Data sources and methodology

Target population

The target population for the Census of Agriculture is all 'census' farms in Canada. In 2016, a census farm was defined as an agricultural operation that produces at least one of the following products intended for sale: crops (hay, field crops, tree fruits or nuts, berries or grapes, vegetables, seed); livestock (cattle, pigs, sheep, horses, game animals, other livestock); poultry (hens, chickens, turkeys, chicks, game birds, other poultry); animal products (milk or cream, eggs, wool, furs, meat); or other agricultural products (Christmas trees, greenhouse or nursery products, mushrooms, sod, honey, maple syrup products).

The observed population starts with establishments on Statistics Canada's Business Register that responded to the 2016 Census of Agriculture or had reported to a recent Statistics Canada agriculture survey and confirmed agricultural activity. In addition, establishments that have indicators of agricultural activity in their tax remittances or have reported a main business activity of agriculture to the Canada Revenue Agency are also eligible for the Census. However due to operational and budgetary constraints, only those establishments with strong indicators were included. During collection, the new potential farms identified through the Census of Population questionnaire were also added to the observed population.

Instrument design

User consultations

A series of workshops were held across Canada in 2012 with users such as federal departments and provincial ministries, agricultural associations, academics and agriculture service providers. Users submitted recommendations for the types of questions they would like to see on the 2016 Census questionnaire which were used to develop the content and design of the census questionnaire.

Evaluating the suggestions

Before going any further, the submitted recommendations had to meet certain criteria before being judged suitable for inclusion in the Census of Agriculture:

- Is this topic of national interest?
- Are data worthwhile at more detailed geographic levels than provincial or national?
- Will farmers easily understand the question?
- Can the question be answered-that is, do the farm operators have the information to answer the question?
- Will farm operators be willing to answer it?
- Will there be a broad demand for the data generated by the question?
- Can the question be answered by either "Yes," "No," or a quantitative response?

Questionnaire content and development

Although the questionnaire is updated every census to reflect users' changing requirements as identified

through the submission process, certain questions appear on every census. These questions-such as those on farm operators, land area, livestock numbers and crop areas-are considered essential by Statistics Canada and other major users of Census of Agriculture data. Repeating basic questions allows the census to measure change over time, while adding new questions and dropping others allows data to be collected that reflect new technologies and structural changes in the agriculture industry. Four new topics, notably the adoption of technologies, direct marketing, succession planning and renewable energy systems were added to the 2016 questionnaire. These topics reflect changes in the industry and strong user demand for this new information. Also, to reduce respondent burden the detailed questions related to farm operating expenses have been replaced with one question asking for total farm operating expenses.

New or changed questions were developed in Head Office in consultation with industry experts and tested a number of times with farm operators across Canada through one-on-one interviews on their farms and in focus groups. Farm operators selected for testing reflected regional diversity-in types of agriculture, production techniques, farm size, language and age. This testing proved that some questions would not perform well on the census, and that the wording of other questions would require fine-tuning. Respondent burden, content-testing results, user priorities and budgets were all taken into consideration in determining the final content of the 2016 Census of Agriculture questionnaire. It was approved by Cabinet in the spring of 2015.

For more information about the 2016 Census of Agriculture consultation process and results, the "Census of Agriculture: Content Consultations" (http://www.statcan.gc.ca/pub/95-635-x/95-635-x2014001-eng.htm) publication is available from the "Browse by key resources" module of our website under "Publications".

Sampling

This survey is a census with a cross-sectional design.

The Census of Agriculture is designed to obtain complete and accurate data from all farms in Canada. Data are collected for all units of the target population, therefore no sampling is done.

Data sources

Data collection for this reference period: 2016-04 to 2016-07 (census day: May 10, 2016)

Responding to this survey is mandatory.

Data are collected directly from survey respondents.

Collection

In 2016, Census of Agriculture forms were delivered to farm operations by Canada Post. Once completed, the questionnaire was mailed back to the Data Operations Centre for processing. Respondents also had the option to complete and submit a questionnaire via the Internet. If it was determined that a questionnaire had not been received, or if data were missing, a follow-up was conducted by telephone. For a more detailed description of the collection process, please refer to Data collection.

