

# TOPICAL PAST PAPER QUESTIONS WORKBOOK

---

## AS & A Level Mathematics (9709) Paper 5 [Probability & Statistics 1]

---

Exam Series: May 2015 – May 2022

Format Type B:

Each question is followed by its answer scheme



**EXAMINENT.COM**  
Eminent Exam Preparation Resources



# Introduction

Each topical past paper questions workbook consists of hundreds of questions and their answer schemes, in the form of worksheets. Questions are assigned to each chapter according to their corresponding topic. Topics, in turn, are based on the items of the latest Cambridge IGCSE or AS/A level syllabus content. This book's specifications are as follows:

Title: AS & A Level Mathematics (9709) Paper 5 Topical Past Paper Questions Workbook

Subtitle: Exam Practice Worksheets With Answer Scheme

Examination board: Cambridge Assessment International Education (CAIE)

Subject code: 9709

Years covered: May 2015 – May 2022

Paper: 5 (Probability & Statistics 1)

Number of pages: 773

Number of questions: 362



# Contents

<b>1</b>	<b>Representation of data</b>	<b>7</b>
<b>2</b>	<b>Permutations and combinations</b>	<b>169</b>
<b>3</b>	<b>Probability</b>	<b>285</b>
<b>4</b>	<b>Discrete random variables</b>	<b>415</b>
<b>5</b>	<b>The normal distribution</b>	<b>581</b>



# Chapter 1

## Representation of data

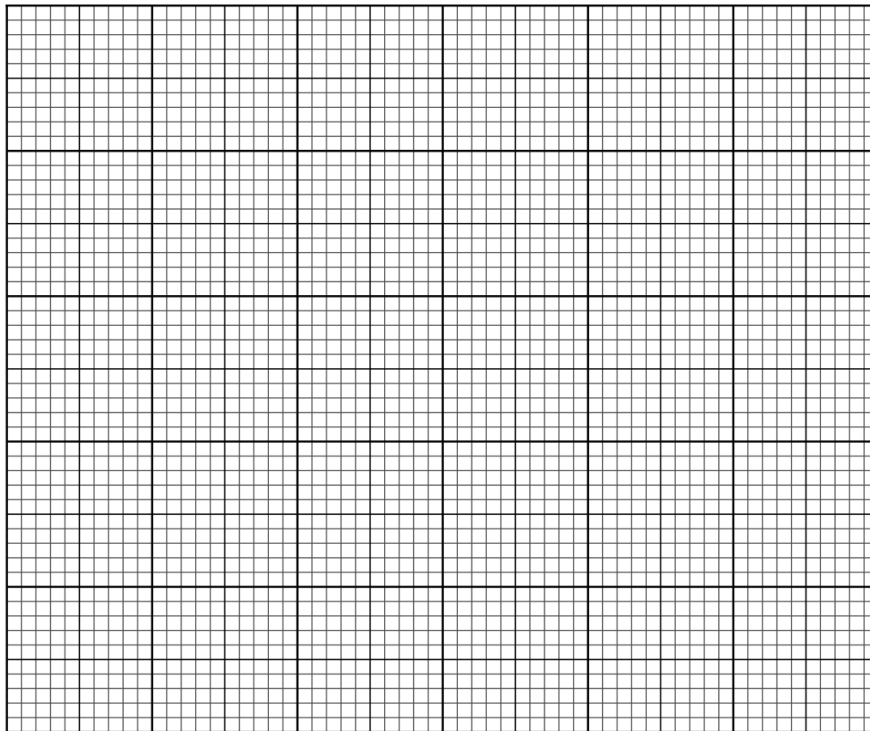
1. 9709\_m22\_qp\_52 Q: 3

At a summer camp an arithmetic test is taken by 250 children. The times taken, to the nearest minute, to complete the test were recorded. The results are summarised in the table.

Time taken, in minutes	1 – 30	31 – 45	46 – 65	66 – 75	76 – 100
Frequency	21	30	68	86	45

(a) Draw a histogram to represent this information.

[4]



(b) State which class interval contains the median.

[1]

.....

.....

.....

(c) Given that an estimate of the mean time is 61.05 minutes, state what feature of the distribution accounts for the median and the mean being different.

[1]

.....

.....

.....



Answer:

Question	Answer						Marks	Guidance
(a)							<b>M1</b>	At least 4 frequency densities calculated
	Class Width	30	15	20	10	25	<b>A1</b>	All heights correct on graph
	Frequency Density	0.7	2	3.4	8.6	1.8	<b>B1</b>	Bar ends at 0.5, 30.5, 45.5, 65.5, 75.5, 100.5 (at axis), 5 bars drawn, condone 0 in first bar $0.5 \leq \text{time axis} \leq 100.5$ , linear scale with at least 3 values indicated.
							<b>B1</b>	Axes labelled: Frequency density (fd), time (t) and mins (or appropriate title). Linear fd scale, with at least 3 values indicated $0 \leq \text{fd axis} \leq 8.6$
							<b>4</b>	
(b)	66 – 75						<b>B1</b>	Condone 65.5 – 75.5
							<b>1</b>	
(c)	Distribution is not symmetrical						<b>B1</b>	Or skewed, ignore nature of skew
							<b>1</b>	

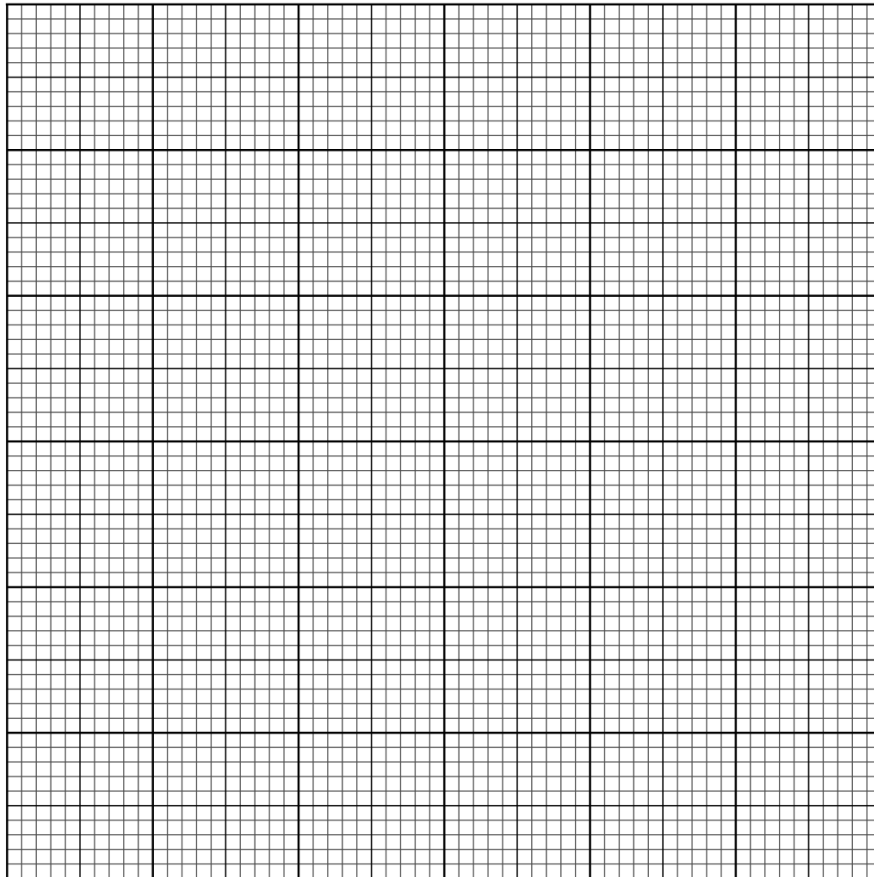
2. 9709\_s22\_qp\_51 Q: 3

The times taken to travel to college by 2500 students are summarised in the table.

Time taken ( $t$ minutes)	$0 \leq t < 20$	$20 \leq t < 30$	$30 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 90$
Frequency	440	720	920	300	120

(a) Draw a histogram to represent this information.

