An Introduction

Chapter 1: An Introduction

- The Meaning of Research
- The Characteristics of Research
- The Objectives of Research
- The Motivation for Research
- Types of Research
- Significance of Research
- Choosing a Research Topic
- Research Methods versus Methodology
- Research Ethics

What is Research ??

https://www.youtube.com/watch?v=mV0bUQpz468

The Meaning of Research

What is Research?

What is a famous study you have heard of ?

Meaning of Research ?

The Word ‘Research includes two parts = Re + Search.

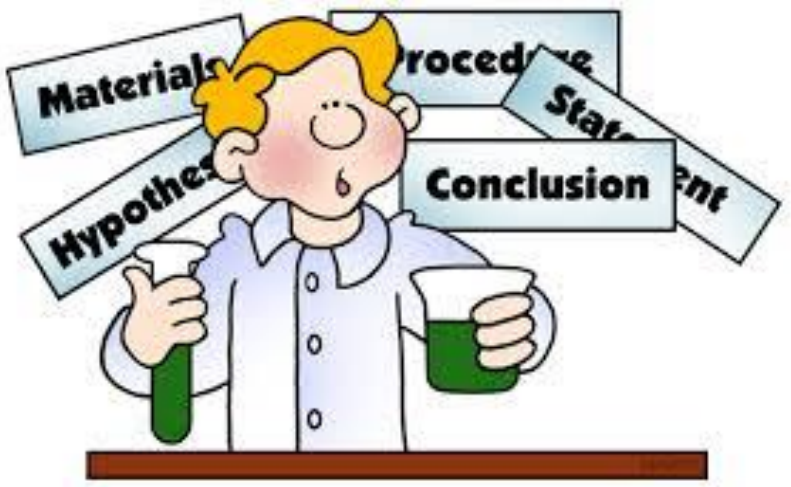
It means to search again

“ Arranged effort to gain knowledge”

- Research is a process of “ manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge.
- Scientific Research is the hunting for facts or truth about a subject
- Scientific Research is the Organized scientific investigation to :
  - Solve problems.
  - Test hypotheses.
  - Develop or discover new products.

Formally:

Scientific research is a systematic, controlled, experimental, and critical investigation of natural or social phenomena using a commonly accepted methodology to (in general):

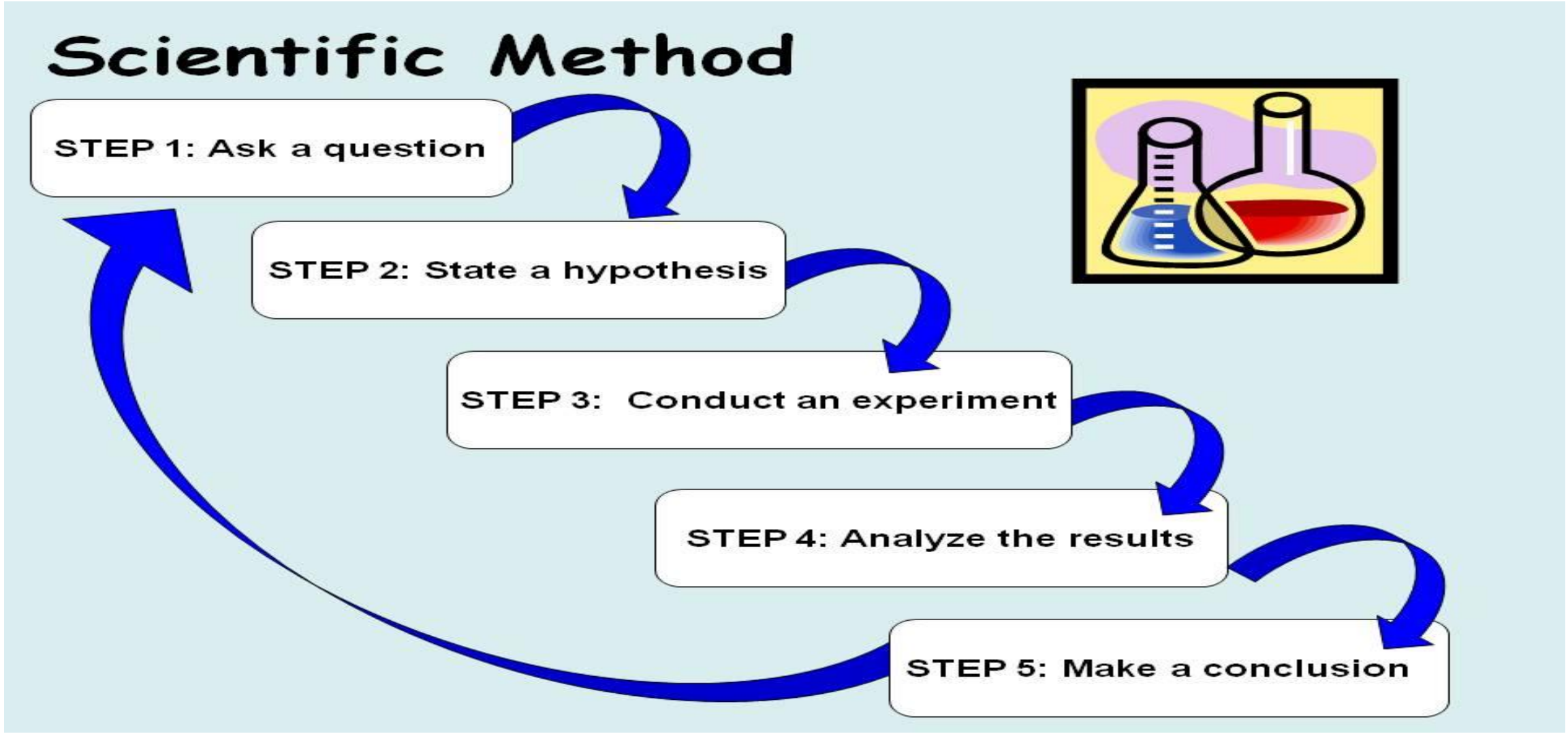


- ✓ Identify the problems
- ✓ Collect data
- ✓ Analyze and draw valid conclusions?

Scientific Research is systematic

because it follows certain steps that are logical in order, i.e.

- **Understanding the nature of the problem to be studied and identifying the related area of knowledge.**
- **Reviewing literature to understand how others have approached or dealt with the problem.**
- **Collecting data in an organized and controlled manner so as to arrive at valid decisions.**
- **Analyzing data appropriate to the problem.**
- **Drawing conclusions and making generalizations.**



The research process will be covered in details in Chapter 2

Important Characteristics of Scientific Research

1. Aim له هدف
2. Rigor الدقة
3. Testability إمكانية الفحص
4. Reliability يمكن الاعتماد عليه
5. Accuracy and Confidence الدقة والثقة
6. Objectivity الموضوعية
7. Generalizability إمكانية التعميم

Characteristics of Scientific Research

1. Aim

- The aim/purpose of scientific research must be written clearly for example:

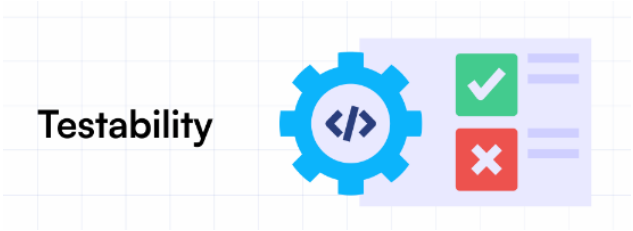
  - To reduce the interaction between iron and oxygen
  - To speed up the search process
  - To optimize an algorithm
  - To determine the factors that influence customer satisfaction towards a specific service

Characteristics of Scientific Research

2. Rigor (دقيق)

- Based on a good theoretical base and sound methodology
- Rigorous research means that it **applies the appropriate instruments to meet the stated objectives of the investigation** for the level of precision in the analysis and to ensure unbiased and well-controlled experimental design, methodology, analysis, interpretation and reporting of results..
- Examples:
  - To measure air pressure, we use an **unbiased** barometer.

Characteristics of Scientific Research



3. Testability

- Developing a set of research questions or hypotheses to be tested, must be **testable** and **analyzable.**

Examples:

- Increasing the processor speed in a computer will lead to faster program execution times compared to the current processor speed.
- For instance, the researcher might hypothesize that those employees who perceive greater opportunities for participation in decision making would have a higher level of commitment

Characteristics of Scientific Research

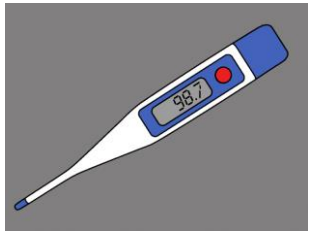
4. Reliability

- The same relationship is achieved repeatedly under the same design parameters.
- Results of the test of research objectives should be supported again and again when the same type of research is being repeated in other similar circumstances

Examples:

- You have a reliable measurement if you dip the thermometer into the water multiple times and get the same reading each time.

- In a study on encryption algorithms, the researchers employed a diverse set of test cases. The same relationships where obtained each time



Characteristics of Scientific Research

5. Accuracy & Confidence

- **Accuracy** refers to the closeness of the findings to reality (e.g. $g = 9.8\ m/s^2$)
- **Confidence** refers to the probability that our estimations are correct. (الثقة بالنتيجة)

In research, there are two types of errors type I and type II.

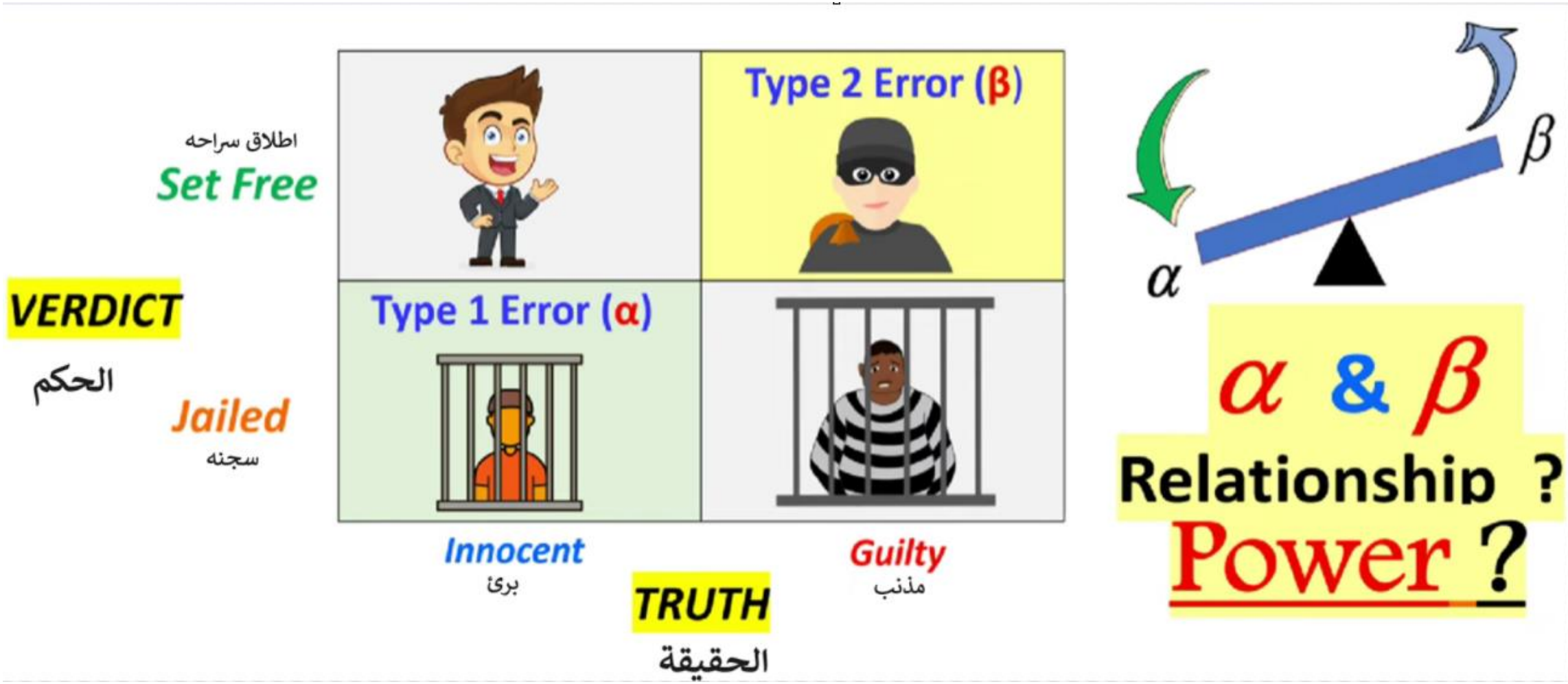
	Null Hypothesis is TRUE	Null Hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct Outcome! (True positive)
Fail to reject null hypothesis	Correct Outcome! (True negative)	Type II Error (False negative)

Type I (α): Reject when it is True

Type II (ß): Accept (Fail to reject) when it is False

In statistical hypothesis testing, researchers aim to control and minimize these errors by choosing appropriate significance levels and sample sizes.

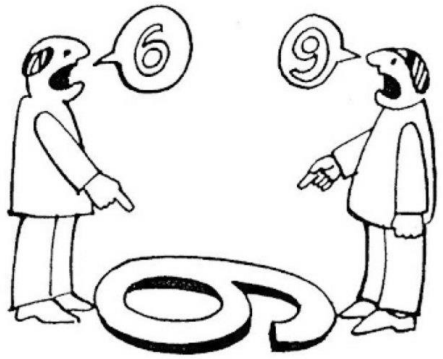
الفرضية: المتهم برئ حتى تثبت ادانته



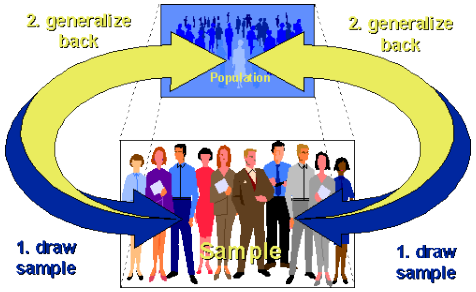
Characteristics of Scientific Research

6. Objectivity

Conclusions drawn from the analysis of data must be objective. It is proved from development of data and **Supported by the theoretical model.**



Characteristics of Scientific Research



7. Generalizability

Scope of applicability:

- The wider, and better in the applicability.

For example, Research in computer science should be applicable beyond specific cases. For example, in studying cybersecurity measures, a focus on generalizability involves testing security protocols across different operating systems to ensure effectiveness in various computing environments.

The objectives of Research

1. Revealing a new, never-before-seen truth. الكشف عن حقيقة جديدة لم يسبق لها مثيل (استكشافية)
2. Completing research or studies that have not been completed by their author, or that need to be completed in some aspects.
3. Explaining something that is not clear and difficult to understand from what was written by predecessors. شرح شيء غير واضح ويصعب فهمه مما كتبه السلف
4. To portray accurately the characteristics of a particular individual, situation or a group (descriptive) تصوير دقيق لخصائص فرد أو موقف أو مجموعة معينة
5. Increase the knowledge acquired by humanity زيادة المعرفة المكتسبة للبشرية
6. Innovation and renewal الابتكار و التجديد
7. To test a hypothesis of a causal relationship between variables (hypothesis-testing research studies)

The Motivation for Research

Researchers conduct research for several reasons including:

1. Desire to face the challenge in solving the unsolved problems, i.e., cancer treatment
   الرغبة في مواجهة التحدي في حل المشاكل التي لم يتم حلها، مثلا علاج السرطان

2. Desire to get intellectual joy of doing some creative work; الرغبة في الحصول على المتعة الفكرية
   من خلال القيام ببعض الأعمال الإبداعية

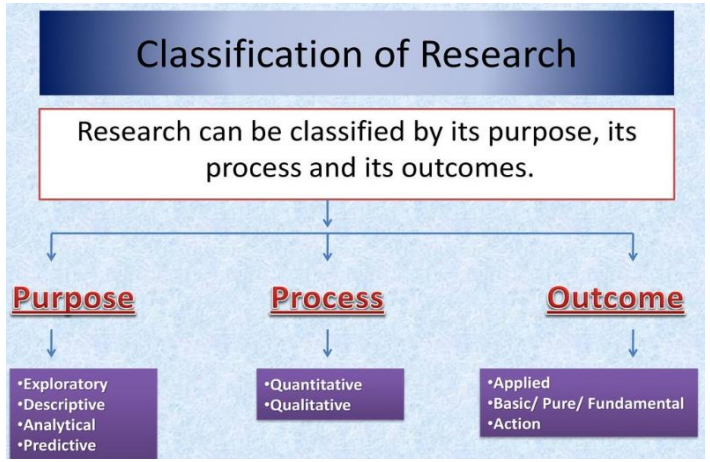
3. Desire to be of service to society; الرغبة في خدمة المجتمع

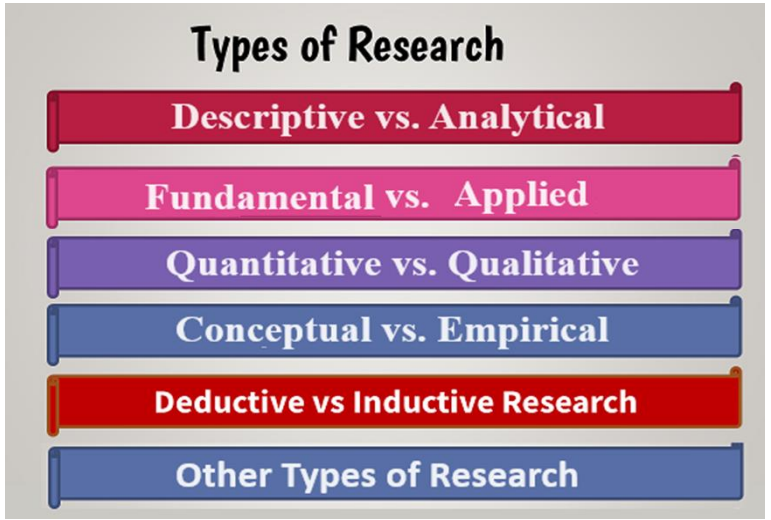
4. Desire to get a research degree along with its consequential benefits; الرغبة في الحصول على درجة
   بحثية مع ما يترتب عليها من فوائد

- Many more factors such as directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking and awakening, and the like may as well motivate (or at times compel) people to perform research operations.

Types of Research

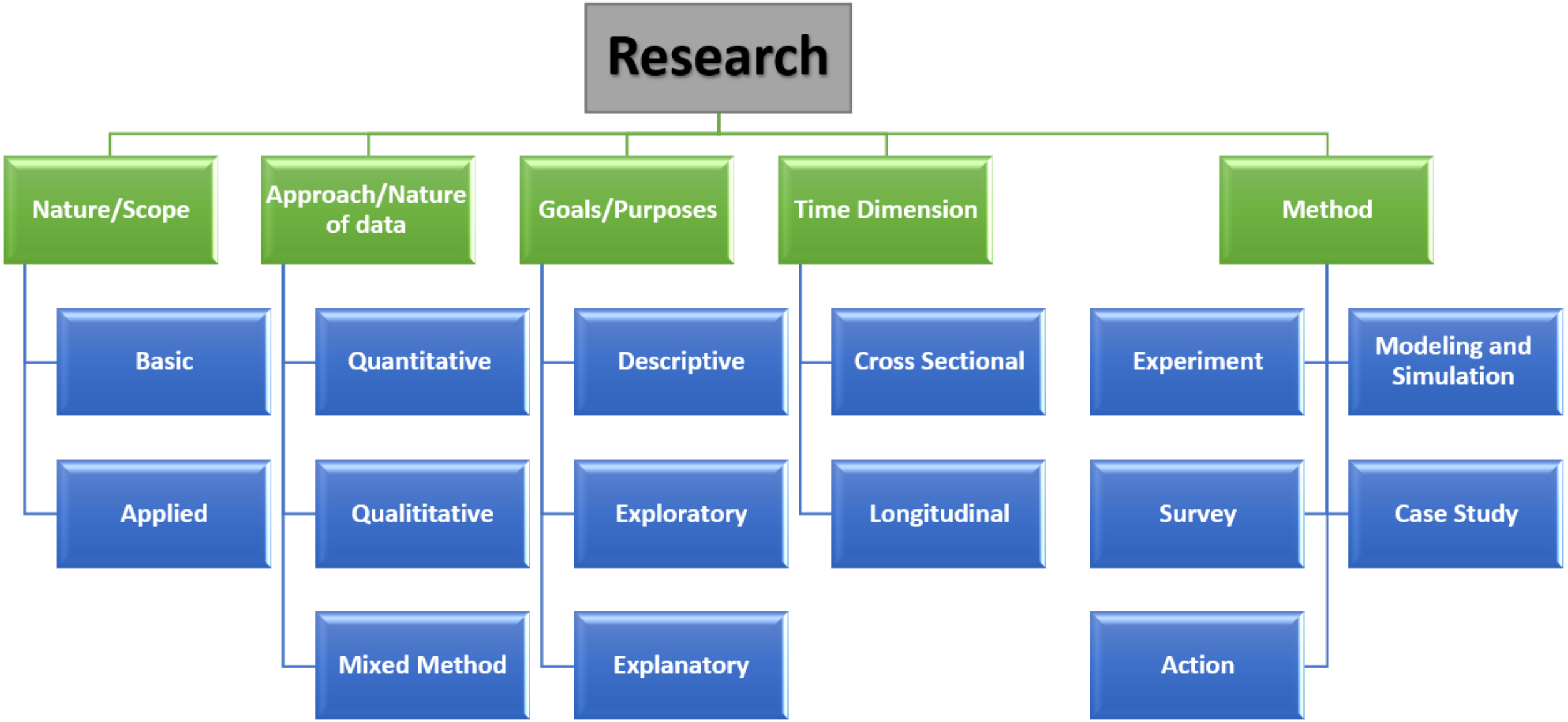
- Several classifications of research exist based on different criteria including the nature, the purpose, the time frame of the research, in addition to other criteria as well.
- These classifications may use different terms to address research types.
- Some of these classifications are :





- **The following classification will be adopted in this course:**

Types of Research



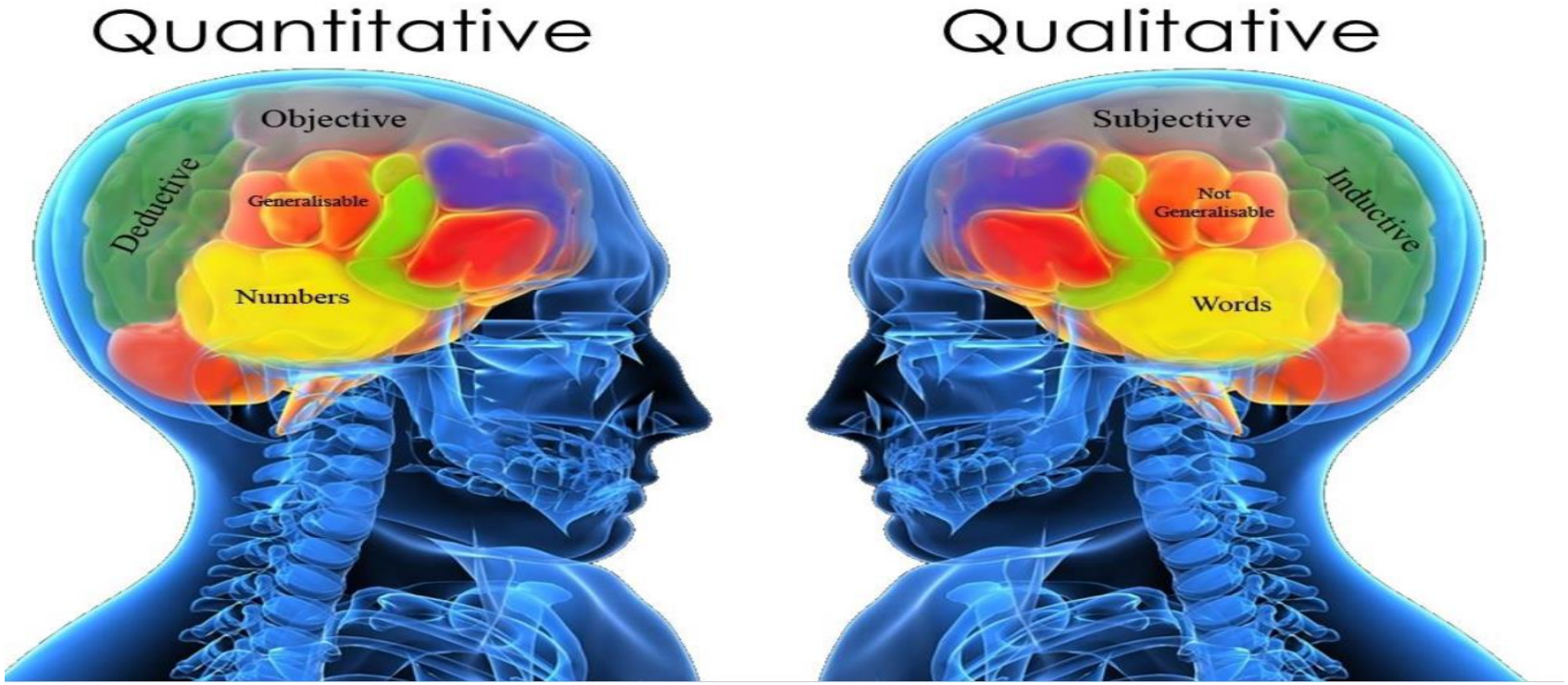
1. Types of Research Based on the Nature/Scope

Aspect	Basic (Fundamental)الأساسي	Applied تطبيقي
Definition	It is used to understand and extend our knowledge about a specific phenomenon or field. It is also accepted as pure investigation or fundamental	Applied research is a non-systematic way of finding solutions to specific research problems or issues. These problems or issues can be on an individual, group, or societal level. It is called "non-systematic" because it goes straight to finding solutions.
Focus	It is Interested in answering "Why" and "How"	It is Interested in answering "What"
Objective	Mainly concerned with generalizations and with the formulation of a theory that adds to the already existing organized body of scientific knowledge.	Aims at finding a solution for an immediate problem facing a society or an industrial /business organization.
Examples	• Studying the best factors of pricing strategies. • Natural phenomenon	• In Human-Computer Interaction (HCI), developing new input devices, designing user interfaces for specific applications, or creating accessibility features to make technology more inclusive. • Applied study to create a plan to keep employees coming to work regularly.

2. Types of Research Based on the Approach / Nature of data

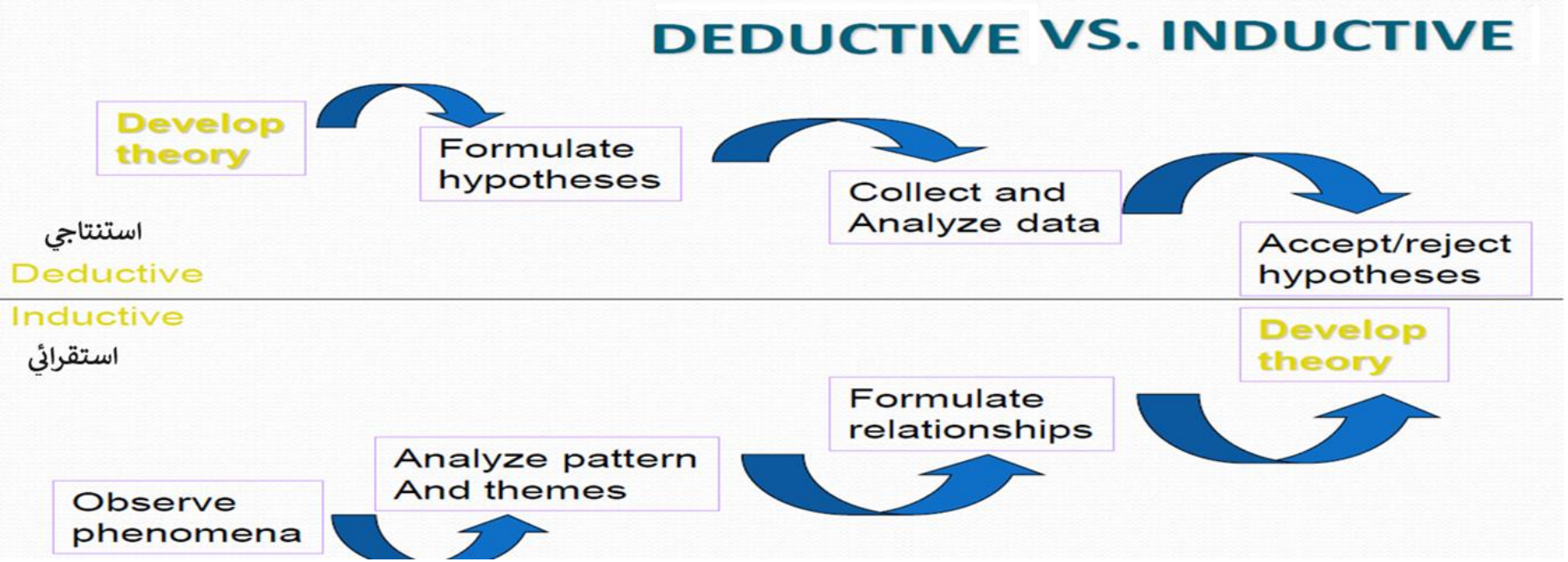
Aspect	Quantitative الكمي	Qualitative النوعي	Mixed Methods
Definition	It is the analyzing and gathering of numerical data to uncover trends, calculate averages, evaluate relationships, and derive overarching insights	It the research that allows for in-depth and further probing and questioning of respondents based on their responses. It is based on the disciplines of social sciences like psychology, sociology, and anthropology.	• Mixed methods research combines the elements of two types of research: quantitative and qualitative. • Mixed methods research allows for a more thorough exploration of a research question. • It can answer complex research queries that cannot be solved with either qualitative or quantitative research. • An example: How have work outputs at XYZ Company changed since the shift to fully remote work?
Focus	Applicable to phenomena that can be expressed in terms of **quantity**	Finding out how people feel or what they think about a particular subject or institution. As well as discovering the underlying motives of human behavior.	
Data collection Instrument	Use highly structured methods such as structured observation, surveys, experiments	Use semi-structured methods such as in-depth interviews, focus groups, and participant observation	
Form of data produced	Numerical data (Reduced to numeric codes)	Descriptive data (Verbal or pictorial. Reduced to verbal codes)	
Examples	Researchers may measure factors like query execution time, database throughput, and resource utilization to optimize database systems.	Case Studies of Technology Implementation within an organization, including challenges, successes, and the impact on workflows.	

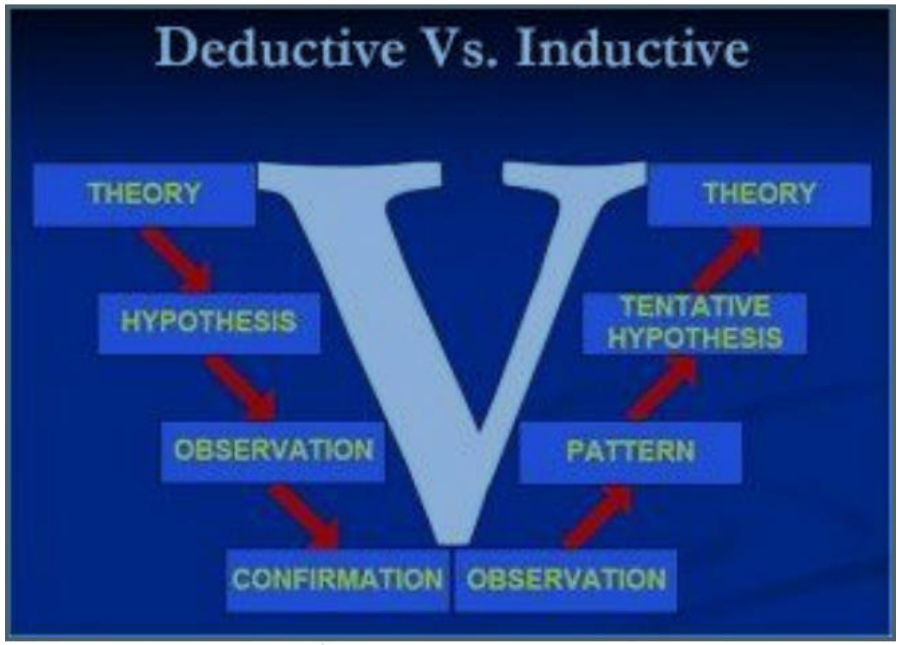
The Quantitative approach is deductive and objective, while the Qualitative approach is inductive and subjective



Objective vs. Subjective

Deductive vs. Inductive


DEDUCTIVE VS. INDUCTIVE
Develop theory
Formulate hypotheses
Collect and Analyze data
Accept/reject hypotheses
استنتاجي
Deductive
Inductive
استقرائي
Develop theory
Formulate relationships
Analyze pattern And themes
Observe phenomena


Deductive Vs. Inductive
THEORY
THEORY
HYPOTHESIS
TENTATIVE HYPOTHESIS
OBSERVATION
PATTERN
CONFIRMATION
OBSERVATION

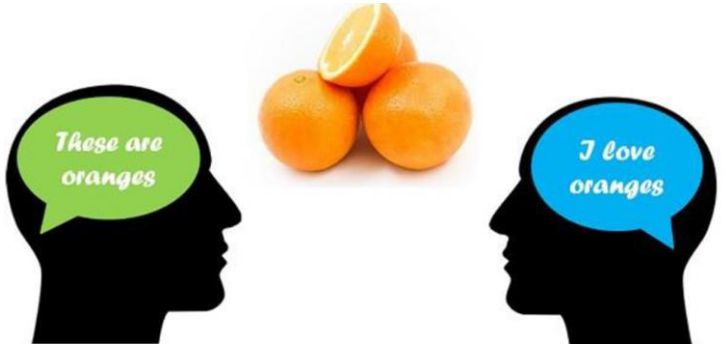
Deductive vs Inductive Research

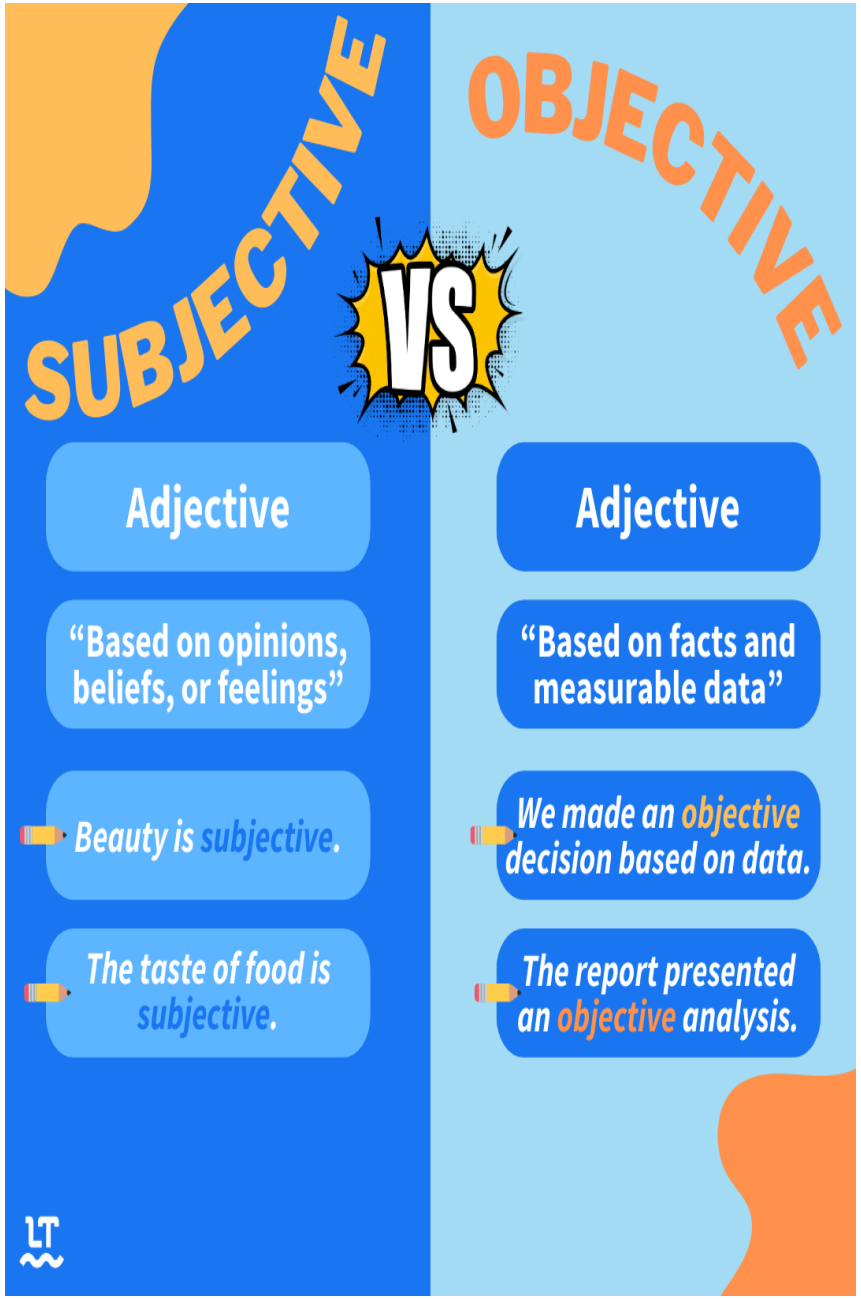
Aspect	Deductive استنتاجي	Inductiveاستقرائي
Definition	Deductive research is a type of research in which the researcher starts with a theory, hypothesis, or generalization and then tests it through observations and data collection.	Inductive research is a method in which the researcher collects and analyzes data **to develop theories**, concepts, or hypotheses based on patterns and observations seen in the data.
Approach	It uses a top-down method in which the researcher starts with a general idea and then tests it through specific observations	It uses a "bottom-up" method in which the researcher starts with specific observations and then moves on to more general theories or ideas.
Usage	Deductive research is often used to confirm a theory or test a well-known hypothesis.	It is often used in exploratory studies or when not much research has been done on a topic before.
Methods used	Deductive research uses more quantitative methods, like statistical analysis, to test and confirm the theory or hypothesis.	Inductive research uses more qualitative analysis, like textual or visual analysis, to find patterns and themes in the data.

