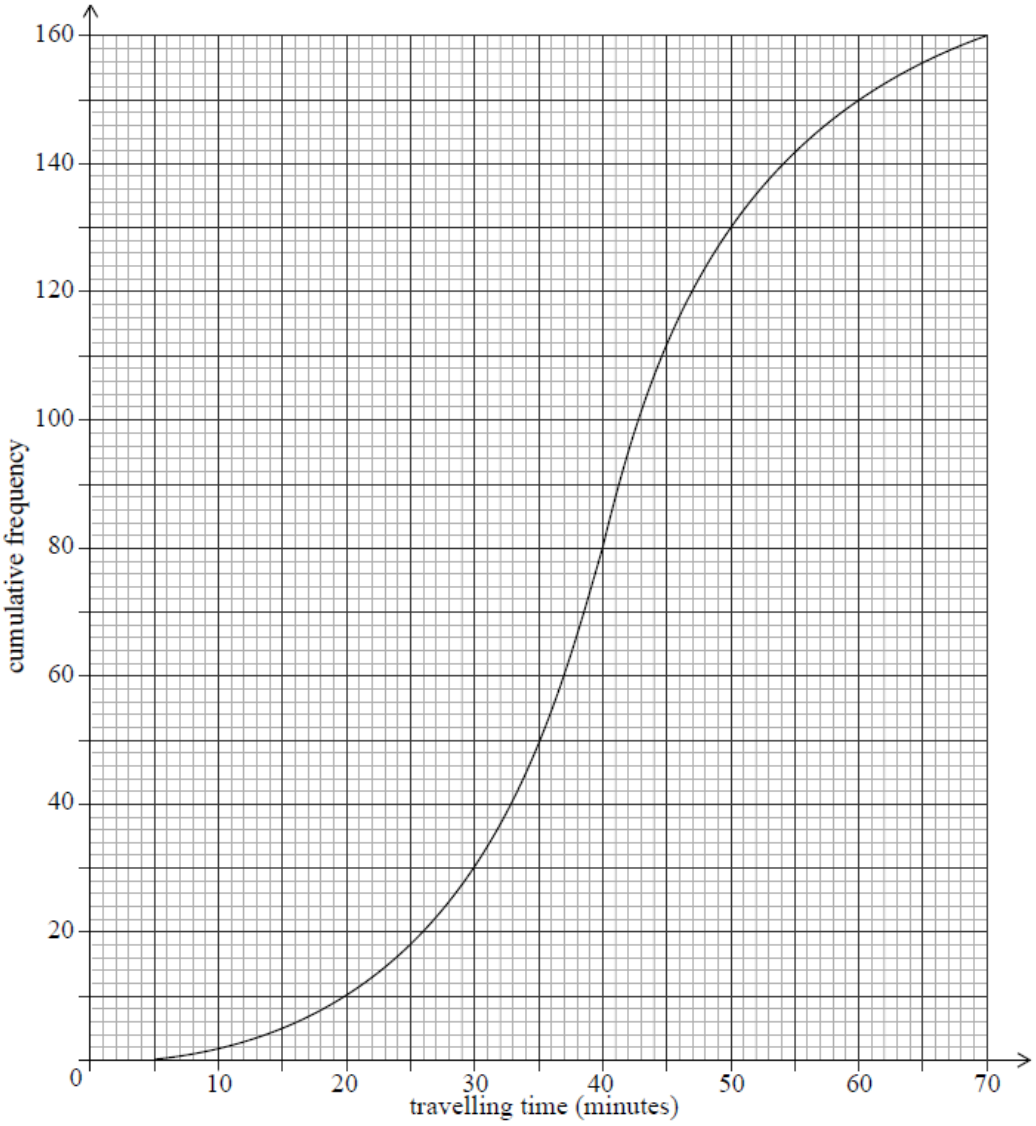
Statistics Test [63 marks]

1. [Maximum mark: 15] SPM.1.SL.TZ0.7

A large company surveyed 160 of its employees to find out how much time they spend traveling to work on a given day. The results of the survey are shown in the following cumulative frequency diagram.



(a) Find the median number of minutes spent traveling to work. [2]

Markscheme
evidence of median position ***(M1)***

80th employee

40 minutes ***A1***

[2 marks]

(b) Find the number of employees whose travelling time is within 15 minutes of the median. [3]

Markscheme

valid attempt to find interval (25–55) ***(M1)***

18 (employees), 142 (employees) ***A1***

124 ***A1***

[3 marks]

Only 10% of the employees spent more than k minutes traveling to work.

(c) Find the value of k. [3]

Markscheme

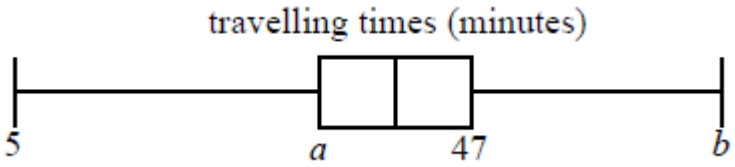
recognising that there are 16 employees in the top 10% ***(M1)***

144 employees travelled more than k minutes ***(A1)***

$k = 56$ ***A1***

[3 marks]

The results of the survey can also be displayed on the following box-and-whisker diagram.



