

Data

Solutions



Curriculum Ready



www.mathletics.com

Page 3 questions

1. The data below represents how many fish some fisherman caught in a week:

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

Complete this frequency and cumulative frequency table:

Score (x)	Frequency (f)	Cumulative Frequency (cf)
0	1	1
1	2	3
2	1	4
3	3	7
4	3	10
5	2	12
6	1	13
7	1	14
8	1	15
$\sum f = 15$		

2. This data was collected for the different ages of some random university students.

19 18 20 21 22 20 19 24 21 22 23 20 23 21 19

a Complete the table below and use it to answer the questions that follow:

Score (x)	Frequency (f)	Cumulative Frequency (cf)
18	1	1
19	3	4
20	3	7
21	3	10
22	2	12
23	2	14
24	1	15
$\sum f = 15$		

b How old was the youngest and oldest student?

The youngest student was 18 (the lowest score in the table)

The oldest student was 24 (the highest score in the table)

c How many students were 22 or younger?

12 were 22 or younger (the cumulative frequency for a score of 22)

Page 4 questions

3. Draw a frequency distribution table for the set of data: 3 6 5 2 1 8 4 3 5 3 4 7 0 4 1

Score (x)	Frequency (f)	Cumulative Frequency (cf)
0	1	1
1	2	3
2	1	4
3	3	7
4	3	10
5	2	12
6	1	13
7	1	14
8	1	15
$\sum f = 15$		

4. Complete the tables below:

a Use the frequency to complete the cumulative frequency column:

f	cf
2	2
3	5
1	6
2	8
3	11
2	13
1	14

b Use the cumulative frequency to complete the frequency column:

f	cf
3	3
1	4
2	6
2	8
2	10
3	13
3	16

Page 5 questions

5. Is it possible for:

a The cumulative frequency of a score to be **larger than** its frequency? Briefly explain.

Yes. The cumulative frequency is the sum of the preceding frequencies, so it must be at least as big as the score

b The cumulative frequency of a score to be **the same as** its frequency? Briefly explain.

Yes. The cumulative frequency will be the same as the frequency all rows up to first row where the score is not zero.

c The cumulative frequency of a score to be **smaller than** its frequency? Briefly explain.

No. The cumulative frequency of a score is the sum of the frequency, and all previous frequencies in the table. These frequencies are all positive integers or zero, so the sum can never be less than the frequency

Page 8 questions

1. Answer these questions in your own words.

- a What is the mean of a data set?

The mean is what most people mean by average. Think of a batting average, you add up all the elements of a data set (the scores), and divide by the number of elements in the data set (the number of innings).

- b What is the median of a data set?

The median is the middle value of a sorted data set. If there is an even number of members, calculate the average of the 2 middle scores.

- c What is the mode of a data set?

The mode is the most frequent member(s) of the data set, the score(s) that occurs the most number of times. If all scores have the same frequency, the data set has no mode.

- d What is the range of a data set?

The range of the data set is the difference between the largest and the smallest members of a data set.

- e What do $\sum f$ and $\sum fx$ mean?

x refers to the members of the data set

f refers to the frequency of a member of the data set

\sum means 'add up all of the'

so $\sum f$ means add up all of the frequencies in a data set

$\sum fx$ means multiply each value in the data set by its frequency then add all of these up.

Page 9 questions

2. A teacher records the percentages that a group of students achieved in a test.

75 80 62 65 71 91 88 55 48 42 63 80 78 66 52

- a Find \bar{x} , the mean of the data set

First build a frequency table

Score (x)	Frequency (f)	fx	Cumulative Frequency (cf)
42	1	$42 \times 1 = 42$	1
48	1	$48 \times 1 = 48$	2
52	1	$52 \times 1 = 52$	3
55	1	$55 \times 1 = 55$	4
62	1	$62 \times 1 = 62$	5
63	1	$63 \times 1 = 63$	6
65	1	$65 \times 1 = 65$	7
66	1	$66 \times 1 = 66$	8
71	1	$71 \times 1 = 71$	9
75	1	$75 \times 1 = 75$	10
78	1	$78 \times 1 = 78$	11
80	2	$80 \times 2 = 160$	13
88	1	$88 \times 1 = 88$	14
91	1	$91 \times 1 = 91$	15
$n = \sum f = 15$		$\sum fx = 1016$	

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{1016}{15} = 67.7\dot{3}$$

- b Arrange the data set in ascending order.

