

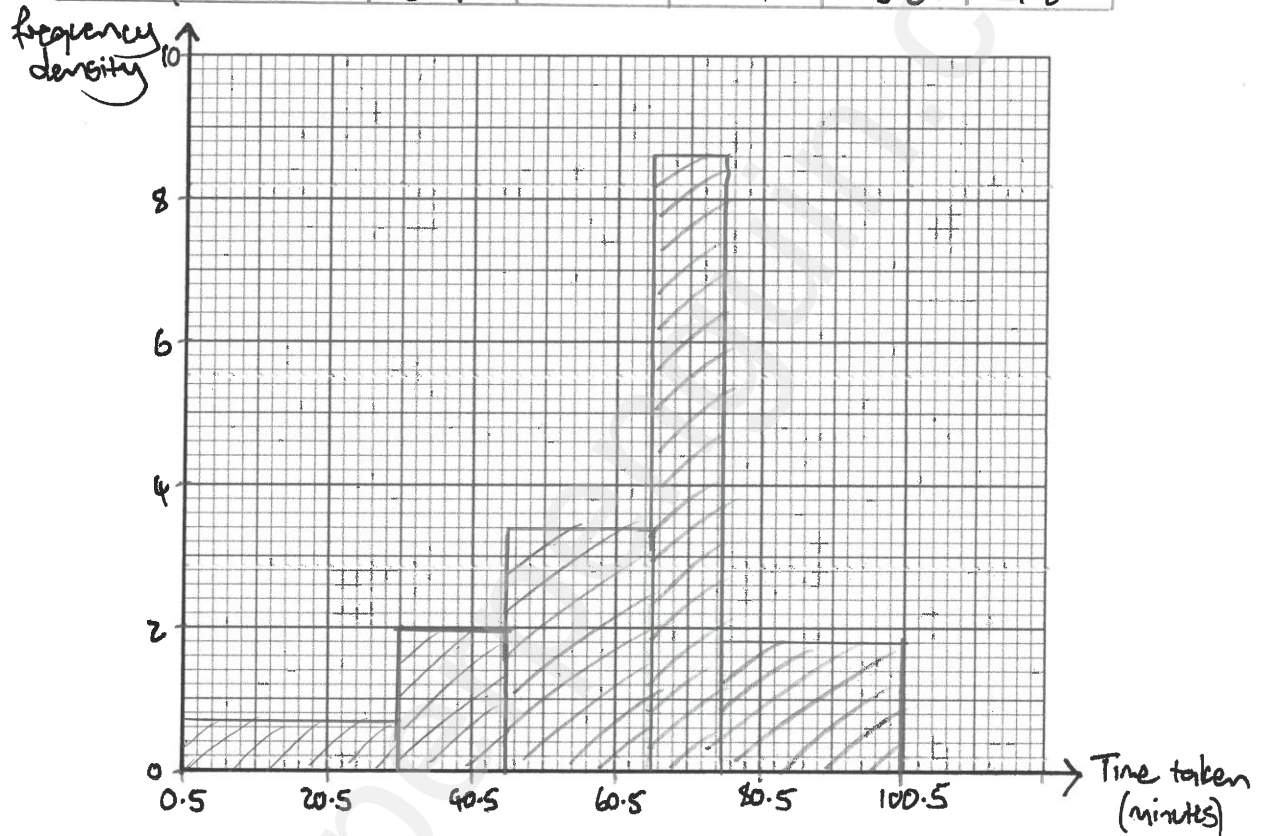
- 3 At a summer camp an arithmetic test is taken by 250 children. The times taken, to the nearest minute, to complete the test were recorded. The results are summarised in the table.

Time taken, in minutes	1 – 30	31 – 45	46 – 65	66 – 75	76 – 100
Frequency	21	30	68	86	45

- (a) Draw a histogram to represent this information.

[4]

Class width	30	15	20	10	25
f.d.	0.7	2	3.4	8.6	1.8



- (b) State which class interval contains the median.

[1]

Time	1–30	31–45	46–65	66–75	76–100
cumulative freq.	21	51	119	205	250

$$Q_2: \frac{250}{2} = 125$$

125 Q_2 Q_2 lies in 66–75

- (c) Given that an estimate of the mean time is 61.05 minutes, state what feature of the distribution accounts for the median and the mean being different.

[1]

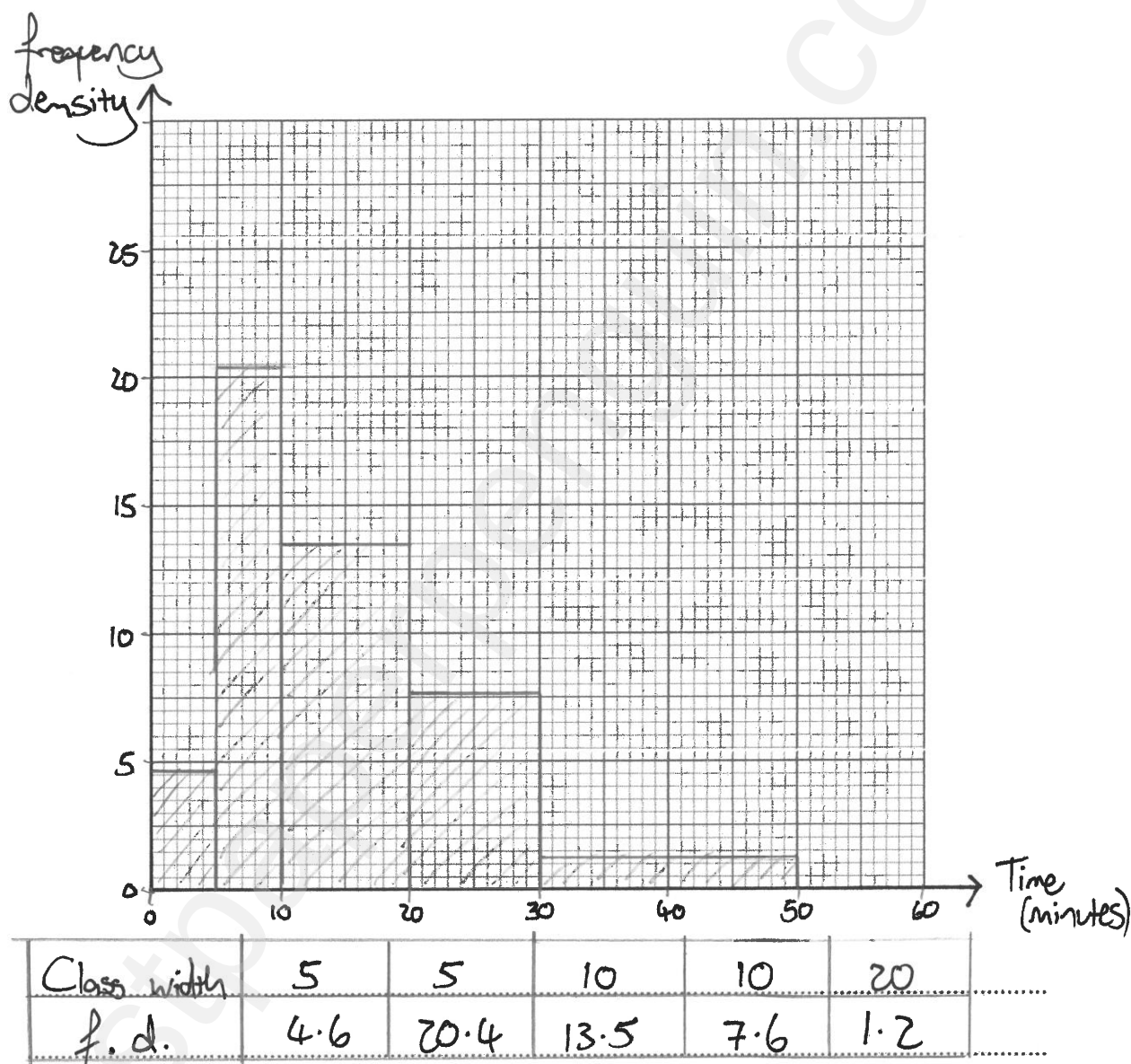
The distribution is negatively skewed
(median > mean)

