

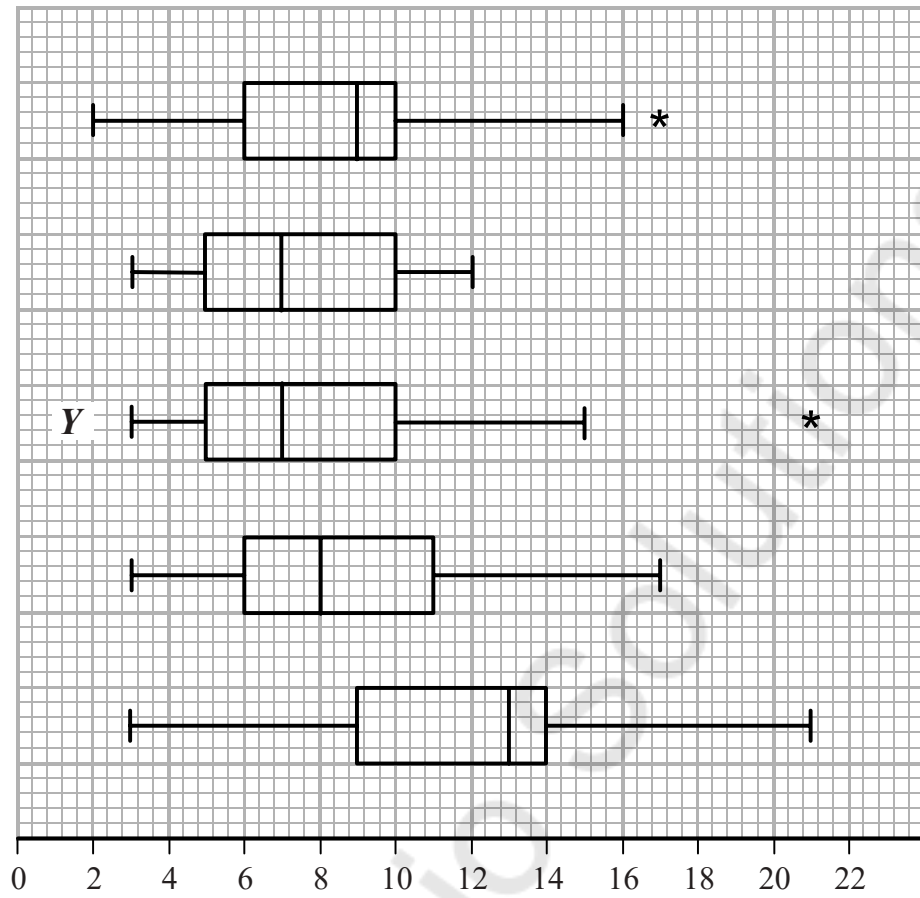
Y1S3 XMQs and MS

(Total: 31 marks)

1. P31(AS)_2018 Q4 . 8 marks - Y1S1 Data collection
2. P31(AS)_2020 Q1 . 4 marks - Y1S3 Representations of data
3. P31(AS)_2021 Q2 . 9 marks - Y1S3 Representations of data
4. P31(AS)_2022 Q3 . 8 marks - Y1S3 Representations of data
5. P31(AS)_2022 Q4 . 2 marks - Y1S3 Representations of data

Mathvault.io Solutions

Question 4 continued



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 4 is 8 marks)



