Build your skills: Interpreting data – Part 1

This task has three parts to it.

Part 1

In this part you will find information and activities to help you work more confidently when you need to extract, interpret and make use of information from tables, graphs and charts, or to work out the median or mean for a set of data.

HD1/ L2.1
L2.3
tutor notes

Part 2

In Part 2 you can find suggestions of other free resources you can use to practise your skills.

Part 3

In Part 3 you can try out your skills on some practical activities and check your progress on some typical questions from the Level 2 National Test. Part 3 also contains the answers to all the activities in Parts 1–3.

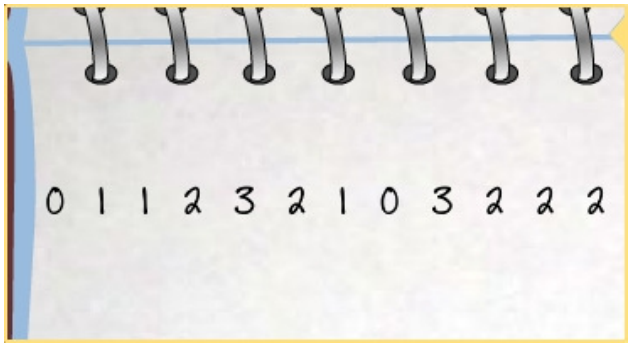
Thinking about interpreting data

When you collect information about a particular item in which you are interested, the information collected is called a set of '*data*' (numbers). Each of the numbers in the set of data is called a '*value'.*

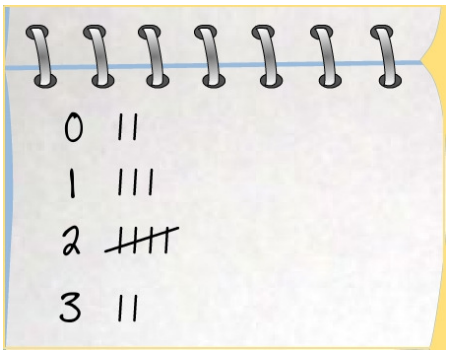
Each set of data can be represented in a variety of different ways.

Example:

We might want to know about the number of cars owned by different households in a particular street. If we go around the street to collect the information and ask each household how many cars they own, we might get a *set of data* that looks like this:



Or we might have collected the data using a tally as we go around. So it would look like this:

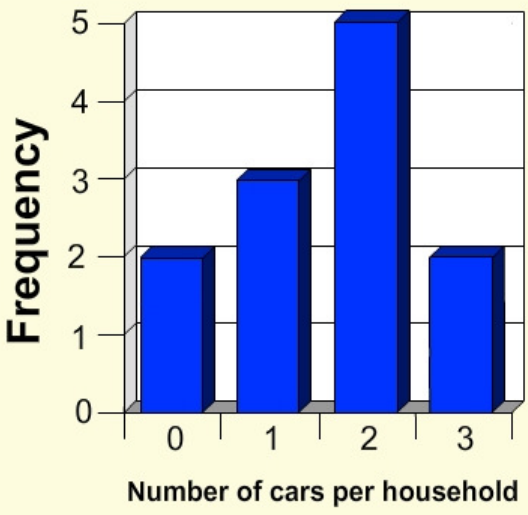


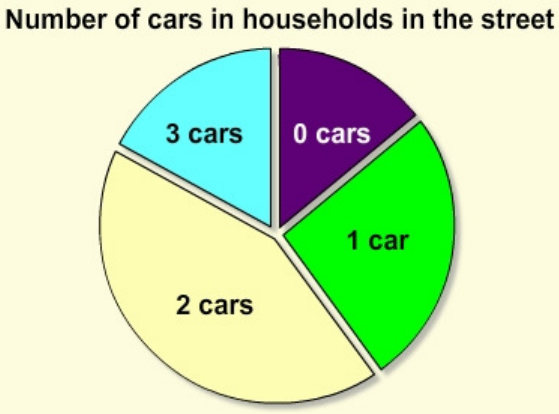
This data might later be presented in a table:

Number of cars owned	0	1	2	3
Number of households	2	3	5	2

The bottom row of numbers is sometimes called the *'frequency'*, i.e. how many times does this value appear in the data set?

. . . or as a chart or graph:

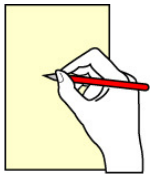




Build your skills: Interpreting data – Part 1

When interpreting data from graphs and charts, it is important to look carefully at the information you are given, in particular:

- the title(s) – this tells you what information is shown on the chart
- any key or other information provided (activity 1 gives examples of this)
- the scale used – (activities 2 and 3 will look at this in more detail.).



Activity 1

Example:

This pie chart shows the make-up of the members of a particular club. It shows the proportion of the members that are men, women, boys and girls respectively. So, from looking at the chart, we can say things like:

'About a quarter of the members are men.' ⟵ *The (blue) wedge that represents the 'men' is about ¼ of the whole chart.*

or

'There are about twice as many girls as boys.' ⟵ *The (cream) wedge representing the 'girls' is about twice the size of the (turquoise) wedge showing the 'boys'.*

Without additional information, we do not know *how many* there are of each type of member because the chart (as it is) does not tell us anything about the total number of

members. However, if we know something about the size of the membership, we can begin to work out absolute numbers for different types of member.

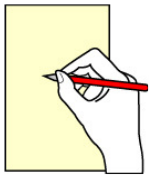
For example:

If the total number of members is 200 ← *The number of 'men' will be about 50 (¼ of 200).*

or

If the number of girls is 200 ← *The number of 'boys' will be about 100 (½ of 200: the number of 'girls').*

For each of the following, fill in the table.



