Census of Population, 2021



Release date: February 8, 2023



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 infostats@statcan.gc.ca

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

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

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.

Published by authority of the Minister responsible for Statistics Canada
© His Majesty the King in Right of Canada as represented by the Minister of Industry, 2023
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.

Release date: February 8, 2023 Catalogue number 98-500-X, issue 2021008

ISBN 978-0-660-42457-6

Table of contents

Definitions and concepts	4
Questions	4
Classifications	
Concepts over time	6
Collection and processing methods	7
Data quality	8
Variability due to sampling and total non-response Non-response bias	8
Data quality indicators	
Certification of final counts	
Data quality for ethnic or cultural origin	11
Comparability over time	11
Comparability with other data sources	11

Definitions and concepts

The 2021 Census of Population's question on <u>ethnic or cultural origins</u> collected information on the ancestral origins of the population, providing information about the composition of Canada's diverse population.

Ethnic or cultural origin refers to the ethnic or cultural origins of a person's ancestors. Ancestors may have Indigenous origins, origins that refer to different countries or other origins that may not refer to different countries. Often referred to as a person's ancestral "roots," ethnic or cultural origins should not be confused with citizenship, nationality, language or place of birth. For example, a person who has Canadian citizenship, speaks Hindi and was born in the United States may report having Guyanese ancestry.

Responses to the ethnic or cultural origins question on the census reflect respondents' perceptions of their background. As such, many factors can influence changes in responses over time, including the contemporary social environment, the respondents' knowledge of their family history, and their understanding of and views on the topic.

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

In the 2021 Census, the terms "origins" and "ancestry" are used interchangeably.

For additional information, please see the <u>Dictionary, Census of Population, 2021</u>, Statistics Canada, Catalogue no. 98-301-X.

Questions

The 2021 Census of Population data on ethnic or cultural origins were obtained from Question 23 on form <u>2A-L</u> and form <u>2A-R</u>. For the 2021 Census, the <u>2A</u> short-form questionnaire was used to enumerate all usual residents of 75% of private dwellings. The 2A-L long-form questionnaire, which also includes the questions from the 2A short-form questionnaire, was used to enumerate a 25% sample of private households in Canada. For private households in First Nations communities, Métis settlements, Inuit regions and other remote areas, the 2A-R questionnaire was used to enumerate 100% of the population.

On both versions of the questionnaire, the ethnic or cultural origins question asked, "What were the ethnic or cultural origins of this person's **ancestors**?" Below the question, a note indicated that "Ancestors may have Indigenous origins, or origins that refer to different countries, or other origins that may not refer to different countries."

To help respondents better understand the question, a link to a list of over 500 examples of ethnic and cultural origins was included in both the <u>2A-L</u> and <u>2A-R</u> electronic questionnaires. On the paper versions of the 2A-L and 2A-R questionnaires, a note indicated that respondents should visit <u>Examples of ethnic or cultural origins</u> to view these examples of ethnic or cultural origins.¹

The linked page with over 500 different examples of ethnic and cultural origins included origins that had a response frequency of 40 or more in 2016. Examples included in the list were also based on stakeholder and expert engagement. A note at the top of the page indicated that these are examples of different origins and that there may be other origins that are not on the list. The examples of different origins were arranged in three columns (Indigenous origins, origins referring to countries and other origins) and ordered alphabetically in each column.

Additional instructions on how to complete the 2021 ethnic or cultural origins question were provided to respondents via a help button accessed in the electronic questionnaire:

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

For all persons, report the specific ethnic or cultural origin or origins of their ancestors.

For examples, refer to the list of ethnic or cultural origins. If applicable, you may report an ethnic or cultural origin that is not on the list of examples.

For persons with **South Asian** origins, report a specific origin or origins. Do **not** report "Indian". Instead, report "Indian" or a specific South Asian origin, such as "Punjabi" or "Tamil".

For persons with North American **Indigenous** or **Aboriginal** ancestry, report a specific origin or origins. Do **not** report "Aboriginal", "Indigenous", "Native" or "Indian". Instead, report "First Nations", "North American Indian", "Métis", "Inuit," or a specific First Nations origin, such as "Cree" or "Mi'kmag".

For persons with **Indigenous** or **Aboriginal** ancestry from **outside** North America, report a specific origin or origins. Do **not** report "Aboriginal", "Indigenous", "Native" or "Indian". Instead, report "Central or South American Indigenous" or a specific Indigenous origin, such as "Arawak" or "Maya".

