Research Methodology

Sushmitha P

April 2025

Census

- All items in any field of inquiry constitute a 'Universe' or 'Population.'
- A complete enumeration of all items in the 'population' is known as a census inquiry.
  - All items covered and highest accuracy obtained
  - Practically difficult, time consuming, resource consuming
  - Method beyond reach of ordinary researchers
  - Even government adopts this rarely

Sample Survey

- When the universe is a small one, no use of a sample survey.
- Practically, considerations of time and cost lead to selection of a few items.
- Representatives of the total population called 'sample'. The selection process is called 'sampling technique.' The survey so conducted is known as 'sample survey'.
  - A smaller group of n members is selected from the population of N according to some rule for studying some characteristic of the population
  - Researcher must plan how a sample should be selected and of what size such a sample would be

Sample Design

- (What?) A definite plan to obtain a sample from a given population.
- (What?) The technique or the procedure adopted in selecting items for the sample.
- (How?) Number of items to be included (size of the sample)
- (When?) Determined before collecting data.
- (How?) Choose the one that is most appropriate to your study.

Steps in sample design

- Type of universe
  - Define the set of objects (Universe)
  - Universe may be finite or infinite (population of a city, listeners of a specific radio programme)
- Sampling unit
  - Decision to be taken about a sampling unit before selecting the sample
  - Sampling unit may be geographical (state, city, village) or a construction unit (house, flat) or a social unit (family, club) or individual

Steps in sample design

- Source list
  - Also known as 'sampling frame' from which sample is to be drawn
  - contains the names of all items of a universe (for finite universe)
  - If not available, researcher should prepare it
  - Such a list should be comprehensive, correct, reliable and appropriate
  - important for the list to be as representative of the population as possible.
- Size of sample
  - The number of items to be selected from the universe
  - A major problem before a researcher. Sample size should be optimum.
  - An optimum sample fulfills the requirements of efficiency, representativeness, reliability and flexibility.
  - Decide the acceptable confidence level for the estimate
  - Larger variance, bigger sample
  - Size of population, parameters of interest in the research study, budgetary restrictions to be kept in view

Steps in sample design

- Parameters of interest
  - specific population parameters which are of interest (proportion of persons with some characteristic in the population)
- Budgetary constraint
  - Affect not only size of the sample but also the type of the same
- Sampling procedure
  - Decide the type of sample (the technique to be used in selecting the items)
  - select the design which, for a given sample size and for a given cost, has a smaller sampling error

Criteria of selecting a sampling procedure

- Two costs involved : cost of collecting the data (systematic bias), cost of an incorrect inference resulting from the data (sampling error)
- Systematic bias
  - results from errors in the sampling procedures
  - cannot be reduced or eliminated by increasing the sample size
  - Best that be done is identify the causes responsible and try to correct
- Sampling error
  - Random variations in the sample estimates around the true population parameters.Occur randomly and are equally likely
  - Compensatory nature and the expected value of these errors are zero
  - Can be measured for a given sample. Measurement of sampling error : 'precision of the sampling plan'
  - Decreases with increase in sample size which in turn increases precision but comes with its own limitations.
  - Select a better sampling design which has a smaller sampling error for a given sample size at a given cost

Factors for Systematic Bias

- Inappropriate sampling frame (biased representation of the universe)
- Defective measuring device (the questionnaire or the interviewer is biased)
- Non-respondents (unable to sample all the individuals initially included in the sample)
- Indeterminancy principle (individuals act differently when kept under observation)
- Natural bias in the reporting of data (Income information based on survey (for tax we tend to give lower income, for social status higher) Generally in psychological surveys, people tend to give what they think is the 'correct' answer rather than revealing their true feelings)

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

- Sample design must result in a truly representative sample.
- Sample design must be such which results in a small sampling error.
- Sample design must be viable in the context of funds available for the research study.
- Sample design must be such so that systematic bias can be controlled in a better way.
- Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

Types of Sample Design

- Different types of sample designs based on two factors : representation basis, element selection technique
- On the representation basis, the sample may be probability sampling (random selection) or it may be non-probability sampling (non-random selection).
- On element selection basis, the sample may be either unrestricted (each sample element is drawn individually from the population at large) or restricted (all other forms of sampling)

