



University
of Glasgow

Excel – Array Formula

V1.1)

e-mail: training@glasgow.ac.uk

web: gla.ac.uk/services/it/training

copyright © University of Glasgow
Course content created by Blair
Last edited by Blair Thompson on 08/01/25

Introduction	iii
Objectives	iii
Excel – Array Formula.....	4
1 What is an array in Excel?	4
2 What is an array formula in Excel?	4
a. Key Features of Array Formulas:	4
b. Types of Array Formulas:	5
c. Multi-cell Array Formula.....	5
d. Referencing a Multi-cell Array	5
e. Sum of Products (Single-cell array formula).....	6
f. Why use array formulas in Excel?	7
g. Use the F9 key to evaluate portions of an array formula	7
h. Single-cell and multi-cell array functions in Excel.....	8
3 Using Excel array functions to alter the layout of datalists	9
a. SORT	9
b. SORTBY.....	9
c. UNIQUE	10
d. TRANSPOSE	10
e. TOROW.....	11
f. TOCOL.....	11
g. FILTER.....	11
h. VSTACK.....	12
i. ARRAYTOTEXT	12
4 Excel array constants	13
a. Horizontal array constant.....	13
b. Vertical array constant	13
c. Two-dimensional array constant	13
d. Working with Excel array constants	14
e. Naming array constants.....	14
5 Data Analysis Using Array Functions.....	14
a. GROUPBY	14
b. PIVOTBY.....	14
Useful Shortcut keys	16

Introduction

This course aims to offer an introduction to a set of formulas that operate upon arrays.

Objectives

On successful completion of this course participants will be able to:

- Insert bulleted list here, style H Bullets

Excel – Array Formula

Text with no indentation at all is created using the Normal Style. This is also used in Tables

1 What is an array in Excel?

Before we start on array functions and formulas, let's figure out what the term "array" means. Essentially, an *array* is a collection of items. The items can be text or numbers and they can reside in a single row or column, or in multiple rows and columns. An array can even exist within a formula result, but more of that later.

For example, if you put your weekly grocery list into an Excel array format, it would look like:

`{"Milk", "Eggs", "Butter", "Corn flakes"}`

Then, if you select cell A1 enter `={"Milk", "Eggs", "Butter", "Corn flakes"}` in the formula bar or cell and press ENTER, you will get the following result:

	A	B	C	D
1	={"milk", "eggs", "butter", "cornflakes"}			
2				

	A	B	C	D
1	milk	eggs	butter	cornflakes
2				

What you see is a one-dimensional horizontal array.

2 What is an array formula in Excel?

An **array formula** in Excel is a powerful type of formula that allows you to perform multiple calculations on one or more items in an array (a range of cells) and then return either a single result or multiple results. Unlike standard formulas, array formulas can work with a range of values rather than a single cell, which makes them useful for performing operations on large datasets or for complex tasks that would normally require multiple steps.

a. Key Features of Array Formulas:

Multiple Calculations: They allow you to perform multiple calculations on data, often across ranges of cells, and return a single result or an array of results.

Single or Multiple Results: An array formula can return a single value or multiple values depending on how it's structured. For example, you can sum products, perform conditional logic across multiple rows, or manipulate arrays.

b. Types of Array Formulas:

Single-cell array formulas: Return a single result after performing calculations on multiple values.

Multi-cell array formulas: Return multiple results in multiple cells after performing calculations on an array.

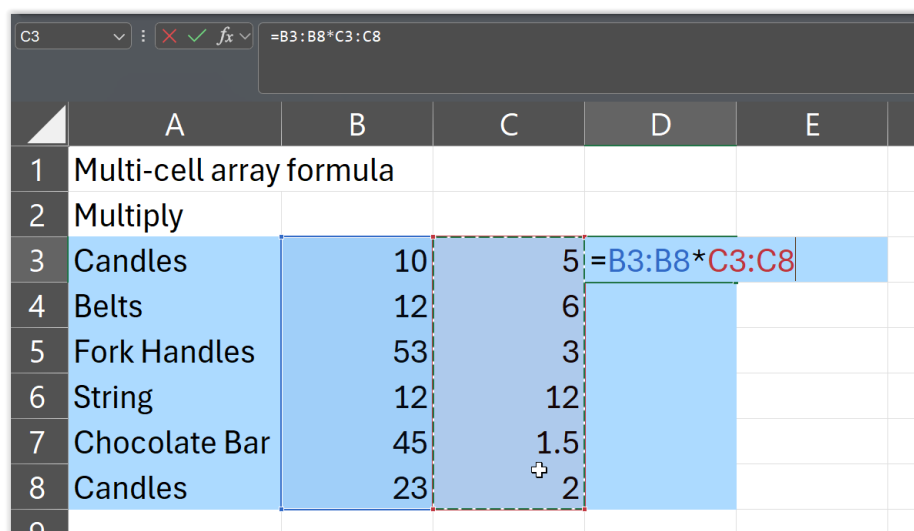
c. Multi-cell Array Formula

To return each product of two cells within a row in separate cells:

Imagine you have two columns of numbers (A1:A3 and B1:B3), and you want to find the sum of their products.