Final Note:

- **Subjective information or writing** is based on personal opinions, interpretations, points of view, emotions and judgment.
- **Objective information or analysis** is fact-based, measurable and observable.

- **Subjective is an opinion**
- **Objective is a fact**





3. Types of Research Based on Goals / Purposes

Aspect	Descriptive وصفي	Exploratory استكشافي	Explanatory تفسيري / تحليلي
Definition	Describing the characteristics of the population or phenomenon studied.	It investigates research questions that have not previously been studied in depth.	It calls for using critical thinking abilities and assessing data and information pertinent to the project at hand.
Focus	It is Interested in answering "What"	It is Interested in answering what is the problem? What is the purpose of the study? And what topics could be studied?	It is Interested in answering "Why" and "How"
Objective	It helps in documenting and characterizing phenomena by describing the state of at present	It is an interpretive research or a grounded theory approach due to its flexible and open-ended nature.	Analytical research involves a deeper analysis to derive insights, solve problems, or make predictions
Data Analysis	Summarizing the information	It is often qualitative and primary in nature.	Statistical research, hypothesis testing, qualitative research
Causal Relationships العلاقات السببية	Not the primary focus	Not the primary focus	Examining underlying factors, causes, or effects
Methods used	Surveys, observations, case-control study, content analysis	Surveys, observations, focus groups, interviews, secondary data.	Experiments, statistical research, qualitative analysis

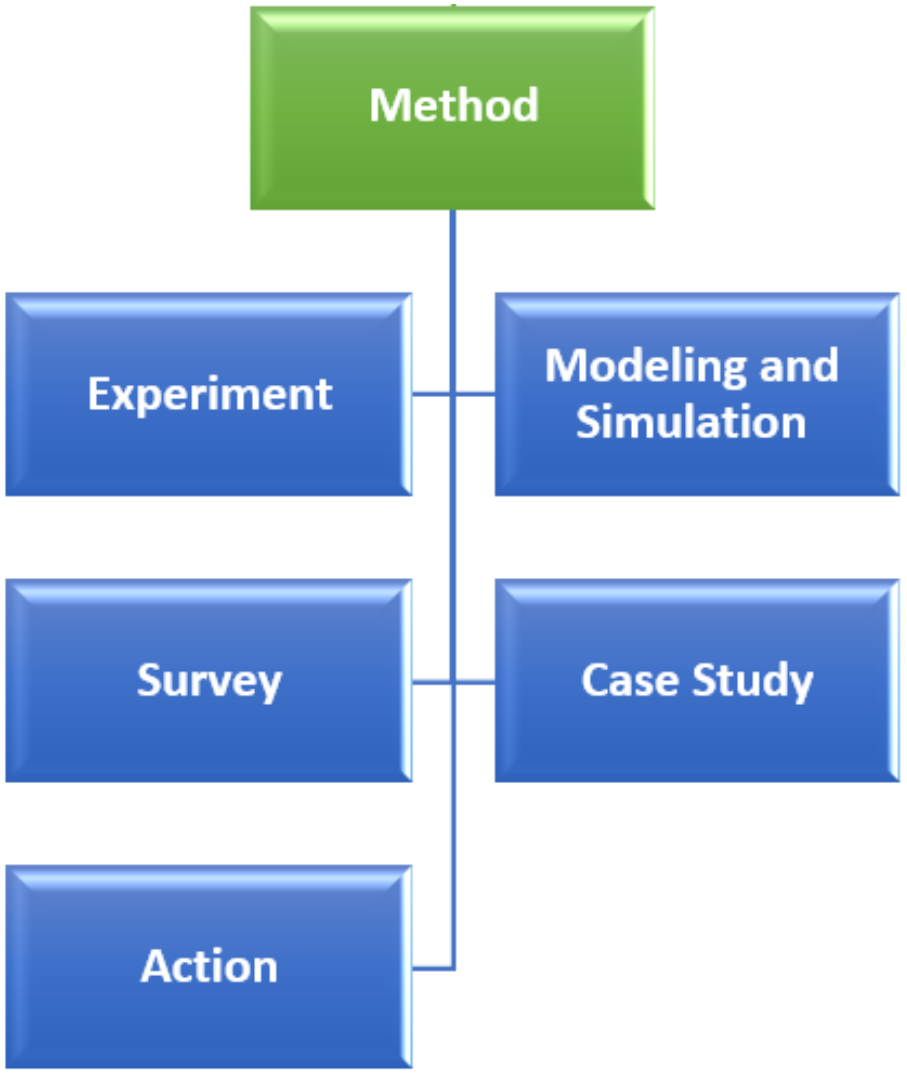
Examples:

Descriptive وصفي	Exploratory استكشافي	Explanatory
User Interface (UI) Analysis Researchers may conduct surveys, interviews, and usability tests to describe the preferences and behaviors of users when interacting with the UI. The goal is to provide a detailed understanding of how users interact with the system.	Interaction with Emerging AI Interfaces To explore how users engage with newly developed AI interfaces that leverage natural language processing (NLP) for various tasks, such as voice commands, chatbots, or virtual assistants.	Algorithm Performance Comparison Analyze the time complexity and space complexity of various sorting algorithms to determine which one is more efficient under specific conditions.
Market researchers want to observe the habits of consumers.	To understand user feedback on new features added to a mobile app.	Why the value of the Japanese Yen has decreased. This is so that an analytical study can consider “how” and “why” questions.
A school district wants to understand if students will access online lessons rather than textbooks.	A study into the role of social networking sites as an effective marketing communication channel	Why students are not satisfied with the university portal?

4. Types of Research Based on Time Dimension

Aspect	Cross Sectional	Longitudinal
Definition	Is the study where data is collected as a whole to study a population at a single point in time to examine the relationship between variables of interest.	Is a study in which data is gathered from the same sample repeatedly over an extended period. It can last from a few years to even decades depending on what kind of information needs to be obtained.
Focus	Cross-sectional studies cannot pin down cause-and-effect relationship.	Longitudinal study can justify cause-and-effect relationship.
Sample	It is conducted with different samples.	Longitudinal study is conducted with the same sample over the years.
Cost	Cross-sectional study is comparatively cheaper.	Since the study goes on for years longitudinal study tends to get expensive.
Speed	Cross-sectional studies are quick to conduct as compared to longitudinal studies.	Longitudinal studies may vary from a few years to even decades
Example	A cross-sectional study on the prevalence of cancer among women may give a generalized opinion that the illness often occurs in middle-aged subjects.	Conversely, the longitudinal study may observe how cancer affects women when they eat certain foods or perform specific activities.

Types of Research Based on Method used



5. Types of Research Based on Method

Aspect	Survey	Case Study	Action
Definition	A method for collecting data from a large group of people through questionnaires or interviews to gather quantitative or qualitative information about specific topics or phenomena.	A detailed examination of a single instance or a small number of instances (such as an organization, event, or individual) to explore complex issues in real-life contexts.	A participatory research method where the researcher works collaboratively with participants to identify a problem, develop a solution, and implement and evaluate actions for change
Features	• Surveys can collect both quantitative data (numerical) and qualitative data (opinions, attitudes, and beliefs). • Surveys can reach a large number of respondents, making it easier to generalize findings	• Provides an in-Depth Analysis. • Contextual Understanding: Provides insights into how and why certain phenomena occur within their specific context. • Flexible Design: Could be exploratory, explanatory, or descriptive, depending on the research objective	• Involves close collaboration between researchers and participants (e.g., community members, employees). • Problem-Solving Focus: Aims to solve practical problems and improve practices
Pros. +ve	• Can collect data from a large, diverse population • Cost-effective	• Provides rich, detailed data and a holistic view of the subject. • Can generate hypotheses for further study and contribute to theory building.	• Directly applicable results that can improve practices and processes. • Empowers participants in the research process.
Cons. -ve	• Bias in self-reported data • Low response rate	• Limited generalizability due to the focus on a single or small number of cases. • Can be time-consuming and resource-intensive.	• May face resistance from participants who are not open to change. • Results may not be easily generalizable beyond the specific context studied.
Example	• National censuses that collect demographic information. • Customer satisfaction surveys	• Studying the development of a new technology within a specific industry.	• Collaborating with businesses to improve workplace practices and employee satisfaction

5. Types of Research Based on Method

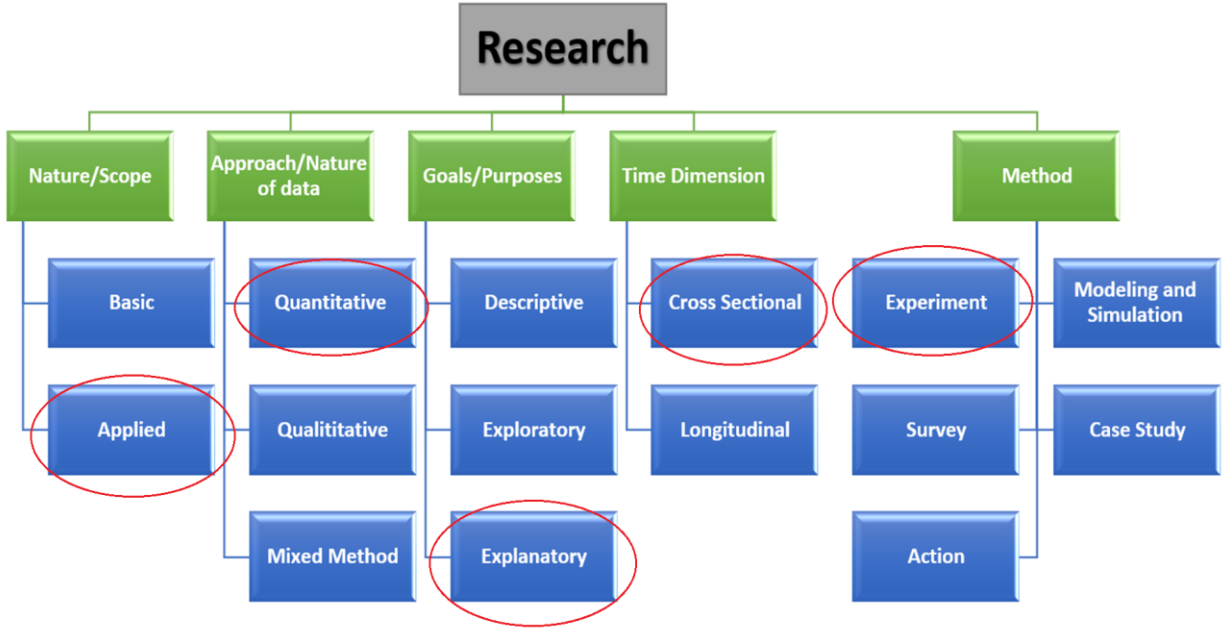
Aspect	Experiment	Modeling and Simulation
Definition	A method involving controlled testing of hypotheses, typically in a laboratory or controlled environment, to establish cause-and-effect relationships.	A research method that uses mathematical models, simulations, and computational techniques to represent real-world processes, allowing researchers to study complex systems in a controlled, replicable manner
Features	• Controlled Environment: Experiments are usually conducted in a laboratory, where extraneous variables can be minimized. • Random Assignment: Subjects are randomly assigned to different groups to ensure that any observed effects are due to the manipulation of the independent variable, not preexisting differences between groups. • Replication: Experiments can be replicated by other researchers	• Mathematical or Computational Models: Uses mathematical equations or algorithms to represent the key components and relationships of the system. • Dynamic Analysis: Simulations allow for the analysis of how systems behave over time, providing insights into complex dynamics that may be difficult to study directly. • Scalability: Models can be adjusted to represent systems at different scales, from small to large-scale phenomena.
+ve	• Allows for precise control over variables. • Can establish causal relationships between variables.	• Cost-effective way to study complex systems • Enables testing of scenarios that are impractical or impossible to experiment with directly (e.g., climate change models).
-ve	• May lack ecological validity (real-world applicability) due to the artificial environment. • Ethical constraints may limit the types of experiments • Results may not generalize well to real-world settings.	• The accuracy of the simulation depends heavily on the quality of the model and assumptions made. • Can be computationally intensive and require specialized software • May not account for all real-world variables, leading to potential oversimplifications.
Example	• Clinical trials testing the effectiveness of a new drug.	• Climate modeling to predict weather patterns and climate change.

Types of Research Based on Method

How to select the proper research type for your study?

The selection depends on the objective and the research questions.

Example:
In a study, the research question was: "What is the effect of different sorting algorithms on the processing time of large datasets?"
Here, the aim is to study the effect of one factor on another using quantitative data that is collected by conducting an experiment at a specific time. Then,

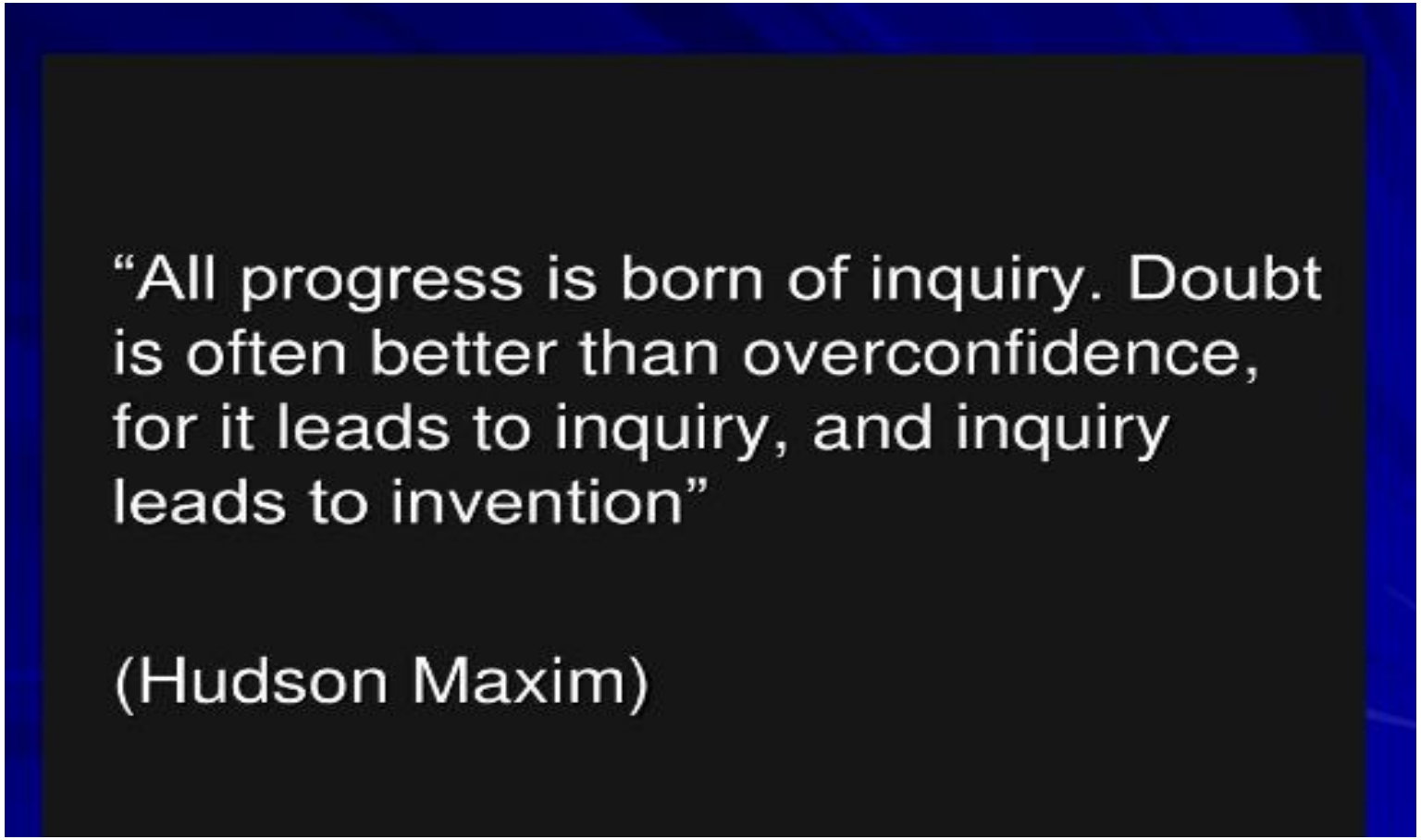


The experimental approach and the Survey approach will explained in detail in Unit 3 (Research Design) as they are most commonly used in computer science and information systems studies.

Other Types of Research

- Research can be *field-setting research or laboratory research or simulation research*, depending upon the environment in which it is to be carried out.
- And other types

Significance of Research




Research Significance
It builds knowledge and promotes learning
Research promotes success in business and other fields.
It helps to increase public awareness
It promotes confidence and passion in reading, sharing information, analyzing, and writing.
Research is a means for discovering opportunities and helps build credibility

Choosing a Research Topic

Choosing **an interesting research topic** is your first **challenge**. Here are some tips:

1. Choose a topic that you are **interested in**! The research process is more relevant if you care about your topic.
2. Narrow your topic to something manageable (**Keep the topic focused**)**.** If your topic is too broad, you will find too much information and not be able to focus.
3. **Ensure the topic is relevant:** Background reading can help you choose and limit the scope of your topic.
4. Ensure you have **enough resources** to support your research.
5. Talk about research ideas with a friend. S/he may be able to help focus your topic by discussing issues that didn't occur to you at first.

Choosing a Research Topic

- Think of the who, what, when, where and why questions:
  - **WHY** did you choose the topic? What interests you about it? Do you have an opinion about the issues involved?
  - **WHO** are the information providers on this topic? Who might publish information about it? Who is affected by the topic? Do you know of organizations or institutions affiliated with the topic?
  - **WHAT** are the major questions for this topic? Is there a debate about the topic? Are there a range of issues and viewpoints to consider?
  - **WHERE** is your topic important: at the local, national or international level? Are there specific places affected by the topic?
  - **WHEN** is/was your topic important? Is it a current event or an historical issue? Do you want to compare your topic by time periods?

Choosing a Research Topic

Characteristics of a good research topic

(1) Appropriateness الملاءمة

- ✔ Does the research topic fit the specifications and meet the standards set by the examining institution?
- ✔ Does the research topic contain issues that have a clear link to theory?
- ✔ Are you able to state your research question(s), aim and objectives clearly?
- ✔ Will the proposed research be able to provide fresh insights into this topic?
- ✔ Are the findings for this research topic likely to be symmetrical: that is, of similar value whatever the outcome?

(2) Capability المقدرة

- ✔ Do you have, or can you develop within the project time frame, the necessary research skills to undertake the research topic?
- ✔ Is the research topic achievable within the available time?
- ✔ Is the research topic achievable within the financial resources that are likely to be available?
- ✔ Are you reasonably certain of being able to gain access to data you are likely to require for this research topic?

(3) Fulfilment الاشباع

- ✔ Does the research topic really interest and motivate you?
- ✔ Will the research topic help towards the achievement of your future aspirations or career goals?

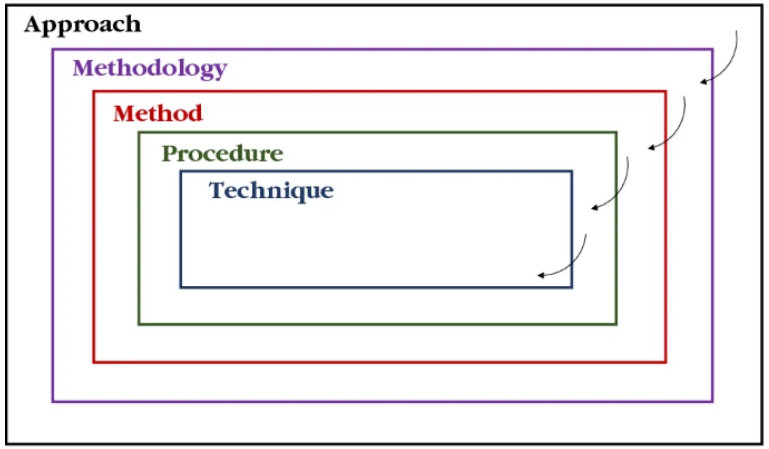
Research Methods versus Methodology

❖Research methodology

- Is a way to systematically solve the research problem.
- The scope of research methodology **is wider than** that of research methods.
- Research methodology involves
  - The research methods
  - The logic behind the methods we use,
  - Explain why we are using a particular method or technique, and why we are not using others

❖ Research methods

- Refer to the methods the researchers use in performing research operations, Like
  - Collection data methods
  - Statistical techniques
  - Evaluation methods.



Methodology vs. Method

Methodology

- It is the beginning.
- The technique(s)/how to conduct research.
- The study about the tools of research.
- Explains the methods by which you may proceed with your research.

Method

- The end of any scientific or non-scientific research.
- The actual tools/steps taken by which you conduct research into a subject or a topic.

What are research ethics?

1. A code of guidelines on how to conduct scientific research in a morally acceptable way.

OR

2. Principles and standards that help researchers to uphold the value and standards of knowledge construction.

https://www.youtube.com/watch?v=mtLPd2u4DiA

Ethical considerations in the research process

Ethical considerations come into play at six stages of research

1. **Conceptualisation and design of the study (scientific merit, identify risks and ways to mitigate the risks)**
2. **When participants are recruited (the process of informed consent, right to privacy)**
3. **During the intervention or measurement procedure to which participants are subjected (management of risk)**
4. **In the release of results obtained**
5. **(protection of confidentiality and anonimity)**
6. **After the release of results (ensure that participants and communities involved in the research benefit)**

Why are ethics important in research?

- ✓ We should protect the rights of people who participate in research.
- ✓ Unethical research can harm people and communities.
- ✓ Unethical research can lead to distrust of science and research, and ultimately hinders scientific progress.

How do research ethics apply to me?

Ethics apply to all aspects of research:

- ✓ How we interact with participants
- ✓ How we collect and store data
- ✓ How we report findings

EVOLUTION OF RESEARCH ETHICS, CODES AND REGULATIONS: INTERNATIONAL LANDSCAPE

A VIDEO ON ETHICS

https://www.youtube.com/watch?v=kA1dL6NqVyw

Three basic principles

While conducting the research enterprise, the following three fundamental principles of research must be upheld.

- Respect for persons
- Beneficence (“do no harm”)
- Justice

Respect for persons

- Obtaining informed consent.
  - ✓ Make sure that people know that the research is voluntary.
  - ✓ Give people time to ask questions so that they can decide whether they want to participate.
  - ✓ Do not unfairly influence people to participate.
  - ✓ Even after someone decides to participate, make sure that he/she understands that he/she can decline to answer a question or stop at any time.
- Protecting vulnerable people and groups.
  - ✓ Protecting privacy and confidentiality.

Beneficence and non-maleficence (“do no harm”)

- Balancing the potential benefits of the research with the risks of harm.

- Examples of harm are physical harm (e.g., violence and discrimination owing to breach of privacy), psychological harm (e.g., emotional distress), and social harm (e.g., loss of a job).
- Benefits may be to society rather than to individual participants.
- The goal is to minimize the risk of harm to individuals and groups.

Justice

- Choosing research participants and groups fairly
- Making sure that everyone who could benefit from the research has the opportunity to participate
- Working to bring beneficial results of research to people and groups who have contributed to the research

Research Process

Unit 2: Research Process

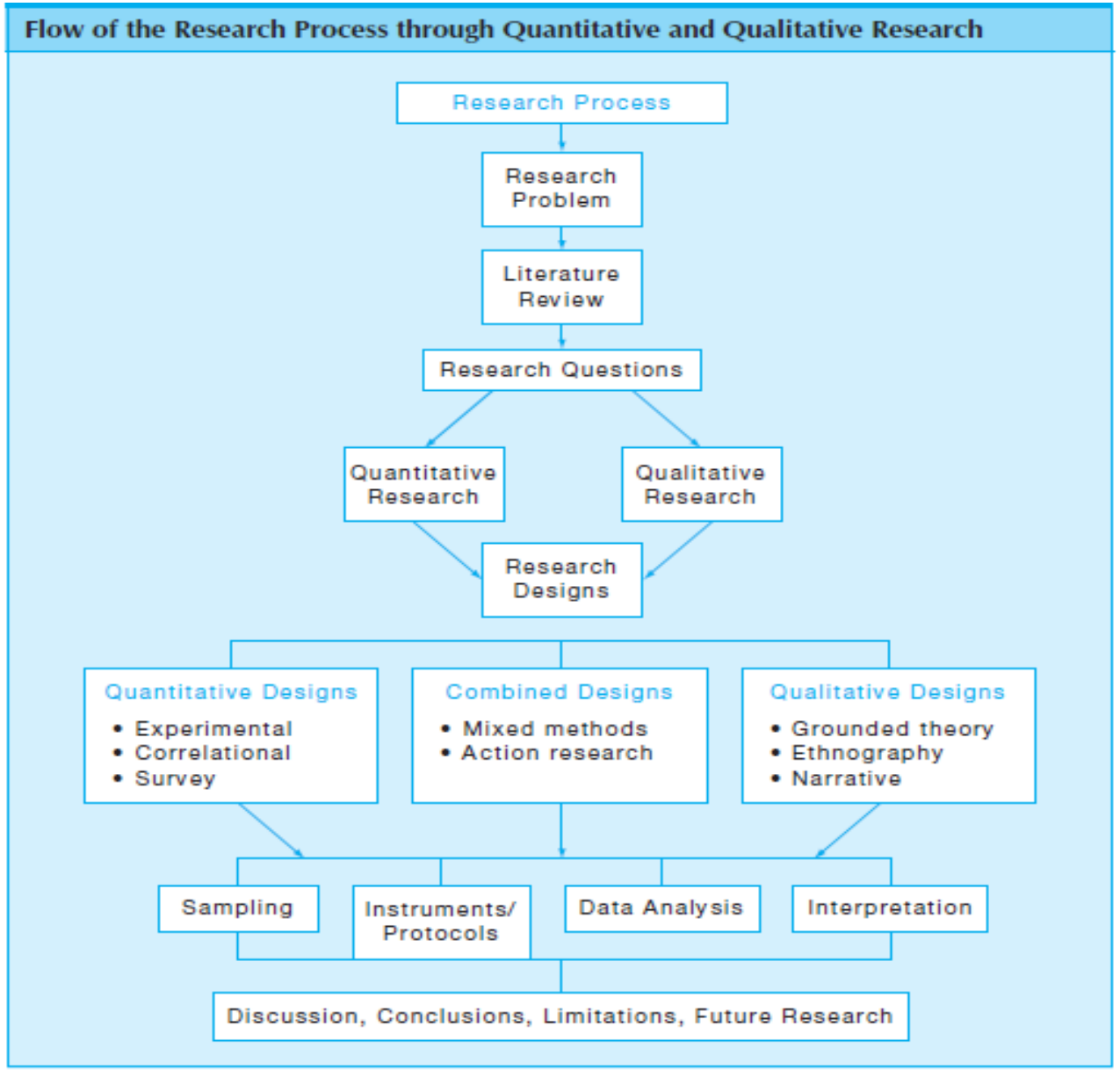
- Research Process
- Criteria of Good Research

Research Process

1. Defining the research problem (research problem, objectives and questions)
2. Extensive literature survey (Literature Review)
3. Developing the hypotheses
4. Preparing the research design
5. Determining sample design
6. Collecting the data
7. Execution of the project
8. Analysis of data
9. Hypothesis testing
10. Generalizations and interpretation
11. Preparation of the report or presentation of the results, i.e., Formal write-up of conclusions reached.

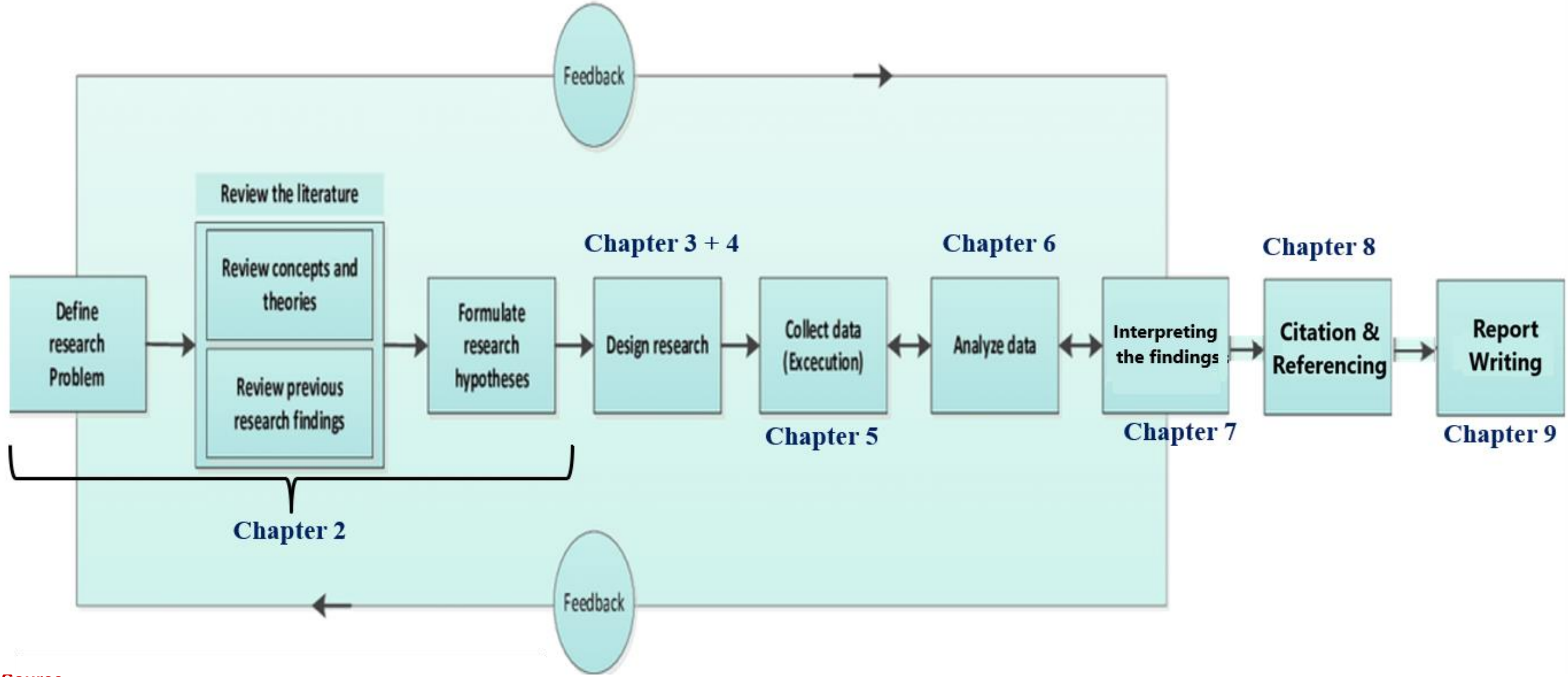
The research Process Flow chart

- **Several Charts of the research process steps are available.**
- **They all share the same core steps with minor differentiations**



Source:
Creswell, J. W., GUETTERMAN, T. C. (2018). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. pearson.

The research Process Flow Chart (Adopted for this course)



Source:
Kothari, C. R. (2009). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. pearson.

Next….

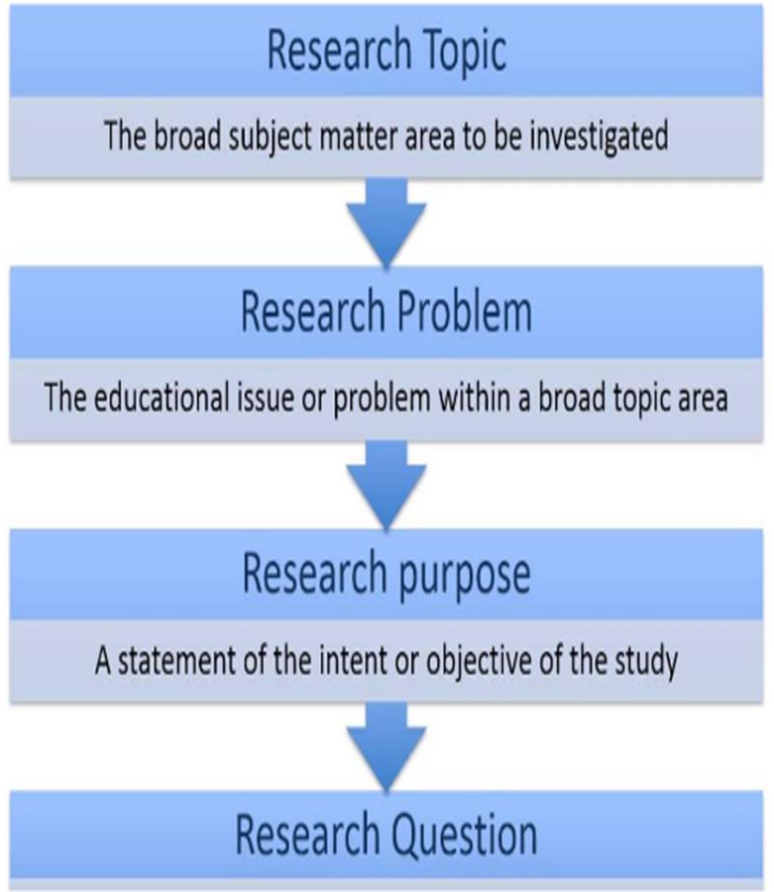
- Throughout this course, each of these steps is explained.
- The current Unit (Unit 2) explains the first three steps (Problem definition, Literature Review, and Hypothesis development) in detail.
- It briefly describes the remaining steps

The flow from Research Topic to Research Questions

After the researcher identifies the research topic he is interested in as explained in Unit 1 the research process starts.

