2013 CENSUS

Statistics New Zealand
TATAURANGA AOTEAROA

# 2013 Census data user guide

New Zealand Government

**Citation**
Statistics New Zealand (2013). *2013 Census data user guide.* Available from www.stats.govt.nz.

ISBN 978-0-478-40865-2 (online)

**Published in December 2013 by**
Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

Updated 21 January 2014

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

# Contents

# List of tables

# 1 Purpose

The *2013 Census data user guide* is an easy-to-use reference for people using data from the 2013 Census of Population and Dwellings. It answers some of the common questions you may have about census data. The questions relate to the collection and processing of census data, as well as its interpretation and use.

**Note:** On 21 January 2014 we updated the variable name 'Country of birth' to 'Birthplace', to reflect the name used in 2013 Census products and services.

# 2 Background to census processes and procedures

This chapter answers some common questions about the collection and processing of census data. For more information see Introduction to the New Zealand Census.

## What areas are covered by the census?

The geographic coverage of the census includes the North Island, South Island, Stewart Island, and the Chatham Islands, plus offshore islands including the Kermadec Islands, Three Kings Islands, Mayor Island, Motiti Island, White Island, Moutohora Island, Bounty Islands, Snares Islands, Antipodes Islands, Auckland Islands, and Campbell Island.

Counts of people by sex on Ross Dependency are taken but are not included in the New Zealand population count.

## How is everyone counted in the census?

Producing high-quality data from the census starts with an efficient and successful collection process. The collection phase is a large logistical exercise that involves detailed planning, identifying and getting access to dwellings, locating respondents, and making sure respondents complete the census. There are three main elements to successful data collection:

- delivery of census forms and an Internet access code for completion online to every person and occupied dwelling in New Zealand by census night
- high response rate from those present in New Zealand on census night
- efficient collection of census forms from every person and occupied dwelling that has not completed forms online.

Statistics New Zealand recruits a team of census collectors before each census. Each collector works an area, delivering and collecting census forms to individuals and households across New Zealand. They are fully trained and supported.

The collectors leave two types of forms: a dwelling form, and an individual form for everyone who will be in the dwelling on census night. Each dwelling is provided with an Internet access code so that people can fill in the census online if they choose to. Doing the forms online is secure, quick, and easy. If everyone in a household completes their forms online, a collector may not need to call back.

Collectors make every effort to deliver census forms to all dwellings. In addition to private dwellings, collectors deliver census forms to non-private dwellings, such as hotels, motels, prisons, hospitals, camping grounds, and even cruise ships.

It can be hard to find people at home in inner-city apartments, especially in Auckland and Wellington. Collectors use strategies for apartments including identifying high-rise inner-city apartments in advance, networking with building managers to access the apartments, and using extra resources to increase awareness among residents of these apartments. It is important to make the effort to reach people in apartments, as this can have significant impacts on data quality, particularly at the area unit or meshblock levels.

People can complete census forms in English or Māori using an English form or a Māori/English form. Collectors deliver the Māori/English form to all dwellings in Northland, Whakatane, Gisborne, and the Chatham Islands. In other areas, individuals can request a Māori/English form in three ways: from the collector, by ringing the toll-free census number, or through the census website.

To help individuals fill in census forms, we:

- provide guide notes with the print and Internet census form to help respondents answer specific questions in the form, as well as questions that individuals may have while filling in the form
- set up a toll-free census helpline number well before the census date, and offer this service in eight languages: English, Māori, Samoan, Tongan, Mandarin, Cantonese, Korean, and Hindi
- offer online help through the census website.

Collection of paper census forms starts immediately after census day. Collectors return to dwellings to pick up the forms, unless everyone in a dwelling has filled in their forms online. After three attempts to contact dwelling occupants, the collector leaves a freepost envelope for the occupants to return forms.

Collectors will check (while collecting forms) that individuals have answered the key questions (foremost variables) on the first page of the individual form. This is a small but significant part of our quality management plan.

## Who must fill in a census form?

Everyone in New Zealand on census night is required to complete a census form, under the Statistics Act 1975.

