



Support / Excel

# Locate duplicates by using conditional formatting

Applies to: Microsoft Office Excel 2003

---

## APPLIES TO

---

Microsoft Office Excel 2003



This article was adapted from MrExcel.com. Visit the [MrExcel.com](http://MrExcel.com) Web site for more tips and information.

You can locate duplicates in a range of data by using conditional formatting and the COUNTIF function. Here are the details on how to make that work.

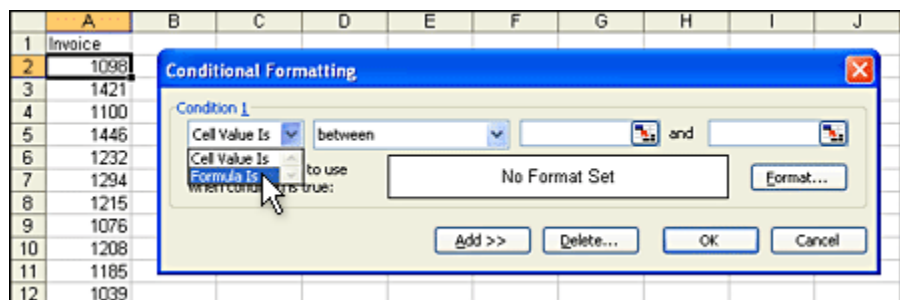
---

## Set up the first conditional formatting formula

---

I'll start by setting up a conditional format for the first data cell. Later, I'll copy that conditional format for the whole range.

In my example, cell A1 contains a column heading (Invoice), so I will select cell A2, and then click **Conditional Formatting** on the **Format** menu. The **Conditional Formatting** dialog box opens. The first box contains the text, **Cell Value Is**. If you click the arrow next to this box, you can choose **Formula Is**.



After you click **Formula Is**, the dialog box changes appearance. Instead of boxes for **between x and y**, there is now a single formula box. This formula box is incredibly powerful. You can use it to enter any formula that you can dream up, as long as that formula will evaluate to TRUE or FALSE.

In this case, we need to use a **COUNTIF** formula. The formula to type in the box is:

```
=COUNTIF(A:A,A2)>1
```

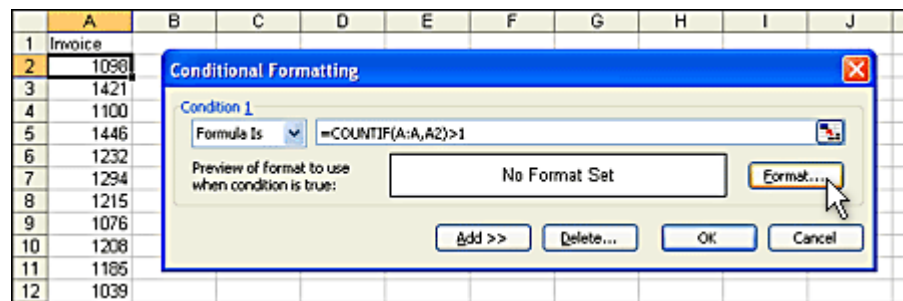
This formula says: Look through the entire range of column A. Count how many cells in that range have the same value as cell A2. Then, compare to see if that count is greater than 1.

When there are no duplicates, the count will always be 1; because cell A2 is in the range, we should find exactly one cell in column A that contains the same value as A2.

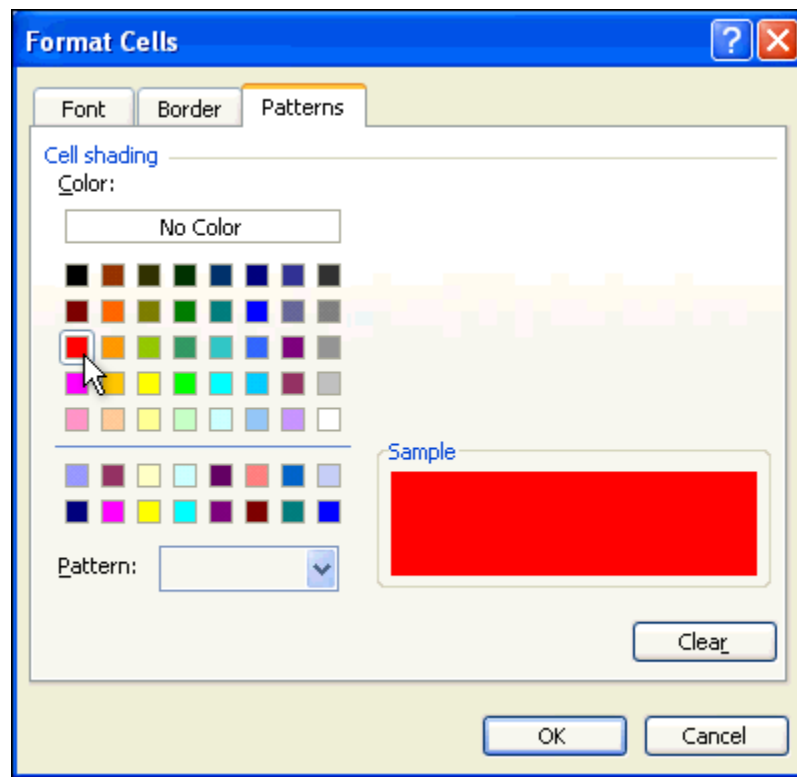
**NOTE** In this formula, A2 represents the current cell — that is, the cell for which you are setting up the conditional format. So, if your data is in column E and you are setting up the first conditional format in cell E5, the formula would be `=COUNTIF(E:E,E5)>1`.