[4]



From the data, the estimate of the mean value of  $t$  is 31.44.

- (b)** Calculate an estimate of the standard deviation of the times taken to travel to college. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (c)** In which class interval does the upper quartile lie? [1]

.....

.....

.....

.....

It was later discovered that the times taken to travel to college by two students were incorrectly recorded. One student's time was recorded as 15 instead of 5 and the other's time was recorded as 65 instead of 75.

- (d)** Without doing any further calculations, state with a reason whether the estimate of the standard deviation in part **(b)** would be increased, decreased or stay the same. [1]

.....

.....

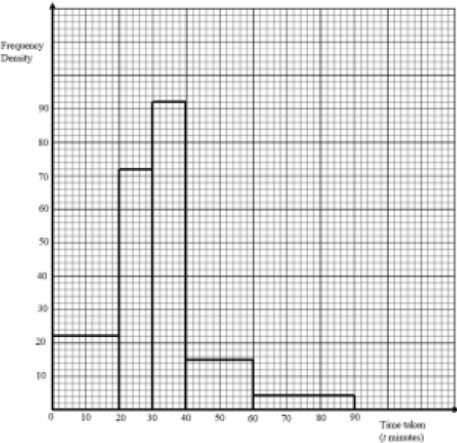
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance												
(a)	<table><tr><td>Class width</td><td>20</td><td>10</td><td>10</td><td>20</td><td>30</td></tr><tr><td>Frequency density</td><td>22</td><td>72</td><td>92</td><td>15</td><td>4</td></tr></table>	Class width	20	10	10	20	30	Frequency density	22	72	92	15	4	M1	At least 4 frequency densities calculated (Frequency $\div$ class width, e.g. $\frac{440}{20}$ (condone $\frac{440}{19.5}$ , $\frac{440}{20.5}$ ) Accept unsimplified, may be read from graph using <i>their</i> scale
	Class width	20	10	10	20	30									
	Frequency density	22	72	92	15	4									
		A1	All heights correct on graph NOT FT												
B1		Bar ends at [0,] 20, 30, 40, 60, 90 at axis with a horizontal linear scale with at least 3 values indicated. $0 \leq$ horizontal scale $\leq 90$													
B1		Axes labelled frequency density (fd), time (t) and minutes (mins) or in a title. Linear vertical scale, with at least 3 values indicated $0 \leq$ vertical axes $\leq 92$ (condone 90 used).													
		4													
Question	Answer	Marks	Guidance												
(b)	<table><tr><td>Midpoints</td><td>10</td><td>25</td><td>35</td><td>50</td><td>75</td></tr></table>	Midpoints	10	25	35	50	75	B1	At least 4 correct midpoints seen						
	Midpoints	10	25	35	50	75									
	<p>[Mean = 31.44 given]</p> $\text{[Variance} = \frac{440 \times 10^2 + 720 \times 25^2 + 920 \times 35^2 + 300 \times 50^2 + 120 \times 75^2}{2500} - 31.44^2]$ $= \frac{44000 + 450000 + 1127000 + 750000 + 675000}{2500} - 31.44^2$ $[= \frac{3046000}{2500} - 31.44^2 = 229.9264]$ <p>Or</p> $\text{Variance} = \frac{440(10 - 31.44)^2 + 720(25 - 31.44)^2 + 920(35 - 31.44)^2 + 300(50 - 31.44)^2 + 120(75 - 31.44)^2}{2500}$ $= \frac{202256 + 29860 + 11659 + 103342 + 227697}{2500} = \frac{574814}{2500} = 229.9264$	M1	Correct formula for variance or standard deviation ( $- \text{mean}^2$ included with <i>their</i> midpoints (not upper bound, lower bound, class width, frequency density, frequency or cumulative frequency) and <i>their</i> $\sum f$ if calculated. Condone 1 data error.												
	Standard deviation = 15.2	A1	WWW, allow 15.16[3...]												
			3												
(c)	30–40	B1													
		1													
(d)	Stays the same, data still in same intervals	B1	Frequencies unchanged												
		1													

3. 9709\_s22\_qp\_52 Q: 1

For  $n$  values of the variable  $x$ , it is given that

$$\Sigma(x - 200) = 446 \quad \text{and} \quad \Sigma x = 6846.$$

Find the value of  $n$ .

[3]

[illegible]

Answer:

Question	Answer	Marks	Guidance
	$\Sigma x - \Sigma 200 = \Sigma (x - 200)$	<b>B1</b>	Forming a correct 3-term (linear) equation from $\Sigma x$ , $\Sigma 200$ and $\Sigma (x - 200)$ . Accept $6846 - 200n = 446$ OE. Condone 1 sign error.
	$\Sigma 200 = 200n$	<b>B1</b>	SOI
	$[200n = 6846 - 446 = 6400] \quad n = 32$	<b>B1</b>	WWW
		<b>3</b>	

4. 9709\_s22\_qp\_52 Q: 3

The back-to-back stem-and-leaf diagram shows the diameters, in cm, of 19 cylindrical pipes produced by each of two companies, *A* and *B*.

Company <i>A</i>							Company <i>B</i>					
					4	33	1	2	8			
9	8	3	2	0		34	1	6	8	9	9	
8	7	5	4	1	1	35	1	2	2	3		
		9	6	5	2	36	5	6				
			4	3	1	37	0	3	4			
						38	2	8				

Key: 1 | 35 | 3 means the pipe diameter from company *A* is 0.351 cm and from company *B* is 0.353 cm.

- (a) Find the median and interquartile range of the pipes produced by company *A*. [3]

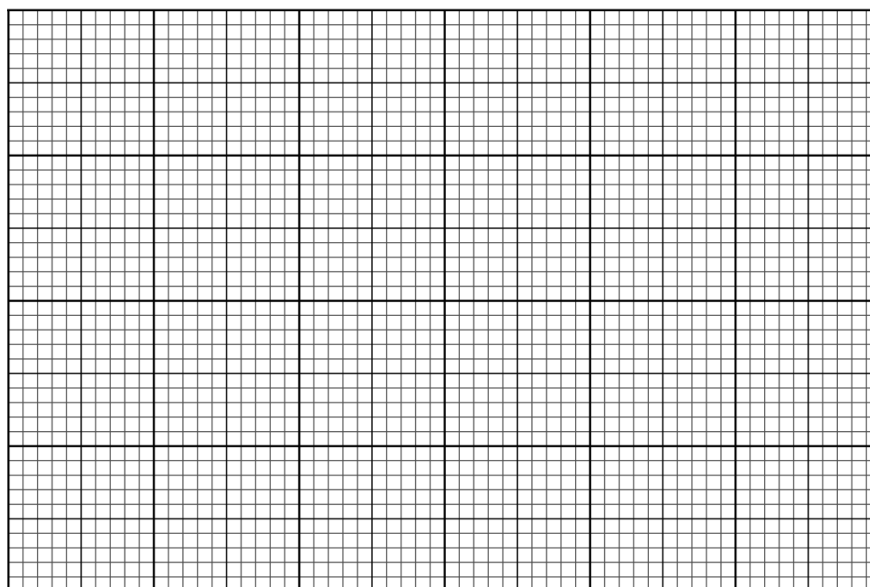
.....

.....

.....

It is given that for the pipes produced by company *B* the lower quartile, median and upper quartile are 0.346 cm, 0.352 cm and 0.370 cm respectively.

