

OPA Excel Tips: Merging in data from another sheet (XLOOKUPS)

There are a number of situations where it is helpful to merge two datasets or copy a variable from one dataset to another. The XLOOKUP function in Excel is a very handy way to do this.

A XLOOKUP looks for a specified value in an array of data (usually in another worksheet or workbook) and returns the value from a specified column. XLOOKUP is meant to be a replacement for a number of previously used look ups including VLOOKUP, HLOOKUP, and MATCH. XLOOKUP is more efficient because like VLOOKUP, you are able to have the simplistic formula but have the flexibility of MATCH (no requirement of where the cells are in the sheet).

The XLOOKUP command requires only three pieces of information:

Look up value: The value you want to look up e.g., and Application ID (Appl ID). This is usually a cell reference and changes for each line of the dataset. It is possible to look up any cell value (text or numeric) but the method is most reliable when using an ID number as text (e.g. institution name) may have different spellings or a space after the final letter for example. These differences would prevent Excel from making the match.

It is also important to consider the two datasets you are using e.g. if there are multiple rows per project number Excel will just copy across the value from the first match.

Table array: The range of data that you want to look to for the variable of interest.

Return array: The range of data that contains the variable you are interested in. This can be on the left or right of the value of interest and works independently of the table array.

A general example referring values both on the right and left of the variable of interest:

SUM : ✕ ✓ fx =XLOOKUP(F2,B2:B7,C2:C7)							
	A	B	C	D	E	F	G
1	Organization	Acronym	Location			What is the location?	
2	National Institutes of Health	NIH	MD			FDA	MD
3	Centers for Disease Control and Prevention	CDC	MD				
4	Food and Drug Administration	FDA	MD			What is the full organization name?	
5	Environmental Protection Agency	EPA	DC			FDA	Food and Drug Administration
6	Federal Bureau of Investigation	FBI	DC				
7	Central Intelligence Agency	CIA	VA				

G5 : ✕ ✓ fx =XLOOKUP(F5,B2:B7,A2:A7)							
	A	B	C	D	E	F	G
1	Organization	Acronym	Location			What is the location?	
2	National Institutes of Health	NIH	MD			FDA	MD
3	Centers for Disease Control and Prevention	CDC	MD				
4	Food and Drug Administration	FDA	MD			What is the full organization name?	
5	Environmental Protection Agency	EPA	DC			FDA	Food and Drug Administration
6	Federal Bureau of Investigation	FBI	DC				
7	Central Intelligence Agency	CIA	VA				

Example 1: Merging Foreign grant flag onto the Transplantation dataset

In this example, the Transplantation_iSearch dataset is used.

Title	Abstract	Specific Aim	Appl Stat	Appl Ty	Award Stat	Appl ID	Grant Numbe	Full Grant Numbe	Fiscal Year	RFA Notice	Th RFA PA Numb	Organization	Total Cost
The role of Allogeneic Graft-versus-			5	1	Awarded	8364475	R21AI101932	R21AI101932-01	2012	NIH EXPLORAT	PA11-261	The University	237000
THE IMM Corneal ti Corneal trans			5	2	Awarded	8437711	R01EY007641	R01EY007641-25	2013	RESEARCH PR	PA11-260	UT Southweste	397500
Gene-Eng Current a : Current und			5	1	Awarded	7590067	R01CA133737	R01CA133737-01	2009	RESEARCH PR	PA07-070	University of Al	291538
Glioblasti Gliomas : Glioblastom			5	1	Awarded	7736551	R01NS064983	R01NS064983-01	2009	RESEARCH PR	PA07-070	Massachusetts	366301
Marrow-li Disease r Blood and m			5	2	Awarded	8291667	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025	Johns Hopkins	338706
Administ, The purp, The purpose			5	2	Awarded	8291673	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025	Johns Hopkins	298550
Targeted, Despit, ti Bone marrow			6	1	Awarded	8594206	F32HL115963	F32HL115963-01	2013	RUTH L. KIRSC	PA11-113	Stanford Univer	53942
Targeting The impo Therapeutic i			5	2	Awarded	8291662	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025	Johns Hopkins	343404
Evaluatin, This is a, The project ti			5	2	Awarded	7987158	R01CA145207	R01CA145207-05	2010	RESEARCH PR	PA07-070	Beckman Rese	405299
Improved The liver Orthotopic li			5	2	Awarded	7579583	R01DK063275	R01DK063275-06	2009	RESEARCH PR	PA07-070	Beth Israel Dec	408000

Assume we want to add on the 'Foreign grant' QVR flag. e.g., to find out how many of the applications have been awarded as a foreign grant and that we have this data in a separate Excel file.

