

Principle Concepts of Statistical Data Analysis Software

Dr. Salem Saleh

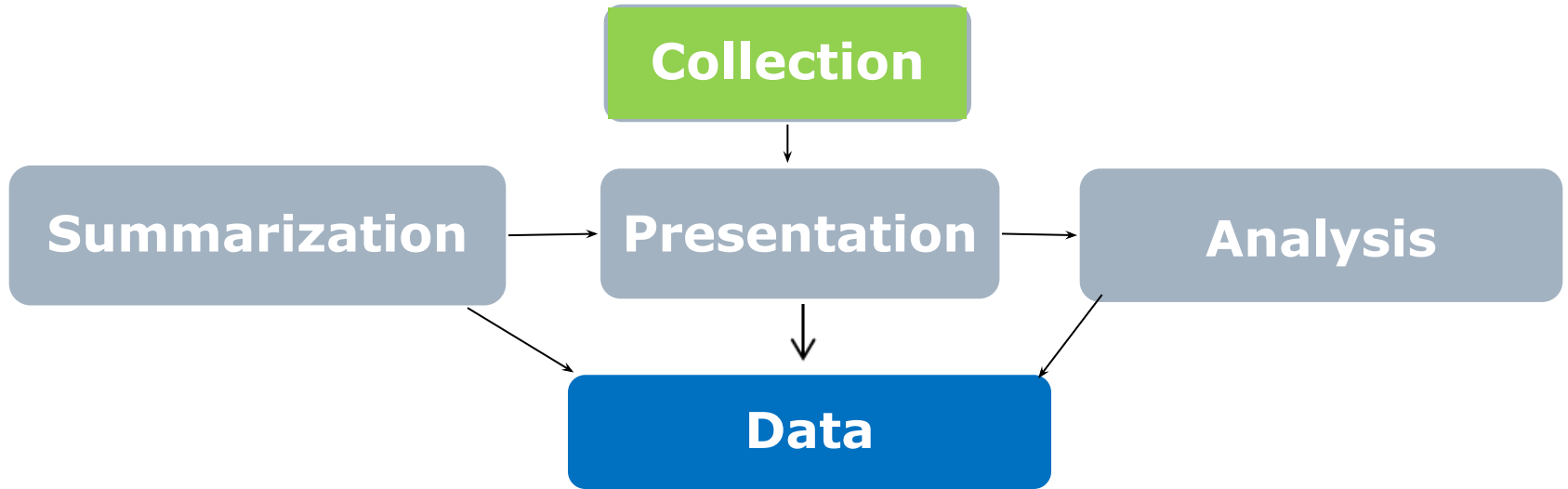
28. Nov. 2022

OUTLINE

- ✓ What is statistics- its sections and fields?
 - ✓ Data and Variables with their types.
 - ✓ Collection Data Methods.
 - ✓ Measurement Levels
 - ✓ Population and Sample with their types.
 - ✓ Some statistical data analysis software.
-

