Chapter 14

Collection of Data

Research tools are administered on the sample subjects for collecting evidences or data. Most educational research will lead to the gathering of data by means of some standardized test or self-constructed research tools. It should provide objective data for interpretation of results achieved in the study. The data may be obtained by administering questionnaires, testing, personal observations, interviews and many other techniques of collecting quantitative and qualitative evidence.

The researcher must know how much and what kind of data collection will take place and when. He must also be sure that the types of data obtainable from the selected instruments will be usable in whatever statistical model he will latter use to bring out the significance of the study. The data collection is the accumulation of specific evidence that will enable the researcher to properly analyse the results of all activities by his research design and procedures. The main purpose of data collection is to verify the research hypotheses.

NEED FOR DATA COLLECTION

The data are needed in a research work to serve the following purposes:

1. Collection of data is very essential in any educational research to provide a solid foundation for it.
2. It is something like the raw material that is used in the production of data. Quality of data determines the quality of research.
3. It provides a definite direction and definite answer to a research inquiry. Whatever inquiry has to give a definite answer to an investigation. Data are very essential for a scientific research.
4. The data are needed to substantiate the various arguments in research findings.
5. The main purpose of data collection is to verify the hypotheses.
6. Statistical data are used in two basic problems of any investigation:
   (*a*) Estimation of population parameters, which helps in drawing generalization.
   (*b*) The hypotheses of any investigation are tested by data collection procedure.
7. The qualitative data are used to find out the facts and quantitative data are employed to formulate new theory or principles.
8. Data are also employed to ascertain the effectiveness of new device for its practical utility.
9. Data are necessary to provide the solution of the problem.

MEANING OF DATA

Data means observations or evidences. The scientific educational researches require the data by means

of some standardized research tools or self-designed instrument. Data are both qualitative and quantitative in nature.

Score is the numerical description of an individual with regard to some characteristics or variables. Measurement process is employed to quantify a variable. Data are collected for both variables as well as attributes. These are gathered in terms of frequency and scores. It depends on the type of instrument employed for its measurement. Generally tests yield the data in the form of scores and questionnaires provide the data in the form of frequency. Data are things with which we think of.

Data and facts are used in educational research, therefore, it is essential to understand them clearly.

DIFFERENCE BETWEEN FACTS AND DATA

The facts and data have been distinguished in the following manner:

1. The facts are organized in their original form whereas data are organized in systematic order.
2. The facts do not have any coherence of system whereas data have an organic unity like body.
3. The facts are difficult to interpret. The interpretation of facts is usually subjective and employs the imagination of the researcher. Data can be interpreted easily and most objectively.
4. The facts are mysterious in nature we have to explore the facts but data have no mystery at all.
5. The facts are descriptive in nature whereas data are explanatory.
6. The facts are not amenable to objective statistical treatments whereas data can be easily subjected to in objective statistical treatment.
7. The facts may not be directly the basis of findings or research conclusions, but data are directly linked with research conclusions.
8. The facts are usually too broad and are not linked and a purposive way whereas data are always collected with a sense of purpose.
9. The facts are collected in historical or survey research whereas data are gathered in a scientific and experimental research.

NATURE OF DATA

The research studies in behavioural science or mainly concerned with the characteristics or traits. Thus, tools are administered to quantify these characteristics, but all traits or characteristics can not be Quantified.

The data can be classified into two broad categories:

1. Qualitative data or attributes.
2. Quantitative data or variables.

1. Quantitative Data or Attributes: The characteristics or traits for which numerical value can not be assigned, are called attributes, e.g. motivation, confidence, honesty integrity etc.

2. Quantitative Data or Variables: The characteristics or traits for which numerical value can be assigned, are called variables, e.g. Achievement Intelligena, Aptitude Height, Weight etc.

The distinction is based on the process of measurement rather than on the properties inherent in the phenomenon or trait, for generally properties considered qualitative can be made quantitative by measuring them with an instrument designed to assign numerical values to the various degrees to which they exist.

