Stats NZ
Tatauranga Aotearoa

2018 Census data user guide

New Zealand Government

Citation
Stats NZ (2019). *2018 Census data user guide*. Retrieved from www.stats.govt.nz.

ISBN 978-1-98-858361-7 (online)

Published in September 2019 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

Purpose

2018 Census data user guide provides information about collecting, processing, interpreting, and ways of using 2018 Census data.

Census processes and procedures

Below are answers to some questions people commonly ask Stats NZ about how we collect and process census data. See also Introduction to the New Zealand Census for more information.

What geographic area does the census cover?

2018 Census covered the North Island, South Island, Stewart Island, and the Chatham Islands, plus offshore islands, including the Kermadec Islands, Manawatāwhi / Three Kings Islands, Mayor Island (Tuhua), Motiti Island, White Island, Moutohora Island, Bounty Islands, Snares Islands / Tini Heke, Antipodes Islands, Auckland Islands, and Campbell Island / Motu Ihupuku.

How is everyone counted in the census?

The 2018 Census used new processes for delivering, following up, and counting everyone in New Zealand.

For more information on the changes to delivery and follow-up, see Operational phases of the 2018 Census.

For more information on how we counted everyone in New Zealand, see 2018 Census: How we combined administrative data and census forms data to create the census dataset.

Who must fill in a census form?

Everyone who is in New Zealand on census night is required to complete a census form under the Statistics Act 1975. People who do not participate in the census or who provide false information can be prosecuted.

However, the legal requirement is not itself enough to achieve the high response rate desired. For more information on communication programmes, community engagement, and targeted strategies see Communication, marketing, and engagement strategies for 2018 Census and Targeted field strategies for the 2018 Census.

How do you protect people's privacy?

We use several methods to protect the privacy and confidentiality of individuals who fill in census forms or whose administrative data has been included in the census dataset.

Under the provisions of the Statistics Act 1975, we must ensure that any statistical information published does not identify any particulars about any individual or entity. To comply with this, we make the data confidential, balancing the need to protect individuals' details while providing useful information for users. For more information on how we collect and process data, see Processing and evaluating the quality of 2018 Census data.

We also consider privacy when we add administrative data to the census dataset. For more information on the measures and actions we take to ensure privacy, see Creating the 2018 Census dataset by combining administrative data and census forms data: Our privacy impact assessment.

We usually review our confidentiality rules after each census and update them when necessary. For more information on the most recent confidentiality rules see Applying confidentiality rules to 2018 Census data and summary of changes since 2013.

How do you know how many people the census missed?

We undertake a post-enumeration survey (PES) after each census (since 1996). The PES is the official measure of how many people were missed or counted more than once in the census. The results of the 2018 PES will be available in 2020.

For 2018, we also used interim coverage rates – these are provisional and unofficial counts of how many people the census missed based on the best estimate of the population on census day. For more information on interim coverage rates and how their use compares with previous censuses, see 2018 Census: Interim coverage rates, collection response rates, and data sources.

How do you decide what topics to include in the census?

We reviewed the topics to be included in the 2018 census to better reflect our customers' information needs. We added 'Usual residence one year ago', questions on housing quality (ie 'Access to basic amenities', 'Dwelling dampness indicator', and 'Dwelling mould indicator'), and 'Main means of travel to education' and 'Educational institution address'.

We removed the question on 'Usual residence five years ago' from the form but will use administrative data to produce information on this variable.

For more information on changes to topics, including consultations and the decision-making process, see 2018 Census report on final content. For more information on how we developed questions, response options, and online and paper forms, see 2018 Census: Design of forms and Developing the forms for 2018 Census. For the English and bilingual individual and dwellings paper forms (including guide notes), see the Stats NZ Store House.

Understanding data quality

Below are answers to some questions people commonly ask Stats NZ about how we manage quality across all phases of the census; from planning, collecting, processing, and evaluating data to creating products and services, ensuring the data is 'fit for use' before we release it.

How do you manage data quality?

