



eClassroom

INTERNATIONAL A LEVEL MATHEMATICS • STATISTICS 1

Statistics 1

Exam Questions by Topic: Mark Scheme

Pearson Edexcel International Advanced Level (WST01/01) • Calculator and Statistical Tables allowed

Every question in the booklet, marked line by line the way an examiner marks it

Chapters (tap a chapter to jump to it)	Qs	Marks	Page	My score
1 Mathematical Modelling Not examined on its own: modelling ideas are tested within Chapters 2 to 7	—	—	Not applicable	
2 Measures of Location and Spread Averages, quartiles, interpolation, standard deviation, coding	10	83	own PDF	
3 Representations of Data Histograms, stem and leaf, box plots, outliers, skewness, comparisons	10	80	3	
4 Probability Venn and tree diagrams, conditional probability, independence	10	80	own PDF	
5 Correlation and Regression PMCC, least squares regression, interpretation, coding	10	84	own PDF	
6 Discrete Random Variables Probability and cumulative distributions, $E(X)$, $\text{Var}(X)$, uniform	10	82	own PDF	
7 The Normal Distribution Standardising, tables, finding μ and σ , conditional problems	10	83	own PDF	

[Question booklet](#)

[Free courses](#)

[More worksheets](#)



Join students already acing their exams.

Book your FREE lesson at www.eclassroom.org.uk

How to use this mark scheme

The same marking rules Pearson examiners use for WST01.

Mark every line, not just the answer. Method marks are given for a correct method even when the final answer is wrong. If you lost marks, find the *first* line where your working differs from the scheme. That is the thing to fix.

Code	Meaning
M1	Method mark: for knowing a correct method and attempting to apply it. Small slips in the working do not lose this mark.
dM1	Dependent method mark: can only be given if the method mark before it has been given.
A1	Accuracy mark: for a correct answer or line of working. It can only be given if the method mark it depends on has been given.
B1	Independent mark: for a correct result or statement, whatever method was used.
*	The answer is given in the question. Every step must be shown, and the answer must follow from correct working (cso).
cao	Correct answer only.
cso	Correct solution only: no errors anywhere in the working.
ft	Follow through: the mark can be given for correct work using your own earlier (wrong) answer.
oe	Or equivalent: any correct equivalent form is accepted.

“Solutions relying on calculator technology are not acceptable.” When a question says this, a correct final answer with no algebraic working scores **no marks**. The notes under each scheme line say when this applies.



Join students already acing their exams.

Book your FREE lesson at www.eclassroom.org.uk

3

Chapter 3: Representations of Data

Mark scheme. Marks are shown on the right of each line.

Question 1

5 marks

Part	Scheme	Marks
(a)	Width: 2 minutes $\leftrightarrow$ 1 cm, so 8 minutes gives 4 cm Frequency densities: $\frac{20}{2} = 10$ and $\frac{40}{8} = 5$; height $\propto$ frequency density $5 \times \frac{5}{10} = \mathbf{2.5 \text{ cm}}$ <i>Scaling the height by the frequencies alone ($5 \times 2 = 10 \text{ cm}$) ignores the change in class width: M0.</i>	B1 M1 A1 (3)
(b)	$\frac{2}{5} \times 32 + 20 + \frac{3}{8} \times 40$ 48 (47.8)	M1 A1 (2)
Total for Question 1: 5 marks		

Question 2

5 marks

Part	Scheme	Marks
(a)	Median = 30 (10th value) Q_1: $\frac{19}{4} = 4.75$, 5th value 23 Q_3: 15th value 41	B1 B1 B1 (3)
(b)	Limits: $41 + 1.5 \times 18 = 68$ and $23 - 27 = -4$ 72 is above 68 (63 is not), so 72 is the only outlier *	M1 A1* (2)
Total for Question 2: 5 marks		



Question 3

5 marks

Part	Scheme	Marks
(a)	Median 48	B1
	$IQR = 60 - 35 = 25$	B1
		(2)
(b)	Location: school B has a slightly higher median (50 compared with 48)	B1
	Spread: school B's marks are less spread, IQR 13 compared with 25	B1
	<i>Each comparison must be in context (marks) with the values.</i>	(2)
(c)	$Q_3 - Q_2 = 5 < Q_2 - Q_1 = 8$, so negative skew	B1
		(1)
Total for Question 3: 5 marks		

Question 4

5 marks

Part	Scheme	Marks
(a)	$\frac{3(29.4 - 26)}{11.2} = 0.911$	M1
	Positive skew	A1
		(2)
(b)	$\frac{38 - 52 + 21}{17} = \frac{7}{17}$	M1
	0.41	A1
		(2)
(c)	Median and IQR, because the data are skewed (and they are not affected by extreme values)	B1
		(1)
Total for Question 4: 5 marks		



Question 5

5 marks

Part	Scheme	Marks
(a)	Stems 2 to 5 with leaves in order: 2 7 8 9; 3 1 3 4 6 7 9; 4 0 1 1 4 5 8; 5 2 Correct diagram with a key, e.g. 3 1 means 31 minutes	M1 A1 (2)
(b)	$n = 16$: median = $\frac{37 + 39}{2} = 38$ $Q_1 = \frac{31 + 33}{2} = 32$ and $Q_3 = \frac{41 + 44}{2} = 42.5$ IQR = 10.5	B1 M1 A1 (3)
Total for Question 5: 5 marks		

