8k

Chapter Four

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:(Avarge)**

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) $X_1, X_2, \ldots, X_N$ defined as:

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

Sample mean:

The mean of sample for (n) observations (values) $X_1, X_2, \ldots, X_n$ 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 $X_1, X_2, \ldots, X_m$ are the mid points and $f_1, f_2, \ldots, f_m$ are frequencies, then:

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

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
Total	20		147

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

2- Weighted Arithmetic Mean (Weighted Average)

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	**61*4=244**
Total	**15**	**1114**

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

3- - The 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

Example: For discrete frequency distribution

Example (1): Find the mode 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- Select the largest frequency

Classes	f_i
12 –1 4	5
15 –1 7	7
18 – 20	3
21 – 23	4

largest frequency (f_k)

24 – 26	1

2- Select the modal class that corresponds to the largest frequency (15 –1 7)

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

Classes	f_i
100 – 104	22
105 – 109	6
110 – 114	**7**
115 – 119	**10**
120 – 124	22
125 – 129	3

Largest frequency (f_k)

<u>Solution:</u>

1- Select the largest frequency

2- Select the modal class that corresponds to the largest frequency (115 –1 19)

B- The Mode in <u>continuous</u> 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: <u>For continuous frequency distribution</u>

bExample (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	← f_{k-1}
15 –1 7	18	← **largest frequency (f_k)**
17 – 19	12	← f_{k+1}
19 – 21	6	
21– 23	1	
23– 25	1	

2- Select the modal class that corresponds to the largest frequency **(15 –1 7)**

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

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

$$=15+\frac{(18-10)\ 2}{(18-10)+(18-12)}$$

$$=15+\frac{(8)\ 2}{(8)+(6)}$$

$$=15+\frac{16}{14}=15+1.43=16.43$$

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	← f_{k-1}
40–50	**9**	← **largest frequency (f_k)**

50–60	8 ← f_{k-1}
60–70	6
70–80	2

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

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

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

$$= 40 + \frac{(9-5)\;10}{(9-5)+(9-8)}$$

$$= 40 + \frac{(4)\;10}{(4)+(1)} == 40 + \frac{40}{5}$$

$$= 40 + 8 = 48$$

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=9 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 is corresponds to the next cumulative frequency(F_k)

4- Calculate the median:

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

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) .

25

F_{k+1}

Classes	f_i	F_i
1 – 3	4	**4**
3 – 5	10	**14** ← F_{k-1}
5 – 7	f_k **=16**	**30**
7 – 9	12	**42**
9 – 11	6	**28**
11 – 13	1	**49**
13 – 15	1	**50**
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 (5 – 7)

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

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

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

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

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

$$Median = 5 + 1.375 = 6.375$$

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	**3**

300–400	5	**8**	F_{k-1}
400–500	**9**	17	
500–600	8	**25**	16.5
600–700	6	**31**	
greater than 700	2	**33**	
Total	$\sum f_i = 33$		

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

3- Select the median class which is correspond to F_k (400 – 500)

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

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

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

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

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

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 – 14	5	**5**
15 – 17	**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 (15 – 17)

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**
110 – 114	**18**	**28**
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 (110 – 114)

5-The Lower Quartile

$$Q_1 = Lb_{Q_1} + \frac{\frac{1}{4}n - f_c}{f_{Q_1}}.C$$

Q_1 = the lower quartile

Lb_{Q_1} = the lower limit of the lower quartile

n = the sum of data

f_c = the cumulative frequency before the lower quartile class

f_{Q_1} = the frequency of the lower quartile class

C = the width of interval class

- **The Upper Quartile**

$$Q_3 = Lb_{Q_3} + \frac{\ldots n - f_c}{f_{Q_3}}.C$$

Q_3 = the upper quartile

Lb_{Q_3} = the lower boundary of the upper quartile

n = the sum of data

f_c = the cumulative frequency before the upper quartile class

f_{Q_1} = the frequency of the upper quartile class

C = the width of interval class

- The **Percentile**

$$P_x = Lb_{P_x} + \frac{\frac{x}{100}n - f_c}{f_{P_x}}.C$$

P_x = the x^{th} percentile

Lb_{P_x} = the lower boundary of the x^{th} percentile

n = the sum of data

f_c = the cumulative frequency before the x^{th} percentile class

f_{P_x} = the frequency of the x^{th} percentile class

C = the width of interval class

Example 37

The length of 40 insects of a certain species were measured correct to the nearest millimeter.

