

## IN THIS CHAPTER

- » Learning about data analysis
- » Analyzing data by applying conditional formatting
- » Adding subtotals to summarize data
- » Grouping related data
- » Combining data from multiple worksheets

# Chapter **1**

# Learning Basic Data-Analysis Techniques

**Y**ou are awash in data. Information multiplies around you so fast that you wonder how to make sense of it all. You think, “I know what to do. I’ll paste the data into Excel. That way, at least the data will be nicely arranged in the worksheet cells, and I can add a little formatting to make things somewhat palatable.” That’s a fine start, but you’re often called upon to do more with your data than make it merely presentable. Your boss, your customer, or perhaps just your curiosity requires you to divine some inner meaning from the jumble of numbers and text that litter your workbooks. In other words, you need to *analyze* your data to see what nuggets of understanding you can unearth.

This chapter gets you started down that data-analysis path by exploring a few straightforward but useful analytic techniques. After discovering what data analysis entails, you investigate a number of Excel data-analysis techniques, including conditional formatting, data bars, color scales, and icon sets. From there, you dive into some useful methods for summarizing your data, including subtotals, grouping, and consolidation. Before you know it, that untamed wilderness of a worksheet will be nicely groomed and landscaped.

# What Is Data Analysis, Anyway?

Are you wondering, “What is data analysis, anyway?” That’s an excellent question! Here’s an answer that I unpack for you as I go along: *Data analysis* is the application of tools and techniques to organize, study, reach conclusions, and sometimes make predictions about a specific collection of information.

For example, a sales manager might use data analysis to study the sales history of a product, determine the overall trend, and produce a forecast of future sales. A scientist might use data analysis to study experimental findings and determine the statistical significance of the results. A family might use data analysis to find the maximum mortgage it can afford or how much it must put aside each month to finance retirement or the kids’ education.

## Cooking raw data

The point of data analysis is to understand information on some deeper, more meaningful level. By definition, *raw data* is a mere collection of facts that by themselves tell you little or nothing of any importance. To gain some understanding of the data, you must manipulate the data in some meaningful way. The purpose of manipulating data can be something as simple as finding the sum or average of a column of numbers or as complex as employing a full-scale regression analysis to determine the underlying trend of a range of values. Both are examples of data analysis, and Excel offers a number of tools — from the straightforward to the sophisticated — to meet even the most demanding needs.

## Dealing with data

The *data* part of *data analysis* is a collection of numbers, dates, and text that represents the raw information you have to work with. In Excel, this data resides inside a worksheet, which makes the data available for you to apply Excel’s satisfyingly large array of data-analysis tools.

Most data-analysis projects involve large amounts of data, and the fastest and most accurate way to get that data onto a worksheet is to import it from a non-Excel data source. In the simplest scenario, you can copy the data from a text file, a Word table, or an Access datasheet and then paste it into a worksheet. However, most business and scientific data is stored in large databases, so Excel offers tools to import the data you need into your worksheet. I talk about all this in more detail later in the book.

After you have your data in the worksheet, you can use the data as is to apply many data-analysis techniques. However, if you convert the range into a *table*, Excel treats the data as a simple database and enables you to apply a number of database-specific analysis techniques to the table.

## Building data models

In many cases, you perform data analysis on worksheet values by organizing those values into a *data model*, a collection of cells designed as a worksheet version of some real-world concept or scenario. The model includes not only the raw data but also one or more cells that represent some analysis of the data. For example, a mortgage amortization model would have the mortgage data — interest rate, principal, and term — and cells that calculate the payment, principal, and interest over the term. For such calculations, you use formulas and Excel’s built-in worksheet functions.

## Performing what-if analysis

One of the most common data-analysis techniques is *what-if analysis*, for which you set up worksheet models to analyze hypothetical situations. The “what-if” part means that these situations usually come in the form of a question: “What happens to the monthly payment if the interest rate goes up by 2 percent?” “What will the sales be if you increase the advertising budget by 10 percent?” Excel offers four what-if analysis tools: data tables, Goal Seek, Solver, and scenarios, all of which I cover in this book.

# Analyzing Data with Conditional Formatting

Many Excel worksheets contain hundreds of data values. You could try to make sense of such largish sets of data by creating complex formulas and wielding Excel’s powerful data-analysis tools. However, just as you wouldn’t use a steam-roller to crush a tin can, sometimes these sophisticated techniques are too much tool for the job. For example, what if all you want are answers to simple questions such as the following:

- » Which cell values are less than 0?
- » What are the top 10 values?
- » Which cell values are above average, and which are below average?

These simple questions aren't easy to answer just by glancing at the worksheet, and the more numbers you're dealing with, the harder it gets. To help you eyeball your worksheets and answer these and similar questions, Excel lets you apply conditional formatting to the cells. Excel applies this special format only to cells that satisfy some condition, which Excel calls a *rule*. For example, you could apply formatting to display all negative values in a red font, or you could apply a filter to show only the top 10 values.

## Highlighting cells that meet some criteria

A *conditional format* is formatting that Excel applies only to cells that meet the criteria you specify. For example, you can tell Excel to apply the formatting only if a cell's value is greater or less than some specified amount, between two specified values, or equal to some value. You can also look for cells that contain specified text, dates that occur during a specified time frame, and more.

When you set up your conditional format, you can specify the font, border, and background pattern. This formatting helps to ensure that the cells that meet your criteria stand out from the other cells in the range. Here are the steps to follow:

### 1. Select the range you want to work with.

Select just the data values you want to format. Don't select any surrounding data.

