GRASSROOTS DATA COLLECTION

INCLUSIVE LABOR INSTITUTE | GLOBAL FAIRNESS INITIATIVE

GRASSROOTS DATA COLLECTION

The Inclusive Labor Institute (ILI) is a worker-led, Global South-based knowledge center on the conditions and experience of work for the 2 billion+ essential workers who power the informal economy and the future of work. ILI provides grassroots worker organizations and grasstops advocates a platform for sharing information and collaborating on opportunities to improve the standing, and strengthen the voice of informal workers. ILI empowers workers by expanding knowledge through a data-driven, grassroots-led approach and partners with organizations throughout the Global South to provide a comprehensive understanding on the conditions of informal worker, especially for women workers. Launched by the Global Fairness Initiative (GFI) and a coalition of grassroots partners, ILI provides a platform for engagement and access to data and information for Global South and Global North organizations alike. This includes online training tools, digital technology, and a catalogue of Institute-led and partner-sourced studies. Through the Inclusive Labor Institute, GFI give individuals and organizations tools, information and a collaborative space to advance labor rights, women's empowerment and social and economic progress so that communities of promise can become centers of prosperity.

ACKNOWLEDGMENTS

This report was developed through the GFI's Grassroots Center on Women in Work, a pilot initiative supported by the Bill & Melinda Gates foundation aimed at strengthening the rights and economic empowerment of women in the informal economy. This report was created in partnership with Dataworld Research Ltd, GFI program Director, Asha Metcalf, GFI Program Manager, Anais Dicrose, with support from Sharifah Mwajabu. GFI would like to thank our local partners for their continuous support, guidance, and invaluable insights throughout the course of this curriculum development.

Authors
This report was written by Caleb Ongong'a, Dr. Andrew Dibo, Zipporah Kitambulisho, Scovia Odhiambo, Ronny Otieno and Vera Achieng' .A thanks to Dataworld Research Limited (DRL) who were instrumental in composing, writing, and testing the contents of this report.

Disclaimer
The views expressed in this report are those of authors and GFI. They do not represent those of the institutions referred to in the report. All errors and omissions remain the responsibility of the team.

15 June 2025

Table of Contents

Contents

CHAPTER 1.

Grassroots Data Model and How it Aligns with the Data Collection Process

Section 1: Why Data Matters, and Its Impact on Decision-Making and the Future of Work

1.1 Introduction

This module introduces the participant to various types of data such as grassroots data and digital data. The participant will also gain an understanding of the importance of data and how it influences decision-making at the grassroots level. And finish by understanding areas where grassroots data can be used and ethical considerations in data collection.

1.2 Training Objectives

By the end of this session, the participants should be able to:

a. Understand digital data, methods of collecting big data and their limitations;
b. Define grassroots data model;
c. Determine the importance of data in decision making;
d. Give examples where grassroots data can be applied.

1.3 Definition of terms

a. **Data** simply means raw facts, figures or information collected for analysis, reference or decision making. Data is a crucial tool for decision making especially at the grassroots level, where communities and organizations rely on information to drive sustainable development, improve livelihoods and create inclusive opportunities.

b. **Grassroots data model** is a structured approach to collecting, managing, and utilizing data at the local or community level to inform decision-making, improve service delivery, and empower communities. It emphasizes participatory, inclusive and content specific data processes that reflect the unique needs and realities of the local population.

c. **Digital data or 'Big data'** is the information that is stored, processed and transmitted in digital form. It is collected through various digital means including the internet, mobile devices and databases. It is the foundation of modern information systems, analytics and decision making.

d. **Research:** Is a systematic process of collecting, analyzing and interpreting data to generate new knowledge, validate existing theories or solve a specific problem. Research is essential in decision making, policy formulation and advancing knowledge in various fields.

1.4 Methodology: Group discussions, lectures, brainstorming, exercises, and stories.

1.5 Duration: 20 Minutes

1.6 Materials: Felt pens, masking tape, flipchart, flipchart board/black board, chalk, computer, projector

1.7 Content

a. Why data matters and its Impact on Decision Making
b. Types of Data
c. Grassroots data
d. Importance of data at grassroots level
e. Digital Data
f. Methods of collecting digital data
g. Ethical considerations when undertaking grassroots data collection

1.8 Training Content

1.8.1 Why Data Matters, and its Impact on Decision-Making

In today's world, data is a powerful asset that influences policies, businesses and community development. When used effectively it helps organizations, individuals and governments in making informed, strategic and impactful decisions. Some of the reasons why data matters include;

- **Data provides evidence-based insights:** Data helps decision makers to make decisions based on real world evidence instead of relying on assumptions and intuitions.
- **Data helps to identify trends and patterns:** Real time data helps organizations to anticipate market shifts, economic trends and social changes.
- **Community empowerment:** Relevant and localized data helps communities to advocate for their rights, demand services and develop grassroot solutions.
- **Data supports accountability and transparency:** Public access to data promotes fairness, equity and transparency in governance and business.

1.8.2 Different types of data

There are different types of data and can be categorized based on its nature, and method of data collection. This includes:

a) Based on Nature

- Qualitative data (Descriptive): Qualitative data is non-numerical and captures subjective aspects like opinions, experiences and narratives.
- Quantitative data (Numerical): Quantitative data is expressed in numbers and can be measured.

b) Based on Source

There are two sources of data i.e. primary (firsthand data collected by researchers for a specific purpose) and secondary data (is data that has already been collected and available from other sources).

NOTE: More information on types of data is presented in chapter 2.

1.8.3 Grassroots data

Grassroots data is community driven, locally collected data that is gathered by local organizations, activists or researchers. It provides first-hand insights into the experiences, challenges and needs of the marginalized communities often filling gaps left by large scale, top-down data collection efforts.

1.8.4 Importance of Data at the Grassroots Level

Data serves as a powerful tool at the grassroots level for decision-making, planning and development. It enables communities, organizations and policy makers to make informed choices based on real needs and evidence. Proper data collection and analysis helps in understanding local challenges, tracking progress and improving interventions as discussed below.

1. **Informed decision making**: Data helps local governments, NGOs, and community groups to identify priority areas and design programs that directly address the needs of the community. Data also ensures that resources are distributed efficiently to areas that require them the most.
2. **Improved service delivery**: Real-time data helps in tracking progress and evaluating the impact of development projects. It also ensures that social programs reach the right people and achieve their intended goals.
3. **Strengthening community voices**: Data amplifies the needs and concerns of marginalized groups leading to more inclusive policies. Data also helps in advocacy efforts, enabling communities to demand for accountability and transparency.
4. **Empowering community participation**: Grassroots data involves community participation which fosters ownership and engagement, leading to better buy-in for projects and policies.
5. **Accurate understanding of local context**: Community level data captures nuances and specific situations that might be missed by broader analysis, allowing tailored solutions to local problems.
6. **Monitoring and evaluation**: Data helps to track progress on development initiatives. It allows the stakeholders to measure progress, assess effectiveness, and make necessary adjustments to improve outcomes.

1.8.5 Examples Where Grassroots Data Can Be Used

Here are some of the examples where grassroots data can be used;

- **Economic development:** Analyzing income levels and economic trends of informal workers to improve their economic status

- **Education:** Assessing learning outcomes in technical schools to identify areas requiring additional support for skill acquisition
- **Environmental monitoring:** Collecting data on climate change to identify how it affects agricultural production
- **Health:** Grassroot data can be used to track members who have access to health insurance i.e. physical and mental health, as well as safety measures at the work place.

1.8.6 Digital Data

Digital data is data that is collected, stored, processed and transmitted in digital form, and can be collected through direct and indirect means.

1.8.6.1 Methods of collecting digital/big data

Direct method: This is where data is collected with the knowledge, consent and participation of the user. Examples include:

- Online surveys
- Social media interactions, e.g. user generated interactions on platforms like Facebook, X, etc.

Indirect method: Data is collected without the user's knowledge. Examples include:

- Transactional data, e.g. digital payment, financial transactions, etc.
- Location and mobile data, e.g. GPS tracking from mobile devices and transportation Apps
- Web & App tracking: tracking browsing behavior via cookies and analytics

1.8.7 Ethical Considerations when Undertaking Grassroots Data Collection

As with any research undertaking, it is essential for the researcher to uphold ethical research practices that safeguard the rights, dignity, and welfare of participants, while also ensuring the credibility and validity of the findings. This includes obtaining informed consent, ensuring voluntary participation, minimizing harm, and maintaining data confidentiality and anonymity. Additionally, researchers must ensure that data is securely stored and handled with integrity throughout the research process.

NOTE: Chapter 4 of this report offers a comprehensive overview of the key ethical considerations involved in conducting research.

1.9 Resources

1. https://sustineo.com.au/sites/default/files/news/documents/ideas_in_brief_7_pankhurst.pdf
2. https://www.wvi.org/sites/default/files/Grassroots%20to%20Global_Report_FINAL.pdf

3. https://www.opensocietyfoundations.org/uploads/de7781bb-c2a5-4f1a-84ca-7143447ad2e6/grassroots_08022012.pdf.

1.10 Assessment

1. Which of the following statements best describes grassroots research?

a. Research conducted by academic institutions to address global issues
b. A community led research to address local challenges
c. A top-down approach where experts design, and collect data
d. A research that focuses on urban population to address their needs

2. **Which of the following is not digital data?**

a. Social media e.g. Facebook interactions
b. Online transactions
c. Location and mobile data Apps
d. Key Informant Interviews

3. **Why is grassroots data important?**

a. It improves service delivery
b. It promotes biases in fund allocation
c. It promotes discrimination among the marginalized communities
d. It does not allow partnerships and collaborations

4. **Which one of the following is a method of collecting digital data?**

a. Key informant interviews
c. Surveys and questionnaires
c. Direct method and indirect method
d. Focus group discussion

5. **Which one of the following is not an example where grassroot data can be used?**

a. Economic development
b. Education sector
c. Informal sector
d. For national security and defense

Section 2: Conventional Data Collection Methodologies and Life Cycles

2.1 Introduction

This module introduces conventional data collection methods and their application. It equips the participant with knowledge of structured data gathering techniques used in research, studies and surveys. The participant will further look at data manufactured and managed by developed countries institutions and examine how unequal representation can influence decision making and we will conclude by looking at the data lifecycle.

2.2 Training Objectives:

By the end of this session, the participants should be able to:

a. Identify and describe different conventional data collection methods;
b. Understand when and how to apply each method;
c. Describe data managed by developed countries and its limitations;
d. Understand the impact of unequal representation;
e. Explain conventional data collection life cycles.

2.3 Definition of terms

a. **Developed countries** are nations with high economic outputs advanced infrastructure and living standards.
b. **Developing countries** are nations that have growing economies but face challenges in infrastructure income inequalities, and social services.
c. **Data colonialism:** It is the practice of collecting and owning data from users without their consent or knowledge and using the same data to exploit and control them.
d. **Conventional data collection method:** Conventional data collection methods refer to modern methods used to gather information from various sources. These methods are structured, systematic and widely used in research studies.
e. **Raw data (Also referred as primary data):** This is the information/data that is collected directly from a source before its cleaned or analyzed.

2.4 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.5 Duration: 30 min

2.6 Materials: Felt pens, masking tape, flipchart, scissors, flipchart board, newsprints, black board, chalk, posters, computer, projector

2.7 Contents

a. Conventional data collection methods
b. Types of conventional data collection methods
c. Developed countries dominated data systems and its limitations
d. How to ensure equal voices and inclusive decision making

e. Data lifecycle

2.8 Training Content

2.8.1 Types of Conventional Data Collection Methods

These methods include:

- **Document review:** This includes reviewing existing documents, reports or records to extract relevant information.
- **Surveys and questionnaires:** Surveys are a structured set of questions used to gather information from individuals.
- **Interviews:** Interviews are a one-on-one conversation where a researcher asks a respondent questions to get detailed insights. They include, Focus Group Discussions and Key Informant Interviews (KII).
- **Observations:** Here, the researcher watches and records behavior, events or interactions in a natural setting.
- **Case studies:** Case studies are an in-depth examination of an organization or individual to understand their experiences, and/or success as a result of a program.

These methods will be discussed in depth in the next module.

2.8.2 Data Manufactured and Managed by Developed Countries Institutions

Data manufactured and managed by Developed Country Institutions leads to data colonialism, which in turn hurts grassroots organizations. To grow, there must be a

shift from the developed country systems to a more decentralized community led data collection systems to ensure there is equality, equity, and meaningful impact in research and policy making and decisions. Below are limitations of relying on the developed countries dominated data systems.

2.8.3 Limitations of Developed Countries Data Dominated Systems

Developed countries often collects, analyzes and manages data from the developing countries which leads to biases, inequalities and misinterpretation of grassroot organizations. These issues include:

- **Bias:** Data collected by institutions from developed countries often reflect their priorities, perspectives and methodologies which do not align with the local reality.
- **Limited local ownership:** Decision making at the grassroots level is external without meaningful participation of the local communities because they lack control over data.
- **Limited resources and technological gaps:** Grassroots organizations rely on developed countries to produce data because they lack funding, technical expertise and infrastructure to collect, analyze and manage their own data.
- **Funding and decision-making disparities:** Donors and policy makers in the developed countries base their funding decisions on their indicators which may not reflect the realities on the ground, which can lead to inequalities in resource distribution and misalignment of priorities.

2.8.4 How to Ensure Equal Voices and Inclusive Decision Making

True resilience and sustainable development can only be achieved when all voices are heard at the table. Actively including formal and informal workers, men, women, youth, Persons with Disabilities (PWDs), and other marginalized groups ensures that development moves from ineffective, top-down solutions to community-driven transformative change. This can be achieved through encouraging:

- Participatory policy making
- Gender responsive and youth inclusive strategies
- Empowering grassroot leadership, and
- Funding grassroot initiatives

2.8.5 Data Lifecycle

Data goes through different stages from collection to its final use. This structured cycle ensures that data is accurate, accessible and actionable throughout its lifecycle. The conventional data lifecycle consists of:

a. **Planning and designing**: This stage ensure that the study is well structured, methodologically sound and capable of generating reliable insights. This phase lays the groundwork for efficient data gathering and meaningful

analysis. It involves defining the research problem statement, identifying the population and the data collection method.

b. **Data collection**: Data collection process is a critical step in any research. It involves gathering information from various sources such as surveys, interviews and observations in order to analyze and draw conclusions. To ensure the success of the data collection process, it is important to carefully plan and organize the methods and tools that will be used. This includes determining the target population, selecting appropriate data collection techniques and establishing a timeline for gathering and analyzing the data.

   This process is also essential in maintaining the integrity and confidentiality of the data collected. This includes ensuring that data is collected ethically and securely and that it is stored and managed in a way that protects the privacy of individuals.

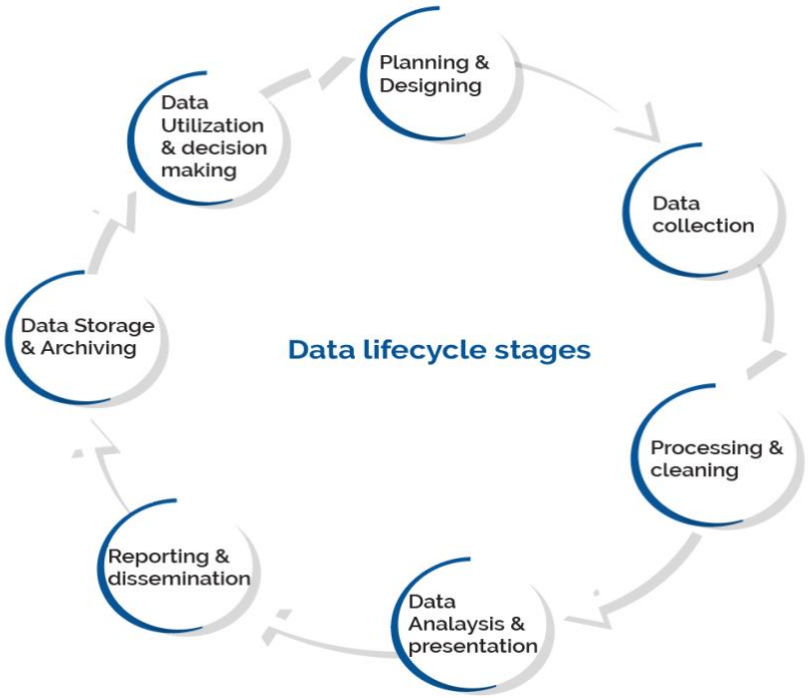
c. **Data processing and cleaning:** This process involves removing inconsistencies, errors or missing values in the dataset to ensure accurate and reliable results. It also involves checking typos, misclassification and ensuring collected data is consistent.

d. **Data analysis and interpretation**: This is a crucial stage where raw data is transformed into meaningful information for decision making. It involves arranging data into tables, graphs or databases for easy access, and summarizing the descriptive findings in order to draw conclusions and make recommendations.

e. **Reporting and dissemination**: This stage involves the communication and sharing of data findings and insights with relevant stakeholders. During this stage, it is important to consider the audience and tailor the presentation of the findings accordingly. This may involve the use of visual aids such as charts, diagrams and tables to convey complex information in a more accessible format. Additionally, feedback from stakeholders should be sought out to ensure data is meeting their needs and addressing any concerns they have.

f. **Data storage and archiving**: Data storage and archiving ensures that data remains accessible, secure and well managed for future use. Data can either be stored on cloud using remote servers that provide scalable storage solutions (e.g. Google Drive, One Drive, etc.), or using physical devices like hard drives, USB and SSD. Other storage options include storing the data on shared storage systems within an organization and databases like SQL or data warehouses. Stored data will require backups, access control and compliance with regulations.

g. **Data utilization and decision making**: Data utilization is the last stage of data lifecycle. This involves using collected, processed and analyzed data to inform actions, strategies and policies. It ensures that insights derived from data drive effective decision-making at all levels.



2.9 Resources

1. https://www.surveycto.com/resources/guides/data-collection-methods-guide/#:~:text=Surveys%2C%20interviews%2C%20observations%2C%20focus,and%20relevant%20data%20is%20crucial.
2. https://www.researchgate.net/publication/325846997_METHODS_OF_DATA_COLLECTION
3. https://www.researchgate.net/publication/11842651_Data_Collection_Techniques
4. https://www.researchgate.net/publication/383155577_Data_Collection_Fundamentals_A_Guide_to_Effective_Research_Methodologies_and_Ethical_Practices
5. https://www.sciencedirect.com/science/article/pii/S1877050920315465

2.10 Assessment

1. Which of the following is not a limitation of data managed by developed countries?

a. Limited local ownership
b. Biasness
c. Funding and decision-making disparities

d. Promotes equality in resource allocation at the grassroot level

2. **Which of the following is not a data lifecycle process?**

    a. Data collection
    b. Data analysis and presentation
    c. Data collection using surveys
    d. Data storage and archiving

3. **What is the importance of data cleaning and processing in the data lifecycle?**

    a. To reduce the bulk of data collected
    b. To ensure data collected is consistent
    c. To delete unnecessary information
    d. To add information that was not collected

4. **How can organizations ensure that grassroots voices are included in the decision-making process?**

    a. By ensuring there is inclusive participatory in decision making
    b. By funding individuals to do business
    c. Promoting patriarchy
    d. Electing the right leaders to represent the marginalized during decision making

5. **Which of the following is not a conventional data collection method?**

    a. Surveys and questionnaires
    b. Interviews and observations
    c. Case studies and experiments
    d. Document review and data analysis

Section 3: Grassroots Research-Worker Led, Bottom-Up, and Participatory [Decentralized] Approach

3.1 Introduction

This section introduces participants to data driven approaches. They will also gain understanding of grassroots research approaches, and key principles of participatory approaches.

3.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Define concept on grassroots research approach;
b. Describe -worker led, bottom-up research approaches;
c. Explain participatory research [Decentralized] approach;

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration: 20 Min

3.5 Materials: Felt pens, masking tape, flipchart, scissors, flipchart board, newsprints, black board, chalk, posters, computer, projector

3.6 Contents:

a. Grassroots research approaches
b. Participatory approach
c. Key principles of participatory research

3.7 Training Content

3.71 Grassroots Research Approaches

Grassroots research refers to research that originates directly from within a community or group of people from a local level. It emphasizes inclusive, community driven knowledge production, ensuring that local perspective, experiences and expertise shape the research outcomes. Unlike top-down research approach, bottom-up research ensures that solutions are locally relevant, and empowering communities to take ownership of the process and its outcomes.

Grassroots research is also known as “Community-Based Research”, “Bottom-Up Research” or “Participatory Research”.

3.7.2 Participatory Approach

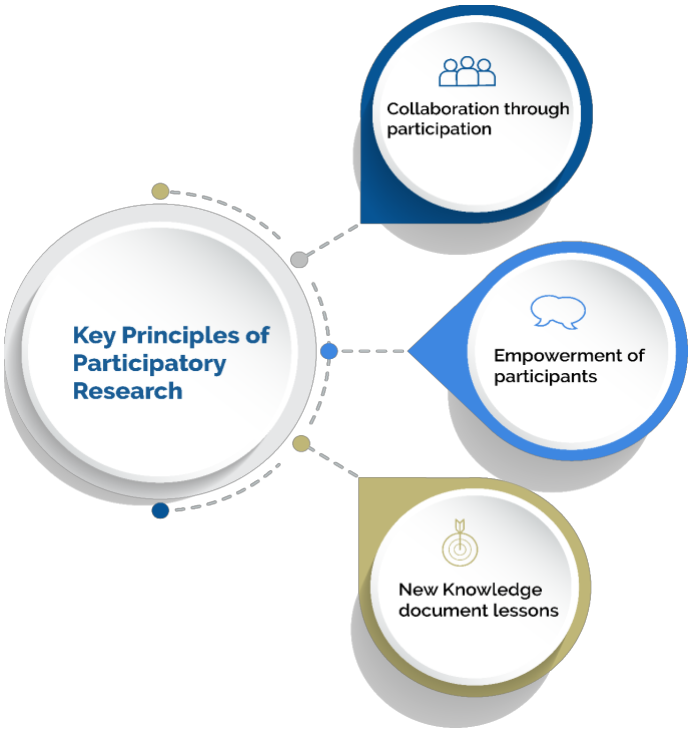
Participatory research is an approach that actively involves community members, stakeholders, and beneficiaries in the research process. Unlike the traditional research methods where data is collected by external researchers and analyzed

separately, participatory research emphasizes on collaboration, shared learning and empowerment of local communities

3.7.3 Strategies for Grassroots Data Collection

3.7.4 Key Principles of Participatory Research

1. **Community involvement:** local communities are actively engaged in designing, conducting and analyzing the research findings.
2. **Co-learning:** Participants share knowledge, ensuring mutual understanding.
3. **Empowerment:** The process strengthens the capacity of the community members to address their own challenges.
4. **Inclusivity and equality:** All voices, especially marginalized groups, are considered in the research process.



3.8 Resources

1. https://ijoc.org/index.php/ijoc/article/download/893/458
2. https://www.undp.org/sites/g/files/zskgke326/files/migration/acceleratorlabs/Grassroots-Innovation---An-Inclusive-Path-to-Development.pdf
3. https://www.tandfonline.com/doi/full/10.1080/09614524.2021.1937554#d1e177

3.9 Assessment

1. Which of the following statements best defines grassroots research?

a. *It is research conducted by international organizations on grassroot research*
b. Studies conducted by government agencies on local development
c. It's a type of research that uses quantitative data
c. A community-driven research focusing on local challenges and solutions

2. **What is a major challenge of grassroots research?**

    a. Limited access to funding and resources
    b. Heavy reliance on external data
    c. Lack of relevance to community issues
    d. Low engagement from stakeholders

3. **Which of the following is not a key principle of participatory research?**

    a. Inclusivity and equality
    b. Community empowerment
    c. Community involvement
    d. Labor division

4. **Which of the following is a key feature of grassroot research?**

    a. It solely relies on academic research
    b. It excludes indigenous knowledge in favor of external expertise
    c. It engages local communities in research design and implementation
    d. It prioritizes top-down decision making

5. **How can you best describe how grassroot research can impact policy?**

    a. It removes the need for government intervention
    b. It highlights community specific needs that needs inform to policy adjustment
    c. It replaces traditions scientific research methods
    d. It provides generalized findings applicable to all regions

Section 4: Organizational Capacity for Data Collection

4.1 Introduction

In this section, the participant will learn about organizational capacity for data collection, challenges in building data collection capacity and strategies for strengthening organizational capacity.

4.2 Training Objectives:

By the end of this Session, the participant should be able to:

a. Describe the organizational opportunities;
b. Understand the gaps and challenges in grassroots research.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 20 min

4.5 Materials: Felt pens, masking tape, flipchart, scissors, flipchart board, newsprints, black board, chalk, posters, computer, projector

4.6 Contents:

a. Organizational capacity for data collection
b. Challenges in building data collection capacity
c. Strategies for strengthening organizational capacity

4.7 Training Content

4.7.1 Organizational Capacity for Data Collection

Grassroots data collection requires strong organizational capacity to ensure accuracy, efficiency and ethical integrity. Below are key elements that determine an organization's ability to collect and manage data effectively at the grassroots level. They include:

i. **Institutional readiness**: Institutions should develop clear objectives on data collection that align with the organization's mission and strategy goals. They also need to develop a strong leadership that can prioritize data collection and ensure ethical practices are adhered to.

ii. **Human resource capacity:** The organization should ensure that personnel are well trained and equipped with skills in data collection methods, survey administration and digital tools. This can be achieved through regular capacity building programs to enhance technical and analytical skills.

iii. **Technological and infrastructure support:** There is a need to use appropriate technologies such as Mobile-Based Data Collection Apps and ensure that data is well stored to prevent data loss and breaches.

iv. **Financial and logistical capacity:** The organization should be able to provide adequate funding for fieldwork, equipment, training, and personnel. Streamline logistics to optimize costs and maximize efficiency and collaborate with donors, government agencies, and research institutions for sustainability.

v. **Monitoring, evaluation and learning:** The organization should be able to develop a real-time monitoring system to track the data collection progress, share findings with stakeholders and communities for validation and action and use the collected data to refine methodologies and improve future efforts.

4.7.2 Challenges in Building Data Collection Capacity

a) Limited funding and resources: Limited funding and resources can impact the quality, accuracy and efficiency of data collection efforts.
b) Shortage of trained personnel in data collection: They would affect the quality, accuracy, and reliability of data across different sectors.
c) Technological barriers: Poor access to technology can affect accuracy, efficiency and accessibility of data. This may include poor internet connection in rural and remote areas, and/ lack of digital infrastructure such as reliable power supply and proper devices.
d) Community mistrust: In situations where community members do not trust the process and therefore may not provide accurate information or refuse to participate in the data collection process. This may be due to lack of transparency, misuse of personal data etc.

4.7.3 Strategies for Strengthening Organizational Capacity

i. **Invest in training and capacity building:** Conduct regular workshops on data methodologies and ethics.

ii. **Enhance collaborations and partnerships:** Engage and create more partnerships with other grassroot organizations to improve resource and knowledge sharing.

iii. **Develop a culture of data use**: Organizations should encourage evidence-based decision-making at all levels.

4.8 Resources

1. https://www.ncjfcj.org/wp-content/uploads/2021/11/Building-Organizational-Data-Capacity.pdf
2. https://msh.org/wp-content/uploads/2015/09/2015_08_msh_organizational_capacity_assessment_tool_ocat.pdf
3. https://tools4dev.org/blog/organizational-capacity/

4.9 Assessment

1. Which of the following is not a key requirement for organizational capacity to collect data?

a. Number of years the organization has been in existence
b. Finances
c. Data storage and security systems
d. Trained personnel

2. **An organization wants to improve its data collection process. What is the first step they need to do?**

a. Purchase an expensive data collection software
b. Conduct a capacity assessment to identify strengths and gaps
c. Eliminate manual data collection methods
d. Collect random data to see what works best

3.**Which of the following factors can contribute to an organization's ability to collect high quality data?**

a. Use of data collection tools from other organizations
b. Avoid stakeholder engagement
c. Availability of trained staff and resources
d. Only use secondary data

4. **How can organizations strengthen their capacity to collect and manage data?**

a. Invest in training and capacity building
b. Involve research expertise to collect and manage data
c. Exclude the marginalized during the study
d. Avoid data validation

5. **Which of the following best describes data governance in an organization**

a. The process of collecting data from different sources
b. It is an exclusive responsibility of the ICT department
c. It is a system of rules, policies and procedures to ensure data quality and security
d. It is the elimination of traditional records

CHAPTER 2.

Primary and Secondary [qualitative and quantitative] Data Collection Methods

Section 1: Data Methodologies Overview

1.1 Introduction

This module focuses on the principles, techniques and best practices in data collection. It explores both qualitative and quantitative data collection methods and provides guidance on selecting appropriate methodologies for different research contexts.

1.2 Training Objectives:

By the end of this modules, the participant should be able to:

a. Understand the difference between qualitative and quantitative data collection methods;
b. Differentiate between primary and secondary data collection methods;
c. Understand the importance of different methodologies in research.

