Qu 3	Scheme	Marks	AO
(a)	Hectopascal or hPa	B1 **(1)**	1.2
(b)	$\bar{x} = \bar{y} + 1010$ or $\frac{214}{30} + 1010$	M1	1.1b
	= 1017.1333… awrt **1017**	A1 **(2)**	1.1b
(c)	$\sigma_x = \sigma_y$ (or statement that standard deviation is not affected by this type of coding)	M1	3.1b
	$\left[\sigma_y =\right] \sqrt{\frac{5912}{30} - (\text{"7.13[33...]"})^2}$ or $\sqrt{146.1822...}$	M1	1.1b
	= 12.0905… awrt **12.1**	A1 **(3)**	1.1b
(d)	High pressure (since approx. mean + sd) so clockwise Locations are (from North to South): Leuchars, Heathrow, Hurn	B1	2.4
	Wind direction is direction wind blows from So: Heathrow (**NE**) Hurn (**E**) Leuchars (**W**)	B1 **(2)**	2.2a
		(8 marks)	

	Notes
FYI	1 hPa = 100 Pa; 10hPa = 1 kPa; 1Pa = 1 Nm^{-2}
(a)	B1 for "hectopascal" or hPa (condone pascals, allow millibars or mb) o.e. Do NOT allow kPa or kilopascals or Pa on its own
(b)	M1 for a strategy to find $\bar{x}$ Allow an attempt to find $\sum x$ that gets as far as $\sum x = \sum y + 30 \times 1010$ [= 30 514] A1 for awrt 1017 (accept 1020) [Ignore incorrect units]
(c)	1st M1 for an overall strategy using the fact $\sigma_x = \sigma_y$ (can be implied by correct final ans) or for $\sum x = 30\,514$ and $\sum x^2 = 31\,041\,192$ (both seen and correct) 2nd M1 for a correct expression (with $\sqrt{}$)(ft their $\bar{y}$ to 3sf) allow awrt 146 for 146.1822.. or for correct expression in x can ft their $\sum x > 30\,000$ or their answer to (b) A1 (dep on 2nd M1) for awrt 12.1 [Ignore incorrect units]
Final answer	Final ans of awrt 12.1 scores 3/3 **but** if they then adjust for x e.g. add 1010 (M0M1A1)
(d)	1st B1 for at least one of these reasons (these 2 lines) clearly stated (may see diagram) Need "high pressure" **and** "clockwise" to score on 1st line Contradictory statements B0 e.g. correct N~S list but say "anticlockwise" 2nd B1 (indep of 1st B1) for deducing the 3 correct directions either in the table or stated as above If the answers in table and text are different we take the table (as question says)

9MA0-31 Mark Scheme October 2021(Final)

Qu 1	Scheme	Marks	AO
(a)	**Disadvantage:** e.g. Not random; cannot use (reliably) for inferences	B1 **(1)**	1.1b
(b)	[Sight or correct use of] $X \sim B(36, 0.08)$	M1	3.3
(i)	$P(X = 4) = 0.167387\ldots$ awrt **0.167**	A1	1.1b
(ii)	$[P(X \ldots 7) = 1 - P(X \text{„} 6) =]$ 0.022233… awrt **0.0222**	A1 **(3)**	1.1b
(c)	P(In dance club and dance tango) = 0.4×0.08 = **0.032** or $\frac{4}{125}$ or **3.2%**	B1 **(1)**	1.1b
(d)	[Let T = those who can dance the Tango. Sight or use of] $T \sim B(50, \text{“}0.032\text{”})$	M1	3.3
	$[P(T < 3) = P(T \text{„} 2) =]$ 0.7850815… awrt **0.785**	A1 **(2)**	1.1b
		(7 marks)	

Notes

(a) B1 for a suitable disadvantage:

Allow (B1)	Do NOT allow (B0)
Not random or less random (o.e.)	Not representative
Cannot use (reliably) for inferences	Less accurate
(More likely to be) biased	Any comment based on time or cost
	Any mention of skew
	Any mention of non-response

