Statistics 1 Correlation and Regression Answers

1(a)	Gradient, $b = 0.886$ to 0.887 $\quad\quad\quad b = 0.88$ to 0.89	B2 (B1)		AWFW AWFW
	Intercept, $a = 2.31$ to 2.33 $\quad\quad\quad a = 2.3$	B2 (B1)		AWFW AWRT
	Attempt at $\Sigma x \ \ \Sigma x^2 \ \ \Sigma y \ \ \Sigma xy$ **or** Attempt at $S_{xx} \ \ S_{xy}$ Attempt at a **correct** formula for b $\quad\quad\quad b = 0.886$ to 0.887 $\quad\quad\quad a = 2.31$ to 2.33	(M1) (m1) (A1) (A1)		72, 624, 87, 720 105.6, 93.6 AWFW AWFW
	Accept a & b interchanged only if $y = ax + b$ stated or subsequently used correctly in either (b) or (c)		4	
(b)	a: average **waiting time** of 2.32 minutes (139 seconds) when entering **empty** **restaurant**	B1		OE; accept minimum waiting time
	b: average **increase in waiting time** of 0.886 minutes (53 seconds) **for each** **customer in restaurant** on entry	B1	2	OE
(c)	Use of $y = a + 5b$ or $y = a + 25b$	M1		
(i)	For $x = 5 \quad\quad y = 6.6$ to 6.8			
(ii)	For $x = 25 \quad\quad y = 24.3$ to 24.6	A1	2	Both; AWFW
(d)(i)	**Reliable as interpolation** and **small** **residuals** **or** Reliable as interpolation but large percentage residuals so inconclusive **or** Large percentage residuals so unreliable	B1 B1 (B1) (B1) (B1)		Within range OE OE
(ii)	Unreliable as extrapolation	B1	3	Outside range OE
	Total		11	

5(a)	Scatter Diagram	B2		4 labelled points plotted	
	or	(B1)		3 labelled points plotted	
	or	(B1)	2	4 unlabelled points plotted	
(b)(i)	Positive/linear correlation/relationship	B1		OE	
	except for				
	two unusual values/results	B1	2	OE	
(ii)	0.462	B1	1	CAO; accept 3rd/final/last value	
(c)	C and D	B1		CAO	
	C is likely **freestyle** champion	B1		Style identified	
	D is likely **backstroke** champion				
	or				
	C is likely freestyle champion	(B1)			
	D is likely backstroke champion	(B1)	2		

(d)(i)	$r = 0.912$ to 0.913	B3		AWFW	
	or				
	$r = 0.91$ to 0.92 or 0.46 to 0.47	B2		AWFW	
	or				
	$r = 0.9$	B1		AWRT	
	Attempt at $\quad \Sigma x \quad \Sigma x^2$			270.4, 9188.46	
	$\Sigma y \quad \Sigma y^2$			301.6, 11437.84	
	Σxy			10246.53	
	or				
	Attempt at $\quad S_{xx} \quad S_{yy} \quad S_{xy}$	(M1)		48.94, 67.52, 52.45	
	Attempt at a **correct** formula for r	(m1)			
	$r = 0.912$ to 0.913	A1	3	AWFW	
(ii)	Boys are **faster/slower** at **both strokes**			OE; do not accept freestyle times	
	or	B1	1	are proportional to backstroke times	
	Boys are **equally good** at **both strokes**				
	Total		11		

