QUANTITATIVE ANALYSIS

Quantitative monitoring and evaluation methods involve collecting and analysing data in the form of numbers rather than words. There are two main types of quantitative analysis. Descriptive statistics are used to describe or present data in an easily accessible form. More complex statistical analysis is used to show changes resulting from a project or programme, and to draw conclusions.

Quantitative monitoring and evaluation (M&E) methods are designed to collect and analyse data in the form of numbers rather than words. For simplicity, quantitative data can be divided into two types.

- ***Administrative data*** is generated through basic monitoring processes. It is often concerned with activities or outputs, such as the number of training sessions conducted, or the number of children immunised. It may also cover finances or logistics.
- ***Sampled*** or ***population data***, on the other hand, is often collected to assess changes resulting from a project or programme. It usually includes information taken from a sample (of people, households, communities, events, etc.) or an entire population, such as all the farmers working in a region.

Broadly, there are two different forms of quantitative analysis. ***Descriptive statistics*** are used to describe or present data in an easily accessible form. They can be used with both administrative and sampled (or population) data. Examples of descriptive statistics include financial reports, or simple tabulations showing outputs such as trainings conducted, seeds delivered to farmers, or water points installed.

More complex statistical analysis is normally only carried out on sampled data. Within M&E, the purpose of more complex statistical analysis is usually to show changes resulting from a project or programme, and then to use that information to draw wider conclusions. This is sometimes known as ***inferential statistics***. This means that conclusions or findings for wider populations are based on (or inferred from) results obtained in a sample. For example, if information collected from a sample of people shows that assets have increased in line with support provided, then it may be reasonable to suppose that this is also true for the wider target group.

Within CSOs, administrative data is usually collected through basic record-keeping, such as financial transactions, records of trainings delivered, etc. Sometimes, administrative data is generated automatically. For example, most websites automatically capture data on how many people are viewing web pages or downloading copies of reports. Sampled or population data, on the other hand, is usually collected through data collection methods such as surveys. Surveys may be based around interviews, observations or direct measurements (e.g. the height and weight of children). Some of the key terms used in sampled or population data are shown in the box below.

Terms Used in Sampled or Population Data

The **dataset** is a single file that contains the data you are going to analyse. It is normally organised into cases (usually rows) and variables (usually columns).

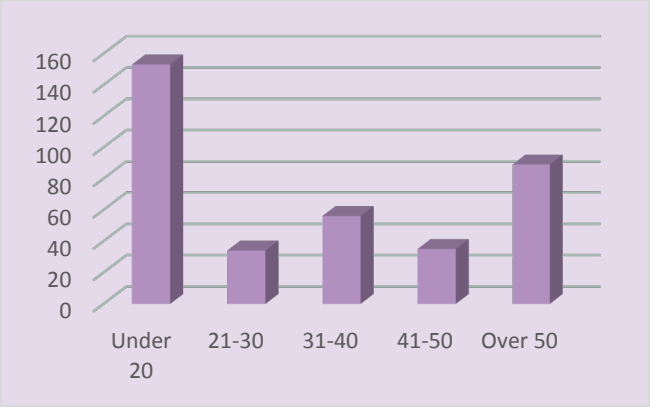
A **case** refers to a single unit in a dataset about which different information is collected. Examples might include individual survey respondents, a community, a project, a school, a publication or an event.

A **variable** refers to a single piece of information that has been collected across the different cases. Examples might include height and weight of children, income levels of farmers, school exam scores, training satisfaction levels, the number of times a publication has been downloaded, or any other piece of information that can be quantified.

Basic statistical processes

Many different types of statistical processes can be used for quantitative analysis. Some of the more common ones used for descriptive statistics are described below (see Trochum 2006 for a fuller description).

