Spearman's Rank Correlation Coefficient – Excel Guide

What is Spearman's Rank correlation coefficient?
Spearman's Rank correlation coefficient is used to identify and test the strength of a relationship between two sets of data. It is often used as a statistical method to aid with either proving or disproving a hypothesis e.g. the depth of a river does not progressively increase the further from the river bank.

The formula used to calculate Spearman's Rank is shown below.

$$r = 1 - \frac{6 \sum d^2}{n^3 - n}$$

NB. Sometimes $n^3 - n$ is written as $n(n^2 - 1)$. Both mean the same thing.

How can the calculation be carried out in Excel?
Once the data has been collected, Excel can be used to calculate and graph Spearman's Rank correlation to discover if a relationship exists between the two sets of data, and how strong this relationship is. Please note this example uses a dataset of 10 samples, but your dataset should include a minimum of 15 to be valid.

Step 1: Create a table in Excel and enter your data sets.

Sample	Width (cm)	Width Rank	Depth (cm)	Depth Rank
1	0		0	
2	50		10	
3	150		28	
4	200		42	
5	250		59	
6	300		51	
7	350		73	
8	400		85	
9	450		104	
10	500		96	

[1]

---

[1] Data for this example has been taken from www.geography-site.co.uk/pages/skills/fieldwork/stats/spear.html

Step 2: Rank each set of data (width rank and depth rank). Rank 1 will be given to the **largest** number in column 2. Continue ranking till all widths have been ranked. Once all the widths have been ranked then do exactly the same for depth.

Sample	Width (cm)	Width Rank	Depth (cm)	Depth Rank
1	0	10	0	10
2	50	9	10	9
3	150	8	28	8
4	200	7	42	7
5	250	6	59	5
6	300	5	51	6
7	350	4	73	4
8	400	3	85	3
9	450	2	104	1
10	500	1	96	2

If there are two samples with the same value, the mean (average) rank should be used – for example if there were 3 samples all with the same depth ranked 6th in order you would add the rank values together (6+7+8=21) then divide this number by the number of samples with the same depth number, in this case 3 (21/3=7) so they would all receive a rank of 7. The next greatest depth would be given a value of 9.

Step 3: The next stage is to find d (the difference in **rank** between the width and depth). First, add a new column to your table, and then calculate d by subtracting the depth rank column (column 5) from the width rank column (column 3). For example, for sample 6 width rank is 5 and the depth rank is 6 so $d = 5 - 6 = -1$. To calculate d in Excel, select the cell you wish to enter the information into and type =. Now click on the width rank cell you want to use and type –. Finally, click on the depth rank cell and press enter. The value of d should appear in the first box you selected.

Sample	Width (cm)	Width Rank	Depth (cm)	Depth Rank	d
1	0	10	0	10	0
2	50	9	10	9	0
3	150	8	28	8	0
4	200	7	42	7	0
5	250	6	59	5	1
6	300	5	51	6	-1
7	350	4	73	4	0
8	400	3	85	3	0
9	450	2	104	1	1
10	500	1	96	2	-1

Step 4: The next step is to calculate d^2. Add another column to your table and label it. To calculate d^2 type in the first cell =POWER(number,power). In this case the number is the value of d and the power is 2 as we are trying to find the square value e.g. for sample 6 the value of d is -1 so you would enter into the cell =POWER(-1,2) then press enter and the value you should get is 1.

Sample	Width (cm)	Width Rank	Depth (cm)	Depth Rank	d	d^2
1	0	10	0	10	0	
2	50	9	10	9	0	
3	150	8	28	8	0	
4	200	7	42	7	0	
5	250	6	59	5	1	
6	300	5	51	6	-1	1
7	350	4	73	4	0	
8	400	3	85	3	0	
9	450	2	104	1	1	
10	500	1	96	2	-1	

Repeat the same process until all of your samples have a value of d^2. Once all the d^2 values have been calculated add them together to calculate Σd^2. The quickest way to do this in Excel is to click on the cell underneath your last entry into the d^2 column, click on the autosum symbol Σ (which you can find on the tool bar at the top of the page), and press enter. (Depending on which version of Excel you are using, you may have to select the column you wish to add together before you press enter.)

Step 5: Now we have the Σ d^2 values, but to complete the equation we still need to calculate n^3 - n. n is the number of samples, so in this case is 10. As in step 4, type into the cell you wish to use =POWER(number,power) which will give you a value for n^3. Remember this time that 'number' is the number of samples and 'power' is 3 as you are cubing not squaring the value. Once n^3 has been calculated, subtract the value of n from it.

