DATA GOVERNANCE AND MANAGEMENT TOOLKIT

Data Collection & Analysis Tool

SGIGs Community Well-being Tools Project

The Community Well-being Tools Project is an initiative of the Data Steering Committee (DSC). The DSC is comprised of Self-Governing Indigenous Government (SGIG) representatives and was formed to support SGIGs in building their data governance and management capacity. Since 2018, the DSC has overseen data capacity-building initiatives for the SGIGs, including an online toolkit

The DSC is building upon this previous work by developing a suite of tools to support SGIGs in understanding and reporting on community well-being:

Tool	Description
DATA CAPACITY ASSESSMENT	Support SGIGs navigating the toolkit by highlighting relevant components of those tools to an assessment of an SGIGs' readiness, capacity, and objectives
DATA GOVERNANCE FRAMEWORK	Overarching framework describing and providing guidance with respect to the component parts of a data governance program
CONCEPTUAL WELL-BEING FRAMEWORK	Starting point Well-being Framework with guidance on specific customization options
FACILITATION TOOL	Guided steps for an optional process to create a customized Conceptual Well-being Framework
INDICATOR GUIDE & RECOMMENDED INDICATORS	A list of well-being indicators aligned within the Conceptual Well-being Framework with guidance about indicator selection and customization
DATA COLLECTION AND ANALYSIS	Options and recommendations for data collection and for creating and reporting statistics
DATA VISUALIZATION DASHBOARD	Interactive platform for reporting indicator-affiliated statistics organized around the customized Well-being Framework
WORKPLAN AND COSTING GUIDANCE	Milestones and budget guidelines for implementing these tools

Table of contents

About this tool 3
- What are primary data? 3
- When to collect primary vs. secondary data? 3

Scoping a primary data collection exercise 5
- Getting started 5
- Should you run a census or a survey? 5

Survey design 7
- Defining a sample 7
- Sources of bias 10

Questionnaire design 11
- Overview 11
- Questionnaire best practices 11
- Testing 13
- Customization 14
- Privacy 14
- Choosing software 14

Communication strategy 17
- Key considerations 17
- Communications methods 18
- Incentives 19

Enumeration 20
- Enumeration methods 20
- Enumeration planning 21
- Survey enumeration 23

Processing data 25
- Data cleaning 26
- Weighting 27

Analysis 28

Reporting 29

External support and RFP guidelines 32

References 34

About this tool

This tool outlines key considerations for effective primary data collection. It is closely related to the *Indicator Guide & Recommended Indicators Tool*, which describes a suite of indicators developed by SGIGs to measure well-being, many of which may need to be informed by primary data.

Additional resources

Some of the content included in this tool has been adapted from Statistics Canada (2010), *Survey Methods and Practices* (Catalogue no. 12-587-X. ISBN 978-1-100-16410-6), which provides detailed guidance on primary data collection and can serve as a useful companion guide for users of this tool.

What are primary data?

Primary data are information that do not already exist and that need to be collected for specific research purposes. In the context of socioeconomic research, primary data are useful when precision about the experiences of members of a specific population is important and when the data need to be current. In contrast, secondary data are pre-existing information collected for purposes other than the specific research in question. Common sources of secondary data include the Census of the Population by Statistics Canada and administrative data like property tax records. While secondary data are more general and not precisely tailored to specific research interests, they can offer cost-effective insights. Secondary data can be used for comparisons, context setting, and identifying additional research questions. Sometimes secondary data can match the research question well and they would displace the need for primary data.

When to collect primary vs. secondary data?

When choosing between primary and secondary data, consider budget, timeline, research needs, and whether existing data can answer the research question. In the case of the well-being indicators, consider whether administrative data can inform those indicators. Limited resources usually forces a reliance on secondary data, while a larger budget and time may allow for primary data collection. Primary data collection can be

difficult, but when it's required, the benefits can be tremendous! After SGIGs have gone through a proper primary data collection exercise, they will have moved along the data use continuum towards expertise, and primary data collection exercises can serve as a catalyst for organizational change more broadly.

On using administrative data

While this Toolkit focuses on the creation and use of primary data, it may be useful for many SGIGs to draw on administrative data to inform their indicators. Administrative data are collected and used to perform usual government operations and functions – for example:

- A post-secondary program may maintain records about students receiving financial assistance. These administrative data can be used to report on indicators about educational attainment, or areas of study among these students.
- A land and resource program may maintain a harvesting app. These administrative data can be used to report on indicators about traditional food harvesting activities, and locations of importance for harvesting.
- The registration program maintains a Citizenship database. These administrative data can be used to report on key demographic information of the population, such as age, where people live, and household composition.

The benefit of administrative data is that they already exist and therefore significant efficiencies are gained in utilizing these data. The drawback of administrative data sources is that they aren't designed specifically to create indicators and measure well-being. This means that they may not be precisely aligned with the ideal indicators governments have in mind, and there may be variability in data quality and completeness that will need to be addressed. Here are some questions to ask when considering using a specific set of administrative data for an indicator(s) of interest:

1. *Accuracy:* Are there systems in place that ensure data are accurate? Do the data reflect reality? How are mistakes addressed?
2. *Completeness:* Does the data set have a lot of missing values? How will this affect analysis?
3. *Consistency:* Are there any instances of conflicting information within the

dataset?

4. *Timeliness*: Does the timeline on which the data can be accessed align with project objectives and reporting plan?
5. *Validity*: Is there consistent formatting that makes the data usable? For example, are birth dates reported in a consistent format?
6. *Uniqueness*: Does each unique person only appear once in the dataset? For example, if someone is listed as "Robin A. Mackey" in one instance "Robin Mackey" in another instance, will this show up as two entries?

Use of administrative data also requires consideration about whether consent or other authorization has been provided to use the data for these other purposes.

An additional benefit of using administrative data is that they are usually collected on a regular or continuous cycle. That means, if they're a good fit to inform an indicator, that indicator can be updated in alignment with the cycle of the administrative data generating process, whereas when relying on primary data, indicators can only be updated when the entire primary data collection process is conducted.

More information about using administrative data can be found on the Indigenous Data Toolkit site.

Scoping a primary data collection exercise

Getting started

The most important step in acquiring primary data is deeply understanding the information needs that are driving data collection. At this stage, SGIGs may have already scoped, selected, and customized indicators in alignment with the approach described in the Scope section of the Indicator Guide & Recommended Indicators Tool. However, selecting indicators before undertaking a data collection exercise is not always necessary. Whether or not indicators have been selected, the two most important guiding questions to answer ahead of undertaking a primary data collection exercise are:

1. What are our priority areas for data collection?
    a. Are there specific domains of well-being that are the priority for measurement? Or is this project specifically about a comprehensive picture of well-being?
2. What do we have the capacity (people, time, and funding) to undertake?

For guidance on answering the latter question, a workplan and costing tool is available.

Should you run a census or a survey?

Once a decision has been made to collect primary data and a clear understanding exists of the information needs, resources, and capacity, a subsequent key decision is whether to collect data from the entire population or from a sample of the population. Collecting data from the entire population is called a census. The main alternative to a census is a survey. A survey is used to collect information from a subset, or a fraction, of the population. This subset of the population is called a sample. With both a census and a survey, the information collected is used to calculate statistics for the population as a whole.

Surveys are used when it is not feasible or necessary to collect information or observations from the entire population of interest. A survey is a smaller scale operation than a census, and as such it provides a faster and more economical way of collecting information. A census might not be feasible if the population is very large, geographically dispersed, or difficult to reach, or if limited resources are available, including people, time, and funding. From a technical perspective, if a complete sample frame (a full list of the population of interest) exists, a census is only required when there are important strategic reasons to conduct a census. For example, if people would feel like their voices weren't heard because they were excluded from the sample in a survey. Otherwise, the size of the sample in a survey can always be adjusted to accommodate the information needs in any primary data collection exercise. If the information needs are very detailed (e.g. an SGIG wants to be able to understand, very precisely, some things about small sub-populations) then the size of the sample may be almost the same as the entire population and therefore the exercise essentially becomes a census. Even the long-form of the Census of the Population from Statistics Canada is actually a survey, sampling 1 in 4 Canadians.