- 3 The times taken, in minutes, by 360 employees at a large company to travel from home to work are summarised in the following table.

Time, t minutes	$0 \leq t < 5$	$5 \leq t < 10$	$10 \leq t < 20$	$20 \leq t < 30$	$30 \leq t < 50$
Frequency	23	102	135	76	24

- (a) Draw a histogram to represent this information.

[4]



- (b) Calculate an estimate of the mean time taken by an employee to travel to work.

[2]

Mid-point (t)	Frequency (f)	$f \times t$
2.5	23	57.5
7.5	102	765
15	135	2025
25	76	1900
40	24	960
	$\Sigma f = 360$	$\Sigma ft = 5707.5$

$$\bar{t} = \frac{5707.5}{360}$$

$$= \underline{\underline{15.9 \text{ mins}}}$$

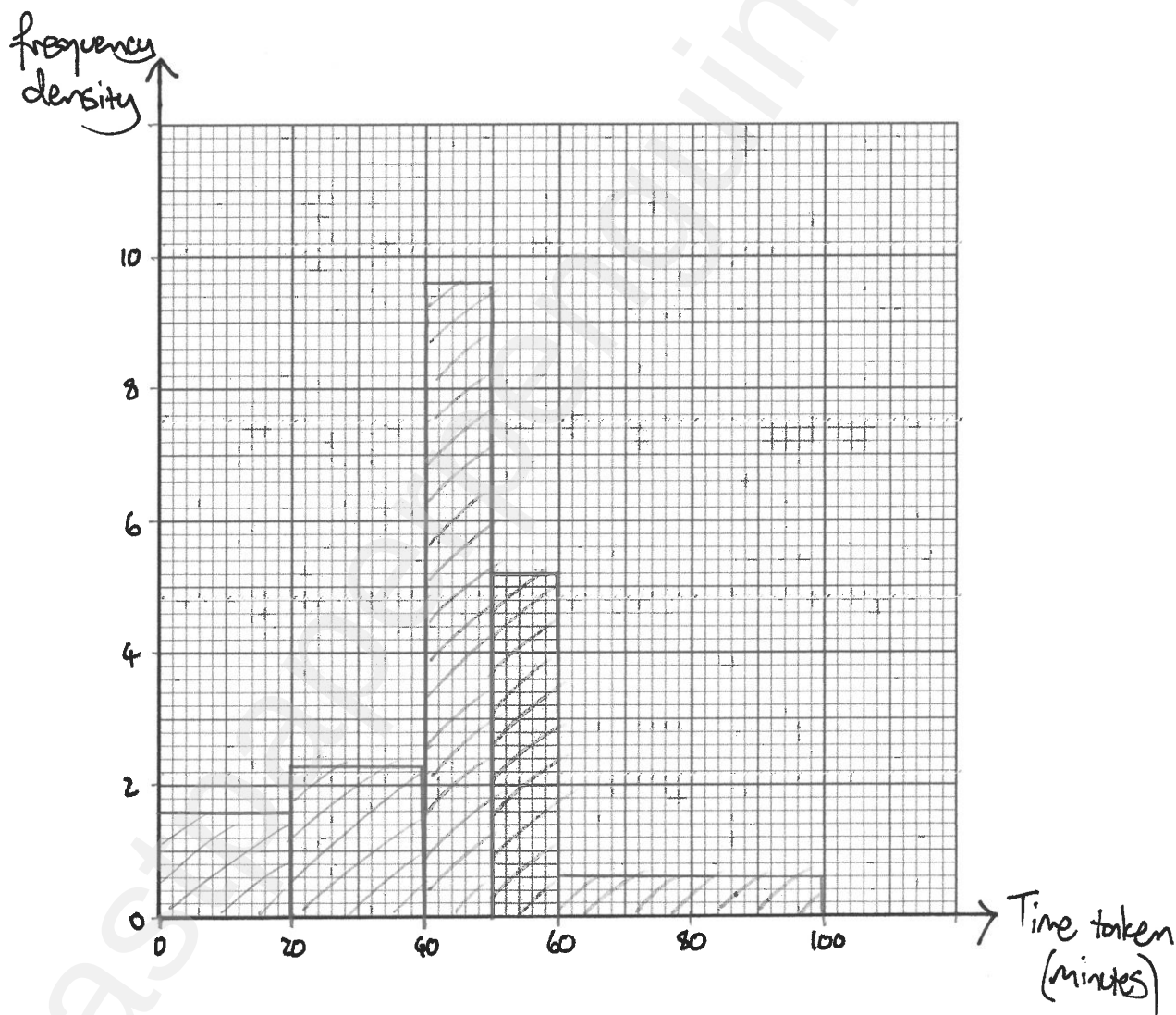
- 4 The times taken, in minutes, to complete a word processing task by 250 employees at a particular company are summarised in the table.

Time taken (t minutes)	$0 \leq t < 20$	$20 \leq t < 40$	$40 \leq t < 50$	$50 \leq t < 60$	$60 \leq t < 100$
Frequency	32	46	96	52	24

- (a) Draw a histogram to represent this information.

[4]

Class width	20	20	10	10	40
f.d.	1.6	2.3	9.6	5.2	0.6



From the data, the estimate of the mean time taken by these 250 employees is 43.2 minutes.

(b) Calculate an estimate for the standard deviation of these times.

