6 Reordering and Summarizing Data

In this chapter, you will learn how to

- ✔ Sort worksheet data.
- ✔ Organize data into levels.
- ✔ Look up information in a worksheet.

Most of the time, when you enter data in a Microsoft Excel worksheet, you will enter it in chronological order. For example, you could enter hourly shipment data in a worksheet, starting with the first hour of the day and ending with the last hour. The data would naturally be displayed in the order you entered it, but that might not always be the best arrangement to answer your questions. For instance, you might want to sort your data so that the top row in your worksheet shows the day of the month with the highest package volume, with subsequent rows displaying the remaining days in decreasing order of package volumes handled. You can also sort based on the contents of more than one column. A good example is sorting package handling data by week, day, and then hour of the day.

After you have sorted your data into the desired order with Excel, you can find partial totals, or subtotals, for groups of cells within a given range. Yes, you can create formulas to find the sum, average, or standard deviation of data in a cell range, but you can do the same thing much more quickly by having Excel calculate the total for rows that have the same value in one of their columns. For example, if your worksheet holds sales data for a list of services, you can calculate subtotals for each product category.

When you calculate subtotals in a worksheet, Excel creates an outline that marks the cell ranges used in each subtotal. For example, if the first 10 rows of a worksheet contain overnight shipping data, and the second 10 rows contain second-day shipping data, Excel divides the rows into two units. You can use the markers on the worksheet to hide or display the rows used to calculate a subtotal; in this case, you can hide all the rows that contain overnight shipping data, hide all the rows that contain second-day shipping data, hide both, or show both.

Excel also has a capability you might expect to find only in a database program—in Excel, you can type a value in a cell and have Excel look in a named range to find a corresponding value. For instance, you can have a two-column named range with one column displaying customer identification numbers and the second column displaying the name of the company assigned each number. By using a *VLOOKUP* formula that references the named range, you can let colleagues using your workbook type a customer identification number in a cell and have the name of the corresponding company appear in the cell with the formula.

In this chapter, you'll learn how to sort your data using one or more criteria, calculate subtotals, organize your data into levels, and look up information in a worksheet.

> **Practice Files** Before you can complete the exercises in this chapter, you need to copy the book's practice files to your computer. The practice files you'll use to complete the exercises in this chapter are in the Chapter06 practice file folder. A complete list of practice files is provided in "Using the Practice Files" at the beginning of this book.

Sorting Worksheet Data

Although Excel makes it easy to enter your business data and to manage it after you've saved it in a worksheet, unsorted data will rarely answer every question you want to ask it. For example, you might want to discover which of your services generates the most profits, which service costs the most for you to provide, and so on. You can discover that information by sorting your data.

When you sort data in a worksheet, you rearrange the worksheet rows based on the contents of cells in a particular column or set of columns. For instance, you can sort a worksheet to find your highest-revenue services.

You can sort a group of rows in a worksheet in a number of ways, but the first step is to identify the column that will provide the values by which the rows should be sorted. In the revenue example, you could find the highest revenue totals by sorting on the cells in the Revenue column. First you would select the cells in the Revenue column and display the Home tab. Then, in the Editing group, in the Sort & Filter list, click Sort Largest

To Smallest. Clicking Sort Largest To Smallest makes Excel put the row with the highest value in the Revenue column at the top of the worksheet and continue down to the lowest value.


If you want to sort the rows in the opposite order, from the lowest revenue to the highest, select the cells in the Revenue column and then, in the Sort & Filter list, click Sort Smallest To Largest.

Tip The exact set of values that appears in the Sort & Filter list changes to reflect the data in your column. If your column contains numerical values, you'll see the options Sort Largest To Smallest, Sort Smallest To Largest, and Custom List. If your column contains text values, the options will be Sort A To Z (ascending order), Sort Z To A (descending order), and Custom List. And if your column contains dates, you'll see Sort Newest To Oldest, Sort Oldest To Newest, and Custom List.

The Sort Smallest To Largest and Sort Largest To Smallest options enable you to sort rows in a worksheet quickly, but you can use them only to sort the worksheet based on the contents of one column, even though you might want to sort by two columns. For example, you might want to order the worksheet rows by service category and then by total so that you can see the customers that use each service category most frequently. You can sort rows in a worksheet by the contents of more than one column by using the Sort dialog box, in which you can pick any number of columns to use as sort criteria and choose whether to sort the rows in ascending or descending order.

