

Excel's Lists, Names, and Filters

3

Creating Lists

In Excel, a *list* is a set of data, arranged in a certain way. Many of Excel's features—the Data Form, pivot tables, filters, and others—will not work properly if data is not arranged as a list. And while other features such as charts will work with data that's not in list format, they become more difficult to use.

Furthermore, it's typical for data that has been imported into Excel from an external data source to enter the workbook in the form of a list. So, it's useful to know what a list is and how to set one up.

A list is a rectangular range of cells on a worksheet. It has one or more adjacent columns and two or more rows. The list is usually separated from other data on the worksheet by blank rows and columns.

In versions of Excel prior to Excel 2003, a list is an informal structure. It's more a way of arranging data and following a few conventions than something readily identifiable, such as a print area or a pivot table. There's no command to select or button to click to create a list.

Still, as informal as they are, list structures make it much easier to manage your data. For just one example of several in this chapter, see the section titled "Using Data Forms."

In Excel 2003, list structures remain informal, but a set of commands has been added to Excel's worksheet menu structure. These commands make it easier to set up lists, edit their data, and extend their reach. This section describes lists in general, and Excel 2003's new commands in particular.

IN THIS CHAPTER

Creating Lists	55
Sorting Lists	62
Working with Names	64
Filtering Data with the AutoFilter	74
Using the Advanced Filter	78
Looking Ahead	82



Understanding List Structures

Lists have three fundamental characteristics, shown in Figure 3.1.

Figure 3.1

The data in cells A1:C21 make up a *list*.

	A	B	C	D	E	F
1	Party	Sex	Age			
2	Democrat	Female	40			
3	Democrat	Female	52			
4	Republican	Male	75			
5	Democrat	Male	47			
6	Republican	Male	56			
7	Republican	Female	27			
8	Republican	Male	46			
9	Republican	Male	66			
10	Democrat	Female	72			
11	Republican	Male	34			
12	Democrat	Male	54			
13	Republican	Female	56			
14	Democrat	Female	78			
15	Democrat	Female	42			
16	Democrat	Male	36			
17	Democrat	Male	41			
18	Republican	Female	29			
19	Democrat	Male	88			
20	Republican	Female	24			
21	Republican	Female	39			
22						
23						

NOTE

If you've worked with a database management system such as Access, SQL Server, or even dBASE, you probably recognize the layout shown in Figure 3.1. In datasheet view, most database management systems display data in the same fashion: records occupying separate rows and fields occupying separate columns.

Notice the following in Figure 3.1:

- The variable (also called a *field*) named Party is in column A, the variable Sex is in column B, and the variable Age is in column C. The variables' names aren't important, nor are the particular columns. It *is* important that each variable is in a different column, and that the columns are adjacent.
- Each record is in a different row. Just by looking at the data, you can infer that row 3 represents a 52-year-old female who is a Democrat, row 11 represents a 34-year-old Republican male, and so on. No matter whether the data describe people, products, or plant life, in a list, each person, product, or plant is in a different row.
- Variable names occupy the list's first row. In Figure 3.1, the variable names are Party, Sex, and Age, and they are in the first row of the list. They do not have to be in row 1 of the worksheet, but they must be in row 1 of the list.

NOTE

Excel's Help, and other Microsoft documentation, variously use the terms *column labels* and *header row* (among others) to represent the first row of a list. To avoid confusion with the letters at the top of each worksheet column (which Excel terms *column headings*), this book uses the term *variable names* to mean the values, normally text, in a list's first row.

Figure 3.2 shows two data ranges that are not lists. In the range A1:C22, the first row does not contain variable names in each column. In the range F1:I21, there is an empty column so that not all the columns are adjacent.

Figure 3.2
Neither A1:C22 nor
F1:I21 is a list.

1	Party	Sex				Party	Sex		Age		
2	Democrat	Female	Age			Democrat	Female			40	
3	Democrat	Female	40			Democrat	Female			52	
4	Republican	Male	52			Republican	Male			75	
5	Democrat	Male	75			Democrat	Male			47	
6	Republican	Male	47			Republican	Male			56	
7	Republican	Female	56			Republican	Female			27	
8	Republican	Male	27			Republican	Male			66	
9	Republican	Male	66			Republican	Male			46	
10	Democrat	Female	46			Democrat	Female			72	
11	Republican	Male	72			Republican	Male			34	
12	Democrat	Male	34			Democrat	Male			54	
13	Republican	Female	54			Republican	Female			56	
14	Democrat	Female	56			Democrat	Female			78	
15	Democrat	Female	78			Democrat	Female			42	
16	Democrat	Male	42			Democrat	Male			36	
17	Democrat	Male	36			Democrat	Male			41	
18	Republican	Female	41			Republican	Female			29	
19	Democrat	Male	29			Democrat	Male			88	
20	Republican	Female	88			Republican	Female			24	
21	Republican	Female	24			Republican	Female			39	
22			39								
23											
24											

However, just because these ranges violate a couple of rules for list making doesn't necessarily mean that you'll get an error message, or that Excel will quit unexpectedly. It just means that the tools you want to use with lists won't work as readily.

For example, suppose that you click in cell A2 of Figure 3.1; then you choose Filter from the Data menu and click AutoFilter. A dropdown will appear next to each variable name: Party, Sex, and Age.

But if you do the same with the data as shown in columns A:C of Figure 3.2, Excel ignores the first row and puts dropdowns in the second row, next to Democrat, Female, and Age. Excel puts the dropdowns in the first row in the range that has nonblank values, making the assumption that they are variable names. (You'll find much more information about Excel's data filters in this chapter's sections titled "Filtering Data with the AutoFilter" and "Using the Advanced Filter.")

But that behavior is not consistent. Suppose that you click in cell A2 as shown in Figure 3.2, and choose Sort from the Data menu. If you tell Excel that your sort range has a header row, it will sort the range A3:C22. If you specify *no* header row, it will sort starting one row higher: A2:C22. In neither case will it pick up the first row. This isn't the behavior you want. When you structure your worksheet, be sure to put the names of the variables in the same row.

You can force the sort to pick up the first row by selecting the entire range before choosing Sort from the Data menu, but you've still inconvenienced yourself. Excel will then let you sort on Party, Sex, and column C—and if you sort on column C, Excel sorts what it regards as the value “Age” to the bottom of the range.

TIP

Excel's ascending sort order puts numbers first, and then text values, and then logical values (TRUE comes before FALSE), and then error values such as #REF!, and finally blanks. Except blanks, the order is reversed for a descending sort. Regardless of the sort order, blanks always come last.

To further illustrate the point, in Figure 3.2, click in cell G5 and choose Data, Filter, AutoFilter. Excel puts the dropdowns in F1 and G1, but ignores I1. Column I is not regarded as part of the list because it's separated from the rest of the data by a blank column. You can select the entire range F1:I21 before you start the AutoFilter, and then you'll get dropdowns in cells F1:I1. But what's the point of doing that? Where possible, make the columns in your list adjacent.

Again, the poorly designed lists cause no error messages in these examples, but Excel does not behave as you'd want when it encounters list structures that it doesn't expect.

On the other hand, try selecting A1:C22 as shown in Figure 3.2, and choose PivotTable and PivotChart Report from the Data menu. In step 2 of the wizard, make sure that A1:C22 is in the Range box: Excel will try to avoid using a range that contains a blank variable name, and the way it resolves its difficulty depends on the version you're using.

At some point (again, the point at which this occurs depends on the version that you have installed), Excel complains that The PivotTable field name is not valid and you won't be able to complete the pivot table. All the columns in a list that you use for a pivot table have to have variable names. Another way to violate list structure appears in Figure 3.3.

