

0580

IGCSE Maths

Contents	Pages
1 Stem-and-leaf diagram	1
2 Histogram	2-3
3. Mean, median, mode and range of discrete data	4
4. Estimate of Mean of grouped data	5-7
5 Box-and-whisker plots	8
6. Cumulative frequency diagrams, median, quartiles, percentiles, inter-quartile range	9-16

Statistics

Revision

SP-20/M-20/S-20/W-19

Suresh Goel

(Former Director)

Alliance World School

Noida, Delhi-N.C.R.

INDIA

(+91 9810444804)

1. The number of cars parked in a car park at 9 am is recorded for 10 days.

124 130 129 116 132 120 127 107 118 114

Complete the stem-and-leaf diagram.

10	
11	
12	
13	

Key: 12/3 represents 123 cars.

---[2]

SP-20/02/Q6

Solution:

10	7			
11	4	6	8	
12	0	4	7	9
13	0	2		

2. The number of people swimming in a pool is recorded each day for 12 days.

24 28 13 38 15 26 45 21 48 36 18 38

(a) Complete the stem-and-leaf diagram.

1	
2	
3	
4	

Key: 1/3 represents 13 swimmers.

---[2]

(b) Find the median number of swimmers.

---[1]

Solution (a)

1	3	5	8	
2	1	4	6	8
3	6	8	8	
4	5	8		

(b) Median number = $\frac{12+1}{2}th = 6.5th = \frac{6th+7th}{2} = \frac{26+28}{2} = 27$

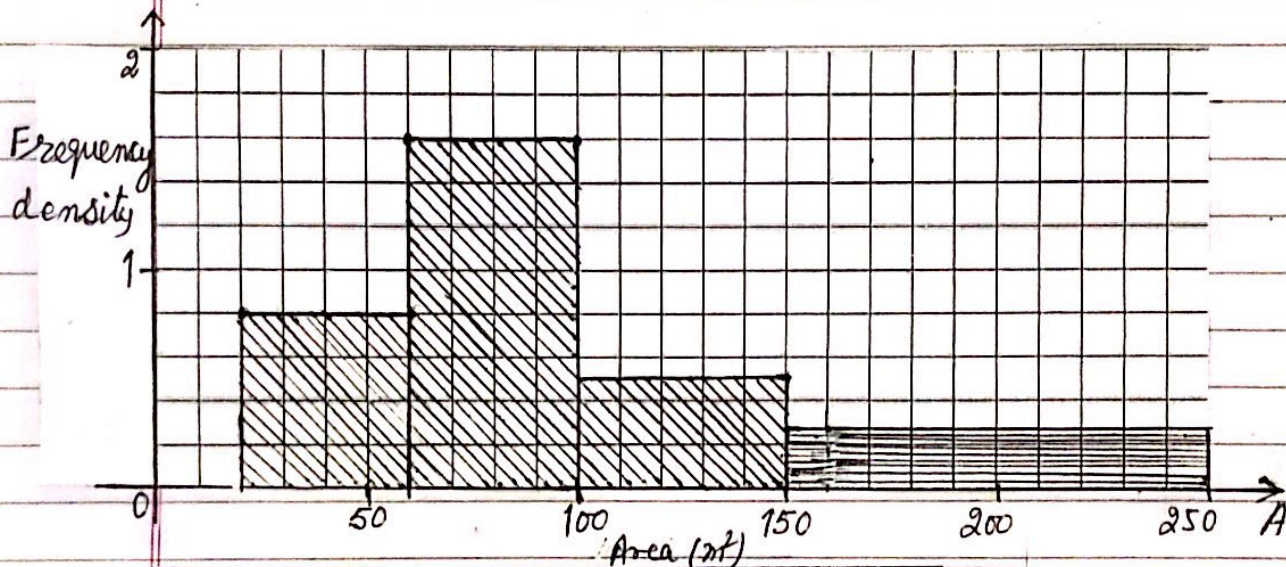
Histogram

3. The 200 students also estimate the total area, $A m^2$, of the windows in the class room. The table shows their results:

Area ($A m^2$)	$20 < A \leq 60$	$60 < A \leq 100$	$100 < A \leq 150$	$150 < A \leq 250$
Frequency	32	64	80	24

(i.) Complete the histogram to the information in the table. --[4]

/SP-20/04/2(b)(i)(ii)



Classes	frequency f	Frequency density = $\frac{\text{frequency}}{\text{class size}}$
$20 < A \leq 60$	32	$\frac{32}{40} = 0.8$
$60 < A \leq 100$	64	$\frac{64}{40} = 1.6$
$100 < A \leq 150$	80	$\frac{80}{150} = 0.53$
$150 < A \leq 250$	24	$\frac{24}{100} = 0.24$

(ii) Two students are chosen at random from those students that estimated area of the windows to be more than $100 m^2$.

Find the prob. that one of the students estimates the area to be greater than $150 m^2$ and the other student estimates the area to be $150 m^2$ or less. --[3]

Solution: Number of students who estimate area more than $100 m^2 = 80 + 24 = 104$

----- more than $150 m^2 = 24$
----- $150 m^2$ or less = 80

$$\therefore \text{Required Prob} = 2 \times \frac{24}{104} \times \frac{80}{103} = \frac{3840}{10712} = \frac{480}{1339} \checkmark$$

4 Given frequency table:

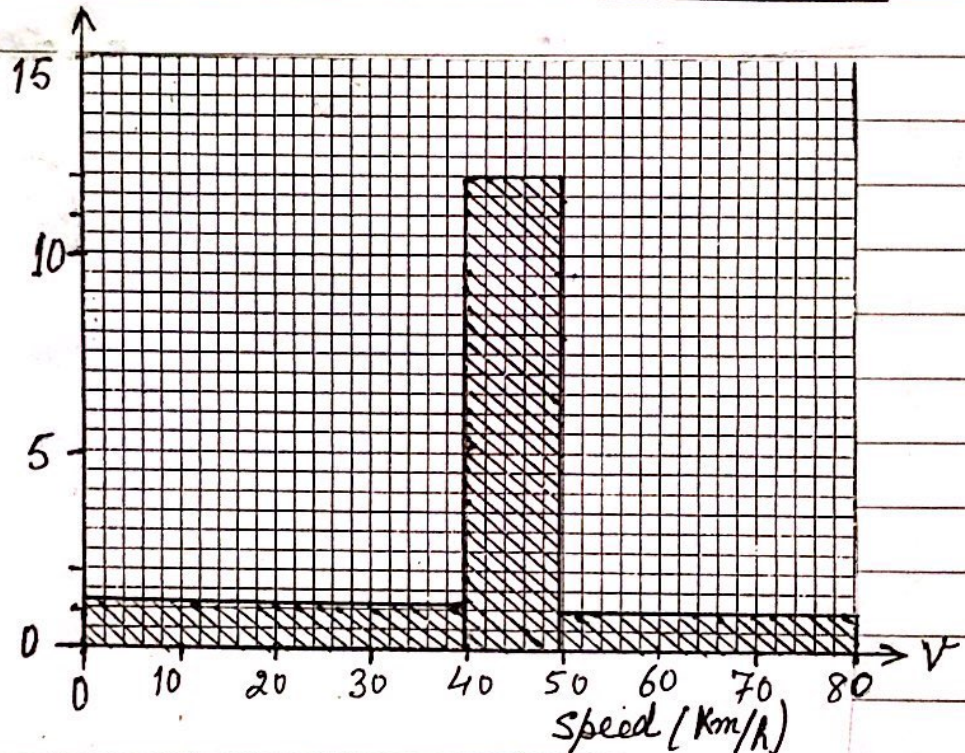
Speed (V Km/R)	$0 < V \leq 40$	$40 < V \leq 50$	$50 < V \leq 80$
Frequency	50	120	30