Information on the historical comparability of the 2021 Census ethnic or cultural origins question with questions asked in earlier censuses is provided in the sections of this document entitled Concepts over time and Comparability over time.

For more information on the reasons why the census questions are asked, please refer to the five <u>fact sheets</u> found on The road to the 2021 Census web page.

Classifications

Data from the ethnic or cultural origins 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 and cultural origins disseminated in the 2021 Census and their comparability with the origins from the 2016 Census and the 2011 National Household Survey are available in Appendix 2.5 of the *Dictionary, Census of Population, 2021*,

^{1.} As people completing the 2A-R paper questionnaire may face more challenges accessing the list of examples on the Statistics Canada website, examples were included directly on this version of the 2021 questionnaire to help them better understand the question. Specifically, the following examples of ethnic or cultural origins were included as part of the question: First Nations, Métis, Cree, Plains Cree, Mi'kmaq, Qalipu Mi'kmaq, Ojibway, Inuit, Inuvialuit, Mohawk, Iroquois [Haudenosaunee], Innu, Montagnais, Dene, Chipewyan [Denesuline], Blackfoot, Blood [Kainai], Salish, Secwepemc [Shuswap], Huron [Wendat], Cherokee, Scottish, Chinese, Jamaican, Moroccan, Jewish, Acadian and Punjabi. These examples include the most frequently reported Indigenous origins in 2016, and efforts were made to include Indigenous examples from different regions of Canada. Non-Indigenous examples in the list include the most frequently reported non-Indigenous origins in 2016 from different regions of the world, including examples of both country-based origins and other origins.

Statistics Canada, Catalogue no. 98-301-X. The 2021 Census includes data for almost 500 ethnic and cultural origins reported by people living in Canada. For each origin published, total single and multiple response counts are provided.

A single ethnic or cultural origin response occurs when a respondent reports having only one origin. For example, in the 2021 Census, about 559,575 people stated that their only origin was Scottish.

A multiple response occurs when a respondent reports having two or more origins. For example, in the 2021 Census, about 3,832,625 people gave a response that included Scottish and one or more other origins.

Total response counts (also called "Total – Single and multiple ethnic or cultural origin responses" in some data tables) indicate the number of people who reported a specified origin, either as their only origin or in addition to one or more other origins. Total responses are the sum of single and multiple responses for each ethnic or cultural origin. For example, in 2021, a total of about 4,392,200 people reported having Scottish ancestry (the sum of the 559,575 people who reported Scottish as their only origin and the 3,832,625 people who reported Scottish in combination with other origins). Because multiple origins can be reported, the sum of all ethnic or cultural origin responses is typically greater than the total population of a geographic area.

Ethnic or cultural origin is a difficult concept to measure, and there is no internationally recognized classification for this concept. In general, 2021 Census of Population data for an ethnic or cultural group are published by Statistics Canada if the estimate is approximately 500 or higher.

Concepts over time

Over time, there have been differences in the wording, format, examples and instructions of the ethnic or cultural origins question used in the census. There have also been changes in what is considered to be a valid ethnic or cultural origin response. The historical comparability of data on ethnic or cultural origins has been affected by these factors.

Changes to the ethnic or cultural origins question

The ethnic or cultural origins question asked in the 2021 Census ("What were the ethnic or cultural origins of this person's **ancestors**?") was the same question that was asked in 2016, 2011 and 2006. 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?"

However, despite the wording of the question itself not changing for 2021, other elements of the question were revised since the 2016 Census, including the following:

- In 2021, no examples of ethnic or cultural origins were listed in either the 2A-L or the 2A-R electronic questionnaire, as well as the 2A-L paper questionnaire. In 2016, a list of 28 examples of different origins was included in the 2A-L question, and 12 examples were included in the 2A-R question.
- Instead, in electronic versions of the 2021 question, there was a link to a page with over 500 different
 examples of ethnic and cultural origins. In paper versions of the 2021 question, respondents were
 instructed to go to a website Examples of ethnic or cultural origins.
- In 2021, there was a note describing different types of origins, specifically "Ancestors may have Indigenous origins, or origins that refer to different countries, or other origins that may not refer to different countries." The 2016 Census question included two notes that were removed in 2021, specifically "This question collects information on the ancestral origins of the population and provides information about the composition of Canada's diverse population," and "An ancestor is usually more distant than a grandparent."

