



Grade 12 Data Handling

2025 Trial Papers

QUESTION 1

The ages (in years) of visitors that visited a bed and breakfast (B&B) on a specific day were recorded as follows.

15	17	20	20	20	21	22	24
29	29	30	$11t$	$11t + 2$	36	38	55

- 1.1 Identify the mode of the data. (1)
 - 1.2 It is further given that interquartile range of the ages of visitors is 14. Calculate the value of t . (3)
 - 1.3 Calculate the mean ages of the visitors if $t = 3$. (2)
 - 1.4 Calculate the standard deviation of the data. (1)
 - 1.5 How many visitors have ages that are less than one standard deviation of the mean? (3)
- [10]**

QUESTION 2

An athlete's finishing times and the daily temperatures were tracked during 10 years of Comrades Marathon to investigate the relationship between temperature and race completion time. The table below shows the data from these 10 years.

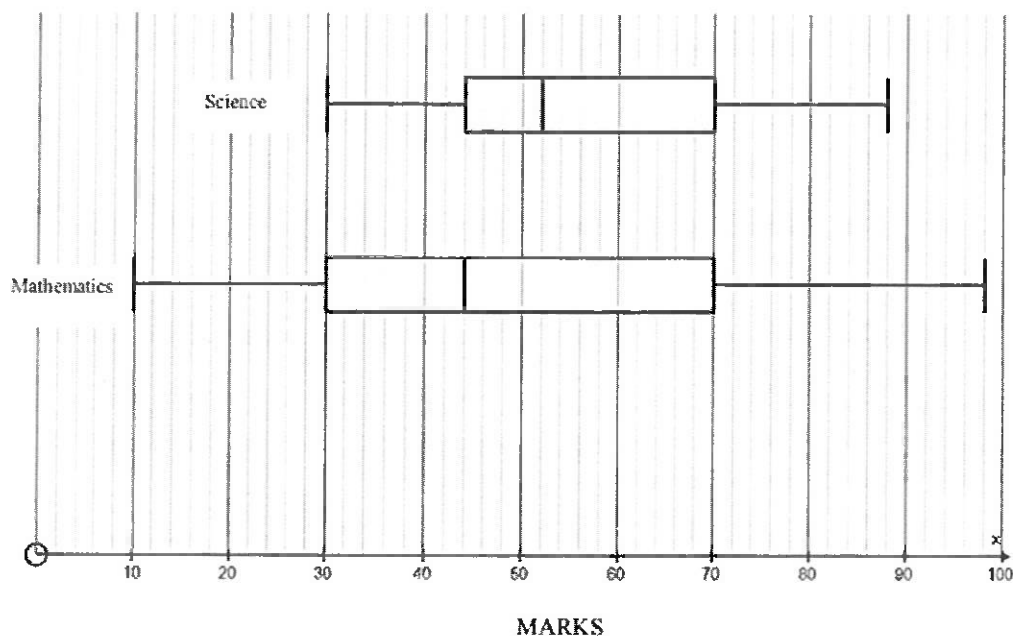
Year	1	2	3	4	5	6	7	8	9	10
Temperatures (degrees Celsius) x	41	9	30	15	25	20	20	35	39	16
Finishing times of athlete (in minutes) y	72	30	66	29	45	43	41	66	68	31

- 2.1 Calculate the correlation coefficient of the data. (1)
- 2.2 Comment on the strength of the relationship between temperatures and finishing time. (1)
- 2.3 Calculate the equation of the least squares regression line of the data. (3)
- 2.4 Use the equation of the least squares regression line to predict the temperatures if an athlete takes 57 minutes to finish the race. Round the answer off to the nearest whole number. (2)
- 2.5 Draw the least squares regression line on the grid provided in the SPECIAL ANSWER BOOK. (2)

[9]

QUESTION 1

The box and whisker diagrams below show the distribution of test scores obtained by a sample of students in Mathematics and Science.



Use the diagram to determine the following:

- 1.1 The range for Mathematics. (2)
 - 1.2 The median for Science. (1)
 - 1.3 The interquartile range for Science. (2)
 - 1.4 Which examination was easier for the students? Motivate your answer. (2)
 - 1.5 Which examination has a weaker spread of scores? Justify by giving TWO reasons. (2)
 - 1.6 If a learner from the Science sample is selected randomly, find the probability that the learner achieved a mark greater than 70%. (1)
- [10]**

