

# BASIC DATA COLLECTION METHODS, TOOLS AND ANALYSIS TECHNIQUES

By

**David Onen (Ph.D.)**

College of Education and External Studies

Makerere University

A Paper Presented on 7<sup>th</sup> January 2020 to the Participants at an Academic Staff Seminar held in the School of Research and Postgraduate Studies of Amoud University, Somaliland



# Introduction...

2

- Scientific research, generally, has an elaborate process.
- Nonetheless, this process still needs to be systematic and logical.
- As a result, all aspects of research should be justifiably determined.
- For example, the researcher **MUST** logically choose from the array of data collection method(s), tools and data analysis techniques that suit the achievement of him/her study objectives.
- But, the choice of data collection methods and tools is determined by several factors like the:
  1. type of research;
  2. kind of data the researcher intends to collect and use;
  3. degree of effectiveness and efficiency to be achieved, etc.
- All these must be guided by the data collection design/plan.

# The Data Collection Plan...

3

- The data collection plan must take care of the following:
  - 1) Research questions/objectives to be achieved.
  - 2) Kind of data that are available.
  - 3) How much data are yet needed.
  - 4) Determine how to measure the data.
  - 5) Decide who is going to collect the data.
  - 6) Determine where the data will be collected from.
  - 7) Project how the data will be analysed and presented.

# But First, what is Data?

4

- ❑ Data are a collection of facts, such as values or measurements.
- ❑ It can be numbers, words, measurements, observations or even just descriptions of things.
- ❑ With regards to research, data is any facts, values or useful information collected, observed, or created, for purposes of analysis to produce original research results.
- ❑ In short, data is any recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings.

# Types of Data by Source: Primary Vs Secondary Data

5

- *Primary data* are generated by a researcher who is responsible for the design of the study and the collection, analysis and reporting of the data.
- *Secondary data* are the raw data that have already been collected by someone else, either for some general information purpose, such as a government census or another official purpose, or for a specific research project.
- *Tertiary data* have been analyzed by either the researcher who generated them or an analyst of secondary data. In this case the raw data may not be available, only the results of this analysis.

# Types of Data by Form: Quantitative Versus Qualitative Data

6

- Data can be given in numerical or narrative form.
- Data obtained/transformed into numerical form are quantitative while in narrative form are qualitative.
- Quantitative data therefore, is the data type presented as numerical figures e.g. amount of goods sold per day, performance of students in marks scored, etc.
- While, qualitative data are data that cannot be measured in amounts; but reflect some quality which is being observed e.g. indicating that one is male or female, etc.

