of data sources used in the 2023 Census

The use of alternative data sources has increased the number of 2023 Māori descent census records that have a value for iwi affiliation, from 716,208 people (73.2 percent) to 954,042 people (97.5 percent). The alternative data source that

provides the most data is parental data from 2023 Census responses (12.1 percent), followed by historical census (8.5 percent), and admin data (1.9 percent). Table 8 provides a breakdown of the counts for each data source as well as whether the record was a census response or an admin enumeration.

Table 8

Iwi affiliation count for each data source in 2023 Census

Data source	Census individual form	Individuals on the household listing only	Admin enumeration	Total
2023 Census form	716,208	0	0	716,208
	(97.5%)	(0.0%)	(0.0%)	(73.2%)
Parental data from 2023 Census response	6,393	18,429	93,267	118,086
	(0.9%)	(45.3%)	(46.0%)	(12.1%)
Parent	5,337	14,769	60,603	80,709
	(0.7%)	(36.3%)	(29.9%)	(8.3%)
Grandparent	1,032	3,567	31,560	36,162
	(0.1%)	(8.8%)	(15.6%)	(3.7%)
Great-grandparent	24	90	1,101	1,215
	(0.0%)	(0.2%)	(0.5%)	(0.1%)
Historical census response	7,896	8,898	66,552	83,346
	(1.1%)	(21.9%)	(32.8%)	(8.5%)
Parental data from historical census response	537	1,209	14,547	16,296
	(0.1%)	(3.0%)	(7.2%)	(1.7%)
Parent	468	1,044	12,501	14,016
	(0.1%)	(2.6%)	(6.2%)	(1.4%)
Grandparent	69	156	1,995	2,220
	(0.0%)	(0.4%)	(1.0%)	(0.2%)
Great-grandparent	C	6	54	60
	(C%)	(0.0%)	(0.0%)	(0.0%)
Admin data	747	1,479	16,131	18,354
	(0.1%)	(3.6%)	(8.0%)	(1.9%)
Combined MoE Tertiary and MSD SLA data set	618	1,107	12,759	14,484
	(0.1%)	(2.7%)	(6.3%)	(1.5%)
MoE Schools	129	369	3,372	3,867
	(0.0%)	(0.9%)	(1.7%)	(0.4%)
Parental data from admin data	45	123	1,587	1,752
	(0.0%)	(0.3%)	(0.8%)	(0.2%)
Parent	39	117	1,521	1,677
	(0.0%)	(0.3%)	(0.7%)	(0.2%)
Grandparent	C	C	66	75
	(C%)	(C%)	(0.0%)	(0.0%)
Great-grandparent	C	C	C	C
	(C%)	(C%)	(C%)	(C%)

Iwi affiliation count for each data source in 2023 Census				
Data source	Census individual form	Individuals on the household listing only	Admin enumeration	Total
No information	2,922	10,530	10,749	24,201
	(0.4%)	(25.9%)	(5.3%)	(2.5%)
Total	734,745	40,668	202,833	978,246
	(100%)	(100%)	(100%)	(100%)

Note:
MoE Ministry of Education
MSD Ministry of Social Development
SLA Student Loans and Allowances
Symbol: C confidential
Source: Stats NZ

Average number of iwi affiliations per person

In the 2023 Census, the average number of iwi affiliations per person who responded to the iwi affiliation question with at least one iwi (excluding 'Don't know') was 1.7. The average number of iwi affiliations for people who reported at least one iwi reduced from 1.6 in the 2013 Census to 1.4 in the 2018 Census (responses to 2018 Census only, not counting any mitigations applied) (Stats NZ, 2021). After applying the 2023 Census iwi affiliation methodology, the average number of iwi affiliations per person in the 2023 Census file is 1.8.

The average number of iwi affiliations per person varies by the source of the iwi affiliation variable. Except for historical censuses, all the alternative data sources have a higher mean number of iwi affiliations per person than for people whose iwi affiliation is sourced from the 2023 Census. These averages are shown in table 9.

Table 9

Average number of iwi affiliations per person by source of iwi affiliation variable For people who reported at least one iwi affiliation (excluding 'Don't know')		
Iwi affiliation variable source	Number of people with at least one iwi affiliation	Average number of iwi affiliations per person
2023 Census form	598,308	1.7
Parental data from 2023 Census response	106,080	2.0
Historical census response	69,792	1.7
Parental data from historical census response	13,812	1.8
Admin data	17,811	2.0
Parental data from admin data	1,665	2.2
Total	807,468	1.8

Source: Stats NZ

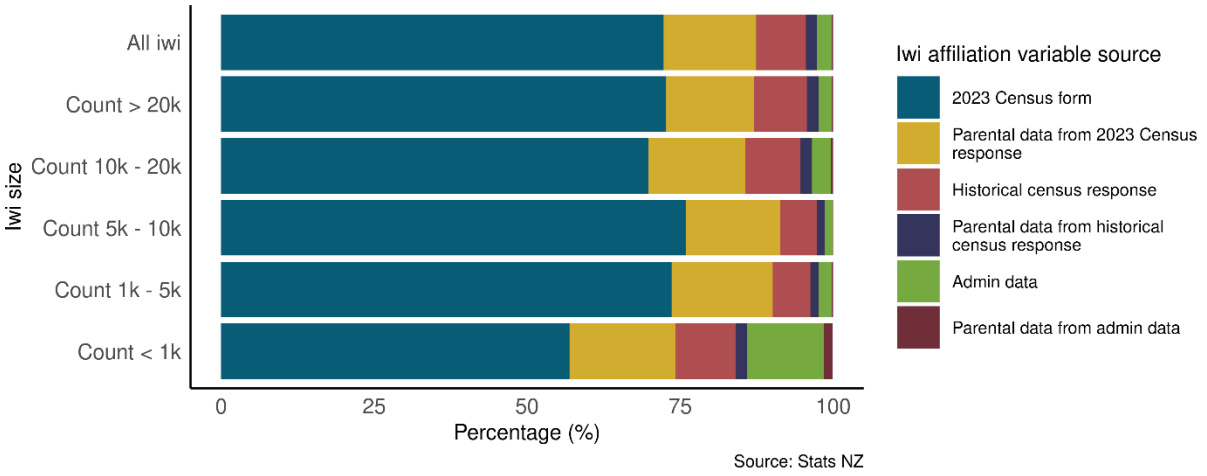
Distribution of alternative data sources in iwi of different sizes

