A Practitioners' Manual on Monitoring and Evaluation of Development Projects

Kultar Singh, Dharmendra Chandurkar and Varun Dutt

A Practitioners' Manual on Monitoring and Evaluation of Development Projects

A Practitioners' Manual on Monitoring and Evaluation of Development Projects

By

Kultar Singh,

Dharmendra Chandurkar

and Varun Dutt

Cambridge
Scholars
Publishing

A Practitioners' Manual on Monitoring and Evaluation of Development Projects

By Kultar Singh, Dharmendra Chandurkar and Varun Dutt

This book first published 2017

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-8992-X
ISBN (13): 978-1-4438-8992-6

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

AHS	Annual Health Survey
AI	Appreciative Inquiry
AIP	Annual Implementation Plan
ANC	Antenatal Care
APL	Above Poverty Line
ASCII	American Standard Code for Information Interexchange
BPL	Below Poverty Line
BPY	*Beti Padhao Yojana*
CAPI	Computer-Assisted Personal Interviewing
CASI	Computer-Assisted Survey Interviewing
CATI	Computer-Assisted Telephone Interviewing
CAWI	Computer-Assisted Web Interviewing
CBA	Cost Benefit Analysis
CCDB	Christian Commission for Development in Bangladesh
CEA	Cost Effectiveness Analysis
CHC	Community Health Centre
CSPro	Census and Survey Processing System
CSR	Corporate Social Responsibility
DID	Difference in Differences
FE	Fixed Effect
FGD	Focused Group Discussion
FLW	Frontline Health Worker
GAD	Gender and Development
Hb	Haemoglobin
HH	Household
HU	Hermeneutic Unit
IDI	In-depth Interview
IE	Inclusive Education
IFA	Iron Folic Acid
IGA	Income Generating Activities

IHHL	Individual Household Latrines
IOCE	International Organisation for Cooperation in Evaluation
IV	Instrumental Variable
JSY	*Janani Suraksha Yojana*
LFA	Logical Framework Approach
M&E	Monitoring and Evaluation
MDG	Millennium Development Goal
MEL	Monitoring, Evaluation and Learning
MIS	Management Information System
MLE	Monitoring Learning and Evaluation
MMR	Maternal Mortality Rate
MSC	Most Significant Change
NGO	Non-Governmental Organisation
ODBC	Open Database Connectivity
ODF	Open Defecation Free
PAPI	Paper- Assisted Personal Interviewing
PAS	Project Annual Survey
PDA	Personal Digital Assistant
PHC	Primary Health Centre
PMF	Performance Measurement Framework
PMP	Performance Monitoring Plan
PPM	Project Planning Matrix
PSM	Propensity Score Matching
RCT	Randomised Controlled Trial
RD	Regression Discontinuity
RMNCH	Reproductive, Maternal, Newborn and Child Health
RWE	Real World Evaluation
SBA	Skilled Birth Attendants
SC	Significant Change
SPSS	Statistical Package for Social Science
TE	Transformative Evaluation
ToC	Theory of Change
UFE	Utilisation Focused Evaluation

VHC	Village Health Committee
WHO	World Health Organization
ZOPP	Zielorientierte Projektplanung
GOPP	Goal Oriented Project Planning

PREFACE

This book is an effort to provide handholding support to professionals working in areas of monitoring and evaluation (M&E), programme management or other related domains in the development sector.

A Practitioner's Manual on Monitoring and Evaluation of Development Projects attempts to take development professionals through the life cycle of the M&E process and explain its fundamentals and concepts. Practical examples have been used throughout the manual to assist professionals to visualise and internalise the key concepts and processes that need to be followed for monitoring and evaluating development interventions.

Chapter 1 of the manual aims at developing a fundamental understanding of projects which is a prerequisite for understanding how to monitor and evaluate the same. Next, the steps of the project cycle are explained sequentially, after which the critical concepts like project objective and goal and the difference between these terms are clarified.

Chapter 2, which focuses on understanding how to design projects, begins with a discussion on the results chain or the series of end to end causal pathways that need to be formulated to understand how the short- term and the long-term project results are to be achieved. This is followed by enumerating the key tools that are used in designing projects, after which the reader is taken step by step through the logical framework analysis, a popular tool used in designing projects, illustrated with an example.

Chapter 3 explains the key concepts, the fundamentals of M&E and also the key differences between them. The key typologies and different types of M&E are explained, in addition to the critical concept of indicators and how to design good indicators.

Chapter 4 takes the readers through the process of developing a Performance Monitoring Plan (PMP) which acts as an overarching framework for monitoring and evaluating a development project. PMP is helpful in monitoring the performance of the project and the achievement of its intended results. Subsequently, the key activities that should be done to ensure the data quality are explained.

Chapter 5 describes different approaches to monitoring including progress, process and participatory monitoring. It also explains the most significant change (MSC) tool used to monitor projects without using indicators.

Chapter 6 delineates in detail the various evaluation designs by explaining the fundamental concepts of evaluation as well as elucidating the different types of evaluation designs by using practical examples. Randomised Controlled Trial (RCT), Cluster RCT and Pipeline design are explained under experimental designs, followed by different designs like Difference in Differences (DID), Propensity Score Matching (PSM), Regression Discontinuity (RD), Instrumental Variable (IV) and Reflexive Comparison, all of which fall under the category of Quasi-Experimental designs. The non-experimental evaluation designs are explained next, followed by different types of impact evaluation designs compared through a matrix to understand when any of these evaluation designs should be used and the advantages and disadvantages of using the same. Chapter 6 ends with an exercise which has different situations in which the practitioner is required to choose an evaluation design which will be suitable for each situation.

Chapter 7 explains the fundamental concepts of the different approaches which can be adopted for evaluating projects based on the context and the objective of a developmental project. It should be remembered that these approaches are at the macro level philosophy of evaluations and they will use the evaluation designs explained above as part of their methodology.

Chapter 8 delineates data basics and how the data may be analysed usefully for monitoring and evaluating a project. Univariate and multivariate analysis are covered as part of the data analysis. Basic quantitative data analysis using MS Excel and SPSS software has been demonstrated using screenshots. The fundamentals of qualitative data analysis and the popular software that can be used to conduct the same have also been touched upon.

In Chapter 9, practitioners are taken through a step by step process of executing the Monitoring Learning and Evaluation (MLE) function. This chapter focuses on the operational dimension of executing the MLE process. Concepts discussed in the earlier chapters are summarised through an operational lens. Steps like pilot testing, developing data collection and quality assurance protocols, getting ethical approvals, developing a learning framework etc., are covered in this chapter.

To summarise, this book will act as a hands-on manual for professionals working in the development sector to understand the fundamentals of M&E from the point of view of a practitioner. Attempt has been made to sequence the book in such a way that it covers all the key concepts related to M&E in a simple, lucid and coherent manner. However, it is quite possible that despite our best efforts, certain shortcomings may have remained and we look forward to removing these in the future versions. We will be grateful to our readers for their valuable suggestions and feedback.

CHAPTER ONE

UNDERSTANDING PROJECTS

The mainstay of work in the development sector is in the form of projects, which are targeted at various areas like health, education, sanitation, livelihoods, child rights, climate change etc., depending on the mandate and the objective of project implementation and funding organisation. In this chapter, the fundamental concepts related to projects are explored and a project-based approach is discussed, which is imperative for an in depth understanding of M&E.

On completion of the chapter the reader will be able to:

- ✍ Describe what is a project
- ✍ Understand and explain the elements of a project cycle
- ✍ Define the project objective and goal

1.1 Defining a Project

As we embark upon our journey to understand how to monitor and evaluate projects, it is important to first understand the fundamentals and underlying concepts of projects and project management. At the outset of every project, it is envisaged that several activities will be performed over the course of the project's implementation. These activities constitute the work that will be done during the project and they form the mainstay of the action that will take place.

Every project has a *specific objective* and it is envisioned that through these activities the project will achieve its objective. The example of a five-year project of making its target villages open defecation free (ODF) is used to illustrate this point. To achieve this objective, the project engages in several activities like construction of household (HH) and community toilets, conducting awareness campaigns to motivate people not to defecate in the open, educating people about the technologies that should be used for toilet construction etc. The types of activities performed as part of the project vary depending on the project objective and the implementing organisation's capacity. These activities form the key work that is done as part of the project implementation.

Another important aspect of every project is that it has a specific start date and a specific end date i.e., *a specific time period* within which it has to be executed. The project is expected to achieve its desired objective within this specific time period, which in the example quoted above, is a duration of five years. Last but not the least, it is very critical to understand that each project is allocated a *limited set of resources*. Resources, which may be financial, human and physical, are allocated to a project so that its activities may be implemented and its objectives achieved within a specific time period. Accordingly, the example project is also allotted a fixed budget, human resources and fixed physical resources with which its activities may be implemented and its objective achieved in a specific period of time.

Hence, a project may be defined as:

> ***"A set of activities implemented within a specific period of time and with specific resources to achieve a specific objective."***

1.2 Project Cycle

Usually, social development interventions are formulated and implemented in the form of a project and follow a cycle or a sequence which is known as the *project cycle*. From its inception to its closure, every project has its unique cycle of operation though the fundamental project cycle remains the same. Therefore, it is essential to understand the project cycle in order to better conceptualise, design, plan and implement it and also to monitor and evaluate it effectively.

From the beginning till the end of the project, the project cycle comprises of various phases or stages. All the stages in the project cycle are delineated and implemented successively in a phased manner. Each of these stages are defined by their objective, information requirements, responsibilities and key outputs. The various stages of a generic project cycle are:

Figure 1.1: Project Cycle

Stage I: Situation Analysis

It is a well-known fact that a project does not exist in vacuum. It is formulated to respond to a negative situation or condition. The stage of identifying and understanding this existing negative situation which needs to be responded to through intervention is called *situation analysis*. This stage consists of understanding the prevalent situation and identifying the cause(s) of this situation. Situation analysis is useful in the later stages when the strategy and subsequently, the specific activities to target these causes are defined. A good situation analysis serves as an entry point for the project by throwing light on what needs to be done to address the negative conditions in each context. For example, in the sanitation project referred to earlier, the situation of open defecation (OD) and toilet usage in the target area are assessed, apart from identifying the key reasons why people are defecating in the open.

Stage II: Gap Analysis

A project always works towards achieving its desired objective. By the time it is completed, the project envisages reaching the intended or desired situation as opposed to the situation from where it had started. The project works towards bridging the "gap" that exists between the present and the desired situation. *Gap analysis* is thus done to identify the gap between the current situation and the desired situation. In the example, the project aims to make the target area ODF from the current situation, where it is assumed that 35 per cent of the target population defecates in the open. If in terms of actual numbers, 35 per cent works out to one lakh families, the gap is 35 per cent or one lakh families. Therefore, the intensity of project activities, the requirement of resources and the implementation plan are made keeping in mind that the target population the project must reach is one lakh families.

Stage III: Project Planning

Project planning follows the gap that needs to be bridged through the project that has been identified. During the *project planning* stage, objectives are defined, strategies by which to achieve this objective are formulated, activities are identified, timeline based targets are set and resources are allocated to the project. A detailed implementation plan with the activity schedule and milestone timelines is also prepared as part of the project planning. During this phase, a project monitoring plan (PMP) is also devised to assess its achievement.

Discussing the project planning phase in reference to the example, the first step is to define the project objective, which should be specific and realistic. For this project, the objective is to make the project area ODF in the next five years. The second step is to identify the activities that are undertaken as part of the project to achieve its intended objective. The project can undertake activities like building HH and community toilets, creating awareness in the community about the ill effects of OD, tracking people who defecate in the open and counselling them, and providing information and technical support for building the right type of toilet etc. After identifying the activities, targets are set, i.e., the number of HH and community toilets that could be built within each year, and the number of communication campaigns that need to be undertaken to motivate people to not defecate in the open. This is followed by deployment of resources for the project, primarily in the form of finances available for implementation of the project. The money is utilised to recruit human resources (project staff) based on the defined roles and responsibilities. Physical and infrastructure resources like office, equipment etc., required for the project are also purchased. Timeline targets are set for communication campaigns and constructing toilets in a phased manner.

Stage IV: Implementation and Monitoring

The next stage is project *implementation* during which the formulated plan is executed. *Monitoring* of project activities is done concurrent to their implementation to ensure that the project is on track and as per the formulated plan. Monitoring helps to identify deviations, if any, from the project plan and also to introduce mid-course corrections. While executing a project, its quality, time, cost and risk management needs to be considered to ensure that it is successfully implemented within its predefined resources and timeline.

To take the case of the ODF project, the various activities that were identified in the planning stage are executed during this stage. The project staff is recruited and deployed, awareness campaigns are conducted about the ill effects of OD, and subsidy is provided for toilet construction etc. These activities are also simultaneously monitored to assess whether the toilets are being constructed as planned, both in terms of quantity and quality and whether the communication campaigns are being conducted as planned etc.

Stage V: Evaluation

After project activities are completed, many stake holders like project implementers, policy makers, the government, and the external audience, among others, want to know whether there is any change in the 'situation'. The stake holders also want to know whether this change is due to the project intervention or other external factors. An *evaluation,* helps to systematically assess the impact, effectiveness and the contribution of the project. Mid-term evaluations are helpful because they provide timely learning which helps in course correction. Post project evaluations help in getting insights that are helpful in formulation of other similar projects. Various techniques or designs are thus adopted for different projects in different situations. These evaluation designs are explained in detail in the following chapters.

Taking the current example of the ODF project, an evaluation should be conducted to know if the rate of OD has reduced in the project area. Also, in case it has reduced, would it be right to say that it has reduced due to the activities undertaken by the project?

To consider another example of a project which aims to increase the rate of institutional delivery, and examining its project cycle through the M&E lens, the first question at the start of the project would be to inquire about the proportion of women who had opted for institutional delivery. At the end of the project, the same question is asked once again regarding the proportion of women that had institutional deliveries. Also, whether it could be confidently asserted that the institutional delivery rate had increased only because of the project activities and not because of any other external factor. The change from the initial status (e.g., 65%) is measured and whether the desired situation (say 80%) has been reached or by what margin the gap has reduced.

After the various stages of the project cycle are examined, the practitioner needs to assess the stage at which M&E needs to be conducted and the specific M&E activities that should be performed at each stage of the project cycle.

1.3 Project Goal and Objective

One of the most critical and fundamental steps in designing a project is to define its objective and goal. Many people often get confused and use these terms interchangeably without realising that they are two different though interrelated concepts. Poorly defined goals and objectives cause ambiguity in project planning and implementation. Therefore, it is essential for any project to lucidly define its *goal* and *objective* and make sure that the entire project planning and implementation is aligned towards achieving them.

Objective

The *objective* of a project is the specific condition that the project targets to achieve and that too by itself. An objective is derived from a goal, has the same intention as a goal, but it is more specific, quantifiable and verifiable than the goal (SMART, Characteristics of Good Objectives, 2016). Usually, the project tries to address the inverted image of the core problem. For example, if the core problem in a specific village is the ‘high rate of OD’, then its corresponding objective would be to ‘reduce the rate of OD from 35 per cent to zero per cent in a period of five years’.

The SMART criteria is widely used to judge project objectives. A project objective is said to be SMART if it fulfils the following criteria

- Specific
- Measurable
- Attainable
- Realistic
- Time-bound

Using the SMART criteria to examine the objective to 'reduce the rate of OD from 35 per cent to zero per cent in a period of five years:

- Specific: The objective is said to be specific as it specifically targets the issue of OD.
- Measurable: It can be safely stated that this objective is measurable using the indicator 'rate of OD'. The data for this indicator can be collected using the HH survey data collection tool.
- Attainable: To assess if the target of reducing the OD rate from the current 35 per cent to zero per cent, for one lakh families is attainable or not the availability of resources, time in hand and other relevant factors are considered.
- Realistic: While setting an objective, the project team and other experts analyse whether the project would realistically be able to achieve its objective, keeping in mind the resources available for the project and the external factors related to the project.
- Time bound: It needs to be ensured that the project has defined timelines. In the case of the current example, the project is expected to achieve its objective in a time frame of five years.

If the project objective fulfils all the five criteria as part of the SMART framework, it is said to be a good objective. Another question which may be considered is whether a project should have a single objective or multiple objectives? By its definition, an objective is supposed to be precise, specific and definite. A project should thus simply have a single objective. In common parlance, the term 'project objectives' is often used. However, this term is incorrect and this usage is against the spirit of a 'project-based approach'. A project is essentially about the breaking down of a large problem into its smaller constituent parts, such that each problem is addressed one at a time in a single project. It is not about tackling all problems at once by trying to find a general solution for the various aspects of the problem. A project which has more than one objective is likely to lose focus, while it renders the objective open to ambiguity and subjectivity.

Therefore, a well-designed project should ideally have one objective which is clearly articulated.

Goal

By now the practitioner knows that the planned activities have to lead to fulfilment of the project objective. The objective also has a higher purpose which extends beyond the precincts of the project. This higher purpose is the *goal* which the project aims to contribute towards. Each project, by achieving its objective contributes towards a larger goal. Thus, it can be rightly said that while a project 'achieves' its objective, it 'contributes' towards the goal. In other words, the goal is the macro-level change that the project contributes to at the micro-level.

Thus, the objective of our example project, ‘reducing the rate of OD from 35 per cent to zero per cent in a period of five years’, at the same time contributes to the higher goal of 'improving the sanitation conditions in rural India’. Another current example would be of a project which aims to contribute towards the goal of improving the standard of living of the rural poor. Increasing the HH income of the rural poor, improving their access to social schemes, improving the sanitation conditions, improving their health, providing safe drinking water, increasing the HH assets etc., could be a few of the project objectives which would contribute towards the above mentioned larger goal. Thus, a goal is the intended change we seek; it is expansive in its scope and all-encompassing in its vision. In this way, many projects may refer to and contribute towards a single goal while achieving their respective objectives.

Hence, it is of paramount importance to define the objective and goal lucidly to ensure clarity in project planning and implementation.

CHAPTER TWO

DESIGNING PROJECTS

Having understood what is project and project cycle, it is important to understand what *project designing* is in order to ensure effective and useful M&E. M&E per se is not an activity that starts at the time the project is nearing completion. Monitoring starts from the day the project is rolled out, therefore, the M&E system needs to be conceived during the project design itself.

At the completion of the chapter, the practitioner will be able to:

- ✍ Develop the Results Chain for a project
- ✍ List and briefly describe Project Designing tools
- ✍ Describe Logical Framework Analysis for designing projects
- ✍ Describe the Logical Framework Analysis and Project Planning Matrix

2.1 Results Chain

During the project designing phase, it is essential to specify all activities and objectives that are to be achieved through the project. The *results chain* helps to manage projects, and at the same time, to understand the causal linkage between project intervention and its desired impact rather than managing the project based solely on activities. It helps to formulate a roadmap to the envisioned change, while highlighting the necessary conditions and assumptions required for ushering in a change in each situation (Foundations of Success, 2007).

Operations are based on the 'if-and-then' logic. For example, ***if*** we put fuel in a car’s fuel tank, ***then*** only can we drive and go somewhere in it. This 'if-and-then' logic is the means-to-an-end relation or a cause-and-effect connection between the system components. So, what implication does this system model have for projects?

Every project, as we know, has its own rationale of intervention, one that clearly addresses the nuts-and-bolts of the problem of 'what', 'when', 'why', 'how', 'who' and 'where'. The clearer a project is about the logic of change underpinning its project activities or processes, the better it can deliver the results or achieve the objective it has in mind.

A results chain thus describes the causal pathways of the activities translating into expected results i.e., the outputs, outcomes and impacts of a project. The results chain helps to track the progress of the project from its more immediate results (outputs), to a result more proximate to the achievement of the objective (outcome) and finally to a long-lasting result or goal (impact). A basic results chain has the following components:

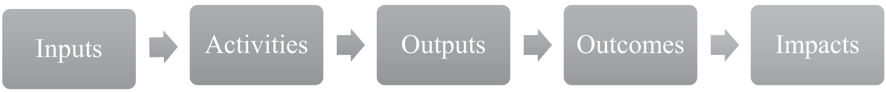

Figure 2.1: Results Chain

Inputs: This includes the resources that are available or allocated for the project. Input resources may be natural, human, and financial, depending upon the nature of the project. For example, funds allocated, human resources deployed, laptops allotted etc.

Activities: Activities are actions undertaken using the resources. In simpler terms, this is the work performed that converts inputs into outputs. For example, the training of frontline health workers (FLWs) on the counselling of women, building of separate toilets for girls in schools etc.

Outputs: Outputs are the immediate effect of the activities of a project. Outputs are also defined as the short-term results and often form the deliverables of the project. For example, counselling of mothers on institutional delivery is the output achieved from the activity training of FLWs on counselling. Also, the increased attendance rate of girls is an output of the activity of building separate toilets for girls in schools.

Outcomes: The mid-term results likely to be achieved from outputs are called outcomes. Outcomes are generally the objective which the project aims to achieve. For example, ‘increase in the rate of institutional delivery’ is an outcome achieved through the output of ‘effective counselling of women on institutional delivery'. Also, 'increased female literacy' is an outcome achieved through the output of 'increased female attendance rate'.

Impact: The final desired goal or the macro level goal that the project envisages to achieve is defined as its impact. Impact is what the project aims to contribute towards rather than trying to claim that it is what it would achieve by itself. For example, 'decreasing the Maternal Mortality Rate (MMR)' is the impact which the project aims to contribute to by providing the outcome, which is, 'increase in the rate of institutional delivery'. Also, 'increase in the empowerment level of women' is the impact which the project aims to achieve through its outcome of 'increased female literacy rate'.

Governing the interrelationships between inputs, activities, outputs, outcome and impact are several assumptions or enabling pre-conditions that are necessary for the delivery of project results and achievement of the project objective. They provide the necessary if not sufficient preconditions without which the project cannot hope to achieve its results. These assumptions are the causal inferences that govern the change processes in a project and lay the groundwork based on which correlations between the results, chain of inputs, outputs, outcome and impact are sought to be made explicit (Dharmendra Chandurkar, 2014).

Although generic in character, this framework can be fine-tuned to understand and unpack the non-linear, multi-contextual and multi-layered nature of change that defines and determines the landscape of a project. In general, it captures the project's broad canvas of change in one sweep, while it sheds light on the causal relationship among the various levels of change termed as outputs, outcome and impact.

Examples of the results chain are presented below in Figures 2.2 and 2.3 respectively.

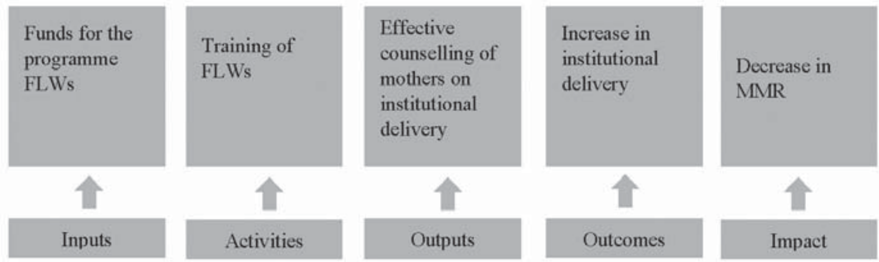

Figure 2.2: Results Chain - Example 1

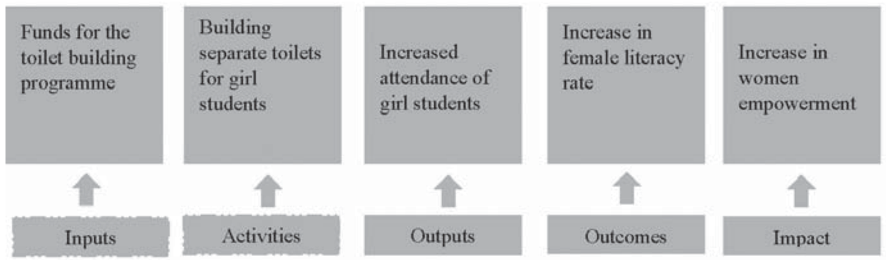

Figure 2.3: Results Chain - Example 2

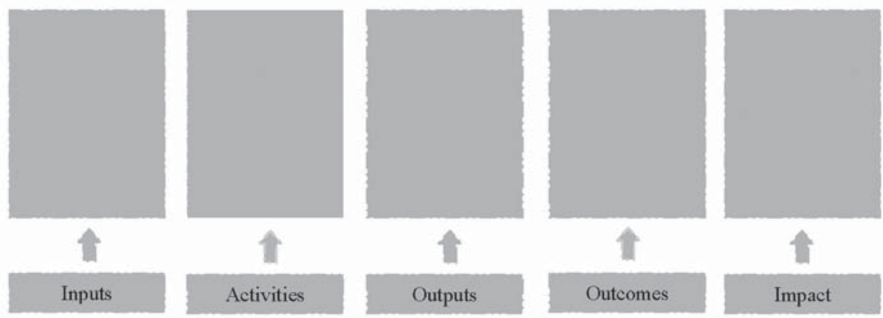

Figure 2.4: Results Chain - Example 3

Exercise: Can you think of other examples of results chain? Fill the results chain given below with your example.

Usually, in a project, there are multiple results that are envisioned and therefore the project has multiple chains. The example of a project with multiple results chain is presented below:

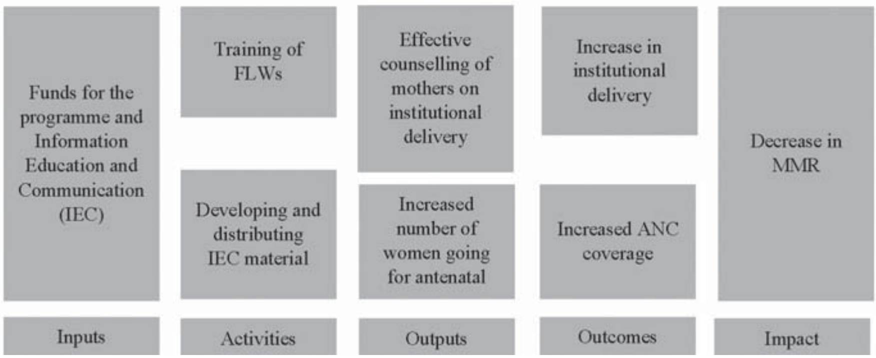

Figure 2.5: Results Chain - Example 4

2.2 Project Designing Tools

Various tools are used to structure and facilitate the process of project designing. In the development sector, the most commonly used tools are the Logical Framework Approach, Zielorientierte Projektplanung/Goal Oriented Project Planning (ZOPP/GOPP) and the Result Framework.

Logical Framework Approach or LFA is a simple project design tool that can helps to organise and structure our thinking while designing a project. As the name suggests, it is a logical approach for designing an efficient and effective project. On the one hand, it facilitates optimal resource allocation, while on the other hand, it sets performance measures and standards that provide a framework for M&E. All this helps in efficient project management.

ZOPP or GOPP is an adapted form of LFA that is suitable for the development sector. It also uses the same logical approach of LFA, but it is more flexible in accommodating the qualitative and subjective nature of issues inherent in the development sector.

Result Framework is more output oriented, as it focuses on the 'things that would be on ground' after completion of the project. These are basically the results that we want to achieve and the underlying assumption is that achievement of these results would lead to achievement of the envisaged objective.

All the above are tools for facilitating project development and each one has got its strengths and limitations. LFA is discussed in the following sub-chapter in greater detail.

2.3 Logical Framework Analysis

LFA is an analytical and project management tool which is widely used by funding agencies, international nongovernmental organisations (NGOs) and many government agencies for the designing and management of development projects. It was developed in the late 1960s to assist USAID to improve its project planning and evaluation systems (Republic of Serbia, 2011). LFA supports objective oriented planning and management. It can be used to perform systematic and structured analysis of a project or programme. This analytical process consists of a set of tools or techniques which can be used in managing development projects. Logframe matrix is the documented product at the end of conducting the logical framework analysis. On one hand, it facilitates optimal resource allocation, while on the other hand, it sets performance measures and standards that provide for a framework for M&E. It also takes into account the assumptions and risks envisaged while implementing the project (NORAD, 1999).

The steps that are followed in conducting an LFA are explained below. An example is used to reinforce the reader's understanding of this concept.

2.3.1 Situation analysisA project, as the practitioner is aware, does not exist in a vacuum, rather, it is a response to a situation or negative condition that it attempts to change. In the first stage of LFA, which is known as *situation analysis,* attempt is made to understand the existing situation or condition to identify what is 'wrong' with the given context that needs to be addressed through programme interventions. Along with this, the causes for the situation are analysed so that appropriate strategies may be adopted and specific activities designed targeting the causes. A good situation analysis serves as an entry point to the project by throwing light on what needs to be done under the project to address the negative conditions in each context.

2.3.2 Problem analysis

After the situation analysis, it is necessary to understand the problem at hand. Development projects are usually proposed as a response to address and overcome the identified social development issues. Problem analysis involves identifying the key problems and then establishing the cause and effect relationships between these problems. The primary purpose of this analysis is to try and ensure that it is not just the symptoms of the problem(s), but the 'root causes' that are identified and subsequently addressed (Kari, 2004).

For example, if a person goes to the doctor with high fever and the doctor prescribes medicine for high fever without diagnosing the cause for the fever, then the doctor is simply treating the symptom but not the cause of the problem. If the real cause of the fever is not discovered, it is quite possible that the fever will persist after the effect of the medicine wears out. Similarly, projects which only address the effects of the problems and not the underlying causes are unlikely to bring about sustainable benefits.

A common method used to identify the core problem is the *problem tree,* which helps to visually depict the connection between the various possible causes of the problem. This is best explored in a participatory workshop with key project stakeholders, who are asked to write down both the causes and effects of the problem in a structured way. This procedure makes it possible to clearly visualise the causes of the focal problem and its effects and to find out how different problems are related to each other. The next step is to plan activities and implement them within the framework of the project to treat the causes of these problems.

A clear and comprehensive problem analysis helps to build a sound foundation on which relevant and focused project objectives can be built.

A simplified example of a problem tree for a project which envisages to increase the HH income and thereby contribute to improving the standard of living in rural areas is presented below. The two key causes for the problem of low income in rural areas are identified and stated, viz., non-remunerative agriculture and lack of income generating activities (IGA). The further causes of each of these problems have also been listed in the problem tree.

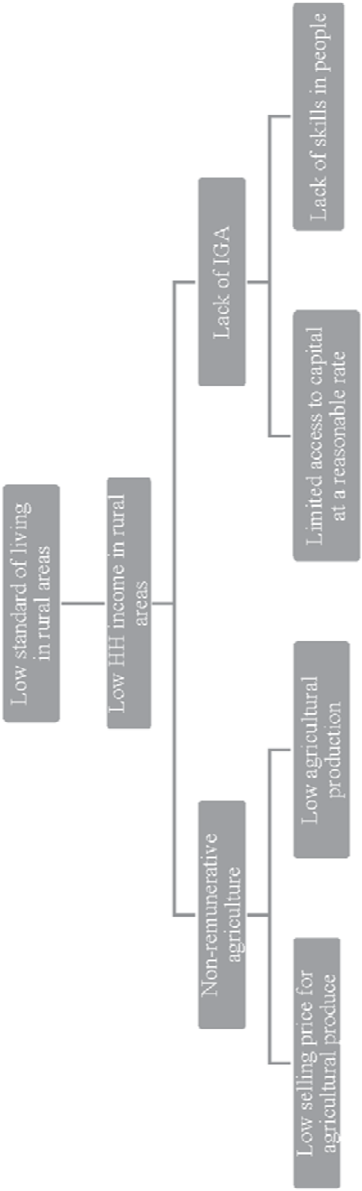

Figure 2.6: Problem Analysis - Example

2.3.3 Objective analysis

After identifying the problems and developing the problem tree, the next step is to conduct the *objective analysis*. Objective analysis is done to formulate the objectives of the project by developing an objective tree. The objective tree uses the same structure as the problem tree, but with the problem statements (negatives) turned into objective statements (positives). Based on insights from discussions with key stakeholders, the priority problems are identified. This means that the original problem statement does not need to be translated into objective statements.

While the problem tree shows the cause and effect relationship between problems, the objective tree shows the means to the end relationship between the objectives. This leads directly to developing the project's narrative description in the form of the Logical Framework Matrix.

The objective tree needs to be developed for the problems which the project aims to solve through its intervention. The objective tree delineates a logical sequence or cause and effect relation that needs to be followed to solve the problems as presented in the problem tree.

The objective tree of the example project intervention is presented in the diagram given below. In this case, 'Increasing HH income in rural areas' becomes the key objective/result which the project wants to achieve, and this is called the project outcome, which further contributes towards the intended impact i.e., 'Increasing standard of living in rural areas'. The short-term results i.e., the outputs and the activities planned to be undertaken in the project are the branches of the objective tree.

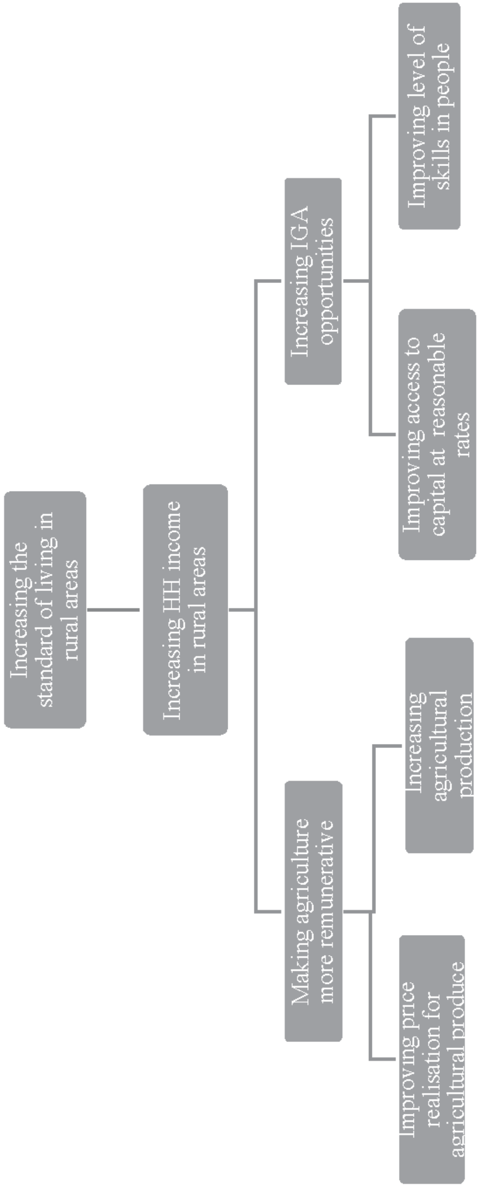

Figure 2.7: Objective Analysis - Example

2.3.4 Alternative analysis

The objective of *alternative analysis* is to identify the possible alternative options, assess their feasibility and select one project strategy. The different means-to-end branches in the objective tree, which constitute of alternate options, are then identified. It is advisable to have a set of assessment criteria against which the alternative strategies are ranked or compared. Criteria like benefits to the target group, resource availability, sustainability of the intervention, technical feasibility, relevance, likelihood of success, implementability etc., can be used for the assessment of the alternatives. Based on the assessment, the most feasible and appropriate alternative is selected which would further constitute the project strategy (NORAD, 1999).

Alternative analysis is a challenging stage as it requires synthesising a significant amount of information and then making complex judgements about the most suitable implementation strategy that should be pursued.

In the example used, after assessing the two alternative strategies available, the project management team selects ‘Increasing IGA opportunities’. The project activities, thus, will focus on the objective of ‘Increasing IGA opportunities' instead of the objective of working on 'Making agriculture more remunerative', assuming that the other stakeholders will work on the same.

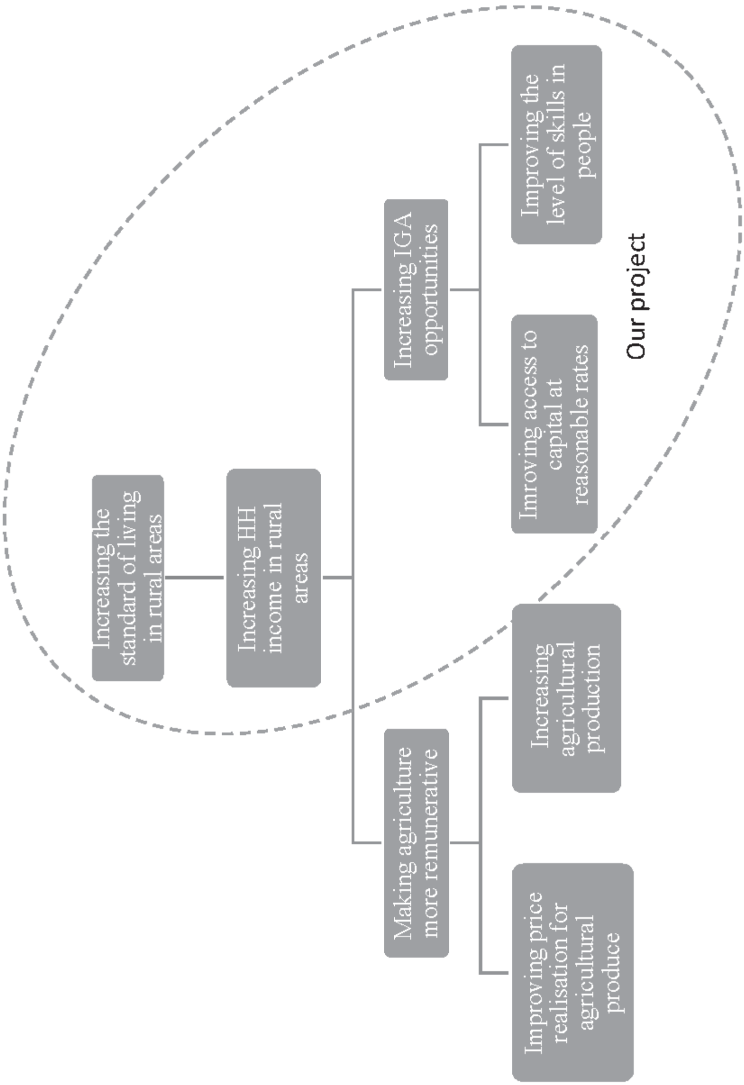

Figure 2.8: Alternative Analysis - Example

2.3.5 Risk and assumption analysis

After the project intervention is selected based on the analysis of the alternatives, it is very important to assess and address (to the extent possible) the potential project risks. Projects are always subject to influence by factors beyond the control of the project implementation team. This is applicable in the context of development projects too, where among other factors, the cooperation of various stakeholders, behaviour change or the target group's willingness to participate are essential for the project to fulfil its desired objective. Analysis of possible critical external and internal factors/risks gives the opportunity to assess the conditions that the project is working under.

Based on the *risk analysis*, the project management develops a risk management plan i.e., a plan of how to avoid potential risks. Some logframes prefer to use the term *'assumptions'* instead of 'risks', the distinction being that while risks are negative statements about what might go wrong, assumptions are positive statements about the conditions which need to be met to ensure that the project activities are being implemented as planned. Whether the term 'risks' or the term 'assumptions' is used, the purpose is the same, i.e., to assess and adapt the external impacts of the project to improve where possible, the robustness of the design.

After identifying the key risks associated with the project logic model, they can be plotted on a 2*2 matrix consisting of Impact and Likelihood on its axis. The project management should aim to mitigate the risks which lie in the top-right most quadrant i.e., risks which can have high impact and high likelihood as these risks can jeopardise the project in achieving its objective. The 2*2 matrix used for plotting potential risks is presented in the figure given below.

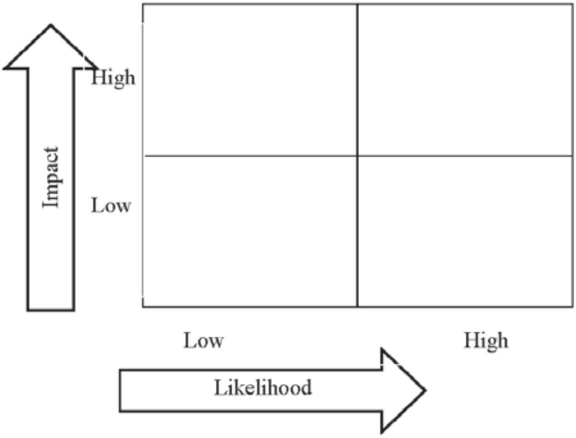

Figure 2.9: Risk Analysis

Also, in the case of the example used, several risks associated with the project can be identified. It is a good practice to identify risks, which can be both internal and external, at each level of programme implementation, i.e., from activity to output level, output to outcome level, outcome to impact level. The risks identified with the example project are listed below:

External Risks:

1. There is not enough demand of the goods produced under IGAs taken up by the target group.
2. Raw material cost is very high making the IGA unviable.
3. Shortage of availability of raw material.

