

Improving Data Quality

Why is data quality important?

Victim service providers can use good-quality program data (or information) to increase their effectiveness in improving the lives of crime victims and their communities. Unfortunately, poor quality program data is of limited use and can even hamper program effectiveness – if the errors or omissions in data entry skew the results of data analysis and do not accurately reflect the program’s work.

For example, if a program collects clients’ ages, but ages are missing for 30 percent of clients, the program might be under- or overestimating the proportion of young or older adults served, information that can affect outreach, planning, and funding.

Leaders can lay the foundation for good data quality by ensuring all staff understand the value of the data collected and how it can be used to provide insights into their work. (For more, see CVR’s short animation, [Make the Most of Your Data.](#))

How do we examine data quality?

Using the following simple steps, a program can regularly review and improve its own data quality.¹

Step 1: Assemble the most recent 100 data records used or created for some part of your work. This could be client intake data, satisfaction survey results, or other administrative records. Focus on 10 to 15 critical data elements within these records that your team most relies upon for operations (e.g., client age, service need, satisfaction response). Place these 10-15 data elements for these records in a table or spreadsheet, so the quality of this information can easily be reviewed.

Step 2: Convene a two-hour meeting with several people with knowledge of the data, people who understand how the data is collected or used or who will use this data in their work.

Step 3: Working record by record, highlight obvious errors—impossible ages, incorrectly formatted dates, information in the wrong column, or missing data. Do this quickly, spending about 30 seconds per record.

Step 4: Summarize the results by including a “perfect record” column on your spreadsheet. Put YES if there were no errors in the record across the selected data elements reviewed or a NO if you found errors. Add up the number of accurate records after your review.

After all records have been reviewed, discuss the results. Overall, what portion of records were perfect? What data points had the most errors? How might those errors limit the value of your data? What might be some common causes of errors, and what strategies could you use to reduce these errors moving forward?

¹Redman, T.C. (2016). Assess Whether You Have a Data Quality Problem. Harvard Business Review. <https://hbr.org/2016/07/assess-whether-you-have-a-data-quality-problem>.

How do we improve data quality?

Once you have a sense of your data quality, you can take some organization-wide actions to improve consistency and quality.

- Share results with all program staff. For example, let staff know that *“Two-thirds of our cases are missing important data,”* or *“We have no information on the original victimization for 25 percent of people we serve.”* Keep staff updated on your improvements over time.
- Create a data dictionary, or written guide to data entry. This document should clarify what is meant for each data element. For example, indicate whether *“Start of Service”* is meant to capture the date of the first contact with a victim, the date an intake form is completed, or the date the first service was provided. The more precise your data dictionary, the more consistent your data will be. Update your dictionary as issues arise.
- Limit opportunities for error. Spreadsheet programs (e.g., Excel, Google Sheets) allow you to set *“data validation rules”* for the range of responses, which can limit data entry errors. For example, you can require entries in a certain column to be formatted as dates (e.g., MM/DD/YYYY). You can also create a simple drop-down menu of preset choices for a column, such as *“Requested,” “In Process,”* or *“Completed.”* Search for *“data validation”* or *“create a drop-down list”* in your spreadsheet program’s Help feature to learn how to set these rules.
- Train staff on data collection. Make sure your staff are familiar with the data dictionary and expectations for timely and accurate data entry. Designate someone to be available to answer questions and support staff on data quality issues.

Note: if you are using OVC’s Performance Measurement Tool, start with the [Performance Measure Dictionary and Terminology Resource](#).

Improving data quality using these simple steps will pay dividends over time—helping better inform your victim services, supporting funding requests, and enabling program staff to see the benefits of their work with victims and survivors.

Additional Resources



The Urban Institute’s [Introduction to Data Quality](#) provides a helpful guide for collecting reliable program data.

Find us online at VictimResearch.org [@VictimResearch](https://twitter.com/VictimResearch) [@CenterVictimResearch](https://facebook.com/CenterVictimResearch)