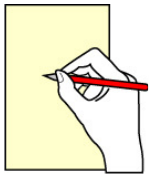
1 If the **total number of members is 200** (as in the first example above), work out the approximate number of each type of member and fill in the table below showing this data.

The number of men has already been filled in for you.

men	women	boys	girls
50			

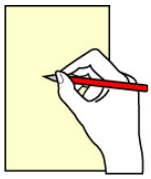
***Tip**: You can work this out in whatever way makes sense to you when you look at the chart. If you are not sure, however, it might help you to think about:*

- *How does the number of women compare with the number of men?*
- *What proportion of the total number of members are either girls or boys?*
- *Roughly what fraction do the boys make up of the left half of the chart (which shows the boys and girls together)?*



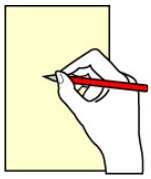
2 If the **number of girls is 200** (as in the second example above), fill in the table below showing the approximate number of each type of member.

men	women	boys	girls
			200



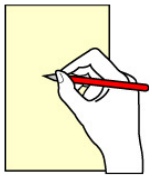
3 If the **total number of members is 48**, fill in the table below showing the approximate number of each type of member.

men	women	boys	girls



4 If the **number of boys is 25**, fill in the table below showing the approximate number of each type of member.

men	women	boys	girls



Activity 2

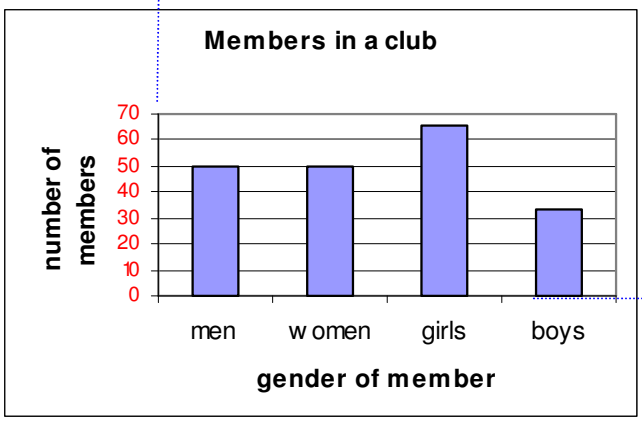
As mentioned above, it is important to look carefully at charts or graphs and notice what the different axes show and which scale is used on each axis.

Example:

Let's look at the data from question 1 of Activity 1 in a bar chart format.

The line going up the chart is called the *'y axis'*.

The *'scale'* is the numbers used on the axis to show the data (the intervals between these numbers).



The line going across the chart is called the *'x axis'*.

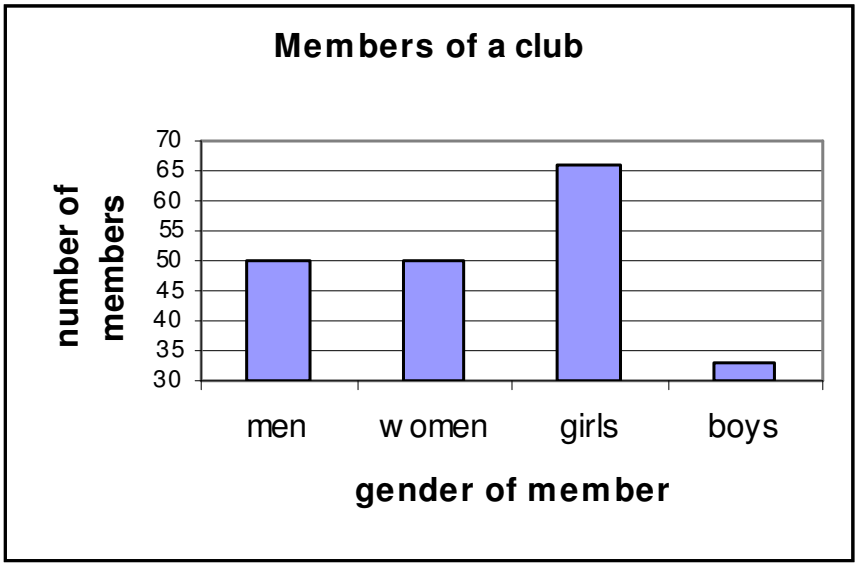
Look at the bar chart on the previous page and check that you would feel confident to read off the number of members in each category from it.

***Note**: This bar chart makes it easy to see that:*

- *the numbers of men and women are the same*
- *there are fewer boys and more girls.*

But it is harder to see the relative proportions of each type of member in the overall membership.

What if the data was shown like this:



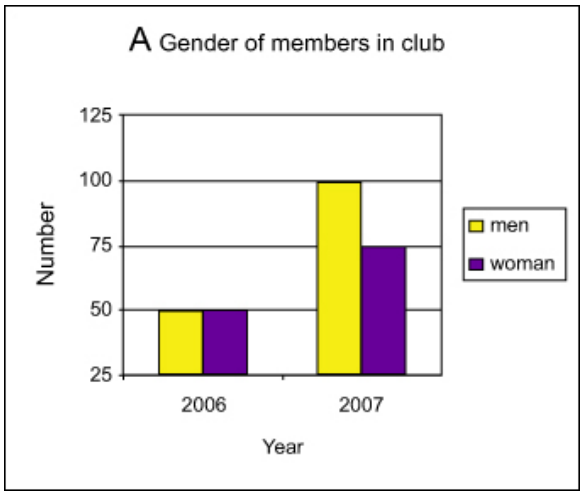
This bar chart makes it *look like* there are relatively *very few boys*, compared with lots of girls, but notice the scale on the y axis:

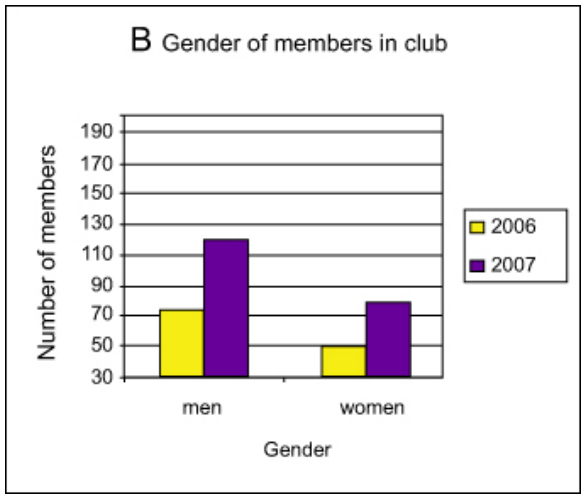
The scale does not start at 0.

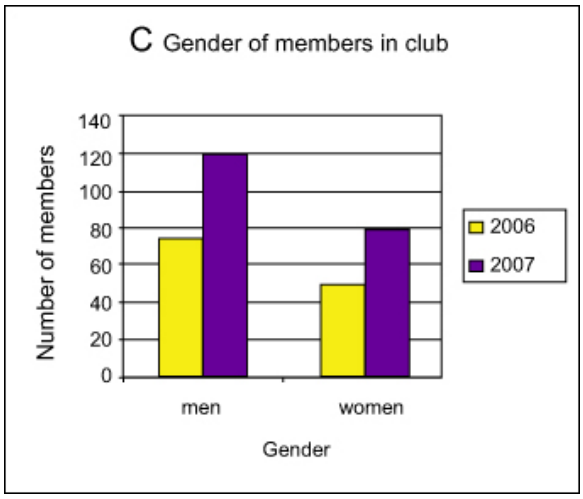
So although this chart shows the same data as the previous one, it gives quite a different impression. It is important, therefore, to look carefully at the scale used on graphs and charts to make sure that you interpret the information shown correctly.

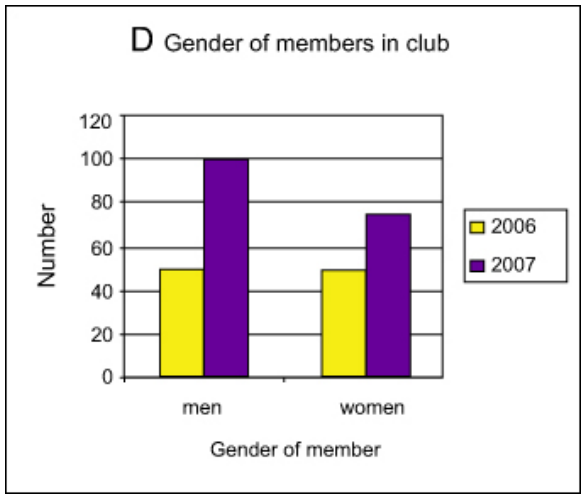
Build your skills: Interpreting data – Part 1

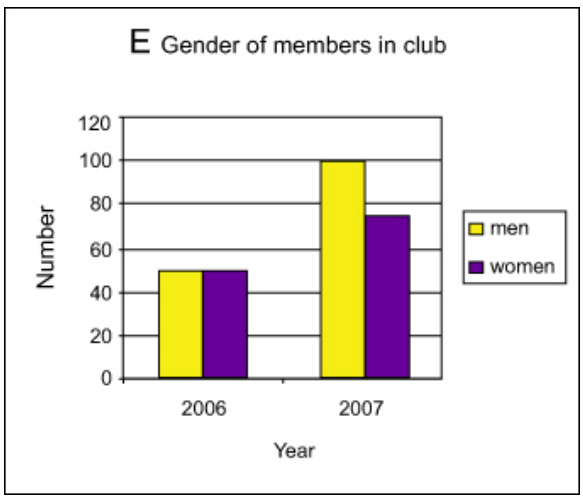
Look at these charts and identify which of these charts show the same data.

