

Research Methodology in Computer Science (AI800)

Lecture 6

Research Methods in Computer Science – part 2



Prepared by
Dr. Mohammed Anbar



Outlines

- Learning Outcomes
- Research Design
- Research Design Parts
- Important Features of a Research Design
- FEATURES OF A GOOD DESIGN
- Research Design Concepts
- Research Design in Research process
- Data Collection Definition and Terms
- Primary Data Vs Secondary Data
- Primary Data Collection
- Quantitative Data Collection
- Qualitative Data Collection
- Secondary Data Collection
- Rules for Collecting Data
- Developing a Data Collection Plan
- Strategies for Rapid Reporting
- Data Analysis Definition, important and Types
- Techniques for data analysis
- Data Analysis Methods and Process
- Tools to support data analysis
- Results Interpretation and Discussion
- How to Organize the Results Section
- Tips to Write the Results Section
- How to Organize the Discussion Section
- Tips to Write the Discussion Section



Learning Outcomes

- To understand the research design parts, features and concepts.
- To Understand the importance of data collection in computer Science research.
- To Learn about different methods and techniques used for data collection.
- To Gain knowledge on how to design and implement effective data collection strategies.
- To Understand the process and Rules for Collecting Data
- To Understand the basic concepts and principles of data analysis.
- To Learn various methods and techniques used in data analysis.
- To Gain knowledge on different types of statistical analyses and when to use them.
- To Understand the importance of interpreting and discussing research results.
- To Learn how to effectively interpret and discuss research findings.



Research Design

- A research design is the planning of data collection and analysis in a way that seeks to balance efficiency with relevance to the research purpose.
 - Comprises a schedule outlining the researcher's steps, from formulating the hypothesis to considering its operational ramifications to conducting the final data analysis.
- In more precise terms, the design choices pertain to:

What?

the study about?

type of data is required?

periods of time will the study include?

will be the sample design?

techniques of data collection will be used?

style will the report be prepared?

Where?

will the study be carried out?

can the required data be found?

Why?

the study being made?



Research Design Parts

Sampling Design

- Concerned with the procedure for selecting elements to be observed
- Elaborate on the different types of sampling techniques : random sampling, or cluster sampling.

Observational Design

- Relates to the conditions under which the observations are to be made.
- Includes various factors such as the location, time, and environment

Statistical Design

- Concerns with the question of how many items are to be observed and how the information and data gathered are to be analyzed.
- Considers with practical considerations such as time, resources, and feasibility of data collection methods.

Operational Design

- Translates the theoretical framework into practical actions and procedures.
- Involves determining the specific tools, techniques, and resources needed to carry out the sampling, statistical, and observational designs effectively.



Important Features of a Research Design

- (i) It is a **strategy outlining** the **kinds** and sources of **data pertinent** to the research **issue**.
- (ii) It is a **plan outlining** the **methodology** for **collecting** and **analyzing** the data.
- (iii) It also includes the time and cost budgets.



FEATURES OF A GOOD DESIGN

There are several key features that define a good design in computer science research
Should be:

- *Clear* and *concise*, ensuring that it is easily understandable to both experts and non-experts in the field.
- *Flexible* and *adaptable*, allowing for modifications and improvements as new technologies and methodologies emerge.
- *Well-documented*, providing thorough explanations of the research process and methodology used, enabling others to replicate or build upon the findings.



1. **Dependent and independent variables:**
 - ☐ Dependent variables **refer** to the outcomes or results that are being measured or observed.
 - ☐ Independent variables **refer** to the factors or conditions that are manipulated or controlled by the researcher.
 - ☐ **Identifying** and understanding the relationship between dependent and independent variables to effectively design experiments and analyze data to draw meaningful conclusions in computer science research.
2. **Extraneous variable:**
 - ☐ **refers** to any variable that may influence the relationship between the independent and dependent variables in a research study.
3. **Control:**
 - ☐ **refers** to the ability to manipulate and regulate variables to observe their effects on the outcome of an experiment or study.



