© WFP/Anna Eriksen

INNOVATION

Baseline and Endline Survey Guidelines

A short guide to collecting, analyzing and using outcome data to review your innovation project.
By Bethany Plant, Joy Matwale and Vrinda Kabra

Contents

1

Introduction

A baseline and endline survey are important review processes for any innovation initiative. Conducting these surveys will allow you to evaluate the outcomes of your project on your target population. The information you gather can be used to show accountability to current donors and stakeholders, prove your project's success to potential future partners and funders, and most importantly, understand what worked well and what didn't, so you can make improvements.

Baseline Survey

The baseline survey is conducted at the beginning of a project, before any interventions or activities are implemented. It establishes a reference point, or "baseline," for measuring changes and impacts throughout the project.

Endline Survey

The endline survey is conducted at the end of a project or project phase, after all interventions and activities have been completed. It assesses the project's outcomes and impact by collecting data that can be compared to the baseline responses as part of the review and evaluation process.

Example: Project A aims to improve educational outcomes in a rural area. At the start, they conduct a baseline survey to collect initial data on key metrics such as school enrollment and attendance rates for the past academic year. They then implement an intervention, providing healthy school meals for all students. After one year, they conduct an endline survey, collecting the same key metrics on school enrollment and attendance rates. By comparing these two data points, they can determine whether providing school meals has increased enrollment and attendance, thereby assessing the effectiveness of their intervention.

Help! I am Midway Through My Project and Didn't Do a Baseline.

If you find yourself midway through a project without a baseline survey, it's important to still conduct an endline survey and attempt to gather accurate baseline data retrospectively to support the project review process.

- **Accessing Existing Data:** Look for pre-existing data that may have been collected before the project began. However, this approach may have limitations if the desired data is inaccessible or not in the format needed.

Example: Project A could reach out to the schools they are working with to see if they have historical data on attendance and enrollment from previous years.

Recollecting Past Information:
Alternatively, you can ask endline survey participants to recall their situation before the project's implementation. Keep in mind that relying on participants' memories can introduce inaccuracies or biases into the data.

Example: Project A could ask participants whether they were enrolled in school before the intervention and if they can recall their attendance rate for the previous year.

WFP/Agron Dragaj

Defining the Survey

The first stage of preparing for a baseline and endline survey is defining the parameters of your review. This includes defining why you want to do the survey, what key questions you would like to answer, what indicators you want to measure, who you are going to survey, and how you will conduct the survey?

Why? Setting Survey Objectives

The general goal of the baseline and endline survey is to assess the effectiveness of your project. However, before designing a survey, it's essential to clearly define a specific objective for your evaluation to ensure that it yields meaningful and actionable insights. This involves stating exactly what you aim to achieve through the survey and identifying the key questions you want to answer with the data collected. Your specific objective should be shaped by your project's goals, the nature of the intervention, and any requirements set by donors or stakeholders at the start of the project.

If you're finding it difficult to define an objective for the survey, it may be helpful to revisit your *Theory of Change* which outlines how you assumed your project would impact your users or beneficiaries. These assumptions are often what need to be tested through a baseline or endline survey. For additional guidance on using your *Theory of Change* to establish project indicators, please refer to this video[1] on impact measurement for innovation.

What? Defining Your Indicators

Once you are satisfied with your objective, it is time to create relevant measurable indicators. An objective is a broad statement whereas indicators are specific variables that you will measure to fulfil your survey's objective.

There are many great compendiums of indicators that can be used for inspiration.

- WFP Indicator Compendium
- IRIS Catalog of Metrics
- SDG Indicators

When defining indicators you should follow the SMART indicator advice. This states that a good indicator is:

- **Specific:** Clearly defined so that everyone understands what is being measured.
- **Measurable:** You should be able to easily measure and collect data on the variable.
- **Achievable:** Realistic given the resources and context of the project.
- **Relevant:** Directly linked to the project's objectives and goals.
- **Time-bound:** Measurable within the project implementation timeline.

---

[1] https://www.youtube.com/watch?v=IQ94yYVzhbQ

Who? Sampling Methods

The term *sampling* refers to the selection of a limited number of individual units from a population of interest, with the purpose of inferring something about that population through the individual units selected.

When you are doing a baseline and endline survey for an innovation project that has a large number of beneficiaries you should choose a sample of individuals to reduce your workload.

All participants included in a sample must have direct experience with the intervention being evaluated.

As a general rule the minimum sample size for any level of statistical significance is 100 participants If you have less than 100 beneficiaries, you should survey them all. A maximum sample size is 10% of your population or 1000 participants whichever is smaller.