[3]

Mid-points: 10, 30, 45, 55, 80

$$\text{Var} = \frac{10^2 \times 32 + 30^2 \times 46 + 45^2 \times 96 + 55^2 \times 52 + 80^2 \times 24}{250} - 43.2^2$$

$$= 333.36$$

$$\sigma = \sqrt{333.36}$$

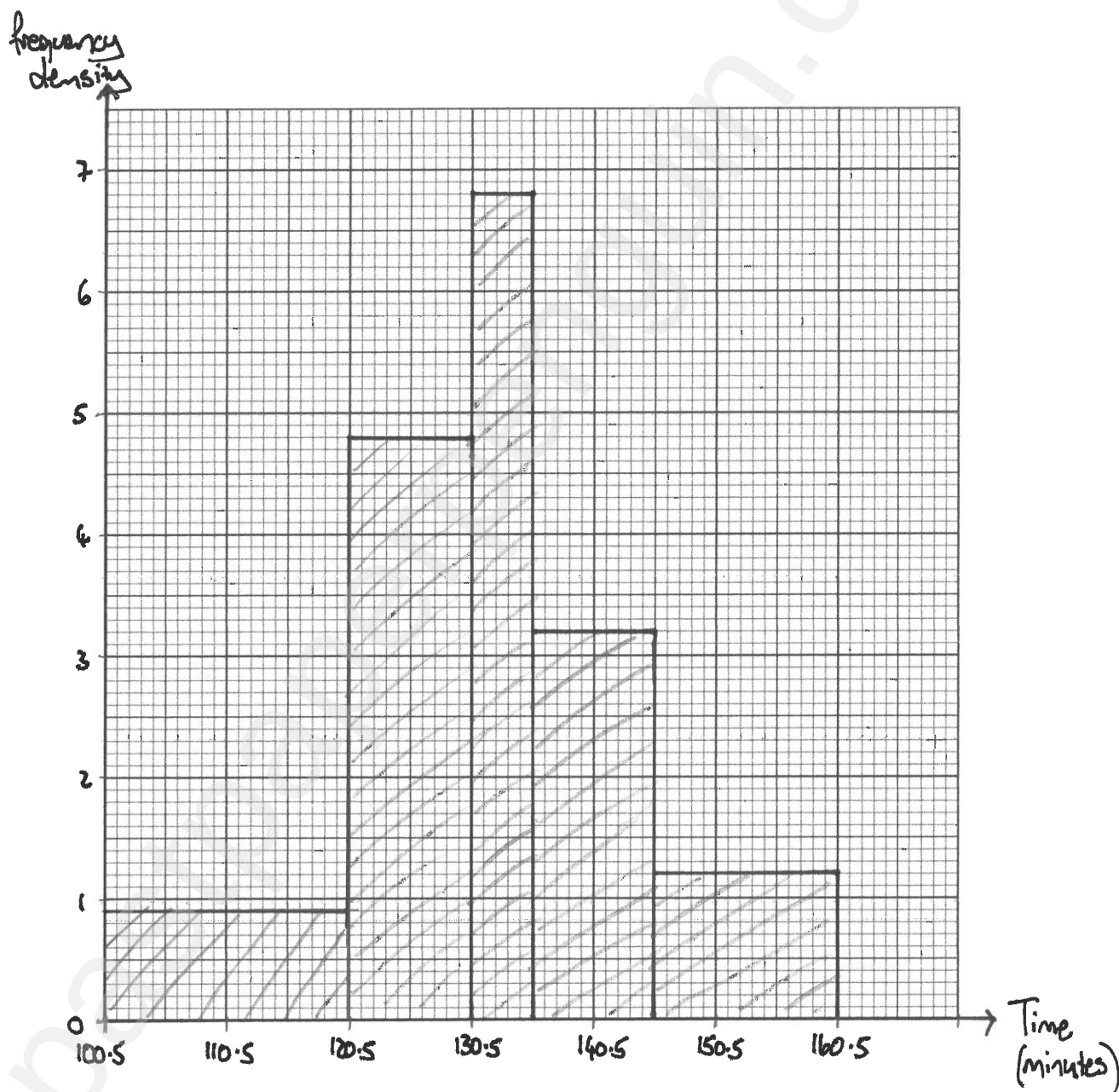
$$= \underline{\underline{18.3}}$$

- 4 The times, to the nearest minute, of 150 athletes taking part in a charity run are recorded. The results are summarised in the table.

	100.5	120.5	130.5	135.5	145.5	160.5
Time in minutes	101 – 120	121 – 130	131 – 135	136 – 145	146 – 160	
Frequency	18	48	34	32	18	
class width	20	10	5	10	15	
f.d.	0.9	4.8	6.8	3.2	1.2	

- (a) Draw a histogram to represent this information.

[4]



- (b) Calculate estimates for the mean and standard deviation of the times taken by the athletes. [5]

Mid-point (x)	Frequency (f)	$f \times x$
110.5	18	1989
125.5	48	6024
133	34	4522
140.5	32	4496
153	18	2754
	$\Sigma f = 150$	$\Sigma fx = 19785$

$$\begin{aligned}\bar{x} &= \frac{\Sigma fx}{\Sigma f} \\ &= \underline{131.9}\end{aligned}$$

$$\text{Var} = \frac{110.5^2 \times 18 + 125.5^2 \times 48 + 133^2 \times 34 + 140.5^2 \times 32 + 153^2 \times 18}{150} - 131.9^2$$