Trends in larger iwi are likely to be comparable to overall Māori descent population trends due to the effect of their larger proportional contribution to that population than that of smaller iwi. To investigate whether there is a relationship between the size of iwi and the data sources used in the 2023 Census, we examined the proportion of counts from different data sources for iwi grouped into five broad size groupings.

Figure 12 shows that the proportion of iwi counts coming from alternative data sources is higher for smaller iwi, and this is mainly driven by admin sources. However, the proportion of alternative data sources for iwi with a count more than 1,000 remains stable at around 25 percent. The breakdown by source for different iwi sizes is similar for those iwi with counts more than 1,000, suggesting the impact on those iwi is not strongly affected by the size of the iwi.

Figure 12

Proportion of different data sources for iwi of different sizes (excluding 'Don't know' responses)

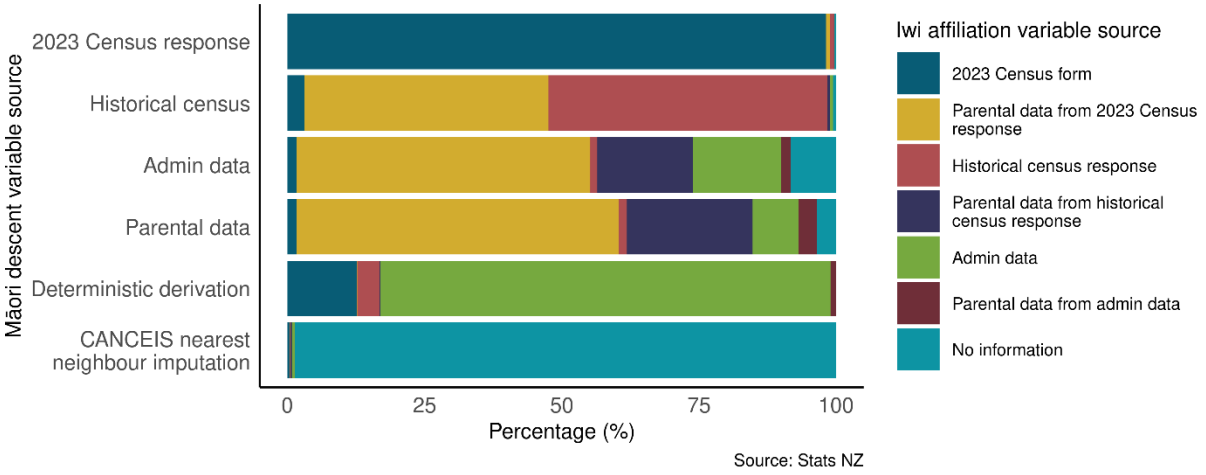


Distribution of iwi affiliation data sources by Māori descent source and unit record data source

The proportion of iwi affiliation variable source broken down by Māori descent source is shown in figure 13. This figure shows that when an individual's Māori descent source is a 2023 Census form, their iwi affiliation variable is largely sourced from 2023 Census responses as well. However, the other rows show differing breakdowns.

Figure 13

Source of iwi affiliation variable, broken down by source of Māori descent variable, for people of Māori descent

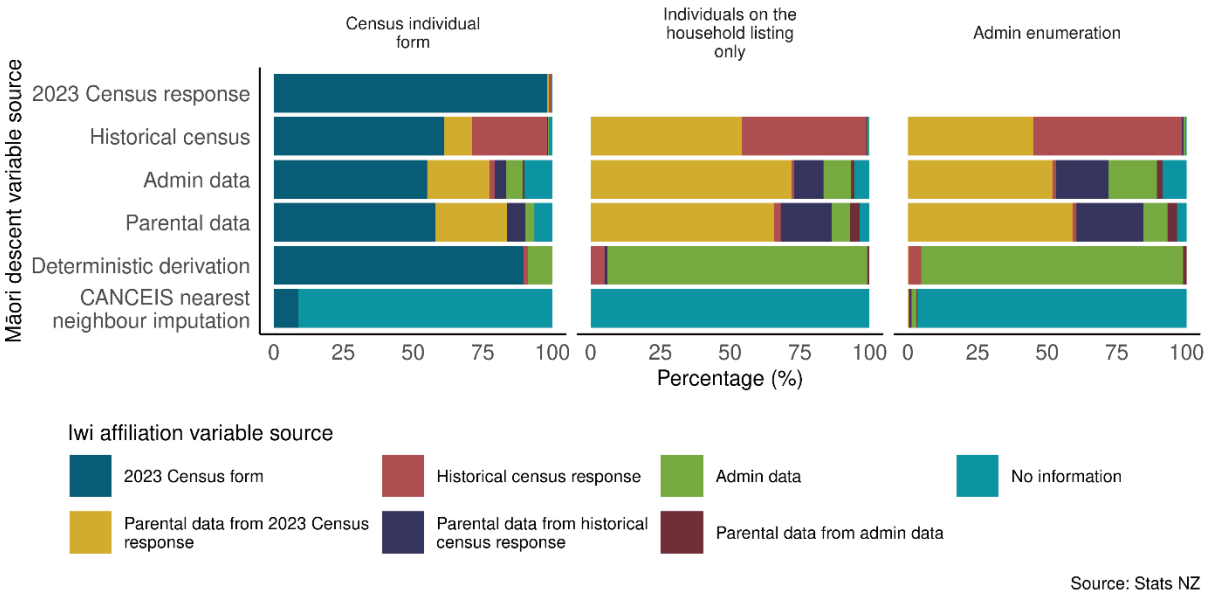


Source: Stats NZ

The above is broken down further by unit record source in figure 14. The unit record source explains how the person is added to the census file. This can be by filling in an individual form, individuals on the household listing where we did not receive an individual form, or by being admin-enumerated. The iwi affiliation question is included on the individual form, so people who are only on the household listing will not have responded to the question. Figure 14 shows that people added from the household listing have a similar source breakdown as people counted in admin enumerations. This highlights the impact different alternative data sources have on those who did not fill in an individual form.

