Census of Agriculture: Data quality report, 2021

Release date: February 6, 2024

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   1-800-263-1136
- National telecommunications device for the hearing impaired   1-800-363-7629
- Fax line   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.

Table of Contents

Census of Agriculture: Data quality report, 2021

Introduction

Every five years, the Census of Agriculture collects and disseminates statistical information on a wide and detailed range of farm commodities, practices and characteristics as well as on the operators of these farms. Numerous quality assurance steps are implemented throughout the planning, collection, processing and dissemination processes to ensure that census data are reliable and that they meet user needs.

The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. An integral part of each Census of Agriculture is the implementation of new or enhanced methods, procedures and technologies that ultimately contribute to ensuring that Census of Agriculture data are as accurate as they can be, by improving all areas of the process including collection, processing, validation and dissemination of the data.

The objective of this report is to provide data quality information pertaining to the 2021 Census of Agriculture, such as sources of error, error detection, disclosure control methods, data quality indicators, response rates and collection rates.

For a more detailed description of the Census of Agriculture concepts and methods, please refer to the [Guide to the Census of Agriculture, 2021](#).

Reference date

The 2021 Census of Agriculture reference date is May 11, 2021.

Target population

The target population for the Census of Agriculture is all "[census farms](#)" in Canada. In 2021, a "census farm" is defined a unit that produces agricultural products and reports revenues or expenses for tax purposes to the Canada Revenue Agency. Agricultural products include:

- Crops: grains, oilseeds, leguminous crops, potatoes, vegetables, fruits, berries, greenhouse products, mushrooms, sod, nursery, Christmas trees, maple tree taps, hay and fodder crops, hemp and other crops.
- Livestock: dairy and beef cattle (including feedlots), pigs, poultry and eggs (including hatcheries), turkeys, ducks, geese, sheep, goats, horses and other equines, bison (buffalo), elk (wapiti), deer, llamas and alpacas, rabbits, mink, bees and other animals.

Not included are: forestry and logging, hunting and trapping, fishing and aquaculture, support activities for agriculture and post-harvest activities, horse boarding and riding lessons, operations producing products that are not for human consumption (e.g., genetic operations, insect farms for pet food).

The observed population is selected from Statistics Canada's Business Register in conjunction with information from the latest set of tax remittances. The selection process uses the detailed tax information of the establishments on the Business Register to select those which have reported agricultural commodity revenues and/or expenses, signaling that they are involved in agriculture. To ensure more complete coverage, additional data and methods are used to include establishments which report their fiscal data differently. Because the latest available tax data are from 2019, an additional set of modelled records was added to this population to represent newer farms and reduce undercoverage.

Collection

Census of Agriculture data are collected directly from survey respondents, with the exception of those modelled units for which the data are imputed rather than collected.

In 2021, the Census of Agriculture focused on electronic questionnaire collection. Invitation letters were delivered to farm operations by Canada Post. Farm operators were asked to complete the 2021 Census of Agriculture online by using the secure access code provided in the invitation letter. If it was determined that a questionnaire was not received, follow-up was conducted by telephone.

The Census of Agriculture reduces response burden by replacing questionnaire data with administrative data where possible. The utilization of high-quality data sources —such as tax data from the Canada Revenue Agency — eliminates the need to ask respondents questions about the operating arrangement, and revenues and expenses because this information can be obtained from their tax forms. Alternative data sources can also be used to populate questionnaire data in some cases.

The age and sex of the farm operators come from what is reported in the Census of Population for each farm operator. The farm operators from the Census of Agriculture are linked to the Census of Population database using a probabilistic linkage method which matches personal and household information provided on both questionnaires (such as name, birthdate, telephone number, etc.). Operators on the Census of Agriculture for which no link is found will have their information imputed with that of another Census of Population person having similar characteristics.

Error detection

Error detection is an integral part of both collection and data processing activities. Edits were applied to microdata records during collection to identify reporting and capture errors, as well as data inconsistencies. Totals in key variables that do not equal the sum of their parts and ratios that exceed tolerance thresholds were flagged for respondents to review.

Data from paper questionnaires were captured through the electronic questionnaire application and were subjected to the same rigorous quality control and processing edits as the electronic responses, to identify and resolve problems related to inaccurate, missing or inconsistent data.

During data processing, additional edits were used to automatically detect errors or inconsistencies that remain in the microdata following collection. These edits include value edits (e.g., values which fall outside of expected ranges), linear equality edits (e.g., the sum of parts is equal to the total), linear inequality edits (e.g., a value for one question is always expected to be larger than the value of another), and consistency edits (e.g., an amount is reported for the value of trucks, but no trucks are reported, or the vegetables screening question is flagged as 'yes' but no area is reported for any vegetables). When errors were found, they were corrected using the data editing and imputation processes, or during the data validation process.