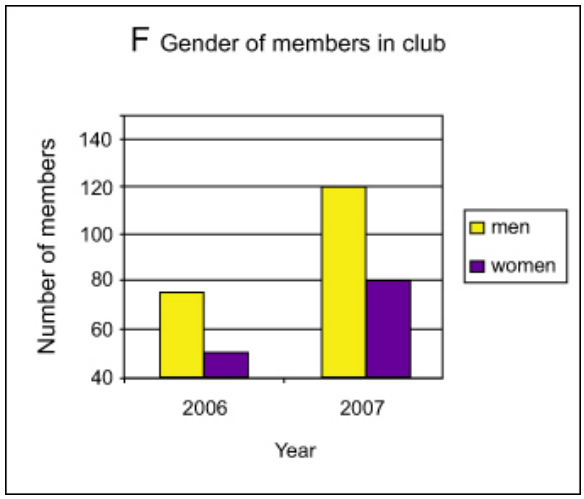


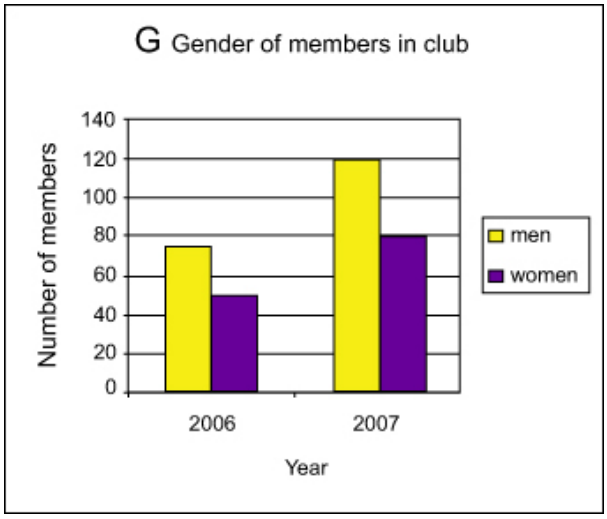


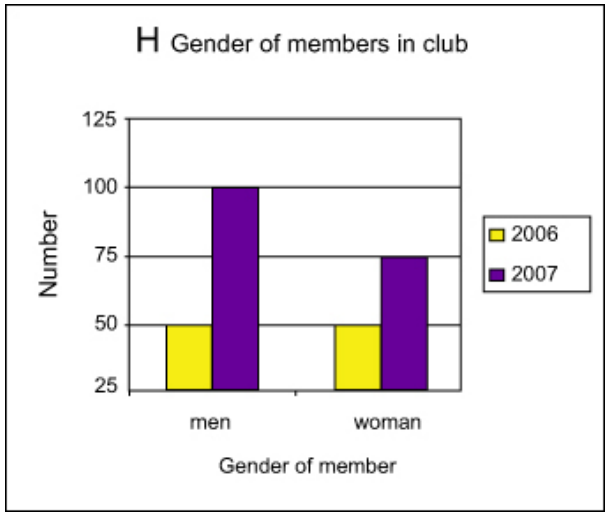


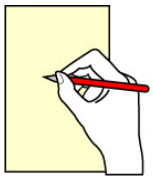












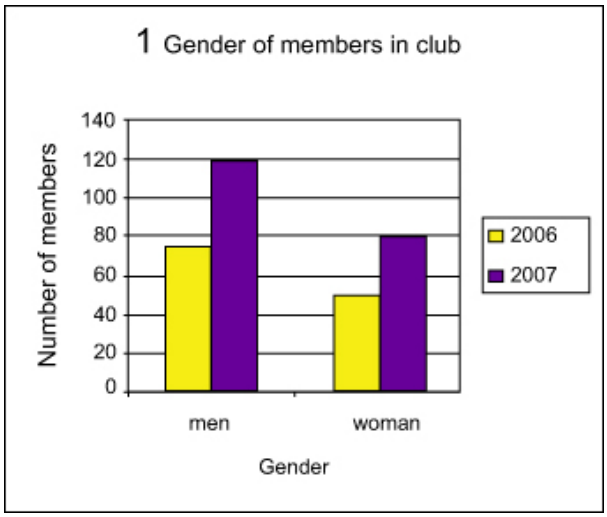
Activity 3

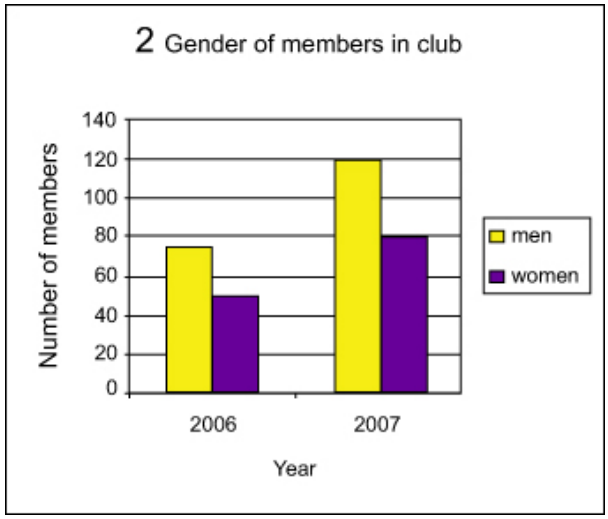
Sometimes questions ask you to compare data and draw conclusions or summarise what the data tells you.

Example:

Look at the two bar charts below.

We are going to think about what conclusions you can draw from the charts.





Which chart would you find it easiest to use to draw each of these conclusions?

1 The number of male and female members both increased between 2006 and 2007.

2 There were more male than female members in both years.

3 The total membership increased from 2006 to 2007.

As both bar charts show the same data set, you could get this information from either chart. However:

Conclusion 1 *– Chart 1 makes it easy to compare the number of both male and female members from year to year.*

Conclusion 2 *– Chart 2 makes it easy to compare the relative numbers of men and women in each year.*

Conclusion 3 *– To find the total membership for each year you would need to work this out. You could do this using whichever chart you prefer.*

Which of these is NOT true?

4 The number of women members was about the same proportion of the total membership in both years.

5 The percentage increase in both male and female members was about the same from 2006 to 2007.

6 The total number of members in 2007 was double the membership in 2006.

Conclusion 4 *– On Chart 1 you can see that the number of women members is about two thirds the number of men in both years, so this statement is* ***TRUE.***

