

Analyzing and interpreting data

Evaluation resources from Wilder Research

Once data are collected, the next step is to analyze the data. A plan for analyzing your data should be developed well before it is time to conduct analysis. The best time to develop your analysis plan is when you are first identifying your key evaluation questions and determining how you will collect the needed information. It's important to match the analysis strategy to the type of information that you have and the kinds of evaluation questions you are trying to answer.

Organizing data

Depending on the scope of your evaluation and your available resources, you may need to organize the data before you can begin analysis. A database or spreadsheet can be helpful in organizing your data. Readily available computer programs, such as Excel and Access, may be useful. Excel tends to be easily accessible for most people who have access to a computer with Microsoft products.

Other software is available to help you analyze both quantitative and qualitative evaluation results, such as SPSS or Atlas-TI. Some of this software is expensive, however, and you may be able to analyze your findings without it. Before investing in this software, consider seeking outside consultation to determine if it is needed.

If you create an electronic database, it is important to be thoughtful about how the data is organized. Decisions made as you design and begin to enter information will influence how easy or difficult it will be for you to analyze your results. Tips for designing your database include:

- Assign a unique identifier to each individual in your dataset.
- Include all information about an individual in one row of your database, rather than having the same person appear in multiple places.
- Limit responses so that incorrect information cannot be entered (such as not allowing numbers that fall outside of your response choices).
- Code text responses into numerical form so that they are easier to analyze (e.g., 1=Yes, 2=No).
- Enter data in a consistent format, such as always using a "1" to reflect female gender, rather than using various labels (e.g., "F," "female," "girl," etc.)

Analyzing quantitative data

Quantitative data is information you collect in numerical form, such as rating scales or documented frequency of specific behaviors. For example, typically, close-ended survey questions are coded into numbers so they can be analyzed quantitatively. While statistical analysis of quantitative information can be quite complex, some relatively simple techniques can provide useful information.

Descriptive statistics can help summarize your data and identify key findings while inferential analyses can help you draw conclusions about your results. Descriptive analysis is used to reduce your raw data down to an understandable level. Common methods include:

- *Frequency distribution*: tables or charts that show how many of your evaluation participants fall into various categories of interest.
- *Central tendency*: the number that best represents the “typical score,” such as the mode (the number or category that appears most frequently), median (the number in the exact middle of the data set), and mean (the arithmetic average of your numbers).
- *Variability*: amount of variation or disagreement in your results. Common measures of variability include range (difference between the highest and lowest scores) and standard deviation (a more complicated calculation based on a comparison of each score to the average).

Mean, median, mode, and range

The mean, median and mode are measures of central tendency, while the range shows how your responses vary from one end of the spectrum to the other. Typically, there will not be a need to figure these out by hand, since database programs such as Excel, SPSS and SAS already have functions to calculate them.

The "mean" is simply the "average." It is calculated by adding up all of the numbers in your data and dividing by the total number of numbers.

Example

Dataset: 1, 9, 5, 6, 9

$$\text{Mean} = \frac{\text{Sum of data}}{\text{Total \# of numbers}} = \frac{1 + 9 + 5 + 6 + 9}{5} = \frac{30}{5} = 6$$

The mean for this dataset is 6.

The "median" is the "middle" value of your data. To obtain the median, you must first organize your data in numerical order. In the event you have an even number of responses, the median is the mean of the middle two values.

Example

Dataset: 1, 9, 5, 6, 9

- 1) Organize data in numerical order = 1, 5, 6, 9, 9
- 2) Find the middle value = 6

The median is 6.

The "mode" is the value that occurs most often. If no number is repeated, then there is no mode for the list. When you have a large amount of data, it may be helpful to sort your data in numerical order.

Example

Dataset: 1, 9, 5, 6, 9

The mode is 9.

The "range" is the difference between the lowest and highest number. To calculate the range, subtract the lowest number from the highest number. Again, it may be helpful to first organize your data in numerical order when you have a large amount of data.

Example

Dataset: 1, 9, 5, 6, 9

- 1) Organize data in numerical order = 1, 5, 6, 9, 9
- 2) Calculate the range by subtracting the lowest value from the highest value;
Range = $9 - 1 = 8$

The range is 8.

Inferential analysis

Inferential analysis is used to help you draw conclusions about your results. Many statistical tests can be used to explore the relationships found in your data. Common statistical tests include chi-squares, correlations, t-tests, and analyses of variance. If these statistics are not familiar to you, seek consultation to ensure that you select the right type of analysis for your data and interpret the findings appropriately.

The overall goal of inferential analysis is to determine whether results are meaningful. For example, did participants change in important ways over time? Were participants really different from people who did not receive services? In statistical terms, the meaningfulness of findings is typically described in terms of “significance.” There are two common forms of significance:

- *Statistical significance*: Using probability theory, statistical significance indicates whether a result is stronger than what might have occurred due to chance or random error. To be considered significant, there has to be a high probability that the results were not due to chance or random variation. When this occurs, we can infer that a relationship between two variables is strong and reliable. Several factors influence the likelihood of significance, including the strength of the relationship, the amount of variability in the results, and the number of people in the sample.
- *Clinical significance*: Statistical significance can be difficult to obtain, especially when data are available for a relatively small number of people. As a result, some evaluations focus instead on clinical significance. Clinical significance is sometimes seen as having more practical value, but only when there is a clear and justified rationale for establishing the underlying standards. Clinical significance does not necessitate statistical significance, but rather is based on what appears to be practically significant based on existing knowledge, expertise, and its meaning.