Extreme values were also identified using automated methods based on the distribution of the collected information. Following their detection, these values were reviewed by subject-matter analysts in order to assess their validity. Macro-level totals were also reviewed to make sure they line up with expectations and economic market trends. During this process, provincial or agricultural experts are consulted. In general, every effort was made to minimize the non-sampling errors of omission, duplication, misclassification, reporting and processing.

Imputation

Non-response occurs when respondents do not answer a portion of the questionnaire or the questionnaire as a whole, or when reported data are considered erroneous during the error detection steps. In those situations, imputation is used to fill in the missing information and modify the erroneous information. Many methods of imputation may be used to complete a questionnaire, including manual changes made by an analyst. The automated, statistical techniques used to impute the missing data mainly include: deterministic imputation and replacement using data from a similar unit in the sample (known as donor imputation). Usually, important variables are imputed first and are used as anchors in subsequent steps to impute other related variables. In some cases, ratio imputation and historical imputation are also used to complete the data for some specific types of units.

Manual imputation of missing data is done only for some cases when the collected data does not align with historical data or with a known data relationship. These are generally done on rare occasions during the data validation process, after thorough investigation.

Estimation

The Census of Agriculture collects or imputes a set of values for each census farm in the population. For this reason, no sampling weights are required for tabulations. Totals and averages can be calculated by simply summing or taking the average value of the variables in question for the records in the database within a desired domain.

The quality of the resulting estimates is represented through the use of a variance measure. In the case of the Census of Agriculture, this variance estimate represents the amount of uncertainty in the point estimate due to the imputation that took place during the data processing step and any additional variance required to maintain the confidentiality of a respondent's data.

Quality evaluation: Validation and certification

The initial data validation stage is undertaken by a team of subject-matter analysts. They review the estimates by comparing them to the results of previous censuses or estimates from other data sources. Although the validation of all individual records is not feasible, the analysts review the most important contributors individually, especially when the estimates vary significantly from those of other data sources.

For the Census of Agriculture, the final data validation process is the certification of the data. At this stage, a wider range of analysts and experts review and compare the results to estimates from previous censuses or estimates from other data sources. During data certification, response rates, edit failure rates, coverage rates and a comparison of the data before and after imputation are among the measures used to evaluate the accuracy and coherence of the data and possibly explain differences with other sources. Detailed cross-tabulations are also checked for consistency and accuracy.

Some estimates are not comparable with those of previous censuses. This may be due to wording or conceptual changes in the questions in 2021, or the addition or removal of questions between 2016 and 2021. After thoroughly investigating each case, notes were developed to identify the affected questions and explain the reasons that users should use caution when comparing the results.

Data for cannabis operations were collected for the first time during the 2021 Census of Agriculture. Due to the complexity of these operations' activities and organizational structure, these respondents were not able to provide responses that precisely captured the agricultural activity of cannabis cultivation in its entirety and/or disassociated from non-agricultural activities. Furthermore, the response rate for cannabis operations was low, and this consequently resulted in a high imputation rate. As a result, cannabis operations were excluded from the Census of Agriculture databases and its data releases. As an alternative, the Census of Agriculture published cannabis

data extracted from administrative files received from Health Canada. These data provide the number of licensed cannabis cultivators and their production areas at the national and provincial levels of geography. Separate information is available for operations growing cannabis under cover and in open fields. Cannabis operations included in this release are not included in all other Census of Agriculture releases.

Disclosure control

Statistics Canada is prohibited by law from releasing any information it collects that could identify any person, business, or organization, unless consent has been given by the respondent or as permitted by the *Statistics Act*. Various confidentiality rules are applied to all data that are released or published to prevent the publication or disclosure of any information deemed confidential.

Data disclosure occurs when the value in a tabulation cell is composed of or dominated by a few census farms. In order to prevent any data disclosure, all published tables are analysed using a method known as Random Tabular Adjustment. This technique aims to increase the amount of data made available to users compared to traditional data suppression approaches, while protecting the confidentiality of respondents. In cases where the direct publication of an estimate would potentially lead to the disclosure of an individual's data, the estimate is adjusted by a random amount in order to provide additional uncertainty to the estimate of the individual's responses. For a more detailed description of the Random Tabular Adjustment process, please refer to the [Frequently asked questions on random tabular adjustment (RTA)](#).