42 48 52 55 62 63 65 66 71 75 78 80 80 88 91

- c Find the median of the data set.

The median is the 'middle' value of the sorted data set. There are 15 values. The middle value is then therefore the eighth one, 66.

- d Find the mode(s) of the data set.

The mode is the most frequent (common) value in the data set. 2 is the largest value in the frequency column above, and the mode is the corresponding f value, 80.

Page 10 questions

3. Complete the following frequency distribution table.

Score (x)	Frequency (f)	fx	Cumulative Frequency (cf)
11	10	$10 \times 11 = 110$	10
14	11	$11 \times 14 = 154$	21
16	9	$9 \times 16 = 144$	30
17	7	$7 \times 17 = 119$	37
19	11	$11 \times 19 = 209$	48
20	15	$15 \times 20 = 300$	63
21	8	$8 \times 21 = 168$	71
		$n = \sum f = 71$	$\sum fx = 1204$

a What is the range of this data set?

The range of the data set is the difference between the largest (21) and the smallest (11) members of the data set. The range is $21 - 11 = 10$

b Find \bar{x}

$$\begin{aligned}\bar{x} &= \frac{\sum fx}{\sum f} \\ &= \frac{1204}{71} \\ &= 16.96\end{aligned}$$

c How many scores are in the data set?

The last cumulative frequency value or the $\sum f$ value give the total number of scores, 71

d In which position is the median score and what is the median?

The median is in the middle position. There are 71 positions so the middle score is in position 36. Using the cumulative frequency column shows that scores of 17 are in positions 31 to 37 inclusive. The median therefore is 17.

e What is the mode of the data set?

The mode is the most frequent score(s). By inspecting the frequency table, we can see 20 occurs the most (15 times) so 20 is the mode.

Page 11 questions

4. Look at this data set: 5 3 8 4 8 1 7 9 0 5

a Find \bar{x}

$$\begin{aligned}\bar{x} &= \frac{\text{Sum of Scores}}{\text{Number of scores}} \\ &= \frac{5 + 3 + 8 + 4 + 8 + 1 + 7 + 9 + 0 + 5}{10} \\ &= \frac{50}{10} \\ &= 5\end{aligned}$$

b If an 11th score of 6 was added to the data set, is the new mean greater than, equal to or less than the original mean? Why?

The new member of the data set is larger than the original mean, so the new mean will be larger. $\frac{56}{11} = 5.\dot{0}\dot{9} > 5$

c If an 11th score of 5 was added to the data set, is the new mean greater than, equal to or less than the original mean? Why?

The new value is equal to the original mean, so the new mean remains the same. $\frac{55}{11} = 5$.

d If an 11th score of 4 was added to the data set, is the new mean greater than, equal to or less than the original mean? Why?

The new value is smaller than the original mean, so it makes the new mean smaller also. $\frac{54}{11} = 4.\dot{9}\dot{0} < 5$

Page 12 questions

5. This data set shows how many points a team scored each week over a 4 week period: 210 320 100 220. How many points do they need to score in the 5th week to ensure that the mean score over 5 weeks is 200 points?

Let A be the score for the 5th week.

$$\begin{aligned}\bar{x} &= 200 \\ \frac{\text{Sum of Scores}}{\text{Number of scores}} &= 200 \\ \frac{210 + 320 + 100 + 220 + A}{5} &= 200 \\ \frac{850 + A}{5} &= 200 \\ 850 + A &= 1000 \\ A &= 150\end{aligned}$$

The team must score 150 in the 5th week to have a mean of 200 over the 5 week period.

Page 12 questions

6. How many scores are in a data set if:

- a The median is the 17th score in an ordered data set?

In a data set that has been sorted from lowest value to highest value, there will 16 values before the median, 16 after, and the median itself, for a total number of $16 + 1 + 16 = 33$ scores in the data set.

- b The median is the average of the 20th and 21st score in the data set?

This data set has an even number of members. 20 numbered from 1 to 20, and 20 more numbered from 21 to 40, for a total of 40 members.

