



Remove Duplicates Quickly and Painlessly in Excel

By Nate Moore, CPA, MBA, CMPE

It is often helpful to review a unique list of patients, procedure codes, or diagnosis codes when analyzing data. However, you may sometimes find that you have duplicate information within the list you are reviewing. Fortunately, Excel makes it easy to quickly remove duplicates, and in this column, I will show you how to do it.

Removing Duplicates from a List

The easiest way to remove duplicate values is by using a list in which all data are in a single column.

For example, let us assume that Dr. Johnson has charges for the patients listed in Figure 1. Because his practice management system records ancillary charges separately from visits and includes voided charges, patients may be listed more than once. In this example, Colleen McCormick is shown twice, and Karl Crawford's name appears three times in the list.

To remove the duplicate patients from the list, select the entire list of names. (Note that including a column title in the selected list is optional.) With the list selected, choose "Remove Duplicates" from the Data tab, as shown in Figure 2. Your screen should now look like Figure 3. If you included a column title in your list, be sure the box "My data has headers" is checked. Otherwise, the box should be unchecked. Click "OK" and Excel will remove the duplicate patients. It will also show you the number of duplicate patients found and the number of unique patients remaining, as shown in Figure 4.

Removing Duplicates from a Table

It is almost as easy to remove duplicates from a table that has multiple columns of data.

Consider the table in Figure 5, which shows patients and evaluation and management (E&M) codes for a specific diagnosis code for the year. Notice that Audrey Barker is included four times in the table but with different E&M codes. To get a list of unique patients and E&M codes for the year, select both the patients and E&M code columns, and then choose "Remove Duplicates"

from the Data tab. Your screen should now look like Figure 6.

Now, compare Figure 6 to Figure 3. We only selected one column of data in the first example, so Figure 3 only includes one column. Since this example includes two columns, Excel shows patients and E&M codes in Figure 6. (Excel renames our E&M code column "EM Code" by removing the ampersand.)

Since our table includes multiple adjacent columns, we can check the box next to each column we want to evaluate for duplicate information. The "Select All" and "Unselect All" buttons make it easy to select or unselect all of the columns at once. Click "OK," and your screen should look like Figure 7, with a list of unique patients and E&M codes.

There are a variety of ways to use the "Remove Duplicates" function in Excel to get unique lists for a medical practice. Questions like "How many hypertension patients did we see last month?" or "What are all of the procedure codes we billed this year?" are easily answered by getting lists of the patients or codes and removing the duplicates. A table showing diabetic patients and foot exams could even answer this question: "How many of our diabetic patients had a foot exam last year?"

This function also helps you ensure you answer these questions accurately. If you don't remove duplicates from a table, a patient who had multiple foot exams in the same year, for example, would skew the results of your analysis, making it look like a higher percentage of patients received foot exams.

The next time you need a unique list or table, try the "Remove Duplicates" tool. It is fast and easy, even when using lots of data. ■

*Nate Moore, CPA, MBA, CMPE, writes custom SQL server code to mine practice management data for analysis in Excel, web pages, and via email. Nate's first book, **Better Data, Better Decisions: Using Business Intelligence in the Medical Practice**, written with Mona Reimers, was recently published by MGMA. His free Excel Videos have been viewed more than 700,000 times and are available at mooresolutionsinc.com. "Like" PivotTableGuy on Facebook and follow @PivotTableGuy on Twitter to be notified each time a new Excel Video is released.*



FIGURE 1

Today's Patients
Santos, Matthew
Chapman, Janis
McCormick, Colleen
Brown, Susie
Crawford, Karl
Kim, Stephanie
Walton, Chad
Gilbert, Margarita
Crawford, Karl
Bowen, Jill
Pierce, Belinda
Ruiz, Irving
Hayes, Kevin
Harrington, Rolando
McCormick, Colleen
Goodwin, Felix
Gilbert, Blanche
Osborne, Donnie
Robertson, Ruben
Crawford, Karl