$$= 137.54$$

$$\sigma = \sqrt{\text{Var}}$$

$$= \underline{11.7}$$

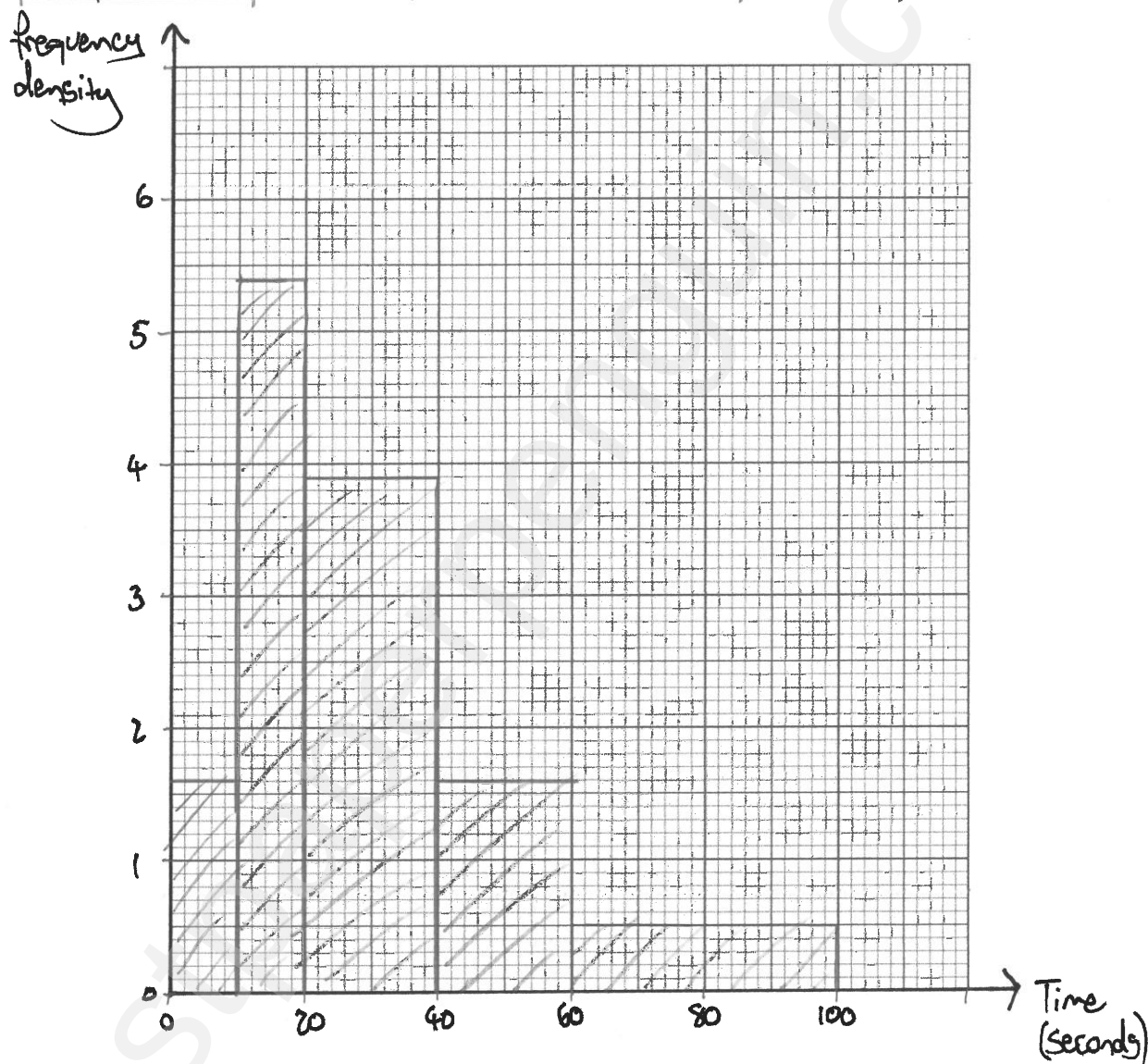
- 5 The times taken by 200 players to solve a computer puzzle are summarised in the following table.

Time (t seconds)	$0 \leq t < 10$	$10 \leq t < 20$	$20 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 100$
Number of players	16	54	78	32	20

- (a) Draw a histogram to represent this information.

[4]

class width	10	10	20	20	40
f.d.	1.6	5.4	3.9	1.6	0.5



- (b) Calculate an estimate of the mean time taken by these 200 players.

[2]

Mid-point (t)	Frequency (f)	f × t
5	16	80
15	54	810
30	78	2340
50	32	1600
80	20	1600
	$\Sigma f = 200$	$\Sigma ft = 6430$

$$\bar{t} = \frac{6430}{200}$$

$$= \underline{\underline{32.15}}$$

- (c) Find the greatest possible value of the interquartile range of these times.

[2]

Time	0-10	10-20	20-40	40-60	60-100
cumulative freq	16	70	148	180	200

$$Q_1: \frac{200}{4} = 50$$

↑
50
Q₁

↑
150
Q₃

$$Q_3: \frac{3 \times 200}{4} = 150$$

Greatest possible IQR = biggest possible Q₃ - smallest possible Q₁

$$= 60 - 10$$

$$= \underline{\underline{50}}$$

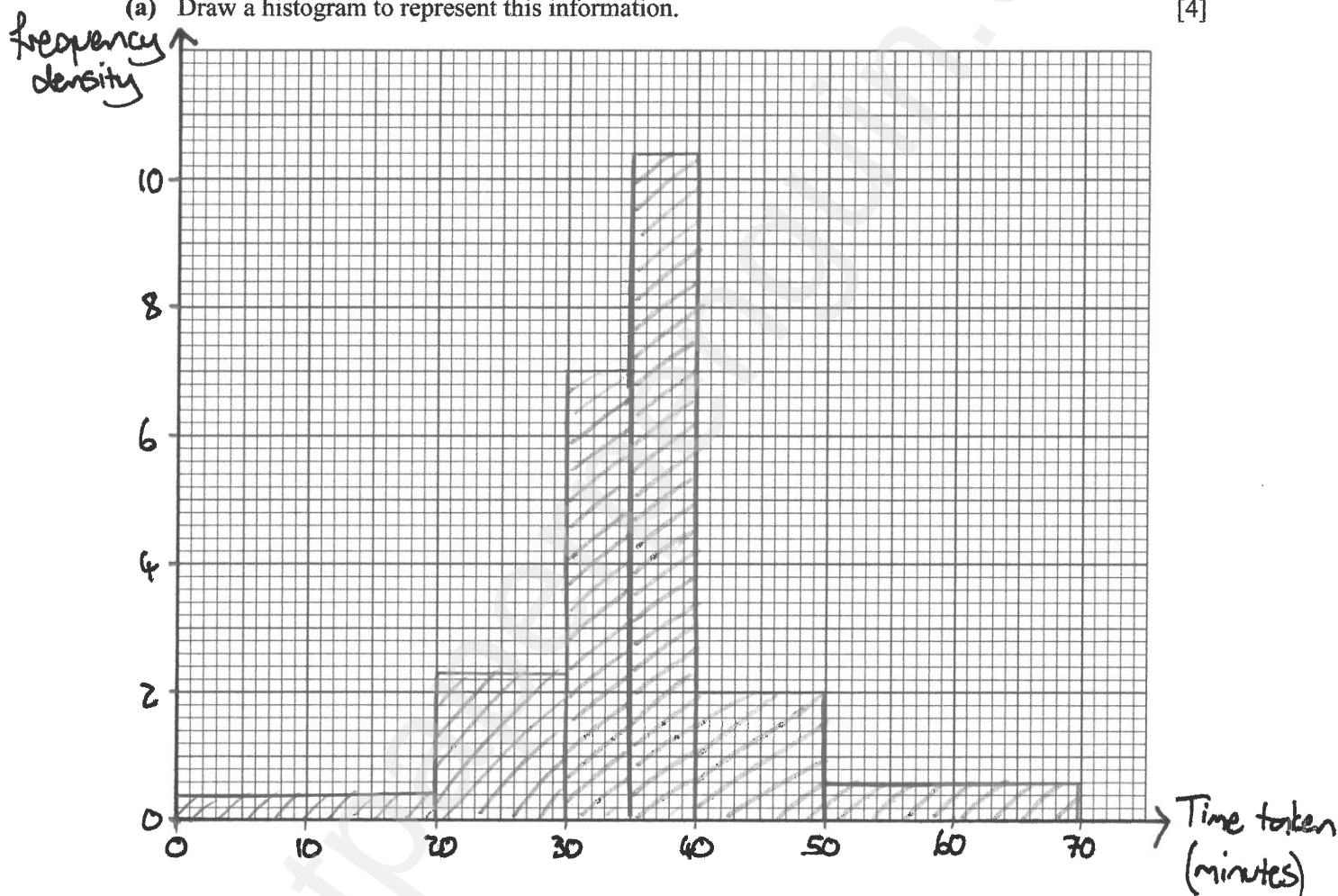
- 3 The times taken, in minutes, by 150 students to complete a puzzle are summarised in the table.