(b) M1 for sight of B(36, 0.08) Allow in words: binomial with $n = 36$ and $p = 0.08$
may be implied by one correct answer to 2sf or sight of $P(X \text{„} 6) = 0.97776...$ i.e. awrt 0.98
Allow for $36C4 \times 0.08^4 \times 0.92^{32}$ as this is "correct use"

(i) 1st A1 for awrt 0.167 NB An answer of just awrt 0.167 scores M1($\Rightarrow$)1st A1

(ii) 2nd A1 for awrt 0.0222

(c) B1 for 0.032 o.e. (Can allow for sight of 0.4×0.08)

(d) M1 for sight of B(50, “0.032”) ft their answer to (c) provided it is a probability $\neq 0.08$
may be implied by correct answer
or sight of $[P(T \text{„} 3)] = 0.924348\ldots$ i.e. awrt 0.924 or $P(T \text{„} 2)$ as part of $1 - P(T \text{„} 2)$ calc.
A1 for awrt 0.785

MR Allow MR of 50 (e.g. 30) provided clearly attempting $P(T \text{„} 2)$ and score M1A0

Qu 4	Scheme	Marks	AO
(a)	Convenience or opportunity [sampling]	B1 (1)	1.2
(b)	Quota [sampling]	B1	1.1a
	e.g. Take 4 people every 10 minutes	B1 (2)	1.1b
(c)	Census	B1 (1)	1.2
(d)	[58 – 26 =] **32** (min)	B1 (1)	1.1b
(e)	$\mu = \frac{4133}{95} = 43.505263\ldots$ **awrt 43.5 (min)**	B1	1.1b
	$\sigma_x = \sqrt{\frac{202\,294}{95} - \mu^2} = \sqrt{236.7026...}$	M1	1.1b
	= 15.385… **awrt 15.4 (min)**	A1 (3)	1.1b
(f)	There are outliers in the data (or data is skew) which will affect mean and sd	B1	2.4
	Therefore use median and IQR	dB1 (2)	2.4
(g)	Value of 20, LQ at 26 and outliers will not change or state that median and upper quartile are the values that do change	B1	1.1b
	More values now below 40 than above so Q_2 or Q_3 will change and be lower	M1	2.1
	Both Q_2 and Q_3 will be lower	A1 (3)	2.4
		(13 marks)	

	Notes
(b)	1st B1 for quota (sampling) mentioned ("Stratified" or "systematic" or "random" are B0B0) 2nd B1 for a description of how such a system might work, requires suitable strata or categories e.g. time slots, departments, gender, age groups, distance travelled etc Suggestion of randomness is B0
(e)	B1 for a correct mean (awrt 43.5) M1 for a correct expression for the sd (including $\sqrt{\ }$)ft their mean A1 for awrt 15.4 (Allow s = 15.4667… awrt 15.5)
(f)	1st B1 for acknowledging outliers or skewness are a problem for mean and sd "extreme values"/"anomalies" OK May be implied by saying median and IQR not affected by.. We need to see mention of "outliers", "skewness" and the problem so "data is skewed so use median and IQR" is B0 unless mention that they are not affected by extreme values or mean and standard deviation can be "inflated" by the positive skew etc 2nd dB1 dep on 1st B1 for therefore choosing median and IQR
(g)	B1 for identifying 2 of these 3 groups of unchanged values or stating only Q_2 and Q_3 change M1 for explaining that median or UQ should be lower. E.g. the 2 values have moved to below 40 (or 58) and therefore more than 50% below 40 or (more than 75% below 58) or an argument to show that the other 3 values are the same. (o.e.) Allow arrows on box plot provided statement in words about increased % below 40 or 58 etc A1 for stating median and UQ are both lower with clear evidence of M1 scored [If lots of values on 40 then median might not change but, since two values do change then UQ would change. If this meant that 92 became an outlier then we would have a new value for upper whisker and an extra outlier so effectively 3 values are altered. So median changes]