1.3 Definition of Terms

a. **Qualitative data:** is non-numerical information that describes characteristics, qualities or attributes. It is used to understand concepts, thoughts and experiences in-depth to provide subjective insights. It is about gauging people's feelings, emotions and attitudes.

b. **Quantitative data:** is numerical information that can be measured, counted and analyzed statistically. It focuses on quantities, frequencies and patterns to provide objective insights.

c. **Primary source:** are original, firsthand information collected directly from the source without any prior interpretation. They provide raw and unfiltered data for analysis and decision making.

d. **Secondary source:** refers to information that has been collected and analyzed by someone else for a different purpose. The sources can include articles, textbooks, journals, etc.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 20 mins

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents

a. Overview of data collection methodologies
b. Qualitative data collection methods
c. Quantitative data collection methods
d. Difference between qualitative and quantitative data collection methods
e. Primary and secondary data

1.7 Training Content

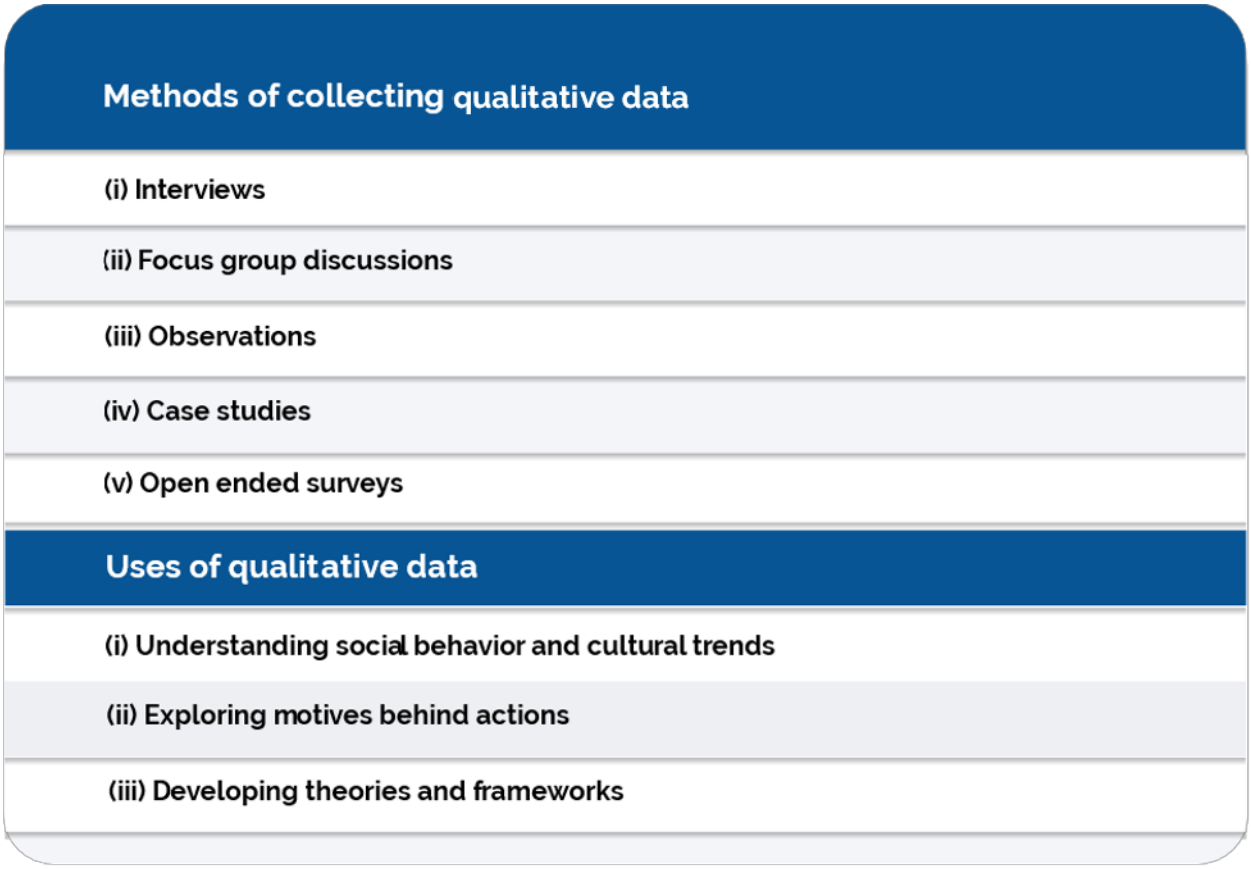
1.7.1 Data Collection Methodologies Overview

Data methodologies refer to structured techniques used to collect, analyze and interpret data. Over centuries, researchers have categorized data into either qualitative or quantitative. The nature of data collected is dependent on the purpose of the study, population and sample of the subject, and the problem focus of the research. Data as mentioned in the passage means values of qualitative or quantitative variables, belonging to a set of items.

A well selected research methodology ensures clarity, precision and relevance in achieving research goals, and shapes the study's direction, data quality and credibility while determining how well results align with the hypothesis or objectives of the study. On the other hand, poor methodology can introduce bias, leading to misleading conclusions.

1.7.2 Qualitative data

Information is **qualitative** when it is not countable or measurable. *For instance, the attitude of an individual, color, sex, qualification, etc*. are good examples of qualitative information. These intangible concepts are merely described or explained.



1.7.3 Quantitative data

Unlike qualitative data, **quantitative** information is countable and measurable. This enables the correct values to be obtained. *For example, number of acres owned by an individual, revenue generated per month, number of hospitals in an area, etc*.

Methods of collecting quantitative data
(i) Surveys and questionnaires
(ii) Experiments
(iii) Structured observation
(iv) Census and administrative data
Uses of quantitative data
(i) Identifying trends and correlations
(ii) Making evidence-based policy decisions
(iii) Conducting impact assessment
(iv) Measuring economic and social indicators

1.74 Difference between qualitative and quantitative data collection methods

Difference between qualitative and quantitative data collection methods	
Qualitative	Quantitative
Descriptive and Subjective	Numerical and Measurable
Capture detailed insights, emotions & behaviors	They are objective & standardized
Can be measured in numbers but categorized	Can be analyzed using statistical tool
Collected using open ended & explanatory methods	Collected using structured methods

1.7.5 Primary and Secondary Data Collection Methods

A lot of times, studies are geared towards achieving a specific purpose. These studies may involve the individuals concerned creating their own data. Other times, there may be no data that exist to assist the researcher, or the information available may not be adequate for the study. This may necessitate the generation of data for the purposes of the study. This first-hand information collected is what is known as primary data. There are different methods of collecting primary data and they allow data to be collected in line with the study needs.

On the other hand, secondary data collection methods involve gathering existing data that had been collected and processed by individuals or organizations and only extracted for use by present researchers.

NOTE: Types of primary and secondary data collection methods are discussed in the subsequent sections

1.8 Resources

1. Johnston, M. P. (2014). Secondary data analysis: A method of which the time has come. Qualitative and quantitative methods in libraries, 3(3), 619-626.
2. Ajayi, V. O. (2017). Primary sources of data and secondary sources of data. Benue State University, 1(1), 1-6.
3. Taherdoost, H. (2021). Data collection methods and tools for research; a step-by-step guide to choose data collection technique for academic and business research projects. International Journal of Academic Research in Management (IJARM), 10(1), 10-38.

1.9 Assessment

1. **Which of the following is not a method of collecting qualitative data?**
   - *a. Focus group discussion*
   - *b. Key informant interviews*
   - *c. Observations*
   - *d. Surveys and questionnaires*
2. **How can you best describe quantitative data collection method?**
   - *a. It is collected using structured method*
   - *b. It is relatively subjective*
   - *c. It is collected from key informants and focus group discussions*
   - *d. It does not add value in research*
3. **Which statement best describes primary data sources?**
   - *a. They can be collected from journals and textbooks*
   - *b. They provide first hand and original information*
   - *c. They lack credibility and are unreliable*
   - *d. They only contain quantitative data*
4. **Which of the following is not a secondary data source?**
   - *a. Literature review*
   - *b. Newspaper articles*

c. Government reports
d. Surveys

5. What is the main advantage of using qualitative data collection method?

a. They provide numerical data for statistical analysis
b. They allow in-depth understanding of a subject
c. They focus on structured responses
d. They are an alternative source of secondary data

Section 2: Use of Surveys

2.1 Introduction

In this session the participant will learn surveys as data collection methods and explore various types of surveys. Additionally, the participant will be introduced to advantages and disadvantages of using surveys.

2.2 Training Objectives

By the end of this session, the participant should be able to:

a. Understand different types of surveys as data collection method;
b. Understand the advantages and challenges of using surveys in data collection.

2.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.4 Duration: 10 Min

2.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents:

a. Overview of surveys
b. Advantages of using surveys
c. Challenges of using surveys

2.7 Training Content

2.7.1 Overview of Surveys

Surveys are a primary data collection method used to gather information from individuals, groups or communities. They involve a structured or semi-structured set of questions aimed at understanding opinions, behaviors, preferences or factual information. Surveys can be conducted in different ways depending on the target population, research objectives and available resources. Common ways of conducting surveys include:

- **Face-to-face surveys:** These are conducted in person by an interviewer. They can either be paper based or mobile supported. Examples include household surveys, individual surveys, enterprise/business surveys, etc.
- **Through the telephone:** Telephone surveys are conducted over the phone and are used to reach respondents remotely. They are cost effective but have lower response rates. Examples include customer satisfaction surveys.
- **Online surveys:** Which are conducted via web-based platforms e.g. Google Forms. They are fast and convenient for large population. For example, employee feedback.

Example of structured question

Which source do you rely on for news update?

a. *Television/radio*
b. *Social media*
c. *Newspapers*
d. *Online news websites*

Example of semi structured question

What challenges do you face in accessing farm inputs?

Semi-structured questions allow for more discussions/opinions on the subject.

2.7.2 Advantages of Surveys

1. They can reach a large number of respondents
2. They provide both qualitative and quantitative data
3. They are cost effective compared to other research methods
4. Standardized questions ensure comparability

2.7.3 Challenges of using Surveys

1. Survey participants can filter responses
2. Researchers cannot probe in surveys
3. Surveys only collect attitudinal data

2.8 Resources

1. 10 Survey Challenges and How to Avoid Them
2. http://law.gtu.ge/wp-content/uploads/2017/02/Berg-B.-Lune-H.-2012.-Qualitative-Research-Methods-for-the-Social-Sciences.pdf

3. https://www.researchgate.net/publication/325846997_METHODS_OF_DATA_COLLECTION
4. https://uomustansiriyah.edu.iq/media/lectures/6/6_2021_02_10!11_52_26_AM.pdf

2.9 Assessment

1. **Which of the following is an example of a survey question?**
   - *a. What is your opinion on the competency-based curriculum?*
   - *b. Describe your experience with the banking sector*
   - *c. Are their gender-based violence cases in your area? Yes/No*
   - *d. How has climate change affected agricultural productivity in your area?*
2. **What is the advantage of using surveys in data collection?**
   - *a. They provide deep qualitative insights*
   - *b. They do not require a structured questionnaire*
   - *c. They allow data collection from a large sample size*
   - *d. They eliminate all forms of biases during data collection*
3. **Which one is an example of surveys?**
   - *a. Online surveys*
   - *b. Focus group discussions*
   - *c. Field trials and experiments*
   - *d. Observation*
4. **Which of the following is false about telephone surveys?**
   - *a. They have low response rate*
   - *b. They can be used to reach respondents remotely*
   - *c. They are cost effective*
   - *d. They can be used to collect secondary data*
5. **Which of the following sentiments is not true about surveys?**
   - *a. They include both structured and semi-structured questions*
   - *b. They have only semi-structured questions*
   - *c. They provide qualitative data*
   - *d. They are a primary source of data*

Section 3: Use of Key Informant Interviews

3.1 Introduction

In this session the participant will learn about Key Informant Interviews. The participant will also learn about advantages and limitations of using Key Informant Interviews in data collection.

3.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand key informant interviews;
b. Differentiate structured and semi structured interviews;
c. Understand advantages and limitations of KIIs.

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration: 20 Min

3.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents

a. Overview of Key Informant Interviews (KIIs)
b. Purpose of KIIs
c. Advantages and limitations of KIIs

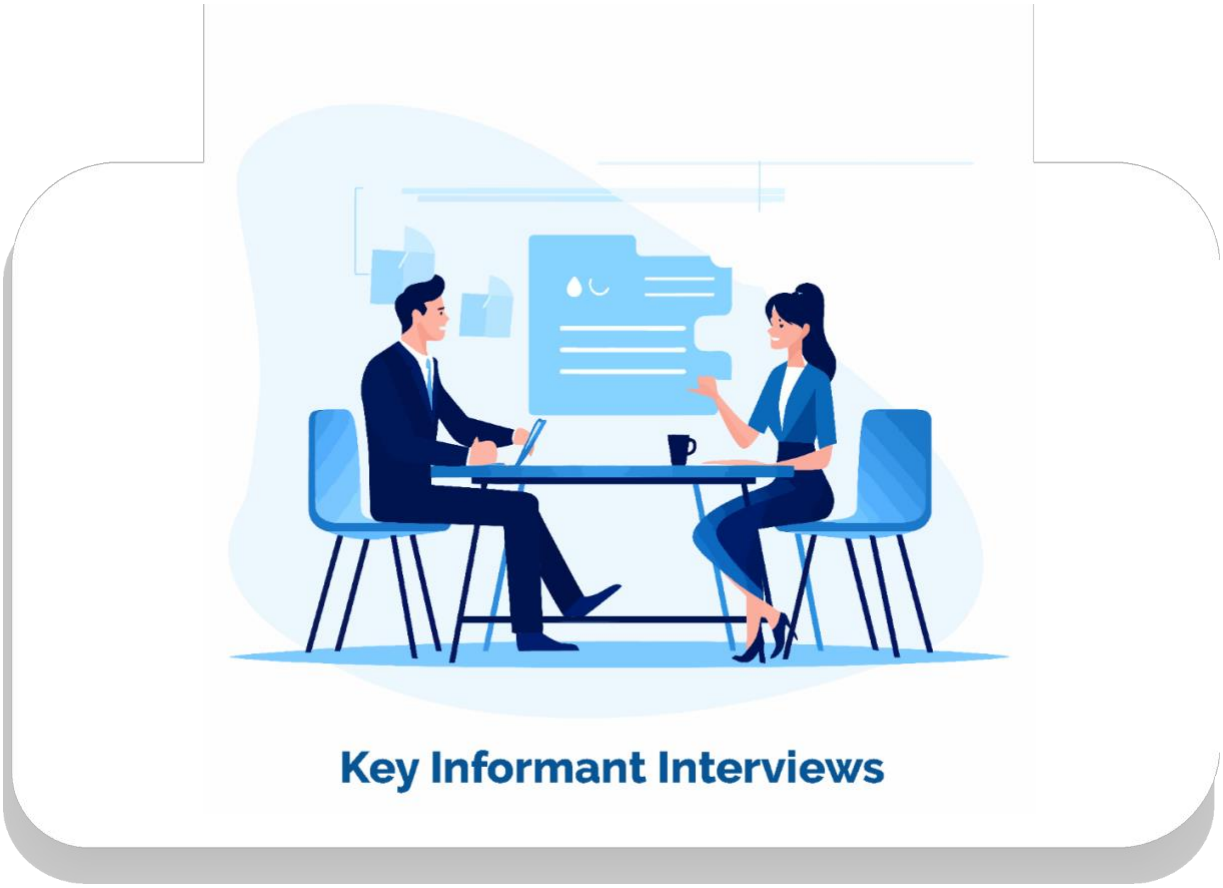
3.7 Training Content

3.7.1 Key Informant Interviews

Key informant interviews (KIIs) are qualitative, in-depth interviews conducted with individuals who have specialized knowledge or expertise on a specific topic. Key informant interviews are semi structured i.e. they use open-ended questions to encourage discussion, and the interviewer can probe deeper into the responses in order to get rich, detailed and contextual information. Interviews can either be structured or semi-structured.

I. **Structured interview**: is a systematic approach to interviewing where the same predetermined questions are asked to all candidates in the same order and are rated in a standardized scoring system. They are often closed-ended, or open-ended structured interviews.

II. **Semi-structured interview:** is a blend of structured and unstructured interviews. They have a general plan for what to be asked, but the questions do not have to follow a particular phrasing or order. They are often open-ended, allowing for flexibility, but follow a predetermined thematic framework, giving a sense of order.



3.7.2 Purpose of KIIs

The main purpose of KIIs include:

- **Gathering expert opinion:** professionals, stakeholders and community leaders have specialized knowledge that can enhance credibility and relevance of the research.
- **Identifying needs and priorities:** experts' opinions can help researchers and organizations to understand the needs of specific populations.
- **Understanding context and issues:** KIIs have knowledge, experience and perceptions that can provide deep insights into social, economic, political and cultural issues.
- **Informing program design and policy making**: KIIs can provide information that can shape policies, interventions and development programs to ensure that they align with real needs and challenges faced by the target group.

3.7.3 Advantages and Limitations of KIIs

Below are the advantages and limitations of KIIs:

Advantages and Limitations of KIIs

Advantages	Limitations
They provide deep, contextual insights	KIIs are time consuming and requires skilled interviewers
They are flexible, allowing clarification and probing	There can be potential bias from informants
Useful in exploring sensitive or complex issues	Findings may not be generalizable

3.8 Resources

1. http://law.gtu.ge/wp-content/uploads/2017/02/Berg-B.-Lune-H.-2012.-Qualitative-Research-Methods-for-the-Social-Sciences.pdf
2. https://www.researchgate.net/publication/325846997_METHODS_OF_DATA_COLLECTION
3. https://uomustansiriyah.edu.iq/media/lectures/6/6_2021_02_10!11_52_26_AM.pdf
4. https://books.kdpublications.in/index.php/kdp/catalog/download/268/262/2250?inline=1
5. https://www.measureevaluation.org/resources/training/capacity-building-resources/data-quality-portuguese/DATA_COLECTION.pdf

3.9 Assessment

1. **Which of the following statements is true about Key Informant Interviews?**
   a. *They are conducted with an individual who has special knowledge of a subject*
   b. *They are conducted with a group who have special knowledge of a subject*
   c. *They provide numeric information*
   d. *Its information gathered from social media opinions*
2. **Which of the following is a characteristic of a key informant?**
   a. *They must have a leadership position in the community*
   b. *They represent the views of the general population*
   c. *They must be willing to provide quantitative data*
   d. *They possess specialized knowledge on a specific topic*
3. **What technique is most effective in ensuring success when conducting KIIs?**
   a. *Conducting group discussions*

b. *Avoiding probe questions*
c. *Prepare an interview guide with open-ended questions*
d. *Prepare a guide with closed questions*

4. **Which challenge is one likely to encounter when conducting Key Informant Interviews?**
   a. *They are time consuming and may require skilled personnel*
   b. *The quantitative data gathered may not be easy to analyze*
   c. *The results might affect analysis*
   d. *The respondents may not able to provide secondary data required*
5. **Which of the following is not a reason why we collect information from KIIs?**
   a. *They provide information that can shape policies*
   b. *They provide deep insights into social, economic, political and cultural issues*
   c. *They provide information to understand needs of specific population*
   d. *They provide general information to help a country*

Section 4: Use of Focus Group Discussions

4.1 Introduction

In this session the participant will learn background information about Focus Group Discussions (FGDs), and their specific benefits and challenges. It will also clarify when and how to use Focus Group Discussions in data collection.

4.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Explain what is meant by Focus Group Discussion approach of data collection;
b. State the benefits and challenges of using Focus Group Discussions in data collection;
c. Discuss when and how to use Focus Group Discussions in data collection.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 10 mins

4.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

4.6 Contents

a. The background information about focus group discussion (FGDs), and their specific benefits and challenges
b. Discuss if and how to use FDGs

4.7 Training Content

4.7.1 Focus Group Discussions

Focus group discussions (FGDs) are used to gather insights from a group of people who share common characteristics or experiences related to a research topic, for example; farmers, artisans, etc. The conversations are guided to allow participants to discuss their perceptions, opinions and experiences in a structured yet open manner. The focus groups should consist of 6-12 participants to ensure active participation and should have a facilitator to guide the discussion and keep it focused, encourage free expression of ideas and allowing participants to build on each other participants' responses.

Use of FGDs enable researchers to explore complex topics through interactive dialogue, allowing for the emergence of diverse perspectives and group dynamics that may not surface in individual interviews. FGDs are particularly useful for understanding social norms, testing new ideas, and identifying community needs or

concerns. By facilitating open and guided conversations, FGDs contribute to richer data, improved program design, and more responsive policy development.

Advantages and Limitations of FDGs

Advantages	Limitations
In-depth insights and rich qualitative data	Group dynamics may influence individual responses
Group dynamics and synergy	Sensitive topics may lead to discomfort in group setting
Real-time feedback and instant reaction	They can be time-consuming to organize & analyze
Cost-effectiveness and efficiency	

Focus Group Discussions

To mitigate potential concerns about group dynamics in FGDs, it is important to create a safe, inclusive, and respectful environment where all participants feel comfortable sharing their views. The facilitator should know how to manage dominant voices, encourage quieter participants to contribute, and remain neutral to avoid influencing responses.

Establishing clear ground rules at the beginning, ensuring confidentiality, and grouping participants based on shared characteristics (e.g., age, gender, background) can also help reduce power imbalances and promote open, honest dialogue. Careful

planning and skilled facilitation are key to minimizing bias and ensuring balanced participation.

4.8 Resources

1. https://www.intrac.org/app/uploads/2017/01/Focus-group-discussions.pdf
2. https://www.eiu.edu/ihec/Krueger-FocusGroupInterviews.pdf
3. https://www.swisstph.ch/fileadmin/user_upload/SwissTPH/Topics/Society_and_Health/Focus_Group_Discussion_Manual_van_Eeuwijk_Angehrn_Swiss_TPH_2017.pdf

4.9 Assessment

1. **What is the primary purpose of conducting a Focus Group Discussion?**
   a. *To explore group opinions, experiences and perceptions on a specific topic*
   b. *To conduct a one-on-one interview with the participants*
   c. *To ensure data collected is anonymous*
   d. *To collect large scale statistical data*
2. **Which of the following is a characteristic of an effective FGD?**
   a. *It involves at least 20 – 50 participants to ensure diverse viewpoints*
   b. *Participants are selected randomly without consideration of their relevance to a topic*
   c. *It is conducted in a formal conference*
   d. *It is a structured discussion with 6-12 participants led by a facilitator*
3. **Which of the following is a challenge commonly faced in FGDs?**
   a. *Lack of diverse perspective within the group*
   b. *Inability to obtain qualitative information*
   c. *Few participants may dominate the discussion*
   d. *They provide bulk information which is difficult to synthesize*
4. **How can a researcher ensure they yield high quality data from an FGD?**
   a. *By allowing one participant to speak a time without limited time*
   b. *By avoiding controversial topics to prevent disagreements*
   c. *By using well -structured guides and allowing all participants to contribute with time limit.*
   d. *By keeping the discussion short to avoid participant fatigue*
5. **What is the role of a facilitator while conducting an FGD?**
   a. *To control the discussion and ensure participants agree on a single view*
   b. *Guide the discussion, encouraging participants to speak and ensuring all voices are heard*
   c. *Providing expert opinion and adding contribution to the discussion*
   d. *Guide the discussion and steer the group to a predefined conclusion*

Section 5: Other Primary Data Collection Methods (Non-Conventional Approaches)

5.1 Introduction

In this session the participant will learn the non-conventional approaches of data collection and how to capture either qualitative or quantitative information and their appropriateness.

5.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand the non-conventional data collection approaches and their advantages and disadvantages;
b. Know when to apply the non-conventional approaches.

5.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

5.4 Duration: 10 mins

5.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

5.6 Contents

a. Identify non-conventional approaches
b. Advantages and disadvantages of observations
c. Case studies
d. Experiments and field trials

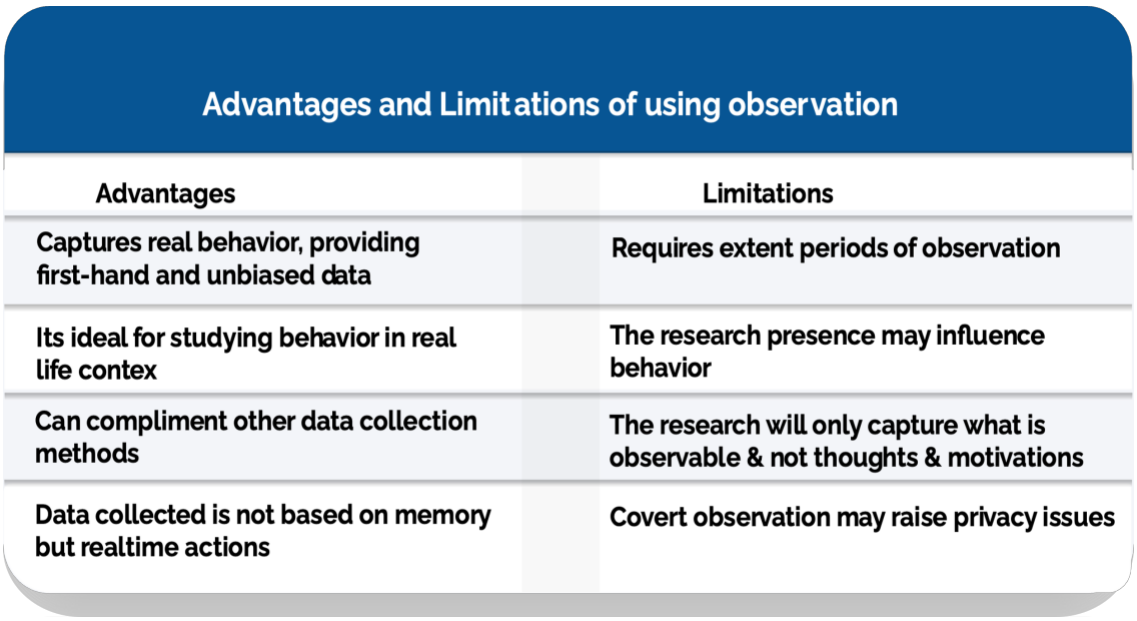
5.7 Training Content

5.7.1 Non-Conventional Data Collection Approaches

I. **Observation:** Observation is a data collection method where researchers systematically watch, record and analyze behaviors, events, or physical characteristics in their natural settings. It is commonly used in qualitative research but can also support quantitative findings. For example, observation can be used to study how customers browse a store.

Observation data collection methods have advantages and limitations as indicated below:

Advantages and limitations of using observation



Advantages and Limitations of using observation

Advantages	Limitations
Captures real behavior, providing first-hand and unbiased data	Requires extent periods of observation
Its ideal for studying behavior in real life contex	The research presence may influence behavior
Can compliment other data collection methods	The research will only capture what is observable & not thoughts & motivations
Data collected is not based on memory but realtime actions	Covert observation may raise privacy issues

II. **Case Studies:** A case study is an in-depth research method that investigates a single unit (a group, individual or organization) and may include a success story. Case studies focus on specific case and explores real world situations to understand causes and effects. It also provides a detailed narrative of a situation or phenomenon.

A case study may rely on multiple data collection methods to ensure a rich and valid understanding of the subject. These methods include case narratives where a person gives a story of a detailed account of experience related to the subject, observation and interviews.

Example of a case study
Documenting the journey of a women group who have been trained on climate smart agriculture and are producing vegetables for home consumption and sale to generate income.

III. **Experimental and Field Trials:** Experiments are field trials used to test hypotheses, evaluate interventions and measure outcomes under controlled and natural conditions. Experiments are aimed to establish cause and effect relationships between different factors.

Field trials is a research method that tests new techniques, interventions or innovations in a real-world setting, while trials take place in natural conditions and involve real participants.

Example of Experiments/ field trials
Experiments and field trials are commonly used to test interventions in real-world settings. For example, a randomized controlled trial (RCT) may be conducted to assess the effectiveness of a new agricultural technique on crop yield, where farmers

are randomly assigned to either adopt the new method or continue with traditional practices. Another example is a health intervention trial testing the impact of community health education on improving hygiene practices and reducing disease prevalence.

5.8 Resources

1. https://iris.who.int/bitstream/handle/10665/41795/0963552228.pdf
2. https://es.ircwash.org/sites/default/files/125-10840.pdf
3. https://archive.unu.edu/unupress/food2/UIN03E/UIN03E06.HTM
4. https://www.swisstph.ch/fileadmin/user_upload/SwissTPH/Topics/Society_and_Health/Focus_Group_Discussion_Manual_van_Eeuwijk_Angehrn_Swiss_TPH_2017.pdf
5. https://jet.jips.org/wp-content/uploads/Manual-FGD-Facilitators-Delhi-Example-JIPS-Phase3-JET.pdf

5.9 Assessment

1. **Which of the following is an advantage of using case studies?**
   a. *They provide in-depth understanding of a specific phenomenon*
   b. *They allow for statistical generalization to a large population*
   c. *They eliminate all forms of research bias*
   d. *They focus on numerical data collection*
2. **Which source is commonly used when conducting case studies?**
   a. *Randomized surveys using closed ended questions*
   b. *Interviews, documents and observation*
   c. *Use of social media*
   d. *Field experiments*
3. **Which of the following is an example of participant observation?**
   a. *A scientist recording temperature changes in a lab*
   b. *A researcher conducting interviews over the phone*
   c. *A researcher collecting information from social media*
   d. *A researcher joining farming community to understand their agricultural practices*
4. **What is a limitation of using observation as a data collection method?**
   a. *Observer bias can influence the interpretation of data*
   b. *Its only useful to laboratory scientists*
   c. *It does not allow researchers to study real-world behavior*
   d. *It provides quality quantitative data*
5. ***Which of the following is an example of field trial?***
   a. *Testing a new education program in technical school to measure its impact*
   b. *Studying agricultural historical records*
   c. *Interviewing experts*
   d. *Conducting a consumer survey for a new product*

Section 6: Use of Secondary Data Collection

6.1 Introduction

In this session the participant will learn Secondary Data Collection Methods, their benefits and challenges. This section also outlines when and how to use the secondary research methods

6.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand secondary data collection methods, benefits and challenges;
b. Understand when and how to use secondary research methods.

6.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

6.4 Duration: 10 mins

6.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

6.6 Contents

a. Overview of secondary data collection
b. Types of secondary data sources
c. Advantages and disadvantages of secondary data sources

6.7 Training Content:

6.7.1 Secondary Data Collection Sources

Secondary data collection sources are data that has been collected and processed by individuals or organizations and only extracted for use by present researchers. Such data are usually obtained from either published or unpublished sources and they are useful for supplementing primary data and providing background information.