Census variables and topics are ranked by three 'priority levels' – priority 1, 2, and 3. We use these levels to guide the amount we spend on quality control at all phases of the census. That is, when considering quality (accuracy, relevance, timeliness, consistency, interpretability, and accessibility), time, and resources, priority 1 variables take precedence over priority 2 and priority 3 variables.

For more information on quality management, including priority rankings of variables, see 2018 Census data quality management strategy.

All variables by priority rating and quality rating are available in Data quality ratings for 2018 Census variables.

How do you ensure census data is fit for use?

For the 2018 Census, we followed a quality assurance framework to identify and assess data quality concerns. This included evaluating the variables using a quality rating scale composed of the three metrics:

- metric 1 – data sources and coverage
- metric 2 – consistency and coherence
- metric 3 – data quality.

For more information on the quality assurance framework, including the 2018 Census quality rating scale, see Data quality assurance for 2018 Census. For more information on the evaluation phase, see Processing and evaluating the quality of the 2018 Census data.

The results of the evaluation process, including a variable's quality rating, are published in 2018 Census information by variable and quality.

What are possible sources of error?

The census covers the entire population of New Zealand and is not subject to sampling error. Sampling error occurs when a sample of people in the population is surveyed and the responses from that sample are used to estimate the results of a survey of the whole population. A number of errors may be introduced due to how the sample is drawn, the sample size, and population variability.

However, census data may be subject to non-sampling errors resulting from respondent errors, collection or processing errors, and undercounts. We strive to reduce each of these error types and provide data that is fit for use.

Being self-administered (ie designed to be completed by the respondent without an interviewer), the census may be subject to errors made by individuals when filling in census forms. These could happen because individuals misunderstand the question, accidentally mark the wrong box, or give a partial response or no response to census questions that were relevant to them. To minimise these errors, census forms have been designed so that questions are as easy to understand and as simple to answer as possible. Online census forms also help minimise respondent errors by 'piping' previous responses (ie copying an earlier response – for example, copying their usual residence address across to relevant sections later in the same form so respondents don't need to type it in multiple times), directing respondents to appropriate questions, alerting the respondent when their response is not valid, and providing suggestions as respondents start typing a text response. For more information on the online form, see 2018 Census: Design of forms.

To minimise individuals intentionally distorting information, the importance of the census is communicated through a variety of media channels – such as television, radio, the internet (including social media), and newspapers – and through community engagement and targeted strategies. Guide notes (included with paper forms and available online), other online help (including e-queries), the toll-free census helpline number, and assisted completion events help individuals complete their census forms.

Collection errors are errors made when census information (including access codes) and forms are delivered to, or collected from, dwellings in the census dwelling frame (a list of all private and non-

private dwellings in New Zealand). These could include assigning a dwelling to an incorrect meshblock, misidentifying a dwelling as occupied or unoccupied, or incorrectly classifying a dwelling as private or non-private. We have checks and balances at different stages of the collection, processing, and evaluation phases to identify and fix these errors. For more information on the new processes around the census dwelling frame, see Creating the census dwelling frame for the 2018 Census.

Examples of errors that can occur during data processing include incorrectly classifying responses and misrecognising written responses (processing online forms results in fewer of these types of errors). Checks are made during data processing to identify possible errors and correct them if necessary. The data processing phase is followed by a data evaluation phase, where the data is checked further to ensure that it meets quality standards and is fit for use. More information on these phases is available in Processing and evaluating the quality of 2018 Census data. More information on the inclusion of administrative data is available in Overview of statistical methods for adding admin records to the 2018 Census dataset.

While we aim to collect information on everyone living in New Zealand, some people may be missed, and some may be counted more than once. Our collection processes seek to minimise these errors. In most censuses, more people are missed than overcounted, which results in a net undercount. This is measured through the PES (see How do you know how many people the census missed? above).

What about missing forms or questions that aren't answered?

Census aims to give complete coverage of New Zealand's population. For the 2018 Census, we used data from alternative sources to adjust for missing information, including missing forms, by including:

- real information from administrative sources
- real information from the 2013 Census
- imputed data.