The size of the sample between the minimum and maximum is your choice. A larger sample gives more accuracy but also incurs more cost and takes more time. If possible, your baseline and endline survey should use the same sample. This ensures the results can be reliably compared.
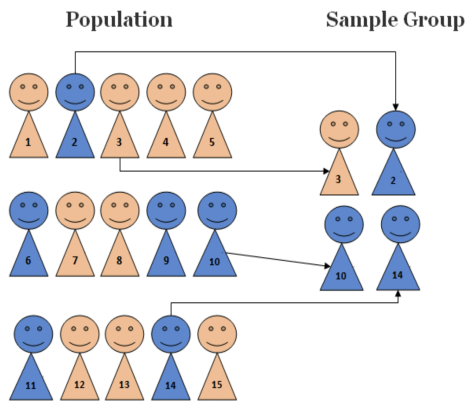
Your sample should be selected randomly and each individual should have an equal chance of being included. This reduces the

chance of bias and increases the reliability of the results.

There are multiple ways to choose a random sample:

> ➢ **Simple random sampling**

This should be used if:

→ you have a complete list of the individuals you would like to survey
→ all individuals are contactable
→ all individuals are located in a similar geographic area.

The list of individuals should be inputted into a spreadsheet and each record given a number. Then a random number generator or random number function (within your spreadsheet software) is used to select participants from the list.

> ➢ **Systematic random sampling**
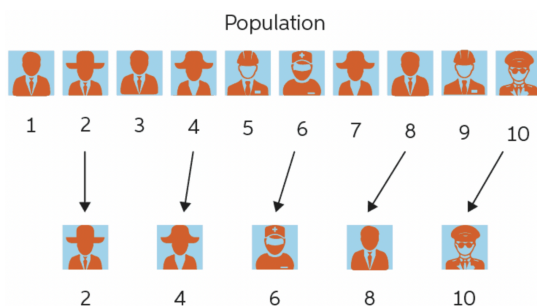
In this case, every second person in systematically selected

This should be used if:

→ you interact and are able to contact all of your beneficiaries BUT
→ you do not have a registered list of names and contact details.

In systematic sampling, instead of choosing randomly from a list, individuals are chosen at regular intervals to make up the sample.

*Example: Project C wants to sample 100 people out of the 1000 they are targeting. They don't register users but they do interact with every user at the point of sale.
They divide 1000 by 100 and get 10. Therefore, they decide that they will survey every 10th customer who buys their solution. Once they have sold 1000 solutions they will have their desired sample of 100.*

> ➢ **Cluster sampling**
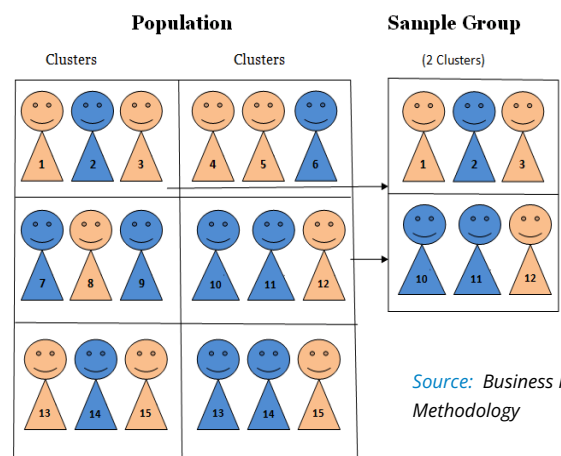
This should be used if:

→ your survey area is too large
→ and/or the population is dispersed.

The first step is to identify small geographic areas such as villages or settlements and input them into a spreadsheet. Then use a random number generator or random number function (within your spreadsheet software) to select areas from the list. Once you have selected areas complete simple

random sampling or systematic sampling only in those areas. Limiting your sample to randomly chosen geographic areas reduces your travel time and costs whilst minimising bias.

Example: Project C has sales points in 5 districts which are very far apart. They randomly choose 2 districts to be part of their sample. They still want 100 people in their sample so they must sample 50 people in each district. They use systematic sampling to select 50 people in each district.

➢ **Stratification**

This should be used if:

→ you want to compare the outcomes for different customer segments OR
→ you are testing two types of solutions.

To do stratification, you must first split your target group into subgroups and then apply simple random sampling or systematic random sampling to each subgroup.

Example: Project C is supplying urban farmers and rural commercial farmers. They want to see if the type of farmer makes a difference to their outcomes. They first divide their users into urban farmers and rural farmers, then they use simple random sampling to select 100 farmers in each group.

Note: *stratification usually doubles the size of your required sample. A sample with four groups would increase your sample size by four times.*

Hence, consider stratification only if the need for this information is important.