Element selection technique	Representation basis: Probability sampling	Representation basis: Non-probability sampling
Unrestricted sampling	Simple random sampling	Haphazard sampling or convenience sampling
Restricted sampling	Complex random sampling (such as cluster sampling, systematic sampling, stratified sampling etc.)	Purposive sampling (such as quota sampling, judgement sampling)

Non-probability Sampling

- does not afford any basis for estimating the probability that each item in the population has of being included in the sample
- different names: deliberate, purposive, judgement
- If economic conditions of people living in a state are to be studied, a few towns and villages may be purposively selected for intensive study on the principle that they can be representative of the entire state
- judgement of the organisers plays an important part in this sampling design
- investigator may select a sample which shall yield results favourable to his point of view
- no assurance that every element has some specifiable chance of being included
- Sampling error in this type of sampling cannot be estimated and the element of bias, great or small, is always there

Probability Sampling

- every item of the universe has an equal chance of inclusion
- also known as ‘random sampling’ or ‘chance sampling’
- Results obtained can be assured in terms of probability, that is errors of estimation can be measured
- Follows law of Statistical Regularity (if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe)

HOW TO SELECT A RANDOM SAMPLE?

- write the name of each element of a finite population on a slip of paper, put the slips of paper so prepared into a box or a bag and mix them thoroughly and then draw (without looking) the required number of slips for the sample one after the other without replacement
- This procedure will also result in the same probability for each possible sample
- What happens when the sample is infinite? Same as sampling with replacement from a finite population and our sample would be considered as a random sample if in each draw all elements of the population have the same probability of being selected and successive draws happen to be independent.

COMPLEX RANDOM SAMPLING DESIGNS

Systematic sampling

- Select every i^{th} item on a list
- Element of randomness is introduced by using random numbers to pick up the unit with which to start
- For instance, if a 4 per cent sample is desired, the first item would be selected randomly from the first twenty-five and thereafter every 25th item would automatically be included in the sample.
- Only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals
- Not a random sample in the strict sense, but it is considered reasonably random.

COMPLEX RANDOM SAMPLING DESIGNS

Systematic sampling

- Improvement over a simple random sample in as much as the systematic sample is spread more evenly over the entire population
- Easier and less costlier for large populations
- If there is a hidden periodicity in the population, systematic sampling will prove to be an inefficient method of sampling. (say every 25th item produced is defective)
- If the population list is in random order, systematic sampling is considered equivalent to random sampling
- Used when lists of population are available and they are of considerable length.

COMPLEX RANDOM SAMPLING DESIGNS

Stratified sampling

- Used when population from which a sample is to be drawn does not constitute a homogeneous group
- The population is divided into several sub-populations that are individually more homogeneous than the total population (the different sub-populations are called 'strata') and then we select items from each stratum to constitute a sample.
- Since each stratum is more homogeneous than the total population, we are able to get more precise estimates for each stratum
- By estimating more accurately each of the components, we get a better estimate of the whole.
- More reliable and detailed information is obtained.

Stratified sampling

- How to form strata?
- How should items be selected from each stratum?
- How many items be selected from each stratum or how to allocate the sample size of each stratum?

Stratified sampling

- How to form strata?
  - formed on the basis of common characteristic(s) of the items to be put in each stratum
  - various strata be formed in such a way as to ensure elements being most homogeneous within each stratum and most heterogeneous between the different strata
  - purposively formed and are usually based on past experience and personal judgement of the researcher
  - pilot study may be conducted for determining a more appropriate and efficient stratification plan

Stratified sampling

- How should items be selected from each stratum?
  - Simple random sampling
  - Systematic sampling if it is considered more appropriate

COMPLEX RANDOM SAMPLING DESIGNS

Stratified sampling

- How many items be selected from each stratum or how to allocate the sample size of each stratum?
  - the method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata
  - if P_i represents the proportion of population included in stratum i, and n represents the total sample size, the number of elements selected from stratum i is nP_i.
  - Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic

COMPLEX RANDOM SAMPLING DESIGNS

Stratified sampling

- How many items be selected from each stratum or how to allocate the sample size of each stratum?
  - in case the purpose happens to be to compare the differences among the strata, then equal sample selection from each stratum would be more efficient even if the strata differ in sizes
  - In cases where strata differ not only in size but also in variability and it is considered reasonable to take larger samples from the more variable strata and smaller samples from the less variable strata, we can then account for both (differences in stratum size and differences in stratum variability)

