Catalogue no. 89-653-X2024002
ISBN 978-0-660-72740-0

Indigenous Peoples Survey

Concepts and Methods Guide for the Indigenous Peoples Survey and Indigenous Peoples Survey – Nunavut Inuit Supplement, 2022

by Marie-Hélène Toupin and Amanda Varias

Release date: August 14, 2024

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 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.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of contents

Concepts and Methods Guide for the Indigenous Peoples Survey and Indigenous Peoples Survey – Nunavut Inuit Supplement, 2022

by **Marie-Hélène Toupin** and **Amanda Varias**

1 Introduction

1.1 Survey overview

The 2022 Indigenous Peoples Survey (IPS) is a national survey on the social and economic conditions of First Nations people living off reserve, Métis and Inuit, aged 1 and over as of April 27, 2022. The objectives of the IPS are to identify the needs of these Indigenous groups and to inform policy and programming activities aimed at improving the well-being of Indigenous peoples. The IPS aims to provide current and relevant data for a variety of stakeholders, including Indigenous organizations, communities, service providers, researchers, governments and the general public. Funding was provided by three federal departments: Crown-Indigenous Relations and Northern Affairs Canada, Indigenous Services Canada, and Employment and Social Development Canada.

The IPS was formerly called the Aboriginal Peoples Survey (APS). For the 2022 cycle, the name has been updated to the Indigenous Peoples Survey. “Indigenous” is a collective term used to encompass a variety of original peoples of North America and their descendants. The term “Indigenous” is increasingly replacing the term “Aboriginal,” in response to calls from within Indigenous communities. Following consultations with Indigenous communities and organizations, Statistics Canada has adopted the term “Indigenous” to refer to First Nations people, Métis and Inuit.

The IPS has been conducted by Statistics Canada since 1991, providing a range of social and economic indicators about Indigenous peoples. It is a postcensal survey, designed to follow and complement the census of population.

For the 2022 IPS, a supplemental survey called the Indigenous Peoples Survey–Nunavut Inuit Supplement (IPS–NIS), targeted at Inuit enrolled under the Nunavut Agreement, was added, as it was for the 2017 APS. This included a large supplementary sample of Inuit aged 15 and over in Nunavut, as well as an additional set of questions designed to learn more about the availability, interest and level of preparedness of Inuit enrolled under the Nunavut Agreement for government employment. The purpose of the 2022 IPS–NIS is to provide information for the Nunavut Inuit Labour Force Analysis (NILFA) project and find ways to increase Inuit employment in government, as required by Article 23 of the Nunavut Agreement. The IPS–NIS was funded by Crown-Indigenous Relations and Northern Affairs Canada (formerly Indigenous and Northern Affairs Canada) and Employment and Social Development Canada.

The 2022 IPS represents the sixth cycle of the survey and follows the thematic approach that was first introduced in the 2012 APS. The focus for the 2022 IPS is families and children. This focus will help to provide a deeper understanding of the challenges Indigenous families and children face at different times in their lives and culturally appropriate support for families to promote the safety and well-being of children, youth and adults. The 2022 IPS includes questions about child care, access to services, family stability, intergenerational trauma and discrimination, sense of belonging, and Indigenous languages and culture. This is in contrast with the 2017 APS, which focused on participation in the Canadian economy. Additional information collected in the 2022 IPS will ensure continuity with past iterations of the survey and provide core indicators in the areas of education, employment, housing, health and access to services. The IPS will continue to serve as an important source of information for First Nations, Métis and Inuit organizations; community planners; service providers; governments; and researchers.

This cycle of the IPS and IPS–NIS was conducted from May 11 to November 30, 2022, with in-person follow-up occurring from January 16 to March 31, 2023. Approximately 74,000 people were selected to participate in the survey, and the final response rate was 52.3%. For the 2022 IPS, there was an oversample of veterans women, as well as Inuit aged 15 and over enrolled under the Nunavut Agreement who received the IPS–NIS. The survey design allowed for the production of reliable data for each of the provinces and territories (Atlantic provinces grouped), as well as for each of the four Inuit regions: Nunatsiavut (Northern coastal Labrador), Nunavik (Northern Quebec), the

territory of Nunavut and the Inuvialuit region of the Northwest Territories. The survey also targeted five particular age groups: 1 to 5 years, 6 to 14 years, 18 to 24 years, 25 to 54 years and 55 years and over, and although not targeted, 15- to 17-year-olds were also included in the sample. For those children under the age of 15 included in the 2022 IPS sample, the parent, guardian or the person most knowledgeable about the child completed the survey for the child. This contrasts with the 2017 APS, which selected only those aged 15 and over for its sample.

It is important to note that unlike the 2017 IPS and IPS–NIS, when there were two separate data files, there is now one shared masterfile for the 2022 IPS and IPS–NIS. This analytical file will include individuals selected for the main IPS sample as well as individuals who were selected for the IPS–NIS.

This guide describes both the IPS and the IPS–NIS. When describing both, it will be written as "the 2022 IPS and IPS–NIS," or the shorter term "the survey" will be used. The same notions and processes are used for both, unless otherwise stated explicitly within each section of this guide.

1.2 Purpose of the Concepts and Methods Guide

This guide is intended to provide a detailed review of both the 2022 IPS and IPS–NIS with respect to subject matter and methodological approaches. It is designed to help data users by serving as a guide to the concepts and measures of the survey as well as the technical details of the survey's design, field work and data processing. This guide is meant to provide users with helpful information on how to use and interpret survey results. The discussion on data quality also allows users to review the strengths and limitations of the data for their particular needs.

Chapter 1 of this guide provides an overview of the 2022 IPS and IPS–NIS by introducing the survey background and objectives. Chapter 2 outlines the survey's themes and explains the key concepts and definitions used for the survey. Chapters 3 to 6 cover important aspects of the survey methodology, sampling design, data collection and processing. Chapters 7 and 8 review issues of data quality and caution users about comparing 2022 IPS or IPS–NIS data with data from other sources. Chapter 9 outlines the survey products available to the public, including data tables, analytical articles and reference material. The appendices provide a comprehensive list of survey indicators, extra coding categories and standard classifications used on both the IPS and the IPS–NIS. Lastly, a glossary of survey terms and information on confidence intervals is also provided.

2 Survey content: concepts and questions

2.1 Content development process

The content for the 2022 Indigenous Peoples Survey (IPS) and IPS–NIS was developed by Statistics Canada in collaboration with three federal funding departments: Crown-Indigenous Relations and Northern Affairs Canada, Indigenous Services Canada, and Employment and Social Development Canada (ESDC). To identify Indigenous-specific needs, feedback was requested from the following six national Indigenous organizations (NIOs) on proposed new topics and any other key areas of interest: Assembly of First Nations, Congress of Aboriginal Peoples, Inuit Tapiriit Kanatami, Métis National Council, National Association of Friendship Centres and Native Women's Association of Canada. The content was designed to meet the commitments put forward in the Treasury Board submission, while also meeting the data needs and priorities of the funding departments and NIOs. The 2022 IPS also drew on many key indicators from previous cycles of the IPS, which were developed in collaboration with NIOs for historical compatibility and cross-sectional analysis.

With respect to new content requirements for 2022, relevant standardized and well-established measures used on other Statistics Canada surveys were gathered and reviewed as potential indicators. These indicators increase opportunities to compare responses between the IPS and other Statistics Canada surveys. In addition, to recognize the new theme of Indigenous families and children, new IPS content was sought that would allow for the measurement of a diverse range of social and economic outcomes related to education, employment, health and access to services among First Nations people living off reserve, Métis and Inuit. With the goal of identifying demographic and socioeconomic characteristics of the Indigenous population, the new content was developed to cover topics such as access to services; child care; family stability; trauma and discrimination; missing and murdered Indigenous women, girls and Two-Spirit people; and food security.

As previously mentioned, Inuit aged 15 and over enrolled under the Nunavut Agreement received the IPS and the IPS–NIS questions. The IPS–NIS questions were developed by the NILFA Technical Working Group, which included representatives from Nunavut Tunngavik Incorporated (NTI), the Government of Nunavut, Pilimmaksaivik, ESDC and Statistics Canada. The questions were designed to learn more from Inuit about interest and availability in government employment, future employment plans, interest in training and relevant skills experience. The 2022 IPS–NIS also drew on many key indicators from the previous cycle of the Aboriginal Peoples Survey–Nunavut Inuit Supplement in 2017.

Once the questions were developed for the IPS and IPS–NIS, they underwent qualitative testing to ensure that they were clearly understood by respondents and would yield valid results. The IPS questionnaire, excluding the IPS–NIS content, was tested with First Nations people living off reserve, Métis and Inuit across Canada. For the IPS–NIS specifically, testing was conducted in Ottawa, Ontario, and Iqaluit, Nunavut. Observations and feedback received from this testing allowed the final questionnaire to be determined.

The questions in the 2022 IPS and IPS–NIS were designed with an electronic questionnaire (EQ) environment to replace previous cycles of this survey, which had only been collected by computer-assisted interview (CAI) or on paper. EQ incorporated many features that served to maximize the quality of data collection. An EQ was developed for three different modes of collection used for this survey: computer-assisted telephone interview (CATI), computer-assisted personal interview (CAPI) and respondent electronic questionnaire (rEQ). Refer to Section 4 for more details on the data collection.

To supplement data for this survey, variables from the census were added to the analytical file. Respondents were informed that their census data would be linked and they were able to choose not to link their census data, but they all implicitly agreed.

2.2 Indigenous identity groups

A definition of Indigenous identity

The 2022 IPS and IPS–NIS identifies the Indigenous identity population as anyone who self-reported being at least one of the following:

- an Indigenous person, that is, First Nations (North American Indian), Métis or Inuk (Inuit)
- a Status Indian, that is, a Registered or Treaty Indian as defined by the *Indian Act* of Canada
- a member of a First Nation or Indian band.

The first criterion above is referred to as “Indigenous self-reporting.” A respondent may self-report as belonging to one or more of the particular Indigenous groups mentioned: First Nations, Métis or Inuit. First Nations includes Status and non-Status Indians. It should also be noted that some respondents use the term *First Nations* while others use the term *North American Indian*.

As the survey definition implies, a person does not need to self-report as Indigenous (e.g., as First Nations, Métis or Inuk) to be considered part of the Indigenous identity population. If a person has reported being a Status Indian or a member of a First Nation or Indian band, they are considered to be part of the Indigenous identity population, regardless of their responses to the Indigenous group questions. Individuals who are not First Nations (North American Indian), Métis or Inuk (Inuit) but who have Registered or Treaty Indian status or membership in a First Nation or Indian band are considered as “Indigenous responses not included elsewhere.”

With respect to the measurement of “Status Indian,” the 2022 IPS includes everyone who said that they are a Registered or Treaty Indian as defined by the *Indian Act*. Registered Indians are persons who are registered under the *Indian Act* of Canada. Treaty Indians are persons who belong to a First Nation or Indian band that signed a treaty with the Crown.

The universe or target population for the 2022 IPS are persons who meet the definition of Indigenous identity. Although the survey data will support distinct analyses for First Nations living off reserve, Métis and Inuit, the IPS was not designed to produce estimates for the group “Indigenous responses not included elsewhere.” This

is a very small group and caution should be used when conducting analyses for this group (see Chapter 8). The universe or target population for the 2022 IPS–NIS are Inuit aged 15 and over who are enrolled under the Nunavut Agreement.

The 2022 IPS did not include First Nations people who were living in First Nations communities (on reserve). For current information on First Nations people living on reserve, please refer to the 2021 Census of Population or to a similarly themed survey, the First Nations Regional Early Childhood, Education and Employment Survey (FNREEES), conducted on reserve by the First Nations Information Governance Centre. More information on this survey is available on the First Nations Information Governance Centre website.

Questionnaire items for Indigenous identification

Table 1 below lists the three Indigenous identification questions asked in the 2022 IPS and shows the Indigenous identity classification derived from the answers provided to these questions. As shown, in order for someone to be part of the Indigenous identity population, respondents needed to have a YES response to ID_Q05 (self-reported Indigenous person), to ID_Q15 (Status Indian) or to ID_Q20 (Member of First Nation or an Indian band). Note that those who answered NO to all these questions were classified as non-Indigenous and considered out of scope for the survey. Adjustments to survey weights were made accordingly.

Note that ID_Q10 (Inuit land claims agreement) and ID_Q25 (Métis Settlement or organization) are asked in the questionnaire but not used to determine Indigenous identity.

Table 1
The Indigenous identification questions asked in the 2022 IPS and IPS–NIS questionnaire

Questionnaire flow	Indigenous identity classification
ID_Q05 - Are you First Nations, Métis or Inuk (Inuit)? First Nations (North American Indian) includes Status and Non-Status Indians.	
ID_Q05_2 - Yes, First Nations (North American Indian)	First Nations
ID_Q05_3 - Yes, Métis	Métis
ID_Q05_4 - Yes, Inuk (Inuit)	Inuit
No … see subsequent questions	
ID_Q15 - Are you a Status Indian (Registered or Treaty Indian as defined by the Indian Act of Canada)?	
• No	
• Yes	Indigenous responses not included elsewhere
ID_Q20 - Are you a member of a First Nation or Indian band? If “Yes,” specify name of First Nation or Indian band.	
• No	
• Yes	Indigenous responses not included elsewhere
No to ID_Q05, and ID_Q15, and ID_Q20	Non-Indigenous (out of scope)

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

Indigenous identity variables available to data users

Data made available from the 2022 IPS and IPS–NIS will provide analytical variables for each aspect of Indigenous identity. These variables will be central to data users for conducting their analyses of subject matter themes for each group— First Nations people living off reserve, Métis, Inuit and Indigenous responses not included elsewhere. Indigenous identity variables will include indicators of both single and multiple identities (for example, persons who reported being both a First Nations person and Métis). Variables on Status Indians (Registered or Treaty) will also be available for analysis. A variable on membership in a Métis organization or Settlement will be available for the first time in the 2022 IPS and IPS–NIS file. Derived variables for Indigenous ancestry, based on data from the 2021 Census, will be provided since Indigenous ancestry was not measured directly by the 2022 IPS and IPS–NIS.

2.3 Levels of geography

The 2022 IPS ensured coverage of certain core geographic domains. These included provinces and territories (with the Atlantic provinces grouped) and the four Inuit regions of Inuit Nunangat. The 2022 IPS–NIS sample ensured coverage at lower-level geographies in the territory of Nunavut.

Inuit Nunangat is the homeland of Inuit of Canada. It includes the communities located in the four Inuit regions:

- Nunatsiavut (Northern coastal Labrador);
- Nunavik (Northern Quebec);
- the territory of Nunavut;
- the Inuvialuit region of the Northwest Territories.

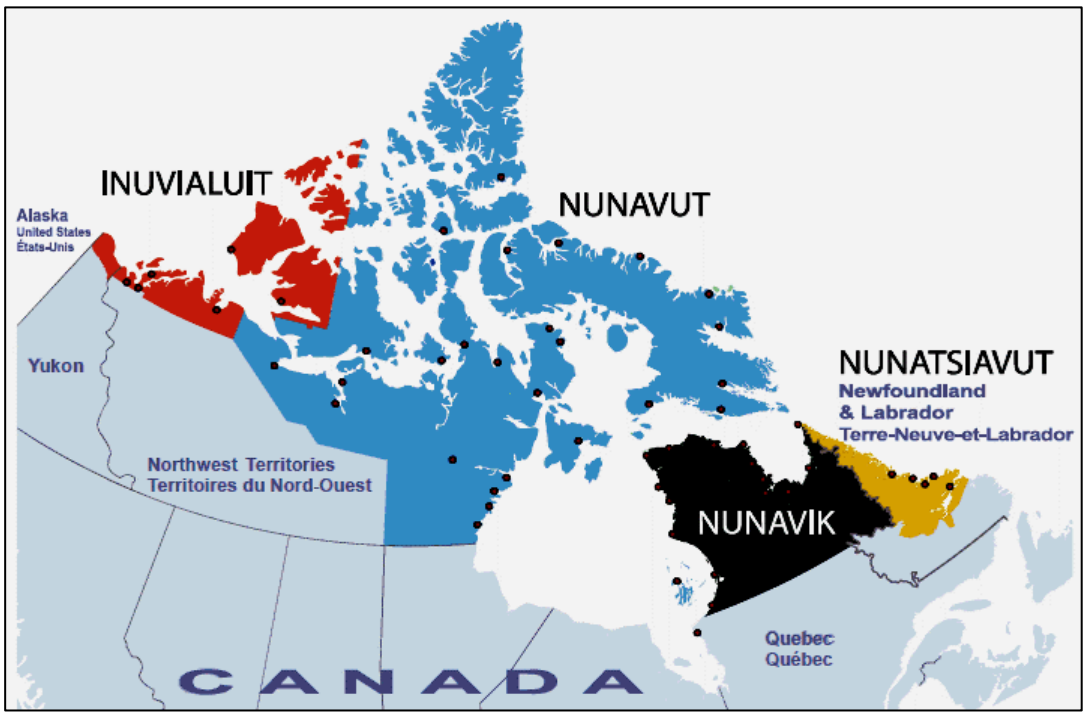
These regions collectively encompass the area traditionally occupied by Inuit in Canada (see Map 1).

These geographic domains were targeted by the 2022 IPS to ensure that adequate data estimates would be available at these levels of geography. Other geographic variables are also available in the 2022 IPS database, based on geographies from the 2021 Census of Population. These include census metropolitan areas (CMAs), census subdivisions (CSDs) and population centres, among others. In addition, geographies will include health regions across Canada that represent administrative areas or regions of interest to health authorities.

Users should note that not all survey data can be cross-tabulated or analyzed at detailed levels of geography. Some data tables may be available for more detailed geographies, but the reliability of data estimates at each level will need to be examined on a case-by-case basis.

Data users are directed to the 2022 IPS and IPS–NIS Data Dictionary for a complete list of geographies available from the survey. For details on how to obtain the data dictionary, please contact the Centre for Indigenous Statistics and Partnerships data support unit at Statistics Canada at statcan.cispdatasupport-cspasoutienauxdonnees.statcan@statcan.gc.ca. The *2021 Census Dictionary* also defines geographies relevant to the IPS and IPS–NIS.

Map 1
The four regions of Inuit Nunangat



Source: Inuit Tapiriit Kanatami.

2.4 Survey content themes

Reflecting a focused thematic approach, the 2022 IPS collected data falling into nine content areas:

1. Demographics
2. Family stability
3. Economic wellbeing
4. Education
5. Health
6. Support to families
7. Trauma and discrimination
8. Labour
9. Belonging, language and culture.

These blocks are placed throughout the survey, and contain questions specifically related to Indigenous people in Canada. A comprehensive description of all the variables available from the survey data is provided in the 2022 IPS and IPS–NIS Data Dictionary.

Demographics

The demographics component of the 2022 IPS is intended to help better understand the current living situations of those 1 year and over. It was built on the Identification component of the 2017 Aboriginal Peoples Survey (APS) as well as that of the 2021 Census, and further delves into the identities of Indigenous peoples. These demographic blocks consist of questions about Indigenous status and land claims, household composition, gender and sexual orientation and person most knowledgeable profile.

Family stability

This component was designed to obtain information on family history and mobility. It consists of questions about the living arrangements of children, the number of times a child moved homes since birth or adoption, family ties, and mobility of respondents aged 15 and over.

Economic well-being

The content of the economic well-being component includes the following modules: Food security, Basic needs, Housing, Neighbourhood safety, Information technology, and Income. Continued inclusion of questions measuring food security, housing, basic needs, income and information technology will allow for ongoing tracking of economic well-being for First Nations people living off reserve, Métis and Inuit.

Education

The education blocks collect information about adult, child and youth education, and postsecondary education. The education module dedicated to children and youth asks questions about current level of school they are attending, highest level completed, number of schools attended, reason for changing schools, description of school in terms of Indigenous culture, child's school performance, family involvement with school and child's studies, leaving school and returning, and reasons for doing each. The module about postsecondary education collects information on topics such as education taken, education completed, types of programs completed, major field of study, and full-time or part-time enrolment. These data allow for in-depth analysis of characteristics of Indigenous people who are most likely to successfully complete postsecondary education and help better understand the labour market returns of postsecondary education and types of programs that provide the best returns.

Health

The 2022 IPS Health component is designed to obtain information about the health of Indigenous people. As such, data are collected on a variety of topics: general health, general mental health, pregnancy and childbirth, maternal health, chronic conditions, eating habits, smoking, alcohol consumption, drug use, disease screening, consultations about mental health, suicide, general health, oral health, disability, and the COVID-19 pandemic. As a result of the addition of children under the age of 15 in the 2022 IPS, questions asked only to this population were added on chronic conditions, injuries, eating habits, consultations about mental health, oral health and physical activities.

Support to families

This component contains two modules: Access to services and Child care. For the first time, the 2022 IPS collects information about access to services funded under Jordan's Principle and Inuit Child First Initiative programs. These programs help children access government-funded health, social and educational products; services; and supports they need when they need them. The module collects information about knowledge of the programs, whether respondents ever applied, reasons for not applying, types of products, services and supports applied for under the programs, and status of application. The Child care module collects information about types and costs of child care currently in use (or reasons for not using child care); satisfaction with child care currently in use, especially in terms of cultural appropriateness; and barriers to preferred child care.

Trauma and discrimination

New content on trauma and discrimination was added to the 2022 IPS, asking questions about family separation, missing and murdered Indigenous women, girls, and Two-Spirit people, discrimination because of Indigenous identity, and victimization. Information was collected about personal or familial residential school attendance, the experience as a child being under the custody of a child welfare agency or being placed in foster care, victims' experiences of physical assault, as well as the presence of missing or murdered Indigenous women, girls and Two-Spirit people in one's life. This information can be used to analyze the impacts of the systematic oppression of Indigenous women, girls and Two-Spirit people on those directly or indirectly related to the victims.

Labour

The Labour component on the 2022 IPS contains the following modules: Labour market activities, Labour force status, Class of worker, Industry, Occupation, Usual hours of work, and Child care impact on employment. The Child care impact on employment module collects information about types of child care currently in use and the extent to which caring for children impacts respondents' employment.

Belonging, language and culture

This component gathers information on the following topics: sense of belonging; harvesting, handcrafting and cultural activities; activities for children; and language. These questions were introduced to measure respondents' cultural connectedness and sense of belonging to their Indigenous identity and to Canada.

The 2022 IPS-NIS, Inuit aged 15 and over enrolled under the Nunavut Agreement were asked an additional set of questions designed around three themes:

1. Availability for government employment
2. Interest in government employment
3. Preparedness for government employment

The IPS–NIS questions were asked at the end of the IPS questionnaire.

2.5 Questionnaire modules

This section provides a list of the modules on the 2022 IPS and the IPS–NIS questionnaire. Appendix A1 provides a detailed list of what is measured in each of the IPS modules. Appendix A2 provides a detailed list of what is measured in each module of the IPS–NIS. Appendix B lists the extra categories created during survey coding. Appendix C describes the standard classifications used to create indicators for open-ended survey questions.