The flowchart illustrates the relationship between the **Research Topic**, **Research Problem**, **Research Purpose**, and **Research Question**.

- The Research topic is the field of the study. It represents the broader scope.
- Examples of research topics:
    - Machine learning in education.
    - Social media advertising
    - Health care applications.



Each of these represents a wide area of investigation. For instance, in social media advertising, researchers may study different fields such as :

- **Measuring the performance of social media advertising campaigns**
- **Collaborating with social media influencers to promote products or services.**
- **Navigating the ethical and legal considerations in social media advertising.**

The flow from Research Topic to Research Questions

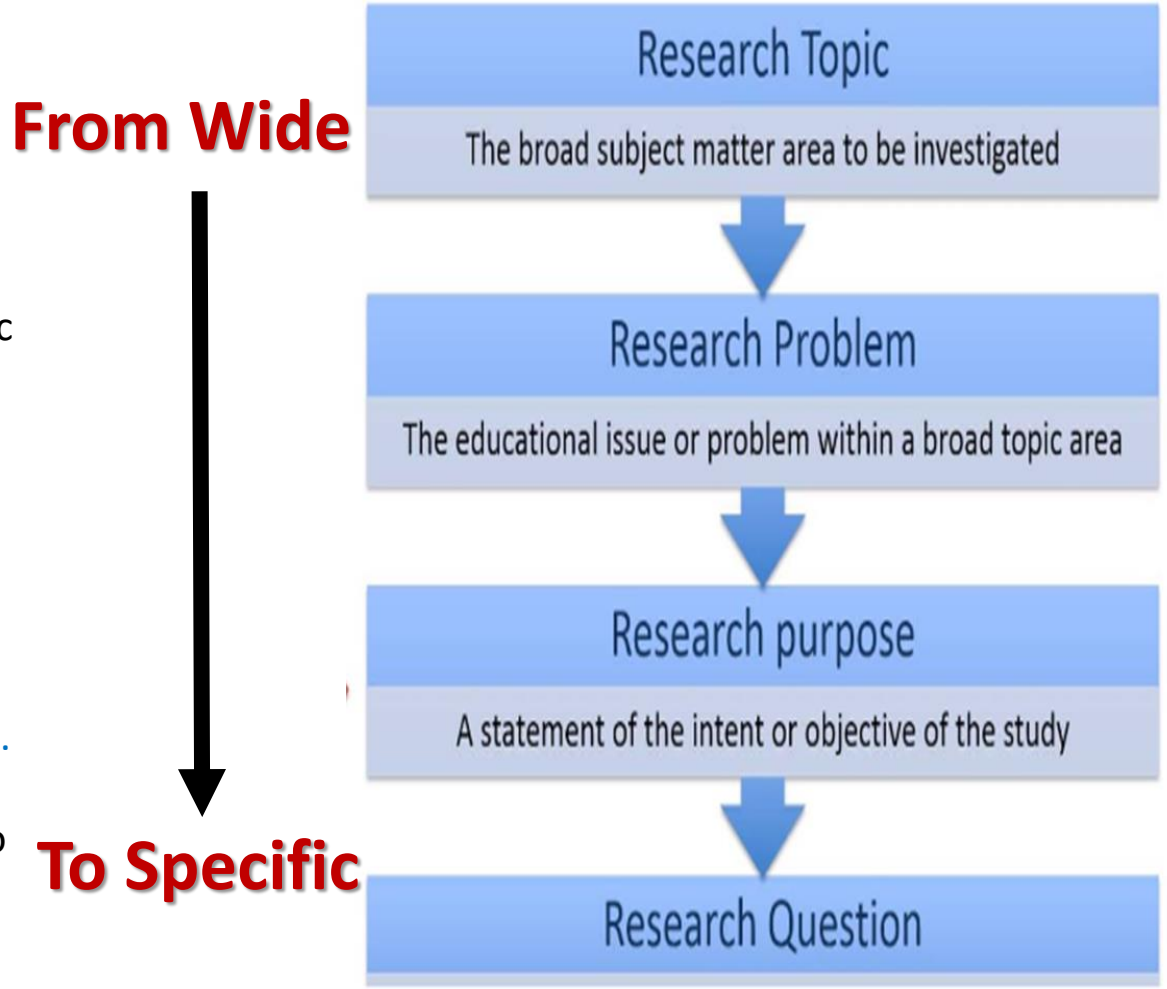
Here's an example to briefly explain how they are all related:

1.Research Topic:
1. **Definition**: The broad subject matter area to be investigated.
2. **Example**: Artificial Intelligence in Healthcare.

2.Research Problem:
1. **Definition**: The specific issue or problem within the broad topic area that the research will address.
2. **Example**: The challenge of accurately diagnosing rare diseases using artificial intelligence.

3.Research Purpose:
1. **Definition**: A statement of the intent or objective of the study. It outlines what the research aims to achieve.
2. **Example**: To develop a machine learning model that can improve the accuracy of diagnosing rare diseases in healthcare.

4.Research Questions:
1. **Definition**: The specific questions that the research will seek to answer. These questions guide the research process.
2. **Example**: How effective is the proposed machine learning model in diagnosing rare diseases compared to existing diagnostic methods?



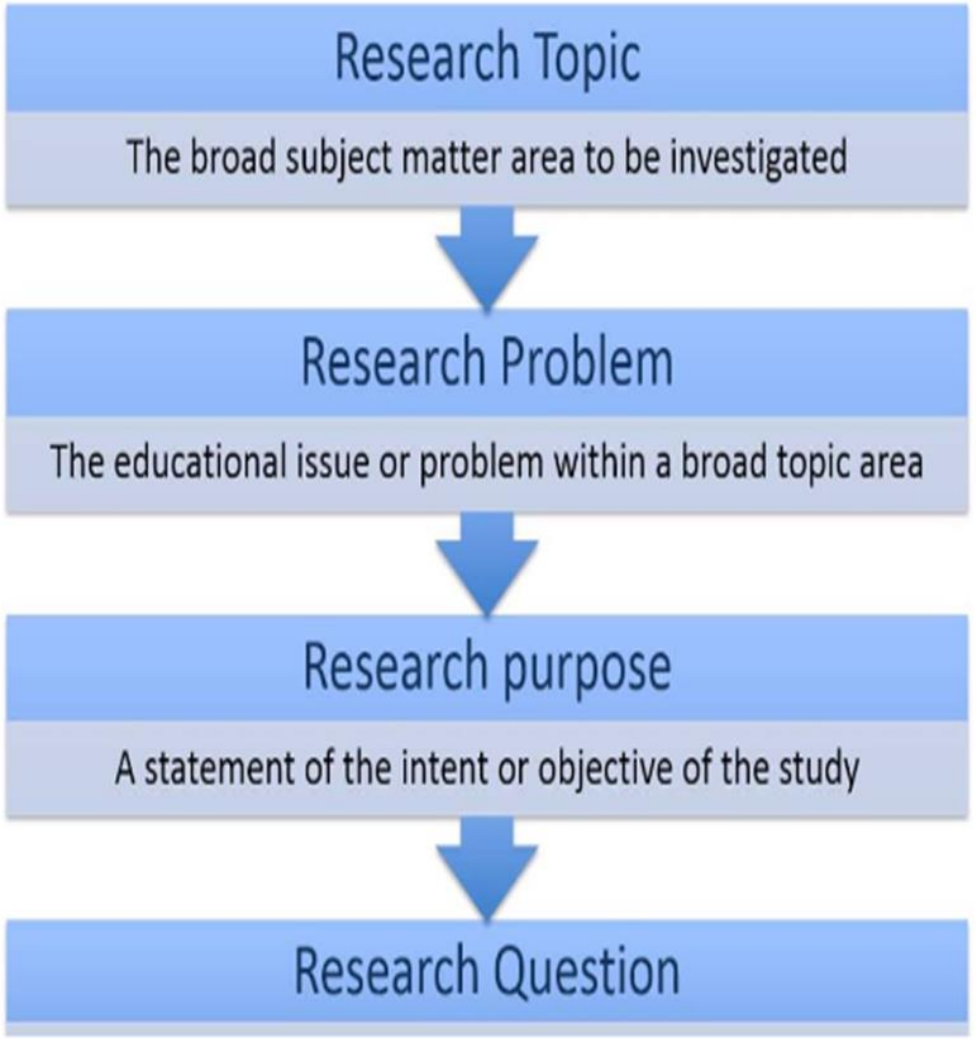
The flow from Research Topic to Research Questions

Artificial Intelligence in Healthcare.

The challenge of accurately diagnosing rare diseases using artificial intelligence.

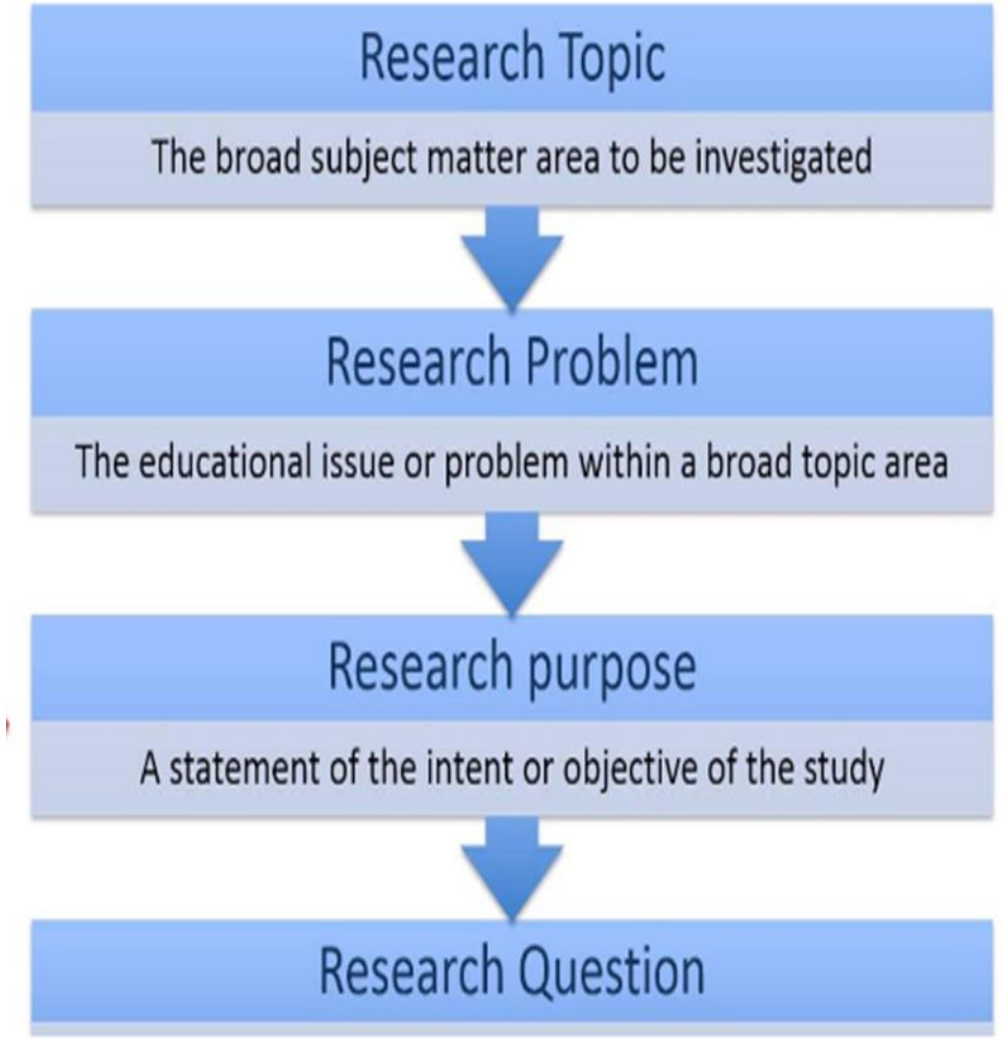
To develop a machine learning model that can improve the accuracy of diagnosing rare diseases in healthcare.

How effective is the proposed machine learning model in diagnosing rare diseases compared to existing diagnostic methods?



The flow from Research Topic to Research Questions

Next, the Research problem, research objectives, and research questions are explained in detail



1. Defining the Research Problem

What is a research problem?

A research problem is a specific question, problem, or difficulty that needs to be investigated or analyzed.

Examples:

- Why turnover rate is increasing these days (لماذا نسبة ترك العمل بين الموظفين تتزايد؟)
- Students use of the university portal is limited (استخدام الطلاب لبورتال الجامعة محدود)
- Marketing the new product is not possible!

So:

It is a concise statement that expresses the difference between what is currently known and what needs to be known or the difference between a current situation and a desired state.

Defining the Research Problem…. Cont.

- Examining research problems helps to identify the key concepts and terms of research.
- A research problem should be **clear, measurable, concise** مختصر **, and specific enough** to guide the process and contribute to the definition of research project objectives, methods, and outcomes.
- It is the foundation of any research project, and a well-formulated research problem is required for any research study to be successful.

Defining the Research Problem.. Cont.

A problem well defined is a problem half solved!

Failure to properly identify where you are headed and why will inevitably lead you to wonder where you are and how you got there! Wren

Research Problem examples and corrections

Example 1:

Incorrect: What are the effects of social media on people?

In the above example, the first research question is not specific enough to capture accurate feedback. Nobody knows what social media you're talking about and what 'people' you're referring to.

Correct: What effect does use Facebook every day have on teenagers?

Here, the research problem is more specific

Example 2:

Incorrect: Is advertising copy A better than advertising copy B?

Correct: Which advertising copy has a higher day-after recall score?

ما هي النسخة الإعلانية التي حصلت على درجة استدعاء أعلى في اليوم التالي؟

Research Problem

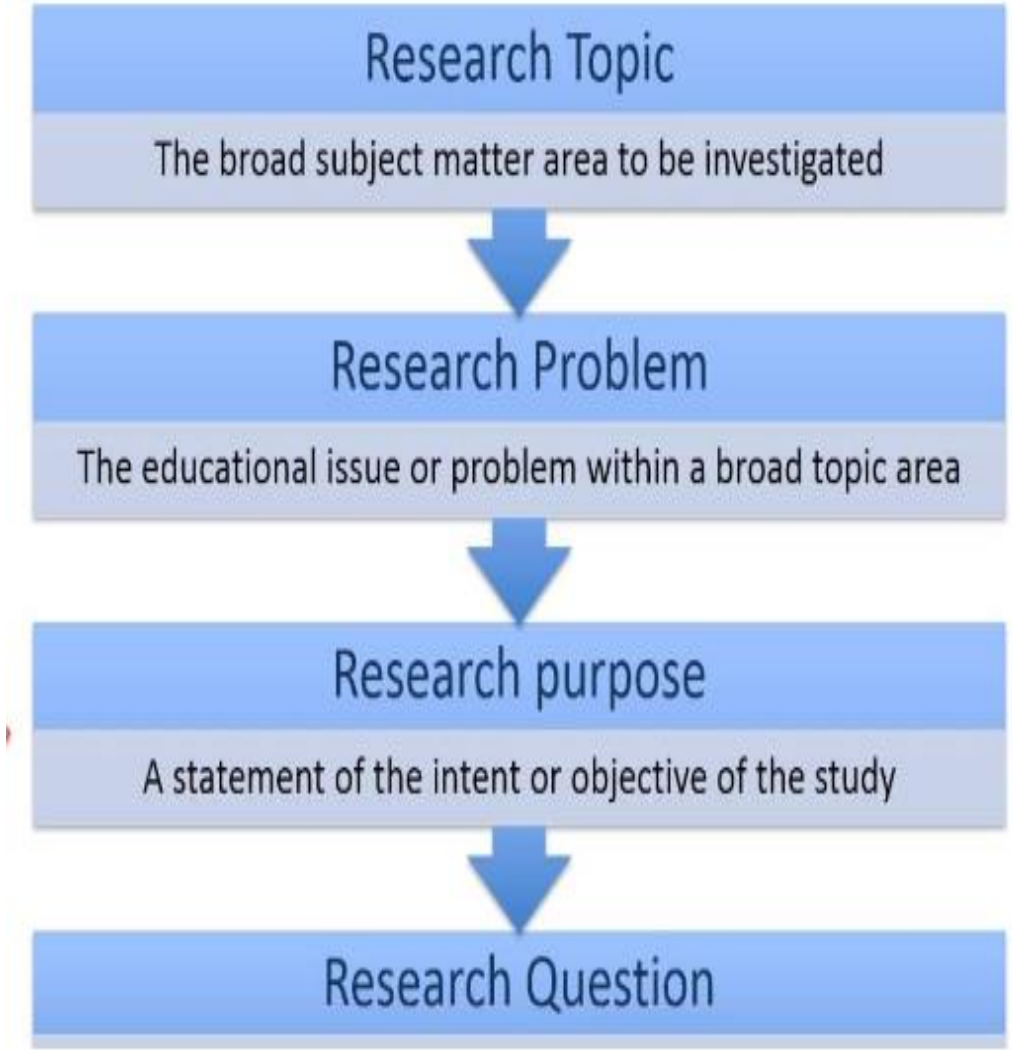
A **good research problem** should have the following characteristics:

1. It should **address a gap** in knowledge.
2. It should be **significant enough** to contribute to the existing body of research
3. It should **lead to further research**
4. The problem should **render itself to investigation through collection of data**
5. It should be of **interest to the researcher** and suit his/her **skills, time, and resources**
6. The approach towards solving the problem should be **ethical**

1.1. Defining the Research Objectives

- After defining the research problem, the research objectives (purposes) and the research questions are formulated.
- **Research objectives** are more specific than your research problem and indicate the particular focus and approach of your project.
- Though you will only have one research problem, you will likely have several research objectives.

The path from Research Topic To Research Questions



From Research Problem To Research Questions

Example:

Research Problem

To examine contributory factors to muscle retention in a group of elderly people

دراسة العوامل المساهمة في احتباس العضلات في مجموعة من كبار السن

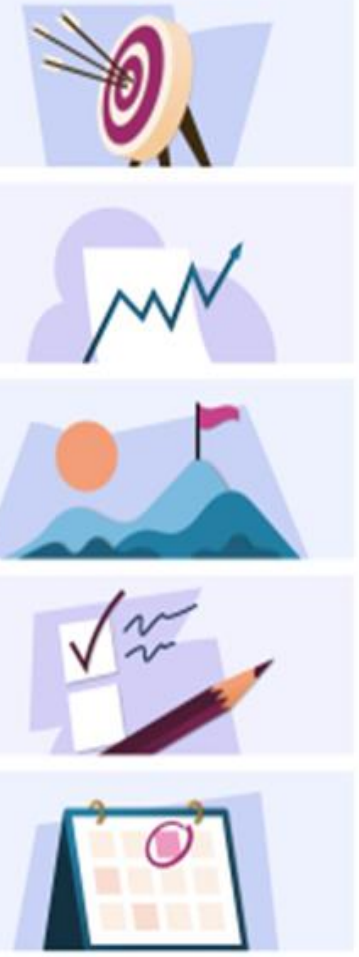
Research objectives

- To determine the effect of physical activity on the participants' muscular health.
- To determine the impact of dietary factors, particularly protein consumption, on the muscular health of the participants.

Characteristics of Good Research Objectives

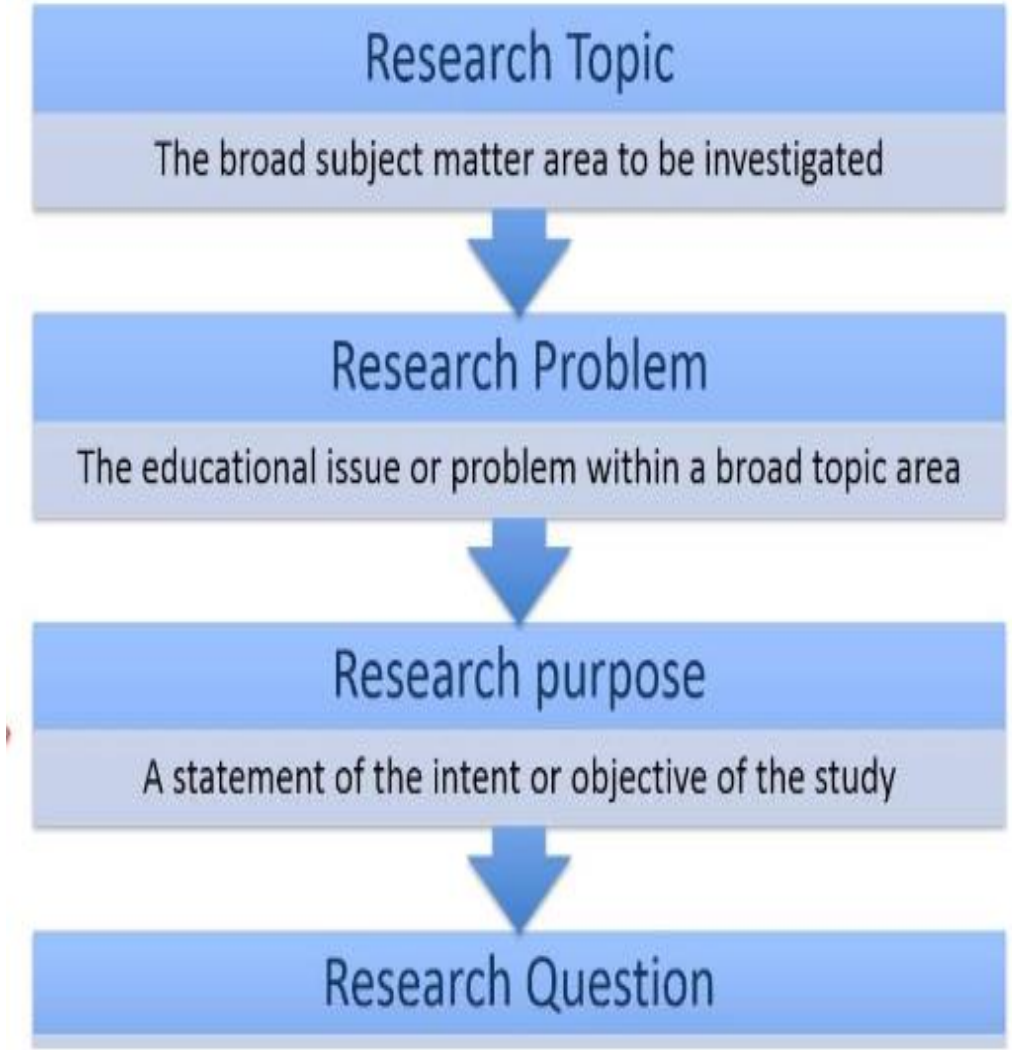
- **Good Objectives should be Specific, Measurable, Achievable, Relevant, and Time-based.**

S	Specific	Make your goal specific and narrow for more effective planning
M	Measurable	Make sure your goal and progress are measurable
A	Achievable	Make sure you can reasonably accomplish your goal within a certain time frame
R	Relevant	Your goal should align with your values and long-term objectives
T	Time-based	Set a realistic but ambitious end date to clarify task prioritization and increase motivation



1.2. Defining the Research Questions

- Research questions are more specific. They state the specific issue or problem that your assignment will focus on.
- It also outlines the task that you will need to complete.
- They are correspondent with research objectives



Characteristics of Good Research Questions

Research Problem

A *research problem* can relate to a specific topic or opportunity

Research Question

A *research question* provides an area to focus on regarding a research problem.

For Example

Climate Change

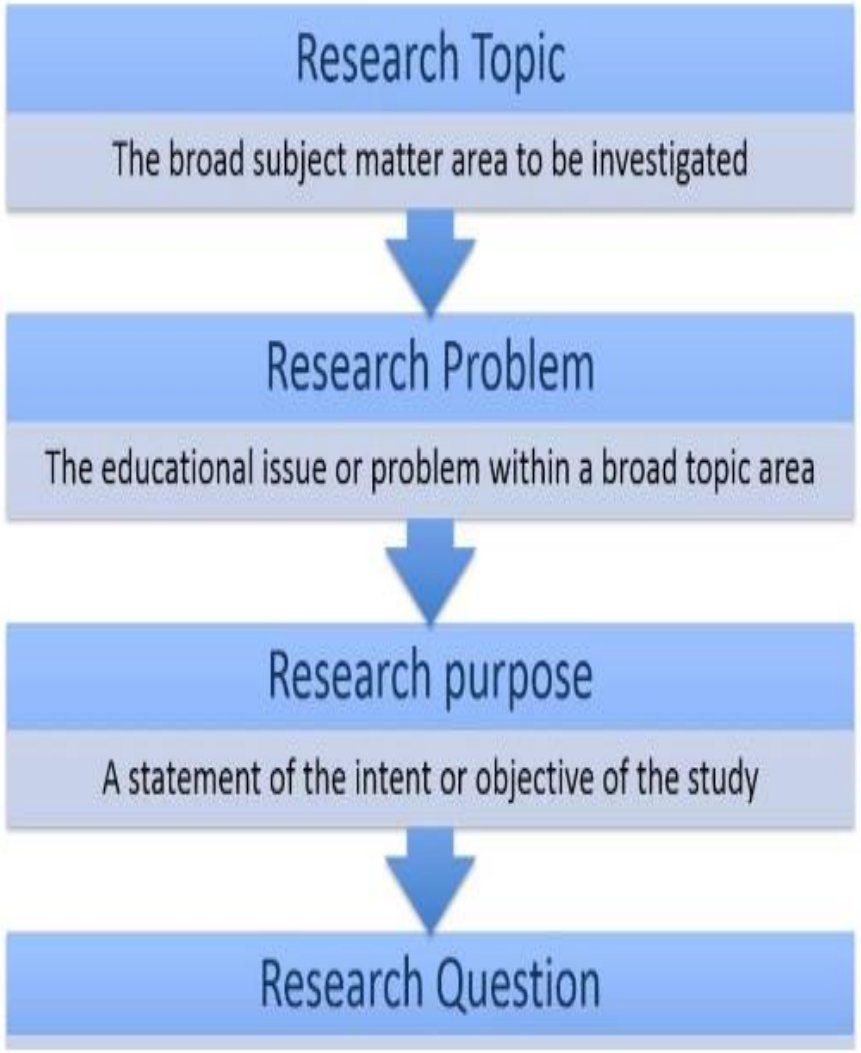
How do ozone levels impact global temperature levels?

From Research Problem To Research Questions

Examples of research questions:

•Has there been an increase in homelessness in San Francisco in the past ten years? ☒ •How have economic, political, and social factors affected patterns of homelessness in San Francisco over the past ten years? ☑	The first question is too simple: it can be answered with a simple yes or no. The second question is more complex, requiring in-depth investigation and the development of an original argument.
•What effect does social media have on your mind? ☒ •What effect does daily use of Twitter have on the attention span of 12- to 16-year-olds? ☑	The first question is not specific enough. The second question is more researchable, using qualitative and quantitative data collection.

A comprehensive Example from Research Topic to Research Questions



Research Topic: Differential tuition at Arizona State University

Research Problem:

Differential tuition limit students' access to and persistence/graduation from those (affluent) majors with higher tuition

Research Purpose:

- To evaluate the impact of differential tuition on students' choice of major as well as retention and graduation (Quant)
- To understand how ASU students perceive differential tuition in pursuing those majors (Qual)

Research Questions:

- What is the impact of differential tuition on ASU students' choice of major as well as retention and graduation?
- How do students perceive differential tuition?
- What is the role of differential tuition during the process that students make choices about academic major?

Understanding Research Objectives and Research Questions

Translated: https://www.youtube.com/watch?v=JilzYGir3Ys

2. Literature Review

Literature Review is the second step of the research process.

What is Literature Review مراجعة الادبيات ?

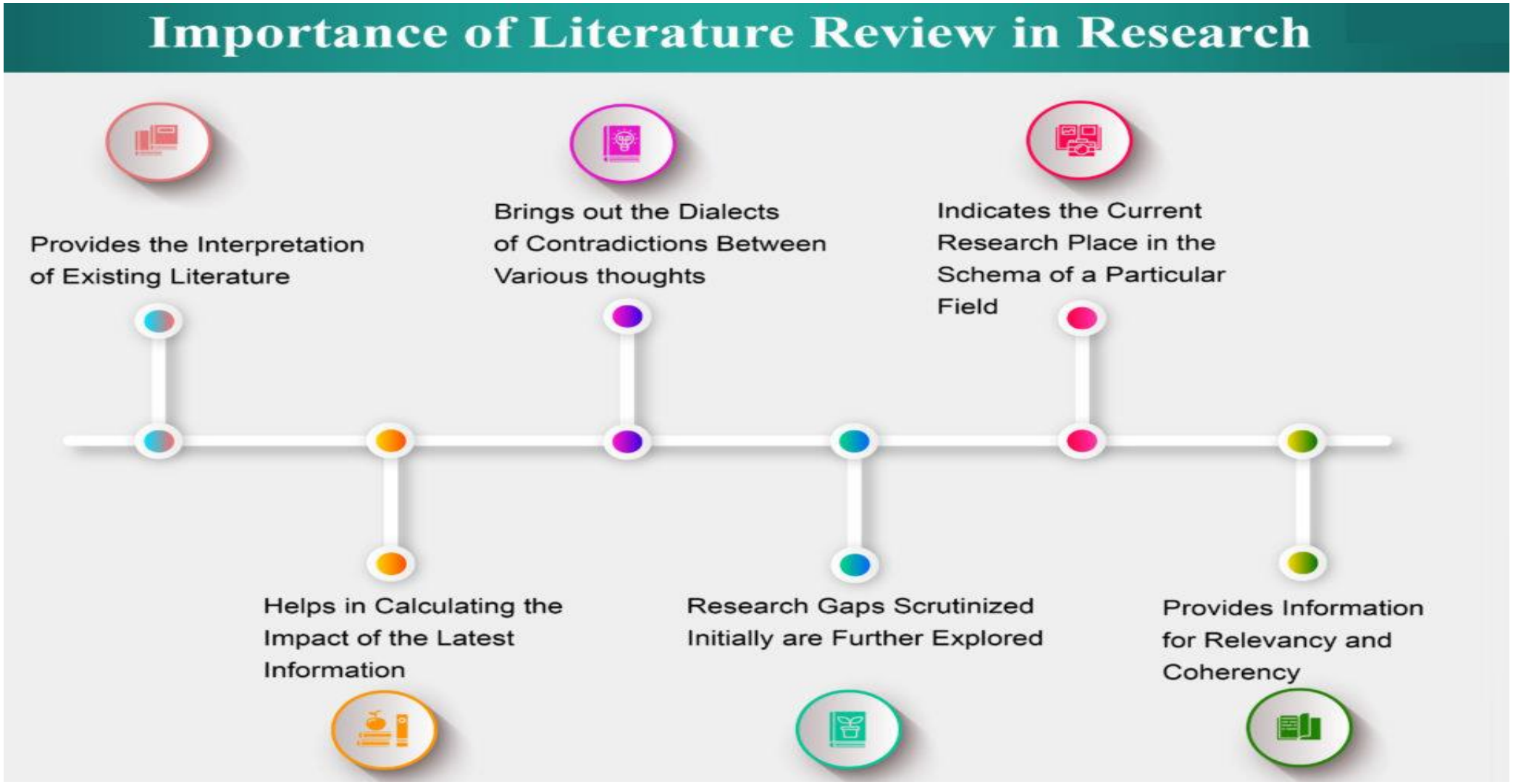
A literature review is a survey of scholarly sources on a specific topic. It provides an overview of current knowledge, allowing you to identify relevant theories, methods, and gaps in the existing research that you can later apply to your research

There are four key steps to writing a literature review:

1. Search for relevant literature
2. Evaluate sources
3. Identify debates and gaps
4. Write your literature review

A good literature review doesn't just summarize sources—it analyzes, synthesizes, and critically evaluates to give a clear picture of the state of knowledge on the subject.

Why is it important to do a literature review in research?



2. Literature Review

What are the Literature Review sources?

1. Peer reviewed journal articles.
2. Edited academic books.
3. Articles in professional journals.
4. Statistical data from government websites.
5. Website material from professional associations (use sparingly and carefully).

3. Developing the Hypotheses صياغة الفرضيات

What does a hypothesis mean?

Hypothesis: An idea or explanation for something that is based on known facts but has not yet been proved.
It's a statement that provides an **explanation** for why or how something works, based on **facts** (or some reasonable assumptions), but that has **not yet been specifically tested**.
For example, a hypothesis might look something like this:

Hypothesis: sleep impacts academic performance.

This statement **predicts** that academic performance will be influenced by the amount and/or quality of sleep a student engages in – sounds reasonable, right? It's based on **reasonable assumptions**, underpinned by what we currently know about sleep and health (from the existing literature). So, loosely speaking, we could call it a hypothesis, at least by the dictionary definition.
But that's not good enough…
In the world of academic research, a statement needs a few more criteria to constitute a true **research hypothesis**.

3. Developing the Hypotheses صياغة الفرضيات

What does a research (scientific) hypothesis mean?

A research hypothesis (also called a scientific hypothesis) is a statement that introduces a research question and proposes an expected result.

- Research hypotheses are formulated **after** the literature review
- The researcher then tries to verify the validity of his hypotheses using appropriate materials and methods in which he puts all his experiences and the information he has collected to reach correct research solutions, which may or may not be compatible with the hypotheses.

3. Developing the Hypotheses صياغة الفرضيات

Examples: (The language of hypotheses always discusses **variables**)

In Natural language:

- If I water plants daily they will grow faster.
- Facebook usage is not good for university students'.
- When turning off your phone, it recharges faster.

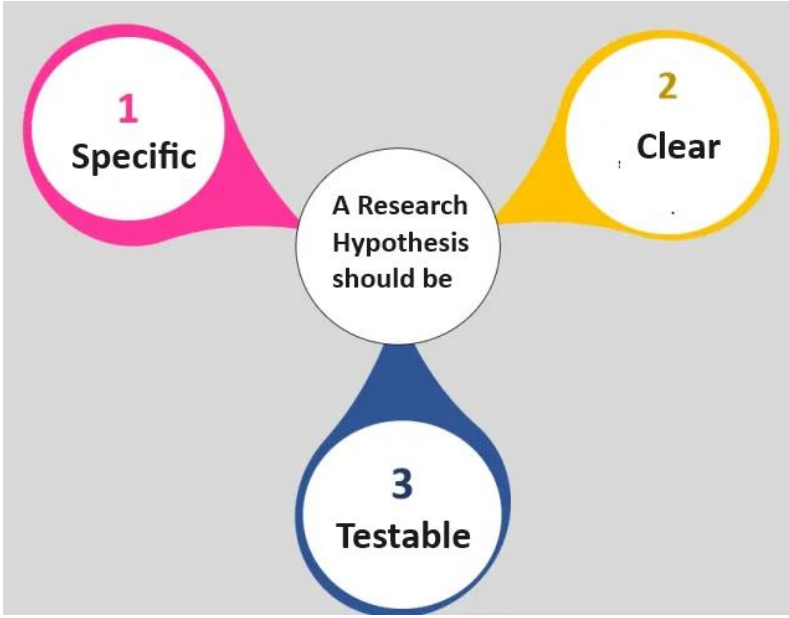
In scientific syntax:

- Daily irrigation has a positive impact on plants growth.
- Facebook negatively affects university students' willingness to study the courses required of them.
- Turning off your phone negatively influence time needed to recharge.

3. Developing the Hypotheses صياغة الفرضيات

To constitute a quality hypothesis, the statement needs to have three attributes:

- **Specificity:** the statement is very specific as it identifies the variables involved
- **Clarity**: A good research hypothesis needs to be extremely clear and articulate about both what's being assessed (who or what variables are involved) and the expected outcome (for example, a difference between groups, a relationship between variables, etc.).
- **Testability**: A statement must be testable to qualify as a research hypothesis. In other words, there needs to be a way to prove (or disprove) the statement. If it's not testable, it's not a hypothesis – simple as that.