The revised question addressed concerns regarding the impact on response patterns of the list of examples on the questionnaire. The 2021 question allowed respondents to report their origins without the influence of examples listed on the questionnaire, resulting in better-quality data that are more representative of the population.

However, the new version of the question for the 2021 Census produces results that are not comparable to the 2016 Census results for many ethnic and cultural origins. The origins that are particularly affected are those that were among the 28 examples listed directly on the 2016 Census questionnaire (e.g., Canadian).² For the 2021 Census, these origins were included as part of the much more extensive list of examples of ethnic and cultural origins, mitigating the prompting effect they had in the past.

It should be noted that the results on ethnic or cultural origins were not fully comparable between censuses in the past, in part because of changes to the list of examples over time.³ Now that examples have been removed from the questionnaire, results will no longer be affected by this factor, and this will help improve comparability between censuses moving forward.

Changes to the ethnic or cultural origins variable

The revised approach for the 2021 Census, along with changing immigration patterns and increasing diversity in Canada, yields more varied and diverse responses than in past censuses. To better reflect the range of responses received, a greater number of origins have been disseminated for the 2021 Census.

For 2021, data for over 150 additional origins have been included in the ethnic or cultural origins variable for the first time. As a result, the variable now includes almost 500 detailed origins. Many of these newly included origins are Indigenous origins that were included in a separate variable in 2016, Aboriginal ancestry responses. However, the 2021 ethnic or cultural origins variable also includes many additional non-Indigenous origins, based on the wider range of responses received.

Of particular note is the inclusion of many origins that are more cultural, ethno-religious, ethno-racial or ethno-linguistic in nature than strictly ethnic. In 2016 and previous censuses, many of these origins were included under different categories (e.g., the 2021 origin "Sikh" was included under "Punjabi" in 2016). However, the census question clearly asks respondents to report ethnic **or** cultural origins, so these types of origins were included in the 2021 variable if they had sufficient response estimates.

For the detailed list of ethnic or cultural origins disseminated in the 2021 Census and their comparability with ethnic origins from the 2016 Census and 2011 National Household Survey, please refer to <u>Appendix 2.5</u> of the *Dictionary, Census of Population, 2021*, Statistics Canada, Catalogue no. 98-301-X.

Information on the historical comparability of the 2021 Census data on ethnic or cultural origins is provided in the Comparability over time section.

Collection and processing methods

The COVID-19 pandemic emerged in Canada in early 2020 and affected all steps of the 2021 Census process, from data collection to dissemination. Please refer to the <u>Guide to the Census of Population</u>, 2021, Statistics Canada Catalogue no. 98-304-X, for more detailed information on this topic.

^{2.} Please refer to Ethnic or cultural origins: Technical report on changes for the 2021 Census.

^{3.} Please refer to the Technical report on changes in response related to the census ethnic origin question: Focus on Jewish origins, 2016 Census integrated with 2011 National Household Survey.

Data quality

The 2021 Census of Population underwent a thorough data quality assessment. The different certification activities conducted to evaluate the quality of the 2021 Census data are described in Chapter 9 of the Guide to the Census of Population, 2021, Statistics Canada Catalogue no. 98-304-X.

The data quality assessment was conducted in addition to the regular verifications and quality checks completed at key stages of the census. For example, throughout data collection and processing, the accuracy of specific steps such as data capture and coding was measured, the consistency of the responses provided was checked, and the non-response rates for each question were analyzed. As well, the quality of imputed responses was assessed during data editing and imputation.

During the data quality assessment, a number of data quality indicators were produced and used to evaluate the quality of the data. These indicators are briefly described below. Finally, resulting census counts were compared with other data sources and certified for final release.

The main highlights of this assessment of the data pertaining to Ethnic or Cultural Origin 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 populations (such as Indigenous peoples and immigrants) that are generally referred to in this document as "populations 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 total non-response introduce variability into 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 considerably depending on the domain or geography of interest, in particular because of the variation in response rates. For more information on variability due to sampling and total non-response in long-form census questionnaire estimates, please refer to the <u>Guide to the Census of Population</u>, 2021, Statistics Canada 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.

