Topical Past Paper Questions Workbook

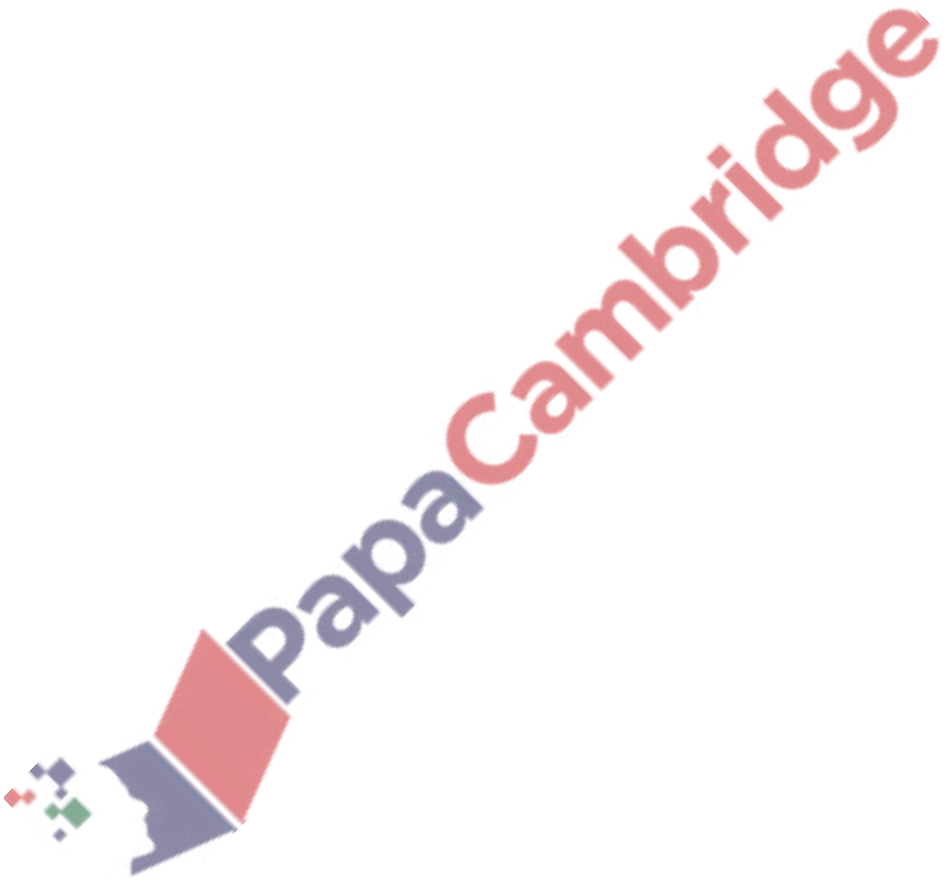
AS & A Level Mathematics (9709) Paper 5 [Probability & Statistics 1]

Exam Series: May 2015 – May 2022

Format Type B:
Each question is followed by its answer scheme

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:

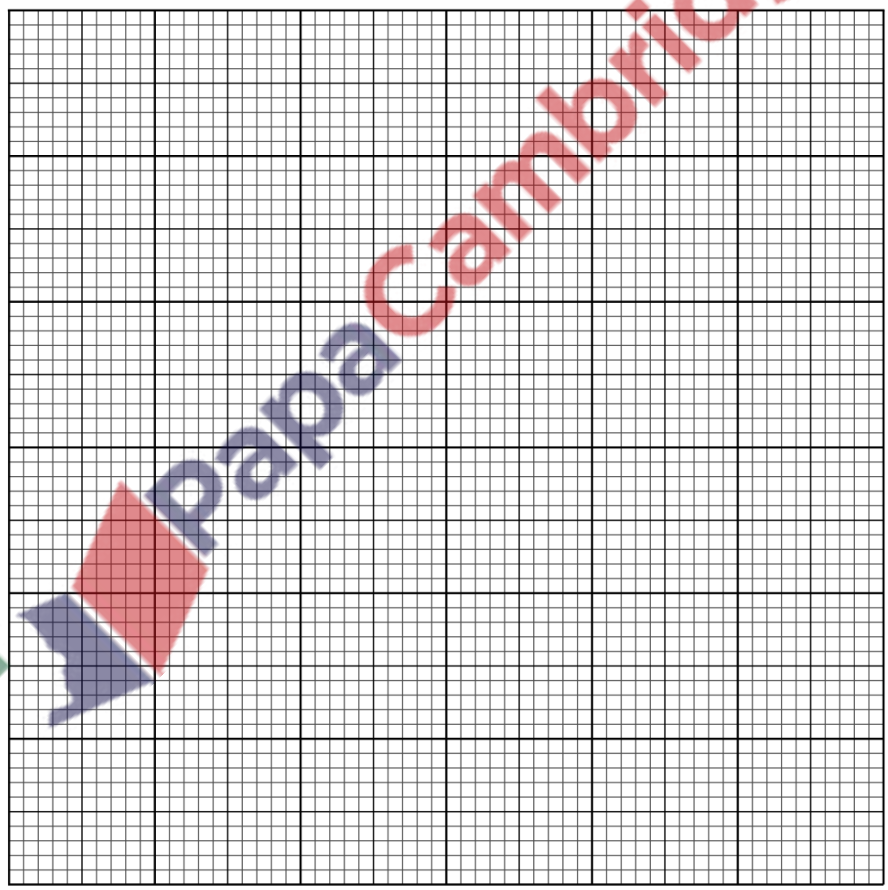
<table>
<tr><th>Question</th><th>Answer</th><th>Marks</th><th>Guidance</th></tr>
<tr><td rowspan="4">(a)</td><td rowspan="3"><table><tr><td>Class Width</td><td>30</td><td>15</td><td>20</td><td>10</td><td>25</td></tr><tr><td>Frequency Density</td><td>0.7</td><td>2</td><td>3.4</td><td>8.6</td><td>1.8</td></tr></table></td><td>M1</td><td>At least 4 frequency densities calculated</td></tr>
<tr><td>A1</td><td>All heights correct on graph</td></tr>
<tr><td>B1</td><td>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 ⩽ time axis ⩽ 100.5, linear scale with at least 3 values indicated.</td></tr>
<tr><td>B1</td><td>Axes labelled: Frequency density (fd), time (t) and mins (or appropriate title). Linear fd scale, with at least 3 values indicated 0 ⩽ fd axis ⩽ 8·6</td></tr>
<tr><td></td><td>4</td><td></td></tr>
<tr><td rowspan="2">(b)</td><td>66 – 75</td><td>B1</td><td>Condone 65.5 – 75.5</td></tr>
<tr><td></td><td>1</td><td></td></tr>
<tr><td rowspan="2">(c)</td><td>Distribution is not symmetrical</td><td>B1</td><td>Or skewed, ignore nature of skew</td></tr>
<tr><td></td><td>1</td><td></td></tr>
</table>

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)	$0 \leqslant t < 20$	$20 \leqslant t < 30$	$30 \leqslant t < 40$	$40 \leqslant t < 60$	$60 \leqslant t < 90$
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]

.....

.....

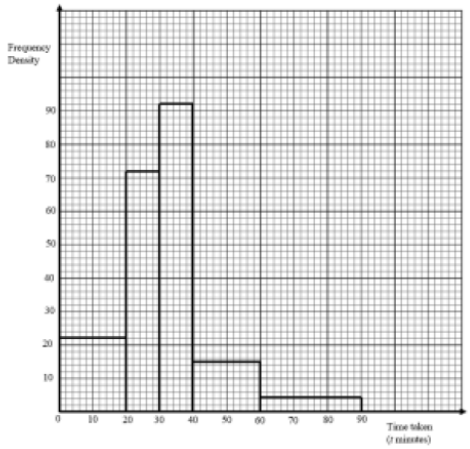
.....

.....

.....

.....

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> 	M1	At least 4 frequency densities calculated (Frequency ÷ class width, e.g. $\frac{440}{20}\left(condone\ \frac{440}{19.5}, \frac{440}{20.5}\right)$ Accept unsimplified, may be read from graph using *their* scale
		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 ⩽ horizontal scale ⩽ 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 ⩽ vertical axes ⩽ 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>	B1	At least 4 correct midpoints seen
	[Mean = 31.44 given] [Variance = $\frac{440\times10^2+720\times25^2+920\times35^2+300\times50^2+120\times75^2}{2500}-31.44^2$] $=\frac{44000+450000+1127000+750000+675000}{2500}-31.44^2$ [$=\frac{3046000}{2500}-31.44^2=229.9264$] Or 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 (− mean^2 included with *their* midpoints (not upper bound, lower bound, class width, frequency density, frequency or cumulative frequency) and *their* $\sum f$ 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]

Answer:

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

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 *A*							Company *B*				
					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]

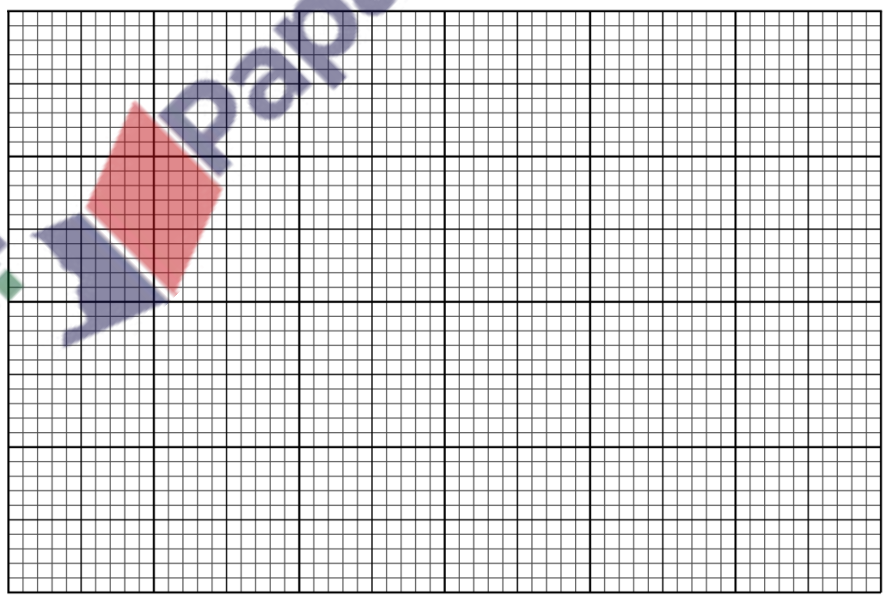
.....

.....

.....

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]

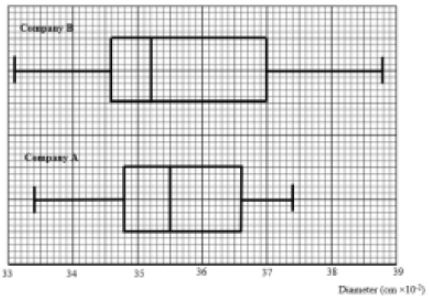


(c) Make one comparison between the diameters of the pipes produced by companies *A* and *B*. [1]

.....

.....

Answer:

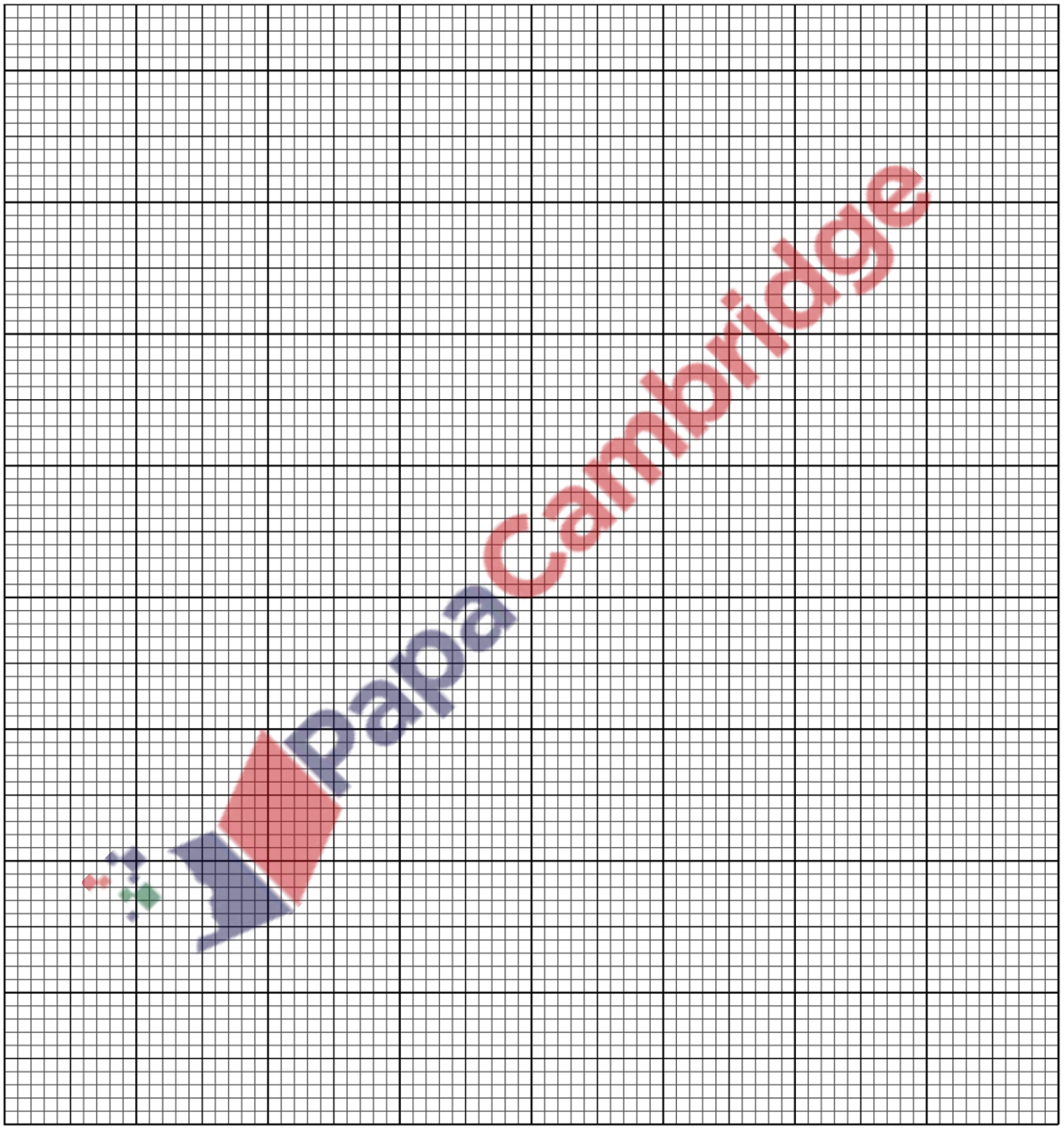
Question	Answer	Marks	Guidance
(a)	Median = 0.355	**B1**	Identified condone Q2.
	[IQR =] 0.366 – 0.348	**M1**	0.365 ⩽ UQ ⩽ 0.369 – 0.343 ⩽ LQ ⩽ 0.349. Subtraction may be implied by answer.
	0.018	**A1**	If 0/3 scored **SC B1** for figs Median = 355 IQR = 18.
		3	
(b)	Box-and-whisker plot on provided grid 	**B1**	All 5 key values for *B* plotted accurately in standard format using *their* scale. Labelled *B*. Check accuracy in the middle of vertical line.
		B1 FT	All 5 key values for *A*, FT from part **3(a)**, plotted in standard format accurately using *their* scale. Labelled *A*. Check accuracy in the middle of vertical line.
		B1	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 (*d* etc) and cm. Accept as a title.
		3	If both plots attempted and plot(s) not labelled, **SC B1** for at least 1 fully correct set of values plotted.
(c)	A comparison in context	**B1**	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.
		1	

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)	$t \leqslant 25$	$t \leqslant 50$	$t \leqslant 75$	$t \leqslant 100$	$t \leqslant 150$	$t \leqslant 200$
Cumulative frequency	16	44	86	104	132	150

(a) Draw a cumulative frequency graph to illustrate the data. [2]

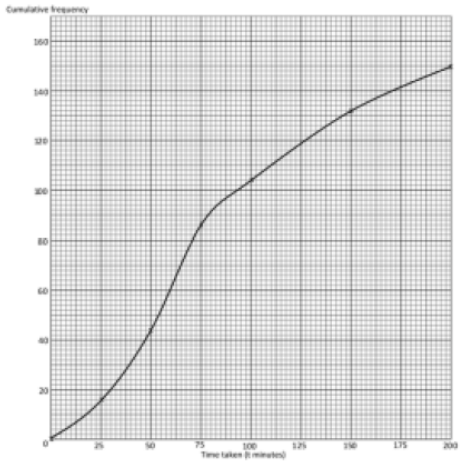


(b) Use your graph to estimate the 20th percentile of the data. [1]

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Cumulative frequency (cf) graph	M1	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 $0 \leqslant cf \leqslant 150$ and linear time scale $0 \leqslant$ time(mins) $\leqslant 200$ with at least 3 values identified on each axis.
		A1	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.
		2	
(b)	Line from cumulative frequency = 30 to meet graph at t is between 37.5 and 42	B1 FT	Not from wrong working. Must be an increasing cumulative frequency graph.
		1	

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:

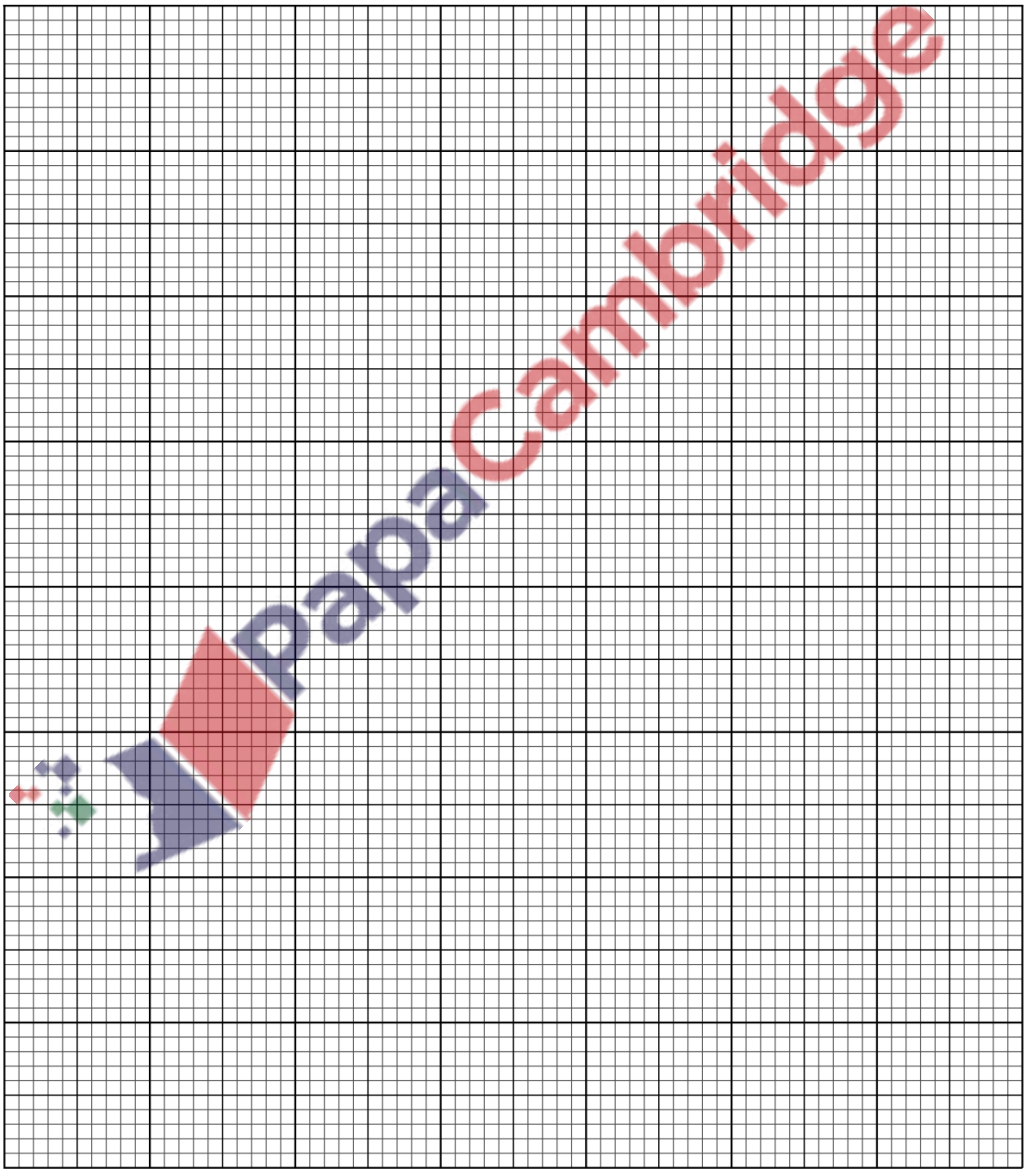
Question	Answer	Marks	Guidance
(a)	$\left[\frac{123.4}{20} = \right] 6.17$	B1	Accept 6 m 17 cm, $\frac{1234}{200}$.
		1	
(b)	$\frac{10\text{th} + 11\text{th}}{2} = \frac{5.4 + 5.5}{2} = 5.45$ (m)	B1	Accept 5 m 45 cm.
		1	
(c)	The mean is unduly influenced by an extreme value, 19.4.	B1	Comment must be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



(b) For 30% of these journeys the distance travelled is d km or more.

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

.....

.....

.....

.....

.....

.....

.....

(c) Calculate an estimate of the mean distance travelled for the 150 journeys. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

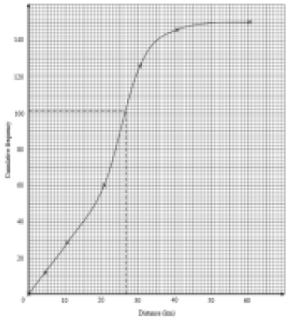
.....

.....

.....

.....

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>	**B1**	Correct cumulative frequencies seen (may be by table or plotted accurately on graph), condone 12 not stated.
		**B1**	Axes labelled 'distance (or d) [in] km' from 0 to 60 **and** 'cumulative frequency' (or cf) from 0 to 150.
		M1	At least 5 points plotted at upper end points for *d* (allow upper boundary ±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.
		A1	All plotted correctly at correct upper end points (4.5 etc.) with both scales linear ($0 \leqslant d \leqslant 60$, $0 \leqslant$ cf $\leqslant 150$), curve drawn accurately joined to (0,0), cf line>150, no daylight if >150.
		4	
(b)	70% of 150 = 105	**M1**	105 seen or implied by indication on grid.
	Approx. 27	**A1 FT**	Strict FT *their* increasing cumulative frequency graph, use of graph must be seen. If no clear evidence of use of graph: **SC B1 FT** correct value from *their* increasing cumulative frequency graph.
		2	

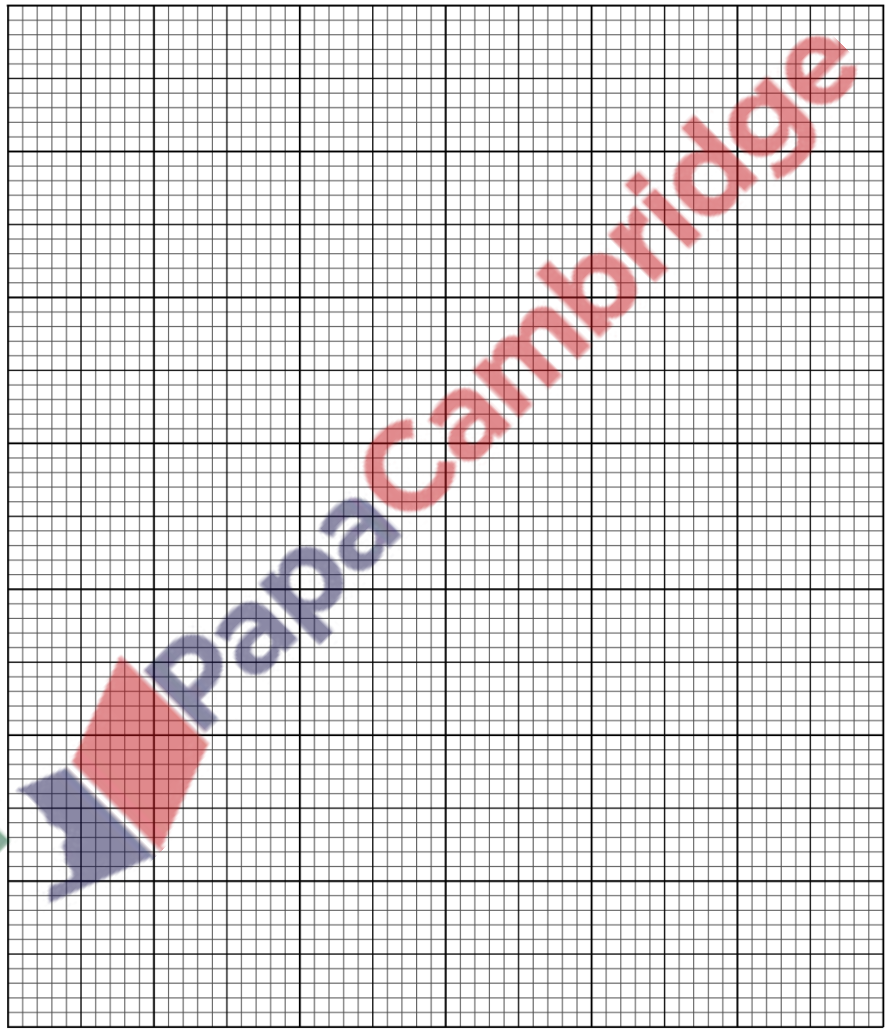
Question	Answer	Marks	Guidance
(c)	Midpoints: 2.25, 7.5, 15.5, 25.5, 35.5, 50.5	**B1**	At least 5 correct midpoints seen.
	$\text{Mean} = \frac{2.25\times12+7.5\times16+15.5\times32+25.5\times66+35.5\times20+50.5\times4}{150}$ $= \frac{27+120+496+1683+710+202}{150}$	**M1**	Using 6 midpoint attempts (e.g. 2·25 ±0·5), condone one error not omission, multiplied by frequency, accept unevaluated, denominator either correct or *their* Σ frequencies.
	$\left[= \frac{3238}{150}\right] = 21.6,\ 21\frac{44}{75}$	**A1**	Evaluated, WWW, accept 21·5[866…].
		3	

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)	$0 \leqslant t < 10$	$10 \leqslant t < 20$	$20 \leqslant t < 40$	$40 \leqslant t < 60$	$60 \leqslant t < 100$
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]

.....