(d) Write down the value of b. [1]

Markscheme

$b = 70$ ***A1***

[1 mark]

(e.i) Find the value of a. [2]

Markscheme

recognizing a is first quartile value ***(M1)***

40 employees

$a = 33$ ***A1***

[2 marks]

(e.ii) Hence, find the interquartile range. [2]

Markscheme

$47 - 33$ ***(M1)***

IQR = 14 ***A1***

[2 marks]

(f) Travelling times of less than p minutes are considered outliers.

Find the value of p.

[2]

Markscheme

attempt to find 1.5 × **their** IQR *(M1)*

33 – 21

12 *(A1)*

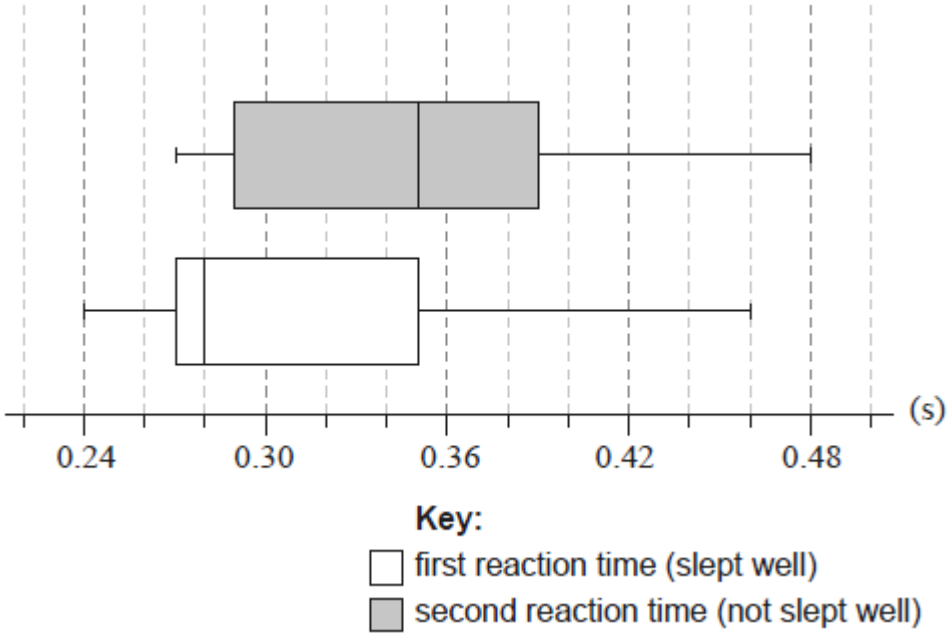
[2 marks]

2. [Maximum mark: 6] 22M.2.SL.TZ2.5

A random sample of nine adults were selected to see whether sleeping well affected their reaction times to a visual stimulus. Each adult's reaction time was measured twice.

The first measurement for reaction time was taken on a morning after the adult had slept well. The second measurement was taken on a morning after the same adult had not slept well.

The box and whisker diagrams for the reaction times, measured in seconds, are shown below.



Consider the box and whisker diagram representing the reaction times after sleeping well.

(a) State the median reaction time after sleeping well. [1]

Markscheme

0.28 (s) **A1**

[1 mark]

(b) Verify that the measurement of 0.46 seconds is not an outlier. [3]

Markscheme

$\text{IQR} = 0.35 - 0.27\ (= 0.08)$ (s) ***(A1)***

substituting **their** IQR into correct expression for upper fence ***(A1)***

$0.35 + 1.5 \times 0.08\ (= 0.47)$ (s)

$0.46 < 0.47$ ***R1***

so 0.46 (s) is not an outlier ***AG***

[3 marks]

(c) State why it appears that the mean reaction time is greater than the median reaction time. [1]

Markscheme

EITHER

the median is closer to the lower quartile (positively skewed) ***R1***

OR

The distribution is positively skewed ***R1***

OR

the range of reaction times below the median is smaller than the range of reaction times above the median ***R1***

Note: These are sample answers from a range of acceptable correct answers. Award ***R1*** for any correct statement that explains this.
Do not award ***R1*** if there is also an incorrect statement, even if another statement in the answer is correct. Accept a correctly and clearly labelled diagram.

[1 mark]

(d) Now consider the two box and whisker diagrams.

