SA-STUDENT

To pass high school please visit us at:

https://sa-student.com/

"You have to ask yourself how badly do you want something? If you really, really want something then put in the work". -Lewis Hamilton

QUESTION 1

Truck drivers travel a certain distance and have a rest before travelling further. A driver kept record of the distance he travelled (in km) on 8 trips and the amount of time he rested (in minutes) before he continued his journey. The information is given in the table below.

Distance travelled (in km) (x)	180	200	400	600	170	350	270	300
Amount of rest time (in minutes) (y)	20	25	55	120	15	50	40	45

1.1 Determine the equation of the least squares regression line for the data. (3)

1.2 If a truck driver travelled 550 km, predict the amount of time (in minutes) that he should rest before continuing his journey. (2)

1.3 Write down the correlation coefficient for the data. (1)

1.4 Interpret your answer to QUESTION 1.3. (1)

1.5 At each stop, the truck driver spent money buying food and other refreshments. The amount spent (in rands) is given in the table below.

100	150	130	200	50	180	200	190

1.5.1 Calculate the mean amount of money he spent at each stop. (2)

1.5.2 Calculate the standard deviation for the data. (1)

1.5.3 At how many stops did the driver spend an amount that was less than one standard deviation below the mean? (2)

[12]

QUESTION 2

At a certain school, the staff committee wanted to determine how many glasses of water the staff members drank during a school day. All teachers present on a specific day were interviewed. The information is shown in the table below.

NUMBER OF GLASSES OF WATER DRANK PER DAY	NUMBER OF STAFF MEMBERS
$0 \le x < 2$	5
$2 \le x < 4$	15
$4 \le x < 6$	13
$6 \le x < 8$	5
$8 \le x < 10$	2

2.1 Complete the cumulative frequency column provided in the table in the ANSWER BOOK. (2)

2.2 How many staff members were interviewed? (1)

2.3 How many staff members drank fewer than 6 glasses of water during a school day? (1)

2.4 The staff committee observed that k teachers were absent on the day of the interviews. It was found that half of these k teachers drank from 0 to fewer than 2 (that is $0 \le x < 2$) glasses of water per day, while the remainder of them drank from 4 to fewer than 6 (that is $4 \le x < 6$) glasses of water per day. When these k teachers are included in the data, the estimated mean is 4 glasses of water per staff member per day.

How many teachers were absent on the day of the interviews? (4)
[8]

QUESTION 1

1.1 The owner of a small company wishes to establish whether advertising in a regional newspaper is effective. The table below shows the amount spent on advertising and the corresponding sales figures for the last 9 years.

Amount spent on advertising (in rands) (x)	21 300	23 700	24 800	30 540	24 100	40 680	22 400	35 250	29 110
Sales (in rands) (y)	311 500	326 700	349 200	470 000	316 100	564 200	314 000	487 300	392 900

1.1.1 Determine the equation of the least squares regression line for the data. (3)

1.1.2 Predict the sales for a year in which the company will spend R28 500 on advertising. (2)

1.1.3 Write down the correlation coefficient of the data. (1)

1.1.4 Describe the association between the amount spent on advertising in the regional newspaper and the sales of this company. (1)

1.2 The profit that the small company made over the same 9 years is given in the table below.

Profit (in rands)	110 750	107 376	152 338	244 480	144 021	275 994	121 900	207 636	187 700

1.2.1 Calculate the mean profit made over the 9 years. (2)

1.2.2 Write down the standard deviation for the data. (1)

1.2.3 Determine the number of years in which the company made a profit that was greater than one standard deviation above the mean. (2)
[12]

QUESTION 2

The ages of the people who attended a music concert was summarised in the table below.

AGE	NUMBER OF PEOPLE
$5 < x \leq 15$	20
$15 < x \leq 25$	25
$25 < x \leq 35$	60
$35 < x \leq 45$	90
$45 < x \leq 55$	55
$55 < x \leq 65$	40
$65 < x \leq 75$	30

2.1 Write down the modal class of the data. (1)

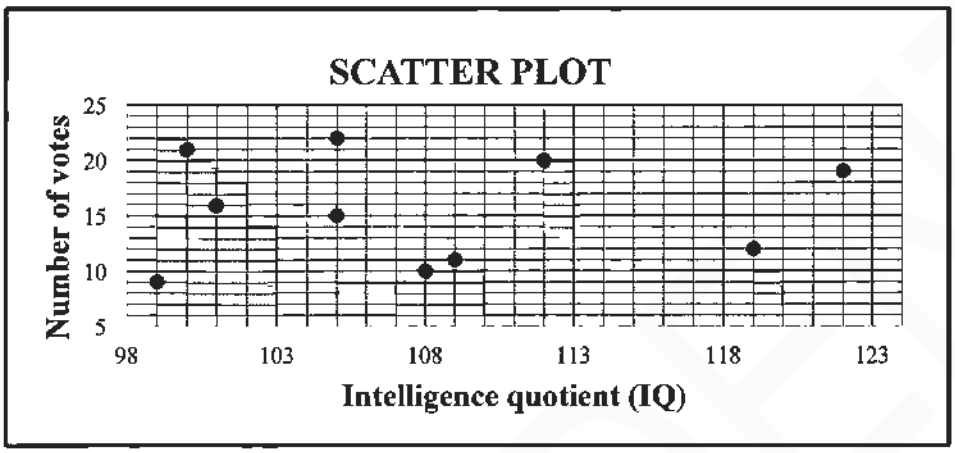
2.2 How many people attended the music concert? (1)

2.3 On the grid provided in the ANSWER BOOK, draw a cumulative frequency graph (ogive) to represent the above data. (4)

2.4 Use the cumulative frequency graph to determine the median age of the people who attended the music concert. (2)
[8]

QUESTION 1

The matric class of a certain high school had to vote for the chairperson of the RCL (representative council of learners). The scatter plot below shows the IQ (intelligence quotient) of the 10 learners who received the most votes and the number of votes that they received.



Before the election, the popularity of each of these ten learners was established and a popularity score (out of a 100) was assigned to each. The popularity scores and the number of votes of the same 10 learners who received the most votes are shown in the table below.

Popularity score (x)	32	89	35	82	50	59	81	40	79	65
Number of votes (y)	9	22	10	21	11	15	20	12	19	16

1.1 Calculate the:

1.1.1 Mean number of votes that these 10 learners received (2)

1.1.2 Standard deviation of the number of votes that these 10 learners received (1)

1.2 The learners who received fewer votes than one standard deviation below the mean were not invited for an interview. How many learners were invited? (2)

1.3 Determine the equation of the least squares regression line for the data given in the table. (3)

1.4 Predict the number of votes that a learner with a popularity score of 72 will receive. (2)

1.5 Using the scatter plot and table above, provide a reason why:

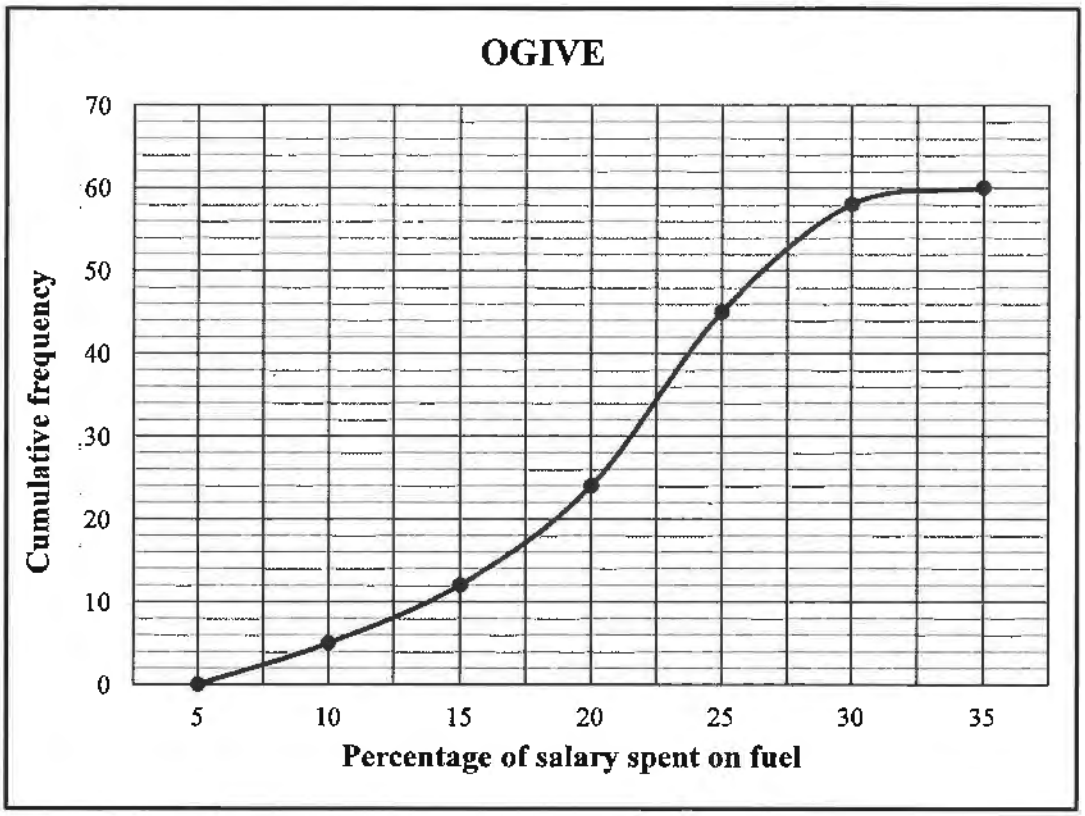
1.5.1 IQ is not a good indicator of the number of votes that a learner could receive (1)

1.5.2 The prediction in QUESTION 1.4 is reliable (1)

[12]

QUESTION 2

A company conducted research among all its employees on what percentage of their monthly salary was spent on fuel in a particular month. The data is represented in the ogive (cumulative frequency graph) below.



2.1 How many people are employed at this company? (1)

2.2 Write down the modal class of the data. (1)

2.3 How many employees spent more than 22,5% of their monthly salary on fuel? (2)

2.4 An employee spent R2 400 of his salary on fuel in that particular month. Determine the monthly salary of this employee if he spends 7% of his salary on fuel. (2)

2.5 The monthly salaries of these employees remains constant and the number of litres of fuel used in each month also remains constant. If the fuel price increases from R21,43 per litre to R22,79 per litre at the beginning of the next month, how will the above ogive change? (2)
[8]

QUESTION 1

The table below shows the mass (in kg) of the school bags of 80 learners.

