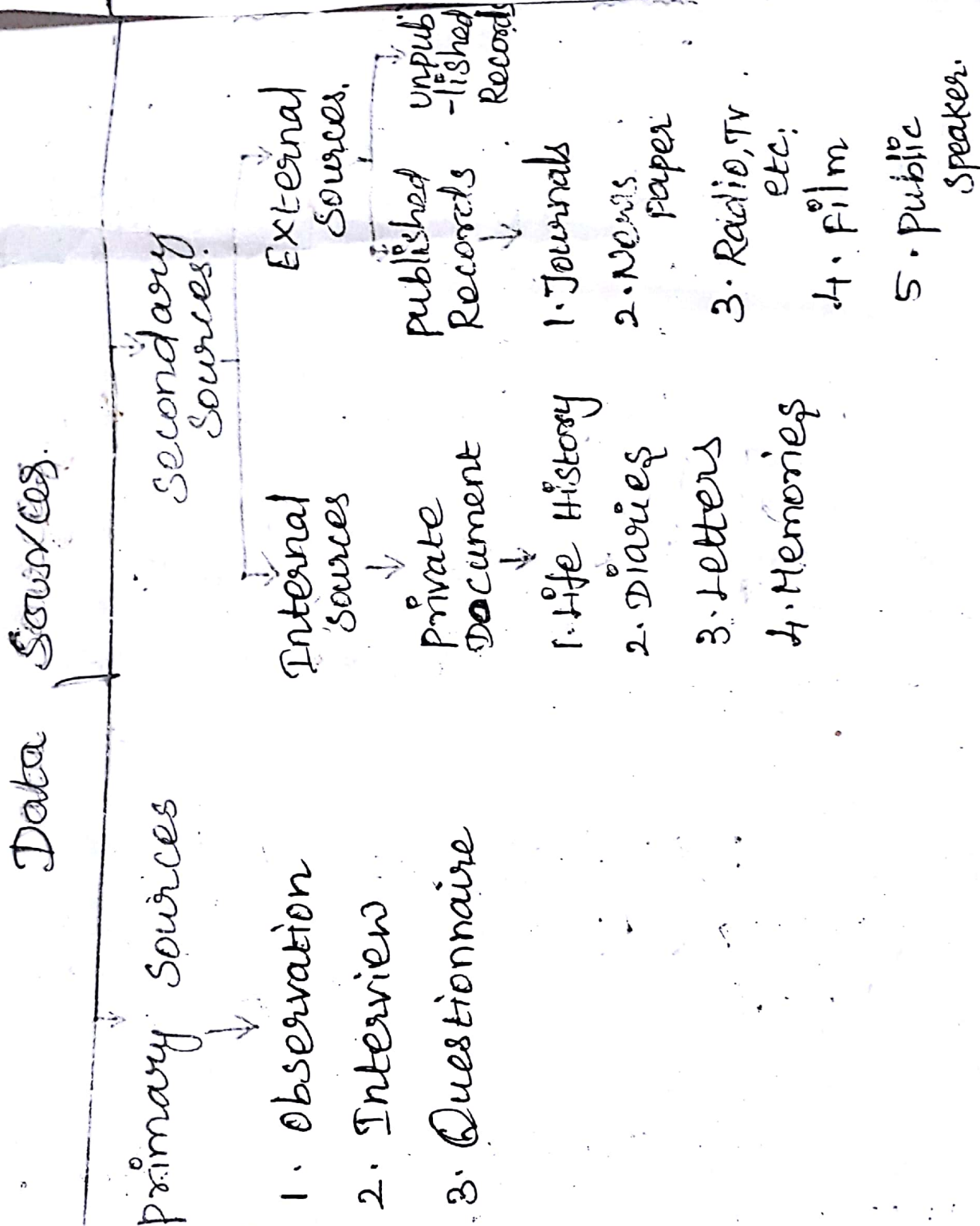


## Module - II

### Sources of Data, Types and Collecting Method



17) 1. Observation:-  
\* Observation means acquiring knowledge through use of sense of the organs.

\* In strict sense, sight (eye) is rather than ear and voice.  
\* Observation means it is watching minutely, slowly and systematically and occurrences with a situation and derived results.

\* Observation is a primary tool of scientific enquiry.

1. Onthropologist
2. Psychologist
3. Sociologist
4. Educationalist.

1. Onthropologist:-

\* Onthropologist observe the activities of folk ways, and models of tribe people.

2. Psychologist:-

\* Psychologist observe the psychological behaviour of child and adult.

