

# Perform a Nested ANOVA in Excel (Step-by-Step)

Authored by  
**Mohammed loot**

November 5, 2025

## RECOMMENDED CITATION

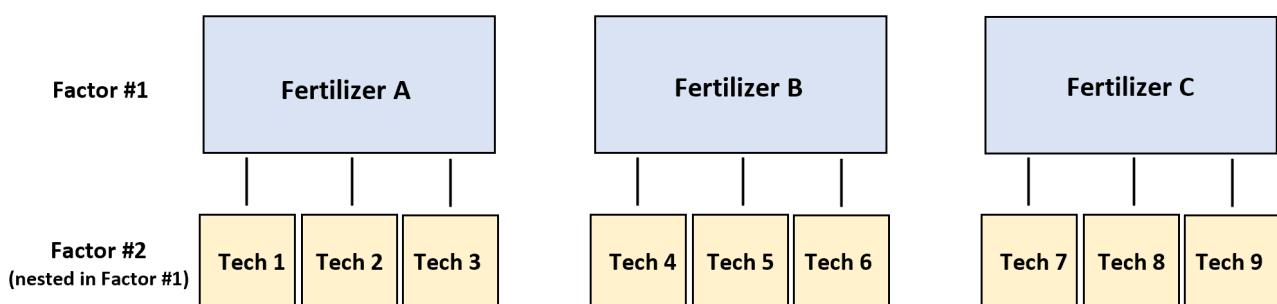
Mohammed loot (2025). *Perform a Nested ANOVA in Excel (Step-by-Step)*.  
PSYCHOLOGICAL STATISTICS. Retrieved from  
<https://statistics.arabpsychology.com/?p=10428>

A [Nested ANOVA](#), also referred to as a hierarchical design, is a specialized and powerful statistical method within the [Analysis of Variance](#) framework. This technique is indispensable when the levels of one or more experimental [factors](#) are not fully crossed but are systematically contained, or **nested**, within the levels of a higher-order factor. This structure is fundamentally different from a standard factorial ANOVA, which assumes all factor levels interact with all other factor levels.

Determining the appropriate context for applying a nested design is critical for ensuring the accuracy and validity of statistical inferences. To illustrate this necessity, let us consider a practical research scenario focused on comparing the efficacy of three distinct fertilizer types (A, B, and C) on overall plant growth.

For rigorous execution, the experiment employs nine different technicians. Crucially, three technicians are exclusively assigned to apply Fertilizer A, three others are dedicated solely to Fertilizer B, and the final three handle only Fertilizer C. Each technician treats four individual plants (replications). In this setup, the identity of the technician is entirely dependent upon--or **nested within**--the specific type of fertilizer being tested, establishing the hierarchical structure required for a [Nested ANOVA](#).

In this experimental configuration, the resulting plant growth measurement acts as the [dependent variable](#). The two independent factors are the primary factor (Fertilizer) and the nested factor (Technician). The key insight here is that Technician 1 operating under Fertilizer A is statistically and physically distinct from Technician 1 operating under Fertilizer B. Thus, the levels of the technician factor do not cross all levels of the fertilizer factor, solidifying the need for a nested approach:



While specialized statistical software packages are typically recommended for complex hierarchical designs, this comprehensive guide will demonstrate a method to perform a robust [Nested ANOVA](#) calculation using the standard tools available within **Microsoft Excel**. This method leverages Excel's built-in analysis capabilities, combined with essential manual adjustments, to yield statistically reliable results.

## Setting Up Excel: Prerequisites for Analysis

Before initiating the analysis, it is vital to acknowledge a significant limitation: Excel lacks a dedicated, one-click function for conducting a [Nested ANOVA](#). Consequently, we must utilize the **Anova: Two-Factor With Replication** option found within the [Data Analysis ToolPak](#). This tool will serve as a proxy, and its resulting output will subsequently require manual modification to correctly address the nested hypothesis structure.

To ensure access to the necessary analytical functions, you must first verify that the [Data Analysis ToolPak](#) add-in is active. This is accomplished by navigating to the **Data** tab on the main Excel ribbon. If the **Analyze** group contains the **Data Analysis** button, you are ready to proceed. If it is missing, the add-in must be loaded through the Excel Options menu.

Activating the ToolPak is a standard requirement for executing virtually any advanced statistical procedure within the Excel environment, ensuring that the essential computational algorithms are available before attempting the complex ANOVA calculations detailed in the following steps.

