EXCEL FILTERS AND PIVOT TABLES

Sarah Cohen, The Washington Post
 (Updated for Excel 2007 by Jaimi Dowdell)

Some reporters choose to use Excel as a little database. The reason is that it's simple and flexible for summarizing and filtering small data sets.

The two key tools are filters and pivot tables. Filters let you see just the items you want, without changing the underlying data. Pivot tables are powerful summary tools, much like statistical software for large databases. You'll use them to create crosstabs.

SETTING UP YOUR SPREADSHEET

There are two steps to get ready for a filter or pivot table. One you have to do. The other is optional, but will make your life easier.

Make sure it looks like a database

To make a filter or pivot table work, you need your data in columns all next to one another, and rows without blanks. You also need field or column names directly above the data in a single cell:

Make sure any totals, averages, notes or other words you placed at the bottom are separated by a blank row:

Name the database

You don't have to do this, but it makes later work easier.

Select the data, including the titles but excluding the totals at the bottom (click in your data and type ctrl+shift+8 to quickly make the selection). Right click on the selected data and choose "Name a range."
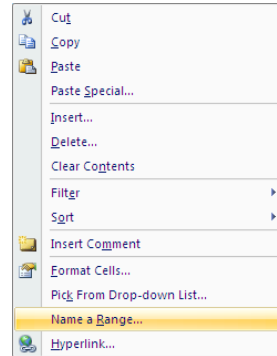
Now, type in a name to represent the data range. I usually call it "Database."

Now whenever you want to refer to that area, you can refer to the name Database instead of its address.

Try it. Select any cell on your worksheet. Now drop down the menu on the upper left corner of your worksheet (where the cell address usually appears) and pick Database:
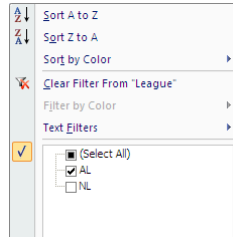

You'll be taken to your data.

FILTERING

Select your database using its name, and click on the Data tab. Under the Data ribbon, select Filter. It looks like a funnel. After you've clicked on it, drop-down

menus will appear next to each of the column headings:

Word Filtering

Use the drop-down menu to show only rows with certain words, like AL here:

Your row numbers will turn blue to give you a visual clue that you're filtering. The drop-down arrow turns blue next to League to show that it's the field that's filtered.

To return to all rows, drop down the filter again and click on the box next to (Select All).

Number Filtering

To choose numbers, you'll usually want to choose values greater than or less than a certain value.

Drop down your menu in the appropriate column and choose Number Filters. Then choose "is greater than" and type in the value:
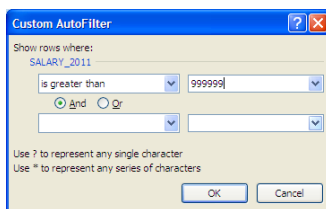
REMOVING A FILTER

Make sure to get rid of any filters before you sort or make any new calculations. To remove the filter, choose Data tab and click on Filter. Now all of the drop-downs will disappear.

SETTING UP A PIVOT TABLE

You have powerful summary tools with a Pivot Table. It's much like a Totals, Group By or Summary query in a database manager, but it allows even more flexibility.

To start a Pivot Table, select your data, including the headings. Then choose the Insert tab, then Pivot table and pivot table again.

You'll get a Pivot Table Wizard. Make sure Excel recognized your database. If there aren't marching ants going around your data, then type in Database if you named it.

Now make sure Excel recognized your database:

Type in Database if you named the range. You can also click in somewhere in your data and type ctrl+shift+8 to highlight the selection.
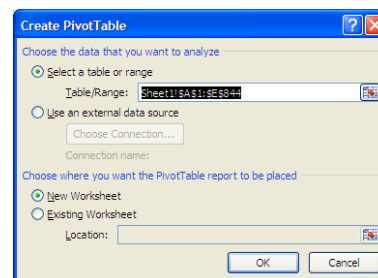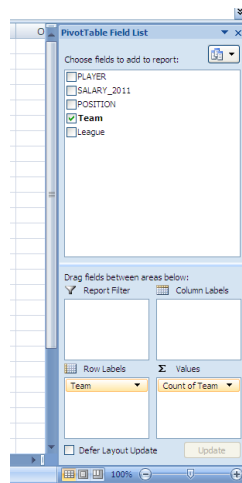

There's no reason why you should put your pivot table over your current work, so make sure "New Worksheet" is select. Then click "Ok."

