Advanced Excel 2007: Dates, times and lookup tables

Updated by MaryJo Webster, St. Paul Pioneer Press (Feb. 2010)

You cover the county commission in Gabalot, R.I. You've been to a few NICAR conferences, and you are convinced that the cellular phone logs of commissioners offer a great resource for an enterprising reporter.

So you get copies of four commissioners' cellular phone bills and type them into a spreadsheet.

· Open the worksheet called CarPhone2007.xls.
What you see is a spreadsheet of four Gabalot commissioners' cellular phone bills. Take a look at the structure. It looks like a database, with fields (columns) and records (rows) of each phone call made.

Zoom in on the area that includes the date and time information. Note there's a column for the day of the month, the month, the year, and the time. Separating them out was wise so that the data has more flexibility.

So what are you going to do with all of these phone calls?

If you only wanted to know which telephone numbers are called the most by each commissioner, you probably wouldn't use a spreadsheet at all. Database managers like Access and FoxPro make easy work of counting calls and adding minutes.

But you're looking for something different. How often do commissioners bill taxpayers for personal use of their cellular phones?

One way to examine that question is to examine the individual phone numbers. You'd look for out-of-state calls and calls to known friends or family members. Another way is to look at the timing of the calls.

Of course, knowing whether the commissioners used their phones during normal work hours won't tell you for sure whether the calls were for personal or professional work. Public officials often work well past closing time, on weekends or very early in the morning. But you can get a feel for the patterns by determining what day of the week and what time of day the calls are made. At least you'll be able to ask the commissioners about late night and weekend phone use in your next interview.

This exercise will show you some powerful functions of spreadsheets that can summarize date and time values. You'll use date functions, a lookup table and a pivot table to summarize the cell phone use in this spreadsheet.

Understanding dates and times

Before you start, you have to understand how Excel works with date and time values.

Virtually all personal computer applications treat dates and times in a special way. They display a date or a time that looks like one you recognize, but they are actually storing something very different. In Excel, a date is the number of days since Jan. 1, 1900. Time works the same way. Excel stores a fraction, or decimal representation, of a day. Here's how a spreadsheet sees different values:

What you see:	What a spreadsheet sees:
January 23, 2005	38375
12:05 PM	0.503472
23-Jan-2005 12:05 PM	38375.503472

This is a powerful way to think of time for three reasons:

1. No matter how you look at your dates and times, they can be sorted in chronological order. Anyone who's tried to sort a text representation of a date, like "2/1/2005," knows how powerful that can be.
2. Excel can keep information about the date. For instance, it knows whether the time is in the afternoon or the morning. It knows that a date is the 25th of December, no matter what year it is. And it knows that December 25, 2005 was a Sunday. It stores the day of the week as a number, too: Sunday is 1, Monday is 2, etc.
3. You can do math on the results. This means that you can add a date to a time. Once you do that, you can find out the number of minutes, days, weeks or months that separate the values. So your spreadsheet will be able to calculate the number of minutes it took for an ambulance to respond to a call, even when the call was made at 11:55 on Saturday night and the response took 20 minutes.

Using number formats

You've probably used number formats in Excel before. When you calculated a percent change, for example, Excel displayed the results as a fraction. To get it to appear as a percentage, you had to go to Format Cells to change it to a percentage.

You can do the same thing with dates and times. In Excel, you format a number in the Format, Cells, Number dialog box. Special codes tell Excel how to display the date. Here are some ways that Excel can display the value 38375.503472:

Date format
What you see

Yyyy	**1995**
mmmm d, yyyy	January 23, 2005
mm/dd/yy	01/23/2005
m/d/yy	1/23/2005
dddd, mmm d, yyyy hh:mm AM/PM	Tuesday, Jan 23, 2005 12:05 PM

What day is it?

In our cellular phone log, you have to create one of these special date values before you can work with information like the day of the week. You'll use the DATEVALUE() worksheet function to figure it out.

The DATEVALUE() function works on a string of text that looks like a date. So the function command =DATEVALUE("23-Jan-2005") would return the value 37994.

• Select the entire column E. Insert a blank column there. Once it's inserted, select cell E1, type "Date" as the column header and hit enter.

Now there's a problem. Your spreadsheet has the day (8), month (Jan) and year (2004) in separate cells. We'll put them together using text operators, separated by dashes, in our formula.

