Qu	Scheme	Marks	AO
2. (a)	From [5,20) fd = 3 or 1 large square = 2.5 passengers o.e.	M1	2.2a
	Correct bar above [0, 5)	A1	1.1b
	Correct bar above [20, 40)	A1	1.1b
		(3)	
(b)	For [40, 65) **130** passengers or for [65, 80) **60** passengers	M1	2.1
	For attempt to find total number of passengers = **331**	A1ft	1.1b
	[Median =] $40+\frac{\frac{1}{2}("331")-140}{"130"}\times 25$ or $65-\frac{270-\frac{1}{2}("331")}{"130"}\times 25$ (o.e.)	M1	1.1b
	$= 44.9038\ldots$ = awrt **44.9**	A1	1.1b
		(4)	
(c)	Upper outlier limit = $58.9 + 1.5\times(58.9 - 27.3) = 106\ (.3) > 90$	M1	2.4
	So oldest passenger is not an outlier	A1	2.2a
		(2)	
		(9 marks)	

	Notes
(a)	M1 for attempt at fd or a suitable method to deduce the scale for the histogram May be implied by one correct bar. 1st A1 for first bar [0, 5) with fd = 1 or 2 large squares high 2nd A1 for third bar with fd = 4.5 or 9 large squares high
(b)	1st M1 for an attempt using their fd to find the missing frequencies. May be in table 1st A1ft for a clear attempt to find the total number of passengers (ft their 130 and 60) 2nd M1 for any expression/equation leading to correct Q_2 Must be using 40-65 class 2nd A1 for awrt 44.9 (allow $(n + 1)$ leading to 45)
(c)	M1 for finding the upper outlier limit (expression or awrt 106) and stating or implying > 90 A1 dep on M1 seen for deducing NOT an outlier

Qu	Scheme	Marks	AO
4 (a)	$\bar{x} = 10.2\,(2222\ldots)$ awrt **10.2**	B1 (1)	1.1b
(b)	$\sigma_x = 3.17\,(20227\ldots)$ awrt **3.17**	B1ft	1.1b
	Sight of "knots" or "kn" (condone knots/s etc)	B1 (2)	1.2
(c)	October ….. since	B1	2.2b
	it is windier in the autumn or month of the hurricane or latest month in the year	B1 (2)	2.4
(d)(i)	They represent outliers	B1	1.2
(ii)	*Y* has low median so expect lowish mean (but outlier so > 7) and *Y* has big range/IQR or spread so expect larger st.dev	M1	2.4
	Suggests *B*	A1 (3)	2.2b
		(8 marks)	

Notes

NB $\bar{x} = \frac{184}{18}$ and $\sigma_x = \sqrt{\frac{2062}{18} - \bar{x}^2}$

(a) B1 for $\bar{x} = 10.2$ (allow exact fraction)

(b) 1st B1ft allow 3.2 from a correct expr' accept $s = 3.26(3984\ldots)$ [ft use of n/a]
Treating n/a as 0 May see $n = 31$ or $\bar{x} = 5.9354...$ which is B0 in (a) but here in (b) it gives $\sigma_x = 5.59(34\ldots)$ or $s = 5.6858\ldots$(awrt 5.69) and scores 1st B1
2nd B1 accept kn accept in (a) or (b) (allow nautical miles/hour)

(c) 1st B1 choosing October but accept September.
2nd B1 for stating that (Camborne) is windier in autumn/winter months "because it is winter/autumn/windier/colder in "month" " Sep $\leqslant$ "month" $\leqslant$ Mar
scores B1B1 for "month" = Sep or Oct and B0B1 for other months in range

(d)(i) B1 for outlier or the idea of an extreme value allow "anomaly"

(ii) M1 for a comment relating to location that mentions both median and mean and a comment relating to spread that mentions both range/IQR and standard deviation and leads to choosing *B*, *C* or *D*

Question	Scheme	Marks	AOs
4 (a)	Tr(ace) (data needs to be converted to numbers before the calculation can be carried out)	B1	2.4
		(1)	
(b)	$[1+]\frac{138-131}{24}\times 4$	M1	2.1
	$= 2.1666\ldots$ awrt **2.17**	A1	1.1b
		(2)	
(c)	$\sigma = \sqrt{\frac{7704.1875}{184} - \left(\frac{539.75}{184}\right)^2} = 5.7676...$ $\sigma =$ awrt **5.77**	M1 A1	1.1b 1.1b
		(2)	
(d)(i)	Using class midpoints to estimate the mean assumes that the values are uniformly distributed **within the class(es)**.	B1	2.4
(ii)& (iii)	This is not the case here as the majority of the data (in the first class) are 0.	B1	2.3
	The actual mean is likely to be smaller than the estimate (since the first group has more values at 0 and close to 0)	dB1	2.2b
		(3)	
			(8 marks)

