Excel Tips #2: Many to One (altavia.com, 2016)

A co-worker called me the other day with a conundrum. The Excel VLOOKUP function was the perfect tool to solve his problem, but he had two columns to match not one column, as required by VLOOKUPs. The solution here, and for other Excel functions that require using a single column, is simple. But first let's review the powerful VLOOKUP function.

VLOOKUP Function

In database terms, VLOOKUP joins a key field in two tables and returns a field from the second table. Perhaps the easiest way to explain it is by way of an example. Let's say you have a column of numeric Cost Centers in your table, but for your final report you also want to include each Cost Center's description. You have another table with Cost Center description which you can manually look up each Cost Center, but if you have hundreds of Cost Centers this approach will never work. This is our starting point:

Table 1

	A	B	C	D
1	**Employee Name**	**Cost Center**	**CC Description**	**Other Data**
2	Employee 1	40059518	?	
3	Employee 2	40059553	?	…
4	Employee 3	40202952	?	…
5	Employee 4	40059519	?	…
6	Employee 5	40203686	?	…

We need the Cost Center Description

Which is available in a second table

Table 2 [Lookup Table Containing CC Description]

	A	B	C	D
1	**Cost Ctr**	**Description**	**Paragraph**	**UIC**
2	40059499	OFFICE OF THE DIRECTOR	200	W4T8AA
3	40059500	INNOVATIVE VENTURES OFFICE	201	W4T8AA
4	40059502	DIR OF STRATEGIC DEVELOPMENT	210	W4T8AA
5	40059503	TACTICS FORMULATIONS DIV	211	W4T8AA
6	…	…	…	…

We can use the VLOOKUP function in column C on Table 1 to return the Description from Table 2. The syntax is:

VLOOKUP (lookup_value, table_array, col_index_num, range_lookup)

Lookup_value → the Cost Center in table 1
Table_array → the second table (note that the Cost Center is in the leftmost column)
Col_index_num → 2, the second column in the table array holds the Description
Range_lookup → ALWAYS False or 0 for our purposes

This is what it looks like:

Table 1 Revised [showing formula and result]

	A	B	C	D
1	**Employee Name**	**Cost Center**	**Formula**	**Result**
2	Employee 1	40059518	=VLOOKUP($B2,'Table 2'!$A:$D,2,0)	STUDIES & ANALYSIS DIV
3	Employee 2	40059553	=VLOOKUP($B3,'Table 2'!$A:$D,2,0)	STRATEGIC ANALYSIS AND DESIGN
4	Employee 3	40202952	=VLOOKUP($B4,'Table 2'!$A:$D,2,0)	MANAGEMENT SUPPORT BR
5	Employee 4	40059519	=VLOOKUP($B5,'Table 2'!$A:$D,2,0)	MODELS & SIMULATIONS DIV
6	Employee 5	40203686	=VLOOKUP($B6,'Table 2'!$A:$D,2,0)	SECURITY DIVISION

The formula [=VLOOKUP ($B2,'Table2'!$A:$D,2,0)] says:

> Take the value in B2, lookup that value in Table 2, column A (the leftmost column in the table array), return the value in the second column of columns A:D in Table 2, and finally the "0" tells Excel to fetch an exact match only.

Not too bad, right? And extremely useful. Note that the key field, in this case Cost Center, does not need to be column A in Table 2 as in our example, but it must be the leftmost column in the table array used in the formula. There can be other columns to the left and right of the table array, but the key or lookup field (Cost Center above) must be the leftmost column of the table array (columns A:D above).

One final note on VLOOKUPs. The values in the key fields being matched between the two tables must match exactly. Two common problem areas, which you may need to address in one or both of the tables to get a match, are:

> (1) extra spaces before or following one of the values (use the TRIM function), or
> (2) comparing a number value with a text value – although they may look alike, a number and a text string with the same numeric characters are different to Excel (use the TEXT TO COLUMNS function).

Many to One

Now back to my co-worker's original conundrum where he wanted to use two columns for the key (or lookup value) in the VLOOKUP formula. However, that function is based on using a single key column in each of the two tables. So the easy trick is to combine the two columns into one column using the CONCATENATE function, which can be abbreviated using an "&".

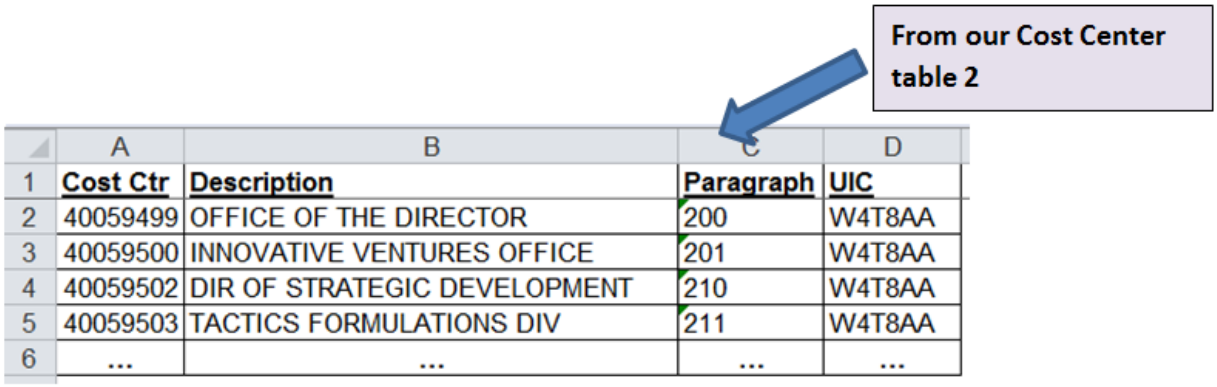
As an example, a problem from one assignment was a table where an employee's UIC and Paragraph were available in two columns, but the Cost Center was not. Since we knew that a UIC and Paragraph combination points to one unique Cost Center, we can use the UIC and Paragraph columns and our lookup table 2 from the above example:

Table 1

	A	B	C	D
1	**Employee Name**	**UIC**	**Paragraph**	**Cost Center**
2	Employee 1	W4T8AA	222C	**?**
3	Employee 2	W4T8AA	430A	**?**
4	Employee 3	W6HUAA	033B	**?**
5	Employee 4	W4T8AA	222B	**?**
6	Employee 5	W6HUAA	062A	**?**

The UIC and Paragraph can be used to point to a unique Cost Center

Table 2



	A	B	C	D
1	**Cost Ctr**	**Description**	**Paragraph**	**UIC**
2	40059499	OFFICE OF THE DIRECTOR	200	W4T8AA
3	40059500	INNOVATIVE VENTURES OFFICE	201	W4T8AA
4	40059502	DIR OF STRATEGIC DEVELOPMENT	210	W4T8AA
5	40059503	TACTICS FORMULATIONS DIV	211	W4T8AA
6	…	…	…	…

All we need to do is insert a new column A and concatenate the two columns using a delimiter of your choice; I like the plus sign "+". So the Concatenate formula looks something like =C2&"+"&D2, which says take whatever is in cell C2, append a + sign (the quotes before and after the + are needed), then append whatever is in cell D2.

I don't advise using the "+" delimiter if you are concatenating numbers (try an underscore). And be sure to *create the new concatenated column in both of your tables* (remember, it must be the leftmost column in the table array you use in the VLOOKUP formula). Here's the result:

Table 3

	A	B	C	D	E	F
1	**Key**	**Employee Name**	**UIC**	**Paragraph**	**Formula**	**Result (Cost Ctr)**
2	=C2&"+"&D2	Employee 1	W4T8AA	212	=VLOOKUP($A2,'Table 2'!$A:$E,2,0)	40059504
3	W4T8AA+220	E… 2	W4T8AA	220	=VLOOKUP($A3,'Table 2'!$A:$E,2,0)	40059509
4	W6HUAA+033B	Employee	W6HUAA	033B	=VLOOKUP($A4,'Table 2'!$A:$E,2,0)	40202952
5	W4T8AA+222	Employee 4	W…		…($A5,'Table 2'!$A:$E,2,0)	40059512
6	W6HUAA+005	Employee 5	W…		…($A6,'Table 2'!$A:$E,2,0)	40202940

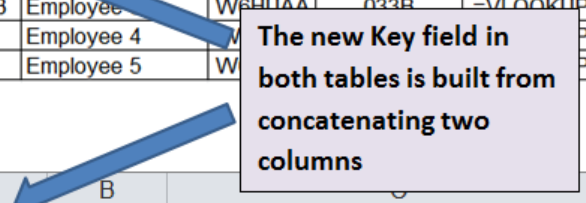


Table 2

	A	B	C	D	E
1	**Key**	**Cost Ctr**	**Description**	**Paragraph**	**UIC**
2	=E2&"+"&D2	40059499	OFFICE OF THE DIRECTOR	200	W4T8AA
3	W4T8AA+201	40059500	INNOVATIVE VENTURES OFFICE	201	W4T8AA
4	W4T8AA+210	40059502	DIR OF STRATEGIC DEVELOPMENT	210	W4T8AA
5	W4T8AA+211	40059503	TACTICS FORMULATIONS DIV	211	W4T8AA
6	…+…	…	…	…	…

Create a single column from multiple columns using the Concatenate function (i.e., &), then use that single column in Excel commands requiring a single column.

The formula in column E in Table 3 [=VLOOKUP ($A2,'Table2'!$A:$E,2,0)] says:

> Take the value in A2 in Table 3, lookup that value in Table 2, columns A through E (note that the Key field is the leftmost column in the table array), return the value in the second column of Table 2, and finally the "0" tells Excel to fetch an exact match only.

Last Things

The above concept can be used to combine 2, 3, or 10 columns, or as many as you like, into one column for Excel to manipulate. Excel commands other than VLOOKUP where this may come in handy are SUBTOTAL, COUNTIF, SUMIF (although the new SUMIFS does it better), and pivot tables.

If you use this approach with pivot tables, you will end-up with only the concatenated Key column for your variables after you pivot. You can break the Key column back into its original columns using the TEXT TO COLUMNS command on the DATA ribbon, selecting the delimiter option and providing the delimiter used. This will avoid creating the "stacked" or hierarchical table a pivot table generates by default when multiple key columns are used for row labels. (Note that options in Excel 2010 and 2013 can create flattened Pivot tables and get around the default hierarchical table structure).

Multiple columns to one. Simple. Formula on, dudes…