We recognise that the legal requirement is not itself enough to achieve the high response rate desired. We encourage people to participate in the census through a broad public communications programme – using advertising, news media, community programmes, and social media websites. The programme aims to tell people why the census is important, when it is, and how to take part. We can prosecute individuals who do not participate in the census or provide false information.

## How is an individual's privacy protected?

We use several methods to protect the privacy and confidentiality of individuals who fill in census forms. These are in place from the time collectors deliver census forms through to the publication of census data.

2013 Census definitions and forms has more information on what happens to your census forms and how we maintain an individual’s privacy and confidentiality during the collection, transportation, processing, and storage of census forms.

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 confidentialise the data, balancing the need to protect the details of individuals while providing useful information to users. We usually review confidentiality rules after each census and update them when necessary.

2013 Census confidentiality rules and how they are applied has the most recent information.

## How do we know how many people were missed?

We undertake a Post-enumeration Survey (PES) after each census (since 1996). This estimates how many people were missed or counted more than once in the census.

The 2006 PES showed a net undercount of New Zealand residents in the country on census night of about 2 percent, which means the 2006 Census counted 98 percent of all

New Zealand residents in the country on census night, compared with 97.8 percent for the 2001 Census.

Recent census coverage results in New Zealand are similar to those in Australia and Canada. For example, in Australia, 97.3 percent of people were counted in the 2006 census and 98.2 percent in 2001 (Census of Population and Housing – Details of Undercount – Australian Bureau of Statistics, 2011).

Results from the PES are used to adjust the population base for deriving post-censal population estimates and demographic projections. However, census counts are not altered.

The results of the 2013 PES are available in 2014.

# How do topics get included in the census?

Census information covers a wide variety of topics, spanning many different areas, for example, income, education, work, housing, and culture and identity. Statutory requirements ensure certain information must be collected in each census:

- the name and address, sex, age, and ethnic origin of every occupant of the dwelling
- particulars of the dwelling – location, number of rooms, ownership, and number of occupants on census night (Statistics Act,1975).

Comparability of census information over time is one of its greatest strengths. Any change to census content reduces the ability of users to make these valuable comparisons and must be carefully considered.

New topics are added and existing questions changed to reflect real-world changes and emerging information needs. When including topics in the census, we are guided by these criteria:

- Is the census the most appropriate information source?
- Is there public acceptance of the topic?
- Will the topic produce quality information?
- Does the topic have significant community value?

There are constraints on what can and cannot be included. These are:

- some topics are required by law
- the collection of data must be cost effective
- the length of the forms is limited
- respondent burden and resistance must be minimised
- continuity and relevance need to be balanced
- the need for data consistency must be considered.

We invite submissions from the public and special interest groups and also consult users when considering changes to census content.

Following content decisions, the form-development phase involves redesigning the census forms, and balancing the requirements of content, printing, scanning, and the size of the forms.

The questions identified for redevelopment are usually those affected by data quality concerns from previous censuses (for example, growing numbers of people not

answering the question or real-world changes that require revisions to questions to reflect new legislation or changes in terminology.

For more details on the content and form design phases of the 2013 Census, see:

- 2011 Census content report
- 2011 Census form design report.

(Please note that the 2011 Census was cancelled after the Canterbury earthquake on 22 February 2011, but the content and form design was used for the 2013 Census.)

# 3 Understanding data quality

The census programme actively manages quality across all phases of the census, from planning, collection, processing, and data evaluation, to creation of products and services. This chapter explains how we ensure data is ‘fit for use’ before we release it.

## How do we manage data quality?

Census variables and topics are ranked by three 'quality levels' – foremost, defining, and supplementary. These levels are used to guide the amount of resources spent on quality control at all phases of the census.

**Level 1** – foremost variables / topics are core census variables. Their outputs are the main reason for conducting a census and include information on age, sex, ethnicity, and location. Some of these variables produce the key outputs used for maintaining the accuracy of population estimates. Across all phases of the census, foremost variables are given the highest priority in terms of quality control, time, and resources.

**Level 2** – defining variables / topics describe the key subject populations that the census provides measures for. They are important for policy development, evaluation, or monitoring. Defining variables are used frequently in cross-tabulations with foremost variables. They represent key sub-populations and the measures that are of high public interest, for example, birthplace and labour force status. These variables are closely linked to the main purpose of the census, and in the New Zealand context may only be available in detail, for example, at the subnational level. These variables have second priority in terms of quality control, time, and resources across all phases of the census.