.....

.....

.....

.....

.....

.....

.....

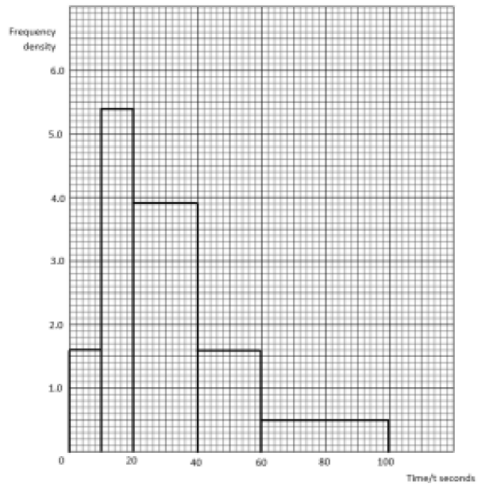
.....

.....

.....

.....

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> 	M1	At least 4 frequency densities calculated, accept unsimplified. May be read from graph using *their* scale, 3SF or correct
		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, $0 \leqslant$ horizontal axis $\leqslant 100$
		B1	Axes labelled: Frequency density (fd), time (t) and seconds. Linear vertical scale, with at least 3 values indicated $0 \leqslant$ vertical axis $\leqslant 5.4$
		4	

Question	Answer	Marks	Guidance
(b)	$\text{Mean} = \left[\frac{16\times5+54\times15+78\times30+32\times50+20\times80}{200}\right]$ $= \frac{80+810+2340+1600+1600}{200}$	M1	Uses at least 4 midpoint attempts (e.g. 5 ± 0.5). Accept unsimplified expression, denominator either correct or *their* Σfrequencies
	$\left[\frac{6430}{200} =\right] 32\frac{3}{20}$ 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]

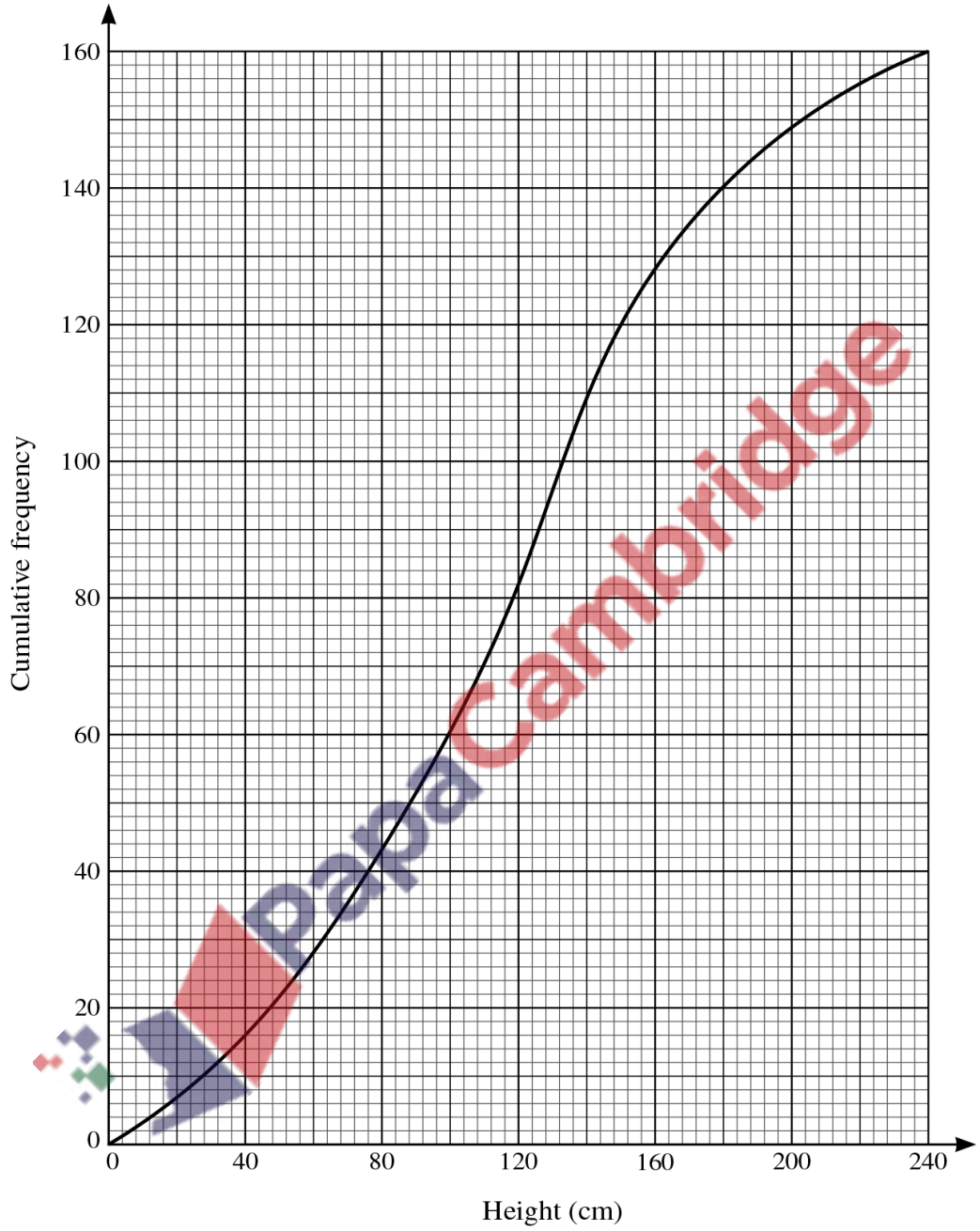
Answer:

Question	Answer	Marks	Guidance
(a)	Includes all data	B1	Reference to *either* 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 **not** only median, IQR, range or spread which can be found from both.
		1	
(b)	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 Key: 1|18|2 means 181 cm for Amazons and 182 cm for Giants	B1	Correct stem can be upside down, ignore extra values
		B1	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.
		B1	Correct Giants labelled on same diagram, leaves in order and lined up vertically (less than halfway to next column), no commas or other punctuation.
		B1	Correct single key for their diagram, need both teams identified and 'cm' stated at least once here or in leaf headings or title. **SC** for if 2 separate diagrams drawn, award **SCB1** if both keys meet these criteria (Max B1, B0, B0, B1)
		4	
(c)	[UQ = 202 (cm), LQ = 182 (cm)] [IQR =] 202 – 182 = 20 (cm)	M1	$201 \leqslant UQ \leqslant 205 - 181 \leqslant LQ \leqslant 184$
		A1	WWW
		2	

Question	Answer	Marks	Guidance
(d)	$[\Sigma_{11} = 2132$ $\Sigma_{15} = 191.2 \times 15 = 2868\ ]$	B1	Both Σ_{11} and Σ_{15} found. Accept unevaluated.
	their 2868 = *their* 2132 + (180 + 185 + 190) + h	M1	Forming an equation for the height using *their* Σ_{11} and Σ_{15}.
	181 (cm)	A1	
	Alternative method for Question 7(d)		
	$[\Sigma_{15} = 191.2 \times 15 = 2868$ $\Sigma_{15} = 2687 + h\ ]$	B1	Σ_{15} found using the mean and raw data methods. Accept unevaluated.
	their 2868 = *their* 2687 + h	M1	Forming an equation for the height using *their* Σ_{15} expressions.
	181 (cm)	A1	
	Alternative method for Question 7(d)		
	$[\Sigma_{15} = 2687 + h$ $\frac{\Sigma_{15}}{15} = 191.2\ ]$	B1	Σ_{15} found using raw data method and statement on calculating new mean. Accept unevaluated.
	$\frac{\text{their } 2687 + h}{15} = 191.2$	M1	Forming an equation for the height using *their* Σ_{15} expressions
	181 (cm)	A1	
		3	**N.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	**B1**	Accept 60 or 61. No decimals
		1	
(b)	65% of 160 = 104	**M1**	0.65 × 160 (=104) seen unsimplified or implied by use on graph
	136 (cm)	**A1**	Use of graph must be seen. **SCB1** correct value (136 only) if neither 104 nor use of graph are evident
		2	
(c)	UQ: 150 LQ: 76 IQR = 150 – 76 = 74 [cm]	**M1**	UQ – LQ ; 148 ⩽ UQ ⩽ 152; 74 ⩽ LQ ⩽ 78.
		A1	Must be from 150 - 76
		2	

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.
	$\text{Sd}^2 = \frac{560100}{17} - \textit{their}\,178.88^2 \left[= 948.289\right]$	**M1**	Appropriate variance formula using *their* mean^2, 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:

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

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)	$\left[\frac{\Sigma x}{40} - k = \frac{\Sigma(x-k)}{40}\right]$ $\frac{40\times 34}{40} - k = \frac{520}{40}$	M1	Forms an equation involving Σx, $\Sigma(x\text{-}k)$ and k. Accept at a numeric stage with k.
	$k\,[=34-13]=21$	A1	Evaluated.
		2	

Question	Answer	Marks	Guidance
(b)	$\text{Var} = \left[\frac{\Sigma(x-k)^2}{40} - \left(\frac{\Sigma(x-k)}{40}\right)^2\right] = \frac{9640}{40} - \left(\frac{520}{40}\right)^2 = [241 - 13^2 =]$	M1	Values substituted into an appropriate variance formula, accept unsimplified.
	72	A1	
		2	

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]

.....

.....

.....

.....

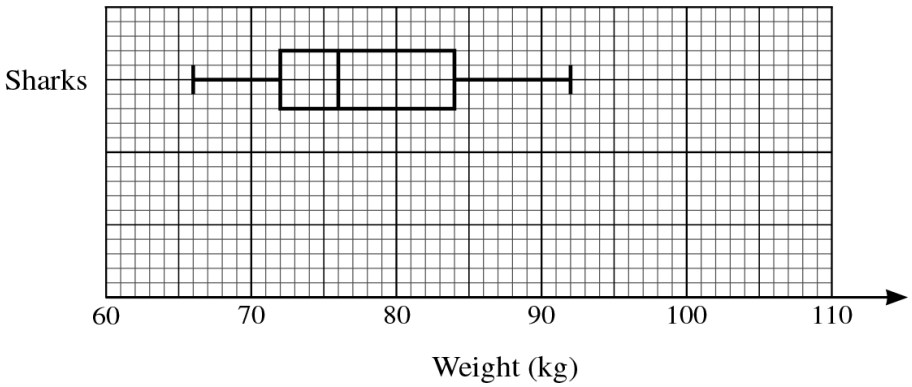
.....

.....

.....

.....

A box-and-whisker plot for the Sharks is shown below.



(c) On the same diagram, draw a box-and-whisker plot for the Rebels. [2]

(d) Make one comparison between the weights of the players in the Rebels club and the weights of the players in the Sharks club. [1]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Rebels | Sharks 6 | 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 Key: 8 | 7 | 2 means 78 kg for Rebels and 72 kg for Sharks	**B1**	Correct stem, ignore extra values (not in reverse).
		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	

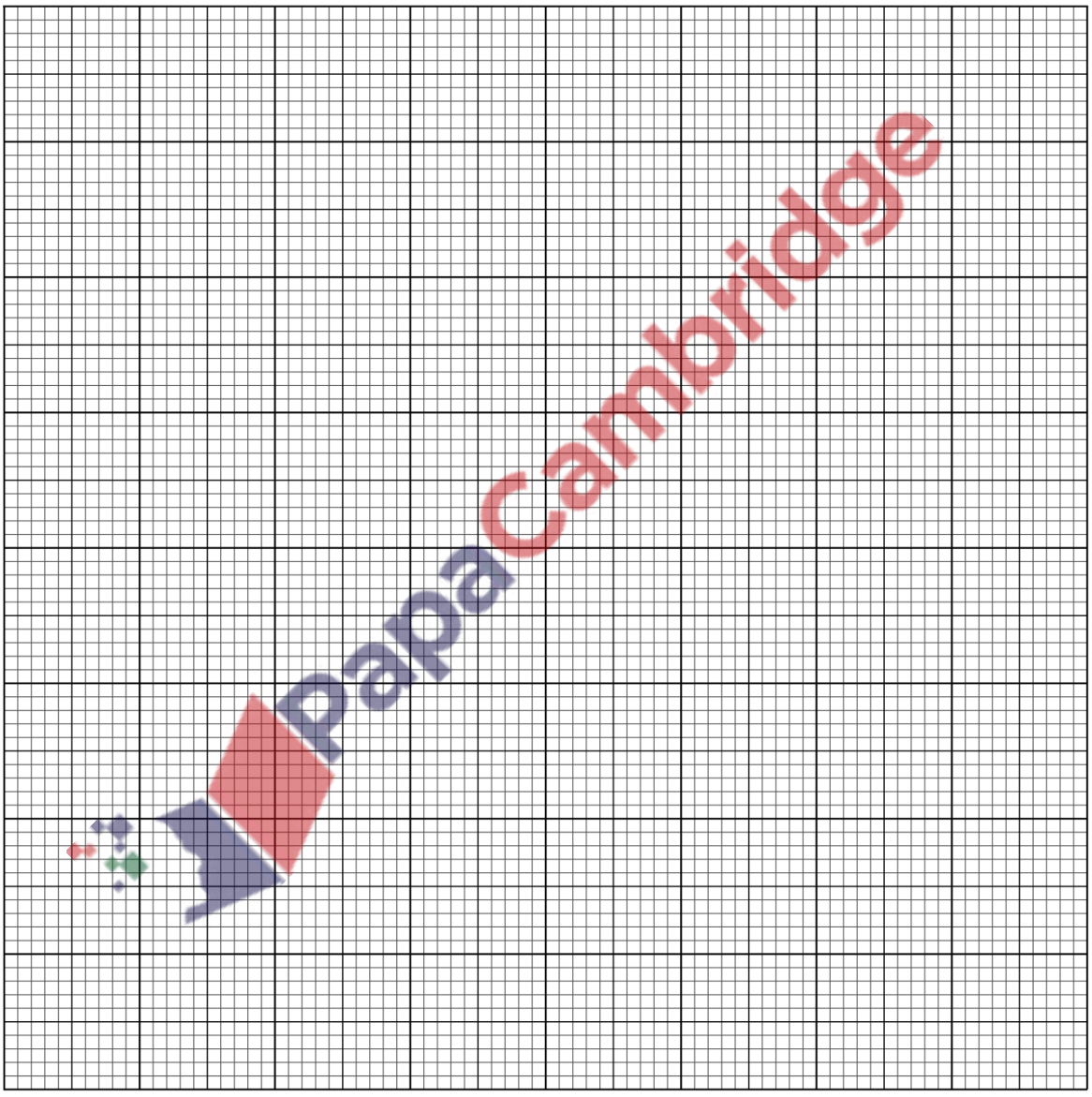
Question	Answer	Marks	Guidance
(b)	Median = 84 (kg)	**B1**	
	[UQ = 93, LQ = 80] 93 – 80	**M1**	$95 \leqslant UQ \leqslant 89 - 79 \leqslant LQ \leqslant 82$
	[IQR =] 13 (kg)	**A1**	WWW
		3	
(c)	Box and whisker with end points 75 and 102	**B1**	Whiskers drawn to correct end points not through box, not 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	$x \leqslant 200$	$x \leqslant 300$	$x \leqslant 500$	$x \leqslant 900$	$x \leqslant 1200$	$x \leqslant 1600$
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]

.....

.....

.....

.....

.....

.....

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

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Cumulative frequency graph drawn	**B1**	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.
		B1	All plotted correctly at correct upper end points (200 etc.) **curve** drawn accurately joined to (0, 0) (straight line segments B0) but no daylight above 140. Cf scale no less than 2 cm = 20 children .
		2	

Question	Answer	Marks	Guidance
(b)	[UQ at 75% of 140 = 105, LQ at 25% of 140 = 35] [IQR:] 700 – 260	**M1**	Accept $660 \leqslant \text{UQ} \leqslant 720 - 240 \leqslant \text{LQ} \leqslant 290$. If values are outside our range, FT providing scales linear and increasing cf drawn.
	440	**A1**	Accept correct evaluation of $660 \leqslant$ *their* UQ $\leqslant 720 - 240 \leqslant$ *their* LQ $\leqslant 290$ with clear indication that graph has been used for at least one of 105 or 35.
		2	

Question	Answer	Marks	Guidance
(c)	[Mean =] $\frac{16\times100+30\times250+42\times400+34\times700+12\times1050+6\times1400}{140}$	**B1**	Frequencies: 16, 30, 42, 34, 12, 6 Mid-points: 100, 250, 400, 700, 1050, 1400 5 or 6 correct frequency values seen.
		B1	5 or 6 correct midpoint values seen.
		M1	Values substituted into mean formula using *their* 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 **M1**.
	505	**A1**	WWW
	Variance = $\frac{16\times100^2+30\times250^2+42\times400^2+34\times700^2+12\times1050^2+6\times1400^2}{140}$ -505^2	**M1**	Values substituted into variance formula using (*their* mean)^2 and *their* midpoints and *their* frequencies (including for denominator). Accept unsimplified. Condone 1 data error. Accept: $[\frac{160\,000+1\,875\,000+6\,720\,000+16\,660\,000+13\,230\,000+11\,760\,000}{140}$ or $\frac{50\,405\,000}{140}$ or 360 035.7143] $-[505^2$ or 255 025$]$ If formula stated accept 105 010 or 105 011 WWW.
	S.d. = $[\sqrt{105\,010.7} =]$ 324	**A1**	WWW
		6	

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:

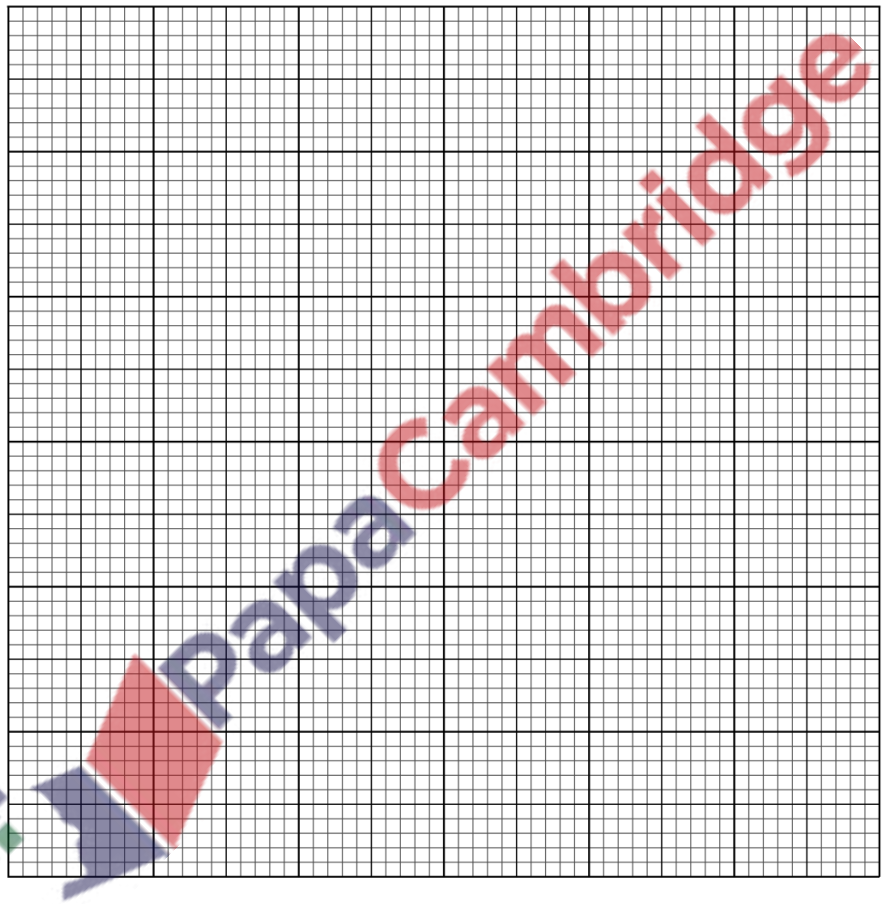
Question	Answer	Marks	Guidance
(a)	Lakeview | Riverside 9 4 0 | 1 | 8 8 8 7 6 2 | 2 | 0 1 3 4 5 5 3 2 0 | 3 | 0 6 7 1 | 4 Key: 6|2|3 means 26m for Lakeview and 23m for Riverside	B1	Correct stem, ignore extra values.
		B1	Correct Lakeview labelled on left, leaves in order from right to left and lined up vertically, no commas.
		B1	Correct Riverside labelled on same diagram, leaves in order and lined up vertically, no commas.
		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.
		4	
(b)	UQ = 32, LQ = 19	M1	$(30 \leqslant UQ \leqslant 33) - (14 \leqslant LQ \leqslant 22)$
	IQR = 32 – 19 = 13	A1	WWW
		2	

16. 9709_w21_qp_53 Q: 3

The times taken, in minutes, by 360 employees at a large company to travel from home to work are summarised in the following table.

Time, t minutes	$0 \leqslant t < 5$	$5 \leqslant t < 10$	$10 \leqslant t < 20$	$20 \leqslant t < 30$	$30 \leqslant t < 50$
Frequency	23	102	135	76	24

(a) Draw a histogram to represent this information. [4]



.....

.....

.....

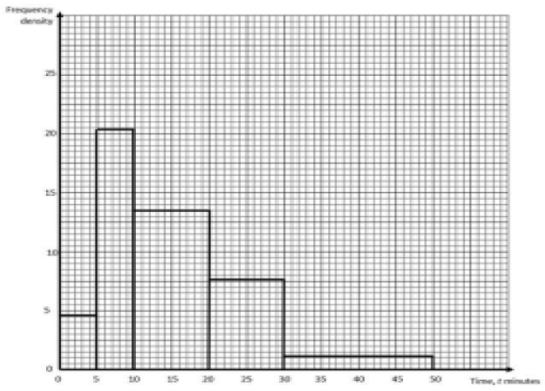
.....

.....

.....

(b) Calculate an estimate of the mean time taken by an employee to travel to work. [2]

Answer:

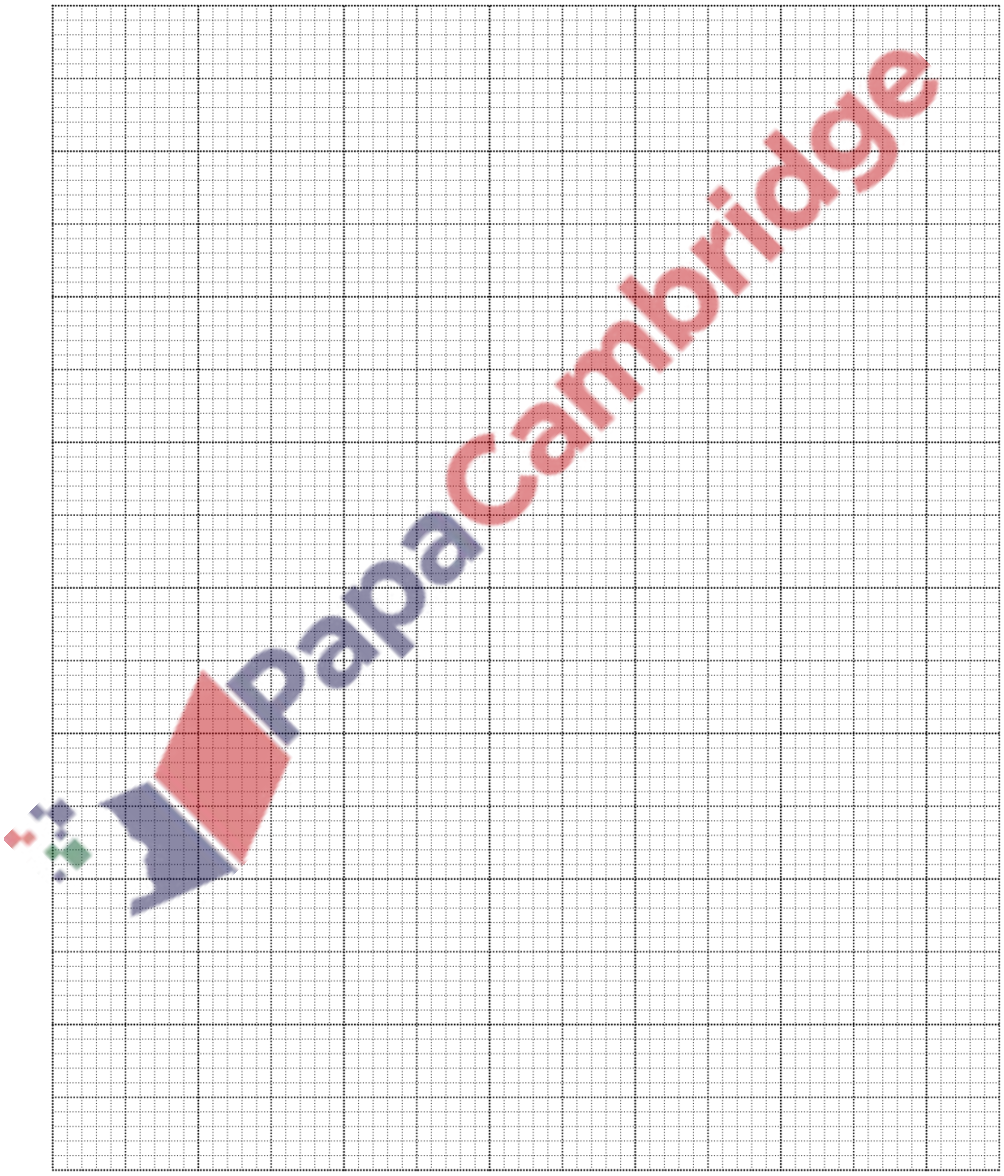
Question	Answer	Marks	Guidance
(a)	Cw: 5 5 10 10 20	**M1**	At least 4 frequency densities calculated (f/cw), accept unsimplified and class widths ±1 of true values. May be implied by graph.
	Fd: 4.6 20.4 13.5 7.6 1.2	**A1**	All heights correct on graph **NOT FT**
		**B1**	Bar ends at 0, 5, 10, 20, 30, 50 clear intention not to draw at 4.5 or 5.5 etc.
		B1	Axes labelled: Frequency density (fd), time (t) and mins (or appropriate title). Linear scales between 0 and 20.4 or above on vertical axis, and 0 and 50 or above on the horizontal axis. (Axes may be reversed.)
		4	
(b)	$\frac{2.5\times23+7.5\times102+15\times135+25\times76+40\times24}{360}$	**M1**	Uses at least 4 midpoint attempts (e.g. 2.5 ± 0.5) in correct formula, accept unsimplified expression, denominator either correct or *their* Σfrequencies .
	$\left[\frac{5707.5}{360}=\right]15.9,\ 15\frac{41}{48}$	**A1**	Evaluated.
		2	

17. 9709_m20_qp_52 Q: 7

Helen measures the lengths of 150 fish of a certain species in a large pond. These lengths, correct to the nearest centimetre, are summarised in the following table.

