

Catalogue no. 98-500-X2016008
ISBN 978-0-660-05516-9

Census of Population Reference Guide

Ethnic Origin Reference Guide

Census of Population, 2016

Release date: October 25, 2017



Statistics
Canada

Statistique
Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at STATCAN.infostats-infostats.STATCAN@canada.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under “Contact us” > “Standards of service to the public.”

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard table symbols

The following symbols are used in Statistics Canada publications:

- . not available for any reference period
- .. not available for a specific reference period
- ... not applicable
- 0 true zero or a value rounded to zero
- 0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
- ^P preliminary
- ^r revised
- X suppressed to meet the confidentiality requirements of the *Statistics Act*
- ^E use with caution
- F too unreliable to be published
- * significantly different from reference category ($p < 0.05$)

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2017

All rights reserved. Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of Contents

Definitions and concepts	1
Classifications	1
Questions	2
Data quality	3
Variability due to sampling and total non-response	3
Non-response bias	3
Data quality indicators	4
Certification of final estimates	4
Comparability	5
Historical comparability	5
Comparability with other data sources	6
Data quality notes	6

Definitions and concepts

The 2016 Census of Population question on [ethnic origin](http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop030-eng.cfm) (<http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop030-eng.cfm>) collects information on the ancestral origins of the population and provides information about the composition of Canada's diverse population.

Ethnic origin refers to the ethnic or cultural origins of the person's ancestors. An 'ancestor' is usually more distant than a grandparent. Other than Aboriginal persons, most people can trace their origins to their ancestors who first came to this continent. A person may have only a single ethnic origin, or may have multiple ethnicities.

Ethnic origin refers to a person's 'roots' and should not be confused with citizenship, nationality, language or place of birth. For example, a person who has Canadian citizenship, speaks Punjabi (Panjabi) and was born in the United States may report Guyanese ethnic origin.

It is important to note that ethnic origin responses are a reflection of each respondent's perception of their ethnic ancestry. Consequently, the measurement of ethnicity is affected by changes in the social environment in which the question is asked and changes in the respondent's understanding or views about the topic. Awareness of family background or length of time since immigration can affect responses to the ethnic origin question as well.

This means that two respondents with the same ethnic ancestry could have different response patterns and thus could be counted as having different ethnic origins. For example, a respondent could report 'East Indian' as an ethnic origin while another respondent, with a similar ancestral background, could report 'Punjabi' or 'South Asian' instead. Therefore, ethnic origin data can be fluid. Nevertheless, ethnic origin data in the Census of Population are a reflection of the respondent's perception of his or her ethnic ancestry at the time of collection. Users who wish to obtain broader response estimates may wish to combine data for one or more ethnic origins together or use estimates for ethnic categories (e.g., 'South Asian origins').

In the 2016 Census, the terms 'ethnic origin,' 'ethnic group' and 'ethnic ancestry' are used interchangeably.

Classifications

Data from the ethnic origin question in the census are used to derive summary and detailed variables that provide an ethnocultural portrait of the population of Canada. The detailed list of ethnic origins disseminated in the 2016 Census and their comparability with ethnic origins from the 2011 National Household Survey and 2006 Census is available in [Appendix 5.1](http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/app-ann/a5_1-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/app-ann/a5_1-eng.cfm) of the [Dictionary, Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/index-eng.cfm) (<http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/index-eng.cfm>), Catalogue no. 98-301-X. The 2016 Census includes data for more than 250 ethnic origins reported by people living in Canada. For each ethnic origin published, total, single and multiple response counts are provided.

A single ethnic origin response occurs when a respondent provides one ethnic origin only. For example, in 2016, about 475,575 people stated that their only ethnic origin was Scottish.

A multiple response occurs when a respondent provides two or more ethnic origins. For example, in 2016, about 4,323,430 people gave a response which included Scottish and one or more other ethnic origins.