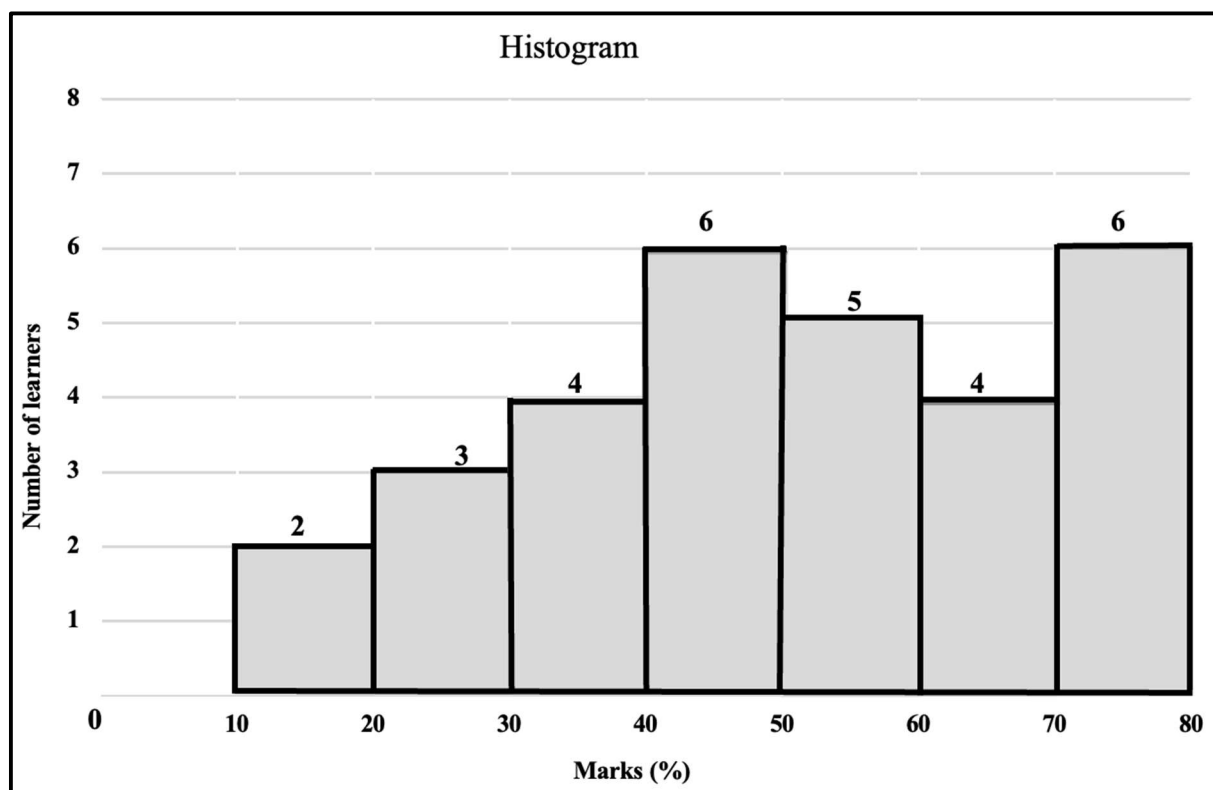
QUESTION 2

Refer to the statements below and answer the questions that follow:

- A least squares regression line (a line of best fit) has an equation of $y = 49 - 3x$
 - There is one outlier that has been identified as (9 ; 4)
 - The correlation coefficient is very strong
- 2.1 If the outlier was removed, would the correlation coefficient be stronger?
(Explain your answer.) (1)
- 2.2 Is the correlation coefficient closer to one or negative one? (Explain your answer.) (2)
- 2.3 If a line of best fit is used, would it be perfectly accurate? (Explain your answer.) (2)
- 2.4 An individual predicts that if x has a value of 30, y will have a value of -41 . However, his friend explains that he is extrapolating and that his result is inaccurate. Explain what his friend is referring to. (2)
- [7]