COMPLEX RANDOM SAMPLING DESIGNS

Cluster sampling

- If the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these smaller areas (usually called clusters), with the ultimate sample consisting of all (or samples of) units in these small areas or clusters
- the total population is divided into a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample

COMPLEX RANDOM SAMPLING DESIGNS

Cluster sampling

- Suppose we want to estimate the proportion of machine parts in an inventory which are defective. Also assume that there are 20000 machine parts in the inventory at a given point of time, stored in 400 cases of 50 each. Now using a cluster sampling, we would consider the 400 cases as clusters and randomly select 'n' cases and examine all the machine parts in each randomly selected case.
- reduces cost by concentrating surveys in selected clusters
- Less precise than random sampling
- Cluster sampling is used only because of the economic advantage it possesses

Area sampling

- If clusters happen to be some geographic subdivisions, in that case cluster sampling is better known as area sampling.
- The plus and minus points of cluster sampling are also applicable to area sampling.

COMPLEX RANDOM SAMPLING DESIGNS

Multi-stage sampling

- A further development of the principle of cluster sampling.
- Suppose we want to investigate the working efficiency of nationalised banks in India and we want to take a sample of few banks for this purpose. The first stage is to select large primary sampling unit such as states in a country. Then we may select certain districts and interview all banks in the chosen districts. This would represent a two-stage sampling design with the ultimate sampling units being clusters of districts. If instead of taking a census of all banks within the selected districts, we select certain towns and interview all banks in the chosen towns. This would represent a three-stage sampling design.

COMPLEX RANDOM SAMPLING DESIGNS

Multi-stage sampling

- If instead of taking a census of all banks within the selected towns, we randomly sample banks from each selected town, then it is a case of using a four-stage sampling plan.
- If we select randomly at all stages, we will have what is known as 'multi-stage random sampling design'.
- multi-stage sampling is applied in big inquires extending to a considerable large geographical area, say, the entire country.
- easier to administer because sampling frame is developed in partial units
- A large number of units can be sampled for a given cost under multistage sampling because of sequential clustering

COMPLEX RANDOM SAMPLING DESIGNS

Sampling with probability proportional to size

- In case the cluster sampling units do not have the same number or approximately the same number of elements, it is considered appropriate to use a random selection process where the probability of each cluster being included in the sample is proportional to the size of the cluster.
- For this purpose, we have to list the number of elements in each cluster irrespective of the method of ordering the cluster. Then we must sample systematically the appropriate number of elements from the cumulative totals.
- The actual numbers selected in this way do not refer to individual elements, but indicate which clusters and how many from the cluster are to be selected by simple random sampling or by systematic sampling
- Results are equivalent to those of a simple random sample and the method is less cumbersome and relatively less expensive.

COMPLEX RANDOM SAMPLING DESIGNS

Sequential sampling

- The ultimate size of the sample under this technique is not fixed in advance, but is determined according to mathematical decision rules on the basis of information yielded as survey progresses.
- Usually adopted in case of acceptance sampling plan in context of statistical quality control.
- When a particular lot is to be accepted or rejected on the basis of a single sample, it is known as single sampling; when the decision is to be taken on the basis of two samples, it is known as double sampling and in case the decision rests on the basis of more than two samples but the number of samples is certain and decided in advance, the sampling is known as multiple sampling.
- But when the number of samples is more than two but it is neither certain nor decided in advance, this type of system is often referred to as sequential sampling. One can go on taking samples one after another as long as one desires to do so.

MEASUREMENT IN RESEARCH

- We measure characteristics of physical objects as well as abstract concepts.
- Measurement is a relatively complex and demanding task, specially so when it concerns qualitative or abstract phenomena.
- By measurement we mean the process of assigning numbers to objects or observations.
- It is easy to assign numbers in respect of properties of some objects (age, weight), but it is relatively difficult in respect of others (intelligence).
- This causes less accuracy of results of measurement.
- measurement is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence.

MEASUREMENT IN RESEARCH

- In measuring, we devise some form of scale in the range (in terms of set theory, range may refer to some set) and then transform or map the properties of objects from the domain (in terms of set theory, domain may refer to some other set) onto this scale.
- For eg. measuring the number of girls who attend the class. In terms of set theory, this process is one of mapping the observed physical properties of those coming to the show (the domain) on to a sex classification (the range)
- In this artificial or nominal way, categorical data (qualitative or descriptive) can be made into numerical data.
- Nominal data (say 1 for boy and 0 for girl) doesn't have any significance with the numbers 1 and 0. We cannot do 1+1=2 or say $1 < 0$.