On the grid draw a histogram to show the information in this table:

S-20/42/23(d) --- [3]

Solution:

Frequency density



Classes V (Km/R)	Frequency f	Frequency density
$0 \leq V \leq 40$	50	$\frac{50}{40} = 1.25$
$40 < V \leq 50$	120	$\frac{120}{10} = 12$
$50 < V \leq 80$	30	$\frac{30}{30} = 1$

Frequency density
= $\frac{\text{frequency}}{\text{class width}}$.

Mean, Median, Mode and Range of Discrete data

5. The table shows the marks scored by 40 students in a test.

Mark	5	6	7	8	9	10	
Frequency	8	5	11	7	5	4	...[3]

Calculate mean mark

M-20 | 22 | Q6

Solution:

Mark x	Frequency f	f.x	
5	8	40	Mean $\bar{x} = \frac{\sum fx}{\sum f}$ $= \frac{288}{40}$ $\therefore \text{Mean} = 7.2 \checkmark$
6	5	30	
7	11	77	
8	7	56	
9	5	45	
10	4	40	
$\sum f = 40$		$\sum fx = 288$	

6. Roberto records the value of each of the coins he has at home.

Value (Cents)	1	2	5	10	20	50
Frequency	3	1	3	2	4	2

- (i) Find the range. W-19 | 41 | Q6(b) -- [1]
- (ii) Find the mode. ... [1]
- (iii) Find the median. ... [1]
- (iv) Total value of Roberto's coins. ... [2]
- (v) Work out the mean. ... [1]

Solution:

(v)

Value (Cents) x	Frequency f	f.x	
1	3	3	(i) Range = $50 - 1 = 49 \checkmark$ (ii) Mode = 20 \checkmark (Max freq is 4) (iii) Median = $\frac{15+1}{2} = 8_{th} = 10 \checkmark$ (iv) Total Value = $\sum fx = 220 \checkmark$ (v) Mean = $14.67 \checkmark$
2	1	2	
5	3	15	
10	2	20	
20	4	80	
50	2	100	

$\sum f = 15$ $\sum fx = 220$

Mean = $\frac{\sum fx}{\sum f} = \frac{220}{15} = 14.67 \checkmark$

Estimate of Mean of the Grouped data

DATE: / /

PAGE: P-5

7. The table shows the amount of money, \$x, given to charity by each of 60 people.

Amount (\$x)	$0 < x \leq 20$	$20 < x \leq 25$	$25 < x \leq 35$	$35 < x \leq 50$	$50 < x \leq 100$
Frequency	21	16	6	10	7

Calculate an estimate of the mean.

[S-20/23/Q15] --- [4]

Solution:

Amount	Frequency f	Mid Value x	f.x
$0 < x \leq 20$	21	10	210
$20 < x \leq 25$	16	22.5	360
$25 < x \leq 35$	6	30	180
$35 < x \leq 50$	10	42.5	425
$50 < x \leq 100$	7	75	525

$$\text{Mean} = \frac{\sum fx}{\sum f}$$

$$= \frac{1700}{60}$$

$$= 28.33$$

$$\therefore \text{Estimated Mean} = 28.33 \checkmark$$

$$\sum f = 60$$

$$\sum fx = 1700$$

8. The 200 students estimate the total area, $A \text{ m}^2$, of the windows in the class room. The table shows their results.

Area ($A \text{ m}^2$)	Frequency f	Mid Value x	f.x
$20 < A \leq 60$	32	40	1280
$60 < A \leq 100$	64	80	5120
$100 < A \leq 150$	80	125	10000
$150 < A \leq 250$	24	200	4800

$$\text{Estimated Mean} = \frac{\sum fx}{\sum f}$$

$$= \frac{21200}{200}$$

$$= 106 \text{ m}^2 \checkmark$$

$$\sum f = 200$$

$$\sum fx = 21200$$

* Note: (i) Explain why your answer is an estimated mean?

We assume that all the values of data (in any interval) are at the mid value of each class interval.

(ii) Can we find the range of the grouped frequency distribution?
No, as the raw data is unavailable, not knowing the individual values, we do not know the highest or the lowest values.

[W-17/43/Q4(C)(i)]

9.(a) The heights, h metres, of the 120 boys in a athletics club are recorded. The table shows this information:

Height (in metres)	$1.3 < h \leq 1.4$	$1.4 < h \leq 1.5$	$1.5 < h \leq 1.6$	$1.6 < h \leq 1.7$	$1.7 < h \leq 1.8$	$1.8 < h \leq 1.9$
Frequency	7	18	30	24	27	14

(i) write down the modal class ---[1]

(ii) calculate an estimate of the mean. [S-20/41/Q2(a)(b)]---[4]

(b)(i) one boy is chosen at random from the club.

Calculate the prob. that this boy has a height greater than 1.8m. ---[1]

(ii) Three boys are chosen at random from the club.

Calculate the prob that one of the boys has a height greater than 1.8m and the other two boys each have a height of 1.4m or less. ---[4]

Solution (a)(i) Modal class $1.5 < h \leq 1.6$

Classes	Frequency f	Mid Value x	$f \cdot x$
$1.3 < h \leq 1.4$	7	1.35	9.45
$1.4 < h \leq 1.5$	18	1.45	26.10
$1.5 < h \leq 1.6$	30	1.55	46.50
$1.6 < h \leq 1.7$	24	1.65	39.60
$1.7 < h \leq 1.8$	27	1.75	47.25
$1.8 < h \leq 1.9$	14	1.85	25.90

$$\begin{aligned} \text{Estimated mean} &= \frac{\sum fx}{\sum f} \\ &= \frac{194.8}{120} \\ &= \underline{\underline{1.623}} \checkmark \end{aligned}$$

$$\sum f = 120 \quad \sum fx = 194.8$$

b(i) $P(x > 1.8) = \frac{14}{120} \checkmark$

(ii) $P = \frac{3}{120} \times \frac{14}{119} \times \frac{7}{118} \times \frac{6}{117}$
 $= \frac{21}{20060} \checkmark$

$$\begin{aligned} & \left[\begin{aligned} P_1(x > 1.8) &= \frac{14}{120} \\ \text{and} \\ P_2(x \leq 1.4) &= \frac{7}{120} \end{aligned} \right. \\ & (P_1 \checkmark P_2 \checkmark P_2 + P_2 \checkmark P_1 \checkmark P_2 + P_2 \checkmark P_2 \checkmark P_1) \end{aligned}$$

10. The speed, V km/h, of each of 200 cars passing a building is measured. The table shows the results.

Speed (V km/h)	$0 < V \leq 20$	$20 < V \leq 40$	$40 < V \leq 45$	$45 < V \leq 50$	$50 < V \leq 60$	$60 < V \leq 80$
Frequency	16	34	62	58	26	4

(a) Calculate an estimate of mean.