- (b) Draw box-and-whisker plots for companies *A* and *B* on the grid below. [3]

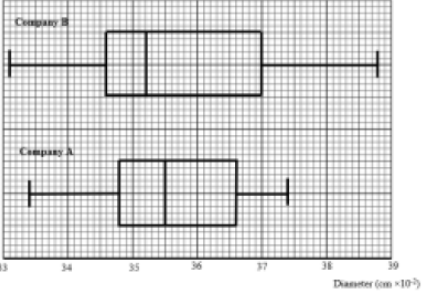


- (c) Make one comparison between the diameters of the pipes produced by companies *A* and *B*. [1]

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Median = 0.355	<b>B1</b>	Identified condone Q2.
	[IQR =] 0.366 – 0.348	<b>M1</b>	$0.365 < UQ \leq 0.369 - 0.343 < LQ \leq 0.349$ . Subtraction may be implied by answer.
	0.018	<b>A1</b>	If 0/3 scored <b>SC B1</b> for figs Median = 355 IQR = 18.
		<b>3</b>	
(b)	Box-and-whisker plot on provided grid 	<b>B1</b>	All 5 key values for <i>B</i> plotted accurately in standard format using <i>their</i> scale. Labelled <i>B</i> . Check accuracy in the middle of vertical line.
		<b>B1 FT</b>	All 5 key values for <i>A</i> , FT from part 3(a), plotted in standard format accurately using <i>their</i> scale. Labelled <i>A</i> . Check accuracy in the middle of vertical line.
		<b>B1</b>	Whiskers not through box for both, not drawn at corners of boxes, single linear scale with at least 3 values stated, covering at least 0.34 to 0.38 and labelled diameter ( <i>d</i> etc) and cm. Accept as a title.
		<b>3</b>	If both plots attempted and plot(s) not labelled, <b>SC B1</b> for at least 1 fully correct set of values plotted.
(c)	A comparison in context	<b>B1</b>	Single comment comparing spread or central tendency in context. Must reference either diameter or pipes. Not a simple numerical comparison of statistical values such as median, range, IQR or min/max.
		<b>1</b>	



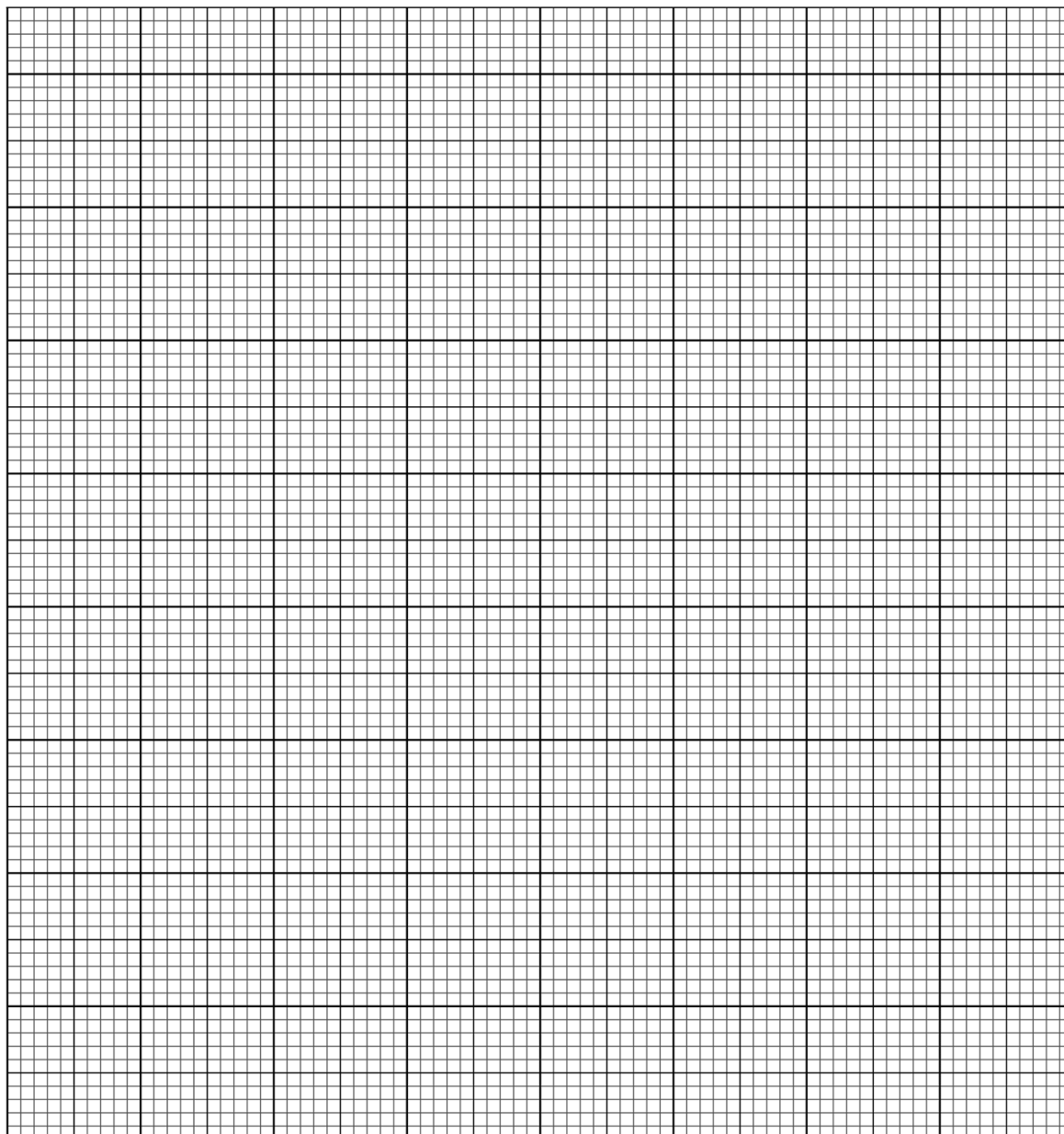
5. 9709\_s22\_qp\_53 Q: 1

The time taken,  $t$  minutes, to complete a puzzle was recorded for each of 150 students. These times are summarised in the table.

Time taken ( $t$ minutes)	$t \leq 25$	$t \leq 50$	$t \leq 75$	$t \leq 100$	$t \leq 150$	$t \leq 200$
Cumulative frequency	16	44	86	104	132	150

(a) Draw a cumulative frequency graph to illustrate the data.

[2]



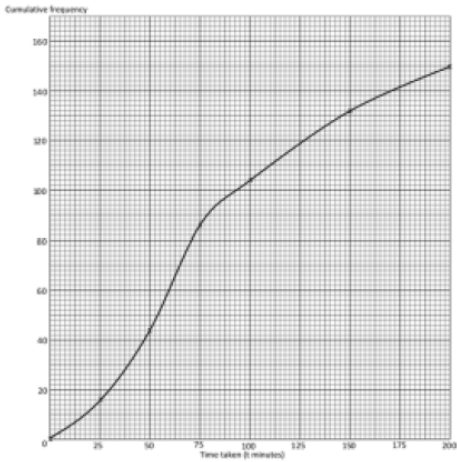
(b) Use your graph to estimate the 20th percentile of the data.

[1]

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Cumulative frequency (cf) graph	<b>M1</b>	At least 3 points plotted accurately at class upper end points (25,16) (50,44) (75,86) (100,104) (150, 132) (200, 150). Linear cf scale $0 \leq cf \leq 150$ and linear time scale $0 \leq \text{time}(\text{mins}) \leq 200$ with at least 3 values identified on each axis.
		<b>A1</b>	All points plotted correctly, curve drawn (within tolerance) and joined to (0,0). Axes labelled cumulative frequency (cf), time ( $t$ ) and minutes (min), or a suitable title.
		<b>2</b>	
(b)	Line from cumulative frequency = 30 to meet graph at $t$ is between 37.5 and 42	<b>B1 FT</b>	Not from wrong working. Must be an increasing cumulative frequency graph.
		<b>1</b>	

6. 9709\_s22\_qp\_53 Q: 2

Twenty children were asked to estimate the height of a particular tree. Their estimates, in metres, were as follows.

4.1	4.2	4.4	4.5	4.6	4.8	5.0	5.2	5.3	5.4
5.5	5.8	6.0	6.2	6.3	6.4	6.6	6.8	6.9	19.4

- (a) Find the mean of the estimated heights. [1]

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find the median of the estimated heights. [1]

.....

