# 18 Quantitative Data

## 18.1 Presentation

In this section we look at how vertical line diagrams can be used to display discrete quantitative data. (Remember that discrete data can only take specific numerical values.)

### Example 1

The marks below were scored by the children in a class on their maths test. The marks are all out of a possible total of 10 marks.

| | | | | |
|---|---|---|---|---|
| 8 | 6 | 8 | 7 | 7 |
| 7 | 10 | 9 | 6 | 8 |
| 8 | 4 | 3 | 2 | 5 |
| 8 | 8 | 6 | 5 | 6 |
| 4 | 9 | 8 | 4 | 7 |
| 7 | 5 | 3 | 7 | 6 |

Draw a vertical line diagram to illustrate these data.

Use your diagram to answer these questions:

(a) What is the *most common* mark?

(b) What is the *highest* mark?

(c) What is the *lowest* mark?

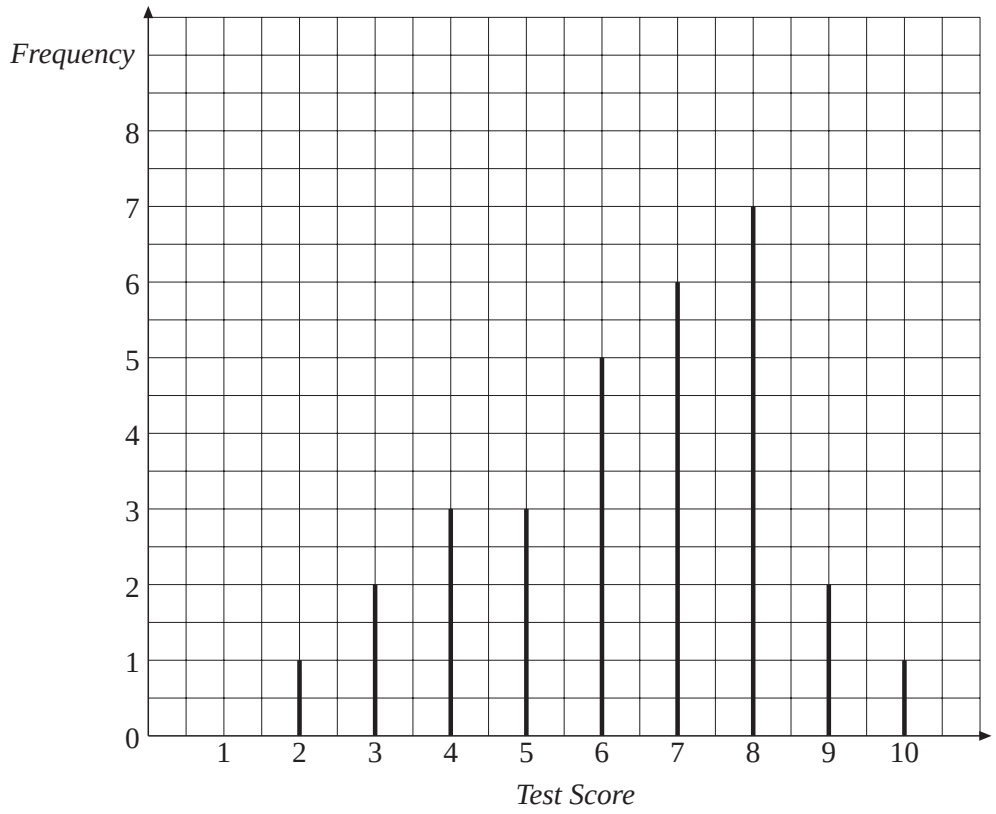
(d) What is the difference between the *highest* and *lowest* marks?

### Solution

The first step is to organise the data using a *tally chart*, as shown here:

| *Mark* | *Tally* | *Frequency* |
|---|---|---|
| 2 | \| | 1 |
| 3 | \|\| | 2 |
| 4 | \|\|\| | 3 |
| 5 | \|\|\| | 3 |
| 6 | 𝍸 | 5 |
| 7 | 𝍸 \| | 6 |
| 8 | 𝍸 \|\| | 7 |
| 9 | \|\| | 2 |
| 10 | \| | 1 |

The diagram can then be drawn as shown below. The height of each line is the same as the frequency; that is, the number of times it occurs in the data list.




(a) The *most common* mark is 8, which occurred 7 times.

(b) The *highest* mark is 10.

(c) The *lowest* mark is 2.

(d) The difference between the *highest* and *lowest* marks is $10 - 2 = 8$.

*Note:* a *vertical line diagram* is an appropriate way to represent information that consists of distinct, single values, each with its own frequency. A *bar graph* is more suitable for grouped numerical data.

## Exercises

1. A teacher gives the children in her class a test, and lists their scores in this table:

| *Mark* | *Frequency* |
|---|---|
| 1 | 1 |
| 2 | 4 |
| 3 | 1 |
| 4 | 3 |
| 5 | 6 |
| 6 | 8 |
| 7 | 4 |
| 8 | 2 |

   (a) Draw a vertical line diagram to illustrate these results.

   (b) What is the *most common* mark?

   (c) How many children are there in the class?