The decision to research a given phenomenon on the basis of its attributes or on the basis of its quantitative aspects is frequently a matter of choice, depending on such considerations as the need of precision and the ease of manipulation of data.

In fact, the quantification of phenomenon is generally considered essential to the progress of a science particularly at the more advanced levels. Quantification provides a greater refinement and possesses definite advantages by virtue of its statistical treatment.

The quantitative data provide the nature of the characteristic or trait. They have the verbal exposition of the trait. There is much scope for logical manipulation is the interpretation of result. The trait is not quantifiable.

The qualitative data provide the extent and nature of the distribution of the trait or variable measured. The tools are available to measure the variable. In the experimental research data are collected in the controlled situation to study the functional relationship of variables.

Quantification is the process of assigning numerical values to the trait of the subjects of sample which normally would be quantitative. This can be done by :

(*a*) Observations or information by first hand experience. It is used in small children and animals.

(*b*) Systematic collection and analysis of factual data. This is done in historical research.

(*c*) Scales and inventories are designed to explore or reveal the interests, attitude and personality. It is used in the case of study and survey research.

(*d*) Questionnaire, interview and opinionnaire are designed to gain information. This is employed in survey research.

(*e*) Educational and psychological tools are administered to quantify the variables more accurately. These tools are used in scientific research studies.

Various types of research tools are employed to collect the data. These tools yield different types of data.

CONSTANTS

A constant is all characteristic or condition that is the same for all the observed units or sample subjects of a study.

A variable, on the other hand is a characteristic which takes on different values for different sample subjects or for all the observed units.

The use of variable and constant has been illustrated with the help of the following example. Suppose a study is conducted for determing the effect of three different teaching methods upon the achievement in secondary mathematics. Each of three ninth grade maths sections in the same school, are taught by the same teacher, is taught using one of the methods. Both boys and girls are included in each method.

In this study grade level, school and teacher are recognized as constants. It assumes that the teacher can hold constant teaching effectiveness except for the method. The independent variable in the study is the teaching method and achievement in maths is known as the dependent variable or criterion variable.

A dependent variable is the consequent of the independent variable. The functional relationship is analysed between the two variables. The precision of the data is governed by the constants, if these conditions remain unchanged during experimentation. A researcher should be careful or conscious enough about the constants of his investigation.

VARIABLES

The variables are those which vary from person to person and can be quantified by employing measuring instrument. The sample or group variation can be ascertained in terms of numerical values. The characteristic or the trait in the behavioural science which can be quantified is termed as variable.

Variables can be classified into two categories:

1. Continuous variables.
2. Discrete variables.
1. Continuous variables are those for which fractional value exists and have meaning e.g. age, weight, achievement, where 14.5 years, 62.75 kgs and 45.50 scores or any other fractional of a whole unit is logical and measurable within the precision of the instrument used.
2. Discrete variables are those on the other hand, which exist only in units not the fractional value (usually units of one) e.g. 30 boys, 25 girls, 40 Indians and 24 Americans.

This distinction is somewhat more complicated in practice. The typical problem in educational research deals with test scores. These are generally reported as discrete variable though they are often fundamentally continuous. Intelligence is recorded in terms of I.Q.s. as discrete thought by their very computation they are technically continuous.

In research, where the concern is with group measures which almost invariably are fractional, continuous variables appear somewhat more acceptable than discrete variables.

The variables can be classified with regard to their roles or functions In particular study. The assumptions of an investigation determine the role of the variables. The following are the types of variables: independent, dependent or criterion, experimental, control, moderator and intervening variables. When the investigator is concerned with the teachers attitude toward teaching in relation to their classroom verbal interaction. The teaching attitude is the independent variable. In another study the investigator intends to analyse the relationship between classroom interaction and student's achievement. The classroom interaction is the independent variable (whereas in earlier study it is dependent variable) and student's achievement is the dependent variable. Thus, this type of classification is important from a particular research point of view. Every research worker must understand his variables and their roles in his investigation.