Total response counts (also called 'Total – Single and multiple ethnic origin responses' in some data tables) indicate the number of persons who reported a specified ethnic origin, either as their only ethnic origin or in addition to one or more other ethnic origins. Total responses are the sum of single and multiple responses for each ethnic origin. For example, in 2016, a total of about 4,799,010 persons reported having Scottish as an ethnic origin (the sum of the 475,575 persons who reported Scottish as their only ethnic origin and the 4,323,430 persons who reported Scottish in combination with other ethnic origins).

Ethnic Origin Reference Guide

Ethnicity is a difficult concept to measure, and there is no internationally recognized classification for this concept. In general, 2016 Census of Population data for an ethnic group are published by Statistics Canada if the count is approximately 800 or higher.

Questions

The 2016 Census of Population data on ethnic origin were obtained from Question 17 on the [2016 Census of Population 2A-L questionnaire](http://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=getInstrumentList&Item_Id=295122&UL=1V&) (http://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=getInstrumentList&Item_Id=295122&UL=1V&) or the [2016 Census of Population 2A-R questionnaire](http://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=getInstrumentList&Item_Id=295299&UL=1V&) (http://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=getInstrumentList&Item_Id=295299&UL=1V&). The 2A-L questionnaire was used to enumerate a 25% sample of all private households in Canada, except for private households on Indian reserves, Indian settlements and other remote areas, which all received the 2A-R questionnaire.

To assist respondents whose first language was neither English nor French, census questions were translated into 22 other languages, including 11 Aboriginal languages.

On both versions of the questionnaire, the 2016 Census ethnic origin question asked: 'What were the ethnic or cultural origins of this person's **ancestors**?'. A note provided above the question stated that 'This question collects information on the ancestral origins of the population and provides information about the composition of Canada's diverse population.' Below the question, a second note indicated that 'An ancestor is usually more distant than a grandparent.'

The 2016 Census 2A-L question provided the following 28 examples in following order: Canadian, English, Chinese, French, East Indian, Italian, German, Scottish, Cree, Mi'kmaq, Salish, Métis, Inuit, Filipino, Irish, Dutch, Ukrainian, Polish, Portuguese, Vietnamese, Korean, Jamaican, Greek, Iranian, Lebanese, Mexican, Somali and Colombian. In contrast, the 2A-R questionnaire provided 12 examples, starting with Canadian Aboriginal groups. The examples, in the following order, were: Cree, Ojibway, Mi'kmaq, Salish, Dene, Blackfoot, Inuit, Métis, Canadian, French, English, and German.

It is not possible to list all the ethnic or cultural origins on the census questionnaires, and examples were provided only to guide respondents as to how to answer the question. The list of examples was based on Statistics Canada's long established methodology. For the most part, the 2A-L examples were based on the most frequent single origins reported in the 2011 National Household Survey (NHS) and were arranged in order of size as reported in 2011, beginning with the largest group. Examples were also included which represent Aboriginal peoples in Canada (Cree, Mi'kmaq, Salish, Métis and Inuit). The last four examples (Lebanese, Mexican, Somali and Colombian) were included so that an example was provided for each world region, ensuring that recently arrived groups in Canada, who might not be the most numerous, were also represented in the list of examples.

Similarly, on the 2A-R questionnaire, the most frequently reported Aboriginal origins were included as examples, with an effort being made to ensure that Aboriginal examples from different regions of Canada were included. Non-Aboriginal examples in the list included the most frequently reported origins in the 2011 NHS. For more information on the Aboriginal ancestry variable in the 2016 Census of Population, please refer to the [Aboriginal Peoples Reference Guide, Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/guides/009/98-500-x2016009-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/guides/009/98-500-x2016009-eng.cfm), Catalogue no. 98-500-X2016009.