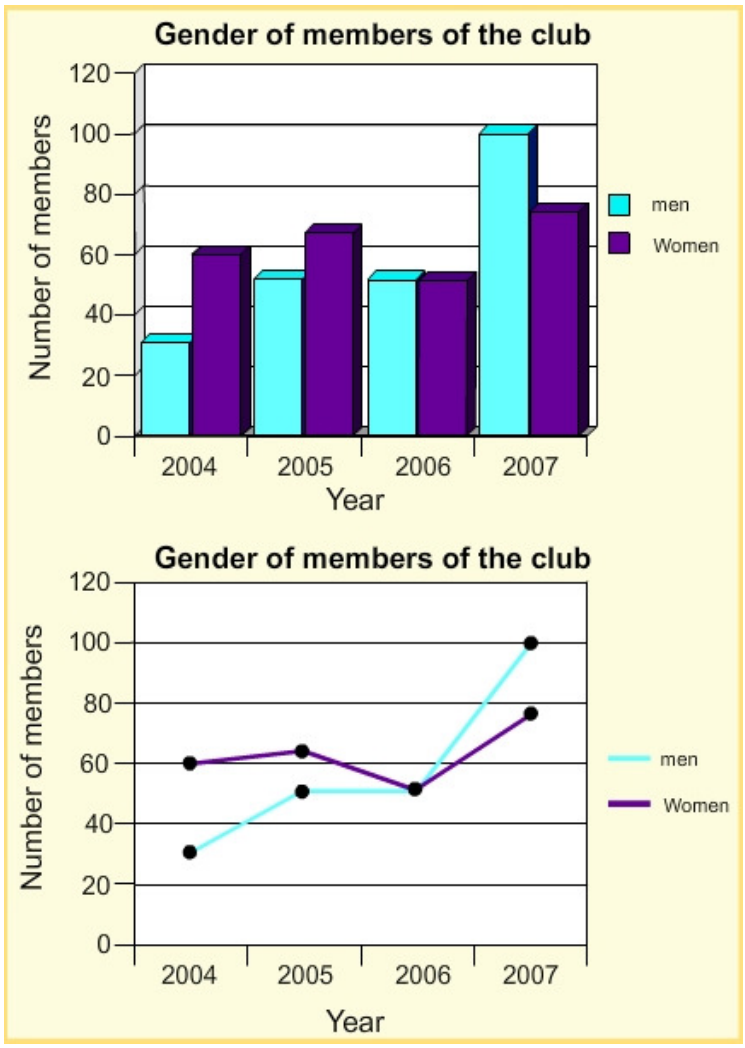
Conclusion 5 *– On Chart 2 you can see that the numbers of men and women both increased by about 50% – so this statement is also* ***TRUE.***

Conclusion 6 *– The total membership in 2006 was 75 men + 50 women = 125 members. The total membership in 2007 was 120 men + 80 women = 200 members. So this statement is* ***NOT TRUE.***

Build your skills: Interpreting data – Part 1

The two diagrams below both show the same data in different forms.

Using either chart, **identify whether each of the statements beneath is true or false.**



1 The number of female members was greater than the number of male members each year.

2 In 2006 the numbers of male and female members was equal.

3 The number of members who are men has gradually increased over the years.

4 There were more female members in 2007 than there were in 2005.

5 There were more male members in 2006 than there were in 2005.

6 The balance of male and female members has changed over the years.

Activity 4

This section will offer a quick reminder about mean, median and mode. If you already feel confident about what each of these is and how to work it out, move straight on to the next activity.

Working out mean, median and mode

Reminder:

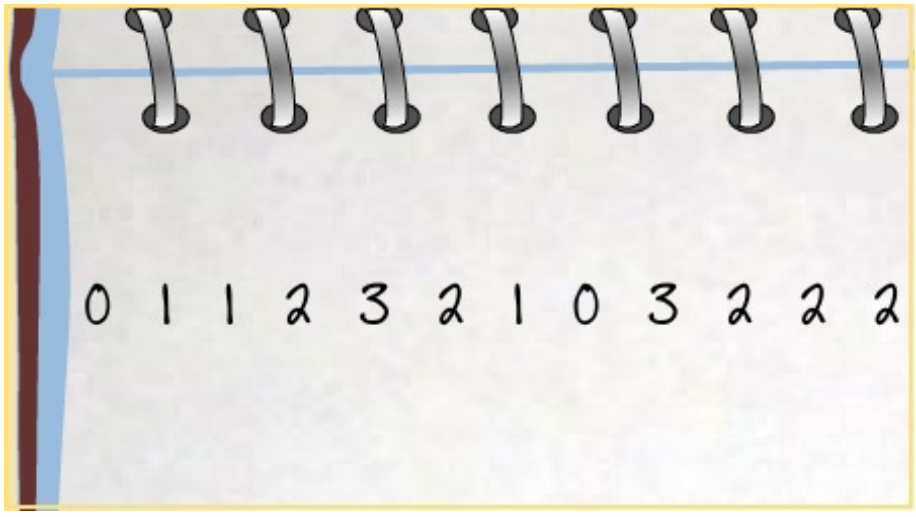
$$\textit{Mean} = \frac{\textit{Total of all the values}}{\textit{Number of values in the data set}}$$

Median is the middle value (when they are ordered from smallest to largest).

Mode is the value that appears most frequently.

