Statistical Diagrams and Measures

Q1, (Edexcel 6683, Jan 2010, Q2)

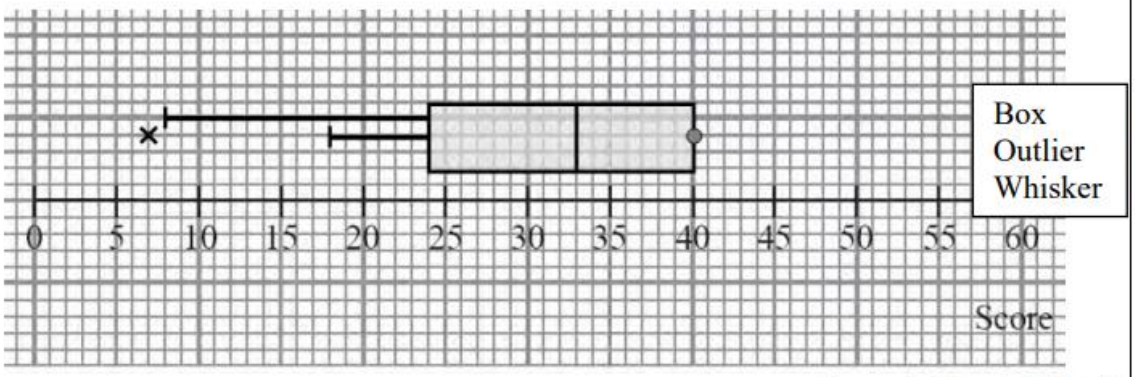
(a) Median is 33 — B1 (1)

(b) $Q_1 = 24, Q_3 = 40, \text{IQR} = 16$ — B1 B1 B1ft (3)

(c) $Q_1 - \text{IQR} = 24 - 16 = 8$ — M1

So 7 is only outlier — A1ft (2)

(d)



B1ft B1 B1ft (3)

(accept either whisker)

Total [9]

Q2, (Edexcel 6683, Jun 2007, Q5)

(a) 18-25 group, area=7x5=35 — B1

25-40 group, area=15x1=15 — B1 (2)

(b) (25-20)x5+(40-25)x1=40 — M1A1 (2)

(c) Mid points are 7.5, 12, 16, 21.5, 32.5 — M1

$\sum f = 100$ — B1

$\frac{\sum ft}{\sum f} = \frac{1891}{100} = 18.91$ — M1A1 (4)

(d) $\sigma_t = \sqrt{\frac{41033}{100} - \bar{t}^2}$ $\sqrt{\frac{n}{n-1}\left(\frac{41033}{100} - \bar{t}^2\right)}$ alternative OK — M1

M1

$\sigma_t = \sqrt{52.74...} = 7.26$ — A1 (3)

(e) $Q_2 = 18$ or 18.1 if (n+1) used — B1

$Q_1 = 10 + \frac{15}{16} \times 4 = 13.75$ or 15.25 numerator gives 13.8125 — M1A1

$Q_3 = 18 + \frac{25}{35} \times 7 = 23$ or 25.75 numerator gives 23.15 — A1 (4)

Q3, (Edexcel 6683, Jun 2012 Q5)

(a)	One large square = $\frac{450}{"22.5"}$ or one small square = $\frac{450}{"562.5"}$ (o.e. e.g. $\frac{"562.5"}{450}$)	M1	
	One large square = 20 cars or one small square = 0.8 cars or 1 car = 1.25 squares	A1	
	No. > 35 mph is: 4.5×"20" or 112.5×"0.8" (or equivalent e.g. using fd)	dM1	
	= **90** (cars)	A1	(4)
(b)	$[\bar{x}] = \frac{30\times12.5+240\times25+90\times32.5+30\times37.5+60\times42.5}{450}$ $\left[=\frac{12975}{450}\right]$	M1 M1	
	$= 28.83...$ or $\frac{173}{6}$ **awrt 28.8**	A1	(3)
(c)	$[Q_2 =]\ 20+\frac{195}{240}\times10$ (o.e.) [Allow use of $(n+1)$ giving 195.5 instead of 195]	M1	
	$= 28.125$ [Use of $(n+1)$ gives 28.145…] **awrt 28.1**	A1	(2)
(d)	$Q_2 < \bar{x}$ So positive skew [Condone $Q_2 \approx \bar{x}$] [so (almost) symmetric]	B1ft dB1ft	(2)
(e)	[If chose skew in (d)] **median (Q_2)** Since the data is skewed or median not affected by extreme values [If chose symmetric in (d)] **mean ($\bar{x}$)** Since it uses all the data	B1 dB1	(2)
			[13]

Q4, (OCR 4732, Jun 2005, Q5)

