Box Plots Mark Scheme
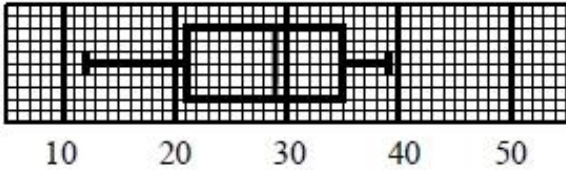
Q1.

Question	Answer	Additional guidance	Mark
(a)	\n\nMemory test score	M1 for a box with two whiskers drawn with at least 3 values correct\nA1 fully correct	(2)
(b)	B1 for eg median for teenagers greater than median for parents\n\nB1 for eg IQR/range for parents greater than IQR for teenagers\n\nB1 for eg both negative skew\n\ndepB1 for eg "The teenagers did better on the memory test than their parents" OR "The teenagers were more consistent on the memory test"	B1 for a correct statistical statement comparing the medians\n\nB1 for a correct comparison of the IQRs or ranges\n\nB1 for a correct comparison of the skews\n\ndepB1 for a correct contextual interpretation comparing medians or IQR/ranges (dependent upon previous B1 being scored)	(4)
(c)	M1 56×0.75\nA1 42	M1 for identifying 75% or $\frac{3}{4}$ or 0.75\nA1 for 42	(2)

Q2.

Question	Scheme	Marks
(a)	$9.5 - 3 =$ 6.5	M1 A1cao (2)

(b)	Box plot drawn box with at least one whisker
	                                  2, 3 and 10 plotted correctly
	                                  all correct (0, 2, 3, 10 and 15)
		B1
		 B1
		 B1
		 (3)

*(c)	The distribution is not symmetrical since… It has (positive) skew(ness)	B1 B1 (2)
		[7]
	Notes	
(a)	M1 for $k - 3$ where $9 < k < 10$	
(c)	1st B1 for not symmetrical/not evenly distributed/no PLUS any reason 2nd B1 for <u>skew</u> or a correct description of skewness which involves the <u>median</u> and a <u>quartile</u> (e.g. 'The median is closer to the lower quartile '). No/Not symmetric **on its own** is 1st B0 Negative skew here is 2nd B0 Must use correct statistical language. Condone poor spelling if intention is clear.	

Q3.

Question	Scheme	Marks
*(a)	Any two from: • The <u>median</u> is higher in January. • The <u>ranges</u> are the same/the <u>IQR</u> is lower in January. • They both have <u>positive skew</u>.	B1 B1 (2)
(b)	25% complete the race in less than 24 minutes/24 is the lower quartile in January <u>or</u> 50%/24 is the median in July. A greater proportion finish the race in less than 24 minutes in July	B1 B1 (2) [4]
	Notes	
(a)	Allow each bullet point once. Must use correct statistical language to score the marks here. Must be a comparison (not just statement of values) Allow e.g. 'July is more positively skewed than January' for 3rd bullet point Ignore extraneous non-contradictory comparisons	
(b)	1st B1 for a correct proportion for either month 2nd B1 for a correct comparison Condone 'more athletes' for greater proportion	

Q4.

Question	Answer	Additional guidance	Mark
(a)	M1 for box with two whiskers AND at least two values plotted correctly A1 for all correct		(2)
(b)	B1 for eg • median reaction time for 30-39 year olds is greater than for 20-29 year olds • median reaction times increase as age increases	B1 for a correct statistical statement comparing the medians	(4)
	B1 for eg • IQR for 40-49 year olds greater than IQR for 20-29 year olds • 30-39 year olds have the greatest IQR • range for 30-39 year olds greater than range for 20-29 year olds • range of reaction times increases as age increases	B1 for a correct comparison of the IQRs or ranges	
	B1 for eg • 20-29 year olds positively skewed, 30-39 year olds negatively skewed	B1 for a correct comparison of the skews Allow positive or negative skew for 40-49 year olds box plot	
	B1 for eg reaction times increase as you get older (on average) OR reaction times for the 20-29 year olds most consistent	B1 for a correct contextual interpretation comparing medians or IQR/ranges/skews	

Q5.

