

Excel RANK & PERCENTRANK Problems

Excel also has a strange way to calculate ranks and percentiles. Excel's built-in RANK function does not take into account 'tied' ranks. For example, in a series of measurements 100, 120, 120, 125 Excel gives two times rank 2 to the value of 120 and value 125 gets the rank 4. When 'tied' ranks are taken into account, the rank of 120 should be $(2 + 3)/2 = 2.5$ and the value of 125 should indeed get rank 4. Excel assigned

	A	B	C
1	Score	Rank	Excel Rank
2	120	1	1
3	125	2.5	2
4	125	2.5	2
5	145	4.5	4
6	145	4.5	4
7	150	6.5	6
8	150	6.5	6
9	160	8	8
10	170	9	9
11	175	10	10

the lowest of the two ranks to both observations, giving each a rank of 2. Because Excel doesn't consider 'tied' ranks it is impossible to calculate the correct non-parametric statistics from the obtained ranks. For this reason, a User Defined Function will take into account 'tied' ranks.

Function Ranking(V As Double, R As Range) As Double

Dim No As Integer

Ranking = Application.WorksheetFunction.Rank(V, R, 1)

No = Application.WorksheetFunction.CountIf(R, V)

Ranking = Ranking + (No - 1) / 2

End Function

The way Excel calculates percentiles is also not the way most statistical packages calculate them. In general, the differences are most obvious in small data sets. As an example, let's take a sample of test scores from 10 students sorted in ascending order: 120, 125, 125, 145, 145, 150, 150, 160, 170, 175. The lower quartile (or 25% percentile) as calculated with Excel's built-in function QUARTILE (or PERCENTILE) is 130 and the upper quartile is 157.5. **A statistical package, however, will give 125 and 170 as lower and upper quartile, respectively. Apparently, Excel calculates the lower quartile $130 = 125 + (145-125)*0.25$ and the upper quartile as $157.5 = 150 + (160-150)*0.75$. This is an interpolation between the values below and above the 25% or 75% observation.** Normally, the pth percentile is obtained by first calculating the rank $l = p(n+1)/100$, rounded to the nearest integer and then taking the value that corresponds to that rank. In case of lower and upper quartiles, the ranks are $0.25*(10+1) = 2.75 \approx 3$ and $0.75*(10+1) = 8.25 \approx 8$ which corresponds to 125 and 170 resp.

A formula approach would be:

$$=(\text{COUNT}(\$A\$2:\$A\$11)+1+\text{RANK}(A2,\$A\$2:\$A\$11,1)-\text{RANK}(A2,\$A\$2:\$A\$11,0))/2$$

Reference data can be viewed at this site:

<http://www.coventry.ac.uk/ec/~nhunt/pottel.pdf>