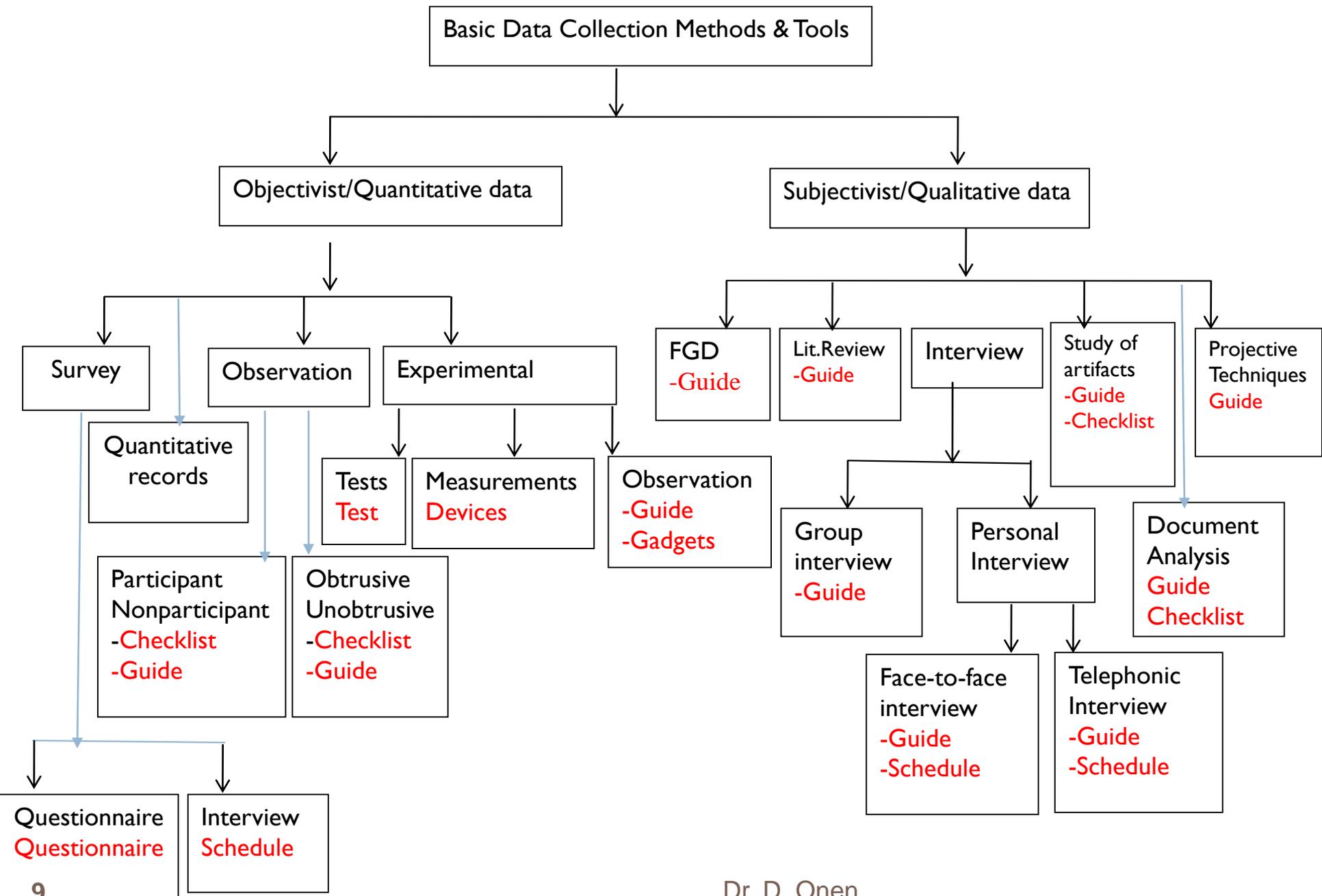
# Types of Data by Levels of Measurement

Level	Description	Types of categories	Examples
Nominal	A set of categories for classifying objects, events or people, with no assumptions about order.	Categories are homogeneous, mutually exclusive and exhaustive.	<ul style="list-style-type: none"><li>• Marital status</li><li>• Religion</li><li>• Ethnicity</li></ul>
Ordinal	As for nominal-level Categories except the categories are ordered from highest to lowest.	Categories lie along a continuum but the distances between them cannot be assumed to be equal.	<ul style="list-style-type: none"><li>• Frequency scale</li><li>• Likert scale</li></ul>
Interval	A set of ordered and equal- Categories on contrived measurement scale.	Categories may be discrete or continuous with arbitrary intervals and zero point.	<ul style="list-style-type: none"><li>• Attitude score</li><li>• IQ score</li><li>• Celsius scale</li></ul>
Ratio	As for interval-level measurement	Categories may be discrete or continuous but with an absolute	<ul style="list-style-type: none"><li>• Age</li><li>• Income.</li><li>• No. of</li></ul>

# What is meant is by Data Collection, Data Collection Methods, and Tools?

8

- **Data collection** refers to the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes (Northern Illinois University, 2005).
- **Data collection methods** meanwhile refer to the strategies, processes and procedures that a researcher employs to gather data during research (Kothari, 2004; 2009).
- **Data collection tools** refer to the devices/instruments used to collect data.



