Grade 10 – Book D
(Revised CAPS Edition)

CONTENT:

This book was compiled and processed by E.J. Du Toit in 2011.
Revised edition 2023.

Website: www.abcbooks.co.za

Also visit www.abcmathsandscience.co.za for extra exercises, tests and exam papers.

Chapter D1

Data handling – calculations

D1.1 Important terminology:

D1.1.1 Discrete data:

It is information collected by **counting**, for instance, the number of people who visit a clinic every day in example 2 on page 5.

D1.1.2 Continuous data:

It is information obtained through **measurement**. These data are usually measured to an acceptable degree of accuracy and therefore cannot necessarily be expressed precisely.

D1.1.3 Ungrouped data:

Ungrouped data is data that may each be classified on its own as in example 1 below. The data may be arranged to make counting easier, but it is not necessary.

E.g. 1 The following table shows the number of hours (X) per week a certain group of people watch TV.

5	7	2	4	3	3	5	7	6	5	6	2	3
3	4	7	2	5	6	3	5	7	3	4	2	

Here are only 6 different possibilities, $2, 3, 4, 5, 6$ or 7. Therefore, it is not necessary to group the data. It may be arranged if necessary:

2	2	2	2	3	3	3	3	3	3	4	4	4
5	5	5	5	5	6	6	6	7	7	7	7	

The above data may then be presented in an ungrouped frequency distribution table:

X	F	C_F	%	With:
7	4	4	16	X → number of hours per week
6	3	7	12	F → frequency
5	5	12	20	C_F → cumulative frequency
4	3	15	12	N → total frequency
3	6	21	24	% → percentage
2	4	25	16	
	$\sum F = N = 25$		100	

It is important to interpret such a Frequency Table in one glance, e.g.:
* From the above table most people participating in the survey spend 3 hours per week watching TV. They represent 24% of the participants.

D1.1.4 **Grouped data:**

Grouped data are arranged and divided into various class intervals because the count of the data on their own represent too large a variety.

E.g.2 The following table indicates the number of patients per day who visited Clinic A for a certain month:

12	23	15	38	26	5	38	11	25	27	42	11	23
23	45	26	18	9	14	33	24	36	20	6	24	

The above data may then be presented in a grouped frequency distribution table:

Class interval	Class boundaries	Class mid-point	Count	F	C_F	%
1 – 10	0, 5 – 10, 5	5, 5	///	3	3	12
11 – 20	10, 5 – 20, 5	15, 5	~~HH~~ //	7	10	28
21 – 30	20, 5 – 30, 5	25, 5	~~HH~~ ////	9	19	36
31 – 40	30, 5 – 40, 5	35, 5	////	4	23	16
41 − 50	40, 5 – 50, 5	45, 5	//	2	25	8
				N = 25		100

This Frequency table may also be interpreted at one glance:

* **The largest number of people visiting the clinic every day falls within the interval 21 – 30.**
* **The smallest number of people visiting the clinic every day falls within the interval 41 – 50.**

 # Other deductions may also be made that we will discuss later!

D1.1.5 **Range:**

The range is the difference between the highest and the lowest values of the applicable set of data. The range is also important in the interpretation of data as will be discussed later in the examples.

E.g.3 Determine the range of the data in E.g.2

****

Range = 45 − 5 = 40

D1.2 **Central tendencies:**

Sometimes it is necessary to compare one set of data to another set of data. Then it is difficult to compare only the raw sets of data. Each set of data's central tendencies are then determined and the central tendencies are then compared. The central tendencies therefore only have meaning when they are compared or interpreted. In some cases, it is better to use only one of the central tendencies.

D1.2.1 <u>Mode</u>:

This is the count that occurs the **most**. It is also known as the modal value.

E.g.4 Consider the following data: 2, 2, 2, 3, 4, 4, 4, 5, 5, 6

Therefore the mode is 2 and 4 as it is the counts that occurs the most.

D1.2.2 <u>Median</u>:

This is the count that occurs precisely in the **middle.** If there is an uneven number of counts, it is easy to determine the median, but if there is an even number of counts, the average between the two middle counts are calculated.

