# 1

# Range Names

- I want to add up sales in Arizona, California, Montana, New York, and New Jersey. Can I use a formula to compute total sales in a form such as *AZ+CA+MT+NY+NJ* instead of *SUM(A21:A25)* and still get the right answer?
- To compute total sales for a year, I often average the 12 monthly sales below the current cell. Can I name my formula *annualaverage* so that when I enter *annualaverage* in a cell the appropriate average is computed?
- How can I easily select a cell range?
- How can I paste a list of all range names (and the cells they repre sent) into my spreadsheet?

You have probably received spreadsheets that use formulas such as *SUM(A5000:A5049)*. Then you have to struggle to understand what's contained in cells A5000:A5049. If cells A5000:A5049 contain sales in each U.S. state, wouldn't the spreadsheet be easier to understand if the formula was *SUM(USSales)*? In this chapter, I'll teach you how to name individual cells, ranges of cells, constants, and formulas. I'll also show you how to use range names in formulas.

# How Can I Create Range Names?

There are three ways to create range names:

- Entering a range name in the Name box
- Choosing the Name, Create command from the Insert menu
- Choosing the Name, Define command from the Insert menu

## Using the Name Box to Create a Range Name

The Name box is located directly above the label for column A, as you can see in Figure 1-1. (To see the Name box, you need to display the Formula bar.) To create a range name using the Name box, simply select with the mouse the cell or range of cells that you want to name, click in the Name box, and then type the range name you want to use. Press Enter, and you've created the range name. Clicking on the drop-down arrow for the Name box displays the range names defined in the current workbook. You can also display all the range names in a workbook by pressing the F3 button, which displays the Paste Name dialog box. When you select a range name from the Name box, Excel selects the cells corresponding to that range name. This enables you to check that you've chosen the cell or range that you intended to.

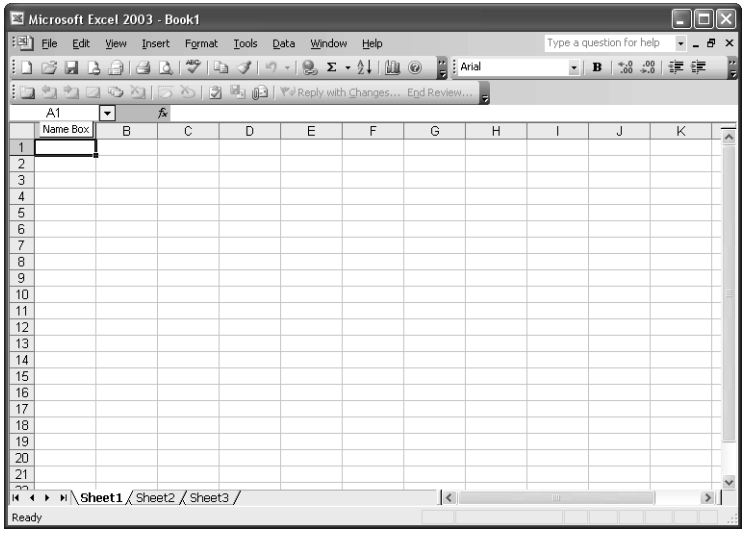



**Figure 1-1** You can create a range name by selecting the cell range you want to name and then typing the range name in the Name box.

## Creating Range Names by Using the Name Create Command

The spreadsheet States.xls contains sales during March for each of the 50 U.S. states. Figure 1-2 shows a subset of this data. We would like to name each cell in the range B6:B55 with the correct state abbreviation. To do this, select the range A6:B55, choose Insert, Name Create, and then choose the Create Names In Left Column option, as indicated in Figure 1-3.

| | A | B |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | State | March Sales |
| 6 | AL | $ 915.00 |
| 7 | AK | $ 741.00 |
| 8 | AZ | $ 566.00 |
| 9 | AR | $ 754.00 |
| 10 | CA | $ 687.00 |
| 11 | CO | $ 757.00 |
| 12 | CT | $ 786.00 |
| 13 | DE | $ 795.00 |
| 14 | FL | $ 944.00 |
| 15 | GA | $ 624.00 |
| 16 | HI | $ 663.00 |
| 17 | ID | $ 895.00 |
| 18 | IL | $ 963.00 |
| 19 | IN | $ 854.00 |

**Figure 1-2** By naming the cells that contain state sales with state abbreviations, you can use the abbreviation when you refer to the cell rather than the cell's column and row number.

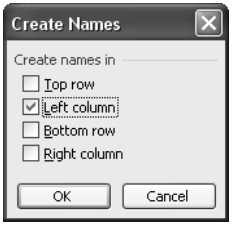



**Figure 1-3** Use the Name, Create command on the Insert menu to name cell ranges. The Create Names dialog box provides options for naming cell ranges.

Excel now knows to associate the names in the left column of the selected range with the cells in the second column of the selected range. Thus B6 is assigned the range name *AL*, B7 is named *AK*, and so on. Note that creating these range names in the Name box would have been incredibly tedious! Click on the drop-down arrow in the Name box if you don't believe that these range names have been created.

## Creating Range Names by Using the Name Define Command

If you choose the Insert, Name Define command, the dialog box shown in Figure 1-4 comes up.

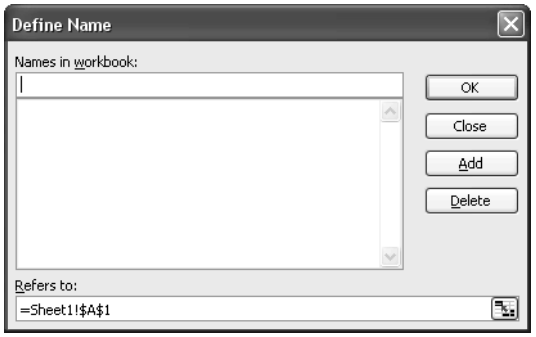



**Figure 1-4** The Define Name dialog box before creating any range names.

Suppose you want to assign the name *range1* (range names are not case sensitive) to the cell range A2:B7. Simply type *range1* in the Names In Workbook box and then go down to the Refers To area and point to the range or type in *=A2:B7*. Click Add, and you're done. The Define Name dialog box will now look like Figure 1-5.

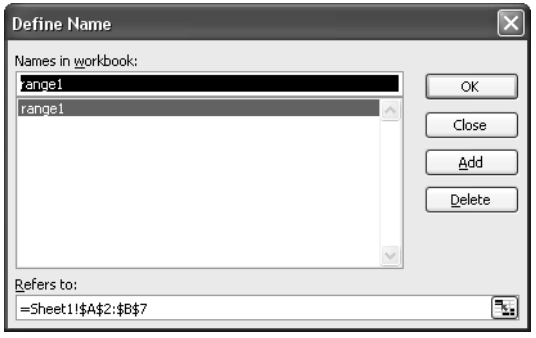



**Figure 1-5** Define Name dialog box after creating a range name.

Of course, if you now click in the Name box, the name *range1* will appear. Now let's look at some specific examples of how to use range names.

## I want to add up sales in Arizona, California, Montana, New York, and New Jersey. Can I use a formula to compute total sales in a form such as *AZ+CA+MT+NY+NJ* instead of *SUM(A21:A25)* and still get the right answer?

Let's return to the file States.xls in which we assigned each state's abbreviation as the range name for the state's sales. If we want to compute total sales in Alabama, Alaska, Arizona, and Arkansas, we could clearly use the formula *SUM(B6:B9)*. We could also point to cells B6, B7, B8, and B9, and the formula would be entered as *=AL+AK+AZ+AR*. The latter formula is, of course, much easier to understand.

As another illustration of how to use range names, look at the file HistoricalInvest.xls, shown in Figure 1-6, which contains annual percentage returns on stocks, T-Bills, and bonds. (The rows for years 1935–1996 are hidden.)

| | A | B | C | D |
|---|---|---|---|---|
| 4 | | | | |
| 5 | | | | |
| 6 | | Annual Returns on Investments in | | |
| 7 | Year | Stocks | T.Bills | T.Bonds |
| 8 | 1928 | 43.81% | 3.08% | 0.84% |
| 9 | 1929 | -8.30% | 3.16% | 4.20% |
| 10 | 1930 | -25.12% | 4.55% | 4.54% |
| 11 | 1931 | -43.84% | 2.31% | -2.56% |
| 12 | 1932 | -8.64% | 1.07% | 8.79% |
| 13 | 1933 | 49.98% | 0.96% | 1.86% |
| 14 | 1934 | -1.19% | 0.30% | 7.96% |
| 77 | 1997 | 31.86% | 4.91% | 9.94% |
| 78 | 1998 | 28.34% | 5.16% | 14.92% |
| 79 | 1999 | 20.89% | 4.39% | -8.25% |
| 80 | 2000 | -9.03% | 5.37% | 16.66% |
| 81 | 2001 | -11.85% | 5.73% | 5.57% |
| 82 | | | | |
| 83 | | stocks | tbills | bonds |
| 84 | averages | 12.05% | 3.96% | 5.21% |

**Figure 1-6** Historical investment data.

After selecting the cell range B7:D81 and choosing Insert, Name Create, we choose to create names in the top row of the range. The range B8:B81 is named *Stocks*, the range C8:C81 *T.Bills*, and the range D8:D81 *T.Bonds*. Now we no longer need to remember where our data is. For example, in cell B84, after typing *=AVERAGE(*, we can press F3 and the Paste Name dialog box appears, as shown in Figure 1-7.

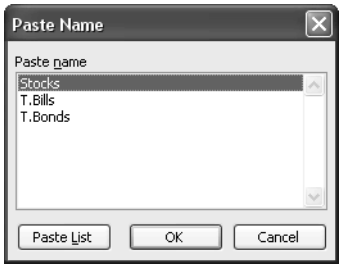



**Figure 1-7** You can add a range name to a formula by using the Paste Name dialog box.

Then we can select Stocks in the Paste Name list and click OK. After entering the closing parentheses, our formula, *=AVERAGE(Stocks)*, computes the average return on stocks (12.05 percent). The beauty of this approach is that even if we don't remember where the data is, we can work with the stock return data anywhere in the workbook!

If we use a column name (in the form A:A, C:C, and so on) in a formula, Excel treats an entire column as a range name. For example, entering the formula *=AVERAGE(A:A)* will average all numbers in column A. Using a range name for an entire column is very helpful if you frequently enter new data into a column. For example, if column A contains monthly sales of a product, as new sales data is entered each month, our formula always computes an up-to-date monthly sales average. I caution you, however, that if you enter the formula *=AVERAGE(A:A)* in column A, you will get a circular reference message because the value of the cell containing the average formula depends on the cell containing the average. You will learn how to resolve circular references later in the book, in Chapter 11.

## How Do I Delete a Range Name?

To delete a range name, simply choose Insert, Name Define to bring up the Define Name dialog box. (See Figure 1-5.) Select the range name you want to delete, and then click Delete. Unfortunately, Excel lets you delete only one range name at a time.

## How Do I Change a Range Name?

To change a range name, select Insert, Name Define to bring up the Define Name dialog box. (See Figure 1-5.) Select the range whose name you want to

change, and then enter a new name or the cell range to which the range name refers. You can change the range by typing in the new range or by selecting the new range in the worksheet. Click OK, and the range is changed.

## How Do I Name a Constant?

In many spreadsheets a value occurs so often that you might want to name it to use it as a constant. For example, suppose sales during the current year are $10 million and sales are growing 10 percent a year. (See the file Growth.xls.) We can name the annual growth rate *Growth* by choosing Insert, Name Define, filling in the Names In Workbook box and the Refers To area as shown in Figure 1-8, and then clicking OK.

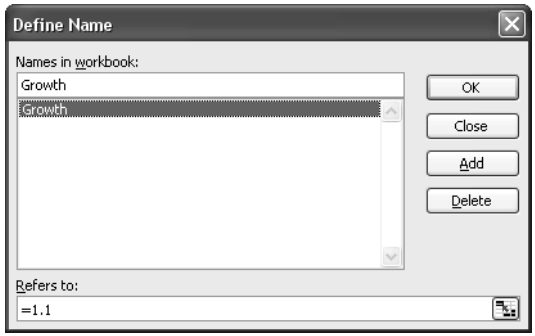



**Figure 1-8** You can name a value that occurs frequently in a workbook and use it as a constant.

Now everywhere in your spreadsheet, *Growth* will be interpreted as 1.1.

If you haven't already, open the file Growth.xls. Copying from F11 to G11:H11 the formula *E11*Growth* generates the sales for years 1–6 that grow at 10 percent a year, as you can see in Figure 1-9.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | | | | | | |
| 4 | | | | | | | | |
| 5 | | | | | | | | |
| 6 | | | | | | | | |
| 7 | | | | | | | | |
| 8 | | | | | | | | |
| 9 | | | | | | | | |
| 10 | Year | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 11 | | 10000000 | 11000000 | 12100000 | 13310000 | 14641000 | 16105100 | 17715610 |

**Figure 1-9** Sales growth was calculated using a named constant.

### To compute total sales for a year, I often average the 12 monthly sales below the current cell. Can I name my formula *annualaverage* so that when I enter *annualaverage* in a cell the appropriate average is computed?

Often you perform the same type of operation over and over. Wouldn't it be nice if you could name the formula that performs the operation? Then, whenever you needed to use the formula, you could call up the formula from the Paste Name dialog box. Here's an example.

Suppose that we repeatedly average 12 cells below the current cell to compute average monthly sales for several products during a calendar year. (See the file Sales.xls and Figure 1-10.) We would like to start 2 rows below the current cell and average 12 cells to get an average of monthly sales. To name this operation, we begin by going to any cell. I used cell B2 in this example. Choose Insert, Name Define to bring up the Define Name dialog box. Select the name *annualaverage,* and then look at the Refers To area, as indicated in Figure 1-11.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | average | 261.75 | 254.3333 | 256.3333 | 218.1667 |
| 3 | Month | Product 1 | Product 2 | Product 3 | Product 4 |
| 4 | Jan | 171 | 231 | 281 | 257 |
| 5 | Feb | 292 | 226 | 303 | 223 |
| 6 | Mar | 285 | 393 | 175 | 196 |
| 7 | Apr | 100 | 260 | 370 | 153 |
| 8 | May | 358 | 238 | 317 | 105 |
| 9 | Jun | 253 | 205 | 242 | 180 |
| 10 | Jul | 317 | 301 | 398 | 312 |
| 11 | Aug | 397 | 385 | 296 | 269 |
| 12 | Sep | 393 | 163 | 350 | 259 |
| 13 | Oct | 155 | 387 | 113 | 132 |
| 14 | Nov | 169 | 150 | 131 | 321 |
| 15 | Dec | 251 | 113 | 100 | 211 |

**Figure 1-10** Data for naming a formula to average 12 months of sales.

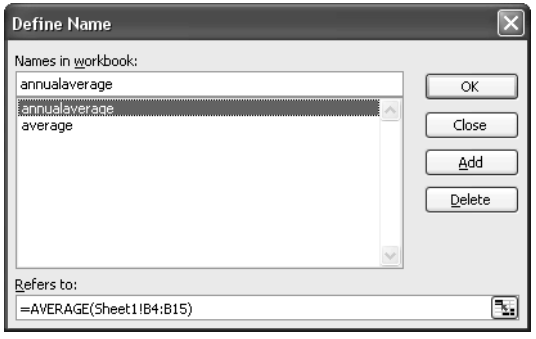



**Figure 1-11** Naming a formula to average 12 months of sales.

You enter the formula that performs the operation given the current cell location. Because we are in cell B2, we want to average cells B4:B15, so we enter the formula *=AVERAGE(B4:B15)*. You must type in the formula you want rather than enter the formula by pointing. Now, if we enter the name *annualaverage* in any cell, the formula will drop down 2 cells and average the next 12 cells below the current cell. To illustrate how this named formula works, I've entered the formula *=annual average* in cells C2:E2. You now obtain the annual average of sales for products 2–4. Note that if you are in cell E2 and you select Insert, Name Define, the Refers To area indicates that the range *annualaverage* refers to the formula *=Average(E4:E15)*, which is the correct formula!

### How can I easily select a cell range?

If you have selected a cell within a cell range, press Ctrl+Shift+* to select the entire range.

### How can I paste a list of all range names (and the cells they represent) into my spreadsheet?

Press F3 to display the Paste Name box, and then click the Paste List button. A list of range names and the cells each corresponds to will be pasted into your spreadsheet, beginning at the current cell location.

## Remarks

- Excel does not allow you to use the letters r and c as range names. Instead it will create a range name r_ or c_.
- If you use Insert, Name Create to create a range name and your name contains spaces, Excel inserts an underscore (_) to fill in the spaces. For example, the name Product 1 is created as Product_1.
- Range names cannot begin with numbers or look like a cell refer ence. For example, 3Q or A4 are not allowed as range names.
- The only symbols allowed in range names are periods (.) and underscores (_).

# Problems

1. The file Stock.xls contains monthly stock returns for General Motors and Microsoft. Name the ranges containing the monthly returns for each stock and compute the average monthly return on each stock.
2. Create a formula to add up the six numbers below a given cell.
3. Suppose you want to name the range containing the cells A1:B3 and A6:B8. You can select multiple ranges in Excel by holding down the Ctrl key between the selection of different ranges. Use this technique to give the cell range consisting of A1:B3 and A6:B8 the name *Red*.

# 2

# Natural Language Range Names

- How do I create a name for a range of cells based on a spreadsheet label and use that name in formulas?
- How do I use natural language range names to incorporate a relationship such as *Month(t) Ending Inventory = Month(t) Beginning Inventory + Month(t) Production –Month(t) Demand* in a spreadsheet formula?

Microsoft Excel 2002 and Microsoft Office Excel 2003 give you the ability to create "natural language" range names based on labels already present in your spreadsheet. I'll illustrate the use of natural language range names with two examples.

> **Note** This feature is available only in Excel 2002 and Excel 2003.

## How do I create a name for a range of cells based on a spreadsheet label and use that name in formulas?

In the file NaturalRange.xls, shipments from three plants to four cities are listed. We want to find the total amount shipped from each plant and the total shipped into each city. Figure 2-1 shows the information in NaturalRange.xls.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | City 1 | City 2 | City 3 | City 4 | |
| 5 | Plant 1 | 174 | 102 | 197 | 100 | 573 |
| 6 | Plant 2 | 162 | 197 | 103 | 159 | 621 |
| 7 | Plant 3 | 189 | 185 | 163 | 175 | 712 |
| 8 | | 525 | 484 | 463 | 434 | |
| 9 | | | | | | |
| 10 | Plant 1 City 1 | 174 | | | | |

**Figure 2-1** Shipments from plants to cities.

To use the natural language range feature, choose Tools, Options. On the Calculation tab, select the Accept Labels In Formulas option. Clearing this option makes the natural language range names invalid.

Next we enter in cell F5 the formula *SUM(Plant 1)*. Note that the total shipments from Plant 1 (573 units) have been tabulated. Excel sensed from the Plant 1 label in row 5 that you wanted to name the range B5:E5 *Plant 1*. This is amazing, but there's more! Copy the formula in F5 to the range F6:F7. In F6 you will see the formula *SUM(Plant 2)*, and in F7 you will see *SUM(Plant 3)*. You can see that Excel is smart enough to determine the "natural" way to copy the formula in F5, and it adjusted the range names used in F6 and F7 accordingly. In a similar fashion, copying the formula *SUM(City 1)* from B8 to C8:E8 computes the total shipments into each of the four cities. This use of natural range names makes the logic of the spreadsheet much easier to understand.

Finally, suppose we want to enter the amount shipped from Plant 1 to City 1 into a cell. We could just use the formula *=B5*. Alternatively, we could use the formula *=Plant 1 City 1*. Excel now looks for the cell at the *intersection* of the natural language names defined by Plant 1 and City 1. Of course, this is the spreadsheet entry in cell B5.

## How do I use natural language range names to incorporate a relationship such as *Month(t) Ending Inventory = Month(t) Beginning Inventory + Month(t) Production –Month(t) Demand* in a spreadsheet formula?

Often, spreadsheets contain modeling relationships that relate quantities at different points in time. For example, if we are trying to update our ending inventory each month, an appropriate formula is:

```
Month(t)Ending Inventory = Month(t)Beginning Inventory + Month(t)Production -
Month(t)Demand
```

If we are updating a toy store's cash on hand at the end of every month, an appropriate formula is:

```
Month(t)Ending Cash = Month(t)Beginning Cash + Month(t)Cash Inflow -Month(t)Cash
Outflow
```

Natural language range names make it easy for a person viewing a spreadsheet to see how this logic was included in the spreadsheet. The file NatRangeII.xls (see Figure 2-2) shows how to model the computation of monthly inventory levels with natural language range names. We begin month 1 with 100 units in inventory and produce 150 units each month. Monthly demands are given in column D of the spreadsheet.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | Month | Beg Inv | Prod | Demand | End Inv |
| 5 | 1 | 100 | 150 | 149 | 101 |
| 6 | 2 | 101 | 150 | 172 | 79 |
| 7 | 3 | 79 | 150 | 175 | 54 |
| 8 | 4 | 54 | 150 | 146 | 58 |
| 9 | 5 | 58 | 150 | 136 | 72 |
| 10 | 6 | 72 | 150 | 126 | 96 |

**Figure 2-2** Computing inventory with natural language range names.

In cell B6 we set month 2 beginning inventory equal to month 1 ending inventory by entering the formula *=E5*. Copying this formula to the range B7:B10 computes beginning inventory for months 3–6. We use natural language range names to compute ending inventory by copying from E5 to E6:E10 the following formula:

```
=Beg Inv + Prod - Demand
```

To see what this formula does, focus on any of rows 5–10. For example, to evaluate the formula in cell E5, Excel finds the number in row 5 from the column headed Beg Inv, adds this number to the number in row 5 from the Prod column, and then subtracts the number in row 5 from the Demand column. Copying this formula down to rows 6 through 10 now computes the correct ending inventory level for each month. Note that if we had copied *=B5+C5-D5* down to E6:E10, few people would have understood the logic of our computations. The formula *=Beg Inv+ Prod - Demand* makes our logic more apparent.

# Problems

**1.** A small candy store has $6,000 in cash on hand on January 1. Cash inflows and outflows for the months January–June are shown in the following table. Use natural language range names to compute each month's ending cash position.

| Month | Inflow | Outflow |
|---|---|---|
| January | $5,000 | $4,500 |
| February | $6,000 | $5,000 |
| March | $5,200 | $6,500 |
| April | $7,200 | $7,000 |
| May | $8,000 | $8,000 |
| June | $9,000 | $7,500 |

**2.** The following table gives the dollar sales of three products in four regions of the country: East, Midwest, South, and North. Use natural language range names to compute total sales of each product and the total dollar volume of sales in each region.

| | East | Midwest | South | North |
|---|---|---|---|---|
| Product 1 | $500 | $600 | $700 | $650 |
| Product 2 | $600 | $536 | $665 | $888 |
| Product 3 | $999 | $834 | $432 | $326 |

# 3

# Lookup Functions

- How do I write a formula to compute tax rates based on income?
- Given a product ID, how can I look up the product's price?
- Suppose that a product's price changes over time. I know the date the product was sold. How can I write a formula to compute the product's price?

## Syntax of the LOOKUP Functions

Lookup functions enable you to "look up" values from worksheet ranges. Excel allows you to perform both vertical lookups (using VLOOKUP functions) and horizontal lookups (using HLOOKUP functions). In a vertical lookup, the lookup operation starts in the first column of a worksheet range. In a horizontal lookup, the operation starts in the first row of a worksheet range. Because the great majority of formulas using lookup functions involve vertical lookups, we'll concentrate on VLOOKUP functions.

### VLOOKUP Syntax

The syntax of the VLOOKUP function is as follows. The brackets ([ ]) indicate optional arguments.

```
VLOOKUP(lookup value, table range, column index,[range lookup])
```

- *Lookup value* is the value that we want to look up in the first column of the table range.

- *Table range* is the range that contains the entire lookup table. The table range includes the first column, in which we try and match the lookup value, and any other columns in which we will look up formula results.
- *Column index* is the column number in the table range from which the value of the lookup function is obtained.
- *Range lookup* is an optional argument. If *range lookup* is True or is omitted, the first column of the table range must be in ascending numerical order. If *range lookup* is True or omitted and an exact match to the lookup value is found in the first column of the table range, Excel keys the lookup off the row of the table in which the exact match is found. If *range lookup* is True or omitted and an exact match does not exist, Excel keys the lookup off the largest value in the first column that is less than the lookup value. If *range lookup* is False and an exact match to the lookup value is found in the first column of the table range, Excel keys the lookup off the row of the table in which the exact match is found. If no exact match is obtained, Excel returns an #N/A response (Not Available).

## HLOOKUP Syntax

For an HLOOKUP function, simply change "column" to "row." Thus, in an HLOOKUP function, Excel tries to locate the lookup value in the first row (not the first column) of the table range.

Let's explore some interesting examples of lookup functions.

### How do I write a formula to compute tax rates based on income?

The following example shows how a VLOOKUP function works when the first column of the table range consists of numbers in ascending order. Suppose that the tax rate depends on income as follows:

| Income Level | Tax Rate |
|---|---|
| $0-$9,999 | 15 percent |
| $10,000-$29,999 | 30 percent |
| $30,000-$99,999 | 34 percent |
| $100,000 and over | 40 percent |

To see how to write a formula that computes the tax rate for any income level, open the file Lookup.xls, shown in Figure 3-1.

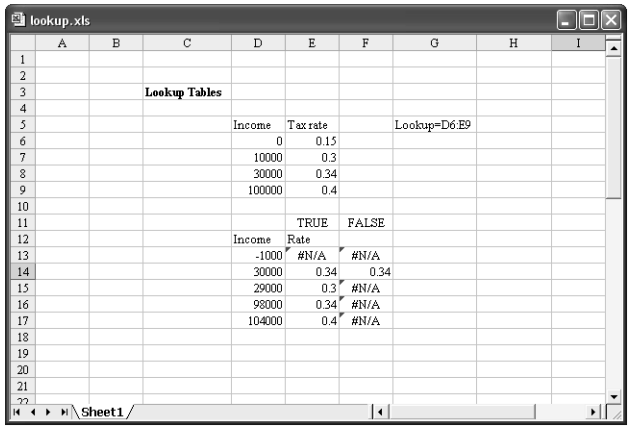



**Figure 3-1** Using a lookup function to compute a tax rate. The numbers in the first column of the table range are sorted in ascending order.

I began by entering the relevant information (tax rates and break points) in cell range D6:E9. I named the table range D6:E9 *lookup*. I recommend that you always name the cells you're using as the table range. Then you need not remember the exact location of the table range, and when you copy any formula involving a lookup function, the lookup range will always be correct. To illustrate how the lookup function works, I entered some incomes in the range D13:D17. By copying from E13:E17 the formula *VLOOKUP(D13,Lookup,2,True)*, we compute the tax rate for the income levels listed in D13:D17. Let's examine how the lookup function worked in cells E13:E17. Note that because the column index in the formula is 2, the answer always comes from the second column of the table range.

- In D13, the income of -$1,000 yields #N/A because -$1,000 is less than the lowest income level in the first column of the table range.
- In D14, the income of $30,000 exactly matches a value in the first column of the table range, so the function returns a tax rate of 34 percent.
- In D15, the income level of $29,000 does not exactly match a value in the first column of the table range, which means the lookup function stops at the largest number less than $29,000 in the first column of the range; $10,000 in this case. This function returns the tax rate in column 2 of the table range opposite $10,000, or 30 percent.
- In D16, the income level of $98,000 does not yield an exact match in the first column of the table range. The lookup function stops at the largest number less than $98,000 in the first column of the table range. This returns the tax rate in column 2 of the table range opposite $30,000, 34 percent.

- In D17, the income level of $104,000 does not yield an exact match in the first column of the table range. The lookup function stops at the largest number less than $104,000 in the first column of the table range, which returns the tax rate in column 2 of the table range opposite $100,000, 40 percent.
- In F13:F17, we changed the value of the *range lookup* argument from True to False and copied from F13 to F14:F17 the formula *VLOOKUP(D13,Lookup,2,False)*. Cell F14 still yields a 34 percent tax rate because the first column of the table range contains an exact match to $30,000. The other entries in F13:F17 all display #N/A because none of the other incomes in D13:D17 have an exact match in the first column of the table range.

## Given a product ID, how can I look up the product's price?

Often the first column of a table range does not consist of numbers in ascending order. For example, the first column of the table range might list product ID codes or employee names. In my experience teaching thousands of financial analysts, I've found that many people don't know how to deal with lookup functions when the first column of the table range does not consist of numbers in ascending order. In these situations, you need to remember only one simple rule: use False as the value of the *range lookup* argument.

Here's an example. In the file Lookup.xls (see Figure 3-2), you can see the prices for 5 products, listed by their ID code. How do you write a formula that will take a product ID code and return the product price?

