SURVEY AND COLLECTION OF DATA: COMPILATION, TABULATION AND ANALYSIS OF DATA

Meaning of Survey

A survey is the scientific study of conditions and needs of a community for the purpose of presenting a constructive programme of social advance. In other words, it is a process of collecting information about the living and working conditions of a given community during a particular period of time.

Purpose of Survey

- To collect detailed information about a social problem under investigation.
- To obtain reliable information for the purpose of determining the needs of a particular community.
- To explore relationships between different social incidence.
- To make plan of a programme to be implemented in a particular area.
- To measure progress of a particular developmental programme.

Areas usually to be covered during an Agricultural Survey

- People: educational level, family size, occupation (primary and secondary), farming experience, group affiliation.
- Economic: Individual income, family income, sources of income, farm size.
- Land: topography, types, land utilization.
- Enterprises: variety of crop grown, nature and number of livestock, poultry, and their management style and nature and number of homestead trees.
- Farm implements: possession of different machinery, implements, and equipments.
- Media exposure: use of personal group, group and mass media.
- Realistic/fatalistic behavior.
- Nature of housing and sanitation.
- Awareness of recent issues.
- Skills in farm & home management.
- Knowledge, attitude and behavior towards any developmental aspect.

Qualities of a Good Enumerator

An enumerator must be:

- Clear about the nature of investigation.
- Acquainted with the type of instrument(s) used for collecting information.
- Clear about the fields and areas of survey.
- Mentally prepared to do field work and ready to use both primary and secondary sources of information.
- Able to create a congenial atmosphere with the respondents so that they provide correct information.

Advantages of Survey Method

- Social problems can be deeply and specifically identified.
- Relationships between different social incidences can be explained.
- Very useful for administrators and policy makers. In many cases, important remedial measures are taken based on findings of survey.
- The information can be dependable and reliable if the investigators are well trained and themselves comes in contact with the people for whom the survey is conducted.
- Facts of hidden problems can also be obtained.

Limitations of Survey Method

- A considerable amount of fund and trained enumerators are required, to conduct a dependable survey.
- Very time consuming; if efforts are made to complete time work in a hurry, obviously it will become undependable and unreliable.
- If the respondents are not prepared to co-operate with correct information, collected data can become unreliable.
- Result of the survey report is applicable only to the particular area of survey and cannot be generalized.
- Cannot be useful for historical problems as people cannot provide much information from their memory.

Data are the basis of hypothesis testing. Data are collected with the help of various methods for finding out the causes of a problem. It is usually done in an impartial manner to avoid any bias.

Type of Data

Based on quality and numerical values

- Qualitative data e.g. honesty, intelligence, color, Yes/No responses, etc.
- Quantitative data: discrete and continuous

Discrete data

Data having values of full numbers and not in fraction, for example, number of children in a family, number of workers in a factory, etc.

Continuous data

Data having a highest and lowest level with equal value of interval. For an example: number of farmers of a given area according to their farm income.

Farm Income	No. of farmers
3000 – 4000	35
4001 – 5000	30
5001 – 6000	25
6001 – 7000	20

Based on source

- ✓ Primary data
- ✓ Secondary data

Primary Data

Data are collected first time through direct investigation and they are original in nature.

Secondary Data

Data which are already collected by someone for some purpose and are kept as records for future needs. For instance, the data collected during agricultural census are primary data to the department conducted the census, however when the same data are used by a researcher for another study considered as secondary data.

Compilation of Data

The act of selection and putting together the gathered data is referred to as compilation of data. Compilation includes (a) coding, and (b) categorization or classification of raw data.

Coding

The purpose of coding is to translate raw data into symbols which may be counted and tabulated. Coding involves the assigning of symbols or numerals to each response so that responses can be put into a limited number of categories or classes. Such classed should be appropriate to the research problem under consideration. They must also possess the characteristic of exhaustiveness (i.e. there must be a class for every data item), and that of mutual exclusivity, i.e. a specific answer can be placed in one and only one cell in a given category set. Another rule to be observed is that of unidimensionality, that is, every class is defined in terms of only one concept.

Categorization or classification of raw data

Categorization of raw data is not only essential but it is unavoidable, if it is desired that the data should be put to any use. Without such a categorization, whole data will remain jumbled, complex and incomprehensible. The main object of categorization is to arrange the data in some scientific manners. Categorization or classification of data is the process of, arranging data into homogeneous groups or classes according to resemblances and similarities. Data having a common characteristic are placed in one class and in this way the entire data get divided into a number of groups or classes. Proper classification helps proper tabulation.