Internal Risks:

1. Enterprises are not able to develop market linkages for selling their products.
2. Banks do not easily provide loans to individuals or self-help groups (SHGs) supported by the project.
3. The project target group is not self-motivated and enterprising enough to adopt new IGAs.

2.3.6 Development of Project Planning Matrix

Results of the situation, problem, objective, alternative and risk/assumption analysis are used as the basis for developing the Project Planning Matrix (PPM) or logframe. PPM or logframe, which is said to be a snapshot of the project intervention, is explained in detail in the next sub chapter.

2.3.7 The Logframe

The logframe or the PPM is a four-by-four matrix that details the logical connect between the various components of the project as well as the framework for assessing the performance.

The basic philosophy of the logical framework approach (LFA) is the logical approach in achievement of the impacts, where inputs given through activities leading to outputs and further to outcomes, which finally contributes towards impacts. This hierarchy forms the rows of the PPM matrix and the logical connect between various levels is termed as the vertical logic (FAO).

Performance measures called *indicators* are set for each hierarchical level, with sources of information listed for each indicator. Finally, assumptions that make activities translate into outputs, and outputs to purpose, or purpose to goal are also listed. This forms the horizontal logic of the PPM. Columns and rows of the PPM are given below.

Vertical Logic	**Horizontal Logic**			
	Narrative Summary	**Objectively Verifiable Indicators (OVI)**	**Means of Verification (MOV)**	**Assumptions/ Risks**
	Goal			
	Outcome			
	Outputs			
	Activities	**Inputs**		

Figure 2.10: Logframe Format

The logframe of the example project which aims to contribute towards increasing the HH income in rural areas is presented below. Objectively verifiable indicators (OVI) and means of verification (MOV) are discussed in detail in the following chapters.

Vertical Logic	Horizontal Logic			
	Narrative Summary	**Objectively Verifiable Indicators**	**Means of Verification**	**Assumptions/ Risks**
	Goal Increasing the HH income in rural areas	• HH income	Structured Interview Schedule	**Goal to Outcome** Non-profitability of the enterprises due to reasons like no market linkages, low demand and competition from the external market.
	Outcome 1.Increasing IGA opportunities	• Proportion of targeted beneficiaries engaged in micro-enterprises	Structured Interview Schedule	**Outcome to Impact** There is not sufficient uptake of IGA opportunities by target beneficiaries
	Outputs 1.1 Improving access to capital at reasonable rates 1.2 Improving skill level in people	• Amount of credit taken from financial institutions • Percentage of people acknowledging improved access to finances • Percentage of people acknowledging access to credit/ finance through financial	Structured Interview Schedule	**Activity to Output** Financial institutions are not supportive of providing finance to target beneficiaries

		institutions • Rate of interest for credit		
	Activity: 1.1.1 Forming SHGs 1.1.2 Performing all pathway activities in formation and ensuring quality of SHGs 1.1.3 Linking SHGs with financial institutions for access to credit 1.2.1 Providing skill based vocational training to women 1.2.2 Providing life skills training to women	**Inputs** • Number of SHGs formed under the project • Number of SHGs for which all pathways activities were performed • Amount of credit provided through linkages with banks • Number of people given vocational skills training • Number of people trained in life skills	Project MIS/Reports	

Figure 2.11: Logframe - Example

The logframe should be used as the basis of the funding application and then throughout the project lifecycle to track progress and adapt to changing situations. It can be used to review assumptions and implications, and to keep donors and other stakeholders informed of significant changes related to the project.

CHAPTER THREE

UNDERSTANDING MONITORING AND EVALUATION: BASICS AND KEY CONCEPTS

In this chapter, the practitioner will develop an understanding of the fundamental concepts of M&E, its importance for project management and learn the difference between the two. The practitioner will also understand how M&E is juxtaposed with the results chain of a project. Finally, at the end of the chapter, they will understand about indicators and learn how to design them. At the completion of the chapter, the practitioner will be able to:

- Define M&E
- Understand the key differences between M&E
- Define indicators and the important characteristics of a good indicator
- Learn how to design good indicators

3.1 Understanding M&E

Monitoring is defined as the concurrent process of tracking the implementation of activities of the project and attaining its planned outputs. It helps to provide real time information of the progress of the project in terms of completing its activities and achieving its immediate outputs, both in terms of quality and target. Monitoring, thus, is an activity to see if an ongoing project is proceeding on track. It involves the process of systematically collecting data to provide real time information for all stakeholders (managers, funders, participants) on the progress of implementation and the achievement of desired outcomes.

The critical functions of monitoring are: to gather feedback from the participants; collect data; observe the implementation of activities of the project; analyse contextual changes; and provide an early warning system of potential challenges. Analysis of monitoring data is critical to ensure that the project is being implemented in the right direction for it to achieve its intended outcomes. In case the project is not moving in its intended direction, midcourse correction should be done. Monitoring is applicable to all programme levels (from input, process, output and outcome). Most commonly, the focus is on output data, although it is also important to track the goals and the objectives. Monitoring should ideally be an internal function of the project management team. Monitoring, thus, plays a critical role in the success of a project.

Monitoring of results helps to:

- Improve strategies and targeting. Enabling decision makers to focus the project resources on areas where they can get the maximum output.
- Understand project implementation barriers or challenges in real time and suggest course correction measures.
- Ensure that the project is more effective and result oriented. It also focuses on impact level changes throughout the project, rather than just at the end of project evaluation.

Evaluation is defined as systematic research to see if a programme can achieve its intended outcomes and impacts. Evaluation is done firstly to see whether the envisaged objectives and goals have been achieved or not, and secondly, to see whether the achievement is because of the project interventions. It should assess the magnitude of change in the outcome and impact and whether the change in the outcome or the impact can be attributed to the project intervention. Evaluation assesses if there is any deviation from the goals and the objectives, and whether it can confidently be said that the objectives are achieved only because of project intervention. Evaluation, then, is a type of causal research that establishes the cause-effect relationship between the activities and the outputs on the one hand and the objectives and the goals on the other.

While monitoring facilitates mid-course correction in attainment of project outcomes, evaluation helps analyse variances from envisioned objectives and goals. By providing feedback to the project functionaries, M&E facilitates learning by doing. Development and enhancement of in-house capacities to anchor the M&E functions is, thus, a prerequisite for learning organisations.

3.2 Monitoring vs Evaluation

Having understood the definition of M&E, the practitioner can now list the key differences between the two.

Some of the distinctions between monitoring, which is to see ‘what we are doing’ and evaluation, which is to assess ‘what we have done’ are given in the matrix below (KEPA, 2015).

Table 3.1: Monitoring vs Evaluation

	Monitoring	Evaluation
Definition	Concurrent analysis of project progress towards achieving the planned results with the purpose of improving management decision making (Aquaknow, 2016).	Assessment of the magnitude of change in the results proposed by the project that may be attributed to the project.
When is it done?	Systematic activity should be done regularly throughout the project implementation.	It should be done only at specific points of time like in the middle of the project, at the change of phase, and at the end of the project etc.
Scope	Focuses on activities, outputs and indicators of progress and change.	Focuses on delivery of project outcomes and impacts. It assesses the progress towards the project objectives and goals.
Who does it?	Ideally, it should be an internal activity. This should be done by project staff or its target beneficiaries.	Ideally, it should be an external activity to avoid conflict of interest. It should be conducted by external evaluators while involving donors, project staff and project users.
Why is it done?	It is done to report project progress to the management, to identify the bottlenecks, take remedial action and modify the project implementation plans.	It is done to ensure accountability of the project, learn broad lessons and provide recommendations to similar projects. It highlights the potential and the achievements of the project.

3.3 M&E Levels

The matrix in the sub-chapter above makes it clear that monitoring is a day to day activity of assessment of project progress, whereas evaluation is the episodic assessment of overall achievement.

By juxtaposing M&E with the various stages of the results chain, the practitioner should be able to gauge at which stage of the results chain or theory of change of the project, should the M&E be focused on.

Table 3.2: M&E at Various Levels of the Results Chain

Stages in the Results Chain	M&E
Impact	**Evaluation**
Outcomes	
Outputs	**Monitoring**
Activities/Processes	
Inputs	

With respect to the Logical Framework, impact and outcomes fall within the domain of evaluation, whereas outputs, activities and inputs fall within the domain of monitoring.

3.4 M&E Typologies

Some popular typologies and terminologies commonly used in M&E are discussed below.

Monitoring Typologies

Monitoring is a task that is inherently undertaken by doers or implementers of the project themselves. Therefore, when monitoring is performed by the project team itself or internally by the project implementing team, it is called ***internal monitoring***. Sometimes, when projects involve parties or organisations external to the project for facilitating the monitoring functions, it is classified as ***external monitoring***.

If the project implementers restrict the monitoring process to themselves, it is called ***non-participatory monitoring***. Project stakeholders, including the communities remain mere information providers and have no role in analysing the information and providing inputs for project implementation. When functional participation of key stakeholders of the project, including the target community for the project is solicited, it is called ***participatory monitoring***. In the case of the sample project about working to make its target area ODF, a committee from the local community can be involved in formulating the monitoring plan, collecting information and analysing it. The community monitoring committee can keep a check on activities like communication campaigns and toilet construction (both in terms of quantity and quality).

Evaluation Typologies

While monitoring is inherently an internal activity, evaluation is an external activity usually done by those external (individuals/agencies/ institutions) to the project. Generally speaking, evaluation is external evaluation. However, when project implementers choose to undertake evaluation by themselves, it is called ***internal evaluation***.

Evaluation can be best defined based on the timing of conducting the evaluation. Evaluation *per se* is a less frequent activity generally undertaken at the completion of a project for assessment of attainment of the project objective. This is the ***post-project*** or ***post-facto*** evaluation. In the case of many long duration projects, evaluation is also conducted midway through the project implementation. ***Mid-term evaluations*** help to ascertain the level of achievement of long duration projects half way through the project. When the programme funders or the management is interested in a more regular assessment of the achievements of the project, the outcome and impact level assessment is done at a six monthly or a yearly period also. This time series design of evaluation is commonly known as ***concurrent evaluation***.

To relate the achievement of objectives and goals directly to the project, it may also be necessary to compare the status in the project area with an identical non-project area, which forms the control group while the project villages are the treatment group. This kind of evaluation design is called the ***control-experiment design***. Depending upon the time when the evaluation is implemented, it can be concurrent, mid-term or post-facto and internal or external.

3.5 Indicators

The concept of indicators is pivotal to M&E. As per its dictionary definition, an indicator is defined as a sign or a signal. In the context of M&E, an indicator is said to be a quantitative standard of measurement or an instrument which gives us information (UNAIDS, 2010). Indicators help to capture data and provide information to monitor performance, measure achievement, determine accountability and improve the effectiveness of projects or programmes.

Designing indicators is one of the key steps in developing an M&E system. As mentioned above, indicators are units which measure information over time to document changes in the specific conditions. With respect to the various M&E levels and the result chain of the project, specific indicators need to be developed for each stage of the results chain. Thus, there should be a different set of indicators at the impact level, at the outcome level as well as at the output, activity and input level. Also, for each level, there can be more than one indicator.

An indicator may be quantitative or qualitative based on the characteristics of information that it provides. Those that deal with information that can be expressed in numbers are quantitative indicators, while those dealing with information units expressed in any form other than in numbers, e.g., statements, are qualitative indicators. Another important attribute of quantitative indicators is that arithmetic functions can be applied to its corresponding data while this is not possible in the case of qualitative indicators. For qualitative indicators, their count or frequency may be considered. Income measured in rupees, the weight of a baby measured in kilograms and the number of toilets built are examples of quantitative indicators. If the same information of income or weight is collected in categories of high, medium and low, they are qualitative indicators.

SMART and SPICED Indicators

In order that development interventions are more result oriented, projects must be made SMART. The acronym SMART for developing project objectives is mentioned in Chapter 1, sub-chapter1.3. At the same time, it is important to make sure that the indicators or the performance measures also fit the SMART criteria. SMART indicators play an important role in results based project management as they ensure accountability (MDF Training & Consultancy, 2016). They have the following characteristics:

- Specific
- Measurable
- Attainable
- Realistic
- Timebound

Another school of thought advocates qualitative indicators represented by the acronym SPICED, which stands for the attributes listed below:

- Subjective
- Participatory
- Interpreted and communicable
- Cross-checked and compared
- Empowering
- Diverse and disaggregated

3.6 Design Indicators

Indicators are essential instruments for M&E, thus, practitioners need to keep in mind some of the critical points while designing or formulating them.

Firstly, creating new indicators or reinventing the wheel should be avoided till the time it is absolutely required. Over the years, development professionals have worked to provide M&E practitioners with sets of well tested and proven indicators and these sets of indicators should be referred to while formulating indicators for any project.

Secondly, while designing indicators, it should be made sure that they fulfil either the SMART or the SPICED criteria. Indicators can document change therefore, any indicator finalised should essentially be able to capture change in the condition that is being assessed using the indicator. A good indicator is therefore:

- Simple: As all the good things in the world are
- Measurable: Provides a measure for depicting change
- Precise: Has a definition so that it can be defined in the same way by all
- Consistent: Has consistent measurement results. On measuring the same thing, its value remains consistent and does not change over time

- Sensitive: Can capture the smallest amount of change in the indicator value
- Action Focused: Captures information that is eventually useful for stakeholders and leads to some action.

While designing indicators, it is very important to collectively brainstorm to identify candidate indicators for a specific condition. Once several indicators are listed for a given specific condition, the next step is to assess each of the indicators using the characteristics of a good indicator to find out whether the candidate indicator is simple, measurable, precise, consistent and sensitive. The source of data for the indicator and the reliability of the sources is also considered. The cost incurred in collecting data for this indicator is also considered while finalising the indicators. Candidate indicators that satisfy the criteria are then taken as indicators for assessment of that condition. Candidate indicators are also modified till they acquire the characteristics of a good indicator.

For example, in the case of the project which aims to make its target area ODF, the output level indicator is 'The number of individual household latrines (IHHL) constructed'. As constructing toilets is one of the key outputs expected from the project, this indicator helps in measuring the same. Similarly, as creating awareness about sanitation is another key activity, 'The number of village level meetings conducted to create awareness about sanitation' is another output level indicator.

Considering another example of a project which aims to improve maternal and child health (MCH) in its target area, 'Maternal Mortality Ratio' (MMR) and 'Infant Mortality Ratio'(IMR) are the result level indicators for this project. At the outcome level, the indicators are, 'The number of women with incidences of serious health problems related to child birth' and 'The number of women consuming iron fortified food or iron supplements during pregnancy'. At the output level, the indicators are 'The number of deliveries conducted by skilled health professionals', and 'The number of women receiving at least three antenatal care (ANC) visits'.

Defining Indicators

After selecting suitable indicators, it is very important to fully define them. No indicator should be deployed without fully defining it and making sure its essential components are lucid and concrete (UNAIDS, 2010). Each indicator definition should have the following components:

- Title: A brief heading that captures the summary or focus of the indicator.
- Definition: A lucid and to the point definition of each indicator so that everyone can interpret it in the same way.
- Source: The source i.e., the tool used for getting this indicator value and the respondent from whom this information is collected is also defined.
- Data Collection Frequency: The frequency at which the data is collected to derive at the indicator value is defined. This could be at quarterly, half yearly or annual intervals etc.
- Numerator: The variable that is included above the line in a common fraction.
- Denominator: The variable that is included below the line in a common fraction.
- Calculation Method: The method for calculating the indicator value is defined.

For instance, the complete definition of the indicator, 'The number of deliveries conducted by skilled health professionals', is stated below:

Title: The number of deliveries conducted by skilled health professionals

Definition: This indicator measures the number of deliveries conducted by skilled health professionals. The term skilled health professional refers exclusively to people with midwifery skills (auxiliary nurse midwives (ANMs), doctors and nurses) who are trained in the skills necessary to manage normal delivery cases and diagnose, manage or refer obstetric complications

Source: The sample survey conducted as part of the baseline, midline and endline survey

Data Collection Frequency: In the first quarter of the first year and in the last quarter of the third and in the fifth year of the project.

Numerator: The total number of deliveries attended by skilled birth attendants (SBA) as reported in the sample survey.

Denominator: The overall sample size of the number of women who had deliveries in the last two years.

Calculate: This indicator value is calculated by dividing the number of births attended by skilled health professionals by the total sample size of the number of deliveries conducted in the last two years.

CHAPTER FOUR

DEVELOPING A PERFORMANCE MONITORING PLAN

To ensure that a development project achieves its objective, it is very important to monitor its performance simultaneously. For this, a systematic collection and analysis of the project performance data is done, and based on this, feedback is provided to the project management staff. In this chapter the practitioner will understand what performance monitoring is and why it is important. Subsequently, the practitioner will learn what a *Performance Monitoring Plan* (PMP) is and about the tools used by development practitioners for project performance monitoring. The practitioner is taken step by step through the process of developing a PMP while understanding its key elements. Finally, the practitioner will learn about some key aspects which are taken care of to ensure the quality of the monitoring data.

At the completion of the chapter the practitioner will be able to:

- ✍ Understand the fundamentals and key concepts of a PMP
- ✍ Define the key components of a PMP
- ✍ Develop a PMP
- ✍ Understand the key levers for ensuring the quality of the monitoring data

4.1 What is Performance Monitoring Plan?

Project management is about managing 'what we are doing' in the domain of activities. The results of these activities are outputs, outcomes and impact which depict the performance of the project. *Performance management* is, therefore, geared towards 'what has been done' in terms of the results achieved in a project. It is to ensure that project objectives and goals are being met consistently in an effective and efficient manner.

Unlike project management which is limited to process management or managing project activities, time and cost, performance management is about managing project results. In project management, the starting point is activities and how to better manage or organise them to achieve the required results. However, in the case of performance-based results management, results become the basis on which all project planning is done, wherein the result that is to be achieved determines the activities that are to be carried out.

By developing the results chain, a conceptual map of the process of change is developed within the project. To track the performance of the project, it is necessary to track the project results i.e., project outputs, outcomes and impact. Thus, a *PMP* is needed to capture and monitor project performance. PMP, therefore, is a tool by which the outcomes or objectives are monitored and evaluated concurrently.

Once the results chain is developed, it is essential to translate it into a PMP that can monitor and assess the achievement of project results. This helps in managing the results by continuously providing evidence on the performance of the project.

PMP is also called *Performance Measurement Framework (PMF)* or the *Results Framework*. PMP serves as the reference document for concurrent performance monitoring of the project. A typical PMP format is shown in the matrix below.

Key components of the PMP are mentioned below:

1. Statement of expected results in their hierarchy-outputs, outcome and impact - or the conditions that are to be achieved
2. Indicators informing what information to seek so that the practitioner knows that the condition is achieved
3. Periodicity describing when the measurement of the indicator is made
4. Source from where the information is received at the defined periodicity
5. Baseline or the starting value of the indicator before the intervention situation or condition
6. Milestones that are planned for the changed condition, as the project moves forward
7. Target of the condition to be achieved at the completion of the project
8. Responsibility of who will fetch the information at the desired periodicity from the defined source.

Table 1.1: Project Monitoring Plan Format

Expected results	Indicator	Periodicity	Source	Base line Value	Milestones			Target	Responsibility
					Y1	Y2	....		
Impact									
Outcome									
Outputs									
1............									
2............									

4.2 How to develop a Performance Monitoring Plan

The practitioner has understood by now that PMP is a tool for developing an overarching plan for managing the performance of a project, which includes activities for monitoring and evaluating the project. Using this tool a plan is formalised for collecting and monitoring the project performance related data (USAID Centre for Development Information and Evaluation, 1996). PMP, at the bare minimum, includes the following details:

1. Identifying performance indicators at various levels and subsequently defining them.
2. Indicator-wise details like data sources, method of collection, frequency, timing etc.
3. Building teams and delineating roles and responsibilities for supervision and collection of performance data.
4. Plan for analysis of performance data.
5. Developing processes to ensure that it is reported and further used to provide feedback to the programme for informed decision making.

A step by step process is given below for developing a PMP:

Step 1: Populate the Expected Results Column

In the Results Chain developed in Chapter 2, all levels of results viz., impact, outcome and outputs were defined. As the first step of developing a PMP, the outputs and outcome from the Results Chain or the PPM (Project Planning Matrix) are put in the respective Expected Results column at the assigned level. Outputs and outcomes are focused on as the project is accountable for their achievement. Inclusion of impact in the Expected Results column is kept optional as a project always contributes towards the impact but does not achieve it by itself.

An example of a Reproductive, Maternal, Newborn and Child Health (RMNCH) focused project is used to develop its PMP. The project aims to improve MCH in its target area by increasing community awareness and sensitising them on maternal health issues as well as by increasing the accountability of government institutions in implementation of their programmes. The project will create village level monitoring committees for spreading awareness about the critical MCH care practices. It will also hold the government health system accountable through advocacy to ensure that beneficiaries have access to their entitlements under the national flagship health programmes. The project also aims to encourage participatory formulation of the Annual Implementation Plan (AIP) for the government plan at the district level.

The first step of populating the expected results column is given below:

Table 1.2: Developing PMP - Example

Expected Results	Indicators	…......................	Responsibility
Impact Ensuring universal access to maternal health services for achievement of MDG-5			
Outcome Improved status of women's health in the project target area			
Output 1. Women have increased access to obstetric care			
2.Increased capacity of community to demand for the rights and entitlements under government health programmes			
3. Increased awareness and knowledge about important maternal and child practices and contraception methods			
4.Participatory formulation of AIP in consultation with village health committees (VHCs) (Practice Level Change)			

Step 2: Develop Indicators for Results at Each Level

As mentioned above, during its implementation, change in result level indicators of the project is tracked and measured through the PMP. Indicators provide information regarding change in the condition and its magnitude of change. In Chapter 3, the practitioner learnt about the essential characteristics of a good indicator and how to design and define them.

Developing or selecting suitable indicators remains a challenging task in any performance measurement framework. For example, poverty-reduction can be measured by using multiple indicators ranging from per capita income to calorie consumption. Depending upon the indicator selected, information that is collected is either income or kilocalories consumed at time t1 (start of the project) and time t2 (end of project), which depicts whether the condition of poverty had increased or decreased or remained the same over a period. It needs to be reiterated here that it is only the condition that is changing, with the indicator simply pointing out whether it has increased or decreased or stayed the same over a period of time. Being a unit of information, an indicator, by its very nature, is value-neutral and can at best be an approximation. Based on the SMART criteria of developing an indicator, a suitable indicator for each expected result is selected. It is also important that there are complete definitions of each indicator in order to avoid confusion and create consistency. For each level of result, it is suggested that there should be a minimum of one indicator and a maximum of three. More than three indicators imply that it is not clear what the project is trying to achieve as its result.

In the example used, indicators are defined in the next step for developing the PMP and at each level of the results, they are populated with impact, outcome and output.

An example of the PMP after identification of suitable indicators is given below:

Table 1.3: Developing PMP-Example

Expected Results	Indicators	....	Responsibility
Impact Ensuring universal access to maternal health services for achievement of MDG-5	Maternal Mortality Ratio (MMR)		
Outcome Improved status of women's health in the target project area	Number of women reporting serious health problems related to child birth		
	Number of women consuming iron rich/ fortified foods or supplements		
Output 1. Women with increased access to obstetric care	Number of births attended by skilled health professionals		
	Number of women benefitted under the *Janani Surkasha Yojana* (JSY) scheme		
	Number of primary health clinics (PHCs) with referral services for complicated pregnancies		

	Number of women with access to contraceptives and safe abortion services		
2.Increased capacity of community to demand for the rights and entitlements under government health programmes	Number of village level communities formed which are capable of demanding entitlement under the national government health programme		
	Number of VHCs which are capable of registering complaints with the health department in case any entitlement under the government health programme is not received		
3. Increased awareness and knowledge about important MCH practices and contraception methods	Number of women who are able to articulate the key danger signs during pregnancy		
	Number of women who have knowledge of temporary method of contraception		
	Number of women who are aware about the availability of		

	contraceptive and abortion services at various levels of the public health system		
4.Participatory formulation of annual implementation plan (AIP) in consultation with VHCs) (Practice Level Change)			

Step 3: Assign Periodicity, Identify Sources and Assign Responsibility

The next step is to decide the period or frequency at which each indicator in the PMP is to be measured. Periodicity is decided based on factors like the minimum time that is required for change to be visible in the condition under observation, the cost of data collection of a specific indicator, requirement of indicator etc. For e.g., if the change is expected within a short duration, the periodicity can be low such as quarterly or half-yearly. If the project intervention is expected to take a longer duration to have an effect on the indicator value, then the periodicity would be higher, say annual or biannual. In the case of indicators for which data collection is a complex process or where special investigating skills are required which make the data collection costly, e.g., in agriculture production, testing the haemoglobin (Hb) level in blood etc., it is better to collect this data over a longer periodicity or for a smaller sample.

Apart from assigning the periodicity, it is important to identify the source from where the indicator information is drawn. The exact source of this information is reported so that anyone can access and understand this information. Also, if there are multiple sources for availing the indicator information, then at least two of those sources, which seem to be the most credible and accessible, are mentioned. The primary sources of indicator value can be census surveys, sample surveys or concurrently collected project MIS data. The secondary sources of information can be past sample surveys or government census surveys or some secondary MIS data. While using secondary data, it is important to consider the gap between the time of the actual data collection of the secondary survey and the time when it is being used. In case the time gap is substantial, it is not advisable to use this secondary data source. Data sources are decided while considering factors like availability, reliability, cost etc.

It is important to assign the responsibility of gathering information specific to each indicator to someone from the project team, to a stakeholder or to an external M&E team if deployed.

The PMP of the sample project with its periodicity, source and the person/agency responsible for gathering information is populated in the table below:

Table 1.4: Developing PMP - Example

Expected Results	Indicators	Periodicity	Source	....	Responsibility
Impact **Ensuring universal access to maternal health services for achievement of MDG-5**	MMR	Annual	Sample Registration Survey		**Monitoring, Evaluation and Learning (MEL) Manager**
Outcome **Improved status of women's health in the project target area**	Number of women reporting serious health problems related to childbirth	Annual	HH Structured Interview, Project Annual Survey (PAS)		**Designated Research Agency**
	Number of women consuming iron rich/ fortified foods or supplements	Annual	HH Structured Interview, PAS		**Designated Research Agency**
Output **1. Women with increased access to obstetric care**	Number of births attended by skilled health professionals	Annual	HH Structured Interview, PAS		**Designated Research Agency**
	Number of women who benefitted under the JSY scheme	Annual	HH Structured Interview, PAS		**Designated Research Agency**

	Number of primary health clinics (PHCs) with referral services for complicated pregnancy	Annual	CHC/PHC Register Audit, PAS		**Designated Research Agency**
	Number of women with access to contraceptives and safe abortion services	Annual	HH Structured Interview, PAS		**Designated Research Agency**
2.Increased capacity of the community to demand for the rights and entitlements under government health programmes	Number of VHCs formed which are capable of demanding entitlement under the national government health programme	Annual	Project MIS		**MEL Manager**
	Number of VHCs capable of registering complaints with the health department in case any entitlement under government health programme is not received	Annual	VHC Representative in-depth interviews (IDIs), PAS		**Designated Research Agency**
3. Increased awareness and knowledge about important MCH practices and	Number of women who can articulate the key danger signs during pregnancy	Annual	HH Structured Interview, PAS		**Designated Research Agency**

contraception methods	Number of women with knowledge of temporary methods of contraception	Annual	HH Structured Interview, PAS		**Designated Research Agency**
	Number of women who are aware about the availability of contraceptives and abortion services at various levels of the public health system	Annual	HH Structured Interview, PAS		**Designated Research Agency**
4.Participatory formulation of AIP in consultation with VHCs (Practice Level Change)					

Step 4: Put Milestone Values for each Indicator from the Baseline to the Endline of the Project

For monitoring the performance of a project and ensuring that it is result oriented, it is very important to define a starting point for gauging or calculating the change and tracking it intermittently. The baseline is the value of the indicator at or just before the start of the project. This acts as a reference point for assessing change in the condition after the project has been initiated.

After putting in the baseline value, key milestones and target values are assigned at specific time gaps for the indicators. The target is *a posteriori* condition that is envisaged at or after completion of the project. In other words, it is the concrete result that is sought at the end of the project efforts. Milestones are the benchmark values that are hoped to be achieved at specific points of time within the project implementation years. The milestone values are to be assigned based on the way the project implementation has been planned. For instance, for a five-year project, formative work is undertaken or the foundation is built during the first year of project implementation. The implementation intensity picks up from the second year and peaks during the third or fourth year. The fifth year is usually reserved for the handing over of the project after ensuring its sustainability. Therefore, the project milestones are set according to the planned pace or intensity of work.

The PMP of the sample project after the key milestones values are assigned is given below:

Table 1.5: Developing PMP - Example

Expected Results	Indicators	Periodicity	Source	Baseline	Milestone 1 (Year2)	Milestone 2 (Year 3)	Milestone 3 (Year 5)	Respons-ibility
Impact Ensuring universal access to maternal health services for achievement of MDG-5	MMR	Annual	Sample Registration Survey	200	MMR to be brought down by 5% from the baseline value in the target area	MMR to be brought down by 10% from the baseline value in the target area	MMR to be brought down by 25% from the baseline value in the target area	**MEL Manager**
Outcome Improved status of women's health in the target project area	Number of women who are reporting serious health problems related to childbirth	Annual	HH Structured Interview, PAS	30%	27%	20%	10%	**Designated Research Agency**
	Number of women who are consuming iron rich/fortified foods or supplements	Annual	HH Structured Interview, PAS	5%	8%	15%	30%	**Designated Research Agency**
Output 1. Women with increased access to obstetric care	Number of births which are attended by skilled health professionals	Annual	HH Structured Interview, PAS	70%	75%	80%	90%	**Designated Research Agency**
	Number of women who benefited under the JSY scheme	Annual	HH Structured Interview, PAS	40%	45%	50%	60%	**Designated Research Agency**

	Number of PHCs with referral services for complicated pregnancy	Annual	CHC/PHC Register Audit, PAS	2 PHCs out of 30 PHCs have referral transport support	5 PHCs out of 30 PHCs have referral transport support	10 PHCs out of 30 PHCs have referral transport support	20 PHCs out of 30 PHCs have referral transport support	**Designated Research Agency**
	Number of women with access to safe abortion services	Annual	HH Structured Interview, PAS	5%	10%	15%	25%	**Designated Research Agency**
2.Increased capacity of community to demand for the rights and entitlements under government health programmes	Number of VHCs formed which are capable of demanding entitlement under the national government health programme	Annual	Project MIS	Out of 500 interventi on villages, none have VHCs	Out of 50 intervention villages, none have VHCs	Out of 100 interventi on villages, none have VHCs	Out of 250 intervention villages, none have VHCs	**MEL Manager**
	Number of Village Health Committees that are capable of registering complaints with the health department in case any entitlement under government health programme is not received	Annual	VHC Representative, IDIs, PAS	Zero	50	100	250	**Designated Research Agency**

3.Increased awareness and knowledge about important MCH practices and contraception methods	Number of women who can articulate the key danger signs during pregnancy	Annual	HH Structured Interview, PAS	15%	20%	25%	50%	**Designated Research Agency**
	Number of women with knowledge of at least two temporary methods of contraception	Annual	HH Structured Interview, PAS	30%	35%	45%	70%	**Designated Research Agency**
4.Participatory formulation of AIP in consultation with VHCs (Practice Level Change)								

Step 5: Develop Measurement Protocols for Behavioural Change

By following Step 2 to Step 4, physical and financial changes sought through the project are taken care of and tracked via their respective indicators. However, different measurement protocols or systems for behavioural outputs are needed, which are essentially practice-level changes that are sought through the sample project. At the same time, it is to be noted that measurement protocols are developed for practice level indicators only in cases where the project aims to achieve institutional changes or behavioural changes.

In the case of projects where the perceived behavioural change falls in the realm of outputs or outcomes which are not physically observable or verifiable, *outcome mapping* serves as a vital tool to map and track process-level changes along the pathway of change. Using a graded measure of change known as 'progress markers', it seeks to unpack the multiple layers of change.

Progress markers are statements describing the practice or process- level changes that lead to the envisioned outcome. For example, the key progress markers for a project aiming to improve participative governance and budgeting at the *panchayati* level are: transparency in public budgets, public participation in the budget formulation process, simplification of budget documents, social audit of expenditure by the public etc. All these constitute key practice-level demands that require institutional change. These are best mapped through progress markers and ideally a project should not have more than five progress markers.

These progress markers define and describe the milestone changes that are required to achieve the behavioural change from the current situation to the envisaged situation. These are progressive milestones, ranging from the low-hanging fruit to the higher-order or difficult to attain changes. The first milestone is an 'expect to see' yardstick, which refers to the minimally acceptable level of change or first-level change expected from the intervention. The second milestone is what is called 'like to see' or the change that could realistically be achieved, while the final milestone is known as ‘love to see’ or the most desired or ambitious change that is foreseen. To use a real-life example, passing an exam is the minimally acceptable or what we 'expect to see', whereas getting a first division is what we would 'like to see' and finally, to get a first-class distinction is what we would ‘love to see’.

These progress markers are analogous to milestones and targets for indicators and their periodicity, source and the person/agency responsible for gathering information is assigned accordingly.

The PPM of the example project, is presented below after adding the its behavioural outcomes and related progress markers

Table 1.6: Developing PMP Example

Expected Results	Indicators	Periodicity	Source	Baseline	Milestone 1(Year2)	Milestone 2(Year 3)	Milestone 3(Year 5)	Responsibility
Impact Ensuring universal access to maternal health services for achievement of MDG-5	MMR	Annual	Sample Registration Survey	200	MMR to be brought down by 5% from the baseline value in the target area	MMR to be brought down by 10% from the baseline value in the target area	MMR to be brought down by 25% from the baseline value in the target area	**MEL Manager**
Outcome Improved status of women's health in the target project area	Number of women who are reporting serious health problems related to childbirth	Annual	HH Structured Interview, PAS	30%	27%	20%	10%	**Designated Research Agency**
	Number of women who are consuming iron rich/fortified foods or supplements	Annual	HH Structured Interview, PAS	5%	8%	15%	30%	**Designated Research agency**

Output **1. Women with increased access to obstetric care**	Number of births which are attended by skilled health professionals	Annual	HH Structured Interview, PAS	70%	75%	80%	90%	**Designated Research agency**
	Number of women who benefitted under the JSY scheme	Annual	HH Structured Interview, PAS	40%	45%	50%	60%	**Designated Research Agency**
	Number of PHCs with referral services for complicated pregnancy	Annual	CHC/PHC Register Audit, PAS	2 out of 30 PHCs have referral transport support	5 out of 30 PHCs have referral transport support	10 out of 30 PHCs have referral transport support	20 out of 30 PHCs have referral transport support	**Designated Research Agency**
	Number of women with access to safe abortion services	Annual	HH Structured Interview, PAS	5%	10%	15%	25%	**Designated Research Agency**
2.Increased capacity of the community to demand for rights and entitlements under government health programmes	Number of VHCs formed which are capable of demanding entitlement under the national government	Annual	Project MIS	Out of 500 intervention villages, none have VHCs	Out of 50 intervention villages, none have VHCs	Out of 100 intervention villages, none have VHCs	Out of 250 intervention villages, none have VHCs	**MEL Manager**

	health programme							
	Number of VHCs which are capable of registering complaints with the health department in case any entitlement under the government health programme is not received	Annual	VHC Representative IDIs, PAS	Zero	50	100	250	**Designated Research Agency**
3.Increased awareness and knowledge about important MCH practices and contraception methods	Number of women who can articulate the key danger signs during pregnancy	Annual	HH Structured Interview, PAS	15%	20%	25%	50%	**Designated Research Agency**
	Number of women with knowledge of at least	Annual	HH Structured Interview, PAS	30%	35%	45%	70%	**Designated Research Agency**

	two temporary methods of contraception							
Output(Practice Level Change)	**Progress Marker**	**Periodicity**	**Source**	**Expect to See**		**Like to See**	**Love to See**	
4.Participatory formulation of AIP in consultation with VHCs	**District wise AIP planning**			**State government officials including VHC representative, AIP development meetings**		**Inputs of VHC representatives are taken while developing an AIP**	**Representative of VHC is included in the executive body which develops the district level AIP**	

To sum up, in this sub chapter, the practitioner has learnt to develop a PMP, which is a plan to monitor and manage results, by following a step-by-step process.

4.3 Ensuring Quality of Monitoring Data

It is said that the quality of project monitoring is as good as the quality of data collected. It is very important to collect good quality data so that the programme management team and other important decision makers can trust it and make use of it for tracking and improving the programme. Good quality data should be accurate, complete, consistent (across different sources), timely, useful, precise, and accessible.

Robust data quality assurance mechanisms and regular data auditing ensures the veracity of the monitoring data.

Data Quality Assurance

For data quality assurance, systems, document protocols and guidelines for ensuring monitoring data quality are developed throughout the various project stages. Checks and procedures are also defined right across the various stages viz., designing of monitoring formats, translation of monitoring formats, during data collection, data entry or digitisation etc. Key aspects that should be included while developing a data quality assurance plan are listed below:

1. Listing quality assurance mechanisms to be followed at all project stages.
2. Assigning roles and responsibilities for data collection, data auditing and deciding their frequency.
3. Conducting concurrent data auditing and investigating the reasons for data variance if any.
4. Calibration of the monitoring and reporting formats to check their precision, and translation, while ensuring they are not biased in any way.
5. Validation of data and data cleaning systems.
6. Conducting training to build the capacities of the data collection team and to maintain consistency in the way the tools are administered.
7. Defining roles and responsibilities for data collection and data auditing.

8. Defining data quality benchmarks and action to be taken in case the collected data falls below these benchmarks.

Data Auditing

As part of quality assurance, it is also necessary to devise the plan for conducting data auditing. Data auditing is an exercise by means of which the veracity of data is checked. Data auditing is done primarily using two techniques-one, by conducting spot-checks and two, by conducting back-checks. In spot-checks, the data quality auditor is present at the time when the data collector is collecting data. In this way, real time feedback can be provided to the data collector on tool administration or other related issues. *Spot-checks* done during the starting phase of data collection help to assess how survey tools are being administered by the data collectors. In back-checks, the data auditor goes to a randomly selected sample of respondents from whom data has been collected by the main survey team. It is suggested that there should be a separate team for conducting back-checks, and these should be conducted immediately, i.e., on the very next day of the data collection. Both data sets are then matched to find if there is any variance between the two sets. This helps in gauging the quality of the data being collected, and based on this, improvement or rectification measures are taken up as and when required. At the same time, it is important to keep in mind that mostly factual questions should be used as part the back-check tool so that the possibility of variance due to different responses by the respondent or any other external factor is at a minimum. Triangulation with other data sources is also useful for auditing the quality of the data.

CHAPTER FIVE

APPROACHES TO MONITORING

In the previous chapter the practitioner learnt how to develop a project PMP. In this chapter, the practitioner will learn about the different approaches that are adopted for monitoring a project, about progress and process monitoring, how monitoring can be done in a participative way while involving all project stakeholders, and how to conduct monitoring without using indicators.

At the completion of the chapter, the practitioner will be able to:

- ✍ Understand and explain the definition and key concept of progress and performance monitoring
- ✍ Understand the definition and concept of Participatory Monitoring
- ✍ Explain the concept of Most Significant Change (MSC) and Outcome Mapping
- ✍ Understand and explain the key levers for ensuring quality of monitoring data

5.1 Progress and Process Monitoring

In this section, the practitioner will learn why progress and process monitoring are important, what the key differences between them are and how they can be implemented for monitoring a project.

Progress Monitoring

Progress monitoring, as the name suggests, aims to assess the progress of a project towards its objectives and target milestones. It is advisable to do progress monitoring concurrently or intermittently along with the project implementation to ensure that the project is on track. Progress monitoring is an ongoing process which involves collecting and analysing the output level or even outcome level indicators related to the project to see if the project is on track to deliver its planned results.

