How do we **process and analyse** data in a profiling process?

**WE ARE HERE!**

- Initiating the profiling exercise
- Establishing the collaborative process
- Designing the methodology
- Implementing data collection
- Processing and analysing the data
- Validation, reporting and dissemination
Table of contents

Why is this phase important for a profiling process?
page 3

What are the key components of this phase?
page 4

How does JIPS recommend planning and implementing this phase?
page 6

What are the background concepts I need to know?
page 8

Unit of analysis
Unit of measurement
Disaggregation
Correlation
Causation

Guidance:
- Data processing
- Analysing qualitative data*

Templates:
- Workshop for analysing data
- Processing checklist*
- Sample table of summary statistics*
- Concept note and agenda for a joint analysis workshop*

* Coming in 2018

www.jet.jips.org/phase/processing-and-analysing-the-data
Why is this phase important for a profiling process?

This step is critical to make actionable information out of the data; without processing and analysis the data would just be numbers or words that would not mean anything on their own.

Processing the data involves organising it into a form that can be analysed. Analysing the data means identifying patterns, trends, and anomalies that make up the basis of the evidence of a profiling exercise. In other words, this is the process of turning facts and figures into information. With this information, profiling partners can agree on the most relevant findings and make recommendations for their context.
For qualitative data:

- Reading the transcripts and debriefing with facilitators and ensure interpretability of the transcripts

- Apply social science and contextual reasoning to assess and decide on the most important findings

- Compare and contrast emerging results with other data sources

- Identify categories according to the indicators from the Analysis Plan; organize and summarise the categories

Once ready to tackle the data, note that the approach is generally the same whether working with qualitative (words) or quantitative (numerical) data. It is the way that you complete each step, the techniques you use, that differ. The basic components are:

For qualitative data:

- Identify and correct omissions and errors in the data, then calculate or create new variables to measure the indicators from the Analysis Plan

- Create tables showing the results for each indicator

- Identify similarities and differences between the units of analysis (i.e. target groups, locations)

- Apply social science and contextual reasoning to assess and decide on the most important findings

For quantitative data:

- Decide who within the Profiling Working Group will take forward the technical tasks of the phase, from the data processing to the preparations for an analysis workshop. This is typically a small group with relevant, complementary technical capacity (for instance on both the quantitative and qualitative sides) working in close communication with each other.

- If additional technical capacity is needed, seek expertise from partners both within and outside the Profiling Working Group.

- Revisit the Work plan to establish clear timeframes for this phase.

- Decide on how the process for data processing and analysis will be documented, for instance by adding a section to the Methodology Document describing the steps taken and problems encountered along the way that could affect the profiling results.
A few more steps need to be taken before arriving at findings (in other words, the result after jointly interpreting the data):

☑ Consolidate the “preliminary results” - so tables and graphs from the quantitative data and summaries from the qualitative data - into one manageable format, for instance a presentation or a document. Don’t forget to include the findings from the secondary data review.

☑ Present them in a joint analysis workshop to the Profiling Working Group, inviting outside expertise or other interested stakeholders as relevant, to generate discussion and feedback that could contribute to a deeper and more relevant interpretation of the data.

☑ Plan and conduct visits or workshops with displaced communities to share results of the profiling and discuss their perspectives, again to contribute to a deeper and more relevant interpretation of the data.

☑ Revisit the raw data to generate more analysis based on the feedback.

☑ Discuss with the Profiling Working Group next steps for developing a final report of the results.

<secondary data

interviews

household survey

focus group discussions

triangulating data is fun!
How does JIPS recommend planning and implementing this phase?

Follow the established Analysis Plan and Tabulation Plan

This tells you which indicators to measure with the data and is a useful starting point to any analysis. Also, because these were discussed by partners early on in the process, and revisited throughout, referring back to them in the analysis phase ensures that the analysis will be carried out in a transparent and agreed-upon way.

Review the data collaboratively through workshops

This reaffirms the collaborative nature of the profiling process and helps partners understand what they can use the data for. Concretely, this works by creating a series of tables and summaries of the data, and then splitting up profiling partners into groups according to their interest and areas of expertise to review and discuss the results.