How? Data Collection Methods

To strengthen the ability of your endline survey to evaluate the effectiveness of your intervention and to measure the extent to which the intended outcomes and objectives were achieved, you should consider using a combination of quantitative and qualitative data collection methods. This will allow you to present clear, measurable success accompanied by participants' stories, both of which are essential for donor reporting.

Quantitative methods capture numerical data (e.g. percentages, rates, averages) which can be used to objectively measure your defined indicators.

Qualitative methods generate deeper insights into the experiences, perceptions, and attitudes of participants. This is valuable for understanding the *why* and *how* behind the results of your intervention.

Example: Quantitative data shows that 80% of users had increased their income. Interview data highlighted that the initiative also had an unintended benefit of improving nutrition as participants used vegetables they had grown to supplement their households meals.

Recommended Quantitative Data Collection Tool: Structured Questionnaires

A structured questionnaire is a research tool designed with fixed, predetermined questions that remain consistent across all respondents, ensuring the comparability and reliability of the data collected. The questions are arranged in a logical order, typically starting with general topics and gradually

moving towards more specific or sensitive issues, to maintain a coherent flow and keep respondents engaged. This structured approach ensures that the data is standardized, making it suitable for drawing generalizable conclusions.

The questions can be asked to participants through various modalities, including online forms (using tools such as Google or Microsoft forms), physical questionnaires, or in person. The latter is useful if translation is needed or if your population has limited literacy.

For information on how to write questions for your structured questionnaire see section 3 "Writing Survey Questions"

Recommended Qualitative Data Collection Tool 1: Focus Group Discussions (FGDs)

FGDs typically consist of a small, diverse group of participants and a moderator. The moderator should guide an open conversation among the participants based on predetermined topics or open questions. For example, what was the most significant change in your daily life after starting to use the solution?

FGDs should be used for:

- ***Understanding Community Sentiment***: Focus groups are ideal for exploring collective opinions and shared experiences because conversations between participants show you if the opinion is shared or is controversial.

- ***Generating Ideas and Concepts***: The group setting often stimulates creativity, presents diverse perspectives, and allows people to build on the insights of others. This can help you to think of ideas to improve your innovative solution.

Recommended Qualitative Data Collection Tool 2: Key Informant Interviews (KIIs)

KIIs are in-depth, one-on-one interviews with individuals who have specialized knowledge, experience or insights on your survey objectives.

KIIs should be used for:

- ***Expert Insights***: KIIs allow you to capture the views of those with deep knowledge of the topic, offering valuable expert opinions and recommendations.

- ***Sensitive Topics***: Individuals are more likely to open up about sensitive or personal topics in a one-on-one interview than in a group setting.

- ***Understanding Individual Perspectives:*** KIIs provide a more focused and personalized approach so they enable a deeper understanding of a single person's experiences, attitudes, or motivations.

WFP/Anna Eriksen

Writing Survey Questions

Once you have decided on the parameters of your survey, it is time to write the questions. These should be based on your indicators and any additional information that you will need to answer your survey objectives.

Below is a list of tips for writing good survey questions.

- **Use language that is clear and understandable**

  You should avoid jargon or technical terms in survey design. If respondents cannot understand your questions or response options, you will introduce bad data into your dataset. While we should strive to keep survey questions short and simple, it is sometimes necessary to provide brief definitions or descriptions when asking about complex topics, to prevent misunderstanding. Always test your questionnaires with the target audience to ensure that all jargon has been removed.

- **Stay neutral and do not suggest a desired response**

  Sometimes the way we phrase a question can influence the response we receive. When designing a survey it is important that we keep questions neutral and do not push the responder to respond in a certain way, as this introduces bias.

  Example of a bad question: In what ways has the hydroponic solution decreased your workload? This question assumes that the solution has decreased the workload of the participant and only allows them to give a positive response. Therefore the data will be biased towards positive responses.

Example of a good question: Since using the hydroponic solution, has your workload decreased, stayed the same, or increased? Please explain why in your answer. This question invites negative and neutral responses as well as positive responses and therefore, it does not bias the participant.

- **Use balanced scales**

Often, a respondent is asked to rate their agreement with a statement on an agreement scale (e.g., *Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree*), or otherwise to rate something using a scale of adjectives (e.g., *Excellent, Good, Neutral, Fair, Poor*). You'll notice that, in both of the examples given above, there is an equal number of positive and negative options (2 each), surrounding a neutral option. The equal number of positive and negative options means that the response scale is balanced and eliminates a potential source of bias or error.