(i) Read at 300 or 300.25 and 900 or 900.75	M1		or 44-46 and 68-70 incl.
44.5 to 45.5 and 69 to 69.9	A1		
IQR 23.5 to 25.4	A1	3	dep A1 Must look back, see method. No wking, ans in range: M1A1A1
(ii) 0.6 or 60%	M1		Seen or implied
CF 720	M1		Seen or implied
63 to 64	A1	3	55.5 to 56: SC B1
(iii) 1200 – 860	M1		Allow 1200 – (850 to 890)
= 340	A1	2	310 to 350
(iv) 340/1200	M1		their (iii)/1200
0.283^5	M1dep		[their (iii)/1200]^5 exactly
= 0.00183	A1	3	Allow 0.00114 to 0.00212 ≥ 2 sfs
			${}^{340}C_5/{}^{1200}C_5$ M1
(v) Incorrect reason or ambiguity: B0B0. Otherwise:			eg IQR = 55–35 = 20 or IQR = value >27
Too low,			or new info' implies straight line: B1
or should be 26 or 27 or 2 or 3 higher	B2	2	or originally, majority in range 35 – 55 are at top of this range: B1

Q5, (OCR 4732, Jun 2008, Q6)

(i)(a)	256	B1 1	
			(i)(b) & (ii)(abc): ISW ie if correct seen, ignore extras
(b)	Total unknown or totals poss diff or Y13 may be smaller or similar or size of pie chart may differ	B1 1	pie chart shows only proportions oe or no. of students per degree may differ not "no. of F may be less" not "Y13 may be larger"
(ii)(a)	B&W does not show frequencies oe	B1 1	or B&W shows spread or shows mks or M lger range
(b)			1 mk about overall standard, based on median or on F's IQR being "higher" 1 mk about spread (or range or IQR) or about skewness. must be overall, not indiv mks must be comparison, not just figures Examples: not F higher mean not M have hiest and lowest mks condone F +ve skew
	F generally higher or median higher F higher on average or F better mks F IQR is above M IQR	B1	
	F more compact M wide(r) range or gter IQR or gter variation or gter variance or more spread or less consistent M evenly spread or F skewed	B1 2	
(c)	<u>Advantage</u>: B&W shows med or Qs or IQR or range or hiest & lowest or key values	B1	not B&W shows skewness not B&W shows info at a glance not B&W easier to compare data sets not B&W shows mean not B&W shows spread not B&W easier to calculate or easier to read
	<u>Disadvantage</u>: B&W loses info' B&W shows less info' B&W not show freqs B&W not show mode B&W: outlier can give false impression hist shows more info hist shows freqs or fds hist shows modal class (allow mode) hist shows distribution better can calc mean from hist	B1 2	not B&W does not give indiv (or raw) data not B&W does not show mean not hist shows freq for each mark not hist shows all the results not hist shows total allow adv of hist as disadv of B&W
(iii)	102 x 51 + 26 x 59 ÷ 128 = 52.6 (3 sfs)	M1 M1dep A1 3	or 5202 + 1534 or 6736
Total		**10**	

Q6, (OCE 4732, Jan 2009, Q5i-iii)

(i)	68 75 – 59 = 16	B1 M1 A1	3	 attempt 6th & 18th or 58-60, 74-76 & subtr must be from 75 – 59
(ii)	Unaffected by outliers or extremes (allow less affected by outliers) sd can be skewed by one value	B1	1	NOT: … by anomalies or freaks easier to calculate
(iii)	Shows each data item, retains orig data can see how many data items can find (or easier to read) mode or modal class can find (or easier to read) frequs can find mean Harder to read med (or Qs or IQR) Doesn't show med (or Qs or IQR) B&W shows med (or Qs or IQR) B&W easier to compare meds	 B1 B1	2	NOT: shows freqs shows results more clearly B&W does not show freqs NOT: B&W easier to compare B&W shows spread or variance or skew B&W shows highest & lowest Assume in order: Adv, Disadv, unless told Allow disadv of B&W for adv of S&L & vice versa Ignore extras

Q7, (OCR 4732, Jun 2014, Q1)

(i)		Median = 7.45 (m) IQR = 7.75 – 6.7 = 1.05 (m) allow 1.175 or 1.18 NOT 1.3	B1 M1 A1 **[3]**	cao allow 7.775 – 6.6 or 77.5 – 67 or 77.75 – 66 or 7.8 – 6.5 even though this is an incorrect method or 78 – 65 allow 10.5 or 11.75 or 11.8 but only if med = 74.5	These pairs of values only, and subtract, for M1 eg 7.45, 7.75 – 6.7 = 1.05 B1M1A1 7.45, 7.775 – 6.6 = 1.175 B1M1A1 7.45, 7.8 – 6.5 = 1.3 B1M1A0 7.45, 7.7 – 6.5 = 1.2 B1M0A0 7.45, 77.5 – 67 = 10.5 B1M1A0 74.5, 77.5 – 67 = 10.5 B0M1A1 74.5, 7.75 – 6.7 = 10.5 B0M1A1 74.5, 77.75 – 66 = 11.75 B0M1A1 7.45, 78 – 65 = 13 B1M1A0 74.5, 78 – 65 = 13 B0M1A0 74.5, 77 – 65 = 12 B0M0A0
(ii)		4 2 2 | 5 5 3 3 0 | 6 8 7 7 | 6 4 3 2 | 7 6 5 | 7 8 5 | 8 Complete correct diag including order and	 B1* B1dep	 correct digits in correct leaves, ignore order, allow one omitted or extra or misplaced or incorrect digit key: eg 8|6|4 means 6.8 (*B*) and 6.4 (*A*)	Allow a separate diag with leaves to left of stem. If only a separate diag is drawn, with leaves to right of stem: all correct including order, alignment and key: B1 If all digits are in correct rows and orders,
		key and alignment **[2]**		allow just 8 | 6 means 6.8 NOT 8 | 6 means 8.6 Allow 8 | 6 means 68, if consistent with (i)	& correct key, award this mark unless EITHER: 1. eg a 2nd digit in one row is clearly aligned with a 3rd digit in another OR 2. 1st, 3rd, 4th & 5th rows are very different lengths, eg because of crossing out and replacement
(iii)		One correct comment on size: B1. One correct comment on spread or shape: B1. The following are examples only. Ignore any working; mark the statements only. Allow "First set" or "Right" for A, "Second set" or "Left" for B.			