Formula:

- `=B3:B8*C3:C8`



	A	B	C	D	E
1	Multi-cell array formula				
2	Multiply				
3	Candles	10	5	=B3:B8*C3:C8	
4	Belts	12	6		
5	Fork Handles	53	3		
6	String	12	12		
7	Chocolate Bar	45	1.5		
8	Candles	23	2		
9					

- After typing this formula, press **Enter**. Excel will automatically spill the multiple answers into the cells below.
- If you see a #SPILL! error when creating an array formula, Excel is telling you that there is not enough empty cells below for all the answers to spill into.

Note: In previous versions of Excel, Array formulas were formatted differently. The formula above would display as `{=B3:B8 * C3:C8}` and required the user to use **CTRL + SHIFT + ENTER** when completing a cell to generate the `{}` brackets. This is no longer the case, but you will still see lots of articles online that instruct you to do this.

d. Referencing a Multi-cell Array

A multi-cell array spills into the empty cells around it as it has multiple values in its answer. However, if you look at one of the cells its formula is greyed out in the formula

bar. This is because the formula was not entered into that cell, the values were spilled from the one above.

	A	B	C	D	E
1	Referencing an Array (#)				
2					
3	Multiply				
4	Candles	10	5	50	
5	Belts	12	6	72	
6	Fork Handl	53	3	159	
7	String	12	12	144	
8	Chocolate	45	1.5	67.5	
9	Candles	23	2	46	
10					

But what if you want to use those answers within another (array) formula?

You can do so, with a small change to the cell reference

	A	B	C	D	E	F
1	Referencing an Array (#)					
2						
3	Multiply					Sum of Sales
4	Candles	10	5	50		=sum(D4#)
5	Belts	12	6	72		
6	Fork Handl	53	3	159		
7	String	12	12	144		
8	Chocolate	45	1.5	67.5		
9	Candles	23	2	46		
10						

If you reference the cell that you input your first array formula, in our example Cell D4, then you add a # to the end i.e.

=SUM(D4#)

This will reference the array created at cell D4 and in this example, sum the 6 values returned.

e. Sum of Products (Single-cell array formula)

This formula multiplies each corresponding value in columns A and B and then sums the results.

To perform the same operation but return each product in a single cell:

- Select a cell, for example, C10
- Enter the formula:

=SUM(B4:B9*C4:C9)

	A	B	C	D
1	Sum			
2				
3	Sum Product (The hard way)			
4	Candles	10	5	
5	Belts	12	6	
6	Fork Hand	53	3	
7	String	12	12	
8	Chocolate	45	1.5	
9	Candles	23	2	
10			=sum(B4:B9*C4:C9)	

- Press **Enter**.
- This will return the SUM of all the products of each pair of corresponding values in columns A and B, with the result appearing in a single cell.

Common Use Cases:

- Summing or averaging a filtered dataset.
- Performing complex mathematical operations like matrix multiplication.
- Extracting values based on multiple criteria (advanced conditional formulas).

Important Notes:

- Array formulas are essential for handling more complex data analysis tasks and can save time by reducing the need for multiple helper columns or manual calculations.
- Simple example of Excel array formula
- Suppose you have some items in column B, their prices in column C, and you want to calculate the grand total of all sales.

f. Why use array formulas in Excel?

Excel array formulas are the handiest tool to perform sophisticated calculations and do complex tasks. A single array formula can replace literally hundreds of usual formulas. Array formulas are very good for tasks such as:

g. Use the F9 key to evaluate portions of an array formula

When working with array formulas in Excel, you can observe how they calculate and store their items (internal arrays) to display the final result you see in a cell. To do this, select one or several arguments within a function's parentheses, and then press the F9 key. To exit the formula evaluation mode, press the Esc key.

In the previous example, to see the sub-totals of all products, you select B4:B9*C4:C9, press F9 and get the following result.