3. Developing the Hypotheses صياغة الفرضيات

- Not all research in computer science necessarily involves research hypotheses.
- The nature of research in computer science can vary widely, and the approach taken often depends on the specific subfield, the research questions, and the goals of the study.
    - **Theoretical Research**: Focuses on proving theories or algorithms, often without explicit hypotheses.
    - **Experimental Research**: Typically involves hypotheses to test performance or effectiveness.
    - **Empirical Research**: Usually involves hypotheses when analyzing collected data.
    - **Engineering and Design Research**: Emphasizes building systems; hypotheses may not be necessary.
    - **Descriptive Research**: Focuses on observation and description, often without formal hypotheses.

How does one go about developing working hypotheses?

a) Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;

b) Examination of data and records, if available, concerning the problem for possible trends,

c) Review of similar studies in the area or of the studies on similar problems;

d) Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Hypotheses writing in Research

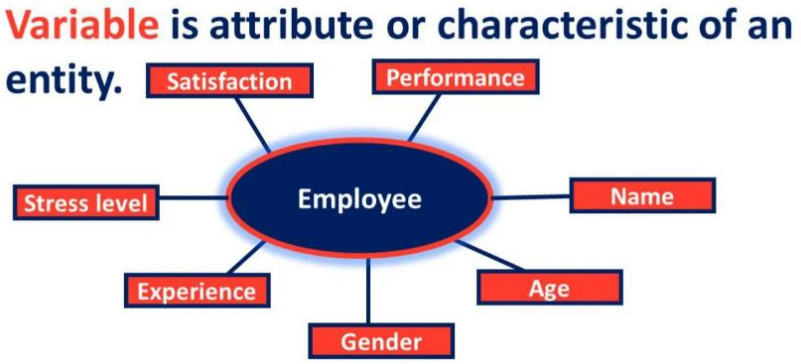
For formulating the hypotheses, **first, we have to understand the term "variables"** in research and their role in the study.

What is a variable?

A variable is any kind of attribute or characteristic that you are trying to measure, manipulate and control in research.

All studies analyze a variable, which can describe a person, place, thing or idea. A variable's value can change between groups or over time.

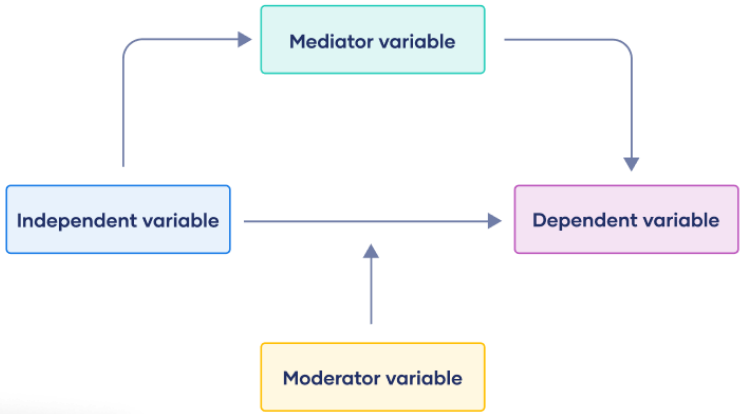
Example: Employee is also an entity. Name, Age, Gender, Experience, Stress level, satisfaction, Performance are the attributes of an employee



Hypotheses writing in Research

Variables are classified into the following most popular types:

- **Independent** variable (المتغير المستقل) -- وهو المتغير المؤثر
- **Dependent** variable (المتغير التابع) وهو المتغير المتأثر
- **Moderating** variable (المتغير المساعد)
- **Mediating** variable (المتغير الوسيط)
- **Control** variable (المتغير المتحكم)



Types of variable in Research

1. **Independent Variable:**

- Is the one that influences the dependent variable in a certain (positive or negative, linear or non-linear) way.
- It is the condition that you change in an experiment.
- It is the variable you control.

Example 1:

(Time Spent Studying) influences (Test Score).

We see that "Time Spent Studying" must be the independent variable

Example 2:

Enjoyment while using a mobile App may encourage the user to its continuance usage of the App.

Then, enjoyment is the independent variable

Types of variable in Research

2. Dependent Variable (هو محور الدراسة الذي نريد فهمه)

- Is the variable of primary interest to the researcher.
- The researcher's goal is to understand and describe the dependent variable, or to explain its variability, or predict it.
- In other words, it is the main variable that lends itself for investigation.

Example 1:

(Time Spent Studying) influences (Test Score).

"Test Score" is the dependent variable

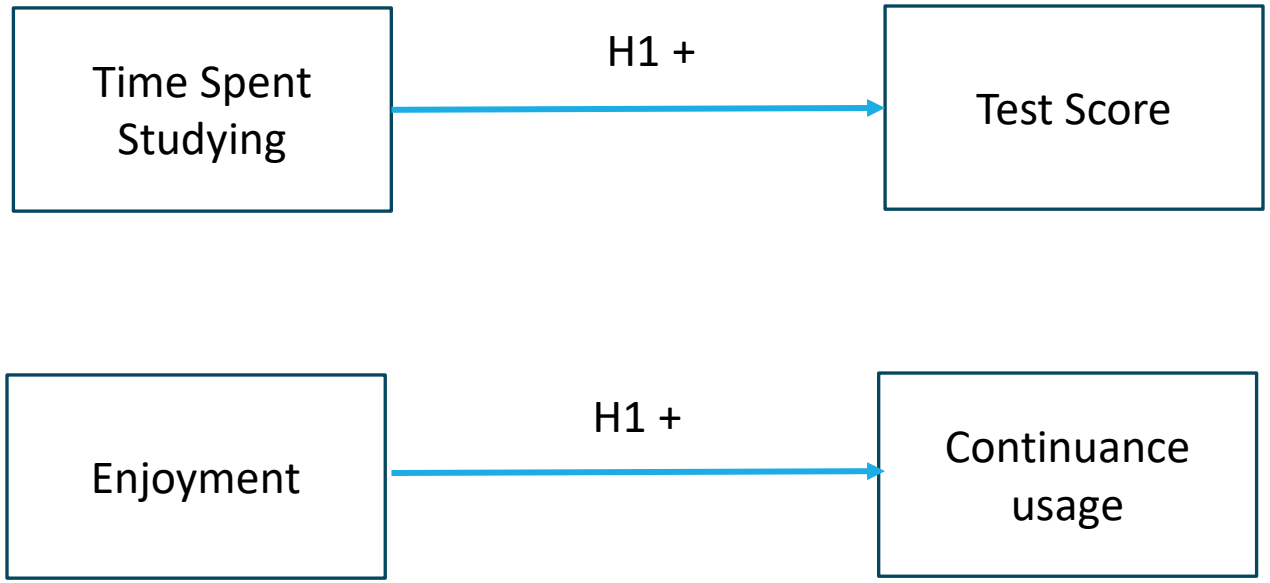
Example 2:

Enjoyment while using a mobile App may encourage the user to its continuance usage of the App.

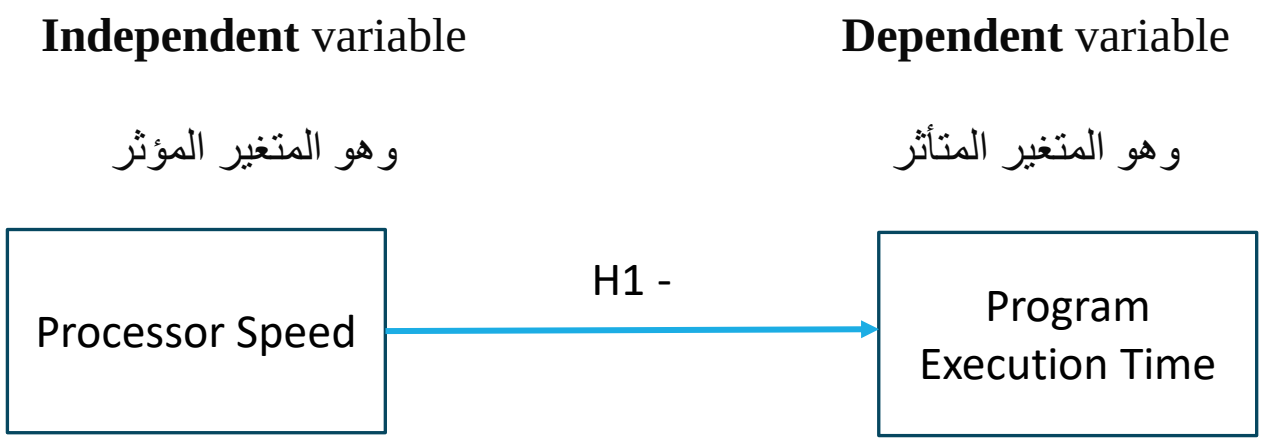
Then, continuance usage is the dependent variable

Types of variable in Research

Usually, these relationships are graphically represented as follows:



Example 3 : Processor Speed and Program Execution Time: In a study examining the effect of processor speed on the time it takes to execute a program, the independent variable is the processor speed (measured in GHz), while the dependent variable is the program execution time (measured in seconds). The hypothesis might suggest that as the processor speed increases, the program execution time decreases.



Types of variable in Research

3. Moderating Variable:

- The moderating variable is one that has a contingent effect on the independent variable–dependent variable relationship.
- That is, the presence of a third variable (the moderating variable) modifies the original relationship between the independent and the dependent variables.
- It either affects the **strength** or **direction** of the relationship between the independent and dependent variable

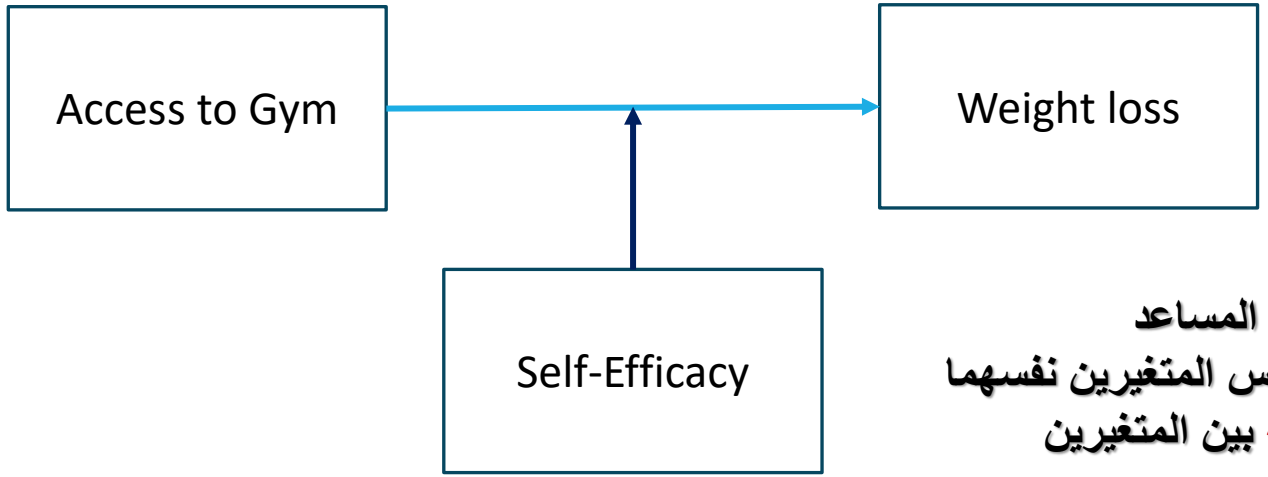
Types of variable in Research

3. Moderating Variable:

Example 1 (influencing the strength of relationship):

Self-Efficacy (الكفاءة الذاتية) in Exercise

The relationship between access to a gym (independent variable) and weight loss (dependent variable) might be **stronger** for individuals with high self-efficacy compared to those with low self- efficacy. Here, self-efficacy is a moderating variable.



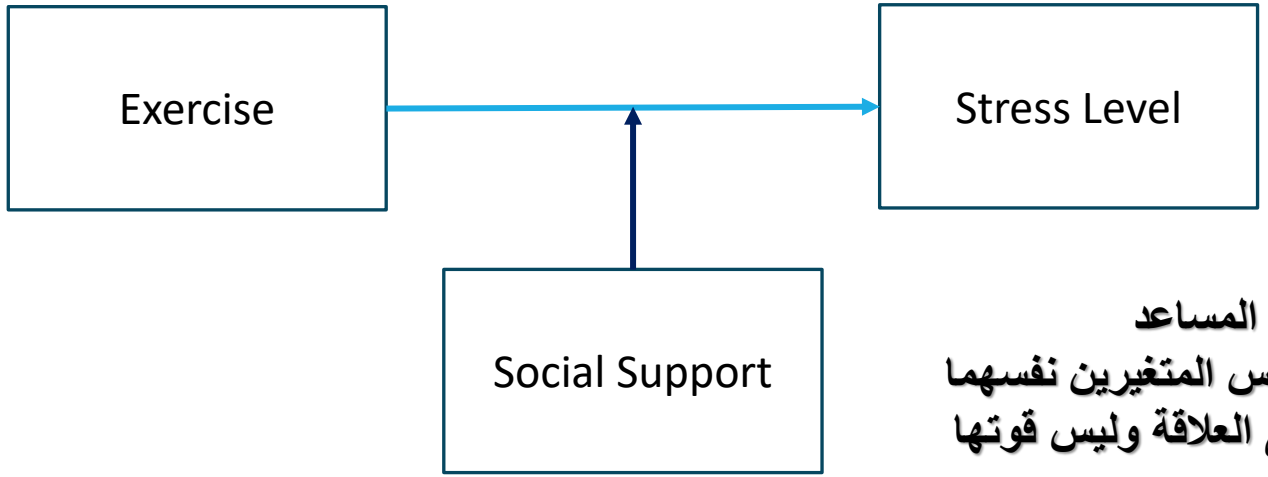
نلاحظ هنا ان تأثير العامل المساعد
يكون على العلاقة بين المتغيرين وليس المتغيرين نفسهما
ووجوده قد يزيد من قوة العلاقة بين المتغيرين

Types of variable in Research

3. Moderating Variable:

Example 2 (influencing the direction of relationship):

The presence or absence of social support from friends alters the direction of the relationship between exercise and stress levels among college students.



نلاحظ هنا ان تأثير العامل المساعد
يكون على العلاقة بين المتغيرين وليس المتغيرين نفسهما
ووجوده قد يؤثر إيجابيا او سلبيا على العلاقة وليس قوتها

Types of variable in Research

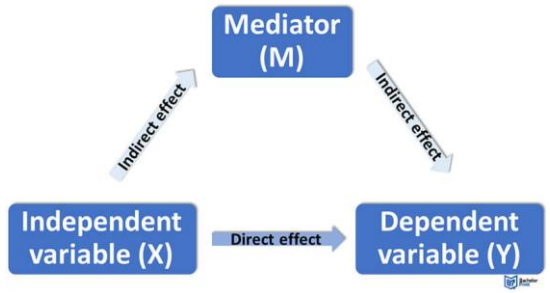
3. Moderating Variable:

Example 3: In a study examining the relationship between computer-based training programs and job performance, computer experience could moderate this relationship. Individuals with more computer experience might benefit more from the training programs compared to those with less experience.

Example 4: In a study investigating the impact of website loading times on user satisfaction, internet connection speed could moderate this relationship. Users with faster internet connections may be less sensitive to website loading times compared to users with slower connections.

Example 5: consider the relationship between sun exposure and skin damage. The primary relationship is more sun exposure leads to increased skin damage. A moderator would be the use of sunscreen.

Types of variable in Research



4. Mediating Variable:

A **mediator** is a way in which an independent variable impacts a dependent variable. It's part of the causal pathway of an effect, and it tells you how or why an effect takes place.

If something is a mediator:

1.It's caused by the independent variable.

2.It influences the dependent variable

3.When it's taken into account, the statistical correlation between the independent and dependent variables is higher than when it isn't considered.

Types of variable in Research

4. Mediating Variable

1.Example 1:

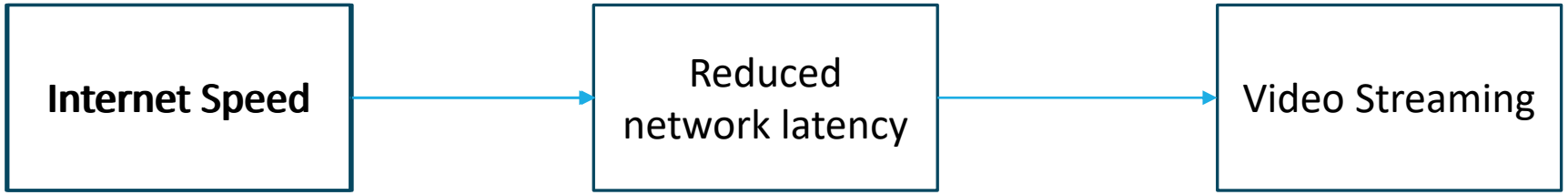
Exercise and Mood: Suppose you're studying the relationship between exercise and mood. The independent variable is exercise frequency (e.g., number of times exercised per week), and the dependent variable is mood (e.g., measured using a mood scale). A potential mediator could be endorphin release. The hypothesis might suggest that exercise frequency leads to an increase in endorphin release, which in turn improves mood.

Types of variable in Research

4. Mediating Variable

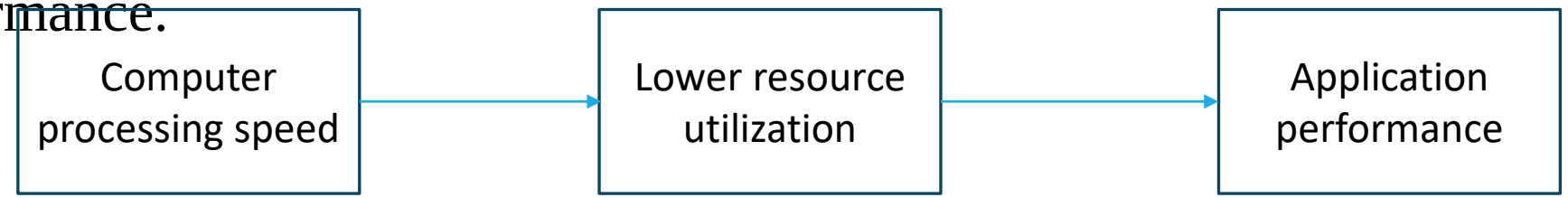
Example 2:

Internet Speed and Video Streaming Quality: Faster internet speed (in Mbps) → Reduced network latency → Better video streaming quality.



Example 3:

Computer Processing Speed and Application Performance: Faster processing speed (e.g., CPU clock speed) → Lower resource utilization → Faster application performance.



Types of variable in Research

5. Control Variable

A control variable is anything that is held constant or limited in a research study. It's a variable that is not of interest to the study objectives', but is controlled because it could influence the outcomes.

Variables may be controlled directly by holding them constant throughout a study (e.g., by controlling the room temperature in an experiment), or they may be controlled indirectly through methods like randomization or statistical control (e.g., to account for participant characteristics like age in statistical tests).

Types of variable in Research

5. Control Variable

1.Example 2 (Control variables in observational study):

You want to investigate whether there's a relationship between the variables of income and happiness. You hypothesize that income level predicts happiness, but it's not practically possible to manipulate the variable of income. Instead, you use a survey with Likert scale questions to collect data about income and happiness.

To account for other factors that are likely to influence the results, you also measure these control variables:

- Age
- Marital status
- Health

How to write a hypothesis in 5 steps

1. Ask a question. ...

2. Conduct preliminary research. ...

3. Define your variables. ...

4. Phrase it as an if-then statement. ...

5. Refine and rewrite using variables' names

These five steps will be applied in the following example

Developing a hypothesis (with example)

Step 1. Ask a question

Writing a hypothesis begins with a research question that you want to answer. The question should be focused, specific, and researchable within the constraints of your project.

Example: (Research question)

Do students who attend more lectures get better exam results?

هل الطلبة الذين يحضرون المحاضرات بانتظام يحصلون على نتائج افضل؟

Step 2. Do some preliminary research

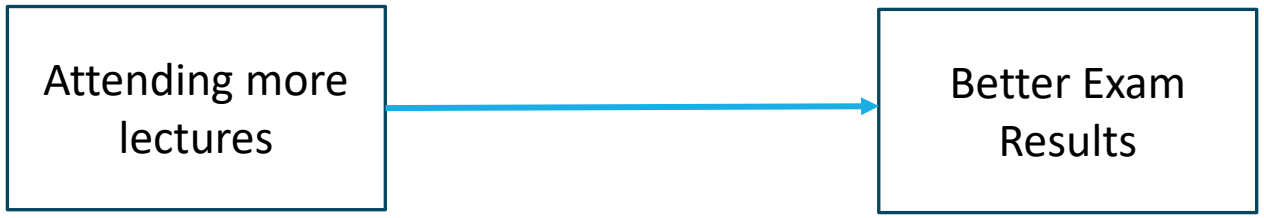
- Your initial answer to the question should be based on what is already known about the topic.
- Look for **theories and previous studies** to help you form educated assumptions about what your research will find. (**Literature Review**)
- At this stage, you might construct a conceptual framework to ensure that you're embarking on a relevant topic.
- This can also help you identify which variables you will study and what you think the relationships are between them.

Step 3. Formulate your hypothesis

Now you should have some idea of what you expect to find based on what you read in literature. Write your initial answer to the question in a clear, concise sentence.

Example: Formulating your hypothesis

Attending more lectures leads to better exam results.



4. Phrase your hypothesis

To identify the variables, you can write a simple prediction in if…then form. The first part of the sentence states the independent variable and the second part states the dependent variable.

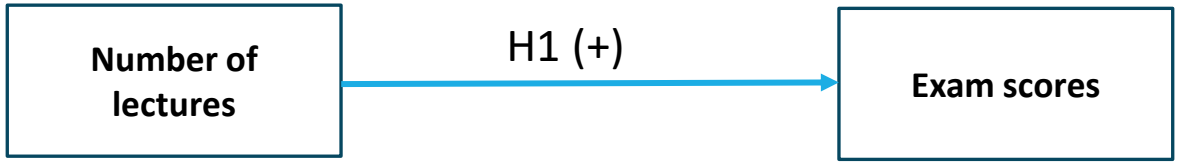
If students attend more lectures, then their exam scores will improve.

5. Refine your hypothesis

In academic research, hypotheses are more commonly phrased in terms of **correlations** or effects, where you directly state the predicted relationship between variables (**Directional**)

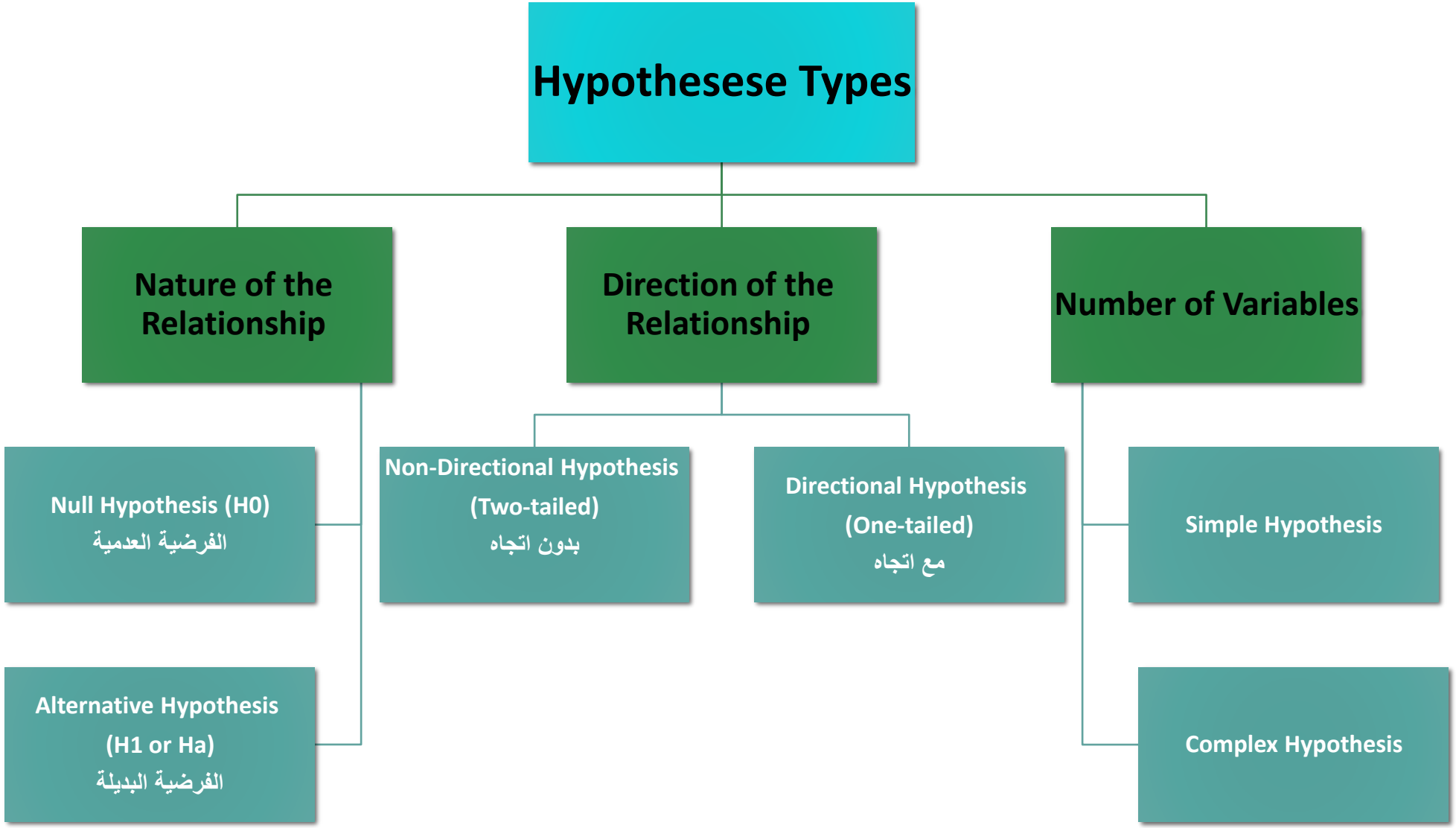
(Directional means : **positive** or **negative**)

The number of lectures attended by the students has a positive effect on their exam scores.



- On the arrow, we write H (for Hypothesis), followed by a number that indicates the hypothesis number. It is preferable to use parentheses with a + or – to indicate the direction (which is positive correlation in this case)
- The previous hypothesis is read as :
  - The number of lectures positively influence exams scores, **or**
  - As the number of lecture increases, the exam scores increases, **or**
  - The number of lectures has a positive impact on exams scores

Types of scientific research hypotheses



Types of scientific research hypotheses

1. Based on the nature of relationships:

Null hypothesis: A hypothesis that is to be tested.

Alternative hypothesis: The alternative to the null hypothesis.

Null Hypothesis الفرضية الصفرية (تفترض عدم وجود التأثير في العلاقة)	Alternative hypothesis الفرضية البديلة (تفترض وجود التأثير في العلاقة)
There is **no** relationship between academic achievement and Facebook.	There **is an apparent** relationship between addiction to smoking and the resulting diseases of the lungs and heart
There is **no** relationship between a person's gender and his educational attainment.	There **is an impact** on the child's upbringing and the academic achievement he achieves
There are **no** significant relationships between intelligence and height.	There **is a relationship** between the level of poverty and the occurrence of the phenomenon of divorce in the Syrian Arab Republic

Writing a null hypothesis

If your research involves statistical hypothesis testing, you will also have to write a null hypothesis. The null hypothesis is the default position that there is no association between the variables. The null hypothesis is written as H_0, while the alternative hypothesis is H_1 or Ha.

- H_0: The number of lectures attended by first-year students has **no effect** on their final exam scores.
- H_1: The number of lectures attended by first-year students has **an effect** on their final exam scores.

Types of scientific research hypotheses

2. Based on the direction:

Directed hypothesis: Directed hypotheses are the hypotheses that describe the direct relationships between the various research variables, or describe the effect of the variable or variables on another variable or variables.

Non-Directed hypothesis: These are the hypotheses that confirm the existence of relationships between the research variables, in addition to confirming the differences that exist between them, and here the researcher does not know the directions of the research relationships.

Directed hypothesis	Non-Directed hypothesis
The more hours a student studies, the higher grades and evaluations he will receive (**positive relationship**).	There **is a relationship** between low temperatures and people not leaving their homes.
The higher the illiteracy rate, the less development of society (**negative relationship**).	There **is an effect** of environmental pollution on plant growth.
The more promotions an employee gets in his job, the more active and ambitious his career becomes (**positive relationship**).	There **is a relationship** between an employee's love for his work, his regular attendance, and work productivity.

Types of scientific research hypotheses

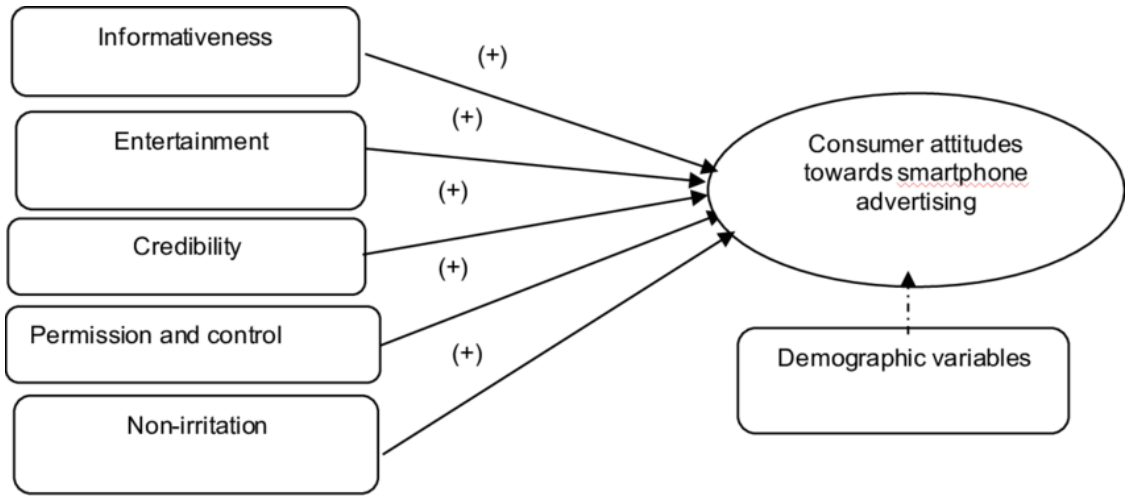
3. Based on the number of variables:

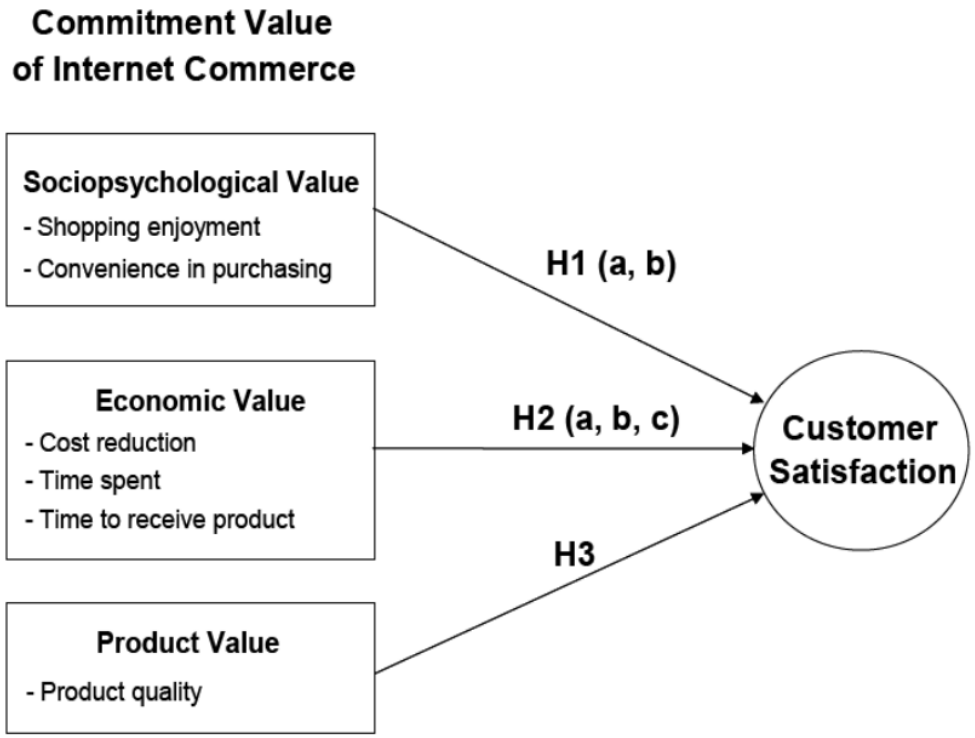
Simple Hypothesis: It involves a relationship between two variables: one independent variable and one dependent variable

Complex Hypothesis: It involves multiple variables, such as multiple independent and dependent variables.

Simple hypothesis	Complex hypothesis
Increasing the number of threads in a parallel computing task will reduce the total execution time	Implementing both data encryption and user authentication mechanisms will increase data security and reduce unauthorized access in cloud storage systems
The use of light mode in an Integrated Development Environment (IDE) reduces the strain on the developer's eyes compared to dark mode.	The combination of agile methodologies and continuous integration (CI) practices leads to faster software development cycles and higher product quality.
Using an optimized binary search algorithm reduces search time in a sorted array compared to a linear search.	Using both unit testing and automated regression testing during the software development process will decrease the number of post-release bugs and reduce overall development costs.

Examples of Research models





Hypothesis examples

Research question	Hypothesis	Null hypothesis
What are the health benefits of eating an apple a day?	Increasing apple consumption in over-60s will result in decreasing frequency of doctor's visits.	Increasing apple consumption in over-60s will have no effect on frequency of doctor's visits.
Which airlines have the most delays?	Low-cost airlines are more likely to have delays than premium airlines.	Low-cost and premium airlines are equally likely to have delays.
Can flexible work arrangements improve job satisfaction?	Employees who have flexible working hours will report greater job satisfaction than employees who work fixed hours.	There is no relationship between working hour flexibility and job satisfaction.
What effect does daily use of social media have on the attention span of under-16s?	There is a negative correlation between time spent on social media and attention span in under-16s.	There is no relationship between social media use and attention span in under-16s.

4. Preparing The Research Design

- The researcher will have to state the conceptual structure within which research would be conducted.
- The function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money.
- Research purposes may be grouped into four categories,
  - Exploration,
  - Description,
  - Diagnosis,
  - Experimentation

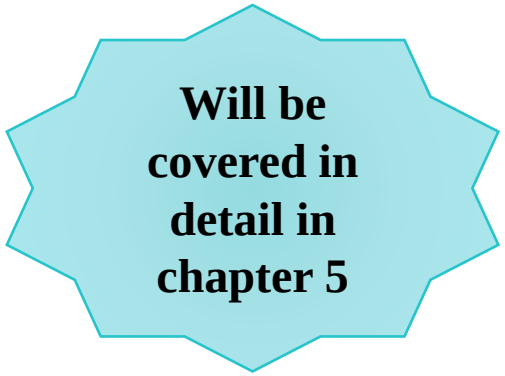
Will be covered in detail in chapter 3 + chapter 4

5. Collection The Data

Several ways are used for data collection, including:

a) By observation:

- the collection of information by way of investigator's own observation, without interviewing the respondents.
- an expensive method
- the information provided by this method is very limited.
- Not suitable in inquiries where large samples are concerned.



b) Through personal interview:

- The investigator seeks answers to a set of preconceived questions through personal interviews.
- output depends upon the ability of the interviewer to a large extent.

c) Through telephone interviews:

- collecting information involves contacting the respondents on telephone itself.
- not a very widely used.
- it plays an important role when the survey has to be accomplished in a very limited time.

d) By mailing of questionnaires:

- Questionnaires are mailed to the respondents with a request to return after completing the same.
- It is the most extensively used method in various economic and business surveys.

e) Through schedules:

- Under this method the enumerators are appointed and given training.
- The enumerators are provided with schedules containing relevant questions.
- These enumerators go to respondents with these schedules.
- Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents.