Question	Scheme	Marks	AOs
4 (a)	$\frac{132}{184} = 0.71739...$ awrt **0.717**	B1	1.1b
		(1)	
(b)(i)	$P(X \geq 6) = 1 - P(X \leq 5)$ or $P([X =]6) + P([X =]7) + P([X =]8)$	M1	3.4
	$= 1 - 0.296722...$ awrt **0.703**	A1	1.1b
		(2)	
(b)(ii)	$184 \times P(X = 7)$ $[= 184 \times 0.2811\ldots]$	M1	1.1b
	$= 51.7385...$ awrt **51.7**	A1	1.1b
		(2)	
(c)	Part (a) and part (b)(i) are similar **and** the expected number of 7s (51.7 or 0.281) matches with the number of 7s found in the data set (52 or 0.283) so Magali's model is supported.	B1ft	3.5a
		(1)	
(d)	$\frac{23}{28} = 0.82142...$ awrt **0.821**	B1	1.1b
		(1)	
(e)	Any one of… • Part (d)/'0.821' differs from part (a)/(b)(i)/(0.7…) • there is a **greater/different probability of high cloud cover**/**more likely to have high cloud cover** if the previous day had high cloud cover • **independence(o.e.)** does not hold	B1	2.4
	…therefore Magali's (binomial) model may not be suitable.	dB1	3.5a
		(2)	
			(9 marks)

Notes	
	Allow fractions, decimals or percentages throughout this question.
(a)	Allow equivalent fraction, e.g. $\frac{33}{46}$
(b)(i)	**M1:** for writing or using $1 - P(X \leq 5)$ or $P(X = 6) + P(X = 7) + P(X = 8)$ **A1:** awrt 0.703 (correct answer scores 2 out of 2)
(b)(ii)	**M1:** for $184 \times P(X = 7)$ o.e. e.g., $184 \times [P(X \leq 7) - P(X \leq 6)]$ **A1:** awrt 51.7
(c)	**B1ft:** comparing '0.717' with '0.703' and '51.7 or '0.281' with 52 or 0.283 **and** concluding that Magali's model **is supported** (must be comparing prob. with prob. and days with days). Allow not supported or mixed conclusions if consistent with their f.t. answers in (a) and (b)
(e)	**B1:** Any bullet point **dB1:** (dep on previous B1) for Magali's model may not be suitable (o.e.) Condone not accurate for not suitable **SC:** part (d) is similar to part (a)/(b)(i) **and** a compatible conclusion (i.e. Magali's model is supported) to score B1B1.

Qu 3	Scheme	Marks	AO
(a)	[68 – 7 =] **61** (only)	B1 **(1)**	1.1b
(b)	[25 – 14] = **11**	B1 **(1)**	1.1b
(c)	$\left[\mu \text{ or } \bar{x} = \frac{607.5}{27} = \right]$ = **22.5**	B1 **(1)**	1.1b
(d)	$\sigma = \sqrt{\frac{17\,623.25}{27} - "22.5"^2}$ or $\sqrt{146.4629...}$	M1	1.1b
	= 12.10218… awrt **12.1**	A1 **(2)**	1.1b
(e)	$\mu + 3\sigma = "22.5" + 3 \times "12.1..."$ = awrt 59 so only **one** outlier	B1ft **(1)**	1.1b
(f)	Median increases implies that both values must be > 20	M1	3.1b
	Mean is the same means that $a + b = 45$	M1	1.1b
	So possible values are: e.g. $b = 21$ and $a = 24$ (o.e.)	A1 **(3)**	2.2b
(g)	Both values will be less than 1 standard deviation from the mean and so the standard deviation of all 29 values will be smaller	B1 **(1)**	2.4
		(10 marks)	

Notes

(a) B1 for correctly interpreting the box plot to find the range (more than 1 answer is B0)

(b) B1 for correct understanding of IQR and answer of 11

(c) B1 for 22.5 only (or exact equivalent such as $\frac{45}{2}$). Allow 22 mins and 30 secs.