Table 1. Census vs. survey

	Census	Survey
Scope	Data are collected from the entire population. If the population is small, or there are strategic considerations, a census may be preferable.	Data are collected from a subset of the population. If the population is large, accurate results can usually be derived from a relatively small sample.
Timeliness	Time-intensive. A census can take several years to complete and so it is typically conducted only once or twice in a decade.	Can be completed in a short timeframe, ranging from weeks to months. Can be conducted frequently.
Cost	More costly. Primary data collection exercise costs typically scale with the size of the sample because enumeration is usually a main cost-driver.	Less costly overall, and more resources can be devoted to each individual response.

It is entirely possible to conduct both a census and a survey. For example, SGIGs may consider conducting a census on-lands in tandem with a survey off-lands. Several considerations inform whether to run a census or survey on-lands only, or both on-lands and off-lands. Running a census or survey on-lands is easier because enumeration is easier when the population is close together. For example, enumerators can actually go door-to-door to reach the population of interest. When running a census or survey off-lands, it is likely that the population will be very dispersed, forcing a reliance on electronic enumeration methods that can introduce errors in the sample. Generally, the enumeration requirements for people living off-lands will be greater and more expensive, so current personnel, time frame, and funding should be assessed to determine the feasibility of enumerating off-lands. Additionally, if the content of the questionnaire is program-focused and those programs are primarily or exclusively offered on-lands, including citizens or members living off-lands may not be necessary.

Scenario: Decision to undertake a census

While the Sulingituk Government has decided to undertake a comprehensive primary data collection exercise, they have not determined the exact form that this exercise will take - the big decision they are faced with now is whether to undertake a census of their entire population (members plus those living on lands), a census of members only, or a survey among a representative set of members.

To make this decision, the Sulingituk team connects with both leadership and senior administration about their desired outcomes of this exercise. Leadership emphasizes their interest in connecting with all members, as the last time a similar effort was undertaken was in the decision to assume self-government seven years previously. Administration is clear that they have interests in understanding the needs of smaller groups – including those with unstable housing and gender diverse persons – who are otherwise not well represented in the planning and programming of the Sulingituk Government. As both leadership and senior administration confirm the availability of budget for a comprehensive project, the decision is made to undertake a census, attempting to reach all Sulingituk members and those living in their households.

Survey design

Defining a sample

The goal in running a survey is to ensure that the sample is representative of the population of interest. In other words, the data collected should accurately reflect the characteristics of the entire population. The process of defining a sample can be broken down into four main steps: (1) identify the survey frame, (2) decide the sample design, (3) determine the sample size, and (4) produce the sample randomly.

A survey frame is the means of identifying and contacting the units of the survey population (e.g., individuals or households). It is essentially the list or database from which the sample is drawn for the survey. The frame can be a physical list such as a data

file of contact information, a geographic list which includes units within specific areas, or a conceptual list that provides conditions delineating who is included. For SGIGs, the survey frame will often be their citizenship or membership list. The frame chosen ultimately determines the definition of the survey population and can affect the methods of data collection, sample selection, and estimation, the cost of the survey, and the quality of its outputs. It should ideally include all (and only) the units (e.g. people, households) of interest in the population.

On the citizenship/membership list

For SGIGs, their citizenship or membership list will be a critical input to conducting either a census or a survey. Evaluating the list's data quality will be a first step to understanding what kind of data collection is possible and at what cost. If there are many errors or discrepancies in the list, addressing these before getting started may be necessary.

Here are some questions to consider when reviewing the citizenship/membership list:

- How regularly is the list maintained? How current is the information it contains?
- Does the list clearly distinguish between living or deceased individuals?
- Are there any duplicate entries?
- Is contact information accurate and available?
- Is it used for other important purposes like elections?

Once the survey frame has been established, the next step is to plan the sampling design. This step in the survey design determines how the units from the frame will be selected for participation in the survey. The sampling design should be chosen to ensure that the sample is representative of the population of interest and that the survey results will have the desired level of precision. Broadly speaking, there are two types of sampling designs: probability and non-probability sampling. At a high-level, probability sampling employs randomness when selecting a sample from a population, whereas non-probability sampling refers to a wide range of sampling techniques that do not

employ randomness. The choice of sampling design can be influenced by a variety of factors, including the nature of the population, the objectives of the survey, the statistical methods that will be employed during analysis, and the resources available. Generally, non-probability sampling is faster, easier, and less expensive, but that comes at a cost: the risk of introducing uncertainty and bias into findings. Table 2 presents common types of sampling designs and their descriptions.

Once the sampling design has been chosen, the next step is to determine the sample size. Sample size refers to the number of participants or data points selected from a larger population to be studied. It is a critical factor that impacts the precision of research findings. A larger sample size generally leads to more accurate and generalizable results, while a smaller sample size may introduce more uncertainty into the findings. The sample size should be large enough to provide the desired level of precision for the survey estimates, but not so large as to be unnecessarily costly or difficult to manage.

Technically, the size of the sample should be calculated for each statistic of interest. Since there is often interest in many statistics that will be informed by the set of questions in a questionnaire, the sample size needs to be set for the type of statistic that will require the largest sample and then used for the entire survey.

When using a probability sampling design, it is possible to compute the necessary sample size with respect to (1) the overall population size, (2) the desired margin of error, and (3) the desired confidence level. The margin of error measures the degree to which the sample responses could differ from the entire population's responses. The confidence level measures how confident one can be (in a mathematical sense) that the true parameter across the entire population differs from the sampled estimate by at most the given margin of error. For example, suppose one selects a margin of error of 2% and a confidence level is 95%. If 65% of those sampled answer "yes" to a "yes/no" survey question, then one can be 95% confident that between 63% and 67% of those in the entire population would answer yes to this question.

For non-probability sample designs it is not possible to mathematically determine how the general population's opinions could differ from the opinions of those in the sample. Nonetheless, it is common practice to employ the same formulas used to estimate the needed sample size for probability samples.

In the internet age, the formula is not so important, as it is possible to simply plug the population size, statistic of interest, margin of error, and confidence level into one of the many online sample size calculators, like the Qualtrics sample size calculator.

Table 2. Common types of sampling designs

Sampling Design Types		Description
Probability Sampling	**Simple Random Sample**	A type of probability sampling, a simple random sample is a method where everyone has the same probability of being selected. Meant for homogeneous populations that are equally reachable. For example, a survey from a cohort of students with similar background and motivation. *Strengths* • Simple to implement, analyze, and understand. • Population members are all assigned to the sample in one step. *Weaknesses* • Limited range of use cases. If the population is non-homogeneous, some members may be harder to include in the sample than others, and a simple random sample will give biased results.
	Stratified Sample	A type of probability sampling, a stratified sample is a method where a fixed number of individuals are selected from each stratum, or "group", within a population. Strata should be something with a previously known distribution. For example, age or sex in the population of people. It is important to note that, when using stratified sampling, the sample size for each strata must be calculated separately. *Strengths* • Wider use of cases • Accounts for some types of under- and over-representation *Weaknesses* • Requires additional setup and planning
Non-probability Sampling	**Snowball or Referral Sampling**	Encourage participants to share the survey with individuals in their network that match the target audience.

Sampling Design Types		Description
		Strengths • Effective for studying hard-to-reach or hidden populations. • Can result in a more diverse and knowledgeable sample as referrals are made by initial participants. *Weaknesses* • Higher risk of inaccuracy due to selection, response bias, and non-response error, caused by survey respondents not representing the population of interest.

	Convenience Sample	Convenience sampling is the method of collecting survey responses by approaching individuals that are conveniently located around a location or Internet service. For example, phoning randomly selected people for a voluntary survey.
		
		Strengths
		• Most realistic scenario for any sort of opinion survey.
		• Quick data collection, saving time and resources.
		• Easy implementation, cost-effective.
		
		Weaknesses
		• Without adjustments, the sample will be over representative of people with the time and inclination to answer the survey.
		  ○ Will not represent people who never answer their phones.
		• Higher risk of inaccuracy due to selection, response bias, and non-response error, caused by survey respondents not representing the population of interest.
		  ○ Can be made more representative with adjustments for known variables like age and sex.

Sources of bias

When selecting a sampling design, it is important to keep in mind the types of error that commonly arise in the sampling process. The two most common types of errors are non-response bias and selection bias. Non-response bias is when the individuals who respond to a survey or whom you observe in a study differ significantly from those who do not respond or opted not to participate. Selection bias occurs when the selection of the sample is non-random and some individuals had a higher or lower probability of being selected. When conducting either a census or a survey, non-response bias is the form of bias that is most concerning. High response rates will tend to reduce non-response bias, which is part of the reason to consider a sample over a census.