• Move your cursor to cell E2 and select it. Type =datevalue(b2&"-"&c2&"-"&d2) and hit enter.

What you've done is create the text string "8-Jan-94" inside the parentheses. Excel can then recognize that string as a date, and has converted it into the number of days since Jan. 1, 1900.
(Don't worry if you get a number slightly different than this. There are two date systems. One counts the number of days since 1904, and someone might have put that option onto your computer. As long as you get a number close to this, you've probably done it correctly.)

You'd like to see this number as a date anyway.

• Select column E. On the Home menu ribbon, use the pull-down menu to change the format to "Short Date"

In this example, 37994 will become 1/8/2004, a format that saves space while giving you all the information you need about your date.

• Copy your new date down through the rest of the rows. You now have a new field that is made up entirely of dates that Excel understands as dates.
We're interested in finding out whether our commissioners use their cell phones on the weekend. Our first step is to figure out what day of the week each call was made.

• Select column F and insert another column. Type "Weekday" as the column header in cell F1 and hit enter. Now move your cursor to cell F2, select it and type =weekday(e2) and hit enter.

What happened? When you inserted a new column next to Column E, the spreadsheet assumed you wanted it formatted the same way. It's trying to figure out a date that corresponds to 7, which is the day of the week that was calculated. Again, all you need is a little formatting help.

• Select column F, go to Format Cells (either right-mouse click or under the Home ribbon). Choose Custom under Category and type ddd (for 3-letter days of the week) into the Type box provided:
Now you should see "Thurs" in the weekday column, which is the 5th day of the week.

• Select F2 and copy it down through your spreadsheet.

What are your hours?

Commissioners often work long hours. In your town, county commission meetings are held during the day, so you don't have to worry much about night meetings. But how can you tell whether a commissioner is working at night or on weekends?

You can't. But you can set up some rules that will give you a hint about how much time they're spending on the cell phone at times that they're less likely to be working. In Gabalot County, most commissioners get to work around 8:00 a.m., unless they have an early breakfast meeting. They leave around 5:30, unless they have an early evening meeting.

So let's set up some rules about whether they're likely to be working or not:

Before 6:00 a.m., work is doubtful.
Between 6:00 a.m. and 8:00 a.m. it's a tossup.
Between 8:00 a.m. and 5:30 p.m., work is likely.

Between 5:30 p.m. and 7:30 p.m., it's a tossup.
After 7:30 p.m., work is doubtful.

You could probably make a huge formula in your spreadsheet that would test for these conditions. There's an easier way to look up information based on rules like these in a spreadsheet. It's called a lookup formula.

The simplest lookup formulas work like telephone directories: Once you find the name of the person you want to call, you glance over and pick up the corresponding phone number. Anyone who's used a database manager like FoxPro or Access has probably used lookup tables by joining one table of data, like campaign contributions, with details about a field, like PACs.

In this case, you'll use the spreadsheet equivalent of a database join using a formula called VLOOKUP. It looks for your value in another part of the spreadsheet and gives you back the corresponding value from an adjacent column.
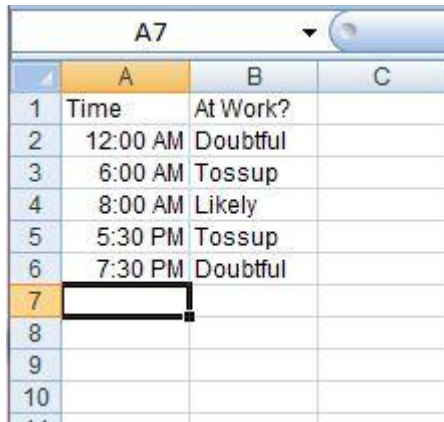
VLOOKUP goes a little further than this simple join. If you've used a database manager, you know that joining two tables requires an exact match. If my Social Security number were off by one digit in one table, the database would treat me as if I didn't exist. That's good. But in this case, we don't want to have to type all of the possible times that commissioners used their phones. Instead, we want to set up slots of time within which the day.

Let's set this up on another worksheet within our CarPhone workbook. That'll keep it separate from the rest of our analysis.

• Insert a new worksheet by clicking on the "Insert worksheet" button in the lower right corner where the worksheet names are located. You'll be taken to a new worksheet named Sheet1 within the same spreadsheet.