6.7.2 Types of Secondary Data Sources

a. **Published sources**: They include government publications (e.g. policy documents, census reports etc.), international organization reports, books and newspapers, academic research and journals.
b. **Institutional and organizational records**: Include company reports, NGOs and Civil Society reports and educational institutional and research centers reports.
c. **Digital and online sources**: Incorporate websites and online databases, social media and blogs and open data platforms.

6.7.3 Advantages and Disadvantages

Advantages and Disadvantages

Advantages	Disadvantages
They save time and resources compared to collecting primary data	Existing data may not fully, match the research needs
They provides access to large scale national, or global data	Errors in original data collection may lead to misleading conclusion
Allow trend analysis over time	Lack of control on how data was collected
Its always useful when primary data is not feasible	Some information may be outdated

6.8 Resources

1. https://www.unhcr.org/handbooks/assessment/sites/assessment/files/2023-11/How%20to%20Conduct%20a%20Secondary%20Data%20Review.pdf
2. https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/impact-evaluation/planning-your-impact-and-evaluation/choosing-evaluation-methods/using-secondary-data/

6.9 Assessment

1. **What is secondary data?**
   a. *Data collected firsthand by researchers for a specific study*
   b. *Data that is obtained through direct observation*
   c. *Data gathered through experiments*
   d. *Data collected from previously existing sources*
2. **What is the major advantage of using secondary data?**
   a. *It is more accurate than primary data*
   b. *It is less expensive and saves time*
   c. *It requires sampling*

d. It eliminates need for further data analysis

3. What is the main limitation of secondary data?

a. It is always more reliable than primary data
b. It may not be specific to the researcher's study objective
c. It's always up-to-date and relevant
d. The researcher always has control on how the data was collected

4. Which of the following source provides secondary data?

a. Focus group discussions
b. Field experiments
c. Government reports and national surveys
d. Key informant interviews

5. In what scenario would a researcher most likely use secondary data?

a. When one wants to use existing data to support decision making and further analysis of the data
b. When collecting real-time data
c. When conducting interviews with local community members
d. When looking for personal opinions

Section 7: Stakeholder Mapping

7.1 Introduction

This section introduces participants to stakeholder mapping, understanding different types of stakeholders and how to identify and categorize them according to their power interest and engagement strategies.

7.2 Training Objectives

By the end of this session, the participant should be able to:

a. To understand different types of stakeholders;
b. To be able to identify and classify stakeholders according to their influence/interest.

7.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

7.4 Duration: 10 mins

7.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

7.6 Content

a. Introduction to stakeholder mapping
b. Stakeholder identification and categorization
c. Influence vs interest matrix

7.7 Training Content

7.7.1 Introduction to Stakeholder Mapping

Stakeholders are individuals, groups or organizations that have interest in or affected by a project or initiative. They can influence or be influenced by the outcomes of an initiative. On the other hand, stakeholder mapping entails identifying, analyzing and prioritizing stakeholders who have interest in or influence over a particular project or initiative. Effective stakeholder mapping can ensure better engagement, collaboration and decision making.

There are different types of stakeholders including decision makers, beneficiaries, implementors, donors and information disseminators.

When mapping stakeholders, researchers should understand their purpose, how they can influence the project, their potential impact on the project, understand the needs of each stakeholder and how to address issues from their perspective.

7.7.2 Stakeholder Categorization

Once stakeholders have been identified, they can be categorized based on various attributes to determine their level of engagement and influence. The following methods can be used to categorize stakeholder:

a. **Power Dynamics:** This involves evaluating stakeholders based on their ability to influence decisions (power) and their like hood of them changing positions (dynamics).

b. **Internal vs External Stakeholders:** Internal stakeholders are individuals or groups within an organization such as managers, and employees. While external stakeholders are groups outside the organization that may be affected by or influence the project such as government agencies, NGOs, and Community groups.

c. **Primary, Secondary and Tertiary stakeholders:**

- **Primary stakeholders** are people who are directly affected by the project and they include beneficiaries, employees or customers.
- **Secondary stakeholders** are stakeholders who are indirectly affected by a project/initiative. They include government agencies, advocacy groups, or media.
- **Tertiary stakeholders** are individuals with significant influence such as policy makers, funders or industry leaders.

d. **Functional Categorization:** Stakeholders can also be grouped based on their roles in the project, i.e.:

- **Decision makers** e.g. government officials and executives
- **Implementers** e.g. Project managers, staff and field workers
- **Beneficiaries** e.g. community members

7.7.3 Influence vs Interest Matrix (Power interest grid)

This matrix classifies stakeholders into four groups based on their level of influence/power and interest of the project.

Influence vs Interest Matrix (Power interest grid)

Influence level	Description	Category/ examples	Engagement strategy
High Influence/ High Power.	These are stakeholders who can significantly impact the project.	Government agencies, donors, regulatory bodies.	Manage closely.
High Influence/ Low Interest.	They have high power but low interest. While they cannot be actively engaged, their influence requires careful management to avoid resistance.	Policy makers, influential leaders, and high-ranking officials.	Keep satisfied.
Low Influence / High Interest.	These are stakeholders who are interested in the project but lack significant power.	Local communities, advocacy groups, civil society organizations etc.	Keep informed.
Low Influence/ Low interest.	These are stakeholders with/ require minimal effort.	General public, media (if they are not actively reporting on the issue.)	Monitor closely in case they change their interest

7.8 Resources

1. https://www.projectmanager.com/blog/stakeholder-mapping-guide
2. https://cdn.who.int/media/docs/default-source/reproductive-health/contraception-family-planning/stakeholder-mapping-tool.pdf

7.9 Assessment

1. **What is the primary purpose of stakeholder mapping in project planning and implementation?**
   a. *To exclude non-relevant stakeholders*
   b. *To identify, analyze, and prioritize stakeholders for effect engagement*
   c. *To create a static list of all potential stakeholders*
   d. *To assign financial resources to all stakeholders*
2. **Which of the following is not a common method for categorizing stakeholders?**
   a. *Influence vs interest matrix*
   b. *Internal vs external stakeholders*
   c. *Power-dynamics matrix*
   d. *Revenue contribution matrix*
3. **When conducting a stakeholder analysis, why is it important to assess both interest and influence?**
   a. *To determine how stakeholders may oppose the project*

b. *To ensure all stakeholders receive equal engagement*
c. *To focus solely on stakeholders with high financial investment*
d. To assign legal responsibility to stakeholders

4. **Which stakeholder engagement strategy would be most appropriate for stakeholders with high influence and high interest?**
    a. Keep informed
    b. Consult occasionally
    c. Collaborate and involve in decision making
    d. Avoid engaging to prevent interference
5. **A stakeholder mapping exercise reveals that a community group has high but low influence in policy change. What is the best approach for engaging with this group?**
    a. Ignore them as they have low influence
    b. Keep them informed in decision making
    c. Give them decision making power
    d. Engage them after implementation

Section 8: Data Quality and Data Consistency

8.1 Introduction

In this session the participant will learn data quality and data consistency, key dimensions and how to ensure that data collected is of high quality and consistent.

8.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Define data quality and data consistency;
b. Explain why data quality and consistency is important when collecting data.

8.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

8.4 Duration: 10 mins

8.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

8.6 Contents

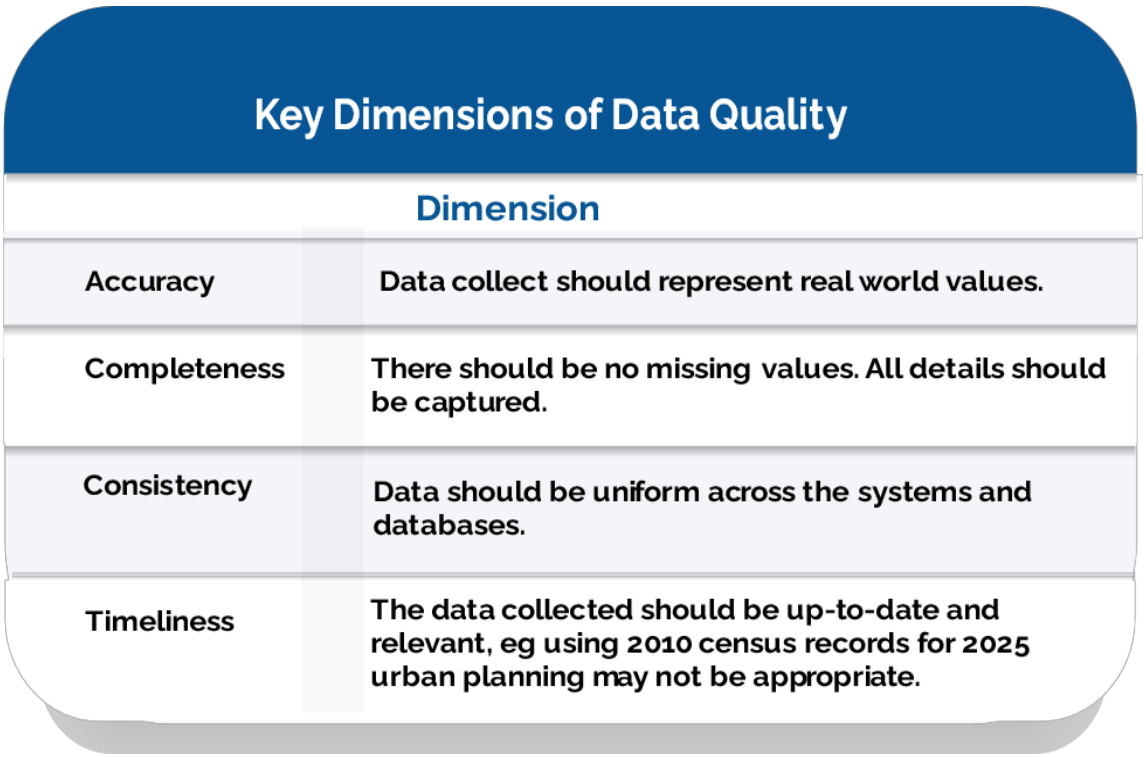
a. Data quality control and data consistency
b. Key dimensions of data quality
c. How to ensure data quality and consistency

8.7 Training Content

8.7.1 Data Quality and Data Consistency

Ensuring data quality is crucial for reliable research, decision making and analysis. Data **quality** refers to the accuracy, completeness, and reliability of data while **consistency** ensures that data remains uniform across different sources, systems and timeframes.

8.7.2 Key Dimensions of Data Quality



8.7.3 How to Ensure Data Quality and Consistency

Below are some of the factors to consider for ensuring data quality and consistency.

- **Use standardized data collection methods**: Develop and use structured data collection tools (i.e. questionnaire guides, checklists etc.) to maintain uniformity.
- **Use automated validation checks to detect errors**: Use software tools like Excel, SPSS or data management systems that flag errors automatically.
- **Ensure the data collected is properly documented:** Maintain detailed and clear field records with descriptions of data including dates, location, responsible individuals for easy traceability.
- **Logical consistency.** For example, if a participant's response to two similar questions contradicts each other, it could mean they didn't understand the questions properly and the data is inaccurate.
- **Check the data regularly for inconsistency and errors**: Conduct periodic audits and reviews of data to identify and correct inconsistencies.
- **Store and manage data in a unified system:** Use centralized data storage systems like databases (e.g. Google Drive, Microsoft SharePoint). And assign unique identifiers (e.g. respondent IDs, survey codes etc.) to avoid duplication.

- **Train and capacity build data collectors:** Conduct regular training sessions on data collection ethics, accuracy and best practices as well as training on use of digital tools and software to enhance efficiency.

8.8 Resources

1. Fan, W., GEERTS, F., & Jia, X. (2007). Improving data quality: Consistency and accuracy. ACM.
2. Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, *45*(4), 211-218.
3. Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM*, *40*(5), 103-110.

8.9 Assessment

1. **What is data quality?**
   a. *The amount of data collected in research*
   b. The number of sources from which data is collected
   c. The speed at which data is processed
   d. The degree of which data meets accuracy, reliability and relevance
2. **What is the main impact of poor data quality in decision making?**
   a. *It can lead to inaccurate conclusions and poor policy choices*
   b. It improves decision making by increasing data volume
   c. It reduces data costs in data management
   d. It has no significant impact in decision making
3. **Which practices can help in ensuring data accuracy?**
   a. *Collecting data from unverified sources*
   b. Regular data validation and cross checking
   c. Ignoring missing values and proceeding to analysis
   d. Using secondary data
4. **Which of the following is not a dimension of data quality?**
   a. Subjectivity
   b. Completeness
   c. Timeliness
   d. Consistency
5. **What is the best approach to maintaining high quality data?**
   a. Conducting regular data audits and validations
   b. Allowing data entry from multiple unverified sources
   c. Avoiding any modification of data once it has been collected
   d. Limiting data collection to small sample size

CHAPTER 3.

Sampling Techniques

Section 1: Overview of Sampling

1.1 Introduction

This session outlines the general understanding of sampling; purpose of sampling in research and the role of randomization in research.

1.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Define sampling and understand different sampling techniques;
b. State the importance of sampling;
c. Discuss the role of randomization in research.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 15 mins

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents

a. Overview of sampling
b. Importance of sampling
c. The purpose of sampling when researching and the role of randomization

1.7 Training Content

1.7.1 Overview of Sampling

Example – *Weighing Basu's Elephants* -

"National Park Manager would like to transport 50 adult elephants and so he needs a rough estimate of the total weight of the elephants. As weighing an elephant is a cumbersome process, the owner wants to estimate the total weight by weighing just one elephant. Which elephant would he choose?"

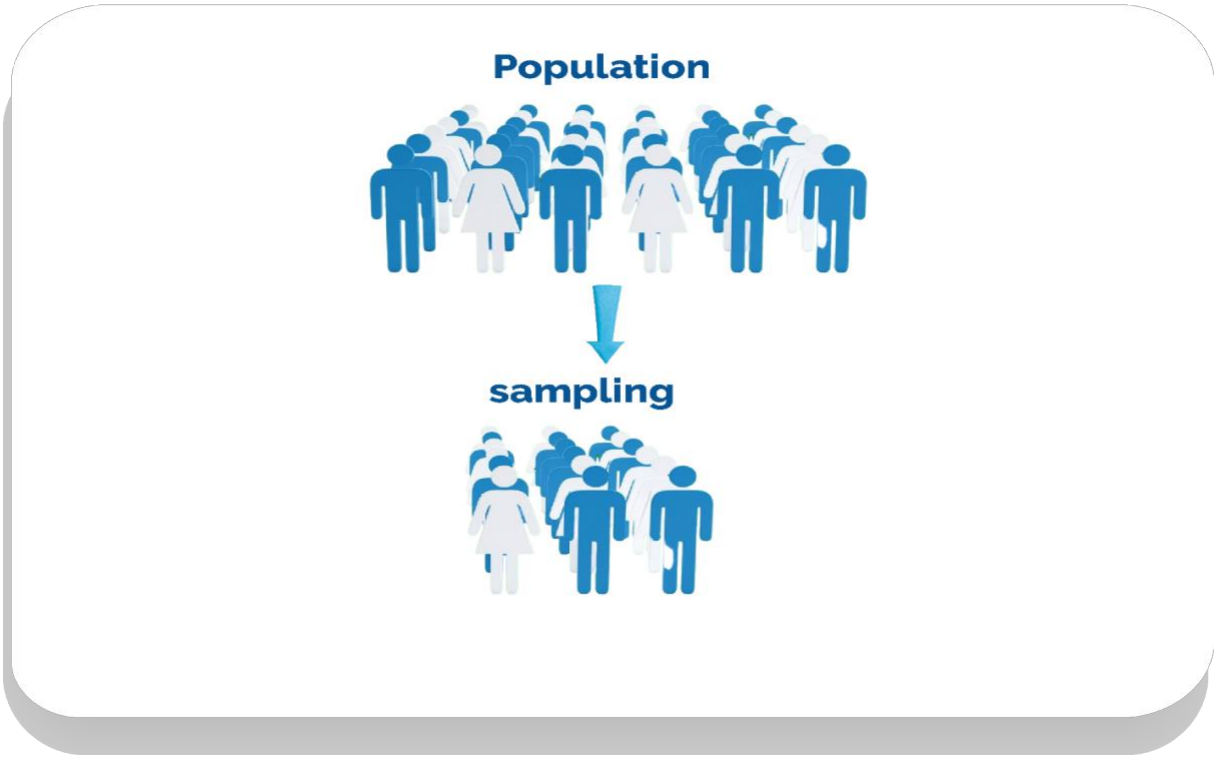
Sampling is about using all the resources efficiently. The park manager must work out how to estimate the total weight as economically as possible and achieve the best balance between accuracy and cost.

In research, **sampling** is the process of using a subset of a population to represent the whole population. For instance, assume a person wanted to do some research on everyone in Kenya. To ask every person would not be possible. Even if everyone is

allowed to participate in the survey, carrying out data collection and processing all the results, will take a long time and be very costly. Sampling is about using resources efficiently. There is need to work a balance between accuracy and cost as presented in the example of Basu's Elephant above.

Sampling allows large-scale research to be carried out with a more realistic cost and time-frame because it uses a smaller number of individuals in the population with representative characteristics to stand in for the whole.

However, when a researcher decides to sample, a new task is taken. And therefore, they must decide on the sample and how to choose the people who will best represent the whole population. How a researcher goes about this is what sampling is all about.



Sampling can sometimes be a complicated process, requiring a large amount of careful planning and expertise in statistics. However, at other times it may involve something as simple as selecting a few community groups to visit during a field trip. In either case, it is always important to thoroughly understand how results will be analyzed, and how the analyses will be used, before a sample can be developed.

Sampling can be used at four different stages;

a. Before a project or program starts, to contribute to design and planning;
b. At the start of a project or program, to form a baseline;
c. During a project or program, to establish what has changed and make modifications if necessary; and

d. At the end of a project or program, or sometime after completion, to establish what has changed.

1.7.2 Why Sampling is Important

Here are some of the benefits of sampling

a. **Reducing population to a manageable size/number:** Sampling ensures that even when populations are large or dispersed, research can be conducted feasibly.
b. **Reduces cost:** By studying a **smaller** group, researchers significantly reduce the expenses associated with data collection, making large-scale studies more manageable and affordable.
c. **Minimizing errors**: Properly selected samples yield accurate and representative results. By ensuring that the sample reflects the characteristics of the larger population, researchers can confidently make inferences and draw conclusions about the entire group.
d. **Time saving:** Gathering data from a subset of participants allows researchers to complete studies faster than attempting to study the entire population.

1.7.3 Role of Randomization in Research

Randomization is a statistical process in which a random mechanism is employed to select a sample from a population or assign subjects to different groups. *For example, if you want to compare the effects of two types of exercise on blood pressure, you can randomly assign half of your participants to do aerobic exercise and the other half to do resistance exercise.* Randomization is essential in research to improve the reliability of research findings and the quality of the research. This is mainly done to reduce selection bias. In simple terms, every potential participant has equal chance of being assigned to either of the groups. With randomization the results can be extrapolated from the general population where the subjects are enrolled from. Randomization also increases statistical validity by helping to determine whether observed differences in outcomes are due to the studied treatment or chance.

1.8 Resources

1. Ken Brewer – Combined Survey Sampling Inference, Oxford University Press Inc, New York
2. Bakewell, O; Adams, J and Pratt, B (2003). Sharpening the Development Process; A practical guide to monitoring and evaluation. INTRAC, UK.
3. Patton, M (1990). Qualitative evaluation and research methods. Beverly Hills, CA, Sage.
4. https://uca.edu/psychology/files/2013/08/Ch7-Sampling-Techniques.pdf
5. https://www.measureevaluation.org/resources/training/capacity-building-resources/data-quality-portuguese/ANOSTRAGEM1.pdf

1.9 Assessment

1. **Which of the following is not a stage where sampling can be used?**
   a. *At the start of a program*
   b. *Before the beginning of the project*
   c. *After data collection*
   d. *At the end of the project*
2. **What is the primary purpose of sampling in research?**
   a. *To select a representative subset of a population for the study*
   b. *To eliminate the need for data analysis*
   c. *To ensure only qualitative data is collected*
   To collect data from the entire population

3. **Which of the following is not a role of randomization in research?**
    a. *It reduces selection bias*
    b. *It ensures the entire population is studied*
    c. *It ensures that potential participants have a chance to participate*
    d. *It improves the reliability of research findings*
4. **Which of the following is a key characteristic of a good sample size?**
    a. *It should be large regardless of the population size*
    b. *It should only include participants who agree with the research objectives*
    c. *It should include all members of the target population*
    d. *It should be a representative of the population*
5. **Which of the following statements does not describe sampling?**
    a. *It's the process of using a subset of a population to represent the whole population*
    b. *It ensures resources are utilized efficiently*
    c. *It ensures that you use all the project resources to reach all the project participants*
    d. *It allows large scale research to be carried out within the time-frame*

Section 2: Target Population vs. Sample Size

2.1 Introduction

This session outlines the difference between population vs sample, and target population and sampling size.

2.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Define sample, population, sample size and target population;
b. Explain how target population and sample size are determined.

2.3 Definition of Terms

a. **Target population** means the entire group of units from which data could theoretically be collected. Always denoted by N
b. **Sample frame** is a list of all the known units in the population.
c. **Sample** is a list of specific units from which data will be collected. Denoted by n
d. **Sample size** means the total number of units from which data will be collected and analyzed.

2.4 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.5 Duration: 15 mins

2.6 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents

a. Target populations
b. Sampling Frame
c. Sample size

2.7 Training Content

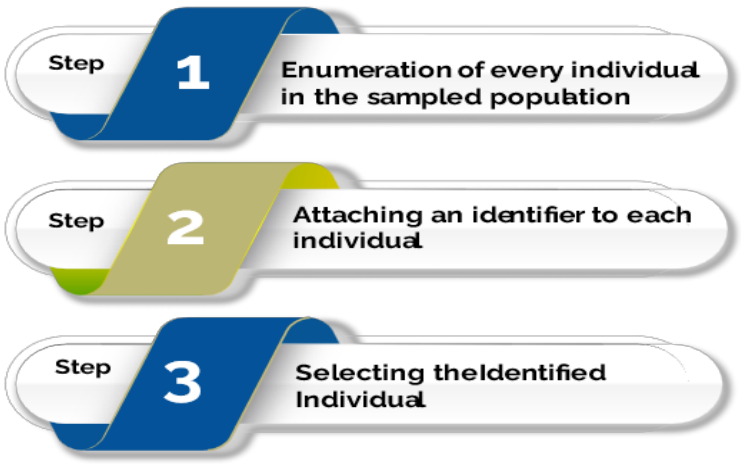
2.7.1 Target Population

Target population refers to the entire group of individuals or elements that a researcher wants to study and draw conclusions about in their research, while **sample size** is the specific number of individuals or elements selected from that target population to collect data from, representing a smaller subset used to make inferences about the larger group.

If the population is very large, demographically mixed, and geographically dispersed, it might be difficult to gain access to a representative sample. A lack of a representative sample may affect the validity of your results, and can lead to several research biases, particularly sampling bias.

2.7.2 Sampling Frame

The sampling frame is the actual list of individuals that the sample will be drawn from. Ideally, it should include the entire target population (and nobody who is not part of that population). For example, all farmers, informal sector workers etc. To construct a Sampling Frame requires three steps as highlighted below:



2.7.3 Sample sSize

The number of individuals you should include in your sample depends on various factors, including the size and variability of the population and your research design. There are different sample size calculators and formulas depending on what you want to achieve with statistical analysis.

Choosing the statistically significant sample size depends on a number of things such as the size of the population, how precise you want your estimates to be, how confident you want to be in the results, how different the population is likely to be and how much money and time you have for the study

2.7.4 Difference between target population and sample size

Difference between target population and sample size	
Target Population	**Sample Size**
It is the entire group of interest.	It is a selected portion of the population.
It is difficult to study as a whole.	It is easy & manageable to collect data from.
It is defined by research scope.	Provides findings that can be generalized.

2.8 Resources

1. https://www.ojp.gov/sites/g/files/xyckuh241/files/media/document/peer5.pdf
2. https://appspenang.uitm.edu.my/sigcs/2023-2/Articles/20234_UnderstandingPopulationAndSampleInResearch.pdf
3. https://www.ndi.org/sites/default/files/samplesizecalculation.pdf
4. https://library2.lincoln.ac.nz/documents/sample-size.pdf
5. https://ihopejournalofophthalmology.com/content/132/2022/1/1/pdf/IHOPEJO-1-009.pdf
6. https://www.slideshare.net/slideshow/sample-size-determination23112021pdf/264611481

2.9 Assessment

1. **What is a sampling frame in research?**
   a. *The list or database that contains all the elements of the population*
   b. *The process of selecting sample from the population*
   c. *It's the final report of a research analysis*
   d. *It's a statistical method used to analyze data*
2. **Which of the following is not a factor to be considered when determining a sample size?**
   a. *The size of the population*
   b. *How much money and time you have for the study*
   c. *How precise you want your estimates to be*

d. *How long you want the report to be*

3. **Which of the following can be used as a sampling frame?**
   a. *A hypothesis study*
   b. *A summary of research findings*
   c. *A list of registered voters for an election study*
   d. *A general description of the target population*
4. **What is a target population?**
   a. *A specific number of individuals a researcher wants to study*
   b. *The entire group of individuals that a researcher wants to study*
   c. *Both A and B above*
   d. *It's a subset used to make inferences about a population*
5. **What is the difference between a target population and sample size?**
   a. *The sample size is difficult to study as a whole while the target population is easy and manageable to collect data from*
   b. *The target population provides findings that can be generalized while the sample size is defined by the research scope*
   c. *The sample size is the entire group of interest while the target population is a selected portion of the population*
   d. *The target population is difficult to study as a whole while the sample size is easy to manage and collect data from*

Section 3: Sample Size Calculations and Errors

3.1 Introduction

This session highlights how sample size is determined and how to account errors.

3.2 Training Objectives

By the end of the session, the participant should be able to:

a. Be able to calculate sample size;
b. Explain variables in sample size determination;
c. Discuss error analysis.

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration: 20 min

3.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents:

a. Sampling size calculation
b. Sampling error

3.7 Training Content:

3.7.1 Sampling Size Calculation

A **sample size** calculation determines the number of participants needed in a study to achieve a desired level of accuracy, while **sampling error** refers to the difference between the results obtained from a sample and the true population value, which is directly impacted by the sample size; a larger sample generally leads to a smaller sampling error.

The sample size is determined based on the target population list using the formula (Cochran equation) below:

$$n_0 = \frac{Z^2 pq}{e^2}$$

Where:

Z = Critical value of the normal distribution at the required confidence level,
p = Sample proportion,
e = Margin of error

The z-score is the number of standard deviations a given proportion is away from the mean. To find the right z-score to use, refer to the table below:

The z-score is the number of standard deviations a given proportion is away from the mean. To find the right z-score to use, refer to the table below:	
Desired confidence level	**z-score**
80%	1.28
85%	1.44
90%	1.65
95%	1.96
99%	2.58

In social science research, 95% confidence level with a z-score of 1.96 is preferred.

With the online sample size calculator (https://www.calculator.net/sample-size-calculator.html), determining the appropriate sample size has become simple and accessible. Even individuals without a statistical background can easily calculate the required sample size, ensuring accuracy and efficiency in research planning.

3.7.2 Calculating Sampling Error

The accuracy of a sample can affect the results of a study if a researcher selects a sample that does not reflect the real composition of the population being studied. Calculating sampling error can help research professionals determine a sample's efficacy by measuring how close it is to the targeted community. A sampling error is a calculation that measures statistical error when a researcher uses a sample that does not properly reflect the population under consideration. Sampling error decreases as the sample size increases.

The sampling error is calculated by dividing the standard deviation of the population by the square root of the size of the sample and then multiplying the resultant with the Z-score value, which is based on the confidence interval.

Sampling error = $z.\frac{\sigma}{\sqrt{n}}$

Where:

Z is the Z score value based on the confidence interval (approx = 1.96)

σ is the population standard deviation
n is the size of the sample

3.7.3 Types of sampling errors

a. **Population-specific error**: A population-specific error can happen when a researcher does not understand who to survey. Avoid this mistake by understanding the research question before selecting your sample.

b. **Selection error:** This error occurs when respondents elect to participate in a study but then only those interested in the survey answer its questions. Overcoming this selection error by encouraging participation from the sample population.

c. **Sample frame error:** This refers to errors that happen when a researcher selects a sample from incorrect population data. Sample frame errors also occur when researchers accidentally include respondents from outside the population of interest.

d. **Nonresponse error:** Nonresponse errors arise when researchers cannot contact potential respondents or participants do not engage in the study.

3.8 Resources

1. https://mics.unicef.org/sites/mics/files/MICS3_Chapter_4__Designing_and_Selecting_the_Sample_060219.pdf
2. https://www.opalco.com/wp-content/uploads/2014/10/Reading-Sample-Size1.pdf
3. https://www.tarleton.edu/academicassessment/wp-content/uploads/sites/119/2022/05/Samplesize.pdf
4. https://www.ndi.org/sites/default/files/samplesizecalculation.pdf
5. https://m.ecios.org/Synapse/Data/PDFData/0157CIOS/cios-5-235.pdf

3.9 Assessment

1. **Which of the following is not true about a sampling error?**
   a. *It is obtained from the entire population*
   b. *It's the difference between the results obtained from a sample and a true population value*
   c. *It's a larger sample generally leading to a smaller sampling error*
   d. *It is directly impacted by the sample size*
2. **Which of the following is not a common type of sampling error?**
   a. *Selection error*
   b. *Population-specific error*
   c. *Sample frame error*
   d. *Statistics error*
3. **What is the consequence of not selecting an accurate sample?**

a. *The sample size will be reduced*
b. *The study results will be affected*
c. *The results will still be the same*
d. *It will affect the validity of the findings*

4. **Which of the following is true about the population specific error?**
   a. *When a researcher understands who to survey*
   b. *When a researcher does not understand who to survey*
   c. *When a researcher understands the research questions*
   d. *When a researcher selects only interested participants*
5. **When does nonresponse error occur?**
   a. *When participants engage in the study*
   b. *When researchers contact all the respondents*
   c. *When the researcher cannot reach all the respondents*
   d. *When the researcher reschedules the dates of interview*

Section 4: Determination Sample Size

4.1 Introduction

Sampling techniques are essential in research as they help in data gathering that is representative of the population. This session will provide an understanding of various sampling techniques and different approaches. The participant will further understand the importance of selecting the right sampling method based on the research objectives and population characteristics.