Time taken (t minutes)	$0 \leq t < 20$	$20 \leq t < 30$	$30 \leq t < 35$	$35 \leq t < 40$	$40 \leq t < 50$	$50 \leq t < 70$
Frequency	8	23	35	52	20	12

Class width	20	10	5	5	10	20
f.d.	0.4	2.3	7	10.4	2	0.6

- (a) Draw a histogram to represent this information.

[4]



- (b) Calculate an estimate for the mean time for these students to complete the puzzle.

[3]

Mid-point (t)	Frequency (f)	$f \times t$
10	8	80
25	23	575
32.5	35	1137.5
37.5	52	1950
45	20	900
60	12	720
	$\Sigma f = 150$	$\Sigma ft = 5362.5$

$$\begin{aligned}\bar{t} &= \frac{5362.5}{150} \\ &= \underline{35.75}\end{aligned}$$

- (c) In which class interval does the lower quartile of the times lie?

[1]

Time	0-20	20-30	30-35	35-40...
Cumulative freq	8	31	66	118

$$Q_1: \frac{150}{4} = 37.5^{\text{th}} \text{ value}$$

Q_1 lies in $30 \leq t < 35$

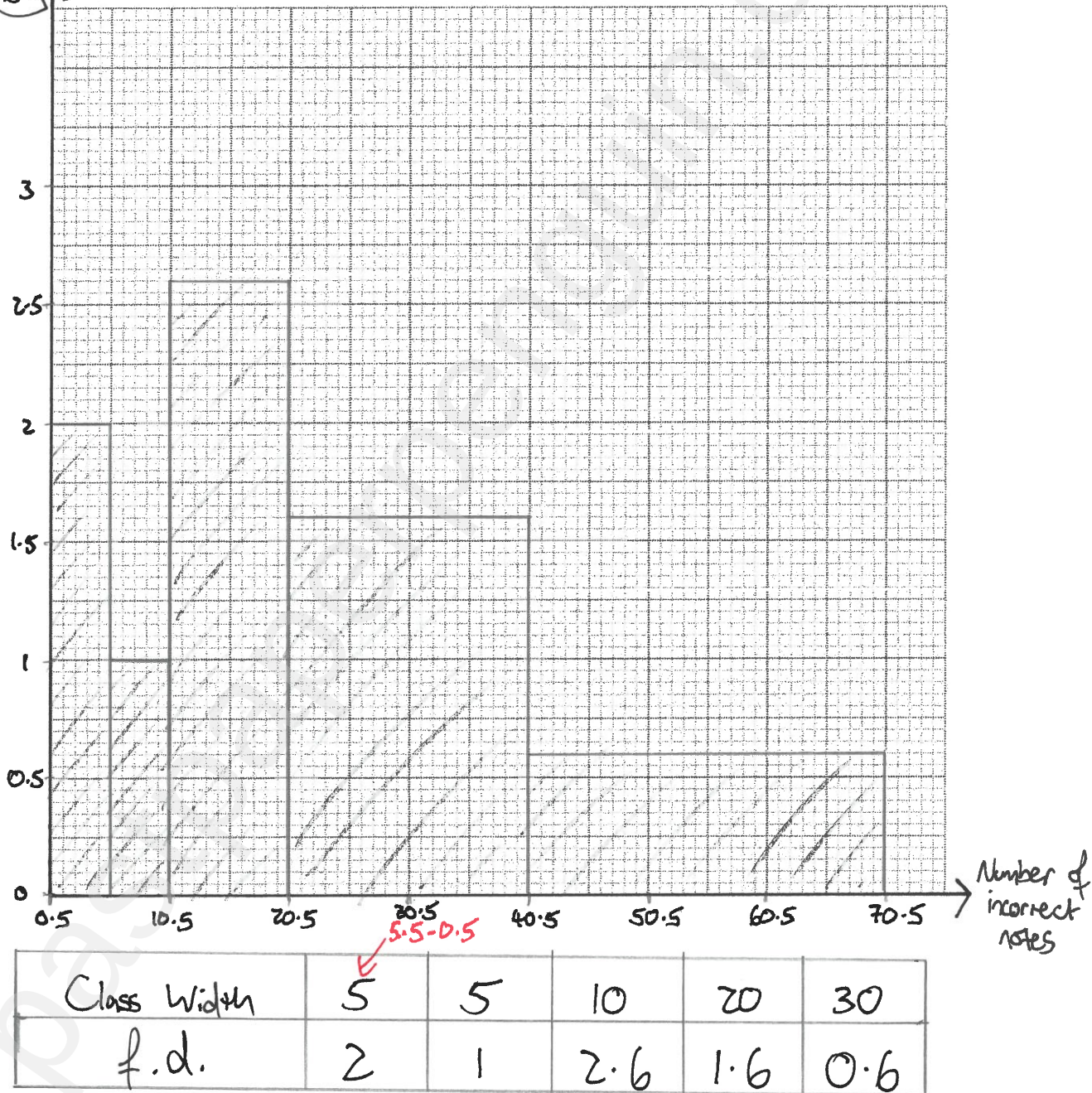
- 7 A particular piece of music was played by 91 pianists and for each pianist, the number of incorrect notes was recorded. The results are summarised in the table.

Number of incorrect notes	1 – 5	6 – 10	11 – 20	21 – 40	41 – 70
Frequency	10	5	26	32	18

(a) Draw a histogram to represent this information.

[5]

Frequency
density



- (b) State which class interval contains the lower quartile and which class interval contains the upper quartile.

Hence find the greatest possible value of the interquartile range.

[2]

Number	1-5	6-10	11-20	21-40	41-70
cumulative freq	10	15	41	73	91

$$Q_1: \frac{91}{4} = 22.75$$

$$Q_3: \frac{3 \times 91}{4} = 68.25$$

↑
22.75
Q₁

↑
68.25
Q₃

Q₁ is in 11-20

Q₃ is in 21-40

$$\begin{aligned} \text{Greatest IQR} &= \text{biggest possible } Q_3 - \text{smallest possible } Q_1 \\ &= 40 - 11 \\ &= \underline{29} \end{aligned}$$

- (c) Calculate an estimate for the mean number of incorrect notes.

