

How to Create a Contingency Table in Excel

Authored by
stats writer

December 23, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Create a Contingency Table in Excel*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=108560>

The ability to effectively summarize and visualize relationships within datasets is fundamental to sound data analysis. Among the most powerful tools for this purpose is the contingency table, which provides a straightforward way to examine the interplay between two or more dimensions of classification.

While statistical software packages are often used for complex analysis, Microsoft Excel offers robust functionality for generating these tables quickly and accurately, leveraging its powerful PivotTable feature. This guide will walk you through the process of creating, configuring, and interpreting a two-way contingency table in Excel, ensuring your data is organized for immediate insight and conclusion drawing.

Understanding how to structure your data and utilize the PivotTable tool is essential for anyone dealing with categorical data. This method simplifies complex counting and cross-tabulation tasks, allowing even novice users to perform advanced statistical summaries without writing formulas or complex code.

Introduction to Contingency Tables

A **contingency table**--often referred to as a cross-tabulation or "crosstab"--is a matrix format used in statistics to display the frequency distribution of the variables. Specifically, it summarizes the relationship between two or more categorical variables. Each cell in the table represents the count (or frequency) of observations that satisfy the categories defined by that particular row and column intersection.

The primary purpose of generating a contingency table is to determine if there is a relationship or association between the variables being studied. For instance, if you are analyzing market research, you might create a contingency table showing the relationship between "Product Type Purchased" (Column Variable) and "Customer Region" (Row Variable). By viewing the counts in the cells, you can immediately identify if certain product types are disproportionately popular in specific regions. This visualization is crucial before proceeding to advanced statistical tests, such as the Chi-Squared test, which formally assesses independence.

In Excel, the flexibility of the PivotTable allows for rapid creation of these summaries. Instead of manually counting occurrences across various categories, the PivotTable function automatically aggregates the raw data based on the chosen row and column fields. This automation minimizes errors and significantly speeds up the preliminary stage of data analysis.

Prerequisites and Data Preparation

Before initiating the PivotTable creation process, your raw data must be structured correctly. Excel requires data to be in a **flat format**, where each row represents a single observation or record, and

each column represents a specific variable. For a two-way contingency table, you must ensure that the two variables you wish to cross-tabulate are clearly defined in separate columns.

Key requirements for source data include:

Clean Data: Ensure there are no missing values (blanks) in the columns you intend to use as row or column headers, as this can lead to incorrect aggregation. Consistency in spelling and capitalization is also critical, as Excel treats "TV" and "tV" as two different categories.

Headers Required: Every column must have a unique header name. The PivotTable tool uses these headers to label the fields you drag into the Rows, Columns, and Values areas.

Categorical Structure: The variables used for the rows and columns must be categorical variables. These variables classify data into groups or categories (e.g., Country A, Country B, TV, Radio). The values area will typically use a numeric or unique identifier field (like an Order ID) which is then counted to produce the frequencies.

Adhering to these preparation steps ensures that the resulting contingency table accurately reflects the relationships within your dataset, setting the stage for reliable interpretation.

Example Dataset Overview

To demonstrate the creation of a contingency table, we will use a hypothetical dataset representing product order information. This dataset tracks 20 unique orders and includes two primary categorical variables of interest: the type of product purchased and the country from which the order originated. The goal is to cross-tabulate these two variables to understand the distribution of sales across countries and products.

The raw data includes three columns: Order Number (a unique identifier), Product (TV, Computer, or Radio), and Country (A, B, or C). Analyzing this data manually to determine how many computers were purchased in Country A, for example, would be tedious and prone to error, especially with thousands of records. The PivotTable simplifies this instantly.