Figure 3.3

This arrangement transposes the list from records in rows to records in columns.

	A	B	C	D	E	F	G	H	I
1	Party	Democrat	Democrat	Republican	Democrat	Republican	Republican	Republican	Republican
2	Sex	Female	Female	Male	Male	Male	Female	Male	Male
3	Age	40	52	75	47	56	27	66	46
4									
5									
6									

Suppose that you began by selecting the entire range seen in Figure 3.3. Now, if you try to use AutoFilter on the range as shown, Excel will put dropdowns in each cell in the first row. The assumption will be that you have a variable named Party, one named Democrat, another named Democrat, another named Republican, and so on.

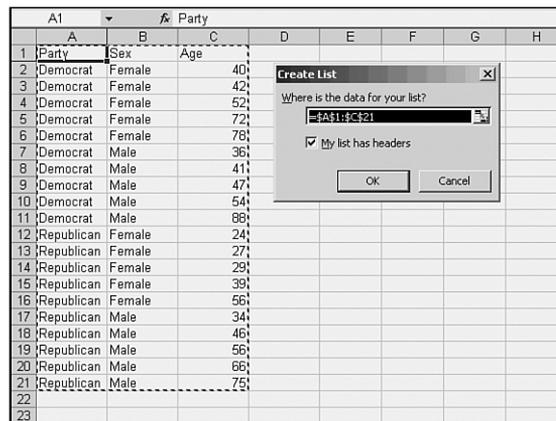
In other words, although you won't cause an error message using AutoFilter with this layout, you won't get what you're after, either.

Your life with lists will be much easier if you put variable names in the first row, different records in different rows, and different variables in different, adjacent columns.

Setting Up Lists in Excel 2003

The title of this section is a little misleading. You set up lists in Excel 2003 exactly as you do in earlier versions. The difference is that after you've arranged your list, you can click any cell in the list, and then choose List from the Data menu and Create List from the cascading menu. The window shown in Figure 3.4 appears.

Figure 3.4
Excel automatically proposes all adjacent, non-blank columns and rows for your list.



If Excel finds values in the first row that it can interpret as headers, it fills the My List Has Headers check box for you. Use the window to edit the list's range address if necessary, and use the check box to describe the list accurately. Then click OK. When you do so, several things happen:

- A border is drawn around the list, including its header row.
- The AutoFilter is turned on; you can tell this from the drop-down arrows in the header cells. (See this chapter's section titled "Filtering Data with the AutoFilter" for more information.)
- A row for entering additional records is established at the bottom of the list. Excel terms this the *insert row*. You can identify it from the asterisk in the list's first column. (If you're familiar with Microsoft Access, you'll recognize the asterisk as the indicator for adding records.)
- If you did not provide variable names, Excel supplies them for you, using the labels Column1, Column2, and so on.
- With a list active, right-click in it, choose List from the shortcut menu, and click Total Row in the cascading menu. Excel adds a row to the list that can show eight different types of total, including Sum, Average, and Count. Click a cell in the Total row to choose the total you want from a drop-down list.

See Figure 3.5 for an example of how a list appears after you've used the Create List command.

Figure 3.5

When you add a new value in any column in row 22, Excel expands the border and moves the insert row.

	A	B	C	D	E	F
1	Party	Sex	Age			
2	Democrat	Female	40			
3	Democrat	Female	42			
4	Democrat	Female	52			
5	Democrat	Female	72			
6	Democrat	Female	78			
7	Democrat	Male	36			
8	Democrat	Male	41			
9	Democrat	Male	47			
10	Democrat	Male	54			
11	Democrat	Male	88			
12	Republican	Female	24			
13	Republican	Female	27			
14	Republican	Female	29			
15	Republican	Female	39			
16	Republican	Female	56			
17	Republican	Male	34			
18	Republican	Male	46			
19	Republican	Male	56			
20	Republican	Male	66			
21	Republican	Male	75			
22	*					
23						
24						

You can expand the list directly by clicking and dragging the resize handle in the bottom-right corner of the list.

Excel 2003 also provides automatic subtotals for your list. To get them, select any cell in the list, choose Data, List, and then click Total Row in the cascading menu. It's a toggle, so to remove the total row, just click Total Row again. (You can also get to the List menu by right-clicking any cell in the list.)

Using Data Forms

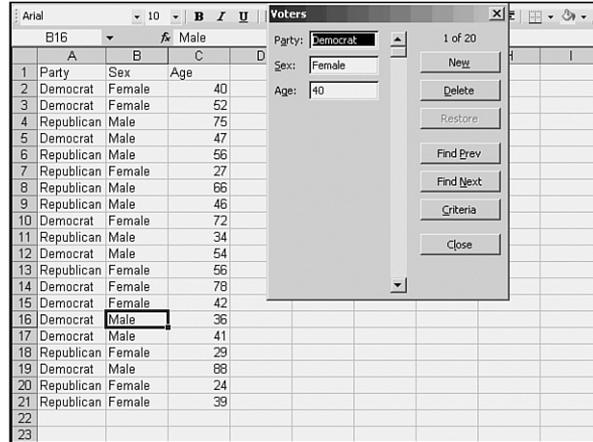
After you have a list set up, you can immediately start browsing through records, adding records, deleting records, and editing fields. You can use a form that Excel constructs for you automatically (see Figure 3.6).

All you need to get the Data Form to appear is to have a list, select any cell in the list, and choose Data, Form. A form that looks like the one shown in Figure 3.6 appears, and using it you can take any of the following actions:

- Click New to establish a new record in the list.
- See which record you're currently viewing by glancing at the record counter just above the New button.

Figure 3.6

The Data Form is automatically tailored to your list's variable names and number of records.



- Click Delete to delete the selected record. You're prompted to confirm that you want to delete it, and you can cancel the deletion if you want.
- Change a value in one or more of the edit boxes. After you've done so, click Restore to return all variables in the current record to their prior values. After you move to another record, you can no longer use Restore on the edited record.
- Move from edit box to edit box by using hot keys. Notice in Figure 3.6 that the edit box labels have hot keys on the form, indicated by the underscores. To move from, say, Party to Age, hold down Alt and simultaneously press Age's hot key, g.
- Click Criteria to set a selection criterion on any of the variables in your list (see Figure 3.7).

Figure 3.7

Clicking the Clear button clears all the boxes.



- Enter a value in one or more boxes to establish selection criteria.
- With criteria established, Find Prev takes you to an earlier record that matches the criteria and Find Next takes you to a subsequent matching record. If no match is found, the currently selected record remains selected.

- Return to Form view by clicking Form.
- Scroll through records by using the scrollbar.
- Click Close to remove the Data Form.

The Data Form is a handy way to manage records and variables that are set out in list format. Using it requires only a list structure and knowing to choose Data, Form (and it's an easy way to impress someone who doesn't know it's there).

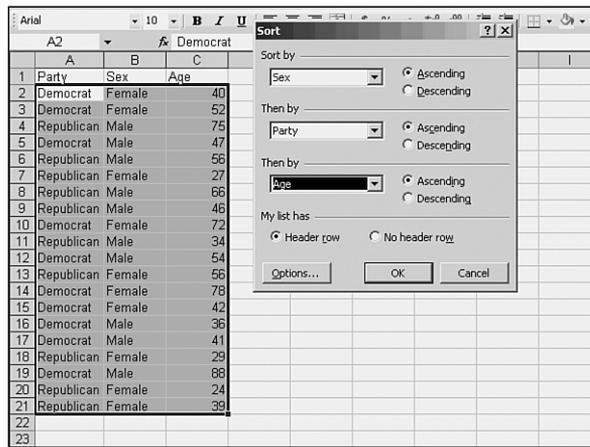
Sorting Lists

This chapter wouldn't have insisted that you be so fastidious about structuring lists if there weren't plenty of good reasons. Many of those reasons have to do with data management, but they don't stop there.