Notes	
(a)	**B1:** Identifying tr(ace) data Ignore comments about n/a, missing data, anomalies, etc.
(b)	**M1:** Correct fraction $\frac{7}{24}\times 4$ allow working down $[5]-\frac{155-138}{24}\times 4$ allow a correct equation leading to a correct fraction e.g. $\frac{x-1}{5-1}=\frac{138-131}{155-131}$ for M1 Use of $(n + 1)$ with 138.75 allow $\frac{7.75}{24}\times 4$ **A1:** awrt 2.17 (condone $\frac{13}{6}$) awrt 2.29 from $(n + 1)$ (condone $\frac{55}{24}$)
(c)	**M1:** Correct expression for standard deviation (allow mean = awrt 2.93) **A1:** awrt 5.77 correct answer only scores M1A1 (allow $s = 5.78$) **SC:** 5.76 with no working scores M1A0
(d)(i)	**B1:** Explaining that data assumed to be spread evenly across each class (o.e.) e.g. The midpoint of each class is the mean of each class or all the values in the class are located at the midpoint condone normally distributed within each class
Mark together (ii)&(iii)	**B1:** Demonstrating an understanding of the LDS that the majority of data values (in the first class) are at 0 or close to 0 (trace). **dB1:** (dependent upon 2nd B1) Correct inference based on knowledge of the LDS **SC:** If B1 is scored in (i) for 'The data are spread evenly across each class,' then in (ii) 'The data are not evenly distributed in the classes' scores B1 but in (iii) 'the actual mean is smaller' with no further justification scores B0

Question	Scheme	Marks	AOs
1	1 square is $\frac{78}{12\times3+3\times4+2\times2}=\left[\frac{78}{52}=1.5\right]$ **and** $(8\times1+1\times8)\times"1.5"$	M1	3.1a
	24 students took less than 11 minutes	A1	1.1b
	Percentage of students = $\frac{"24"}{78+"24"+1\times8\times"1.5"+3\times4\times"1.5"}\times100$	M1	3.1b
	= 18.18… awrt 18%	A1	1.1b
		(4)	
			Total 4

Notes		
1	**M1:**	For clear use of frequency density to establish the fd scale and then use the area to find frequency of <11 minutes. Allow maximum of 3 errors in either the heights or widths in total if working shown. They may calculate the area using other size squares. Allow for realising they need to find the total number of squares (88) maximum of 4 errors in either the heights or widths and number < 11 minutes(16) - must have a maximum of 1 error in either the heights or widths (and not use the 78 as part of calulation)
	A1:	For correct values seen. Allow for 88 and 16
	M1:	For realising the need to find the total and calculating a percentage. (with "their 24" as the numerator). Allow $(8\times1+2\times8)\times"1.5"$ instead of $"24"+1\times8\times"1.5"$ If working shown can allow maximum of 2 errors in either the heights or widths in the calculation of the total. Allow "their 24" / 132 oe
	A1:	awrt 18

Question	Scheme	Marks	AOs
4(a)	It is not possible to have a sampling frame	B1	2.3
		(1)	
(b)	Quota sampling **and** (catch 85 common carp, 45 mirror carp and 30 leather carp) **or** (ignore any fish caught of a type where the quota is full)	M1	1.1a
	Quota sampling **and** catch 85 common carp, 45 mirror carp and 30 leather carp **and** ignore any fish caught of a type where the quota is full	A1	1.1b
		(2)	
(c)	$\sigma = \sqrt{\frac{3053}{160} - \left(\frac{692}{160}\right)^2}$	M1	1.1b
	$= 0.6129\ldots$ awrt 0.613	A1	1.1b
		(2)	
(d)(i)	This would have no effect as the piece of data would remain in the same class	B1	2.2a
(ii)	This would increase the standard deviation as change in mean is small and $6.4 - 4.6 \approx 3\sigma$ therefore estimate of standard deviation will increase	B1	2.2a
		(2)	
			(7 marks)

Notes		
(a)	**B1:**	For the idea there cannot be a sampling frame/list
(b)	**M1:**	Quota sampling **and** either for the correct numbers of each type **or** for the idea that if quota full ignore the fish.
	A1:	Quota sampling **and** both the correct numbers of each type **and** for the idea that if quota full ignore the fish or sample until all quotas are full
(c)	**M1:**	A correct expression for σ
	A1:	Awrt 0.613 allow s = awrt 0.615
(d)	**B1:**	Correct deduction with suitable explanation Allow range for class. Do not allow there is no differences
	B1:	Correct deduction with suitable explanation. so would increase the standard deviation and a suitable reason. Allow the value is bigger than any others in the table **oe**

