

MATHEMATICS IN THE MODERN WORLD



BASIC CONCEPTS IN STATISTICS

STATISTICS

It is a **collection** of methods for planning experiments, **obtaining data**, and then **analyzing, interpreting** and drawing conclusions based on the data.



BASIC TERMS IN STATISTICS

- **Data** are the values that the variables can assume.
- A **variable** is a characteristics that is observable or measurable in every unit of universe.
- **Population** is the set of all possible values of a variable.
- **Sample** is a subgroup of a population.

CLASSIFICATION OF VARIABLES

- **Qualitative Variables**

- ❖ words or codes that represent a class or category.
- ❖ express a categorical attribute
 - ✓ gender
 - ✓ religion
 - ✓ marital status
 - ✓ Highest educational attainment

CLASSIFICATION OF VARIABLES

- **Quantitative Variables**
 - ❖ number that represent an amount or a count.
 - ❖ numerical data, sizes are meaningful and answer questions such as "how many" or "how much".
 - ✓ Height
 - ✓ Weight
 - ✓ household size
 - ✓ number of registered cars

Quantitative Variables classified as:

- **Discrete Variables**

- ❖ data that can be counted
 - ✓ number of days
 - ✓ Number of siblings
 - ✓ Usual number of text messages sent in day
 - ✓ Usual daily allowance in school

Quantitative Variables classified as:

- **Continuous Variables**

- ❖ it can assume all values between any two specific values like 0.5, 1.2. etc. and data that can be measured.

- ✓ weight
 - ✓ height
 - ✓ body temperature

Levels of Measurement

- **Nominal Level**

- ❖ This is characterized by data that consists of names, labels, or categories only.

- ✓ gender
 - ✓ most preferred color
 - ✓ usual sleeping time
 - ✓ civil status

Levels of Measurement

- **Ordinal Level**

- ❖ This involves data that arranged in some order, but differences between data.
 - ✓ happiness index for the day
 - ✓ highest educational attainment
 - ✓ the rankings of tennis player
 - ✓ Academic Excellence Awards

Levels of Measurement

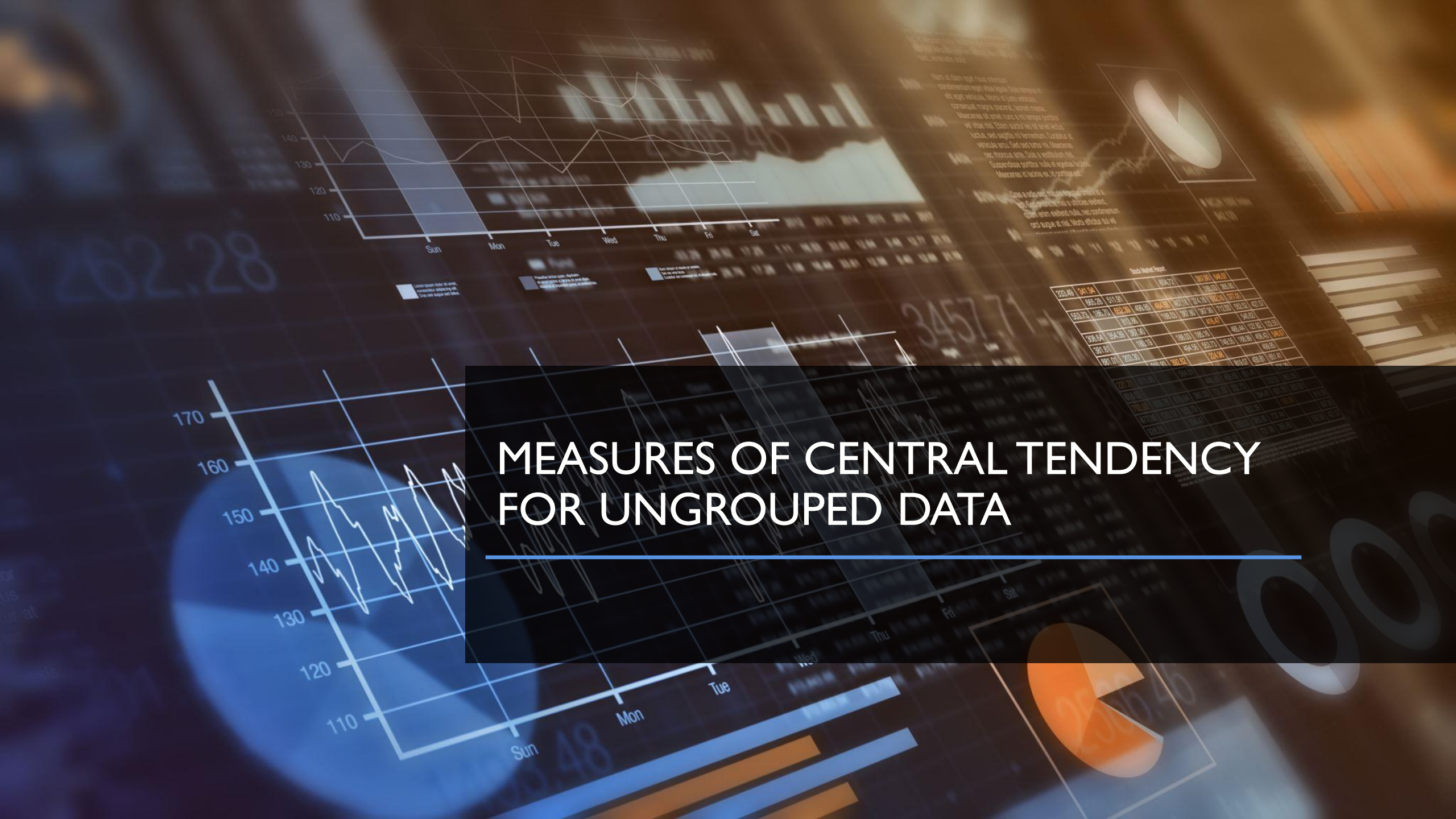
- **Interval Level**
 - This is the same in ordinal level, with an additional property that we can determine meaningful amounts of differences between the data.
 - ✓ body temperature
 - ✓ Intelligence Quotient