**Level 3** – supplementary variables / topics do not fit directly with the primary purpose of the census but are important to some groups of users. Examples include occupation, language, and religious affiliation. These variables have third priority in terms of effort and resources.

A list of census output variables / topics, together with the quality level assigned to each variable, is given in chapter 5, ‘Census variables / topics by quality level’.

The evaluation phase assesses each variable’s quality for its ‘fitness for use’. Every variable needs to meet minimum standards before we release any output data.

2013 Census information by variable has information on the data quality of variables.

## How do we ensure census data is fit for use?

A comprehensive testing programme is undertaken before the census to ensure that its key processes and systems are working efficiently and to specification. The programme consists of different types of tests. These usually include form tests (cognitive and usability tests), field tests, system tests, integration tests, and a dress rehearsal.

Before a new question is included on the census form, or before a change is made to an existing question, the questions undergo cognitive testing. Cognitive testing shows how well a question meets the combined needs of respondents and users. It confirms that the question is collecting good quality data.

Field tests ensure that the census forms collect reliable information and that field procedures are working effectively. The data processing system is thoroughly tested to ensure that the system treats the responses on the census form correctly. Individual systems are tested, and then combined and tested as a group.

After the data has been processed it is evaluated before it is published. This evaluation phase ensures that the data provided by respondents has not been changed and that errors have not been introduced by the systems or processes used to input the data. The results of this evaluation are published in 2013 Census information by variable.

2011 Census testing strategy was adopted for the 2013 Census.

# What are the 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 their responses 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, 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 at all 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. Built-in editing functionality directs individuals to the appropriate questions and ensures that their responses are valid.

To minimise purposeful distortion of information by individuals, the importance of the census is communicated through a variety of media channels – such as television, radio, the Internet, and newspapers – and through a community outreach programme. Guide notes (delivered with both paper and online census forms), other online help, and the toll-free census helpline number help individuals complete their census forms.

Collection errors are errors made by collectors when forms are delivered to, or collected from, dwellings. These could include assigning dwellings 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.

Examples of errors that can occur during data processing include incorrectly classifying responses and misrecognition of written responses (processing of Internet forms is less subject to 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 further checks on the data are done to ensure that it meets quality standards and is fit for use.

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 Post-enumeration Survey discussed in How do we know how many people were missed?

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

Census aims to give complete coverage of New Zealand's population. The previous chapter explains how everyone is counted in the census. The collection phase is a key

part of the process of ensuring high-quality data. Our collection programme includes strategies for hard-to-reach groups such as those living in apartments. However, some people will not fill in census forms that have been delivered and some people will not fill in all the questions that are relevant to them.

These uncollected forms or non-responses to questions affect data quality. The higher the non-response the more it means we cannot be sure we have information on the characteristics of all the people in an area. We use statistical techniques, called imputation, to fill these data gaps with information that is most likely to have been on the forms. These imputed records are created to improve census coverage. Imputed records are included in the final population and dwelling counts.

## What is imputation?

There are two types of imputation: imputation of records to create a substitute individual or dwelling records, and imputation of variables where respondents have not answered a question.

### What are substitute records?

Substitute dwelling records are created where there is sufficient evidence that an occupied dwelling exists but we have no corresponding dwelling form. Similarly, substitute individual records are created where there is sufficient evidence that a person exists but we have no corresponding individual form.

These imputed records are created to maintain the high level of census coverage. The first step in ascertaining if there are missing dwelling forms or individual forms is to attach individual forms to dwelling forms in household groups. The dwelling form shows the number of people present in a dwelling on census night. There are three key processes within this household balancing step.

First, a three-way check is done on the number of occupants, number of names on the household grid from the dwelling form, and the number of individual forms received. Most households pass this check.

Second, collectors' field books are checked to resolve errors and flag households with unresolved errors as 'unbalanced'. Further checks, to look for missing dwelling or individual forms, extra dwelling forms, and unoccupied dwellings, result in more households being balanced.

Third, substitute dwelling and individual records are created to complete balancing.