Will be covered in detail in chapter 5

6. Data Analysis

- The analysis of data requires a number of closely related operations such as
  - establishment of categories,
  - the application of these categories to raw data through coding, tabulation
  - drawing statistical inferences.
- **Coding** operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted.
- **Editing** is the procedure that improves the quality of the data for coding.
- **Tabulation** is a part of the technical procedure wherein the classified data are put in the form of tables.

Will be covered in detail in chapter 6

7. Interpreting the findings

- If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalization, i.e., to build a theory.
- As a matter of fact, the real value of research lies in its ability to arrive at certain generalizations.
- If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as ***interpretation***.

Will be covered in detail in chapter 7

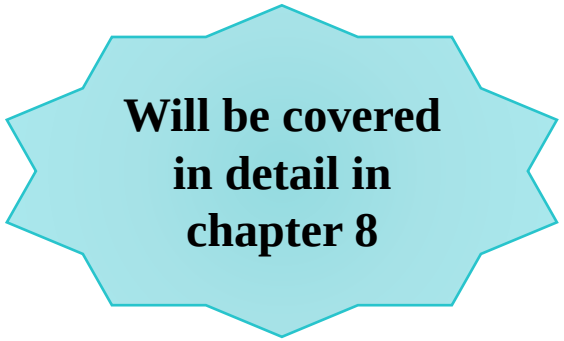
7. Interpreting the findings

- If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalization, i.e., to build a theory.
- As a matter of fact, the real value of research lies in its ability to arrive at certain generalizations.
- If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as ***interpretation***.

Will be covered in detail in chapter 7

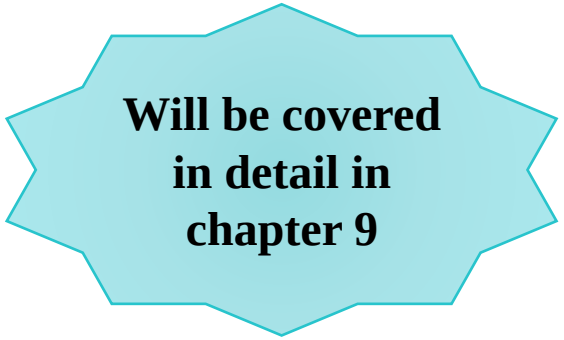
8. Citation and Referencing

- Documentation of each of your sources is essential
- This is referred to as citation and referencing



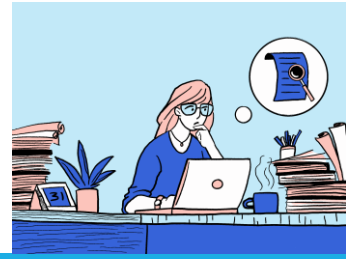
9. Report Writing

- Writing the final report of the research should be done with great care



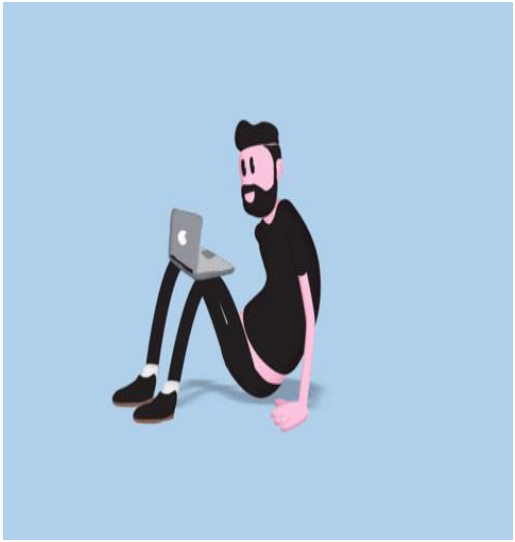
Criteria of Good Research

1. The research's objective should be clearly defined.
2. The research process should be outlined in enough detail for another researcher to repeat it for further advancement while maintaining the continuity of what has already been accomplished. (It can be replicable)
3. The research's procedural design should be meticulously planned in order to produce objective results.
4. Data analysis should be sufficient to reveal the significance of the data, and the methods of analysis used should be relevant. The data's validity and reliability should be doublechecked.
5. Conclusions should be limited to those that are supported by the study data and for which the data provide an adequate foundation.
6. It is generalizable to other settings
7. If the researcher is seasoned, has a good research reputation, and is a person of integrity, greater trust in the research is justified.
8. It is based on some logical rationale and tied to theory.
9. It is achievable!



What is Bad Research?

1. The opposites of what has been discussed.
2. Looking for something when it simply is not to be found.
3. Plagiarizing other people's work.
4. Falsifying data to prove a point.
5. Misrepresenting information and misleading participants.



Research Design Part I

Chapter 3: Research Design: Part I

- What is Research Design
- Selecting the proper research design
- Some Basic terms
- Theoretical Computer Science (TCS) Research
- Empirical Computer Science Research
    - Quantitative Research
        - ✓ Modeling
        - ✓ Experiments
        - ✓ Simulation
        - ✓ Survey Research
    - Qualitative Research (In Unit 4)

What Is a Research Design?

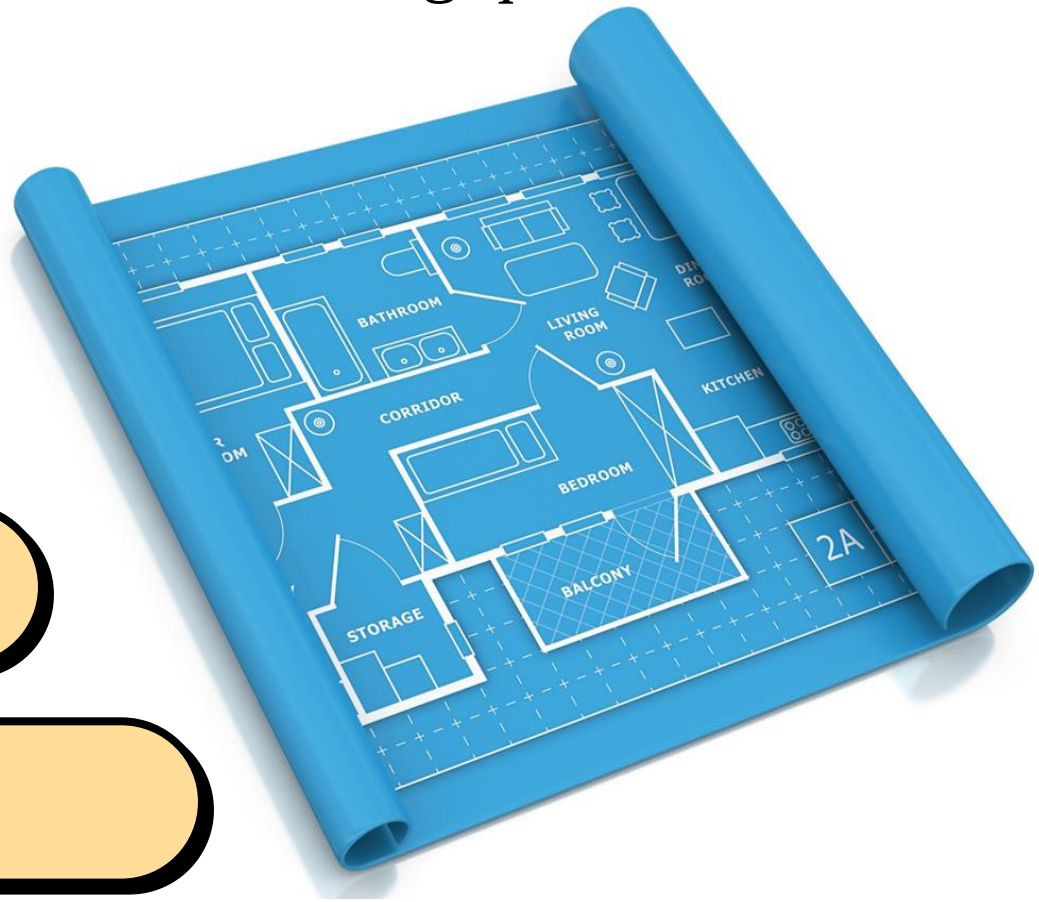
- There are many definitions of research design. **Research design** is the **blueprint** for fulfilling research objectives and answering questions.

Blueprint

Plan

Guide

Framework



What Is a Research Design?

- **Data Management Blueprint:** Provides structured methodologies for data collection, measurement, and analysis.
- **Resource Allocation Guidance:** Aids in efficient resource use and selection among methods like experiments, observations, interviews, and simulations.
- **Data Collection Strategy:** Helps choose between structured or unstructured approaches and determine appropriate sample sizes.
- **Research Approach Decision:** Assists in selecting between quantitative or qualitative research methods.
- **Systematic Research Execution:** Ensures methodical, rational, and efficient research addressing key questions or hypotheses.
- **Method Refinement:** Enables researchers to tailor methods to best fit the subject matter, enhancing study success.

Time-based procedural plan

Based on Research Questions

Guide for selecting sources of information & methods

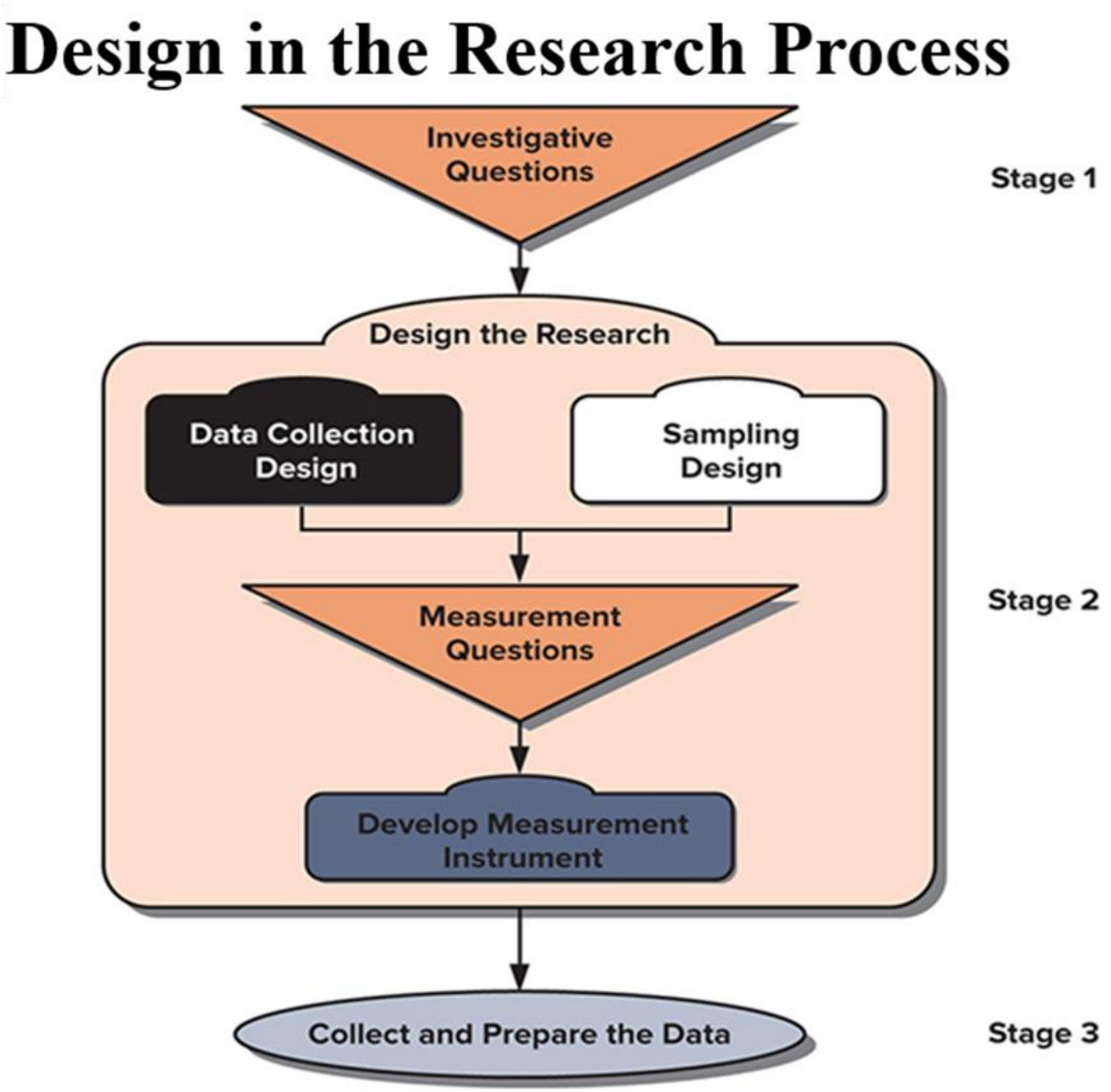
Framework for specifying relationships among variables

A procedural outline for every research activity

What Is a Research Design?

Formally:

A research design is a detailed and systematic plan or blueprint for carrying out a research study to ensure that the study is valid, reliable, and produces meaningful results.



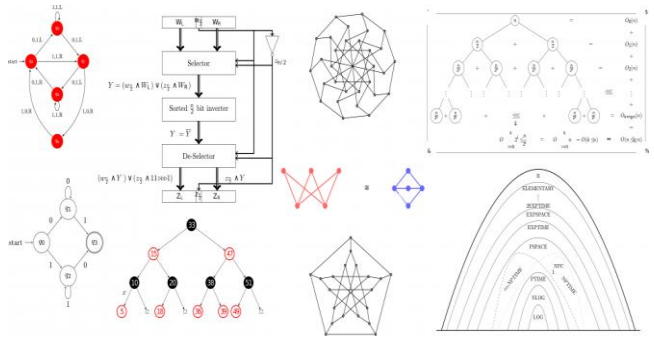
Selecting the proper research design

- Once the research questions are clearly defined, the researcher must choose the most appropriate research design to address them.
- Selecting the proper research design involves several steps and considerations to ensure that the research questions are answered effectively and that the results are valid and reliable.
- These include:
    1. The nature of the study (Exploration, Description, Diagnosis, Experimentation…)
    2. The assessment of the available resources (time, budget, and personnel available for your research. Some designs require more resources than others)
    3. Ethical Considerations: Ensuring the research design adheres to ethical standards, including informed consent, confidentiality, and minimizing harm.
    4. Choosing a Methodology:
        - **Theoretical Research:**
        - **Empirical Research:** which classified into:
            - ✓ **Quantitative Research**: Focuses on quantifying variables and analyzing statistical relationships. Methods include surveys, simulations, and experiments.
            - ✓ **Qualitative Research**: Focuses on understanding concepts, experiences, or social phenomena. Methods include interviews, focus groups, and content analysis.
            - ✓ **Mixed Methods**: Combines both qualitative and quantitative approaches
    5. Selecting the most suitable method to answer the research questions
    6. Identifying your population and sampling method.

1. Theoretical Computer Science (TCS) Research:

- The primary goal of TCS is to advance or refine formal models and algorithms within computing science.
- TCS is based on formal rules derived from logic and mathematics like Boolean logic and number theory.

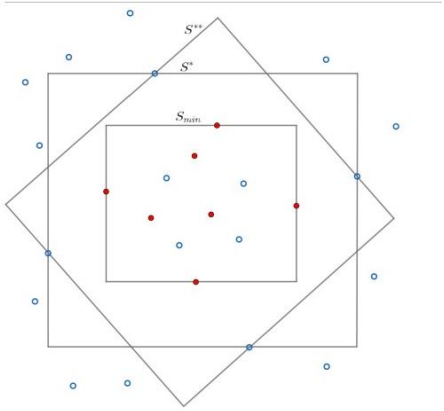
  Hence, it is referred to as “Formal Method” as well.
- Research focus in TCS often involves forming hypotheses that lead to generalizations.
- Some of the main techniques when dealing with problems are iteration, recursion, and induction.
- The outcome of research could be:
  - Theorems: Formal proofs derived from hypotheses.
  - Formal Models: New or improved conceptual frameworks.
  - Algorithm Improvements: Enhancements to existing algorithms.



Examples of research papers in TCS:

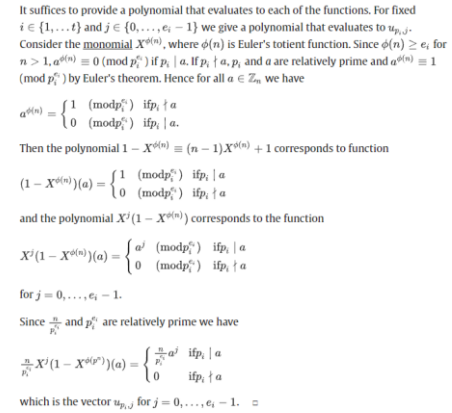
1. Armaselu, B. (2024). Finding the Largest Separating Rectangle among Two Point Sets. Theoretical Computer Science, 114778.

The research discusses the problem of finding a rectangle that contains all red points, minimizes the inclusion of blue points, and has the maximum area, termed as a maximum-area separating rectangle (MSR).



2. Guha, A., & Dukkipati, A. (2015). A faster algorithm for testing polynomial representability of functions over finite integer rings. Theoretical Computer Science, 579, 88-99."

A faster algorithm for testing polynomial representability of functions over finite integer rings - This research is also an improvement to an already devised algorithm in polynomial representability. Reading the article that describes this research, one, at the first glance, would say that is a research in mathematics However, when it is read carefully, the algorithms that have been provided explain why this research has happened in the computer science area.



It suffices to provide a polynomial that evaluates to each of the functions. For fixed $i \in \{1, \ldots t\}$ and $j \in \{0, \ldots, e_i - 1\}$ we give a polynomial that evaluates to $u_{p_i,j}$. Consider the monomial $X^{\phi(n)}$, where $\phi(n)$ is Euler's totient function. Since $\phi(n) \geq e_i$ for $n > 1, a^{\phi(n)} \equiv 0 \pmod{p_i^{e_i}}$ if $p_i \mid a$. If $p_i \nmid a$, p_i and a are relatively prime and $a^{\phi(n)} \equiv 1 \pmod{p_i^{e_i}}$ by Euler's theorem. Hence for all $a \in \mathbb{Z}_n$ we have

$$a^{\phi(n)} = \begin{cases} 1 \pmod{p_i^{e_i}} & \text{if } p_i \nmid a \\ 0 \pmod{p_i^{e_i}} & \text{if } p_i \mid a. \end{cases}$$

Then the polynomial $1 - X^{\phi(n)} \equiv (n-1)X^{\phi(n)} + 1$ corresponds to function

$$(1 - X^{\phi(n)})(a) = \begin{cases} 1 \pmod{p_i^{e_i}} & \text{if } p_i \mid a \\ 0 \pmod{p_i^{e_i}} & \text{if } p_i \nmid a \end{cases}$$

and the polynomial $X^j(1 - X^{\phi(n)})$ corresponds to the function

$$X^j(1 - X^{\phi(n)})(a) = \begin{cases} a^j \pmod{p_i^{e_i}} & \text{if } p_i \mid a \\ 0 \pmod{p_i^{e_i}} & \text{if } p_i \nmid a \end{cases}$$

for $j = 0, \ldots, e_i - 1$.

Since $\frac{n}{p_i^{e_i}}$ and $p_i^{e_i}$ are relatively prime we have

$$\tfrac{n}{p_i^{e_i}} X^j(1 - X^{\phi(p^{e_i})})(a) = \begin{cases} \frac{n}{p_i^{e_i}} a^j & \text{if } p_i \mid a \\ 0 & \text{if } p_i \nmid a \end{cases}$$

which is the vector $u_{p_i,j}$ for $j = 0, \ldots, e_i - 1$. □

2. Empirical Research in Computer Science and Information

- What does “empirical” mean?

Terminological Prelude

- **Empirical**: relying upon or derived from observation or experiment
- **Empiricism**: the employment of empirical methods in science.
- **Empirical Data**: data produced by an experiment or observation.
- **Empirical Finding**: Something that is observed from real-world observation or data, in contrast to something that is deduced from theory.
- **Empirical Method**: Simplified method of design justified by experience or testing.

2. Empirical Research in Computer Science and Information

- Empirical work should complement theoretical work
    - Theories often have holes (e.g., How big is the constant term? Is the current problem a “bad” one?)
    - Theories are suggested by observations
    - Theories are tested by observations
    - Conversely, theories direct our empirical attention
- In addition, empirical means “wanting to understand behavior of complex systems”
- Empirical methods in computer science aim to answer questions like these:
    - Is program A really more efficient than program B, in practice?
    - Does A’s run time vary more widely than B’s on different inputs?
    - Is program A really more accurate than program B, when applied to some task?
    - What are the best parameters for compiling or running a particular program?

2. Empirical Research in Computer Science and Information

- Empirical methods could be used in either quantitative or qualitative research

Criteria	Quantitative Research	Qualitative Research
Methods Used	Usually collects data through surveys, experiments, content analyses or structured interviews	Usually collects evidence through unstructured interviewing, observation or participant observation or content analysis

- **Next…. The most commonly used methods for each of these methodologies are explained next**

First: Quantitative Research Methods in CS and IS

- CS is characterized as an empirical field ,where each new program is considered an experiment. The structure and behavior of these programs can be studied
- Among the most commonly used methods are:

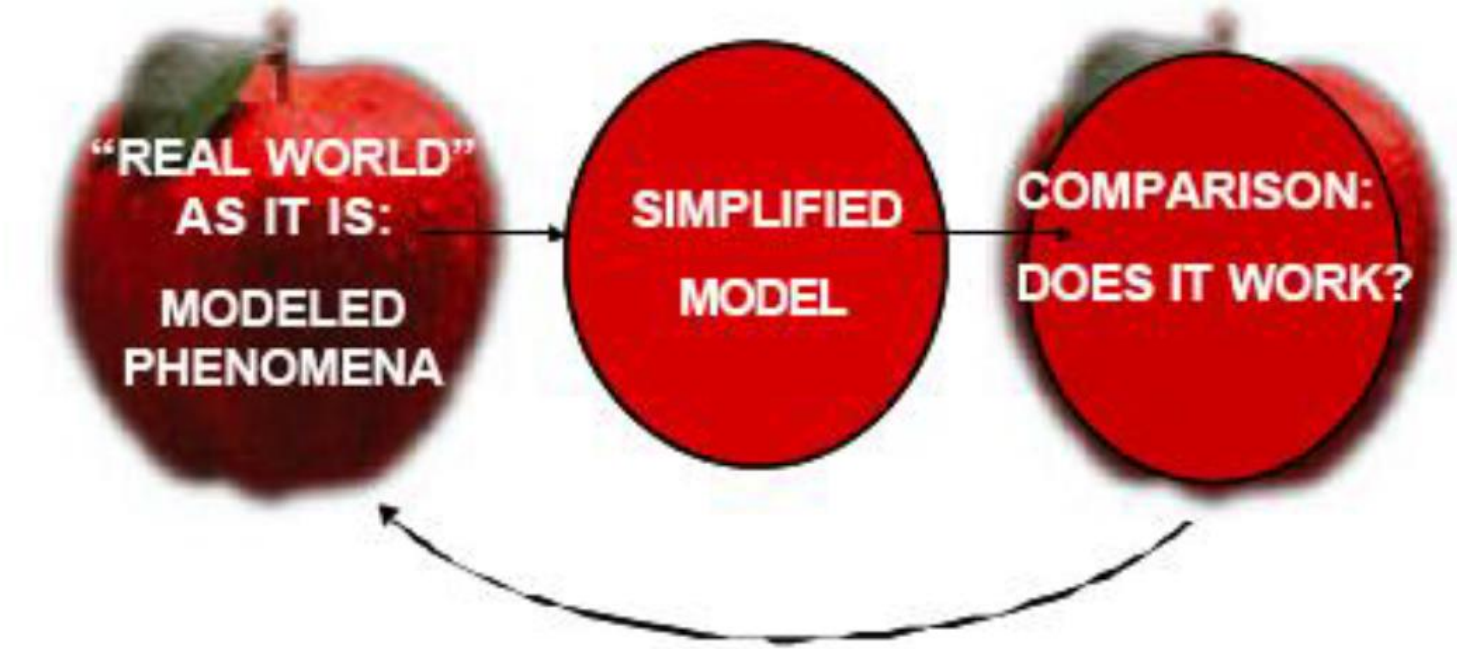
1. Modeling
2. Experimental Computer Science
3. Computer Simulation
4. Survey (Sometimes partially required combined with other methods)
5. System Development Methodologies (An Overview)

1. Modeling

- Modeling is the process of constructing an abstract model for an actual system motivated by the intended application of the model in the study.
- This model will possess a lower level of complexity compared to the system it represents. Consequently, it will enable the researcher to gain a deeper understanding of the system and utilize the model to conduct experiments that were not otherwise feasible within the system due to financial or accessibility constraints.
- Modeling is **not the object** of the research, it **is part of an arsenal of instruments** used by the researchers to study and understand the research's object. Therefore, the modeling approach is frequently employed when combined with the other approaches.
- As modeling is the process of making a model based on how it will be used in a study itself. there isn't a single modeling method that works for all systems. Instead, scientists make models that accurately show the important parts of a system while ignoring or only roughly approximating the parts that aren't important to their main study.
- Figuring out the relative importance of the factors and how they affect each other is an important part of the modeling method. Models that leave out important information or focus too much on less important factors give wrong results. In groups that depend on modeling, there is a lot of disagreement about which model should be used.

1. Modeling

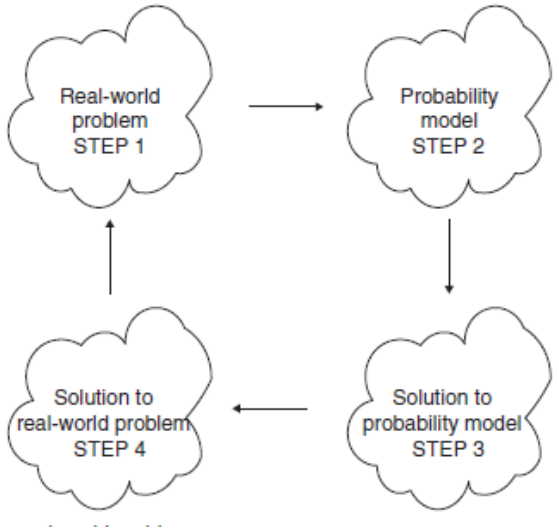
This figure emphasizes the iterative nature of modeling, where the real-world system is abstracted into a model, tested for accuracy, and adjusted to improve its representation of the original phenomenon.



"Real World"	Model
Program	Compiler theory
Artificial Neural Networks	Experiments testing ANN
Computer hardware	Simulation

1. Modeling

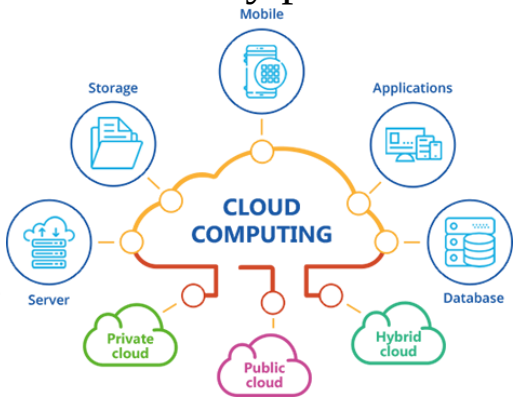
- **Real World (Modeled Phenomena)**:
  - Represents the actual system or phenomenon being studied, labeled as the "Real World as it is." This is the complex, raw reality that researchers aim to understand, abstract, and model.
- **Simplified Model**:
  - A simplified representation of the real-world system, created to focus on essential components and interactions. The model abstracts away complexities, highlighting the critical aspects necessary for study and analysis.
  - The goal of this model is to capture the most relevant features of the real-world system while ignoring less important or secondary aspects.
- **Comparison (Does It Work?)**:
  - This step involves comparing the simplified model to the real-world phenomena. Researchers check if the model accurately represents and predicts behaviors in the real system. If it works well, the model can be validated or refined based on discrepancies found during comparison.
- **Cycle**:
  - The process is cyclical, with continuous refinement. If the model doesn't accurately capture the real-world behavior, it is adjusted and tested again.



1. Modeling Example: The modeling of a distributed system to analyze its performance under network failure conditions.

- **Real World (Modeled Phenomena)**: The real-world system is a **distributed computing network** where multiple computers communicate over a network to complete tasks. The network may experience issues like node failures, latency, or bandwidth limitations.
- **Simplified Model**: A simplified model could abstract the distributed system into a **graph** where nodes represent computers and edges represent network connections. The model may include failure probabilities for nodes and bandwidth constraints but ignore lower-level details like physical hardware specifications or exact network protocols.
- The model might focus on how **network failures** impact **task completion time** and **system reliability** under various conditions.
- **Comparison (Does It Work?)**: The model's predictions (e.g., system performance under node failures) are compared against actual experiments or historical data from the real-world distributed system. If the model successfully predicts the system's behavior under failure conditions, it’s considered a valid model.
- **Cycle:** If not, adjustments to the model are made

(e.g., modifying failure rates or communication delays).



The process of generating a formal description of a model to validate the functionality or accuracy of a system is referred to as **model-checking**

2. Experimental Computer Science

- Experiments are research studies **designed to explain why something happens**; in these studies the researcher proposes a **causal hypothesis** (a speculation that one variable causes a change in another variable).
- The researcher then **manipulates the independent variable (IV)** and **observes whether the dependent variable (DV) is affected** as hypothesized.
- Experimental design is an approach that allows the researcher to take control of all variables that could affect the outcome of an experiment.

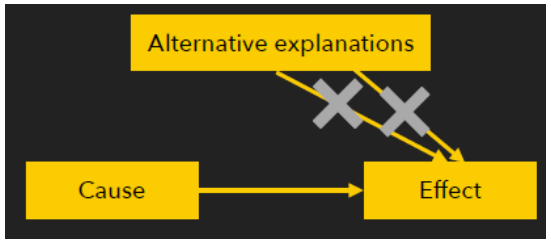
- Three **conditions** have to be met simultaneously as evidence of a causal relationship:

1. The cause preceded the effect

2. The cause was related to the effect

3. Extraneous variables did not influence DVs

We can find no plausible alternative explanation for the effect other than the cause



2. Experimental Computer Science

Example 2:

Research Question:

Do malware attacks influence individuals?

Conditions 1: The cause (malware attacks) precedes the influence on individuals

Conditions 2: There is a correlation (relation) between being exposed to malware attacks and experiencing the influence of this attack (by compromising their data, financial security, and overall digital trust, leading to psychological and economic consequences)

Conditions 3: During the experiment, the researcher has to make sure that he is examining the effect of malware attack only by neutralizing other possibilities)

2. Experimental Computer Science

Key concepts in experiments:

- **An Experiment:**
  - A study in which an intervention (treatment) is deliberately introduced to observe its effects
- **Treatment**:
  - Experimental stimulus was given to some participants but not others.
  - Successful when treatment group responses differ from the control group.
  - Example: Interface design differences in e-commerce research.
- **Treatment manipulation**:
  - Involves controlling the cause-effect relationship by varying stimulus levels.
  - Typically includes pre-test and post-test measures.
- **Experimental controls**:
  - Ensure observed effects are due to the treatment, not confounding factors.
  - Example: Placebo groups in medical experiments.
  - Controls help rule out rival explanations.

2. Experimental Computer Science

Experiments could be:

- **Lab experiments**:
  - Conducted in controlled environments.
  - Examples include students performing tasks on a computer or rats undergoing treatments.
  - Provide high control over the situation.
- **Field experiments:**
  - Performed in real-world settings.
  - Example: Manipulating elements on an e-commerce website like Amazon.
  - Require collaboration with organizations hosting the technology.
  - Achieve higher ecological and internal validity.
  - Becoming more popular in information systems research.

2. Experimental Computer Science

What are the principles of Experimental Design?

There are three principles, as explained by R.A. Fisher, of Experimental Design:

- **Replication**

  Replication involves repetition of the basic experiment. The principle is that even when the same treatment is used in other experiments, the output would differ. Replication in experimental design helps to study the variation in the yield of different experiments.

  “r” refers to a number of replicates which implies the no. of experimental units per treatment.

- **Randomization**

  Randomization involves distributing the treatment to different experimental units using probability. This ensures that each experimental unit is likely to receive the treatments. Randomization eliminates the probability of bias from the result of the experimental research design.

- **Local Control**

  Local control is the method to control the error variation and thus reducing the error by arranging the experimental units. The value of a variable is kept constant to keep it from affecting the conclusion of the experiment.

2. Experimental Computer Science

An experimental design must deal with four issues:

1. The people (subjects) who participate in the experiment. وهي العناصر او الافراد الذين سيشاركون في التجربة
2. The **independent variable or variables**, which are also **called the treatment variables**. These are the variables the researchers manipulate during the experiment. تسمى متغيرات المعالجة وهي التي نريد فحص تطبيق تأثيرها على المشاركين
3. The **dependent variable**, or **the effect** that the researchers measure. وهو (المتغير التابع) التأثير الذي نريد قياسه او الوصول اليه من خلال التجربة
4. The plan for controlling extraneous variables. وسيلة الباحث للتحكم بالعوامل الأخرى التي قد تؤثر على نتائج التجربة ونحن غير مهتمين بدراستها حاليا في التجربة

2. Experimental Computer Science

Experiment Example: Testing the Impact of Different Sorting Algorithms on Execution Time

Subjects:

- Different datasets of various sizes and complexities.

Independent Variable (IV):

- The sorting algorithms used (e.g., QuickSort, MergeSort, and BubbleSort).
- These are the variables that the researcher manipulates during the experiment, i.e., the specific sorting algorithm applied to the datasets.

Dependent Variable (DV):

- The execution time (measured in seconds or milliseconds) taken by each sorting algorithm.
- This is the effect that the researchers aim to measure, which depends on the sorting algorithm used.

