

How to Easily Perform a Nested ANOVA in Excel

Authored by
stats writer

December 6, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a Nested ANOVA in Excel*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=106256>

Introduction to Nested ANOVA: Understanding the Concept

A Nested ANOVA, also known as a Hierarchical Analysis of Variance, is a powerful statistical technique used when factors in an experiment are arranged in a hierarchical structure. Unlike standard factorial designs where all factors cross each other, in a nested design, one or more levels of a Factor are contained entirely within the levels of another primary Factor. This structure means that not all combinations of factor levels are possible or observed, leading to a unique interpretation of variance components. Understanding the distinction between crossed and nested factors is crucial for correctly modeling experimental data, especially in industrial or biological studies where logistics often dictate hierarchical sampling.

The core purpose of the Nested ANOVA is to determine if significant variance exists at the level of the primary, or "nesting," factor, and also if there is significant variability introduced by the secondary, "nested," factor within those primary levels. Failing to recognize a nested design and treating it as a standard two-way interaction can lead to inaccurate hypothesis testing and inflated error terms. This statistical method helps isolate variation attributable to different levels of the hierarchy, providing a clearer picture of where the effects are truly originating.

Unfortunately, most statistical software packages, including **Microsoft Excel**, do not offer a direct, built-in function specifically labeled "Nested ANOVA." Therefore, analysts must employ a strategic workaround, utilizing the existing **Anova: Two-Factor With Replication** tool and performing subsequent manual calculations. This detailed guide demonstrates precisely how to navigate these steps within the Excel environment, ensuring that the critical adjustments needed for accurate interpretation of the nested model are properly executed.

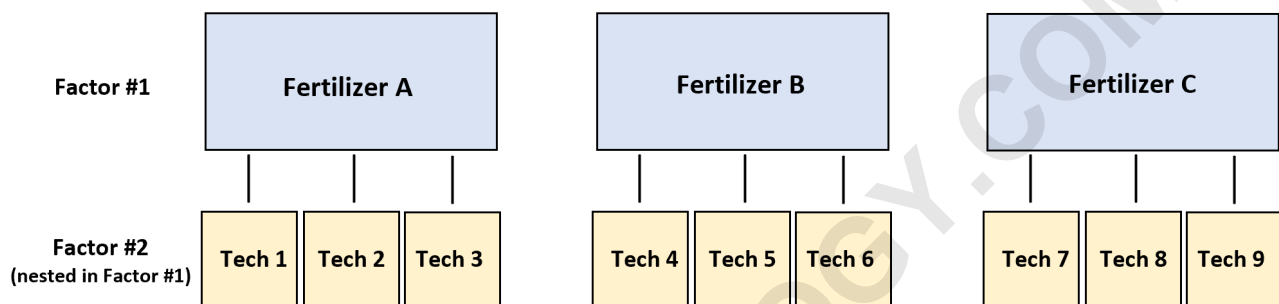
The Case Study: Fertilizer and Technician Experiment

To illustrate the application of a Nested ANOVA, consider a common research scenario in agricultural science. Suppose a researcher aims to investigate whether three distinct types of fertilizer (A, B, and C) yield different average levels of plant growth. To conduct this experiment efficiently and account for potential variability in application, the researcher employs nine technicians in total--three technicians assigned exclusively to Fertilizer A, three to Fertilizer B, and three to Fertilizer C. This specific setup creates the nested structure we need to analyze.

In this experimental design, the primary response variable--the variable being measured--is **plant growth**. The two factors under investigation are **Fertilizer Type** and **Technician**. The crucial observation here is that the Technician factor is entirely contained, or *nested*, within the Fertilizer factor. Technician 1, 2, and 3 who apply Fertilizer A are distinct and independent from Technician 4, 5, and 6 who apply Fertilizer B. A specific technician never applies more than one type of fertilizer. If the technicians were crossed with the fertilizers (meaning every technician applied every fertilizer), it would be a standard two-way ANOVA. Because of the exclusive assignment, the

nested relationship holds true.

Specifically, the researcher instructs each of the three technicians assigned to a given fertilizer to sprinkle that fertilizer on four individual plants. This results in a total of 36 plants (3 fertilizers * 3 technicians/fertilizer * 4 plants/technician = 36 observations). The visual representation of this hierarchical dependency clearly shows how technicians are grouped under fertilizers, reinforcing that the effect of the technician can only be assessed *within* the level of fertilizer they were assigned.