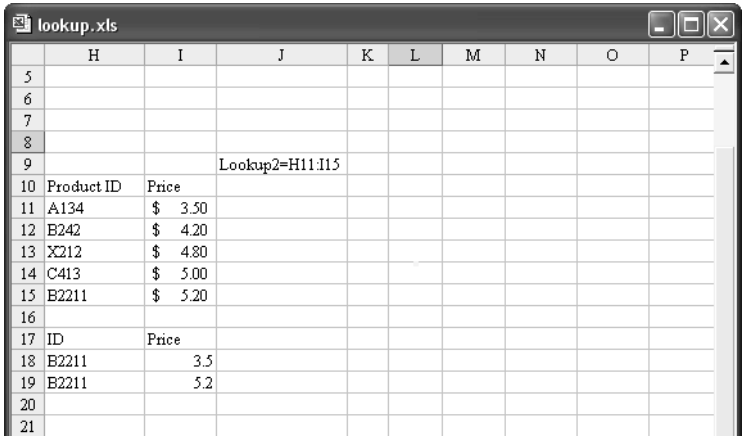



**Figure 3-2** Looking up prices from product ID codes. In cases like these, when the table range isn't sorted in ascending order, enter False as the last argument in the lookup function formula.

Many people would enter the formula as I have in cell I18, *VLOOKUP(H18,Lookup2,2)*. However, note that when you omit the fourth argument (the *range lookup* argument), the value is assumed to be True. Because the product IDs in the table range *Lookup2* (H11:I15) are not listed in alphabetical order, an incorrect price ($3.50) is returned. If we enter in cell I18 the formula *VLOOKUP(H18,Lookup2,2, False),* the correct price ($5.20) is returned.

You would also use False in a formula designed to look up an employee's salary using the employee's last name or ID number.

## Suppose that a product's price changes over time. I know the date the product was sold. How can I write a formula to compute the product's price?

Suppose the price of a product depends on the date the product was sold. How can you use a lookup function in a formula that will pick up the correct product price? More specifically, suppose the price of a product is as follows.

| Date Sold | Price |
|---|---|
| January-April 2005 | $98 |
| May-August 2005 | $105 |
| September-December 2005 | $112 |

We'll write a formula that will determine the correct product price for any date on which the product is sold in the year 2005. For variety, we'll use an *HLOOKUP* function. I've placed the dates when the price changes in the first row of the table range. See the file DateLookup.xls, shown in Figure 3-3.

datelookup.xls

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | Date | 1/1/2005 | 5/1/2005 | 8/1/2005 | | |
| 3 | Price | 98 | 105 | 112 | | |
| 4 | | | | | | |
| 5 | | | Lookup:B2:D3 | | | |
| 6 | | | | | | |
| 7 | | Date | Price | | | |
| 8 | | 1/4/2005 | 98 | | | |
| 9 | | 5/10/2005 | 105 | | | |
| 10 | | 9/12/2005 | 112 | | | |
| 11 | | 5/1/2005 | 105 | | | |
| 12 | | | | | | |

**Figure 3-3** Using an HLOOKUP function to determine a price that changes depending on the date it's sold.

I copied from C8 to C9:C11 the formula *HLOOKUP(B8,lookup,2,TRUE)*. This formula tries to match the dates in column B with the first row of the range

B2:D3. At any date between 1/1/05 and 4/30/05, the lookup function will stop at 1/1/05 and return the price in B3; for any date between 5/01/05 and 7/31/05, the lookup stops at 5/1/05 and returns the price in C3; and for any date later than 8/01/05, the lookup stops at 8/01/05 and returns the price in D3.

# Problems

1. The file Hr.xls gives employee ID codes, salaries, and years of experience. Write a formula that takes a given ID code and yields the employee's salary. Write another formula that takes a given ID code and yields the employee's years of experience.
2. The file Assign.xls gives the assignment of workers to four groups. The suitability of each worker for each group (on a 0-10 scale) is also given. Write a formula that gives the suitability of each worker for the group to which the worker is assigned.
3. You are thinking of advertising Microsoft products on a sports telecast. As you buy more ads, the price of each ad drops as described in the following table.

| Number of Ads | Price Per Ad |
|---|---|
| 1-5 | $12,000 |
| 6-10 | $11,000 |
| 11-20 | $10,000 |
| More than 20 | $9,000 |

   For example, if you buy 8 ads, you pay $11,000 per ad, but if you buy 14 ads, you pay $10,000 per ad. Write a formula that yields the total cost of purchasing any number of ads.

4. You are thinking of advertising Microsoft products on a popular TV music program. You pay one price for the first group of ads, but as you buy more ads, the price per ad drops as described in the following table:

| Number of Ads | Price Per Ad |
|---|---|
| 1-5 | $12,000 |
| 6-10 | $11,000 |
| 11-20 | $10,000 |
| More than 20 | $9,000 |

For example, if you buy 8 ads, you pay $12,000 per ad for the first 5 ads and $11,000 for each of the next 3 ads. If you buy 14 ads, you pay $12,000 for each of the first 5 ads, $11,000 for each of the next 5 ads, and $10,000 for each of the last 4 ads. Write a formula that yields the total cost of purchasing any number of ads. Hint: You will probably need at least three columns in your table range and your formula might involve two lookup functions.

**5.** The annual rate your bank charges you to borrow money for 1, 5, 10, or 30 years is given in the following table.

| Duration of Loan (Years) | Annual Loan Rate (Percent) |
|---|---|
| 1 | 6 |
| 5 | 7 |
| 10 | 9 |
| 30 | 10 |

If you borrow money from the bank for any duration between 1 and 30 years that's not listed in the table, your rate is found by interpolating between the rates given in the table. For example, let's say you borrow money for 15 years. Because 15 years is one quarter of the way between 10 years and 30 years, the annual loan rate would be calculated as follows:

$$\frac{3}{4}(9) + \frac{1}{4}(10) = 9.25\%$$

Write a formula that will return the annual interest rate on a loan for any period between 1 and 30 years.

**6.** The distance between any two U.S. cities (excluding cities in Alaska and Hawaii) can be approximated by the formula

$$69^* \sqrt{(\textit{difference in latitudes})^2 + (\textit{difference in longitudes})^2}$$

The file CityData.xls contains the latitude and longitude of selected U.S. cities. Create a table that gives the distance between any two of the listed cities.

# 4

# The INDEX Function

- I have a list of distances between U.S. cities. How do I write a function that returns the distance between, say, Seattle and Miami?
- Is there any way I can write a formula that references the entire column containing the distances of each city to Seattle?

## Syntax of the INDEX Function

The INDEX function allows you to return the entry in any row and column within a rectangular array of numbers. The most commonly used syntax for the INDEX function is:

```
INDEX(Array, Row Number, Column Number)
```

To illustrate, the formula *INDEX(A1:D12,2,3)* would return the entry in the second row and third column of the array A1:D12. This entry is the one in cell C2.

### I have a list of distances between U.S. cities. How do I write a function that returns the distance between, say, Seattle and Miami?

The file Index.xls (see Figure 4-1) contains the distances between eight U.S. cities. The range C10:J17, which contains the distances, is named Distances.

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |
| 4 | | Boston-Denve | 1991 | | | T Dist to Seattle | 15221 | | | |
| 5 | | Seattle- Miami | 3389 | | | T Dist to Seattle | 15221 | | | |
| 6 | | | | | | | | | | |
| 7 | | | | | | | | | | |
| 8 | | | | | | | | | | |
| 9 | | | Boston | Chicago | Dallas | Denver | LA | Miami | Phoenix | Seattle |
| 10 | 1 | Boston | 0 | 983 | 1815 | 1991 | 3036 | 1539 | 2664 | 2612 |
| 11 | 2 | Chicago | 983 | 0 | 1205 | 1050 | 2112 | 1390 | 1729 | 2052 |
| 12 | 3 | Dallas | 1815 | 1205 | 0 | 801 | 1425 | 1332 | 1027 | 2404 |
| 13 | 4 | Denver | 1991 | 1050 | 801 | 0 | 1174 | 2100 | 836 | 1373 |
| 14 | 5 | LA | 3036 | 2112 | 1425 | 1174 | 0 | 2757 | 398 | 1909 |
| 15 | 6 | Miami | 1539 | 1390 | 1332 | 2100 | 2757 | 0 | 2359 | 3389 |
| 16 | 7 | Phoenix | 2664 | 1729 | 1027 | 836 | 398 | 2359 | 0 | 1482 |
| 17 | 8 | Seattle | 2612 | 2052 | 2404 | 1373 | 1909 | 3389 | 1482 | 0 |

**Figure 4-1** You can use the INDEX function to calculate the distance between U.S. cities.

Suppose that you want to enter the distance between Boston and Denver in a cell. Because distances from Boston are listed in the first row of the array named Distances, and distances to Denver are listed in the fourth column of the array, the appropriate formula is *INDEX(distances,1,4)*. We find that Boston and Denver are 1991 miles apart. Similarly, to find the (huge) distance between Seattle and Miami, you would use the formula *INDEX(distances,6,8)*. Seattle and Miami are 3389 miles apart.

Imagine that the Seattle Sonics basketball team is embarking on a road trip in which they play games in Phoenix, Los Angeles, Denver, Dallas, and Chicago. At the conclusion of the road trip, the Sonics return to Seattle. Can we easily compute how many miles they travel on the trip? As you can see in Figure 4-2, we simply list in order the cities the Sonics visit (8-7-5-4-3-2-8), starting and ending in Seattle, and copy from D21 to D26 the formula *INDEX(distances,C21,C22)*. The formula in D21 computes the distance between Seattle and Phoenix (city number 7), the formula in D22 computes the distance between Phoenix and Los Angeles, and so on. The Sonics will travel a total of 7112 miles on their road trip. By the way, I used the INDEX function to show that the Miami Heat log more miles during the NBA season than any other team.

| | A | B | C | D |
|---|---|---|---|---|
| 18 | | | | |
| 19 | | | **Road Trip!!** | |
| 20 | | | City | Distance |
| 21 | | | 8 | 1482 |
| 22 | | | 7 | 398 |
| 23 | | | 5 | 1174 |
| 24 | | | 4 | 801 |
| 25 | | | 3 | 1205 |
| 26 | | | 2 | 2052 |
| 27 | | | 8 | |
| 28 | | | Total | 7112 |

**Figure 4-2** Distances for a Seattle Sonics road trip.

### Is there any way I can write a formula that references the entire column containing the distances of each city to Seattle?

The INDEX function makes it easy to reference an entire row or column of an array. If we set the row number to 0, the INDEX function references the listed column. If we set the column number to 0, the INDEX function references the listed row. To illustrate, suppose we want to total the distances from each listed city to Seattle. We could enter either of the following formulas:

```
SUM(INDEX(distances,8,0))
SUM(INDEX(distances, 0,8))
```

The first formula adds up the numbers in the eighth row (row 17) of the Distances array; the second formula adds up the numbers in the eighth column (Column J) of the Distances array. In either case we find the total distance from Seattle to the other cities is 15,221 miles, as you can see in Figure 4-3.

| | F | G |
|---|---|---|
| 3 | | |
| 4 | T Dist to Seattle | 15221 |
| 5 | T Dist to Seattle | 15221 |

**Figure 4-3** Total distances of cities to Seattle.

# Problems

1. Use the INDEX function to compute the distance between LA and Phoenix and between Denver and Miami.
2. Use the INDEX function to compute the total distance from Dallas to all cities.
3. Mark Cuban and the Dallas Mavericks are embarking on a road trip that takes them to Chicago, Denver, Los Angeles, Phoenix, and Seattle. How many miles will they travel on this road trip?
4. The file Product.xls contains monthly sales for six products. Use the INDEX function to compute the sales of product 2 in March. Use the INDEX function in a formula that computes total sales during April.

# 5

# The MATCH Function

- Given monthly sales for several products, how do I write a formula that will calculate the sales of a product during a given month? For example, how much of product 2 did I sell during June?
- Given a list of baseball player salaries, how do I write a formula that yields the player with the highest salary? How about the player with the fifth highest salary?
- Given the annual cash flows from an investment project, how do I write a formula that returns the number of years required to pay back the project's initial investment cost?

Suppose you have a spreadsheet with 5000 rows containing 5000 names. You need to find the name John Doe, which you know appears somewhere (and only once) in the list. Wouldn't you like to know a formula that would return the row number at which John Doe is located? The Excel MATCH func tion enables you to find within a given array the first occurrence of a "match" to a given text string or number. You should use the MATCH function instead of a LOOKUP function in situations in which you want the position of a number in a range rather than the value in a particular cell. The syntax of the match function is:

```
Match(lookup value, lookup range, match type)
```

In the explanation that follows, we'll assume that all cells in the lookup range are located in the same column. In this syntax,

- *Lookup value* is the value you're trying to match in the lookup range.
- *Lookup range* is the range you're examining for a "match" to the lookup value.

- *Match type =1* requires your lookup range to consist of numbers listed in ascending order. The MATCH function then returns the row location in the lookup range (relative to the top of the lookup range) that contains the largest value in the range that is less than or equal to the lookup value. *Match type = -1* requires the lookup range to consist of numbers listed in descending order. The MATCH function returns the row location in the lookup range (relative to the top of the lookup range) that contains the last value in the range that is greater than or equal to the lookup value. No matter in what order the values or text in the lookup range are listed (ascending, descend ing, or no order), *match type = 0* returns the row location in the lookup range that contains the first exact match to the lookup value. When no exact match exists and *match type = 0*, Excel returns #N/A. The great majority of MATCH function applications use *match type = 0*, but if *match type* is not included, *match type = 1* is assumed.

The file MatchEx.xls, shown in Figure 5-1, contains three examples of the MATCH function's syntax.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| 4 | | Boston | | | -5 | | 6 |
| 5 | | Chicago | | | -4 | | 5 |
| 6 | | Dallas | | | -3 | | 4 |
| 7 | | Denver | | | -1 | | 3 |
| 8 | | LA | | | 3 | | -1 |
| 9 | | Miami | | | 4 | | -3 |
| 10 | | Phoenix | | | 5 | | -4 |
| 11 | | Seattle | | | 6 | | -5 |
| 12 | | | | last number<=0 | 4 | last number>=-4 | 7 |
| 13 | Boston | 1 | | | | | |
| 14 | Phoenix | 7 | | | | | |
| 15 | Pho* | 7 | | | | | |

**Figure 5-1** Examples of using the MATCH function to locate the position of a value in a range.

In cell B13, the formula *MATCH("Boston",B4:B11,0)* returns 1 because the first row in the range B4:B11 contains the value *Boston*. Notice that text values must be enclosed in quotation marks (""). In cell B14, the formula *MATCH("Phoenix",B4:B11,0)* returns 7 because cell B10 (the seventh cell in B4:B11) is the first cell in the range that matches *"Phoenix"*. In cell E12, the for mula *MATCH(0,E4:E11,1)* returns 4 because the last number that is less than or equal to 0 in the range E4:E11 is in cell E7 (the fourth cell in the lookup range). In cell G12, the formula *MATCH(-4,G4:G11,-1)* returns 7 because the last num ber that is greater than or equal to -4 in the range G4:G11 is contained in cell G10 (the seventh cell in the lookup range).

Unlike LOOKUP functions, the MATCH function can work with an inexact match. For example, the formula *MATCH("Pho*",B4:B11,0)* returns 7. The asterisk is treated as a wildcard, which means that Excel searches for the first text string in the range B4:B11 that begins with *Pho*.

If the lookup range is contained in a single row, Excel returns the relative position of the first match in the lookup range, moving left to right. As shown in the following examples, the MATCH function is often very useful when it is combined with other Excel functions, such as VLOOKUP, INDEX, or MAX.

## Given monthly sales for several products, how do I write a formula that will calculate the sales of a product during a given month? For example, how much of product 2 did I sell during June?

The file ProductLookup.xls (shown in Figure 5-2) lists sales of four NBA bobble-head dolls for January through June. How can we write a formula that computes the sales of a given product during a given month? The trick is to use one MATCH function to find the row in which the given product is located and another MATCH function to find the column in which the given month is located. We can then use the INDEX function to return the product sales.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | January | February | March | April | May | June |
| 4 | Shaq | 831 | 685 | 550 | 965 | 842 | 804 |
| 5 | Kobe | 719 | 504 | 965 | 816 | 639 | 814 |
| 6 | MJ | 916 | 906 | 851 | 912 | 964 | 710 |
| 7 | T-Mac | 844 | 509 | 991 | 851 | 742 | 817 |
| 8 | | | | | | | |
| 9 | Product | Month | Row # of product | Column # of month | Product Sales | | |
| 10 | Kobe | June | 2 | 6 | 814 | | |
| 11 | | | | | | | |

**Figure 5-2** The MATCH function can be used in combination with functions such as INDEX and VLOOKUP.

We have named the range B4:G7, which contains sales data for the dolls, as *Sales*. We enter the product we want to know about in cell A10 and the month we're examining in cell B10. In C10, we use the formula *MATCH(A10,A4:A7,0)* to determine which row number in the range Sales contains sales figures for the Kobe doll. Then, in cell D10, we use the formula *MATCH(B10,B3:G3,0)* to determine which column number in the range Sales contains June sales. Now that we have the row and column numbers that contain the sales figures we want, we can use the formula *(INDEX(Sales,C10,D10))* in cell E10 to yield the piece of sales data we want. For more information on the INDEX function, see Chapter 4.

## Given a list of baseball player salaries, how do I write a formula that yields the player with the highest salary? How about the player with the fifth highest salary?

The file Baseball.xls (see Figure 5-3) lists the salaries paid to 401 major league baseball players during the 2001 season. The data is not sorted in order of sal ary, and we want to write a formula that returns the name of the player with the highest salary as well as the name of the player with the fifth highest salary. To find the name of the player with the highest salary we proceed as follows:

- Use the MAX function to determine the value of the highest salary.
- Use the MATCH function to determine the row that contains the player with the highest salary.
- Use a VLOOKUP function (keying off the data row containing the player's salary) to look up the player's name.

We have named the range C12:C412, which includes the players' salaries, as *Salaries*. We've named the range used in our LOOKUP function (range A12:C412) as *Lookup*.

| | A | B | C | D |
|---|---|---|---|---|
| 3 | | | | |
| 4 | | | | |
| 5 | | | | |
| 6 | | name | Alex Rodriguez | dl-Derek Jeter |
| 7 | | | highest | 5th highest |
| 8 | | player position | 345 | 232 |
| 9 | | amount | 22000000 | 12600000 |
| 10 | | | | |
| 11 | | name | salary | |
| 12 | 1 | dl-Mo Vaughn | 13166667 | |
| 13 | 2 | Tim Salmon | 5683013 | |
| 14 | 3 | Garret Anderson | 4500000 | |
| 15 | 4 | Darin Erstad | 3450000 | |
| 16 | 5 | Troy Percival | 3400000 | |
| 17 | 6 | Ismael Valdes | 2500000 | |
| 18 | 7 | Pat Rapp | 2000000 | |
| 19 | 8 | Glenallen Hill | 1500000 | |

**Figure 5-3** This example uses the MAX, MATCH, and VLOOKUP functions to find and display the highest value in a list.

In cell C9, we begin by finding the highest player salary ($22 million) with the formula *MAX(Salaries)*. Next, in cell C8, we use the formula *MATCH(C9,Salaries,0)* to determine the "player number" of the player with the highest salary. We use *match type = 0* because the salaries are not listed in ascending or descending order. We find that player number 345 on the list has the highest salary. Finally, in cell C6, we use the function *VLOOKUP(C8,Lookup,2)* to look up the player's name in the second column of

the lookup range. Not surprisingly, we find that Alex Rodriguez was the highest paid player in 2001.

To find the name of the fifth highest player, we need a function that yields the fifth largest number in an array. The LARGE function does that job. The syntax of the large function is *LARGE(cell range, k)*. When the LARGE function is entered in this fashion, it returns the $k^{th}$ largest number in a cell range. Thus, the formula *LARGE(salaries,5)* in cell D9 yields the fifth largest salary ($12.6 million). Proceeding as before, we find that Derek Jeter is the player with the fifth highest salary. (The *dl* indicates that at the beginning of the season, Jeter was on the disabled list.)

## Given the annual cash flows from an investment project, how do I write a formula that returns the number of years required to pay back the project's initial investment cost?

The file Payback.xls, shown in Figure 5-4, shows the cash flows for an investment project over the next 15 years. We assume that at time 0, the project required a cash outflow of $100 million. During year 1, the project generated a cash inflow of $14 million. We expect cash flows to grow at 10 percent per year. How many years will pass before the project pays back its investment?

The number of years required for a project to payback an investment is called the *payback period.* In high-tech industries, payback period is often used to rank investments. (You'll learn in Chapter 8, however, that payback is flawed as a measure of investment quality because payback ignores the value of money over time.) For now, let's just concentrate on how to determine the payback period for our simple investment model.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Year 1 cf | 14 | | | Payback Period |
| 2 | Growth | 0.1 | | | 6 |
| 3 | Initial investment | 100 | | | |
| 4 | Year | Cash flow | Cum cash flow | | |
| 5 | 0 | -100 | -100 | | |
| 6 | 1 | 14 | -86 | | |
| 7 | 2 | 15.4 | -70.6 | | |
| 8 | 3 | 16.94 | -53.66 | | |
| 9 | 4 | 18.634 | -35.026 | | |
| 10 | 5 | 20.4974 | -14.5286 | | |
| 11 | 6 | 22.54714 | 8.01854 | | |
| 12 | 7 | 24.80185 | 32.820394 | | |
| 13 | 8 | 27.28204 | 60.1024334 | | |
| 14 | 9 | 30.01024 | 90.11267674 | | |
| 15 | 10 | 33.01127 | 123.1239444 | | |
| 16 | 11 | 36.31239 | 159.4363389 | | |
| 17 | 12 | 39.94363 | 199.3799727 | | |
| 18 | 13 | 43.938 | 243.31797 | | |
| 19 | 14 | 48.3318 | 291.649767 | | |
| 20 | 15 | 53.16498 | 344.8147437 | | |

**Figure 5-4** In this example, we'll use the MATCH function to calculate an investment's payback period.

To determine the payback period for the project, we proceed as follows:

- In column B we compute the cash flows for each year.
- In column C we compute the cumulative cash flows (total to date) for each year.
- We use the MATCH function (with *match type =1*) to determine the row number of the first year in which cumulative cash flow is positive. This calculation gives us the payback period.

We gave the cells in B1:B3 the range names listed in A1:A3. Our year 0 cash flow (-Initial_investment) is entered in cell B5. Our year 1 cash flow (Year_1_cf) is entered in cell B6. Copying from B7 to B8:B20 the formula *B6*(1+Growth)* computes the cash flow for years 2-15.

To compute the year 0 cumulative cash flow, we enter the formula *B5* in cell C5. For later years we calculate cumulative cash flow by using a formula such as *Year t cumulative cash flow = Year t – 1 cumulative cash flow + Year t cash flow*. To implement this relationship, simply copy from C6 to C7:C20 the formula *=C5+B6*.

To compute the payback period, we use the MATCH function (with *match type = 1*) to compute the last row of the range C5:C20 containing a value less than 0. This calculation will always give us the payback period. For example, if the last row in C5:C20 that contains a value less than 0 is the sixth row in the range, that means the seventh value marks the cumulative cash flow for the first year the project is paid back. Because our first year is year 0, the payback occurs during year 6. Therefore, the formula in cell E2, *MATCH(0,C5:C20,1)* yields the payback period (6 years). If any cash flows after year 0 were negative, this method would fail because our range of cumulative cash flows would not be listed in ascending order.

# Problems

**1.** Using the distances between U.S. cities given in Chapter 4, write a formula with the MATCH function to determine (based on the names of the cities) the distance between any two of the given cities.

**2.** The file MatchType1.xls lists in chronological order the dollar amount of 30 transactions. Write a formula that yields the first transaction for which total volume to date exceeds $10,000.

**3.** The file MatchTheMax.xls gives the product ID code and unit sales for 265 different products. Use the MATCH function in a formula that yields the product ID of the product with the largest unit sales.

# 6

# Text Functions

- I have a spreadsheet in which each cell contains a product description, a product ID, and a product price. How can I put all the product descriptions in column A, all the product IDs in column B, and all the prices in column C?
- Every day I receive data about total U.S. sales, which is computed in a cell as the sum of East, North, and South region sales. How can I extract East, North, and South sales to separate cells?
- I download quarterly gross national product (GNP) data from the Web. The cell containing first quarter data for 1980 contains the entry 1980.1  5028.8. How can I place the date and GNP value in different cells?
- In the spreadsheet I use for a mailing list, column A contains people's names, column B contains their street address, and column C contains their city and zip code. How can I create each person's full address in column E?

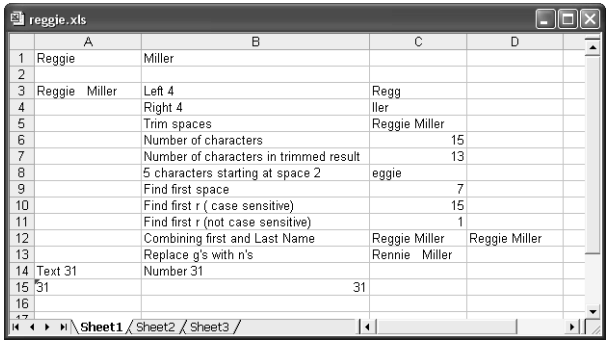
When someone sends you data or you download data from the Web, often the data isn't formatted the way you want. For example, when downloading sales data, dates and sales amounts might be in the same cell but you need them to be in separate cells. How can you manipulate data so that it has the format you need? The answer is to become good at using Excel's set of text functions**.** In this chapter, I'll show you how to use the following

Excel text functions to magically manipulate your data until it looks the way you want:

- LEFT
- RIGHT
- MID
- TRIM
- LEN
- FIND
- SEARCH
- CONCATENATE
- REPLACE
- VALUE

# Text Function Syntax

The file Reggie.xls, shown in Figure 6-1, includes examples of text functions. You'll see how to apply these functions to a specific problem later in the chapter, but let's begin by describing what each of the text functions do. Then we'll combine these functions to perform some fairly complex manipulations of data.


reggie.xls

| | A | B | C | D |
|---|---|---|---|---|
| 1 | Reggie | Miller | | |
| 2 | | | | |
| 3 | Reggie Miller | Left 4 | Regg | |
| 4 | | Right 4 | ller | |
| 5 | | Trim spaces | Reggie Miller | |
| 6 | | Number of characters | 15 | |
| 7 | | Number of characters in trimmed result | 13 | |
| 8 | | 5 characters starting at space 2 | eggie | |
| 9 | | Find first space | 7 | |
| 10 | | Find first r ( case sensitive) | 15 | |
| 11 | | Find first r (not case sensitive) | 1 | |
| 12 | | Combining first and Last Name | Reggie Miller | Reggie Miller |
| 13 | | Replace g's with n's | Rennie Miller | |
| 14 | Text 31 | Number 31 | | |
| 15 | 31 | 31 | | |
| 16 | | | | |

Sheet1 / Sheet2 / Sheet3 /

**Figure 6-1** Examples of text functions.

## The LEFT Function

The function *LEFT(text,k)* returns the first *k* characters in a text string. For example, cell C3 contains the formula *LEFT(A3,4)*. Excel returns *Regg*.

## The RIGHT Function

The function *RIGHT(text,k)* returns the last *k* characters in a text string. For example, in cell C4, the formula *RIGHT(A3,4)* returns *ller*.

## The MID Function

The function *MID(text, k, m)* begins at character *k* of a text string and returns the next *m* characters. For example, the formula *MID(A3,2,5)* in cell C8 returns characters 2-6 from cell A3, the result being *eggie*.

## The TRIM Function

The function *TRIM(text)* removes all spaces from a text string except for single spaces between words. For example, in cell C5 the formula *TRIM(A3)* eliminates two of the three spaces between Reggie and Miller and yields *Reggie Miller*.

## The LEN Function

The function *LEN(text)* returns the number of characters in a text string (spaces are included). For example, in cell C6 the formula *LEN(A3)* returns 15 because cell A3 contains 15 characters. In cell C7, the formula *LEN(C5)* returns 13. Because the trimmed result in cell C5 has two spaces removed, cell C5 contains two less characters than the original text in A3.

## The FIND and SEARCH Functions

The function *FIND(text to find, actual text, k)* returns the location at or after character *k* of the first character of *text to find* in *actual text*. *FIND* is case sensitive. *SEARCH* has the same syntax as *FIND*, but it is not case sensitive. For example, if we enter *FIND ("r",A3,1)* in cell C10, Excel returns 15, the location of the first lowercase *r* in the text string *Reggie Miller*. (The uppercase *R* is ignored because *FIND* is case sensitive.) Entering *SEARCH("r",A3,1)* in cell C11 returns 1 because *SEARCH* matches *r* to either a lowercase or an uppercase character. Entering *FIND(" ",A3,1)* in cell C9 returns 7 because the first space in the string *Reggie Miller* is the seventh character.