MASS (in kg)	FREQUENCY
$5 < m \leq 7$	6
$7 < m \leq 9$	18
$9 < m \leq 11$	21
$11 < m \leq 13$	19
$13 < m \leq 15$	11
$15 < m \leq 17$	4
$17 < m \leq 19$	1

1.1 Write down the modal class of the data. (1)

1.2 Complete the cumulative frequency column in the table in the ANSWER BOOK. (2)

1.3 Draw a cumulative frequency graph (ogive) for the given data on the grid provided in the ANSWER BOOK. (3)

1.4 Use the graph to determine the median mass for this data. (2)

1.5 The international guideline for the mass of a school bag is that it should not exceed 10% of a learner's body mass.

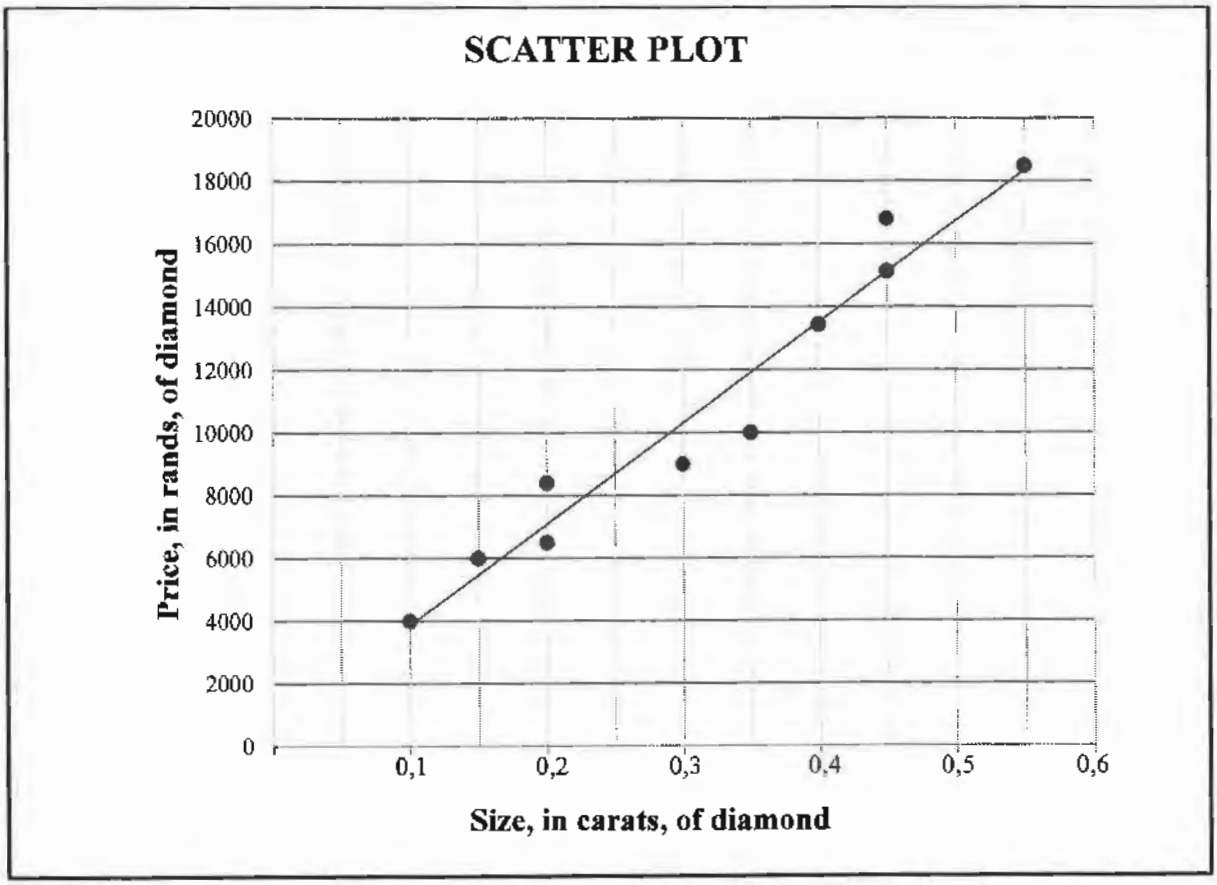
1.5.1 Calculate the estimated mean mass of the school bags. (2)

1.5.2 The mean mass of this group of learners was found to be 80 kg. On average, are these school bags satisfying the international guideline with regard to mass? Motivate your answer. (2)
[12]

QUESTION 2

The table below shows the size (in carats) and the price (in rands) of 10 diamonds that were sold by a diamond trader. This information is also presented in the scatter plot below. The least squares regression line for the data is drawn.

Size, in carats, of diamond (x)	0,1	0,15	0,2	0,2	0,3	0,35	0,4	0,45	0,45	0,55
Price, in rands, of diamond (y)	4 000	6 000	6 500	8 400	9 000	10 000	13 440	15 120	16 800	18 480



2.1 Determine the equation of the least squares regression line for the data. (3)

2.2 If the trader sold a diamond that was 0,25 carats in size, predict the selling price of this diamond in rands. (2)

2.3 Calculate the average price increase per 0,05 carat of the diamonds. (2)

2.4 It was later found that the selling price of the 0,35 carat diamond was recorded incorrectly. The correct price is R11 500. When this correction is made to the data set, the correlation between the size and price of these diamonds gets stronger. Explain the reason for this by referring to the given scatter plot. (1)
[8]

QUESTION 1

A bakery kept a record of the number of loaves of bread a tuck-shop ordered daily over the last 18 days. The information is shown in the table below.

10	11	13	14	14	15	16	18	18
19	19	20	21	35	35	37	40	41

1.1 Calculate the:

1.1.1 Mean number of loaves of bread ordered daily (2)

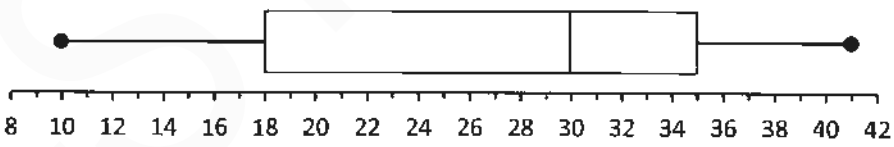
1.1.2 Standard deviation of the data (1)

1.1.3 Number of days on which the number of loaves of bread ordered was more than one standard deviation above the mean (2)

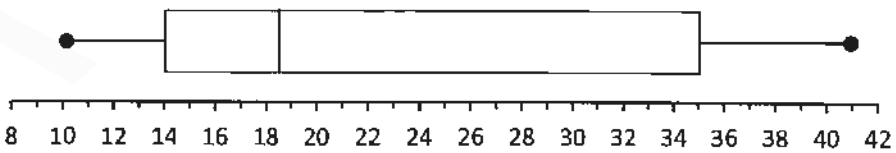
1.2 The tuck-shop owner was not able to sell all the loaves of bread delivered daily. He calculated the mean number of loaves sold over the 18 days to be 20. Calculate the number of loaves of bread which were NOT sold over the 18 days. (2)

1.3 One of the two box and whisker diagrams drawn below represents the data given in the table above.

Graph A:



Graph B:



1.3.1 Which ONE of the two box and whisker diagrams, drawn above, correctly represents the data? Write down a reason for your answer. (2)

1.3.2 Describe the skewness of the data. (1)

[10]

QUESTION 2

A farm stall sells milk in 5-litre containers to the local community. The price varies according to the availability of milk at the farm stall. The price of milk, in rands per 5-litre container, and the number of 5-litre containers of milk sold, are recorded in the table below.

Price of milk in rands per 5-litre container (x)	26	32	36	28	40	33	29	34	27	30
Number of 5-litre containers of milk sold (y)	48	30	26	44	23	32	39	29	42	33

2.1 On the grid provided in the ANSWER BOOK, draw the scatter plot to represent the data. (3)

2.2 Determine the equation of the least squares regression line for the data. (3)

2.3 If the farmer sells a 5-litre container of milk for R38, predict the number of 5-litre containers of milk he will sell. (2)

2.4 Refer to the correlation between the price of 5-litre containers of milk and the number of 5-litre containers of milk sold, and comment on the accuracy of your answer to QUESTION 2.3. (2)
[10]

QUESTION 1

1.1 Sam recorded the amount of data (in MB) that she had used on each of the first 15 days in April. The information is shown in the table below.

26	13	3	18	12	34	24	58	16	10	15	69	20	17	40

1.1.1 Calculate the:

(a) Mean for the data set (2)

(b) Standard deviation for the data set (1)

1.1.2 Determine the number of days on which the amount of data used was greater than one standard deviation above the mean. (2)

1.1.3 Calculate the maximum total amount of data that Sam must use for the remainder of the month if she wishes for the overall mean of April to be 80% of the mean for the first 15 days. (3)

1.2 The wind speed (in km per hour) and temperature (in °C) for a certain town were recorded at 16:00 for a period of 10 days. The information is shown in the table below.

WIND SPEED IN km/h (x)	2	6	15	20	25	17	11	24	13	22
TEMPERATURE IN °C (y)	28	26	22	22	16	20	24	19	26	19

1.2.1 Determine the equation of the least squares regression line for the data. (3)

1.2.2 Predict the temperature at 16:00 if, on a certain day, the wind speed of this town was 9 km per hour. (2)

1.2.3 Interpret the value of b in the context of the data. (1)
[14]

QUESTION 2

The number of days that Grade 8 learners were absent at a certain high school during a year was recorded. This information is represented in the table below.

NUMBER OF DAYS ABSENT	NUMBER OF LEARNERS
$0 \leq x < 5$	34
$5 \leq x < 10$	45
$10 \leq x < 15$	98
$15 \leq x < 20$	43
$20 \leq x < 25$	7
$25 \leq x < 30$	3

2.1 Write down the modal class for the data. (1)

2.2 How many learners were absent from school for less than 15 days? (1)

2.3 How many Grade 8 learners are at the school? (1)

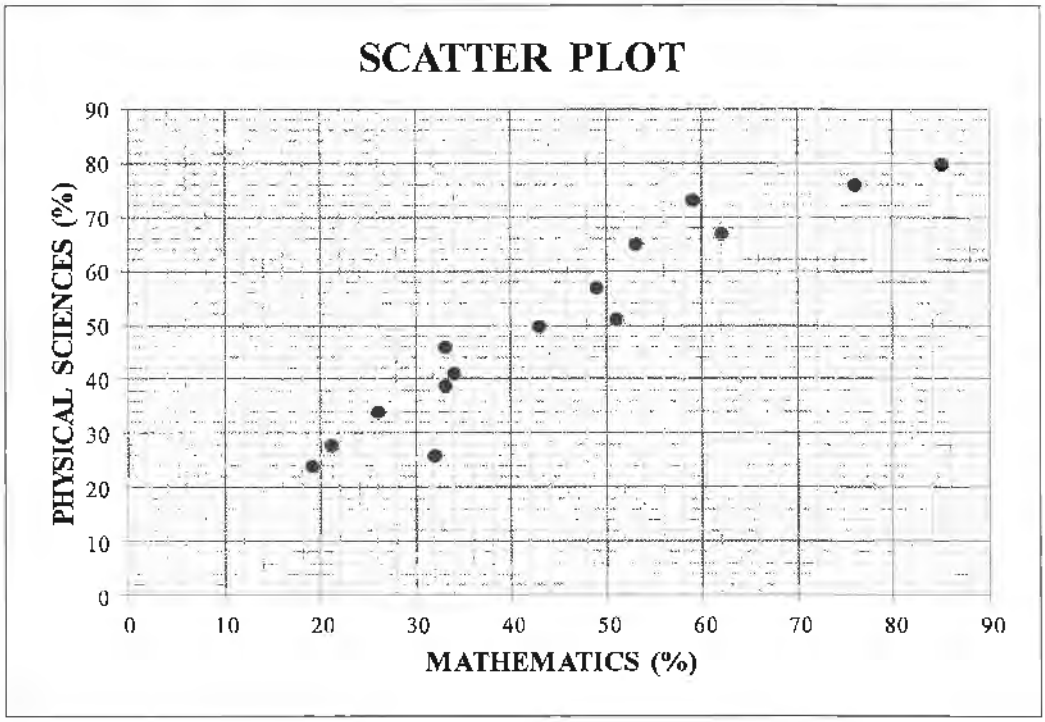
2.4 Draw a cumulative frequency graph (ogive) to represent the data above on the grid provided in the ANSWER BOOK. (4)

2.5 Use the cumulative frequency graph to determine the median number of days the Grade 8 learners were absent. (2)
[9]

QUESTION 1

A Mathematics teacher was curious to establish if her learners' Mathematics marks influenced their Physical Sciences marks. In the table below, the Mathematics and Physical Sciences marks of 15 learners in her class are given as percentages (%).

