

**University of Salahaddin**

**College of Administration and Economy**

**Department of Accounting**

# **Principles of Statistics**

## **First Stage**

**Lecturer:  
Cheman A. Omer**

**Academic Year: 2020 - 2021**

# Chapter One

## Definition of Statistics:

Statistics is the scientific method concerned with collecting, organizing, and analyzing data for a particular phenomenon. As well as making the reasonable decisions.

## The steps of scientific research:



## Types of statistics:

1. **Descriptive statistics:** Consists numerical and graphical technique to summarize and present the data.
2. **Inferential statistics:** Consists hypothesis testing, and estimations for generalizing sample results on population.

## Variable:

Variable is any characteristic that takes different values.

## **Types of Variables:**

**First: Qualitative Variables:** are variables which assume non-numerical values.

### **Example:**

- Gender (male or female).
- Blood group (A , B , AB , O)
- Economic state (v. good , good , bad)
- Level of education (primary , secondary , high school)
- Religious (Muslim , Christian , Jewish , etc.)

## **Types of Qualitative Variables**

**1- Nominal Variable:** There is no order among characteristics of variable.

### **For example:**

- Gender (male , female).
- Blood group (A , B , AB , O)

**2- Ordinal Variable:** There is order among characteristics of variable.

### **Example:**

- Level of education (primary, secondary, high school)
- Economic state (v. good, good, bad)

**Second: Quantitative Variables:** are variables which assume numerical values.

### **For example:**

- The number of students
- Age (54 , 65.6 , 43 , .....).
- Grades of students (78 , 89 , 54 , .....)

## Types of Quantitative Variables:

**1- Discrete Variables:** Usually obtained by counting.

**For example:**

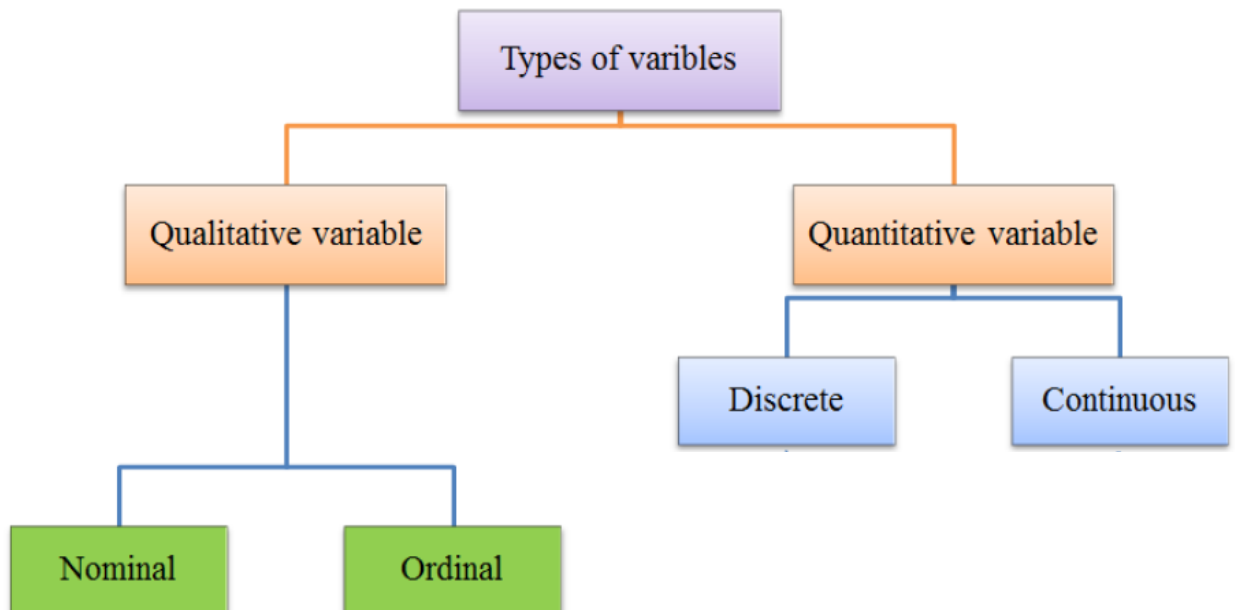
- The number of students
- The number of bacteria on a plate,
- The number of children in a family

**2- Continuous Variables:** Usually obtained by measurement.

**For example:**

- Blood pressure (14 , 17 , 12 , ....)
- Age (54 , 65.6 , 43 , .....).
- Height (154 , 165.6 , 143 , .....).

### Classification for the Types of Variables



## Sources of data:

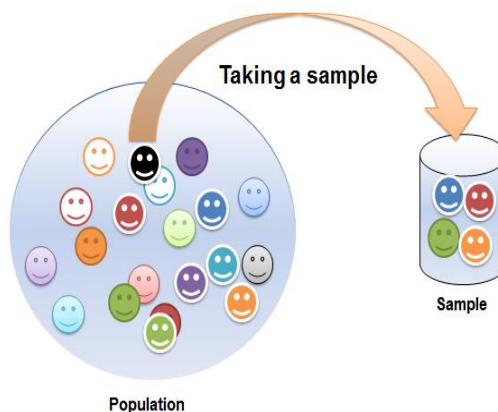
- 1- **Historical sources:** Data and information stored and collected by the organs and institutions of the State or organizations.
- 2- **Field sources:** Data and information can be obtained from their original sources by correspondence, confrontation or any other way of communication.

## Methods of Collection the data:

- 1- **Census method:** the collection of data from every element in a population.
- 2- **Samples method:** the collection of data from a set of elements in a population.

**Population (Statistical population):** It's all the elements or values under a statistical study.

**Sample:** is a part from the population.



**Finite population:** A population is called finite if it is possible to count its individuals.

**Infinite population:** Sometimes it is not possible to count the units contained in the population. Such a population is called infinite

**Homogenous population:** is any population which every element has joint specified characteristics

**Non-homogenous population:** is any population which every element hasn't joint specified characteristics

**Sampling Techniques:** It is the process of selecting a part from the population.

**Types of Sampling Techniques:**

**First: Probability (Random) Sampling:** Taking the sample from the population in such a way that every element in the population has the same probability (chance) at selection.

**Second: Nonprobability Sampling:** Every element in the population does not have equal probability of being chosen.

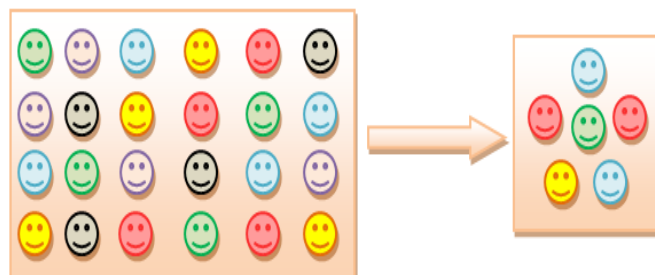
### **Types of Random Sampling**

1. Simple Random Sampling
2. Systematic Sampling
3. Stratified Sampling
4. Multistage Sampling

#### **1- Simple Random Sampling:**

Uses this Technique when the population is homogenous.

**Example: Select 6 students out of 24 students in Statistic class.**



## 2- Systematic Sampling:

- Numbering each element in the population.
- Dividing the population (N) to ( $n^{\text{th}}$ ) group each group contains ( $K=N/n$ ) elements and then one element selection from first group randomly as a starting point.

**Example: Select (6) students out of (24) students by systematic sampling.**

Step 1 : Numbering each element in the population.