Figure 14

Source of iwi affiliation variable, broken down by unit source and source of the Māori descent variable, for people of Māori descent



Source: Stats NZ

Assessing quality

The tables in this section show the consistency scores used in metric 1 of the quality rating scale, but without source weighting applied. There are three metrics used to assess quality in the 2023 Census, and metric 1 calculates a score by rating the overall quality of the data sources used for a census output variable. This provides one measure of the quality of the iwi affiliation variable in the 2023 Census.

Data quality assurance in the 2023 Census provides more information.

Data quality ratings for 2023 Census variables provides the final quality ratings for 2023 Census variables.

How consistency is measured

To determine a consistency score, we take people who have responded to the iwi affiliation question in the 2023 Census and who also have an iwi affiliation value in an alternative data source. This uses the same calculation as the consistency by individual in the earlier section Quality of alternative data sources with the difference that each person is only included in the calculation for one alternative data source. This is because only the value for iwi affiliation from the preferred alternative data source for each individual is used. The preferred alternative data source is the data source that would be used if they didn't have a response to the 2023 Census, and uses the logic shown in the flowchart in figure 8.

We then check what fraction of their 2023 Census responses (valid iwi and 'Don't know') were present in alternative data sources, to give a score between 0 and 1 for each person. The average of this for each alternative data source is the consistency score for that source. This calculation is the same as illustrated in table 5.

By definition, data from the 2023 Census has a consistency score of 1.00.

Consistency of alternative data sources

Table 10 shows the consistency scores of the different alternative data sources and the number of records contributing to that score. These scores are the ones used in metric 1 of the quality rating, though the rows *in italics* are not used directly in the calculations. They are shown here to provide additional breakdowns. For example, the consistency score for 'Parental data from 2023 Census response' is used in the metric 1 calculation, and the scores below for '*Parent*', '*Grandparent*', and '*Great-grandparent*' are not used individually.

Table 10

Consistency scores for each alternative data source		
Alternative data source	Number of records	Consistency score
Parental data from 2023 Census response	396,720	0.73
Parent	*379,953*	*0.74*
Grandparent	*16,554*	*0.46*
Great-grandparent	*216*	*0.32*
Historical census	253,383	0.68
Parental data from historical census response	7,956	0.47
Parent	*7,074*	*0.48*
Grandparent	*858*	*0.37*
Great-grandparent	*21*	*0.57*
Admin data	17,799	0.56
Combined MoE Tertiary and MSD SLA data set	*14,553*	*0.58*
MoE Schools	*3,249*	*0.48*
Parental data from admin data	876	0.37
Parent	*831*	*0.37*
Grandparent	*45*	*0.33*
Great-grandparent	*C*	*C*
Total	676,740	

Note:
MoE Ministry of Education
MSD Ministry of Social Development
SLA Student Loans and Allowances
Symbol: C confidential
Source: Stats NZ

Strengths and limitations of the 2023 Census iwi affiliation methodology

This section provides an overview of the data and guidance to iwi data users on how to use the counts.

Strengths

The strengths of the 2023 Census iwi affiliation methodology are summarised below.

Co-development of methodology

- The methodology was co-developed with Data ILG, resulting in a robust methodology designed with Māori data governance and Māori data sovereignty as key considerations, aiming to balance the appropriate and ethical use of data with the desire to produce meaningful data and statistics for iwi.

- Improved transparency around the methodology and treatment of data from all data sources.

Parental data

- Identifying parental relationships from DIA birth records, rather than from the Census Family and Household coding step, allows adults to be allocated a value from their parents' data, removes the restriction that individuals with missing information must live with their parents, extends to more generations (up to great-grandparent level), and better represents biological relationships.
- The use of parental 2023 Census responses reduces missing iwi responses by 12 percentage points. Use of past census data and admin data (along with the corresponding parental steps) further reduces the missingness by 10 percentage points, and 2 percentage points, respectively.
- The use of parental iwi affiliation from 2023 Census responses has been prioritised over self-responses from historical census and admin data. 2023 Census data has been prioritised as it is the most recent and it aligns with the latest iwi classification. The latter ensures that the impact on the iwi newly added to the classification in 2023 is minimised. Unlike admin data, 2023 Census parental data has also been collected in the same collection context.

Historical census

- The combination of data from the 2013 and 2018 Censuses minimises the potential bias (under-reporting of iwi) caused by the 2018 Census online form design.

Limitations

There are differences in methodology for both the Māori descent and iwi affiliation variables across the 2013, 2018, and 2023 Censuses. This, together with classification changes for the iwi affiliation variable, impact comparability of the iwi affiliation variable between censuses.

The limitations of the 2023 Census iwi affiliation methodology are outlined below.

Combining responses across all sources

- The methodology does not allow combining responses across all sources, that is, the methodology only allows for one source (or limited combinations of sources) for iwi affiliation for an individual.