In an unbalanced scale, you'll see an unequal number of positive and negative options (e.g., *Excellent, Very Good, Good, Poor, Very Poor*). This example contains 3 positive options and only 2 negative ones. It, therefore, biases the participant to select a positive option.

- **Mainly use closed-ended questions**

The most effective surveys focus on closed-ended questions (like multiple choice or rating scales) to gather numerical data, with a few open-ended ones added to provide qualitative context. This makes it easier to process and analyze all of the responses in your sample (remember you may have over

100). If your survey relies too heavily on open-ended questions, it could indicate that a different qualitative method, like Focus group discussions for interviews, might be more appropriate for your review process.

Nevertheless, it is a good idea to include a broad, open-ended question at the end of your questionnaire, such as "Is there anything else you'd like to share?". Many respondents may have feedback or concerns they want to share, but if no question invites this input, they might finish the survey feeling dissatisfied.

- **Avoid double-barreled questions**

A double-barreled question is one that requires respondents to address two separate issues or topics within a single question

Example of a double-barreled question: "*How easy and intuitive was the solution to use?*"

Easy and intuitive, while related, are not synonymous, and, therefore, the question is asking the respondent to use a single rating scale to assess the solution on two elements. As a result, the respondent will either pick one of these words to focus on or try to assess both and estimate a midpoint "average" score. Neither of these will generate fully accurate or reliable data.

Double-barreled questions should always be split up into two separate questions.

- **Have answer options that are all-inclusive and mutually exclusive**

Answer options for a multiple-choice question should include all possible answers (i.e., all inclusive) and should not overlap (i.e., mutually exclusive).

Example of a bad question:
How many people are in your household?

- ☐ *1-3*
- ☐ *3-5*
- ☐ *5-7*

In this formulation, some possible answers are skipped (i.e.,there is no way to answer if your household is over 7). Additionally, some answers overlap (i.e., someone in a household of 3 could choose the first or second response).

- **Provide an opt-out**

No matter how carefully and inclusively you craft your questions, there will always be respondents for whom none of the available answers are acceptable. Maybe they are an edge case you hadn't considered. Maybe they don't remember the answer. Or maybe they simply don't want to answer that particular question. it is good to provide an opt-out answer for these cases to avoid bad data.

Opt-out answers can include things like the following: *Not applicable*, *None of the above*, *I don't know, I don't recall*, *Other*, or *Prefer not to answer*.

Any multiple-choice question should include at least one of these answers. However, avoid the temptation to include one catch-all opt-out answer containing multiple possibilities. For example, an option labelled "*I don't know / Not applicable*" covers two very different responses with different meanings; combining them may confuse your data.

- **Don't ask irrelevant questions**

Based on the purpose of the survey, only ask questions that you need answers to and don't ask questions that you can find

answers to. One of the easiest traps to fall into when writing a survey is to ask about too much. For example, asking many standard demographic questions that are not relevant to your objective. This is tiring for the participant and wastes your time during data analysis.

When writing a question ask yourself: *Why do I need this information? Will I use it to answer my survey objective? Will it support my data analysis?* Even if you will use it, is there another way to capture it besides asking about it in a survey?

Example: Project C is installing hydroponic farming solutions. Their survey objective is to evaluate the solution's ability to increase agricultural outputs and raise farmers income and to find out whether their users like the product..

Bad question: "Where is your farm located?" This question is not relevant as the review is not looking into whether there are different outcomes in different locations. Even if the location is relevant, the question does not need to be asked. The survey was done in person at the farm, therefore the enumerator already knows the farm's location.

- **Avoid sensitive topics when possible**

When asking about any topics that may be deemed sensitive, private, or offensive, first ask yourself: Does it *really* need to be asked? Often, we can get plenty of valuable information while omitting that topic. Other times, it is necessary to delve into potentially sensitive topics. In these cases, be sure to choose your wording carefully. If necessary, also consider providing a brief explanation for why you are asking about that particular topic and what benefit will come from responding.

- **Don't assume negativity**

This is particularly important when discussing health issues or disability. Phrasings such *Do you suffer from a physical disability?* may be perceived as offensive. Instead, use objective wording such as *Do you have any physical disability?*

- **Use inclusive and appropriate wording for demographic questions.**

When asking about topics such as race, ethnicity, sex, or gender identity, use the preferred accurate and sensitive terminology for your target population. This may differ for different locations and communities. An inclusive question is respectful of your users' identities and allows them to answer only if they feel comfortable.