[S-20/42/Q3(a)] -- [4]

(Speed) Classes	Frequency f	Mid Value x	$f \cdot x$
$0 < V \leq 20$	16	10	160
$20 < V \leq 40$	34	30	1020
$40 < V \leq 45$	62	42.25	2619.5
$45 < V \leq 50$	58	47.25	2740.5
$50 < V \leq 60$	26	55	1430
$60 < V \leq 80$	4	70	280

$$\begin{aligned} \text{Estimated mean} &= \frac{\sum f \cdot x}{\sum f} \\ &= \frac{8280}{200} \\ &= \underline{41.4} \checkmark \end{aligned}$$

$$\sum f = 200$$

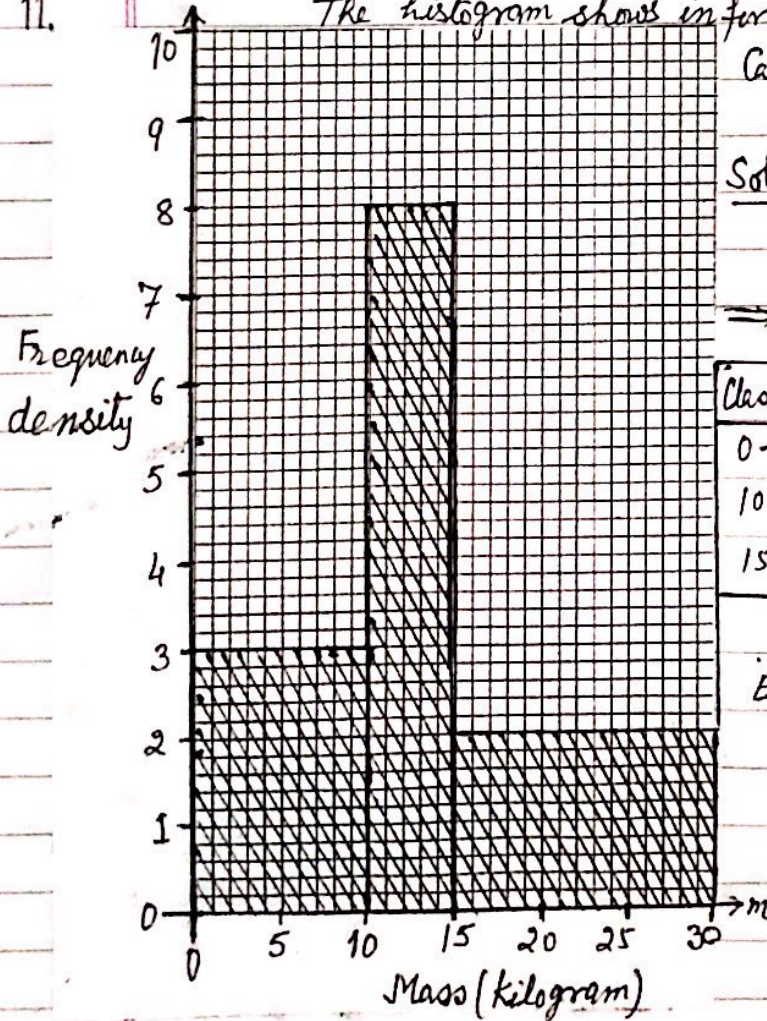
$$\sum f \cdot x = 8280$$

11.

The histogram shows information about the mass of 100 boxes.

Calculate an estimate of the mean. [6]

[W-19/41/Q6(c)]



Solution:

$$\text{Frequency density (F.D)} = \frac{\text{frequency}}{\text{Class size}}$$

$$\Rightarrow \text{frequency} = \text{F.D} \times \text{class size}$$

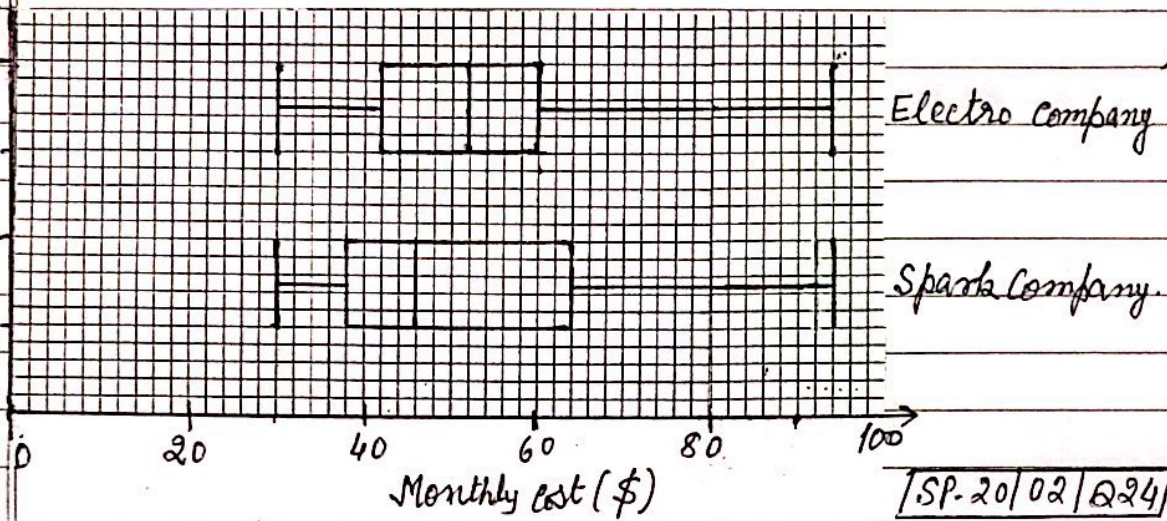
Classes	Freq density	frequency f	Mid Value x	$f \cdot x$
0-10	3	$3 \times 10 = 30$	5	150
10-15	8	$8 \times 5 = 40$	12.5	500
15-30	2	$2 \times 15 = 30$	22.5	675

$$\sum f = 100$$

$$\sum f \cdot x = 1325$$

$$\begin{aligned} \text{Estimated mean} &= \frac{\sum f \cdot x}{\sum f} \\ &= \frac{1325}{100} \\ &= \underline{13.25} \checkmark \end{aligned}$$

12. These box-and-whisker plots show the monthly electricity costs for 100 different household who use Electro company or Spark Company.