Because surveys collect information from a subset of a population, it is possible to devote more resources per each individual response relative to the resources that would be available per response when conducting a census. Higher resources per response tends to significantly improve the response rates, consequently helping to reduce non-response bias.

Questionnaire design

Overview

A questionnaire (or form) is one of the survey instruments through which primary data are systematically collected. It is a group of sequence of questions, prompts, or statements designed to obtain information from a respondent. A questionnaire can include open ended questions, close ended questions, or a combination of both. Questionnaires are both critically important to data quality and create an impression on the survey respondents. The goal is to make sure respondents understand what they are being asked and can provide the answers easily in a form that is suitable for subsequent data processing and analysis. This is achieved through thoughtful questionnaire design, where question topics are determined and precisely worded and ordered. SGIGs that are also referencing the *Conceptual Well-being Framework Tool* and the *Indicator Guide & Recommended Indicators Tool* —including the starting point questions they provide— may use these as guidance in the planning, development, and customization of a questionnaire.

Questionnaire best practices

A well designed questionnaire should be efficient, consistent, and respondent-friendly. An efficient questionnaire increases engagement by taking the minimum time necessary for a respondent to complete and reduces costs by making responses easy for data analysts to process. A consistent questionnaire reduces errors by ensuring that respondents and analysts share a clear interpretation of the questions and answers. A respondent-friendly questionnaire considers the experience of survey respondents in its content, layout, and design, and aims to increase engagement and leaves a good impression on survey respondents. For example, a respondent-friendly questionnaire considers the applicability of different questions for respondents living on- or off-lands.

Questionnaires should be as streamlined as possible. Where possible, use images, Likert Scale questions, randomization, and limit qualitative questions. An open-ended text response question counts for about three multiple choice questions depending, of course, on the difficulty of the question. The funnel approach is a survey flow technique that makes the respondent's job easier, suggesting that a questionnaire: follow a logical order, start with broad and general questions that qualify the respondent and introduce the topic, move into more specific questions, finish with general, easy-to-answer questions (like demographics), and keep sensitive information in the middle or towards the end.

Questionnaire layout

When designing a questionnaire, the layout ensures respondents' ease of understanding and completion. Before presenting the questions, include an introduction that may answer some frequently asked questions. For example, the introduction may include project background and purpose, privacy and security information, the consent statement, and details about timelines or incentives. It is important to also thank participants for their time and make sure they understand the value and purpose of their contributions.

Visually, consistency in formatting, effective use of white space, and well-aligned page layouts contribute to the user experience, whether in print or online. Include headers with the questionnaire title, footers with contact information, and branding elements, if applicable, to help maintain a cohesive visual identity. Care should be taken with the use of colour to ensure readability for all respondents, including those with colour vision deficiencies. Usability testing, mobile-friendly design for online surveys, progress indicators, and numbered questions aid in questionnaire effectiveness. Additionally, response validation, clear language, and adherence to accessibility standards

contribute to a well-structured and user-friendly questionnaire layout, ultimately enhancing data quality.

General rules of thumb

- Always use plain and accessible language, at about the Grade 4 to 6 reading level
  - No technical language should be used
  - The more accessible it is, the better

- Ensure all questions are phrased as true questions and not statements

- Every survey has an attention budget
  - When that budget is spent, respondents will leave early or answer haphazardly
  - Each question costs attention; more complex questions cost more

- Consistency comes first
  - Use the same term for something throughout a survey
  - e.g. Don't use “climate change” and “global warming” interchangeably

- Avoid negatives
  - The word NOT is easy to overlook even when highlighted, and processing negatives costs attention

- Be exact
  - Don’t assume knowledge - provide necessary context / detail to answer accurately/consistently. E.g. Use terms like “in the last 7 days” instead of “in a typical week”

- Pre-empt guessing
  - e.g. Use bins like “Between $20,000 and $30,000” for large quantities that people are unlikely to know exactly

- Neutral, unassuming language gets honest responses
  - Avoid double-barrelled and leading questions
  - Do not introduce bias
  - e.g. Ask about “small dogs”, not “small, yappy dogs”

- Use skip patterns to justify assumptions
    - e.g. Ask “Do you use any time management software?”, and if the answer is no, skip over “What time management software do you use?”
    - Provide respondents with the opportunity to skip a question or (better yet) choose a response that indicates the question does not apply to them, they do not know, or prefer not to answer

- Allow non-responses like “Don't know” or “Prefer not to answer”
    - Missing a value for one variable is much better than missing half a survey to frustrated dropout

- Pre-made Likert scales for questions about degrees of agreement or importance are readily available online

- Consider using knowledge testing questions as an alternative to questions about attitudes and perceptions
    - e.g. Ask respondents to translate words or passages into the Indigenous language, or vice versa, to provide an objective measure of language fluency
    - Knowledge testing questions can provide valuable data less prone to respondents’ transient moods or desire to provide socially desirable answers
    - More information on knowledge testing questions can be found in the Indicator Guide & Recommended Indicators Tool

- Take into account the applicability of questions for on-lands and off-lands residents
    - Consider whether all survey respondents have the information or experience they need to answer a particular question and use skip logic if they don’t
    - E.g., Be mindful about asking off-lands residents questions about programs, services, or events that are only available on-lands
    - Define terms that might be ambiguous. For example, define "community" and whether it refers to the community of the SGIG or the respondent’s community of residency

Testing

Effective questionnaire design requires thorough, iterative testing to collect meaningful data. Designers should first self-administer the draft questionnaire, including reading it out loud, to assess flow, clarity, accuracy, and alignment to analysis needs. To enhance data quality and survey effectiveness, each question should then be evaluated from the perspective of multiple individuals with diverse backgrounds and circumstances to reveal insights on question interpretation and barriers to precise responses. Ensure to refine question wording to eliminate confusing language, leading nature, double negatives, repetition, and absolutes. Finally, ensure to review the responses against expected outcomes and refine questions to align with analysis needs.

Think bigger picture	**Think smaller picture**
Length, energy/thought required, relevance to the research question(s), ability to generate tables/chart in the analysis plan, aesthetic, legibility, flow, longevity	• Layout, font size and colour, spelling • Individual question structure - do's and don'ts and checklist

If feasible given time and budget, designers should next test the questionnaire with a small group of data users and target respondents directly to assess flow, clarity, accuracy, and alignment to analysis needs. This will help identify any problems with the questionnaire that were not apparent during the initial review stage before finalizing it and getting it ready for distribution.

Customization

Pre-made questions are readily available and can serve as a useful starting point when designing a questionnaire. However, questionnaires should also be customized to ensure relevance to each SGIG's cultural and community context and priorities. The customization section in the *Indicator Guide & Recommended Indicators Tool* includes further information on customization relevant to questionnaire design.

Privacy

Data collection carries risks to privacy that must be considered carefully. Best practices dictate that data collectors and analysts should be transparent about how information

will be used, limit collection of personal details, restrict data access, and remove identifying details from aggregate data. Data collection requires balancing analysis benefits with privacy protections. All data privacy and protection policies should be aligned with SGIGs' data governance program. See for reference the *Data Governance Framework Tool*.

Example survey consent statement

"Participation is voluntary, and you can consent, decline, or change your mind at any time. All information will be kept strictly private and confidential. The survey results will be available only in the form of anonymous summary tables that include large numbers of survey respondents to prevent personal identification."

Choosing software

When choosing software to conduct a survey or census, there are several important considerations to keep in mind to ensure that the selected tool aligns with the project's specific needs and objectives. Some key considerations include:

- *Survey Complexity*: Some software options are better suited for simple questionnaires with basic question types, while others offer advanced features for complex questionnaires with branching logic, skip patterns, and randomization.
- *Budget*: Some software options offer free plans with limitations, while others require a paid subscription. Consider the cost in relation to the features needed and the value these bring to the project.
- *Data Security*: If you're collecting sensitive information, prioritize data security. Look for software that offers encryption, compliance with data protection regulations, and access controls to protect respondent data.
- *Ease of Use*: Consider the user-friendliness of the software. Will the platform be easy to use for both administrators and respondents? Intuitive interfaces can save time and reduce the risk of errors.
- *Reporting and Analytics*: Evaluate the reporting and analytics capabilities of the software. Is in-depth data analysis required, or should the tool be able to produce any reports? The tool selected should align with analysis and reporting needs.