Modules in the 2022 Indigenous Peoples Survey questionnaire, grouped by theme:

- Demographics
  - Indigenous identity
  - Household composition
  - Sex and gender
  - Sexual orientation
  - Person most knowledgeable profile
- Family stability
  - Family history
  - Mobility – place of residence
- Belonging, language and culture
  - Sense of belonging
  - Harvesting and handcrafting and cultural activities
  - Activities – Child
  - Language
  - Language – Child ages 1 to 5
- Economic well-being
  - Food security
  - Basic needs
  - Housing
  - Neighbourhood safety
  - Sources of personal income
  - Total personal income
  - Employment income
  - Information technology
- Health
  - General health 1
  - General mental health
  - Pregnancy and childbirth
  - Maternal health
  - Height and weight
  - Birth weight of the child
  - Chronic conditions – Adult
  - Chronic conditions – Child
  - Injuries
  - Eating habits – Adult
  - Eating habits – Child ages 6 to 14

    - Eating habits – Child ages 1 to 5
    - Eating habits – Child ages 1 to 14
    - Smoking
    - Alcohol consumption
    - Drug use
    - Disease screening
    - Consultations about mental health – Adult
    - Consultations about mental health – Child 6 to 14
    - Suicide
    - General health 2
    - Oral health – Adult
    - Oral health – Child
    - Physical activity
    - Disability
    - COVID-19 pandemic
- Support to families
    - Child care
    - Access to services
- Education
    - Education – Adult
    - Education – Children and youth
    - Postsecondary education
- Trauma and discrimination
    - Family separation
    - Missing and murdered Indigenous women, girls and Two-Spirit people
    - Discrimination
    - Victimization
- Labour
    - Labour market activities
    - Labour Force Status
    - Class of worker
    - Industry
    - Occupation
    - Usual hours of work
    - Child care impact on employment

Modules in the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement questionnaire:

- Trust in public institutions
- Availability for government employment
- Past government employment
- Relief workers and substitute teachers
- Permanent work
- Looking for work

- Labour market attachment
- Labour mobility
- Interest in government employment
- Previous applications to government employment
- Plans to apply for government employment in the next 12 months
- Postsecondary preparedness and distance education
- Skills and training
- Interest in training for government employment
- Pre-employment training
- Plans for further education
- Skill-relevant experiences
- Language fluency for work
- Retirement income

2.6 Linked content from the 2021 Census of Population

The IPS and IPS–NIS is a postcensal survey, with its sample drawn from census respondents. In 2001, 2006, 2012 and 2017, IPS data were linked with each corresponding Census of Population (or National Household Survey). For 2022, the sample was drawn from 2021 Census of Population respondents, and the final microdata file was linked with the 2021 Census of Population dissemination database.

The benefits of a record linkage between the IPS and IPS–NIS and the census are reduced response burden for the target population of the survey; the establishment of survey weights, which are crucial to providing valid estimates; and the creation of a comprehensive microdata file, which can be used by data analysts to extend their learning and to inform policy and program development for Indigenous peoples in Canada. Together, data from the two sources will provide a detailed statistical portrait of First Nations people living off reserve, Métis and Inuit in Canada, as well as a detailed statistical portrait of Inuit enrolled under the Nunavut Agreement. These data are not available from any other source.

At the time of data collection, all census respondents were informed that the information they provided might be used to support other Statistics Canada surveys. Specifically, the message on the cover of the 2021 Census questionnaire stated, "The data may also be used by Statistics Canada for other statistical and research purposes or be combined with other survey or administrative data sources." As well, 2022 IPS and IPS–NIS questionnaire contained a Getting Started page, in which respondents were told about the purpose of the survey, its voluntary nature and that:

> "To enhance the data from this survey and to reduce the reporting burden, Statistics Canada will combine the information you provide with information from the 2021 Census of Population.
>
> Statistics Canada may also combine the information you provide with other survey or administrative data sources."

It should be noted that respondents can choose not to have their data linked from the IPS and IPS–NIS to other surveys or administrative data sources. All products containing linked data are disseminated in accordance with Statistics Canada's policies, guidelines and standards. Only aggregate statistical estimates that conform to the confidentiality provisions of the *Statistics Act* can be released outside of Statistics Canada. All data requests for survey tabulations from the IPS and IPS–NIS are screened for confidentiality and the aggregate data are rounded before being released to clients.

More than 250 variables from the 2021 Census were linked to the final 2022 IPS and IPS–NIS file. The list below indicates the type of census variables that have been appended to the analytical file. It is important to note that these census variables, having been obtained from the 2021 Census responses for 2022 survey respondents, refer to each respondent's situation on the census reference day, May 11, 2021. Users should be aware that in some

cases, the respondent may have moved, had a change in the composition of their household or had a change in employment between the date of the 2021 Census and the date of the IPS interview. In other words, some of the information provided by the 2021 Census may not reflect the respondent's situation when the IPS questionnaire was completed.

A complete list of linked census variables and their accompanying notes are provided in the *2022 IPS and IPS–NIS Data Dictionary*, which accompanies the analytical file.

Household-level variables

- Geography - including CMAs, CSDs and health regions
- Housing - including tenure, number of rooms in dwelling and condition of dwelling
- Family - including family status, presence of grandparents and multigenerational status of the household

Person-level variables

- First Nation or Indian band
- Labour - including, class of worker, full- or part-time weeks worked in 2020 and hours worked for pay or in self-employment
- Place of work - including province, census division or CSD
- Mobility - including place of residence one and five years ago, inter- and intra-CMA mobility status, and inter- and intra-provincial mobility status
- Income - including family income, employment income, number of earners and COVID-19 benefits
- Language - including first official language spoken, home language, Indigenous home language and Inuit home language

2.7 Content input from other surveys

The 2022 IPS drew on questionnaire content from multiple sources. The 2017 APS was used as a starting point and served as a key source of well-established questions for First Nations people living off reserve, Métis and Inuit. Questions were also drawn from other surveys as much as possible to enhance the robustness of the data and to allow for comparability across data sources, where appropriate. Sometimes these questions were modified for a better fit with the IPS. In such instances, the data are not directly comparable.

This section outlines the primary sources used for 2022 IPS questionnaire development in relation to the main survey themes. It should be noted that these sources do not represent a comprehensive list of all the Statistics Canada surveys that include indicators found in the IPS.

Demographics

- 2021 Census of Population
- 2017 Aboriginal Peoples Survey
- 2006 Aboriginal Children's Survey

Family stability

- 2017 and 2012 Aboriginal Peoples Survey
- 2019 Canadian Health Survey on Children and Youth

Economic wellbeing

- 2017 Aboriginal Peoples Survey
- 2021 Canadian Community Health Survey
- 2020 Canadian Internet Use Survey

Education

- 2017 Aboriginal Peoples Survey
- 2012 Aboriginal Peoples Survey
- 2006 Aboriginal Children's Survey

Health

- 2017, 2012 and 2006 Aboriginal Children's Survey
- 2006 Canadian Internet Use Survey, 2006 Maternity Experiences Survey
- 2022 and 2021 Canadian Community Health Survey
- 2015-2016 First Nations Regional Health Survey
- 2019 Canadian Health Survey on Children and Youth
- 2020/2021 Canadian Perspectives Survey Series - Impacts of COVID-19

Support to families

- 2006 Aboriginal Children's Survey
- 2018 Survey on Early Learning and Child Care Arrangements

Trauma and discrimination

- 2017 and 2006 Aboriginal Peoples Survey
- 2006 Aboriginal Children's Survey
- 2018 Survey of Safety in Public and Private Spaces

Labour

- 2017 Aboriginal Peoples Survey
- 2019 Survey on Early Learning and Child Care Arrangements

Belonging, language and culture

- 2017 Aboriginal Peoples Survey
- 2006 Aboriginal Peoples Survey
- 2006 Aboriginal Children's Survey
- 2020 General Social Survey - Social identity
- 2006 Aboriginal Peoples Survey, Arctic Supplement

The survey questions for the 2022 IPS–NIS were developed by a Technical Working Group which included NTI, the Government of Nunavut, Pilimmaksaivik, ESDC and Statistics Canada. The 2017 APS–NIS was used as a starting point, with some questions modified after consulting the 2021 Nunavut Government Employee Survey questionnaire. In addition, certain questions related to the theme of labour that were part of the 2017 APS but dropped from the 2022 IPS were included in the 2022 IPS–NIS to ensure availability, interest and preparedness for government employment could be properly measured.

3 Survey design

3.1 Target populations and coverage

The 2022 IPS target population was composed of Canada's Indigenous identity population aged 1 year and over as of April 27, 2022, living in private dwellings, excluding people living on Indian reserves and settlements, and in certain First Nations communities in Yukon and the Northwest Territories. The list of census subdivisions (CSDs) included in Yukon and the Northwest Territories is presented in Table 2.

Table 2
List of census subdivisions in Yukon and Northwest Territories included in the Indigenous Peoples Survey

CSD No.	Name of census subdivision	Census subdvision type
Yukon		
6001003	Watson Lake	T
6001004	Faro	T
6001009	Whitehorse	CY
6001029	Dawson	T
6001044	Mt. Lorne	HAM
6001045	Yukon, Unorganized	NO
6001046	Swift River	SÉ
6001049	Destruction Bay	SÉ
6001050	Stewart Crossing	SÉ
6001052	Keno Hill	SÉ
6001055	Ibex Valley	HAM
6001058	Marsh Lake	NO
6001059	Macpherson-Grizzly Valley	NO
6001060	Whitehorse, Unorganized	NO
6001065	North Slope	NO
Northwest Territories		
6101014	Paulatuk	HAM
6101017	Inuvik	T
6101025	Aklavik	HAM
6101036	Tuktoyaktuk	HAM
6101041	Sachs Harbour	HAM
6101063	Region 1, Unorganized	NO
6101095	Ulukhaktok	HAM
6102007	Norman Wells	T
6102063	Region 2, Unorganized	NO
6103097	Region 3, Unorganized	NO
6104097	Region 4, Unorganized	NO
6105003	Enterprise	HAM
6105016	Hay River	T
6105026	Reliance	SET
6105097	Region 5, Unorganized	NO
6106023	Yellowknife	CY
6106097	Region 6, Unorganized	NO

Source: Statistics Canada, Canadian Census of the Population, 2021.

The CSD types can be found in Table 1.5 of the *Census Dictionary*.

The target population for the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement (IPS–NIS) was Inuit aged 15 and over as of April 27, 2022, who are enrolled under the Nunavut Agreement, living in Nunavut at the time of the 2021 Census in private dwellings, but excluding those living on Indian reserves or Indian settlements.

3.1.1 Identifying the Indigenous population

The IPS sample was selected from respondents who reported either Indigenous identity or Indigenous ancestry on the 2021 Census long-form questionnaire. More precisely, the sample was selected based on responses to Questions 23, 24, 26 and 27, on either Form 2A-L or Form 2A-R.

The sample for the Nunavut Inuit Supplement, the IPS–NIS, was selected from respondents who reported an Inuit identity on the 2021 Census long-form questionnaire and were enrolled under the Nunavut Agreement. More specifically, the sample was selected on the basis of responses to questions 24 and 29 of forms 2A-L and 2A-R, which are described below.

Form 2A-L was completed by self-enumeration and was administered to approximately one in four households in most parts of Canada (2A-L regions). Other than the basic census demographic questions (name, sex at birth and gender, date of birth, legal marital status, common-law status, relationship to Person 1, and languages), the 2021 Census Form 2A-L included questions on labour market activity, income, education, activity limitations, citizenship, housing and ethnic origin.

Form 2A-R – identical in content to the 2A-L form, except for some adapted examples and two additional questions on band housing – was completed by self-enumeration or administered by personal interview to all households in remote areas, Inuit communities and Indian reserves and settlements (2A-R regions).

The four screening questions used to identify the Indigenous population were the ones on ethnic or cultural origin (Question 23), Indigenous self-reporting (Question 24), whether the person reported being a Status Indian (Registered or Treaty Indian, Question 26), and membership in a First Nation or Indian band (Question 27).

The reporting of an Indigenous origin in Question 23 determines the Indigenous ancestry population (or ancestry population).

In the 2021 Census, the population with Indigenous identity is derived from three questions: 24, 26 and 27. The concept of Indigenous identity refers to people who have (a) self-reported as at least one Indigenous group, namely, First Nations (North American Indian), Métis or Inuit; (b) reported being a Status Indian (Registered Indian or Treaty Indian, as defined by the *Indian Act* of Canada); or (c) reported being a member of a First Nation or Indian band.

Within the context of the IPS, the Indigenous ancestry-only population (or ancestry-only population) is made up of individuals with Indigenous ancestry but who do not have an Indigenous identity according to the census. The Indigenous identity concept in the IPS is the same as that in the census, and is defined based on a very similar set of questions (see Table 1 in Section 2.2).

2021 Census — 2A-L, Question 23
What were the ethnic or cultural origins of this person's ancestors?

Ancestors may have Indigenous origins, or origins that refer to different countries, or other origins that may not refer to different countries.

For examples of ethnic or cultural origins, visit Examples of ethnic or cultural origins.

- Specify as many origins as applicable using capital letters.

2021 Census — 2A-L, Question 24
Is this person First Nations, Métis or Inuk (Inuit)?

Note: First Nations (North American Indian) includes Status and Non-Status Indians.

If "Yes," mark the circle(s) that best describe(s) this person now with an "x."

- No, not an Indigenous person.
  - ▶ Continue with the next question.

- Yes, First Nations (North American Indian).
    - ▶ Go to Question 26.
- Yes, Métis.
    - ▶ Go to Question 26.
- Yes, Inuk (Inuit).
    - ▶ Go to Question 26.

2021 Census — 2A-L, Question 26

Is this person a Status Indian (Registered or Treaty Indian as defined by the *Indian Act* of Canada)?

- No
- Yes, Status Indian (Registered or Treaty)

2021 Census — 2A-L, Question 27

Is this person a member of a First Nation or Indian band?

If "Yes," which First Nation or Indian band?

For example, Atikamekw of Manawan, Sturgeon Lake First Nation, Soowahlie Indian Band.

- No
- Yes, member of a First Nation or Indian band
    - ▶ Specify name of First Nation or Indian band:

2021 Census — 2A-L, Question 29

Is this person enrolled under, or a beneficiary of, an Inuit land claims agreement?

- No
- Yes

  Which Inuit land claims agreement?
    - ▶ Inuvialuit Final Agreement
    - ▶ Nunavut Agreement (Nunavut Land Claims Agreement)
    - ▶ James Bay and Northern Quebec Agreement (Nunavik)
    - ▶ Labrador Inuit Land Claims Agreement (Nunatsiavut)
    - ▶ or
        - ▪ Specify agreement:

For the purposes of 2022 IPS sampling, the Indigenous population includes both the Indigenous identity population and the ancestry-only population. Although the Indigenous ancestry-only population was not part of the target population for the 2022 IPS, like for the 2012 and 2017 Aboriginal Peoples Survey (APS), it was still sampled, because it had been noted in previous cycles of the survey that about a quarter of the Indigenous ancestry-only population in the census reported an Indigenous identity in the APS.

For the 2022 IPS–NIS, only individuals with Inuit identity (single or multiple) and enrolled under the Nunavut Agreement (Question 29) were considered in the sample. Note that individuals with Inuit-only ancestry are not included in the supplement.

3.1.2 Survey reference date

April 27, 2022, was chosen as the reference date for the 2022 IPS and IPS–NIS. This date corresponds to the planned launch date of data collection activities. Age is established based on this reference date.

3.1.3 2021 Census frame

The IPS and IPS–NIS sample was selected from the unedited, non-imputed database of the 2021 Census, which is the census database referred to as the Response Database (RDB). Before selecting the sample, the survey methodology team developed an editing and imputation strategy to deal with missing values in any of the four census screening questions or in the variables used during stratification (including age and certain education variables), as well as for individuals with certain characteristics that appeared to be inconsistent with being Indigenous.

The first step in selecting the sample was to include all individuals who reported Indigenous identity or ancestry on the 2021 Census in the survey frame. In the second step, all individuals who were part of the same households as the units in the initial frame were added to the frame. These additional individuals were kept in the frame only if their responses to screening questions were missing and the characteristics of the people in the household who had answered the screening questions indicated that they would have had a good chance of having Indigenous identity or ancestry. Hence, an individual with missing responses to the screening questions on identity would normally have been imputed as having Indigenous identity if at least 50% of the members of the same household who completed the identity screening questions had Indigenous identity. Similarly, an individual with non-response to the ancestry screening question would normally have been imputed as having Indigenous ancestry if at least 50% of the members of the same household who completed the ancestry screening question had Indigenous ancestry.

Once processing was complete, individuals under 1 year of age as of April 27, 2022, those living on reserves or in certain communities in Yukon and the Northwest Territories, and those who no longer self-identified as Indigenous were excluded from the survey frame following processing.

The 2022 IPS–NIS survey frame is a subset of the 2022 IPS survey frame. In fact, only individuals aged 15 and over living in Nunavut at the time of the 2021 Census reporting Inuit identity and enrolled under the Nunavut Agreement were kept in the survey frame.

3.2 Sampling design

3.2.1 Domains of estimation

Domains of estimation are groups of units for which estimates are targeted and must be produced with an "acceptable" level of precision. To be effective, stratification needs to be based on the different domains of estimation.

3.2.1.1 Domains of estimation of the Indigenous Peoples Survey

In the case of the IPS, the domains of estimation corresponded to geographical regions for which estimates with an "acceptable" level of precision for a particular Indigenous group (i.e., First Nations, Métis or Inuit) and particular age group were targeted.

An example of a domain of estimation would be Métis in Alberta aged 25 to 54 years. During stratification, the Métis Indigenous group comprised individuals reporting a Métis identity only to Question 24 on the census or, for individuals without Indigenous identity (Indigenous ancestry-only population), reporting Métis ancestry only to Question 23 (with or without non-Indigenous ancestry). In fact, people with Indigenous ancestry only were not part of the survey's target population. However, they were sampled because they had a relatively high chance of reporting an Indigenous identity in the survey, as described in Subsection 3.1.1. This is why the term "stratification-specific domains of estimation" is used rather than the term "survey-specific domains of estimation."

More precisely, the stratification-specific domains of estimation were created by cross-tabulating the following variables:

- Geography
    - Inuit regions

        - Outside Inuit regions
            - Province or territory
            - Atlantic provinces grouped
- Age group[1]
    - 1 to 5 years of age
    - 6 to 14 years of age
    - 18 to 24 years of age
    - 25 to 54 years of age
    - 55 years and over
- Indigenous group
    - Inuit in Inuit regions
    - Inuit outside Inuit regions (rest of Canada)
    - Indigenous groups (excluding Inuit) combined for Atlantic Canada (outside Nunatsiavut), Quebec (outside Nunavik), Yukon and Northwest Territories (outside Inuvialuit)
    - For Ontario, Manitoba, Saskatchewan, Alberta and British Columbia
        - First Nations people with Indian Status living off reserve
        - First Nations people without Indian Status living off reserve
        - Métis

Below is how the Indigenous groups were defined during stratification based on responses to the census:

- First Nations people with Indian Status living off reserve — individuals answering First Nations only to Question 24 and answering YES to Question 26 on the census.
- First Nations people without Indian Status living off reserve — individuals answering First Nations only to Question 24 and answering NO to Question 26, or individuals without Indigenous identity but answering First Nations only to Question 23 (with or without non-Indigenous ancestry).
- Métis only — individuals answering Métis only to Question 24, or individuals without Indigenous identity but answering Métis only to Question 23 (with or without non-Indigenous ancestry).
- Inuit only — individuals answering Inuit only to Question 24, or individuals without Indigenous identity but answering Inuit only to Question 23 (with or without non-Indigenous ancestry).
- Multiple Indigenous groups — individuals reporting more than one identity to Question 24 or individuals without Indigenous identity but reporting more than one Indigenous ancestry to Question 23 (with or without non-Indigenous ancestry).
- Person with Indian Status or member of a First Nation or Indian band only — individuals answering YES to Question 26 or to Question 27 but NO to Question 24.

In Atlantic Canada outside Nunatsiavut, in Quebec outside Nunavik, and in Yukon and the Northwest Territories outside Inuvialuit, all Indigenous groups (excluding Inuit) were combined during stratification. This is because the number of Métis is generally too small to produce estimates by age group in these regions.

Note that estimates were targeted for Inuit outside Inuit regions nationally; therefore, a domain of estimation was created for this group.

In total, 120 target domains of estimation and 78 supplementary domains of estimation were created, for a total of 198 domains of estimation. Supplementary domains of estimation included domains where there was no plan to produce estimates for the IPS but where a particular group should still be represented in the sample. Examples of supplementary domains of estimation included the non-Inuit population living in Inuit regions and the population aged 15 to 17 years.

1. Although the estimates were not specificially targeted for those aged 15 to 17 years, they were included in the sample.

For each target domain, the goal was to estimate a characteristic present for a certain proportion of individuals in the domain, with a given maximum coefficient of variation (CV). Depending on the domain sizes, CV values were set to 20%, 25% or 33%. The CV is a measure of the precision of the estimate, which is described in Section 7.2. The minimum proportion associated with the targeted CV is referred to as "min-p." For the 2022 IPS, the min-p is 10%, except for seven areas identified in Table 3 for which the min-p value is either 12% or 15%. The ability to achieve the targeted CVs for a given min-p value depended on the following factors:

- the size of the population
- the number of respondents available from the 2021 Census
- the expected response rate
- the expected number of false positives (individuals reporting Indigenous identity or ancestry in the census without Indigenous identity in the IPS, a very specific type of respondent outside the target population)
- the sample loss associated with the constraint of selecting a maximum of three individuals per household
- the loss associated with reducing overlap with other surveys.

The following table gives the targeted CVs and min-p for each target domain of estimation.

Table 3
Coefficients of variation in percentages and min-p values (in brackets when other than 10%) by domain of estimation for the 2022 Indigenous Peoples Survey

	Age group				
	1 to 5	6 to 14	18 to 24	25 to 54	55 and over
Indigenous group and Region	Coefficients of variation (min-p values)				
Inuit					
Nunatsiavut	33 (15%)	33	33	20	25
Nunavik	25	25	20	20	20
Inuvialuit	33	25	33	20	25
Nunavut	25	25	20	20	20
Canada, outside Nunangat	25	25	25	20	20
Status First Nations					
Ontario	25	25	25	25	25
Manitoba	25	25	25	25	25
Saskatchewan	25	25	25	25	25
Alberta	25	25	25	25	25
British Columbia	25	25	25	25	25
Non-Status First Nations					
Ontario	25	25	25	25	25
Manitoba	25	25	33 (12%)	25	33
Saskatchewan	25	25	33 (15%)	25	33
Alberta	25	25	25	25	25
British Columbia	25	25	25	25	25
Métis					
Ontario	25	25	25	25	25
Manitoba	25	25	25	25	25
Saskatchewan	25	25	25	25	25
Alberta	25	25	25	25	25
British Columbia	25	25	25	25	25
All Indigenous group except Inuit					
Atlantic outside Nunatsiavut	25	25	20	20	20
Quebec outside Nunavik	25	25	20	20	20
Yukon	33 (15%)	33	33 (15%)	25	33
Northwest Territories outside Inuvialuit	33 (12%)	25	33 (12%)	20	20

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

3.2.1.2 Domains of estimation of the Indigenous Peoples Survey–Nunavut Inuit Supplement

In the case of the IPS–NIS, the domains of estimation consisted of combinations of communities and education groups. More specifically, there were two education groups:

- those with a high school diploma or some postsecondary education (including trades certificate or apprenticeship, college, university)

- those without a high school diploma and any postsecondary education.

When possible, estimates were targeted for each education group within a community. Otherwise, community-level estimates were targeted. For the IPS–NIS, the domains of estimation were as follows:

- 12 communities with education-level estimates within the community
- 11 communities with community-level estimates only.

In total, 35 target domains of estimation and 3 supplementary domains of estimation were created, for a total of 38 domains of estimation. Supplementary domains of estimation are domains where there was no plan to produce estimates but where a particular group should still be represented in the sample. For the IPS–NIS, the supplementary domains correspond to three communities in Nunavut that were too small to produce community-level estimates. These communities were still sampled so that estimates for all of Nunavut can be produced.

For each target domain, the goal was to estimate a characteristic present for a minimum proportion of individuals in the domain, with a given CV of 25%. The minimum proportion targeted is referred to as "min-p." The ability to achieve the targeted CV for a given min-p value depended on such factors as the size of the population, the number of respondents available from the 2021 Census, the expected response rate, the expected number of individuals reporting being Inuit enrolled under the Nunavut Agreement on the IPS–NIS, the sample loss associated with the constraint of selecting a maximum of three individuals per household, and the loss associated with reducing overlap with other surveys.

In addition, to manage response burden in Nunavut, the min-p value was adjusted so that a maximum of 75% of a domain of interest was sampled. If, with a min-p of 10%, the required sample size represented more than 75% of people in the domain, then the minimum proportion was adjusted (to 12%, 15% or 20%) to manage the response burden in Nunavut. If it was not possible to produce estimates for a minimum proportion of 20% by education group, then community-level estimates were targeted (initially for a minimum proportion of 10%, which was then adjusted to 12%, 15% or 20% as required).

