Excel Lesson: Table Lookup Functions

Topics Covered:

VLookup() [Look across] HLookup() [Look down] Lookup() [Look almost anywhere] Related Functions (a list)

We will not be examining all forms of these functions – a single example will be given of them and variations won't be. Examples will use the "find closest match" version of them, rather than the alternative "find exact match" versions.

Real World Example Application:

The various lookup functions can be used to find related information contained in a data table. The data table may be a single column or row of information or it may be a complex table containing many rows and columns. For example, you might have a table of classroom student information containing information such as days absent, individual test scores, average of all test scores, etc. With the HLookup() and VLookup() functions digging out that information for use elsewhere in your workbook becomes almost a trivial exercise.

Lesson Support Workbook: ExcelLookupsLesson.xls¹ (at same web site/page that this lesson came from).

VLOOKUP()

We will tackle VLOOKUP() first because it seems to be the one most often used. VLOOKUP() uses the data in the leftmost column of a table to perform the lookup or match to know which row to look across to find the related information.

VLOOKUP() takes three mandatory parameters and one default/optional parameter:

Mandatory parameter 1: Called Lookup_value in Excel Help, it is the value to be found in the first column of the Table_array. Lookup_value can be a value, a reference, or a text string.

Mandatory parameter 2: Called Table_array in Excel Help, it is the table of information in which data is looked up. Use a reference to a range or a range name. If you are unfamiliar with working with named ranges, read this document <u>ExcelBasicsLesson.pdf</u>². We will be using named ranges in this lesson later.

Mandatory parameter 3: Called Col_index_num in Excel Help, it is the column number in table_array from which the matching value must be returned. A col_index_num of 1 returns the value in the first column in table_array; a col_index_num of 2 returns the value in the second column in table_array, and so on. If col_index_num is less than 1, VLOOKUP returns the #VALUE! error value; if col_index_num is greater than the number of columns in table_array, VLOOKUP returns the #REF! error value.

¹ <u>http://www.jlathamsite.com/teach/excellookupslesson.xls</u>

² <u>http://www.jlathamsite.com/teach/excelbasicslesson.pdf</u>

Optional Parameter: Called Range_lookup in Excel Help, it is a logical value that specifies whether you want VLOOKUP to find an exact match or an approximate match. If <u>TRUE or</u> <u>omitted</u>, an approximate match is returned. In other words, if an exact match is not found, the next largest value that is less than lookup_value is returned. If FALSE, VLOOKUP will find an exact match. If one is not found, the error value #N/A is returned.

NOTE: If the optional parameter, Range_lookup, is either TRUE or omitted (thus defaulting to TRUE) then the data in the first column of the array *must* be sorted in ascending order for VLOOKUP() to work properly.

NOTE: If you use the FALSE value for the optional parameter, then the search is done for an exact match to the **Mandatory parameter 1.**

Example Discussion: You are a teacher and during parent's visiting days you wish to be able to show a parent their child's test scores, but without exposing the scores of the other students to them. You keep your class roster and test scores, including the average on a worksheet called TestScores, in a table you have named GradeTable, which includes a range from B2 to I12 on that sheet. The first/leftmost column (B) contains the student names.

To solve your problem, you need to set up a second worksheet that you will call OneStudent that allows you to type in a student's name and show only that student's grades.

	A	В	С	D	E	F	G	Н	I
1		Name	Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade
2		Arthur	80	78	84	92	82	83	83.17
3		Bea	72	100	64	84	78	88	81.00
4		Cindy	68	74	70	72	74	70	71.33
5		Harold	88	92	94	96	98	100	94.67
6		Helen	100	100	96	96	100	98	98.33
7		Joe	58	62	60	70	66	68	64.00
8		Mabel	98	98	98	100	98	96	98.00
9		Mary	74	84	86	92	76	78	81.67
10		Ralph	42	60	66	56	60	58	57.00
11		Sam	86	84	100	92	84	88	89.00
12		Tim	42	60	66	56	60	58	57.00
40									

Lets begin by looking at the GradeTable on sheet TestScores:

First, realize that everything in color (both red and light teal) has been named GradeTable, so when we use the phrase GradeTable as a range, it is the same as writing TestScores!B2:I12 Either of those can be used as the **Second Mandatory parameter** in the VLOOKUP() function.

The column in that table with the names is the first column of the table and so it is column 1 of the table, columns C through I are columns 2 through 8 of the table, in order. Notice that the list of names has already been sorted alphabetically in ascending order.