### 2. Choose Home ⇨ Conditional Formatting.

### 3. Choose Highlight Cells Rules and then select the rule you want to use for the condition.

You have six rules to play around with:

- **Greater Than:** Applies the conditional format to cells that have a value larger than a value that you specify.
- **Less Than:** Applies the conditional format to cells that have a value smaller than a value that you specify.
- **Between:** Applies the conditional format to cells that have a value that is greater than or equal to a minimum value that you specify and less than or equal to a maximum value that you specify.
- **Equal To:** Applies the conditional format to cells that have a value that is the same as a value that you specify.
- **Text that Contains:** Applies the conditional format to cells that include the text that you specify.
- **A Date Occurring:** Applies the conditional format to cells that have a date value that meets the condition that you specify (such as Yesterday, Last Week, or Next Month).

(I cover a seventh rule — Duplicate Values — later in this chapter.) A dialog box appears, the name of which depends on the rule you click in Step 3. For example, Figure 1-1 shows the dialog box for the Greater Than rule.

	A	B	C	D	E	F	G	H	I	J	K
2	Country Name	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
3	Afghanistan	13.7	3.6	21.0	8.4	6.1	14.4	3.9	2.7	1.3	2.4
4	Albania	5.9	3.8	3.4	3.7	2.6	1.4	1.0	1.8	2.2	3.4
5	Algeria	3.4	2.4	1.6	3.6	2.9	3.4	2.8	3.8	3.8	3.3
6	American Samoa	2.0	-2.6	-4.2	0.4	0.3	-4.4	-2.8	0.9	1.2	-2.6
7	Andorra	0.0	-8.6	-3.7	-5.4	-4.6	-1.6	0.4	2.3	0.8	1.2
8	Angola	22.6	13.8	2.4	3.4	3.9	5.2	6.8	4.8	3.0	-0.7
9	Antigua and Barbuda	9.3	0.0							4.1	5.3
10	Arab World	4.6	5.8							3.4	3.2
11	Argentina	9.0	4.1							2.6	-2.2
12	Armenia	13.7	6.9							3.2	0.2
13	Australia	3.7	3.7							2.4	2.8
14	Austria	3.6	1.5	-5.0	1.9	2.9	0.7	0.0	0.0	1.1	1.5
15	Azerbaijan	25.0	10.8	9.4	4.9	0.1	2.2	5.8	2.0	1.1	-3.1
16	Bahamas, The	1.4	-2.3	-4.2	1.5	0.6	3.1	-0.6	-1.2	-3.1	0.2
17	Bangladesh	7.1	6.0	5.0	5.6	6.5	6.5	6.0	6.1	6.6	7.1
18	Barbados	1.8	0.4	-4.0	0.3	0.5	0.3	0.0	0.0	0.9	2.0
19	Belarus	8.6	10.2	0.2	7.8	5.5	1.7	1.0	1.7	-3.8	-2.6

**FIGURE 1-1:** The Greater Than dialog box and some highlighted values.

#### 4. Type the value to use for the condition.

You can also click the up arrow button that appears to the right of the text box and select a worksheet cell that contains the value. Also, depending on the operator, you might need to specify two values.

#### 5. Use the right drop-down list to select the formatting to apply to cells that match your condition.

If you're feeling creative, you can make up your own format by selecting the Custom Format command.

#### 6. Click OK.

Excel applies the formatting to cells that meet the condition you specified.



**TIP**

Excel enables you to specify multiple conditional formats for the same range. For example, you can set up one condition for cells that are greater than some value and a separate condition for cells that are less than some other value. You can apply unique formats to each condition. Keep the range selected and follow Steps 2 through 6 to configure the new condition.

## Showing pesky duplicate values

You use conditional formatting mostly to highlight numbers greater than or less than some value, or dates occurring within some range. However, you can use conditional formatting also to look for duplicate values in a range. Why would you

want to do that? The main reason is that many range or table columns require unique values. For example, a column of student IDs or part numbers shouldn't have duplicates.

Unfortunately, scanning such numbers and picking out the repeat values is hard. Not to worry! With conditional formatting, you can specify a font, border, and background pattern that ensures that any duplicate cells in a range or table stand out from the other cells. Here's what you do:

1. **Select the range that you want to check for duplicates.**
2. **Choose Home ⇨ Conditional Formatting.**
3. **Choose Highlight Cells Rules ⇨ Duplicate Values.**

The Duplicate Values dialog box appears. The left drop-down list has Duplicate selected by default, as shown in Figure 1-2. However, if you want to highlight all the unique values instead of the duplicates, select Unique from this list.

	A	B	C	D
1	Code	Product Name	Cost	Quantity
2	NWTB-1	Northwind Traders Chai	\$13.50	25
3	NWTCO-3	Northwind Traders Syrup	\$7.50	50
4	NWTCO-4	Northwind Traders Cajun Seasoning	\$16.50	0
5	NWTO-5	Northwind Traders Olive Oil	\$16.01	15
6	NWTJP-6	Northwind Traders Boysenberry Spread	\$18.75	0
7	NWTJP-6	Northwind Traders Marmalade	\$60.75	0
8	NWTDFN-7	Northwind Traders...	2.50	0
9	NWTS-8	Northwind Traders...	0.00	0
10	NWTFN-14	Northwind Traders...	7.44	40
11	NWTCFV-17	Northwind Traders...	9.25	0
12	NWTBGM-19	Northwind Traders...	5.90	0
13	NWTBGM-21	Northwind Traders Scones	\$7.50	0
14	NWTB-34	Northwind Traders Beer	\$10.50	23
15	NWTCM-40	Northwind Traders Crab Meat	\$13.80	0
16	NWTSO-41	Northwind Traders Clam Chowder	\$7.24	0
17	NWTB-34	Northwind Traders Coffee	\$34.50	325
18	NWTCA-48	Northwind Traders Chocolate	\$9.56	0
19	NWTFN-51	Northwind Traders Dried Apples	\$39.75	0

**FIGURE 1-2:** Use the Duplicate Values rule to highlight worksheet duplicates.

4. **In the right drop-down list, select the formatting to apply to the cells with duplicate values.**

You can create your own format by choosing the Custom Format command. In the Format Cells dialog box, use the Font, Border, and Fill tabs to specify the formatting you want to apply, and then click OK.