- *Mobile Responsiveness*: In today's mobile-centric world, make sure the software offers mobile responsiveness so that respondents can complete surveys on various devices.

Here are some software options:

- Qualtrics
- SurveyMonkey Enterprise
- QuestionPro Enterprise
- LimeSurvey (self-hosted)

Example of how to apply these considerations when choosing between softwares

Qualtrics and SurveyMonkey each have their distinct advantages and disadvantages, primarily differing in terms of cost, limitations, survey complexity, and user experience (UX). Qualtrics stands out with its advanced features, making it the go-to choice for organizations with complex survey needs. It excels in managing intricate survey designs and data analytics but may be less budget-friendly for smaller organizations. In contrast, SurveyMonkey offers affordability and user-friendliness, making it ideal for quick and straightforward surveys. However, it is less capable compared to Qualtrics, limiting its suitability for organizations seeking in-depth data analysis or complex survey logic. Ultimately, the choice between these platforms depends on the specific survey requirements, budget constraints, and the desired level of complexity of the surveys. Qualtrics caters to complex projects and advanced analytics, while SurveyMonkey is tailored for simpler surveys and ease of use.

Scenario: Drafting a questionnaire

The decision to conduct a census has sparked significant interest, with numerous staff members putting forth their information needs in the hopes it will be incorporated into the questionnaire. However, given all of these incoming information needs, the team anticipates a challenge in creating an efficient, consistent, and respondent-friendly questionnaire.

To address the multitude of competing interests, the team initiates the process by prioritizing questions related to the previously agreed-upon indicators chosen from the *Indicator Guide & Recommended Indicator List*. This step is crucial, as these indicators form the core of their previously agreed-upon information needs.

The team then establishes a set of guiding principles endorsed by the senior administrator to facilitate the systematic evaluation of other proposed questions and information requirements in a transparent manner. These principles encompass:

1. *Conciseness*: Ensuring that the survey does not exceed 20 minutes.
2. *Durability*: Crafting questions that remain replicable and relevant even 5, 10, and 20 years into the future.
3. *Actionable Benefit*: Ensuring that the questions generate data that inform government activities.
4. *Reportability*: Creating questions that are easy to analyze and likely to be reportable within privacy thresholds.

Applying these guiding principles results in a set of questions that undergoes a thorough review by the senior administrator. Careful consideration is given to the 'rules of thumb' to shape the final questionnaire.

Once the draft questionnaire is loaded into the survey software, the Sulingituk team seeks to test the questionnaire to ensure it meets the four guiding principles and the 'rules of thumb' - and to ensure that their decisions on question selection aren't leaving any major gaps in the eyes of the community. They also want to ensure that the technology, links, and survey structure work as intended.

Given that many staff of the Sulingituk Government are also members, they decide to use this group to test the questionnaire. Senior administration provides one hour for Sulingituk members on staff to complete the questionnaire, and prizes are offered for their participation. Participants are asked to identify any questions that are unclear or uncomfortable to answer. They are also asked to point out any broken links or other

technical snags in the survey. Finally, they are asked to think about whether the questionnaire covers the big things that the community will want to share or talk about with the Sulingituk Government.

Based on the feedback, the Sulingituk team makes a number of changes to the questionnaire:

- Some terminology is adjusted to reflect the community's worldview and common language;
- Additional information is built into the introduction to the questionnaire to explain privacy and confidentiality for all respondents;
- One question is added to reflect a gap in the questionnaire related to child care;
- Some unclear questions are refined to enhance accuracy, reliability, and clarity.

Communication strategy

Key considerations

Effective communication when conducting primary data collection ensures potential survey participants understand the importance of the data collection exercise, how the data will be used and safeguarded, and important timelines. This is the opportunity to connect with potential survey participants, express gratitude for their time and participation, and let them know if you are offering any incentives. The communication strategy may also address concerns, build trust, and maximize participation by providing transparency to participants. When the respondents feel that their input is valued and their concerns are being addressed, they may be more likely to participate fully in the survey. To achieve this, consider these guiding questions while creating the communication strategy:

- What is the primary goal of the data collection efforts and why is it important?
- Who is the target audience for the collected data and what are their information coneeds and preferences?
- What is the level of trust amongst participants in data collection and in the SGIG holding citizen data? How can trust be built or reinforced?

- Who are the opinion leaders in the community and within families that can encourage people to participate?
- Who are trusted messengers or people with large followings that respondents will be connected to?
- What potential barriers or misconceptions may participants have about the data collection and how can the concerns be addressed proactively? What are the messages people need to hear about how the information will be treated?
- How can the messaging be tailored to resonate with the values and interests of participants?
- What are the key events and times of year to target the messaging and audience?
- What are the likely questions people will ask? What feedback mechanisms can be implemented to continually gather input and insights from participants to refine the communication strategy?
- What channels of communication do participants use to obtain information? What content formats are most engaging? (e.g., reports, infographics, videos)

Think of the communication timeline as being composed of multiple waves, each with its own time frame, target responses, and audience. As the project moves through each wave, there is the opportunity to refine the approach based on successes and challenges.

Communications methods

There are many possible communication products to develop and pathways to leverage. These should be chosen based on the answers to the questions in the section above and the key preferences of the target audiences. Options to consider:

- In-person gatherings, particularly for those audiences that prefer face-to-face communications, such as Elders, or through family-based meetings.
- Newsletters, particularly if the SGIG has a regular newspaper or newsletter that has regular community readership.
- Production of a video that describes why this project is important and why people should participate, with messaging delivered by key opinion leaders in the SGIG.

- A strong social media campaign that reflects the platforms used most by the SGIG members (e.g. Facebook). This may include paid advertising where appropriate. Another effective strategy is to identify any members or pages with large followings among the target respondent group, and work with those influencers to promote the project (with appropriate compensation for their time). Partnering with other organizations with large followings among the SGIG's members should be considered to leverage their channels to promote participation.

Various materials are useful to share at community events or door-to-door. This includes materials such as posters, presentation slides, door knockers, and Frequently Asked Questions (FAQs).

Respondent FAQs

When finalizing communications materials, make sure that they individually or collectively address these frequently asked respondent questions:

Purpose of the survey/study

- Why am I being asked to participate in this survey/study?
- What is the goal or objective of this survey?

Confidentiality and privacy

- How will my personal information be protected?
- Will my responses be kept confidential?

Data anonymity

- Will my responses be anonymous?
- How will my data be aggregated and reported?

Data usage

- How will the collected data be used?

Survey length and time commitment

- How long will it take to complete the survey?
- Can I save my progress and come back to it later?

Incentives and compensation

- Will I receive any incentives or compensation for participating?
- How and when will I receive any incentives?

Contact information

- Who can I contact if I have questions or concerns about the survey?
- Is there a dedicated support team for respondents?

Survey deadline

- Is there a deadline for completing the survey?
- Can I request an extension if needed?

**Note that many of these FAQs can be answered in the questionnaire introduction.*

Incentives

Financial incentives are an effective way to increase survey responses. Develop an approach to incentives that aligns with the specific goals of the project and caters to the preferences of the audience. Consider: the types of incentives to offer; how frequently to issue incentives; and, whether there are ways for respondents to increase their chances of winning.

Consider if the target respondent group is:

- Most likely to be motivated by a high chance of winning. If so, consider offering a greater number of smaller prizes on a regular basis, such as a weekly gift card draw for those that have participated.
- Most likely to be motivated by a major large prize. If so, consider offering a top three prize draw at the end of the data collection process, with major incentives (e.g. trips, electronics).
- Motivated by a certain type of incentive unique to the SGIG's culture and community, such as traditional foods or clothing items.

Consider if there are ways to increase the response or completion rates by establishing different levels of incentives, or increasing opportunities for respondents to win. For example, respondents could earn more entries in a prize draw through completing more sections of the survey in a comprehensive way, or earn more entries through referring eligible friends and family members.

Scenario: Engagement strategies

A key reason for Sulingituk Government to undertake a census rather than a survey was that they are interested in hearing from groups that typically do not participate otherwise in community engagement efforts. These groups are challenging to reach, as they feel stigmatized, have low trust in sharing their challenges and needs with the Sulingituk Government, and some have very limited access to technology and no fixed address. The Sulingituk Government develops tailored outreach and engagement strategies to collect accurate representation of these populations:

Friends and family approach

- The Sulingituk team offers prizes to incentivize members to promote the census among their friends and family, and to share updated contact information with the enumeration team.