Admin data

- The methodology reuses data collected by the Ministry of Education and Ministry of Social Development outside the collection context it was collected in.
- Admin data does not have raw responses from individuals filling in the forms. The data received only has the responses coded to the classification in use at the time. This means data using prior iwi classifications might not be comparable with the data in the current classification, which can have different impacts for different iwi. For example, this could lead to underrepresentation of iwi newly added to the classification from those records whose iwi affiliation is sourced from admin data, or parental data from admin data (these sources comprise 2 percent of records).

Parental data

- The methodology of deriving parental relationships from DIA birth records does not capture relationships for Māori born overseas. Additionally, there is no distinction made between adoptive parents and parents at birth.
- Please note also:
    - Whāngai (where a legal adoption has not occurred) and foster care arrangements are not reflected in the DIA birth records.
    - Parental relationships for Māori born overseas can be captured through census household forms, for those 14 years old and younger who are living with their parents (as used in the methodology for the 2018 iwi affiliation estimated counts), although adoption, whāngai, and foster care cannot be distinguished there either.

Points to note

- As noted before, care must be taken due to differences in methodology between censuses and classification changes when comparing with iwi affiliation results from previous censuses.
- This methodology has been specifically designed for use in the 2023 Census, which uses a combined model by design. Further research and input from iwi will be required to potentially create a new methodology for a future census.
- The methodology favours donated information from parental relationships in a higher priority data source (for example, the 2023 Census) over an individual's own responses in a lower priority data source (for example, historical census). As a result, a parent's 'Don't know' response can be prioritised over an

individual's own iwi affiliation response (that is not 'Don't know') from historical census data or admin data.

Overall assessment

Iwi affiliation information produced in the census is used to provide iwi, iwi-related groups, and Māori with the information they need to support the wellbeing and development of their people. It also provides the government with information to support policymaking.

For the 2023 Census, Stats NZ together with Data ILG and other groups took many measures to improve the data quality for iwi affiliation. These measures included improvements to form and question design, iwi-led collections, changes to the classification, and, for responses that could not be automatically matched, coding by the manual intervention team.

Additionally, we used alternative data sources to fill missing information for the iwi affiliation variable. This included data from historical census, admin data, and responses from parents.

These measures have increased the quality of iwi affiliation counts, which allows for meaningful cross-tabulations with other variables of interest. However, the impact of this methodology varies for different iwi. There are a number of factors that lead to this, including the size of iwi and the recency of addition to classification. For these reasons, it is imperative that Stats NZ continues to analyse and improve the methodology, build a deeper understanding of the impacts of missing iwi affiliation data, and work with iwi and Māori to develop accurate and viable long-term solutions.

Abbreviations

Data ILG	Data Iwi Leaders Group
DIA	Department of Internal Affairs
IDI	Integrated Data Infrastructure
MoE	Ministry of Education
MSD	Ministry of Social Development
SLA	Student Loans and Allowance

Glossary

Census glossary provides definitions of terms used in the 2023 Census.

References and further reading

References

Stats NZ (2016). Recommendations from the 2016 review of the statistical standard for iwi. Retrieved from cdm20045.contentdm.oclc.org.

Stats NZ (2021). Methodology for the 2018 iwi affiliation estimated counts. Retrieved from www.stats.govt.nz.

Stats NZ (2024a). Interim coverage rates, collection response rates, and data sources for the 2023 Census. Retrieved from www.stats.govt.nz.

Stats NZ (2024b). Linking 2023 Census responses to the Integrated Data Infrastructure. Retrieved from www.stats.govt.nz.

Further reading

Jack, M, & Graziadei, C (2019). Report of the Independent Review of New Zealand's 2018 Census. Retrieved from www.stats.govt.nz.

Ministry of Education (n.d.). Iwi Data: Collection and Use. Retrieved 27 June 2024 from www.educationcounts.govt.nz.

Stats NZ (2005). Iwi - New Zealand Standard Classification V5.0.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2009). Iwi - New Zealand Standard Classification V6.0.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2017). Iwi and iwi-related groups statistical classification V1.0.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2018a). 2018 Census: Design of forms. Retrieved from www.stats.govt.nz.

Stats NZ (2018b). Census Iwi grouping recode V1.0.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2020). Ngā Tikanga Paihere: a framework guiding ethical and culturally appropriate data use. Retrieved from www.data.govt.nz.

Stats NZ (2021a). 2023 Census: Final content report. Retrieved from www.stats.govt.nz.

Stats NZ (2021b). 2023 Census: High Level Design. Retrieved from www.stats.govt.nz.

Stats NZ (2021c). Iwi 2013 Census (information by variable and its quality). Retrieved from datainfoplus.stats.govt.nz.

Stats NZ (2021d). Iwi (information about this variable and its quality). Retrieved from datainfoplus.stats.govt.nz.

Stats NZ (2021e). Mana Ōrite Relationship Agreement. Retrieved from www.stats.govt.nz.

Stats NZ (2022a). Design of forms for the 2023 Census. Retrieved from www.stats.govt.nz.

Stats NZ (2022b). How we keep integrated data safe. Retrieved from www.stats.govt.nz.

Stats NZ (2023a). Census iwi and iwi-related groups V2.1.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2023b). Census Iwi grouping recode V2.1.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2023c). Iwi and iwi-related groups statistical classification V2.1.0. Retrieved from aria.stats.govt.nz.

Stats NZ (2024c). Applying confidentiality rules to 2023 Census data and summary of changes since 2018 and 2013 Censuses. Retrieved from www.stats.govt.nz.

Stats NZ (2024d). Data sources and imputation for Māori descent in the 2023 Census. Retrieved from www.stats.govt.nz.

Stats NZ (2024e). Editing, data sources, and imputation in the 2023 Census. Retrieved from www.stats.govt.nz.

Stats NZ (2024f). Iwi – 2023 Census: Information by concept. Retrieved from datainfoplus.stats.govt.nz.

Stats NZ (2024g). Measuring the success of the 2023 Census. Retrieved from www.stats.govt.nz.

Te Kāhui Raraunga (n.d.). Te Whata. Retrieved from tewhata.io.