2. The staff in a shoe shop keep a record of the sizes of all the shoes they sell in one day. These are listed below:

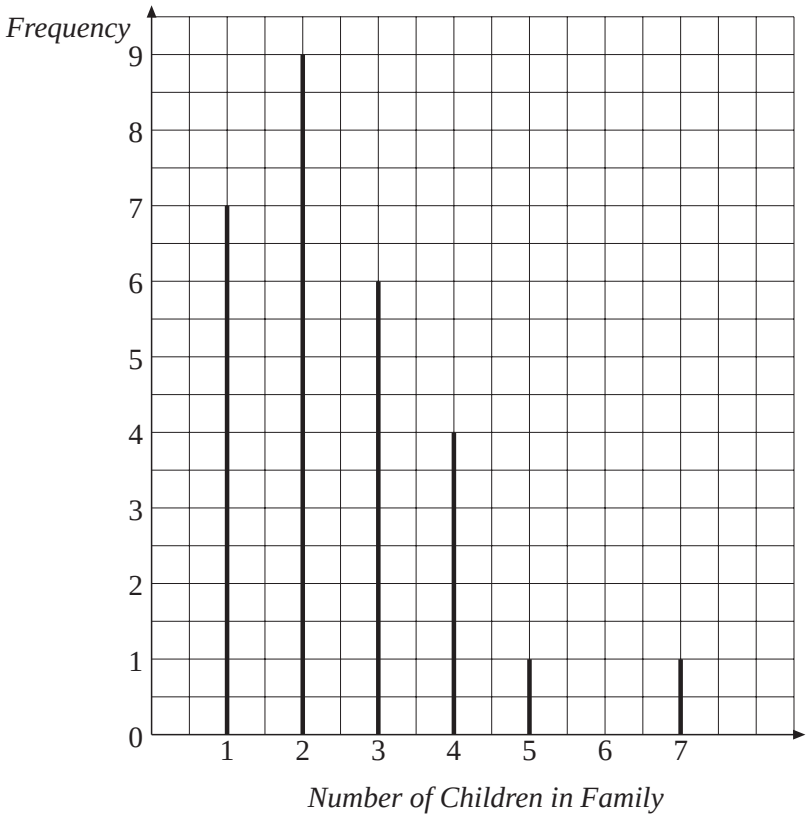
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 8 | 7 | 6 | 6 | 8 | 7 | 5 | 4 | 3 | 1 |
| 11 | 7 | 8 | 9 | 5 | 6 | 6 | 5 | 6 | 4 |
| 3 | 10 | 8 | 9 | 7 | 6 | 6 | 5 | 4 | 2 |
| 6 | 9 | 11 | 3 | 5 | 6 | 7 | 8 | 8 | 3 |
| 4 | 6 | 7 | 8 | 9 | 8 | 8 | 7 | 6 | 4 |

(a) Complete a tally chart for these data.

(b) Draw a vertical line diagram for these data.

(c) What advice could you give the shop staff about which size shoes they should keep in stock?

3. The vertical line diagram below is based on data collected by a class about the number of children in their families:




(a) What is the *most common* number of children per family?

(b) How many children are there in the class?

4. (a) Collect data on the number of children in the families of the pupils in your mathematics class.

   (b) Draw a vertical line diagram like the one in question 3.

   (c) Compare your vertical line diagram with the one for question 3. What *similarities* are there? What *differences* are there?

5. Mr Graddon says that his class is better at tables than Mr Hall's class. The two classes each take a tables test, and the results are given below. The scores are out of 10.

| *Mr Graddon's Class* | | | | | | *Mr Hall's Class* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 6 | 7 | 8 | 9 | 10 | 4 | 7 | 8 | 3 | 5 | 6 |
| 0 | 1 | 3 | 6 | 9 | 2 | 7 | 4 | 5 | 6 | 6 | 5 |
| 5 | 1 | 2 | 2 | 0 | 1 | 5 | 5 | 6 | 7 | 4 | 3 |
| 6 | 4 | 0 | 1 | 10 | 9 | 4 | 5 | 6 | 6 | 7 | 8 |
| 1 | 2 | 3 | 5 | 10 | 9 | 6 | 7 | 5 | 6 | 4 | 5 |

   (a) Draw a vertical line diagram for each class.

   (b) Which features of the two diagrams would Mr Graddon use to support his claim that his class is better at tables?

   (c) How would Mr Hall use the diagrams to argue the other way?

   (d) Which class do *you* think is better at tables?

6. A gardener keeps a record of the number of tomatoes he picks from the plants in his greenhouse during August. The number of tomatoes picked each day is listed below:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 10 | 3 | 6 | 8 | 9 | 5 | 10 | 4 | 7 | 9 |
| 6 | 10 | 11 | 12 | 13 | 7 | 8 | 4 | 3 | 6 | 9 |
| 7 | 9 | 10 | 11 | 14 | 13 | 7 | 8 | 9 | | |

   (a) Draw a vertical line diagram for these data.

   (b) What is the *largest* number of tomatoes picked on one day?

   (c) What is the *smallest* number of tomatoes picked on one day?

   (d) What is the number of tomatoes that was picked *most often*?

7. A sample of children were asked how many pets they had, and their responses are listed below:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 1 | 0 | 2 | 0 | 1 | 3 | 4 | 0 |
| 1 | 0 | 1 | 2 | 0 | 1 | 1 | 3 | 0 | 5 |

(a) Draw a vertical line diagram for these data.

(b) How many pets were in the sample?

(c) How many children owned at least one pet?

(d) Is it true that, in this sample, there are more children who own pets than children who do not?

8. A rail company keeps a record of how many trains are late each day. The data for January are listed below:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 0 | 3 | 0 | 1 | 1 | 2 | 0 | 3 | 0 | 4 |
| 6 | 1 | 0 | 0 | 0 | 2 | 1 | 3 | 1 | 0 | |
| 0 | 0 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 | |

The data for February are listed below:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 2 | 4 | 7 | 0 | 1 | 2 | 0 | 1 | 2 |
| 0 | 0 | 0 | 1 | 0 | 1 | 2 | 1 | 2 | 0 |
| 0 | 2 | 1 | 3 | 1 | 2 | 1 | 1 | | |

(a) Draw vertical line diagrams for each month.

(b) Comment on whether the trains were on time more often in February than in January.