The list first shown in Figure 3.1 is repeated in Figure 3.8. Suppose that you wanted to sort the list, first by Sex, and then by Party within Sex, and finally by Age within Party within Sex. You would click any cell inside the list—as shown in Figure 3.8, that might be C8 or B18—and choose Sort from the Data menu.

Figure 3.8

You can choose No Header Row if you're not sorting a true list: one without variable names, perhaps.



In the first row of the range of contiguous cells, if Excel can find values that could be variable names, it treats those names as a header row. As Figure 3.8 shows, Excel excludes the variable names from the range of cells to be sorted. It lets you know by *not* highlighting the first row and by choosing (on your behalf) the Header row option button on the Sort dialog box.

This makes sorting very convenient. You use the dropdowns in the Sort dialog box to select the variables you want to use as the first, second, and third sort keys. If, as in Figure 3.8, you choose to sort by Age within Party within Sex, your choices appear in the dropdowns.

Suppose that you choose to sort in ascending order for all three variables: first by Sex, and then by Party, and then by Age. This pattern would sort all Females into one group of adjacent rows. Within the group of Females, it would sort the Female Democrats together and, in a different group of rows, the Female Republicans. Finally, within the Female Democrats, the pattern would sort the records in ascending order by Age.

When you have a range of cells that's so wide that its columns disappear off your screen, using a header row becomes especially convenient. Suppose that your range runs from column A to column J, and that you can see only columns A through E on your screen. You want to sort on, say, Product (column B), Sales Office (column G), and Revenue (column I). Without a header row that has variable names, you need to remember which columns are occupied by Sales Office and by Revenue—because you can't see them, you need to remember that your sort involves columns G and I. But if you're using a header row, you can see the variable names in the Sort dropdowns, and you don't need to remember where anything is.

Here's yet another reason to set your data up as a list. Assume that the data layout is as shown in Figure 3.9.

Figure 3.9
The layout is a simple transposition of the data in Figure 3.8.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Party	Democrat	Democrat	Republican	Democrat	Republican	Republican	Republican	Republican	Democrat	Republican	Democrat
2	Sex	Female	Female	Male	Male	Male	Female	Male	Male	Female	Male	Male
3	Age	40	52	75	47	56	27	66	46	72	34	51
4												
5												
6												

As you might know, you can choose to sort left-to-right as well as top-to-bottom. Use the following to do so:

1. Choose **Data, Sort**.
2. Click the **Options** button.
3. Fill the **Sort Left to Right** option button.
4. Click **OK** to return to the Sort menu and continue as before.

The problem is that left-to-right sorts don't support header rows. Notice in Figure 3.9 that the variable names are in the first column of each row instead of in the first row of each column. If you don't exclude that column from the range to be sorted, the headers will be sorted as though they were values.

NOTE

Notice that if you specify a left-to-right sort, Excel disables the Header Row and the No Header Row option buttons.

To preserve the headers in this layout, you would have to begin by selecting the entire range of values, excluding the headers in A1:A3. If you start by selecting a single cell only, Excel insists on treating the first column's variable names as values to be sorted.

Working with Names

Names are enormously useful in Excel. They make it easier to work with everything from worksheet ranges to arrays to constants to formulas and more.

There's no brief, crisp definition of the term *name* as it's used in Excel. It's best to look at examples of names to see how they're used and how they function. This book is concerned mainly with names as they apply to worksheet ranges, but keep in mind that names have a variety of uses.

Naming Formulas

Suppose that you frequently work with data that consists of a person's first name and last name: For example, you might have *George Washington* in cell A1 and *John Adams* in cell A2. You might want to strip off the person's first name, perhaps for use in a salutation. One way would be to use this combination of functions:

```
=LEFT(A1,FIND(" ",A1)-1)
```

If *George Washington* is in cell A1, this formula would return *George*. In words, it finds a blank space in the value in cell A1, and notes the position of that character in the string (here, that's 7). It subtracts 1 from that position, and returns that many characters. To simplify the formula:

```
=LEFT(A1,6)
```

or *George*.

That's useful, of course, but it's not very intuitive. Here's a way, involving a name, that's more cumbersome at first but lots easier in the long run:

1. Select cell **B1**.
2. Choose **Insert, Name, Define**. You'll see the window shown in Figure 3.10—this is a window you'll become very familiar with if you make effective use of names in Excel.
3. In the Names in Workbook box, type a handy mnemonic such as **FirstName**.
4. In the Refers To box, type

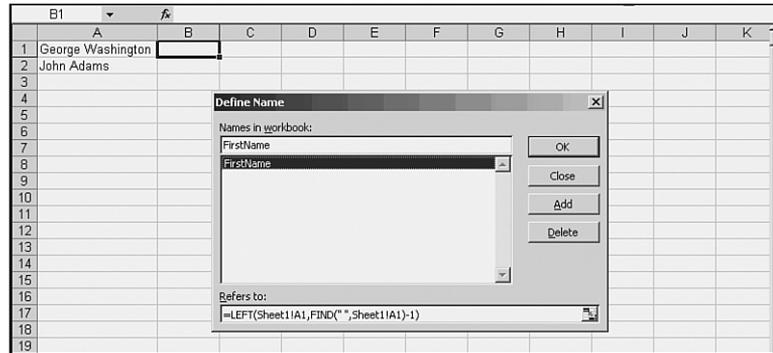
```
=LEFT(A1,FIND(" ",A1)-1)
```

If, instead of typing the cell address **A1**, you click in the cell, Excel will fill in the address for you. But Excel will add the dollar signs that make the reference absolute (**\$A\$1**). For the present purpose, you don't want that. Either type the address yourself or remove the dollar signs supplied by Excel.

5. Click **OK**.

Figure 3.10

You can put a combination of worksheet ranges, functions, and even other names in the Reference box.



Now, in cell B1, type `=FirstName`. If *George Washington* is in A1, you'll see *George* in B1. Select cell B2 and type `=FirstName`. If *John Adams* is in A2, you'll see *John* in B2.

The name *FirstName* is standing in for the formula that combines the LEFT and FIND functions. You have defined a name that refers to a formula.

Furthermore, it's a formula whose results depend on—are relative to—where you enter it. This is the reason that you took in step 1, and step 4 where you avoided using dollar signs. You selected cell B1, and the formula you entered refers to A1. By selecting a cell immediately to the right of the one you want the formula to refer to, you arrange for any instance of the formula to refer to the cell immediately to its left. If you enter `=FirstName` in cell AC27, it will refer to the value in cell AB27.

Especially if it's been some time since you used this worksheet, it's a lot easier to recognize, remember, and understand this

```
=FirstName
```

than this

```
=LEFT(A1,FIND(" ",A1) - 1)
```

and that's typical of names in Excel. That is, you can usually choose a name for something that's much easier to remember and use than is the thing itself.

Naming Constants

Another use for names is to refer to constants. You might establish the name `DollarsPerMile` by typing that name in the Names in Workbook box and `=.365` in the Refers To box. This would let you use the name `DollarsPerMile` in any calculation where you wanted to know how much to expense (in this example, 36.5 cents) per mile driven.

P 65For example

```
=100*DollarsPerMile
```

to return \$36.50. (If you're really nuts and have a scientific disposition, you might enter `Planck` in the Names in Workbook box, and `=6.62606891 * 10^(-34)` in the Refers To box.)

