

CUMULATIVE FREQUENCY DIAGRAMS

When working with continuous grouped data we estimate the different averages as the individual data values are not known, only the class interval they fall into. A cumulative frequency diagram is a good way of estimating one of those averages the median and also the interquartile range.

Check first that you:

- understand the difference between the averages as a measure of central tendency and the range as a measure of the spread of the data
- understand how data is grouped into class intervals
- can find averages of grouped data
- can read scales on a graph
- can plot points on a graph
- can find $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ of a number.

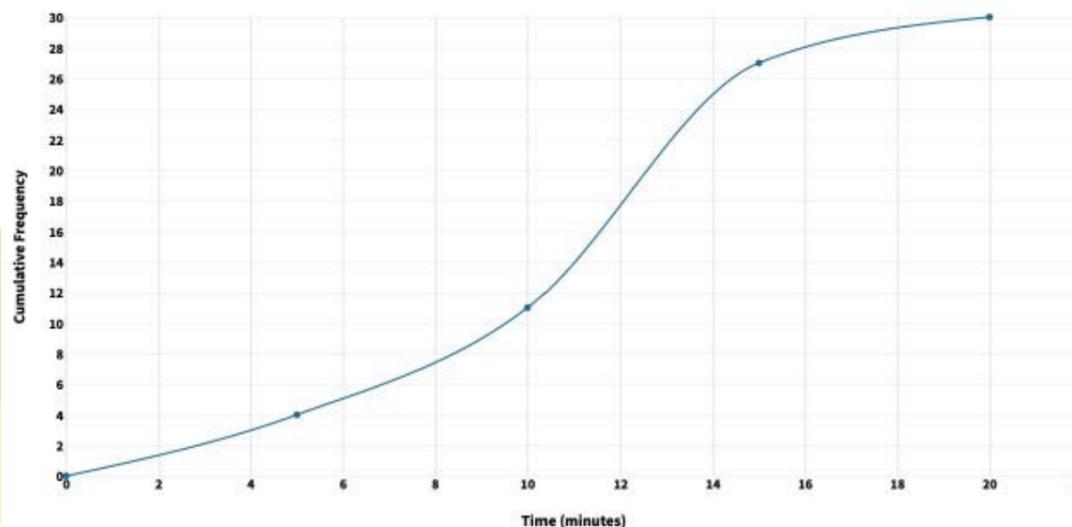
Drawing a Cumulative Frequency Diagram

E.g. The frequency table shows the length of time in minutes it took pupils to complete a puzzle. Draw a cumulative frequency diagram for this data.

Time (minutes)	$0 < t \leq 5$	$5 < t \leq 10$	$10 < t \leq 15$	$15 < t \leq 20$
Frequency	4	7	16	3

We draw a cumulative frequency diagram by plotting the upper boundary of each class against the cumulative frequency therefore we use the information in the table to create a cumulative frequency table.

Time (minutes)	≤ 0	≤ 5	≤ 10	≤ 15	≤ 20
Cumulative frequency	0	$0 + 4 = 4$	$4 + 7 = 11$	$11 + 16 = 27$	$27 + 3 = 30$



In this example time should be on the horizontal axis and the cumulative frequency should be on the vertical axis.

Plot each upper group boundary against the cumulative frequency ensuring that the cumulative frequency of the last point is equal to the total frequency.

Connect the points using a ruler or by free hand to get a smooth curve.

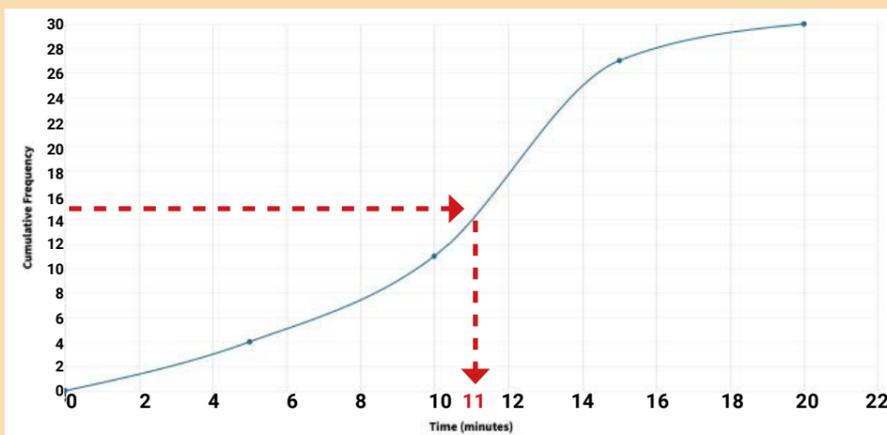
Using a Cumulative Frequency Diagram

We use a cumulative frequency diagram to find the following:

Median

This is the middle value of the data. It is located at $\frac{1}{2}$ of the total frequency.