=SUM({50;72;159;144;67.5;46})						
	A	B	C	D	E	
1	Sum					
2						
3	Sum Product (The hard way)					
4	Candles	10	5			
5	Belts	12	6			
6	Fork Hand	53	3			
7	String	12	12			
8	Chocolate	45	1.5			
9	Candles	23	2			
10			=SUM({50;72;159;144;67.5;46})			

h. Single-cell and multi-cell array functions in Excel

Excel array formula can return a result in a single cell or spilled into multiple cells. An array formula with returned values that spills into a range of cells is called a **multi-cell formula**. An array formula with a single returned value in a single cell is called a **single-cell formula**.

There exist a few Excel array functions that are designed to return multi-cell arrays, for example TRANSPOSE, UNIQUE, TREND, FREQUENCY, LINEST, etc.

Other functions, such as SUM, SUMPRODUCT, AVERAGE, AGGREGATE, MAX, MIN, can calculate array expressions when entered into a single cell.

12						
13	Sum Product (The easy way?)					
14	Candles	10	5			
15	Belts	12	6			
16	Fork Hand	53	3			
17	String	12	12			
18	Chocolate	45	1.5			
19	Candles	23	2			
20		=SUMPRODUCT(B14:B19,C14:C19)				
21		SUMPRODUCT(array1, [array2], [array3], [array4], ...)				
22						

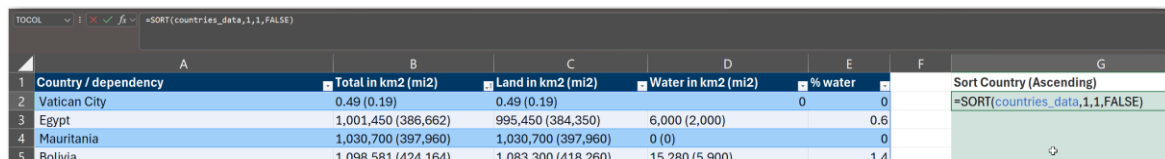
3 Using Excel array functions to alter the layout of datalists

As already mentioned, Microsoft Excel provides a few so called "array functions" that are specially designed to work with multi-cell arrays.

a. SORT

The SORT function can be used to create an array of values sorted in either ascending or descending order. Often you will see it used as part of a nested formula to "process" the results of another array calculation.

The formula is =SORT(countries_data,1,1,FALSE)



The screenshot shows an Excel spreadsheet with a table of country data. The table has columns: Country / dependency, Total in km2 (mi2), Land in km2 (mi2), Water in km2 (mi2), and % water. The data is sorted by the 'Country / dependency' column in ascending order. The formula bar shows the formula =SORT(countries_data,1,1,FALSE).

	A	B	C	D	E	F	G
1	Country / dependency	Total in km2 (mi2)	Land in km2 (mi2)	Water in km2 (mi2)	% water		Sort Country (Ascending)
2	Vatican City	0.49 (0.19)	0.49 (0.19)	0	0		=SORT(countries_data,1,1,FALSE)
3	Egypt	1,001,450 (386,662)	995,450 (384,350)	6,000 (2,000)	0.6		
4	Mauritania	1,030,700 (397,960)	1,030,700 (397,960)	0 (0)	0		
5	Bolivia	1,098,581 (424,164)	1,083,300 (418,260)	15,280 (5,900)	1.4		

Parameters

Syntax =SORT(array, [sort_index], [sort_order], [by_col])

- **array**: The range or array to be sorted (required).
- **sort_index**: The column (or row) number to sort by (optional, default is 1).
- **sort_order**: The order of sorting (optional):
 - 1 for ascending (default).
 - 1 for descending.
- **by_col**: A logical value (optional):
 - FALSE or omitted sorts by rows.
 - TRUE sorts by columns.

In this case in our first argument, **array**, countries_data is the name of a Table within our practice file, in Excel, formulas that reference **tables** are called **structured references**. **Structured references** use special syntax to refer to specific table elements by name rather than by cell address. This makes formulas easier to read and more dynamic when the table changes.

The next argument [sort_index] contains the index 1. This tells the SORT function we want to sort on the 1st column in the array

The [sort_order] argument contains 1, this sorts the table in Ascending order

Finally the [by_col] is set to FALSE so that the rows of our array are sorted

b. SORTBY

The **SORTBY** function in Excel allows you to sort a range or array **based on the values in a different range or array**. It provides more flexibility than the **SORT** function by using multiple sort ranges and criteria.

The Formula is: =SORTBY(countries_data[Country / dependency],countries_data[% water],1)

Parameters

Syntax=SORTBY(array, by_array1, [sort_order1], [by_array2], [sort_order2], ...)