Additional instructions on how to complete the 2016 ethnic origin question were provided to respondents in the [2016 Census of Population Long-form Guide](http://www23.statcan.gc.ca/imdb-bmdi/document/3901_D18_T1_V1-eng.htm) (http://www23.statcan.gc.ca/imdb-bmdi/document/3901_D18_T1_V1-eng.htm):

This question refers to the ethnic or cultural origin or origins of a person's ancestors. Other than Aboriginal persons, most people can trace their origins to their ancestors who first came to this continent. Ancestry should not be confused with citizenship or nationality.

Ethnic Origin Reference Guide

For all persons, report the specific ethnic or cultural group or groups to which their ancestors belonged, not the language they spoke.

For persons of East Indian or South Asian origins, report a specific origin or origins. Do not report “Indian.” For example, report “East Indian from India,” “East Indian from Guyana,” or indicate the specific group, such as “Punjabi” or “Tamil.”

For persons with Aboriginal ancestors, report a specific origin or origins. For example, report “Cree,” “Mi’kmaq,” “Ojibway,” “Métis,” or “North American Indian.” Do not report “Indian.”

Information on the historical comparability of the 2016 Census ethnic origin question with questions asked in earlier censuses is provided in the section of this document entitled [Historical comparability](#).

Data quality

The 2016 long-form census questionnaire underwent a thorough data quality assessment, similar to what was done for the 2011 National Household Survey (NHS) and past censuses. A number of data quality indicators (briefly described below) were produced and used to evaluate the quality of the data.

The data quality assessment was done in addition to the regular quality checks completed at key stages of the survey. For example, during data collection and processing, the consistency of the responses provided was checked and the non-response rates for each question were analysed. As well, the quality of imputed responses was examined as part of the data editing and imputation steps. Finally, long-form census questionnaire estimates were compared with other data sources, and certified for final release.

For information about data quality for the census subdivision of Wood Buffalo, the data collection methodology and the use of administrative data sources, please refer to [Appendix 1.4](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/app-ann1-4-eng.cfm) (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/app-ann1-4-eng.cfm>) of the *Guide to the Census of Population, 2016*, Catalogue no. 98-304-X.

The main highlights of this assessment for the ethnic origin data are presented below.

Variability due to sampling and total non-response

The objective of the long-form census questionnaire is to produce estimates on various topics for a wide variety of geographies, ranging from very large areas (such as provinces and census metropolitan areas) to very small areas (such as neighbourhoods and municipalities), and for various subpopulations (such as Aboriginal peoples and immigrants) that are generally referred to as ‘domains of interest’. In order to reduce response burden, the long-form census questionnaire is administered to a random sample of households.

This sampling approach and the total non-response introduce variability in the estimates that needs to be accounted for. This variability also depends on the population size and the variability of the characteristics being measured. Furthermore, the precision of estimates may vary appreciably depending on the domain or geography of interest, in particular because of the variation in response rates. For more information on the variability due to sampling and total non-response in long-form census questionnaire estimates, please refer to the [Guide to the Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm) (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm>), Catalogue no. 98-304-X.

Non-response bias

Non-response bias is a potential source of error for all surveys, including the long-form census questionnaire. Non-response bias arises when the characteristics of those who participate in a survey are different from those who do not.

Ethnic Origin Reference Guide

In general, the risk of non-response bias increases as the response rate declines. For the 2016 long-form census questionnaire, Statistics Canada adapted its collection and estimation procedures in order to mitigate, to the extent possible, the effect of non-response bias. For more information on these mitigation strategies, please refer to the [Guide to the Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm), Catalogue no. 98-304-X.

Data quality indicators

A number of quality indicators were produced and analysed during the data quality assessment of the long-form census questionnaire. Three of these are presented to users: the global non-response rate (GNR), the standard error, and the imputation rate by question.

The GNR combines non-response at the household level (or total non-response) and non-response at the question level (partial non-response). It is calculated and presented for each geographic area. The GNR is the key criterion that determines whether or not the long-form census questionnaire results are released for a given geographic area – data are suppressed for geographic areas with a GNR equal to or greater than 50%. More information on the GNR is available in the [Guide to the Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm), Catalogue no. 98-304-X.