Naming Ranges

As useful as named formulas and named constants are, it's likely that the most frequent use of names in Excel is to refer to ranges of cells, including single cells. Figure 3.11 repeats the lookup situation originally shown in Figure 2.9.

Figure 3.11

A little care in naming ranges goes a long way toward clarifying what your formulas are intended to do.

C12		=INDEX(CommissionTable,MATCH(B12,QuantitySold,1),MATCH(A12,ProductLine,D))									
A	B	C	D	E	F	G	H	I	J	K	
1	Quantity	Product Line									
2	Sold	Laptops	Desktops	Scanners	Routers	Monitors	Printers				
3	1	2.5%	8.6%	3.8%	3.3%	7.2%	5.7%				
4	6	3.6%	10.8%	6.2%	5.4%	9.4%	8.3%				
5	11	4.6%	13.0%	8.6%	7.6%	11.6%	10.9%				
6	16	5.7%	15.2%	11.0%	9.7%	13.9%	13.5%				
7	21	6.8%	17.4%	13.4%	11.8%	16.1%	16.1%				
8											
9											
10											
11	Product	Quantity	Commission Percent								
12	Desktops	8	10.8%								
13	Routers	2	3.3%								
14	Printers	4	5.7%								
15	Monitors	12	11.6%								
16	Laptops	14	4.6%								
17											

In Figure 3.11, as in Figure 2.9, the value in cell C12 is 10.8%. In Figure 2.9, the formula in C12 is based on the INDEX function, and it uses columns and rows as its arguments:

```
=INDEX($C$3:$H$7,MATCH(B12,$B$3:$B$7,1), _
MATCH(A12,$C$2:$H$2,0))
```

That's not a formula that's rich in intuitive meaning. If you entered it on Monday and had another look at it on Friday, you'd spend a few seconds figuring out what it's intended to do.

Now suppose that you define some names, so that

- The name `CommissionTable` refers to `C3:H7`.
- The name `ProductLine` refers to `C2:H2`.
- The name `QuantitySold` refers to `B3:B7`.

Notice that the ranges that the names refer to are the ranges used in the INDEX formula, repeated from Figure 2.9. But now those ranges have names, and you can use this formula as shown in Figure 3.11:

```
=INDEX(CommissionTable,MATCH(B12,QuantitySold,1), _
MATCH(A12,ProductLine,0))
```

That's a lot easier to interpret. You can look at it and see almost at once that it returns from the commission table the value that is found at the intersection of a particular quantity and a particular product.

There are several good methods you can use to define these names. The method shown next gives you the most control. Using the layout shown in Figure 3.11, follow these steps:

1. Choose **Insert, Name, Define**.
2. In the Names in Workbook box, type `CommissionTable`.

3. Click in the Refers To box and then, using the mouse pointer, drag through the worksheet range **C3:H7**. (Notice that when you use the worksheet in this way to establish a reference, Excel makes the reference absolute.)
4. Click **OK**, or click **Add** if you're not through defining names.

Another convenient method to establish a named range involves the Name box. Begin by selecting the worksheet range C3:H7. Now click in the Name box—that's the box with the drop-down arrow, at the left edge of the Excel window and on the same row as the Formula Bar. Type the name `CommissionTable` and then press Enter.

Similarly, you could begin by selecting B3:B7, clicking in the Name box, typing `QuantitySold`, and pressing Enter.

TIP The Name box is a convenient way to tell whether the active range or cell has a name and, if so, what the name is. After naming the `QuantitySold` range, for example, the Name box shows that name if you select the range B3:B7. Turning it around, you can choose a name from the Name box's dropdown in order to select that range on the worksheet.

The Name box displays only names that refer to worksheet cells and ranges, and you can use it to define only range and cell names.

Using Implicit Intersections

Suppose that you define the name `Quantity` as referring to the range B12:B16 in Figure 3.11. Now, select a cell in some column other than B, and in a row anywhere from 12 through 16, you enter this formula:

```
=Quantity
```

That formula will return the corresponding value in the range named `Quantity`. For example, suppose that you entered that formula in cell F14. It would return the value 4: the value in the cell where the range named `Quantity` (that is, B12:B16) intersects the row where you enter the formula (here, row 14). That value is 4, so that's what the formula returns.

This is an example of an implicit intersection. It's implicit because the row is *implied* by the location of the formula. If you entered the same formula in row 16, it would return the value 14, where row 16 intersects the range named `Quantity`.

Suppose that the range that you name `Quantity` occupies several columns in one row, rather than several rows in one column. You enter the same `=Quantity` formula in one of its columns but in a row that's outside the named range. In that case, you would get the same effect: an implicit intersection, but with one of the range's columns instead of one of its rows.

The implicit intersection is useful in the current example on sales commissions. If you give the name `Quantity` to the range `B12:B16` in Figure 3.11, you can enter this formula in `C12`:
`=INDEX(CommissionTable,MATCH(Quantity,QuantitySold,1),MATCH(A12,ProductLine,0))`

Note the difference from the earlier example. In the first `MATCH`, the argument `B12` has been replaced with a reference to `Quantity`. Because you have entered it in cell `C12`, the implicit intersection picks up the value 8 from cell `B12` and returns the original result, 10.8%. When you copy and paste that formula into `C13:C16`, the implicit intersection again gets the necessary values from `B13:B16`, and again returns the proper results.

In the same way, you could give the name `Product` to the range in `A12:A16`. Then you could completely dispense with cell and range references:

```
=INDEX(CommissionTable,MATCH(Quantity,QuantitySold,1), _
MATCH(Product,ProductLine,0))
```

Now the formula has become self-documenting. You need not go back and forth between cell references in the formula and their contents in the worksheet to figure out what's going on.

NOTE

If a formula that relies on an implicit intersection is entered *outside* the rows or columns that the named range occupies, the formula returns the #VAL! error.

Defining Static Range Names

A name is *static* if it refers directly to a cell or range of cells. It's useful to distinguish a static name from a *dynamic* name, which refers to a range that can change size automatically as new data arrives (see the next section for more information).

You've already seen a couple of ways to define static range names: using the Name box, and using Insert, Name, Define. You can also use the Create item in the Name menu.

If you have an Excel list, you can easily create static range names based on the list's variable names. Select the entire list, choose Name from the Insert menu, and click Create. The window shown in Figure 3.12 appears.

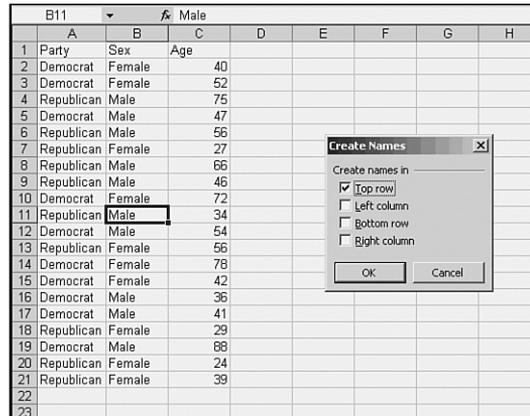
By filling the Top Row check box and clicking OK, you create three names: the name `Party` refers to `A2:A21`, `Sex` refers to `B2:B21`, and `Age` refers to `C2:C21`.

If your variable names are in the range's left column, instead of its top row, fill the Left Column check box. To create names that occupy rows as well as names that occupy columns, fill both the Left Column and the Top Row check boxes before clicking OK.

The Create Names window even allows for eccentrically placed variable names: If you've put them in the rightmost column or bottommost row, just fill the appropriate check boxes.

Figure 3.12

Filling the Left Column check box misleads Excel into treating the values *Democrat* and *Republican* as Names.