## CONCATENATE and & Functions

The function *CONCATENATE(text1, text2, . . ., text30)* can be used to join up to 30 text strings into a single string. The & operator can be used instead of *CONCATENATE*. For example, entering in cell C12 the formula *A1&" "&B1* returns *Reggie Miller*. Entering in cell D12 the formula *CONCATENATE(A1," ",B1)* yields the same result.

## REPLACE Function

The function *REPLACE(oldtext, k,m, newtext)* begins at character *k* of *oldtext* and replaces the next *m* characters with *newtext*. For example, in cell C13 the formula *REPLACE(A3,3,2,"nn")* replaces the third and fourth characters (*gg*) in cell A3 with *nn*. This formula yields *Rennie Miller*.

## VALUE Function

The function *VALUE(text)* converts a text string that represents a number to a number. For example, entering in cell B15 the formula *VALUE(A15)* converts the text string *31* in cell A15 to the numerical value 31. You can tell that the value 31 in cell A15 is text because it is left-justified. Similarly, you can tell that the value 31 in cell B15 is a number because it is right-justified.

# Text Functions in Action

You can see the power of text functions by using them to solve some actual problems that were sent to me by former students working in Fortune 500 corporations. The key to solving each problem is to combine multiple text functions in a single formula.

### I have a spreadsheet in which each cell contains a product description, a product ID, and a product price. How can I put all the product descriptions in column A, all the product IDs in column B, and all the prices in column C?

In this example, the product ID is always defined by the first 12 characters and the price is always indicated in the last 8 characters (with two spaces following the end of each price). Our solution, contained in the file Lenora.xls and shown in Figure 6-2, uses the LEFT, RIGHT, MID, VALUE, TRIM, and LEN functions.

It's always a good idea to begin by trimming excess spaces, which we do by copying from B4 to B5:B12 the formula *TRIM(A4)*. The only excess spaces in

column A turn out to be the two spaces inserted after each price. The results of using the TRIM function are shown in Figure 6-2.

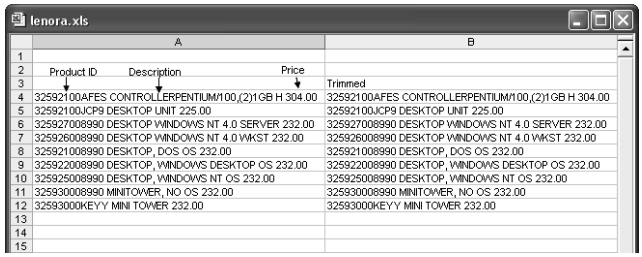



**Figure 6-2** Use the TRIM function to trim away excess spaces.

To capture the product ID, we need to extract the 12 leftmost characters from column B. We copy from C4 to C5:C13 the formula *LEFT(B4,12)*. This formula extracts the 12 leftmost characters from the text in cell B4 and the following cells, yielding the product ID, as you can see Figure 6-3.

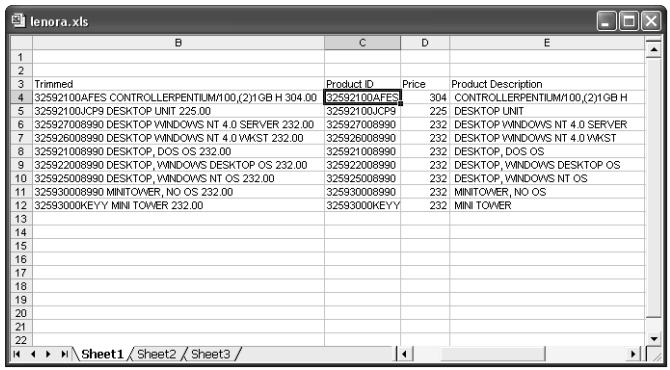



**Figure 6-3** Text functions extract the product ID, price, and product description from a single text string.

To extract the product price, we first note that the price occupies the last 6 digits of each cell, so we need to extract the rightmost 6 characters in each cell. I copy from cell D4 to D5:D12 the formula *VALUE(RIGHT(B4,6)*. I use the *Value* function to turn the extracted text into a numerical value. Without converting the text to a numerical value, you couldn't perform mathematical operations on the prices.

Extracting the product description is much trickier. By examining the data, we can see that if we begin our extraction with the 13th character and continue until we are 6 characters from the end of the cell, we will have the data we want. Copying the following formula from E4 to E5:E12 does the job: *MID(B4,13,LEN(B4)-6-12)*. *Len(B4)* returns the total number of characters in

the trimmed text. This formula (*MID* for Middle) begins with the 13th character and then extracts the number of characters equal to the total number less the 12 characters at the beginning (the product ID) and the 6 characters at the end (price). This subtraction simply leaves the product description!

Now suppose we are given the data with product ID in column C, the price in column D, and the product description in column E. Can we put these values together to recover our original text?

Text can easily be combined by using the CONCATENATE formula. Copying from F4 to F5:F12 the formula *CONCATENATE(C4," ",E4," ",D4)* recovers our original (trimmed) text, which you can see in Figure 6-4.

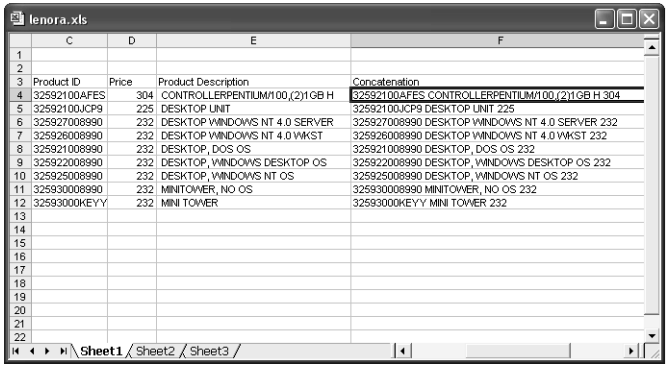



**Figure 6-4** Using concatenation to recombine product ID, product description, and price.

The concatenation formula starts with the product ID in cell C4. The empty quotation marks (" ") insert a space. Next we add the product description from cell E4 and another space. Finally, we add the price from cell D4. We have now recovered the entire text describing each computer! Concatenation can also be performed by using the & sign. We could also recover the original product ID, product description, and price in a single cell with the formula *C4&" "&E4&" "&D4*.

If the product IDs did not always contain 12 characters, this method of extracting the information would fail. We could, however, extract the product IDs by using the FIND function to discover the location of the first space. Then we could obtain the product ID by using the LEFT function to extract all characters to the left of the first space. The example in the next section will show how this approach works.

If the price did not always contain precisely six characters, extracting the price would be a little tricky. You would need to use the OFFSET function, which you will learn about in Chapter 21.

## Every day I receive data about total U.S. sales, which is computed in a cell as the sum of East, North, and South region sales. How can I extract East, North, and South sales to separate cells?

This problem was sent to me by an employee in Microsoft's finance department. She received a spreadsheet each day containing formulas such as *=50+200+400, =5+124+1025*, and so on. She needed to extract each number into a cell in its own column. For example, she wanted to extract the first number (East sales) in each cell to column C, the second number (North sales) to column D, and the third number (South sales) to column E. What makes this problem challenging is that we don't know the exact location of the character at which the second and third numbers start in each cell. In the first cell in our example, the second number begins with the fourth character. In the second cell, the second number begins with the third character. The data we're using in this example is in the file SalesStripping.xls, shown in Figure 6-5. By combining the FIND, LEFT, LEN, and MID functions, we can easily solve this problem as follows:

- Use the Edit, Replace command to replace each equal sign with a space. This converts each formula into text.
- Use the FIND function to locate the two plus (+) signs in each cell.
- East sales are represented by every character to the left of the first plus sign.
- North sales are represented by every character between the first and second plus sign.
- South sales are represented by every character to right of the second plus sign.

salesstripping.xls

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Extracting Sales in Three Regions | | | | | | | | |
| 2 | East+North+South | First + | Second + | East | North | Total Length | South | | |
| 3 | 10+300+400 | 3 | 7 | 10 | 300 | 10 | 400 | | |
| 4 | 4+36.2+800 | 2 | 7 | 4 | 36.2 | 10 | 800 | | |
| 5 | 3+23+4005 | 2 | 5 | 3 | 23 | 9 | 4005 | | |
| 6 | 18+1+57.31 | 3 | 5 | 18 | 1 | 10 | 57.31 | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |

**Figure 6-5** Extracting East, North, and South sales with a combination of the FIND, LEFT, LEN, and MID functions.

We begin by finding the location of the first plus sign for each piece of data. By copying from B3 to B4:B6 the formula *FIND("+",A3,1)*, we can locate the first plus sign for each data point. To find the second plus sign, we begin

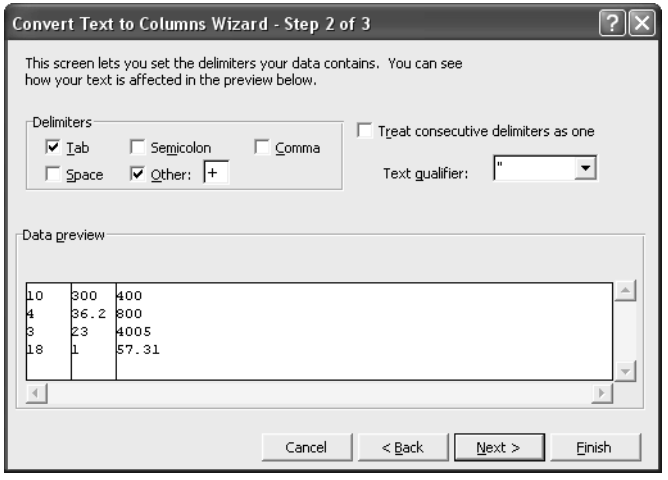
*one character* after the first plus sign, copying from C3 to C4:C6 the formula *FIND("+",A3,B3+1)*.

To find East sales, we use the LEFT function to extract all the characters to the left of the first plus sign, copying from D3 to D4:D6 the formula *LEFT(A3,B3-1)*. To extract the North sales, we use the MID function to extract all the characters between the two plus signs. We begin one character after the first plus sign and extract the number of characters equal to *(Position of 2nd plus sign)-(Position of 1st plus sign) – 1*. If you leave out the –1, you'll get the second + sign. (Go ahead and check this.) So, to get the North sales, we copy from E3 to E4:E6 the formula *MID(A3,B3+1,C3-B3-1)*.

To extract South sales, we use the RIGHT function to extract all the characters to the right of the second plus sign. South sales will have the number of characters equal to *(Total characters in cell) – (Position of 2nd plus sign)*. We compute the total number of characters in each cell by copying from F3 to F4:F6 the formula *LEN(A3)*. Finally, we obtain South sales by copying from G3 to G4:G6 the formula *RIGHT(A3,F3-C3)*.

## Using the Text To Columns Command to Extract Data

There is an easy way to extract East, North, and South sales (and data similar to this example) without using text functions. Simply select cells A3:A6, and then choose Text To Columns on the Data menu. Select Delimited, click Next, and then fill in the dialog box shown here:



Entering the plus sign in the Delimiters area directs Excel to separate each cell into columns, breaking at each occurrence of the plus sign. The result is shown here.

salesstripping.xls

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | Extracting Sales in Three Regions | | | | |
| 2 | Data Text to Columns Results | | | | |
| 3 | 10 | 300 | 400 | | |
| 4 | 4 | 36.2 | 800 | | |
| 5 | 3 | 23 | 4005 | | |
| 6 | 18 | 1 | 57.31 | | |
| 7 | | | | | |

# Problems

**1.** Cells B2:B5 of the workbook Showbiz.xls contain the fictitious addresses of some of our favorite people, as shown here. Use text functions to extract each person's name to a single column and each person's street address to a single column.

| | A | B |
|---|---|---|
| 1 | | |
| 2 | | Ally McBeal 1200 Lawyer Drive |
| 3 | | Drew Carey 1000 Hollywood Lane |
| 4 | | Britney Spears 300 Singer Road |
| 5 | | Peyton Manning 500 QB Street |

**2.** The workbook IDPrice.xls contains the product ID and prices for various products. Use text functions to put the product IDs and prices in separate columns. Use the Text To Columns command on the Data menu to accomplish the same goal.

| | A | B |
|---|---|---|
| 1 | | |
| 2 | | ID and Price |
| 3 | | xa1 304 |
| 4 | | za23 23 |
| 5 | | xa13 4123 |
| 6 | | zzx 12 |
| 7 | | a12q 374 |

**3.** The workbook QuarterlyGnpData.xls contains quarterly GNP data for the United States (in billions of 1996 dollars) in the format shown here. Extract this data to three separate columns, where the first column contains the year, the second column contains the quarter number, and the third column contains the GNP value.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | http://forecasts.org/data/index.htm | | | | |
| 2 | | | | | |
| 3 | Real Gross National Product | | | | |
| 4 | | | | | |
| 5 | Billions of Chained 1996 Dollars, Seasonally Adjusted Annual Rate | | | | |
| 6 | Source: U.S. Department of Commerce, Bureau of Economic Analysis | | | | |
| 7 | | | | | |
| 8 | DATE GNPC96 | | | | |
| 9 | | | | | |
| 10 | 1950.1 1618.4 | | | | |
| 11 | 1950.2 1667.2 | | | | |
| 12 | 1950.3 1733.1 | | | | |
| 13 | 1950.4 1763.9 | | | | |
| 14 | 1951.1 1782.9 | | | | |
| 15 | 1951.2 1814.9 | | | | |
| 16 | 1951.3 1851.6 | | | | |

**4.** The file TextStylesData.xls contains information about the size, color, and style for a variety of shirts. For example, the first shirt is style 100 (indicated by digits between the colon and the hyphen). Its color is 65, and its size is L. Use text functions to extract the style, color, and size of each shirt.

| | A |
|---|---|
| 1 | **Item** |
| 2 | 100's:100-65L |
| 3 | 100's:100-65XL |
| 4 | 100's:100-65XXL |
| 5 | 100's:100-06M |
| 6 | 100's:100-06L |
| 7 | 100's:100-06XL |
| 8 | 100's:100-06XXL |
| 9 | 100's:100-05M |
| 10 | 100's:100-05L |
| 11 | 100's:100-05XL |
| 12 | 100's:100-05XXL |

# 7

# Dates and Date Functions

- When I enter dates into Excel, I often see a number such as 37625 rather than a date such as 1/4/2003. What does this number mean, and how do I change 37625 to a normal date?
- Can I enter a formula that automatically displays today's date?
- How do I determine a date that is 50 workdays after another date? What if I want to exclude holidays?
- How do I determine the number of workdays between two dates?
- I have 500 different dates entered in Excel. What formulas can I use to extract the month, year, day of the month, and day of the week from each date?

To illustrate the most commonly used month-day-year formats in Excel, suppose today is January 4, 2004. We could enter this date as any of the following:

- 1/4/2004
- 4-Jan-2004
- January 4, 2004
- 1/4/04

> **Note** If you enter a year with only two digits, for years entered as 30 or higher, Excel assumes the digits represent years in the twentieth century. For example, 1/1/29 is treated as January 1, 2029, but 1/1/30 is treated as January 1, 1930. Each year, the year treated as dates in the twentieth century increases by one.

If you want to walk through the explanations for the problems in this chapter in Excel, open the file Date.xls. (Be sure you open the correct file. There is a similarly named file named Dates.xls, which is used in the next section.)

### When I enter dates into Excel, I often see a number such as 37625 rather than a date such as 1/4/2003. What does this number mean, and how do I change 37625 to a normal date?

The way Excel treats calendar dates is sometimes confusing to the novice. The key is understanding that Excel can display a date in a variety of month-day-year formats, or it can display a date in *serial format*. A date in serial format, such as 37625, is simply a positive integer that represents the number of days between the given date and January 1, 1900. Both the current date and January 1, 1900, are included in the count. For example, Excel displays January 3, 1900, in serial format as the number 3, which means there are three days between January 1, 1900, and January 3, 1900 (including both days).

> **Note** Excel assumes that 1900 was a leap year containing 366 days. In reality, 1900 contained only 365 days.

Figure 7-1 shows the worksheet named Serial Format in the file Dates.xls. Suppose you are given the dates shown in cells D5:D14 in serial format. For example, the value 37622 in cell D5 indicates a date that is 37,622 days after January 1, 1900 (including both January 1, 1900, and the current day). To display these dates in month-day-year format, copy them to E5:E14. Select the cell range E5:E14, choose Format Cells, and then select the date format you want from the list shown in Figure 7-2. The dates in E5:E14 will now be displayed in date format, as you can see in Figure 7-1. If you want to format dates in the serial number format, select E5:E14 and then choose Format Cells General.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | Dates | Reformatted | |
| 5 | | | | 37622 | 1/1/2003 | |
| 6 | | | | 37623 | 1/2/2003 | |
| 7 | | | | 37624 | 1/3/2003 | |
| 8 | | | | 37625 | 1/4/2003 | |
| 9 | | | | 37626 | 1/5/2003 | |
| 10 | | | | 37627 | 1/6/2003 | |
| 11 | | | | 37628 | 1/7/2003 | |
| 12 | | | | 37629 | 1/8/2003 | |
| 13 | | | | 37630 | 1/9/2003 | |
| 14 | | | | 37631 | 1/10/2003 | |

**Figure 7-1** Use the Format Cells command to change dates from serial number format to month-day-year format.

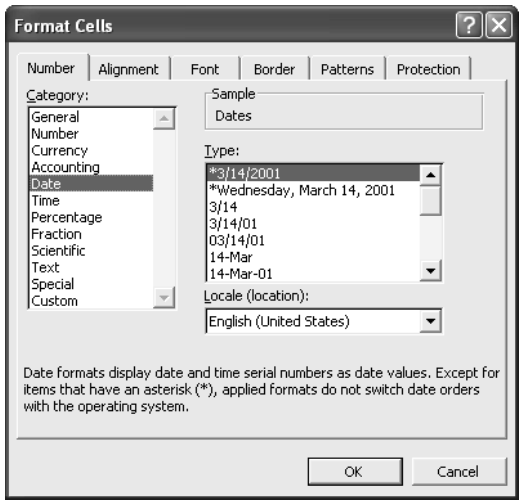



**Figure 7-2** Reformatting a serial number to month-day-year format.

As I mentioned previously, simply changing the format of a cell in a date format to General will yield the date in serial format. Another way to obtain the date in serial format is to use the DATEVALUE function. Enclose any date in quotation marks, and this function will return the date's serial number. For example, cell I5 in the worksheet named Date Format contains the formula *DATEVALUE("1/4/2003")*. Excel yields 37625, which is the serial format for January 4, 2003.

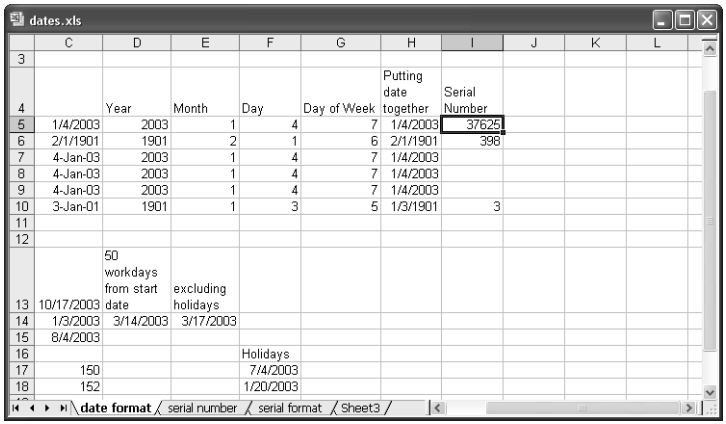



**Figure 7-3** You can use the DATEVALUE function to format a date in serial format.

## Can I enter a formula that automatically displays today's date?

Displaying today's date with a formula is easy, as you can see by looking at the worksheet Date Format, shown in Figure 7-4. Entering the *TODAY()* function in a cell will display today's date. Of course, whenever you open the workbook, the cell displays the current date. If you update a worksheet every day and want the current date to be displayed, use *TODAY()*.

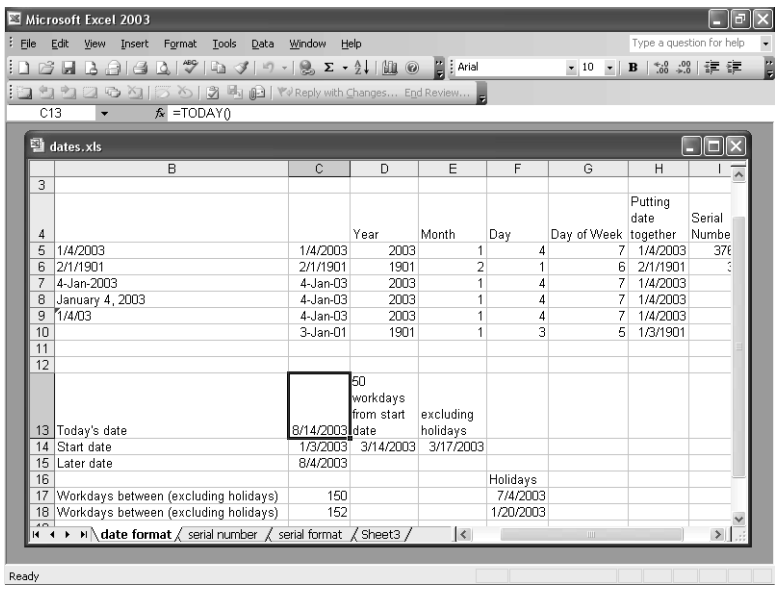



**Figure 7-4** Use the TODAY function to display today's date. Use the WORKDAY and NETWORKDAYS functions to calculate dates a certain number of days from today.

## How do I determine a date that is 50 workdays after another date? What if I want to exclude holidays?

The function *WORKDAY(start_date, #days, [holidays])* displays the date that is the number of workdays indicated in *#days* (a workday is a nonweekend day) after a given start date. *Holidays* is an optional argument for the function that allows you to exclude from the calculation any dates that are listed in a cell range. Thus, entering the formula *WORKDAY(C14,50)* in cell D14 of the Date Format worksheet tells us that 3/14/2003 is 50 workdays (Monday–Friday dates) after 01/03/2003. If we believe that the only two holidays that matter are Martin Luther King Day and the 4th of July, we can change the formula to *WORKDAY(C14,50,F17:F18)*. With this addition, Excel does not count 01/20/2003 in its calculations, making 03/17/2003 the 50th workday after 01/03/2003.

> **Important** To use the WORKDAY function, you must have the Analysis Toolpak installed. To install the toolpak, run Tools Add-Ins and select Analysis ToolPak and Analysis ToolPak VBA.

## How do I determine the number of workdays between two dates?

The key to solving this problem is to use the NETWORKDAYS function. The syntax for this function is *NETWORKDAYS(start_date, end_date, [holidays])*, where *holidays* is an optional argument identifying a cell range that lists the dates you want to count as holidays. The NETWORKDAYS function returns the number of working days between *start_date* and *end_date* excluding weekends and any listed holidays. As an illustration of the NETWORKDAYS function, look at cell C18 in the Date Format worksheet, which contains the formula *NETWORKDAYS(C14,C15)*. This formula yields the number of working days between 1/3/2003 and 8/4/2003, which turns out to be 152. The formula *NETWORKDAYS(C14,C15,F17:F18)*, which is entered in cell C17 of the worksheet Date Format, yields the number of workdays between 1/3/2003 and 8/4/2003 excluding Martin Luther King Day and the 4th of July. The answer is 152 -2 = 150. The NETWORKDAYS function also requires installation of the Analysis Toolpak.

### I have 500 different dates entered in an Excel worksheet. How do I write formulas that will extract from each date the month, year, day of the month, and day of the week?

The worksheet Date Format (see Figure 7-4) lists several dates in the cell range C5:C10. In C5 and C7:C9, I've used four different formats to display January 4, 2003. In columns D:G, I've extracted the year, month, day of the month, and day of the week for each date. By copying from D5 to D6:D9 the formula *YEAR(C5)*, we extract the year for each date. By copying from E5 to E6:E9 the formula *MONTH(C5)*, we extract the month (1 = January, 2 = February, and so on) part of each date. By copying from F5 to F6:F9 the formula *DAY(C5)*, we extract the day of the month for each date. Finally, by copying from G5 to G6:G9 the formula *WEEKDAY(C5,1)*, we extract the day of the week for each date.

> **Note** When the last argument of the Weekday function is 1, 1 = Sunday, 2 = Monday, and so on. When the last argument is 2, 1 = Monday, 2 = Tuesday, and so on. When the last argument is 3, 0 = Monday, 1 = Tuesday, and so on.

Suppose you are given the year, month, and day of the month for a date. Is there an easy way to recover the actual date? The DATE function, whose arguments are *DATE(year,month,day)*, will return the date with the given year, month, and day of the month. In the worksheet Date Format, copying from cell H5 to cells H6:H9 the formula *DATE(D5,E5,F5)* recovers the dates we started with.

# Problems

1. What is the serial format for January 25, 2006?
2. What is the serial format for February 14, 1950?
3. To what actual date does a serial format of 4526 correspond?
4. To what actual date does a serial format of 45000 correspond?
5. Determine the day that occurs 74 workdays after today's date (including holidays).
6. Answer question 5 but exclude the current year's Christmas, New Year's Day, and the 4th of July.
7. How many workdays (including holidays) are there between July 10, 2005, and August 15, 2006?
8. Answer question 7 but exclude Christmas, New Year's Day, and the 4th of July.

File DateP.xls contains several hundred dates. Use this file for the next set of problems.

1. Determine the month, year, day of the month, and day of the week for each date.
2. Express each date in serial format.
3. A project will begin on December 4, 2005. The project consists of three activities: Activities 1, 2, and 3. Activity 2 can start the day after Activity 1 finishes. Activity 3 can start the day after Activity 2 finishes. Set up a spreadsheet that will accept as inputs the duration (in days) of the three activities and output both the month and year during which each activity is completed.

# 12

# IF Statements

- If I order up to 500 units of a product, I pay $3.00 per unit. If I order between 501 and 1200 units, I pay $2.70 per unit. If I order between 1201 and 2000 units, I pay $2.30 per unit. If I order more than 2000 units, I pay $2.00 per unit. How can I write a formula that expresses the purchase cost as a function of the number of units purchased?
- I've just purchased 100 shares of stock at a cost of $55 per share. To hedge the risk that the stock might decline in value, I purchased 60 six-month European put options. Each option has an exercise price of $45 and costs $5. How can I develop a spreadsheet that indicates the six-month percentage return on my portfolio for a variety of possible future prices?
- Many stock market analysts believe that moving-average trading rules can outperform the market. A commonly suggested moving-average trading rule is to buy a stock when the stock's price moves above the average of the last 15 months and to sell a stock when the stock's price moves below the average of the last 15 months' price. How would this trading rule have performed against buying and holding the Standard and Poor's Index?
- In the game of craps, two dice are tossed. If the total of the dice on the first roll is 2, 3, or 12, you lose. If the total of the dice on the first roll is 7 or 11, you win. Otherwise, the game keeps going. How can I write a formula to determine the status of the game after the first roll?
- In most pro forma financial statements, cash is used as the plug to make assets and liabilities balance. I know that using debt as the plug would be more realistic. How can I set up a pro forma statement having debt as the plug?

The five situations listed above seem to have little, if anything, in common. However, setting up Excel models for each of these situations requires the use of an IF statement. I believe that the IF formula is the single most useful formula in Excel. IF formulas let you conduct conditional tests on values and formulas, mimicking (to a limited degree) the conditional logic provided by computing languages such as C, C++, and Java.

An IF formula begins with a condition, for example, *A1>10*. If the condition is true, the formula returns the first value listed in the formula; otherwise, we move on within the formula and repeat the process. The easiest way to show you the power and utility of IF formulas is to use them to help answer each of our five questions.

### If I order up to 500 units of a product, I pay $3.00 per unit. If I order between 501 and 1200 units, I pay $2.70 per unit. If I order between 1201 and 2000 units, I pay $2.30 per unit. If I order more than 2000 units, I pay $2.00 per unit. How can I write a formula that expresses the purchase cost as a function of the number of units purchased?