Length (cm)	0 – 9	10 – 14	15 – 19	20 – 30
Frequency	15	48	66	21

(a) Draw a cumulative frequency graph to illustrate the data. [4]



(b) 40% of these fish have a length of d cm or more. Use your graph to estimate the value of d. [2]

.....

.....

.....

.....

The mean length of these 150 fish is 15.295 cm.

(c) Calculate an estimate for the variance of the lengths of the fish. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

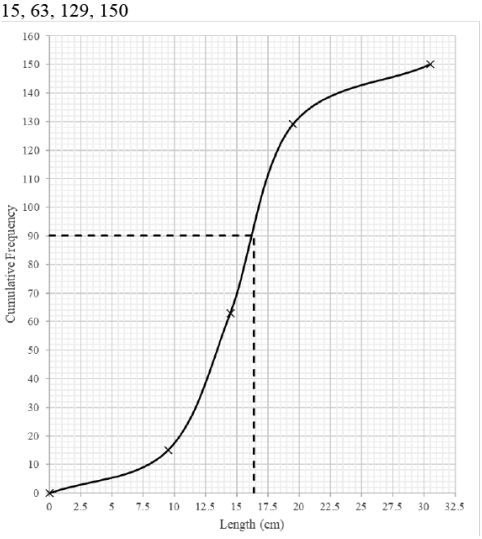
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	15, 63, 129, 150 	**B1**	Correct cumulative frequencies seen (may be on graph)
		B1	0 ⩽ Horizontal axis ⩽ 30, 0 ⩽ vertical axis ⩽ 150 Labels correct: length cm, cf
		M1	At least 3 points plotted at upper end points (e.g. allow 9, 9.5, 10) with a linear horizontal scale.
		A1	Linear vertical scale, all points at correct upper end points (9.5 etc.), curve drawn accurately, joined to (0,0) (condone (–0.5, 0))
		4	
(b)	60% of 150 = 90	**M1**	90 seen or implied by use on graph
	Approx. 16.5 [cm]	**A1FT**	FT *their* increasing cumulative frequency graph, Use of graph must be seen. If no clear evidence of use of graph **SCB1FT** correct value from *their* graph
		2	

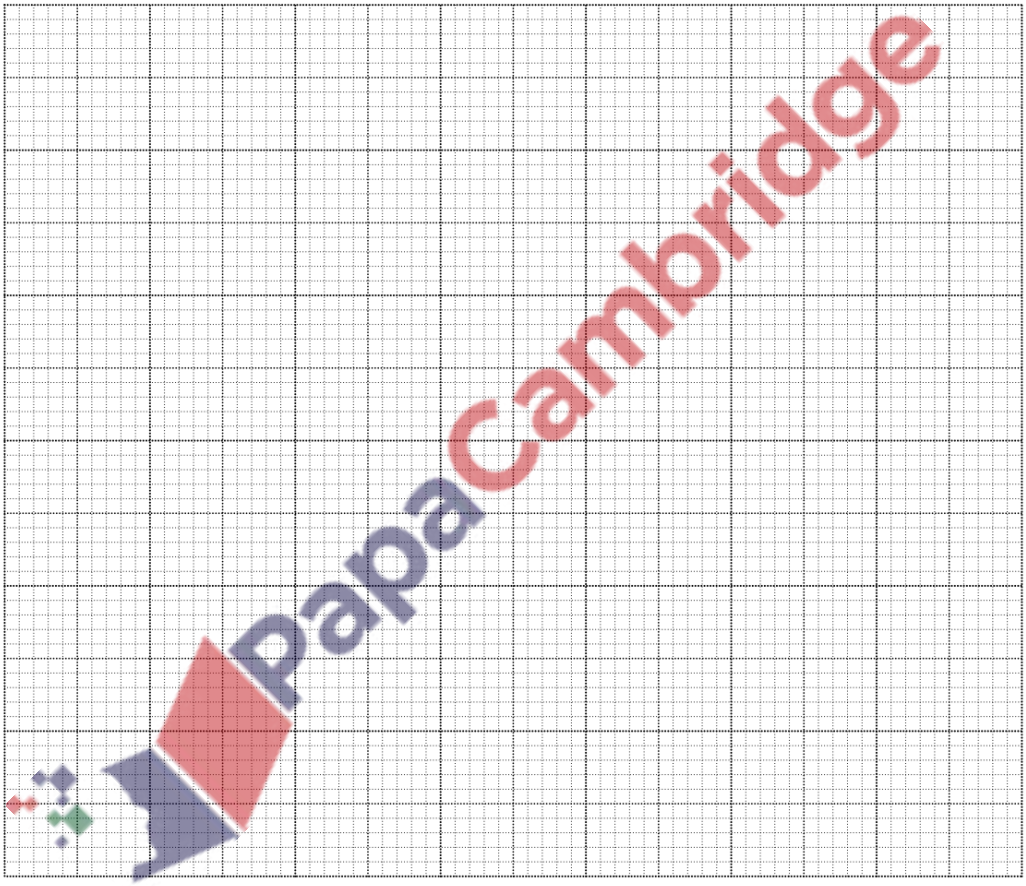
Question	Answer	Marks	Guidance
(c)	Midpoints: 4.75, 12, 17, 25	**M1**	At least 3 correct midpoints used (39449.4375 implies M1)
	Var = $\frac{4.75^2 \times 15 + 12^2 \times 48 + 17^2 \times 66 + 25^2 \times 21}{150} - 15.295^2$	**M1**	Using midpoints ±0.5 in correct var formula, including subtraction of *their* μ^2.
	= 29.1	**A1**	
		3	

18. 9709_s20_qp_51 Q: 7

The numbers of chocolate bars sold per day in a cinema over a period of 100 days are summarised in the following table.

Number of chocolate bars sold	1 – 10	11 – 15	16 – 30	31 – 50	51 – 60
Number of days	18	24	30	20	8

(a) Draw a histogram to represent this information. [5]



(b) What is the greatest possible value of the interquartile range for the data? [2]

.....

.....

.....

.....

.....

.....

.....

(c) Calculate estimates of the mean and standard deviation of the number of chocolate bars sold. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks
(a)	Class widths: 10, 5, 15, 20, 10	**M1**
	Frequency density = frequency/*their* class width: 1.8, 4.8, 2, 1, 0.8	**M1**
	All heights correct on diagram (using a linear scale)	**A1**
	Correct bar ends	**B1**
	Bar ends: 10.5, 15.5, 30.5, 50.5, 60.5	**B1**
		5
(b)	11 – 15 and 31 – 50	**B1**
	Greatest IQR = 50 – 11 = 39	**B1**
		2
(c)	Mean $= \frac{18\times5.5+24\times13+30\times23+20\times40.5+8\times55.5}{100} = \frac{2355}{100} = 23.6$	**B1**
	Var $= \frac{18\times5.5^2+24\times13^2+30\times23^2+20\times40.5^2+8\times55.5^2}{100} - \text{mean}^2$	**M1**
	$\frac{77917.5}{100} - \text{mean}^2 = 224.57$	**A1**
	Standard deviation = 15.0 (**FT** *their* variance)	**A1 FT**
		4

19. 9709_s20_qp_52 Q: 1

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

$$\Sigma(x - 50) = 144 \quad \text{and} \quad \Sigma x = 944.$$

Find the value of n. [3]

Answer:

Question	Answer	Marks
	$\Sigma x - 50n = 144$	B1
	$50n + 144 = 944$	M1
	$n = 16$	A1
		3

20. 9709_s20_qp_52 Q: 3

Two machines, A and B, produce metal rods of a certain type. The lengths, in metres, of 19 rods produced by machine A and 19 rods produced by machine B are shown in the following back-to-back stem-and-leaf diagram.

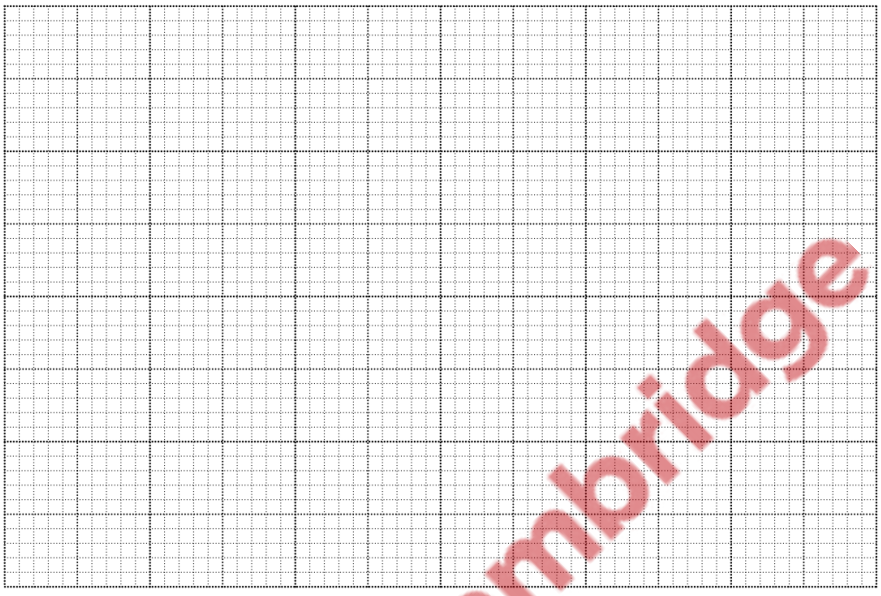
A							B					
						21	1	2	4			
		7	6	3	0	22	2	4	5	5	6	
8	7	4	3	1	1	23	0	2	6	8	9	9
	5	5	5	3	2	24	3	3	4	6		
		4	3	1	0	25	6					

Key: 7 | 22 | 4 means 0.227 m for machine A and 0.224 m for machine B.

(a) Find the median and the interquartile range for machine A. [3]

It is given that for machine B the median is 0.232 m, the lower quartile is 0.224 m and the upper quartile is 0.243 m.

(b) Draw box-and-whisker plots for A and B. [3]



(c) Hence make two comparisons between the lengths of the rods produced by machine A and those produced by machine B. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	*Answer*	Marks
(a)	Median = 0.238	**B1**
	UQ = 0.245, LQ = 0.231, So IQR = 0.245 – 0.231	**M1**
	0.014	**A1**
		3

Question	Answer	Marks
(b)		
	Medians and quartiles correctly plotted for *A* or *B*	**B1**
	End points correct for *A* or *B*	**B1**
	Completely correct, including scale	**B1**
		3
(c)	Lengths of rods produced by machine *A* are longer. (**B1** for comparison of central tendency)	**B1**
	Lengths of rods produced by machine *A* are less spread out (**B1** for comparison of spread)	**B1**
		2

		LQ	M	UQ	
A	0.220	0.231 **FT**	0.238 **FT**	0.245 **FT**	0.254
B	0.211	0.224	0.232	0.243	0.256

21. 9709_s20_qp_53 Q: 6

The annual salaries, in thousands of dollars, for 11 employees at each of two companies *A* and *B* are shown below.

Company *A*	30	32	35	41	41	42	47	49	52	53	64
Company *B*	26	47	30	52	41	38	35	42	49	31	42

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

(b) Find the median and the interquartile range of the salaries of the employees in company *A*. [3]

A new employee joins company *B*. The mean salary of the 12 employees is now $38 500.

(c) Find the salary of the new employee. [3]

Answer:

Question	Answer	Marks
(a)	A | | B 2 | 6 5 2 0 | 3 | 0 1 5 8 9 7 2 1 1 | 4 | 1 2 2 7 9 3 2 | 5 | 2 4 | 6 KEY 1 | 4 | 2 means \$41 000 for A and \$42 000 for B	
	Correct stem	B1
	Correct A on LHS	B1
	Correct B on same diagram	B1
	Correct key for *their* diagram, both companies identified and correct units	B1
		4
(b)	Median = [\$]42 000	B1
	LQ = [\$]35 000 UQ = [\$]52 000	B1
	IQR = [\$]17 000 (**FT** if $49000 \leqslant UQ \leqslant 53000 - 32000 \leqslant LQ \leqslant 41000$)	B1 FT
		3

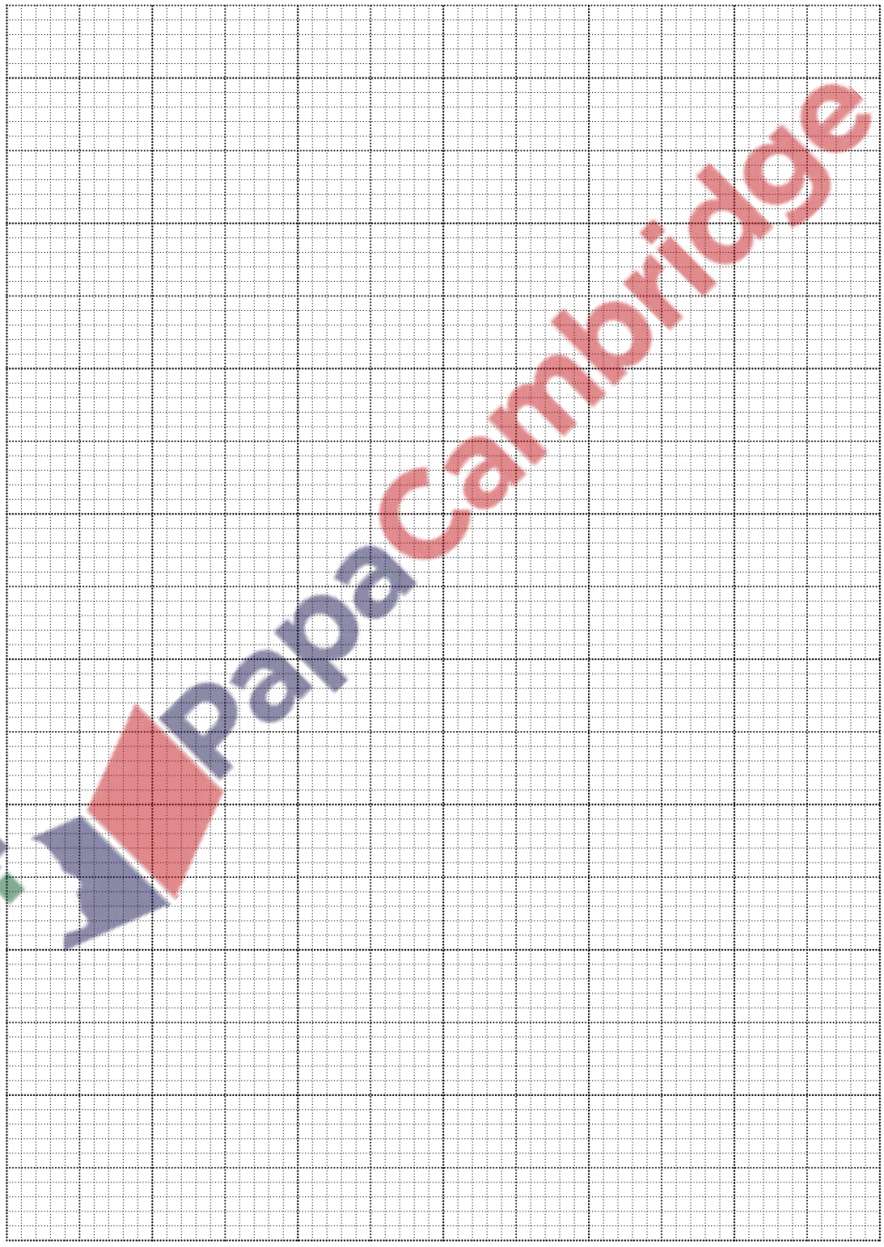
Question	Answer	Marks
(c)	Sum of given 11 numbers is 433 000	M1
	Sum of 12 numbers, including new = $38\,500 \times 12 = 462\,000$	M1
	Difference = new salary = [\$]29 000	A1
		3

22. 9709_w20_qp_51 Q: 6

The times, t minutes, taken by 150 students to complete a particular challenge are summarised in the following cumulative frequency table.

Time taken (t minutes)	$t \leqslant 20$	$t \leqslant 30$	$t \leqslant 40$	$t \leqslant 60$	$t \leqslant 100$
Cumulative frequency	12	48	106	134	150

(a) Draw a cumulative frequency graph to illustrate the data. [2]



(b) 24% of the students take k minutes or longer to complete the challenge. Use your graph to estimate the value of k. [2]

.....

.....

.....

.....

.....

(c) Calculate estimates of the mean and the standard deviation of the time taken to complete the challenge. [6]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

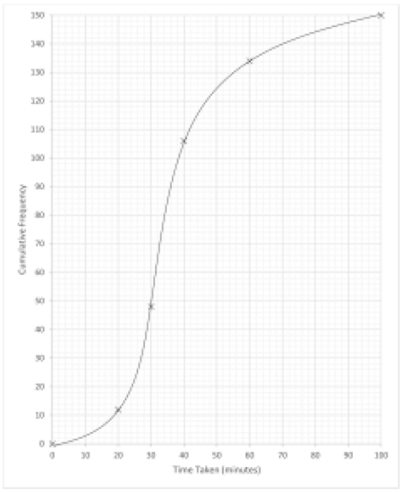
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)		M1	At least 4 points plotted at upper end points, with both scales linear with at least 3 values indicated
	Correct cumulative frequency curve	A1	All plotted correctly with curve drawn joined to (0, 0), axes labelled cumulative frequency, time, minutes
		2	
(b)	$150 \times 0{\cdot}76 = 114$	M1	114 SOI, may be on graph
	$k = 45$ (mins)	A1 FT	Clear indication that *their* graph has been used, tolerance ±1mm
		2	

Question	Answer	Marks	Guidance
(c)	Frequencies: 12 36 58 28 16	B1	Correct frequencies seen
	Mean $= \dfrac{10\times12+25\times36+35\times58+50\times28+80\times16}{150}$	B1	At least 4 correct midpoints seen and used
	$\dfrac{120+900+2030+1400+1280}{150}$	M1	Correct formula with *their* midpoints (not upper boundary, lower boundary, class width or frequency density).
	$38.2, 38\dfrac{1}{5}$	A1	
	Variance $= \dfrac{12\times10^2+36\times25^2+58\times35^2+28\times50^2+16\times80^2}{150} - mean^2$ $= \dfrac{1200+22500+71050+70000+102400}{150} - mean^2$	M1	Substitute *their* midpoints and frequencies (condone use of cumulative frequency) in correct variance formula, must have '– *their* mean^2'
	(Standard deviation $= \sqrt{321.76}$) $= 17.9$	A1	
		6	

23. 9709_w20_qp_52 Q: 5

The following table gives the weekly snowfall, in centimetres, for 11 weeks in 2018 at two ski resorts, Dados and Linva.

Dados	6	8	12	15	10	36	42	28	10	22	16
Linva	2	11	15	16	0	32	36	40	10	12	9

(a) Represent the information in a back-to-back stem-and-leaf diagram. [4]

(b) Find the median and the interquartile range for the weekly snowfall in Dados. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) The median, lower quartile and upper quartile of the weekly snowfall for Linva are 12, 9 and 32 cm respectively. Use this information and your answers to part **(b)** to compare the central tendency and the spread of the weekly snowfall in Dados and Linva. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

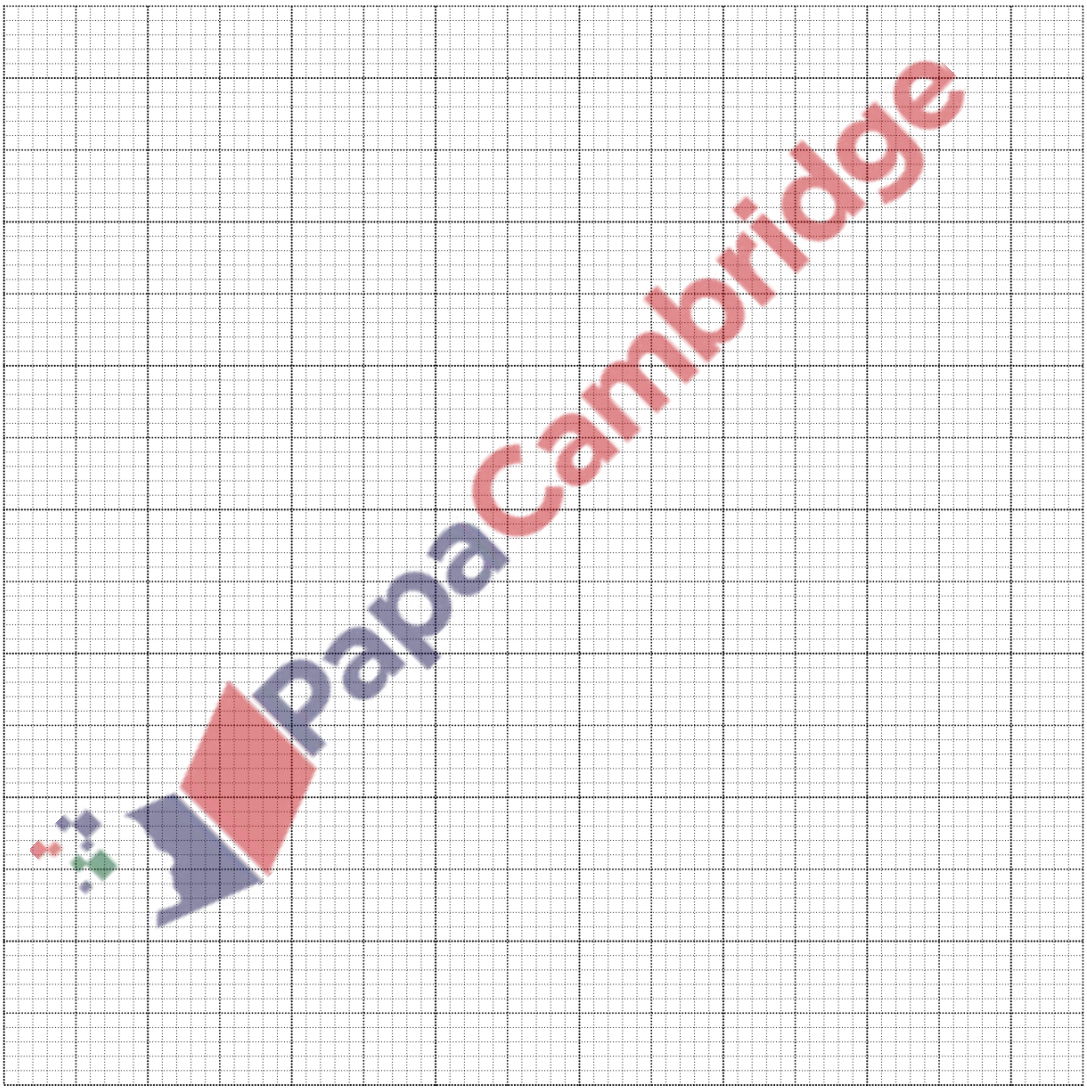
Question	Answer	Marks	Guidance
(a)	*Dados* | | *Linva* 8 6 | 0 | 0 2 9 6 5 2 0 0 | 1 | 0 1 2 5 6 8 2 | 2 | 6 | 3 | 2 6 2 | 4 | 0 KEY 6| 3| 2 means 36 cm (snow) in Dados and 32 cm (snow) in Linva	**B1**	Correct stem can be upside down, ignore extra values
		B1	Correct Dados labelled, leaves in order and lined up vertically (less than midway to next column), no commas etc, no extra terms
		B1	Correct Linva on opposite side of stem labelled, leaves in order and lined up vertically (less than midway to next column), no commas etc, no extra terms
		B1	Correct single key for their diagram, need both resorts identified and 'cm' stated at least once here or in leaf headings or title. **SC** If 2 separate diagrams drawn, **SCB1** if both keys meet these criteria B0B1B0SCB1 max.
		4	
(b)	Median or Q2 = 15 (cm)	**B1**	Correct
	UQ or Q3 = 28 cm, LQ or Q1 = 10 cm IQR = 28 – 10	**M1**	$22 \leqslant UQ \leqslant 36 - 8 \leqslant LQ \leqslant 10$
	18 (cm)	**A1**	WWW
		3	
(c)	On average the snowfall in Davos is higher	**B1 FT**	FT from *their* **5(b)** values for Dados. Statement comparing central tendency in context
	The amount of snowfall in Linva varies more than in Davos	**B1 FT**	Statement comparing spread in context Note: simply stating and comparing the values is not sufficient.
		2	

24. 9709_w20_qp_53 Q: 7

A particular piece of music was played by 91 pianists and for each pianist, the number of incorrect notes was recorded. The results are summarised in the table.

Number of incorrect notes	1 – 5	6 – 10	11 – 20	21 – 40	41 – 70
Frequency	10	5	26	32	18

(a) Draw a histogram to represent this information. [5]



(b) State which class interval contains the lower quartile and which class interval contains the upper quartile.