For more information on these processes, see:

- 2018 Census: Technical support information
- 2018 Census: How we combined administrative data and census forms data to create the census dataset
- *Data sources, editing, and imputation for the 2018 Census* (Stats NZ, in press)
- Overview of statistical methods for adding admin records to the 2018 Census dataset
- Processing and evaluating the quality of 2018 Census data

For information on how missing information was adjusted for a specific variable, see 2018 Census information by variable and quality.

Using census data

Below are answers to some questions people commonly ask Stats NZ about using the census data.

How is census data different from survey data?

Census data is different from other survey data in a number of ways. The most important difference is that a census sets out to include information from every person in the country. Therefore, it is not subject to sampling errors that occur in other methods (see What are possible sources of error? above).

The census includes a broad range of topics providing good contextual information for individuals, families, and households, unlike other surveys, which have a narrower focus. However, in order to cover such a broad range of topics and maximise response rates, census questions are quick and simple and may not gather information in as much detail or in as much depth as other methods.

The population coverage of census means information is available for much smaller geographic areas – down to statistical area one (SA1) (see Statistical standard for geographic areas 2018) – and for small population groups, for example ethnic groups. Sample surveys only cover a small proportion of the population.

Respondents complete the forms themselves. Like other self-administered questionnaires, this can lead to more truthful responses because the interviewer cannot influence the respondent. However, respondents might not complete the form correctly, which may lead to issues or errors with the data.

For 2018 Census, we also used administrative data, historic data from the 2013 Census, and imputed data to fill in missing information. While there are improvements to the dataset, there are also some limitations for certain variables (see 2018 Census information by variable and quality for more information).

We advise you to understand the strengths and limitations of census data compared with other survey methods before deciding which to use.

Is 2018 data comparable to 2013 and 2006 Census data?

The 2018 Census included a number of changes, such as new and changed questions on the forms, different collection operations, the introduction of repatriation, and the inclusion of data sources other than census forms for the final dataset.

Due to these changes and the lower than expected response rate, time series data should be interpreted with care. We recommend users consult 2018 Census information by variable and quality for their variable of interest and consider, when performing time series comparisons and analysis, using proportions as the point of comparison rather than specific numbers or percentage change across censuses.

What does it mean for time series comparisons if you used 2013 Census data to fill in gaps in the 2018 Census?

Most of the variables that use 2013 Census data are unlikely to have changed since we collected the data. For example, we used responses in the 2013 Census to ‘Birthplace’ and ‘Years since arrival in New Zealand’ to fill in gaps in the 2018 Census. We also used the usual residence provided in the 2013 Census as the respondent’s location for ‘Usual residence five years ago’. This means that, for these variables, the inclusion of data from the 2013 Census improves the quality of the dataset (as compared with no response or imputing the response).

There are, however, some variables where the inclusion of 2013 Census data may be an issue for characteristics that can change over time. For example, we used 2013 Census data to fill in gaps for ‘Cigarette smoking behaviour’. We acknowledge that some people may have changed their smoking behaviour since the data was collected, and we note this in the 2018 Census information by variable for Cigarette smoking behaviour.

For more information on data sources and issues with comparability for specific variables, see 2018 Census information by variable and quality.

Does data quality differ depending on the data source?

During the evaluation phase of 2018 Census, we evaluated the quality of the dataset as a whole. That is, regardless of the source of the data, we completed checks and corrections for errors in the dataset to achieve the highest data quality possible. While most of the dataset comes from census forms, any administrative data, historic data, or imputed data included in the dataset has been investigated for data quality.

2018 Census information by variable and quality provides detailed information on data sources and data quality for each of the variables/topics.

For more information on the evaluation phase and administrative data, see:

- Processing and evaluating the quality of 2018 Census data
- Census transformation – research papers
- Overview of statistical methods for adding admin records to the 2018 Census dataset.

Should I treat data differently depending on its source?

Unless noted as a data quality concern for a specific variable (see 2018 Census information by variable and quality), we recommend that the dataset be treated as a whole, regardless of data source. That is, while datasets are composed of information from census forms, administrative data, historic data, and/or imputed data, we don't recommend separating the data by source unless specifically noted in the variable's information.