The subsequent steps will utilize this precise data structure to demonstrate how to execute the necessary calculations in **Excel**, ultimately allowing us to statistically test whether fertilizer type is significant and, separately, whether there is significant variability among the technicians *within* each fertilizer type.

Step 1: Preparing Your Data in Excel

The first and most critical step in performing a Nested ANOVA simulation in Excel involves correctly structuring your raw data. Although the data naturally falls into three fertilizer groups (A, B, C), the Excel tool we use (Two-Factor With Replication) assumes a crossed design. To force it to recognize the nested structure implicitly, we must organize the data such that the nested factor (Technician) is clearly represented, while ensuring the primary factor (Fertilizer) dictates the overall column grouping.

For the Two-Factor with Replication tool to work, the data must be organized into distinct columns representing the primary factor levels (Fertilizer A, B, C) and rows representing the individual replicates (plant growth measurements) organized by the nested factor (Technician). Since Technician is nested, we must label the rows such that each technician is distinct within their fertilizer group. If Technician 1, 2, and 3 applied Fertilizer A, they appear as three separate groups within the "Fertilizer A" column. Since we have four replicates per technician, each row grouping must contain exactly four entries.

The structured input must follow this strict format. The column headers should represent the

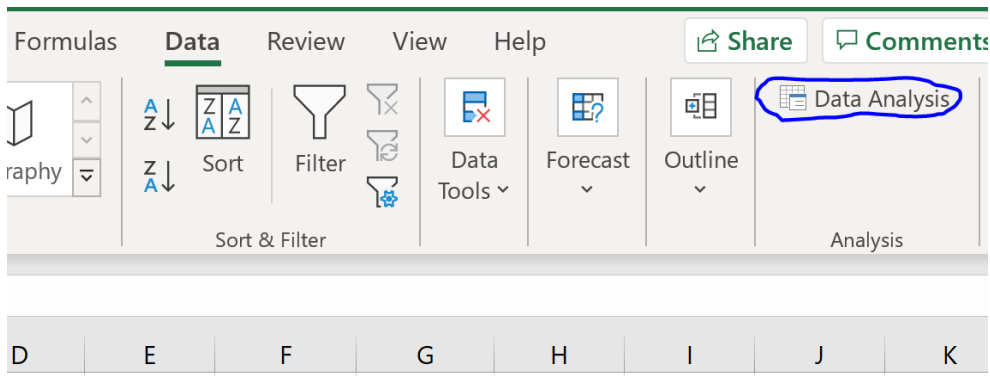
primary factor levels (A, B, C), and the rows should be labeled to indicate the unique nested combinations (Technician 1 within A, Technician 2 within A, etc.). This specific layout is required by the **Anova: Two-Factor With Replication** function, which expects the number of samples (replicates) to be uniform across all row groups and columns. Careful attention must be paid to ensure the counts are accurate, as any deviation will cause the analysis tool to fail or produce incorrect results.

	A	B	C	D	E	F	G
1		Technician 1	Technician 2	Technician 3			
2	Fertilizer A	13	15	15			
3		16	16	15			
4		16	19	12			
5		12	16	15			
6	Fertilizer B	19	23	19			
7		19	18	20			
8		20	16	21			
9		22	18	21			
10	Fertilizer C	21	25	24			
11		23	20	22			
12		24	20	25			
13		22	22	26			
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							

Step 2: Configuring Excel for Analysis

Since Excel lacks a dedicated Nested ANOVA function, we rely on the versatile **Data Analysis ToolPak**. This suite of analysis tools is essential for performing advanced statistical procedures in Excel. If you have not previously enabled it, you must first load the Data Analysis ToolPak through the Excel Add-ins manager (File > Options > Add-ins > Excel Add-ins > Go). Once activated, the **Data Analysis** button will appear in the **Analyze** group under the **Data** tab on the main ribbon.

