Excel Tips: Simple Ways to Compare Lists without VLOOKUPS (altavia.com, 2018)

I can't believe it's been almost two years since I've published an Excel Tips article. Time flies when you're busy and having fun! In this issue of Excel Tips we are going to discuss simple ways to compare similar lists without using VLOOKUP.

During analysis or while preparing ad-hoc reports, I often want to combine data from two short but similar lists or compare the records of two similar lists. Now, the VLOOKUP function can always help in these situations, but oftentimes I am doing things on the fly and don't want to take the time to set-up VLOOKUPs. In addition, it takes two VLOOKUPS going both ways to be sure two lists contain the same data. So here are few simple tricks to combine data or compare records from two similar lists without using VLOOKUPS.

To combine data from two short but similar lists, say fewer than 20 – 30 records, one can visually scan the two lists to see if the key field is in the same order, and if so, then merely copy columns of data from one list to the other. Although that works, I want to be sure my eyes aren't deceiving me and that every record key truly matches.

Table 1 is a list of Full Time Equivalent (FTE) employees in each Resource Pool (RP), broken into Direct FTE's and Salaried FTE's. Direct employees are the ones on the plant floor operating the machinery (the worker bees), whereas Salaried employees are generally management and administrative personnel.

Table 1 [FTE's by Resource Pool]

Resource Pool	RP Pool Name	Direct FTE's	Salaried FTE's	Total FTE's
RP_17001A	Quality Assurance Manager	1.00	2.00	3.00
RP_17001B	Quality Assurance Administration		3.75	3.75
RP_17001C	QA Engineer	2.00		2.00
RP_17001D	QA Manufacturing Metrics	4.00	1.00	5.00
RP_17001E	QA Molding	14.00		14.00
RP_17001F	QA Painting	4.00		4.00
RP_17001G	QA Assembly	3.75		3.75

Table 2 is the capacity in employee hours (Hrs) of each Resource Pool (RP), again broken into Direct and Salaried employees.

Table 2 [Capacity in Hours by Resource Pool]

Resource Pool	RP Pool Name	Direct Hrs	Salaried Hrs	Total Capacity
RP_17001A	Quality Assurance Manager	2,043	3,394	5,437
RP_17001B	Quality Assurance Administration		7,995	7,995
RP_17001C	QA Engineer	4,086		4,086
RP_17001D	QA Manufacturing Metrics	8,172	1,697	9,869
RP_17001E	QA Molding	20,430		20,430
RP_17001F	QA Painting	6,129		6,129
RP_17001G	QA Assembly	4,086		4,086

We want to quickly combine the data from Table 1 and Table 2. First use the SORT function on the key field in both lists, the Resource Pool field in this case. Then, when we visually scan the lists and the key field Resource Pool looks identical in the two tables, we can copy the last three columns of either table

to the other. Although that works, let's be sure every record's key field actually matches between the two tables.

Table 3 [Tables 1 + 2]

	A	B	C	D	E	F	G	H	I	J
1	**Table 1 [FTE's by Resource Pool]**						**Table 2 [Capacity in Hours by RP]**			
2	**Resource Pool**	**RP Pool Name**	**Direct FTE's**	**Salaried FTE's**	**Total FTE's**	**TEST**	**Resource Pool**	**Direct Hrs**	**Salaried Hrs**	**Total Capacity**
3	RP_17001A	Quality Assurance Manager	1.00	2.00	3.00	=A3=G3	RP_17001A	2,043	3,394	5,437
4	RP_17001B	Quality Assurance Admin.		3.75	3.75	TRUE	RP_17001B		7,994.54	7,994.54
5	RP_17001C	QA Engineer	2.00		2.00	TRUE	RP_17001C	4,086.00		4,086.00
6	RP_17001D	QA Manufacturing Metrics	4.00	1.00	5.00	TRUE	RP_17001D	8,172.00	1,697.00	9,869.00
7	RP_17001E	QA Molding	14.00		14.00	TRUE	RP_17001E	20,430.00		20,430.00
8	RP_17001F	QA Painting	4.00		4.00	TRUE	RP_17001F	6,129.00		6,129.00
9	RP_17001G	QA Assembly	3.75		3.75	TRUE	RP_17001G	4,086.00		4,086.00

First copy the lists side by side (with the key field in both lists sorted A-Z). A simple formula then does the trick. The formula "=A3=G3" will return TRUE if A3 equals G3 (the Resource Pool key fields), otherwise returns FALSE. A visual scan of column F for any FALSE's is a quick way to ensure the key fields in every record match before combining the data into Table 4.