This is the meat of your pivot table. You'll drag the buttons for each element you want to see from the right-hand side to the boxes.
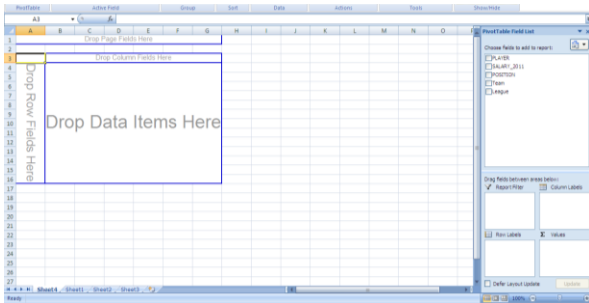
(When you drag a column that contains only numbers into the center, Data, area, Excel assumes you want to add up the values. Otherwise it assumes you want to count up the lines, or players per team in this case.)

EDITING A PIVOT TABLE

To change column or row headings

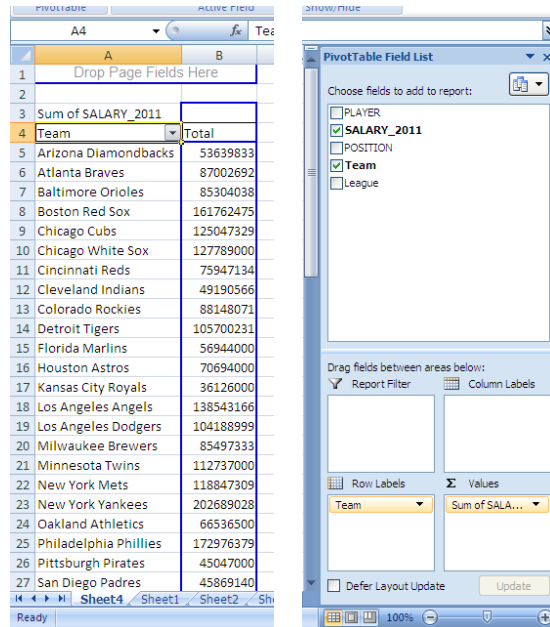Type right over what Excel gave you:

To add or delete elements

If the pivot table field list along the right disappears, simply click back on your pivot table along the left.

This brings you back to your choices for the table.

AUTOMATICALLY CALCULATING PERCENTAGES

You can add data elements more than once, and change how Excel presents them to you. For example, let's say you had a pivot table with total salary by team as shown in the following figure:
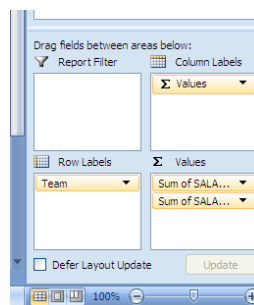

Drag another copy of the salary field into your data area by dragging it under "Values" in the lower right-hand corner of the field list box:
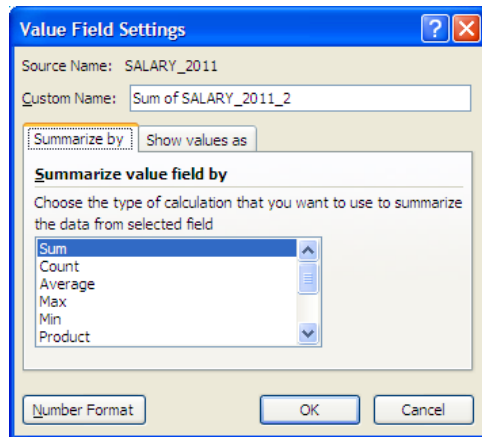
Then click on the drop-down arrow next to the "Sum of SALA…" you just added. Select, "Value field settings…" Under
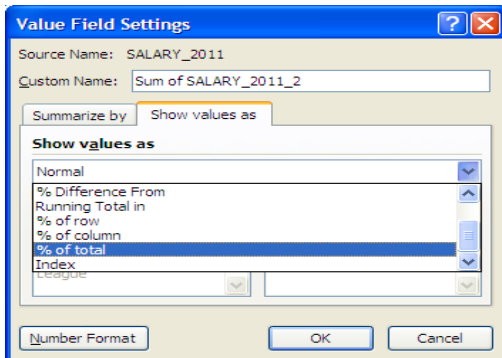

