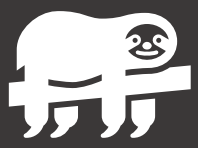
myOnlineTraininghub

How to Improve Excel Performance

Microsoft® Most Valuable Professional

By Mynda Treacy

Are your Excel files slow?



16 common causes of Excel performance issues and how to solve them

myOnlineTraininghub

Large Amounts of Data

Too much data is one of the most obvious causes of slow Excel files, although this is rarely the sole contributor.

Solution: Be diligent and only store data in the file that is required for the task.

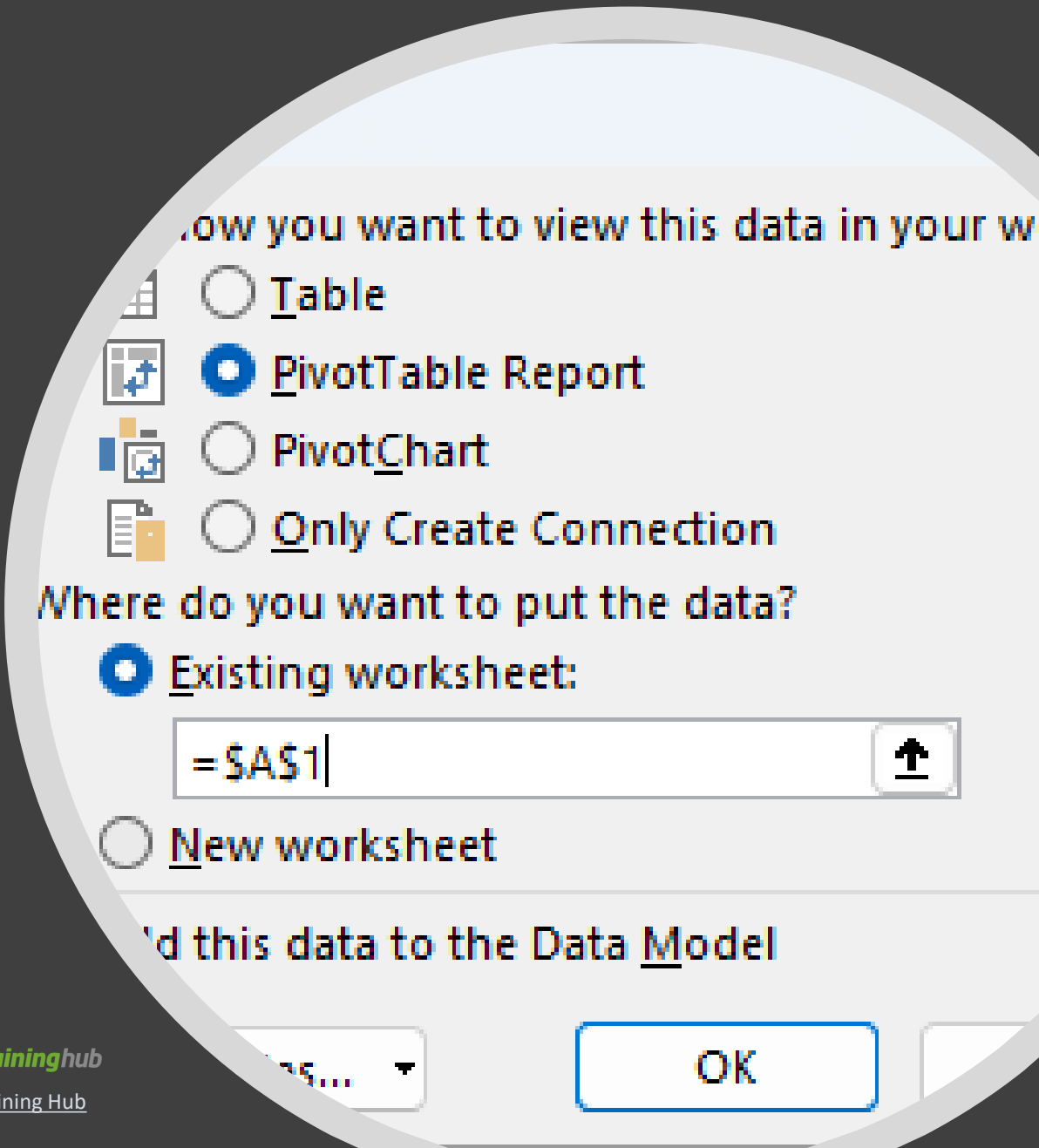
Use Power Query to filter out the unnecessary data before loading it to the Excel file.

myOnlineTraininghub

Duplicated Data

When **data is used in a PivotTable it is typically stored in the file twice**, once in the worksheet and again in the Pivot Cache.