# How are Data Collected?

10

## Methods of Data Collection

1. Measurement
2. Testing
3. Surveys
4. Interviews
5. Focused Group Discussion
6. Observation
7. Document analysis
8. Literature review
9. A study of artifacts or relics
10. Oral testimony, etc.

## Instruments of Data Collection

1. Measuring Devices
2. Tests
3. Questionnaires
4. Interview Guide or Schedule
5. FGD Guide
6. Observation Checklist/Guide
7. Document analysis  
Checklist/Guide
8. Literature review Guide
9. A Guide for study of Relics
10. Oral testimony Guide, etc.

# **Validity and Reliability of Data Collection Instruments**

# Validity of Study Instruments

12

- ❑ Validity refers to the ability of the instrument to collect justifiable and truthful data i.e. the extent to which the instrument measures what it is intended to measure (Amin, 2004).
- ❑ Validity is therefore about soundness or effectiveness of study instruments.
- ❑ There are two aspects of validity of study instruments; namely: internal and external validity.
- ❑ Internal validity is the extent to which the data collected by the instruments can be accurately interpreted.
- ❑ While external validity is the extent to which the data collected by the instrument can be generalized to similar situations elsewhere.

# Approaches of Ensuring Validity

13

- ❑ There are several approaches for ensuring validity of study instruments.
- ❑ The choice of the appropriate approach depends on the type of validity. For examples:
  1. Face Validity – is mere appearance that a measure has validity. It involves using judgment to decide the extent of face validity.
  2. Content validity – is the extent to which the items in the instrument represent the content of the attribute being measured. It can also be ensured through judgment of the items by experts.
  3. Criterion validity – is the extent to which the data collected by the instruments can predict performance on some value measured other than the score on the instrument. This may be classified into predictive and concurrent validity.

# Commonest Approach for Ensuring Validity

14

- ❑ Most often researchers compute the Content Validity Index (CVI) for each item in the instrument as rated by two or more experts in order to determine how valid the study instrument is.
- ❑ The CVI = 
$$\frac{\text{Number of items rated relevant by all judges}}{\text{Total number of items in the instrument}}$$
- ❑ Note: If the CVI is 0.7 and above, the instrument can then be considered valid (Amin, 2004; Kathuri & Palls, 1994).

# Reliability of Study Instruments

15

- ❑ Reliability of an instrument is the ability of the instrument to collect the same data consistently under similar conditions (Amin, 2004; Odiya, 2009).
- ❑ Reliability is about accuracy and consistency of the instruments.
- ❑ Reliability enhances repeatability and generalization of study findings.
- ❑ It can be ensured through: test re-test method, split-half method, parallel form reliability method and internal consistency method.

# 1. Test Retest Method

16

- ❑ In the **test retest method**, the researcher pretests twice the instrument on a specific sample and then correlates the recorded scores of the two administrations ( $T_1$  and  $T_2$ ) to check consistency.
- ❑ A correlation coefficient of 0.7 and above is often recommended in most studies (Amin, 2004).

## 2. Split-Half Method

17

- ❑ In the **Split-Half method**, the researcher pretests the instruments only once on a selected sample and then divides the scores into two halves in order to test the correlation of the responses.
- ❑ It is cheaper than the test re-test because administration is done only once; but it may not be reliable in guaranteeing reliability.

# 3. Parallel Form Method

18

- The **parallel form reliability test** involves the use of two instruments that are similar – one may be a standardized one and the other developed by the researcher.
- Both instruments are administered to a sample of participants and the scores of responses obtained from the two instruments are correlated to measure reliability.
- The aim is to ensure that the new instrument is closely related to the standardized one in terms of the data collected.

# 4. Internal Consistency Method

19

- ❑ While **the internal consistency method** involves pre-testing the instrument once to a sample of respondents, and the scores of the responses are correlated using:
  1. Chronbach's alpha coefficient (for Likert scale items involving more than two response categories or
  2. Kudar-Richardson (KR) coefficient (for items involving two response categories e.g. yes/no).