Table 4 lists the domains of estimation for the IPS–NIS as well as the min-p that was used.

Table 4
Coefficients of variation in percentages and targeted min-p values by domain of estimation for the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement

Domains of estimation: community and education group	Min-p in percentage (c.v. = 25%)
Arctic Bay	
High school diploma or some postsecondary education	20
No high school diploma or postsecondary education	20
All	12
Arviat	
High school diploma or some postsecondary education	10
No high school diploma or postsecondary education	10
All	10
Baker Lake	
High school diploma or some postsecondary education	15
No high school diploma or postsecondary education	15
All	10
Cambridge Bay	
High school diploma or some postsecondary education	10
No high school diploma or postsecondary education	10
All	10
Cape Dorset	
High school diploma or some postsecondary education	15
No high school diploma or postsecondary education	15
All	10
Chesterfield Inlet	
All	15
Clyde River	
All	10

Table 4
Coefficients of variation in percentages and targeted min-p values by domain of estimation for the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement

Domains of estimation: community and education group	Min-p in percentage (c.v. = 25%)
Coral Harbour	
High school diploma or some postsecondary education	20
No high school diploma or postsecondary education	20
All	12
Gjoa Haven	
All	10
Grise Fiord	
Supplementary Domain[1]	…
Hall Beach	
High school diploma or some postsecondary education	20
No high school diploma or postsecondary education	20
All	12
Igloolik	
High school diploma or some postsecondary education	12
No high school diploma or postsecondary education	12
All	10
Iqaluit	
High school diploma or some postsecondary education	10
No high school diploma or postsecondary education	10
All	10
Kimmirut	
All	12
Kivalliq, Unorganized	
Supplementary Domain[1]	…
Kugaaruk	
All	15
Kugluktuk	
High school diploma or some postsecondary education	20
No high school diploma or postsecondary education	20
All	12
Naujaat	
High school diploma or some postsecondary education	20
No high school diploma or postsecondary education	20
All	12
Pangnirtung	
All	12
Pond Inlet	
All	10
Qikiqtarjuaq	
All	15
Rankin Inlet	
High school diploma or some postsecondary education	10
No high school diploma or postsecondary education	10
All	10
Resolute	
Supplementary Domain[1]	…
Sanikiluaq	
All	10
Taloyoak	
All	10
Whale Cove	
All	15

... not applicable
1. This is a domain for which there was no plan to produce estimates (therefore no targeted min-p) but where a particular group should still be represented in the sample.
Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

3.2.2 Stratification

Stratification produces more precise estimates if the characteristic of interest is homogeneous within strata and heterogeneous between strata. In addition, the estimation weights associated with survey respondents should ideally be as close as possible within strata.

3.2.2.1 Stratification of the IPS

Two variables were used to create the 2022 IPS strata:

- the region associated with the census form type (2A-L or 2A-R)
- the type of Indigenous identification.

One of the assumptions used in the IPS sample allocation method was that the census weights would vary as little as possible within an IPS stratum. As previously mentioned, the region associated with the census form type (2A-L or 2A-R) created a difference in census weights. In regions where Form 2A-L was administered, a systematic sample of one in four households was drawn. In contrast, in regions where Form 2A-R was used, the form was administered to all households.

The other factor considered for IPS stratification, i.e., the type of Indigenous identification (Indigenous ancestry only or Indigenous identity), is unrelated to the variability of census weights. It was important to consider this variable as a stratification factor, given that the Indigenous identity (ID) group and the Indigenous ancestry only (AO) group also have very different characteristics. Incorporating this variable in the stratification will increase the homogeneity of the strata. Moreover, considering the type of Indigenous identification (AO or ID) will allow the allocation of the sample to reflect the probability of each unit being part of the target population (having Indigenous identity on the IPS). Nationally, based on past survey data, individuals having Indigenous identity on the census have a probability of about 85% of having Indigenous identity on the IPS, while individuals with Indigenous ancestry only on the census have a probability of about 26% of having Indigenous identity on the IPS.

Combining the form type (2A-L or 2A-R) and type of Indigenous identification (AO or ID) factors gave a maximum of four strata per domain of estimation:

- 2A-L, AO
- 2A-L, ID
- 2A-R, AO
- 2A-R, ID.

For certain domains such as Status First Nations, there were only two possible strata because included individuals had Indigenous identity by definition. Some strata may also have been empty, especially in the supplementary domains. For example, for a specific Inuit region, it was possible that no one classified as Indigenous ancestry only had Indigenous ancestry other than Inuit.

3.2.2.2 Stratification of the Indigenous Peoples Survey–Nunavut Inuit Supplement

To increase sampling efficiency for analytical purposes, each Nunavut community was stratified by education group:

- those with a high school diploma or some postsecondary education (including trades certificate or apprenticeship, college, university)
- those without a high school diploma and any postsecondary education.

For 12 communities, the stratification corresponds to the domain of interest. For the remaining 14 communities, this will ensure that both education groups are represented in the sample.

3.2.3 Sampling design and sample allocation

The 2022 IPS and IPS–NIS sample was selected from 2021 Census long-form questionnaire respondents. Thus, the IPS and IPS–NIS sampling design can be considered a two-phase design where the first phase corresponds to the census long-form questionnaire sample and the second phase corresponds to the IPS and IPS–NIS sample.

3.2.3.1 Indigenous Peoples Survey sampling design

Once the survey frame was constructed, it was stratified according to domain of estimation and then substratified by form type (2A-L versus 2A-R) and type of Indigenous identification (ID versus AO). A systematic random sample was then selected within each stratum, the frame having been presorted by household and person number. The purpose of this was to ensure proper distribution of the sample geographically within the strata as well as across as many households as possible. However, the effectiveness of this procedure was limited by the fact that members of the same household could be in different strata.

A method for optimal allocation between the substrata of a particular domain was used by taking into account the different types of sample size loss as described in Subsection 3.2.4, as well as the probability of each unit belonging to the target population in a given stratum. This allocation depended in part on the census weights. It should be noted that at the time of the allocation, these weights had not yet been calculated. Preliminary weights were therefore derived solely for the allocation. The definitive weights derived by the methodology team working on the census estimates were used during weighting (see Chapter 6).

Allocation was done in the survey's target domains first. For the non-targeted or supplementary domains (for example, non-Inuit in Inuit regions or individuals aged 15 to 17 years), the sample size in a given stratum was calculated using a sampling fraction equal to or less than the sampling fraction of the corresponding stratum in the corresponding target domain. These supplementary domains had to be retained to cover the entire target population but did not require a specific sample size since no estimates were to be derived for them. For non-targeted Indigenous groups, such as non-Inuit in Inuit regions, the same sampling fraction as for Inuit was used. For individuals aged 15 to 17 years, half the sampling fraction of all other age groups combined was used.

Veterans Affairs Canada was interested in an oversample of veterans in the 2022 IPS. They wanted to produce estimates of acceptable quality by gender at the Canada level, for all Indigenous groups combined among veterans. The required sample sizes of male and female veterans were calculated to obtain a CV of 25% for a min-p value of 7.5%. For male veterans, the number of individuals already selected from the main IPS sample was sufficient for the purpose. In the case of female veterans, on the other hand, there was a shortfall in the number of units required to achieve the desired level of quality. An additional systematic random sample of veteran women was therefore drawn across Canada, the frame having been previously sorted by household and person number.

3.2.3.2 Indigenous Peoples Survey–Nunavut Inuit Supplement sampling design

The main sample for the 2022 IPS was selected first. All Inuit aged 15 and over living in Nunavut and enrolled under the Nunavut Agreement who had been selected in the main sample were identified and included in the 2022 IPS–NIS sample. For each IPS–NIS domain, the oversample size required after taking into account the main IPS sample was calculated. The oversample was selected directly from Inuit having Indigenous identity, enrolled under the Nunavut Agreement, aged 15 and over and living in Nunavut.

The sample frame, excluding individuals already selected for the IPS, was designed and stratified by community and education group. A systematic random sample was then selected among the remaining units within each stratum, since the frame had previously been sorted by household and person number. The purpose of this was to ensure a distribution of the sample across as many households as possible. The effectiveness of this procedure was, however, limited by the fact that members of the same household can be in different strata.

Allocation was done in the survey's target domains first. This allocation depended on the size of the domain, the min-p value and different types of sample size loss.

For communities where education-level estimates were targeted, the allocation provided the required sample size for each education stratum. For communities where only community-level estimates were targeted, the allocation provided the required sampling fraction for the community, which was then applied in both education strata in the community.

For the supplementary or non-targeted domains, the overall sampling fraction for target domains in Nunavut was calculated and applied in both education strata within each supplementary domain.

3.2.4 Sample size adjustment

Probability of belonging to the target population

Based on previous iterations of the survey, it was expected that 85% of those who reported an Indigenous identity in the census would report an Indigenous identity in the IPS, and 26% of those who reported Indigenous ancestry only in the census would report an Indigenous identity in the IPS. These probabilities of belonging to the target population were also expected to be lower in 2022 than in 2017. Accordingly, the probabilities of belonging to the specified 2022 target population were adjusted downward by multiplying 2017 figures by 90% in the census Indigenous ancestry-only strata, and by 95% in the census Indigenous identity strata.

For the IPS–NIS, the probability of belonging to the target population was established at approximately 90%.

Response rate

Response rates from the 2017 Aboriginal Peoples Survey (APS) were used to allocate the sample for the 2022 IPS. As a preventive measure, these rates were reduced by multiplying them by an adjustment factor of 70% for the four Inuit regions, and 80% elsewhere in Canada. These parameters were set to target a response rate of approximately 60% overall for the 2022 IPS. For the 2022 IPS–NIS, a response rate of 60% (uniformly across all domains) was assumed for the calculation of sample sizes. Note that increasing the sample size based on an expected response rate is a protective measure to reduce sampling error. However, this does nothing to reduce non-response bias, a particular type of non-sampling error (see Section 7.3). Clearly, it is preferable to obtain a higher response rate on a smaller sample size than a lower response rate on a larger sample size.

Households with more than three individuals selected

Because the sample unit was the individual and not the household, the sample may have included several individuals from a single household. To limit response burden, one of the constraints imposed during collection was to select no more than three individuals per household. As a way to estimate this loss, preliminary samples were selected and the average number of individuals who would have been removed from the sample because of this constraint was calculated. The relative loss obtained was multiplied by 1.15 for the IPS and by 1.2 for the IPS–NIS for each combination of domains and strata, and was incorporated into the sample size adjustment factor.

Overlap with other surveys

The IPS and IPS–NIS data collection period overlapped with that of several other Statistics Canada surveys. It is generally understood that an individual contacted for one survey is less likely to agree to participate in that survey if they have just been interviewed for another survey.

Several of the overlapping surveys also cover the territories where the Indigenous population represents a high portion of the total population. Therefore, there was a higher possibility of overlap between the different surveys in the territories. Special attention was given to this matter to reduce overlap as much as possible to reduce the potential burden on respondents.

First, the overlap with the Canadian Community Health Survey (CCHS) and the Labour Force Survey (LFS) was of particular concern in all three territories. Thus, the identifiable overlap in households in the territories to be interviewed between April and September 2022 for the CCHS, and between April and June 2022 for the LFS, who then must complete the Canadian Income Survey (CIS) supplement, was eliminated. The households concerned were removed from the IPS and IPS–NIS survey frames, so that they would have no chance of being selected.

Finally, two other post-censal surveys, the 2022 Survey on the Official Language Minority Population (SOLMP) and the 2022 Canadian Survey on Disability (CSD), were collecting data at approximately the same time as the IPS and the IPS–NIS. To reduce the response burden for all of Canada, it was decided that an individual could be selected by no more than one of the three post-censal surveys (SOLMP, CSD and IPS and IPS–NIS). In addition, a maximum

of three individuals per household could be selected for these three surveys. To compensate for any loss due to overlap with the SOLMP and the CSD, the total sample size for the IPS and the IPS–NIS was increased.

3.3 Sample size and response rate

The initial combined IPS and IPS–NIS sample included 76,230 individuals across Canada. This included the additional sample of 411 female veterans and the additional sample of 6,380 Inuit enrolled under the Nunavut Agreement of the IPS–NIS. Of these individuals, 73,984 were sent to collection (see Section 6.2). Table 5 shows the allocation of the 73,984 units of the IPS sampled by geographical domain and the corresponding response rates.

Table 5
Sample size and response rate by geographical domain for the 2022 Indigenous Peoples Survey using two definitions

Geographical domain	Total	Eligible[1]	Eligible[2]	Respondents[3]	Respondents[4]	Response rates[5]	Response rates[6]
	number of units					percentage	
Nunatsiavut	1,041	1,014	1,028	359	373	35.4	36.3
Nunavik	2,159	2,052	2,085	1,291	1,324	62.9	63.5
Inuvialuit	1,732	1,680	1,697	937	954	55.8	56.2
Nunavut	8,112	7,898	7,950	4,363	4,415	55.2	55.5
Nunangat (total)	**13,044**	**1, 644**	**12,760**	**6,950**	**7,066**	**55.0**	**55.4**
Atlantic excluding Nunatsiavut	5,002	4,155	4,958	2,125	2,928	51.1	59.1
Quebec excluding Nunavik	6,312	3,748	6,282	1,951	4,485	52.1	71.4
Ontario	10,758	9,112	10,690	4,142	5,720	45.5	53.5
Manitoba	8,442	7,798	8,383	3,086	3,671	39.6	43.8
Saskatchewan	8,111	7,449	8,053	2,893	3,497	38.8	43.4
Alberta	9,594	8,454	9,538	3,704	4,788	43.8	50.2
British Columbia	9,377	8,437	9,333	3,880	4,776	46.0	51.2
Yukon	1,335	1,234	1,323	486	575	39.4	43.5
Northwest Territories excluding Inuvialuit	2,009	1,891	1,990	725	824	38.3	41.4
Rest of Canada (total)	**60,940**	**52,278**	**60,550**	**22,992**	**31,264**	**44.0**	**51.6**
Canada (total)	**73,984**	**64,922**	**73,310**	**29,942**	**38,330**	**46.1**	**52.3**

1. Sample size excluding out of scope on the IPS (non-Indigenous individuals are out of scope under this definition).
2. Sample size excluding out of scope on the IPS (non-Indigenous individuals are in scope under this definition).
3. Number of respondents excluding non-Indigenous individuals on the IPS.
4. Number of respondents including non-Indigenous individuals on the IPS.
5. Response rate excluding non-Indigenous individuals on the IPS.
6. Response rate including non-Indigenous individuals on the IPS.

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

The initial sample for the 2022 IPS–NIS consisted of 7,968 individuals, 710 of whom were not surveyed (see Section 6.2). Table 6 shows the allocation of the 7,258 units of the IPS–NIS sampled by community and the corresponding response rates.

The response rate is defined as the number of eligible respondents divided by the number of eligible units in the sample. Individuals living outside Canada or in an institution at the time of the survey and those under the age of 1 as of April 27, 2022, are examples of ineligible units for the IPS and the IPS–NIS.

For the IPS, two definitions of eligible units were used. In the first definition, individuals not reporting Indigenous identity on the IPS were deemed ineligible (in addition to the other types of ineligible units). In the second definition, these individuals were deemed eligible and were included as respondents. These individuals agreed to participate in the survey and completed all of the questions administered to them, that is, the screening questions determining whether or not they had an Indigenous identity.

For the IPS–NIS, two definitions of eligible units were also used. In the first definition, individuals who were not enrolled under the Nunavut Agreement on the IPS–NIS (including non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement) were deemed ineligible (in addition to the other types of ineligible units). In the second definition, these individuals were deemed eligible and were included as respondents.

These two definitions of eligible units and respondents therefore offer two ways to measure the response rate. The first response rate (46.1% for the IPS and 51.6% for the IPS–NIS) is generally the one used at the time of collection, and is associated with sampled units belonging to the target population. The second response rate (52.3% for

the IPS and 55.1% for the IPS–NIS) is the one used by the methodology team, and corresponds more closely to the response rate established on the basis of all sampled units (with some exclusions). Of course, in both cases, non-responding units cannot always be classified as being eligible or not eligible. The probability of reporting an Indigenous identity on the IPS differs considerably depending on whether respondents reported an Indigenous identity or Indigenous ancestry only on the census. As a result, the response rate used during collection is highly influenced by the allocation of sample units between these two groups, which is not the case for the response rate used by the methodology team. In Inuit regions, since most individuals have Indigenous identity, the two response rates presented in Table 4 are very similar, which is not the case in the other regions. As shown in Table 4, the maximum difference between the two rates is observed in Quebec, excluding Nunavik. In this region, the Indigenous ancestry-only population forms a large part of the total Indigenous population and, according to past APS surveys, the probabilities of having Indigenous identity on the survey are the lowest in the country regardless of whether the individual has Indigenous ancestry-only or Indigenous identity on the census.

Table 6
Sample size and response rate by community for the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement using two definitions

	Total	Eligible[1]	Eligible[2]	Respondents[3]	Respondents[4]	Response rates[5]	Response rates[6]
Community	number of units					percentage	
Arctic Bay	243	196	240	84	128	42.9	53.3
Arviat	589	549	585	396	432	72.1	73.8
Baker Lake	376	345	364	180	199	52.2	54.7
Cambridge Bay	526	483	511	283	311	58.6	60.9
Cape Dorset	350	322	342	98	118	30.4	34.5
Chesterfield Inlet	151	131	149	75	93	57.3	62.4
Clyde River	268	235	254	92	111	39.1	43.7
Coral Harbour	252	223	245	75	97	33.6	39.6
Gjoa Haven	274	235	268	187	220	79.6	82.1
Grise Fiord	48	46	48	27	29	58.7	60.4
Hall Beach	238	205	235	106	136	51.7	57.9
Igloolik	478	442	472	88	118	19.9	25.0
Iqaluit	627	588	613	326	351	55.4	57.3
Kimmirut	183	171	178	76	83	44.4	46.6
Kivalliq, Unorganized	5	3	5	1	3	33.3	60.0
Kugaaruk	177	142	175	65	98	45.8	56.0
Kugluktuk	258	231	253	169	191	73.2	75.5
Naujaat	251	228	250	126	148	55.3	59.2
Pangnirtung	231	209	224	86	101	41.1	45.1
Pond Inlet	277	249	265	169	185	67.9	69.8
Qikiqtarjuaq	161	151	156	61	66	40.4	42.3
Rankin Inlet	591	559	579	317	337	56.7	58.2
Resolute	55	50	53	42	45	84.0	84.9
Sanikiluaq	255	239	252	25	38	10.5	15.1
Taloyoak	241	221	238	142	159	64.3	66.8
Whale Cove	153	138	147	104	113	75.4	76.9
Total — Nunavut	**7,258**	**6,591**	**7,101**	**3,400**	**3,910**	**51.6**	**55.1**

1. Sample size excluding individuals out of scope of the IPS–NIS (according to this definition, non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement are out of scope of the survey).
2. Sample size excluding individuals out of scope of the IPS–NIS (according to this definition, non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement are in scope of the survey).
3. Number of respondents of the IPS–NIS excluding non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement.
4. Number of respondents of the IPS–NIS including non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement.
5. Response rates of the IPS–NIS excluding non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement.
6. Response rates of the IPS–NIS including non-Indigenous individuals, non-Inuit and Inuit not enrolled under the Nunavut Agreement.
Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

4 Data collection

4.1 Time frame

The 2022 IPS and IPS–NIS was conducted from May 11 to November 30, 2022, with in-person follow-up occurring from January 16 to March 31, 2023. As a postcensal survey, it followed the 2021 Census of Population, which was held in May 2021. The IPS–NIS was conducted as a supplement at the same time as the IPS.

4.2 Mode of collection

Six collection methods were used to collect data for the 2022 IPS and the IPS–NIS. The 2022 survey was the first cycle to introduce the use of respondent electronic questionnaire (rEQ), where respondents receive a secure access code to log in and complete the survey online. The survey continued the use of computer-assisted interview (CAI) methods used in previous iterations: computer-assisted telephone interview (CATI) and computer-assisted personal interview (CAPI). Additionally, Canada still had various restrictions in place related to the COVID-19 pandemic; therefore, Statistics Canada implemented CAPI Lite Plus collection (CLP; where interviewers visit selected individuals in person to schedule an appointment with them to later complete the questionnaire via CATI), and Knock, Talk and Call (KTC, similar to CLP but a CATI interview is scheduled immediately). Finally, a pilot collection method was conducted for approximately a month in Iqaluit, where selected individuals were invited to complete their questionnaire via rEQ or CAPI at a municipal building. In person follow up via CAPI was used for all Inuit Nunangat regions, except for Nunatsiavut so as to avoid overlap with the Qanuippitaa? National Inuit Health Survey. rEQ and CATI follow-up were the primary modes of collection for the remainder of the sample.

Respondents could choose to complete the questionnaire in English or French. The full 2022 IPS and IPS–NIS questionnaire was available on Statistics Canada's website in Inuktitut (South Baffin syllabics), Inuinnaqtun, and Labrador Inuktitut, for respondents to reference as required.

The time required to complete the survey varied from person to person. In some cases, the questionnaire took up to an hour or more to finish, but on average, the survey took about 47 minutes to complete.

Before collection started, as well as during, every effort was made by Statistics Canada's western regional office to hire interviewers who were fluent in Inuktitut. When in-person follow-up was implemented, local guides were hired whenever possible. Local guides played a crucial role in the success of the in-person follow-up collection in Inuit Nunangat, since they were able to use their language skills to explain the importance of the survey to respondents in the local dialect.

4.3 Supervision and quality control

All Statistics Canada interviewers were under the supervision of senior interviewers who were responsible for ensuring that interviewers were familiar with the concepts and procedures of the survey to which they were assigned. Senior interviewers were also responsible for periodically monitoring the interviewers.

Interviewers were trained on the survey content, the CAI application, and on how to proceed with the new methods CLP and KTC. In addition to online or classroom training, the interviewers completed a series of mock interviews to become familiar with the survey and its concepts and definitions. They were also required to complete cultural awareness and sensitivity training related specifically to Indigenous peoples.

4.4 Proxy interviews

Because the 2022 survey has children as part of the target population, many proxy interviews were conducted. For children aged 1 to 14, interviews were conducted by proxy with a parent, guardian or the person most knowledgeable about the child.

The questionnaire for individuals aged 15 and over was designed to be answered by the selected respondent. Proxy interviews were acceptable in some circumstances, such as when the selected respondent was not able to answer because of mental or physical health or a language barrier, or because the selected respondent was absent from home for the duration of the survey. Anyone over the age of 18 can act as a proxy for the selected respondent and answer the survey for them.

For individuals between the ages of 15 and 17, interviews were conducted directly with the youth. If the parent or guardian denied permission for interviewers to speak to their youth aged 15 to 17, then the data were collected through proxy from the parent or guardian.

4.5 Communications strategy

Prior to collection of the 2022 IPS and IPS–NIS, introductory letters (containing a Secure Access Code [SAC]) and a handout (outlining the purpose of the survey as well as emphasizing the importance of participating) were mailed to selected respondents. These were available in English, French, Inuktitut and Inuinnaqtun for those living in Nunavut, and in English and French elsewhere. During the collection period, efforts were made to raise awareness of the 2022 IPS and IPS–NIS and to encourage participation. Activities and products were prepared in advance, and messaging was adapted based on any ongoing needs and challenges that arose during the collection.

The promotional products included handouts (brochures), posters, electronic templates, radio placements and public service announcements (PSA), and a promotional video. In addition to this, a social media campaign for the IPS and IPS–NIS was also launched on the official Statistics Canada Facebook, X (formerly Twitter), Instagram, YouTube, and LinkedIn pages. The survey was also promoted on Statistics Canada's website.

A toolkit was made available on the IPS's Information for Survey Participants page, where these promotional materials, along with additional social media content, web banners and hashtags could be downloaded and shared by community supporters, governments and Inuit organizations.