Comment on whether these box and whisker diagrams provide any evidence that might suggest that not sleeping well causes an increase in reaction time. [1]

Markscheme

EITHER

the distribution for 'not sleeping well' is centred at a higher reaction time ***R1***

OR

The median reaction time after not sleeping well is equal to the upper quartile reaction time after sleeping well ***R1***

OR

75% of reaction times are < 0.35 seconds after sleeping well, compared with 50% after not sleeping well *R1*

OR

the sample size of 9 is too small to draw any conclusions *R1*

Note: These are sample answers from a range of acceptable correct answers. Accept any relevant correct statement **that relates to the median and/or quartiles shown in the box plots**. **Do not accept** a comparison of means. Do not award *R1* if there is also an incorrect statement, even if another statement in the answer is correct.

Award *R0* to "correlation does not imply causation".

[1 mark]

3. [Maximum mark: 6] 20N.1.SL.TZ0.T_3

Hafizah harvested 49 mangoes from her farm. The weights of the mangoes, w, in grams, are shown in the following grouped frequency table.

Weight (g)	$100 \leq w < 200$	$200 \leq w < 300$	$300 \leq w < 400$	$400 \leq w < 500$	$500 \leq w < 600$
Frequency	4	7	14	16	8

(a) Write down the modal group for these data. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure. It appeared in a paper that permitted the use of a calculator, and so might not be suitable for all forms of practice.

$400 \leq w < 500$ ***(A1) (C1)***

Note: Accept alternative notation $[400,\ 500)$ or $[400,\ 500[$.
Do not accept "$400\text{-}500$".

[1 mark]

(b) Use your graphic display calculator to find an estimate of the standard deviation of the weights of mangoes from this harvest. [2]

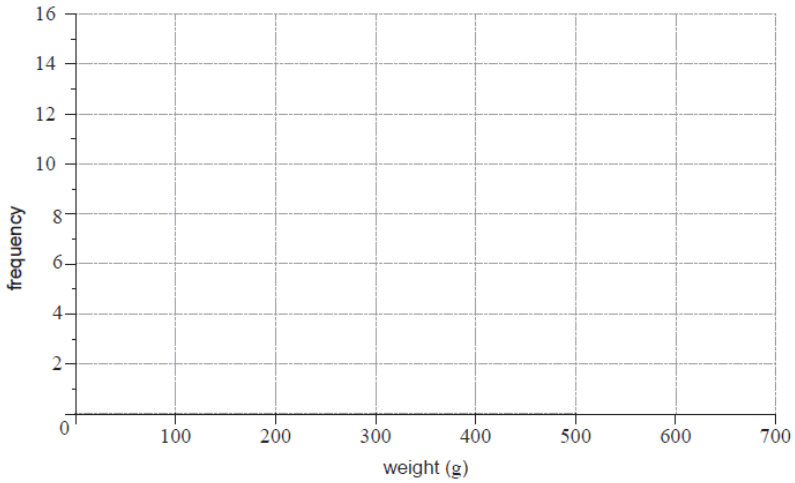
Markscheme

$115 \quad (115.265\ldots\ (g))$ ***(A2) (C2)***

Note: Award ***(A1)(A0)*** for an answer of $116\ (116.459\ldots)$.

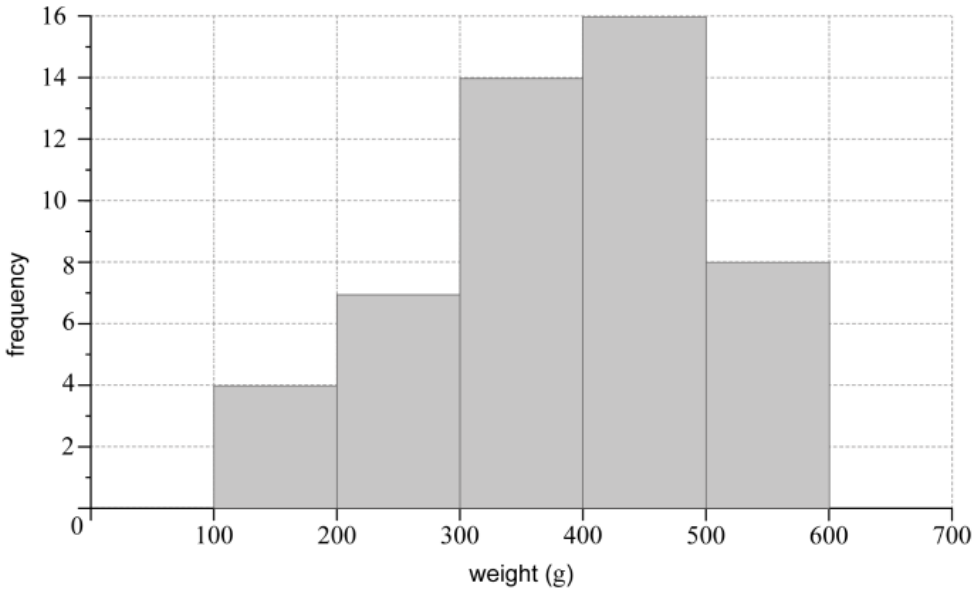
[2 marks]

(c) On the grid below, draw a histogram for the data in the table.