Local enumerators and enumeration

- Enumerators are recruited from underrepresented groups to ensure familiarity, trust, and cultural sensitivity.
- Enumerator teams travel to different locations and urban centres to provide on-site support to collect data in these areas and from underrepresented groups.
- Census period:
- Recognizing the geographic challenges of reaching members off-lands, the census timeline accounts for ample time for data collection.

Inclusivity and sensitivity

- The census questions are inclusive and respectful of diverse identities, including gender and sexual orientation.
- Enumerators and staff are culturally aware, sensitive, and respectful when interacting with underrepresented populations, having been trained and provided with tools to address concerns and questions effectively.

Enumeration

Enumeration methods

Enumeration refers to the process of collecting information from survey participants, usually through questionnaires or interviews. Having varied enumeration methods allows people to participate in the way most convenient and comfortable for them. Consider providing various options, such as:

- Self-directed online survey
- Phone enumeration appointments where staff enumerate respondents
- Door-to-door
- Events
- Email, phone, social media, text
- Other approaches like mail-in questionnaires

When selecting an enumeration method, SGIGs should consider the trade-offs between respondent preference and cost, data quality, and timely responses. Depending on the project objectives and constraints, SGIGs may find certain methods are more suitable than others. For example, respondents often prefer mail-in questionnaires for their convenience and privacy, as individuals can complete them at their own pace and in the comfort of their own space. However, they are not always recommended due to several reasons, including:

- *Low response rates*: Mail-in questionnaires typically have lower response rates compared to other data collection methods resulting in a less representative sample and may require additional efforts and costs to increase participation.
- *Limited control and lengthy response times*: There is little control over the timing and completion of mail-in questionnaires. Respondents may take longer to complete mail-in questionnaires, leading to delays in obtaining data, which may not be suitable for the projected timeline.
- *Costs*: Printing, postage, and data entry costs can be significant when using mail-in questionnaires, especially for large-scale surveys, possibly straining the budget.
- *Limited interaction and quality of response*: Mail-in questionnaires lack the opportunity for enumerators to clarify questions or provide assistance in real-time, potentially leading to misunderstandings or incomplete responses.

While mail-in questionnaires can be a valuable data collection method in specific contexts, other methods like online surveys, phone interviews, or in-person interviews may be more fitting alternatives for many projects.

Enumeration planning

Recruitment

In the pre-enumeration phase, the first step is to hire enumerators. The hiring process involves developing job descriptions, specifying requirements and qualifications, and outlining the necessary skills. To recruit suitable candidates, the job may be posted on local job boards or obtained through network connections. Conduct interviews to assess the candidates' suitability for the role. SGIGs can consider these questions to shape their enumerator recruitment strategy:

Community knowledge

- How much community knowledge is necessary within enumerators to effectively engage with and understand the target population?
- Are there specific language or cultural competencies required for the enumerators to establish rapport and ensure cultural safety during interactions with respondents?

Community dynamics and representation

- Are there specific dynamics within the community that may be better managed by ensuring a diverse group of enumerators representing different groups within the community?
- Should efforts be made to match enumerators with respondents who have preferences regarding the demographic background of the enumerator they interact with?

Trust factors

- Are there trust issues within the community that may affect the level of trust respondents have in enumerators?
- Are there strategies or considerations to address trust issues, such as engaging community leaders or involving trusted community organizations in the recruitment process?

Community employment opportunities

- Is it important to prioritize hiring enumerators from the community, and what are the potential benefits or challenges associated with this approach?
- How can the recruitment strategy balance the need for community representation with potential issues of trust, and/or the need for specific skills or qualifications?

Leveraging existing staff

- Are there opportunities to utilize existing staff or resources within the organization to support the enumerator workforce, such as training or logistical support?
- How can existing staff be integrated into the recruitment and training process effectively?

Age and gender

- What age and gender mix among enumerators might be most effective for engaging with the target population?
- Are there cultural norms or preferences within the community that should guide the composition of the enumerator team in terms of age and gender?

Onboarding and training

Once the enumerators are selected, proceed with enumerator onboarding and project orientation to ensure they are well-prepared for their tasks. Onboarding and training should include training on privacy and confidentiality. Enumerators must understand the significance of these principles to maintain respondents' trust and uphold data quality. Enumerators must understand that safeguarding respondents' personal information is ethical and necessary for data integrity. Enumerators should also be aware of legal obligations and how to handle confidential information appropriately, including secure data storage practices. They must be instructed not to share or disclose any personal details of respondents with individuals outside the data collection process.

Enumerators should also be trained on the survey instrument, data collection techniques, and any specific protocols or instructions related to the project. In this training, enumerators should be aware of any potential biases, cultural sensitivities, or ethical issues that may arise during data collection and how to handle them appropriately.

Lastly, verify that enumerators have the necessary equipment, such as tablets or paper surveys, and that this equipment is in good working condition.

For large-scale enumeration projects, it is helpful to develop an enumeration guide to provide enumerators the necessary background and protocols to successfully conduct survey fieldwork. The guide should include an overview of the survey background, required materials and resources, and step-by-step guidelines for preparation, contact, introduction, administration, and follow-up. The guide should also detail the required protocols for response capturing, record updating, and data quality. This should include a template for enumerators to consistently log each attempt to complete the survey with an eligible respondent – this log should include who the attempt was with (including whether this respondent is someone on their enumeration list), the date and time of the attempt, how the attempt was made (e.g. door-to-door, by phone), reasons for unsuccessful attempts, and any additional comments or noteworthy observations. More information on enumeration lists and logging attempts is provided in the Survey enumeration section.

Alongside hiring enumerators, SGIGs should appoint or hire someone to oversee enumeration and enumerators. The enumerator supervisor will be responsible for safety management, ensuring the safety of enumerators in the field. This involves conducting risk assessments, establishing safety guidelines, and ensuring enumerators are equipped to handle emergencies. The enumeration supervisor will also coordinate logistics such as planning enumerator deployment and managing transportation and equipment distribution.

Survey enumeration

In the survey enumeration phase, accurate enumeration lists need to be produced and maintained to allow the enumeration supervisor to track the overall progress in logging attempts and completions of eligible respondents. To produce the enumeration lists, the membership list will be cleaned by removing duplicate entries and restructuring the data set. There may be additional information for other sources added, such as from an in-community housing list. The list will then be divided into smaller lists organized by location and the presence of contact information—members without any contact details will be put into a separate list. A template will be utilized to format the lists into a more usable layout for the enumerators. The initial versions of the lists will then be refined based on feedback and input from the enumerators. The lists will be consistently updated and reproduced as surveys are completed in order to track progress.

When assigning lists to enumerators, consider their familiarity with the area, language proficiency, and capacity to manage the workload effectively. The enumeration

supervisor will provide ongoing monitoring to redeploy enumerators, consider new incentive approaches, and make recommendations to increase productivity and the response rate.

Enumerators should log attempts, including details such as the date and time of the attempt, the specific location involved, reasons for unsuccessful attempts, and any additional comments or noteworthy observations. Additionally, enumerators should execute any required follow-up communications, such as sending follow-up emails, and log those as well. Real-time reporting or regular check-ins with the enumeration supervisor is encouraged. Performing daily checks can ensure the completion of surveys and validate participant responses, removing those who decline to participate or provide invalid information. To gauge the overall progress, diligently track response rates and the advancement of enumeration on a weekly basis. This comprehensive approach ensures the smooth execution of the survey distribution process.

Supervisors should monitor progress by tracking the number of completed surveys and identifying areas where enumeration may be falling behind schedule or experiencing excess capacity. To ensure effective data collection, they should be prepared to pivot and redistribute resources, such as reallocating enumerators from areas with lower workloads to those in need or adjusting the deployment schedule as required. Since they will work closely with enumerators, motivating the team fosters a positive work environment, enhancing data quality and efficiency. Careful quality control throughout enumeration provides a strong foundation for analysis and mitigate the risk of inaccuracy.