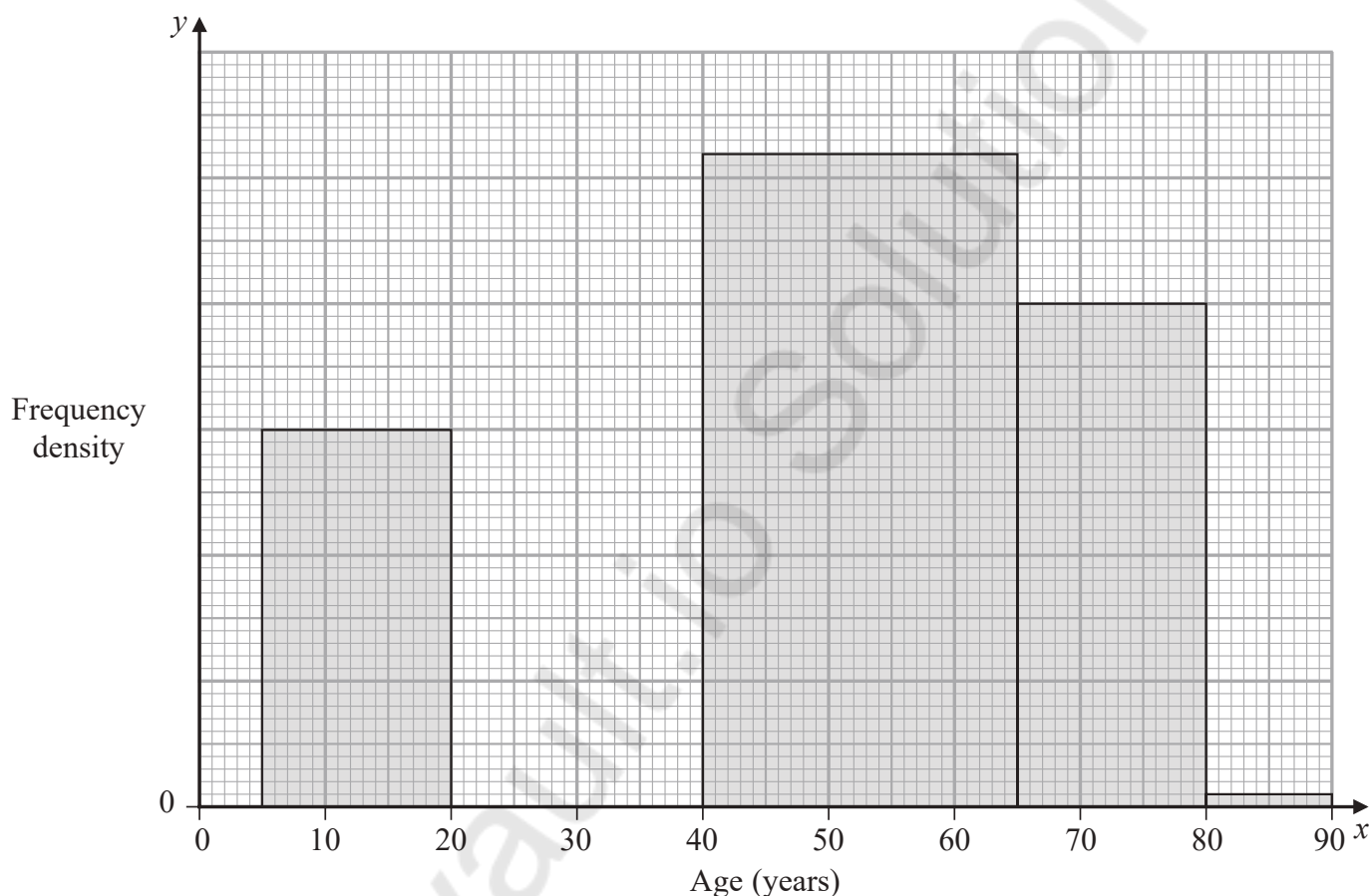
Qu	Scheme	Marks	AO	
4	(a) $\bar{x} = 10.2$ (2222...)	awrt 10.2	B1	1.1b
			(1)	
	(b) $\sigma_x = 3.17$ (20227...)	awrt 3.17	B1ft	1.1b
	Sight of “knots” <u>or</u> “kn” (condone knots/s etc)		B1	1.2
			(2)	
	(c) October since it is windier in the autumn <u>or</u> month of the hurricane <u>or</u> latest month in the year		B1	2.2b
			B1	2.4
			(2)	
	(d)(i) They represent <u>outliers</u>		B1	1.2
	(ii) Y has low median so expect lowish mean (but outlier so > 7) <u>and</u> Y has big range/IQR or spread so expect larger st.dev Suggests B		M1	2.4
		A1	2.2b	
		(3)		
(8 marks)				
Notes				
NB	$\bar{x} = \frac{184}{18}$ and $\sigma_x = \sqrt{\frac{2062}{18} - \bar{x}^2}$			
(a)	B1 for $\bar{x} = 10.2$ (allow exact fraction)	[This is an LDS mark]		
(b)	1 st B1ft allow 3.2 from a correct expr' accept $s = 3.26(3984...)$ [ft use of n/a] <u>Treating n/a as 0</u> May see $n = 31$ or $\bar{x} = 5.9354...$ which is B0 in (a) but here in (b) it gives $\sigma_x = 5.59(34...)$ or $s = 5.6858...$ (awrt 5.69) and scores 1 st B1 2 nd B1 accept kn accept in (a) or (b) (allow nautical miles/hour)	[This is an LDS mark]		
(c)	1 st B1 choosing October but accept September. 2 nd B1 for stating that (Camborne) is windier in autumn/winter months “because it is winter/autumn/windier/colder in “month” ” Sep \leq "month" \leq Mar scores B1B1 for “month” = Sep or Oct and B0B1 for other months in range	[This is an LDS mark]		
(d)(i)	B1 for outlier or the idea of an extreme value allow “anomaly”			
(ii)	M1 for a comment relating to location that mentions both median and mean <u>and</u> a comment relating to <u>spread</u> that mentions both range/IQR and standard deviation and leads to choosing B , C or D			
	Choosing A or E is M0			
	Incorrect/false statements score M0 e.g. $Q_3 = (\text{mean} + \sigma)$ or identify $Q_2 = \text{mean}$ or Y has small spread			
ALT	Use of outliers: outlier is $(\text{mean} + 3\sigma)$ ($B = 19.9$), ($C = 18.95$), ($D = 20.2$) Must <u>see</u> at least one of these values and compare to Y 's outlier [leads to D or B]			
	A1 for suitable inference i.e. B (accept D <u>or</u> B or D) M1 must be scored			

