UNIT-III
Statistics & Biostatistics
(Part-1)

Compiled By;
Dr. Rajesh K. Kesharwani,
Dr. Rudra P. Ojha
Nehru Gram Bharati (Deemed to be University), Prayagraj

(The content is only for academic purpose)

DEFINITION

STATISTICS: *Principles and methods for collection, presentation, analysis and interpretation of numerical data.*

BIOSTATISTICS: *Tools of statistics applied to the data that is derived from biological sciences.*

Descriptive statistics: *methods of producing quantitative summaries of information.*

- ***Measures of central tendency***
- ***Measures of dispersion***

Inferential statistics: *methods of making generalizations about a larger group based on information about a subset (sample) of that group.*

3

DATA

COLLECTIVE RECORDING OF OBSERVATIONS

TYPES OF DATA

Depending on the source of Collection:

Primary data: Interview
                  Examination
                  Questionnaire

Secondary data: Hospital records
                  Census data

DATA

COLLECTIVE RECORDING OF OBSERVATIONS

TYPES OF DATA

Depending on the source of Collection:

Primary data: Interview
                    Examination
                    Questionnaire

Secondary data: Hospital records
                    Census data

DATA

QUALITATIVE DATA

- SUBJECTS WITH SAME CHARECTERISTICS ARE COUNTED TO FORM SPECIFIC GROUPS OR CLASSES.

e.g. ATTACKED, DIED, MALES

QUANTITATIVE DATA

- THEY HAVE A MAGNITUDE.

- CHARACTERISTICS AND FREQUENCY VARY

e.g. HEIGHT OF A PERSON

    WEIGHT OF A PERSON

Scales of Measurement of Data

Four principal scales are used to measure data:

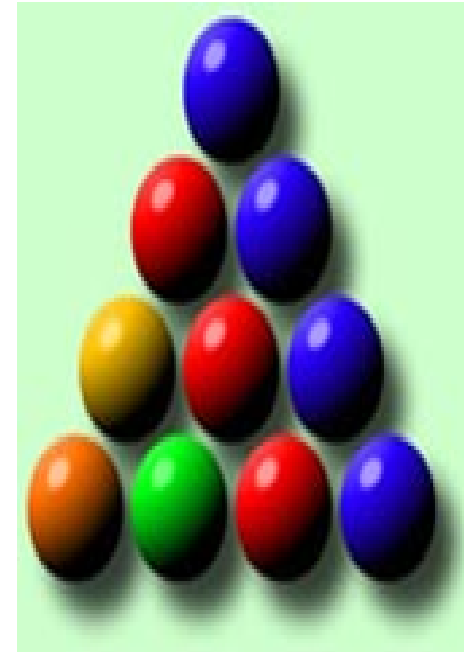1. Nominal scale
2. Ordinal scale
3. Interval scale
4. Ratio scale

Ratio — Absolute zero

Interval — Distance is meaningful

Ordinal — Attributes can be ordered

Nominal — Attributes are only named; weakest
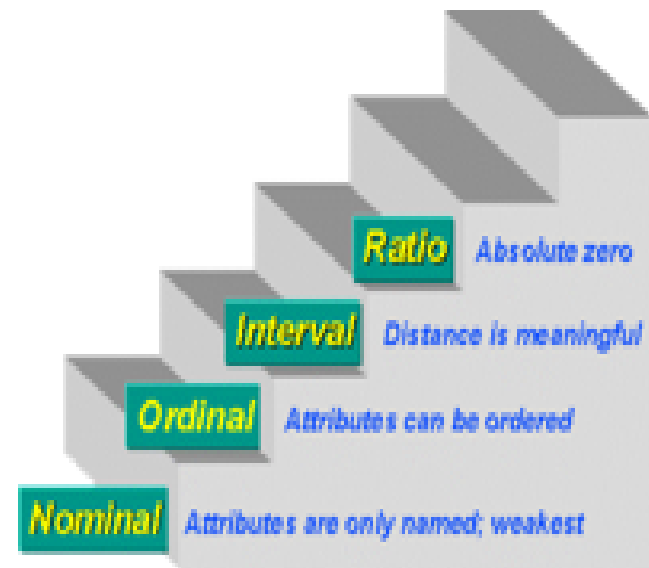
1. Nominal scale

This scale uses names or tags to distinguish one measurement from another. It does not imply magnitude of individual measurements.