Appendix 1: Changes to the iwi classification

The iwi classification has changed significantly over the last 20 years, and the different data sources used in the methodology are classified against different versions of the classification:

- The 2023 Census uses the Census iwi and iwi-related groups V2.1.0 classification (referred to as 'V2.1.0'), which is based on the Iwi and iwi-related groups statistical classification V2.1.0.
- The 2018 Census used the Iwi and iwi-related groups statistical classification V1.0.0. In the following tables this is referred to as 'V1.0.0'.
- The 2013 Census used the Iwi - New Zealand Standard Classification V6.0.0. In the following tables this is referred to as 'V6'.
- The current data from MoE also uses the Iwi and iwi-related groups statistical classification V1.0.0. In the following tables this is referred to as 'V1.0.0'.
- Current data from MSD uses the Iwi - New Zealand Standard Classification V5.0.0. In the following tables this is referred to as 'V5'.

In this appendix, we highlight differences between these previous classification versions and V2.1.0 of the Census iwi and iwi-related groups classification.

Iwi added to the classification

Table 11 shows the categories that have been added to the classification since the Iwi - New Zealand Standard Classification V5.0.0 (2005). This section outlines changes since that version, as it is the oldest version used in the 2023 Census iwi affiliation methodology. Therefore, we cover the classifications used in all the data sources used in the methodology. This table indicates which version of the classification each category first appeared in, for categories not present in V5 (2005). This is shown with the current (V2.1.0, 2023) code and descriptor.

Table 11

Iwi that have been added to the classification since V5 (2005)		
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Present since version
0113	Ngāti Whātua o Kaipara	V1.0.0 (2017)
0114	Ngāti Whātua o Ōrākei	V1.0.0 (2017)
0116	Ngāti Hine (Te Tai Tokerau)	V1.0.0 (2017)
0117	Te Paatu	V1.0.0 (2017)
0118	Ngāti Manuhiri	V1.0.0 (2017)
0119	Ngāti Rehua	V1.0.0 (2017)
0120	Ngāti Torehina ki Mata-ure ō Hau	V2.1.0 (2023)

Iwi that have been added to the classification since V5 (2005)		
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Present since version
0121	Ngāti Kahu ki Whangaroa	V2.1.0 (2023)
0213	Ngāti Huarere	V2.1.0 (2023)
0305	Ngāti Te Ata	V1.0.0 (2017)
0306	Ngāti Hikairo	V1.0.0 (2017)
0307	Rereahu	V1.0.0 (2017)
0308	Ngāti Tiipa	V1.0.0 (2017)
0309	Ngāti Korokī Kahukura	V1.0.0 (2017)
0310	Ngāti Tamaoho	V1.0.0 (2017)
0311	Te Ākitai-Waiohua	V1.0.0 (2017)
0312	Tainui Awhiro	V2.1.0 (2023)
0313	Ngāti Hinerangi	V2.1.0 (2023)
0413	Ngāti Mākino	V1.0.0 (2017)
0414	Ngāti Kearoa / Ngāti Tuarā	V1.0.0 (2017)
0415	Ngāti Rongomai (Te Arawa)	V1.0.0 (2017)
0511	Ngā Pōtiki ā Tamapahore	V1.0.0 (2017)
0512	Te Upokorehe	V1.0.0 (2017)
0513	Ngāti Tūwharetoa ki Kawerau	V1.0.0 (2017)
0605	Te Aitanga ā Hauiti	V1.0.0 (2017)
0712	Ngāti Hineuru	V1.0.0 (2017)
0713	Maungaharuru Tangitū	V1.0.0 (2017)
0714	Rangitāne o Tamaki nui ā Rua	V1.0.0 (2017)
0715	Ngāti Ruapani ki Waikaremoana	V1.0.0 (2017)
0716	Te Hika o Pāpāuma	V1.0.0 (2017)
0717	Ngāti Hinemanu (Heretaunga)	V2.1.0 (2023)
0905	Ngāti Whitikaupeka (Rangitīkei)	V1.0.0 (2017)
0906	Ngāi Te Ohuake (Rangitīkei)	V1.0.0 (2017)
0907	Ngāti Tamakōpiri (Rangitīkei)	V1.0.0 (2017)
0908	Ngāti Rangi (Ruapehu, Whanganui)	V1.0.0 (2017)
0909	Uenuku (Ruapehu, Waimarino)	V1.0.0 (2017)
0910	Tamahaki (Ruapehu, Waimarino)	V1.0.0 (2017)
0911	Tamakana (Ruapehu, Waimarino)	V1.0.0 (2017)
0912	Ngāti Hinemanu (Rangitīkei)	V2.1.0 (2023)
1008	Ngāti Kauwhata	V6 (2009)
1009	Ngāti Tukorehe	V1.0.0 (2017)
2018	Mōkai Pātea, iwi not named	V1.0.0 (2017)
2111	Ngāi Tai, region not known	V1.0.0 (2017)
2113	Ngāti Tūwharetoa, region not known	V1.0.0 (2017)
2114	Ngāti Hinemanu, region not known	V2.1.0 (2023)
2201	Hapū affiliated to more than one iwi	V1.0.0 (2017)
2301	Te Tai Tokerau/Tāmaki-makaurau Region, Iwi not named	V1.0.0 (2017)
2303	Waikato/Te Rohe Pōtae Region, Iwi not named	V1.0.0 (2017)
2304	Te Arawa/Taupō Region, Iwi not named	V1.0.0 (2017)

Iwi that have been added to the classification since V5 (2005)		
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Present since version
2305	Tauranga Moana/Mātaatua Region, Iwi not named	V1.0.0 (2017)
2307	Te Matau-a-Māui/Wairarapa Region, Iwi not named	V1.0.0 (2017)
2308	Taranaki Region, Iwi not named	V1.0.0 (2017)
2312	Rēkohu/Wharekauri Region, Imi/Iwi not named	V2.1.0 (2023)
Source: Stats NZ		

Iwi with code changes

Some categories have had their codes changed but are otherwise unchanged. These are shown in table 12. There were no code changes between V5 (2005) and V6 (2009).