• Type in your cutoff hours from the rules above into the first column. Begin with midnight, since that's 0 to a spreadsheet. (Remember, times are actually fractions of a day. So midnight is zero, noon is .5, and 10:00 p.m. is .91 2/3.)

(To type a time, just type it in, with a space between the time and the AM or PM indicator. If the formula bar shows extra precision, like 12:00:00 AM instead of 12:00 AM, and if the spreadsheet justifies the number to the right, you've successfully typed in a time.)

Here's what it should look like when you're done.

VLOOKUP will check each time in your cell phone log against this table. If it doesn't find an exact match, it will give you back the value corresponding to its next lowest value. In this example, 6:30 p.m. will be treated the same way as 5:30 p.m; 7:00 a.m. will be treated as if it's 6:00 a.m To fill out the lookup table, we'll associate these times with the words Doubtful, Likely, Tossup. This means, for instance, that any time between 6:00 a.m. to 7:59:59 a.m. will be treated as a Tossup.

The trick to these kinds of lookups is that your leftmost column – the one you're looking up – must be sorted.

• Just to be sure it's all been typed in correctly, you need to sort your lookup table. Select A1:B6, choose Data, Sort. Then choose Time and Ascending in the first sort area. (If it doesn't look like the table shown above, you've done something wrong. Keep trying until it looks like this table. It must be sorted on the first column for this lookup function to work properly. The worst part is, you'll get an answer but it'll be wrong if it's not sorted.)

Finally, let's make it a little easier on ourselves later by naming this area in the spreadsheet.

• Select cells A2:B6 (the data area of your lookup table). Go to the Formulas Ribbon and go to the section called "Name Manager" and choose "Define Name." It will bring up a box like this, where you can name the area. Let's name this "LkupTime" (all one word).
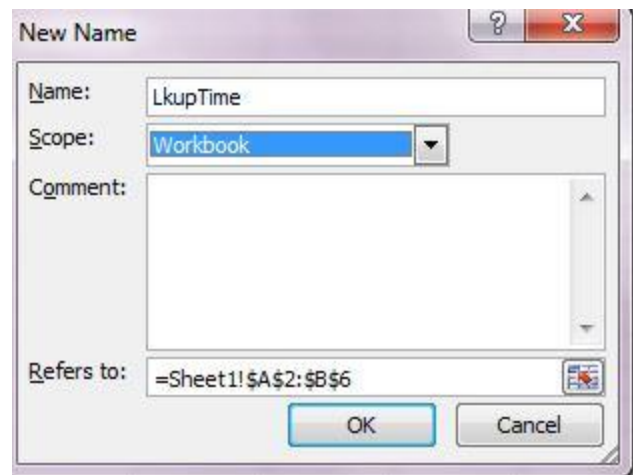
From now on, anytime you use the term LkupTime in this spreadsheet, it will refer to this area, or the lookup table.

• Now you're ready to go back to the original database. Choose the tab at the bottom of the worksheet called RawData.

• Select column H in the original worksheet, and insert a blank column. Type "TimeType" as the column header in column H1, press enter, and select cell H2. Before you type in the lookup formula, you have to know how it's structured. You'll use the VLOOKUP formula, which tells the spreadsheet to search for your value down a column of sorted values. It's structured like this:

=VLOOKUP(what, where, answer, how)

The "what" is the value you want looked up. In this case, we want to look up the value in G2, or 11:50 PM.
The "where" is where you want to look it up. In this case, we want to look it up in the table we just typed in, which our spreadsheet now recognizes as LkupTime.

The "answer" is the column number in your lookup table that holds the value you want back. We want the only other information in the table, or the corresponding word in column 2 of our table.

And the "how" (which is optional) tells the spreadsheet whether you want to check ranges in a sorted table versus an exact match in an unsorted table. This is TRUE for us, since we have a sorted table and want the formula to give us the next lower value if there isn't an exact match on the time. (Use FALSE here if you want to get an error value when there's no exact match or if you can't sort your lookup table for some reason.)

• Type it this formula now: =vlookup(g2,lkuptime,2,true)
· Hit enter, and here is what you get back.

• Copy the formula down through your database.

At work or not?

We have one last problem. Scroll through your data, and soon you'll find a case in which we've labeled the commissioner "Likely" to be at work on a Saturday or Sunday. Our last step in setting up our cell phone logs is to put together the day of the week with the time of day.