Paper 2: Statistics and Mechanics Mark Scheme

Question	Scheme	Marks	AOs
1(a)	Systematic (sample) cao	B1	1.2
(b)	In LDS some days have gaps because the data was not recorded	B1	2.4
(c)	$\left[\bar{t} = \frac{374}{20} = 18.7\right]$ $\sigma_t = \sqrt{\frac{7600}{20} - \bar{t}^2}$ $[= \sqrt{30.31}]$	M1	1.1a
	= 5.5054… awrt **<u>5.51</u>** (Accept use of $s_t = \sqrt{\frac{7600 - 20\bar{t}^2}{19}}$ = 5.6484…)	A1	1.1b
			(4 marks)

Notes:

(b)
B1: A correct explanation

(c)
M1: For a correct expression for $\bar{t}$ and σ_t or s_t
ft an incorrect evaluation of $\bar{t}$
A1: For σ_t = awrt 5.51 or s_t = awrt 5.65

Question	Scheme	Marks	AOs
2	$17 + 45 + \frac{1}{3} \times 9$ $[= 65]$	M1	2.2a
	(7 – 8) **<u>14</u>** or (16 – 20) **<u>5</u>** [Values may be seen in the table]	M1 A1	3.1a 1.1b
	Percentage of motorists is $\frac{"65"}{6+"14"+17+45+9+"5"} \times 100$	M1	3.1b
	= **<u>67.7%</u>**	A1	1.1b
			(5 marks)

Notes:

M1: For a fully correct expression for the number of motorists in the interval
M1: For clear use of frequency density in (4- 6) or (13- 15) cases to establish the fd scale. Then use of area to find frequency in one of the missing cases
A1: For both correct values seen
M1: For realising that total is required and attempting a correct expression for %
A1: For awrt 67.7%

Paper 2: Statistics and Mechanics Mark Scheme

Question	Scheme	Marks	AOs
1(a)	Systematic (sample) cao	B1	1.2
(b)	In LDS some days have gaps because the data was not recorded	B1	2.4
(c)	$\left[\bar{t} = \frac{374}{20} = 18.7\right]$ $\sigma_t = \sqrt{\frac{7600}{20} - \bar{t}^2}$ $[= \sqrt{30.31}]$	M1	1.1a
	= 5.5054… awrt **<u>5.51</u>** (Accept use of $s_t = \sqrt{\frac{7600 - 20\bar{t}^2}{19}}$ = 5.6484…)	A1	1.1b
			(4 marks)

Notes:

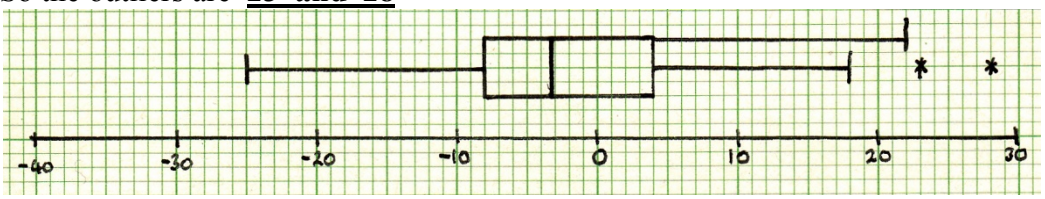
(b)
B1: A correct explanation

(c)
M1: For a correct expression for $\bar{t}$ and σ_t or s_t
ft an incorrect evaluation of $\bar{t}$
A1: For σ_t = awrt 5.51 or s_t = awrt 5.65

Question	Scheme	Marks	AOs
2	$17 + 45 + \frac{1}{3} \times 9$ $[= 65]$	M1	2.2a
	(7 – 8) **<u>14</u>** or (16 – 20) **<u>5</u>** [Values may be seen in the table]	M1 A1	3.1a 1.1b
	Percentage of motorists is $\frac{"65"}{6+"14"+17+45+9+"5"} \times 100$	M1	3.1b
	= **<u>67.7%</u>**	A1	1.1b
			(5 marks)

Notes:

M1: For a fully correct expression for the number of motorists in the interval
M1: For clear use of frequency density in (4- 6) or (13- 15) cases to establish the fd scale. Then use of area to find frequency in one of the missing cases
A1: For both correct values seen
M1: For realising that total is required and attempting a correct expression for %
A1: For awrt 67.7%