Research Design Concepts (cont.)

4. Confounded relationship

- ☐ The framework and methodology used to conduct studies and gather data.
- ☐ These concepts help researchers establish clear objectives, identify variables, and determine the most appropriate research methods

5. Research hypothesis:

- ☐ A statement that proposes a relationship between variables or predicts an outcome based on existing knowledge or theories.

6. Experimental and non-experimental hypothesis-testing research

- ☐ Experimental research design involves manipulating variables and measuring their effects on the outcome.
- ☐ Non-experimental hypothesis-testing research design does not involve manipulation of variables but rather focuses on observing and analyzing existing data to test hypotheses.



Research Design Concepts (cont..)

7. Experimental and control groups

- ☐ Experimental groups are exposed to a specific treatment or intervention, while control groups do not receive the treatment and serve as a baseline for comparison.

8. Treatments

- ☐ The different conditions under which experimental and control.

9. Experiment

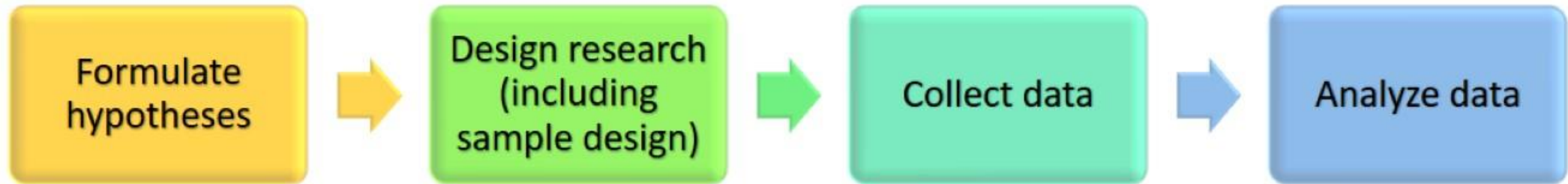
- ☐ The process of examining the truth of a statistical hypothesis, relating to some research problem.

10. Experimental unit(s):

- ☐ The pre-determined plots or the blocks, where different treatments are used.



Research Design in Research process





In summary

A research design must include :

- (a) a statement of the research problem;
- (b) the methods and techniques to be used in data collection;
- (c) the study population; and
- (d) a plan for processing and analyzing the data



Data Collection





What is Data Collection?

- Data collection is the *process* of *gathering* information for a *specific purpose*
 - *Collecting*, *quantifying*, and *evaluating* precise research insights through standard, validated methods.
- This *involves* various methods such as *surveys*, *experiments*, *observations*, or *existing datasets*.
- The *collected data serves* as the foundation for *analysis* and *drawing* meaningful conclusions to *contribute* to the advancement of computer science knowledge and innovation.





Data Collection Terms

➤ *Data Collection Strategies:*

There is no one ideal strategy; the choice depends on

1. What you **should** know: **narratives** or numbers
2. Where the environment, files, **and** people **contain** the data
3. **Time and resources** available
4. **Difficulty** of the **information** to be **gathered**
5. **How often** data **is collected** ?



Data Collection Terms(cont.)

➤ *Data sources:*

- ❑ A data source is a location where information is gathered.
- ❑ Include academic journals, conference proceedings, and online databases.
- ❑ Provide researchers with valuable information and insights into various topics such as algorithms, artificial intelligence, data mining, software engineering.

➤ *Data sources divided* into two groups based on *technology*:

- Machine data sources.
- File data sources.

➤ *Data sources classified* into two types:

- Primary data.
- Secondary data.



Data Collection Terms(cont.)

➤ *Data Collection Methods:*

- ☐ Techniques and procedures used to gather information for research purposes.
- ☐ Involve either quantitative or qualitative approaches to data gathering.