### Step 1: Structuring and Inputting the Data Correctly

The success of the nested analysis in Excel hinges almost entirely on the accurate structuring of the raw data to meet the specific demands of the **Two-Factor With Replication** tool. This tool is designed for fully crossed designs, so we must format the nested data to trick the tool into calculating the required intermediate sums of squares (SS) and degrees of freedom (DF).

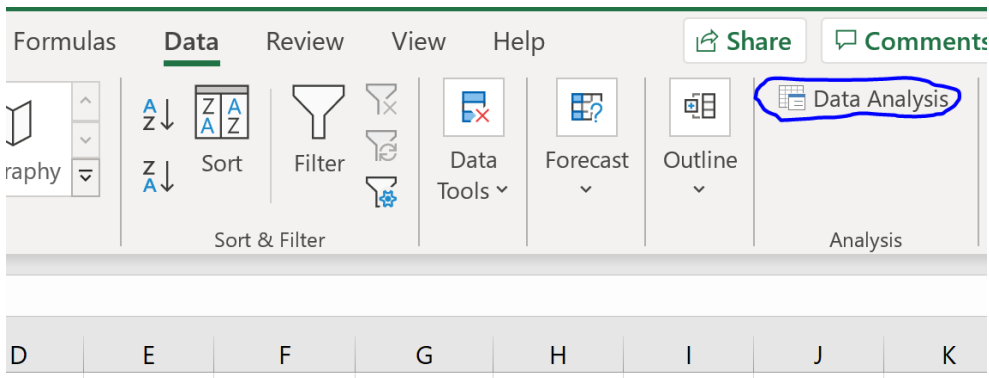
The required structure mandates that the levels of the primary factor (Fertilizer) must be arranged across the **columns**, while the levels of the nested factor (Technician) must be arranged across the **rows**. Furthermore, all replicate measurements associated with each unique nested level must be listed consecutively beneath that level. For our plant growth example, the data must adopt the following tabular arrangement:

	A	B	C	D	E	F	G
1		<b>Technician 1</b>	<b>Technician 2</b>	<b>Technician 3</b>			
2	<b>Fertilizer A</b>	13	15	15			
3		16	16	15			
4		16	19	12			
5		12	16	15			
6	<b>Fertilizer B</b>	19	23	19			
7		19	18	20			
8		20	16	21			
9		22	18	21			
10	<b>Fertilizer C</b>	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							

Observe how the nested factor (Technician 1, 2, 3) is repeated for each main factor (Fertilizer A, B, C). This repetition, combined with the four plant growth measurements (replications) taken for each technician, forms the precise input block required by Excel. This layout ensures that the Row Factor will correspond to the Technician (nested factor) and the Column Factor will correspond to the Fertilizer (main factor) in the ANOVA output. It is crucial to accurately define the **Replication** count--which is four in our case--as this is used by the tool to correctly estimate the within-cell variance, a value absolutely vital for our subsequent manual calculations.

## Step 2: Executing the Two-Factor ANOVA Proxy

Since we are using the **Anova: Two-Factor With Replication** as a computational proxy for the true [Nested ANOVA](#), precise configuration of the input parameters is mandatory. Begin by navigating back to the **Data** tab, clicking the **Data Analysis** button, selecting the **Anova: Two-Factor With Replication** option, and confirming your selection with **OK**.



Within the dialog box, meticulously define the input parameters: The **Input Range** must span all data points, including the row and column labels for both the nested factor and the main factor. The **Rows per Sample** setting must be set to 4, accurately reflecting the number of plants (replications) measured under each unique technician/fertilizer combination. You should also set the desired [Alpha](#) level (usually 0.05) and specify an **Output Range** where the results table will be generated.

Ensure your inputs mirror the following configuration, adjusting the input range to match your specific spreadsheet cell references:

	A	B	C	D	E	F	G	H	I	J
1		<b>Technician 1</b>	<b>Technician 2</b>	<b>Technician 3</b>						
2	<b>Fertilizer A</b>	13	15	15						
3		16	16	15						
4		16	19	12						
5		12	16	15						
6	<b>Fertilizer B</b>	19	23	19						
7		19	18	20						
8		20	16	21						
9		22	18	21						
10	<b>Fertilizer C</b>	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										

Anova: Two-Factor With Replication

Input

Input Range:  ↑

Rows per sample:

Alpha:

Output options

Output Range:  ↑

New Worksheet Ply:

New Workbook

OK Cancel Help

Upon clicking **OK**, Excel will produce the standard two-factor output table. It is paramount to

recognize that this initial table is based on the false assumption that Technician and Fertilizer are fully crossed. Consequently, it incorrectly calculates an interaction term and uses the within-cell variability as the sole error term for all tests. The next step is dedicated to correcting these assumptions.