- **array**: The range or array to sort (required).
- **by_array1**: The first range or array to sort by (required).
- **sort_order1**: The order to use for the first array (optional):
1 for ascending (default).
-1 for descending.
- **by_array2, sort_order2**: Additional ranges and sort orders for secondary sorting (optional).

c. UNIQUE

The **UNIQUE** function in Excel returns a list of unique values from a range or array. It eliminates duplicates and can optionally return unique values by columns or by rows.

The formula is: =UNIQUE(stock_data5[Category])

	I	J	K	
lue		UNIQUE		UNIQUE
02.00		=UNIQUE(stock_data5[Category])		
23.00				
30.00				

Parameters

Syntax =UNIQUE(array, [by_col], [exactly_once])

- **array**: The range or array from which to extract unique values (required).
- **by_col**: A logical value (optional):
FALSE (default) to compare by rows.
TRUE to compare by columns.
- **exactly_once**: A logical value (optional):
FALSE (default) returns all unique values.
TRUE returns values that appear only once (distinct values).

In this case we only supply the array, the default options (compare by rows and return all unique values are fine)

d. TRANSPOSE

The **TRANSPOSE** function in Excel converts a vertical range of cells to a horizontal range, or vice versa. It rearranges the layout of the data without changing the content.

Parameters

Syntax =TRANSPOSE(array)

- **array**: The range or array of cells to transpose (required).

e. TOROW

The **TOROW** function in Excel transforms an array or range into a **single row**, combining all values into a horizontal sequence. It is a dynamic array function available in newer versions of Excel (Excel 365 and Excel 2021).

Parameters

Syntax=TOROW(array, [ignore], [scan_by_column])

- **array**: The range or array to transform (required).
- **ignore**: Specifies what to ignore (optional):
0 to include all values (default).
1 to ignore blanks.
2 to ignore errors.
- **scan_by_column**: A logical value (optional):
FALSE or omitted scans by row (default).
TRUE scans by column.

f. TOCOL

The **TOCOL** function in Excel converts an array or range into a **single column**, stacking all values vertically. It is part of the dynamic array functions available in **Excel 365** and **Excel 2021**.

Parameters

Syntax =TOCOL(array, [ignore], [scan_by_column])

- **array**: The range or array to transform (required).
- **ignore**: Specifies what to ignore (optional):
0 to include all values (default).
1 to ignore blanks.
2 to ignore errors.
- **scan_by_column**: A logical value (optional):
FALSE or omitted to scan by row (default).
TRUE to scan by column.

g. FILTER

The **FILTER** function in Excel extracts data from a range or array that meets specified conditions. It is a **dynamic array** function available in **Excel 365**, **Excel 2021**, and **Excel 2019** (with updates).

Parameters

Syntax =FILTER(array, include, [if_empty])

- **array**: The range or array to filter (required).
- **include**: A condition or logical expression to filter by (required).
- **if_empty**: The value to return if no results meet the criteria (optional).

h. VSTACK

The VSTACK function in Excel vertically stacks arrays or ranges into a single array. It is part of the dynamic array functions available in Excel 365 and Excel 2021.

Parameters

- Syntax =VSTACK(array1, [array2], ...)
- array1: The first range or array to stack (required).
- array2, ...: Additional arrays to stack (optional). You can specify multiple arrays.

i. ARRAYTOTEXT

The ARRAYTOTEXT function in Excel converts an array or range of data into text. It is particularly useful for viewing data as a single text string representation of values, preserving structure for display or debugging purposes. This function is available in Excel 365 and Excel 2021.

Parameters

Syntax =ARRAYTOTEXT(array, [format])

- array: The array or range to convert to text (required).
- format: The text output format (optional):

0 for compact (default) — outputs values as text without extra formatting.

1 for strict — includes quotes and detailed formatting for each value.

Use Cases

- Debugging dynamic array results by converting them into readable text.
- Storing array outputs as static text representations for documentation or analysis.

Notes:

- ARRAYTOTEXT can handle both 2D and dynamic arrays.
- Compact formatting is simpler and more suitable for basic displays, while strict formatting provides more detailed outputs.

SEQUENCE

The **SEQUENCE** function in Excel generates a sequence of numbers in an array, either in rows, columns, or both. It is a dynamic array function available in Excel 365, Excel 2021, and Excel 2019 (with updates).

Parameters

Syntax =SEQUENCE(rows, [columns], [start], [step])

- rows: The number of rows in the sequence (required).
- columns: The number of columns (optional, default is 1).
- start: The starting number of the sequence (optional, default is 1).
- step: The increment between each number (optional, default is 1).

Use Cases

- Quickly creating number sequences for indexing or calculations.
- Generating date sequences by using custom start and step values.
- Automating grid-based numbering for forms or tables.