MATHEMATICS (AS %)	26	62	21	33	53	76	32	59	43	33	49	51	19	34	85
PHYSICAL SCIENCES (AS %)	34	67	28	46	65	76	26	73	50	39	57	51	24	41	80



1.1 Determine the equation of the least squares regression line for the data. (3)

1.2 Draw the least squares regression line on the scatter plot provided in the ANSWER BOOK. (2)

1.3 Predict the Physical Sciences mark of a learner who achieved 69% for Mathematics. (2)

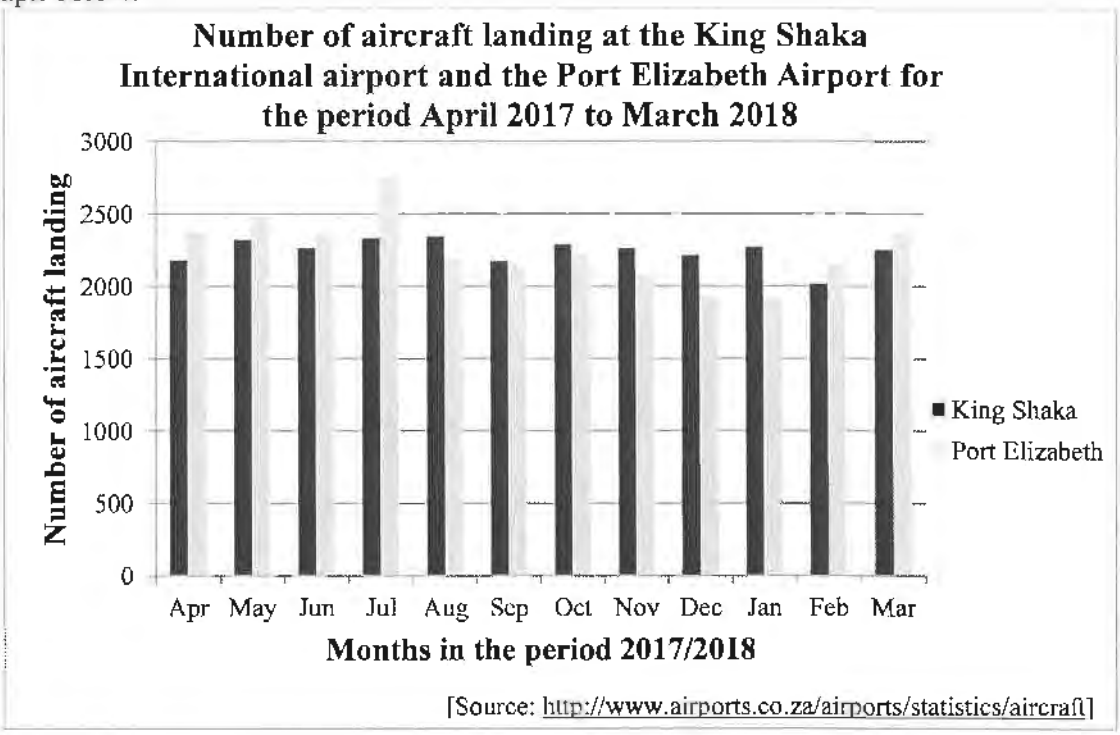
1.4 Write down the correlation coefficient between the Mathematics and Physical Sciences marks for the data. (1)

1.5 Comment on the strength of the correlation between the Mathematics and Physical Sciences marks for the data. (1)

1.6 What trend did the teacher observe between the results of the two subjects? (1)
[10]

QUESTION 2

The number of aircraft landing at the King Shaka International Airport and the Port Elizabeth Airport for the period starting in April 2017 and ending in March 2018, is shown in the double bar graph below.



[Source: http://www.airports.co.za/airports/statistics/aircraft]

2.1 The number of aircraft landing at the Port Elizabeth Airport exceeds the number of aircraft landing at the King Shaka International Airport during some months of the given period. During which month is this difference the greatest? (1)

2.2 The number of aircraft landing at the King Shaka International Airport during these months are:

2 182	2 323	2 267	2 334	2 346	2 175
2 293	2 263	2 215	2 271	2 018	2 254

Calculate the mean for the data. (2)

2.3 Calculate the standard deviation for the number of aircraft landing at the King Shaka International Airport for the given period. (2)

2.4 Determine the number of months in which the number of aircraft landing at the King Shaka International Airport were within one standard deviation of the mean. (3)

2.5 Which ONE of the following statements is CORRECT?

A. During December and January, there were more landings at the Port Elizabeth Airport than at the King Shaka International Airport.

B. There was a greater variation in the number of aircraft landing at the King Shaka International Airport than at the Port Elizabeth Airport for the given period.

C. The standard deviation of the number of landings at the Port Elizabeth Airport will be higher than the standard deviation of the number of landings at the King Shaka International Airport. (1)

[9]

QUESTION 1

The table below shows the monthly income (in rands) of 6 different people and the amount (in rands) that each person spends on the monthly repayment of a motor vehicle.

MONTHLY INCOME (IN RANDS)	9 000	13 500	15 000	16 500	17 000	20 000
MONTHLY REPAYMENT (IN RANDS)	2 000	3 000	3 500	5 200	5 500	6 000

1.1 Determine the equation of the least squares regression line for the data. (3)

1.2 If a person earns R14 000 per month, predict the monthly repayment that the person could make towards a motor vehicle. (2)

1.3 Determine the correlation coefficient between the monthly income and the monthly repayment of a motor vehicle. (1)

1.4 A person who earns R18 000 per month has to decide whether to spend R9 000 as a monthly repayment of a motor vehicle, or not. If the above information is a true representation of the population data, which of the following would the person most likely decide on:

A Spend R9 000 per month because there is a very strong positive correlation between the amount earned and the monthly repayment.

B NOT to spend R9 000 per month because there is a very weak positive correlation between the amount earned and the monthly repayment.

C Spend R9 000 per month because the point (18 000 ; 9 000) lies very near to the least squares regression line.

D NOT to spend R9 000 per month because the point (18 000 ; 9 000) lies very far from the least squares regression line. (2)

[8]

QUESTION 2

A survey was conducted among 100 people about the amount that they paid on a monthly basis for their cellphone contracts. The person carrying out the survey calculated the estimated mean to be R309 per month. Unfortunately, he lost some of the data thereafter. The partial results of the survey are shown in the frequency table below:

AMOUNT PAID (IN RANDS)	FREQUENCY
$0 < x \leq 100$	7
$100 < x \leq 200$	12
$200 < x \leq 300$	a
$300 < x \leq 400$	35
$400 < x \leq 500$	b
$500 < x \leq 600$	6

2.1 How many people paid R200 or less on their monthly cellphone contracts? (1)

2.2 Use the information above to show that $a = 24$ and $b = 16$. (5)

2.3 Write down the modal class for the data. (1)

2.4 On the grid provided in the ANSWER BOOK, draw an ogive (cumulative frequency graph) to represent the data. (4)

2.5 Determine how many people paid more than R420 per month for their cellphone contracts. (2)

[13]

QUESTION 1

Each child in a group of four-year-old children was given the same puzzle to complete. The time taken (in minutes) by each child to complete the puzzle is shown in the table below.

TIME TAKEN (t) (IN MINUTES)	**NUMBER OF CHILDREN**
$2 < t \leq 6$	2
$6 < t \leq 10$	10
$10 < t \leq 14$	9
$14 < t \leq 18$	7
$18 < t \leq 22$	8
$22 < t \leq 26$	7
$26 < t \leq 30$	2

1.1 How many children completed the puzzle? (1)

1.2 Calculate the estimated mean time taken to complete the puzzle. (2)

1.3 Complete the cumulative frequency column in the table given in the ANSWER BOOK. (2)

1.4 Draw a cumulative frequency graph (ogive) to represent the data on the grid provided in the ANSWER BOOK. (3)

1.5 Use the graph to determine the median time taken to complete the puzzle. (2)
[10]

QUESTION 2

Learners who scored a mark below 50% in a Mathematics test were selected to use a computer-based programme as part of an intervention strategy. On completing the programme, these learners wrote a second test to determine the effectiveness of the intervention strategy. The mark (as a percentage) scored by 15 of these learners in both tests is given in the table below.

LEARNER	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**	**L8**	**L9**	**L10**	**L11**	**L12**	**L13**	**L14**	**L15**
TEST 1 (%)	10	18	23	24	27	34	34	36	37	39	40	44	45	48	49
TEST 2 (%)	33	21	32	20	58	43	49	48	41	55	50	45	62	68	60

2.1 Determine the equation of the least squares regression line. (3)

2.2 A learner's mark in the first test was 15 out of a maximum of 50 marks.

2.2.1 Write down the learner's mark for this test as a percentage. (1)

2.2.2 Predict the learner's mark for the second test. Give your answer to the nearest integer. (2)

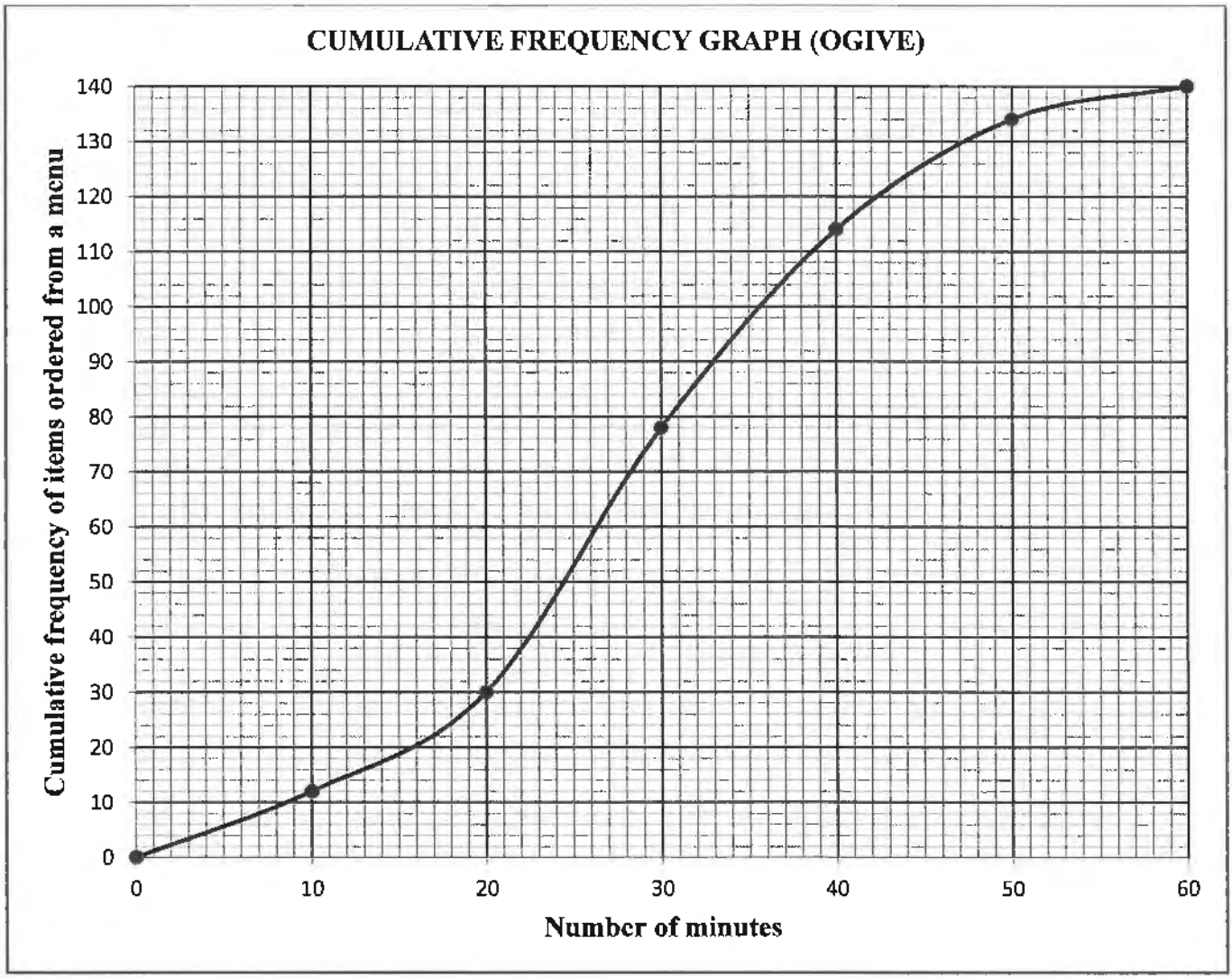
2.3 For the 15 learners above, the mean mark of the second test is 45,67% and the standard deviation is 13,88%. The teacher discovered that he forgot to add the marks of the last question to the total mark of each of these learners. All the learners scored full marks in the last question. When the marks of the last question are added, the new mean mark is 50,67%.