Swimming Times

Scatter diagram plotting Backstroke (y seconds) against Freestyle (x seconds) with points:
- C (≈25, 37.4) — circle
- A (≈30, 33.5) — circle
- E (≈31, 34.6) — cross
- B (≈32.5, 35.4) — circle
- D (≈31.8, 27.2) — circle
- G (≈32.5, 37.6) — cross
- J (≈34, 38.4) — cross
- F (≈35.5, 38.1) — cross
- I (≈36, 42.3) — cross
- H (≈38, 41.4) — cross
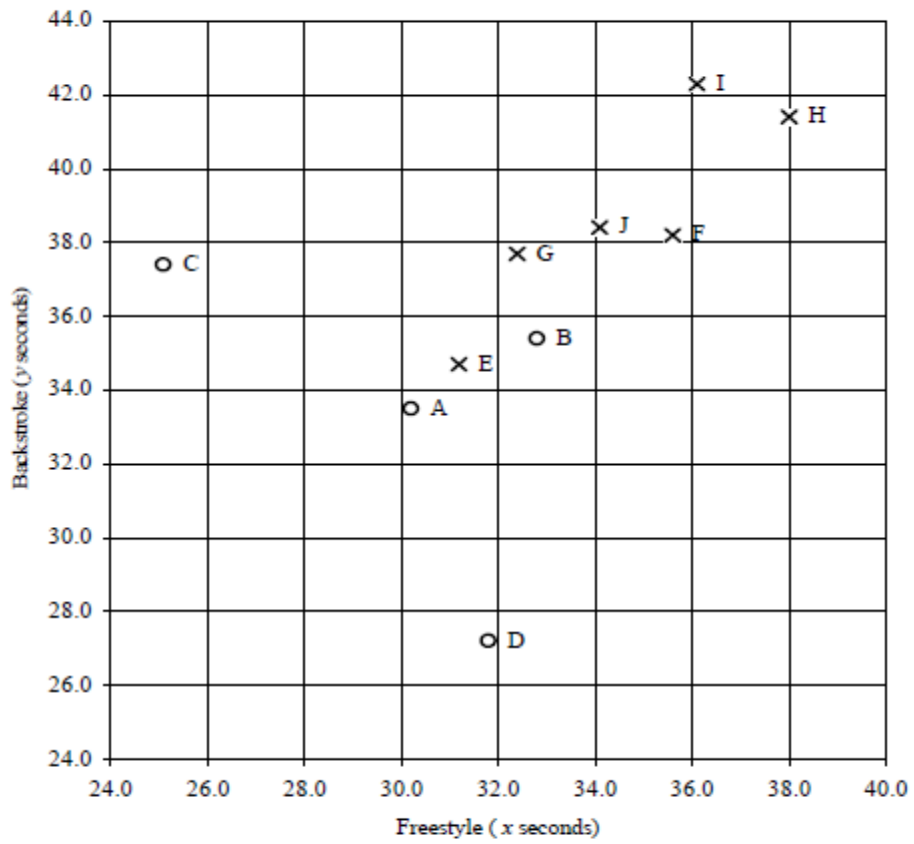
(a) Scatter Diagram

4 labelled points plotted	B2
3 labelled points plotted	(B1)
4 unlabelled points plotted	(B1)

Graph = 2

1(a)(i)	$r = 0.143$ to 0.1432	B3		AWFW	
	or				
	$r = 0.142$ to 0.144	B2		AWFW	
	or				
	$r = 0.1$ to 0.2	B1		AWRT	
	Attempt at $\quad \Sigma x \; \Sigma x^2$			3952, 2228282	
	$\Sigma y \; \Sigma y^2$			47.00, 292.0000	
	Σxy			23517.50	
	or				
	Attempt at $\quad S_{xx} \; S_{yy} \; S_{xy}$	M1		275994, 15.875, 299.5	
	Attempt at a correct formula for r	m1			
	$r = 0.143$ to 0.1432	A1	3	AWFW	
(ii)	Little/weak/no correlation/relationship/association between number of pages and (retail)	B1		or equivalent; but not poor	
	price	B1	2	context	
(iii)	Size (page, thickness), author, ranking, publicity/marketing, cover design, recommendations on back, publisher, font, popularity, quality, print-run, etc	B1	1	or any sensible variable but not pictures, coloured pictures, age, words, weight, mass	
(b)	(Very) strong/almost exact positive/perfect correlation/relationship/association between	B1		or equivalent	
	number of pages and sale/new price	B1	2	context	
	Sale price appears to be determined by number of pages	B2		or equivalent	
	Total		8		

3(a)(i)	Gradient, $b = -3.24$ to -3.26	B2		AWFW $\quad (-3.25)$	
	$b = -3.2$ to -3.3	B1		AWFW	
	Intercept, $a = 262$ to 264	B2		AWFW $\quad (262.8\dot{8})$	
	$a = 260$ to 270	B1		AWFW	
	Attempt at $\Sigma x \; \Sigma x^2 \; \Sigma y \; \Sigma xy$			108, 1836, 2015, 22425	
	or				
	Attempt at $S_{xx} \; S_{xy}$	M1		540, -1755	
	Attempt at a correct formula for b	m1			
	$b = -3.24$ to -3.26	A1		AWRT	
	$a = 262$ to 264	A1	4	AWFW	
	Accept a & b interchanged only if identified correctly in (b) and (c)				
(ii)	Gradient, b:				
	Decrease in pressure per month	B2		or equivalent	
	Change in pressure	B1	2	or better	
(iii)	Intercept, a:				
	Initial pressure or pressure at $x = 0$	B1		or equivalent; not y-intercept	
	Reference to 265, actual or expected value	B1	2		