9. A traffic warden keeps a record of the number of parking tickets that she issues on 20 working days.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 3 | 7 | 8 | 12 | 0 | 1 | 3 | 4 | 5 |
| 6 | 5 | 4 | 0 | 1 | 3 | 4 | 6 | 7 | 5 |

(a) Draw a vertical line diagram for these data.

(b) How many blank parking tickets do you think she should take with her when she starts her daily traffic patrol? Explain your answer.

10. Graham uses his calculator to generate random numbers. He decides to investigate if the numbers are really random. Using his calculator, he produces the following numbers:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 9 | 9 | 1 | 5 | 4 | 7 | 0 | 3 | 9 | 2 |
| 7 | 9 | 2 | 3 | 0 | 9 | 1 | 0 | 5 | 8 |
| 9 | 2 | 2 | 1 | 0 | 7 | 0 | 4 | 3 | 9 |
| 0 | 8 | 6 | 2 | 9 | 7 | 3 | 2 | 9 | 9 |

(a) Draw a vertical line diagram for these data.

(b) Do you think that the numbers that Graham's calculator produces are really random? Explain your answer.

# 18.2 Measures of Central Tendency

In this section we will consider three different types of 'average'. These are the *mean*, the *median* and the *mode*, and statisticians refer to them as *measures of central tendency.*

$$\textit{Mean} = \frac{\textit{sum of all values}}{\textit{total number of values}}$$

*Median* = *middle value (when the data are arranged in order)*

*Mode* = *most common value*

Measures of central tendency are single values chosen as being representative of a whole data set. When we select which of the mean, the median or the mode to use, we choose the one that we think is most typical of the data and appropriate for the context.

## Example 1

What is:

(a) the *mean*, (b) the *median* and (c) the *mode*

of the numbers:

4, 7, 8, 4, 5

## Solution

(a) $\textit{Mean} = \dfrac{4+7+8+4+5}{5}$

$= \dfrac{28}{5}$

$= 5.6$

(b) To calculate the *median*, write the numbers in order,

4, 4, (5), 7, 8

The middle number is 5,

*median* = 5

(c) The most common number is 4, so

*mode* = 4

## Example 2

What number is the *median* of the numbers:

4, 7, 11, 4, 6, 7, 2, 9

## Solution

First write the numbers in order:

2, 4, 4, 6, 7, 7, 9, 11

In this case there are two middle numbers, 6 and 7. The *median* is the mean of these two numbers:

$$\textit{Median} = \frac{6+7}{2}$$

$$= 6.5$$

*Note:* where there is an *odd* number of data items, there will be a single value in the middle and that will be the median – provided you have arranged the data in order. When there is an *even* number of data items, there will be two values in the middle and you must find their mean to get the median of the full data set.

## Example 3

David keeps a record of the number of carrier bags that he is given when he does his weekly shopping. The data he collects over 10 weeks is listed below:

9 8 5 9 12 8 7 6 5 9

(a) Calculate: (i) the *mean*, (ii) the *median*, (iii) the *mode*?

(b) Explain why the mean is not very useful in this context.

(c) Which value might be used by an environmental group who think that supermarkets cause pollution by giving out too many carrier bags?

(d) Which value might be used by a shopper who thinks that the supermarket doesn't give him enough carrier bags for his shopping?

## Solution

(a) (i) $\textit{Mean} = \dfrac{9+8+5+9+12+8+7+6+5+9}{10}$

$$= \frac{78}{10}$$

$$= 7.8$$

(ii) To find the *median*, put the numbers in order, and find the middle numbers:

5 5 6 7 8 8 9 9 9 12

$$\textit{Median} = \frac{8+8}{2}$$

$$= 8$$

(iii) The most common number is 9:

$\textit{Mode} = 9$

(b) The mean is not very useful as no one would ever actually use 7.8 plastic bags.

(c) The mode, as this is the largest of the three values.

(d) The mean, as this is the smallest of the three values.

## Exercises

1. Find the *mean, median* and *mode* of each set of numbers:

(a) 4 4 6 8 5

(b) 6 7 7 7 7 5 6 2 9 8

(c) 8 4 3 3 5 7

(d) 6 6 7 7 4 9 1 7 10

2. The owner of a shoe shop recorded the sizes of the feet of all the customers who bought shoes in his shop in one morning. These sizes are listed below:

8 7 4 5 9 13 10 8 8 7

6 5 3 11 10 8 5 4 8 6

(a) What are the *mean, median* and *mode* shoe sizes?

(b) Which of these values would be most sensible for the shop owner to use when ordering shoes for his shop? Explain your choice.

3. Eight people work in a shop. They are paid hourly rates of

£4 £15 £6 £5 £4 £5 £4 £4

Would you use the *mean, median* or *mode* to show that they were:

(a) *well* paid, (b) *badly* paid?

4. A newspaper reports that the average number of children per family is 2.4.

   (a) Which type of value has the newspaper used?

   (b) Explain how you can tell which value was used.

   (c) Would your answer to (b) be the same if the newspaper had reported the average as 2.5 children?

5. The mean of six numbers is 9. If five of the numbers are 10, 12, 7, 6 and 9, what is the sixth number?

6. The table below gives the number of accidents each year at a particular road junction:

| 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 |
|---|---|---|---|---|---|---|---|
| 4 | 5 | 4 | 2 | 10 | 5 | 3 | 5 |

   (a) Calculate the *mean*, *median* and *mode*.

   (b) Describe which value would be most sensible for a road safety group to use, if they want the junction to be made less dangerous.

   (c) The council do not want to spend money on the road junction. Which value do you think they should use?

7. One day the number of minutes that trains were late to arrive at a station was recorded. The times are listed below:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 7 | 0 | 0 | 1 | 2 | 5 | 0 | 0 | 0 |
| 6 | 0 | 1 | 52 | 0 | 10 | 1 | 1 | 8 | 22 |

   (a) Calculate the *mean*, *median* and *mode* of these data.