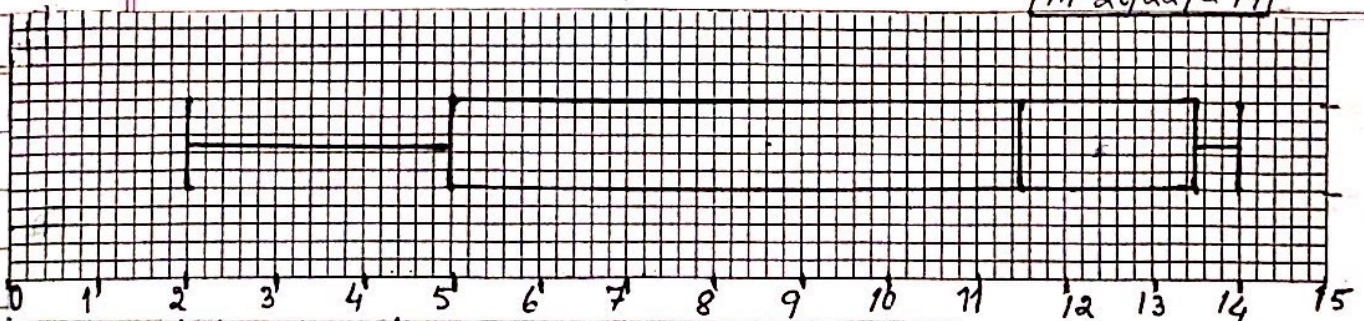
Tom says that the monthly costs with Electro company are lower and very less than with spark company.
Is Tom correct.

Justify your answer with reference to the box-and-whisker plots.

Answer: Median is greater for Electro company so Tom is wrong, as spark is cheaper. (spark median = 46 and Electro median = 52)
IQR is greater for Spark company so Tom is right spark is more varied. (for Spark IQR = 26 and Electro IQR = 18)

13. The box-and-whisker plot gives information about the height, in cm, of some plants.

[M-20/22/Q-14]



(a) write down the median.

⇒ Answer (a) Median = 11.5 ✓ --[1]

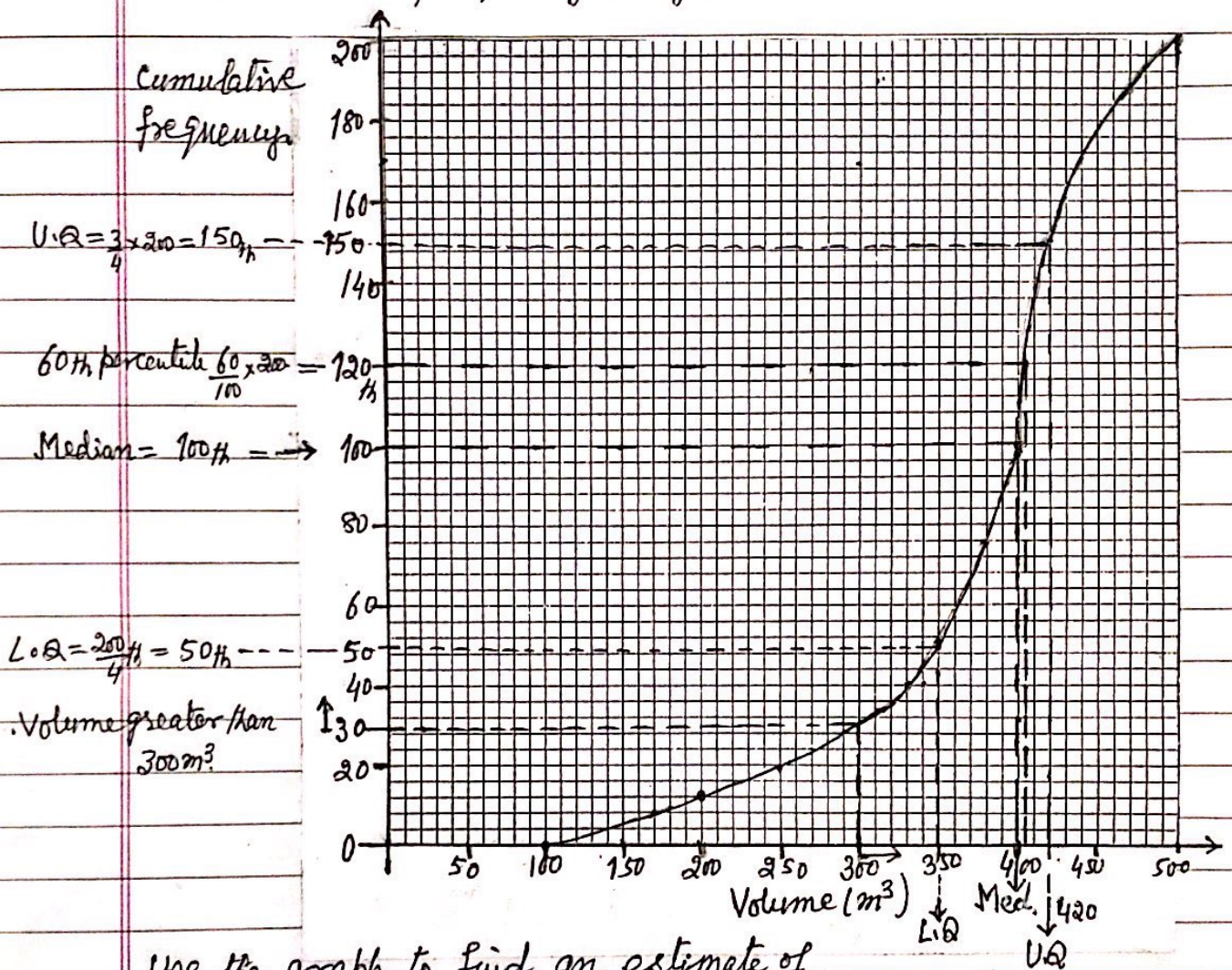
(b) Find (i) the range.

(b) (i) the range = 14 - 2 = 12 ✓ --[1]

(ii) the interquartile range.

(ii) the interquartile range = 13.5 - 5 = 8.5 ✓ --[1]

14. 200 students estimate the volume, $V \text{ m}^3$, of a classroom. The cumulative frequency diagram shows their results.



Use the graph to find an estimate of

- (i) the median, [1]
- (ii) the interquartile range, [2]
- (iii) the 60th percentile, [1]
- (iv) the number of students who estimate that the volume is greater than 300 m^3 . [2]

$N = 200$

SP-20/04/Q2(a)

