TechRepublic : A ZDNet Tech Community

How to find duplicates in Excel

Date: January 12th, 2009
Author: Susan Harkins
Category: Microsoft Excel
Tags: Column, Cell, Microsoft Excel, Microsoft Office, Office Suites, Software, Susan Harkins

In the duplicate world, definition means everything. That's because a duplicate is subjective to the context of its related data. Duplicates can occur within a single column, across multiple columns, or complete records. There's no one feature or technique that will find duplicates in every case.

To find duplicate records, use Excel's easy-to-use Filter feature as follows:

> Select any cell inside the recordset.
> From the Data menu, choose Filter and then select Advanced Filter to open the Advanced Filter dialog box.
> Select Copy To Another Location in the Action section.
> Enter a copy range in the Copy To control.
> Check Unique Records Only and click OK.

Excel will copy a filtered list of unique records to the range you specified in Copy To. At this point, you can replace the original recordset with the filtered list (the copied list) if you want to delete the duplicates.

	A	B	C	D	E	F	G	H	I	J
1	First Name	Last Name	Date							
2	Susan	Harkins	1/8/2009							
3	William	Harkins	1/8/2009							
4	Susan	Harkins	1/9/2009				First Name	Last Name	Date	
5	Susan	Harkins	1/8/2009				Susan	Harkins	1/8/2009	
6							William	Harkins	1/8/2009	
7							Susan	Harkins	1/9/2009	
8										
9										

Finding duplicates in a single column or across multiple columns is a bit more difficult. Use conditional formatting to highlight duplicates in a single column as follows:
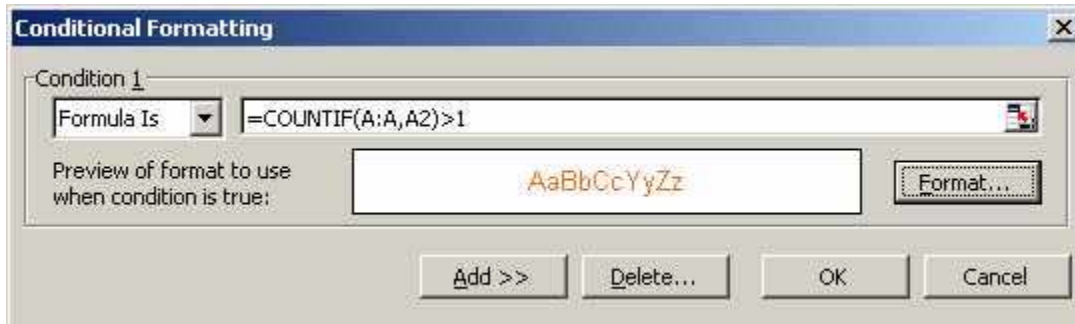
Using the example worksheet, select cell A2. When applying this to your own worksheet, select the first data cell in the list (column).

Choose Conditional Formatting from the Format menu.

Choose Formula Is from the first control's drop-down list.

In the formula control, enter $=COUNTIF(A:A,A2)>1$.

Click the Format button and specify the appropriate format. For instance, click the Font tab and choose Red from the Color control and click OK. At this point, the Conditional Formatting dialog box should resemble the following figure:

Click OK to return to the worksheet.

With cell A2 still selected, click Format Painter.

Select the remaining cells in the list (cells A3:A5 in the example worksheet).

The conditional format will highlight any value in column A that's repeated. If you want Excel to highlight only the copies, leaving the first occurrence of the value unaltered, enter the formula $=COUNTIF(\$A\$2:\$A2, A2)>1$ in step 4.

The conditional format works great for a single column. To find duplicates across multiple columns, use two expressions: One to concatenate the columns you're comparing; a second to count the duplicates. For example, if you wanted to find duplicates of both first and last names in the example worksheet, you'd enter the following formula in cell D2 to concatenate the first and last name values:

=A2&B2

You could insert a space character between the two names if you liked, but it isn't necessary. Copy the formula to accommodate the remaining list items.

	D2		fx =A2&B2			
	A	B	C	D	E	F
1	First Name	Last Name	Date			
2	Susan	Harkins	1/8/2009	SusanHarkins		
3	William	Harkins	1/8/2009	WilliamHarkins		
4	Susan	Harkins	1/9/2009	SusanHarkins		
5	Susan	Harkins	1/8/2009	SusanHarkins		
6						
7						

Next, in cell E2 enter the following formula and copy it to accommodate the remaining list:

=IF(COUNTIF(D\$2:D\$7,D2)>1,"Duplicated","")

	E2		fx =IF(COUNTIF(D\$2:D\$7,D2)>1,"Duplicated","")				
	A	B	C	D	E	F	G
1	First Name	Last Name	Date				
2	Susan	Harkins	1/8/2009	SusanHarkins	Duplicated		
3	William	Harkins	1/8/2009	WilliamHarkins			
4	Susan	Harkins	1/9/2009	SusanHarkins	Duplicated		
5	Susan	Harkins	1/8/2009	SusanHarkins	Duplicated		
6	Susan	Smith	1/9/2009	SusanSmith			
7							
8							

Notice that the worksheet has a new record (row 6). This record duplicates the first name, Susan, but not the last name. The conditional format highlights the first name because it's a duplicate in column A. However, the formula in column E doesn't identify the combined values across columns A and B as a duplicate because the first and last names together aren't duplicated.

People who read this, also read...

Mark duplicate entries automatically with Excel's COUNTIF function
Use Word to type on a printed form without a typewriter
Clean up your data with this easy-to-use Excel function
Quickly find matching records in Excel with a few keystrokes
Entering leading zeroes in Excel

Print/View all posts
Comments on this blog

An easier way greg.williams@... | 01/12/09

Duber is not free. slange@... | 01/13/09

Free for personal use.... gordon.cooper@... | 01/14/09

ASAP Utilities Excel.Instructor | 01/14/09

I've never been a fan of the add-ins. mattohare@... | 12/23/09

ASAP and Spreadsheet Assistant janly@... | 12/23/09

Another way... mattohare@... | 12/23/09

Another way = simple bkfriesen | 12/23/09

Do the same with a slight variance damiross@... | 12/23/09

I've used blanks in the past... mattohare@... | 12/24/09

Another way... dkidd23@... | 12/23/09

RE: How to find duplicates in Excel naoufel.mami@... | 12/23/09

MS-Access is nice, but not always available mattohare@... | 12/23/09

Another Easy way Clanphier | 12/23/09

Another Easy Way lkeppel@... | 12/23/09

Yet another way (Excel 2007 only) kkroon_ftb | 12/23/09

My Updates
My Contacts

Popular on CBS sites: SEC Football | NFL | Video Game Cheats | iPhone | Video Game Reviews | Notebooks | Antivirus Software

About CBS Interactive | Jobs | Advertise | Mobile | Site Map

© 2010 CBS Interactive Inc. All rights reserved. | Privacy Policy (updated) | Terms of Use