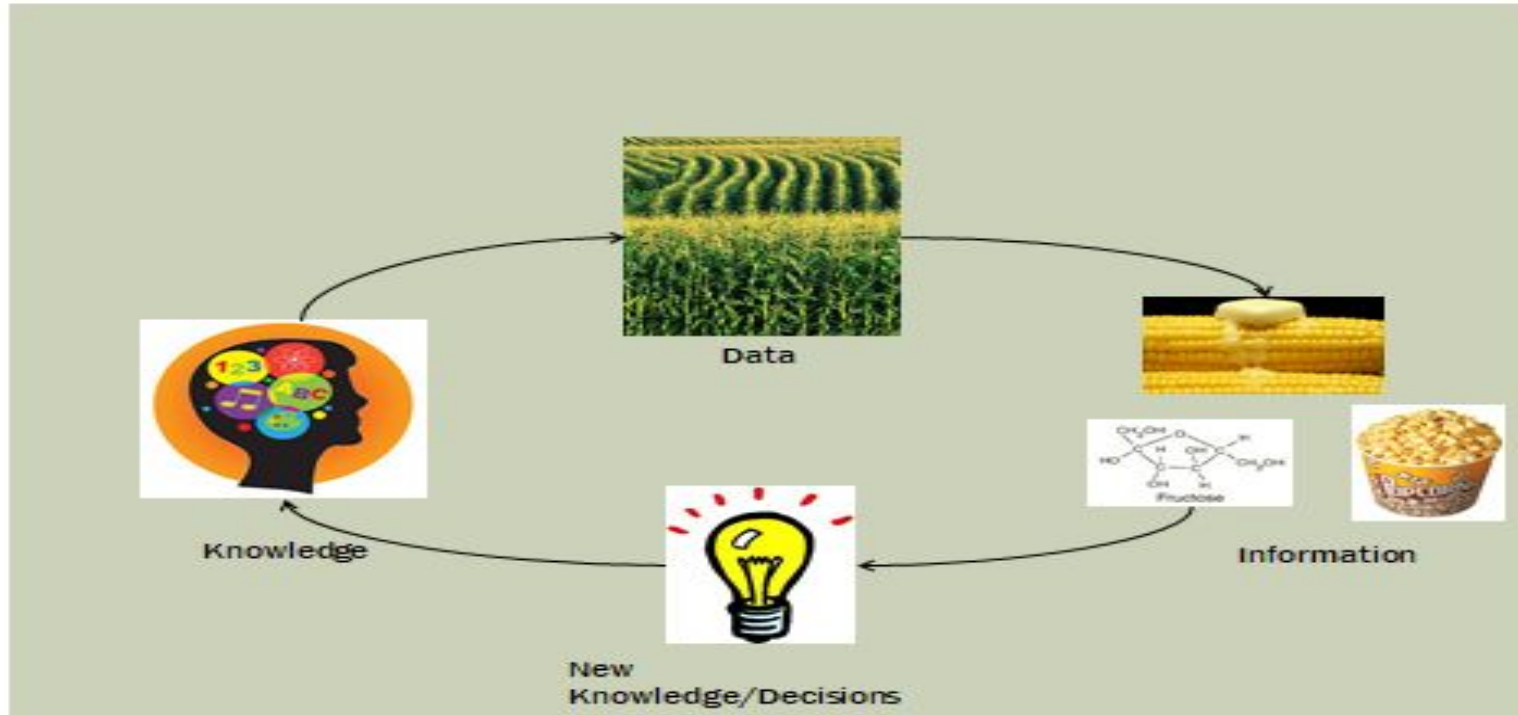
Basics of Statistics

Statistics: is the science concerned with:



In order to obtain acceptable scientific results,
to make appropriate decisions

Knowledge, Data, Information, Analyze, Decisions

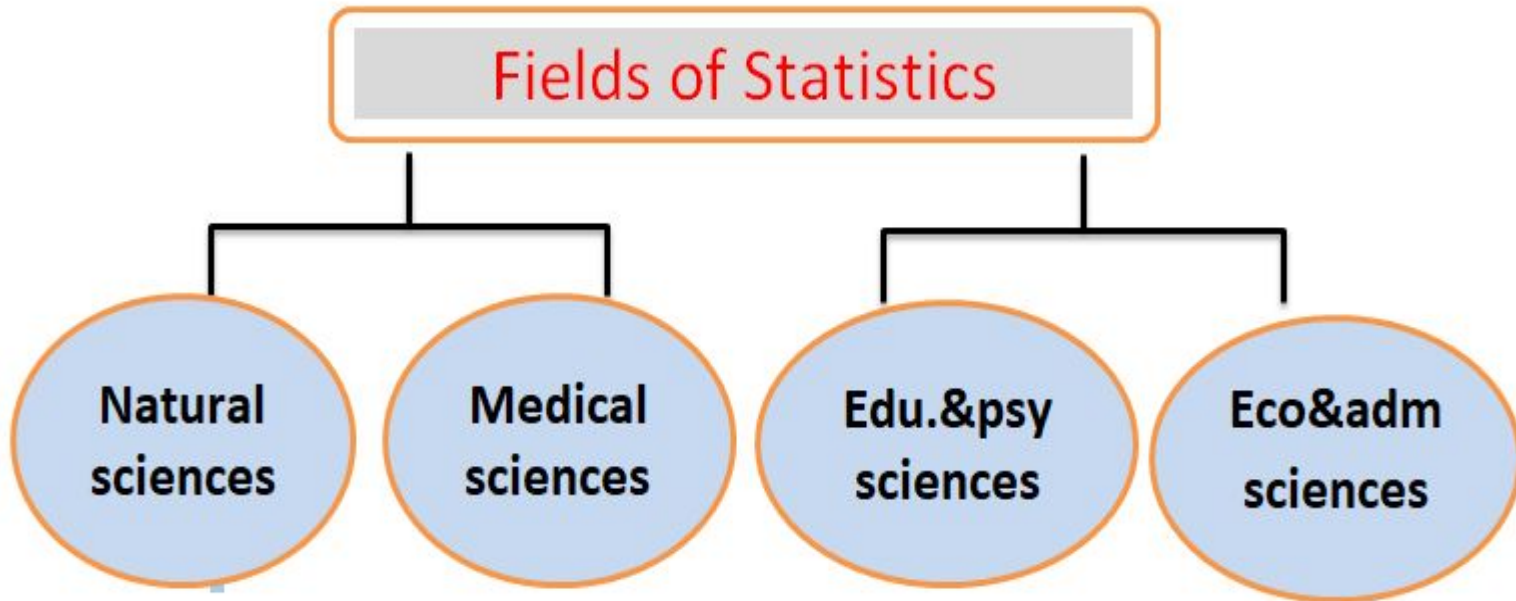


importance of statistics

The importance of statistics is can be summarized in the following points:

- ✓ Statistics is the main tool in scientific research.
- ✓ It is considered one of the important tools for drawing strategic policies for development.
- ✓ It is used in various areas of life on the personal and institutional level in managing life affairs.

Fields of statistics



Main sections of Statistics

Main sections of Statistics

Descriptive statistics

It is the branch that is concerned with:

- collecting, summarizing and presenting data. To,
- Extract some results and statistical indicators for the cases studied.

Inference statistics

It is the branch that is concerned with:

- generalizing the results obtained from the sample study on the population,
- prediction and estimating population parameters.
- Testing the validity of scientific hypotheses.

Some principle concepts

Variable:

is a characteristic or attribute that can have different values or attributes, such as age, gender, level education, Marital status.

- The variables whose values are determined by chance they are called random variables.