To find the median draw a straight line from this $\frac{1}{2}$ way value on the vertical axis (cumulative frequency) across to the curve. Where it meets the curve, draw a line down to the horizontal axis and read off the value. In the diagram below, the total frequency is 30 so we draw a line from 15 on the vertical axis to the curve.



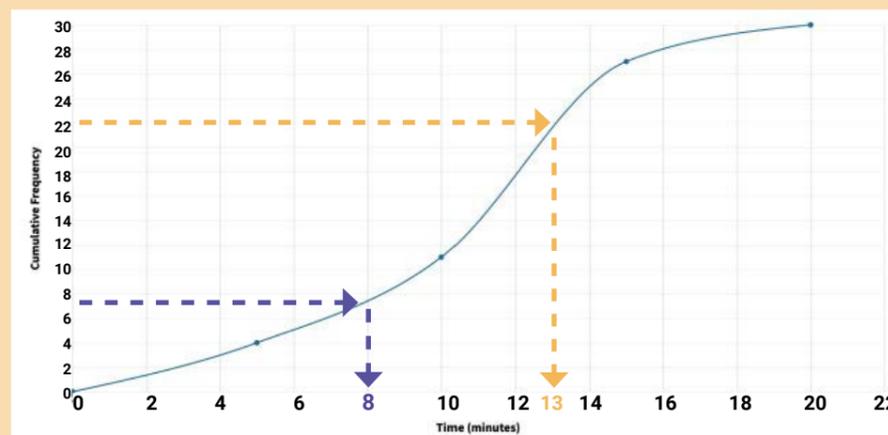
We can see that the median length of time it took to solve the puzzle was **11** minutes to the nearest minute.

Interquartile Range

This is a measure of the spread of the middle 50% of the data.

Interquartile Range = Upper Quartile – Lower Quartile.

The lower quartile is found at a $\frac{1}{4}$ of the total frequency and the upper quartile is found at $\frac{3}{4}$ of the total frequency. In the diagram below the total frequency is 30 so we draw a line from 7.5 (LQ) and 22.5 (UQ) on the vertical axis, to the curve and then draw a line down to the horizontal axis to read off the two values.



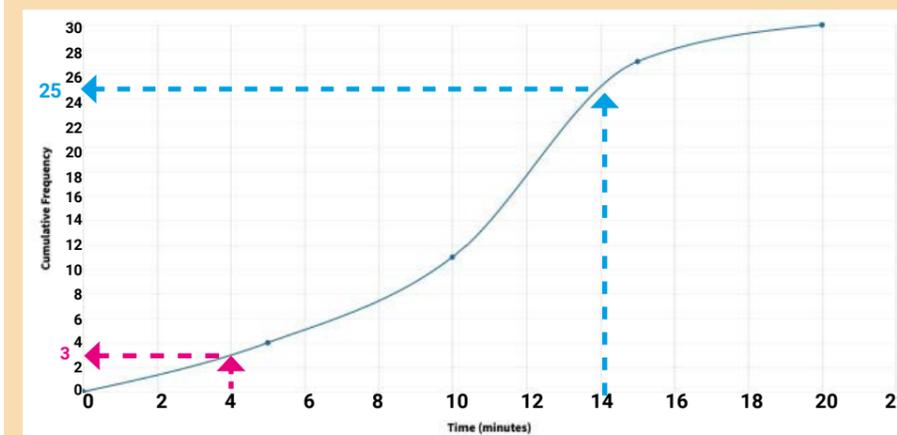
We can see that the LQ is **8** and the UQ is **13** therefore the interquartile range is **13 – 8 = 5** minutes.

Less than and more than values

In this example, if we wanted to find the number of pupils that took

- less than 4 minutes,
- more than 14 minutes to complete the puzzle

we draw lines from the horizontal axis (time) to the curve and read off the values on the vertical axis to get the cumulative frequency. The diagram always shows values for less than therefore to obtain a more than value, the less than value needs to be subtracted from the total frequency.



- 3** students
- $30 - 25 = 5$ students

You may be asked to use cumulative frequency diagrams to compare two sets of data, **remember** to compare the median values for each set and also the interquartile range.

Remember a smaller interquartile range shows more consistent data and a larger interquartile range less consistent data.