[3]

Mid-point (x)	Frequency (f)	fxx
3	10	30
8	5	40
15.5	26	403
30.5	32	976
55.5	18	999
	$\Sigma f = 91$	$\Sigma fx = 2448$

$$\bar{x} = \frac{2448}{91}$$

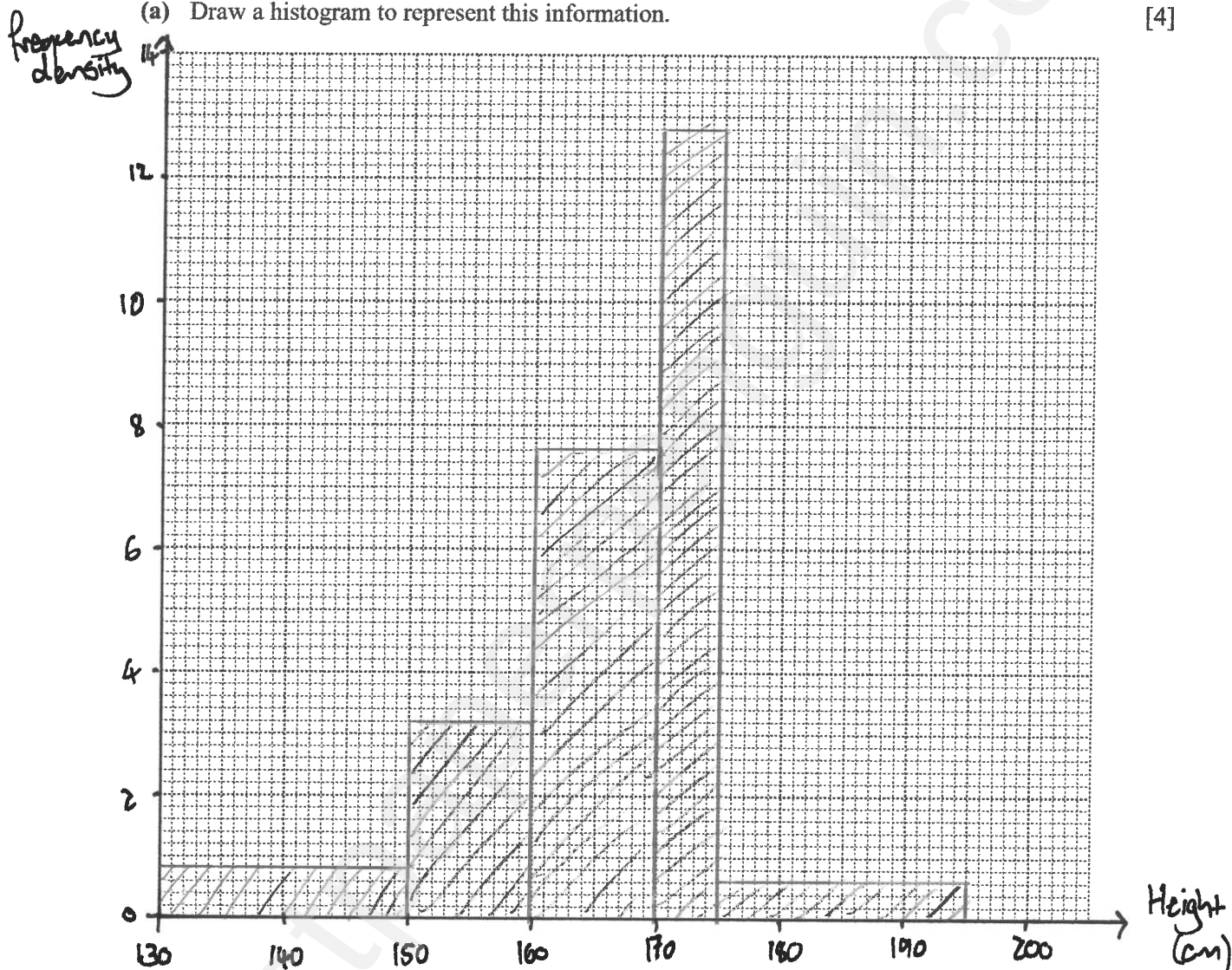
$$= \underline{26.9}$$

3 The heights, in cm, of 200 adults in Barimba are summarised in the following table.

Height (h cm)	$130 \leq h < 150$	$150 \leq h < 160$	$160 \leq h < 170$	$170 \leq h < 175$	$175 \leq h < 195$
Frequency	16	32	76	64	12
Class width	20	10	10	5	20
f.d.	0.8	3.2	7.6	12.8	0.6

(a) Draw a histogram to represent this information.

[4]



- (b) The interquartile range is R cm. Show that R is **not** greater than 15.

[2]

200 adults, so $Q_1 = 50^{\text{th}}$ person
and $Q_3 = 150^{\text{th}}$ person

so Q_1 lies in $160 \leq h < 170$
and Q_3 lies in $170 \leq h < 175$

Max. value for IQR would use the lowest possible value of Q_1 and the highest possible value of Q_3 :

$$\begin{aligned} & Q_3(\text{max}) - Q_1(\text{min}) \\ &= 175 - 160 \\ \text{Max IQR} &= 15 \end{aligned}$$

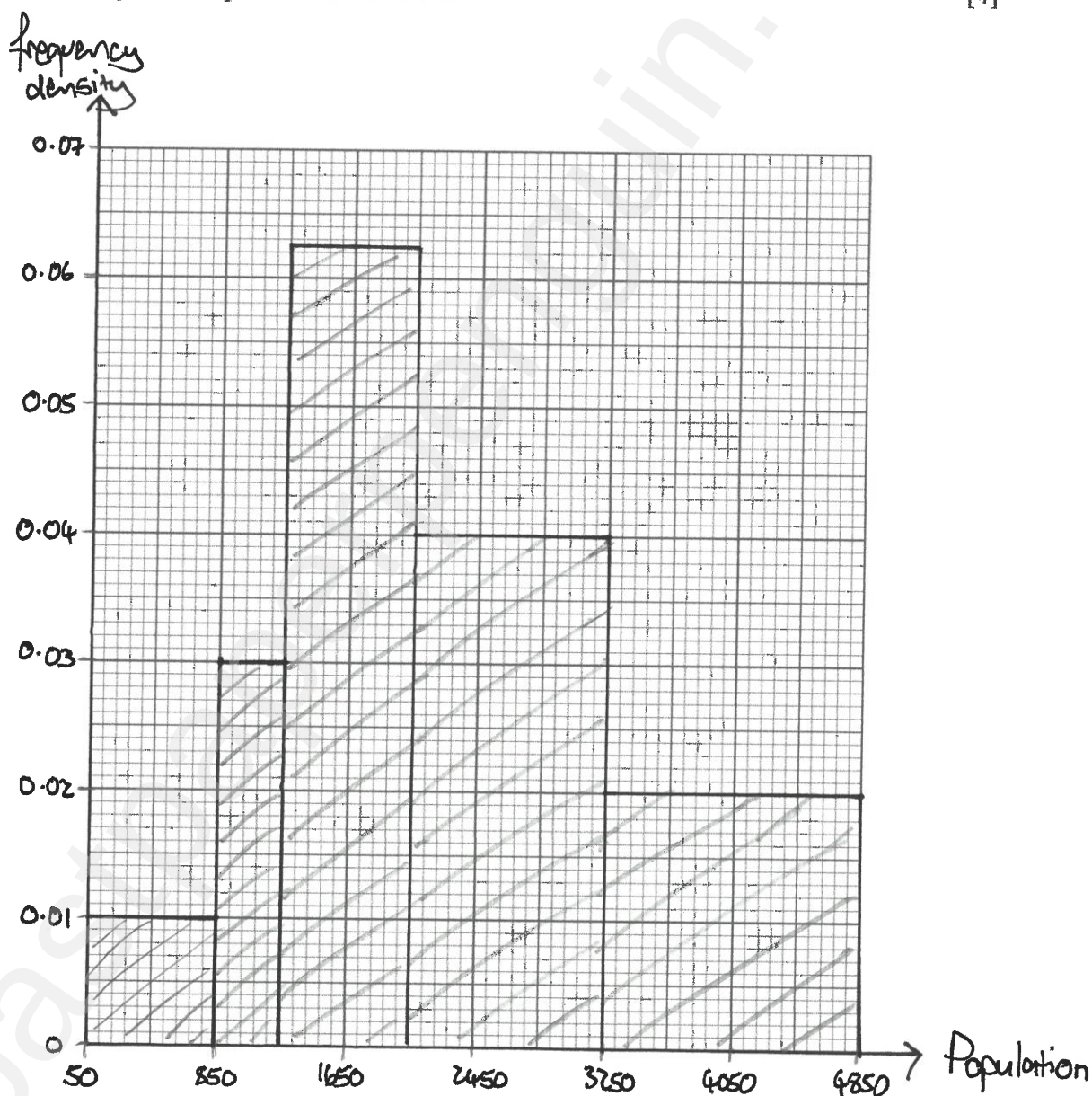
$\rightarrow$ so R cannot be greater than 15.