Hence find the greatest possible value of the interquartile range. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(c) Calculate an estimate for the mean number of incorrect notes. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

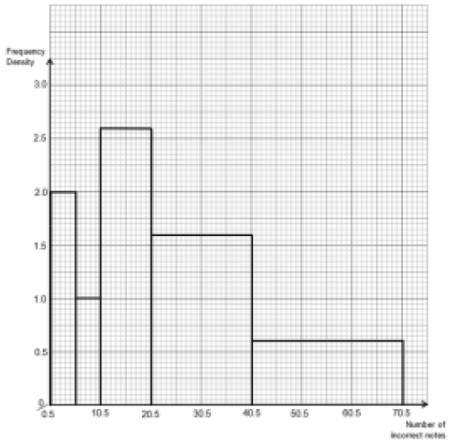
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(a)	Class widths: 5, 5, 10, 20, 30 Frequency density: 2, 1, 2.6, 1.6, 0.6	**M1**	At least 3 class widths correct and used in a calculation
		M1	At least 3 correct frequency densities unsimplified – FT *their* class widths
		**A1**	All correct heights on a histogram using a linear vertical scale from zero – no FT
		B1	Correct upper bar ends (5.5, 10.5, 20.5, 40.5, 70.5) and 4 correct lower bar ends of 5.5, 10.5, 20.5, 40.5. Condone 0 or 1.
		B1	Linear scales with at least 3 values indicated on each axis, vertical scale from 0, axes labelled 'fd' and 'no. of (incorrect) notes', or better.
		5	
(b)	LQ: 11 – 20 UQ: 21 – 40	**B1**	Both UQ and LQ correct
	Greatest IQR = 40 – 11 = 29	**B1 FT**	Subtract lower end of *their* LQ interval from upper end of *their* UQ interval
		2	

Question	Answer	Marks	Guidance
(c)	Midpoints: 3 8 15.5 30.5 55.5	**M1**	At least 4 midpoints correct and used
	Mean $= \frac{3\times10+8\times5+15.5\times26+30.5\times32+55.5\times18}{91}$ $= \frac{30+40+403+976+999}{91}$ $= \frac{2448}{91}$	**M1**	Correct formula with *their* midpoints (not upper boundary, lower boundary, class width, frequency density, frequency or cumulative frequency)
	$26.9,\ 26\frac{82}{91}$	**A1**	Accept 26 or 27
		3	

25. 9709_m19_qp_62 Q: 2

For 40 values of the variable x, it is given that $\Sigma(x - c)^2 = 3099.2$, where c is a constant. The standard deviation of these values of x is 3.2.

(i) Find the value of $\Sigma(x - c)$. [3]

(ii) Given that $c = 50$, find the mean of these values of x. [1]

Answer:

Question	Answer	Marks	Guidance
(i)	$\sigma^2 = \frac{\Sigma(x-c)^2}{n} - \left(\frac{\Sigma(x-c)}{n}\right)^2$ $3.2^2 = \frac{3099.2}{40} - \left(\frac{\Sigma(x-c)}{40}\right)^2$	**M1**	Use correct formula with values substituted
	$\left(\frac{\Sigma(x-c)}{40}\right)^2 = 67.24$: $\Sigma(x-c) = 40 \times \sqrt{67.24}$	**M1**	Rearrange to make *their* $\left(\frac{\Sigma(x-c)}{40}\right)^2$ the subject, unsimplified.
	$= 328$	**A1**	Exact value, cao
		3	
(ii)	$\Sigma x - 40c =$ *their* **(i)** Mean $= \frac{\textit{their (i)}}{40} + 50$ $= 58.2$	**B1FT**	FT *their* **(i)**
		1	

26. 9709_m19_qp_62 Q: 5

The weights, in kg, of the 11 members of the Dolphins swimming team and the 11 members of the Sharks swimming team are shown below.

Dolphins	62	75	69	82	63	80	65	65	73	82	72
Sharks	68	84	59	70	71	64	77	80	66	74	72

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

(ii) Find the median and interquartile range for the Dolphins. [3]

Answer:

Question	Answer	Marks	Guidance
(i)	Dolphins | Sharks 5 | 9 9 5 5 3 2 | 6 | 4 6 8 5 3 2 | 7 | 0 1 2 4 7 2 2 0 | 8 | 0 4 Key: 3|6|4 means 63 kg for Dolphins and 64 kg for Sharks	**B1**	Correct stem can be upside down, ignore extra values,
		B1	Correct Dolphin must be on LHS,
		B1	Correct Sharks on either LHS or RHS of back-to-back. Alignment ± half a space, no late entries squeezed in, no crossing out if shape is changed. Condone a separate RHS stem-and-leaf diagram
		B1FT	Correct single key for *their* single diagram, need both teams identified and 'kg' stated at least once here or in leaf headings or title.
		4	
(ii)	Median = 72 LQ = 65, UQ = 80,	**B1**	72<UQ<82 – 62<LQ<72
	IQR = 80 – 65	**M1**	nfww
	= 15	**A1**	SCB1 if M0 scored for LQ = 65 **and** UQ = 80
		3	

27. 9709_s19_qp_61 Q: 1

The times, t seconds, taken to swim 100 m were recorded for a group of 9 swimmers and were found to be as follows.

95 126 117 135 120 125 114 119 136

(i) Find the values of $\Sigma(t - 120)$ and $\Sigma(t - 120)^2$. [2]

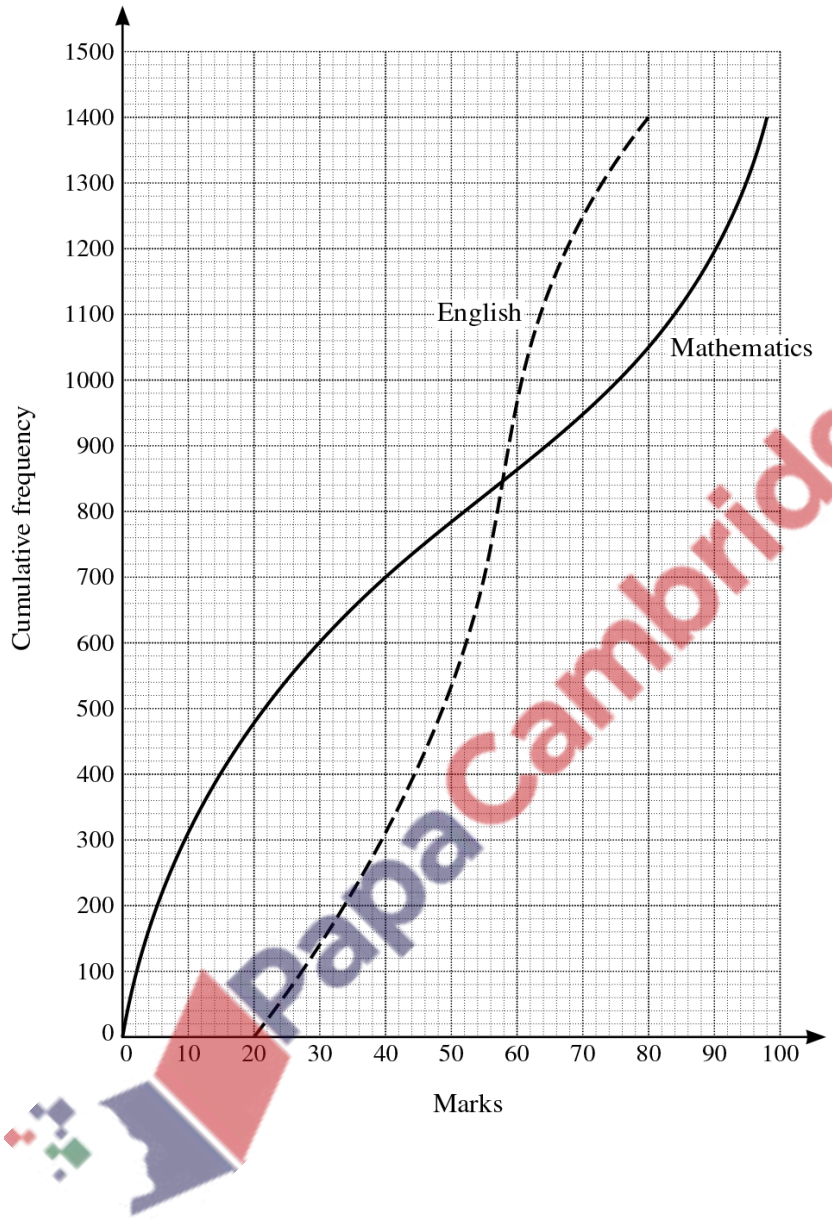
(ii) Using your values found in part **(i)**, calculate the variance of t. [2]

Answer:

Question	Answer	Marks	Guidance
(i)	$\Sigma(t-120) = -25+6-3+15+0+5-6-1+16 = 7$	**M1**	Attempt to sum both $(t-120)$ and $(t-120)^2$ Correct ans using $\Sigma t - 9\times 120$ and $\Sigma (t-120)^2$ M1A1
	$\Sigma(t-120)^2 = 25^2+6^2+3^2+15^2+0^2+5^2+6^2+1^2+16^2$ $= 1213$	**A1**	Both correct, www, **SC** correct ans no working **B1B1**
		2	
(ii)	$\text{Var} = \frac{\Sigma(t-120)^2}{9} - \left(\frac{\Sigma(t-120)}{9}\right)^2 = \frac{their\,1213}{9} - \left(\frac{their\,7}{9}\right)^2$	**M1**	Using two coded values in correct formula including finding Σt from 7 etc
	$= 134(.2)$	**A1**	Correct answer **SC** if correct variance obtained by another method from raw data give **SCB1**
		2	

28. 9709_s19_qp_61 Q: 4

The Mathematics and English A-level marks of 1400 pupils all taking the same examinations are shown in the cumulative frequency graphs below. Both examinations are marked out of 100.



Use suitable data from these graphs to compare the central tendency and spread of the marks in Mathematics and English. [6]

Answer:

Question	Answer	Marks	Guidance
	Median Maths = 40	**M1**	Indication of finding medians, such as mark on graph or reference marks to 700 pupils, condone poor terminology such as 'mean'
	Median English = 55	**A1**	Both values correct, condone 54<English<56 but 54, 56 get A0
	Median of English is larger than median of Maths	**B1**	Correct statement, median must be referenced within answer. No credit if statement references 'means'
	Range Maths is 100 or IQ range Maths = 80 – 12 = 68	**M1**	Evidence of finding either both ranges or both IQ ranges i.e. see a minus
	Range English is 60 or IQ range English = 62 – 42 = 20	**A1**	Both ranges or IQR correct
	Maths marks have more spread then English marks	**B1**	Correct conclusion. Accept standard deviation but must see some figures
		6	

29. 9709_s19_qp_62 Q: 6

(i) Give one advantage and one disadvantage of using a box-and-whisker plot to represent a set of data. [2]

.....

.....

.....

.....

.....

.....

.....

.....

(ii) The times in minutes taken to run a marathon were recorded for a group of 13 marathon runners and were found to be as follows.

180 275 235 242 311 194 246 229 238 768 332 227 228

State which of the mean, mode or median is most suitable as a measure of central tendency for these times. Explain why the other measures are less suitable. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(iii) Another group of 33 people ran the same marathon and their times in minutes were as follows.

190	203	215	246	249	253	255	254	258	260	261
263	267	269	274	276	280	288	283	287	294	300
307	318	327	331	336	345	351	353	360	368	375

(a) On the grid below, draw a box-and-whisker plot to illustrate the times for these 33 people. [4]

.....

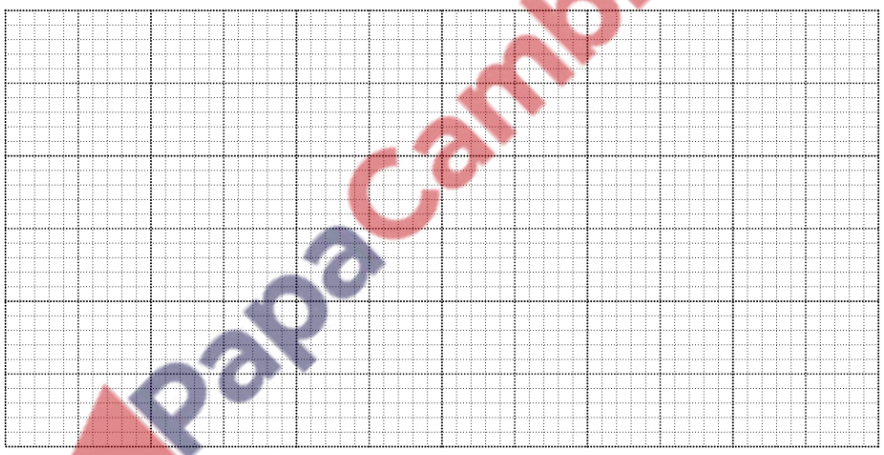
.....

.....

.....

.....

.....



(b) Find the interquartile range of these times. [1]

.....

.....

.....

.....

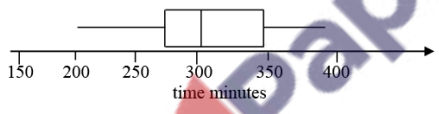
.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	Advantage: comment referring to spread or range or shape	**B1**	Comments referring to quartiles, IQR, Range, median, shape, skewness, data distribution, spread score B1 Any comments with reference to mean or standard deviation or any other 'disadvantage' will score B0 Comments referring to '5-value plot', comparison with another data set, overview or ease of drawing/plotting/reading require an appropriate advantage statement.
	Disadvantage: comment referring to limited data information provided	**B1**	Comments referring to no individual data, no information about the number of values, unable to calculate mean, standard deviation, variance and mode score B1 Any comments with reference to median, shape or any other 'advantage' will score B0 Comments referring to 'size of data set' or 'average' require an appropriate disadvantage statement. Comments referring to outliers are ignored in all cases (as outliers are not in the syllabus content) unless supported by an appropriate advantage / disadvantage statement. If comments not clearly identified, assume first comment is the advantage.
		2	

Question	Answer	Marks	Guidance
(ii)	Not mean as data skewed by one large value	**B1**	Comment which identifies 768 (or 'a very large number') as the problem. Condone the use of 'outlier'
	Not mode as frequencies all the same	**B1**	Comment which indicates that no mode exists (e.g. all the data is different, there is no repeated number, all the values are different)
	Median	**B1**	Median identified as choice, dependent upon statements for mean and mode being given, even if incorrect or very general.
	SC: Mean is identified as most suitable		
	Not mode as frequencies all the same	**SCB1**	Comment which indicates that no mode exists
	Not median as not all values used	**SCB1**	Comment which indicates limitation of median e.g. median is not in middle of range.
		3	
(iii)(a)	LQ = 256 or 256.5 Med = 280 UQ = 329 Min 190 max 375  	**B1**	Median, UQ and LQ values seen, may not be identified or identified correctly. (Not read from box plot unless value stated)
		B1	**FT** Median and quartiles plotted in box on graph, linear scale
		B1	Correct end points, whiskers from ends of box but not through box, not at top or bottom of box
		B1	Uniform scale from 190 to 375 (need at least 3 linear identified points min) and labelled 'time' and 'minutes' (can be in title) **No time axis or time axis with no scale attempt,** Max B1B0B0B0
		4	

Question	Answer	Marks	Guidance
(iii)(b)	IQR = *their* 329 – *their* 256 = 73 or 72.5	**B1**	**FT Must** follow through only from ***their*** stated values (condone if correct quartiles stated here), not reading from graph.
		1	

30. 9709_s19_qp_63 Q: 7

The times in minutes taken by 13 pupils at each of two schools in a cross-country race are recorded in the table below.

Thaters School	38	43	48	52	54	56	57	58	58	61	62	66	75
Whitefay Park School	45	47	53	56	56	61	64	66	69	73	75	78	83

(i) Draw a back-to-back stem-and-leaf diagram to illustrate these times with Thaters School on the left. [4]

(ii) Find the interquartile range of the times for pupils at Thaters School. [2]

.....

.....

.....

.....

.....

.....

.....

.....

The times taken by pupils at Whitefay Park School are denoted by x minutes.

(iii) Find the value of $\Sigma(x - 60)^2$. [2]

(iv) It is given that $\Sigma(x - 60) = 46$. Use this result, together with your answer to part **(iii)**, to find the variance of x. [2]

If you use the following lined page to complete the answer(s) to any question(s), the question number(s) must be clearly shown.

Answer:

Question	Answer	Marks	Guidance
(i)	Thaters School | Whitefay Park School 8 | 3 8 3 | 4 | 5 7 8 8 7 6 4 2 | 5 | 3 6 6 6 2 1 | 6 | 1 4 6 9 5 | 7 | 3 5 8 8 | 3 Key 8 | 4 | 5 represents 48 minutes for Thaters School and 45 minutes for Whitefay Park School.	B1	Correct stem can be upside down, ignore extra values,
		B1	Correct Thaters School labelled on left, leaves in order from right to left and lined up vertically, no commas
		B1	Correct Whitefay Park School labelled on same diagram on right hand side in order from left to right and lined up vertically, no commas
		B1	**FT** Correct key for *their* diagram, need both teams identified and 'minutes' stated at least once here or in leaf headings or title. **SC** If 2 separate diagrams drawn, **SCB1** if both keys meet these criteria
		4	
(ii)	LQ = 50 UQ = 61.5	B1	Both quartiles correct
	IQ range = 61.5 – 50 = 11.5	B1	**FT** $61 \leqslant UQ \leqslant 62 - 48 \leqslant LQ \leqslant 52$
		2	
(iii)	$\Sigma(x-60)^2 = (-15)^2 + (-13)^2 + (-7)^2 + (-4)^2 + (-4)^2 + 1^2 + 4^2 + 6^2 + 9^2 + 13^2 + 23^2 + 15^2 + 18^2$	M1	Summing squares with at least 5 correct unsimplified terms
	= 1856	A1	Exact value
		2	

Question	Answer	Marks	Guidance
(iv)	Var = mean of coded squares – (coded mean)^2 $= \frac{\Sigma(x-60)^2}{13} - \left(\frac{\Sigma(x-60)}{13}\right)^2$	M1	Using two coded values in correct formula (variance or sd)
	Var $= \frac{their\,1856}{13} - \left(\frac{46}{13}\right)^2$ = 130	A1	Correct answer **SC** if correct variance obtained by another method give **SCB1**
		2	

31. 9709_w19_qp_61 Q: 3

The mean and standard deviation of 20 values of x are 60 and 4 respectively.

(i) Find the values of Σx and Σx^2. [3]

Another 10 values of x are such that their sum is 550 and the sum of their squares is 40 500.

(ii) Find the mean and standard deviation of all these 30 values of x. [4]

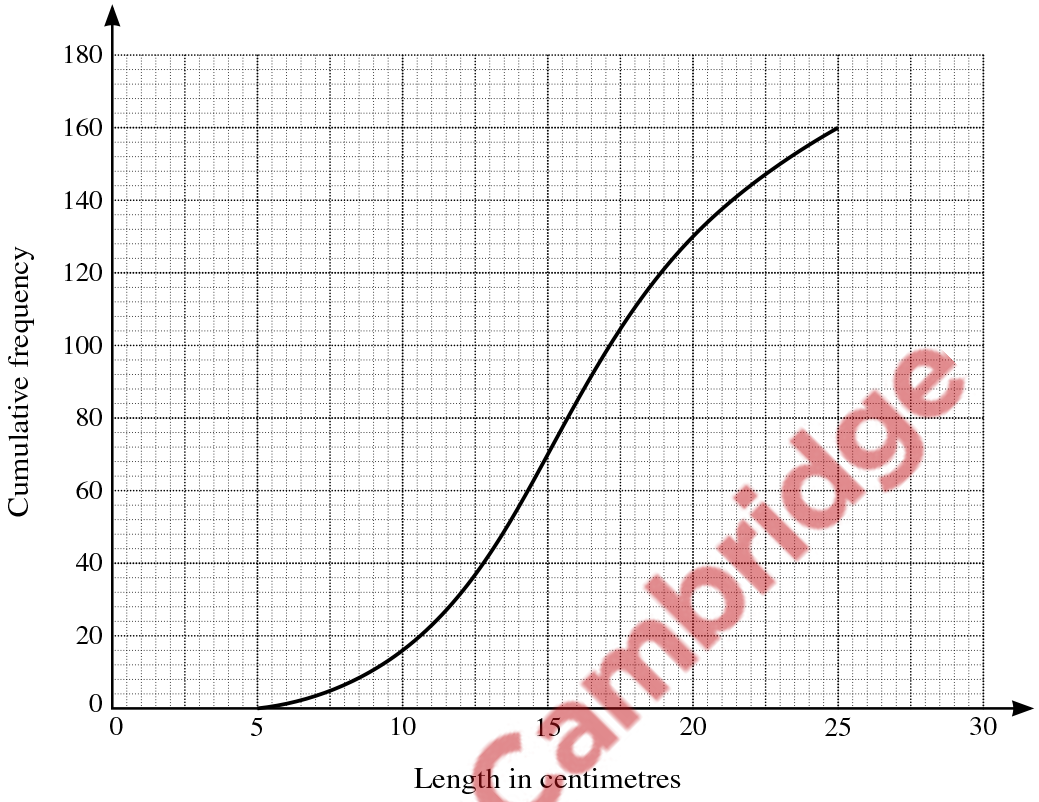
Answer:

Question	Answer	Marks	Guidance
(i)	$\Sigma x = 60 \times 20 \quad = 1200$	**B1**	
	$\frac{\Sigma x^2}{20} - 60^2 = 4^2$	**M1**	Correct variance formula used, condone = 4
	$\Sigma x^2 = 3616 \times 20 \quad = 72320$	**A1**	Exact value
		3	

Question	Answer	Marks	Guidance
(ii)	$\Sigma x = 1200 + 550 = 1750$ $\Sigma x^2 = 72320 + 40500 = 112800$	**M1**	Summing both values of Σx and Σx^2
	Mean $= \frac{\textit{their}1750}{30} = 58.3$	**B1FT**	FT *their* 1750 (not 550 or 1200)/*their*(20+10), accept unsimplified
	Variance $= \frac{\textit{their}112820}{30} - \left(\frac{\textit{their}1750}{30}\right)^2 \quad (= 357.89)$	**M1**	substitute *their* Σx and Σx^2 into correct variance formula
	s.d. = 18.9	**A1**	
		4	

32. 9709_w19_qp_61 Q: 5

Ransha measured the lengths, in centimetres, of 160 palm leaves. His results are illustrated in the cumulative frequency graph below.



(i) Estimate how many leaves have a length between 14 and 24 centimetres. [1]

.....

.....

.....

.....

.....

(ii) 10% of the leaves have a length of L centimetres or more. Estimate the value of L. [2]

.....

.....

.....

.....

.....

(iii) Estimate the median and the interquartile range of the lengths. [3]

.....

.....

.....

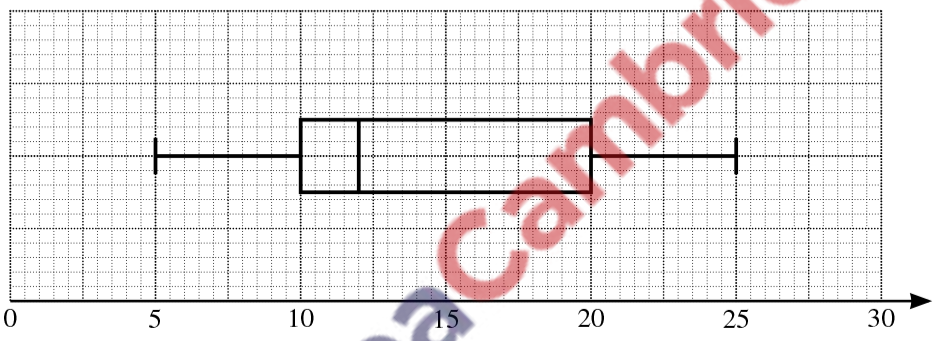
.....

.....

.....

.....

Sharim measured the lengths, in centimetres, of 160 palm leaves of a different type. He drew a box-and-whisker plot for the data, as shown on the grid below.



(iv) Compare the central tendency and the spread of the two sets of data. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	156 – 55 = 99	**B1**	98 ⩽ answer < 100
		1	
(ii)	90% of 160 = 144	**M1**	144 seen, may be marked on graph
	(L =) 22	**A1**	
		2	
(iii)	Median = 15.6 UQ = 18.8, LQ = 12.7	**B1**	15.5 < median < 15.8
	IQR = 18.8 – 12.7	**M1**	18.5 < UQ < 19 – 12.5 < LQ < 13
	6.1	**A1**	6.0 ⩽ IQR ⩽ 6.2
		3	
(iv)	The Median higher for Ransha (1st set of data)	**B1**	Any correct comparison of central tendency, must mention median
	IQR lower for Ransha (1st set of data)	**B1**	Any correct comparison of spread, must refer to IQR
		2	

33. 9709_w19_qp_62 Q: 1

Twelve tourists were asked to estimate the height, in metres, of a new building. Their estimates were as follows.

