CHAPTER FOUR

Research Design

Research Design (1)

- The arrangement of conditions for the collection and analysis of data in a manner that aims to address the **research problem.**
- **The research design should be inline with:**
  - What is the study about? (Problem definition)
  - Why is the study being made? (Justification)
  - Where will the study be carried out? (Location)
  - What type of data is required? (Quanti, Qual, Pri, Sec)
  - Where can the required data be found (target population)
  - What will be the sample design (technique chosen)
  - What techniques of data collection will be used? (observation, interview, questionnaire, or document analysis).
  - How will the data be analysed (Data analysis techniques & tools to be employed).

Research Design (2)

- Research methodology involves such general activities as identifying problems, review of the literature, formulating hypotheses, procedure for testing hypotheses, measurement, data collection, analysis of data, interpreting results and drawing conclusions.
- We may split the overall research design into three:
  - **The sampling design** - which deals with the method of selecting items to be observed for the given study.
  - **The statistical design** - which concerns with the question of how many items are to be observed and how the information and data gathered are to be **analysed**;
  - **The operational design** - which deals with the techniques by which the procedures specified in the **sampling, Statistical and observational** designs can be carried out.

Important concepts in Research Design

- **Variable**: A concept which can take on different values.
  - **Continuous variable**: A variable which can assume any numerical value within a specific range. (e.g. Age, weight).
  - **Discrete variable**: A variable for which the **individual values** fall on the scale only with distinct gaps. (e.g. Number of **children**).

Important concepts in Research Design

- **Dependent variable**: A consequence of another variable (effect)
- **Independent variable**: The presumed cause of the dependent variable

  Example

  Behavioral changes: dependent variable

  Films, lectures, …: independent variable

- **Extraneous variables**: Those uncontrolled variables that may have a **significant influence** on the dependent variable.

- Research design can be
  - Quantitative
  - Qualitative

Quantitative Research Design (1)

- **Aim**: Determine relationship between quantitative variables or compare groups.
- Types of **Quantitative** design:
  - Descriptive: Collect information and statistically analyze it
  - Experimental
  - Quasi-experiment: attempts to establish a cause-and-effect relationship

Qualitative Research Design (2)

- **Qualitative research methods** goes through participant observation, in-depth interviewing that yield **descriptive** data.
- **Qualitative** research notions:
  - Are concerned with the meanings people attach to things in their lives.
  - Qualitative researches develop concepts, insights and understandings from patterns in the data rather than collecting data to assess preconceived models, **hypothesis** or **theories**.

Qualitative Research Design (3)

- Qualitative research are concerned with how people think and act in their everyday lives
  - In qualitative interviewing, researchers model their interviews after a normal conversation rather than a formal question and answer exchange.
  - For the qualitative researcher, there is something to be learned in all settings and groups.

Qualitative Research Design (4)

- Research setting
  - The ideal research setting is one in which the observer obtains **easy access**
- **Literature**
  - You can not be sure what literature might be relevant to your study until you have completed your research.