Step 2: Dividing the population to ( $n=6$ ) groups

|            |            |            |            |            |            |
|------------|------------|------------|------------|------------|------------|
| 1          | 5          | 9          | 13         | 17         | 21         |
| 2          | 6          | 10         | 14         | 18         | 22         |
| 3          | 7          | 11         | 15         | 19         | 23         |
| 4          | 8          | 12         | 16         | 20         | 24         |
| Group<br>1 | Group<br>2 | Group<br>3 | Group<br>4 | Group<br>5 | Group<br>6 |

Step 3 :  $K = \frac{N}{n} = \frac{24}{6} = 4$

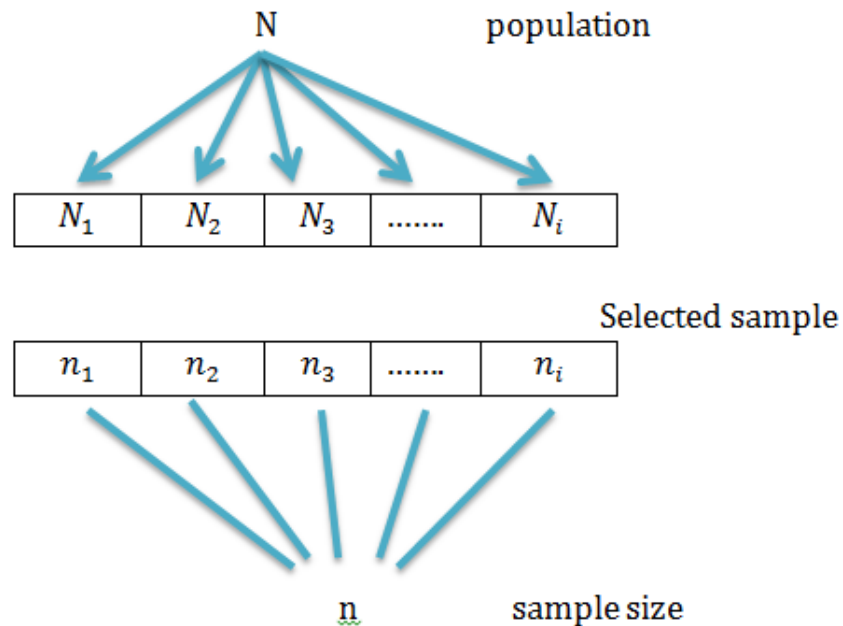
Step 4 : Select a number in group (1) randomly and suppose it is 2 as a starting point.

Step 5 : The second number is  $2+4=\underline{6}$  , the third number is  $6+4=\underline{10}$  , .....

Then your sample is  $2^{\text{nd}}$ ,  $6^{\text{th}}$ ,  $10^{\text{th}}$ ,  $14^{\text{th}}$ ,  $18^{\text{th}}$ , and  $22^{\text{nd}}$  students.

### 3- Stratified random sampling:

Uses this method when the population is non-homogenous. The sample select by dividing the population into groups (strata) according to some characteristic and then taking samples from each group by using simple random sampling according to the weight of the each group.



The sample of each group  $n_i = \frac{N_i}{N} * n$

**Example:** In a company there are the following staffs:

Male (full time) = 90

Male (part time) = 18

Female (full time) = 9

Female (part time) = 63

**Select (40) of the staff by stratified Sampling:**

**Solution:**  $N_1 + N_2 + N_3 + N_4 = N \Rightarrow 90 + 18 + 9 + 63 = 180$



The sample of Male (full time)  $n_1 = \frac{N_1}{N} * n = \frac{90}{180} * 40 = 20$

The sample of Male (part time)  $n_2 = \frac{N_2}{N} * n = \frac{18}{180} * 40 = 4$

The sample of Female (full time)  $n_3 = \frac{N_3}{N} * n = \frac{9}{180} * 40 = 2$

The sample of Female (part time)  $n_4 = \frac{N_4}{N} * n = \frac{63}{180} * 40 = 14$

**Note:**  $n_1 + n_2 + n_3 + n_4 = n \Rightarrow 20 + 4 + 2 + 14 = 40$

## Chapter Two

### Presentation of the data:

After collecting data, the next step should be expressing it in some form such as:

- 1- Frequency distribution
- 2- Graphical presentation

### Frequency distribution:

Frequency distribution is the organizing of **raw data** in table form, used classes and frequencies.

**Example (1):** The following data (**raw data**) represent weights of (30) children:

20 , 10 , 8.6 , 7 , 5.9 , 6 , 12 , 8 , 14 , 5 , 9 , 16.5 , 25 , 16 , 19 ,  
20.9 , 7.8 , 15 , 18 , 24 , 9 , 7 , 15 , 16.9 , 11 , 8 , 26 , 17.6 , 21 , 10

**Example (2):** The following data (**frequency distribution**) represent weights of (30) children:

| Classes      | Frequency ( $f_i$ ) |
|--------------|---------------------|
| 5 – 9        | 9                   |
| 9 – 13       | 6                   |
| 13 – 17      | 6                   |
| 17 – 21      | 5                   |
| 21 – 25      | 2                   |
| 25 – 29      | 2                   |
| <b>Total</b> | <b>30</b>           |

Lower Limit – Upper Limit

There are (6) classes; each class has a lower class limit and upper class limit

**Frequency distribution for qualitative data:**

Frequency distribution for qualitative data lists all classes and the number of elements that belong to each of the classes

**Example (1):** The following data represent eye's color for 10 students:

|       |       |       |       |       |
|-------|-------|-------|-------|-------|
| Blue  | Brown | Black | Brown | Black |
| Black | Brown | Blue  | Black | Black |

**Construct the frequency distribution.**

**Solution:**

| Classes (Eyes color ) | Tally | Frequency ( $f_i$ ) |
|-----------------------|-------|---------------------|
| Blue                  |       | 2                   |
| Brown                 |       | 3                   |
| Black                 | /     | 5                   |
| <b>Total</b>          |       | <b>10</b>           |

**Example (2):** The following data represents blood group for 20 students:

|   |    |   |    |   |    |   |    |   |   |
|---|----|---|----|---|----|---|----|---|---|
| A | AB | O | AB | B | A  | O | AB | A | B |
| B | AB | O | A  | O | AB | O | O  | B | A |

**Construct the frequency distribution.**

**Solution:**

| Classes (Blood group) | Tally | $f_i$     |
|-----------------------|-------|-----------|
| A                     | /     | 5         |
| B                     |       | 4         |
| AB                    | /     | 5         |
| O                     | /     | 6         |
| <b>Total</b>          |       | <b>20</b> |

**Relative frequency** ( $f_i^*$ ): The ratio of frequency of class to the total of frequencies.

$$f_i^* = \frac{f_i}{\sum_{i=1}^m f_i}$$

**Example (3):** The following data represents the types of laptops used by (20) students:

|             |               |             |               |               |             |           |             |               |               |
|-------------|---------------|-------------|---------------|---------------|-------------|-----------|-------------|---------------|---------------|
| <i>hp</i>   | <i>lenovo</i> | <i>Dell</i> | <i>Hp</i>     | <i>Dell</i>   | <i>Dell</i> | <i>hp</i> | <i>hp</i>   | <i>lenovo</i> | <i>Sony</i>   |
| <i>Dell</i> | <i>Sony</i>   | <i>hp</i>   | <i>lenovo</i> | <i>lenovo</i> | <i>Sony</i> | <i>hp</i> | <i>Dell</i> | <i>Dell</i>   | <i>lenovo</i> |

- 1- Construct the frequency distribution.
- 2- Find Relative frequency ( $f_i^*$ ).

**Solution:**

| Classes<br>(types of laptops) | Tally | Frequency<br>( $f_i$ ) | Relative frequency<br>( $f_i^*$ ) |
|-------------------------------|-------|------------------------|-----------------------------------|
| <i>Hp</i>                     |       | 6                      | $\frac{6}{20} = 0.3$              |
| <i>Lenovo</i>                 |       | 5                      | $\frac{6}{20} = 0.25$             |
| <i>Dell</i>                   |       | 6                      | $\frac{6}{20} = 0.3$              |
| <i>Sony</i>                   |       | 3                      | $\frac{3}{20} = 0.15$             |
| <b>Total</b>                  |       | <b>20</b>              | <b>1</b>                          |

## Graphical presentation:

It is another method for presenting and summarizing grouped data.

### Graphical presentation for qualitative data:

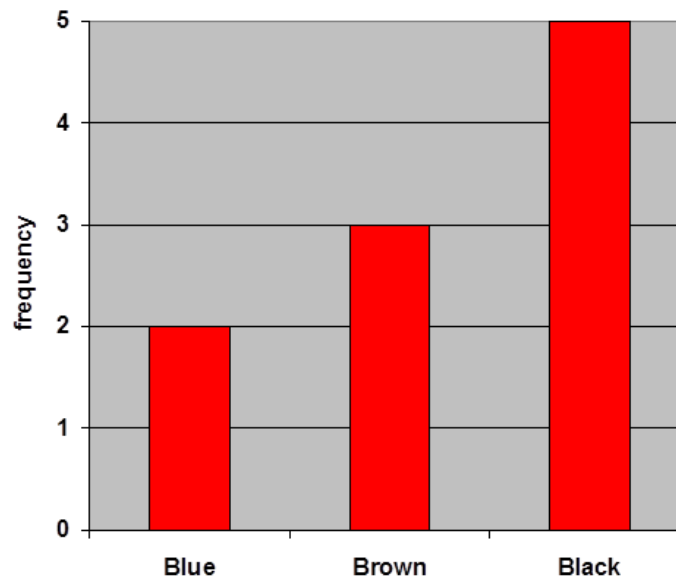
#### 1- Bar chart:

Bar chart is a graph that displays the classes (characteristics) on the horizontal axis and the frequency on the vertical axis.