Example:

Let's think again about the data used in our original example. Households in a street were surveyed and asked how many cars they each own. The results from this survey were:



The mean

The formula for working out the mean can be summarised as:

$$\frac{\textit{Total of all values}}{\textit{Number of values}}$$

This means that there are two steps:

1 Add up all the *values.*

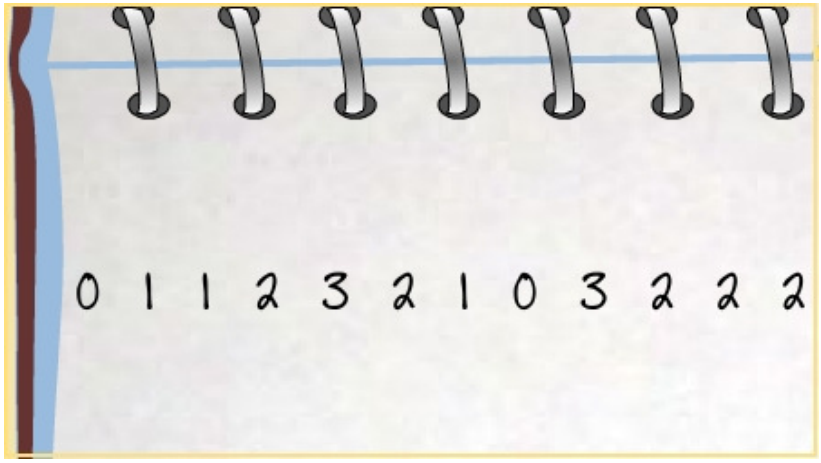
2 Divide this number by the number of values in the data set.

So:

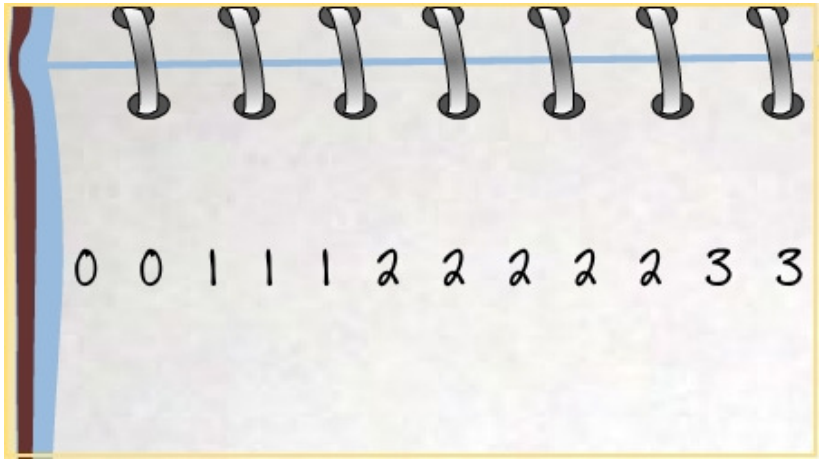
$$\frac{0+0+1+1+1+2+2+2+2+2+3+3}{12} = \frac{19}{12} = \mathbf{1.58}$$

The median

We need to write the values in order of size, and then count through them to find the middle value(s).



When written in order (starting with the smallest) they will look like this:



Tip: If you cross the numbers off from your original list of values as you order them, it will help to make sure that you don't lose any numbers in the process of getting them in order.

(It is also useful to count how many values you have in the re-ordered list to check that you have them all. Remember to count each individual value: for example, in the above data set, the two zeros will be values 1 and 2 respectively.)

You can then see which are the middle value(s).

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

Once you have found the middle value(s):

If there is only one middle value, this will be the median.

If there are two middle values (because the data set has an even number of values in it), the median will be the midway point between the two. (You can work this out by adding them together and then dividing this number by two.)

In this example, both middle values are 2 – so the **median** is **2**.

Tip: when you work out the middle values like this, you can check that you have the middle value(s) by counting the number of values on either side – there should be the same number on each side of the middle value(s). In this example there are five numbers to the left of values 6 and 7 – and five to the right.

***Note: If the number of values is odd**, the position of the middle value will be half of the number of values (rounded up).*

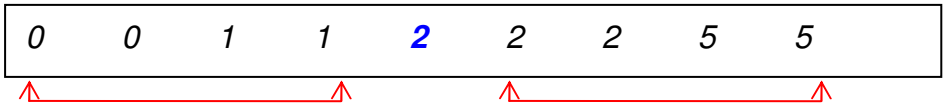
Example:

0 0 1 1 2 2 2 5 5

There are 9 values in this data set.

Half of this is 4.5.

*Rounded up, this tells us that the **5th value will be the middle one.***



If the number of values is even, *there will be two middle values – the one that represents half of the number of values* ***and*** *the next value.*

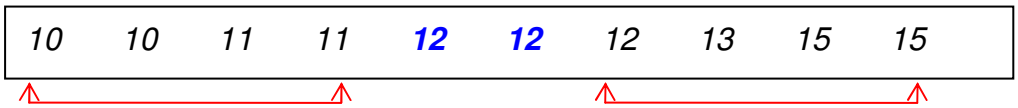
Example:

10 10 11 11 12 12 12 13 15 15

There are 10 values in this data set.

Half of this is 5.

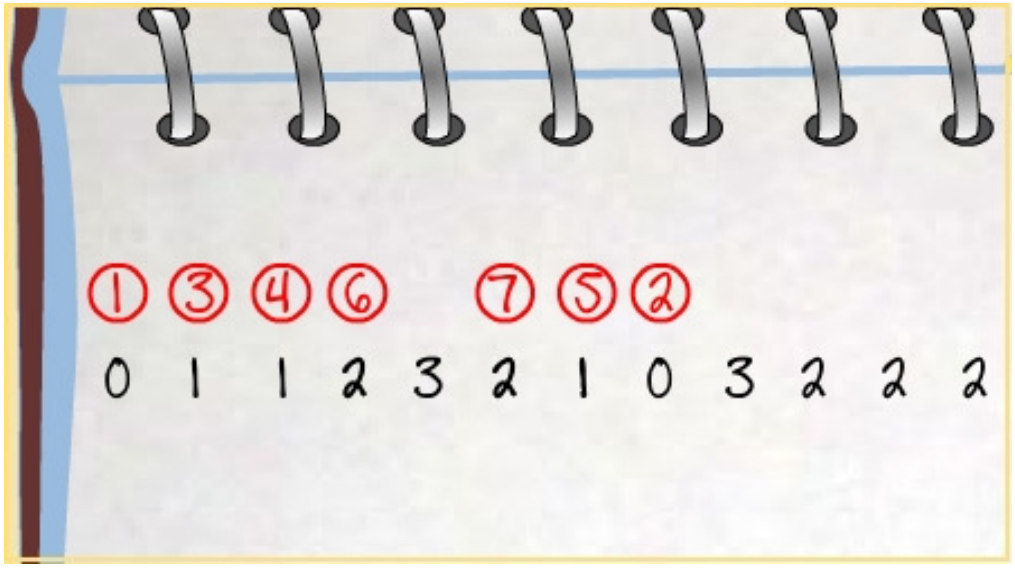
So, ***the 5th*** *(half of the number of values)* ***and the 6th*** *(the next value on)* ***will be the middle ones****.*