Table 4 [Combined Data]

	A	B	C	D	E	F	G	H
1	**Resource Pool**	**RP Pool Name**	**Direct FTE's**	**Salaried FTE's**	**Total FTE's**	**Direct Hrs**	**Salaried Hrs**	**Total Capacity**
2	RP_17001A	Quality Assurance Manager	1.00	2.00	3.00	2,043	3,394	5,437
3	RP_17001B	Quality Assurance Admin.		3.75	3.75		7,994.54	7,994.54
4	RP_17001C	QA Engineer	2.00		2.00	4,086.00		4,086.00
5	RP_17001D	QA Manufacturing Metrics	4.00	1.00	5.00	8,172.00	1,697.00	9,869.00
6	RP_17001E	QA Molding	14.00		14.00	20,430.00		20,430.00
7	RP_17001F	QA Painting	4.00		4.00	6,129.00		6,129.00
8	RP_17001G	QA Assembly	3.75		3.75	4,086.00		4,086.00

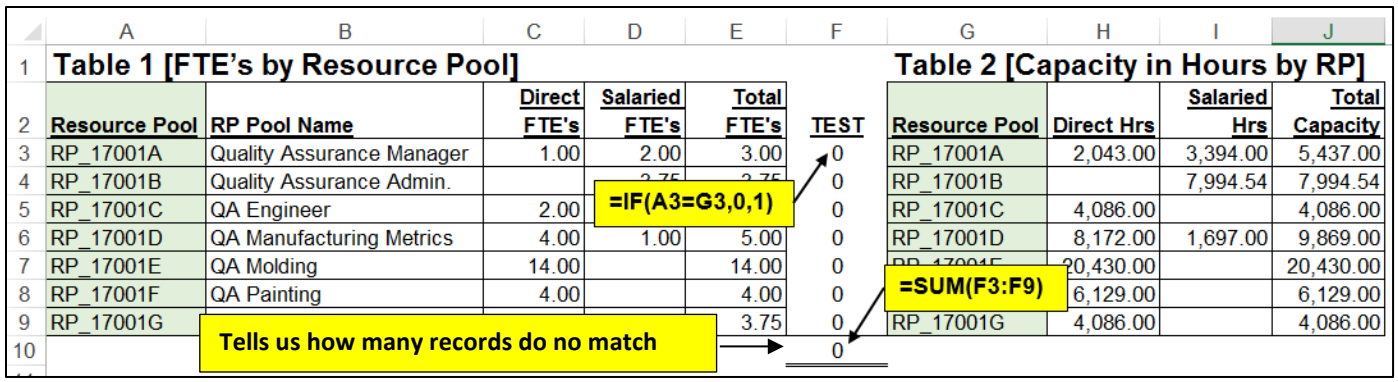
The TRUE/FALSE approach displayed in Table 3 above works well for a limited number of records (say, less than 30), but what if you want to quickly compare two similar lists with 500 records, or 2,000, or 100,000? Visually scanning becomes burdensome, or you could use the FIND function to look for any FALSE's, but there is a better way using the IF function.

We discussed the IF function in AVC's March 2016 Newsletter, which you might want to review if you're hesitant to use logic in your spreadsheets. The syntax for the IF function is:

IF(logical_test, value_if_true, value_if_false)

We want to test to see if the key field is identical in both lists, and if so return a 0, if not return a 1. Thus, the formula in the above example for Table 3 becomes "=IF(A3=G3,0,1)". Copy the formula to the last row of data and then sum the column as in the figure below. If the sum at the bottom of the Test column is 0, all records match. If not, the sum tells you how many records do not match.

Table 5 [Using the IF Function]