Eg. Classification of sex
Classification of religion

2. ORDINAL SCALE

It is like nominal scale but there exists an implicit graded order relationship among the categories.

Eg. Pain measured as:

Mild   – 1
Moderate – 2
Severe   – 3

or

Socio economic status measured as

Class I
Class II
Class III
Class IV

Hottest

Hotter

Hot

The " Hot" Scale
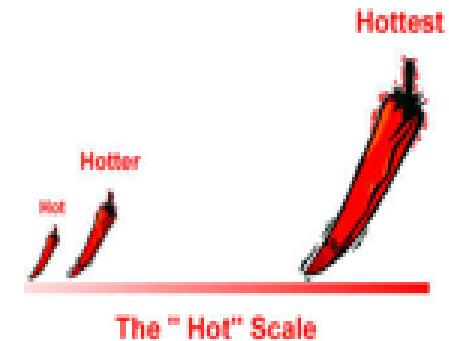
3. INTERVAL SCALE

A numerical unit of measurement is used in this scale.

The difference between any two measurements can be clearly identified in terms of an interval between two points of scale.

This scale has no true zero point.

Eg. Measurement of body temperature in degrees of Celsius.

4. Ratio scale

It is as same as interval scale in every aspect except that measurement begins at a true or absolute zero.

E.g. Weight in Kgs,

Height in Mts.

Are you the right weight for your height?
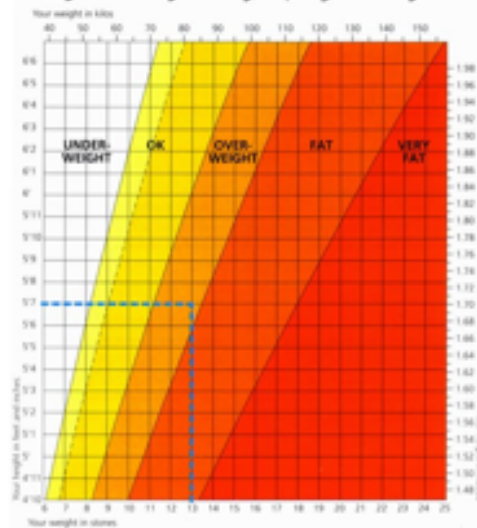
There cannot be negative measurements.

MEASURES OF CENTRAL TENDENCY

Central Tendency:

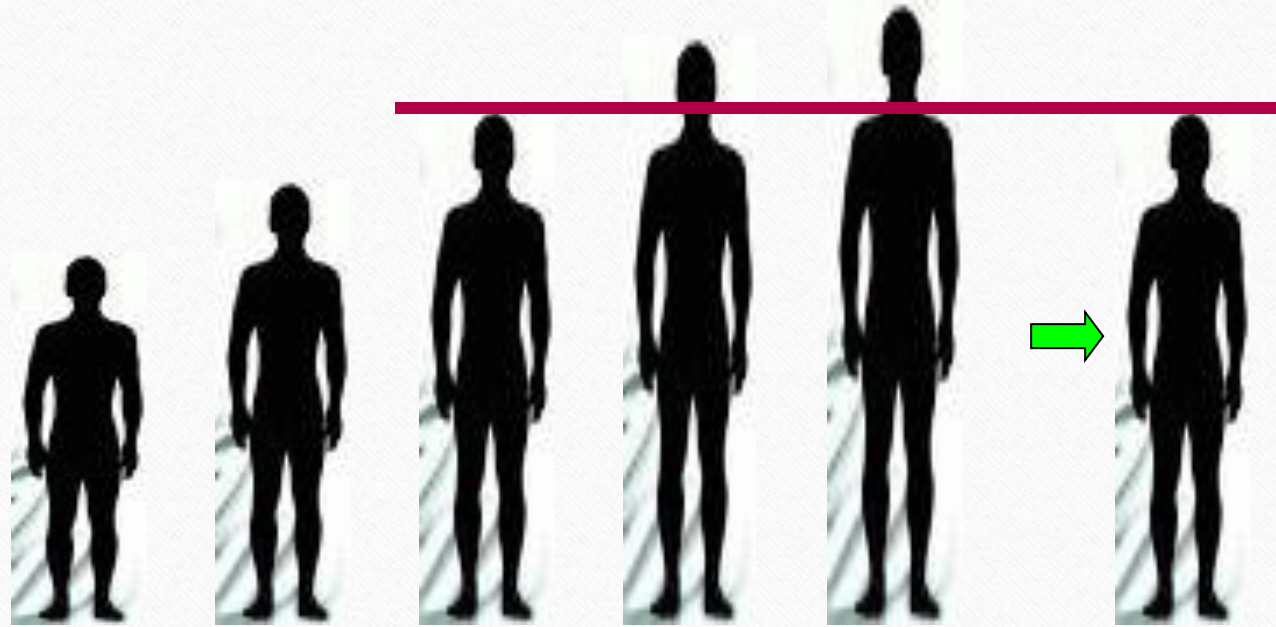General tendency for the observations from a sample to cluster around a central value.