FIGURE 3

Today's Patients
Santos, Matthew
Chapman, Janis
McCormick, Colleen
Brown, Susie
Crawford, Karl
Kim, Stephanie
Walton, Chad
Gilbert, Margarita
Crawford, Karl
Bowen, Jill
Pierce, Belinda
Ruiz, Irving
Hayes, Kevin
Harrington, Rolando
McCormick, Colleen
Goodwin, Felix
Gilbert, Blanche
Osborne, Donnie
Robertson, Ruben
Crawford, Karl

Remove Duplicates

To delete duplicate values, select one or more columns that contain duplicates.

☒ My data has headers

Columns

☒ Today's Patients

FIGURE 2

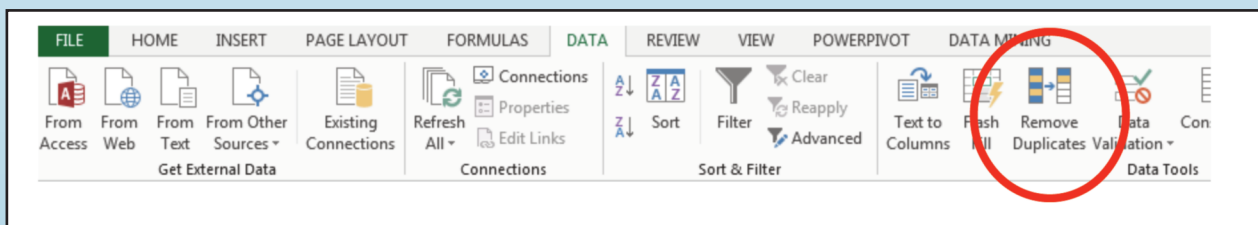




FIGURE 4

Today's Patients
Santos, Matthew
Chapman, Janis
McCormick, Colleen
Brown, Susie
Crawford, Karl
Kim, Stephanie
Walton, Chad
Gilbert, Margarita
Bowen, Jill
Pierce, Belinda
Ruiz, Irving
Hayes, Kevin
Harrington, Rolando
Goodwin, Felix
Gilbert, Blanche
Osborne, Donnie
Robertson, Ruben

Microsoft Excel

3 duplicate values found and removed; 17 unique values remain.

OK

FIGURE 5

Patients	E&M Code
Gilbert, Sean	99212
Barker, Audrey	99214
Hernandez, Joshua	99214
Mendez, Agnes	99215
Barker, Audrey	99213
Collins, Thelma	99214
Murray, Kristy	99214
Estrada, Geneva	99213
Riley, Ian	99214
Barker, Audrey	99213
Underwood, Yvonne	99213
Rivera, Irene	99213
Barker, Audrey	99212

FIGURE 6

Patients	E&M Code
Gilbert, Sean	99212
Barker, Audrey	99214
Hernandez, Joshua	99214
Mendez, Agnes	99215
Barker, Audrey	99213
Collins, Thelma	99214
Murray, Kristy	99214
Estrada, Geneva	99213
Riley, Ian	99214
Barker, Audrey	99213
Underwood, Yvonne	99213
Rivera, Irene	99213
Barker, Audrey	99212

Remove Duplicates

To delete duplicate values, select one or more columns that contain duplicates.

Select All Unselect All ☒ My data has headers

Columns

☒ Patients

☒ E&M Code

OK Cancel

FIGURE 7

Patients	E&M Code
Gilbert, Sean	99212
Barker, Audrey	99214
Hernandez, Joshua	99214
Mendez, Agnes	99215
Barker, Audrey	99213
Collins, Thelma	99214
Murray, Kristy	99214
Estrada, Geneva	99213
Riley, Ian	99214
Underwood, Yvonne	99213
Rivera, Irene	99213
Barker, Audrey	99212

Microsoft Excel

1 duplicate values found and removed; 12 unique values remain.

OK