4.2 Training Objectives:

By the end of this session, the participant should be able:

a. Understand sampling techniques;
b. Be able to differentiate different categories of sampling;
c. Understand when to use different sampling techniques.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 30 minutes

4.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

4.6 Contents:

a. Discuss different approaches to sampling for qualitative vs. quantitative research
b. Cover simple random sampling, systematic sampling, stratified sampling and cluster sampling

4.7 Training Content:

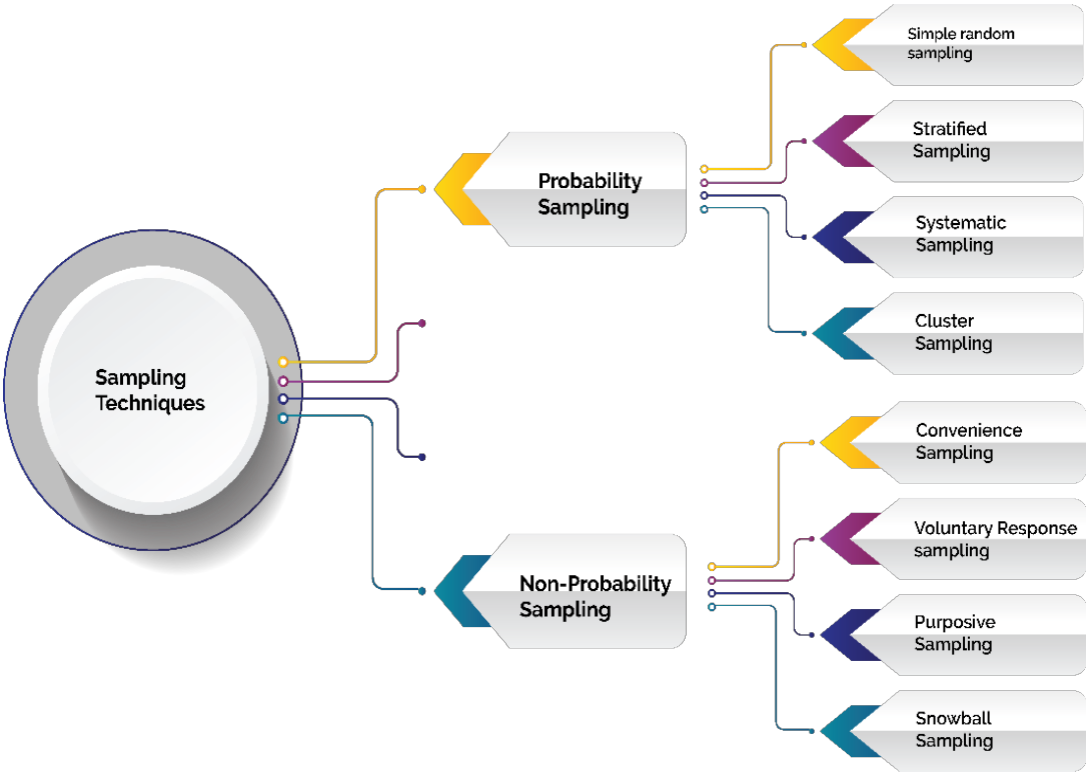
4.7.1 Overview of Sampling

Sampling is a critical process in research that allows researchers to draw conclusions about a larger population by examining a smaller, manageable subset. Sampling methods are essential for producing reliable, representative data without needing to survey an entire population. This topic covers various types of sampling methods, key techniques and practical examples that can help one select the most suitable method for their research.

The fundamental aim for sampling is to draw conclusions about the entire population without having to engage with every individual data point, hence saving time, resources and effort while still achieving accurate results.

5.7.2 Sampling Techniques

Sampling can be categorized into two (2) groups: probability and non-probability sampling.



I. Probability/ Random Sampling

In probability (Random) sampling, every individual or item in the population has a known, non-zero chance of being selected. This type of sampling is often used when researchers aim for unbiased, generalizable results.

It involves random selection, allowing for strong statistical inferences about the whole group. There are four main types of probability sampling.

a. Simple Random Sampling

In a simple random sample, every member of the population has an equal chance of being selected. The sampling frame should include the whole population. To conduct this type of sampling, you can use tools like random number generators or other techniques that are based entirely on chance. Random samples require naming or numbering the target population and then using some raffle method to choose those to make up the sample.

For example, if there are 500 students in a school and a researcher wants to select 50 students to participate in a survey about learning experience, he/she can assign unique numbers from 1 – 500 and select 50 numbers randomly.

b. Systematic Random Sampling

Systematic sampling is similar to simple random sampling, but it is usually slightly easier to conduct. Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals. A starting point is randomly selected, and then every k^{th} individual is chosen from a list. This method is often used when there's a fixed pattern or order in the population list.

For example, from our previous example of students to be surveyed in a school the researcher, if k is 10 then it will mean that every 10^{th} student on the list is selected for the survey.

c. Stratified Random Sampling

Stratified sampling involves dividing the population into subpopulations that may differ in important ways. It allows researchers to draw more precise conclusions by ensuring that every subgroup is properly represented in the sample. The population is divided into subgroups (strata) based on a characteristic (e.g., age, gender), and random samples are taken from each subgroup.

d. Cluster Random Sampling

Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample. Instead of sampling individuals from each subgroup, researchers randomly select entire subgroups. The population is divided into clusters (groups) that are randomly selected. All individuals within selected clusters are then included in the sample.

For example, is a government wants to evaluate the performance of students in a country, instead of surveying all students, they can divide the population into clusters based on geographical regions. i.e. school, districts, counties/provinces.

II. Non-probability Sampling

Non-probability sampling involves non-random selection based on convenience or other criteria, allowing for one to easily collect data. In non-probability sampling, individuals are selected based on specific characteristics or convenience rather than random selection. This method is suitable for exploratory research where generalizability is less critical.

This type of sample is easier and cheaper to access, but it has a higher risk of sampling bias. That means the inferences made about the population are weaker than with probability samples, and conclusions may be more limited. When using a non-

probability sample, the aim should be to make it as representative of the population as possible.

Non-probability sampling technique hues are often used in exploratory and qualitative research. In these types of research, the aim is not to test a hypothesis about a broad population, but to develop an initial understanding of a small or under-researched population.

a. Convenience Sampling

A convenience sample simply includes the individuals who happen to be most accessible to the researcher.

This is an easy and inexpensive way to gather initial data, but there is no way to tell if the sample is representative of the population, so it cannot produce generalizable results. Convenience samples are at risk for both sampling bias and selection bias. Participants are selected based on availability or ease of access, making it a fast and easy sampling method.

b. Voluntary Response Sampling

Similar to a convenience sample, a voluntary response sample is mainly based on ease of access. Instead of the researcher choosing participants and directly contacting them, people volunteer themselves (e.g. by responding to a public online survey).

Voluntary response samples are always somewhat biased as some people will inherently be more likely to volunteer than others, leading to self-selection bias.

c. Purposive Sampling

This type of sampling, also known as judgement sampling, involves the researcher using their expertise to select a sample that is most useful to the purposes of the research.

For example, if a research team is conducting a study on leadership on women's career, instead of randomly selecting participants, they will deliberately choose female executives to participate in the study.

Purposive sampling is often used in qualitative research, where the researcher wants to gain detailed knowledge about a specific phenomenon rather than make statistical inferences, or where the population is very small and specific. An effective purposive sample must have clear criteria and rationale for inclusion. Researchers must always make sure to describe their inclusion and exclusion criteria and beware of observer bias affecting their arguments.

Participants are selected based on specific criteria or characteristics relevant to the study's purpose.

d. Snowball Sampling

If the population is hard to access, snowball sampling can be used to recruit participants via other participants. The number of people you have access to "snowballs" as researchers get in contact with more people.

For example, if a researcher studying experiences of shea butters traders in Tamale Ghana, is having difficulties to identify participants, to overcome this the researcher can request the first participant to refer him/her to other traders in the region.

The downside here is also representativeness, as there is no way of knowing how representative the sample is due to the reliance on participants recruiting others. This can lead to sampling bias. Participants recruit other participants, making it useful for studying hard-to-reach populations.

e. Quota Sampling

Quota sampling relies on the non-random selection of a predetermined number or proportion of units. This is called a quota. The population is divided into categories (e.g., age, gender), and a specified number of participants from each category is chosen non-randomly.

5.7.3 Advantages and Disadvantages of Data Sampling Methods

Advantages and Disadvantages of Data Sampling Methods

Sampling Method	Advantages	Limitations
Random Sampling	Provides unbiased representation of the population. Allows for generalization to the population. Representative	Requires complete list of population members. May be impractical for large populations. Time consuming
Stratified Sampling	Ensures representation of subgroups within the population. Increases precision and reduces sampling error.	Requires knowledge of population characteristics. Complex to implement for heterogeneous populations. Time consuming
Systematic Sampling	Simple and easy to implement. Suitable for ordered populations.	Can introduce bias if there is a periodic pattern in data. May miss out on variability present in the population.
Cluster Sampling	Cost-effective for large and geographically dispersed populations. Reduces logistical challenges.	Requires accurate clustering information. May lead to increased sampling error compared to other methods. Less precision
Convenience Sampling	Quick and easy to implement. Cost-effective for small-scale studies.	Prone to selection bias and lack of representativeness. May not generalize to the broader population.
Purposive Sampling	Allows for targeted selection of specific groups. Useful for studying rare or hard-to-reach populations.	Limited generalizability. Subject to researcher bias.
Snowball Sampling	Facilitates access to hidden or hard-to-reach populations. Cost-effective for studying social networks or sensitive topics.	Relies on referrals and may lead to biased samples. Sample may lack diversity or representativeness.
Quota Sampling	Ensures proportional representation of key population segments. Simplifies sampling process compared to probability sampling.	Requires careful selection of quotas and sampling method. May lead to biased samples if quotas are not well-designed.

4.8 Resources

1. Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications.
2. Babbie, E. (2020). The Practice of Social Research. Cengage Learning.
3. Fowler, F. J. (2014). Survey Research Methods. Sage Publications.
4. Lohr, S. (2021). Sampling: Design and Analysis. Chapman and Hall/CRC.
5. Patton, M. Q. (2015). Qualitative Research & Evaluation Methods. Sage Publication

4.9 Assessment

1. **Which of the following sentences best describe probability sampling?**
   a. It involves nonrandom selection of participants based on convenience
   b. It has a higher risk of sampling bias
   c. They are mostly used in qualitative research
   d. Its where every individual in the population has a chance to be selected
2. **What is systematic sampling?**
   a. *Its where the population is divided into subgroups with similar characteristics*
   b. *Its where people volunteer to be interviewed*
   c. *Its where individuals are listed with a number and chosen at regular intervals*
   d. *Its where the researcher uses their expertise to select the sample that is most useful to the study*
3. **Which of the following is not a non-probability sampling technique?**
   a. *Snowball sampling*
   b. *Cluster sampling*
   c. *Purposive sampling*
   d. *Voluntary response sampling*
4. **What is the limitation of Snowballing?**
   a. *Sampling may lack diversity or representative*
   b. *It ensures representation of subgroups within a population*
   c. *It requires complete list of population members*
   d. *It requires knowledge of the population*
5. **Which of the following is not an advantage of purposive sampling?**
   a. *It allows for selection of specific target group*
   b. *It's useful for studying a hard-to-reach population*
   c. *It's useful for gathering qualitative data*
   d. *It relies on referrals*

CHAPTER 4.

Research Ethics

Section 1: Overview of Research Ethics

1.1 Introduction

This section outlines the overview of research ethics; the participant will also explore the principles that govern ethical research practices.

1.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Define research ethics;
b. Understand the principles of research ethics.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 10 mins

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents

a. What is research ethics and why it is important

1.7 Training Content

1.7.1 Research Ethics

Research ethics refers to the set of principles, standards and practices that guide ethical behavior in conducting research. Research ethics are essential because they guarantee that scientific studies are carried out with accountability, integrity and respect. Unethical research techniques can have serious repercussions, including injury to participants, a decline in public confidence and harm to the image of the research community.

1.7.2 Ethical Principles in Research

Research ethics are based on the following fundamental principles:

i. **Informed consent** refers to a situation in which all participants receive and understand all the information they need to decide whether they want to participate. This includes information about the study's benefits, risks, funding and institutional approval.

ii. **Voluntary participation** means that all participants/respondents are free to choose to participate without any pressure or coercion from the researcher. All participants are able to withdraw from the study at any point without feeling an obligation to continue, and participants do not need to provide a reason for leaving the study.

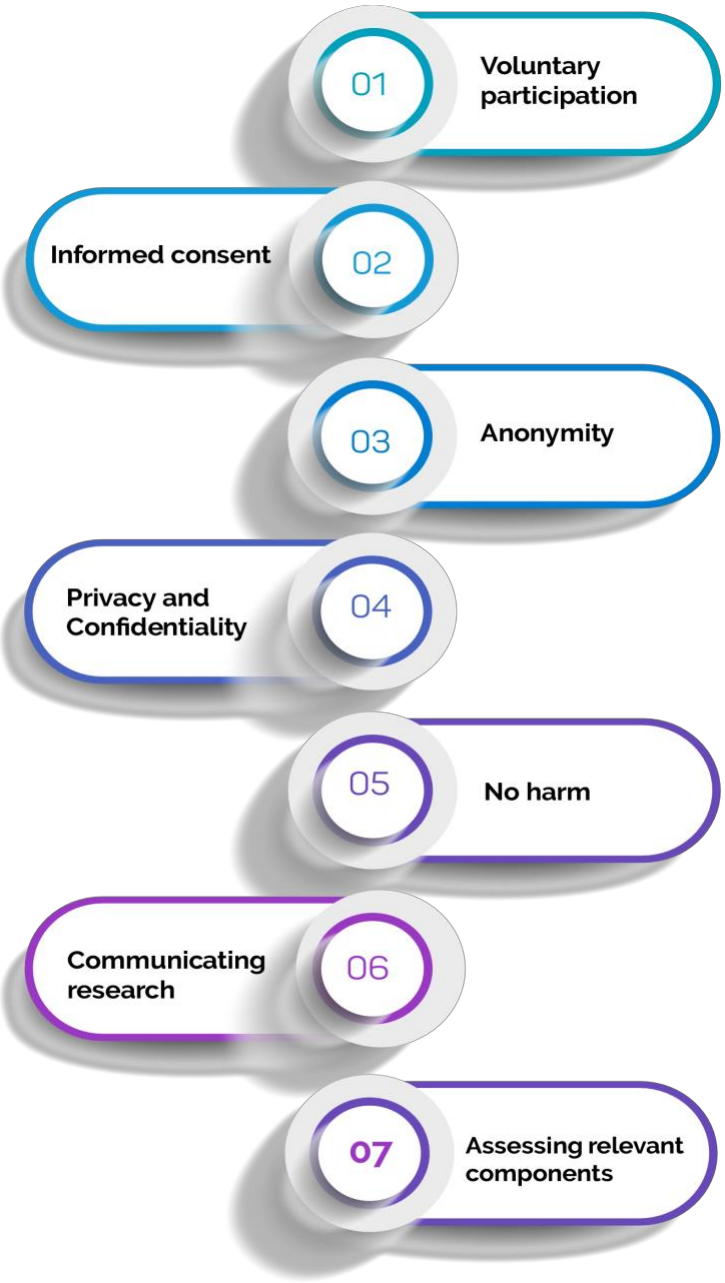
iii. **No harm** means the researcher has to consider all possible sources of harm to participants. Harm can come in forms psychological, social, physical and legal harm.

iv. **Privacy and Confidentiality** means that the researcher knows who the participants are, but removes all identifying information from the report. All participants have a right to privacy, so researchers should protect their personal data for as long as it is stored or used. Even when the researcher cannot collect data anonymously, the information should be secured.

v. **Anonymity** means that the researcher does not know who the participants are and can't link any individual participant to their data.

vi. **Communicating research** results ethically is present findings honestly without manipulation or bias, clearly disclose limitations and potential conflicts of interest, ensure accessibility by sharing results through appropriate channels and protect participant confidentiality when reporting sensitive data. This is essential to maintain transparency, credibility and public trust.

vii. **Only assessing relevant components** in research ensures ethical integrity and methodological soundness.



1.8 Resources

1. https://online225.psych.wisc.edu/wp-content/uploads/225-Master/225-UnitPages/Unit-10/Resnik_NIH_2015.pdf
2. https://edwebcontent.ed.ac.uk/sites/default/files/atoms/files/research_ethics_and_integrity_awareness.pdf
3. https://www.niehs.nih.gov/research/resources/bioethics/whatis

1.9 Assessment

1. **Which of the following sentences is not true about research ethics?**
   a. *Research ethics entails taking photos of the participants without their consent and publishing them on social media.*
   b. *Research ethics is a set of principles, standards and practices that guide ethical behavior in conducting research*
   c. *Unethical research ethics have serious repercussions*
   d. *Research ethics ensures that the studies are carried out with integrity and respect*
2. **Which of the following is not a principle of research ethics?**
   a. *Voluntary participation and anonymity*
   b. *Communicating research findings and doing no harm*
   c. *Privacy and confidentiality*
   d. *Coarse/uninformed consent and public opinion*
3. **Why is informed consent important when conducting research?**
   a. *It informs the participants on how knowledgeable they are*
   b. *It reduces bias when collecting data*
   c. *It ensures that the participants understand about the study, and allows them to decide whether they want to participant or not*
   d. *To ensures the participants understand that they will be paid after the study*
4. **Why should information provided by the respondent be treated with confidentiality?**
   a. *It ensures that the participants provide information forcefully*
   b. *It ensures that participants are not victimized*
   c. *It reduces errors in data collection*
   d. *It enhances the result findings*
5. **Which of the following is false about communicating research findings?**
   a. *It ensures that that there is transparency of the findings*
   b. *It clearly discloses limitations of the study*
   c. *It presents findings without manipulation*
   d. *It ensures that all participants have a right to privacy*

Section 2: Building Interviewees Trust

2.1 Introduction

This section highlights the barriers to participant engagement, how to mitigate them, and the importance of transparency to interviewees. It also entails the process of acquiring and documenting participants' consent. It elaborates further on incentives to participants and their benefits and drawbacks.

2.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Explain the barriers to participant engagement and how to mitigate them;
b. State the importance of transparency to interviewees;
c. Discuss the rights of participants, and how to protect those rights;
d. Explain how to acquire and document participants consent;
e. Explain incentives and their benefits and drawbacks.

2.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.4 Duration: 30 min

2.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents

a. The barriers to participant engagement, and how to understand them, their context, and build trust with the research context
b. How to mitigate barriers to data collection
c. The importance of transparency to interviewees while in the field
d. Rights of participants, and how to protect those rights while gathering data
e. How to acquire and document participants consent
f. Incentives, benefits, and drawback in research studies
g. Reasonable incentives

2.7 Training Content

2.71 Building Participants' Trust and Transparency

I. **The barriers to participant engagement, and how to understand them, their context, and build trust with the research context.**

Barriers to participant engagement in research are challenges that hinder respondents from participating fully or at all in studies. Understanding these barriers is crucial for researchers to design inclusive studies, build trust with diverse communities, and ensure ethical practices that promote meaningful

participation, while addressing participants' needs and concerns. Below are some of the barriers to participant participation:

a. **Lack of time and commitment:** Participants may struggle to allocate time for research activities due to personal or professional responsibilities.
b. **Concerns about the research process:** Fear of potential risks, privacy concerns or skepticism about research intentions can deter participation.
c. **Lack of awareness about the study:** Some potential respondents may not be informed about the research, leading to low participation rates.
d. **Inconvenience of study location:** If the research requires in-person visits at difficult-to-reach locations, participation may be limited.
e. **Strict eligibility criteria:** Narrow participant selection criteria may exclude willing respondents.
f. **Mistrust of researchers:** Previous unethical research practices or cultural distrust can discourage participation.
g. **Transportation issues:** Difficulty accessing research sites may prevent potential participants from engaging in the study.
h. **Financial constraints:** Participation may require travel, time off work, or other expenses that pose financial burdens.
i. **Lack of incentives:** Without compensation or meaningful incentives, individuals may lack motivation to participate.
j. **Difficulty understanding research materials:** Complex language or jargon in research documents may make it hard for participants to comprehend their role and rights.

II. How to mitigate barriers to data collection

To improve research participation and inclusivity, researchers can implement the following strategies:

a. **Incorporating researchers from marginalized groups:** Including researchers who face systematic barriers due to gender, geographical local, socio-economic status, disability etc. This can foster trust and comfort, especially in studies involving women and sensitive topics.
b. **Interviewing in safe locations:** Conducting interviews in secure, neutral, and accessible environments ensures participant safety and encourages participation.
c. **Raising awareness about the study:** Using multiple outreach methods (e.g. social media, community leaders, flyers, etc.) to inform potential participants about the research.
d. **Offering flexible participation schedules:** Providing evening, weekend or virtual options can accommodate participants' time constraints.
e. **Reducing financial burdens:** Compensating for transportation costs, lost wages or providing childcare support can remove financial barriers.
f. **Building trust through transparency:** Clearly explaining the research purpose, methodology and expected outcomes to address skepticism and concerns.

g. **Simplifying research materials:** Using simple terms, visual aids and multilingual translations to improve comprehension.
h. **Relaxing strict eligibility criteria:** Broadening inclusion criteria while maintaining research integrity can help increase participant numbers.
i. **Providing reasonable incentives:** Offering incentives to interviews to encourage participation while maintaining ethical standards.

III. The Importance of Transparency to Interviewees While in The Field

Transparency is important in fostering trust and encouraging open participation in research. Researchers should:

- **Clearly explain the purpose of the study:** Ensuring interviewees/respondents understand the objectives and how their responses will be used.
- **Disclose potential risks and benefits:** Informing participants about any risks involved and the possible benefits of the research.
- **Clarify confidentiality measures:** Assuring participants that their information will be protected and anonymized when necessary.
- **Encourage voluntary participation:** Making it clear that respondents have the right to decline participation or withdraw at any time without consequences.
- **Provide contact information:** Offering a way for participants to follow up with questions or concerns about the research.
- **Use simple and accessible language:** Avoiding jargon and using clear, culturally appropriate language to ensure full comprehension.

2.7.2 Getting Informed Consent from Participants

I. Rights of Participants, and How to Protect Those Rights While Gathering Data

Participants in research have rights to:

1. Voluntary participation: no participant should be coerced into taking part in a study.
2. Withdraw from the study at any time without consequences.
3. Privacy and confidentiality: Personal data must be protected and anonymized.
4. Informed Consent: participants must be fully informed about the study before agreeing to participate.
5. Protection from harm: researchers must minimize risks and ensure participants' safety.

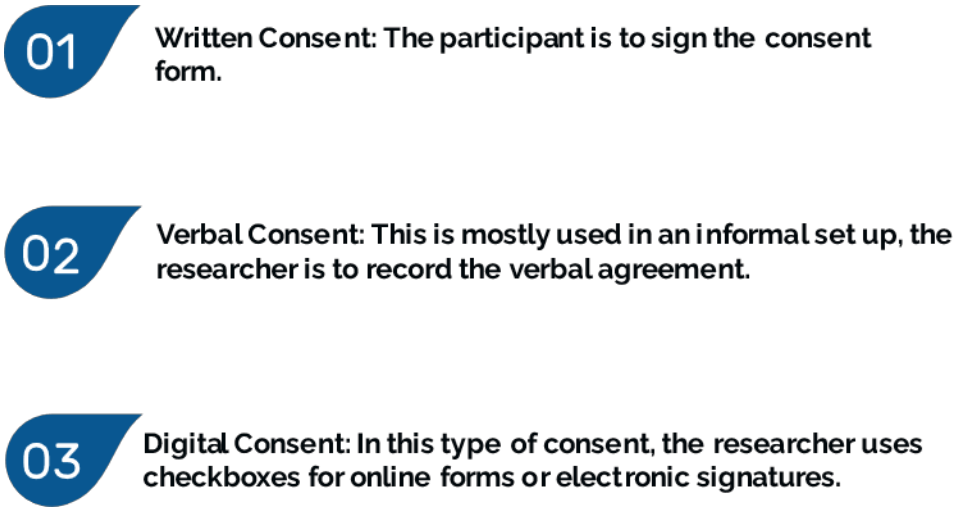
II. How to Acquire and Document Participants Consent

To acquire participant consent, the researcher should provide a clear and detailed explanation of their research study, including its purpose, methods, potential

risks and benefits and the participant's right to withdraw at any time. After, researchers need to have participants/respondents sign a written consent form that documents their agreement to participate to ensure they fully understand what they are consenting to and are not being coerced into participation.

The consent should explain how the participant's data will be protected. The consent should be written in a simple language to ensure clarity and allow participants to ask questions.

The consent can be obtained in three different ways:



When seeking consent to involve children in research, it must be demonstrated that comparable research cannot be done with adults. Once it has been determined that the research should be permissible, researchers must obtain parental/guardian consent on an Informed Consent Form for all children to be involved in the research.

2.7.3 Understand Incentives and Their Limits

I. What are Incentives, what are the Benefits and Drawback, and How They are Used in Research Studies

In research studies, incentives are rewards or motivations given to research participants or respondents to attract them, enhance engagement and ensure the completion of tasks required for the study. Incentives can be monetary (cash, gift cards), non-monetary (certificates, gifts, access to resources), or intrinsic (recognition, learning opportunities).

While incentives play a crucial role in research, they can also have negative impacts that may influence the integrity of the study. Some of these disadvantages include:

1. **Coercion or Undue Influence**: Large incentives may pressure respondents to participate, even if they are uncomfortable or unwilling, which can compromise voluntary participation and ethical standards.
2. **Reduced Motivation for Genuine Participation**: Respondents may focus on obtaining the incentive rather than genuinely engaging with the study, potentially leading to dishonest responses.
3. **Perception of Exploitation:** In cases where incentives are seen as too small relative to the effort or risks involved, participants may feel exploited or undervalued.
4. **Distortion of Responses:** Respondents might tailor their answers to align with what they perceive as desirable to researchers, hoping this will increase the likelihood of receiving the incentive or avoiding disqualification.
5. **Dependency on Incentives:** Regular use of incentives can create an expectation, making it difficult to recruit participants for future studies without similar or larger rewards.
6. **Ethical and Cultural Concerns:** In some cultures, offering incentives might be seen as inappropriate or coercive, potentially reducing trust in the research process.

II. Reasonable Incentives

Reasonable incentives are rewards or compensations offered to research participants that are appropriate in value and type, ensuring they motivate participation without exerting undue influence or compromising ethical and regulatory standards. The determination of what constitutes a "reasonable incentive" depends on several factors, including the research context, participant demographics, and the nature of the study.

Factors Determining Reasonable Incentives

1. **Financial Limitations:** The available budget for the study influences the value and type of incentives offered. Researchers must balance attracting participants with ensuring cost-effectiveness for the study.
2. **Ethical Considerations:** Incentives provided to the participants by the researcher must follow all the ethical principles and guidelines.
3. **Nature and Risk of the Study:** Studies involving significant time, effort, or discomfort, higher compensation may be justified to acknowledge participants' contributions and burdens. Low-risk studies, such as surveys, simple incentives should be provided to match the limited demands on participants.
4. **Transparency and Consistency:** Incentives should be clearly communicated during recruitment and remain consistent for all participants to avoid perceptions of favoritism or bias.

2.8 Resources

1. https://www.who.int/docs/default-source/ethics/process-seeking-if-printing.pdf?sfvrsn=3fac5edb_4

2.9 Assessment

1. **Which of the following is not a barrier to participants participation during a study?**
    a. *Lack of incentives*
    b. *Lack of time and commitment due to personal responsibilities*
    c. *Inconvenience of the study location*
    d. *Understanding research questions*
2. **Which of the following is not true about the rights of participants when conducting research?**
    a. *Protecting the participants from harm*
    b. *Protecting participants personal data*
    c. *Coercing participants into taking a study*
    d. *Participants being allowed to withdraw from a study without consequences*
3. **What are the disadvantages of giving incentives to the study participants?**
    a. *Large pressure participants to participate even when they are unwilling or not comfortable*
    b. *Respondents might tailor their answer to align with what the research wants*
    c. *Regular use of incentives can create an expectation making it difficult for future studies*
    d. *All of the above*
4. **Which of the following is not a factor when determining a reasonable incentive?**
    a. *Consider friends and family members first*
    b. *Consider the budget to balance attracting participants with ensuring cost effectiveness of the study*
    c. *They should be clearly communicated to avoid favoritism*
    d. *You should consider the nature and risk of the study*
5. **Which of the following is not a way of obtaining informed consent?**
    a. *Using written consent*
    b. *Using a police abstract*
    c. *Use of digital consent*
    d. *Use of verbal consent*

Section 3: Confidentiality and Privacy Policies, including Data Protection Policies

3.1 Introduction

This session explains how researchers can protect participant's information, privacy policies and to utilize them in research. The section further entails regulatory review boards such as Institutional Review Board (IRB) in USA or Ethics Committees (ECs) or Research Ethics Committees (RECs) in Europe and other parts of the world and other agencies that protect "human subjects in accordance with federal, institutional and ethical guidelines.

There is no single global IRB, but internationally, research ethics oversight is guided by various ethical frameworks and regional/national bodies as elaborated in this section. Data protection laws are now operationalized in many countries with an aim of safeguarding individuals' personal data, ensuring its lawful, fair, and transparent processing, while also enabling individuals to exercise control over their information.

