Statistical Diagrams and Measures Mark Scheme

Q1.

Question	Scheme	Marks	AOs
(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 $= 5\times0.5$ $=$ **2.5 (cm)**	B1cao	1.1b
		(3)	
(b)	$[\bar{y}=]\ \frac{205.5}{31}=$ awrt 6.63	B1cao	1.1b
	$[\sigma_y=]\ \sqrt{\frac{1785.25}{31}-\bar{y}^2}$ $=\sqrt{13.644641}$ **= 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 not 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)	

Q2.

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

	Notes
(a)	1st M1 for an attempt to form frequency table (at least 1st 4 rows and freq <u>or</u> cum freq seen must have the frequency of 75 correct and can condone one error/omission in 15, 35, 55) Frequencies or cum freq may be seen on bars of the histogram 1st A1 for identifying class, freq and cum freq (i.e. highlighted values from the table) <u>or</u> sight of 3.5-4.5, freq of 55 and "128" – 125 or 180 – "128" <u>or</u> diagram with 125, "128", 180, 3.5 & 4.5 May be implied by values in 2nd M1expression 2nd M1 for a correct calculation for Q_2 (condone error in end point e.g. 3.45 or 3.49 etc) Can ft their "125" (provided > 100) and their "55" Allow use of $(n + 1)$, usually see $128.5 - \ldots$ leading to 3.5636… or awrt 3.56 2nd A1 awrt 3.55 but 3.555 is fine (allow 3.56 if $(n + 1)$ being used …need sight of $\frac{257}{2}$ etc) Correct answer with no incorrect working scores 4/4
(b)	1st M1 for identifying the need to find the area under the curve by integrating 2nd M1 for correct integration and = 256 (condone missing limits) A1 for $k = 3$ [May see use of calculator for the integration so score 2nd M1A1 together]
(c)	**NB The answer to part (c) may be written within the question.** B1 for 4 (Independent of their value of k but must be their "x" value) NB when $k = 0.25$ and $x = 4$ gives $y = 4$ so must be clear they intend median = 4 The statement in part (c) "$k = 4$" is B0

Q3.

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

Part	Notes
	1st M1 for a fully correct expression for the number of motorists in the interval
	2nd 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.
	1st A1 for both correct values seen
	3rd M1 for realising that total is required and attempting a correct expression for %
	2nd A1 for awrt 67.7%

Q4.

Qu	Scheme	Marks	AO
(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 =] $40+\dfrac{\frac{1}{2}("331")-140}{"130"}\times 25$ or $65-\dfrac{270-\frac{1}{2}("331")}{"130"}\times 25$ (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

Q5.

	Scheme	Marks	AO
(a)	[68 – 7 =] **61** (only)	B1 (1)	1.1b
(b)	[25 – 14] = **11**	B1 (1)	1.1b
(c)	$\left[\mu \text{ or } \bar{x} = \frac{607.5}{27} = \right] = \mathbf{22.5}$	B1 (1)	1.1b
(d)	$\sigma = \sqrt{\frac{17\,623.25}{27} - "22.5"^2}$ or $\sqrt{146.4629...}$ = 12.10218… awrt **12.1**	M1 A1 (2)	1.1b 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 Mean is the same means that $a + b = 45$ So possible values are: e.g. $b = 21$ and $a = 24$ (o.e.)	M1 M1 A1 (3)	3.1b 1.1b 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)	

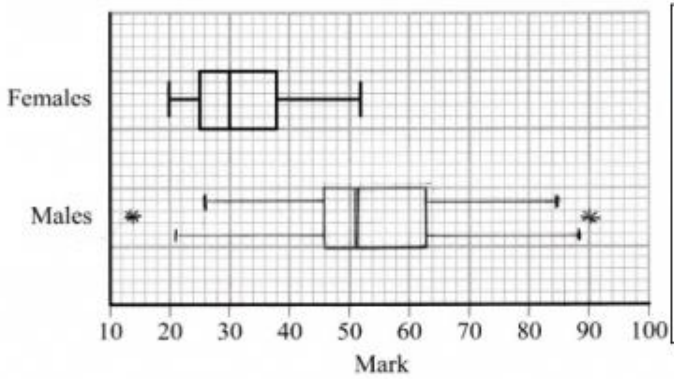
	Notes
(a)	B1 for correctly interpreting the box plot to find the range (more than 1 answer is B0)
(b)	B1 for correct understanding of IQR and answer of 11
(c)	B1 for 22.5 only (or exact equivalent such as $\frac{45}{2}$). Allow 22 mins and 30 secs.
(d)	M1 for a correct expression including square root. Allow $\sqrt{146}$ or better. Ft their mean A1 for awrt 12.1 NB Allow use of $s = 12.3327\ldots$ 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

Q6.