   (b) Explain which value would be the best to use to argue that the trains arrive late too often.

   (c) Explain who might use the mode and why it might be an advantage to them.

8. Mr Hall grows two different types of tomato plant in his greenhouse. One week he keeps a record of the number of tomatoes he picks from each type of plant.

| *Day* | *Mon* | *Tues* | *Wed* | *Thurs* | *Fri* | *Sat* | *Sun* |
|---|---|---|---|---|---|---|---|
| *Type A* | 5 | 5 | 4 | 1 | 0 | 2 | 5 |
| *Type B* | 3 | 3 | 3 | 3 | 7 | 9 | 6 |

(a) Calculate the *mean*, *median* and *mode* for each type of plant.

(b) Use one value to argue that type A is the best plant.

(c) Use a different value to argue that type B is the best plant.

9. The heights of eight children are given below, to the nearest cm:

158 162 142 155 163 157 160 112

(a) Explain why the mode is *not* a suitable value to use for these data.

(b) Calculate the median and the mean of these data.

(c) Explain why the mean is less than the median.

10. A set contains four positive numbers.

The *mode* of these numbers is 1.

The *mean* of these numbers is 2.5.

The *median* of these numbers is 1.5.

What are the four numbers?

# 18.3 Measures of Dispersion

The *range* of a set of data is the difference between the largest and the smallest values in the data set. The range gives a measure of the dispersion of the data, or, more simply, describes the spread of the data.

## Example 1

Calculate the *range* of this set of data:

4 7 6 8 3 9 14 22 3

### Solution

The largest value is 22.

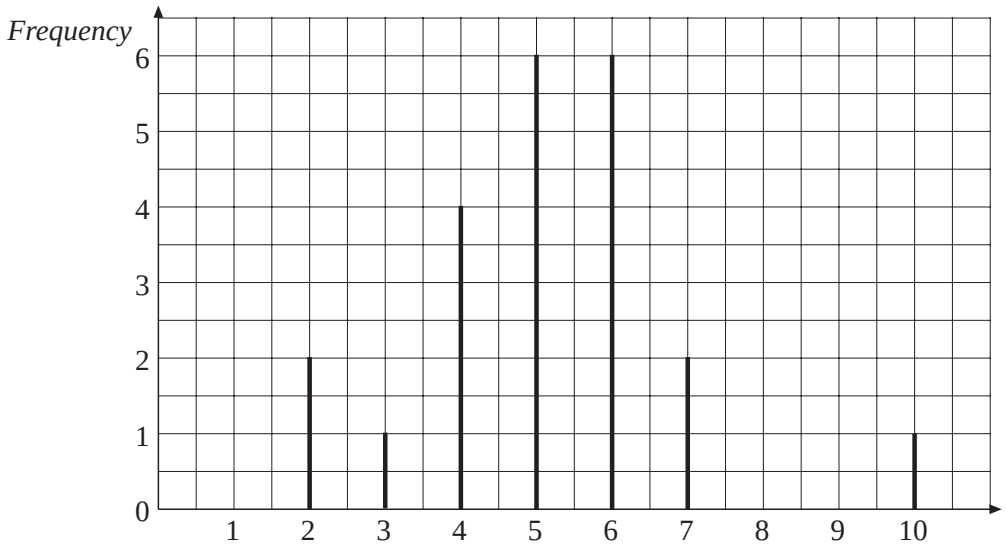
The smallest value is 3.

*Range* = 22 – 3

= 19

## Example 2

What is the *range* of the data illustrated in this vertical line diagram?




## Solution

Largest value = 10

Smallest value = 2

*Range* = 10 – 2

= 8

## Exercises

1. Calculate the *range* of each of these sets of data:

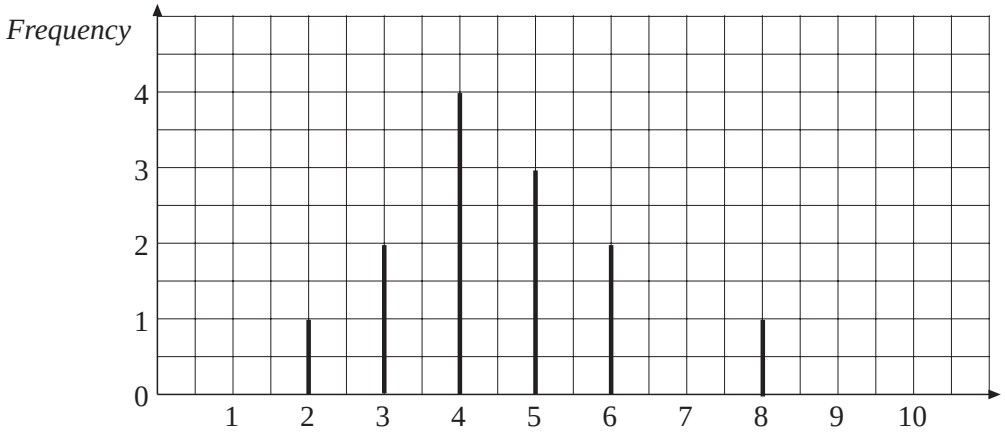
   (a) 4 7 6 3 9 12 7 12

   (b) 6 5 5 16 12 21 42 7

   (c) 0 2 4 1 3 0 6

   (d) 3 7 8 9 4 7 11

2. Calculate the *range* of the data illustrated in this vertical line diagram:




3. The range of a set of data is 12 and the smallest number in the set of data is 5.

   What is the *largest* number in the set of data?

4. The largest number in a set of data is 86. The range of the set of data is 47. What is the *smallest* number in the set of data?