Levels of Measurement

- **Ratio level**
 - This is an interval level modified to include the inherent zero starting point.
 - It possesses a meaningful absolute, fixed zero point and allows all arithmetic operations.
 - ✓ number of siblings
 - ✓ weight
 - ✓ height

Four Basic Methods of Sampling

- **Random Sampling**, this is done by using chance or random numbers.
- **Systematic Sampling**, this is done by numbering each subject of the population and then selecting n th number.
- **Stratified Sampling**, If a population has a distinct groups, it is possible to divide the population into these groups and to draw SRS's from each of the groups
- **Cluster Sampling**, This method uses intact groups called clusters.



MEASURES OF CENTRAL TENDENCY FOR UNGROUPED DATA

MEAN

- The mean is the most commonly used measure of central tendency. When we speak of average, we always refer to the mean.

$$\bar{x} = \frac{\sum x}{N}$$

EXAMPLES:

- Six friends in a biology class of 20 students receives test grades of 92, 84, 65, 76, 88, and 90. Find the mean of these test scores.

EXAMPLES:

- Six friends in a biology class of 20 students receives test grades of 92, 84, 65, 76, 88, and 90. Find the mean of these test scores.

$$\begin{aligned}\bar{x} &= \frac{\sum x}{N} = \frac{92 + 84 + 65 + 76 + 88 + 90}{6} \\ &= \frac{495}{6} \Rightarrow \bar{x} = 82.5\end{aligned}$$

EXAMPLES:

- The ages of five contestants in a Statistics Quiz Bee are the following: 18, 17, 18, 19, and 18. Find their average age.

MEDIAN

- The median is the midpoint of the data array. Before finding this value, the data must be arranged in order, from least to greatest or vice versa. The median will either be a specific value or will fall between two values.



MEDIAN

- The median is the midpoint of the data array. Before finding this value, the data must be arranged in order, from least to greatest or vice versa. The median will either be a specific value or will fall between two values.

$$\tilde{x} = x_1, x_2, x_3, x_4, x_5$$

$$\tilde{x} = x_1, x_2, x_3, x_4, x_5, x_6$$

$$\tilde{x} = \frac{x_3 + x_4}{2}$$

EXAMPLES:

- Seven mothers were selected and given a blood pressure check. their blood pressure were recorded below.

135 121 119 116 130 121 131

EXAMPLES:

- ❑ Eight novels were randomly selected and the numbers of pages were recorded as follows:

415, 398, 402, 400 420, 415, 407, 425

MODE

- It is the value that occurs most often in the data set.
- The number/value/observation in a data set which appears the most number of times.

EXAMPLES:

- Find the mode of the given data set: 15, 28, 25, 48, 22, 43, 39, 44, 43, 49, 34, 22, 33, 27, 25, 22 and 30.

EXAMPLES:

- Find the mode of the given data set: 15, 28, 25, 48, 22, 43, 39, 44, 43, 49, 34, 22, 33, 27, 25, 22 and 30.

Arrange the data set in ascending or descending order.

15, 22, 22, 22, 25, 25, 27, 28, 30, 33, 34, 39, 43, 43, 44, 48, 49.

EXAMPLES:

- The speed of ten stenographers in typing per minute are as follows: 121 110 120 119 112 121 118 115 107 115.