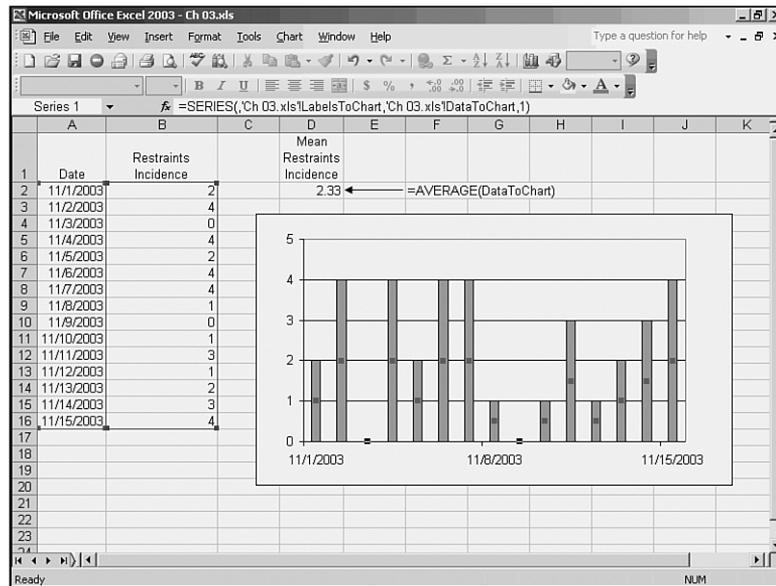


Defining Dynamic Range Names

Dynamic names are those that change the dimensions of the ranges they refer to, depending on how much data the ranges contain. Figure 3.13 gives an example.

Figure 3.13

Dynamic range names are effective in formulas and charts based on data that you update frequently.



There are two named ranges in Figure 3.13: one named `LabelsToChart` and one named `DataToChart`. The name `LabelsToChart` refers to the date values in column A. These are dates on which observations were made. The name `DataToChart` refers to the counts in column B. These count the incidents of the use of restraints in a hospital on a given date. The user wants to know the average daily incidence of the use of restraints, and to chart the actual daily incidence over time.

Both range names were defined by choosing Insert, Name, Define. This is the only way to define a dynamic range name; the Name box won't help you here.

Here's the definition of the range LabelsToChart, as found in the Refers To box of the Define Names window:

```
=OFFSET(Restrains!$A$1,1,0,COUNT(Restrains!$A:$A),1)
```

TIP

The easiest way to enter this reference is to click in the Refers To box. Type =**OFFSET**(and then click in cell A1. Excel will put `Restrains!A1` in the formula for you. When you get to the argument to the **COUNT** function, click the column label A. Excel automatically puts the reference `Restrains!$A:$A` into the formula.

This is another useful instance of the **OFFSET** function, already discussed in Chapter 2, "Excel's Data Management Features." In words, here's what it does:

- The **COUNT** function, as used in the name definition, returns the number of numeric values found in column A of the worksheet named Restrains. In the case shown in Figure 3.11, that result is 15. The 15 dates found in A2:A16 are all numbers, and the one label in cell A1 is a text value.
- The definition can now be simplified to
`=OFFSET(Restrains!A1,1,0,15,1)`
- Using the syntax of the **OFFSET** function, the definition refers to the range that's offset from `A1` by one row and zero columns, that's 15 rows high and one column wide—in other words, A2:A16.

Suppose now that you have reached November 16 on the calendar and it's time to enter another day's worth of data. You type `11/16/2003` in cell A17. Notice what happens to the definition of the dynamic range name LabelsToChart: The **COUNT** function in the definition now finds 16, not 15, numeric values in column A. So the definition now refers to the range that's offset from `A1` by one row and zero columns, *that's 16 rows high* and one column wide—that is, A2:A17.

This is why the name is termed a *dynamic* range name. The **COUNT** function makes it sensitive to the number of numeric values in column A. The more values in that column, the more rows in the named range.

The other named range in Figure 3.13, DataToChart, is defined as

```
=OFFSET(LabelsToChart,0,1)
```

in the Define Names window's Refers To box. This makes the name dependent on the name LabelsToChart: It is offset from that range by zero rows and one column. Because the (optional) height and width arguments are not provided, the DataToChart range automatically assumes the same number of rows and columns as LabelsToChart. So, as the number

of rows in `LabelsToChart` increases (or decreases), so does the number of rows in `DataToChart`.

After all this hand-waving, you're in a position to take advantage of the dynamic range names. In cell D2 of Figure 3.13, you find the formula `=AVERAGE(DataToChart)`, which—with the data as shown in column B—returns the value 2.33.

Suppose that you now enter 11/16/2003 in cell A17 and 4 in B17. This increases the average restraints incidence from 2.33 to 2.44. It also causes both ranges to increase by one row, and the formula `=AVERAGE(DataToChart)` recalculates accordingly.

Another effect appears in the chart. An additional column appears on the chart to reflect the new values entered in A17:B17. This is because the charted series is defined in the chart as `=SERIES('Ch 03.xls'!LabelsToChart,'Ch 03.xls'!DataToChart,1)`

TIP To view or edit what a chart's data series refers to, click on the series to select it. You can then see what it refers to, and edit that information, in the Formula Bar.

So, as each range gets more data, the names are dynamically redefined to capture the new information, and the chart updates to show more labels on its x-axis and more values in its columns.

There are a couple of aspects to dynamic range names that it pays to keep in mind, and we'll discuss them in the following sections.

Looking Out for Extraneous Values

Notice in Figure 3.13 that the formula `=AVERAGE(DataToChart)` is outside column A. If it were in column A, it would count as a numeric value, and would contribute to the number of numeric values returned by the `COUNT` function in the definition of `LabelsToChart`.

Suppose, for example, that `=AVERAGE(DataToChart)` were in column A. Then `=AVERAGE(DataToChart)` would involve a circular reference: The formula would contribute to the definition of the range it refers to. (There are situations in which this can be a good thing, but this isn't one of them.)

Or suppose that you somehow let an extraneous numeric value get into column A—as far away, perhaps, as cell A60000. Then column A would have 15 dates and one extra, unwanted number, each counting as a numeric value. The `COUNT` function would return 16, not 15, and `LabelsToChart` would extend from A2:A17.

To use dynamic range names effectively, you need to make sure to keep extraneous values out of the range that `COUNT` looks at.

Selecting Dynamically Defined Ranges

Static range names are available in the Name box: You can click the Name box's dropdown and choose a range name to select that range. They're also available via the Go To item in the Edit menu. The list box shows all accessible range names; just select one of them and then click OK to select its range.

Dynamic range names don't behave that way. You'll never see one in the Name box—not, at least, through the 2003 version of Excel. And if you choose Go To from the Edit menu, you won't see dynamic range names in the list box. You can, however, choose Go To from the Edit menu, and type an already existing dynamic range name in the Reference box. When you click OK, Excel selects the range that's currently defined by that dynamic name.

Understanding the Scope of Names

Names can be either *workbook-level* or *worksheet-level* names. Workbook-level names are the default, and are the type that this chapter has discussed so far.

Workbook-level names (often referred to more briefly as *book-level names*) are accessible from any worksheet or chart sheet in a workbook. So, if the book-level name DataToChart refers to a range on Sheet1, you can use that name in any sheet in the workbook. For example, the formula =AVERAGE(DataToChart) could be used on Sheet2 or Sheet3, and would return the same result each time.

You define a book-level name using any of the methods discussed so far in this chapter: by means of the Name box, using the Define Names dialog box, or with the Create Names dialog box.

