Statistics 6
Histograms & cumulative frequency SOLUTIONS

Section 1

1, a,

1	4 7
2	5 8
3	1 1 3 4 5 6
4	0 4 9

$(n=13)$

$Q_1 = \frac{13}{4} = 3.25 \rightarrow 4^{th} = 28$

$Q_2 = \frac{13}{2} = 6.5 \rightarrow 7^{th} = 33$

$Q_3 = \frac{3(13)}{4} = 9.75 \rightarrow 10^{th} = 36$

(Also allow $Q_1 = 26.5$, $Q_2 = 33$, $Q_3 = 38$)

b, $Q_1 - 1.5(IQR) = 28 - 1.5(36-28)$

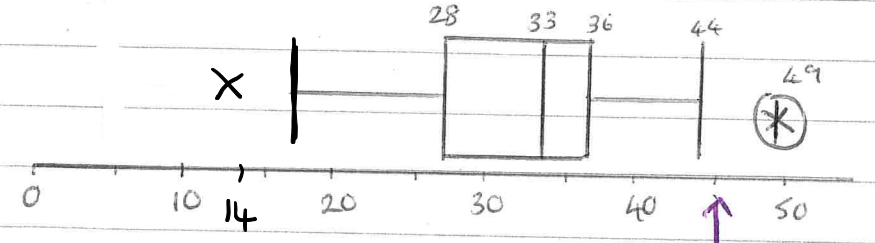
$= 16$ ∴ 14 is an outlier

$Q_3 + 1.5(IQR) = 36 + 1.5(36-28)$

$= 48$ ∴ 49 is an outlier

(Also allow boundaries ⇒ $26.5 - 1.5(38-26.5) = 9.25$

⇒ $38 + 1.5(38-26.5) = 55.25$



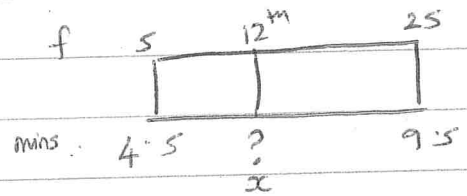
allow top of line to go to 48 (outlier boundary)

(Also allow with values from different quartiles as given before.)

2,

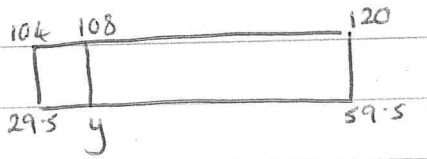
mins	f	cf
0.5 - 4.5	5	5
4.5 - 9.5	20	25
9.5 - 19.5	28	53
19.5 - 29.5	51	104
29.5 - 59.5	16	120

10% of 120 = 12th → 12th person lies in the group 4.5 - 9.5



$$\frac{x-4.5}{9.5-4.5} = \frac{12-5}{25-5} \qquad x = 6.25.$$

90% of 120 = 108th → 108th person lies in group 29.5 - 59.5



$$\frac{y-29.5}{59.5-29.5} = \frac{108-104}{120-104} \qquad y = 37.$$

So Inter-percentile range = 37 - 6.25 = 30.75

3, $\bar{x} = £10$ $\quad$ $s.d_x = £1$

Scheme A : + 50p per hour $\quad$ Scheme B : 5% increase $= \times 1.05$

$\bar{y} = £10.50$ $\quad$ $\bar{z} = £10.50$

$s.d_y = £1$ $\quad$ $s.d_z = £1.05$

4, $y = 2x + 11$ $\quad$ $x^2 + y^2 - 6x - 4y = 32$

$$x^2 + (2x+11)^2 - 6x - 4(2x+11) = 32$$
$$x^2 + 4x^2 + 44x + 121 - 6x - 8x - 44 = 32$$
$$5x^2 + 30x + 45 = 0$$
$$x^2 + 6x + 9 = 0$$
$$(x+3)^2 = 0$$
$$x = -3$$

(Since there is only 1 solution for x the line only touches the circle once)

Can also use discriminant

$b^2 - 4ac^2$

$= 6^2 - 4(1)(9)^2$

$= 36 - 36$

$= 0$ $\quad \therefore$ one solution for x.