**Example (1): Draw the bar chart from the following data:**

| Classes (Eyes color ) | Frequency ( $f_i$ ) |
|-----------------------|---------------------|
| Blue                  | 2                   |
| Brown                 | 3                   |
| Black                 | 5                   |
| <b>Total</b>          | <b>10</b>           |

**Solution:**



**2- Pie chart:** is a circle divided to number of sectors which are representing characteristics of variable.

The angle of each sector  $= \frac{f_i}{\sum_{i=1}^m f_i} * 360$

**Example (1): Draw the bar chart from the following data:**

| Classes (Eyes color ) | Frequency (f <sub>i</sub> ) |
|-----------------------|-----------------------------|
| Blue                  | 2                           |
| Brown                 | 3                           |
| Black                 | 5                           |
| <b>Total</b>          | <b>10</b>                   |

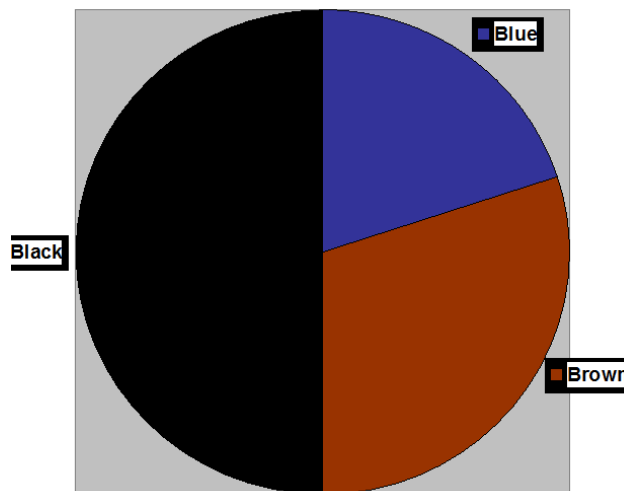
**Solution:**

The angle of each sector  $= \frac{f_i}{\sum_{i=1}^m f_i} * 360$

The angle of Blue  $= \frac{2}{10} * 360 = 72$

The angle of Brown  $= \frac{3}{10} * 360 = 108$

The angle of Black  $= \frac{5}{10} * 360 = 180$



## Frequency distribution for quantitative data :

General rules for constructing frequency table:

**1- Find the range (R):**

$$R = X_L - X_S + 1$$

**2- Find the number of classes (m):**

$$m = 2.5 * \sqrt[4]{n}$$

**3- find the length of classes (L):**

$$L = \frac{R}{m}$$

**4- Writing class limits**

**a- For continuous variable**

| Lower Limit    | Upper Limit |
|----------------|-------------|
| $X_S$          | $X_S + L$   |
| $X_S + L$      | $X_S + 2L$  |
| $X_S + 2L$     | $X_S + 3L$  |
| .              | .           |
| .              | .           |
| .              | .           |
| $X_S + (m-1)L$ | $X_S + mL$  |

**b- For discrete variable**

| Lower Limit    | Upper Limit    |
|----------------|----------------|
| $X_S$          | $X_S + L - 1$  |
| $X_S + L$      | $X_S + 2L - 1$ |
| $X_S + 2L$     | $X_S + 3L - 1$ |
| .              | .              |
| .              | .              |
| .              | .              |
| $X_S + (m-1)L$ | $X_S + mL - 1$ |

**5- Find frequencies of any classes**

**Examples for continuous variables:**

**Example (1):** The following data represents the times (minutes) for fifty (50) students, which are working on the Internet during one day.

|      |    |    |      |    |    |    |    |    |      |
|------|----|----|------|----|----|----|----|----|------|
| 50.5 | 40 | 41 | 17.2 | 11 | 7  | 22 | 44 | 28 | 21.1 |
| 19   | 23 | 37 | 51   | 54 | 42 | 20 | 41 | 78 | 56   |
| 72   | 56 | 17 | 7    | 69 | 30 | 80 | 56 | 29 | 33   |
| 46   | 31 | 39 | 88   | 18 | 29 | 34 | 59 | 73 | 77   |
| 36   | 39 | 30 | 62   | 54 | 67 | 39 | 31 | 53 | 44   |

**Construct the frequency distribution**

**Solution:**

**1- Find the range ( $R$ ):**

$$R = X_L - X_S + 1$$
$$R = 88 - 7 + 1 = 82$$

**2- Find the number of classes ( $m$ ):**

$$m = 2.5 * \sqrt[4]{n}$$
$$m = 2.5 * \sqrt[4]{50}$$
$$m = 2.5 * 2.66 = 6.65 \cong 7$$

**3- Find the length of classes ( $L$ ):**

$$L = \frac{R}{m} = \frac{82}{7} = 11.7 \cong 12$$



#### 4- Writing class limits

$$X_S = 7 \quad L = 12$$

| Classes         |                 |
|-----------------|-----------------|
| Lower Limits    | Upper Limits    |
| $X_S = 7$       | $X_S + L = 19$  |
| $X_S + L = 19$  | $X_S + 2L = 31$ |
| $X_S + 2L = 31$ | $X_S + 3L = 43$ |
| $X_S + 3L = 43$ | $X_S + 4L = 55$ |
| $X_S + 4L = 55$ | $X_S + 5L = 67$ |
| $X_S + 5L = 67$ | $X_S + 6L = 79$ |
| $X_S + 6L = 79$ | $X_S + 7L = 91$ |

#### 5- Find frequencies of any classes

| Classes      | Tally | Frequency ( $f_i$ ) |
|--------------|-------|---------------------|
| 7 – 19       |       | 6                   |
| 19 – 31      | /     | 10                  |
| 31 – 43      | /     | 13                  |
| 43 – 55      | /     | 8                   |
| 55 – 67      | /     | 5                   |
| 67 – 79      |       | 6                   |
| 79 – 91      |       | 2                   |
| <b>Total</b> |       | <b>50</b>           |

**Example (2):** The following data represent weights (30) children:

20 , 10 , 8.6 , 7 , 5.9 , 6 , 12 , 8 , 14 , **5** , 9 , 16.5 , 25 , 16 , 19 ,  
20.9 , 7.8 , 15 , 18 , 24 , 9 , 7 , 15 , 16.9 , 11 , 8 , **26** , 17.6 , 21 , 10

**Construct the frequency table.**

**Solution:**

**1- Find the range:**

$$R = X_L - X_S + 1 = 26 - 5 + 1 = 22$$

**2- Find the number of classes ( $m$ ):**

$$m = 2.5 * \sqrt[4]{n}$$

$$m = 2.5 * \sqrt[4]{30} = 5.6 \cong 6$$

**3- Find the length of classes:**

$$L = \frac{R}{m} = \frac{22}{6} = 3.66 \cong 4$$

**4- Writing class limits**

**5- Find frequencies of any classes**

| Classes      | Tally | Frequency ( $f_i$ ) |
|--------------|-------|---------------------|
| 5 – 9        |       | 9                   |
| 9 – 13       |       | 6                   |
| 13 – 17      |       | 6                   |
| 17 – 21      |       | 5                   |
| 21 – 25      |       | 2                   |
| 25 – 29      |       | 2                   |
| <b>Total</b> |       | <b>30</b>           |

**Examples for discrete variables:**

**Example (1):** The following data represent persons number of family in a small town:

3    5    8    10    12    7    5    11    2    9  
5    6    13    15    12    4    5    4    3    5

**Construct the frequency table.**

**Solution:**

**1- Find the range ( $R$ ):**

$$R = X_L - X_S + 1 = 15 - 2 + 1 = 14$$

**2- Find the number of classes ( $m$ ):**

$$m = 2.5 * \sqrt[4]{n}$$

$$m = 2.5 * \sqrt[4]{20} = 5.2 \cong 5$$

**3- Find the length of classes ( $L$ ):**

$$L = \frac{R}{m} = \frac{14}{5} = 2.8 \cong 3$$

**4- Writing class limits**

**5- Find frequencies of any classes**

| Classes      | Tally | $f_i$     |
|--------------|-------|-----------|
| 2 – 4        |       | 5         |
| 5 – 7        |       | 7         |
| 8 – 10       |       | 3         |
| 11 – 13      |       | 4         |
| 14 – 16      |       | 1         |
| <b>Total</b> |       | <b>20</b> |