E.g.5 Consider the following data: 2, 2, 2, 3, 3, 4, 4, 4, 5, 6

The median will therefore be between the 5th and 6th count because the number of counts even.

$$\therefore \text{ Median } = \frac{3 + 4}{2} = \frac{7}{2} = \ 3,5$$

D1.2.3 <u>Mean</u>:

The mean may be a precise calculation, but sometimes only an estimated mean is calculated with grouped data.
The mean is calculated as follows:

$$\text{Mean } = \frac{\text{The sum of all the counts}}{\text{The number of counts}}$$

E.g.6 Consider the following data: 2, 2, 2, 3, 3, 4, 4, 4, 5, 6

$$\therefore \text{ Mean } = \frac{2 + 2 + 2 + 3 + 3 + 4 + 4 + 4 + 5 + 6}{10} = \frac{35}{10} = 3,5$$

D1.2.4 <u>Modal class interval</u>:

It is the interval with the highest frequency. This tendency is only used in grouped data.

D1.3 <u>Application of central tendencies - ungrouped data</u>:

E.g.7 Consider the following table and then answer the questions below. The table shows the number of cars Mr. Bopela, a car dealer, sold per week.

7	4	6	8	7	4	5	6	6	8	1	7	5	6	8

(a) Arrange the data.
(b) Are these data discrete or continuous? Motivate.
(c) Present the data above as an ungrouped frequency distribution table.
(d) Determine the central tendencies of the data.
(e) Which of the three central tendencies is the best to use? Motivate your answer.

****

(a)

1	4	4	5	5	6	6	6	6	7	7	7	8	8	8

(b) These are discrete data because the motor sales could have been counted weekly.

(c)

X	F	C_F	%
8	3	3	20
7	3	6	20
6	4	10	27
5	2	12	13
4	2	14	13
3	0	14	0
2	0	14	0
1	1	15	7
	N = 15		100

(d) **Mode = 6** → occurs the most times

Median = 6 → 8th count is the middle count

Mean = $\dfrac{1+4+4+5+5+6+6+6+6+7+7+7+8+8+8}{15}$ = $\dfrac{88}{15}$ = **5,9**

(e) because all three the central tendencies are close to each other, any of the three may be used!

Exercise 1:   Date:_____

(1) Determine the central tendencies of the following set of data:

24 ; 45 ; 34 ; 23 ; 54 ; 34 ; 25 ; 56 ; 36 ; 44

_____

_____

_____

_____

_____

_____

_____

(2) The daily temperature, for a certain month, in Bloemfontein measured at 12 o'clock in the afternoon is as follows:

18	21	19	22	18	19	20	24	22	18	17	17	18	18
20	20	22	24	20	18	19	17	16	15	15	15	16	19
19	18												

(a) Represent the data above as an ungrouped frequency distribution table without arranging the data first. (Use page 8.)

(b) Use the table and predict the average temperature for this month.

_____

(c) Are these data discrete or continuous? Motivate.

_____

_____

_____

(d) Calculate the central tendencies for the data above.

_____

_____

_____

_____

_____

_____

(e) Which central tendency is the best to use? Motivate.

_____

_____

(f) Which possible month(s) of the year do you think is relevant here? Motivate fully.

_____

_____

_____

_____

D1.4 Application of central tendencies - grouped data:

E.g.8 Consider the following table and then answer the questions below. The table shows the daily rainfall (in mm.) of City X as measured over a period of 5 weeks:

4	31	21	0	0	45	23	3	11	17	4	0	2	12
13	37	48	7	11	49	0	12	0	0	40	17	9	15
0	33	45	12	6	4	47							

(a) **Is this data discrete or continuous? Motivate.**
(b) **Present the data above as a grouped frequency distribution table.**
(c) **Determine the modal class interval of the data.**
(d) **Determine the estimated average rainfall of City X.**
(e) **Which is the best to use – modal class or estimated mean?**
     Motivate your answer.
                              ******
(a) **Continuous, because rainfall is measured to a certain degree of accuracy.**

(b)

Class interval	Class boundaries	Class mid-point	Count	F	C_F	%
0 – 9	0 – 9,5	4,5	~~HHH HHH HHH~~	15	15	42,9
10 – 19	9,5 – 19,5	14,5	~~HHH~~ IIII	9	24	25,7
20 – 29	19,5 – 29,5	24,5	II	2	26	5,7
30 – 39	29,5 – 39,5	34,5	III	3	29	8,6
40 – 49	39,5 – 49,5	44,5	~~HHH~~ I	6	35	17,1
				N = 35		100

(c) Modal class interval is 0 – 9, because this class has the highest frequency, namely 15.