Substitute forms have some variables imputed, but the majority are given a non-response code. Substitute dwelling forms have all variables coded to the non-response category apart from number of occupants, dwelling address, dwelling type, and dwelling occupancy status. Substitute individual forms have all variables coded to the non-response category apart from age, sex, usual residence address, census night address, and individual record type.

### How are variables imputed?

Imputation is the replacement of missing information with a best estimate of what the true value might be. The estimate is typically based on available non-missing information. In the 2013 Census, four variables are imputed on the output dataset. These are: sex, age, census usual residence meshblock, and work and labour force status.

The census uses three main types of imputation. They are:

- Deterministic imputation. This involves gathering information from other responses on the census form to determine a response to a question without a valid answer. For example, using age information on the dwelling form to impute a missing age on the individual form.

- Stochastic imputation. This involves imputing missing values according to an existing distribution. For example, using New Zealand's age distribution to impute a person who has not responded to the age question. Available information on the respondent from their census forms narrows the stochastic distribution range. For example, if a respondent lives in a rest home, they are likely to be aged over 58 years.
- Donor method imputation. This involves matching the non-respondent (recipient) to a respondent (donor) for a particular question, based on a set of matching variables that are closely related to the missing variable. The method copies the missing information from the donor to the recipient. For example, if the work and labour force status information is missing for a 35-year-old male, a male in the same age group is found as a donor of this information.

In practice, we use one or several of these imputation approaches depending on the level of information respondents provide.

## What about questions that aren't imputed?

We only impute the answer for sex, age, census usual residence meshblock, and work and labour force status, as these variables are used as the base for detailed population estimates and projections. For questions that are not answered and not imputed, we classify the answer as 'not stated'.

We calculate non-response rates by counting the number of responses entered as 'not stated' as a proportion of the subject population for that question, unless otherwise stated. For the 1981–96 Censuses, these were called 'not specified'.

High non-response rates may affect data quality. There is no standard scale for deciding what low, moderate, or high non-response rates are. The scale shown in the table below is based on the non-response rate together with information about the likely effects of different non-response rates. You can use this scale to think about the impacts of non-response on data quality.

**Effects of different non-response rates on data quality[1]**

| Non-response rate | Description of non-response level | Likely effect on data quality |
|---|---|---|
| <3.0 percent | Low | Little or none |
| 3.0–4.9 percent | Relatively low | Low |
| 5.0–6.9 percent | Moderate | Some reduction in data quality may have occurred |
| 7.0–8.9 percent | Relatively high | Some reduction in data quality is likely |
| 9.0+ percent | High | Data quality will have been reduced |

1. Adapted from: A guide to using data from the New Zealand census: 1981–2006 (Errington, Cotterell, van Randow, & Milligan, 2008).

Response rates to questions may vary. Where non-response rates are available, these are provided in 2013 Census information by variable.

We suggest you refer to the information about variables you intend to use to be aware of any data you need to interpret with caution.

# 4 Using 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 the possible sources of error?).

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 rate, 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 the meshblock and area unit levels (see Geographic definitions) – and for small population groups, for example ethnic groups. Sample surveys only cover a small proportion of the population (as outlined in chapter 3).

Respondents fill in 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 fill the form correctly, which may lead to issues or errors with the data.

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

## To what geographic levels can I get census data?

Census data is available in two main ways:

- as standard published outputs available from the Statistics NZ website
- as customised data available on request from Statistics NZ's Client Services team.

The geographic levels available from these two sources are summarised in the table below.

**Geographic levels available from census data**

| Geographic level | Standard published output | Customised request |
|---|---|---|
| Meshblock[1] | X | X |
| Area unit | X | X |
| Ward | | X |
| Territorial authority area | X | X |
| Regional council area | X | X |
| Urban area | | X |
| Statistical area | | 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[2] | | X |

1. The Meshblock dataset will be available in March 2014 for selected variables from the 2013, 2006, and 2001 Censuses, rebased to 2013 Census boundaries. These counts are at the highest level of each variable's classification. The Meshblock dataset also contains counts for area units, wards, territorial authority areas, and regional council areas.