# 5. Reliability and Validity for Qualitative Research Instruments

20

- For qualitative research instruments, their reliability and validity are more difficult to guarantee. But, the researcher can still enhance the validity and reliability of such instruments through:
  1. Triangulation – Using different tools or approaches to collect the same data.
  2. Appropriately clarifying purpose and objectives of the research.
  3. Adequately explaining data collection and data analysis procedures to promote consistency of the data collected.

# Basic Data Analysis Techniques

# What is Data Analysis?

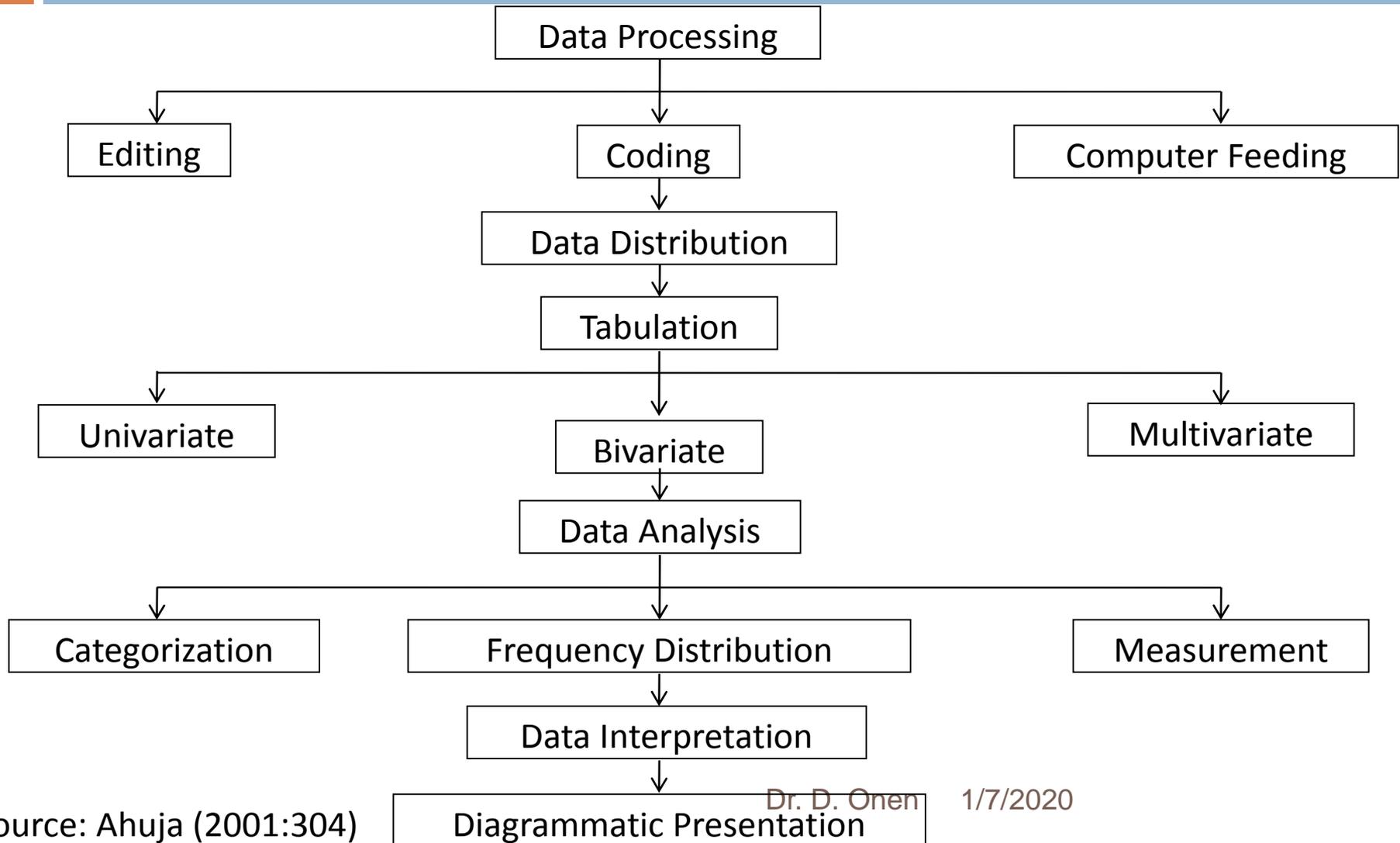
22

- The word Analysis ordinarily means detailed examination of the elements or structure of something, typically as a basis for discussion or interpretation.
- With regard to data, Analysis refers to the practice in which raw data is ordered and organized so that useful information can be extracted from it.
- It is therefore the process of converting raw data into meaningful statements.
- Broadly, data analysis include: Data processing, presentation, analysis and interpretation.