You can find the solution to this question on the worksheet named Quantity Discount in the file IfStatement.xls. The worksheet is shown in Figure 12-1.

ifstatement.xls

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | cutoff | price | | |
| 2 | cut1 | 500 | $ 3.00 | price1 | |
| 3 | cut2 | 1200 | $ 2.70 | price2 | |
| 4 | cut3 | 2000 | $ 2.30 | price3 | |
| 5 | | | $ 2.00 | price4 | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | order quantity | cost | per unit cost | | |
| 9 | 450 | $1,350.00 | $ 3.00 | | |
| 10 | 900 | $2,430.00 | $ 2.70 | | |
| 11 | 1450 | $3,335.00 | $ 2.30 | | |
| 12 | 2100 | $4,200.00 | $ 2.00 | | |
| 13 | | | | | |

**Figure 12-1** You can use an IF formula to model quantity discounts.

Suppose cell A9 contains our order quantity. We can compute an order's cost as a function of the order quantity by implementing the following logic:

- If A9<=500, the cost is 3*A9.
- If A9 is between 501 and 1200, the cost is 2.70*A9.
- If A9 is between 1201 and 2000, the cost is 2.30*A9.
- If A9 is more than 2000, the cost is 2*A9.

We begin by linking the range names in A2:A4 to cells B2:B4, and linking the range names in cells D2:D5 to cells C2:C5. We can implement this logic in cell B9 with the formula:

```
IF(A9<=cut1,price1*A9,IF(A9<=cut2,price2*A9,IF(A9<=cut3,price3*A9,price4*A9)))
```

To understand how Excel computes a value from this formula, recall that IF statements are evaluated from left to right. If the order quantity is less than or equal to 500 (*cut1*), the cost is given by *price1*A9*. If the order quantity is not less than 500, the formula checks to see whether the order quantity is less than 1200. If this is the case, the order quantity is between 501 and 1200, and the formula computes a cost of *price2*A9*. Next we check whether the order quantity is less than or equal to 2000. If this is true, the order quantity is between 1201 and 2000, and our formula computes a cost of *price3*A9*. Finally, if the order cost has not yet been computed, our formula defaults to the value *price4*A9*. In each case, the IF formula returns the correct order cost. Note that I typed in three more order quantities in cells A10:A12 and copied our cost formula to B10:B12. For each order quantity, our formula returns the correct total cost.

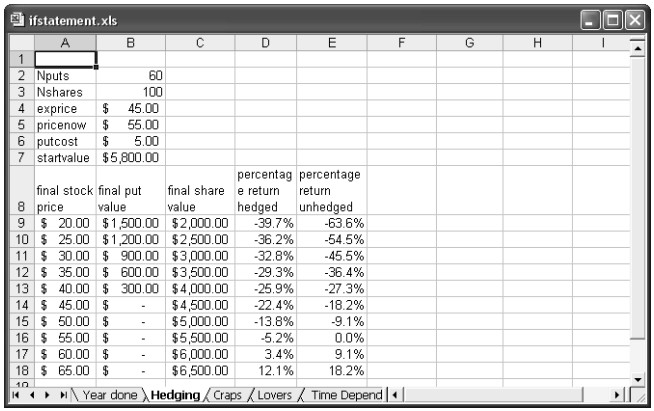
An IF formula containing more than one IF statement is called a *nested IF formula*. A nested IF formula can contain up to seven IF statements and up to 256 characters. More complex formulas need to be broken into multiple cells.

## I've just purchased 100 shares of stock at a cost of $55 per share. To hedge the risk that the stock might decline in value, I purchased 60 six-month European put options. Each option has an exercise price of $45 and costs $5. How can I develop a spreadsheet that indicates the six-month percentage return on my portfolio for a variety of possible future prices?

Before tackling this problem, I want to review some basic concepts from the world of finance. A European option allows you to sell at a given time in the future (in this case six months) a share of a stock for the exercise price (in this case $45). If our stock's price in 6 months is $45 or higher, the option has no value. Suppose, however, that the price of the stock in 6 months is below $45. Then you can make money by buying a share and immediately reselling the stock for $45. For example, if in 6 months our stock is selling for $37, you can make a profit of $45 - $37, or $8 per share, by buying a share for $37 and then using the put to resell the share for $45. You can see that put options protect you against downward moves in a stock price. In this case, whenever the stock's price in six months is below $45, the puts start kicking in some value. This cushions a portfolio against a decrease in value of the shares it owns. Note also that the percentage return on a portfolio is computed by taking the change

in the portfolio's value (*final portfolio value – initial portfolio value*) and dividing that number by the portfolio's initial value.

With this background, let's look at how the six-month percentage return on our portfolio, consisting of 60 puts and 100 shares of our stock, varies as the share price varies between $20 and $65. You can find this solution on the sheet named Hedging in the file IfStatement.xls. The worksheet is shown in Figure 12-2.



ifstatement.xls

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | Nputs | 60 | | | | | | | |
| 3 | Nshares | 100 | | | | | | | |
| 4 | exprice | $ 45.00 | | | | | | | |
| 5 | pricenow | $ 55.00 | | | | | | | |
| 6 | putcost | $ 5.00 | | | | | | | |
| 7 | startvalue | $5,800.00 | | | | | | | |
| 8 | final stock price | final put value | final share value | percentage return hedged | percentage return unhedged | | | | |
| 9 | $ 20.00 | $1,500.00 | $2,000.00 | -39.7% | -63.6% | | | | |
| 10 | $ 25.00 | $1,200.00 | $2,500.00 | -36.2% | -54.5% | | | | |
| 11 | $ 30.00 | $ 900.00 | $3,000.00 | -32.8% | -45.5% | | | | |
| 12 | $ 35.00 | $ 600.00 | $3,500.00 | -29.3% | -36.4% | | | | |
| 13 | $ 40.00 | $ 300.00 | $4,000.00 | -25.9% | -27.3% | | | | |
| 14 | $ 45.00 | $ - | $4,500.00 | -22.4% | -18.2% | | | | |
| 15 | $ 50.00 | $ - | $5,000.00 | -13.8% | -9.1% | | | | |
| 16 | $ 55.00 | $ - | $5,500.00 | -5.2% | 0.0% | | | | |
| 17 | $ 60.00 | $ - | $6,000.00 | 3.4% | 9.1% | | | | |
| 18 | $ 65.00 | $ - | $6,500.00 | 12.1% | 18.2% | | | | |

Year done | Hedging | Craps | Lovers | Time Depend

**Figure 12-2** Hedging example that uses IF statements.

The labels in A2:A7 are linked to the cells B2:B7. The initial portfolio value is equal to 100($55) + 60($5) = $5,800, shown in cell B7. By copying from B9 to B10:B18 the formula *IF(A9<exprice,exprice-A9,0)*Nputs*, we compute the final value of our puts. If the six-month price is less than our exercise price, we value each put as *exercise price-six-month price*. Otherwise, each put will in six months have a value of $0. Copying from C9 to C10:C18 the formula *Nshares*A9*, we compute the final value of our shares. Copying from D9 to D10:D18 the formula *((C9+B9)-startvalue)/startvalue)* computes the percentage return on our hedged portfolio. Copying from E9 to E10:E18 the formula *(C9-Nshares*pricenow)/(Nshares*pricenow)* computes the percentage return on our portfolio if we are unhedged (that is, buy no puts).

In Figure 12-2, you can see that if the stock price drops below $45, our hedged portfolio has a larger expected return than our unhedged portfolio. Also note that if the stock price does not decrease, the unhedged portfolio has a larger expected return. This is why the purchase of puts is often referred to as *portfolio insurance*.

### Many stock market analysts believe that moving-average trading rules can outperform the market. A commonly suggested moving-average trading rule is to buy a stock when the stock's price moves above the average of the last 15 months and to sell a stock when the stock's price moves below the average of the last 15 months' price. How would this trading rule have performed against buying and holding the Standard and Poor's Index?

In this example, we'll compare the performance of the moving-average trading rule (in the absence of transaction costs for buying and selling stock) to a buy-and-hold strategy. The strength of a moving-average trading rule is that it helps you follow market trends. A moving-average trading rule lets you ride up with a bull market and sell before a bear market destroys you. Our data set contains the monthly value of the Standard and Poor's (S&P) Index for the time period January 1871 through October 2002. To track the performance of our moving-average trading strategy, we need to track the following information each month:

- The average of the S&P Index over the last 15 months.
- Do we own stock at the beginning of each month?
- Do we buy stock during the month?
- Do we sell stock during the month?
- What is our cash flow for the month (positive if we sell stock, negative if we buy stock, and 0 otherwise)?

The file MaTradingRule.xls, shown in Figure 12-3, includes the work I did to answer this question. Tackling this problem requires several IF formulas, and some of the IF formulas will require an AND operator. For example, we'll buy the stock during a month if and only if we don't own the stock at the beginning of the month and the current month's price is larger than the 15-month moving average for the stock's price. The first month for which we can compute a 15-month moving average is April 1872, so we begin our calculations in row 24.

matradingrule.xls

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 6 | | S&P | | | | MA profit | $1,319.75 | |
| 7 | | Comp. | | | | Buy and hold profit | $849.45 | |
| 8 | Date | P | Own? | MA | Buy? | Sell? | Cash flow | |
| 24 | 1872.04 | 5.18 | Yes | 4.739 | no | no | 0 | |
| 25 | 1872.05 | 5.18 | Yes | 4.788 | no | no | 0 | |
| 26 | 1872.06 | 5.13 | Yes | 4.833 | no | no | 0 | |
| 27 | 1872.07 | 5.1 | Yes | 4.868 | no | no | 0 | |
| 28 | 1872.08 | 5.04 | Yes | 4.892 | no | no | 0 | |
| 29 | 1872.09 | 4.95 | Yes | 4.904 | no | no | 0 | |
| 30 | 1872.1 | 4.97 | Yes | 4.913 | no | no | 0 | |
| 31 | 1872.11 | 4.95 | Yes | 4.929 | no | no | 0 | |
| 32 | 1872.12 | 5.07 | Yes | 4.939 | no | no | 0 | |
| 33 | 1873.01 | 5.11 | Yes | 4.955 | no | no | 0 | |
| 34 | 1873.02 | 5.15 | Yes | 4.989 | no | no | 0 | |
| 35 | 1873.03 | 5.11 | Yes | 5.023 | no | no | 0 | |
| 36 | 1873.04 | 5.04 | Yes | 5.048 | no | yes | 5.04 | |
| 37 | 1873.05 | 5.05 | no | 5.06 | no | no | 0 | |
| 38 | 1873.06 | 4.98 | no | 5.071 | no | no | 0 | |
| 39 | 1873.07 | 4.97 | no | 5.067 | no | no | 0 | |
| 40 | 1873.08 | 4.97 | no | 5.053 | no | no | 0 | |

Stock Data

**Figure 12-3** Moving-average trading rule beats buy and hold!

- Let's assume we first owned the stock in April 1872, so we entered Yes in cell C24.
- By copying from D24 to D25:D1590 the formula *AVERAGE(B9:B23)*, we compute the 15-month moving average for each month.

> **Tip** An easy way to copy the formula from D24 to D25:D1590 is to point at the lower right corner of cell D24 (the pointer is displayed as a crosshair) and then double-click the left mouse button. Double-clicking copies the formula in a column to match the number of filled rows in the column to the left of the current column. This trick can also be used to copy formulas in multiple columns.

- By copying from E24 to E25:E1590 the formula *IF(AND(C24="No", B24>D24),"yes","no")*, we determine for each month whether our S&P share is purchased during the month. Remember that we purchase the share only if we did not own the stock at the beginning of the month and the current value of the S&P exceeds its 15-month moving average. Notice the AND portion of the formula. It contains two conditions (more than two are allowed) separated by a comma. If both conditions are satisfied, the formula returns Yes; otherwise, it returns No. For an IF formula to recognize text, you need to place quotation marks (" ") around the text.

- By copying from F24 to F25:F1590 the formula *IF(AND(C24="Yes", B24<D24),"yes","no")*, we determine for each month whether our S&P share is sold. The stock is sold if and only if we owned the S&P share at the beginning of the month and the current value of the S&P share is below the 15-month moving average. April 1873 is the first month in which we sell our S&P stock.
- During any month before October 2002, if we buy a share of the S&P during the month, our cash flow is negative the value of the S&P share we bought. If we sell a share of the S&P during the month, our cash flow equals the value of the S&P. Otherwise, the cash flow is 0. During October 2002, we sell any S&P share we own to get credit for its value. Therefore, by copying from G24 to G25:G1589 the formula *IF(E24="yes",-B24,IF(F24="yes",B24,0))*, we record our cash flow for all months before October 2002. Entering in cell G1590 the formula *IF(C1590="yes",B1590,0)* gives us credit for selling any stock we own at the beginning of October 2002.
- In cell G6 we compute our total profit from the moving-average trading strategy with the formula *SUM(G24:G1590)*. We find the 15-month moving average strategy earns a profit of $1319.75.
- The profit from buying and holding shares is simply the October 2002 S&P value minus the April 1872 S&P value. We compute the profit from the buy-and-hold strategy in cell G7 with the formula *B1590-B24*. We find that the buy and hold profit of $849.45 is far worse than the profit from the moving-average trading rule.

### In the game of craps, two dice are tossed. If the total of the dice on the first roll is 2, 3, or 12, you lose. If the total of the dice on the first roll is 7 or 11, you win. Otherwise, the game keeps going. How can I write a formula to determine the status of the game after the first roll?

The fact that you lose in craps if you throw a 2, 3, or 12 can be conveniently modeled by placing an OR formula within an IF formula. In cell B5 of the worksheet Craps, shown in Figure 12-4 and found in the file IfStatement.xls, we enter the formula *IF(OR(A5=2,A5=3,A5=12),"Lose",IF(OR(A5=7,A5=11),"Win","Keep going"))*. This formula is then copied from B5 to B6:B7. The formula displays *Lose* if a 2, 3, or 12 is entered in cell A5. It displays *Win* if a 7 or 11 is entered, and it displays *Keep going* for any other value.

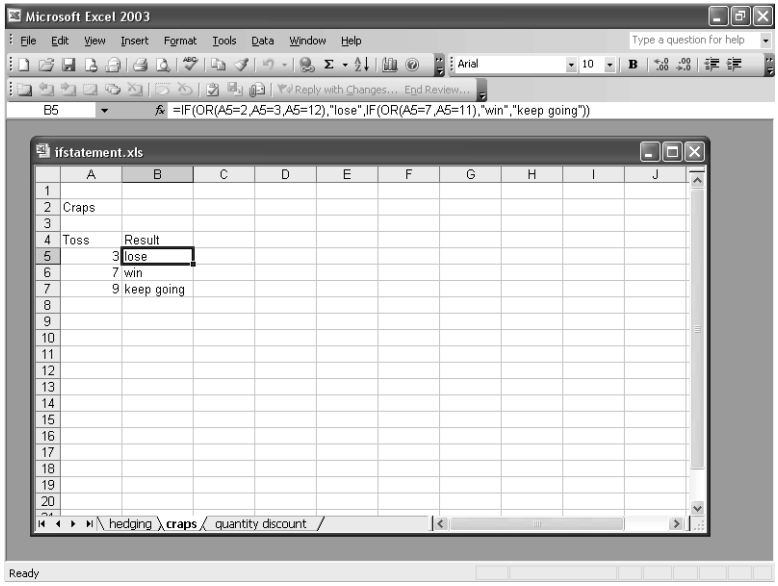
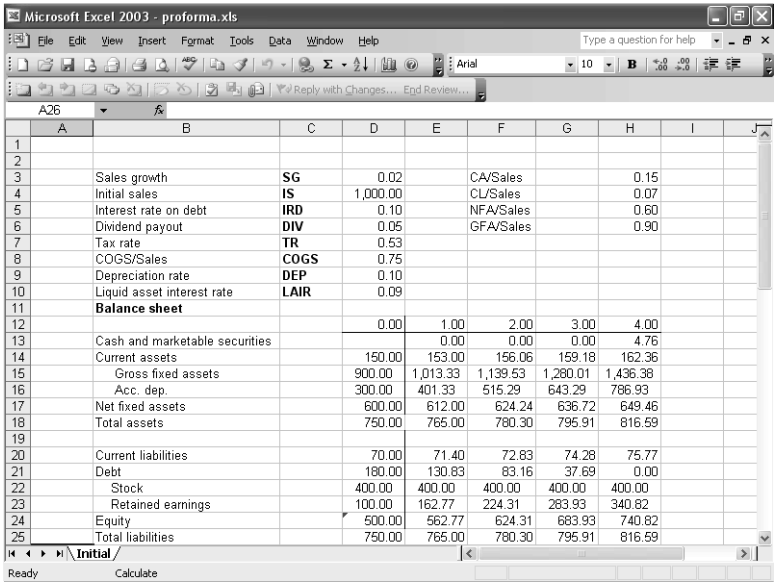



**Figure 12-4** Using IF statements to model the first roll in craps.

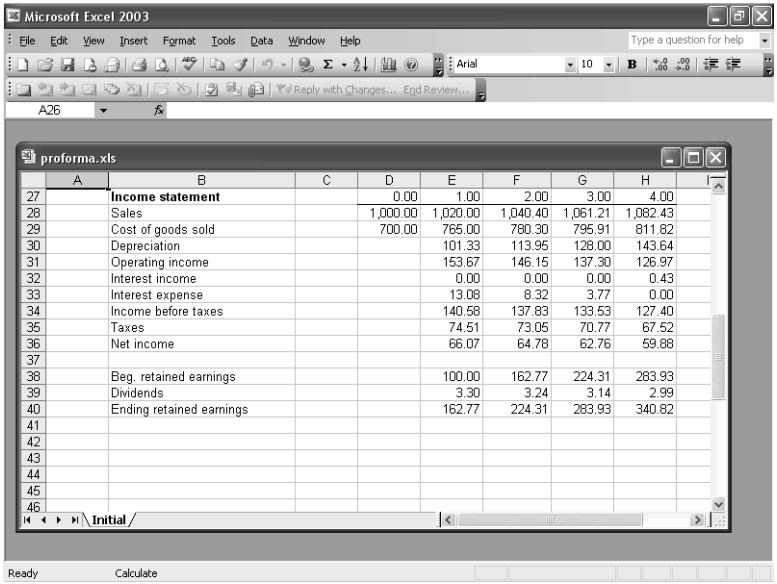
## In most pro forma financial statements, cash is used as the plug to make assets and liabilities balance. I know that using debt as the plug would be more realistic. How can I set up a pro forma statement having debt as the plug?

A pro forma is basically a prediction of a company's financial future. A pro forma requires construction of a company's future balance sheets and income statements. The balance sheet provides a snapshot of the company's assets and liabilities at any point in time. An income statement tells us how the company's financial status is changing at any point in time. Pro forma statements can help a company determine its future needs for debt and are also key parts of valuation models that stock analysts use to determine whether a stock is properly valued. In the file Proforma.xls, I've generated the free cash flows (FCFs) for the next four years for a firm. Figure 12-5 shows the balance sheet, and Figure 12-6 shows the income statement.




| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | | Sales growth | **SG** | 0.02 | | CA/Sales | | 0.15 | |
| 4 | | Initial sales | **IS** | 1,000.00 | | CL/Sales | | 0.07 | |
| 5 | | Interest rate on debt | **IRD** | 0.10 | | NFA/Sales | | 0.60 | |
| 6 | | Dividend payout | **DIV** | 0.05 | | GFA/Sales | | 0.90 | |
| 7 | | Tax rate | **TR** | 0.53 | | | | | |
| 8 | | COGS/Sales | **COGS** | 0.75 | | | | | |
| 9 | | Depreciation rate | **DEP** | 0.10 | | | | | |
| 10 | | Liquid asset interest rate | **LAIR** | 0.09 | | | | | |
| 11 | | **Balance sheet** | | | | | | | |
| 12 | | | | 0.00 | 1.00 | 2.00 | 3.00 | 4.00 | |
| 13 | | Cash and marketable securities | | | 0.00 | 0.00 | 0.00 | 4.76 | |
| 14 | | Current assets | | 150.00 | 153.00 | 156.06 | 159.18 | 162.36 | |
| 15 | | Gross fixed assets | | 900.00 | 1,013.33 | 1,139.53 | 1,280.01 | 1,436.38 | |
| 16 | | Acc. dep. | | 300.00 | 401.33 | 515.29 | 643.29 | 786.93 | |
| 17 | | Net fixed assets | | 600.00 | 612.00 | 624.24 | 636.72 | 649.46 | |
| 18 | | Total assets | | 750.00 | 765.00 | 780.30 | 795.91 | 816.59 | |
| 19 | | | | | | | | | |
| 20 | | Current liabilities | | 70.00 | 71.40 | 72.83 | 74.28 | 75.77 | |
| 21 | | Debt | | 180.00 | 130.83 | 83.16 | 37.69 | 0.00 | |
| 22 | | Stock | | 400.00 | 400.00 | 400.00 | 400.00 | 400.00 | |
| 23 | | Retained earnings | | 100.00 | 162.77 | 224.31 | 283.93 | 340.82 | |
| 24 | | Equity | | 500.00 | 562.77 | 624.31 | 683.93 | 740.82 | |
| 25 | | Total liabilities | | 750.00 | 765.00 | 780.30 | 795.91 | 816.59 | |

**Figure 12-5** Pro forma assumptions and balance sheet.




| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 27 | | **Income statement** | | 0.00 | 1.00 | 2.00 | 3.00 | 4.00 |
| 28 | | Sales | | 1,000.00 | 1,020.00 | 1,040.40 | 1,061.21 | 1,082.43 |
| 29 | | Cost of goods sold | | 700.00 | 765.00 | 780.30 | 795.91 | 811.82 |
| 30 | | Depreciation | | | 101.33 | 113.95 | 128.00 | 143.64 |
| 31 | | Operating income | | | 153.67 | 146.15 | 137.30 | 126.97 |
| 32 | | Interest income | | | 0.00 | 0.00 | 0.00 | 0.43 |
| 33 | | Interest expense | | | 13.08 | 8.32 | 3.77 | 0.00 |
| 34 | | Income before taxes | | | 140.58 | 137.83 | 133.53 | 127.40 |
| 35 | | Taxes | | | 74.51 | 73.05 | 70.77 | 67.52 |
| 36 | | Net income | | | 66.07 | 64.78 | 62.76 | 59.88 |
| 37 | | | | | | | | |
| 38 | | Beg. retained earnings | | | 100.00 | 162.77 | 224.31 | 283.93 |
| 39 | | Dividends | | | 3.30 | 3.24 | 3.14 | 2.99 |
| 40 | | Ending retained earnings | | | 162.77 | 224.31 | 283.93 | 340.82 |
| 41 | | | | | | | | |
| 42 | | | | | | | | |
| 43 | | | | | | | | |
| 44 | | | | | | | | |
| 45 | | | | | | | | |
| 46 | | | | | | | | |

**Figure 12-6** Pro forma income statement.

Column D contains information about the company's current status (during year 0). Our basic assumptions are as follows:

- Sales growth is 2 percent per year
- Initial sales are $1,000
- Interest rate on debt is 10 percent
- Dividend payout is 5 percent of net income
- The tax rate is 53 percent
- Cost of goods sold (COGS) are 75 percent of sales
- Depreciation is 10 percent of gross fixed assets
- Liquid assets earn 9 percent
- Current assets are 15 percent of sales
- Current liabilities are 7 percent of sales
- Net fixed assets are 60 percent of sales

I've assigned the names in the cell range C3:C10 to the cells in the range D3:D10. Then, during each year *t*, basic finance and accounting imply the following relationships, which are then implemented in a series of formulas.

- *Year t+1 sales = (Year t sales)*(1+SG)*. I've computed sales during each year by copying from E28 to F28:H28 the formula *D28*(1+SG)*. I'll refer to this formula as formula 12.1.
- *Year t COGS = COGS*(Year t sales)*. Each year's COGS are computed by copying from E29 to F29:H29 the formula *COGS*E28*. I'll refer to this formula as 12.2.
- If *Year t assets>Year t liabilities*, *Year t debt* must be set equal to *Year t total assets - Year t current liabilities - Year t equity*. Otherwise, *Year t debt = 0*. I've computed each year's debt in E21:H21 with the formula *IF((E18>E20+E24,E18-E20-E24,0)*. If Year *t* total assets are greater than Year *t* total liabilities, this formula sets Year *t* debt to *Year t total assets - Year t current liabilities - Year t equity*. This equalizes, or balances, assets and liabilities. Otherwise, we set Year *t* debt equal to 0. In this case, Year *t* cash and marketable securities will be used to balance assets and liabilities. (12.3)
- *Year t current liabilities = (CL/Sales ratio)*(Year t sales)*. In E20:H20, we use the formula *$H$4*E28* (formula 12.4) to compute current liabilities for each year (copying this formula from E20 to F20:H20).

- *Year t equity = Year t stock + Year t retained earnings*. In E24:H24, we compute equity by copying from E24 to F24:H24 the formula *SUM(E22:E23)*. (12.5)
- If Year *t* debt is greater than 0, Year *t* cash and marketable securities equals 0. Otherwise, Year *t* cash and marketable securities equals *MAX(0, Year t total liabilities - Year t current assets - Year t net fixed assets)*. In E13:H13, I compute cash and marketable securities for each year by copying from E13 to F13:H13 the formula *IF(E21>0,0,MAX(0,E25-E14-E17))*. If Year *t* debt is greater than 0, we need not use Year *t* cash and marketable securities to balance assets and liabilities. In this case we set Year *t* cash and marketable securities equal to 0. Otherwise, we set Year *t* cash and marketable securities equal to *Year t total assets - Year t current liabilities - Year t equity*. This balances assets and liabilities if Year *t* assets (without cash and marketable securities) are less than Year *t* liabilities. If debt does not balance assets and liabilities, this creates liquid assets as the plug that does balance assets and liabilities. (12.6)
- *Year t interest expense = (Year t Debt)*IRD*. In E33:H33, I compute interest expense by using the formula *IRD*E21*, copying this formula again to F33:H33. (12.7)
- *Year t interest income = (Year t cash and marketable securities)*LAIR*. In E32:H32, I compute interest income by copying from E32 to F32:H32 the formula *LAIR*E13*. (12.8)
- *Year t operating income = Year t sales - Year t COGS - Year t depreciation*. In E31:H31, operating income is computed by copying from E31 to F31:H31 the formula *E28-E29-E30*. (12.9)
- *Year t dividends = (Year t net income)*DIV*. In E39:H39, I copy from E39 to F39:H39 the formula *E36*DIV* to compute dividends for each year. (12.10)
- *Year t + 1 beginning retained earnings = Year t ending retained earnings*. I compute beginning retained earnings each year in F38:H38, copying from F38 to G38:H38 the formula *E40*. (12.11)
- *Year t end of year retained earnings = Year t beginning retained earnings + Year t net income - Year t dividends*. In E40:H40, I compute each year's ending retained earnings by copying from E40 to F40:H40 the formula *E38+E36-E39*. (12.12)

- *Year t income before taxes = Year t operating income - Year t interest expense + Year t cash income*. By copying from E34 to F34:H34 the formula *E31-E33+E32*, I compute income before taxes. (12.13)
- *Year t taxes = (Year t income before taxes)*TR*. I compute each year's taxes in E35:H35 by copying from E35 to F35:H35 the formula *TR*E34*. (12.14)
- *Year t net income after taxes = (Year t income before taxes) - (Year t taxes)*. In E36:H36, I compute each year's net income by copying from E36 to F36:H36 the formula *E34-E35*. (12.15)
- *Year t net fixed assets = (Year t sales)*(NFA/Sales)*. In E17:H17, I compute each year's net fixed assets by copying the formula *$H$5*E28* from E17 to F17:H17. (12.16)
- *Year t gross fixed assets = Year t net fixed assets + Year t accumulated depreciation*. In cells E15:H15, I compute gross fixed assets for each year by copying the formula *E17+E16*. (12.17)
- *Year t depreciation = (Year t net fixed assets)*DEP*. Each year I use the formula *DEP*E15* to compute depreciation, copying the formula from E30 to F30:H30. (12.18)
- *Year t accumulated depreciation = Year t-1 accumulated depreciation + Year t depreciation*. Each year I use the formula *D16+E30* to compute accumulated depreciation by copying the formula from E16 to F16:H16. (12.19)
- *Year t net fixed assets = Year t gross fixed assets - Year t accumulated depreciation*. In row 17, to compute net fixed assets, I copy from E16 to F16:H16 the formula *D15-D16*. (12.20)
- *Year t total assets = Year t liquid assets + Year t net fixed assets + Year t cash and marketable securities*. By adding liquid assets, current assets, and net fixed assets, I compute our total assets by copying from E18 to F18:H18 the formula *SUM(E13,E14,E17)*. (12.21)
- *Year t total liabilities = Year t current liabilities + Year t debt + Year t equity*. By copying from E25 to F25:H25 the formula *SUM(E20,E21,E24)*, I compute total liabilities for each period. Each year will balance because of our debt and liquid asset statements.

Formulas 12.3 and 12.6 require the use of IF statements. This spreadsheet will also contain circular references. (For more information about solving circular references, see Chapter 11). For example, the following relationships create a circular reference:

- Year *t* cash affects Year *t* total assets.
- Year *t* total assets affect Year *t* debt.
- Year *t* debt affects Year *t* cash.