Taking the example of a project which aims to eradicate OD and improve sanitation by building individual household latrines (IHHL)in the project area, one lakh HHs are identified without toilets using a baseline study. The project aims to cover this deficit over a period of five years. During the project planning stage, it was decided that the project would build ten thousand toilets in the first year, twenty thousand in the second year, twenty-five thousand in the third and the fourth year respectively and twenty thousand in the fifth year.

In this case, through progress monitoring, a tab can be kept concurrently on the number of toilets that are constructed per year. Through progress monitoring, the number of toilets constructed is assessed and then it needs to be seen if the project can achieve the numbers it has planned to achieve. Based on this, feedback is provided to the project implementers to check whether the project is on track or whether it is lagging. In progress monitoring, the focus is more on the quantity of target achievement in comparison to its quality.

Project MIS (Management Information System) data is generally the source of data required for progress monitoring. This should ideally be done by the project management team itself.

Process Monitoring

Process monitoring, as the name suggests, includes monitoring of the processes and the activities done as part of the project implementation. Its objective is to focus on the quality of the implementation rather than focusing only on the targets or the milestones achieved by the project.

As an initial step, the processes that need to be adopted for implementation of each project activity are delineated and listed. Activity-wise, this is an ideal process and sequence in which each activity that is to be implemented is first envisaged so that the desired results are achieved. A process intensive approach is required for ensuring quality and sustainability of project outcomes.

A process check-list is developed by making a list of all the steps that are followed as part of the ideal process implementation. Process monitoring is usually done using these process checklists. The activities and the processes are observed and recorded on the checklist. In case any deviation is observed from the ideal required process, it is recorded.

For example, to monitor the process of subsidy transfer to beneficiaries for construction of IHHL, at first, all the steps are listed in a sequential order.

1. The beneficiary should take the required photographs of the toilet site before, during and after its construction.
2. Get the toilet constructed.
3. Go to the project district office and intimate them about the toilet construction and get the form for availing subsidy.
4. Fill the form requesting for the subsidy amount and attach the following documents:
    a. Three photographs of the various stages of toilet construction as mentioned above.
    b. Identification card of the beneficiary to prove that the beneficiary is a resident of the mentioned area.
    c. A verification letter from the *sarpanch* that the beneficiary's HH does not have a toilet.
    d. Bank account number in the name of the beneficiary.
5. The form is submitted to the district project office after being duly filled.
6. The subsidy should be transferred to the beneficiary's account within 15 days of the receipt of the application form at the district office.

Based on the above-mentioned steps, a checklist for process monitoring is prepared. Using this check- list as a reference, the processes that need to be followed as part of the project activity are monitored. Deviation from the ideal process is also recorded. As part of process monitoring, the reason for this deviation is assessed and the feedback is given to the project management for improvement of the programme.

5.2 Participatory Monitoring and Evaluation

Participatory M&E, as the name suggests, is a process through which all the project stakeholders at various levels are engaged in monitoring and evaluating a project or a programme. Unlike conventional M&E, in participatory M&E, all stakeholders of the project including donors, implementation agencies, primary stakeholders and other stakeholders share control over the process, content and results of the M&E activity. Also, all stakeholders are involved in identifying or taking corrective action in case any issues or deviation is observed (Forster, 2002).

Participatory M&E's core principle is that it considers the project beneficiaries as active participants and not just information providers. It believes that their capacities should be built so that they can be actively engaged in analysing and reflecting on the project performance so that the performance can be improved based on the inputs from the ground. During participatory M&E, all stakeholders are engaged throughout all stages of the M&E, which include formulating the M&E framework, information collection, collation, analysis, interpretation and finally, decision making.

Why is Participatory M&E necessary?

Why is Participatory M&E (PM&E) necessary and what are its advantages? Engaging all stakeholders has increasingly been considered as important for the M&E process as it offers new ways of assessment and deriving learning from the project. PM&E makes the M&E process more inclusive i.e., internal and contextual to the programme and also more responsive to the needs and expectations of the primary stakeholders of the project.

PM&E aims not only to measure the effectiveness of a project but also ensures accountability to its beneficiaries and increases transparency in the M&E process of the project. Making the process participatory empowers the project beneficiaries and inculcates a sense of ownership amongst the stakeholders of the project. Also, making the process participatory buttresses the process of formulating corrective action for improving the performance and outcomes of the project.

PM&E also helps to build the beneficiaries' commitment and increases their understanding of the design, planning and implementation of the project.

Steps in conducting Participatory M&E

Key steps involved in conducting PM&E are:

Step I: Identify the Key Stake holders

As the first step, all the stakeholders in the project, right from the project funders, project implementers, government agencies, other key boundary partners and its target beneficiaries are identified to be included in the M&E process.

Step II: Developing a framework for PM&E

After identifying the key stakeholders, the next step is to develop a framework for PM&E. This is one of the most important and challenging stages as it involves getting all the stakeholders together on board for the first time to provide their inputs, raise their concerns and expectations from the process (Gaventa). At this stage, the objective of the PM&E is established. Usually there are lot of deliberations and negotiations between the stakeholders at this stage. This process can answer key fundamental questions like:

- What are the resources and limitations at hand for the PM&E?
- What is the role and the responsibility of each stake holder?
- What information is required?
- From whom is this information to be collected?
- How is it to be incorporated to improve the programme implementation?

Step III: Developing the Indicators and Performance Monitoring Plan

As in the case of conventional M&E, after formulating the objective, the next step is to develop the indicators at each level of the Results Chain and define them fully. While developing the indicators in PM&E, special thought needs to be given to who the end users of this information are, and how it can be used to further improve the programme. The indicators should be developed keeping in mind the SMART criteria which was elucidated in Chapter 3.

Subsequently, the project PMP and the matrix will be developed as it has been done in Chapter 4. All the stakeholders are jointly involved in developing the PMP. Also, the responsibilities are distributed among all stakeholders including the beneficiaries, unlike in the case of the conventional M&E.

Step IV: Collecting Data

After the objective and framework are finalised, the next critical step is to determine how to go ahead with the data collection. For this, the tools and techniques used for data collection are identified. A wide range of tools is available for this and a tool based on the context and requirements of the project is selected. After the tool is selected, the actual process of data collection is started. In the PM&E framework, it is already detailed who will do the data collection.

Step V: Data Analysis

The next step after the data collection is to analyse this collected data. Unlike in conventional M&E where analysis is usually done by a third party or by funders or implementers of the project, in PM&E, all stakeholders, including the project beneficiaries, are involved in conducting the data analysis. Workshops or joint meetings are conducted in which all stakeholders analyse the data collectively, mull over the problems and constraints in achieving the project objective, look for solutions and examine the results of the project activities.

Step VI: Programme Reporting and Feedback

After the data is analysed and insights derived from it, the last step is to document and report the process and the information generated from the PM&E process. It is also very important to share and disseminate this information to all stakeholders and relevant external audience. Based on the insights derived from this process, feedback and recommendations are given to the project implementation team for further improvement and course correction.

Limitations

Though there are various advantages in conducting PM&E, there are also some practical limitations while applying the same. Conducting a participatory evaluation which fulfils the core objectives requires a lot of commitment and focus from the donors and the implementers. Conducting such an evaluation is very resource intensive, both in terms of time and finance. Moreover, adept professionals are required to facilitate this kind of process.

5.3 Monitoring without Indicators using the Most Significant Change technique

Most Significant Change

The Most Significant Change (MSC) technique is a participatory technique that is used in both monitoring and evaluation. It was invented by Rick Davis in an attempt to meet some of the challenges associated with monitoring and evaluating a complex participatory rural development programme (Davis, 2005). This programme, run by the Christian Commission for Development in Bangladesh (CCDB), had diversity in both implementation and outcomes. MSC is a participatory technique in which all stake holders are first involved in deciding the criteria of change to be recorded, then in collecting these stories of change and lastly, in analysing them. It is essentially a type of monitoring tool as this activity is done through the project cycle to provide feedback to programme managers about the direction of the programme. MSC also contributes to the evaluation as it provides information about the project i.e., whether it is moving towards achieving its intended long term results while commenting on the effectiveness of the programme.

The process involves collection of significant change (SC) stories from the grassroots level. The MSC stories are then shortlisted by the panel of designated stakeholders or the project team. Subsequently, various representatives sit together and read out these SC stories and further discuss the value of these stories.

When should MSC be used?

While learning about a technique, it is of utmost importance to know when the technique should be used and when it should not be used. Effectiveness or suitability of MSC for a programme depends upon the programme context, its objective and its implementation strategy. MSC is suitable for programmes that:

- Are complex and have multiple outcomes
- Aim at creating behaviour and social change
- Struggle with conventional monitoring systems.

MSC is more suitable for monitoring because it is focused on deriving learning for the improvement of the programme rather than just its accountability.

The MSC technique is implemented in the programme using the ten steps formulated by Rick Davis in 2005. They are:

1. Getting started: Establishing champions and getting familiar with the approach. This step includes introducing this technique to the key people and groups and convincing them how useful it can be for the programme while being easy to implement. After all stakeholders have agreed to adopt this technique then the

champions, who will act as facilitators and catalysts to involve others and further facilitate the process, need to be identified.

2. Establishing the domains of change: The second step is to identify the domains of change in the programme, which are broad categories in which SC stories are identified. Unlike in conventional M&E where the indicators need to be SMART, domains of change are left open ended and fuzzy so that different stakeholders can interpret them in their own way.
3. Defining the reporting period: The next step is to define the period or frequency of collecting and reporting these stories. This period could vary from a fortnight to a year depending on the rigour required by the programme team and donors, although quarterly reporting is considered the most common.
4. Collecting stories of change: The next critical step is collecting these stories of SC by asking the respondents what according to them has been the MSC that the project has brought about in the quality of life of its intended beneficiaries. These stories are captured in different ways by fieldworkers during work, through interviews conducted by people designated for the purpose, through group discussions or by encouraging beneficiaries to write their stories directly etc. The key information documented includes the background information of the person who collected the story, the context of the story, the narrative, and why the storyteller thinks this story describes a SC.
5. Reviewing the stories within the organisational hierarchy: The next step is to review and shortlist the most significant stories from the pool of stories collected. The stories are reviewed by a group of people who select the most significant stories amongst them and forward them to their seniors in the hierarchy. It is very important to decide and document the criteria, the process, and the people involved in the process.
6. Providing stakeholders with regular feedback about the review process: As an objective that is imperative in any MLE tool, MSC also aims at providing useful feedback to the project team. Information regarding the MSC observed on the field is fed back into the system. This helps the team in handling the programme to qualitatively understand the perceived significant changes that have been brought into the people's lives. This feedback is provided using different methods like formal reports, newsletters, seminars or through public meetings etc.

7. Verifying the authenticity of stories if necessary: As a lot of stories are collected, there is merit in verifying some of them. Getting details would help to ensure that the reported stories are authentic. Also, if a process of verification is put in place, people who are collecting these stories will be more careful while collecting them. Before verifications are made, it is decided, who will do the verification, what content will be verified, how many verifications will be done etc.
8. Quantification: Though MSC focuses on reporting qualitative change, there is also scope for reporting quantification of change using three methods. The first is at the level of the individual story, in which the number of people involved in that activity, the amount of spent, the amount of money that was saved etc. is considered. The second way of quantification is by quantifying the frequency of instances where this change got selected and the MSC was observed. The third way of quantification is to collect all the SC stories, and amongst them, to keep a count of the number of times a specific type of change was observed.
9. Conducting secondary analysis of the stories: Besides doing the participatory selection and analysis of the stories, additional analysis is done which includes secondary analysis and meta-monitoring of data. Though this is not a mandatory step under MSC, the information generated could be useful in making the process more rigorous. Both these techniques involve analysing the complete set of SC stories. Secondary analysis involves going through the SC stories and then classifying and analysing them. Meta-monitoring involves analysing all the SC stories based on the attributes of the study. This macro level monitoring involves finding out about the number of SC stories written during each reporting period, identifying the people who had written these stories, recognising whether there was any pattern in the stories that got selected and if any action had been taken based on these stories etc.
10. Revising the MSC process: The last step is to revise the MSC process for the next rounds based on the feedback and learnings, while or after the first round has been conducted. The change could be related to any of the steps from 1 to 9 that have been explained above. The change could be related to the stories collected, the frequency at which the stories were reported, it could be associated with the people involved in the story collection, the method of analysis and shortlisting of stories, providing feedback to the

programme based on the MSC process and so on. The aim of reviewing is to improve the MSC post after every round of implementation.

CHAPTER SIX

CHOOSING EVALUATION DESIGNS

This chapter deals with evaluation and its fundamental concepts, and experimental and quasi experimental evaluation designs. At the completion of the chapter, the practitioner will be able to:

- ✍ Describe evaluation and its fundamental concepts
- ✍ Describe experimental evaluation design, its advantages and disadvantages
- ✍ Describe various quasi experimental evaluation designs, their advantages and disadvantages
- ✍ Choose the most suitable evaluation design depending upon the project situation.

Evaluation is defined as a type of research, which measures an effect and attributes the measured effect to a project or intervention. Impact evaluation determines the impact of the programme on the project outcome levels and assesses whether these effects are intended or unintended. Thus, there are two key questions that are to be answered in an evaluation:

1. Is there a change? What is the magnitude of this change?
2. Can this change be attributed to the programme?

Evaluation designs help to answer these questions with some degree of certainty. Evaluations are important to understand and assess whether the programme has achieved its intended goal. Understanding the efficacy of the programme helps policy makers, the government, development funding agencies and other organisations to take decisions regarding expansion, modification or elimination of the programme.

Evaluation designs are determined by the choice of methods used to identify a *counterfactual*, or in other words, a group of comparable non-participants in a project. This counterfactual should be as similar to the target group as possible, apart from the fact that its members do not participate in the project or receive the intervention. The more similar the counterfactual built through the evaluation design is to the target group, the more robust the evaluation becomes in attributing the impact of the intervention to the programme. An estimate of impact can then be derived by comparing the outcome of interest, e.g., HH income, MMR across the counterfactual group and the treatment/intervention group.

Evaluation designs can be broadly classified into experimental, quasi-experimental and non-experimental designs. In the case of experimental evaluations, the counterfactual is known as the *control* group while in the case of quasi experimental evaluations, the counterfactual is known as the *comparison* group.

These evaluation designs vary in feasibility, cost, generalisability of results and the degree of selection bias in selecting the treatment and control groups. Only after understanding the situation at hand and considering the resources at disposal for conducting the evaluation, an appropriate evaluation design should be selected.

Different types of evaluation designs are explained as below:

6.1 Experimental Design

The Experimental design involves gathering a set of individuals or other units of analysis who are equally eligible and willing to participate in the project and then randomly dividing them into two groups: those who will receive the intervention, i.e., the treatment group and those who will not receive the intervention, i.e., the control group.

Experimental or randomised designs are the gold standard of evaluation designs and are the most robust of evaluation methodologies. By randomly allocating the intervention among eligible beneficiaries, the assignment process itself creates comparable treatment and control groups that are statistically equivalent to one another, given appropriate sample sizes (Paul J. Gertler S. M., 2011). This is a very powerful outcome because in theory, the control groups generated through the random assignment process serve as a perfect counterfactual, free from the troublesome selection bias issues that exist in all evaluations. Experimental evaluation design needs to be embedded in the programme design phase as the randomisation has to be in place while deciding on the programme intervention area.

The main benefit of this technique is that it interprets results in a simple way, which means that the project impact on the outcome which is being evaluated can be measured by the difference between the mean of the samples of the treatment group and the control group.

While experimental design is considered the optimum approach towards estimating the project impact, in practice, there are several problems while implementing this design.

1. Randomisation may be unethical owing to the denial of benefits or services to otherwise eligible members of the population for the purposes of the study.
2. It can be politically difficult to provide an intervention to one group and not to the other.
3. The scope of the intervention may rule out the possibility of selecting a control group, as with a nationwide project or policy change.
4. Individuals in treatment or control groups may change certain identifying characteristics during the experiment that could invalidate or contaminate the results. For example, people may move in and out of the treatment or control group. Alternatively, people who were denied a project benefit may seek it through alternative sources, or even those being offered a project may not take up the intervention.
5. It may be difficult to ensure that the assignment is truly random. An example of this might be administrators who exclude high-risk applicants to achieve better results.
6. Experimental designs can be expensive and time consuming in certain situations, particularly in the collection of new data.

The above mentioned challenges exist while conducting evaluations using the Randomised Control Trial (RCT) design. However, as it is the gold standard of evaluation and the most robust of all designs, evaluators prefer to use this design whenever it is feasible.

Example of Evaluation using Cluster RCT:

To illustrate an example of an evaluation design using Cluster RCT, one may take into consideration a programme, in which separate toilets are built for girl students. This programme is implemented in schools where the attendance rate of girls is less than 50 per cent. Based on this criterion, 300 schools are identified in the programme area. Out of these 300 schools which are comparable and eligible for the programme, intervention is provided to 150 randomly selected schools, while the balance 150 schools act as the control group. This evaluation design is called Cluster RCT as the programme is implemented at the cluster level, i.e., at the school. For evaluating this intervention, information regarding the relevant indicators at baseline and endline in both project and control clusters is collected. The impact of the programme is calculated as the double difference across baseline and endline in both project and control clusters.

Pipeline Design

A practitioner can use the experimental Pipeline evaluation design approach for programmes which are supposed to have nationwide or universal coverage, but are rolled out in a phased manner. Pipeline evaluation design helps to exploit the delay in the programme implementation as the areas/groups where the programme is implemented at a later phase can be used as the comparison area. However, here too it is important to ensure that the treatment and control areas are similar and comparable.

Also, it is essential to ensure that there is appropriate gap between the rollout phases so that the project can achieve its expected outcome.

Example of Evaluation using Pipeline design:

Taking the example of a state flagship programme which is to be implemented in 20 low performing similar districts of State X. Due to the scarcity of funds, the programme was initially rolled out in 10 districts and is to be further expanded in the next 10 districts after three years. In such a case, the second phase districts which are like the first phase districts can be used as a control group.

6.2 Quasi-experimental Design

The Quasi-experimental or semi-experimental designs are those in which the allocation to treatment and comparison groups is not randomised. As discussed above, due to various constraints, it is not always feasible to conduct a pure experiment in the form of RCT. Quasi-experimental designs are the best available option for evaluations, in cases where it is not feasible to apply experimental designs. Unlike experimental designs, quasi-experimental designs are used when the programme intervention area is pre-decided. This design consists of various techniques which can be used for constructing a comparison group. The quasi- experimental design used depends upon the type of programme, the stage of programme implementation, data availability and resources available for evaluation. Various quasi-experimental techniques are explained in detail below.

6.2.1 Difference in Difference

The Difference in Differences (DID) design, as its name suggests, compares the change in outcome over time between the treatment group and the comparison group. In a situation where the programme implementation area is pre-decided but the programme has not started, DID can be used as a powerful statistical tool to measure and attribute the impact of the programme to the intervention. The DID design requires two cross sections of data, viz., the baseline or the pre-programme data for both the treatment and the control group and the post-programme data for the treatment and the control group.

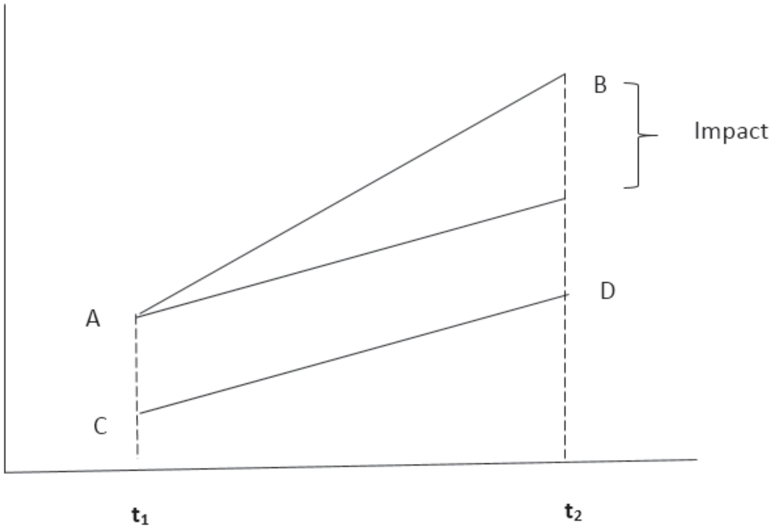

Figure 1.1: Difference in Difference

The figure above illustrates the DID method where A is the value of the output indicator of the treatment group at time t_1 i.e., before the intervention and B is the value of the output indicator of the treatment group at time t_2, i.e., after the intervention. Similarly, C is the value of the output indicator of the comparison group at time t_1 and D is its value at time t_2. DID helps to separate the impact on the output indicator due to time and estimates the change in output variable which can be attributed to the project. The computation of DID is estimated to be (B-A)-(D-C) as shown in the table given below.

Table 1.7: Computation of DID

	After	Before	Difference
Treatment Area	B	A	B-A
Comparison Area	D	C	D-C
Difference	B-D	A-C	DD=(B-A)-(D-C)

Example of Evaluation using the DID design:

Using the example of a programme which aims to improve MCH, the project management, based on some predefined criteria, has already selected the project intervention districts. In such a case, DID is the most suitable technique to measure and attribute the change in outcome indicators of a project. In this case, based on some predefined characteristics like education level, SC/ST population ratio, sex ratio etc., comparison districts are selected which are similar to the intervention districts. Data regarding useful indicators are collected at both the baseline (before the project) and at the endline (after the project), based on which the impact evaluation of the programme is done.

6.2.2 Propensity Score Matching

Matching involves using statistical techniques to construct an artificial comparison group which is comparable in essential characteristics to the treatment group. Both groups should be matched on the basis of observed characteristics that are believed to influence project outcomes. It can be rightly said that matching is as good as the characteristics used for matching (Paul J. Gertler S. M., 2011). This method is applied in the context of almost any programme assignment rule, as long as there is a group that has not participated in the programme. Matched comparison groups can be selected before project implementation (prospective studies) or afterwards (retrospective studies).

The key advantage of evaluations using matching methods is that they draw on existing data sources and are thus often quicker and cheaper to implement. The principal disadvantage is that the reliability of the results is often lower, as the methodology may not completely solve the problem of selection bias; moreover, as the matching methods can be statistically complex, considerable expertise is required in the design of the evaluation and in analysis and interpretation of results.

The most widely used type of matching is propensity score matching (PSM), in which the comparison group is matched to the treatment group based on the propensity score. The *propensity score* is the predicted probability score of each unit for participating/enrolling in the programme. This score varies between zero to one, where zero depicts zero per cent and one depicts one hundred per cent probability of enrolling in a programme. Using this method, a comparison group is found from a sample of non-participants who are closest in terms of observable characteristics to a sample of project participants.

PSM is a very useful method when there are many potential characteristics to match between a sample of project participants and a sample of non-participants. Instead of aiming to ensure that the matched control for each participant has the same value of the control variable X, the same result can be achieved by matching the predicted probability of project participation P, given X, which is the propensity score of X. The range of propensity scores estimated for the treatment group should correspond closely to that of the retained sample of non-participants. The closer the propensity score, the better the match. A good comparison group comes from the same economic environment and is administered the same questionnaire as the treatment group by the same group of trained interviewers.

Example of Evaluation using the PSM design:

To illustrate evaluation using the PSM design, one may take the example of a programme which was started four years ago, and is currently on the verge of completion. The objective of the programme was to create SHGs and work towards giving its members greater access to financial resources. In such cases, where the baseline data is not available, the matching propensity score is used to identify the comparison units for the treatment units. The propensity score is calculated based on variables which are expected to affect the probability of the person becoming a member of the SHG. Variables like the education level, above poverty line (APL)/below poverty line (BPL) status, caste, engagement in agricultural activities, and participation in *gram sabha* activities can be used to calculate the propensity score.

6.2.3 Regression Discontinuity

The *Regression Discontinuity* (RD) design is used to evaluate programmes in which an index or a continuous variable is used to decide the eligibility of participation in a programme. In such programmes, all potential participants are ranked based on a continuous index or variable and the eligible participants are selected based on a cut-off point.

The RD design exploits the discontinuity around the cut-off score in the regression line to estimate the counterfactual. The cases just above and below the cut-off point are similar in their characteristics and the intervention is the main difference between them. Development interventions like subsidy programmes based on poverty score, pension programmes based on age, scholarship programmes based on marks etc., can be evaluated by using the RD design.

Although this design is useful in evaluating programmes where eligibility is decided based on the cut off score, there are a few limitations to this design. This design is less robust and has weak external validity i.e., less generalisability to the population, as only a few units across the cut off line are compared to estimate the impact on the population. Also, in many cases, assignment rules are not strict, thus, a few ineligible beneficiaries can get the intervention while a few beneficiaries who are eligible may not get the intervention. Evaluation of these programmes using the RD design becomes challenging in such cases.

Example of Evaluation using Regression Discontinuity design:

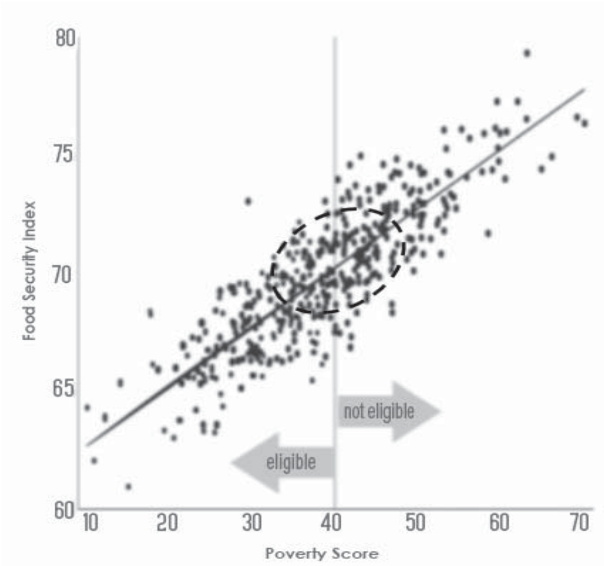

Figure 1.2: Regression Discontinuity - Example

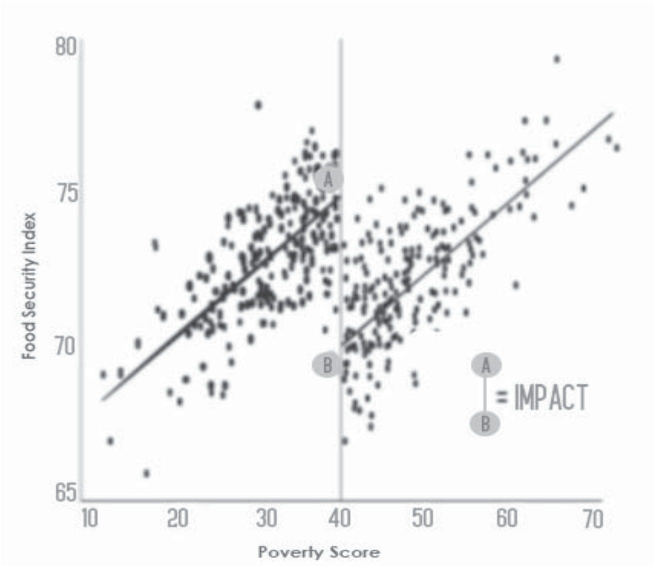

Figure 1.3: Regression Discontinuity - Example

The practitioner may consider the example of a food subsidy programme which aims to improve the food security of its target families having its eligibility criterion based on a poverty score. For this programme, the impact can be evaluated by comparing HHs which are just below and above the poverty line. The figures given above show the distribution of the population with respect to the poverty score and the food security index. All families with a poverty score below 40 are eligible for the programme. The cases just below and above the cut off score (circled in the diagram) can be compared for evaluating the impact of the programme. During the endline survey, the difference in the food security index between the treatment group and the comparison group is assessed. This difference is said to be the attributable impact of the programme.

6.2.4 Instrumental Variable

The Instrumental Variable evaluation design is used to evaluate programmes which have universal coverage and voluntary or open enrolment. For such programmes, the programme administrators do not have control over who will and who will not participate in the programme. Comparing people who have enrolled in the programme and those who did not enrol in the programme is not appropriate for attributing the impact, as both these sets of people are not similar. There are some endogenous factors like motivation, talent, access to information, opportunity cost etc., because of which participants who enrolled in the programme and those who did not enrol in the programme are different.

The *Instrumental Variable* (IV) design is one of the econometric techniques that can be used to compare project participants and non-participants correcting for selection bias (Paul J. Gertler S. M., 2011). It consists of using one or more variables (instruments) that matter to participation but not to outcomes given participation. This identifies the exogenous variation in outcomes attributable to the project, recognising that its placement may not be random but purposive. The instrumental variables are first used to predict project participation, then the project impact is estimated using the predicted values from the first equation.

Randomised Promotion can be used as an instrument variable that allows the practitioner to create variation between units and further use these units to create a valid comparison group. The target population of the programme is classified into three groups. The first is the ‘Never Enrol’ group, i.e., people who will never enrol for the programme whether the programme is promoted amongst them or not. The second is the ‘Always

Enrol' group, which constitutes of people who will enrol in the programme even if the programme is not promoted. The third group is 'Enrol if Promoted', and these are the people who will enrol in the training programme if the training programme is promoted amongst them. The lattermost are the set of participants on which the impact of the programme can be evaluated (Paul J. Gertler S. M., 2011).

The fundamental assumptions in using randomised promotion as the instrument variable is that the promoted and the non-promoted group are similar. The instrument variable used in this case viz., randomised promotion, should be effective enough to increase the enrolment rate substantially amongst the 'Enrol if Promoted' group. Also, it is to be made sure that randomised promotion does not have an effect on the outcome of the programme.

Example of evaluation using the Instrument Variable design:

The practitioner may take the example of evaluation of a voluntary skill development training programme which aims to improve the monthly income of its participants. In this example randomised promotion can be used as an instrument variable to measure and attribute the impact of this programme. A set of similar target participants is chosen to constitute the evaluation sample. The evaluation sample is then randomly divided into two clusters. The programme is promoted in one cluster by an outreach worker who encourages the target participants to enrol in the programme but no promotional activity is done in the other cluster. It is expected that there will be a considerable increase in enrolment to the training programme in the cluster where the programme was promoted.

The impact of the vocational skill training programme is calculated by dividing the change in the monthly income in the promoted and non-promoted cluster divided by the increase in the enrolment rate to the programme. One may assume that the monthly income in the promoted and the non-promoted group is found to be INR 6,000 and INR 5,000 respectively, and that there is an increase in the enrolment rate from 40 per cent to 80 per cent due to the promotion. In this case, the estimated impact will be change in the monthly income divided by the increase in the programme enrolment i.e., 1000 divided by 40 per cent which is equal to INR 2,500. Therefore, there is an increase of INR 2,500 in the monthly income because of the training programme.

6.2.5 Reflexive Comparison

The *Reflexive Comparison design* is another type of a quasi-experimental design. In a reflexive comparison, the counterfactual is constructed on the basis of the situation of project participants before the project. Thus, project participants are compared to themselves before and after the intervention i.e., they function as both treatment and comparison groups.

This type of design is particularly useful in evaluations of full-coverage interventions such as nationwide policies and projects in which the entire population participates and there is no scope for a control group.

There is, however, a major drawback in using the reflexive comparison design. The situation of project participants before and after the intervention may change owing to myriad reasons independent of the project. For example, participants in a training project may have improved their employment prospects after the project. While this improvement may be due to the project, it may also be due to the fact that the economy is recovering from a past crisis and employment is growing again. Unless they are carefully done, reflexive comparisons may not be able to distinguish between the project and other external effects, thus compromising the reliability of results.

6.3 Non-Experimental Designs

Non-Experimental designs are impact evaluation designs that do not include a matched comparison group. These days, an increasing proportion of development interventions fall into categories where it is difficult to apply conventional i.e., experimental and quasi-experimental designs. In situations when the development programmes are complex and it is not possible to define a comparison group, non-experimental designs seem to be potentially the best possible and feasible designs that can be adopted (Richard Blundell, 2000).

In these designs, outcomes and impact are assessed without having a conventional counterfactual to address the attribution question, i.e., "What would have been the situation of the target population if the project had not taken place?"

Creating a counterfactual for a longitudinal design is an example of a non-experimental design. The Longitudinal design is a design in which more than one cross-section of data is collected. A counterfactual can be created in case of longitudinal research design to attribute impact before and after the intervention. In order to analyse the longitudinal/panel data, the fixed effect and random effect model is used to create the counterfactual.

Fixed Effect Model

The *Fixed Effect* (FE) model explores the relationship between predictor and outcome variables within an entity. It assumes that something within the individual may impact or bias the predictor or outcome variables and during the study, the practitioner needs to control for the same.

The FE model removes the effect of those time-invariant characteristics from the predictor variables so the predictor's net effect can be assessed. The equation for the FE model is:

$Yit= \beta 1\ Xit+ \alpha i+ uit$

Random Effect Model

In the case of the *Random Effect model*, the key rationale is that unlike in the FE model, the variation is assumed to be random and uncorrelated with the independent variables included in the model or to quote "…the crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not" (Greene, 2008).

One of the advantages of the random effect method over the FE method is that one can include time invariant variables. The equation is written as:

$Yit= \beta Xit+ \alpha + uit+ \varepsilon\ it$

Fixed Effect or Random Effect Model

To decide whether to use the Fixed or the Random Effect model, the Hausman test can be carried out wherein the null hypothesis is that the preferred model is the random effect vs the alternative, the fixed effect (Greene, 2008).

Example of Non-Experimental Design:

The practitioner may consider the example of a programme aiming to increase awareness about health and hygiene being spread to eliminate OD at the village level. All the villages are covered as part of the intervention and it is not possible to select a control or comparison group. In this case, a cohort design can be used to evaluate the performance of the intervention in which the same HHs are surveyed as part of the baseline and the endline. This is the best possible option to evaluate the project impact considering the current situation.

6.4 Comparison Matrix

As discussed above in detail, each impact evaluation design has its own merits and demerits. It is important for an evaluator to understand the project and the present condition and then consider the availability of critical resources like time, data, and financial resources. Based on these factors, an evaluation expert can choose the best possible design. The table given below gives a detailed comparison between the various impact evaluation designs, when an evaluator should use them and what their key advantages and disadvantages are.

Table 1.8: Comparison Matrix of Evaluation Designs

Design	When to use	Advantages	Disadvantages
Randomised Control Trial	• Whenever feasible	• Gold standard • Most powerful	• Not always feasible • Not always ethical • Costly
Instrument Variable-Randomised Promotion	• When an intervention is universally implemented	• Provides external variation for a sub-set of beneficiaries	• Only looks at the sub-group of the sample • Power of encouragement design only known ex post
Regression Discontinuity	• If an intervention has a clear, sharp assignment rule	• Project beneficiaries often must qualify through established criteria	• IE can be generalised only to a sub-group i.e., cases near the cut-off line • Assignment rule in practice often not implemented strictly

Difference-in-Differences	• If two groups are growing at similar rates • Baseline and follow-up data are available	• Eliminates fixed differences not related to treatment	• Can be biased if trends change • Ideally have two pre-intervention periods of data
Matching	• When other methods are not possible	• Overcomes observed differences between treatment and comparison	• Assumes no unobserved differences (often implausible) • Large sample required
Non-Experimental Design	• When it is possible to select a different control/comparison group	• The only possible option in the case of many real world complex development projects	• Does not have a robust design • Does not control for external factors causing observable changes

6.5 Practice Exercise- Choosing the appropriate Impact Evaluation Design

Project: - *Beti Padhao Yojna* – Improving Educational Outcomes

Background:

According to the 2010 education survey in YY state in XXX country, primary school attendance in the country was estimated at 45 per cent nationally, while it was only 40 per cent for rural areas and 36 per cent for girl children. To help address this problem, the corporate social responsibility (CSR) foundation of XXXX organisation funded a four year, twenty-million-dollar project focused on increasing education outcomes and hygiene conditions in schools that were especially for girls. The *Beti PadhaoYojana* (BPY) programme was implemented in 200 villages across the 10 districts which had the lowest levels of enrolment rates for girls.

It focused on building a school (with separate toilets for girls, and access to drinking water) in each selected village. Also, some additional interventions were introduced in these schools which included:

- *Mid-day Meals*: Free mid-day meals provided to all students daily
- *Menstrual Hygiene Kits:* Free menstrual hygiene kits are provided to adolescent girls
- *Bridge Classes: Bridge classes in the evening for children who had dropped out of school*
- *Take-home Rations*: Girls with high attendance are given food to take home to their families
- *School Kits and Textbooks*: Free material provided to all students

An Impact Evaluation was commissioned to address three key questions:

1. What was the impact of the programme on school enrolment?
2. What was the output and outcome level impacts of the programme?
3. Was the impact on girl students different to the impact on boy students?

Team Tasks for Practitioners

You are an employee of XYZ consulting services which has been contracted as the Impact Evaluation agency for this evaluation. You will be provided a few different scenarios. Please respond to the questions for each scenario given below.

Scenario 1:

It has been decided by the programme team to make inclusive education (IE) an integral part of the programme design. Your team has been called during the design phase to conduct an impact evaluation to answer the evaluation questions. Based on eligibility criteria, 450 villages were identified initially for this programme, though due to a resource constraint, only 200 out of these 450 villages which are to be chosen will be covered under this phase of BPY. Please provide your response to the following questions:

1. What IE design will you recommend?
2. What will be the level of assignment of the design?
3. What could be the practical challenges and disadvantages in implementing this design?
4. What are the strengths of this design?

Scenario 2:

A new CSR head joined the foundation of XXXX organisation six months after the commencement of the BPY project. He appreciates the importance of having a robust IE and wants IE to be conducted for this project. A baseline was not conducted initially, so the programme team is uncertain about the IE design that should be adopted. The CSR head, who has had a positive experience of working with your organisation, requests you to conduct an IE. In your background research, you find that the YY state Education Ministry had conducted a village level survey in the 20 worst performing districts which includes the 10 districts covered by BPY. This data includes characteristics like the number of girls in the village, the enrolment rate of girl students, the distance to the nearest school from the village, the number of students in the nearest school etc. This application data had been used by the implementing agency to rank each village of 450 initially shortlisted villages in terms of need, with the 200 villages ranked as the neediest being selected to receive the programme interventions. Consider the following questions and prepare your responses.

1. Which IE design will you recommend?
2. What will be the level of assignment of the design?
3. What could be the practical challenges in implementing this design?
4. What are the strengths of this design?
5. Can this design be adopted in case the programme team has the option to collect the baseline data? What are the advantages and the disadvantages of applying this design in that case?

Scenario 3:

At the time when the five year BPY programme is about to end, the CSR board realises that they should assess how effective the programme has been. The programme team is left with a limited budget, making it difficult to do large scale primary data collection. However, it is found that the state Ministry of Education has annually been collecting very good village level data in 20 low performing districts, including in the 10 project districts. The data includes important indicators like the enrolment rate of girl and boy students, the number of toilets in the school, and the learning test scores of both girl and boy students. Consider the following questions and prepare your responses.

1. Which IE design will you recommend?
2. What will be the level of assignment of the design?
3. What could be the practical challenges in implementing this design and also, what are the weaknesses of this design?
4. What are the strengths and advantages of this design?

Scenario 4:

The BPY programme is about to start. The board of the CSR foundations is considering conducting an IE for the project. Due to strategic reasons and political pressure, the 200 villages where the programme will be implemented have already been decided. Your organisation is being called to conduct an IE for this project. Please present your responses to the questions mentioned below:

1. Which IE design will you recommend?
2. What would be the practical challenges in implementing this design? What are the weaknesses of this design?
3. What are the strengths and advantages of this design?
4. Would you also be able to use this design if you were called in to evaluate after the programme had already started?