- 5 The populations of 150 villages in the UK, to the nearest hundred, are summarised in the table.

Population	100 – 800	900 – 1200	1300 – 2000	2100 – 3200	3300 – 4800
Number of villages	8	12	50	48	32
class width	800	400	800	1200	1600
f.d.	0.01	0.03	0.0625	0.04	0.02

- (a) Draw a histogram to represent this information.

[4]



- (b) Write down the class interval which contains the median for this information.

[1]

	50	450	1250	2050	3250	4850
Population	100-800	900-1200	1300-2000	2100-3200	3300-4800	
Cumulative freq	8	20	70	118	150	

$$Q_2: \frac{150}{2} = 75^{\text{th}} \text{ value}$$

Q_2 lies in 2100 - 3200

- (c) Find the greatest possible value of the interquartile range for the populations of the 150 villages.

[2]

$$Q_1: \frac{150}{4} = 37.5^{\text{th}} \text{ value (1300 - 2000 interval)}$$

$$Q_3: \frac{3(150)}{4} = 112.5^{\text{th}} \text{ value (2100 - 3200 interval)}$$

$$\text{Greatest IQR} = \max(Q_3) - \min(Q_1)$$

$$= 3249 - 1250$$

$$= \underline{1999}$$

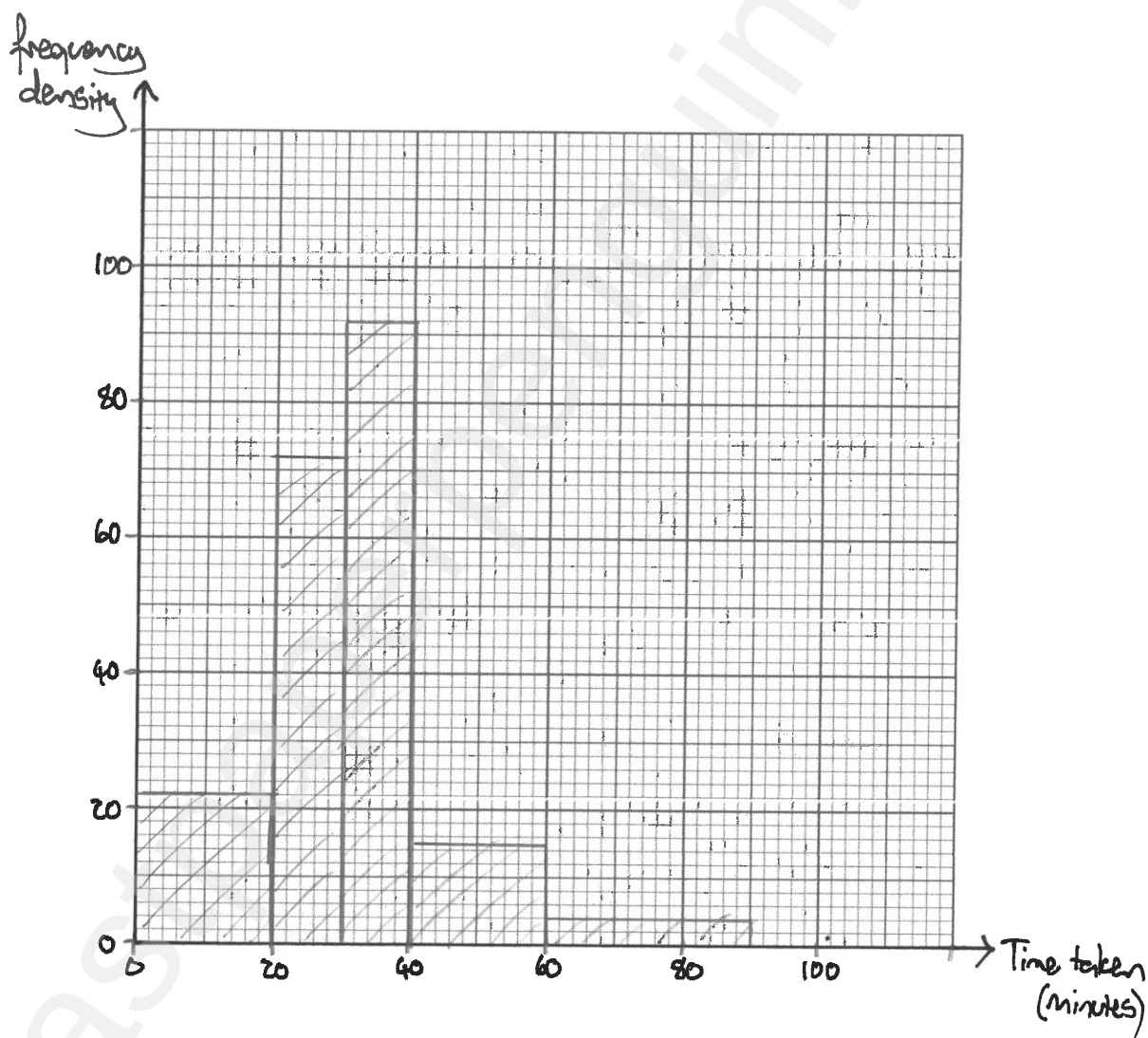
- 3 The times taken to travel to college by 2500 students are summarised in the table.

Time taken (t minutes)	$0 \leq t < 20$	$20 \leq t < 30$	$30 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 90$
Frequency	440	720	920	300	120

- (a) Draw a histogram to represent this information.