For this spreadsheet to return sensible results, we need to select the Iteration option on the Calculation tab of the Options dialog box. (Click Tools, Options to open the dialog box.) Note that for each year *t*, total assets in row 18 equal total liabilities in row 25. This shows the power of IF formulas and circular references.

# Problems

**1.** Suppose the price of a product will change at dates in the future as follows:

| Date | Price |
|---|---|
| On or before February 15, 2004 | $8.00 |
| Between February 16, 2004, and April 10, 2005 | $9.00 |
| Between April 11, 2005, and January 15, 2006 | $10.00 |
| After January 15, 2006 | $12.00 |

Write a formula that will compute the price of the product based on the date the product is sold.

**2.** The Flybynight Airline flight from Seattle to New York has a capacity of 250 people. 270 tickets have been sold for the flight at a price of $300 per ticket. Tickets are non-refundable. The variable cost of flying a passenger (mostly food costs and fuel costs) is $30 per passenger. If more than 250 people show up for the flight, the flight is overbooked and Flybynight must pay overbooking compensation of $350 per person to each overbooked passenger. Develop a spreadsheet that computes Flybynight's profit based on the number of customers who show up for the flight.

3. A major drug company is trying to determine the correct plant capacity for a new drug. A unit of annual capacity can be built for a cost of $10. Each unit of the drug sells for $12 and incurs variable costs of $2. The drug will be sold for 10 years. Develop a spreadsheet that computes the company's 10-year profit given the chosen annual capacity level and the annual demand for the drug. We will assume demand for the drug is the same each year. You may ignore the time value of money in this problem.

4. Our drug company is producing a new drug. They have made the following assumptions:

   - 100,000 units will be sold during year 1.
   - Sales will grow for 3 years and then decline for 7 years.
   - During the growth portion of the product life cycle, sales will grow at 15 percent per year. During the decline part of the product life cycle, sales will drop at a rate of 10 percent per year.

   Develop a spreadsheet that takes values for year 1 sales, length of growth cycle, length of decline cycle, growth rate during growth cycle, and rate of decrease during decline cycle and computes units sales for years 1-11.

5. We are bidding on a construction project. The low bid will get the project. We estimate our project cost at $10,000. Four companies are bidding against us. It costs $400 to prepare the bid. Write a formula that (given the bids of our four competitors and our bid) computes our profit.

6. We are bidding on a valuable painting. The high bid will get the painting. We estimate the painting's value at $10,000. Four companies are bidding against us. It costs $400 to prepare the bid. Write a formula that (given the bids of our four competitors and our bid) determines whether we get the painting.

7. Our drug company believes a new drug will sell 10,000 units during 2004. They expect two competitors to enter the market. The year in which the first competitor enters, our company expects to lose 30 percent of their market share. The year in which the second competitor enters, they expect to lose 15 percent of their market share. The size of the market is growing at 10 percent per year. Given values of the years in which the two competitors enter, develop a spreadsheet that computes the annual sales for the years 2004-2013.

**8.** A clothing store has ordered 100,000 swimsuits. It costs $22 to produce a swimsuit. They plan to sell them until August 31 at a price of $40 and then mark the price down to $30. Given values for demand through August 31 and after August 31, develop a spreadsheet to compute the profit from this order.

**9.** In a game of craps, on each roll of the dice after the first roll, the rules are as follows: If the game has not ended and the current roll matches the first roll, we win the game. If the game has not ended and the current roll is a 7, we lose. Otherwise, the game continues. Develop a spreadsheet that tells us (given knowledge of the first four rolls) the status of the game after four dice rolls.

**10.** On our S&P moving-average example, suppose we still buy shares if the current price exceeds the 15-month moving average, but we sell if the current price is less than the 5-month moving average. Is this strategy more profitable than selling if the current price is less than the 15-month moving average?

**11.** A European call option gives us the right to *buy* at a specified future date a share of stock for a given exercise price. A *butterfly spread* involves buying one call option with a low exercise price, buying one call option with a high exercise price, and selling two call options with an exercise price midway between the low and high exercise prices. Here is an example of a butterfly spread. The current stock price is $60. We buy a $54 six-month European call option for $9, buy a $66 six-month European call option for $4, and sell two $60 European call options for the price of $6. Compute the profit (in dollars, not percentage) for this transaction as a function of six-month stock prices ranging from $40-$80. When a trader purchases a butterfly spread, which type of movement in the stock price during the next six months is the trader betting on?

**12.** Suppose a stock is currently selling for $32. We buy a six-month European call option with an exercise price of $30 for $2.50 and sell a six-month European call option with an exercise price of $35 for $1. Compute the profit of this strategy (in dollars) as a function of a six-month stock price ranging from $25-$45. Why is this strategy called a *bull spread*? How would you modify this strategy to create a *bear spread*?

**13.** Let's reconsider our pro forma example. Suppose the interest rate on our debt depends on our financial well-being. More specifically, suppose that if our earnings before interest and taxes (EBIT) is negative, our interest rate on debt is 16 percent. If our interest expense is more than 10 percent of EBIT and EBIT is positive, our interest rate on debt is 13 percent. Otherwise, our interest rate is 10 percent. Modify our pro forma to account for this variable interest rate.

**14.** Do this problem independently of problem 13. Suppose our firm wants a debt/equity ratio of 50 percent during each year. How would you modify our pro forma? Hint: You must keep each year's stock nonnegative and use stock and cash or marketable securities to balance assets and liabilities.

**15.** Martin Luther King Day occurs on the third Monday in January. Write a formula that computes (given the year) the date of Martin Luther King Day. Hint: First determine the day of the week for January 1 of the given year.

**16.** Thanksgiving occurs on the fourth Thursday in November. Write a formula that computes (given the year) the date of Thanksgiving. Hint: First determine the day of the week for November 1 of the given year.

**17.** The first quarter of the year is January-March; the second quarter, April-June; the third quarter, July-September; and the fourth quarter, October-December. Write a formula that returns (for any given date) the quarter of the year.

# 13

# The Paste Special Command

- How can I move the results of calculations (not the formulas) to a different part of a worksheet?
- I have a list of names in a single column. How can I make the list appear in one row instead of one column?
- I've downloaded U.S. Treasury bill interest rates from a Web site into Excel. The data lists a 5 when the interest rate is 5 percent, 8 when the interest rate is 8 percent, and so on. How can I easily divide my results by 100 so that a 5 percent interest rate, for example, is listed as .05?

The Paste Special command in Excel enables you to easily perform certain manipulations of spreadsheet data. In this chapter, I'll show how you can use the Paste Special command to perform the following types of operations:

- Pasting only the values in cells (not the formulas) to a different part of a spreadsheet.
- Transposing data in columns to rows and vice versa.
- Transforming a range of numbers by adding, subtracting, dividing, or multiplying each number in the range by a given constant.

## How can I move the results of calculations (not the formulas) to a different part of a worksheet?

In the worksheet named Paste Special Value in the file PasteSpecial.xls, the cell range E4:H9 contains the names, games, total points, and points per game for five 10–11-year-old basketball players from Bloomington, Indiana. In the cell range H5:H9, I've used the data in cells F5:G9 to compute each child's points per game, as shown in Figure 13-1. Suppose we want to copy this data and the calculated points per game—but not the formulas that perform the calculations—to a different cell range (E13:H18, for example). All you do is select the range E4:H9, choose Edit, Copy, and then move to the upper left corner of the range where you want to copy the data (cell E13 in this example). Next, select Edit, Paste Special, and then fill in the Paste Special dialog box as indicated in Figure 13-2. After clicking OK, the range E13:H18 contains the data but not the formulas from the cell range E4:H9. You can check this by going to cell H16. You will see a value (7) but not the formula that was used to compute Gregory's average points per game. Note that if you use the Paste Special command, select Values, and then paste the data into the same range from which you copied the data, your formulas will disappear from the spreadsheet.

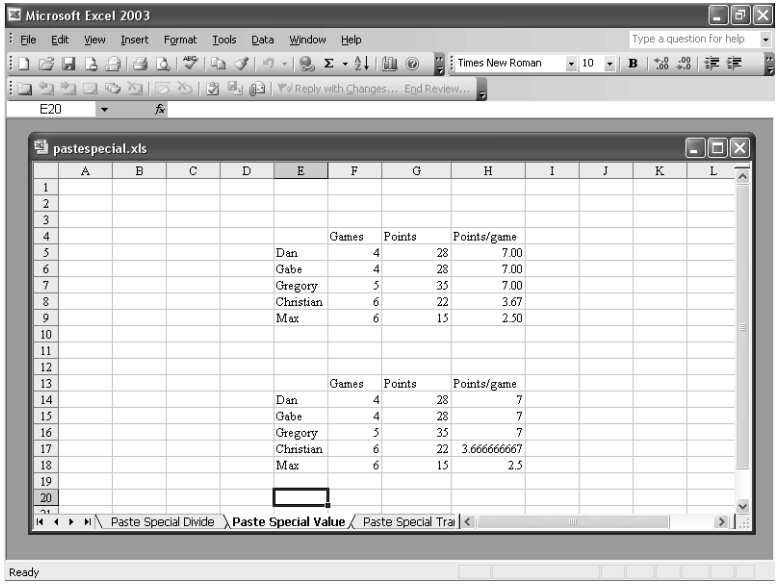



**Figure 13-1** An example of using the Paste Special command to paste only values.

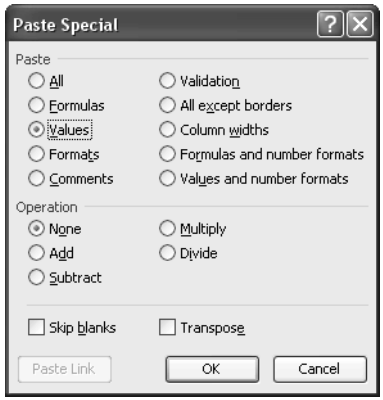
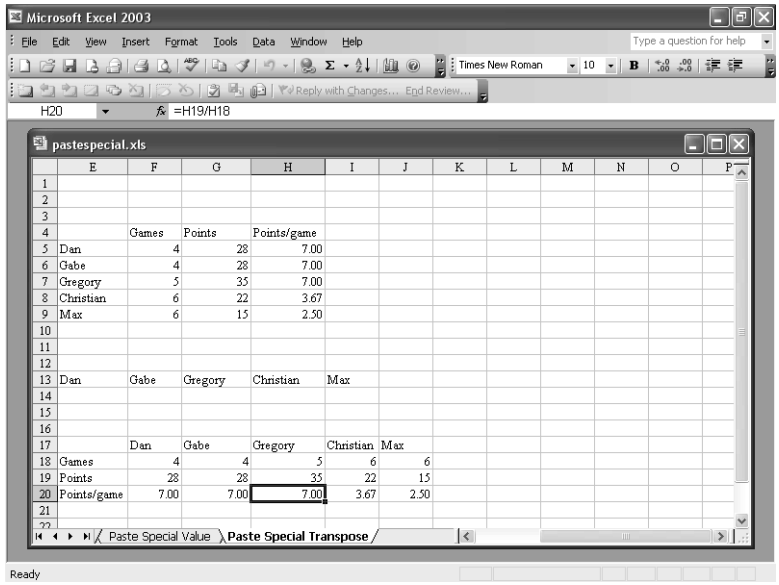



**Figure 13-2** The Paste Special dialog box with Values selected. Selecting Values pastes only values and not any formulas.

## I have a list of names in a single column. How can I make the list appear in one row instead of one column?

To realign data from a row to a column (or vice versa), the key is to use Edit, Copy and then Paste Special, Transpose. Essentially, the Transpose option in the Paste Special dialog box "flips" selected cells around so that the first row of the copied range becomes the first column of the range you paste data into, and so on. Look at the worksheet named Paste Special Transpose in the file PasteSpecial.xls, shown in Figure 13-3.




| | E | F | G | H | I | J |
|---|---|---|---|---|---|---|
| 4 | | Games | Points | Points/game | | |
| 5 | Dan | 4 | 28 | 7.00 | | |
| 6 | Gabe | 4 | 28 | 7.00 | | |
| 7 | Gregory | 5 | 35 | 7.00 | | |
| 8 | Christian | 6 | 22 | 3.67 | | |
| 9 | Max | 6 | 15 | 2.50 | | |
| 13 | Dan | Gabe | Gregory | Christian | Max | |
| 17 | | Dan | Gabe | Gregory | Christian | Max |
| 18 | Games | 4 | 4 | 5 | 6 | 6 |
| 19 | Points | 28 | 28 | 35 | 22 | 15 |
| 20 | Points/game | 7.00 | 7.00 | 7.00 | 3.67 | 2.50 |

**Figure 13-3** Use the Transpose option in the Paste Special dialog box to transpose a row of data into a column of data or a column into a row.

Suppose that you want to list the players' names in a single row (starting in cell E13). Simply select the range E5:E9, and then choose Edit, Copy. Move to cell E13, and then choose Edit, Paste Special and check Transpose in the Paste Special dialog box. After clicking OK, the players' names are transposed into a single row.

Suppose you want to transpose the spreadsheet content in E4:H9 to a range beginning in cell E17. Begin by selecting the range E4:H9. Next, choose Edit, Copy and then move to the upper left corner of the range where you want to put the transposed information (E17). Choose Edit, Paste Special, check Transpose, and then click OK. You'll see that the content of E4:H9 is transposed (turned on its side), as shown in Figure 13-3. Note that in F20:J20, Excel was smart enough to adjust the points-per-game formula so that the average for each player is now computed from data in the same column instead of the same row.

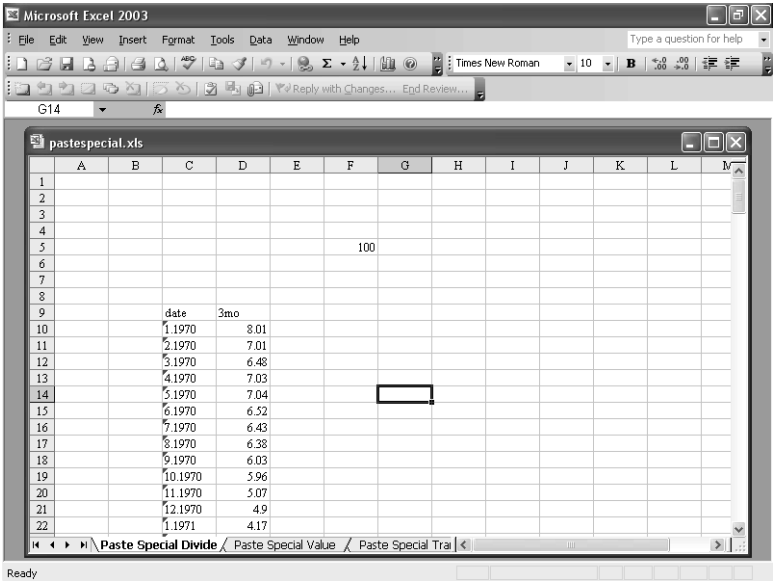
**Note** When you select Paste Special and click Paste Link instead of OK, the transposed cells are linked to the original cells, and changes you make to the original data are reflected in the copy. By changing the value in cell F5 to 7, the value in cell F18 becomes 7 as well, and cell F20 would display Dan's average as 4 points per game.

## I've downloaded U.S. Treasury bill interest rates from a Web site into Excel. The data lists a 5 when the interest rate is 5 percent, 8 when the interest rate is 8 percent, and so on. How can I easily divide my results by 100 so that a 5 percent interest rate, for example, is listed as .05?

The worksheet Paste Special Divide in the file PasteSpecial.xls (see Figure 13-4) contains the annual rate of interest paid by three-month U.S. Treasury bills for each month between January 1970 and February 1987. In January 1970, the annual rate on a three-month Treasury bill was 8.01 percent. Suppose we want to earn annual interest on $1 invested at the current T-bill rate. The formula to calculate the rate is *(1 + (annual rate)/100)*. It would be easier to compute earned interest if our column of annual interest rates were divided by 100.

The Operations portion of the Paste Special dialog box lets you add, subtract, multiply, or divide each number in a range by a given number, providing an easy way to divide each interest rate by 100. Here we want to divide each number in column D. To begin, I entered our given number (100). You can enter it anywhere in the spreadsheet. I chose F5. With F5 selected, choose Edit, Copy. Next select the range of numbers you want to modify. To select all the

data in column D, move the cursor to cell D10 and press Ctrl+Shift and then the down arrow. This shortcut is a useful trick for selecting a "tall" cell range. (By the way, to select a "wide" set of data listed in a single row, move to the first data point and then press Ctrl+Shift and the right arrow.) Next, select Edit, Paste Special, and then select Divide, as shown in Figure 13-5.




| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | |
| 4 | | | | | | | | | | | | | |
| 5 | | | | | | 100 | | | | | | | |
| 6 | | | | | | | | | | | | | |
| 7 | | | | | | | | | | | | | |
| 8 | | | | | | | | | | | | | |
| 9 | | | date | 3mo | | | | | | | | | |
| 10 | | | 1.1970 | 8.01 | | | | | | | | | |
| 11 | | | 2.1970 | 7.01 | | | | | | | | | |
| 12 | | | 3.1970 | 6.48 | | | | | | | | | |
| 13 | | | 4.1970 | 7.03 | | | | | | | | | |
| 14 | | | 5.1970 | 7.04 | | | | | | | | | |
| 15 | | | 6.1970 | 6.52 | | | | | | | | | |
| 16 | | | 7.1970 | 6.43 | | | | | | | | | |
| 17 | | | 8.1970 | 6.38 | | | | | | | | | |
| 18 | | | 9.1970 | 6.03 | | | | | | | | | |
| 19 | | | 10.1970 | 5.96 | | | | | | | | | |
| 20 | | | 11.1970 | 5.07 | | | | | | | | | |
| 21 | | | 12.1970 | 4.9 | | | | | | | | | |
| 22 | | | 1.1971 | 4.17 | | | | | | | | | |

**Figure 13-4** Use the Divide option in the Paste Special dialog box to divide a data range by a constant.

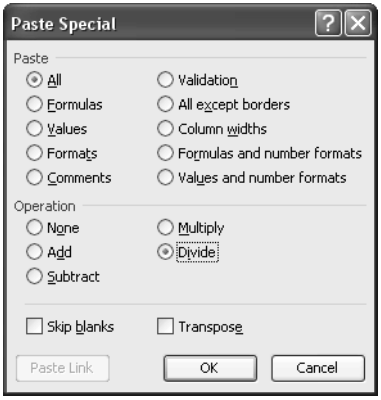



**Figure 13-5** You can apply an option in the Operation area of the Paste Special dialog box to a range of cells.

After you click OK, Excel divides each selected number in column D by 100. The results are shown in Figure 13-6. If we had selected Add, D10 would have displayed 108.01; if we had selected Subtract, D10 would have displayed -91.99; and if we had selected Multiply, D10 would have displayed 801.

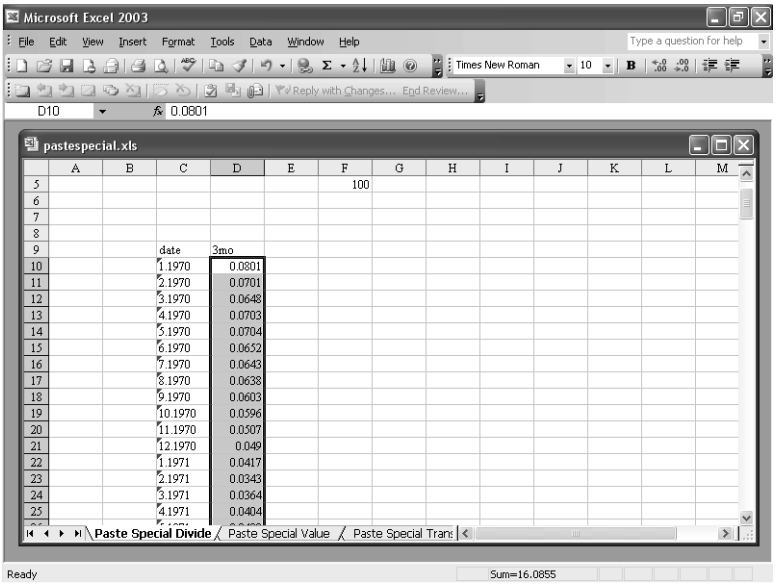



**Figure 13-6** Results of using the Divide option in the Paste Special dialog box.

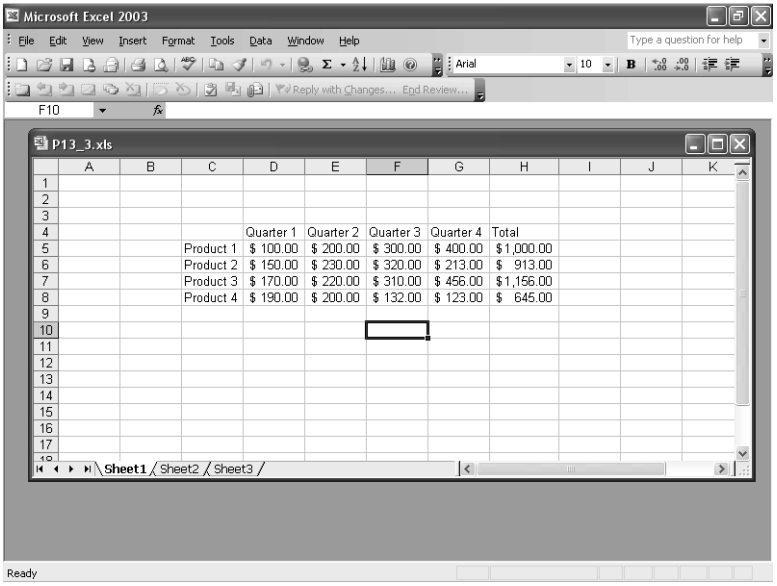
# Problems

The workbook Mavs.xls contains statistics for the great 2002-2003 Dallas Mavericks basketball team. Player names are listed in column A, statistical categories are listed in row 3, and statistical results are listed in rows 4-20.

**1.** Change the spreadsheet so that all player names are listed in the same row and all statistical categories are listed in the same column.

**2.** Field goal percentage, free throw percentage, and three point percentage are listed as decimals. For example, Steve Nash makes 91.9 percent of his free throws, which is listed as .919. Change the spreadsheet so that all shooting percentages are listed as a number between 1 and 100.

**3.** The file P13_3.xls, shown here, contains data on quarterly sales for four products.




| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | | | | | | | | | | | |
| 4 | | | | Quarter 1 | Quarter 2 | Quarter 3 | Quarter 4 | Total | | | |
| 5 | | | Product 1 | $ 100.00 | $ 200.00 | $ 300.00 | $ 400.00 | $1,000.00 | | | |
| 6 | | | Product 2 | $ 150.00 | $ 230.00 | $ 320.00 | $ 213.00 | $ 913.00 | | | |
| 7 | | | Product 3 | $ 170.00 | $ 220.00 | $ 310.00 | $ 456.00 | $1,156.00 | | | |
| 8 | | | Product 4 | $ 190.00 | $ 200.00 | $ 132.00 | $ 123.00 | $ 645.00 | | | |
| 9 | | | | | | | | | | | |
| 10 | | | | | | | | | | | |
| 11 | | | | | | | | | | | |
| 12 | | | | | | | | | | | |
| 13 | | | | | | | | | | | |
| 14 | | | | | | | | | | | |
| 15 | | | | | | | | | | | |
| 16 | | | | | | | | | | | |
| 17 | | | | | | | | | | | |

Copy this data to another range so that quarterly sales are read across instead of down. Link your copied data to the original data so your computation of annual sales in the copied data range will reflect changes entered in rows 5–8.

# 19

# The COUNTIF, COUNT, COUNTA, and COUNTBLANK Functions

Suppose I have a list of songs that are played on the radio. For each song, I know the singer, the date the song was played, and the length of the song. How can I answer questions such as

- How many songs were sung by each singer?
- How many songs were not sung by Eminem?
- How many songs lasted at least 4 minutes?
- How many songs were longer than average?
- How many songs were sung by a singer whose last name begins with S?
- How many songs were sung by a singer whose last name contains six letters?
- How many songs were sung after June 15, 2005?
- How many songs were sung before the beginning of 2009?
- How many songs lasted exactly 4 minutes?

In a more general context, how do I perform operations such as

- Count the number of cells in a range containing numbers?
- Count the number of blank cells in a range?
- Count the number of nonblank cells in a range?

We often want to count the number of cells in a range that meet a given criterion. For example, if a spreadsheet contains information about makeup sales, we might want to count how many sales transactions were made by the salesperson named Jennifer or how many sales transactions occurred after June 10. The COUNTIF function lets you count the number of cells in a range that meet criteria that are defined on the basis of a single row or column of the spreadsheet.

The syntax of the COUNTIF function is *COUNTIF(range, criteria)*.

- *Range* is the range of cells in which you want to count cells meeting a given criterion.
- *Criteria* is a number, date, or expression that determines whether a given cell in the range is counted.

The key to using the COUNTIF function (and other similar functions) suc cessfully is understanding the wide variety of criteria that Excel will accept. The types of criteria you can use are best explained through the use of examples. In addition to examples of the COUNTIF function, I'll provide examples of the COUNT, COUNTA, and COUNTBLANK functions. The COUNT function counts the number of cells in a range containing numbers. The COUNTA function counts the number of nonblank cells in a range. The COUNTBLANK function counts the number of blank cells in a range.