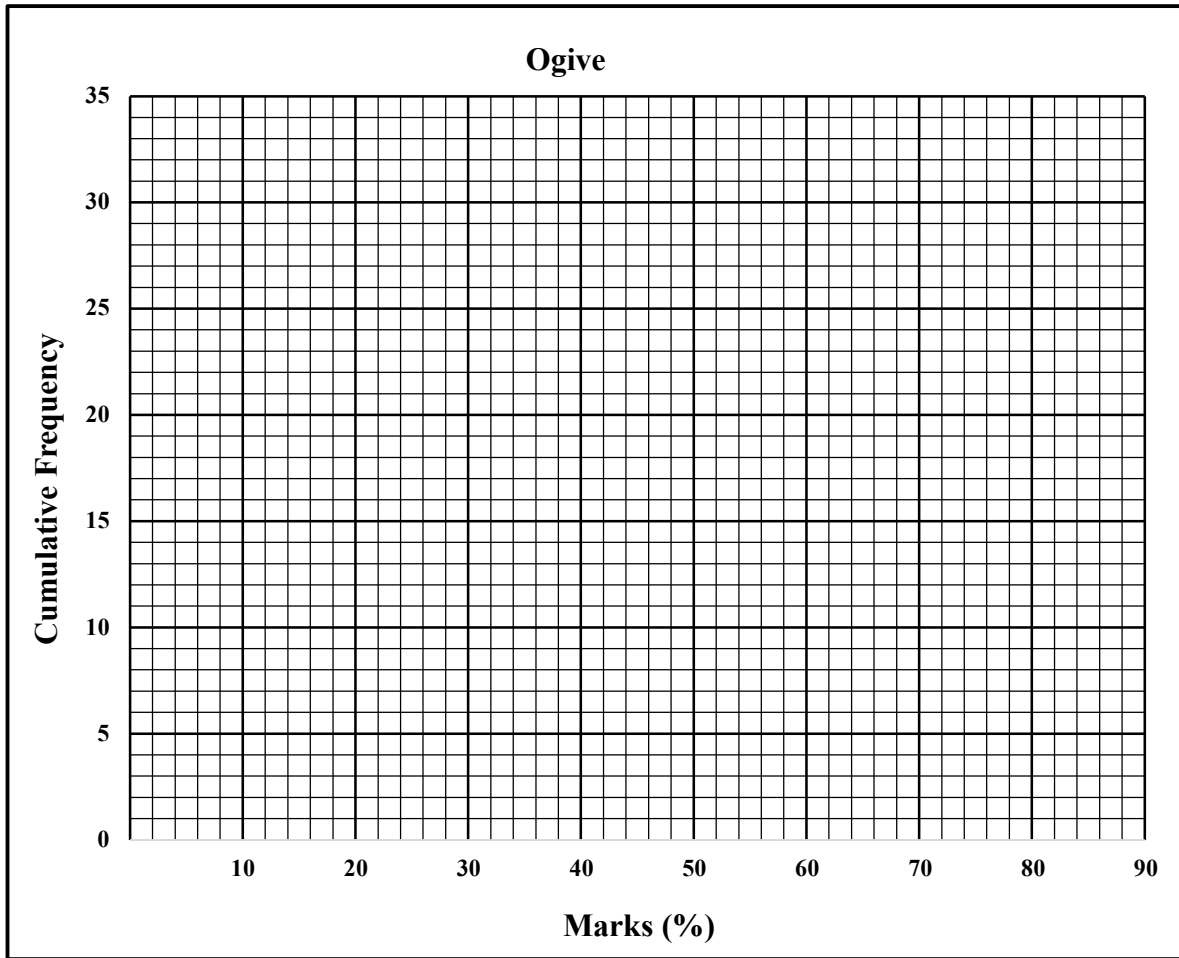
QUESTION 1

The histogram below shows the marks (in %) of Grade 11 learners in their November 2024 examination.



1.1	Complete the table below.																												
	<table> <tr> <th>Class Interval (Marks in %)</th><th>Frequency</th><th>Cumulative frequency</th></tr> <tr> <td>$10 \leq x < 20$</td><td></td><td></td></tr> <tr> <td>$20 \leq x < 30$</td><td></td><td></td></tr> <tr> <td>$30 \leq x < 40$</td><td></td><td></td></tr> <tr> <td>$40 \leq x < 50$</td><td></td><td></td></tr> <tr> <td>$50 \leq x < 60$</td><td></td><td></td></tr> <tr> <td>$60 \leq x < 70$</td><td></td><td></td></tr> <tr> <td>$70 \leq x < 80$</td><td></td><td></td></tr> <tr> <td>TOTAL</td><td></td><td></td></tr> </table>	Class Interval (Marks in %)	Frequency	Cumulative frequency	$10 \leq x < 20$			$20 \leq x < 30$			$30 \leq x < 40$			$40 \leq x < 50$			$50 \leq x < 60$			$60 \leq x < 70$			$70 \leq x < 80$			TOTAL			(2)
Class Interval (Marks in %)	Frequency	Cumulative frequency																											
$10 \leq x < 20$																													
$20 \leq x < 30$																													
$30 \leq x < 40$																													
$40 \leq x < 50$																													
$50 \leq x < 60$																													
$60 \leq x < 70$																													
$70 \leq x < 80$																													
TOTAL																													

1.2 Draw an ogive (cumulative frequency graph) to represent the data.

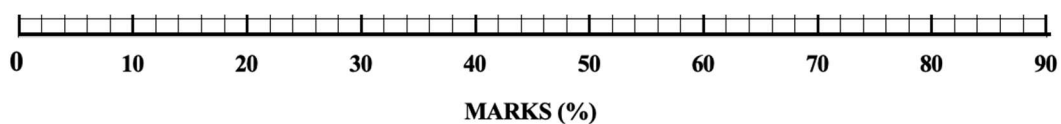


(3)

1.3 Use the ogive to estimate how many learners obtained at least 54% in the examination.

(2)

- 1.4 It is further given that the minimum mark is 12% and the range of the data is 66%.
Use the grid provided to draw a box-and-whisker diagram.



