# Chapter 18: Measures of Central Tendency for Grouped Data

OBJECTIVES
At the end of the chapter, students should be able to determine the mean, median and mode of grouped data.

## I. Determination of the Mean of a Grouped Data

### (i) Formulae for calculating the mean of a grouped data

The mean of a grouped data can be determined by applying any of the following formulae.
(i)

$x = \frac{\sum fx}{\sum f}$ where x is the mean, f is the frequency and Î£ is the sum.

(ii) $x = A + \frac{\sum fd}{\sum f}$ where d = $d = \bar{x} - A$ and A is the assumed mean.

(iii) $\bar{x} = A + c\frac{\sum fu}{\sum f}$

### (ii) Application of the formulae

**Worked Example 1**

1. The following data shows the lengths of 30 pieces of iron rods in centimetres:

45 55 65 60 61 68 59 54 64 76
50 68 72 68 80 67 70 62 79 67
64 63 71 59 64 53 57 74 55 57

Using class intervals of 45–49, 50–54, 55–59, etc., construct a frequency table of the data. Calculate the mean of the distribution. (WAEC June 2002:12)

**SOLUTION**
Using the first formula

| Class interval | Tally | $f$ | $x$ | $fx$ |
|---|---|---|---|---|
| 45–49 | / | 1 | 47 | 47 |
| 50–54 | /// | 3 | 52 | 156 |
| 55–59 | ~~////~~ / | 6 | 57 | 342 |
| 60–64 | ~~////~~ // | 7 | 62 | 434 |
| 65–69 | ~~////~~ / | 6 | 67 | 402 |
| 70–74 | //// | 4 | 72 | 288 |
| 75–79 | // | 2 | 77 | 154 |
| 80–84 | / | 1 | 82 | 82 |
| | | 30 | | 1 905 |

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{1\,905}{30} = 63.5$$

Using the second formula (assumed mean)

| Class interval | $f$ | $x$ | $d = x - A$ | $fd$ |
|---|---|---|---|---|
| 45–49 | 1 | 47 | –15 | –15 |
| 50–54 | 3 | 52 | –10 | –30 |
| 55–59 | 6 | 57 | –5 | –30 |
| 60–64 | 7 | 62 | 0 | 0 |
| 65–69 | 6 | 67 | 5 | 30 |
| 70–74 | 4 | 72 | 10 | 40 |
| 75–79 | 2 | 77 | 15 | 30 |
| 80–84 | 1 | 82 | 20 | 20 |
| | 30 | | | 45 |

$$\bar{x} = A + \frac{\sum fd}{\sum f} = 62 + \frac{45}{30}$$

$$= 62 + 1.5$$

$$= 63.5$$

Using the third formula (coding method)

| Class interval | $f$ | $x$ | $d = x - A;$ $A = 62$ | $u = \frac{d}{c};$ $c = 5$ | $fu$ |
|---|---|---|---|---|---|
| 45–49 | 1 | 47 | –15 | –3 | –3 |
| 50–54 | 3 | 52 | –10 | –2 | –6 |

| | | | | | |
|---|---|---|---|---|---|
| 55–59 | 6 | 57 | −5 | −1 | −6 |
| 60–64 | 7 | 62 | 0 | 0 | 0 |
| 65–69 | 6 | 67 | 5 | 1 | 6 |
| 70–74 | 4 | 72 | 10 | 2 | 8 |
| 75–79 | 2 | 77 | 15 | 3 | 6 |
| 80–84 | 1 | 82 | 20 | 4 | 4 |
| | 30 | | | | 9 |

$$\bar{x} = A + c\frac{\sum fu}{\sum f}$$

$$= 62 + 5\left(\frac{9}{30}\right)$$

$$= 62 + 5 \times \frac{9}{30}$$

$$= 62 + 1.5$$

$$= 63.5$$

**Exercise**

1. The table below shows the distribution of distances (in kilometres) of 60 villagers from niger state of Nigeria:

| Distances | Number of villagers |
|---|---|
| 20–29 | 10 |
| 30–39 | 9 |
| 40–49 | 12 |
| 50–59 | 8 |
| 60–69 | 6 |
| 70–79 | 7 |
| 80–89 | 5 |

Calculate the mean of the distribution using an assumed mean.