Observe the dataset structure below. Note that the data is organized neatly with distinct headers and consistent categorical entries:

	A	B	C	D	E	F
1	Order Number	Product	Country			
2	1	TV	A			
3	2	TV	A			
4	3	Computer	A			
5	4	TV	A			
6	5	TV	B			
7	6	Computer	B			
8	7	Computer	B			
9	8	Computer	B			
10	9	TV	B			
11	10	Radio	B			
12	11	TV	B			
13	12	Radio	B			
14	13	Radio	C			
15	14	Radio	C			
16	15	Computer	C			
17	16	Computer	C			
18	17	TV	C			
19	18	TV	C			
20	19	Radio	C			
21	20	TV	C			
22						
23						
24						
25						
26						
27						

Our task is to transform this list of 20 individual records into a concise 3x3 matrix that summarizes the **joint frequencies** of Product Type and Country. We will use the Order Number as the field to count, thereby determining the frequency of orders per combination.

Step 1: Initiating the PivotTable Feature

The first crucial step in generating the contingency table is invoking the **PivotTable** function within Excel. This feature is located within the main application ribbon and is designed to handle large amounts of raw data efficiently.

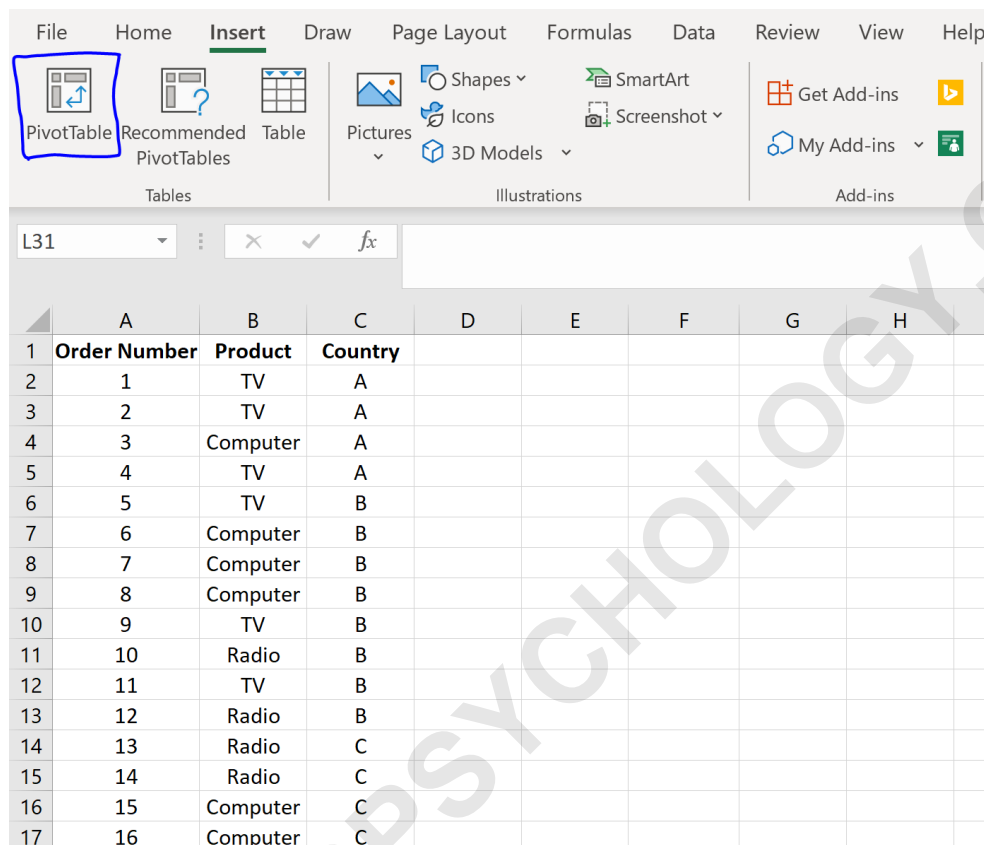
To begin, follow these instructions precisely:

Ensure any cell within your source data range is selected (e.g., cell A1 or C20). While not strictly necessary, selecting a cell within the data helps Excel automatically detect the surrounding data range.

Navigate to the **Insert** tab located on the Excel ribbon at the top of the window.

Locate the **PivotTable** option, usually found on the far left of the Insert tab, and click it. This action will launch the "Create PivotTable" dialog box, prompting you to define the source data and the location for the resulting table.