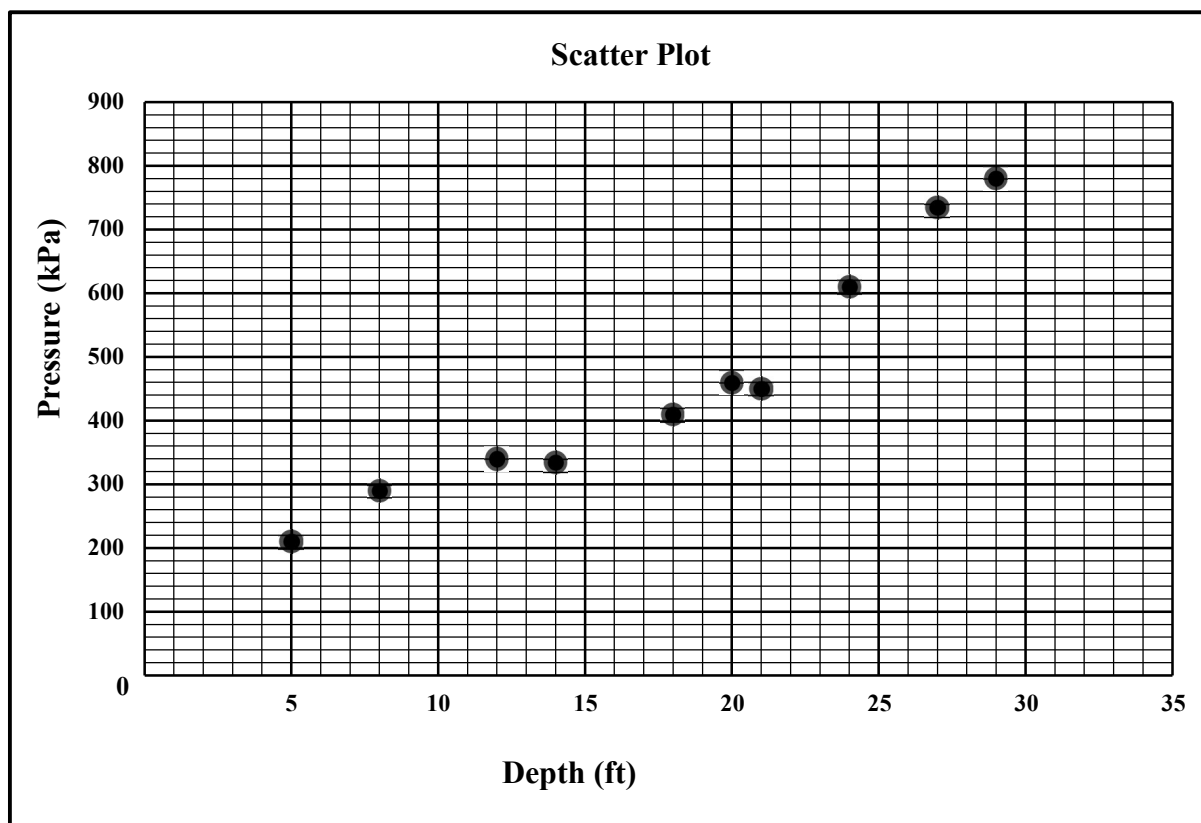
(4)

[11]

QUESTION 2

- 2.1 A town manager conducted a survey to measure the pressure of the sea level in his coastal city. The table and the scatter plot below shows the results of the survey:

Depth (ft)	5	8	12	14	18	20	21	24	27	29
Pressure (kPa)	210	290	340	335	410	460	450	610	735	780



2.1.1	Determine the equation of the least squares regression line for the data.	(3)
2.1.2	Use the scatter plot provided above to draw the least squares regression line.	(2)
2.1.3	Write down the correlation coefficient.	(1)

Research was conducted to measure the testosterone levels in runners at the Olympic Games. The table below shows the results of five runners who were tested for testosterone.

Runners	1	2	3	4	5
Level of testosterone	6,7	$2k + 2,4$	p	3,8	4,3

The mean of the five runners is 7,7 and the values of the last three runners form an arithmetic sequence. Determine the value of k and p .

(5)

[11]

QUESTION 1

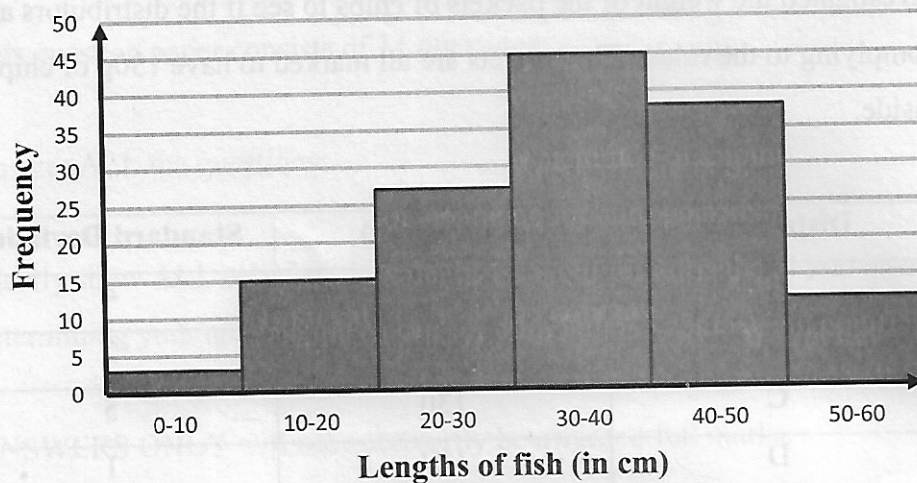
- 1.1 The table below shows the average weight and standard deviation of packets of chips from different distributors. Independent quality controllers investigated the weight of the packets of chips to see if the distributors are complying to the rules. The packets are all marked to have 130g of chips inside.

Distributor	Average (g)	Standard Deviation
A	125	4
B	132	2
C	130	5
D	129	1
E	134	3

- 1.1.1 If 2000 packets of chips were investigated from Distributor C, what was the total weight of chips tested? (1)
- 1.1.2 Which one of the Distributors is the best to buy from? Give a reason for your answer. (2)
- 1.1.3 Which one of the Distributors will you rather not buy from? Give a reason for your answer. (2)

- 1.2 The following histogram represents the lengths and number of fish caught in a freshwater angling competition. The shortest fish was 9cm and the longest one was 58cm.

Fish caught in freshwater angling competition



- 1.2.1 Determine the range of the lengths of the fish.

(1)

- 1.2.2 Complete the cumulative frequency table in the ANSWER BOOK.