Page 16 questions

1. A single die was rolled 15 times and these were the results: 1 6 4 2 4 1 5 3 3 2 6 5 4 3 1

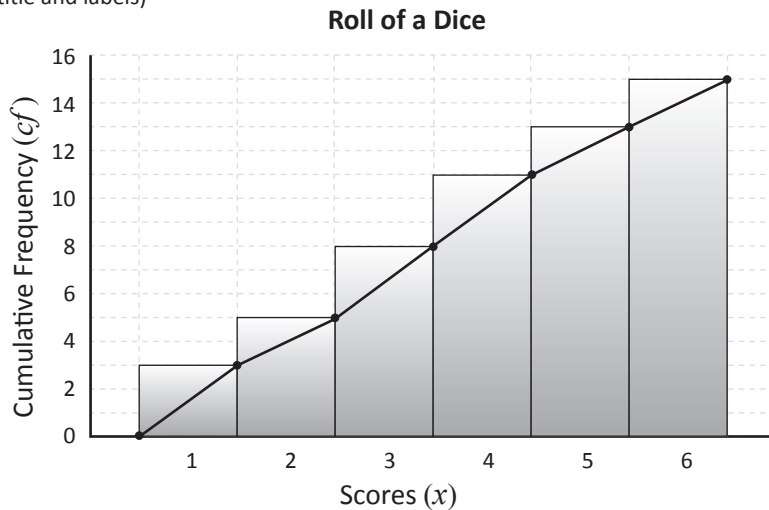
a Complete this table for the data:

Score (x)	Frequency (f)	Cumulative Frequency (cf)
1	3	3
2	2	5
3	3	8
4	3	11
5	2	13
6	2	15
$\sum f = 15$		

b Use the table to complete a frequency histogram and polygon on the same axes below.
(hint: don't forget the title and labels)



c Use the table to draw a histogram for cumulative frequency and an ogive on the same axes.
(hint: don't forget the title and labels)

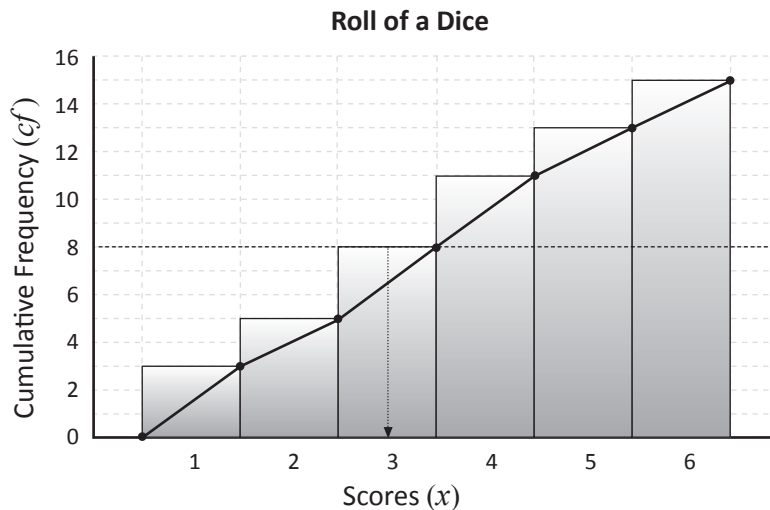


Page 17 questions

- d There are 15 scores in the data set. In which position of the ordered data set is the median?

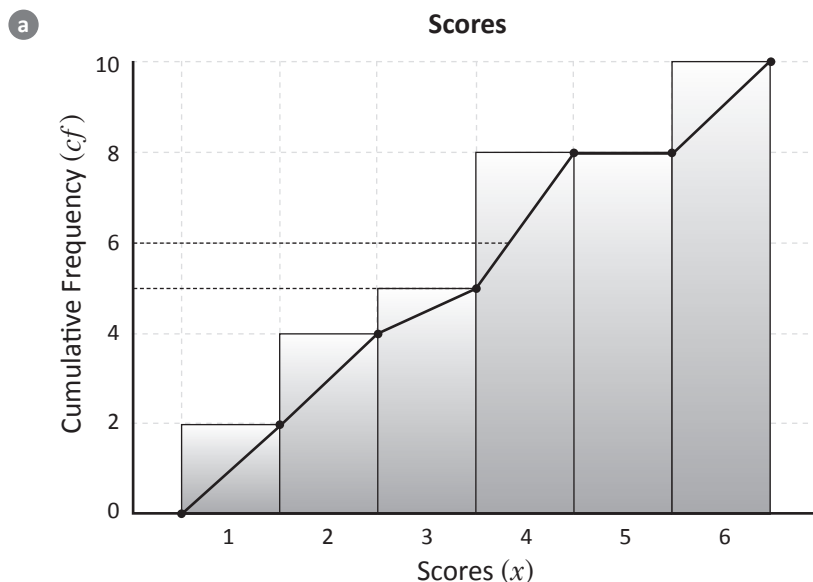
8th position. There are 7 before, 7 after and the median make a total of 15.