50 45 62 30 40 55 110 38 52 60 55 40

(i) Find the median and the interquartile range for the data. [3]

(ii) Give a disadvantage of using the mean as a measure of the central tendency in this case. [1]

Answer:

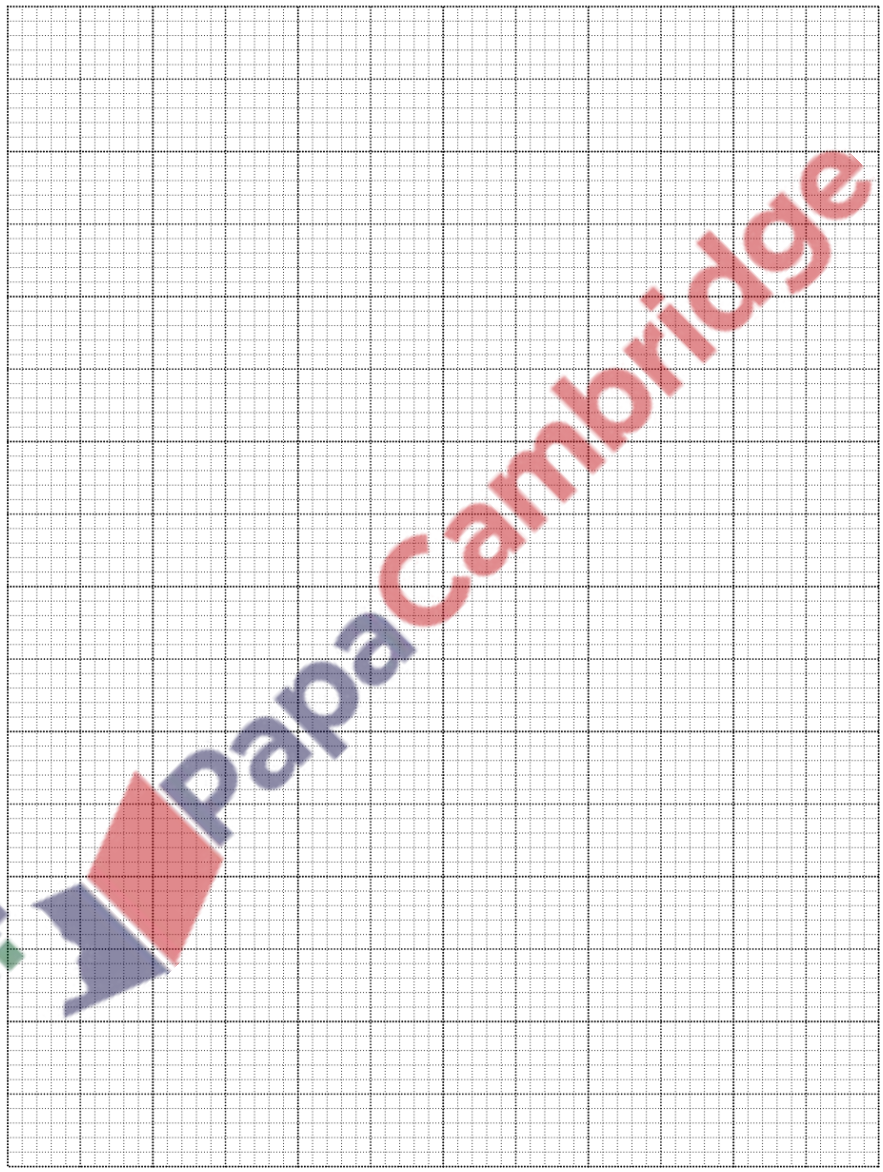
Question	Answer	Marks	Guidance
(i)	Median = 51 UQ = 57.5, LQ = 40	**B1**	
	IQR = UQ – LQ	**M1**	$55 \leqslant UQ \leqslant 62 - 38 \leqslant LQ \leqslant 45$
	17.5	**A1**	NFWW
		3	
(ii)	Result will be disproportionately affected by 110	**B1**	Affected by an extreme/large value There is a large outlier …contains outliers such as 110… Not 'mean affected by extreme values'
		1	

34. 9709_w19_qp_62 Q: 3

The speeds, in km h^{-1}, of 90 cars as they passed a certain marker on a road were recorded, correct to the nearest km h^{-1}. The results are summarised in the following table.

Speed (km h^{-1})	10 – 29	30 – 39	40 – 49	50 – 59	60 – 89
Frequency	10	24	30	14	12

(i) On the grid, draw a histogram to illustrate the data in the table. [4]



(ii) Calculate an estimate for the mean speed of these 90 cars as they pass the marker. [2]

Answer:

Question	Answer	Marks	Guidance
(i)	0.5 2.4 3 1.4 0.4	**M1**	At least 3 frequency densities calculated (frequency ÷ class width) e.g. $\left(\frac{10}{20}, \frac{10}{19} \text{ or } \frac{10}{19.5}\right)$ may be read from graph using *their* scale, 3SF or exact
	All heights correct on graph.	**A1**	
	Bar ends of 9.5, 29.5, 39.5, 59.5, 89.5	**B1**	
	Axes labelled: Frequency density (fd) and speed/km h^{-1} (or appropriate title). Linear scales 9.5 ⩽ horizontal axis ⩽ 89.5, 0 ⩽ vertical axis ⩽ 3, 5 bars with no gaps	**B1**	
		4	

Question	Answer	Marks	Guidance
(ii)	$\frac{19.5\times10+34.5\times24+44.5\times30+54.5\times14+74.5\times12}{\textit{their}\,90}$ $=\frac{195+828+1335+763+894}{90}$ $=\frac{4015}{90}$ or $\frac{803}{18}$	**M1**	Uses at least 4 midpoint attempts (e.g. 19.5 ± 0.5). Allow unsimplified expression.
	$44\frac{11}{18}$ or 44.6 (km h^{-1})	**A1**	Final answer not an improper fraction NFWW
		2	

35. 9709_w19_qp_63 Q: 5

Last Saturday, 200 drivers entering a car park were asked the time, in minutes, that it had taken them to travel from home to the car park. The results are summarised in the following cumulative frequency table.

Time (t minutes)	$t \leqslant 10$	$t \leqslant 20$	$t \leqslant 30$	$t \leqslant 50$	$t \leqslant 70$	$t \leqslant 90$
Cumulative frequency	16	50	106	146	176	200

(i) On the grid, draw a cumulative frequency graph to illustrate the data. [2]

(ii) Use your graph to estimate the median of the data. [1]

.....

.....

.....

.....

.....

.....

(iii) For 80 of the drivers, the time taken was at least T minutes. Use your graph to estimate the value of T. [2]

.....

.....

.....

.....

.....

(iv) Calculate an estimate of the mean time taken by all 200 drivers to travel to the car park. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

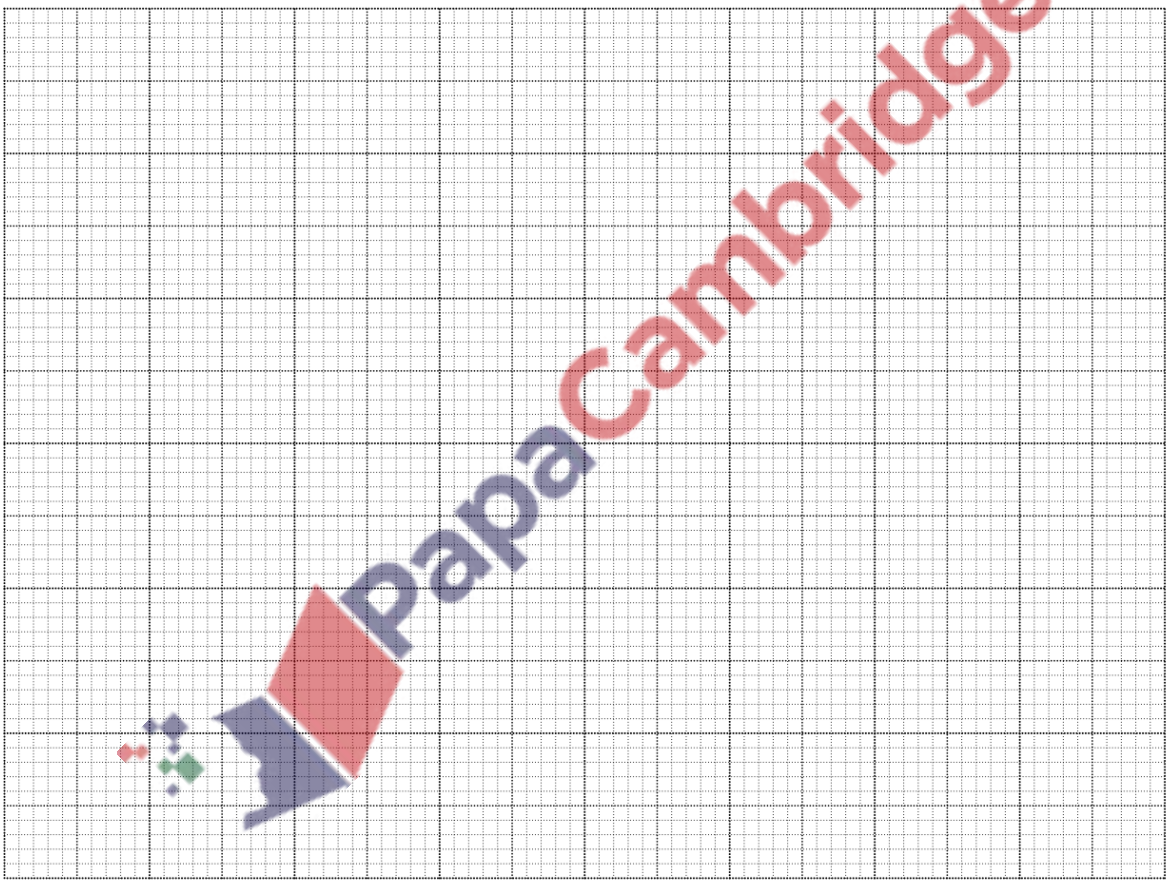
Question	Answer	Marks	Guidance
(i)	Correct labels and scales	**B1**	Axes labelled 'cumulative frequency' (or cf) and 'time (or t) [in] min(utes)', linear scales from 0 to 90 and 0 to 200 with at least 3 values marked on each axis.
	7 correctly plotted points above upper boundaries joined in a curve or line segments	**B1**	(0, 0); (10, 16); (20, 50); (30, 106); (50, 146); (70,176); (90,200)
		2	
(ii)	29	**B1**	$28 \leqslant$ median $\leqslant 30$
		1	
(iii)	120 seen	**M1**	For seeing 120 in a calculation or marked on the graph
	37	**A1FT**	$36 \leqslant$ Ans $\leqslant 39$ or FT from *their* graph **SC1** unsupported answer in range
		2	
(iv)	Frequencies 16 34 56 40 30 24	**B1**	Seen. Allow unsimplified
	Est. Mean = $\frac{5\times16+15\times34+25\times56+40\times40+60\times30+80\times24}{200}$	**M1**	At least 4 correct midpoints (5, 15, 25, 40, 60, 80) used in a calculation
	$\frac{7310}{200}$	**M1**	Summing products of *their* 6 mid-points (not lower or upper bound or class width) × *their* frequencies / 200 (or *their* $\sum$f), unsimplified
	36.55	**A1**	Accept 36.6
		4	

36. 9709_m18_qp_62 Q: 1

There are 900 students in a certain year-group. An identical puzzle is given to each student and the time taken, t minutes, to complete the puzzle is recorded. These times are summarised in the following frequency table.

Time taken, t minutes	$t \leqslant 3$	$3 < t \leqslant 4$	$4 < t \leqslant 5$	$5 < t \leqslant 6$	$6 < t \leqslant 8$	$8 < t \leqslant 10$	$10 < t \leqslant 14$
Frequency	120	180	200	160	110	80	50

On the grid, draw a cumulative frequency graph to represent the data. Use your graph to estimate the median time taken by these students to complete the puzzle. [4]



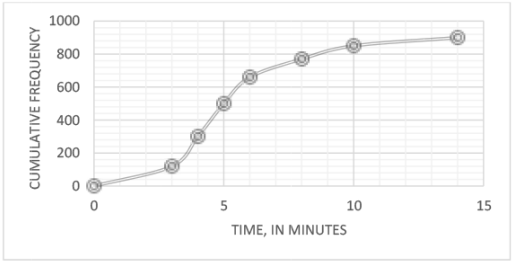
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
	 t 0 3 4 5 6 8 10 14 cf 0 120 300 500 660 770 850 900	M1 A1	Attempt to plot cumulative frequencies at ucb and all points joined between $(3,y_1)$ and $(14,y_2)$. Cf table not required. Linear scales starting at (0,0) and axes labelled cf and time in mins, all points correct; (allow straight lines or curves)
		M1	450 seen in median attempt on increasing CF graph (independent);
	Median value: 4.8 (minutes)	A1 FT	Correct ($4.7 \leqslant m < 4.9$) or FT from reading their increasing graph at cf = 450
		4	

37. 9709_m18_qp_62 Q: 5

A summary of n values of x gave the following information:

$$\Sigma(x - 20) = 136, \qquad \Sigma(x - 20)^2 = 2888.$$

The mean of the n values of x is 24.25.

(i) Find the value of n. [2]

(ii) Find Σx^2. [4]

Answer:

Question	Answer	Marks	Guidance
(i)	$24.25n - 20n = 136$ **Or** $\frac{136}{n} + 20 = 24.25$	**M1**	Unsimplified correct equation
	$n = 32$	**A1**	
		2	
(ii)	Using coded information: Variance = $\frac{2888}{32} - \left(\frac{136}{32}\right)^2$	**M1**	unsimplified expression for variance
	= 72.1875 = 72.19	**A1**	accept answers 72.2 SOI
	Using uncoded information: Variance = $\frac{\Sigma x^2}{32} - 24.25^2$ Equate with 72.1875 to give	**M1**	Equate two expressions for variance and solve
	$\Sigma x^2 = 21128$	**A1**	
		4	

38. 9709_s18_qp_61 Q: 1

In a statistics lesson 12 people were asked to think of a number, x, between 1 and 20 inclusive. From the results Tom found that $\Sigma x = 186$ and that the standard deviation of x is 4.5. Assuming that Tom's calculations are correct, find the values of $\Sigma(x - 10)$ and $\Sigma(x - 10)^2$. [3]

Answer:

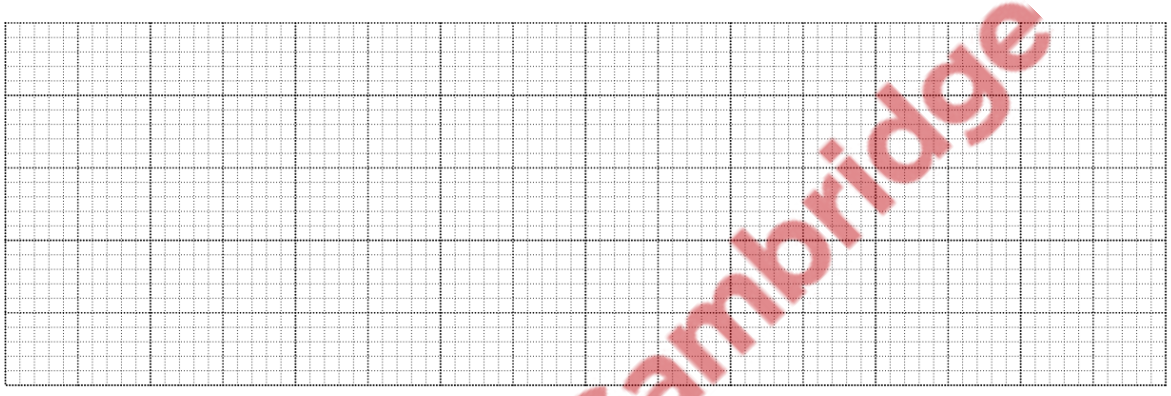
Question	Answer	Marks	Guidance
	$\Sigma(x-10) = 186 - 12 \times 10 = 66$	**B1**	Correct answer
	$\frac{\Sigma(x-10)^2}{12} - \left(\frac{\Sigma(x-10)}{12}\right)^2 = 4.5^2$	**M1**	Consistent substituting in the correct coded variance formula OR Valid method for Σx^2 then expanding $\Sigma(x-10)^2$, 3 terms with at least 2 correct
	$\Sigma(x-10)^2 = 606$	**B1**	Correct answer
		3	

39. 9709_s18_qp_61 Q: 2

In a survey 55 students were asked to record, to the nearest kilometre, the total number of kilometres they travelled to school in a particular week. The results are shown below.

5	5	9	10	13	13	13	15	15	15	15
16	18	18	18	19	19	20	20	20	20	21
21	21	21	23	25	25	27	27	29	30	33
35	38	39	40	42	45	48	50	50	51	51
52	55	57	57	60	61	64	65	66	69	70

(i) On the grid, draw a box-and-whisker plot to illustrate the data. [5]



An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

(ii) Show that there are no outliers. [2]

.....

.....

.....

.....

.....

.....

.....

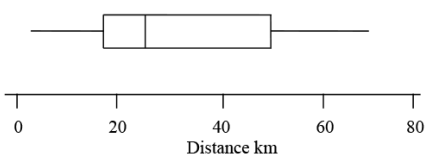
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	LQ = 18, Median = 25, UQ = 50	B1	median correct
		B1	LQ and UQ correct
		B1	Quartiles and median plotted as box graph with linear scale min 3 values
		B1ft	Whiskers drawn to correct end points with linear scale, not thr' box, not joining at top or bottom of box. Ft their UQ and LQ. Whiskers must be with ruler If scale non-linear or non-existent SCB1if all 5 data values (quartiles and end points) have values shown and all are correct numerically and fulfil the 'box' and 'whiskers ruled line' requirements
		B1	Label to include 'distance or travelled' and 'km,' allow 'total km', linear scale, numbered at least 5 – 70.
		5	

Question	Answer	Marks	Guidance
(ii)	1.5 × IQR = 48 **Method 1** LQ – 48 = –ve, (i.e. < 0) UQ + 48 = 98 (i.e. > 70)	M1	Attempt to find 1.5 × their IQR and add to UQ **or** subt from LQ
	hence no outliers	A1	Correct conclusion from correct working, need both ends. No need to state comparisons.
	Method 2 LQ – 5 = 13 (< 48) 70 – UQ = 20 (< 48)	M1	Compare their 1.5 × IQR (= 48) > gap (20) between UQ and max 70 **or** LQ and min 5
	Hence no outliers	A1	Correct conclusion from correct working, need both ends. No need to state comparisons
		2	

40. 9709_s18_qp_62 Q: 1

Each of a group of 10 boys estimates the length of a piece of string. The estimates, in centimetres, are as follows.

37 40 45 38 36 38 42 38 40 39

(i) Find the mode. [1]

(ii) Find the median and the interquartile range. [3]

Answer:

Question	Answer	Marks	Guidance
(i)	38	**B1**	
		1	
(ii)	Median = 38.5	**B1**	CAO
	IQR = 40 – 38	**M1**	$39 < UQ < 45 - 36 < LQ \leqslant 38$
	= 2	**A1**	If M0 awarded **SCB1** for **both** UQ = 40 or 40.5 **and** LQ = 38 or 37.75 seen
		3	

41. 9709_s18_qp_62 Q: 5

The lengths, t minutes, of 242 phone calls made by a family over a period of 1 week are summarised in the frequency table below.

Length of phone call (t minutes)	$0 < t \leqslant 1$	$1 < t \leqslant 2$	$2 < t \leqslant 5$	$5 < t \leqslant 10$	$10 < t \leqslant 30$
Frequency	14	46	102	a	40

(i) Find the value of a. [1]

.....

.....

.....

.....

.....

(ii) Calculate an estimate of the mean length of these phone calls. [2]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

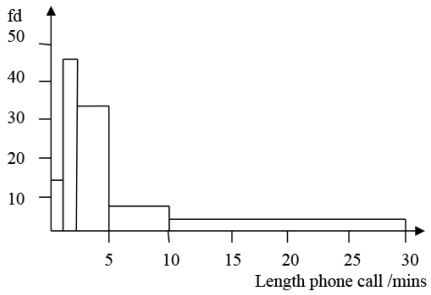
.....

.....

.....

(iii) On the grid, draw a histogram to illustrate the data in the table. [4]

Answer:

Question	Answer	Marks	Guidance
(i)	$a = 40$	**B1**	
		1	
(ii)	Mean $= \dfrac{0.5\times14+1.5\times46+3.5\times102+7.5\times\textit{their}\,40+20\times40}{242}$ $= \dfrac{1533}{242}$	**M1**	Numerator: 5 products with at least 3 acceptable mid-points × appropriate frequency FT **(i).** Denominator: 242 CAO $\dfrac{1533}{242}$ implies M1, but if **FT** an unsimplified expression required
	$= 6\dfrac{81}{242}$ or 6.33	**A1**	CAO (6.3347… rounded to 3 or more SF)
		2	
(iii)	fd = 14, 46, 34, ($\dfrac{\textit{their}\,(i)}{5}$ =) 8, 2	**M1**	Attempt at fd [f/(attempt at cw)] or scaled freq
		**A1FT**	Correct heights seen on diagram with linear vertical scale from (x, 0) **FT** their $\dfrac{a}{5}$ only
		B1	Correct bar widths (1:1:3:5:20) at axis, visually no gaps, with linear horizontal scale from ($0, y$), first bar starting at (0,0)
		B1	Labels (time, mins, and fd(OE) seen, some may be as a title) and a linear scale with at least 3 values marked on each axis. (Interval notation not acceptable)
		4	

42. 9709_s18_qp_63 Q: 1

The masses in kilograms of 50 children having a medical check-up were recorded correct to the nearest kilogram. The results are shown in the table.

Mass (kg)	10 – 14	15 – 19	20 – 24	25 – 34	35 – 59
Frequency	6	12	14	10	8

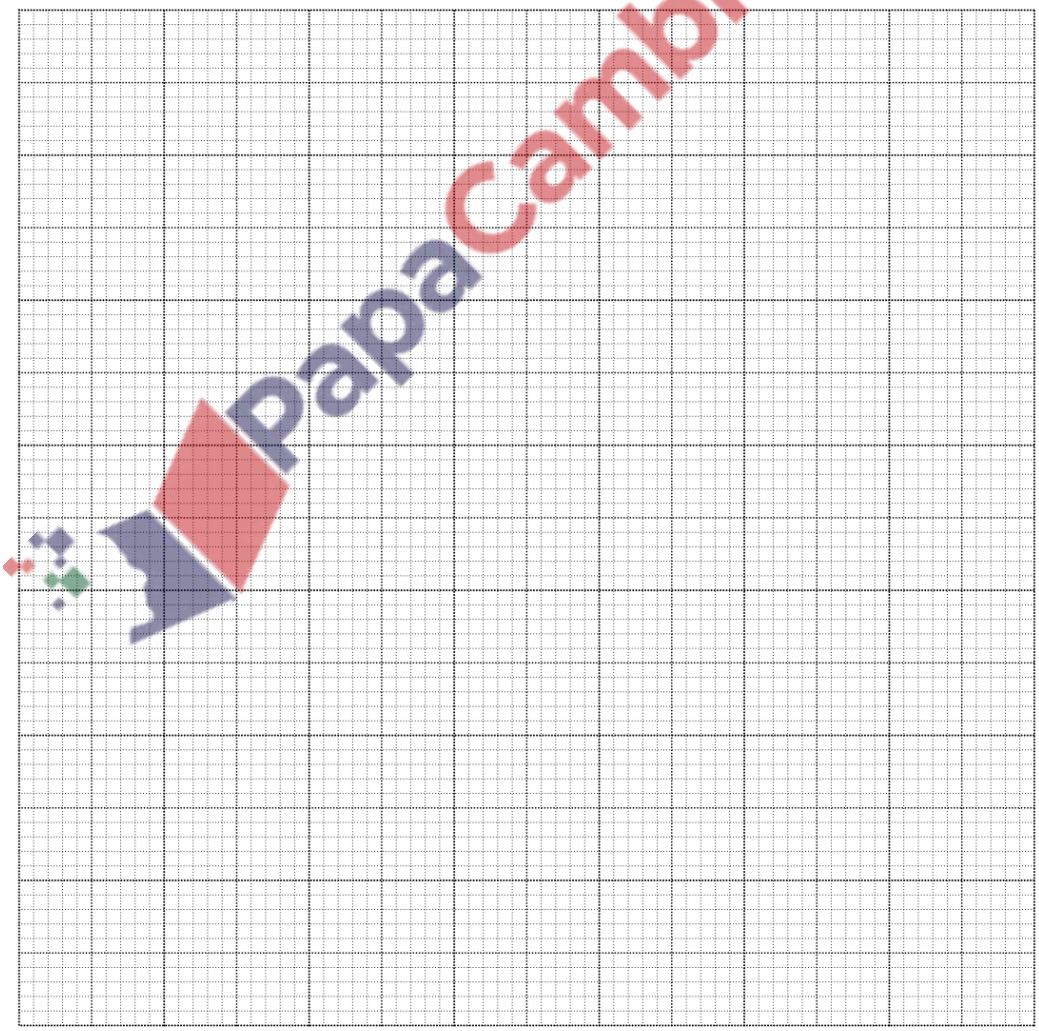
(i) Find which class interval contains the lower quartile. [1]

.....

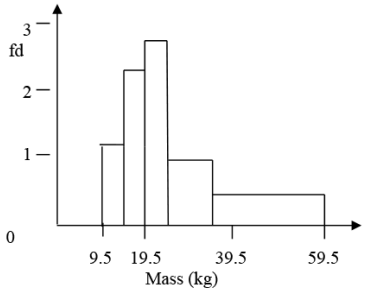
.....

.....

(ii) On the grid, draw a histogram to illustrate the data in the table. [4]



Answer:

Question	Answer	Marks	Guidance
(i)	15–19 (kg) cao	**B1**	kg not necessary; condone 14.5 – 19.5
	Total:	**1**	
(ii)	fd = 1.2, 2.4, 2.8, 1, 0.32  	**M1**	Attempt at fd [f/(attempt at cw)] or scaled freq (may be implied by 4 correct)
		A1	Correct heights seen on diagram with linear vertical scale from (x, 0)
		B1	Correct bar widths (1:1:1:2:5) visually no gaps with linear horizontal scale from (9.5,y) and first bar starting at (9.5, y)
		B1	Histogram, using attempted fds, with labels (mass, kg and fd seen) and at least 3 linearly spaced values on each axis. Horizontal axis must range from at least 9.5 to 59.5 If horizontal axis clearly starts from zero, either a break in the scale must be indicated or the scale must be linear from zero.

43. 9709_s18_qp_63 Q: 4

Farfield Travel and Lacket Travel are two travel companies which arrange tours abroad. The numbers of holidays arranged in a certain week are recorded in the table below, together with the means and standard deviations of the prices.

	Number of holidays	Mean price ($)	Standard deviation ($)
Farfield Travel	30	1500	230
Lacket Travel	21	2400	160

(i) Calculate the mean price of all 51 holidays. [2]

.....

.....

.....