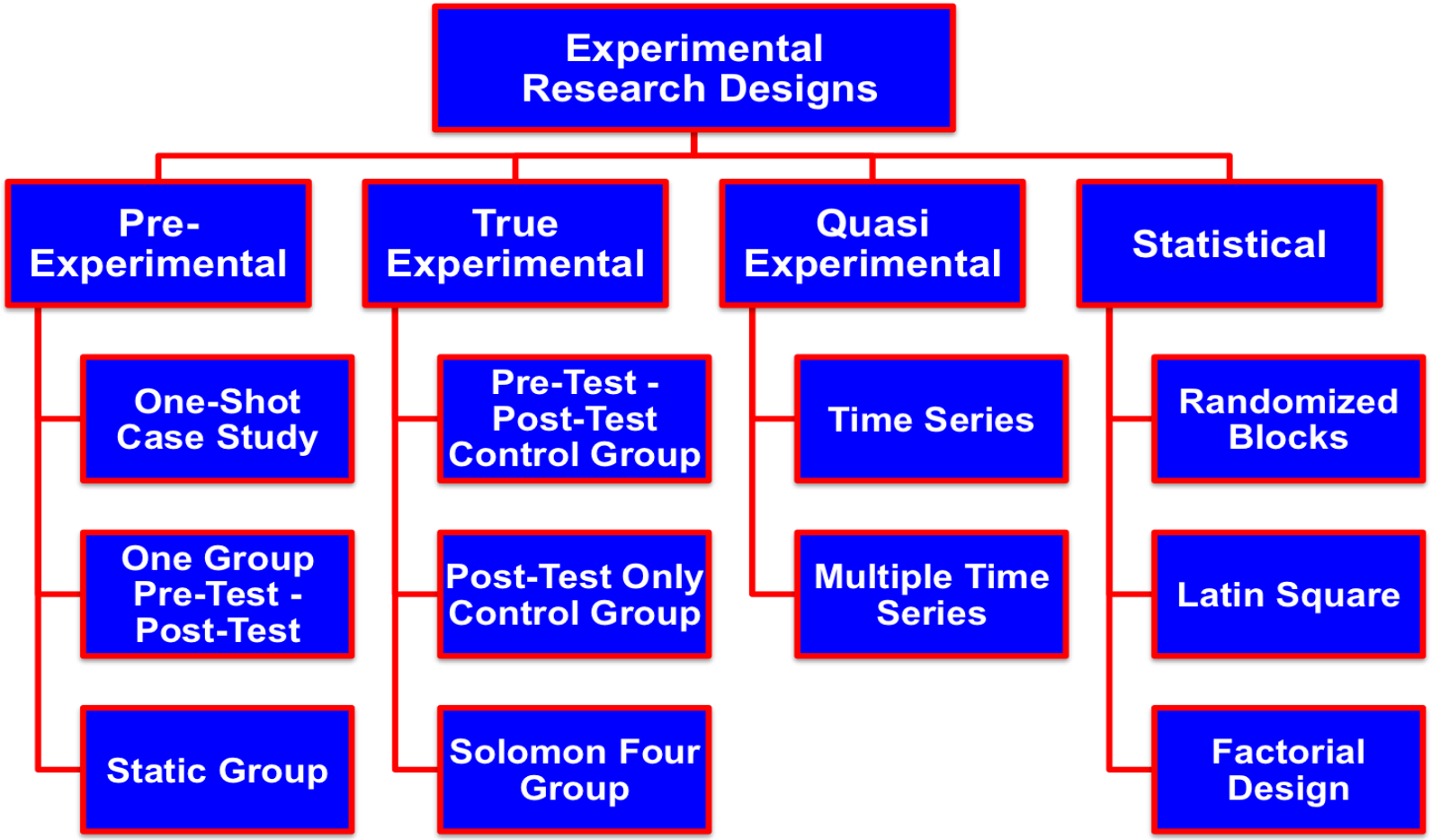
Extraneous Variables (Factors to Control):

1. **Hardware specifications; Programming language and compiler; Dataset characteristics; and Background processes on the machine**

In this experiment, the **independent variable** is the type of sorting algorithm, and the **dependent variable** is the execution time. The extraneous variables, such as hardware and programming language, need to be controlled to ensure the experiment's results are valid and reliable

2. Experimental Computer Science

There are different types (designs) of experiments as illustrated in the graph



2. Experimental Computer Science

This figure represents different types of **Experimental Research Designs**:

- **Pre-Experimental**: Basic designs with no random assignment.
  - **One-Shot Case Study**: One group was observed after treatment.
  - **One Group Pre-Test - Post-Test**: A group tested before and after treatment.
  - **Static Group**: Comparison of treated and untreated groups, but no pre-testing.
- **True Experimental**: Random assignment with control and experimental groups.
  - **Pre-Test - Post-Test Control Group**: Groups tested before and after treatment with a control group.
  - **Post-Test Only Control Group**: Post-test given after treatment without a pre-test.
  - **Solomon Four Group**: Combines both pre-test/post-test and post-test only groups.
- **Quasi-Experimental**: No random assignment but comparison is made.
  - **Time Series**: Repeated observations over time.
  - **Multiple Time Series**: Observations with a comparison group over time.
- **Statistical**: Uses statistical methods for group assignment or variable control.
  - **Randomized Blocks**: Randomly assign groups within blocks.
  - **Latin Square**: Control for two variables.
  - **Factorial Design**: Study multiple factors at once

2. Experimental Computer Science

Evaluation of Experiments

Advantages

1. Ability to manipulate IV
2. Use of control group
3. Control of extraneous variables
4. Replication possible
5. Field experiments possible

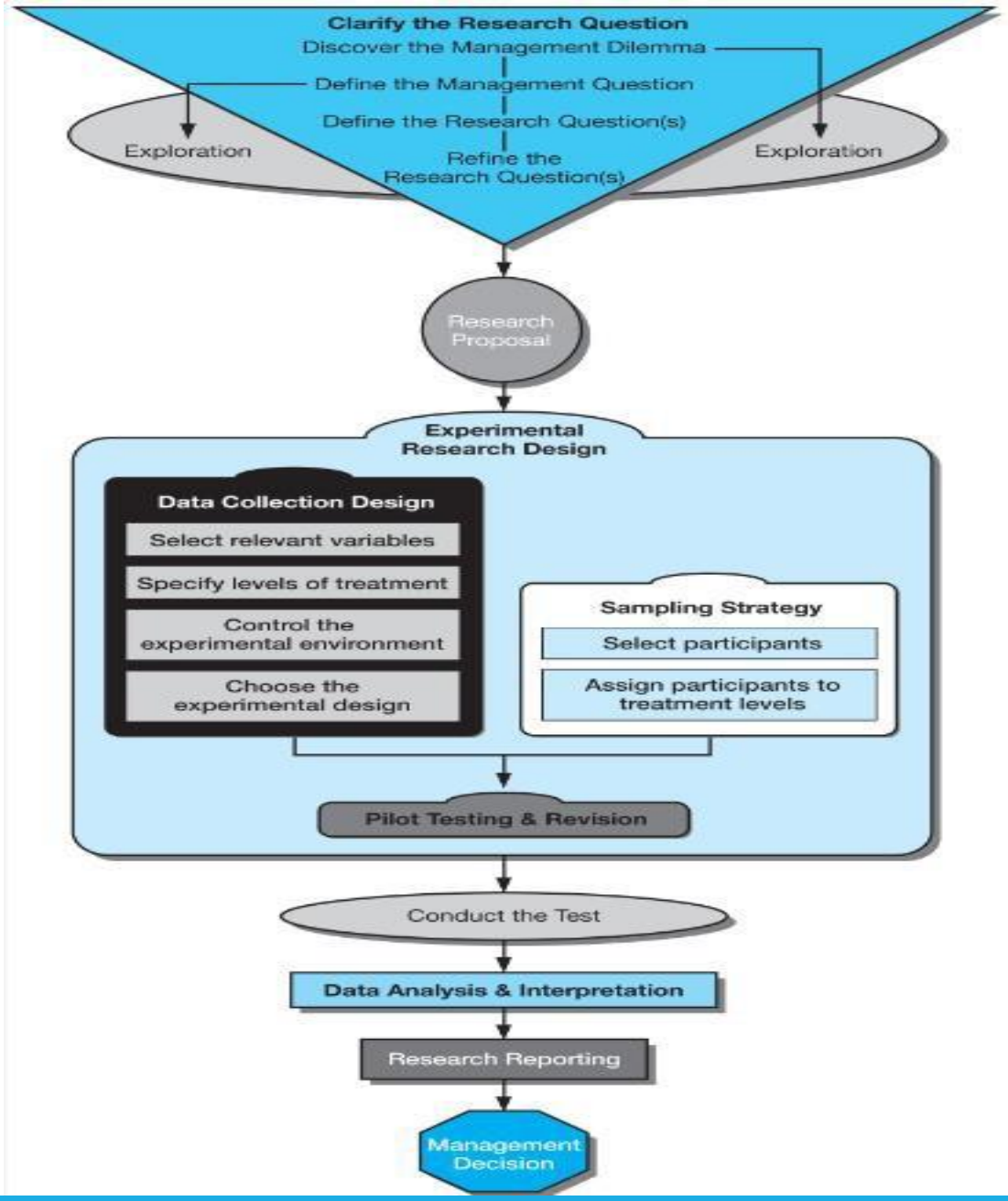
Disadvantages

1. Artificiality of labs
2. Non-representative sample
3. Expense
4. Focus on present and immediate future
5. Ethical limitations

2. Experimental Computer Science

Steps for conducting an Experiment

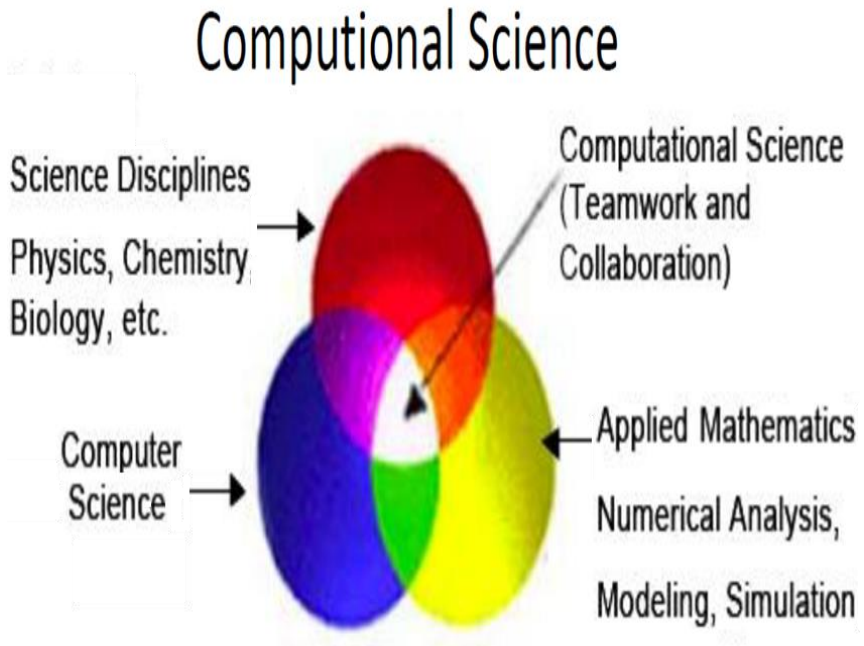
1. Select relevant variables
2. Specify treatment/intervention levels
3. Control the experimental environment
4. Choose an experimental design
5. Select and assign participants
6. Pilot test data collection protocols and revise
7. Conduct Experiment



3. Computer Simulation

- Simulation is a method of study commonly employed in computer science (CS) to explore systems or regimes that are not part of the experimental domain or the systems being developed or constructed.
- The emergence of the universe is a collection of intricate events that are typically not feasible to replicate in laboratory settings.
- Computer simulation techniques are used in fields such as astronomy, physics, and economics.
- These techniques are also used in specialized areas such as the analysis of non-linear systems, virtual reality, and artificial life.
- Simulation approaches can be readily used in many projects, such as the investigation of a newly established network protocol. In order to evaluate this protocol, it is necessary to construct a vast network featuring several costly network equipment. However, achieving this network is not a straightforward task. Therefore, the simulation approach can be employed.

3. Computer Simulation

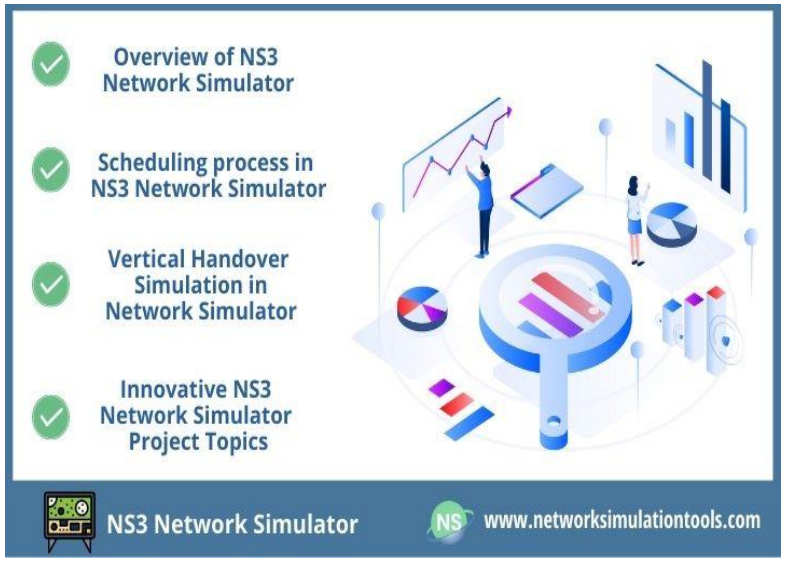


- This picture shows how Computational Science is multidisciplinary. It involves people from different fields working together to use computer simulations to answer difficult problems.
- Science Disciplines (Physics, Chemistry, Biology, etc.) are the different natural sciences that use computer models and programs to figure out how complicated things work. For instance, scientists may use computer models to simulate how living things work, while physicists may use computer models to simulate how particles behave.
- The field of computer science is responsible for creating the algorithms, data structures, and hardware that are needed to run simulations and handle large amounts of data.
- Applied mathematics uses math methods like numerical analysis, modeling, and simulation to help come up with and answer the equations that explain scientific phenomena.
- Teamwork and Collaboration in Computational Science in the middle of the picture emerges from the integration of these fields.
- To use computational methods and models to solve real-world scientific problems, people need to work together.

3. Computer Simulation

- For example, a company might be developing a new routing protocol aimed at improving performance in highly dynamic mobile networks (such as vehicular networks). Physically setting up and managing a network of hundreds of vehicles communicating across several cities is impractical. Instead, they use a simulator like NS-3 to create a virtual environment where nodes represent cars and access points, and test how efficiently the protocol routes data in the presence of moving vehicles, varying signal strengths, and network congestion.

ns-3 is a discrete-event network simulator for Internet systems, targeted primarily for research and educational use. ns-3 is free, open-source software, licensed under the GNU GPLv2 license and maintained by a worldwide community.



3. Computer Simulation

- Simulation approaches can be readily used in many projects, such as the investigation of a newly established network protocol. In order to evaluate this protocol, it is necessary to construct a vast network featuring several costly network equipment. However, achieving this network is not a straightforward task. Therefore, the simulation approach can be employed.

- The most commonly used simulators are :
  - QualNet: A high-performance, commercial simulator ideal for real-time, large-scale network simulations, especially for wireless and military applications.
  - NS-3: An open-source simulator known for its flexibility and extensive protocol support, widely used in research and academic projects.
  - OMNeT++: A modular, extensible open-source simulator with a user-friendly GUI, commonly used in academic research for layered systems like vehicular and sensor networks.
  - Mininet: A lightweight emulator specialized in Software Defined Networking (SDN), allowing real-time SDN and OpenFlow protocol testing on a single machine.
  - Packet Tracer: A network simulation tool developed by Cisco, primarily used for educational purposes to simulate and visualize network configurations and troubleshoot Cisco-based network setups. It's widely used for CCNA training.

Simulation = Experiment based on a model / Experiments based on a model are called simulations.

Experiments based on a model are called simulations.

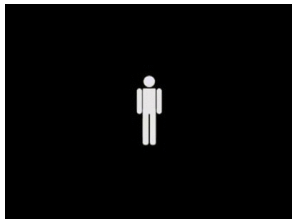
Simulation = Experiments based on a model

4. Survey Research

Survey research is a method of collecting data by asking a predefined set of questions to a sample of individuals, typically in the form of questionnaires or interviews. This approach aims to gather information about people's opinions, behaviors, attitudes, or characteristics. Survey research is often used in various fields, such as social sciences, market research, and public health, to understand trends and make generalizations about a larger population based on the sample responses.

- Survey research is particularly **effective for studying a large population**, **gathering a wide range of opinions**, **cost effective,** and **exploring relationships between variables.**
- However, it also has limitations, such as **potential response biases** and **the reliance on self-reported data**.
- There are some basic terms that should be identified. These are **Target Population**, **Study Population**, **Sample, Sampling design or technique, Sample Size, Questionnaire Design** (will be explained in detail in data collection chapter)

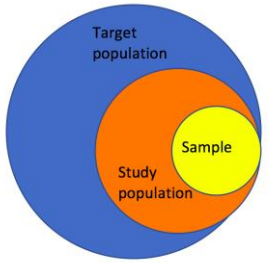
Some Basic Concepts



1. Target Population (المجتمع)

Is "The entire population, or group, that a researcher is interested in researching and analysing. A sampling frame is then drawn from this target population".

2. Study Population (مجتمع الدراسة)



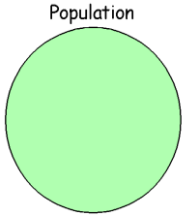
Is "The subset of the target population available for study and meet the criteria of the study".

3. Study Sample (عينة الدراسة)

Is "The selected group of individuals or elements in the population from which data are collected for a study".

Why do we select a sample?

- We select a sample since in most cases it is not applicable to study the whole study population.
- Samples are easier to collect data from because they are practical, cost-effective, convenient, and manageable.



4. **Sampling Design or Technique (طريقة تحديد العينة)**

Is " The technique of selecting individual members or a subset of the population to make statistical inferences from them and estimate the characteristics of the whole population".

Different methods are available and will be all covered in detail later

4. Survey Research

5. Sampling Technique or Design

A sampling design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample design may as well lay down the number of items to be included in the sample i.e., the size of the sample

Characteristics of a good sample design

From what has been stated in Chapter 2, the characteristics of a good sample design are :

(a) Sample design must result in a truly representative sample (عينة ممثلة فعليا لمجتمع الدراسة).

(b) Sample design must be such which results in a small sampling error (تقليل الأخطاء الناتجة من عملية اختيار العينة قدر الامكان).

(c) Sample design must be viable in the context of funds available for the research study (ضمن الإمكانات المتوفرة).

(d) Sample design must be such so that systematic bias can be controlled in a better way (ان لا يكون هناك تحيز).

4. Survey Research

Types of sample design

Sample designs are basically of two types: probability sampling and on-probability sampling.

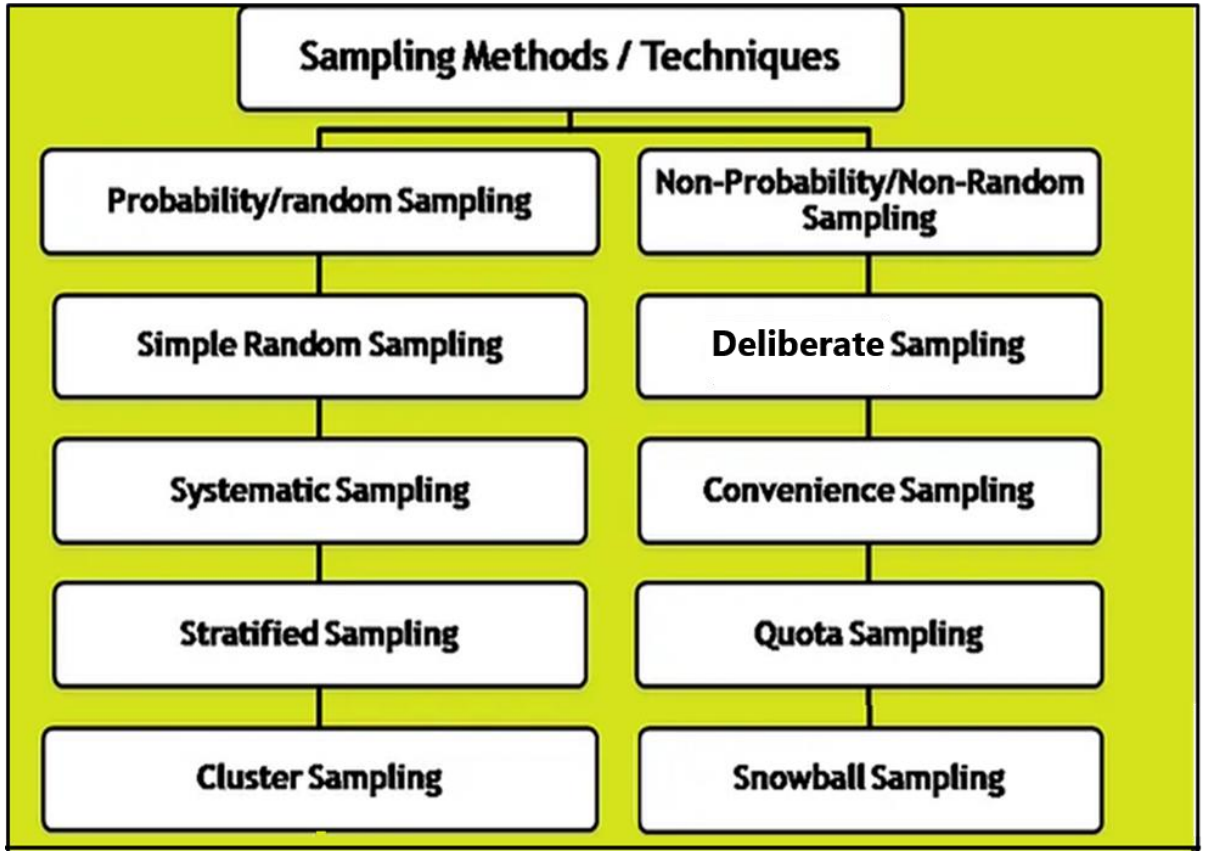
- **Probability sampling**: Probability sampling is also known as ‘random sampling’ or ‘chance sampling’. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample.

- **Non-probability sampling:** Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling, purposive sampling and judgement sampling.

Probability sampling	Non-probability sampling
The sample is selected at random.	Sample selection based on the subjective judgment of the researcher.
Everyone in the population has an equal chance of getting selected.	Not everyone has an equal chance to participate.
Used when sampling bias has to be reduced.	The researcher does not consider sampling bias.
Useful when the population is diverse.	Useful when the population has similar traits.
Used to create an accurate sample.	The sample does not accurately represent the population.
Finding the right respondents is not easy.	Finding respondents is easy.

4. Survey Research

- Each type of sampling design has different techniques. The most common for each type are shown in the figure.



4. Survey Research

Main Types of Probability Sampling

1.Simple Random Sampling.

2.Systematic Random Sampling.

3.Stratified Random Sampling.

4.Cluster sampling

To be able to use the random sampling, the researcher should have access to the sampling frame.

Sampling frame is the actual list of individuals that the sample will be drawn from. Ideally, it should include the entire target population (and nobody who is not part of that population).

If the researcher doesn't have the sampling frame , using the random sampling is not allowed

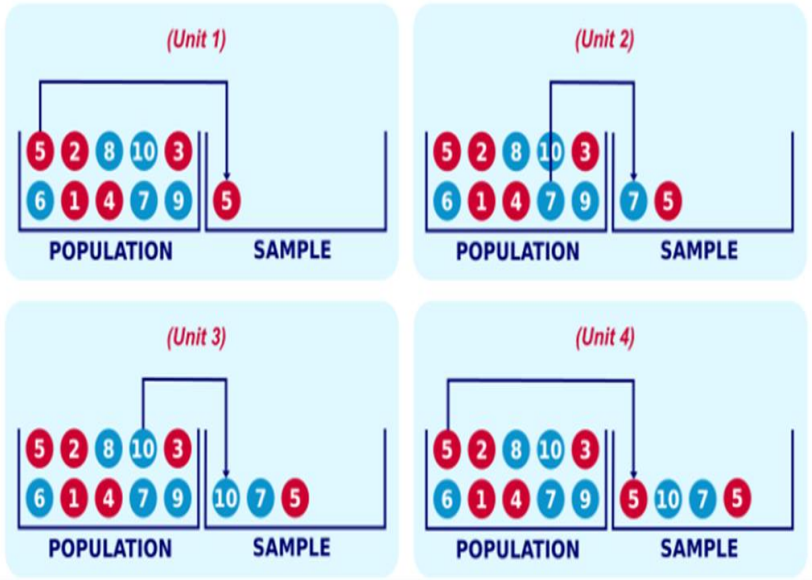
4. Survey Research

A simple random sample is a randomly selected subset of a population. In this sampling method, each member of the population has an exactly equal chance of being selected.

- **Example 1:** The names of 25 employees being chosen out of a company of 250 employees.
- **Example 2:** Using the lottery method is one of the oldest ways and is a mechanical example of random sample.

(1) Simple random sampling:

- known as chance sampling or **probability sampling** where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected.
- In case of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.



4. Survey Research

2. Systematic sampling is a probability sampling method where the researcher chooses elements from a target population by selecting a random starting point and selecting sample members after a fixed 'sampling interval.'

(2) Systematic sampling

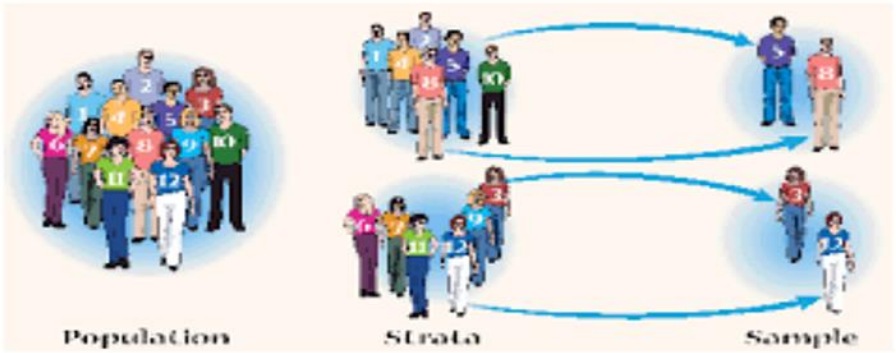
- An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start.
- This procedure is useful when sampling frame is available in the form of a list.
- In such a design the selection process starts by picking some random point in the list and then every nth element is selected until the desired number is secured.

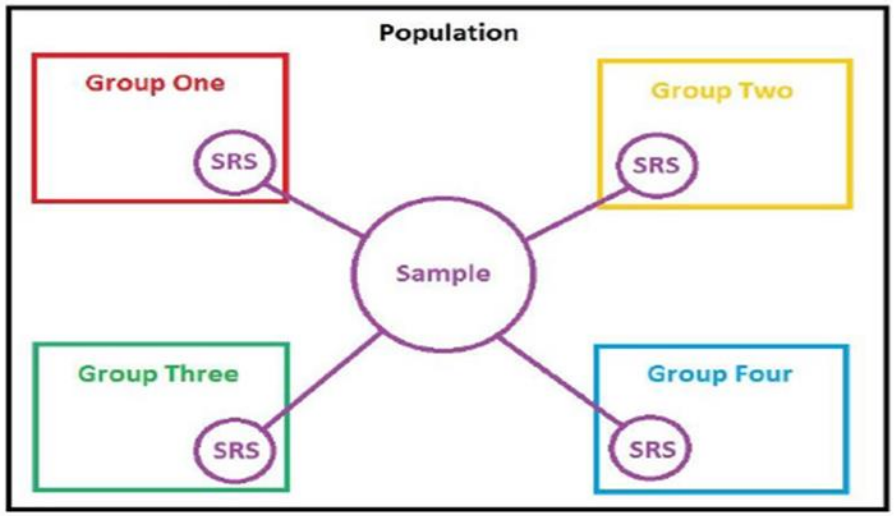
- **Example 1:** As a hypothetical example of systematic sampling, assume that, in a population of 10,000 people, a statistician selects every 100th person for sampling.
- **Example 2:** Sampling on the spot (choosing a sample of the customers of a certain store)
  As you cannot get a complete list of your store's customers, you instead choose to sample every kth customer as they exit the store. This allows you to include both those who buy items and those who do not.

4. Survey Research

(3) Stratified sampling

- If the population from which a sample is to be drawn does **not constitute a homogeneous group**, then stratified sampling technique is applied so as to obtain a representative sample.
- In this technique, the population is stratified into a number of non-overlapping subgroups and sample items are selected from each subgroup.
- If the items selected from each subgroup is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as *stratified random sampling*.





4. Survey Research

- Stratified random sampling is a method of sampling that involves the division of a population into smaller subgroups known as strata. In stratified random sampling, or stratification, the strata are formed based on members' shared attributes or characteristics, such as income or educational attainment. Stratified random sampling has numerous applications and benefits, such as studying population demographics
- The stratification could be either : (1) proportionate (نسبي) or (2) disproportionate (غير نسبي)

What Is the Difference Between Proportionate and Disproportionate Stratification?

In a proportionate stratified method, the sample size of each stratum is proportionate to the population size of the stratum. This type of stratified random sampling is often a more precise metric because it's a better representation of the overall population. In a disproportional stratified sample, the size of each stratum is not proportional to its size in the population

Example: suppose we want to conduct a study on the students of PTUK given that we have access to the sampling frame. As the number of students in each college is not the same, then, in order to have a representative sample of all students we can use stratified random sampling where the number of students in the drawn sample should be proportional to the total number of each college's students number to the whole number of students.

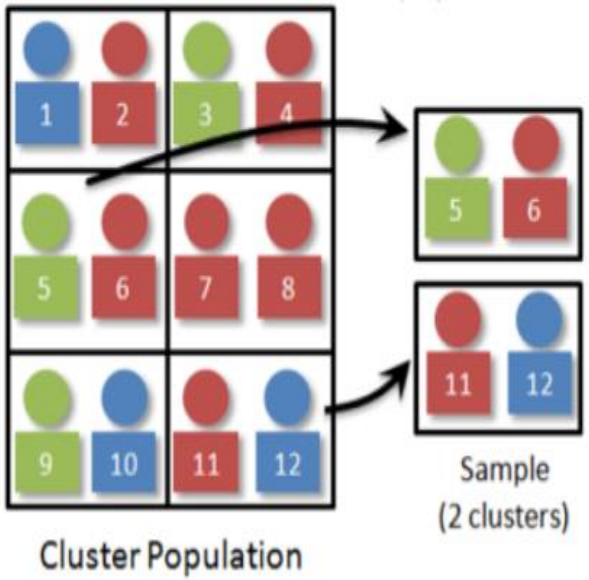
If the number of IT college students is 400 , and the total number of students is 8000, and the sample size to be drawn is 200. then out of the 200 students in the sample a total of (400/8000) * 200 = 10 students should be from IT college.

4. Survey Research

4. Cluster sampling is a sampling method where the researcher creates multiple clusters of people from a population where they are indicative of homogeneous characteristics and have an equal chance of being a part of the sample.

(4) Cluster sampling

- Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample.
- The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, specially in the case of personal interviews.



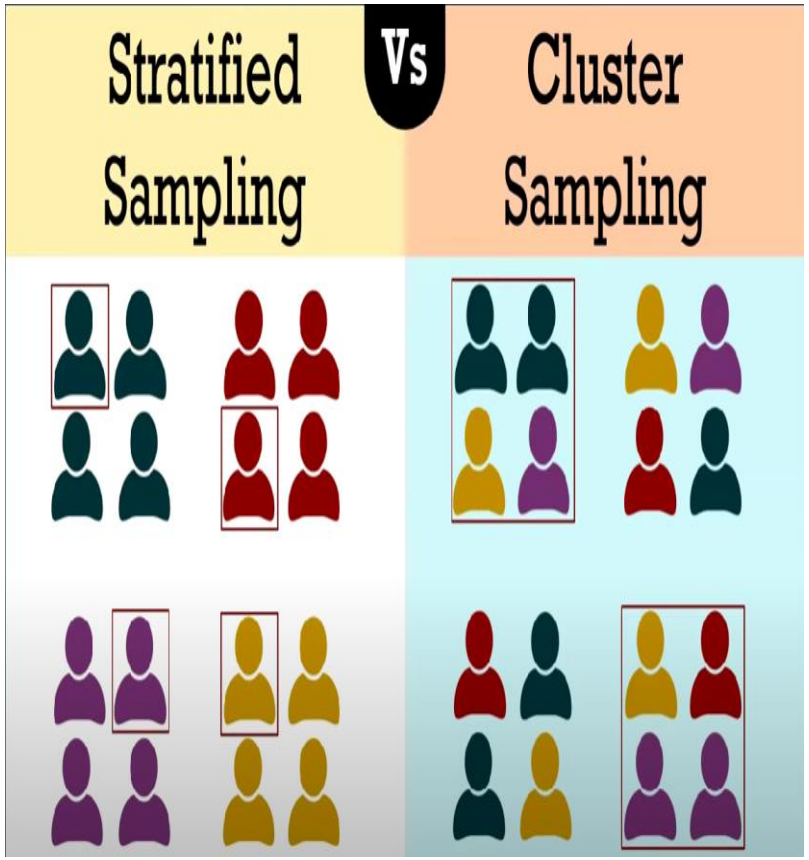
Sample Designs

(4) Cluster sampling

What is the difference between stratified random sampling and cluster sampling ?

Relatedly, in cluster sampling you randomly select entire groups and include all units of each group in your sample. However, in stratified sampling, you select some units of all groups and include them in your sample. In this way, both methods can ensure that your sample is representative of the target population.

See the video in the next slide



Sample Designs

(4) Cluster sampling

Example 1:

https://www.youtube.com/watch?v=1XFU1d9XIWM&t=2s

Example 2 : Consider a scenario where a data organization is looking to survey the performance of smartphones across Germany. They can divide the entire country's population into cities (clusters), select further towns with the highest population, and filter those using mobile devices.

4. Survey Research

Main Types of non-Probability Sampling include:

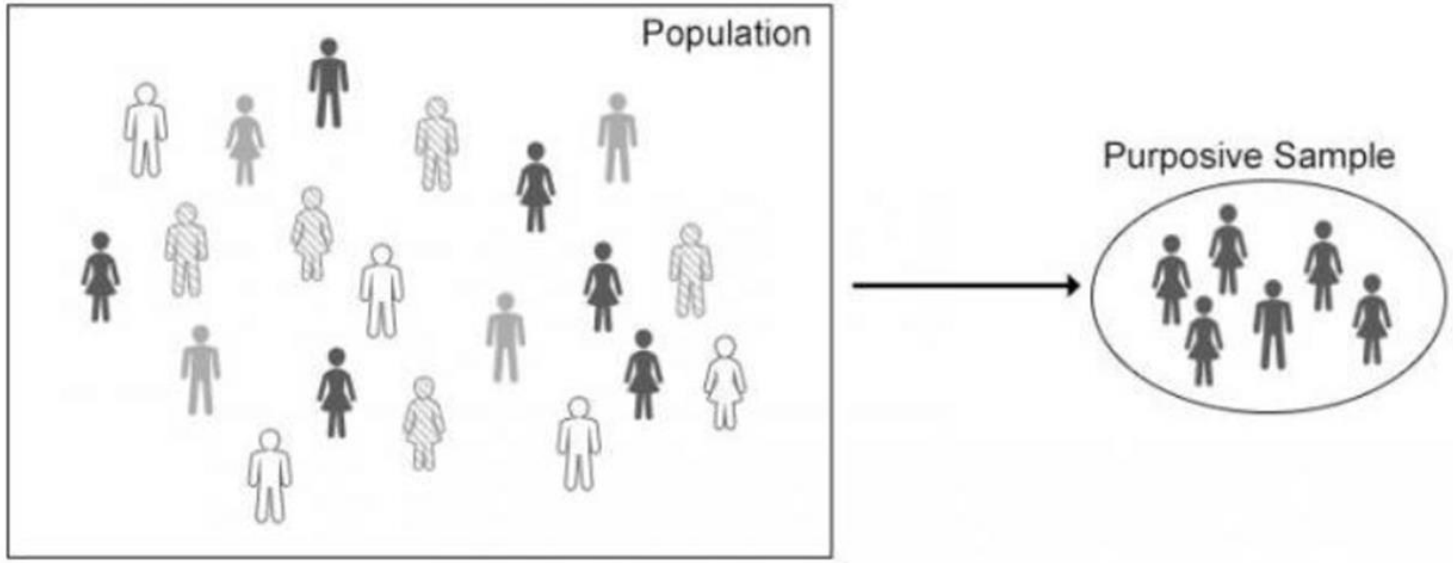
1. Deliberate Sampling.
2. Convenience Sampling.
3. Judgment Sampling.
4. Quota sampling

If the researcher doesn't have the sampling frame , using the random sampling is not allowed and the non-probability sampling should be used

4. Survey Research

(1) Deliberate sampling

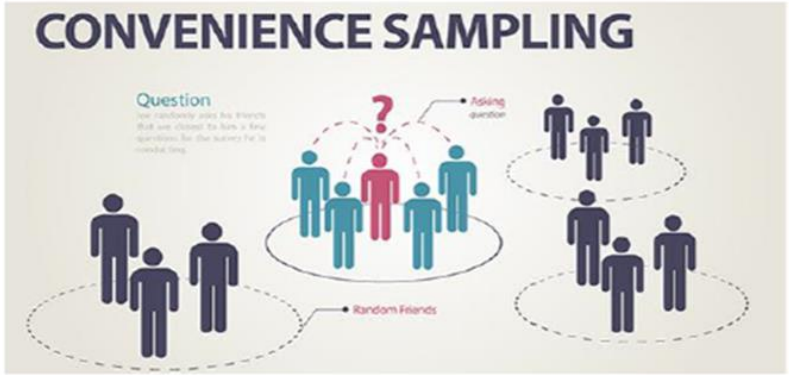
- known as purposive or **non-probability** sampling.
- involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe.