A proper set of categories must satisfy the following three conditions: i) there should be a single classificatory principle, ii) the categories must be exhaustive and sufficient for classifying all responses, and iii) the different categories must be distinct, separate, and hence, mutually exclusive. The categories have to be independent and not overlapping.

Depending upon the nature of the phenomenon involved, classification can be one of the following two types:

Classification according to attributes

As stated above data are classified on the basis of common characteristic which can either be descriptive (e.g. literacy, gender, honesty) or numerical (e.g. weight, height, income). Descriptive characteristics refer to quantitatively only their presence or absence in an individual item can be noticed. Data obtained in this way, on the basis of certain attributes, are known as "statistics of attributes" and their classification is said to be classification according to attributes.

Classification according to class-intervals

Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomenon which can be measured through some statistical units. Data relating to income, age, production, etc. come under this category. Such data are known as "statistics of variables" and are classified on the basis of class intervals. For example, the average monthly income of the marginal farmers in a community can be classified as follows:

Average monthly income (Taka)	Category
Upto – 1,000	Very low
1,001 – 2,000	Low
2,001 – 3,000	Medium
3,001 – 4,000	High
4,001 – 5,000	Very high

Tabulation of Data

Tabulation is the process of presenting data in tables. It has been differently defined by different scholars but the basic-objective of tabulation remains to clarify and simplify the data collected so as to make it socially understandable.

Objectives of the tabulation

- ✓ Tabu1ation provides clarification on the characteristics of data,
- ✓ It conserves space and reduces explanatory and descriptive statement to a minimum,
- ✓ It facilitates the process of comparison,
- ✓ It facilitates the summation of items and the detection of errors and omissions,
- ✓ It depicts the trend and tendencies of the problem under consideration,
- ✓ It provides a basis for various statistical computations.

Rules for tabulation

The construction of a good statistical table is a specialized art and requires great skill, experience and common sense on the part of the tabulator. Tabulation can be done by hand or by mechanical or electronic devices. The choice depends on the size and type of study, cost considerations, time pressures and the availability of tabulating machines or computers. Tabulation rules may be divided into two groups:

(a) Rules relating to table structure and (b) General rules.

Rules relating to table structure

The following parts must be present in all tables: (i) Table number, (ii) Title, (iii) Head note, (iv) Caption, (v) Stubs, (vi) Body of the table, (vii) Foot-note, (viii) Source-note.

General Rules

- ✓ Tabu1ation provides clarification on the characteristics of data.
- ✓ The table should be simple and compact. It should not be overloaded with details.
- ✓ Every table should be given a distinct number to facilitate easy reference.
- ✓ The captions and stubs in the tables should be arranged in a systematic manner. This arrangement could be alphabetical, chronological, geographical, conventional, etc.
- ✓ The units of measurement under each heading or sub-heading must always be indicated; for example, height in meters, weight in kilograms, etc.
- ✓ Explanatory footnotes should be placed directly beneath the table with the reference symbols used in the tables.
- ✓ Suitable approximation may be adopted.
- ✓ Ditto (“ ”) marks should not be used that may be mistaken.
- ✓ Miscellaneous and exceptional items, if any, should usually be placed in the last row of the table.
- ✓ A table should be complete and self-explanatory.
- ✓ As the table forms a basis for statistical analysis, it should be accurate and free from all sorts of errors.
- ✓ Proper lettering will help adjust the size of the table.
- ✓ If it is a big table, it will lose its simplicity and understandability, and in such a case, it should be broken into two or three tables.

Analysis of Data

Statistical analysis means the computation of certain indices or measures along with searching of patterns of relationship that exist among the data groups. Research findings very much depend on the analysis of data. Analyses of data sums up the findings and provide answers to questions for which research was conducted. It is therefore, essential that the data should be carefully analyzed and the researcher should see whatever comes out of the analysis is properly presented. It is at the analysis stage that the decision on retention or discarding of hypothesis will take place. At this stage the researcher also may develop hypothesis of his own and a smallscale research that may lead to very important findings. Usually there are two types of analysis: descriptive analysis and inferential analysis/statistical analysis.

Descriptive analysis

It largely the study of distributions of one variable and is helpful to summarize the collected data. It also provides us with the profiles of companies, work groups, individuals and other subjects on any of a multitude of characteristics such as size, composition, efficiency, preferences, etc.

Inferential analysis/Statistical analysis

It is concerned with the estimation of population values. It is also concerned with the various tests of significance for testing hypotheses. The task of interpretation (that is, the task; of drawing inferences and conclusion) is performed on the basis of inferential analysis.