➤ *Data Collection Methods* **categorized** into two :

- ✓ **Primary Data Collection Methods** divided into two categories:
 - ☐ Quantitative methods
 - ☐ Qualitative methods.
- ✓ **Secondary Data Collection Methods** include data collected from previous experiments or studies conducted:
 - ☐ Internal data sources.
 - ☐ External data sources.

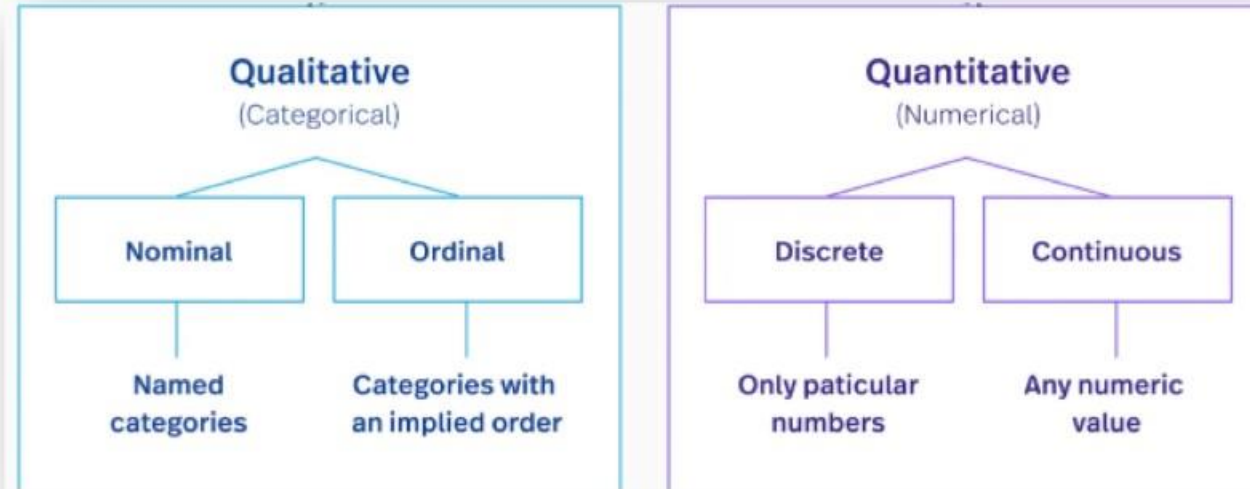


Primary Data vs Secondary Data

- *Primary data refers* to the original data that is collected firsthand by the researcher for a specific research project.
 - In computer science research, primary data can be obtained through various methods such as surveys, experiments, observations, or interviews.
 - Allows researchers to have control over the data collection process and ensures that the data is specific to their research needs.
- *Secondary data refers* to existing data that has been collected by someone else for a different purpose but can be utilized by researchers for their own studies.
 - Include sources like books, articles, databases, or previous research studies.
 - Readily available and can save time and resources.



Primary Data Collection



Quantitative methods involve the collection and analysis of numerical data, often through surveys or experiments, to measure and quantify variables.

Qualitative methods focus on gathering non-numerical data, such as interviews or observations, to gain in-depth insights and understand complex phenomena.



Quantitative Data Collection

Quantitative Data Collection Methods

Probability Sampling

Simple random sampling.
Cluster sampling.
Systematic sampling
Stratified sampling.

Interviews

Telephone interviews
Face-to-face interviews
Computer-Assisted Personal Interviewing (CAPI)