2. Such as police districts, radius from a specific point, any combination of standard geographies.

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 identity the particulars about a person, dwelling, or household. This means that the microdata is anonymised for use in the Data Laboratory (Data Lab) facilities in Statistics NZ's offices in Auckland, Wellington, and Christchurch or by a secure remote access system. Applicants applying for Data Lab access follow an application process with strict eligibility criteria based on the requirements of the Statistics Act 1975. Find out more about access to the Data Lab and associated costs.

# 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 area where they usually live and form part of the census usually resident population count of that 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 are the impacts of the longer-than-usual time period between the 2006 and 2013 Censuses?

There was a seven-year gap between the 2013 Census and the 2006 Census because the planned 2011 Census was cancelled due to the Canterbury earthquakes. Generally, New Zealand has had a five-year gap between censuses. For example, censuses were held every five years from 1951 to 2006, although on different dates in either March or April. Before 1951, the census was held on different dates throughout the year, and there were longer gaps than five years between 1926, 1936, 1945, and 1951.

History of the census in New Zealand has more about previous censuses.

You need to be aware of the irregular time period when interpreting results. Issues to be aware of include:

- an inconsistent time series that means care is needed in comparing trends over time
- the ‘usual residence five years ago’ indicator remains the same and so, unlike with previous censuses, does not align with the last census
- there may appear to be greater changes in data due to the longer intercensal period.

Annualising intercensal changes (for example average annual change) is one approach to comparing periods of different length. You should consider adding a footnote to any tables or graphs you produce comparing the 2013 data to earlier periods to ensure your readers are aware of the longer gap.

2013 Census products, including graphs and tables with time series data, will have a footnote highlighting the seven-year gap between data points. Some 2013 Census tables included annualised change.

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

Between censuses, we prepare population estimates by age and sex to show changes annually. The estimates use the latest census data with adjustments for net census undercount, residents temporarily overseas on census night, and births, deaths, and migration since the last census. Population estimates are usually higher than the census usually resident population count because of the adjustments for net census undercount and residents who are temporarily overseas at the time of the census.

National population estimates are produced quarterly (reference dates at 31 March, 30 June, 30 September and 31 December) and provisional results are available within six weeks of the reference date.

Subnational population estimates are produced annually (reference date at 30 June) and are available in October. After each census, the population estimates for the preceding intercensal period are revised. For example, following the release of results from the 2013 Census and the 2013 Post-enumeration Survey, a new estimated resident population at 30 June 2013 is calculated. This new base is used to revise the previously published population estimates for the period 2006–13.

We also update and release population projections every two to three years to give an indication of the future size and composition of the population. Population projections are

available at national and subnational levels by age, sex, and major ethnic group. We produce multiple projection series by using different combinations of assumptions about future births (fertility), deaths (mortality), and migration. Population projections use population estimates as a starting point or base.

Find out more about population measures including estimates and projections.

## Why is the subject population important?

The subject population is the individuals, families, households, or dwellings to which variables apply.

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. Chapter 6, 'Census variables / topics by subject population', lists the subject population(s) for each census variable.

2013 Census information by variable has more on variables and their subject populations.

## 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 a 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, 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, 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.

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

## How are occupied and unoccupied dwellings defined?

Dwellings are classed as occupied if the dwelling is occupied at the midnight on the night of the census or occupied up to 12 hours after midnight of the night of the census.

A dwelling is defined as unoccupied if it was unoccupied at all times during the 12 hours after midnight on the night of the census and was suitable for habitation.

The count of occupied dwellings includes:

- private occupied dwellings such as houses, flats, and apartments

- non-private occupied dwellings such as hotels and hospitals.

## What is an absentee?

An absentee is identified on the census dwelling form as someone who usually lives in a particular dwelling, but has 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.

Included as absentees in the census are children away at boarding school, people away on business or holiday, in hospital, and so on.

Excluded are long-term hospital patients and tertiary (including university) students who live away from the dwelling for most of the year

Absentees are only recorded in dwellings where a dwelling form was completed, therefore there are no absentees recorded for unoccupied dwellings.

Statistics NZ uses information on absentees to work out the composition of households, for example the total number of absentees for each household and whether they were elsewhere in New Zealand or overseas.

The census can also tell us how long an absentee who is overseas on census night is away from New Zealand. A 'New Zealand resident temporarily overseas' is an absentee who is overseas and away from New Zealand for less than 12 months.