3.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Explain how protect participants' information;
b. Discuss privacy policies and elaborate how to apply them to research;
c. Discuss regulatory review boards such as Institutional Review Board (IRB) and other agencies (i.e. ECs and RECs) that protect "human subjects...in accordance with federal, institutional, and ethical guidelines."

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration: 20 minutes

3.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents

a. Understand Confidentiality and Privacy Policies, including Data Protection Policies
b. Understand Legal and Regulatory Compliance
c. How to protect participants' information, maintain anonymity of participants and protect data
d. What are privacy policies and how to apply them to research
e. Background on general legal and regulatory compliance considerations
f. Regulatory review boards such as Institutional Review Board (IRB) and other agencies that protect "human subjects in accordance with federal, institutional, and ethical guidelines."

3.7 Training Content

3.7.1 Confidentiality and Privacy Policies, Including Data Protection Policies

How to Protect Participants' Information, Maintain Anonymity of Participants and Protect Data

Anonymity means that researchers do not know who the participants are and researchers cannot link any individual participant to their data.

Privacy and Confidentiality means that researchers know who the participants are, but remove all identifying information from the report. All participants have a right to privacy, so researchers should protect their personal data for as long as they store or use it. Even when researchers cannot collect data anonymously, they should secure confidentiality whenever they can.

A researcher can protect participant's privacy by doing the following:

i. Obtaining an **informed consent** clearly explaining the purpose, objectives, methodologies, potential benefits and risks of the research study.

ii. **Anonymity and confidentiality**- collecting data in a way that the participant will not be identified by the researcher (anonymous surveys). In cases where anonymity is not possible, ensure only the researcher can link data to an individual participant to maintain confidentiality.

iii. **Data collection and storag**e- the researcher should only collect necessary data that is relevant for the research, store participants information on password secured systems and limit access to the system.

iv. **Data anonymization**: when sharing data, the researcher should remove all identifiable information and instead use codes to identify participants.

v. **Data sharing and dissemination**: the researcher should establish clear agreements when sharing data with stakeholders (outlining data protection measures). The study findings should be presented in a way that protects participant privacy, avoiding identifiable details in publications.

Privacy Policies

Read more about data privacy via the below link

https://privacy.upd.edu.ph/revised-privacy-policy-for-researchers-and-research-subjects/

3.7.2 Legal and Regulatory Compliance [IRB and equivalency]

Regulatory bodies have been constituted to uphold the safety of participants involved in research. It is imperative to obtain approval from the appropriate regulatory authorities before proceeding to any research. The constitution and the types of these bodies vary between nations. The researchers are expected to be aware of these authorities and some of the international bodies are:

1. **The TRUST Code** is a resource for all research stakeholders who want to ensure that international research is equitable and carried out without 'ethics dumping' (conducting research without adhering to ethical considerations).
2. **Code of Conduct and Ethical Guidelines for Social Science Research** which ensures that research is conducted with integrity, respect and responsibility.

Examples of research compliance review boards include:

a. **Institutional Review Board (IRB):** Evaluates proposed research projects' ethical, methodological and regulatory aspects, ensuring that participants' rights, safety and well-being are upheld. It also assesses the risks and benefits of the research and the informed consent process.

b. **National Ethics Committees (NECs):** Many countries have national-level ethics authorities that set guidelines and accredit local ethics committees. For instance, In Kenya, National Scientific and Ethics Committee (NSEC) was created by NACOSTI to offer guidelines and create accreditation standards for Institutional Ethics Review Committees (ISERCs) to adhere to for their accreditation process to be successful. Researchers are encouraged to check with their countries on their local NECs.

c. **Data Safety Monitoring Board (DSMB):** Independently monitors and assesses the safety, scientific validity and ethical conduct of a research.

d. **Institutional Biosafety Committee (IBC):** Focuses on biosafety containment and worker safety. It assesses and manages potential risks associated with research endeavors, thereby safeguarding the well-being of researchers, the public and the environment.

3.8 Resources

1. https://www.hhs.gov/sites/default/files/ohrp-international-compilation-intl-orgs-2024.pdf
2. Home - The TRUST Code
3. https://unesdoc.unesco.org/codeofconduct/ethicalguidelines/socialscienceresearch
4. Different Types of Research Compliance Review Boards | Blog
5. https://sharing.nih.gov/data-management-and-sharing-policy/protecting-participant-privacy-when-sharing-scientific-data/principles-and-best-practices-for-protecting-participant-privacy

3.9 Assessment

1. **A researcher can protect the participants by doing the following except?**
   a. *Obtaining informed consent*
   b. *Ensuring anonymity when sharing data*
   c. *Establishing clear agreements when sharing data with stakeholders*

d. *Quoting the names of respondents in the report*

2. **Which of the following is not an ethical guideline for social science research?**
   a. *Abuse*
   b. *Responsibility*
   c. *Integrity*
   d. *Respect*
3. **What is the role of legal and regulatory compliance bodies?**
   a. *They ensure the software used for analysis is friendly*
   b. *They help to develop standard tools for data collection*
   c. *They provide guidelines and safety of the participants in research*
   d. *They provide the right sample size for the study*
4. **Why is data collection and storage an important component in ensuring that the participants privacy is protected?**
   a. *It ensures that only relevant data is collected and participants information is stored in secured systems*
   b. *It helps to remove duplicates*
   c. *It makes data manageable*
   d. *It ensures that analysis is done in a timely manner and the information is communicated to the grassroots stakeholders.*
5. **What is ethics dumping?**
   a. *It is understanding and applying research ethics during data collection*
   b. *It is removing duplicates from the data*
   c. *It is conducting research without adhering to ethical considerations*
   d. *Its collecting data from secondary sources*

Section 4: Risk Minimization and Maintaining Scientific Integrity

4.1 Introduction

Under this section, the participant will understand the purpose of removing personal or organizational bias, and how to protect data integrity in research. This section will also explain on how to create a document with the needed compliance and regulations to acquire to conduct research, and the organizations' plan to uphold informed consent, privacy and confidentiality, scientific integrity and manage incentives systems within research.

4.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Explain the purposes of removing personal or organizational bias;
b. Elaborate how to protect data integrity.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 15 mins

4.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

4.6 Contents

a. Understand Risk Minimization and Maintaining Scientific Integrity
b. Determine Research Compliance

4.7 Training Content:

4.7.1 Understand Risk Minimization and Maintaining Scientific Integrity

I. **Purposes of Removing Personal or Organizational Bias, and Protection of Data Integrity**

Researcher Bias

Researcher bias refers to the subjective beliefs, values, attitudes or preferences of a researcher that can influence the design, conduct or interpretation of a study's results. It occurs when the personal biases of the researcher inadvertently or intentionally influence the research process, potentially leading to distorted or inaccurate results. To minimize researcher bias, researchers employ rigorous research methodologies to maintain transparency.

The presence of bias in research affects the validity and reliability of research findings, which may lead to false conclusions and misinterpretation. All research depends on unbiased identification of personal and organizational preferences in order to achieve fairness and objective research. The main goals of bias reduction in research consist of the following elements:

- **Enhancing Validity**: Ensuring research findings accurately represent reality and are not influenced by subjective opinions or institutional agendas.
- **Promoting Fairness:** Ensuring diverse perspectives are considered, particularly when studying marginalized or underrepresented populations.
- **Improving Reproducibility:** Allowing other researchers to replicate and verify results without inconsistencies caused by bias.
- **Building Public Trust:** Demonstrating ethical responsibility by producing balanced and objective research outcomes.
- **Reducing Conflicts of Interest:** Avoiding potential manipulation of data due to financial, political or institutional influences.

Strategies to minimize bias include:

- Using diverse research teams with varied or different perspectives.
- Employing double-blind study designs where neither participants nor researchers know group assignments.
- Pre-registering research studies to avoid post-hoc alterations of hypotheses.
- Utilizing standard and uniform approaches in data collection and analysis.
- Conducting peer reviews and independent audits to assess objectivity.

4.7.2 Determining Research Compliance

Research compliance is the process of ensuring that research is conducted legally and ethically. It involves adhering to laws, regulations and institutional policies. It plays a crucial role in research that is:

- Ensuring objectivity and transparency
- Ensures the research is conducted in an ethical and responsible way
- Helps in the protection of the rights of research participants
- It ensures that research complies with privacy and data protections

In Kenya, the National Commission for Science, Technology, and Innovation (NACOSTI) requires all researchers conducting studies within the country to register and obtain a permit before undertaking any research activities. Researchers are also required to pay the necessary fees as part of the approval process. Upon completion of the study, the findings must be shared with NACOSTI to ensure compliance with national research regulations and contribute to the country's knowledge base.

4.8 Resources:

1. How To Avoid Bias In Research: Navigating Scientific Objectivity - Mind the Graph Blog
2. https://www.scribbr.com/category/research-bias/#:~:text=Why%20is%20bias%20in%20research,of%20treatment%20may%20be%20evaluated.

4.9 Assessment:

1. **How does research bias affect the validity of the research?**
   a. *It may lead to providing false conclusion and misinterpretation of the study findings*
   b. *It leads to gathering quality data*
   c. *It ensures that the data collected is valid*
   d. *It may lead providing viable conclusions and recommendations*
2. **Which of the following statements is false about research bias?**
   a. *It can lead to distorted or inaccurate results*
   b. *It can be minimized by employing rigorous research methodologies to maintain transparency*
   c. *It occurs when the researcher intentionally influences the research process*
   d. *It does not affect the research findings*
3. **How can bias be minimized in a study?**
   a. *By utilizing standard and uniform approaches in data collection and analysis*
   b. *By using data collection tools from other organizations which have been tested*
   c. *By using the appropriate software for data analysis*
   d. *By employing qualified data collectors*
4. **Which of the following is not true about research compliance?**
   a. *It ensures that research is conducted legally and ethically*
   b. *It involves adhering to laws, regulations and policies*
   c. *It ensures that all the key stakeholders are involved in the study*
   d. *It ensures that the research is transparent*
5. **What is the primary role of research compliance?**
   a. *It ensures that all participants are researched*
   b. *It eliminates bias during a study*
   c. *It ensures that study findings are well communicated to the government*
   d. *It ensures that research complies with privacy and data protection*

CHAPTER 5.

Problem Statement and Research Questions

Section 1: Organizational Context

1.1 Introduction

In this topic, the participant will understand how to understand background information of a project and the region to be studied. This includes understanding socio-political structures, economic factors, gender differences, and urban-rural dynamics, all of which influence the lived experiences of the population and workers.

1.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand local organizational context in research;
b. Recognize the role of worker movements in shaping regional labor conditions;
c. Examine gender lens and identify challenges related to gender disparities;
d. Understand factors affecting research process.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 20 min

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Training Content

1.6.1 Understanding Local Organizational context

Understanding local organizational and union factors is essential for effective grassroot research. It ensures that the study is contextual, inclusive and actionable. When collecting data within an organization or union, it is important to understand their structure, operations and socio-political context.

Some of the key areas you need to assess in an organization or union include:

a. **Organization structure and governance**: Assess the organizational hierarchy, the decision-making process and the transparency of the formal and informal guiding policies.
b. **Representation:** Assess gender representation, employment type, age, level of participation in decision making, any systemic barriers or underrepresentation of a specific group.

c. **Economic and political environment:** Understand how local policies affect the organizations operations and how they influence decision making, and the sustainability strategies of the organization
d. **Engagements and partnerships:** Understand how organizations engage with external stakeholders such as the government, NGOs, private sector and global unions.

1.6.2 Gender Differences in Lived Experiences and Barriers

When collecting data gender differences can influence the accuracy, reliability and inclusivity of data. Cultural and gender norms can influence the findings. Recognizing and addressing these differences ensures that data collection methods capture diverse perspectives and lead to more equitable decision making. Some of the gender-based barriers in data collection include:

i. **Access to data collection activities**: social and cultural norms can affect findings.
ii. **Gender bias in data collection instruments**: survey instruments not including gender disaggregated indicators.
iii. **Power dynamics in data collection settings:** social norms that may hinder women from expressing opinions in public.
iv. **Safety and ethical concerns**: women providing data on sensitive topics and may face stigma or retaliation.

It is therefore important to ensure respectful and ethical data collection by employing gender-sensitive methodologies such as utilizing same gender interviewers for sensitive topics to create a safe space for discussion.

1.6.3 Rural versus urban challenges and opportunities

Data collection in rural and urban areas present unique challenges and opportunities due to differences in socioeconomic factors, infrastructure, population density and technological access. Understanding these distinctions is crucial for designing effective and inclusive data collection methodologies.

Some of the challenges and opportunities in data collection in urban vs. rural areas may include:

Challenges in Data Collection	
Rural Setting	**Urban Setting**
Low literacy levels.	Unavailability of respondents due to their busy schedule.
Cultural and social barriers.	Data saturation and survey fatigue.
Dispersed population.	Diverse and complex population.
Limited awareness and trust.	Privacy and security concerns.
Poor infrustructure may hinder access.	
Opportunities in Data Collection	
Community and local support.	High literacy rates.
Willingness to participate.	Diverse data source eg social media data analytics.
	Better access to technology and digital tool (can do online surveys.)

1.6.4 Factors Affecting Research Processes and their Mitigation Measures

Economic, political and social impacts affecting the population and workers can affect quality data gathering. These factors can either facilitate or hinder the research process hence affect the reliability and inclusiveness of data.

Some of these factors and mitigation measures are presented in the table below:

Factors Affecting Research Processes & Their Mitigation Measures

Challenges	Mitigation measures
Economic Factors	
Limited financial resources for research activities.	Seek multiple funding sources including grants, partnerships and sponsorships.
Economic instability may affect research activity.	Plan budget research activitieseffectively.
High cost of data collection.	Collaborate with institutionsto share resources and expertise.
Social Factors	
Cultural barriers andresistance to participate.	Engage local leaders and influencers to build trustwith communities.
Language barrier between researchers and participants.	Provide translated materials and interpretation where necessary.
Low literacy levels can affect comprehension of the research materials.	Use culturally sensitive and context appropriate research approach.
Mistrust of researchers dueto past negative experiences.	Conductawareness and educate participants about the purposeof the research.
Political Factors	
Government regulations that restrict data collection.	Ensure compliance with national and local research regulations.
Political instability affecting access to research areas.	Develop contingency plans in caseof political disruption.
Bureaucratic processes affecting research approvals.	Obtain necessary research permits and approvals in advance.

1.7 Resources

1. https://edwebcontent.ed.ac.uk/sites/default/files/atoms/files/research_ethics_and_integrity_awareness.pdf
2. Dooly, M., Moore, E., & Vallejo, C. (2017). Research ethics. Research-publishing. Net.
3. Israel, M., & Hay, I. (2006). Research ethics for social scientists. Sage.
4. Fujii, L. A. (2012). Research ethics 101: Dilemmas and responsibilities. PS: Political Science & Politics, 45(4), 717-723.

1.8 Assessment

1. **What are the key areas you need to assess in a local organization when conducting research?**
   a. *Organizational structure and governance, economic and political environment, and engagements and partnerships*
   b. *Time employees report to work and their salaries*
   c. *Available facilities and the technology they use*
   d. *Security measures in place*
2. **Which of the following is not a gender barrier in data collection process?**
   a. *Socio-cultural norms*
   b. *Safety and ethical concerns*
   c. *Lack of disaggregation in data collection instruments*
   d. *Lack of partnerships and engagements*
3. **What challenge are you likely to encounter when collecting data in a rural setting?**
   a. *Cultural and social barriers, dispersed population and limited awareness and trust*
   b. *Data saturation and survey fatigue*
   c. *Community and local support*
   d. *High literacy levels*
4. **Which of the following is not a political factor that can hinder/ delay research process?**
   a. *Government mistrust*
   b. *Delays in acquiring research permits due to policies and regulations*
   c. *Political instability and security concerns*
   d. *Financial constraints on research team*
5. **Which of the following is not an advantage of collecting data in urban setting?**
   a. *High literacy rates*
   b. *Readily available sample pools*
   c. *Dispersed population*
   d. *Diverse data sources*

Section 2: Understanding Study Objectives

2.1 Introduction

In this subtopic the participant will learn how to identify key objectives which ensures that the collected data is meaningful, relevant and aligned with the research goals. Analytical objectives help guide data collectors in structuring their methodologies, selecting key variables and interpreting findings in a way that accurately reflects the local context.

2.2 Training Objectives:

By the end of this session, the participant should be able to

a. Be able to align analytical objectives with the broader study goals and outcomes;
b. Identify organizational needs;
c. Be able to link analytical objectives to organizational objectives.

2.3 Definition of Terms

a. **Objective:** A clear statement of what a researcher aims to achieve in a study.
b. **Analytical Objectives:** A specific goal that guides an analysis.
c. **Organizational Objectives:** Medium- and short-term goals that help an organization to achieve long term goals. They help an organization or project to evaluate its performance and operations.

2.4 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.5 Duration: 15 min

2.6 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.7 Contents

a. Define the purpose and objectives of the research
b. Identify analytical objectives
c. Assess organizational needs
d. Link analytical objectives to organizational objectives

2.8 Training Content

2.8.1 Defining the Purpose and Objectives of the Research

Before conducting research, it is important to determine the purpose and the main objectives of the study. This will ensure that the research is focused, relevant and aligned with the organization's mission. First, researchers need to define the purpose of the study, which should answer why the research is being conducted, challenges the research is addressing and how the findings will be used. Second, researchers

need to identify the main objectives which should be SMART (Specific, Measurable, Achievable, Relevant, and Time-bound).

Example:

If the purpose of the study is "*To improve agricultural productivity of small-scale farmers".* The objectives would include:

i. *Identifying key barriers to accessing extension services*
ii. *To improve crop yield by 50%*
iii. *To assess the effectiveness of new farming technologies*

Thirdly, researchers will need to define what the organization wants to learn. For instance, what new interventions can the organization provide, and how the findings can support decision-making and policy development.

2.8.2 Identify Analytical Objectives

Analytical objectives are specific goals that guide the process of analyzing data. In other words, they are the objectives to be achieved through data analysis. They guide the research process in ensuring that data is collected, processed and interpreted in a meaningful way.

Example of an analytical objective:

Assessing the impact of Information Communication Technology (ICT) based agricultural extension services in small scale farmer' productivity.

Steps of formulating analytical & objectives include:

Steps	Description
Define the problem	What specific challenge is the researcher addressing researchers objectives is "Assessing the impact of ICT farmer productivity".
Identify key variables & indicators	What points does the researcher need to measure the problem? Example of indicators"Access to ICT services, crop yield, income levels, adoption of new technologies".
Determine the type of analysis	What points does the researcher need to measure the problem? Example of indicators "Access to ICT services, crop yield, income levels, adoption of new technologies". (Use descriptive analysis or predictive (forecast trends)"eg comparing crop yields of farmers who have adopted ICT services and those who don't.
Set measurable outcomes	What specific results is the researcher are you trying to obtain? eg if farmers using ICT services have an increase in productivity.
Relevance to stakeholders	Who will use the findings? Provide understanding to policy makers on how to expand digital agriculture services.

2.8.3 Assess Organizational Needs

Assessing organizational needs before a study helps to ensure the data collection efforts align with the organization's goal, improve efficiency and support decision-making. This process involves evaluating an organization's capacity, gaps, opportunities and resources.

This will involve:

- ✓ Defining the organizational goals and priorities and how data can support the goals.
- ✓ Identifying key data needs, and determine whether the organization needs new data or if they can use existing data.
- ✓ Determine the data collection methods you intend to use, and if they are efficient, reliable and scalable.

- ✓ Assess organizational capacity to determine if the organization has trained staff, financial resources and the digital tools to collect and manage data effectively.
- ✓ Identify barriers and challenges that might affect the data collection and utilization process. This may include poor stakeholder engagement, limited funding, lack of technical expertise and inconsistent data quality.

2.8.4 Link Analytical Objectives to Organizational Objectives

For data collection and analysis to be meaningful, analytical objectives must align to the organization's broader goal. This will ensure that findings generated from data contribute directly to strategic decision-making, program improvement and resource optimization.

Steps to link analytical objectives to organizational objectives.

1. **Identify organizational and key analytical objectives**
   Using the organizational objectives and the analytical objectives generated from the previous sessions: **Organizational objective:** *"To improve the agricultural productivity of small-scale farmers"*
   Analytical objective: *"The impact of ICT based agricultural extension services in small scale farmer' productivity"*
2. **Map analytical objectives to organizational goals**
   Next to create direct links between data findings and decision-making process

Here is an example

Organizational Objective	Analytical Objective	Required Data	Expected Outcome
Improve agriculture productivity.	The impact of ICT based agriculture extension service in small scale farmers productivity.	Access to ICT services crop yield, income levels, level of adoption of new technologies.	Develop digital based agricultural initiatives.

2.8 Resources

1. Competing on Analytics by Thomas H. Davenport – A guide to data-driven decision-making.
2. The Data Warehouse Toolkit by Ralph Kimball – Insights into structuring analytics.
3. Harvard Business Review (HBR) articles on data-driven decision-making.
4. Kaggle (for datasets to practice defining analytical objectives).

2.9 Assessment

1. **What is the primary purpose of an analytical objective?**
   a) *To define the overall research purpose*
   b) To determine what data is needed and how it will be analyzed
   c) To create a marketing strategy
   d) To replace decision-making with automation
2. **Which characteristic is not essential for a strong analytical objective?**
   a*) Specificity*
   b) Measurability
   c) Alignment with personal opinions
   d) Relevance to the organization's needs
3. **When defining analytical objectives, which of the following is most important?**
   a) *Data availability*
   b) Company budget
   c) Employee preference
   d) Market competition
4. **Why is it important to link the organization objectives to the analytical objectives of a study?**
   a. *It ensures that findings generated contribute directly to strategic decision-making*
   b. *It ensures that you chose the right software for analysis*
   c. *It helps to generate the purpose of the study*
   d. *It ensures that you write a good report*
5. **What are the steps for formulating analytical objective?**
   a. *Identify key variables and indicators, set measurable outcomes, relevance to stakeholders, define the problem*
   b. *Define the problem, identify key variables and indicators, determine the type of analysis, set measurable outcomes, relevance to stakeholders*
   c. *Define the problem, determine the type of analysis, identify key variables and indicators, set measurable outcomes, relevance to stakeholders*
   d. *Define the problem, identify key variables and indicators, set measurable outcomes, determine the type of analysis, relevance to stakeholders*

Section 3: Define the Research Problem

3.1 Introduction

This topic will teach the participant on how to define a research problem, how to determine who will benefit from the research objectives being completed and how to concisely and clearly develop the hypotheses to the research.

3.2 Training Objectives

By the end of this session, the participant should be able to

a. Be able to define a study objective;
b. Understand the research audience and stakeholders;
c. Be able to determine a research hypothesis.

3.3 Definition of Terms

a. **Research Problem:** A statement about a concern, a condition to be improved or a troubling question that exists within an existing practice that points to a need for meaningful understanding and deliberate investigation.
b. **Research Audience:** The process of gathering information about people who use a service or engage with an organization. This information can be used to inform or measure success of a strategy or evaluate performance.
c. **Stakeholders:** Any individual, group or institution with vested interest or stake in the decision making or activities of a business, organization or project. Stakeholders can have direct or indirect influence in activities of an organization or project.
d. **Stakeholder mapping:** A process of identifying and analyzing people or a group who have an interest in a project. It helps to understand their perspective, influence and how they impact the project's success.
e. **Hypothesis:** An assumption that is made based on some piece of evidence. It is the initial point of any investigation that translates the research question into predictions.
f. **Variable**: Any characterized number or quantity that can be measured or counted and takes different values within a dataset.

3.4 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.5 Duration: 15 Minutes

3.6 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.7 Contents

a. Defining research problem
b. Assess research audience and stakeholders

c. Determining research hypothesis

3.8 Training Content

3.8.1 Defining Research Problem

A well-defined research problem ensures that the study remains relevant, focused and actionable. Researchers may therefore be required to use background information and organizational needs to identify problems.

Here is the step to follow when defining a problem:

i. **Access Background Information/Desk Review:** To have a well-defined problem, there is a need to examine the trends within the area of study, gaps or challenges emerging from the target population. It will also be required to examine inequalities, inefficiencies and any stakeholder concerns or historical patterns highlighting a recurring issue.
ii. **Align with Organizational Objectives:** Ensure that the objectives align with the organizational objectives and be relevant to the stakeholder, policy makers or donors' goals.
iii. **Framing the Problem Statement:** The problem statement should be clear, concise and specific. It should be able to identify root causes of a problem and not just symptoms. Further, it should highlight who is affected and how and be able to connect the issue to the broader social, economic, political and environmental factors.
iv. **Prioritize Problem Based on Availability of Resources:** After defining the research problem, ensure that it is manageable and within the available resources and timeframe. It should also be ensured that there is evidence that supports the problem i.e. secondary data that show there is a gap. Finally, the organization or institute in charge of the research should be able to realistically influence change within the target area.

3.8.3 Defining Hypothesis

Hypothesis is a testable statement or assumption that explains a phenomenon or relationship between variables. It is often based on prior knowledge, observation, or research and serves as the basis for further investigation. A hypothesis can be formulated as a prediction that can be tested through experiments, surveys or data analysis.

Types of Hypotheses

i. **Null hypothesis (Ho):** It assumes no relationship or effect between variables. It assumes that any observed difference or patterns in data are due to random chance rather than real effect. The null hypothesis is tested using statistical analysis to determine whether to reject or not. If the evidence strongly contradicts with the null hypothesis it is rejected in favor of an alternative hypothesis.

Example: "There is no significant impact of business skills training and income levels of SMEs"

ii. **Alternative hypothesis (H1):** It is a statement that suggests that there is a significant relationship between two or more variables. It is what a researcher aims to support through data collection and analysis.

Example: "Business skills training significantly increases income levels of SMEs"

iii. **Directional hypothesis:** Directional hypothesis not only predicts that a relationship exists but it also specifies the direction of that relationship. It states whether one variable decrease or increases because of another.

Example: *"SMEs who do marketing are likely to receive more customers"*

"Students who receive extra tuition are likely to perform better"

iv. **Non-directional hypothesis:** predicts there is a relationship between two variables but does not specify direction of the relationship. It only highlights that there is an association that exists but it doesn't indicate whether one variable increase or decreases as a result of the other.

Example: *"There is a relationship between access to microfinance and small business success rate"* This statement does not specify *how* access to microfinance can impact businesses.

3.9 Resources

1. https://libguides.usc.edu/writingguide/introduction/researchproblem#:~:text=A%20research%20problem%20is%20a,for%20meaningful%20understanding%20and%20deliberate

3.10 Assessment

1. What is the difference between a null hypothesis (H_0) and an alternative hypothesis (H_1)?

a. The null hypothesis predicts an effect, while the alternative hypothesis predicts no effect.

b. The null hypothesis assumes no significant effect, while the alternative hypothesis suggests a relationship, effect or significance difference between variables.

c. The null hypothesis is always correct.

d. The alternative hypothesis must always be rejected.

2. Which of the following is an example of a directional hypothesis?

a) "There is a relationship between exercise and weight loss."

b) "Students who get more sleep perform better on exams than those who get less sleep."

c) "There is no significant difference in productivity between remote and in-office workers."

d) "Employee engagement levels vary by department."

3. Why should a hypothesis be based on existing research or observations?

a) To ensure personal opinions guide the research

b) To increase the chances of proving it correct

c) To make it easier to justify the study's relevance

d) To avoid any need for data collection

Section 4: Applying Multiple Lenses: A Framework for Inclusive and Holistic Analysis

4.1 Introduction

This section will enlighten the participant on how to apply lenses during a research study. This knowledge will help to be able to identify biases, address inequalities and design effective interventions.

4.2 Training Objectives

By the end of this session, the participant should be able to

a. Understand gender lens;
b. Be able to apply gender lens during a study.

4.3 Definition of Terms

i. **Gender Lens:** An analytical framework that considers gender when making decisions. The goal of gender lens is to promote equality and reduce gender inequality.
ii. **Trauma-informed Lens:** A trauma informed lens is a perspective that recognizes the widespread impact of trauma and integrates this understanding into policies, practices, and interactions to avoid re-traumatization and promote healing. It involves viewing situations, behaviors and systems through an awareness of how past trauma may shape a person's responses and experiences.
iii. **Climate Lens:** A climate lens in data collection is an approach that integrates climate consideration into research evaluation, and decision-making processes. It ensures that that data collection efforts account for the impacts of climate change, environmental sustainability, and resilience in policies, programs and interventions.
iv. **Human Rights Lens:** Human rights lens in data collection is an approach that ensures that data is gathered, analyzed and used in ways that uphold human dignity, equality, and fundamental rights. It focuses on fairness, non-discrimination, participation, and accountability, ensuring data collection does not reinforce inequalities or violate individuals' rights.
v. **Conformity Lens:** Conformity lens in data collection is an approach that ensures that data collection aligns with established standards, regulations, and norms within a specific field or context. It emphasizes on consistency, accuracy, and compliance with legal, ethical, and institutional frameworks.