While the crux of the analysis would be done by a smaller group of technical partners as delegated by the Profiling Working Group, opening up the discussions to a broader group ensures broader interest and relevance of the results, and provides useful input to the analysis. Request feedback on the results by asking questions such as:

› Are these results surprising for this context?
› What additional questions do these results raise?
› Is there further analysis that could be useful to contextualising this result?
› How does this look to different government ministries, development or humanitarian actors working here?
› This can then give new directions for further analysis for the drafting of the final report.

For example, in Sittwe, Myanmar (2017), the Profiling Coordinator produced a series of tables showing the results of the household questionnaire and did not include any interpretation of that data. She presented those results in various analysis workshops with the Profiling Working Group members, who brought in different expertise from each of the humanitarian sectors to discuss the results. This was done first by sector to go into depth and then

Read about the findings and conclusions from the profiling exercise in Sittwe, Myanmar.

www.jet.jips.org/phase/processing-and-analysing-the-data
all together for a more holistic understanding and to see how the sector-specific information related to the bigger picture. This produced a rich analysis of the different types of vulnerability faced by subgroups of the displaced populations living in camps, showing for example that elderly-headed households tended to have higher levels of humanitarian need.

**Document the process**

The documentation should make it possible to review what the results are based on and to replicate the study at a later stage. This means keeping track of any changes made to the raw data. This is critical for transparency; partners can double-check that the data was well processed and can replicate the same analysis. This can be done manually or automatically using statistical software.

For example, the household questionnaire in a profiling exercise in El Salvador (2017) was conducted using pen and paper to avoid security risks. This meant that all of the questionnaires had to be manually entered into a database before they could be processed and analysed. The small group designated by the Profiling Working Group to take forth the analysis of the data documented every change made to the raw dataset after receiving it using the **syntax function in SPSS**. This proved critical when the implementing partners discovered new boxes of completed questionnaires that they had overlooked during data entry. Instead of having to process the data all over again, the partners analysing could take the new raw dataset with the questionnaires added and apply all the same changes they made before onto the larger dataset. Thus the documentation not only made the processing more transparent, it also saved a lot of time in the event of unexpected issues (also see our [guidance on data processing](https://www.jet.jips.org/phase/processing-and-analysing-the-data)).

**Automate processing as much as possible**

We all make mistakes - by using Excel formulas or statistical software like SPSS or R, we can avoid making manual changes to the data and make the processing more accurate and efficient. This may not be easy to learn without prior background in data analysis or programming, but this expertise can be found through partnerships with technical organisations, for instance information management officers in the humanitarian sector, implementing partners specialising in data collection, or National Statistics Offices.
What are the background concepts I need to know?

**Unit of analysis**

These are the main groups being analysed for similarities and differences. For instance, a profiling exercise may be carried out to understand the differences between IDPs and host populations; alternately, it could focus on differences between certain geographical areas.

**Unit of measurement**

This is the unit described by the data. For instance, a profiling exercise analysing the difference between two population groups could do so using the household and/or individuals as the unit of measurement. The indicators should specify which unit it seeks to measure.

**Disaggregation**

This is when a larger group is broken down into subgroups according to certain characteristics. For instance, disaggregating the target population by sex, age, or displacement status would give more in-depth information about those subgroups.
Causation

There are several theories of causation. It could be that A is always followed by B; A is a necessary and/or sufficient condition for B; there is a causal mechanism from A to B; or that the occurrence of A raises the probability of B.

In profiling, we are not able to determine what variables cause a specific effect. But we can demonstrate correlations, or general trends within a population and see general differences between population groups. When we are comparing different population groups based on their displacement status, we are essentially exploring whether the groups differ from each other as a result of having been displaced. But we cannot say for sure that it was exclusively displacement that led to these different results. In other words, we can show that the overall differences correlate with their displacement status, but because there are so many other variables that interact with displacement, for instance ethnicity, age-group, place of origin, etc., we cannot say for sure that it was the act of having been displaced that caused the differences, and that the same households would have been better or worse off had they not been displaced.

Correlation

This describes the degree to which two variables are related; when one variable goes up, does another variable tend to also go up? For instance, the height of parents and their child are correlated: taller parents tend to have taller children, and shorter parents tend to have shorter children. There does not need to be causality for there to be correlation. In profiling, we can in some cases show how certain variables correlate with each other to identify general tendencies within a population.