5. **Click OK.**

Excel applies the formatting to any cells that have duplicate values in the range.

# Highlighting the top or bottom values in a range

When analyzing worksheet data, looking for items that stand out from the norm is often useful. For example, you might want to know which sales reps sold the most last year, or which departments had the lowest gross margins. To quickly and easily view the extreme values in a range, you can apply a conditional format to the top or bottom values of that range.

You can apply such a format by setting up a *top/bottom rule*, in which Excel applies a conditional format to those items that are at the top or bottom of a range of values. For the top or bottom values, you can specify a number, such as the top 5 or 10, or a percentage, such as the bottom 20 percent. Here's how it works:

1. **Select the range you want to work with.**
2. **Choose Home → Conditional Formatting.**
3. **Choose Top/Bottom Rules and then select the type of rule you want to create.**

You have six rules to mess with:

- **Top 10 Items:** Applies the conditional format to cells that rank in the top  $X$ , where  $X$  is a number that you specify (the default is 10).
- **Top 10 %:** Applies the conditional format to cells that rank in the top  $X\%$ , where  $X$  is a number that you specify (the default is 10).
- **Bottom 10 Items:** Applies the conditional format to cells that rank in the bottom  $X$ , where  $X$  is a number that you specify (the default is 10).
- **Bottom 10 %:** Applies the conditional format to cells that rank in the bottom  $X\%$ , where  $X$  is a number that you specify (the default is 10).
- **Above Average:** Applies the conditional format to cells that rank above the average value of the range.
- **Below Average:** Applies the conditional format to cells that rank below the average value of the range.

A dialog box appears, the name of which depends on the rule you selected in Step 3. For example, Figure 1-3 shows the dialog box for the Top Ten Items rule.

4. **Type the value to use for the condition.**

You can also click the spin buttons that appear to the right of the text box. Note that you don't need to enter a value for the Above Average and Below Average rules.

	A	B	C	D
1		<b>Sales Rep</b>	<b>2017 Sales</b>	<b>2018 Sales</b>
2		Nancy Freehafer	\$996,336	\$960,492
3		Andrew Cencini	\$606,731	\$577,983
4		Jan Kotas	\$622,781	\$967,580
5		Mariya Sergienko	\$765,327	\$771,399
6		Steven Thorpe	\$863,589	\$827,213
7		Michael Neinner	\$795,518	\$669,394
8			740	\$626,945
9			059	\$574,472
10			380	\$827,932
11			623	\$569,609
12			777	\$558,601
13		Paul Durbin	\$685,091	\$692,182
14		Andrea Granek	\$540,484	\$693,762
15		Charles Aster	\$650,733	\$823,034
16		Karen Reilly	\$509,863	\$511,569
17		Karen Munson	\$503,699	\$975,455
18		Vince Voyatzis	\$630,263	\$599,514
19		Paul Sellars	\$779,722	\$596,353

**FIGURE 1-3:** The Top 10 Items dialog box with the top 5 values highlighted.

- In the right drop-down list, select the formatting to apply to cells that match your condition.**



REMEMBER

When you set up your top/bottom rule, select a format that ensures that the cells that meet your criteria will stand out from the other cells in the range. If none of the predefined formats suits your needs, you can always choose Custom Format and then use the Format Cells dialog box to create a suitable formatting combination. Use the Font, Border, and Fill tabs to specify the formatting you want to apply, and then click OK.

- Click OK.**

Excel applies the formatting to cells that meet the condition you specified.

## Analyzing cell values with data bars

In some data-analysis scenarios, you might be interested more in the relative values within a range than the absolute values. For example, if you have a table of products that includes a column showing unit sales, you might want to compare the relative sales of all products.

Comparing relative values is often easiest if you visualize the values, and one of the easiest ways to visualize data in Excel is to use *data bars*, a data visualization feature that applies colored horizontal bars to each cell in a range of values; these bars appear “behind” (that is, in the background of) the values in the range. The length of the data bar in each cell depends on the value in that cell: the larger the value, the longer the data bar.



Follow these steps to apply data bars to a range:

1. **Select the range you want to work with.**
2. **Choose Home → Conditional Formatting.**
3. **Choose Data Bars and then select the fill type of data bars you want to create.**

You can apply two type of data bars:

- **Gradient fill:** The data bars begin with a solid color and then gradually fade to a lighter color.
- **Solid fill:** The data bars are a solid color.

Excel applies the data bars to each cell in the range. Figure 1-4 shows an example in the Units column.