WFP/Shehzad Noorani

Pre-survey Checklist

Before you go to the field to conduct your survey please run through the following checklist:

☐ **Final survey design**

Review your survey design to ensure no last-minute changes are needed. Oftentimes a mistake is spotted only after having conducted several interviews. Therefore it is advisable to conduct a test run with a colleague or friend.

☐ **The right team members**

Ensure that you have all necessary team members for your data collection. For in person data collection you may need a moderator, note taker and translator. It is advisable to identify the personnel needs beforehand to ensure the effectiveness of the process.

☐ **Onboarding and training of the interviewers**

For in-person data collection, it is imperative that the interviewer is properly trained and onboarded. Training should focus on best practices for avoiding bias, the data collection methodology, the purpose of the evaluation and any cultural sensitivities present in the target population

☐ **Digital and physical tools**

Ensure all your tools are organized and easy to find. For in-person methods, this may be pens/pencils, printed questionnaires, and clipboards. For digital methods this may be a a laptop or tablet and a centralized database that automatically stores responses.

☐ **Data storage**

Have a designated place to store physical and digital survey responses. The storage location should be known and accessible to all members of the reviewing team. This will ensure that data does not get misplaced. Online stores should be password protected so that your participants' data is not accessible to unauthorized individuals

☐ **Ethical considerations and consent form**

A consent form or a verbal consent agreement should be drafted to ensure that participants are fully informed about the purpose of the data collection, how their data will be used, and that their participation is voluntary. Confidentiality should also be guaranteed in this agreement, especially when dealing with sensitive issues. It is essential that all participants sign the form or give verbal consent to being part of your endline survey.

Data Preparation and Analysis

This section of the guide will provide some guidance on how to prepare and analyze the data you have collected. To make this process simple for those without specialized skills or access to advanced tools like Stata and R the guidelines will focus on Excel, which is widely available and relatively user-friendly.

Data preparation

This is the process of cleaning, organizing, and transforming raw data into a usable format ahead of analysis. This crucial step ensures that the data is accurate, complete, and ready for meaningful analysis, by checking for errors and organizing the data in a structured way. Good preparation will enhance the quality of insights and conclusions drawn from the analysis.

The following steps guide you in preparing your data for analysis:

1. ***Data entry:*** Enter collected data into an Excel spreadsheet Use separate columns for each question/variable.

    Tip: assigning numerical codes for categorical data (e.g., 1 for "Yes", 0 for "No") can help with totals and averages later on.

2. ***Data validation:*** Check for completeness and consistency. Ensure all responses fall within expected ranges.

    Tip: Use Excel's "Data Validation" tool to restrict input ranges (e.g. ensure all responses are numeric) and identify outliers.

3. ***Handling missing data:*** Decide how to handle missing values. Options include leaving them blank or filling with a default value (e.g., the average response).

    Tip: Use the "Filter" feature to quickly identify missing data.

4. ***Data transformation:*** Adjust data to the required format for analysis or indicator reporting. This might include creating new variables or converting text to numerical codes.

    Tip: You could use Excel formulas to calculate "age" from birth dates or to calculate 'income increase' from income in year one and income in year two.

What are some best practice tips I can use to prepare my data for analysis?

- ***Documentation:*** Keep detailed records of all data cleaning and transformation steps for transparency and reproducibility.

- ***Consistency:*** Apply consistent methods and criteria for handling missing values, outliers, and other data issues.

- ***Automation:*** Use automated tools and scripts to handle repetitive tasks and reduce the likelihood of human error.

- ***Validation:*** Regularly validate the prepared data to ensure it meets the required standards and is suitable for analysis.

- *Video Tutorial: How to Download & Prepare Survey Data for Analysis in Excel (youtube.com)*[2]

---

[2] https://www.youtube.com/watch?v=LSm_rS8-6QM

Data Analysis

Data analysis is the process of systematically applying statistical and logical techniques to describe, summarize, and evaluate data with the goal of extracting useful information and identifying patterns to support decision making. In this section, the guide will walk you through three main types of data analysis methods: 1) descriptive analysis, 2) comparative analysis, and 3) thematic analysis (qualitative data)

➢ Descriptive Analysis

Descriptive analysis refers to the process of summarizing and describing the main features of a dataset. It typically involves calculating measures of central tendency (mean, median, mode), measures of variability (range, variance, standard deviation), and frequencies (count, percentage). This provides simple summaries about the sample and the measures, giving insights into the basic patterns in the data.

Below are some examples of descriptive analyses:

- *Measures of central tendency*: Use Excel functions =AVERAGE(range), =MEDIAN(range), and =MODE(range).

   Example: =AVERAGE(C2:C100) to find the average age.

- *Measures of variability*: Use Excel functions =MAX(range) - MIN(range) for range, and =STDEV.P(range) for standard deviation.