The IPS and IPS–NIS social media content and toolkit were also amplified on the social media pages of survey partners and funders. National Indigenous organizations were also able to share content on their social media accounts for the intended audience of the IPS and IPS–NIS.

Statistics Canada's Indigenous Liaison Advisors promoted awareness of the survey in regions across the country. They did this through their distribution lists, via emails and newsletters, encouraging participation using the materials provided.

Paid advertising tactics were also used for the survey. The main media objective was to build awareness about the survey and its importance, encouraging those who have been selected to participate. The secondary objective was to generate traffic to the Information for Survey Participants page to encourage respondents to complete the survey online. This campaign was unique in that it was not directed at a general population audience, but rather at Indigenous adults living off reserve, with a focus on northern communities.

A mixed-media approach was used for the advertising strategy, complementing the approach used for promotion. This included traditional media (community/static TV and radio ads in northern communities due to digital connectivity issues). In addition, digital advertising was used including social media ads (Facebook and Instagram), web banners and search engine marketing (Google and Bing). This allowed for local contextual ad placements and a means of reaching the target population during their time online.

Where possible, the promotional and advertising products were made available in English, French, Inuktitut (South Baffin syllabics) and Inuinnaqtun.

The advertising campaign was split into two phases to effectively target participants at the beginning and end of collection.

A third phase supported the second round of collection in the North (Nunavut and Nunavik).

For IPS–NIS only: the Nunavut Inuit Labour Force Analysis (NILFA) Technical Working Group representatives from Nunavut Tunngavik Inc. (NTI), the Government of Nunavut (GN), and Pilimmaksaivik, with the support of Statistics Canada and Employment and Social Development Canada, prepared a PSA to circulate internally to their employees containing information about the IPS–NIS collection in Nunavut, as well as specific examples on how survey data can be used to inform policy related to Article 23 of the Nunavut Agreement. The NILFA Technical Working Group representatives from the GN, NTI and Pilimmaksaivik also participated in interviews with the media regarding survey collection in Nunavut.

4.6 Special issues in the field

Interviewers were instructed to make all reasonable attempts to obtain a completed interview with the selected member of the household. Those who refused at first to participate were re-contacted up to two more times to explain the importance of the survey and to encourage their participation. For cases where the timing of the interviewer's call or personal visit was inconvenient, an appointment was arranged to call back at a more convenient time. For cases where no one was home, numerous call-backs were made. All cases in Inuit Nunangat were first contacted by CLP, KTC or CATI before having their final contact attempts made via CAPI (with the exception of Nunatsiavut where respondents did not receive CAPI collection).

Special issues arose in relation to data collection for the survey, which were addressed with extra coordination in the field and corrective adjustments to survey methods. For instance, the collection of the 2022 IPS and IPS–NIS occurred during the same time period as several other surveys conducted by Statistics Canada, creating a potentially heavy response burden for individuals who might be selected for participation in more than one survey. Careful planning and adjustments to survey design were implemented to address and minimize this type of respondent burden. Please refer to Subsection 3.2.4 for more detail about the survey overlap.

4.7 Final response rates

Around 38,330 individuals took part in the 2022 IPS. Of these individuals, 3,910 were eligible for the IPS–NIS. Including individuals who reported being non-Indigenous, the final response rate of the IPS was 52.3%,while that of the IPS–NIS was 55.1%. Excluding the 8,388 non-Indigenous respondents, the total number of Indigenous respondents included in the 2022 IPS database was 29,942, including 3,400 respondents who also completed the IPS–NIS.

A detailed description of the sampling design used and response rates obtained is provided in Chapter 3 of this guide. Table 5 (IPS) and Table 6 (IPS–NIS) in Section 3.3 provide the final response rates and sample sizes for each of the geographic domains covered by the survey (e.g., provinces, Inuit regions and territories).

5 Data processing

5.1 Data capture

Responses to survey questions were captured in the electronic questionnaire (EQ), either directly by the respondent (rEQ) or by an interviewer.

Some editing of data was done directly at the time of the interviews via CATI or CAPI. Specifically, where a particular response appeared to be inconsistent with previous answers or outside of expected values, the EQ prompted the respondent or interviewer, through message screens on the computer, to confirm answers and, if needed, to modify the information.

5.2 Social survey processing steps

Data processing involves a series of steps to convert the electronic questionnaire responses from their initial raw format to a high-quality, user-friendly database involving a comprehensive set of variables for analysis. A series of data operations are executed to clean files of inadvertent errors, rigorously edit the data for consistency, code open-ended questions, create useful variables for data analysis, and finally to systematize and document the variables for ease of analytical usage.

The 2022 IPS and IPS–NIS used a set of social survey processing tools developed at Statistics Canada called the Social Survey Processing Environment (SSPE). The SSPE uses SAS software, custom applications and manual processes for performing the following systematic steps:

Processing steps:

- Receipt of raw data
- Combination of raw data from two collection periods
- Reformatting
- Clean up
- Recoding
- Coding
- Flows
- Edits and imputations
- Census linkage
- Derived variables
- Conversion ofspecs
- Creation of final processing file
- Creation of dissemination files
- Data certification
- Codebook (data dictionary)
- Public use microdata file (PUMF)
- Share files

5.3 Receipt of raw data to coding

Following the receipt of raw data from the 2022 IPS and IPS–NIS electronic questionnaire application, the data from the first collection period needed to be combined with that from the in-person follow-up. Once one combined file was created, a number of preliminary cleaning procedures were implemented at the individual record level. These included the removal of all personal identifier information from the files, such as names and addresses, as part of a rigorous set of ongoing mechanisms for protecting the confidentiality of respondents. Duplicate records were resolved at this stage, and standard coding classifications were applied. Also, part of clean-up procedures was the review of all respondent records to ensure each respondent was "in-scope" and had a sufficiently completed questionnaire. Note that the criteria to determine whether or not a respondent was in scope was applied before any edit or imputation was done. Specific criteria for determining who would be a final survey respondent and who would not be a final survey respondent are provided below.

5.3.1 Definition of a respondent

To be "in scope", respondents must have been at least 1 year of age as of April 27, 2022, and have met a minimum of one Indigenous identity criterion (see Section 2.2 for complete criteria).

To receive the IPS–NIS questions, "in scope" respondents must have been at least 15 years of age as of April 27, 2022, and indicated they are Inuit and enrolled under, or a beneficiary of, the Nunavut Agreement.

To have a "complete" questionnaire, respondents aged 1 and over must have provided valid responses (i.e., not left a question unanswered) to specific key questions in the area of health **and** either the areas of labour, activities **or** education.

Those who did not meet the above criteria were removed from the database. As per the rules above, all "partial" respondents who were in-scope according to the definition but who did not fulfill the content completion requirements, were among those removed from the final database. Please refer to Section 6.3 of this document for more information on partial respondents.

5.4 Variable recodes and multiple response questions

This stage of processing involved changes at the level of individual variables. Variables could be dropped, recoded, resized or left as is. Formatting changes were intended to facilitate processing as well as analysis of the data by end-users. One such change was the conversion of multiple-response questions ("Select all that apply" questions) to corresponding sets of single-response variables, which are easier to use. For each response category associated with the original question, a variable was created with YES/NO response values. An example is provided below.

Original multiple-response question:

FHA_Q10 - Where does this child live now?

Select all that apply if multiple children are living outside of the household or if there are multiple living arrangements for the child.

1. With their other birth parent
2. With an adoptive family
3. With grandparents or other relatives
4. In foster care or group home
5. Away at school
6. Child is living independently
7. Other living arrangement

   Please specify living arrangement - FHA_S10

Final variables in single-response YES/NO format:

FHA_Q10A - Where does this child live now?

- With their other birth parent

1. Yes
2. No
6. Valid skip
9. Not stated

FHA_Q10B - Where does this child live now?

- With an adoptive family

1. Yes
2. No
6. Valid skip
9. Not stated

FHA_Q10C - Where does this child live now?

- With grandparents or other relatives

1. Yes
2. No
6. Valid skip
9. Not stated

and so on until the final single response:

FHA_Q10G - Where does this child live now?

- Other living arrangement – please specify

1. Yes
2. No

6. Valid skip
9. Not stated

5.5 Flows: response paths, valid skips and question non-response

Another set of data processing procedures for the survey was the verification of questionnaire flows or skip patterns. All response paths and skip patterns built into the questionnaire were verified to ensure that the universe or "target population" for each question was accurately captured during processing. Special attention was paid to distinctions between valid skips and non-response. These concepts are explained below to help users better understand question universes as well as statistical outputs for IPS and IPS–NIS survey variables.

Table 7
Flows: response paths, valid skips and question non-response

Answer type	Description
Response	An answer directly relevant to the content of the question that can be categorized into pre-existing answer categories, including "Other-specify."
Valid skip	This indicates that the question was skipped because it did not apply to the respondent's situation, as determined by a) specific valid answers to an earlier question, b) no valid answer provided in an earlier question that required specific answers to avoid skipping the question, and c) other criteria such as the respondent's age, sex, or whether the survey was completed by a proxy for the respondent. In such cases, the respondent is not considered to be part of the target population or universe for that question.
Don't know	For the 2022 IPS, "Don't know" was provided as a valid category in a few selected questions, but in all other questions, there was no specific code for "Don't know" in the collection system and it is treated as non-response.
Not stated	This indicates that the question response is missing. For example, the respondent may not have known what response to give, or they refused to provide a response or simply missed the question. Any question could remain unanswered since it was possible to continue to the next question without providing a response to the current question.

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

Special codes have been designated to each of these types of responses to facilitate user recognition and data analysis. For instance, "valid skip" codes are set to "6" as the last digit, with any preceding digits set to "9" (for example, code would be "996" for a threedigit variable). All "Don't know" responses end in "7," with any preceding digits set to "9" (for example, "997"). "Not stated" values end in 9, with any preceding digits set to "9" also (for example, "999"). These reserve codes are different for each variable, depending on how many categories the variable has and the length of the variable.

5.6 Coding

5.6.1 "Other-specify" items

Data processing also includes the coding of "Other-specify" items, also referred to as "write-in responses." In the event that a respondent's answer could not be easily assigned to an existing category, many questions allowed the respondent or interviewer to enter a long-answer text response in the "Other-specify" category.

All questions with "Other-specify" categories were closely examined during processing. Based on a qualitative review of the types of text responses given, coding guidelines were developed for each question. Based on these coding guidelines, many of the long answers provided were recoded back into one of the pre-existing listed categories. Responses that were unique and different from existing categories were kept as "Other." For some questions, one or more new categories were created when there were sufficient numbers of responses to warrant them. In the case of questions where "Other-specify" responses constituted less than roughly 5% of overall responses to the question, coding to new categories was not done; responses were instead left in "Other."

Approximately 17,000 responses across 62 questionnaire items were recorded under "Other-specify" and reviewed for coding; 30 of these needed new categories created. These will be taken into account when refining the answer categories for future cycles of the survey.

5.6.2 Open-ended questions and standard classifications

A few questions on the survey were asked in a completely open-ended format. These included questions related to the respondent's occupation and industry of work, as well as their major field of postsecondary study, where applicable. These responses were coded using a combination of automated and interactive coding procedures. Standardized classification systems were used to code these responses. Appendix C provides details of these classifications.

A standardized classification was also used to code Indigenous languages that respondents spoke or understood as well as the first language learned in childhood. The questionnaire also includes language questions for children such as languages the child can understand or speak, languages in which the child can express needs and languages the child can understand when someone speaks to them. For language, a comprehensive drop-down menu of 60 Indigenous languages, English and French, as well as 1 non-verbal language, 3 sign languages and 1 other language were provided to choose from, but write-in responses were also captured as needed. When coding was final, 67 language categories were used to code the surveys language data.

5.7 Edit and imputation

After the coding stage of processing, a series of customized edits were performed on the data. These consisted of validity checks within and across variables to identify gaps, inconsistencies, extreme outliers and other problems in the data. To resolve the problematic data identified by the edits, corrections were performed based on logical edit rules. In some cases, corresponding data were taken from the respondent's answers to the census. This is referred to as imputation.

An example of a validity check within a single question is the question asking for the respondent's weight, which allowed any numerical value between one and three digits to be entered. To remove outlier responses that were suspected of being invalid, an edit was built to ensure that the reported weight was not equal to 0 and did not exceed 450, whether the respondent had indicated the weight in pounds or kilograms. An additional edit ensured that weight given in pounds was not less than 30 for respondents aged 15 and over.

Additionally, some consistency edits across questions were performed to avoid any inconsistencies. For example, the person most knowledgeable for a respondent aged less than 15 is asked whether the respondent had ever been removed or separated from their family by child welfare agencies. If the answer is "Yes," the next question asked what the longest period of time was that the respondent had been removed or separated from their family by child welfare agencies, for which a response could be given in number of years, number of months or number of weeks. A value of 15 years or more would be invalid since the respondent's age would be less than that, and the response would be blanked out by the edit.

Although all of these edits across topics were performed systematically using computer-programmed edits, there were some cases for which very complex combinations of information were reviewed and corrected manually.

For the 2022 IPS and IPS–NIS, a series of important imputations was conducted in relation to Indigenous identity classifications. These imputations were the following:

a) First, those with missing data for questions ID_Q15 on Status Indian or ID_Q20 on membership in a First Nation or Indian band were imputed values based on their responses to the equivalent questions in the 2021 Census.

b) Next, those with missing data for question ID_Q05 on Indigenous self-identification and Indigenous identity group would still have been asked questions ID_Q15 and ID_Q20, described above. However, because of the IPS respondent definition, these respondents would have had to identify as either a Status Indian in ID_Q15 or a member of a First Nation or Indian band in ID_Q20 to be considered an IPS respondent. If these respondents had self-identified as Indigenous persons on the census, then their Indigenous identity group(s) for question ID_Q05 were imputed from their identity group(s) on the census. Those who did not self-identify as an Indigenous person on the census were imputed as not having Indigenous identity in ID_Q05. These people are still considered to be IPS respondents because of their affirmative response in ID_Q15 or ID_Q20 and the IPS respondent definition. In the

derived variable for Indigenous identity, they are grouped as "Indigenous responses not included elsewhere."

c) Next, those who self-identified as Inuit in ID_Q05 but had missing data for question ID_Q10, which asked if they were enrolled under, or were a beneficiary of an Inuit land claims agreement, had a response to the question imputed from their data from the equivalent question in the 2021 Census. Additionally, for those who answered ID_Q10 in the affirmative but had missing data for ID_Q10A, which asked which Inuit land claims agreement the respondent was enrolled under or was a beneficiary of, a value for ID_Q10A was imputed based on their response to the equivalent question in the census.

d) Finally, respondents who self-identified as Métis in ID_Q05 but had missing data for question ID_Q25A, which asked if they were a registered member of a Métis organization or Settlement, had a response imputed from their response from the equivalent question in the 2021 Census. Likewise, if they had answered ID_Q25 in the affirmative but had missing data in ID_Q25A, which asks the respondent to specify the name of the Métis organization or Settlement, the name was imputed from the census data.

5.8 Derived variables and census linkage

To facilitate more in-depth analysis of the rich IPS and IPS–NIS dataset, approximately 270 derived variables were created by combining items on the questionnaire. Derived variables (DVs) were created across all major content domains. In addition, approximately 260 variables from the 2021 Census were linked to the final 2022 IPS and IPS–NIS analytical file.

Some simple DVs involved the collapsing of categories into broader categories. In other cases, two or more variables were combined to create a new or more complex variable that would be useful for data analysts. Some of the DVs were based on linked variables from the census, including multiple census geographies and Inuit regions. Indigenous ancestry was also taken from the census since it is not measured directly by the 2022 IPS. If a respondent refuses census linkage, then their data are suppressed for census and census-based variables.

For most DVs, there is a residual category labelled "Not stated" for when the responses to the DV source questions do not meet the conditions to place a respondent in any of the valid categories for the DV. In many, but not all cases, a respondent is included in the "Not stated" category if any part of the equation was not answered (that is, if any question used in the DV had been coded to "Don't know" or "Not stated").

Most DV names have a "D" in the first character position of the name. One exception is the geography DVs, since they reflect the corresponding census variable name. The other exceptions are the DVs indicating levels 1, 2 and 3 of the North American Industry Classification System (NAICS) Canada 2022, based on responses to the IPS industry questions, and levels 1, 2, 3, 4 and 5 of the National Occupation Classification (NOC) 2021, based on responses to the IPS occupation questions. For all linked census variables, the census variable name was preserved as much as possible on the IPS and IPS–NIS database. However, some census variables were required to be renamed since the survey variable names are restricted to eight characters, whereas some census variable names exceed this limit. In these cases, there is a note in the data dictionary to indicate the original census variable name that it was shortened from.

The 2022 IPS and IPS–NIS Data Dictionary identifies in detail which variables were derived and indicates the source variables from which the DVs were derived. A complete list of linked census variables and their accompanying notes are provided in the 2022 IPS and IPS–NIS Data Dictionary, which accompanies the survey analytical file.

5.9 Creation of final data files and data dictionary (codebook)

Four final data files are created in data processing:

- final processing file
- analytical file for use in research data centres (RDCs) and with the Real Time Remote Access (RTRA) tool

- Public Use Microdata File (PUMF)
- Inuit share files, as per data sharing agreement with the four Inuit regions.

The final processing file is an in-house file that includes a number of temporary variables used exclusively for processing purposes. The analytical file, PUMF and the Inuit share files are dissemination files that are processed further for release purposes. Dissemination files are scheduled for distribution at various points in time following the IPS release day, August 14, 2024.

The analytical file is distributed in RDCs across Canada but can be accessed only by researchers who fulfill certain requirements. Researchers associated with an academic institution, a government department or a non-profit organization can also use the RTRA tool to access the file. The analytical file is also used at Statistics Canada to produce data tables in response to client requests. The PUMF is constructed for wider public distribution. The Inuit share files are produced in accordance with data sharing agreements with the Inuit regions: Nunatsiavut, Nunavik, Nunavut and the Inuvialuit region. On all of these dissemination files, many steps are taken to ensure respondent confidentiality. For more detail on dissemination, see Chapter 9.

To transform the final, cleaned processing file to a final analytical file for researchers, a number of steps were taken. First, several measures were taken for the enhanced protection of respondent confidentiality. Next, person-level weights were added to the file. Finally, all temporary variables or variables used exclusively for processing purposes were removed from the final processing file.

Accompanying the 2022 IPS and IPS–NIS analytical file is the record layout, SAS, SPSS and Stata syntax to load the file, and metadata in the form of a data dictionary that describes each variable and provides weighted and unweighted frequency counts.

To ensure the non-disclosure of confidential information, the level of detail of the PUMF is not as exhaustive as that of the analytical files kept by Statistics Canada. Actions are taken to protect against the recognition of respondents with potentially identifiable combinations of characteristics. These protective actions include the restriction of geographies included in the file, adjustments to survey weights, review of overlaps with other PUMFs being published, exclusion of variables, grouping of categories for some variables, capping of some extreme numerical values, as well as identification of unique records at risk and rare occurrences.

6 Weighting

In a sample survey, each person selected represents not only themselves but also others who were not sampled. Consequently, a weight is associated with each person selected to indicate the number of people they represent. This weight must be used for all estimates. For example, in a simple random sample of 2% of the population, each individual in the sample represents 50 individuals. The initial weight is then adjusted to take account of factors such as non-response and gaps between sample characteristics and known totals for the target population (calibration). The weighting process for the 2022 IPS and IPS–NIS consists of five steps.

In the 2017 cycle, the order of the questionnaire varied depending on whether the person was selected for the Aboriginal Peoples Survey (APS) or the APS–NIS. This difference led to the creation of two set of weights, one for the main APS sample and one for the supplement (APS–NIS). For the 2022 cycle, the questionnaire no longer has this attribute; the order of questions is the same for all participants, and the block of questions related to the Nunavut supplement is asked at the end of the questionnaire if the respondent meets the conditions of eligibility for it. As a result, only one set of weights is required and will take into account the 2022 IPS and IPS–NIS.

In this section, the term “census” will be used to refer to the census long-form questionnaire.

6.1 Initial weights

In many surveys, the initial weight corresponds to the inverse of the probability of selecting a unit. In simple cases, this selection probability is calculated in each stratum as the number of individuals selected from the sample in

each stratum, divided by the total number of individuals available in the stratum. This was the case for the main sample of the 2022 IPS.

However, calculating these selection probabilities is much more complex for the two oversamples, i.e., that of the female veterans and that of the IPS–NIS. Firstly, stratification was different for individuals from the main IPS sample than for those from the two oversamples. Then, the selection probabilities of the two oversamples depended on the individuals previously selected. The main sample for the 2022 IPS was drawn first, then the selected individuals were removed from the frame before selecting the oversample of female veterans. Finally, all those already selected were removed from the survey frame to allow selection of the additional sample for the 2022 IPS–NIS. Thus, the selection probability of a given unit for both oversamples depended on which units had already been selected, adding a degree of complexity to the selection probability calculation.

To properly calculate the selection probability of a given unit in both oversamples, a simulation study was conducted. The sampling process for the 2022 IPS and IPS–NIS was repeated 25,000 times. The selection probabilities in one of the two oversamples were calculated as the number of repetitions in which a unit was selected for the oversample in question divided by the total number of repetitions in the simulation (i.e., 25,000). For the 2022 IPS main sample, the selection probability was calculated according to the theoretical formula.

Finally, the inverse of the selection probability of all selected individuals could then be multiplied by the census long-form questionnaire weight to obtain the initial weight of units in the combined sample. The census weight used is the census survey weight corrected for non-response and for overlap with other surveys (see Subsection 3.2.4).

6.2 Adjustment for units not sent to collection

For various reasons, some sampled units of the 2022 IPS and IPS–NIS were not sent to collection, including:

1. units for which three members of the same household had already been selected for the IPS or IPS–NIS
2. units selected from a household where more than three individuals had been selected when considering other post-censal surveys, i.e., the 2022 Canadian Survey on Disability (CSD) and the 2022 Survey on the Official Language Minority Population (SOLMP)
3. units from the oversample of female veterans selected for the CSD
4. units with contact information deemed insufficient (based on name, gender and age)
5. units selected for the 2022 IPS or IPS–NIS and for the 2022 SOLMP
6. duplicates identified by overlap of names, dates of birth, and addresses following sample selection.

In the first four cases, an initial ratio adjustment was carried out. The weights of the units removed in the first four cases were set to zero, and the weights of the remaining units were increased proportionally (ratio adjustment) within each combination of the IPS, IPS–NIS and veteran domains (eligible for the oversample of veteran women or not). A few small domains were combined.

In the fifth case, a second adjustment was made for units selected for both the 2022 IPS or IPS–NIS and the 2022 SOLMP. The weight of these units was set to zero after being redistributed among those in the sample who were eligible for the SOLMP, and according to IPS domains where possible.

In the sixth case, units identified as duplicates were considered out of scope for the survey and their weight was set to zero during a third adjustment.

In total, 2,246 units were not sent to collection for the 2022 IPS and IPS–NIS.

6.3 Adjustment for non-response and partial respondents

Two adjustments were made for two types of non-response: selected individuals with whom no contact was made (non-contact: 10,713 individuals) and individuals who were contacted but did not (or could not) provide information about themselves (non-response with contact: 21,624 individuals). The second type of non-response

is mainly associated with refusals or disguised refusals. An example of a disguised refusal could be a person who was contacted several times and continually postponed the interview. Two adjustments were made since the characteristics of the people who could not be reached are often different from those who refuse to respond when contacted.

An adjustment was also made for the 2,642 partial respondents. Partial respondents are individuals who reported Indigenous identity in the IPS but who did not provide sufficient information to meet the definition of respondent, as set out in Chapter 5. These partial respondents are considered as non-response cases, since they are not included in the final analytical file.

Weights were first adjusted for "non-contact" cases, then for "non-response with contact" cases and finally for "partial respondent" cases. In the rest of this document, the term "non-response" will be used for all three types of non-response. Note that all adjustments for non-response were made separately for children (aged 1 to 14) and adults (aged 15 and over).