(d) Estimated average $= \dfrac{4,5 \times 15 + 14,5 \times 9 + 24,5 \times 2 + 34,5 \times 3 + 44,5 \times 6}{35}$

$= \dfrac{67,5 + 130,5 + 49 + 103,5 + 267}{35}$

$= \dfrac{617,5}{35}$

$\approx 17,6$ **mm per day**

(e) The estimated mean is a better indication of the central tendency as there are many counts of 0 in the modal class interval.

<u>Exercise 2:</u>   Date: _____

(1) (a) Represent the table below as a grouped frequency distribution table. Choose four suitable class intervals.

3	5	12	11	11	2	16	11	19	7	11
15	11	15	17	1	13	9				

(b) Determine the data's range in (a): _____

(c) Determine the mode of the data in (a): _____

(d) Determine the modal class interval of the data in (a): _____

(e) Does the mode lie in the modal class interval? _____

(f) Determine the median of the data in (a): _____

_____

_____

_____

(g) Determine the mean of the data in (a): _____

_____

(h) Which of the central tendencies is the best indication of the data in the table?
Motivate your answer.

_____

_____

(2) The table shows the ages of all the members of Green Trees Golf Club whose birthdays are in January.

12	34	24	41	50	56	32	21	19	27	39	45
22	12	35	71	64	23	56	14	47	25	27	31
40	19	33	38	43	30	61	38	24	32	39	47
16	35	23	45	38	26	70	28	33	38	23	29
36	41	12	38	35	51	43					

(a) Calculate the range of the data in the table: _____

(b) Represent data above as a grouped frequency distribution table. Choose suitable class intervals. Clearly indicate the class boundaries and class mid-points.

(c) Determine the data's modal class interval: _____

(d) Determine the estimated mean of the data in the grouped frequency distribution table in (b). Show all calculations.

_____

_____

_____

(e) Is the modal class interval or estimated mean the best to use? Motivate your answer.

_____

_____

(3) Study the grouped frequency distribution table below and answer the questions below. The table shows a summary of grade 10C's Mathematics marks.
The class intervals are in terms of percentage.

Class interval	Class boundaries	Class mid-point	F	C_F	%
0 − 9			1		
10 − 19			0		
20 − 29			1		
30 − 39			3		
40 − 49			5		
50 − 59			6		
60 − 69			7		
70 − 79			5		
80 − 89			5		
90 − 99			2		
			N =		

(a) Complete the table. The percentage column must be rounded off to the closest integer. Show all calculations.

_____

_____

_____

(b) Calculate the estimated class average for grade 10C's Mathematics test.

_____

_____

_____

_____

(c) Determine how many learners passed if the passing rate is 40%.

_____

(d) Determine the modal class interval: _____

(e) Discuss the measurements for central tendencies.

_____

_____

_____

D1.5 Stem-and-leaf representations:
D1.5.1 Single stem-and-leaf representation:

E.g.9 The heights (in cm) of the members of a school's choir are as follows:

132	145	142	121	144	176	165	150	132	140
157	172	120	133	167	171	144	149	134	146
151	166	154	134	176	144	154	133	151	135

(a) **Draw a stem-and-leaf representation of the table.**
(b) **Determine the mode of the data in the table.**
(c) **Determine the median of the data in the table.**
(d) **Calculate the mean. Show all calculations.**
(e) **Which of the central tendencies are best to determine the average height of the members of the choir?**
(f) **How tall is the tallest member of the choir?**

****

(a) **Unarranged:**

	Heights of choir members:								
Stem:	**Leaves:**								**Frequency (F):**
12	1	0							2
13	2	2	3	4	4	3	5		7
14	5	2	4	0	4	9	6	4	8
15	0	7	1	4	4	1			6
16	5	7	6						3
17	6	2	1	6					4
	Key: 15 − 7 represents 157								30

If you like, you may first do the representation above as unarranged and then as arranged below or you may arrange it immediately and therefore draw the representation below immediately. The arranged representation may be later used to determine the mode and median.