Surveys/Questionnaires

Web-based questionnaire
Mail Questionnaire

Observations

Naturalistic observation
Structured observation

Document Review

Public Records
Personal Documents
Physical Evidence



Qualitative Data Collection

Qualitative Data Collection Methods



Individual Interview



Qualitative Surveys



Focus Group Discussions



Record Keeping



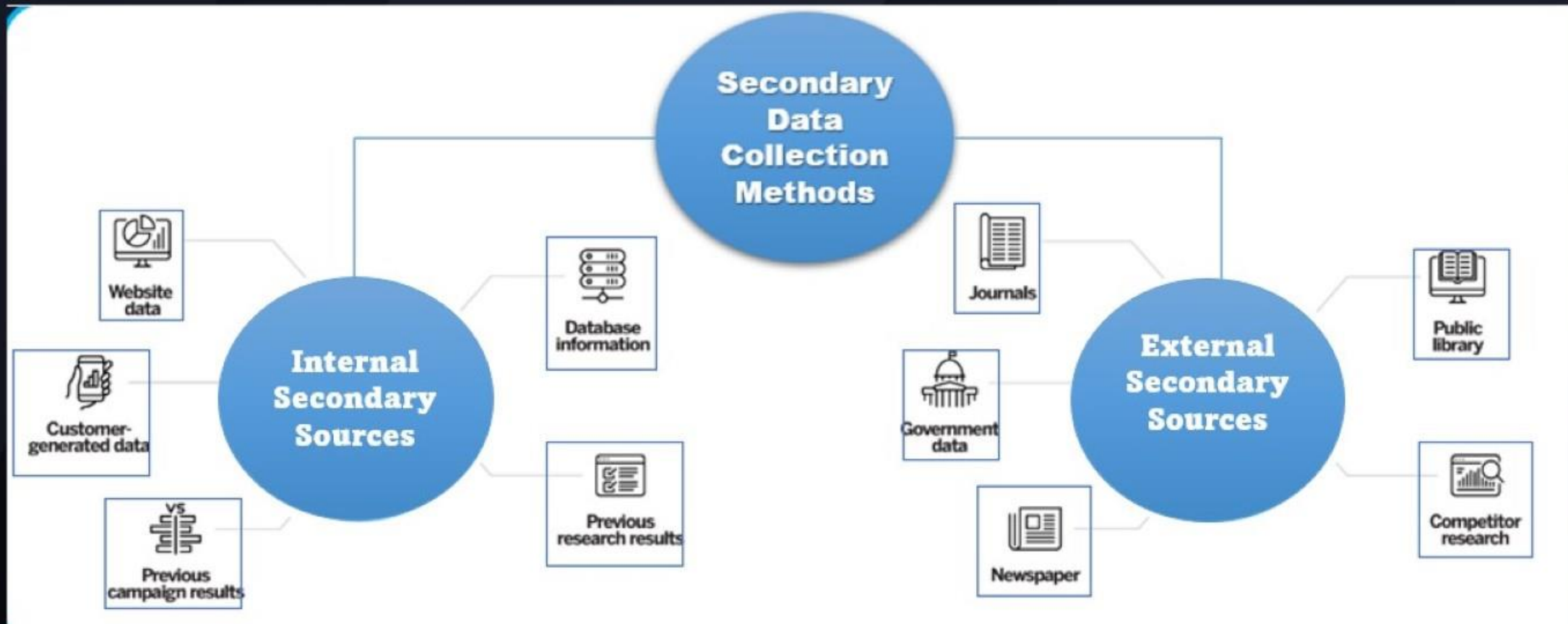
Case Studies



Observations



Secondary Data Collection





Rules for Collecting Data

- Apply a **variety of** data collection **techniques**.
- Utilize the data at **hand**, but **be aware that**:
 - The **definition of** the measures.
 - How the **information was gathered** and **cleansed**.
 - The **amount** of data missing.
 - How the data's accuracy was **guaranteed**.
- If **real data collection is required**:
 - ✓ Establish **protocols** and **adhere** to them.
 - ✓ **keep up-to-date** definition and coding **data**.
 - ✓ Check the data **entry and** coding accuracy.