5. The heights of 10 students were measured to the nearest centimetre and are listed below:

   144 162 173 158 143
   159 164 182 162 158

   What is the *range* of this set of data?

6. Rafiq keeps a record of the amount of money he spends each day. The amounts for one week are listed below:

   47p 10p 36p 85p 22p 30p

   There are only 6 amounts because he forgets to include one day.

   (a) What is the *range* of the numbers listed above?

   (b) If the range was 90p, what was the missing amount?

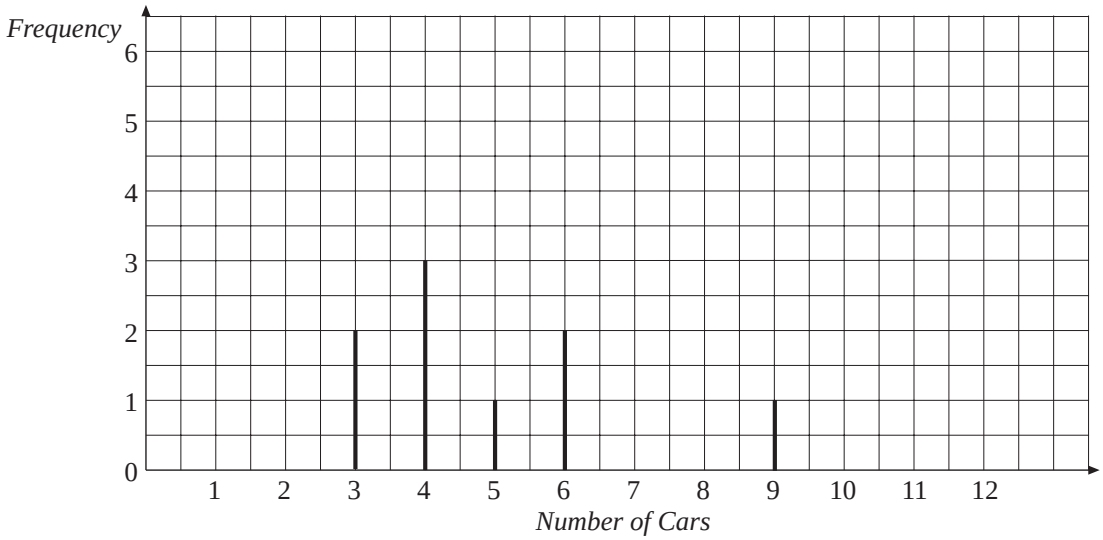
   (c) If the range was double your answer to (a), what was the missing amount?

   (d) Explain why the range must be equal to or greater than your answer to part (a).

7. The vertical line diagram on the following page is for a data set that has one missing value.

   What can you say about the missing value if the range is:

   (a) 7, (b) 9, (c) 6 ?




8. What is the range of this set of temperatures:

$-4\,°C \qquad 3\,°C \qquad 5\,°C \qquad -1\,°C \qquad -3\,°C \qquad 6\,°C$ ?

9. The range of a set of temperatures is $8\,°C$. If the *maximum* temperature in the set is $6\,°C$, what is the *minimum* temperature?

10. The range of a set of temperatures is $7\,°C$. If the *minimum* temperature in the set is $-11\,°C$ what is the *maximum* temperature?

# 18.4 Comparing Data

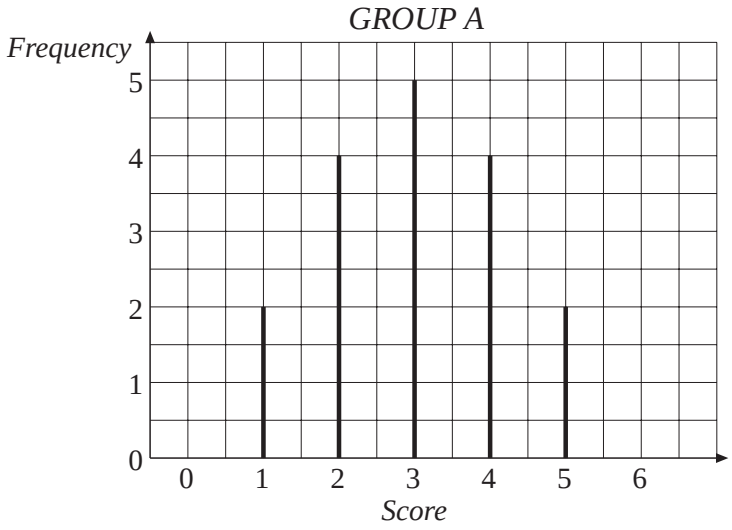
In this section we consider how averages and the range can be used to compare sets of data.

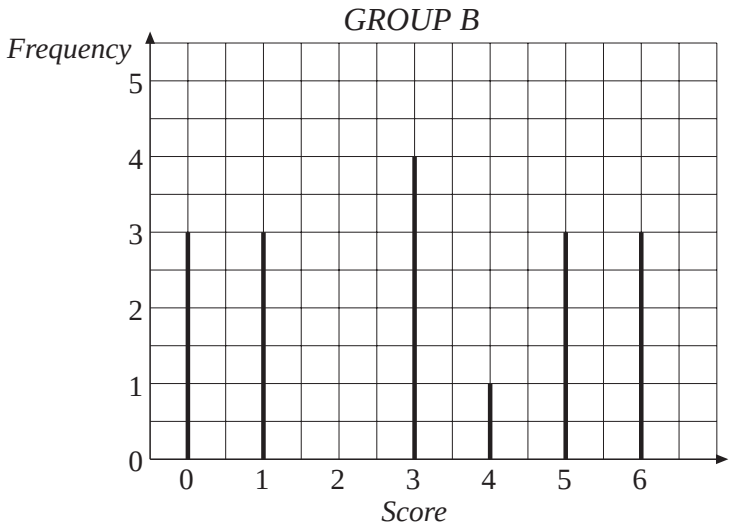
## Example 1

The two line diagrams on the next page illustrate data that was collected about the scores of two groups of children in a short test.