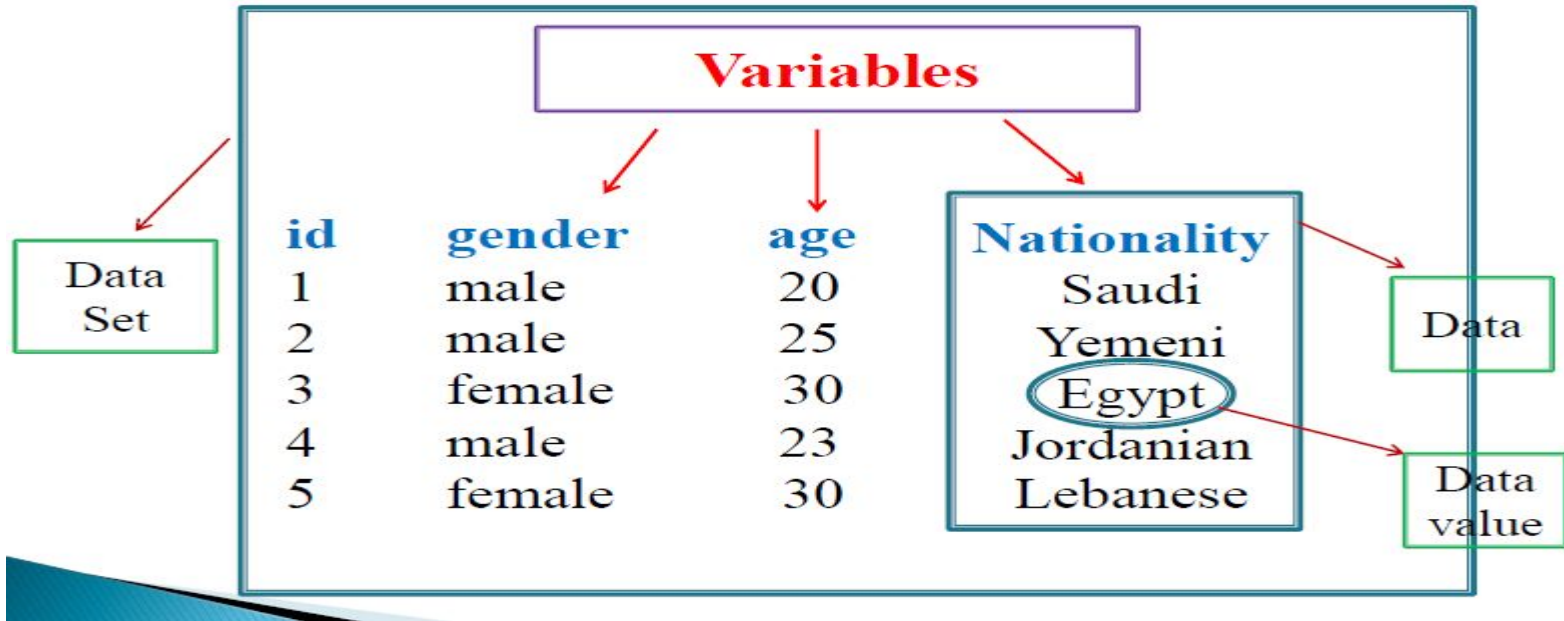
Data:

is the values that a variable can take it.

- A collection of data values forms a data set.

Some principle concepts

For example: A data set in table

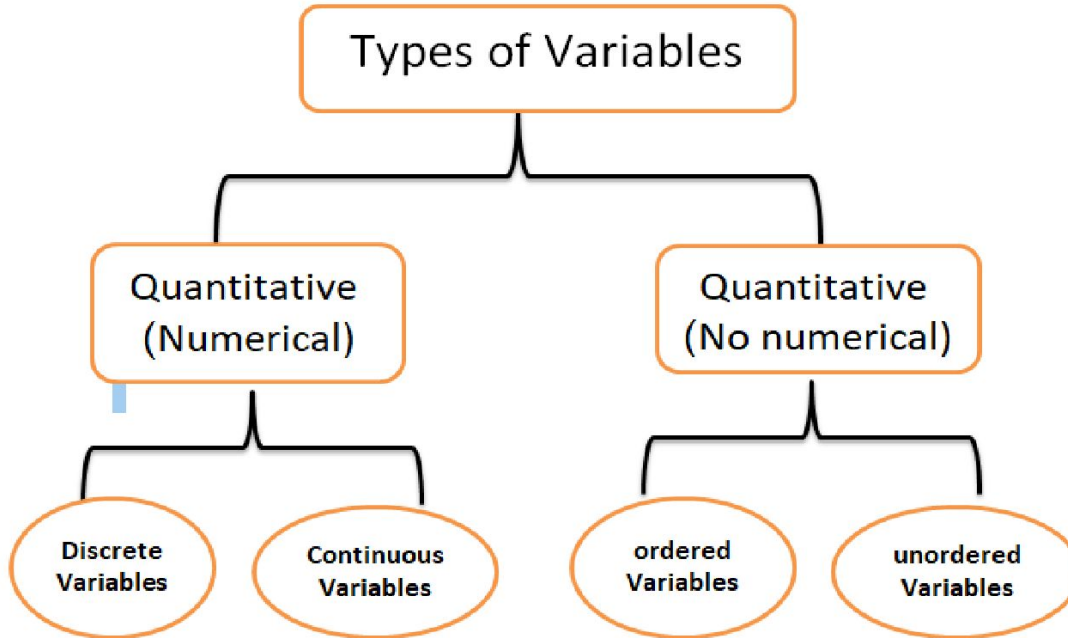


Some principle concepts

- **A population:** consists of all subjects (human or elements or otherwise) that are studied.
- **A sample:** is a subset of the population or (is a group selected from a population).



Some basic concepts



Some basic concepts

Types of Variables

Qualitative Variables

are variables that can be placed into distinct **categories**, according to some characteristic or attribute.