3. Sociologist:-

\* Sociologist observe different aspects of the community.

#### 4. Educationalist:-

\* Educationalist observe the teacher - Pupil interaction within classroom and teacher effectiveness.

#### 17) Observation - Three Types of Components

1. Sensation (receiving)

2. Attention (focus)

3. Perception (interpretation)

#### 1. Sensation:- (నిర్గుణ భావ)

Man has six senses.

1. Sight

2. Hearing

3. Touch

4. Smell

5. Taste

6. Rationality.

#### 2. Attention:- (నిర్గుణ భావ)

1. Interest

2. Motivation

3. Novelty - కొత్త, అసాధారణ

4. Will power

5. Adequate training and

experience are determine factors of Concentration.

#### 3. Perception:- (నిర్గుణ భావ)

It is the interpretation of the Sensory report.



Interview:-

Importance of Observation:-

i) \* It is a simplest method than other method.

ii) \* It is non-technical method.

iii) \* It requires less training for collection of data.

iv) \* It is useful for formulating

hypothesis.

v) \* It is associated with greater accuracy.

vi) \* It is considered as more convenience results.

vii) \* Observation method is a common method to all sciences.

Interview:-

### Interview

\* It is one of the important tool of social science.

\* The person who is interviewing is called as interviewer.

\* An interviewee is called respondents.

\* Under this method all types of data can be gathered intensively and extensively.

\* So an interview is assessing the personality of the respondents.

### \* Advantages of Interview:-

1. Measuring personality.
2. Full information collected
3. Accuracy is possible.

#### Measuring personality:-

\* It is very useful tool to the social researcher.

\* It is assessing the opinion, motivation and reaction of the respondents.

#### Full information collected:-

\* Different types of data, information are gathered through interview.

\* Interviewer has questions getting clear answer.

#### Accuracy is possible:- (100%)

\* In questionnaire method it is accuracy is not possible.

\* Unwanted responses can be eliminated.

## Types of Interview:-

1. Structured interview
2. Unstructured interview
3. Focused interview
4. Repetitive interview.

### Structured Interview:-

\* It is controlled, guided  
 also known as direct interview  
 \* Here schedule is completely used.  
 \* Under this method the questions are ask and getting answers.

### Unstructured Interview:-

\* It is uncontrolled unguided and undirect interview.

\* It is also study about form of free discussion or free stories.

### Focused Interview:-

\* It is used for the study about social and psychological affects from mass communication.

### Example:-

Reaction of film show  
or Radio programme.

\* It is similar to unstructured interview.



## 17 Repetitive Interview:-

\* It is one of the <sup>more to</sup> cost of interview.

\* Some person living in different places and different houses.

\* They may be interviewed so many times.

## Schedule and Questionnaire:-

### Questionnaire:-

\* The term schedule and questionnaire are used as well as the same meaning by most of the researcher.

\* But there is a thin line difference between them.

\* This difference is based on Electronic differences.

\* The term schedules are those used in personal interviews and questionnaire is the form of distribution through mailed or post.

\* Schedule and the questionnaire are the indispensable tool for social science research.

### Collection of data through Questionnaire

\* In this method the question is quite properly particularly in the case of big enquires.

\* In this method a questionnaire sent to the person by post or mailed concerned with the request the answer to the question and return the questionnaire.

\* In this method a questionnaire contents of number of questions printed or typed or definit in a order.

### Characteristics of Good Questionnaire

1. Size of questionnaire

2. Appearance

3. Clarity

4. Sequence of questions

5. catching interest.

#### 1. size of Questionnaire:-

\* The size of questionnaire is less than the size of schedule.

\* If the questionnaire is too length the respondent wants to be poor.

\* whereas the questionnaire may be shaped as post card.



## 2. Appearance:-

\*The questionnaire contents

1. A decent paper and printed
2. Attractive layout.
3. Grand letters and head.

\*The number of haphazard questions fail to create the interest of the respondent.

3. clarity:- <sup>It requires special attention.</sup>  
\* The question should be very clear & straight.

\*Required special attention in questionnaire is the problem in clarity.

\*In schedule personally present the researcher and interpretation of data.

\*But in questionnaire method it is not possible or avoided.

## 4. Sequence of questions:-

\*In questionnaire sequence of question are very essential than schedule. <sup>It is order in question and sequence of question.</sup>

\*while the question is very boring to the respondent in filling the question. when it is not in order or sequence.