OBJECTIVES:

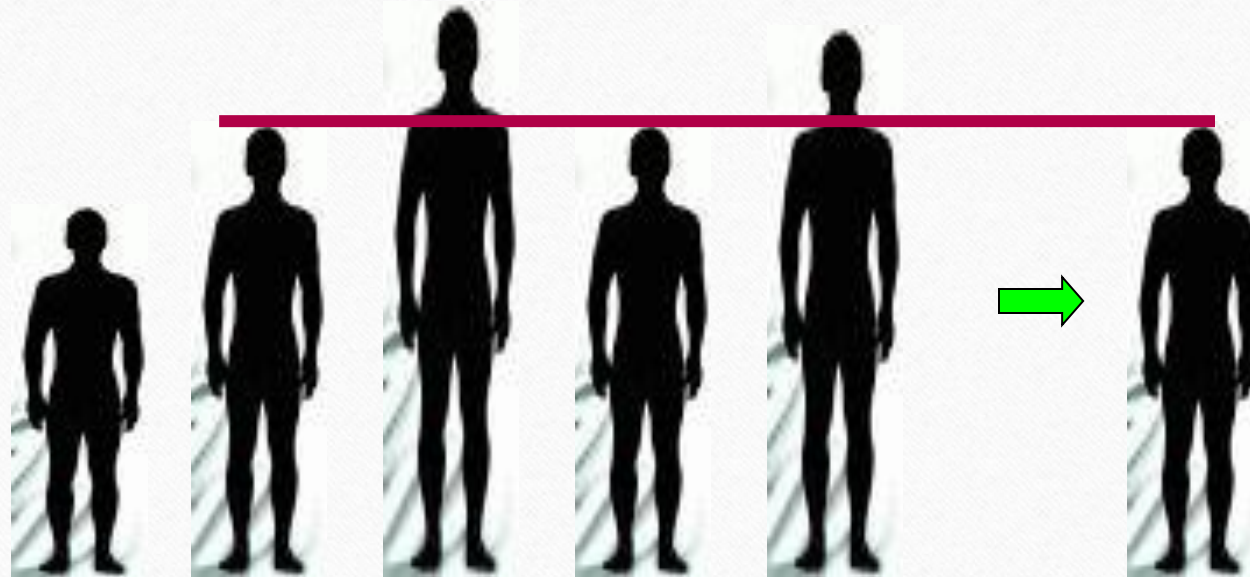- To condense the entire mass of data

- To facilitate comparison
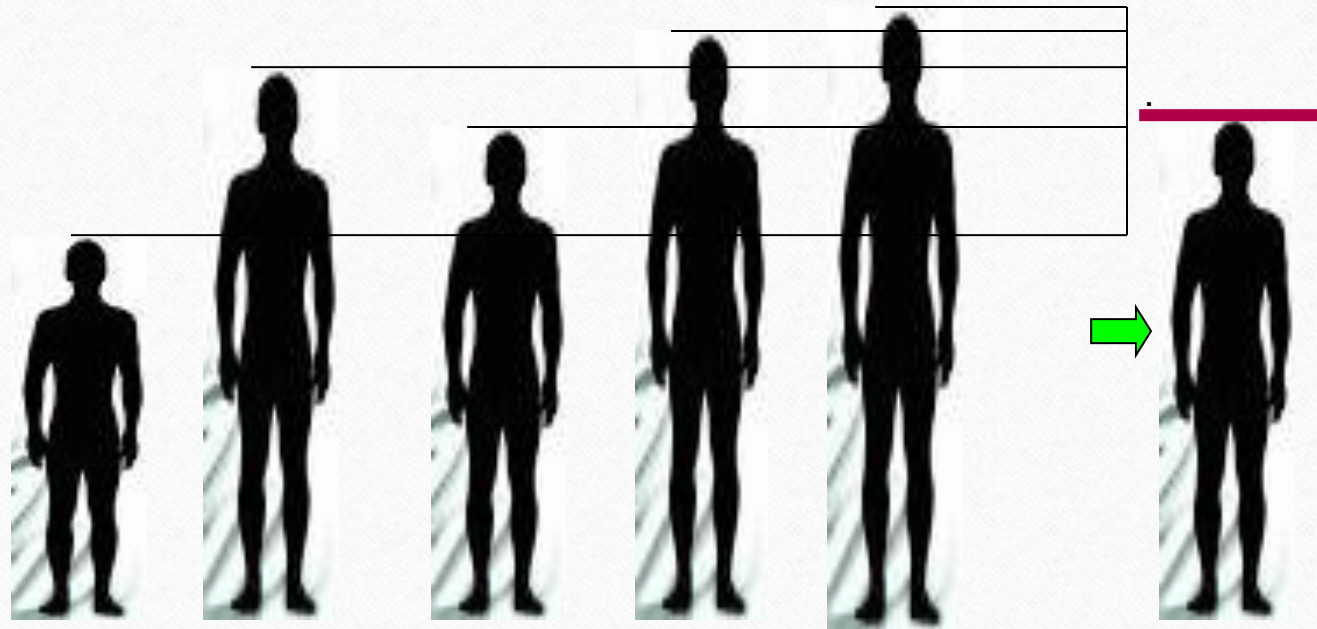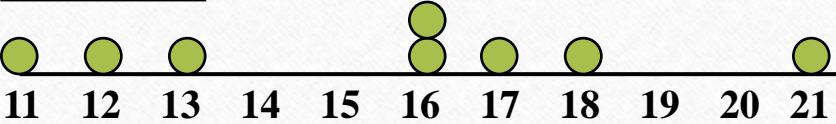
MEAN