Can I still use the data even if the data quality is moderate / poor / very poor?

Any data that is released has gone through an evaluation phase and the quality assurance framework. That is, we have checked and corrected errors and have used the quality rating scale and quality assurance framework to determine that the data is fit for use. Where necessary, we have indicated any limitations to the use of the data in 2018 Census information by variable and quality.

For more information on the evaluation phase, see Processing and evaluating the quality of 2018 Census data.

For more information on the quality rating scale and quality assurance framework, see Data quality assurance for 2018 Census.

To what geographic levels can I get census data?

The 2018 Census uses new geographies and will be using these geographies for all 2018 Census data, as well as data from the 2013 and 2006 Censuses. For more information on the new geographies, see Statistical standard for geographic areas 2018.

Census data is available in two main ways:

- as standard published outputs available from our website, www.stats.govt.nz
- as customised data available on request, see Customised data services.

The geographic levels available from these two sources are summarised in Table 1 below.

Table 1

Geographic levels available from census data		
Geographic level	**Standard published output**	**Customised request**
Meshblock		X
Statistical area 1 (SA1)[1]	X	X
Statistical area 2 (SA2)[1]	X	X
Ward		X
Territorial authority area	X	X
Regional council area	X	X
Urban area		X
Statistical area (provincial districts)[2]		X
Regional council constituency		X
Community board		X
Auckland local board area	X	X
General electoral district	X	X
Māori electoral district	X	X
District health boards	X	X
User defined[3]		X

1. These are part of the new geographic areas, with SA2 replacing area units; for more information see Statistical standard for geographic areas 2018.
2. Statistical area (Provincial districts) V2.0.0
3. Such as police districts, radius from a specific point, any combination of standard geographies.

Source: Stats NZ

Can I get microdata / unit record level data?

In addition to standard products and customised requests, we may provide accredited researchers with access to microdata. Microdata is unit-record-level data or data corresponding to information at the respondent level. We present all statistical data in a way that does not identify the particulars about a person, dwelling, or household. This means the microdata is anonymised for use in our Data Lab facilities. The application process for Data Lab access has strict eligibility criteria based on the requirements of the Statistics Act 1975. We are still preparing the data for this use, but for more information on applying for access to microdata in the Data Lab (including costs), see Apply to use microdata for research.

Which population count should I use?

For some output variables, data about individuals/people can be reported in two ways:

- census usually resident population count
- census night population count.

Most often, the 'census usually resident population count' is used. This is the count of all people who usually live in an area of New Zealand and are present in New Zealand on census night. This count excludes visitors from overseas and residents who are temporarily overseas on census night. New Zealand residents who are away from their usual address on census night are allocated back to the address or area where they usually live and form part of the 'census usually resident population count' of the area.

The 'census night population count' is a count of all people present in a given area of New Zealand on census night. This count includes visitors from overseas who are in New Zealand on census night and people who usually live elsewhere in New Zealand but excludes New Zealand residents who are temporarily overseas on census night.

What is the difference between census counts, population estimates, and projections?

See Frequently asked questions – Population statistics for more detailed information on the differences between census counts, population estimates, and projections.

Why is the subject population important?

The subject population is the individuals, families, households, or dwellings to which variables apply. For example, while the subject population for 'Birthplace' is the census night population (as the question applies to everyone in New Zealand on census night), the subject population for 'Years since arrival in New Zealand' is the overseas born census usually resident population (as the question does not apply to people born in New Zealand or overseas visitors).

When interpreting census data, it is important for users to know what subject population the data is based on, so that any inferences drawn from that data are restricted only to that population group and not generalised outside that population group. Table 2 lists the subject population(s) for each census variable.