One possible drawback to a book-level name is that only one instance of that name can exist in a workbook. For example, you can't use the book-level name DataToChart to refer to A1:A20 on Sheet1 and also to some other range, such as C1:C20 on Sheet1, or A1:A20 on Sheet2, or to a constant or a formula. A book-level name can exist only once in a workbook and can have one reference only.

In contrast, a worksheet-level name (also termed a *sheet-level name*) can exist once in a workbook for each sheet in that workbook. One sheet-level name DataToChart can exist for Sheet1, and another sheet-level name DataToChart can exist for Sheet2, and so on.

Here's how to define the sheet-level name DataToChart for Sheet1 and Sheet2 (you can extend it to as many worksheets as you like):

1. Activate Sheet1.
2. Choose **Insert, Name, Define**.
3. In the Names in Workbook box, type Sheet1!DataToChart.

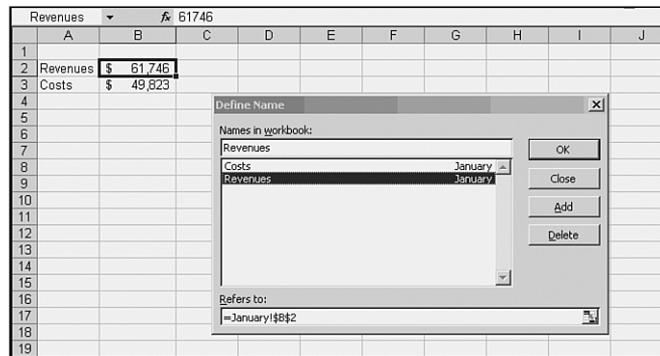
That is, qualify the range name by the name of the worksheet it is to belong to. Separate the name of the worksheet from the range name itself with an exclamation point.

- In the Refers To box, assign whatever reference you want: a constant, formula, or worksheet range. (If you cause the name to refer to a worksheet range, bear in mind that you can choose a range on any worksheet—not just Sheet1. That is, the name Sheet1!DataToChart can refer to B1:B10 on Sheet2.)
- Click **OK**, or click **Add** to continue defining names.

You could also use the Name box: Select the range you want to refer to, then type, for example, Sheet1!DataToChart in the Name box.

When you're through entering sheet-level names, it's a good idea to double-check them in the Define Names window (see Figure 3.14).

Figure 3.14
Which sheet-level names are visible depends on which sheet is active when you choose Insert, Name, Define.



Notice that the name of the sheet to which the name belongs (in Figure 3.14, that's January) appears to the right of the sheet-level name in the Names in Workbook list box.

Sheet-level names are very handy when you assign similar kinds of data to different sheets in a workbook. For example, you might place a different income statement for each month in a year on a different worksheet. Each worksheet might be named according to its month. Then, you could have January!Revenues, February!Revenues, March!Revenues, and so on.

Keep these points in mind as you work with sheet-level names:

- You don't need to qualify a sheet-level name when you use it on the sheet that it's defined for. That is, if the worksheet named January is active, the formula =SUM(Revenues) is equivalent to =SUM(January!Revenues).
- If you want to refer to a sheet-level name, and the sheet that it's defined for is not active, you must qualify the name. If the February worksheet is active, and you want it to show the sum of January's revenues, you would need to enter =SUM(January!Revenues). This is true even if there is no sheet-level name February!Revenues.

Filtering Data with the AutoFilter

When you're working with a large amount of data, you sometimes want to hide certain records so that you can focus on others. This is called *filtering*: You filter out the records that you want to ignore.

Excel's worksheets offer two approaches to filtering data: the AutoFilter and the Advanced Filter. (There's nothing really complicated about the Advanced Filter; it just takes an extra step to set up.) Both filters require that you arrange data in list form: different records in different rows, different variables in different columns, and variable names in the first row of each column.

NOTE

If you don't put variable names in the list's first row, Excel's filters ask you whether you want to treat the first row as names. Excel calls these *column names*. To avoid confusion with the A, B, C, D, . . . at the top of worksheet columns, we call them *variable names*.

Fast Filtering: Using the AutoFilter's Dropdowns

The idea behind AutoFilter is to make it easy to focus on a subset of records in your list. AutoFilter does this by temporarily hiding the records that don't belong to the subset you're interested in (see Figure 3.15).

Figure 3.15

Notice that the records in the unfiltered list are not sorted, nor need they be.

	A15	Branch			
	A	B	C	D	E
1	Branch	Product	Revenue		
2	Northwest	Laptop	\$33,208.98		
3	Southeast	Desktop	\$45,768.62		
4	Southeast	Desktop	\$24,280.12		
5	Northwest	Desktop	\$27,214.79		
6	Southeast	Laptop	\$38,538.92		
7	Southeast	Laptop	\$37,246.16		
8	Northwest	Laptop	\$31,567.80		
9	Northwest	Desktop	\$33,721.53		
10					
11					
12					
13					
14					
15	Branch	Product	Revenue		
16	Northwest	Laptop	\$33,208.98		
19	Northwest	Desktop	\$27,214.79		
22	Northwest	Laptop	\$31,567.80		
23	Northwest	Desktop	\$33,721.53		
24					
25					
26					

In Figure 3.15, the list in the range A1:C9 is unfiltered. That data is repeated in A15:C23, where the user has applied AutoFilter on the Branch variable so as to display only the Northwest branch's records. Notice that rows 17, 18, 20, and 21 are hidden: AutoFilter does this by setting their height to zero. And, although you can't tell in a black-and-white figure, the drop-down triangle in a field that has been used as a filter turns from black to blue.

TIP You can easily activate a cell even if it's in a hidden row. Click in the Name box, type the cell's address, and press Enter. You'll see the cell's contents, if any, in the Formula Bar.

Suppose that you have a worksheet with the data shown in the range A1:C9 of Figure 3.15. To prepare to use AutoFilter, just do this:

1. Select any cell in the list.
2. Choose **Data, Filter, AutoFilter**.

This puts a series of dropdowns into the first row of your list. Each dropdown, when clicked, displays the *unique* values found in its column (see Figure 3.16).

Figure 3.16
Each unique value in a dropdown appears once only.

	A15	Branch			
	A	B	C	D	E
1	Branch	Product	Revenue		
2	Northwest	Laptop	\$33,208.98		
3	Southeast	Desktop	\$45,768.62		
4	Southeast	Desktop	\$24,280.12		
5	Northwest	Desktop	\$27,214.79		
6	Southeast	Laptop	\$38,538.92		
7	Southeast	Laptop	\$37,246.16		
8	Northwest	Laptop	\$31,567.80		
9	Northwest	Desktop	\$33,721.53		
10					
11					
12					
13					
14					
15	Branch	Product	Revenue		
16	Sort Ascending	Laptop	\$33,208.98		
17	Sort Descending	Desktop	\$27,214.79		
18					
19					
20	(All)	Laptop	\$31,567.80		
21	(Top 10...)	Desktop	\$33,721.53		
22	(Custom...)				
23	Northwest				
24	Southeast				
25					
26					

By clicking the Northwest value in column A's dropdown, you can filter the records in the entire list so that only the Northwest branch's records appear.

If your list has more than one column, you can use AutoFilter to select two or more values simultaneously, one from each column's dropdown (see Figure 3.17).

In Figure 3.17, the user has focused on both the Northwest branch and the Desktop product line. The two dropdowns act as though they were connected by an *and*: "If Branch is Northwest *and* Product is Desktop ...".

Bear in mind that the effect of AutoFilter is to hide the rows that don't meet the filtering criteria that you set by means of the dropdowns. In Figure 3.17, for example, you can no longer see rows 16 through 18 and 20 through 22. But the values in the hidden records are still there. Formulas that depend on values in the hidden records are unchanged by AutoFilter.