Each non-response adjustment was done in three steps. In the first step, a logistic regression model was used to predict the response probability (probability of getting a response) for each unit (both responding and non-responding units) from a series of explanatory variables. Explanatory variables are "person" or "household" characteristics from the 2021 Census (e.g., the selected person's Indigenous group, the number of people in the selected person's household) for the selected person or the person most knowledgeable in the case of children.

In the second step, respondents and non-respondents with similar predicted response probabilities were grouped into adjustment classes using classification analysis. A simulation was carried out to determine approximately the optimal number of classes and the minimum number of respondents per class. The response rate was derived for each class based on the number of respondents and non-respondents in the class. The calculated response rate was then weighted using the weights from the previous adjustment step.

In the third step, the weights of the responding units within each class were adjusted using the inverse of the weighted response rate in that class. The weights of the non-responding units were set to zero.

It should be noted that at this step, the adjustment of partial respondents is done by excluding all units considered to be out of scope. However, for the adjustments of non-contacts and non-respondents with contact, all units considered to be out of scope were classified as respondents. In fact, all the information required to determine that they were out of scope for the survey was obtained from these individuals. The weights of these units were set to zero in the second calibration step (see Section 6.4), and these units were removed from the analytical file. Retaining them until that step makes it possible to internally produce weighted estimates of different groups of units outside the target population. This will be very useful in estimating certain parameters in a future survey cycle, for example.

6.4 Calibration and post-stratification

The purpose of the calibration or post-stratification step is to minimize the sampling variability of estimates derived from the IPS and IPS–NIS. Calibration is a weighting step that ensures the consistency of estimates of certain IPS and IPS–NIS totals with the corresponding totals estimated from the census long-form questionnaire. These totals are referred to as "control totals" in the rest of this document.

A calibration consists of modifying the weights obtained in the previous step as little as possible so that the weighted estimates are equal to the estimated census control totals. Statistics Canada's Generalized Estimation System (G-Est) was used to perform calibration under constraints related to control totals.

Post-stratification (also known as post-stratification adjustment) is a special case of calibration that ensures the sum of the adjusted weights of the respondent units matches the census estimates, according to different groups called post-strata. For the IPS and IPS–NIS, two separate calibrations (or post-stratifications) were carried out.

6.4.1 First calibration

For the first calibration phase, the Response Database (RDB) variables used to select the sample (and not the variables measured by the IPS) were used. Census control totals were defined on the basis of variables in the census dissemination database. This calibration is used to achieve the control totals of the census dissemination database, which may differ from those of the RDB. In fact, an individual of Indigenous identity or origin in the RDB may no longer be on the dissemination database. In addition, the Indigenous group of individuals may have been changed from one database to the other. The first calibration phase therefore corrects this problem, among others. This ensures that the sample does not underrepresent or overrepresent certain combinations of census characteristics (for example, Indigenous groups, regions, and age groups).

The census estimates to which the IPS weights were adjusted in this calibration phase correspond exactly to the IPS coverage, i.e., the population of Indigenous identity or ancestry aged 1 and over on April 27, 2022, excluding individuals living on reserves and in certain First Nations communities in the territories.

6.4.1.1 Outside Nunavut

Outside the territory of Nunavut, calibration corresponds to a post-stratification, that aimed to adjust the weights, by post stratum, to the population of Indigenous identity or ancestry estimated by the census, using the identity and ancestry variables taken from the RDB (see Subsection 3.1.3) at the time of sample selection (and not the variables measured in the IPS, which were instead the subject of the second calibration phase).

For this first post-stratification, post-strata were defined on the basis of certain combinations of region, Indigenous type, Indigenous group and age group:

- geographic region (Nunatsiavut, Nunavik, Inuvialuit, Atlantic other, Quebec other, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia, Yukon, Northwest Territories other)
- Indigenous type (Indigenous ancestry only [AO] versus Indigenous identity [ID])
- Indigenous group (First Nations [Status and non-Status Indians], Métis, Inuit, Other: multiple and members of a First Nation or Indian band only)
- age groups (1 to 5, 6 to 14, 15 to 17, 18 to 24, 25 to 54, 55 or over).

The weights were adjusted according to the ratio of the weighted census estimate to the weighted IPS sample estimate for each post-stratum.

6.4.1.2 In Nunavut

In Nunavut, an initial calibration was carried out to ensure that both IPS control totals and those of the IPS–NIS were met. First, the IPS control totals are combinations of these characteristics:

- Indigenous type (AO versus ID)
- Indigenous group (Inuit or Other)
- age groups (1 to 5, 6 to 14, 15 to 17, 18 to 24, 25 to 54, 55 or over).

Then, the IPS–NIS control totals are combinations of these characteristics:

- Inuit enrolled under the Nunavut Agreement
- community
- age groups (under 15, 15 and over).

6.4.2 Second calibration

The second calibration phase ensured that the Indigenous identity population counts of the IPS and IPS–NIS were the same as those obtained from the census. More specifically, the census estimates to which the IPS weights were adjusted in this second calibration phase correspond to the in-scope population, i.e., individuals with an Indigenous identity aged 1 and over on April 27, 2022, excluding those living on reserves and in certain First Nations communities in the territories.

Unlike the first calibration phase, the second was based on responses to questions from the IPS and IPS–NIS. For various reasons, these could have differed from those obtained in the census (Section 8.1). This was a calibration to ensure that the IPS and IPS–NIS population counts were the same as those obtained from the census. Indigenous origin is not measured by the IPS and IPS–NIS, and unless a unit reporting only Indigenous ancestry in the census declares an Indigenous identity in the IPS and IPS–NIS, these units are considered out of scope and are not included in the second calibration. After this step, only respondents with Indigenous identity according to the IPS and IPS–NIS had positive weights. The weights of out-of-scope units had been set to zero.

6.4.2.1 Outside Inuit regions

Outside the four Inuit regions (Nunatsiavut, Nunavik, Inuvialuit and Nunavut), the second calibration corresponds to a post-stratification aimed at adjusting, by means of post-strata, the weights to the Indigenous identity population estimated by the census, using identity variables that were taken from the variables measured in the IPS. For this second post-stratification, the post-strata were defined on the basis of combinations of the following characteristics:

- geographic region (Atlantic other, Quebec other, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia, Yukon, Northwest Territories, other);
- Indigenous identity group (First Nations [Status and non-Status Indian], Métis, Inuit, Other: multiple and members of a First Nation or Indian band only);
- Age groups (1 to 5, 6 to 14, 15 to 17, 18 to 24, 25 to 54, 55 and over).

The weights were adjusted according to the ratio of the weighted census estimate to the weighted IPS sample estimate for each post-stratum.

6.4.2.2 In Nunatsiavut, Nunavik and Inuvialuit

In Nunatsiavut, Nunavik and Inuvialuit, a second calibration was carried out to respect two sets of control totals: those of the IPS and those related to the territorial agreements of these three Inuit regions. The IPS control totals are combinations of the following characteristics:

- geographic region (Nunatsiavut, Nunavik and Inuvialuit)
- Indigenous identity group (Inuit or Other)
- age groups (1 to 5, 6 to 14, 15 to 17, 18 to 24, 25 to 54, 55+).

The second set of control totals are combinations of the following characteristics:

- Inuit under the Inuvialuit Final Agreement
- Inuit under the James Bay and Northern Quebec Agreement (Nunavik)
- Inuit under the Labrador Inuit Land Claims Agreement (Nunatsiavut).

6.4.2.3 In Nunavut

In Nunavut, a second calibration was carried out to ensure consistency with two sets of control totals: those of the IPS and those of the IPS–NIS. The IPS control totals were obtained by cross-tabulating the Indigenous identity group with the age group (1 to 5, 6 to 14, 15 to 17, 18 to 24, 25 to 54, 55 and over), while the IPS–NIS control totals were the same as those used for the first calibration and are described in Subsection 6.4.1.2.

6.5 Adjustment for extreme weights — Sigma-gap method

When previous weight adjustments were carried out, it was possible for some weights to have very large values compared with others. This could have created problems during estimation if, in addition to having very large weights, these units also had very different characteristics to units with smaller weights. The sigma-gap method was used to detect extreme weights within various groups closely related to the survey's domains of estimation (see Subsection 3.2.1). Bernier and Nobrega (1998) describe one application of the sigma-gap method.[2] The

2. Bernier, J., and K. Nobrega. 1998. Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method. Annual Meeting of the Statistical Society of Canada, June 1998.

sigma-gap method was applied separately for the Inuit regions (Nunatsiavut, Nunavik, Inuvialuit and Nunavut) and the rest of Canada, since the weight distribution was very different for these two groups.

The sigma-gap method used here was intended to detect extreme weights by calculating the difference between two successive weights after being sorted in descending order. This difference was compared with *n*standard deviation* of the weights within each group corresponding to the different control totals used in the second calibration phase described in Subsection 6.4.2. If the difference exceeded *n*standard deviation* of the weights, the largest of the two weights was declared extreme. Once a weight was declared extreme, all other larger weights in its group were automatically considered to be extreme as well. Those weights were then truncated to the value of the first weight that was not extreme. The mass of the truncated weights was then redistributed within the groups using a ratio adjustment. After examining several scenarios, a value of 2 was finally chosen for *n* for the Inuit regions and 1.5 for the rest of Canada. These specific *n* values enabled the identification of an acceptable number of extreme weights (16 extreme weights for the Inuit regions and 85 extreme weights elsewhere in Canada). In fact, most weights that should have been intuitively identified as extreme following a manual review were identified by the sigma-gap method used with these two *n* values. As well, a small number of weights per group were identified as extreme, thereby leaving the vast majority of the weights calculated in the previous steps unchanged.

7 Data quality

7.1 Overview of the data quality evaluation

The objective of the 2022 IPS and IPS–NIS is to produce quality estimates in the areas of employment, health, education and other core indicators for off-reserve First Nations people, Métis, and Inuit aged 1 and over (IPS) and for Inuit aged 15 and over living in Nunavut and enrolled under the Nunavut Agreement (IPS–NIS).

Sections 7.2 and 7.3 below explain the two types of errors that occur in surveys: sampling and non-sampling errors. Each type of error is evaluated in the context of the 2022 IPS and IPS–NIS. Sampling error is the difference between a statistical result obtained from the sample and that which would have been obtained under a census of the whole population conducted under similar conditions. Thus, sampling error can be described as differences arising from sample-to-sample variability. Non-sampling errors are all other errors that are caused by factors other than sampling. Non-sampling errors can occur at any stage of the survey process, and include non-response errors as well as errors introduced during data collection or computer processing. Errors may appear in the responses of respondents who are trying to recall facts from the past, for example, or they may occur when the responses are obtained through a proxy. A response may have been incorrectly captured because of interviewer fatigue or a computer malfunctioning. An error may have been made during programming, when the data was being processed or tabulated. These are all examples of non-sampling errors.

This section describes the various measures adopted to prevent errors from occurring wherever possible, and to adjust for any errors found throughout the various stages of the IPS and IPS–NIS.

7.2 Sampling errors

The estimates that can be derived from the 2022 IPS and IPS–NIS are based on a sample of individuals. Somewhat different estimates might be obtained if a complete census had been carried out using the same questionnaires, interviewers, supervisors, processing methods, etc. as those actually used.

To provide estimates of sampling error for statistics produced by the IPS and the IPS–NIS, a particular type of bootstrap method (the bootstrap being itself a specific resampling method), was used. Several bootstrap methods exist in the literature, but none of them were appropriate for the IPS and IPS–NIS sampling design. The following particularities of the sampling design made the estimation of sampling errors difficult:

- A two-phase sampling design in which households (or dwellings) were selected in the first phase and individuals in the second phase was used (Subsection 3.2.3).

- The sampling fraction of the first phase sample (census long-form questionnaire sample) was non-negligible (about one fourth in the 2A-L questionnaires), and the sampling fraction of the IPS and IPS–NIS was generally quite high in most strata.
- The IPS and IPS–NIS strata (Subsection 3.2.2) were not nested within the census strata (collection units).
- The method used had to be flexible enough to produce standard statistics such as proportions, totals, means and ratios but also more sophisticated statistics, including percentiles, logistic regression coefficients, etc.

For the 2006 Aboriginal Peoples Survey (APS), a general bootstrap method for two-phase sampling was developed by Langlet, Beaumont, and Lavallée (2008).[3] The underlying idea of the general bootstrap method is that the initial bootstrap weights can be seen as the product of the initial sampling weights and a random adjustment factor. In the case of a two-phase sample, the variance can be split into two components, each associated with one sampling phase. The general two-phase bootstrap methodology produced a random adjustment factor for each phase of sampling. In this case, the initial bootstrap weight of a given unit is the product of its initial sampling weight and the two random adjustment factors. Once initial bootstrap weights have been calculated, all weight adjustments applied to the initial sampling weights are applied to the initial bootstrap weights to obtain the final bootstrap weights. Therefore, the final bootstrap weights capture the variance associated not only with the particular sampling design but also with all weight adjustments applied to the full sample to derive the final weights.

For the 2022 IPS and IPS–NIS, the method developed for the 2006 APS has been adapted to account for non-response to the 2021 Census long-form questionnaire. In terms of calculating variance, the 2021 Census sampling design is considered a two-phase design: the first phase involves the initial selection of approximately one in four households, while the second is the census respondent sample. Although the 2021 Census had a very high collection response rate (97.4% for the long-form questionnaire), the second phase accounts for non-response when calculating variance for the census. That said, to use the general two-phase method for the IPS and IPS–NIS, the two census phases have been combined into a single phase, with the 2022 IPS and IPS–NIS sample constituting the second phase.

There is a major advantage to having two sets of random adjustment factors. The first set of adjustment factors can be used for estimates based on the first phase only, i.e., estimates based on the census long-form questionnaire sample. These estimates are used when the weights are adjusted to the census totals during calibration (Section 6.4). This method produces census totals that vary for each bootstrap sample. This reflects the fact that census totals are based on a sample and not on known fixed totals.

For the IPS and IPS–NIS, 1,000 sets of bootstrap weights were generated using the general bootstrap method. The method used is slightly biased in that it slightly overestimates the variance. However, the amount of overestimation was found to be negligible for the IPS and IPS–NIS. The method can also lead to negative bootstrap weights. To remedy this problem, bootstrap weights were transformed, but this reduced their variability. Consequently, the variance calculated with these transformed bootstrap weights has to be multiplied by a factor that is a function of a certain parameter. The parameter's value is chosen as the smallest integer that makes all bootstrap weights positive. For the IPS and IPS–NIS, the value of this parameter is 4. The variances calculated on the transformed bootstrap weights must therefore be multiplied by $4^2 = 16$. Similarly, the standard error (square root of the variance) must be multiplied by 4. However, most software applications that produce sampling error estimates from bootstrap weights have an option to specify this adjustment factor, so that the correct variance estimate is obtained without the extra step of multiplying by the constant.

It is extremely important to use the appropriate multiplicative factor for any estimate of sampling error such as variance, standard error or coefficient of variation (CV). Omission of this factor would lead to erroneous results and conclusions. This factor is often specified as the "Fay adjustment factor" in software, which produces sampling error estimates from bootstrap weights.

3. Langlet, É., Beaumont, J.-F., and Lavallée, P. (2008). Bootstrap Methods for Two-Phase Sampling Applicable to Postcensal Surveys. Paper presented at Statistics Canada's Advisory Committee on Statistical Methods, May 2008, Ottawa.

Note that if C is the variance multiplicative factor, some software uses the parameter k instead, where $k = 1 - 1/\sqrt{C}$. In the present situation, since $C = 16$, then $k = 0.75$. For examples of procedures using the Fay adjustment, see the 2022 *Indigenous Peoples Survey: User's Guide to the Main Analytical File*.

7.3 Non-sampling errors

Besides sampling errors, non-sampling errors can occur at almost every stage of a survey. Respondents may misunderstand the questions and answer them inaccurately, responses may be inadvertently entered incorrectly during data capture, and errors may be during data processing. These are all examples of non-sampling errors.

Over a large number of observations, randomly occurring errors will have little effect on estimates drawn from the survey. However, errors occurring systematically may contribute to biases in the survey estimates. Thus, much time and effort was devoted to reducing non-sampling errors in the survey. At the content development stage, extensive activities were undertaken to develop questions and response categories that would be well understood by respondents. The questionnaire was tested thoroughly during several rounds of qualitative testing. In addition, many initiatives were completed in the field to encourage participation and reduce the number of non-response cases. Quality assurance measures were applied during the data collection, coding and processing stages to verify and correct errors in the data. Adjustments to the survey weights were made to take into account the different characteristics of non-respondents and respondents, and thus minimize any potential bias that might have resulted.

In the following paragraphs, the various types of non-sampling errors are presented, along with the measures used to reduce or correct these errors in the 2022 IPS and IPS–NIS.

7.3.1 Non-response errors

Non-response errors result from a failure to collect complete information on all units in the selected sample. Non-response produces errors in the survey estimates in two ways. First, non-respondents often have different characteristics from respondents, which can result in biased survey estimates if non-response is not corrected properly. The larger the non-response rate, the larger the risk of potential bias will be. Second, having a larger number of non-respondents reduces the effective size of the sample. As a result, the precision of the estimates decreases (the sampling error on the estimates will increase). This second aspect can be overcome by selecting a larger sample size initially. However, this will not reduce the potential bias in the estimates.

There are many types of non-response. One form of non-response is partial non-response, where the respondent does not respond to one or more questions, but has completed a significant portion of the overall questionnaire. Partial non-response can be caused by difficulty in understanding a particular question. Generally, the extent of partial non-response was relatively small in the IPS and IPS–NIS. Extensive qualitative reviews and questionnaire testing were done before the survey, which reduces the extent of partial non-response. A response to key pre-defined questions was required before a case was classified as "respondent" as described in Subsection 5.3.1. There were, however, a significant number of cases where a large proportion of responses to key questions were missing. These cases were eliminated from the database of respondents (did not satisfy the definition of respondent) and were treated during weighting as special cases of total non-response (see Section 6.4). Finally, total non-response occurs when the person selected to participate in the survey could not be contacted or did not participate once contacted. Weights of respondents were inflated to compensate for those who did not respond as described in Section 6.3.

To mitigate the number of non-response cases, many initiatives were also undertaken before and during data collection (as mentioned in Chapter 4). Statistics Canada's website included a web page dedicated to the IPS and IPS–NIS, which contained a series of questions and answers for respondents, as well as general information about the survey. At the outset of collection, each selected respondent received an introductory letter providing an overview of the survey and a full-colour brochure explaining the importance of participating. During data collection, messages containing graphics and information were regularly posted on Statistics Canada's various social media accounts (X, Facebook, Instagram, etc.) to promote the IPS and IPS–NIS.

In addition, in-depth interviewer training was conducted by experienced Statistics Canada staff. Detailed interviewer manuals were provided as a reference, in addition to the training. As well, all interviewers were under the direction of senior interviewers, who oversaw activities in the regional offices. Interviewers made significant efforts to reach non-respondents through call-backs and follow-ups. Whenever possible, more than one phone number was provided for each selected respondent to maximize the chance of reaching the person during the collection period.

During the collection period, several reminder letters, emails and text messages were sent out to encourage respondents to respond. Interviewers also gave respondents their personal secure access code if they preferred to respond online rather than on the phone.

Finally, face-to-face interviews were also conducted to encourage non-respondents to take part in the survey in certain Inuit regions. Since many respondents may have moved between the time of census data collection and that of the IPS and IPS–NIS, as many local guides as possible were hired in communities in Nunavut to help interviewers locate respondents who had moved.

Detailed tables of the final response rates obtained for the 2022 IPS and IPS–NIS can be found in Section 3.3 of this guide (Tables 5 and 6).

7.3.2 Coverage errors

As mentioned in Section 3.1, the target population for the 2022 IPS was Canada's Indigenous identity population aged 1 and over and living in private dwellings. People living on Indian reserves and settlements and in certain First Nations communities in Yukon and the Northwest Territories were excluded. The target population for the IPS–NIS included individuals aged 15 and over, of Inuit identity, living in Nunavut, and enrolled under the Nunavut Agreement.

First, the population sampled or covered by the 2022 IPS corresponded to respondents to the 2021 Census long-form questionnaire who reported either Indigenous ancestry or identity (see Subsection 3.1.1). Then, the population sampled or covered by the 2022 IPS–NIS corresponded to respondents to the 2021 Census long-form questionnaire who reported Inuit identity and were enrolled under the Nunavut Agreement. In both cases, the same age and geographic restrictions applied to the sampled or covered population as to the target population.

Coverage errors occur when there are differences between the target population and the sampled population (population covered by the survey frame). Over-coverage is generally not an issue since out-of-scope units in the sample are typically identified during data collection and can be estimated for the entire survey frame. However, under-coverage can exist. Because the IPS and IPS–NIS sample was selected from individuals who had participated in the 2021 Census, individuals who did not participate in the census could not be sampled for the IPS and the IPS–NIS. If this group of individuals is significantly different than the those who participated in the census with respect to the characteristics measured in the IPS and IPS–NIS, it could introduce a bias. This bias is assumed to be relatively small given the very high response rate obtained in the census (97.4% response rate for the long-form questionnaire) and the adjustments made to the initial census sampling weights.

For more information about coverage errors on the census, please see the *2021 Census Coverage Technical Report*, to be released on the Statistics Canada's website in 2024. For more information on the quality of census data, please refer to Chapter 9 of the Guide to the Census of Population, 2021.

7.3.3 Measurement errors

Measurement errors occur when a provided response differs from the real value. Such errors may be attributable to the respondents, the interviewer, the questionnaire, the collection method, or the data processing system. Significant efforts were made to develop questions for the 2022 IPS and IPS–NIS that would be understandable, relevant and culturally sensitive.

Following the release of the 2017 APS and APS–NIS data, an extensive content review was conducted. This review drew on the expertise of a diverse group of researchers and subject matter experts from within and outside of

Statistics Canada. An analysis was conducted on which questions worked the best and which were most effective in producing valid indicators. This process also extended into an extensive search for relevant questions from other standardized survey questions at Statistics Canada.

Questions selected for potential inclusion in the 2022 IPS and IPS–NIS questionnaire then underwent several rounds of qualitative testing using one-on-one virtual interviews with respondents in various communities across Canada, including Iqaluit and Yellowknife. Testing was done among First Nations people, Métis and Inuit. Qualitative testing of the survey questionnaire was carried out by Statistics Canada's Questionnaire Design Resource Centre (QDRC). Based on the results of these tests, adjustments were made to question wording and sequences to reduce measurement error.

Many other measures were also taken to specifically reduce measurement error, including the use of skilled interviewers, extensive training of interviewers in survey procedures and content, and observation and monitoring of interviewers to detect problems of questionnaire design or misunderstanding of instructions.In addition, incoming data were evaluated in real time to detect problems of questionnaire design or misunderstanding of instructions. In fact, changes were made based on errors found in the data received, which was a new monitoring method for the 2022 iteration of the IPS and IPS–NIS.

7.3.4 Processing errors

Processing errors may occur at various stages of the survey process, including programming of the electronic questionnaire, data capture by the interviewer or the respondent, coding, and data editing. Quality control procedures were applied to every stage of the 2022 IPS and IPS–NIS data processing to minimize this type of error. Data collection was carried out using an electronic questionnaire, either administered by an interviewer or completed by self-reporting over the Internet. A number of verifications were built into the system to warn the respondent or the interviewer in the event of inconsistencies or unusual values, making it possible to correct them immediately (see Section 5.7).

At the data processing stage, a detailed set of procedures and edit rules were used to identify and correct any inconsistencies between the responses provided. For every step of data cleaning, a set of thorough, systematized procedures was developed to assess the quality of every variable in the file and correct every error found. A snapshot of the output files was taken at each step and verification was done by comparing files at the current and previous step. The programming of all edit rules were tested before being applied to the data. Examples of data processing verifications included:

- the review of all question flows, including very complex sequences, to ensure skip values were accurately assigned and distinguished from different types of missing values
- an in-depth qualitative review of open-ended and "Other — Specify" responses to ensure accurate and rigorous coding
- completion of coding to standardized classifications
- the review of all derived variables against their component variables to ensure accurate programming of derivation logic, including very complex derivations.