Chapter Seven

Approaches to Evaluation

In this chapter, the different approaches used in evaluations are briefly touched upon. It is very important to understand that there are different approaches to evaluation which are adopted while being integrated with the conventional evaluation methods and designs, as per their suitability to a particular project context. At the end of the chapter various approaches to evaluation will be introduced and described, including:

- Utilisation Focused Evaluation
- Feminist Evaluation
- Transformative Evaluation
- Real World Evaluation
- Using Appreciative Enquiry in Evaluation
- Realistic Evaluation
- Equity Focused Evaluation
- Developmental Approach to Evaluation

7.1 Utilisation Focused Evaluation

Utilisation focused evaluation (UFE) is an evaluation approach where the focus is on its intended users, the 'utilisation' of the evaluation findings and recommendations. This approach, developed by Michael Quinn Patton, is based on the premise that the evaluation should be judged for its worth, based on its usefulness to its intended users. UFE aims to facilitate a learning process through which the users of the evaluation can use it to improve and buttress the programme (Ramirez, 2013). It is just a guiding framework for conducting evaluations and does not prescribe any particular methodology or theory. It includes various evaluation methods, but the key is that they should be implemented in a participatory approach while including the users of the evaluation. The evaluation is planned and implemented in such a way that the users are involved in the designing of the evaluations and it is made sure that the evaluation findings are utilised for the improvement of the project or the programme.

This evaluation approach has two key components. Firstly, the specific and real life primary intended users of the evaluation are identified and engaged from the beginning to design the evaluation and decide the key outputs of the evaluations based on their requirement. Secondly, the decisions about the evaluation process should be made keeping in mind the intended use of the evaluation by its intended users.

The seventeen-step framework listed below, was outlined by M.Q. Patton in the year 2012.It is used as a guide while adopting UFE.

1. Assess and build programme and organisational readiness for UFE
2. Assess and enhance evaluator readiness and competence to undertake a UFE
3. Identify, organise, and engage primary intended users: the personal factor
4. Situation analysis is conducted jointly with primary intended users
5. Identify and prioritize primary intended uses by determining priority purposes
6. Consider and build in-process uses if and as appropriate
7. Focus on priority evaluation questions
8. Check that fundamental areas for evaluation inquiry: implementation, outcomes, and attribution questions are being adequately addressed
9. Determine what intervention model or theory of change is being evaluated
10. Negotiate appropriate methods to generate credible findings that support intended use by intended users
11. Make sure intended users understand potential methods of controversies and their implications
12. Simulate use of findings; this is the evaluation's equivalent of a dress rehearsal
13. Gather data with ongoing attention to use
14. Organise and present the data for interpretation and use by primary intended users, viz., analysis, interpretation, judgment, and recommendations
15. Prepare an evaluation report to facilitate use and disseminate significant findings to expand influence
16. Follow up with primary intended users to facilitate and enhance use
17. Meta-evaluation of use: be accountable, learn, and improve.

7.2 Feminist Evaluation

Feminist evaluation and gender approaches offer researchers methods of evaluation to be applied to particular circumstances. Feminist evaluation is based on feminist theory. According to feminist researchers, a story based on only a man's experience is equivalent to missing half the picture, while adding a woman's perspective enables a researcher to fully construct the picture of reality (M. Connelly, 2000). Feminist evaluation emphasises participatory, empowering, and social justice agendas (Patton M. , 2008). Denise Seigart (2005) suggests that feminist theory and feminist research, influenced by the women's movement, encourages researchers and evaluators 'to question what it means to do research, to question authority, to examine gender issues, to examine the lives of women, and to promote social change'. She further explains that over the years, feminist research has moved from feminist empiricism to standpoint theory, and finally to postmodern feminism. Feminist evaluation is often described as 'fluid, dynamic, and evolving' (Brisolara, 2002). Feminist evaluation theorists tend to describe feminist evaluation as flexible and do not recommend a strict approach or provide a framework; rather, they describe it as a way of thinking about evaluation. Although feminist evaluation follows a dynamic structure, it stands on six basic tenets. Sielbeck-Bowen et al., (2002) defined these six tenets as follows:

1. Feminist evaluation has, as a central focus, the gender inequities that lead to social injustice.
2. Discrimination or inequality based on gender is systemic and structural.
3. Evaluation is a political activity; the contexts in which evaluation operates are politicised and the personal experiences, perspectives, and characteristics that evaluators bring to evaluations (and with which we interact) lead to a particular political stance.
4. Knowledge is a powerful resource that serves an explicit or implicit purpose.
5. Knowledge should be a resource of and for the people who create, hold, and share it. Consequently, the evaluation or the research process can lead to significant negative or positive effects on the people involved in the evaluation/research. Knowledge and values are culturally, socially, and temporally contingent. Knowledge is also filtered through the knower.
6. There are multiple ways of knowing; some ways are more privileged than others.

Difference between Feminist Evaluation and Gender Approaches

Table 1.9: Gender Approach vs Feminist Evaluation

Gender Approaches	Feminist Evaluation
Gender approaches do not challenge women's position in society (Gender and Development (GAD) approach)	Feminist evaluation challenges and attempts to strategically improve women's lives. Feminist evaluation is used to guide the evaluation methodology if the evaluation questions seek to understand *why* differences exist between men and women
A technical fix that leaves the prevailing and unequal power relations intact	A feminist evaluation would take a more activist approach and through its evaluation design, explore the possibility that perhaps all women do not want what men have
Gender approaches provide more concrete guidelines and prescriptive methods	Feminist evaluation offers broad guidance that encourages an evaluator how to think about an evaluation and how to use that reflection

7.3 Transformative Evaluation

Transformative evaluation (TE) is an evaluation approach which was proposed by Donna Martens in the year 2009. Transformative evaluation believes that equity and social justice should be the central values of the evaluation approach. This approach places importance on the lives and experiences of the socially marginalised groups such as women, ethnic minorities, people with disability etc., and advocates that they should be given special focus and priority in this type of evaluation approach (Kosheleva, 2016). This approach also believes that the impact evaluation should be methodologically inclusive. By having a methodology which is inclusive, evaluators can create opportunities for all social groups to participate in the evaluation. This helps in making the evaluation contribute towards increasing social inclusion. This approach is congruent with the International Organization for Cooperation in Evaluation's (IOCE) philosophy on evaluation.

The transformative paradigm believes that evaluators should establish an interactive and trusting relationship with the communities involved in the evaluation so that they understand the realities and the actual condition of these communities and social groups. This approach believes that all social groups and communities have their own strengths and assets and people from socially excluded groups are experts of their own situation and context. Hence, this approach treats people as investors who can provide valuable knowledge about their community and their life experiences, and in this way contribute significantly to the evaluation. Therefore, TE requires evaluators to develop a network with these socially excluded groups which the project or programme evaluates as targets. For its implementation, the TE approach requires the involvement of all stake holders including programme donors, implementers and the intended and unintended beneficiaries of the programme.

There are a few critical points that need to be taken care of while implementing TE. It is suggested that the researchers should have a discussion at the beginning of the planning phase in order to understand and ascertain the background and cultural context of the programme. Also, it is good to use mixed-method designs, as quantitative and qualitative data facilitates the responsiveness of different participants to different issues. Also, the methodology adopted should be able to capture the cultural complexity and it should be appropriate for the cultural groups in the research study area.

7.4 Real World Evaluation

Real world evaluation (RWE) was developed to address the conundrum of conducting a methodological rigorous evaluation under real-world scenarios which face a lot of practical constraints. In real world scenarios, there are constraints like time constraints, budget constraints, and data constraints like non-availability of data or poor quality data. Evaluators are also often required to modify the evaluation design as per political pressures and influences. Also, sometimes evaluations need to be tailored according to organisational systems involving different agencies and actors, as well as administrative procedures that are not well suited to conduct a rigorous evaluation. Very often, evaluators are brought in to conduct the evaluation only towards the closure of the programme. Such situations lead to a poor quality baseline data or no comparison group during the baseline. The RWE approach helps evaluators to adapt to these constraints and come up with the best possible methodology in the context of the programme.

RWE envisages that the highest level of methodological rigour is adopted with respect to the circumstances and the constraints under which the project evaluation is conducted. Michael Bamberger, Jim Rugh and Linda Mabry first authored a book on this evaluation approach which was published in 2006 (Real World Evaluation, 2016). The figure below summarises the seven steps of the RWE approach (Michael Bamberger, 2006) (Real World Evaluation, 2016).

RWE tries to address the quality challenge to ensure that the study has the minimum acceptable methodological rigour and adapts the evaluation design to the level of rigour required by decision makers. One of the key advantages of RWE is that it can lead the increase in the uptake of evidence into policy making as project stakeholders are involved across all phases including design, implementation, analysis and dissemination. This would also ensure that the evaluation focuses on the key issues highlighted by the stakeholders and that it uses the client's preferred style of communication.

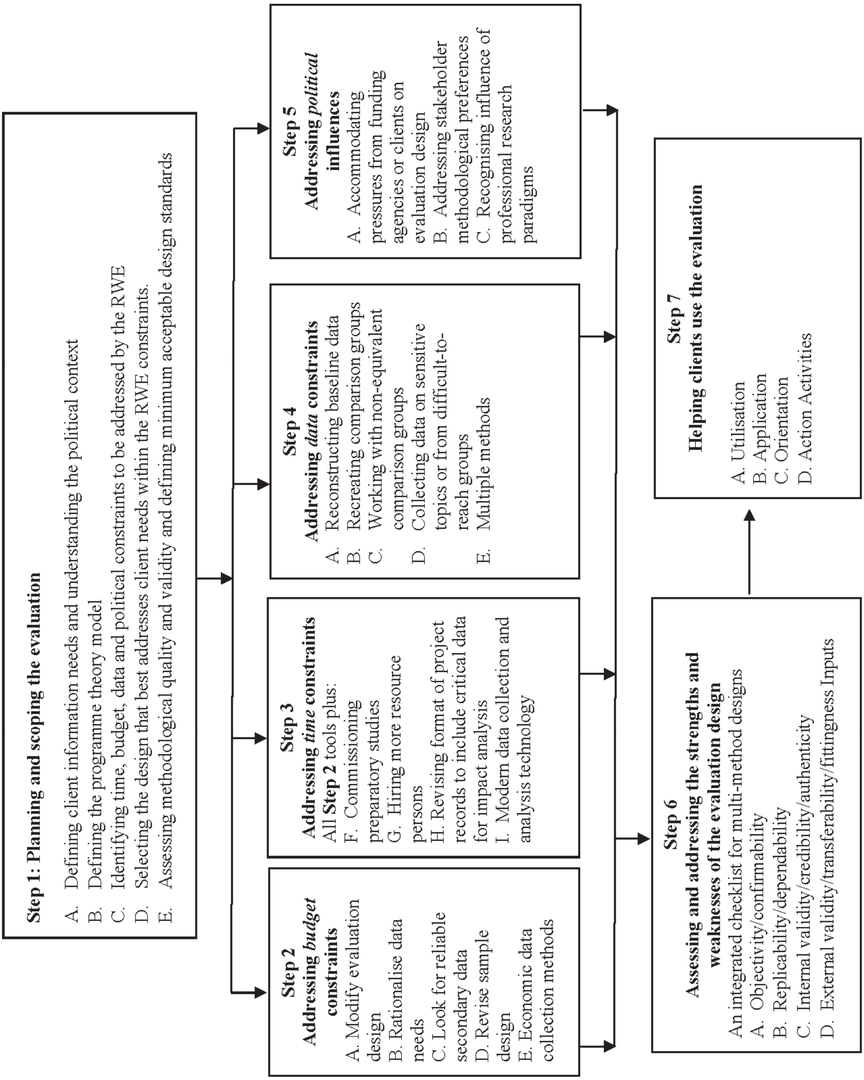

Figure 1.4: Steps in Real World Evaluation

7.5 Evaluation Approach Using Appreciative Inquiry

Appreciative Evaluation is an evaluation approach which is influenced by *Appreciative Inquiry* (AI), which is originally from the field of organisational development. It is a change management tool developed in the 1980s. To explain briefly AI is a process that inquiries into, identifies and further develops the best of what is present in organisations in order to improve them further. This approach aims to focus on the strengths, positive experiences, best practices etc., of the organisation and is implemented using the 4-D model which includes four stages:

- Discovery: Discover what is the best in the present system and identify the system strengths
- Dream: Dream about what can be the best case scenario and envision the probable results in such a case
- Design: Prepare a detailed action plan at this stage to convert dreams or imagination into action
- Delivery: Implement the overall vision and specific activities of the design phase.

Using AI as part of the evaluation approach is a recent development. In terms of evaluation, AI can use the 4-D model to measure change, develop the programme logic model, clarify the evaluation purpose, determine the key evaluation questions, develop indicators and evaluation plans etc. The AI approach can also be used to modify the existing survey instruments or developing new ones to include the AI questions. An evaluation adopting the AI framework would focus on stories of best practices, positive learnings, successful processes etc. This enables the system to look in the direction of success and to look forward to building a future on these positive experiences. An assessment of the programme using the method of AI evaluation seeks answers to what is working well in the programme and how to enhance or replicate those aspects. The AI approach to evaluation is a participatory approach which involves all stakeholders in the evaluation process.

This approach is apposite in cases where previous evaluation methods have failed, the relationship among individuals or various stakeholders have deteriorated, where there is a desire to build the evaluation capacity to help others learn from evaluation practice, and to develop specific data collection methodologies etc.

The advantages of using this approach is that it focuses on developing the positive aspects of the programme that is being evaluated rather than by simply focusing on the problems. Being a participatory technique, it involves various stakeholders and maximises their ownership to the evaluation process and its results. It helps to make the evaluation mode democratic, participatory, deliberative, engaging and enlightening.

7.6 Realist Evaluation

Realist Evaluation is a theory driven evaluation approach. This approach believes that the most durable and practical programme recommendations that evaluators can offer come from research that commences with programme theory and ends with refinement of the same. Therefore, apart from determining whether and how the programme worked in a particular context, realist evaluation also works to test and further refine the programme theory.

This approach was proposed by Ray Pawson and Nick Tilley in 1997. While conventional evaluations aim to answer the question, 'What works?' or 'Does the programme work?' realist evaluation aims to answer the question 'What works for whom in what circumstances and in what respect, and how?' (Tilley, 2004). To understand this, realist evaluation aims to identify the mechanisms and processes and explain how the output was achieved under the influence of the programme milieu.

Realist evaluation believes that the fundamentals of a programme is a theory which can be practically implemented. Every programme is based on a theory and the efficacy of the programme depends upon the efficacy of the theory. In this evaluation approach, various programme theory elements such as mechanisms/processes, programme context, outcomes etc., are made explicit at the evaluation design stage as it enables design of data collection tools for testing of the programme theory.

Realist evaluation is method neutral i.e., it does not impose any particular method. As with other evaluation approaches, realist evaluation too believes that the choice of data collection and analysis methods and tools depends upon the type of data that is needed to answer the evaluation questions and to test the programme theory. Quantitative data is collected to test the various levels of results and the programme context, whereas qualitative data is collected on the processes and generative mechanisms.

This evaluation approach is useful in cases where complex situations need to be assessed. Thus, this approach is more useful to policy makers as it helps them to understand how the intervention works and the conditions or the catalysts that are required to make it work.

7.7 Equity-Focused Evaluation

Before discussing equity-focused evaluations, it is important to first understand what equity is. As defined by the World Health Organization (WHO), *equity* is the absence of avoidable or remediable differences among groups of people, whether these groups are defined socially, economically, demographically, or geographically. Often, the term equity is used as a synonym for equality. However, it is important to recognise that these terms are not synonymous. Equality means sameness or no differences at all, whereas, the aim of equity is not to eliminate all differences so that everyone has the same level of income, health, and education. Rather, the goal is to eliminate the unfair and avoidable circumstances that deprive people of their rights. Therefore, inequities generally arise when certain population groups are unfairly deprived of basic resources that are available to other groups. A disparity is 'unfair' or 'unjust' when its cause is due to the social context, rather than biological factors.

In recent times, there has been increased demand and stronger focus on equity in human and social development. A greater number of organisations working in the development sector including national governments are focusing on achieving equitable development results for the vulnerable or disadvantaged groups amongst the target population. While the focus is on equity focused interventions, there is also the challenge of how to evaluate the effect of interventions on equitable outcomes.

Equity-focused evaluation, as the name suggests, focuses on evaluating projects, policies or programmes which aim to achieve equity focused results. As a standard definition, *equity-focused evaluation* is judgement made on the relevance, effectiveness, efficiency, impact and sustainability of policies, programmes and projects concerned with achieving equitable developmental results (UNICEF Evaluation Office, 2006). This type of evaluation involves going through a systematic and rigorous process of designing, analysing and interpreting what works and what does not work in reducing inequality. It highlights intended and unintended results of the project/programme on the vulnerable or disadvantaged group. It also highlights the gaps between the various strata of beneficiaries viz., 'the best-off group', 'the average group' and 'the worst-off group'.

As the name suggests, the key difference in equity-focused evaluations is that it places a special focus on evaluating the impact of the programme from the perspective of equity. It aims to assess if the programme has had any impact to decrease the inequities, as well as to evaluate the change in the outcome indicators for 'the worst-off group' against the stand-alone index and in comparison, to 'the average group' and 'the best-off group'. In case of equity evaluation, special focus is given to sampling to ensure the representation of disadvantaged or vulnerable groups in the evaluation sample. Also, evaluation questions are explicitly included to assess the impact of the programme on 'the worst-off group'.

Equity-focused evaluation can be used to evaluate the sanitation programme to assess if the programme has been able to ensure participation of the poorest and the most vulnerable groups. The central goal of this evaluation then is to assess the equity outcomes of the project.

7.8 Developmental Approach to Evaluation

Not all projects and the milieu in which they work are simple, controlled or easily predictable at the onset. Sometimes the situation at hand is complex and the project team cannot say with any degree of certainty what they will achieve through the project, although they are confident that by doing their planned activities they will find a way forward. Sometimes these programmes have to be very innovative and dynamic in order to respond to situations, i.e., they can be modified while being implemented or have emerging outcomes. The *developmental evaluation* approach has been developed to respond to such situations. The focus of developmental evaluation is to assist social innovators to develop innovative social change programmes to respond to complex or uncertain situations. In developmental evaluation, the evaluator is part of a team whose members collaborate to conceptualise and design new approaches in the long-term, on-going process of continuous improvement, adaption and intentional change (Gamble, 2008). In this approach, the evaluator’s primary role is to embed evaluative questions, data and logic models in the team discussions and also to facilitate data based decision making in the development process.

Such situations are suitable for developmental evaluation if:

- The project does not have an intervention model yet i.e. it needs to be developed
- The model is not fool proof and needs to be developed or modified
- The programme situation or context is complex i.e., there is no simple cause and effect situation. Most of the important relationships between cause and effect are not clear
- When a rapid response needs to be developed in case of a sudden major change in the project or situation.

Instead of single loop learning where it is first assessed whether the programme is working out or not and then giving feedback to the system, developmental evaluation aims to focus on double loop learning. Double loop learning aims to question the assumptions in the theory model, while it does a deeper assessment of the situation and the strategy. More than problem solving, double loop learning re-evaluates and re-frames the goals, values or assumptions about the strategies that will be successful in a given situation.

The key characteristics of developmental evaluation is that it envisages a tight integration between evaluators and the programme staff. It also aims to use the data concurrently for continuous programme improvement. In this evaluation approach, measures and tracking mechanisms are developed quickly as outcomes emerge. Also, as the process unfolds, these measures can be changed during the evaluation.

Thus, developmental evaluation is suitable in innovative settings where programme goals are evolving and emergent rather than pre-determined and fixed, time periods are fluid and forward looking and not imposed by external deadlines and the objective of evaluation is to facilitate innovation and change and derive learning rather than external accountability.

CHAPTER EIGHT

ANALYSING M&E INFORMATION

The M&E system generates a substantial amount of programmatic data which needs to be analysed and interpreted for programmatic improvement and effective decision making. This chapter explores ways in which M&E data is analysed and interpreted meaningfully for programmatic decision-making. The chapter starts with revisiting data basics and goes on to explore the fundamentals of qualitative and quantitative analysis conceptually and it also provides an overview of doing analysis using statistical software. At the end of this chapter the practitioner will be able to:

- ✍ Understand and explain the fundamentals of data basics
- ✍ Understand and explain the fundamentals of univariate data analysis
- ✍ Understand and explain the fundamentals of bivariate data analysis
- ✍ Use MS-Excel for data analysis
- ✍ Use SPSS for basic and advanced data analysis
- ✍ Understand the fundamental concepts of Qualitative Data Analysis

8.1 Data Basics

At the outset, some of the basics of M&E and data collection and collation, which is done as part of the M&E effort, are revised. The M&E process is an attempt to collect information about variables of interest and assessing change in those variables as a function of the internal and external environment. It is the process of grouping observations about variables of interest in a systematic and coherent way to provide data, which could be qualitative or quantitative in nature depending on the nature and type of variables. Qualitative data is segregated by words, pictures or images, while quantitative data is segregated by numbers on which are based basic mathematical operations.

A variable is defined as an attribute or a characteristic of a case, which is different for different cases. Its variability is usually captured on the measurement scale, varying between two scale values to potentially an infinite number of scale values for the binary scale or continuous metric scale.

Variables can be classified into three categories depending on typology:

1. **Independent Variable**: A variable, which at a particular point of time, can take values on its own. It can also be described as a factor that is selected and manipulated or controlled by the experimenter.
2. **Dependent variable**: A variable whose value is dependent upon one or more independent variable or on other dependent variables. This measurable behaviour is exhibited by the participant in the research.
3. **Extraneous variables**: Extraneous variables are variables that may influence the outcome of a study but are not directly related to the study.

There are four types of measurement scales that are used in the measurement of variables: nominal, ordinal, interval, and ratio scales, which follow a hierarchy, with nominal being at the lowest rung of the hierarchy. These four types of measurement scales are briefly defined below:

1. **Nominal variables**: The values of the nominal variable data have no numeric meaning as no mathematical operation except counting can be done on this type of data. They are in fact used to classify whether the individual items belong to distinctively different categories. Typical examples of nominal variables are gender, race, colour, city, etc.
2. **Ordinal variables**: Ordinal variables, unlike nominal variables, allow the measured items to be ranked in terms of order, where the higher order item that is specified represents more of the quality represented by the variable, though it may not be possible to specify exactly how much more than the other items. A typical example of an ordinal variable is the rating assigned to the impact of the programme, i.e., whether it is excellent, average or poor. Even if the x% rated programme is considered to be excellent and the y% rated programme is average while the one rated z% is poor, the researcher is not able to estimate accurately whether the difference between excellent and average is the same as that between average and poor. In the case of ordinal variables, only certain mathematical variables such as 'greater than' or 'less than'

are feasible and only measures such as median and range can be calculated on this type of data.

3. **Interval variables**: Interval variables provide more flexibility in terms of measurement as they not only allow measured items to be ranked but also help to quantify the sizes of differences between them. For example, temperature that is measured in degrees Fahrenheit or Celsius constitutes an interval scale. Thus, although the temperature of 80 degrees is higher than the temperature of 40 degrees, the temperature at 80 degrees is not twice as hot as 40 degrees. Another example of the interval variable is that of the time system such as B.C. or A.D., wherein time is measured taking B.C. or A.D. as the reference point (initial point of reference is assumed as zero). However, this does not mean that time did not exist before B.C. or A.D. A reference scale to measure time has been constructed, but it does not have a true or rational zero.
4. **Ratio variables**: Ratio variables are measured by scale and they have an equidistant point as well as a rational zero. Thus, in addition to all the properties of interval variables, ratio variables feature an identifiable absolute zero point. A typical example of the ratio scale is the Kelvin temperature scale, where not only is the temperature of 60 degrees higher than the temperature of 20 degrees, but it can be specified that the temperature of 60 degrees is thrice as high as the temperature at 20 degrees. Similarly, an object weighing 20 kg is twice as heavy than an object weighing 10 kg. Most of the variables used for measuring in field situations conform to ratio scale properties, though most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales.

Data Classification

Data, as explained earlier, is a set of observations of variables, characteristics or indicators. Based on the nature, source, values and scale of measurement, data has a different typology. The typology and basis of classification of data are mentioned in the following table.

A broad snapshot of Data Types

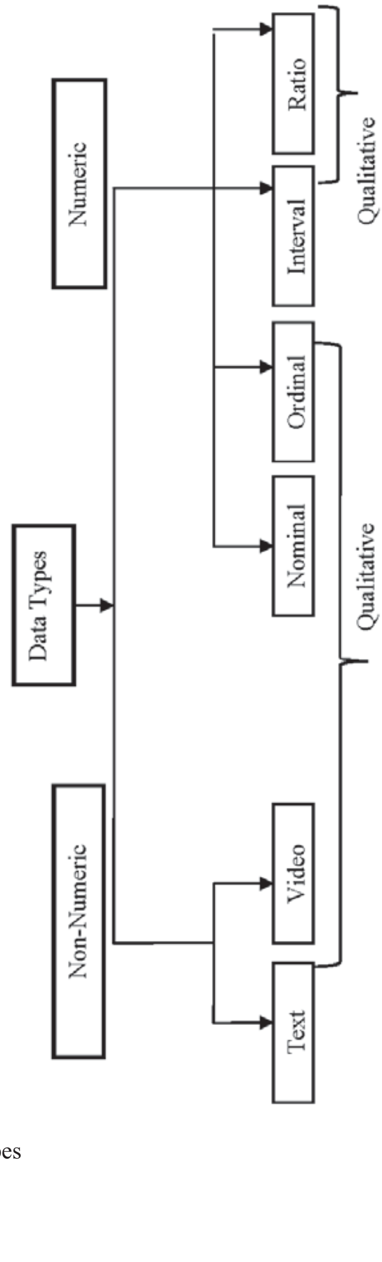

Figure 1.5: Data Types

Table 1.10: Data Classification

Basis	Classification
Nature	**Qualitative**
	Quantitative
Source	**Primary**
	Secondary
Value	**Discrete**
	Continuous

Qualitative Data

Data which describes features or attributes, such as the eye colour of a group of individuals, that cannot be measured or computed by arithmetic relations is termed as qualitative data. They are categorical variables that serve as labels which advise in which category or class an individual, object, or process falls.

Quantitative data

Numerical data which results from a process of measurement is termed quantitative data. Measurement is commonly defined as the assignment of numbers to represent the properties of objects as per rule. Quantitative data can be further classified into metric and non-metric data based on the metric scale that is associated with the measurement of data.

1. Non-Metric Data: The measurement scale used by such data does not possesses a metric scale with which the distance between the scale values can be measured. Nominal and ordinal data falls under this classification.
2. Metric Data: Unlike non-metric data, this data can be measured on a metric scale, thus, the researcher can define distances between the scale values while measuring the data. Metric data can be further divided into two groups:
    - Discrete data: This is countable data which can take values designated by natural numbers, e.g., population, age, etc.
    - Continuous data: Continuous data can take values that can be expressed in rational numbers. This data is measured on a continuous scale, e.g., height, weight etc.

Primary Data

Information collected directly by researchers through first hand investigation for the specific purpose of the research project is termed primary data. Primary data is original data obtained from a primary source which provides direct evidence related to the issue, e.g., – data collected using questionnaires, interviews, focus groups etc., during baseline survey.

Secondary Data

Secondary data is data that has already been collected and is available from other sources. When primary data derived from primary sources is made available to the researcher, it is called secondary data, e.g., data collected from the Census of India, AHS (Annual Health Survey), data that is already published in literature etc. Secondary data is readily available so it saves time, as the researcher is provided with a larger database. It also helps the researcher to identify the gaps and deficiencies in the available data and the additional information that is required.

8.2 Univariate Analysis

When only one variable is being analysed, it is called univariate analysis. The two main types of analyses that are undertaken for a single variable are called measures of central tendency and measures of variability.

8.2.1 Measures of central tendency

Measures of central tendency are descriptive measures that indicate where the centre of the data lies. The goal of descriptive statistics is to describe the data with one single measurement or statistic. This should be an average or typical score if it is to accurately describe or represent the group. There are three measures of central tendency used in different occasions, the mean, the median and the mode. These can be easily calculated and help to simplify the understanding of raw data in a variety of situations.

Mode

The mode is the simplest measure of central tendency; it is the frequency in a data set that occurs for the maximum number of times. In the case of nominal scale data, it is the only analysis that can be done along with percentages.

Median

The median is the middle value in a data set that divides the *ordered* data set exactly in half, with one half of the data being greater than the median and one half being less than the median. The goal here is to determine the precise midpoint of the data set. When the number of observations (N) is an odd number, the median is the middlemost number in the ordered data set. When N is an even number, the median is the average of the two middlemost scores in an ordered data set. One very important property of the median is that it is less sensitive than the mean to extreme scores.

Mean

The mean is the arithmetic *average,* i.e., the sum of all data divided by the number of cases.

Mean = (sum of observations)/ (total number of observations)

The mean is the 'balancing point' of a distribution. If one score is changed, then the balance point shifts (unlike the median, the mean is affected by extreme scores). Since this involves mathematical calculations, it is *only used for interval/ratio* data.

Taking an example of the monthly savings of seven SHGs in a village, and assuming that their monthly saving data is to the tune of 120, 130, 150, 160, 120, 140 and 120. Here the mode is 120 (most occurring frequency as it appears three times), the median is the mid value when the data set is arranged in ascending or descending order (this being 120, 120, 120, 130, 140, 150, 160), therefore, 130, while the mean is 134.3, which is the sum of all values, i.e., 940 divided by the number of observations, i.e., 7.

How is it decided which measure of central tendency should be used?

Often it is best to calculate more than one measure since each measure tells us something different, although all refer to the centre of the data. Generally, the mean is the most preferred measure of central tendency because it uses every score in the distribution. The practitioner needs to remember that the goal is to use one score to best represent the entire data set though it might not always be the best.

When the Mode is the best: Since the mean has to use the interval or the ratio data, the mode is often the best in nominal or ordinal data though it can be used for all types. Moreover, it is very easy to use, since no calculations are necessary.

When the Median is preferred over the Mean: There are several instances when the median is the preferred statistic:

1. When there are extreme scores: An extreme score pulls the mean towards that score, which means that an extremely high score pulls the mean higher, while an extremely low score pulls the mean lower. If, for example, one of the groups in the data given above saved INR 1000, the mean would not be the best way by which to represent the central tendency.
2. If some scores are missing.
3. If one has ordinal data; since this data is discrete, the mean is generally not used.

8.2.2 Measures of variability

A measure of central tendency alone does not suffice to tell the complete story if one is to understand data distribution completely. Unlike the measures of central tendency which capture similarities, a measure of variability is a measure of the dispersion of the scores around the mean, i.e., how spread out is the data, or how different is the data? In the data set given above, measures of variability are required to know the minimum or the maximum savings, the range of the group's savings, and the average deviation from the mean savings etc. The goal here is to determine how much variability is there in the data set or the distribution.

Measures of variability have three main purposes:

1. To describe the distributions (how disperse or variable they are)
2. It gives an idea of how accurately the mean describes the distribution
3. It gives an indication of how well a sample represents the entire population (distribution).

The key measures of variability are:

Range

The range is the difference between the highest and lowest score in the distribution. Expressed as a single number, the range is the result of the following formula:

Range = Highest score - Lowest score

For the above example, Range = 160 – 120 = 40

Range checks are a very straight-forward procedure used on numeric fields and can be a good initial check on the data. In practice, however, people usually list the minimum and the maximum rather than the single number range.

Variance

The variance, s^2, is the mean of the squared deviations:

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{N} \quad \text{or} \quad \frac{SS}{N}$$

where SS = the sum of squares, which refers to the sum of squared deviations from the mean $\Sigma(x - \bar{x})^2$.

The variance detects the differences that the range does not detect because it uses all the data points in the calculation and not just the minimum and the maximum.

The key to variance is the SS since if the SS is large, the variance is large, etc.

The variance is very useful in more sophisticated inferential statistics, but its use in descriptive statistics has one major flaw, as the calculated value is expressed in *squared* units of measurement. Thus, it is seldom used for interpreting descriptive variance.

Standard Deviation

To work around this problem of using the variance, the researcher should go back to the original units, and take the square root

$$\sigma = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n}} \quad \text{or} \quad \sqrt{\frac{SS}{n}}$$

This is the standard deviation, which is the most commonly used measure of variability. It is frequently cited with the mean as the main piece of information about a distribution.

It considers every piece of data in the distribution (as does the mean) as it measures, on an average, how much each piece of data deviates from the mean.

This is done in several steps:

1. Calculate the mean (134.3 in the example)
2. Measure the deviation of each score (X) from the mean ($\overline{X}$)
3. Square the deviations to get rid of the (-) signs
4. Divide the sum of the squares by the number of observations
5. Take the square root of the total value

The standard deviation has been calculated below taking the SHG example discussed above.

Table 1.11: Calculating the Standard Deviation - Example

S. no.	Savings	(X- $\overline{X}$)	(X- $\overline{X}$)^2
1	120	-14.3	**204.49**
2	130	-4.3	**18.49**
3	150	15.7	**246.49**
4	160	25.7	**660.49**
5	120	-14.3	**204.49**
6	140	5.7	**32.49**
7	120	-14.3	**204.49**
	Total	**0**	**1571.43**

If the sum of the square of deviations is divided by the number of observations (1571.43/7), the figure 224.49 is obtained. The square root of this figure is 14.98. This is the standard deviation. Even though the variance provides some idea of deviation, there is still another problem to deal with. The square of the numbers magnifies the numbers considerably and the units are NOT in their original form. Therefore, the square root of the variance is taken, which takes one back to the original units; it makes the correction for having squared the distances from the mean (the deviations) and gives the standard deviation.

$$\text{Standard deviation} = \sqrt{\text{variance}}$$

The standard deviation (and the variance) is very sensitive to extreme scores (as is the mean). Therefore, one needs to be very careful in interpreting the standard deviation or the variance for a distribution that contains even a few very extreme scores.

This should give the practitioner an idea of how much dispersion around the mean there is in a particular distribution. It provides the overall notion of dispersion in a distribution needed to complete the description of the distribution.

8.3 Bivariate analysis: Correlation and Regression

Correlation

Correlation is one of the most widely used measures of association between two or more variables. In its simplest form, it signifies the relationship between two variables, i.e., whether an increase in one variable results in increase in the other variable. The measure of correlation is employed to explore presence or absence of correlation, i.e., whether there is correlation between the variables in an equation. Correlation coefficient also describes the direction of correlation, i.e., whether it is positive or negative. At the same time, it also describes the strength of the correlation, i.e., whether an existing correlation is strong or weak.

There are various measures of correlation between nominal and ordinal data. Pearson product-moment correlation coefficient is a measure of linear association between two interval-ratio variables. The measure, represented by the letter ‘r’ varies from -1 to +1. A zero correlation indicates that there is no correlation between the variables.

A correlation coefficient indicates both types of correlations as well as the strength of the relationship. The coefficient value determines the strength, whereas the sign indicates whether variables change in the same direction or in the opposite direction. A positive correlation indicates that as one variable increases, the other variable also increases in a similar way. A negative correlation signified by a negative sign indicates that there is an inverse relationship between two variables, i.e., an increase in one variable is associated with decrease in the other variable. A zero correlation suggests that there is no systematic relationship between the two variables and any change in one variable is not associated with change in the other variable.

Regression

Regression is one of the most frequently used techniques in social research. It is used in estimating the value of one variable based on the value of the other variable. It does so by finding a line of best-fit using the ordinary least square method. The relation between the variables could be linear or non-linear, correspondingly, the regression equation could also be linear or non-linear.

Regression analysis extends correlation by asserting a direction of causality (from explanatory or independent variable to dependent or endogenous or explained variable) and measuring the influence of the explanatory variable(s) on the explained variable.

By convention, Y is used for the dependent variable, and X for the independent variable. The aim is to find the 'line of best-fit' to the data. Initially, the relationship between Y and X is defined as a linear one.
$Y = a+bX$

In regression equation, 'a' is defined as the intercept and 'b' is known as the regression coefficient. The value of b indicates the change in the dependent variable for every unit change in the independent variable.

Regression coefficient is another widely used measure of association between two interval-ratio variables. Regression coefficient is an asymmetric measure of association, which is why the regression coefficient of the dependent variable on the independent variable is different from the regression coefficient of the independent variable on the dependent variable. Further, depending on the application of the regression method, the practitioner can use an asymmetric measure of association or a symmetric measure of association. However, when the practitioner needs to predict one variable by another variable, then an asymmetric measure is preferred.

8.4 Quantitative Data Analysis using Computer-aided Analytical Software

8.4.1 Performing analysis using MS-Excel

MS-Excel provides a vast range of functions for data management and analysis. If the practitioner knows the functions needed and how to use them, they can enter the functions into a cell and start analysing the data. However, in case they are not sure of which function to use for a particular task or analysis or even how to use any particular function, MS-Excel can assist them with that too.

Finding the Right MS-Excel Function

To identify the functions available in MS-Excel, the user needs to click the on *Formulas* tab on the Ribbon and then click the *Insert Function* button in the *Function Library* group. Alternatively, the user clicks the button on the formula bar. Either approach provides the *Insert Function* dialog box shown in the figure below:

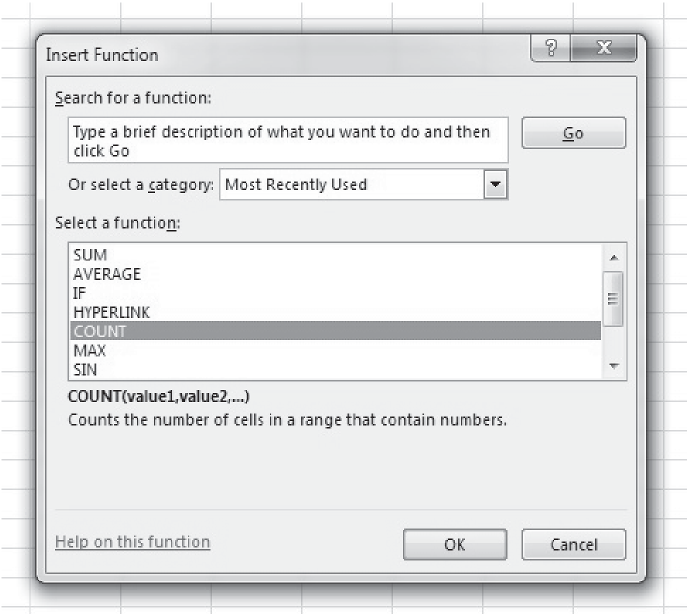

Figure 1.6: Insert Function

The *Search for a function* box at the top of the Insert Function dialog box enables the user to type a brief description of what they want to do. After that, when the *Go* button is clicked, MS-Excel searches for and displays the functions that may accomplish the task in the *Select a function* box. In many situations, however, the user may want to browse through an entire category of functions to see what is available. For this task, the user has the choice of *Or select a category* box which contains a drop-down list of several categories of functions provided by MS-Excel. The figure below shows that the Statistical category is selected. As a result of this, MS-Excel's statistical functions appear in alphabetic order in the Select a function box. The user will find the AVEDEV function listed first, followed by the AVERAGE function, and so on.

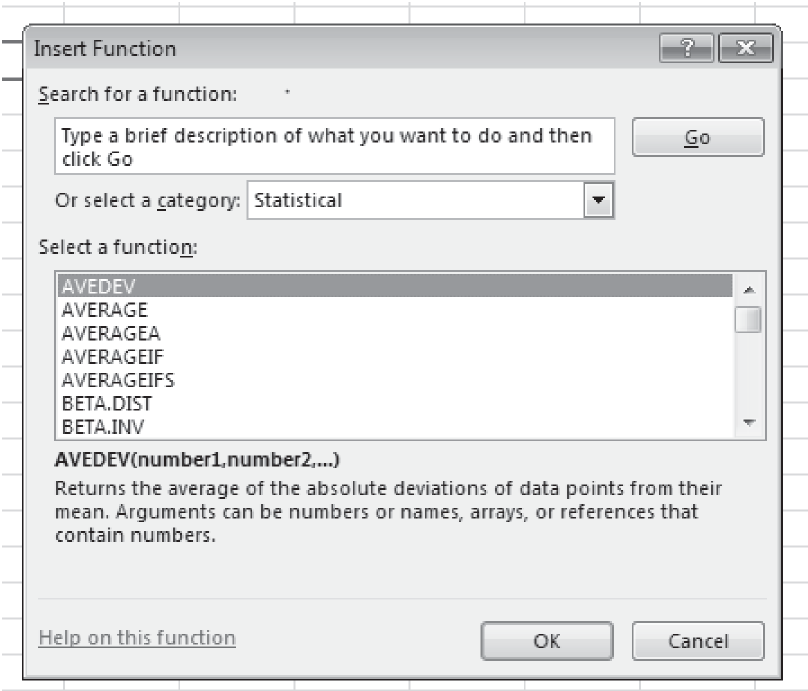

Figure 1.7: Insert Function in Statistical Category

For data analysis, the user can choose out of the key MS-Excel functions displayed. For this purpose, the sample dataset *Health_Data_2013-14.xlsx* is used.