We add a new sheet to the workbook and name it OneStudent and begin setting it up. We will take it step by step, but I'll show you what it looks like when all is said and done:

В	С	D	E	F	G	Н	
Enter Student's Name	Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade
Bea	72	100	64	84	78	88	81
	72	100	64	84	78	88	81

There is a reason that the scores are shown twice – we will cover that later. In the 'real world' there would only need to be a single row of information. Now you know what we are looking to create, lets move on.

But first, a quick look at a 'cheat' for this sheet (making it a cheat sheet?) just for this example. Ordinarily you'd probably set up some kind of pull-down list with the student names so you could easily choose one. But for this, I just copied all the student names from the TestScores sheet over to this one so we could have a handy reference to them without jumping back and forth – remember, we don't want the parents to ever see that TestScores sheet with all student test scores on it!

Enter Student's Name	Test 1		Test 2	Mid
Bea		72	100	
		72	100	
Reminder: Students				
Arthur				
Bea				
Cindy				
Harold				
Helen				
Joe				
Mabel				
Mary				
Ralph				
Sam				
Tim				

Alright, now to backup some. You have entered the column labels (Enter Student's Name, Test 1, etc) on this new OneStudent sheet and are at a bit of a loss as to what to do next. How about we begin by entering a VLOOKUP() formula for a student whose name will appear in cell B2 on this sheet into the Test 1 score location (cell C2).

Type the following into cell C2 and end it with the [Enter] key: =VLOOKUP(\$B\$2,GradeTable,2)

Oops – looks like an error? ...

	A	В	С
1		Enter Student's Name	Test 1
2			#N/A
2			

Well, maybe not – what is missing here? We told it to use the value in cell B2 to match within the GradeTable, but we don't have anything in B2 right now. Maybe if we put a valid entry in B2 then things might look different? Let's use Bea since it's short to type and is a valid entry.

В	C	
Enter Student's Name	Test 1	
Bea		72

Ok that fixed it. Having a missing piece of the puzzle found sure helps evaluate whether or not things are working. Onward...

Let us examine the formula we entered before moving on.

=VLOOKUP(\$B\$2,GradeTable,2)

The equals symbol (=) tells Excel that this is a formula, all Excel formulas start with one of those.

VLOOKUP(

Tells Excel that we're setting up to use that built-in function of Excel.

\$B\$2 – again, if the format of the cell reference confuses you, please go read that linked to .pdf document I mentioned earlier. Basically using the \$ sign in front of each part of the address (letter and number) we're telling Excel not to automatically change either one as we move or copy it around on this sheet or within this workbook. We are saying that the value we want to use for the **First Mandatory parameter** is going to always be located in cell B2 on this sheet.

The comma separates the first parameter from the second.

The **Second Mandatory parameter** is the range that defines the data table. We gave that table the name GradeTable, so we can type that in here.

That next comma is there to separate the second parameter from the last one we're using.

2 - this tells Excel that when a match is found, we want the information from the second (2^{nd}) column of the table. And with that done, and knowing we want an exact match and that the information in the first column of the table is in ascending order, we end things up by closing the parameter list with a) (right parenthesis symbol).

Alright, seems we are ready to move on and fill up the rest of the cells in this row with that formula by extending the formula across them. How you fill/extend it is up to you. If you're without clue on this, go back and read that referenced .pdf file before continuing!!!

В	С	D	E	F	G	Н	
Enter Student's Name	Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade
Bea	72	72	72	72	72	72	72

But wait! Once again, things are not as we expected. All of Bea's scores were NOT the same!

Well, well. We finally found a time when Excel got stupid on us! If you look at any of the formulas, you will find that they are all exactly the same. They all use the value **2** for the **Third Mandatory parameter**. So you'll have to fix that manually. Change the one in C2 to a 3, in D2 to a 4, in E2 to a 5, F2 to a 6, G2 to a 7 and the one in I2 to an 8. Then it should look like this:

В	С	D	E	F	G	Н	1
Enter Student's Name	Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade
Bea	72	100	64	84	78	88	81

Now that wasn't really so hard was it? The principles shown here work for any table set up on a sheet of any size that Excel can handle. And with VLOOKUP() that is a lot – since you could have up to 65,535 rows in the table to hold things to be matched to, along with 254 items to select from based on the match! For the curious, that is a table with some 16,711,425 cells in it!