**Example (2) : Construct a frequency distribution** from the following data which represent the number of apple trees of 20 farmers:

29    40    50    24    20    25    50    60    71    15  
 20    45    55    8    69    61    33    30    20    31

**Solution:**

**1- Find the range:**

$$R = X_L - X_S + 1 = 71 - 8 + 1 = 64$$

**2- Find the number of classes:**

$$m = 2.5 * \sqrt[4]{n}$$

$$m = 2.5 * \sqrt[4]{20} = 5.2 \cong 5$$

**3- Find the length of classes:**

$$L = \frac{R}{m} = \frac{64}{5} = 12.8 \cong 13$$

**4- Writing class limits**

**5- Find frequencies of any classes**

| Classes | Tally | Frequency |
|---------|-------|-----------|
| 8 - 20  |       | 5         |
| 21 - 33 |       | 6         |
| 34 - 46 |       | 2         |
| 47 - 59 |       | 3         |
| 60 - 62 |       | 4         |
| Total   |       | 20        |

**Relative frequency:** The ratio of frequency of class to the total of frequencies.

$$f_i^* = \frac{f_i}{\sum_{i=1}^m f_i}$$

**Example: Find relative frequency for this table:**

| Classes      | $f_i$     | $f_i^*$               |
|--------------|-----------|-----------------------|
| 2 – 4        | 5         | $\frac{5}{20} = 0.25$ |
| 5 – 7        | 7         | $\frac{7}{20} = 0.35$ |
| 8 – 10       | 3         | $\frac{3}{20} = 0.15$ |
| 11 – 13      | 4         | $\frac{4}{20} = 0.2$  |
| 14 – 16      | 1         | $\frac{1}{20} = 0.05$ |
| <b>Total</b> | <b>20</b> | <b>1</b>              |

**Center of class (Midpoint):**

Midpoint is the arithmetic mean for lower limit ( $L. L.$ ) and upper limit ( $U. L.$ ) of each class

$$X_i = \frac{L.L. + U.L.}{2}$$

**Example: Find the midpoint** from the following frequency table:

|         |       |       |        |         |         |
|---------|-------|-------|--------|---------|---------|
| Classes | 2 – 4 | 5 – 7 | 8 – 10 | 11 – 13 | 14 – 16 |
| $f_i$   | 5     | 7     | 3      | 4       | 1       |

**Solution:**

| Classes | $f_i$ | Midpoint ( $X_i$ )     |
|---------|-------|------------------------|
| 2 – 4   | 5     | $\frac{2+4}{2} = 3$    |
| 5 – 7   | 7     | $\frac{5+7}{2} = 6$    |
| 8 – 10  | 3     | $\frac{8+10}{2} = 9$   |
| 11 – 13 | 4     | $\frac{11+13}{2} = 12$ |
| 14 – 16 | 1     | $\frac{14+16}{2} = 14$ |

**Example:** From the following frequency table:

| Classes | 5 – 9 | 9 – 13 | 13 – 17 | 17 – 21 | 21 – 25 | 25 – 29 |
|---------|-------|--------|---------|---------|---------|---------|
| $f_i$   | 9     | 6      | 5       | 6       | 2       | 2       |

**Find:**

- 1- Relative frequency ( $f^*$ )
- 2- Percentage Relative frequency ( $f^*\%$ )
- 3- Mid points of the classes ( $X_i$ )

**Solution:**

| Classes      | $f_i$     | 1- $f_i^*$     | 2- $f_i^*\%$      | 3- $X_i$              |
|--------------|-----------|----------------|-------------------|-----------------------|
| 5 – 9        | 9         | <b>0.3</b>     | <b>0.3*100=30</b> | $\frac{5+9}{2} = 7$   |
| 9 – 13       | 6         | <b>0.2</b>     | <b>0.2*100=20</b> | $\frac{9+13}{2} = 11$ |
| 13 – 17      | 5         | <b>0.16666</b> | <b>16.666</b>     | <b>15</b>             |
| 17 – 21      | 6         | <b>0.2</b>     | <b>20</b>         | <b>19</b>             |
| 21 – 25      | 2         | <b>0.06666</b> | <b>6.666</b>      | <b>23</b>             |
| 25 – 29      | 2         | <b>0.06666</b> | <b>6.666</b>      | <b>27</b>             |
| <b>Total</b> | <b>30</b> |                |                   |                       |

### The Cumulative Frequency Distribution:

The cumulative frequency of a class is the sum of the frequency for that class and all previous classes.

### Examples for continuous frequency distribution

**Example (1): Find cumulative frequency distribution ( $F_i$ ) from the following frequency distribution:**

| Classes | $f_i$ | Upper limits | $F_i$          |
|---------|-------|--------------|----------------|
| 5 – 9   | 9     | Less than 9  | 9              |
| 9 – 13  | 6     | Less than 13 | 9+6=15         |
| 13 – 17 | 5     | Less than 17 | 9+6+5=20       |
| 17 – 21 | 6     | Less than 21 | 9+6+5+6=26     |
| 21 – 25 | 2     | Less than 25 | 9+6+5+6+2=28   |
| 25 – 29 | 2     | Less than 29 | 9+6+5+6+2+2=30 |
| Total   | 30    |              |                |

**Example (2): Find cumulative frequency distribution ( $F_i$ ) from the following frequency distribution:**

| Classes | $f_i$ | Upper limits | $F_i$ |
|---------|-------|--------------|-------|
| 1 – 3   | 4     | Less than 3  | 4     |
| 3 – 5   | 10    | Less than 5  | 14    |
| 5 – 7   | 16    | Less than 7  | 30    |
| 7 – 9   | 12    | Less than 9  | 42    |
| 9 – 11  | 6     | Less than 11 | 48    |
| 11 – 13 | 2     | Less than 13 | 50    |
| Total   | 50    |              |       |

**Examples for discrete frequency distribution**

**Example (1): Find cumulative frequency distribution ( $F_i$ ) from the following frequency distribution:**

| Classes | $f_i$ | Upper limits          | $F_i$        |
|---------|-------|-----------------------|--------------|
| 2 – 4   | 5     | Less than or equal 4  | 5            |
| 5 – 7   | 7     | Less than or equal 7  | 5+7=12       |
| 8 – 10  | 3     | Less than or equal 10 | 5+7+3=15     |
| 11 – 13 | 4     | Less than or equal 13 | 5+7+3+4=19   |
| 14 – 16 | 1     | Less than or equal 16 | 5+7+3+4+1=20 |
| Total   | 20    |                       |              |

**Example (2): Find cumulative frequency distribution ( $F_i$ ) from the following frequency distribution:**

| Classes | $f_i$ | Upper limits          | $F_i$ |
|---------|-------|-----------------------|-------|
| 50 – 54 | 1     | Less than or equal 54 | 1     |
| 55 – 59 | 2     | Less than or equal 59 | 3     |
| 60 – 64 | 11    | Less than or equal 64 | 14    |
| 65 – 69 | 10    | Less than or equal 69 | 24    |
| 70 – 74 | 12    | Less than or equal 74 | 36    |
| 75 – 79 | 21    | Less than or equal 79 | 57    |
| 80 – 84 | 8     | Less than or equal 84 | 65    |
| 85 – 89 | 9     | Less than or equal 89 | 74    |
| 90 – 94 | 4     | Less than or equal 94 | 78    |
| 95 – 99 | 4     | Less than or equal 99 | 82    |
| Total   | 82    |                       |       |



**Class boundaries:**

Class boundaries are the numbers used to separate classes without gaps. It's obtained as follows:

$$\text{Lower class boundary} = \text{Lower Limit} - 0.5$$

$$\text{Upper class boundary} = \text{Upper Limit} + 0.5$$

**Example: Find Class boundaries** from the following frequency table:

| Classes | $f_i$ |
|---------|-------|
| 2 – 4   | 5     |
| 5 – 7   | 2     |
| 8 – 10  | 3     |
| 11 – 13 | 9     |
| 14 – 16 | 2     |
| 17 – 19 | 3     |
| 20 – 22 | 1     |