The interface is designed to guide you through the process, but paying attention to the initial selections is vital for accurate setup.



This initial step confirms your intent to summarize the data, moving beyond simple filtering or sorting into true cross-tabulation.

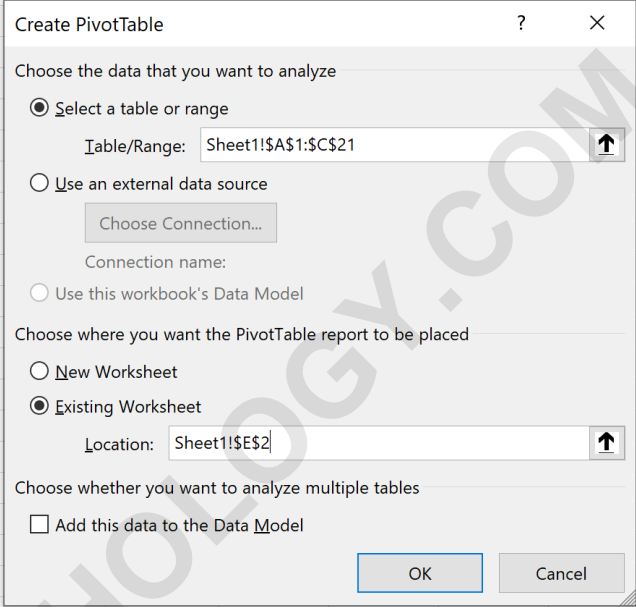
Step 2: Defining Data Range and Location

Once the "Create PivotTable" dialog box appears, you must accurately define two key parameters: the source data range and where you want the resulting contingency table to be placed. Defining the range ensures that all relevant data points are included in the calculation, while specifying the location keeps your analysis separate from the raw data.


For our example dataset, we have 21 rows of data (including the header row) spanning columns A, B, and C. Therefore, the data range to select is **A1:C21**. It is generally recommended to select the entire dataset, including the headers, so that Excel recognizes all variables.

Next, you must specify the location for the new PivotTable. You have two primary choices: a **New Worksheet** or an **Existing Worksheet**. For clarity and convenience in this tutorial, we will choose the latter. We will place the top-left corner of the PivotTable in cell **E2** within the existing worksheet.

	A	B	C	D	E	F	G	H	I	J
1	Order Number	Product	Country							
2	1	TV	A							
3	2	TV	A							
4	3	Computer	A							
5	4	TV	A							
6	5	TV	B							
7	6	Computer	B							
8	7	Computer	B							
9	8	Computer	B							
10	9	TV	B							
11	10	Radio	B							
12	11	TV	B							
13	12	Radio	B							
14	13	Radio	C							
15	14	Radio	C							
16	15	Computer	C							
17	16	Computer	C							
18	17	TV	C							
19	18	TV	C							
20	19	Radio	C							
21	20	TV	C							
22										
23										
24										
25										
26										
27										



After clicking **OK**, an empty framework for the PivotTable will appear in the designated starting cell (E2), and the PivotTable Fields pane will open on the right side of the screen. This pane contains all the headers (fields) detected from your selected range, ready to be organized.

	A	B	C	D	E	F	G	H
1	Order Number	Product	Country					
2	1	TV	A		<div style="border: 1px solid gray; padding: 5px; text-align: center;"> PivotTable3 Click in this area to work with the PivotTable report </div> 			
3	2	TV	A					
4	3	Computer	A					
5	4	TV	A					
6	5	TV	B					
7	6	Computer	B					
8	7	Computer	B					
9	8	Computer	B					
10	9	TV	B					
11	10	Radio	B					
12	11	TV	B					
13	12	Radio	B					
14	13	Radio	C					
15	14	Radio	C					
16	15	Computer	C					
17	16	Computer	C					
18	17	TV	C					
19	18	TV	C					
20	19	Radio	C					
21	20	TV	C					
22								
23								
24								
25								
26								