EXAMPLES:

- Find the mode of the given data: 2, 5, 8, 9, 11, 4, 7, 23.

THE WEIGHTED MEAN

- The weighted mean of the n numbers $x_1, x_2, x_3, \dots, x_n$ with the respective assigned weights $w_1, w_2, w_3, \dots, w_n$ is

$$\text{Weighted Mean} = \frac{\sum(x \cdot w)}{\sum w}$$

where $\sum(x \cdot w)$ is the sum of the products formed by multiplying each number by its assigned weight, and $\sum w$ is the sum of all weights.

EXAMPLE:

- Many colleges use the 4-point grading system:

A = 4, B = 3, C = 2, D = 1, F = 0. Find the grade point average of Dillon's Grades, in the given semester course grade.

Course	Grade	Units
English	B	4
History	A	3
Chemistry	D	3
Algebra	C	4

EXAMPLE:

- Many colleges use the 4-point grading system:

A = 4, B = 3, C = 2, D = 1, F = 0. Find the grade point average of Dillon's Grades, in the given semester course grade.

Course	Grade	Units
English	B	4
History	A	3
Chemistry	D	3
Algebra	C	4

$$\text{Weighted Mean} = \frac{\sum(x \cdot w)}{\sum w}$$

$$\begin{aligned}\text{Weighted Mean} &= \frac{(3 \cdot 4) + (4 \cdot 3) + (1 \cdot 3) + (2 \cdot 4)}{14} \\ &= \frac{35}{14} \rightarrow = 2.5\end{aligned}$$

Dillon's GPA For the given semester is 2.5

FREQUENCY DISTRIBUTION

- A frequency distribution, which is a table that list observed events and the frequency occurrence of each observed event, is often used to organized raw data.

Number of Computers per Household

2	0	3	1	2	1	0	4
2	1	1	7	2	0	1	1
0	2	2	1	3	2	2	1
1	4	2	5	2	3	1	2
2	1	2	1	5	0	2	5

Find the mean of the data.

Frequency
Distribution

Number of Computers x	Number of Households f
0	5
1	12
2	14
3	3
4	2
5	3
6	0
7	1
	$N = 40$

Number of Computers per Household

2	0	3	1	2	1	0	4
2	1	1	7	2	0	1	1
0	2	2	1	3	2	2	1
1	4	2	5	2	3	1	2
2	1	2	1	5	0	2	5

Frequency
Distribution

Number of Computers x	Number of Households f
0	5
1	12
2	14
3	3
4	2
5	3
6	0
7	1
	$N = 40$

Find the mean of the data.

$$\bar{x} = \frac{\sum fx}{N}$$

$$\bar{x} = \frac{(0 \cdot 5) + (1 \cdot 12) + (2 \cdot 14) + (3 \cdot 3) + (4 \cdot 2) + (5 \cdot 3) + (6 \cdot 0) + (7 \cdot 1)}{40}$$

Number of Computers per Household

2	0	3	1	2	1	0	4
2	1	1	7	2	0	1	1
0	2	2	1	3	2	2	1
1	4	2	5	2	3	1	2
2	1	2	1	5	0	2	5

Frequency
Distribution

Number of Computers x	Number of Households f
0	5
1	12
2	14
3	3
4	2
5	3
6	0
7	1
	$N = 40$

Find the mean of the data.

$$\bar{x} = \frac{\sum fx}{N}$$

$$\bar{x} = \frac{(0 \cdot 5) + (1 \cdot 12) + (2 \cdot 14) + (3 \cdot 3) + (4 \cdot 2) + (5 \cdot 3) + (6 \cdot 0) + (7 \cdot 1)}{40}$$

$$\bar{x} = \frac{79}{40} \rightarrow \bar{x} = 1.975$$

Number of Computers per Household

2	0	3	1	2	1	0	4
2	1	1	7	2	0	1	1
0	2	2	1	3	2	2	1
1	4	2	5	2	3	1	2
2	1	2	1	5	0	2	5

Frequency
Distribution

Number of Computers x	Number of Households f
0	5
1	12
2	14
3	3
4	2
5	3
6	0
7	1
	$N = 40$

Find the mean of the data.

$$\bar{x} = \frac{\sum fx}{N}$$

$$\bar{x} = \frac{(0 \cdot 5) + (1 \cdot 12) + (2 \cdot 14) + (3 \cdot 3) + (4 \cdot 2) + (5 \cdot 3) + (6 \cdot 0) + (7 \cdot 1)}{40}$$

$$\bar{x} = \frac{79}{40} \rightarrow \bar{x} = 1.975$$

The mean number of computers per household is 2

QUESTIONS?