- The central tendency of a distribution (better known as the ***average***) is used to estimate the centre of a distribution of values. The most common form of average is the *'mean'*, which is calculated by adding together a variable across all the different cases, and then dividing the total by the number of cases.
- ***Dispersion*** is used to show how variables are spread across a range of values. The simplest method of showing dispersion is the range, which shows the difference between the highest and lowest values. A more useful method is known as standard deviation. This describes the relationship between a set of values and the *'mean'* average of those values.
- A ***frequency distribution*** shows a breakdown of individual variables according to different criteria. For example, the chart on the following page shows a simple breakdown of the ages of people living within a village.



- Whilst the three examples above are all based around examination of a single variable, ***correlations*** are used to describe the relationship between two variables. A correlation results in a single number between 1 and -1 that shows how two variables are related. Correlations are often accompanied by statistical significance tests. These show how likely it is that the correlation is a matter of chance.

Statistical processes for inferential statistics (the kind used when applying randomised control trials or quasi-experimental approaches) are much more complicated, and usually require a degree of statistical expertise.

Common elements in quantitative analysis

Whichever way the information was generated, many elements of quantitative analysis used within M&E are similar. Some common elements are described below.

Common Elements in Quantitative Analysis

Data collection: Within M&E, quantitative analysis is based around data collection tools and methodologies that generate numbers. Sometimes numeric data is generated through simple record-keeping or other kinds of administrative process. Sometimes it is collected deliberately in order to assess changes resulting from a project or programme. The most common collection methods for quantitative information are surveys based on interviews, structured observation, checklists and/or direct measurements.

Data storage: Raw data needs to be stored both manually and (if necessary) electronically to make sure it can be retrieved when necessary.

Data entry: Normally, raw data is first placed into a dataset and structured according to the needs. Nowadays, the dataset is usually developed on a computer, using a spreadsheet or simple database. If using a spreadsheet, information is normally sorted into cases (rows) and variables (columns).

Data preparation and cleaning: The aim of this stage is to ensure that data can be manipulated easily. The data needs to be inspected for completeness and accuracy. This may mean dealing with incomplete or wrong data. Sometimes, qualitative data needs to be coded in order to transfer it into numeric form.

Tabulation and summary statistics: The next step is to describe and summarise the data. This will normally involve some of the processes described in the section on basic statistical processes (e.g. frequency distributions, averages, measures of dispersion, correlations). Tabulation means presenting information in a table form, with clearly labelled rows and columns. Data can also be shown as charts or graphs. These are often most useful when the communication of trends and patterns is considered more important than the presentation of exact figures.

Descriptive analysis: Descriptive analysis is used to identify and show patterns in the data. Descriptive analysis may involve cross-tabulations, showing how different variables compare to each other. It may also involve analysis of sub-groups (such as boys and girls) within the data. Descriptive analysis may show how variables change over time, for example how many children turn up to school during different seasons.

Statistical analysis of differences and associations: These methods, including the calculation of confidence intervals and the statistical testing of differences, are only normally used for inferential statistics. Their aim is to test hypotheses, and confirm any patterns identified. Statistical analysis is routinely used when CSOs use experimental approaches, such as randomised control trials or quasi-experimental approaches. However, statistical analysis may also be used when comparing change against a baseline, or in any other circumstances where data is collected for the purpose of assessing numerical change, or contribution to change.

More complex analysis can be carried out in some circumstances. The aim is to explore underlying patterns and account for complexities in the structure of the data. More complex analysis techniques, such as multivariate analysis and modelling, are beyond the scope of this paper, and require specialist knowledge.

Presentation of data and analysis: Finally, data and findings need to be presented. The type of presentation depends very much on the audience. Some people cannot understand tables and statistics, and need to have findings explained clearly in descriptive form. Other people like to see exactly how results were produced, so that they can check whether statistical procedures have been properly followed. Larger studies tend to present data and analysis in many different ways to suit different audiences.

Challenges when working with quantitative analysis

Many stakeholders prefer quantitative to qualitative data as a basis for decision-making. This is for several reasons. Firstly, the rules for quantitative analysis are well known and well established. Provided these rules are properly followed, quantitative analysis should yield the same results whoever carried out the work. This contrasts with qualitative analysis where a lot rests on the skills and integrity of the person carrying out the analysis.