	A	B	C
1	Product Name	Units	\$ Total
2	Northwind Traders Almonds	20	\$ 200
3	Northwind Traders Beer	487	\$ 6,818
4	Northwind Traders Boysenberry Spread	100	\$ 2,500
5	Northwind Traders Cajun Seasoning	40	\$ 880
6	Northwind Traders Chai	40	\$ 720
7	Northwind Traders Chocolate	200	\$ 2,550
8	Northwind Traders Chocolate Biscuits Mix	85	\$ 782
9	Northwind Traders Clam Chowder	290	\$ 2,799
10	Northwind Traders Coffee	650	\$ 29,900
11	Northwind Traders Crab Meat	120	\$ 2,208
12	Northwind Traders Curry Sauce	65	\$ 2,600
13	Northwind Traders Dried Apples	40	\$ 2,120
14	Northwind Traders Dried Pears	40	\$ 1,200
15	Northwind Traders Dried Plums	75	\$ 263
16	Northwind Traders Fruit Cocktail	40	\$ 1,560
17	Northwind Traders Gnocchi	10	\$ 380
18	Northwind Traders Green Tea	275	\$ 822
19	Northwind Traders Long Grain Rice	40	\$ 280
20	Northwind Traders Marmalade	40	\$ 3,240
21	Northwind Traders Mozzarella	90	\$ 3,132
22	Northwind Traders Olive Oil	25	\$ 534

**FIGURE 1-4:**  
The higher the value, the longer the data bar.



TIP

If your range includes right-aligned values, gradient-fill data bars are a better choice than solid-fill data bars. Why? Because even the longest gradient-fill bars fade to white toward the right edge of the cell, so your range values will mostly appear on a white background, making them easier to read.

## Analyzing cell values with color scales

Getting some idea about the overall distribution of values in a range is often useful. For example, you might want to know whether a range has many low values

and just a few high values. Color scales can help you analyze your data in this way. A *color scale* compares the relative values in a range by applying shading to each cell, where the color reflects each cell's value.

Color scales can also tell you whether your data includes *outliers*: values that are much higher or lower than the others. Similarly, color scales can help you make value judgments about your data. For example, high sales and low numbers of product defects are good, whereas low margins and high employee turnover rates are bad.

To apply a color scale to a range of values, do the following:

1. **Select the range you want to format.**
2. **Choose Home ⇨ Conditional Formatting.**
3. **Choose Color Scales and then select the color scale that has the color scheme you want to apply.**

Color scales come in two varieties: three-color scales and two-color scales.



TIP

If your goal is to look for outliers, go with a three-color scale because it helps the outliers stand out more. A three-color scale is also useful if you want to make value judgments about your data, because you can assign your own values to the colors (such as positive, neutral, and negative). Use a two-color scale when you want to look for patterns in the data, because a two-color scale offers less contrast.

Excel applies the color scale to each cell in your selected range.

## Analyzing cell values with icon sets

Symbols that have common or well-known associations are often useful for analyzing large amounts of data. For example, a check mark usually means that something is good or finished or acceptable, whereas an X means that something is bad or unfinished or unacceptable. Similarly, a green circle is positive, whereas a red circle is negative (think traffic lights). Excel puts these and other symbolic associations to good use with the *icon sets* feature. You use icon sets to visualize the relative values of cells in a range.



REMEMBER

With icon sets, Excel adds a particular icon to each cell in the range, and that icon tells you something about the cell's value relative to the rest of the range. For example, the highest values might be assigned an upward-pointing arrow, the lowest values a downward-pointing arrow, and the values in between a horizontal arrow.

Here's how you apply an icon set to a range:

1. **Select the range you want to format with an icon set.**
2. **Choose Home → Conditional Formatting.**
3. **Choose Icon Sets and then select the type of icon set you want to apply.**

Icon sets come in four categories:

- **Directional:** Indicates trends and data movement
- **Shapes:** Points out the high (green) and low (red) values in the range
- **Indicators:** Adds value judgments
- **Ratings:** Shows where each cell resides in the overall range of data values

Excel applies the icons to each cell in the range, as shown in Figure 1-5.

	A	B
1	Student ID	Grade
2	64947	✓ 82
3	69630	⚠ 66
4	18324	✘ 52
5	89826	✓ 94
6	63600	✘ 40
7	25089	⚠ 62
8	89923	✓ 88
9	13000	⚠ 75
10	16895	⚠ 66
11	24918	⚠ 62
12	45107	⚠ 71
13	64090	✘ 53
14	94395	⚠ 74
15	58749	⚠ 65

**FIGURE 1-5:**  
Excel applies an icon based on each cell's value.

## Creating a custom conditional-formatting rule

The conditional-formatting rules in Excel — highlight cells rules, top/bottom rules, data bars, color scales, and icon sets — offer an easy way to analyze data through visualization. However, you can tailor your formatting-based data analysis also by creating a custom conditional-formatting rule that suits how you want to analyze and present the data.



REMEMBER

Custom conditional-formatting rules are ideal for situations in which normal value judgments — that is, that higher values are good and lower values are bad — don't apply. In a database of product defects, for example, lower values are better than higher ones. Similarly, data bars are based on the relative numeric values in a range, but you might prefer to base them on the relative percentages or on percentile rankings.

