

Chapter 5

Methodology and Tools

The deliberation on the methodology has been made to understand the concept, methods and techniques which were utilized to design the study, collection of information, analysis of the data and interpretation of the findings for revelation of truths and formulation of theories. This chapter deals with the method and a procedure used in the study and consists of eight main parts-

- A. Locale of research.
- B. Pilot study.
- C. Sampling Design.
- D. Empirical measurement of the variables.
- E. Preparation of interview Schedule.
- F. Pre-testing of Interview Schedule.
- G. Techniques of Data collection.
- H. Statistical Tools used for Analysis of Data.

A) Locale of research

Kotowali Gram Panchayat of the English Bazarblock of Malda district in West Bengal was purposively selected for the study. The village

namely Kotowali was selected by random sampling. The area had been selected for the study because of-

- a) There is ample scope for collecting relevant data for the present study.
- b) Acquaintance with the local people as well as the local language.
- c) The concern area was easily accessible to the researcher in terms of place of residence.
- d) The area was very easily accessible to the researcher in terms of transportation
- e) The closure familiarities of the student researcher with the area, people, officials and local dialects.

B) Pilot Study

Before taking up actual field work a pilot study was conducted to understand the area, its people, institution, communication and extension system and the knowledge, perception and attitude of the people towards mango cultivation concept. The components of pilot study are:

- General information;
- Specific information;
- Prevalence of variables;
- Body languages of the prospective respondents;
- Access to physical location;
- The type, level and intensity of responsiveness;
- Related information including Agriculture.

C) Sampling Design

Purposive as well as simple random sampling techniques were adopted for the study. For selection of state, district, block and gram panchayat purposive sampling techniques was adopted because the area was ideal for mango cultivation, production function study

convenient for researcher and having the infrastructural facilities and in case of selection of villages and respondents simple random sampling technique was taken up.

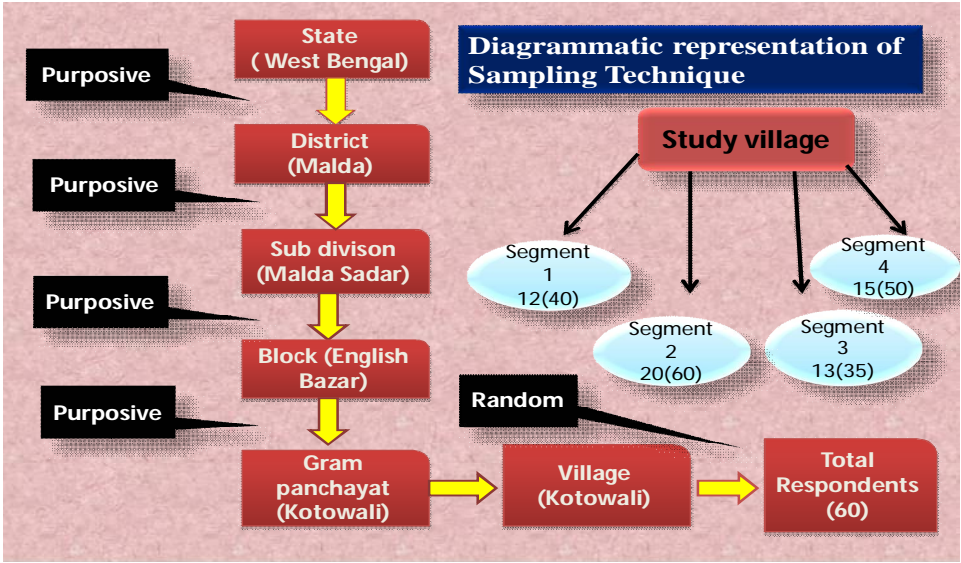


Figure no. 15: Sampling Technique and Sampling Design

D) Empirical measurement of the variables

Appropriate operationalization and measurement of the variables help the researcher to land upon the accurate conclusion. Therefore, the selected variables for this study had been operationalised and measured in following manner.

Variables in the present study have been categorized into two main categories.

- 1) Independent variables.
- 2) Dependent variables.