2.3.1 What is the standard deviation after the marks for the last question are added to each learner's total? (2)

2.3.2 What is the total mark of the last question? (2)
[10]

QUESTION 1

1.1 The cumulative frequency graph (ogive) drawn below shows the total number of food items ordered from a menu over a period of 1 hour.



1.1.1 Write down the total number of food items ordered from the menu during this hour. (1)

1.1.2 Write down the modal class of the data. (1)

1.1.3 How long did it take to order the first 30 food items? (1)

1.1.4 How many food items were ordered in the last 15 minutes? (2)

1.1.5 Determine the 75^{th} percentile for the data. (2)

1.1.6 Calculate the interquartile range of the data. (2)

1.2 Reggie works part-time as a waiter at a local restaurant. The amount of money (in rands) he made in tips over a 15-day period is given below.

35	70	75	80	80
90	100	100	105	105
110	110	115	120	125

1.2.1 Calculate:

(a) The mean of the data (2)

(b) The standard deviation of the data (2)

1.2.2 Mary also works part-time as a waitress at the same restaurant. Over the same 15-day period Mary collected the same mean amount in tips as Reggie, but her standard deviation was R14.

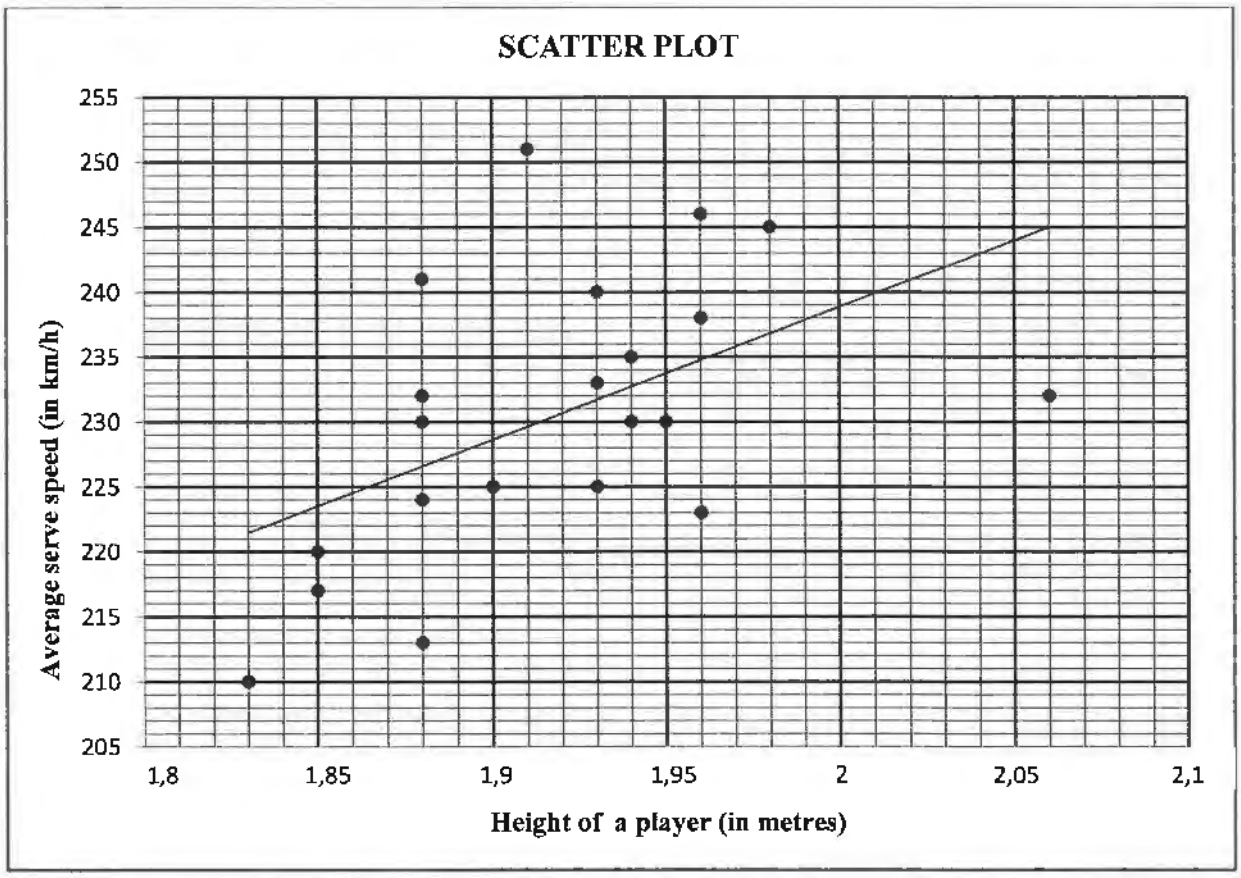
Using the available information, comment on the:

(a) Total amount in tips that they EACH collected over the 15-day period (1)

(b) Variation that EACH of them received in daily tips over this period (1)
[15]

QUESTION 2

A familiar question among professional tennis players is whether the speed of a tennis serve (in km/h) depends on the height of a player (in metres). The heights of 21 tennis players and the average speed of their serves were recorded during a tournament. The data is represented in the scatter plot below. The least squares regression line is also drawn.



2.1 Write down the fastest average serve speed (in km/h) achieved in this tournament. (1)

2.2 Consider the following correlation coefficients:

A. $r = 0,93$ B. $r = -0,42$ C. $r = 0,52$

2.2.1 Which ONE of the given correlation coefficients best fits the plotted data? (1)

2.2.2 Use the scatter plot and least squares regression line to motivate your answer to QUESTION 2.2.1. (1)

2.3 What does the data suggest about the speed of a tennis serve (in km/h) and the height of a player (in metres)? (1)

2.4 The equation of the regression line is given as $\hat{y} = 27,07 + bx$.
Explain why, in this context, the least squares regression line CANNOT intersect the y-axis at $(0 ; 27,07)$. (1)
[5]

QUESTION 1

An organisation decided that it would set up blood donor clinics at various colleges. Students would donate blood over a period of 10 days. The number of units of blood donated per day by students of college X is shown in the table below.

DAYS	1	2	3	4	5	6	7	8	9	10
UNITS OF BLOOD	45	59	65	73	79	82	91	99	101	106

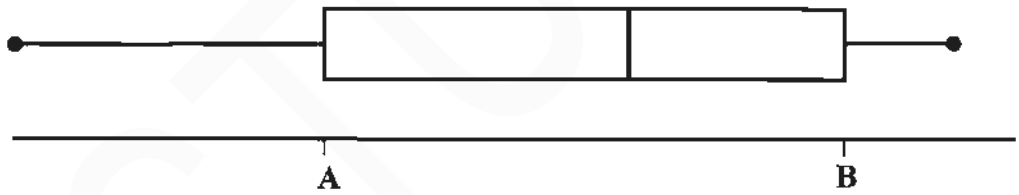
1.1 Calculate:

1.1.1 The mean of the units of blood donated per day over the period of 10 days (2)

1.1.2 The standard deviation of the data (2)

1.1.3 How many days is the number of units of blood donated at college X outside one standard deviation from the mean? (3)

1.2 The number of units of blood donated by the students of college X is represented in the box and whisker diagram below.



1.2.1 Describe the skewness of the data. (1)

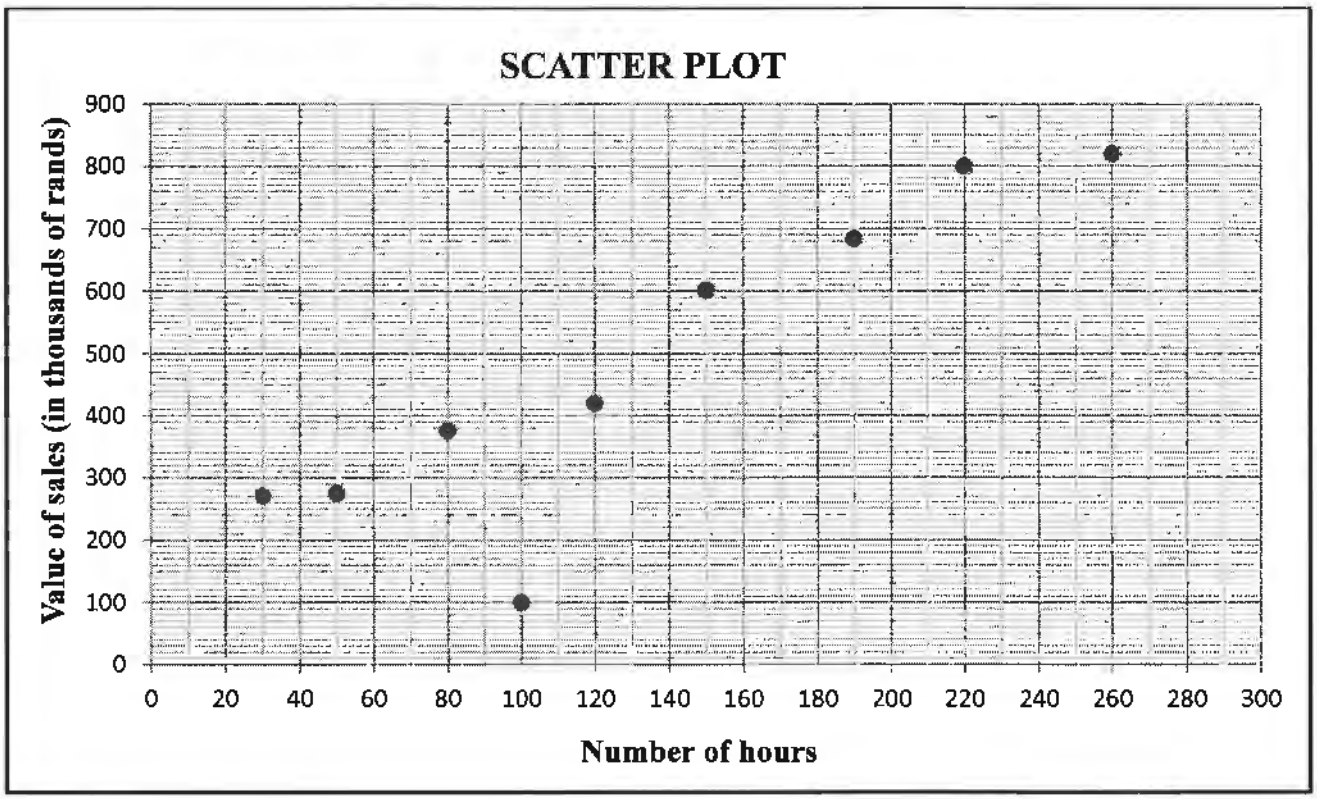
1.2.2 Write down the values of **A** and **B**, the lower quartile and the upper quartile respectively, of the data set. (2)

1.3 It was discovered that there was an error in counting the number of units of blood donated by college X each day. The correct mean of the data is 95 units of blood. How many units of blood were NOT counted over the ten days? (1)
[11]

QUESTION 2

The table below shows the number of hours that a sales representative of a company spent with each of his nine clients in one year and the value of the sales (in thousands of rands) for that client.

