

A quick demonstration of three
methods of handling similar, but not
identical spreadsheets using
Microsoft Excel

Anthony J. Castellani

Quick visualization

METHOD 1

Two spreadsheets side-by-side (or on different sheets or even books)

	A	B	C	D	E	F	G	H	I	J	K
1	Spreadsheet#	UniqueID	Attribute1	Attribute2	Attribute3		Spreadsheet#	UniqueID	Attribute1	Attribute2	Attribute3
2	S1	a	1	1	1		S2	a	1	1	1
3	S1	b	2	2	2		S2	b	2	2	2
4	S1	c	3	3	3		S2	c	3	3	3
5	S1	d	4	4	4		S2	d	4	4	4
6	S1	e	5	5	5		S2	g	7	7	7
7	S1	f	6	6	6		S2	h	8	8	8

Copy the two spreadsheets into one block, highlight a column of unique values, and use the “Duplicate Values” tool

The screenshot shows the Microsoft Excel interface. The 'Conditional Formatting' menu is open, and the 'Duplicate Values...' option is selected. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	Spreadsheet#	UniqueID	Attribute1	Attribute2	Attribute3						
2	S1	a	1	1	1						
3	S1	b	2	2	2						
4	S1	c	3	3	3						
5	S1	d	4	4	4						
6	S1	e	5	5	5						
7	S1	f	6	6	6						
8	S2	a	1	1	1						
9	S2	b	2	2	2						
10	S2	c	3	3	3						
11	S2	d	4	4	4						
12	S2	g	7	7	7						
13	S2	h	8	8	8						

	A	B	C	D	E
1	Spreadsheet#	UniqueID	Attribute1	Attribute2	Attribute3
2	S1	a	1	1	1
3	S1	b	2	2	2
4	S1	c	3	3	3
5	S1	d	4	4	4
6	S1	e	5	5	5
7	S1	f	6	6	6
8	S2	a	1	1	1
9	S2	b	2	2	2
10	S2	c	3	3	3
11	S2	d	4	4	4
12	S2	g	7	7	7
13	S2	h	8	8	8

Pivot Tables

METHOD 2

With the two spreadsheets copied together, add a column containing the rows concatenated together. Use a symbol not otherwise found in the data to separate values. Select the “PivotTable” tool.

File

Home

Insert

Page Layout

Formulas

Data

Review

View

PivotTable

Table

Tables

Picture

Clip Art

Illustrations

Shapes

SmartArt

Screenshot

Illustrations

Column

Line

Pie

Bar

Charts

F13

fx

=CONCATENATE(B13,"|",C13,"|",D13,"|",E

	A	B	C	D	E	F
1	Spreadsheet#	UniqueID	Attribute1	Attribute2	Attribute3	Concatenated
2	S1	a	1	1	1	a 1 1 1
3	S1	b	2	2	2	b 2 2 2
4	S1	c	3	3	3	c 3 3 3
5	S1	d	4	4	4	d 4 4 4
6	S1	e	5	5	5	e 5 5 5
7	S1	f	6	6	6	f 6 6 6
8	S2	a	1	1	1	a 1 1 1
9	S2	b	2	2	2	b 2 2 2
10	S2	c	3	3	3	c 3 3 3
11	S2	d	4	4	4	d 4 4 4
12	S2	g	7	7	7	g 7 7 7
13	S2	h	8	8	8	h 8 8 8
14						

In the pivot table screen, arrange as seen here. Copy the contents of the pivot table.

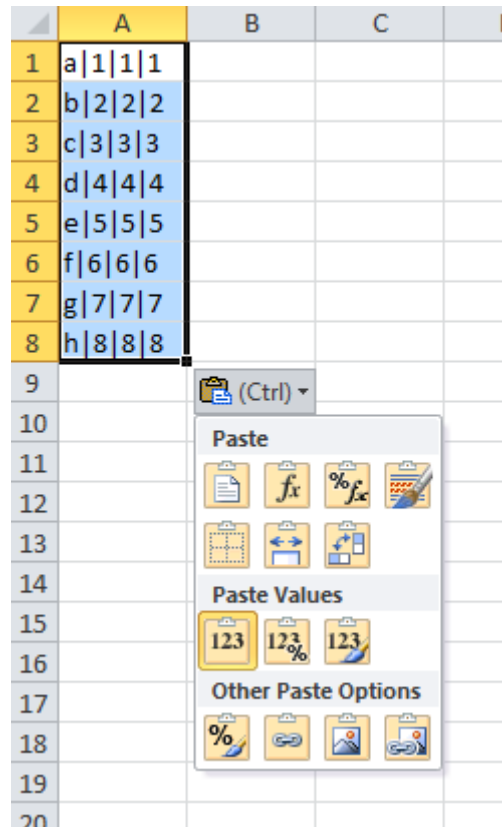
The screenshot shows an Excel PivotTable with the following data:

	A	B	C	D
1				
2				
3	Row Labels			
4	a 1 1 1			
5	b 2 2 2			
6	c 3 3 3			
7	d 4 4 4			
8	e 5 5 5			
9	f 6 6 6			
10	g 7 7 7			
11	h 8 8 8			
12	Grand Total			
13				
14				
15				
16				
17				
18				
19				
20				
21				

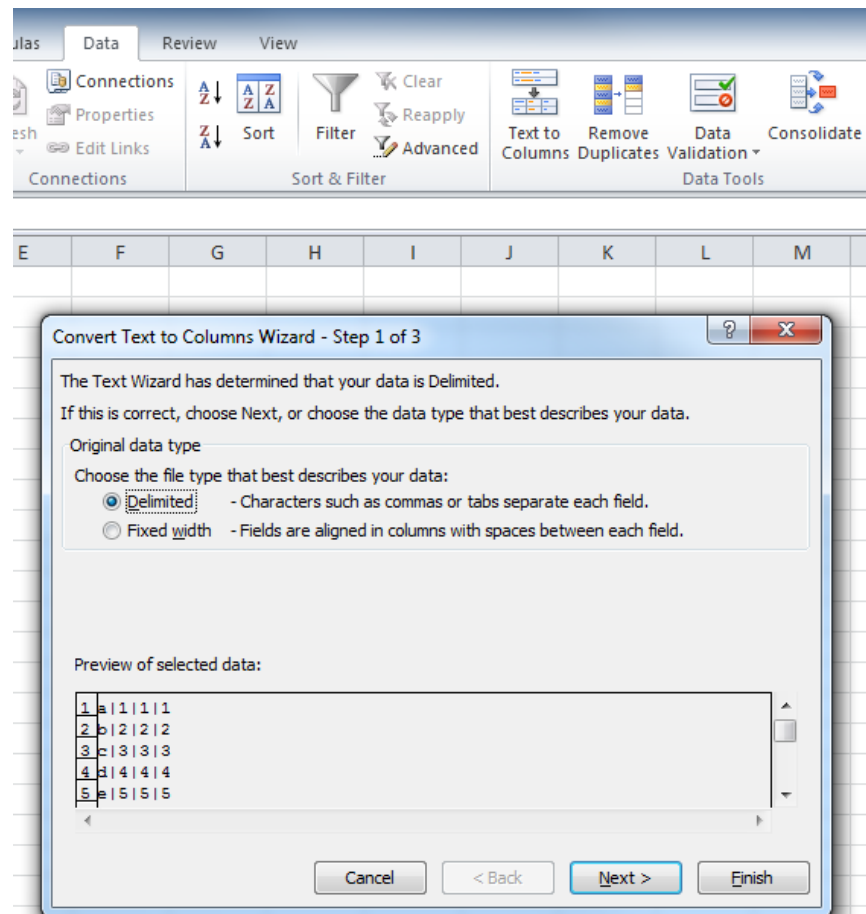
The PivotTable Field List task pane on the right shows the following configuration:

- Choose fields to add to report:
 - ☐ Spreadsheet#
 - ☐ UniqueID
 - ☐ Attribute1
 - ☐ Attribute2
 - ☐ Attribute3
 - ☒ Concatenated
- Drag fields between areas below:
 - Report Filter: (empty)
 - Column Labels: (empty)
 - Row Labels: Concatenated
 - Values: (empty)

In the location of your choosing, use the “Paste Values” function to paste your pivot table less the formulas.



With your data still highlighted, use the “Text to Columns” tool.
Follow the wizard steps seen on the next slides.



Wizard continued

Convert Text to Columns Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

☐ Tab

☐ Semicolon

☐ Comma

☐ Space

☒ Other: ||

☐ Treat consecutive delimiters as one

Text qualifier: " " ▾

Data preview

a	1	1	1
b	2	2	2
c	3	3	3
d	4	4	4
e	5	5	5

Cancel < Back Next > Finish

Wizard continued

Convert Text to Columns Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

Column data format

☒ General
☐ Text
☐ Date: MDY
☐ Do not import column (skip)

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Advanced...

