Unit 3 Sampling process and data collection

Meaning and definition of sampling

- Selecting a few cases out of some larger grouping fo study is called sample
- A sample is defined as a smaller set of data that a researcher chooses or selects from a larger population by using a pre-defined selection method
- Sampling is the method or process of selecting samples from the larger population
- Sampling allows us to study a workable number of cases from the large group to derive findings that are relevant for all members of the group

- Sampling is a process used in statistical analysis in which a predetermined number of observations are taken from a larger population
- The methodology used to sample from a larger population depends on the type of analysis being performed
- Sampling is a technique of selecting individual members or a subset of the population to make statistical inferences from them and estimate characteristics of the whole population
- A sample consists of one or more elements or cases selected from larger grouping or population(Sullivan 2001)

Characteristics of sampling

1. Population
2. Size
3. Process/method
4. Representation
5. Neutrality
6. Careful
7. Systematic

Significane of sampling

1. Representation
2. Neutrality
3. Coverage
4. Manageable
5. Reduced cost and time
6. Reduced resource deployment
7. Accuracy of data
8. Intensive and exhaustive data
9. Apply properties to a larger population

Proesses

1. Identify the population of interest
2. Specify a sampling frame
3. Specify a sampling method
4. Determine the sample size
5. Implement the plan

Types of samples

1. Probability sampling
2. Non-probability sampling

1. Probability Sampling is a sampling technique in which samples from a larger population are chosen using a method based on the theory of probability

- Probability sampling is a method of deriving a sample where the objects are selected from a population-based on the theory of probability
- This method includes everyone in the population, and everyone has an equal chance of being selected
- There is no bias whatsoever in this type of sample. Each person in the population can subsequently be a part of the research
- Probability sampling can be further classified into four distinct types of samples as;

A. Simple random sampling: In this method, each member has an equal chance of being a part of the study

- The objects in this sample population are chosen purely on a random basis, and each member has the same probability of being selected

B. Cluster/multistage sampling: Cluster sampling is a type of sampling method where the respondent population is divided into equal clusters

- Clusters are identified and included in a sample based on defining demographic parameters such as age, location, sex, etc.
- This makes it extremely easy for a survey creator to derive practical inferences from the feedback
- Research studies are then administered to respondents in these clusters. This type of generating a sample makes the data collection in-depth and provides easy to consume and act upon, insights

C. Systematic sampling: Systematic sampling is a sampling method where the researcher chooses respondents at equal intervals from a population

- The approach to select the sample is to pick a starting point and then pick respondents at a pre-defined sample interval. For example, making list of 100 and then pick 10 on the basis of interval 10

D. Stratified random sampling: Stratified random sampling is a method of dividing the respondent population into distinctive but pre-defined parameters in the research design phase

- In this method, the respondents don't overlap but collectively represent the whole population
- This forms smaller groups of people or samples, and then some objects from these samples can be used for the research study.

2. The non-probability sampling

- Non-probability sampling is a sampling technique in which the researcher selects samples based on the researcher's subjective judgment rather than random selection
- The non-probability sampling method uses the researcher's discretion to select a sample
- This type of sampling is used for preliminary research where the primary objective is to derive a hypothesis about the topic in research
- Here each member does not have an equal chance of being a part of the sample population, and those parameters are known only post-selection to the sample.
- Non-probability sampling is further into four distinct types of samples which are:

A. Convenience sampling: Convenience sampling stands for the convenience of a researcher accessing a respondent.

- There is no scientific method of deriving this sample. Researchers have nearly no authority over selecting the sample elements, and it's purely done on the basis of proximity and not representativeness
- In this sampling method, the sample respondents are chosen purely on their proximity to the survey desk and their willingness to participate in the research.

B. Judgemental/purposive sampling: This is a method of developing a sample purely on the basis and discretion of the researcher purely on the basis of the nature of study along with his/her understanding of the target audience

- In this sampling method, people who only fit the research criteria and end objectives are selected, and the remaining are kept out

C. Snowball sampling: Snowball sampling or chain-referral sampling is defined as a non-probability sampling technique in which the samples have traits that are rare to find

- This is a sampling technique, in which existing subjects provide referrals to recruit samples required for a research study. In this case, the researcher can recruit people with an understanding or knowledge of such people and collect information from them or ask them to collect information.