Solution: Use Power Query to get the data and load it direct to the Pivot Cache by selecting 'PivotTable Report' in the Import Data dialog box.



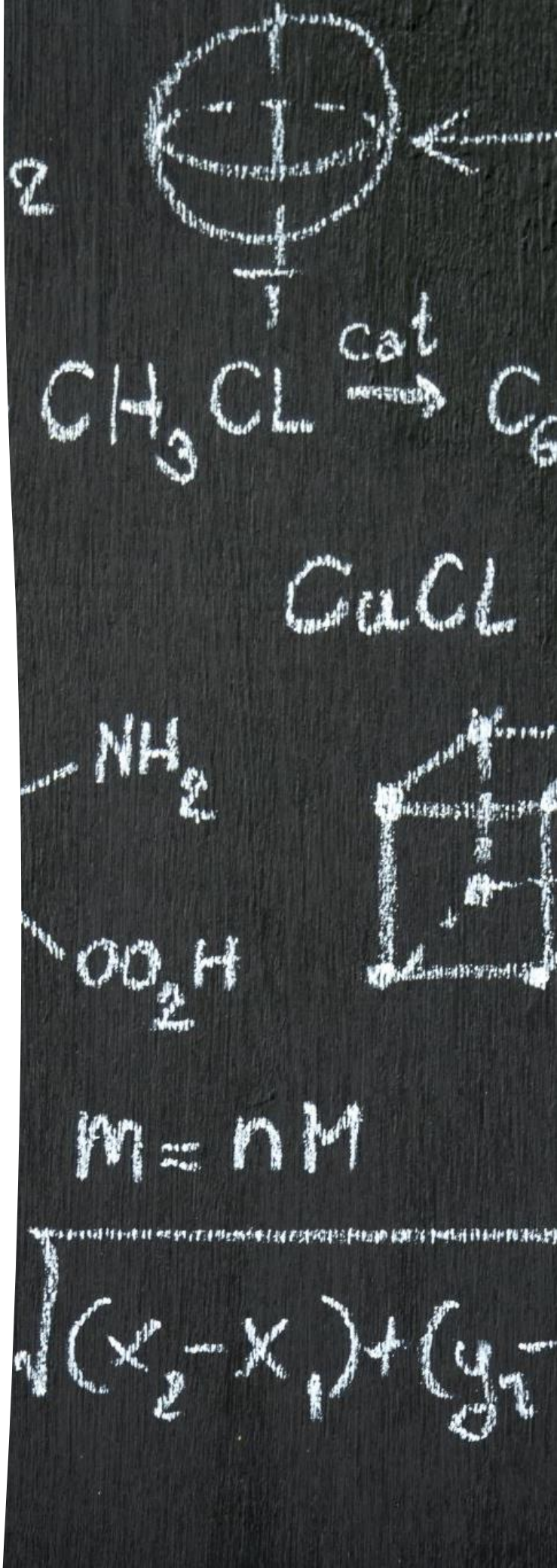
Array Formulas

Single cell array formulas can be processed several times, depending on the number of cells referenced in the formula.

Solutions:

- Avoid mixing row and column references or overlapping array references.
- Separate the calculation into multiple cells rather than an array formula.
- Move the calculation to Power Query. Power Query only calculates when you load the data or refresh the data, whereas formulas calculate every time something they reference changes, and in the case of volatile functions, every time ANYTHING changes.

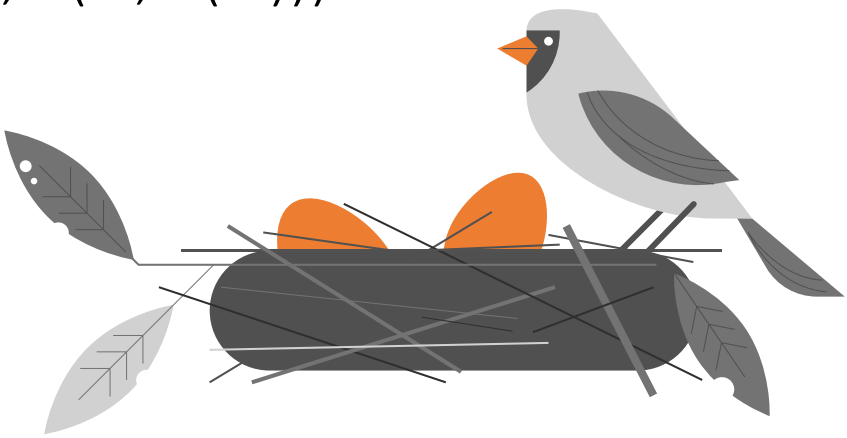
Note: Modern dynamic array functions are more efficient than their older CTRL+SHIFT+ENTER counterparts.