VARIATE

The variable is quantified by using an instrument. The quantified variable is termed as variate. When sample subjects I.Q.s or scores of achievement are collected, it is known as variate. The statistical analysis involves variate analysis: uni-variate, bi-variate, multi-variate analysis. It is evident that data collection means to convert variables into variates so that data can be subjected to an appropriate statistical analysis for obtaining the results.

QUANTIFICATION OF VARIABLE

Quantification is the process of assigning numeral value to the extend or amount of a variable of an individual. The quantification is done by employing the process of measurement. This process yields data and scores.

Many studies in education produce data for the verification of research hypotheses so as to draw conclusions. Some other studies in education produce evidences that require evaluation by subjective methods which do not readily permit the use of statistical analysis are termed as qualitative data. These qualitative data may be converted first into their quantitative data. The statistical techniques may be applied to test the significance of data. If qualitative data cannot be converted into quantitative data, a descriptive interpretation is done.

CHARACTERISTICS OF QUANTITATIVE DATA

The quantitative data are collected by administering the research tools. These should possess the following characteristics:

1. The quantitative data should be collected through standardized tests. If self-made test is used it should be reliable and valid.
2. They are highly reliable and valid. Therefore, generalization and conclusions can be made easily with certain level of accuracy.
3. The obtained results through quantitative data can be easily interpreted with scientific accuracy. The level of significance can also be determined.
4. The scoring system of quantitative data is highly objective.
5. The use of quantitative data is always based upon the purpose of the study. The specific psychometric tests are used in difficult investigation.
6. The inferential statistical can be used with the help of quantitative data.
7. The precision and accuracy of the results can be obtained by using quantitative data in an educational research.

TYPES OF DATA

There are four basic ways of quantifying the variables. They are also called levels of measurement or scales of measurement. These are commonly referred to as:

1. Nominal scale,
2. Ordinal or rank scale,
3. Equal-interval scale, and
4. Ratio scale.

1. Nominal Scale

The nominal scale is the least precise or crude of the four basic scales of measurement. It simply implies the classification of an item into two or more categories without any extent or magnitude. There is no particular order assigned to them. The frequency or numbers are used to give a name to something that may be used for determining per cent, mode. For example boys and girls; pass and fail; rural and urban.

In classroom observation the measurement is done at nominal scale. The teaching and instruction are organized considering the mode of the students, because teacher cannot pace with each and every student in his teaching and learning process.

2. Ordinal Scale

The ordinal scale is more precise scale than the nominal scale. It allows the teacher to assign values by placing of arranging the observations in relative rank order. No value is assigned to the distances to the positions of ranking. This scale assigns observations to categories by number and arranges them in some logical order. It does not require the relationship of equivalence but also requires one observation to be greater or lesser than the other.

This scale is used frequently in the schools for prize distribution and to provide the motivation by the technique of competition. In asking the questions teacher considers the place of students in the class.

3. Equal Interval Scale

The equal interval scale is more precise and refined scale than nominal and ordinal scales. This scale has all the characteristics and relationship of the ordinal scale, besides which distances between any two numbers on the scale are known. The zero point and the unit of measurement used on the scale are arbitrary assumed. A linear relationship is established in the equal-interval scale.

The equal-interval scale has the greater use in teaching-learning situation, educational administration, educational guidance and counselling and educational research. The effectiveness of any instructional procedure, can be evaluated precisely by collecting the data on this scale. The measurement in education is usually done on equal interval scale. The dependable inferences are drawn in educational research by collecting evidences on equal interval scale.

These three types of data are mainly used in behavioural researches. These have been illustrated in the following table:

Subject	Nominal Sex	Ordinal Achievement in ranks	Equal-interval Achievement in scores
A	Boy	3rd	64
B	Girl	1st	76
C	Girl	2nd	68
D	Boy	5th	58
E	Girl	4th	60

These scales have not absolute 'zero'. The group performance is the reference point.