Sample	Width (cm)	Width Rank	Depth (cm)	Depth Rank	d	d^2
1	0	10	0	10	0	0
2	50	9	10	9	0	0
3	150	8	28	8	0	0
4	200	7	42	7	0	0
5	250	6	59	5	1	1
6	300	5	51	6	-1	1
7	350	4	73	4	0	0
8	400	3	85	3	0	0
9	450	2	104	1	1	1
10	500	1	96	2	-1	1
					Σd^2	4
					n	10
					n^3	1000
					$n^3 - n$	990

Step 6: All that is left to do now is to insert the values into the equation to calculate r.

$$r = 1 - \frac{6 \sum d^2}{n^3 - n}$$

The formula you should enter into a cell in this case would be =1-((6*4)/990)

(N.B * is the symbol for multiply)

End result

The result at the end should always be between the value of +1 and -1. In this case the value is 0.9757 to 4 decimal places, or 0.98 to 2 decimal places.

Step 7: To see if this r value is significant, a Spearman's Rank significance table or graph must be used. To do this you need to calculate the degrees of freedom.

This is calculated using the sum n - 2, which in this example = 8 (remember n is the number of samples).
Using this number and your value for r you can use the table below to work out the significance level of your data (the yellow lines).

The significance of the Spearman's rank correlation coefficients and degrees of freedom
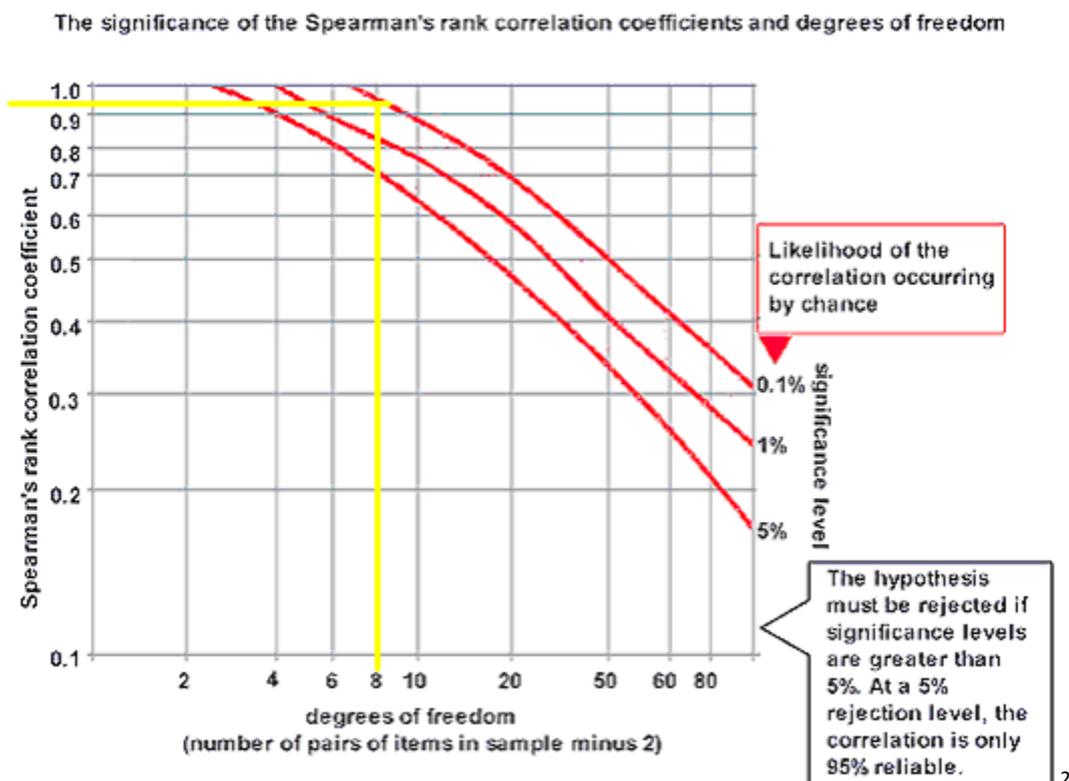
As you can see the yellow line meets the red line at 0.1% meaning that there is a greater than 99% chance that the relationship is significant and not random. Thus we can **reject** the hypothesis and conclude that in this case the depth of a river does progressively increase the further the distance from the river bank.