4. Survey Research

(2) *Convenience Sampling*

- it is when population elements are selected for inclusion in the sample based on the ease of access.
- give very biased results particularly when the population is not homogeneous.

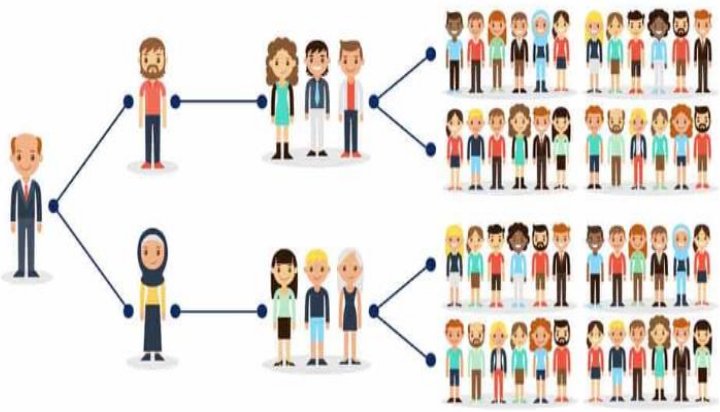


(3) Quota sampling

- interviewers are simply given quota to be filled from different subgroup, the actual selection of items for sample being left to the interviewer's judgement.
- Quota sampling is thus an important form of non-probability sampling.
- Quota samples generally happen to be judgement samples rather than random samples.
- The size of the quota for each subgroup is generally proportional to the size of that subgroup in the population.

4. Survey Research

(4) Snowball Sampling



Snowball sampling or chain-referral sampling is defined as a non-probability sampling technique in which the samples have rare traits. This is a sampling technique, in which existing subjects provide referrals to recruit samples required for a research study.

“Snowball Sampling" reflects an analogy to a snowball increasing in size as it rolls downhill.

4. Survey Research

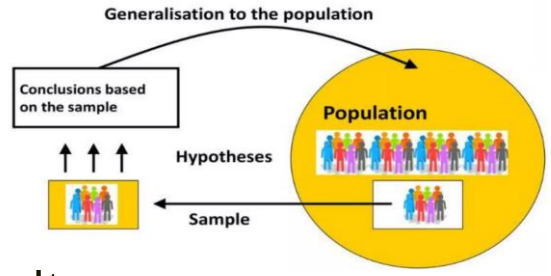
Snowball Sampling:

- **For example,** if you are studying the level of customer satisfaction among the members of an elite country club, you will find it extremely difficult to collect primary data sources unless a member of the club agrees to have a direct conversation with you and provides the contact details of the other members of the club.
- This sampling method involves a primary data source nominating other potential data sources that will be able to participate in the research studies. Snowball sampling method is purely based on referrals and that is how a researcher is able to generate a sample. Therefore this method is also called the chain-referral sampling method.
- Snowball sampling is a popular business study method. The snowball sampling method is extensively used where a population is unknown and rare and it is tough to choose subjects to assemble them as samples for research.

4. Survey Research

SAMPLE SIZE AND ITS DETERMINATION

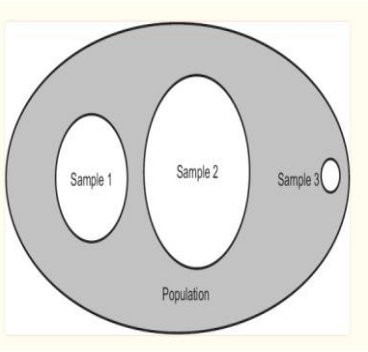
- Sample size: refers to the number of individuals included in a research study as a representative of the total population.
- The sample size is determined from the target study population.
- Sample size determination is the most crucial methodological part of a research study.
- It is essential for an investigator to estimate optimum sample size to produce reliable results
- In sampling analysis, usually ‘N’ refers to the target population, and ‘n’ to the sample.
- If the sample size (‘*n*’) is too small, it may not serve to achieve the objectives. And if it is too large, we may incur huge cost and waste resources.
- As a general rule, one can say that the sample must be **of an optimum size** i.e., it should neither be excessively large nor too small.



4. Survey Research

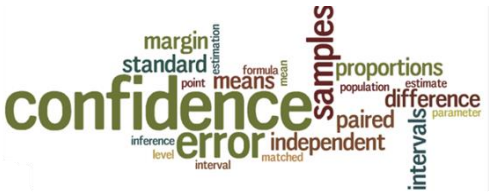
Problems if:

Your sample size is too small….	Your sample size is too big...
It won't be representative of the target population.	It could lead to ethical problems.
The investigation may be considered a waste of time, money or resources.	It may use more time, money and resources than is necessary.
It may not be possible to detect any differences between study groups, or the noted differences won't have enough power to be considered significant.	Statistical testing may be affected.



Concepts that should be understood in sample size determination:

1. **The size of the study population**
2. **Margin of Error:** amount of error you wish to allow in your results (α). (usually 5%)
3. **Confidence Level:** It tells us how confident we are that if a study was repeatedly done, we would get the same results. Usually 90%, 95% & 99%.



4. Survey Research

How can the researcher determine the sample size ?

Sample size can be determined in different ways:

1. Using statistical formulas
2. Using sample size determination' tables
3. Using online sample size calculators

1. Determining the sample size using statistical formulas:

- Basically, there are different statistical formulas for sample size calculations, an example of these is Yamane's Formula :

Yamane's formula

It is used when population is known

$$n= N/(1+Ne^2)$$

Where n= sample size

N=study population

e-error of margin

For a study whose confidence level is 95%, e=5% 90%, e=10% etc.

Example:

Suppose we want to evaluate a program where 1000 medical doctors were encouraged to adopt a new practice; i.e. if Confidence level is 95% and margin of error +/- 5%:

n= N/(1+Ne2)

n= $1000/(1+1000*0.05^2)$

n= 1000/2.5

n=400

So, the sample size that should be collected is equal to 400

4. Survey Research

2. Determining the sample size using tables

- Many statisticians have developed tables to determine the required sample size based on population size, precision rate and confidence intervals like Krejcie and Morgan Table (1970).
- These tables were calculated using the statistical formulas, so they are reliable and will produce similar results like formulas.
- However, using the table is much simpler than employing a formula.

Required Sample Size†

Population Size	Confidence = 95% Margin of Error				Confidence = 99% Margin of Error			
	5.0%	3.5%	2.5%	1.0%	5.0%	3.5%	2.5%	1.0%
10	10	10	10	10	10	10	10	10
20	19	20	20	20	19	20	20	20
30	28	29	29	30	29	29	30	30
50	44	47	48	50	47	48	49	50
75	63	69	72	74	67	71	73	75
100	80	89	94	99	87	93	96	99
150	108	126	137	148	122	135	142	149
200	132	160	177	196	154	174	186	198
250	152	190	215	244	182	211	229	246
300	169	217	251	291	207	246	270	295
400	196	265	318	384	250	309	348	391
500	217	306	377	475	285	365	421	485
600	234	340	432	565	315	416	490	579
700	248	370	481	653	341	462	554	672
800	260	396	526	739	363	503	615	763
1,000	278	440	606	906	399	575	727	943
1,200	291	474	674	1067	427	636	827	1119
1,500	306	515	759	1297	460	712	959	1376
2,000	322	563	869	1655	498	808	1141	1785
2,500	333	597	952	1984	524	879	1288	2173
3,500	346	641	1068	2565	558	977	1510	2890
5,000	357	678	1176	3288	586	1066	1734	3842
7,500	365	710	1275	4211	610	1147	1960	5165
10,000	370	727	1332	4899	622	1193	2098	6239
25,000	378	760	1448	6939	646	1285	2399	9972
50,000	381	772	1491	8056	655	1318	2520	12455
75,000	382	776	1506	8514	658	1330	2563	13583
100,000	383	778	1513	8762	659	1336	2585	14227
250,000	384	782	1527	9248	662	1347	2626	15555
500,000	384	783	1532	9423	663	1350	2640	16055
1,000,000	384	783	1534	9512	663	1352	2647	16317
2,500,000	384	784	1536	9567	663	1353	2651	16478
10,000,000	384	784	1536	9594	663	1354	2653	16560
100,000,000	384	784	1537	9603	663	1354	2654	16584
300,000,000	384	784	1537	9603	663	1354	2654	16586

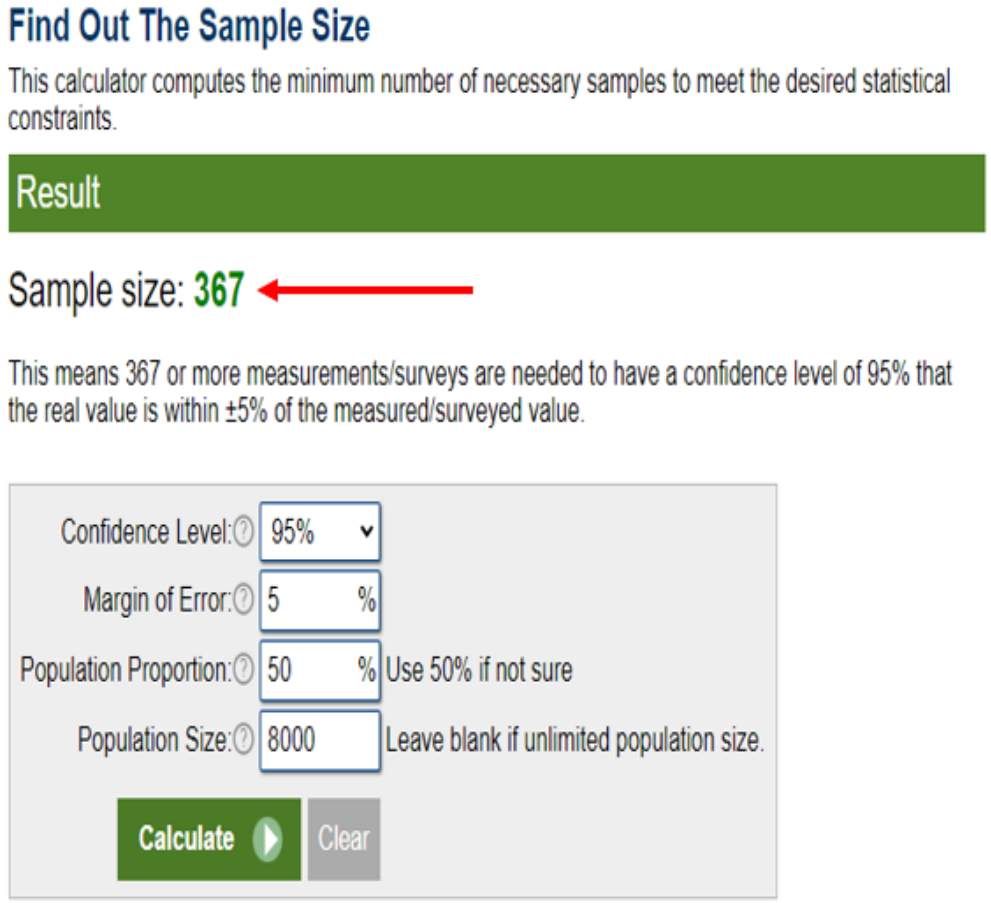
4. Survey Research

3. Determining the sample size using online calculators

Many websites provide sample size calculations such as https://www.calculator.net/sample-size-calculator.html , or http://www.raosoft.com/ and many others.

Here, the researcher have to enter the population size, the confidence interval level which is usually 95% and margin of error (usually 5%) and press calculate.

For example, if a researcher want to conduct a study on the students of PTUK (population = 8000), then he should collect a sample of 367 responses to have a representative sample



4. Survey Research

How to choose the proper sampling technique?

For any research, it is essential to choose a sampling method accurately to meet the goals of your study. The effectiveness of your sampling relies on various factors. Here are some steps expert researchers follow to decide the best sampling method.

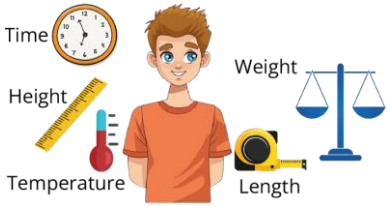
- Jot down the research goals. Generally, it must be a combination of cost, precision, or accuracy.
- Identify the effective sampling techniques that might potentially achieve the research goals.
- Test each of these methods and examine whether they help achieve your goal.
- Select the method that works best for the research.

Research Design Variables, Part II

Chapter 4: Research Design Part II

- What is Measurement in Research
- Primary scales of measurement
  - ✓ Nominal Scale
  - ✓ Ordinal Scale
  - ✓ Interval Scale
  - ✓ Ratio Scale
- Sources of Error in Measurement
- Tests of Sound Measurement
  - ✓ Validity
  - ✓ Reliability
  - ✓ Practicality
- Quantitative Research Methods in CS and IS
- Mixed Methods
- Methodologies for Software Development

What is Measurement in Research?



After studying and understanding the different types of quantitative research methods, sampling design, and selecting the proper type of sampling for the study in hand, the researcher has to decide and select the measurement of the study variables.

- **Measurement** is the process of describing some property of a phenomenon under study and assigning a numerical value to it. Measurement requires a scale.
- **A scale** provides a range of values—a yardstick—that corresponds to the presence of the properties of the concept under investigation.
- A scale provides the rules that associate values on the scale to the concept we are studying.
- Properties like weight and height can be measured directly with some standard unit of measurement
- However, it is not that easy to measure properties like motivation to succeed, ability to stand stress and the like

In these cases, the researcher uses predefined and tested scales for these properties.

The Meaning of Scaling

We can measure concepts through direct and indirect observations:

1. **Direct Observation:** We can measure someone's weight or height. And, we can record the color of their hair or eyes.
2. **Indirect Observation:** We can use a questionnaire in which respondents provide answers to our questions about gender, income, age, attitudes, and behaviors.

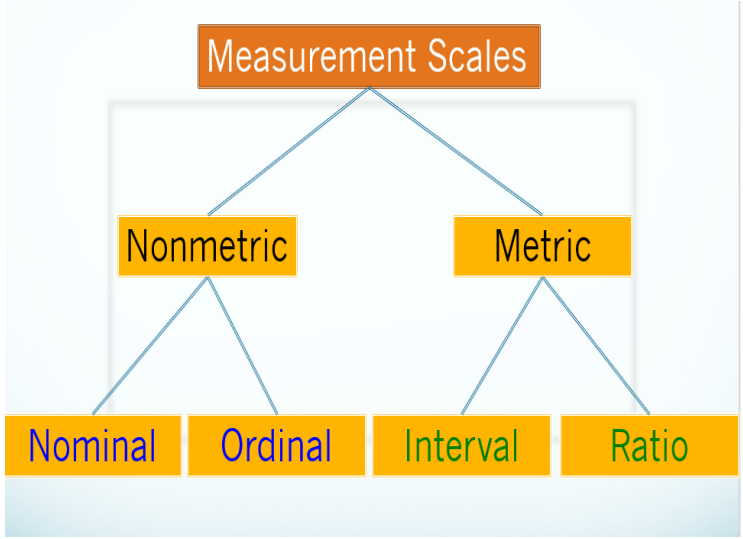
هناك بعض المتغيرات يمكن قياسها كمياً مثل حجم أو مدى التعرض والمشاهدة للتلفزيون بتحديد عدد الساعات والدقائق، أو مرات المشاهدة، وهناك بعض المتغيرات لا يمكن قياسها إلا من خلال مؤشرات تدل عليها، مثلاً لا يمكن قياس المنافع التي يحققها الفرد من قراءة الصحف إلا من خلال مقياس يتضمن مؤشرات متوقعة من القراءة، ويجيب الفرد بالموافقة أو عدم الموافقة على عبارات تعتبر مؤشرات عن هذه المنافع.

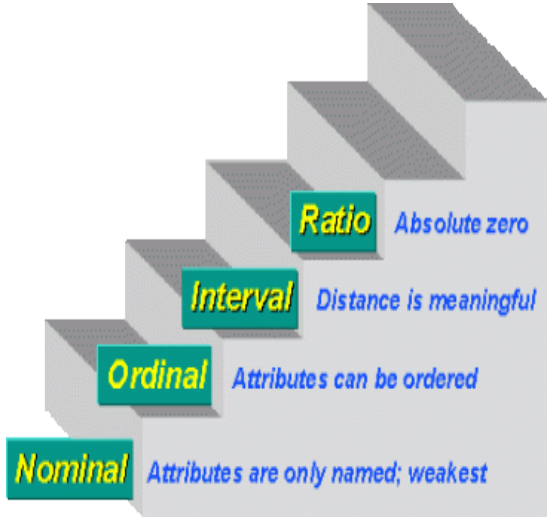
Primary scales of measurement

- Scales of measurement can be considered in terms of their mathematical properties.
- The most widely used classification of measurement scales are:

  (a) nominal scale; (b) ordinal scale;

  (c) interval scale; and (d) ratio scale.
- The First two are nonmetric the other two are metric.





Measurement in Research

2. Nominal scale

- Are numerical in name only, because they do not share any of the properties of the numbers.
- For instance if we record marital status as 1, 2, 3, or 4 for (single, married, widowed or divorced respectively),
  - we cannot write 4 > 2 or 3 < 4 ; and we cannot write 3 – 1 = 4 – 2, 1 + 3 = 4 or 4 / 2 = 2.
  - أي انه قد تم استخدام الأرقام للتمييز فقط ولا يمكن اجراء حسابات او مقارنات بينها / مثل ارقام الشعب ش1 و ش2 او ارقام الأسئلة في الامتحانات س1 او س2 وهكذا
  - Common examples of nominal data include:
    - Student registration numbers at their college or university
    - Numbers assigned to football players or jockeys in a horse race
    - Numbers assigned to gender, marital status, nationality, etc …

Examples

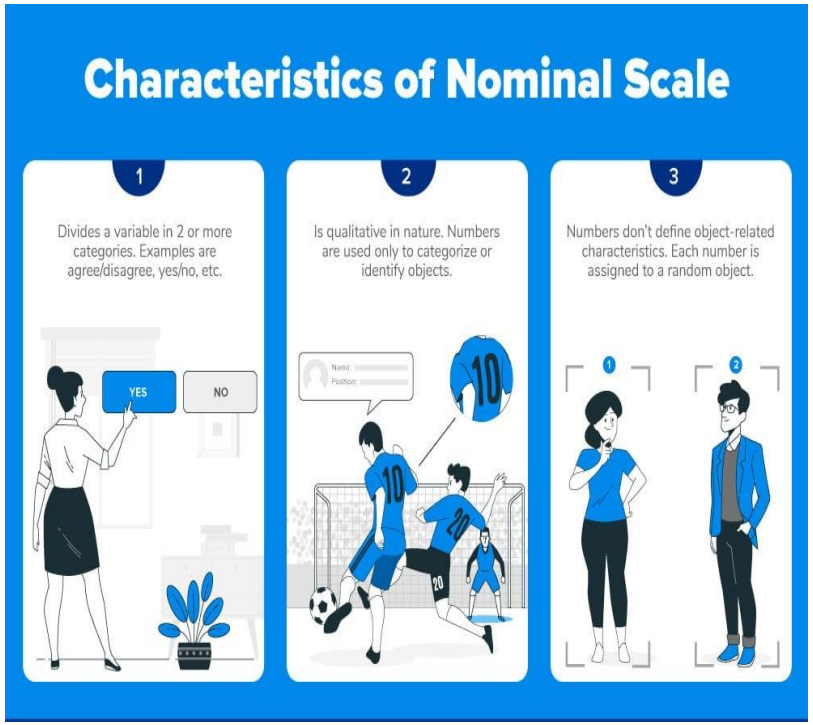
Gender ( ) Female ( ) Male

Marital Status ( ) Married ( ) Single( ) Divorced

Nationality ( ) Turk ( ) Arab ( ) Germen ( ) English

Characteristics of Nominal Scale:

- In nominal scale a variable is divided into two or more categories, for example, agree/disagree, yes or no etc. It's is a measurement mechanism in which answer to a particular question can fall into either category.
- Nominal scale is qualitative in nature, which means numbers are used here only to categorize or identify objects.
- In nominal scale, numbers don't define the characteristics related to the object, which means each number is assigned to one object. The only permissible aspect related to numbers in a nominal scale is "counting."



Measurement in Research

2. Ordinal Scale

- An ordinal scale is employed when data may be logically **ranked** or **ordered**, but the distinctions between the values are **unknown or unequal**. It is frequently utilized in surveys, rating systems, and ranking exercises.
- An ordinal scale is a ranking scale in which numbers are assigned to objects to indicate the relative extent to which the objects possess some characteristic. An ordinal scale allows you to determine whether an object has more or less of a characteristic than some other object, but not how much more or less.
- Thus, an ordinal scale indicates relative position, **not** the magnitude of the differences between the objects.
- Grocery stores generally rank their hot sauce as, “mild”, “medium”, and “spicy”. We know which one is the hottest and which one is less hot, but we don’t know whether the increase in hotness from “mild” to “medium” is the same as it is from “medium” to “spicy.”

Measurement in Research

- An Example:

If we have four minerals: gypsum, fluorite, quartz, and diamond. If one mineral can scratch another, it receives a higher hardness number. Then theses will be assigned the numbers from 1 to 4 respectively. Diamond will have number 4 since it has the highest hardness among all, whereas gypsum is assigned the value of 1.

- The numbers we can write 4>1 as diamond is harder than gypsum.
- We can write 2<4 as fluorite is softer than diamond.
- But we cannot write for example 3-1=4-2 because the difference in hardness between are not the same (we don't know the magnitude of extra hardness).

1

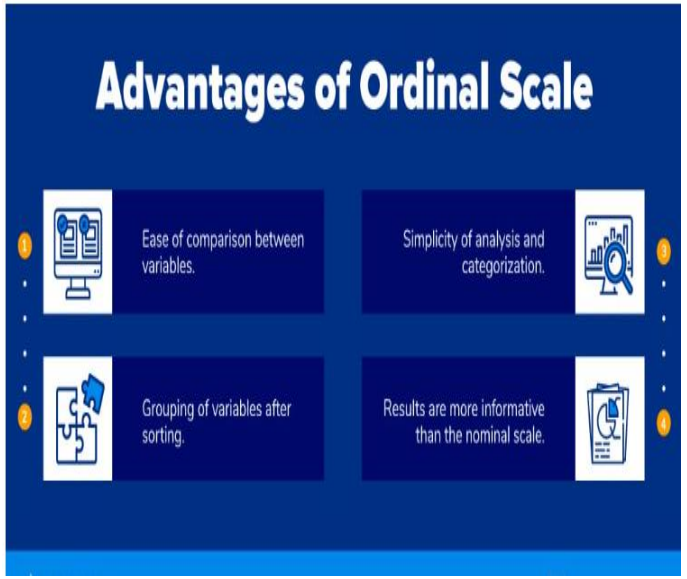
2

3

4

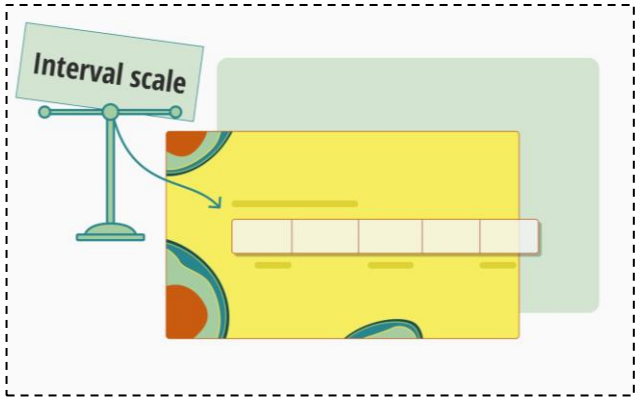
Characteristics of Ordinal Scale (Advantages):

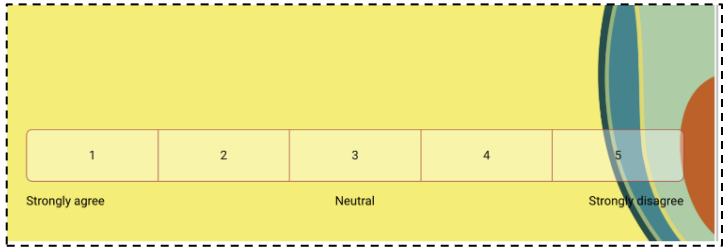
1. The primary advantage of using ordinal scale is the ease of comparison between variables.
2. Extremely convenient to group the variables after ordering them.
3. Effectively used in surveys, polls, and questionnaires, due to the simplicity of analysis and categorization. Collected responses are easily compared to draw impactful conclusions about the target audience.
4. As the values are indicated **in a relative manner** using a linear rating scale the results are more informative than the nominal data.

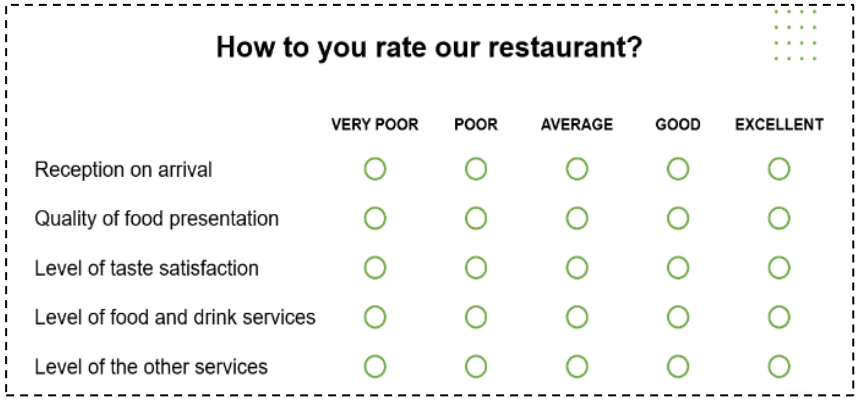


3. Interval scale

- In an interval scale, numerically equal distances on the scale represent equal values in the characteristic being measured because it is metric.
- When in addition to setting up inequalities we can also form differences. An interval scale contains all the information of an ordinal scale, but it also allows you to compare the differences between objects.
- The difference between any two scale values is **identical** to the difference between any other two adjacent values of an interval scale.
- There is a constant or equal interval between scale values. The difference between 1 and 2 is the same as the difference between 2 and 3, which is the same as the difference between 5 and 6
- Likert scale is an example of interval scale in research

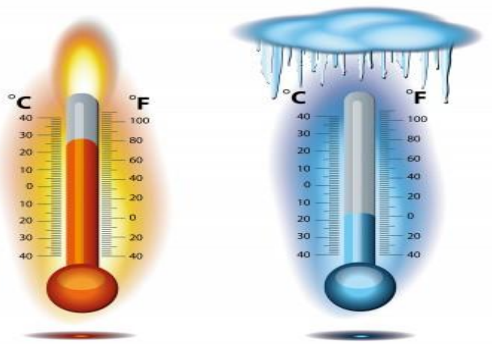






Example 1 on Interval Scale:

- A common example in everyday life is a temperature scale.
- Suppose we are given the following temperature readings 58°, 63°, 70°, 95°, 110°, 126° and 135°.
- In this case, we can write 110° > 70° or 95° < 135° which simply means that 110° is warmer than 70° and that 95° is cooler than 135°.
- We can also write for example 95° – 70° = 135° – 110°, si
temperature differences are equal in the sense.



Thermometer

Example 2:

As a Palestinian, I am proud of my country.

(5) Strongly agree

(4) Agree

(3) Neither agree nor disagree

(2) Disagree

(1) Strongly disagree

- Interval is a metric scale, so it is possible to calculate the arithmetical means of interval scale.
- Let us assume that 200 citizens have answered the question above as shown in the adjacent table:

Participation Degree	Frequency
5 Strongly agree	75
4	64
3	38
2	18
1 Strongly disagree	5
Total	**200**

- to calculate the weighted mean which is the mean of all their responses we use the following weighted mean equation:

$$\bar{x} = \frac{w_1x_1 + w_2x_2 + \ldots\ldots + w_nx_n}{w_1 + w_2 + \ldots\ldots w_n}$$

Where,

x
is the repeating value
w
is the number of occurrences of
x
weight
$\bar{x}$
is the weighted mean

The weighted arithmetic means is = (5*75) + (4*64) + (3*38) + (2*18) + (1*5) / 200 = 3.93

Interpretation of the result

To interpret the result we have to compare the resulting weighted average with a standard value. We create five intervals as follows:

Here, we have 5 anchors, the largest value is 5 and the lowest value is 1, so :

- Range of the value = max value – min value = 5-1= 4
- Interval length = range / number of values = 4 / 5 = 0.8

then we can create the following scale:

Weighted average	Result	Result Interpretation
1 – 1.79	Strongly disagree	Very uninfluential
1.80 – 2.59	Disagree/	Uninfluential
2.60 – 3.39	Neutral	Neutral or do not know
3.40 – 4.19	Agree	Influential
4.20 – 5	Strongly agree	Very influential

The arithmetical means (the average) of the population is 3.93 which means that the population agree that they are proud of their country to an influential level. (درجة كبيرة)

Weighted average	Result	Result Interpretation
1 – 1.79	Strongly disagree	Very uninfluential
1.80 – 2.59	Disagree/	Uninfluential
2.60 – 3.39	Neutral	Neutral or do not know
3.40 – 4.19	Agree	Influential
4.20 – 5	Strongly agree	Very influential

3.93

- Interval scale can be five, seven or nine-level scale.

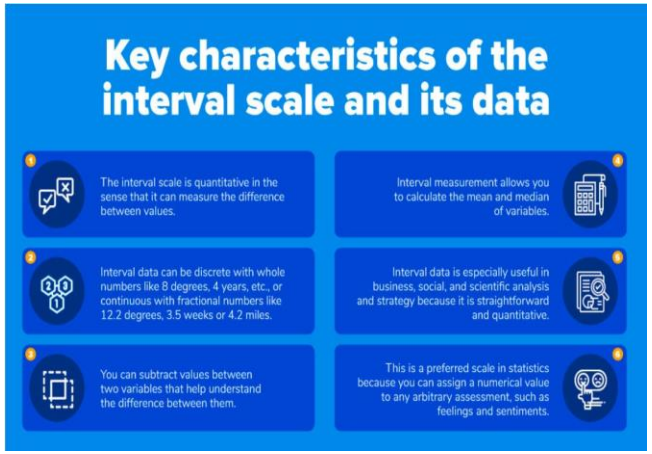
Strong positive	Positive	Neutral	Negative	Strong negative
(1)	(2)	(3)	(4)	(5)

Strongly agree	Agree	More or less agree	Undecided	More or less disagree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)	(5)	(6)	(7)

An example of a 5-point and 7-point Likert Scale.

Characteristics of Interval Scale :

1. The interval scale is preferred to nominal scale or ordinal scale because the latter two are qualitative scales. The interval scale is quantitative in the sense that it can quantify the difference between values.
2. Interval data can be discrete with whole numbers like 8 degrees, 4 years, 2 months, etc., or continuous with fractional numbers like 12.2 degrees, 3.5 weeks or 4.2 miles.
3. You can subtract values between two variables that help understand the difference between two variables.
4. Interval measurement allows you to calculate the mean and median of variables.
5. Interval data is especially useful in business, social, and scientific analysis and strategy because it is straightforward and quantitative.



Characteristics of Interval Scale :

6. Interval scale **still lacks** one property. **It lacks a true zero.**

In other words, a value of zero does not indicate the absence of the attribute being measured; it simply represents an arbitrary point on the scale. For example, 0°C on the Celsius scale doesn't mean the absence of temperature; it's just a point chosen as the freezing point of water.
Because of the absence of a true zero, operations such as multiplication and division are not meaningful on an interval scale. For instance, if we have temperatures of 20°C and 10°C, it wouldn't make sense to say that 20°C is "twice as hot" as 10°C because there's no true zero point to anchor such comparisons.

(**Zero cant be used as a reference point**)

4. Ratio scale

- Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples.
- 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).
- Ratio data includes all the usual measurement (or determinations) of length, height, money amounts, weight, volume, area, pressures etc.

Characteristics of Ratio Scale:

1.Ratio scale, as mentioned earlier **has an absolute zero** characteristic. It has orders and equally distanced value between units. The zero point characteristic makes it relevant or meaningful to say, “one object has twice the length of the other” or “is twice as long.” (A zero value indicates the absence of this characteristic)

2.Ratio scale **doesn’t have a negative number,** unlike interval scale because of the absolute zero or zero point characteristic. To measure any object on a this scale, researchers must first see if the object meets all the criteria for interval scale plus has an absolute zero characteristic. (so it is used for variables that cant be negative)

3.Ratio scale provides unique possibilities for statistical analysis plan. In this scale, **variables can be systematically added, subtracted, multiplied and divided (ratio).** All statistical analysis including mean, mode, the median can be calculated using it.

3. Ratio scale **has units** which have several unique and useful properties. One of them is they allow unit conversion. Take an example of calculation of energy flow. Several units of energy occur like Joules, gram-calories, kilogram-calories, British thermal units. Still more units of energy per unit time (power) exist kilocalories per day, liters of oxygen per hour, ergs, and Watts.

Examples of ratio scale questions

- **Miles per hour:** Speed is an easy example of the ratio scale. Because there is a true zero value—the absence of forward motion—it's easy to plot and visualize speed data
- **Time:** Time spent can be measured on a ratio scale, since "negative time" doesn't exist. If you're researching how long it takes customers to make a purchase on your website, you would measure from the moment they initiate the purchasing process to the final "place your order" command. (the second they entered, is the reference point)
- **Weight:** Weight also has a true zero, making it appropriate for ratio scales. Whether you're measuring in milligrams, kilograms, ounces, or pounds, a ratio scale allows you to understand the relationship between different weights.
- **Age:** Age is one of the most common attributes that can be plotted on ratio scales. Companies often use this in market research and customer demographic studies. Because no one can be younger than zero, you'll know the true value of each age group, and how they relate to others. For example, 40-year-olds are twice as old as 20-year-olds, but half the age of 80-year-olds.

Examples of ratio scale questions

The following are the most commonly used examples:

1. What is your height in feet and inches?

- Less than 5 feet.
- 5 feet 1 inch – 5 feet 5 inches
- 5 feet 6 inches- 6 feet
- More than 6 feet

2. What is your weight in kgs?

- Less than 50 kgs
- 51- 70 kgs
- 71- 90 kgs
- 91-110 kgs
- More than 110 kgs

3. How much time do you spend daily watching television?

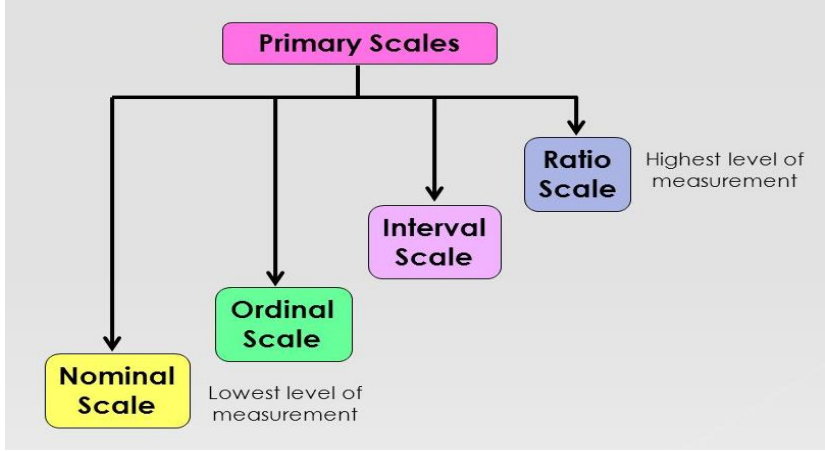
- Less than 2 hours
- 3-4 hours
- 4-5 hours
- 5-6 hours
- More than 6 hours

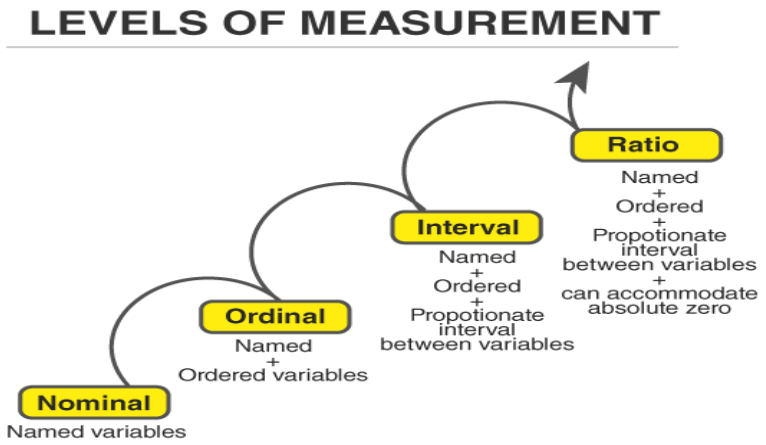
To summarize the scales of measurement, here are their properties.