- 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
- When in addition to setting up inequalities we can also form differences, we refer to the data as interval data. Temperature for instance.
- 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. Money amounts for instance
- The 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.

MEASUREMENT SCALES

Nominal Scale

- A system of assigning number symbols to events in order to label them
- Assignment of numbers of basketball players in order to identify them
- The numbers are just convenient labels for the particular class of events and as such have no quantitative value.
- Convenient ways of keeping track of people, objects and events.
- counting 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.

MEASUREMENT SCALES

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.
- 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.
- Ordinal scales only permit the ranking of items from highest to lowest.
- 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.
- 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.

MEASUREMENT SCALES

Interval Scale

- In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal.
- 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
- 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° to 40° involves the same increase in temperature as an increase from 60° to 70°, but one cannot say that the temperature of 60° is twice as warm as the temperature of 30° 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.

MEASUREMENT SCALES

Interval Scale

- Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval.
- Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion.
- The generally used tests for statistical significance are the 't' test and 'F' test.

MEASUREMENT SCALES

Ratio Scale

- Ratio scales have an absolute or true zero of measurement. For example, the zero point on a centimeter scale indicates the complete absence of length or height.
- With ratio scales involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu."
- Ratio scale represents the actual amounts of variables.
- 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.
- Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated.

Sources of Error in Measurement

- Respondent
  - Reluctance to express strong negative feelings or admit ignorance
  - Transient factors like fatigue, boredom, anxiety, etc. may limit the ability to respond accurately or fully
- Situation
  - Conditions have serious effects on the interviewer-respondent rapport
  - Presence of someone else, anonymity not assured
- Measurer
  - Interviewer can distort responses by rewording or reordering questions
  - behaviour, style and looks may encourage or discourage certain replies
  - Errors due to incorrect coding, faulty tabulation and/or statistical calculations
- Instrument
  - defective measuring instrument
  - use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, etc.

Tests of Sound Measurement

- Sound measurement must meet the tests of validity, reliability and practicality.
- Validity refers to the extent to which a test measures what we actually wish to measure.
- Reliability has to do with the accuracy and precision of a measurement procedure
- Practicality is concerned with a wide range of factors of economy, convenience, and interpretability

Test of Validity

- Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure.
- Validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested.
- How can one determine validity without direct confirming knowledge?
- We seek other relevant evidence that confirms the answers we have found with our measuring tool.
- Three types : Content validity, Criterion-related validity, Construct validity

Test of Reliability

- A measuring instrument is reliable if it provides consistent results.
- Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument.
- For instance, a scale that consistently overweighs objects by five kgs., is a reliable scale, but it does not give a valid measure of weight.
- A valid instrument is always reliable.
- Reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity.

Test of Reliability

- Two aspects of reliability
  - stability
    - concerned with securing consistent results with repeated measurements of the same person and with the same instrument
    - We usually determine the degree of stability by comparing the results of repeated measurements.
    - By standardising the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimised to the extent possible.
  - equivalence
    - considers how much error may get introduced by different investigators or different samples of the items being studied
    - A good way to test for the equivalence of measurements by two investigators is to compare their observations of the same events.
    - By carefully designed directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used.

Test of Practicality

- The practicality characteristic of a measuring instrument can be judged in terms of economy, convenience and interpretability.
- Economy consideration suggests that some trade-off is needed between the ideal research project and that which the budget can afford.
- For instance, a questionnaire, with clear instructions (illustrated by examples), is certainly more effective and easier to complete
- Interpretability consideration is specially important when persons other than the designers of the test are to interpret the results.
- The measuring instrument, in order to be interpretable, must be supplemented by
  - detailed instructions for administering the test
  - scoring keys
  - evidence about the reliability
  - guides for using the test and for interpreting results.

Technique of developing measurement tools

Involves four stages

- Concept development (the researcher should arrive at an understanding of the major concepts pertaining to his study)
- Specification of concept dimensions (adopting an intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts. For instance, one may think of several dimensions such as product reputation, customer treatment, corporate leadership, concern for individuals, sense of social responsibility and so forth when one is thinking about the image of a certain company.)
- Selection of indicators (specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc.)
- Formation of index (When we have several dimensions of a concept or different measurements of a dimension, we may need to combine them into a single index.)