2. The distribution below shows the reading scores of 20 JSS3 students:

| Score | Number of students |
|---|---|
| 10–14 | 1 |
| 15–19 | 3 |
| 20–24 | 1 |
| 25–29 | 2 |
| 30–34 | 1 |
| 35–39 | 5 |
| 40–44 | 3 |
| 45–49 | 4 |
| 50–54 | 2 |

Find the mean of the distribution above using the formulae below:

$$\text{Mean} = \frac{\sum fx}{\sum f} \text{ and}$$

$$\text{Mean} = A + c\frac{\sum fu}{\sum f}$$

3. The following marks were scored by 25 students in an examination:

| | | | | |
|---|---|---|---|---|
| 50 | 20 | 10 | 40 | 50 |
| 40 | 10 | 20 | 20 | 40 |
| 30 | 30 | 40 | 50 | 60 |
| 20 | 60 | 40 | 50 | 50 |
| 10 | 50 | 50 | 40 | 60 |

Find the mean.

4. The table below shows the frequency distribution of marks obtained by 70 students in a model examination:

| Marks | Frequency |
|---|---|
| 10–19 | 3 |
| 20–29 | 7 |
| 30–39 | 15 |
| 40–49 | 20 |

| | |
|---|---|
| 50–59 | 15 |
| 60–69 | 10 |
| 70–79 | 0 |

Calculate the mean.

## II. Determination of the Median of a Grouped Data

### (i) Formula for calculating the median of a grouped data

To find the median of grouped data, the formula below is employed:

$$\text{Md} = L + \left( \frac{\frac{N}{2} - n_b}{n_w} \right) i$$

where,
L = real lower limit or lower boundary of the median class
N = number of cases
$n_b$ = cumulative frequency of scores just below the median
$n_w$ = frequency of interval containing the median
i = class size

### (ii) Application of the formula

**Worked Example 2**

The frequency distribution below shows the marks of 120 students in a Biology test:

| Marks | Frequency |
|---|---|
| 30–39 | 11 |
| 40–49 | 25 |
| 50–59 | 45 |

| Marks | Frequency |
|---|---|
| 60–69 | 15 |
| 70–79 | 10 |
| 80–89 | 14 |

Calculate the median of the distribution.

**SOLUTION**

From the frequency table,

$$\text{Median class} = 50 - 59$$

$$L = 50 - 0.5 = 49.5$$

$$N = \Sigma f = 120$$

$$n_b = 11 + 25 = 36$$

$$n_w = 45$$

$$i = 59.5 - 49.5 = 10$$

$$\text{Md} = L + \left(\frac{\frac{N}{2} - n_b}{n_w}\right) i$$

$$= 49.5 + \left(\frac{\frac{120}{2} - 36}{45}\right) 10$$

$$= 49.5 + \frac{60 - 36}{45} \times \frac{10}{1}$$

$$= 49.5 + \frac{24}{45} \times \frac{10}{1}$$

$$= 49.5 + \frac{16}{3}$$

$$= 49.5 + 5.33$$

$$= 54.83$$

**Worked Example 3**

The table below shows the distribution of the heights of trees in a farm:

| Heights of trees (cm) | Number of trees |
|---|---|
| 151–175 | 3 |
| 176–200 | 6 |
| 201–225 | 19 |
| 226–250 | 18 |
| 251–275 | 13 |
| 276–300 | 4 |

Calculate the median height.