[3]

Markscheme



(A2)(A1) (C3)

Note: Award *(A2)* for all correct heights of bars or *(A1)* for three or four correct heights of bars.

Award ***(A1)*** for rectangular bars all with correct left and right end points ($100,\ 200,\ 300,\ 400,\ 500$ and 600) and for no gaps; the bars do **not** have to be shaded.
Award at most ***(A2)(A0)*** if a ruler is not used for all lines.

[3 marks]

4. [Maximum mark: 6] 19M.1.SL.TZ1.T_2

The fastest recorded speeds of eight animals are shown in the following table.

Animal	Speed (km h^{-1})
Golden eagle	300
Swordfish	97
Hare	80
Lion	80
Horse	71
Zebra	64
Komodo dragon	21
Tiger beetle	6

(a) State whether **speed** is a continuous or discrete variable. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

continuous ***(A1)(C1)***

[1 mark]

(b) Write down the median speed for these animals. [1]

Markscheme

75.5 (km h^{-1}) ***(A1)(C1)***

Note: Answer must be exact.

[1 mark]

(c) Write down the range of the animal speeds. [1]

Markscheme

294 (km h^{-1}) ***(A1) (C1)***

[1 mark]

(d.i) For these eight animals find the mean speed. [2]

Markscheme

$\frac{300+97+80+80+71+64+21+6}{8}$ **OR** $\frac{719}{8}$ ***(M1)***

Note: Award ***(M1)*** for correct sum divided by 8.

89.9 (89.875)(km h^{-1}) ***(A1) (C2)***

[2 marks]

(d.ii) For these eight animals write down the standard deviation. [1]

Markscheme

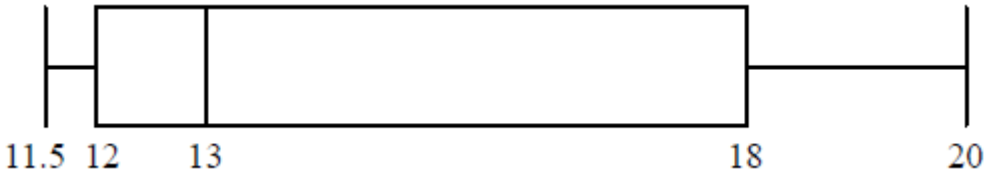
84.6 (84.5597…)(km h^{-1}) ***(A1) (C1)***

Note: If the response to part (d)(i) is awarded zero marks, a correct response to part (d)(ii) is awarded ***(C2)***.

[1 mark]

5. [Maximum mark: 6] 19M.1.SL.TZ2.T_6

A health inspector analysed the amount of sugar in 500 different **snacks** prepared in various school cafeterias. The collected data are shown in the following box-and-whisker diagram.



Amount of sugar per snack in grams

(a) State what 13 represents in the given diagram. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

median ***(A1) (C1)***

[1 mark]

(b.i) Write down the interquartile range for this data. [2]

Markscheme

18 – 12 ***(A1)***

Note: Award ***(M1)*** for correct quartiles seen.

6 (g) ***(A1) (C2)***

[2 marks]

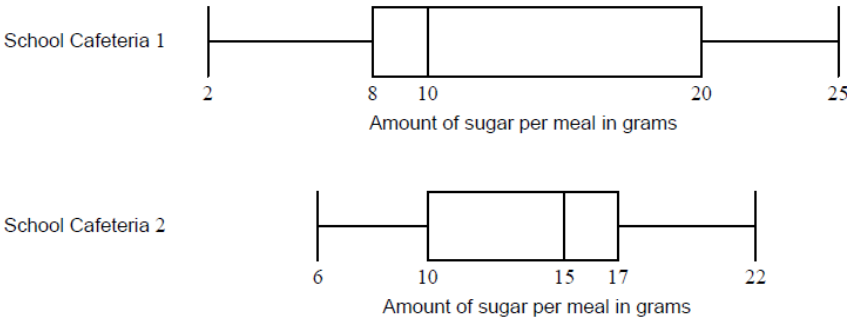
(b.ii) Write down the approximate number of snacks whose amount of sugar ranges from 18 to 20 grams. [1]

Markscheme

125 ***(A1) (C1)***

[1 mark]

(c) The health inspector visits two school cafeterias. She inspects the same number of **meals** at each cafeteria. The data is shown in the following box-and-whisker diagrams.



Meals prepared in the school cafeterias are required to have less than 10 grams of sugar.

State, giving a reason, which school cafeteria has more meals that **do not** meet the requirement. [2]

Markscheme

Cafeteria 2 ***(A1) (C1)***

75 % > 50 % (do not meet the requirement) ***(R1) (C1)***

OR

25 % < 50 % (meet the requirement) ***(R1) (C1)***

Note: Do not award ***(A1)(R0)***. Award the ***(R1)*** for a correct comparison of percentages for both cafeterias, which may be in words. The percentage values or fractions must be seen. It is possible to award ***(A0)(R1)***.