The figure below provides a screenshot of the dataset:

A1 District Name

	A	B	C	D	E
1	District Name	Month of reporting	Total number of pregnant women Registered for ANC	Deliveries conducted at the facility (Including C-Sections)	Number of newborns breast fed within 1 hour of birth
2	District-1	January	8825	1608	2651
3	District-1	February	6697	2336	3630
4	District-1	March	7723	2429	4375
5	District-1	April	12432	3481	5103
6	District-1	May	12095	4781	6772
7	District-1	June	9429	4604	6914
8	District-1	July	7304	3898	7924
9	District-1	August	6056	1029	2637
10	District-1	September	5754	1776	2936
11	District-1	October	7866	2051	3689
12	District-1	November	8203	2676	3933
13	District-1	December	5941	2703	4960
14	District-2	January	6293	3591	4687
15	District-2	February	3258	774	1618

Figure 1.8: Dataset - Example

The first step in any data analysis is to understand the dataset. In the example used, it can be observed that the dataset consists of the following variables:

- District Name
- Month of reporting
- Total number of pregnant women registered for ANC
- Deliveries conducted at facility (including C-sections)
- The number of newborns breastfed within one hour of birth

The data is culled from three districts viz., District-1, District-2 and District-3 and contains 36 cases in all. Each case contains the monthly data for a district and based on these observations, the user analyses the data.

The AVERAGE function is used on the variables listed above to compute the average number of deliveries in a month.

The first step is to select the cell in which the output is to be displayed. The cell next to the last entered value at the end of column D is selected. In this case, the cell number is D38. After this cell has been selected, the user needs to:

1. Go to the Formulas tab in the Ribbon and select the Insert Function
2. Select *All* from the drop-down menu of *Or select a category*
3. Select *AVERAGE* from the list of functions and press OK.

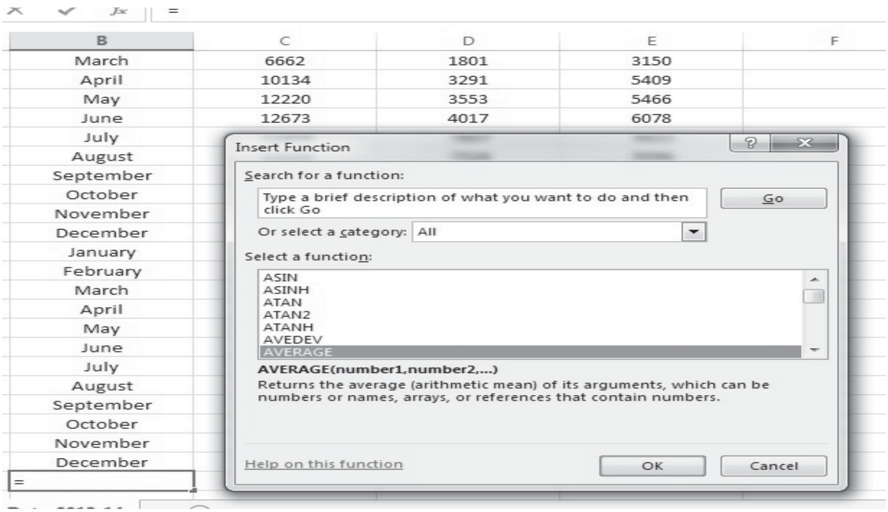

Figure 1.9: AVERAGE Function

The dialog box given above will appear.

The *AVERAGE* function requires the user to give an input of the list of numbers that are used to compute the average. To estimate the average number of deliveries in facilities in a month, the entire range of the variable for institutional deliveries i.e., D2 to D37 is selected. The user needs to click OK to get the average number of deliveries conducted at the facility per month across the three districts, which is 2684.7.

Using the *AVERAGE* function, the average number of deliveries in a month is obtained. However, this average was computed for all the three districts. If, for instance, District-1 had a higher delivery load than the other districts, the estimation of the average number of deliveries is inaccurate.

One way to tackle this issue is to compute the average number of deliveries separately for each of the districts. As in the figure above, the dialog box for the AVERAGE function asks the user to give an input of the range of numbers to be used to calculate the average. To calculate the averages of institutional deliveries separately for the districts, the process is repeated three times, and each time the range of cells belonging to a particular district is specified to calculate the average.

The following are the functions to calculate the average number of deliveries in a month for each of the districts:

Average number of deliveries in a month for District 1 = AVERAGE (D2:D13)

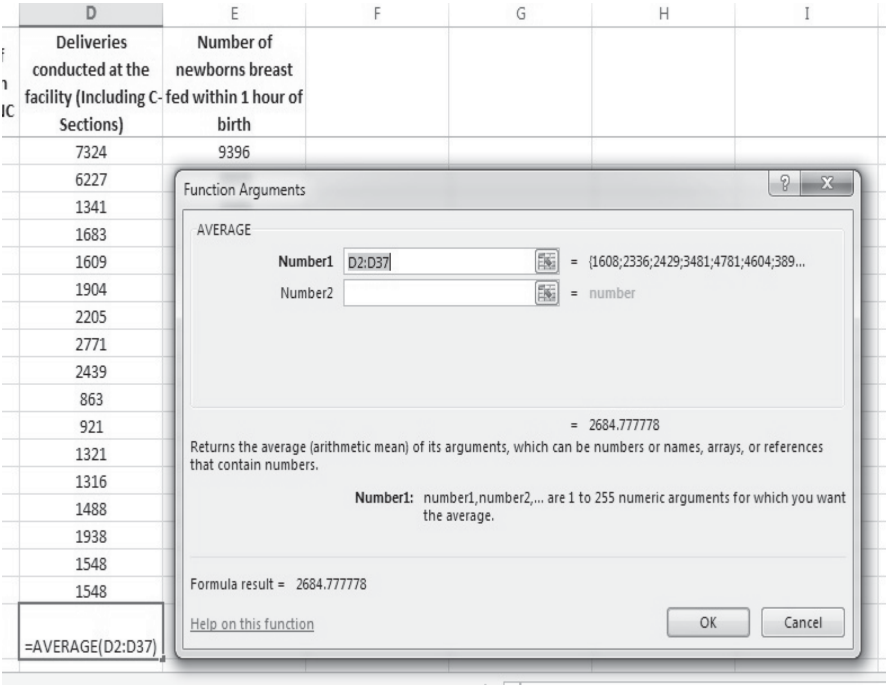

Figure 1.10: AVERAGE Function

Average number of deliveries in a month for District 2 = AVERAGE (D14:D25)

Average number of deliveries in a month for District 3 = AVERAGE (D26:D37)

Though the required average figures may be obtained using this process, several steps are involved in it. An easier way to do this is by using MS-Excel Pivot Tables.

For the process of using MS-Excel functions, the user needs to:

1. Click on the cell from which the output is required
2. Insert the appropriate function using the formula list and specify the range
3. Click on OK

The desired output can be obtained using this process and subsequently the results can be interpreted. Given here is the list of functions that are most frequently used:

Table 1.12: Frequently Used Functions

Function	Use
MAX	Returns the largest value from a list of supplied numbers
MIN	Returns the smallest value from a list of supplied numbers
SUM	Returns the sum of a list of supplied numbers
IF	Tests a user-defined condition and returns one result if the condition is TRUE, and another result if the condition is FALSE
COUNT	Returns the number of numerical values in a supplied set of cells or values
COUNTIF	Returns the number of cells (of a supplied range) that satisfy a given criterion
CORREL	Returns the correlation coefficient between array 1 and array 2 cell ranges
AVERAGE	Returns the Average of a list of supplied numbers
MEDIAN	Returns the Median (the middle value) of a list of supplied numbers
MODE	Returns the Mode (the most frequently occurring value) of a list of supplied numbers *(replaced by MODE.SNGL function in MS-Excel 2010)*
STDEV.S	Returns the standard deviation of a supplied set of values (which represents a sample of a population) *(New in MS-Excel 2010 - replaces STDEV function)*
STDEV.P	Returns the standard deviation of a supplied set of values (which represents an entire population) *(New in MS-Excel 2010 - replaces STDEVP function)*

Using Pivot Tables in MS-Excel

Pivot Tables are one of the most powerful features of MS-Excel. Pivot tables provide an easy way to summarise, analyse, explore and present data. A Pivot Table is a versatile and user friendly reporting tool that makes it easy to extract information from large tables of data without the use of formulae. By moving or pivoting fields of data from one location to another or by using drag and drop, the same data can be seen in several different ways.

To analyse the data using Pivot Tables, the first step for the user is to insert a Pivot Table for the dataset and click on any single cell in the dataset.

The user goes to the insert tab and clicks Pivot Table.

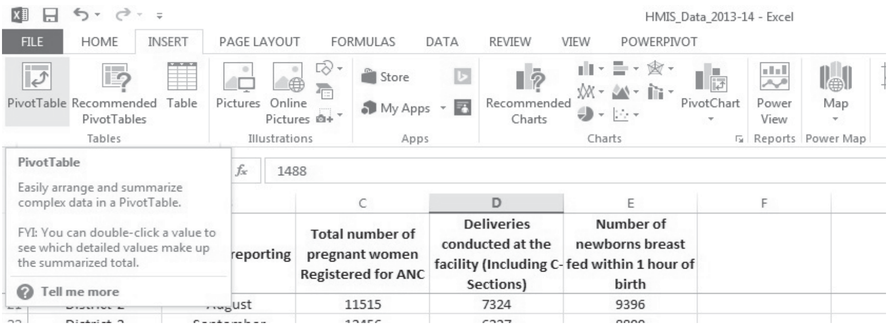

Figure 1.11: Insert Pivot Table

The following dialog box appears. MS-Excel automatically selects the data and the default location for a new pivot table is a New Worksheet.

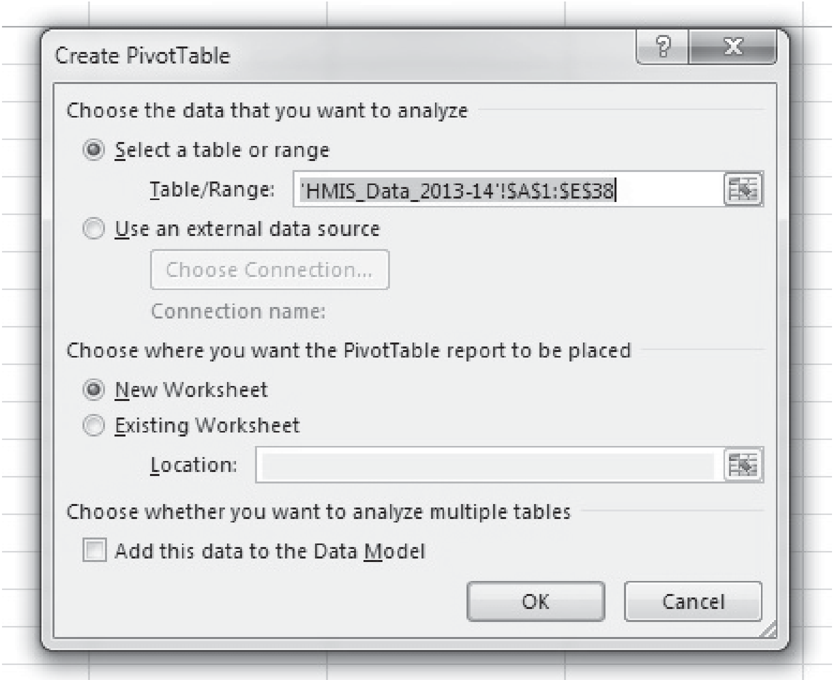

Figure 1.12: Creating a Pivot Table

The Pivot Table Fields will appear on the right side of the screen. For example, to get the total number of women registered for ANC for each district, the user drags the following fields to different areas on the Pivot Table Field:

1. District Name to *Rows* area.
2. Number of women registered for ANC to *Values* area.

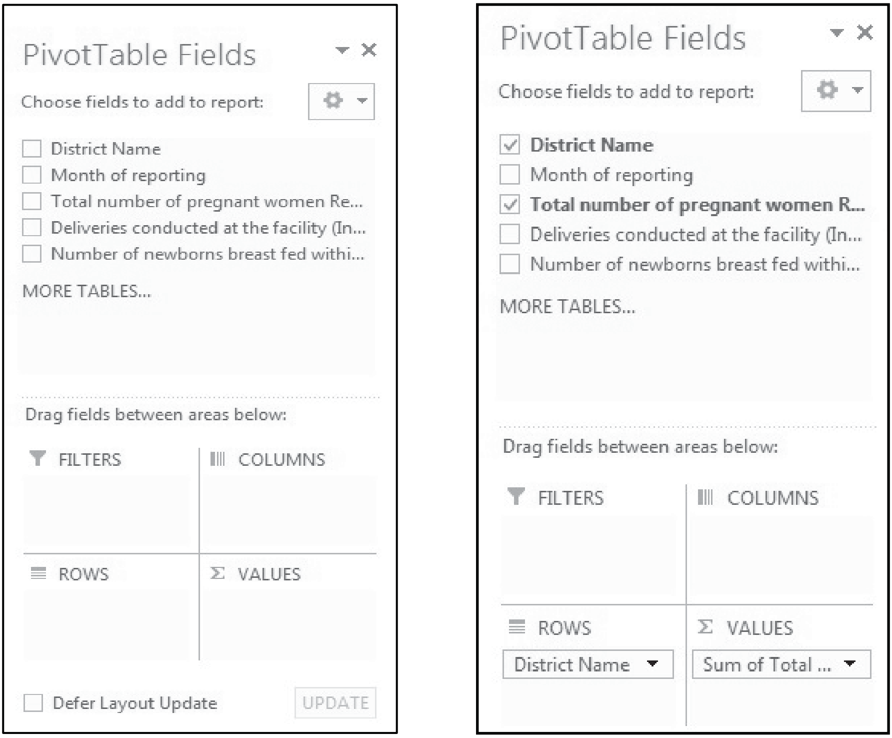

Figure 1.13: Selecting Pivot Table Fields

With these settings, an output showing the total number of pregnant women registered for ANC for each of these three districts is obtained.

Row Labels	Sum of Total number of pregnant women Registered for ANC
District-1	98325
District-2	99244
District-3	42840
Grand Total	**240409**

Figure 1.14: Pivot Table Output

Pivot Tables also provide options to summarise the data using functions such as count, average, standard deviation etc. These options can be chosen from the value field settings in the values area.

To summarise data with the help of the Pivot Tables tool, the user needs to:

1. Insert Pivot Tables for the data set
2. Drag columns in the appropriate fields based on the type of output required
3. Select the appropriate function from value field settings as per the summary required.

Installing the MS-Excel Analysis ToolPak add-in

The Analysis ToolPak is an MS-Excel add-in programme that provides data analysis tools for a variety of statistical analyses. The steps that need to be followed to load the Analysis ToolPak are mentioned below:

Step 1: Click the *Office Button*
Step 2: Click *Excel Options* (on the taskbar at the bottom)
Step 3: When the Excel Options dialog box appears:
- Select *Add-Ins* from the list of options (on the pane on the left)
- In the *Manage* box, select *Excel Add-Ins* (in the second last row)
- Click *Go*

Step 4: When the Add-Ins dialog box appears:
- Select *Analysis ToolPak*
- Click *OK*

The data analysis functions in MS-Excel are ready to be used now.

Calculating descriptive statistics using ToolPak

Suppose the descriptive statistics for the number of children breastfed within one hour of birth from the dataset are to be examined, the user needs to:

1. Go to the Data tab
2. Select *Data Analysis*
3. Wait for the dialog box to appear

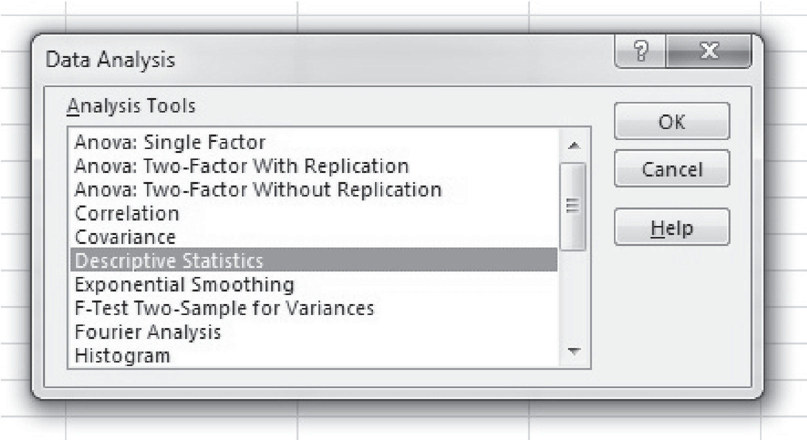

Figure 1.15: Analysis using the ToolPak-Data Analysis

4. Select *Descriptive Statistics* from the list of analysis tools and click *OK*. The following dialog box appears.

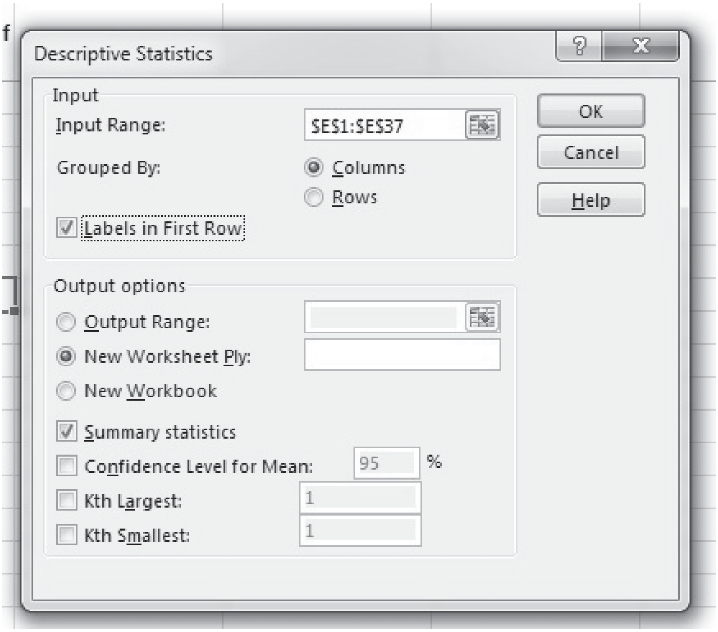

Figure 1.16: Selecting Descriptive Statistics

The user gives the input about the data to be analysed in the *Input Range* box. To analyse the number of children breastfed within one hour of birth, E1:E37 is selected as the *Input Range*. Each column in the dataset represents a variable, the user clicks on the *Columns* in the *Grouped By* checkbox. Also, because Row 1, which includes the variable name is included, the user will click on the *Labels in First Row* checkbox.

Next, under *Output Options,* the user decides where the results are to be written and gives the command for a new sheet by clicking on the *New Worksheet Ply* radio button so that accidentally some of the data is not overwritten.

In the example given above, the *Summary statistics* checkbox is selected.

The user needs to click *OK* for the new results to appear in a New Worksheet.

Calculating correlation using ToolPak

Using the ToolPak add-in programme, the relationship between two variables can be calculated. To see whether there is any relationship between the number of institutional deliveries and the number of children given breast milk within one hour of birth, the user: needs to:

1. Go to *Data* and select *Data Analysis*.
2. Select *Correlation* from the list of analysis tools and tick the checkbox *OK*.

The following dialog box will appear.

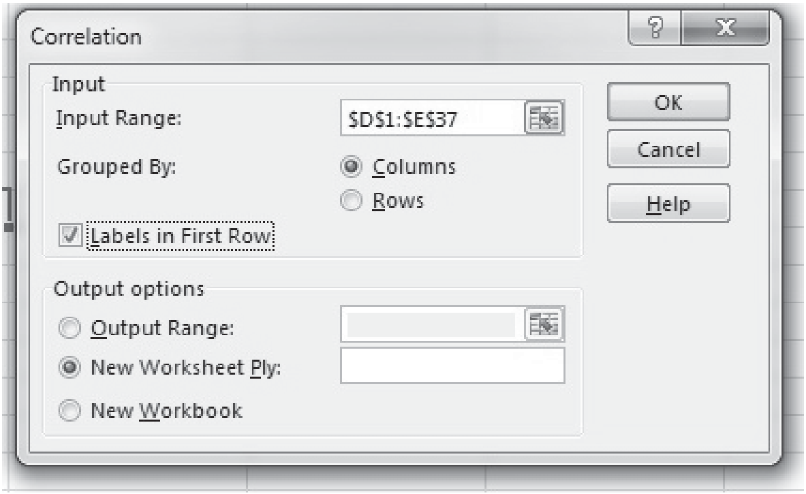

Figure 1.17: Correlation using the Analysis ToolPak programme

The first box is the *Input Range,* where the user gives the input about the variables for which the correlation is to be analysed.

To check the relationship between the number of institutional deliveries and the number of children breastfed within one hour of birth, D1:E37 is selected as the Input Range. In the dataset, each column represents a variable, so, the *Columns* radio button in the *Grouped By* option is clicked. Also, because Row 1 which includes the variable name is included, the Labels *in First Row* checkbox is ticked. Finally, the user clicks *OK* to view the output.

Performing Regression using ToolPak

To check how a single dependent variable is affected by one or more independent variables, the ToolPak is used to perform a linear regression analysis.

In the dataset, to see how early initiation of breastfeeding is affected by institutional deliveries, the user needs to:

1. Go to *Data* and select *Data Analysis*
2. Select *Regression* from the list of analysis tools and click on *OK*

The following dialog box appears:

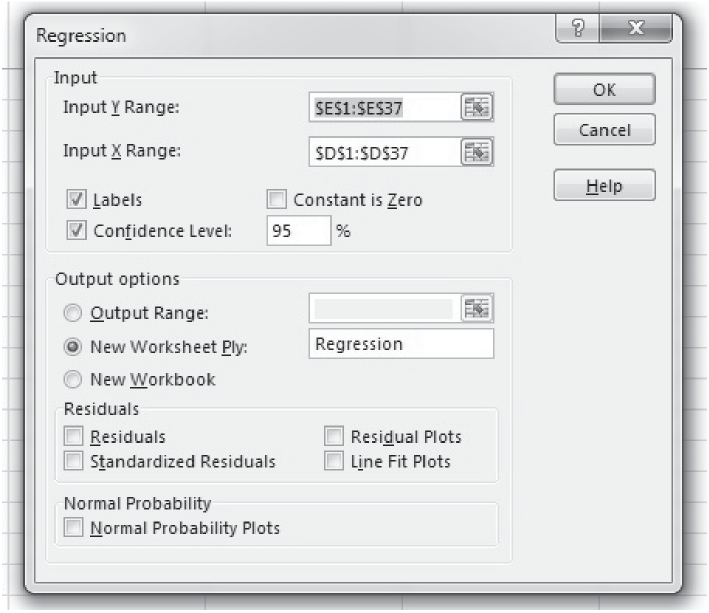

Figure 1.18: Regression using the Analysis ToolPak programme

For the *Input Y Range*, all the values of the dependent variable (the number of children breastfed within one hour of birth) is selected.

In the next step the independent variable (the number of deliveries in the facility) is entered in the *Input X Range* box.

The *Labels* checkbox is ticked since the first row of data is the variable label. Once again, the output is entered on a new spreadsheet.

When the user selects *OK,* the result appears in the New Worksheet.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.954831953							
R Square	0.911704059							
Adjusted R Square	0.909107119							
Standard Error	723.3673942							
Observations	36							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	183700324.5	183700324.5	351.068663	1.70967E-19			
Residual	34	17790853.16	523260.387					
Total	35	201491177.6						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	422.1899738	225.7265258	1.870360482	0.070063012	-36.54151898	880.9214665	-36.54151898	880.9214665
Deliveries conducted at the facility (Including C-Sections)	1.331810629	0.07107984	18.73682639	1.70967E-19	1.187359014	1.476262245	1.187359014	1.476262245

Figure 1.19: Regression Output

Interpretation of the Output

As explained above, the dependent variable in this case is the number of children breastfed within one hour of birth, while the independent variable is the number of deliveries in the facility. The coefficient of the independent variable denotes the size of the effect that this variable has on the dependent variable and the sign on this coefficient denotes the direction of the effect. In this case, it is observed that with the change in each unit of the independent variable, the dependent variable changes by 1.33 times in the positive direction. Also, as the p- value is 1.709E-19, less than this proves that this value is significant. The value of intercept in this case is 422.18 but as its p value .07 is greater than .05, it is not significant.

Therefore, the equation for our model for this example would be:

Y (The number of children breastfed within one hour of birth) = 1.33 X (the number of deliveries in the facility)

Another important result that is derived from this model is the R squared value. The R squared value denotes the fraction of the variance in the dependent variable that is accounted for by the independent variable. In this case, the R square value is 0.91, which suggests that 91 per cent of the variation in the dependent variable is accounted for the independent variable.

The process of using the Analysis ToolPak programme is as follows:

1. Go to Options → Add-Ins and select the Analysis ToolPak programme to be added to MS-Excel
2. Go to Data and select Data Analysis to open the ToolPak programme
3. Select the appropriate analysis option mentioned in the list
4. Select the input range and other parameters to get the output.

8.4.2 Overview of data analysis using Statistical Package for Social Science

Statistical Package for Social Science (SPSS) is the most popular quantitative analysis software used today in social research. SPSS is a comprehensive and flexible statistical analysis and data management software, which can utilise data from almost every type of dataset format to generate tabulated reports, distribution charts, trends, descriptive statistics and complex statistical analyses. Besides this, SPSS is compatible with almost every operating system such as Windows, Macintosh and open sources systems like UNIX or LINUX.

SPSS, unlike other statistical analysis software, provides a window based user interface which makes it very user friendly. Researchers can use simple menus, pop up boxes or dialog boxes to perform complex analyses without writing even a single line of syntax. SPSS is an integrated software which also provides a spreadsheet-like utility function for entering data and browsing the working data file. SPSS and spreadsheet SPSS have an advantage over other prevalent statistical software such as SAS and STATA because of its window driven options. Its operation is pivoted around three windows as mentioned below:

1. The Data Window, with a blank data sheet ready for analysis, is the first window the user encounters. It is used to define and enter data and to perform statistical procedures. Data files have the suffix '. sav' at the end of the file.
2. The Syntax Window is used to keep a record of all commands issued in a session. The researcher does not necessarily have to know the language for writing syntax, instead they can just select the appropriate options from the menu and the dialog box, after which they can select the Paste function option. This command pastes the equivalent syntax of the selected operation in the Syntax window. Beyond serving as a log for operations, it is possible to run commands directly from the syntax window. It is recommended that a syntax file is maintained with the syntax of all data analyses. This ensures regeneration of the required tables and analysis quickly. The syntax files have the suffix '.sps' at the end of the file.
3. Output window: Whenever a procedure is run, the output is directed to a separate window called the Output Window. The tables can be directly copied from the Output Window and be saved in MS-Excel or MS-Word files. Output files have the suffix '. spo' at the end of the file.

Entering Data

The researcher can create a data file simply by entering the data. The present section describes the step-by-step procedure for creating a data file by entering variable information about subjects or cases. The SPSS main data window provides two options on the bottom left hand corner of the screen and the researcher can access either the *Data View* or the *Variable View* window as shown in the figure given below.

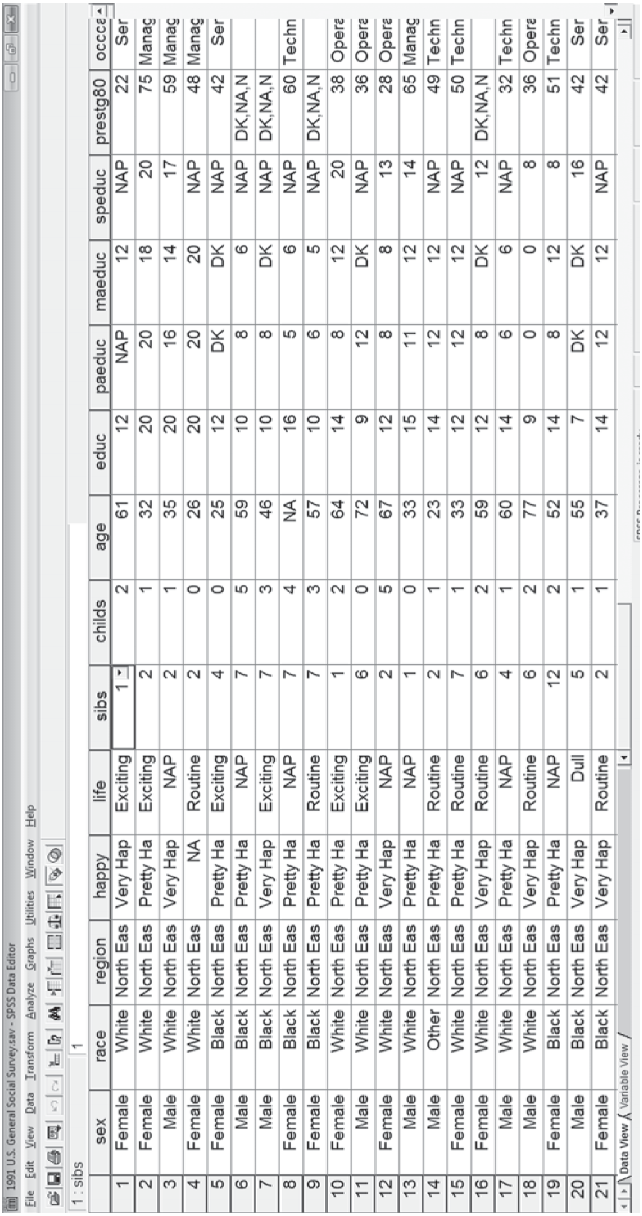

Figure 1.20: Data Editor Window in SPSS

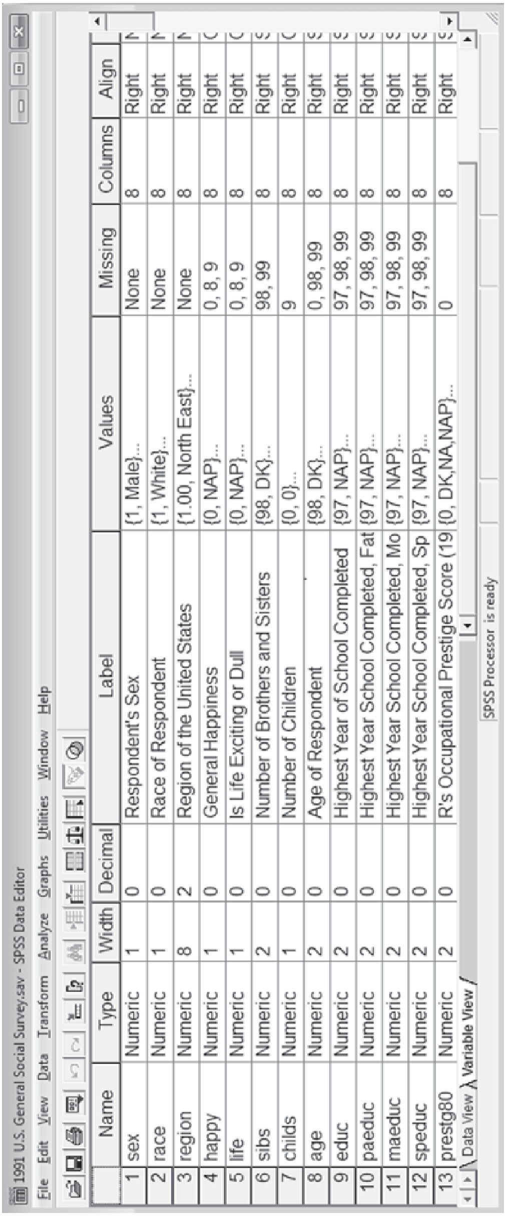

Figure 1.21: Variable View in SPSS

The Variable View window depicts the characteristic of the variable in terms of name, type and nature as mentioned below:

- *Name*: The variable names in SPSS are not case sensitive but they must begin with a letter. Further, the variable name should not exceed eight characters.
- *Type*: SPSS also provides the facility to specify the data type, i.e., whether the data is in numeric, string or date format etc., though the default format is numeric.
- *Labels*: SPSS also provides the facility of attaching labels to a variable name. A variable name is limited to a length of eight characters but by using a variable label, a researcher can use up to 256 characters to attach a label to a variable name. This provides the ability to have very descriptive labels that appear as the output. The researcher can enhance the readability of the output by using the Labels option.
- *Values*: The researcher can assign values, e.g., the male and the female can be coded as 1 and 2, respectively.

Researchers should familiarise themselves with the data characteristic at the initial stage. Though SPSS also provides the facility by which the researcher, at later stages of data analysis, can list information about data nature, type and characteristic by just selecting the data dictionary option which is explained below.

Data Dictionary

A *data dictionary* provides information about the nature, type and characteristics of data. It provides all information i.e., variable name, type, variable label, value labels, missing value definition, and display format for each variable in the data set. In other words, it documents how each variable in the data set is defined.

A data dictionary is easily produced by selecting the File Info option from the Utilities menu. The data dictionary first provides a list of variables on the working file, wherein the variable name appears on the left-hand side and the column number of the variable appears on the right side of the output window. The data dictionary also provides the print and write format after the variable name, followed by special characteristics of the variable such as value labels.

Opening Data Files

Data can be exported in several ways, though the simplest is the Open option. In this case, the requisite database file can be directly opened by simply selecting the ‘Open’ option as displayed in the figure given below. Secondly, data files of other formats (dBase, Access) can be imported to SPSS through database capture. It can also be used to read ASCII data as it is the third option to read files that are saved in ASCII format, which further provides two options - Freefield and Fixed Columns.

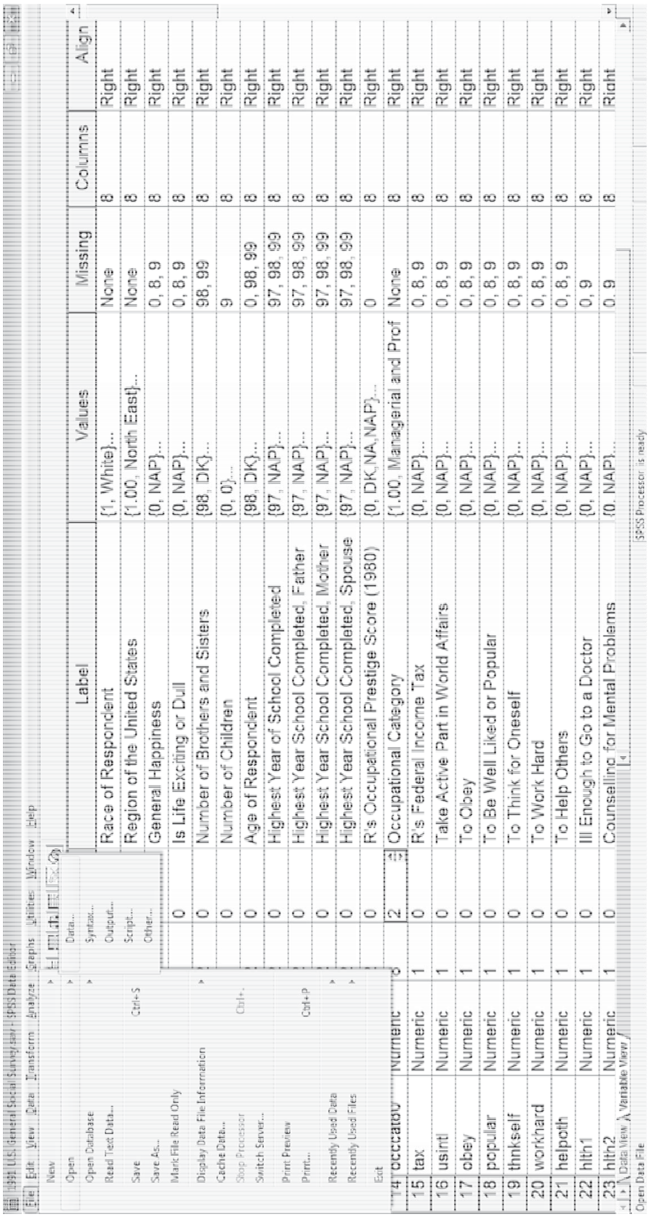

Figure 1.22: Opening Data File in SPSS

a) Opening an SPSS file

An SPSS file can be opened quite easily by clicking on the menu item and selecting the option File -> Open -> Data. The SPSS file type (having a '. sav' extension) can then be selected and the checkbox ticked to open it. Similarly, an SPSS output file can be opened by selecting File -> Open -> Output document files type (having a '.spo' extension).

b) Opening an Excel file

An Excel workbook can be opened by selecting the Read Text Data option by navigating through File->Read Text Data. Excel file types (.xls) can be selected in file type and any excel file can be selected to be opened as shown in the figure given below.

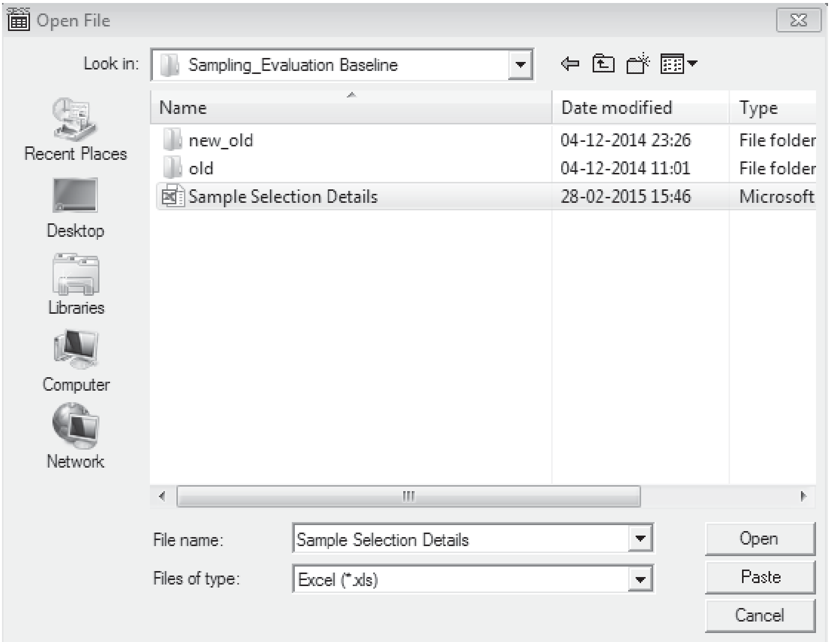

Figure 1.23: Option of Opening an Excel File

After this, the Excel workbook can be selected which can be opened by selecting the option 'Open and Continue'. This opens a screen like the one shown in the figure below.