An Excel file downloaded from QVR with this foreign grant flag and appl ID is shown below. It is also available on the OPA training pages as Transplantation_additional QVR data.xlsx.

Appl ID	Actv	Admin IC	Foreign Grant	Institution
7791728	U01	DK	1	UNIVERSITY HEALTH NETWORK
8190759	R01	DK	1	MCGILL UNIVERSITY
8391285	R21	OD	1	UNIVERSITY OF CAMBRIDGE
8426987	R21	NS	1	UNIVERSITY HEALTH NETWORK
8513552	R01	OD	1	UNIVERSITY OF CALGARY
7581353	R01	HL	2	JOSLIN DIABETES CENTER
7649736	R01	CA	2	IMMUNE DISEASE INSTITUTE, INC.
7715411	U01	AI	2	BRIGHAM AND WOMEN'S HOSPITAL
7730010	R01	AG	2	STANFORD UNIVERSITY

Note both datasets have the same identifying variable (Appl ID) which is vital to enable XLOOKUP to work.

First, add a header to new column to denote the field you are adding in.

Title	Abstract	Specific Aim	Appl Stat	Appl Ty	Award Stat	Appl ID	Grant Numbe	Full Grant Numbe	Fiscal Year	RFA Notice	Th RFA PA Numb	Organization	Total Cost	Foreign Grant
The role of Allogeneic Graft-versus-			5	1	Awarded	8364475	R21AI101932	R21AI101932-01	2012	NIH EXPLORAT	PA11-261	The University	237000	
THE IMM Corneal ti Corneal trans			5	2	Awarded	8437711	R01EY007641	R01EY007641-25	2013	RESEARCH PR	PA11-260	UT Southweste	397500	
Gene-Eng Current a : Current und			5	1	Awarded	7590067	R01CA133737	R01CA133737-01	2009	RESEARCH PR	PA07-070	University of Al	291538	
Glioblasti Gliomas : Glioblastom			5	1	Awarded	7736551	R01NS064983	R01NS064983-01	2009	RESEARCH PR	PA07-070	Massachusetts	366301	

Next, enter the XLOOKUP formula in the first blank cell of the newly created column (O2). The formula is:

=XLOOKUP([cell with ID value of interest],[group of cells to look to for values],[group of cells with number of interest])

- In this instance, we want to look up based on Appl ID so in the lookup formula in cell O2, the value to look up is cell G2.
- The group of cells (array) we're interested in is within the Transplantation_additional QVR data worksheet, and this sheet needs to have Appl ID in the group of referenced cells. It does not matter if these reference cells are on the right or left of the column you want added into the other sheet. Select the data in the column (A).
- The QVR Foreign grant flag is in column D of this same reference worksheet, so for the return array, we select this column (D).

