Central Tendency

When analysing a group of data, it is often useful to consider what would be a 'typical' value. The typical value of a set of data is also known as the central tendency of the data.

The Mode

Perhaps the simplest way of making an estimate of the typical value of a set of data is to take it as the value that occurs most often. This is known as the mode, or modal value of the data. For example, consider the following set of values:

1,1,2,3,4,4,4,4,5,6,6,7,7,8,9

Here 4 occurs most frequently. The mode of the data is therefore 4.
There are several problems with the mode as an average measure, as shown in the following example:

1,1,3,5,10,12

Here the modal value is 1, but it is clear that 1 is not a good estimate of the typical value of the data.

Instinctively, we understand the typical value of a set of data to be 'in the middle'. We must therefore consider how we can find the middle value of our data.

The Median

One way to find the middle value is by writing out all the data in order and then taking the value that lies in the middle of the list. For example, for a set of 7 statistics, we would take the fourth highest value. This is known as the median of the data.
Where the number of values on a set of data is even, a figure half way between the two central values is taken as the mode.
Consider the following set of data:

2,5,3,7,8,6,9,3,1,2

Writing this in order, we get:

1,2,2,3,3,5,6,7,8,9

The two middle values are 3 and 5, so we take 4 as the median of the data.

This is clearly a better estimate of typical value than the mode. However, it fails to take into account the actual values of the data, only their order, as illustrated in the following example:
2,3,5,6,15,18,24

Here the median of the data is 6, but we instinctively feel that a typical value should be higher, to take into account the highest three statistics.
This leads us to the third of our measures of central tendency.

The Mean

The mean $\overline{x}$ of a set of data $x_1, x_2 \ldots x_n$, of size n is defined as follows:

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$
$$= \frac{\sum x_i}{n} \text{ for } i = 1, 2, \ldots, n$$

Note: The symbol $\sum$ means "sum of" and is read "sigma".
So $\sum x_i$ for $i = 1, 2, \ldots, n$ means "the sum of the values $x_1, x_2, \ldots, x_n$ "i.e. $x_1 + x_2 + \ldots + x_n$

You may encounter the following alternative notations, which mean the same thing:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ or } \overline{x} = \frac{1}{n}\sum x \text{ or } \overline{x} = \frac{\sum x}{n}$$

The mean is what people usually think of when considering the 'average' of a set of data.

Example (1)

Find the mean of the following set of numbers;

24,27,31,32,38

We can write down immediately the fact that $n = 5$, as there are five numbers.

$\sum x = 24 + 27 + 31 + 32 + 38 = 152$

$\bar{x} = \frac{\sum x}{n}$

$= \frac{152}{5}$

$= 30.4$

The mean of the set of numbers is 30.4

Example (2)

To obtain a grade A in maths, Samuel must obtain an average of at least 70 over five exams. His marks in the first four exams are as follows:

67,65,72,70

What mark must he obtain in the fifth exam to obtain a grade A overall?

The mean of the five scores must be 70, i.e.

$\frac{\sum x}{5} = 70$

$\sum x = 350$

$67 + 65 + 72 + 70 + x_5 = 350$, where x_5 is Samuel's score in the fifth exam.

Rearranging:

$x_5 = 350 - 67 - 65 - 72 - 70$

$= 76$

To obtain a grade A, Samuel must get a score of at least 76 in his fifth exam.