Scale Classification Bases

- Subject orientation
  - a scale may be designed to measure characteristics of the respondent who completes it or to judge the stimulus object which is presented to the respondent.
- Response form
  - we may classify the scales as categorical and comparative.
  - Categorical scales are also known as rating scales. These scales are used when a respondent scores some object without direct reference to other objects.
  - Under comparative scales, which are also known as ranking scales, the respondent is asked to compare two or more objects.

Scale Classification Bases

- Degree of subjectivity
  - With this basis the scale data may be based on whether we measure subjective personal preferences or simply make non-preference judgements.
  - In the former case, the respondent is asked to choose which person he favours or which solution he would like to see employed, whereas in the latter case he is simply asked to judge which person is more effective in some aspect or which solution will take fewer resources without reflecting any personal preference.
- Scale properties
  - classify the scales as nominal, ordinal, interval and ratio scales.
- Number of dimensions
  - In respect of this basis, scales can be classified as 'unidimensional' (only one attribute of the respondent or object) and 'multidimensional' scales (recognizes that an object might be described better by using the concept of an attribute space of 'n' dimensions).

Scale construction techniques

- Arbitrary approach

  - developed on ad hoc basis, most widely used approach
- Consensus approach

  - a panel of judges evaluate the items chosen for inclusion in terms of whether they are relevant to the topic area and unambiguous in implication.
- Item analysis approach

  - Individual items are developed into a test given to respondents. Total scores are calculated for every one after test. Individual items are analysed to determine which items discriminate between persons or objects with high total scores and those with low scores.
- Cumulative scales

  - chosen on the basis of their conforming to some ranking of items with ascending and descending discriminating power.
- Factor scales

  - may be constructed on the basis of intercorrelations of items which indicate that a common factor accounts for the relationship between items.

Important Scaling Techniques

- Rating Scales
  - qualitative description of a limited number of aspects of a thing or of traits of a person.
  - When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria
  - judge properties of objects without reference to other similar objects.
  - "like-dislike", "above average, average, below average"
  - three to seven points scales are generally used for the simple reason that more points on a scale provide an opportunity for greater sensitivity of measurement.
  - Rating scale may be either a graphic rating scale (The meanings of the terms like "very much" and "some what" may depend upon respondent's frame of reference so much so that the statement might be challenged in terms of its equivalency) or an itemized rating scale (statements are ordered progressively in terms of more or less of some property).

Errors in Rating Scales

- The error of leniency occurs when certain respondents are either easy raters or hard raters.
- When raters are reluctant to give extreme judgements, the result is the error of central tendency.
- The error of hallo effect or the systematic bias occurs when the rater carries over a generalised impression of the subject from one rating to another.

Ranking Scales

- relative judgements against other similar objects. The respondents under this method directly compare two or more objects and make choices among them.
- Method of paired comparisons
  - the respondent can express his attitude by making a choice between two objects
  - If there are more than 2 objects, then a paired comparison is made that leads to $\frac{n(n-1)}{2}$ choices to be made.
  - Too many questions to answer may discourage respondent to participate
  - May use law of transitivity to avoid certain questions
- Method of rank order
  - the respondents are asked to rank their choices
  - Faster and easier than paired comparisons

Scale Construction Techniques

Points to be kept in mind before framing an opinionnaire

- That the statements must elicit responses which are psychologically related to the attitude being measured
- That the statements need be such that they discriminate not merely between extremes of attitude but also among individuals who differ slightly.

Limitations

- People may conceal their attitudes and express socially acceptable opinions.
- They may not really know how they feel about a social issue.
- People may be unaware of their attitude about an abstract situation; until confronted with a real situation, they may be unable to predict their reaction.
- Even behaviour itself is at times not a true indication of attitude.
- we only try to measure the expressed opinion and then draw inferences from it about people's real feelings or attitudes.

Different Scales for Measuring Attitudes of People

Assignment : Try and understand the similarities and differences between each of these approaches

Name of the scale construction approach	*Name of the scale developed*
1. Arbitrary approach	Arbitrary scales
2. Consensus scale approach	Differential scales (such as Thurstone Differential scale)
3. Item analysis approach	Summated scales (such as Likert Scale)
4. Cumulative scale approach	Cumulative scales (such as Guttman's Scalogram)
5. Factor analysis approach	Factor scales (such as Osgood's Semantic Differential, Multi-dimensional Scaling, etc.)