In general, the risk of non-response bias increases as the response rate declines. For the 2021 long-form census questionnaire, Statistics Canada adapted its collection and estimation procedures to mitigate the effect of non-response bias to the extent possible. For more information on these mitigation strategies, please refer to the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Data quality indicators

A number of quality indicators were produced and analyzed during the 2021 Census of Population data quality assessment. Four indicators are available to data users for long-form content: the total non-response (TNR) rate; the confidence interval; as well as the non-response rate and the imputation rate per question.

The **total non-response (TNR) rate** is the primary quality indicator that accompanies each disseminated 2021 Census of Population product, and is calculated for each geographic area. It measures total non-response at the dwelling level. Non-response is said to be total when no questionnaire is returned from a dwelling or when

a returned questionnaire does not meet the minimum content. More information on the TNR rate is available in Chapter 9 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

The confidence interval was selected as a variance-based quality indicator to accompany the 2021 Census of Population long-form estimates because it helps users easily make a statistical inference. This indicator provides a measure of the accuracy of the long-form estimates. Using a science-based approach, research and simulations were done to ensure that confidence intervals are constructed using adequate statistical methods for the Census of Population data and areas of interest.

A confidence interval is associated with a confidence level, generally set at 95%. A 95% confidence interval is an interval constructed around the estimate so that, if the process that generated the sample were repeated many times, the value of the interest parameter in the population would be contained in 95% of these intervals. The confidence interval consists of a lower bound and an upper bound. These two bounds accompany the long-form estimates in most data tables.

Further details on the different methods used to construct confidence intervals and their assumptions are provided in the <u>Sampling and Weighting Technical Report, Census of Population, 2021</u>, Statistics Canada Catalogue no. 98-306-X.

The **non-response rate per question** is a measure of missing information due to non-response to a question. It measures only the non-response that is resolved through imputation during data processing (as opposed to weighting when a sample is used). For the long-form questionnaire, the non-response rate per question includes only partial non-response to the question, except for First Nations communities, Métis settlements, Inuit regions and other remote areas where both partial and total non-response are taken into account. Partial non-response is when answers to certain questions are not provided for a respondent household.

The non-response rate per question for a question on the long-form questionnaire is defined as the sum of the weights of in-scope units in the population of interest who did not respond to the question divided by the sum of the weights of in-scope units in the population of interest. Here "units" refers to the statistical units for which data are collected or derived (e.g., persons or households, depending on whether the question is about a person-level characteristic or a household-level characteristic). A unit is considered to be in scope for a given question if the question is applicable to that unit and the unit belongs to the population of interest related to the question.

The **imputation rate per question** measures the extent to which responses to a given question were imputed. Imputation is used to replace missing data in the event of non-response or when a response is found to be invalid (e.g., multiple answers are provided when a single answer is expected). Imputation is conducted to eliminate data gaps and to reduce bias introduced by non-response. Imputation is generally done by identifying persons or households in the same geographical area with similar characteristics to the incomplete record and copying their values to fill in the missing or invalid responses.

The imputation rate for a question on the long-form questionnaire is defined as the sum of the weights of in-scope units in the population of interest for which the response to the question was imputed divided by the sum of the weights of in-scope units in the population of interest (see the definition of "units" provided in the above section on the non-response rate per question).

For long-form content, imputation for most areas is done to resolve partial non-response—not total non-response, which instead is treated by weighting. However, in First Nations communities, Métis settlements, Inuit regions and other remote areas, whole household imputation (WHI) is used to resolve total non-response. It first imputes the occupancy status of non-respondent dwellings and further imputes all the data for those dwellings resolved as occupied in the first step. WHI is included in the imputation rate per question, including the use of administrative data to impute non-responding households in areas with low response rates; see Appendix 1.7 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X. As with the non-response rate, a unit is considered to be in scope if the question is applicable to that unit and the unit belongs to the population of interest related to the question.

The non-response and imputation rates per question can be interpreted as the proportion of in-scope units in the population of interest for which information was not reported or was imputed, respectively. The long-form rates are weighted to reflect the fact that the long-form questionnaire is only distributed to a sample of the population, so in this case, the proportion is estimated.

The non-response and imputation rates for a question are often similar, but some differences can be observed for a given question because of additional data processing steps that may have been required. These rates were regularly checked during data assessment, and a detailed analysis was done if there was a difference between the two rates for a question, to ensure the appropriateness of the processing steps taken and the quality of the data. A difference between the non-response rate and the imputation rate for a question can generally be explained by one of the following two factors:

- Some responses were considered invalid or inconsistent during the edit stage and imputation was needed, which is why the imputation rate is higher than the non-response rate for a question.
- Some non-responses were resolved in a straightforward manner early during data processing because
 a single resolution was possible based on the answers provided to other questions, making imputation
 unnecessary. This may explain why the non-response rate is higher than the imputation rate for
 a question.

