Manage data

Good data management includes developing effective processes for consistently collecting and recording data, storing data securely, backing up data, cleaning data, and modifying data so it can be transferred between different types of software for analysis.

Good data management is inextricably linked to data quality assurance –the processes and procedures that are used to ensure data quality. Using data of unknown or low quality may result in making the wrong decisions about policies and programmes. Data quality assurance (DQA) should be built into each step in the data cycle ? data collection, aggregation and reporting, analysis and use, and dissemination and feedback.

Even when data have been collected using well-defined procedures and standardised tools, they need to be checked for any inaccurate or missing data. This “data cleaning” involves finding and dealing with any errors that occur during writing, reading, storage, transmission, or processing of computerised data.

Ensuring data quality also extends to presenting the data appropriately in the evaluation report so that the findings are clear and conclusions can be substantiated. Often, this involves making the data accessible so that they can be verified by others and/or used for additional purposes such as for synthesising results across different evaluations.

Commonly referred to aspects of data quality are:

- *Validity*: The degree to which the data measure what they are intended to measure.
- *Reliability*: Data are collected consistently; definitions and methodologies are the same when doing repeated measurements over time.
- *Completeness*: Data are complete (i.e., no missing data or data elements).
- *Precision*: Data have sufficient detail.
- *Integrity*: Data are protected from deliberate bias or manipulation for political or personal reasons
- *Availability*: Data are accessible so they can be validated and used for other purposes.
- *Timeliness*: Data are up-to-date current and available on time.

Methods

- Consistent data collection and recording

  An important aspect of data quality is to ensure data is collected consistently across different sites and different data collectors.

- Data backup

  Data backup refers to onsite and offsite, automatic and manual processes to guard against the risk of data being lost or corrupted.

- Data cleaning

  Data cleaning involves the detection and removal (or correction) of errors and inconsistencies in a data set or database due to data corruption or inaccurate entry.

- Effective data transfer

Effective data transfer involves processes to move data between systems, including between software packages, to avoid the need to rekey data.

- Secure data storage

  Processes to protect electronic and hard copy data in all forms, including questionnaires, interview tapes and electronic files from being accessed without authority or damaged.

- Archive data for future use

  Putting systems in place to store de-identified data so that they can be accessed for verification purposes or for further analysis and research in the future, researchers can extend the range of the data collection efforts and encourage future innovati

Resources

- Data management

  Supports the design of quality data management systems. (Food and Agriculture Organization, Fisheries and Aquaculture Department)

- Data quality tools and mechanisms (archive link)

  Guides to three tools that can be used to assess the quality of data and reporting systems. (The Global Fund)

- Data Quality

  This online course from the Global Health Learning Centre is designed to help learners understand what data quality is, why it is important, and what programs can do to improve it.