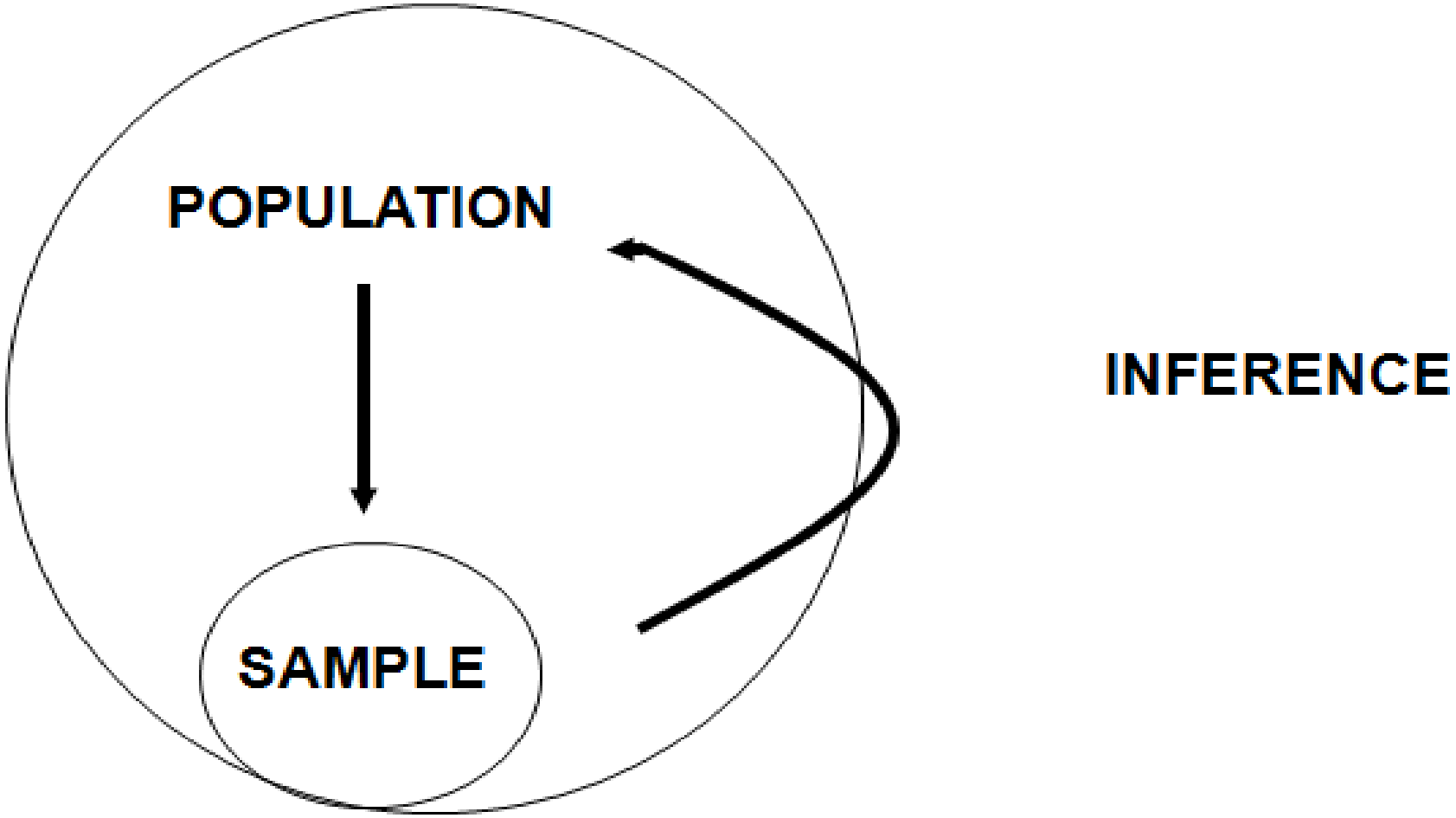
Research Method and Methodology

- **Research Method** describes the specific technique used in a given study.
- **Research Methodology**
  - A science of studying how research is done scientifically.
  - A way to systematically solve the research problem by logically adopting various steps.
  - Methodology helps to understand not only the products of **scientific inquiry** but the **process itself**.
  - Aims to **describe** and **analyze** methods, throw light on their limitations and resources, clarify their presuppositions and consequences, relating their potentialities to the twilight zone at the 'frontiers of knowledge'

Sampling Design

- Sampling: The act, process, or technique of selecting a suitable sample or representative part of a population for the purpose of determining parameters or characteristics of the whole population.
- **Population**: The entire group under study as defined by research objectives
  - Sometimes called the "**universe**." or Reference population
- Researchers define populations in **specific terms** such as heads of households, individual person types, families, types of retail outlets, etc.
- Population, geographic location and time of study are also considered.