See the data processing section of this guide for further details (Chapter 5).

8 Differences between the Indigenous Peoples Survey and other data sources

Because of a number of differences in methodology between the 2022 IPS and IPS–NIS, previous Aboriginal Peoples Survey (APS) and APS–NIS cycles, and other Statistics Canada surveys, caution should be exercised when comparing data from various sources. The following sections provide information on the factors affecting data comparability, providing data users with important information on the factors to consider when carrying out analyses using IPS and IPS–NIS data.

8.1 Differences between the 2022 Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement and the 2021 Census

The 2022 IPS and IPS–NIS sample is selected from respondents who provided specific responses in the census. More detailed information on how census responses were used to determine the population of interest for the IPS and the IPS–NIS is provided in Chapter 3 (Survey design).

The census and the IPS are both rich sources of information on Indigenous peoples that complement each other. The IPS takes concepts that are touched on in the census and asks questions that dig deeper to provide more detailed information. For instance, questions 33 to 37 on the census provide information on the topic of education (which includes high school completion or equivalent; completion of registered apprenticeship or other trades certificate; completion of college or CEGEP; university certificate, diploma or degree completion; major field of study completed; and province, territory or country of completion).). Adding information from the IPS provides an opportunity to learn more about secondary (high) school education as well as primary school education: for example, whether children aged 5 to 18 attended an early childhood development or preschool program, whether they have ever attended school on reserve, or whether they were taught an Indigenous language at school.

The IPS and IPS–NIS also cover entire topics or themes that are not included in the census. For example, the IPS can provide detailed information on the education and health of Indigenous peoples. The IPS–NIS provides an opportunity to obtain more detailed information from Inuit enrolled under the Nunavut Agreement about their experiences with public sector employment, their interest in such employment, their plans to apply, their interest in training, and the language of work.

While both the 2021 Census and the 2022 IPS cover the “Indigenous identity population” in their sampling designs, the 2022 IPS did not cover the “ancestry-only population” described in Chapter 3. (Census respondents who reported Indigenous ancestry only were included in the IPS sample because they had a non-negligible probability of reporting Indigenous identity in the IPS. These respondents were retained in the IPS dataset only if they had reported an Indigenous identity in this survey.)

Population counts from the 2022 IPS or IPS–NIS for certain subpopulations may differ from those obtained from the census, even if the population universe for the census is restricted to that of the IPS or IPS–NIS. The second calibration described in Section 6.4 ensured that the number of individuals with Indigenous identity was the same in the census and the IPS, but this applied only to certain combinations of Indigenous group, region and age group. Similarly, the number of people aged 15 and over reporting Inuit identity and enrolled under the Nunavut Agreement will be identical whether calculated from the census or the IPS–NIS, but only for certain combinations of Nunavut regions. However, population counts may differ for other subpopulations that were not controlled for during calibration.

On the other hand, for the same individual, reported Indigenous identity and enrolment under the Nunavut Agreement may, in some cases, differ between the IPS and IPS–NIS and the census. There are many reasons why responses to these surveys may differ.

8.1.1 Collection methods and impact of proxy reporting

In most regions, 2021 Census data were collected by self-enumeration. Questionnaires were filled in over the Internet or returned by mail. For Indian reserves and remote areas, including all Inuit communities, census interviewers were used. Often one member of the household completed the census form for all members of the household. This is called proxy reporting.

As described in Section 4.2, IPS and IPS–NIS data were collected, in most cases, by self-reporting via an electronic questionnaire from the selected person. Because the person contacted for the IPS and IPS–NIS may not be the same person who filled in the census questionnaire, there may be some differences in responses to similar questions.

8.1.2 Different questionnaires

The Indigenous self-reporting questions on the 2022 IPS and IPS–NIS and the 2021 Census were identical. However, while the IPS and IPS–NIS did not have a question on Indigenous ancestry in 2022, Question 23 of the census asked about ethnic or cultural origins, and was immediately followed by Question 24 on self-reported Indigenous identity. Several Indigenous origins are included as examples in the census ethnic or cultural origins question. This might affect the responses provided to Question 24 on self-reported Indigenous identity.

8.1.3 Different contexts

The census questionnaire asks a limited number of general questions to the Canadian population as a whole, while the IPS and IPS–NIS questionnaire is designed specifically for Indigenous people. Consequently, given the more specific context of the IPS and IPS–NIS, the concept of identity may have been better understood. Hence, it is possible that individuals who reported being Indigenous on the census may have reported being non-Indigenous in the IPS and IPS–NIS. Conversely, individuals who had reported Indigenous ancestry only in the census may have later reported an Indigenous identity in the IPS and IPS–NIS.

8.1.4 Effect of time

The concept of Indigenous identity may not be a static characteristic in time. Events affecting the rights of certain Indigenous groups or changes in the general population's perception of Indigenous people may affect the way in which Indigenous identity questions are answered. Individuals who see themselves as having only Indigenous ancestry at one point in time may later self-report as being Indigenous. Alternatively, the same individual may identify as a First Nations person at one point in time and Métis at another.

8.2 Differences between the 2017 Aboriginal Peoples Survey and Aboriginal Peoples Survey–Nunavut Inuit Supplement and the 2022 Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement

Several significant changes in content were made between the 2017 Aboriginal Peoples Survey (APS) and APS–NIS and the 2022 IPS and IPS-NIS. However, a similar methodology was used in 2017 and 2022, particularly for creating the survey frame, drawing the sample and weighting.

8.2.1 Methodological differences

The 2017 APS target population included individuals aged 15 and over only, while the 2022 target population also included children aged 1 to 14. Estimates should be comparable between the 2017 APS and 2022 IPS for individuals aged 15 and over in the target population.

As for the IPS–NIS, one major difference between the 2017 and 2022 cycles concerns weighting, and more specifically, the calibration stage. In 2017, the weighting of the APS and the APS–NIS was carried out separately. Thus, the APS–NIS had different control totals to those of the APS. In fact, the control totals for the 2017 APS–NIS were defined by combining the Nunavut community and education group (people with a high school diploma or some postsecondary education and those without). For the 2022 cycle, a single weighting (and therefore a single calibration) was carried out for the IPS and IPS–NIS together. Control totals therefore had to be consistent for the IPS and the IPS–NIS. Thus, for the 2022 IPS–NIS, the education group was not included in defining the control totals (see Section 6.4 for more details on the 2022 calibration). This could have an impact on comparisons of estimates from the two cycles, particularly in the case of education-related estimates.

In terms of methodology, the biggest change between the 2017 and 2022 cycles concerns quality indicators. Before 2022, the coefficient of variation (CV) was used to report the quality of estimates in terms of their sampling error. For the 2022 IPS and IPS–NIS, the 95% confidence interval (CI) is used for this purpose.

A CI is associated with a confidence level, which is generally set at 95%. A 95% CI is an interval constructed around the estimate so that, if the process that generated the sample were repeated many times, the value of the

parameter of interest in the population would be contained in 95% of these intervals. The CIs presented with the 2022 IPS and IPS–NIS estimates are produced using appropriate methods so that their coverage is close to the nominal rate, i.e., 95%. For further details on the methods used to construct confidence intervals for the 2022 IPS and IPS–NIS, please refer to Appendix E.

There are many advantages in using CIs to report quality. They are appropriate for all types of estimates. However, CVs are not suitable for estimated proportions and estimates of differences, among others. As well, CIs provide an objective measure of the sampling error in a form that is easy to interpret: see – see Appendix E for more details. Lastly, CIs convey that there is uncertainty around the estimate, and that there is a range of other possible values that could have been obtained if different samples had been selected.

Readers may wish to consult the *Indigenous Peoples Survey and Indigenous Peoples Survey – Nunavut Inuit Supplement User Guide to the Main Analytical File* for related information on the use of CIs to report the precision of survey estimates, and quality-related release rules.

8.2.2 Changes to Indigenous identity questions

Inuit enrolled under an Inuit land claim agreement

This question was added into the Identification module of the 2017 Aboriginal Peoples Survey and remained in the 2022 IPS. Only those who self-identified as Inuit are asked this question. This question was moved before the Registered Indian and band membership questions to improve data quality.

Métis registered as members of a Métis organization or Settlement

This question was added into the Identification module of the 2022 IPS. Only those who self-identified as Métis are asked this question.

9 Data dissemination

9.1 An overview of 2022 Indigenous Peoples Survey dissemination

Data for the 2022 IPS and IPS–NIS were released on August 14, 2024. There was an announcement in *The Daily*, Statistics Canada's official release bulletin, with highlights of the analytical article, and a separate announcement in *The Daily* for the IPS–NIS. A set of data tables designed to accompany the analytical article has been made available on the Statistics Canada website. An IPS and IPS–NIS analytical file is also available through Research Data Centres (RDCs) and the Real Time Remote Access (RTRA) tool. Additional custom data tables are also available on a cost-recovery basis. In addition, Statistics Canada's network of Indigenous Liaison Advisors across Canada will be distributing IPS-focused newsletters and responding to IPS data needs in their regions.

9.2 Information products and services

An analytical report on the theme of families and children, with a focus on children, has been made available with the survey release on August 14, 2024, providing data at the national and, where possible, regional levels for First Nations people living off reserve, Métis and Inuit.

An analytical file for the 2022 IPS and IPS–NIS has been made available in Statistics Canada's RDCs across the country. To access the file, researchers must undergo a research and ethics committee review for approval. Their use of the data must be conducted according to Statistics Canada policies, guidelines and standards (for instance, only aggregate statistical estimates that conform to the confidentiality provisions of the *Statistics Act* may be released outside of Statistics Canada). For more information on the Research Data Centres Program, please refer to the Frequently asked questions.

Accompanying the analytical file is the record layout, SAS, SPSS and Stata syntax to load the file, as well as metadata in the form of a data dictionary that describes each variable and provides weighted and unweighted

frequency counts. In addition, a detailed user's guide provides researchers with guidelines they need for conducting statistical analysis of 2022 IPS and IPS–NIS data.

A public use microdata file (PUMF) will also be disseminated (expected to be available in 2025) to allow for wider and direct use of the data by researchers. To provide extra assurance with respect to the non-disclosure of confidential information, the level of detail of the PUMF is not as fine as that of the analytical file kept by Statistics Canada.

Statistics Canada plans to have four data sharing agreements for the 2022 IPS and IPS–NIS with each of the Inuit regions of Canada: Nunatsiavut (Northern coastal Labrador), Nunavik (Northern Quebec), the territory of Nunavut and the Inuvialuit region of the Northwest Territories. These data sharing agreements are an excellent method to increase data use and make the data more accessible. Only those respondents who are Inuit and have agreed to share their information with the respective region will be part of the data sharing file.

Statistics Canada offers a service called the Indigenous Statistical Capacity Building Initiative (ISCBI). The ISCBI is aimed at providing expanded services and building statistical capacity, grounded in the needs of First Nations people, Métis and Inuit. The Initiative offers statistical training, engagement and outreach, enabling Indigenous organizations to develop and sustain statistical capacity.

Another means of access to the data file is the Real Time Remote Access (RTRA) tool at Statistics Canada. This is a paid subscription service for clients associated with an academic institution, a government department or a non-profit organization. RTRA is an online remote access tool allowing users to run SAS software, in real-time, against microdata sets located in a central and secure location. Researchers using the RTRA system do not gain direct access to the microdata and cannot view the content of the microdata file. The RTRA provides around-the-clock access to survey results to users with a secure username and password from any computer with Internet access. Confidentiality rules and reliability guidelines are applied to all requests in an automated way by the RTRA system, eliminating the need for manual intervention and allowing for rapid access to results. An RTRA agent can be reached at: statcan.dad-rtra-dad-adtr.statcan@statcan.gc.ca, or please visit Real Time Remote Access for more information on how to obtain a user account.

In addition to these data products and services, clients can request custom data tables from Statistics Canada's Centre for Indigenous Statistics and Partnerships Client Services by email at statcan.cispdatasupport-cspasoutienauxdonnees.statcan@statcan.gc.ca. All custom requests are screened for confidentiality and aggregate data are rounded before being released to clients.

9.3 Analytical products

A detailed analytical report on the theme of families and children, with a focus on children, has been made available with the 2022 IPS and IPS–NIS release on August 14, 2024. This report uses data from the nationally representative 2022 IPS and IPS–NIS to broadly examine self-reported demographic information, living arrangements of Indigenous children, socioeconomic conditions, food insecurity and basic needs, Indigenous languages and culture, child care, and health of children.

9.4 Reference products

Information about the survey is available on Statistics Canada's website. Statistics Canada provides an Integrated Metadata Base (IMDB) online for all surveys that it conducts, including the 2022 IPS and IPS–NIS. The purpose of the IMDB is to provide information that will help the public interpret Statistics Canada's published data. The information (also known as metadata) is provided to ensure an understanding of the basic concepts that define the data, including variables and classifications, the underlying statistical methods and surveys, and key aspects of the data quality. Direct access to the 2022 IPS and IPS–NIS questionnaire is also provided.

As well, this guide is provided online for a detailed discussion of survey content; sampling design; data collection and processing; weighting of the data; data quality; differences between the 2022 IPS and IPS–NIS and other data sources, including the 2021 Census of Population; and dissemination products for the IPS and IPS–NIS.

For researchers using the analytical file in Statistics Canada's RDCs, the *Indigenous Peoples Survey and Indigenous Peoples Survey – Nunavut Inuit Supplement, 2022: User Guide to the Main Analytical File* is available, which details the concepts and methods of the survey along with detailed step-by-step instructions for using the data file. It also provides detailed guidelines for tabulation and statistical analysis, how to apply the necessary weights to the data, information of software packages available and guidelines for the release of data, such as rounding rules. The process of estimating the reliability of estimates, both quantitative and qualitative, is covered in detail. The RDC user's guide describes the structure of the data file in detail, including all core variables, derived variables and linkages to the census. A detailed data dictionary provides information for all variables available.

Finally, a separate *Public Use Microdata File User Guide* will be created for users of the public use microdata file for the survey. This will cover variables available, a data dictionary (codebook), the process of estimation, use of weights and guidelines for tabulation, statistical analysis and dissemination of data.

9.5 Disclosure control

Statistics Canada is prohibited by law from releasing any information it collects which 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.

Appendix A1 2022 Indigenous Peoples Survey questions summarized

This appendix lists all themes, modules and a summary of associated questions of the 2022 Indigenous Peoples Survey (IPS).

Please note that the details for each question in the survey are not listed. Please refer to the 2022 IPS and IPS–NIS questionnaire for a comprehensive list of all questions.

Themes

- Demographics
- Family stability
- Belonging, language and culture
- Economic well-being
- Health
- Support to families
- Education
- Trauma and discrimination
- Labour

Summary of Indigenous Peoples Survey questions by theme

Demographics

- Indigenous identity: First Nations (North American Indian), Métis or Inuk (Inuit)
- Inuit enrolled under an Inuit land claims agreement
- Status Indian (Registered or Treaty) as defined by the *Indian Act* of Canada
- Member of a First Nation or Indian band
- Registered member of a Métis organization or Settlement
- Household composition
    - ▶ Number of household members at the address
    - ▶ Relationship between household members and respondent
- Sex and gender of the selected child, adult and proxy
- Sexual orientation of the selected adult
    - ▶ Profile of the person most knowledgeable

Family stability

- Family history
    - ▶ Children under the age of 18 living away from home
    - ▶ Reasons for their absence
    - ▶ Ties to extended family
    - ▶ Moved from home since birth or adoption
- Mobility – place of residence
    - ▶ Moved in last five years

Belonging, language and culture

- Sense of belonging
    - ▶ To your geographical surroundings
    - ▶ About your Indigenous identity

- Harvesting, handcrafting and cultural activities – Adults
    - Activities - harvesting, handcrafting and cultural activities
    - Income from harvesting, handcrafting and art
- Activities – Child
    - Activities
    - Indigenous cultural activities
- Language - adult and child

Economic well-being

- Food security
    - Worried money would not last for more food or to eat balanced meals
    - Ate less than you felt you should
    - Lost weight because you didn't have enough money for food
    - Hungry but did not eat for a whole day
- Basic needs
    - Income enough to meet household's needs
    - Cover an unexpected expense of $500
- Housing
    - Dwelling in need of repairs
    - Number of rooms in the dwelling
    - Subsidized dwelling
- Neighbourhood safety
- Perceived safety from crime when walking alone in your area at night or after dark
- Sources of personal income
    - Income sources
- Total personal income
    - Best estimate of total personal income
- Employment income
    - Total wages and salaries from your job or business
- Information technology
    - Internet service and use

Health

- General health and general mental health
- Pregnancy and childbirth
    - Currently pregnant
    - Number of children given birth to
    - Number of children fathered
    - Age at first child's birth
- Maternal health
    - Birthing experiences
- Height and weight
    - Height

    - Weight
    - Considers themselves overweight/underweight/just about right
- Birth weight
- Chronic conditions - Adult
    - Long-term conditions
        - Asthma
        - Arthritis, excluding fibromyalgia
        - High blood pressure
        - Chronic bronchitis, emphysema, or chronic obstructive pulmonary disease (COPD)
        - Diabetes
        - Heart disease
        - Intestinal or stomach ulcers
        - Bowel disorder
        - Mood disorder
        - Anxiety disorder
        - Cancer
- Chronic conditions - Child
    - Long-term conditions
        - Asthma
        - Respiratory allergies
        - Lactose intolerance or trouble digesting milk
        - Food or digestive allergies
        - Ear infection or ear problems
        - Visual impairment
        - Speech or language difficulties
        - Emotional, psychological or nervous difficulties
        - Learning disability
        - Attention deficit disorder or attention deficit hyperactivity disorder, also known as ADD or ADHD
- Injuries
    - Injured in the past 12 months
    - Type of injury
- Eating habits adult and child
- Smoking
    - Cigarettes smoking habits
    - Smoking inside home
    - Use of e-cigarette or vaping device
- Alcohol consumption
    - Drinking habits
- Drug use
    - Cannabis use (past 12 months)
    - Street drugs (past 12 months)
    - Prescription drugs for recreational purposes (past 12 months)

- Disease screening
    - Mammogram
    - Pap smear
    - Screened for colorectal cancer
- Consultations about mental health – Adults, children aged 6 to 14
- Suicide
    - Ever seriously contemplated suicide in the past 12 months
    - Ever seriously attempted suicide in the past 12 months
- General Health
    - Regular medical doctor
    - Routine (regular) vaccinations or immunizations (child)
    - Child been diagnosed with or tested positive for COVID-19
- Oral health – Adult and child
- Physical activity
- Disability
    - Difficulties you may have doing certain activities (long-term conditions)
    - How often these difficulties limit your daily activities
- COVID-19 pandemic health impacts

Support to families

- Child care
    - Types and costs of child care currently in use (or reasons for not using child care)
    - Satisfaction with child care currently in use, especially in terms of cultural appropriateness
    - Barriers to preferred child care
    - Hours per week the child is in each type of child care
- Access to services
    - Jordan's Principle
    - Inuit Child First Initiative

Education

- Education – Adult
    - Highest grade of elementary or high school
    - High school diploma or its equivalent
    - Currently attending high school or a high school equivalency program
    - Saving money for child's postsecondary education
- Postsecondary education
- Education – Children and youth
    - Current level attending, or highest level completed
    - Number of schools attended, reason for changing schools
    - Description of the school in terms of Indigenous culture (e.g., Indigenous languages taught)
    - Child's school performance
    - Family's involvement with the school and child's studies
    - Leaving school and returning, reasons for doing each

Trauma and discrimination

- Family separation
  - Personal or familial residential school attendance
  - Experience as a child being under the custody of a child welfare agency or being placed in foster care
- Missing and murdered Indigenous women, girls and Two-Spirit people
  - Presence of missing or murdered Indigenous women, girls, and Two-Spirit people in one's life as well as relationship to them.
- Discrimination experienced because of Indigenous identity
- Victimization
  - Determine whether the respondent has been the victim of physical or sexual assault in the past 12 months
  - What is the relationship of the perpetrator to the respondent

Labour

- Labour market activities
  - Worked at a job or business last week
  - Job or business from which you were absent last week
- Labour force status
  - Finding work
  - Wanting a job with more or less than 30 hours per week
- Class of worker
  - Employee or self-employed
- Industry and occupation
- Usual hours of work
- Child care impact on employment
- Impacts of the COVID-19 pandemic
- Disability screening questions

Appendix A2 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement questions summarized

This appendix lists a summary of the questions of the 2022 Indigenous Peoples Survey–Nunavut Inuit Supplement (IPS–NIS). Please note that the details for each question in the survey are not listed. Please refer to the 2022 IPS and IPS–NIS questionnaire for a comprehensive list of all questions.

Summary of Indigenous Peoples Survey–Nunavut Inuit Supplement questions

- Trust in public institutions
- Availability for government employment
    - ▶ Worked in a government or hamlet job in Nunavut last week
    - ▶ Planning to stay within the government for the next two years
- Past government employment
- Relief workers and substitute teachers
- Permanent work
- Looking for work
- Labour market attachment
    - ▶ Wanted a job during the last week
- Labour mobility
    - ▶ Willing to move to improve your job or career opportunities
- Interest in government employment
    - ▶ Interest in working for the government or hamlet in Nunavut in the next two years
- Previous applications to government employment
    - ▶ Ever applied for a governmental or hamlet job in Nunavut
    - ▶ Experienced difficulties or challenges when applying for a government or hamlet job in Nunavut
- Plans to apply for government employment in the next 12 months
    - ▶ Plan to apply for a government or hamlet job in Nunavut in the next 12 months
- Postsecondary preparedness and distance education
    - ▶ Taken part in any workshops, programs or activities specifically designed to help Indigenous students adjust to or succeed in postsecondary education
    - ▶ Taken any postsecondary courses through the Internet or some other form of distance education
- Skills and training
    - ▶ Taken any courses, workshops, seminars or training to develop your job skills
- Interest in training for government employment
    - ▶ Interest in skills training in a specific area to help get a government or hamlet job
- Pre-employment training
    - ▶ Type of support that would be helpful to pursue this type of training
- Plans for further education
    - ▶ Plans on taking formal schooling counted towards a certificate, diploma or degree
    - ▶ Plans on taking the Nunavut Sivuniksavut program
    - ▶ Experiencing any difficulties or challenges that prevent pursuing further formal schooling
    - ▶ Difficulties or challenges

- Skill-relevant experiences
    - Experience:
        - Leadership and project management
        - Using numerical skills
        - Reading and writing
        - Researching or analyzing information
        - Instructing or teaching
        - Interpreting or translating
        - Going out on the land
- Language fluency for work
- Retirement income

Appendix B Extra response categories created for "Other-specify" questions

The table below summarizes the extra categories added to certain IPS survey questions during the coding of data from questions involving an "Other-specify" category.