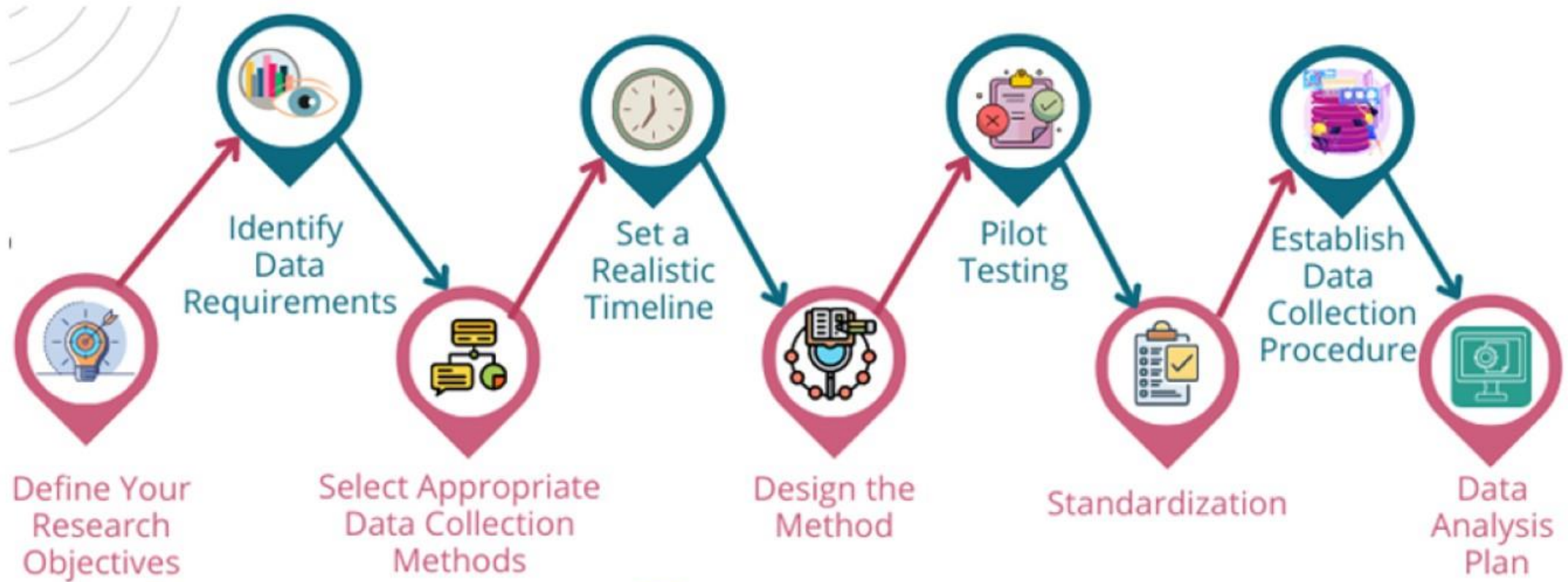


Developing a Data Collection Plan

- A crucial step in ensuring the accuracy and reliability of the collected data.
- Developing a Data Collection Plan involves
 - Identifying the variables to be measured.
 - Determining the appropriate data collection methods.
 - Establishing a timeline for data collection.
- Important to consider ethical considerations and potential biases that may impact the results of the research.



Developing a Data Collection Plan (cont.)





Developing a Data Collection Plan (cont.)

- The *importance* of *developing* a clear research question and objectives *before creating* a data collection plan in computer science research.
- *Techniques* for *identifying* the most appropriate data sources and variables *to include* in a data collection plan, *considering* the specific requirements and goals of the study.
- *Methods* for efficiently *collecting* large amounts of data in computer science research, such as *utilizing* automated tools or *leveraging* existing datasets.
- *Best practices* for *ensuring* data quality and accuracy during the collection process, *including* techniques for validating collected data.



Data Analysis

Data Analysis





What is Data Analysis?

- *Data analysis* is the *process* of *inspecting*, *cleaning*, *transforming*, and *modeling data* to discover useful information, draw conclusions, and support decision-making.
- In computer science research, *data analysis* plays a crucial role in *extracting meaningful* insights from *large datasets* and *identifying patterns* or *trends* that can contribute to the advancement of knowledge in various domains.
- *Data analysis involves* applying statistical techniques and computational algorithms to *analyze* data and uncover valuable findings that can be used for further research or practical applications.