Thoughts on a possible alternative approach

As an alternative approach to find the middle value(s), you could choose not to physically re-order the values by writing them out again, but instead could use the original (unordered) list of data and write what place in the order each value in the sequence would have by writing this above that value in the list.

To do this, scan through the list of values looking for the smallest; label this value 1 by putting a (1) above it. Then search for the next biggest value and label this with a (2), etc.

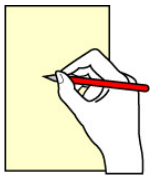


(When you get as far as the middle value(s) you don't need to label the remaining values.)

Which value comes next in the order?

The mode

The **mode** will be **the most frequent value**, i.e. the number that appears most often in the data set. In this example, this is the number **2** (which appears five times).



For each of these sets of data, work out the:

- **mean**
- **median**
- **mode.**

1 The cost of an item at six different shops is:

£1.25 £1.15 £1.19 £1.21 £1.17 £1.23.

2 The number of houses in different streets in a new housing estate is:

74 83 72 94 97 88 81 79 92 80.

3 In a class, the number of grade A–C passes gained by different pupils was:

3 5 6 4 7 8 4 3 5 5.

4 The time taken for members of a staff team to travel to work each day:

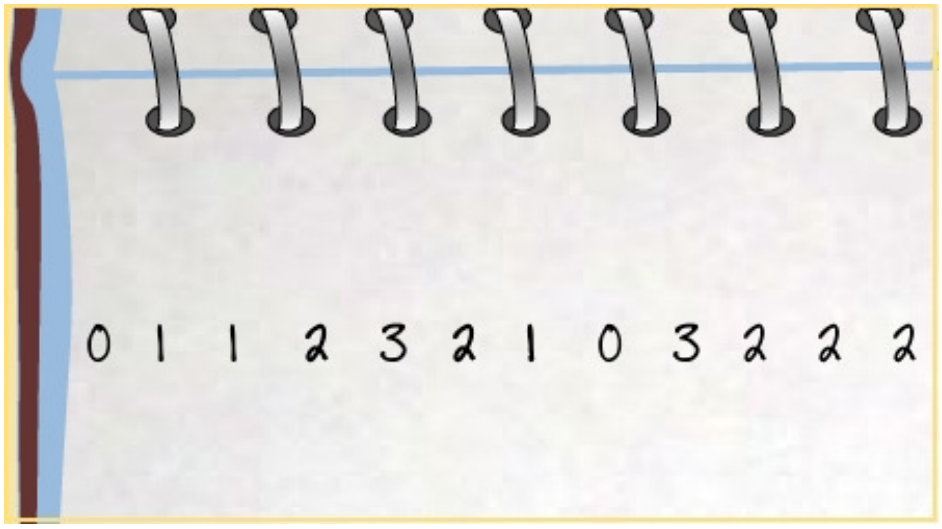
10 mins, 30 mins, 45 mins, 20 mins, 15 mins, 25 mins, 5 mins, 10 mins.

Thinking about mean, median and mode – and what they tell us

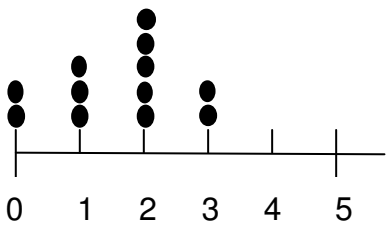
In this next section we are going to think about the mean, median and mode in a slightly different way. Thinking about each using a visual approach might help you to understand better and/or remember what they each represent.

Example:

Using the same data in the example above about the number of cars in households in a street:



We could think of the data represented on a diagram like this:

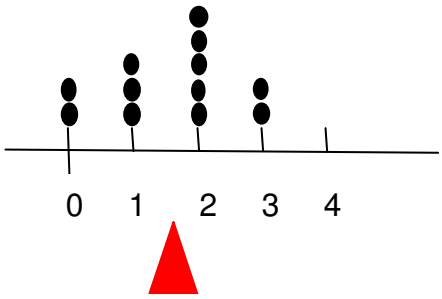


Each dot on the line represents one value of the data set.

The number ‘1’ appears three times in the list of data values, so there are three dots above the point on the line marking 1.

Similarly, the number '2' appears five times in the list of data values, so it has five dots above the point on the line marking 2.

Thinking about the mean visually:



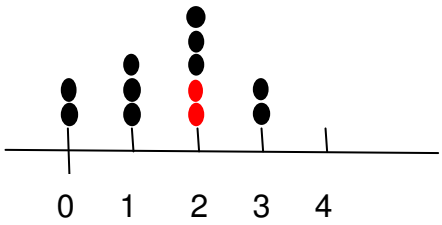
Mean of the values = **1.58**

The mean of the data is like the 'balancing point'.