## 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 date of birth
- years since arrival in New Zealand is derived from month and year first arrived in New Zealand
- household composition is derived from the relationship to the reference person, absentees' relationship to the reference person, and living arrangements.

Total family and household incomes are also derived. The following section details how these variables are worked out.

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 maybe greater when two or more census questions feed into the derived variable. We comment on any issues or errors with derived variables in 2013 Census information by variable.

## 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 31 March 2013. 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.

## Why use income bands?

The census question that asks about the total personal income of individuals provides the respondent with a list of income ranges or bands to choose from. 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.

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

## Why do totals for some geographic areas from previous censuses change after the latest census?

Population changes throughout New Zealand lead to changes in geographic boundaries. This means that totals for geographic areas, for example meshblocks, area units, and local 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 is called rebasing.

In the process of rebasing, each dwelling and individual within a meshblock or area unit 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’, ‘don't know', ‘response outside scope’, ‘response unidentifiable’ and ‘not elsewhere included’).

3. Where a ‘total stated’ population specifically appears in the census table, we advise you to use this ‘total stated’ population as the denominator.

4. Where a ‘total stated’ population does not appear in the table, we advise you to calculate the total population to use as the denominator (as in in point 2), by subtracting the residual categories from the total mentioned in the table.

5. A number of variables have categories that are valid responses and should not be excluded from the total population such as:

- 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)
- iwi affiliation – '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' categories from the total 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. The calculation is:

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

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

$$\frac{\text{latest year figure – base year census figure}}{\text{base year census figure}} \text{ X } 100$$

Note that 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 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 are:

- 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

- fuel type used to heat dwellings
- access to telecommunication systems.

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
- fuel type used to heat dwellings.

# More information about the 2013 Census

The following publications give you more information on the 2013 Census and past censuses, and definitions, metadata, and other explanatory information about 2013 census data:

Introduction to the New Zealand Census has information about the 2013 Census. It tells you what happens during each phase of the census cycle, and how the census has changed over time. *Introduction to the New Zealand Census* has three parts:

- What happens to your census forms? tells you how your census forms are collected and processed after a census.
- Overview of the 2013 Census explains how we plan, conduct, advertise, and process the census.
- History of the census in New Zealand maps the history of census-taking in New Zealand.

2013 Census definitions and forms has definitions for terminology used in outputs from the census. It also provides copies of the forms from the 2013 Census, including the online form and forms for previous censuses dating back over 100 years (where available).

2013 Census data dictionary has a list of the all variable information available from the 2013 Census. The information is categorised into parts including dwelling, household, family, personal, and geographic information.

2013 Census information by variable has useful information that helps with understanding our census data, covering matters such as non-response rates, comparability over time, and data quality.

You may 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

# 5 Census variables / topics by quality level

**Census variables / topics by quality level**

| Census output variable / topic | Quality level |
|---|---|
| Absentees | Defining |
| Access to telecommunication systems | Supplementary |
| Age | Foremost |
| Birthplace | Defining |
| Census night address | Foremost |
| Census night population count | Foremost |
| Cigarette smoking behaviour | Supplementary |
| Dwelling occupancy status | Foremost |
| Ethnicity | Foremost |
| Extended family type | Defining |
| Family type | Defining |
| Fuel type used to heat dwelling | Supplementary |
| Hours worked in employment per week | Defining |
| Household composition | Defining |
| Industry | Supplementary |
| Iwi | Defining |
| Language | Supplementary |
| Relationship status (legal and social) | Defining |
| Main means of travel to work | Supplementary |
| Māori descent | Defining |
| Number of children born alive | Supplementary |
| Number of motor vehicles | Supplementary |
| Number of rooms and bedrooms | Supplementary |
| Occupation | Supplementary |
| Occupied dwelling type | Defining |
| Qualifications | Defining |
| Religious affiliation | Supplementary |
| Sector of landlord | Defining |
| Sector of ownership | Supplementary |
| Sex | Foremost |
| Sources of income | Defining |
| Status in employment | Defining |
| Study participation | Defining |
| Tenure holder | Supplementary |
| Tenure of household | Defining |
| Total income | Defining |

| | |
|---|---|
| Unpaid activities | Supplementary |
| Usual residence | Foremost |
| Usual residence five years ago | Defining |
| Weekly rent paid by household | Defining |
| Work and labour force status | Defining |
| Workplace address | Supplementary |
| Years at usual residence | Supplementary |
| Years since arrival in New Zealand | Supplementary |