To illustrate how to use these functions, consider a database that gives the following information for each song played on radio station WKRP:

- Singer
- The date the song was played
- The length of the song

The file Rock.xls, shown in Figure 19-1, shows a subset of the data.

rock.xls

| | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | | | | | | | | | |
| 4 | | | | average | | | | | |
| 5 | | | | 3.483701 | | | | | |
| 6 | Song Number | Singer | Date | Minutes | | | | | |
| 7 | 1 | Eminem | 5/21/2004 | 4 | | | | | |
| 8 | 2 | Eminem | 4/15/2004 | 2 | | | | | |
| 9 | 3 | Cher | 1/28/2005 | 2 | | | | | |
| 10 | 4 | Eminem | 1/28/2005 | 4 | | | | | |
| 11 | 5 | Moore | 11/5/2004 | 2 | | | | | |
| 12 | 6 | Cher | 9/18/2004 | 4 | | | | | |
| 13 | 7 | Spears | 4/15/2004 | 3 | | | | | |
| 14 | 8 | Spears | 3/17/2005 | 3 | | | | | |
| 15 | 9 | Manilow | 1/16/2005 | 4 | | | | | |
| 16 | 10 | Eminem | 4/10/2005 | 4 | | | | | |
| 17 | 11 | Madonna | 2/15/2004 | 3 | | | | | |
| 18 | 12 | Eminem | 1/10/2004 | 4 | | | | | |
| 19 | 13 | Springsteen | 4/10/2005 | 2 | | | | | |
| 20 | 14 | Spears | 4/15/2004 | 4 | | | | | |
| 21 | 15 | Moore | 7/8/2004 | 3 | | | | | |

Sheet1

**Figure 19-1** A song database that we'll use for COUNTIF examples.

## How many songs were sung by each singer?

To begin, I select the first row of the database, the range D6:G6. Then I select the whole database by pressing Ctrl+Shift+Down Arrow. Next I choose Insert, Name, Create, and then choose Top Row. We have now named the range D7:D957 Song Numb, the range E7:E957 Singer, the range F7:F957 Date, and the range G7:G957 Minutes. To determine how many songs were sung by each singer, we copy from C5 to C6:C12 the formula *COUNTIF(Singer,B5)*. In cell C5, this formula now displays the number of cells in the range Singer that match the value in B5 (Eminem). The database contains 114 songs sung by Eminem. Sim ilarly, Cher sang 112 songs, and so on, as you can see in Figure 19-2. I could have also found the number of songs sung by Eminem with the formula *COUNTIF("Singer,eminem")*. Note that you must enclose text such as *eminem* in quotation marks (" ")and that criteria are not case sensitive.

rock.xls

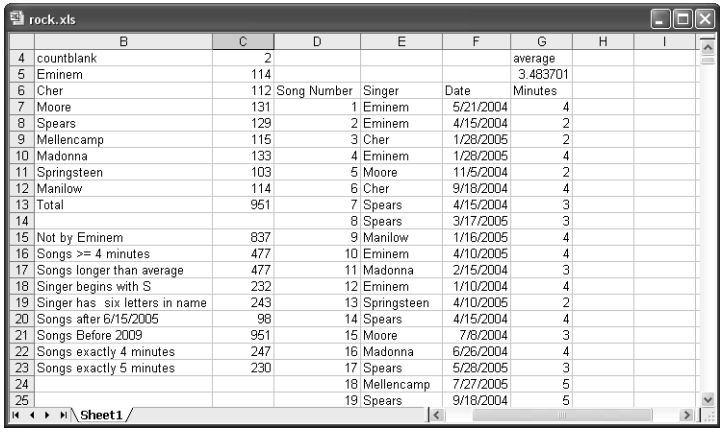
| | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|
| 1 | | | | | | | |
| 2 | count | 9 | | | | | |
| 3 | counta | 18 | | | | | |
| 4 | countblank | 2 | | | | average | |
| 5 | Eminem | 114 | | | | 3.483701 | |
| 6 | Cher | 112 | Song Number | Singer | Date | Minutes | |
| 7 | Moore | 131 | 1 | Eminem | 5/21/2004 | 4 | |
| 8 | Spears | 129 | 2 | Eminem | 4/15/2004 | 2 | |
| 9 | Mellencamp | 115 | 3 | Cher | 1/28/2005 | 2 | |
| 10 | Madonna | 133 | 4 | Eminem | 1/28/2005 | 4 | |
| 11 | Springsteen | 103 | 5 | Moore | 11/5/2004 | 2 | |
| 12 | Manilow | 114 | 6 | Cher | 9/18/2004 | 4 | |
| 13 | Total | 951 | 7 | Spears | 4/15/2004 | 3 | |
| 14 | | | 8 | Spears | 3/17/2005 | 3 | |

Sheet1

**Figure 19-2** Using COUNTIF to determine how many songs were sung by each singer.

## How many songs were not sung by Eminem?

To solve this problem, you need to know that Excel interprets the character combination <> as "not equal to." The formula *COUNTIF(Singer,"<>Eminem")*, entered in cell C15, tells us that 837 songs in the database were not sung by Eminem, as you can see in Figure 19-3. I need to enclose *<>Eminem* in quotation marks because Excel treats the character combination <> as text and *Eminem* is, of course, text. You could obtain the same result by using the formula *COUNTIF(Singer,"<>"&B5)*, which uses the *&* symbol to concatenate the reference to cell B5 and the <> operator.



rock.xls

| | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|
| 4 | countblank | 2 | | | | average | | |
| 5 | Eminem | 114 | | | | 3.483701 | | |
| 6 | Cher | 112 | Song Number | Singer | Date | Minutes | | |
| 7 | Moore | 131 | 1 | Eminem | 5/21/2004 | 4 | | |
| 8 | Spears | 129 | 2 | Eminem | 4/15/2004 | 2 | | |
| 9 | Mellencamp | 115 | 3 | Cher | 1/28/2005 | 2 | | |
| 10 | Madonna | 133 | 4 | Eminem | 1/28/2005 | 4 | | |
| 11 | Springsteen | 103 | 5 | Moore | 11/5/2004 | 2 | | |
| 12 | Manilow | 114 | 6 | Cher | 9/18/2004 | 4 | | |
| 13 | Total | 951 | 7 | Spears | 4/15/2004 | 3 | | |
| 14 | | | 8 | Spears | 3/17/2005 | 3 | | |
| 15 | Not by Eminem | 837 | 9 | Manilow | 1/16/2005 | 4 | | |
| 16 | Songs >= 4 minutes | 477 | 10 | Eminem | 4/10/2005 | 4 | | |
| 17 | Songs longer than average | 477 | 11 | Madonna | 2/15/2004 | 3 | | |
| 18 | Singer begins with S | 232 | 12 | Eminem | 1/10/2004 | 4 | | |
| 19 | Singer has six letters in name | 243 | 13 | Springsteen | 4/10/2005 | 2 | | |
| 20 | Songs after 6/15/2005 | 98 | 14 | Spears | 4/15/2004 | 4 | | |
| 21 | Songs Before 2009 | 951 | 15 | Moore | 7/8/2004 | 3 | | |
| 22 | Songs exactly 4 minutes | 247 | 16 | Madonna | 6/26/2004 | 4 | | |
| 23 | Songs exactly 5 minutes | 230 | 17 | Spears | 5/28/2005 | 3 | | |
| 24 | | | 18 | Mellencamp | 7/27/2005 | 5 | | |
| 25 | | | 19 | Spears | 9/18/2004 | 5 | | |

Sheet1

**Figure 19-3** You can combine the COUNTIF function with the not-equal-to operator (<>).

## How many songs lasted at least 4 minutes?

In cell C16, I've computed the number of songs played that lasted at least 4 minutes by using the formula *COUNTIF(Minutes,">=4")*. You need to enclose *>=4* in quotation marks because >=, like <>, is treated as text. We find that 477 songs lasted at least 4 minutes.

## How many songs were longer than average?

To answer this question, I first computed in cell G5 the average length of a song with the formula *AVERAGE(Minutes)*. Then, in cell C17, I compute the number of songs that last longer than the average with the formula *COUNTIF(Minutes,">"&G5)*. I can refer to another cell (in this case G5) in the criteria by using the *&* sign. You can see that 477 songs lasted longer than average, which matches the number of songs lasting at least 4 minutes. The reason these numbers match is that I assumed the length of each song was an integer. For a song to last at least 3.48 minutes, it has to last at least 4 minutes.

### How many songs were sung by a singer whose last name begins with S?

To answer this question, I use a wildcard character, the asterisk (*), in the criteria. An asterisk represents any sequence of characters. Thus the formula *COUNTIF(Singer,"S*")* in cell C18 picks up any song sung by a singer whose last name begins with S. (The criteria are not case sensitive.) Two hundred thirty-two songs were sung by singers with a last name that begins with S. This number is simply the total of the songs sung by either Bruce Springsteen or Britney Spears (103 + 129 = 232).

### How many songs were sung by a singer whose last name contains six letters?

In this example, I used the question mark (?) wildcard character. The question mark matches any character. Therefore, entering the formula *COUNTIF(Singer,"??????")* in cell C19 counts the number of songs sung by singers having six letters in their last name. The result is 243. (Two singers have last names with six characters, Britney Spears and Eminem, who together sang a total of 129 + 114 = 243 songs.)

### How many songs were sung after June 15, 2005?

The criteria you use with COUNT functions handle dates on the basis of a date's serial number. (A later date is considered larger than an earlier date.) The formula *COUNTIF(Date,">6/15/2005")* in cell C20 tells us that 98 songs were sung after June 15, 2005.

### How many songs were sung before the beginning of 2009?

We want our criteria to pick up all dates on or before December 31, 2008. I've entered in cell C21 the formula *COUNTIF(Date,"<=12/31/2008")*. We find that 951 songs (which turns out to be all the songs) were sung before the start of 2009.

### How many songs lasted exactly 4 minutes?

In cell C22, I compute the number of songs lasting exactly 4 minutes with the formula *COUNTIF(Minutes,4)*. This formula counts the number of cells in the range G7:G957 containing a 4. We find that 247 songs lasted exactly 4 minutes. In a similar fashion, we found in cell C23 that 230 songs lasted exactly 5 minutes.

### How do I count the number of cells in a range containing numbers?

The COUNT function counts the number of cells in a range containing a numeric value. For example, the formula *COUNT(B5:C14)* in cell C2 displays 9 because nine cells (the cells in C5:C13) in the range B5:C14 contain numbers. (See Figure 19-2.)

### How do I count the number of blank cells in a range?

The COUNTBLANK function counts the number of blank cells in a range. For example, the formula *COUNTBLANK(B5:C14)* entered in cell C4 returns a value of 2 because two cells (B14 and C14) in the range B5:C14 contain blanks.

### How do I count the number of nonblank cells in a range?

The COUNTA function returns the number of nonblank cells in a range. For example, the formula COUNTA(B5:C14) in cell C3 returns 18 because 18 cells in the range B5:C14 are not blank.

## Remarks

Unless you use array functions (see Chapter 72), the COUNTIF function can handle only a single criterion. For example, without using an array function, you cannot use COUNTIF to answer a question that depends on two or more criteria, such as how many Britney Spears songs were played before June 10, 2005. The easiest way to answer questions involving multiple criteria is to use *database statistical functions*. See Chapter 37 for a discussion of database statistical functions.

## Problems

The following questions refer to the file Rock.xls.

1. How many songs were not sung by Britney Spears?
2. How many songs were sung before June 15, 2004?
3. How many songs were sung between June 1, 2004, and July 4, 2006? Hint: Take the difference between two COUNTIF functions.
4. How many songs were sung by singers whose last name begins with *M*?
5. How many songs were sung by singers whose names contain an *e*?
6. Create a formula that will always yield the number of songs played today. Hint: Use the TODAY() function.
7. For the cell range D4:G15, count the cells containing a numeric value. Count the number of blank cells. Count the number of nonblank cells.

**8.** The file NBA.xls contains the following information:

- In columns A and B, the name of each NBA team and a code number for each team. For example, Team 1 is Atlanta.
- Column C contains the home team for each game.
- Column D contains the visiting team for each game.
- Column E contains points scored by the home team.
- Column F contains points scored by the visiting team.

From this data, compute for each team the number of games played by each team.

# 20

# The SUMIF Function

- I'm a sales manager for a makeup company and have summarized for each sales transaction the following information: salesperson, date of sale, units sold (or returned), total price received (or paid out for returns). How can I answer the following questions?
  - What was the total dollar amount of merchandise sold by each salesperson?
  - How many units were returned?
  - What was the total dollar volume sold in 2005 or later?
  - How many units of lip gloss were sold?
  - How much revenue did lip gloss sales bring in?
  - What dollar amount of sales were not made by Jen?

If you want to sum all the entries in one column (or row) that match criteria that depend on another column (or row), the SUMIF function gets the job done. The syntax of the SUMIF function is *SUMIF(range, criteria, [sum range])*.

- *Range* is the range of cells that you want to evaluate with a criterion.
- *Criteria* is a number, date, or expression that determines whether a given cell in the sum range is added.
- *Sum* range is the range of cells that are added. If *sum range* is omitted, it is assumed to be the same as *range*.

The rules for criteria you can use with the SUMIF function are identical to the rules used with the COUNTIF function. For information about the COUNTIF function, see Chapter 19.

### What was the total dollar amount of merchandise sold by each salesperson?

Our work for the problems in this chapter is in the file Makeup.xls. Figure 20-1 shows a subset of the data.

makeup.xls

| | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|
| 3 | | | | | | | | | |
| 4 | Trans Number | Name | Date | Product | Units | Dollars | | | |
| 5 | 1 | Betsy | 4/1/2004 | lip gloss | 45 | $ 137.20 | | | |
| 6 | 2 | Hallagan | 3/10/2004 | foundation | 50 | $ 152.01 | | | |
| 7 | 3 | Ashley | 2/25/2005 | lipstick | 9 | $ 28.72 | | | |
| 8 | 4 | Hallagan | 5/22/2006 | lip gloss | 55 | $ 167.08 | | | |
| 9 | 5 | Zaret | 6/17/2004 | lip gloss | 43 | $ 130.60 | | | |
| 10 | 6 | Colleen | 11/27/2005 | eye liner | 58 | $ 175.99 | | | |
| 11 | 7 | Cristina | 3/21/2004 | eye liner | 8 | $ 25.80 | | | |
| 12 | 8 | Colleen | 12/17/2006 | lip gloss | 72 | $ 217.84 | | | |
| 13 | 9 | Ashley | 7/5/2006 | eye liner | 75 | $ 226.64 | | | |
| 14 | 10 | Betsy | 8/7/2006 | lip gloss | 24 | $ 73.50 | | | |
| 15 | 11 | Ashley | 11/29/2004 | mascara | 43 | $ 130.84 | | | |
| 16 | 12 | Ashley | 11/18/2004 | lip gloss | 23 | $ 71.03 | | | |
| 17 | 13 | Emilee | 8/31/2005 | lip gloss | 49 | $ 149.59 | | | |
| 18 | 14 | Hallagan | 1/1/2005 | eye liner | 18 | $ 56.47 | | | |
| 19 | 15 | Zaret | 9/20/2006 | foundation | -8 | $ (21.99) | | | |
| 20 | 16 | Emilee | 4/12/2004 | mascara | 45 | $ 137.39 | | | |
| 21 | 17 | Colleen | 4/30/2006 | mascara | 66 | $ 199.65 | | | |
| 22 | 18 | Jen | 8/31/2005 | lip gloss | 88 | $ 265.19 | | | |
| 23 | 19 | Jen | 10/27/2004 | eye liner | 78 | $ 236.15 | | | |
| 24 | 20 | Zaret | 11/27/2005 | lip gloss | 57 | $ 173.12 | | | |

data / sumif rows

**Figure 20-1** Data about makeup sales we'll use for SUMIF examples.

As usual, we begin by naming the data in columns G through L with the corresponding names in cells G4:L4. For example, the range name Product corresponds to the range J5:J1904. To compute the total amount sold by each salesperson, I simply copy from cell B5 to B6:B13 the formula *SUMIF(Name,A5,Dollars)*. This formula adds up every entry in the Dollars column that has the name *Emilee* in the Name column. We find that Emilee sold $25,258.87 worth of makeup, as you can see in Figure 20-2.

makeup.xls

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 3 | | | | | |
| 4 | Name | Dollar Volume | | | |
| 5 | Emilee | $ 25,258.87 | | | |
| 6 | Hallagan | $ 28,705.16 | | | |
| 7 | Ashley | $ 25,947.24 | | | |
| 8 | Zaret | $ 26,741.31 | | | |
| 9 | Colleen | $ 24,890.66 | | | |
| 10 | Cristina | $ 23,849.56 | | | |
| 11 | Betsy | $ 28,803.15 | | | |
| 12 | Jen | $ 29,050.53 | | | |
| 13 | Cici | $ 27,590.57 | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | Units returned | 922 | | | |
| 17 | Total dollars sold 2005 or later | $ 157,854.32 | | | |
| 18 | Units of lip gloss sold | 16333 | | | |
| 19 | $s of lip gloss sold | $ 49,834.64 | | | |
| 20 | $s sold not by Jen | $ 211,786.51 | | | |
| 21 | | | | | |
| 22 | | | | | |
| 23 | | | | | |

data / sumif rows

**Figure 20-2** Results of SUMIF computations for makeup sales.

### How many units were returned?

In cell B16, the formula *SUMIF(Units,"<0",Units)* adds up every cell in the Units column (column K) that has a number less than 0. The result is -922. After inserting a minus sign in front of the *SUMIF* formula, we see that 922 units were returned. (Recall that when the *sum range* argument is omitted from a SUMIF function, Excel assumes that *sum range* equals *range*. Therefore, the formula *-SUMIF(Units,"<0")* would yield 922 as well.)

### What was the total dollar volume sold in 2005 or later?

In cell B17, the formula *SUMIF(Date,">= 1/1/2005",Dollars)* adds up every cell in the Dollar column (column L) that has a date on or after 1/1/2005. We find that $157,854.32 worth of makeup was sold in 2005 or later.

### How many units of lip gloss were sold? How much revenue did lip gloss sales bring in?

In cell B18, the formula *SUMIF(Product,"lip gloss",Units)* adds up every cell in the Units column that has the text "lip gloss" in the Product column (column J). You can see that 16,333 units of lip gloss were sold. Transactions in which units of lip gloss were returned are counted as negative sales.

In similar fashion, in cell B19 the formula *SUMIF(Product,"lip gloss",Dollars)* tells us that $49,834.64 worth of lip gloss was sold. This calculation counts refunds associated with returns as negative revenue.

### What dollar amount of sales were not made by Jen?

In cell B20, the formula *SUMIF(Name,"<>Jen",Dollars)* sums up the dollar amount of each transaction that does not include *Jen* in the Name column. We find that saleswomen other than Jen sold $211,786.51 worth of makeup.

## Remarks

Unless you use array functions (see Chapter 72), the SUMIF function can handle only a single criteria. For example, if you do not use array functions, you cannot use SUMIF to answer a question that depends on two or more criteria, such as how many units of lip gloss Jen sold before June 10, 2005? The easiest way to answer questions involving multiple criteria is to use database statistical functions. See Chapter 37 for a discussion of database statistical functions.

## Problems

**1.** Determine the total number of units and dollar volume sold for each product.

2. Determine the total revenue earned before December 10, 2005.
3. Determine the total units sold by salespeople whose last name begins with C.
4. Determine total revenue earned by people having five letters in their name.
5. How many units were sold by people other than Colleen?
6. How many units of makeup were sold between January 15, 2004, and February 15, 2005?
7. The file NBA.xls contains the following information:
   - In columns A and B, the name of each NBA team and a code number for each team. For example, team 1 is Atlanta, and so on.
   - Column C lists the home team for each game.
   - Column D lists the visiting team for each game.
   - Column E lists points scored by the home team.
   - Column F lists points scored by the visiting team.

   From this data, compute for each team the average number of points the team scored per game and the average of how many points the team gave up.
8. The file Toysrus.xls contains sales revenue (in millions) during each quarter for the years 1997–2001 and the first two quarters of 2002. Use this data to compute a *seasonal index* for each quarter of the year. For example, the first quarter would have a seasonal index of .8 if average sales during the first quarter were 80 percent of the overall average sales per quarter.
9. How much revenue was made on sales transactions involving at least 100 units of makeup?
10. The file SumIfRows.xls contains sales data during several winter, summer, spring, and fall quarters. Determine average sales during the winter, summer, spring, and fall quarters.

# 21

# The OFFSET Function

- How can I create a reference to a rectangular range of cells that is a specified number of rows and columns from a cell or another range of cells?
- How can I perform a lookup operation by keying off the right-most column instead of the left-most column in a table range?
- I often download sales of a software product listed by country. I need to track revenues from Iran as well as costs and units sold, but the data about Iran isn't always in the same location in the spreadsheet. Can I create a formula that will always pick up Iran's revenues, costs, and units sold?
- Each drug developed by my company goes through three stages of development. I have a list of the cost by month for each drug, and I also know how many months each development stage took for each drug. Can I create formulas that compute for each drug the total cost incurred during each stage of development?
- I run a small video store. In a spreadsheet, my accountant has listed the name of each movie in stock and the number of copies in stock. Unfortunately, for each movie this information is in a single cell. How can I extract the number of copies of each movie in stock to a separate cell?
- I am charting my company's monthly unit sales. Each month I download the most recent month's unit sales. I would like my chart to automatically update. Is there an easy way to accomplish this?

The OFFSET function is used to create a reference to a rectangular range that is a specified number of rows and columns away from a cell or range of cells. Basically, to create a reference to a range of cells, you first specify a reference cell. You then indicate to Excel the number of rows and column away from the reference cell that you want to create your rectangular range. For example, by using the OFFSET function I can create a reference to a cell range that contains two rows and three columns and begins two columns to the right and one row above the current cell. The OFFSET function is very useful because you can copy it as you can any other Excel function, and the specified number of rows and columns you move from a reference cell can be calculated by using other Excel functions.

The syntax of the OFFSET function is

```
OFFSET(reference, rows moved, columns moved, [height], [width])
```

- *Reference* is a cell or range of cells from which the offset begins. If you specify a range of cells, the cells must be adjacent.
- *Rows moved* is the number of rows away from the reference that you want the range reference to start (the upper left cell in the offset range). A negative number of rows moves you up from the reference; a positive number of rows moves you down. For example, if *reference* equals C5 and *rows moved* equals -1, you move to row 4. If *rows moved* equals +1, you move to row 6. If *rows moved* equals 0, you stay at row 5.
- *Columns moved* is the number of columns away from the reference that you want the range reference to start (the upper left cell in the offset range). A negative number of columns moves you left from the reference; a positive number of columns moves you right. For example, if *reference* equals C5 and *columns moved* equals -1, you move to column B. If *columns moved* equals +1, you move to column D. If *columns moved* equals 0, you stay at column C.
- *Height* is an optional argument that gives the number of rows in the offset range. *Width* is an optional argument that gives the number of columns in the offset range. If *height* or *width* is omitted, the OFFSET function creates a range for which the value of *height* or *width* equals the height or width of *reference*.

## How can I create a reference to a rectangular range of cells that is a specified number of rows and columns from a cell or another range of cells?

The file OffsetExample.xls, shown in Figure 21-1, provides some examples of the OFFSET function in action.

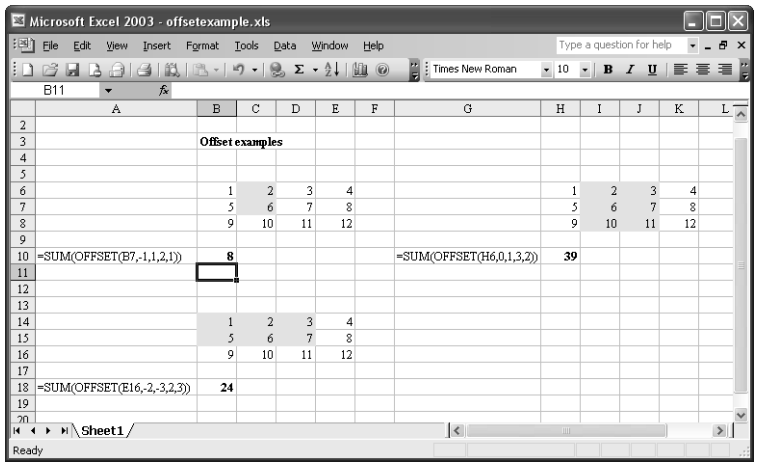



**Figure 21-1** Examples of using the OFFSET function.

For example, in cell B10 I entered the formula *SUM(OFFSET(B7,-1,1,2,1))*. This formula begins in cell B7. We move up one row and one column to the right, which brings us to cell C6. We now pick out a range consisting of 2 rows and 1 column, which yields the range C6:C7. The SUM function adds the numbers in this range, which yields 2 + 6 = 8. The other two examples shown in Figure 21-1 work the same way. In the following sections, we'll apply the OFFSET function to solve some problems that were sent to me by MBAs working at major U.S. companies.

## How can I perform a lookup operation keying off the right-most column instead of the left-most column in a table range?

In Figure 21-2, I've listed the members of the Dallas Mavericks NBA team and their field goal percentages. If I'm asked to "look up" the player with a given field goal percentage, I could easily solve that problem by using a VLOOKUP function. But what I really want to do is a "left-hand lookup," which involves finding the field goal percentage for a player by using his name. A VLOOKUP function can't perform a left-hand lookup, but a left-hand lookup is a snap if you combine the MATCH and OFFSET functions.

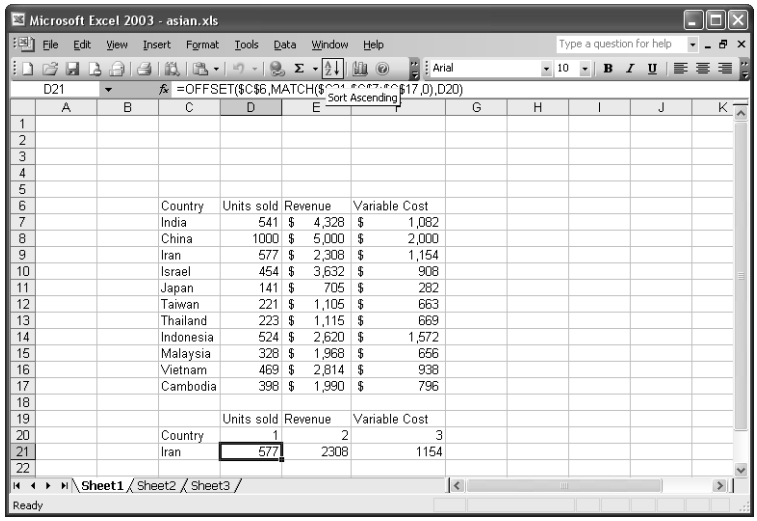
| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | **Left-hand lookup** | | |
| 6 | | | | Name | FG %age |
| 7 | | FG% | Player | Walt Williams | 0.397 |
| 8 | | 45.8% | Dirk Nowitzki | | |
| 9 | | 41.8% | Michael Finley | | |
| 10 | | 46.3% | Steve Nash | | |
| 11 | | 39.5% | Nick Van Exel | | |
| 12 | | 53.5% | Raef LaFrentz | | |
| 13 | | 60.2% | Eduardo Najera | | |
| 14 | | 51.2% | Shawn Bradley | | |
| 15 | | 39.7% | Walt Williams | | |
| 16 | | 44.4% | Adrian Griffin | | |
| 17 | | 48.4% | Avery Johnson | | |
| 18 | | 47.6% | Raja Bell | | |
| 19 | | 66.7% | Evan Eschmeyer | | |
| 20 | | 41.0% | Popeye Jones | | |
| 21 | | 40.0% | Mark Strickland | | |
| 22 | | 23.5% | Adam Harrington | | |

**Figure 21-2** You can do a left-hand lookup by using the MATCH and OFFSET functions.

I enter the player's name in cell D7. Then I use a reference cell of B7 in the OFFSET function. To find the player's field goal percentage, we need to move down to the row below row 7 where the player's name appears. This is a job for the MATCH function. The portion *MATCH(D7,$C$8:$C$22,0)* of the formula *OFFSET(B7,MATCH(D7,$C$8:$C$22,0),0)* picks up the number of rows below row 7 where our player's name occurs. Then we move that many rows below row 7 and move over no columns from column B. Because the reference consists of a single cell, omitting the *height* and *width* arguments of the OFFSET function ensures that the range returned by this formula consists of one row and column (a single cell). Thus we pick up the player's field goal percentage.

## I often download sales of a software product listed by country. I need to track revenues from Iran as well as costs and units sold, but the data about Iran isn't always in the same location in the spreadsheet. Can I create a formula that will always pick up Iran's revenues, costs, and units sold?

The file AsianSales.xls (see Figure 21-3) contains data for the units sold, sales revenue, and variable cost for software sold to several countries in Asia and the Middle East. Each month, when we download the monthly financials, the location of each country in the spreadsheet changes, so we want formulas that will always return (for a given country) the correct units sold, revenue, and variable cost.



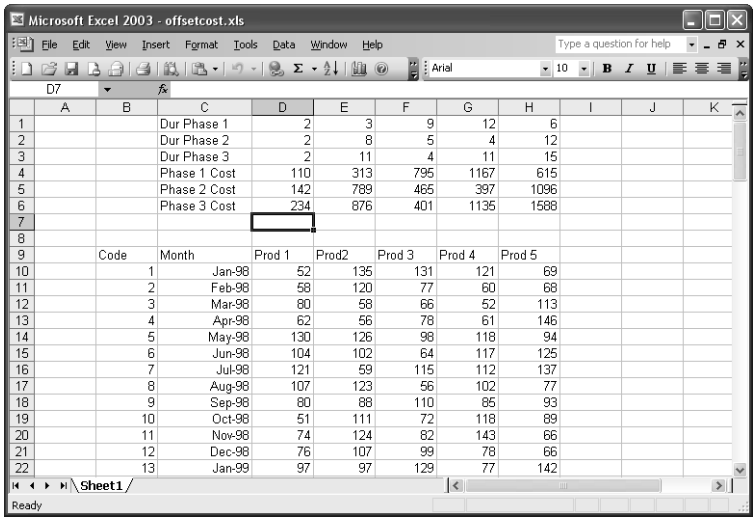

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | | | | | | | | | | | |
| 4 | | | | | | | | | | | |
| 5 | | | | | | | | | | | |
| 6 | | | Country | Units sold | Revenue | Variable Cost | | | | | |
| 7 | | | India | 541 | $ 4,328 | $ 1,082 | | | | | |
| 8 | | | China | 1000 | $ 5,000 | $ 2,000 | | | | | |
| 9 | | | Iran | 577 | $ 2,308 | $ 1,154 | | | | | |
| 10 | | | Israel | 454 | $ 3,632 | $ 908 | | | | | |
| 11 | | | Japan | 141 | $ 705 | $ 282 | | | | | |
| 12 | | | Taiwan | 221 | $ 1,105 | $ 663 | | | | | |
| 13 | | | Thailand | 223 | $ 1,115 | $ 669 | | | | | |
| 14 | | | Indonesia | 524 | $ 2,620 | $ 1,572 | | | | | |
| 15 | | | Malaysia | 328 | $ 1,968 | $ 656 | | | | | |
| 16 | | | Vietnam | 469 | $ 2,814 | $ 938 | | | | | |
| 17 | | | Cambodia | 398 | $ 1,990 | $ 796 | | | | | |
| 18 | | | | | | | | | | | |
| 19 | | | | Units sold | Revenue | Variable Cost | | | | | |
| 20 | | | Country | 1 | 2 | 3 | | | | | |
| 21 | | | Iran | 577 | 2308 | 1154 | | | | | |
| 22 | | | | | | | | | | | |

**Figure 21-3** You can use the OFFSET function in calculations when you're working with data that isn't in a fixed location in a spreadsheet.

By copying from D21 to E21:F21 the formula *OFFSET($C$6,MATCH ($C21,$C$7:$C$17,0),D20)* we compute the result we want. This formula sets *reference* equal to cell C6 (which contains the word *Country*). Then it moves over one column (the value in cell D20) to pick up units sold and down the number of rows that match the number of rows below row 6 in which the country listed in C21 is found. In cell E21, the reference to D20 now refers to E20 and becomes a 2, so we move over two columns to the right of column C and pick up revenue. In cell E21, the reference to D20 now refers to F20 and becomes a 3, so we move three columns to the right of column C and pick up variable cost.