Question 6

11 marks

Part	Scheme	Marks
(a)	Median is halfway between the 15th and 16th values: $\frac{2w + 26}{2} = 25.5$, so $w = 5$ $x = 16$ (8th value) $y = 37$ (23rd value)	B1 B1 B1 (3)
(b)	$37 + 1.0 \times (37 - 16) = 58$ 68 is the only outlier <i>Using 1.5 instead of the rule given (limit 68.5, so no outliers) scores M0 A0.</i>	M1 A1 (2)
(c)	Box from 16 to 37 with the median at 25.5 Whiskers to 7 and 52 Outlier 68 marked with a cross	M1 A1 A1 (3)
(d)	$\frac{(37 - 25.5) - (25.5 - 16)}{37 - 16} = \frac{2}{21}$ 0.095	M1 A1 (2)
(e)	The total is estimated by mean $\times$ number of days; the mean uses all the data (including the high values)	B1 (1)
Total for Question 6: 11 marks		



Question 7

11 marks

Part	Scheme	Marks
(a)	Area of the 5–10 bar represents 40 people, so 1 unit of area ($1 \text{ km} \times 1 \text{ unit of density}$) is 1 person Frequencies 20, 40, 50, 30, 20 160	M1 M1 A1 (3)
(b)	Part of the 5–10 bar: $\frac{2}{5} \times 40 = 16$ Whole 10–20 bar, and half of the 20–30 bar: $50 + \frac{5}{10} \times 30 = 65$ 81	M1 M1 A1 (3)
(c)	80th value, in the 10–20 class $10 + \frac{80 - 60}{50} \times 10$ 14 km	M1 M1 A1 (3)
(d)	Midpoints 2.5, 7.5, 15, 25, 40: $\frac{2650}{160}$ 16.6 km	M1 A1 (2)
Total for Question 7: 11 marks		



Question 8

11 marks

Part	Scheme	Marks
(a)	$Q_2 = 72$ (8th value) $Q_1 = 66$ (4th value) $Q_3 = 79$ (12th value) <i>Read the non-athletes' leaves left to right. Reading the wrong side of the diagram is the common error.</i>	B1 B1 B1 (3)
(b)	Upper limit $79 + 1.5 \times 13 = 98.5$ and lower limit $66 - 19.5 = 46.5$ 101 is above 98.5 and no value is below 46.5, so 101 is the only outlier *	M1 A1* (2)
(c)	The athletes have lower pulse rates: median 56 compared with 72 The athletes' pulse rates are less spread: IQR 9 compared with 13 Both comparisons in context with values	B1 B1 B1 (3)
(d)	Athletes: $Q_3 - Q_2 = 5 > Q_2 - Q_1 = 4$, so positive skew Non-athletes: $Q_3 - Q_2 = 7 > Q_2 - Q_1 = 6$, so positive skew	B1 B1 (2)
(e)	A non-athlete : 69 is above the athletes' Q_3 (61) but within the non-athletes' IQR (66 to 79)	B1 (1)
Total for Question 8: 11 marks		



Question 9

10 marks

Part	Scheme	Marks
(a)	$\bar{x} = \frac{227}{12} = 18.9 \text{ cm}$ $\sigma^2 = \frac{4949}{12} - 18.916\dots^2$ $\sigma = 7.39 \text{ cm}$	B1 M1 A1 (3)
(b)	$\bar{x} + 2\sigma = 18.9\dots + 2 \times 7.38\dots = 33.7$ $42 > 33.7$, so 42 is an outlier *	M1 A1* (2)
(c)	$n = 11, \sum x = 185, \sum x^2 = 4949 - 42^2 = 3185$ Mean 16.8 cm Standard deviation 2.59 cm	M1 A1 A1 (3)
(d)	Tray P seedlings are taller on average (mean 16.8 compared with 15.2 cm) Tray P heights are less spread (sd 2.59 compared with 3.9 cm)	B1 B1 (2)
Total for Question 9: 10 marks		



Question 10

12 marks

Part	Scheme	Marks
(a)	Limits: $22 + 1.5 \times 8 = 34$ and $14 - 12 = 2$ 36 is above 34 and 33 is not; 8 is above 2. So exactly one outlier, 36 *	M1 A1* (2)
(b)	Box from 14 to 22 with the median at 18 Whiskers to 8 and 33 Outlier at 36 marked with a cross	M1 A1 A1 (3)
(c)	Type B grows more: median 24 compared with 18 cm Type B growth is more spread: IQR 11 compared with 8 cm	B1 B1 (2)
(d)	$g > 28$ (above type B's Q_3) and $g \leq 34$ (not an outlier for type A) $28 < g \leq 34$	M1 A1 (2)
(e)	Type B IQR = 11: limits $28 + 16.5 = 44.5$ and $17 - 16.5 = 0.5$ Two ranges $h > 44.5$ or $h < 0.5$ <i>The final mark needs both ranges with the correct inequality signs.</i>	M1 M1 A1 (3)
Total for Question 10: 12 marks		