Table 2

Census variables/topics by subject population	
Census variable / topic	**Subject population**
Absentees	Subject populations for the absentee variables are: 'Number of absentees': occupied private dwellings 'Absentee in New Zealand on census night': absentees from occupied private dwellings 'Absentee – time away from New Zealand': absentees from occupied private dwellings who are away from New Zealand on census night
Access to telecommunications systems	Households in occupied private dwellings
Activity limitations	Census usually resident population aged 5 years and over

Age	Census night population – all people in New Zealand on census night Age is also output for the census usually resident population count.
Birthplace	Census night population Birthplace is usually output for the census usually resident population.
Census night address	Census night population
Census night population count	Census night population count
Cigarette smoking behaviour	Census usually resident population aged 15 years and over
Dwelling occupancy status	All dwellings
Dwelling type	Occupied dwellings (private and non-private)
Educational institution address	Census usually resident population studying part time or full time in any educational institute, from early education (childcare) to tertiary education
Ethnicity	Census night population Ethnicity is usually output for the census usually resident population.
Extended family type	Extended families in households in occupied private dwellings
Family type	Families in households in occupied private dwellings
Field of study	Census usually resident population aged 15 years and over
Highest secondary school qualification	Census usually resident population aged 15 years and over
Highest qualification	Census usually resident population aged 15 years and over
Hours worked in employment per week	Employed census usually resident population aged 15 years and over
Household composition	Households in private occupied dwellings (visitor-only private dwellings are excluded)
Housing quality	Occupied private dwellings
Individual home ownership	Census usually resident population aged 15 years and over
Industry	Employed census usually resident population aged 15 years and over
Iwi affiliation	Māori descent census usually resident population count Iwi affiliation data is also collected where provided for respondents who selected ‘don’t know’ at the Māori descent question but provide an iwi affiliation. However, these respondents are not included in the subject population for the Iwi affiliation variable.
Languages spoken	Census usually resident population
Main means of travel to education	Census usually resident population studying part time or full time in any educational institute, from early education (childcare) to tertiary education
Main means of travel to work	Employed census usually resident population aged 15 years and over
Main types of heating	Occupied private dwellings
Māori descent	Census usually resident population count
Number of children born	Female usually resident population aged 15 years and over

Number of motor vehicles	Households in occupied private dwellings
Number of rooms and bedrooms	Occupied private dwellings
Occupation	Employed census usually resident population aged 15 years and over
Post-school qualification	Census usually resident population aged 15 years and over
Relationship status	Census usually resident population aged 15 years and over
Religious affiliation	Census usually resident population
Sector of landlord	Households in rented occupied private dwellings
Sector of ownership	Employed census usually resident population aged 15 years and over
Sex	Census night population Sex is usually output for the census usually resident population.
Sources of personal income	Census usually resident population aged 15 and over
Status in employment	Employed census usually resident population aged 15 years and over
Study participation	Census usually resident population Study participation is also output for the census usually resident population aged 15 years and over.
Tenure of household	Households in occupied private dwellings
Total personal income	Census usually resident population aged 15 years and over
Unpaid activities	Census usually resident population aged 15 years and over
Usual residence	Census night population count and census usually resident population count However, data on usual residence is generally output for the census usually resident population.
Census usual resident population count	Census usually resident population count
Usual residence five years ago	Census usually resident population
Usual residence one year ago	Census usually resident population
Weekly rent paid by household	Households in rented occupied private dwellings
Work and labour force status	Census usually resident population aged 15 years and over
Workplace address	Employed census usually resident population aged 15 years and over
Years at usual residence	Census night population Years at usual residence is usually output for the census usually resident population.
Years since arrival in New Zealand	Overseas born census usually resident population

See 2018 Census information by variable and quality for more information on variables and subject populations, including any changes from previous censuses.

What is the difference between a dwelling and a household?

A dwelling is any building or structure – or its parts – that is used, or intended to be used, for human habitation. Dwellings can be permanent or temporary and include structures such as houses, motels, hotels, prisons, motor homes, huts, and tents.

There can be more than one dwelling within a building; for example, in an apartment building, each separate apartment or unit is considered an independent dwelling.

There are two types of dwellings:

- private (for example houses, flats, or apartments)
- non-private (for example hotels, hospitals, prisons).

'Dwellings under construction' includes all houses, flats, and groups or blocks of flats being built.