.....

.....

.....

- (c) Give a reason why the median is likely to be more suitable than the mean as a measure of the central tendency for this information. [1]

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	$\left[ \frac{123.4}{20} = \right] 6.17$	B1	Accept 6 m 17 cm, $\frac{1234}{200}$ .
		1	
(b)	$\frac{10\text{th} + 11\text{th}}{2} = \frac{5.4 + 5.5}{2} = 5.45 \text{ (m)}$	B1	Accept 5 m 45 cm.
		1	
(c)	The mean is unduly influenced by an extreme value, 19.4.	B1	Comment must be within context.
		1	

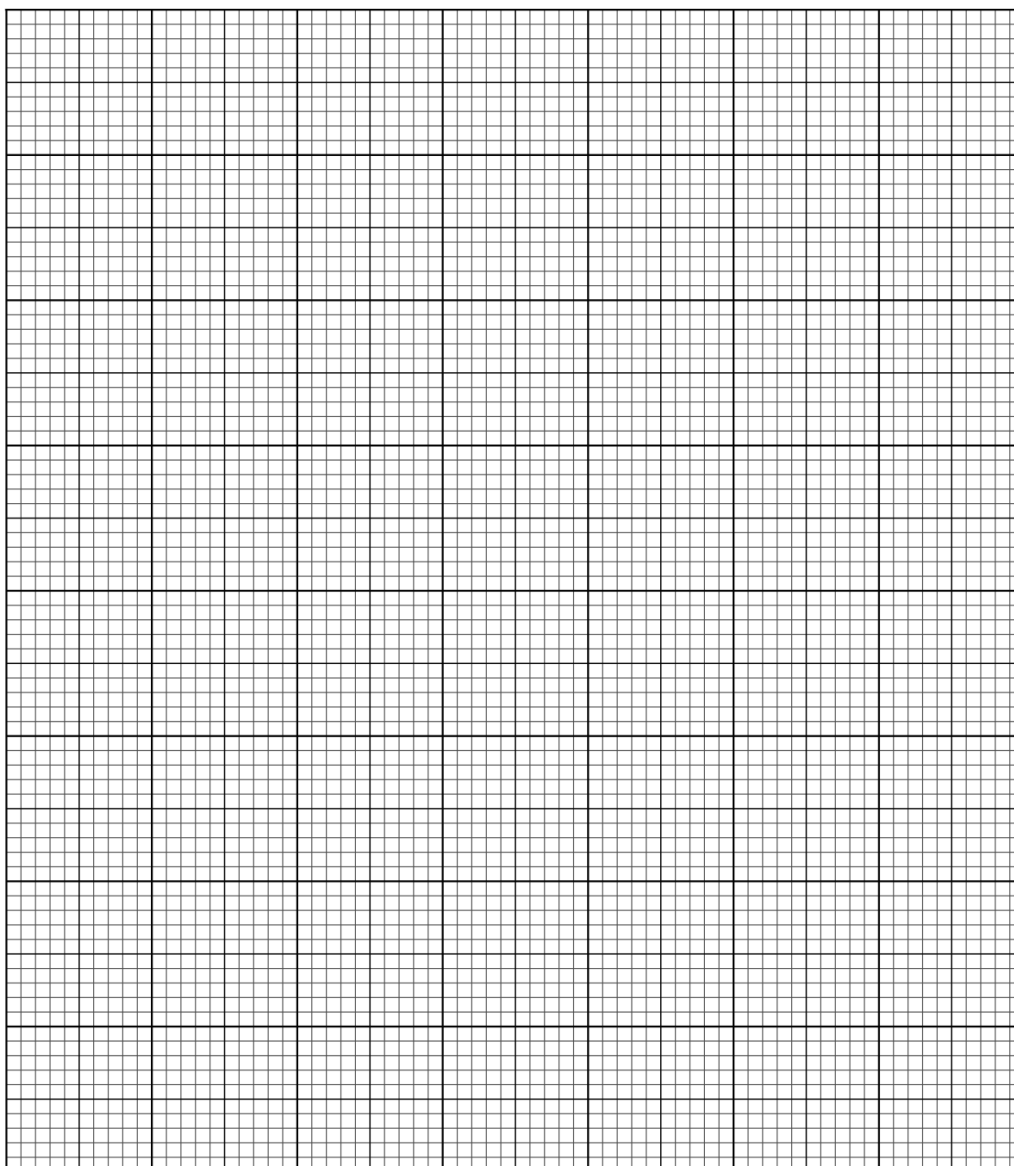
7. 9709\_m21\_qp\_52 Q: 5

A driver records the distance travelled in each of 150 journeys. These distances, correct to the nearest km, are summarised in the following table.

Distance (km)	0 – 4	5 – 10	11 – 20	21 – 30	31 – 40	41 – 60
Frequency	12	16	32	66	20	4

(a) Draw a cumulative frequency graph to illustrate the data.

[4]

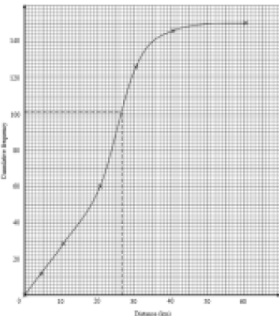


- Use your graph to estimate the value of  $d$ . [2]

[illegible]

- [illegible]

Answer:

Question	Answer	Marks	Guidance																					
(a)	<table><tr><td>Distance</td><td>0-4</td><td>5-10</td><td>11-20</td><td>21-30</td><td>31-40</td><td>41-60</td></tr><tr><td>Upper boundary</td><td>4.5</td><td>10.5</td><td>20.5</td><td>30.5</td><td>40.5</td><td>60.5</td></tr><tr><td>Cumulative frequency</td><td>12</td><td>28</td><td>60</td><td>126</td><td>146</td><td>150</td></tr></table>	Distance	0-4	5-10	11-20	21-30	31-40	41-60	Upper boundary	4.5	10.5	20.5	30.5	40.5	60.5	Cumulative frequency	12	28	60	126	146	150	<b>B1</b>	Correct cumulative frequencies seen (may be by table or plotted accurately on graph), condone 12 not stated.
	Distance	0-4	5-10	11-20	21-30	31-40	41-60																	
	Upper boundary	4.5	10.5	20.5	30.5	40.5	60.5																	
	Cumulative frequency	12	28	60	126	146	150																	
	<b>B1</b>	Axes labelled 'distance (or d) [in] km' from 0 to 60 <b>and</b> 'cumulative frequency' (or cf) from 0 to 150.																						
	<b>M1</b>	At least 5 points plotted at upper end points for $d$ (allow upper boundary $\pm 0.5$ ) with a linear scale for distance, condone 0 – 4 interval inaccurate, no scale break on axis. Not bar graph/histogram unless clear indication of upper end point only of each bar.																						
		<b>A1</b>	All plotted correctly at correct upper end points (4.5 etc.) with both scales linear ( $0 \leq d \leq 60$ , $0 \leq cf \leq 150$ ), curve drawn accurately joined to (0,0), cf line $>150$ , no daylight if $>150$ .																					
		<b>4</b>																						
(b)	70% of 150 = 105	<b>M1</b>	105 seen or implied by indication on grid.																					
	Approx. 27	<b>A1</b> <b>FT</b>	Strict FT <i>their</i> increasing cumulative frequency graph, use of graph must be seen. If no clear evidence of use of graph: <b>SC B1 FT</b> correct value from <i>their</i> increasing cumulative frequency graph.																					
		<b>2</b>																						