(ii) The prices of individual holidays with Farfield Travel are denoted by $\$x_F$ and the prices of individual holidays with Lacket Travel are denoted by $\$x_L$. By first finding Σx_F^2 and Σx_L^2, find the standard deviation of the prices of all 51 holidays. [5]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

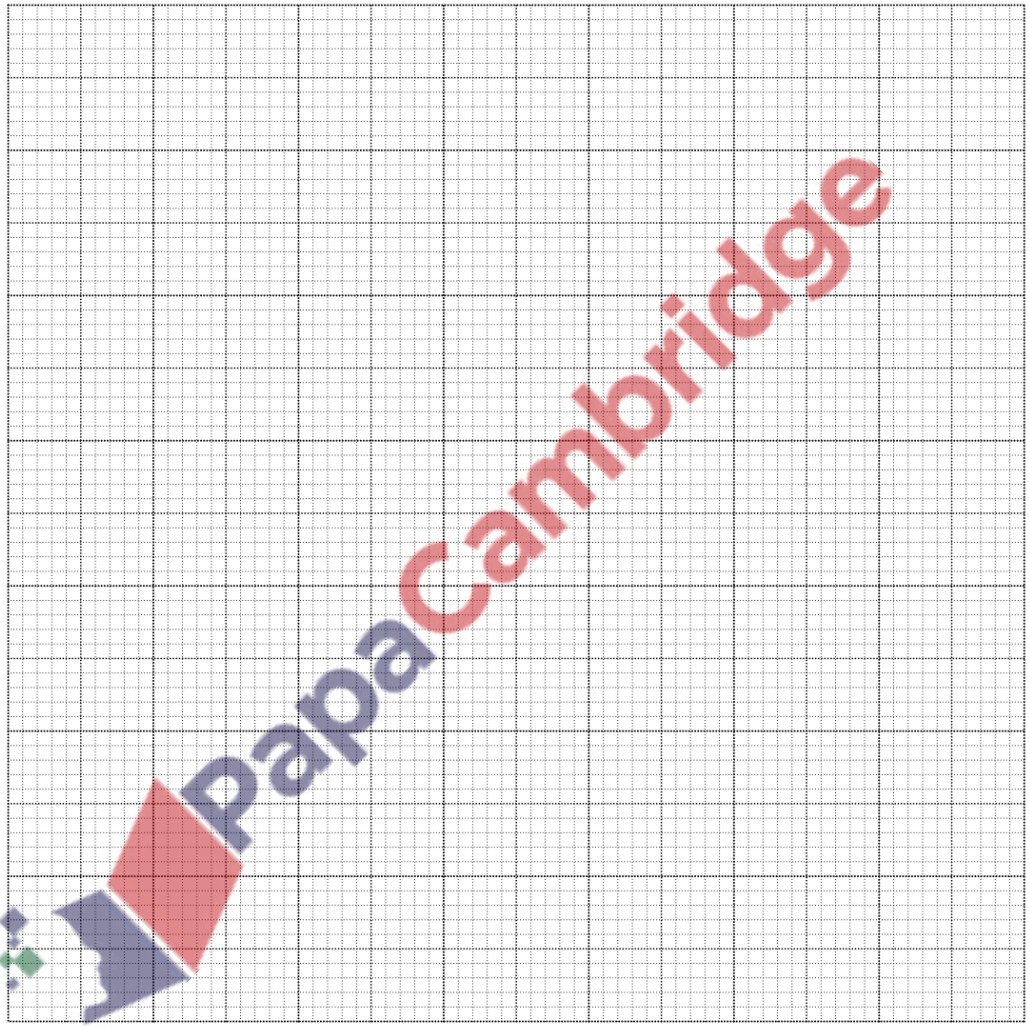
Question	Answer	Marks	Guidance
(i)	Mean = (30 × 1500 + 21 × 2400)/51	**M1**	Multiply by 30 and 21, summing and dividing total by 51 $\left(\frac{45000+50400}{51}\right)$
	= 1870 (1870.59)	**A1**	correct answer (to 3sf)
	Total:	**2**	
(ii)	$230^2 = \frac{\Sigma x_F^{\ 2}}{30} - 1500^2$ so $\Sigma x_F^{\ 2} = 69\,087\,000$	**M1**	One correct substitution into a correct variance formula
		A1	Correct Σx_F^2 (rounding to 69 000 000 2sf)
	$160^2 = \frac{\Sigma x_L^{\ 2}}{21} - 2400^2$ so $\Sigma x_L^{\ 2} = 121\,497\,600$	**A1**	Correct Σx_L^2 (rounding to 121 000 000 3sf)
	New var = $\frac{69087000+121497600}{51} - 1870.588^2 = 237\,853$	**M1**	using 'Σx_F^2'+ '' Σx_L^2 dividing by 51 and subtracting 'i' squared. (Correct 'Σx_F^2' + '' Σx_L^2 = 190 584 600)
	New sd = 488	**A1**	Correct answer accept anything between 486 and 490
	Total:	**5**	

44. 9709_w18_qp_61 Q: 6

The daily rainfall, x mm, in a certain village is recorded on 250 consecutive days. The results are summarised in the following cumulative frequency table.

Rainfall, x mm	$x \leqslant 20$	$x \leqslant 30$	$x \leqslant 40$	$x \leqslant 50$	$x \leqslant 70$	$x \leqslant 100$
Cumulative frequency	52	94	142	172	222	250

(i) On the grid, draw a cumulative frequency graph to illustrate the data. [2]



(ii) On 100 of the days, the rainfall was k mm or more. Use your graph to estimate the value of k. [2]

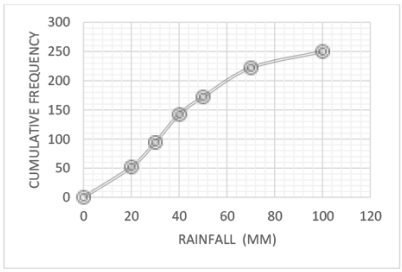
.....

.....

.....

(iii) Calculate estimates of the mean and standard deviation of the daily rainfall in this village. [6]

Answer:

Question	Answer	Marks	Guidance
(i)		B1	Appropriate linear scales starting at (0,0), axes labelled cf and Rainfall, mm
		B1	Correct graph, points plotted at ucb, allow straight lines or curve
		2	
(ii)		M1	Read off from increasing graph at cf = 150
	42	A1	Correct answer ($41 \leqslant r \leqslant 43$)
		2	

Question	Answer	Marks	Guidance
(iii)	Frequencies 52, 42, 48, 30, 50, 28	B1	Correct frequencies
	Mean age = $(10\times52+25\times42+35\times48+45\times30+60\times50+85\times28)/250$	B1	Correct midpoints (allow one error)
	$=9980/250$	M1	Using $\Sigma fx/250$ with mid-points attempt, not cf, cw, lb, ub
	$= 39.9(2)$ oe	A1	Correct answer
	Variance = $10^2\times52+25^2\times42+35^2\times48+45^2\times30+60^2\times50+85^2\times28)/250-\text{mean}^2$ $= 539.59$	M1	Attempt at variance using their midpoints and their mean
	$\sigma = 23.2$	A1	Correct answer for sd
		6	

45. 9709_w18_qp_62 Q: 2

The following back-to-back stem-and-leaf diagram shows the reaction times in seconds in an experiment involving two groups of people, *A* and *B*.

A			*B*	
(4)	4 2 0 0	20	5 6 7	(3)
(5)	9 8 5 0 0	21	1 2 2 3 7 7	(6)
(8)	9 8 7 5 3 2 2 2	22	1 3 5 6 6 8 9	(7)
(6)	8 7 6 5 2 1	23	4 5 7 8 8 9 9 9	(8)
(3)	8 6 3	24	2 4 5 6 7 8 8	(7)
(1)	0	25	0 2 7 8	(4)

Key: 5 | 22 | 6 means a reaction time of 0.225 seconds for *A* and 0.226 seconds for *B*

(i) Find the median and the interquartile range for group *A*. [3]

.....

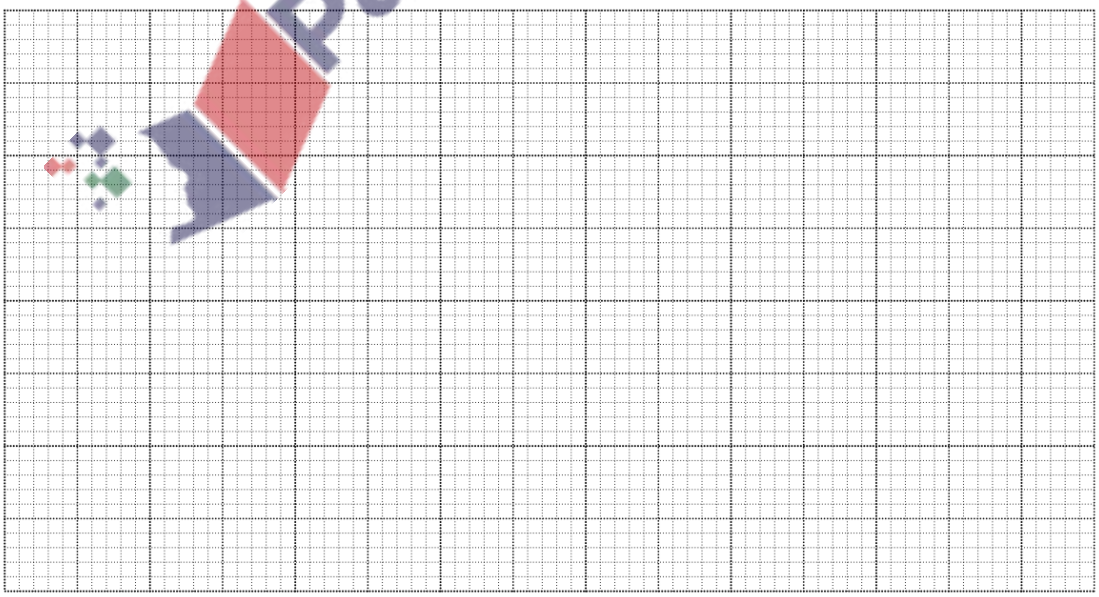
.....

.....

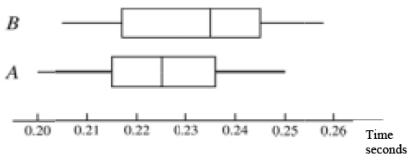
.....

The median value for group *B* is 0.235 seconds, the lower quartile is 0.217 seconds and the upper quartile is 0.245 seconds.

(ii) Draw box-and-whisker plots for groups *A* and *B* on the grid. [3]



Answer:

Question	Answer	Marks	Guidance
(i)	median = 0.225; LQ = 0.215: UQ = 0.236	**B1**	Correct median (Q_2)
	IQR = 0.236 – 0.215	**M1**	0.232 < UQ (Q_3) < 0.238 – 0.204 < LQ (Q_1) < 0.219
	= 0.021	**A1**	www Omission of all decimal points **MR-1** If M0 awarded **SCB1** for both LQ = 0.215: UQ = 0.236 seen
		3	
(ii)		**B1**	Linear scale between 0.20 to 0.26 (condone omission of 0.26) axis labelled (time and seconds), at least one box plot attempted, no lines through boxes, whiskers not at corner of boxes
		B1 ft	Labelled correct graph for A, (ft their median/quartiles), condone lines through boxes, whiskers at corner of boxes
	<table><tr><td>**A**</td><td>0.200</td><td>0.215</td><td>0.225</td><td>0.236</td><td>0.250</td></tr><tr><td>**B**</td><td>0.205</td><td>0.217</td><td>0.235</td><td>0.245</td><td>0.258</td></tr></table>	**B1**	Labelled correct graph for B, condone lines through boxes, whiskers at corner of boxes **SC** If B0B0 scored because graphs not labelled/labels reversed SCB1 if both 'correct' Penalty **MR-1** if graphs plotted on separate axes unless both scales align exactly.
		3	

46. 9709_w18_qp_62 Q: 5

The Quivers Archery club has 12 Junior members and 20 Senior members. For the Junior members, the mean age is 15.5 years and the standard deviation of the ages is 1.2 years. The ages of the Senior members are summarised by $\Sigma y = 910$ and $\Sigma y^2 = 42\,850$, where y is the age of a Senior member in years.

(i) Find the mean age of all 32 members of the club. [2]

(ii) Find the standard deviation of the ages of all 32 members of the club. [4]

Answer:

Question	Answer	Marks	Guidance
(i)	$\frac{15.5\times12+910}{12+20}$	**M1**	Unsimplified total age divided by *their* total members (not 12, 20 or 2)
	=34.25 or 34¼ (years)	**A1**	Correct exact answer (isw rounding), oe (34 years 3 months)
		2	
(ii)	Considering Juniors: variance $=\frac{\Sigma x^2}{12}-15.5^2=1.2^2$	**M1**	$\frac{\Sigma x^2}{k}-15.5^2=1.2^2$, $k=12$ or 20
	$\Sigma x^2=2900.28$	**A1**	Answer wrt 2900
	Considering whole group: $\Sigma z^2=\Sigma x^2+\Sigma y^2=2900.28+42850=45750$ Variance $=\frac{\Sigma z^2}{32}-\mu^2=\frac{\textit{their}\ 45750}{12+20}-(\textit{their}\ 34.25)^2$ $(=256.63)$	**M1**	*Their* 45750 > 42850 (not 85700 or rounding to 1.8×10^9) in correct variance or std deviation formula (Σx^2 and addition may not be seen)
	s d = 16.0(2)	**A1**	Correct final answer, condone 16.03
		4	

47. 9709_w18_qp_63 Q: 7

The heights, in cm, of the 11 members of the Anvils athletics team and the 11 members of the Brecons swimming team are shown below.

Anvils	173	158	180	196	175	165	170	169	181	184	172
Brecons	166	170	171	172	172	178	181	182	183	183	192

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

(ii) Find the median and the interquartile range for the heights of the Anvils. [3]

.....

.....

.....

.....

The heights of the 11 members of the Anvils are denoted by x cm. It is given that $\Sigma x = 1923$ and $\Sigma x^2 = 337\,221$. The Anvils are joined by 3 new members whose heights are 166 cm, 172 cm and 182 cm.

(iii) Find the standard deviation of the heights of all 14 members of the Anvils. [4]

If you use the following lined page to complete the answer(s) to any question(s), the question number(s) must be clearly shown.

Answer:

Question	Answer	Marks	Guidance
(i)	Anvils | Brecons 8 | 15 9 5 | 16 | 6 5 3 2 0 | 17 | 0 1 2 2 8 4 1 0 | 18 | 1 2 3 3 6 | 19 | 2 Key: 5|16|6 means 165 cm for Anvils and 166 cm for Brecons	B1	Correct stem, up or down
		B1	Correct Anvils labelled on left, leaves in order from right to left and lined up vertically, no commas
		B1	Correct Brecons labelled on same diagram on right hand side in order from left to right and lined up vertically, no commas
		B1	Correct key, not split, both teams, at least one with cm
		4	
(ii)	Median = 173	B1	Correct median (or Q2)
	LQ = 169; UQ = 181 IQR = 181 – 169	M1	Either UQ = 181 ± 4, or LQ = 169 ± 4 and evaluating UQ – LQ
	= 12	A1	Correct answer from 181 and 169 only
		3	

Question	Answer	Marks	Guidance
(iii)	$\Sigma x = 1923 + 166 + 172 + 182$ (= 2443) $\Sigma x^2 = 337221 + 166^2 + 172^2 + 182^2$ (= 427485)	M1	Correct unsimplified expression for Σx and Σx^2, may be implied
	Mean = $\frac{\Sigma x}{14} = \frac{2443}{14} = 174.5$	M1	Correct unsimplified mean
	Variance = $\frac{\Sigma x^2}{14} - \left(\frac{\Sigma x}{14}\right)^2 = \frac{427485}{14} - \left(\frac{2443}{14}\right)^2$	M1	Correct unsimplified variance using 14, their Σx and their Σx^2, not using 1923 and/or 337221
	S d = 9.19	A1	Correct answer
		4	

48. 9709_m17_qp_62 Q: 1

Twelve values of x are shown below.

1761.6	1758.5	1762.3	1761.4	1759.4	1759.1
1762.5	1761.9	1762.4	1761.9	1762.8	1761.0

Find the mean and standard deviation of $(x - 1760)$. Hence find the mean and standard deviation of x. [4]

Answer:

Question	Answer	Marks	Guidance
	1.6 –1.5 2.3 1.4 –0.6 –0.9 2.5 1.9 2.4 1.9 2.8 1.0	**M1**	Subtracting 1760, allow max 2 slips
	Mean = 1.23	**A1**	
	sd = 1.39	**A1**	
	Mean of x = 1761.23 , sd of x = 1.39	**A1**√	ft their coded mean and sd.
			SR B1 correct mean and sd without use of coded process
	Total:	**4**	

49. 9709_s17_qp_61 Q: 1

Kadijat noted the weights, x grams, of 30 chocolate buns. Her results are summarised by

$$\Sigma(x-k) = 315, \qquad \Sigma(x-k)^2 = 4022,$$

where k is a constant. The mean weight of the buns is 50.5 grams.

(i) Find the value of k. [2]

(ii) Find the standard deviation of x. [2]

Answer:

Question	Answer	Marks	Guidance
(i)	*EITHER:* $\frac{\sum x}{30} - k = \frac{315}{30} = 10.5$	**(M1**	Dividing 315 by ±30 and + or – from 50.5 need both and no more
	$k = 5.5 - 10.5 = 40$	**A1)**	Correct answer from correct working
	OR: $\sum x = 50.5 \times 30 = 1515$, $1515 - 30k = 315$	**(M1**	Mult by 50.5 by 30 and + or – 315 and dividing by ±30 need all these
	$k = 40$	**A1)**	Correct answer from correct working. 1200 gets **M0**
	Total:	**2**	
(ii)	*EITHER:* var = 4022/30–10.5^2(=23.817)	**(M1**	Subst in correct coded variance formula
	sd = 4.88	**A1)**	
	OR: $\sum x^2 - 2(40)\sum x + 30(40)^2 = 4022$, $\sum x^2 = 77222$ Var = 77222/30 – 50.5^2 (= 23.817)	**(M1**	Expanding with ± 40Σx and ± 30(40)^2 seen
	sd = 4.88	**A1)**	
	Total:	**2**	

50. 9709_s17_qp_61 Q: 4

The times taken, t seconds, by 1140 people to solve a puzzle are summarised in the table.

Time (t seconds)	$0 \leqslant t < 20$	$20 \leqslant t < 40$	$40 \leqslant t < 60$	$60 \leqslant t < 100$	$100 \leqslant t < 140$
Number of people	320	280	220	220	100

(i) On the grid, draw a histogram to illustrate this information. [4]

(ii) Calculate an estimate of the mean of t. [2]

.....

.....

.....

.....

.....

.....

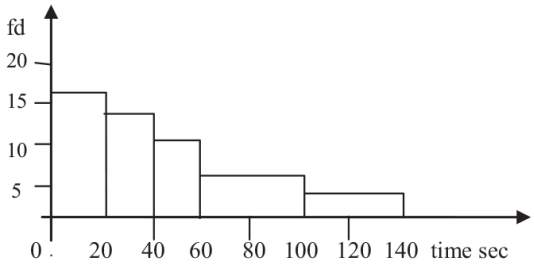
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	fd 16, 14, 11, 505, 2.5	**M1**	Attempt at fd (must be at least 3 freq/cw) – may be implied by graph
		**A1**	Correct heights seen on graph i.e. must see a gap for fd = 2.5 etc.
		B1	Correct end points of bars and correct widths
		B1	labels fd, sec. Time can be optional. Linear axes, condone $0 \leqslant t < 20$ etc.
	Total:	**4**	

Question	Answer	Marks	Guidance
(ii)	$(10 \times 320 + 30 \times 280 + 50 \times 220 + 80 \times 220 + 120 \times 100) / 1140$	**M1**	using $\Sigma fx / n$ with mid-point attempt ±0.5, not ends not class widths
	$= 45.8$	**A1**	
	Total:	**2**	

51. 9709_s17_qp_62 Q: 1

Rani and Diksha go shopping for clothes.

(i) Rani buys 4 identical vests, 3 identical sweaters and 1 coat. Each vest costs \$5.50 and the coat costs \$90. The mean cost of Rani's 8 items is \$29. Find the cost of a sweater. [3]

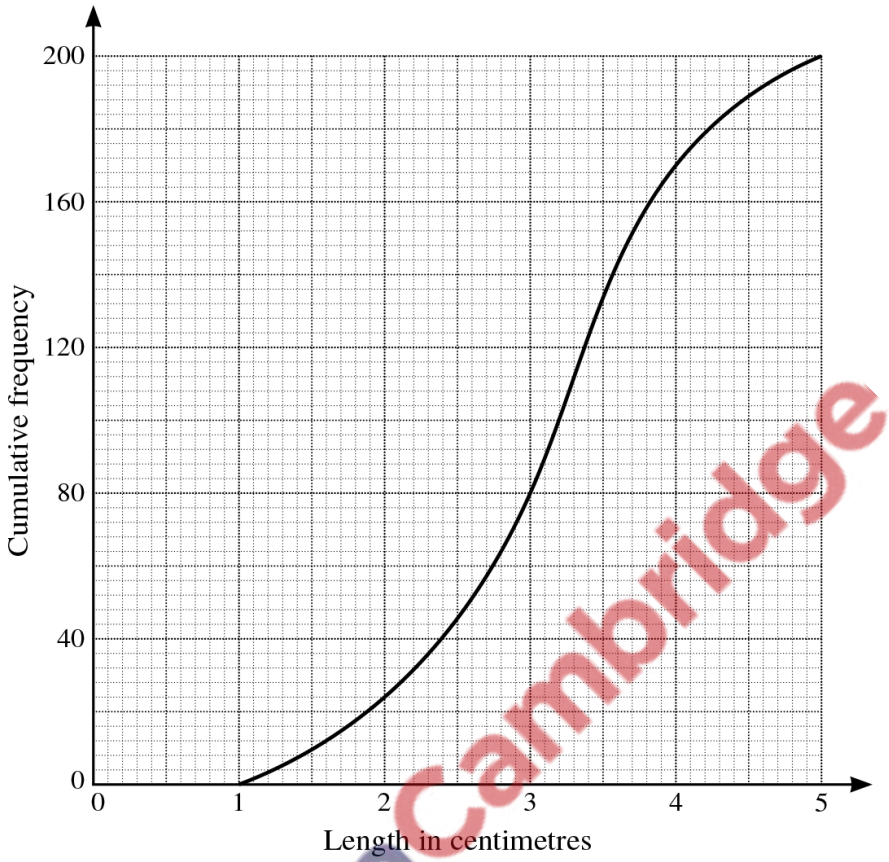
(ii) Diksha buys 1 hat and 4 identical shirts. The mean cost of Diksha's 5 items is \$26 and the standard deviation is \$0. Explain how you can tell that Diksha spends \$104 on shirts. [2]

Answer:

Question	Answer	Marks	Guidance
(i)	$4 \times 5.5 + 3x + 90 = 8 \times 29$	**M1**	An expression to work out total cost of individual items = 8 × mean, x may be implied.
	$112 + 3x = 232$ $x = 40$	**A1**	Correct complete unsimplified expression / calculation
	(Cost = $)40	**A1**	Units not required
	Total:	**3**	
(ii)	sd = 0 so all cost the same	**M1**	Must see comment interpreting sd = 0, OE
	shirts cost 4 × $26 = $104 **AG**	**A1**	See 4 × $26, $130 – $26 OE. Must have a final value of $104 stated
	Total:	**2**	

52. 9709_s17_qp_62 Q: 2

Anabel measured the lengths, in centimetres, of 200 caterpillars. Her results are illustrated in the cumulative frequency graph below.



(i) Estimate the median and the interquartile range of the lengths. [3]

.....

.....

.....

(ii) Estimate how many caterpillars had a length of between 2 and 3.5 cm. [1]

.....

.....

(iii) 6% of caterpillars were of length l centimetres or more. Estimate l. [2]

.....

.....

.....

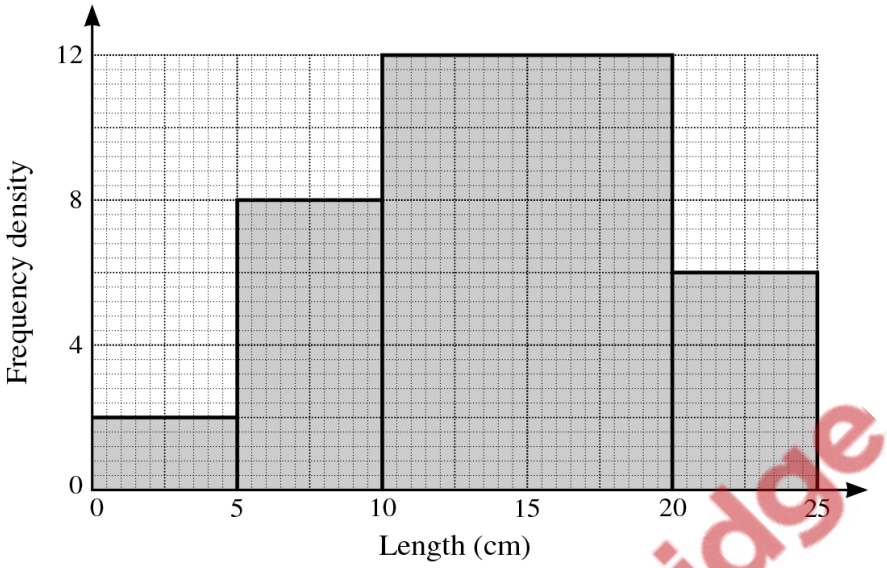
Answer:

(i)	med = 3.2	**B1**	Accept 3.2 ± 0.05
	UQ = 3.65 ⩽ uq ⩽ 3.7 LQ = 2.55⩽ lq ⩽ 2.6	**M1**	UQ – LQ, UQ greater than *their* 'median', LQ less than *their* 'median'
	IQR = 1.05 ⩽ iqr ⩽ 1.15	**A1**	Correct answer from both LQ and UQ in given ranges
	Total:	**3**	
(ii)	134 – 24 = 110	**B1**	Accept 108 ⩽ n ⩽ 112, n an integer
	Total:	**1**	

Question	Answer	Marks	Guidance
(iii)	200 – 12 = 188 less than length l	**M1**	188 seen, can be implied by answer in range, mark on graph.
	l = 4.5 cm	**A1**	Correct answer accept 4.4 ⩽ l ⩽ 4.5
	Total:	**2**	

53. 9709_s17_qp_63 Q: 7

The following histogram represents the lengths of worms in a garden.



(i) Calculate the frequencies represented by each of the four histogram columns. [2]

.....

.....

.....

.....

.....

.....

.....

.....

(ii) On the grid on the next page, draw a cumulative frequency graph to represent the lengths of worms in the garden. [4]

(iii) Use your graph to estimate the median and interquartile range of the lengths of worms in the garden. [3]

.....

.....

.....

.....

.....

.....

.....

.....

.....

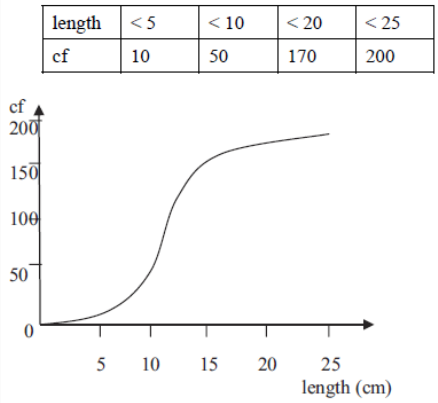
.....