Question	Scheme	Marks	AOs
1	1 square is $\frac{78}{12 \times 3 + 3 \times 4 + 2 \times 2} = \left[\frac{78}{52} = 1.5 \right]$ and $(8 \times 1 + 1 \times 8) \times "1.5"$	M1	3.1a
	24 students took less than 11 minutes	A1	1.1b
	Percentage of students = $\frac{"24"}{78 + "24" + 1 \times 8 \times "1.5" + 3 \times 4 \times "1.5"} \times 100$	M1	3.1b
	= 18.18... awrt 18%	A1	1.1b
		(4)	
Total 4			
Notes			
1	M1:	For clear use of frequency density to establish the fd scale and then use the area to find frequency of <11 minutes. Allow maximum of 3 errors in either the heights or widths in total if working shown. They may calculate the area using other size squares. Allow for realising they need to find the total number of squares (88) maximum of 4 errors in either the heights or widths and number < 11 minutes(16) - must have a maximum of 1 error in either the heights or widths (and not use the 78 as part of calculation)	
	A1:	For correct values seen. Allow for 88 and 16	
	M1:	For realising the need to find the total and calculating a percentage. (with "their 24" as the numerator). Allow $(8 \times 1 + 2 \times 8) \times "1.5"$ instead of $"24" + 1 \times 8 \times "1.5"$ If working shown can allow maximum of 2 errors in either the heights or widths in the calculation of the total. Allow "their 24" / 132 oe	
	A1:	awrt 18	

2. The partially completed table and partially completed histogram give information about the ages of passengers on an airline.

There were no passengers aged 90 or over.

Age (x years)	$0 \leq x < 5$	$5 \leq x < 20$	$20 \leq x < 40$	$40 \leq x < 65$	$65 \leq x < 80$	$80 \leq x < 90$
Frequency	5	45	90			1



- (a) Complete the histogram. (3)

- (b) Use linear interpolation to estimate the median age. (4)

An outlier is defined as a value greater than $Q_3 + 1.5 \times$ interquartile range.