Level of Measurement	Can categorize data?	Can order/rank data?	Can add or Subtract data?	True or Meaningful Zero?	Preciseness
Nominal	Yes	No	No	No	Least precise
Ordinal	Yes	Yes	No	No	Little precise
Interval	Yes	Yes	Yes	No	Precise
Ratio	Yes	Yes	Yes	Yes	Most precise

الرقم	نوع المقياس	الهدف منه (خصائص الأرقام)	أمثلة	وحدة القياس	حالة الصفرفي المقياس
1	الاسمي	التميز بين الأشياء	الذكر 1 الأنثى 2	لا يوجد	لا تنعدم لسمة المقاسة بوجود الصفر لأنه ليس حقيقي
2	الترتيبي (الرتب)	التمييز والترتيب بفروقات ليس شرط أن تكون متساوية لسمة المقاسة	1 ممتاز 2 جيد جدا 3 جيد	لا يوجد	
3	الفئوي (المسافة)	التمميز والترتيب بفروقات متساوية لسمة المقاسة ويمكن استخدام المتوسط الحسابي والإنحراف المعياري	قياس درجات الحرارة المئوية	يوجد وحدة قياس	
4	النسبي	أفضل أنواع المقاييس المعروفة لأنه يحتوي على جميع ما سبق بالإضافة إلى نقطة صفر مطلق و تستخدمه معظم العلوم الطبيعية	قياس الطول قياس الوزن	يوجد	الصفر هنا حقيقي وتنعدم السمة المقاسة بوجود الصفر

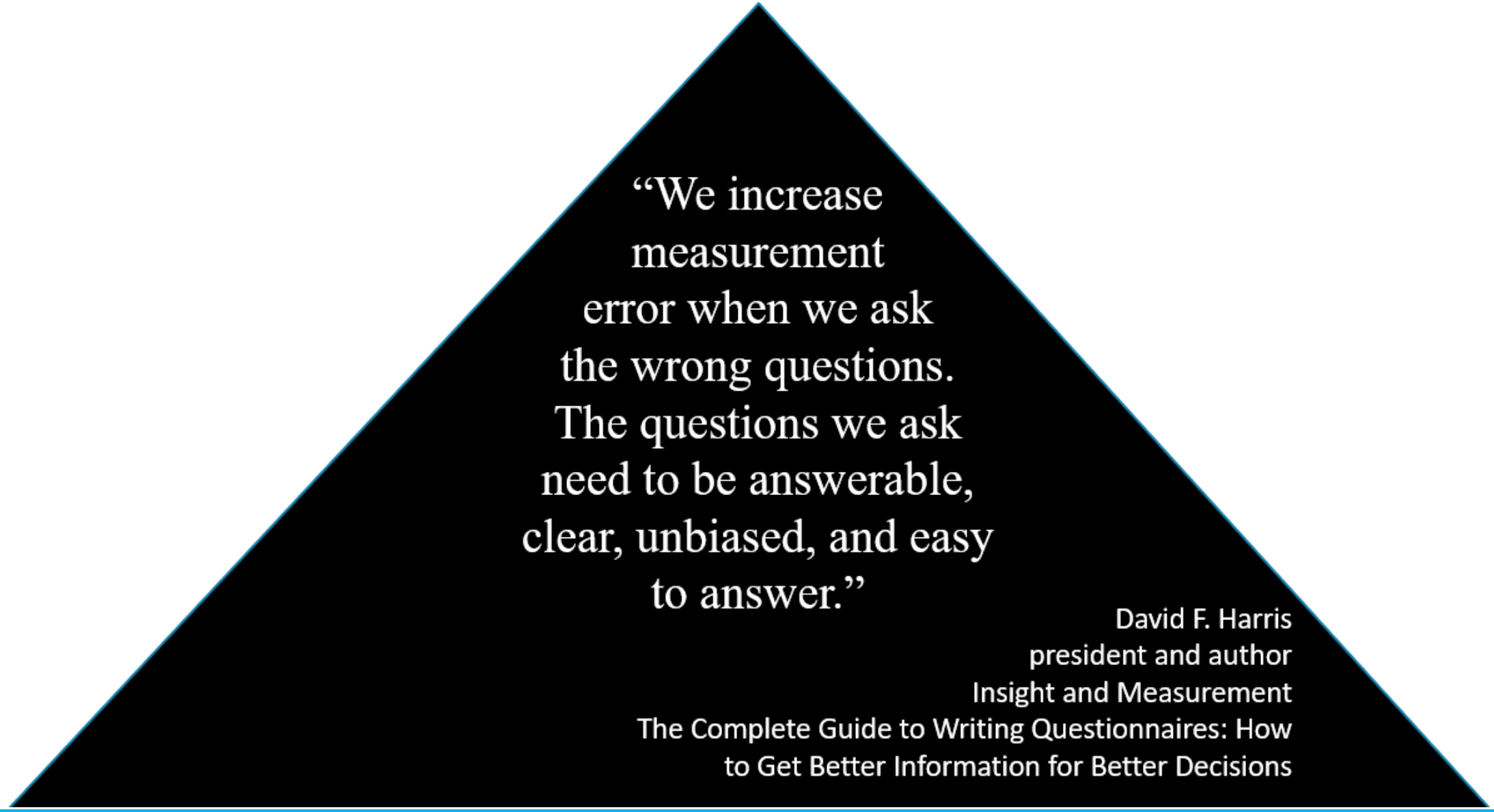
Levels of Measurement





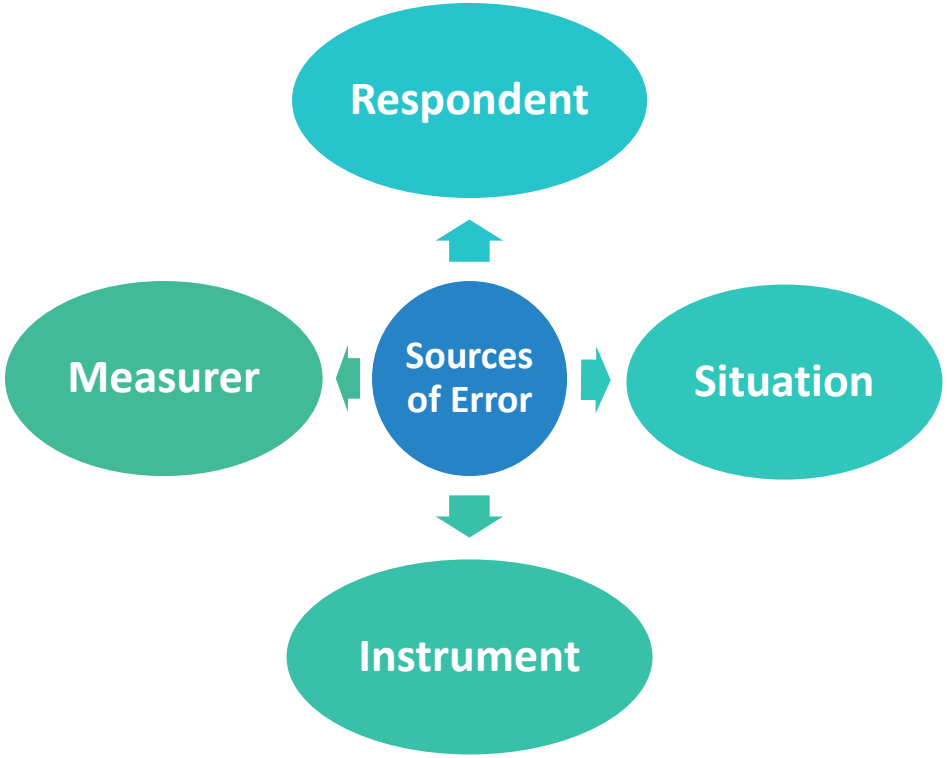
Nominal	Classification	
Ordinal	Classification	
	Order	
interval	Classification	Distance
	Order	
Ratio	Classification	Distance
	Order	Natural Origin

Issues with data collection and format



Sources of Error in Measurement

Measurement should be precise and unambiguous in an ideal research study. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following are the possible sources of error in measurement



Sources of Error in Measurement

1. The Respondent

- Is the Person who provide the response or fill the questionnaire
- May feel reluctant in expressing strong negative or extreme feelings.
- Not admitting ignorance then guessing
- Transient factors: boredom, fatigue, anxiety...etc. lead to inaccurate or incomplete response

2. The Situation

- Is the Conditions around the respondent
- Condition which places a strain on interview.

  For instance, if anonymity or privacy.
- Uncomfortable situation
- Presence of your boss, Camera…

Sources of Error in Measurement

3. The Measurer

- Is the Data collector, Data encoder, Data analyst,
- Error in coding, tabulation or statistical calculations.
- Rewording or reordering questions can distort the response.
- The overall looks of interviewer may encourage or discourage.

4. The Instrument

- It is the tool for data collection (questionnaire)
- Defective measuring instrument/questioners.
- Complex words, poor printing, response choice omissions.
- Difficult to understand
- Ambiguous meaning
- Inadequate space for the response
- Not having clear instruction, lack of logical sequence

Tests of Sound Measurement (https://www.youtube.com/watch?v=JaVkldof8l4)

- It is an important issue in research.
- It must correctly measure the concepts so that the conclusions & decision will be accurate.
- To understand business research, or really any concept, we must be able to measure it.

A concept is an abstraction or idea formed by the perception of phenomena like awareness, enjoyment, ease of use …

Tests of Sound Measurement

- In research we have to ensure that we are using the correct measure for measuring the variables or concepts.
- Sound measurement must meet the tests of Validity, Reliability and Practicality.

Tests of Sound Measurement

1. Validity (الصدق)

- Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure.

Tests of Sound Measurement

- For example, if you want to measure the mass of an abject you use a weighing scale because it gives you the desired measure. **It measures the mass**
- **أي ان المقياس المستخدم يقيس ما صمم لقياسه**

Tests of Sound Measurement

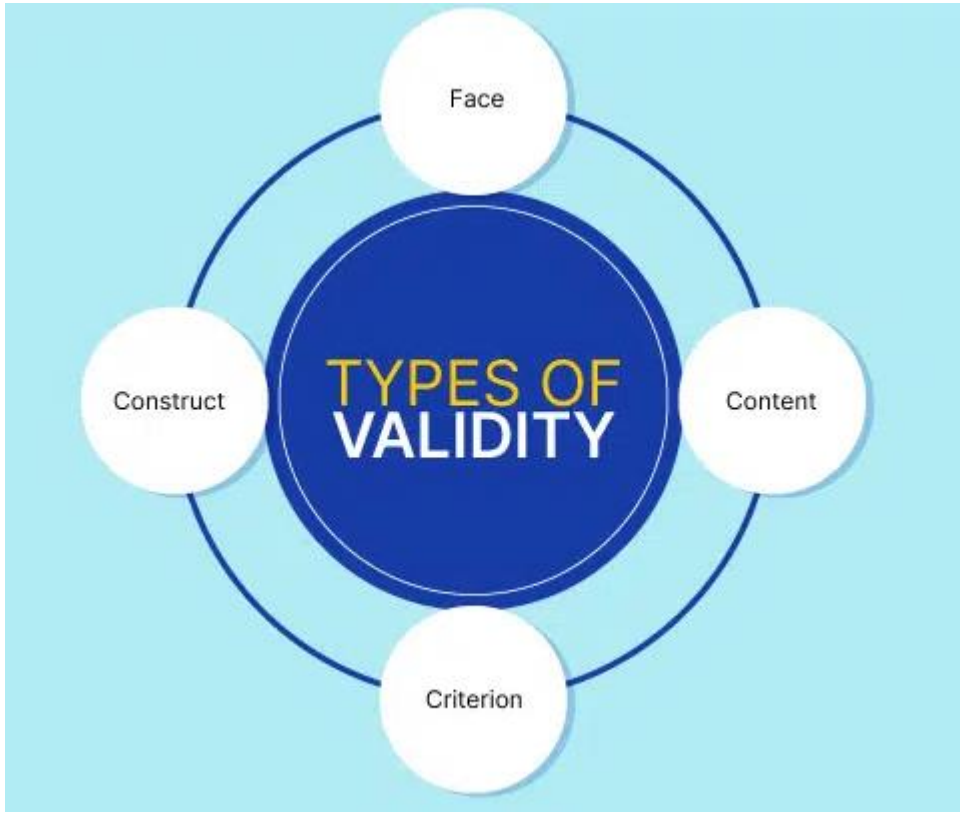
- If you want to measure the temperature, you use a thermometer.

  Which is Valid.

- But you cant use the thermometer to measure air pressure since it does not measure it. It is invalid to measure the air pressure

Tests of Sound Measurement

There are four types of validity tests



Tests of Sound Measurement

1. **Content validity (صدق المحتوى):** is the extent to which a measuring instrument provides adequate coverage of the topic under study (يجب ان نقيس كل الجوانب).

Example:

1. The exam would have high content validity if the questions asked cover every possible topic in the course
2. A mobile phone company wants to conduct a customer satisfaction survey about a specific cell phone model they recently launched. The survey should have questions about the product such as – features, quality, performance, color, design, price, etc. This means that to protect the validity of the result, the content of the survey must cover all relevant questions about the product.

Tests of Sound Measurement

2. Construct validity (صدق المفهوم): is about how well a test measures the concept it was designed to evaluate.

Example:

In a test measuring intelligence, if individuals who score higher on the test also perform better academically or are perceived as more intelligent by experts, it indicates good construct validity. In simpler terms, construct validity confirms that a test or measurement tool is indeed capturing what it claims to measure, based on its alignment with other established measures or theoretical frameworks.

Tests of Sound Measurement

3. Face validity (الصدق الظاهري): considers how suitable the content of a test seems to be on the surface. It's similar to content validity, but face validity is a more informal and subjective assessment.

it's often considered the weakest form of validity. However, it can be useful in the initial stages of developing a method.

Example:

1. (مثل تحكيم المحكمين للاستبيان)
2. A questionnaire designed to measure customer satisfaction at a restaurant. If the questions in the questionnaire directly ask about aspects such as food quality, service speed, and overall dining experience, without any complex or ambiguous language. On its surface, the survey seems like a good representation of what you want to test, so you consider it to have high face validity.

Tests of Sound Measurement

4. Criterion validity (صدق المعيار): How well the results of your test approximate the results of another test.

Example:

1. A university professor creates a new test to measure applicants' English writing ability. To assess how well the test really does measure students' writing ability, she finds an existing test that is considered a valid measurement of English writing ability, and compares the results when the same group of students take both tests. If the outcomes are very similar, the new test has high criterion validity. face validity.

Tests of Sound Measurement

2. Reliability

Reliability is the degree to which a scale is able to produce consistent results when repeated measurements are taken under the same situation.

- This simply means that if the results appear similar again and again, provided that the condition is the same, then we refer to the measurement as reliable.

(اي انه لو تم إعادة الدراسة تحت ظروف مشابهة ستعطي نفس النتائج)

Tests of Sound Measurement

2. Reliability

Example:

When you weigh the same quantity of rice on a balance repeatedly with the result remaining the same all through the measurement process, you refer to this weight as reliable. This also means that the quantity neither decrease nor increase on being measured multiple times.


Reliable
Not Valid
Low Validity
Low Reliablity
Not Reliable
Not Valid
Both Reliable
and Valid

Tests of Sound Measurement

3. Practicality

- Practicality means that the test is easy to design, easy to administer and easy to score.
- Practicality is concerned with a wide range of factors of economy, convenience, and interpretability

Reliability in Experimental Computer Science

Reliability is a critical aspect of any experimental research. Therefore, to ensure reliability, researchers must focus on several key areas.

- **Replication:** Repeating the same experiment under the same conditions verifies consistency and reliability.
- **Consistency in Procedures**: All aspects of the research process should be standardized to ensure equal treatment of all participants.
- **Accurate and Consistent Data Collection**: Valid and reliable measurement tools should be used to ensure consistent data recording and analysis.
- **Statistical Tests:** Test-retest reliability method can be used to assess data reliability.
- **Reliability:** Reliability should be monitored throughout the research process, ensuring consistency and repeatability of results.

Qualitative research in CS and IS

Focus of Research

Qualitative
- Understanding
- Interpretation

Quantitative
- Description
- Explanation

- Qualitative research involves the use of qualitative data, such as interviews, documents, and participant observation, to understand and explain a specific phenomena. It focuses on exploring **how** and **why** certain outcomes occur rather than just measuring them.
- Qualitative research techniques include:
  - **Most commonly used**: individual depth interviews (IDIs) and group interviews (focus group), Case studies, and action research. **(breiefly covered covered in this course)**
  - **Less commonly used**: ethnography, phenomenology, hermeneutics, and grounded theory **(Not Covered in this course)**

Example: In User Experience Research, studying how users interact with a new software application to identify usability issues and improve interface design.

Qualitative research in CS and IS

1. Individual Interviews

Individual Interviews are a qualitative research method where a researcher engages one participant at a time to explore their perspectives, experiences, or insights. These interviews are often open-ended or semi-structured, allowing for in-depth discussions and personalized responses.

Advantages	Disadvantages
In-depth Data: Interviews allow for detailed exploration of participants' thoughts, providing rich data.	**Time-Consuming:** Conducting, transcribing, and analyzing interviews takes a significant amount of time and effort
Flexibility: The interviewer can adjust questions or probe deeper based on the participant's responses.	**Limited Generalizability:** Insights are often specific to individual participants
Personalized Insights: Tailored questions can be asked to suit the specific participant's expertise	**Subjectivity:** Responses are influenced by personal biases or interpretations
Confidentiality: Participants might feel more comfortable sharing sensitive information	**Interviewer Bias:** The interviewer's own views or manner of questioning may unintentionally influence participants' responses

Example: Conducting interviews with system administrators to learn about their strategies and challenges in maintaining cybersecurity in cloud environments.

Qualitative research in CS and IS

2. Group Interviews

Group Interviews, also known as **focus groups**, are a qualitative method where multiple participants engage in a structured or semi-structured discussion led by a facilitator. The goal is to gather collective insights, ideas, or experiences from a group of individuals on a particular topic.

Advantages	Disadvantages
Interaction and Synergy: Participants build on each other's ideas, leading to richer data and diverse perspectives.	**Dominant Voices:** Some participants may dominate the conversation, leading to a skewed dataset and the suppression of quieter voices.
Efficient Data Collection: Multiple participants are interviewed at once, making it more time-efficient	**Limited Depth:** May not allow for the same level of in-depth exploration as IDIs, as time is divided among participants.
Observing Group Dynamics: Allows researchers to observe how individuals interact and discuss topics in a group setting, which can reveal important dynamics.	**Groupthink:** Participants may conform to the group's opinion rather than share their true thoughts, leading to less authentic responses.
Emerging Ideas: New ideas or issues may emerge during the group discussion that individual participants might not have considered on their own.	**Complexity in Analysis:** Analyzing data from group discussions can be challenging due to overlapping conversations and multiple perspectives.

Example: Conducting a focus group with a team of developers to identify issues with communication tools and processes during a sprint in Agile software development.

Qualitative research in CS and IS

3. Case Study

A case study is a qualitative research method that involves an in-depth investigation of a single case or multiple cases within their real-world context. The case can be an individual, organization, project, or system.

Advantages	Disadvantages
In-Depth Understanding: Provides comprehensive insights into a specific instance	**Limited Generalizability:** Findings from a single case or a small number of cases may not be generalizable to broader populations
Real-World Context: Captures phenomena as they occur in practice, offering valuable practical implications	**Time-Consuming:** Collecting and analyzing detailed data over an extended period can be labor-intensive.
Flexibility: Allows for a variety of data collection methods (e.g., interviews, observations, documents), making it adaptable to different research needs	**Subjectivity:** The interpretation of case study data may be influenced by the researcher's biases or perspectives.
Generates Rich Data: Case studies yield detailed qualitative data, helping to understand not just the "what" but also the "how" and "why" behind a phenomenon	**Analysis:** Due to the volume of data and multiple variables involved, analyzing case study findings can be difficult and may require a systematic approach.

Example: A case study of a banking system failure due to a server crash, examining the technical and organizational factors that contributed to the event and the solutions implemented.

Qualitative research in CS and IS

4. Action Research

Action Research is a participatory and iterative research method that involves researchers working closely with participants to address a problem or improve a process in a real-world setting. The focus is on generating practical solutions while simultaneously studying the process. c.

Advantages	Disadvantages
Practical Solutions: It is aimed at solving real-world problems while contributing to academic knowledge.	**Time-Intensive:** The iterative cycles require continuous engagement and time investment, which can slow down the research process.
Collaborative Approach: Researchers work alongside participants (e.g., developers, users, managers) to develop solutions, promoting buy-in and relevance.	**Researcher Bias:** Close collaboration with participants may introduce subjectivity or bias in the interpretation of data
Iterative Process: The cyclical nature (plan, act, observe, reflect) allows continuous improvement and adaptation based on real-time feedback.	**Complex Data Collection:** Balancing the dual goals of solving a practical issue and generating research findings can make data collection and analysis more complex.
Immediate Impact: The outcomes can have immediate benefits for the participants involved.	**Limited Generalizability:** Findings are often context-specific, making it difficult to generalize results beyond the immediate setting.

Example: Action research could be applied to develop and continuously improve a cybersecurity awareness program within an organization.

Qualitative research in CS and IS

- The collected data is analyzed using the following **Commonly used Qualitative Analysis Techniques**:
  - **Coding:** A widely used method to analyze and reduce qualitative data. It organizes raw data into conceptual categories or "bins" by assigning labels to pieces of data (words, phrases, or documents).
  - **Memoing:** Involves writing reflective notes or commentaries during or after data collection. Memos can capture observations, ideas, or initial interpretations of the data.
  - **Content Analysis:** Examines text for the presence and frequency of dominant concepts (e.g., words or constructs).

Qualitative research in CS and IS- Data analysis Software tools

- **Qualitative data analysis tools** help you make sense of customer feedback so you can focus on improving the user and product experience and creating customer delight.
- Some of the most commonly used tools are:

  - NVivo is one of the most popular qualitative data analysis tools and probably the most expensive. NVivo's Transcription tool transcribes and analyzes audio and video files from recorded calls—like interviews, and product demos—and lets you automatically transfer text files into NVivo for further analysis to:
    - ✓ Find recurring themes in customer feedback
    - ✓ Analyze different types of qualitative data, like text, audio, and video
    - ✓ Code and visualize customer input
    - ✓ Identify market gaps based on qualitative and consumer-focused research

  - MAXQDA: It is a data analysis software that can analyze and organize a wide range of data, from handwritten texts, to video recordings, to Tweets.

Several other tools are available.

Mixed Method research in CS and IS

Mixed Methods Research in Computer Science (CS) and Information Systems (IS) refers to a research approach that combines both **qualitative** and **quantitative** methods within a single study. This approach is valuable because it allows researchers to draw from the strengths of both methods, addressing complex research questions that require multiple types of data and analysis.

Key Characteristics:

- **Integration of Data:** The qualitative and quantitative data are combined, often by comparing, validating, or expanding the findings from one method with the other.
- **Complementarity:** Qualitative data can provide a deeper understanding of quantitative results, while quantitative data can offer generalizability to qualitative insights.
- **Sequential or Concurrent Design:** Mixed methods can take one of the three following forms:

Mixed Method research in CS and IS

Sequential Explanatory Strategy	Sequential Exploratory Strategy	Concurrent Triangulation Strategy
		
• Collection and analysis of quantitative data followed by collection and analysis of qualitative data. • Use qualitative results to assist in explaining and interpreting the findings of a quantitative study **Example: Evaluation the effectiveness of a new educational app** The researchers first gather quantitative data through surveys to assess user satisfaction and learning outcomes among students. They then explore these results further through qualitative methods like interviews and focus groups, delving into the reasons behind the statistics, guiding more targeted improvements in app design and functionality.	• Collection and analysis of qualitative data followed by collection and analysis of quantitative data. • Use quantitative data and results to assist in the interpretation of qualitative findings. **Example: Development of a New Feature for Programming IDEs** The researchers first conduct qualitative research through interviews and observations with novice programmers to identify their challenges and needs. Insights gained guide the development of a quantitative survey aimed at a broader audience to validate the demand for new features.	• Uses different methods concurrently -- Why? ▸ "What people say" could be different than "what people do." S, Collecting data from multiple sources helps improve validity. **Example: Assessing a new algorithm's efficiency** The researchers might conduct experiments to gather quantitative performance metrics (like execution time and resource usage) while also interviewing expert users to gain qualitative insights into the algorithm's practical usability and effectiveness

All methods have limitations. The strengths of one method can compensate the weaknesses of other methods.

- In Software development in particular, there are lots of different methodologies, and each of them has its own advantages and disadvantages. The most commonly used are:
    1. Waterfall Methodology
    2. Agile Methodology. It has different types, the most common are:
        - 2.1 Scrum
        - 2.2 Kanpan
        - 2.3 Lean Software Development (LSD)

1.Waterfall Methodology (Traditional Approach) (1 of 3)

- Waterfall methodology is a structured approach to project management, guiding through clear stages.
- **Phases** are initiation – design – implementation – verification - deployment.
- Each phase is clear and concise, making it ideal for projects with clear goals.
- However, it may not be flexible for sudden changes.
- Waterfall is a guide for thoroughness and keeping focus on the goal.
- Early software development projects often used a waterfall approach

When to use it:

- The waterfall approach is great for manufacturing and construction projects, which are highly structured, and when it's too expensive to pivot or change anything after the fact. The waterfall method makes use of Gantt charts for planning and scheduling.

Examples:

- CS Example: Developing a desktop software application with rigid specifications.
- IT Example: Implementing a network infrastructure upgrade with predetermined steps and deliverables.

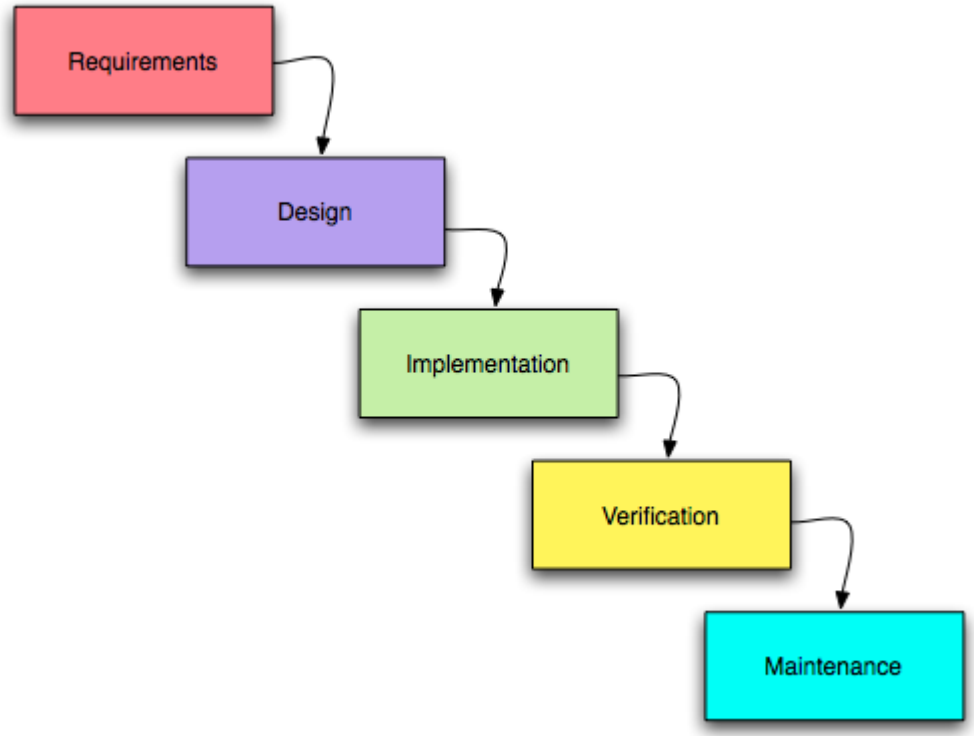
1.Waterfall Methodology (Traditional Approach) (2 of 2)

Some of the advantages of Waterfall are:

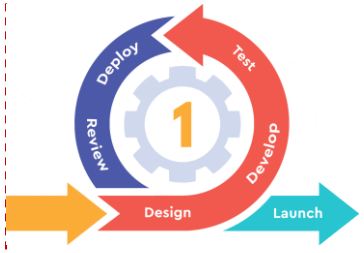
2. Waterfall Model is very simple and easy to understand.
3. Waterfall Model works well for smaller projects and projects where requirements are well understood.
4. In each phase detailed documentation is required.
5. Phases in the Waterfall model are processed one at a time.

Some of the disadvantages of Waterfall are:

2. It is not well-suited where requirement are continuously changing and updating.
3. Also there is no error detection or feedback present in each stage.
4. Waterfall Model is not well-suited for complex projects.
5. Waterfall Model can result in a lengthy development cycle, as each phase must be completed before moving on to the next. This can result in delays and increased costs if requirements change or new issues arise.
6. There is no collaboration and innovation among different team members



2. Agile Methodology (1 of 5)



• Agile is a collaborative, evolving approach to self-organization across teams.
•Characteristics: Flexible, fast, user feedback-driven , unlike waterfall project management.
•Its project planning and work management are adaptive, evolutionary, and seek early delivery.
•Originated in 2001 with the publication of the "Manifesto for Agile Software Development.“

When to use it:

•Originated from software development culture.
• Applicable to non-software products with innovation and uncertainty.
• Used in projects requiring responsive, fast-paced production schedules.
• Suitable for projects like computers, motor vehicles, medical devices, food, clothing, music, and marketing.

Examples:

- CS Example: Developing a machine learning model iteratively based on user feedback and testing at each sprint.
- IT Example: Building a customer-facing web application with regular user feedback to modify features.

Visit: https://www.youtube.com/watch?v=eVDZJ60c_7c

2. Agile Methodology (2 of 5)

- Different Agile frameworks are available.
- What all agile frameworks have in common is that they continuously iterate on the work process itself and aim to deliver value to customers quickly and frequently.
- But there are now countless frameworks and each one offers its own flavor of agile.
- The most common Agile framework that will be discussed are:

  2.1 Scrum (https://www.youtube.com/watch?v=pl9l4YS9AT0))

  2.2 Kanpan (https://www.youtube.com/watch?v=eBXsr7wdWH4)

  2.3 Lean Software Development (LSD) (https://www.youtube.com/watch?v=LKL41cC5pxA))

2. Agile Methodology (3 of 5)

How to Pick an Agile Framework?

It’s crucial to evaluate several key dimensions before choosing:

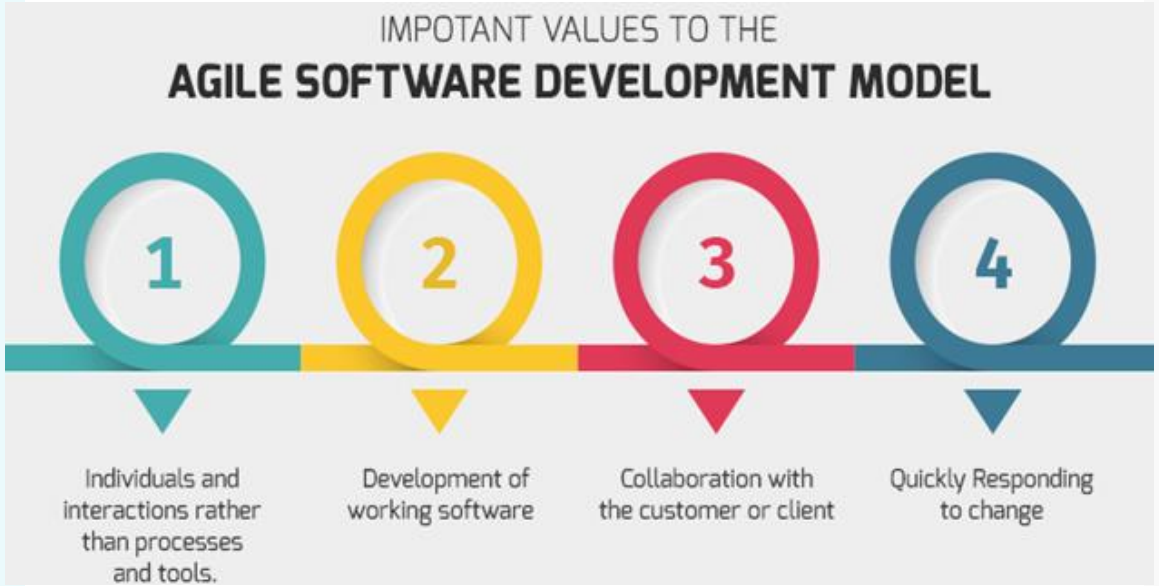
- **Team Size:** The framework you choose should match your team's size. Large frameworks can overwhelm small teams, making them inefficient, while lightweight frameworks might not provide enough structure for larger groups, leading to disorganization..
- **Project Complexity:** When handling projects with multiple complex layers, it’s important to select a framework that helps manage this complexity effectively without halting progress. The framework should simplify project management without removing necessary flexibility.
- **Company Culture:** What’s your company vibe? Your chosen framework must fit into how things run around your office.
- **Customer Engagement:** If customer feedback is crucial to your operations, choose a framework that incorporates this feedback consistently into the development process. Opt for models that facilitate regular inclusion of customer insights in sprint reviews or planning sessions.
- **Adaptability:** Given the rapid pace of change in today’s business environments, it’s essential to adopt a framework that can swiftly adapt to new conditions. Ensure that the framework can support quick decision-making and adjustments without falling apart, especially when immediate shifts are necessary.

2. Agile Methodology (4 of 5) The Agile Manifesto

The Agile Manifesto is a document that identifies four key values and 12 principles that its authors believe software developers should use to guide their work. Formally called the *Manifesto for Agile Software Development*,

The Authors

Kent Beck
Mike Beedle
Arie van Bennekum
Alistair Cockburn
Ward Cunningham
Martin Fowler
Robert C. Martin
Steve Mellor
Dave Thomas
James Grenning
Jim Highsmith
Andrew Hunt
Ron Jeffries
Jon Kern
Brian Marick
Ken Schwaber
Jeff Sutherland



The Agile Manifesto

We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:

Individuals and interactions over processes and tools

Working software over comprehensive documentation

Customer collaboration over contract negotiation

Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more.

2. Agile Methodology (5 of 5) The Agile Manifesto

Some of the advantages of Agile are:

2. It provides faster delivery of software products and features.
3. Agile methodologies prioritize customer satisfaction.
4. In Agile methodology the daily interactions are required between the business people and the developers.
5. Changes in the requirements are accepted even in the later stages of the development.
6. Better adaption to rapidly changing requirements and respond faster.

Some of the disadvantages of Agile are:

2. Agile development models require a high degree of expertise from team members.
3. Agile model not suitable for larger, and complex project.
4. For complex projects, the resource requirement and effort are difficult to estimate.
5. Sometimes in Agile methodology the requirement is not very clear hence it's difficult to predict the expected result.