# Stages of Data Analysis

- According to Ahuja (2001), data analysis occurs in four stages...

23



# Basic Data Analysis Techniques

## Quantitative Data Analysis Techniques

## Qualitative Data Analysis Techniques

### Descriptive Statistics

### Inferential Statistics

### Content Analysis

### Narrative Analysis

### Discourse Analysis

### Framework Analysis

### Grounded Theory

#### Measures of Central Tendency

#### Measures of Variability

#### Correlational measures

#### Prediction/Effect Measures

- Means
- Median
- Median

- SD
- Range
- Minimums
- Maximums

- Pearson r
- Spearman rho
- Kendal Tau
- Chi-square

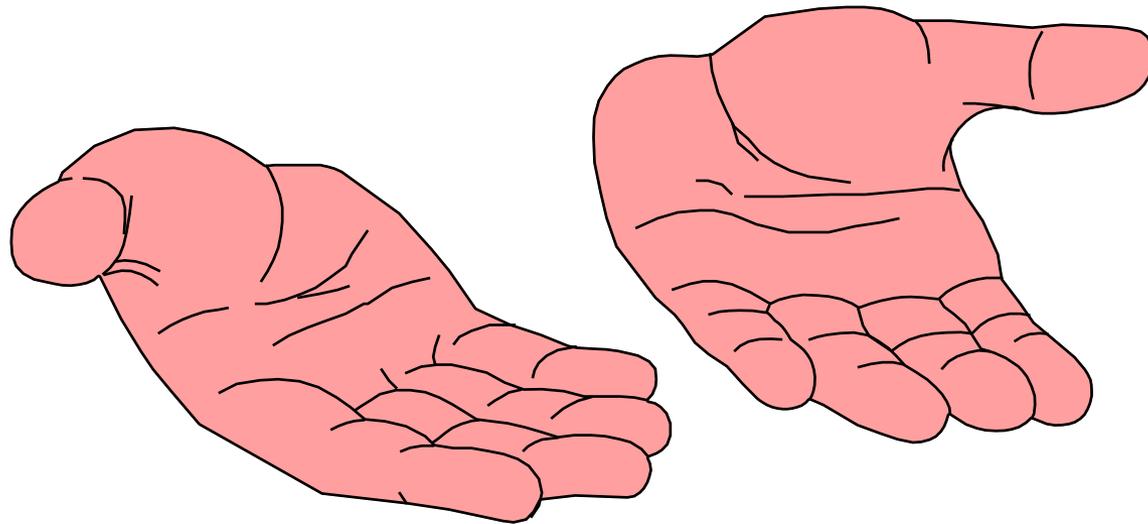
- Regression
- Student-t
- ANOVA
- Factor Analysis
- SEM

# Conclusion...

25

- There are many methods, instruments and techniques for collecting and analyzing both quantitative and qualitative data especially in business and educational researches.
- Each of these data collection methods, instruments and analysis techniques have their own strengths and weaknesses.
- A researcher must be aware of these strengths and limitations while choosing and using them.

**Thank you for your attention!**



**Any questions or comments ?**