Question	Scheme	Marks	AOs
	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

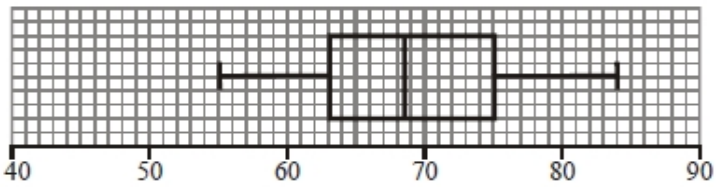
Notes		
	M1:	For clear use of frequency density to establish the fd scale and then use the area to find frequency of <11 minutes. Allow maximum of 3 errors in either the heights or widths in total if working shown. They may calculate the area using other size squares. Allow for realising they need to 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

Q7.

Question	Scheme	Marks
(a)	25 (allow any x where $24 < x < 26$)	B1 **(1)**
(b)	Q_2 (or median or m) = **51** **IQR** = 63 – 46 .= **17** (or $Q_3 - Q_1 = 17$)	B1 M1, A1 **(3)**
(c)	Outliers given by 46 – 1.5 ×17 = 20.5 or 63 + 1.5 ×17 = 88.5 Outliers limits are **20.5** and **88.5**  Allow lower whisker to 20.5 and upper whisker to 88.5 Do **not** allow a mix of whiskers e.g 20.5 and 85 Do **not** allow both sets of whiskers	M1 A1 M1 A1ft B1 **(5)**
(d)	**Medians:** Median for females lower than males **IQR:** IQR for females smaller than males. Allow "lower/higher" but not "wider" **Range:** Range of females is less than males **Skewness:** Male and female marks are both positively skew Ignore other statements about average, spread, mean, st. Dev, variation, outliers etc	B1ft B1ft **(2)** **(11 marks)**

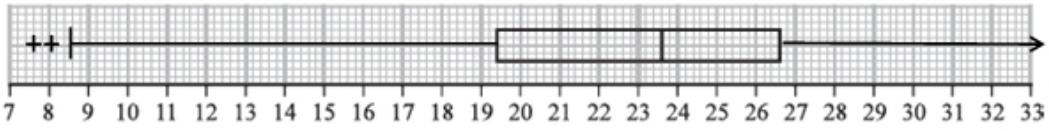
	Notes
	Mark (b) and (c) together BUT must see clear statement that median (or m or Q_2) = 51 and IQR = 17
(b)	M1 for 2 quartiles (at least one correct) and attempt to find the difference. Must see their 63 – their 46 A1 for 17 only. [Answer only of IQR= 17 scores M1A1]
(c)	**A fully correct box-plot (either version) with no supporting work scores 5/5. Otherwise:** 1st M1 for correct attempt to calc' at least one limit for outliers, ft their quartiles or IQR or award for sight of 20.5 or 88.5 1st A1 for identifying both limits of 20.5 and 88.5 2nd M1 for a box with an upper and a lower whisker(s) with at least 2 correct values (or correct ft) (condone no median marked) (condone 2 upper or 2 lower whiskers) 2nd A1ft for their 20.5 or 26 ,46, 51, 63 and 85 or their 88.5 in appropriate places and readable off their scale. Follow through their 20.5 and their 88.5 only, other values need to be correct If there are 2 upper or 2 lower whiskers A0 B1 for only 2 outliers appropriately marked at 14 and 90 Do not award if whiskers go beyond these values. **Apply ± 0.5 square accuracy for diagram** A box plot not on the graph paper can only score the 1st M1A1
(d)	**In (d) ft from their diagrams (if no diagram then use their values)** 1st B1ft for one correct comment comparing median, IQR , range or skewness 2nd B1ft for a second correct comment comparing median, IQR, range or skewness Do not allow contradictory statements

Q8.