Tips for analysis

- Review and clean data to make sure everything is accurate, complete, and any inconsistencies have been resolved before beginning your analysis.
- Leave enough time and money for analysis – it is easy to focus so much on data collection that you do not leave enough resources to analyze the results.
- Identify the appropriate statistics for each key question – get consultation if needed.
- Do not use the word “significant” when describing your findings unless it has been tested and found to be true either statistically or clinically.
- Keep the analysis simple.

Analyzing qualitative data

Qualitative data is non-numerical information, such as responses gathered through interviews, observations, focus groups, written documents or journals, or open-ended survey questions. On its own, or in combination with quantitative information, qualitative data can provide rich information about how programs work. At the simplest level, qualitative analysis involves examining your data to determine how they answer your evaluation questions. Meaningful analysis of qualitative information can be challenging, however.

The first step in analyzing qualitative information is to reduce or simplify the information. Because of its verbal nature, this simplification may be difficult. Important information may be interspersed throughout interviews or focus group proceedings. During this first stage of analysis, you must often make important choices about which aspects of the information should be emphasized, minimized, or left out of our analysis altogether. While it can be difficult to remove comments provided directly by your informants, it is important to remain focused on the questions that you are trying to answer and the relevance of the information to these questions.

When analyzing qualitative data, look for trends or commonalities deeply-rooted in the results. Depending on the amount and type of data that you have, you might want to assign codes to the responses to help you group the comments into categories. You can begin to develop a set of codes before you collect your information, based on the theories or assumptions you have about the anticipated responses. However, it is important to review and modify your set of codes as you proceed to ensure that they reflect the actual findings. When you report the findings, the codes will help you identify the most prevalent themes that emerged. You might also want to identify quotes that best illustrate the themes, for use in reports.

Interpreting your results and drawing conclusions

Both quantitative and qualitative analysis only gets you so far. While the analysis can help you to summarize and identify key findings, you still need to interpret the results and draw your conclusions. Drawing conclusions involves stepping back to consider what the results mean and to assess their implications. During this phase, consider the following types of questions:

- What patterns and themes emerge in the results?
- Are there any deviations from these patterns? If yes, are there any factors that might explain these deviations?

- Do the results make sense?
- Are there any findings that are surprising? If so, how do you explain these results?
- Are the results significant from a clinical or statistical standpoint? Are they meaningful in a practical way?
- Do any interesting stories emerge from the responses?
- Do the results suggest any recommendations for improving the program?
- Do the results lead to additional questions about the program? Do they suggest that additional data may need to be collected?

Involve stakeholders. While evaluation findings must be reported objectively, interpreting those findings and reaching conclusions can be a challenging process. Consider including key stakeholders in this process by reviewing findings and preliminary conclusions with them prior to writing a formal report.

Consider practical value, not just statistical significance. Do not be discouraged if you do not obtain statistically significant results. While a lack of significant findings may suggest that a program was not effective in promoting change, other factors should also be considered. You may have chosen to measure an outcome that was too ambitious, such as a behavioral change that takes longer to emerge. In interpreting your results, consider whether there are alternate explanations for the lack of significance. It is also important to consider the practical significance of the findings. Some statistically significant results do not yield important information to guide program enhancements, while some findings that are not significant are still useful.

Watch for, and resolve, inconsistencies. In some cases, you may obtain contradictory information. For example, you may find that stakeholders describe important benefits of the services, but these improvements do not appear in pre-posttest comparisons. Various stakeholders may also disagree. For instance, staff may report improvements in participants that are not reported by the participants themselves. It can be challenging to determine which information is accurate, especially when comparing different peoples' viewpoints or perspectives. It is important to consider the validity of each information source, and to remember that various stakeholders can have valid viewpoints that vary based on their unique perspectives and experiences. Try to resolve these discrepancies and reflect them in your findings to the extent possible.

Summary

Organizing and analyzing your data, while sometimes daunting, is how you learn the meaning and results of an evaluation. Plan ahead for this phase of the project, taking into consideration the time, budget and expertise necessary to complete this essential phase of any evaluation. Tips for analyzing quantitative and qualitative data are identified throughout this section.

- Quantitative analysis can be descriptive, such as using frequency distributions, mean, median, mode, and variability. Quantitative analysis can also be inferential, a process that requires greater technical skills.
- Qualitative analysis consists of analyzing non-numerical data, such as information gathered from interviews, focus groups, or open-ended survey questions. The key to analyzing qualitative information is to make it manageable.
- Take care to interpret results accurately and to report sound conclusions. Get input from stakeholders to ensure that everyone is drawing similar conclusions from the information available.

**Wilder
Research**

Information. Insight. Impact.

451 Lexington Parkway North
Saint Paul, Minnesota 55104

651-280-2700

www.wilderresearch.org



For more information

To learn more about evaluation www.wilderresearch.org

AUGUST 2009