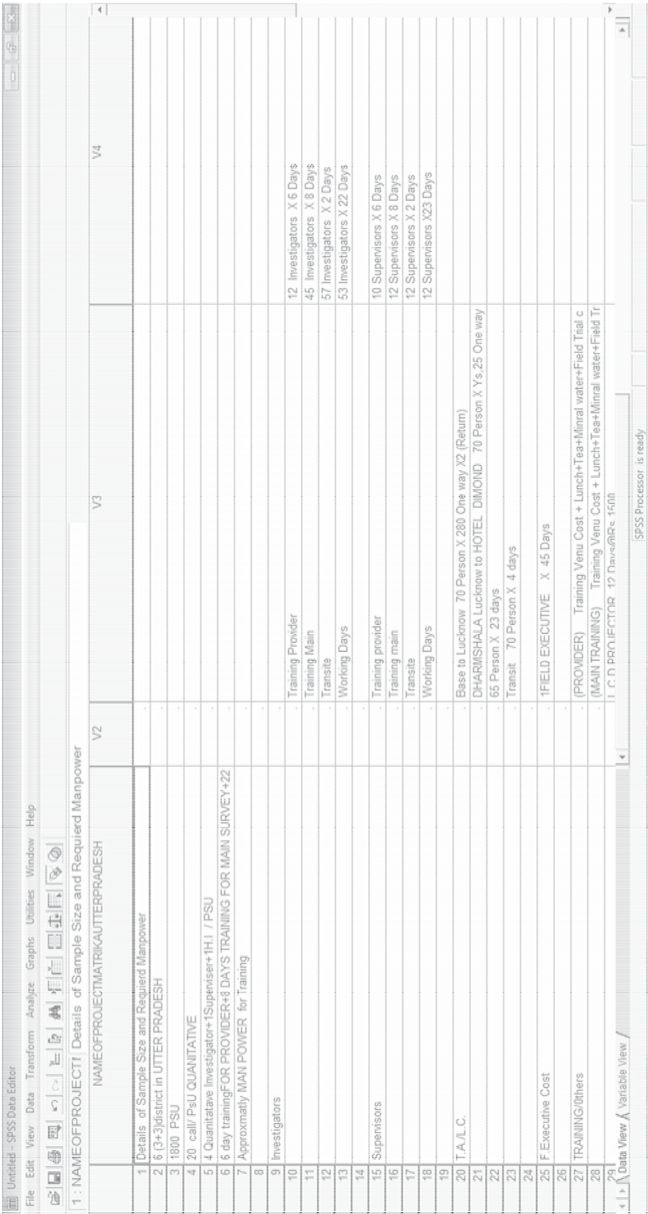

Figure 1.24: MS Excel File Opened in SPSS

It is important to point out that in case the first row of the Excel workbook contains the variable names, then it should be ensured that the checkbox for this option is ticked. In case the data is not on the first worksheet, the worksheet can be changed using the down arrow to select the worksheet which has the data. SPSS also provides the facility to select the range from the worksheet which the researcher wants to use, in case the researcher is not interested in using all the data in the worksheet.

c) Importing Text File

SPSS provides the facility to import the text file through 'Text Wizard'. Text data files can be read in a variety of formats mentioned below:

- Tab-delimited files
- Space-delimited files
- Comma-delimited files
- Fixed-format files.

At the first stage, either a predefined format that can be applied is specified, or the remaining steps in the Text Wizard are followed.

Text Import Wizard at Step 2 requests information about how variables are arranged, i.e., whether the variables are defined by a specific character (spaces, tabs or commas) or whether the variables have a fixed width, where each variable is recorded in the same column. In the next step, it further needs to be seen if the variable names are included at the top of the file as the file may or may not contain variable names. In case the file contains variable names of more than eight characters, then the variable names are truncated. Further, in case the files do not contain variable names, SPSS can allocate default names.

In the third step, the Text Import Wizard requests information about how cases are represented and the number of cases that are to be imported. Usually the first case begins with line 1 if no variable name is supplied and line 2 if the variable name is supplied. Further, each line normally represents a case. In this step, it is also necessary to specify the number of cases the researcher wants to import, i.e., whether the researcher wants to import all the cases, the first n cases, or a random sample.

Step 4 of the Text Import Wizard requests information about the file, i.e., whether it is comma delimited or space delimited. This step, in the case of delimited files, allows selection of the character or symbol used as a delimiter, whereas in case of fixed width files, the step displays vertical lines on the file which can be moved if required.

Step 5 of the Text Import Wizard requests information about variable specifications. This step allows the researcher to control variable names and the data format to read each variable which is included in the data file. Values which contain invalid characters for the selected format are treated as missing.

In the last step, i.e., step 6 of the Text Import Wizard, SPSS provides a facility to save the file specification for later use, or to paste the syntax.

Basic Data Management Function

a) Recoding Variables

Sometimes, while analysing and reporting, fewer categories are required by the researcher than as envisaged in the survey.

In such a situation, the option of recoding variables is used as a way of combining the values of a variable into fewer categories.

For this, there are two options provided by SPSS for recoding of the variable, viz., Recode into Different Variables and Recode into Same Variables. It is strongly recommended that initially the researcher uses only the Recoding into Different Variables option because even if the researcher makes any error, the original variable would still be in the file and the variables can be recoded.

The data can be recoded using several options. A particular value can be changed into a new value by entering the value to be changed into the Old Value box, while the new value is entered into the New Value box.

b) Creating New Variables using Compute

In SPPS, new variables can be created by using the *Compute* command as demonstrated in the figure below. This is done by typing the name of the variable that is to be created in the Target Variable field.

After typing the target variable, the computation that involves the variables in the list is specified to create the **target variable**. The researcher can use all the operations listed at the bottom of the screen, though operations within parentheses are performed first.

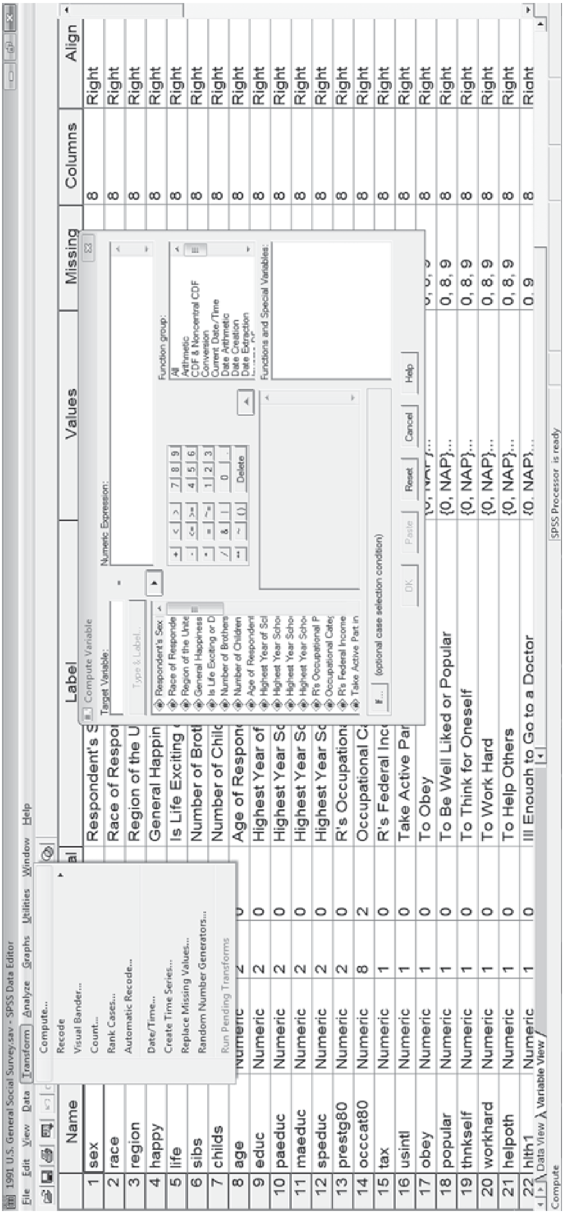

Figure 1.25: Using Compute to Create New Variables

c) Creating New Variables Using If

The *If* command can also be used to create new variables based on the old variables. This can be done by selecting *Transform* and then selecting the *Compute* function. Further, the *Include if Case Satisfies Condition* dialog box should be selected to select subsets of cases using conditional expressions.

There are several operators which can be used, though in the majority of cases, one or more of the six relational operators (<, >, <=, >=, =, and ~=) are used.

d) Spilt File option

SPSS provides the facility of splitting a data file into separate groups for analysis based on the values of one or more grouping variables. In case multiple grouping variables are to be selected, then cases are grouped by each variable within categories of the first grouping variable, based on the groups selected. For example, if Occupation is selected as the first grouping variable and Education as the second grouping variable, then cases will be grouped by Education classification within each Occupation category.

Split a Data File for Analysis

The Data-> Split File option can be very easily chosen from the menu as shown in the figure below. Further, at the next stage, groups are selected to organise output by groups. Before splitting the file, it is important to select the *Split File* option from the data editor window.

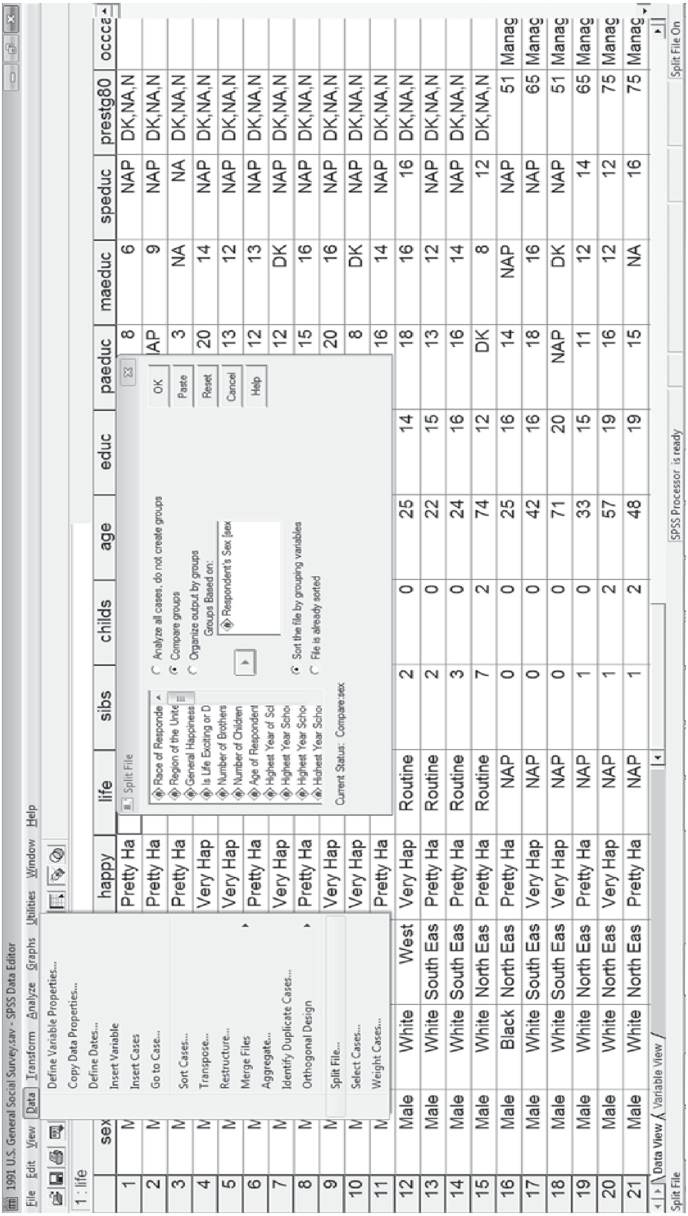

Figure 1.26: Splitting a Data File for Analysis

The Split File option can be turned off by selecting Data and then the Split File option from the Data Editor window by clicking on the *Analyze all cases* option.

e) Using Select Cases

SPSS also provides the facility to select subsets of cases for further analysis. To arrive at this option, first *Data* is selected and then the option *Select Cases*. This opens the Select Cases box and if the researcher wants to select a subset of these cases, they can select the option *If condition is satisfied* to select the subset of cases.

Further, at the bottom of the window, there is a check box to specify that the unselected cases are filtered. By selecting this option, the unselected cases can be used later by clicking on *All Cases*. However, if the *Delete* option is selected, the unselected cases are permanently deleted.

f) Weighting Cases: In particular situations, it may be necessary to weight some cases in the data more heavily than others. For instance, there is a HH in a survey which has an equal probability of selection as other HHs, and there is more than one person eligible in the HH out of which one individual is randomly selected.

To weight the data using this variable just created, the researcher selects Data and then *Weight Cases*. In the next step, the researcher selects the circle to the left of Weight Cases.

g) Missing Values

Missing data can be due to various factors, such as the way in which the interviewer has administered the question, or left certain questions blank, the respondent has declined to respond to certain questions, or due to some human error in data coding and data entry etc. There are a few techniques to handle data with missing values: (1) complete case analysis (list wise deletion), (2) available case methods (pair-wise deletion), and (3) filling in the missing values with estimated scores (imputation).

In case a response is missing, it is recommended that the researcher fixes a pre-determined code for the same, for instance, 999 so that it is understood that the data is missing rather than it being a data entry error. *Missing values* can be coded in the data by recoding the missing values. This can be done by navigating through the following tabs Transform->Recode Into Same Variable. The recoding can be done into the same variable or different variables as required. As shown in the figure below, the *System – missing* radio button can be checked under the old value category and the value can be assigned, which is 999 in this case, after checking the *Value* radio button under the *New Value* category. After this the *Add* button can be checked and the system Missing values can continue to be recoded, as demonstrated in the figure below.

Besides this, the researcher may use various functions and simple arithmetic expressions to extract missing values in different ways.

Thus, the researcher can use the expression AVERAGE in the example (var1+var2+var3)/3 where the result is treated as missing, only if any of the three variables i.e., var1, var2 or var3 has a missing value.

The researcher can use the expression MEAN (var1, var2, var3) where the result is treated as missing only if a case has missing values for all three variables.

SPSS also provides the facility of specifying a statistical function, where the researcher can specify the minimum number of variables that have no missing values, e.g., the function MEAN.2 (var1, var2, var3).

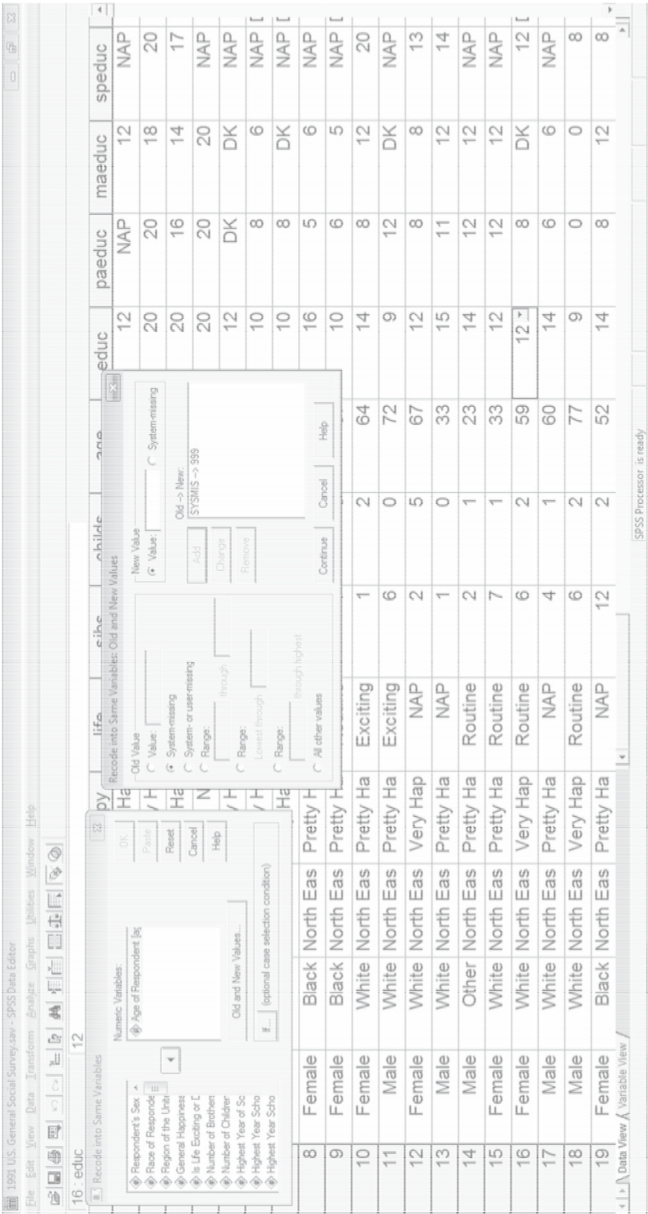

Figure 1.27: Coding Missing Values

h) Reliability Analysis

Reliability signifies the issue of consistency of measures i.e., the ability of a measurement instrument to measure the same thing with consistency each time the instrument is used. Reliability inherently depends on three key elements, such as stability, internal reliability and inter-observer consistency.

SPSS provides the option of doing reliability analysis to assess the additive nature of individual items. This option can be accessed by going to Analyze->Scale->Reliability analysis as demonstrated in the figure below.

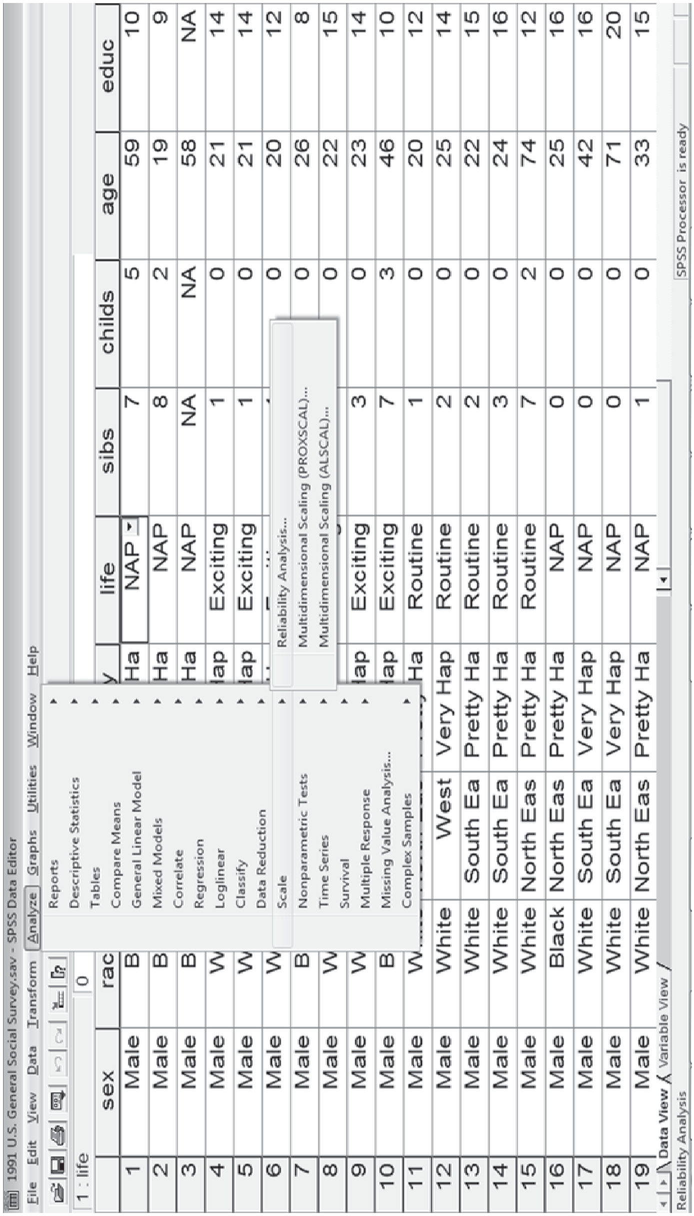

Figure 1.28: Reliability Analysis

As the next step, the selected indicators should be entered as items to assess the reliability of scale. To do so, it is necessary to have an idea about the scale mean, scale variance and cronbach alpha for each item, in case that particular item is to be deleted from the scale. These statistics can be easily computed by selecting the *Descriptives for Scale* option from the Reliability Analysis window as demonstrated in the figure below.

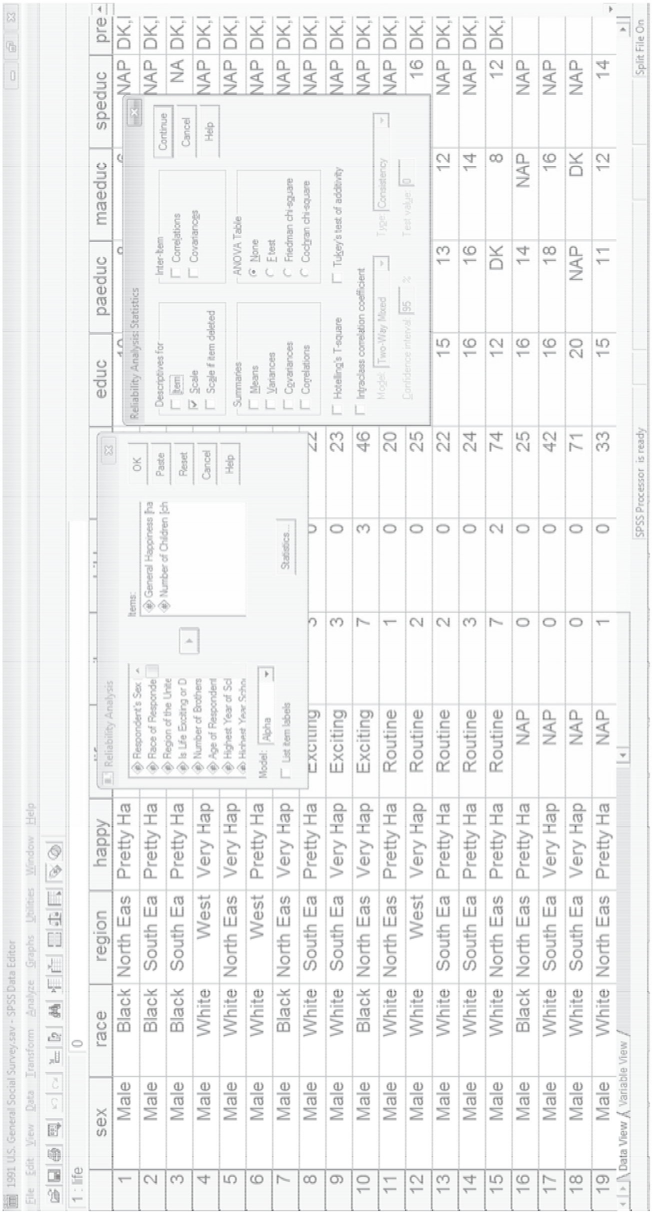

Figure 1.29: Reliability Analysis

The figure above details the reliability coefficients for an awareness scale, which involves eleven awareness variables about the leprosy symptom, having a high value of alpha, i.e., .7672. A rule of thumb for alpha value states that 0.60 is the lower level of acceptability for the alpha, though alphas in the 0.70s and 0.80s range are preferable.

Further, the *Inter-Item Correlations* and Scale if item deleted are also very important indicators to assess the reliability of scale. The *Inter-Item Correlation* option allows the researcher to see if any of the items are negatively correlated with the other items in the scale, while the *Scale if item deleted* option reveals the alpha if each item is deleted from the scale.

Data Analysis

a) Univariate Statistics

Univariate analysis, as its name suggests and explained above in this chapter, provides analytical information about one variable, which could be metric or non-metric in nature. SPSS provides the facility for univariate analysis such as Frequencies, Descriptives and Explore, all of which are located under the Analyze menu as shown in the figure below.

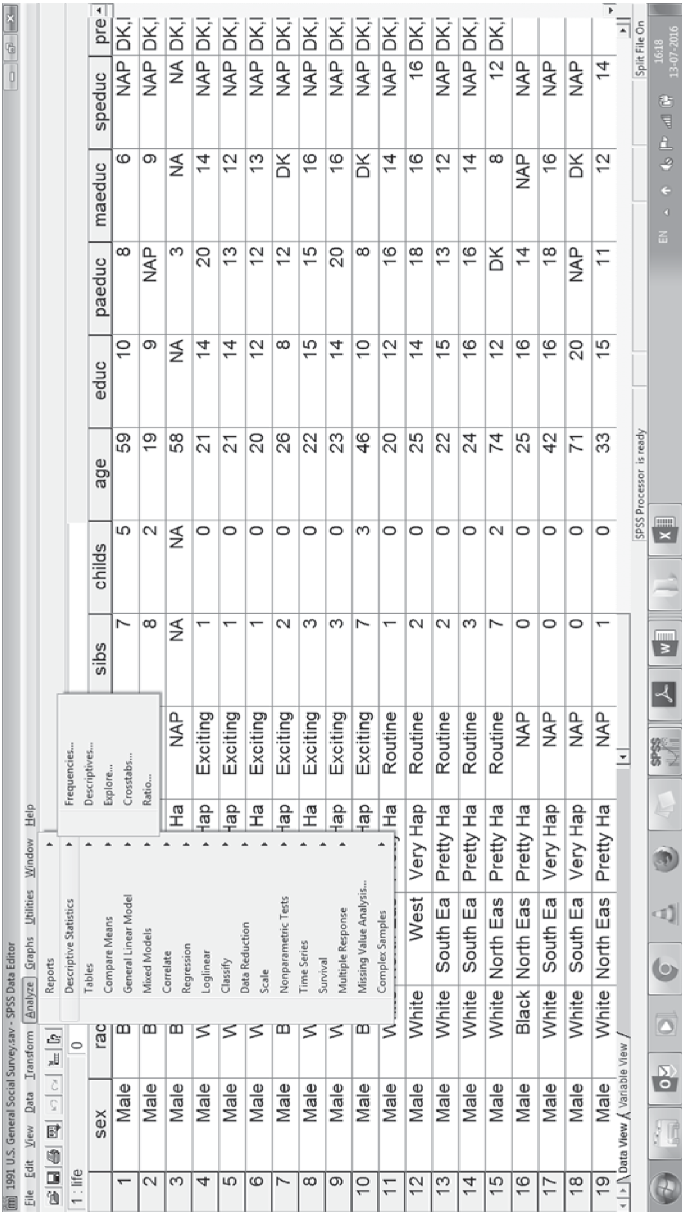

Figure 1.30: Analysing Frequencies

a) If the *Statistics* option at the bottom of the window is selected, this option offers several summary statistics. Based on the options selected, the summarised information gets displayed in the output window. The *Frequencies* option generates the frequency information in addition to the measures of central tendencies and dispersion. Each time when statistical procedures like *Frequencies* and *Descriptives* are selected, the results are immediately displayed in an Output window.

b) Frequencies

The *Frequencies* option can be used to list detailed information of the selected variables. The option is extremely useful for nominal and categorical data such as in the case of gender, wherein data is coded in two categories. The *Frequencies* option provides a table, which shows counts, percentages and statistics including percentile values, central tendency, dispersion and distribution.

c) Descriptives

The SPSS package provides the facility to obtain summary information about the distribution, variability, and central tendency of continuous variables by using the *Descriptives* option. The researcher can select various measures such as mean, sum, standard deviation, variance, range, minimum, maximum, standard error (SE), mean, kurtosis and skewness by using the *Descriptives* option.

d) Explore

The SPSS package provides the option to examine the central tendency and distributional characteristics of continuous variables by using the *Explore* option. The researcher can select statistics such as M estimators, outliers, and percentiles by using the *Explore* option. It also provides options like grouped frequency tables, displays, as well as stem and leaf and box plots.

e) Cross Tabulations

Cross tabulation is one of the easiest ways of summarising the data which can be of any size in terms of rows and columns. It generally allows the researcher to identify relationships between the cross tabulated variables based on their cell values.

The SPSS package provides the facility for generating bivariate cross tabulations. A cross tabulation helps in analysing the association of one variable with another variable and is extremely useful in cases where each variable contains only a few categories.

The crosstab option is found by navigating through Analyze->Descriptive Statistics->Crosstab. Subsequently, the dependent and the independent variables are selected to generate the crosstab option. After selecting the crosstab option, the Statistics button is selected for the Chi-square test to obtain a measure of statistical significance, i.e., Phi and Cramer’s V as shown in the figure below. If both variables are dichotomous, the Phi Correlation Coefficient is used. Cramer’s V is preferred over the Phi Square Coefficient in the case of larger tables.

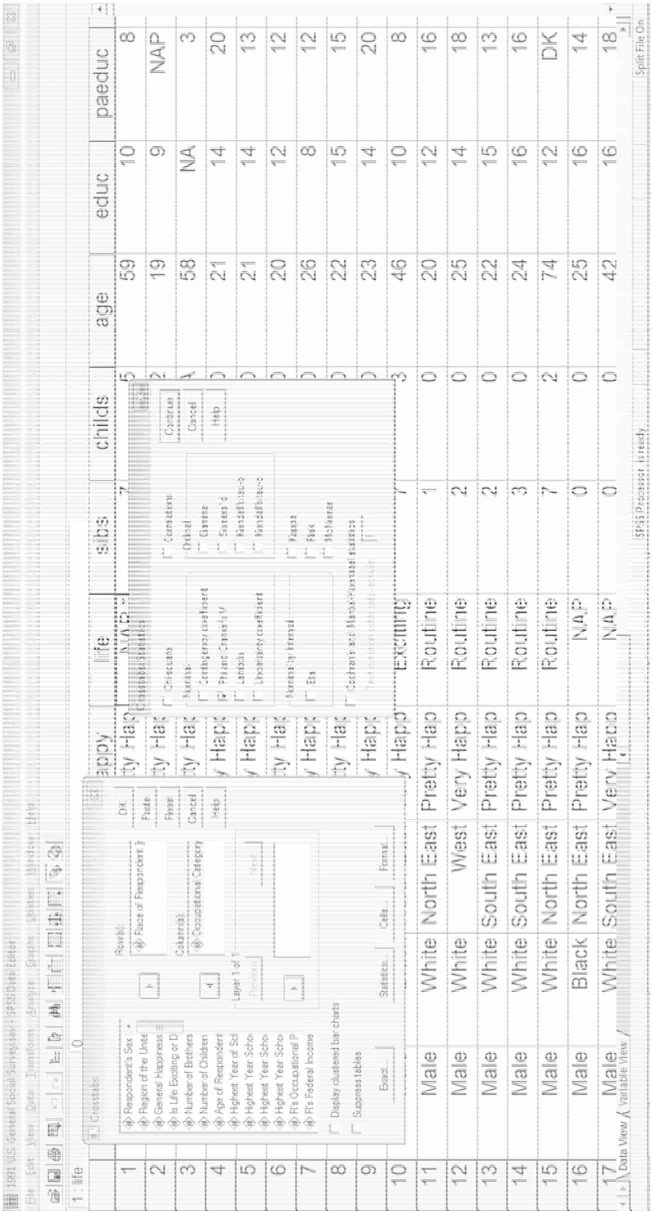

Figure 1.31: Crosstab Statistics

f) The Mean

The mean is a very important measure of the central location of distribution, especially in the case of interval and numerical data. Researchers can use two methods from the SPSS package to produce the mean. One option is to use the *Frequencies* sub option from the *Descriptive Statistics* option.

Alternatively, the *Compare Means* option is selected from the Analyze menu option. Subsequently, the variable which needs to be put in the Independent List box and the one which needs to be put in the Dependent List box can be selected accordingly.

g) The Chi-square Test

The chi-square test is only used with measures, which place cases into categories. This test indicates whether the results from the two measures are about what one would expect if the two were not related. This test can be conducted by selecting Analyze from the top menu and Descriptive Statistics and Crosstabs to open the Crosstabs dialog box. Further, in the Crosstabs dialog box, the variables which the researcher wishes to crosstab can be selected. Further, the researcher can select the Crosstabs: Statistics box option, and then select the Chi-square box option to continue.

h) Independent-Samples t-Test

The Independent-Samples t-test procedure compares the means for two groups of cases. In fact, there are two variants of unpaired t-tests based on the assumption of equal and unequal variances between the two groups of cases. In the case of an unpaired t-test, subjects are randomly assigned to two groups, so that after employing the Significance-test, it can conclude if the difference in response is due to the treatment and not due to other factors.

The Independent-Samples t-Test can be accessed by selecting Analyze and then pointing the mouse at Compare Means and clicking on the Independent-Samples t-Test option. To explain the test further with an example, TV viewership is compared across the gender variable. This test can be done by putting the TV viewership frequency variable into the Test Variable Box and the gender variable in the Grouping Variable Box.

At the next stage, the groups are defined by selecting the *Define Groups* button, which opens the *Define Groups* box. Since males are coded as 1 and females are coded as 2, the researcher types 1 in the Group 1 box and 2 in the Group 2 box.

Table 1.13:Independent Samples t-test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
TV	Equal variances assumed	34.915	.000	-8.4	878	.000	-1.13	.134	-1.390	-.865
	Equal variances no assumed			-8.4	868.501	.000	-1.13	.133	-1.389	-.865

The table above shows the results of the two t tests. The table also gives the values for the degrees of freedom and the observed significance level. It is important to point out that in this test the null hypothesis is that men and women have the same TV viewership frequency. The null hypothesis can be easily tested by comparing the p value with the specified significance level. This significance level value is the probability that the t value would be this big or bigger simply by chance if the null hypothesis were true. Since this probability is less than .05 (the significance level researchers use by convention in social studies), the researcher can reject the null hypothesis and conclude that probably there is a difference between men and women in terms of TV viewership.

i) Paired-Samples t-Test

The *paired-samples t-test* is very similar to the unpaired-samples t-test, except for the difference that the paired-samples t-test is related to matched samples. It tests the difference between the raw scores and assumes that the data is measured in the interval or the ratio scale. This test assumes that the observed data from the matched samples is drawn from a population with a normal distribution.

The test statistic is t with n-1 degrees of freedom. If the p-value associated with t is low (< 0.05), there is evidence to reject the null hypothesis. This proves that there is evidence that there is a difference in the means across paired observations.

The researcher can access the Independent-Samples t-test by selecting Analyze on the menu option and then pointing the mouse at Compare Means and finally selecting the Paired -Samples t-test option.

j) Correlation

Correlation, as described earlier in this chapter, is one of the most widely used measures of association between two or more variables. In its simplest form, it signifies the relationship between two variables i.e., to show whether increase in one variable results in increase in the other variable.

For example, one could hypothesise that as education increases, the level of prestige of one's occupation also increases. The Pearson Correlation Coefficient could tell us the correlation between these two variables. This can be assessed by selecting the Analyze menu option, then Correlate and Bivariate sub-options.

Though there are various measures of correlation between nominal or ordinal data, Pearson product-moment correlation coefficient is a measure of the linear association between two interval-ratio variables. The measure, represented by the letter r, varies from -1 to +1. A zero correlation indicates that there is no correlation between the variables. The SPSS package includes another correlation test, Spearman’s rho, besides Pearson's Correlation Coefficient to analyse variables that are not normally distributed, or ranked. The researcher can opt for both Pearson correlation, and Spearman’s rho. The output screen shows two tables: one for Pearson's Correlation Coefficient, and one for Spearman’s rho.

A correlation coefficient indicates both types of correlation as well as the strength of the relationship. The coefficient value determines the strength, whereas the sign indicates whether the variables change in the same direction or in the opposite direction.

If the coefficient is away from 0, regardless of whether it is positive or negative, it signifies a stronger relationship between the two variables. Thus, a coefficient of .685 is exactly as strong as a coefficient of -.685. The only difference lies in the fact that positive coefficients tell us there is a direct relationship and as one variable increases, the other variable also increases. Negative coefficients indicate that there is an inverse relationship and so when one variable increases, the other one decreases.

k) Regression

Regression is widely used in estimating the value of one variable based on the value of the other variable. This is done by finding a line of best fit using the ordinary least square method. The relation between variables could be linear or non-linear, which would make the regression equation also linear or non-linear.

Further, depending on the number of variables, the regression technique is classified into simple regression and multiple regression. In simple regression, there is one dependent variable and one independent variable, whereas in multiple regression, there is one dependent variable and many independent variables. Beta values (partial regression coefficients) can be computed in case of multiple regression, which gives an idea of the relative impact of each independent variable on the dependent variable.

The output generates the R squared value, which is a summary statistic of the impacts of all the independent variables taken together. It is important to point out that regression assumes that the dependent variable is measured on an interval, continuous scale, though an ordinal scale can also be used. Another condition for multiple regression is to ensure normality i.e., the distribution of all the variables should be normal.

Regression analysis in SPPS can be accessed by selecting the menu option Analyze-> Regression-> Linear. Subsequently, the dependent variable can be selected from the variable list. After the dependent variable is selected from the list, the researcher selects *Continue* and then *Options* as shown in the figure below, which shows the default options.

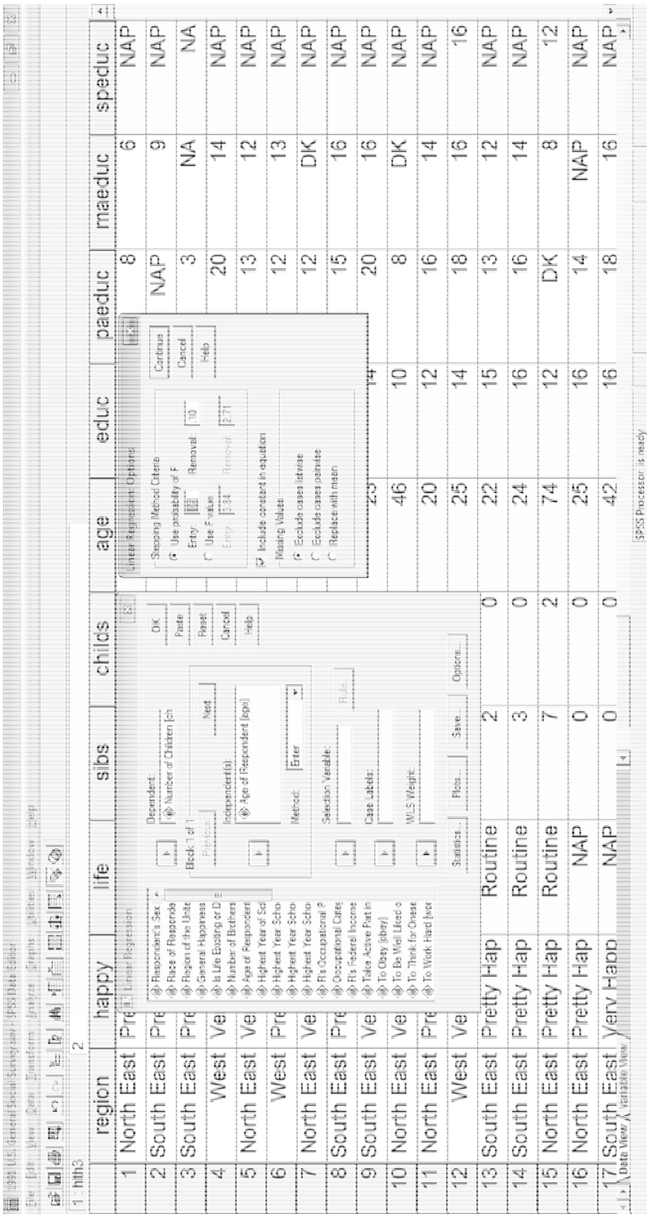

Figure 1.32: Regression

At the next stage, the method of data analysis is selected by clicking on the *Method* button which is right below the *Independent*(s): box. The SPSS statistics package provides several choices for doing regression analysis though this is the one which is used most frequently.

Graphs

Graphical representation of data is a more effective medium for representing data, not only because of its visual appeal but also for its ease of interpretation by users. There are various ways in which data can be represented, viz., by using bar graphs, line graphs and pie graphs. SPSS also, like other statistical software, offers graphical presentations in various formats like bar graphs, stacked bar graphs, pie charts and line graphs etc. However, not all graphic formats are appropriate for all types of data. Bar graphs and pie charts are suitable for all nominal, ordinal and ratio data types, whereas line graphs are appropriate to show time series data, i.e., how variables vary over a continuous period.

a) Bar graphs

To select the bar graph option, first *Graphs* is selected from the menu, after which the *Bar* option is selected, which includes the options of *Simple*, *Clustered* or *Stacked* bar graphs. After this, the independent variable is placed on the vertical axis into the category box, while the dependent variable is placed on the horizontal axis into the Y box.

b) Pie charts

The pie chart sub-option can be selected by going to the '*Graph*' tab through the menu bar. After selecting pie chart as the graphical representation option, the variable for which the pie chart needs to be created can be selected.

c) Line graphs

In the same way, the line graph chart option can also be selected from the *Graph* tab in the menu bar option. SPSS provides the option of creating different types of line graphs like simple line graph, multiple line graph or drop line graph.

In the data view in SPSS, if the grey column heading is clicked on, the whole column gets selected or highlighted. Similarly, if the grey row number is clicked on, the whole row gets selected.

Writing Syntax

One of the key advantages of SPSS is that it provides both options of performing analysis using Windows menu driven selection and by using the programming option to write syntax codes. SPSS saves the syntax file with the extension .sps and is the source of the commands it runs. Any file can be opened with the extension .sps by going to File->Open Menu path, or new syntax files can be created with File->New. Commands typed or pasted into a syntax file.

Directly Editing Syntax Files

Syntax can also be directly edited into syntax files in much simpler terms than menu-generated syntax. Users usually find it more efficient to program by directly typing in the commands.

Like any other language, programming in SPSS is subject to a specific set of rules. First and foremost, programming in SPSS is not case sensitive and a command begins with the Command name keyword. Each command ends with '.', which is the default terminator. SPSS subcommands have two delimiters: the '=' sign distinguishes additional specifications which follow the subcommand and the '/' slash distinguishes one subcommand from another. If the researcher wishes to type in a comment, the researcher can use the asterisk '*'.