- *Measures of frequency*: use Excel "COUNTIF" function to tally responses. Create tables or charts using the "Insert" tab.

   Example: =COUNTIF(B2:B100, "Yes") to count the number of "Yes" responses.

When can I use descriptive analyses?

- *Initial Data Exploration:* when you want to understand the basic features of the data you have collected, as well as identify any patterns, trends and anomalies occurring in the data.

- *Summary Statistics*: when you want to provide a summary of the data for reports, presentations, and dashboards, you could use descriptive analyses to present an overview of key metrics such as average scores, distribution of responses, and demographic breakdowns.

- *Comparative Analysis Preparation*: descriptive analysis can help you prepare your data for comparative analysis (*in the next section*).

   Video Tutorial: [How To Perform Descriptive Statistics In Excel (Very Easy!) (youtube.com)](#)[3]

---

[3] https://www.youtube.com/watch?v=6osDRHWZtK8

➢ **Comparative Analysis**

Comparative analysis involves comparing two or more datasets, variables, or groups to identify differences and similarities. This can be used to compare the results of your baseline and endline surveys or different sub-groups if you have used *stratified sampling*.

The simplest form of comparative analysis is descriptive which involves summarizing the data of two groups and describing differences. Comparative analysis can also be used in a more scientific manner using simple statistical methods to test hypotheses about differences between groups. The easiest test to use is the simplified t-test which can be carried out in excel. However, even this level of statistical test will not usually be required unless your project is at a very large scale.

Below are some examples of comparative analyses:

- *Cross-tabulation*: compare two variables.

  How-to: Use a PivotTable to easily view two data sets together

  > *Tip: Highlight data, go to "Insert" > "PivotTable", and drag variables to rows and columns to create a table only showing your elements of interest..*

- *Compare means* (Simplified): use excel to do a simplified t-test

  How-to: Go to formulas, more functions, statistical, t-test

  > *Tip: Go to Scribbr.com to understand how to do and interpret a simple t-test[4]*

When can I use comparative analyses?

- ***Evaluating Changes Over Time***: when you want to identify trends or changes you could compare the same data from different time periods. This would be particularly useful if you are looking to evaluate the impact of an intervention by comparing your baseline and endline results.

- ***Comparing Groups***: when you want to compare different groups within a dataset (e.g.,urban vs. rural farmers) in order to assess the differences in outcomes, behaviors, or characteristics between the groups.

- ***Benchmarking***: when you want to compare your organization's performance or outcomes against benchmark or industry standards.

- ***Identifying Patterns and Relationships***: when you want to explore relationships between variables and how these differ across groups. This could potentially uncover patterns that can inform decision-making and strategy.

  > *Video Tutorial: [How to Construct a Contingency Table using a Pivot Table in Excel (youtube.com)](youtube.com)[5]*
  >
  > *Video Tutorial: [t-test in Microsoft Excel (youtube.com)](youtube.com)[6]*

---

[4] https://www.scribbr.com/statistics/t-test/

14

[5] https://www.youtube.com/watch?v=HvmpNQ0iJYc
[6] https://www.youtube.com/watch?v=BlS11D2VL_U

➢ **Thematic Analysis**

Thematic Analysis is used to identify, analyze, and report patterns (themes) in qualitative data, such as that collected in focus group discussions, interviews or open-ended survey questions.

The following steps will guide you through your thematic analysis:

1) **Data familiarization**: First, prepare and organize your data by transcribing your interviews and organizing your notes.

   Tip: this can be done using an AI transcription tool such as otter.ai.

   Next, read and re-read the qualitative data you've collected to get an overall sense. Take notes as you go, highlighting initial patterns you observe.

   Tip: import key phrases or ideas directly in the Excel spreadsheet you will use to analyze the data.

2) **Generate initial codes**: Highlight important features of your data and systematically label them with key words. These codes will help you categorize your data in meaningful ways to support performing the analyses.

   Example: You could use "User Experience Challenge" to group any sentence that explains challenges your users face whilst using the solution.

   Tip: Use a new column, in Excel, to tag each response with keywords or codes.

3) **Search for themes and refine**: Next, from the codes you have generated in the previous step, start grouping those that are similar into broad themes.

   Example: Within the " "User Experience Challenge" code you may have repeated challenges such as "material is too heavy" that you want to group.

   Tip: Use the "Sort" and "Filter" features in Excel to organize and group similar themes.

4) **Assess theme frequency**: Once you have refined the themes, the next step will be to count how often each theme emerges in the data. The more common a theme is, the more significant it is for your review.