Table 12

Iwi with code changes					
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Code in V1.0.0 (2017)	Descriptor in V1.0.0 (2017)	Code in V6 (2009)	Descriptor in V6 (2009)
0115	Ngāi Tai ki Tāmaki	0115	Ngāi Tai ki Tāmaki	0209	Ngāi Tai (Hauraki)
1201	Moriori	1105	Moriori	1105	Moriori
1202	Ngāti Mutunga (Wharekauri/Chatham Islands)	1106	Ngāti Mutunga (Wharekauri/Chatham Islands)	1106	Ngāti Mutunga (Wharekauri/Chatham Islands)
2112	Ngāti Kahungunu, region not known	2112	Ngāti Kahungunu, region not known	0705	Ngāti Kahungunu, region unspecified
2201	Hapū affiliated to more than one iwi	2201	Hapū affiliated to more than one iwi	2200	Hapū Affiliated to More Than One Iwi
2301	Te Tai Tokerau/Tāmaki-makaurau Region, Iwi not named	2301	Te Tai Tokerau/Tāmaki-makaurau Region, Iwi not named	0100	Te Tai Tokerau/Tāmaki-makaurau (Northland/Auckland) Region, not further defined
2302	Hauraki Region, Iwi not named	2302	Hauraki Region, Iwi not named	0200	Hauraki (Coromandel) Region, not further defined
2303	Waikato/Te Rohe Pōtae Region, Iwi not named	2303	Waikato/Te Rohe Pōtae Region, Iwi not named	0300	Waikato/Te Rohe Pōtae (Waikato/King Country) Region, not further defined
2304	Te Arawa/Taupō Region, Iwi not named	2304	Te Arawa/Taupō Region, Iwi not named	0400	Te Arawa/Taupō (Rotorua/Taupō) Region, not further defined
2305	Tauranga Moana/Mātaatua Region, Iwi not named	2305	Tauranga Moana/Mātaatua Region, Iwi not named	0500	Tauranga Moana/Mātaatua (Bay of Plenty) Region, not further defined

Iwi with code changes					
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Code in V1.0.0 (2017)	Descriptor in V1.0.0 (2017)	Code in V6 (2009)	Descriptor in V6 (2009)
2306	Te Tai Rāwhiti Region, Iwi not named	2306	Te Tai Rāwhiti Region, Iwi not named	0600	Te Tai Rāwhiti (East Coast) Region, not further defined
2307	Te Matau-a-Māui/Wairarapa Region, Iwi not named	2307	Te Matau-a-Māui/Wairarapa Region, Iwi not named	0700	Te Matau-a-Māui/Wairarapa (Hawke's Bay/Wairarapa) Region, not further defined
2308	Taranaki Region, Iwi not named	2308	Taranaki Region, Iwi not named	0800	Taranaki (Taranaki) Region, not further defined
2309	Whanganui/Rangitīkei Region, Iwi not named	2309	Whanganui/Rangitīkei Region, Iwi not named	0900	Whanganui/Rangitīkei (Wanganui/Rangitīkei) Region, not further defined
2310	Manawatū/Horowhenua/Te Whanganui-a-Tara Region, Iwi not named	2310	Manawatū/Horowhenua/Te Whanganui-a-Tara Region, Iwi not named	1000	Manawatū/Horowhenua/Te Whanganui-a-Tara(Manawatū/Horowhenua/Wellington)Regionnfd
2311	Te Waipounamu Region, Iwi not named	2311	Te Waipounamu/Wharekauri Region, Iwi not named	1100	Te Waipounamu/Wharekauri (South Island/Chatham Islands) Region, nfd
Source: Stats NZ					

Iwi with descriptor changes

Some categories' descriptors have changed. Often the changes are to correct misspellings, or to describe the iwi as a partially coded iwi. Table 13 shows iwi whose descriptor has changed.

Table 13

Iwi with descriptor changes				
Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Descriptor in V1.0.0 (2017)	Descriptor in V6 (2009)	Descriptor in V5 (2005)
0109	Ngāti Whātua (not Ōrākei or Kaipara)	Ngāti Whātua (not Ōrākei or Kaipara)	Ngāti Whātua	Ngāti Whātua
0110	Te Kawerau ā Maki	Te Kawerau ā Maki	Te Kawerau	Te Kawerau
0119	Ngāti Rehua	Ngati Rehua	…	…
0203	Ngāti Maru (Hauraki)	Ngāti Maru (Hauraki)	Ngāti Maru (Hauraki)	Ngāti Maru (Marutuahu)
0303	Raukawa (Waikato)	Raukawa (Waikato)	Ngāti Raukawa (Waikato)	Ngāti Raukawa (Waikato)