4.4 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.5 Duration: 10 min

4.6 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

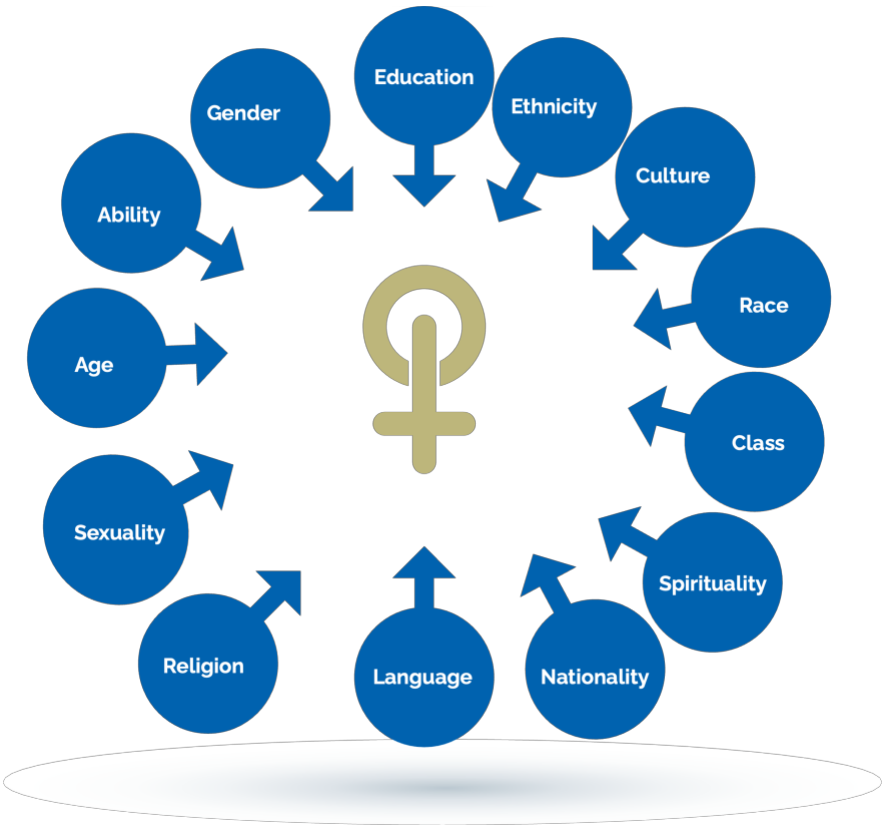
4.7 Contents

a. Gender lens
b. Trauma Informed Lens
c. Climate lens
d. Human rights lens
e. Conformity lens

4.8 Training Content

4.8.1 Gender Lens

Gender lens examines how social, cultural and economic factors affect people differently based on gender. When examining gender lens, researchers need to look at how policies, programs and decisions affect men, women and gender diverse individuals differently, and how systematic barriers affect their participation and any gender specific challenges in issues being studied. Issues of power dynamics and intersectionality should be considered when applying gender lens.



Example: *In agriculture gender lens might reveal that female farmers have less access and control over land than men"*

This lens will ensure there are gender equitable policies, interventions and resource distribution. The gender lens ensures that gender dynamics, disparities, and power

relations are systematically analyzed and addressed, hence leading to inclusivity and equitable decision making.

4.8.2 Trauma Informed Lens

Trauma informed lens recognizes the impact of past or ongoing trauma on an individual or a community. This is important when working with vulnerable populations affected by discrimination, poverty, violence and displacement.

Researchers may need to consider how trauma has shaped the lives of the participants, if the research method is respectful and a non-trigger, and if the research is safe, supportive and empowering.

> *Example: Gender based violence survivors require privacy, sensitivity and ethical data collection methods.*

Such lenes will ensure compassionate, ethical and respectful research practices. Hence avoid assumptions that blame or stigmatize affected persons. The trauma informed lens helps to adopt collaborative and participatory approaches that empower trauma affected population for instance engaging trusted community leaders and mental health professionals to provide counseling and ensure safe participation.

4.8.3 Climate Lens

Climate lens examines how environmental factors intersect with economic, political and social systems. This is essential in addressing climate change and sustainability challenges. When examining climate lens, researchers need to consider how climate change has impacted different communities, the environmental risks associated with the research topic and how the findings can contribute to climate resilience and sustainability

> *Example: A study on rural livelihood might reveal climate change has affected smallholder farmers through soil degradation, drought or floods.*

Climate lens can be applied while developing hypothesis and stakeholder engagement in climate related risks. Applying climate lens ensures that the research and engagement processes are aligned with sustainability, resilience and equity goals. Hence produce more relevant findings, drive climate-smart policy recommendations, and foster discussions on adaptation and mitigation efforts.

4.8.4 Human Rights Lens

Human rights lens ensures that research and development are equitable, inclusive and sustainable. This helps to safeguard marginalized groups and ensure that environmental, social and economic policies are rights based, fair, and impactful. For instance, rights lens will ensure equitable access to essential services such as healthcare, education and social protection.

Example: Kenya introduced the free education policy in 2003 to ensure that all children regardless of gender or economic background have access to basic education

4.8.5 Conformity Lens

Conformity lens ensures that policies, projects, behavior align with established standards, norms, regulations and social expectations. This lens helps to prevent legal consequences and reputational damage. Since they help to streamline compliance to avoid regulatory bottleneck. The conformity lens can be applied in governance and public policy, climate action and environmental sustainability and human rights and social protection.

Example: The United Nations convention on rights of Persons with Disabilities ensures that countries modify policies, infrastructure and services to enable accessibility and equal participation of persons with disabilities.

4.9 Resources

1. https://www.idinsight.org/article/how-were-using-a-gender-lens-to-increase-social-sector-partners-impact/

4.10 Assessment

1. **Why is applying a gender lens important in research?**
   a. *To ensure that research findings are equally applicable to all genders*
   b. To exclude gender-related factors from research outcomes
   c. To focus only on women's experiences
   d. To ignore gender differences in data
2. **When using a trauma-informed lens in research, researchers should?**
   a. *Avoid sensitive topics at all costs*
   b. Ensure voluntary participation and provide mental health support
   c. Ask detailed personal questions to gather more data
   d. Assume that all participants are comfortable sharing trauma-related experiences
3. **What is the importance of climate lens in research?**
   a. *To ensure research findings consider resilience and environmental sustainability*
   b. To exclude communities affected by climate change
   c. To reduce the role of environmental factors in social studies
   d. To focus only on the economic impact of climate change
4. **How does applying conformity lens in research impact policy recommendations?**
   a. It encourages diverse perspectives and challenge traditions
   b. It disregards social expectations to promote radical change
   c. It prevents stakeholders in decision making
   d. It ensures all findings align with pre-existing norms and standards
5. **How can applying human rights lens improve research on marginalized community?**

a. *By focusing only on legally recognized rights and ignoring cultural context*
b. *By limiting data collection to official government reports*
c. *By ensuring that ethical considerations, inclusivity and equality shape the research*
d. *By prioritizing quantitative data over qualitative lived experience*

Section 5: Data Collection Plan and Establishment of Research Team

5.1 Introduction

This session introduces the participant to advantages of proper planning for data collection during research activities.

5.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand why planning is essential in data collection;
b. Identify critical elements of planning such as timelines, personnel, tools and methodologies;
c. Examine common challenges of failed data collection due to inadequate preparation;
d. Outline a structured approach to planning, including defining objectives; resource allocation, risk assessment, and contingency planning.

5.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

5.4 Duration: 15 min

5.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

5.6 Contents:

a. Purpose of planning
b. Key components of planning

5.7 Training Content

5.7.1 Purpose of Planning

Planning is essential in data collection because it lays the groundwork for a systematic and structured approach. It ensures that data collection efforts are efficient, accurate and aligned with the intended objectives. The key reasons why planning is crucial include:

i. **Enhancing Coordination:** Planning for data collection provides a roadmap for data collectors, ensuring consistency and reducing confusion in the data collection process. Planning helps structure the data collection process, ensuring that all necessary resources, tools and personnel are available and allocated appropriately. It also prevents unnecessary delays and optimizes the use of time and manpower.
ii. **Establishing Clear Objectives:** Planning defines the purpose of data collection, ensuring that all team members understand what data is needed and why. A structured plan clarifies the roles and responsibilities of each data

collector. This ensures that training is tailored to meet specific needs, equipping collectors with the necessary skills, knowledge and tools for effective data collection.

iii. **Resource Allocation:** It helps in organizing resources such as personnel, equipment, and time, reducing waste and maximizing efficiency. A well-planned approach minimizes resource wastage and reduces the likelihood of costly errors. It also helps allocate budget and resources efficiently, therefore, ensuring that data collection is conducted within the financial constraints.

iv. **Minimizing Errors:** A well-structured plan helps in identifying potential pitfalls and implementing quality control measures to improve data accuracy and reliability. With a well-defined plan, data collectors can follow standardized procedures, minimizing errors and inconsistencies. Pre-established protocols and guidelines ensure that data is collected uniformly across different teams and locations.

v. **Ensuring Compliance:** Planning incorporates ethical considerations and legal requirements, ensuring that data collection adheres to industry regulations and standards. Ethical guidelines and legal regulations, such as informed consent, confidentiality and data protection laws helps to build trust with respondents and ensures compliance with industry standards.

vi. **Improving Decision-Making:** With a well-planned approach, data collectors and decision-makers can analyze information more effectively, leading to better-informed decisions. Planning provides a framework that allows for adjustments when unforeseen circumstances arise. By having contingency plans in place, data collectors can adapt to challenges without compromising the integrity of the data. Planning allows potential challenges to be identified in advance, such as logistical constraints, technical issues, or data privacy concerns. Mitigation strategies can be developed to address these risks before they impact the data collection process.

5.7.2 Key Components of Planning

To ensure an effective and structured data collection process, planning must incorporate several key components. These components include:

a. **Timelines:** Establishing a clear schedule for data collection, including start and end dates, deadlines for specific tasks and contingency time for unforeseen challenges.

b. **Personnel:** Defining roles and responsibilities for each team member involved in the data collection process to ensure accountability and efficiency.

c. **Tools and Equipment**: Identifying the necessary resources, such as data collection forms, mobile devices, software, and other tools required for accurate data gathering.

d. **Methodologies:** Selecting appropriate data collection methods, such as surveys, interviews, observations or automated data capture based on the research objectives.

e. **Data Quality Control Measures:** Implementing mechanisms such as double-checking entries, regular progress reviews and validation processes to ensure accuracy and consistency.

f. **Logistics and Accessibility:** Planning for transportation, internet connectivity and availability of respondents to minimize disruptions in the data collection process.

g. **Risk Assessment and Mitigation:** Identifying potential obstacles such as technical failures, respondent unavailability or environmental factors and preparing strategies to address them.

5.8 Resources

1. Phillips, P. P., & Stawarski, C. A. (2008). *Data collection: Planning for and collecting all types of data*. John Wiley & Sons.
2. Bina, A. (2023). Developing a Data Collection Plan for Pre-tender Cost Estimation in Design Phase.
3. https://www.fldoe.org/core/fileparse.php/7771/urlt/0084814-phaseiii.pdf
4. https://hal.science/hal-03741847/document
5. https://www.jbassoc.com/wp-content/uploads/2018/03/Module-2-Data-Collection-Home.pdf
6. https://research.trademarkafrica.com/wp-content/uploads/2022/11/Data-standardisation-guidelines2-1.pdf

5.9 Assessment

1. **Which of the following is not a reason for planning for a data collection activity?**
   a. *It helps in organizing resources*
   b. *It helps to identify potential pitfalls and implement quality control measures*
   c. *It helps to define the purpose of data collection and ensures that the team members understand the objective of the study*
   d. *It ensures that the respondents are well compensated*
2. **Why is budgeting important when planning for data collection?**
   a. *It helps to minimize resource wastage and reduce the likelihood of costly errors*
   b. *It helps to identify the respondents who should be compensated*
   c. *It ensures that the organization improves the decision-making process*
   d. *It reduces the errors that can be encountered in the field*
3. **Which of the following is not a key component of planning for data collection?**
   a. *Planning for transport and availability of respondents*

b. *Identifying the necessary resources required for data collection*
c. *Implementing data quality measures*
d. *Using data collection tools from other organization because it has been tested*

4. **Why is it important to define roles and responsibilities for each team member before the data collection process?**
   a. *It ensures that the team members know how much they will be compensated*
   b. *To understand their qualifications*
   c. *To ensure there is accountability and efficiency*
   d. *To understand which questions to ask them during the interview*
5. **Which of the following is not an advantage of planning for data collection?**
   a. *Planning helps to clarify the roles and responsibilities of each team member*
   b. *Planning ensures that the project managers are well compensated*
   c. *Planning helps to improve data accuracy and consistency*
   d. *Planning helps to structure the data collection process ensuring all the necessary resources are appropriately allocated*

Section 6: Project Team Identification

6.1 Introduction

A well-structured project team is essential for successful data collection. This section introduces the participant on how to team members based on their expertise, experience, and ability to work collaboratively within a structured framework. This section outlines the key positions, skills required, team composition and administrative capacity necessary for a high-functioning data collection team.

6.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand the importance of a well-structured team;
b. Identify key positions and responsibilities;
c. Determine essential skills for each role.

6.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

6.4 Duration: 10 min

6.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

6.6 Contents:

a. Having teams decide on the needed roles and skillsets for each role, composing their teams, creating roles and responsibilities, administrative capacities and organizational structure within each team (reporting structure and team leaders)

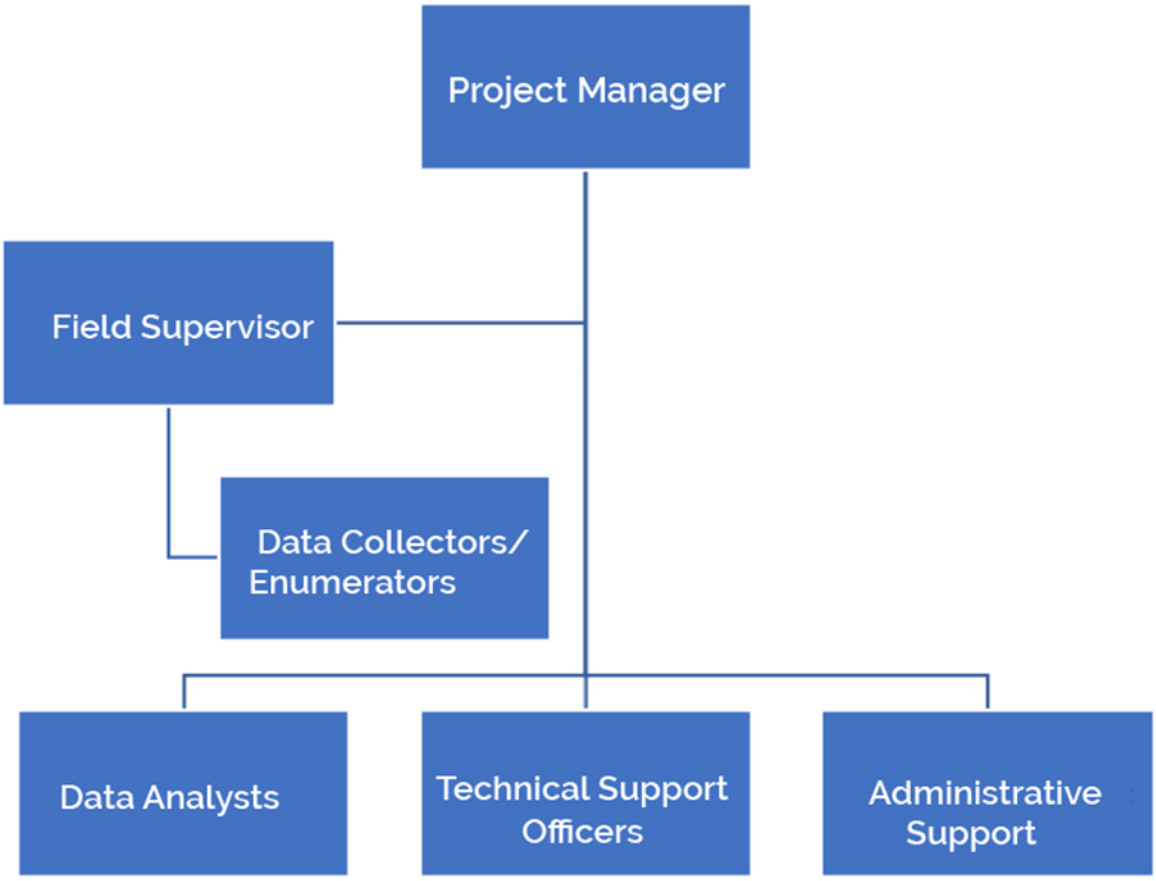
6.7 Training Content

6.7.1 Key Roles in a Data Collection Team

Key Roles in a Data Collection Team		
Team Member	Roles	Skills set
Project manager	Oversees the entire data collection process. Ensures objectives, timelines andresources are alignedwith project goals. Coordinates with stakeholders and ensures compliance with ethical and regulatory standards.	Leadership, strategic planning, problem-solving, communication, budgeting.
Field Supervisor	Manages data collectors in the field and ensures adherence to protocols. Troubleshoots issues encountered during data collection. Conducts quality control checks to maintain data accuracy.	Team coordination, quality control, conflict resolution, adaptability.
Data Collectors /Enumerators	Conduct surveys, interviews and field observations. Ensure accurate data entry and recording. Adhere to ethical guidelines and ensure participant confidentiality.	Attention to detail, data entry proficiency, interpersonal skills, ethical awareness.
Data Analyst	Processes and analyzes collected data. Identifies trends, inconsistencies, and insights from the data. Works closely with project managersto interpret findings.	Statistical analysis, data visualization, software proficiency (Excel, SPSS, R, etc.).
Technical Support Officer	Provides IT and software support for digital data collection tools. Ensures proper functionality of data collection equipment and platforms. Troubleshoots technical issues faced by field staff.	IT troubleshooting, database management, software knowled.

6.7.2 Structuring the Data Collection Team

A well-structured data collection team is critical for efficiency and accuracy. The team should be organized based on project scope and complexity. Below is a recommended structure:



6.7 Resources

a. Phillips, P. P., & Stawarski, C. A. (2008). *Data collection: Planning for and collecting all types of data*. John Wiley & Sons.
b. Bina, A. (2023). Developing a Data Collection Plan for Pre-tender Cost Estimation in Design Phase.
c. https://www.fldoe.org/core/fileparse.php/7771/urlt/0084814-phaseiii.pdf
d. https://hal.science/hal-03741847/document
e. https://www.jbassoc.com/wp-content/uploads/2018/03/Module-2-Data-Collection-Home.pdf
f. https://research.trademarkafrica.com/wp-content/uploads/2022/11/Data-standardisation-guidelines2-1.pdf

6.8 Assessment

1. **What is the role of enumerators in research?**
   a. *They conduct surveys, interviews and field observations*
   b. *They oversee the data collection process*
   c. *They manage the data collection*

d. They ensure proper functioning of the data collection equipment

2. **Which of the following position is not considered key in data collection?**
   a. Project manager
   b. Field supervisor
   c. Data collectors
   d. Donors
3. **Why are technical support officers important team members in data collection?**
   a. They ensure that the field data collection runs smoothly
   b. They process and analyze data
   c. They help to identify inconsistencies from the data collected
   d. They conduct quality checks to maintain quality data
4. **What skill sets does an enumerator required to possess?**
   a. He/she should have budgeting skills
   b. He/she should be attentive to details, and possess data entry proficiency
   c. Should have statistical analysis skills
   d. Should have strategic planning skills
5. **What is the role of a data analyst?**
   a. To troubleshoot technical issues faced by field teams
   b. To manage data collection and ensure adherence to protocols
   c. To ensure objectives are aligned to the project goals
   d. To process and analyze data

Section 7: Developing A Data Collection Plan

7.1 Introduction

A well-structured data collection plan is essential for ensuring accuracy, efficiency and reliability in data gathering. This section will focus on the critical components of a data collection plan, including defining objectives, selecting methodologies, allocating resources and addressing potential challenges. A strong plan lays the groundwork for successful data collection efforts and minimizes risks associated with incomplete or inconsistent data.

7.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Define the purpose and scope of a data collection plan;
b. Identify key components required for data collection;
c. Develop a structured approach to planning data collection activities;
d. Assess and allocate resources efficiently;
e. Anticipate challenges and establish mitigation strategies.

7.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

7.4 Duration 10 min

7.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

7.6 Contents:

a. Create a documented plan of the organization's or partner's research including dates and timeline, action items for each person, recruiting and contracts, budget and other key components to conducting research and building a team
b. Develop a system of communication and collaboration within research teams and amongst different teams to ensure efficiency, and quality control of data and the research process

7.7 Training Content

3.7.1 Developing a Data Collection Plan

A data collection plan helps in mapping out the type of data that is needed for collection and the ways in which they will be collected. The core purpose is to ensure that researcher's data collection is targeted, efficient and reliable, providing meaningful insights for the study.

Data collection plans should be developed at the start of a project or study, before any data is collected. Typically, this responsibility falls to project leaders, researchers, data analysts or a designated team member with expertise in data management.

Creating a plan for data collection involves six steps:

1. **Assemble the data collection team.** Establish an interdisciplinary team of stakeholders as subject matter experts to ensure the planning reflects various perspectives, which will help build consensus about how to collect data and monitor performance.
2. **Define what to measure.** Decide what question the data will answer and why the data are important to collect.
3. **Identify what type of data to collect to answer the question.** Will data be quantitative (numbers or statistics) or qualitative (words or meanings)? Qualitative data often requires a different data collection approach (i.e. focus groups or interviews) compared to the collection of quantitative data (i.e., surveys.
4. **Decide how to collect the data**. Consider whether data of interest already exists. New data may require a different method for collecting information such as a survey or focus group.
5. **Identify the resources needed to collect the data**. In some cases, engaging with community partners or agencies will be necessary to enrich the data. If data collection involves a new process or report, consider what technology, staffing and financial resources you will need.
6. **Decide who will collect the data**. Identifying who is responsible for data collection will help to solidify a schedule and ensure everyone is on the same page about roles and responsibilities for collecting data. Consider whether there are multiple points of data collection and how to manage handoffs.

Example of a data collection plan

Project Name						Project Manager		
Who	What			When	Why	How		Other
Responsible	Operational Definitions	Data Type	Sample size/ Frequency	Date and Time	Questions to be answered	Recording method	Collection Method	Comments

7.8 Resources

1. CDC. "Using Ordered Response Options to Collect Evaluation Data." Evaluation brief no. 23. 2018. https://www.cdc.gov/healthyyouth/evaluation/pdf/brief23.pdf.

7.9 Assessment

1. **Why is data collection plan necessary?**
   a. *It helps in mapping out the type of data needed and ways in which they can collected*
   b. *It's important during analysis*
   c. *It helps to locate respondents*
   d. *It is important during report writing*
2. **When should a data collection plan be developed?**
   a. *Before the data collection process*
   b. *At the end of project /after the data collection*
   c. *During the data collection process*
   d. *Before the project begins*
3. **What is the primary purpose of deciding how to collect data when planning for data collection?**
   a. *It helps to decide what questions the data will answer*
   b. *Its helps to know whether the data of interest already exists*
   c. *It helps to identify who is responsible for data collection*
   d. *It helps to enrich the data collected*
4. **Which of the following is not considered a type of data?**
   a. *Qualitative data*
   b. *Quantitative data*
   c. *Observation*
   d. *Both A and B above*
5. **Which of the following is not a core purpose for developing a data collection plan?**
   a. *It ensures that the data collected is efficient*
   b. *It ensures that the data collected is reliable*
   c. *It ensures that the data collected provides a meaningful insight*
   d. *It ensures that all the target population are reached*

CHAPTER 6.

Pilot Testing

Section 1: Overview of Pilot Testing

1.1 Introduction

Pilot testing in data collection is a critical step in ensuring the accuracy, reliability, validity and efficiency of a research tool before implementation. It involves conducting a test run of the data collection process to identify potential issues, such as unclear survey questions, logistical challenges or data quality concerns. By testing the methodology, tools and procedures in real-world conditions, researchers can refine their approach, minimize errors and enhance the validity of their findings.

1.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understanding the purpose of pilot testing;
b. Explain variables in sample size determination;
c. Discuss error analysis.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 10 min

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents:

a. Importance of pilot testing

1.7 Training Content

1.7.1 Importance of Pilot Testing

Proper preparation is essential for a successful pilot test, ensuring that the data collection process runs smoothly and effectively. The key steps in preparing for pilot testing include defining objectives, developing data collection tools, training personnel and establishing logistics.

Pilot testing:

- Detects flaws in survey/questionnaire design (e.g. ambiguous questions, confusing formats).
- Identifies operational issues such as logistical challenges, respondent engagement and data recording errors.

- Ensures data collectors are well trained and follow standardized procedures.
- Improves the accuracy, reliability, and consistency of the final data set.
- Saves time and resources by preventing costly errors in full-scale implementation

1.8 Resources

1. https://apstraining.com/wp-content/uploads/Rev3_APS-Training-Manual-PPRRC-2007.pdf
2. https://bibliotekanauki.pl/articles/2033043.pdf

1.9 Assessment

1. **Which of the following statements is false about pilot testing?**
   a. *It involves conducting a test run of the data collection process*
   b. *It helps to refine the methodology and approach*
   c. *It helps to enhance the validity of the study findings*
   d. *It helps to reach all the stakeholders*
2. **Which of the following is not a factor to be considered when preparing for pilot testing?**
   a. *Defining objectives*
   b. *Collecting data*
   c. *Developing data collection tools*
   d. *Training personnel*
3. **Why is pilot testing important?**
   a. *It helps to identify errors and inconsistencies in the data collection tools*
   b. *It helps to identify and recruit enumerators*
   c. *It helps to analyze data*
   d. *It helps to identify the right respondents for the study*

Section 2: Identify Pilot Testing Sample

2.1 Introduction

Selecting the right sample for pilot testing is a crucial step in ensuring the effectiveness of the data collection process. A well-chosen sample helps identify potential issues in survey design, question clarity and overall feasibility before full-scale implementation.

2.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand the Importance of Sample Selection in Pilot Testing;
b. Develop a Plan for Recruiting and Managing Pilot Test Participants;
c. Evaluate the Effectiveness of the Selected Sample.

2.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.4 Duration: 20 min

2.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents:

a. Methods for Identifying Pilot Testing Sample
b. Steps in Selecting the Sample
c. Ethical Considerations in Pilot Testing Sample Selection
d. Evaluating the Pilot Testing Effectiveness

2.7 Training Content

2.7.1 Methods for Identifying Pilot-Testing Sample

- **Random Sampling**: Selecting participants randomly from the study population.
- **Purposive Sampling**: Choosing participants who are **most relevant** to the study.
- **Convenience Sampling**: Selecting participants based on **ease of access** (e.g. those nearby).
- **Stratified Sampling**: Ensuring key subgroups (e.g. age, gender, region) are included.

2.7.2 Steps in Selecting the Sample

- Define the study population and identify key characteristics.
- Choose the sampling method based on the study needs.
- Determine the sample size appropriate for testing.
- Recruit participants and obtain their informed consent.
- Conduct the pilot test and gather feedback.

2.7.3 Ethical Considerations in Pilot Testing Sample Selection

- Ensuring voluntary participation and obtaining informed consent.
- Maintaining confidentiality and anonymity of respondents.
- Avoiding bias in sample selection to ensure fair representation.

2.7.4 Evaluating the Pilot Testing Sample Effectiveness

- Assess whether the sample captures key challenges in data collection.
- Identify any gaps or biases in the sample selection.
- Make necessary adjustments before proceeding to full-scale data collection.

a. Resources

1. Kunselman, A. R. (2024). A brief overview of pilot studies and their sample size justification. Fertility and sterility, 121(6), 899-901.
2. Badia-Fabregat, M., Oller, I., & Malato, S. (2018). Overview on pilot-scale treatments and new and innovative technologies for hospital effluent. Hospital Wastewaters: Characteristics, Management, Treatment and Environmental Risks, 209-230.

b. Assessment

1. **Which of the following sampling methodology is suitable when identifying pilot-testing sample?**
   a. *Purposive sampling*
   b. *Convenience sampling*
   c. *Random sampling*
   d. *All of the above*
2. **Why is it important to evaluating pilot testing results?**
   a. *To identify any gaps and challenges that may be encountered during data collection*
   b. *To discard and develop new tools*
   c. *To help in budgeting for the data collection process*
   d. *To ensure voluntary participation of the respondents*
3. **Which of the following is not considered a key step when selecting pilot test sample?**
   a. *Define the study population and identify key characteristics and choose the sampling method based on the study needs.*
   b. *Conduct collection and data cleaning*
   c. *Determine the sample size appropriate for testing and recruit participants and obtain their informed consent.*
   d. *Conduct the pilot test and gather feedback.*
4. **What ethical considerations should you consider when selecting sample for pilot testing?**
   a. *Avoiding bias in sample selection to ensure fair representation.*
   b. *Maintaining confidentiality and anonymity of respondents.*
   c. *Ensuring voluntary participation and obtaining informed consent.*

d. All of the above

5. **Why is convenience sampling important when selecting sample size for pilot testing?**
   a. It ensures that all data is gathered for analysis
   b. It fastens the data collection process
   c. Its cost effective
   d. It helps to identify the best data analysis software to use

Section 3: Orientation and Practice of Data Software and Tools

3.1 Introduction

This training is designed to equip data collectors with the essential knowledge and practical skills needed to effectively use data collection software and tools.

3.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand the importance of accurate and ethical data collection;
b. Navigate and operate key data software tools efficiently;
c. Collect, enter, and manage data using digital platforms;
d. Identify and troubleshoot common software-related issues;
e. Ensure data security and compliance with best practices.

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration 10 min

3.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents:

a. Introduction to Data Software & Tools

3.7 Training Content

3.7.1 Choosing the Right Data Software for the Team

When selecting a data collection software, consider:

- **Type of Data to Be Collected** – Surveys, GPS, images, qualitative responses, etc.
- **Project Scale & Scope** – Number of respondents, locations and team members.
- **Offline vs. Online Capability** – Ensures data collection continues without an internet connection.
- **Security & Compliance** – Protecting sensitive data and meeting ethical and legal requirements.
- **Integration & Compatibility** – Must work with other tools (Excel, SPSS, R, GIS, etc.).

Popular Data Collection Software Options

- Google Forms – Simple surveys, requires internet access.
- KoboToolbox – Offline-friendly, used for humanitarian and field data collection.

- ODK (Open Data Kit) – Offline data entry, supports multimedia.
- REDCap – Used for research and medical data collection.
- SurveyCTO – Secure, flexible and allows remote data monitoring.