**Solution:**

| Class boundaries | $f_i$ |
|------------------|-------|
| 1.5 – 4.5        | 5     |
| 4.5 – 7.5        | 2     |
| 7.5 – 10.5       | 3     |
| 10.5 – 13.5      | 9     |
| 13.5 – 16.5      | 2     |
| 16.5 – 19.5      | 3     |
| 19.5 – 22.5      | 1     |

## Graphical presentation for quantitative data:

### 1- Frequency Histogram:

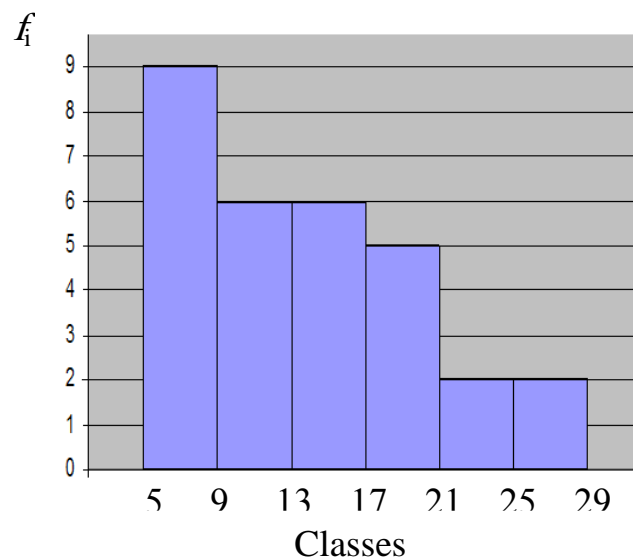
Frequency histogram is a graph that displays the classes on the horizontal axis and the frequencies ( $f_i$ ) on the vertical axis. The frequency of each class is represented by vertical bar.

### Examples for continuous frequency distribution

Draw the frequency histogram from the following frequency distribution:

| Classes | $f_i$ |
|---------|-------|
| 5 – 9   | 9     |
| 9 – 13  | 6     |
| 13 – 17 | 5     |
| 17 – 21 | 6     |
| 21 – 25 | 2     |
| 25 – 29 | 2     |

Solution:



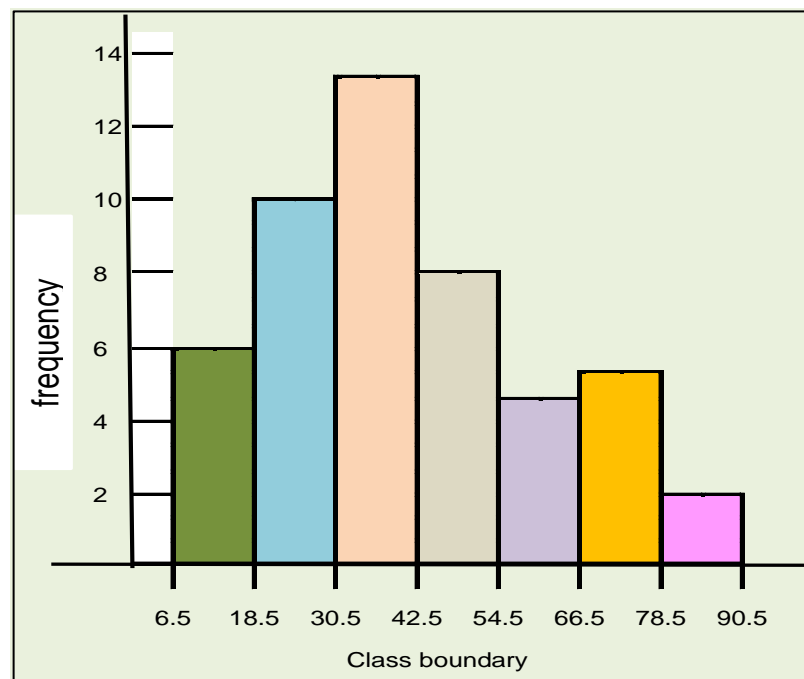
## Examples for discrete frequency distribution

**Example (1): Draw a histogram form the following frequency distribution:**

| Classes | 7 – 18 | 19 – 30 | 31 – 42 | 43 – 54 | 55 – 66 | 67 – 78 | 79 – 90 |
|---------|--------|---------|---------|---------|---------|---------|---------|
| $f_i$   | 6      | 10      | 13      | 8       | 5       | 6       | 2       |

**Solution:** In discrete frequency table must be using the class boundaries for each class to construct histogram.

| Class boundaries | $f_i$ |
|------------------|-------|
| 6.5 – 18.5       | 6     |
| 18.5 – 30.5      | 10    |
| 30.5 – 42.5      | 13    |
| 42.5 – 54.5      | 8     |
| 54.5 – 66.5      | 5     |
| 66.5 – 78.5      | 6     |
| 78.5 – 90.5      | 2     |

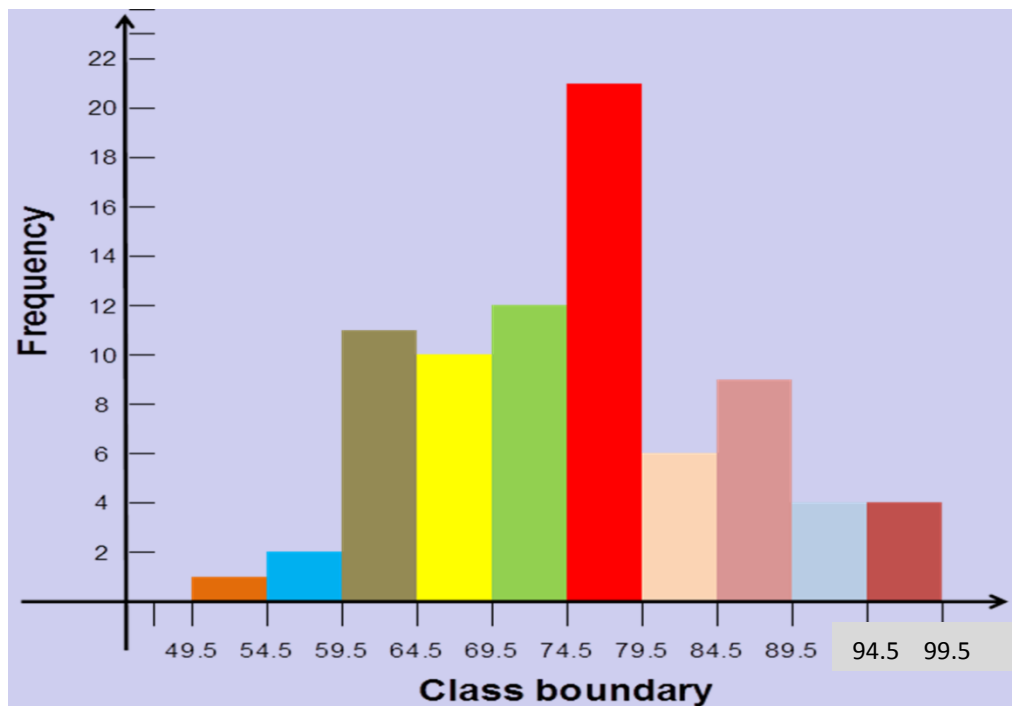


**Example (2): Draw the Histogram from the following frequency table:**

| Classes | 50 – 54 | 55 – 59 | 60 – 64 | 65 – 69 | 70 – 74 | 75 – 79 | 80 – 84 | 85 – 89 | 90 – 94 | 95 – 99 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| $f_i$   | 1       | 2       | 11      | 10      | 12      | 21      | 8       | 9       | 4       | 4       |

**Solution:**

| Class boundaries | $f_i$ |
|------------------|-------|
| 49.5 – 54.5      | 1     |
| 54.5 – 59.5      | 2     |
| 59.5 – 64.5      | 11    |
| 64.5 – 69.5      | 10    |
| 69.5 – 74.5      | 12    |
| 74.55 – 79.5     | 21    |
| 79.5 – 84.5      | 8     |
| 84.55 – 89.5     | 9     |
| 89.5 – 94.5      | 4     |
| 94.5 – 99.5      | 4     |



## 2- Frequency polygon:

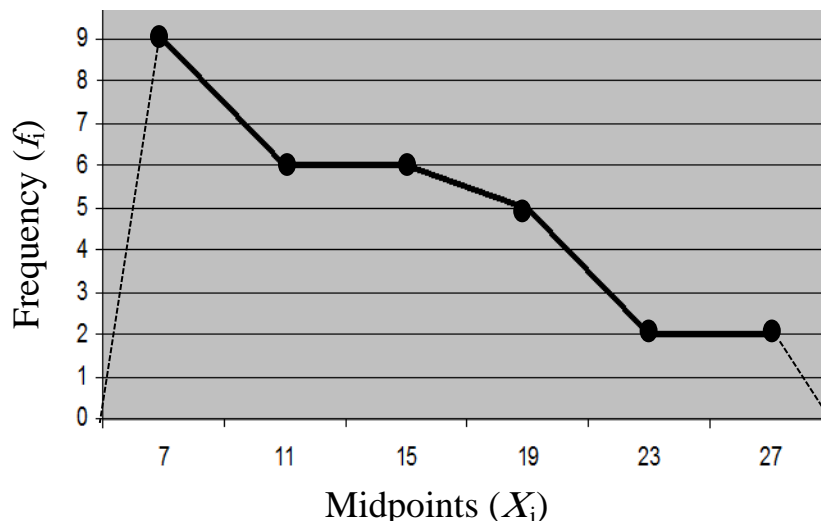
Frequency polygon is a graph that display the midpoints of classes ( $X_i$ ) on the horizontal axis and the frequencies ( $f_i$ ) on the vertical axis. The points are connected by lines.