- e Use the diagram in c to find the median of this data set.



The median of this data set is 3.

2. Draw in the ogive in each of the following and use it to find the median



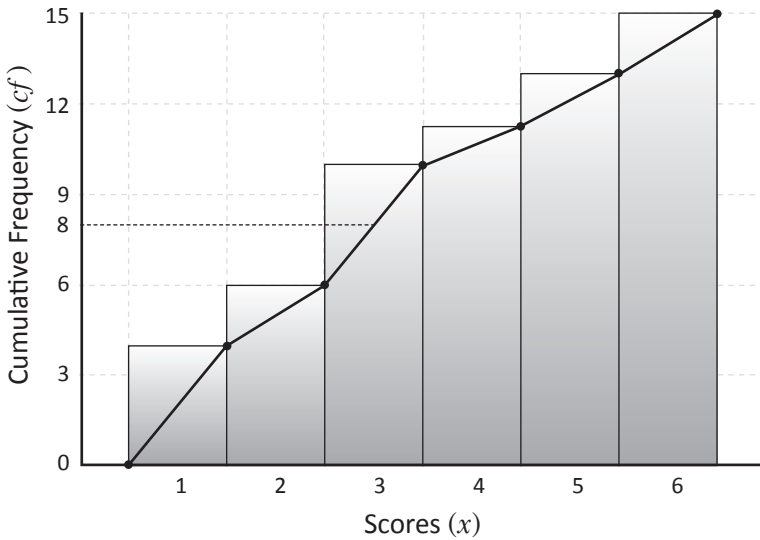
The cumulative frequency shows there are 10 scores in total. As this is even the median is the average of the 5th and 6th scores. The intersection of the ogive and the horizontal line from $y = 5$ shows the 5th score is 3. The intersection of the ogive and the horizontal line from 6 show the 6th score is 4. Calculating the average of the scores to find the median.

$$\begin{aligned} \text{median} &= \frac{5^{\text{th}} \text{ score} + 6^{\text{th}} \text{ score}}{2} \\ &= \frac{3 + 4}{2} \\ &= \frac{7}{2} \\ &= 3\frac{1}{2} \end{aligned}$$

Page 17 questions

b

Scores

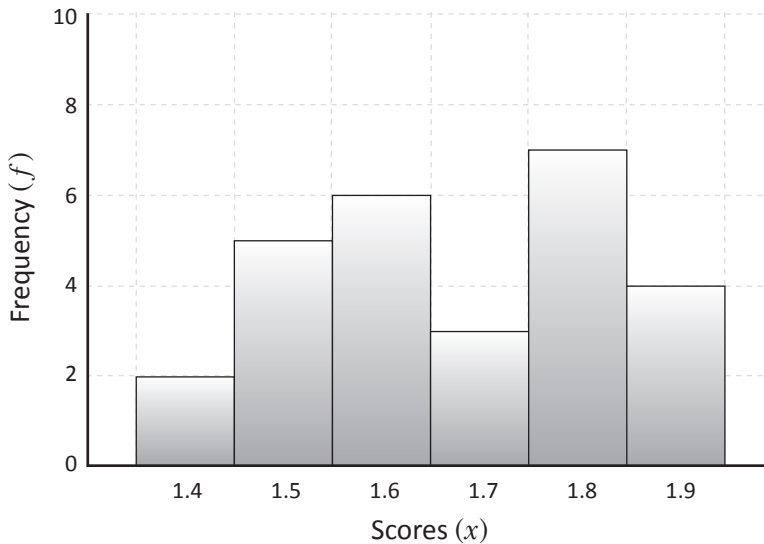


The cumulative frequency shows there are 15 scores in total. The median is the middle 8th score. The intersection of the ogive and the horizontal line from 8 occur in the column for a score of 3, so 3 is the median.