MEDIAN

MODE

MEASURES OF DISPERSION

Range

- The difference between the largest and smallest values in a distribution

- EXAMPLE: 1, 2, 2, 2, 3, 4, 5, 6, 7, 8
- RANGE = 8-1 = 7

Standard deviation

- Measure of dispersion (or scatter) of the values about the mean

  - If the numbers are near the mean, variance is small

  - If numbers are far from the mean, the variance is large

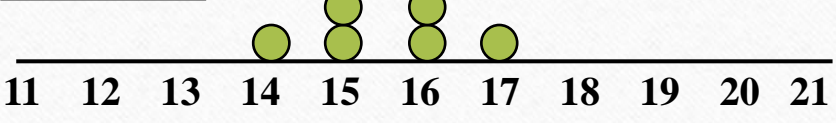# Comparing Standard Deviations

Data A

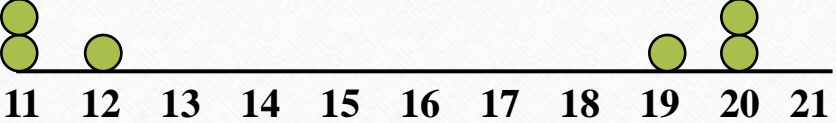11 12 13 14 15 16 17 18 19 20 21

Mean = 15.5
S = 3.338

Data B

11 12 13 14 15 16 17 18 19 20 21

Mean = 15.5
S = .9258

Data C

11 12 13 14 15 16 17 18 19 20 21

Mean = 15.5
S = 4.57

STANDARD DEVIATION

- EX: 1,2,3,4,5

- MEAN= 1+2+3+4+5/5= 15/5=3
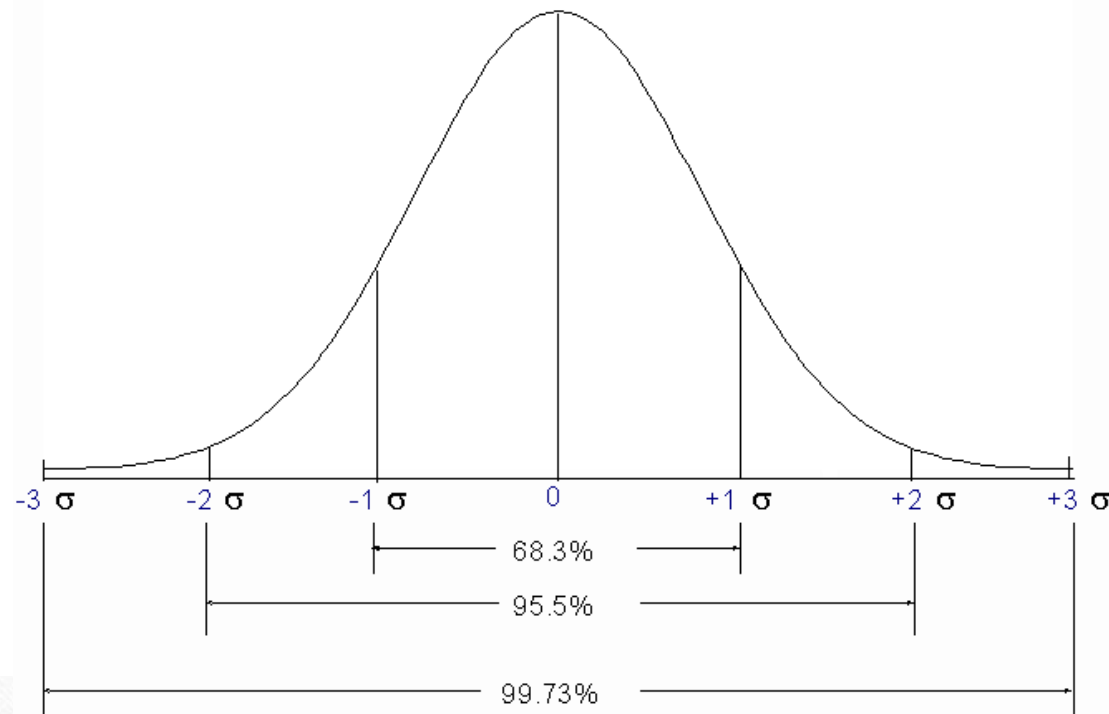
$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{10}{5-1}} = 1.58$$

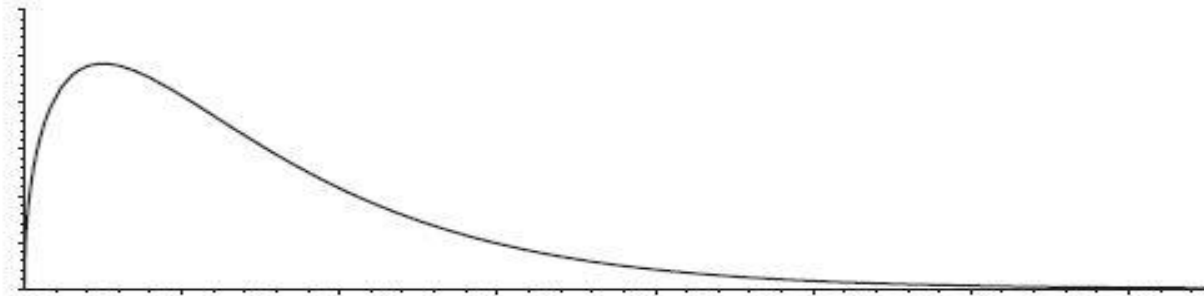$$= \boxed{3 \pm 1.58}$$

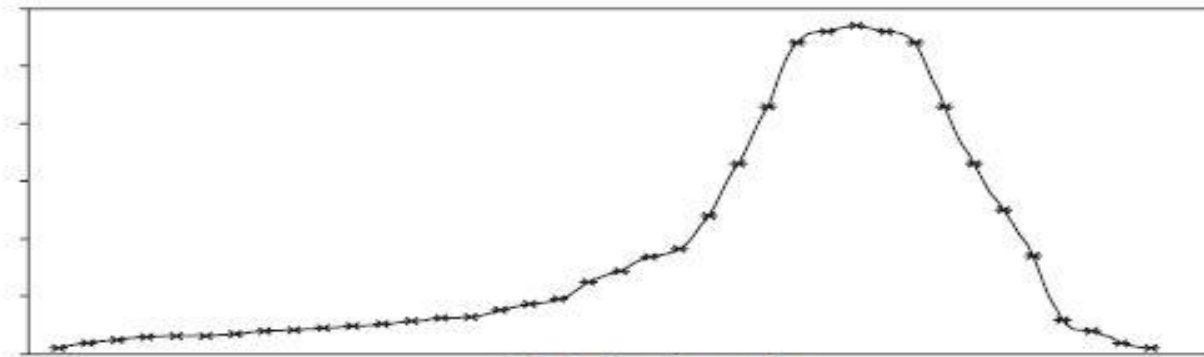| x	$\bar{x}$	$x-\bar{x}$	$(x-\bar{x})^2$
1	3	-2	4
2	3	-1	1
3	3	0	0
4	3	1	1
5	3	2	4