Question Number	Scheme	Marks
1. (a)	[Range =] **63**	B1 (1)
(b)	[IQR =] **18**	B1 (1)
(c)	[Q_2 =] $(-8)+\frac{20}{33}\times 8$ or $(0)-\frac{13}{33}\times 8$ [NB (n +1) will have 20.5 or 12.5] = –3.1515... awrt **– 3.15**	M1 A1 (2)
(d)	[Q_3 =] mid-point of [0, 8] group so therefore = 4	B1cso (1)
(e)(i)	IQR = 4 – – 8 = 12 so upper limit is 4 + 1.5×12 = **22** lower limit is – 8 – 1.5×12 = **– 26** So the outliers are **23 and 28**	M1 A1 A1
(ii)		M1 A1 A1 (6)
(f)	Interquartile range is smaller (12 compared to 18) or range is smaller (53 v 63) Median is closer to zero (– 3.15 is closer than 5) So they have improved	B1 B1 dB1 (3) **[Total 14]**

	Notes
(c)	M1 for a correct fraction and ×8 (ignore end point) A1 for awrt – 3.15 (allow use of n + 1 leading to – 3.03) Accept $-\frac{104}{33}$ if box plot is OK or 3sf value is quoted in (f)
(d)	B1cso for a clear argument with no incorrect working seen. Allow 4.14… from 7.25 for (n + 1) case
(e)(i)	M1 for at least one correct calculation e.g. 4 + 1.5(8 – –4) (implied by one correct limit) 1st A1 for one correct limit 2nd A1 for both correct limits and the two correct outliers identified
(ii)	M1 for a box with 2 whiskers (one at each end) 1st A1 for – 8 and 4 and ft Q_2 between them and lower whisker ending at – 25 **no** outliers 2nd A1 for upper whisker ending at 18 **or** 22 and 2 outliers marked at 23 and 28
(e)(ii)SC	Two incorrect outliers in (e)(i), ft both A1s in (ii) using their outliers provided in [– 25, 28]
(f)	1st B1 for a statement about range or IQR saying that 2nd estimates are better Allow range or IQR" has decreased" or "is smaller" o.e. 2nd B1 for a statement about medians saying that 2nd one is closer to zero Don't allow "decreased" or "smaller" unless clearly using |median| or say e.g. 3.15 < 5 3rd dB1 dep on at least one other B1 for concluding that they have improved based on change in median or range/IQR. Must clearly state "improved" not just "yes"

Question	Scheme	Marks
1. **(a)**	$[Q_2 =] \quad (59.5) + \frac{10}{22} \times 5$	M1
	$= 61.7727\ldots$ awrt **<u>61.8</u>**	A1
		(2)
(b)	$[\bar{x} =] \frac{\sum fx}{50} = \frac{3085}{50}$	M1
	$=$ **<u>61.7</u>**	A1cao
		(2)
(c)	$[\sigma_x =] \sqrt{\frac{192\,102.5}{50} - \bar{x}^2} = \sqrt{35.16}$	M1
	$= 5.929586\ldots$ awrt **<u>5.93</u>**	A1
		(2)
(d)	[Interpolation from above]	
	$\frac{4.5}{10} \times 13$ (= 5.85)	M1
	So probability is $\frac{5.85}{50} =$ **<u>0.117</u>**	A1cao
		(2)
		(8 marks)

	Notes
(a)	M1 for a correct expression (oe) without endpoint. Allow "n+1" so e.g. $(59.5) + \frac{10.5}{22} \times 5$ Allow working down e.g. $(64.5) - \frac{12}{22} \times 5$ Allow $\frac{m - 59.5}{64.5 - 59.5} = \frac{25 - 15}{37 - 15}$ oe for M1 A1 for awrt 61.8 or, if (n + 1) is used, allow awrt 61.9
(b)	M1 for a correct expression for the mean $\frac{49.5 \times 5 + 57 \times 10 + 62 \times 22 + 69.5 \times 13}{50}$ <u>or</u> an attempt at $\frac{\sum fx}{50}$ with at least 3 correct products <u>or</u> $\frac{3000 \leq \sum fx \leq 3200}{50}$ A1 for 61.7 from correct working
(c)	M1 for a correct expression including square root. Ft their 61.7. Allow use of s A1 for awrt 5.93 Allow s = 5.989787… = awrt 5.99
(d)	M1 for $\frac{4.5}{10} \times 13$ (use of interpolation to find the number of carrots weighing more than 70g) (may be implied by sight of 5.85 may also be implied by 50 – 44.15 (Allow 50 – 44 (=6) or 50 – 45 (=5) coming from 44.15 or 44.2 seen in working to score M1) A1 for an answer of 0.117 Note: Use of normal distribution scores M0A0.