myOnlineTraininghub

myOnlineTraininghub

=IF(A1>B1,C1,IF(…,IF(…)))



Nested Formulas

Nested formulas in Excel, like nested IFs can cause performance issues because they involve multiple levels of calculations that require more processing power and memory.

Nested formulas can be especially problematic when they involve large data sets, such as when performing lookups or other operations across multiple worksheets or workbooks.

Solutions:

- Breaking down complex formulas into smaller, more manageable parts can help improve performance and make the worksheet easier to read and debug.
- Move the calculations to Power Query where they are calculated once on loading the data, rather than multiple times during the use of the file.
- Copy and paste the formulas as values if they are not expected to change.

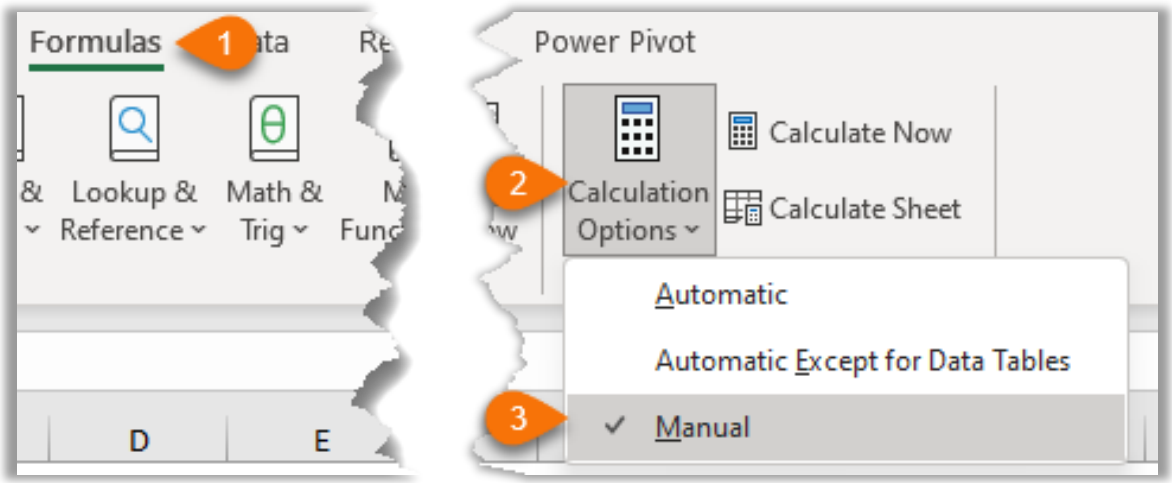
Volatile Functions

Volatile functions can cause performance issues because they recalculate every time anything changes in the workbook, even if the change has no direct impact on the function's output. Examples of volatile functions include:

- INDIRECT
- OFFSET
- NOW
- RAND
- RANDARRAY
- RANDBETWEEN
- TODAY
- CELL (depending on arguments)
- INFO (depending on arguments)
- SUMIF (when the size of the first range is not the same as the second 'sum_range')

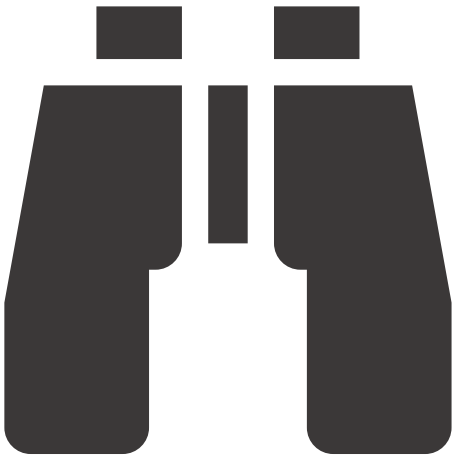
Solutions:

- Use volatile functions sparingly and only when necessary. If possible, use non-volatile alternatives, such as INDEX instead of OFFSET for dynamic named ranges.
- Switch to manual calculation mode. This way you only recalculate the workbook when necessary.