$$\Sigma(x-\bar{x})^2 = 10$$

NORMAL CURVE
[GAUSSIAN CURVE]

(b) Right-skewed

(c) Left-skewed

HEIGHT OF STUDENTS IN CLASS (CMS)

140　　145　　150　　155　　160　　165　　170
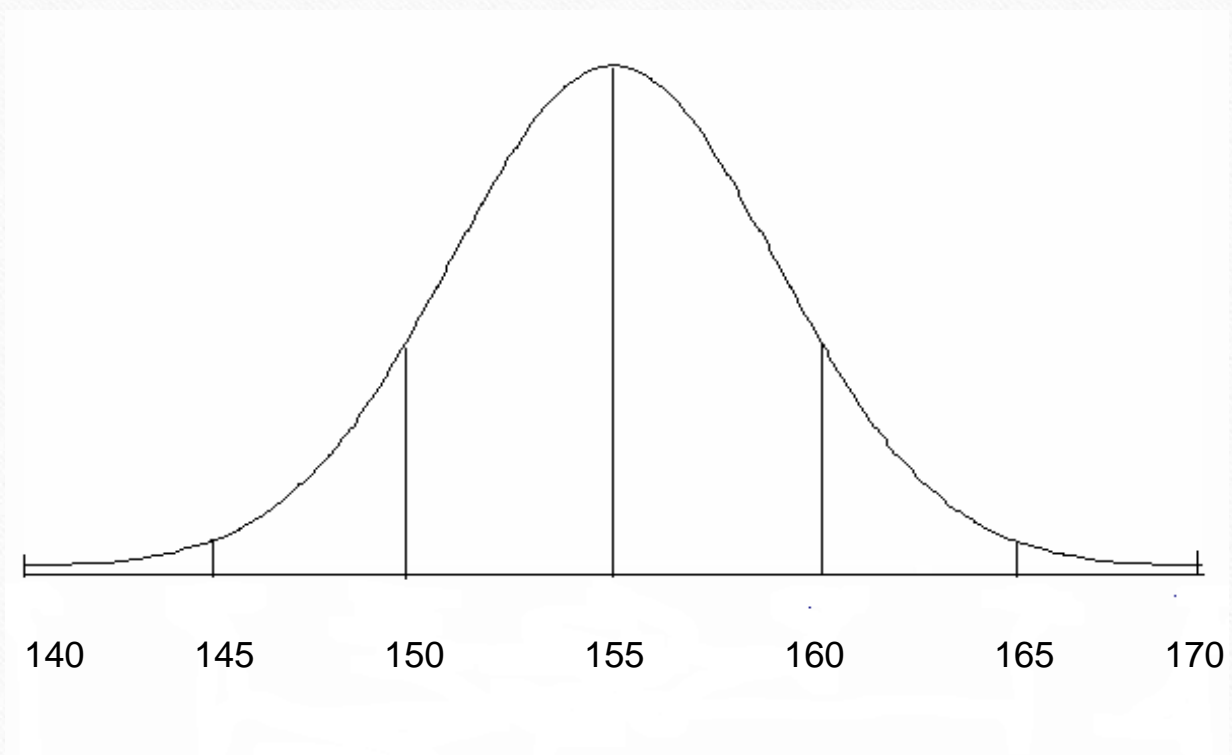
PROPERTIES OF NORMAL CURVE
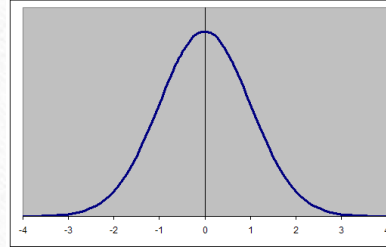
- *BELL SHAPED AND SYMMETRICAL ABOUT THE MID POINT*