Before copying the formula to the entire column, make sure that there are '\$' symbols before the letters and numbers in the array cell references, otherwise when copying and pasting the array may change. The '\$' symbol tells Excel to fix the column and row values in the formula. When using an array in a different workbook, Excel does this automatically, when referring to a different worksheet in the same workbook, you need to do it manually.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Title	Abstract	Specific Aim	Appl Stat	Appl Ty	Award Stat	Appl ID	Grant Numb	Full Grant Numb	Fiscal Yr	RFA Notice	Ti RFA	PA Numb	Organization	Total Cost	Foreign Grant		
2	The role of Allogeneic Graft-versus-			5	1	Awarded	8364475	R21AI101932	R21AI101932-01	2012	NIH EXPLORAT	PA11-261		The University	237000			
3	THE IMM Corneal t Corneal trans			5	2	Awarded	8437711	R01EY007641	R01EY007641-25	2013	RESEARCH PR	PA11-260		UT Southweste	397500			
4	Gene-Eng Current a : Current und			5	1	Awarded	7590067	R01CA133737	R01CA133737-01	2009	RESEARCH PR	PA07-070		University of Al	291538			
5	Glioblasti Gliomas : Glioblastom			5	1	Awarded	7736551	R01NS064983	R01NS064983-01	2009	RESEARCH PR	PA07-070		Massachusetts	366301			
6	Marrow-Ii Disease r Blood and me			5	2	Awarded	8291667	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025		Johns Hopkins	338706			
7	Administ The purp The purpose			5	2	Awarded	8291673	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025		Johns Hopkins	298550			
8	Targeted Despite ti Bone marrow			6	1	Awarded	8594206	F32HL115963	F32HL115963-01	2013	RUTH L. KIRSC	PA11-113		Stanford Univei	53942			
9	Targeting The impo Therapeutic			5	2	Awarded	8291662	P01CA015396	P01CA015396-37	2012	NATIONAL CAI	PAR09-025		Johns Hopkins	343404			
10	Evaluatin This a r The project ti			5	2	Awarded	7987158	R01CA145207	R01CA145207-05	2010	RESEARCH PR	PA07-070		Beckman Rese.	405299			
11	Improved The liver Orthotopic liv			5	2	Awarded	7579583	R01DK063275	R01DK063275-06	2009	RESEARCH PR	PA07-070		Beth Israel Dee	408000			
12	Derivatior Human et Pluripotent st			5	1	Awarded	7565096	R01GM08649	R01GM086497-01	2009	HUMAN PLURI	PA08-043		University of W	312000			
13	Engineeri Every yea Principal Inve			5	2	Awarded	7736701	R01EY015500	R01EY015500-05	2009	RESEARCH PR	PA07-070		Northeastern U	405824			
14	Optimizin The objec Our central h			5	1	Awarded	7736767	R01HL093131	R01HL093131-01	2009	RESEARCH PR	PA07-070		Massachusetts	736855			
15	Cellular a Systemic RESEARCH PI			5	2	Awarded	7992940	R01AR047988	R01AR047988-10	2010	RESEARCH PR	PA07-070		University of Vi	487613			
16	Tissue En There are Bone allograf			5	1	Awarded	8476900	R01AR064200	R01AR064200-01	2013	NANOSCIENCE	PA11-148		University of Ri	293569			

To copy the formula down the entire column, you can hover over the bottom right corner of the cell in which you just added the formula to and double click on the cross hair that come up. This will copy the formula all the way to the end of the adjacent data, changing the lookup cell (G2) to each corresponding ID in that row, while keeping the other parts of the formula intact because of the column aspect and '\$' symbol.

Example 2: Merging Activity code onto the Transplantation dataset.

If you have worked through the first example, you can copy and paste the formula from cell O2 into cell P2.

The screenshot shows an Excel spreadsheet with a formula in cell O2: `=XLOOKUP(G2, [Transplantation_additional QVR data.xlsx]!$A:$A, [Transplantation_additional QVR data.xlsx]!$D:$D)`. The formula is looking for the value in cell G2 (Grant Number) in the 'Appl ID' column (A:A) of the 'Transplantation_additional QVR data.xlsx' spreadsheet. The result is an error message: #N/A. A context menu is open over the error, showing options like Cut, Copy, Paste Options, etc.

This gives an error message as it's looking for the value in cell 'H2' (Grant Number) in the Appl ID column (A:A) of the Transplantation_additional QVR data, which we do not have in the spreadsheet.

The screenshot shows the same Excel spreadsheet with the formula in cell P2 corrected to: `=XLOOKUP(H2, [Transplantation_additional QVR data.xlsx]!$A:$A, [Transplantation_additional QVR data.xlsx]!$D:$D)`. The result is now #N/A. The formula bar shows the corrected formula.

Change the lookup value from 'H2' to 'G2' and the return array from column D to column B to pull in the activity code. Because we had '\$' in front of the reference columns in the formula (\$D:\$D), these copied over unchanged, and it was using the Appl ID as a unique identifier to match in the additional data.