Question	Answer	Marks	Guidance
(c)	Midpoints: 2.25, 7.5, 15.5, 25.5, 35.5, 50.5	<b>B1</b>	At least 5 correct midpoints seen.
	$\text{Mean} = \frac{2.25 \times 12 + 7.5 \times 16 + 15.5 \times 32 + 25.5 \times 66 + 35.5 \times 20 + 50.5 \times 4}{150}$ $= \frac{27 + 120 + 496 + 1683 + 710 + 202}{150}$	<b>M1</b>	Using 6 midpoint attempts (e.g. $2.25 \pm 0.5$ ), condone one error not omission, multiplied by frequency, accept unevaluated, denominator either correct or <i>their</i> $\Sigma$ frequencies.
	$\left[ = \frac{3238}{150} \right] = 21.6, 21\frac{44}{75}$	<b>A1</b>	Evaluated, WWW, accept $21.5[866\dots]$ .
		<b>3</b>	

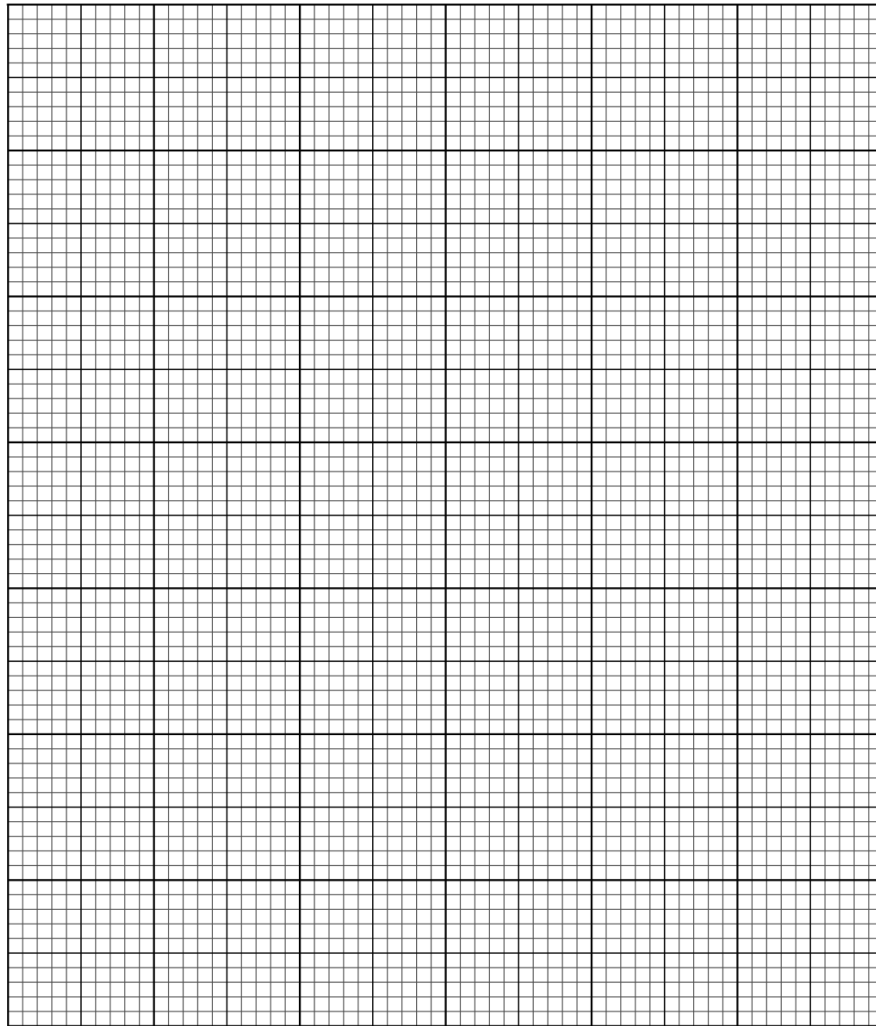
8. 9709\_s21\_qp\_51 Q: 5

The times taken by 200 players to solve a computer puzzle are summarised in the following table.

Time ( $t$ seconds)	$0 \leq t < 10$	$10 \leq t < 20$	$20 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 100$
Number of players	16	54	78	32	20

(a) Draw a histogram to represent this information.

[4]





- (b) Calculate an estimate of the mean time taken by these 200 players. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (c) Find the greatest possible value of the interquartile range of these times. [2]

.....

.....

.....

.....

.....

.....

.....

.....

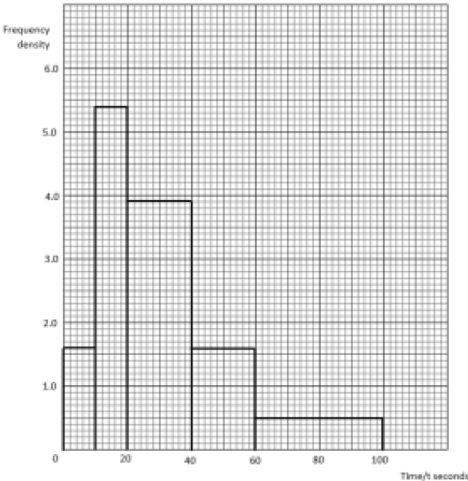
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance												
(a)	<table><tr><td>Class width</td><td>10</td><td>10</td><td>20</td><td>20</td><td>40</td></tr><tr><td>Frequency Density</td><td>1.6</td><td>5.4</td><td>3.9</td><td>1.6</td><td>0.5</td></tr></table>	Class width	10	10	20	20	40	Frequency Density	1.6	5.4	3.9	1.6	0.5	M1	At least 4 frequency densities calculated, accept unsimplified. May be read from graph using <i>their</i> scale, 3SF or correct
	Class width	10	10	20	20	40									
	Frequency Density	1.6	5.4	3.9	1.6	0.5									
		A1	All heights correct on graph												
		B1	Bar ends at 0, 10, 20 ..., etc. with a horizontal linear scale with at least 3 values indicated, $0 \leq \text{horizontal axis} \leq 100$												
B1		Axes labelled: Frequency density (fd), time (t) and seconds. Linear vertical scale, with at least 3 values indicated $0 \leq \text{vertical axis} \leq 5.4$													
	4														
Question	Answer	Marks	Guidance												
(b)	$\text{Mean} = \left[ \frac{16 \times 5 + 54 \times 15 + 78 \times 30 + 32 \times 50 + 20 \times 80}{200} \right]$ $= \frac{80 + 810 + 2340 + 1600 + 1600}{200}$	M1	Uses at least 4 midpoint attempts (e.g. $5 \pm 0.5$ ). Accept unsimplified expression, denominator either correct or <i>their</i> $\Sigma$ frequencies												
	$\left[ \frac{6430}{200} \right] = 32 \frac{3}{20} \text{ or } 32.15$	A1	Accept 32.2												
		2													
(c)	A value in correct UQ (40–60) – a value in correct LQ (10–20)	M1													
	Greatest possible value is $60 - 10 = 50$	A1	Condone 49.9												
		2													

9. 9709\_s21\_qp\_52 Q: 7

The heights, in cm, of the 11 basketball players in each of two clubs, the Amazons and the Giants, are shown below.

Amazons	205	198	181	182	190	215	201	178	202	196	184
Giants	175	182	184	187	189	192	193	195	195	195	204

- (a) State an advantage of using a stem-and-leaf diagram compared to a box-and-whisker plot to illustrate this information. [1]

.....

.....

.....

.....

.....

- (b) Represent the data by drawing a back-to-back stem-and-leaf diagram with Amazons on the left-hand side of the diagram. [4]

- (c) Find the interquartile range of the heights of the players in the Amazons. [2]

.....

.....

.....

.....

.....

.....

.....

.....

Four new players join the Amazons. The mean height of the 15 players in the Amazons is now 191.2 cm. The heights of three of the new players are 180 cm, 185 cm and 190 cm.