Notes

- Automatically spills into adjacent empty cells.
- Can be combined with other functions for advanced calculations or dynamic ranges.

4 Excel array constants

In Microsoft Excel, an array constant is simply a set of static values. These values never change when you copy a formula to other cells or values.

We already looked at Array constants in our very first example of an array

There exist 3 types of array constants:

a. Horizontal array constant

A horizontal array constant resides in a row. To create a row array constant, type the values separated by commas and enclose them in braces, for example
`={"Jan","Feb","Mar"}.`

Note. When creating an array constant, you should type the opening and closing {} braces.

To enter a horizontal array in a spreadsheet, select a cell, type the formula `={1,2,3,4}` in the formula bar, and press Enter.

b. Vertical array constant

A vertical array constant resides in a column. You create it in the same way as a horizontal array with the only difference that you delimit the items with semicolons, for example:

`={11; 22; 33; 44}`

c. Two-dimensional array constant

To create a two-dimensional array, you separate each row by a semicolon and each column of data by a comma.

`={"a", "b", "c"; 1, 2, 3}`

d. Working with Excel array constants

Array constants are one of the cornerstones of an Excel array formula. The following information and tips might help you use them in the most efficient way.

e. Naming array constants

To make an array constant easier to use, give it a name:

- Switch to the *Formulas tab* > *Defined Names* group and click **Define Name**. Alternatively, press Ctrl + F3 and click *New*.
- Type the name in the *Name*
- In the *Refers to* box, enter the items of your array constant surrounded in braces with the preceding equality sign (=). For example:
- Click OK to save your named array and close the window.

To enter the named array constant in a sheet, type the array's name in the formula bar preceded with the = sign and press Enter.

5 Data Analysis Using Array Functions

a. GROUPBY

The GROUPBY function in Excel enables you to group data by one or more columns and perform aggregations, similar to the functionality of PivotTables, but directly within a formula. This function is available in Excel for Microsoft 365.

Parameters

Syntax =GROUPBY(row_fields, values, function, [field_headers], [total_depth], [sort_order], [filter_array])

- row_fields: The range or array containing the values to group by (required).
- values: The range or array of data to aggregate (required).
- function: The aggregation function to apply, such as SUM, AVERAGE, COUNT, etc. (required).
- field_headers: Specifies whether your dataset has headers and whether to include them in the results (optional).
- total_depth: Determines whether to show totals and subtotals (optional).
- sort_order: Specifies the sort order of the grouped data (optional).
- filter_array: Allows filtering of data before grouping (optional).

b. PIVOTBY

The PIVOTBY function in Excel enables you to create dynamic summary tables directly within a formula, functioning similarly to PivotTables but with enhanced flexibility and integration into your worksheets. This function is available in Excel for Microsoft 365.

Parameters

Syntax = PIVOTBY(row_fields, col_fields, values, function, [field_headers], [row_total_depth], [row_sort_order], [col_total_depth], [col_sort_order], [filter_array], [relative_to])

- **row_fields:** The range or array containing the values to group by in rows (required).
- **col_fields:** The range or array containing the values to group by in columns (required).
- **values:** The range or array of data to aggregate (required).
- **function:** The aggregation function to apply, such as SUM, AVERAGE, COUNT, etc. (required).
- **field_headers:** Specifies whether your dataset has headers and whether to include them in the results (optional).
- **row_total_depth:** Determines the level of row totals to display (optional).
- **row_sort_order:** Specifies the sort order of the row fields (optional).
- **col_total_depth:** Determines the level of column totals to display (optional).
- **col_sort_order:** Specifies the sort order of the column fields (optional).
- **filter_array:** Allows filtering of data before grouping (optional).
- **relative_to:** Defines the reference point for relative calculations (optional).

Useful Shortcut keys

Using keyboard shortcuts can help you become more efficient when creating documents in Microsoft applications. Most keyboard shortcuts require you to use two or more keys at the same time. To use a keyboard shortcut first press and hold down the modifier key or keys (i.e. SHIFT, CTRL, ALT) and then press the corresponding standard key on your keyboard.

Function	Shortcut
Go to "Tell me what you want to do"	ALT+Q
Open	CTRL+O
Save	CTRL+S
Close	CTRL+W
Cut	CTRL+X
Copy	CTRL+C
Paste	CTRL+V
Select all	CTRL+A
Bold	CTRL+B
Italic	CTRL+I
Underline	CTRL+U
Cancel	Esc
Undo	CTRL+Z
Re-do	CTRL+Y
Evaluate selected portion of formula	F9
Newline (in formula bar)	ALT + =