Revisions and seasonal adjustment

Data from the Census of Agriculture are not subject to revisions or seasonal adjustments.

Data accuracy

The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. Numerous traditional and enhanced quality assurance steps are put into place to ensure that Census of Agriculture data are as accurate as they can be.

With projects as large and complex as the Census of Agriculture, the estimates produced are inevitably subject to a certain degree of error. Knowing the types of errors that can occur and how they affect specific variables can help users assess the usefulness of the data for their particular applications, as well as assess the risks involved in making conclusions or decisions based on these results.

The quality assurance steps and details about the types of error that can affect the quality of the Census of Agriculture estimates are described in greater detail in the [Guide to the Census of Agriculture, 2021](#).

In addition to these quality assurance steps, in 2021 for the first time, the Census of Agriculture is providing a quality indicator for most published value estimates. These quality indicators take into account the variance in the estimate resulting from the imputation step during data processing and any extra adjustment required by the tabular disclosure avoidance method to protect the confidentiality of census respondents. Quality indicators are represented by letters—ranging from A through F—with each letter defined by a specific coefficient of variation (CV) range (see Table 1).

Table 1
Data quality indicators, 2021 Census of Agriculture

Quality indicator	Coefficient of variation value	Description
A	< 5.00%	Excellent
B	5.00% to 9.99%	Very good
C	10.00% to 14.99%	Good
D	15.00% to 24.99%	Acceptable
E	25.00% to 49.99%	Use with caution
F	≥ 50.00%	Too unreliable to be published

Note: Only one of the letter quality indicators above is published for most estimates included in the Census of Agriculture tabulations; the exact coefficients of variation are not disclosed.

The quality indicators for farm count estimates are calculated using a different method than coefficients of variation, but use a similar A to F scale to represent the quality of the estimate.

Response rates

Response rates are one of the key data quality measures for the Census of Agriculture. The response rates were calculated at the national level and for each province after the data processing and certification steps (see Table 2).

Table 2
Response rate, 2021 Census of Agriculture, Canada, provinces and territories

Geography	Overall response[1] percentage
Newfoundland and Labrador	77.6
Prince Edward Island	78.9
Nova Scotia	82.5
New Brunswick	83.3
Quebec	82.1
Ontario	82.3
Manitoba	77.4
Saskatchewan	74.6
Alberta	75.4
British Columbia	77.6
Yukon	83.0
Northwest Territories	75.0
Canada[2]	**78.6**

1. The term "overall response" refers to the completion of a census questionnaire with or without assistance from a Statistics Canada representative.
2. No census farms were enumerated in Nunavut.

Collection rates

The response rate between the 2021 census and previous censuses are not directly comparable, due to the inclusion of the modelled records in 2021 that were added to the initial population to represent new census farms since 2019 and reduce undercoverage. Since these census farms were modelled and were not sent a direct questionnaire, they are considered to effectively be non-respondents in these calculations. If excluded from the calculations to provide a more equivalent comparison of response rates with previous censuses, the rates look more similar. These collection rates represent the response rate among those census farms for which direct collection was attempted (see Table 3).

Table 3
Collection rate, 2021 Census of Agriculture, Canada, provinces and territories

Geography	Collection rate[1] percentage
Newfoundland and Labrador	85.0
Prince Edward Island	82.8
Nova Scotia	91.4
New Brunswick	91.2
Quebec	90.1
Ontario	88.6
Manitoba	83.2
Saskatchewan	79.7
Alberta	82.6
British Columbia	87.3
Yukon	83.0
Northwest Territories	75.0
Canada[2]	**85.4**

1. The term "collection" refers to the completion of a collected census questionnaire with or without assistance from a Statistics Canada representative.
2. No census farms were enumerated in Nunavut.

Coverage evaluation

Coverage errors occur when there is difference between the target population and the survey population and they may affect the quality of all estimates. For the Census of Agriculture, coverage errors occur when census farms are missed, incorrectly included or double counted. Estimating these errors is one way to assess the quality of the Census of Agriculture estimates.

The Census of Agriculture processes involved in the creation of the frame, data collection, and data processing are not perfect and can contribute to these coverage errors. For example, when creating the frame, real census farms might be missed because they were simply not part of one of the sources used to create the Census of Agriculture frame. Also, at the end of the collection period, non-responding units might be erroneously classified as census farms or as non-census farms during Census of Agriculture data processing.

The overall coverage of the Census of Agriculture was measured using two components. The first component measured the misclassification of non-respondents; that is, the accuracy of the processing step in which it was decided whether a non-responding unit was a census farm or not. Estimates of both undercoverage (non-enumerated agricultural operations) and overcoverage (units incorrectly enumerated as agricultural operations) were calculated.