### Each drug developed by my company goes through three stages of development. I have a list of the cost by month for each drug, and I also know how many months each development stage took for each drug. Can I create formulas that compute for each drug the total cost incurred during each stage of development?

The file OffsetCost.xls contains the monthly costs incurred in developing five drugs. Each drug goes through three stages of development. For each drug, the data also includes the number of months required to complete each phase. A subset of the data is shown in Figure 21-4.




| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | Dur Phase 1 | 2 | 3 | 9 | 12 | 6 | | | |
| 2 | | | Dur Phase 2 | 2 | 8 | 5 | 4 | 12 | | | |
| 3 | | | Dur Phase 3 | 2 | 11 | 4 | 11 | 15 | | | |
| 4 | | | Phase 1 Cost | 110 | 313 | 795 | 1167 | 615 | | | |
| 5 | | | Phase 2 Cost | 142 | 789 | 465 | 397 | 1096 | | | |
| 6 | | | Phase 3 Cost | 234 | 876 | 401 | 1135 | 1588 | | | |
| 7 | | | | | | | | | | | |
| 8 | | | | | | | | | | | |
| 9 | | Code | Month | Prod 1 | Prod2 | Prod 3 | Prod 4 | Prod 5 | | | |
| 10 | | 1 | Jan-98 | 52 | 135 | 131 | 121 | 69 | | | |
| 11 | | 2 | Feb-98 | 58 | 120 | 77 | 60 | 68 | | | |
| 12 | | 3 | Mar-98 | 80 | 58 | 66 | 52 | 113 | | | |
| 13 | | 4 | Apr-98 | 62 | 56 | 78 | 61 | 146 | | | |
| 14 | | 5 | May-98 | 130 | 126 | 98 | 118 | 94 | | | |
| 15 | | 6 | Jun-98 | 104 | 102 | 64 | 117 | 125 | | | |
| 16 | | 7 | Jul-98 | 121 | 59 | 115 | 112 | 137 | | | |
| 17 | | 8 | Aug-98 | 107 | 123 | 56 | 102 | 77 | | | |
| 18 | | 9 | Sep-98 | 80 | 88 | 110 | 85 | 93 | | | |
| 19 | | 10 | Oct-98 | 51 | 111 | 72 | 118 | 89 | | | |
| 20 | | 11 | Nov-98 | 74 | 124 | 82 | 143 | 66 | | | |
| 21 | | 12 | Dec-98 | 76 | 107 | 99 | 78 | 66 | | | |
| 22 | | 13 | Jan-99 | 97 | 97 | 129 | 77 | 142 | | | |

**Figure 21-4** Using the OFFSET function to compute development costs for phases 1–3.

The goal is to determine for each drug the total cost incurred during each development phase. In cells D4:D6, I compute the total development costs for phases 1–3 for drug 1. I compute phase 1 costs for drug 1 by using a cell reference of D10, with *rows moved* and *columns moved* equal to 0. Setting *height* equal to the number of months in phase 1 and *width* equal to 1 captures all first-phase costs. I compute phase 1 costs for drug 1 in cell D4 with the formula *SUM(OFFSET(D10,0,0,D1,1))*. Next, in cell D5, I compute phase 2 total costs for drug 1 by using the formula *SUM(OFFSET(D10,D1,0,D2,1))*. Note that I start with a cell reference of D10 (the first month of costs) and move down the number of rows equal to the length of phase 1. This brings me to the beginning of phase 2. Setting *height* equal to the value in cell D2 ensures that we pick up all phase 2 costs. Finally, in cell D6, I pick up the phase 3 development costs for drug 1 with the formula *SUM(OFFSET(D10,D1+D2,0,D3,1))*. In this formula, I start from the first month of sales and move down the number of rows equal to the total time needed for phases 1 and 2. This brings us to the beginning of phase 3, and we add up the number of rows in cell D3 to capture phase 3 costs. Then, by copying the formulas in D4:D6 to E4:H6, I can compute total costs for phases 1–3 for drugs 2 through 5. For example, we find that for drug 2, total phase 1 costs equal $313, total phase 2 costs equal $789, and total phase 3 costs equal $876.

### I run a small video store. In a spreadsheet, my accountant has listed the name of each movie in stock and the number of copies in stock. Unfortunately, for each movie this information is in a single cell. How can I extract the number of copies of each movie in stock to a separate cell?

The file Movies.xls, shown in Figure 21-5, contains the movies in stock and the number of copies of each movie in stock.

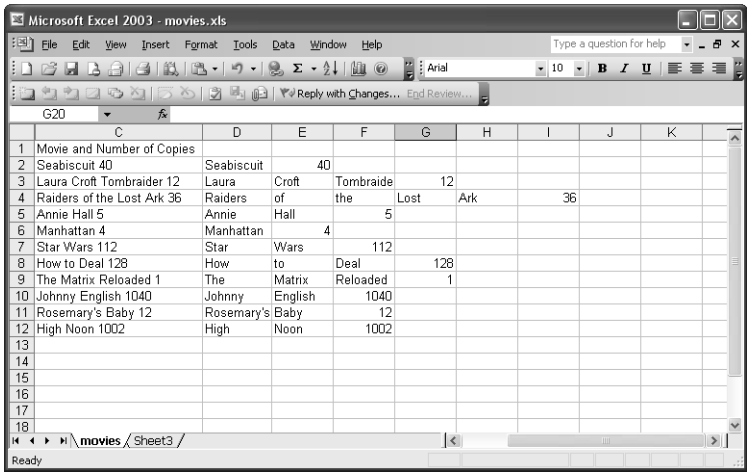
| | C |
|---|---|
| 1 | Movie and Number of Copies |
| 2 | Seabiscuit 40 |
| 3 | Laura Croft Tombraider 12 |
| 4 | Raiders of the Lost Ark 36 |
| 5 | Annie Hall 5 |
| 6 | Manhattan 4 |
| 7 | Star Wars 112 |
| 8 | How to Deal 128 |
| 9 | The Matrix Reloaded 1 |
| 10 | Johnny English 1040 |
| 11 | Rosemary's Baby 12 |
| 12 | High Noon 1002 |

**Figure 21-5** Movies owned by a video store, with data contained in a single cell.

We want to extract the number of copies owned of each movie to a separate cell. If the number of copies were listed to the left of a movie's title, this problem would be easy. We could use the FIND function to locate the first space and then use the LEFT function to pick up all the data to the left of the first space. (See Chapter 6 for a discussion of how to use the LEFT and FIND functions as well as other functions you can use to work with text.) Unfortunately, this technique doesn't work when the number of copies is listed to the right of the movie title. For a one-word movie title, for example, the number of copies is to the right of the first space, but for a four-word movie title, the number of copies is to the right of the fourth space.

One way to solve this problem is to use the Data, Text To Columns command to place each word in a title and the number of copies in separate columns. We can then use the COUNTA function to count the total number of words in a title, including the number of items as a word, for each movie. We can then use the OFFSET function to "pick off" the number of items.

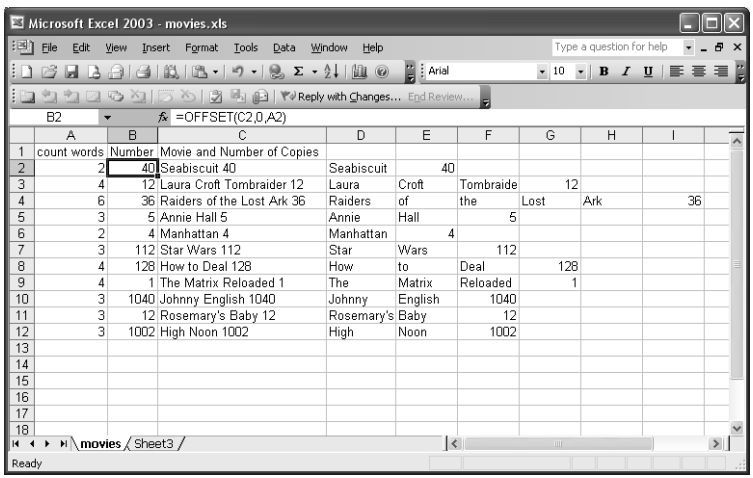
To begin, insert enough columns to the right of the data to allow each word in the movies' titles and the number of items to be extracted to a separate column. I used six columns (*Raiders of the Lost Ark* requires 6 columns), as you can see in Figure 21-6. Then I select the cell range C2:C12 and chose Data, Text To Columns. I select Delimited in the Convert Text To Columns Wizard and use the space character as the delimiting character. After selecting cell D2 as the destination cell, I have the results shown in Figure 21-6.




| | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Movie and Number of Copies | | | | | | | | |
| 2 | Seabiscuit 40 | Seabiscuit | 40 | | | | | | |
| 3 | Laura Croft Tombraider 12 | Laura | Croft | Tombraide | 12 | | | | |
| 4 | Raiders of the Lost Ark 36 | Raiders | of | the | Lost | Ark | 36 | | |
| 5 | Annie Hall 5 | Annie | Hall | 5 | | | | | |
| 6 | Manhattan 4 | Manhattan | 4 | | | | | | |
| 7 | Star Wars 112 | Star | Wars | 112 | | | | | |
| 8 | How to Deal 128 | How | to | Deal | 128 | | | | |
| 9 | The Matrix Reloaded 1 | The | Matrix | Reloaded | 1 | | | | |
| 10 | Johnny English 1040 | Johnny | English | 1040 | | | | | |
| 11 | Rosemary's Baby 12 | Rosemary's | Baby | 12 | | | | | |
| 12 | High Noon 1002 | High | Noon | 1002 | | | | | |
| 13 | | | | | | | | | |
| 14 | | | | | | | | | |
| 15 | | | | | | | | | |
| 16 | | | | | | | | | |
| 17 | | | | | | | | | |
| 18 | | | | | | | | | |

**Figure 21-6** The Data To Text command extracts data in a single cell to multiple cells.

Now we count the number of words, including the number of items as one word, for each movie by copying from A2 to A3:A12 the formula *COUNTA(D2:I2)*. The results are shown in Figure 21-7.




| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | count words | Number | Movie and Number of Copies | | | | | | |
| 2 | 2 | 40 | Seabiscuit 40 | Seabiscuit | 40 | | | | |
| 3 | 4 | 12 | Laura Croft Tombraider 12 | Laura | Croft | Tombraide | 12 | | |
| 4 | 6 | 36 | Raiders of the Lost Ark 36 | Raiders | of | the | Lost | Ark | 36 |
| 5 | 3 | 5 | Annie Hall 5 | Annie | Hall | 5 | | | |
| 6 | 2 | 4 | Manhattan 4 | Manhattan | 4 | | | | |
| 7 | 3 | 112 | Star Wars 112 | Star | Wars | 112 | | | |
| 8 | 4 | 128 | How to Deal 128 | How | to | Deal | 128 | | |
| 9 | 4 | 1 | The Matrix Reloaded 1 | The | Matrix | Reloaded | 1 | | |
| 10 | 3 | 1040 | Johnny English 1040 | Johnny | English | 1040 | | | |
| 11 | 3 | 12 | Rosemary's Baby 12 | Rosemary's | Baby | 12 | | | |
| 12 | 3 | 1002 | High Noon 1002 | High | Noon | 1002 | | | |
| 13 | | | | | | | | | |
| 14 | | | | | | | | | |
| 15 | | | | | | | | | |
| 16 | | | | | | | | | |
| 17 | | | | | | | | | |
| 18 | | | | | | | | | |

**Figure 21-7** Final result for the movie example.

Finally, copying from B2 to B3:B12 the formula *OFFSET(C2,0,A2)*, I can pick up the number of copies of each movie in stock. This formula begins at the cell reference containing the movie title and moves over the number of columns equal to the number of "words" in the title cell. Because the cell reference

contains only one cell, we can omit the last two arguments of the OFFSET function so that the function picks up only the single cell containing the last "word" (which is the number of copies) of the title cell.

### I am charting my company's monthly unit sales. Each month I download the most recent month's unit sales. I would like my chart to update automatically. Is there an easy way to accomplish this?

The workbook DynamicRange.xls (see Figure 21-8) contains units sold for our company's product. As you can see, the units sold have been charted using an XY (Scatter) chart.

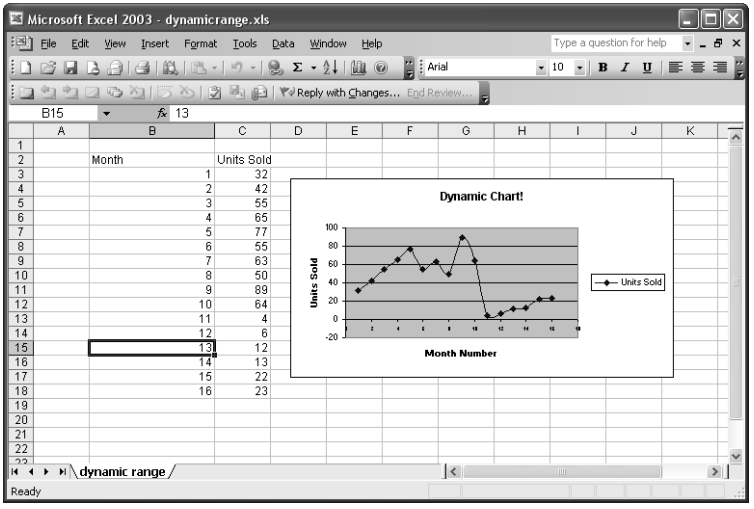



**Figure 21-8** We can use the OFFSET function to update this chart dynamically.

Beginning in row 19, we will download new sales data. Is there an easy way to ensure that the chart automatically includes the new data?

The key to updating the chart is to use the OFFSET function to create a dynamic range name for both the Months column and the Units Sold column. As new data is entered, the dynamic range for unit sales will automatically include all sales data and the dynamic range for months will include each month number. After creating these ranges, I modify the chart, replacing the data ranges used in the chart with the dynamic ranges. The chart will now be updated as new data is entered.

To begin, select Insert, Name Define and create a range named Units by filling in the dialog box as shown in Figure 21-9.

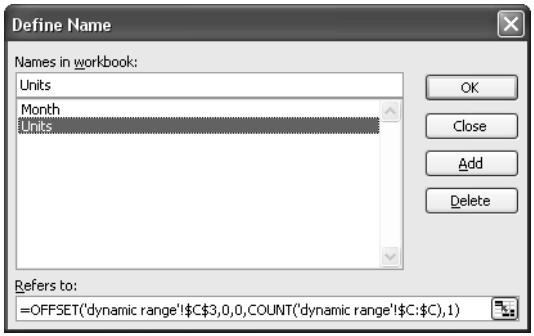



**Figure 21-9** Creating dynamic range name for the units sold.

Entering *=OFFSET('dynamic range'!$C$3,0,0,COUNT('dynamic range'!$C:$C),1)* in the Refers To area of the dialog box creates a range one column wide beginning in cell C3, which contains the first units sales data point. The range will contain as many numbers as there are in column C, which is derived by the portion of the formula that reads *COUNT('dynamic range'!$C:$C)*. As new data is entered into column C, the data is automatically included in the range named Units.

Next we create a dynamic range named Months for the months entered in column B. The formula is shown in Figure 21-10.

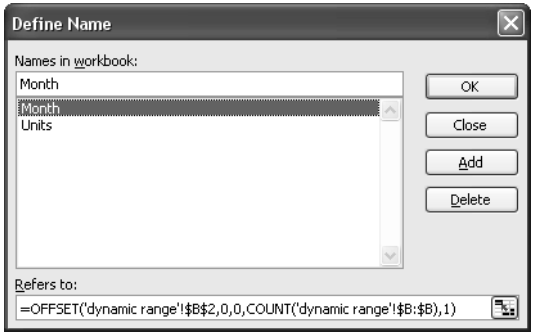



**Figure 21-10** The formula used to define a dynamic range named Month.

Now we go to our chart and click on any selected point. In the formula box you'll see the formula *SERIES('dynamic range'!$C$2,'dynamic range'!$B$3:$B$18,'dynamic range'!$C$3:$C$18,1)*. This formula is Excel's version of the data used to originally set up the chart. Replace the ranges $B$3:$B$18 and $C$3:$C18 with our dynamic range names as follows: *SERIES('dynamic range'!$C$2,dynamicrange.xls!Month,dynamicrange.xls!Units,1)*. Of course, if a blank space is listed above any new data, this method won't work. Enter some new data and you'll see that the data is included in the chart.

# Problems

1. The file C21p1.xls supplies data about the units sold for 11 products during the years 1999–2003. Write a formula using the MATCH and OFFSET functions that picks up the sales of a given product during a given year. Can you think of another way to solve this problem without using the MATCH and OFFSET functions?

2. A commonly suggested moving average trading rule is to buy a stock when the stock's price moves above the average of the last *D* months and to sell a stock when the stock's price moves below the average of the last *D* month's price. In Chapter 12, we showed that for *D* = *15*, this trading rule outperformed the S&P by a substantial amount. By combining a one-way data table with the OFFSET function, determine the value of *D* that maximizes trading profit (excluding transactions costs). You can find data in the file MaTradingRule.xls.

3. A commonly suggested moving average trading rule is to buy a stock when the stock's price moves above the average of the last *B* months and to sell a stock when the stock's price moves below the average of the last *S* month's price. In Chapter 12, we showed that for *B* = *S* = *15*, this trading rule outperformed the S&P by a substantial amount. By combining a two-way data table with the OFFSET function, determine the values of *B* and *S* that maximize trading profit (excluding transactions costs). Data is in the file MaTradingRule.xls.

4. The file Lagged.xls contains data about the number of magazine ads placed by Army recruiting during each of 60 consecutive months. For each month, we define the *k* month lagged number of ads to equal the number of ads placed *k* months ago. For months 7–60, we would like to compute the 1-month lagged, 2-month lagged… 6-month lagged values of the number of ads. Use the OFFSET function to efficiently compute these lagged values.

5. Suppose you are continually entering monthly sales in column C of a spreadsheet. The first month's sales is entered in cell C2. Assume that as data is entered, no rows are left blank. Thus, if the last month of sales is in cell C50, the next month's sales will be entered in cell C51, and so on. Write a formula that will always return sales during the most recent month. Hint: The expression *COUNTA(C:C)* will always yield the number of nonblank entries in column C.

# 22

# The INDIRECT Function

- My spreadsheet formulas often contain references to cells or ranges or both. Rather than change these references in my formulas, I'd like to know how I can place the references in their own cells so that I can easily change my cell or range references without changing my underlying formulas.
- Each worksheet in a workbook lists monthly sales of a product in cell D1. Is there an easy way to write and copy a formula that lists each month's product sales in a single worksheet?
- Suppose I'm adding up the values in the range A5:A10 with the formula *SUM(A5:A10)*. If I insert a blank row somewhere between row 5 and row 10, my formula updates itself to read *SUM(A5:A11)*. Can I write a formula so that when I insert a blank row between row 5 and row 10, my formula still adds the values in the range A5:A10?

The INDIRECT function is probably one of the most difficult functions to master in Excel. Knowing how to use the INDIRECT function, however, enables you to solve many seemingly unsolvable problems. Essentially any reference to a cell within the INDIRECT portion of a formula results in the cell reference being immediately evaluated to equal the content of the cell. To illustrate the use of the INDIRECT function, look at the file IndirectSimpleEx.xls, which is shown in Figure 22-1.

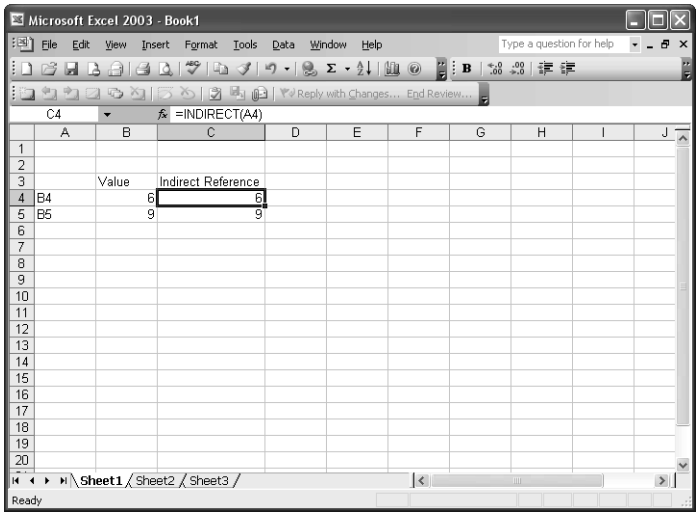


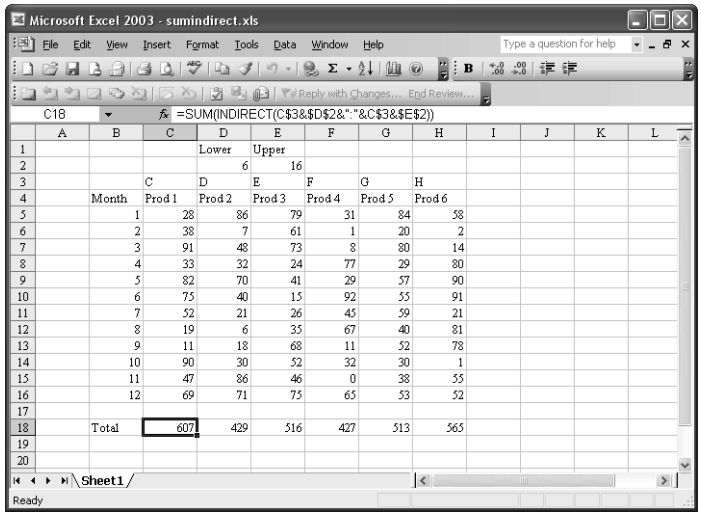

**Figure 22-1** A simple example of the INDIRECT function.

In cell C4, I've entered the formula *=INDIRECT(A4)*, and Excel returns a value of 6. Excel returns this value because the reference to A4 is immediately replaced by the text string *B4*. Therefore, the formula is evaluated as *=B4*, which yields a value of 6. Similarly, entering in cell C5 the formula *=INDIRECT(A5)* returns the value in cell B5, which is 9.

### My spreadsheet formulas often contain references to cells or ranges or both. Rather than change these references in my formulas, I'd like to know how I can place the references in their own cells so that I can easily change my cell or range references without changing my underlying formulas.

In this example, the data we'll use is contained in the file SumIndirect.xls, shown in Figure 22-2. The cell range B4:H16 lists monthly sales data for six products during a 12-month period.

I currently compute total sales of each product during months 2–12. An easy way to make this calculation is to copy from C18 to D18:H18 the formula *SUM(C6:C16)*. Suppose, however, that you want to be able to change which months are totaled. For example, you might want to total sales for months 3–12. You could change the formula in cell C18 to *SUM(C7:C16)* and then copy this formula to D18:H18, but using this approach is a pain because you have to copy the formula in C18 to D18:H18 and, without looking at the formulas, nobody knows which rows are being added.




| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | Lower | Upper | | | | | | | |
| 2 | | | | 6 | 16 | | | | | | | |
| 3 | | | C | D | E | F | G | H | | | | |
| 4 | | Month | Prod 1 | Prod 2 | Prod 3 | Prod 4 | Prod 5 | Prod 6 | | | | |
| 5 | | 1 | 28 | 86 | 79 | 31 | 84 | 58 | | | | |
| 6 | | 2 | 38 | 7 | 61 | 1 | 20 | 2 | | | | |
| 7 | | 3 | 91 | 48 | 73 | 8 | 80 | 14 | | | | |
| 8 | | 4 | 33 | 32 | 24 | 77 | 29 | 80 | | | | |
| 9 | | 5 | 82 | 70 | 41 | 29 | 57 | 90 | | | | |
| 10 | | 6 | 75 | 40 | 15 | 92 | 55 | 91 | | | | |
| 11 | | 7 | 52 | 21 | 26 | 45 | 59 | 21 | | | | |
| 12 | | 8 | 19 | 6 | 35 | 67 | 40 | 81 | | | | |
| 13 | | 9 | 11 | 18 | 68 | 11 | 52 | 78 | | | | |
| 14 | | 10 | 90 | 30 | 52 | 32 | 30 | 1 | | | | |
| 15 | | 11 | 47 | 86 | 46 | 0 | 38 | 55 | | | | |
| 16 | | 12 | 69 | 71 | 75 | 65 | 53 | 52 | | | | |
| 17 | | | | | | | | | | | | |
| 18 | | Total | 607 | 429 | 516 | 427 | 513 | 565 | | | | |
| 19 | | | | | | | | | | | | |
| 20 | | | | | | | | | | | | |

**Figure 22-2** You can use the INDIRECT function to change cell references in formulas without changing the formulas themselves.

The INDIRECT function provides another approach. I've placed in cells D2 and E2 the starting and ending rows of our summation. Then, using the INDIRECT function, all I need to do is change the starting and ending row references in D2 and E2 and the sums are updated to include the rows we want. Also, by looking at the values in D2 and E2, it is obvious which rows (months) are being added! All I need to do is copy from C18 to D18:H18 the formula *SUM(INDIRECT(C$3&$D$2&":"&C$3&$E$2))*. Every cell reference within the INDIRECT portion of this formula is evaluated to equal the contents of the cell. *C$3* is evaluated as C, *$D$2* is evaluated as 6, and *$E$2* is evaluated as 16. Using the concatenation symbol (&), Excel evaluates this formula as *SUM(C6:C16)*, which is exactly want we want. The formula in C18 returns the value 38 + 91 + … 69 = 607. In cell D18, our formula evaluates as *SUM(D6:D16)*, which is the result we want. Of course, if we want to add up sales during, say, months 4 through 6, we simply enter 8 in D2 and 10 in cell E2. Then the formula in C18 returns 33 + 82 + 75 = 190. (For information about using the ampersand to concatenate values, see Chapter 6.)

## Each worksheet in a workbook lists monthly sales of a product in cell D1. Is there an easy way to write and copy a formula that lists each month's product sales in a single worksheet?

The file IndirectMultiSheet.xls (see Figure 22-3) contains seven worksheets. In each worksheet, cell D1 contains data about the sales of a product during a par-

ticular month. Let's suppose Sheet1 contains month 1 sales, Sheet 2 contains month 2 sales, and so on. For instance, sales in month 1 equals 1.

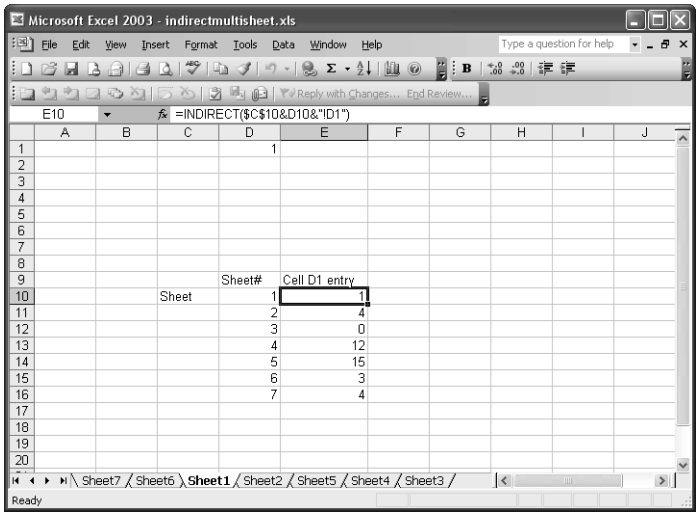



**Figure 22-3** Monthly sales (months 1–7) of a product listed by using the INDIRECT function.

Suppose you want to list each month's sales in a single worksheet. A tedious approach would be to list month 1 sales with the formula *=Sheet1!D1*, list month 2 sales with the formula *=Sheet2!D1*, and so on until you've listed month 7 sales with the formula *=Sheet7!D1*. If you have 100 months of data, this approach would be a royal pain! A much more elegant approach is to list month 1 sales in cell E10 of Sheet1 with the formula *INDIRECT($C$10&D10&"!D1")*. Excel evaluates C10 as "Sheet", D10 as 1, and "!D1" as the text string *!D1*. The whole formula is evaluated as *=Sheet1!D1*, which, of course, yields month 1 sales, located in cell D1 of Sheet1. Copying this formula to the range E11:E16 lists the entries in cell D1 of sheets 2 through 7. Note that when the formula in cell E10 is copied to cell E11, the reference to D10 changes to D11, and cell E11 returns the value located at Sheet2!D1 and so on.

### Suppose I'm adding up the values in the range A5:A10 with the formula *SUM(A5:A10)*. If I insert a blank row somewhere between row 5 and row 10, my formula updates itself to read *SUM(A5:A11)*. Can I write a formula so that when I insert a blank row between row 5 and row 10, my formula still adds the values in the range A5:A10?