It is suggested that the syntax should be saved in syntax files for analysis and tabulations done on a data set. This is very helpful as the analysis and tables can be regenerated instantly multiple times using this syntax. It is very easy to run the syntax commands. To run any syntax command, the appropriate syntax which needs to be run in the Syntax Editor Window is highlighted or selected. After the Syntax Editor Window is selected, the Run button on the Syntax window toolbar is clicked or the same choice is selected from the Run menu and the shortcut Control+R is used.

Interpreting SPSS Output

The output of the analysis or the tabulations done can be viewed in the SPSS Output Viewer window. The SPSS Viewer window displays output, including the warning and error messages. The left-hand side of the window contains the log history highlighting navigation of all the outputs run at a single output format in the opened output file.

It is important to explore both warning and error messages to debug syntax errors. Warning messages should be carefully examined to prevent earlier programming problems from confounding later analyses in ways that can be frustratingly difficult to debug.

SPSS Help and Documentation

The functionality of Help and Documentation is very elaborate and helpful in SPSS. The SPSS Help Index can be opened by selecting *Help->Topics* on the menu bar. The researcher can then go through the several aspects of Help such as the basic steps of Syntax, Command Syntax and SPSS.

To get Help in the SPSS Viewer for interpreting any output table, the researcher right-clicks an output table and selects *Results Coach* in the popup menu. Results Coach is a well-organised, context specific introduction to interpreting output items.

8.5 Qualitative Data Analysis

Like quantitative data analysis, *qualitative data analysis* also aims to describe and summarise the qualitative data generated in the form of text, graphics or audio through interviews and observations. It envisages revealing hidden theories and phenomena by relating them to concepts, behaviour and ideas.

The initial preparation is like that of quantitative analysis regardless of the formal analysis procedure and methodology used for data analysis. Qualitative analysis pays a lot of emphasis on observation during data collection, procedures and the initial process of analysis which starts from the first day of data collection.

Qualitative data is usually in the form of written texts, though nowadays, a lot of graphic and audio data is also collected to complement the textual data. Textual data is in form of transcripts of field notes, observations, or findings of interviews and in-depth discussions. Non-textual data, which nowadays is analysed in volumes, could be in the form of pictures, musical scores, tape recordings or video films.

In qualitative research, it is believed that data analysis is as good as the transcription of the raw data. Transcription is an important stage in qualitative data analysis and some degree of transcription is involved in almost all qualitative research studies. Transcription is not about jotting down or summing up what a researcher, interviewer or transcriber feels, it's all about noting down what the respondent has reported about what he or she feels. Thus, it is important to point out that there is the potential problem of a researcher's bias during transcription, due to which the researcher notes only those sections that seem relevant or interesting to them. During qualitative data analysis, the researcher should take special care to avoid this issue. Transcription is not an act which is done in isolation; it involves lot of processes, which ultimately contribute to collection of quality raw data that is ready for analysis. Transcription coupled with content analysis, takes care of qualitative data analysis and usually includes the following steps:

- At the first stage, all relevant documents that need to be analysed in the project are collected.
- The documents and even the segments of text within the document should be indexed.
- Interesting passages or segments within the passages should be coded or noted.
- Based on codes and notes in the document, the researcher should look out for frequently occurring text and phrases which are of analytical importance.
- The researcher should make notes and memos about the emerging concepts, ideas and theories as part of relationship and network building process.
- The researcher should revisit previously collected data after getting some lead into theory building to further support the emerging theory.

Analysis Methods

There are various distinct approaches to undertake qualitative analysis and none of these can be defined as the right way or the wrong way of analysing qualitative data. Some of the commonly used data analysis techniques are listed below:

a) *Typology:* Typology, as the name suggests, is a classification system which classifies patterns, themes, or other kinds of groups of data (Patton M. , 1990). As these categories are often not mutually exclusive and exhaustive, data analysis or the classification system may not be very accurate.

b) Grounded Theory: Grounded theory was developed in the late 1960s. It uses an inductive method of qualitative research and today it has become by far the most widely used framework for analysing qualitative data. Grounded theory has been defined as 'theory that was derived from data, systematically gathered and analysed through the research process. In this method, data collection, analysis and eventual theory stand in close relationship with one another' (Corbin, 1998). It consists of 'plausible relationships' among sets of concepts, which are directly developed from data analysis. Thus, two central features of grounded theory are that it is concerned with the development of theory out of data, and while the approach is iterative, it means that data collection and analysis proceed in tandem.

Grounded theory starts with a clear research question and it owes a lot of its analytical skill to its constant comparative method. In this method, concepts or categories emerging from one stage of the data analysis are compared with concepts emerging from the next. In a way, by constantly comparing the concepts and categories, the researcher looks for the relationship between them, while looking out for the emerging theory.

This continues until the stage of theoretical saturation has been reached, i.e., when there are no new significant categories or concepts which are of interest. At the initial stage of data analysis, after familiarisation with field notes, the researcher looks for indicators of categories in events and behaviour, which are subsequently coded.

The coding of the notes of text passages which are of interest is done and codes are assigned (mentally or by scribbling) to text passages, which are then compared to find similarities and differences between the codes to reveal concepts and categories. At the next stage, memos are developed on the comparisons and emerging concepts and categories till the point of saturation. Thus, eventually the emerging concepts and theories can be combined in a logical way to find the relationship between them.

c) Framework Analysis:

Framework relationship, in some way, is like grounded theory as it also entails coding and indexing for the preparation of charts and maps, while looking out for the relationship. It is also different in many ways, for unlike grounded theory, framework analysis was developed explicitly for applied policy research and is used widely nowadays in health related research.

The process of framework analysis starts by getting familiar with the raw data. After this, the thematic framework needs to be developed by reviewing and analysing the emerging issues. At the next stage, numeral and text codes are assigned to specific pieces of data based on the developed thematic framework. The data is represented pictorially under the broad heads of the thematic framework by using charts and maps, which later pave the way for identifying patterns, associations and relationships. Framework analysis thus involves a few key stages, which are mentioned below:

- Familiarisation with the data
- Identifying and developing a thematic or conceptual framework
- Reading, indexing and coding
- Charting, mapping and interpretation

d) *Analytic Induction:* Analytic induction is a data analysis approach in which universal explanations of a phenomenon are sought through data collection. This is done until no cases that are inconsistent with the hypothetical explanation are found. At the next stage, the hypothesis is revised to encompass all observed examples to develop a hypothesis that accounts for all the observed cases.

It begins with a rough definition of a research question and continues to formulate a hypothetical explanation of the research question and then proceeds to look out for deviant cases to reformulate or confirm the hypothesis.

e) *Logical Analysis/Matrix Analysis*: Logical analysis, as the name suggests, is based on the logical reasoning process and uses flow charts, diagrams, etc., to pictorially represent ideas, theories and concepts.

f) *Event Analysis/Microanalysis:* Event analysis focuses on finding precise beginnings and endings of events by finding specific boundaries and things that mark boundaries or events. It has found its application mostly in the case of video data.

g) *Metaphorical Analysis:* Metaphorical analysis takes its cue from various metaphors and analyses how well they fit in an observed scenario. The process entails asking participants for metaphors and to jot down their spontaneous answers, which can be linked with the observed phenomenon.

h) *Domain Analysis:* Domain analysis describes a social situation and the cultural patterns within a domain. It emphasises describing the meanings of the social situation to participants and interrelates the social situation with the cultural meanings.

Domain analysis focuses on different kinds of domains, viz., folk domains, mixed domains and analytic domains. The various steps integrating domain analysis are mentioned below:

- Select semantic relationships
- Prepare domain analysis worksheet
- Suggest broad and narrow terms to describe semantic relationships
- Formulate questions about semantic relationships
- Repeat the process to look out for a different semantic relationship
- Make a list of all domains covered

i) *Hermeneutical Analysis:* Hermeneutical analysis not only looks for the objective meaning of text, but looks for the meaning of the text in the context of the people involved. Hermeneutical analysis looks at different layers of interpretation of text.

j) *Semiotics:* Semiotics is the study of signs. It is defined as an approach which seeks to analyse and document phenomenon to reveal deeper meanings of that phenomenon. It explores the meaning of signs and symbols encapsulated in a relationship or phenomenon and further analyses how signs and symbols have an impact on the prospective users of those signs.

k) *Meta-analysis:* Meta-analysis is the process of combining results of several different studies on the same subject or question to look for answers to the current research problem. Though it is quantitative in nature, it involves sorting and coding techniques found in qualitative research.

i) *Content Analysis:* Content analysis is an approach for the analysis of documents and texts, which seeks to quantify content into meaningful and predetermined (in line with the objectives of the study) categories in a coherent way. It involves manual or automated coding of the text in the form of documents, transcripts or even of the audio of the video media to obtain phrases and theories for analysis.

Qualitative Research using Computers

In contrast to quantitative research, qualitative research uses a logical structure to make an inference about the context and meaning of the qualitative data generated. It tries to reveal the underlying theme encapsulated in a plethora of words or images generated through data collection. It does so by establishing theories and the relationship between various identified themes to look for relationships and networking. The strength of the question lies in their effectiveness in ensuring depth and detail and the ability to generate new theories and reveal subtle phenomena. Thus, it is imperative to avoid pre-judgements and preconceived notions to capture people's perspectives and opinions and avoid bias.

Computer-Assisted Qualitative Data Analysis

In practical situations, and even in qualitative studies nowadays, a high volume of data needs to be analysed and that too quickly. It's this quest for data analysis which has led to the development of a range of specialist qualitative packages with text retrieval and data handling capacities. In qualitative data analysis, whether it is beneficial to use a software package depends upon various factors like - the objectives of the study, the kind of data that needs to be analysed and an analytic framework that the researcher wants to adhere to. For instance, if the researcher wants to build theory and relationships, they should use ATLAS/TI and N Vivo series software programs, which use advanced search and query functions to retrieve data.

Software packages offering qualitative analysis facility depend upon procedures for coding the text. Indexing and coding (marking or attaching tags to relevant segments of text) form an integral part of any qualitative research process. These coded texts are then analysed to reveal the hidden or the emerging theory. Qualitative computer software programs facilitate the attachment of these codes to the strips of data and allow the researcher to retrieve all those instances in the data that share a code by formulating a query or search function.

Popular Software for Data Analysis

Qualitative data analysis using statistical software packages is gaining ground and is now used widely by researchers, not only because of the ease with which it can be used but also because it saves time. Now there is a whole range of statistical softwares that is available for data analysis. Some of the popular qualitative softwares available in the market are described below:

a) Ethnography

Ethnography uses segment as the basic unit to retrieve the text from documents and can be described as a text retrieval statistical software.

Segments are then coded in the form of twelve code words, which can be nested or overlapped upto seven levels deep as a sort of query function to retrieve text. It is important to point out that the search results depend on the query function in the form of nesting using codes.

b) HyperQual

HyperQual is an integrated software, which encompasses all features for data entry, storage and retrieval. It thus helps in entering data collected from interviews, observations and documents for further analysis.

c) Hypersoft

Hypersoft provides an integrated facility for data entry and data retrieval for theory building. This software also provides the options of indexing, searching and analysing textual data. Hypersoft does so by providing facilities for annotating, categorising, and coding of raw data.

e) Textbase Alpha

Textbase Alpha provides the facility for coding of both narrative text as well as data which has internal structure. Textbase Alpha also provides the facility for assembling coded segments, which can be represented by frequency counts and data matrix output. Textbase Beta, the new and modified version of Textbase Alpha provides the facility to code longer portions of lines.

f) SONAR

SONAR is basically a text retrieval software which also provides the facility for coding. While using SONAR, it is not required to extensively prepare the text files, while the data can be retrieved instantly based on the coded segments. Data can be edited by putting comments or an index of comments. SONAR also provides the facility for conditional searches based on Boolean operators.

g) ATLAS/ti

ATLAS/ti is perhaps the most powerful software that is available today and has an impressive user interface. ATLAS/ti works on the same logic for searching, coding and indexing as NUDIST, another qualitative data analysis system. However, in the case of ATLAS/ti, the concept of node and hierarchical logical system which is the key to the analysis process in NUDIST is not followed. Rather, it focuses on the inter-code relationships and theory building, rather than straight code and retrieve. Codes and super codes which represent a series of statements expressing textual relationships are used to build theory, concept and relationship through operators.

There are various qualitative data analysis softwares available, however, as stated above, it should be cautiously decided on which software should be used in a project depending on the data that is to be analysed and the analysis theory that is to be followed.

ATLAS/ti is probably the most powerful qualitative analysis tool for text, video and audio data. It can work with text, video and audio files of various formats. Though it follows the analysis paradigm of Grounded Theory and the qualitative content analysis, it is used universally by qualitative researchers from all realms of analysis theory. It helps to reveal the complex phenomena hidden in qualitative data in an exploratory way.

Qualitative data analysis using ATLAS/ti usually involves the following sequence of steps for systematic data analysis:

Step I: At first, a Hermeneutic Unit (HU), which signifies a project or an idea container, is created to contain all qualitative data including findings, codes, memos, and structures under a single name. ATLAS/ti is started by clicking on the software icon. It opens the main window known as Hermeneutic Unit (HU) Editor, which is the main work space area. It is the main editing tool which provides access to all other tools like the title bar, main menu, tool bar, main tool bar and the primary document toolbar.

HU contains all the primary documents, which can be texts, graphics, or sound files for a given project. Further, within each primary document, the analyst selects specific quotations and assigns codes. Codes can be grouped into families and quotations and coded objects can be further documented in memos or graphically displayed as networks of relationships.

Step II: After creating a HU, all the data files relevant to the project, including the primary documents in the form of text, graphics and audio is associated with the created project or the HU.

Step III: After associating all relevant data files or Primary Documents with the HU, the relevant text passages or pieces of special interest are selected and further assigned codes or memos. This step transforms all relevant documents in codes and memos, which are used in later stages of query indexing or theory building.

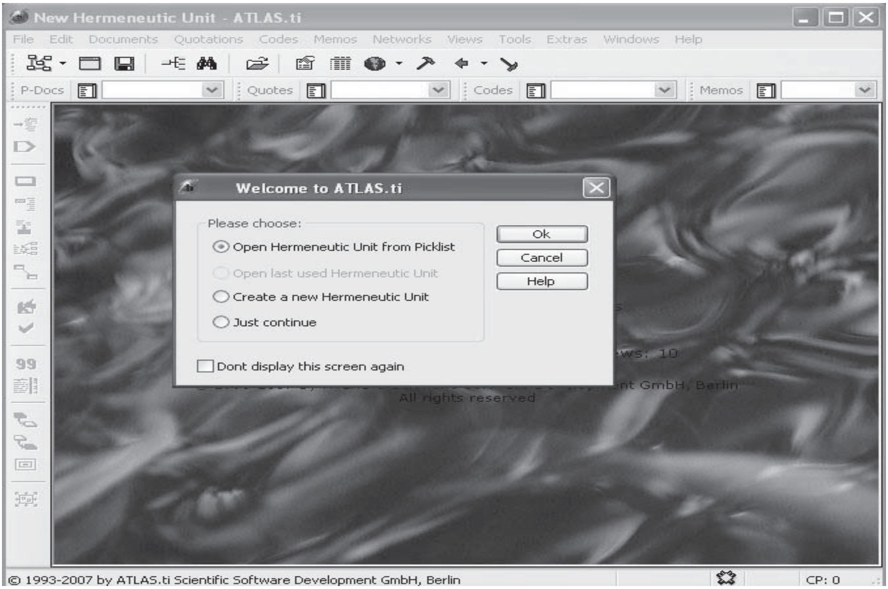

Figure 1.33: ATLAS/ti Main Window

Step IV: After the initial stages of assigning codes and memos, the coded data segments are reviewed thoroughly to ascertain whether it is necessary to associate more files with the HU.

Step V: After assigning codes and memos to all relevant data files associated with the project, all the primary documents, codes, and memos are organised using 'Families'. This is used later in developing theory.

Step VI: After organising the primary documents, codes and memos, as the next step, the semantics, conceptual or terminological networks are built from the codes and memos which have been created and organised. These conceptual networks along with codes, memos, super codes and families help in crystallisation of the emerging theory.

Step VII: As the last step, based on the emerging theory, a report is prepared. In case any further analysis is required, the data is exported for further analysis using the statistical software.

h) NVivo

The NVivo series software developed by QSR International Corporation is also one of the qualitative analysis softwares widely used by researchers. It is based on the concept of building a logical hierarchical system of Nodes where each Node is logically related to other Nodes, which it does by flagging and text search to construct a possibly large and highly structured hierarchical database indexing into the documents to be analysed. Node search uses seventeen operators including Boolean, proximity and other operators to generate new nodes to build hierarchical indexing categories encouraging generation of new ideas.

After the NVivo program is started by clicking in the software icon, one of the three options provided can be selected viz., Run a Tutorial, Start a New Project or Open an Existing Project (Figure 8.31). To start a new project, the dialog box of '*New Project'* is checked.

The NVivo menu provides the option to start a *New Project*, or to *Open* an *Existing Project*.

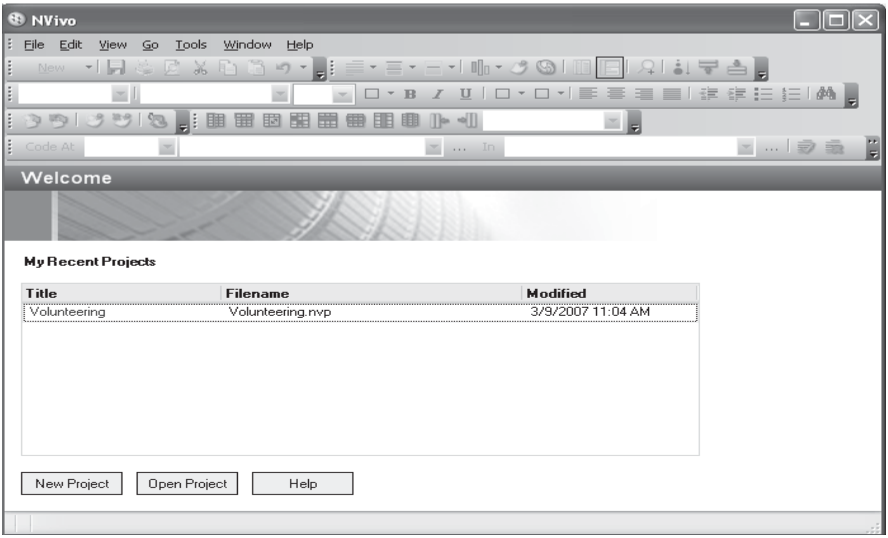

Figure 1.34: NVivo Main Window

Starting a New Project

A project in NVivo can be said to be like a big container which contains all unstructured data, analysis and reporting tools relevant to a specific project.

A project in NVivo encapsulates the following broad themes:

- All Documents: 1) Internal documents such as transcripts of focus group discussions (FGDs) and field/project notes 2) External documents such as books, audios and videos
- Relevant Nodes which represent the emerging themes, cases or relationships
- Attributes (such as age or gender) assigned to cases and Memos which represent the researcher's notes
- Text Queries to ascertain patterns and pursuing ideas
- Models which can visually demonstrate and explore themes

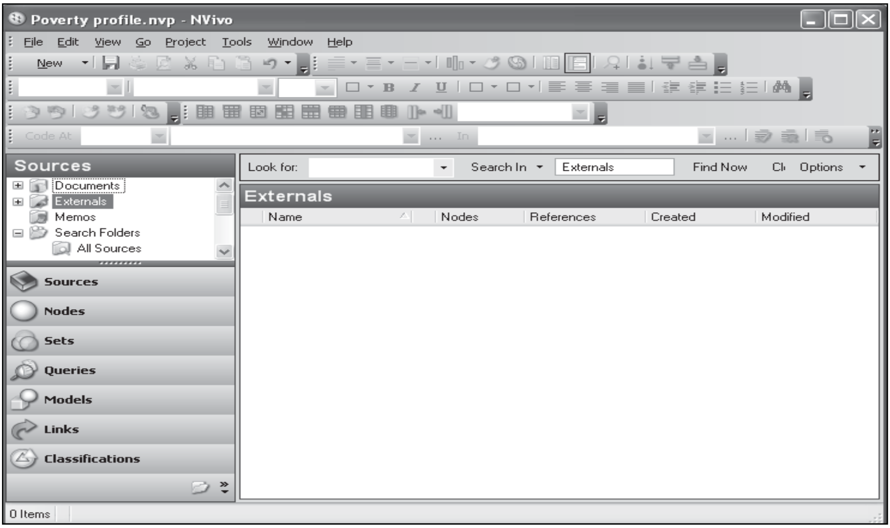

Figure 1.35: NVivo Main Window

The NVivo workspace is categorised into three different spaces viz., Navigation View, List View and Details View. The first upper side pane of Navigation View provides information about the all project related documents.

Reports

Using the NVivo program, it is possible to produce a range of different reports such as:

Project summary – As the name suggests, this is a summary of the entire project, e.g., the number of documents included, the number of attributes, cases, memos etc.

Node summary –This provides a summary of each node, e.g., the number of words coded for this node, the number of sources and cases.

Coding summary –This summarises the allocation of words to each node as a percentage of the overall word count of the document.

To summarise, the following activities can be done using the NVivo software:

- Associating all documents, viz., 1) Internal such as transcripts of FGDs and field/project notes 2) External documents such as books, audios and videos
- Creating relevant Nodes which represent the emerging themes, cases or relationships
- Assigning Attributes (such as age or gender) to cases and Memos which represents the researcher's notes
- Text Queries to ascertain patterns and pursuing ideas
- Models which can visually demonstrate and explore themes
- Preparation of reports based on emerging findings.

CHAPTER NINE

EXECUTING MONITORING EVALUATION AND LEARNING

The final chapter of this manual presents the operational dimensions of executing the Monitoring, Evaluation and Learning (MLE) component of developmental projects and programmes. The MLE function of a project/programme involves its planning, management, execution and finally using its results. Executing MLE involves several logical and sequential steps starting from finalising the theory of change, aligning it to project goals and objectives, defining clear hypotheses, selecting the M&E design, selecting a sampling scheme, implementing the M&E framework and analysis, and finally communicating the results through a well-defined learning plan. The chapter also discusses the quality assurance plan along with analysis and representation of the M&E result.

At the end of this chapter the researcher will be able to:

- ✍ List the steps involved in executing MLE for development projects or programmes
- ✍ Understand the operational dimensions involved in executing MLE for developmental projects or programmes

9.1 Finalising the Theory of Change

The first and foremost answer that an evaluator seeks is to the question, 'What needs to be evaluated?' Different stakeholders, including funders and implementers, have different views on what needs to be evaluated. Moreover, sometimes even within a project, different people have different evaluation priorities. This holds true in real world situations also, primarily because different stakeholders have different priorities and look at evaluation from different vantage points. To have a common view point on evaluation that all stakeholders are agreeable with, it is essential to have a theory of change which presents the evaluation objective and its reasoning equally to all stakeholders.

The theory of change (ToC) provides the blueprint or pathway by mapping out long-term goals and linking them to existing preconditions, besides specifying the causal link for each precondition. It also lists the basic assumptions about achieving a specific set of outcomes underpinning the importance of the context, which in turn helps in agreeing with the logical narrative for the intervention. In some cases, the programme may not have a ToC or even if the ToC exists, there is no agreement on the ToC among all stakeholders. Thus, in both cases, it is important for an evaluator to have detailed discussions with all stakeholders to develop/align the ToC and to attempt to articulate the reasons for the sequence of results to occur.

It is also important to point out that in real world situations, considering the dynamic nature of the interventions, the ToC is built-in for the underlying assumptions, risks, mechanisms, and external factors that may affect the ability of the interventions to achieve its intended results. Hence, the ToC developed at the start of the intervention should be flexible in nature with scope of constant evolution with the changing environment.

An assessment of the ToC during project implementation helps not only in the course correction of the project implementation which may be necessary due to change in ground realities and assumptions, but it also helps in redefining the evaluation objective and design thereafter.

9.2 Finalising the Logic Model

Different evaluators have different viewpoints on whether to first finalise the ToC or the Logic Model. The ToC provides a roadmap or a blue print of achieving an outcome, whereas the Logic Model provides intervention inputs and activities, the outputs they will produce, and the connections between these outputs and the desired outcomes. Thus, while the ToC summarises work at the strategic level defining causality, the Logic Model unpacks the implementation paradigm of the ToC by defining a specific set of interventions.

After building on an agreed ToC, the evaluator works on the Logic Model with the programme team and in case the programme does not have a clear Logic Model, it is essential that the evaluator develops one in collaboration with the relevant stakeholders. While working on the logic model, it is important to have an agreement about its specific components, viz., the inputs or the resources required to execute the intervention, activities that need to be carried out, outputs, outcomes and the impact to be achieved.

Further, as explained in the previous chapters, the Logic Model must also outline the specific indicators for measuring the aspects of each component.

9.3 Finalising the Purpose of Evaluation

At the next stage, building on the agreed ToC and the Logic Model, evaluators work in collaboration with stakeholders to define the purpose of evaluation. To do so, evaluators re-confirm about the need of evaluation to define the evaluation's purpose and scope. It is important to probe whether the focus is on evaluating the project comprehensively or whether the focus is on a specific component? If the focus is on a specific component, the reason for the same is investigated and information regarding the components of the programme that should be evaluated first is solicited.

Based on the purpose of the evaluation, the different terminologies that are used to describe evaluations are as follows:

- Comprehensive Evaluation
- Formative Evaluation
- Process Evaluation
- Outcome Evaluation/ Impact Evaluation

In an ideal situation, one should go in for comprehensive evaluation of the project/programme. *Comprehensive Evaluation* captures the change in the intermediate output, process and outcome level indicators to assess the attribution questions and links the outcome level change to the change in the process/output level to demystify the story of contribution. Comprehensive evaluation further ascertains the cost-effectiveness, scalability and sustainability of the intervention.

The *Cost-Effectiveness Analysis (CEA)* compares the costs in monetary units with the outcomes measured in quantitative terms. The outcomes are measured through rigorous evaluation and at the same time, the study also captures both financial and non-financial relevant costs incurred for the implementation. This is also extended to include the benefits to compute the *Cost Benefit Analysis (CBA)*. Kalpan-Meier (KM) or product limit method is used to assess changes in outcome using STATA or specialised software such as TreeAge in mutual consultation with the key stakeholders.

Comprehensive evaluation is not complete without commenting on the sustainability and scalability aspect of the intervention. The evaluation should also focus on finding out whether the programme is sustainable and scalable. The CBA also helps in providing critical information required for answering questions regarding sustainability and scalability.

While comprehensive evaluation summarises the intervention result, *formative evaluation* is an assessment of efforts usually done at the start of the evaluation and prior to their completion for improving the efforts. Its core objective is to provide feedback regarding the efficacy of the intervention and its progress.

Integrating Mix Method Design

The objective of comprehensive evaluation is not only to answer the attribution question but also the contribution aspect of the programme. It is essential that as part of evaluation, mix method approaches are used. Mix method approaches help in understanding the storyline of the effect of the intervention. Mix method approaches can help us in understanding the change in a specific context and culture.

One can use participative approaches or traditional qualitative methods such as FGDs and IDIs. Additionally, to map the process of change, one can build on qualitative approaches such as the success case method and MSC (Dart, 2003) stories method to collect stories of change to understand and analyse the process of change. As explained in the previous chapters, these stories are of changes both positive and negative, and attempt to document the outliers. Analyses of these stories across domains help the researcher to understand what and why some things work while others do not.

Process monitoring and observation data can also be very useful in mapping the process of change and in understanding the *context* in which they are operating. This provides answers to how and why the programme impacts were or were not achieved. Additionally, through this process, the practitioner aims to identify the facilitating and restraining factors and synthesise the lessons learnt for successful replication. This also ensures that feed-forward is provided for implementation across different interventions for improving systems and processes.

Outcome Evaluation and Impact Evaluation usually follow a similar trajectory of change. While outcome evaluation attributes the change in the outcome because of project intervention, impact evaluation looks at the long-term outcomes or impact that might have resulted from that programme. For example, in the case of specific health interventions, while outcome evaluation primarily captures changes at the ANC level, impact evaluation ascertains change in mortality.

Process Evaluation ascertains whether the programme activities have been implemented as intended or planned. Process evaluation plays a critical role in evaluating how a specific set of processes as enshrined in the project document have been implemented at the field level and whether it has further resulted in change in the outcome. Process evaluation is usually done at the end of the project but a concurrent process monitoring exercise can also be carried out to provide specific feedback about the course correction. Process evaluation helps in identifying processes and how or whether the processes have resulted in the desired outcome.

Besides ascertaining the process of change, it is also critical to identify success stories and delineate the lessons learnt for adoption, adaptation and replication. This is done using *documentation and exploration techniques*. Documentation is focused towards a practice or the process. In the implementation context, process innovations and institutional innovations are documented. Apropos the content of documentation, the context or background of the issue is explained first as it is important to explain the context of the problem before discussing how the problem is to be addressed. Documentation should also include detailing the thought process behind the development of the practice and then the actual process. Also, the favourable factors that have helped in the establishment of the practice and the restraining factors that have been overcome are detailed.

Based on the above tasks, the lessons learnt and best practices are delineated so that they can further be adopted, adapted and replicated across other projects for facilitating their efficient and effective implementation.

In addition to finalising the purpose of evaluation, it is also important to finalise the scope of the evaluation, which means, it is necessary to determine what in terms of resources, i.e., budget, staff, time etc., are available for the evaluation. It is also necessary to look at the long term learning and dissemination plan and whether the same has been factored in as part of the evaluation.

9.4 Finalising the Evaluation Question

After finalising the purpose of the evaluation, the next step is to finalise the evaluation questions that are highly relevant to the aspect(s) being evaluated. The evaluation questions are:

Result Questions

The result questions focus on assessing the impact of interventions at different levels. At each level, the result questions try to assess if there has been any improvement in the coverage and quality of intervention. Further, depending on the hierarchy of the result, they can be further defined as impact or outcome level questions.

Process Questions

The process questions are also related to the inter-linkage and sustainability questions. While the process questions at each level try to draw causal inferences about the intervention, the inter-linkage questions attempt to analyse the outcome of interventions at one level as well as at the next level of implementation. The questions framed under sustainability are cross-cutting in nature and seek answers to issues like the viability of interventions, the possibility of taking the interventions to scale and the scope of internalisation.

9.5 Finalising the Key Areas of Enquiry

Building on the broad question of evaluation, it is important to list the broader areas of inquiry across different components for evaluation. Though only the key questions are listed as part of the evaluation questions, it is important to ensure that all areas of enquiry are captured comprehensively including the intervention and confounding variables.

9.6 Finalising Indicators

After finalising the research questions, at the next stage, the evaluator is enabled with a better understanding and appreciation of the indicators mentioned in the Logic Model. During this process, each of the indicators are critically reviewed with respect to its harmonisation with the Logic Model, practical considerations of collecting data on the indicators and the nature of the indicators, i.e., whether they are **S**pecific, **M**easurable, **A**chievable, **R**elevant and **T**ime bound as explored in detail in Chapter 3.

This process leads to a better presentation of the logframe for stakeholder consultations. This set of indicators is presented to the stakeholders and through a consultative process, the final list of indicators is established.

Table 1.14: Review of the Indicators with Respect to the Logic Model and Theory of Change

Inputs →	Process →	Results/ Outputs →	Outcomes →	Impact
Efficiency			Effectiveness	
Tracking resources	Monitoring activities progress and quality	Monitoring immediate outputs achieved	Attainment of objectives	Contribution to the envisioned goal

As part of the MEL planning process, it is essential for the MEL team to review the indicators regarding the various stages of Logic Model and the ToC. It should be assessed if they are appropriate to each stage and whether they fulfil the SMART characteristics which each indicator should ideally possess.

In addition to the tracking of the impact trajectory, it is also essential to have an estimated timeline for interventions to have a measurable impact. Absence of a timeline makes it difficult to have realistic expectations of the outcomes and makes it hard to differentiate between a failed outcome and an outcome that has not occurred as yet. Hence, an evaluation framework and impact timeline should be developed to assess the impact of policy interventions.

Definitions of Key Indicators

As a part of the evaluation exercise, another step is to finalise the definitions of the key indicators for analysis and reporting. Indicators have been defined in Chapter 3 and it is crucial that the definitions of the indicators is specified beforehand to avoid different interpretations of the meaning of an indicator or a sub-indicator. For instance, an indicator that calculates the percentage of women who have received complete ANC is defined by the number of ANC check-ups, components of the ANC check-up like height, weight, abdomen check-up etc., the number of Tetanus Toxoid (TT) injections and the number of Iron Folic Acid (IFA) tablets that have been consumed to compute the percentage of women who have received complete ANC. It is important to reach a concurrence on the definitions of such key indicators for computing the coverage level estimates from the data and then validating the data.

The method of analysis of key indicators is also decided keeping in mind the definitions of key indicators.

9.7 Finalising Research Instruments

After finalising the indicators, the next task is to develop and finalise the research instruments. It is important that the research instrument, tool or questionnaire is developed in a systematic way. Not following appropriate and systematic procedures in questionnaire development, testing, and evaluation may undermine the quality and utilisation of data (Esposito, 2002).

Translation of Study Tools

After the research tools are developed, they are translated into local languages focusing on three basic aspects: semantic equivalence, conceptual equivalence, and normative equivalence of items. For verification of the translation quality, the translation/back translation method is used, having independent translators translate from one language to another, and then back again, to see if the original and re-translated item remains the same.

9.8 Pilot Testing

After developing the research tools, it is very important to thoroughly test them in order to ascertain their suitability in actual field conditions. Field testing of the developed tools is conducted in the non-sampled areas. Pilot testing is not only critical for identifying the issues in the questionnaire but also helpful in removing ambiguities and other sources of bias and error. Both pre-field and field type of pre-testing should be done. In field type of pre-testing, the study tools are tested under field conditions by professionals and investigators. The team members should be involved in the pre-testing exercise for at least two days. After the field type of pre-testing sessions, a debriefing session is conducted with the field team members and necessary suggestions are incorporated in the tools. It is important that not only the questionnaire but the entire survey protocol is pilot tested.

One of the key objectives of pilot testing is to ascertain the validity to analyse systematic error in measurement (Norland, 1990). Evaluators should try to assess all kinds of validity errors, viz., content, construct, criterion, and face. One of the key aspects of this assessment is to ensure that the questionnaire comprehensively captures all indicators and information with respect to the finalised evaluation objective. Additionally, the evaluator should also assess whether the questionnaire is measuring what it intends to measure. Besides this, the evaluator should assess whether the questionnaire adequately represents the content.

In addition to validity, it is also important to establish the reliability of the questionnaires. Practitioners can measure the accuracy of the research instrument using the reliability assessment test (Norland, 1990). They can use different types of measures such as reliability types (test-retest, split half, alternate form, internal consistency) depending on the nature of the data. Pilot testing is the key to conduct a reliability assessment test. In terms of the number of interviews, pilot testing should be conducted on approximately 20-30 subjects. Reliability can be measured by reviewing the initial data and by calculating the reliability coefficient (alpha). If the reliability coefficient is .70 or more, the research tool can be considered reliable.

There are other important questions that need to be answered as part of the pilot test regarding the research tool, viz.,

1. Do the questions convey the meaning they are supposed to communicate and whether the language used is appropriate with respect to the local context?
2. Do all respondents interpret the question in the same way?
3. Are the required instructions included as part of the question?

Additionally, it is also important to assess the time taken to complete the questionnaire. It is important to assess this to make sure the research tool does not become too long and does not exhaust the respondent, for if tired, the respondent will not be able to provide quality information. Thus, it is important to ascertain at what level respondents feel fatigued and structure the length of the questionnaire accordingly.

Revising the Research Instruments

Based on the results of pilot testing, it is essential to revise the research instruments. Sometimes, the pilot testing findings may call for major changes. In that case, it is suggested that another round of pilot testing be conducted to finalise the research instruments.

The field-tested tools are finalised in consultation with the client and serve the following purposes:

- Finalisation of the areas of inquiry: To enable the team to touch base with all the areas of inquiry enlisted in the ToR and give the communities an opportunity to add-on new areas of inquiry to the study.
- Finalisation of the activity schedule: This helps in getting the timeline right for field-level processes.

9.9 Preparation for Electronic Data Collection

Parallel to finalising the research instruments, it is important to also finalise the data collection protocol. The data collection protocol can be broadly classified into two broad categories of Paper-Assisted Personal Interviewing (PAPI) and Computer-Assisted Personal Interviewing (CAPI). In addition, in specific cases where resources are limited, Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Web Interviewing (CAWI)/Online Data collection approaches are also used.

Regarding the questionnaire's design in both CAPI and PAPI, a standard template should be developed for entering the data. The only difference is that in the case of CAPI, the template or data entry design is the data entry programme. However, in the case of PAPI, one has the choice of making the data entry template later as interviews are conducted using paper based research instruments.

In the case of CAPI, all enumerators require a personal digital assistant (PDA) for conducting the interviews, which turns out to be a substantial cost. It is suggested that over time, research organisations should build an inventory of PDAs so that they can be used for different projects. In the case of pen and paper based interviewing, PDAs are not required. Thus, only a few PDA are required for data entry for which PDAs with minimal configuration can also suffice. Most institutions outsource the data entry task for efficiency. One of the key challenges regarding the hardware is the rapidly changing technology. As the wear and tear of PDAs while the practitioners are working in the field is inevitable, in most cases, mobiles/tablets used for an evaluation do not last for more than two to three projects and hence, it is prudent to budget and plan accordingly.

One great advantage of CAPI is that the enumerators need to spend less time thinking about what question to ask next as the skips are automated in the programme. Thus, the CAPI interview has a more natural conversational flow than a PAPI interview and it also tends to move faster, reducing the total time it takes to administer the questionnaire. Further, it is observed that respondents (and enumerators) are less likely to become tired and/or frustrated with the interview.

Another critical benefit in the case of CAPI is the timeliness of data. One can look and analyse the data in almost real time and there is no time lag between data collection, cleaning and analysis. This saves lot of time, which may be to the tune of one to three months depending on the sample size. More importantly, timely availability of quality data helps to speed up decision making. It also helps in building ownership, as all stakeholders can look at the data on real time basis.

Overall, the key differentiating factor in favour of CAPI based survey administration is the assurance of data quality, which is worth more than the expense incurred on hardware and the time taken to build the right kind of software for CAPI. Different kind of checks such as validity checks, range checks and logical checks can be incorporated in the CAPI programme. These checks function on a real-time basis and provide cues to the enumerators to probe further and restrict the possibility of entering incorrect or illogical data. For example, if a value is out of range, the programme shows an error in a prompt window.

In developing countries, most of the evaluation organisations have already shifted to CAPI and with respect to long term investment, CAPI will pay off as it ensures good-quality data. Institutions are using state of the art technological interfaces viz., PDAs, **laptops** etc., as part of Computer Assisted Survey Interviewing (CASI). Census and Survey Processing System (CSPro) programmes which are custom designed are used on mobile phones or tablets or laptops. Some institutions also use survey platforms like Survey CTO for collecting data electronically.

Once the survey is completed, the complete (survey) data is electronically transferred back to the database server. Automated reports regarding the status of data collection can be generated from the database on real time basis. By reviewing the data, issues in the data quality can be identified and feedback can be shared with field managers, supervisors and enumerators on real time basis.