Table B.1
The 2022 Indigenous Peoples Survey: Extra response categories created for "Other-specify" questions

Theme/Indicator	Questions	Extra categories
Demographic		
Relationship between household members and respondent	HC_Q15D - Other unrelated - Specify the relationship	Friend
		Dating partner
Person most knowledgeable - Currently working, going to school or doing something else	PMK_Q20 - Specify other activity	On financial assistance
		Unemployed
Person most knowledgeable, spouse or partner - Currently working, going to school or doing something else	PMK_Q35 - Specify other activity	Unemployed
Belonging, language and culture		
Anyone helping the child to understand First Nations, Métis or Inuit culture and history	ACT_Q095 - Someone else - Specify the person	Foster family
		Community organizations or programs
		Personal friend or family friend
		Other relatives
Health		
Type of cancer you were diagnosed with	CC1_Q95 - Specify other type of cancer	Leukemia
		Lymphoma
		Thyroid
		Kidney
		Bladder
		Stomach
Other long-term conditions - Child	CC2_Q70 - Specify the other conditions	Developmental disability or disorder
		Fetal alcohol spectrum disorder)
		Heart conditions
		Eczema or psoriasis
Types of help the child received for problems with emotional, behavioural, mental or spiritual health	CMH2_Q20 - Specify other	Support from teachers, school staff
		Medical professional
Reasons of not having a regular medical doctor	GH2_Q10 - Specify the other reason	Moved recently, have not yet found one
		On waiting list or looking for new doctor
		Does not need regular doctor
Reasons for not getting care	GH2_Q20 - Specify other reasons	Dissatisfied with quality of available care
		Waiting for or never received response to health care request
Reasons child has not received their routine (regular) vaccinations or immunizations	GH2_Q35 - Specify other reasons	Parents opted not to have them given to child
		Medical reasons to avoid vaccines
Child ever fed any of the following in their bottle	EH3_Q25 - Specify other	Juice (not specified if 100% fruit juice)

Table B.1
The 2022 Indigenous Peoples Survey: Extra response categories created for "Other-specify" questions

Theme/Indicator	Questions	Extra categories
Support to families		
Main reason why the child is not receiving regular child care	CHC_Q15 - Specify the other main reason	Would like child care, do not currently need
Relative provides child care for the child	CHC_Q35 - Other - Specify the relative	Aunt or uncle
Main barriers to using this preferred type of child care for the child	CHC_Q80 - Specify the other reason	Not available in the community
Reason you did not apply for any services under Jordan's Principle	SER_Q15 - Specify the other reason	Not aware or did not know how to apply
		Waiting for Status under *Indian Act*
Types of products, services or supports applied under Jordan's Principle	SER_Q20 - Specify the other types of products, services or supports	Social supports
		Household assistance
		Transportation and accommodation
		Food and clothing
Reason for not applying for Inuit Child First Initiative services	SER_Q40 - Specify other reasons	Not aware or did not know how to apply
Types of products, services or supports applied under the Inuit Child First Initiative	SER_Q45 - Specify the other types of products, services or supports	Household assistance
		Food and clothing
Education		
Reasons why the child is not receiving the additional help or tutoring required	EDC_Q095 - Specify the other reason	Child doesn't want any tutoring
		Looking for tutor, have not found help yet
		Child receiving help from family or school
Reasons why the child missed school	EDC_Q130 - Specify the other reason	For mental health, wellness
		Attending or participating in other extracurricular activity
		Tired, having sleeping issues
		COVID-19
Reasons why the child left school	EDC_Q140 - Specify the other reason	COVID-19
Main reason you did not finish your postsecondary education	PSE_Q40 - Specify the main reason	COVID-19
Trauma and discrimination		
Types of situations where you have experienced this discrimination in the past five years	DIS_Q05 - Specify the other situations	While travelling, including at airport, hotel, etc.
		Social settings, gatherings and events
		In public, by society in general
Relationship between missing and murdered Indigenous women, girls and Two-Spirit people and the respondent	MMIW_Q10 - Other - Specify the relationship	Acquaintance
		Extended relative
		Father, grandfather, son, brother, grandson
		Uncle, nephew
Relationship of the offender to the respondent - All in one incident	VIC_Q35A - Specify the other relationship	Client at work
Relationship of the offender to the respondent- Over multiple incidents	VIC_Q35B - Specify the other relationship	Client at work
Labour		
Child care responsibilities impact on employment	CCIE_Q25 - Specify the other impacts	Parental leave or maternity leave
		Stay at home with child
		Flexible working arrangements
		Take leave

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

The table below summarizes the extra categories added to certain IPS–NIS survey questions during the coding of data from questions involving an "Other-specify" category.

Table B.2
Indigenous Peoples Survey–Nunavut Inuit Supplement: Extra response categories created for "Other-specify" questions

Theme	Questions	Extra categories
Skills		
Reasoning that explains why they did not take any training in the past 12 months	SAT_Q020	COVID-19
Nunavut Inuit Supplement		
Reasons why they are not interested in working in a government or hamlet job in Nunavut	NIGE_Q35	Health, mental health or disability
		Aging or retired
		Current or future school attendance
Nunavut Inuit Supplement		
Kind of difficulties or challenges when applying for a government or hamlet job in Nunavut	NPAG_Q15	Not considered because of a criminal record
		Limited jobs available

Source: Statistics Canada, Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement, 2022.

Appendix C Standard classifications used for the 2022 Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement

Major field of study: Classification of Instructional Programs

Respondents of the survey provided information on the major field of study of the highest certificate, diploma or degree that they completed. Responses were coded according to the Classification of Instructional Programs (CIP) Canada 2021.

The CIP was developed in 1980 by the National Center for Educational Statistics (NCES) in the United States. NCES released updates in 1985, 1990, 2000, 2010 and 2020. CIP Canada 2021 is the fourth Canadian version of this classification; the others being CIP Canada 2000, 2011 and 2016. CIP is designed to classify "instructional programs." The organizing principle behind CIP is "field of study." At Statistics Canada, a field of study is defined as a "discipline or area of learning or training."

Industry and occupation: North American Industry Classification System and National Occupational Classification

Survey respondents provided information on their occupation and the industry in which they worked during the reference week of the survey. These responses were coded using the North American Industry Classification System (NAICS) Canada and the National Occupational Classification (NOC). Data were coded to the NAICS four-digit level of detail and NOC five-digit level and were coded to current or emerging versions of these classification systems, as follows:

NAICS - The third level of the NAICS 2022 industry group consisting of a four-digit code.

NOC - The fifth level of the NOC 2021 unit group consisting of a five-digit code.

NAICS is an industry classification system developed by the statistical agencies of Canada, Mexico and the United States. Created against the background of the North American Free Trade Agreement, it is designed to provide common definitions of the industrial structure of the three countries and a common statistical framework to facilitate analysis of the three economies. NAICS is a comprehensive system encompassing all economic activities. It has a hierarchical structure. NAICS Canada 2022 Version 1.0 consists of 20 sectors, 99 subsectors, 323 industry groups, 694 industries and 923 Canadian industries, and replaces NAICS Canada 2017 Version 3.0. NAICS 2022 Version 1.0 was the biggest revision to NAICS since 2002. It includes a large number of revisions related to the digital economy, in addition to other changes in a variety of areas. Some classification items were taken over, while others were split off, broken down, merged or transferred. This standard was approved as a departmental standard on July 30, 2021.

The NOC is the nationally accepted taxonomy and organizational framework of occupations in the Canadian labour market. It is deigned to classify occupational information from statistical surveys. The NOC provides a systematic classification structure that categorizes the entire range of occupational activity of Canada. An occupation is defined as a collection of jobs, sufficiently similar to be grouped under a common label for classification purposes. Occupations are identified and grouped primarily according to the work performed, as determined by the tasks, duties and responsibilities of the occupation. The NOC 2021 updates the NOC 2016. The NOC has been developed and maintained as part of a collaborative partnership between Employment and Social Development Canada and Statistics Canada. The NOC 2021 Version 1.0 overhauls the "skill level" structure by introducing a new categorization representing the degree of training, education, experience and responsibilities required for an occupation. The NOC 2021 Version 1.0 also introduces a new five-digit hierarchical structure, compared with a four-digit hierarchical structure in the previous versions of the classification.

Indigenous languages

For the coding of Indigenous languages captured on the 2022 Indigenous Peoples Survey (IPS) and Indigenous Peoples Survey–Nunavut Inuit Supplement (IPS–NIS), the survey used the language categories used in 2021

Census. A total of 60 Indigenous language categories were used, coming from a combination of the high level and the detailed level of languages used by the census. The following link provides more information: List of languages 2021 (statcan.gc.ca).

The expansive list of languages that respondents were able to select from and that were promoted by a detailed section instead of just high-level language categories (e.g., Algonquian languages versus Cree language) has resulted in an increased number of Indigenous languages being reported in the survey.

To note from the 2021 Census, the number of Indigenous people in Canada who reported they could speak an Indigenous language well enough to conduct a conversation in 2021 was down from what was reported in 2016. This is the first decline since comparable data started being collected in 1991. See Indigenous languages across Canada.

The dataset from the IPS and IPS–NIS will include additional data on those with some knowledge of Indigenous languages but who have less than a conversational ability. Further, it will include information on exposure to Indigenous languages, how Indigenous languages are learned and self-rated importance of Indigenous languages.

Appendix D Glossary of Indigenous Peoples Survey and Indigenous Peoples Survey–Nunavut Inuit Supplement survey terms

A

Administrative data

Administrative data are information that is collected by other government agencies and private sector companies for their own purposes, which is then used by Statistics Canada to efficiently accomplish its mandated objectives. Statistics Canada treats all data that can identify a person, a business or an organization with strict confidentiality.

Advertising

Advertising consists of any message (regardless of format) conveyed in Canada or abroad and paid for by the government for placement in media, including but not limited to newspapers, television, radio, cinema, billboards and other out-of-home media, mobile devices, the Internet, and any other digital medium. All advertising activities must be vetted and approved by the Privy Council Office (PCO) and Public Services and Procurement Canada (PSPC) and coordinated by Statistics Canada's advertising manager.

Analytical file

A Statistics Canada **microdata** set for a given survey, available for use in research data centres (RDCs) across Canada. RDCs provide researchers with access, in a secure setting, to **microdata** from population and household surveys. The centres are staffed by Statistics Canada employees. They are operated under the provisions of the ***Statistics Act*** in accordance with all the **confidentiality** rules and are accessible only to researchers with approved projects who have been sworn in under the ***Statistics Act*** as "deemed employees."

B

Bootstrap method

The bootstrap method is an approach for estimating **error** in a **dataset** related to **sampling**. **Sampling** introduces **error** because **data** are not taken from the entire **population**, but only a subsection, called a sample, which is then used to make estimates for the whole population. There are several methods for **estimating** the level of **sampling error**. The bootstrap method usually selects a number of subsamples from the main sample and produces estimates for each subsample. The **sampling error** is estimated as a function of the observed differences between estimates from the different subsamples and estimates from the complete sample.

C

CAPI Lite Plus (CLP)

CAPI Lite Plus (CLP) is a data collection method where interviewers visit selected individuals in person to have their information verified and updated. An appointment is scheduled with them to later complete the questionnaire via CATI (computer-assisted telephone interview).

CEGEP

(Collège d'enseignement général et professionnel) Also written cégep, CÉGEP and cegep. A publicly funded college providing technical, academic, vocational or a mix of programs; exclusive to the province of Quebec's education system.

Census metropolitan area

A census metropolitan area (CMA) is formed by one or more adjacent municipalities centred on a **population centre** (known as the core). A CMA must have a total **population** of at least 100,000 of which 50,000 or more must live in the core based on adjusted data from the previous **Census of Population** Program.

Census of Population

A census is the collection of information about all units in a **population**, sometimes also called a 100% sample survey. Under the ***Statistics Act*** of 1971, it is a statutory requirement to conduct a nationwide census every five years. The Census of Population provides information needed to plan community services such as schools, day care, police services and fire protection, to forecast consumer demand and to conduct market research studies.

Census subdivision

Census subdivision (CSD) is the general term for municipalities (as determined by provincial and territorial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories).

Coefficient of variation

In a sample survey, results from the sample are used to estimate what the findings would be if the whole population were to be measured. In this process of estimation, some level of **error** is inevitable. The coefficient of variation (CV) is a way of expressing the **sampling error** associated with an **estimate**. First a **standard error** or "average" **error** of the **estimate** is calculated. The CV is obtained by dividing the **standard error** of the **estimate** by the estimate itself and expressing the resulting fraction as a percentage. The lower the CV, the higher the **data quality** (see **margin of error**).

Colorectal cancer

A malignant tumour that arises in the colon (large intestine) or rectum.

Confidence interval

A confidence interval (CI) around an estimate indicates the degree of confidence that the interval contains the true population value. The CI places upper and lower bounds around a point estimate. It is affected by sample size and variability of the characteristic studied. The greater the sample and the lower the variability, the narrower the interval and thus the more precise the estimate. It is considered a best practice at Statistics Canada to report the sampling error of an estimate through its 95% confidence interval.

Confidential information

This is a term used within Statistics Canada to describe information that is subject to the secrecy provisions of the ***Statistics Act***. Information is deemed confidential either because it directly identifies a **responding unit**, for example, by name, or because it could permit specific **responding units** to be identified, even when the data is stripped of identifiers, due to the information's detail or its geographical structure or format.

Confidentiality

Confidentiality denotes an implied trust relationship between the person providing the information and the individual or organization collecting it. This relationship is built on the assurance that the information will not be disclosed without the person's permission. Under the ***Statistics Act***, information that would identify an individual, business or institution cannot be disclosed without their knowledge or consent.

Coverage

Coverage is the extent to which every person or unit intended for inclusion in a survey or census is in fact counted and counted only once. Coverage **errors** refer to when persons or units of the survey or census are missed (under-coverage) or over-counted (over-coverage). Studies are often conducted by Statistics Canada to provide **estimates** of under-coverage and over-coverage of a given survey or census or to examine related issues. For example, Statistics Canada has studied and analyzed the extent to which cell-phone use affects coverage for telephone surveys.

COVID-19

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus.

For detailed information about COVID-19, see: Coronavirus disease (COVID-19).

D

Data

Observations and measurements collected during a survey, census or other study. Facts or figures from which conclusions can be drawn.

Data quality

A degree or level of confidence that the **data** and statistical information are "fit for use." The particular issues of quality or fitness for use that must be addressed by Statistics Canada are relevance, accuracy, timeliness, accessibility, **interpretability** and coherence.

Dataset, database

An organized and sorted list of facts or information about a set of individuals, households, businesses, or other relevant units. A Statistics Canada dataset is usually generated by a survey or administrative **data**, stored on a computer, and organized in such a way that it may be accessed easily by a wide variety of statistical application programs.

Derived variable

A new **variable** constructed by applying logical or mathematical operations to one or more existing **variables** to meet particular **data** needs. For example, an age variable can be derived from date of birth information. As another example, a derived variable could be obtained called "presence of a chronic health condition" based on whether or not a **respondent** answered "yes" at least once to a series of questions asking about specific chronic health conditions such as asthma, diabetes, heart disease, etc.

Dissemination

The process of providing statistical products and services to the general public and to specific **data** users. Statistics Canada disseminates **data** and analysis in the form of survey results, research reports, technical papers, periodical magazines, census products and research compendia. Online products date from 1996 to the present. Historical material can be located using the library catalogue. Statistics Canada information is also distributed to an approved network of depository libraries. The objective of dissemination activities is to provide relevant information in a timely fashion, in useful formats, and through accessible channels. Activities in place to support the dissemination of products include client consultation services, marketing, promotions, user-training and other client services.

Dwelling

A dwelling is defined as a set of living quarters. It refers to a separate set of living quarters with a private entrance either from outside the building or from a common hall, lobby, vestibule or stairway inside the building. The entrance to the dwelling must be one that can be used without passing through the living quarters of some other person or group of persons.

E

Editing

Editing is a process that ensures survey **data** are accurate, complete and consistent. A set of editing rules or conditions is applied to a **dataset**. **Data** which do not meet the conditions are examined and corrected where appropriate.

Electronic template

Supplementary promotional handouts containing additional information for participants about the IPS and IPS–NIS such as relevant data points from past survey cycles and additional reference materials.

Errors

In a sample survey, results from the sample are used to estimate what the findings would be if the whole **population** were to be measured. The accuracy of such an **estimate** is a measure of how much the **estimate** differs from

the correct or "true" figure. Departures from true figures are known as errors. Errors can arise from many sources, but can be grouped into a few broad categories: **coverage errors, non-response errors, response errors, processing errors** and **sampling errors**.

Coverage errors

Coverage errors refer to when persons or units of the survey are missed (undercoverage) or over-counted (overcoverage).

Non-response errors

Non-response errors occur when it proves impossible to obtain a complete questionnaire from a person, household, or organization. Although certain adjustments for missing **data** can be made during processing, non-response means that some loss of accuracy is inevitable.

Processing errors

Processing errors include mistakes made during **data** entry, coding, tabulation or other forms of **data** manipulation.

Response errors

Response errors indicate that a response may not be entirely accurate. The **respondent** may have misinterpreted the question or may not know the answer, especially if it is given for an absent household member, for example.

Sampling error

Sampling error refers to the fact that the results of the weighted sample differ somewhat from the results that would have been obtained from the total **population**. The difference is known as sampling error. The actual sampling error is, of course, unknown, but it is possible to calculate an "average" value, known as the "**standard error**."

Estimate, estimation

Using results of the weighted sample to estimate the characteristics of the total **population**.

Electronic questionnaire

An electronic questionnaire (EQ) is a series of questions to gather data about a target population, conducted through a web-based application.

F

Fibromyalgia

A condition that usually presents as chronic aching and pain in the muscles, usually of the back and neck, typically accompanied by stiffness, especially in the morning.

First Nations, First Nations people

A term that came into common usage in the 1970s to replace the word "Indian," which many people found offensive. Although the term First Nations is widely used, no legal definition of it exists. Among its uses, the term "First Nations people" refers to the **North American Indian** people in Canada, both Status and Non-Status. Many people have also adopted the term "First Nation" to replace the word "band" in the name of their community.

Frame

A list, map, or conceptual specification of the units comprising the survey **population** from which persons can be selected. For example, a telephone or city directory, or a list of members of a particular association or group.

Frequency

The number of times an event or item occurs in a **dataset**.

G - H - I

Gender

Gender refers to an individual's personal and social identity as a man, woman or non-binary person (a person who is not exclusively a man or a woman).

Gender includes the following concepts:

- gender identity, which refers to the gender that a person feels internally and individually
- gender expression, which refers to the way a person presents their gender, regardless of their gender identity, through body language, aesthetic choices or accessories (e.g., clothes, hairstyle and makeup), which may have traditionally been associated with a specific gender.

A person's gender may differ from their sex at birth, and from what is indicated on their current identification or legal documents such as their birth certificate, passport or driver's licence. A person's gender may change over time.

Hamlet

A municipal corporation in Nunavut. A municipality is similarly an "area within the boundaries of a hamlet, as described in the order establishing or continuing the hamlet." All of Nunavut's 25 municipalities are hamlets except for the City of Iqaluit, which is the territory's capital.

Health regions

Health regions are legislated administrative areas defined by provincial ministries of health. These administrative areas represent geographic areas of responsibility for hospital boards or regional health authorities. Health regions, being provincial administrative areas, are subject to change.

Imputation

Imputation involves replacing either missing or invalid **data** with valid **data**. This is normally performed using predetermined rules or with the use of data from a "statistical neighbour"—another **responding unit** who has similar characteristics. Imputation is often combined with data **editing**.

Indian Act

The Canadian federal legislation, first passed in 1876, that sets out certain federal government obligations, and regulates the management of Indian reserve lands. The act has been amended several times, most recently in 2017.

Indian band

A group of **North American Indian** people for whom lands have been set apart and money is held by the Crown. Each band has its own governing band council, usually consisting of one or more chiefs, and several councillors. Community members choose the chief and councillors by election, or sometimes through traditional custom. The members of a band generally share common values, traditions and practices rooted in their ancestral heritage. Today, many bands prefer to be known as **First Nations**.

Information

Data that have been recorded, classified, organized, related or interpreted within a framework so that meaning emerges.

Information product

Organization of results from Statistics Canada activities, including **data** files, **databases**, tables, graphs, maps, and text. This organization can be either pre-defined (standard information product) or made in response to special requests (customized information product). Information products can be made available on either print or electronic media.

Indigenous peoples

Indigenous peoples in Canada include three distinct groups: **First Nations (North American Indian)**, **Métis** and **Inuk** (Inuit), each recognized in the *Constitution Act*. There are many cultural, historical, regional, political and socioeconomic differences between these groups, as well as within each of the three groups.

Interpretability

Interpretability reflects the ease with which the user may understand, properly use and analyze the data or information. The degree of interpretability is largely determined by the adequacy of definitions on concepts, target populations and variables; terminology underlying the data; and information on any limitations of the data.

Inuit

A word that means "people" in Inuktitut, the most commonly used dialect of the Inuktut language of Inuit people. The majority of Inuit live in the Northwest Territories, Nunavut, Northern Quebec and Labrador. **Inuk** is the singular form of the word Inuit (i.e., a person).

Inuit Child First Initiative

The Inuit Child First Initiative ensures Inuit children have access to the essential government funded health, social and educational products, services and supports they need, when they need them.

Inuit Nunangat

Inuit Nunangat is the homeland of Inuit of Canada. It includes the communities located in the four Inuit regions: Nunatsiavut (Northern coastal Labrador), Nunavik (Northern Quebec), the territory of Nunavut and the Inuvialuit region of the Northwest Territories. These regions collectively encompass the area traditionally occupied by **Inuit** in Canada.

Inuk

The singular form of the word **Inuit** (i.e., a person).

Inuktitut

A dialect of the Inuktut language, one form of which is written with Indigenous syllabics. It is used in several regions across Inuit Nunangat and is the most commonly used Inuktut dialect. It is recognized as an official language in Nunavut.

Inuinnaqtun

A dialect of the Inuktut language, spoken mainly in Nunavut but also in the Northwest Territories. It is recognized as an official language in Nunavut.

Inuvialuit Final Agreement

An agreement that came into effect on July 25, 1984, and was the first land claim agreement settled in the Northwest Territories. The agreement defines the Inuvialuit Settlement Region (ISR), which covers approximately 435,000 square kilometres in the Mackenzie Delta, Beaufort Sea, and Amundsen Gulf area and includes a part of Yukon.

J – K – L

James Bay and Northern Quebec Agreement (Nunavik)

An agreement that addresses the use, management and ownership of land and resources in James Bay, Hudson Bay, Hudson Strait and Ungava Bay, as well as a portion of northern Labrador and an offshore area adjacent to Labrador. The agreement was ratified on December 1, 2006. It came into effect on February 14, 2008.

Jordan's Principle

Jordan's Principle makes sure all First Nations children living in Canada can access the products, services and supports they need, when they need them. Funding can help with a wide range of health, social and educational

needs, including the unique needs that First Nations Two-Spirit and LGBTQQIA children and youth and those with disabilities may have.

Jordan's Principle is named in memory of Jordan River Anderson. He was a young boy from Norway House Cree Nation in Manitoba.

Knock, Talk and Call

Knock, Talk and Call (KTC) is a method of collection similar to CAPI [computer-assisted personal interview] Lite Plus (CLP), where interviewers visit selected individuals in person, but unlike CLP, for KTC, a computer-assisted telephone interview is scheduled with the respondent immediately.

Labrador Inuit Land Claims Agreement (Nunatsiavut)

An agreement that serves as a modern-day treaty between the Inuit of Labrador, the province of Newfoundland and Labrador and Canada. The Agreement addresses land use, management and ownership of resources of regions in Northern Labrador. The agreement was signed in January 2005 and came into effect in December 2005.

Land claims beneficiary, enrolled under a land claims agreement – An Inuk who is on the enrolment list of one of the four Inuit land claim regions (see below). This person is entitled to the benefits outlined within each specific land claims agreement.

Learning disorder or difficulty

A condition thought to be associated with neurological dysfunction that is marked by inability to master a skill, such as reading or numerical calculation. This condition affects learning in individuals who otherwise demonstrate at least average abilities essential for thinking or reasoning. As such, learning disabilities are distinct from global intellectual deficiency.

Logistic regression

A form of **regression** analysis used when the response **variable** is a binary **variable** (a **variable** having two possible values).

M

Margin of error

In a sample survey, results from the sample are used to estimate what the findings would be if the whole **population** were to be measured. In this process of **estimation**, some level of error is inevitable. The margin of error, a measure used to build confidence intervals, serves as a rough indicator of the precision of an estimate. For example, pollsters often say that a certain percentage of the **population**, plus or minus the margin of error (expressed in percentage points), is likely to vote for a certain candidate, 19 times out of 20. To calculate the margin of error, which in this example corresponds to a 95% confidence interval, the pollster would use the equivalent of plus or minus two **standard errors** of the **estimate** (see **Standard error**).

Methodology

A set of research methods and techniques applied to a particular field of study. At Statistics Canada, methodology refers to survey methodology.

Métis

There is no single definition of Métis. To some respondents, Métis refers to the Métis Nation; to others, it might refer to a person of mixed Indigenousand European ancestry who self-identifies as Métis.