Oh, about that second row of reported scores, what was the deal there?

Well, try typing this formula into Cell C3 (right below the 72 score that Bea had): =VLOOKUP(**\$B\$2,TestScores!B2:I12,2,TRUE**)

Notice that I've used the 'raw' name for the data table on the TestScores sheet, and I explicitly included that fourth, optional, parameter for the type of lookup to perform. Just as I promised you the result was the same. You can extend this across, remembering to fix the offset value into the table for the return value, and see that this format of the formula works just like the other does.

Summary:

First, let me say that using Named Ranges for things like tables is much easier on maintenance of the workbook in the long run. You can move that table anywhere in the workbook without having to change any of the formulas on the new sheet. You would have to 'fix them up' if you moved them and used the second, 'raw', form of the addressing of that table.

Of all the various Lookup functions, VLOOKUP is probably the most versatile because most people can see how it works right away, and because the key column, that first column with the list to match, is a column, you can usually see more of it on a page than you could with an HLOOKUP() setup where the to-be-matched values are in columns across the sheet.

Finally, you can get more items in your list with VLOOKUP than you can with HLOOKUP. You can have as many as 65,535 with VLOOKUP(), but you have a maximum of 255 with HLOOKUP() because Excel only deals with up to 255 columns. But sometimes, it turns out that maybe turning things on edge and then using HLOOKUP() is a better choice. It just all depends.

HLOOKUP()

There really isn't too much to say about HLOOKUP() that hasn't been said about VLOOKUP(). HLOOKUP() is the twisted twin of VLOOKUP(). Instead of the lookup, to-be-matched items being in the first column, they are in the first row, and the returned information is in rows below that.

But a quick discussion would seem appropriate - based on our previous VLOOKUP() example.

First, the GradeTable gets set up like this on a sheet named HTestScores, and it was given the name HGradeTable:

	A	B	C	D	E	F	G	Н		J	K	L
1												
2	Name	Arthur	Bea	Cindy	Harold	Helen	Joe	Mabel	Mary	Ralph	Sam	Tim
З	Test 1	80	72	68	88	100	58	98	74	42	86	42
4	Test 2	78	100	74	92	100	62	98	84	60	84	60
5	Mid Term	84	64	70	94	96	60	98	86	66	100	66
6	Test 4	92	84	72	96	96	70	100	92	56	92	56
7	Test 5	82	78	74	98	100	66	98	76	60	84	60
8	Final	83	88	70	100	98	68	96	78	58	88	58
9	Avg Grade	83.17	81.00	71.33	94.67	98.33	64.00	98.00	81.67	57.00	89.00	57.00
40												

And then our lookup page looks like this:

	A	В	С	D	E	F	G	Н		
1		Enter Student's Name	Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade	
2		Bea	72	100	64	84	78	88	81	
3			72	100	64	84	78	88	81	
4			=HLOOKU	P(\$B\$2,HG	FradeTable,					
5			=HLOOKUP(\$B\$2,HTestScores!B2:L9,2)							
_				1						

I've shown the modified formulas for the two lookups done in column C in the graphic.

Some people find looking at information in this layout easier than the layout needed for VLOOKUP() to work with. It's a personal choice if you are designing the tables from the ground up. There are times when someone may hand you a table in one format and you don't want to have to rebuild it, so you 'go with the flow'. However, **a hint**: check out the Edit / Copy and Edit /Paste Special functions with the "transpose" option checked! It will rotate tables like these two back and forth for you!!

One limit of HLOOKUP() is that your lookup list can only have a maximum of 255 entries in it. More than enough for most people. But it can have up to 65,534 elements below each of those lookup entries.

Do try to plan ahead a little. I once worked on a project where the list consisted of equipment names with their part numbers, costs, and some two or three other pieces of information about them provided in a list setup to do an HLOOKUP() with. Since that's the way it was, I 'went with the flow' and designed the rest of the package around that layout. Later on they kept adding equipment to the list until they finally asked me to add in #256. Oops! Had to turn the whole thing on edge, and then go rewrite all the HLOOKUP() formulas to then use VLOOKUP().

Copyright © 2004, Jerry L. Latham, All Rights Reserved. Some portions Copyright © by The Microsoft Corporation. Page 6 of 9

LOOKUP()

What's so great about this function? Well, it lets you do some things you cannot do with either VLOOKUP() or HLOOKUP(), such as:

- > Look up things in one 'table' or list and pull things out of another one.
- > Letting you look left, right, above or below the lookup table or list.
- Letting you lookup data on one sheet and pull information from another sheet!
- Letting you look up in a horizontally laid out table/list and look up in a vertically laid out table/list.
- Letting you look up in a vertically laid out table/list and look up in a horizontally laid out table/list.