- **Independent variables :-**

Table no.14: Independent variables

SL No.	Variables	Notation
1	Age	X ₁
2	Education	X ₂
3	Family size	X ₃
4	Income from intercroops (per capita)	X ₄
5	Income from mango crop (per capita)(Rs)	X ₅
6	Size of mango orchard (Acre)	X ₆
7	Homestead area (katta)	X ₇
8	Distance Matrix	X ₈
9	Fuel consumption (Per capita, per month)	X ₉
10	Electricity consumption (Per capita)	X ₁₀
11	Frequency of visit to the market (in a week)	X ₁₁
12	Group interaction	X ₁₂
13	Problems related to mango cultivation	X ₁₃

1. Age (x1)

In the present study, the number of years rounded in the nearest whole number the respondent lived since birth at the time of interview, was taken as a measure of age of the respondent.

2. Education(x2)

Education may be operationalized as the amount of formal schooling attained/literacy acquired by the respondent at the time of interview.

3. Family Size (x3)

The influence of family members on the decision-making process of farm operation is inevitable. It refers to the number of members present in the family of farmers.

4. Income from intercroops (per capita) (Rs) (x4)

The Annual Income of a person is an important parameter to assess the economic status of the person in the society. It was operationally defined as the gross income from all the viable source of income in a single year. It is measured in terms of rounded of rupees. Total

income of the family from intercrop (Chick pea, vegetables mustard, lentil etc.) farming is divided by the family size.

5. Income from mango crop (per capita) (Rs) (x5)

It was operationally defined as the gross income from all the viable source of income in a single year. It is measured in terms of rounded of rupees. In the present study it has been calculated with the formula as follows.

(Total annual income from mango crop ÷ Family size)

6. Size of mango orchard (Acre) (x6)

The amount of land owned by a person is an important parameter to access the economic status of the person in the society. In the present study, actual area under mango cultivation is in acre.

7. Homestead area (katta)(x7)

Amount of land acquired by the home building and surrounding. In the present study, homestead area is in katta.

8. Distance Matrix(x8)

In this study distance matrix calculated as distance from respondent's house to -

- Health centre
- Pacca road
- Market/ hat
- Knowledge centre
- Financial institution

All these distances are taken in km unit.

9. Fuel consumption (per month, per capita) (x9)

Consumption of diesel, petrol, LPG and kerosene oil in a month (in Rs) / size of the family.

10. Electricity consumption (per capita) (x10)

Consumption of electricity in house (in Rs)/ size of the family.

11. Frequency of visit to the market (in a week) (x11)

Market research or survey is any organized effort to gather information about target markets or customers. It is a very important component of business strategy. So, this variable is very important character for the study.

12. Group interaction (x12)

Interaction is a kind of action that occur as two or more objects have an effect upon one another. The idea of a two-way effect is essential in the concept of **interaction**, as opposed to a one-way causal effect. **Group interaction** refers to the dynamics of the team and the way individuals in the **group** interact with one another. In this study group interaction is calculated on the basis of **10 point scale**.

STATEMENT	10- POINT SCALE
1.Frequency of interaction with panchayat leader	
2.Frequency of interaction with farmers group	
3.Frequency of interaction with development programmes	
4.Frequency of interaction with social service	

13. Problems related to mango cultivation (x13)

To find the problems related to mango cultivation in the study area Kotowali village questions are raised for inquiry, consideration, or solution. The scores are given on the basis of **5 point scale**.

Problems	Ranking (5 point scale)
1.Insect	
2. Lack of good quality sampling	
3. Hail stone and Norwester	
4. High price of input	
5. Lack of credit	
6. Disease	
7. Lack of irrigation facility	
8. High transport cost	
9. Low price offered	

(* 5= more serious, 4= serious, 3= moderate, 2= a little bit, 1= the least serious)

• **Predicted Variables :-**

Table no. 15: Dependent variables

SL No.	Variables	Notation
1	Yield of mango (unit area) (kg)	Y1
2	Yield marketed (kg)	Y2
3	Yield consumed (kg)	Y3
4	Cost of cultivation (Rs)	Y4
5	Net return (Rs)	Y5
6	GI Perception Index	Y6

1. Yield of mango(unit area) (kg) (Y1)

Total yield of mango in unit area is calculated and it is recorded in kg unit.

2. Yield marketed (kg) (Y2)

The amount of produce taken to the market for sell.

3. Yield consumed (kg) (Y3)

The amount of produce kept at home for consumption purpose.