[4]

Class width	20	10	10	20	30
f.d.	22	72	92	15	4



From the data, the estimate of the mean value of t is 31.44.

- (b) Calculate an estimate of the standard deviation of the times taken to travel to college. [3]

Mid-points: 10, 25, 35, 50, 75

$$\text{Var} = \frac{10^2 \times 440 + 25^2 \times 770 + 35^2 \times 970 + 50^2 \times 300 + 75^2 \times 120}{2500} - 31.44^2$$

$$= 229.9264$$

$$\sigma = \sqrt{229.9264}$$

$$= \underline{15.2}$$

- (c) In which class interval does the upper quartile lie? [1]

Time	0-20	20-30	30-40	40-60	60-90
Cumulative freq.	440	1160	2080	2380	2500

$$Q_3: \frac{3 \times 2500}{4} = 1875$$

↑
1875 (Q_3)

Q_3 lies in $30 \leq t < 40$

It was later discovered that the times taken to travel to college by two students were incorrectly recorded. One student's time was recorded as 15 instead of 5 and the other's time was recorded as 65 instead of 75.

- (d) Without doing any further calculations, state with a reason whether the estimate of the standard deviation in part (b) would be increased, decreased or stay the same. [1]

It would stay the same: 15 and 5 would both be in the $0 \leq t < 20$ class, 65 and 75 would both be in $60 \leq t < 90$, so it doesn't affect the calculations in part (b).

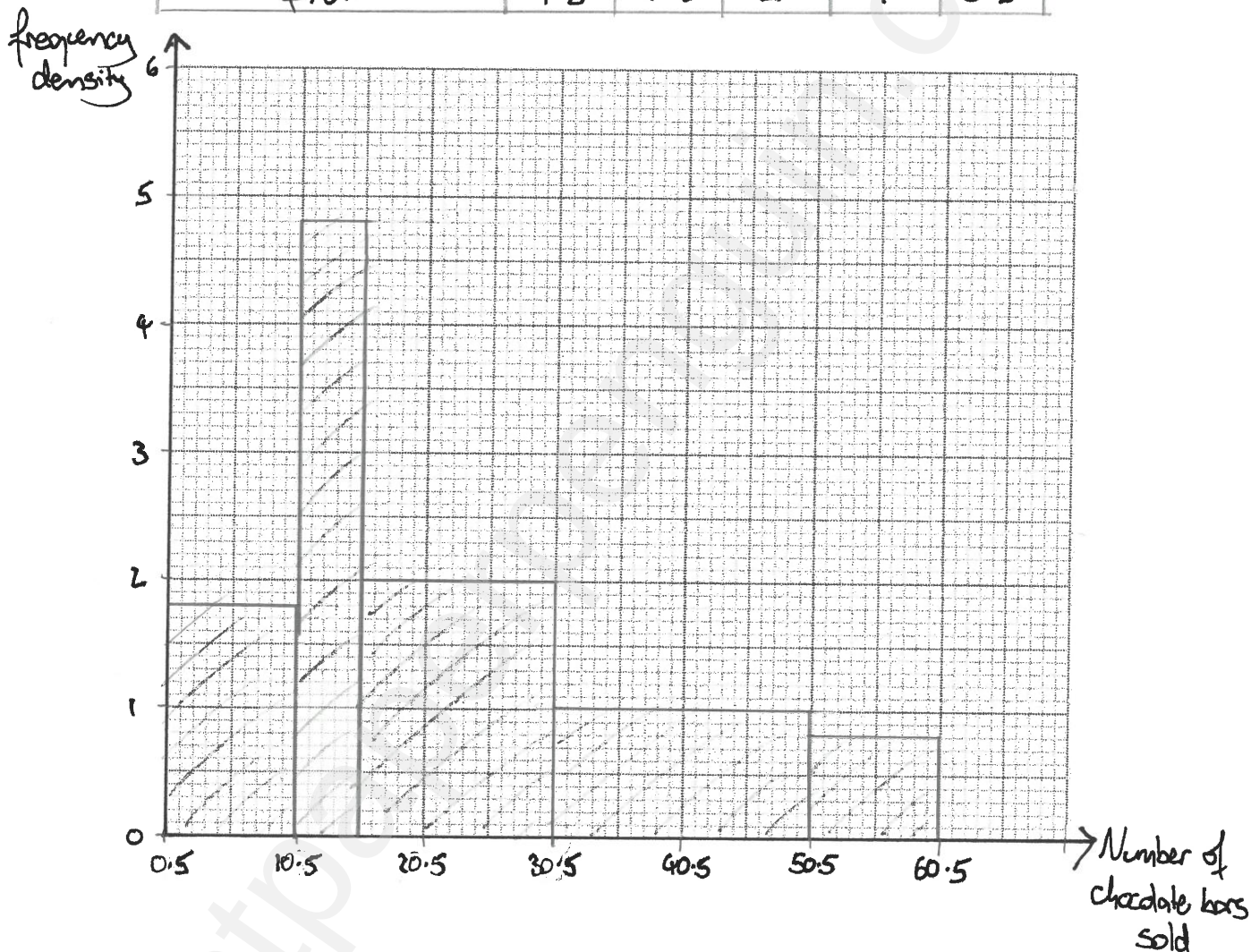
- 7 The numbers of chocolate bars sold per day in a cinema over a period of 100 days are summarised in the following table.

	0.5	10.5	15.5	30.5	50.5	60.5
Number of chocolate bars sold	1 – 10	11 – 15	16 – 30	31 – 50	51 – 60	
Number of days	18	24	30	20	8	

- (a) Draw a histogram to represent this information.

[5]

class width	10	5	15	20	10
f.d.	1.8	4.8	2	1	0.8



- (b) What is the greatest possible value of the interquartile range for the data? [2]

Number of bars	1-10	11-15	16-30	31-50	51-60
cumulative frequency	18	42	72	92	100

$$Q_1: \frac{100}{4} = 25$$

↑
25
Q₁

↑
75
Q₃

$$Q_3: \frac{3 \times 100}{4} = 75$$

$$\begin{aligned} \text{Greatest IQR} &= \max(Q_3) - \min(Q_1) \\ &= 75 - 42 \\ &= 33 \end{aligned}$$

- (c) Calculate estimates of the mean and standard deviation of the number of chocolate bars sold. [4]

Mid-point (x)	Frequency (f)	f × x
5.5	18	99
13	24	312
23	30	690
40.5	20	810
55.5	8	444
	$\Sigma f = 100$	$\Sigma fx = 2355$

$$\bar{x} = \frac{2355}{100}$$

$$= 23.55$$

$$\text{Var} = \frac{5.5^2 \times 18 + 13^2 \times 24 + 23^2 \times 30 + 40.5^2 \times 20 + 55.5^2 \times 8}{100} - 23.55^2$$

$$= 224.57$$

$$\sigma = \sqrt{224.57}$$

$$= 15.0$$