# 6 Census variables / topics by subject population

**Census variables / topics by subject population**

| Census output variable / topic | Subject population |
|---|---|
| Absentees | For number of absentees: households in private occupied dwellings – private dwellings where visitors were the only occupants are excluded from this data<br>For absentee in New Zealand on census night: all absentees<br>For absentee – time away from New Zealand: absentees away from New Zealand on census night |
| Access to telecommunication systems | Households in private occupied dwellings – private dwellings where visitors were the only occupants are excluded from this data |
| Age | Census night population count<br>Census usually resident population count |
| Birthplace | Census night population count<br>Census usually resident population |
| Census night address | Census night population count |
| Census night population count | Census night population count |
| Cigarette smoking behaviour | Census usually resident population count aged 15 years and over |
| Dwelling occupancy status | All dwellings |
| Ethnicity | Census night population count<br>Census usually resident population count |
| Extended family type | Extended families in private occupied dwellings |
| Family type | Families in private occupied dwellings |
| Fuel type used to heat dwelling | Private occupied dwellings |
| Hours worked in employment per week | Employed census usually resident population count aged 15 years and over |
| Household composition | Households in private occupied dwellings – private dwellings where visitors were the only occupants are excluded from this data |
| Industry | Employed census usually resident population count aged 15 years and over |
| Iwi | Census usually resident population count who are of Māori descent |
| Language | Census usually resident population count |
| Relationship status (legal and social) | Census usually resident population count aged 15 years and over |
| Main means of travel to work | Employed census usually resident population count aged 15 years and over |
| Māori descent | Census usually resident population count |

| Census output variable / topic | Subject population |
|---|---|
| Number of children born alive | Female census usually resident population count aged 15 years and over |
| Number of motor vehicles | Households in private occupied dwellings – private dwellings where visitors were the only occupants are excluded from this data |
| Number of rooms and bedrooms | Private occupied dwellings |
| Occupation | Employed census usually resident population count aged 15 years and over |
| Occupied dwelling type | Occupied dwellings |
| Qualifications | Census usually resident population count aged 15 years and over |
| Religious affiliation | Census usually resident population count |
| Sector of landlord | Households in rented private occupied dwellings, that is, households who do not own their home or have it in a family trust and who are paying rent |
| Sector of ownership | Employed census usually resident population count aged 15 years and over |
| Sex | Census night population count<br>Census usually resident population count |
| Sources of income | For sources of personal income: census usually resident population count aged 15 years and over<br>For sources of family income: families in private occupied dwellings<br>For sources of extended family income: extended families in private occupied dwellings<br>For sources of household income: households in private occupied dwellings |
| Status in employment | Employed census usually resident population count aged 15 years and over |
| Study participation | Census usually resident population count aged 15 years and over |
| Tenure holder | Census usually resident population count aged 15 years and over |
| Tenure of household | Households in private occupied dwellings – private dwellings where visitors were the only occupants are excluded from this data |

| Census output variable / topic | Subject population |
|---|---|
| Total income | For total personal income: census usually resident population count aged 15 years and over<br>For total family income: families in private occupied dwellings<br>For combined parental income for couples with child(ren): couples with children in private occupied dwellings<br>For total extended family income: extended families in private occupied dwellings<br>For total household income: households in private occupied dwellings |
| Unpaid activities | Census usually resident population count aged 15 years and over |
| Usual residence | Census night population count<br>Census usually resident population count |
| Usual residence five years ago | Census night population count<br>Census usually resident population count |
| Weekly rent paid by household | Households in rented private occupied dwellings, that is households who do not own their home or have it in a family trust and who are paying rent |
| Work and labour force status | Census usually resident population count aged 15 years and over |
| Workplace address | Employed census usually resident population count aged 15 years and over |
| Years at usual residence | Census night population count<br>Census usually resident population count |
| Years since arrival in New Zealand | Overseas born census usually resident population count |