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			

### Step 3: Manually Adjusting the Nested Factor F-Ratio

The central statistical distinction in a [Nested ANOVA](#) is the determination of the correct error term for the main factor. The variability associated with the nested term (Technician within Fertilizer) must serve as the denominator (error term) when testing the main factor (Fertilizer). However, the Excel output incorrectly tests the main factor against the interaction term.

In the Excel output, the row labeled **Sample** corresponds to the Fertilizer factor, and the row

labeled **Columns** corresponds to the Technician factor. To correctly test the significance of the nested [factor](#) (Technician), we must calculate a new F-ratio by utilizing the mean square (MS) of the within-cell variation (labeled **Within**) as the denominator. This MS<sub>Within</sub> is the true error term for the nested factor.

To determine if the nested factor "technician" is statistically significant, perform the following manual calculations directly on the output table, focusing on the Sum of Squares (SS) and Degrees of Freedom (DF) columns:

**Calculate the Mean Square for Technician (MSTechnician):** This value, found in the MS column, corresponds to the row labeled **Columns** in the Excel output.

**Calculate the Correct F-statistic for Technician (F<sub>Technician</sub>):** The correct F-statistic is obtained by dividing the MSTechnician by the Mean Square Error (MS<sub>Within</sub>). This MS<sub>Within</sub> is the true error term for the nested factor.

**Determine the P-value:** Use the built-in Excel function `F.DIST.RT`, referencing the newly calculated F-statistic, the degrees of freedom for the numerator (DF<sub>Technician</sub>), and the degrees of freedom for the denominator (DF<sub>Within</sub>).

The subsequent illustration demonstrates these critical adjustment formulas applied directly to the standard Excel output, yielding the corrected F and [P-value](#):

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
<b>Technician</b>	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)

After applying the correct F-ratio calculation (MSC<sub>Columns</sub> / MS<sub>Within</sub>) and the corresponding Excel function to determine the probability, the final corrected [P-value](#) for the nested factor (Technician) in this example is calculated to be **0.211**.

## Step 4: Interpreting the Final Statistical Results

The final stage involves interpreting the results for both the main factor (Fertilizer), which is calculated correctly by the tool, and the nested factor (Technician), which required the manual F-ratio adjustment. Interpretation must always be conducted relative to the predefined [Alpha](#) level (typically 0.05).

We first focus on the primary factor, Fertilizer (labeled **Sample** or **Rows** in the output). The corresponding [P-value](#) is 4.27031E-10. Since this value is exponentially small and substantially less than our threshold of 0.05, we decisively conclude that the main factor is **statistically significant**. This finding provides robust evidence to reject the null hypothesis, asserting that the three distinct fertilizer types yield significantly different average levels of plant growth.

Next, we evaluate the nested factor, Technician, utilizing the manually calculated [P-value](#) of **0.211**. Because this calculated value is significantly greater than the standard 0.05 threshold, we conclude that the Technician factor is **not statistically significant**. This outcome implies that the minor differences observed in plant growth among the individual technicians are likely attributable to random experimental error, rather than systematic variations in their technique, skill, or bias.

These combined results offer clear, actionable insights for the researcher: the variation in plant growth is overwhelmingly driven by the type of fertilizer used, not by the specific individual applying it. Consequently, resources aimed at optimizing the experiment should be directed toward selecting the superior fertilizer, rather than focusing efforts on standardizing or retraining personnel regarding application technique.

## Summary: The Role of Excel in Nested Designs

Successfully executing a [Nested ANOVA](#) in Excel is certainly possible, but it demands a thorough understanding of the underlying statistical methodology and the limitations of the software. The essential process requires meticulously restructuring the data to satisfy the requirements of a crossed-design tool and then applying precise manual adjustments to the F-ratio and [P-value](#) calculation for the nested [factor](#).

While this Excel method offers a convenient alternative, researchers engaged in complex nested or hierarchical designs involving numerous variables or more than two [factors](#) are strongly advised to utilize professional statistical software (such as R, SPSS, or SAS). Specialized software automatically ensures that the error terms for all factors are correctly calculated and tested against the appropriate sources of variance, thereby eliminating the risk inherent in manual calculation errors.

Nevertheless, for straightforward two-factor nested experiments, this detailed approach provides a

reliable and highly accessible method for analysis when dedicated software access is limited, successfully confirming the relative importance of both the main factor and its nested components.

## **Additional Resources for Statistical Analysis**