Page 18 questions

3. The frequency histogram of the different heights (in m) of a group of students is represented below.

Student Height Measurement



a How many students were measured for their height? (How many scores are in the data set?)

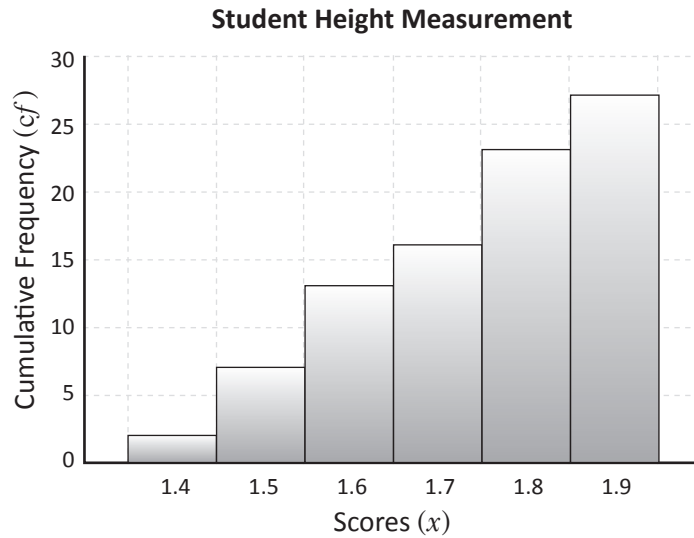
Read the frequency values for each height from the frequency graph, and add to find the total number in the data set.

2 at 1.4m, 5 at 1.5m, 6 at 1.6m, 3 at 1.7m, 7 at 1.8m and 4 at 1.9m

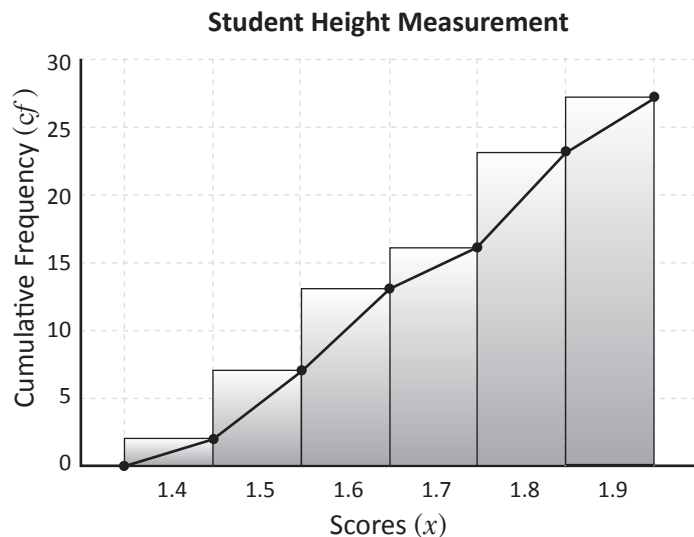
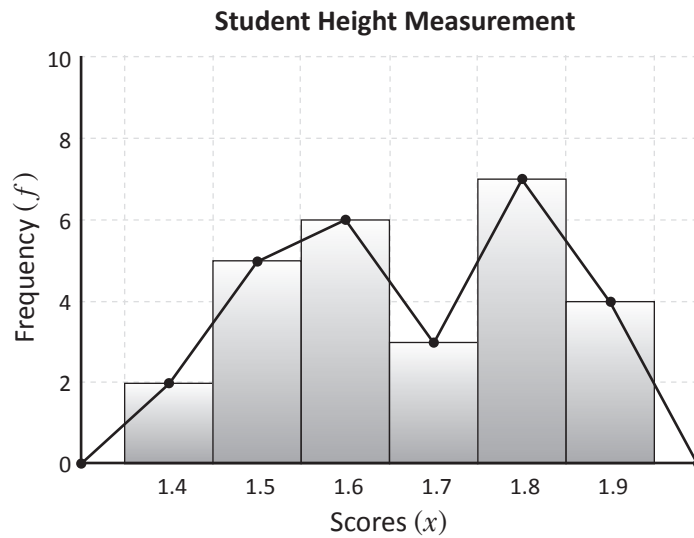
$2 + 5 + 6 + 3 + 7 + 4 = 27$ student height measurements in the data set

Page 18 questions

b Complete the cumulative frequency histogram on the other set of axes.