For example: Gender ,Marital status ,Color.....etc

Quantitative variables

are **numerical** and can be ordered or ranked.

For example: Age ,Height , Weight ,temperatureetc

Some basic concepts

Quantitative variables classified into two groups

```
graph TD; A[Quantitative variables classified into two groups] --> B[Discrete Variables]; A --> C[Continuous Variables]; B --- D[assume values that can be counted. For example: Number of children in a family, Number of student in classroom, Number of DVDs rented each day .....etc]; C --- E[assume an infinite number of values between any two specific values. For example: Temperature, Height, Weight, Time .....etc];
```

Discrete Variables

assume values that can be **counted** .

For example:

- Number of children in a family ,
- Number of student in classroom,
- Number of DVDs rented each dayetc

Continuous Variables

assume **an infinite number** of values between any two specific values.

For example:

- Temperature ,
 - Height
 - Weight
 - Timeetc
-

Presentation Methods of Data

Presentation Methods of Data

Graphical Methods

- Bar charts
- Histogram
- Pie charts
- Box-Plot
- Frequency Curve
- Broken Line

Tabular Methods

- Frequency distributive table
 - R. frequency distribution table
 - P. frequency distribution table
 - A. Cumulative Frequency table
 - D. Cumulative Frequency table
-

For examples of graphical Methods

Figure 1: Bar Chart of Subjects in Treatment Groups

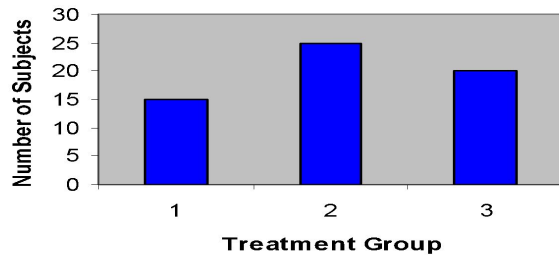


Figure 2: Pie Chart of Subjects in Treatment Groups

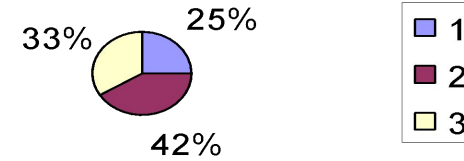
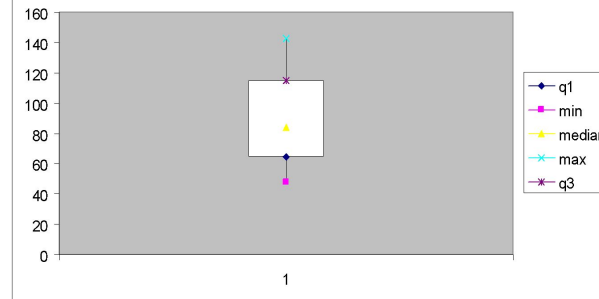
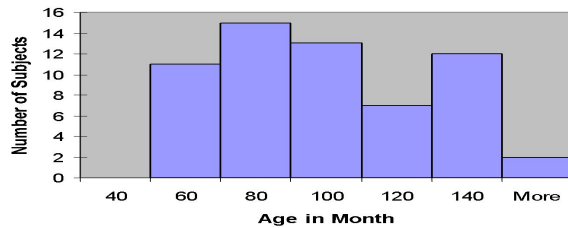
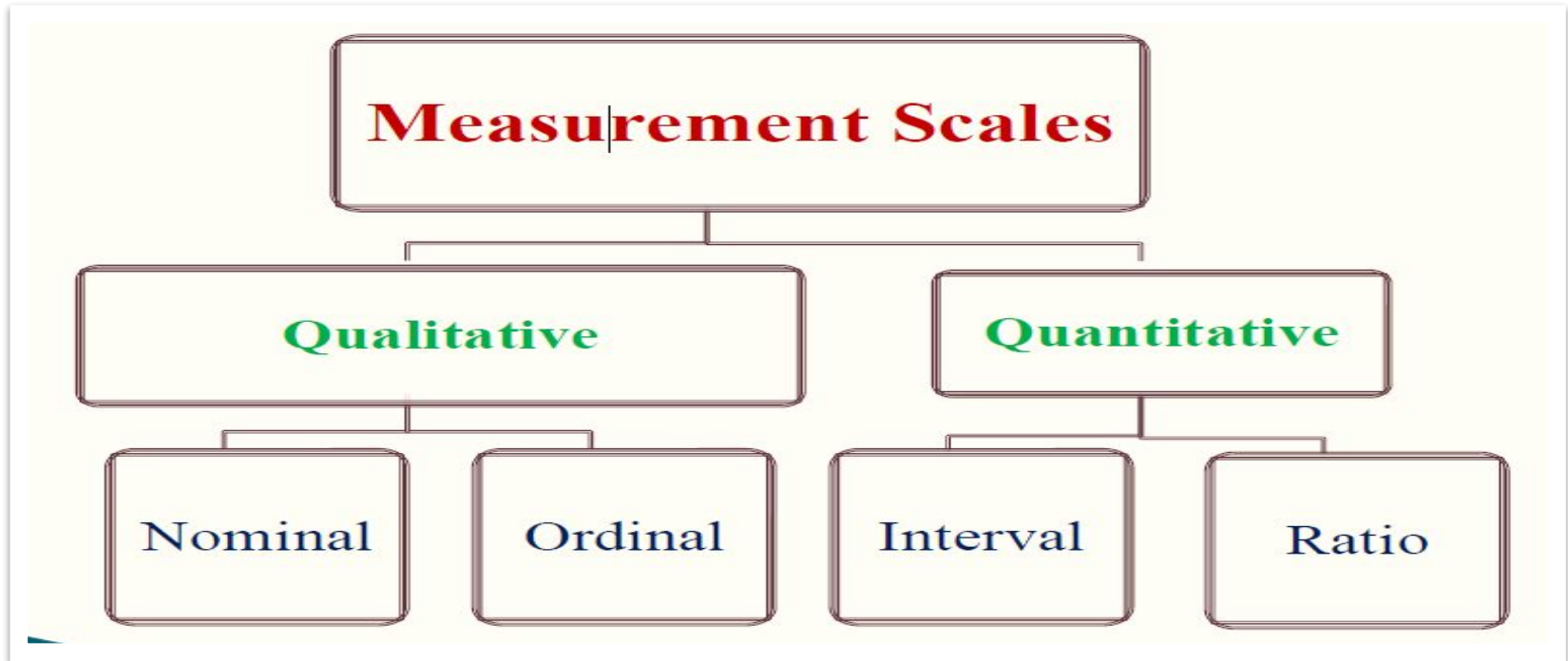