To display the Sort dialog box, click Custom Sort in the Sort & Filter list.
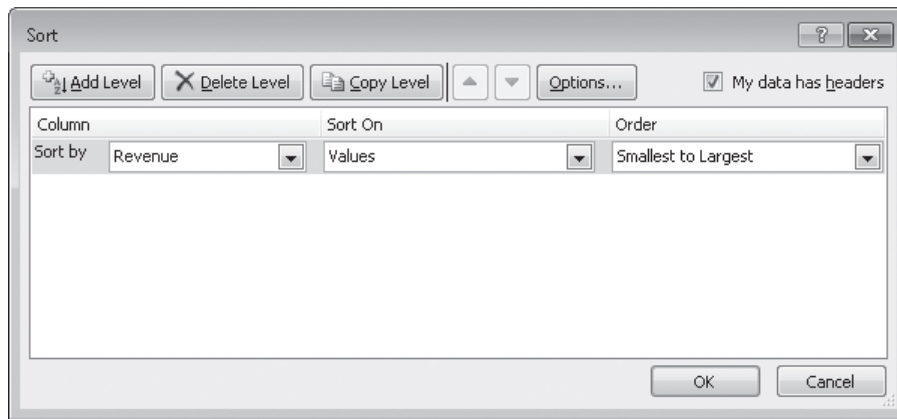
If your data has a header row, select the My Data Has Headers check box so the column headers will appear in the Sort By list. After you identify the column by which you want to sort, the Sort On list enables you to select whether you want to sort by a cell's value (the default), a cell's fill color, a cell's font color, or an icon displayed in the cell.

See Also For more information about creating conditional formats that change a cell's formatting or display icon to reflect the cell's value, see "Changing the Appearance of Data Based on Its Value" in Chapter 4, "Changing Workbook Appearance."

Finally, from the Order list, you can select how you want Excel to sort the column values. As with the Sort & Filter button's list, the exact values that appear in the Order list change to reflect the data you want to sort.

Adding, moving, copying, and deleting sorting levels are a matter of clicking the appropriate button in the Sort dialog box. To add a second level to your sort, click the Add Level button.

Tip In Excel 2003 and earlier versions of the program, you could define a maximum of three sorting levels. You can create up to 64 sorting levels in Excel 2010.

To delete a level, click the level in the list, and then click Delete Level. Clicking the Copy Level button enables you to put all the settings from one rule into another, saving you some work if you need to change only one item. The Move Up and Move Down buttons, which display an upward-pointing arrow and a downward-pointing arrow, respectively, enable you to change a sorting level's position in the order. Finally, clicking the Options button displays the Sort Options dialog box, which you can use to make a sorting level case sensitive and to change the orientation of the sort.

The default setting for Excel is to sort numbers according to their values and to sort words in alphabetical order, but that pattern doesn't work for some sets of values. One example in which sorting a list of values in alphabetical order would yield incorrect results is the months of the year. In an "alphabetical" calendar, April is the first month and September is the last! Fortunately, Excel recognizes a number of special lists, such as days of the week and months of the year. You can have Excel sort the contents of a worksheet based on values in a known list; if needed, you can create your own list of values. For example, the default lists of weekdays in Excel start with Sunday. If you keep your business records based on a Monday–Sunday week, you can create a new list with Monday as the first day and Sunday as the last.

To create a new list, type the list of values you want to use as your list into a contiguous cell range, select the cells, click the File tab, and then click Options. On the Advanced page of the Excel Options dialog box, in the General group near the bottom of the page, click the Edit Custom Lists button to display the Custom Lists dialog box.
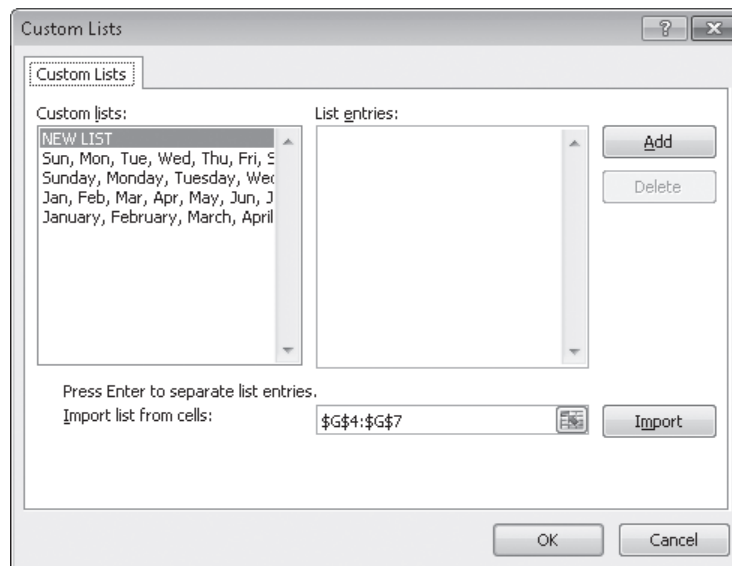
The selected cell range's reference appears in the Import List From Cells field. To record your list, click the Import button.