Destination: \$A\$1

Data preview

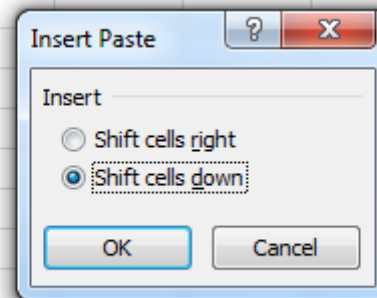
	General	General	General	General
a	1	1	1	1
b	2	2	2	2
c	3	3	3	3
d	4	4	4	4
e	5	5	5	5

Cancel < Back Next > Finish

You can now copy-and-paste your column headers above your new spreadsheet containing only unique values.

	A	B	C	D
1	a	1	1	1
2	b	2	2	2
3	c	3	3	3
4	d	4	4	4
5	e	5	5	5
6	f	6	6	6
7	g	7	7	7
8	h	8	8	8
9				

	A	B	C	D
1	a	1	1	1
2	b	2	2	2
3	c	3	3	3
4	d	4	4	4
5	e	5	5	5
6	f	6	6	6
7	g	7	7	7
8	h	8	8	8



	A	B	C	D
1	UniqueID	Attribute1	Attribute2	Attribute3
2	a	1	1	1
3	b	2	2	2
4	c	3	3	3
5	d	4	4	4
6	e	5	5	5
7	f	6	6	6
8	g	7	7	7
9	h	8	8	8
10				

VLOOKUP

METHOD 3

For my demonstration, my uniqueID field needed to be in Text format.

The screenshot shows the Microsoft Excel interface. The ribbon at the top includes File, Home, Insert, Page Layout, Formulas, Data, Review, and View. The Home ribbon is active, showing the Clipboard, Font, and Alignment groups. The Font group shows Calibri, size 11, and various formatting options. The Alignment group shows options like Wrap Text, Merge & Center, and text alignment. The spreadsheet has two tables. The first table (rows 1-7) has columns UniqueID, Attribute1, Attribute2, Attribute3, and Spreadsheet#. The second table (rows 10-16) has the same columns. The UniqueID field in the second table is highlighted, and the 'Text' format is selected in the Number Format dropdown menu.

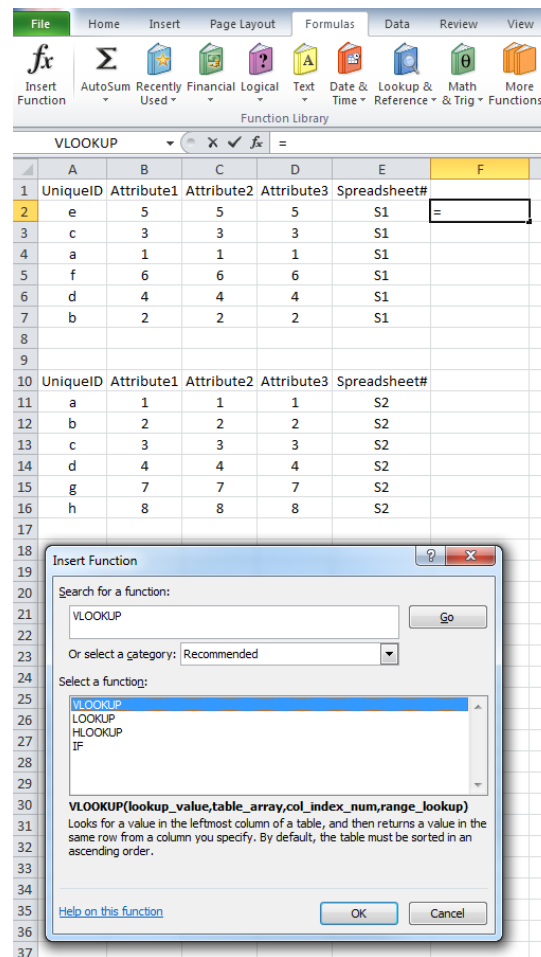
	A	B	C	D	E	F	G	H
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
2	e	5	5	5	S1			
3	c	3	3	3	S1			
4	a	1	1	1	S1			
5	f	6	6	6	S1			
6	d	4	4	4	S1			
7	b	2	2	2	S1			
8								
9								
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
11	a	1	1	1	S2			
12	b	2	2	2	S2			
13	c	3	3	3	S2			
14	d	4	4	4	S2			
15	g	7	7	7	S2			
16	h	8	8	8	S2			
17								
18								
19								
20								
21								

Number Format dropdown menu options:

- General: No specific format
- Number: a
- Currency: a
- Accounting: a
- Short Date: a
- Long Date: a
- Time: a
- Percentage: a
- Fraction: a
- Scientific: a
- Text: a**

More Number Formats...