Inferential or statistical analyses are of two types: a difference testing statistical analysis and a relationship testing statistical analysis.

In difference testing statistical analysis, comparisons of two groups are made, that is, significant differences between two groups are observed. The usual differences testing statistical analysis are: (a) χ^2 (Chi-squared) test, (b) Mc.Nemar test, (c) Sign test, (d) the Walsh test, etc.

In a relationship testing statistical analysis, statistical tests are computed to determine the relationship between two or more variables. Also, it is possible to determine the effects of one variable on the other through this analysis. The usual relationship testing statistical analysis are: (a) Correlation analysis, (b) Regression analysis (c) Analysis of variance (ANOVA). (c) Factor analysis, (e) Path analysis, etc.

CODE PLAN

Variables	Code plan	Column used
Age	Actual figure	4,5
Education	1 for 1 year of schooling	6-9
	.5 for can sign only	
Farm size	Actual figure	10-14
Annual income	1 for 1000 Tk.	15-17
Cosmopoliteness	0 for not at all	18
	1 for rarely	
	2 for occasionally	
	3 for often	
	4 for regularly	
Extension contact	0 for not at all	19
	1 for rarely	
	2 for occasionally	
	3 for often	
	4 for regularly	
Organization participation	1 for 1 year as general member	20,21
	2 for 1 year as executive member	
	3 for 1 year as executive officer	
Agril. knowledge	Actual figure	22-23

N.B. Column 1, 2, 3 are used for responding number

CODE SHEET

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
0	0	7	2	7	0	7	.	0	0	0	.	7	0	0	7	0	1	1	0	4	6	7			
0	1	9	4	2	0	0	.	0	0	1	.	5	0	1	2	0	0	0	0	0	4	0			
0	3	8	5	2	0	0	.	5	0	0	.	6	5	0	6	5	2	2	0	2	5	2			
0	5	2	3	8	1	0	.	0	0	1	.	0	0	1	5	0	3	3	1	5	7	7			
0	6	9	3	3	1	4	.	0	1	0	.	0	0	5	0	0	4	4	0	6	8	2			
0	8	1	2	7	0	7	.	0	0	0	.	7	0	0	7	0	1	1	0	4	6	7			
0	9	2	4	2	0	0	.	0	0	1	.	5	0	1	2	0	0	0	0	0	4	0			
0	9	9	5	2	0	0	.	5	0	0	.	6	5	0	6	5	2	2	0	2	5	2			
1	0	7	3	8	1	0	.	0	0	1	.	0	0	1	5	0	3	3	1	5	7	7			
1	1	9	3	3	1	4	.	0	1	0	.	0	0	5	0	0	4	4	0	6	8	2			

CATEGORIZATION

Variables	Categories	Base
Age	Young age	Up to 30
	Middle age	31-45
	Old age	Above 45
Education	No education	Can't read and write
	Primary education	Up to class V
	Secondary education	Class VI to S.S.C
	Higher education	Above S.S.C
Farm size	Small	Below 1.00 ha.
	Medium	1.00 – 2.00ha
	Large	Above 2.00ha
Annual income	Low	Below 60,000Tk.
	Medium	60,0000-120,000
	High	Above 120,000
Organizational participation	Low	0-3
	Medium	4-10
	High	Above 10
Extension contact	Not at all	0
	Rarely	1
	Occasionally	2
	Often	3
	Regularly	4
Cosmopoliteness	Not at all	0
	Rarely	1
	Occasionally	2
	Often	3
	Regularly	4
Vegetables cultivation knowledge	Low	Below 60%
	Medium	60-80%
	High	Above 80%

TABULATION

Table 1. Distribution of the respondents according to their age

Categories	Number	Percent
Young age	2	20
Middle age	4	40
Old age	4	40
Total	10	100

Table 2. Distribution of the respondents according to their education

Categories	Number	Percent (%)
No education	2	20
Primary education	2	20
Secondary education	4	40
Higher education	2	20
Total	10	100

Table 3. Distribution of the respondents according to their farm size

Categories	Number	Percent (%)
Small	4	40
Medium	4	40
Large	2	20
Total	10	100

Table 4. Distribution of the respondents according to their annual income

Categories	Number	Percent (%)
Low	0	0
Medium	4	40
High	6	60
Total	10	100

Table 5. Distribution of the respondents according to their cosmopoliteness

Categories	Number	Percent (%)
Not at all	2	20
Rarely	2	20
Occasionally	2	20
Often	2	20
Regularly	2	20
Total	10	100