[2 marks]

6. [Maximum mark: 6] 19M.1.SL.TZ2.T_12

University students were surveyed and asked how many hours, h, they worked each month. The results are shown in the following table.

Hours per month, h	Frequency	Cumulative frequency
$0 < h \leq 10$	3	3
$10 < h \leq 20$	7	10
$20 < h \leq 30$	10	20
$30 < h \leq 40$	14	34
$40 < h \leq 50$	p	44
$50 < h \leq 60$	6	50
$60 < h \leq 70$	4	54
$70 < h \leq 80$	2	q

Use the table to find the following values.

(a.i) p. [1]

Markscheme

$p = 10$ *(A1) (C1)*

Note: Award *(A1)* for each correct value.

[1 mark]

(a.ii) q. [1]

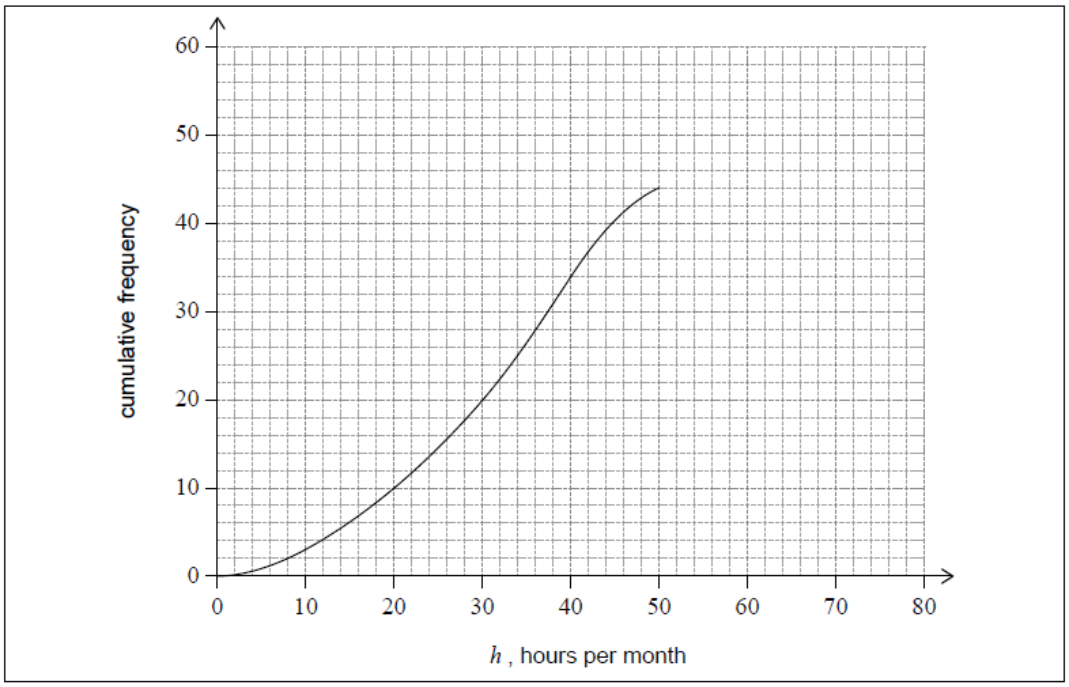
Markscheme

$q = 56$ *(A1) (C1)*

Note: Award *(A1)* for each correct value.

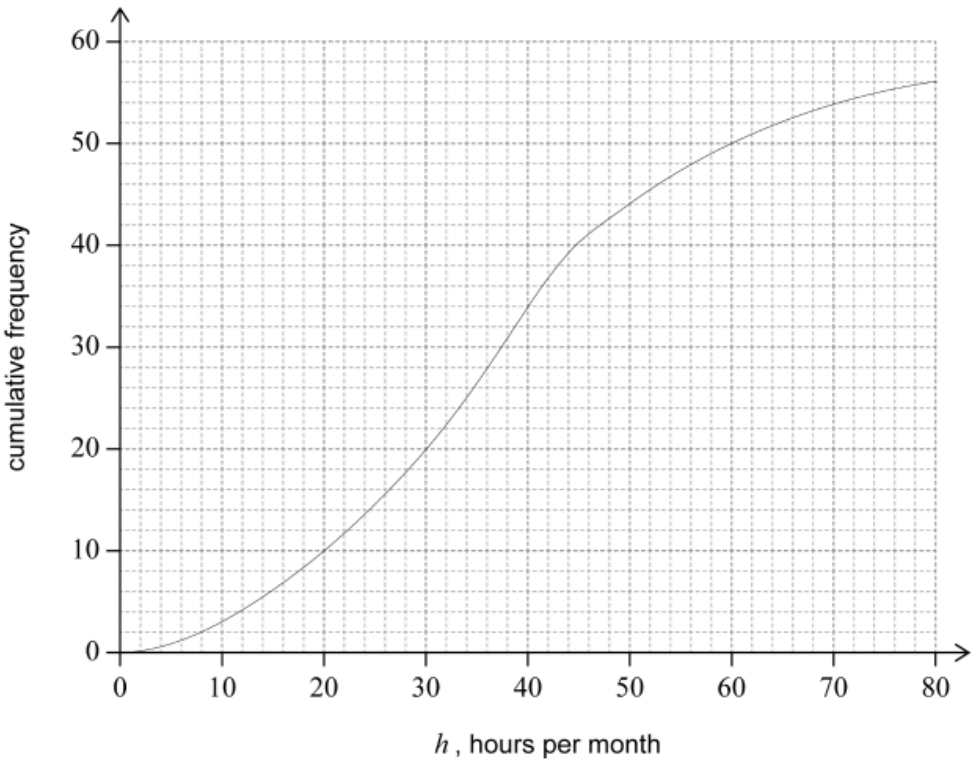
[1 mark]