.....

(iv) Calculate an estimate of the mean length of worms in the garden. [2]

Answer:

(i)	freq = fd × cw 10, 40, 120, 30	**M1** **A1**	Attempt to multiply at least 3 fds by their 'class widths'
	Totals:	**2**	

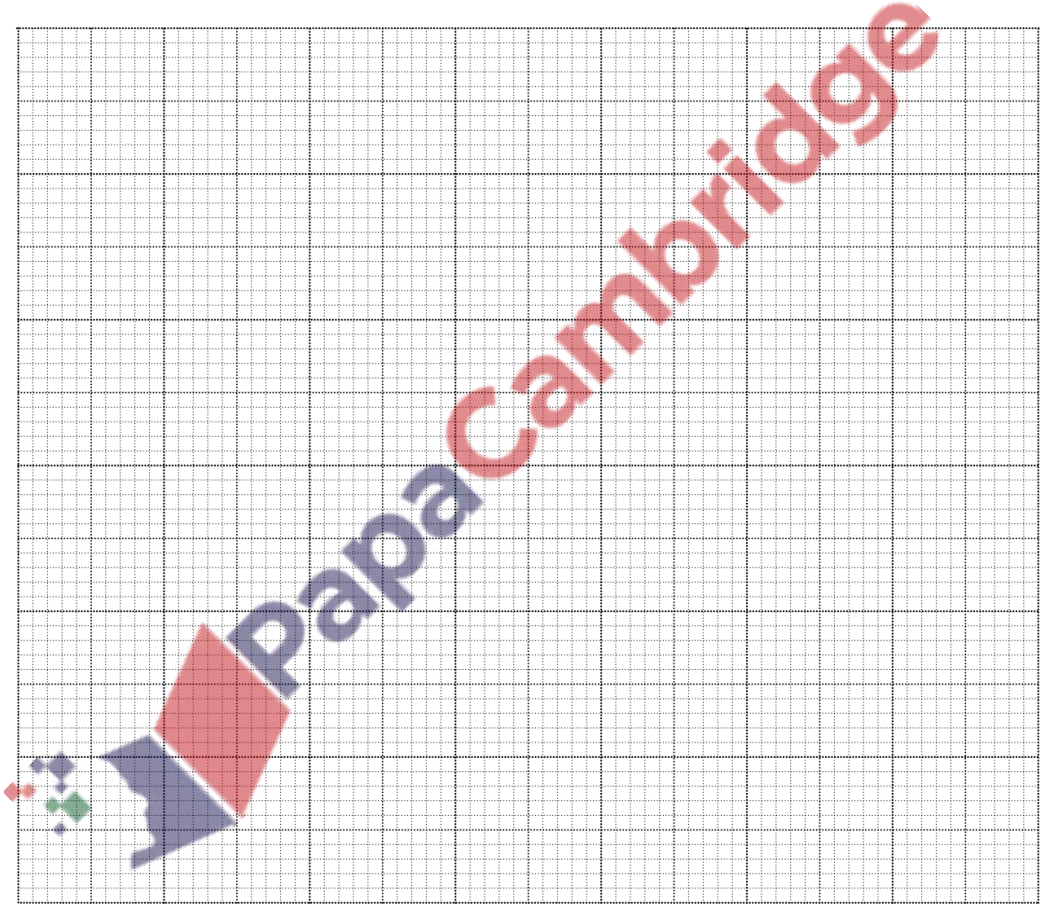
Question	Answer	Marks	Guidance
(ii)		**B1** **B1** **M1** **A1**	3 or more correct cfs **heights** on graph 10, 50, 170, 200 Labels correct cf and length(cm), linear scales from zero (allow 0.5 on horizontal axis) Attempt (at least three) at plotting at upper end points (either 5 or 5.5, 10 or 10.5 etc.) Starting at (0, 0) polygon or smooth curve increasing with plotted points at lengths 5, 10, 20 and 25
	Totals:	**4**	
(iii)	median = 14.2	**B1**	Median (accept 13.2 – 15.2)
	'18.5' – '10'	**M1**	Subt their LQ from their UQ if reasonable from their graph
	IQ range = 8.5	**A1FT**	Correct FT using LQ = 10 and UQ between 17.5 and 19.5
	Totals:	**3**	
(iv)	mean = (2.5×10 + 7.5×40 + 15×120 + 22.5×30) / 200	**M1**	Using mid points (± 0.5) and their frequencies from 7(i) in correct formula
	= 14	**A1**	
	Totals:	**2**	

54. 9709_w17_qp_61 Q: 2

The time taken by a car to accelerate from 0 to 30 metres per second was measured correct to the nearest second. The results from 48 cars are summarised in the following table.

Time (seconds)	3 – 5	6 – 8	9 – 11	12 – 16	17 – 25
Frequency	10	15	17	4	2

(i) On the grid, draw a cumulative frequency graph to represent this information. [3]



(ii) 35 of these cars accelerated from 0 to 30 metres per second in a time more than t seconds. Estimate the value of t. [2]

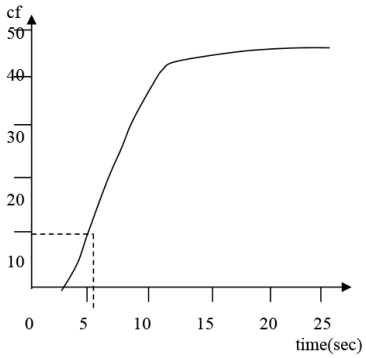
.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	Points (5.5,10), (8.5,25), (11.5,42), (16.5,46), (25.5,48)  	**B1**	Correct cfs values seen listed, in or by table or on graph, 0 not required
		B1	Axes labelled "cumulative frequency" (or cf) and "time [or t etc.] (in) seconds (or sec etc.)". Linear scales – cf 0–48, time 2.5 – 25.5 (ignore <2.5 on time.) At least 3 values stated on each axis, but (0,0) can be implied without stating.
		B1	All points plotted accurately, (5, 10) etc. scores **B0**. Curve or line segments drawn starting at (5.5,10) and passing within '1 scale unit' vertically and horizontally of plotted points
		3	

Question	Answer	Marks	Guidance
(ii)	$48 - 35 = 13$ $t = 6.5$ sec	**M1**	Subt 35 (checked ±1 mm on graph) from 48 or 50,
		A1	$6 \leqslant$ Ans $\leqslant 7$
		2	

55. 9709_w17_qp_61 Q: 4

The ages of a group of 12 people at an Art class have mean 48.7 years and standard deviation 7.65 years. The ages of a group of 7 people at another Art class have mean 38.1 years and standard deviation 4.2 years.

(i) Find the mean age of all 19 people. [2]

.....

.....

.....

.....

.....

.....

(ii) The individual ages in years of people in the first Art class are denoted by x and those in the second Art class by y. By first finding Σx^2 and Σy^2, find the standard deviation of the ages of all 19 people. [4]

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	$\frac{(48.7\times12+38.1\times7)}{19}$	**M1**	Accept unsimplified (may be separate calculations)
	$= 44.8$	**A1**	
		2	
(ii)	$7.65^2 = \frac{\Sigma x^2}{12} - 48.7^2$ $\Sigma x^2 = 29162.55$	**M1**	Substitution in one correct variance formula
	$4.2^2 = \frac{\Sigma y^2}{7} - 38.1^2$ $\Sigma y^2 = 10284.75$	**A1**	One Σx^2 or Σy^2 correct (can be rounded to 4sf))
	Combined var $= \frac{(29162.55+10284..75)}{19} - 44.79^2$ $= \frac{39447.3}{19} - 44.79^2$	**M1**	Using their Σx^2 and Σy^2 and their **4(i)** in the variance formula
	Combined $\sigma = 8.37$ or 8.36	**A1**	
		4	

56. 9709_w17_qp_62 Q: 1

Andy counts the number of emails, x, he receives each day and notes that, over a period of n days, $\Sigma(x - 10) = 27$ and the mean number of emails is 11.5. Find the value of n. [3]

Answer:

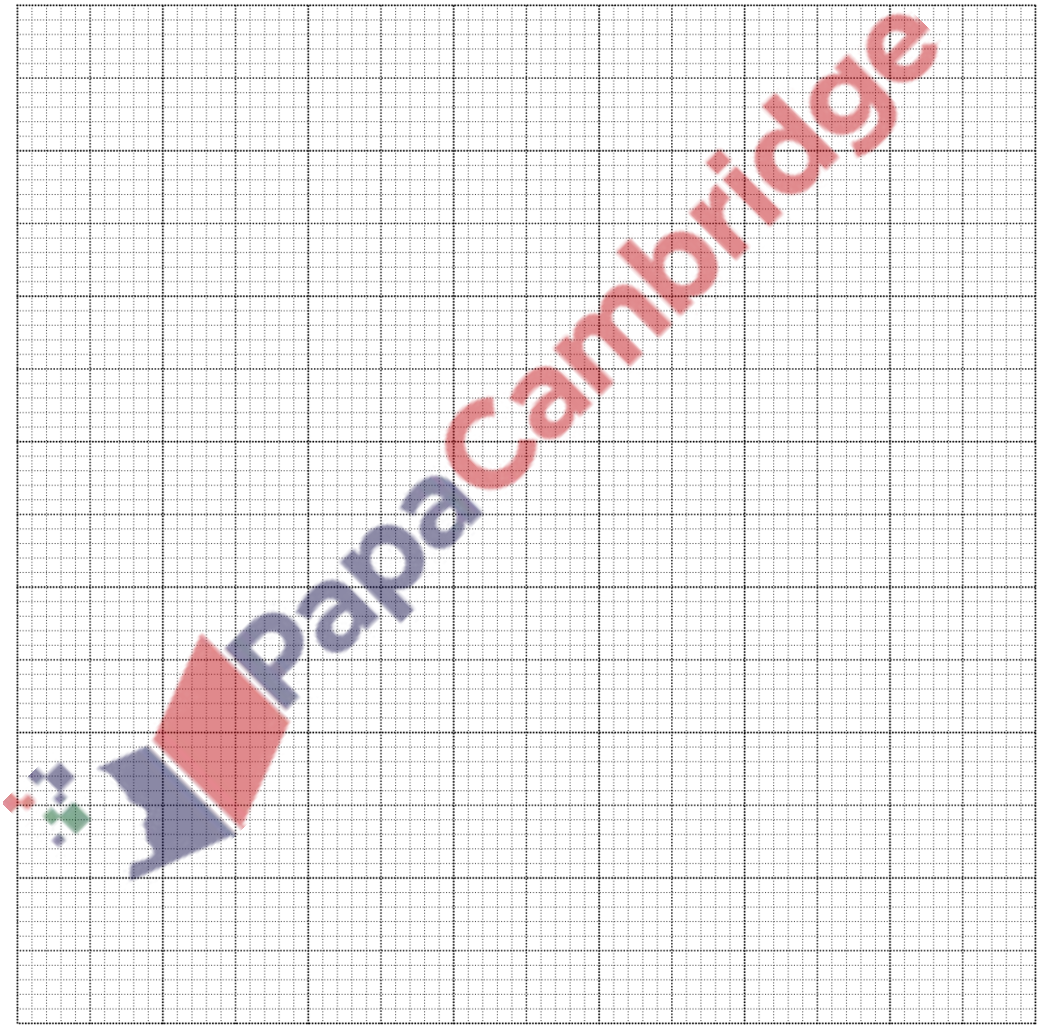
Question	Answer	Marks	Guidance
	EITHER: $(\Sigma x =)\ 11.5n = 27 + 10n$	**(M1**	Expanding brackets and forming a three term equation involving 27 and at least one term in n, without x
		M1	$10n$ or $11.5n$ seen in expression without x ($1.5n = 27$ implies **M2**)
	$n = 18$	**A1)**	
	OR: $11.5 = \frac{27}{n} + 10$	**(M1**	Dividing coded sum by n and forming a three term equation involving 11.5 and at least one term in n, without x
		M1	$27/n$ seen in expression without x ($1.5 = \frac{27}{n}$ implies **M2**)
	$n = 18$	**A1)**	
		3	

57. 9709_w17_qp_62 Q: 2

The circumferences, c cm, of some trees in a wood were measured. The results are summarised in the table.

Circumference (c cm)	$40 < c \leqslant 50$	$50 < c \leqslant 80$	$80 < c \leqslant 100$	$100 < c \leqslant 120$
Frequency	14	48	70	8

(i) On the grid, draw a cumulative frequency graph to represent the information. [3]



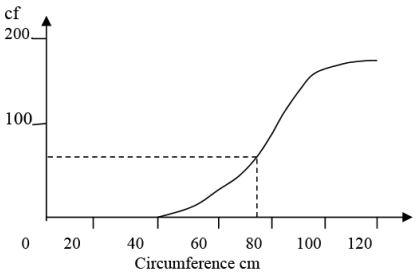
(ii) Estimate the percentage of trees which have a circumference larger than 75 cm. [2]

.....

.....

.....

Answer:

Question	Answer	Marks	Guidance
(i)	points (50, 14), (80, 62), (100, 132), (120, 140)	**B1**	Correct cfs values seen listed, in or by table or on graph, 0 not required
		**B1**	Axes labelled 'cumulative frequency' (or cf) and 'circumference [or cir or c etc.] (in) cm'. Linear scales – c.f. 0–140 circumference 40–120 (ignore <40 on circ.) At least 3 values stated on each axis, but (0,0) can be implied without stating.
		B1	All points plotted accurately
		3	
(ii)	140 – 54 = 86	**M1**	Finding correct value from graph (checked ±1 mm) or linear interpolation. Subtraction from 140 can be implied
	Percentage = 61.4%	**A1**	60.5% ⩽ Ans ⩽ 64.5%
		2	

58. 9709_w17_qp_63 Q: 2

Tien measured the arm lengths, x cm, of 20 people in his class. He found that $\Sigma x = 1218$ and the standard deviation of x was 4.2. Calculate $\Sigma(x - 45)$ and $\Sigma(x - 45)^2$. [3]

Answer:

Question	Answer	Marks	Guidance
	$\Sigma(x-45) = 1218 - 20 \times 45 = 318$	B1	
	$\frac{\Sigma(x-45)^2}{20} - \left(\frac{\Sigma(x-45)}{20}\right)^2 = 4.2^2$	M1	Fully correct substitution in the correct coded variance formula with their $\Sigma(x-45)$ **OR** valid method for $\Sigma x^2 = 74\,529$ ($4.2^2 = \frac{\Sigma x^2}{20} - \left(\frac{1218}{20}\right)^2$) and expanding $\Sigma(x-45)^2$ correctly $= \Sigma x^2 - 90\Sigma x + 20 \times 45^2 = '74\,529' - 90 \times 1218 + 40\,500 = 5409$
	$\Sigma(x-45)^2 = 5409$	A1	
		3	

59. 9709_w17_qp_63 Q: 5

The number of Olympic medals won in the 2012 Olympic Games by the top 27 countries is shown below.

104	88	82	65	44	38	35	34	28
28	18	18	17	17	14	13	13	12
12	10	10	10	9	6	5	2	2

(i) Draw a stem-and-leaf diagram to illustrate the data. [4]

(ii) Find the median and quartiles and draw a box-and-whisker plot on the grid. [5]

.....

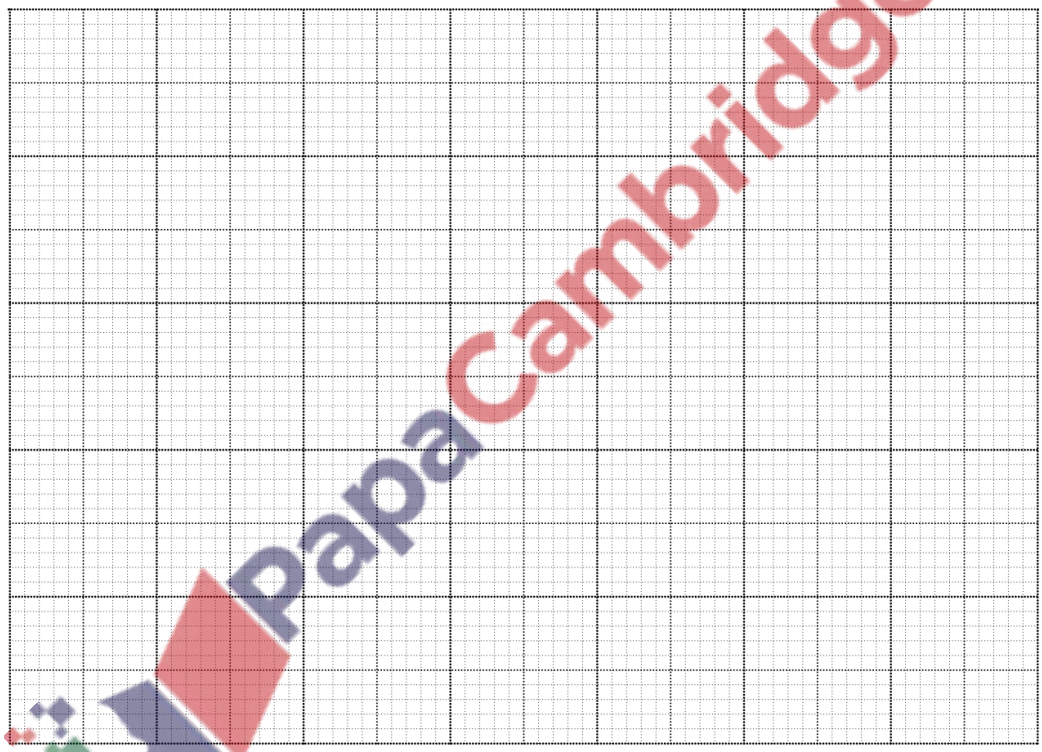
.....

.....

.....

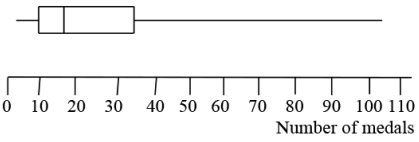
.....

.....



Answer:

Question	Answer	Marks	Guidance
(i)	0 | 2 2 5 6 9 1 | 0 0 0 2 2 3 3 4 7 7 8 8 2 | 8 8 3 | 4 5 8 4 | 4 5 | 6 | 5 7 | 8 | 2 8 9 | 10 | 4 key 2 | 8 means 28 medals	**B1**	Stem, digits 5, 7, 9 can be missing here, can be upside down
		B1	All leaves in correct order increasing from stem, (5, 7 and 9 can be missing), condone commas
		B1	Reasonable shape, requires all values of the stem, only one line for each stem and leaves must be lined up. Can be upside down or sideways. No commas. Condone one 'leaf' error.
		B1	Correct key must state 'medals' or have 'medals' in leaf heading or title
		4	

Question	Answer	Marks	Guidance
(ii)	Med = 17 LQ = 10 UQ = 35  	**B1**	Median correct
		B1	LQ and UQ correct
		B1	Uniform scale from 2 to 104 (need 3 identified points min) and label including medals (can be in title)
		B1 FT	Correct box med and quartiles on diagram, FT their values
		B1	Correct end-whiskers from ends of box but not through box
		5	

60. 9709_m16_qp_62 Q: 1

For 10 values of x the mean is 86.2 and $\Sigma(x - a) = 362$. Find the value of

(i) Σx, [1]

(ii) the constant a. [2]

Answer:

(i)	$\Sigma x = 862$	**B1** 1	Must be stated or replaced in (ii) Can see **(i)** and **(ii)** in any order
(ii)	$362/10 + a = 86.2$ $a = 50$	**M1** **A1** 2	86.2 ± 36.2 seen oe Correct answer, nfww

61. 9709_m16_qp_62 Q: 4

A survey was made of the journey times of 63 people who cycle to work in a certain town. The results are summarised in the following cumulative frequency table.

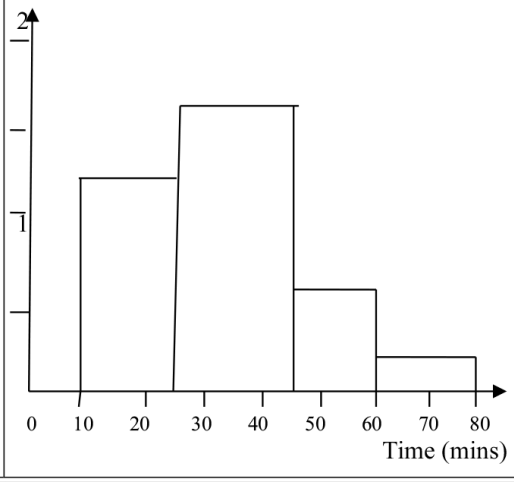
Journey time (minutes)	$\leqslant 10$	$\leqslant 25$	$\leqslant 45$	$\leqslant 60$	$\leqslant 80$
Cumulative frequency	0	18	50	59	63

(i) State how many journey times were between 25 and 45 minutes. [1]

(ii) Draw a histogram on graph paper to represent the data. [4]

(iii) Calculate an estimate of the mean journey time. [2]

Answer:

(i)	32	**B1**	1	
(ii)	freqs 0 18 32 9 4 fd 0 1.2 1.6 0.6 0.2 cf  	**M1** **A1** **B1** **B1**	 4	attempt at fd or scaled freq (at least 3 f/cw attempt) correct heights seen on diagram Correct bar ends Labels fd and time (mins) and linear axes or squiggle
(iii)	$(17.5 \times 18 + 35 \times 32 + 52.5 \times 9 + 70 \times 4)/63$ $= 2187.5/63 = 34.7$	**M1** **A1**	 2	$\Sigma fx/63$ where x is midpoint attempt not end pt or cw Correct answer

62. 9709_s16_qp_61 Q: 7

The amounts spent by 160 shoppers at a supermarket are summarised in the following table.

Amount spent ($\$x$)	$0 < x \leqslant 30$	$30 < x \leqslant 50$	$50 < x \leqslant 70$	$70 < x \leqslant 90$	$90 < x \leqslant 140$
Number of shoppers	16	40	48	26	30

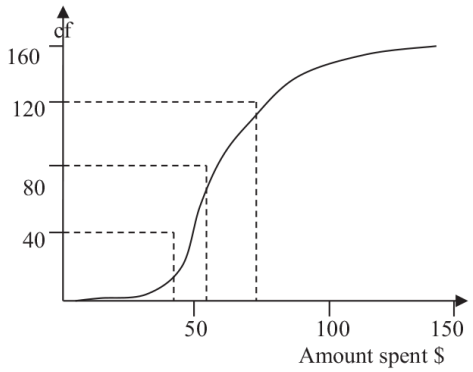
(i) Draw a cumulative frequency graph of this distribution. [4]

(ii) Estimate the median and the interquartile range of the amount spent. [3]

(iii) Estimate the number of shoppers who spent more than \$115. [2]

(iv) Calculate an estimate of the mean amount spent. [2]

Answer:

(i)	cf 16, 56, 104, 130, 160  	**M1** **B1** **M1** **A1** [4]	Attempt at cf table (up to 160) no graph needed accept %cf but give final linear scale minimum 0 to 160 and 0 to 120 Attempt to plot points at (30, 16), (50, 56), (70,104), (90, 130), (140, 160) up to 2 errors can have a polygon All points correct from their scale and joined up, with (0,0) as well
(ii)	median $59 IQR = 82 – 43 = $39	**B1**√ **M1** **A1**√ [3]	accept 57–60 or ft their graph if used lb, midpts instead of ub or assume linear interpolation. Subt a (sensible) LQ from a sensible UQ (generous) Ans ft need a cf graph and UQ 80–84, LQ

Question	Answer	Marks	Guidance
(iii)	160 – 149 = 11 OR 115 is mid pt of last interval so # of shoppers is 30/2 = 15 (can be implied)	**M1** **A1** [2]	41–46 Subtracting from 160 can be implied Correct answer accept 9–16
(iv)	mean = (15×16+ 40×40 +60×48+ 80×26 + 115×30)/160 = 10250/160 = $64.1= $64.1	**M1** **A1** [2]	Using $\Sigma xf/160$ with mid-points

63. 9709_s16_qp_62 Q: 5

The following are the maximum daily wind speeds in kilometres per hour for the first two weeks in April for two towns, Bronlea and Rogate.

Bronlea	21	45	6	33	27	3	32	14	28	24	13	17	25	22
Rogate	7	5	4	15	23	7	11	13	26	18	23	16	10	34

(i) Draw a back-to-back stem-and-leaf diagram to represent this information. [5]

(ii) Write down the median of the maximum wind speeds for Bronlea and find the interquartile range for Rogate. [3]

(iii) Use your diagram to make one comparison between the maximum wind speeds in the two towns. [1]

Answer:

(i)	Bronlea Rogate 6 3 0 4 5 7 7 7 4 3 1 0 1 3 5 6 8 8 7 5 4 2 1 2 3 3 6 3 2 3 4 5 4 Key 3| 1 | 5 represents 13 kph for Bronlea and 15 kph for Rogate	**B1** **B1** **B1** **B1** **B1** [5]	Correct single stem Correct ordered leaves Bronlea Correct ordered leaves Rogate Correct overall shape Single key must have both towns and units consistent with their values
(ii)	median Bronlea = 23 km per hour IQ range Rogate = 23 – 7 = 16	**B1** **M1** **A1** [3]	Units not necessary Subt their LQ <14 from their UQ>14 from Rogate leaf
(iii)	Rogate is less windy than Bronlea	**B1** [1]	Not a comparison of a statistic but interpretation of information

64. 9709_s16_qp_63 Q: 2

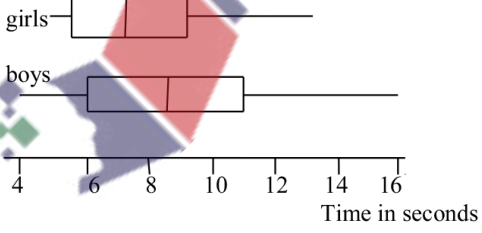
A group of children played a computer game which measured their time in seconds to perform a certain task. A summary of the times taken by girls and boys in the group is shown below.