Question Number	Scheme	Marks
2.	$1.5 \times 12 = 18$ 20 people represented by 18 (cm^2) or 1 person is represented by 0.9 (cm^2) $x = \dfrac{20 \times 94.5}{18}$ oe = 105 (people)	M1 M1 A1cao (3) **Total 3**
	Notes	
	1st M1 for an attempt to relate area to frequency (e.g. $\frac{20}{18}$ or $\frac{18}{20}$ seen) 2nd M1 for a correct expression/equation for total frequency e.g. $\frac{18}{20} = \frac{94.5}{x}$ A1 for 105cao	

Question	Scheme	Marks
8.(a)	(Time is) <u>continuous</u>	B1 (1)
(b)	40 people = 8 large squares/200 small squares 200 people = 40 large squares/1000 small squares 40/(21 – 11) or correct scale on f.d. axis	B1
	$\frac{x}{40}=\frac{180}{200}$ or $\frac{x}{40}=\frac{7.2}{8}$ or $(21-18)\times 4+(25-21)\times 6$	M1
	36 people (spent between 18 and 25 minutes shopping in the supermarket)	A1 (3)
(c)	$\text{Median}=26+\frac{[30]}{36}\times 5=$ awrt **<u>30.2</u>**	M1A1 (2)
(d)	$\sum fx = 16\times 40+23.5\times 30+28.5\times 36+33.5\times 40+38.5\times 14+46\times 20+61\times 20$ $=6390$ **	M1 A1cso (2)
(e)	i $\bar{x}=\frac{6390}{200}=31.95$	B1
	ii $\sigma=\sqrt{\frac{238430}{200}-31.95^2}=\sqrt{171.3475}=13.09\,(\text{or } s=13.122)$ awrt **<u>13.1</u>**	M1A1 (3)
(f)	0.409... awrt **<u>0.4</u>**	B1 (1)
(g)	Method 1: (positive) skew or median ≠ mean oe; not a good decision Method 2 (see note): (almost) symmetric oe; a good decision	B1 dB1 (2)
		Total 14

	Notes
(a)	Allow not discrete. Condone misspellings if intention of 'continuous' is clear.
(b)	B1 for establishing a ratio (usually 5 or 1/5) between people and area <u>or</u> calculating f.d. (may be implied by M1) M1 for a correct ratio <u>or</u> expression using areas for the people from 18 to 25 A1 36 cao (Answer of 36 scores 3 out of 3).
(c)	M1 for an attempt at the medians (should have 26, 36 and 5). If working down $31-\frac{[6]}{36}\times 5$ A1 awrt 30.2 (can come from using (n+1))
(d)	M1 for a correct expression for $\sum fx$ condone one incorrect product A1cso for 6390 and all correct
(e)(i)	B1 31.95 or equivalent fraction
(ii)	M1 for correct expression for standard deviation including root A1 awrt 13.1 (answer of awrt 13.1 scores 2 out of 2) [NB (s = 13.122)]
(g)	1st B1 for comment on skew (may be seen in part (f)). Method 1: skew or median ≠ mean Only allow method 2 if |their(f)| < 0.45. Method 2: ~symmetric (any mention of correlation is B0) 2nd dB1 for a correct compatible comment about the manager's decision

Question	Scheme	Marks
2. (a)	Width = $\frac{5}{3}\times 1.5$ = **2.5 (cm)**	B1
	Area = $6\times 1.5 = 9$ cm^2 has frequency = 12 so 1.5 cm^2 = 2 people (o.e.)	M1
	Frequency of 10 corresponds to area of 7.5 so height = **3 (cm)**	A1 (3)
(b)	$Q_2 = [2.5+]\frac{(25/25.5-16)}{12}\times 3 = 4.75$ (or 4.875 if use n + 1) awrt **4.75**	M1 A1 (2)
(c)(i)	$[\bar{x} =]\ \frac{394}{50} = 7.88$ **(*)**	B1cso
(ii)	$[\sigma_x =]\ \sqrt{\frac{6500}{50}-\bar{x}^2}\ = \sqrt{67.9056}$	M1A1
	= **awrt 8.24** (Accept s = awrt 8.32)	A1 (4)
(d)	$\bar{x} > Q_2$	B1ft
	So positive (skew)	dB1 (2)
(e) (i)	There is no effect on the mean	B1
(ii)	The median will increase	B1
(iii)	The standard deviation will decrease	B1 (3)
		[14]

