Calculating Statistical Significance

The correlation coefficient is a number between −1 and +1 that **shows how closely what you tested is related to what you changed**. +1 means the two always occur together. -1 means they never occur together. We are going to use this number to calculate statistical significance, which will tell us what the chances are that the results you saw were random. Scientists like a significance of 5% or less. That means there is only a 1 in 20 chance the data you saw just randomly looked like your experiment proved something. When that happens, you can report that you got a statistically significant result.

Step 1: Calculating the Correlation Coefficient

1. Record **x** (the control variable, or the one you didn't changed to run the experiment) and **y** (the dependent variable, or the amount of thing you tested for).
2. Calculate **x** times **y** for each.
3. Calculate x^2 for each line.
4. Calculate y^2 for each line.
5. Sum up **x**, and put your answer in the total line at the bottom of the x column.
6. Do the same with **y**, **x·y**, x^2 and y^2 so all the totals are filled in.

η is the number of trials in your experiment

	x	y	X·Y	X^2	Y^2
1					
2					
3					
...					
η					
Totals ->					

Σ means total in math, so you can just plug numbers into the formula and punch it into your calculator.

$$ r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}} $$

r is the correlation coefficient, a number between +1 and -1 that tells you how closely what you changed is related to what you saw.

Please be careful to remember that **it is not necessarily the case that one caused the other**. Both the independent variable and the dependent variable may have been caused by something else or be parts of something bigger.
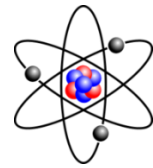
Step 2: Testing the Null Hypothesis

Testing the null hypothesis is the math way of saying we are going to find out if what we saw was just happened randomly. To do this we are going to need **r**, the correlation coefficient and **η**, the number of trials you conducted from step 1. Just plug the numbers into this formula. The number **t**, is the statistical significance of your findings.

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

If you get a number less than 0.05, you have a **statistically significant** result. Here is a work sheet to help:

1	η - 2 =	
2	r =	
3	r^2 =	
4	1 - r^2 =	
5	box1 / box4 =	
6	Sq root box5 =	
7	R times box 6 =	