- *MEAN, MEDIAN AND MODE COINCIDE*

- *MAXIMUM NUMBER OF OBSERVATIONS ARE AT VALUE CORRESPONDING TO THE MEAN AND OBSERVATIONS GRADUALLY DECREASE ON EITHER SIDES.*

MEAN ± 1 SD
COVERS 68.3% OF
OBSERVATIONS

MEAN ± 2 SD
COVERS 95.4% OF
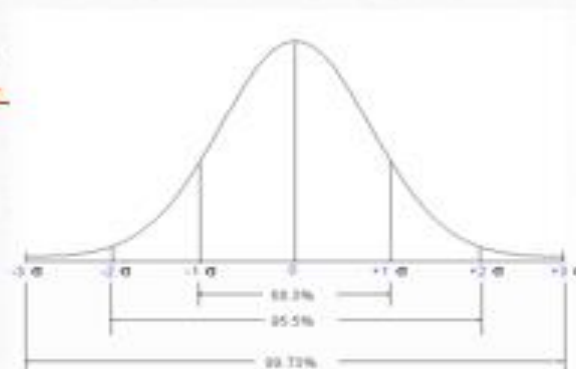OBSERVATIONS

MEAN ± 3 SD
COVERS 99.7% OF
OBSERVATIONS
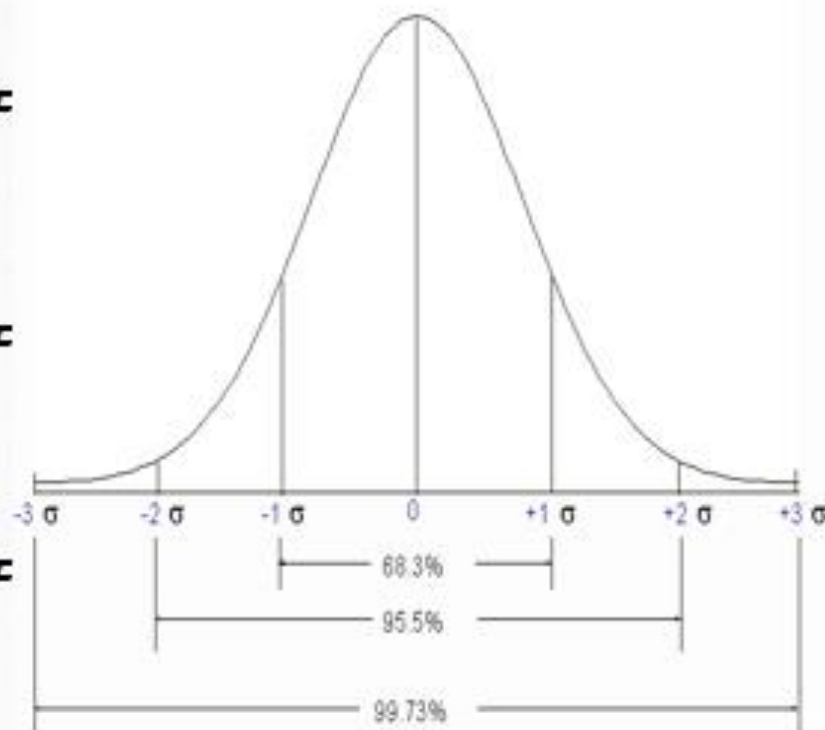
NORMAL CURVE TELLS THE PROBABILITY OF OCCURRENCE BY CHANCE

OR

HOW OFTEN AN OBSERVATION CAN OCCUR NORMALLY IN A POPULATION.

STANDARD NORMAL CURVE

▪ TOTAL AREA IS 1

▪ MEAN IS 0, STANDARD DEVIATION IS 1

▪ MEAN, MEDIAN AND MODE COINCIDE

▪ CURVE HAS NO BOUNDARIES IN EITHER DIRECTION & NEVER TOUCHES THE BASELINE

References

- Kim and Dailey. "Biostatistics for oral health care"

- Park and Park Text book of Community Medicine 23rd edition

- B K Mahajan. Biostatistics

- Armitage. Medical Biostatistics

- Edex Online Resources