Methods of Data Collection

- The task of data collection begins after a research problem has been defined and research design/ plan chalked out.
- two types of data
  - primary : collected afresh and for the first time, and thus happen to be original in character
  - secondary : which have already been collected by someone else and which have already been passed through the statistical process

COLLECTION OF PRIMARY DATA

- observation method
- interview method
- through questionnaires
- through schedules
- other methods which include
  - warranty cards
  - distributor audits
  - pantry audits
  - consumer panels
  - using mechanical devices
  - through projective techniques
  - depth interviews
  - content analysis

Observation Method

- the most commonly used method specially in studies relating to behavioural sciences
- Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability.
- information is sought by way of investigator's own direct observation without asking from the respondent.

Observation Method

- Advantage
  - subjective bias is eliminated, if observation is done accurately.
  - information obtained relates to present; it is not complicated by either the past behaviour or future intentions or attitudes.
  - independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents
  - suitable if subjects are not capable of giving verbal reports of their feelings for one reason or the other
- Limitations
  - expensive method
  - information is very limited
  - unforeseen factors may interfere with the observational task
  - some people are rarely accessible to direct observation

Observation Method

- Keep in mind
  - What should be observed?
  - How the observations should be recorded?
  - how the accuracy of observation can be ensured?
- Types of observation
  - structured
    - characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation
    - appropriate in descriptive studies
  - unstructured
    - when observation is to take place without these characteristics to be thought of in advance
    - appropriate in exploratory study

Observation Method

- Participant
  - observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience
  - enabled to record the natural behaviour of the group.
  - can even gather information which could not easily be obtained if he observes in a disinterested fashion
  - may lose the objectivity to the extent he participates emotionally
- Non-participant
  - observer observes as a detached emissary without any attempt on his part to experience through participation what others feel
  - When the observer is observing in such a manner that his presence may be unknown to the people he is observing, such an observation is described as disguised observation

Observation Method

- Controlled
  - when observation takes place according to definite pre-arranged plans, involving experimental procedure
  - we use mechanical (or precision) instruments as aids to accuracy and standardisation
  - has a tendency to supply formalised data upon which generalisations can be built with some degree of assurance
- Uncontrolled
  - If the observation takes place in the natural setting
  - no attempt is made to use precision instruments
  - major aim is to get a spontaneous picture of life and persons
  - downfall : subjective interpretation
  - danger of having the feeling that we know more about the observed phenomena than we actually do

Interview method - Personal Interviews

- interviewer asking questions generally in a face-to-face contact to the other person or persons.
- direct personal investigation
  - interviewer has to collect the information personally from the sources
  - Has to be on the spot and meet people from whom data have to be collected.
  - suitable for intensive investigations
- indirect oral investigation.
  - in certain cases it may not be possible or worthwhile to contact directly the persons concerned or on account of the extensive scope of enquiry
  - the interviewer has to cross-examine other persons who are supposed to have knowledge about the problem under investigation and the information, obtained is recorded
  - Most of the commissions and committees appointed by government

Interview method - Personal Interviews

- Structured
  - method of collecting information through personal interviews in a structured way
  - involve the use of a set of predetermined questions and of highly standardised techniques of recording
  - interviewer follows a rigid procedure laid down, asking questions in a form and order prescribed
  - preferred in case of descriptive studies
  - requires relatively lesser skill on the part of the interviewer

Interview method - Personal Interviews

- Unstructured
  - characterised by a flexibility of approach to questioning
  - do not follow a system of pre-determined questions and standardised techniques of recording information
  - interviewer is allowed much greater freedom to ask, in case of need, supplementary questions or at times he may omit certain questions if the situation so requires
  - this sort of flexibility results in lack of comparability of one interview with another and the analysis of unstructured responses becomes much more difficult and time-consuming
  - demand deep knowledge and greater skill on the part of the interviewer
  - happens to be the central technique of collecting information in case of exploratory or formulative research studies

Interview method - Personal Interviews

- Focussed
  - meant to focus attention on the given experience of the respondent and its effects
  - the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives
  - main task of the interviewer is to confine the respondent to a discussion of issues with which he seeks conversance
  - used generally in the development of hypotheses and constitute a major type of unstructured interviews
- Clinical
  - concerned with broad underlying feelings or motivations or with the course of individual's life experience
  - The method of eliciting information under it is generally left to the interviewer's discretion

Interview method - Personal Interviews

- Non-directive
  - the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning
  - interviewer often acts as a catalyst to a comprehensive expression of the respondents' feelings and beliefs and of the frame of reference within which such feelings and beliefs take on personal significance
- advantages
  - More information and that too in greater depth can be obtained.
  - Interviewer by his own skill can overcome resistance
  - can be made to yield an almost perfect sample of the general population.
  - non-response generally remains very low.
  - The interviewer can control who answers the questions
  - The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions
  - misinterpretations of questions can be avoided.