Interval	Frequency	Cumulative frequency
$0 < x \leq 10$	3	
$10 < x \leq 20$	15	
$20 < x \leq 30$	27	
$30 < x \leq 40$	45	
$40 < x \leq 50$	38	
$50 < x \leq 60$	12	

(2)

- 1.2.3 Draw an ogive in the ANSWER BOOK to represent the data.

(3)

- 1.2.4 60% of the fish are less than x cm long. Use the ogive to determine the value(s) of x .

(2)

[13]

QUESTION 2

The top 10 athletes in a certain province took part in the Provincial Biathlon championships. They had to participate in 10 km cycling and 3 km swimming. The following table represents the times for each athlete, for cycling and swimming respectively, recorded in the championship.

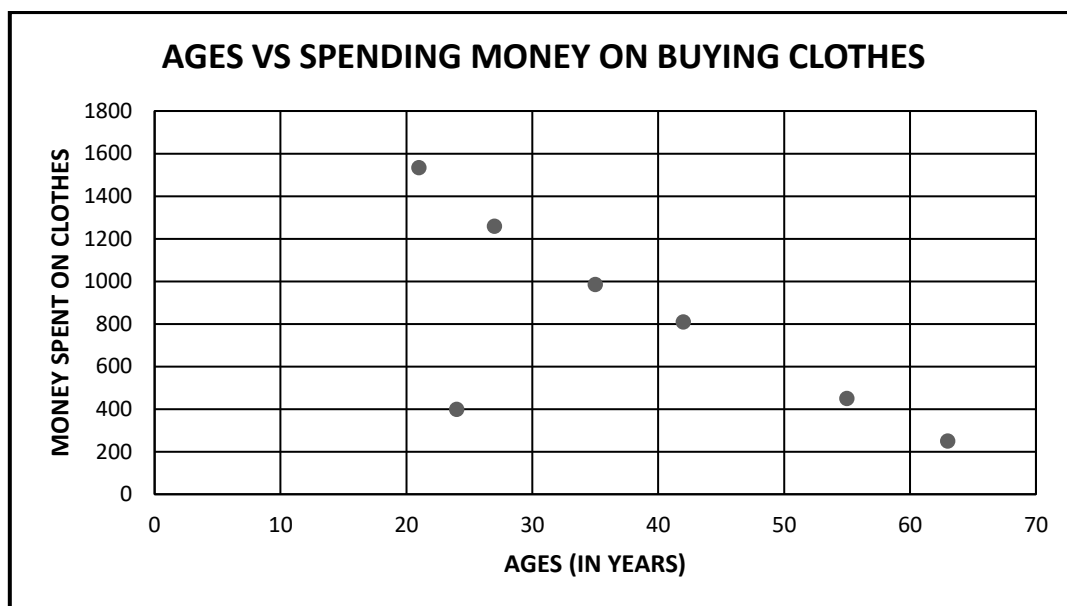
Athlete	Cycling (in minutes)	Swimming (in minutes)
1	22	39
2	15	40
3	24	47
4	19	38
5	22	55
6	19	45
7	26	58
8	30	50
9	20	44
10	35	65

- 2.1 Determine the average cycling time for the 10 athletes. (2)
- 2.2 Determine the equation for the least squares regression line of the data. (3)
- 2.3 Calculate the expected time for an athlete to complete the swimming event if he takes 29 minutes to complete the 10km cycling. (2)
- [7]

QUESTION 1

The relationship between age and spending money on buying clothes per month has been studied for years. Research has shown the following results:

Ages (in years)	21	24	27	35	42	55	63
Spending money on buying clothes (in rands)	1534	400	1260	986	810	450	250