**Example: Draw the Frequency polygon from the following frequency table:**

| Classes | 5 – 9 | 9 – 13 | 13 – 17 | 17 – 21 | 21 – 25 | 25 – 29 |
|---------|-------|--------|---------|---------|---------|---------|
| $f_i$   | 9     | 6      | 5       | 6       | 2       | 2       |

**Solution:**

| Classes | $f_i$ | $X_i$ |
|---------|-------|-------|
| 5 – 9   | 9     | 7     |
| 9 – 13  | 6     | 11    |
| 13 – 17 | 5     | 15    |
| 17 – 21 | 6     | 19    |
| 21 – 25 | 2     | 23    |
| 25 – 29 | 2     | 27    |



### 3- Frequency curve:

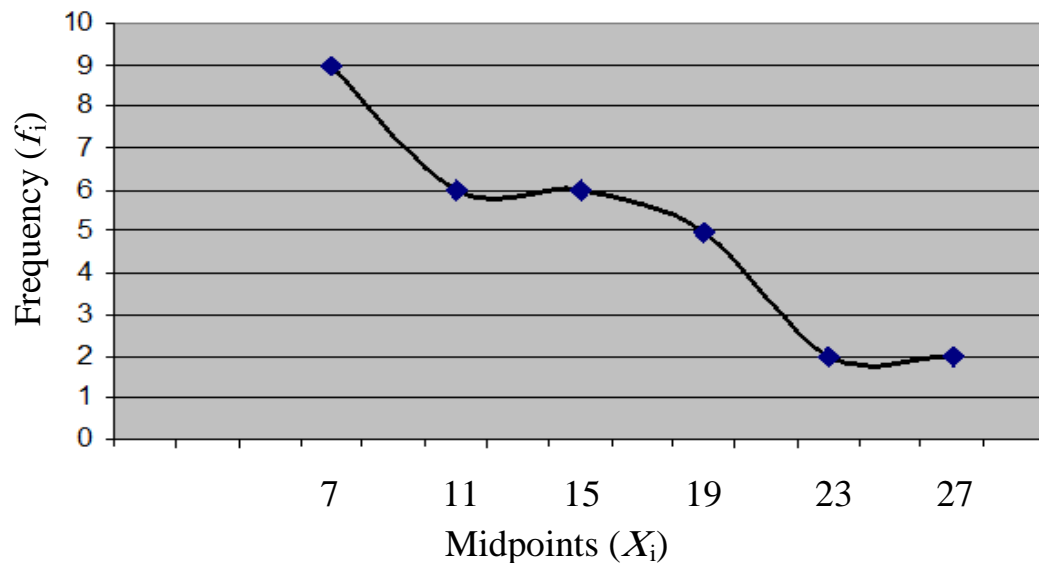
Frequency curve is a graph that displays the midpoints of classes ( $X_i$ ) on the horizontal axis and the frequencies ( $f_i$ ) on the vertical axis. The points are connected by curve.

**Example: Draw the Frequency curve from the following frequency table:**

| Classes | 5 – 9 | 9 – 13 | 13 – 17 | 17 – 21 | 21 – 25 | 25 – 29 |
|---------|-------|--------|---------|---------|---------|---------|
| $f_i$   | 9     | 6      | 5       | 6       | 2       | 2       |

**Solution:**

| Classes | $f_i$ | $X_i$ |
|---------|-------|-------|
| 5 – 9   | 9     | 7     |
| 9 – 13  | 6     | 11    |
| 13 – 17 | 5     | 15    |
| 17 – 21 | 6     | 19    |
| 21 – 25 | 2     | 23    |
| 25 – 29 | 2     | 27    |



# *Chapter Three*

## Measures of central tendency

A measure of central tendency is a value at the center or middle of a data set. This value represents all data of the group.

## Types of central tendency

There are some important types of measures of central tendency such as:

### 1- Arithmetic mean:

The arithmetic mean (generally called mean) is the sum of the values divided by the total number of values.

### Arithmetic mean for ungrouped data:

#### Population mean:

The mean of population for ( $N$ ) observations (values)  $X_1, X_2, \dots, X_N$  defined as:

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

#### Sample mean:

The mean of sample for ( $n$ ) observations (values)  $X_1, X_2, \dots, X_n$  defined as:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

**Example: Find the sample mean** from the following data:

4 , 5 , 8 , 10 , 12 , 6 , 5 , 14

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{4+5+8+10+12+6+5+14}{8} = \frac{64}{8} = 8$$



**Arithmetic mean for frequency distribution:**

If  $X_1, X_2, \dots, X_m$  are the mid points and  $f_1, f_2, \dots, f_m$  are frequencies, then:

$$\bar{X} = \frac{\sum_{i=1}^m f_i X_i}{\sum_{i=1}^m f_i}$$

**Example: Find the mean from the following frequency table:**

|         |       |       |        |         |         |
|---------|-------|-------|--------|---------|---------|
| Classes | 2 – 4 | 5 – 7 | 8 – 10 | 11 – 13 | 14 – 16 |
| $f_i$   | 5     | 7     | 3      | 4       | 1       |

**Solution:**

| Classes      | $f_i$     | $X_i$ | $f_i X_i$  |
|--------------|-----------|-------|------------|
| 2 – 4        | 5         | 3     | 15         |
| 5 – 7        | 7         | 6     | 42         |
| 8 – 10       | 3         | 9     | 27         |
| 11 – 13      | 4         | 12    | 48         |
| 14 – 16      | 1         | 15    | 15         |
| <b>Total</b> | <b>20</b> |       | <b>147</b> |

$$\bar{X} = \frac{\sum_{i=1}^m f_i X_i}{\sum_{i=1}^m f_i} = \frac{147}{20} = 6.35$$

## 2- Weighted Arithmetic Mean:

One of the limitations of the arithmetic mean that is gives equal importance to all the elements. But there are cases where the relative importance of the difference elements is not the same. In this case uses the weighted Arithmetic Mean:

$$\bar{X}_w = \frac{\sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i}$$

$W_i$ : the weights of the elements

**Example:** The following data represents the degrees and number of hours for a set of subjects for a student:

|                               |    |    |    |    |    |
|-------------------------------|----|----|----|----|----|
| Degrees of subjects ( $X_i$ ) | 75 | 87 | 63 | 94 | 61 |
| Number of hours ( $W_i$ )     | 3  | 2  | 3  | 3  | 4  |

### Solution:

| Degrees of subjects ( $X_i$ ) | Number of hours ( $W_i$ ) | $X_i W_i$   |
|-------------------------------|---------------------------|-------------|
| 75                            | 3                         | 225         |
| 87                            | 2                         | 174         |
| 63                            | 3                         | 189         |
| 94                            | 3                         | 282         |
| 61                            | 4                         | 244         |
| <b>Total</b>                  | <b>15</b>                 | <b>1114</b> |

$$\bar{X}_w = \frac{\sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i} = \frac{1114}{15} = 74.2667$$

### 3- The median:

The median of data set is the value that divides it into two equal parts.