Métis organization or Settlement

A Métis organization or Settlement includes organizations that are a signatory under the Canada-Métis Nation Accord of 2017 (Métis Nation of Ontario, Manitoba Metis Federation, Métis Nation – Saskatchewan, Métis Nation of Alberta, Métis Nation British Columbia) or one of the eight Métis Settlements established by the Alberta *Métis Settlements Act* in 1990 (Buffalo Lake Metis Settlement, East Prairie Metis Settlement, Elizabeth Metis

Settlement, Fishing Lake Metis Settlement, Gift Lake Metis Settlement, Kikino Metis Settlement, Paddle Prairie Metis Settlement, Peavine Metis Settlement).

Microdata

Files of **records** pertaining to individual **responding units**.

N - O

National Household Survey

This survey took place in 2011 as a replacement for the long census questionnaire, more widely known as Census Form 2B and 2D. The National Household Survey (NHS) was designed to collect social and economic **data** about the Canadian population. The objective of the NHS was to provide **data** for small geographic areas and small population groups. For more information on the 2011 NHS, please visit: Census of Population.

Non-Status Indians

Non-Status Indians include all persons other than those who belong to the four categories specified in "Status Indians" (see "Status Indians" above). They include persons who may be entitled to register under provisions of the *Indian Act*, but for some reason have not registered.

North American Indian

A term that describes all Indigenous people in Canada who are not **Inuit** or **Métis**. North American Indian peoples are one of three groups of people recognized as Indigenous in the *Constitution Act, 1982*. This also refers to **First Nations** people, consisting of Status and non-Status **Indians**.

Nunavut Agreement (formerly known as the Nunavut Inuit Claims Agreement)

Signed on May 25, 1993, by representatives of the Tunngavik Federation of Nunavut (now Nunavut Tunngavik Incorporated), the Government of Canada and the Government of the Northwest Territories. This agreement gave the Inuit of the central and eastern Northwest Territories a separate territory called Nunavut. It is the largest Indigenous land claims settlement in Canadian history.

Nunavut Sivuniksavut

The term means "our land is our future." This Nunavut postsecondary education program is dedicated to providing Inuit youth with academic as well as cultural learning experiences.

Observation

Data collected for a given **variable** about a particular **responding unit**. Examples include the specific values for a **responding unit** on characteristics such as age, gender or marital status—the observation might be "77," "woman" and "widowed."

On or off reserve

"On reserve" includes eight census subdivision (CSD) types legally affiliated with First Nations or Indian bands, i.e., Indian reserve (IRI), Indian settlement (S-É) (except for the two Indian settlements of Champagne Landing 10 and Kloo Lake, located in Yukon), Indian government district (IGD), Terres réservées aux Cris (TC), Terres réservées aux Naskapis (TK), Nisga'a land (NL), Tsawwassen lands (TWL) and Tla'amin lands (TAL).

"Off reserve" includes all CSDs in Canada not defined as "on reserve."

P

Partner

Refers to two people of the opposite sex or of the same sex, who live together as a couple, but who are not legally married.

Population centre

The term population centre replaces the term urban area (as used in the **Census of Population** until 2006). A population centre is defined as an area with a population of at least 1,000 and no fewer than 400 persons per square kilometre. Population centres are classified into three groups, depending on the size of their population:

- small population centres, with a population between 1,000 and 29,999;
- medium population centres, with a population between 30,000 and 99,999; and
- large urban population centres, with a population of 100,000 or more.

Postcensal survey

A postcensal survey is one where **surveyed units** are selected based upon their responses to the **Census of Population**. These surveys are generally conducted shortly after the census.

Promotion

Promotion refers to communications activities that inform the public of programs and services. These can include posters, handouts, public service announcements and organic social media posts.

Proportion

A proportion refers to how many responses fall into a given response category in relation to the total responses. It is calculated by dividing the **frequency** of the response category by the total number of responses to the question.

Public use microdata file (PUMF)

Public use microdata files provide access to **responding units** so that users can conduct their own research or analysis. They involve a non-identifiable **data set** containing characteristics pertaining to the units of the survey (e.g., individuals, households or businesses). All such **datasets** have been authorized for release to the public by the Statistics Canada Microdata Release Committee. The **dataset** contains no **confidential information** in that individual identifiers have been removed and any **data** combination or geography which could potentially reveal the identity of a **responding unit** has been modified.

Q - R

Record

A record is the data for an individual **responding unit** in a file containing **data** for all of a survey's **responding units**.

Regression

A statistical method which tries to predict the value of a characteristic by studying its relationship with one or more other characteristics. This relationship is expressed through the means of a regression equation.

Research data centres

The Research Data Centre Program provides researchers with access, in a secure Statistics Canada governed setting, to **microdata** from population and household surveys. The RDC program is part of an initiative by Statistics Canada, the Social Sciences and Humanities Research Council and university consortia to help strengthen Canada's social research capacity and to support the policy research community. The program is also supported by the Canadian Foundation for Innovation and the Canadian Institutes of Health Research.

Respondent, responding unit

The respondent is the person providing the information for the **surveyed unit**, which could be a person, household, business or institution. In the case of the 2022 Indigenous Peoples Survey, in general, the respondent is the selected person aged 1 and older.

Response rate

The proportion of a sample for which a response to a questionnaire is obtained, usually expressed as a percentage. Non-response covers those who refused to participate as well as persons whom the survey was unable to reach.

Rooms

Rooms refers to enclosed areas within a private dwelling which are finished and suitable for year-round living. The number of rooms in a private dwelling includes kitchens, bedrooms and finished rooms in the attic or basement. The number of rooms in a private dwelling excludes bathrooms, halls, vestibules and rooms used solely for business purposes. Partially divided rooms are considered to be separate rooms if they are considered as such by the respondent (e.g., L-shaped dining room and living room arrangements).

S

Sample design, sampling design

A set of specifications that describe the **sampling** elements of a survey in detail. These elements include **population, frame, surveyed units**, sample size, sample selection and **estimation** method.

Sampling

The process of selecting some part of a **population** to observe so as to estimate something of interest about the whole **population**. Examples of different sampling methods include simple random sampling, stratified random sampling, cluster sampling, multiple-phase sampling and multi-stage sampling.

Sampled unit

The unit selected by the **sample design** and from which measurements are taken for a survey. Examples include persons, households, families or businesses. For the IPS, the sampling unit is the person.

Sampling fraction

Sample size divided by the **population** size.

Sexual orientation

Romantic and sexual attraction for people of the same or another sex or gender.

- Heterosexual: A person who is sexually or romantically attracted to people of a different gender than themselves.
- Lesbian: Typically a woman who is sexually or romantically attracted to other women.
- Gay: A person who is sexually or romantically attracted to people of their same sex or gender identity. Traditionally, this identity was reserved for men, but it has been adopted by people of all gender identities.
- Bisexual: A person who is sexually or romantically attracted to two or more genders.

Standard deviation

Standard deviation measures the dispersion of a **dataset** around the mean. It is the most widely used measure of dispersion. Mathematically, the standard deviation is the square root of the **variance**.

Standard error

In a sample survey, results from the sample are used to estimate what the findings would be if the whole **population** were to be measured. **Sampling error** refers to the fact that the results of the weighted sample differ somewhat from the results that would have been obtained from the total **population**. The difference is known as **sampling error**. The actual **sampling error** is, of course, unknown, but it is possible to calculate an "average" value, known as the "standard error."

Statistics Act

An act regarding statistics of Canada. It includes the definition of Statistics Canada's mandate: "There shall continue to be a statistics bureau under the Minister, to be known as Statistics Canada, the duties of which are:

- to collect, compile, analyze, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and condition of the people;
- to collaborate with departments of government in the collection, compilation and publication of statistical information, including statistics derived from the activities of those departments;

- to take the census of population of Canada and the census of agriculture of Canada as provided in this Act;
- to promote the avoidance of duplication in the information collected by departments of government; and,
- generally, to promote and develop integrated social and economic statistics pertaining to the whole of Canada and to each of the provinces thereof and to coordinate plans for the integration of those statistics."

Status Indian

"Status Indians" (Registered or Treaty) are persons who:

- are Registered Indians under the *Indian Act*
- are Treaty Indians, only if they are Registered Indians under the *Indian Act*
- have become registered since June 1985, when Bill C-31 changed the *Indian Act*
- have become registered since April 2010, when Bill C-3 changed the *Indian Act*.

Stratification A **sampling** procedure in which the **population** is divided into homogeneous subgroups or strata and the selection of samples is done independently in each stratum.

Suppress, suppression

The process by which particular **data** are prevented from being released based on criteria designed to protect **confidentiality**. "Cell" suppression refers to procedures used to protect sensitive tabular data from disclosure; a cell being an individual entry in a table. For the IPS, **data** were also suppressed for reasons of data quality (coefficient of variation larger than 33.3%).

Surveyed unit

The selected unit from which measurements are taken for a sample survey or a census. Examples include persons, households, families or businesses. For the IPS, the surveyed unit (which is also the **sampled unit** since the IPS is a sample survey) includes persons aged 1 and over.

T - U - V

Target population

The complete group of units to which survey results are to apply. These units may be persons, households, businesses, institutions, etc. This is the population for which information is wanted.

University diploma or certificate below bachelor's level

University certificates are usually one-year programs and are offered in most fields of study. This designation applies when a teaching certificate is awarded by a provincial department of education at an approved institution. This includes certificates or diplomas awarded by a professional association if the courses were taken at a university but a bachelor's degree was not a prerequisite, for example in fields such as accounting, banking, insurance or public administration. In Quebec, a "diplôme de premier cycle" would be included here.

User guide

These guides accompany Statistics Canada survey **datasets**, such as **analytical files** and **public use microdata files (PUMF)**, providing the detailed technical information required to use the **data** appropriately. The guide typically contains important information to know prior to data analysis: weighting **variables** to use, procedures related to the estimate of **variance**, and precautions to take in the dissemination of the **data**.

Variable

A characteristic that may assume more than one value to which a numerical measure can be assigned (e.g., income, age and weight).

Variance

A measure of dispersion for a given characteristic or **variable** in a **dataset**. It indicates how much variability exists for that characteristic. Technically, it is calculated as the average squared deviation from the mean of each **observation** in the **dataset** for a particular **variable**.

Victimization

Includes any crimes (as defined by the *Criminal Code*) against a person (e.g., robbery, sexual assault, physical assault).

W - X - Y - Z

Weight

A weight is the average number of units in the **population** that a unit in the survey represents. Examples of a unit include a person or a household. Weights are applied to **responding units** in a sample **database** to ensure that, when making inferences from the survey **data** to population parameters, **estimates** of characteristics for the total **population** are obtained.

Appendix E Confidence Intervals Construction

This annex provides general information about confidence intervals as well as details about the methods used to construct the confidence intervals disseminated by Statistics Canada for the 2022 postcensal surveys, namely the Canadian Survey on Disability (CSD), the Survey on the Official Language Minority Population (SOLMP), the Indigenous Peoples Survey (IPS) and the Nunavut Inuit Supplement (IPS–NIS). These methods are the modified Wilson confidence interval for proportion-type statistics and the bootstrap percentile method for other statistics. This document also provides high-level information about the underlying assumptions and the properties of these methods. Users interested in theoretical details are invited to consult the referenced papers.

About confidence intervals

Confidence intervals are provided along postcensal surveys estimates so that users can perform valid statistical inferences. Statistical inference is the process of drawing conclusions about the population based on data collected from survey respondents. Survey estimates produced from a sample are subject to uncertainty because of the sampling process—different estimates could have been obtained if a different sample had been selected—and of non-response. This uncertainty must be considered to do proper statistical inference.

As described in the documentation of each survey, variance estimates are produced to quantify the uncertainty of estimates. These variance estimates can be used to produce variance-based quality measures such as standard errors, coefficients of variation and confidence intervals. Among these measures, confidence intervals were chosen to accompany estimates for the 2022 postcensal surveys because they convey that there is uncertainty around the estimate by providing a range of possible values. They also provide an objective measure of the sampling error in a form that is easy to interpret, since wider confidence intervals are naturally associated with greater uncertainty. Moreover, confidence intervals are appropriate for all types of estimates, which is not the case for some other variance-based quality measures.

A confidence interval is associated with a confidence level. The confidence level describes the degree to which one can be confident that the true population parameter is contained in the confidence interval. The confidence level is often written as $1-\alpha$, where α is the significance level of the corresponding hypothesis test. For the postcensal surveys, the confidence level is set to 95% and α is thus equal to 0.05. The confidence level corresponds to the expected coverage of confidence intervals. In other words, if the process that generated the sample data was repeated a very large number of times and corresponding confidence intervals were constructed for each estimate using the same method, the proportion of these confidence intervals containing the true population value should be close to the stated (or nominal) confidence level. Using simulation studies, confidence intervals methods can be evaluated to ensure that their actual coverage is close to their nominal confidence level.

Several different methods can be used to construct confidence intervals. The most commonly used and default in most statistical software are the Wald and the Student's methods.[4] Both of these methods assume a known distribution for the estimator, i.e., the statistical function used to produce estimates. The distribution of the estimator can be seen as the geometrical shape that would appear if all possible samples were drawn from the population and estimation was carried out for each of them, yielding a different estimate each time. While the Wald and the Student's confidence intervals are easy to compute and generally valid, they rely on assumptions that do not hold in some contexts, which can lead to undercoverage.

Other methods, such as modified Wilson for proportions, make different assumptions of the distribution of the estimator. Their improved coverage makes them a better choice for valid statistical inference. Interestingly, the bootstrap percentile confidence interval is obtained in a completely different way. Rather than assuming a shape for the distribution of the estimator, the bootstrap method uses the empirical distribution that arises from the estimates computed from each of the bootstrap replicates. Both methods are described in more detail in the next sections.

4. Wald and Student's confidence intervals are discussed in Sections 7.2 and 7.3 of Sampling and Weighting Technical Report, Census of Population, 2021.

Modified Wilson confidence interval

For the postcensal surveys, this method is used for all proportion-type statistics (e.g., proportions,[5] percent distribution,[6] shares[7]).

The modified Wilson confidence interval method was chosen because of its generally superior coverage for proportion-type estimators and for its practicality of implementation. The method is based on the Wilson confidence interval for a simple random sampling with replacement (SRSWR) sample design (Wilson 1927). For the postcensal surveys, a modified version of the Wilson confidence interval that has been adapted to complex sample designs is used (Kott and Carr 1997). Theoretical developments and extensive simulation studies have shown that this method has good properties in most situations. It performs better than the Wald and Student's confidence intervals in situations where those confidence intervals exhibit undercoverage for proportion-type statistics because the assumptions they rely on do not hold (Neusy and Mantel 2016; Statistics Canada 2023).

Method

The lower bound and the upper bound of a 95% modified Wilson confidence interval for a proportion-type statistic *p* are given by:

$$lower\ bound = \frac{\hat{p} + z^2/2n_e}{1 + z^2/n_e} - \frac{z\sqrt{\hat{p}(1-\hat{p}) + z^2/4n_e}}{\sqrt{n_e}\left(1 + z^2/n_e\right)}$$

$$upper\ bound = \frac{\hat{p} + z^2/2n_e}{1 + z^2/n_e} + \frac{z\sqrt{\hat{p}(1-\hat{p}) + z^2/4n_e}}{\sqrt{n_e}\left(1 + z^2/n_e\right)}$$

where:

- $\hat{p}$ is the estimate of *p*;
- z is the $1-\alpha/2 = 97.5^{th}$ percentile of the standard normal distribution;
- $n_e = \min\left(\frac{n}{deff(\hat{p})}, n\right)$ is the effective sample size;
- $deff(\hat{p}) = \frac{\hat{V}(\hat{p})}{\hat{p}(1-\hat{p})/n}$ is the estimated design effect of $\hat{p}$ with respect to the SRSWR sample design;
- n is the sample size (i.e., the unweighted count of respondents) in the denominator of the proportion;
- $\hat{V}(\hat{p})$ is the estimated variance of $\hat{p}$.

When $\hat{p} = 0$ or $\hat{p} = 1$, the value for n_e is undefined. In this case, the bounds for the Wilson confidence interval are computed as:

- For $\hat{p} = 0$, the confidence interval bounds are $\left(0, \frac{1}{1 + n/z^2}\right)$
- For $\hat{p} = 1$, the confidence interval bounds are $\left(\frac{1}{1 + n/z^2}, 1\right)$.

5. A proportion is a relationship between two quantities, often expressed as a fraction. It usually represents the comparison of one part to the whole.
6. Percent distribution refers to the breakdown of a total into its constituent parts, expressed as a percentage of the whole.
7. A share typically refers to the portion or percentage of a particular category or segment within a dataset. In the context of income distribution, a share refers to the portion or percentage of the total income held by a specific group.

Properties

In addition to achieving better coverage than the Wald and Student's intervals for small sample sizes or when the population parameter is near zero or one, the modified Wilson confidence interval for a proportion has the desirable range-preserving property, meaning that its lower bound is never less than zero and its upper bound is never greater than one. Since proportions cannot take on values outside of the interval between zero and one, it is reasonable that confidence intervals for proportions would exclude negative values and values greater than one.

It should also be noted that, unlike the Wald and Student's intervals, the modified Wilson confidence interval for proportions may be asymmetric, meaning that the estimate will not be exactly at the centre of the interval. The asymmetry is small when the effective sample size is large or when the estimated proportion is near 0.5.

Much like the Wald and Student's confidence intervals, the modified Wilson confidence interval for proportions may suffer from some undercoverage, particularly when the sample size is very small, the value of the proportion is near zero or one, or there is high correlation between members of the same household. However, the modified Wilson method generally achieves nominal coverage rates in extreme situations, compared with the Wald and Student's methods. It generally maintains coverage as good as or better than the Wald and Student's methods in those situations.

Bootstrap percentile confidence interval

For the postcensal surveys, this method is used for all statistics except the proportion-type statistics.

The bootstrap percentile method was chosen because of its straightforward way to derive confidence intervals for various and complex estimators of population parameters. Estimation is repeated with each bootstrap replicate weight and the resulting estimates are used to approximate the distribution of the estimator instead of making assumptions on its shape. The appropriate percentiles of this approximate distribution are then used to delimit an area around the estimate that corresponds to a 95% confidence interval. Theoretical developments and simulation studies have shown that this method has good performance in most situations and can be applied for complex estimators (Efron and Tibshirani 1986; Tibshirani 1984).

Method

The lower bound and the upper bound of a 95% bootstrap percentile confidence interval for a non-proportion-type statistic *Y* are given by:

$$lower\ bound = \hat{Y} + \sqrt{R}\left(\hat{Y}_{(LB)} - \hat{Y}\right)$$

$$upper\ bound = \hat{Y} + \sqrt{R}\left(\hat{Y}_{(UB)} - \hat{Y}\right)$$

where:

- $\hat{Y}$ is the estimate[8]
- R is the bootstrap adjustment factor[9]
- $\hat{Y}_{(LB)}$ is the LB^{th} non-missing ascending bootstrap estimate
- $\hat{Y}_{(UB)}$ is the UB^{th} non-missing ascending bootstrap estimate
- $LB = \frac{\alpha}{2} \times B$
- $UB = \left(1 - \frac{\alpha}{2}\right) \times B$

8. $\hat{Y}$ can be a total, a quantile, the difference of two estimates, etc.
9. *R* is often called the Fay adjustment factor for postcensal surveys. For CSD, IPS and SOLMP, R = 16.

- α is the significance level
- B is the number of replicates (excluding the number of replicates where $\hat{Y}_j = .$, where $j = 1,...,B$)[10].

If LB and UB are not integers, then $\hat{Y}_{(LB)}$ and $\hat{Y}_{(UB)}$ is the average of the two contiguous ascending bootstrap estimates. For the postcensal surveys, the values LB and UB are respectively equal to 25 and 975.

For postcensal surveys, **it is critical to apply the bootstrap adjustment factor R**, otherwise the margin of error associated with the estimate will be underestimated. In other words, the length of the confidence interval will be underestimated if the Fay adjustment factor is omitted in the calculation of the lower and upper bounds.

Properties

Unlike the Wald and Student's intervals, the bootstrap percentile confidence interval is transformation-respecting, which means that the method is valid if there exists a monotone transformation of the estimator that normalizes the distribution of the estimator. This transformation does not have to be known; it only has to exist. In the case of the Wald and Student's intervals, such a transformation would have to be explicitly specified.

Moreover, like the modified Wilson confidence interval, the bootstrap percentile interval is range-preserving, which means that the method always produces intervals that fall within the allowable range of values for the parameter.

Much like other confidence intervals, the bootstrap percentile confidence interval does not perform well when the estimator is biased. Bias in the bootstrap distribution leads to bias in the confidence interval. It is also known that the coverage of the bootstrap percentile confidence interval tends to be lower than the nominal rate when the sample size is small. Nevertheless, it usually achieves better balance in the left and right sides compared with the Wald and Student's intervals (Efron and Tibshirani 1994).

10. There are B=1,000 bootstrap replicates for each postcensal survey.

References – IPS and IPS-NIS

Bernier, J. and Nobrega, K. (1998). Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method. Statistical Society of Canada Annual Meeting, June 1998.

Canadian Council on Learning (CCL). (2009). The State of Aboriginal Learning in Canada: A Holistic Approach To Measuring Success. Ottawa.

Efron, B. and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, Statistical Science, Vol. 1, No. 1, 54-77.

Efron, B. and Tibshirani, R. J. (1994). An Introduction to the Bootstrap. Chapman & Hall/CRC.

Gagné, C., Roberts, G., and Keown, L.A. (2014). "Weighted estimation and bootstrap variance estimation for analyzing survey data: How to implement in selected software". *The Research Data Centres Information and Technical Bulletin*. (Winter) 6(1):1-72. Statistics Canada Catalogue no. 12-002-X.

Haddou, M. (2013). Bootstrap Variance Estimation Specifications - Aboriginal Peoples Survey. Internal document, January 2013.

Kott, P.S. and Carr, D.A. (1997). "Developing an Estimation Strategy for a Pesticide Data Program." Journal of Official Statistics, Vol. 13, No. 4, 367-383.

Langlet, É., Beaumont, J.-F., and Lavallée, P. (2008). Bootstrap Methods for Two-Phase Sampling Applicable to Postcensal Surveys. Paper presented at Statistics Canada's Advisory Committee on Statistical Methods, May 2008, Ottawa.

Neusy, E. and Mantel, H. (2016). "Confidence Intervals for Proportions Estimated from Complex Survey Data." Proceedings of the Survey Methods Section. SSC Annual Meeting, June 2016.

Phillips, O. (2004). "Using Bootstrap Weights with WesVar and SUDAAN". *The Research Data Centres Information and Technical Bulletin*. (Fall) 1(2):1-10. Statistics Canada Catalogue no. 12-002-XIE.

Savard, Sarah-Anne (2023). Overview of the quality component of the dissemination strategy for the 2022 Canadian Survey on Disability, October 2023, Internal document.

Statistics Canada. 2020. GTAB specifications V21. Internal Document. Ottawa, Ontario.

Statistics Canada. 2023. Canadian Survey on Disability, 2022: Concepts and Methods Guide. Catalogue no. 89-654-X2023004. Ottawa, Ontario. (accessed December 12, 2023).

Statistics Canada. 2023. Evaluating confidence interval methods for the 2021 Census using the Census Monte Carlo simulation environment. Working document. Internal Document. Ottawa, Canada.

Statistics Canada. 2023. Sampling and Weighting Technical Report, Census of Population, 2021. Catalogue no. 98-306-X2021001. Ottawa, Ontario. (accessed December 12, 2023).

Statistics Canada. 2024. Canadian Survey on Disability, 2022: User Guide to the Analytical Data Files. Internal Document available in Research Data Centers. Ottawa, Ontario.

Tibshirani, Robert (1984). Bootstrap confidence intervals, Department of Statistics Stanford University, Technical report, No. 3.

Verret, F. (2013). The estimation methodology of the 2011 National Household Survey. Paper presented at the 2013 Joint Statistical Meetings.

Wilson, E.B. 1927. "Probable Inference, the Law of Succession, and Statistical Inference." Journal of the American Statistical Association, Vol. 22, no. 158. p. 209-212.