

# data

## HOW TO COLLECT IT AND WHAT TO DO WITH IT

a tool by

**QUAKE  
LAB**

&



**OLIP**

OTTAWA LOCAL  
IMMIGRATION PARTNERSHIP



# table of contents

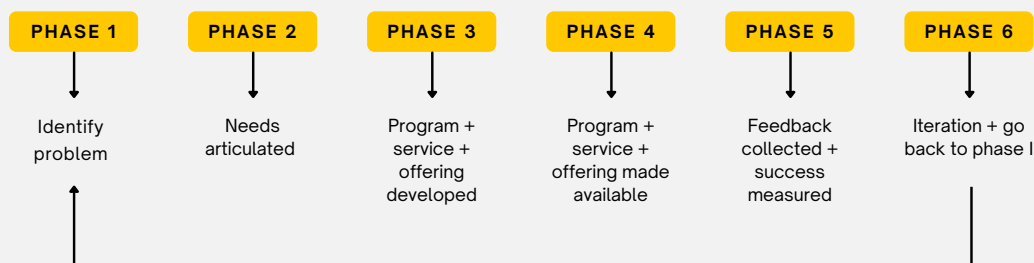
- 2 HOW TO USE THIS RESOURCE
- 3 DEFINITIONS
- 4 SO YOU WANT TO COLLECT DATA
- 7 COLLECTING NEW DATA
- 12 ANALYZING EXISTING DATA

# how to use this resource

This guide is designed for staff and decision makers within the service delivery sector a starting point as they begin to think about data collection, a companion piece as they move through data collection processes, and a reference for ethical data practices when handling data sets from previous collection processes. It familiarizes the reader with key terminology for navigating the resource itself and terminology needed for navigating their wider understanding of community data collection.

The service planning and delivery cycle comprises the following phases:

## service planning and delivery lifecycle




Data collection is the bridge between the first and final stages of this cycle. Good data collection processes allow for service organizations to effectively assess the needs of their service users, and the gaps in their service delivery. Placing us firmly at **Phase 1: Identify the problem** and **Phase 2: Needs Articulated**. These same processes are necessary when collecting feedback and assessing whether interventions, programming, funding, and general organizational decisions effectively met your service users' needs. This puts us back at **Phase 5: Feedback collected and success measured**. And as you iterate and improve upon your service delivery you will move back to phase 1, aided by the data you collected as you moved through the cycle.



# definitions

Before getting started, review the following definitions to ensure you have an understanding of how they will be used throughout this document.

Definitions that you can find on this page will be highlighted throughout the document in [blue text](#).

- **Data Mining** is the act or process of sorting through large data sets to find patterns or relationships between variables. Data mining becomes unethical when the individuals whose data is being used have not consented to their data being used for and/or have not been given information about its use.
  - **Collect:** when we speak about collecting data in this resource we are referring to a wide range of techniques from surveys, service use monitoring, questionnaires, interviews, and focus groups. Be sure to read the "Community Engagement Framework" tool for a fulsome understanding of community engagement considerations.
  - **Analyze:** Data analyses refers to the process of cleaning, transforming and inspecting data and putting variables in conversation with each other to find information and draw conclusions.
  - **Present:** Data presentation encompasses the sharing of conclusions, visual representations and cross analysis taken from data collection.
  - **Expertise:** Within this resource expertise will refer to the capacity or skill to collect, analyze, and present data and familiarity with the tools and software needed to do these things.
  - **Double Barreled Question:** A type of question that offers only one answer or solution, while touching upon multiple issues.
  - **Statistical significance:** The determination and the likelihood that the relationship between two variables is caused by something other than chance.
  - **Spurious relationships:** The occurrence of two or more variables being related (correlated) without one causing the other(s). In cases of spurious relationships the variables are likely related through the existence of a third unknown variable.
- 

# so you want to collect data

The background features a light green and white color scheme. On the left, there are stylized, overlapping hand shapes in shades of green. On the right, there are several large, stylized arrows pointing to the right in white, green, and blue. The main title 'so you want to collect data' is in a bold, orange, sans-serif font.

As service providers, you know the importance of collecting good data and having decisions based on reliable data - but we can sometimes jump ahead to data collection before evaluating our internal capacity for the project of data collection.

## BEFORE COLLECTING DATA ASK:

Do we have the internal resources to collect, analyze and present data?