Iwi with descriptor changes

Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Descriptor in V1.0.0 (2017)	Descriptor in V6 (2009)	Descriptor in V5 (2005)
0306	Ngāti Hikairo	Ngāti Hīkairo	…	…
0406	Ngāti Tarāwhai (Te Arawa)	Ngāti Tarāwhai (Te Arawa)	Tarāwhai (Te Arawa)	Tarāwhai (Te Arawa)
0411	Ngāti Tūwharetoa (ki Taupō)	Ngāti Tūwharetoa (ki Taupō)	Ngāti Tūwharetoa	Ngāti Tūwharetoa
0502	Ngāi Te Rangi	Ngāi Te Rangi	Ngaiterangi	Ngaiterangi
0509	Te Whānau-ā-Apanui	Te Whānau-ā-Apanui	Te Whānau-a-Apanui	Te Whānau-a-Apanui
0805	Ngāruahine	Ngā Ruahine	Ngā Ruahine	Ngā Ruahine
0904	Ngāti Hauiti (Rangitīkei)	Ngāti Hauiti (Rangitīkei)	Ngāti Hauiti	Ngāti Hauiti
1110	Ngāti Tama ki Te Tauihu	Ngāti Tama (Te Waipounamu/South Island)	Ngāti Tama (Te Waipounamu/South Island)	Ngāti Tama (Te Waipounamu/South Island)
2001	Tainui, iwi not named	Tainui, iwi not named	Tainui	Tainui
2002	Te Arawa, iwi not named	Te Arawa, iwi not named	Te Arawa	Te Arawa
2003	Tākitimu, iwi not named	Tākitimu, iwi not named	Tākitimu	Tākitimu
2004	Aotea, iwi not named	Aotea, iwi not named	Aotea	Aotea
2005	Mātaatua, iwi not named	Mātaatua, iwi not named	Mātaatua	Mātaatua
2006	Mahuru, iwi not named	Mahuru, iwi not named	Mahuru	Mahuru
2007	Māmari, iwi not named	Māmari, iwi not named	Māmari	Māmari
2008	Ngātokimatawhaorua, iwi not named	Ngātokimatawhaorua, iwi not named	Ngātokimatawhaorua	Ngātokimatawhaorua
2009	Nukutere, iwi not named	Nukutere, iwi not named	Nukutere	Nukutere
2010	Tokomaru, iwi not named	Tokomaru, iwi not named	Tokomaru	Tokomaru
2011	Kurahaupō, iwi not named	Kurahaupō, iwi not named	Kurahaupō	Kurahaupō
2012	Muriwhenua, iwi not named	Muriwhenua, iwi not named	Muriwhenua	Muriwhenua
2013	Hauraki / Pare Hauraki, iwi not named	Hauraki / Pare Hauraki, iwi not named	Hauraki / Pare Hauraki	Hauraki / Pare Hauraki
2014	Tūranganui a Kiwa, iwi not named	Tūranganui a Kiwa, iwi not named	Tūranganui a Kiwa	Tūranganui a Kiwa
2015	Te Tauihu o Te Waka a Māui, iwi not named	Te Tauihu o Te Waka a Māui, iwi not named	Te Tauihu o Te Waka a Māui	Te Tauihu o Te Waka a Māui
2016	Tauranga Moana, iwi not named	Tauranga Moana, iwi not named	Tauranga Moana	Tauranga Moana

Iwi with descriptor changes

Code in V2.1.0 (2023)	Descriptor in V2.1.0 (2023)	Descriptor in V1.0.0 (2017)	Descriptor in V6 (2009)	Descriptor in V5 (2005)
2017	Horouta, iwi not named	Horouta, iwi not named	Horouta	Horouta
2101	Te Atiawa, region not known	Te Atiawa, region not known	Te Atiawa, region unspecified	Te Atiawa, region unspecified
2102	Ngāti Haua, region not known	Ngāti Haua, region not known	Ngāti Haua, region unspecified	Ngāti Haua, region unspecified
2103	Ngāti Maru, region not known	Ngāti Maru, region not known	Ngāti Maru, region unspecified	Ngāti Maru, region unspecified
2104	Ngāti Mutunga, region not known	Ngāti Mutunga, region not known	Ngāti Mutunga, region unspecified	Ngāti Mutunga, region unspecified
2105	Rangitāne, region not known	Rangitāne, region not known	Rangitāne, region unspecified	Rangitāne, region unspecified
2106	Ngāti Raukawa, region not known	Ngāti Raukawa, region not known	Ngāti Raukawa, region unspecified	Ngāti Raukawa, region unspecified
2107	Ngāti Tama, region not known	Ngāti Tama, region not known	Ngāti Tama, region unspecified	Ngāti Tama, region unspecified
2108	Ngāti Toa, region not known	Ngāti Toa, region not known	Ngāti Toa, region unspecified	Ngāti Toa, region unspecified
2109	Waitaha, region not known	Waitaha, region not known	Waitaha, region unspecified	Waitaha, region unspecified
2110	Ngāti Apa, region not known	Ngāti Apa, region not known	Ngāti Apa, region unspecified	Ngāti Apa, area unspecified
2311	Te Waipounamu Region, Iwi not named	Te Waipounamu/Whakaurī Region, Iwi not named	…	…

Symbol: … not applicable

Source: Stats NZ

Appendix 2: Data quality over time for MoE Tertiary and MSD Student Loans and Allowances

The MoE Tertiary and MSD Student Loans and Allowances datasets have date information for each instance iwi affiliation was collected. There is a small level of missingness of date information.

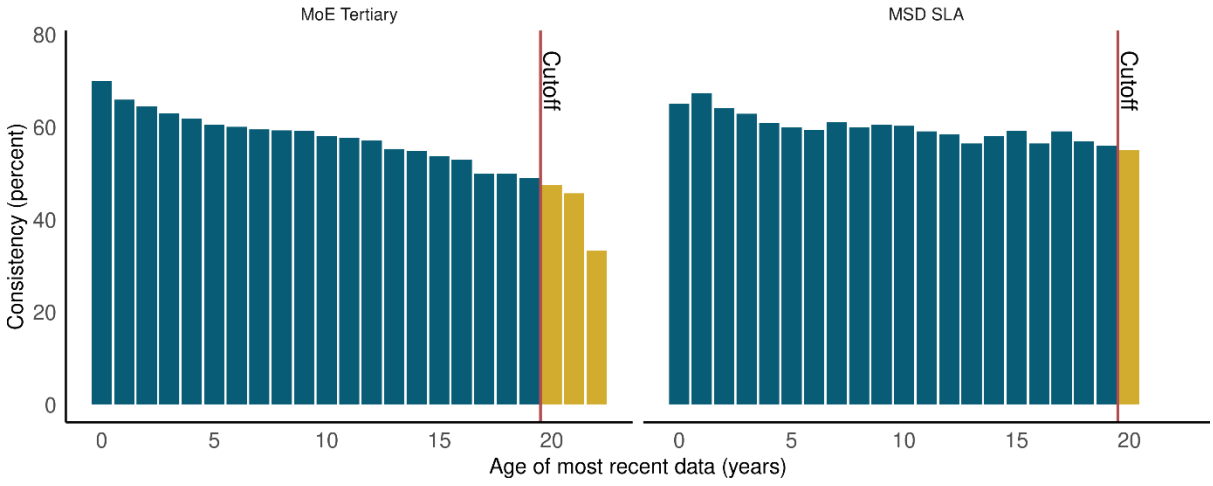
For every individual in each dataset, we define the age of their data as the age of the most-recent data for that individual. Age of data cannot be identified for individuals with missing date information.

