The Excel Quick Fix

R. Jason Weiss
Development Dimensions International

In this issue, Leading Edge takes a step back from considering the future of technology and offers a quick survey of ways to use Microsoft Excel for I-O data management. The inspiration for this column is in the often-heard comment, “I wish I knew more about Excel. I know there’s a lot you can do with it.” It’s true: Excel is a powerful data-management machine that hides behind an inscrutable grid interface. Though this column makes no effort to replace the ten pound, thousand-page book entitled Excel Monster Compendium of Expertise that you will find at your local bookstore, it is my hope to show you a few tools you can use immediately and to positive effect.

Typographic Conventions

The following typographic conventions are used as an aid to comprehension:

Data and code in Excel are shown in a Courier font.

Some actions require you to select an option from a menu. The menu name and option are displayed separated by a pipe symbol: |

Example: To save a file under a different filename, select File | Save As… and enter the new filename into the ensuing dialog box.

Entering Data

Excel is frequently used in I-O as a data entry tool. There are several benefits to this approach. First, Excel files are easily imported for analysis by dedicated programs such as SPSS and SAS. Second, it is easy to catch gross errors, such as the entry of data into the wrong column. The following are several tools that will make data entry easier and more powerful.

Quick Data Entry Form

Excel offers a simple data entry and search form that provides some relief from entering data directly into cells. Though it’s not the prettiest, it’s simple to implement. Plus, you’ll find it invaluable if you ever need to find records that meet multiple criteria across variables.
To use the data entry and search form, you need to follow the convention of organizing your variables in the columns of the spreadsheet, and your records (or cases) in the rows, as shown in Figure 1.

To pull up the data entry and search form, first make sure to enter your variable labels in the first row, if they are not there already. Next, select **Data | Form**… What you will see is a simple form like that shown in Figure 2.

To enter a new record, click the **New** button. Navigate between variables by hitting the **Tab** key on your keyboard, or by holding the **Alt** key and typing the underlined letter for the variable shown on the form (e.g., type **Alt+A** to skip straight to the “Last Name” variable). It is very important to note that the data are not transferred from the form to your spreadsheet until you click the **New** button.

You can also use this form as a simple search tool by clicking the **Criteria** button. Enter the data you wish to find in one or more fields and click the **Find Prev** or **Find Next** buttons to find each instance of the search informa-
tion. To return to the data entry function of the form, click the **Form** button, which you’ll find in place of the **Criteria** button selected earlier.

**Data Validation**

Mis-entered data have plagued every psychologist at some point in time. Excel offers an easy-to-use tool for ensuring that data entered into a spreadsheet meet established criteria. To explore this tool, select a cell or range of cells and click on **Data | Validation…** The *Data Validation* window is shown in Figure 3 with the Validation Criteria listbox opened to show the range of options available.

![Data Validation window](image)

*Figure 3. Data Validation window.*

Each of the options allow you to define a discrete or continuous range of valid data. The “List” option allows you to include an in-cell, drop-down listbox from which the user can use to select from a set of valid choices. Note that the “List” option does not work in the data entry and search form described above.

As an aid to the person doing the data entry, you can create an input message that provides instructions and an error alert that warns of invalid data. Each of these messages can be accessed by clicking on its associated tab on the *Data Validation* window.

**Working With Data**

In my experience, one of the prime uses of Excel among I-O psychologists is as a tool for organizing and “cleaning up” data prior to shifting it to a dedicated data analysis software package. Here, too, Excel offers some read-
ily-accessible power tools to help the user quickly review, sift through, and manage data.

**Automatic Data Filtering**

Excel’s AutoFilter is one of its handiest tools. Simply put, the AutoFilter allows you to filter which rows are displayed by their match to selected variable values. Click on **Data | Filter | AutoFilter** to turn the AutoFilter on. You will see a slight change in the spreadsheet, as drop-down buttons are placed along each variable name in the first row. Figure 4 shows a portion of a spreadsheet window with the AutoFilter applied and the Gender drop-down button selected.

![Figure 4. Spreadsheet with AutoFilter active.](image)

As shown, AutoFilter automatically includes all variable values in its drop-down list. A quick look at the values in the drop-down list can be an easy way of finding miscoded or misspelled entries that would otherwise trip up your data analysis software. The main function of AutoFilter, however, is to allow you to easily filter the rows displayed. For example, selecting “Female” from the Gender drop-down list would filter out John Smith’s and John McGillicuddy’s records. You can create more tailored filters by choosing the “(Custom...)” option.

It is important to note that AutoFilter permits you to filter each variable separately. As such, if you wanted to filter for males named John, AutoFilter makes it easy. AutoFilter can be especially powerful for reviewing experimental data, as you can filter to concentrate on the participants in each group, separately. To reset the filter, click **Data | Filter | Show All.**

**Conditional Formatting**

Conditional formatting allows you to format cells based on a set of criteria. For example, I have a spreadsheet in which I am tracking the progress of different files to be generated and entered into a database. Each file can be in one of three stages: Not Started, In Progress, or Complete. I’ve ensured that these are the only three values that can be entered by using the data validation window, described above. The cells for files that are Not Started are automatically formatted with a red background. Cells for In Progress files
are yellow, and those that are *Complete* are green. A second column indicates whether the files have been entered into the database or not, and these values, too, are color-coded. My programming partners can then quickly scan for files where there is a mismatch of colors, and promptly enter those files into the system.