"Show Data As", click on the drop-down menu and choose % of Total:

Notice under "Summarize by" that you can change Sum to count, average, etc.

SORTING PIVOT TABLES

Newer versions of Excel have made this very simple. If you'd like to sort by numbers created in the pivot table, like the total salary per team, simply right click on one of the values and select sort then "smallest to largest" or "largest to smallest."
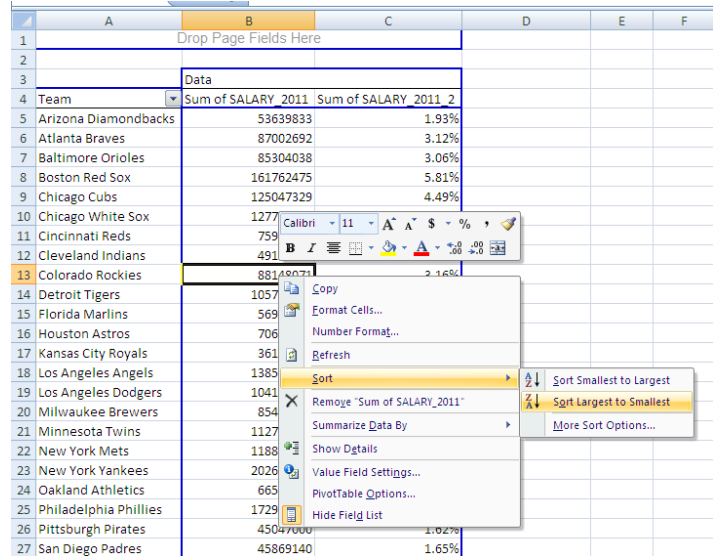
A WORD ON PIVOT TABLE SETUPS

Pivot tables are easy to read when they follow Philip Meyer's rule of thumb: Put the independent field as columns, the dependent field as rows, and calculate a column percent.

Translation: Put whatever comes first in time in the COLUMNS area, put whatever comes last in time in ROWS and change "Normal" to "% of Column" in the Show Data As area.

That way you can read it normally, from left to right.

You can use this for all kinds of stories: Do women with expensive insurance get c-sections more frequently than those covered under Medicaid or charity cases? Do upper-income minority homeowners get refused loans more frequently than similarly situated whites? How frequently do public officials use their cell phones during work hours?

TIPS, TRAPS AND FAQS

How do I get back to my original data?

At the bottom of your screen are little tabs, like index markers. Click on the appropriate one. If you haven't named it, choose Sheet 1.
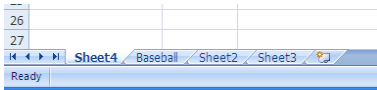
How do I get Pivot Tables to calculate medians?

You don't. It's a key element missing from both Pivot Tables and database management programs. There are ways to calculate medians for groups, but they're quite difficult.

How do I get a Pivot Table to update data I've changed in the original database?

Right click on the pivot table and select "refresh."

I get an error when I try to create the pivot table.
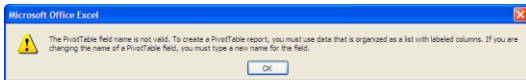
You probably have a missing heading. Check to make sure that the row above your data is filled in with titles.

Creating bigger categories:

Rather than using formulas, it may be simplest to create an "Other" category or other groupings through a pivot table.

Put the values you want together next to one another. Then select them, and choose the Group button on your query and pivot toolbar. You can also find "group" if you right click on the values.