• Choose column I and insert one last column. Type "AtWork" as column header into cell I1 and hit enter. Select cell I2. Type =if(or(F2=7,F2=1),"Doubtful",H2)

Let's walk through that formula, taking it from the inside out:

OR(F2=7,F2=1) is checking the Weekday column. Even though we've formatted it to tell us the day of the week in words, this column really holds a numerical day of the week, beginning with 1 for Sunday and ending with 7 for Saturday. So this part of the formula checks to see if the value is Saturday (7) or Sunday (1).

The next phrase, "Doubtful," is what we want the formula to give us if either of those two conditions is true. In other words, we doubt a commissioner is at work on Saturday or Sunday. The last part of the formula tells us what to do the rest of the time. We'll assume that if it's not Saturday or Sunday, we'll use the value we figured out in the Time Type column.

• Go ahead and hit enter if you haven't already, and copy this formula down.

Summarizing your results

The last step is to summarize your results using one of Excel's most powerful features, the Pivot table. You could skip this part and import your information directly into a database to accomplish something very close to the same thing.

You're going to compute the amount of time that each commissioner spends on the phone during each category of work time: Doubtful, Likely and Tossup. You'll also see what percentage of each commissioner's cell phone use time falls within each category.

Select all of your data. Go to the Insert Ribbon and choose "Pivot Table."

Select "Who" to put it in the Row section.

Select "AtWork" and put it in the Column section.

Select "Mins" and put it in the Values section. Excel will change this to the "Sum of Mins."

Select "Mins" again and put it in the Values section. Click on it and go to "value field settings." Under the "show values as" tab choose from the pulldown menu "% of row" (see example to the right)
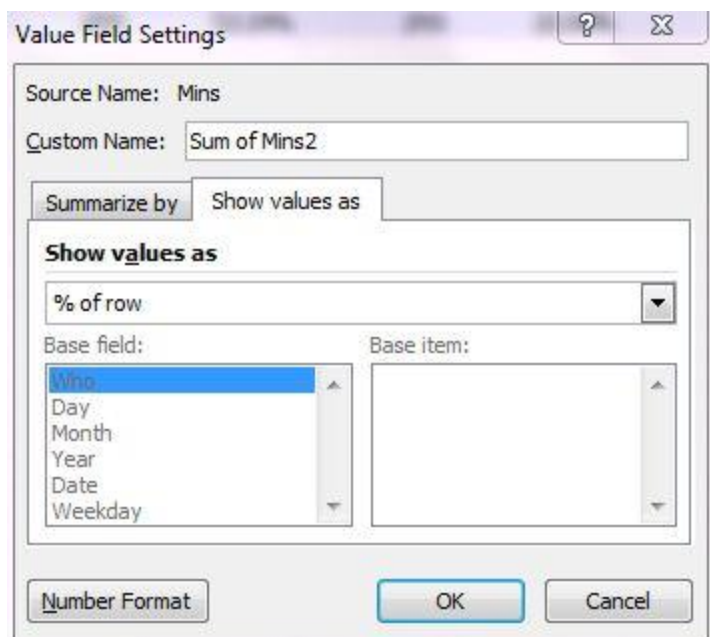
We're interested in the percent of time each commissioner spent on the phone during doubtful, likely or tossup times and days. Since we'll look at each commissioner across a row, we want to find out the percent of the row in each category. Your answer will look similar to this (although the percentages and totals might have different numbers).

| Row Labels | Column Labels | | | | | | Total Sum of Mins | Total Sum of Mins2 |
| | Doubtful | | Likely | | Tossup | | | |
	Sum of Mins	Sum of Mins2	Sum of Mins	Sum of Mins2	Sum of Mins	Sum of Mins2		
Jones	180	30.46%	301	50.93%	110	18.61%	591	100.00%
Remington	4	4.30%	67	72.04%	22	23.66%	93	100.00%
Smith	49	37.12%	46	34.85%	37	28.03%	132	100.00%
Wesson	38	22.49%	85	50.30%	46	27.22%	169	100.00%
Grand Total	271	27.51%	499	50.66%	215	21.83%	985	100.00%

ASSIGNMENT:
#1 Turn in your Excel file that shows your work
#2 Write a sentence or paragraph summarizing your findings.
#3 Explain what you would do next if you were going to use this for a story. Include as much detail as possible.