3.7 Resources

1. Züllighoven, H. (2004). Object-oriented construction handbook: Developing application-oriented software with the tools & materials approach. Elsevier.
2. https://www.audit.gov.cn/WGBD/n1525/c10296921/part/10299823.pdf

3.8 Assessment

1. **Kobo collect software is best for collecting which type of data?**
   a. *Primary data*
   b. *Both qualitative and quantitative data*
   c. *Observation data*
   d. *Secondary data*
2. **What should be considered when selecting a software for data collection?**
   a. *The type of data to be collected*
   b. *The project scope*
   c. *The online and offline capability*
   d. *The type of analysis*
3. **Which software's are not suitable for data collection?**
   a. *Open Data Kit*
   b. *Kobo toolbox*
   c. *NVIVO*
   d. *Survey CTO*
4. **Which type of data does data collection software's' support?**
   a. *Surveys*
   b. *GPS and images*
   c. *Qualitative responses*
   d. *All of the above*
5. **Why is Kobo toolbox and ODK popular for data collection?**
   a. *They are offline friendly and supports multimedia*
   b. *They require internet access to collect data*
   c. *They can be used to collect medical data*
   d. *They are used to analyze qualitative data*

Section 4: Conducting Pilot Testing and Review

4.1 Introduction

Pilot testing involves running a small-scale test of the study's tools, methods, and procedures before full implementation. This section outlines how the pilot testing should be conducted and how to review and adjust to the initial plan.

4.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand how pilot testing is undertaken.
b. Ensure accuracy, completeness, and reliability of pilot data;
c. Help identify errors, inconsistencies or missing responses in the dataset;
d. Allow adjustments to survey design and data collection tools before full implementation.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 20 min

4.5 Materials: Felt pens, masking tape, flipchart, scissors, flipchart board, newsprints, black board, chalk, posters, computer, projector

4.6 Contents:

a. Hold meetings or events to practice questionnaires and data software
b. Convene with research team(s) to discuss results and adjustments to research plans
c. Discuss with researchers on initial data findings, and determine if adjustments are needed
d. Convene with partners to discuss changes

4.7 Training Content

4.7.1 Field Simulation & Data Collection

Field simulation and data collection are critical steps in ensuring the success of a research project. Field simulation involves mimicking real-world conditions to test data collection tools and processes before full implementation. This allows researchers to identify any logistical challenges, refine instruments and improve team readiness. Data collection, on the other hand, is the actual gathering of information from the target population, following the established methodology.

A one-day field level pretest is conducted after the training of the enumerators. This process serves to ensure that the data collection tools are fully understood by the enumerators as well as to test the enumerators' grasp of the training conducted and methods of conducting the interviews. The result of this exercise is then to be

evaluated by the trainers and suggestions reflected in the final version of the questionnaires.

4.7.2 Discuss with Team on Pilot Data Findings

Reviewing pilot data findings with the team is essential for refining the research process before full-scale data collection. This discussion helps identify inconsistencies, errors and challenges encountered in the field, ensuring necessary adjustments are made. By analyzing data accuracy, completeness and respondent understanding, the team can pinpoint areas needing improvement, such as questionnaire clarity, methodology or enumerator training.

Team members should also share their field experiences, highlighting any difficulties faced and proposing solutions. Based on these insights, the team can refine data collection tools, enhance procedures and address logistical issues. This collaborative review process strengthens data quality, ensuring a smoother and more reliable final data collection phase that aligns with research objectives.

4.7.3 Discussing with Partners on Changes

Engaging partners in discussions about proposed changes is essential for ensuring alignment and collaboration in the research process. These discussions provide an opportunity to review pilot data findings, address identified challenges, and collectively decide on necessary adjustments. By involving key stakeholders, the team can gain diverse perspectives, validate proposed modifications and ensure that any changes enhance data quality and research efficiency.

Effective partner discussions should focus on refining data collection tools, improving methodologies and addressing logistical concerns. Open communication and shared decision-making help build consensus and strengthen the research framework. By incorporating partner feedback, the research process becomes more robust, ensuring successful implementation and reliable results.

4.8 Resources

1. https://edpolicyinca.org/sites/default/files/2024-09/pilot-study-resource_acc.pdf
2. https://www.kcaa.or.ke/sites/default/files/manuals/CAA-M-OPS021%20DESIGNATED%20CHECK%20PILOT%20MANUAL.pdf

4.9 Assessment

1. **What is the main purpose of conducting pilot testing before the actual data collection?**
    a. *To finalize the research findings*
    b. *To test the effectiveness of the research tools and process*
    c. *To replace the need for full data collection*
    d. *To reduce the sample size of the study*

2. **Which of the following is not a key aspect evaluated during pilot testing?**

    a. *Clarity and relevance of the survey questions*
    b. *Reliability and validity of data collection tools*
    c. *The statistical analysis of the full study*
    d. *Logistics and feasibility of the study*

3. **How can researchers address issues identified during a pilot?**

    a. *Ignore minor issues and proceed with the study*
    b. *Modify research tools and methods based on the pilot test feedback*
    c. *Restart the entire study with new research design*
    d. *Reduce the number of respondents in the full study*

4. **Why is it important to review pilot test results before proceeding to full-scale data collection?**

    a. *To confirm whether the research objectives are still relevant*
    b. *To finalize the research report before full data collection*
    c. *To ensure errors and inconsistencies are identified and corrected*
    d. *To increase the complexity of data collection*

5. **Which of the following is not a reason for discussing pilot findings with partners?**

    a. *To provide an opportunity to review pilot data findings,*
    b. *To be able to address identified challenges, and collectively decide on necessary adjustments.*
    c. *To validate proposed modifications and ensure that any changes enhance data quality and research efficiency.*
    d. *To get a reason for more funding*

CHAPTER 7

Conducting Research

Section 1: Data Preparation Process

1.1 Introduction

This session will cover key aspects of data preparation. The participant will learn best practices to ensure data quality, eliminate errors, and optimize datasets for analysis. By mastering the data preparation process, researchers can generate credible insights and make informed decisions based on well-structured data.

1.2 Training Objectives:

By the end of the session, the participant should be able to:

a. Understand the Importance of Data Preparation;
b. Perform Data Cleaning;
c. Handle Missing Data.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 20 min

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents:

a. What data preparation is and implementing data planning

1.7 Training Content

1.7.1 Overview of Data Preparation Process

Data preparation is the process of cleaning, transforming and organizing data before analysis. It ensures that data is accurate, consistent and ready for meaningful use. This step is crucial in when data often comes from diverse sources such as surveys, interviews and observations.

1.7.2 Key Steps in Data Preparation

i. **Data Planning:** Before collecting data, it is important to create a structured plan (Refer to chapter 6) to ensure the information collected is relevant and useful. This will involve:

- Defining the research objectives i.e. what questions the data will answer

- Identify the sources of data one intends to use i.e. Primary or Secondary or mixed method.
- Choose the suitable data collection methods, for instance surveys, focus group discussions, interviews, etc.
- Determine data storage and security measures

ii. **Data Collection:** After planning one can go ahead and collect data using the method that best suits your objective.

iii. **Data Cleaning and Processing**: Raw data always has errors, missing values or inconsistencies that need to be corrected before analysis. In case of this, researchers will be required to remove duplicate entries, fill gaps or remove incomplete records, standardize formats and eliminate outliers or inconsistencies where necessary.

iv. **Data Transformation and Structuring:** Once data has been cleaned it will need restructuring to make it easier for analysis. This will involve converting qualitative data into coded categories, summarizing results by groups and merging multiple sources for a complete picture.

v. **Data Verification:** This involves cross checking the data for consistency and running preliminary analysis to detect anomalies. This ensures that the data is accurate before proceeding with the analysis.

vi. **Data Storage and Documentation:** Proper storage of data is necessary to preserve the data integrity. This involves securely storing data in databases, cloud or archives.

1.8 Resources

1. https://www.techtarget.com/searchbusinessanalytics/definition/data-preparation
2. https://firsteigen.com/blog/data-preparation/

1.9 Assessment

1. **Why is data planning important?**
   a. *It helps to create a structured plan to ensure the information collected is relevant and useful.*
   b. *It ensures that data is accurate, consistent and ready for meaningful use*
   c. *None of the above*
   d. *Both A and B above*
2. **What is data preparation?**
   a. *It's the process of collecting qualitative and quantitative data*
   b. *It's the process of writing the report*
   c. *It's the process of cleaning, transforming and organizing data before analysis*
   d. *It's the process of training research teams*
3. **What is the importance of data cleaning and processing?**
   a. *It helps to remove duplicate entries and remove incomplete records*

b. It helps to make the report small
c. It helps to identify codes for data
d. It helps to summarize study finds

4. Which of the following statements is false about data transformation and structuring?

a. It involves coding data
b. It involves summarizing data results by groups
c. It involves merging multiple sources of data
d. It involves collecting data

5. What does data planning involve?

a. Defining objectives, identifying resources, choosing data collection method and determining data storage measures.
b. Selecting the sampling framework to use
c. Cleaning and analyzing data
d. None of the above

Section 2: Data Preparation of Tools and Trainings

2.1 Introduction

This session will focus on the key aspects of preparing data collection tools and training field teams. The participant will explore best practices for designing and testing data collection instruments, ensuring they align with research objectives. By the end of this session, participants will understand how well-prepared tools and trained teams contribute to high-quality data and credible research outcomes.

2.2 Training Objectives:

By the end of this session, participants will be able to:

a. Preparation for data collection and training teams
b. Test and validate tools;
c. Implementing quality control measures.

2.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.4 Duration: 20 min

2.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents:

a. Preparation of tools for data collection;
b. Review of research team plans and coordination;
c. Training data collection teams;
d. Common data collection and management tools;
e. Software preparation.

2.7 Training Content

2.7.1 Preparation of Tools for Data Collection and Training of Teams

Effective data preparation is essential in ensuring high quality data is collected and intended research outcomes are achieved. This involves reviewing research team plans, coordinating logistics and training personnel before you proceed with data collection. In addition, if researchers will be using data collection software, it will be important to select and set up the software to test the efficiency and accuracy.

2.7.2 Reviewing Research Team Plans and Coordination

Before launching a data collection, ensure that the research team aligns in their plans and coordinate activities to ensure efficiency, consistency and effectiveness.

Key steps in Team Preparation

i. **Define roles and responsibilities:** Clearly assign roles to the team members i.e. field enumerators, supervisors and coordinators and ensure all the team members understand their specific tasks and chain of command.
ii. **Confirm and finalize the data collection plan:** Confirm the methodology you intend to adopt such as surveys, interviews, focus groups, etc. Set timelines for the data collection process and establish mechanisms to track progress.
iii. **Logistics and coordination:** Ensure that the field teams will have field access, means of transportation and accommodation are planned prior to the field day. Prepare backup plans in case of unforeseen challenges e.g. security issues, weather or communication.
iv. **Ethical compliance**: Inform the local administration about the research to be allowed to move around their locality. And also ensure that the teams are well trained on privacy, confidentiality and ethical research guidelines.

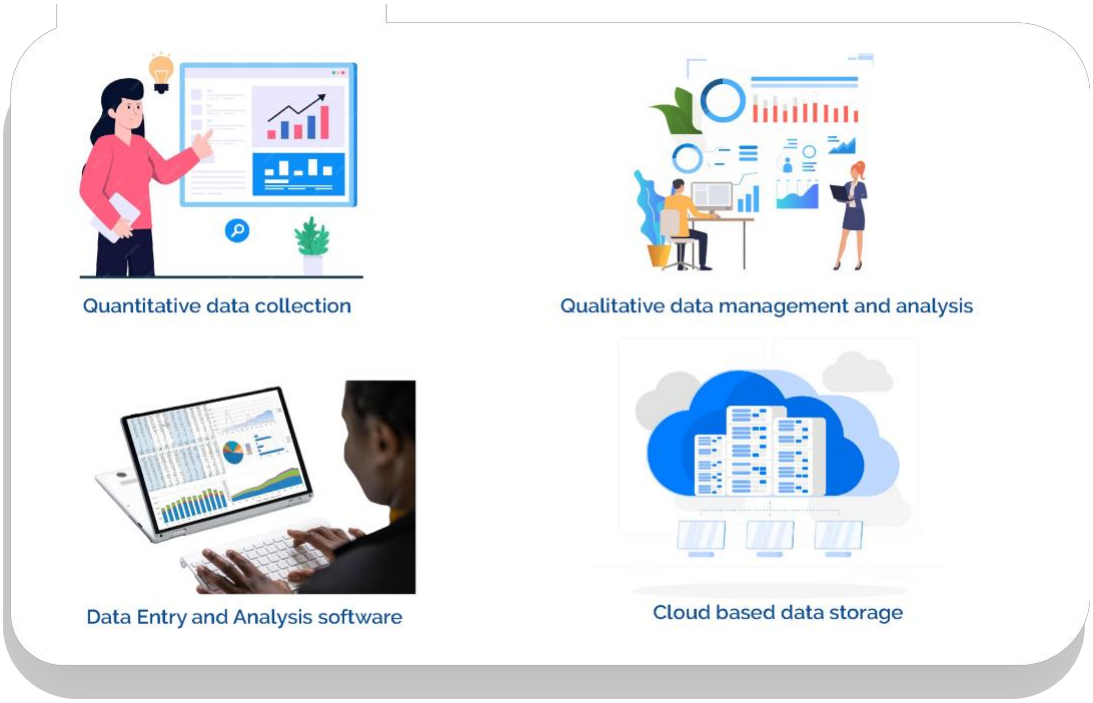
2.7.3 Training the Data Collection Team

To get high quality data, teams must be well-trained in data collection techniques, research ethics and use of mobile data collection applications. The training components should include:

- **Understanding the research objectives**: Ensure that the teams understand the research questions and expected outcomes and let them understand how data collected will be used for decision making.
- **Familiarization with data collection tools**: Train the enumerators on how to ask structured surveys, interview guides and conduct mock interviews and role play exercises for practice.
- **Standardizing data entry and recording methods**: Establish clear guidelines on how to record responses to ensure consistency during the data collection process. Emphasize on proper note taking techniques for qualitative data collection including writing all responses from all participants despite their accuracy.
- **Error reduction:** Teach the enumerators on how to avoid leading questions and ensure objectivity as they ask questions.
- **Ethics and community engagement:** Train the enumerators on how to obtain consent and maintain participant confidentiality. Teach the team on how to build trust with the local communities. This is essential as it encourages participant participation.
- **Field practice sessions:** Conduct pilot tests to troubleshoot any issues before the actual data collection exercise. Allow the team to practice using the mobile data collection application.

2.7.4 Common data collection and management tools

1. **Quantitative data collection:** Kobo Toolbox, ODK (Open Data Kit), Survey CTO, Google forms etc.
2. **Qualitative data management and analysis:** NVivo and Alesti
3. **Analysis software:** MS Excel, SPSS, STATA, Python
4. **Cloud based data storage:** Google drive, Dropbox and OneDrive



2.7.5 Software Preparation

1. Test all software before field deployment.
2. Ensure teams are trained on how to use the tools effectively.
3. Set up automatic backups to prevent data loss.
4. Establish data security measures such as passwords and restricted access.

2.8 Resources

1. https://www.intrac.org/app/uploads/2017/01/Basic-tools-for-data-collection.pdf
2. https://www.dataguard.com/blog/data-collection-process-tools-and-software/

2.9 Assessment

1. **Why is it important to train the data collection team?**
   a. *To ensure that they understand the research objectives*
   b. *To ensure that they have the right qualifications*
   c. *To ensure that they can write a good report*
   d. *To ensure that they understand the analysis tools to be used*
2. **What is the importance of defining roles and responsibilities of the research teams?**

a. *To ensure that the enumerators collect the right data*
b. *To ensure that the teams understand their specific tasks and chain of command*
c. *To ensure that the teams are well facilitated*
d. *To ensure that the data collection process is successful.*

3. **Which components should be included when training the data collection team?**
   a. *How to record responses, how to obtain ethics and how to avoid leading questions*
   b. *How to write a report and present research findings*
   c. *How to clean data and develop a data collection software*
   d. *How to prepare a research proposal*
4. **Which software is ideal for quantitative data collection?**
   a. *MS Excel*
   b. *Google drive*
   c. *Kobo Toolbox*
   d. *Nvivo*
5. **Why should the research team understand the research objectives?**
   a. *To ensure that they understand how to analyze data*
   b. *To ensure that they understand how to obtain informed consent*
   c. *To ensure that they understand how to troubleshoot issues.*
   d. *To ensure that they understand the research questions and expected outcomes*

Section 3: Implementing a Data Research Plan

3.1 Introduction

This session will cover key aspects of implementing a data research plan, including organizing field activities, training data collection teams, ensuring ethical compliance and monitoring progress. By the end of this session, they will be equipped with the knowledge and strategies needed to successfully implement a data research plan.

3.2 Training Objectives:

By the end of this session, the participant will be able to:

a. Understand the Key Components of a Research Plan;
b. Organize and Coordinate Field Activities.

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration 20 min

3.5 **Materials:** Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents:

c. Conducting data researching, monitoring progress and checkpoints, and data integrity

3.7 Training Content

3.7.1 Implementing a Data Research Plan

Implementing a data research plan involves putting the design research framework into action. Once the data collection framework is in place, the research team can begin to gather data. This will involve capturing real-time data through digital tools or manual methods.

Best practices for conducting data research

i. Ensure diversity in data collection i.e. geographic and demographic (age, gender, education level, etc.)
ii. Adopt the right sampling technique to avoid bias
iii. Collect both quantitative and qualitative data for a holistic perspective

3.7.2 Monitoring Progress and Checkpoints

Setting monitoring milestones and checkpoints is essential to keeping the research on track. Regular daily check-ins with data collectors and field teams allow for timely feedback on any challenges encountered. Real-time data dashboards enable continuous progress tracking as data is uploaded. Additionally, conducting random

field visits help ensure that data is collected accurately and from the appropriate respondents.

3.7.3 Ensuring Data Integrity

Maintaining integrity is crucial to ensuring that research findings are reliable and valid. This can be achieved through the following measures:

i. Accurately recording, categorizing and securely storing data.
ii. Addressing missing data through follow-ups or validation processes.
iii. Collecting data transparently to allow for future reference and verification.
iv. Using standardized data collection formats to maintain consistency.
v. Protecting data from unauthorized access or manipulation to safeguard its accuracy and credibility

3.7.4 Example of a research Plan

Activity	Work days	Support & Resources Required
Introduction • Project briefing and initial discussions with the Client	1day	TOR, EOI & Study Plan.

Desk Review of Project documents
• Preliminary review of project documents
• Development of Inception report, including
  ○ Draft tools
  ○ Development of data summary tables
• Submission of inception report
	5.5 days	Relevant project documents, documents and reports;
		Review and feedback

Inception Meeting with Client, revision of the report, logistical planning	0.5 day	
Sampling • Adoption of sampling frame and sample selection, • Initial review of interview guides and the data summary tables/matrices	2 days	Computers and Printing Services; Review and Input from client
Training • Preparing training program/materials • Training of research assistants • Pre-testing & review of Interview Guides	3 days	Transportation, training venue, flip charts, LCD, notebooks & pens, Training schedule and guides.
Field interviews (data collection)	10 days	Transportation, Accommodation and meals; IC Voice Recorders; Camera.
Data capture and analysis • Data capture and cleaning • Data coding and analysis	3 days	Computers, Data summary tables.
Preparation of the Report (draft & final) • Preparation of the draft report • Submission of draft report	5 days	Computer, 2 CDs, Overhead projector

Activity	Work days	Support & Resources Required
• Validation workshop • Finalization of the report • Submission of the final report		
Total number of days	30	

3.8 Resources

1. https://www.fhwa.dot.gov/ohim/handbook/chap2.pdf
2. https://epms.icmr.org.in/extramuralstaticweb/pdf/Adhoc/Adhoc_format_research_proposal_plan.pdf
3. https://dovetail.com/research/how-to-write-a-research-plan/

b. Assessment

1. **What does implementing a data research plan involve?**
   a. *Gathering data*
   b. *Forecasting*
   c. *Deleting data*
   d. *Publishing data*
2. **Why should a researcher conduct random field visits?**
   a. *To ensure data is collected from all the target population*
   b. *To ensure data is collected accurately and from the appropriate respondents*
   c. *To ensure that data is collected from all the government offices*
   d. *To ensure that data collected is well presented*
3. **Which of the following measures should be considered in order to achieve data integrity?**
   a. *Addressing missing data through follow-ups or validation processes.*
   b. *Using standardized data collection formats*
   c. *Collecting data transparently to allow for future reference and verification.*
   d. *All of the above*
4. **Which of the following statements is false about setting checkpoints during the data collection process?**
   a. *To enable continuous tracking of data as it is uploaded*
   b. *To allow timely feedback on any challenges encountered*
   c. *To ensure that the sampling technique used is appropriate during reporting*
   d. *To ensure that data collected is accurate*
5. **Which factors should be put in place when conducting research?**
   a. *Ensuring that you collect both qualitative and quantitative data*
   b. *Ensuring diversity in data collection*
   c. *Adopting the right sampling technique*
   d. *All of the above*

CHAPTER 8.

Data Transformation and Analysis

Section 1: Determining if Data Is Ready for Analysis and Interpretation

1.1 Introduction

This session will focus on key steps to determine whendata is ready for analysis and interpretation. The participant will explore essential quality checks, including verifying data completeness, consistency and accuracy. Additionally, the participant will learn methods for handling missing or inconsistent data, ensuring standardization, and preparing datasets for meaningful analysis.

1.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Assess data completeness;
b. Verify data accuracy and consistency.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 20 mins

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents

a. Completeness of the data for analysis

1.7 Training Contents

1.7.1 Completeness of the Data

After data has been collected, it is necessary to do a cross-check in order to determine if some values are missing or if the data contain outliers that might affect the output of data analysis. This therefore calls for the need of doing a check if the data is clean, complete, consistent, and well-organized and equally aligns with the research objectives.

Below is a snapshot of what is required in determining the readiness of data for analysis:

a. Perform data quality check.

Perform the following functions

1. **Check for missing values:** This requires to check if there exist missing values and to handle the missing data accordingly by either imputing or removing rows and columns with respect to the situation.
2. **Check for outliers**: This involves exploring data to see if there exist outliers which may be extreme or moderate. Extreme outliers refer to those outliers that are considered to extremely differ from other data points, which can disproportionately skew the findings. If not addressed, extreme outliers can impact on the statistical measures such as mean and standard deviation which can further result to wrong visualization during the analysis process. Additionally, the resultant correlation and regression analysis obtained can be misleading when outliers exist in a dataset.
3. **Check on the data type**: It is appropriate to ensure that the data types are correct. Data type can either be classified as numerical which contains values whereas categorical data are those that are organized into groups. For instance, gender where we have male and female, education level (primary, O-level, A level, etc.) among others.
4. **Check for duplicates**: This will require to delete any duplicate entries identified in a dataset.

b. Check on Data Structure

This simply refers to a system in which a dataset is organized, stored and managed to allow easy access and modification when necessary. It is therefore crucial to check and ensure that data is logically organized, and variable names should be clear.

Additionally, at this stage dataset should remove any formatting or coding inconsistencies present in the data.

1.8 Resources

1. https://jointlearningnetwork.org/wp-content/uploads/2022/01/B11-Data-Cleaning-Best-Practice-Guides-and-Tools.pdf
2. https://www.betterevaluation.org/sites/default/files/data_cleaning.pdf
3. https://www.acaps.org/fileadmin/user_upload/acaps_technical_brief_data_cleaning_april_2016_0.pdf

1.9 Assessment

1. **What factors should be considered when performing data quality checks?**
   a. *Accuracy*
   b. *Completeness*
   c. *Integrity*
   d. *All of the above*
2. **Why is it important to cross check data before analyzing?**

a. *It helps to determine if some values are missing*
b. *It helps to separate qualitative and quantitative data*
c. *It helps to code data*
d. *It helps to analyze data*

3. **What is data structure?**
   a. *It's a process of analyzing data*
   b. *It's the systematic way of organizing, storing and managing data*
   c. *It's a way of separating qualitative from quantitative data*
   d. *It the process of gathering data from secondary sources*
4. **Which of the following is not a factor to consider when determining the readiness of data for analysis?**
   a. *Checking the data structure*
   b. *Performing data quality checks*
   c. *Checking and testing the data collection instruments*
   d. *Checking for outliers*

Section 2: Data Cleaning and Transformation Techniques

2.1 Introduction

This session will cover key techniques for cleaning and transforming data, including identifying and correcting errors, handling missing values, standardizing formats and converting data into usable structures. By the end of this session, the participant will be equipped with basic skills to prepare clean, reliable data for analysis.

2.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand the importance of data cleaning;
b. Identify and correct errors;
c. Transform data for analysis;
d. Implement quality control measures.

2.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

2.4 Duration: 20 min

2.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

2.6 Contents

a. Data cleaning and transformation;
b. Methods of correcting errors;
c. Data transformation.

2.7 Training Contents

2.7.1 Data Cleaning and Transformation Techniques

Data cleaning and transformation techniques encompass a set of procedures designed to streamline the data analysis process. They are essential for preparing raw data by enhancing its accuracy, consistency and usability, ensuring that reliable insights can be derived.

Data cleaning is the process of identifying and rectifying errors in a database. Such errors include duplicate records or incorrect data entries, and afterwards implementing appropriate corrective measures. If not handled properly, these errors may lead to inconsistencies in the results.

Data cleaning is crucial for various reasons, including but not limited to:

a. Enhancing data quality to increase the confidence in obtaining accurate information from a database.

b. Ensuring data reliability to enable useful insights that support informed decision-making.
c. Facilitating early detection of errors within a dataset for improving efficiency in the analysis process and ensuring that data accurately reflects trends and gaps.

The following are some of the Data Cleaning Techniques

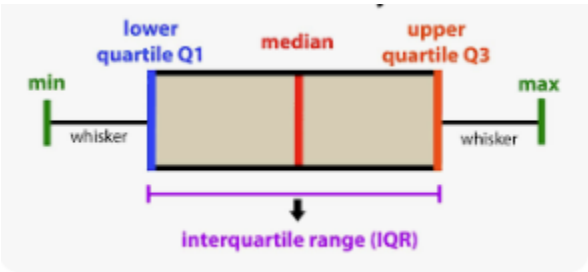
a) ***Identifying and Handling Missing Data****:* Explore the database and to ensure there are no missing values. In case of any, there are two approaches of handling them:

i. **Deletion approach:** Consider removing records with missing values especially if they are few and does not have any significant effect on the dataset
ii. **Imputation approach:** Replace missing values using methods such as mean, median, mode or regression-based predictions.

b) ***Identifying and Handling Outliers:*** Visualize the dataset by exploring by using statistical methods by either calculating the Z-scores or interquartile range. Visualization tools such as box plots can also be useful at this stage.

i. Box plot Visualization

By exploring the data, box plot provides an easier way of identifying outliers as shown below.



Correction: Remove outliers if they arise from errors, or consider transforming them using log transformation with the help of statistical analysis software.

2.7.2 Methods of Correcting Errors

Here are some of the methods for correcting the errors;

i. **Correcting Data Inconsistencies:** This involves checking on variable formats and standardizing them where necessary. For instance, date formats, currency symbols, etc. Additionally, ensure uniform spelling and categorization (e.g. "F" and "Female" should be consistent).

ii. **Removing Duplicates:** Identify rows or columns containing similar entries and delete duplicate records to prevent bias in analysis.

iii. **Validate Data Accuracy***:* This ensures that the data is not only accurate but also reliable. It is therefore recommended to conduct a thorough cross-examination of data entries against source documents to verify correctness. Additionally, implementing automated validation rules in data collection tools can significantly speed up the validation process.

2.7.3 Data Transformation

Data transformation involves converting raw data from one format to a more suitable format. This process can be achieved through the following methods:

i. **Normalization:** This involves transforming numerical figures into a standard range. Normalization ensures that data is organized into a structured and logical format. This is performed to remove unit differences, to ensure consistency and comparability. *For instance, using Dr. instead of Doctor or BSc. Instead of Bachelor's degree.*

ii. **Data Encoding:** This is the process of converting categorical or quantitative data into numerical form. This can either be done through binary encoding, one-hot encoding, label encoding or hash encoding. *For example, when using label encoding, in terms of different genders, Male can be coded as 1 and Female can be coded as 2 for uniformity.*

iii. **Aggregation and Summarization:** Data aggregation involves grouping data into meaningful categories to simplify analysis. *For example, if a dataset contains daily purchases made for seven days, calculating the average purchase for the week provides a summarized view of the data.* Aggregation can be performed using statistical measures such as mean, median and totals, helping to identify trends and patterns efficiently.

2.8 Resources

https://www.betterevaluation.org/sites/default/files/data_cleaning.pdf

2.9 Assessment

1. **Why is data cleaning important before conducting analysis?**
   a. *It enhances data quality*
   b. *Helps to detect errors*
   c. *It ensures data reliability*
   d. *All of the above*
2. **What is the importance of detecting errors within a data set?**
   a. *To improve data efficiency*
   b. *To eliminate the analysis process*
   c. *To ensure that the report is not long*
   d. *Both B and C above*
3. **Which of the following is not a method of correcting errors in a data set?**
   a. *Deleting duplicates*
   b. *Cross checking data entries*
   c. *Deleting the whole dataset*
   d. *Checking on variable formats*
4. **What is data aggregation?**
   a. *It's the process of converting quantitative data into numerical form*
   b. *Its grouping data into meaningful categories*
   c. *Its converting data into a common understandable format*
   d. *Its transforming numerical figures into a stand range*
5. **Which of the following is not a data transformation technique?**
   a. *Data encoding*
   b. *Data aggregation*
   c. *Data standardization*
   d. *Data deletion*

Section 3: Overview of Data Analysis

3.1 Introduction

This session will provide an overview of key data analysis concepts, including different types of analysis, common techniques and best practices for drawing valid conclusions. By the end of this session, the participant will be equipped with essential knowledge to analyze data effectively and make data-driven decisions.

3.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Understand the importance of data analysis;
b. Identify different types of data analysis;
c. Selecting appropriate analytical techniques.