Q8, (OCR 4732, Jan 2011, Q1)

i	38 61	B1 B1 2	Reversed: B1B0	
ii	Paper 2 Higher median or curve is to right	B1 B1dep 2	Indep of reason or similar Higher average or mean or midpoint Paper 2: half $\leq$ 61, cf paper 1: half $\leq$ 38 Paper 1: more students scored lower marks (or lower than eg 40)	Ans "Paper 1", ignore reason: B0B0 unless reversed in (i) More scored higher mks Highest & lowest mks are higher For each cf, the corresponding mark is higher in p2. None get 0-10 Some get 100 Eg 25 scored > 69 in p1, cf 65 scored > 69 in p2 NOT Marks are higher NOT marks seem higher NOT everyone gets higher mks NOT Curve steeper Ignore irrelevant or incorrect SC: If reversed in (i): (ii) p1 because median higher B1B1ft
iii	55, 25 73, 46 Paper 1 IQR = 30 Paper 2 IQR = 27 Suggestion correct or p2 less varied	M1 A1 A1 B1f indep 4	M1 one pair of quartiles p2 more consistent or less spread out Allow "p2 has smaller range (or smaller variance") if IQRs found "It" is less varied: assume p2: B1	Allow 55±1, 25±1 Not necessarily subtracted 73±1, 46±1 30±1 27±1 p1 more varied or more spread out or less consistent Little difference or similarly varied NOT p2 IQR smaller than p1 unless also says less varied oe If quartiles found but not IQRs: max M1A0A0B1 If no quartiles calculated can still score B1 Steeper curve alone M0A0A0B0 If IQRs wrong, with p1 < p2, ft "suggestion wrong": B1f Ignore irrelevant or incorrect

iv	37 (± 3)	B2 2	B1 for 163 (± 3)	Not necessarily integer. B1 for 78-80 mks for min grade A on p2 SC: ans 105 – 110: B1 (from p1 10 mks hier instead of lower)
v	37.5 28.2	B1 B1 2	cao or sd the same	NOT eg 37.51 Ignore all working
Total		**12**		

Q9, (OCR 4732, Jan 2012, Q5)

(i)	(a)	$(\frac{6}{3} =) 2$	B1 [1]	$(\frac{6}{9} \times 3 =) 2$	
(i)	(b)	$^2/_6 \times 2$ $= {}^2/_3$ oe or 0.667 or 0.67 or 0.7	M1 A1[2]	Allow $^2/_5 \times 2$ or ans 0.8 for M1	Can be implied, eg $\frac{1}{3} = 0.3$, ans 0.6: M1A0 Allow 0.66 or 0.666
(ii)		(3.5, 6) (0.5, 0) or (6.5, 15)	B1 B1 [2]	Ignore incorrect	(6, 3.5) AND (15, 6.5): B1
(iii)	(a)	$\frac{\Sigma xf}{21}$ = 5.43 (3 sf) or $\frac{114}{21}$ or $\frac{38}{7}$ oe $\frac{\Sigma x^2 f}{21}$ or $\frac{817.5}{21}$ or 38.9… – "5.43"^{2} or = 9.46 or 9.4592…. ($\sqrt{}$9.4592….) = 3.08 (3 sfs)	M1 A1 M1 M1 A1 [5]	Allow x within classes, incl end pts then ÷5: M0A0 Allow x within class, incl end pt ÷5: M0 dep +ve result; done before $\sqrt{}$; not $-(\bar{x}^2 \div \ldots)$	≥ 2 non-zero terms correct ft their x ≥ 2 non-zero terms correct ft their x Calc 4 values of $(x - \bar{x})^2$ or $(x - \bar{x})^2 f$ or (11.8, 0.184, 6.61, 50) or (70.5, 1.65, 26.4, 100) or 199 M1 $\frac{\Sigma(x-\bar{x})^2 f}{21}$ fully correct method M1
(iii)	(b)	Actual values or exact hours unknown oe Don't have raw data. oe or measured to nearest hour oe	B1 [1]	or Data given in classes or grouped oe or Data evenly distributed in classes oe	Mid-points or medians or averages of class boundaries used oe