NUMBER OF HOURS	30	50	80	100	120	150	190	220	260
VALUE OF SALES (IN THOUSANDS OF RANDS)	270	275	376	100	420	602	684	800	820



2.1 Identify an outlier in the data above. (1)

2.2 Calculate the equation of the least squares regression line of the data. (3)

2.3 The sales representative forgot to record the sales of one of his clients. Predict the value of this client's sales (in thousands of rands) if he spent 240 hours with him during the year. (2)

2.4 What is the expected increase in sales for EACH additional hour spent with a client? (2)
[8]

QUESTION 1

The monthly profit (in thousands of rands) made by a company in a year is given in the table below.

110	112	156	164	167	169
171	176	192	228	278	360

1.1 Calculate the:

1.1.1 Mean profit for the year (3)

1.1.2 Median profit for the year (1)

1.2 On the number line provided in the ANSWER BOOK, draw a box and whisker diagram to represent the data. (2)

1.3 Hence, determine the interquartile range of the data. (1)

1.4 Comment on the skewness in the distribution of the data. (1)

1.5 For the given data:

1.5.1 Calculate the standard deviation (1)

1.5.2 Determine the number of months in which the profit was less than one standard deviation below the mean (2)
[11]

QUESTION 2

It is said that the number of times that a cricket chirps in a minute gives a very good indication of the air temperature (in °C). The table below shows the information recorded during an observation study.

CHIRPS PER MINUTE	AIR TEMPERATURE IN °C
32	8
40	10
52	12
76	15
92	17
112	20
128	25
180	28
184	30
200	35

2.1 Represent the data above on the grid provided in the ANSWER BOOK. (3)

2.2 Explain why the claim, 'gives a very good indication', is TRUE. (1)

2.3 Determine the equation of the least squares regression line of the data. (3)

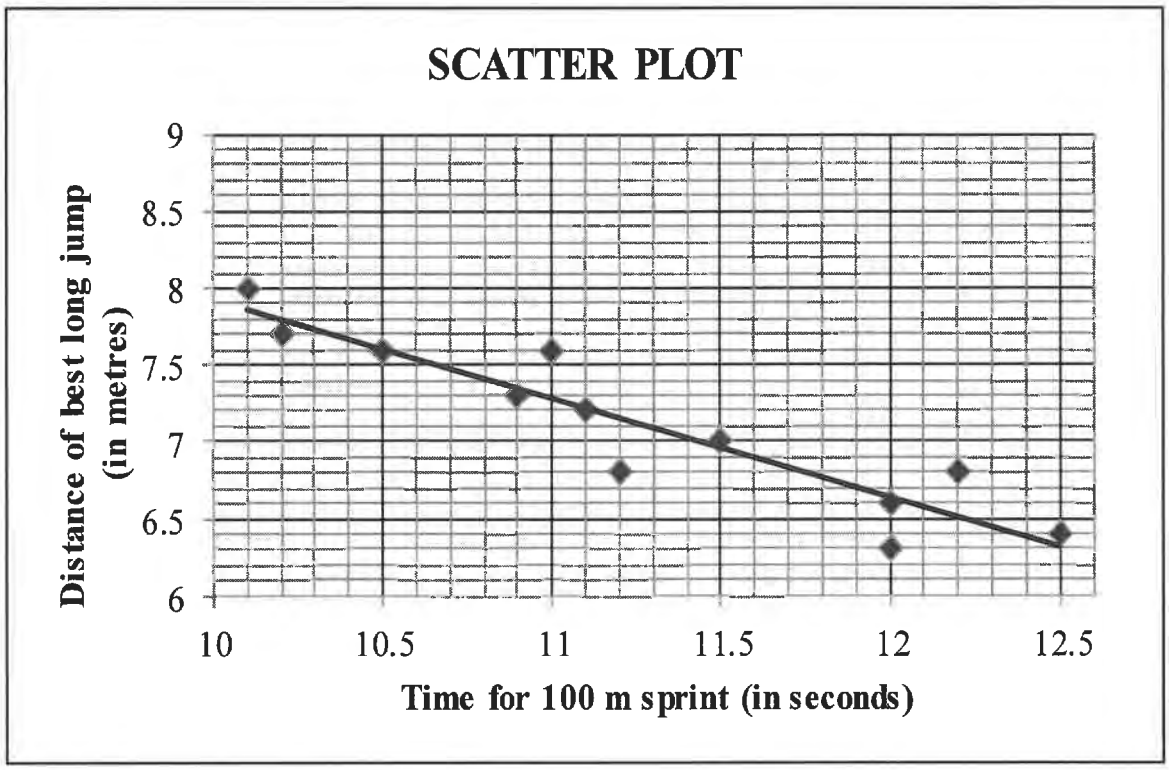
2.4 Predict the air temperature (in °C) if a cricket chirps 80 times a minute. (2)
[9]

QUESTION 1

The table below shows the time (in seconds, rounded to ONE decimal place) taken by 12 athletes to run the 100 metre sprint and the distance (in metres, rounded to ONE decimal place) of their best long jump.

Time for 100 m sprint (in seconds)	10,1	10,2	10,5	10,9	11	11,1	11,2	11,5	12	12	12,2	12,5
Distance of best long jump (in metres)	8	7,7	7,6	7,3	7,6	7,2	6,8	7	6,6	6,3	6,8	6,4

The scatter plot representing the data above is given below.



The equation of the least squares regression line is $\hat{y} = a + bx$.

1.1 Determine the values of a and b. (3)

1.2 An athlete runs the 100 metre sprint in 11,7 seconds. Use $\hat{y} = a + bx$ to predict the distance of the best long jump of this athlete. (2)

1.3 Another athlete completes the 100 metre sprint in 12,3 seconds and the distance of his best long jump is 7,6 metres. If this is included in the data, will the gradient of the least squares regression line increase or decrease? Motivate your answer without any further calculations. (2)
[7]

QUESTION 2

In an experiment, a group of 23 girls were presented with a page containing 30 coloured rectangles. They were asked to name the colours of the rectangles correctly as quickly as possible. The time, in seconds, taken by each of the girls is given in the table below.

12	13	13	14	14	16	17	18	18	18	19	20
21	21	22	22	23	24	25	27	29	30	36	

2.1 Calculate:

2.1.1 The mean of the data (2)

2.1.2 The interquartile range of the data (3)

2.2 The standard deviation of the times taken by the girls is 5,94. How many girls took longer than ONE standard deviation from the mean to name the colours? (2)

2.3 Draw a box and whisker diagram to represent the data on the number line provided in the ANSWER BOOK. (3)

2.4 The five-number summary of the times taken by a group of 23 boys in naming the colours of the rectangles correctly is (15 ; 21 ; 23,5 ; 26 ; 38).

2.4.1 Which of the two groups, girls or boys, had the lower median time to correctly name the colours of the rectangles? (1)

2.4.2 The first three learners who named the colours of all 30 rectangles correctly in the shortest time will receive a prize. How many boys will be among these three prizewinners? Motivate your answer. (2)

[13]

QUESTION 1

An IT company writes programs for apps. The time taken (in hours) to write the programs and the cost (in thousands of rands) are shown in the table below.

TIME TAKEN (IN HOURS)	5	7	5	8	10	13	15	20	18	25	23
COST (IN THOUSANDS OF RANDS)	10	10	15	12	20	25	28	32	28	40	30

1.1 Determine the equation of the least squares regression line. (3)

1.2 Use the equation of the least squares regression line to predict the cost, in rands, of an app that will take 16 hours to write. (2)

1.3 Calculate the correlation coefficient of the data. (1)

1.4 For each app that the company writes, there is a cost that is independent of the number of hours spent on writing the app. Calculate this cost (in rands). (2)
[8]

QUESTION 2

The commission earned, in thousands of rands, by the salesmen of a particular company in a certain month is shown in the table below.

COMMISSION EARNED (IN THOUSANDS OF RANDS)	**FREQUENCY**
$20 < x \leq 40$	7
$40 < x \leq 60$	6
$60 < x \leq 80$	8
$80 < x \leq 100$	10
$100 < x \leq 120$	4

2.1 Write down the modal class of the data. (1)

2.2 Complete the cumulative frequency column in the table given in the ANSWER BOOK. (2)

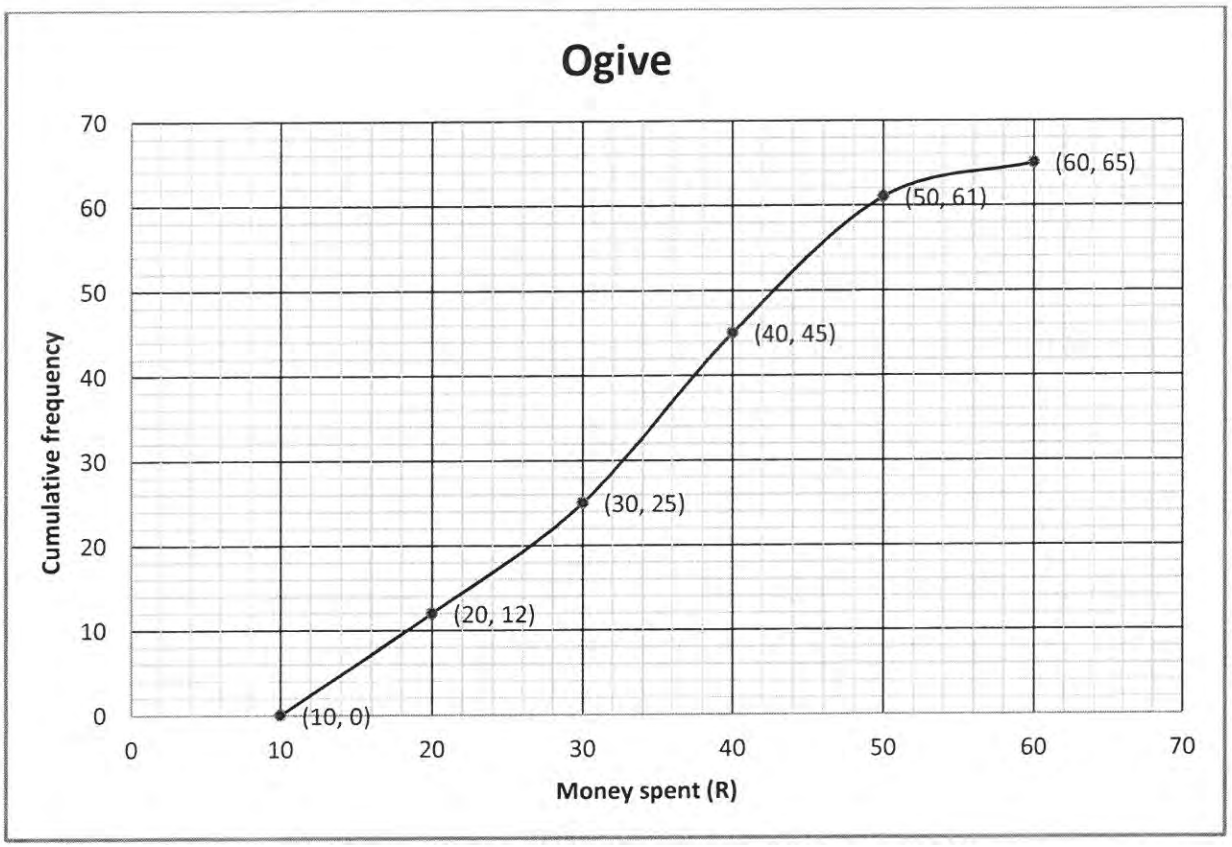
2.3 Draw an ogive (cumulative frequency curve) to represent the data on the grid provided in the ANSWER BOOK. (4)

2.4 A salesman receives a bonus if his commission is more than R90 000 for the month. Calculate how many of the salesmen received bonuses for this month. (2)

2.5 Determine the approximate mean commission earned by the salesmen in this month correct to the nearest thousand rand. (3)
[12]

QUESTION 1

The amount of money, in rands, that learners spent while visiting a tuck shop at school on a specific day was recorded. The data is represented in the ogive below.