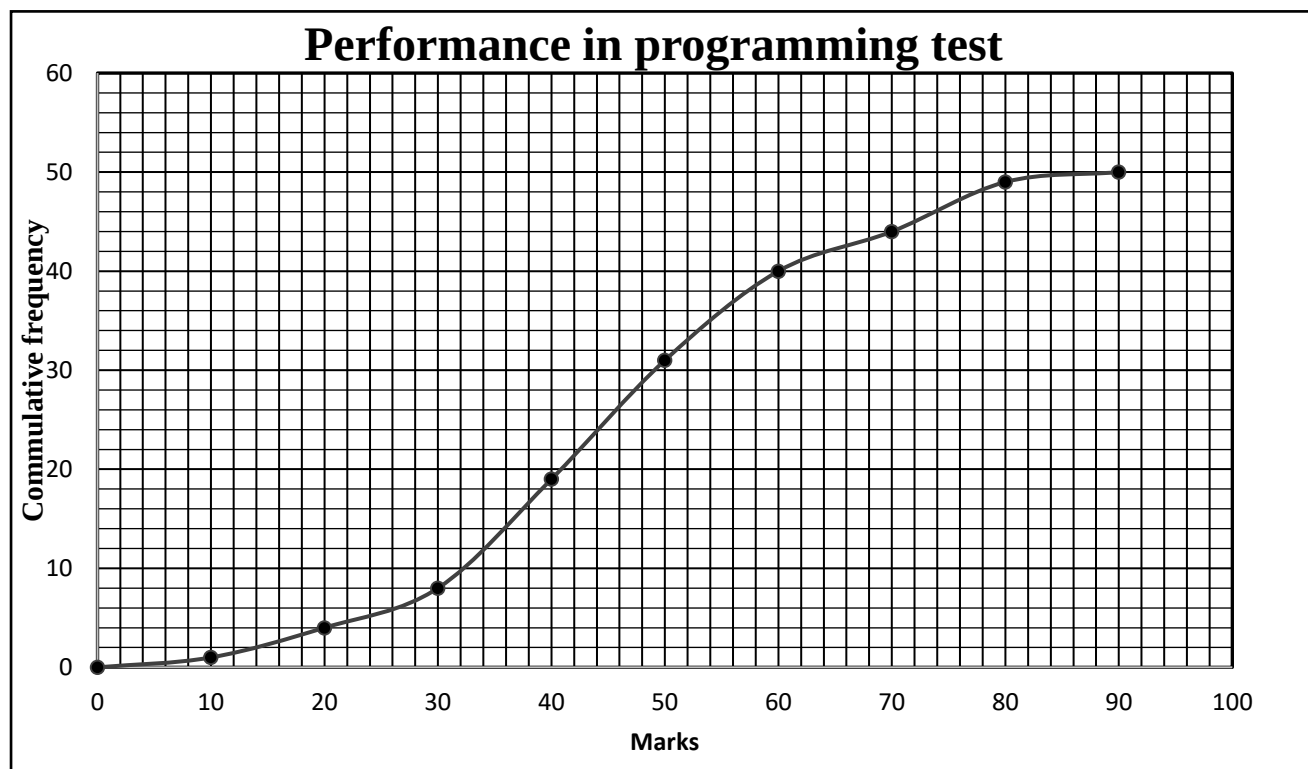


- 1.1 Determine the equation of the least squares regression line. (3)
- 1.2 Draw the least squares regression line on the answer book. (2)
- 1.3 Determine the correlation coefficient. (1)
- 1.4 Predict the money spend by a 50-year-old person on buying clothes. (2)
- 1.5 Is the answer calculated in QUESTION 1.4 reliable? Justify your answer. (2)
- 1.6 Explain the effect on the gradient of the regression line if the point (24; 400) is removed. (1)

[11]

QUESTION 2

- 2.1 The ogive (cumulative frequency graph) shows the performance of the students who took a test in basics programming skills. The test had a maximum of 100 marks.



- 2.1.1 How many students took the test? (1)
- 2.1.2 Only the top 25% of the students are allowed to do an advanced course in programming. Determine the cut-off mark for the top 25%. (2)
- 2.1.3 Complete a frequency table on the ANSWER BOOK provided. (2)
- 2.1.4 Calculate the mean estimate of the data. (3)



NSC

- 2.2 The average time to deliver a pizza ordered telephonically is 25 minutes. The standard deviation of the delivery time allowed, is 6 minutes. The delivery times in a day is summarised below.

Delivery times (minutes)
32
28
20
38
15
24

If the policy of the company is that you will receive the pizza free of charge when the delivery time is more than one standard deviation from the mean, determine the number of free pizzas for the day. Show all supporting calculations. (3)

[11]

QUESTION 1

- 1.1 On June 25, the points of the top 10 Rugby 7s teams on the rankings were as follows:

27	36	50	56	57	65	70	88	96	104
----	----	----	----	----	----	----	----	----	-----

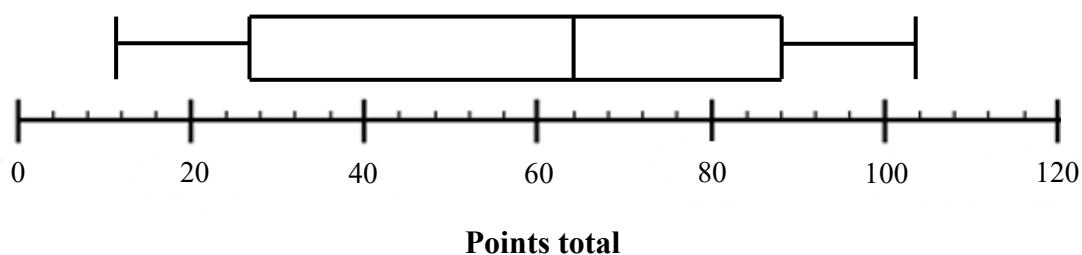
(Source: <https://www.svns.com/en/standings#men>)

- 1.1.1 Calculate the average (mean) of the 10 teams' points. (2)

- 1.1.2 Determine the standard deviation of the points. (2)

- 1.1.3 Determine the number of teams whose points lie within one standard deviation of the mean. (3)

- 1.2 If all 12 teams' point totals on the date are taken into account, the average is 59 points per team. The data is represented in the box-and-whisker diagram below:



- 1.2.1 Comment on the distribution of the data. Provide a reason for your answer. (2)

[9]

QUESTION 2

The term latitude refers to how far a location is from the equator. Latitudes in the Northern Hemisphere range from 0° at the equator to 90° N at the North Pole.

Below are the latitudes of several cities in the Northern Hemisphere along with the average maximum temperature for April in degrees Celsius.