The first five class intervals, indicated in the table, have been used to draw part of a cumulative frequency curve as shown.



(b) On the same grid, complete the cumulative frequency curve for these data. [2]

Markscheme



(A1)(A1) (C2)

Note: Award ***(A1)*****(ft)** for their 3 correctly plotted points; award ***(A1)*****(ft)** for completing diagram with a smooth curve through their points. The second ***(A1)*****(ft)** can follow through from incorrect points, provided the gradient of the curve is never negative. Award ***(C2)*** for a completely correct smooth curve that goes through the correct points.

[2 marks]

(c) Use the cumulative frequency curve to find an estimate for the number of students who worked at most 35 hours per month. [2]

Markscheme

a straight vertical line drawn at 35 (accept 35 ± 1) ***(M1)***

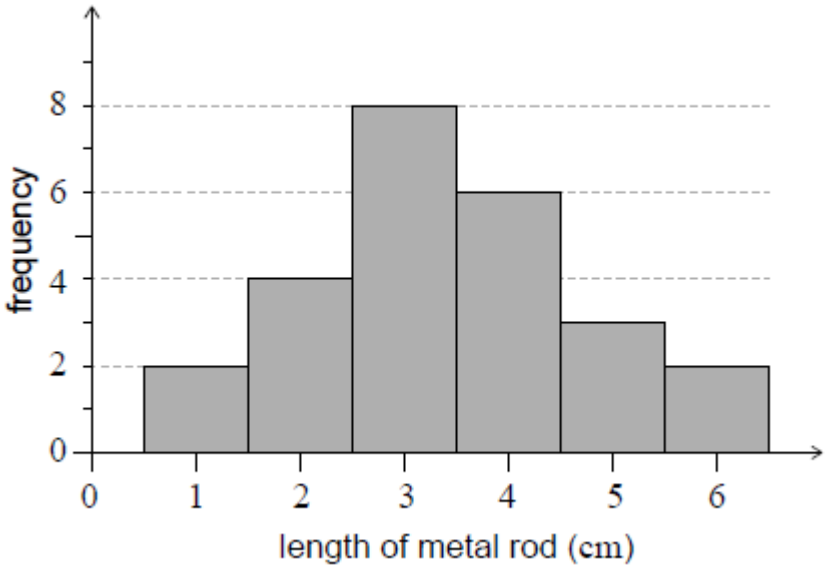
26 (students) ***(A1) (C2)***

Note: Accept values between 25 and 27 inclusive.

[2 marks]

7. [Maximum mark: 6] 18N.1.SL.TZ0.T_2

The histogram shows the lengths of 25 metal rods, each measured correct to the nearest cm.



(a) Write down the modal length of the rods. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

3 ***(A1) (C1)***

[1 mark]

(b) Find the median length of the rods. [3]

Markscheme

median is 13th position ***(M1)***

CF: 2, 6, 14, 20, 23, 25 ***(M1)***

median = 3 **(A1) (C3)**

[3 marks]

The upper quartile is 4 cm.

(c.i) Calculate the lower quartile. [1]

Markscheme

2.5 ***(A1) (C1)***

Note: Award ***(A1)*****(ft)** if the sum of **their** parts (c)(i) and (c)(ii) is 4.

[1 mark]

(c.ii) Calculate the interquartile range. [1]