- (d) Find the height of the fourth new player. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

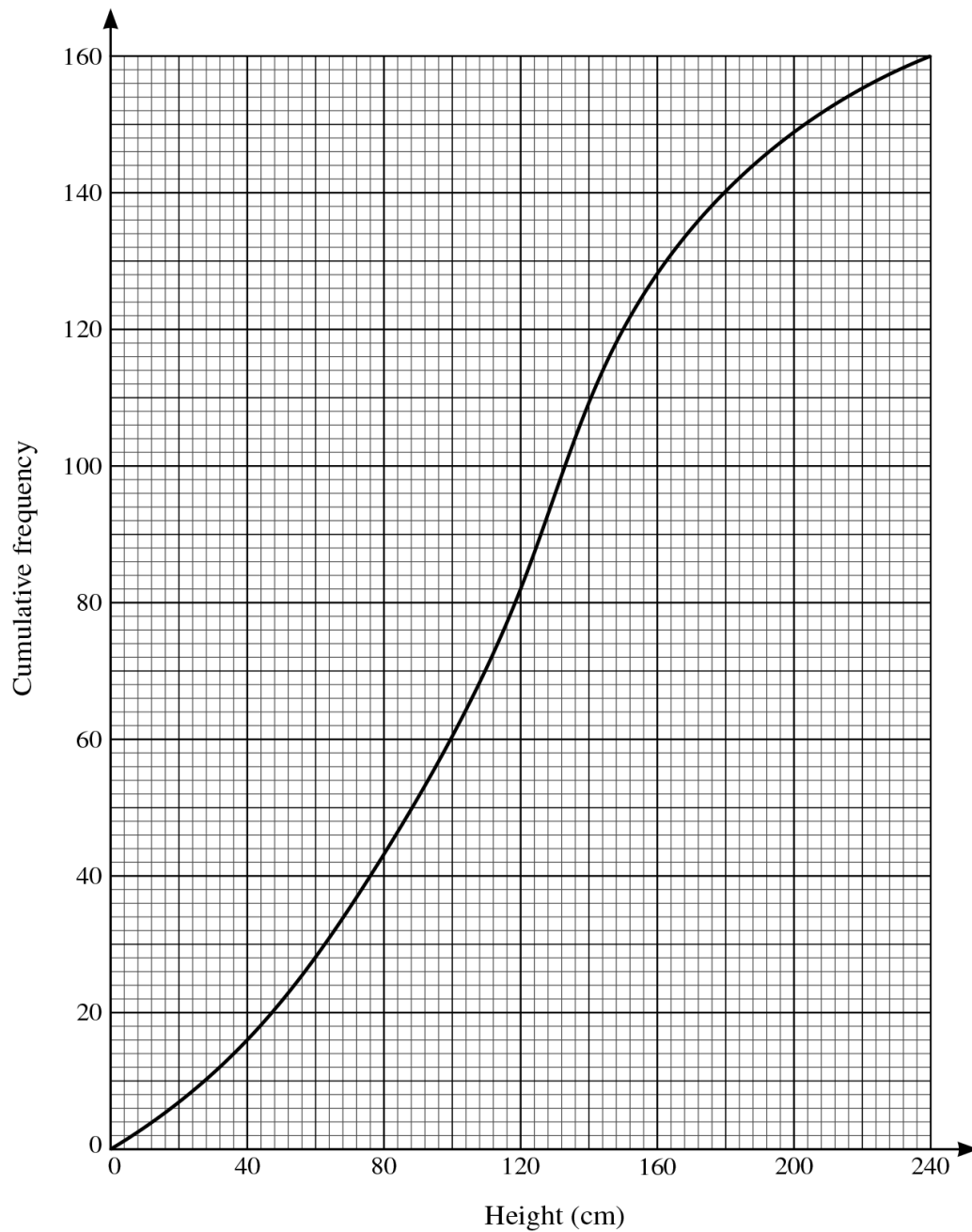
Answer:

Question	Answer	Marks	Guidance																		
(a)	Includes all data	B1	Reference to <i>either</i> including all/raw data or further statistical processes are possible that cannot be found using data from box-and-whisker, eg frequency, mean, mode or standard deviation <b>not</b> only median, IQR, range or spread which can be found from both.																		
		1																			
(b)	<table><tr><td>Amazons</td><td></td><td>Giants</td></tr><tr><td>8</td><td>17</td><td>5</td></tr><tr><td>4 2 1</td><td>18</td><td>2 4 7 9</td></tr><tr><td>8 6 0</td><td>19</td><td>2 3 5 5 5</td></tr><tr><td>5 2 1</td><td>20</td><td>4</td></tr><tr><td>5</td><td>21</td><td></td></tr></table> <p>Key: 1 18 2 means 181 cm for Amazons and 182 cm for Giants</p>	Amazons		Giants	8	17	5	4 2 1	18	2 4 7 9	8 6 0	19	2 3 5 5 5	5 2 1	20	4	5	21		B1 B1 B1 B1	Correct stem can be upside down, ignore extra values Correct Amazons labelled on left, leaves in order from right to left and lined up vertically (less than halfway to next column), no commas or other punctuation. Correct Giants labelled on same diagram, leaves in order and lined up vertically (less than halfway to next column), no commas or other punctuation. Correct single key for their diagram, need both teams identified and ‘cm’ stated at least once here or in leaf headings or title. <b>SC</b> for if 2 separate diagrams drawn, award <b>SCB1</b> if both keys meet these criteria (Max B1, B0, B0, B1)
Amazons		Giants																			
8	17	5																			
4 2 1	18	2 4 7 9																			
8 6 0	19	2 3 5 5 5																			
5 2 1	20	4																			
5	21																				
		4																			
(c)	[UQ = 202 (cm), LQ = 182 (cm)] [IQR =] 202 – 182 = 20 (cm)	M1 A1	201 ≤ UQ ≤ 205 – 181 ≤ LQ ≤ 184 WWW																		
		2																			

Question	Answer	Marks	Guidance
(d)	[Σ <sub>11</sub> = 2132 Σ <sub>15</sub> = 191.2×15 = 2868 ]	B1	Both Σ <sub>11</sub> and Σ <sub>15</sub> found. Accept unevaluated.
	<i>their</i> 2868 = <i>their</i> 2132 + (180 + 185 + 190) + <i>h</i>	M1	Forming an equation for the height using <i>their</i> Σ <sub>11</sub> and Σ <sub>15</sub> .
	181 (cm)	A1	
Alternative method for Question 7(d)			
	[Σ <sub>15</sub> = 191.2×15 = 2868 Σ <sub>15</sub> = 2687 + <i>h</i> ]	B1	Σ <sub>15</sub> found using the mean and raw data methods. Accept unevaluated.
	<i>their</i> 2868 = <i>their</i> 2687 + <i>h</i>	M1	Forming an equation for the height using <i>their</i> Σ <sub>15</sub> expressions.
	181 (cm)	A1	
Alternative method for Question 7(d)			
	[Σ <sub>15</sub> = 2687 + <i>h</i> Σ <sub>15</sub> 15 = 191.2 ]	B1	Σ <sub>15</sub> found using raw data method and statement on calculating new mean. Accept unevaluated.
	<i>their</i> 2687 + <i>h</i> 15 = 191.2	M1	Forming an equation for the height using <i>their</i> Σ <sub>15</sub> expressions
	181 (cm)	A1	
		3	<b>N.B.</b> All methods can be presented as a logical numerical argument which can be condoned if clear.

10. 9709\_s21\_qp\_53 Q: 1

The heights in cm of 160 sunflower plants were measured. The results are summarised on the following cumulative frequency curve.



- (a) Use the graph to estimate the number of plants with heights less than 100 cm.

[1]

.....

.....

.....

.....

- (b) Use the graph to estimate the 65th percentile of the distribution. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (c) Use the graph to estimate the interquartile range of the heights of these plants. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	60	<b>B1</b>	Accept 60 or 61. No decimals
		<b>1</b>	
(b)	65% of 160 = 104	<b>M1</b>	$0.65 \times 160 (=104)$ seen unsimplified or implied by use on graph
	136 (cm)	<b>A1</b>	Use of graph must be seen. <b>SCB1</b> correct value (136 only) if neither 104 nor use of graph are evident
		<b>2</b>	
(c)	UQ: 150 LQ: 76 IQR = $150 - 76 = 74$ [cm]	<b>M1</b>	$UQ - LQ$ ; $148 \leq UQ \leq 152$ ; $74 \leq LQ \leq 78$ .
		<b>A1</b>	Must be from 150 - 76
		<b>2</b>	



11. 9709\_s21\_qp\_53 Q: 3

A sports club has a volleyball team and a hockey team. The heights of the 6 members of the volleyball team are summarised by  $\Sigma x = 1050$  and  $\Sigma x^2 = 193\,700$ , where  $x$  is the height of a member in cm. The heights of the 11 members of the hockey team are summarised by  $\Sigma y = 1991$  and  $\Sigma y^2 = 366\,400$ , where  $y$  is the height of a member in cm.