The sampling process…



- **Sample:** a subset of the population that should represent the **entire** group.
- **A sample** is "a smaller (but hopefully representative) collection of units from a population used to determine truths about that population" (Field, 2005).
- **Sample unit:** the basic level of investigation…consumers, store managers, shelf-facings, adolescence, etc.
  - The research objective should define the sample unit
- **A sampling frame** which has the property that we can identify every single element and include any in our sample
- The **sampling frame** is the list from which the potential respondents are drawn
  - Registrar's office, Class rosters, Must assess sampling frame errors
- The sampling frame must be **representative** of the population
- **Census:** is the counting of the **complete** population

Calculating sample frame error (SFE):

- Subtract the number of items on the sampling list from the total number of items in the population
- Take this number and divide it by the total population. Multiply this decimal by 100 to convert to percent (SFE must be expressed in %)
- If the SFE was 40%, this would mean that 40% of the population was not in the sampling frame

- Practical considerations such as:
  - Cost and population size
  - Nature & purpose of the study
  - Inability of researcher to analyze large quantities of data potentially generated by a census
- Samples can produce sound results if proper rules are followed for the draw.

Why is sampling?

- **Economy**: Save resources such as, money and manpower
- **Timeliness** : Save time
- A sample provides more timely data than a **census**
- **Completeness** and **accuracy**

  –It is possible to collect more **detailed** information, because more time can be spent in **asking** respondents.
- **Destructiveness of the observation/study:** When Experiments are to result in the **destruction** of the materials that **undergo** the experiment.
- When population **size is very large**
- **Inaccessibility of some of the population**: In cases of **inaccessible** population

- The three factors that influence sample representativeness
    - Sampling procedure
    - Sample size
    - Participation (response)
- When might you sample the entire population?
    - When your population is very small
    - When you have extensive resources
    - When you don’t expect a very high response

Two Types of Sampling Methods/Designs:

- **Probability sampling** scheme is one in which every unit in the population has a chance (greater than zero) of being selected in the sample, and this probability can be **accurately** determined.
  - When every element in the population does have the same probability of selection, this is known as an 'equal probability of selection' (EPS) design. Such designs are also referred to as 'self-weighting' because all sampled units are given the **same weight**.
- Each members of the population have a **known chance** (probability) of being selected.
  - Each unit of population will be represented in the **sample**.

Probability sampling includes:

1. Simple random sampling
2. Systematic sampling
3. Stratified sampling
4. Proportional stratified sampling
5. Cluster Sampling

- **Non-probability sampling:** Instances in which the chances (probability) of selecting members from the population are unknown
  - The researcher has no way of forecasting that each member of the population will be represented in the sample.
  - Some member of the population have little or no chance of being selected.
- Any sampling method where some elements of population have no chance of selection (these are sometimes referred to as **'out of coverage'/'under covered'**), or where the **probability of selection** can't be accurately determined

Non-probability sampling includes:

1. Convenience sampling
2. Judgment or purposive sampling
3. Quota sampling

1. Simple random sampling

- Sampling technique in which each element of the population has an equal, known and independent chance of being included in the sample.
- The **least sophisticated** of all sampling designs.
- Applicable when population is **small**, **homogeneous** & readily available/known
- Impractical for very large populations
- A table of random number or lottery system is used to determine which units are to be selected.

1. Simple random sampling….

- In this sampling all subsets of the **frame** are given an equal probability. Each element of the frame thus has an equal probability of selection
- This is done by assigning a number to each unit in the sampling frame
  - Some of examples are
    - Lottery Method
    - Throwing a dice/cube
    - Tossing a coin

2. Systematic sampling:

- Relies on arranging the target population according to some ordering scheme and then selecting elements at regular intervals through that ordered list
- This method is at times more efficient than simple random sampling
- Selecting elements of the population in **predetermined** sequence

2. Systematic sampling….

- Randomness element is in **picking** up the starting point.
  - Involves a random start and then proceeds with the selection of every *k*th element from then onwards. In this case, *k*=(population size/sample size).
- A simple example would be to select every 10th name from the telephone directory (an 'every 10th' sample, also referred to as 'sampling with a skip of 10').