Solution (i) Median = $\frac{N}{2} = \frac{200}{2} = 100$ ✓

(ii) IQR = Upper quartile - lower quartile = $420 - 350 = 70$ ✓

(iii) the 60th percentile = 60% of 200 = 120 = 405 to 410 ✓

(iv) Volume greater than 300 m^3 - No. of students = $200 - 30 = 170$ ✓

15 The heights, h metres, of the 120 boys in an athletics club are recorded. The table shows the information:

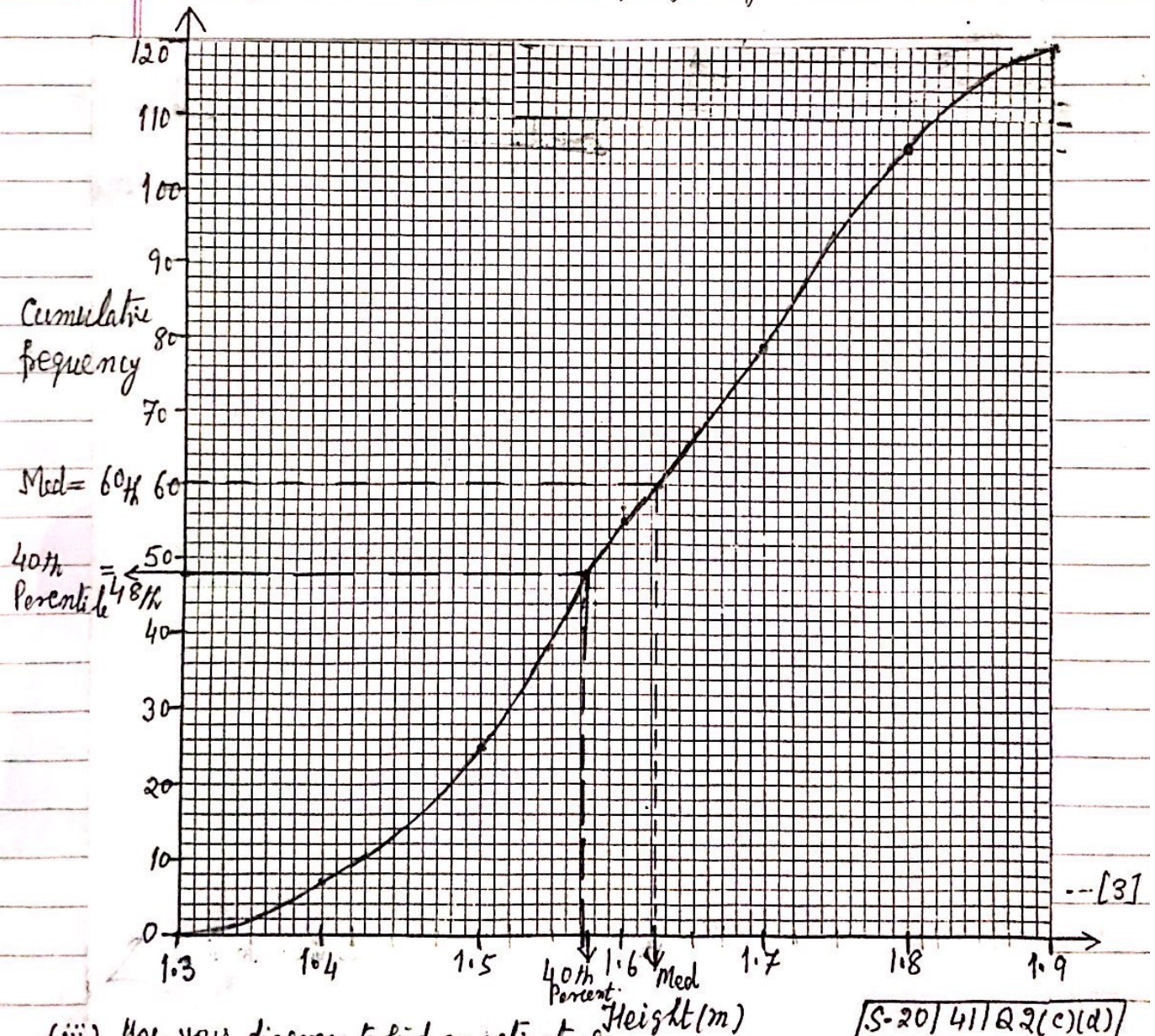
Height (h metres)	$1.3 < h \leq 1.4$	$1.4 < h \leq 1.5$	$1.5 < h \leq 1.6$	$1.6 < h \leq 1.7$	$1.7 < h \leq 1.8$	$1.8 < h \leq 1.9$
Frequency	7	18	30	24	27	14

(i) Use the frequency table above to complete the cumulative frequency table.

Height (h metres)	$h \leq 1.4$	$h \leq 1.5$	$h \leq 1.6$	$h \leq 1.7$	$h \leq 1.8$	$h \leq 1.9$
Cumulative frequency	7	25	55	79	106	120

---[2]

(ii) On the grid draw a cumulative frequency diagram to show this information:



(iii) Use your diagram to find an estimate for:

(a) the median height : $\frac{120}{2} = 60^{\text{th}} = 1.62$ to 1.63 ✓ ---[1]

(b) the 40th percentile : $\frac{40}{100} \times 120 = 48^{\text{th}} = 1.57$ to 1.58 ✓ ---[2]

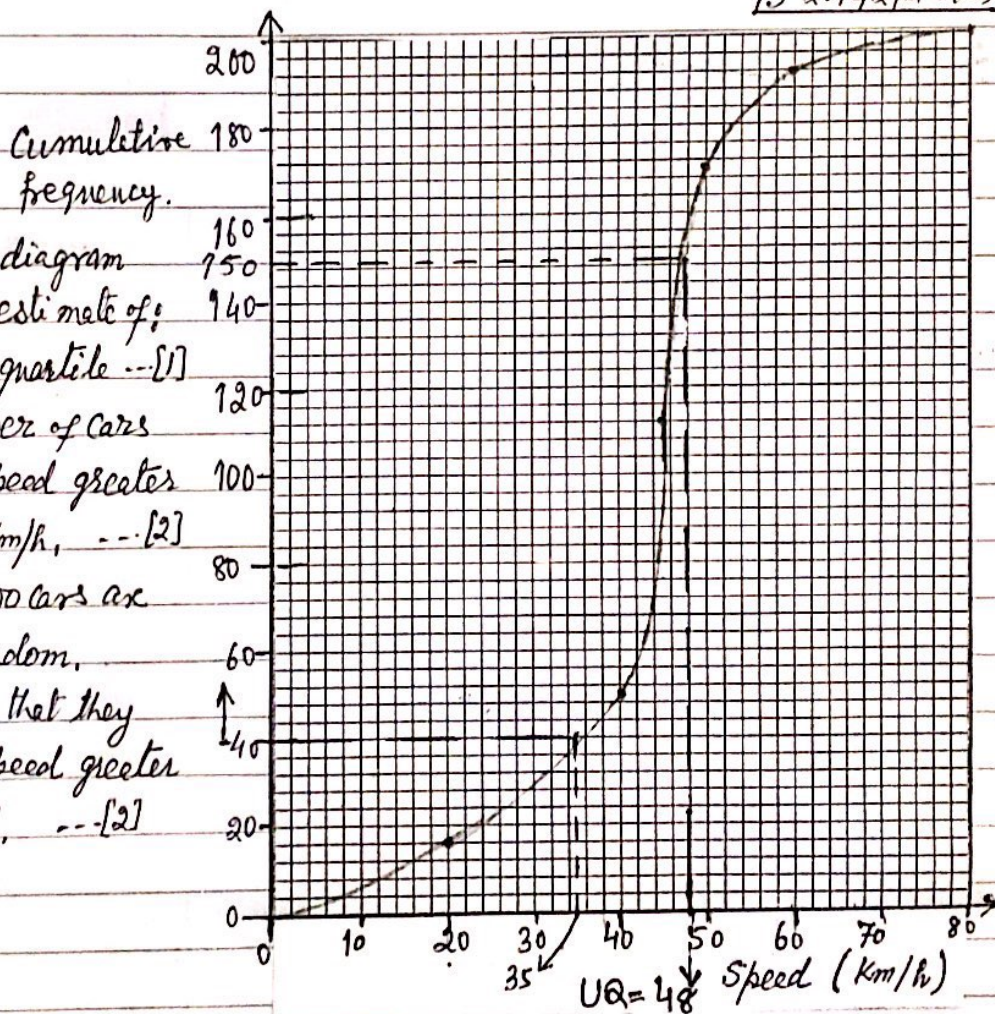
16. The speed, V km/h, of each of 200 cars passing a building is measured. The table shows the results.