Figure 3.17

Only the sales of Desktops at the Northwest branch are shown.

	A	B	C	D	E
1	Branch	Product	Revenue		
2	Northwest	Laptop	\$33,208.98		
3	Southeast	Desktop	\$45,768.62		
4	Southeast	Desktop	\$24,280.12		
5	Northwest	Desktop	\$27,214.79		
6	Southeast	Laptop	\$38,538.92		
7	Southeast	Laptop	\$37,246.16		
8	Northwest	Laptop	\$31,567.80		
9	Northwest	Desktop	\$33,721.53		
10					
11					
12					
13					
14					
15	Branch	Product	Revenue		
19	Northwest	Desktop	\$27,214.79		
23	Northwest	Desktop	\$33,721.53		
24					
25					

You could get a similar effect by sorting the list on the Branch and Product columns. But you might find that approach to be less convenient if you then have to scroll down the worksheet to find the combination that you're interested in.

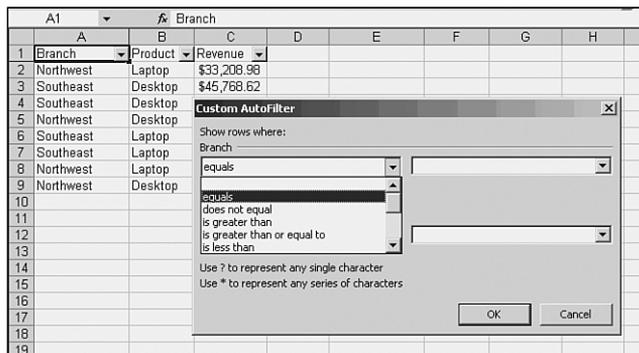
To reveal hidden records, but leave the dropdowns in place, choose **Data, Filter, Show All**. To remove the dropdowns, choose **Data, Filter**. You'll see a check mark by the AutoFilter menu item. Click AutoFilter again to remove the check mark from the menu and the dropdowns from the worksheet.

Using the AutoFilter with Other Criteria

Figure 3.16 shows that the AutoFilter has a Custom item in its dropdowns. Clicking it displays the window shown in Figure 3.18.

Figure 3.18

The AutoFilter enables you to specify two custom criteria for each column in your list.



The Custom AutoFilter window has two dropdowns with operators on the left and two value dropdowns on the right. You choose an operator and a value for it to operate on. For example, if you decided to view records from the Northwest branch only, you could select equals from the left dropdown, and Northwest from the right dropdown.

There are 12 operators available:

- Equals
- Does not equal
- Is greater than
- Is greater than or equal to
- Is less than
- Is less than or equal to
- Begins with
- Does not begin with
- Ends with
- Does not end with
- Contains
- Does not contain

NOTE

The Custom AutoFilter also supports wildcards. As usual, a question mark represents any single character and an asterisk represents any string of characters.

Using the Custom item makes it easy for you to arrange more complex analyses. Suppose that you wanted to filter records by a region that you just made up: North, for example, or East. In the Custom AutoFilter window, you could choose *equals* as the operator, and in the value dropdown, you could enter North*. Using the asterisk after North matches both Northwest and Northeast. You can see the result in Figure 3.19.

Keep in mind that the one or two custom criteria you set using the Custom AutoFilter apply to one column only. If you establish a custom criterion for Branch, and another custom criterion for Product, they act just as if you had selected simple criteria from the AutoFilter dropdowns; that is, they select records as if they were joined by an *and*.

Figure 3.19

You could create an *East* region instead, by entering **east* in the value dropdown.

	A19	fx Branch			
	A	B	C	D	E
1	Branch	Product	Revenue		
2	Northwest	Desktop	\$33,721.53		
3	Northwest	Desktop	\$27,214.79		
4	Northwest	Laptop	\$33,208.98		
5	Northwest	Laptop	\$31,567.80		
6	Southeast	Desktop	\$45,768.62		
7	Southeast	Desktop	\$24,280.12		
8	Southeast	Laptop	\$38,538.92		
9	Southeast	Laptop	\$37,246.16		
10	Northeast	Desktop	\$79,998.88		
11	Northeast	Desktop	\$97,444.60		
12	Northeast	Laptop	\$36,477.68		
13	Northeast	Laptop	\$66,185.68		
14	Southwest	Desktop	\$67,751.96		
15	Southwest	Desktop	\$77,383.85		
16	Southwest	Laptop	\$30,852.08		
17	Southwest	Laptop	\$24,926.69		
18					
19	Branch	Product	Revenue		
20	Northwest	Desktop	\$33,721.53		
21	Northwest	Desktop	\$27,214.79		
22	Northwest	Laptop	\$33,208.98		
23	Northwest	Laptop	\$31,567.80		
28	Northeast	Desktop	\$79,998.88		
29	Northeast	Desktop	\$97,444.60		
30	Northeast	Laptop	\$36,477.68		
31	Northeast	Laptop	\$66,185.68		

Using the Advanced Filter

The Advanced Filter used in Figure 3.20 provides three options that you can't get from AutoFilter:

- You can use it to create an entirely new, filtered list.
- You can obtain a list that contains unique records only—that is, only one instance of each possible combination of values.
- You have greater control over criteria. For example, you can establish more than two custom criteria that apply to a single column in the list.

Figure 3.20

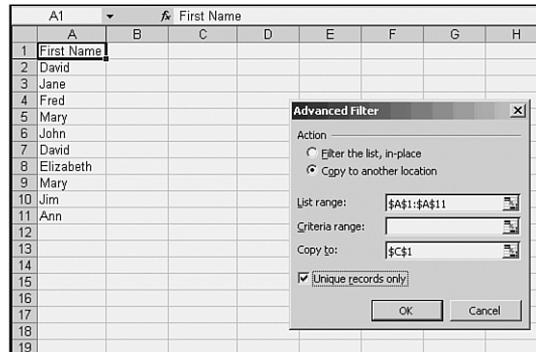
Filtering a list to a new location does not rely on hiding records.

	Extract	fx First Name			
	A	B	C	D	E
1	First Name		First Name		
2	David		David		
3	Jane		Jane		
4	Fred		Fred		
5	Mary		Mary		
6	John		John		
7	David		Elizabeth		
8	Elizabeth		Jim		
9	Mary		Ann		
10	Jim				
11	Ann				
12					

Column A of Figure 3.20 shows the original, unfiltered list. In column C, you see the new list that the user created with Advanced Filter: a list that consists only of unique names in column A. To get the result you see in column C, do this:

1. Click any cell in the original list.
2. Choose **Data, Filter, Advanced Filter**. The window shown in Figure 3.21 appears.

Figure 3.21
Filtering a list to a new location does not rely on hiding records.



3. Click the **Copy to Another Location** option button. This enables the Copy To box.
4. Click in the **Copy To** box, and then in some cell on the worksheet to establish the copy-to location.
5. Fill the **Unique Records Only** check box.
6. Click **OK** to create the filtered list.

CAUTION

Be sure that there's no important data in the columns that the filtered list will occupy (in Figure 3.20, that's columns E through G). The Advanced Filter overwrites existing data in the copy-to columns, either with filtered data or, farther down, with blank cells. **Excel does not provide a warning, and there is no Undo command available to take back the filtering action.**

TIP

If you click the button (termed a *collapse dialog* button) on the right edge of any of the three boxes, the dialog box collapses to give you more room on the worksheet. Click the same button in the collapsed dialog box to restore it to the original size.

NOTE

You cannot copy to a location on a different worksheet. If the list you want to filter is on, say, Sheet2, you cannot cause the Advanced Filter to copy filtered records to Sheet3.