The worksheet named Sum(A5A10) in the file IndirectInsertRow.xls (shown in Figure 22-4) illustrates several ways to add up the numbers in the cell range

A5:A10. In cell A12, I've entered the traditional formula *SUM(A5:A10)*. This formula yields 6 + 7 + 8 + 9 + 1 + 2 = 33.

Similarly, the formula *SUM($A$5:$A$10)* in cell E9 yields a value of 33. As you'll soon see, however, if we insert a row between rows 5 and 10, both formulas will attempt to add up the cells in the range A5:A11.

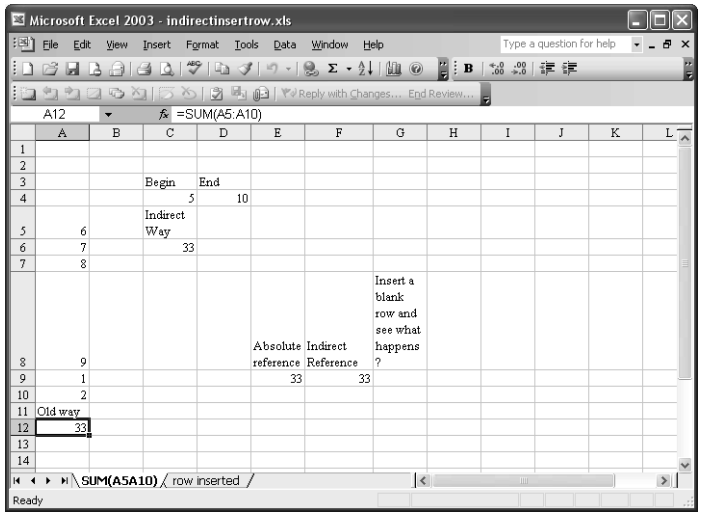



**Figure 22-4** This spreadsheet shows several ways to sum the values in the cell range A5:A10.

With the INDIRECT function, you have at least two ways to sum the values in the range A5:A10. In cell F9, I've entered the formula *SUM(INDIRECT("A5:A10"))*. Because Excel treats *INDIRECT("A5:A10")* as the text string "A5:A10", if I insert a row in the spreadsheet, this formula still adds up the entries in the cell range A5:A10!

Another way to use the INDIRECT function to add up the entries in the range A5:A10 is the formula *SUM(INDIRECT("A"&C4&":A"&D4))*, which is the formula entered in cell C6. Excel treats the reference to C4 as a 5 and the reference to D4 as a 10, so this formula becomes *SUM(A5:A10)*. Inserting a blank row between row 5 and row 10 has no effect on this formula because the reference to C4 will still be treated as a 5 and the reference to D4 will still be treated as a 10. In Figure 22-5, you can see the sums computed by our four formulas after a blank row is inserted below row 7. You can find this data on the worksheet Row Inserted in the file IndirectInsertRow.xls.

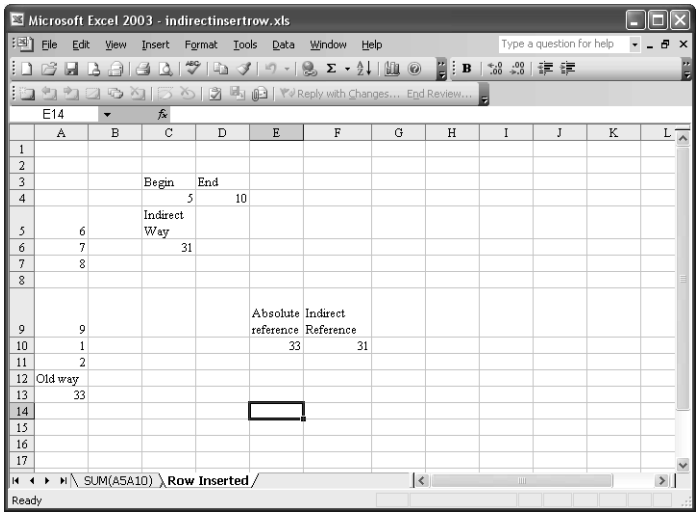



**Figure 22-5** Results of SUM formulas after a blank row is inserted below row 7.

Note that the classic SUM formulas that do not involve the INDIRECT function have changed to add up the entries in the range A5:A11, so these formulas still yield a value of 33. The two SUM formulas involving the INDIRECT function continue to add up the entries in the range A5:A10, so we lose the value of 2 (now in cell A11) when we compute our sum. The SUM formulas that use the INDIRECT function yield a value of 31.

# Problems

1. The ADDRESS function yields the actual cell address associated with a row and column. For example, the formula *ADDRESS(3,4)* yields $D$3. What result would be obtained if you entered the formula *=INDIRECT(ADDRESS(3,4))*?
2. The workbook P22_2.xls contains data for the sales of five products in four regions (East, West, North, and South). Use the INDIRECT function to create formulas that enable you to easily add up the total sales of any combination of consecutively numbered products, such as Products 1–3, Products 2–5, and the like.
3. The file P22_3.xls contains six worksheets. Sheet1 contains month 1 sales for products 1–4. These sales are always listed in the range E5:H5. Use the INDIRECT function to efficiently tabulate the sales of each product by month in a separate worksheet.
4. Write a formula that will add up the entries in the cell range G2:K2 even if one or more columns are inserted between columns G and K.

# 23

# Conditional Formatting

- How can I highlight monthly stock returns so that every good month is highlighted in one color and every bad month is highlighted in another?
- Given quarterly corporate revenues, how can I highlight quarters in which revenues increased over the previous quarter in one color and quarters in which revenues decreased from the previous quarter in another?
- Given a list of dates, how can I highlight the weekend dates in a specific color?
- Our basketball coach has given each player a rating between 1 and 10 for her ability to play guard, forward, or center. Can I set up a spreadsheet that highlights the ability of each player to play the position to which she's assigned?

Conditional formatting lets you specify formatting for a cell range depending on the contents of the cell range. For example, given exam scores for students, you can use conditional formatting to highlight in red the names of students who have a final average of at least 90. For any cell, conditional formatting allows you to create up to three conditions and define formatting for each condition. Basically, when you set up conditions to format a range of cells, Excel checks each cell in the range to see whether any of the conditions you created (such as *exam score > 90*) is satisfied. The format you link to that condition is applied to all cells in the range that satisfy the condition. If the content of the cell does not satisfy any of the conditions, the formatting of the cell is unchanged.

## How can I highlight monthly stock returns so that every good month is highlighted in one color and every bad month is highlighted in another?

The file SandP.xls, shown in Figure 23-1, contains monthly values and returns on the Standard and Poor's stock index. Suppose that you want to highlight in purple each month in which the S&P index went up more than 3 percent and highlight in red each month in which the S&P went down more than 3 percent.

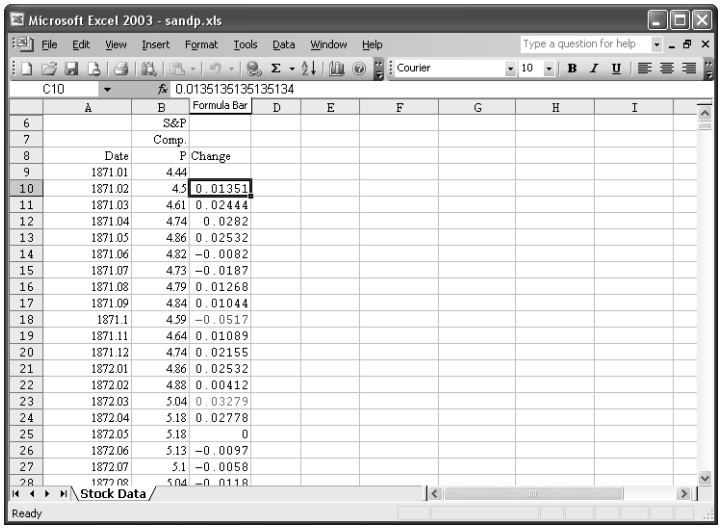



**Figure 23-1** Conditional formatting highlights returns in the S&P stock index.

In this example, I begin by moving to cell C10 (the first month containing an S&P return) and select all monthly returns by pressing Ctrl+Shift+Down Arrow. Next I choose Conditional Formatting from the Format menu. In the Conditional Formatting dialog box, shown in Figure 23-2, I select Cell Value Is and then Greater Than and then fill in .03. These settings mean that the formatting we select will be applied to all cells in the range that contain a value greater than .03.

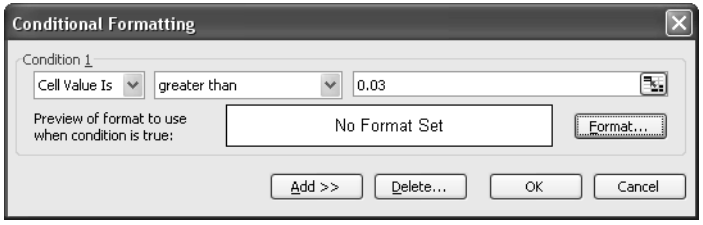



**Figure 23-2** Setting up the condition that will apply special formatting to S&P returns greater than .03.

Now click the Format button, and then select purple (or whatever color or other formatting you'd like) in the Format Cells dialog box, shown in Figure 23-3.

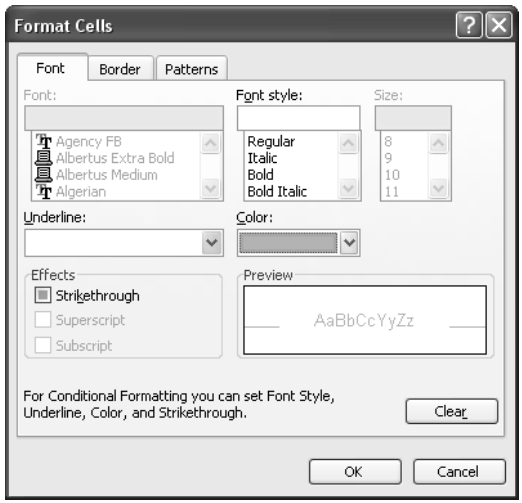



**Figure 23-3** You use the Format Cells dialog box to select the formatting that's applied to the cells that meet the condition you defined.

Notice that the lists for fonts and font size aren't available, so your choice of formatting can't change the font or font size. The Patterns tab provides the option to shade cells in a color you choose, while the Borders tab lets you create a border for cells that satisfy your conditional criteria. After clicking OK in the Format Cells dialog box, you're returned to the Conditional Formatting dialog box, which now appears as it's shown in Figure 23-4.

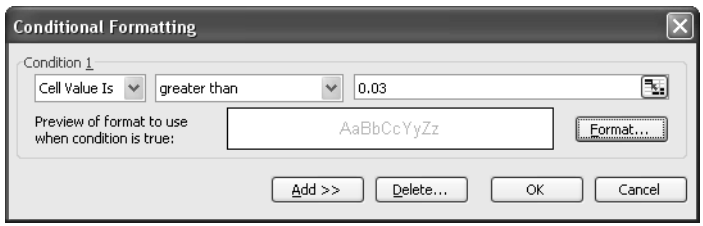



**Figure 23-4** All months with S&P return greater than 3 percent will be colored in purple.

We could stop here, click OK, and months meeting our condition would be highlighted in purple. To define a second condition (in a fashion similar to the first), click Add and then specify that all months in which the S&P return is

less than -3 percent will be colored red. The Conditional Formatting dialog box should now appear as it's shown in Figure 23-5.

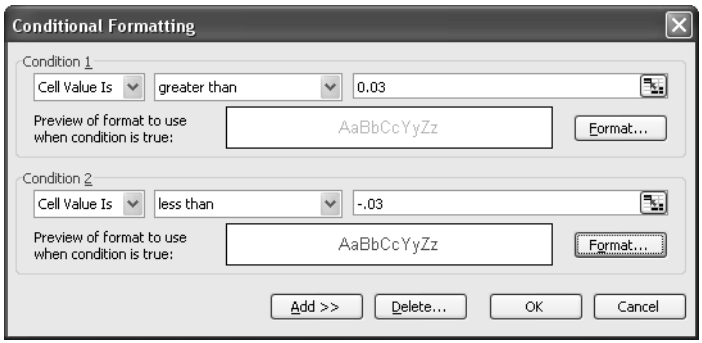



**Figure 23-5** You can define as many as three conditions for a specific cell or cell range.

When you now click OK, all months with an S&P return that's greater than 3 percent (see cell C23, for example) are displayed in purple, and all months with an S&P return less than -3 percent (see cell C18) are displayed in red. Cells in which the monthly returns don't meet either of our conditions maintain their original formatting.

By the way, to delete conditional formatting (or any format) applied to a range of cells, simply select the range of cells and choose Edit, Clear Formats. To select all the cells in a spreadsheet to which conditional formatting applies, press F5 to open the Go To dialog box. In the dialog box, click the Special button, select Conditional Formats, and then click OK.

## Given quarterly corporate revenues, how can I highlight quarters in which revenues increased over the previous quarter in one color and quarters in which revenues decreased from the previous quarter in another?

The file ToysrusFormat.xls contains quarterly revenues (in millions) for Toys "R" Us during the years 1997–2002. (See Figure 23-6.) We'd like to highlight quarters in which revenues increased over the previous quarter in one color and highlight quarters in which revenues decreased over the previous quarter in another.

The Formula Is option in the Conditional Formatting dialog box enables you to use a formula to define conditions that Excel checks before it applies formatting to a cell. We'll use this option in this example, but before we work with the Formula Is option, let's look at how Excel evaluates some logical functions.

What happens when we type a formula such as *=B3<2* in cell B4? If the value in B3 is a number smaller than 2, Excel returns the value True in cell B4; otherwise, Excel returns False. You can refer to the file LogicalExamples.xls,

shown in Figure 23-7, for other examples like this. As you can see in Figure 23-7, you can also use combinations of AND, OR, and NOT in formulas.

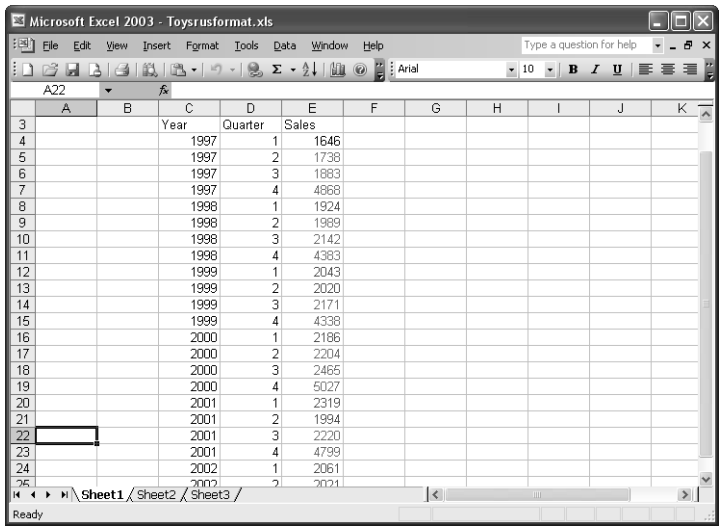



**Figure 23-6** In this spreadsheet, we've highlighted increased sales in purple and decreased sales in red.

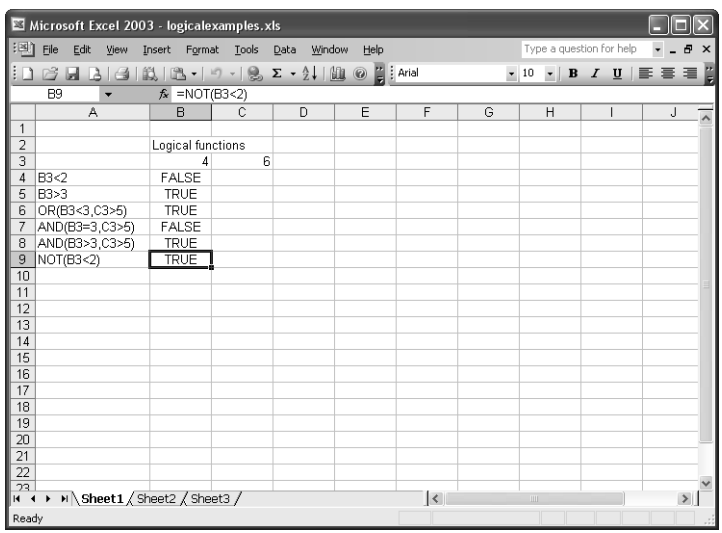



**Figure 23-7** Examples of logical functions.

- In cell B6, the formula *=OR(B3<3, C3>5)* returns the value True if either of the conditions B3<3 or C3>5 is true. Since the value of C3 is greater than 5, Excel returns True.

- In cell B7, the formula *=AND(B3=3,C3>5)* returns True if B3=3 and C3>5. Because B3 is not equal to 3, Excel returns False. In cell B8, however, the formula *=AND(B3>3,C3>5)* returns True because B3>3 and C3>5 are both true.
- In cell B9, the formula *=NOT(B3<2)* returns True because B3<2 would return False, and Not False becomes True.

Now let's look at how to use the Formula Is feature to create a conditional format in a range of cells. Begin by selecting the range of cells to which the conditional format should be applied. Open the Conditional Formatting dialog box, select Formula Is, and then enter the formula that defines the condition (it must start with an equals sign). The formula should be relative to the cell that's selected in the range. Click Format, and then enter the formatting you want. Click OK. After clicking OK in the Conditional Formatting dialog box, your formula and formatting are copied to the whole cell range. The format will be applied to any cell in the selected range that satisfies the condition defined in the formula.

Returning to the file ToysrusFormat.xls, let's focus on highlighting in purple the quarters in which revenues increase. Basically, what we want to do is select the range E5:E25 (there is no prior quarter to which we can compare the revenue figure in cell E4) and then instruct Excel that if a cell's value is larger than the cell above it, highlight the cell in purple. Figure 23-8 shows how to fill in the Conditional Formatting dialog box.

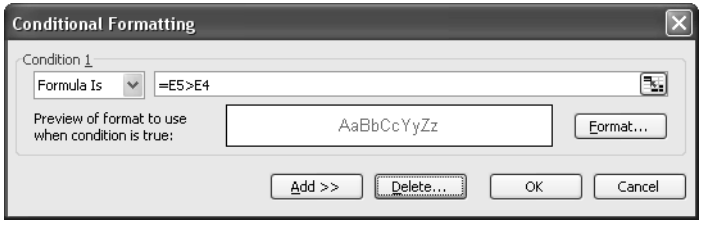



**Figure 23-8** Conditional formatting settings that will display in purple the quarters in which revenue increased.

If you enter *=E5>E4* by pointing to the appropriate cells, be sure you remove the $ signs from the formula in the Conditional Formatting dialog box or the formula won't be copied. The formula in this example ensures that cell E5 is colored purple if and only if sales in that quarter exceed the previous quarter. After clicking OK, you'll find that all quarters in which revenue increased are colored purple. Notice that in cell E6, for example, the formula was copied in the usual way, as *=E6>E5*.

To add the condition for formatting cells in which revenue decreased, select the range E5:E25 again, open the Conditional Formatting dialog box, and then

select Formula Is. Enter the formula *=E5<E4* and change the color to red. The Conditional Formatting dialog box will now appear as it's shown in Figure 23-9.

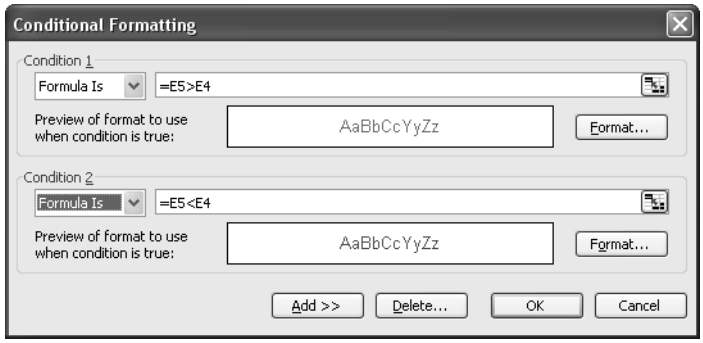



**Figure 23-9** These conditions will display quarters in which revenue increased in purple and quarters in which revenue decreased in red.

## Given a list of dates, how can I highlight the weekend dates in a specific color?

The file WeekendFormatting.xls (see Figure 23-10) contains several dates. We want to highlight all Saturdays and Sundays in red. To do this, I first copied the formula *WEEKDAY(C6,2)* from cell D6 to D7:D69. Choosing *Type = 2* for the WEEKDAY function returns a 1 for each Monday, 2 for each Tuesday, and so on, so that the function returns 6 for each Saturday and 7 for each Sunday.

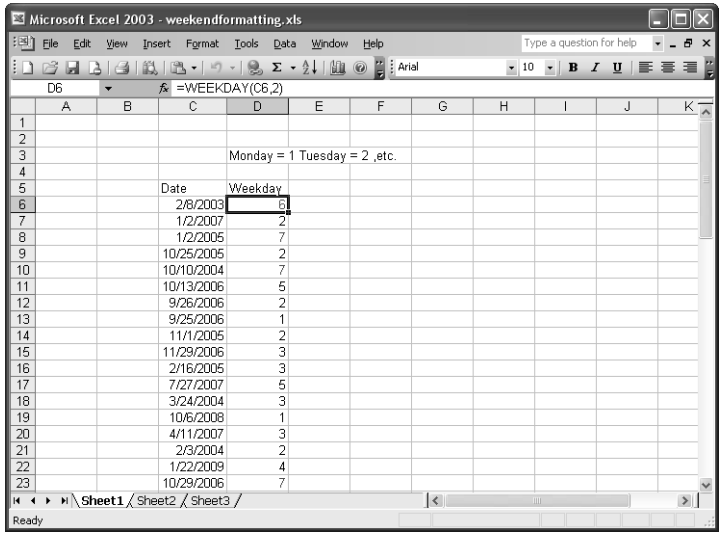



**Figure 23-10** In this spreadsheet, we've used the WEEKDAY function to help highlight weekend days in red.

I now select the range D6:D69 and then choose Format, Conditional Formatting. After selecting Formula Is, I fill in the dialog box as shown in Figure 23-11.

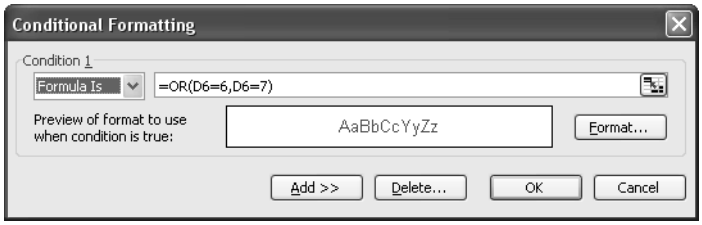
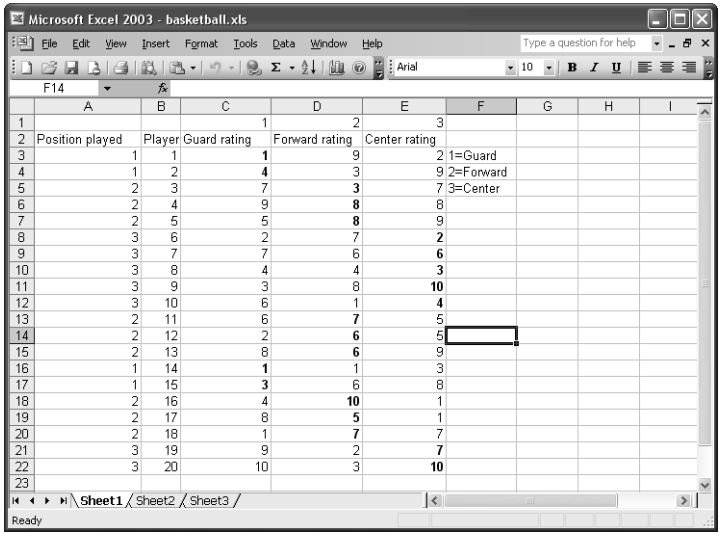



**Figure 23-11** The Conditional Formatting dialog box set up to display weekend days in red.

After clicking OK, each date having its weekday equal to 6 (for Saturday) or 7 (for Sunday) is colored red.

## Our basketball coach has given each player a rating between 1 and 10 for her ability to play guard, forward, or center. Can I set up a spreadsheet that highlights the ability of each player to play the position to which she's assigned?

The file Basketball.xls, shown in Figure 23-12, contains ratings given to 20 players for each position and the position (1 = guard, 2 = forward, 3 = center) played by each player. We would like to display in bold type the rating for each player for the position to which she's assigned.




| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | | | 1 | 2 | 3 | |
| 2 | Position played | Player | Guard rating | Forward rating | Center rating | |
| 3 | 1 | 1 | **1** | 9 | 2 | 1=Guard |
| 4 | 1 | 2 | **4** | 3 | 9 | 2=Forward |
| 5 | 2 | 3 | 7 | **3** | 7 | 3=Center |
| 6 | 2 | 4 | 9 | **8** | 8 | |
| 7 | 2 | 5 | 5 | **8** | 9 | |
| 8 | 3 | 6 | 2 | 7 | **2** | |
| 9 | 3 | 7 | 7 | 6 | **6** | |
| 10 | 3 | 8 | 4 | 4 | **3** | |
| 11 | 3 | 9 | 3 | 8 | **10** | |
| 12 | 3 | 10 | 6 | 1 | **4** | |
| 13 | 2 | 11 | 6 | **7** | 5 | |
| 14 | 2 | 12 | 2 | **6** | 5 | |
| 15 | 2 | 13 | 8 | **6** | 9 | |
| 16 | 1 | 14 | **1** | 1 | 3 | |
| 17 | 1 | 15 | **3** | 6 | 8 | |
| 18 | 2 | 16 | 4 | **10** | 1 | |
| 19 | 2 | 17 | 8 | **5** | 1 | |
| 20 | 2 | 18 | 1 | **7** | 7 | |
| 21 | 3 | 19 | 9 | 2 | **7** | |
| 22 | 3 | 20 | 10 | 3 | **10** | |
| 23 | | | | | | |

**Figure 23-12** This spreadsheet rates each player's ability to play a position.

Begin by selecting the range C3:E22, which contains the players' ratings. Choose Format, Conditional Formatting and then fill in the dialog box as shown Figure 23-13.

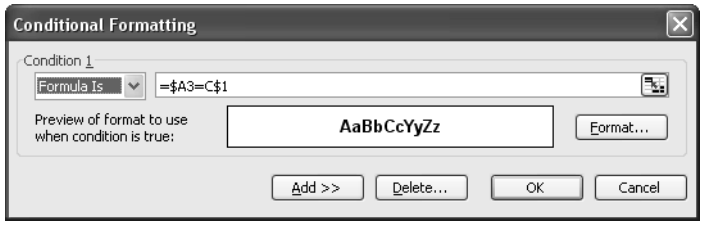



**Figure 23-13** The Conditional Formatting dialog box set up to show player ratings in bold type.

The formula *=$A3=C$1* compares the player's assigned position to the column heading (1, 2, or 3) in row 1. If the player's assigned position equals 1 (guard), her rating in column C, which is her guard rating, appears in bold. Similarly, if the player's assigned position equals 2, the rating in column D, her forward rating, appears in bold. Finally, if the assigned position equals 3, the rating in column E appears in bold.

# Problems

**1.** Using the data in the file SandP.xls, use conditional formatting in the following situations:

- ❑ Format in bold each month in which the value of the S&P increased and underline each month in which the value of the S&P decreased.
- ❑ Highlight in green each month in which the S&P changed by at most 2 percent.
- ❑ Highlight the largest S&P value in red and the smallest value in purple.

**2.** Using the data in the file Toysrus.xls, highlight in red all quarters for which revenue has increased over at least the last two quarters. Highlight all fourth-quarter revenues in blue and first-quarter revenues in red.

**3.** The file Test.xls contains exam scores for students. The top 10 students receive an A, the next 20 students receive a B, and all other students receive a C. Highlight the A grades in red, the B grades in

green, and C grades in blue. Hint: The function *LARGE(D4:D63,10)* gives you the tenth highest grade on the test.

4. In the file WeekendFormatting.xls, highlight all weekdays in red. Highlight all days that occur in the first 10 days of the month.

5. Suppose each worker in Microsoft's finance department has been assigned to one of four groups. The supervisor of each group has rated each worker on a 0–10 scale, and each worker has rated his satisfaction with each of the four groups. (See the file SatisSuper.xls.) Based on the group to which each worker is assigned, highlight the supervisor rating and worker satisfaction rating for each worker.

6. The file VarianceAnalysis.xls contains monthly profit forecasts and monthly actual sales. The sales variance for a month equals

$$\frac{\textit{actual sales} - \textit{forecasted sales}}{\textit{forecasted sales}}$$

Highlight in red all months with a favorable variance of at least 20 percent and highlight in green all months with an unfavorable variance of more than 20 percent.

7. For our drug cost example from Chapter 21, format the worksheet so that all phase 1 costs are displayed in red, all phase 2 costs are displayed in green, and all phase 3 costs are displayed in purple.