How is all of this accomplished? An examination of the syntax of the LOOKUP() function will explain this clearly, I hope:

=LOOKUP(lookup_value, lookup_range, result_range)

The first two parameters are very much like what we've seen before in VLOOKUP() and HLOOKUP().

Parameter #1: lookup_value is exactly as we've been dealing with – it is the value or a reference to a value that we want to look up in the lookup_range.

Parameter #2: lookup_range – almost, but not quite the same as we've used before. In this case it is a range that refers to a single-row or a single-column range that contains the data to match the lookup_value to. Behind the scenes, Excel does the lookup and keeps an index value as to where in this list the match was found.

Parameter #3: result_range – previously we'd entered either a row or column offset value in this position. Now we are entering a single-row or single-column range that contains the values to be returned based on the match found using **lookup_value** and **lookup_range** (that behind the scenes index value I mentioned in the last paragraph). Obviously the order and number of entries in this **result_range** must match those in the **lookup_range** list.

The worksheet shown on the next page shows how some of this works.

	A	В	C	D	E	F	G	Н	I	J	K	L	M
1			Test 1	Test 2	Mid Term	Test 4	Test 5	Final	Avg Grade	Name			
2			80	78	84	92	82	83	83.17	Arthur			
3			72	100	64	84	78	88	81.00	Bea			
4			68	74	70	72	74	70	71.33	Cindy			
5			88	92	94	96	98	100	94.67	Harold			
6			100	100	96	96	100	98	98.33	Helen			
7			58	62	60	70	66	68	64.00	Joe			
8			98	98	98	100	98	96	98.00	Mabel			
9			74	84	86	92	76	78	81.67	Mary			
10			42	60	66	56	60	58	57.00	Ralph			
11			86	84	100	92	84	88	89.00	Sam			
12			42	60	66	56	60	58	57.00	Tim			
13					MTScores					Student N	ames		
14													
15	WideName	eList>	Arthur	Bea	Cindy	Harold	Helen	Joe	Mabel	Mary	Ralph	Sam	Tim
16													
17													
18			Mabel		=LOOKUP								
19				98	B =LOOKUP(\$C\$17,WideNameList,MTScores)								

If you examine various groups of cells on this sheet, you would find that each of the test scores containd in rows 2 through 12 of columns C through I have a unique name, as does the list of names in column J. For example, the scores in the Mid Term column have a name of MTScores, and the list of names out in column J is named StudentNames.

The list of names that run across the sheet in row 15 is named WideNameList.

Down in rows 18 and 19 I've shown two ways of doing lookups with the LOOKUP() function. In row 18, we see the formula for "looking left" from our source data to get a match: you type in a name in cell C18, it looks for the match in StudentNames (cells J2:J12) and returns the relative value from MTScores (cells E2:E12).

In row 19, the formula looks up data in a horizontal list and returns information from a vertical list! You type in a name in cell C18 and it looks up the match in WideNameList (cells C15:M15) and then using the index value determined, looks up the results to return from a vertical list, MTScores, in cells E2:E12 again.

This is potentially a very powerful tool. Heck, all the pieces don't even have to be on the same sheet in a workbook – our table of scores could have been on an entirely different sheet.

But there are some **pitfalls** to watch out for. It is easy for the lists to get 'out of sync'. If you go slip something in the middle of one list and don't adjust for that in places that use it as part of the LOOKUP() function and don't make an appropriate adjustment in either the **lookup_range** or **result_range** portions of LOOKUP() function setups, then your results are going to be WRONG!

Related Functions:

Some related functions to consider when you need to do 'lookups' within Excel:

INDEX() and **MATCH**() – can be combined to emulate **LOOKUP**() and allows you to use wildcards in your match and to perform case-sensitive exact matches. VLOOKUP(), HLOOKUP() and LOOKUP() are not normally case sensitive: dog=DOG=Dog=DoG, etc.

EXACT() needed for that case-sensitive lookup just mentioned.

CHOOSE() Chooses one value from a list of values.

OFFSET() Returns a reference offset from a given reference.

TRANSPOSE() Returns the transpose of an array (turns an array on its side).