4. Ratio Scale

Ratio scale has the properties of equal-interval scale plus two additional characteristics:

(*a*) This scale has a true, rather than arbitrary 'zero'. It is possible to indicate the complete absence of property. The zero point on a centimeter scale indicates the absence of height. However, the zero point on a Fahrenheit temperature scale does not indicate the absence of temperature; this is a scale with zero defined arbitrarily.

(*b*) The ratio scale numerals have the qualities of real numbers, and can be added, subtracted, multiplied; or divided. A may be so many units greater than B and may also be so many times

as great as B. Fifteen grams is three times five grams and fifteen grams is ten more than five grams.

Proceeding from nominal to ratio scale in order each type makes possible more information about the property described. If the variable permit its application, the type of scale provide the maximum amount of information should be used.

DATA COLLECTION

In the behavioural science data are collected by administering various types of research tools of the human sample subjects. The different traits and characteristics are quantified by using measurable instruments. These research tools provide different types of data. It is very essential for a research worker to understand the trait, tools and type of data. This has been illustrated with the help of the following table:

No.	Level	Properties		Assumptions	Examples	Statistics
1.	Nominal Scale	Classification Equate Non-equate	(1)	All members of a set are assigned the same numeral, and no two sets are assigned the same numeral.	Car registration plate numbers Simple questionnaire and interviews data gathered on an all or none basis	Mode Coefficient of association
2.	Ordinal Scale	Classification Order Equate Non-equate	(1) (2)	As above Objects can be rank ordered on the basis of an operationally defined characteristic or property	Moh's scale of hardness Most psychological and educational test scores.	Mode Median Percentiles Coefficient of association Ranking coefficient
3.	Interval Scale	Classification Order equal Units Equate Non-equate Add subtract Multiply divide	(1) (2) (3)	As above As above Distances on the scale represent equal intervals	Temperature scales Very well validated intelligence tests, etc.	Mode Median Mean Percentile Range Standard Deviation Average deviation
4.	Ratio Scale	Classification Order Equal Units Absolute Zero Equate Non-equate Add subtract Multiply divide	(1) (2) (3) (4)	As above As above As above Scale has an absolute or natural Zero	Common, scales of length, Mass and time Some scale of loudness	Standard score Coefficient of association Ranking Coefficient Simple partial and multiple correlation.

A Classification of Scales of Measurement and Common Statistics

Types of Data With Reference to the Traits

Trait	Tool	Type of Data
1. Intelligence	Psychological tests	Equal-interval scale
2. Achievement	Educational tests	"
3. Aptitude	Psychological tests	"
4. Attitude	Scales	Ordinal scale
5. Interest	Inventories	Equal interval scale
6. Personality	"	"
7. Adjustment	"	"
8. Opinions of feelings	Questionnaire or Opinionnaire	Nominal scale

There are other types of research tools which are used to collect' the data. For example, 'observation technique' is most frequently used to collect the data which yields the data at nominal scale and also at equal-interval scale.

In the following table research tools have been classified with regard to scale for measurement and types of statistics may be employed of analysis purpose.

Classification of Tools and Research with Regard to the Level of Measurement

Scale	Tools of Research	Statistics
1. Nominal Scale	Questionnaire Interview Schedule Observation	Model, frequency percentage Simple statistics X_2-Test and 'C'
2. Ordinal scale	Scale observation Rank scale	Median, Spearman's rank correlation X^2 test Median test etc.
3. Equal-interval scale	Psychological and educational tests Observation scale	Mean, Sd. Pearson's correlation 't' test and 'f' test etc.
4. Ratio scale	Physical measurement	Arithmetic mean Pearson's correlation Mathematics is used.

The types of data depend on nature of research tool employed for this purpose. Statistical techniques are virtually selected by considering the nature of research tool and data collected at nominal, ordinal or equal interval scales. It is not the statistical technique but type data which determine the fate of research project.

Ethical considerations in Collection of Data

Any researcher who involves human sample subjects in his research has certain responsibilities towards them. Since the activities of the sample subjects are often closely associated with data collection process, it is appropriate to consider ethical considerations here.

