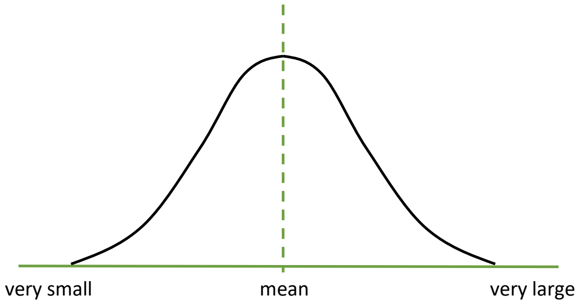
Standard Deviation

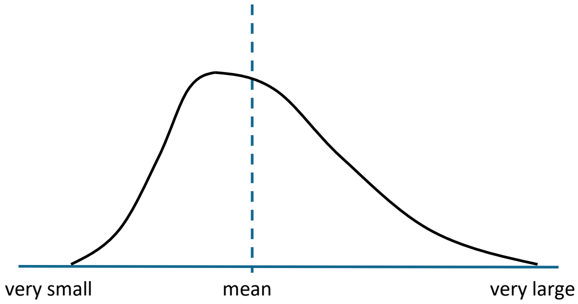
A set of measurements can be analysed in a number of ways. The mean, mode and median all inform the reader of a form of the 'middle value'. However, these 'middle values' do not allow the reader to see how a data set is spread amongst the full range of values available.



Size of bedload samples from a stream

If the spread of the data is even either side of the mean then we say it shows **normal distribution**.

For example, in a sample of pebbles taken from an average stream, one might expect most pebbles to be roughly the same size, with a few being much smaller than the mean size and a few being much larger.



Size of bedload samples from a stream

If the spread of the data is greater to one side of the mean than the other we say it shows an **uneven or skewed distribution**.

For example, in a sample of pebbles taken from a particular part of a stream it is likely that the spread of the data will be more varied or clustered in one part of the distribution.

Drawing a distribution curve is one way to show how the spread of the observed data differs from the mean. A more precise way is to use a calculation known as **Standard Deviation** (shown by the symbol σ). Standard deviation gives a numerical value to the 'spread' of the data. The greater the variability of the data, the larger the standard deviation value.

Worked example:

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

x = the sample data measure

$\bar{x}$ = the mean sample data measure

n = the number of samples

A student collected two sets of bedload, each a set of ten samples. The first was from the centre of a stream on a meander bend and the second was from the inner bank of the same meander. For each piece of bedload, the size of the longest axis was measured. The student wished to see if there was a difference in how each set of samples was distributed about the mean size of the bedload.

Data set 1 (centre of the stream)

Bedload 1:	42mm	Bedload 6:	34mm
Bedload 2:	56mm	Bedload 7:	11mm
Bedload 3:	30mm	Bedload 8:	68mm
Bedload 4:	18mm	Bedload 9:	35mm
Bedload 5:	52mm	Bedload 10:	24mm

Therefore, the mean bedload length ($\bar{x}$) was 37mm.

Data set 2 (inner bank of the meander)

Bedload 1:	5mm	Bedload 6:	35mm
Bedload 2:	39mm	Bedload 7:	12mm
Bedload 3:	20mm	Bedload 8:	7mm
Bedload 4:	12mm	Bedload 9:	44mm
Bedload 5:	71mm	Bedload 10:	16mm

Therefore, the mean bedload length ($\bar{x}$) was 26.1mm.

The initial results from the mean scores appear to show that the bedload on the inner bend is smaller than at the centre of the meander. However, to also show the degree of distribution around these means, the student applies the Standard Deviation calculation.

Data set 1 (central stream):

x	$(x - \bar{x})$	$(x - \bar{x})^2$
42	5	25
56	19	361
30	-7	49
18	-19	361
52	15	225
34	-3	9
11	-26	676
68	31	961
35	-2	4
24	-13	169
	Σ =	3250

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

$$\sigma = \sqrt{\frac{3250}{10 - 1}}$$

$$\sigma = \sqrt{361.1}$$

σ = 19.00

Data set 2 (inner bank):

x	$(x - \bar{x})$	$(x - \bar{x})^2$
5	-21.1	445.21
39	12.9	166.41
20	-6.1	37.21
12	-14.1	198.81
71	44.9	2016.01
35	8.9	79.21
12	-14.1	198.81
7	-19.1	364.81
44	17.9	320.41
16	-10.1	102.01
	Σ =	3928.9

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

$$\sigma = \sqrt{\frac{3928.9}{10 - 1}}$$

$$\sigma = \sqrt{436.5}$$

σ = 20.89

The standard deviation results indicate that although the inner bank has a smaller mean size of bedload, the samples taken there show a wider variation in size against the mean. This means that bedload from the centre of the stream shows more uniformity in size compared to the inner bend.