CITY	LATITUDE (degrees)	AVERAGE MAXIMUM TEMPERATURES FOR APRIL ($^\circ\text{C}$)
Lagos, Nigeria	6	32
London, England	52	13
Calcutta, India	23	36
Rome, Italië	42	20
Moscow, Russia	56	8
Cairo, Egipt	30	28
San Juan, Puerto Rico	18	29

- 2.1 Draw a scatter plot for the information above on the diagram provided in the answer book. (2)
- 2.2 Determine the equation of the least squares regression line. (3)
- 2.3 The city of Madrid has a latitude of 40 degrees. Determine the estimated average maximum temperature for April for this city. (2)
- 2.4 Calculate the correlation coefficient of the data and describe the correlation between latitude and the average maximum temperature for April. (3)

[10]

QUESTION 1

The Road Traffic Inspectorate is doing research on the legibility distance, i.e. the maximum distance at which the driver of a vehicle can read a road sign. For their research they tested 12 drivers of different ages to determine their legibility distances. They hope to improve road safety by examining the relationship between age (x) measured in years and legibility distance (y) measured in metres. The table below lists the data obtained in this way.

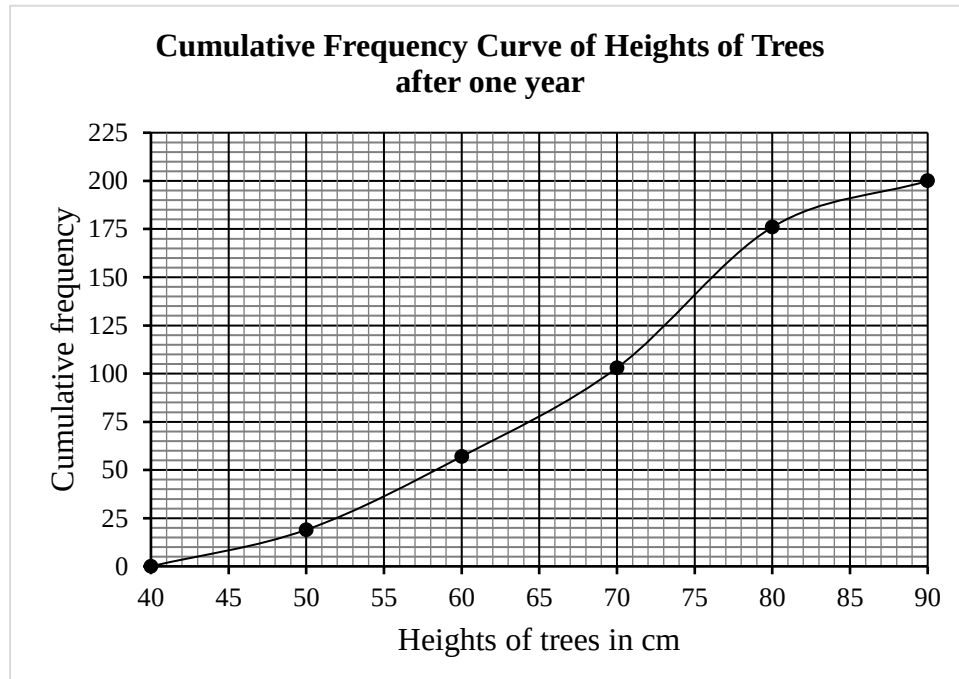
Age of driver in years (x)	19	24	28	29	32	35	49	55	63	74	79	82
Legibility distance in metres (y)	155	149	155	131	128	137	136	128	110	109	94	90

- 1.1 Determine the equation of the least squares regression line for the data. (3)
- 1.2 Write down the correlation coefficient. (1)
- 1.3 By referring to your answer to QUESTION 1.2, comment on the relationship between age and legibility distance. (1)
- 1.4 Predict the legibility distance for a 33-year-old. Give your answer correct to the nearest metre. (2)
- 1.5 Calculate the estimated decrease in legibility distance per age increase of 15 years, for people over 18. Give your answer correct to the nearest metre. (2)

[9]

QUESTION 2

A farmer germinated 200 seeds of a rare and valuable tree species and planted the seedlings in pots. The pots were placed in a greenhouse where they would grow faster than outside. After one year, the heights of the trees in the pots were measured and an ogive was drawn. The shortest tree was 41 cm high and the tallest tree 88 cm.



- 2.1 Draw a box and whisker diagram of the heights of the trees after one year, using the number line provided in the ANSWER BOOK. (3)
- 2.2 Describe the skewness of the data set. (1)
- 2.3 The trees can be sold as soon as they have reached a height of 75 cm. Will the farmer be able to sell more trees in the short term if the data set is skewed to the left or to the right? (1)
- 2.4 The farmer selected 14 of these trees to undergo special experimental treatment to make them grow faster. The heights of these trees, in cm, are as follows:

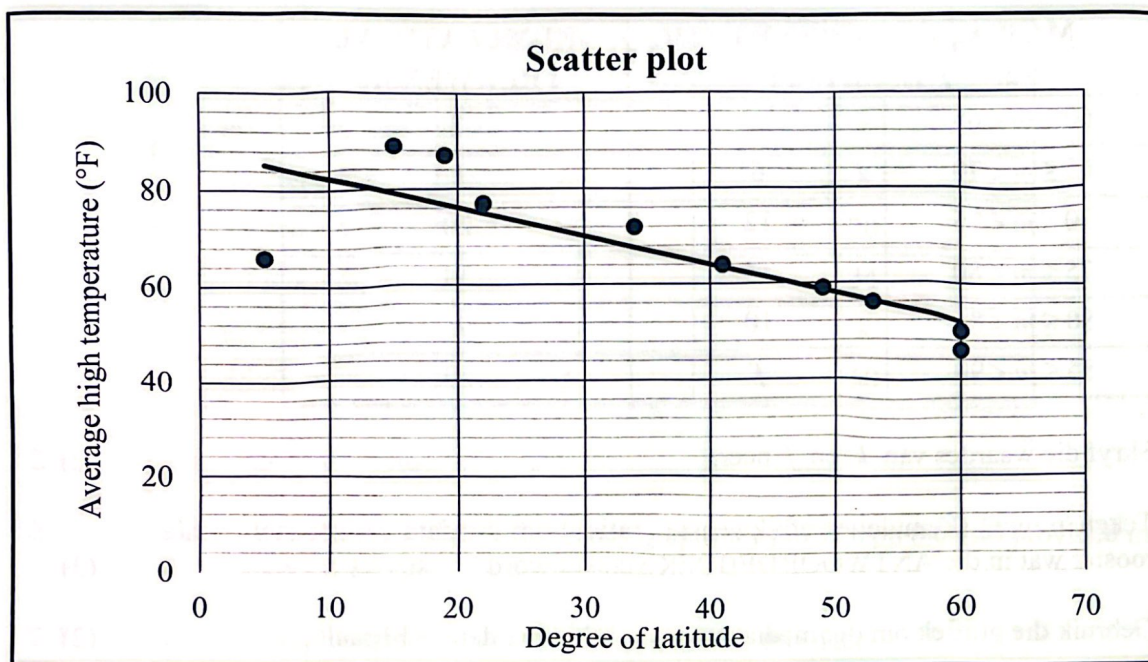
46	47	51	53	53	56	62	68	70	71	74	77	81	86
----	----	----	----	----	----	----	----	----	----	----	----	----	----

- 2.4.1 Calculate the standard deviation of the heights of these trees. (1)
- 2.4.2 How many of the trees have heights outside one standard deviation from the mean? Show all your working. (4)

[10]

QUESTION 1

The scatter plot below shows the degrees of latitude of 10 cities in the Northern Hemisphere and their average high temperatures (in °F). The least squares regression line is also shown.



- 1.1 Use the scatter plot to describe the strength of the correlation between the latitude of a city and its average high temperature. Motivate your answer. (2)

The table below shows the data that was used to draw the scatter plot above.

Latitude in degrees	5	19	34	53	22	41	60	15	60	49
Average high temperature (in °F)	66	87	72	56	77	64	46	89	50	59

- 1.2 Determine the equation of the least squares regression line for the data. (3)
- 1.3 Predict the average high temperature of a city with a latitude of 28 degrees. (2)
- 1.4 Write down the standard deviation of the average high temperature of the 10 cities. (1)
- 1.5 Determine the number of cities of which the average high temperature is greater than one standard deviation above the mean. (3)
- [11]**

QUESTION 2

Fifty boys were weighed and their mass recorded to the nearest kilogram. The results are shown in the table below.

MASS (kg)	FREQUENCY	CUMULATIVE FREQUENCY
$60 \leq m < 65$	2	2
$65 \leq m < 70$	6	8
$70 \leq m < 75$	12	20
$75 \leq m < 80$	14	34
$80 \leq m < 85$	10	k
$85 \leq m < 90$	f	50

- 2.1 Write down the values of k and f . (2)
- 2.2 Draw an ogive (cumulative frequency graph) for the data on the grid provided in the ANSWER BOOK. (3)
- 2.3 Use the graph to determine the median mass for this data. (2)
- 2.4 What percentage of the boys weighed 83 kg or more? (2)
- [9]

QUESTION 1

A survey was done under 100 members of a gymnasium about the cost of their annual membership contract.
Results showed an average annual payment of R3 240.

AMOUNT (R)	FREQUENCY
$0 \leq x < 1000$	8
$1\,000 \leq x < 2\,000$	12
$2\,000 \leq x < 3\,000$	P
$3\,000 \leq x < 4\,000$	30
$4\,000 \leq x < 5\,000$	Q
$5\,000 \leq x < 6\,000$	12

- 1.1 How many members payed less than R2 000 for their annual contract?

(1)
- 1.2 Proof by calculation that $P = 20$ and $Q = 18$.

(5)
- 1.3 Hence, complete the cumulative frequency table in the ANSWER BOOK.

(2)
- 1.4 Use the cumulative frequency table to draw an ogive to illustrate the information above.

(3)
- 1.5 Use the ogive to determine how many members spent more than R4 200 annually on their contracts.

(1)

[12]

QUESTION 2

Eight couples entered a dance competition. Two judges gave marks for their performances. The marks (out of 20) are shown in the table below.

COUPLE	1	2	3	4	5	6	7	8
JUDGE 1	18	4	6	8	5	12	10	14
JUDGE 2	15	6	3	5	5	14	8	15

- 2.1

Determine the equation of the least squares regression line of the marks given by the two judges.

(3)
- 2.2

A ninth couple had a late entry in the competition and received a mark of 15 by JUDGE 1. Calculate to the nearest integer, the mark they possibly would have received from JUDGE 2.

(2)
- 2.3

Are the judges consistent when allocating marks?
Motivate your answer with appropriate calculations.

(3)
[8]