- (a) Find the mean height of all 17 members of the club. [2]

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find the standard deviation of the heights of all 17 members of the club. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Mean height = $\frac{\Sigma x + \Sigma y}{6+11} = \frac{1050+1991}{6+11} = \frac{3041}{17}$	M1	Use of appropriate formula with values substituted, accept unsimplified.
	178.9	A1	Allow 178.88, $178\frac{15}{17}$ , 179
		2	
Question	Answer	Marks	Guidance
(b)	$\frac{\Sigma x^2 + \Sigma y^2}{6+11} = \frac{193700+366400}{6+11}$	M1	Use of appropriate formula with values substituted, accept unsimplified.
	$Sd^2 = \frac{560100}{17} - their 178.88^2 [= 948.289]$	M1	Appropriate variance formula using <i>their</i> mean <sup>2</sup> , accept unsimplified expression.
	Standard deviation = 30.8	A1	Accept 30.7
		3	

12. 9709\_w21\_qp\_51 Q: 2

A summary of 40 values of  $x$  gives the following information:

$$\Sigma(x - k) = 520, \quad \Sigma(x - k)^2 = 9640,$$

where  $k$  is a constant.

- (a) Given that the mean of these 40 values of  $x$  is 34, find the value of  $k$ . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- (b) Find the variance of these 40 values of  $x$ . [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	$\left[ \frac{\sum x}{40} - k = \frac{\sum (x-k)}{40} \right]$	<b>M1</b>	Forms an equation involving $\sum x$ , $\sum (x-k)$ and $k$ . Accept at a numeric stage with $k$ .
	$\frac{40 \times 34}{40} - k = \frac{520}{40}$	<b>A1</b>	Evaluated.
	$k [= 34 - 13] = 21$	<b>2</b>	
Question	Answer	Marks	Guidance
(b)	$\text{Var} = \left[ \frac{\sum (x-k)^2}{40} - \left( \frac{\sum (x-k)}{40} \right)^2 \right] = \frac{9640}{40} - \left( \frac{520}{40} \right)^2 = [241 - 13^2 =]$	<b>M1</b>	Values substituted into an appropriate variance formula, accept unsimplified.
	72	<b>A1</b>	
		<b>2</b>	

13. 9709\_w21\_qp\_51 Q: 6

The weights, in kg, of 15 rugby players in the Rebels club and 15 soccer players in the Sharks club are shown below.

Rebels	75	78	79	80	82	82	83	84	85	86	89	93	95	99	102
Sharks	66	68	71	72	74	75	75	76	78	83	83	84	85	86	92

- (a) Represent the data by drawing a back-to-back stem-and-leaf diagram with Rebels on the left-hand side of the diagram. [4]

- (b) Find the median and the interquartile range for the Rebels. [3]

.....

.....

.....

.....

.....

.....

.....

.....

A box plot for the 'Sharks' category. The x-axis is labeled 'Weight (kg)' and ranges from 60 to 110 with major grid lines every 10 units and minor grid lines every 2 units. The box plot shows a minimum at approximately 66 kg, a first quartile (Q1) at approximately 72 kg, a median at approximately 76 kg, a third quartile (Q3) at approximately 84 kg, and a maximum at approximately 92 kg.

- [illegible]

Answer:

Question	Answer	Marks	Guidance																																																																																																						
(a)	<table><tr><td colspan="8">Rebels</td><td>6</td><td colspan="8">Sharks</td></tr><tr><td colspan="8"></td><td>6</td><td colspan="8">8</td></tr><tr><td colspan="8">9 8 5</td><td>7</td><td colspan="8">1 2 4 5 5 6 8</td></tr><tr><td colspan="8">9 6 5 4 3 2 2 0</td><td>8</td><td colspan="8">3 3 4 5 6</td></tr><tr><td colspan="8">9 5 3</td><td>9</td><td colspan="8">2</td></tr><tr><td colspan="8">2</td><td>10</td><td colspan="8"></td></tr></table> <p>Key: 8   7   2 means 78 kg for Rebels and 72 kg for Sharks</p>	Rebels								6	Sharks																6	8								9 8 5								7	1 2 4 5 5 6 8								9 6 5 4 3 2 2 0								8	3 3 4 5 6								9 5 3								9	2								2								10									B1	Correct stem, ignore extra values (not in reverse).
		Rebels								6	Sharks																																																																																														
										6	8																																																																																														
		9 8 5								7	1 2 4 5 5 6 8																																																																																														
	9 6 5 4 3 2 2 0								8	3 3 4 5 6																																																																																															
9 5 3								9	2																																																																																																
2								10																																																																																																	
B1	Correct Rebels labelled on left, leaves in order from right to left and lined up vertically, no commas.																																																																																																								
B1	Correct Sharks labelled on same diagram, leaves in order and lined up vertically, no commas.																																																																																																								
B1	Correct key for their diagram, need both teams identified and 'kg' stated at least once here or in leaf headings or title.  SC If 2 separate diagrams drawn, SC B1 if both keys meet these criteria.																																																																																																								
		4																																																																																																							