**NO, We don't have funding to hire a data expert or increase staff hours but we have staff with expertise on data collection, analysis, and presentation - and we would be willing to support them in providing them with the tools they need for this project.**

It's great that you have the internal expertise for data collection. It's going to be important to consider your organization's overall capacity when adding this to the responsibilities of a staff members

Consider how much data you want to collect, how robust the data set will be, and how frequently you will be analyzing the data to ensure that you are properly supporting your staff member as they take on a new responsibility. This may look like shifting/easing some of their existing workload to support this project or offering support in the existing scope of their role.

**Internal resources can refer to:** funding, staff hours, software, and/or expertise) to collect, analyze and present data?

Also consider what this responsibility will look like in your organization's succession planning? Will this responsibility be added to a staff role moving forward? Will you ensure that whoever takes this role in the future has the same expertise and is being properly compensated for this work?

**YES, We have the budget to hire a data expert but, we don't have internal expertise, tools, or staff hours**

That's amazing! We recommend reading through the entirety of this resource and recording your answers to the questions and prompts. It will help you find a data expert who can respond to the needs of your organization.

**YES, We have funding to hire a contractor however we would like to have the internal capacity for data collection even though we don't currently have the tools or expertise.**

That's amazing! We recommend reading through the entirety of this resource and recording your answers to the questions and prompts. It will help you find a data expert that can respond to the needs of your organization.

Furthermore, we would recommend providing your employees with professional development opportunities so they can develop the skills needed to collect, analyze, and present data. These skills aren't developed overnight, but investing in the upskilling of your staff will mean that you will be able to do this work internally in the future.

Work with your employees and your contracted data expert to create a data collection guide for your organization so that the basics of your data collection goals, best practices, preferred tools, and consistent questions for year to year reporting are codified as part of your organization's institutional memory.

**NO, We don't have any of these internal resources - but we know data collection is very important.**

It's great that you value data collection! However, collecting data without the resources to do so or a working understanding of data ethics and principles can lead to data being misused. As service providers, many of your clients come from communities who have been harmed by the abuse and misuse of their data. We owe it to them to ensure that we are fully cognizant of the risk that comes with sharing this information so that trust is not lost.

We recommend applying to grants and funding opportunities to specifically collect, analyze and present data AND working with organizations within this community of practice to support Ottawa-wide data collection efforts in areas where your service provision and reporting needs overlap.





# collecting new data

The background features a light green and white color scheme. On the left, there are faint, stylized silhouettes of human figures. On the right, there are three large, overlapping arrows pointing to the right: a white one at the top, a green one in the middle, and an orange one at the bottom.

## STEP 1: HAVE A PLAN FOR YOUR DATA

Not all data is used for the same thing, and knowing the purpose of your data collection will help you in crafting your methodology, hiring external support, getting the appropriate training for your staff and communicating the purpose of the data collection to respondents.

### **1** Year over year reporting

This form of data collection is useful for internal and external reporting as well as reporting to funders. It is often largely quantitative and focuses on the numbers behind resource allocation, service use, spending, and overall demographics of service users.

For internal reporting, the analysis from this data will be shared with your staff, board, and funders. For external reporting, this information will be shared through your website, annual reports, strategies, and any community town halls.

When collecting this type of data you will need to ensure that you are:

- Asking the same questions from year to year to track changes and trends, and are able to create data visualizations
- Being rigorous about maintaining the privacy of your service users
- Ensure that you are communicating the uses of the data to all respondents



## **2 Gaps assessment**

This form of data collection is very useful in identifying areas of improvement. You may use a gap assessment to collect data from your staff and service users. This data will likely be a substantial mix of qualitative and quantitative data collection as you collect sentiment/experience data as well as intersectional demographic data. When collecting this data it is important to note that it works in the comparative and you need to know what you are measuring yourself against.

Are you:

- Setting goals against industry standards?
- Doing proportional analysis against Ottawa demographic data?
- Doing proportional analysis against demographics of service users?

## **3 Community-based research**

This form of data collection is substantially different from the two listed above, because it means that you are collecting data with the intention of supporting a hypothesis or disproving an existing theory. Prior to collecting data for the purposes of community-based research you should:

- have a list of hypotheses you are testing to ensure that you aren't data mining after the data has been collected, and;
- communicate the hypotheses that you are testing to the respondents.

IF you are drawing conclusions for the purpose of community-based research from your data but the respondents think that it is for a gap assessment or year over year reporting you must begin a new data collection process where you disclose the intended use of the data.