#### The median for ungrouped data:

##### 1. If the number of the observations (n) is odd:

The median is the value which arranged in  $\left(\frac{n+1}{2}\right)$

Or  $Me = X_{\frac{n+1}{2}}$

**Example: Find the median** from the following data:

4 , 5 , 8 , 10 , 12 , 6 , 5 , 14 , 2

**Solution:** n=10 is odd

Arrange the data

2 , 4 , 5 , 5 , 6 , 8 , 10 , 12 , 14

└──────────────────┘ └──────────────────┘

$$Me = X_{\frac{n+1}{2}} = X_{\frac{9+1}{2}} = X_5 = 6$$

##### 2. If the number of the observations (n) is even:

The median is average of the two middle numbers.

Or  $Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$

**Example :** Find the median from the following data:

4 , 5 , 8 , 10 , 12 , 6 , 5 , 14 , 2 , 15

**Solution:** n=10 is even and Arrange the data

2 , 4 , 5 , 5 , 6 , 8 , 10 , 12 , 14 , 15

└──────────────────┘ ↑ └──────────────────┘

7

$$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} = \frac{X_5 + X_6}{2} = \frac{6+8}{2} = 7$$

## The median for frequency distribution:

### **A- The median in continuous frequency distribution:**

Steps to find the median value in the frequency distribution:

1- Calculate the cumulative frequency ( $F_i$ ).

2- Calculate  $\frac{\sum f_i}{2}$

3- Select the **Median class** which corresponds to the next cumulative frequency ( $F_{k+1}$ )

4- Calculate the median:

$$\text{Median} = L.L. + \frac{L_k}{f_k} \left( \frac{\sum f_i}{2} - F_{k-1} \right)$$

Where;

$L.L.$  = lower limit of the median class.

$\sum f_i$  = Total frequency.

$F_{k-1}$  = the previous cumulative frequency.

$f_k$  = frequency of the median class.

$L_k$  = Length (width) of median class.

**Examples: For continuous frequency distribution**

**Example (1): Find the median** from the following frequency table:

| Classes | 1 – 3 | 3 – 5 | 5 – 7 | 7 – 9 | 9 – 11 | 11 – 13 | 13 – 15 |
|---------|-------|-------|-------|-------|--------|---------|---------|
| $f_i$   | 4     | 10    | 16    | 12    | 6      | 1       | 1       |

**Solution:** 1- Calculate the cumulative frequency ( $F_i$ ) .

| Classes      | $f_i$                        | $F_i$                 |
|--------------|------------------------------|-----------------------|
| 1 – 3        | 4                            | <b>4</b>              |
| 3 – 5        | 10                           | <b>14</b> ← $F_{k-1}$ |
| <b>5 – 7</b> | <b><math>f_k = 16</math></b> | <b>30</b> ← $F_{k+1}$ |
| 7 – 9        | 12                           | <b>42</b>             |
| 9 – 11       | 6                            | <b>28</b>             |
| 11 – 13      | 1                            | <b>49</b>             |
| 13 – 15      | 1                            | <b>50</b>             |
| Total        | $\sum f_i = 50$              |                       |

2- Calculate  $\frac{\sum f_i}{2} = \frac{50}{2} = 25$

3- Select the **median class** which is corresponds to  $F_{k+1}$  (5 – 7)

4–  $L.L. = 5$        $L_k = 7 - 5 = 2$        $f_k = 16$        $F_{k-1} = 14$

$$Median = L.L. + \frac{L_k}{f_k} \left( \frac{\sum f_i}{2} - F_{k-1} \right)$$

$$Median = 5 + \frac{2}{16} (25 - 14)$$

$$Median = 5 + \frac{2}{16} (11)$$

$$Median = 5 + \frac{22}{16}$$

$$Median = 5 + 1.375 = 6.375$$

**Example (2): Find the median** from the following frequency table:

| Classes | Less than 300 | 300–400 | 400–500 | 500–600 | 600–700 | greater than 700 |
|---------|---------------|---------|---------|---------|---------|------------------|
| $f_i$   | 3             | 5       | 9       | 8       | 6       | 2                |

**Solution:** 1- Calculate the cumulative frequency ( $F_i$ ).

| Classes          | $f_i$           | $F_i$                 |
|------------------|-----------------|-----------------------|
| Less than 300    | 3               | <b>3</b>              |
| 300–400          | 5               | <b>8</b> ← $F_{k-1}$  |
| <b>400–500</b>   | <b>9</b>        | <b>17</b> ← $F_{k+1}$ |
| 500–600          | 8               | <b>25</b>             |
| 600–700          | 6               | <b>31</b>             |
| greater than 700 | 2               | <b>33</b>             |
| Total            | $\sum f_i = 33$ |                       |

2- Calculate  $\frac{\sum f_i}{2} = \frac{33}{2} = 16.5$

3- Select the median class which is corresponds to  $F_{k+1}$  (400 – 500)

4-  $L.L. = 400$        $L_k = 500 - 400 = 100$        $f_k = 9$        $F_{k-1} = 8$

$$\text{Median} = L.L. + \frac{L_k}{f_k} \left( \frac{\sum f_i}{2} - F_{k-1} \right)$$

$$\text{Median} = 400 + \frac{100}{9} (16.5 - 8)$$

$$\text{Median} = 400 + \frac{100}{9} (8.5) = 400 + \frac{850}{9} = 400 + 94.44 = 494.44$$

**B- The median in discrete frequency distribution:**

1- Calculate the cumulative frequency ( $F_i$ ) .

2- Calculate  $\frac{\sum f_i}{2}$

3- Select the median class which is corresponds to  $F_{k+1}$

4- Median= midpoint of median class =  $\frac{L.L. + U.L.}{2}$

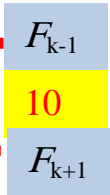
**Examples: For discrete frequency distribution:**

**Example (1): Find the median** from the following frequency table:

|         |         |         |         |         |         |
|---------|---------|---------|---------|---------|---------|
| Classes | 12 –1 4 | 15 –1 7 | 18 – 20 | 21 – 23 | 24 – 26 |
| $f_i$   | 5       | 7       | 3       | 4       | 1       |

**Solution:** 1- Calculate the cumulative frequency ( $F_i$ ) .

| Classes | $f_i$           | $F_i$ |
|---------|-----------------|-------|
| 12 –1 4 | 5               | 5     |
| 15 –1 7 | 7               | 12    |
| 18 – 20 | 3               | 15    |
| 21 – 23 | 4               | 19    |
| 24 – 26 | 1               | 20    |
| Total   | $\sum f_i = 20$ |       |



2- Calculate  $\frac{\sum f_i}{2} = \frac{20}{2} = 10$

3- Select the median class which is corresponds to  $F_{k+1}$  (15 – 17)

4- Median =  $\frac{15 + 17}{2} = 16$

**Example (2): Find the median** from the following frequency table:

| Classes | 100 – 104 | 105 – 109 | 110 – 114 | 115 – 119 | 120 – 124 | 125 – 129 |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|
| $f_i$   | 2         | 8         | 18        | 13        | 7         | 2         |

**Solution:**

1- Calculate the cumulative frequency ( $F_i$ ) .

| Classes          | $f_i$           | $F_i$     |
|------------------|-----------------|-----------|
| 100 – 104        | 2               | 2         |
| 105 – 109        | 8               | 10        |
| <b>110 – 114</b> | <b>18</b>       | <b>28</b> |
| 115 – 119        | 7               | 35        |
| 120 – 124        | 3               | 38        |
| 125 – 129        | 2               | 40        |
| Total            | $\sum f_i = 40$ |           |

2- Calculate  $\frac{\sum f_i}{2} = \frac{40}{2} = 20$

3- Find Median class which is corresponds to  $F_{k+1}$  (110 – 114)

4-  $Median = \frac{110 + 114}{2} = 112$



#### 4- The Mode:

It is the value that occurs more frequently in a set of data.