CAUTION: don't forget to manually calculate by pressing F9

myOnlineTraininghub



Lookup Formulas

VLOOKUP | HLOOKUP | INDEX | XLOOKUP etc.

Lookup formulas can cause performance issues because they often involve searching through large data sets, which can be time consuming and resource intensive*.

Solutions:

- Use Power Query to perform lookups:
  - Power Query lookup exact match
  - Power Query lookup approximate match
  - Power Query lookups using list functions
- If you must use a lookup formula, store the lookup table and the formula on the same sheet.
- Avoid using exact match lookups where possible. Instead, use approximate match with a sorted list.
- Be sure to only reference the cells containing the data being looked up. Including empty rows in the formula will reduce efficiency.

* Office 365 users may not experience performance issues to the same extent as those using earlier versions of Excel due to the new internal cached index Excel creates for lookup functions.

Conditional Functions

Conditional functions in Excel can cause performance issues because they involve testing each cell in a given range against a set of criteria, which can be time consuming and resource intensive, especially for large data sets.

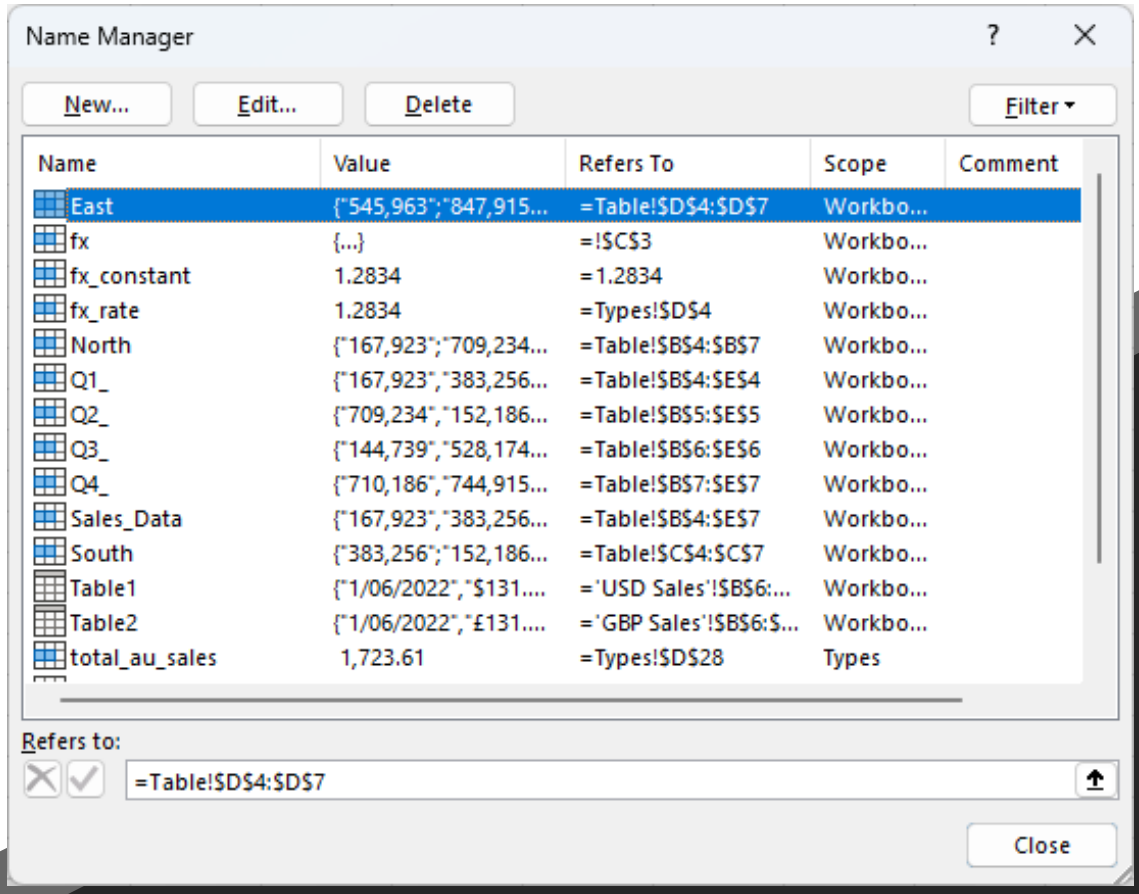
Examples of conditional functions include SUMIF/S, COUNTIF/S, AVERAGEIF/S, among others.