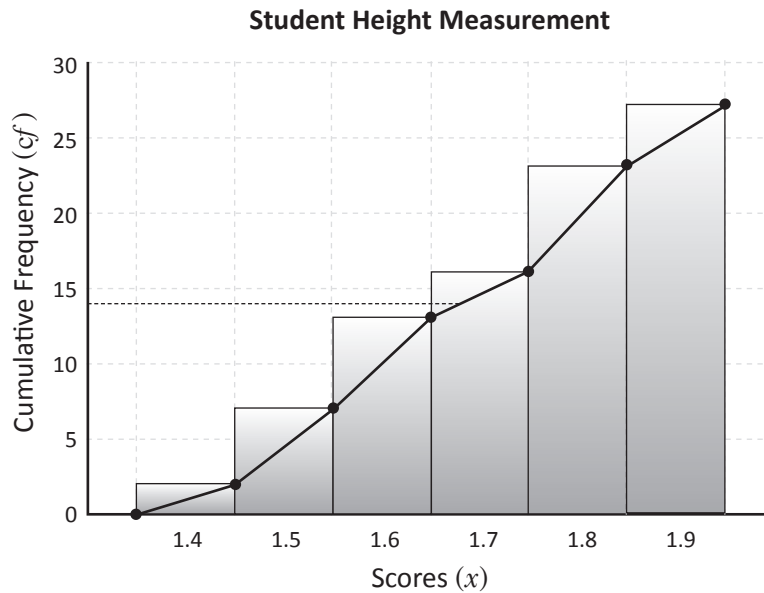
c Complete the polygon in each diagram.



Page 18 questions

- d What is the median of this set of data?

We know from a that there are 27 student heights in the data set, so the median will be in the 14th position. (13 before, 13 after and 1 in the middle make 27).



The horizontal line from $y = 14$ intersects with the ogive in the column for a height of 1.7m. The median of this set of data is therefore a height of 1.7m.

This means, if you arranged these 27 students from shortest to tallest, the middle student would be 14th in the line, and would be 1.7m tall.

Page 22 questions

1. When a frequency polygon is drawn, it returns to 0 at the end, but this does not happen with cumulative frequency polygons. What are some other differences we must remember when drawing these diagrams?

- The frequency polygon passes through the centre of the top of each column. The cumulative frequency polygon passes through the top right corner of each column.
- The frequency polygon starts half a column to the left of the first column, and finishes half a column to the right of the last column. The cumulative frequency polygon starts at the bottom left of the first column and ends at the top right corner of the last column.

2. Explain the difference between the following:

- a 'Class' and 'Class Centre'.

A 'Class' is convenient grouping of scores in a data set, say ages from 21–30. The 'Class Centre' is a value that represents the middle of a 'Class'. For the 21–30 class, the class centre is $\frac{21 + 30}{2} = 25.5$

- b 'Mode' and 'Modal Class'.

The mode of a data set is the score(s) with the highest frequency. The Modal Class is the class that has the highest frequency.

- c 'Frequency' and 'Cumulative Frequency'.

The frequency of a particular score is the number of times that the score occurs in a data set. The cumulative frequency of a score is the number of times that the score or any score smaller than it appears in the data set.

Page 23 questions

3. A magician asked sixty people to pick a number from 1 to 92.

3	56	38	54	23	52	87	66	55	42	29	39	24	74	60
58	61	18	29	51	37	31	78	54	27	73	62	41	25	32
52	48	6	17	23	47	51	88	16	5	12	36	58	81	74
36	59	41	75	63	25	1	78	90	47	53	65	26	20	15

- a List 4 classes to group these scores into.

There are 92 possible values so each class should contain about $\frac{92}{4} = 23$ values.

The 4 classes would be [1–23] (23 values), [24–46] (23 values), [47–69] (23 values), [70–92] (23 values)

Page 23 questions

- b Fill these classes in the first column of the table below:

Class	Class Centre (x)	Frequency (f)	fx	Cumulative Frequency (cf)
1–23				
24–46				
47–69				
70–92				
		$\sum f =$	$\sum fx =$	

- c What is the meaning of ' fx ' in the 4th column?

