

Introduction to Data Sources (M&E)

Understanding where data comes from
And why it matter

ule2: Introductop ule 2 2

Module Overview

- Data Sources = Foundation of Monitoring and Evaluation
- Why? Measures progress, outcome and impact
- Types: Qualitative and Quantitative
- Examples: Surveys, Interviews, Focus Groups, Observations, Records, Reports

Learning Outcomes

- Define types of data sources
- Apply knowledge to select appropriate sources for analysis/ research
- Evaluate reliability, quality and suitability of different sources
- Analyze challenges in accessing, integrating, managing heterogeneous data
- Synthesize and communicate insights derived from data, acknowledging limitations

Introduction to Data Sources and Data Collection Tools

Unit 1: What are Data Sources

Definition:

Fundamental components in Monitoring and Evaluation (M&E) processes, serving as origins of information

Essential Role:

- Evaluate impact – measure outcomes and effectiveness of interventions
- Inform decision-making – Provide evidence for shaping strategies and policies
- Enhance accountability – Ensure transparency with stakeholders and donors
- Optimize resources allocation: Direct resources to areas with greatest need

Unit 2: Classification of Data Sources

Primary Data Sources

- Raw, unprocessed data gathered researcher specifically for M&E purposes
- Collected first-hand through registers, patient forms, EMRs
- Collected directly through activities e.g surveys, interviews, observations

Secondary Data

- Pre-existing processed data collected for other purposes but used in M&E
- Collected through Reports, Statics, Institutional Records

Unit 3: Primary vs Secondary Data

Basis	Primary Data	Secondary Data
Meaning	First hand data gathered by researcher	Collected earlier by someone else earlier
Data	Real-time	Past data
Process	Very involved	Quick and easy
Source	Surveys, observation, experiments, questionnaires, personal interviews	Government publication, websites, books, journal articles
Cost effectiveness	Expensive	Economical
Collection Time	Long	Short
Specificity	Tailored to researcher's needs	May not be specific to researcher's needs
Available in	Raw data	Processed data
Accurarcy and Reliability	Higher	Lower

Unit 4: Data Collection Methods

Primary Data Sources

- **Surveys and Questionnaires:** Structured sets of questions used to gather quantitative and qualitative data directly from respondents.
- **Interviews:** Direct one-on-one or group conversations, which can be structured, semi-structured, or unstructured.
- **Observations:** Systematic watching and recording of behaviors, events, or conditions.
- **Focus Groups:** Guided group discussions facilitated exploring participants' attitudes and opinions.
- **Experiments:** Controlled studies where variables are manipulated to observe effects.
- **Case Studies:** In-depth examinations of specific subjects or events, often supported by qualitative data analysis tools.
- **General Research Management Tools:** Notion for project management.

Secondary Data Sources

- **Identify your needs:** Determine what kind of information is needed for your evaluation questions.
- **Locate existing sources:** Search for relevant data in government reports, previous evaluation studies, academic research, international agency databases (e.g., WHO, UN), administrative records, public statistics, or media sources.
- **Access and retrieve:** Obtain the data, sometimes requiring permission, by downloading, requesting access, or manually extracting it from libraries, online archives, or institutional databases.
- **Assess relevance and quality:** Evaluate the data's credibility, age, methodology, and how well it matches your specific needs.
- **Analyze and use:** Summarize, analyze, or triangulate the data with any primary data collected for a comprehensive evaluation

Data Collection and Tools and Systems

Primary Data Sources

- **Patient Records** - Patient files at records, IDCC filing rooms
- **Electronic Medical Records (EMR) Systems** - PIMS, OpenMRS, DHIS2
- **Dashboards and Data Analytics Platforms** - Power BI, Tableau, Superset
- **Health Information Systems (HIS)** - DHIS2, Service delivery registers, ward records, outpatient/inpatient forms
- **Surveys and Questionnaires** - Patient satisfaction surveys, staff satisfaction surveys
- **Clinical Audit Tools and Checklists** - Procedural checklists, infection control audits
- **Facility Registers and Logbooks** - Outpatient/inpatient registers, Delivery registers (in maternity), Immunization and vaccination logs

Secondary Data Sources

- **Government Records and Reports:** Health information systems (e.g., DHIS2), Vital registration (births, deaths)
- **International Organization Databases:** DATIM, UN (e.g., Global Health Observatory, Spectrum)
- **NGO and Donor Reports:** Baseline or endline survey reports, Annual or programmatic monitoring reports, Project evaluations
- **Academic and Research Publications:** Peer-reviewed journal articles, University studies, Research theses, Case studies.
- **Administrative Data:** Health facility aggregate reports
- **Census and Survey Data:** Demographic and Health Surveys (DHS), National Population Census, Botswana AIDS Impact Survey
- **Media and News Reports:** Newspaper archives, radio and TV reports, Online news platforms
- **Internal Organizational Records:** Past project data, Strategic plans and reports

Unit 2: Data Management Life Cycle

Unit 1: Data Types

Routine Data	Non-Routine
Collected Continuously	Collected at intervals
Collected as part of an ongoing program reporting systems or service points	Surveys, Populaton Census, Research studies

Unit 2: Data Management Life Cycle

There are six key processes in routine data management and reporting systems

- Source
- Collection
- Collation
- Analysis
- Reporting
- Use



Unit 3: Data Quality Management

Practices to maintain high-quality information

- Collection
- Distribution
- Analysis
- Storage

Ensures insights are credible and actionable

Data Quality

Attributes used to measure how trustworthy and usable data is

- Completeness- all the required data
- Uniqueness- no duplicates
- Timeliness – data availability
- Validity- conformity to rules
- Accuracy – data correctness
- Consistency- uniform across database



Data Quality Dimensions

Attributes used to measure how trustworthy and usable data is

- Completeness- all the required data
- Uniqueness- no duplicates
- Timeliness – data availability
- Validity- conformity to rules
- Accuracy – data correctness
- Consistency- uniform across database