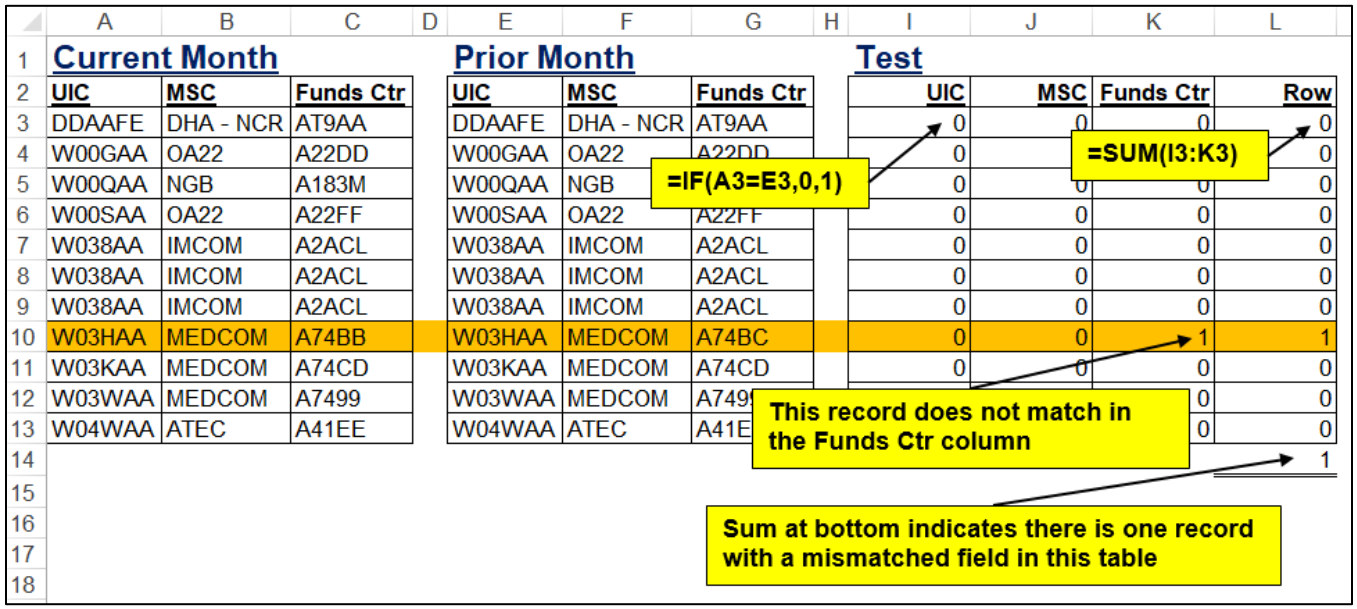
Resource Pool	RP Pool Name	Direct FTE's	Salaried FTE's	Total FTE's	TEST
RP_17001A	Quality Assurance Manager	1.00	2.00	3.00	0
RP_17001B	Quality Assurance Admin.				0
RP_17001C	QA Engineer	2.00			0
RP_17001D	QA Manufacturing Metrics	4.00	1.00	5.00	0
RP_17001E	QA Molding	14.00		14.00	0
RP_17001F	QA Painting	4.00		4.00	0
RP_17001G				3.75	0
					0

Resource Pool	Direct Hrs	Salaried Hrs	Total Capacity
RP_17001A	2,043.00	3,394.00	5,437.00
RP_17001B		7,994.54	7,994.54
RP_17001C	4,086.00		4,086.00
RP_17001D	8,172.00	1,697.00	9,869.00
RP_17001E	20,430.00		20,430.00
RP_17001F	6,129.00		6,129.00
RP_17001G	4,086.00		4,086.00

The IF function approach can be extended to compare all the fields in two tables (not just the key field), to see if all of the records are identical or where there are differences. Lay the two tables to be compared side by side in one worksheet, and sort each by its key field(s). Create a third table with the same column headings but no data.

In the upper left-hand corner of the third table write an IF formula returning a 0 if the first column, first record of the first table matches the first column, first record of the second table. Copy that formula across for all fields in the third table, and then create a SUM across the fields that will be 0 if all the fields match between the two tables for the same row. Here's an abbreviated example:

Table 6 [Comparing All Fields Using IF Functions]



UIC	MSC	Funds Ctr
DDAAFE	DHA - NCR	AT9AA
W00GAA	OA22	A22DD
W00QAA	NGB	A183M
W00SAA	OA22	A22FF
W038AA	IMCOM	A2ACL
W038AA	IMCOM	A2ACL
W038AA	IMCOM	A2ACL
W03HAA	MEDCOM	A74BB
W03KAA	MEDCOM	A74CD
W03WAA	MEDCOM	A7499
W04WAA	ATEC	A41EE

UIC	MSC	Funds Ctr
DDAAFE	DHA - NCR	AT9AA
W00GAA	OA22	A22DD
W00QAA	NGB	A183M
W00SAA	OA22	A22FF
W038AA	IMCOM	A2ACL
W038AA	IMCOM	A2ACL
W038AA	IMCOM	A2ACL
W03HAA	MEDCOM	A74BC
W03KAA	MEDCOM	A74CD
W03WAA	MEDCOM	A7499
W04WAA	ATEC	A41E

UIC	MSC	Funds Ctr	Row
0	0	0	0
0			0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	1	1
0	0	0	0
		0	0
		0	0
			1

So we see that comparing two similar tables on the fly without using VLOOKUPS can save time and, when used correctly, point to any problems with combining the two tables. Formula on, dudes…