Section 2

1,

	freq	freq density
0·5 - 3·5	30	$\frac{30}{3}$ = 10
3·5 - 4·5	96	$\frac{96}{1}$ = 96
4·5 - 5·5	48	$\frac{48}{1}$ = 48
5·5 - 7·5	84	$\frac{84}{2}$ = 42
7·5 - 10·5	27	$\frac{27}{3}$ = 9
10·5 - 15·5	15	$\frac{15}{5}$ = 3

Method ✓ ✓ ✓ (3)

$f.d = \frac{\text{freq}}{\text{class width}}$

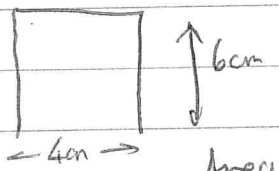
2, a, 21 - 25 $\Rightarrow$ $20.5 \leq h < 25.5$ $\Rightarrow$ Midpoint = 23 ✓

31 - 40 $\Rightarrow$ $30.5 \leq h < 40.5$ $\Rightarrow$ Midpoint = 35·5 ✓ (2)

b, 1 - 10 $\Rightarrow$ $0.5 \leq h < 10.5$ ✓ numbers seen. (1)

c, 11 - 20 $\Rightarrow$ 10·5 - 20·5

freq = 15.



Area = $24cm^2$ = 15 people

$1.6cm^2$ = 1 person ✓

26 - 30 $\Rightarrow$ 25·5 - 30·5 $\Rightarrow$ ∴ width = 2cm ✓

freq = 13

13 people = $13 \times 1.6 = 20.8cm^2$ Some relationship

∴ Height = 10·4cm ✓ (3)

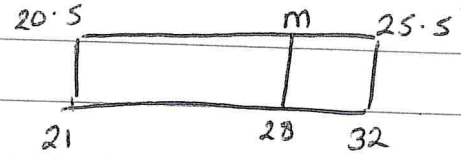
d, mean = 23·5 ✓ (from calculator)

s.d = 10·7 ✓✓

(5)

e, Median $= \frac{56}{2} = 28 \Rightarrow 28.5^{th}$ value ✓

28.5^{th} value is in the "21-25" group $\Rightarrow$ $20.5 \leq h < 25.5$ ✓



$$\frac{m - 20.5}{25.5 - 20.5} = \frac{28 - 21}{32 - 21}$$ (3)

$m = 23.681$ ✓

$Q_2 = 23.7$

3, a, (9·5, 0) ✓ (19·5, 25) ✓ Allow (19·5, 25) and (24·5, 47) (2)

b, Because the data is grouped, we do not know the smallest length pebble and/or the largest length pebble. ✓ (1)

4, a, 1200 — $Q_1 = 300^{th} \approx 46$ ✓

$Q_3 = 900^{th} \approx 70$ ✓

$\therefore$ IQR $= 70 - 46 = 24$ ✓ (3)

b, 40% of 1200 $= 480$ ✓ $1200 - 480 = 720$ ✓ $\rightarrow$ 64% ✓

c, 840 ✓ $1200 - 840 = 360$ ✓ (2)

allow $1200 - 860 = 340$

d, $P(\text{all 5 score} > 68) = \left(\frac{360}{1200}\right)^5 = \frac{243}{100,000} = 0.00243$ ✓ (3)

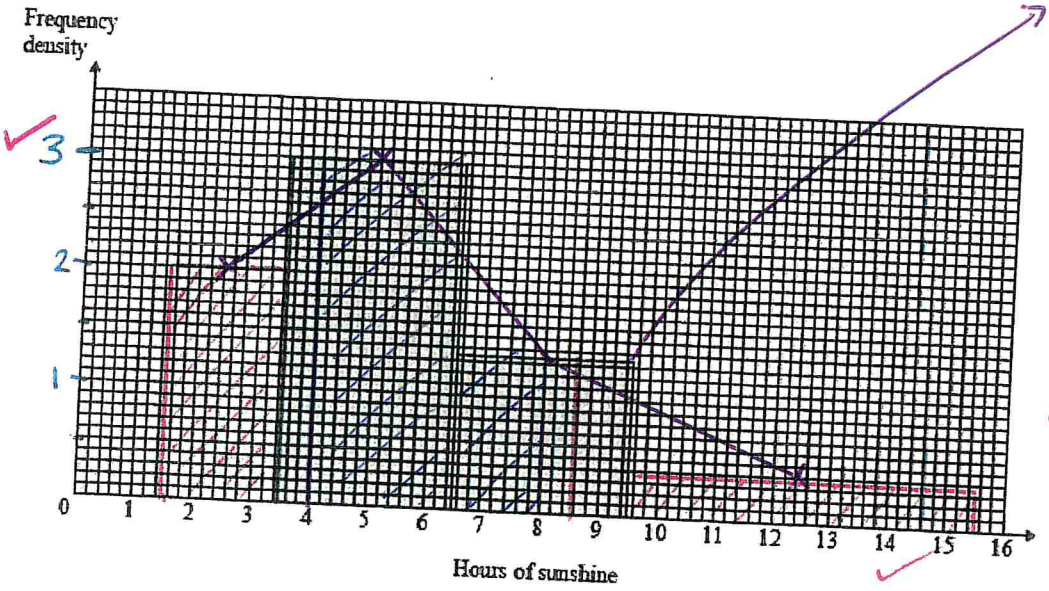
Allow $\left(\frac{340}{1200}\right)^5 = 0.00183$