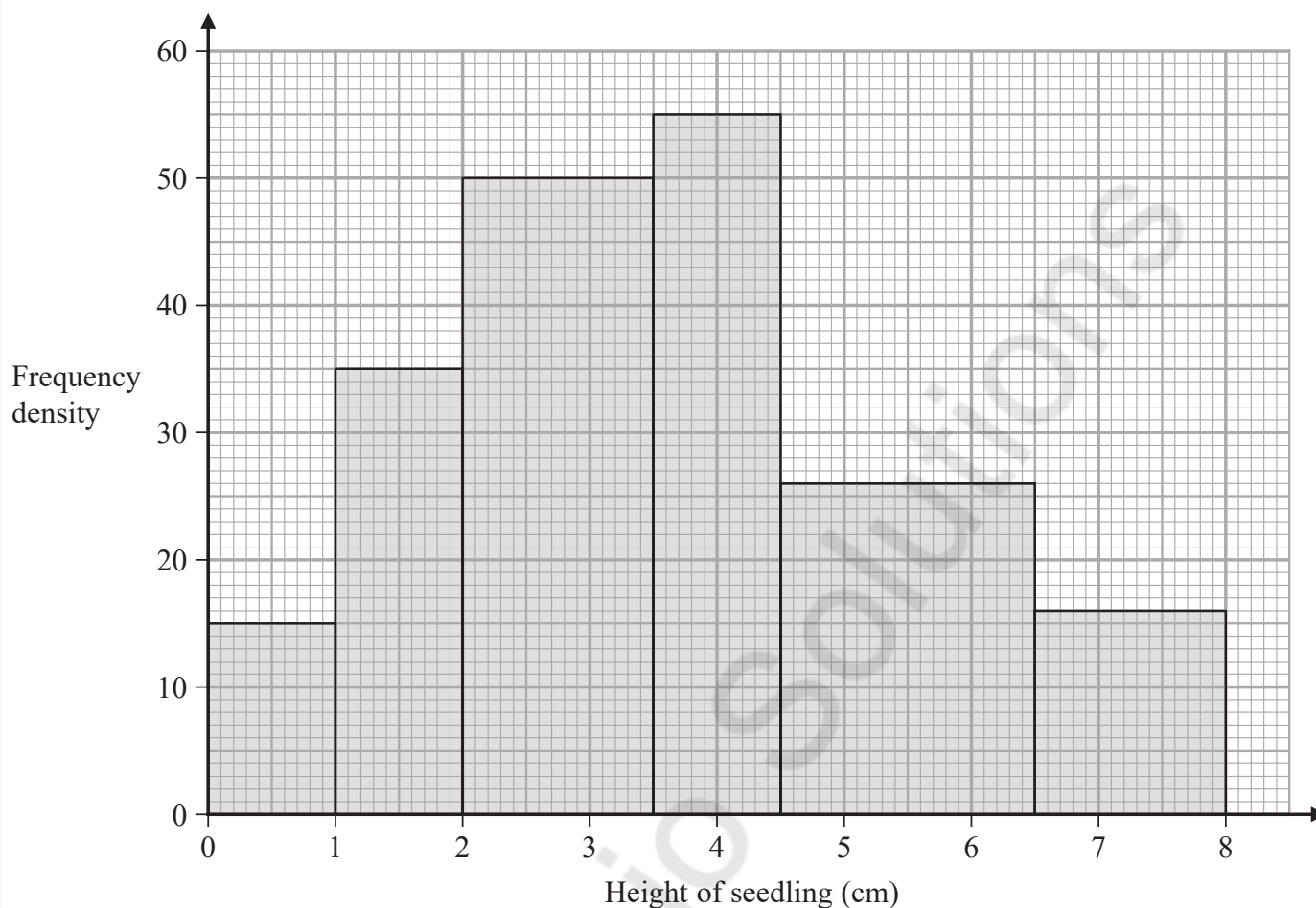
Given that $Q_1 = 27.3$ and $Q_3 = 58.9$

- (c) determine, giving a reason, whether or not the oldest passenger could be considered as an outlier. (2)



Qu	Scheme	Marks	AO
2. (a)	From [5,20) fd = 3 <u>or</u> 1 large square = 2.5 passengers o.e.	M1	2.2a
	Correct bar above [0, 5)	A1	1.1b
	Correct bar above [20, 40)	A1	1.1b
			(3)
	(b) For [40, 65) 130 passengers <u>or</u> for [65, 80) 60 passengers	M1	2.1
	For attempt to find total number of passengers = 331	A1ft	1.1b
	[Median =] $40 + \frac{\frac{1}{2}("331") - 140}{"130"} \times 25$ <u>or</u> $65 - \frac{270 - \frac{1}{2}("331")}{"130"} \times 25$ (o.e.)	M1	1.1b
	$= 44.9038... = \text{awrt } \underline{44.9}$	A1	1.1b
			(4)
	(c) Upper outlier limit = $58.9 + 1.5 \times (58.9 - 27.3) = 106 (.3) > 90$ So oldest passenger is <u>not</u> an outlier	M1	2.4
	A1	2.2a	
		(2)	
		(9 marks)	
Notes			
(a)	M1 for attempt at fd or a suitable method to deduce the scale for the histogram May be implied by one correct bar. 1 st A1 for first bar [0, 5) with fd = 1 <u>or</u> 2 large squares high 2 nd A1 for third bar with fd = 4.5 <u>or</u> 9 large squares high		
(b)	1 st M1 for an attempt using their fd to find the missing frequencies. May be in table 1 st A1ft for a clear attempt to find the total number of passengers (ft their 130 and 60) 2 nd M1 for any expression/equation leading to correct Q_2 Must be using 40-65 class 2 nd A1 for awrt 44.9 (allow $(n + 1)$ leading to 45)		
(c)	M1 for finding the upper outlier limit (expression or awrt 106) <u>and</u> stating or implying > 90 A1 dep on M1 seen for deducing NOT an outlier		

3. The histogram summarises the heights of 256 seedlings two weeks after they were planted.



- (a) Use linear interpolation to estimate the median height of the seedlings.