An incomplete frequency table is also given for the data.

Amount of money (in R)	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$	$50 \leq x < 60$
Frequency	a	13	20	b	4

1.1 How many learners visited the tuck shop on that day? (1)

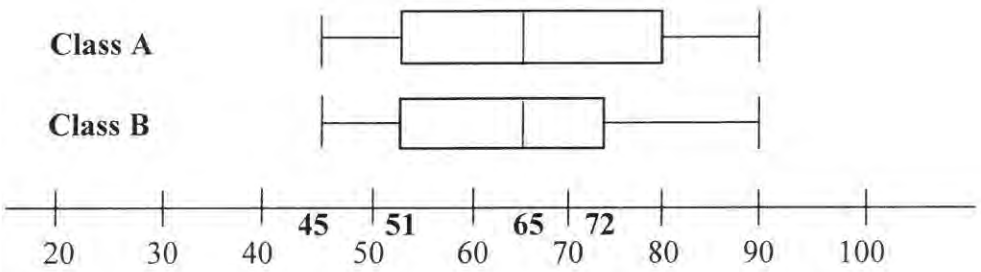
1.2 Write down the modal class of this data. (1)

1.3 Determine the values of a and b in the frequency table. (2)

1.4 Use the ogive to estimate the number of learners who spent at least R45 on the day the data was recorded at the tuck shop. (2)
[6]

QUESTION 2

2.1 Mrs Smith has two classes, each having 30 learners. Their final marks (out of 100) for the year are represented in the box and whisker diagram below.



2.1.1 Determine the interquartile range of Class B. (2)

2.1.2 Explain the significance in the difference of the length of the boxes in the diagram. (2)

2.1.3 Mrs Smith studied the results and made the comment that there was no significant difference in the performance of the two classes. Give TWO reasons you think Mrs Smith will use to prove her statement. (2)

2.2 Eight couples entered a dance competition. Their performances were scored by two judges. The scores (out of 20) are given 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

2.2.1 Determine the equation of the least squares regression line of the scores given by the two judges. (3)

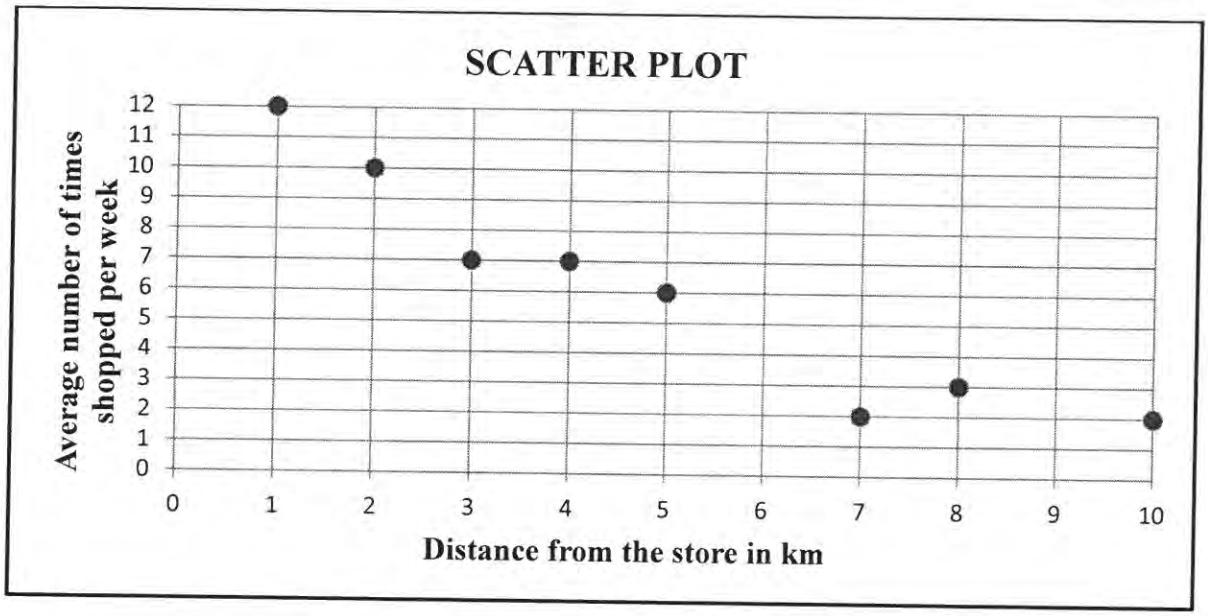
2.2.2 A ninth couple entered late for the competition and received a score of 15 from JUDGE 1. Estimate the score that might have been assigned by JUDGE 2 to the nearest integral value. (2)

2.2.3 Are the judges consistent in assigning scores to the performance of the couples? Prove your answer and support it with relevant statistics. (2)
[13]

QUESTION 1

A survey was conducted at a local supermarket relating the distance that shoppers lived from the store to the average number of times they shopped at the store in a week. The results are shown in the table below.

Distance from the store in km	1	2	3	4	5	7	8	10
Average number of times shopped per week	12	10	7	7	6	2	3	2



1.1 Use the scatter plot to comment on the strength of the relationship between the distance a shopper lived from the store and the average number of times she/he shopped at the store in a week. (1)

1.2 Calculate the correlation coefficient of the data. (1)

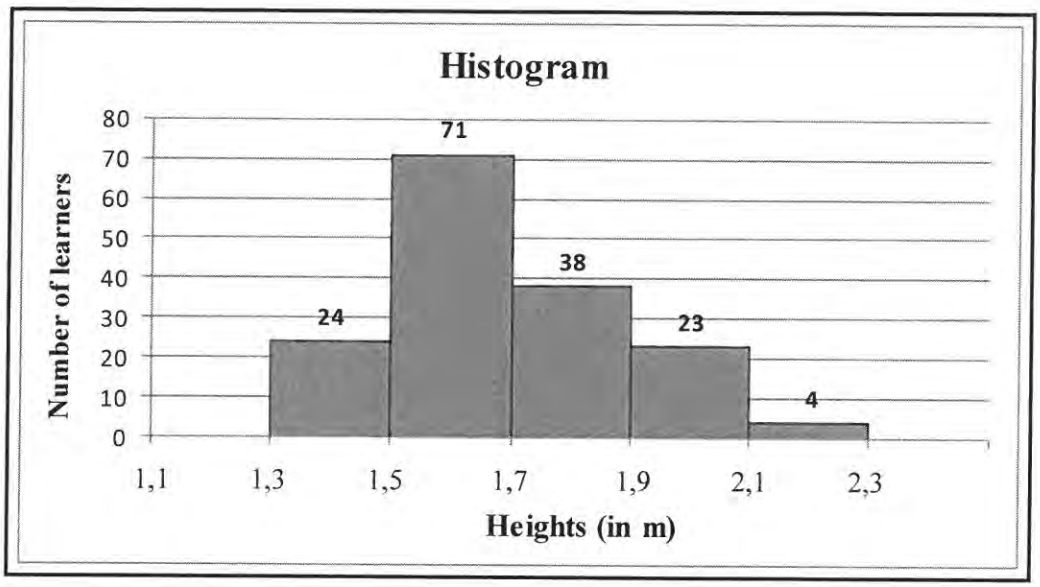
1.3 Calculate the equation of the least squares regression line of the data. (3)

1.4 Use your answer at QUESTION 1.3 to estimate the average number of times that a shopper living 6 km from the supermarket will visit the store in a week. (2)

1.5 Sketch the least squares regression line on the scatter plot provided in the ANSWER BOOK. (2)
[9]

QUESTION 2

The heights of 160 learners in a school are measured. The height of the shortest learner is 1,39 m and the height of the tallest learner is 2,21 m. The heights are represented in the histogram below.



2.1 Describe the skewness of the data. (1)

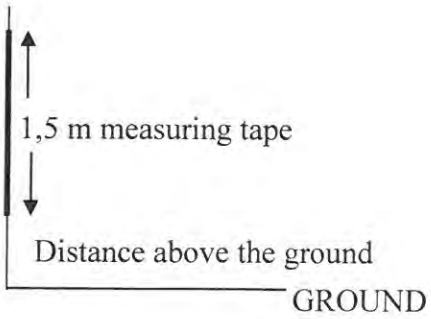
2.2 Calculate the range of the heights. (2)

2.3 Complete the cumulative frequency column in the table given in the ANSWER BOOK. (2)

2.4 Draw an ogive (cumulative frequency curve) to represent the data on the grid provided in the ANSWER BOOK. (4)

2.5 Eighty learners are less than x metres in height. Estimate x. (2)

2.6 The person taking the measurements only had a 1,5 m measuring tape available. In order to compensate for the short measuring tape, he decided to mount the tape on a wall at a height of 1 m above the ground. After recording the measurements he discovered that the tape was mounted at 1,1 m above the ground instead of 1 m.



How does this error influence the following:

2.6.1 Mean of the data set (1)

2.6.2 Standard deviation of the data set (1)

[13]

QUESTION 1

On a certain day a tour operator sent 11 tour buses to 11 different destinations. The table below shows the number of passengers on each bus.

8	8	10	12	16	19	20	21	24	25	26

1.1 Calculate the mean number of passengers travelling in a tour bus. (2)

1.2 Write down the five-number summary of the data. (3)

1.3 Draw a box and whisker diagram for the data. Use the number line provided in the ANSWER BOOK. (2)

1.4 Refer to the box and whisker diagram and comment on the skewness of the data set. (1)

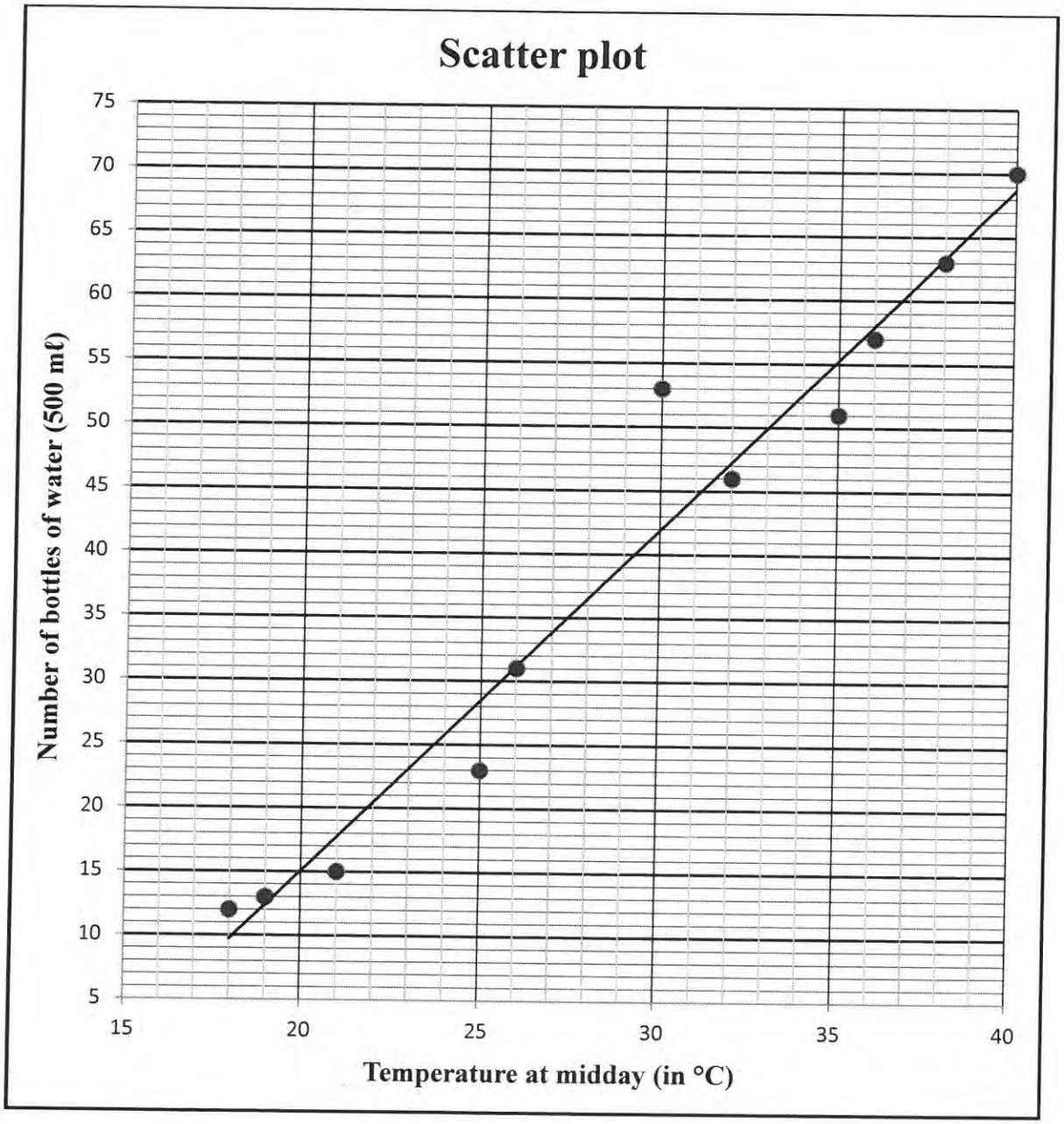
1.5 Calculate the standard deviation for this data set. (2)