To initiate the analysis, click the **Data** tab, then select the **Data Analysis** button. A list of available analytical tools will appear. We must select **Anova: Two-Factor With Replication**. This is the only suitable tool for simulating a nested design in Excel because it generates the necessary Sums of Squares and Mean Squares components for the primary factor, the combined nested factor, and the error term, which are required for subsequent manual adjustment.



Upon selecting this option, a configuration dialog box will appear, prompting for specific input parameters. These parameters must accurately reflect the structure of the prepared data. The **Input Range** should encompass all the raw data and their associated labels (Fertilizer columns and Technician rows). Critically, the **Rows per sample** parameter must be set precisely to the number of replicates collected for each nested factor level. In our case study, since each technician sprinkled fertilizer on four plants, the rows per sample value is **4**. Defining the output range ensures the results are placed in an easily accessible location on the worksheet.

Step 3: Running the Two-Factor ANOVA with Replication (The Proxy Method)

Once the configuration dialog box is correctly populated, clicking **OK** will execute the analysis and generate the standard two-factor ANOVA output table. It is essential to understand how Excel is treating the input data at this stage. By using the Two-Factor With Replication tool, Excel treats the rows (which represent the nested technician groups) and the columns (which represent the fertilizer groups) as if they were crossed factors, producing interaction terms. We must mentally relabel and re-interpret these Excel outputs to fit our nested model.

In the resulting output table, Excel calculates and reports the Sums of Squares (SS), Degrees of Freedom (DF), Mean Squares (MS), F-statistic, and p-value for three main sources of variation: **Sample** (representing the rows, or Technician), **Columns** (representing the primary factor, or Fertilizer), and **Interaction** (the combined effect). For the purpose of the Nested ANOVA simulation, we interpret these categories differently.

Specifically, the Excel output provides the following components necessary for our final calculations:

The **Columns** row correctly provides the statistics for the primary factor (**Fertilizer**).

The **Sample** row provides the total variability due to all Technician groups, irrespective of which fertilizer they used.

The **Interaction** row provides the variability of the technicians *crossed* with the fertilizer, which

mathematically is used to isolate the unique variability of technicians *nested* within fertilizer.

	A	B	C	D	E	F	G	H	I	J
1		Technician 1	Technician 2	Technician 3						
2	Fertilizer A	13	15	15						
3		16	16	15						
4		16	19	12						
5		12	16	15						
6	Fertilizer B	19	23	19						
7		19	18	20						
8		20	16	21						
9		22	18	21						
10	Fertilizer C	21	25	24						
11		23	20	22						
12		24	20	25						
13		22	22	26						
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										

The output also provides the **Within** component, which represents the residual error variance--the variability among the four replicates within each unique technician/fertilizer combination. This is the true error term used for calculating the significance of the primary factor (Fertilizer).

Step 4: Initial Interpretation of Excel Output

Once the analysis is complete, the extensive ANOVA table appears. Before proceeding to the manual adjustments, we can directly interpret the results for the primary factor, **Fertilizer**. In the context of the Excel output, the row labeled **Columns** corresponds to the main effect of Fertilizer. This factor is tested against the lowest level of variation--the residual error term.

F	G	H	I	J	K	L
Anova: Two-Factor With Replication						
SUMMARY	Technician	Technician	Technician	Total		
<i>Fertilizer A</i>						
Count	4	4	4	12		
Sum	57	66	57	180		
Average	14.25	16.5	14.25	15		
Variance	4.25	3	2.25	3.818182		
<i>Fertilizer B</i>						
Count	4	4	4	12		
Sum	80	75	81	236		
Average	20	18.75	20.25	19.66667		
Variance	2	8.916667	0.916667	3.69697		
<i>Fertilizer C</i>						
Count	4	4	4	12		
Sum	90	87	97	274		
Average	22.5	21.75	24.25	22.83333		
Variance	1.666667	5.583333	2.916667	3.969697		
<i>Total</i>						
Count	12	12	12			
Sum	227	228	235			
Average	18.91667	19	19.58333			
Variance	15.17424	9.818182	20.08333			
ANOVA						
<i>Source of Variator</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Sample	372.6667	2	186.3333	53.2381	4.27031E-10	3.354130829
Columns	3.166667	2	1.583333	0.452381	0.640845706	3.354130829
Interaction	28.66667	4	7.166667	2.047619	0.115848461	2.727765306
Within	94.5	27	3.5			

We examine the **F-statistic** and the corresponding **P-value** in the **Columns** row. In this example, the p-value (4.27031E-10) is extremely small, far less than the conventional significance level of 0.05. Therefore, we confidently conclude that the type of fertilizer used (Factor A, B, or C) has a statistically significant effect on plant growth. This immediate finding addresses the primary research question regarding the overall effectiveness of the fertilizers.