Notes

(a)
M1 for forming a relationship between area and no. of people or "their width" × "their height"= 7.5
or for $\frac{3h}{10}=\frac{9}{12}$ oe
A1 for height of 3 (cm)
NOTE: the common incorrect answer width = 3 and height = 2.5 scores B0M1A0

(b)
M1 for a correct fraction $\left[\frac{9}{12}\right.$ or $\left.\frac{9.5}{12}\right]\times 3$. Ignore end point but must be +.
May be seen in an equivalent expression e.g. $\frac{(x-2.5)}{5.5-2.5}=\frac{25-16}{28-16}$
Allow use of (n + 1) giving 4.875
NB May work down so look out for $[5.5]-\frac{28-25}{12}\times 3$, etc.

(c)(i)
B1 for $\frac{394}{50}$ or for fully correct expression seen $\frac{16\times 1.25+12\times 4+10\times 8+8\times 15.5+4\times 30.5}{50}$

(ii)
M1 for a correct expression must have 6500, 50 and 7.88. (square root not necessary for M1)
1st A1 for a correct expression which must have square root
2nd A1 for awrt 8.24 (use of s = awrt 8.32). Condone incorrect labelling if awrt 8.24 is found.

(d)
1st B1ft for a correct comparison of $\bar{x}$ =7.88 and their Q_2 (this may be seen embedded in another formula i.e. 3(mean-median)/s.d.)
$Q_3-Q_2 > Q_2-Q_1$ is B0 unless Q_1 and Q_3 have been found. ($Q_1 = 1.95/1.99$, $Q_3 = 10.25/10.81$)
2nd dB1 Dependent on the 1st B1 and for concluding "positive" skew.
Note: if their Q_2 >7.88, then B0. Positive correlation is B0.

Question Number	Scheme	Marks
1. (a)	$[61\times 15=]$ **915**	**B1** **(1)**
(b)	$[\text{Var}_A]=\frac{59610}{10}-77^2$ $=$ **32** $\quad$ $[\text{Var}_B]=\frac{58035}{15}-61^2$ $=$ **148**	**M1** **A1 A1** **(3)**
(c)	Class B since its variance is larger	**B1ft** **(1)**
(d)(i)	$\text{Mean}_{AB}=\frac{770+"915"}{25}=67.4$ or $\frac{10}{25}\times 77+\frac{15}{25}\times 61=$ **67.4**	**M1 A1**
(ii)	$\text{Var}_{AB}=\frac{59\,610+58\,035}{25}-"67.4"^2=163.04$ **awrt 163**	**M1 A1** **(4)**
(e)(i)	No effect on the variance of class A since addition does not change variance (Var $(X + b)$ = Var (X))	**B1**
(ii)	The mean will increase since the total score has increased or mean of A increased but mean of B stayed the same	**B1**
(iii)	The variance of the entire group will increase since the mean of class A is now further away from the mean of class B	**B1** **(3)** **[12 marks]**

	Notes
(b)	M1 for a correct method for variance for either class. Accept s^2 and allow inside $\sqrt{...}$ 1st A1 for 1 correct answer. NB $s_A^2=35.\dot{5}$ or awrt 35.6 and $s_B^2=158.57...$ or awrt 159 2nd A1 for both correct. [ISW standard deviations following correct variances.]
(c)	B1ft for Class B **and** it has a larger variance/standard deviation (do not allow spread) If $\text{Var}_A > \text{Var}_B$ then allow choice of A since variance is larger. Ft their values if > 0
(d)(i)	M1 for a correct calculation for the mean (or weighted mean), ft their 915 from (a) A1 for 67.4 o.e.
(ii)	M1 for use of correct formula (no $\sqrt{...}$) with total $\sum x^2 = 117\,645$ and their mean.NB $\frac{S_{xx}}{25}=\frac{4076}{25}$ A1 for awrt 163 [Don't ISW standard deviation]
(e)(i)	B1 for no effect/does not change **and** correct supporting reason that mentions addition or subtraction doesn't affect or only affected by multiplication/division. Comment that $(x-\bar{x})$ doesn't change is fine. Just "coding" is not sufficient.
(ii)	B1 for stating the mean will increase **and** correct supporting reason that states or implies that total (of scores) has increased. Allow new mean = $\frac{1715}{25}=68.6$
(iii)	B1 for increase **and** correct supporting reason that mentions A marks and B marks and that they are more spread out. Just saying: "marks are more varied" or "only added 3 to one class" is **not** sufficient
NB	Calc for (iii) gives new $\Sigma x^2 = 64\,320$ and $\text{Var}_{AB} = 188.24$ but no mark