5) **Write Up**: You can now begin writing a narrative that presents your interpretation of the analysis carried out in the previous steps. You can use the themes that you defined to group your findings in a coherent and logical manner.

   Tip: Use quotes from the data to illustrate each theme.

   Video Tutorial: Thematic Analysis | Explanation and Step by Step Example using Delve (youtube.com)[7]

   Video Tutorial: Qualitative analysis using Excel (youtube.com)[8]

---

[7] https://www.youtube.com/watch?v=rvMf1cbctYM
[8] https://www.youtube.com/watch?v=P0gzlWNodKw

Data Visualization

Data visualization is the graphical representation of information and data, providing an accessible way to see and understand your data.

Why to use Data Visualizations

- *Simplify complex data*: when you are looking to make large and complex data sets understandable.

- *Reveal patterns and trends*: when you need to identify correlations, trends, and outliers that may not be apparent in your raw data.

- *Facilitate decision making*: if you are looking to support your team to make informed decisions, data visualizations can help you present your data in a digestible format.

- *Enhance communication*: when you want to communicate your findings clearly and effectively.

Best practices for Data Visualizations

- *Clarity and simplicity*: Avoid clutter and focus on key messages to effectively communicate your findings.

- *Narrative context*: Complement visuals with a narrative that explains the context and significance of the data. Highlight key findings and insights from your analysis.

- *Digital tools:* There are many tools you can use to create exciting and interactive data visualizations such as Power BI or Tableau. However, if you are first starting out with data visualization it is best to keep things simple and affordable by effectively utilizing the visualization functions in Excel or google sheets. This guide focuses on these tools.

Types of Data Visualizations

- **Bar chart:** it uses rectangular bars to show the frequency or value of different categories. It is used to compare discrete categories or groups.

  Example: comparing your venture's sales figures across different regions or household incomes across refugees and host-community groups.

- **Histogram:** it is similar to bar charts but used for continuous data divided into bins. It is used to show the distribution and frequency of data points within a dataset.

  Example: Analyzing the age distribution of your customers.

- **Pie chart:** it is a circular chart divided into sectors to illustrate numerical proportions.

  Example: Showing the types of produce grown by your users

- ***Scatter plot***: it uses dots to represent values for two different variables and is used to show correlations or relationships between variables.

> *Example: Examining the relationship between yield increase and income increase to see if all yield increases are being converted into income.*

- **Choropleth maps:** Visual representations of geographic areas showing various features and trends within these areas. This is best for showing how a variable, like population density or income level, varies across a geographic area.

> *Example: Identifying areas with high demand for your solution.*

To create a choropleth map in excel, first list the locations in one column. Go to data type and choose geographic data. If there is a question mark next to any of the locations, it means that excel has been unable to locate it. Click on any question marks and use the search function to manually assign your desired geography. In a second column list the desired data points. These data points must be quantitative. For example, the number of sales in this area. Highlight the full data-set and click on insert, map, filled map.

> *Tip: If some regions are left in gray on the map it means that Excel has not been able to locate the region or data point. In this case, you should double check your data set for spelling mistakes or missing data.*
> *Tool: Download Excel workbook with several map examples & instructions[9]*

Infographics

Infographics combine images, data visualizations, and minimal text to create an engaging and easily understandable overview of a topic. Infographics use visual elements such as icons, illustrations, and color coding to enhance the presentation of data and concepts.
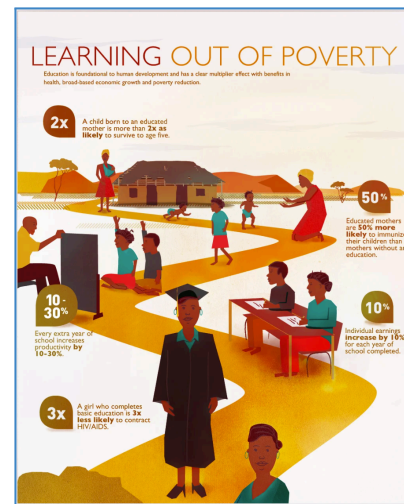
Source: Visme, Example of infographic showing 'education' as a multiplier effect for economic growth & poverty reduction

> *Tip: use Canva, a user-friendly online tool with templates for creating professional infographics.*

When can I use Infographics?

- *To simplify complex information*: Breaking down complex data or concepts makes it more digestible for audiences

> *Example: Explaining the steps of a scientific process or the results of a survey; Visually representing Health & Safety procedures.*

- *To engage audiences*: visually appealing elements capture and maintain the audience's attention.