(a) Calculate the *mode* and *range* for each group.

(b) Describe the differences between the groups.




## Solution

(a)

| *Group A* | *Group B* |
|---|---|
| *Mode* $= 3$ | *Mode* $= 3$ |
| *Range* $= 5 - 1$ | $= 6 - 0$ |
| $= 4$ | $= 6$ |

(b) Both groups have the same mode but different ranges. The range is greater for group B.

The low range for group A indicates that the scores for those students are reasonably similar. The higher range for group B shows that their scores are much more varied. This can be seen from the line diagrams, where none of group A get the extreme scores of 0 and 6, while these are obtained by several students in group B.

## Example 2

Kathryn plants two different types of tomato plant. She records the number of tomatoes that she picks from each plant every day for 10 days. Her records are shown below:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Plant A* | 4 | 6 | 7 | 3 | 5 | 2 | 1 | 3 | 6 | 5 |
| *Plant B* | 5 | 6 | 7 | 6 | 8 | 9 | 6 | 7 | 8 | 9 |

Compare the two plants and recommend which type she should buy next year.

## Solution

First consider the mean and range for each plant:

*PLANT A*

$$\textit{Mean} = \frac{4+6+7+3+5+2+1+3+6+5}{10}$$

$$= \frac{42}{10}$$

$$= 4.2$$

$$\textit{Range} = 7 - 1$$

$$= 6$$

*PLANT B*

$$\textit{Mean} = \frac{5+6+7+6+8+9+6+7+8+9}{10}$$

$$= \frac{71}{10}$$

$$= 7.1$$

$$\textit{Range} = 9 - 5$$

$$= 4$$

As plant B has a higher mean, this suggests that using plant B will produce more tomatoes than using plants of type A. The fact the plant B has the lower range suggests that it will also be more consistent in the number of tomatoes that it produces than type A. Type A will have some productive days but it will also have some poor days.

## Exercises

1. (a) Calculate the *mean* and *range* of these two data sets:

| | | | | | | |
|---|---|---|---|---|---|---|
| A | 5 | 10 | 0 | 1 | 9 | 5 |
| B | 5 | 6 | 4 | 3 | 7 | 5 |

(b) Describe the difference between the two sets.

2. (a) Calculate the *mean* and *range* of these two data sets:

| | | | | | | |
|---|---|---|---|---|---|---|
| A | 4 | 6 | 7 | 8 | 5 | 6. |
| B | 5 | 7 | 7 | 8 | 9 | 6 |

(b) Describe the difference between the two sets.

3. (a) Calculate the *mean* and *range* of these two data sets:

| | | | | | | |
|---|---|---|---|---|---|---|
| A | 4 | 6 | 10 | 3 | 5 | 2 |
| B | 6 | 7 | 9 | 9 | 5 | 3 |

(b) Describe the differences between the two sets.

4. (a) Calculate the *mean* and *range* of these 3 sets of data:

| | | | | | | |
|---|---|---|---|---|---|---|
| A | 4 | 7 | 8 | 6 | 5 | |
| B | 0 | 10 | 12 | 1 | 3 | |
| C | 8 | 8 | 9 | 10 | 9 | 8 |

(b) Describe the differences between the three sets.

5. Roy and Frank are second-hand car salesmen. The following vertical line diagrams show how many cars they have sold per week over a period of time.

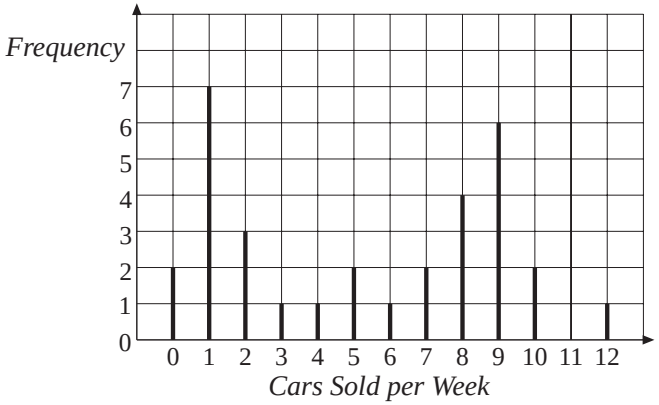
(a) Write down the *mode* for Roy and for Frank.

(b) Calculate the *range* for Roy and for Frank.

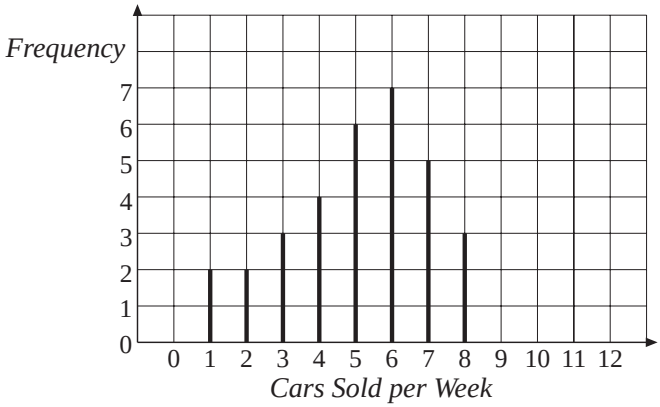
(c) Who sold more cars?

(d) Who you think is the better salesman? Explain why.