	Minimum	Lower quartile	Median	Upper quartile	Maximum
Girls	5	5.5	7	9	13
Boys	4	6	8.5	11	16

(i) On graph paper, draw two box-and-whisker plots in a single diagram to illustrate the times taken by girls and boys to perform this task. [3]

(ii) State two comparisons of the times taken by girls and boys. [2]

Answer:

(i)		**B1** **B1** **B1** [3]	Labels 'time' **and** 'seconds', 'boys' and 'girls' on correct plots and scaled line One box and whisker all correct on graph paper – ignore boy or girl label Second box and whisker all correct (on graph paper and ignore boy/girl label) on SAME scaled line.
(ii)	girls smaller range or IQ range than boys /girls less spread out oe girls generally quicker than boys or girls median<boys median (not mean) oe boys almost symmetrical, girls +vely skewed oe	**B1** **B1** [2]	Any 2 comments – MUST be a comparison

65. 9709_s16_qp_63 Q: 4

The monthly rental prices, $\$x$, for 9 apartments in a certain city are listed and are summarised as follows.

$$\Sigma(x-c) = 1845 \qquad \Sigma(x-c)^2 = 477\,450$$

The mean monthly rental price is \$2205.

(i) Find the value of the constant c. [2]

(ii) Find the variance of these values of x. [2]

(iii) Another apartment is added to the list. The mean monthly rental price is now \$2120.50. Find the rental price of this additional apartment. [2]

Answer:

Qu	Answer	Marks	Guidance
(i)	1845/9 (= 205) c = 2205 - 205 = 2000	M1 A1	Accept (1845± anything)/ 9
	OR $\Sigma x = 2205 \times 9$ (= 19845) $\Sigma x - \Sigma c = 1845$ Σc = 19845 -1845 = 18000 c = 2000	M1 A1 [2]	For 2205×9 seen
(ii)	var = $\frac{477450}{9} - 205^2$ = 11025	M1 A1	For $\frac{477450}{9}$ – (their coded mean)^2
	OR var = $\frac{43857450}{9} - 2205^2$ = 11025	M1 A1 [2]	For their $\Sigma x^2/9 - 2205^2$ where Σx^2 is obtained from expanding $\Sigma(x-c)^2$ with $2c\Sigma x$ seen
(iii)	new total = 2120.5×10 = 21205 new price = 21205 – 19845 = 1360	M1 A1 [2]	Attempt at new total

66. 9709_w16_qp_61 Q: 7

The masses, in grams, of components made in factory A and components made in factory B are shown below.

Factory A	0.049	0.050	0.053	0.054	0.057	0.058	0.058
	0.059	0.061	0.061	0.061	0.063	0.065	
Factory B	0.031	0.056	0.049	0.044	0.038	0.048	0.051
	0.064	0.035	0.042	0.047	0.054	0.058	

(i) Draw a back-to-back stem-and-leaf diagram to represent the masses of components made in the two factories. [5]

(ii) Find the median and the interquartile range for the masses of components made in factory B. [3]

(iii) Make two comparisons between the masses of components made in factory A and the masses of those made in factory B. [2]

Answer:

(i)	Factory *A* | Factory *B* 3 | 1 5 8 9 | 4 | 2 4 7 8 9 9 8 8 7 4 3 0 | 5 | 1 4 6 8 5 3 1 1 1 | 6 | 4 Key: 9 | 4 | 2 represents 0.049g for factory *A* and 0.042 g for factory *B*	**M1** **B1** **B1** **B1** **B1**	[5]	Attempt at ordering factory *B* Correct stem Correct leaves factory *A* Correct leaves factory *B* Correct key need factory *A* and factory *B* and units
(ii)	median factory *B* = 0.048 g IQR = UQ – LQ = 0.055 – 0.04 = 0.015	**B1** **M1** **A1**	[3]	using their key i.e. 48, 0.48 etc or correct Subt their LQ from their UQ for factory *B*
(iii)	generally heavier in factory *A* Masses more spread out in factory *B*	**B1** **B1**	[2]	oe must refer to context, e.g. mass

67. 9709_w16_qp_62 Q: 5

The number of people a football stadium can hold is called the 'capacity'. The capacities of 130 football stadiums in the UK, to the nearest thousand, are summarised in the table.

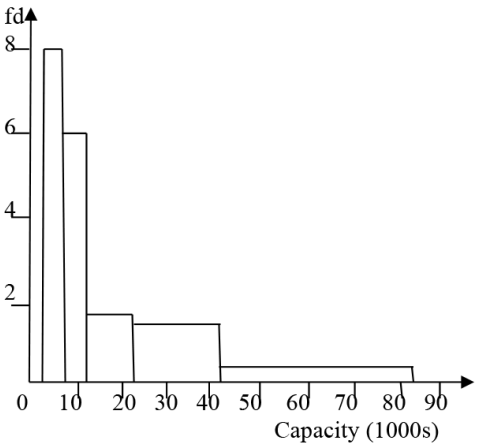
Capacity	3000–7000	8000–12 000	13 000–22 000	23 000–42 000	43 000–82 000
Number of stadiums	40	30	18	34	8

(i) On graph paper, draw a histogram to represent this information. Use a scale of 2 cm for a capacity of 10 000 on the horizontal axis. [5]

(ii) Calculate an estimate of the mean capacity of these 130 stadiums. [2]

(iii) Find which class in the table contains the median and which contains the lower quartile. [2]

Answer:

(i)	cw 5, 5, 10, 20, 40 fd 8, 6, 1.8, 1.7, 0.2  	**M1** **M1** **A1** **B1** **B1**	[5]	cw either 4 or 5 etc fd or scaled freq [f/their cw attempt] fd may be ÷ 1000 Correct heights seen accurately on diagram Correct bar ends, accurately plotted on axis Labels fd and capacity (thousands) Correct horizontal scale required. Vertical scale linear from 0
(ii)	(5×40+10×30+17.5×18+32.5×34+62.5×8)/130 = 2420/130 = 18.6 thousand	**M1** **A1**	[2]	Σ*fx*/130 where *x* is mid point attempt (value within class, not end pt or cw)
(iii)	median group = 8 – 12 thousand LQ group = 3 – 7 thousand	**B1** **B1**	[2]	Thousands not needed

68. 9709_w16_qp_63 Q: 5

The tables summarise the heights, h cm, of 60 girls and 60 boys.

Height of girls (cm)	$140 < h \leqslant 150$	$150 < h \leqslant 160$	$160 < h \leqslant 170$	$170 < h \leqslant 180$	$180 < h \leqslant 190$
Frequency	12	21	17	10	0

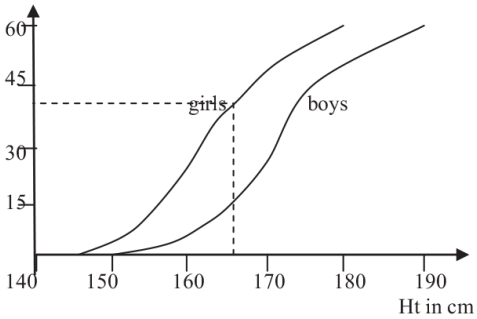
Height of boys (cm)	$140 < h \leqslant 150$	$150 < h \leqslant 160$	$160 < h \leqslant 170$	$170 < h \leqslant 180$	$180 < h \leqslant 190$
Frequency	0	20	23	12	5

(i) On graph paper, using the same set of axes, draw two cumulative frequency graphs to illustrate the data. [4]

(ii) On a school trip the students have to enter a cave which is 165 cm high. Use your graph to estimate the percentage of the girls who will be unable to stand upright. [3]

(iii) The students are asked to compare the heights of the girls and the boys. State one advantage of using a pair of box-and-whisker plots instead of the cumulative frequency graphs to do this. [1]

Answer:

(i)	cf  	**B1** **B1** **B1** **B1**	[4]	Horizontal axis from min of 140 to 190 and vertical axis from 0 to minimum of 60 and two CF graphs on the same set of axes. Labels: CF; height (ht) in cm; girls; boys in correct places CF graph going through (150, 0) , (160, 20), (170, 43), (180, 55) and (190, 60) CF graph going through (140, 0), (150, 12), (160,33), (170,50), (180, 60) [and (190, 60)]
(ii)	42 (± 1) shorter than 165. (18(± 1))/60×100 = 30% (± 1.7%)	**M1** **M1** **A1**	[3]	Line or reading from 165 on their cf graph oe subtracting from 60
(iii)	can see which is taller; see which of boys or girls is more spread out	**B1**	[1]	any sensible comment in context

69. 9709_s15_qp_61 Q: 2

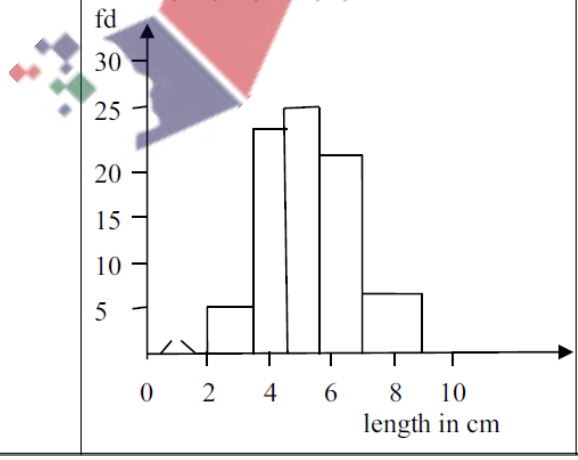
The table summarises the lengths in centimetres of 104 dragonflies.

Length (cm)	2.0 – 3.5	3.5 – 4.5	4.5 – 5.5	5.5 – 7.0	7.0 – 9.0
Frequency	8	25	28	31	12

(i) State which class contains the upper quartile. [1]

(ii) Draw a histogram, on graph paper, to represent the data. [4]

Answer:

(i)	UQ 5.5 – 7.0 cm	**B1 [1]**	
(ii)	fd 5.33, 25, 28, 20.7, 6,  	**M1** **A1** **B1** **B1 [4]**	Attempt at fd or scaled freq [fr/cw] Correct heights seen on graph Correct bar widths no gaps Labels (fd and length/cm) and correct bar ends

70. 9709_s15_qp_61 Q: 5

The table shows the mean and standard deviation of the weights of some turkeys and geese.

	Number of birds	Mean (kg)	Standard deviation (kg)
Turkeys	9	7.1	1.45
Geese	18	5.2	0.96

(i) Find the mean weight of the 27 birds. [2]

(ii) The weights of individual turkeys are denoted by x_t kg and the weights of individual geese by x_g kg. By first finding Σx_t^2 and Σx_g^2, find the standard deviation of the weights of all 27 birds. [5]

Answer:

(i)	new mean $= \dfrac{9\times 7.1+18\times 5.2}{27}$ $= 5.83$	**M1** **A1 [2]**	Mult by 9 and 18 and dividing by 27 correct answer
(ii)	$1.45^2 =$ so $\dfrac{\Sigma x_t^2}{9} = 472.6125$ mm	**M1**	subst in a correct variance formula sq rt or not
		A1	correct Σx_t^2 (rounding to 470)
	$0.96^2 = \dfrac{\sum x_g^2}{18} - 5.2^2$ so $\Sigma x_g^2 = 503.3088$	**A1**	correct Σx_g^2 (rounding to 500)
	New sd^2 $\dfrac{472.6\ldots^2 + 503.3\ldots^2}{27} - 5.83\ldots^2 = 2.117$	**M1**	using $\Sigma x_t^2 + \Sigma x_g^2$, dividing by 27 and subt comb mean^2
	New sd = 1.46	**A1 [5]**	correct answer

71. 9709_s15_qp_62 Q: 2

120 people were asked to read an article in a newspaper. The times taken, to the nearest second, by the people to read the article are summarised in the following table.

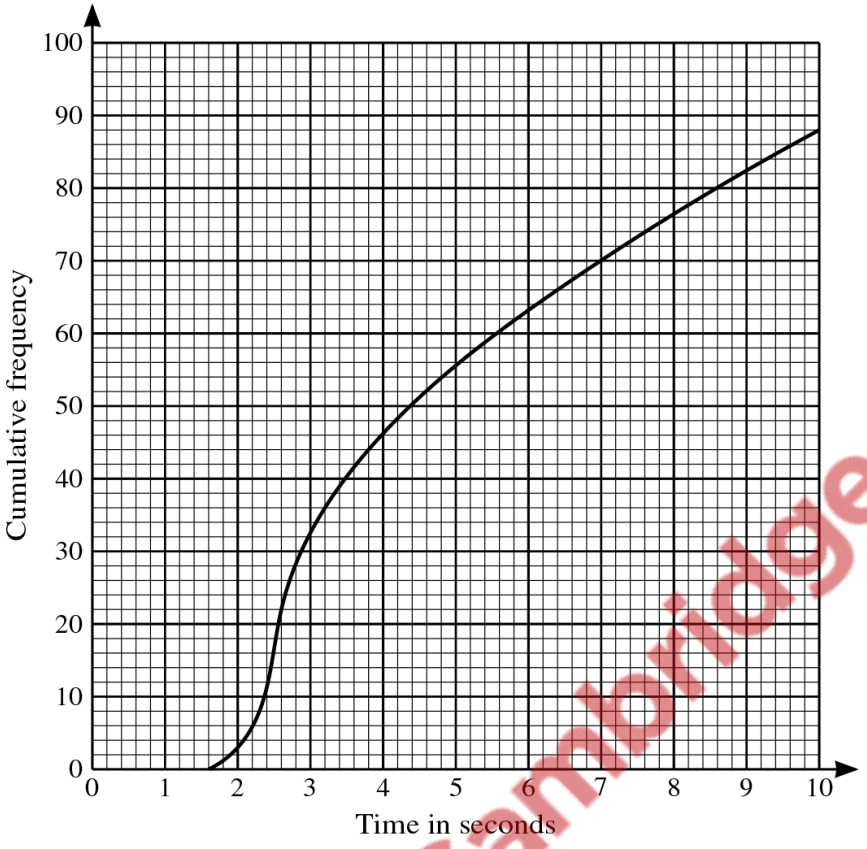
Time (seconds)	1 – 25	26 – 35	36 – 45	46 – 55	56 – 90
Number of people	4	24	38	34	20

Calculate estimates of the mean and standard deviation of the reading times. [5]

Answer:

	mid points 13, 30.5, 40.5, 50.5, 73	M1		Attempt at midpoints at least 3 correct
	Mean = $\frac{4\times13 + 24\times30.5 + 38\times40.5 + 34\times50.5 + 20\times73}{120}$	M1		Using their midpoints i.e. cw, ucb, 1/2 cw and freqs into correct formula must be divided by 120
	$= \frac{5500}{120} = 45.8$	A1		Correct answer from correct working Evaluating
	var = $\frac{4\times13^2 + 24\times30.5^2 + 38\times40.5^2 + 34\times50.5^2 + 20\times73^2}{120} - (45.8\ldots)^2$	M1		$\frac{\Sigma fx^2}{120}$ – their $\bar{x}^2$ must see their 45.8^2 subtracted allow cw etc
	$= \frac{278620}{120} - 45.8\ldots^2$ $= 2321.8333 - 45.8\ldots^2$ sd = 14.9	A1	**5**	Correct answer

72. 9709_s15_qp_62 Q: 3



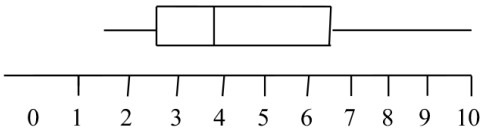
In an open-plan office there are 88 computers. The times taken by these 88 computers to access a particular web page are represented in the cumulative frequency diagram.

(i) On graph paper draw a box-and-whisker plot to summarise this information. [4]

An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

(ii) Show that there are no outliers. [2]

Answer:

(i)		B1 B1 √ B1 B1	**4**	LQ = 2.6 med = 3.8– 3.85, UQ = 6.4– 6.6 Correct quartiles and median on graph ft linear from 2–10 End whiskers correct not through box Label need seconds and linear 2–10 axis or can have 5 values on boxplot no line provided correct
(ii)	1.5 × IQR = 1.5 × 3.8 = 5.7 LQ – 5.7 = –ve, UQ + 5.7 = 12.1 i.e. > 10 So no outliers AG	M1 A1	**2**	Attempt to find 1.5 × IQR and add to UQ or subt from LQ OR compare 1.5 × IQR with gap 3.6 between UQ and max 10 Correct conclusion from correct working need both

73. 9709_s15_qp_63 Q: 6

3

Seventy samples of fertiliser were collected and the nitrogen content was measured for each sample. The cumulative frequency distribution is shown in the table below.

Nitrogen content	⩽ 3.5	⩽ 3.8	⩽ 4.0	⩽ 4.2	⩽ 4.5	⩽ 4.8
Cumulative frequency	0	6	18	41	62	70

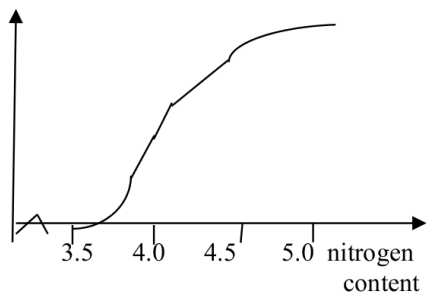
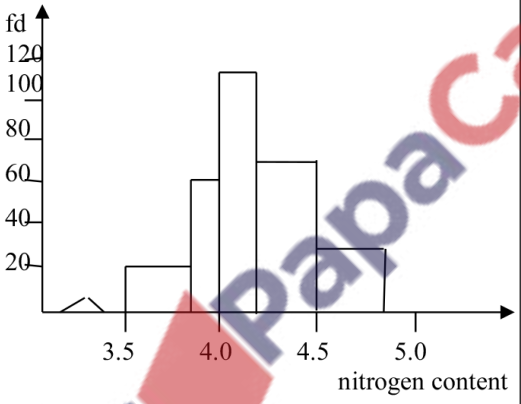
(i) On graph paper draw a cumulative frequency graph to represent the data. [3]

(ii) Estimate the percentage of samples with a nitrogen content greater than 4.4. [2]

(iii) Estimate the median. [1]

(iv) Construct the frequency table for these results and draw a histogram on graph paper. [5]

Answer:

(i)	cf  	B1 M1 A1 **[3]**	Uniform axes cf and nitrogen content labelled, at least 0 to 70 and 3.5 to 4.8 seen 5 points plotted correctly on graph paper 3.5 | 3.8 | 4.0 | 4.2 | 4.5 | 4.8 0 | 6 | 18 | 41 | 62 | 70 All points correct and a reasonable curve (condone 1 missed point) or line segments.
(ii)	70 – their 55 = 15 = 21.4%	M1 A1 **[2]**	Subt a value > 41 from 70 (or n/70, n<29) Correct ans, accept 18.5 – 22
(iii)	median = 4.15	B1 **[1]**	Accept 4.1< median < 4.2, nfww
(iv)	nit cont: 3.5–3.8 | 3.8–4.0 | 4.0–4.2 | 4.2–4.5 | 4.5–4.8 fr: 6 | 12 | 23 | 21 | 8 fd: 20 | 60 | 115 | 70 | 26.7  	M1 M1 A1 B1 B1 **[5]**	Attempt at freqs, at least 3 correct, ignore labelling Attempt at fd as f/cw only at least 3 correct FT (Accept f/cw × k) Correct heights seen on graph (plot at 4.8,27 A0) Graph paper must be used (3 correct relative heights implies M1M1) Correct bar ends seen on graph – graph paper used Correct linear scale and labels.

74. 9709_w15_qp_61 Q: 3

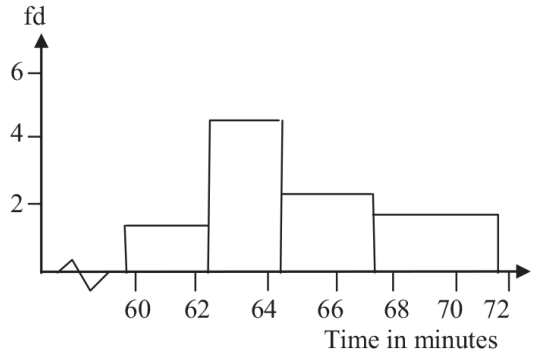
Robert has a part-time job delivering newspapers. On a number of days he noted the time, correct to the nearest minute, that it took him to do his job. Robert used his results to draw up the following table; two of the values in the table are denoted by a and b.

Time (t minutes)	60 – 62	63 – 64	65 – 67	68 – 71
Frequency (number of days)	3	9	6	b
Frequency density	1	a	2	1.5

(i) Find the values of a and b. [3]

(ii) On graph paper, draw a histogram to represent Robert's times. [3]

Answer:

(i)	$a = 9/cw$ $= 9/2 = 4.5$ $1.5 = b/4$ so $b = 6$	**M1** **A1** **A1** [3]	Using fd = f/cw Correct a Correct b
(ii)		**B1**√ **B1** **B1** [3]	Correct heights ft their b Correct widths, ie 3, 2, 3, 4 starting either 60 or 59.5 Labels fd, time or minutes and squiggle and bars from 59.5 to 71.5

75. 9709_w15_qp_62 Q: 1

For n values of the variable x, it is given that $\Sigma(x - 100) = 216$ and $\Sigma x = 2416$. Find the value of n. [3]

Answer:

	$\Sigma x - 100n = 216$ $2416 - 100n = 216$ $n = 22$ OR $\frac{2416}{n} = \frac{216}{n} + 100$ $n = 22$	**B1** **B1** **B1** 3 **B1** **B1** **B1**	$\Sigma x - 100n$ seen Subst 2416 for their Σx Correct answer $2416/n$ seen or $216/n + 100$ oe eg $\Sigma x/n - 100 = 216/n$ correct equation Correct answer

76. 9709_w15_qp_62 Q: 5

The weights, in kilograms, of the 15 rugby players in each of two teams, A and B, are shown below.

Team A	97	98	104	84	100	109	115	99	122	82	116	96	84	107	91
Team B	75	79	94	101	96	77	111	108	83	84	86	115	82	113	95

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

(ii) Find the interquartile range of the weights of the players in team A. [2]

(iii) A new player joins team B as a substitute. The mean weight of the 16 players in team B is now 93.9 kg. Find the weight of the new player. [3]

Answer:

(i)	team *A* |	team *B*	**B1** Correct stem can be upside down, ignore extra values, allow 70, 80 etc with suitable numerical key
		7 | 5 7 9	
	4 4 2 |	8 | 2 3 4 6	**B1** Correct team *A* must be on LHS, alignment ± half a space, no late entries squeezed in, no crossing out if shape is changed
	9 8 7 6 1 |	9 | 4 5 6	
	9 7 4 0 |	10 | 1 8	
	6 5 |	11 | 1 3 5	**B1** Correct team *B* in single diagram can be either LHS or RHS
	2 |	12	
	key 1 | 9 | 4 means 91 kg for team *A* and 94 kg for *B*		**B1** 4 Correct key or keys for their diagram/s, need both teams, at least one kg.
(ii)	LQ = 91 UQ = 109 IQ range = 18		**B1** Both quartiles correct **B1√** 2 Correct IQR ft wrong quartiles, LQ < UQ, not 12 – 4 etc
(iii)	$\Sigma x_{15} = 1399$ $\Sigma x_{16} = 16 \times 93.9 = 1502.4$ New wt = 1502.4 – 1399 = 103 (103.4)		**M1** Attempt at Σx_{15} for either team **M1** Mult 93.9 by 16 attempt **A1** 3 Correct answer

77. 9709_w15_qp_63 Q: 1

The time taken, t hours, to deliver letters on a particular route each day is measured on 250 working days. The mean time taken is 2.8 hours. Given that $\Sigma(t - 2.5)^2 = 96.1$, find the standard deviation of the times taken. [3]

Answer:

	coded mean = 0.3 oe	**B1**	$\Sigma(t - 2.5) = 75$ B0 until ÷ 250
	$\text{sd} = \sqrt{\frac{96.1}{250} - (0.3)^2}$	**M1**	Subst in variance formula both terms coded
	$= 0.543$	**A1** 3	Correct answer
	Alt: $\Sigma(t-2.5)^2$ expanded	**Or**	
	$\Sigma t^2 = 2033.6$	**B1**	
	$\text{sd} = \sqrt{\frac{2033.6}{250} - 2.8^2}$	**M1**	Substituting their Σt^2 from expanded 3-term expression, 250 and 2.8 in variance formula
	$= 0.543$	**A1** 3	

78. 9709_w15_qp_63 Q: 6

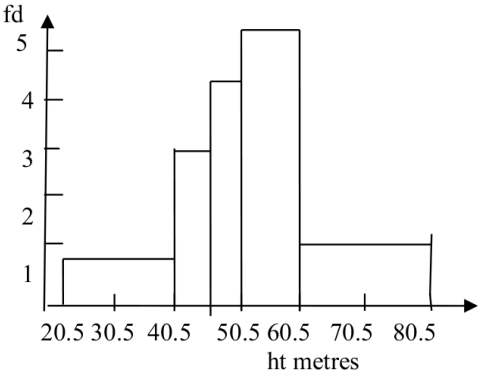
The heights to the nearest metre of 134 office buildings in a certain city are summarised in the table below.

Height (m)	21 – 40	41 – 45	46 – 50	51 – 60	61 – 80
Frequency	18	15	21	52	28

(i) Draw a histogram on graph paper to illustrate the data. [4]

(ii) Calculate estimates of the mean and standard deviation of these heights. [5]

Answer:

(i)	fd 0.9, 3, 4.2, 5.2, 1.4 fd  	**M1** **A1** **B1** **B1**	4	Attempt at scaled freq [f/(attempt at cw)] Correct heights seen on diagram Scale no less than 1cm to 1 unit Correct bar widths visually no gaps Labels (ht/metres and fd or freq per 20 m etc.) and end points at 20.5 etc. condone 2 end point errors, scale no less than 1cm to 5m for 20,30… unless clearly accurate, linear scale between 20.5 and 80
(ii)	(30.5 × 18 + 43 × 15 + 48 × 21 + 55.5 × 52 + 70.5 × 28)/134 $= \frac{7062}{134} = 52.701$ Var = $(30.5^2 \times 18 + 43^2 \times 15 + 48^2 \times 21 + 55.5^2 \times 52 + 70.5^2 \times 28)/134 - 52.701^2$ $= 392203.5/134 - 52.701^2 = 149.496$ sd = 12.2	**M1** **M1** **A1** **M1** **A1**	5	Attempt at unsimplified, mid points (at least 4 within 0.5) Attempt at Σfx their mid points ÷ 134 Correct mean rounding to 53 Attempts at Σfx^2 their mid points ÷ their Σf – mean^2 Correct answer, nfww