> *Example: Marketing campaigns, social media posts, educational materials.*

---

[9] https://download.microsoft.com/download/5/3/C/53 C9D26C-B87D-45BE-A76C-FFCF78703940/Map%20Ch art%20samples.xlsx

- *To highlight key points*: when you want to emphasize essential data points and insights for quick understanding.

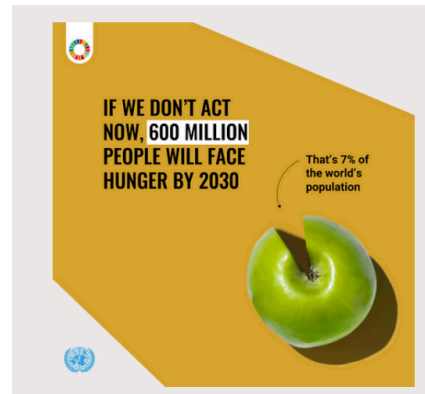   *Example: Annual reports, project summaries, and executive dashboards.*

- *Comparative analysis*: when you want to visually compare different datasets or concepts side by side.

    Example: Comparing yield and income before your intervention vs after your intervention.

- *Storytelling*: when you want to use visual narratives to tell a story or convey a message.

    Example: Chronicling your company's growth, explaining historical events, or presenting case studies.

- **Design Layout:** this includes *Flow and Structure* organized in a logical flow to guide the reader through the information. *Typography* using clear and readable fonts to ensure the text is easily understandable.

What to Consider When Creating an Infographic

- **Visual elements:** these could include *Icons and Illustrations* to simplify and represent data points or concepts. *Colors and Shapes* to enhance readability and visual appeal, as well as categorize information. *Images* to support the data with relevant visuals.

- **Data and Information:** this could include *Charts and Graphs* to present quantitative data in a visually appealing format. *Statistics* to highlight key numbers and figures. *Text* to provide context and explanations, remembering to keep this to a minimum to maintain visual appeal.

Utilization of Findings

Now that you have collected, processed, analyzed and visualized your data it's time to put it to use! Endline survey data has three main purposes, reporting, communications, and most importantly learning.

➢ **Reporting**

Most donors or stakeholders that you have engaged with during the project will want a report on your endline review. These reports provide evidence that you implemented the initiative as planned, and show whether the innovation project was a success.

For your WFP sprint programme, the outcomes of your survey will be asked for in your *project tracker* and the *end of sprint report*. It is mandatory to report this data when asked. For the sprint project, you should not worry if your endline review is not as positive as you would hope or if you did not meet all outcome targets. Sprints are designed to test new innovations, and not every test will be a success. It is more important to provide accurate data and identify learnings, than present positive data.

➢ **Communications**

If your endline review provides quantitative evidence that your innovation works and positively impacts your users, this can be used to advertise your solution on your website, pitch decks or other marketing materials.

Example: Project C found that their hydroponic solution doubled Urban farmers' annual income on average in their sample. They can use this statistic on their website to entice future customers to buy the product.

Furthermore, providing quantitative evidence of impact and demonstrating the ability to conduct quality monitoring and evaluation activities will make your team more attractive to granting

organizations and impact investors. These stakeholders often make decisions based on how much impact they can create with their investment. Therefore, including end-line results as additional materials would be a strong way to improve your chances of success.

➢ **Learning**

Organizational learning is the most important outcome of endline reviews. This is because your surveys will provide you with key insights into what is working in your solution, what you should keep doing, what your users want, and what could be improved to help you better achieve your goals. As such, engaging with your data and using it in decision making can make the difference between a successful and unsuccessful venture.

Example: Project C notes that farmers quadrupled their yields of leafy greens, but their income from leafy greens only doubled. They looked into this trend further during analysis and found that it was due to high post-harvest losses. They ideate on potential solutions and decide in the future they will sell a coldbox alongside their hydroponic solution to reduce post-harvest loss and ensure all extra yield can go to the market.

To ensure engagement with your data and integration into decision making we suggest having a learning session with your whole team after the endline review. This session should begin with a presentation of data and key trends, followed by a team discussion to identify learnings and provide context. Finally, you should brainstorm action steps for how to improve your solution. Your innovation coach should be available to support this session.

Thank you!

About the authors:

Bethany Plant is a Monitoring, Evaluation
and Learning Consultant at the WFP
Innovation Accelerator.
Joy Matwale is an Innovation Community
Manager at the WFP Regional Bureau for
Eastern Africa in Nairobi, Kenya.
Vrinda Kabra is an Innovation Consultant at
the WFP Innovation Accelerator.

World Food Programme Innovation Accelerator

Josephspitalstraße 3, 80331 München

innovation.wfp.org