A household is either one person who usually resides alone, or two or more people who usually reside together and share facilities (such as for eating and cooking, or a living area and bathroom and toilet) in a private dwelling. Included are people who were absent on census night but usually live in a particular dwelling and are members of that household, as long as they were reported as being absent by the reference person on the dwelling form or they have completed their individual form elsewhere and included their usual residence address on that form.

Census collects information on families and households in private occupied dwellings. No family and household data is collected for non-private dwellings.

With the use of administrative data to fill in missing information, the count for 'Total households' is now equal to the count for 'Occupied private dwellings' (excluding visitor-only dwellings, that is, private dwellings that have no usual residents, such as holiday homes that were occupied on census night). See the 2018 Census information by variable for Dwelling occupancy status for more information.

How do you define occupied and unoccupied dwellings?

See the 2018 Census information by variable for Dwelling occupancy status for a detailed definition of 'occupied', 'unoccupied', and 'usually occupied', including the use of administrative data.

What is an absentee?

An absentee is identified on the census dwelling form on paper or the household set-up form online as someone who usually lives in a particular dwelling but had not completed a census individual form there – because the person was elsewhere in New Zealand or overseas on census night. Such a person may have completed a census individual form elsewhere in New Zealand. (Note: People who completed an individual form elsewhere but are not listed as an absentee at their usual residence address are not considered 'absentees'.)

Absentee data is not currently available due to data quality issues. See the 2018 Census information by variable for Absentees for more information on absentees, including changes to how we collect information and data quality issues.

What is repatriation?

On census night, most people complete their individual forms at their usual residence. Some people, however, are somewhere else in New Zealand on census night (for example, staying in a hotel, a campground, hospital, etc), and they complete their individual form at that location. In previous censuses, this meant that we would have limited information about them that would be linked to their usual residence or to their household. This is because we could only assign the information from these individual forms to a meshblock, rather than to their usual residence address.

For 2018 Census, we introduced 'repatriation'. Repatriation is the process of linking individual forms completed by people who were somewhere other than their usual residence on census night to their actual usual residence dwelling.

Some of these people are absentees, that is, listed as an absentee in the paper dwelling form or online household set-up form. In previous censuses, we would only have the absentees' age, relationship to reference person, and whether they were in New Zealand or overseas on census night (and if overseas, for how long). While this would be used for family coding, we would not have any more information about them. With repatriation, we now have all the information on their individual form linked to their actual usual residence dwelling. This includes more complete information about the ethnicity, income, and qualifications, for example, of people in a specific family/household.

Some of the people who were repatriated were not listed in the dwelling form / household set-up form as an absentee for their usual residence. These people are considered 'census night visitors', and we link their information to their actual usual residence dwelling.

The introduction of repatriation has meant that we needed to change person record types and introduce absentee flags.

What are person record types and absentee flags?

Each individual record is assigned a person record type. The person record type determines if the individual record is included in a particular subject population (for example, census usual resident population, census night population).

In previous censuses (see 2013 Census data dictionary), we had five person record types:

1	Absentee
3	New Zealand adult
4	New Zealand child
5	Overseas adult
6	Overseas child.

For the 2018 Census, we expanded the person record types to seven to provide more specific information about the absentee record types:

3	New Zealand adult
4	New Zealand child
5	Overseas adult
6	Overseas child
7	Absentee – in NZ or away from NZ < 12 months adult
8	Absentee – in NZ or away from NZ < 12 months child
9	Absentee – away from NZ >= 12 months

This is because we introduced repatriation, as discussed above. In instances where an absentee has completed an individual form elsewhere and we have linked their individual form to their actual usual residence dwelling, they would now be a person record type 3 (New Zealand adult) or 4 (New Zealand child) instead of an absentee record. They can still be identified as absentees on census night, however, by using absentee flags.

Absentee flags include:

0	Not an absentee
1	Absentee on usual residence dwelling form
2	Census night visitor not linked to usual residence dwelling form

A person who was at their usual residence on census night would have an absentee flag of '0'. A person who was listed on a dwelling form or household summary page as being a usual resident absent on census night would have an absentee flag of '1'. A person who was listed as a visitor elsewhere and provided a usual residence address but was not listed as an absentee at that address would have an absentee flag of '2'.