1.6 A tour is regarded as popular if the number of passengers on a tour bus is one standard deviation above the mean. How many destinations were popular on this particular day? (2)
[12]

QUESTION 2

On the first school day of each month information is recorded about the temperature at midday (in °C) and the number of 500 mℓ bottles of water that were sold at the tuck shop of a certain school during the lunch break. The data is shown in the table below and represented on the scatter plot. The least squares regression line for this data is drawn on the scatter plot.

Temperature at midday (in °C)	18	21	19	26	32	35	36	40	38	30	25
Number of bottles of water (500 mℓ)	12	15	13	31	46	51	57	70	63	53	23



2.1 Identify an outlier in the data. (1)

2.2 Determine the equation of the least squares regression line. (3)

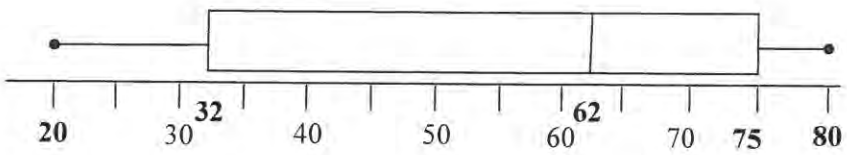
2.3 Estimate the number of 500 mℓ bottles of water that will be sold if the temperature is 28 °C at midday. (2)

2.4 Refer to the scatter plot. Would you say that the relation between the temperature at midday and the number of 500 mℓ bottles of water sold is weak or strong? Motivate your answer. (2)

2.5 Give a reason why the observed trend for this data cannot continue indefinitely. (1)
[9]

QUESTION 1

The box and whisker diagram below shows the marks (out of 80) obtained in a History test by a class of nine learners.



1.1 Comment on the skewness of the data. (1)

1.2 Write down the range of the marks obtained. (2)

1.3 If the learners had to obtain 32 marks to pass the test, estimate the percentage of the class that failed the test. (2)

1.4 In ascending order, the second mark is 28, the third mark 36 and the sixth mark 69. The seventh and eighth marks are the same. The average mark for this test is 54.

	28	36			69			

Fill in the marks of the remaining learners in ascending order. (6)
[11]

QUESTION 2

A company recorded the number of messages sent by e-mail over a period of 60 working days. The data is shown in the table below.

NUMBER OF MESSAGES	NUMBER OF DAYS
$10 < x \leq 20$	2
$20 < x \leq 30$	8
$30 < x \leq 40$	5
$40 < x \leq 50$	10
$50 < x \leq 60$	12
$60 < x \leq 70$	18
$70 < x \leq 80$	3
$80 < x \leq 90$	2

2.1 Estimate the mean number of messages sent per day, rounded off to TWO decimal places. (3)

2.2 Draw a cumulative frequency graph (ogive) of the data on the grid provided in the ANSWER BOOK. (4)

2.3 Hence, estimate the number of days on which 65 or more messages were sent. (2)
[9]

QUESTION 1

The table below shows the total fat (in grams, rounded off to the nearest whole number) and energy (in kilojoules, rounded off to the nearest 100) of 10 items that are sold at a fast-food restaurant.

Fat (in grams)	9	14	25	8	12	31	28	14	29	20
Energy (in kilojoules)	1 100	1 300	2 100	300	1 200	2 400	2 200	1 400	2 600	1 600

1.1 Represent the information above in a scatter plot on the grid provided in the ANSWER BOOK. (3)

1.2 The equation of the least squares regression line is $\hat{y} = 154{,}60 + 77{,}13x$.

1.2.1 An item at the restaurant contains 18 grams of fat. Calculate the number of kilojoules of energy that this item will provide. Give your answer rounded off to the nearest 100 kJ. (2)

1.2.2 Draw the least squares regression line on the scatter plot drawn for QUESTION 1.1. (2)

1.3 Identify an outlier in the data set. (1)

1.4 Calculate the value of the correlation coefficient. (2)

1.5 Comment on the strength of the relationship between the fat content and the number of kilojoules of energy. (1)
[11]

QUESTION 2

A group of 30 learners each randomly rolled two dice once and the sum of the values on the uppermost faces of the dice was recorded. The data is shown in the frequency table below.

Sum of the values on uppermost faces	Frequency
2	0
3	3
4	2
5	4
6	4
7	8
8	3
9	2
10	2
11	1
12	1

2.1 Calculate the mean of the data. (2)

2.2 Determine the median of the data. (2)

2.3 Determine the standard deviation of the data. (2)

2.4 Determine the number of times that the sum of the recorded values of the dice is within ONE standard deviation from the mean. Show your calculations. (3)
[9]

QUESTION 1

The table below shows the distances (in kilometres) travelled daily by a sales representative for 21 working days in a certain month.

131	132	140	140	141	144	146
147	149	150	151	159	167	169
169	172	174	175	178	187	189

1.1 Calculate the mean distance travelled by the sales representative. (2)

1.2 Write down the five-number summary for this set of data. (4)

1.3 Use the scaled line on DIAGRAM SHEET 1 to draw a box-and-whisker diagram for this set of data. (2)

1.4 Comment on the skewness of the data. (1)

1.5 Calculate the standard deviation of the distance travelled. (2)

1.6 The sales representative discovered that his odometer was faulty. The actual reading on each of the 21 days was p km more than that which was indicated. Write down, in terms of p (if applicable), the:

1.6.1 Actual mean (1)

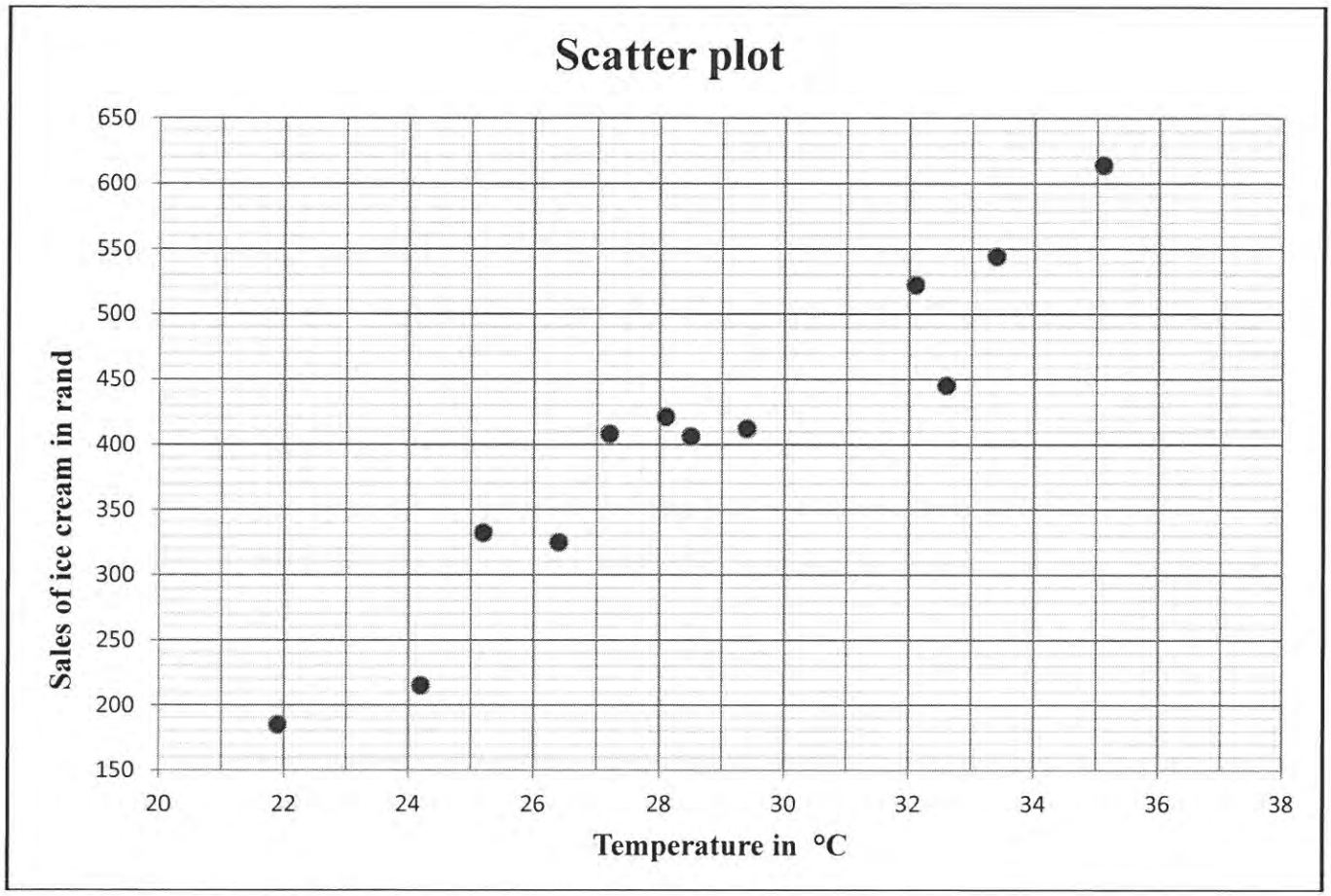
1.6.2 Actual standard deviation (1)

[13]

QUESTION 2

An ice-cream shop recorded the sales of ice cream, in rand, and the maximum temperature, in °C, for 12 days in a certain month. The data that they collected is represented in the table and scatter plot below.