Step 3: Populating the Contingency Table Fields

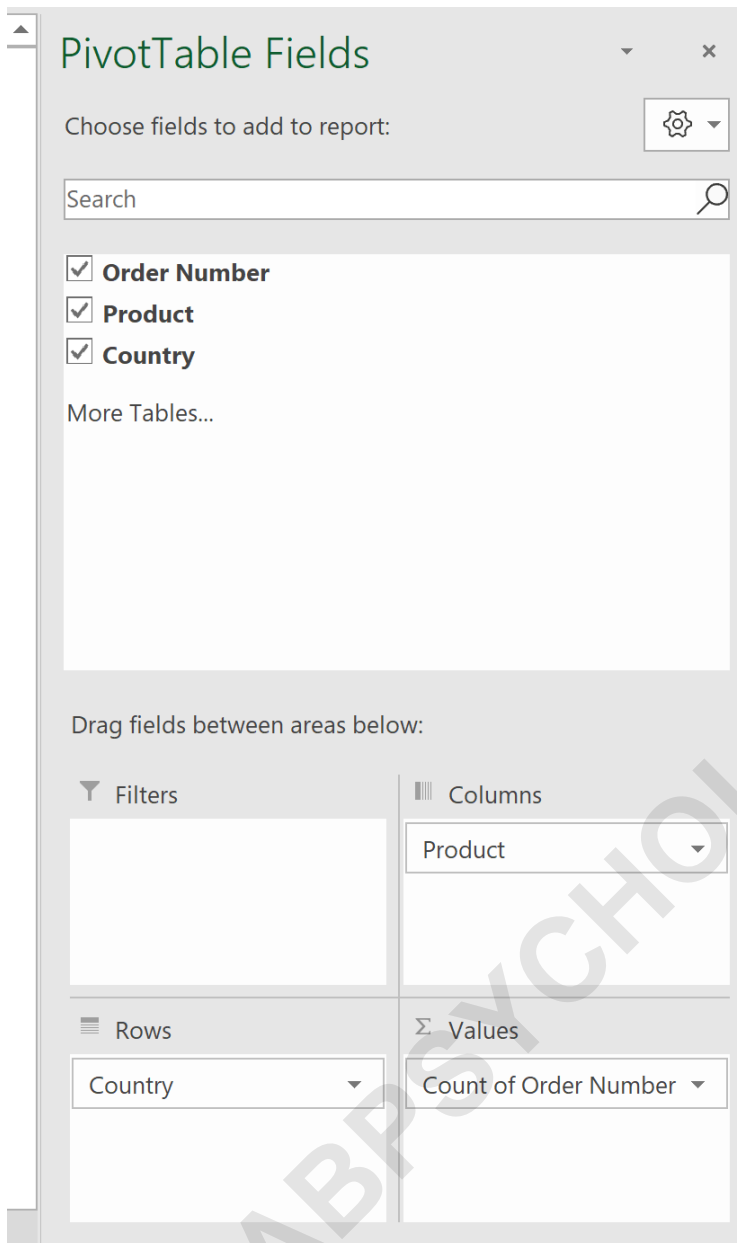
This is the most critical step, where you define the structure of the contingency table by assigning your categorical variables to the appropriate PivotTable areas: Rows, Columns, and Values. The arrangement here determines how the frequencies are calculated and displayed.

To create a standard two-way contingency table showing the frequency of orders by Country and Product, follow the dragging and dropping process in the PivotTable Fields pane:

Rows: Drag the **Country** field into the **Rows** box. This defines the categories that will appear vertically down the left side of the table (A, B, C).

Columns: Drag the **Product** field into the **Columns** box. This defines the categories that will appear horizontally across the top of the table (TV, Computer, Radio).

Values: Drag the unique identifier field, **Order Number**, into the **Values** box. Since we want to count the number of orders for each category intersection, this field will be the basis of the frequency calculation.