Q5, a, The data is grouped, the data is continuous ✓ either (1)

Hours of sunshine	0	1 – 3	4 – 6	7 – 9	10 – 15
Number of days	0	6	9	4	2
f.d	0	2	3 ✓	1·3	0·3

0·5 - 3·5 3·5 - 6·5 6·5 - 9·5 9·5 - 15·5

b,



e, Polygon
(2,2)
(5,3) Plotted and joined (1)
(8,1.3) (1)
(12.5,0.3)

5
• Bars from 0.5-3.5 etc.
• Correct values of axis
• Correct heights

c, 4 - 8 hours ⇒ 4 - 6.5 hours = 2.5 × 3 = 7.5 (1)

6.5 - 8 hours = 1.5 × 1.3 = 2 (1)

⇒ 9.5 days (1)

d, $\Sigma x = 114$ $\Sigma x^2 = 817.5$ $n = 21$

$\bar{x} = 5.4285$

$= 5.43$ hours. ✓

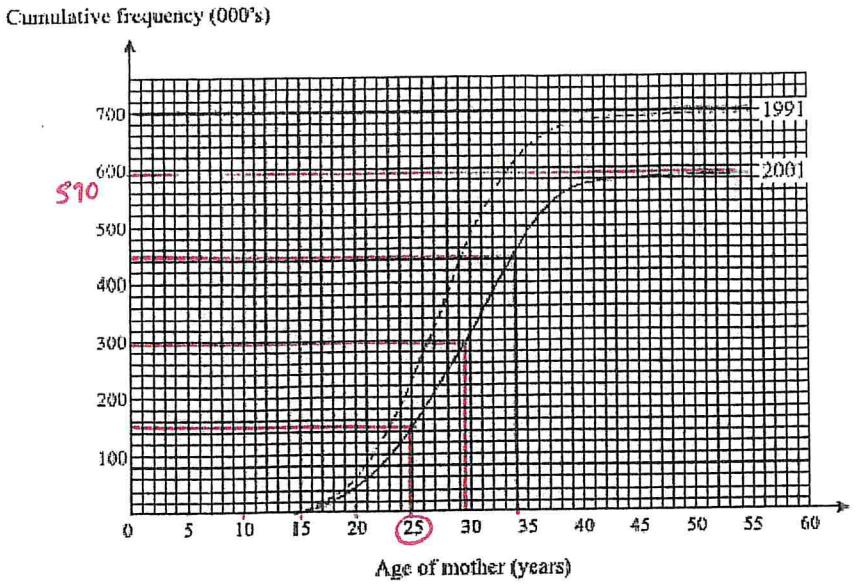
$$s.d = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{817.5}{21} - \left(\frac{114}{21}\right)^2}$$ (2)

$= 3.07557$

$= 3.08$ or 3.1 ✓

e, See above.

6, cont.



Year	Median age (years)	Interquartile range (years)	Proportion of mothers giving birth aged below 25	Proportion of mothers giving birth aged 35 or above
1991	27.5	7.3	33%	9%
2001	29.5 ✓1	9 ✓✓✓3	25% ✓1	18%

others ok ±.9 29.

6, a, 1991 ✓ → more births, by ≈ 110 ✓ (2)

b, i) $Q_2 = \frac{590}{2} = 295^{th} = 29.5$

$Q_1 = \frac{590}{4} = 147.5^{th} = 25$

$Q_3 = \frac{3}{4}(590) = 442.5^{th} = 34$

See table ~~below~~ above.

$\therefore$ IQR = 9

ii) The mothers in 2001 tended to be older than the women who gave birth in 1991. The median age in 2001 was greater than 1991. The proportion of 'younger' mothers (below 25) was lower in 2001, and the proportion of 'older' mothers (over 35) was higher in 2001. However there is a slightly larger spread of data in 2001. ✓

(2) Any 2 sensible comments!