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 were 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 was found had their information imputed with that of another Census of Population person having similar characteristics.

View the [Questionnaire(s) and reporting guide(s)](#) .

Error detection

Error detection is an integral part of both collection and data processing activities. Edits are 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 are flagged for the respondents to review.

Data captured from paper questionnaires are subjected to many rigorous quality control and processing edits to identify and resolve problems related to inaccurate, missing or inconsistent data.

During data processing, other edits are 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 are found, they can be corrected using the failed edit follow up process during collection or via the data editing and imputation processes post collection.

Extreme values are also identified, using automated methods based on the distribution of the collected information. Following their detection, these values are reviewed by subject-matter analysts in order to assess their reliability. Macrodata totals are 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 is made to minimize the non-sampling errors of omission, duplication, misclassification, reporting and processing.

Imputation

When non-response occurs, when respondents do not completely answer the questionnaire, or when reported data are considered incorrect during the error detection steps, imputation is used to fill in the missing information and modify the incorrect 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 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.

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 during the data validation process after thorough investigation.

Estimation

All units in the observed population are surveyed. Estimation of totals is done by simple aggregation of the values of all the farms that are found in a given domain. Estimates are computed for domains of interest such as farm type and provinces/territories.

Quality evaluation

Prior to the release, the data are analyzed for comparability with a variety of other data sources, such as previous census data and recent agriculture survey results. Macro level coherence with results from related economic indicators, historical trends, and information from other external sources (e.g. associations, agriculture publications, newspaper articles) is evaluated. Microdata records may be further validated using available provincial administrative commodity lists, and taxation declarations. Extra scrutiny is given to the largest farms or top commodity contributors.

The data quality assessment done during validation is presented to the Census of Agriculture Certification Committee for review and approval. This Committee's mandate is to approve the accuracy, coherence and reasonable quality of the Census data for each variable prior to publication. The members of this Committee are subject-matter agriculture experts, survey methodology specialists, as well as managers.

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. If necessary, data are suppressed to prevent direct or residual disclosure of identifiable data.

In order to prevent any data disclosure, confidentiality analysis on the tabular data to be published is done using the Statistics Canada generalized confidentiality system (G-CONFID). G-CONFID is used for primary suppression (direct disclosure) as well as for secondary suppression (residual disclosure). Direct disclosure occurs when the value in a tabulation cell is composed of or dominated by few farms while residual disclosure occurs when confidential information can be derived indirectly by piecing together information from different tabulation cells or data series. When the cells of tabular data are not sufficiently protected for confidentiality, suppression of cell values suggested by the G-CONFID system are applied to the tables before publication.

Revisions and seasonal adjustment

This methodology type does not apply to this statistical program.

Data accuracy

An integral part of each Census of Agriculture is the implementation of new or enhanced methods, procedures and technologies that improve not only the collection, but also the processing, validation

and dissemination of the data. Enhanced methods , procedures and technologies adopted for the 2016 Census of Agriculture include significant updates to the Statistics Canada Business Register in preparation of the Census, mailing invitation letters and/or questionnaires to the entire farm population with a recognized mailing address and an enhanced centralized telephone follow-up operation to resolve non-response as well as inconsistencies within questionnaires returned by respondents. In addition, to help ensure that data from the 2016 Census of Agriculture would be of consistently high quality, improved quality assurance and control procedures were   incorporated into each of the collection and data processing stages.

With projects as large and complex as the Censuses of Agriculture and Population, the estimates produced from them 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 data's usefulness for their particular applications as well as assess the risks involved in basing conclusions or decisions on them.

Errors can arise at virtually every stage of the census process, from preparing materials, through collecting data, to processing. Moreover, errors may be more predominant in certain areas of the country or vary according to the characteristic being measured. Some errors occur at random, and when individual responses are aggregated for a sufficiently large group they tend to cancel each other out. For errors of this nature, the larger the group, the more accurate the corresponding estimate. For this reason, data users are advised to be cautious when using estimates based on a small number of responses. Some errors, however, might occur more systematically and result in "biased" estimates. Because the bias from such errors is persistent no matter how large the group for which responses are aggregated, and because bias is particularly difficult to measure, systematic errors are a more serious problem for most data users than random errors.