4. Cost of cultivation (Rs) (Y4)

Cultivation expenses such as planting material, fertilizer, pesticide, labour, irrigation, land development, infrastructure etc.

5. Net return (Rs) (Y5)

Price gained after selling the produce – cost of cultivation

6. Geographical Perception Index (Y6)

The index is calculated on the basis of **5 point scale**. Some questions are asked to the respondents to know the perception about GI. After that scores are given.

- Do you know about Geographical indication?
- If yes, then what are the benefits?
- What are the varieties assigned GI?

E) Preparation of interview schedule

On the basis of the findings of pilot study a preliminary interview schedule was formed with the help of literature and by the assistance of Chairman of Advisory Committee. The interview schedule consisted of three major parts according to the specific objectives of the study.

F) Pre-testing of Interview Schedule

Pretesting or preliminary testing is the process of an advance testing of the study design after the schedule/questionnaire has been prepared. The object of pretesting is to detect the discrepancies that have emerged and to remove them after necessary modification in the schedule. It also helps to identify whether the questions are logically organized, the replies could properly recorded in the space provided for or there is any scope for further improvement. After conducting pretesting appropriate changes and modification of the interview schedule have been made.

G) Techniques of field data collection

The respondents were personally interviewed during June month and January month. The items were asked in Bengali in a simple term so that the members could understand easily. The entries were done in the schedule by student investigator himself at the time of interview.

H) Statistical tools for Analysis and Interpretation of Data

The statistical methods used for analysis and interpretation of raw data were –

1. Frequency
2. Percentage
3. Mean.
4. Standard deviation.
5. Coefficient of Variance.
6. Coefficient of Correlation.

7. Step down regression analysis.
8. Factor analysis.
9. Stepwise Canonical discriminant function analysis.

1. Frequency:

The term frequency is used to denote how frequently a response appears in a class or category.

2. Percentage:

Percentage used for making simple comparisons. For calculating percentage, the frequency of a particular cell was divided by the total number of respondents in that particular category and multiplied by 100.

3. Mean:

The mean is the arithmetic average and is the result obtained when the sum of the of value of individual in the data is divided by the number of individuals in the data. Mean is and relatively stable measure of central tendency. When the data are expressed in a frequency distribution (grouped), the mean is calculated using the following formula -

$$X = \frac{\sum_{i=1}^N f_{ixi}}{N}$$

Where,

x = Mean of the observation.

fi = Frequency of theclass.

xi = Mid value of the class.

N = Total number of observation

4. Standard deviation:

Standard deviation (SD) of a set of observation is the square root of the arithmetic mean of the squares of the deviations. The deviations being measured from the arithmetic mean of the distributions.It is

commonly denoted by the symbol (Sigma). To measure the average deviation from the standard value of the data standard deviation is used. The standard deviation of the data grouped in the form of frequency distribution is computed by using the following formula –

$$SD = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

When,

d = Standard deviation

N = Total No of observation in a particular coll.

X = value of observation in a particular cell

F = Frequency of observation

\bar{X} = mean number of observation

l = any number (e.g. 1, 2, 3) denoting position

5. Coefficient of Variance:

A measure of variation which is independent of the unit of measurement is proved by the coefficient of variation. Being unit free, this is useful For comparison of variability between different populations. The coefficient of variation is standard deviation expressed as percentage of the mean.

Coefficient of variation is measured by using the following formula —

$$C.V. = \frac{S.D.}{Mean} \times 100$$

6. Coefficient of correlation:

When increase or decrease in one variety is accompanied by an increase or in another variety, the two are said to be correlated and

the phenomenon is known as correlation. Correlation coefficient (r) is a measure of the relationship between two variables, which are at the interval or rational level of measurement and are linearly related. A Person product-moment “ r ” is computed by the formula.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

where,

X and Y	=	Original scores in variables X and Y
N	=	Number of paired scores
$\sum XY$	=	Each X multiplied by its corresponding Y, then summed
$\sum X$	=	Sum of X scores
$\sum X^2$	=	Each of X squared, then summed
$(\sum X)^2$	=	Sum of X score squared
$\sum Y$	=	Sum of Y scores
$\sum Y^2$	=	Each of Y squared, than summed
$(\sum Y)^2$	=	Sum of Y score squared

- The value of ‘ r ’ lies between +1 to -1. Positive values of r indicate that positive correlation between the two variables (i.e. changes in both variables take place in the same direction), whereas negative values of ‘ r ’ indicate negative correlation i.e. changes in the two variables taking place in opposite direction. A zero value of ‘ r ’ indicates that there is no association between the two variables.