Figure 3: Age Distribution



Measurement System



Measurement System

Measurement scale of qualitative

Nominal level

- Use only for classification without order
- The operations (+, -, *, /) have not meaningful
- The variables are text not numbers. May be take values but the operation on it not meaningful.
- For example: gender, names, cases numbers etc.

Ordinal level

- Use to classifies and order data in categories.
- The operations (+, -, *, /) have not meaningful .
- For example, education level (low, moderate, high). Can be compared for equality, or greater or less. Grad of course (A,B,C,D), rating of scale.

Measurement System

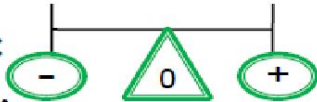
Measurement Scale of Quantitative

Interval level

ranks data and precise differences between units of measure do exist ,however there is **no meaningful zero.**

For example:

Temperature ,
IQ test...etc



Ratio level

possesses all the characteristics of interval and there exist **a true zero.**



For example:

Height , Weight, Time,
Salary , Age ...etc

Measurement System

Measurement scale of quantitative

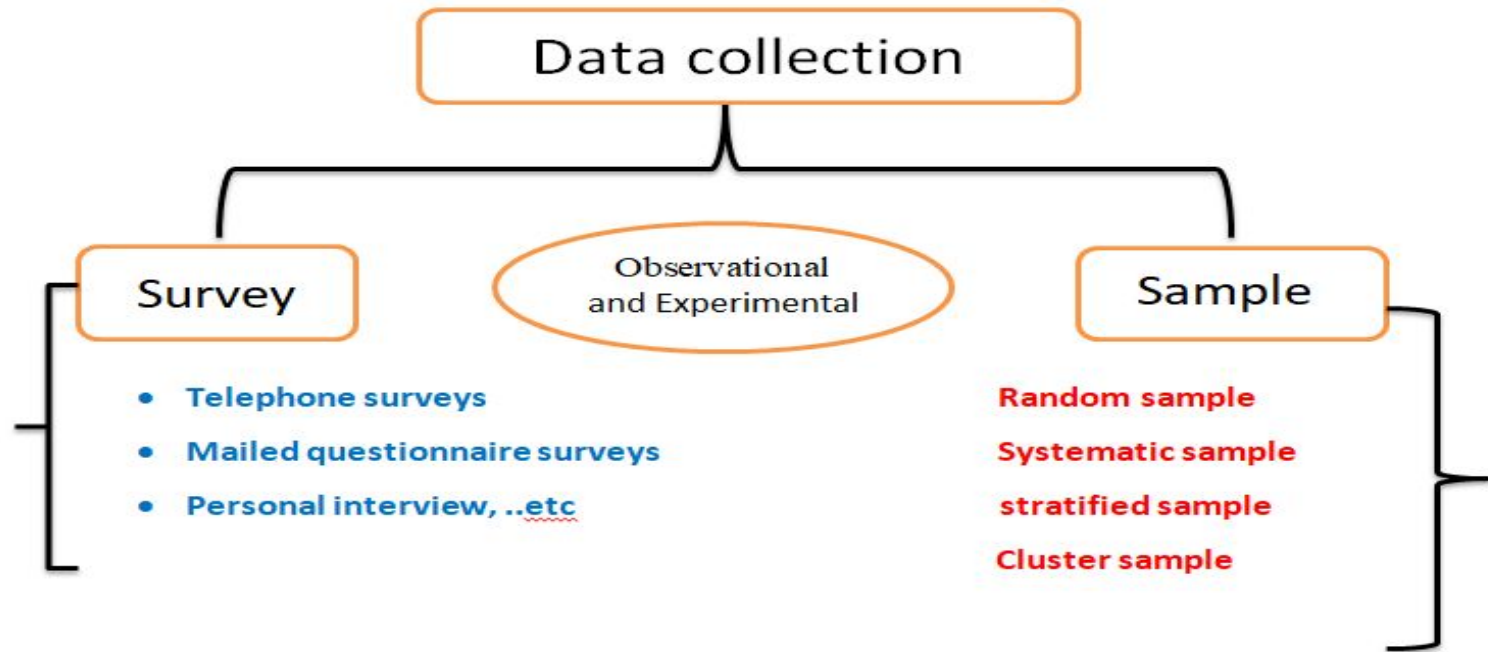
Interval level

- Values of the variable are ordered as in Ordinal,
- differences between units do exist.
- however, the scale is not absolutely fixed.
- For example, Calendar dates and temperatures.
- Zero here has a relative meaning only.