(d) M1 for a correct expression including square root. Allow $\sqrt{146}$ or better. Ft their mean
A1 for awrt 12.1 NB Allow use of s = 12.3327… or awrt 12.3

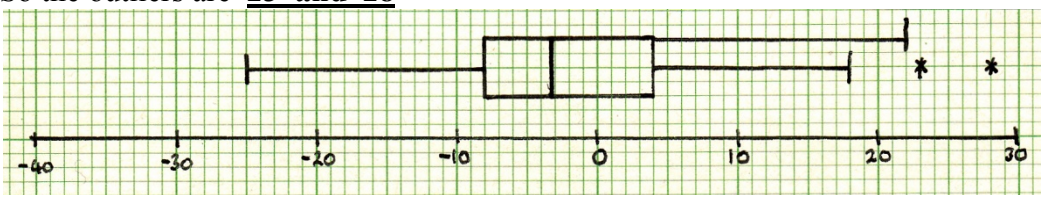
(e) B1ft for a correct calculation or value based on their μ and σ and compatible conclusion

(f) 1st M1 Correct start to the problem and a correct statement about the values based on median
Allow if their final two values are both >20
2nd M1 for a correct explanation leading to equation $a + b = 45$ (o.e. e.g. equidistant from mean)
Allow if their final two values sum to 45
A1 for a correct pair of values (both > 20 with a sum of 45) **and** at least some attempt to explain how their values satisfy at least one of the conditions (both > 20 or $a + b = 45$).
Ignore $a =$ or $b =$ labels
NB The values for a and b do not need to be integers.

(g) B1 for a correct explanation.
Must mention that both values are less than 1 sd (ft their answer to (d)) from the mean

Paper 3: Statistics and Mechanics Mark Scheme

Question	Scheme	Marks	AOs
1(a)	**Area** = $8\times1.5=12$ cm^2 **Frequency** = 8 so 1 cm^2 = $\frac{2}{3}$ **hour (o.e.)**	M1	3.1a
	Frequency of 12 corresponds to area of 18 so height = $18\div2.5$ = **7.2 (cm)**	A1	1.1b
	Width = 5×0.5 = **2.5 (cm)**	B1cao	1.1b
		(3)	
(b)	$[\bar{y}=]\ \frac{205.5}{31}=$ awrt 6.63	B1cao	1.1b
	$\left[\sigma_y=\right]\sqrt{\frac{1785.25}{31}-\bar{y}^2}$ = $\sqrt{13.644641}$ = **awrt 3.69** allow $[s=]\sqrt{\frac{1785.25-31\bar{y}^2}{30}}$ = **awrt 3.75**	M1 A1	1.1a 1.1b
		(3)	
(c)	Mean of Heathrow is higher than Hurn and standard deviation smaller suggesting Heathrow is more reliable	M1	2.4
	Hurn is South of Heathrow so does <u>not</u> support his belief	A1	2.2b
		(2)	
(d)	$\bar{x}+\sigma\approx10.3$ so number of days is e.g. $\frac{(11-"10.3")}{3}\times8\ (+5)$	M1	1.1b
	= 6.86 so **7 days**	A1	1.1b
		(2)	
(e)	[H = no. of hours] P(H > 10.3) or P(Z > 1) = [0.15865…]	M1	3.4
	Predict $31\times0.15865...$ = **<u>4.9 or 5 days</u>**	A1	1.1b
		(2)	
(f)	(5 or) 4.9 days < (7 or) 6.9 days so model may **not** be suitable	B1	3.5a
		(1)	
			(13 marks)

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]

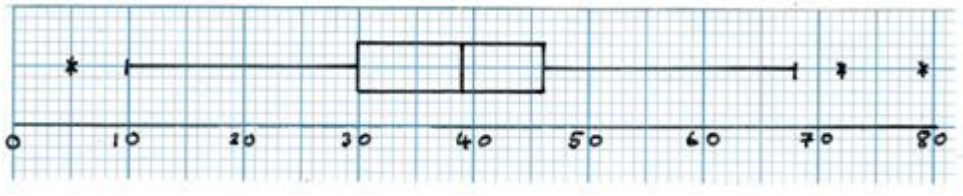
(e) B1 for stating "median > mean" and "negative skew" (independent of (d))

(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.