3. Stratified random sampling

- Where population embraces a number of distinct categories, the frame can be organized into separate "strata."
  - Each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected.
- Every unit in a stratum has same chance of being selected
- Applied when the population has different layers (strata).

- The researcher samples from each one of the layers (stratum) equally
  - Examples
    - Sampling of school children from grades 4, 5 and 6
    - Sampling of customers (customer frame)
      - Corporate customers
      - Residential customers, etc…
- Since each stratum is treated as an independent population, different sampling approaches can be **applied** to different strata.

4. Proportional stratified sampling

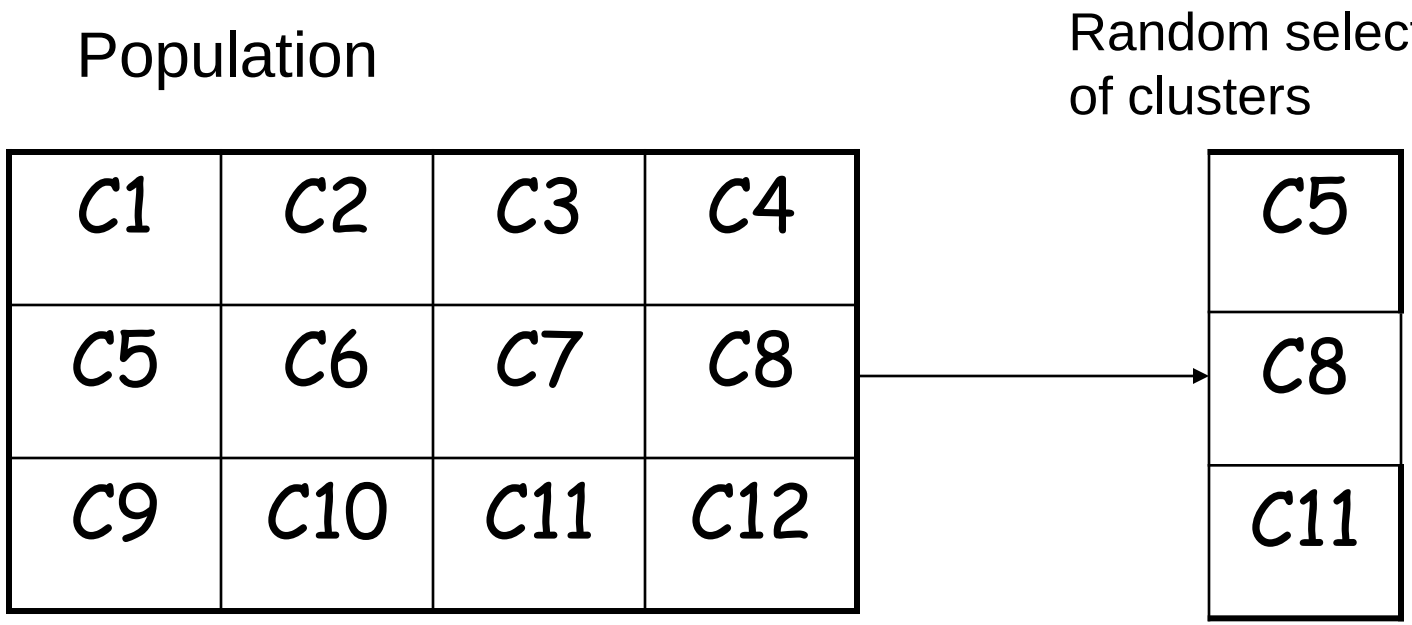
- It is preferable sampling technique when number of the elements of the strata are **different** in size
    - 2,000 corporate customers
    - 400,000 residential customers
- Adequate representation of minority subgroups of interest can be ensured by stratification & varying sampling fraction between strata as required.

5. Cluster sampling

- Once clusters are established, a random draw is done to select one (or more) clusters to represent the population
- Population divided into clusters of homogeneous units, usually based on **geographical** contiguity
- Sampling units are groups rather than individuals.
- A sample of such clusters is then selected.
- All units from the selected clusters are studied.

5. Cluster sampling….

– Grouping the population into clusters and then select members of clusters.



