Average & spread (7–9)

Contents

1 Averages (7)

There are three types of average which are used to find one number which will be a good representation for a set of data.

Mean. Add up all of the numbers and divide by how many there are. E.g. The mean of 3, 4, 7, 9, 11 is

$$\frac{3+4+7+9+12}{5} = \frac{35}{7} = 7.$$

Mode. This is the most frequent number in a list. E.g. The mode of 3, 4, 4, 6, 6, 6, 7, 8, 9 is 6 since it appears three times.

Median. : This is the middle number in a set of ordered numbers: E.g to find the median of 5, 2, 7, 9, 8, put them in ascending order first:

$$2 \quad 5 \quad \mathbf{7} \quad 8 \quad 9$$

So the median is 7.

Having an average on its own to describe a set of data is often not enough. For instance, imagine you were trying to decide where to go on a summer beach holiday and you found two resorts that you liked, both with average temperature of 31°C: which one would you go to? This information on its own is not good enough to make a decision. To back it up, it would be good to know how spread the daily temperatures are.

2 Spread (7)

To find the spread of numbers we generally use the range.

Range. This is the difference between the highest and lowest value.

$$\text{Range} = \text{Highest value} - \text{Lowest value}$$

E.g. The range of 2, 4, 6, 9, 13, 21 is $21 - 2 = 19$.

So, let's take these two resorts you are looking at for your holiday. Imagine that you now know the following:

	Average	Range
Resort A	31°	1°C
Resort B	31°C	10°C

Now, we don't know exactly what the daily temperatures are, but we could predict something like the following:

Resort A	31	31	31.5	30.5	31	30.5	31
Resort B	31	31	36	31	26	28	34

We notice that the temperatures in resort A differ by 1°C (that is, 30.5 up to 31.5) and those in resort B differ by 10°C (that is, 26 up to 36).

Maybe you would prefer resort B since you get some very, very hot days but you also run the risk of some much cooler days. Resort A is probably preferable since the temperatures do not vary too much. That is, they are more consistent. You are assured of lovely temperatures every day.

3 Frequency tables (7)

Since lists of data can be very long, they could be organised into a frequency table which is more concise. Imagine a shop recorded how much it took on sweets each hour during the course of two working days:

£5	£5	£10	£5	£10	£15	£20	£5
£10	£15	£20	£5	£5	£5	£10	£5

These amounts can more clearly be recorded as shown below:

Amount (£)	Frequency
5	8
10	4
15	2
20	2

We can find each average and the range in a more efficient way from this table, rather than referring back to or writing out the original list.

3.1 Mean from a frequency table

Consider the hours in which different amounts were taken

£5	5	5	5	5	5	5	5	5
£10	10	10	10	10				
£15	15	15						
£20	20	20						

To find the mean we add these up first:

$$(5+5+5+5+5+5+5+5)+(10+10+10+10)+(15+15)+(20+20).$$

However, we see that this is the same as:

$$(5\times 8)+(10\times 4)+(15\times 2)+(20\times 2).$$

So we could total these more easily by multiplying across each row in our grid.
We then have to divide by how many data items there were — since there were 8 £5s, 4 £10s, 2 £15s and 2 £20s, this gives 16 data items altogether. This can be worked out quickly by simply adding the frequencies:

$$8+4+2+2=16.$$

Hence, the overall calculation can be set out clearly as follows:

Amount (£)	Frequency	Total
5	8	$5\times 8=40$
10	4	$10\times 4=40$
15	2	$15\times 2=30$
20	2	$20\times 2=40$
	16	150

Therefore

$$\text{Mean}=\frac{150}{16}=9.375=\pounds 9.38$$

You can use more complex notation, but don't worry too much about this:

Amount (£x)	Frequency (f)	Total (fx)
5	8	$5\times 8=40$
10	4	$10\times 4=40$
15	2	$15\times 2=30$
20	2	$20\times 2=40$
	16	150

Therefore

$$Mean=\overline{x}=\frac{\sum fx}{\sum f}=\frac{150}{16}=9.375=\pounds 9.38.$$

3.2 Mode from a frequency table

In most hours £5 was taken — this happened 8 times, more than any other amount:

$$\text{Mode}=\pounds 8.$$

3.3 Median from a frequency table

If there are 16 values, the middle occurs between the 8th and 9th position. Eight data values occur in the £5 row so the 9th is in the £10 row, making the median between £5 and £10. That is,

$$\text{Median} = \pounds 7.50.$$

If this seems too complicated, you can always write out the list in full:

$$5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \,|\, 10 \quad 10 \quad 10 \quad 10 \quad 15 \quad 15 \quad 20 \quad 20$$

3.4 Range from this frequency table

The lowest amount is £5, the highest is £20 so:

$$\text{Range} = 20 - 5 = \pounds 15.$$

4 Grouped frequency tables (8 & 9)

It may be appropriate to put data into groups. For example, imagine the ages of various people at a concert:

$$4 \quad 9 \quad 12 \quad 13 \quad 15 \quad 15 \quad 16 \quad 16 \quad 17 \quad 17 \quad 18 \quad 25 \quad 33 \quad 35 \quad 38 \quad 40 \quad 59$$

If we put these into a table, there would be 14 different rows in this table since there are 14 different ages. Grouping similar people together would be efficient (try and use equal groups and about 5-6 groups):

Age	Frequency
$0 - 10$	2
$10 - 20$	9
$20 - 30$	1
$30 - 40$	3
$40 - 50$	1
$50 - 60$	1

N.B. Notice how we rounded the 40 up into $40 - 50$, not $30 - 40$.

Comment. Perhaps this is a music concert designed for teenagers, but some of them have gone with their parents and some younger brothers or sisters.

4.1 Mean from this grouped frequency table

This is similar to mean from a frequency table, but we don't have an actual value to work with from each group. It seems sensible to choose the midpoint.

Age	Midpoint (x)	Frequency (f)	Total (fx)
$0-10$	5	2	$5\times 2 = 10$
$10-20$	15	9	$15\times 9 = 135$
$20-30$	25	1	$25\times 1 = 25$
$30-40$	35	3	$35\times 3 = 105$
$40-50$	45	1	$45\times 1 = 45$
$50-60$	55	1	$55\times 1 = 55$
		$\sum f = 17$	$\sum fx = 375$

Therefore:

$$\text{Mean} = \overline{x} = \frac{\sum fx}{\sum f} = \frac{375}{17} = 22.0588\cdots = 22.1 \text{ years (to 1 d.p.).}$$

Note. Since we have used midpoints to represent each group, our answer will only be an estimate.

4.2 Mode from a grouped frequency table

We can't find a mode but a modal group. Most people fell in the $10-20$ age category.

The modal group is $10-20$ years.

4.3 Median from a grouped frequency table

This is too difficult to pinpoint exactly so we need to use a cumulative frequency graph to do so (see these notes, year 9 only).

4.4 Range from a grouped frequency table

We could estimate the range as $60-0 = 60$ years. However, we see, by referring to the original list, it is actually $59-4 = 55$. So, our answer from a table will only ever be an estimate since we would not always have the raw data.