Interview method - Personal Interviews

- Disadvantages
  - very expensive, there remains the possibility of the bias of interviewer as well as that of the respondent
  - Certain types of respondents such as important officials or executives or people in high income groups may not be easily approachable
  - more-time-consuming, specially when the sample is large
  - presence of the interviewer may over-stimulate the respondent
- Pre-requisites and basic tenets of interviewing
  - interviewers should be carefully selected, trained and briefed, should be honest, sincere, hardworking, impartial and must possess the technical competence and necessary practical experience.
  - Occasional field checks should be made to ensure that interviewers are neither cheating, nor deviating from instructions
  - interviewing is an art governed by certain scientific principles
  - Every effort should be made to create friendly atmosphere of trust and confidence, so that respondents may feel at ease while talking to

Interview Method - Telephonic interviews

A1 more flexible, faster, cheaper

A2 Recall is easy; callbacks are simple and economical.

A3 higher rate of response

A4 Replies can be recorded without causing embarrassment to respondents.

A5 access can be gained to respondents who otherwise cannot be contacted for one reason or the other.

A6 No field staff is required.

A7 Representative and wider distribution of sample is possible.

DA1 Little time is given to respondents for considered answers

DA2 Surveys are restricted to respondents who have telephone facilities.

DA3 not suitable for intensive surveys where comprehensive answers are required to various questions.

DA4 Questions have to be short and to the point.

COLLECTION OF DATA THROUGH QUESTIONNAIRES

- Advantages
  - There is low cost even when the universe is large and is widely spread geographically
  - It is free from the bias of the interviewer
  - Respondents have adequate time to give well thought out answers.
  - Respondents, who are not easily approachable, can also be reached.
  - Large samples can be made use of and thus the results can be made more dependable and reliable.
- Disadvantages
  - Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.
  - It can be used only when respondents are educated and cooperating.

Before using this method, it is always advisable to conduct 'pilot study' (Pilot Survey) for testing the questionnaires.

Main aspects of a questionnaire - General form

- structured
  - definite, concrete and pre-determined questions
  - The questions are presented with exactly the same wording and in the same order to all respondents
  - The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning
  - highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum
  - simple to administer and relatively inexpensive to analyse. wide range of data and that too in respondent's own words cannot be obtained
  - considered inappropriate in investigations where the aim happens to be to probe for attitudes and reasons for certain actions or feelings

Main aspects of a questionnaire - General form

- unstructured
  - the interviewer is provided with a general guide on the type of information to be obtained, but the exact question formulation is largely his own responsibility and the replies are to be taken down in the respondent's own words to the extent possible
  - widely used when a problem is being first explored and working hypotheses sought

Main aspects of a questionnaire - Question sequence

- The question-sequence must be clear and smoothly-moving, with questions that are easiest to answer being put in the beginning.
- The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation.
- The opening questions should be such as to arouse human interest.
- Following the opening questions, we should have questions that are really vital to the research problem and a connecting thread should run through successive questions.
- Relatively difficult questions must be relegated towards the end so that even if the respondent decides not to answer such questions, considerable information would have already been obtained.

Main aspects of a questionnaire - Question formulation and wording

- should be easily understood, concrete and should conform as much as possible to the respondent's way of thinking. (For instance, instead of asking. "How many razor blades do you use annually?" The more realistic question would be to ask, "How many razor blades did you use last week?")
- two principal forms : multiple choice (respondent selects one of the alternative possible answers) and open-end question (supply the answer in his own words).
- 'closed question' (Yes or no)
- MCQ : easy handling, simple to answer, quick and relatively inexpensive to analyse.
- But "putting answers in people's mouths"
- Simple words, which are familiar to all respondents should be employed. Words with ambiguous meanings must be avoided.
- danger words, catch-words or words with emotional connotations should be avoided

Essentials of a good questionnaire

- short and simple
- Questions should proceed in logical sequence moving from easy to more difficult questions.
- Personal and intimate questions should be left to the end.
- There should be some control questions in the questionnaire which indicate the reliability of the respondent. For instance, a question designed to determine the consumption of particular material may be asked first in terms of financial expenditure and later in terms of weight. The control questions, thus, introduce a cross-check to see whether the information collected is correct or not.
- Questions affecting the sentiments of respondents should be avoided.
- There should always be provision for indications of uncertainty, e.g., "do not know," "no preference" and so on.