January 2017 WST01 Mark Scheme

Question Number	Scheme	Marks
1. (a)	25 small sq' = 5 tomatoes or 1 large square = 5 tomatoes or fd=5 for 2~3 or $\frac{5}{25}\times 20$ or 5×0.8 or 2×2 $= \underline{\mathbf{4}}$	**M1** **A1** **(2)**
(b)	$100-(5+'4')$ or $16+32+25+10+8$, so probability = $\frac{91}{100}$ (condone 91%)	**M1, A1** **(2)**
(c)	$\frac{(7-6.25)\times 16+25+10+8}{100}$ or $1-\frac{(a)+5+16+(6.25-5)\times 16}{100}$ =, $\frac{55}{100}$	**M1, A1** **(2)**
(d)	Since '0.55' > 0.5 (or equivalent reason) and state median > 6.25	**B1** **(1)**
(e)	Median > mean, so negative skew	**B1** **(1)**
(f)	Freq. for $(5.5 < \text{weight} < 7) = (7-5.5)\times '16'$ or $\frac{3}{4}\times '32'$, probability $= \frac{24}{100}$ P (both weigh between 5.5 and 7) = $\frac{24}{100}\times\frac{23}{99}=\frac{46}{825}$ (o.e.) or **awrt 0.056**	**M1, A1** **M1 A1** **(4)** **[12 marks]**

Notes

A correct answer with no working scores M1A1 in parts (a)~(c)

(a) M1 for a correct: statement linking area with frequency or calculation or at least 2 values on the fd scale on axis or at least 2 frequencies on/in histogram bars.
A1 for an answer of 4 (if not in script, can be awarded if 4 seen correctly on histogram). If answers on both diagram and script contradict, the script has preference.

(b) M1 for 100 – (5 + '(a)') ft 0 < 'their (a)' < 10 **or** for a correct method for finding the sum of the areas of all the bars above 3 (condone one slip if 5 terms seen)

(c) M1 fully correct expression (possibly ft their (a)) and need division by 100 (o.e.)

A1 for $\frac{11}{20}$ or 0.55 (o.e.) [Allow 55% or ratio 55:100]

(d) B1 for $Q_2 > 6.25$ with reason based on (c) where 0.5 < 'their (c)' < 1 [comparison of "55" & 50]

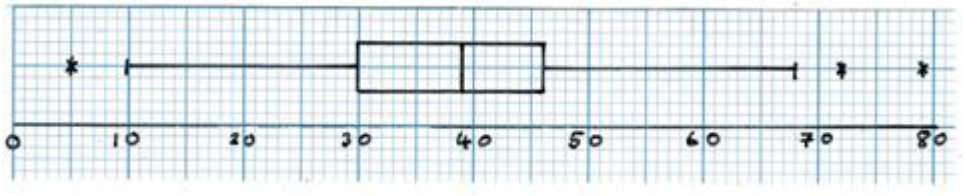
(e) B1 for stating "median > mean" and "negative skew" (independent of (d))

(f) 1st M1 for method to find the frequency between 5.5 and 7 (Implied by the 24 used)
e.g. $(4+5+16+16\times 2)-(4+5+16+16\times 0.5)=57-33$ based on $(\leqslant 7)-(\leqslant 5.5)$

1st A1 for $\frac{24}{100}$ (o.e.)

2nd M1 for $\frac{'24'}{100}\times\frac{'24'-1}{99}$ ft their 24 but must have numerator < denominator of 100×99

2nd A1 for $\frac{46}{825}$ (o.e.) or awrt 0.056 NB $\frac{24}{100}\times\frac{24}{100}$ scores M1A1M0A0 [0.0576 alone 0/4]

Question Number	Scheme	Marks
Question Number	Scheme	Marks
2.(a)	[pass for] **30** (labelled or 1st answer)	B1 (1)
(b)	[merit for] **46** (labelled or 2nd answer)	B1 (1)
(c)	$[1.5(Q_3-Q_1)=1.5\times16=24]$ so $c=$ **70** and $d=$ **6**	B1, B1 (2)
(d)	**68, 72, 79**	B2/1/0 (2)
(e)	$5 < d$ therefore 5 is an outlier  	M1 A1 A1 (3)
(f)	$\frac{1}{2}\times\left(\frac{1}{4}\right)^2\times3$ $=\underline{\frac{3}{32}}$	M1M1 A1 (3) **[12 marks]**