Lengths (mm)	Frequency (f_i)
25 – 29	2
30 – 34	4
35 – 39	7
40 – 44	10
45 – 49	8
50 – 54	6
55 – 59	3

Use the cumulative frequency curve (ogive) to estimate:

a. the median length
b. the upper quartile
c. the lower quartile

Solution

The cumulative frequency table is constructed below. The table shows the cumulative frequency distribution of the length of 40 insects.

a. the median length, 50% of the total frequency = $\frac{50}{100} \times 40 = \ldots\ldots$

From the curve, the median length = ……

b. the upper quartile, 75% of the total frequency = $\frac{75}{100} \times 40 = \ldots\ldots$

From the curve, the upper quartile = ……

c. the lower quartile, 25% of the total frequency = $\frac{25}{100} \times 40 = \ldots\ldots$

From the curve, the lower quartile = ……

By formula:

Lengths (mm)	Frequency (f_i)	The cumulative frequency
25 – 29	2	2
30 – 34	4	
35 – 39	7	
40 – 44	10	
45 – 49	8	
50 – 54	6	
55 – 59	3	

½ n = ½ x 40 = 20, 20 in the class 40 – 44 .

$$Q_2 = Lb_{Q_2} + \frac{\frac{1}{2}n - f_c}{f_{Q_2}}.C$$

$$Q_2 = 39.5 + \frac{\frac{1}{2}.40 - 13}{10}.5$$

$$Q_2 = 39.5 + \frac{20 - 13}{10}.5$$

$$Q_2 = 39.5 + \frac{7}{2} = 39.5 + 3.5 = 43$$

a. ¾ n = ¾ x …… = ……, …… in the class …… – ……

$$Q_3 = Lb_{Q_3} + \frac{\frac{3}{4}n - f_c}{f_{Q_3}}.C$$

$$Q_3 = \ldots\ldots + \frac{\frac{3}{4} \times \ldots\ldots - \ldots\ldots}{\ldots\ldots}.\ldots\ldots$$

$$Q_3 = \ldots\ldots + \frac{\ldots\ldots}{\ldots\ldots} = \ldots\ldots + \ldots\ldots = \ldots\ldots$$

b. ¼ n = ¼ x …… = ……, …… in the class …… – ……

$$Q_1 = Lb_{Q_1} + \frac{\frac{1}{4}n - f_c}{f_{Q_1}}.C$$

$$Q_1 = \ldots\ldots + \frac{\ldots\ldots}{\ldots\ldots} = \ldots\ldots + \ldots\ldots = \ldots\ldots$$

Example:

find the the quartiles Q1, Q2, and Q3 of the following data .

Class Interval	Frequency (fi)
50 - 69	3
70 - 89	7
90 - 109	4
110 – 129	4
130 – 149	9

Solution: 1) find the cumulative frequency and the summation of frequencies and real interval limit.

Class Interval	Frequency (fi)	Cumulative frequency	Real interval	
50 - 69	3	3	49.5 – 69.5	
70 - 89	7	10	69.5 – 89.5	Q1
90 - 109	4	14	89.5 – 109.5	
110 – 129	4	18	109.5 – 129.5	
130 – 149	9	27	129.5 – 149.5	

2) find the arrangement number of quartiles to find quartile interval 1.

q1 = (1/4) x n = (1/4) x 27 = 6.75

The interval of quartile number 1 is have the cumulative frequency = 10

$$Q_1 = L_1 + \frac{\frac{n}{4} - cf}{f} x\, i$$

L_1 = Lower limit of quartiles interval = 69.5

cf = the cumulative frequency of the previous interval of the quartiles interval = 3

f = the frequency of quartiles interval = 7

i = the length of quartiles interval = 20

* $Q_1 = L_1 + \frac{\frac{n}{4} - cf}{f} x\ i$ $Q_1 = 69.5 + \frac{6.75-3}{7} x\ 20 = 80.2$

3. find the arrangement number of quartiles to find quartile interval 2.

$Q_2 = (2/4) \times n = (2/4) \times 27 = 13.5$

The interval of quartile number 2 is have the cumulative frequency = 14

* $Q_2 = L_1 + \frac{2(\frac{n}{4} - cf)}{f} x\ i$ $Q_2 = 89.5 + \frac{13.5-10}{4} x\ 20 = 107$

4. find the arrangement number of quartiles to find quartile interval 3.

$Q_3 = (3/4) \times n = (3/4) \times 27 = 20.25$

The interval of quartile number 3 is have the cumulative frequency = 27

$Q_3 = L_1 + \frac{3(\frac{n}{4} - cf)}{f} x\ i$ $Q_3 = 129.5 + \frac{20.25-18}{9} x\ 20 = 134.5$

Class Interval	Frequency (fi)	Cumulative frequency	Real interval	
50 - 69	3	3	49.5 – 69.5	
70 - 89	7	10	69.5 – 89.5	Q1
90 - 109	4	14	89.5 – 109.5	Q2
110 – 129	4	18	109.5 – 129.5	
130 – 149	9	27	129.5 – 149.5	Q2

Relation between Mean, Median and Mode