The standard error is a measure of the precision of an estimate with respect to sampling and total non-response variability. A small standard error corresponds to a precise estimate. Standard errors are made available to users for certain long-form census questionnaire estimates, except in cases where confidentiality would be compromised. The standard error can be used to derive other indicators of precision such as the coefficient of variation. It can also be used for most types of population parameters of interest (e.g. a count, a proportion or an average) and, using an adequate methodology, to derive margins of errors or confidence intervals for a given confidence level or to perform statistical inference (hypothesis testing). For more information on the long-form census questionnaire standard error and its interpretability and use, please refer to the [Guide to the Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm), Catalogue no. 98-304-X.

The imputation rate by question, excluding global non-response, is a measurement of quality specific to each question in the long-form census questionnaire. It measures the proportion of respondents ('respondents' being defined as those for whom a fully- or partially-completed questionnaire was returned) who did not answer the question, or whose response was invalid and for which a valid value was assigned. Imputation eliminates gaps in the data and, when done appropriately, reduces bias introduced by non-response. This is done by identifying persons or households that have characteristics similar to the incomplete record and by copying their values to fill in the missing or erroneous responses. The imputation rates for the ethnic origin question are presented in [Table 1](#).

Certification of final estimates

Once data processing, editing and imputation were completed, the data were weighted in order for estimates to represent the total Canadian population living in private dwellings. Certification of the final weighted estimates was the last step in the validation process leading to recommendation for release of the data for each geography and domain of interest. Based on the analysis of data quality indicators and the comparison of the long-form census questionnaire estimates with other data sources, the recommendation is for unconditional release, conditional release or non-release for quality reasons. In the case of conditional release or non-release, appropriate notes and warnings are included in this guide. Several data sources were used to evaluate the long-form census questionnaire estimates. However, since the risk of error often increases for lower levels of geography and for smaller populations, and the data sources used to evaluate these results are less reliable (or not available) at these lower levels, it can be difficult to certify the estimates at these levels.

Long-form census questionnaire estimates are also subject to confidentiality rules that ensure non-disclosure of individual respondent identity and characteristics. For more information on confidentiality rules, please refer to the

Guide to the Census of Population, 2016 (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm>), Catalogue no. 98-304-X.

For more information on data processing and the calculation of the estimates and their level of precision, please refer to the *Sampling and Weighting Technical Report, Census of Population, 2016* (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-306/index-eng.cfm>), Catalogue no. 98-306-X.

Comparability

Historical comparability

Over time, there have been differences in the question wording, format, examples and instructions for the ethnic origin question used in the census. The historical comparability of ethnic origin data has thus been affected by these factors, as well as by changes in data processing and the social environment at the time of the census.

The ethnic origin question in the 2016 Census, 2011 National Household Survey and 2006 Census was 'What were the ethnic or cultural origins of this person's ancestors?' In contrast, in the 2001, 1996 and 1991 censuses, the question was 'To which ethnic or cultural group(s) did this person's ancestors belong?' A definition of 'ancestor' has been included directly in the questionnaire since 2006 Census. Before 2006, the definition of ancestor was included only in the census guide. Finally, the list of examples of ethnic origins was updated in 2016 to reflect the frequency of single responses reported in the 2011 National Household Survey. For 2016, 'Iranian' and 'Mexican' were added to the list of examples, while 'Jewish' and 'Salvadorean' were removed. In addition, the order of the examples was modified based on single-response frequency in 2011. For example, 'Chinese' was moved from fourth place in 2011 to third in 2016, while 'French' was moved from third place to fourth.