To get the type of data analysis you prefer, follow these steps to create a custom conditional-formatting rule and apply it to your range:

1. **Select the range you want to analyze with a custom conditional-formatting rule.**
  2. **Choose Home ⇨ Conditional Formatting ⇨ New Rule.**
- The New Formatting Rule dialog box appears.
3. **In the Select a Rule Type box, select the type of rule you want to create.**
  4. **Use the controls in the Edit the Rule Description box to edit the rule's style and formatting.**

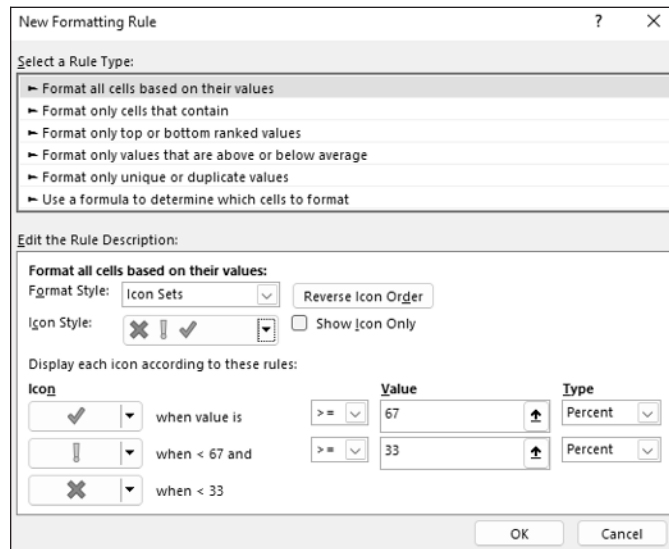
The controls you see depend on the rule type you selected in Step 3. For example, if you select Icon Sets, you see the controls shown in Figure 1-6.



TIP

5. **Click OK.**
- Excel applies the conditional formatting to each cell in the range.

**FIGURE 1-6:** Use the New Formatting Rule dialog box to create a custom rule.





TECHNICAL  
STUFF

## HIGHLIGHT CELLS BASED ON A FORMULA

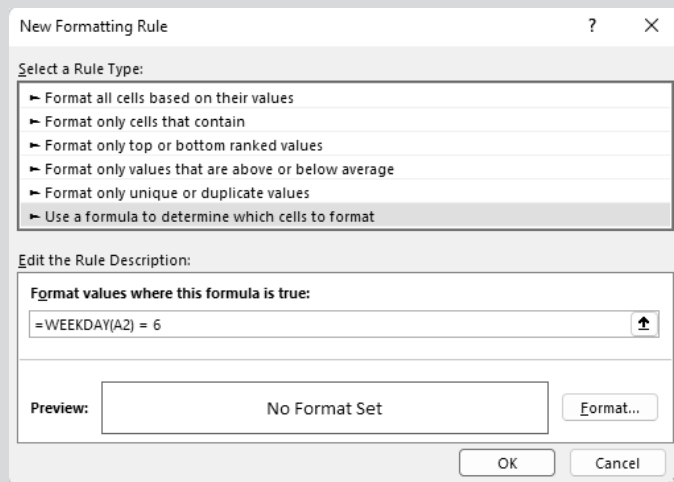
You can apply conditional formatting based on the results of a formula. That is, you set up a logical formula as the conditional-formatting criteria. For each cell in which that formula returns TRUE, Excel applies the formatting you specify; for all the other cells, Excel doesn't apply the formatting.

In most cases, you use a comparison formula, or you use an IF function, often combined with another logical function such as AND or OR. In each case, your formula's comparison value must reference only the first value in the range. For example, if the range you are working with is a set of dates in A2:A100, the comparison formula `=WEEKDAY(A2)=6` would apply conditional formatting to every cell in the range that occurs on a Friday.

The following steps show you how to apply conditional formatting based on the results of a formula:

- 1. Select the range you want to work with.**
- 2. Choose Home ⇨ Conditional Formatting ⇨ New Rule.**  
The New Formatting Rule dialog box appears.
- 3. Select Use a Formula to Determine Which Cells to Format.**
- 4. In the Format Values Where this Formula Is True text box, type the logical formula.**

The figure shows an example of using a formula to apply conditional formatting.



(continued)

(continued)

**5. Choose Format, use the Format Cells dialog box to define the rule's style and formatting, and then click OK.**

**6. Click OK.**

Excel applies the conditional formatting to each cell in the range in which the logical formula returns TRUE.

When you're messing around with formula-based rules, one useful technique is to apply a conditional format based on a formula that compares all the cells in a range to one value in that range. The simplest case is a formula that applies conditional formatting to those range cells that are equal to a cell value in the range. Here's the logical formula to use for such a comparison:

*=range=cell*

Here, *range* is an absolute reference to the range of cells you want to work with, and *cell* is a relative reference to the comparison cell. For example, to apply a conditional format to those cells in the range A1:A50 that are equal to the value in cell A1, you would use the following logical formula:

*=\$A\$1 : \$A\$50=A1*

## Editing a conditional-formatting rule

Conditional-formatting rules are excellent data-visualization tools that can make analyzing your data easier and faster. Whether you're highlighting cells based on criteria, showing cells in the top or bottom of a range, or using features such as data bars, color scales, and icon sets, conditional formatting enables you to interpret your data quickly.

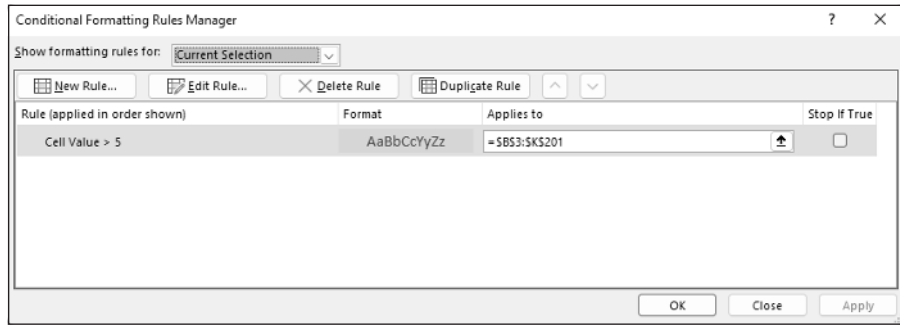
But it doesn't follow that all your conditional-formatting experiments will be successful ones. For example, you might find that the conditional formatting you used isn't working out because it doesn't let you visualize your data the way you'd hoped. Similarly, a change in data might require a change in the condition you used. Whatever the reason, you can edit your conditional-formatting rules to ensure that you get the best visualization for your data. Here's how:

**1. Select a cell in the range that includes the conditional-formatting rule you want to edit.**

You can select a single cell, multiple cells, or the entire range.