3.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

3.4 Duration: 10 min

3.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

3.6 Contents

a. Data analysis
b. Difference between qualitative and quantitative
c. Qualitative coding and technology

3.7 Training Content

3.7.1 Data analysis

After data has been collected in a raw form, there is need to understand and obtain vital information from the collected data. thereby necessitating data analysis. Therefore, data analysis refers to a step-by-step process that involves collecting, cleaning and finally performing a series of analysis aiming at drawing insightful conclusions that can inform decision making.

Data analysis utilizes different techniques in order to draw meaningful conclusions from datasets, which can be used to inform future undertakings.

Quantitative and Qualitative data

- **Quantitative**

Quantitative data refer to datasets that can be numerically measured. They can either be in terms of quantities, amounts or even ranges that are expressed in terms of numbers. Quantitative data can include temperature readings e.g.

30°C, quantities of products purchased or sold, height e.g. 1.5 feet, and weight e.g. 90Kg among others.

- **Qualitative**

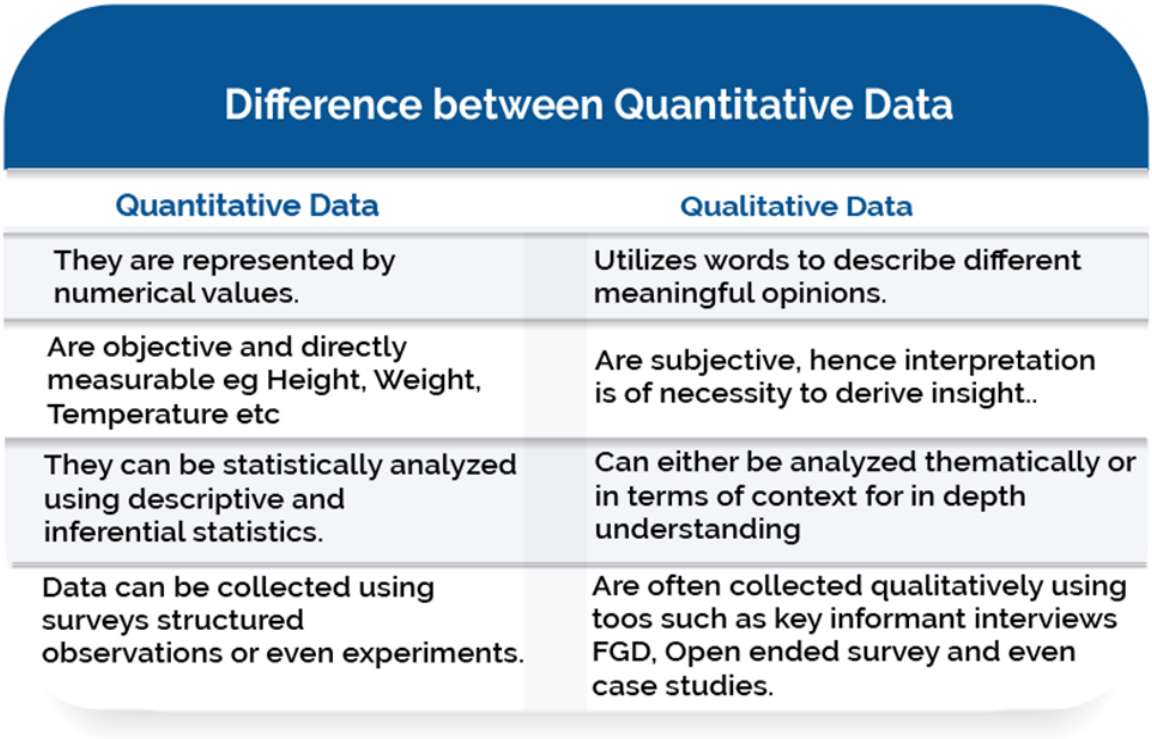
Qualitative data are those datasets that largely express opinions in a textual form. Therefore, qualitative data often gives detailed information that assists a researcher to understand experiences, perceptions or even social phenomenon that cannot be solemnly measured numerically. The information provided can be obtained through a myriad of methods including interviews, observations, Focus Group discussions, open ended surveys, narratives, transcriptions, desk reviews, case studies audio/video recordings etc. When collecting qualitative data, open ended questions are used to allow for description of characteristics, qualities or even attributes. For example, in a case where a researcher wants to get information regarding the impact of a project on household hygiene, a question such as follows;

"How effective were the hygiene promotion and awareness sessions in changing your household's hygiene practices?". This allows the respondents to give his/her opinion with respect to the stem of the question asked.

The response to the above question can be;

"The hygiene promotion awareness has really helped us to improve our household hygiene as we make sure we wash our hand regularly, we also ensure we eat clean food and also making sure that we house and animals are also clean"

Differences between the qualitative and quantitative data;



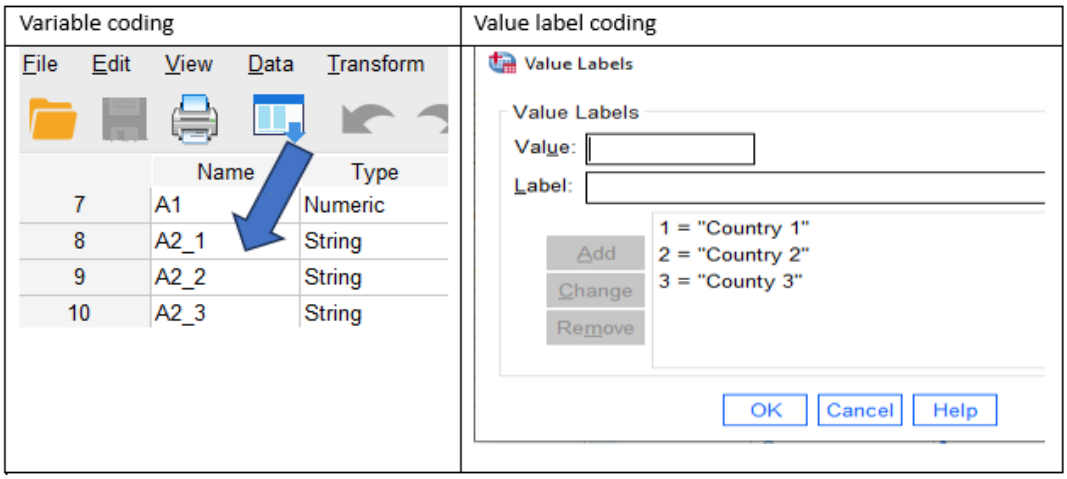
i. Data Coding

The first step towards quantitative data analysis is data coding. This involves systematically assigning numerical, categorical or alphabetical characteristics to responses in order to allow for analysis. Coding can be done for the variable name and variable label.

Variable coding- refers to the process where variables are assigned different characteristics. During variable labelling, numerical, alphabetical or a combination of both can be used to distinguish or identify variables.

Value label coding- This refers to the process of assigning descriptions to values that a variable can take. For instance, in assigning value labels to identify countries as shown in the figure below, country 1, country 2 and country 3 can be assigned a value of 1, 2 and 3 respectively. Another example includes assigning a value of 1 and 2 for Male and Female.

Figure XXX: Variable label coding and value label coding in SPSS



ii. Qualitative Coding

Qualitative data coding is a process whereby qualitative data from interview transcripts or open-ended responses are categorically analyzed by assigning different descriptive codes to record diverse recurring themes and patterns. This allows for effective documentation of crucial information from qualitative data.

Coding can be done in two ways:

a. **Inductive Coding:** This is where the responses guide the reader in coming up with appropriate themes when reading through qualitative data.

For instance, respondents were asked "in your opinion, what could be done to increase the participation of women, youth and PWDs to enable them reach their full potential"
Response:
Provide training and education to men, women, youth and people with disabilities; address barriers to participation such as norms and attitudes;

create opportunities for participation of the marginalized groups in the agricultural sector; raise awareness about the importance of inclusivity; focus on practical needs for marginalized communities; make agriculture profitable; develop outreach strategy that sets quotas for adult and young men and women, and PWDs in farmer groups.

Some of the codes that can be Identified from the response above are;

a. Capacity building (highlighted in yellow); Coded as 1
b. Opportunity Creation (highlighted in light grey); codes as 2
c. Awareness Creation (highlighted in dark grey); coded as 3
d. Outreach strategy (highlighted in dark green); Coded as 4

Theme that can be extracted from the codes above is "***Empowerment and Inclusion***"

b. **Deductive Coding:** This involves a process where predefined codes are applied during qualitative data analysis. Researchers are therefore required to find statements that are in support of the pre-existing theme.

In the question "How can the project further strengthen alignment with government systems to ensure sustainability"

Set of pre-defined themes can include

- Stakeholder collaboration (highlighted in grey)
- Operational M&E support system (highlighted in blue)

Response

Engage key stakeholders including the government agencies to create ownership; organize joint monitoring & evaluation programs; and initiate multistakeholder forums

iii. Categorization of Variables

Variables can be classified into two main categories depending on their nature and the kind of role they play in research:

a) ***Categorical Variables:*** Categorical variables are qualitative in nature and cannot be measured numerically; instead, they are grouped into distinct categories. *Examples include age, which can be categorized as below 18, 18–30 years, 30–60 years, and 60 years and above;* customer satisfaction levels, classified as highly satisfied, satisfied, somewhat satisfied, dissatisfied, and very dissatisfied; and clothing sizes, represented as small, medium, large, and extra-large.

There are two types of categorical variables;

1. **Nominal variables:** These are categorical variables that are not only distinctively labeled but also contain two or more categories that are not in ordered form as shown below.

In the question" How can the project further strengthen alignment with government system to ensure sustainability".	
Nominal Variable	**Categories**
Gender.	Male or Female.
Blood type.	A,B,AB,O.
Nationality.	Rwandese, Sudanese etc.
Favorite color.	Blue, Green, Purple or Red.

2. **Ordinal variables:** Refers to categorical variables that can either ranked or ordered in a particular way.

Categorical Variables	Categories
Level of Satisfaction	Very Dissatisfied, Dissatisfied, Somehow Satisfied, Satisfied, Very Satisfied
Economic Status	Low, Middle, High
Education Level	No Education, Primary, Secondary, Tertiary

b) ***Numeric***: These variables that exhibit numerical characteristics and can serve as proxies for describing measurable quantities. Numerical variables can be classified into two:

Discrete: These are quantitative variables that assume a specific value which can be clearly interpreted.

Continuous: These are quantitative variables that can take an infinite number of values within a defined range. For example, measuring the time taken to travel between two locations 500 km apart can yield varying values, such as 5.25 hours, 5.5 hours, or 5.75 hours.

3.8 Resources

https://www.betterevaluation.org/sites/default/files/2024-02/Introduction%20to%20Qualitative%20Research%20Methodology%20-%20A%20Training%20Manual.pdf

https://www.stat.cmu.edu/~hseltman/309/Book/chapter2.pdf

https://www.betterevaluation.org/sites/default/files/EA_PM%2526E_toolkit_module_5_QDA_for_publication.pdf

3.9 Assessment

1. **Which of the following best describes data analysis?**

a. It's the process of collecting raw data
b. It's the process of storing data securely
c. It's the process of designing survey questions
d. It's the process of inspecting, and transforming data to useful information

2. **Which of the following is a key difference between qualitative and quantitative data analysis**

a. Qualitative analysis focuses on numerical data while quantitative analysis focuses on textual data
b. Qualitative data involves identifying patterns and themes while quantitative data relies on statistical techniques
c. Qualitative analysis is subjective while qualitative data is objective
d. Quantitative analysis does not require data collection, while qualitative analysis does

3. **Which of the following is not an example of nominal variable**

a. Color
b. Gender
c. Economic status
d. Blood type

4. **What is ordinal variable**

a. They are variables that is ranked
b. They are variables that contain two or more categories
c. They are variables that appear several times
d. They are variables that cannot be measured

5. **Which of the following statements best describe nominal variable?**

a. *They are variables that exhibit numerical characteristics and can serve as proxies*
b. *They are variables that are not only distinctively labeled but also contains two or more categories*
c. *They are variables that can either be ranked or ordered in a particular way*
d. *They are quantitative variables that can take an infinite number of values*

Section 4: Conducting Data Analysis

4.1 Introduction

This session will cover the key steps in data analysis, including data preparation, selection of analytical methods, and interpretation of results. The participant will learn how to apply statistical and qualitative techniques to extract valuable insights while ensuring accuracy and reliability.

4.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Conduct data analysis.

4.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

4.4 Duration: 20 min

4.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

4.6 Contents

a. Overview of quantitative and qualitative data analysis.

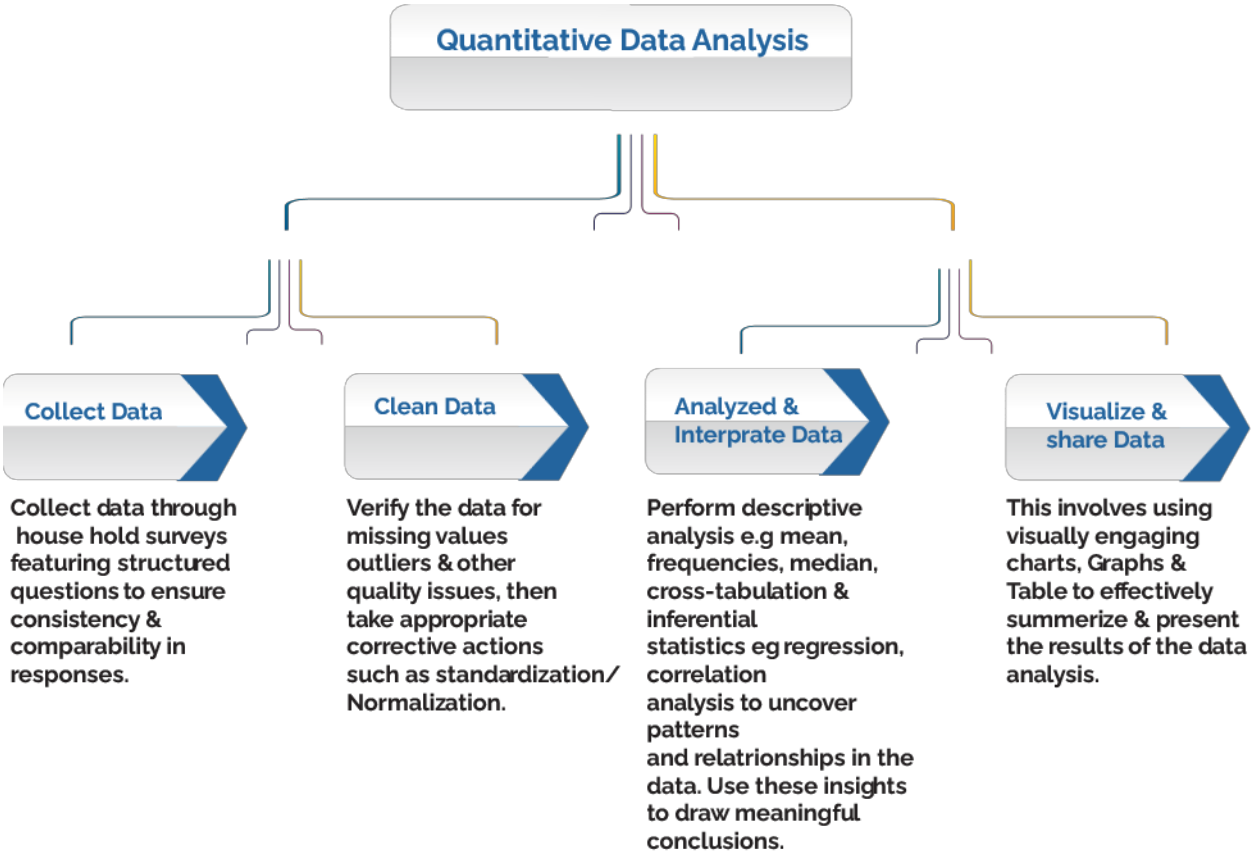
4.7 Training Content

4.7.1 Overview of quantitative and qualitative data analysis

When conducting research or testing a hypothesis, data must be collected objectively. It then undergoes a systematic process that includes cleaning, transformation, description, modeling and interpretation. This comprehensive process defines data analysis, which primarily involves analyzing and interpreting numerical data using statistical methods.

To extract meaningful insights from raw data, conducting data analysis is essential. However, quantitative and qualitative data require different analytical approaches to identify trends, patterns or relationships within a dataset.

With **quantitative data analysis** the following outline highlights the key steps for conducting an effective quantitative data analysis.



Note: These steps have been described in section 2 and 3 above.

4.8 Resources

1. https://www.fao.org/fileadmin/user_upload/food-security-capacity-building/docs/Seeds/Training_Material/DART/06_-_FAO-SSA-DART_-_Qualitative_Data_Analysis.pdf
2. https://rhntc.org/sites/default/files/resources/opa_qual_analysis_2020-07.pdf
3. https://www.ufs.ac.za/docs/librariesprovider68/resources/methodology/uwe_flick_(ed-)-_the_sage_handbook_of_qualitative(z-lib-org)-(1).pdf?sfvrsn=db96820_2

4.9 Assessment

1. **Which of the following is not a step in conducting quantitative data analysis**
   a. *Collecting data and cleaning data*
   b. *Analyzing and interpreting data*
   c. *Data Visualization*
   d. *Data management*
2. **What does data visualization involve?**
   a. *It involves verifying missing values, outliers and other quality issues*
   b. *It involves using charts, graphs and tables to effectively to summarize and present findings*
   c. *It involves collecting data through household surveys*
   d. *It involves performing descriptive analysis*
3. **Which of the following does not involve quantitative data analysis?**
   a. *Data cleaning*
   b. *Data management*
   c. *Descriptive analysis*
   d. *Thematic analysis*

CHAPTER 9.

Communicate Data Results for Decision Making

Section 1: Making Sense of the Data and Making Recommendations

1.1 Introduction

This topic introduces the participant on how to make sense of data gathered, tips to synthesize result findings and the best way to develop recommendations as well as how to communicate the findings to the stakeholders and making use of their feedback.

1.2 Training Objectives:

By the end of this session, the participant should be able to:

a. Synthesize results findings;
b. Develop recommendations from research findings;
c. Communicate findings;
d. Use the feedback to refine the report.

1.3 Methodology: Group discussions, lectures, brainstorming, exercises, case studies, stories, audio-visual aids, etc.

1.4 Duration: 30 mins

1.5 Materials: Felt pens, masking tape, flipchart, flipchart board, black board, chalk, posters, newsprints, computer, projector.

1.6 Contents

a. Overview of making sense of data and making recommendations;
b. Synthesizing data;
c. Developing data driven recommendations.

1.7 Training Content

1.7.1 Making Sense of Data and Making Recommendations

Once data has been collected and analyzed, the next critical step is to synthesize results and translate findings into actionable recommendations. This process ensures that decision making is informed by evidence.

1.7.2 Synthesizing data

Synthesizing results involves combining and analyzing data from various sources as highlighted below:

- Review findings and identify the key points
- Relate findings to the initial key questions or hypothesis
- Identify key patterns and trends

- Contextualize data and considering external factors like social-economic, cultural, etc.
- Highlight key evidence and using graphs, charts, etc. to make findings clear.

Example: If data shows small scale farmers who have adopted ICT extension services have a higher productivity, this suggests that digital tools play a crucial role in improving farming.

1.7.3 Developing Data Driven Recommendations

Once key findings have been developed, the next step is to generate practical recommendations. The following are key areas to consider:

- Address identified gaps: the recommendations should solve problems highlighted in the findings
- Rank recommendations based on the urgency and impact
- Align recommendations with stakeholder needs and priorities of the decision maker and the community
- Ensure that the recommendations are SMART (Specific, Measurable, Achievable, Relevant, and Time-bound).
- Also consider stakeholder buy-ins and availability of resources

Example: If findings indicate that women in the rural walk long distances to access water, recommendation should be:

"*Develop strategies to improve water accessibility in order to reduce the amount of time women spend looking for water*"

3.7.1 Defining and Planning Communication Goals

Effective communication of research findings ensures that insights are actionable, relevant and impactful for the organization and its partners. The process involves determining how, where and why the research is being shared, identifying the effective communication strategies, and setting clear goals for dissemination.

3.7.2 Determine How, Where and Why the Research Findings are Being Shared

Before sharing the research findings, it is essential to consider the following:

i. **Purpose of sharing:** The purpose of sharing the findings could be to inform decision making, influence policy and practices, raise awareness or build partnerships. This will guide you on the stakeholders to be involved.
ii. **Channels of communication:** There are different channels of communicating research findings including reports, policy briefs and executive summaries; presentations, workshops and conferences; social media, newsletters and blogs; and community engagement sessions.

iii. **Benefits of sharing:** Sharing research findings enhances credibility, visibility of the findings, encourages data driven decision-making and strengthens collaborations and supports advocacy and policy change.

3.7.3 Determining the Most Effective Communication Strategies

Choosing the right communication strategies ensures that research findings reach the intended audience in an effective way. The strategies include:

i. Use of digital platforms to expand reach e.g. zoom, and webinar
ii. Ensuring accessibility by using clear language and translations where necessary
iii. Use of case studies to make data relatable and actionable
iv. Tailoring messages to different audiences e.g. policy briefs for policy makers

3.7.5 Creating a Set of Communication Goals

It is important to have SMART (specific, Measurable, Actionable, Relevant and Time-bound) goals. Examples of communication goals include:

i. Present findings to at least one government official agency to help with influencing policy
ii. Organize stakeholder engagements to discuss findings
iii. Publish research findings to increase awareness
iv. Track metrics such as report downloads, and social media shares on the findings

4.7.1 Developing Presentation Material

To effectively communicate research findings and promote key findings, it is essential to develop clear, engaging and well-structured presentation materials. The key steps include:

a. Choose the Right Presentation Format

This could include:

1. Slides such as PowerPoint slides or Google Slides. They are good for structured content with visuals.
2. Handouts or reports which are ideal for detailed or technical information.
3. Videos and animation are good when presenting case studies.

4.7.2 Qualities of Effective Slides

- Keep it simple and use bullet points (five-words across and five-points downwards)
- Graphs and images used that are clear and relevant
- Avoid excessive colors
- Highlight/bold key points
- Minimize the use of animations and keep transition smooth

5.7.1 Awareness Building and Communication of Data

Awareness building is a crucial step in ensuring research findings reach the right audience. The purpose of awareness building varies depending on the stakeholder involved and desired impact.

a. **To inform stakeholders**: awareness building will ensure that relevant stakeholders have access to accurate, timely and relevant information
b. **Drive community action and advocacy:** raising awareness among grassroot communities will ensure that they understand their rights, and have opportunities to demand for accountability and change
c. **Improve program implementation and service delivery:** awareness creation can strengthen collaboration between implementing agencies and beneficiaries which is crucial in enhancing program efficiency and impact

5.7.3 Choosing the Right Communication Channel

Choosing the appropriate platform to share research findings is crucial in ensuring that the audience can easily access and engage with the information. Disseminating the information through the right channel will help that the information reaches the intended audience. Some of the channels include:

- Community engagements through workshops and stakeholder forums
- Digital and social media through using social media and posting the findings on websites
- Traditional media such as publishing the findings in newspapers, press releases and having radio talks or TV shows
- Use of institutional reports and policy briefs to formally document the findings

Methods of documenting audience reactions

Documenting audience reaction is crucial for evaluating the effectiveness of communication efforts, understanding stakeholder perspectives and refining future perspective. Some of the key approaches include:

i. **Observation and note-taking**: record all feedback provided and observe audience engagement
ii. **Audio recording**: obtain the participants consent to record the presentations and review the recording to analyze critical feedback
iii. **Focus group discussions**: organize participants into small groups to share their thoughts on clarity and relevance of the findings
iv. **Social media and digital engagements:** monitor and track shares, like, retweets, comments and reactions to posts related to the presentation
v. **Surveys and feedback forms**: distribute surveys immediately after presentations to gauge participants interests and concerns.

6.7.3 How to Use Feedback to Improve Findings

Once you have collected all the feedback,

- Identify key themes and analyze common feedback points to refine your reporting
- Adapt the content to address misunderstandings
- Share the refined report/findings to inform the next step

Examples of using feedback to improve findings:

- *Addressing misunderstanding: Clarifying ambiguous survey results by providing additional context from the qualitative interviews and stakeholder feedback*
- *Altering findings inappropriately: Removing certain responses because they do not support the desired conclusion.*

1.8 Resources

1. https://www.distillersr.com/resources/systematic-literature-reviews/understanding-the-synthesis-of-results-in-a-systematic-review

1.9 Assessment

1. **What are the key areas to consider when developing recommendations**
   a. *They should be SMART*
   b. *You should consider stakeholder buy-ins*
   c. *Align the recommendations with stakeholder needs and priorities*
   d. *All of the above*
2. **Which of the following is not a key factor to consider when sharing research findings?**
   a. *Understand how much the stakeholders will be paid*
   b. *Channels of communication*
   c. *Benefits of sharing the findings*
   d. *Understand the purpose of sharing the findings*
3. **What is the primary objective of preparing presentation materials for presenting research findings?**
   a. *To make the presentation look more professional*
   b. *To ensure clear, effective communication of key insights to the intended audience*
   c. *To impress stakeholders*
   d. *To replace the need for verbal explanations*
4. **How can documented audience reactions improve the study findings?**
   a. *By making presentations more complex*
   b. *By ensuring that the audience reacts the same way*
   c. *By discouraging audience participation*
   d. *By helping to refine content, address common concerns and improve the report*
5. **Which of the following channels would you consider when communicating findings to the community members?**
   a. *Policy briefs*
   b. *Impact reports*
   c. *Community meetings*
   d. *Journals*

ANNEXTURE

Annex I: Glossary of Terms

Conventional Data Collection method	Modern methods used to gather information from various sources.
Data	Raw facts, figures or information collected for analysis, reference or decision making
Digital data or 'Big data'	Is the information that is stored, processed and transmitted in digital form
Gender lens	An analytical framework that considers gender when making decisions.
Grassroots data model	Is a structured approach to collecting, managing, and utilizing data at the local or community level to inform decision-making, improve service delivery, and empower communities
Hypothesis	An assumption that is made based on some piece of evidence.
Participatory Approach	an approach that actively involves community members, stakeholders, and beneficiaries in the research process
Primary data	data that is collected directly from a source
Research	a systematic process of collecting, analyzing and interpreting data to generate new knowledge, validate existing theories or solve a specific problem
Sampling	A process of using a subset of a population to represent the whole population
Stakeholders	individuals, groups or organizations that have interest in or affected by a project or initiative
Variable	Any characterized number or quantity that can be measured or counted and takes different values within a dataset

Annex II: Proposed timeline for the course

1) Duration of the online experience (Weeks)

- Total content time: 39 subsections × 20 mins (including assessment) = 13 hours.
- Distribution across 9 chapters: Assuming roughly equal distribution, each chapter has about 5 subsections (i.e 5*20 = 1 hour 30 minutes of content per chapter).

- Optimal weekly learning load: Online learners often manage 3-5 hours per week comfortably.

Two possible course structures

- Intensive (5 Weeks):
    - ~5 hours per week
    - Covers 2 chapters per week
    - Suitable for highly engaged learners

- Standard (8-10 Weeks):
    - ~2.5-3.5 hours per week
    - Covers about 1 chapter per week
    - More digestible pacing for participants

2) Estimated Weekly Time Commitment

- Intensive Format: 5 hours/week
- Standard Format: 2.5-3.5 hours/week

Annex III: Responses to the Assessment Questions

CHAPTER 1

Question No.	Sections			
	Section 1	**Section 2**	**Section 3**	**Section 4**
Qs. 1	B	D	C	A
Qs. 2	D	C	A	B
Qs. 3	A	B	D	C
Qs. 4	C	A	C	A
Qs. 5	D	D	B	C

CHAPTER 2

Question No.	Sections							
	Section 1	**Section 2**	**Section 3**	**Section 4**	**Section 5**	**Section 6**	**Section 7**	**Section 8**
Qs. 1	D	C	A	A	A	D	B	D
Qs. 2	A	C	D	D	B	B	D	A
Qs. 3	B	A	C	C	D	B	B	B
Qs. 4	D	D	A	C	A	C	A	A
Qs. 5	B	B	D	B	A	A	B	C

CHAPTER 3

Question No.	Sections			
	Section 1	**Section 2**	**Section 3**	**Section 4**
Qs. 1	C	A	A	D
Qs. 2	A	D	D	C
Qs. 3	B	C	C	B
Qs. 4	D	B	B	A
Qs. 5	C	D	C	D

CHAPTER 4

Question No.	Sections			
	Section 1	Section 2	Section 3	Section 4
Qs. 1	A	A	D	A
Qs. 2	D	C	A	B
Qs. 3	C	D	C	A
Qs. 4	B	A	A	C
Qs. 5	D	B	C	D

CHAPTER 5

Question No.	Sections						
	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Qs. 1	A	B	B	A	D	A	A
Qs. 2	D	C	B	B	A	B	A
Qs. 3	A	A	C	A	D	A	B
Qs. 4	C	A		D	C	B	C
Qs. 5	C	C		C	B	D	D

CHAPTER 6

Question No.	Sections			
	Section 1	Section 2	Section 3	Section 4
Qs. 1	D	D	B	B
Qs. 2	B	A	C	C
Qs. 3	A	B	C	B
Qs. 4		D	D	C
Qs. 5		C	A	D

CHAPTER 7

Question No.	Sections		
	Section 1	Section 2	Section 3
Qs. 1	D	A	A
Qs. 2	C	B	B
Qs. 3	A	A	D
Qs. 4	D	C	C
Qs. 5	A	D	D

CHAPTER 8

Question No.	Sections			
	Section 1	Section 2	Section 3	Section 4
Qs. 1	D	D	D	A
Qs. 2	A	A	B	B
Qs. 3	B	C	C	D
Qs. 4	C	B	A	
Qs. 5		D	B	

CHAPTER 9

Question No.	Sections
	Section 1
Qs. 1	D
Qs. 2	A
Qs. 3	B
Qs. 4	D
Qs. 5	C