## 5. catching Interest:-

\*The questionnaire must be very interest for catching the respondent. <sup>the questionnaire should be very interesting & attractive.</sup>

\*If it is possible that the

respondent carefully filled the questions and properly despatched.

### Advantages of Questionnaire:-

\* This method is a very cheaper in cost to the researcher.

\* Under this method the widely area or largely area may be covered.

\* This method relates with save in time.

\* This method associated with free-flows of respondent.

\* It is validity reliability and accuracy.

### Schedules:-

\* The schedule is the form containing some questions which are to be filled by the worker or researcher after collecting information from the respondent.

\* The schedule contain two type of questions:

1. Those are form of a question.
2. Those are form of table or blanks.

## Types of schedule:

1. Observation schedule.
2. Rating schedule
3. Document schedule.
4. Interview schedule.

### Observation schedule:-

\* Observation this method is used for observation purposes.

\* In this method, clearly pointing what is to be observed and how to be regard

### Rating schedule:-

\* It is used in sociological and psychological research guidance.

\* The attitude or opinions are to be measured.

### Document schedule:-

\* Collecting data from documents

They are:

1. Autobiography
2. Case History
3. Diary
4. Govt. Records.

### Interview schedule:-

\* It is used for the process of interview.

\* It contain some standard question.

\* Also it contain questions and filled answers filled by researcher.



Distinguish between questionnaire & schedule

	Questionnaire	Schedule
1. usage	usually filled up only by the respondents themselves.	Answers are recorded by enumerator.
2. Cost	It is relatively cheaper rate. Because it is sent through mail and post.	It is costlier one because of appointment of enumerators training and meeting.
3. Degree of response	All respondents may not response	All respondents are responsible for questions & answers.
4. Quality of response.	Not good	better quality
5. Time	More time No control of time → Very slow.	Time is controlled and restricted.
6. Coverage area	More area	Restricted area.
7. Personal Contact	personal contact is not possible	personal contact with respondents
8. Information	In complete and wrong information is possible	May be more accurate info from
9. Literate Educated	Need literate	No need for illiterate people to answer.