The following points have to be considered in process of data collection:

1. The researcher must protect the dignity and welfare of human sample subjects.
2. The human sample subjects freedom to decline participation must be respected, and the confidentially of research data must be maintained.
3. The researcher must guard against violation or invasion of privacy.
4. The responsibility for maintaining ethical standard remains with the individual researcher and the principal investigator or supervisor is also responsible for actions of his scholars.

Any researcher anticipating "the use of human sample subjects should consult on 'ethics' statements such as those mentioned above. A researcher should not mention the name of subjects anywhere in the report. If possible name of institutions where sample subjects have selected for data collection should not be mentioned even in the appendix. The code number should be used for this purpose. As a general rule, he must respect the human sample subjects selected in his specific research study.

Precautions in Data Collection

In the data collection the following precautions should be observed:

1. The data must be relevant to the research problem.
2. It should be collected through formal or standardized research tools.
3. The data should be such as these can be subjected to statistical treatment easily.
4. The data should have minimum measurement error.
5. The data must be tenable for the verification of the hypotheses.
6. The data should be such as parameters of the population may be estimated for inferential purpose.
7. The data should be complete in itself and also comprehensive in nature.
8. The data should be collected through objective procedure.
9. The data should be accurate and precise.
10. The data should be reliable and valid.
11. The data should be such that these can be presented and interpreted easily.
12. The scoring procedure of the research tool should be easy and objective.

ORGANIZATION OF DATA

After the data have been collected. it must be organized and analyzed to draw proper inferences.

The mass of data collected through the use of various tools, however reliable, valid and adequate it may be is yet but raw. It needs to be systematized and organized i.e. edited classified and tabulated. before it can serve any worthwhile purpose. Editing implies the checking of gathered data for accuracy, utility and completeness. Classifying refers to the dividing of the information into different categories classes or heads, for the use. Tabulating denotes the recording of the classified material in accurate mathematical terms e.g. marking and counting frequency tallies for different items on which Information is gathered. Tabulation is a tedious and painstaking process and must be accurate. Before, tabulating all raw data should be tested on the basis of the purpose for which they are gathered and only the useful and usable data should be tabulated. Tabulating machines and other mechanical aids for tabulating are becoming current.

Once the data have been collected through the use of the measuring instruments, some type of tabulation and possibly a transformation of data in preparation for the analysis may be necessary. if answer sheets are to be hand scored routine precautions should be taken to practice for scores and accuracy checks while the actual scoring is being done.

Research study which includes the collection of considerable data using standardized tests should make provision for machine scoring.

There are IBM answer sheets with space of responses upto 150 items. Test scoring machines provide for obtaining the actual scores, they commonly provide tabulations, summaries, and conversions to various types of standard score. If some cases the machines are connected to a computer to provide certain kinds of analysis.

Machine scoring of tests is usually less expensive than hand scoring. Other advantages of machine scoring are its accuracy and its preparation of the data for computer analysis if necessary. Not all educational "research data are collected in a form that can be machine is not to be used to transmit data from the answer sheet to the computer is an important part of the research procedure.

The data should be so organized that minimum effort is required to transmit the data from its original form to the IBM card. The format data card, indicating the information that enters into each column, must be defined by the researcher. Information is commonly of two types: identification and responses of the subject. Any confusion should be eliminated to minimize the number of copy errors. The computations on the calculator should be so performed in such a manner that several internal checks can be made during the calculation.

EXERCISES

1. Define the term 'Data', How it is different from 'facts' and 'scores'?
2. Indicate the need for data collection and describe the nature of data.
3. (*a*) Differentiate between questionnaire data and quantitative data,

   (*b*) Distinguish among attribute, variable, constant and variate.
4. Enumerate the types of data and describe these types in detail with reference to tools, types of research and statistical techniques,
5. Indicate the ethical consideration in data collection and their justification in conducting a research.
6. What precautions are to be taken in collection of data? Indicate the limitations in data collection.
7. What do you understand by the term 'Organisation of data'? Indicate the need and importance of data organization in conducting educational research.