D. Quota sampling: Quota sampling is a method of collecting a sample where the researcher has the liberty to select a sample based on their strata

- The primary characteristic of this method is that two people cannot exist under two different conditions

Sampling and Non-sampling error

Sampling error

- sample error is defined as the differences between the population mean and sample mean
- For example, the population mean age of KCC student is 20 and the sample mean age of KCC student is 19. It seems difference which is sample error
- The reasons of sample error are;

1. Faulty selection of sample
2. Substitution(replacement or change)

3. Faulty demarcation of the sampling unit

4. Improper choice of the estimation techniques or error due to bias in the estimation method

5. Variability of the population that is heterogenous in nature

Non-sampling error

- Faulty planning including vague and faulty definitions of the population or the statistical units to be used, incomplete list of population members
- Vague or imperfect questionnaire
- Defective methods of interviewing and asking questions
- Vagueness about the type of data to be collected
- Personal bias of the investigator
- Lack of trained and qualified investigators and lack of supervisory staff

- Failure of respondent's memory to recall the events or happenings in past
- Non-response and inadequate or incomplete response
- Improper coverage
- Publication errors

Control or minimizing non-sampling errors

- Qualified and trained personnel
- More sophisticated statistical tool
- Adequate supervisory
- Pre-testing/pilot survey
- Effective checking in the processing data

Data collection

- Data are information, facts, opinions, texts, or numbers that can be collected
- Data are the building block of research
- Data collection consists of taking systematic information from reality and transferring it into some recording system
- Data collection is the main activity of research
- Data can be defined as the values collected through record keeping or polling, observing or measuring

- Data are knowledge. Data becomes information when it becomes relevant to your decision. Information becomes fact when the data can support it
- Facts are what the data reveals
- Knowledge is more than knowing
- Data convert into information then facts, and ultimately into knowledge
- Data do not speak itself, it needs to interpret
- Data are verifiable

Types of data

1. On the basis of nature

A. Qualitative: Data comes in words

B. Qualitative: Data comes in numbers

2. On the basis of explanation

A. Subjective: Narrative, measurable in words

B. Objective: Measurable in numbers

3. On the basis of sources

A. Primary: First hand or field

B. Secondary: Previous studies done by others

Sources of secondary data

- Literature review
- Books
- Articles
- Journals
- Reports
- Documents
- Videos
- Electronics
- Archives .....

Methods of collecting primary data

1. Observation
2. Interview
3. Questionnaire

1. Observation

- Looking on something systematically and purposefully
- Observation is a way of collecting data through observing. Observation data collection method is classified as a participatory study
- Observation refers to watching what people do
- Observation is a technique that involves systematically selecting, watching, listening, reading, touching, and recording behavior and characteristics of living beings, objects, or phenomena
- The observation method involves human or mechanical observation of what people actually do or what events take place

- The observation method is described as a method to observe and describe the behavior of a subject
- Observation is recording of all forms of sense perception - Johan Galtung
- Science begins with observation and must ultimately return to observation for final validation - Good and Hat
- Observation is systematic and deliberate study through the eyes of the spontaneous occurences at the the time they occur - P.V.Young

Types

1.Participant Observation

- participant observation, which refers to an observation in which an observer gains firsthand knowledge by being in and around the social setting that is being investigated.
- In participant observation researcher joins in and becomes part of the group they are studying to get a deeper insight into their lives
- Participant observations can be either covert or overt. Covert is where the study is carried out 'undercover'. The researcher's real identity and purpose are kept concealed from the group being studied
- The researcher takes a false identity and role, usually posing as a genuine member of the group

Semi/or non participant

- Overt is where the researcher reveals his orher true identity and purpose to the group and asks permission to observe
- The observer takes part in the situation he or she observes.
- With this method, the observer joins in the daily life of the group or organization he is studying.
- He watches what happens to the members of the community and how they behave, and he also engages in conversations with them to find out their reactions to and interpretations of the events that have occurred.

The observer watches the situation openly

- Observers become participating member in the activities of the group by revealing their identities and goals of research
- When the researcher adopts this type of role, s/he informs the group being studied that there is a research agenda
- Research goal is explicitly identified

Direct observation

- Direct observation refers to the situation when the observer remains physically present and personally monitors what takes place
- This approach is very flexible because it allows the observer to react to and report subtle aspects of events as they occur
- During the act of observation, the observer is free to change the focus of observation, concentrate on unexpected events, or even change the place of observation if the situation demands.