## Choose a color to highlight duplicated entries

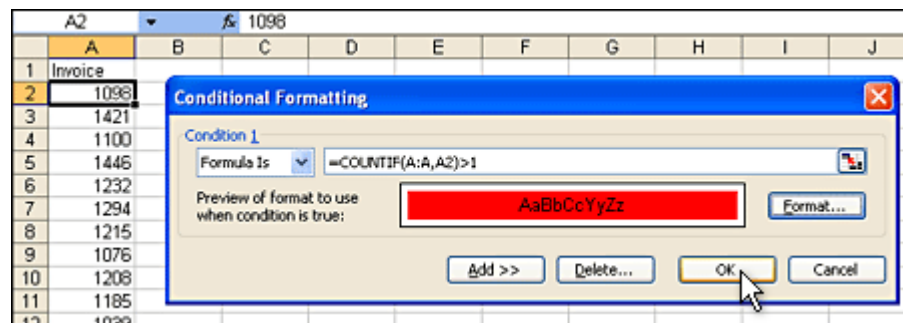
Now it is time to select an obnoxious (that is, obvious) format to identify any duplicates that are found. In the **Conditional Formatting** dialog box, click the **Format** button.



Click the **Patterns** tab and click a bright color swatch, like red or yellow. Then click **OK** to close the **Format Cells** dialog box.



You will see the selected format in the preview box. Click **OK** to close the **Conditional Formatting** dialog box, and...

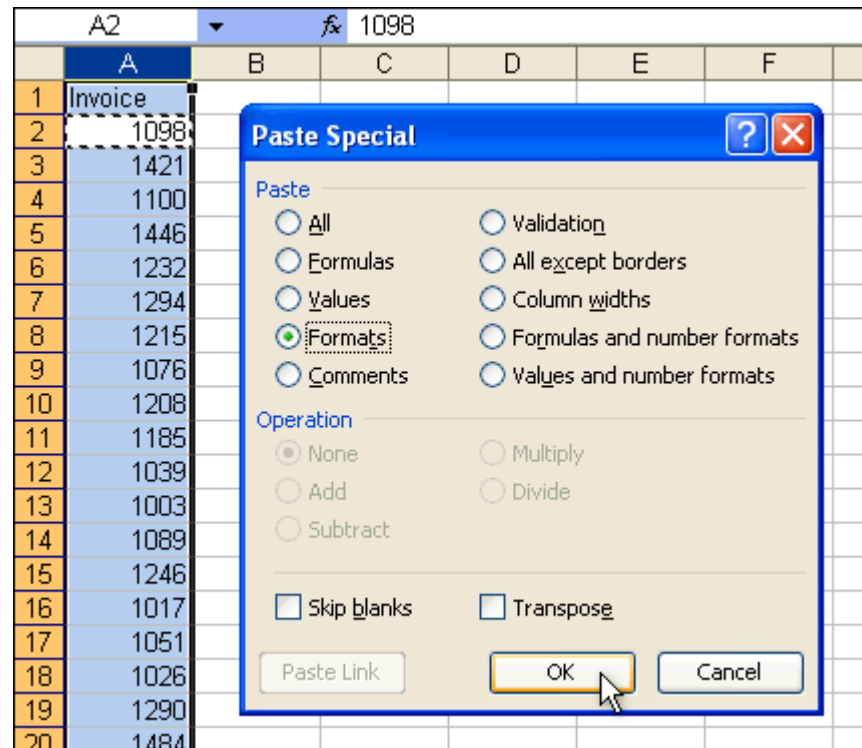


Nothing happens. Wow. If this is your first time setting up conditional formatting, it would be really nice to get some feedback here that it worked. But, unless you are lucky enough that the data in cell A2 is a duplicate of the data in some other cell, the condition is FALSE and no formatting is applied.