Speed (V km/h)	$0 < V \leq 20$	$20 < V \leq 40$	$40 < V \leq 45$	$45 < V \leq 50$	$50 < V \leq 60$	$60 < V \leq 80$
Frequency	16	34	62	58	26	4

(i) Use the frequency table above to complete the cumulative frequency table.

Speed (V km/h)	$V \leq 20$	$V \leq 40$	$V \leq 45$	$V \leq 50$	$V \leq 60$	$V \leq 80$
cumulative frequency	16	50	112	170	196	200

(ii) On the grid, draw a cumulative frequency diagram. [3]



(iii) Use your diagram to find an estimate of:

(a) the upper quartile [1]

(b) the number of cars with a speed greater than 35 km/h. [2]

(iv) Two of the 200 cars are chosen at random. Find the prob. that they both have a speed greater than 50 km/h. [2]

Solution (iii)

(a) upper quartile = $\frac{3}{4} \times 200 = 150 = 48$ km/h.

(b) No. of cars greater than 35 km/h = $(200 - 40) = 160$ cars.

(iv) $P(\text{Two cars chosen at random, } V > 50 \text{ km/h}) = \frac{30}{200} \times \frac{29}{199} = \frac{87}{3980}$

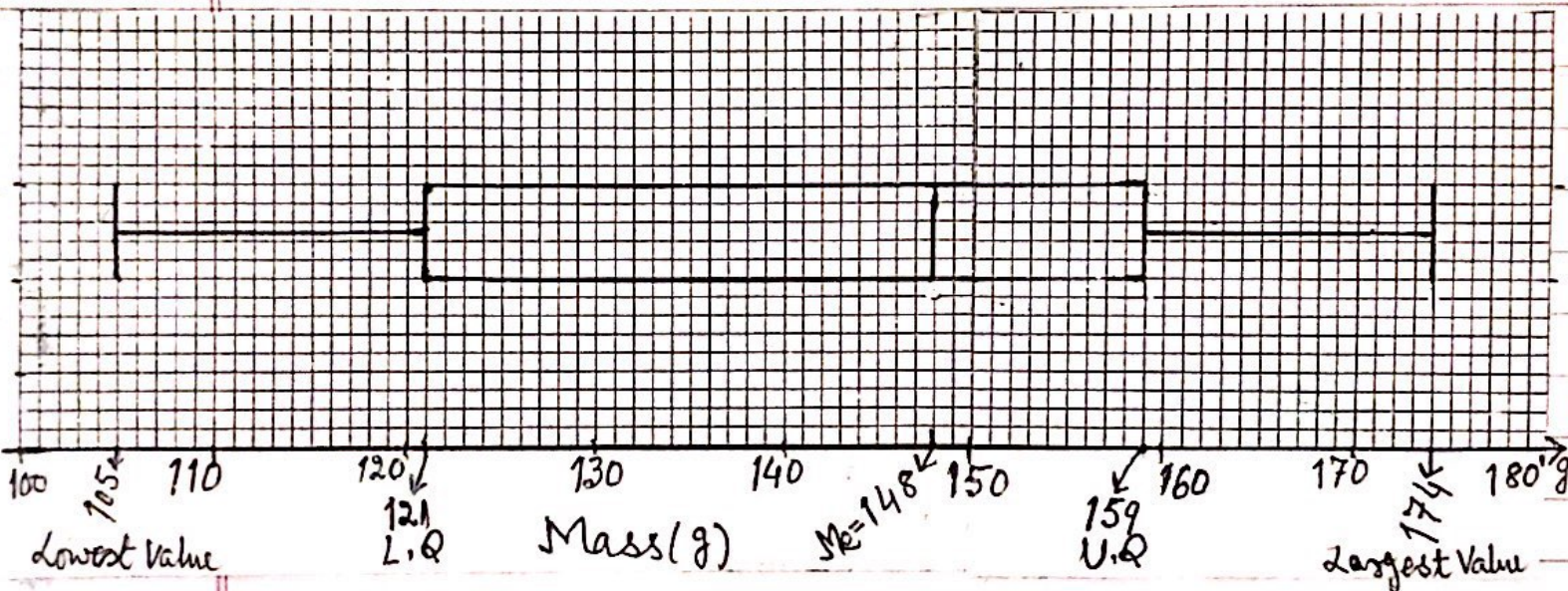
(∵ No. of Car with $V > 50 \text{ km/h} = 200 - 170 = 30$)

16. Here is some information about the masses of potatoes in a sack:

- The largest potato has a mass 174g.
- The range is 69g.
- The median is 148g
- The lower quartile is 121g
- The interquartile range is 38g.

S-20 | 43 | Q3(a) [4]

on the grid below, draw a box-and-whisker plot to show this information.



$$\text{Largest Value} = 174g$$

$$\text{Lowest Value} = 174 - 69 = 105g \quad (\because \text{Range} = \text{Largest Value} - \text{Lowest Value})$$