Indirect observation

- Indirect observation occurs when the recording is done by mechanical, photographic, videotape, cameras, or other electronic means
- Such observation can also be conducted in planning traffic control and redesigning of peripheral streets.

Characteristics

1. Both mental and physical activities
2. Purposive
3. Selective
4. Natural/social environment context
5. Direct method
6. Used for primary datcollection
7. Focuse of significant events

Recording tools

- Notes
- Tape
- Video
- Photographs

Advantages

- Directness
- Natural environment
- Longitudinal analysis
- Non-verbal behavior
- Not necessary to rely on the willingness and ability of respondents
- More objective and generally more accurate
- Provides direct access to research phenomena
- By observing firsthand, the researcher can collect, check and record accurate data
- Greater flexibility in terms of application

Disadvantages

- Lack of control
- Difficulties in quantification
- Smallness in sample size
- No opportunity to learn past
- Inability to observe such things such as attitudes, motivations
- Takes time for the investigator to wait for a particular action to take place
- Cost is the final disadvantage of observation method
- There is a chance of higher observer bias
- Several personal behaviors are not open for observation and this proves a limitation in case of observation method

2. Interview

- An inteview is a conversation between two or more people(the interviewer and interviewee)
- Interview method involves a face to face meeting in which researcher asks an individuals a series of questions
- An interview involves an interviewer reading questions to respondent and recording their answers
- “The interview is a social relationship designed to exchange information between respondents and interviewer” - Sullivan 2001

- “ The interview is a face to face iterpersonal role situation in which one person, the interviewer asks a person being interviewed, the respondent questions designed to obtain answers pertinent to the purpose of the research problem”- Kerlinger, F.D.

Types of interview

1. On the basis of structure

A. Structured: It is formal or closed type of interview

- Predetermined questions are asked
- Close ended questions are asked, it is fairly rigid type
- Directive interview, not flexibility
- Ticks appropriate
- Useful for hypothesis

B. Semi-structured interview

- Both structured and unstructured questions are asked
- It incudes both open ended and close ended nature of questons

C. Unstructured

- It is informal, focused, depth, non-directed and open ended questions
- Freedom of expressions

2. On the basis of procedures

A Face to face

B. Telephone

C. Messenger

D. Skype

E. Video conference

F. E-mail ....

Characteristics of Interview

1. Interviewer's job
2. Personal form of research
3. Direct or face to face
4. Opportunuty to probe
5. Comfortable for respondents
6. Timec onsuming
7. Needs well trained interviewer

Principles of interview

1. The factors help in motivating the respondents to cooperate
2. The respondent must feel that their interaction with the interviewer will be pleasant and satisfying
3. Barriers to the interview in the respondents' mind need to overcome

Procedures of interview

1. Verify the recording instruments
2. Start from easy questions
3. Ask one question at a time
4. Neutrality
5. Dont lose control
6. provide transition metween major topics

3. Questionnaire

- A questionnaire is a formal list of qestions designed to gather responses from respondent on a given topic
- A questionnaire is a set of questions prepared for respondents
- Questionnaires are designed so tha they can be answered without assistance
- A questionnaire is tool for data collection that consists of a number of questions printed or typed in a definite order on a form or a set of forms
- A questionnaire is a device for securing answers to questions using a form which the respondent fills himself" - Good and Hatt
- "A questionnaire contains written questions that people respond to directly on the questionnaire form itself, without the assistance of an interviwer"

  - Sullivan 2001

Types of questionnaire

1. Structured : Those which pose definite, concrete, and preordained questions, that is, these are prepared in advance andt constructed on the spot during the questioning period
2. Unstructured: Frequently referred as interview guides or interview schedule
3. Self administered: Respondents fill themselves
4. Online
5. Mail post
6. Telephone

Type of questions

1. Open ended
2. close ended
3. Contigency: Questions which depend on the responses to earlier questions are referred as contigency questions
4. Matrix questions: Almost simila questions, e.g. extremly support, support
5. Pictorial questions

Considerations to formulate questions

1. Simple and clear questions
2. No double-barrelled questions
3. Not leading questions

Merits

- Less expensive
- Maintains confidentiality

Demerits

1. Response rate is low
2. lack of clarity of issues
3. Time consuming