Markscheme

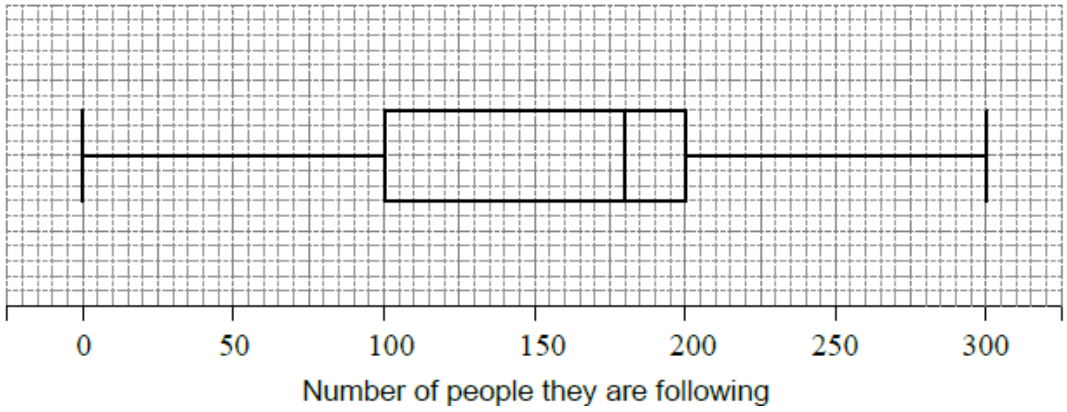
1.5 ***(A1)*****(ft)** ***(C1)***

Note: Award ***(A1)*****(ft)** if the sum of **their** parts (c)(i) and (c)(ii) is 4.

[1 mark]

8. [Maximum mark: 6] 18M.1.SL.TZ1.T_6

In a high school, 160 students completed a questionnaire which asked for the number of people they are following on a social media website. The results were recorded in the following box-and-whisker diagram.



(a) Write down the median. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

180 *(A1)(C1)*

[1 mark]

The following incomplete table shows the distribution of the responses from these 160 students.

Number of people they are following (x)	Number of high school students
$0 \leq x \leq 50$	4
$50 < x \leq 100$	
$100 < x \leq 150$	34
$150 < x \leq 200$	46
$200 < x \leq 250$	
$250 < x \leq 300$	16

(b) Complete the table. [2]

Markscheme

36, 24 ***(A1)(A1) (C2)***

Note: Award ***(A0)(A1)*** for two incorrect values that add up to 60.

[2 marks]

(c.i) Write down the mid-interval value for the $100 < x \leq 150$ group. [1]

Markscheme

125 (accept 125.5) ***(A1)***

(c.ii) Using the table, calculate an estimate for the mean number of people being followed on the social media website by these 160 students. [2]

Markscheme

$\frac{4\times25+36\times75+34\times125+46\times175+24\times225+16\times275}{160}$ ***(M1)***

Note: Award ***(M1)*** for correct substitution of their mid-interval values, multiplied by their frequencies, into mean formula.

=156 (155.625) ***(A1)*****(ft)** ***(C3)***

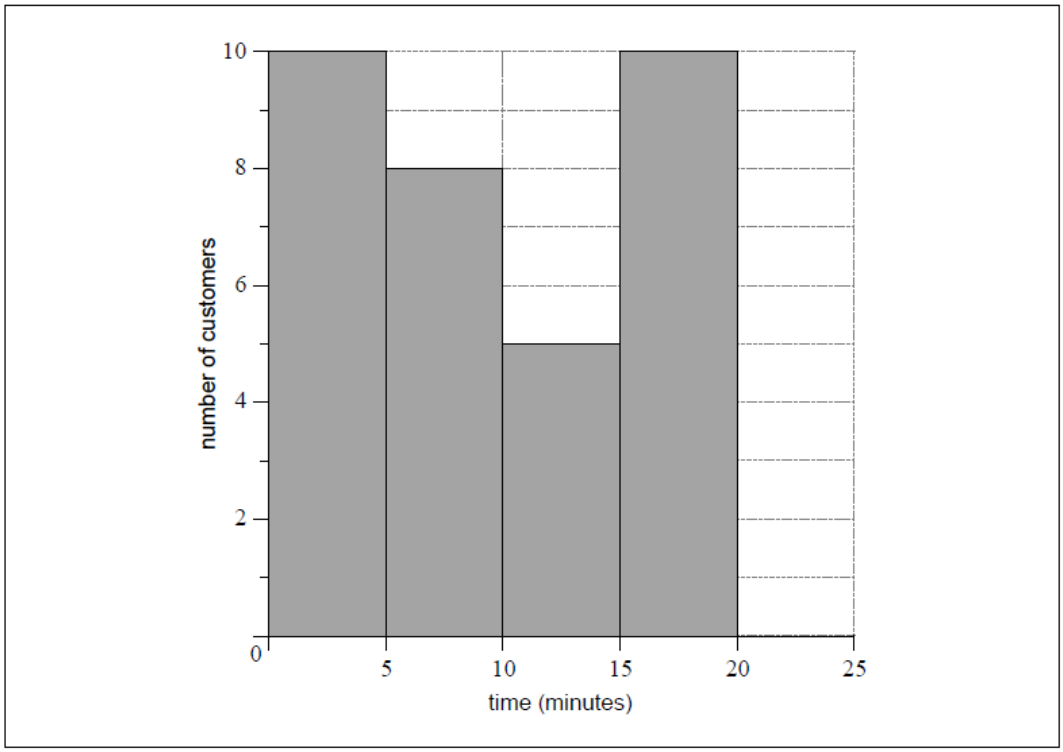
Note: Follow through from parts (b) and (c)(i).