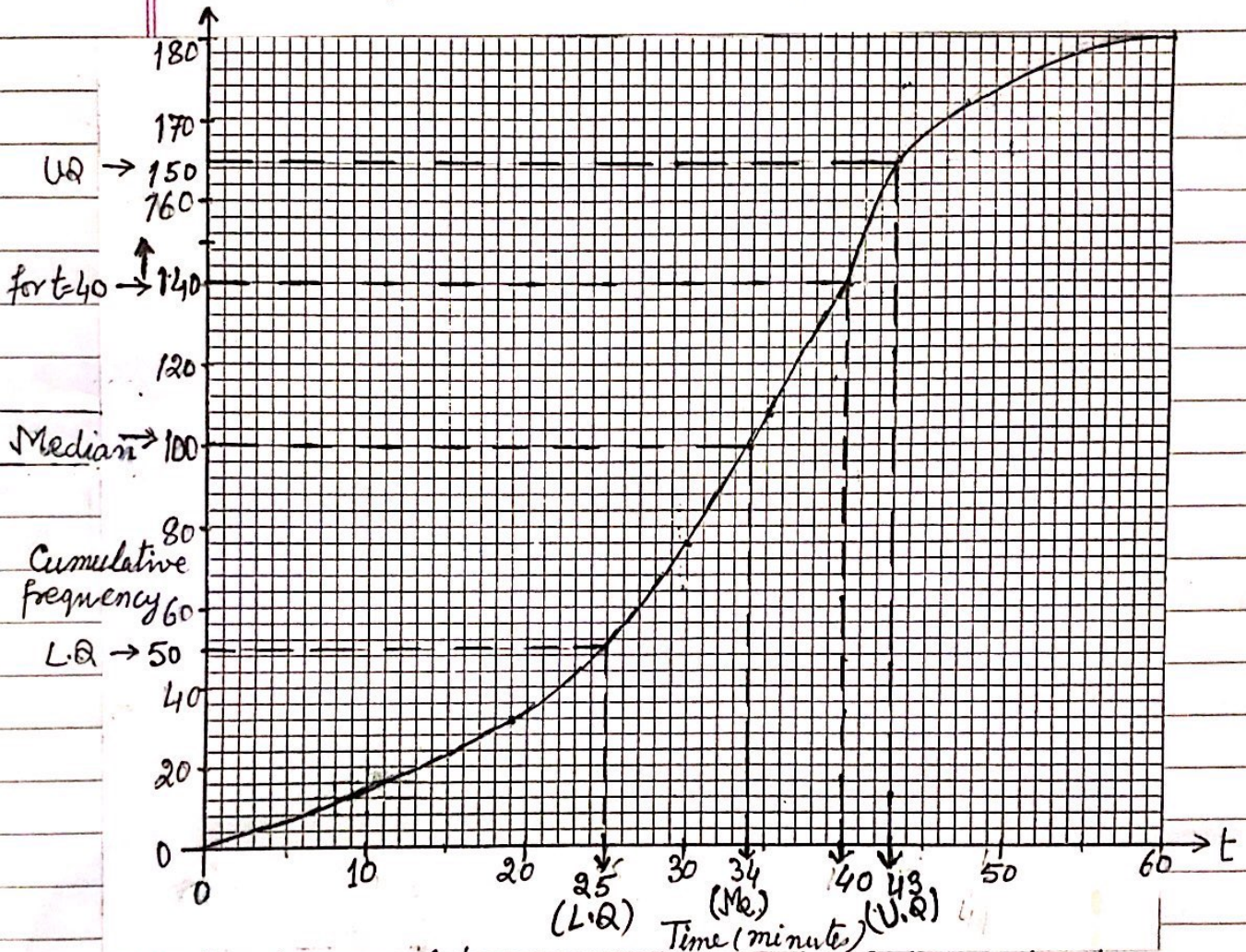
$$\text{Median} = 148g$$

$$\text{L.Q} = 121g$$

$$\text{I.Q.R} = 38.$$

$$\therefore \text{Upper quartile U.Q} = \text{L.Q} + \text{I.Q.R} = 121 + 38 = 159g \quad \checkmark$$

17. The cumulative frequency diagram shows information about the times taken by 200 students to solve a problem. W-19/41/Q6(a)



Use the cumulative frequency diagram to find an estimate for

- (i) the median ---[1]
- (ii) the interquartile range. ---[2]
- (iii) the number of students who took more than 40 minutes, -- [2]

Solution (i) Median = $\frac{200}{2}th = 100th = \underline{34} \checkmark$

(ii) I.Q.R = UQ - LQ

$$= 43 - 25 = \underline{18} \checkmark$$

$$\left\{ \begin{array}{l} U.Q = \frac{3}{4} \times 200th = 150th = 43 \\ L.Q = \frac{1}{4} \times 200th = 50th = 25 \end{array} \right.$$

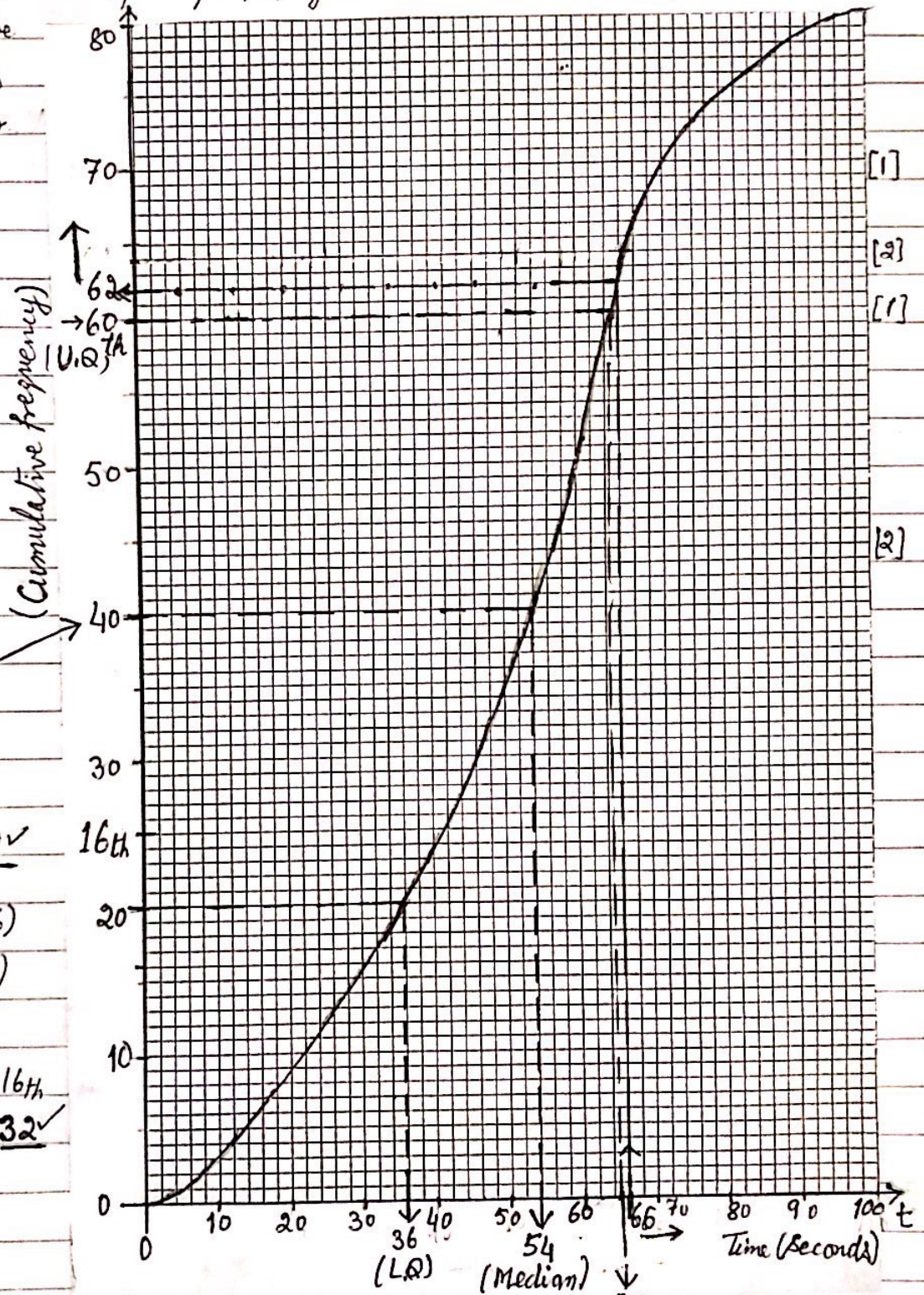
(iii) Number of students who took more than 40 mins = 200 - c.f at t=40 m
= 200 - 140 = 60 \checkmark

18. The cumulative frequency diagram shows information about the time taken, t seconds, for a group of girls to each solve a maths problem.