You want two spreadsheets: the one you will add to, and the one you will reference from. The reference spreadsheet needs to be sorted in ascending order. Select an empty cell next to the first row and find the VLOOKUP function to deploy its wizard.



Select the first column of the spreadsheet that will be added to. Note the use of “\$”.

VLOOKUP X ✓ fx =VLOOKUP(\$A\$2:\$A\$7)

	A	B	C	D	E	F	G	H
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
2	e	5	5	5	S1	\$A\$2:\$A\$7		
3	c	3	3	3	S1			
4	a	1	1	1	S1			
5	f	6	6	6	S1			
6	d	4	4	4	S1			
7	b	2	2	2	S1			
8								
9								
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
11	a	1	1	1	S2			
12	b	2	2	2	S2			
13	c	3	3	3	S2			
14	d	4	4	4	S2			
15	g	7	7	7	S2			
16	h	8	8	8	S2			
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								

Function Arguments

VLOOKUP

Lookup_value: \$A\$2:\$A\$7 = "e"

Table_array: = number

Col_index_num: = number

Range_lookup: = logical

=

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Lookup_value is the value to be found in the first column of the table, and can be a value, a reference, or a text string.

Formula result =

[Help on this function](#) OK Cancel

Select all of the data in the reference spreadsheet. Note the use of “\$”.

VLOOKUP X ✓ fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16)

	A	B	C	D	E	F	G	H
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
2	e	5	5	5	S1	.S11:\$E\$16		
3	c	3	3	3	S1			
4	a	1	1	1	S1			
5	f	6	6	6	S1			
6	d	4	4	4	S1			
7	b	2	2	2	S1			
8								
9								
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
11	a	1	1	1	S2			
12	b	2	2	2	S2			
13	c	3	3	3	S2			
14	d	4	4	4	S2			
15	g	7	7	7	S2			
16	h	8	8	8	S2			
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								

Function Arguments

VLOOKUP

Lookup_value \$A\$2:\$A\$7 = "e"

Table_array \$A\$11:\$E\$16 = {"a",1,1,1,"S2";"b",2,2,2,"S2";"c",3,3,3

Col_index_num = number

Range_lookup = logical

=

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Table_array is a table of text, numbers, or logical values, in which data is retrieved. Table_array can be a reference to a range or a range name.

Formula result =

[Help on this function](#) OK Cancel

Input a 1 for now.

VLOOKUP X ✓ fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1)

	A	B	C	D	E	F	G	H
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
2	e	5	5	5	S1	11:\$E\$16,1)		
3	c	3	3	3	S1			
4	a	1	1	1	S1			
5	f	6	6	6	S1			
6	d	4	4	4	S1			
7	b	2	2	2	S1			
8								
9								
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
11	a	1	1	1	S2			
12	b	2	2	2	S2			
13	c	3	3	3	S2			
14	d	4	4	4	S2			
15	g	7	7	7	S2			
16	h	8	8	8	S2			
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								

Function Arguments

VLOOKUP

Look_up_value \$A\$2:\$A\$7 = "e"

Table_array \$A\$11:\$E\$16 = {"a",1,1,1,"S2";"b",2,2,2,"S2";"c",3,3,3

Col_index_num 1 = 1

Range_lookup = logical

= "d"

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Col_index_num is the column number in table_array from which the matching value should be returned. The first column of values in the table is column 1.

Formula result = d

[Help on this function](#) OK Cancel

“FALSE” is a common choice here.

