Core Maths Revision notes

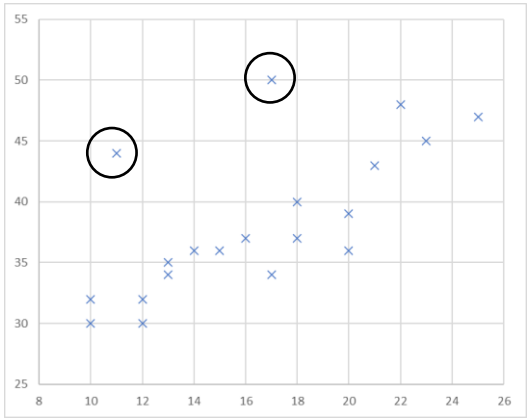
Correlation and Regression

For AQA 2A

Compiled by Dave Gale

Are things connected?

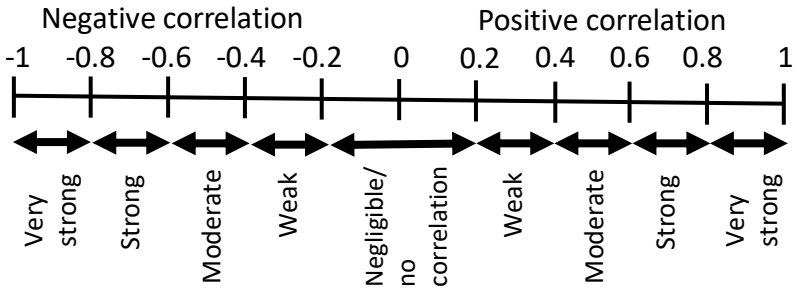
Scatter graphs are very useful to look for correlation.
This one has a strong positive correlation with two outliers circled.



We measure correlation with a coefficient.
These are all names for the same thing:

- **Correlation Coefficient**
- **Product Moment Correlation Coefficient**
- **r**

The correlation coefficient, r, can be between -1 and 1
The words below are a **rough guide**.



Correlation does not imply causation.
Just because things are correlated, doesn't necessarily mean one causes the other to happen

Positive correlation
As one value increases, the other increases
Negative correlation
As one value increases, the other decreases

Outliers
Can have a big impact on your value of r.
Treat them carefully!

Linear Correlation
This course only looks at correlation that gives straight lines. Other types do exist!

Using a line of best fit to make predictions

If the correlation is strong, then a regression line can be used to make predictions.

Regression line is the phrase for the correct **line of best fit**
It might be written as "the regression line of y on x"

You'll be asked to **find the equation** of the regression line and probably be expected to **plot** and **use it**.

$$y = a + bx$$

Useful fact:
The regression line goes through the point

$$(\bar{x}, \bar{y})$$

a is the intercept

b is the gradient

You need to be able to interpret the values of a and b in the context of the question.
a is what happens when $x = 0$.
Look for what that would mean in the question you're looking at.
b is the gradient. It tells you how quickly the y values are changing.
You can interpret it with this **customisable phrase**:
"For every one extra [x thing], the [y thing] [increases/decreases] by [b]."

The variables might not be x and y.
Just replace them with whatever is used in the question.

The values of r, a and b can all be found directly from your calculator.
Exactly how depends on your make of calculator so you must find out how to do it.

Google "correlation on [your calculator model]"