Instructions provided alongside the census ethnic origin question have also been altered over time to suit changing collection needs. Since 1986, an instruction to specify as many ethnic groups as applicable has been included in the ethnic origin question. Prior to the 1981 Census, only the respondent's paternal ancestry was to be reported. If multiple ethnic origins were provided, only one origin was captured, resulting in one ethnic origin per respondent. In 1981, multiple origins were allowed, and a write-in space was added to the question, although respondents were not instructed to provide more than one origin. In 1986, respondents were permitted to write in up to three origins other than those shown in the mark-in circles. In 1991, they were permitted to write in up to two additional origins. Since 1996, four write-in boxes have been provided on the questionnaire, and up to six ethnic origins have been captured.

As a result of changing immigration patterns and increasing diversity in Canada, modifications are made to the specific ethnic groups and categories for which data are released each census. In general, the dissemination list of ethnic and cultural origins grows slightly each year. For 2016, data for 15 ethnic origins were released for the first time: Arawak, Bavarian, Bhutanese, Catalan, Corsican, Djiboutian, Edo, Ewe, Guadeloupean, Hazara, Karen, Kyrgyz, Malinké, Turkmen and Wolof. For the 2011 National Household Survey and previous censuses, these 15 ethnic origins had been included under other categories. For the detailed list of ethnic origins disseminated in the 2016 Census and their comparability with ethnic origins from the 2011 National Household Survey and 2006 Census, please refer to [Appendix 5.1](http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/app-ann/a5_1-eng.cfm) (http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/app-ann/a5_1-eng.cfm) of the *Dictionary, Census of Population, 2016* (<http://www12.statcan.gc.ca/census-recensement/2016/ref/dict/index-eng.cfm>), Catalogue no. 98-301-X.

Finally, the 2011 National Household Survey employed a different methodology than was used for the 2016 Census, the 2006 Census and prior censuses. These differences can affect comparability between 2016 Census and 2011 National Household Survey estimates for the ethnic origin variable. For more information on the comparability between the 2016 Census and the 2011 National Household Survey, please refer to the *Guide to the Census of Population, 2016* (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm>), Catalogue no. 98-304-X.

Comparability with other data sources

The 2016 Census is currently Statistics Canada's primary source of data on ethnic origin. In 2011, the National Household Survey collected information on ethnic origin, while, prior to 2011, previous censuses of population collected this information. Occasionally, other household surveys (e.g., the General Social Survey) also collect data on the ancestral origins of the population. In addition, a one-time postcensal survey, the [Ethnic Diversity Survey](http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4508&lang=en&db=imdb&adm=8&dis=2) (<http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4508&lang=en&db=imdb&adm=8&dis=2>), was conducted in 2002.

Many factors affect comparisons of ethnic origin data across these sources. Among other factors, comparability is affected by:

- differences in survey target populations, reference periods, sampling and collection methods
- question wording, questionnaire format, examples and instructions
- approaches to data processing
- the social and political climate at the time of data collection.

For additional information, please see the [Guide to the Census of Population, 2016](http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm) (<http://www12.statcan.gc.ca/census-recensement/2016/ref/98-304/index-eng.cfm>), Catalogue no. 98-304-X.

Data quality notes

The imputation rates for the 2016 Census ethnic origin question are shown in Table 1. The 2016 Census imputation rate for ethnic origin at the national level is 4.5%, compared with 5.8% in the 2011 National Household Survey and 5.9% in the 2006 Census. At the provincial level, the imputation rate for ethnic origin ranges from 3.7% in British Columbia and Nunavut to 10.7% in Newfoundland and Labrador.

Table 1
Imputation rates in the 2016 Census of Population for ethnic origin, Canada, provinces and territories

Province or territory	Imputation rate (%)
Canada	4.5
Newfoundland and Labrador	10.7
Prince Edward Island	8.2
Nova Scotia	7.0
New Brunswick	6.6
Quebec	5.2
Ontario	3.8
Manitoba	4.1
Saskatchewan	5.1
Alberta	4.6
British Columbia	3.7
Yukon	4.5
Northwest Territories	4.2
Nunavut	3.7

Source: Statistics Canada, Census of Population, 2016.