[3 marks]

9. [Maximum mark: 6] 18M.1.SL.TZ2.T_12

The histogram shows the time, t, in minutes, that it takes the customers of a restaurant to eat their lunch on one particular day. Each customer took less than 25 minutes.

The histogram is incomplete, and only shows data for $0 \leq t < 20$.



(a) Write down the mid-interval value for $10 \leq t < 15$. [1]

Markscheme

* This question is from an exam for a previous syllabus, and may contain minor differences in marking or structure.

12.5 ***(A1) (C1)***

[1 mark]

The mean time it took **all** customers to eat their lunch was estimated to be 12 minutes.

It was found that k customers took between 20 and 25 minutes to eat their lunch.

(b.i) Write down the total number of customers in terms of k. [1]

Markscheme

$33 + k$ **OR** $10 + 8 + 5 + 10 + k$ ***(A1)***

Note: Award ***(A1)*** for "number of customers = $33 + k$".

[1 mark]

(b.ii) Calculate the value of k. [3]

Markscheme

$\frac{2.5\times10+7.5\times8+\ldots+22.5\times k}{33+k} = 12$ ***(M1)(A1)*****(ft)**

Note: Award ***(M1)*** for substitution into the mean formula and equating to 12, ***(A1)*****(ft)** for their correct substitutions.

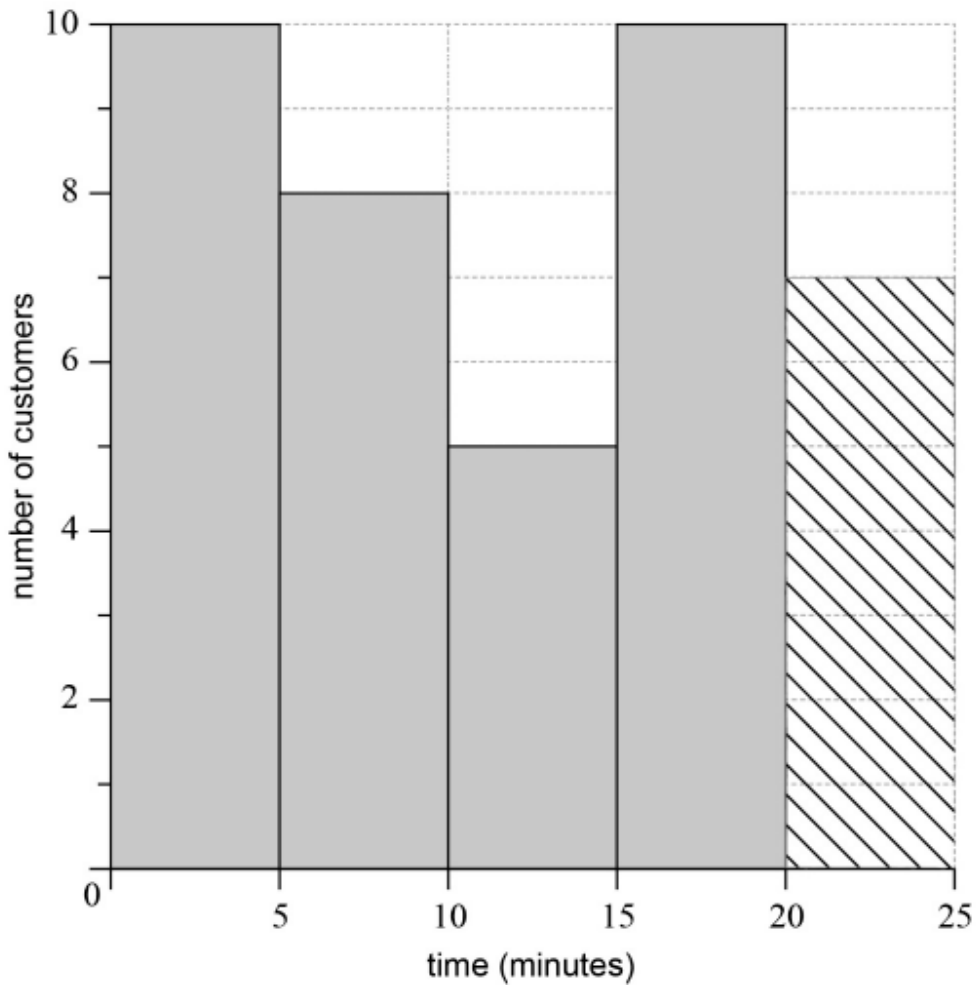
($k =$) 7 ***(A1)*****(ft)** ***(C4)***

Note: Follow through from part (b)(i) and their mid-interval values, consistent with part (a). Do not award final ***(A1)*** if answer is not an integer.

[3 marks]

(c) Hence, complete the histogram. [1]

Markscheme



(A1)(**ft**) *(C1)*

Note: Follow through from their part (b)(ii) but only if the value is between 1 and 10, inclusive.

[1 mark]