#### The Mode for ungrouped data:

**Example (1): Find the mode** from the following data:

4 , 5 , 8 , 10 , 12.5 , 6 , 5 , 4 , 2 , 5

**Solution:** mode = 5

**Example (2):**

4 , 5 , 8 , 5 , 4 , 6 , 5 , 4 , 2 , 7

**Solution:** mode = 4 , 5

**Example (3):**

3 , 5.6 , 8 , 10 , 12 , 6 , 7 , 4 , 2 , 1

**Solution:** there no mode

#### The Mode for frequency distribution:

**A- The mode in discrete frequency distribution:**

1- Select the largest frequency

2- Select the modal class that corresponds to the largest frequency

3- Mode= midpoint of modal class =  $\frac{L.L. + U.L.}{2}$

**Example: For discrete frequency distribution**

**Example (1): Find the mode** from the following frequency table:

| Classes | 12 – 14 | 15 – 17 | 18 – 20 | 21 – 23 | 24 – 26 |
|---------|---------|---------|---------|---------|---------|
| $f_i$   | 5       | 7       | 3       | 4       | 1       |

**Solution:**

1- Select the largest frequency

| Classes | $f_i$ |
|---------|-------|
| 12 – 14 | 5     |
| 15 – 17 | 7     |
| 18 – 20 | 3     |
| 21 – 23 | 4     |
| 24 – 26 | 1     |

**largest frequency ( $f_k$ )**

2- Select the modal class that corresponds to the largest frequency (15 – 17)

$$3- \text{Mode} = \frac{15+17}{2} = 16$$

**Example (2): Find the mode** from the following frequency table:

| Classes   | $f_i$ |
|-----------|-------|
| 100 – 104 | 2     |
| 105 – 109 | 6     |
| 110 – 114 | 7     |
| 115 – 119 | 10    |
| 120 – 124 | 8     |
| 125 – 129 | 3     |

**Largest frequency ( $f_k$ )**

**Solution:**

1- Select the largest frequency

2- Select the modal class that corresponds to the largest frequency (115 – 119)

$$3- \text{Mode} = \frac{115+119}{2} = 117$$

**B- The Mode in continuous frequency distribution:**

- 1- Select the largest frequency.
- 2- Select the modal class that corresponds to the largest frequency.
- 3- Calculate the mode:

$$Mode = L.L. + \frac{(f_k - f_{k-1})L_k}{(f_k - f_{k-1}) + (f_k - f_{k+1})}$$

Where;

$L.L.$  = lower limit of the median class.

$f_k$  = largest frequency.

$f_{k-1}$  = Previous frequency.

$f_{k+1}$  = next frequency.

$L_k$  = Length (width) of median class.

**Examples: For continuous frequency distribution**

**Example (1): Find the mode** from the following frequency table:

| Classes | 1 – 3 | 3 – 5 | 5 – 7 | 7 – 9 | 9 – 11 | 11 – 13 | 13 – 15 |
|---------|-------|-------|-------|-------|--------|---------|---------|
| $f_i$   | 4     | 10    | 16    | 12    | 6      | 1       | 1       |

**Solution:** 1- Select the largest frequency

| Classes        | $f_i$     |
|----------------|-----------|
| 11 – 13        | 4         |
| 13 – 15        | 10        |
| <b>15 – 17</b> | <b>18</b> |
| 17 – 19        | 12        |
| 19 – 21        | 6         |
| 21 – 23        | 1         |
| 23 – 25        | 1         |

$f_{k-1}$ 
 $f_{k-1}$

largest frequency ( $f_k$ )

2- Select the modal class that corresponds to the largest frequency (**15 – 17**)

$$3- \quad L.L. = 15 \quad f_k = 18 \quad f_{k-1} = 10 \quad f_{k+1} = 12 \quad L_k = 17 - 15 = 2$$

$$\begin{aligned} \text{Mode} &= L.L. + \frac{(f_k - f_{k-1})L_k}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \\ &= 15 + \frac{(18 - 10) 2}{(18 - 10) + (18 - 12)} \\ &= 15 + \frac{(8) 2}{(8) + (6)} \\ &= 15 + \frac{16}{14} = 15 + 1.43 = 16.43 \end{aligned}$$

**Example (2): Find the mode** from the following frequency table:

| Classes | Less than 30 | 30–40 | 40–50 | 50–60 | 60–70 | 70–80 |
|---------|--------------|-------|-------|-------|-------|-------|
| $f_i$   | 3            | 5     | 9     | 8     | 6     | 2     |

**Solution:** 1- Select the largest frequency

| Classes      | $f_i$    |
|--------------|----------|
| Less than 30 | 3        |
| 30–40        | 5        |
| <b>40–50</b> | <b>9</b> |
| 50–60        | 8        |
| 60–70        | 6        |
| 70–80        | 2        |

$f_{k-1}$  ←  
**largest frequency ( $f_k$ )** ←  
 $f_{k-1}$  ←

2- Select the modal class that corresponds to the largest frequency (**40 –50**)

$$3- \quad L.L. = 40 \quad f_k = 9 \quad f_{k-1} = 5 \quad f_{k+1} = 8 \quad L_k = 50 - 40 = 10$$

$$\begin{aligned} \text{Mode} &= L.L. + \frac{(f_k - f_{k-1})L_k}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \\ &= 40 + \frac{(9 - 5) 10}{(9 - 5) + (9 - 8)} \\ &= 40 + \frac{(4) 10}{(4) + (1)} = 40 + \frac{40}{5} \\ &= 40 + 8 = 48 \end{aligned}$$

### **Measures of Dispersion:**

Measures of dispersion give us information about how the data are spread about the mean, or describes how the data are spread out.

### **Objectives of Dispersion:**

The main objective of dispersion is to know the homogeneity of the values for a data set, or to compare between the values for two or more than data set.

### **Note:**

1. If all the values are the same  
→ There is no dispersion.
2. If all the values are different  
→ There is a dispersion:
3. If the values close to each other  
→ The amount of dispersion small.
4. If the values are widely scattered  
→ The dispersion is large.

### **Types of Measures of Dispersion:**

1. Range (R).
2. Mean Deviation (*M.D.*)
3. Variance. ( $S^2$ )
4. Standard deviation. (*S*)
5. Coefficient of variation (C.V).

## 1. Range (R):

Range is the difference between the largest value and smallest value.

$$R = X_L - X_S$$

**Example:** Find the range from the following data:

43 , 66 , 61 , 64 , 38, 55 , 50 , 20 , 30

$$R = X_L - X_S = 66 - 20 = 44$$

## 2. Mean Deviation (M.D.)

Mean Deviation is average of the absolute deviations from the mean.

$$M.D = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n-1}$$

**Example:** Find the mean deviation from the following data:

1 , 4 , 7 , 5 , 3

| $X_i$                   | $X_i - \bar{X}$ | $ X_i - \bar{X} $                  |
|-------------------------|-----------------|------------------------------------|
| 1                       | 1 - 4 = -3      | 3                                  |
| 4                       | 4 - 4 = 0       | 0                                  |
| 7                       | 7 - 4 = 3       | 3                                  |
| 5                       | 5 - 4 = 1       | 1                                  |
| 3                       | 3 - 4 = -1      | 1                                  |
| $\sum_{i=1}^n X_i = 20$ |                 | $\sum_{i=1}^n  X_i - \bar{X}  = 8$ |

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{20}{5} = 4$$

$$M.D = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n-1} = \frac{8}{5-1} = 2$$

### 3.The Variance ( $S^2$ ):

The variance is the average of the square deviations of values ( $X_i$ ) from their arithmetic mean ( $\bar{X}$ ).

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

**Example:** Find the variance from the following data:

1 , 4 , 7 , 5 , 3

| $X_i$                   | $X_i - \bar{X}$ | $(X_i - \bar{X})^2$                   |
|-------------------------|-----------------|---------------------------------------|
| 1                       | 1 - 4 = -3      | 9                                     |
| 4                       | 4 - 4 = 0       | 0                                     |
| 7                       | 7 - 4 = 3       | 9                                     |
| 5                       | 5 - 4 = 1       | 1                                     |
| 3                       | 3 - 4 = -1      | 1                                     |
| $\sum_{i=1}^n X_i = 20$ |                 | $\sum_{i=1}^n (X_i - \bar{X})^2 = 20$ |

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{20}{5} = 4$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{20}{5-1} = 5$$

### 4.The Standard Deviation ( $S$ ):

The Standard Deviation is the square root of the variance

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

**Example:** Find the Standard Deviation ( $S$ ) from the following data:

1 , 4 , 7 , 5 , 3

| $X_i$                   | $X_i - \bar{X}$ | $(X_i - \bar{X})^2$                   |
|-------------------------|-----------------|---------------------------------------|
| 1                       | 1 - 4 = -3      | 9                                     |
| 4                       | 4 - 4 = 0       | 0                                     |
| 7                       | 7 - 4 = 3       | 9                                     |
| 5                       | 5 - 4 = 1       | 1                                     |
| 3                       | 3 - 4 = -1      | 1                                     |
| $\sum_{i=1}^n X_i = 20$ |                 | $\sum_{i=1}^n (X_i - \bar{X})^2 = 20$ |

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{20}{5} = 4$$

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{20}{5-1}} = \sqrt{5} = 2.236$$