To understand how the MoE Tertiary and MSD Student Loans and Allowances data quality varies by age, figure 15 shows the consistency by age of data. Each bar in figure 15 represents the consistency of data for people whose age of most-recent record is represented by the x-axis.

The methodology in this paper uses up to the first 20 years of age of most-recent data. This is due to the declining consistency shown in MoE Tertiary data for data older than 20 years.

Figure 15

Consistency of MoE Tertiary and MSD SLA data by age of data

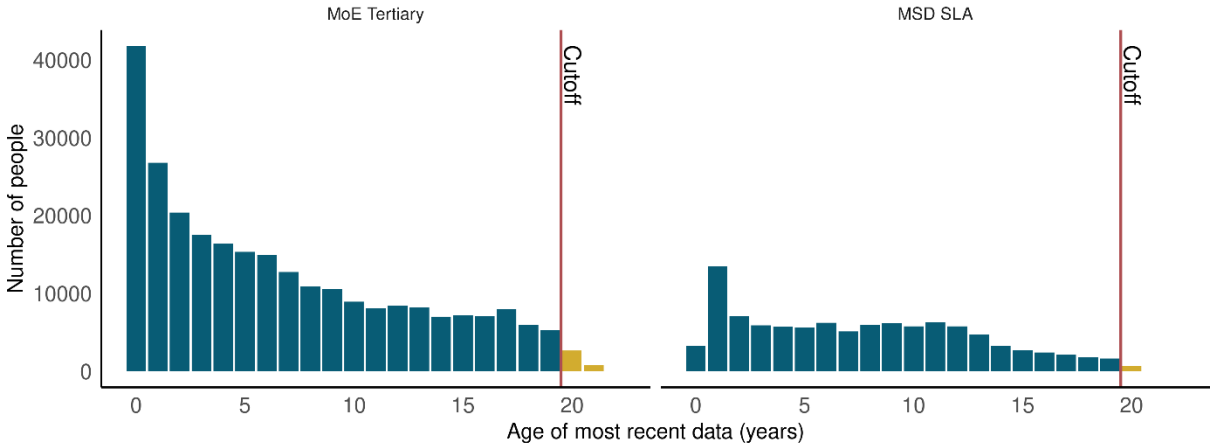


Source: Stats NZ

Figure 16 shows the count of people and the age of their most-recent data. It shows that there are few people whose age of most-recent data is more than 20 years. This implies we do not reduce the coverage of the data by introducing the age of data cut-off of 20 years.

Figure 16

Number of people in 2023 Census who have data in MoE Tertiary and MSD SLA, by age of data



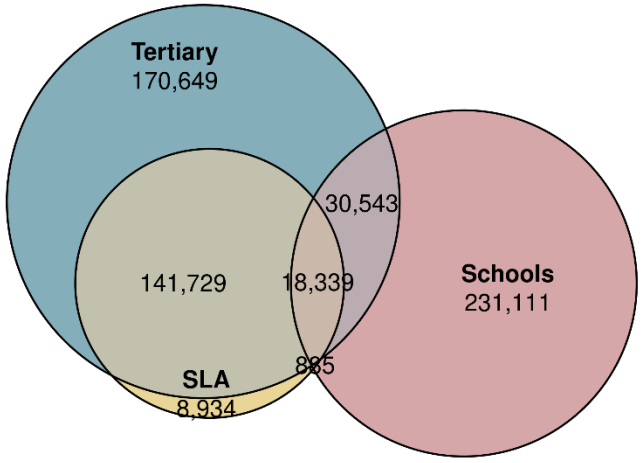
Source: Stats NZ

Appendix 3: Overlap of admin sources

The Venn diagram in figure 17 displays the overlap of MoE Tertiary, MoE Schools, and MSD Student Loan and Allowances data. The total number of people with iwi affiliation across the three datasets is 602,190, which is 61.6 percent of the 2023 Māori descent census usually resident population (978,246).

Figure 17

Venn diagram of overlap of admin sources for the 2023 Māori descent census usually resident population



Source: Stats NZ

The intersection between MSD Student Loan and Allowances data and MoE Tertiary data is high – 44.3 percent of individuals in MoE Tertiary data are also in MSD Student Loan and Allowances data, and 94.2 percent of individuals in MSD Student Loan and Allowances data are also in MoE Tertiary data. There are 160,068 people appearing in both data sources after linking with the 2023 Māori descent census usually resident population.

This high level of overlap, plus the likelihood of data being collected at the same time, means we combine iwi responses from an individual if they have iwi affiliation recorded in both MoE Tertiary data and MSD Student Loan and Allowances data. This creates in effect a ‘new’ data source – the combined MoE Tertiary and MSD Student Loans and Allowances dataset.

In contrast, there is a lower overlap with the MoE Schools data. In addition, it is very likely that the data is collected at a different time to MoE Tertiary data and MSD Student Loans and Allowances data. Hence, we do not combine MoE Schools data with the other two admin data sources.