**SOLUTION**

From the frequency table,

$$\text{Median class} = 226 - 250$$

$$L = 226 - 0.5 = 225.5$$

$$N = \Sigma f = 63$$

$$n_b = 3 + 6 + 19 = 28$$

$$n_w = 18$$

$$i = 250.5 - 225.5 = 25.0$$

$$\text{Formular: Md} = L + \left(\frac{\frac{N}{2} - n_b}{n_w}\right) i$$

$$= 225.5 + \left(\frac{\frac{63}{2} - 28}{18}\right) \times 25$$

$$= 225.5 + \left(\frac{63 - 56}{2} \times \frac{1}{18}\right) \times 25$$

$$= 225.5 + \frac{7}{2} \times \frac{1}{18} \times \frac{25}{1}$$

$$= 225.5 + \frac{7}{36} \times \frac{25}{1}$$

$$= 225.5 + \frac{175}{36}$$

$$= 225.5 + 4.86$$

$$= 230.36$$

## III. Determination of the Mode of a Grouped Data

The mode of a distribution is the most frequently occurring score in the distribution.

For instance, 4, 4, 5, 6, 6, 6, 8, 8, 9, 10 has its mode as 6.

However, in the case of grouped data, the class interval with the largest number of cases is where the mode is located. The mid-point of the modal class where the mode is located constitutes the mode. It is possible to have more than one class interval having the highest frequency.

Frequency distribution with two 2 modal classes is called bi-modal and the one with more than 2 modal classes is referred to as multi-modal.

The mode of a grouped data is obtained by using the formula.

$$\text{Mode} = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) c$$

where,

L = lower class boundary of the modal class

Î”$_1$ = y â^’ x and Î” $_2$ = y â^’ z

Where x = frequency of the next lower class

y = frequency of the modal class

z = frequency of the next higher class

c = size of modal class interval

**Worked Example 4**

The data below represents the distribution of income in sample households in a town in one of the states in Nigeria:

| Monthly basic income | Number of earners ($f$) |
|---|---|
| 5–15 | 19 |
| 15–25 | 30 |
| 25–35 | 20 |
| 35–45 | 12 |
| 45–55 | 6 |
| 55–65 | 6 |
| 65–75 | 3 |

SOLUTION
From the table, modal class = 15–25
y = frequency of the modal class = 30
L = lower class boundary of the modal class = 14.5
x = the frequency of the next lower class to the modal class = 19
z = the frequency of the next higher class to the modal class = 20
c = size of the modal class interval = 25.5 âˆ’ 14.5 = 11
Hence, Î”$_1$ = y âˆ’ x = 30 âˆ’ 19 = 11
Î”$_2$ = y âˆ’ z = 30 âˆ’ 20 = 10

$$\text{Mode} = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)c$$

$$= 14.5 + \left(\frac{11}{11+10}\right)11$$

$$= 14.5 + \frac{11}{21} \times 11$$

$$= 14.5 + \frac{121}{21}$$

$$= 14.5 + 5.76$$

$$= 20.26$$

## Worked Example 5

The following data shows the lengths of 30 pieces of iron rods in centimetres.

64 63 71 59 64 53 57 74 55 57
50 68 72 68 80 67 70 62 79 67
45 55 65 60 61 68 59 54 64 76

Using class intervals of 45–49, 50–54, 55–59, etc., construct a frequency table for the data, then calculate the mode.

**SOLUTION**

| Class interval | Tally | *f* |
|---|---|---|
| 45–59 | / | 1 |
| 50–54 | /// | 3 |
| 55–59 | ~~###~~ / | 6 |
| 60–64 | ~~###~~ // | 7 |
| 65–69 | ~~###~~ / | 6 |
| 70–74 | //// | 4 |
| 75–79 | // | 2 |
| 80–84 | / | 1 |

$$\text{Mode} = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)c$$

$$\Delta_1 = y - x \text{ and } \Delta_2 = y - z$$

Modal class $= 60 - 64$

$$L = 60 - 0.5 = 59.5$$

$$x = 6$$

$$y = 7$$

$$z = 6$$

$$c = 64.5 - 59.5 = 5$$

So, $\Delta_1 = y - x = 7 - 6 = 1$

$\Delta_2 = y - z = 7 - 6 = 1$

$$\text{Mode} = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)c$$

$$= 59.5 + \left(\frac{1}{1+1}\right)5$$

$$= 59.5 + \left(\frac{1}{2} \times \frac{5}{1}\right) = 59.5 + 2.5$$

$$= 62$$