*ROY*

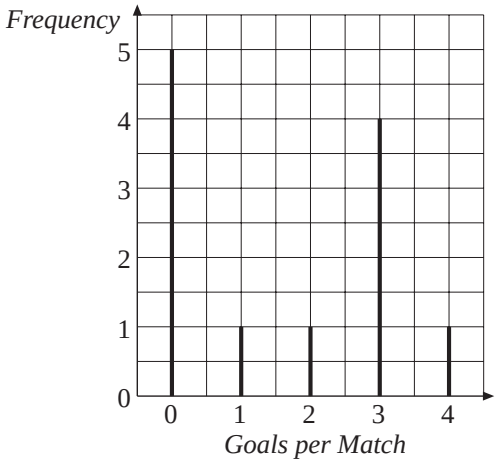



*FRANK*

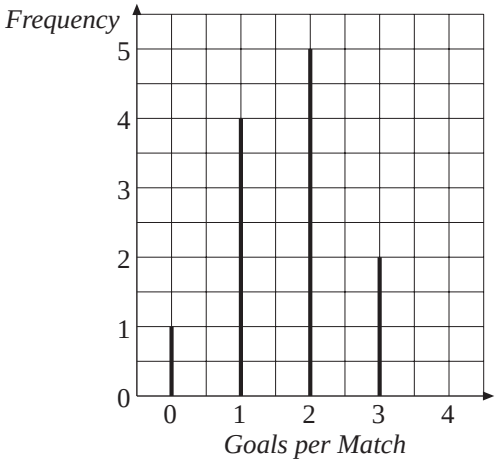



6. The two vertical line diagrams show the number of goals scored per match by two top footballers.

*ANDY GOAL*




*ALAN SCORER*




(a) Calculate the *mean* and *range* for each player.

(b) Describe the differences between the two players.

(c) Which of these players would you like to have on your favourite team? Explain why.

7. Miss Sharp's class decide to have a spelling competition with Mr Berry's class. They have a test and the scores for each class are listed below:

| *Miss Sharp's Class* | | | | | | *Mr Berry's Class* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 5 | 8 | 5 | 7 | 5 | 5 | 7 | 6 | 7 | 8 |
| 2 | 6 | 8 | 7 | 5 | 9 | 5 | 4 | 3 | 3 | 2 | 5 |
| 2 | 4 | 8 | 0 | 5 | 3 | 4 | 5 | 6 | 5 | 4 | 6 |
| 5 | 10 | 2 | 5 | 7 | 1 | 7 | 7 | 6 | 4 | 3 | 5 |
| 5 | 5 | 3 | 3 | 0 | 9 | 3 | 5 | 5 | 6 | 4 | 5 |

(a) Calculate the *mean* for each class.

(b) Calculate the *range* for each class.

(c) Comment on the differences between the two classes.

8. A bus company keeps records of the number of buses that were late each day in February and in July in the same year:

*February*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 7 | 5 | 4 | 3 | 0 | 0 | 1 | 2 | 5 |
| 9 | 10 | 5 | 4 | 3 | 6 | 7 | 1 | 0 | 0 |
| 0 | 0 | 1 | 2 | 1 | 0 | 4 | 1 | | |

*July*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0 | 1 | 0 | 3 | 1 | 2 | 3 | 4 | 9 | 1 |
| 2 | 0 | 4 | 1 | 1 | 2 | 3 | 4 | 1 | 5 | |
| 7 | 2 | 1 | 2 | 3 | 0 | 4 | 1 | 0 | 2 | |

(a) Calculate the *mean, median* and *mode* for each month.

(b) Calculate the *range* for each month.

(c) Do you think the bus company improved its service to customers between February and July? Give reasons for your answer.

9. "Do boys have bigger feet than girls?"

(a) Collect data from your class.

(b) Draw separate vertical line diagrams for the boys' data and the girls' data.

(c) Calculate the *mode, mean, median* and *range* for each set of data.

(d) Use your diagrams and calculations to decide, for your class, the answer to the question above.

10. Investigate whether girls eat more fruit than boys.

# 18.5 Trends

*Moving averages* can be used to make predictions. They do this by smoothing out monthly, seasonal or other periodic variations.
For example, an ice-cream seller might expect to sell more in the summer than he does in the winter. He could use a moving average over the four seasons to find out if his sales are increasing for each 12 month period.

$$\textit{1st moving average} = \frac{\text{spring 1} + \text{summer 1} + \text{autumn 1} + \text{winter 1}}{4}$$

$$\textit{2nd moving average} = \frac{\text{summer 1} + \text{autumn 1} + \text{winter 1} + \text{spring 2}}{4}$$

$$\textit{3rd moving average} = \frac{\text{autumn 1} + \text{winter 1} + \text{spring 2} + \text{summer 2}}{4}$$

$$\textit{4th moving average} = \frac{\text{winter 1} + \text{spring 2} + \text{summer 2} + \text{autumn 2}}{4}$$

and so on. In each case, the oldest piece of data is replaced by the newest one. So, for the *fifth moving average*, the ice-cream seller would replace the winter sales figure for the first year with the winter sales figure for the second year, and so on. Because the mean of four items of data is being found every time, this is called a *4 point* moving average.

## Example 1

(a) Calculate the 4 point moving averages for this list of data:

6 5 7 4 6.1 5.1 7.1 4.1

(b) Estimate the next two values in the list.

## Solution

(a) *1st moving average* $= \dfrac{6 + 5 + 7 + 4}{4}$

$= 5.5$

*2nd moving average* $= \dfrac{5 + 7 + 4 + 6.1}{4}$

$= 5.525$

*3rd moving average* $= \dfrac{7 + 4 + 6.1 + 5.1}{4}$

$= 5.55$

*4th moving average* $= \dfrac{4 + 6.1 + 5.1 + 7.1}{4}$

$= 5.575$

$$\textit{5th moving average} = \frac{6.1 + 5.1 + 7.1 + 4.1}{4}$$
$$= 5.6$$

(b) Note that the moving averages increase by 0.025 at each step.

