Average and Spread

To keep pace with market trends many organisations such as manufactures, employers, college's etc need data to find out if what they are making or delivering needs to be modified or withdrawn. Data is usually collected by undertaking a survey and asking people to complete a questionnaire.

Somewhere along the line a person or group of people (Statisticians) have to sit down and make sense of the data collected. In a perfect world a survey would ask a question of everyone that it needed to in order to produce100% accurate results.

In reality this is rarely possible. Suppose you were a T.V. producer and you wanted to find out what proportion of the population liked watching sport on television, to get an accurate result you would ask everyone with a television to complete an appropriate questionnaire.

This would be an enormous task, which would take a very long time, and be very costly.

In general you would take a sample of the population (say a 1000 people) and measure the responses of this group of people.
Once you have received all the data you need, you then need how to interpret this data to draw some meaningful conclusions.
You need to be aware of the pitfalls of jumping to the wrong conclusion.

Example 1

Lets suppose you asked 10 households the simple question "how many times a week is the television switched to a sports programme"?

The data received was: 0 2 2 4 1 3 3 3 3 9.
There are three different ways of finding an average value that describes this data.

The mean

This is calculated by adding up all of the values and dividing this result by the number of values measured.

$$\frac{0 + 2 + 2 + 4 + 1 + 3 + 3 + 3 + 3 + 9}{10}$$

$$\frac{30}{10} = 3$$

The mode
This is the value that occurs most often. The data received was: 0 2 2 4 1 3 3 3 3 9

The number that occurs most by inspection is 3, it occurs 4 times.

The median
This is the middle value when all data is arranged in order of size. The data received was: 0 2 2 4 1 3 3 3 3 9

Arranged in order this becomes: 0 1 2 2 3 3 3 3 3 9

There are two numbers that sit either side of the middle of this group of numbers, in this case you add the two numbers together and divide by 2. Hence, the median is $\frac{3 + 3}{2} = 3$

	Mode	Median	Mean
Advantages	Very easy to find Not affected by extreme values. Can be used for non numerical values.	Easy to find for ungrouped data. Not affected by extreme values	Easy to find. Uses all the values. The total for a given number of values can be calculated from it.
Disadvantages	Doesn't use all the values.	Doesn't use all the values. Often not understood.	Extreme values can distort it. Has to be calculated.
Used for	Non-numerical data. For finding the most likely value.	Data with extreme values	Data whose values are spread in a balanced way.