Note: as with the lookup functions, in Microsoft 365 these functions now create an internal cached index for the range being searched. This cached index is reused in any subsequent aggregations that are pulling from the same range.

Solutions:

- Avoid selecting more cells than necessary
- Use a PivotTable instead. PivotTables can perform these calculations and you don't need to know how to write the formula.

Order Quantity					
Row Labels	Critical	High	Medium	Low	Grand Total
⊟2010	8,516	11,870	11,348	12,270	44,004
⊞Qtr1	2,421	2,334	3,004	2,788	10,547
⊞Qtr2	1,657	3,277	2,928	3,129	10,991
⊞Qtr3	2,252	3,016	3,022	3,318	11,608
⊞Qtr4	2,186	3,243	2,394	3,035	10,858
⊟2011	10,304	11,807	9,280	10,350	41,741
⊞Qtr1	2,651	2,150	2,272	2,581	9,654
⊞Qtr2	2,559	3,236	2,811	2,574	11,180
⊞Qtr3	2,276	3,291	2,035	2,477	10,079
⊞Qtr4	2,818	3,130	2,162	2,718	10,828
⊟2012	10,466	11,910	9,798	11,321	43,495
⊞Qtr1	2,371	2,730	2,442	3,347	10,890
⊞Qtr2	2,653	3,648	2,738	2,646	11,685
⊞Qtr3	2,542	2,941	2,783	3,131	11,397
⊞Qtr4	2,900	2,591	1,835	2,197	9,523
Grand Total	29,286	35,587	30,426	33,941	129,240



Defined Names

Defined names are recalculated each time a formula that refers to the name is recalculated, even if the value of a cell does not change when calculated.

Solution: defined names are one of the most valuable Excel features, so don't avoid them because they *might* cause performance problems. Instead, check if any of the other causes could be contributing and resolve them first. If you still have performance problems, then try replacing names with direct references to cells.

Handling Errors

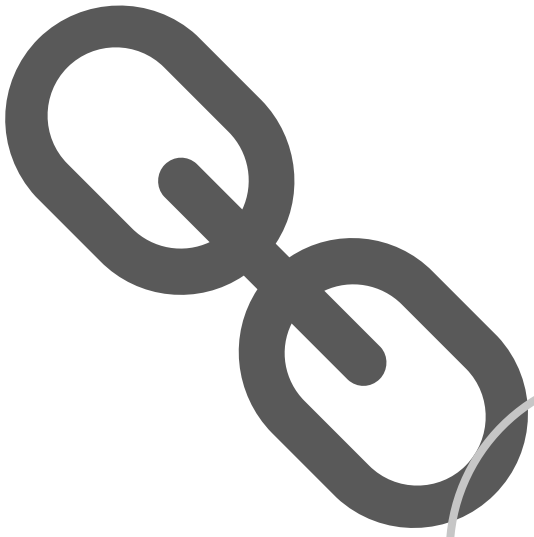
The old IF(ISNA(VLOOKUP(…),0,VLOOKUP(…)) technique requires Excel to do double the work i.e. two lookups.

Solution: Instead, use the IFERROR or IFNA function to handle errors returned by lookup functions.

=IFNA(VLOOKUP(…),0)

=IFERROR(INDEX(…),"")

myOnlineTraininghub



External Links

As a general rule, external links should be avoided. Links to external Excel files are slow to calculate and easily broken. Plus, many functions cannot evaluate on a closed workbook. Internal links can also slow down calculation.

Solution: Use Power Query to bring the data into the current file. If you must use external links, open the file being linked to before opening the file doing the linking.

myOnlineTraininghub

Excessive Formatting

Excessive cell formatting, including font styles, colors, borders, and other visual attributes that are applied to cells can consume a significant amount of processing power and memory, especially when applied to large data sets.

Excessive formatting can also make the file size larger, which can slow down the loading and saving times of the workbook.