**2. Choose Home ⇨ Conditional Formatting ⇨ Manage Rules.**

The Conditional Formatting Rules Manager dialog box appears, as shown in Figure 1-7.



**FIGURE 1-7:**  
Use the  
Conditional  
Formatting Rules  
Manager to edit  
your rules.

### 3. Select the rule you want to modify.

If you don't see the rule, click the Show Formatting Rules For drop-down list and then select This Worksheet. The list that appears displays every conditional-formatting rule that you've applied in the current worksheet.

### 4. Choose Edit Rule.

The Edit Formatting Rule dialog box appears.

### 5. Make your changes to the rule.

### 6. Click OK.

Excel returns you to the Conditional Formatting Rules Manager dialog box.

### 7. Select OK.

Excel updates the conditional formatting.



**WARNING**

If you have multiple conditional-formatting rules applied to a range, the visualization is affected by the order in which Excel applies the rules. Specifically, if a cell already has a conditional format applied, Excel does not overwrite that format with a new one. For example, suppose that you have two conditional-formatting rules applied to a list of student grades: one for grades over 90 and one for grades over 80. If you apply the over-80 conditional format first, Excel will never apply the over-90 format because those values are already covered by the over-80 format. The solution is to change the order of the rule. In the Conditional Formatting Rules Manager dialog box, select the rule that you want to modify and then click the Move Up and Move Down button to set the order you want. If you want Excel to stop processing the rest of the rules after it has applied a particular rule, select that rule's Stop If True check box.

## Removing conditional-formatting rules

Conditional-formatting rules are useful critters, but they don't work in all scenarios. For example, if your data is essentially random, conditional-formatting rules won't magically produce patterns in that data. You might also find that

conditional formatting isn't helpful for certain collections of data or certain types of data. Or you might find conditional formatting useful for getting a handle on your data set but then prefer to remove the formatting.

Similarly, although the data-visualization aspect of conditional-formatting rules is part of the appeal of this Excel feature, as with all things visual, you can overdo it. That is, you might end up with a worksheet that has multiple conditional-formatting rules and therefore some unattractive and confusing combinations of highlighted cells, data bars, color scales, and icon sets.

If, for whatever reason, you find that a range's conditional formatting isn't helpful or is no longer required, you can remove the conditional formatting from that range by following these steps:

**1. Select a cell in the range that includes the conditional-formatting rule you want to trash.**

You can select a single cell, multiple cells, or the entire range.

**2. Choose Home ⇨ Conditional Formatting ⇨ Manage Rules.**

The Conditional Formatting Rules Manager dialog box appears.

**3. Select the rule you want to remove.**

If you don't see the rule, use the Show Formatting Rules For list to select This Worksheet, which tells Excel to display every conditional-formatting rule that you've applied in the current worksheet.

**4. Choose Delete Rule.**

Excel removes the rule from the range.

**5. Click OK.**



TIP

If you have multiple rules defined and want to remove them all, click the Home tab, choose Conditional Formatting, choose Clear Rules, and then select either Clear Rules from Selected Cells or Clear Rules from Entire Sheet.

## Summarizing Data with Subtotals

Although you can use formulas and worksheet functions to summarize your data in various ways — including sums, averages, counts, maximums, and minimums — if you're in a hurry, or if you just need a quick summary of your data, you can get Excel to do the work for you. The secret here is a feature called *automatic subtotals*, which are formulas that Excel adds to a worksheet automatically.





REMEMBER

Excel sets up automatic subtotals based on data groupings in a selected field. For example, if you ask for subtotals based on the Customer field, Excel runs down the Customer column and creates a new subtotal each time the name changes. To get useful summaries, you should sort the range on the field containing the data groupings you're interested in.

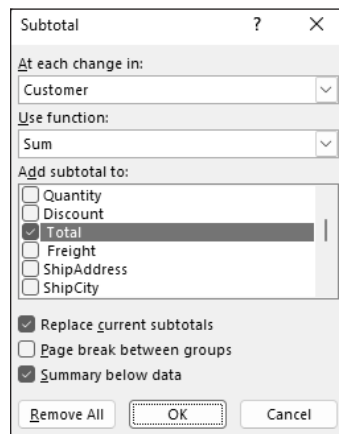
Follow these steps to summarize your data with subtotals:

1. **Select a cell within the range you want to subtotal.**
2. **Choose Data ⇨ Subtotal.**  
If you don't see the Subtotal command, choose Outline ⇨ Subtotal. The Subtotal dialog box appears.
3. **In the At Each Change In list, select the column you want to use to group the subtotals.**
4. **In the Use Function list, select Sum.**
5. **In the Add Subtotal To list, select the check box for the column you want to summarize.**

In Figure 1-8, for example, each change in the Customer field displays the sum of that customer's Total cells.