Tips for efficient door-to-door enumeration logistics

- Ensure membership or citizenship information is up-to-date prior to enumeration
- Review your enumeration list and map households systematically
- Prioritize clusters and dense areas first
- Plan your route in advance for in-person enumeration
- Minimize revisits and travel time
- Choose contact methods based on relationships and proximity
- Be flexible if initial contact fails, trying different methods
- Visit households at convenient times

- Seek commitment or follow-up if the respondent is unavailable
- Consider asking neighbors for advice on availability

Interviewing

Effective interviewing during data collection relies on the enumerator's confidence, listening skills, empathy, and speech. When meeting respondents, the enumerator should introduce themselves and the project courteously, confirm the address if relevant, and ask for the respondent's full name. The enumerator should assure respondents that their information is confidential. Language barriers can be addressed through translation by another household member, so long as the respondent feels comfortable. If respondents seem uncomfortable, the enumerator can emphasize incentives, skipping questions, and the importance of participation. With reluctant respondents, the enumerator can leave contact information to complete online at a different time. Instruct enumerators to keep accurate records of their calls or visits, including date, time, location (if relevant), and any issues encountered during data collection.

Safety

Enumerator safety is a top priority and enumerators should be aware of potential hazards and take steps to avoid and mitigate risks, especially when conducting door-to-door enumeration. Interviews should only occur where the enumerator feels safe and comfortable, including their mental health and safety.

Safety tips for enumerators

- Checking for dogs and avoiding risks
- Conducting most interviews at the door, not entering homes
- Programming supervisor contacts into phones
- Regular check-ins with supervisors
- Avoiding dangerous walking areas
- Do not enumerate door-to-door at night
- Carry Personal Protective Equipment such as masks and hand sanitizer
- Ensuring mental health support and resources

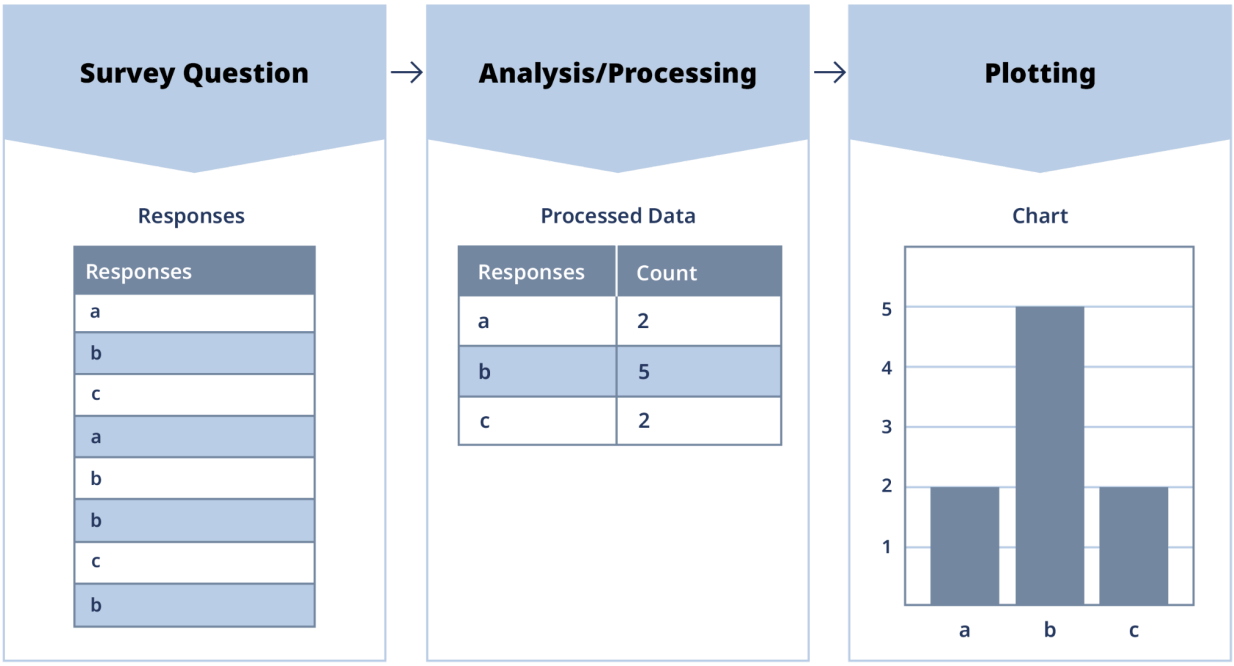
Scenario: Trust building

The Sulingituk team knows that many members have trust issues with sharing their personal data and information with the Sulingituk Government – and this was emphasized strongly by participants in the questionnaire testing phase. Accordingly, Sulingituk implements several measures to address these concerns to build trust and encourage participation in the Census.

- **Training and strict protocols**: Census staff are trained to follow strict protocols to protect respondents' privacy. This training emphasizes building trust with the community and adhering to ethical data collection practices.
- **Informed consent**: The introduction to the questionnaire includes an explanation of all privacy and security protocols, including that individuals' responses will be anonymized and not accessible to Sulingituk Government staff and leadership overall. A consent statement is included for all participants to acknowledge.
- **Communications materials**: Communications materials are prepared, including Frequently Asked Questions and an infographic, specifically to show how individual respondents' privacy and confidentiality are protected.
- **Data steward role:** To mitigate privacy concerns, a Data Steward is appointed to oversee data security. The dataset is secured in encrypted storage, aligning with Sulingituk Government's data governance policies.
- **Policy**: The data collection policy is made public and shared with any interested respondent. This policy describes controls for data access and measures to ensure privacy and confidentiality.
- **Community representation**: The census team includes representatives from the community to build a sense of ownership and trust, including community leaders and trusted individuals that publicly endorse and support the census effort.
- **Accountability and reporting**: Census enumeration progress is regularly reported, building credibility that the Sulingituk Government is taking this effort seriously, and the voices of respondents will be taken into account in future planning.

Processing data

Once the data have been collected and the enumeration phase is closed, analysts should "clean" the data. Data cleaning means identifying and correcting errors and inconsistencies. Then, analysts can move onto data processing, which transforms survey responses obtained during collection into a form that is suitable for tabulation and data analysis.



Data cleaning

Before analyzing data, it is important to ensure that they contain no obvious errors or unusable responses. Some of the most common forms of erroneous survey data are:

- Duplicated entries,
- Incomprehensible responses, and
- Incomplete entries/missing responses.

For example, if a survey is enumerated online then a user may accidentally submit their responses twice, leading to a duplicate entry in the data. It is important to remove

duplicate entries before analysis as otherwise the opinions of the person who accidentally submitted their survey twice will have an inordinately large impact on the results. Duplicated entries are easy to deal with: simply remove one of the copies of the response from the data.

Incomprehensible and missing responses can cause more difficulties. It may be possible to remove such responses from the data, but removing data increases the margin of error and, therefore, leads to less precise results. In many cases it is preferable to impute missing values, i.e. to insert an informed guess of what the missing value should be. The most common form of imputation involves using responses to other, similar questions by the same respondent to guess the most likely response. For example, one could randomly select a response from among those with the same age, gender, and city of residence as the respondent who is missing a response. There are many more complex means of imputing data, with entire textbooks devoted to the subject. One such textbook is freely available online for those interested in the many imputation methods available.

Scenario: Data cleaning

During data processing, the data cleaning team encounters a situation where they discover duplicate names with different birthdates and the use of nicknames that match birth dates but do not align with legal names on the membership list. These discrepancies raise concerns about the accuracy of the data, as the data cleaning team lacks the necessary community knowledge to verify the identity of these individuals in the membership list.

To address this issue, the data cleaning team compiles a list of questionable entries and organizes a meeting with the enumeration team, all of whom are Sulingituk community members. During the meeting, they collaboratively work through the list, confirming the legal names of each person and resolving discrepancies. This effort aims to ensure the accuracy and integrity of the membership data, fostering a more reliable foundation for the analysis phase.

Weighting

It is common, particularly in a census, to not fully reach the target population or sample size. When processing data where the target population was not reached, post-stratification or weighting techniques can be used. Post-stratification techniques adjust survey data to align with the known target population demographics, reducing

bias and improving generalizability. Weighting techniques, which assign differential weights to survey responses based on specific criteria, such as age, gender, and geographic location, aim to ensure results reflect the overall target population.

Weighting survey data is slightly more expensive than using unweighted data. Moreover, it introduces additional uncertainty into the survey results. A survey that is focused on soliciting specific feedback may not need to be weighted because it is safe to assume that people who feel strongly about the subject are quite likely to respond. Nonetheless, there are many cases in which weighting is worthwhile.