Ratio level

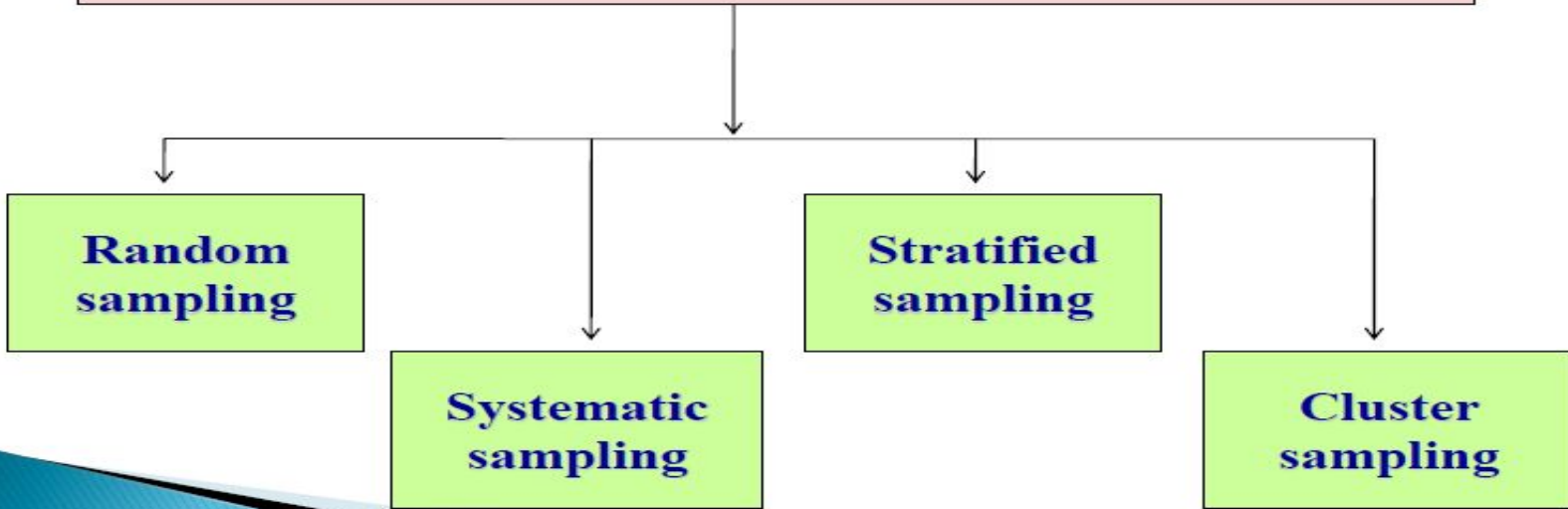
- Her variables take all properties of pervious level scales plus Zero here is real,
- Also, all variables take numerical values e.g. age, weight, size. In addition, all operations (+, -, *, /) have meaningful.
- The ratio is the highest level of measurement.

Methods of Data collection



Some Types of Samples

Some Sampling Techniques



● Random sample:

are selected by using chance methods or random numbers.
For example

79	41	71	93	60	35	04	67	96	04	79	10	86
26	52	53	13	43	50	92	09	87	21	83	75	17
18	13	41	30	56	20	37	74	49	56	45	46	83
19	82	02	69	34	27	77	34	24	93	16	77	00
14	57	44	30	93	76	32	13	55	29	49	30	77
29	12	18	50	06	33	15	79	50	28	50	45	45
01	27	92	67	93	31	97	55	29	21	64	27	29
55	75	65	68	65	73	07	95	66	43	43	92	16
84	95	95	96	62	30	91	64	74	83	47	89	71
62	62	21	37	82	62	19	44	08	64	34	50	11
66	57	28	69	13	99	74	31	58	19	47	66	89
48	13	69	97	29	01	75	58	05	40	40	18	29
94	31	73	19	75	76	33	18	05	53	04	51	41
00	06	53	98	01	55	08	38	49	42	10	44	38
46	16	44	27	80	15	28	01	64	27	89	03	27
77	49	85	95	62	93	25	39	63	74	54	82	85
81	96	43	27	39	53	85	61	12	90	67	96	02
40	46	15	73	23	75	96	68	13	99	49	64	11

Q: Select random sample of 15 subjects out of 85 subjects:

A: 12, 27, 75, 62, 57, 13, 31, 06, 16, 49, 46, 71, 53, 41, 02

- **Systematic sample**

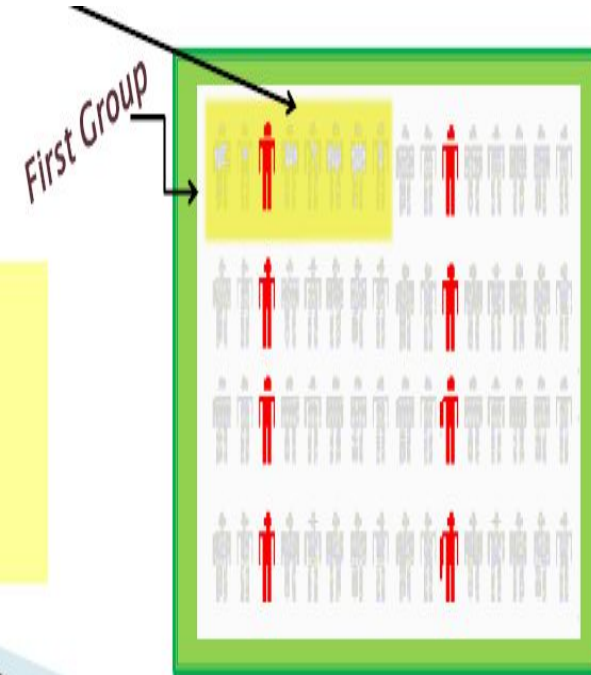
are obtained by numbering each element in the population and then selecting the k^{th} value. **For example,**