7. Step down regression analysis:

- **Regression**

Generally a number of antecedent variables simultaneously contribute to influence the consequent variables, as in the case under study. It is of immense practical value to know the extent to which the antecedent variables, individually or jointly, could predict or contribute towards the consequent variable. This was done by computing multiple regression analysis. If Y is the consequent variable and X1, X2, X3...are the antecedent variables; the multiple regression equation is given by the following formula-

Multiple Regression Analysis

- Multiple Regression:
- $Y = a + b_1X_1 + b_2X_2 + B_3X_3 + \dots + B_tX_t + u$

Where:

Y= the variable that we are trying to predict(DV)

X= the variable that we are using to predict Y(IV)

a= the intercept

b= the slope (Coefficient of X1)

u= the regression residual (error term)

The significance of the b- value was judged by calculating their respective t-values and comparing them to the table, given by Fisher and Yates (1963), with (n-p-1) degree of freedom (where, n = number respondents and p = number of antecedent variables) at 5% and 1% level of significance.

- **Stepwise multiple regression:**

Stepwise regression is a variation of multiple regressions which provides a means of choosing independent variables that yield the best prediction possible with the fewest independent variables. It permits the user to solve a sequence of one or more multiple linear regression problems by stepwise application of the least square

method. At each step in the analysis, a variable is added or removed which results in the greatest production in the error sum of squares (Burroughs Corporation, 1975).

- **Forward selection**, which involves starting with no variables in the model, testing the addition of each variable using a chosen model fit criterion, adding the variable (if any) whose inclusion gives the most statistically significant improvement of the fit, and repeating this process until none improves the model to a statistically significant extent.
- **Backward elimination**, which involves starting with all candidate variables, testing the deletion of each variable using a chosen model fit criterion, deleting the variable (if any) whose loss gives the most statistically insignificant deterioration of the model fit, and repeating this process until no further variables can be deleted without a statistically significant loss of fit.

8. Factor analysis:

The term factor analysis refers to a set of analytical techniques designed to reduce data into smaller, meaningful groups based upon their inter-correlations or shared variance. Factor analysis is a very useful and popular method of multivariate research technique, mostly used in social and behavioral sciences. Factor analysis seeks to resolve a large set of measured variables in terms of relatively few categories, known as factors. In the Factor Analysis the "**Principle Component Method**" was followed.

- **Concepts Used In Factor Analysis**

Some important concepts used in factor analysis are explained, following Kothari (1996).

Factor: A factor is an underlying dimension that accounts for several observed variables. Factor is a hypothetical construct or classification. There may be one or more factors, depending upon the nature of the study and the number of variables involved in it.

Eigen value (or latent root): The sum of squared values of factor loadings relating to a factor is referred to as Eigen value or latent root. Eigenvalue indicates the relative importance of each factor in accounting for the particular set of variables being analyzed.

Factor Analysis is used:

- To reduce the dimensionality of large number of variables to a fewer number of factors.
- To confirm the hypothesized factor structure by way of testing of hypothesis about the structure of variables in terms of expected number of significant factor loading. Hence in factor analysis specific and error variables are excluded and only the common variables are taken into account.

9. Stepwise Canonical discriminant function analysis :

Probably the most common application of discriminant function analysis is to include many measures in the study, in order to determine the ones that discriminate between groups. For example, an educational researcher interested in predicting high school students' choices for further education would probably include as many measures of personality, achievement motivation, academic performance, etc. as possible in order to learn which one(s) offer the best prediction. Put another way, we want to build a "model" of how we can best predict to which group a case belongs. In the following discussion we will use the term "in the model" in order to refer to variables that are included in the prediction of group membership, and we will refer to variables as being "not in the model" if they are not included.

Significance of discriminant functions: One can test the number of roots that add *significantly* to the discrimination between groups. Only those found to be statistically significant should be used for interpretation; non-significant functions (roots) should be ignored.