Two types of cluster sampling methods.

One-stage sampling. All of the elements within selected clusters are included in the sample.

Two-stage sampling. A subset of elements within selected clusters are randomly selected for inclusion in the sample.

Difference Between Strata and Clusters

- Although **strata** and **clusters** are both non-overlapping subsets of the population, they differ in several ways.
  - All strata are represented in the sample; but only a subset of clusters are in the sample.
  - With stratified sampling, the best survey results occur when elements within strata are internally **homogeneous**. However, with cluster sampling, the best results occur when elements within clusters are internally **heterogeneous.**

Summary

- **Population characteristics** • Appropriate sampling technique

1. Homogeneous members	•Simple random sampling •Systematic random sampling
2. Stratified population with approximately equal in size	Stratified random sampling
3. Stratified population, strata different in size	Proportional stratified sampling
4. Population with discrete clusters with similar characteristics	Cluster sampling

Non-probability Sampling Methods

1. Convenience sampling:

- Sometimes known as **grab** or **opportunity sampling or accidental or haphazard sampling.**
- A type of nonprobability sampling which involves the sample being drawn from that part of the population which is close to hand. That is, readily available and convenient.
- Taking People/sample who are available, volunteer, or can be easily recruited.

  E.g. those that arrive on a scene by coincidence
- Appropriate for some less demanding research.
- Use results that are easy to get

Non-probability Sampling Methods…

- The researcher using such a sample cannot scientifically make generalizations about the total population from this sample because it would not be representative enough.
- For example, if the interviewer was to **conduct** a **survey** at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time, which would not represent the views of other members of society in such an area, if the survey was to be conducted at different times of day and several times per week.

2. Judgment or purposive sampling:

- The researcher chooses the sample based on who they think would be appropriate for the study. This is used primarily when there is a limited number of people that have expertise in the area being researched.
- This involves the selection of a group/ indivisuals from the population on the basis of available information thought.
- Samples require a judgment or an “educated guess” on the part of the interviewer.
- Also, “judges” (informed individuals) may be asked to suggest who should be in the sample.
- A **sampling technique** to pick out the sample in relation to some **criterion.**
- Which are considered important for the particular study.
- Units of the sample are chosen purposively.
- Choosing people who we have decided are “**typical**” of a group;

3. Quota sampling

- A variation of convenience **sampling**
- Elements are selected in the same proportion as in the population.
  - Ex: There are equal number of **IT** and SE **students**.
    - Quota sampling would choose **20 IT** and **20 SE** students without any attempt to **random selection**.

Size of Sample (1)

- The size of the sample depends upon the *precision* the researcher desires in estimating the population parameter at a particular confidence level.
- There is no single rule that can be used to determine sample size.
- The best answer to the question of size is to use as large a sample as possible.
- A larger sample is much more likely to be representative of the population.

Size of Sample (2)

- Furthermore with a large sample the data are likely to be more ***accurate*** *(value)* and ***precise*** *(sample error)*.

–It was pointed out in that the larger the sample, the smaller the standard error (high precision). And

–The larger the sample, the higher representative (high accuracy)

- In general, the standard error of a sample mean is inversely proportional to the square root of *n*. Thus, in order to double the precision of one's estimation, the sample size would need to be quadrupled.

Steps in Sampling Process

- Defining the population
- Specifying the sampling unit
- Specifying the sampling frame (the means of representing the elements of the population). i.e. a set of items or events possible to measure
- Specifying the sampling method for selecting items or events from the frame.
- Determining the sampling size
- Specifying the sampling plan
- Selecting the sample

Measurement & Measurement Scales

- Measurement can be defined as a process through which researchers describe, explain, and predict the phenomena and constructs of our daily existence.
- Measurement is often viewed as being the basis of all scientific inquiry, and measurement techniques and strategies are therefore an essential component of research methodology.
- For example, we measure how long we have lived in years, our financial success in dollars, and the distance between two points in miles.
- The concept of measurement is important in research studies in two key areas.
- First, measurement enables researchers to quantify abstract constructs and Variables.