If the line above was a set of balancing scales, and the dots represented weights (each of the same weight) placed at the points shown on the scale, the triangle shows the point at which the weights on either side of the triangle would balance and the scales would become level.

***Note**: If you would like more practice in working out the mean and/or in thinking about what the mean shows using this sort of visual approach, see the Build Your Skills task: 'Working out the mean' on the Move On Learner Route.*

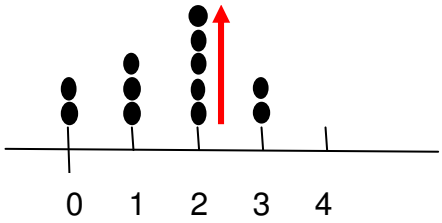
Thinking about the median visually:



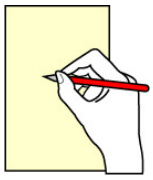
Median = **2** (the middle values)

Thinking about the mode visually:

Mode = **2** (the tallest 'stack')

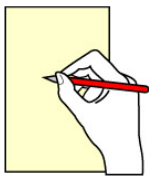


Note: You would not be asked to draw diagrams like these in the National Test. However, thinking about the mean, median and mode visually like this may help you to remember what each is and to think about what they tell you about the data.



Activity 5

If you would like to, you could try creating visual diagrams like those in the examples above for the data sets from each of the questions in Activity 4. Otherwise, just move on to the next activity below.



Activity 6

Working out the mean and range of a set of data presented in a table

It can help you to interpret and use data confidently if, when you look at it, you think about how it relates to real life.

This cell of the table is telling me that 5 households have 2 cars . . .

. . . and this cell that another 2 households have 3 cars

Number of cars owned	0	1	2	3
Number of households	2	3	**5**	**2**

As shown in Activity 4 above, the formula for working out the mean can be summarised as:

$$\frac{\textbf{\textit{Total of all values}}}{\textbf{\textit{Number of values}}}$$

When data is presented in tabular format, to work out the mean you need to think about what the full data set would look like.

In the example of the data presented in the table above:

0 0 1 1 1 2 2 2 2 2 3 3

1 *Adding up all the values gives:*

0 + 0 + 1 + 1 +1 + 2 + 2 + 2 + 2 + 2 + 3 + 3 = ***19***

2 The number of values is 12. (Twelve households were surveyed.)

So, the mean number of cars per household is:

$$\frac{19}{12} = \textbf{\textit{1.58}}$$

Think about how this relates to the data presented in the table format:

Number of cars owned	**0**	**1**	**2**	**3**	**Total**
Number of households	2	3	5	2	**12**
Number of cars	0 × 2 = 0	1 × 3 = 3	2 × 5 = 10	3 × 2 = 6	**19**

3 households have 1 car each, making 3 cars between these 3 households.

***Work out the number of values** **

Add up the numbers in the blue row: 2 + 3 + 5 + 2 = 12 (12 households surveyed)

***Work out the total of all the values** **

Add up the numbers in the red row: 0 + 3 + 10 + 6 = 19 (i.e. 19 cars altogether)

Thinking about how this relates to the data presented (as a list of values) might help to understand the calculation for the total number of cars (shown in the bottom row of the table above):

*0 + 0 + 1 + 1 + 1 + 2 + 2 + 2 + 2 + 2 + 3 + 3 = **19***

*0 + 3 + 10 + 6 = **19***

Working out the mean of the data from the table gives us:

$$\frac{\text{Total of all values *}}{\text{Total number of values *}} = \frac{19}{12} = \textbf{1.58 cars}$$

Take these numbers from the cells in the table where you worked them out

Working out the median for this same set of data – again it helps to think about what the data represents in real life:

Median = middle value(s)

Again, thinking about what the data you have means will help you to interpret data from a table and work out the median.

Example:

Continuing to think about the data on the number of cars per household presented in the table above:

The total number of households is 12.
So the middle values will be the 6th and 7th.

Number of cars owned	0	1	2	3	
Number of households	2	3	5	2	**12**

Add up this row to work out the total number of households

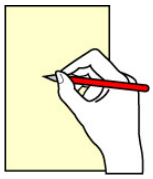
First two values *Next three values*

*So, counting through the values: 1 and 2 3, 4 and 5 **6, 7** – 10*

So, the 6th and 7th values are both 2 cars.

The median is 2.

Build your skills: Interpreting data – Part 1



Try using approaches like those above (or other appropriate approaches) to work out the answers to these questions.

1 Number of children in households in the street

The number of children in the families in one street is:

Number of children	0	1	2	3	4	5
Number of households	2	2	3	1	1	1

a Present this data set as a tally.

b And as a chart or graph.

c What is the mean number of children per family?

d What is the median of the number of children per family?

e What is the mode?

2 The number of computers owned by households in street X

The number of computers owned by households in a street was:

No. of computers	**0**	**1**	**2**	**3**	**4**
1998	10	7	3	0	0
2008	3	10	4	2	1

a What is the mean number of computers per household in each of the two years?

***Tip**: Think about each year (each of the 1998 and 2008 rows of the table) one at a time.*

You might want to add a row beneath the row you are working on to help you do your calculations.

For example, when focusing on working out the mean for 1998:

No. of computers	**0**	**1**	**2**	**3**	**4**	***Total***
1998	10	7	3	0	0	
Total computers (1998)						
2008	3	10	4	2	1	

b Work out the median and mode for each of the two years.

Now print out Part 2 of this task to find suggestions of other free resources you might want to use to practise these skills.