Corrected and copied formula to the end of the dataset in Transplantation_iSearch:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Applicant ID	Activity	Admin ID	Foreign Grant	Institution														
2	7791728	U01	DK		1 UNIVERSITY HEALTH NETWORK														
3	8190759	R01	DK		1 MCGILL UNIVERSITY														
4	8391285	R21	OD		1 UNIVERSITY OF CAMBRIDGE														
5	8426987	R21	NS		1 UNIVERSITY HEALTH NETWORK														
6	8513552	R01	OD		1 UNIVERSITY OF CALGARY														
7	7581353	R01	HL		2 JOSLIN DIABETES CENTER														
8	7649736	R01	CA		2 IMMUNE DISEASE INSTITUTE, INC.														
9	7715411	U01	AI		2 BRIGHAM AND WOMEN'S HOSPITAL														
10	7730010	R01	AG		2 STANFORD UNIVERSITY														

AutoSave [icon] [

Note: You can also use other identifiers in both spreadsheets such as the grant number to match the same information across worksheets. Keep in mind it must be the core grant number or full grant number in both sheets (exact match).

More help

Click in the formula box, on 'XLOOKUP', this will bring up the pop-up box shown below with a hyperlink to the Excel XLOOKUP help text.

AutoSave

Transplantation_Search_062022_Search...Grants-export_2022-06-23-15-5...

Search

Bibb, Katrina (NIH/OD) [C] [K]

Share

Comments

FileHomeInsertPage LayoutFormulasDataReviewViewHelpAcrobat

CutCopyFormat Painter

Font

Alignment

Number

Conditional Formatting

Styles

Cell Styles

InsertDeleteFormat

Cells

AutoSum

Sort & Filter

Find & Select

Sensitivity

ClipboardFontAlignmentNumberStylesCellsEditingSensitivity

SUM

=XLOOKUP(

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Title	Abstract	Specific Aim	Applicant ID	Status	Award Status	Applicant ID	Grant Number	Full Grant Number	Fiscal Year	RFA Notice Title	RFA PA Number	Organization	Total Cost	Foreign Grant	Activity		
2	The role of Allogeneic Graft-versus-			8364475	5	Awarded	R21AI101932	R21AI101932-01		2012	NIH EXPLORAT PA11-261		The University	237000		R21		
3	THE IMM Corneal t Corneal trans			8437711	5	Awarded	R01EY007641	R01EY007641-25		2013	RESEARCH PR(PA11-260		UT Southweste	397500		R01		UP
4	Gene-Eng Current a: Current und			7590067	5	Awarded	R01CA133737	R01CA133737-01		2009	RESEARCH PR(PA07-070		University of Al	291538		R01		
5	Glioblasti Gliomas i: Glioblastom			7736551	5	Awarded	R01NS064983	R01NS064983-01		2009	RESEARCH PR(PA07-070		Massachusetts	366301		R01		
6	Marrow-I Disease r Blood and me			8291667	5	Awarded	P01CA015396	P01CA015396-37		2012	NATIONAL CAI PAR09-025		Johns Hopkins	338706		P01		
7	Administ The purp The purpose			8291673	5	Awarded	P01CA015396	P01CA015396-37		2012	NATIONAL CAI PAR09-025		Johns Hopkins	298550		P01		
8	Targeted Despit t Bone marrow			8594206	6	Awarded	F32HL115963	F32HL115963-01		2013	RUTH L. KIRSC PA11-113		Stanford Univer	53942		F32		
9	Targeting The impo Therapeutic i			8291662	5	Awarded	P01CA015396	P01CA015396-37		2012	NATIONAL CAI PAR09-025		Johns Hopkins	343404		P01		
10	Evaluatin This is a i The project t			7987158	5	Awarded	R01CA145207	R01CA145207-05		2010	RESEARCH PR(PA07-070		Beckman Rese	405299		R01		
11	Improved The liver Orthotopic liv			7579583	5	Awarded	R01DK063275	R01DK063275-06		2009	RESEARCH PR(PA07-070		Beth Israel Dea	408000		R01		
12	Derivat Human ei Pluripotent st			7565096	5	Awarded	R01GM086497	R01GM086497-01		2009	HUMAN PLURI PA08-043		University of W	312000		R01		
13	Engineer Every yea Principal Inve			7736701	5	Awarded	R01EY015500	R01EY015500-05		2009	RESEARCH PR(PA07-070		Northeastern U	405824		2 R01		
14	Optimizin The objec Our central h			7736767	5	Awarded	R01HL093131	R01HL093131-01		2009	RESEARCH PR(PA07-070		Massachusetts	736855		R01		
15	Cellular a Systemic RESEARCH Pl			7992940	5	Awarded	R01AR047988	R01AR047988-10		2010	RESEARCH PR(PA07-070		University of Vi	487613		R01		