Arranged:

	Heights of choir members:							
Stem:	**Leaves:**							**Frequency (F):**
12	0	1						2
13	2	2	3	3	4	4	5	7
14	0	2	4	4	4	5	6	9	8
15	0	1	1	4	4	7		6
16	5	6	7					3
17	1	2	6	6				4
	Key: 15 − 7 represents 157							30

(b) **Mode = 144** ∴ It is the count that occurs the most.

(c) To calculate the median, the middlemost count must be found first. In this case there are 30 counts and the middlemost count will be between count 15 and count 16.

Count 15 = 14⑤ and Count 16 = 14⑥.

$$\therefore \text{ Median } = \frac{145 + 146}{2} = \frac{291}{2} = \boxed{145,5}$$

(d) To calculate the mean, you may use a table as follows:

Stem × 10   F   Sum of leaves
   ↓   ↓   ↓

120	×	2	+	1	=	241
130	×	7	+	23	=	933
140	×	8	+	34	=	1 154
150	×	6	+	17	=	917
160	×	3	+	18	=	498
170	×	4	+	15	=	695
				Total:	=	4 438

$$\therefore \text{ Mean } = \frac{438}{30} \approx \boxed{147,9}$$

(e) All three the central tendencies are remarkably close to each other. Therefore, it might be best here to use the mean to describe the height of the choir members.

(f) The tallest choir member is 176 cm tall.

Exercise 3:   Date: _____

(1) The number of births per day in Gauteng was recorded as follows for March 2002:

385	410	341	390	376	435	431	409	348
411	385	397	388	402	415	390	410	417
385	435	437	342	415	347	438	387	403
410	411	341	400					

(a) Illustrate the data (on page 20) as a stem-and-leaf representation. Give a clear key.

(b) What is the highest number of births per day recorded in Gauteng for March 2002?

_____

(c) What is the lowest number of births per day recorded in Gauteng for March 2002?

_____

(d) For how many days were less than 400 births per day recorded? _____

(e) Determine the median and mode of the data.

_____

(f) Calculate the average number of births recorded per day in Gauteng for March 2002.

_____

_____

(g) Which central tendencies are best to describe the average number of births per day? Motivate.

_____

_____

(2) Bags of tomatoes are sold at John's vegetable stall. He first fills the bags, weighs each bag and sells the tomatoes at R7,25 per kg. After he has weighed the day's bags, he records the weights as a stem-and-leaf representation. Study the representation and answer the questions below:

Weight per bag (in gram):		
Stem:	Leaves:	Frequency (F):
31	0 3 3 6	
32	2 4 4 8 9	
33	3 5 5 5 5 6	
34	0 1 2 4 6 8 9	
35	2 2 5 5 7	
36	5 6 6 6 7 7 8 8	
Key: 33 − 5 represents 335		

(a) Complete the table.

(b) How many bags in total did John fill for the day? _____

(c) How much does: (i) the lightest bag of tomatoes weigh? _____

(ii) the heaviest bag of tomatoes weigh? _____

(d) Calculate the average weight of the bags.

_____

_____

(e) If John sells all the bags for the day, calculate his total profit if he acquired the tomatoes for R58,20 in total.

_____

_____

_____

_____

_____

(3) During an observation assignment, Sandra must count the cars that turn at a certain traffic light (per hour). During the three days of observation, she recorded the following:

56	71	64	33	29	11	23	34	34	45	76
54	56	65	74	23	34	16	11	66	45	39
60	54	70	54	68	23	12	26	34	54	68

(a) Present the data in a stem-and-leaf representation.

(b) Determine the mode and median of the data.

_____

_____

(c) Determine the mean for the data.

_____

_____

(d) Which measure for central tendency is the best to determine the average number of cars (per hour) that moved past this point. Motivate.

_____

_____

_____

(e) If Sandra did the same number of hours of observation the first and second day, but only spent $\frac{3}{4}$ of the previous days' time on the observation on the third day, calculate the number of hours she spent observing per day.

_____

_____

_____

_____

_____

_____