(b)(i)	Value for $b = 2 \times$ [gradient or b from (a)(i)]	M1		accept $2b$; ignore sign	
	$= -6.4$ to -6.6	A1√	2	AWFW (-6.5) √ from (a)(i) but must be < 0	
(ii)	$P_8 = 265 - 6.5 \times 8$	M1		must use 265 and $x = 8$ and $2 \times [b\ (< 0)$ from (a)(i)] or [from (b)(i) (< 0)]	
	$= 212$ to 214 (< 220)	A1	2	AWFW AG	
	Total		**12**		

3(a)	$0.5 \leq$ Value ≤ 0.95 Positive value < 1 (and > 0)	B2 (B1)		Value is actually 0.8
(b)	$-0.2 \leq$ Value $\leq +0.2$	B1		Value is actually 0.0
(c)	$-0.95 \leq$ Value ≤ -0.5 Negative value > -1 (and < 0)	B2 (B1)	5	Value is actually -0.7
	Total		**5**	

7(a)	8 or 7 points plotted accurately (6 or 5 points plotted accurately)	B2 (B1)	2	
(b)	Gradient, $b = 0.114$ to 0.115 $\quad\quad (b = 0.11$ to $0.12)$	B2 (B1)		AWFW (0.11469)
	Intercept, $a = 15.9$ to 16.1 $\quad\quad (a = 13$ to $19)$	B2 (B1)		AWFW (16.00824)
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$ or	(M1)		4420, 3230800, 635 and 441300
	Attempt at S_{xx} and S_{xy}			788750 and 90462.5
	Attempt at correct formula for b	(m1)		
	$b = 0.114$ to 0.115	(A1)		AWFW
	$a = 15.9$ to 16.1	(A1)		AWFW
	Accept a and b interchanged only if then identified correctly later in question			
	Line plotted accurately (Evidence of correct method for ≥ 2 points)	B2 (M1)	6	At least from $x = 200$ to 1000

(c)	$\text{Res}_H = y_H - Y_H = 70 - (a + b \times 480)$	M1		Used; or implied by **correct** answer; allow for $Y_H - y_H$ **shown**
	$= -1.5$ to -0.5	A1		AWFW (-1.06)
	Point H is (almost) **on** / **just below** the line	B1	3	Accept near / close / just above or equivalent
(d)	$Y = a + b \times 560$ or reading from scatter diagram	M1		Used
	$= 79$ to 81	A1		AWFW (80.2)
	$\text{Cost} = Y \times \dfrac{12}{60}$ or $\dfrac{Y}{5}$	M1		Used
	$= £15.8$ to $£16.2$	A1	4	AWFW; ignore units (£16.05)
	Total		15	
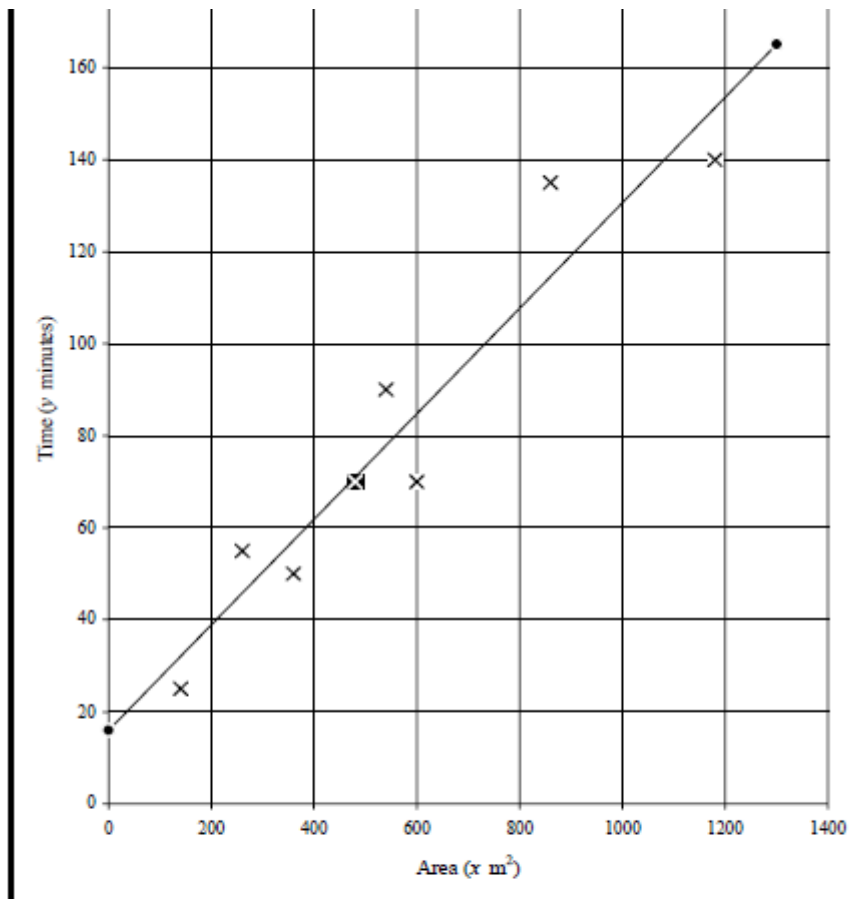
(a) 8 or 7 points plotted accurately   B2
     (6 or 5 points plotted accurately   B1)