Table 1 below presents the non-response and imputation rates for Canada and for each province and territory.

The non-response and imputation rates per question at lower levels of geography are also available in 2021 Census data tables presenting data quality indicators. This information is scheduled for release on August 17, 2022, for short-form questions and on November 30, 2022, for long-form questions.

The <u>2021 Census Data Quality Guidelines</u>, Statistics Canada Catalogue no. 98-26-0006, provides all the information required to understand and interpret the data quality indicators for the 2021 Census, along with guidelines to enable their proper usage. Data quality indicators are provided so that users are informed about the quality of the statistical information and can determine the relevance and the limitations of the data relative to their needs. In general, the quality of the 2021 Census of Population data is very good, but in some cases, data have to be used with caution. It is strongly recommended that users consult all available data quality indicators to get a better sense of the quality of the data products they are interested in.

Certification of final counts

Once data editing and imputation were completed, the data were weighted to ensure that estimates represent the total Canadian population living in private dwellings. Certification of the final weighted estimates was the last step in the validation process, which led to the recommendation to release the data for each level of geography and domain of interest. Based on the analysis of the data quality indicators and the comparison of long-form census questionnaire estimates with other data sources, the recommendation is for unconditional release, conditional release, or non-release (for quality reasons on rare occasions). For conditional release or non-release, appropriate notes and warnings are included in the products and provided to users. Moreover, other 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 counts are less reliable or not available at these lower levels, it can be difficult to certify the counts at these levels.

Long-form census questionnaire estimates are also subject to confidentiality rules that ensure non-disclosure of respondent identity and characteristics. For more information on privacy and confidentiality, please refer to Chapter 1 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X. For information on how Statistics Canada balances the protection of confidentiality and the need for disaggregated census data, with specific attention to new 2021 Census content, please refer to Census of Population, 2021, Statistics Canada Catalogue no. 98-26-0005.

For more information on data processing and the calculation of estimates and their level of precision, please refer to the <u>Sampling and Weighting Technical Report, Census of Population, 2021</u>, Statistics Canada Catalogue no. 98-306-X.

Data quality for ethnic or cultural origin

The non-response and imputation rates for the ethnic or cultural origins question in the 2021 Census are shown in Table 1. At the national level, the non-response rate for the ethnic or cultural origins question was 3.5% and the imputation rate was 8.0%. At the provincial and territorial level, the imputation rate ranged from 6.2% in British Columbia to 26.1% in Nunavut. In the territories, as well as in First Nations communities, Métis Settlements, Inuit regions and other remote areas in the provinces, COVID-19 presented challenges for conducting the Census of Population, including some that affected in-person enumeration, such as travel restrictions and unavailability of local staff.

Table 1 Non-response rate and imputation rate for the ethnic or cultural origins question, Canada, provinces and territories, Census of Population, 2021

	Non-response rate	Imputation rate
Geography	percent	
Canada	3.5	8.0
Newfoundland and Labrador	6.2	17.0
Prince Edward Island	4.0	10.0
Nova Scotia	4.1	10.3
New Brunswick	5.9	13.6
Quebec	4.3	10.1
Ontario	2.8	6.6
Manitoba	4.0	7.0
Saskatchewan	4.2	7.9
Alberta	3.5	8.3
British Columbia	2.7	6.2
Yukon	6.8	10.4
Northwest Territories	10.1	12.1
Nunavut	25.2	26.1

Source: Statistics Canada, Census of Population, 2021.

Comparability over time

The question on ethnic or cultural origins has been revised for the 2021 Census. A description of the changes and their impact on historical comparability is included in the section of this document entitled Concepts over time.

Comparability with other data sources

The 2021 Census is currently Statistics Canada's primary source of data on ethnic or cultural origin. Prior to 2021, the census collected this information in 2016, in 2006 and earlier. In 2011, the National Household Survey 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, was conducted in 2002.

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

- differences in survey target populations, reference periods, sampling and collection methods
- differences in the list of incompletely enumerated reserves and settlements
- 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, 2021*, Statistics Canada Catalogue no. 98-304-X.