(4)

Chris decides to model the **frequency density** for these 256 seedlings by a curve with equation

$$y = kx(8 - x) \quad 0 \leq x \leq 8$$

where k is a constant.

- (b) Find the value of k

(3)

Using this model,

- (c) write down the median height of the seedlings.

(1)



Qu	Scheme	Mark	AO															
3. (a)	<table border="1"> <thead> <tr> <th>Class</th> <th>Frequency</th> <th>Cum. Frequency</th> </tr> </thead> <tbody> <tr> <td>0 – 1</td> <td>15</td> <td>15</td> </tr> <tr> <td>1 – 2</td> <td>35</td> <td>50</td> </tr> <tr> <td>2 – 3.5</td> <td>75</td> <td>125</td> </tr> <tr> <td>3.5 – 4.5</td> <td>55</td> <td>180</td> </tr> </tbody> </table>	Class	Frequency	Cum. Frequency	0 – 1	15	15	1 – 2	35	50	2 – 3.5	75	125	3.5 – 4.5	55	180	M1	2.1
	Class	Frequency	Cum. Frequency															
	0 – 1	15	15															
	1 – 2	35	50															
2 – 3.5	75	125																
3.5 – 4.5	55	180																
$[Q_2 =](3.5) + \frac{\frac{256}{2} - "125"}{"55"} \times (4.5 - 3.5) \text{ or } (4.5) - \frac{"180" - \frac{256}{2}}{"55"} \times 1$ $= 3.5545\dots\dots \text{ awrt } \underline{\underline{3.55}}$	M1	2.1																
	A1	1.1b																
	A1	1.1b																
		(4)																
(b)	Need area under curve to be 256 so $\int_{(0)}^{(8)} kx(8-x) dx = 256$	M1	3.1a															
	$k \left[4x^2 - \frac{x^3}{3} \right]_{(0)}^{(8)} = 256$ $\left\{ k \left[4 \times 8^2 - \frac{8}{3} \times 8^2 \right] = 256 \Rightarrow \right\} \underline{\underline{k = 3}}$	M1	1.1b															
		A1	1.1b															
		(3)																
(c)	[By symmetry median =] <u>4</u>	B1	2.2a															
		(1)																
		(8 marks)																

Notes

(a)	<p>1st M1 for an attempt to form frequency table (at least 1st 4 rows and freq or cum freq seen must have the frequency of 75 correct and can condone one error/omission in 15, 35, 55) Frequencies or cum freq may be seen on bars of the histogram</p> <p>1st A1 for identifying class, freq and cum freq (i.e. highlighted values from the table) or sight of 3.5-4.5, freq of 55 and “128” – 125 or 180 – “128” or diagram with 125, “128”, 180, 3.5 & 4.5 May be implied by values in 2nd M1 expression</p> <p>2nd M1 for a correct calculation for Q_2 (condone error in end point e.g. 3.45 or 3.49 etc) Can fit their “125” (provided > 100) and their “55” Allow use of $(n + 1)$, usually see 128.5 – ... leading to 3.5636... or awrt 3.56</p> <p>2nd A1 awrt 3.55 but 3.555 is fine (allow 3.56 if $(n + 1)$ being used ... need sight of $\frac{257}{2}$ etc) Correct answer with no incorrect working scores 4/4</p>		
(b)	<p>1st M1 for identifying the need to find the area under the curve by integrating</p> <p>2nd M1 for correct integration and = 256 (condone missing limits)</p> <p>A1 for $k = 3$ [May see use of calculator for the integration so score 2nd M1A1 together]</p>		
(c)	<p>NB The answer to part (c) may be written within the question.</p> <p>B1 for 4 (Independent of their value of k but must be their “x” value) NB when $k = 0.25$ and $x = 4$ gives $y = 4$ so must be clear they intend median = 4 The statement in part (c) “$k = 4$” is B0</p>		

	Scheme	Marks	AO
--	--------	-------	----

<p>4. (a)</p> <p>(b)</p>	<p>Accept 990 to 1030 inclusive</p> <p>Any range between 10 and 50 inclusive</p>	<p>B1 (1)</p> <p>B1 (1)</p>	<p>1.1b</p> <p>1.1b</p>
		(2 marks)	
Notes			
<p>(a)</p>	<p>B1 (Median pressures usually around 1000~1020)</p>		<p>[LDS mark]</p>
<p>(b)</p>	<p>B1 Any answer in this range Allow answers in the form $a \sim b$ where $b - a$ is between 10 and 50 Also allow the case where <u>both</u> a and b are in $[10, 50]$</p>		<p>[LDS mark]</p>

Mathvault.io Solutions