This means multiply the frequency by the class centre of the class.

Page 24 questions

- d Find the class centre of each class and fill these in on the table in the second column:

Class	Class Centre (x)	Frequency (f)	fx	Cumulative Frequency (cf)
1–23	$\frac{(1 + 23)}{2} = 12$			
24–46	$\frac{(24 + 46)}{2} = 35$			
47–69	$\frac{(47 + 69)}{2} = 58$			
70–92	$\frac{(70 + 92)}{2} = 81$			
		$\sum f =$	$\sum fx =$	

- e Fill in the rest of the table:

Class	Class Centre (x)	Frequency (f)	fx	Cumulative Frequency (cf)
1–23	$\frac{(1 + 23)}{2} = 12$	12	144	12
24–46	$\frac{(24 + 46)}{2} = 35$	17	595	29
47–69	$\frac{(47 + 69)}{2} = 58$	21	1218	50
70–92	$\frac{(70 + 92)}{2} = 81$	10	810	60
		$\sum f = 60$	$\sum fx = 2767$	

Page 24 questions

- f Estimate the mean. (How would you find the exact value of the mean using the original scores?)

An estimate of the mean can be found with the formula:

$$\begin{aligned}\bar{x} &= \frac{\sum fx}{\sum f} \\ &= \frac{2767}{60} \\ &= 46.11\dot{6}\end{aligned}$$

(The exact mean is found by adding all the values in the data set and dividing by 60 $\frac{2687}{60} = 44.78\dot{3}$)

- g Which class is the modal class? Why?

The modal class is [47–69] because this class has the highest frequency (21) of any class.

- h Which class is the median class? Why?

The median class is [47–69] because this class contains the 30th and 31st values.

Page 25 questions

4. During a recent game of darts some scores were recorded.

a

Class	Class Centre (x)	Frequency (f)	fx	cf
1–5	3	12	36	12
6–10	8	17	136	29
11–15	13	9	117	38
16–20	18	5	90	43
21–25	23	8	184	51
		$\sum f = 51$	$\sum fx = 563$	

- b Estimate the mean

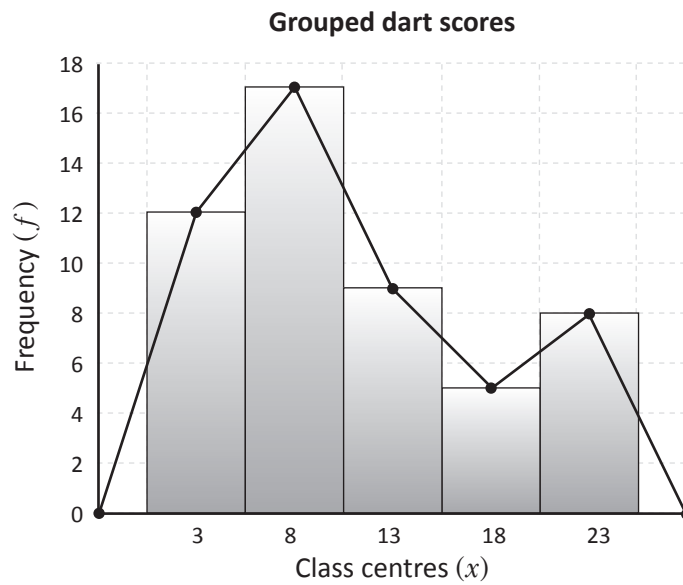
$$\begin{aligned}\bar{x} &= \frac{\sum fx}{\sum f} \\ &= \frac{563}{51} \\ &\approx 11.04\end{aligned}$$

Page 25 questions

c Which is the modal class?