Use the cumulative frequency diagram to find an estimate for

- (i) the median
- (ii) the interquartile range.

- (iii) the 20% percentile
- (iv) the number of girls who took more than 66 seconds to solve the problem.



Solution:

(i) Median = $\frac{80}{2}$
= 40
= 54 ✓

(ii) IQR = UQ - LQ
= $65 - 36 = 29$ ✓

(LQ = $\frac{80}{4} = 20 = 36$)

(UQ = $\frac{3 \times 80}{4} = 60 = 65$)

(iii) 20% percentile
= $\frac{20}{100} \times 80 = 16$
= 32 ✓

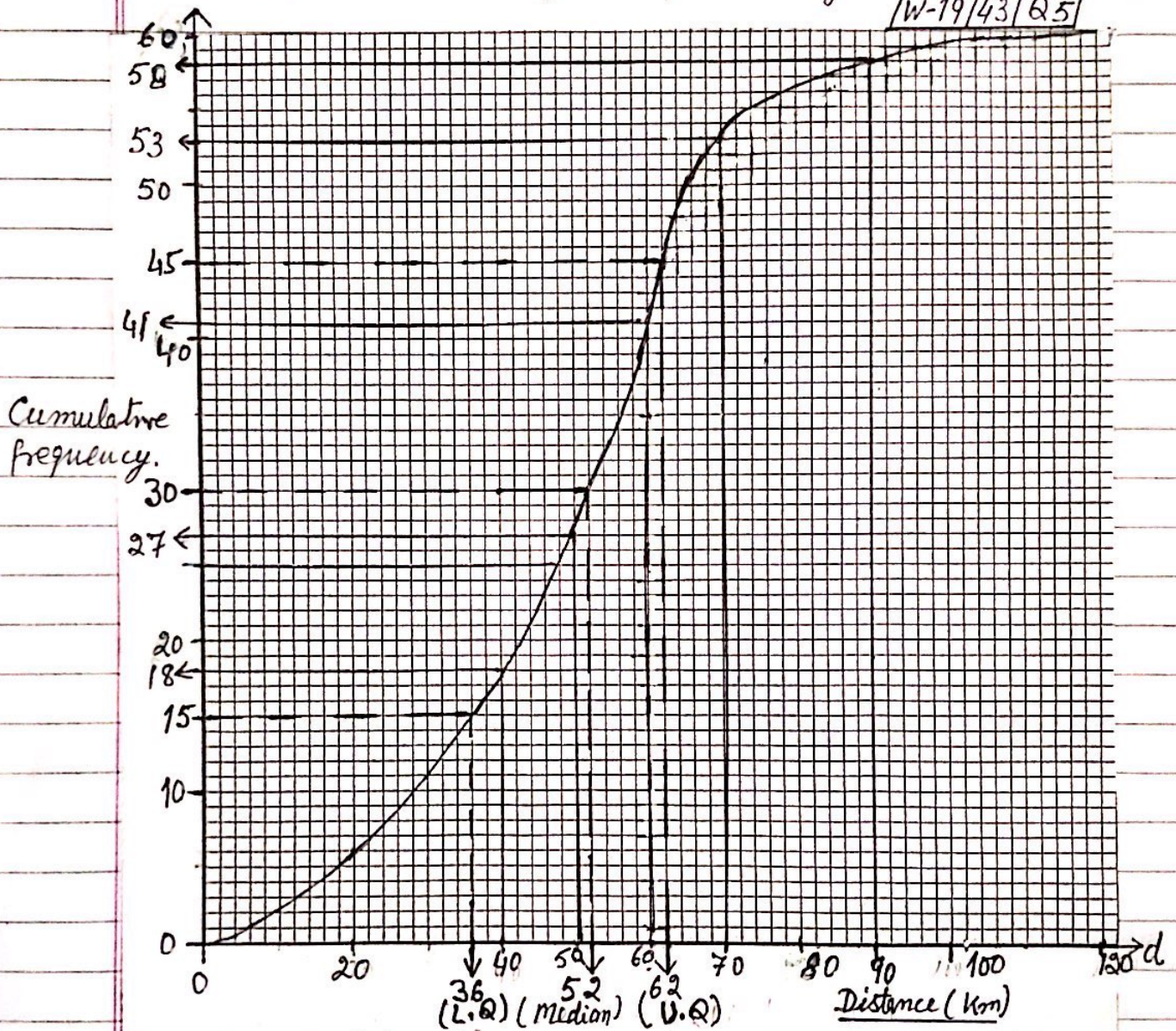
(iv) The number of girls who took more than 66 seconds to solve the problem = $80 - (\text{Number} \leq 66 \text{ seconds})$

= $80 - 62 = 18$ girls ✓

W-19/42/Q2(a)

19. The cumulative frequency diagram shows information about the distance, d km, travelled by each of 60 male cyclists in one weekend.

W-19/43/Q5



- (a) Use the cumulative frequency diagram to find an estimate of,
- (i) the median ...[1]
 - (ii) the lower quartile ...[1]
 - (iii) the interquartile range. ...[1]

Solution (i) Median = $\frac{60}{2} = 30$ = 52 ✓

(ii) Lower quartile = $\frac{1}{4} \times 60 = 15$ = 36 ✓

(iii) UQ = $\frac{3}{4} \times 60 = 45$ = 62 [I.R.R = UQ - LQ]

\therefore Interquartile range = $62 - 36 = 26$ ✓

(Continued →)

(Continued →)

19(b) For the same weekend, the interquartile range for the distances travelled for the distances travelled by a group of female cyclists is 40 km,

Make one comment comparing the distribution of the distances travelled by the males with the distribution of the distances travelled by the females. --- [1]

Solution: Distances for females are more varied.

(c) A male cyclist is chosen at random.
Find the probability that he travelled more than 50 km. --- [2]

Solution: C.f for 50 km, $\therefore (d \leq 50 \text{ km}) = 27$
 \therefore males travelled more than 50 km = $60 - 27 = 33$
 $\therefore P(\text{he travelled more than 50 km}) = \frac{33}{60} = \frac{11}{20} \checkmark$

(d) (i) Use the cumulative frequency diagram to complete this frequency table. --- [2]

Distance (d km)	No. of male cyclists C.f	f	Mid Value $\frac{x}{2}$	f · x
$0 < d \leq 40$	18	18	20	360
$40 < d \leq 50$	27	9	45	405
$50 < d \leq 60$	41	14	55	770
$60 < d \leq 70$	53	12	65	780
$70 < d \leq 90$	58	5	80	400
$90 < d \leq 120$	60	2	105	210
	$\Sigma f = 60$			$\Sigma f \cdot x = 2925$

(ii) Calculate an estimate of mean distance travelled. --- [4]

Solution: Estimate of mean = $\frac{\Sigma f \cdot x}{\Sigma f} = \frac{2925}{60} = \underline{48.7} \text{ km} \checkmark$