Using Criteria with the Advanced Filter

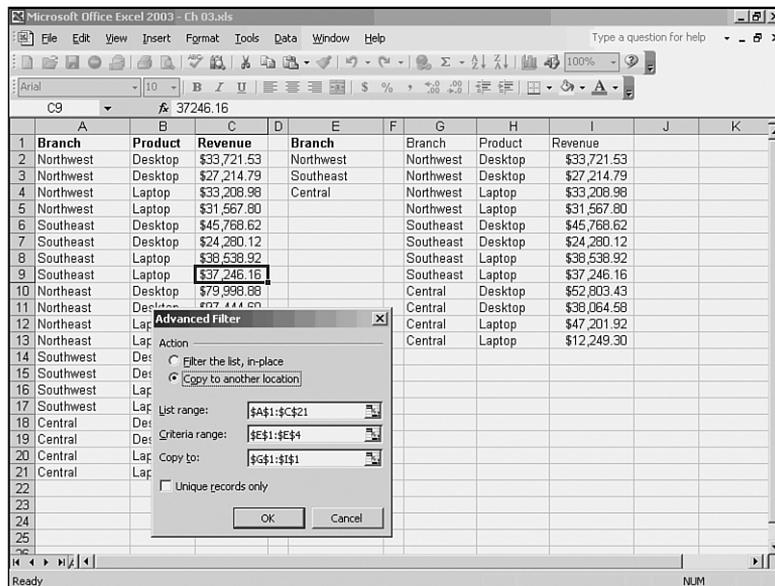
You can specify more criteria for the Advanced Filter than you can with the AutoFilter. Recall that for any column, you can choose one or two criteria using the Custom AutoFilter. For most purposes, this is plenty—especially when you consider that you can use wildcards in the criteria.

But suppose that you need more than two criteria per column, or that the filter you have in mind requires the Advanced Filter for some other reason. In that case, you'll need to specify in another worksheet location the criteria that the Advanced Filter will use. You might want to create a separate, filtered list that contains only records from the Northwest, Southeast, and Central branches. (Note that because this condition requires three criteria on the same field, AutoFilter won't do.) A separate list requires the Advanced Filter, so you would take these steps:

1. In some blank cell, enter Branch.
2. In the cell immediately below, enter Northwest. Below that enter Southeast, and below that enter Central.
3. Click in any cell in the existing list.
4. Choose **Data, Filter, Advanced Filter**. Click in the **Criteria Range** box, and then select the four cells you used in steps 1 and 2 (see Figure 3.22).

Figure 3.22

Any variable names in a criteria range must be identical to those in the list.



5. Continue as usual with the Advanced Filter, specifying a Copy To range, and Unique Values if you want.

The result appears in Figure 3.22, in cells G1:I13.

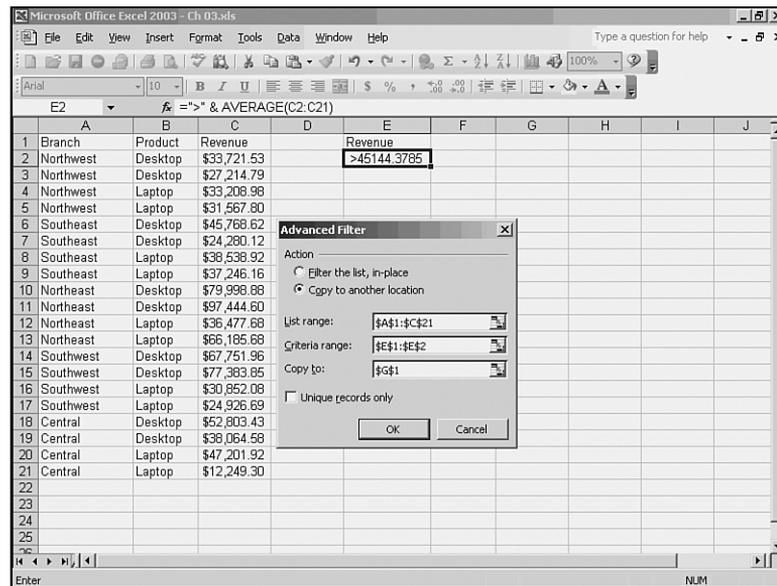
TIP

You can't specify a Copy To range on a different worksheet from the original list. But you can put the Criteria range on a different worksheet—even in a different workbook if you can think of a good reason to do so.

Using Formulas as Filter Criteria

With Advanced Filter, it's also possible to filter a list with the result of a formula as a criterion. Consider the list in A1:C22 of Figure 3.22. Suppose that you wanted to see only those records whose revenue value exceeded the average revenue of all the records. One way to arrange that is to use the Advanced Filter as before, but with a formula as the criterion. Your criteria range could occupy E1:E2, as shown in Figure 3.23.

Figure 3.23
The criterion in cell E2 actually evaluates to a string.

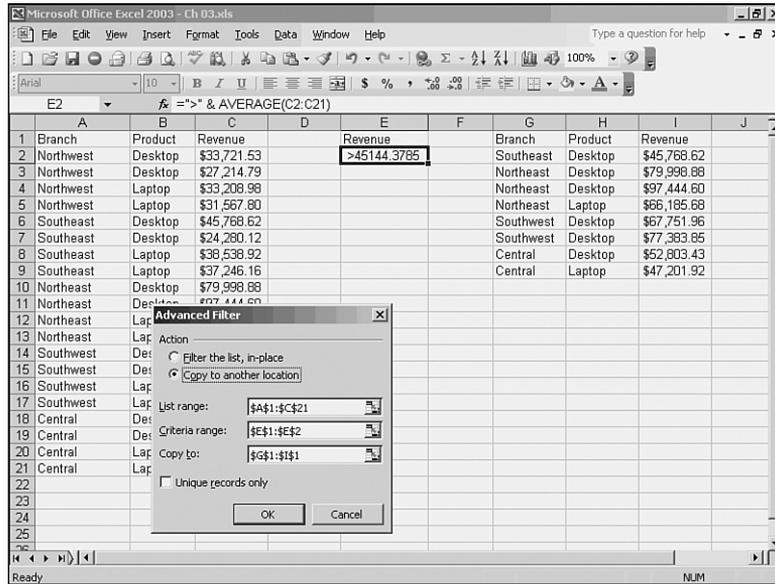


In E2, you enter = ">" & AVERAGE(C2:C21). This is the criterion that Advanced Filter will use: It is to return any value that is larger than the average of the Revenue values in C2:C21. Then take the usual steps:

1. Select any cell in the A1:C21 list.
2. Choose **Data, Filter, Advanced Filter**.
3. Select the **Copy to Another Location** option.
4. Click in the **Criteria Range** box, and drag through E1:E2.
5. Click in the **Copy To** box, and then click cell G1.
6. Click **OK**. The result appears in Figure 3.24.

Figure 3.24

Notice that the filtered list in G1:I9 contains only those records whose revenue exceeds \$45,144.38.



Looking Ahead

In Chapter 3, you've seen how to use three Excel techniques to manage data: lists, names, and filters. You've seen how to use them in conjunction with other Excel features such as the Data Form and the Sort command.

Because of the limitations of the printed page, the examples given in this chapter are necessarily brief and you've had to suspend your disbelief from time to time. (It would be unusual, for example, to apply a unique records filter to a list of 10 names.)

It's when you have hundreds and thousands of records in an Excel workbook that the techniques you learned in Chapter 3 start to become real timesavers. In Chapter 4, "Importing Data: An Overview," you'll start to see how you can quickly and automatically bring large amounts of data into an Excel worksheet from an external database.