## STEP 2: CRAFTING YOUR DATA COLLECTION PROCESS

Once you know what your data is being used for, you can craft your data collection process according to your needs. Keep these steps in mind as you begin creating data collection process:

- 1 Only collect the data that you need.** Adding unnecessary questions makes responding to data collection tedious. Furthermore, it makes your data set unnecessarily bulky, harder to pull key insights from, and more likely to encourage unethical data mining.
- 2 Think carefully about question design.** Ensure that the question format supports the type of data you want. For example: if you are trying to collect quantitative data then you don't want to ask the question in paragraph format.
- 3 Be mindful of word choice.** Explain any concepts that are unclear or have multiple definitions. Think about who the survey/questionnaire is for. Are you using language or words that are familiar and accessible to them?
- 4 Be mindful of the overall accessibility of your data collection tools.** How are you collecting the data? Is it through an online survey or questionnaire? Can respondents save their answers to return to the survey later? Is it through a program that is unfamiliar to your service users? Does the survey or questionnaire need to be translated? Do you have mechanisms of ensuring that the data collection is available to service users of various digital literacies, learning styles, learning disabilities etc.? Is the survey and/or survey questions screen-reader friendly?
- 5 Don't make your questions leading.** Be mindful of tone or creating double barreled questions. The goal of data collection is to ensure that you are collecting trustworthy data that will ultimately inform the decisions within your organization. We do a disservice to our mandate when we collect untrustworthy data.

**6**

**Consider the complexity of the data set you are creating, don't over complicate your questions.** This is particularly important to consider when thinking about the internal capacity of your organization. Think about the employee hours you have committed towards data analysis, the data knowledge of your employees and/or the funds you have available for any external contractors will determine how many questions you ask, the type of questions you ask, and how many avenues for cross-sectional analysis you make available.

**7**

**Build transparency into your data collection.** Communicate to respondents to the usage of the data, the expectations of privacy, anonymization, and your policy around data storage. This information should be available at the beginning of your survey and at the end, along with direct contact information for a staff member at your organization.

### **STEP 3: PRACTICE ETHICAL DATA COLLECTION AND STORAGE**

Once you know what your data is being used for, you can craft your data collection process according to your needs. Keep these steps in mind as you begin creating data collection process:

**1**

**Collect your data anonymously.**

Do not ask your respondents to provide identifying information. This includes, but is not limited to, first or last name, address, email address, and ID numbers such as employee number or health card number.

**2**

**Store your data securely.**

Best practices would dictate that you save your database to a password-protected computer or secure cloud storage location. Limit access to the database by encrypting it with a password or managing access rights, ensuring that the fewest number of people required can access it.

### 3 **Share only what is appropriate.**

You may want to share the research results with a wide range of people for a wide range of reasons. When discussing the research with the key employees directly involved, you can share the raw data and all of the insights during the analysis and reporting process. When sharing the results with the staff as a whole, or when sharing the results with stakeholders (such as board members or funders), never share the raw data; talk about the findings in terms of percentages and proportions. When sharing the results with the public, never share the raw data; data visualizations like charts and graphs can provide a general overview of the findings.

### 4 **Beware accidental de-anonymization.**

Although the data collected is anonymous, DEI research involves the investigation of minority groups in the workplace. If there are sufficiently few people that belong to a particular equity-deserving group, and you report findings for that group, you may accidentally de-anonymize someone's data.

**For example,** if you publish the finding: "100% of Indigenous employees report that the CEO demonstrates racist tendencies" but there is only one Indigenous person at the organization, the staff will know who provided this response.

If your minority groups are so small that there are only a few people in each category, consider merging them together and speaking in more broad terms : "40% of employees from minority racial backgrounds report that the CEO demonstrates racist tendencies"



# analyzing existing data



## STEP 1

### **Are you able to analyze your existing data?**

You may have a large dataset that you are hoping to pull information from, but you have no idea how. First ask yourself the following questions:

- Do you have the time and internal skill set to analyze the dataset?
- Do you have the funds to hire an external contractor to analyze the dataset?
- If you are using an external contractor, are you able to anonymize the data set before handing it off? Have you ensured that the contractor will sign an NDA to protect your respondents?
- Do you have consent to use the data?
- How was the data collected and under what pretense

## STEP 2

### **Have a plan for your analysis.**

Having access to existing data sets or data sets that were collected previously can provide valuable information about past practices and service provision. However, it can be easy to data mine and find spurious relationships IF you do not have a clear understanding of what the data is being used for and whether its collection method supports your organization's current needs.

*See page 8 for more on having a plan for your data.*

## STEP 3

### **Practice ethical data reporting and storage.**

Having an organizational understanding of ethical data reporting and storage is essential even when handling old data sets and ultimately impacts the trust the communities you serve have in future data collection processes.

*See page 11 for more on reporting and storage*