Question	Scheme	Marks
(a)		B1 M1 A1 (3)
(b)	Year 7 is symmetric Year 9 is positively skewed	B1 B1ft (2)
(c)	$76 \times 0.75 = 57$	M1A1 (2) **[7]**

	Notes	
(a)	B1 for median plotted at 1.86 M1 for a box with 2 whiskers drawn with 1.40, 1.68 and 2.26 correct A1 fully correct SC: If 0 scored, 1.86 in the correct place in the table or 2.16 seen scores B1M0A0	
(b)	B1 for Year 7 is symmetric / no skew condone 'symmetrical skew' B1 for Year 9 is positively skewed or allow ft for negative skew from a fully drawn box plot with their (median – lower quartile) > their (upper quartile – median) (Positive) correlation is B0 SC: Both box plots have negative skew B2	
(c)	M1 for identifying 75% or ¾ or 0.75 A1 cao	

Q6.

Question	Scheme	Marks
(a)	The median is (£)240 000 (so more than half the houses cost more than (£)200 000)	B1 (1)
(b)	$2.9 - 1.8 = (1.1)$ $2.9 + 1.1 \times 1.5 (= 4.55)$ $500\,000 > 455\,000$, so it is an outlier.	M1 M1 A1 (3) **[4]**

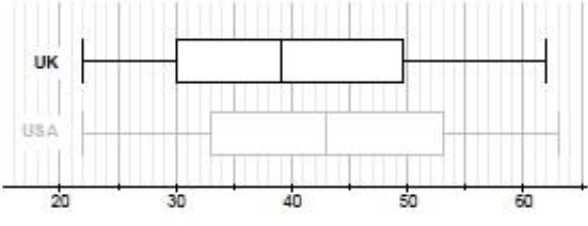|	Notes	
(a)	For either the median is 240 000 <u>or</u> the median is more than 200 000 <u>or</u> for stating that half of the houses cost more than 240 000 (Allow 2.4 for 240 000 and 2 for 200 000)	
(b)	1st M1 for $2.9 - 1.8$ (290 000 – 180 000) or IQR =1.1 (110 000) 2nd M1 for $2.9 + 1.5 \times$ their IQR (or 290 000 + 1.5 × their IQR) A1 dependent on both M1 marks for a correct **comparison** of 500 000 and 455 000 <u>or</u> 5 > 4.55 <u>or</u> for concluding that 500 000 is an outlier	

Q7.

Question	Scheme	Marks
(a)	Median = 84	B1
	IQR = 94 − 80	M1
	$\qquad$ = 14	A1
		(3)
*(b)	Data (for salad potatoes) is <u>skewed</u> or <u>not symmetrical</u> or has <u>outliers</u>, so use median and IQR (or 'option 1')	B2,1,0
		(2)
*(c)	New potatoes have a smaller <u>median</u> than salad potatoes	B1ft
	New potatoes have a smaller <u>IQR</u> than salad potatoes	B1ft
		(2)
	ALTERNATIVE	
	New potatoes have a smaller <u>mean</u> than salad potatoes	B1
	New potatoes have a smaller <u>standard deviation</u> than salad potatoes	B1
		[7]

	Notes	
(a)	M1 Subtraction using attempts at Q_1 and Q_3 (at least one of 80, 94 correct)	
(b)	QWC	
	B2 Correct choice **and** reason, using the correct vocabulary (underlined) (Condone 'anomaly' for outlier)	
	OR	
	B1 for a partially correct answer which includes a sensible reason (e.g. "box plot is skewed" but no conclusion, or "there are *extreme* values so use median & IQR" – i.e. lacking correct vocab.)	
	NB Correct choice with **no reason** scores B0	
(c)	QWC	
	Comparison (not just listing) using the correct vocabulary (underlined - allow s.d. for standard deviation). Follow through their answers in (a).	
	Allow correct converse statements, and allow clear equivalents to 'smaller' etc	
	Condone mention of both types of measure.	
	SC: if compare mean and IQR only, or median and s.d. only, score max B1B0	

Q8.

