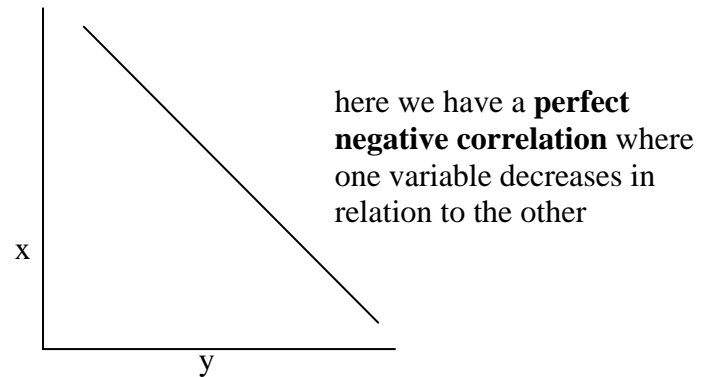
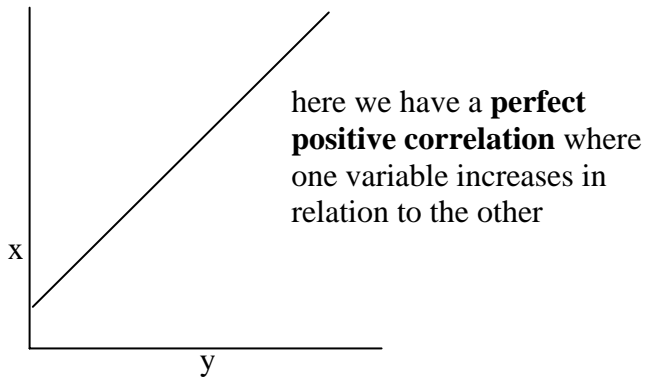
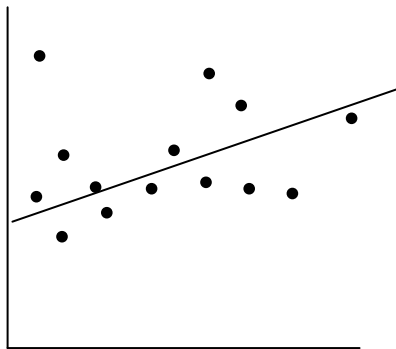


Looking at relationships: correlation analysis

Correlation analysis aims to establish whether there is a significant relationship between two sets of data. Correlations range between **perfect positive** and **perfect negative**.



However in real life it is highly unlikely that all points will fall on a straight line at 45° ! Usually there is a scatter of points and even when a 'line of best fit' is applied it may not be possible to decide whether there is a significant correlation. A **scatter diagram** is a useful tool here.



The **line of best fit** can be inserted by eye to achieve a good balance of points on either side of it. If it is difficult to decide on the path of the line, calculate the mean of both sets of variables and ensure that the line passes through this point.

Although trends in the data can be identified in this way, we need to have some confidence that there really is a significant correlation –a correlation at the 5% probability level (that is, there is less than 5% probability that the relationship has happened by random chance).

Correlation analysis enables us to decide statistically whether there is a significant correlation between two sets of data or not. **Numerical correlation analysis (Pearson's Product-moment test)** can only be used when two data sets are at the **continuous level** of measurement (mm, sec, g, numbers of individuals). However **rank correlation analysis** can be used when two sets of data obtained by categorical or ordinal measurement need to be compared.