However, the row labeled **Sample** and **Interaction** cannot be interpreted directly. These are based on the assumption of a crossed design. In a true Nested ANOVA, the nested factor (Technician within Fertilizer) is tested using the Interaction Mean Square as its error term, not the residual (Within) Mean Square. This requires us to manually recombine the appropriate components to calculate the correct F-statistic and its associated p-value for the nested factor.

Step 5: Calculating the Nested Factor Manually (The Crucial Adjustment)

The most crucial step in simulating a Nested ANOVA in Excel is calculating the correct F-ratio for the nested factor (Technician nested within Fertilizer). In a nested design, the appropriate statistical approach dictates that the variability of the nested factor is tested against the residual error. The challenge is extracting the appropriate Sums of Squares (SS) and Degrees of Freedom (DF) from the Excel output.

The variance component representing the technician effect *within* fertilizer is calculated by summing the SS of the **Sample** row and the SS of the **Interaction** row. This combined sum represents the total variance attributable to the technician factor (nested).

The specific manual calculations required are as follows:

Sum of Squares (Technician(Fertilizer)): $SS_{Sample} + SS_{Interaction}$

Degrees of Freedom (Technician(Fertilizer)): $DF_{Sample} + DF_{Interaction}$

Mean Square (Technician(Fertilizer)): $SS_{Technician(Fertilizer)} / DF_{Technician(Fertilizer)}$

Once the Mean Square for the nested factor is calculated, the final F-ratio for Technician(Fertilizer) is determined by dividing this new Mean Square by the Mean Square of the residual **Within** term. This F-ratio is then used to find the corresponding p-value using the appropriate F-distribution. The numerator Degrees of Freedom (DF) is the calculated DF for Technician(Fertilizer), and the denominator DF is the DF_{Within} (residual error).

ANOVA					
Source of Variation	SS	df	MS	F	P-value
Sample	372.6666667	2	186.3333333	53.23809524	4.27031E-10
Columns	3.166666667	2	1.583333333	0.452380952	0.640845706
Interaction	28.66666667	4	7.166666667	2.047619048	0.115848461
Within	94.5	27	3.5		
Total	499	35			
Source of Variation	SS	df	MS	F	P-value
Technician	31.83333333	6	5.305555556	1.515873016	0.211
<i>Formulas Used</i>	=G34+G33	=H34+H33	=G40/H40	=I40/I35	=F.DIST.RT(J40, H40, H35)

Step 6: Interpreting the Final Results and Drawing Conclusions

Following the manual calculation illustrated above, the calculated p-value for the nested factor, Technician(Fertilizer), is determined to be **0.211**. We compare this value to our established significance level ($\alpha = 0.05$). Since 0.211 is significantly greater than 0.05, we fail to reject the null hypothesis regarding the technician effect.

This statistical outcome leads to a clear and practical conclusion: the variability introduced by the individual technician applying the fertilizer is **not statistically significant** in predicting plant growth. While technicians might vary slightly in their application methods, this variability is not large enough to cause significant differences in the observed plant growth measurements.

When combined with the earlier finding that **Fertilizer Type is statistically significant** ($p < 0.05$), the results provide clear management advice: to increase plant growth, efforts and resources should be focused entirely on optimizing the quality or quantity of the fertilizer being used (the primary factor), rather than attempting to standardize or retrain the individual technicians (the nested factor). This is the power of the Nested ANOVA--it allows researchers to pinpoint the exact source of variation within a complex hierarchical system.

While conducting a nested ANOVA through proxy tools in Excel requires extra diligence and manual calculation of Degrees of Freedom and F-ratios, the method successfully delivers accurate and interpretable results consistent with specialized statistical software. Analysts must ensure they correctly structure the input data and apply the specific formulas for combining the SS and DF components to accurately evaluate the nested factor's influence.