## Copy the conditional formatting to the rest of the cells

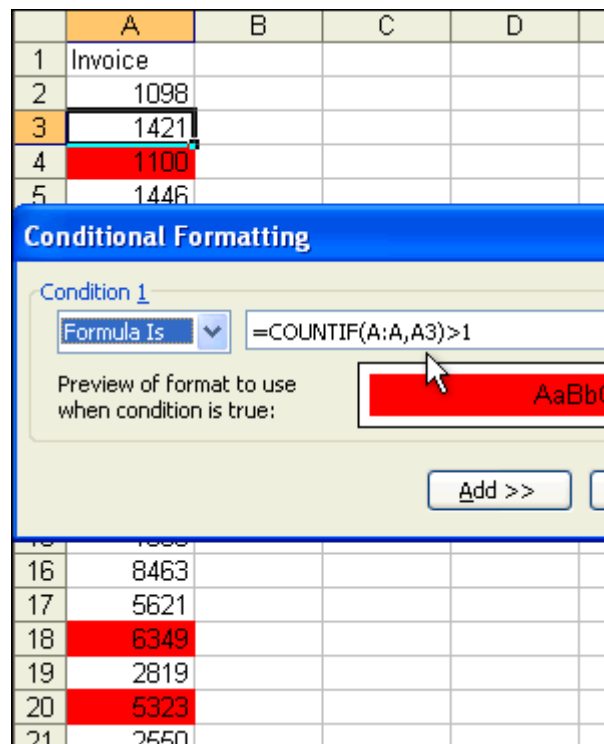
You need to copy the conditional formatting from cell A2 down to the other cells in your range. With the cursor still in A2, click **Copy** on the **Edit** menu. Press CTRL+Spacebar to select the entire column. Then click **Paste Special** on the **Edit**

menu. In the **Paste Special** dialog, click **Formats**, and then click **OK**.



This will copy the conditional formatting to all cells in the column. Now — finally — you may see some cells with the colored fill formatting, indicating that you have a duplicate.

It is informative to go to cell A3 and look at the conditional formula (after you've copied it from A2). Select cell A3 and click **Conditional Formatting** on the **Format** menu. The formula in the **Formula Is** box has changed to count how many times the data in cell A3 appears in column A.



You can have up to 65536 cells with conditional formatting, each cell comparing the current cell to 65535 other cells. Technically, the formula in the first step could have been `=COUNTIF($A$2:$A$1751,A2)>1`.

Also, when copying the conditional format to the entire column, you could have selected just the cells that contained data before using the **Paste Special** command.

## Highlight only the second instance of a duplicate

The previous solution assumes that you want to highlight both of the duplicate invoice numbers so that you can manually figure out which to delete or correct. If you don't want to mark the first occurrence of the duplicate, you can adjust the formula to:

```
=COUNTIF($A$2:$A2,A2)>1
```

**NOTE** It is important to enter the dollar signs exactly as shown.

In the first argument of this formula, only the second cell reference for the data range changes as it is copied down the column. This means the formula will compare only the cells from the current cell up to the first cell in the data range when looking for duplicate entries.

## Sorting the data

---

You really cannot sort a column on the basis of a conditional format. If you want to sort the data so that the duplicates are in one area, follow these steps:

First, type the heading **Duplicate** in cell B1. Then, type this formula in B2:

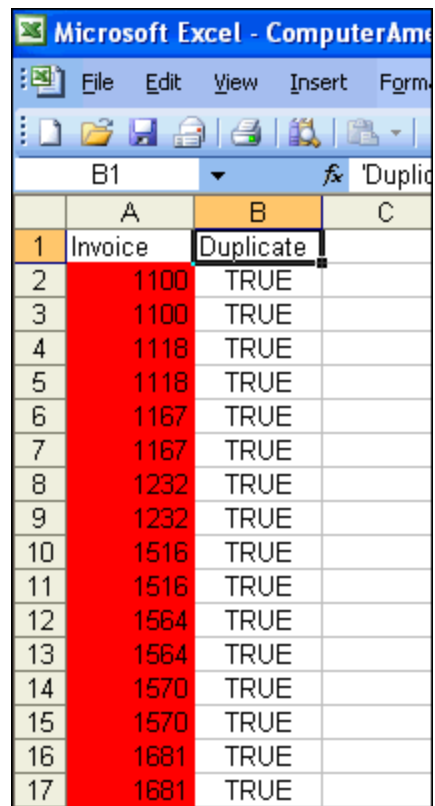
```
=COUNTIF(A:A,A2)>1
```

	A	B	C	D
1	Invoice	Duplicate		
2	1098	=COUNTIF(A:A,A2)>1		
3	1421			
4	1100			
5	1110			

With the cursor in cell B2, double-click the AutoFill handle (the little square in the lower-right corner of the cell) to copy the formula all the way down the column.

	B
	Duplicate
1098	FALSE
1110	

You can now sort the columns by column B (descending), then by column A (ascending), to show the duplicate invoice numbers at the top of the range.



	A	B	C
1	Invoice	Duplicate	
2	1100	TRUE	
3	1100	TRUE	
4	1118	TRUE	
5	1118	TRUE	
6	1167	TRUE	
7	1167	TRUE	
8	1232	TRUE	
9	1232	TRUE	
10	1516	TRUE	
11	1516	TRUE	
12	1564	TRUE	
13	1564	TRUE	
14	1570	TRUE	
15	1570	TRUE	
16	1681	TRUE	
17	1681	TRUE	

© 2010 Microsoft Corporation. All rights reserved.