Absentees who completed an individual form elsewhere in New Zealand and had been repatriated would have a record type of '3' or '4' and an absentee flag of '1'. There are, however, absentees who may not have completed an individual form elsewhere or who may have been overseas at the time of the census. They will have a person record type of '7', '8', or '9', with an absentee flag of '1'.

What are derived variables?

Some census output variables are created from responses to individual questions or from a combination of responses given to two or more questions on the census forms. These are called derived variables. For example:

- age is derived from the census question on the date of birth
- years since arrival is derived from month and year first arrived in New Zealand
- work and labour force status is derived from the questions on job indicator, hours worked, seeking paid work, job search methods, and availability for work.

(For comparison, variables such as 'sex' and 'total income' are not derived – that is, we directly ask for this information in questions specific to these variables.)

Derived variables are dependent on the quality of the input variables. Any errors or issues with the input variables are likely to affect the data quality of the derived variable and may be greater when two or more census questions feed into the derived variable. See 2018 Census information by variable and quality for more information on any issues or errors with derived variables.

How are total family and household incomes worked out?

Total personal income received is the before-tax income of a person in the 12 months ended 6 March 2018. The information is collected as income bands rather than in actual dollars.

'Total family income' is derived by aggregating the total personal income of all members of the family nucleus who are aged 15 years and over. To calculate total family income, a representative income is worked out for each total personal income range. The representative value for each band is the median value (half are above and half below) for those in that band of the more detailed Household Economic Survey (HES). These median values are then added together.

Household income is calculated in a similar way to family income, except that all people in the household who are aged 15 years and over are included in the calculation.

Family and household income derivations are not included in output for the first release of Census 2018 due to data quality issues with family and household data. See also Variables of very poor quality below.

Why use income bands?

The census question that asks about the total personal income of individuals provides the respondent with a choice of income ranges or bands. This is because asking respondents to state their actual income is a sensitive issue and will often result in a higher level of non-response to the question. For consistency, data from other sources (for example, administrative data) are also allocated to a band.

Other total income variables, such as total household income and total family income, are derived from total personal income.

Why have the totals for some geographic areas from previous censuses changed with the 2018 Census?

Population changes throughout New Zealand lead to changes in geographic boundaries. This means that totals for geographic areas, for example SA1s, SA2s, and regional council areas, may change between the censuses.

We produce data from previous census years according to the current census's geographic boundaries to maintain comparability and allow time-series analysis of census data. This statistical process is called rebasing.

In the process of rebasing, each dwelling and individual within a meshblock split since the previous census is identified and allocated to the new meshblock pattern.

This allows users to compare people and dwellings in the same area between different censuses.

How should I calculate percentages?

When you calculate percentages using census data, it is important to follow these steps:

1. Ensure that the data reflects the correct subject population. For example, when calculating the percentage of regular cigarette smokers, the data needs to refer to the 'census usually resident population count' aged 15 years and over, as this is the correct subject population for this variable.
2. Use the 'total stated' population as the denominator for the calculation – this excludes residual categories ('not stated', 'refused to answer', 'response outside scope', 'response unidentifiable', and 'not elsewhere included').
3. Where a 'total stated' population specifically appears in the census table, we recommend you use this 'total stated' population as the denominator.
4. Where a 'total stated' population does not appear in the table, we recommend you calculate the 'total stated' population to use as the denominator (as in point 2) by subtracting the residual categories from the total population (people, families, households, dwellings) mentioned in the table.