External data

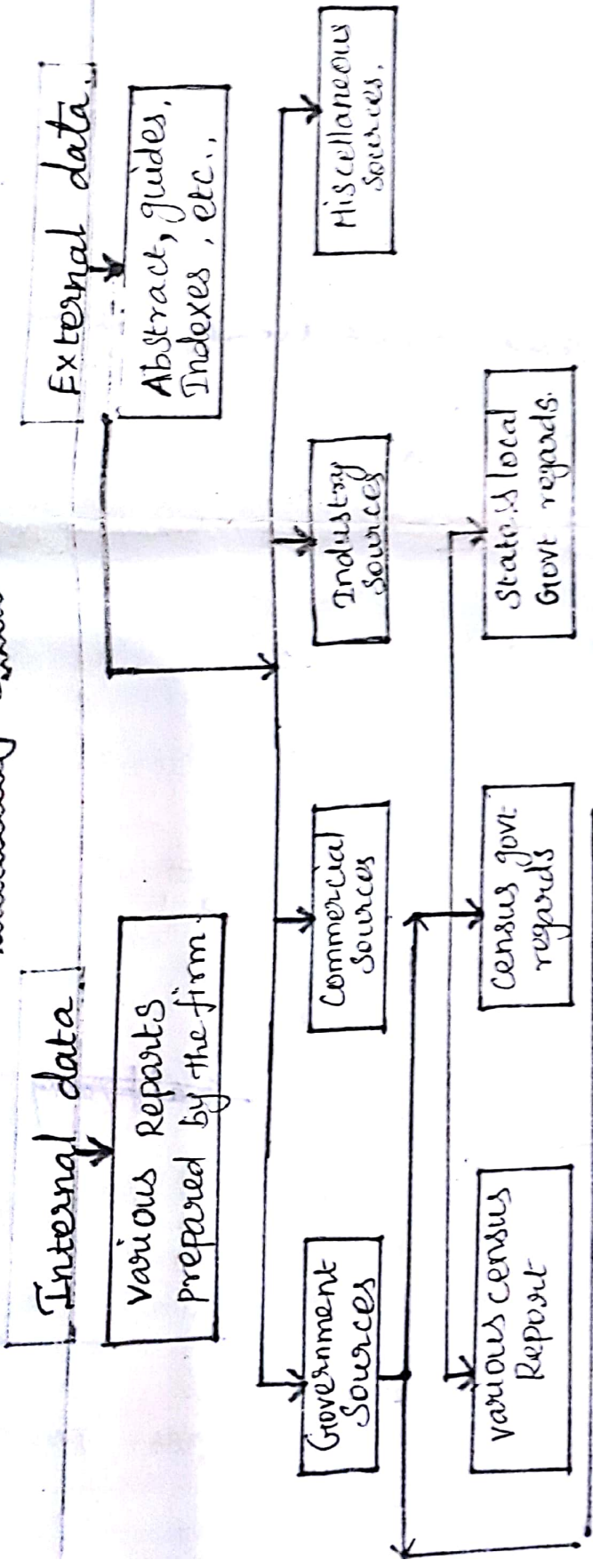
Secondary data

Internal data

date 7/10/21

# Sources of secondary data:

## Secondary Data



State, Counties, information.

→ Census of population.

→ Housing

→ Agriculture

→ Manufacture

→ Minerals

→ Industry.

→ Transportation.



The central government has following department aspects: Report of Control group

- ⇒ on health
- ⇒ Education of social welfare
- ⇒ Industries
- ⇒ Agriculture
- ⇒ Housing.

Commercial Sources:-

- ⇒ Indian Marketing Research Bureau of Hindustan Thomson Associates Limited.
- ⇒ Digital Money Market India
- ⇒ National Advertising Services.

[NAS].

1 Industry sources:- Indian Institute of Foreign Trade (IIFT)

\* Number of researcher visits this Institution.

\* Result releases in publication

\* publication by.

⇒ Monthly

⇒ Quarterly: ~~half yearly~~

⇒ yearly

Ex: Journal

All India Management Association

New Delhi:

\* Published Monthly Journal of

Indian Management.

\* It deals business problems and industrial aspects.



## Miscellaneous items of Sources:- (x)

\* It is related with individual researcher.

\* It includes

1. Dissertation (or) thesis
2. Monographs and periodicals.
3. Maintain records.
4. Maintain reports.

### 5.1 QUANTITATIVE AND QUALITATIVE DATA

Measurement is defined as a process of associating numbers or symbols to observations obtained in a research study. These observations could be qualitative or quantitative.

Most of the analysis can be conducted using quantitative data. For example, mean, standard deviation, etc., can be computed for quantitative characteristics. Qualitative characteristics can be counted and cannot be computed. Therefore, the researcher must have a clear understanding of the type of characteristic or variable before collecting the data. The observations on qualitative variables may also be assigned numbers. For example, we can record a person's marital status as 1, 2, 3 or 4 depending on whether the person is single, married, widowed or divorced. We can as well record "Yes or No" answers to a question as "0" and "1" (or as 1 and 2 or perhaps as 59 and 60). In this artificial or nominal way, categorical data (qualitative or descriptive) can be made into numerical data, and if we thus code the various categories, we refer to the numbers we record as nominal data.



*Nominal data* is numerical in name only, because they do not share any of the properties of the numbers we deal in ordinary arithmetic. For instance if we record marital status as 1, 2, 3, or 4 as stated above, we cannot write  $4 > 2$  or  $3 < 4$  and we cannot write  $3 - 1 = 4 - 2$ ,  $1 + 3 = 4$  or  $4 \div 2 = 2$ .

In those situations when we cannot do anything except set up inequalities, we refer to the data as *ordinal data*. For instance, if one mineral can scratch another, it receives a higher hardness number and on Mohs' scale the numbers from 1 to 10 are assigned respectively to talc, gypsum, calcite, fluorite, apatite, feldspar, quartz, topaz, sapphire and diamond. With these numbers we can write  $5 > 2$  or  $6 < 9$  as apatite is harder than gypsum and feldspar is softer than sapphire, but we cannot write for example  $10 - 9 = 5 - 4$ , because the difference in hardness between diamond and sapphire is actually much greater than that between apatite and fluorite. It would also be meaningless to say that topaz is twice as hard as fluorite simply because their respective hardness numbers on Mohs' scale are 8 and 4. The greater than symbol (i.e.,  $>$ ) in connection with ordinal data may be used to designate "happier than" "preferred to" and so on.