- Determine the sample size: n
- Determine the systematic period k , where $k = N/n$
- Randomly Select any number in k
- Add k to this selected number to obtain the next value.

$$N = 64$$

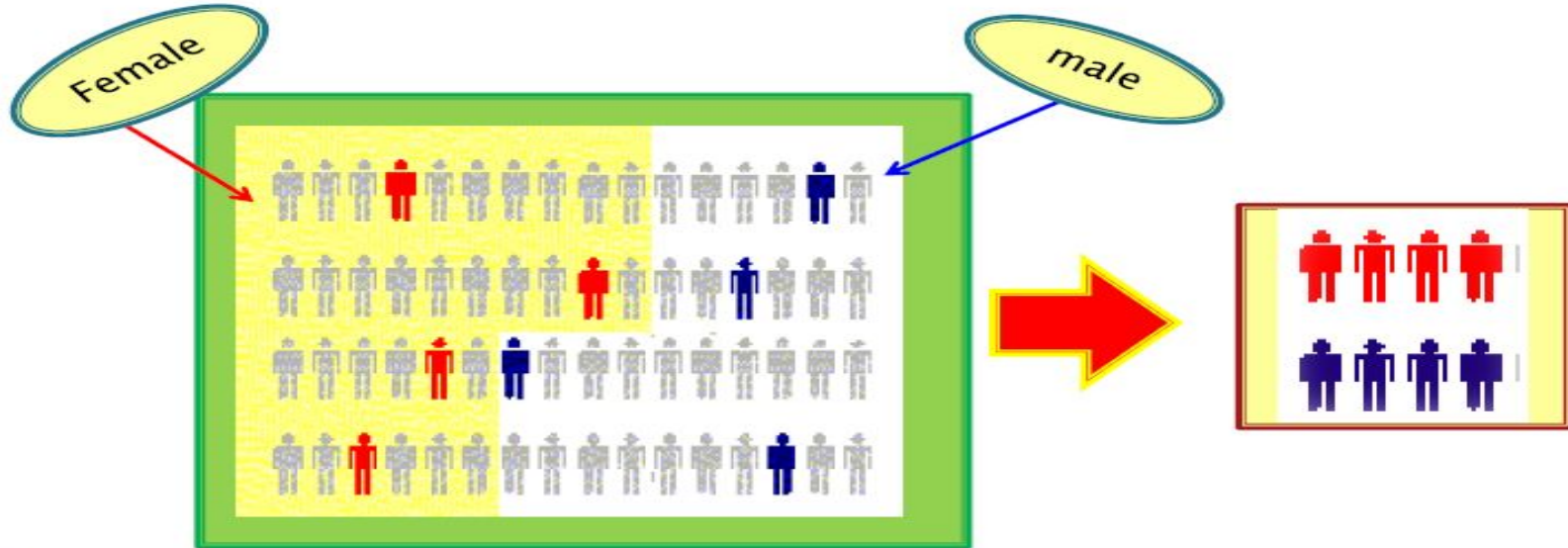
$$n = 8$$

$$k = 64/8 = 8$$



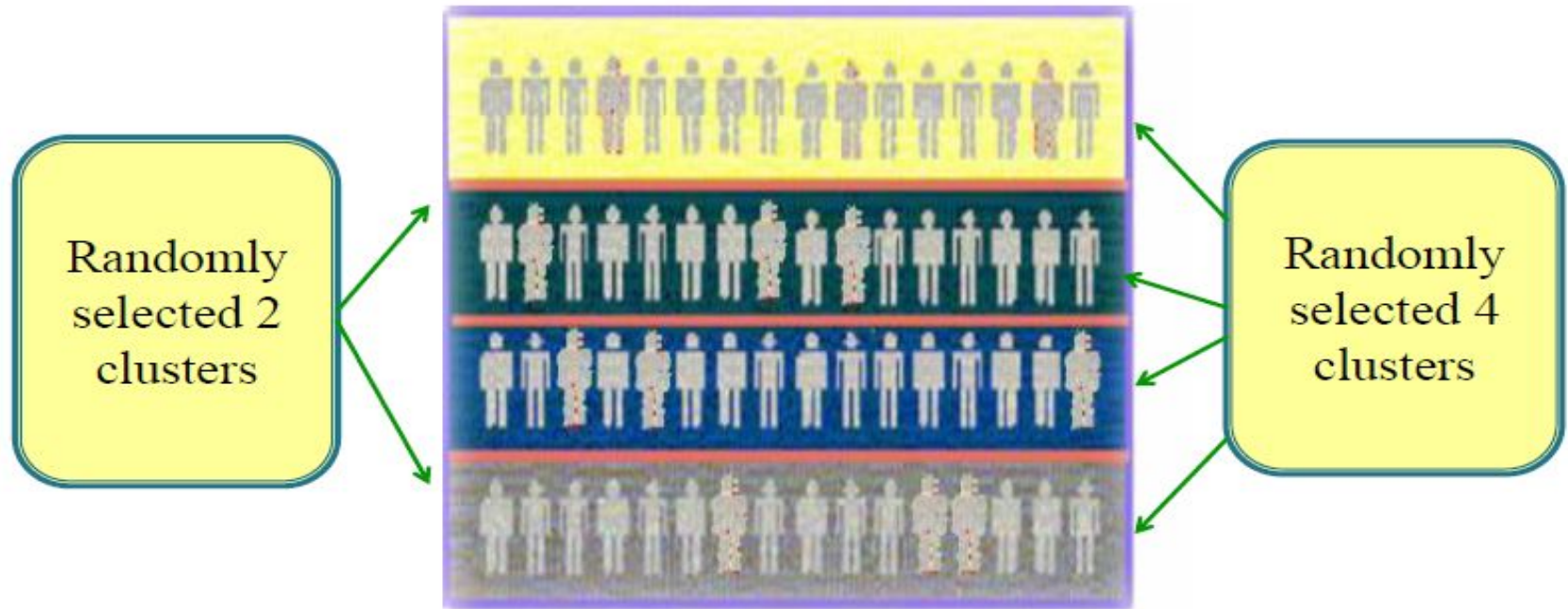
- **Stratified sample:** are selected by dividing the population into groups (strata) according to some characteristic and then taking samples from each group. **For example,**