The schematic diagram presented below describes the processes used in data collection:

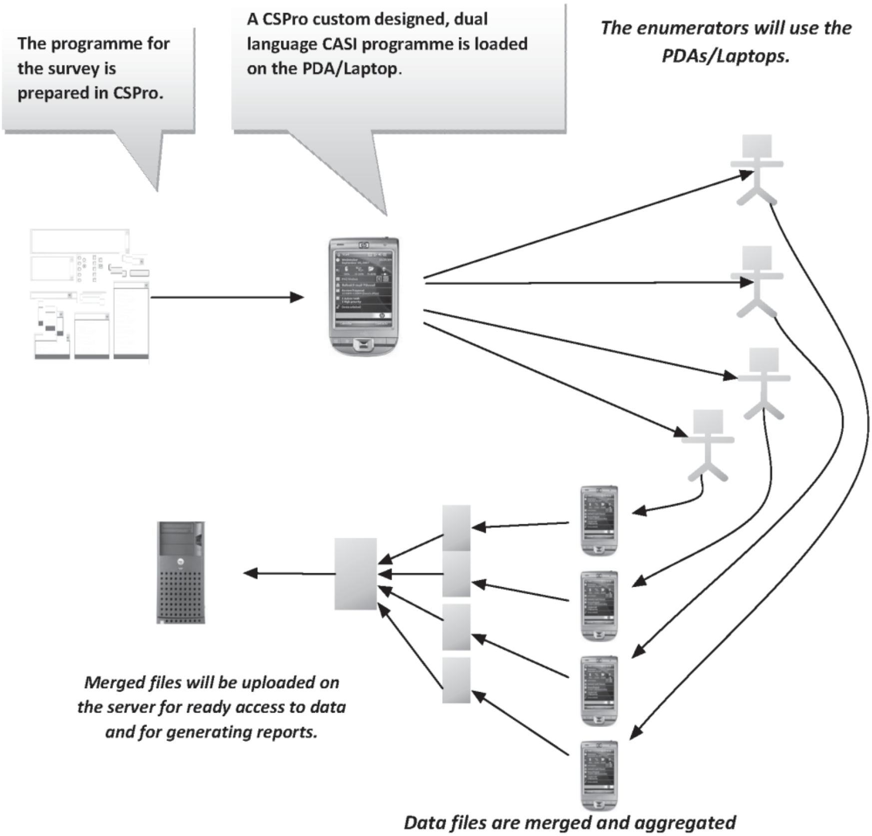

Figure 1.36: Data Collection Process using PDAs

9.10 Recruitment of the Field Team

Recruitment of enumerators, supervisors and field managers in the survey team is based on their educational qualifications and experience in conducting similar surveys.

The quality of the data depends on the skill of the enumerators recruited, trained and managed during the data collection exercise. Enumerators are the fulcrum of the data collection exercise and, hence, they are also the custodians of the data. The field staff consisting of field enumerators and supervisors are primarily responsible for the data collection effort.

Thus, it is essential that the enumerators selected for the data collection exercise possess the requisite skills. The first and foremost skillset required is good communication or people skills as the key job of enumerators is to communicate the purpose of the survey, solicit consent, highlight the importance of the respondent's participation and then solicit responses to research questions listed in the research instrument. It is always helpful to have enumerators who have either basic understanding of thematic areas wherein they are supposed to collect data or have prior experience of collecting data on similar thematic areas viz., public health, nutrition, livelihood etc. For instance, the data collection exercises which involve technical knowledge as in the case of collection of data on the quality of care at the health facility, it is in the interest of the programme to have enumerators who are trained in nursing. Similarly, in case the survey is being done using PDAs, it needs to be assessed if the investigators are comfortable using the same. The composition of the survey team i.e., the number or males or females required is as per the nature of the survey. For example, in the case of surveys related to reproductive health and family planning to be administered to women, it is advisable to have female investigators.

The recruitment process of enumerators should be standardised and according to the evaluation objective. In most cases, enumerators are freelancers and not working on a contract basis with any specific agency. Depending on the availability of work and established work-relations, the majority of the enumerators prefer working with certain organisations. Thus, recruitment for any particular survey becomes easy as the established organisations usually have a pool of enumerators to choose from, based on the nature and the location of the survey. In the case of large surveys or surveys of a different kind, recruitment is also facilitated through advertisements. Subsequently, interviews or written tests or both are conducted to check the suitability for a specific evaluation exercise. In both situations, whether the enumerator is recruited from an existing pool of enumerators or a new set of enumerators is recruited, it is important that recruitments are done based on competence i.e., skills required to conduct the survey effectively.

Another key skill set required for enumerators is the ability to grasp and follow detailed instructions in addition to literacy, basic numeracy, and the ability to absorb training. Language becomes one of the key criteria for the selection of an enumerator and enumerators should be fluent in the local language(s) which is native to respondents. In specific cases, an additional language may come handy, and more so in the case of multi-lingual surveys.

A few personal attributes are also important in the selection of enumerators. Enumerators should be able to approach strangers and build rapport with them, which is a key factor in soliciting consent for the interview. Additionally, enumerators should also be diligent in adhering to the protocol and following instructions. They should be responsible and should act as custodians of data by showing ownership in ensuring quality data.

As most of the enumerators are freelancers, it is important to ascertain that they are available for the entire duration of the training and the data collection effort before their recruitment is finalised.

Generally, as a contingency plan to cater to drop outs and rejection of enumerators who are not found to be adept enough to be part of the data collection, it is advisable to have an additional 15 to 20 per cent of enumerators and supervisors in the training programme so that enough team members are available to conduct the survey at the end of the training.

9.11 Preparation and Implementation of Data Collection

Careful and thoughtful field work enable teams to collect data systematically. The following steps ensure quality assurance of the data.

9.11.1 Development of the field manual

Development of the field guide or manual for data collection is essential to ensure smooth functioning and standardisation of field operations. The training manual serves as a guide for facilitators, enumerators and supervisors for managing the data collection effectively. The manual should be in the local language to serve as a ready reference during the survey for all research and field staff. It broadly consists of:

- Background and the research objective
- Research design and methodology
- Role and task of research and field staff
- Quality control mechanism and protocols which need to be followed
- Brief description of the key questions and explanation of the key concepts which are required for administering the questionnaire.

9.11.2 Training of field survey teams

Different enumerators have different skill sets and they may also have a different understanding of the study. Thus, to minimise interviewer bias and to ensure consistency and accuracy in the data collection exercise, it is important to train the data enumerators well. The objectives of the training exercise are to train enumerators on how to use the survey instrument(s), how to conduct interviews, and how to perform other data collection activities in a coherent and consistent manner according to the evaluation objective. This will help in minimising non-sampling errors and response errors substantially. Training of the survey team to deliberate on survey objectives, survey tools, sampling design, as well as on expected data quality are some of the critical steps that will ensure data quality. It is suggested that all the researchers and field staff go through the field manual carefully before the training session starts.

All members of the data collection team are continuously monitored during the training and based on the internal assessment of their performance, the teams are finalised. Those who are not up to the standard or are not performing well should not be included in the team in order to ensure data quality.

The training should be facilitated by a team of professionals. The field team should also be included in the exercise of identifying the errors in the questionnaire and the plausible questions and responses on which automated checks are applied to minimise data entry errors.

Enumerator training should take place in a systematic way and in a structured learning environment. A training agenda for the training exercise should be made and shared with all stakeholders. All stakeholders should be in agreement regarding the duration, design and method employed for the training.

The training should have a proper mix of classroom sessions and field practice. Classroom training should focus on training the field team on the structured interview schedule and related fundamental concepts. Besides lectures, the training should also include mock exercises and role play demonstration. Field practice sessions should be scheduled in non-sampled areas, once enumerators are trained adequately on research instruments. Debriefing sessions to reiterate and assess whether investigators have understood the questions correctly should be scheduled thereafter.

The training agenda should have sufficient time for classroom training, simulation and mock exercise, field training and then feedback. The training methods selected should include active learning opportunities for the trainee enumerators whenever possible. The training method should provide enumerators with an opportunity to practice, apply, or respond. In case PDAs are being used for conducting the surveys, technical training is also important and sufficient time should be given to train enumerators on the usage of PDAs.

The lead trainer should discuss the survey instrument or questionnaire in detail with the data collectors to point out specific instructions. The training should include an example of a completed instrument or an interview transcript for data collectors. There should be clear, written instructions and/or a script for the data collectors to follow while interviewing respondents.

Sufficient time should be allocated on how interviewers should introduce themselves as it has a strong influence on the respondent's reaction and willingness to cooperate. One of the key and critical aspects of collecting data is the ability to introduce oneself well and to establish rapport with the respondent as soon as possible. It is expected that enumerators should be able to communicate the purpose of the survey, what information is being sought, how the information will be used and to be able to assure respondents about the confidentiality of the information provided. It is important to have interviewers practice in role plays or mock interviews during the training sessions to help the enumerators to prepare for difficult or unusual situations which may arise. It is also important to ensure that ethics and protocols are followed at the field level. Following ethical guidelines such as, providing study related information, soliciting consent and ensuring confidentiality of data not only ensures transparency but it also ensures that the data collected is of good quality and the non-response rate is minimum.

Training using Visual Aids

Training of enumerators is a critical aspect of data collection and the effective use of training aids can enhance the quality of training. Though it is important to start the training on a lecture mode, the monotony of this mode of training can set in quite quickly and after a point, enumerators may not pay sufficient attention. In such situations, audio visual aids along with group or simulation exercises can be a more effective way of imparting training. Even during lectures, it is suggested that presentations are made with interesting graphics and videos. Flip charts can be used to make the training sessions more participatory. Simulation exercises along with mock interviews and role plays are also very helpful in making the training sessions participatory.

Testing and Scoring One way of ensuring the quality of the field team and to ensure that they are ready to go to the field for data collection is to rate and score them. One can use a mock survey exercise where one enumerator chooses another enumerator as a respondent and completes a mock questionnaire. Each investigator can be scored based on their performance.

Finalising the Team Size

The team size, i.e., the number of enumerators required, depends on the productivity of the enumerators, which in turn depends on the length and complexity of the research instruments which determines the time allowed for the survey to be conducted. In addition to this, geographical location or terrain plays a part in determining the team size. Thus, the number of enumerators required may vary from survey to survey.

Dry Run and Initial Feedback

The dry run is another activity which is the key to ensure data quality. During the dry run, the field team is informed that their field work has started, though in reality, it is like a practice session. The dry run helps the research team to understand how the field team will perform in real world situations. The research team and field managers can take stock of the situation based on the dry run.

Logistic Plan including Transportation

Transportation is a key aspect of logistics and can be a critical factor in ensuring data quality. Thus, the team size should be planned in a way that vehicles may be hired to build efficiency.

Generating a Sample Frame

In many cases, to generate a robust sampling frame at the primary sampling unit level, HH listing or mapping exercises are conducted. Enumerators or address listers can begin at the northwest corner of the segment and record all the residential addresses they encounter while travelling in a clockwise direction around the segment. They can proceed around a segment, always turning right when the opportunity presents itself, and in this way, add the address of each structure to the list. To ensure that house-listing is completed just before the survey begins (to maintain integrity of the reference period of cases), it is proposed that house-listing should be conducted at least one day before the actual survey.

9.11.3 Roles and responsibilities of the data collection team

It is very important to make a list of the roles and responsibilities of enumerators, supervisors and field managers or executives. These roles and responsibilities should be explained to the field team in detail and included in the training manual for their reference. The roles and responsibilities of the field team members at various levels are listed below:

Role and Responsibilities of the Enumerator

Enumerators have a central role in the study since they are the ones who collect information from respondents. Thus, the success of the survey depends on the quality of each interviewer's work. Successful interviewing is an art that should not be treated as a mechanical process. The skill of interviewing develops with practice but there are certain basic principles which should be followed by every successful interviewer. One of the key tasks of an enumerator is building rapport with the respondent. The enumerator should keep some of these points in mind while conducting an interview:

Make a good first impression: When approaching the respondent for the first time, the investigator should do their best to make the respondent feel at ease. With a few well-chosen words, the respondent can be put in the right frame of mind for the interview. The interview should be started with a smile and salutation such as '*namaste*' followed by the introduction of the investigator.

The approach should always be positive: An apologetic approach should not be adopted by using opening words such as, 'Are you too busy?', 'Could you spare a few minutes?' or 'Would you mind answering some questions?' Such questions invite refusal even before the process starts. Rather, the investigator should engage with the respondent by saying, 'I'd like to ask you a few questions' or 'I'd like to talk with you for a few moments.'

Stress the confidentiality of responses: The investigator should explain to the respondent that the information that will be collected will remain confidential, and that no individual names will be divulged at any stage. It is important to reassure the respondent about the confidentiality of the survey, especially if the respondent is hesitant about responding to the interview. Also, the interviewer should never mention other interviews or show completed questionnaires to other interviewers or supervisors in front of a respondent or any other person.

Answer the respondent's questions frankly: Before agreeing to be interviewed, the respondent may ask the interviewer some questions about the survey or how the respondent got selected to be interviewed. The interviewer should be honest and answer any queries the respondent may have in a pleasant manner and put the respondent at ease.

Interview the respondent alone: The presence of a third person during an interview is a hindrance in getting frank answers from a respondent. Therefore, it is very important that individual interviews are conducted privately and that all questions are answered by the respondents themselves. The investigator should keep in mind the points mentioned below while conducting the interview:

- Be neutral throughout the interview
- Never suggest answers to the respondent
- Not change the wording or sequence of the questions
- Handle hesitant respondents tactfully
- Not raise the respondent's expectations to believe that they would be getting any direct benefits by participating in the interview
- Not rush through the interview

The interviewer's responsibilities include:

- Locating the HHs in the sample as allocated by the team supervisor.
- Conducting interviews with the identified HH member in each of the HHs assigned to the investigator by their supervisor.
- Administering the complete questionnaire in the required logical order.
- Giving the respondent information about the survey and taking their consent before starting the interview.
- Checking completed interviews to ensure that all questions were asked and the responses neatly and legibly entered.

Roles and Responsibilities of the Supervisor

Controlling and ensuring the quality of data collection is the most important responsibility of the field executive or supervisor. Throughout the fieldwork, they are responsible for observing interviews and carrying out field-editing. It should be noted that it is necessary to observe the interviewers more frequently at the beginning of the survey and again towards the end of the survey. In the beginning, interviewers may make mistakes due to lack of experience or lack of familiarity with the questionnaire. These can be corrected by discussions with the team members in the evening, getting solutions for their problems and getting answers for their queries. Towards the end of the survey, interviewers may become bored or lazy in anticipation of the end of the fieldwork, and their lack of attention to detail may result in carelessness with the data. To maintain the quality of data, the supervisor should check the performance of the interviewers thoroughly.

Collection of high-quality data is crucial for the success of the survey. It is important that the quality of work is monitored through editing, spot check, back check and re-interviews. This is especially important during the initial phases of the fieldwork, when it is possible to eliminate interviewer error patterns before they become habits. Thus, the supervisor should:

- Observe some of the interviews to ensure that the interviewers are conducting themselves well, asking the questions in the right manner, interpreting and recording the answers correctly
- Spot check some of the HHs selected for the interview to be sure that the right HHs are being interviewed

- Review a few of the filled questionnaires to ensure that they are complete and consistent
- Meet each member of the team daily to discuss their performance and point out areas where they can improve
- Help to solve problems that the interviewer might face in finding the assigned HHs, understanding the concepts in the questionnaire or with difficult respondents.

During the fieldwork, the supervisor should assign work to the interviewers after considering the competence of individual interviewers and ensuring that there is an equitable distribution of the workload, allocate the HHs in such a way that all the investigators are working close to each other. This will help the supervisor in proper supervision and monitoring and communicating any problems to the field coordinator/project director.

Assigning work to interviewers

While assigning work to the interviewer/investigator, the supervisor should:

- Assign a maximum of two calls or interviews at a time to an investigator in usual circumstances so that it is easy to back-check the investigator's work.
- Assign more interviews than an interviewer can do in one complete day. This is necessary because some HHs and/or women may not be available for the interview at the time of the interviewer’s visit. Sometimes, there may be as many as three or four such cases in a day for a particular interviewer. Fewer HHs should be assigned at the beginning of the survey to allow time for discussion of problems and for close supervision.
- Provide sample details such as structure number, HH number, name of the head of the HH, the respondent's name etc., to the investigators so that it is easy to locate the sample HH.
- Assign calls or interviews by following the sequence of the list so that no sample is missed.
- Make daily work assignments. The supervisor should make sure that each interviewer has enough work to do for the day, considering the duration of an interview as well as the working conditions in the area.
- Distribute the work fairly among the interviewers, assign work considering the capabilities and strengths of each interviewer but

also make sure that a larger work load is not assigned to a certain set of investigators consistently. The supervisor should also ensure that each interviewer has all the required information and materials for completing the work assignment.

- Make sure that all the selected HHs have been interviewed before the team leaves that area.
- Maintain a checklist of all completed records for each day, and carefully monitor all assignments and work completed by each interviewer and for each work area for completeness and accuracy.
- Make sure that all the mobiles, laptops etc., are charged and are in proper functioning condition in case the survey is being done using PDAs.
- Make sure that the interviewers fully understand the instructions given to them and that they adhere to the work schedule.

Roles and Responsibilities of Field Executive/Field Manager

The responsibility of the field manager is to manage the field work starting with the recruitment of the investigators and supervisors, their training, ensuring data collection is carried out in adherence with the required quality standards and finally providing data to the field team. The field manager should concurrently monitor the data collection, make spot checks and back checks and have a look at the data collected regularly to provide feedback to supervisors and enumerators on data quality. The field manager should also conduct debriefing sessions initially and intermittently to discuss any issues which the field team might be facing and also provide feedback to the investigators based on his and the supervisor's observations.

9.12 Quality Assurance Protocols

The lead researcher and the core team members are directly responsible for ensuring that the data collection norms are adhered to. There should be regular visits from the field managers during data collection to ensure that the quality protocols are followed. The field executives and the supervisor should observe each interviewer multiple times throughout the course of the fieldwork. At first the investigators should be observed during the training sessions and then screened based on their understanding and skill and by observing them during mock sessions, field practice and dry runs. Each interviewer should also be observed during the first two days of fieldwork so that any error made consistently is brought to their notice immediately. Additionally, observation of each interviewer's performance is done during the rest of the fieldwork.

The field manager and executives, in tandem with the supervisors, should ensure best practices are followed at the data collection stage. Along with this, there should be regular visits by the research team. Handholding support should be provided to ensure data quality through spot checks and back checks.

Back checks

The field manager and executives should ensure that back checks are done by them or by the supervisors in all sampled areas wherein the completion rate is found to be low or it seems to be a problem. Back checks are a powerful tool for checking the quality of the data in which the information for a few respondents is systematically checked. This is done by conducting a short re-interview with some respondents and matching the results with what had been collected by the interviewer. Back checks or re-interviews help reduce the types of problems that affect the accuracy of the survey data.

Spot checks and observation

Some of the interviews conducted should be observed by the field managers, executives and supervisors to ensure that the interviewers are administering the survey tool properly. Rigorous spot checks and observations need to be done at the start of the survey so that immediate feedback can be provided to the investigators and the quality of the field work can also be monitored. All observations during spot checks should be noted and discussed with the complete team during the de-briefing sessions.

Validation Checks and Rules

To ensure the data quality and to minimise logical and data entry errors, basic checks should be included in the programme or the survey tool.

Validity checks should be run on one question or field at a time to ensure that the fields do not include invalid characters and that all the essential fields have been completed (e.g., no field is left blank where a number is required). A sample check is presented as follows: the example specifies that the value of the indicator for the question on outcome of delivery can either be 'live birth' or 'still birth'. The value of the indicator for this value cannot be missing or cannot have any other value than the two options of 'live birth' or 'still birth'. A log of cases with such kinds of errors should be generated for each survey.

Example:

Table 1.15: Example Validation Check

Indicator	Variable value check	Variable value
What was the outcome of this delivery?	Live birth-1 Still birth-2	
CHECK	Value=1 or 2	

Range Checks

For data fields containing information about a continuous variable e.g., height, weight, month etc., observations should fall within a specified range. Thus, if the value of a variable falls outside the normal range, it should be checked. Any value outside this range is erroneous and needs to be checked.

Example:

Table 1.16: Example Range Check

Indicator	Variable Value
At which month of your pregnancy did the ASHA visit you for the first time?	
CHECK	Value 1 should be from 0-9

Consistency Checks

Often, certain combinations of values of different variables that are within range are logically impossible or very unlikely. The validation rules are defined to have some checks to ensure data consistency. As mentioned in the table given below, the number of hours after which the woman delivered after reaching the facility should be less than the number of hours that the woman stayed in the facility in total. Such consistency checks are documented for different indicator values for the three types of data.

Table 1.17: Example Consistency Check

Value	Indicator
Value 1	How many hours did you stay in this facility?
Value 2	After how many hours of reaching the facility did you deliver?
CHECK	**Value 1 > Value 2**

Data Cleaning

After the completion of data collection, when the field team hands over the raw data to the research team, data cleaning is done by the research team. Data cleaning consists of checking the data thoroughly and checking to see if the data has all the variables, including the required number of cases. Open ended questions should be post coded as part of data cleaning. It is also necessary to ensure that the data is properly labelled. The data analyst looks for areas of data inconsistency and clarifies with the field team if required.

9.13 Data Analysis

The next step after data cleaning is to prepare a data analysis framework and go ahead with the data analysis. Though the data analysis paradigm is usually finalised at the start of the project, it is important to revisit the data analysis protocol and paradigm. In an ideal situation, an evaluation should have a mixed method approach, where the quantitative method provides the attribution story and the qualitative method contributes to the story line.

Qualitative Data Analysis

In contrast to the quantitative data collection, qualitative data collection through FGDs and IDIs are mostly conducted using paper-based qualitative in-depth research tools. There is a process sheet to document the information which includes the time and place of the interview, the subjects covered, the person providing the information, the field team member conducting and documenting that interview, and other relevant information.

The content analysis format is usually developed beforehand, which is not only helpful in the documentation process but it also helps in supporting data analysis and synthesis.

The qualitative data collected is transcribed and translated into English. Detailed field notes also supplement the transcription and the translation. After the transcription, the detailed coding framework is developed. This serves as a reference to the qualitative analysis software as per the pre-defined set of assigned criteria.

The data/content is appropriately coded and validated. The coding of content is carried out in parallel by two coders and processed to find out the inter-coder reliability. Evaluators use NVivo or ATLAS.ti to code textual data, wherein they attach annotation, indexing and memo at the query stage to build theory. The analysis not only looks for the major themes that cut across groups but also at the key insights. During analysis, different aspects and contextual themes are considered to build theories with respect to the broad areas of a) Context b) Internal Consistency c) Frequency d) Extensiveness e) Intensity and f) Specificity.

In terms of presentation of qualitative data, evaluators base the findings of the research around the categories or themes that have emerged. These are presented as sections with relevant sub-sections. Quotes are used to demonstrate and or to inform or support findings building on reliability and validity of each quote. Further, and if required, qualitative data is represented in a quantitative form.

Data Analysis Plan: Quantitative Protocol

The core team members, under the guidance of the project coordinator and the data analyst, prepare the analysis or tabulation plan. The required tables are generated using the latest versions of analysis software like SPSS or STATA.

The choice and nature of data analysis depends on several factors such as the type of the variable, the nature of the variable and the mode of analysis performed. Evaluators can choose either SPSS 17.0 or STATA 10.0 for analysis. SPSS is the most popular quantitative analysis software used today in social research. SPSS 17.0 can be used for basic analysis and for generating tabulated reports, descriptive statistics, and complex statistical analyses, in addition to which the SPSS package provides easy to use interface for analytics. STATA is another strong statistical software package which provides strong modules for evaluation methodologies such as DID, RD and IV. Nowadays, evaluators are also using R for its strong analytical power. The data analysis is usually categorised into two broad categories:

Basic Analysis

- Simple percentage of all variables of interest
- Descriptive statistics of key variables in socio-economic surveys
- Parametric test of association to ascertain the statistical significance of existing associations between variables of interest

Advanced Analysis

- Significance testing
- Calculating correlation coefficient
- Analysing internal consistency reliability of the scale using Alpha Coefficient
- Inequality Analysis
- Multivariate Analysis

9.14 Ethical Approvals

Ethics plays an important role in any research, and more so in the case of evaluations involving human subjects. Evaluation guidelines prescribed by various international and national organisations place central emphasis on ethics as the key tenet of evaluation. A snapshot of the evaluation guidelines presented in the table below highlights the key tenets of evaluation.

Table 1.18: Snapshot of Evaluation Guidelines

UNDP Norms for Evaluation	UNEG	American Evaluation Association's five guiding principles are	African Evaluation Guidelines
Independent	Intentionality of Evaluation: • Utility • Necessity	Systematic Inquiry	Utility
Intentional - Clarity in Rationale	• Obligation of Evaluators • Independence • Impartiality • Credibility • Conflict of interest • Honesty and integrity • Accountability	Competence	Feasibility
Transparent	Obligation to Participants • Respect for Dignity and Diversity • Rights • Confidentiality • Avoidance of Harm	Integrity/Honesty	Propriety
Ethics - Devoid of personal or sectoral interests	Evaluation Process and Product • Accuracy • Transparency • Reporting • Omission and wrong doing	Respect for People	Accuracy
Impartial		Responsibilities for General and Public Welfare	

High Quality			
Timely			
Used			

In terms of the operational aspects and in case it is required, it is important that approvals are sought from the local Institutional Review Board (IRB) prior to the study. In addition, it is critical that the key tenets of ethics such as taking proper consent, ensuring privacy and confidentiality are adhered to during the evaluation.

9.15 Learning Framework and Structure

One of the critical objectives of M&E is to derive learnings for course correction and to suggest changes with respect to programme replication and scale up. Learning as part of MEL becomes as critical as conducting M&E itself. More so, in practice where different partners and stakeholders are involved in contributing towards the project goal, it is imperative that the learning is shared not only among the M&E team members and the project implementation team but also with the donor and the external audience, including policy makers and evaluators, so that the lessons learnt can be built upon and replicated in other settings. It is in this context that it is imperative that a clear-cut learning structure is spelt out in which the modalities of knowledge sharing are also clearly specified. The learning structure should be in accordance with the reflective and integrated nature of the external MLE support and constitutes both ‘informal’ and ‘formal’ spaces.

There are four aspects to developing learning structures for intervention:

- Develop a learning framework
- Build dialogue/feedback spaces to support the ongoing learning
- Develop structures to support the ongoing feedback
- Pay attention to mechanisms of policy influence

Developing a Learning Framework

The first task after having an initial agreement on an M&E framework is to develop a Learning Framework.

One of the areas on which there is key focus in M&E is to explore questions on the configurations of contexts, mechanisms and outcomes which inherently result in a broad range of learning that includes knowledge development, policy, organisational process and impact learning (Sridharan, 2008). Thus, to formalise the learning, a key action step should be taken to develop a learning framework which will: (i) categorise the type of learning that is needed over the course of the evaluation; (ii) the timing when such learning is needed; (iii) the dissemination vehicle for such learning.

In this context, examples of learning include:

• **Policy learning**

Policy makers typically want to know more than just if a programme works or not (Sanderson, 2003). Based on the evaluation, policy makers also want to know how future versions of similar interventions can better translate the underlying policy theory into more clearly developed programmes.

• **Organisational learning**

'Social inventions are complex systems thrust amidst complex systems' (Pawson, 2004). Thus, the key questions related to organisational learning that need to be answered are:

- What kind of organisational structures are needed to successfully implement complex programmes?
- Which organisational structures support coordination across the different teams implementing the multiple interventions?

• **Process Learning**

It is important that the learnings related to processes are also derived from the programme evaluation. Key areas of non-compliance and bottle necks need to be identified and communicated to the programme team as part of process learning.

• **Knowledge Development**

Programmes often do not have sufficient knowledge about the lives of the programme recipients or beneficiaries they are trying to impact. Evaluations provide an opportunity to build such knowledge. Evaluations also provide opportunities to learn about the barriers that need to be overcome. This knowledge can be fed back into the system with the aim of further improving the programme or improving the future versions of the support programme.

• **Programme Impacts**

One of the most important focus areas of evaluations is impact. Evaluations aim to answer the programme related learnings, viz.,

- Does the programme model work?
- Are there demonstrable impacts of the interventions on programme recipients?
- Is there empirical support for programme impacts on interim and proximal individual-level outcomes based on the programme theory?

Dialogue Spaces for Ongoing Learning

Building on the learning framework, dialogue spaces are created for interaction between multiple stakeholders. Dynamic interaction is required among programme planners, implementers, and evaluators when contexts and mechanisms are being worked out. This means moving away from a purely product focused view of knowledge and encouraging spaces in which dialogue and ongoing learning is encouraged. Learning needs to be supported by well thought out processes and products. Learning processes need active and intentionally planned spaces for the ongoing dialogue.

Develop Clear Structures to Support Feedback and Learning

One key goal is to develop clear structures within the MLE team to support the learning and feedback. This implies:

- Having at least one person whose responsibility is to ensure that feedback is occurring periodically and that the same person is responsible for the ongoing alignment of learning from the evaluation
- Collaboratively developing the learning framework discussed above
- Working collaboratively with the implementing organisation and funders to develop both clarity of the evidence needs of the variety of stakeholders and a policy influence plan to connect the learning with the stakeholders

- Developing a clear responsibility for both a framework of influence and a dissemination strategy for all stakeholders and developing metrics for accountability to learning collaboratively
- Identifying events (both policy meetings, conferences etc.) that can support the dissemination strategy.

Developing a Policy Influence Plan

As one of the key objectives of monitoring is to influence decision making, similarly at the last mile, evaluation aims to provide evidence to replicate and scale-up intervention and to influence policy (Carden, Knowledge to Policy: Making the Most of Development Research, 2009) (Mark, 2004). For policy recommendations that are given based on the evaluation, there should be compliance with the points mentioned below:

- Fit with the latest '*political zeitgeist'* (The defining spirit or mood of a period of history as shown by the ideas and beliefs of the time). For example, the *political zeitgeist* may be conducive to implementing health equity policies. Working closely with policy makers and other policy and programme stakeholders can provide evaluators an avenue by which to understand the policy landscape and in some cases, impact the same through a broadening of policy horizons and enhancing policy capacities (Carden, Knowledge to Policy: Making the Most of Development Research, 2009).
- Raise the salience of the key policy considerations emerging from the evaluations to those with the power to decide (Mark, 2004)
- Be simple enough to be communicable
- Be simple enough for those delivering to be able to apply it
- Be cogent enough to persuade the various agents in the delivery chain that it is worth doing
- Be cogent enough to win over the required leadership - from the start, and from the top, to the end
- Be embedded within the intervention teams, e.g., explore perspectives with action research approaches which are helpful but not captured by the intervention team - with people who can remind the intervention team about the gap between their measures and the intended outcome. Such embedded evaluators can mediate/present the experience of others in similar contexts and advise on what is likely to work
- Educate policy makers and practitioners that an evaluation is not simply an instrument to determine if a programme is working

- Explore loss in translation from policy aspirations to programme design,as translating the policy aspiration into a demonstration project can be a challenge. Thinking about contexts, mechanisms, and outcomes in a focused manner can help raise questions about the match between the programme design and the longer-term policy aspiration. As described by Sanderson, functions of evaluations can include 'influencing the conceptualization of issues, the range of options considered and challenging taken-for-granted assumptions about appropriate goals and activities' (Sanderson, 2003).

The challenge of 'making the evaluation matter' is not just to generate learning(s) but to align such learning with policy priorities. A key focus of the dialogues or feedback is aligning the evaluation learnings to the changing policy landscape. However, there is also a growing body of literature on external validity (Deaton, 2009), which establishes how learning from evaluations that can be spread to impact populations has only a limited evidence base (Rodrik, 2008). Questions of alignment between the evaluations and policy priorities are especially important given dynamic policy landscapes, changes in the government and changes in policy priorities.

To give an example, a policy influence framework that focuses on inequities needs to go considerably beyond the details of a specific programme (Sridharan, 2008). A programme theory usually informs the practitioner why a programme may be effective, though it rarely describes why and how a programme can impact inequities at the population level. More formal, explicit and pragmatic thinking is needed on how and why an intervention can impact the population at large, and how a policy influence plan can help. Questions about contexts and mechanisms should not be confined to the programme level, raising questions about contexts, mechanisms and their relationship to equity outcomes at the population level can help enhance wide scale policy implementation, mainstreaming and sustainability. The dialogue spaces discussed below will help in developing such a policy influence plan.

Operational Mechanism of Learning Structures

In order to have coherent strategy, it is important for every agency and expert to have role clarity about their specific tasks and responsibilities. A reporting management structure is prepared and operationalised to address these issues and ensure clarity among partners.

A simple operational aspect such as initiating a system of conference calls fortnightly between all the key personnel from different partners for updating each other on the progress made so far also adds to the efficiency of the process.

A dropbox or samepage can been created to act as a depository of all relevant documents produced, and all documents and updates made available there. All partners have online access to the created document repository which ensures ease of document retrieval.

For sharing learnings with the implementation organisation, informal sharing of lessons should take place as and when required and at different levels and forums. However, for a systematic sharing of learning, it is important to have monthly meetings where it is expected that routine data issues and MLE learning is shared and discussed.

As the project starts being implemented at the ground level, it is imperative that the progress is shared with the funders and other key stakeholders frequently to ideate, operationalise the design and outcome, besides apprising the team there of the challenges faced by the project as well as the measures being taken to address them. A routine reporting in the form of an activity reporting template as part of reporting requirement is used.

In addition to the above, meetings to share the learnings with the funders should continue to take place as and when required.

To share the project learnings with the outside world, i.e., the policy makers and other evaluators, the project should generate working papers or methodological notes for presentation in local and international conferences as well as in peer reviewed journals. These documents are aimed primarily at policy makers/evaluators and highlight the new approaches attempted in the project.

9.16 Evaluation Report

Evaluation reports present findings, conclusions, and recommendations from an evaluation, including recommendations regarding how evaluation results can be used to guide programme improvement and decision making. The evaluation report can have the following chapterisation plan:

- Title Page
- Executive Summary
- Intended Use and Users
- Programme Description
- Evaluation Design Scope and Focus
- Results and Conclusion
- Programmatic Implication
- Learning
- Annexure

Executive Summary: An Executive Summary provides a summary of the complete evaluation starting from its design, intended users, evaluation focus, key results and its implication on programme application.

Intended Use and Users: In line with the spirit of utilisation focused evaluation, the evaluation report clearly specifies its intended use and the intended users for transparency and better use of evaluation findings. This helps in building the demand for evaluation use and also helps the reader or the user to interpret findings in a better way.

Project Description: Project description forms an important component of the evaluation report. It provides a background section on the project along with a ToC and logic model. In addition, it also provides a literature review to set the context for evaluation findings.

Evaluation Focus, Scope and Design: This chapter presents data sources and methods along with evaluation indicators, performance measures, data sources, and methods used in the evaluation which are described in this section.

Results and Conclusion: The Results and Conclusion section provides the findings derived from the evaluation study. This section also aims to provide a meaningful interpretation of the data, which is more than mere presentation.

Programmatic Implication: Based on the evaluation result, it is critical to provide insights for programmatic improvement or replication.

Use, Dissemination, and Sharing of Research Results: Ensuring the use of evaluation as its key tenet. Building on this key tenet, this section describes plans for the use of evaluation result findings, along with a specific dissemination plan clearly outlining user specific strategies. It is also important to list the ways in which different stakeholders should be involved taking cognisance of the complex and dynamic environment in which the intervention is executed.

BIBLIOGRAPHY

Aquaknow. (2016, 01 06). Retrieved from http://www.aquaknow.net/en/water-toolkit-trainings/monitoring-evaluation-and-audit-water-development-projects/monitoring-evaluation-and-audit-water-dev/15689

Brisolara, D. S. (2002). *Feminist Evaluation:Explorations and Experiences .New Directions for Evaluation.*

Carden, F. (2009). *Knowledge to Policy: Making the Most of Development Research.* New Delhi: Sage.

—. (2009). *Knowledge to Policy: Making the Most of Development Research.* New Delhi: Sage Publications.

Chemonics. (2016, 05 08). *Beyond Indicators: Why We Need Qualitative Monitoring ?* Retrieved from The Compass: http://blog.chemonics.com/beyond-indicators:-why-we-need-qualitative-monitoring

Connelly, L. M. (2000). *Feminism and Developmenet:Theoretical Perspectives.*

Corbin, A. S. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Sage Publications.

Dart, R. D. (2003). *A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique.* American Journal of Evaluation.

Davis, J. D. (2005). *The 'Most Significant Change(MSC)' Technique- A Guide to Its Use.*

Deaton, A. (2009). Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development.

Dharmendra Chandurkar, N. S. (2014). *Developing Monitoring and Evaluation Framework for Budget Work Projects. A How-To Guide for Practitioners.*

FAO. (n.d.). *The logical framework.* Retrieved from FAO Corporate Document Repository: http://www.fao.org/WAIRdocs/x5405e/x5405e0p.htm

Forster, R. (2002, April 19). Participatory Monitoring and Evaluation: Principles, Action Steps, Challenges : Introduction.

Foundations of Success. (2007). *Using Results Chain to Improve Strategy Effectiveness: An FOS How-To Guide.* Foundations of Success, Bethesda, Maryland, USA.

Gamble, J. A. (2008). *A Developmental Evaluation Primer.* The J.W. McConnell Family Foundation.

Gaventa, M. E. (n.d.). Who Counts Reality ? Participatory Monitoring and Evaluation: A Literature Review. *Institute of Development Studies-Working Paper 70*.

Greene, W. H. (2008). *Econometric Analysis.* Upper Saddle River, N.J. 07458: Prentice Hall.

Kari, O. (2004). *The Logical Framework Approach.* SIDA .

KEPA. (2015, 01 06). Retrieved from https://www.kepa.fi/: https://www.kepa.fi/tiedostot/differences-between-monitoring-evaluation.pdf

Kosheleva, N. (2016, 06 17). *Methodologically Inclusive Transformative Evaluation as an Enabler of Social Inclusion.* Retrieved from http://www.unesco.org/: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/SHS/pdf/Workshop-Social-Inclusion_IOCE.pdf

Mark, M. &.-5. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*, 35-37.

MDF Training & Consultancy (2016, 01 07). Retrieved from Urban Reproductive Health: https://www.urbanreproductivehealth.org/sites/mle/files/10%20Indicators.pdf

Michael Bamberger, J. R. (2006). *Real World Evaluation : Working under Budget, Time, Data and Political Constraints.* Sage Publications.

NORAD. (1999). *The Logical Framework Approach: handbook for objectives-oriented planning.*

Patton, M. (1990). *Qualitative evaluation and research methods.* Sage Publications.

—. (2008). *Utilization-Focused Evaluation(4th edition).* Sage Publishing.

Patton, M. Q. (2012). *Essentials of Utilization-Focussed Evaluation.* Sage Publishing.

Paul J. Gertler, S. M. (2011). *Impact Evaluation in Practice.* The World Bank.

—. (2011). *Impact Evaluation in Practice.* World Bank.

Pawson, R. G. (2004). Realist synthesis: an introduction. ESRC Research Methods Programme. Retrieved from http://www.ccsr.ac.uk/methods/publications/documents/RMPmethods2.pdf

Ramirez, D. B. (2013). *Utilzation Focused Evaluation-A primer for evaluators.* Southbound Sdn. Bhd.

Real World Evaluation. (2016, July 2). Retrieved from Real World Evaluation: http://www.realworldevaluation.org/Home_Page.html

Republic of Serbia. (2011). *Guide to Logical Framework Approach: A Key Tool to Project Cycle Management* .

Richard Blundell, M. C. (2000). Evaluation Methods for Non-Experimental Data. *Fiscal Studies* , 427-468.

Rodrik, D. (2008, October). The New Development Economics: We Shall Experiment, but How Shall We Learn? *Faculty Research Working Papers Series-John F. Kennedy School of Government - Harvard University* .

Sanderson, I. (2003). Is it 'What Works' that Matters? Evaluation and Evidence-Based Policy Making. *Journal of Research Papers in Education*, 329-343.

Sielbeck-Bowen, S. B. (2002). *Exploring femenist evaluation:The ground from which we rise.New Directions for Evaluation* .

SMART, Characteristics of Good Objectives. (2016, 01 04). Retrieved from Community Empowermnet Collective:
http://cec.vcn.bc.ca/cmp/modules/pd-smar.htm

Sridharan, S. G. (2008). *Evaluation of Primary Prevention Intervention: Have a Heart Paisley Phase 2.* . Glasgow: NHS Health Scotland.

Tilley, R. P. (2004). *Realist Evaluation.* Sage Publications.

UNAIDS. (2010). *An Introduction to Indicators* .

UNDP Evaluation Office. (2002). *Handbook on Monitoring and Evaluation for Results.*

UNICEF Evaluation Office. (2006). *How to design and manage Equity-focused evaluations.* UNICEF.

USAID Centre for Development Information and Evaluation. (1996). Preparing a Performance Monitoring Plan.