Question	Scheme	Marks
(a)	[Range = 48 − 9] = **39**	B1 (1)
(b)	[IQR = 25 − 12]= **13**	B1 (1)
(c)	Median = $65+\frac{[9]}{13}\times 5 = \frac{890}{13}$ = awrt **68.5**° $\left[\text{Condone: } 65+\frac{[9.5]}{13}\times 5 = 68.7\right]$	M1 A1 (2)
(d)	Lower Quartile = $60+\frac{9}{15}\times 5 = \mathbf{63}$ (*)	M1 A1cso (2)
(e)(i)	$63-1.5\times(75-63)=45$ $75+1.5\times(75-63)=93$ No data above 93 and no data below 45 or 55>45 etc or there are no outliers.	M1A1 A1
(ii)		M1 A1ft (5)
(f)	Median for the 70°angle is closer (to 70°)[than the 20° median is to 20°] The range/IQR for the 70° angle box plot is smaller/shorter Therefore, students were more accurate at drawing the 70° angle.	B1 B1 dB1 (3)
		(14 marks)

	Notes
(c)	M1 for an attempt (should have 65 or 70, 13 and 5)NB working down: $70-\frac{[4]}{13}\times 5$ Allow any correct method leading to $\frac{890}{13}$, the "5" may be implied by 65 and 70 seen A1 awrt 68.5 (condone 68.7 if $(n+1)$ is used). Ans only of 68.5 is 2/2 but 68.7 needs M1
(d)	M1 for correct expression for the lower quartile (condone 9.25 if $(n+1)$ used) Watch out for working down e.g. $65-\frac{6}{15}\times 5$ (M1) but e.g. $\frac{60+65}{2}=62.5=63$ is M0 A1 for correct solution with no incorrect working seen (condone $(n+1)$ giving 63.08..)
(e)(i)	M1 for either correct calculation (may be implied by one correct limit) A1 for either 45 or 93 A1 for 45 and 93 and conclusion
(ii)	M1 for a box with 1 whisker drawn on each side (must see the line drawn) A1ft their median $63 < Q_2 < 75$ but quartiles (63 and 75), 55 and 84 must be correct.
Accuracy	Use 0.5 sq. accuracy so condone median on 68 or 69 if 68.5 seen
(f)	1st B1 for correct comparison of their **medians** (63 < (c) < 75) to true value 2nd B1 for correct comparison of their **range** or **IQR** ("spread" is B0) Allow saying IQRs of 12 and 13 are similar. Ignore mention of "skewness" or "outliers" 3rd dB1 dependent upon at least one previous B1 being scored for choosing 70°

Q9.

Part	Working or answer an examiner might expect to see	Mark	Notes
(a)	IQR = 2.6. – 19.4 = 7.2	B1	This mark is given for finding the interquartile range
	19.4 – (1.5 × 7.2) = 8.6 19.4 + (1.5 × 7.2) = 37.4	M1	This mark is given for a method find the values for the whiskers of the boxplot
			
		A1	This mark is given for plotting the correct whisker (8.6) on the boxplot
		A1	This mark is given for plotting the two correct outliers 7.6 °C and 8.1 °C
(b)	October (since it is the month with the coldest temperatures between May and October in Beijing)	B1	This mark is given for a correct suggestion with a supporting reason.
(c)	$\sigma = \sqrt{\frac{S_{xx}}{n}} = \sqrt{\frac{4952.906}{184}} = \sqrt{26.92} = 5.19$	B1	This mark is given for showing the calculation for the standard deviation to three significant figures
(d)	$z = (\pm)\ 1.2816$	B1	This mark is given for identifying the z-value for the 10th and 90th percentiles (from tables or calculator)
	$2 \times z \times 5.19$	M1	This mark is given for a method to find the interpercentile range between the 10th and 90th value
	= 13.303	A1	This mark is given for finding a correct interpercentile range between the 10th and 90th value
(e)	Daily wind speed (Beaufort) since it is qualitative data	B1	This mark is given for stating a correct variable with a supporting reason
	Rainfall (since it is not symmetric)	B1	This mark is given for stating a correct variable with a supporting reason
			(Total 11 marks)

Q10.

Question Number	Scheme	Marks
(a)	1(cm) cao	B1
(b)	10 cm^2 represents 15 10/15 cm^2 represents 1 or $1cm^2$ represents 1.5 Therefore frequency of 9 is $\frac{10}{15}\times 9$ or $\frac{9}{1.5}$ Require $x\frac{2}{3}$ or $\div 1.5$ height = 6(cm)	M1 A1 [3]
Notes	If 3(a) and 3(b) incorrect, but their (a) x their (b)=6 then award B0M1A0 3(b) Alternative method: f/cw=15/6=2.5 represented by 5 so factor x2 award M1 So f/cw=9/3=3 represented by 3x2=6. Award A1.	