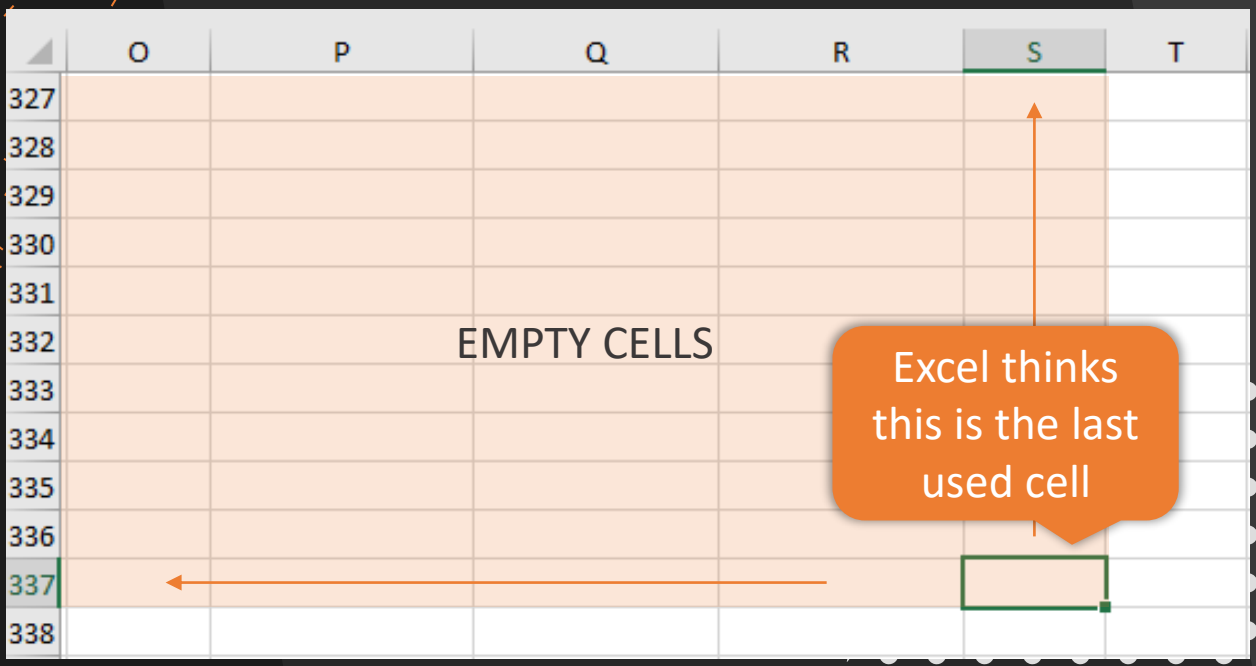
Solution: instead of applying formatting like cell fill colour to a large range of cells, apply it to the whole column/row. It's easier for Excel to know that a whole row or column is formatted in a particular way than it is to keep track of 1000 separate cells.

Redundant Formatting

Redundant formatting (used range) can sometimes linger in cells unbeknownst to you. Cells can appear empty, but Excel is still storing information about those cells in memory.

CTRL+END will take you to the last cell in the sheet that Excel is storing information for. If this isn't the end of your table, then you know you have redundant formatting.

In the image below, Excel thinks the last used cell is S337, but there is no data in the cells to the left or above. See next page for the solution.

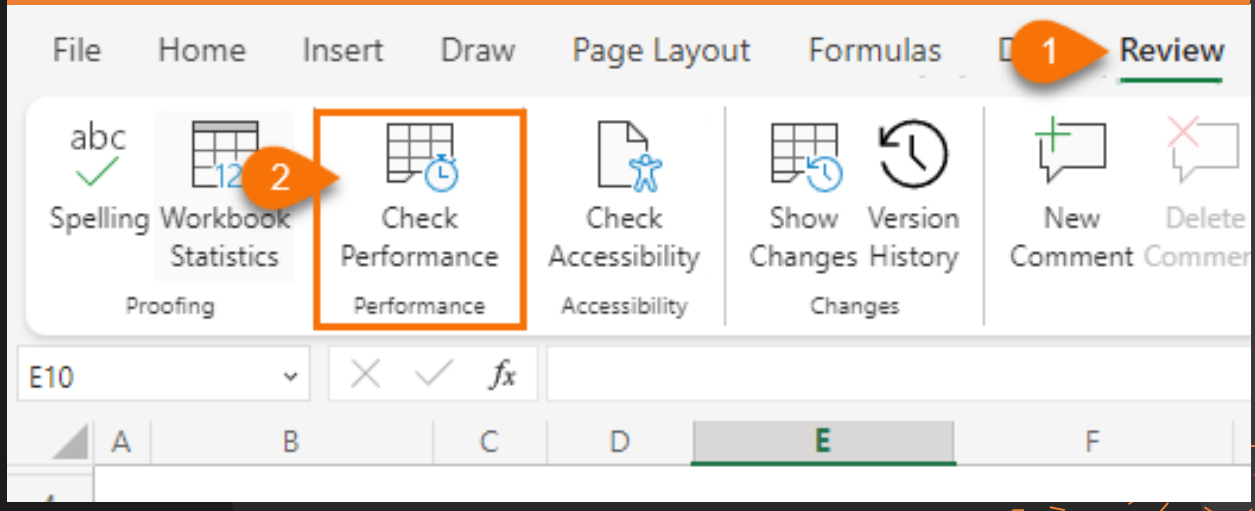


Redundant Formatting Solution

Solution: you can try deleting rows and columns that are empty, but I have found this often doesn't resolve the problem (backup the file before you do).