Important Note on Value Settings: By default, Excel sometimes attempts to summarize numeric fields using the "Sum" function. Since the "Order Number" is a unique ID and not a quantity we wish to sum, we must ensure the calculation is set to **Count**. If the Values box shows "Sum of Order Number," click the dropdown arrow next to it, select **Value Field Settings**, and then choose **Count** before clicking **OK**. This ensures the table displays frequencies (counts) rather than the sum of the order ID numbers.

Upon completing these steps, the contingency table is instantly generated, displaying the cross-tabulated frequencies.

D	E	F	G	H	I	J
	Count of Order Number	Column Labels				
	Row Labels	Computer	Radio	TV	Grand Total	
	A	1	3		4	
	B	3	2	3	8	
	C	2	3	3	8	
	Grand Total	6	5	9	20	

Step 4: Interpreting the Contingency Table Results

The resulting table is a powerful summary of the data, divided into three main components: the individual cell counts (joint frequencies), the row and column totals (marginal frequencies), and the grand total (overall sample size).

A statistical interpretation of a contingency table requires looking at these components holistically to identify patterns, associations, or lack thereof. The count within each cell represents the **joint frequency**--the number of observations that simultaneously possess the characteristics of the corresponding row and column.

We can break down the interpretation as follows:

Analyzing Marginal Frequencies (Row Totals)

The Row Totals, located in the far right column, summarize the distribution of the row variable (Country) independent of the column variable (Product). These are crucial for understanding the overall activity level of each country.

A total of **4** orders were made from country A.

A total of **8** orders were made from country B.

A total of **8** orders were made from country C.

These counts demonstrate that Country A accounted for the smallest volume of orders, while

Countries B and C had equal and higher volumes. These marginal frequencies provide the basis for calculating marginal probabilities.

Analyzing Marginal Frequencies (Column Totals)

Similarly, the Column Totals, located in the bottom row above the Grand Total, summarize the distribution of the column variable (Product) independent of the row variable (Country). They indicate the overall popularity or volume of sales for each product type across all regions combined.

A total of **6** computers were purchased.

A total of **5** radios were purchased.

A total of **9** TV's were purchased.

The marginal frequencies clearly indicate that TVs were the most popular product in this sample, followed by Computers, with Radios being the least frequently purchased item.

Analyzing Joint Frequencies (Individual Cells)

The most informative part of the contingency table lies within the individual cells, which show the joint distribution of the two variables. This allows us to observe specific relationships, such as how sales of a particular product are distributed across different countries.

For example, examining the first row (Country A):

A total of **1** computer was purchased from country A.

A total of **0** radios were purchased from country A. (This suggests a possible association or preference structure, as no radio orders originated here.)

A total of **3** TV's were purchased from country A.

Reviewing the remaining cells provides the complete picture of the joint counts:

A total of **3** computers were purchased from country B.

A total of **2** radios were purchased from country B.

A total of **3** TV's were purchased from country B.

A total of **2** computers were purchased from country C.

A total of **3** radios were purchased from country C.

A total of **3** TV's were purchased from country C.

By comparing these joint frequencies to the marginal totals, analysts can quickly identify potential dependencies. For instance, while Country A had the fewest total orders (4), 75% of those orders (3/4) were for TVs, indicating a strong preference for that specific product in Country A relative to

other products.

Advanced Contingency Table Uses in Excel

Once you have mastered the basic frequency count, the PivotTable allows you to convert the raw counts into relative frequencies, percentages of the row total, percentages of the column total, or percentages of the grand total. This transformation is essential for comparing distributions when the marginal totals are unequal.

To access these options, simply return to the **Value Field Settings** for the "Count of Order Number" field, and navigate to the **Show Values As** tab. Here, you can select options like "% of Row Total" or "% of Grand Total." Displaying the data this way provides context beyond simple counts and is often necessary for rigorous data analysis and hypothesis testing.

For instance, changing the display to "% of Row Total" would clearly show that 75% of Country A's orders were TVs, regardless of Country B's much higher volume of total orders (8). This normalization facilitates true comparative analysis between categories.

The creation of a contingency table using the PivotTable is an indispensable skill in Excel, transforming raw data into structured, interpretable statistical summaries that drive informed decision-making.