Like the other tools presented here, conditional formatting is easy to implement. First, select the cells for which you would like conditional formatting. Second, click on **Format | Conditional Formatting...** You will get the window shown in Figure 5.

![Conditional Formatting window](image)

*Figure 5. Conditional Formatting window*

Conditional formatting is a two-step process. First, you must specify the condition to satisfy for the formatting to be applied. The condition can be specified in one of two ways. First, as shown above, you can specify a condition based on the value of the selected cell. The second application is more advanced, and that is to specify a formula to evaluate in place of the cell value. If the result of that formula is true, the selected formatting is applied. The second step in the process is to designate the formatting to be applied based on this condition. This is done by clicking the **Format...** button, which brings up another window in which you can specify formats for the cell font, border, and shading. The system default for conditional formatting is a single condition. You can add up to two more by clicking the **Add >>** button at the bottom of the *Conditional Formatting* window.

One useful application I have found for conditional formatting is to shade alternate rows so that I can more effectively track data as I read across rows. To do this, select all of the rows to be formatted. Open the *Conditional Formatting* window and change *Condition 1* as shown in Figure 6. The formula shown tests whether the row number for the current row is odd. If so, a gray cell background is applied.
Tools That Make Life Easier

The tools described above offer a lot of power for working with and viewing your data. Presented below are some even more basic tools that make life with Excel much easier.

Comments

Excel permits you to tag cells with pop-up comments that appear when you mouse over that cell. To insert a comment, select your desired cell and click **Insert | Comment.** You will then be given a little yellow box in which to enter a note. I usually find it necessary to resize the note by dragging the little handles at the corner of the note window. When you have finished typing your note, click outside the note window to close it. Notes are indicated by a red triangle in the top, right corner of an annotated cell. I find notes particularly useful for keeping information about variables or data which I am unlikely to remember when I go back to look at the data some months (or even weeks!) later.

Printing Control

If you have a large worksheet and hit the **Print** button, you can find yourself under a dismaying avalanche of paper. Fortunately, Excel offers several handy ways to control what you print and how it appears. For example, I sometimes have both raw data and summary computations on a worksheet, and wish to print only the summary computations. To print only part of a worksheet, select the cells you wish to print and click **File | Print Area | Set Print Area.** Printed output from that worksheet will then be confined to the cells you selected.

It is also useful to control how Excel paginates the data. This is done by selecting **View | Page Break Preview.** You will then see your worksheet with a set of blue outlines. The solid blue outline is the print area you specified in the paragraph above. The dotted blue lines indicate page breaks. Drag these outlines so that pages break in logical places between rows and columns. Be aware, however, that Excel will automatically scale the result on your printout. The more columns and rows you include on a page, the
smaller the output will appear! To return to the standard view of the data, select View | Normal.

**Naming Ranges**

If you perform any data analysis with Excel, you’ll find it wearisome to track the cell addresses of your data. Though there’s nothing wrong with \( \text{SUM(G24:H32)} \), having names for ranges of cells makes it a lot easier to work with them. Simply select the cells you wish to name and select Insert | Name | Define… Type your desired name (without spaces) into the box, and you will now be able to use this name instead of cell addresses in formulas, e.g., \( \text{AVERAGE(ReactionTimes)} \). This simple process for naming ranges can make a night-and-day difference in terms of Excel’s ease of use.

To make it even easier, ranges can be named for you automatically. Select Tools | Options and click on the Calculation tab. Under Workbook options, make sure the Accept labels in formulas checkbox is checked and hit OK. Now you can use the variable labels in the first row and column of your spreadsheet as predefined range names. Using the example in Figure 4 above, I could include Gender and Birthday in my formulas without any further editing. Range names cannot contain spaces, however, so I would change cells A1 and B1 to FirstName and LastName, respectively.

**Useful Resources**

By now, I hope I’ve generated some excitement around the tools in Excel that can be quickly leveraged to valuable effect. If you are interested in learning more about Excel, there are many resources available. Here are a few recommendations:

**Books**

Your local bookstore has plenty of thousand-page behemoths that provide detailed surveys of Excel’s features. I recommend reading the reviews at Amazon.com and leafing through a few to find one that suits you. For straightforward how-to information, I like Peachpit Press’s approach. Their books focus on how to perform tasks and include little narrative fluff. Though I have not reviewed it, I suggest you take a look at their book on Excel, which tops out at a mere 280 pages. They have editions for many versions of Excel across both Windows and Macintosh platforms. The Mr. Excel book is also worth checking out.


When you start to itch for more power with Excel, I highly recommend the following:


**Web Sites**

The following Web sites are useful for help on Excel:

The ExcelTip site is hosted by the folks who wrote *Mr. Excel on Excel,* cited above. I have found it to be an excellent source of conveniently categorized tips. http://www.exceltip.com

John Walkenbach’s Web site has a lot of information on Excel, including a long list of links to other Excel sites. http://www.j-walk.com/ss/excel/

**Questions, Comments, Ideas?**

If you have any questions or comments on this edition of *Leading Edge,* or ideas for what you’d like to see in this column, please don’t hesitate to e-mail me at jason.weiss@ddiworld.com. I regret, of course, that I cannot offer specific user-to-user help on Excel.