	Notes
	In (a), (b), (c) and (d) condone correct numbers with % e.g. 30% for (a)
(c)	1st B1 for $c = 70$ 2nd B1 for $d = 6$ (Allow B1B1 for unlabelled 70 followed by 6)
SC	Award B1B0 for $c = 6$ and $d = 70$ or 6 and 70 in the wrong order
(d)	B2 for all 3 correct values (and no extra value) (B1 for two correct)
(e)	**Fully correct box plot scores M1A1A1**
	M1 for identifying or stating(e.g. on the grid) that 5 is the only outlier or lower whisker ending at 6 or 10 (May be implied by a correct diagram)
	1st A1 for only one outlier correctly marked at 5 (whisker(s) must stop above 5)
	2nd A1 for a single lower whisker stopping at 10 (2 whiskers is A0 here)
	Condone 15 marked (e.g. dash or cross) on an otherwise correct whisker
	If the outlier is at 5 and lower whisker ends at 6 award M1A1A0
(f)	1st M1 for $\frac{1}{2}\times\left(\frac{1}{4}\right)^2$
	2nd M1 for an expression of the form $pq^2\times3$ where p and q are probabilities $(p \neq q)$
	NB $\frac{3}{4}\times\left(\frac{1}{4}\right)^2\times3=\frac{9}{64}=0.140625$ is a common incorrect answer and scores M0M1A0
SC	Can award M0M1A0 if just $\frac{9}{64}$ (o.e.) is seen.
	A1 for $\frac{3}{32}$ or exact equivalent. Allow 0.0937 or 0.0938 following a correct expression.
Warning	$2\times(0.25)^2\times(0.75)$ or $2\times\left(\frac{1}{4}\right)^2\times\frac{3}{4}$ gives the correct answer but is M0M0A0

Question Number	Scheme	Marks
8. (a)	Total area of bars = 400 small squares	B1
	Area required $= 40\times4+20\times6+6\times10=340$ small squares	B1
	No of staff $= "340"\times\frac{40}{"400"}, = 34$	M1, A1
		(4)
(b)	Median is $(2+)\frac{4}{12}\times3=3$ or $(5-)\frac{8}{12}\times3=3$	M1A1
		(2)
(c)	Mean is $\frac{\sum fx}{40}=\frac{1\times16+3.5\times12+7.5\times6+15\times4+25\times2}{40}=,\frac{213}{40}=5.325$	M1,A1
		(2)
(d)	(Positive) skew but not negative or there are outliers (which affect mean)	B1
	Median	dB1
		(2)
		Total 10

	Notes
(a)	1st B1 for a correct attempt to calculate the whole area (400 small squares o.e.) Accept 160+120+60+40+20 or 80+60+30+20+10=200 or frequencies:16 + 12 + 6 + 4 + 2 or cm^2 6.4 + 4.8 + 2.4 + 1.6 + 0.8 = 16 or key: 10 small squares = 1 person (o.e.) 2nd B1 for a correct attempt to calculate required area (Accept 160 + 120 + 60) or frequencies: 16 + 12 + 6 or cm^2 6.4 + 4.8 + 2.4 = 13.6 M1 for a correct expression using their 400 and their 340 A1 for 34 If using frequencies they get M1A1 together when 34 is seen. An answer of 34 will usually score 4/4 unless there is incorrect working seen
(b)	**NB frequencies are:** 16, 12, 6, 4, 2 **and mid-points are:** 1, 3.5, 7.5, 15, 25 M1 for $\frac{20-16}{12}\times(5-2)$ or $\frac{20.5-16}{12}\times(5-2)$ or similar expressions working down Look out for methods based on areas should have 1st bar + $\frac{1}{3}$(2nd bar) if working up or (5th + 4th + 3rd bars) + $\frac{2}{3}$(2nd bar) if working down. E.g. $16 + 4(x - 2) = 20$ A1 for 3 or (if using $n + 1$ accept 3.125 or awrt 3.13)
(c)	M1 for an attempt at $\frac{\sum fx}{40}$ where at least 3 correct products of $\sum fx$ are seen or $\sum fx$ = awrt 200 (1 sf) A1 for 5.325 or any exact equivalent e.g. $\frac{213}{40}$ and accept 5.33 Accept 5 h 19 mins or 5h 20 mins
(d)	1st B1 for a reason e.g. that the data is skewed Allow mention of "extreme values" or "outliers" Do not allow for negative skew or "anomalies" 2nd dB1 dependent on mentioning skew for choosing median
SC	Allow B0B1 for "Choose median since the data has negative skew" o.e.