Why is data analytics important?

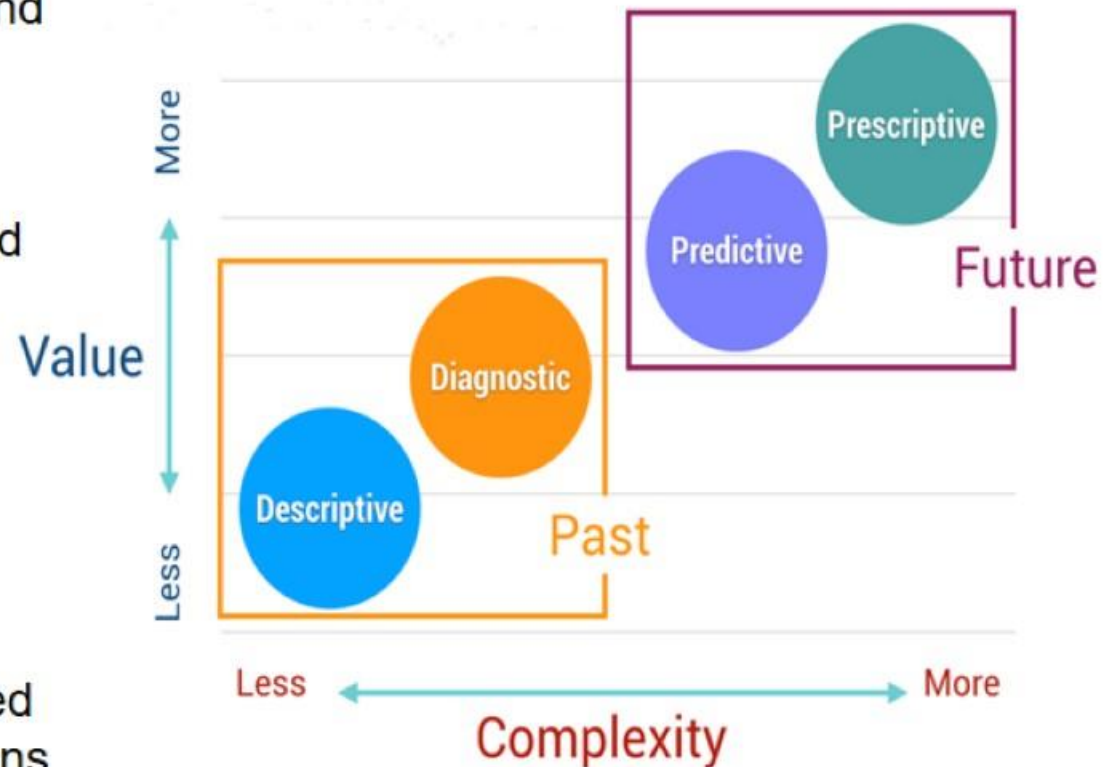
- Informed decision-making
- Identifying opportunities and challenges.
- Improving efficiency and productivity
- Customer understanding and personalization
- Competitive advantage
- Performance tracking and evaluation
- Predictive analytics
- Data-driven innovation
- Fraud detection and security
- Regulatory compliance





Data Analysis Types

- Descriptive Analysis: a method used to analyze and interpret data collected during experiments or studies.
- Diagnostic analysis: a process used to identify and understand problems or issues within computer systems or software.
- Predictive Analysis: a field that involves using historical data and statistical models to make predictions about future events or behaviors.
- Prescriptive analysis: refers to the use of advanced algorithms and models to provide recommendations and make informed decisions.





Techniques for data analysis

Data analysis techniques that depend on the type of information collected:

➤ **Qualitative data:**

- ✓ The underlying emotions, body language, and expressions are the sources of qualitative data for researchers.
- ✓ The most popular methods for gathering this data are focus groups, open-ended interviews, and observation groups.
- ✓ During the data collecting stage, researchers typically examine trends in the observations made by the participants.