VLOOKUP X ✓ fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,FALSE)

	A	B	C	D	E	F	G	H
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
2	e	5	5	5	S1	VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,FALSE)		
3	c	3	3	3	S1			
4	a	1	1	1	S1			
5	f	6	6	6	S1			
6	d	4	4	4	S1			
7	b	2	2	2	S1			
8								
9								
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#			
11	a	1	1	1	S2			
12	b	2	2	2	S2			
13	c	3	3	3	S2			
14	d	4	4	4	S2			
15	g	7	7	7	S2			
16	h	8	8	8	S2			
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								

Function Arguments

VLOOKUP

Lookup_value \$A\$2:\$A\$7 = "e"

Table_array \$A\$11:\$E\$16 = {"a",1,1,1,"S2";"b",2,2,2,"S2";"c",3,3,3

Col_index_num 1 = 1

Range_lookup FALSE = FALSE

= VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,F

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Range_lookup is a logical value: to find the closest match in the first column (sorted in ascending order) = TRUE or omitted; find an exact match = FALSE.

Formula result = VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,FALSE)

[Help on this function](#) OK Cancel

Your first result. Do not fear “#N/A”, as in this case the first row is unique, having no match in the reference spreadsheet.

F2		fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,FALSE)				
	A	B	C	D	E	F
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#	
2	e	5	5	5	S1	#N/A
3	c	3	3	3	S1	
4	a	1	1	1	S1	
5	f	6	6	6	S1	
6	d	4	4	4	S1	
7	b	2	2	2	S1	
8						
9						
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#	
11	a	1	1	1	S2	
12	b	2	2	2	S2	
13	c	3	3	3	S2	
14	d	4	4	4	S2	
15	g	7	7	7	S2	
16	h	8	8	8	S2	
17						

Because of the use of “\$”, you can now just drag down the first formula to the rest of the cells in the first column.

F2		fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,1,FALSE)				
	A	B	C	D	E	F
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#	
2	e	5	5	5	S1	#N/A
3	c	3	3	3	S1	c
4	a	1	1	1	S1	a
5	f	6	6	6	S1	#N/A
6	d	4	4	4	S1	d
7	b	2	2	2	S1	b
8						
9						
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#	
11	a	1	1	1	S2	
12	b	2	2	2	S2	
13	c	3	3	3	S2	
14	d	4	4	4	S2	
15	g	7	7	7	S2	
16	h	8	8	8	S2	
17						

To now move to the second column, copy over the formula, and change the “1” to “2”.

G2		fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,2,FALSE)					
	A	B	C	D	E	F	G
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#		
2	e	5	5	5	S1	#N/A	#N/A
3	c	3	3	3	S1	c	
4	a	1	1	1	S1	a	
5	f	6	6	6	S1	#N/A	
6	d	4	4	4	S1	d	
7	b	2	2	2	S1	b	
8							
9							
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#		
11	a	1	1	1	S2		
12	b	2	2	2	S2		
13	c	3	3	3	S2		
14	d	4	4	4	S2		
15	g	7	7	7	S2		
16	h	8	8	8	S2		
17							

You can now drag the formula down the rest of column 2.

G2		fx =VLOOKUP(\$A\$2:\$A\$7,\$A\$11:\$E\$16,2,FALSE)					
	A	B	C	D	E	F	G
1	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#		
2	e	5	5	5	S1	#N/A	#N/A
3	c	3	3	3	S1	c	3
4	a	1	1	1	S1	a	1
5	f	6	6	6	S1	#N/A	#N/A
6	d	4	4	4	S1	d	4
7	b	2	2	2	S1	b	2
8							
9							
10	UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#		
11	a	1	1	1	S2		
12	b	2	2	2	S2		
13	c	3	3	3	S2		
14	d	4	4	4	S2		
15	g	7	7	7	S2		
16	h	8	8	8	S2		
17							

Repeat this procedure for the remainder of the spreadsheet. It is now clear what rows are shared, and what rows are unique.

[illegible]

A quick sorting action makes this even clearer...

The screenshot displays the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Sort & Filter', 'Filter', 'Advanced', 'Text to Columns', 'Remove Duplicates', and 'Data Validation'. The active worksheet contains a table with the following data:

UniqueID	Attribute1	Attribute2	Attribute3	Spreadsheet#
e	5	5	5	S1
c	3	3	3	S1
a	1	1	1	S1
f	6	6	6	S1
d	4	4	4	S1
b	2	2	2	S1

A 'Sort' dialog box is open, showing the following settings:

- Column: Column F
- Sort On: Values
- Order: A to Z
- My data has headers: ☐

The dialog box also includes buttons for 'Add Level', 'Delete Level', 'Copy Level', 'Options...', 'OK', and 'Cancel'.

... crystal clear.

[illegible]