Secondly, the fact that quantitative studies can be replicated rules out deliberate bias. In theory, anyone with access to the same data could produce the same results. This means work can be checked and verified. This makes it much harder for findings to be manipulated to suit the individual or organisation carrying out the analysis.

Thirdly, when dealing with complex statistical studies, results can be quoted with a known margin for error, which can be accurately calculated. This means there is complete clarity regarding whether, or how far, any results are likely to be accurate.

However, there are a few factors that can seriously undermine the value of quantitative studies. The most important of these are described below.

- To be useful, data first needs to be collected and stored correctly. If information is incorrect before being processed it will result in inaccurate and misleading findings afterwards. Sometimes information can be measured directly (e.g. measuring the weight of new-born infants or measuring pollution in ponds), in which case it should be accurate. But quantitative information is often collected through interviews, and there are many reasons why people will not give honest answers to questions. For example, it is notoriously difficult to get honest answers to questions about household income.
- Sometimes the quantitative information collected does not really represent a desired change. In some sectors (e.g. health, water and sanitation) there are many standard, numeric indicators that can be used to show change. But in other sectors of work, such as governance or capacity development, it is much harder to find numbers that clearly show desired changes.
- In quantitative analysis it is rare for information to emerge over the course of a study. This means it is important to know what information is needed before data collection starts. This contrasts with qualitative analysis, where findings can emerge over time.
- The most common mistakes in statistical analysis are around sampling. It can be very hard to infer results from anything other than straightforward random sampling. Applying results from a sampled population to a wider population often relies on making assumptions that may or may not be justified.
- Even where results can be accurately calculated with known margins of error, some degree of interpretation is still needed. For example, a study might show that livestock ownership amongst farmers has increased by 30% over a two-year period. Whilst the facts may not be in doubt, the implications may still be a matter for debate. Is increased ownership of livestock a good thing? Does a 30% increase warrant the investment? Might there be better or cheaper ways of bringing about the same results?

In reality, as with qualitative analysis, the findings of quantitative analysis studies are always open to dispute to some degree. However mechanical and replicable the process of quantitative analysis, the information still needs to be interpreted by humans.

Electronic analysis

In the past much statistical analysis had to be done by hand, or using slide rules or logarithmic tables. Nowadays there is normally no need to perform calculations manually. The widespread introduction of calculators, spreadsheets and databases has made quantitative analysis much easier. There are also dedicated statistical packages (such as the Statistical Packages for the Social Sciences (SPSS)) which enables non-experts to produce analytical statistics such as standard deviations and confidence levels without needing to know precisely how these are calculated.

However, there are still times when detailed statistical knowledge and judgement are needed. Part of the skill of an M&E practitioner or evaluator is knowing when something can be learned and applied easily, and when it is necessary to call in an expert.

Further reading and resources

Quantitative analysis is further explored in two other papers in the M&E Universe, dealing with randomised control trials and quasi-experimental approaches. Other papers deal with qualitative analysis and the use of rating and scalar tools.

Randomised control trials

Quasi-experimental approaches

Qualitative analysis

Ratings and scales

There is a website article dedicated to social research methods that covers quantitative analysis methods (see Trochum 2006, referenced below). Another useful website is the WISE website (http://wise.cgu.edu) which is a web interface for statistics education, and contains many tutorials on statistics and related subjects.

References

- Trochum, W (2006). *Research Methods Knowledge Base. Descriptive statistics.* https://www.socialresearchmethods.net/kb/statdesc.php

Author(s):
Dan James and Nigel Simister

INTRAC is a not-for-profit organisation that builds the skills and knowledge of civil society organisations to be more effective in addressing poverty and inequality. Since 1992 INTRAC has provided specialist support in monitoring and evaluation, working with people to develop their own M&E approaches and tools, based on their needs. We encourage appropriate and practical M&E, based on understanding what works in different contexts.