When to weight data

- Response rates are much higher in one demographic group than in others.
    - Suppose two-thirds of respondidents are over the age of 50 but more than half of those being surveyed are less than 50 years old. Weighting data by age ensures that the views of younger people are represented fairly.
- A census is being conducted and less than 99% of members respond.
    - Even if most members respond, a census is meant to represent the entire population, including those who did not respond.
- A reliable and complete source of data that overlaps with survey responses is available.
    - Ultimately, weighting relies upon comparing survey data to another data source. If it is not possible to compare demographic responses in the survey to an alternative, trusted source of demographic data then it is not possible to apply weights.

Analysis

Data analysis involves summarizing the data and interpreting their meaning in a way that provides clear answers to the questions that initiated the survey. Often, it consists of interpreting tables and various summary measures, such as frequencies, means and ranges, or more complex statistics when relevant.

The analysis of the data should follow an analysis plan. The analysis plan is produced prior to enumeration and shows how the data generated by the survey will meet the information needs. It includes how the data generated by the survey will be processed, which relationships will be examined, which statistical methods will be used to examine those relationships, what criteria will be used to interpret the results, and how the results will be reported. A successful analysis plan ensures that each aspect of the survey works together to meet the objectives of the survey: that the variables used in the survey meet the needs the analysis, that the planned survey sample will meet the needs of the statistical methods, that the outputs of the statistical methods have objective criteria they can be evaluated against, and that the results will meet the information needs of the survey.

As part of the analysis plan, drafting mock-ups tables that show how the final results will be reported is a good step to focus the team on the results that will be generated. These tables serve as a visual representation of the expected outcomes and provide a clear structure for presenting the survey findings. They can be designed to display the key variables and relationships that will be examined in the analysis, as well as any subgroup or demographic breakdowns that will be included. The mock-up tables also help to ensure that the planned analyses will indeed produce the desired information, and they can be used to identify any additional data processing or analysis steps that may be needed. By visualizing the end product of the analysis in this way, the survey designers can work more efficiently and effectively towards the survey objectives, and stakeholders can have a clearer understanding of what to expect from the survey results.

In general, the development of the analysis plan and the data analysis itself should be conducted by an external service provider that specializes in data analysis. Ideally, this service provider will have been present throughout the process and so would have guided other parts of the project to ensure consistency all the way through to data analysis. There are certainly instances when SGIGs will have internal capacity to analyze the data, but even then it's not a bad idea to have external support to validate the approach taken to analyzing the data.

What kinds of questions should the analysis plan seek to answer?

- Which tables, graphs, tabs, & charts would be of highest value?

- Does the design of your research project actually answer what you set out to understand?
- Are there relevant comparison populations?
- What context is important for understanding this result?
- Does the presentation of the data/analysis articulate the most important information?

Scenario: Weighting

While analyzing the census results, a situation arises where the youth population significantly outperformed other age groups in terms of census participation. This has implications for data analysis, policy formulation, and community engagement. The overrepresentation of youth is linked to their active use of social media, where census-related posts garnered attention and motivated many young residents to participate. This now affects indicators related to education, employment, and substance use, because results are skewed due to their high participation.

Accordingly, the Sulingituk Government has to deliberate between a report that genuinely reflects the voices of those who participated and one that accurately represents the entire population, including those who did not participate. In other words, does the Sulingituk Government want to reflect the voices of participants specifically, or to ensure a comprehensive portrayal of the entire membership?

The Sulingituk Government goes back to the core purpose of their census – to obtain data to inform planning, decision-making, and budgeting. From this perspective, they recognize that a representative report would provide a more wholistic and representative understanding crucial for effective planning, policy formulation, and resource allocation.

To achieve this, they decide to implement a weighting strategy. This involves assigning different weights to the responses from various age groups to account for the overrepresentation of youth. This approach helps ensure that the report represents the voices and needs of all age groups and is, from their perspective, a fair and inclusive approach to take in actioning the findings.

Reporting

Reporting is about taking all of the work that's been done and making it *useful* for different audiences. Importantly, reporting is one of the major cost-drivers of the project overall, so understanding your reporting requirements and then translating those into a set of products that meets those requirements and budget is key. The initial set of reports that will be generated from the dataset should largely be determined by the work done earlier in the project timeline. The questionnaire was developed with a specific set of statistics in mind, the impetus for the project overall would have been clearly articulated, and the government's specific priorities that lead to the desire to conduct the data collection exercise in the first place should all be answered in an initial set of reporting products. That said, new reporting requirements will invariably emerge through the course of the project. Those new reporting requirements should be scoped through a planning process. The planning process does not need to be too involved, but at a minimum, the following questions should be answered:

- What are the key objectives and intended uses for reporting?
- Who are the target audiences and what formats are likely to resonate most?
- What are the main insights and takeaways to communicate?
- What level of detail is appropriate for different audiences?
- How will respondent privacy be protected in public reporting?
- What is the budget and timeline that can be allocated to this product?

Reporting products end up being the most substantive legacy of the entire project. When people talk about the project in the future, they'll be thinking about the reporting products that were generated.

The following are a set of reporting products to consider:

Summary reporting

Written report

Usually, the form of the final report will be confirmed in the course of the project. It could include one final major report composed of subsections that can stand alone, multiple shorter summary reports concisely compiling key findings on major topics into brief

documents, or some combination of the two. Regardless of the form of the final report, topic-based summaries are useful tools for gathering insights into findings on core topics important to the SGIG. Summary reports can provide vital evidence to support funding requests, development projects, and supportive programming across departments. For instance, these summaries could cover topic areas such as cultural vitality and identity, governance and jurisdiction, health and wellness, labour market indicators, housing, and environmental stewardship, depending on the content of the questionnaire and interest.

Example of a Written Report Page

Education

Figure 1.1 shows e-learning and in-person class preferences by age group. Younger individuals (18 to 24 years old) are more likely to prefer e-learning (50%), with a smaller portion favouring in-person classes (40%). For the 40 to 55 age group, the preference shifts significantly towards in-person education (70%), with a relatively smaller percentage embracing e-learning (35%). Interestingly, a substantial number of the older cohort (65%) shows a preference for blended learning options that combine both in-person and online modalities.

Figure 1.1: E-Learning vs. In-Person Class Preferences by Age Group

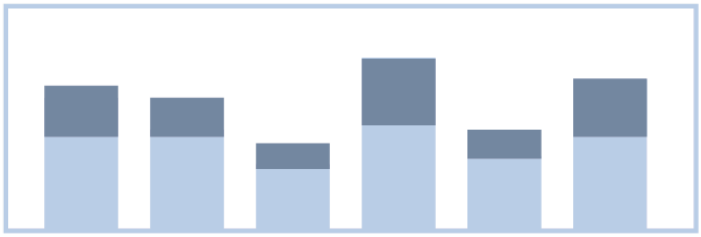


Figure notes: Question: 'Which type of learning environment do you prefer?' The chart reflects preferences among 250 respondents.