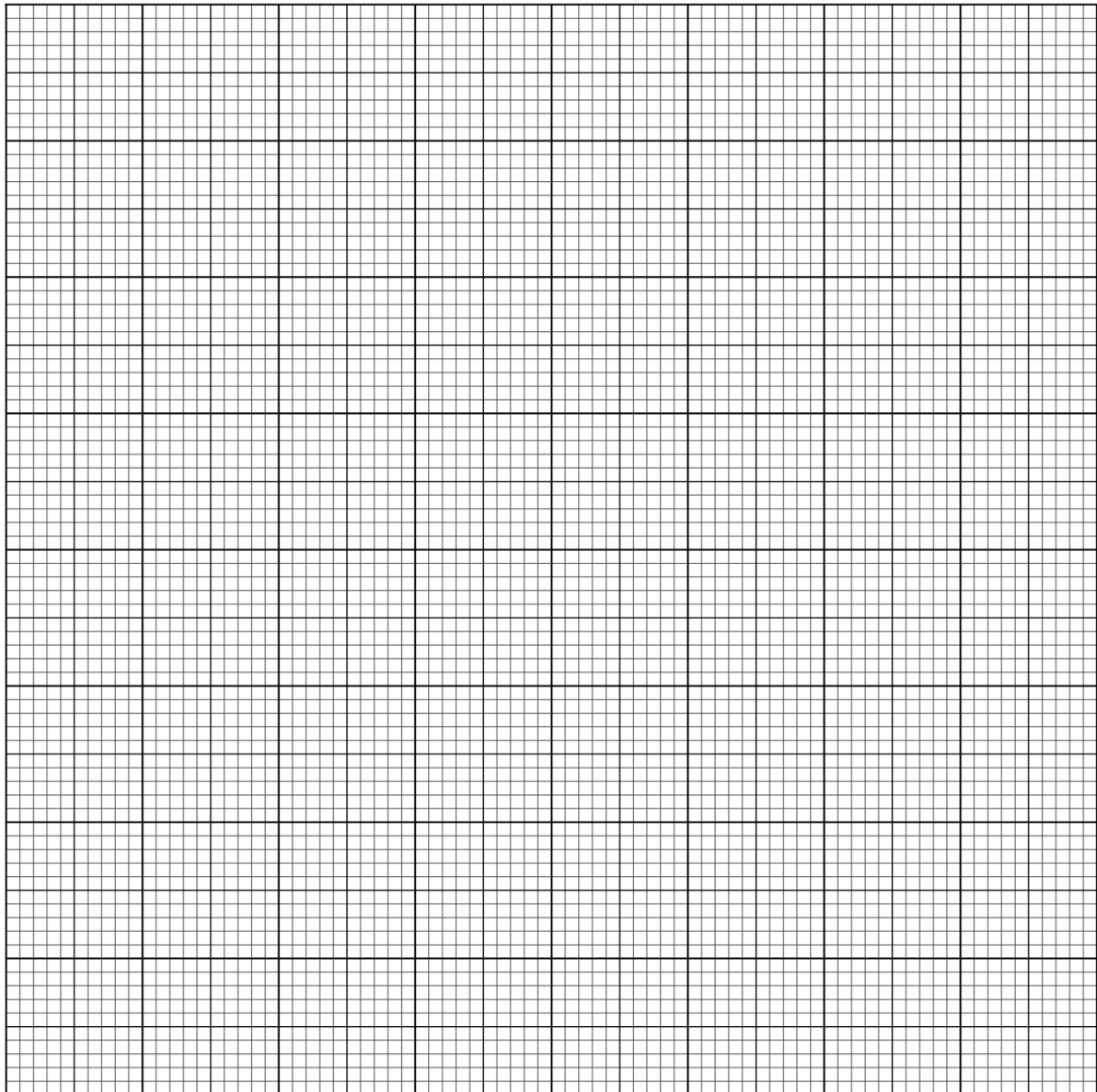
Question	Answer	Marks	Guidance
(b)	Median = 84 (kg)	B1	
	[UQ = 93, LQ = 80] 93 – 80	M1	$95 \leq UQ \leq 89 - 79 \leq LQ \leq 82$
	[IQR =] 13 (kg)	A1	WWW
		3	
(c)	Box and whisker with end points 75 and 102	B1	Whiskers drawn to correct end points not through box, not joining at top or bottom of box.
	Median and quartiles plotted as found in (b)	B1 FT	Quartiles and median plotted as box graph.
		2	
(d)	e.g. Average weight of Rebels is higher than average weight of Sharks	B1	Acceptable answers refer to: Range, skew, central tendency within context. E.g. range of Rebels is greater B0. Range of weights of the rebels is greater B1. Simple value comparison insufficient.
		1	

14. 9709\_w21\_qp\_52 Q: 7

The distances,  $x$  m, travelled to school by 140 children were recorded. The results are summarised in the table below.

Distance, $x$ m	$x \leq 200$	$x \leq 300$	$x \leq 500$	$x \leq 900$	$x \leq 1200$	$x \leq 1600$
Cumulative frequency	16	46	88	122	134	140

- (a) On the grid, draw a cumulative frequency graph to represent these results. [2]





- (b)** Use your graph to estimate the interquartile range of the distances. [2]

[illegible]

- (c) Calculate estimates of the mean and standard deviation of the distances. [6]

This image shows a full page of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page, providing a template for handwriting practice or general writing. There are no margins, text, or other markings on the page.

Answer:

Question	Answer	Marks	Guidance														
(a)	Cumulative frequency graph drawn	<b>B1</b>	Axes labelled ‘cumulative frequency’ (or cf) from 0 to at least 140 and ‘distance (or d) [in] m’ from 0 to at least 1600, linear scales with at least 3 values stated.														
		<b>B1</b>	All plotted correctly at correct upper end points (200 etc.) <b>curve</b> drawn accurately joined to (0, 0) (straight line segments B0) but no daylight above 140. Cf scale no less than 2 cm = 20 children .														
		<b>2</b>															
Question	Answer	Marks	Guidance														
(b)	[UQ at 75% of 140 = 105, LQ at 25% of 140 = 35] [IQR:] 700 – 260	<b>M1</b>	Accept $660 \leq UQ \leq 720 - 240 \leq LQ \leq 290$ . If values are outside our range, FT providing scales linear and increasing cf drawn.														
	440	<b>A1</b>	Accept correct evaluation of $660 \leq \textit{their} UQ \leq 720 - 240 \leq \textit{their} LQ \leq 290$ with clear indication that graph has been used for at least one of 105 or 35.														
		<b>2</b>															
Question	Answer	Marks	Guidance														
(c)	[Mean =] $\frac{16 \times 100 + 30 \times 250 + 42 \times 400 + 34 \times 700 + 12 \times 1050 + 6 \times 1400}{140}$	<b>B1</b>	<table border="1"><tr><td>Frequencies</td><td>16</td><td>30</td><td>42</td><td>34</td><td>12</td><td>6</td></tr><tr><td>Mid-points</td><td>100</td><td>250</td><td>400</td><td>700</td><td>1050</td><td>1400</td></tr></table> 5 or 6 correct frequency values seen.	Frequencies	16	30	42	34	12	6	Mid-points	100	250	400	700	1050	1400
	Frequencies	16	30	42	34	12	6										
	Mid-points	100	250	400	700	1050	1400										
		<b>B1</b>	5 or 6 correct midpoint values seen.														
		<b>M1</b>	Values substituted into mean formula using <i>their</i> midpoints which must be in the class – condone 1 data error. Accept $\frac{1600 + 7500 + 16\,800 + 23\,800 + 12\,600 + 8400}{140}$ or $\frac{70\,700}{140}$ . Condone $\frac{70\,770}{140}$ for <b>M1</b> .														
	505	<b>A1</b>	WWW														
Variance = $\frac{16 \times 100^2 + 30 \times 250^2 + 42 \times 400^2 + 34 \times 700^2 + 12 \times 1050^2 + 6 \times 1400^2}{140}$ – 505 <sup>2</sup>	<b>M1</b>	Values substituted into variance formula using ( <i>their</i> mean) <sup>2</sup> and <i>their</i> midpoints and <i>their</i> frequencies (including for denominator). Accept unsimplified. Condone 1 data error. Accept: $\left[ \frac{160\,000 + 1\,875\,000 + 6\,720\,000 + 16\,660\,000 + 13\,230\,000 + 11\,760\,000}{140} \right]$ or $\frac{50\,405\,000}{140}$ or 360 035.7143] – [505 <sup>2</sup> or 255 025] If formula stated accept 105 010 or 105 011 WWW.															
S.d. = $\left[ \sqrt{105\,010.7} \right] = 324$	<b>A1</b>	WWW															
		<b>6</b>															

15. 9709\_w21\_qp\_53 Q: 2

Lakeview and Riverside are two schools. The pupils at both schools took part in a competition to see how far they could throw a ball. The distances thrown, to the nearest metre, by 11 pupils from each school are shown in the following table.

Lakeview	10	14	19	22	26	27	28	30	32	33	41
Riverside	23	36	21	18	37	25	18	20	24	30	25

- (a) Draw a back-to-back stem-and-leaf diagram to represent this information, with Lakeview on the left-hand side. [4]

- (b) Find the interquartile range of the distances thrown by the 11 pupils at Lakeview school. [2]

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer										Marks	Guidance		
(a)	Lakeview				Riverside							B1	Correct stem, ignore extra values.	
	9	4	0	1	8	8							B1	Correct Lakeview labelled on left, leaves in order from right to left and lined up vertically, no commas.
	8	7	6	2	2	0	1	3	4	5	5			
	3	2	0	3	0	6	7						B1	Correct Riverside labelled on same diagram, leaves in order and lined up vertically, no commas.
			1	4									B1	Correct key for their diagram, need both teams identified and ‘m’ stated at least once here or in leaf headings or title.  SC If 2 separate diagrams drawn: SC B1 if both keys meet these criteria.
Key: 6 2 3 means 26m for Lakeview and 23m for Riverside														
											4			
(b)	UQ = 32, LQ = 19										M1	(30 ≤ UQ ≤ 33) – (14 ≤ LQ ≤ 22)		
	IQR = 32 – 19 = 13										A1	WWW		
											2			