6. **Click OK.**

Excel calculates the subtotals and adds them into the range. Note, too, that Excel also adds outline symbols to the range. I talk about outlining in a bit more detail in the next section.



**FIGURE 1-8:** Use the Subtotal dialog box to apply subtotals to a range.

Figure 1-9 shows some subtotals applied to a range.

1	2	3	A	B	C	E	F	G	H	I
1	Customer	Country	Region	Unit Price	Quantity	Discount	Total	Freight		
2	Cactus Comidas para llevar	Argentina		\$ 46.00	7	0%	\$ 322.00	\$ 19.76		
3	Cactus Comidas para llevar	Argentina		\$ 7.75	20	0%	\$ 155.00	\$ 19.76		
4	Cactus Comidas para llevar	Argentina		\$ 15.00	10	0%	\$ 150.00	\$ 2.84		
5	Cactus Comidas para llevar	Argentina		\$ 45.60	8	0%	\$ 364.80	\$ 31.51		
6	Cactus Comidas para llevar	Argentina		\$ 14.00	20	0%	\$ 280.00	\$ 31.51		
7	<b>Cactus Comidas para llevar Total</b>							\$ 1,271.80		
8	Océano Atlántico Ltda.	Argentina		\$ 6.00	5	0%	\$ 30.00	\$ 1.27		
9	Océano Atlántico Ltda.	Argentina		\$ 21.35	20	0%	\$ 427.00	\$ 49.56		
10	Océano Atlántico Ltda.	Argentina		\$ 30.00	6	0%	\$ 180.00	\$ 49.56		
11	Océano Atlántico Ltda.	Argentina		\$ 34.80	5	0%	\$ 174.00	\$ 49.56		
12	Océano Atlántico Ltda.	Argentina		\$ 21.00	30	0%	\$ 630.00	\$ 217.86		
13	Océano Atlántico Ltda.	Argentina		\$ 81.00	15	0%	\$ 1,215.00	\$ 217.86		
14	Océano Atlántico Ltda.	Argentina		\$ 18.00	10	0%	\$ 180.00	\$ 217.86		
15	Océano Atlántico Ltda.	Argentina		\$ 13.00	15	0%	\$ 195.00	\$ 217.86		
16	<b>Océano Atlántico Ltda. Total</b>							\$ 3,031.00		
17	Rancho grande	Argentina		\$ 81.00	5	0%	\$ 405.00	\$ 90.85		
18	Rancho grande	Argentina		\$ 263.50	2	0%	\$ 527.00	\$ 90.85		
19	Rancho grande	Argentina		\$ 17.45	6	0%	\$ 104.70	\$ 63.77		
20	Rancho grande	Argentina		\$ 32.00	6	0%	\$ 192.00	\$ 63.77		
21	Rancho grande	Argentina		\$ 19.50	20	0%	\$ 390.00	\$ 63.77		
22	<b>Rancho grande Total</b>							\$ 1,618.70		

**FIGURE 1-9:** Some subtotals applied to the Total column for each customer.



REMEMBER

Note that in the phrase, *automatic subtotals*, the word *subtotals* is misleading because it implies that you can summarize your data only with totals. Not even close! Using “subtotals,” you can also count the values (all the values or just the numeric values), calculate the average of the values, determine the maximum or minimum value, and calculate the product of the values. For statistical analysis, you can also calculate the standard deviation and variance, both of a sample and of a population. To change the summary calculation, follow Steps 1 to 3, open the Use Function drop-down list, and then select the function you want to use for the summary.

## Grouping Related Data

To help you analyze a worksheet, you might be able to control what parts of the worksheet are displayed by grouping the data based on the worksheet formulas and data. Grouping the data creates a worksheet outline, which works similarly to the outline feature in Microsoft Word. In a worksheet outline, you can *collapse* sections of the sheet to display only summary cells (such as quarterly or regional totals), or *expand* hidden sections to show the underlying detail. Note that when you add subtotals to a range, as I describe in the preceding section, Excel automatically groups the data and displays the outline tools.



REMEMBER

Not all worksheets can be grouped, so you need to make sure that your worksheet is a candidate for outlining:

- » The worksheet must contain formulas that reference cells or ranges directly adjacent to the formula cell. Worksheets with SUM functions that subtotal cells above or to the left are particularly good candidates for outlining.
- » There must be a consistent pattern to the direction of the formula references. For example, a worksheet with formulas that always reference cells above or to the left can be outlined. Excel won't outline a worksheet with, say, SUM functions where some of the range references are above the formula cell and some are below.

Here are the steps to follow group-related data:

1. **Display the worksheet you want to outline.**
2. **Choose Data ⇨ Group ⇨ Auto Outline.**

If you don't see the Group command, choose Outline ⇨ Group. Excel outlines the worksheet data.

As shown in Figure 1-10, Excel uses *level bars* to indicate the grouped ranges and *level numbers* to indicate the various levels of the underlying data available in the outline.