Analysis

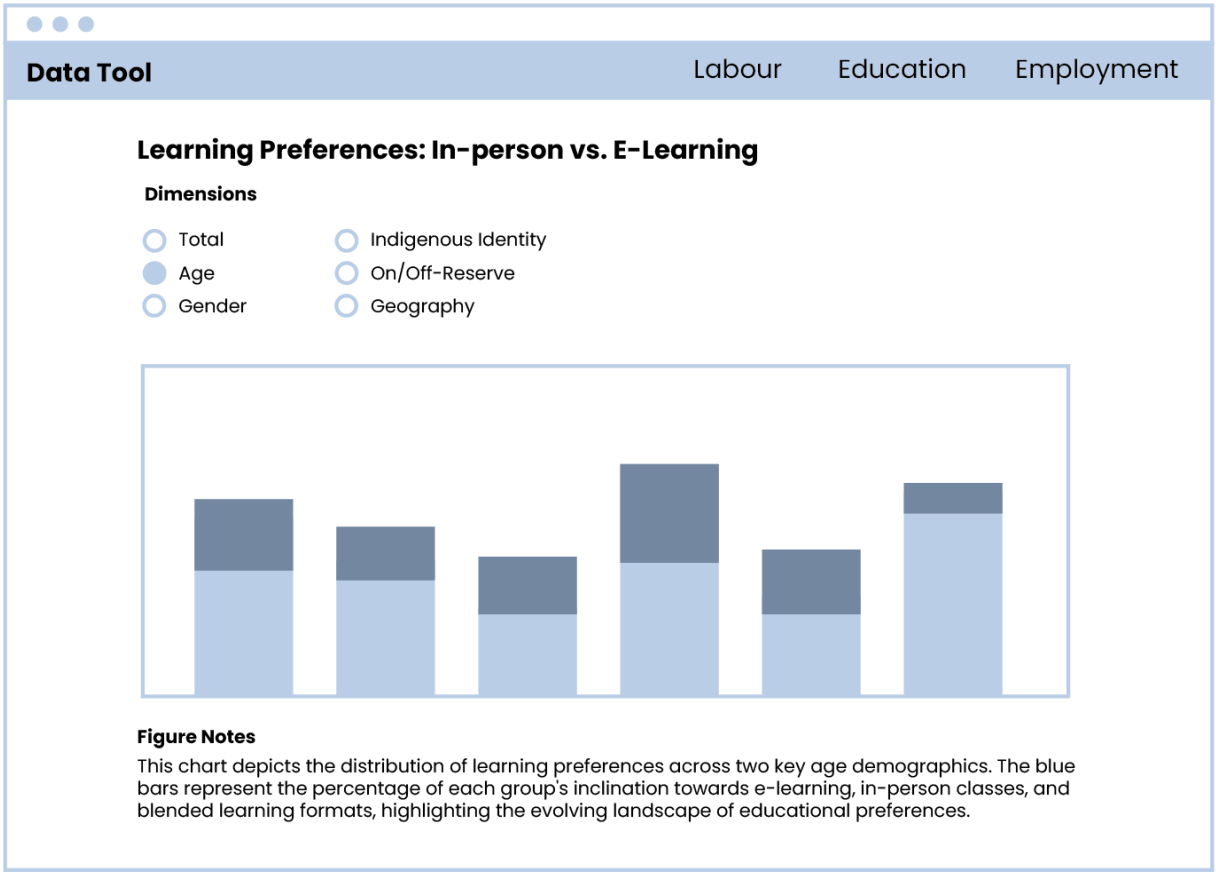
Survey results highlight a digital divide in learning preferences, with 50% of 18-24-year-olds gravitating towards e-learning for its flexibility, while the 40-55 age group remains rooted in in-person learning, underscoring an ongoing generation gap. Yet, a notable 65% of older learners show curiosity in blended educational courses, combining adaptive digital solutions with traditional methods.

Interactive web tools

Immersive and interactive web-based tools are a way to present findings in a visually appealing and user-friendly manner. These tools can provide custom insights and allow users to explore specific indicators through charts, graphs, gauges, maps, and filters. Some of the interactive web tools that can be produced using collected data to support SGIG priorities are:

Data visualization tool: A public-facing website designed to promote transparency and provide a comprehensive overview of major project takeaways, including statistical snapshots and high-level trends and comparisons. An interactive dashboard can allow visitors to filter demographic dimensions, export, or share personalized views. *Note that the SGIGs, through this tool-development process have access to a data visualization tool that can be used to present their data.*

Example of a Data Visualization Tool Screen



Internal analytics portal: A business intelligence platform providing staff with powerful tools to quickly answer research questions related to key planning needs. The platform would allow staff to gain insights into the SGIG's needs and priorities, identify service gaps, and make data-driven decisions.

Ad hoc data requests

Custom data tables and analyses are often ideal for addressing specialized information needs that emerge during and after the project. A streamlined "self-serve" request mechanism can be developed, such as an online intake form or template, in which authorized government staff can submit specific questions or data they require. Analysts would then review the inquiry, follow up with the requester to clarify details, pull the relevant datasets, perform additional analysis if applicable, and deliver the output in the appropriate format.

The deliverables could range from basic cross-tabulation summaries exported into Excel to more advanced integrated data visualizations and reports. While more complex ad hoc requests may require deeper research paired with the existing data inputs, creating a centralized process can facilitate rapid turnaround to address emergent needs.

Knowledge transfer and capacity building

Knowledge Transfer Workshops or similar capacity-building tools can support SGIG staff to increase their capacity to derive insights from the project continuously. The workshops could include detailed discussions and practical activities, including finding and interpreting key indicators, using analysis software tools, and applying insights to program planning.

The knowledge transfer workshops could also focus on practical applications of the insights gained from the project. Attendees can learn how to use the insights to inform program planning and decision-making processes. The knowledge transfer workshops aim to provide SGIG staff with the necessary skills and knowledge to use data for program planning and decision-making.

Scenario: Reporting

At the beginning of the census project, the Sulingituk Government chose a set of key health and wellness indicators, which included asking respondents about their rates of substance use. The data showed elevated rates of substance use among older adult males. Sulingituk Government wants to report these data to the community, but does not want to reinforce stereotypes that perpetuate racism, or create stigma that disincentivizes people from seeking help.

Therefore, in developing the analysis, interpretation, and reporting on these results, Sulingituk Government:

- Presents clearly the rates of substance use among all age groups in a factual manner.
- Describes the broader historical, social, and cultural context surrounding Indigenous peoples and substance use, and ongoing stereotypes in this respect.
- Seeks expertise from addiction and substance abuse professionals to ensure that the representation of the data is factual and non-stigmatizing.
- Includes a description of support services available in the community, and how to access those services.
- Describes the intention for using these data to tailor programs to the demographics in highest need for those programs.
- Invites older adult males who use substances to review the draft report, incorporating their views and perspectives prior to releasing the report to the community.

External support and RFP guidelines

Should you recruit external support?

When considering the hiring of external support for primary data collection, several factors come into play, including the project's scope and objectives, the time and capacity of any in-house expertise, budget, project timelines, and the existence and availability of trusted service providers, among others. Primary data collection is made up of several activities, and it's possible to go to market on some, but not all, of these activities, using one or several service providers:

- Scoping/Statement of objectives
- Survey design and analysis plan
- Questionnaire design
- Communications strategy
- Survey enumeration
- Processing, analysis, and reporting
- Specialized products, including websites or dashboards, and reporting, including audio, video, events, or other communication formats

Drafting a Request for Proposals

When drafting a Request for Proposals (RFP), make sure to include:

Objectives

- Define the project's objectives.
- Explain the desired outcomes and benefits.

Background

- Describe the project's background and context.
- Explain the project's strategic importance.

Main activities and timeframe

- Detail the main activities or tasks that need to be completed as part of the project.
- Specify the project's timeline, including start and end dates, as well as any key milestones.

Deliverables

- Clearly list the expected deliverables.
- Include a description of each deliverable's purpose and format.

Mandatory criteria

- Outline the mandatory requirements that all proposals must meet.
- Specify any qualifications, certifications, or licenses required from potential service providers
- Define the criteria that will be used to evaluate service provider proposals and indicate the relative importance of each criterion.
- Specify the qualifications and experience expected from potential service providers .
- Request references or case studies from previous projects.

Budget and pricing

- Clearly state the budget constraints or expectations.
- Request a breakdown of pricing, including any recurring costs.

Terms and conditions

- Include any legal or contractual terms and conditions that service providers must adhere to.
- Address intellectual property rights, confidentiality, and termination clauses.

Evaluating proposals

SGIGs may follow their pre-existing evaluation process for RFPs while considering the specific requirements for data collection. This includes evaluating the proposed methods, including sampling techniques, data collection tools, and data quality measures. Given the sensitivity of data in many projects, SGIGs should inquire about the proposed data security measures. Questions should be raised regarding how data privacy and confidentiality will be maintained throughout the data lifecycle, from collection to analysis and storage, as well as data ownership.

Similarly, assess the proposed risk mitigation strategies and contingency plans. SGIGs may also inquire about the proposed data quality assurance processes, including data cleaning, validation, and verification. Understand how potential biases will be addressed

and how data integrity will be maintained throughout the analysis. Regarding the interview process, questions might revolve around the experience and expertise of the proposing team, their track record in similar projects, and their approach to handling potential data collection and analysis challenges.

Example interview questions:

- What is your approach to pre-testing your questionnaires? What do you look for in your review?
- What software or tools do you use for data collection and data analysis, and why do you choose these particular tools?
- What measures do you have in place to ensure the quality and integrity of the data throughout data collection and analysis?
- How do you mitigate the risk of low response rates?
- How do you handle and store data securely during the project?