Temperature in °C	24,2	26,4	21,9	25,2	28,5	32,1	29,4	35,1	33,4	28,1	32,6	27,2
Sales of ice cream in rand	215	325	185	332	406	522	412	614	544	421	445	408



2.1 Describe the influence of temperature on the sales of ice cream in the scatter plot. (1)

2.2 Give a reason why this trend cannot continue indefinitely. (1)

2.3 Calculate an equation for the least squares regression line (line of best fit). (4)

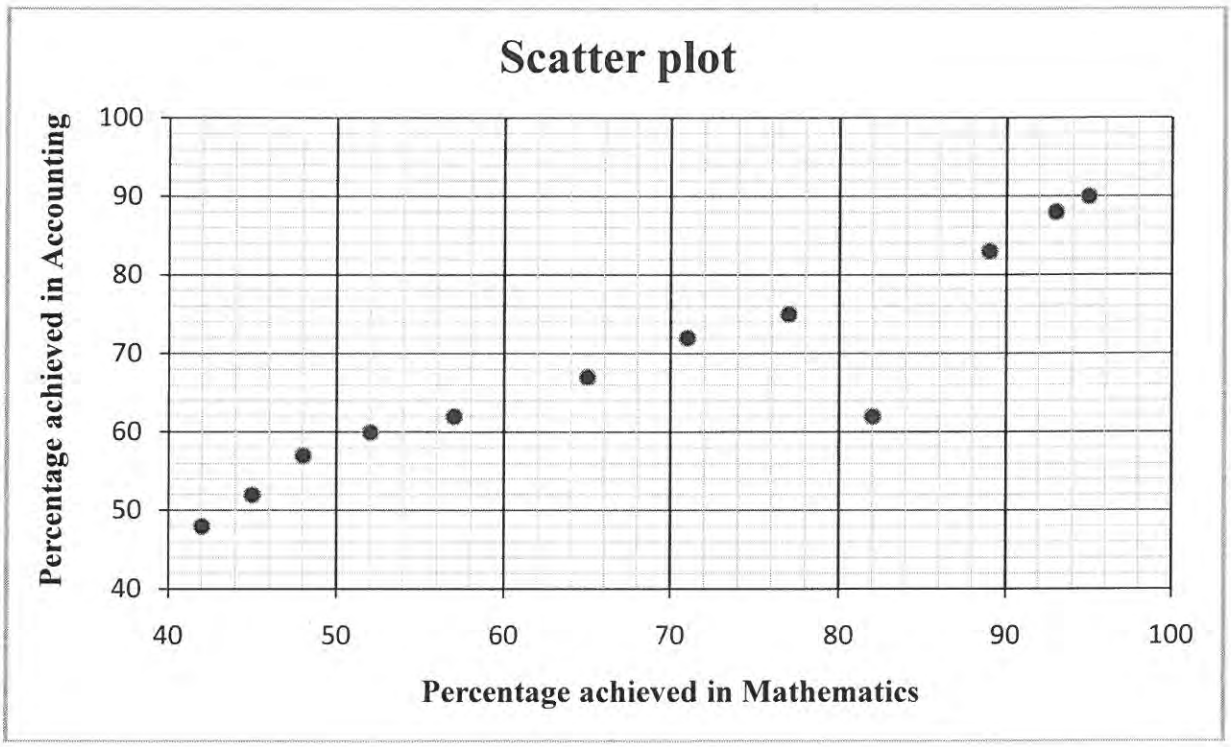
2.4 Calculate the correlation coefficient. (1)

2.5 Comment on the strength of the relationship between the variables. (1)
[8]

QUESTION 1

At a certain school, only 12 candidates take Mathematics and Accounting. The marks, as a percentage, scored by these candidates in the preparatory examinations for Mathematics and Accounting, are shown in the table and scatter plot below.

Mathematics	52	82	93	95	71	65	77	42	89	48	45	57
Accounting	60	62	88	90	72	67	75	48	83	57	52	62



1.1 Calculate the mean percentage of the Mathematics data. (2)

1.2 Calculate the standard deviation of the Mathematics data. (1)

1.3 Determine the number of candidates whose percentages in Mathematics lie within ONE standard deviation of the mean. (3)

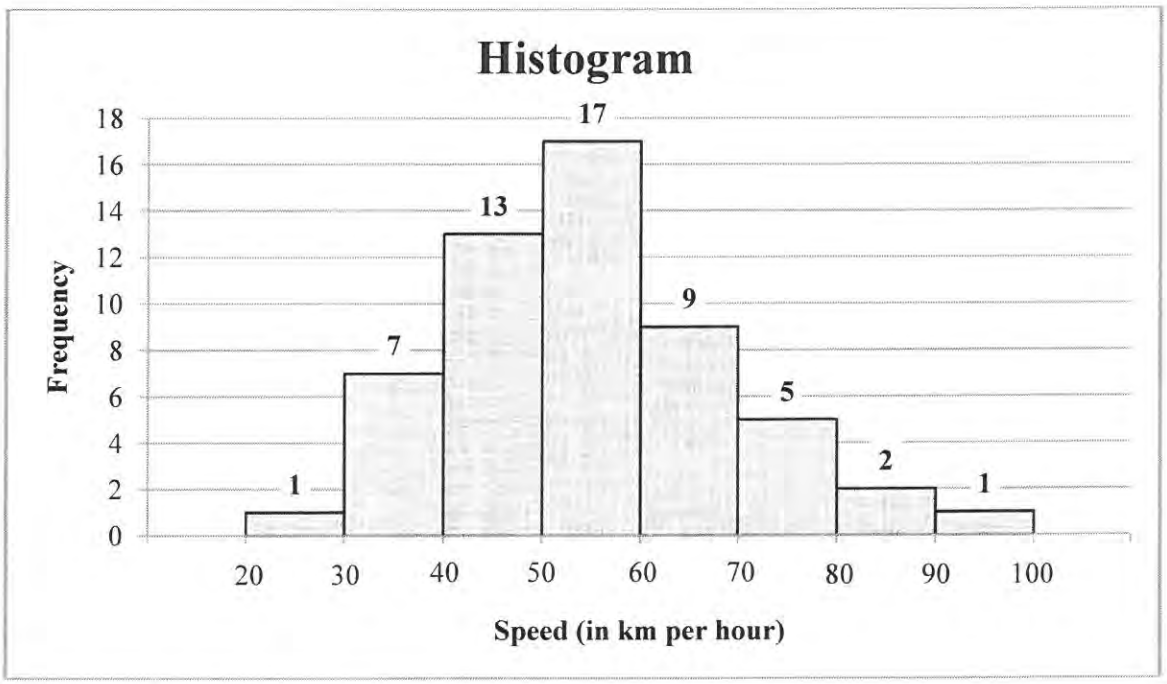
1.4 Calculate an equation for the least squares regression line (line of best fit) for the data. (3)

1.5 If a candidate from this group scored 60% in the Mathematics examination but was absent for the Accounting examination, predict the percentage that this candidate would have scored in the Accounting examination, using your equation in QUESTION 1.4. (Round off your answer to the NEAREST INTEGER.) (2)

1.6 Use the scatter plot and identify any outlier(s) in the data. (1)
[12]

QUESTION 2

The speeds of 55 cars passing through a certain section of a road are monitored for one hour. The speed limit on this section of road is 60 km per hour. A histogram is drawn to represent this data.



2.1 Identify the modal class of the data. (1)

2.2 Use the histogram to:

2.2.1 Complete the cumulative frequency column in the table on DIAGRAM SHEET 1 (2)

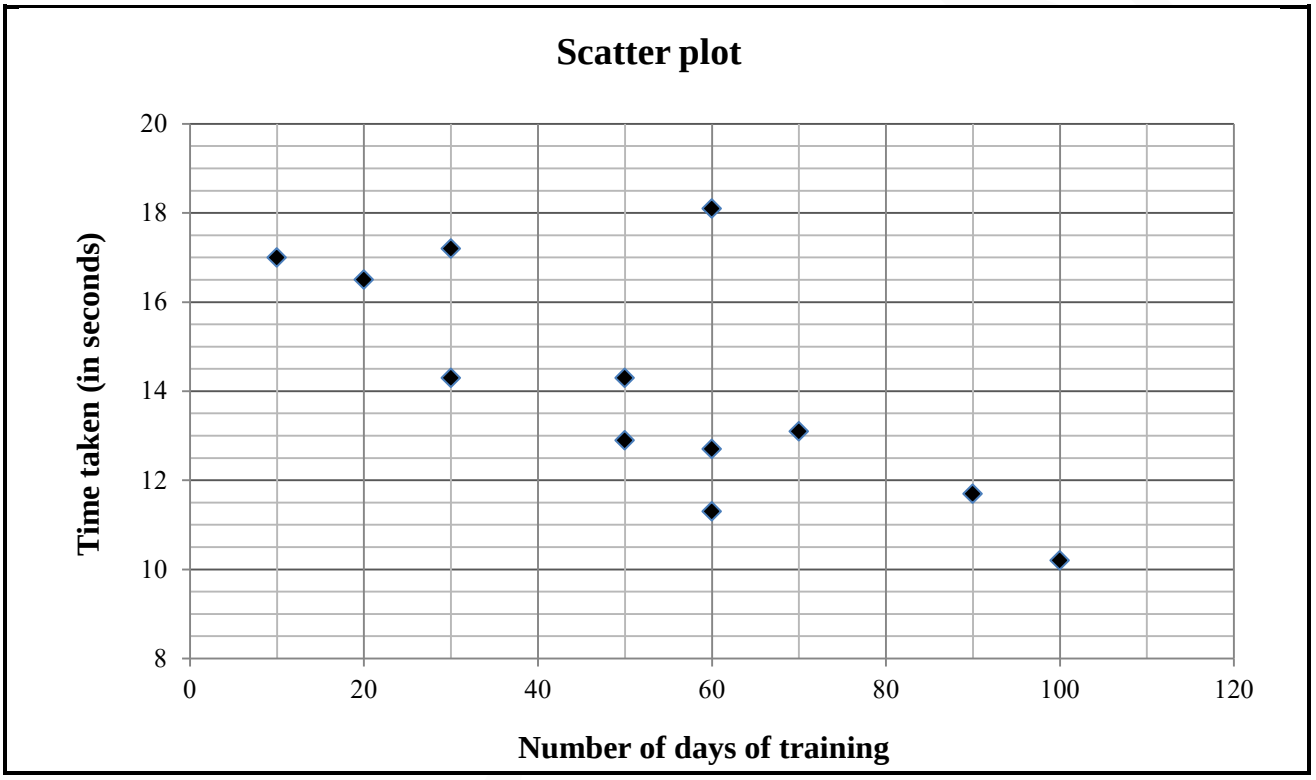
2.2.2 Draw an ogive (cumulative frequency graph) of the above data on the grid on DIAGRAM SHEET 1 (3)

2.3 The traffic department sends speeding fines to all motorists whose speed exceeds 66 km per hour. Estimate the number of motorists who will receive a speeding fine. (2)
[8]

QUESTION 1

Twelve athletes trained to run the 100 m sprint event at the local athletics club trials. Some of them took their training more seriously than others. The following table and scatter plot shows the number of days that an athlete trained and the time taken to run the event. The time taken, in seconds, is rounded to one decimal place.

Number of days of training	50	70	10	60	60	20	50	90	100	60	30	30
Time taken (in seconds)	12,9	13,1	17,0	11,3	18,1	16,5	14,3	11,7	10,2	12,7	17,2	14,3



1.1 Discuss the trend of the data collected. (1)

1.2 Identify any outlier(s) in the data. (1)

1.3 Calculate the equation of the least squares regression line. (4)

1.4 Predict the time taken to run the 100 m sprint for an athlete training for 45 days. (2)

1.5 Calculate the correlation coefficient. (2)

1.6 Comment on the strength of the relationship between the variables. (1)
[11]

QUESTION 2

The table below shows the amount of time (in hours) that learners aged between 14 and 18 spent watching television during 3 weeks of the holiday.

Time (hours)	Cumulative frequency
$0 \leq t < 20$	25
$20 \leq t < 40$	69
$40 \leq t < 60$	129
$60 \leq t < 80$	157
$80 \leq t < 100$	166
$100 \leq t < 120$	172

2.1 Draw an ogive (cumulative frequency curve) on DIAGRAM SHEET 1 to represent the above data. (3)

2.2 Write down the modal class of the data. (1)

2.3 Use the ogive (cumulative frequency curve) to estimate the number of learners who watched television more than 80% of the time. (2)

2.4 Estimate the mean time (in hours) that learners spent watching television during 3 weeks of the holiday. (4)
[10]