The second component measured additional undercoverage (non-enumerated agricultural operations) errors arising from missing census farms on the frame. These were estimated using a post-censal survey called the Agriculture Frame Update Survey. This survey targeted establishments from the Business Register that had some indication of being a census farm, but were not included in the Census of Agriculture frame. Through this survey, additional census farms that had been missed by the Census of Agriculture were identified and used to estimate the total number of such missed farms.

The final net undercoverage estimates combine the estimates of these two components and is calculated using the following formula:

$$\text{Net undercoverage rate (\%)} = 100 * \frac{\text{non-enumerated farms} - \text{incorrectly enumerated farms}}{\text{enumerated farms} + \text{non-enumerated farms} - \text{incorrectly enumerated farms}}$$

Similarly, the net undercoverage estimate can be weighted by census variables to estimate the rate of undercoverage of that characteristic. Undercoverage rates were calculated for three principal measurements in the Census of Agriculture–the number of census farms (see Table 4), the total farm area (see Table 5) and the total operating revenues (see Table 6). Please note that there are no estimates of undercoverage for Yukon, the Northwest Territories and Nunavut as the Agriculture Frame Update Survey was not carried out in the territories.

To improve the accuracy of the coverage estimates in 2021, a more complex and complete estimation approach was used for the components of undercoverage coming from the Agriculture Frame Update Survey, compared to the one used in 2016. This resulted in additional undercoverage which would not have been captured in the 2016 results. Thus, the undercoverage estimates are not directly comparable to those from the 2016 Census of Agriculture.

Table 4
Census farm net undercoverage, 2021 Census of Agriculture, Canada and provinces

Geography	Enumerated farms	Estimated non-enumerated farms	Estimated incorrectly enumerated farms	Estimated net undercoverage
		number of farms		percentage
Newfoundland and Labrador	344	41	8	8.8
Prince Edward Island	1,195	148	30	9.0
Nova Scotia	2,741	347	37	10.2
New Brunswick	1,851	188	24	8.1
Quebec	29,380	2,494	483	6.4
Ontario	48,346	6,092	700	10.0
Manitoba	14,543	1,624	339	8.1
Saskatchewan	34,128	3,205	1,001	6.1
Alberta	41,505	5,639	980	10.1
British Columbia	15,841	2,279	276	11.2
Canada	**189,874**	**22,015**	**3,829**	**8.7**

Table 5
Total farm area net undercoverage, 2021 Census of Agriculture, Canada and provinces

Geography	Enumerated farms	Estimated non-enumerated farms	Estimated incorrectly enumerated farms	Estimated net undercoverage
		acres		percentage
Newfoundland and Labrador	49,425	2,609	2,176	0.9
Prince Edward Island	504,674	18,542	9,761	1.7
Nova Scotia	720,046	41,604	11,217	4.0
New Brunswick	685,377	26,208	7,476	2.7
Quebec	7,770,429	409,487	124,256	3.5
Ontario	11,766,071	980,762	172,248	6.4
Manitoba	17,121,019	1,134,659	451,149	3.8
Saskatchewan	60,265,339	3,540,832	1,997,930	2.5
Alberta	49,157,232	3,986,173	1,368,179	5.1
British Columbia	5,648,161	245,596	79,867	2.9
Canada	**153,687,771**	**10,136,315**	**3,254,968**	**4.3**

Table 6
Total operating revenues net undercoverage, 2021 Census of Agriculture, Canada and provinces

Geography	Enumerated farms	Estimated non-enumerated farms	Estimated incorrectly enumerated farms	Estimated net undercoverage
		dollars		percentage
Newfoundland and Labrador	154,592,361	1,519,570	2,640,653	-0.7
Prince Edward Island	682,912,760	17,436,280	11,118,683	0.9
Nova Scotia	727,873,979	22,072,099	13,212,359	1.2
New Brunswick	739,913,440	12,431,267	8,605,204	0.5
Quebec	13,098,971,426	287,848,433	224,412,358	0.5
Ontario	19,741,314,319	1,257,691,862	370,405,321	4.3
Manitoba	8,188,252,189	347,430,408	229,065,530	1.4
Saskatchewan	16,777,324,532	847,593,121	574,828,454	1.6
Alberta	22,220,826,389	1,059,680,003	688,565,670	1.6
British Columbia	4,804,135,169	181,311,300	107,229,900	1.5
Canada	**87,136,116,565**	**4,030,121,424**	**2,350,355,352**	**1.9**