➤ **Quantitative data:** Numerical data is how quantitative data is shown. It emphasizes observable outcomes.



Data Analysis Methods

Descriptive
Statistics

Inferential
Statistics

Data Visualization

Exploratory Data
Analysis (EDA)

Time Series
Analysis

Cluster Analysis

Factor Analysis
and Principal
Component
Analysis (PCA)

Text Mining and
Natural Language
Processing (NLP)

Qualitative Data
Analysis

Quantitative Data
Analysis

Data Mining

Regression
Analysis



Data Analysis Process



1. Define your questions carefully



2. Establish measurement priorities



3. Collect all the relevant data.



4. Analyze the data you gathered



5. Review and interpret the results.



Tools to support data analysis

- There are a variety of tools available to support data analysis range from spreadsheet software like Microsoft Excel to more advanced statistical software such as R or Python.
- These tools allow analysts to manipulate and visualize data, perform complex calculations, and generate meaningful insights from large datasets.
- Some tools offer machine learning capabilities that can help identify patterns and make predictions based on the data.





Results Interpretation and Discussion

- The *results section contains* a description about the main findings of your research.
- The *discussion section interprets* the results for readers and *provides* the significance of the findings and *should not repeat* the results.
- The *interpretation* and *discussion* of results play a crucial role in advancing knowledge and understanding.
- The *interpretation* and *discussion processes involve* analyzing the data collected, identifying patterns or trends, and drawing meaningful conclusions.
- Researchers often *compare* their findings with existing literature to *validate* their results and *contribute* to the broader scientific community.



How to Organize the Results Section

- Organizing the results section is crucial for effectively presenting the findings of a study.
- One common approach is to start by summarizing the main results and highlighting any significant findings.
 - This can be followed by presenting the results in a logical and coherent manner, using tables, graphs, or figures to enhance clarity and understanding.
 - ✓ Text: to explain about the research data
 - ✓ Figures: to display the research data and to show trends or relationships.
 - ✓ Tables: to represent a large data and exact value.
- Provide sufficient details about the methodology used and any limitations or assumptions made during the study to ensure transparency and reproducibility of the results.



