Measures of dispersion

Measures of central tendency tell you about how spread out (dispersed) or concentrated data are.

Open the **Measures of data dispersion** Excel workbook.

Conditional formatting

Excel allows you to format cells and text based on any formula you wish - **Conditional Formatting**.

Using data from Keswick, the tables below have been conditionally formatted to show the inter-quartile range of the data.

Temperature

Blue cells are the coldest 25%

Orange cells are the warmest 25%

Green cells are the interquartile range (middle 50%)

Rainfall

Light blue cells are the driest 25%

Dark blue cells are the wettest 25%

Green cells are the inter-quartile range (middle 50%)

White cells have no data



Calculating values

Excel can calculate measures of dispersion / central tendency very quickly and easily.

The examples below are from the temperature calculations on the spreadsheet.

1. To find the first quartile, containing the bottom 25% of data, use =QUARTILE.INC(B4:K15,1). BK:K15 is the range of the data you're using, and 1 asks for the value of the first quartile.
2. To find the third quartile, containing the top 25% of data, use =QUARTILE.INC(B4:K15,3). The range is the same, the 3 asks for the value of the third quartile.
3. To calculate the inter-quartile range (IQR) that contains the middle 50% of the data you subtract the first quartile from the third quartile =O5-O4.
4. To calculate the average use the formula =AVERAGE(B4:K15).
5. To calculate the mode of the data use the formula =MODE.SNGL(B4:K15). The mode says #N/A because there is no modal value.
6. To calculate the median of the data use the formula =MEDIAN(B4:K15).

**Challenge Yourself – Standard Deviation**

(this is not an examinable skill)

The Standard Deviation (SD) is a measure of how spread out the data are, i.e. how far away from the average (mean) value they go. To calculate SD the formula to use is =STDEV.P(B4:K15). If you were only using a sample of data (i.e. not all of it), you'd use =STDEV.S().



Interpreting Standard Deviation

The diagram below graphically represents how it works. Your data is all in the blue area.

68.2% of your data will be within **one** SD of the mean.

A further 27.2%, making 95% in total will be within **two** SD of the mean.

A further 4.2%, making 99.2% in total will be within **three** SD of the mean.

So, if you have data with a mean of 100, and a SD of 1, the data are very closely bunch together because over 99% are between 97 and 103.

If you have data with a mean of 100, and an SD of 30, the data are much more spread out. Only 68.2% of your data are between 70 and 130!

Be careful - you can't just remember that a small SD means data are bunched together and vice versa - you have to think about the size of the SD in relation to the mean.