- for example, the amount of weight lost in pounds
- Without measurement, researchers would be able to do little else but make unsystematic observations of the world around us.

- Second, the level of statistical sophistication used to analyze data derived from a study is directly dependent on the scale of measurement used to quantify the variables of interest.

Non metric Data vs. Metric Data

- **Non metric data** (also referred to as qualitative data)-which cannot be quantified and are predominantly used to describe and categorize.
- **Metric data** (also referred to as quantitative data)-are used to examine amounts and magnitudes.

There are four main scales of measurement subsumed under the broader categories of non-metric and metric measurement:

1. Nominal scales
2. Ordinal scales
3. Interval scales, and
4. Ratio scales

- Nominal and ordinal scales are non-numeric measurement scales.

1. Nominal scales:- are the least sophisticated type of measurement and are used only to qualitatively classify or categorize.

- They have no absolute zero point and cannot be ordered in a quantitative sequence, and there is no equal unit of measurement between categories.

- They do not imply amounts of an attribute or characteristic.
  - This makes it impossible to conduct standard mathematical operations such as addition, subtraction, division, and multiplication.
- Examples of nominal scale data include **gender, religious and political affiliation, place of birth, city of residence, ethnicity, Marital status, eye and hair color, and employment status**

2. Ordinal scale:- measurement is characterized by the ability to measure a variable in terms of both identity and magnitude.

- This makes it a higher level of measurement than the nominal scale because the ordinal scale allows for the categorization of a variable and its relative magnitude in relation to other variables
- Variables can be ranked in relation to the amount of the attribute possessed
- In simpler terms, ordinal scales represent an ordering of variables, with some number representing more than another

Distinguishing Characteristics of Ordinal Measurement Scales and Data

- Build on nominal measurement.
- Categorize a variable and its relative magnitude in relation to other variables.
- Represent an ordering of variables with some number representing more than another.
- Information about relative position but not the interval between the ranks or categories.
- Qualitative in nature.

Example - finishing position of runners in a race.

  - Lack the mathematical properties necessary for sophisticated statistical analyses.

3. Interval scale:- of measurement- builds on ordinal measurement by providing information about both order and distance between values of variables.

- The numbers on an interval scale are scaled at equal distances, but there is no absolute zero point. Instead, the zero point is arbitrary.

Distinguishing Characteristics of Interval Measurement Scales and Data

- Quantitative in nature.
- Build on ordinal measurement.
- Provide information about both order and distance between values of variables.
- Numbers scaled at equal distances.
- No absolute zero point; zero point is arbitrary.
- Addition and subtraction are possible.
- Examples include temperature measured in Fahrenheit and Celsius.
- Lack of an absolute zero point makes division and multiplication impossible

4. Ratio scale:- The properties of the ratio scale are identical to those of the interval scale, except that the ratio scale has an absolute zero point, which means that all mathematical operations are possible.

- Example –Money- It is possible to have no (or zero) money—a zero balance in a checking account is an example of an absolute zero point.

Characteristics:

- Identical to the interval scale, except that they have an absolute zero point.
- Unlike with interval scale data, all mathematical operations are possible.

Examples include height, weight, and time.

- Highest level of measurement.
- Allow for the use of sophisticated statistical techniques.

Data collection Techniques/Tools

- Applying the chosen method(s).
- The means to collect data for analysis.
- There are a number of different data collection tools available and one should be selected which is most likely to meet the objective of the research and gather the correct type of information for your study.