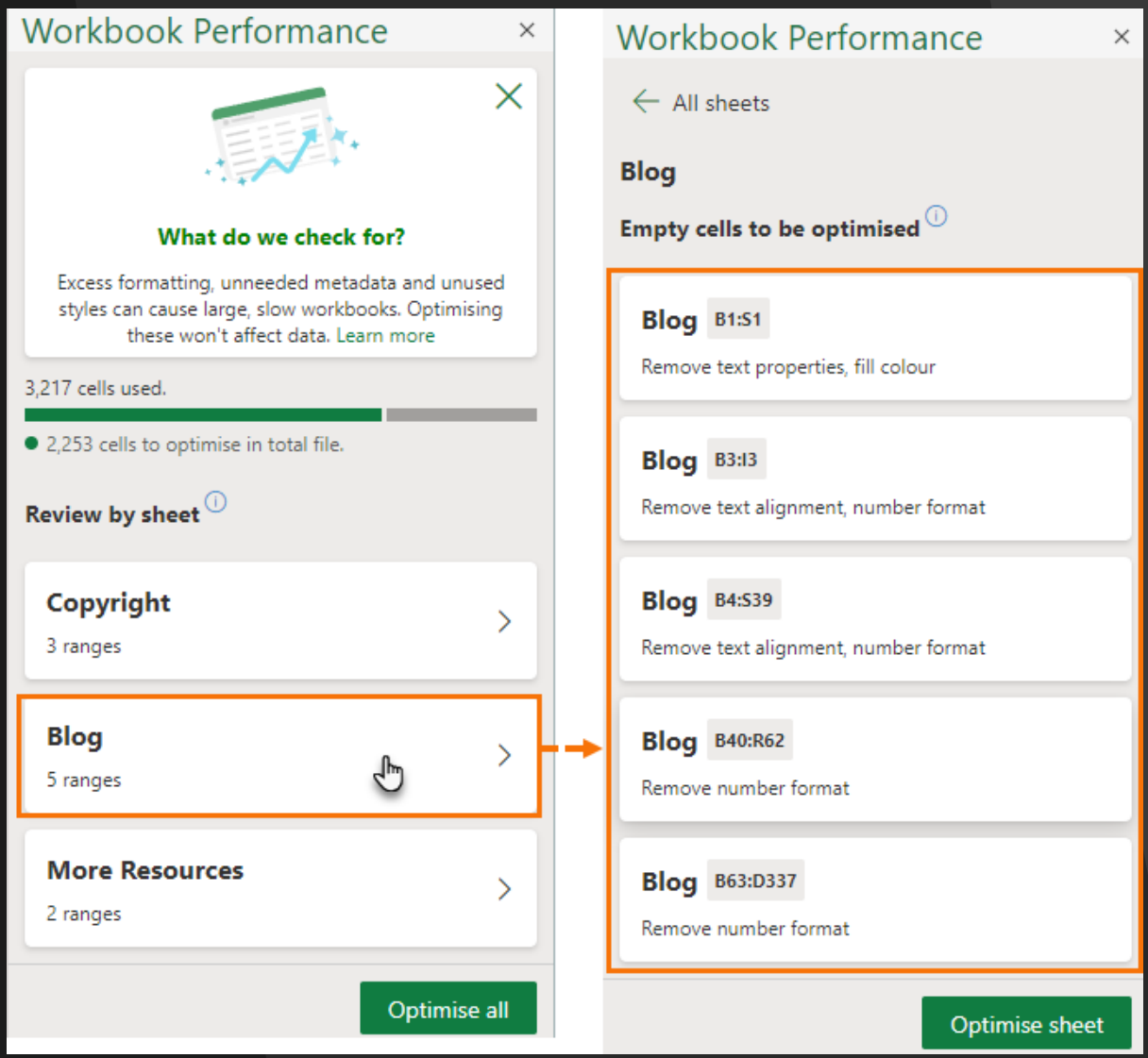
Use the new Optimize Sheet tool available in **Excel Online** for Microsoft 365 to check for performance issues. It's available on the Review tab of the ribbon:

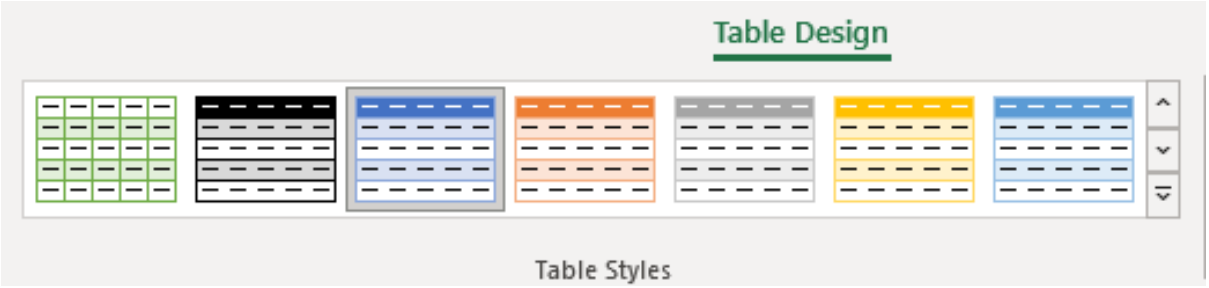
Excel Online



Redundant Formatting Solution

Clicking on ‘Check Performance’ brings up the Workbook Performance pane which summarises the sheets that contain any issues. From there you can see a list of the individual cells/ranges and the issues (see image below). Clicking Optimise all/sheet will remove them for you.





Excel Tables (2013 Only)

Storing large amounts of data in Excel Tables in Excel 2013 and earlier can sometimes yield worse performance than not formatting it in a table. An unofficial cut off is >500k rows x 10 columns. This is sometimes too much to store in Excel 2013 Tables, although it can depend on whether you have a lot of formulas or not.

Solution: If the file is slow, consider storing the data in the Power Pivot data model rather than the worksheet. Power Pivot has an advanced compression algorithm that enables it to store data more efficiently than Excel itself. Power Pivot can also store 10's of millions of rows of data and overcome the row limitations of Excel.

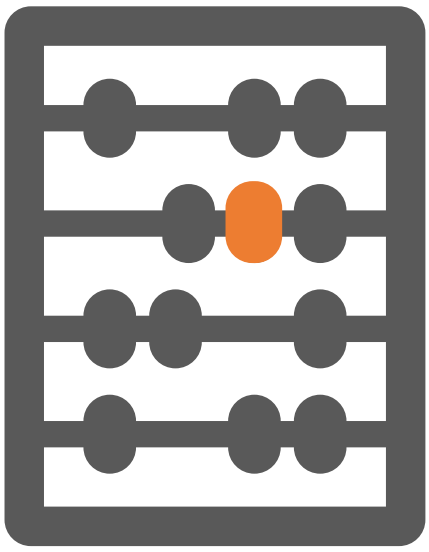
myOnlineTraininghub

Password Protected Structure

Files with password protected workbook structures will be slower to open and close than one without a password.

Solution: Given that Excel passwords can be removed easily, consider whether the password is worth the performance hit.

myOnlineTraininghub



User Defined Functions (UDFs)

UDFs are typically less efficient than the built in Excel functions.

Solution: consider using the built in functions, breaking them into separate calculations if required. Or write a custom function with the new LAMBDA function.

myOnlineTraininghub

Microsoft 365 is 250% Faster* than earlier Excel versions!

Outdated Software

Many improvements to Excel's calc engine were released for Office 365 users for SUMIFS, AVERAGEIFS, COUNTIFS, MAXIFS, MINIFS, and their singular counterparts as well as VLOOKUP, HLOOKUP and MATCH functions.

*The improvement is dramatic: for example, calculating 1200 SUMIFS, AVERAGEIFS, and COUNTIFS formulas aggregating data from 1 million cells on a 4 core 2 GHz CPU that took 20 seconds to calculate using Excel 2010, now only takes 8 seconds, in Excel M365.

Solution: where possible update to Microsoft 365 to take advantage of new improved functions and calc efficiencies. Also use the 64-bit version of Excel. The 32-bit version that is commonly used for compatibility with old add-ins only has 2GBs of virtual memory.

I help you STAND OUT from the crowd with skills that will TRANSFORM your CAREER

Get your skills up to speed with my courses
Follow me for tips and tutorials

Mynda Treacy | Microsoft MVP