Read it and understand the differences between a questionnaire and a schedule!

COLLECTION OF SECONDARY DATA

- Secondary data means data that are already available
- Secondary data may either be published data (available in various publications of the central, state are local governments, publications of foreign governments or of international bodies and their subsidiary organisations, books, magazines, reports) or unpublished data (diaries, letters, unpublished biographies and autobiographies).

Caution before using secondary data

- Reliability of data
  - Who collected the data? What were the sources of data?
  - Were they collected by using proper methods
  - At what time were they collected? Was there any bias of the compiler?
  - What level of accuracy was desired? Was it achieved ?
- Suitability of data
  - The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry.
  - Unsuitable data should not be used.
  - the researcher must very carefully scrutinise the definition of various terms and units of collection used at the time of collecting the data from the primary source
- Adequacy of data
  - If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher.

SELECTION OF APPROPRIATE METHOD FOR DATA COLLECTION

- Nature, scope and object of enquiry
  - The method selected should be such that it suits the type of enquiry.
  - Important in deciding whether the data already available (secondary data) are to be used or the data not yet available (primary data) are to be collected
- Availability of funds
- Time factor
- Precision required

CASE STUDY METHOD

- very popular form of qualitative analysis and involves a careful and complete observation of a social unit, be that unit a person, a family, an institution, a cultural group or even the entire community.
- method of study in depth rather than breadth.
- places more emphasis on the full analysis of a limited number of events or conditions and their interrelations.
- deals with the processes that take place and their interrelationship.
- The object of the case study method is to locate the factors that account for the behaviour-patterns of the given unit as an integrated totality.

CASE STUDY METHOD - Characteristics

- The selected unit is studied intensively
- Generally, the study extends over a long period of time to ascertain the natural history of the unit
- Through this method we try to understand the complex of factors that are operative within a social unit as an integrated totality.
- Under this method the approach happens to be qualitative and not quantitative. For instance, under this method we not only study how many crimes a man has done but shall peep into the factors that forced him to commit crimes when we are making a case study of a man as a criminal. The objective of the study may be to suggest ways to reform the criminal.
- Case study method results in fruitful hypotheses along with the data which may be helpful in testing them, and thus it enables the generalised knowledge to get richer and richer.

CASE STUDY METHOD - Assumptions

- The assumption of uniformity in the basic human nature in spite of the fact that human behaviour may vary according to situations.
- The assumption of studying the natural history of the unit concerned.
- The assumption of comprehensive study of the unit concerned.

CASE STUDY METHOD - Phases involved

- Recognition and determination of the status of the phenomenon to be investigated or the unit of attention.
- Collection of data, examination and history of the given phenomenon.
- Diagnosis and identification of causal factors as a basis for remedial or developmental treatment.
- Application of remedial measures i.e., treatment and therapy (this phase is often characterised as case work).
- Follow-up programme to determine effectiveness of the treatment applied.

CASE STUDY METHOD - Advantages

- enables to understand fully the behaviour pattern of the concerned unit
- can obtain a real and enlightened record of personal experiences which would reveal man's inner strivings, tensions and motivations that drive him to action along with the forces that direct him to adopt a certain pattern of behaviour.
- helps in formulating relevant hypotheses along with the data which may be helpful in testing them.
- facilitates intensive study of social units which is generally not possible if we use either the observation method or the method of collecting information through schedules.
- enhances the experience of the researcher and this in turn increases his analysing ability and skill.

CASE STUDY METHOD - Limitations

- not comparable between subjects.
- danger of false generalisation
- consumes more time and requires lot of expenditure
- based on several assumptions which may not be very realistic at times
- Sampling is not possible under a case study method.
- Response of the investigator is an important limitation of the case study method. He often thinks that he has full knowledge of the unit and can himself answer about it.

THANK YOU FOR THE ATTENTION AND PARTICIPATION!