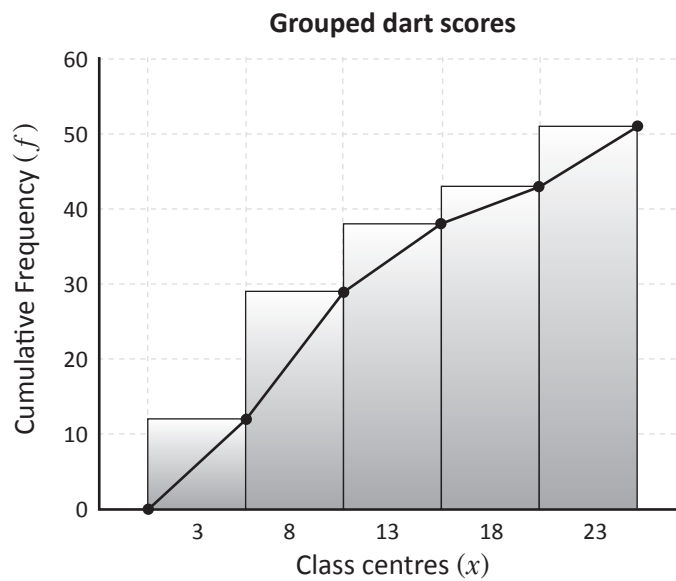
[6–10] is the modal class, as it has the highest frequency (17)

d Draw a frequency histogram and polygon on the same axes below (remember to fill in the class centres):



Page 26 questions

e Complete a cumulative frequency histogram and polygon on the axes below:



Page 26 questions

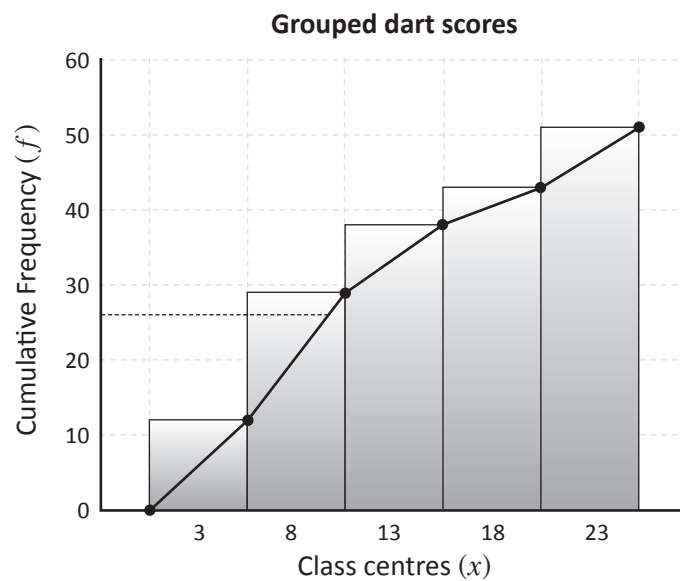
f What is the total number of scores?

The total number of scores is 51, $\sum f$ or the last value in the cumulative histogram.

g Which position will the median be placed in the ordered set of scores?

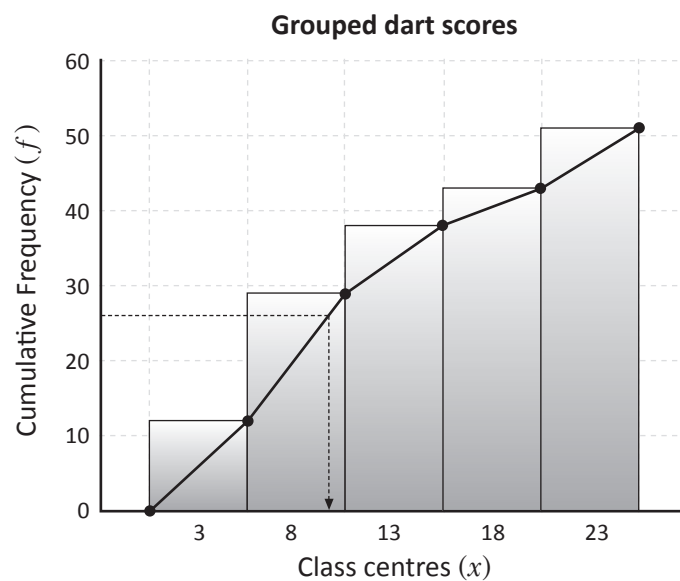
The median is at position 26.

h Use the cumulative frequency polygon to determine which is the median class.



The horizontal line from $y = 26$ intersects the ogive in the second column, so the median class is the second, 6 – 10.

i Use the diagram above to estimate the median.



The arrowhead on the vertical line from the intersection of the ogive and the horizontal line above touches the x -axis between 8 and 13, but closer to the 8 than the 13. 9 or 10 would be good estimates for the median.



www.mathletics.com