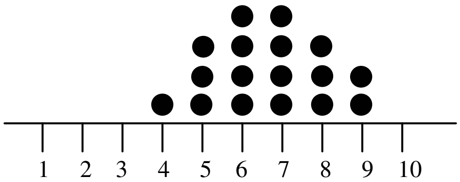
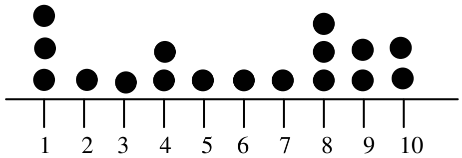
Standard Deviation:

Consider these two dotplots of data





Data: **Set 1** **Set 2**

{ 4, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9 } { 1, 1, 1, 2, 3, 4, 4, 5, 6, 7, 8, 8, 8, 9, 9, 10, 10}

Clearly, the data in set 2 is spread out further than that in set 1.

We are going to look at a method of assigning a number to this idea of the spread of the data.

It is called the **standard deviation**.

Method:

Let us take a smaller data set for illustration.

E.g. Sample of prices of milk in local supermarkets: { 94, 96, 101, 101, 102 }

We draw up a table:

	x		
	94		
	96		
	101		
	101		
	102		
TOTAL	494		

The mean $\bar{x}$ is $\frac{\sum x}{n}$ i.e. $494 \div 5 = 98.8$ Now calculate the deviation from the mean: $x - \bar{x}$

	x	$x - \bar{x}$	
	94	–4.8	
	96	–2.8	
	101	2.2	
	101	2.2	
	102	3.2	
TOTAL	494		

Standard Deviation:

Next, we square the deviation column to get $(x-\bar{x})^2$ Remember this will always be **positive**.
and then we add up this column of $(x-\bar{x})^2$

	x	$x-\bar{x}$	$(x-\bar{x})^2$
	94	–4.8	23.04
	96	–2.8	7.84
	101	2.2	4.84
	101	2.2	4.84
	102	3.2	10.24
TOTAL	494		50.8

We now have all the sum of all the deviations squared.

We are now going to average these, by dividing by the number of items (5 in this case)

So we get: 50.8 ÷ 5 = 10.16, but remember that this is the deviation squared.

So now we take the square root of this figure $\sqrt{10.16} = 3.187...$

We have effectively taken the average of the deviations.

This is called the **standard deviation** and is denoted by *s* or σ (greek letter – sigma)

Important.

In all cases at Standard Grade we will be working with **sample** data sets.

It has been found that a more accurate result with samples is found by dividing the deviation squared by ***n* – 1** instead of *n*.

We can express this in a formula – this is given on the inside cover of the examination paper.

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$ which is the result we have just found.

There is another formula which is more useful when using computers or a large data set. This gives the same result, and may save some calculations.

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

In both cases, n is the sample size (the number of data items)

Standard Deviation:

Example:

Fiona checks out the price of a litre of milk in several shops.

The prices in pence are:

49 44 41 52 47 43

a) Find the mean price of a litre of milk.

b) Find the standard deviation of the prices.

c) Fiona also checks out the price of a kilogram of sugar in the same shops and finds that the standard deviation of the prices is 2.6.

Make one valid comparison between the two sets of prices.

	x	$x-\overline{x}$	$(x-\overline{x})^2$
TOTAL			

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

You try this one on your own:

A group of fifth year students from Scotia High School were asked how many hours studying they did in the week prior to their exams.

The results are shown below.

13 8 10 11 18 9 15

(a) Use an appropriate formula to calculate the mean and standard deviation of these times.

(b) A similar group of students from Scotia Academy were asked the same question

The mean number of hours studied was 14 and the standard deviation was 2·8.

How did the number of hours studied by students from Scotia High School Compare with the number of hours studied by students from Scotia Academy ?

TOTAL			

Now try it using the other formula.

	x	x^2
TOTAL		

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

$\sum x =$ $(\sum x)^2 =$ $\sum x^2 =$

$n =$ $s =$