The next moving average will be expected to be 5.625, so

$$5.625 \times 4 = 5.1 + 7.1 + 4.1 + x$$

where $x$ is the next term.

$$x = 5.625 \times 4 - 5.1 - 7.1 - 4.1$$
$$= 6.2$$

To estimate the next value, we use

$$5.65 \times 4 - 7.1 - 4.1 - 6.2 = 5.2$$

## Example 2

The table below gives the average daytime temperatures for each of the four seasons over a two-year period.

| *Year 1* | | | | *Year 2* | | | |
|---|---|---|---|---|---|---|---|
| *Spring* | *Summer* | *Autumn* | *Winter* | *Spring* | *Summer* | *Autumn* | *Winter* |
| 12.1 | 18.6 | 11.2 | 8.1 | 12.4 | 19.0 | 11.8 | 8.6 |

Use a 4 point moving average to predict the temperature for Spring and Summer of Year 3.

## Solution

(a) $$\textit{1st moving average} = \frac{12.1 + 18.6 + 11.2 + 8.1}{4}$$
$$= 12.5$$

$$\textit{2nd moving average} = \frac{18.6 + 11.2 + 8.1 + 12.4}{4}$$
$$= 12.575$$

$$\textit{3rd moving average} = \frac{11.2 + 8.1 + 12.4 + 19}{4}$$
$$= 12.675$$

$$\textit{4th moving average} = \frac{8.1 + 12.4 + 19 + 11.8}{4}$$
$$= 12.825$$

$$\textit{5th moving average} = \frac{12.4 + 19 + 11.8 + 8.6}{4}$$
$$= 12.95$$

The differences between the moving averages are

0.075, 0.1, 0.15, 0.125

$$\text{The mean difference} = \frac{0.075 + 0.1 + 0.15 + 0.125}{4}$$
$$= 0.1125$$

We can now predict:

$$\textit{6th moving average} = 12.95 + 0.1125$$
$$= 13.0625$$

$$\textit{7th moving average} = 13.0625 + 0.1125$$
$$= 13.175$$

$$\text{Year 3 } \textit{Spring} \text{ temperature} = 13.0625 \times 4 - 8.6 - 11.8 - 19.0$$
$$= 12.85$$

$$\text{Year 3 } \textit{Summer} \text{ temperature} = 13.175 \times 4 - 12.85 - 8.6 - 11.8$$
$$= 19.45$$

## Exercises

1. (a) Calculate the 3 point moving averages for this set of data:

   4 3 5 4 3 5

   (b) What do you notice about the moving averages?

2. (a) Calculate the 4 point moving averages for this set of data:

   6 2 7 1 8 4 9 3 10

   (b) Describe what is happening to the moving average.

   (c) Predict the next *two* values using a 4 point moving average.

3. (a) Calculate the 4 point moving averages for this data:

16 7 20 5 14.2 7.2 19.2 4.2

(b) Use your results to predict the next 2 values.

4. Use a 3 point moving average to estimate the next 2 entries in this list:

4 6 5 5.5 7.5 6.5 . . . . . .

5. The first value from a list of data is missing:

☐ 3.8 6.2 5.8 4.6 4.2 6.6 6.2

(a) Calculate the 4 point moving averages for the data given.

(b) Estimate the missing value.

6. The sales of an ice-cream company are given in the table below, in thousands of ice-creams:

| *1996* | | | | *1997* | | | |
|---|---|---|---|---|---|---|---|
| *Spring* | *Summer* | *Autumn* | *Winter* | *Spring* | *Summer* | *Autumn* | *Winter* |
| 3.6 | 9.7 | 3.2 | 4.1 | 3.6 | 9.8 | 3.4 | 4.4 |

Use a 4 point moving average to estimate the number of ice-creams sold each season in 1998.

7. The value, in pence, of a single share in a company is given in the table below:

| *1997* | | | | *1998* | | | |
|---|---|---|---|---|---|---|---|
| *January* | *April* | *July* | *October* | *January* | *April* | *July* | *October* |
| 58 | 62 | 74 | 81 | 67 | 70 | 81 | 89 |

Use a 4 point moving average to estimate the value of the share for January, April, July and October 1999.

8. A company keeps a record of its total profits, in £10 000's, for the first, second, third and fourth quarters of each year.

| *1997* | | | | *1998* | | | |
|---|---|---|---|---|---|---|---|
| *1st* | *2nd* | *3rd* | *4th* | *1st* | *2nd* | *3rd* | *4th* |
| 24.1 | 26.3 | 28.4 | 20.4 | 29.3 | 31.9 | 35.2 | 28.4 |

Use a 4 point moving average to estimate the profits for:

(a) 1999, (b) 1996.

9. A school tuck shop keeps a record of the number of cans of drink it sells over a 3-week period.

| *Week 1* | | | | | *Week 2* | | | | | *Week 3* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Mon* | *Tues* | *Wed* | *Thurs* | *Fri* | *Mon* | *Tues* | *Wed* | *Thurs* | *Fri* | *Mon* | *Tues* | *Wed* | *Thurs* | *Fri* |
| 18 | 22 | 9 | 7 | 15 | 19 | 23 | 9 | 8 | 16 | 21 | 23 | 10 | 10 | 16 |

Use a 5 point moving average to estimate the sales of cans for week 4.

10. The amount of fuel used in a school in the 4 seasons is shown in the table below (in 1000s of litres).

| *1997* | | | | *1998* | | | |
|---|---|---|---|---|---|---|---|
| Spring | Summer | Autumn | Winter | Spring | Summer | Autumn | Winter |
| 5.3 | 4.4 | 5.4 | 7.3 | 6.6 | 5.6 | 6.5 | 8.3 |

Use an appropriate moving average to estimate the amount of fuel used each season in 1999.