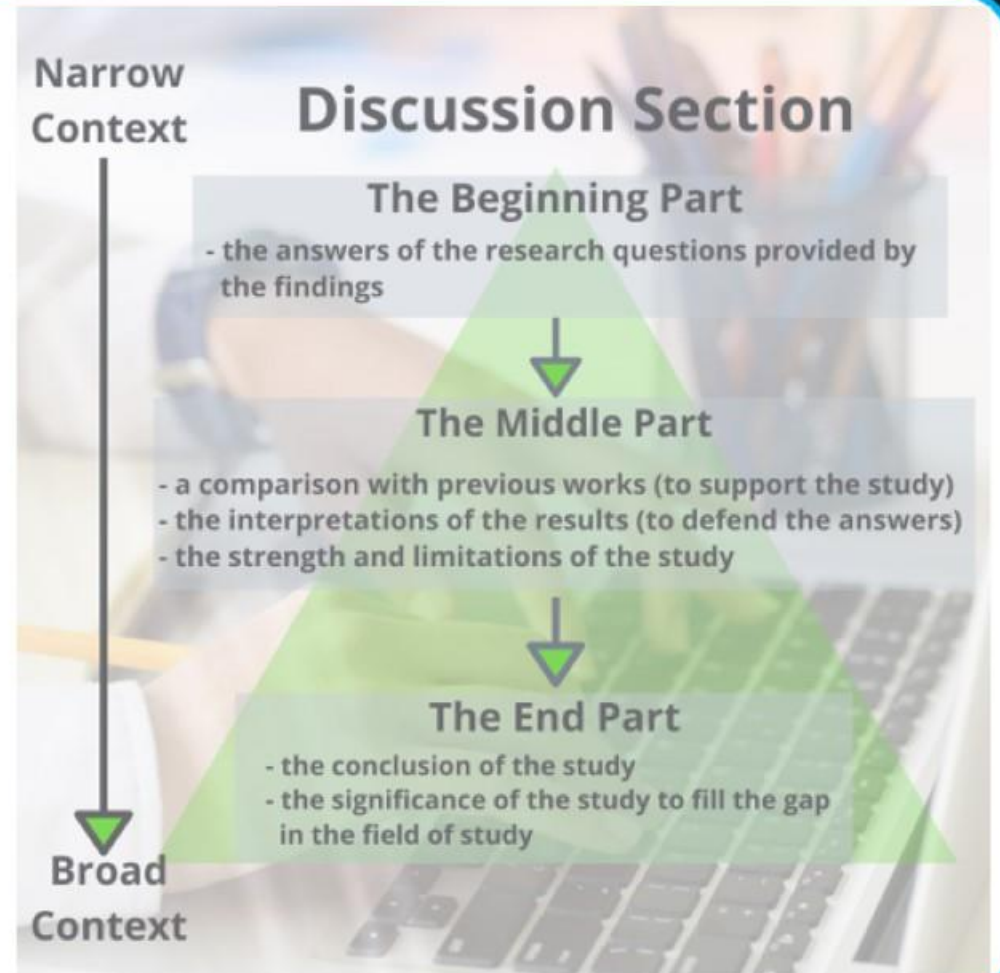
Tips to Write the Results Section

- **Clearly outline** the **key findings** and **contributions** of your research in a **concise** and **structured manner**.
- **Provide** a summary of the **main results**, **highlighting** their **significance** and **relevance** to the field of computer science.
- **Present** your **results** in a **logical** and **organized manner**, using *tables*, *graphs*, or *visualizations* when appropriate to **enhance** clarity and understanding.
- **Make** sure to **include relevant** statistical analyses or **metrics** to support your findings.
- **Discuss** any **limitations** or **potential biases** that may have *influenced the results*, *acknowledging any challenges*



How to Organize the Discussion Section

- Summarize the main results and their significance in relation to the research objectives.
- Provide a clear context for readers to understand the implications of the study.
- Discuss any limitations or potential biases that may have affected the results
- Demonstrates a thorough understanding of the research process and promotes transparency in scientific inquiry.





How to Organize the Discussion Section (cont.)

A list of questions to guide you when organizing the structure of your discussion section (Viera et al., 2018):

- What experiments did you conduct and what were the results?
- What do the results mean?
- What were the important results from your study?
- How did the results answer your research questions?
- Did your results support your hypothesis or reject your hypothesis?
- What are the variables or factors that might affect your results?
- What were the strengths and limitations of your study?
- What other published works support your findings?
- What other published works contradict your findings?
- What possible factors might cause your findings different from other findings?
- What is the significance of your research?
- What are new research questions to explore based on your findings?



Tips to Write the Discussion Section

Begin by summarizing the main findings and contributions of your research.

Analyze and interpret your results in relation to existing literature and previous studies.

Consider any limitations or constraints that may have influenced your results. Discuss potential sources of bias, data collection

Highlight the significance of your findings

Mention how the study filled a gap in knowledge.

Indicate the implication of your research.

Avoid generalizing, misinterpreting your results, drawing a conclusion with no supportive findings from your results.



References

- ❑ LAZAR, Jonathan; FENG, Jinjuan Heidi; HOCHHEISER, Harry. Research methods in human-computer interaction. Morgan Kaufmann, 2017.
- ❑ RICHEY, Rita C.; KLEIN, James D. Design and development research: Methods, strategies, and issues. Routledge, 2014.
- ❑ GODDARD, Wayne; MELVILLE, Stuart. Research methodology: An introduction. Juta and Company Ltd, 2004.
- ❑ Choy, L. T. (2014). The strengths and weaknesses of research methodology: Comparison and complimentary between qualitative and quantitative approaches. IOSR Journal of Humanities and Social Science, 19 (4), 99-104. <https://doi.org/10.9790/0837-194399104>



Thank You