- Each technique is designed to get certain types of information and not others.
- Viable methods/tools should be weighed up in terms of their **advantages** and **disadvantages**.
- Some of the data collection techniques/tools with their advantage and disadvantages are discussed here.
  - Observation
  - Questionnaires,
  - Interviews,
  - Document analysis

Observation

- Observation refers to the process of observing and recording events or situations.
- Particularly useful for discovering how individuals or groups of people or animals (and in some instances inanimate objects) behave, act or react.
- There are two main types of observation:
  - Participant and
  - Non-participant

- **In Participant observation**: The researcher becomes part of the group studied and participates in their daily life and activities:
  - Observing their everyday situations and their behavior in these situations.
  - This observation is usually limited to studies of **human subjects**.
- **In non-participant observation:** The researchers simply observe the activities without taking part themselves.
  - Has the advantage of preventing the researcher from unduly influencing or becoming involved in activities they may not wish to take part in (for example dangerous or criminal actions),
  - But they are less likely to understand fully the meanings behind behavior in the group studied.
  - Beside the study of human subjects, non- participant observation can also be used to study animal behavior.

Advantages and disadvantages of observation

Advantages	Disadvantages
Requires little training or familiarization.	Time consuming.
Can understand meanings behind actions.	Problems with recording data.
Behavior can be observed in its natural environment, the subject is undisturbed.	Can only study a small group.
Can study deviant groups.	Cannot make generalizations - no way of judging whether the group is typical.
Flexibility - researcher may come across conditions and events previously not comprehended.	If covert is it ethical?
	Moral, legal and injury risks associated with this method.

Questionnaires

- A type of survey where respondents write answers to questions posed by the researcher on a question form.
- Questionnaires are extremely flexible and can be used to gather information on almost any topic involving large or small numbers of people.
- The two approaches of questionnaires are:
  - Close-ended questionnaire and
  - Open-ended questionnaire

- **Close-ended questionnaire**: The approach where the respondents are required to answer by choosing an option from a number of given answers,
  - Usually by ticking a box or circling an answer.
- They only gather straightforward, uncomplicated information, and only simple questions can be asked.
- **The open-ended questionnaire:** differs in that it allows the respondent to formulate and record their answers in **their own words.**
  - Are more qualitative and can produce detailed answers to complex problems.

Advantage and disadvantages of questioners:

Advantages	Disadvantages
Quick.	Limited answers only can be given.
Cheap.	Lack of qualitative depth results in superficiality.
Efficient.	No way of probing/searching for more information in superficial responses.
Can reach a large number of people.	Not always accurate - not possible to verify what appears to be an inaccurate answer and little check on honesty of responses. ***Questions may mean different things to different people.***
Consistent format means there is little scope for bias introduced by different researchers.	Predetermined boxes may not be appropriate.
	Low response rate.
	Construction difficult - instructions must be clear and unambiguous and questions carefully ordered.

Interviews

- Are limited to cases where the subjects of study are humans.
- Are a type of survey where questions are delivered in a face-to-face encounter by an interviewer.

Advantages and disadvantages of Interview

Advantages	Disadvantages
High response rate.	Limited sample only.
Can collect complex information.	Can be difficult to analyze (especially in-depth interviews).
High degree of researcher control achieved.	May be a hostile reaction.
Can be made more responsive to early results.	Whole process is time consuming.
Relaxed environment.	Recording techniques may cause problems.
	There is room for interviewer bias - this should be acknowledged.

Document Analysis

- This refers to the process of using any kind of document, films, television programs and photographs as well as written sources, such as books, papers and letters, for analysis in relation to a particular research question.
- Document analysis, also referred to as content analysis, differs from the majority of research methods in two major ways.
  - It is an indirect form of research; it is something that has been produced, so the investigator is not generating original data.
  - It is an 'unobtrusive', or 'non-reactive' method. This refers to the fact that the document will not be affected in any way by your research; it cannot react as a human can.

Advantages and Disadvantages of Document analysis

Advantages	Disadvantages
The data never alters and can be subject to reanalysis.	Subject to bias and subjectivity - impossible to allow for biases introduced by the fact that the document studied has been written for a particular purpose and is the author's own particular account; events may be sensationalized, subject to political bias etc.
Unobtrusive.	Evidence may be out of date.
Events can be compared over time and cultures.	May not be accurately recorded.
Gives an expert understanding.	Documents available may be limited.
Cheap.	Can be laborious and time consuming.
Computers can aid analysis and lead to complete reliability in applying the rules you set down for coding the text.	

THANK YOU!!