When in addition to setting up inequalities we can also form differences, we refer to the data as *interval data*. Suppose we are given the following temperature readings (in degrees Fahrenheit):  $58^\circ$ ,  $63^\circ$ ,  $70^\circ$ ,  $95^\circ$ ,  $110^\circ$ ,  $126^\circ$  and  $135^\circ$ . In this case, we can write  $100^\circ > 70^\circ$  or  $95^\circ < 135^\circ$  which simply means that  $110^\circ$  is warmer than  $70^\circ$  and that  $95^\circ$  is cooler than  $135^\circ$ . We can also write for example  $95^\circ - 70^\circ = 135^\circ - 110^\circ$ , since equal temperature differences are equal in the sense that the same amount of heat is required to raise the temperature of an object from  $70^\circ$  to  $95^\circ$  or from  $110^\circ$  to  $135^\circ$ . On the other hand, it would not mean much if we said that  $126^\circ$  is twice as hot as  $63^\circ$ , even though  $126^\circ \div 63^\circ = 2$ . To show the reason, we have only to change to the centigrade scale, where the first temperature becomes  $5/9 (126 - 32) = 52^\circ$ , the second temperature becomes  $5/9 (63 - 32) = 17^\circ$  and the first figure is now more than three times the second. This difficulty arises from the fact that Fahrenheit and Centigrade scales both have artificial origins (zeros) i.e., the number 0 of neither scale is indicative of the absence of whatever quantity we are trying to measure.

When in addition to setting up inequalities and forming differences we can also form quotients (i.e., when we can perform all the customary operations of mathematics), we refer to such data as *ratio data*. In this sense, ratio data includes all the usual measurement (or determinations) of length, height, money amounts, weight, volume, area, pressures etc.

The above stated distinction between nominal, ordinal, interval and ratio data is important for the nature of a set of data may suggest the use of particular statistical techniques\*. A researcher has to be quite alert about this aspect while measuring properties of objects or of abstract concepts.

## 5.2 CLASSIFICATIONS OF MEASUREMENT SCALES

From what has been stated above, we can write that scales of measurement can be considered in terms of their mathematical properties. The most widely used classification of measurement scales are: (a) nominal scale; (b) ordinal scale; (c) interval scale; and (d) ratio scale.

\*When data can be measured in units which are interchangeable e.g., weights (by ratio scales), temperatures (by interval scales), that data is said to be parametric and can be subjected to most kinds of statistical and mathematical processes. But when data is measured in units which are not interchangeable, e.g., product preferences (by ordinal scales), the data is said to be non-parametric and is susceptible only to a limited extent to mathematical and statistical treatment.



- (a) **Nominal scale:** Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no quantitative value. Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another. The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales. Chi-square test is the most common test of statistical significance that can be utilized, and for the measures of correlation, the contingency coefficient can be worked out.

Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes differences between things by assigning them to categories. Nominal data is, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc. In spite of all this, nominal scales are still very useful and are widely used in surveys and other *ex-post-facto* research when data is being classified by major sub-groups of the population.

- (b) **Ordinal scale:** The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. A student's rank in his graduation class involves the use of an ordinal scale. One has to be very careful in making statement about scores based on ordinal scales. For instance, if Ram's position in his class is 10 and Mohan's position is 40, it cannot be said that Ram's position is four times as good as that of Mohan. The statement would make no sense at all. Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.

Thus, the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than the difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non-parametric methods.

- (c) **Interval scale:** In the case of interval scale, the intervals are adjusted in terms of some rule that have been established as a basis for making the units equal. The units are equal only in so far as one accepts the assumptions on which the rule is based. Interval scales



can have an arbitrary zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin. The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure the complete absence of a trait or characteristic. The Fahrenheit scale is an example of an interval scale and shows similarities in what one can and cannot do with it. One can say that an increase in temperature from  $30^{\circ}$  to  $40^{\circ}$  involves the same increase in temperature as an increase from  $60^{\circ}$  to  $70^{\circ}$ , but one cannot say that the temperature of  $60^{\circ}$  is twice as warm as the temperature of  $30^{\circ}$  because both numbers are dependent on the fact that the zero on the scale is set arbitrarily at the temperature of the freezing point of water. The ratio of the two temperatures,  $30^{\circ}$  and  $60^{\circ}$ , means nothing because zero is an arbitrary point.

Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval. As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the 't' test and 'F' test.

- (d) **Ratio scale:** Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to be. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time. For example, the zero point on a centimeter scale indicates the complete absence of length or height. But an absolute zero of temperature is theoretically unobtainable and it remains a concept existing only in the scientist's mind. The number of minor traffic-rule violations and the number of incorrect letters in a page of type script represent scores on ratio scales. Both these scales have absolute zeros and as such all minor traffic violations and all typing errors can be assumed to be equal in significance. With ratio scales involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu." The ratio involved does have significance and facilitates a kind of comparison which is not possible in case of an interval scale.

Ratio scale represents the actual amount of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values. Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated. Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly. If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.