Question number	Answer	Additional guidance	Mark
(a)	B2 Accept answers in the range $6.0 \leq Q_1 \leq 6.5$, $Q_2 = 10$, $15.0 \leq Q_3 \leq 15.5$	B2 for all three values correct (quartiles in ranges) OR B1 for one value correct	(2)
(b)	M1 for a box with two whiskers AND at least two values plotted correctly from any of 2.0, 22.3, Q_1, Q_2, Q_3 (correct or follow through) A1ft for all correct with 2.0, 22.3 and median and quartiles (median and quartiles correct or follow through from (a))	For M1 and A1 Allow 22.3 plotted at between 22.0 and 22.5 Allow $\pm \frac{1}{2}$ small square accuracy on Q_1, Q_2, Q_3.	(2)

(c)	B1ft e.g. median height for Oak trees greater or median height for maple trees smaller.
	
	B1ft e.g. IQR/range smaller for Maple trees or IQR/range greater for Oak trees.
	
	B1ft e.g. Maple symmetrical and Oak positively skewed
	
	depB1ft for any of the above interpreted in context e.g.
	  • Oak <u>trees are taller</u> (on average) or Maple <u>trees are shorter</u> on (on average)
	  • The <u>heights</u> of Oak trees are more widely dispersed, or the <u>heights</u> of Maple trees are more consistent.
	  • The spread of <u>heights</u> above the average (median) for the Oak trees is greater than the spread of <u>heights</u> below the average, whereas the spread of heights above and below the average for the Maple trees is broadly the same.
		B1ft for a correct statistical statement comparing the medians (ft (a) or (b)) Condone taller for comparison of medians.
		Condone misspellings but medium is B0.
		B1 for a correct comparison of the IQRs or ranges (ft (a) or (b)) Condone wider for comparison of IQR/range.
		
		B1 for a correct comparison of the skews (ft (b)) Condone both positively skewed.
		For symmetrical accept no skew but not neutral skew, normal skew or symmetrical skew.
		
		depB1ft for a correct contextual interpretation comparing medians or IQR/ranges or skew
		Dependent on correct statistical comparison having been made to support the interpretation given.
		
		Note: in this question ignore any numerical values in comparisons.
			(4)

Q9.

Question	Scheme	Marks
(a)	$30, (39 \leq M < 40), (49 < UQ \leq 50)$	B2 (2)
(b)		M1 A1ft (2)
(c)	1) UK have lower **median** OR USA have higher **median**	B1ft
	2) **IQR** is similar / the same for both (or UK slightly lower **IQR**), OR **range** almost the same for both (or UK slightly lower **range**)	B1ft
	3) USA symmetrical/no skew <u>AND</u> UK symmetrical (or slight positive skew)	B1ft
	4) Teachers in UK are younger OR Head Teacher is correct	B1ft (4) [8]
	Notes	
(a)	B2 All three correct (or B1 for at least one correct) Median/UQ in given ranges. Do **not** accept 40 or 49	
(b)	M1 box with two whiskers <u>and two correct (ft) values</u> from five. (condone missing median for this mark) A1ft all correct with 22, 62 and their median+quartiles from (a) ($\pm\frac{1}{2}$ line tolerance)	
(c)	Max one mark from each bullet point. Must be <u>comparisons</u>, not just listing values. **<u>COMMENTS MUST BE CONSISTENT WITH THEIR BOX PLOT</u>**, (if no box plot, ft from table, otherwise cao) Words in bold must be used in comparisons for each of first three marks. (Condone poor spelling, but not 'medium'.) **For point 1**: allow smaller/bigger for lower/higher. (condone younger/older) NB: allow medians stated without an explicit comparison ONLY IF they have a correct conclusion – i.e. if they also score for point 4 **For point 2**: do not accept 'wider'/'narrower' for larger/smaller range or IQR **For point 3**: in description of UK allow skew / slight skew / positive skew <u>if correct ft</u>. (Could be 'negative skew' on ft) **For point 4**: Accept 'yes' if clearly meant as an answer to the stated question. Note, do NOT accept direct comparison of individual max, min, LQ, UQ values	