(b) Line plotted accurately   B2
     (Evidence of correct method for ≥ 2 points   M1)

                                        (Graph = 4)

1(a)	$r = -0.526$ to -0.525 or $r = -0.53$ to -0.52 or $r = -0.6$ to -0.4	B3 (B2) (B1)		AWFW AWFW; ignore sign AWFW; ignore sign
	OR Attempt at $\sum x,\ \sum x^2,\ \sum y,\ \sum y^2\ \text{and}\ \sum xy$ or Attempt at $S_{xx},\ S_{yy}\ \text{and}\ S_{xy}$ Attempt at a correct formula for r $r = -0.526$ to -0.525	 (M1) (m1) (A1)	 3	 260, 6970, 143, 2083 and 3671 210, 38.1 and -47 AWFW
(b)	Weak/some/moderate negative correlation (relationship/association) between length and (maximum) diameter Ignore subsequent comments (as below) only if B1 B1 already scored **OR** Some evidence that large lengths are associated with small diameters **OR** Longer melons tend to have smaller diameters / be thinner	 B1 B1 (B1) (B1) (B1) (B1)	 2	OE; must qualify strength and indicate negative B0 for strong/poor/reasonable/average B0 if $r > 0$ or $r < -1$ B0 if contradictory statements Context OE; must qualify strength and indicate negative OE; must qualify strength and indicate negative
	Total		5	

5(a)	Time taken depends upon temperature	B1	1	OE; **not** x set values
(b)	b (gradient) $= -0.0873$ to -0.087 b (gradient) $= -0.09$ to -0.08	B2 (B1)		AWFW $\quad\quad\quad\quad\quad (-0.08\dot{7}2\dot{7})$ AWFW; $-8.73^{-02} \Rightarrow$ B0
	a (intercept) $= 5.94$ to 5.96 a (intercept) $= 5.6$ to 6.1	B2 (B1)		AWFW $\quad\quad\quad\quad\quad (5.95\dot{0}\dot{9})$ AWFW
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$ or Attempt at S_{xx} and S_{xy} Attempt at correct formula for b $b = -0.0873$ to -0.087 $a = 5.94$ to 5.96	(M1) (m1) (A1) (A1)	4	$396, 16016, 30.9$ and 958.8 1760 and -153.6 AWFW AWFW
	Accept a and b interchanged only if then identified correctly later in question			
(c)(i)	Each 1 °C rise in temperature results in an (average) decrease of 0.087 m (5 s) in time taken for pellets to dissolve	B1 B1	2	Quantified rise in x (results in) Decrease in y OE
(ii)	a is y–value at $x = 0$ at which water is solid/ice/frozen so pellets cannot dissolve	B1 B1	2	Indication that it is y at $x = 0$ Mention of solid or ice or frozen
(ii)	a is y–value at $x = 0$ at which water is solid/ice/frozen so pellets cannot dissolve	B1 B1	2	Indication that it is y at $x = 0$ Mention of solid or ice or frozen
(d)(i)	When $x = 30$ $y = 3.3$ to 3.4 $y = 2.9$ to 3.7	 B2 (B1)		 AWFW $\quad\quad\quad\quad\quad (3.33\dot{2}\dot{7})$ AWFW
	If B0, use of their equation with $x = 30$	(M1)	2	
(ii)	When $x = 75$ $y < 0$ or negative which is impossible	 B1 ↑Dep↑ B1	 2	 OE OE; **not** extrapolation
	Total		13	