5. A number of variables have categories that are valid responses and should not be excluded from the 'total stated' population. For example:
    - Number of children born – 'object to answer' is a valid response and is part of the 'total stated' population (it is a tick-box option on the form)
    - Māori descent – 'don't know' is a valid response and is part of the 'total stated' population (it is a tick-box option on the form)
    - Religious affiliation – 'no religion' and 'object to answer' are valid responses and are part of the 'total stated' population (they are tick-box options on the form).
6. Exclude 'not further defined' and 'not elsewhere included' categories from the 'total stated' population when they are used for cases where the information of interest was not provided. For example, if calculating the percentage of households who own the dwelling they live in with a mortgage, the 'dwelling owned or partly owned, mortgage arrangements not further defined' category is excluded from the calculation. This calculation is:

$$\frac{\text{owned with a mortgage}}{\text{(owned with a mortgage + owned without a mortgage)}} *100$$

7. When calculating percentage change over time, use the following formula:

$$\frac{\text{(latest year figure – base year census figure)}}{\text{base year census figure}} *100$$

Note: In published census data, percentages are usually rounded to one decimal place. When percentages are calculated for categories within total response variables (variables for which there can be more than one valid response), they will most likely add to more than 100 percent.

What is total response data?

Several census questions give individuals the option to provide more than one response. We work out the total response count or percentage by counting each response given, for example, each ethnic group stated. This means that the total response count may add up to more than the count of the subject population for that variable. When calculating percentages for categories within these variables, they will most likely add to more than 100 percent.

Variables that may be output on the basis of total responses include:

- ethnic group
- languages spoken
- iwi
- religious affiliation
- sources of personal income
- job search methods
- unpaid activities
- sources of family income
- sources of extended family income

- sources of household income
- main types of heating
- access to telecommunications systems
- access to amenities.

Total response variables can also be output as single and combined data, so that individuals or dwellings count once in the category that applies to them. For example, for 'ethnic group', the categories may be combined to be European only, European/Māori, or Māori/Pacific peoples. This means that the total population will be equal to the usual subject population for that variable, as we count individuals once only.

Examples of variables that can be output on the basis of single and combination categories are:

- ethnic group
- languages spoken
- main types of heating.

Variables of very poor quality

As discussed in the Overview of data quality ratings, interim coverage and response rates, and data sources for 2018 Census, there are three variables that have been rated as very poor quality and are not currently available:

- Absentees (people not at home on census night): Due to new methodology, absentee data is not used for population counts, but the very poor quality has had an impact on family and households data. See the 2018 Census information by variable for Absentees for details on data quality issues.
- Iwi: We are continuing to develop partnerships with iwi and Māori interest organisations to find solutions for sharing data from the census. See the 2018 Census information by variable for Iwi for details on data quality issues.
- Family and households: This suite of data is currently being investigated as part of the quality assurance framework to determine an output treatment decision. See Data quality assurance for 2018 Census and the 2018 Census information by variable for Families and households: family type and Families and households: household composition for more information.

More information about 2018 Census

The following publications offer more information on the 2018 and past censuses, as well as definitions, metadata, and other explanatory information about 2018 Census data:

- Introduction to the New Zealand Census outlines the changes to this census, including the digital-first model, new and updated questions, and how we counted people. It also gives an overview of the census, including why we conduct a census and the census cycle. See Operational phases of the 2018 Census for more information about census phases.
- Developing the forms for 2018 Census explains how we developed the forms, online help, and guide notes. A more detailed report on questions can be found in 2018 Census: Design of forms, and the 2018 Census report on final content outlines content changes.

- 2018 Census: Technical support information outlines statistical terms and processes used to measure and adjust for missing information in the census.
- 2018 Census: How we combined administrative data and census forms data to create the census dataset outlines our approach for adding administrative data records to the 2018 Census dataset. More technical information on this process is available in:
  - Overview of statistical methods for adding admin records to the 2018 Census dataset
  - Predicting the quality of admin location information for use in the 2018 Census
  - Linking 2018 Census respondents to the Integrated Data Infrastructure.

Data quality assurance for 2018 Census outlines the quality rating scale and quality assurance framework used to assess the quality of data from the 2018 Census and determine whether it is fit for purpose and suitable for release.

2018 Census information by variable and quality has information about our census data, covering matters such as non-response rates, data sources, comparability over time, and data quality.

You can also contact our Information Centre for further assistance:

Statistics NZ Information Centre: info@stats.govt.nz
Phone toll-free: 0508 525 525
Phone international: +64 4 931 4610
www.stats.govt.nz