A researcher select a random sample from each gender to check their blood pressure



- **Cluster sample**

are selected by dividing the population into groups and then taking samples of the groups. **For example,**



Some Types of studies:

Types of Studies

Observational Study

The researcher merely observes what is happening or what has happened in the past and tries to draw conclusions based on these observations.

For example:

- people who sleep 8 hours report better health.
- A researcher counts the number of people living in each house in specific a street .

Experimental Study

The researcher manipulates one of the variables and tries to determine how the manipulation influences other variables.

For examples:

- Patients were randomly assigned to two groups was given drug A and the other group was given drug B to determine if the drug has an effect on patient's blood pressure.
- An Instructor has Three Teaching method ,he want to apply a best method by seeing students grades.

Variables of study:

For example

Independent	temperature of water	exercise
Dependent	time to cook an egg	health

Note :Statistical studies usually include **one or more independent variables** and **one dependent variable**.

Variables of study:

Any Experiment has 2 Variables

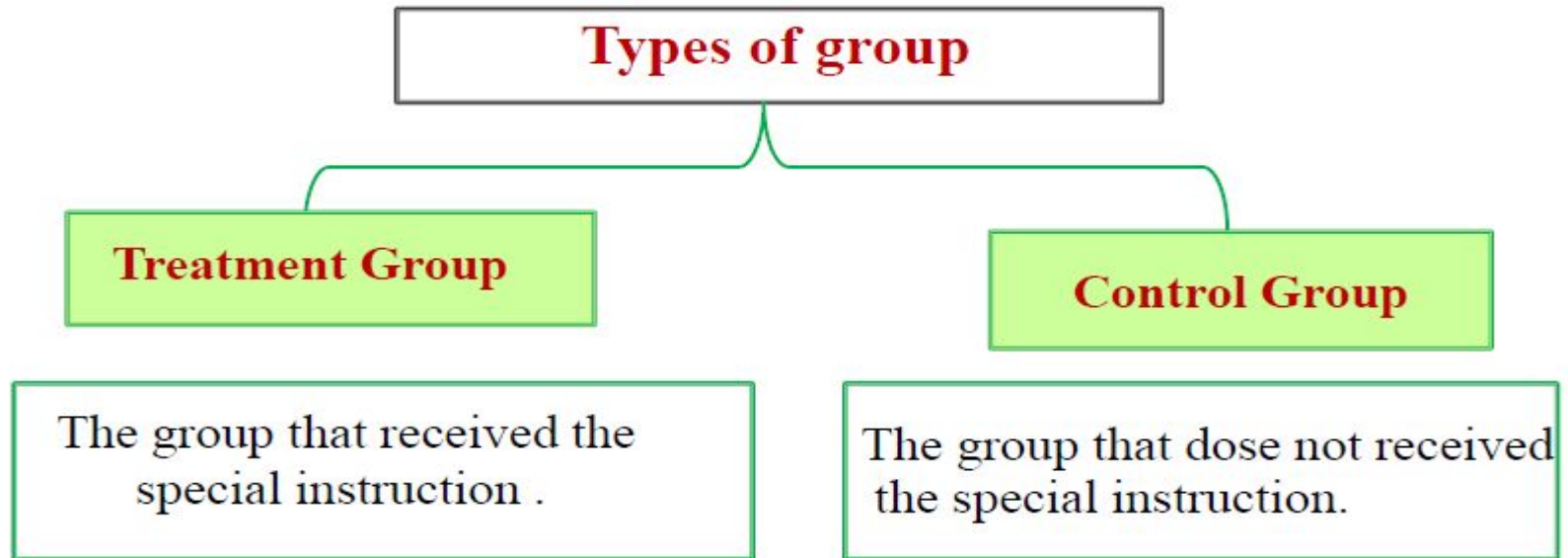
**Independent Variable
or
Explanatory Variable**

(or input) variable is the one that is being manipulated by the researcher.

**Dependent Variable
or
Outcome Variable**

is the resultant variable

Experimental study:



Some Statistical Data Analysis Software

The Best Statistical Software tools:

- IBM SPSS Statistics
 - SAS/STAT
 - Stata
 - Minitab
 - Graph Pad Prism
 - SmartPLS
-

..Thank you for all