Level bar

Level numbers

Collapse button

Expand button

	A	B	C	D	E	I	M	Q	R
1		Jan	Feb	Mar	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter	TOTAL
2	<b>Sales</b>								
3	Division I	23,500	23,000	24,000	70,500	75,500	74,000	74,000	294,000
4	Division II	28,750	27,800	29,500	86,050	91,500	90,000	91,000	358,550
5	Division III	24,400	24,000	25,250	73,650	80,350	77,500	78,500	310,000
6	<b>SALES TOTAL</b>	76,650	74,800	78,750	230,200	247,350	241,500	243,500	962,550
7	<b>Expenses</b>								
8	Cost of Goods	6,132	5,984	6,700	18,816	19,788	19,320	19,480	77,404
9	Advertising	4,600	4,200	5,200	14,000	15,750	15,900	14,900	60,550
10	Rent	2,100	2,100	2,100	6,300	6,300	6,300	6,300	25,200
11	Supplies	1,300	1,200	1,400	3,900	3,950	4,100	4,000	15,950
12	Salaries	16,000	16,000	16,500	48,500	50,000	51,000	52,000	201,500
13	Shipping	14,250	13,750	14,500	42,500	44,250	44,000	46,500	177,250
14	Utilities	500	600	600	1,700	1,800	1,850	1,925	7,275
15	<b>EXPENSES TOTAL</b>	44,882	43,834	47,000	135,716	141,838	142,470	145,105	565,129
16	<b>GROSS PROFIT</b>	31,768	30,966	31,750	94,484	105,512	99,030	98,395	397,421

**FIGURE 1-10:** When you group a range, Excel displays its outlining tools.

Here are some ways you can use the outline to control the range display:

- » Click the – (collapse) button to hide the range indicated by the level bar.
- » Click the + button (expand) for a collapsed range to view it again.
- » Click a level number to collapse multiple ranges on the same outline level.
- » Click a level number to display multiple collapsed ranges on the same outline level.

## Consolidating Data from Multiple Worksheets

Companies often distribute similar worksheets to multiple departments to capture budget numbers, inventory values, survey data, and so on. Those worksheets must then be combined into a summary report showing company-wide totals. Combining multiple worksheets into a summary report is called *consolidating* the data.

Sounds like a lot of work, right? It sure is, if you do it manually, so forget that. Instead, Excel can consolidate your data automatically. You can use the consolidate feature to consolidate the data in either of two ways:

- » **By position:** Excel consolidates the data from two or more worksheets, using the same range coordinates on each sheet. Use this method if the worksheets you're consolidating have an identical layout.
- » **By category:** Excel consolidates the data from two or more worksheets by looking for identical row and column labels in each sheet. Reach for this method if the worksheets you're consolidating have different layouts but common labels.

In both cases, you specify one or more *source ranges* (the ranges that contain the data you want to consolidate) and a *destination range* (the range where the consolidated data will appear).

### Consolidating by position

Here are the steps to trudge through if you want to consolidate multiple worksheets by position:

1. **Create a new worksheet that uses the same layout — including row and column labels — as the sheets you want to consolidate.**

The identical layout in this new worksheet is your destination range.

2. **If necessary, open the workbooks that contain the worksheets you want to consolidate.**

If the worksheets you want to consolidate are in the current workbook, you can skip this step.

3. **In the new worksheet from Step 1, select the upper-left corner of the destination range.**

4. **Choose Data ⇨ Consolidate.**

The Consolidate dialog box appears.

5. **In the Function list, select the summary function you want to use.**

6. **In the Reference text box, select one of the ranges you want to consolidate.**

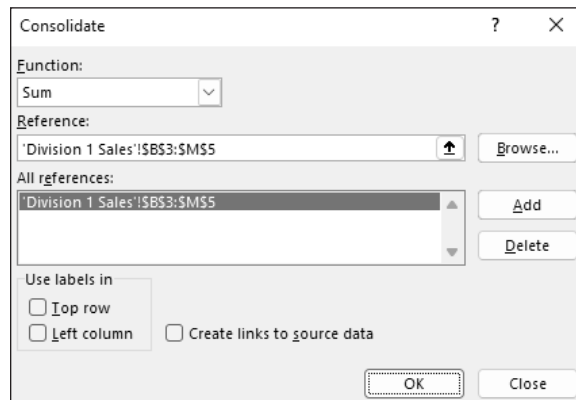
7. **Click Add.**

Excel adds the range to the All References list, as shown in Figure 1-11.

8. **Repeat Steps 6 and 7 to add all the consolidation ranges.**

9. **Click OK.**

Excel consolidates the data from the source ranges and displays the summary in the destination range.



**FIGURE 1-11:** Consolidate multiple worksheets by adding a range from each one.



TIP

If the source data changes, you probably want to reflect those changes in the consolidation worksheet. Rather than run the entire consolidation over again, a much easier solution is to select the Create Links to Source Data check box in the Consolidate dialog box. You can then update the consolidation worksheet by choosing Data ⇨ Refresh All.

## Consolidating by category

Here are the steps to follow to consolidate multiple worksheets by category:

**1. Create a new worksheet for the consolidation.**

You use this worksheet to specify your destination range.

**2. If necessary, open the workbooks that contain the worksheets you want to consolidate.**

If the worksheets you want to consolidate are in the current workbook, you can skip this step.

**3. In the new worksheet from Step 1, select the upper-left corner of the destination range.**

**4. Choose Data ⇨ Consolidate.**

The Consolidate dialog box appears.

**5. In the Function list, select the summary function you want to use.**

**6. In the Reference text box, select one of the ranges you want to consolidate.**

When you're selecting the range, be sure to include the row and column labels in the range.

**7. Click Add.**

Excel adds the range to the All References list.

**8. Repeat Steps 6 and 7 to add all the consolidation ranges.**

**9. If you have labels in the top row of each range, select the Top Row check box.**

**10. If you have labels in the left-column row of each range, select the Left Column check box.**

Figure 1-12 shows a completed version of the Consolidate dialog box.

**11. Click OK.**

Excel consolidates the data from the source ranges and displays the summary in the destination range.

**FIGURE 1-12:**  
When consolidating by category, tell Excel where your labels are located.