If you prefer, you can type the list in the List Entries box, to the right of the Custom Lists box.

Tip Another benefit of creating a custom list is that dragging the fill handle of a list cell that contains a value causes Excel to extend the series for you. For example, if you create the list *Spring, Summer, Fall, Winter*, then type Summer in a cell and drag the cell's fill handle, Excel extends the series as *Fall, Winter, Spring, Summer, Fall*, and so on.

To use a custom list as a sorting criterion, display the Sort dialog box, click the rule's Order arrow, click Custom List, and select your list from the dialog box that opens.

Tip In Excel 2003 and earlier, your custom list had to be the primary sorting criterion. In Excel 2010 and Excel 2007, you can use your custom list as any criterion.

In this exercise, you'll sort worksheet data, sort by multiple criteria, change the order in which sorting criteria are applied, sort data by using a custom list, and sort data by color.

SET UP You need the ShippingSummary_start workbook located in your Chapter06 practice file folder to complete this exercise. Start Excel, open the ShippingSummary_start workbook, and save it as *ShippingSummary*. Then follow the steps.

1. Click cell **C3**.

2. On the **Home** tab, in the **Editing** group, click the **Sort & Filter** button and then, in the list, click **Sort A to Z**.

   Excel sorts the data by season, with the seasons listed in alphabetical order.

3. In the **Sort & Filter** list, click **Custom Sort**.

   The Sort dialog box opens and displays the parameters of the sort you just applied.

4. If it's not already selected, select the **My data has headers** check box.

5. In the **Column** list, click **Customer**. If necessary, in the **Sort On** list, click **Values**; then in the **Order** list, click **A to Z**.

6. Click **Add Level**.

   A new Then By sorting level appears.

7. In the new **Column** list, click **Revenue**.

8. In the new **Order** list, click **Largest to Smallest**.

9. Click **OK**.

Excel closes the Sort dialog box and sorts the data list.

10. In the **Sort & Filter** list, click **Custom Sort**.

    The Sort dialog box opens.

11. Click **Then by**.

    Excel highlights the Revenue sorting rule.

12. Click the **Move Up** button.

    Excel moves the Revenue sorting rule above the Customer sorting rule.
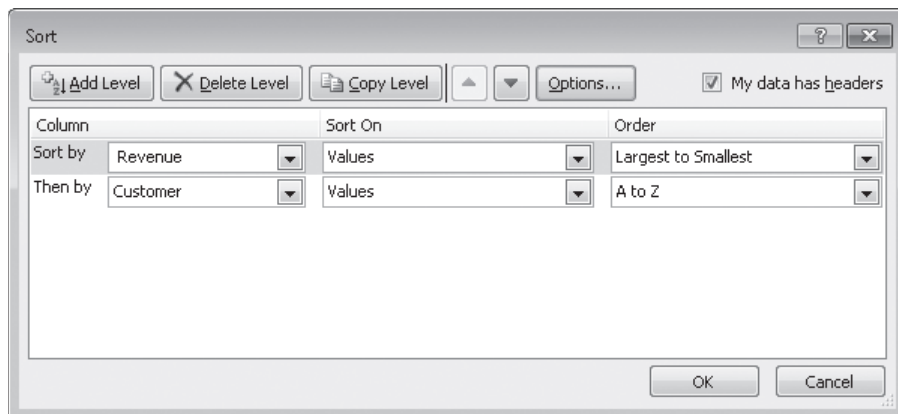
13. Click **OK**.

    Excel closes the Sort dialog box and sorts the data list.

14. Select cells **G4:G7**, click the **File** tab, and then click **Options**.

    The Excel Options dialog box opens.

15. On the **Advanced** page, in the **General** group toward the bottom of the page, click **Edit Custom Lists**.

    The Custom Lists dialog box opens.

16. Verify that the cell range **$G$4:$G$7** appears in the **Import list from cells** field, and then click **Import**.

    The new list appears in the Custom Lists pane.
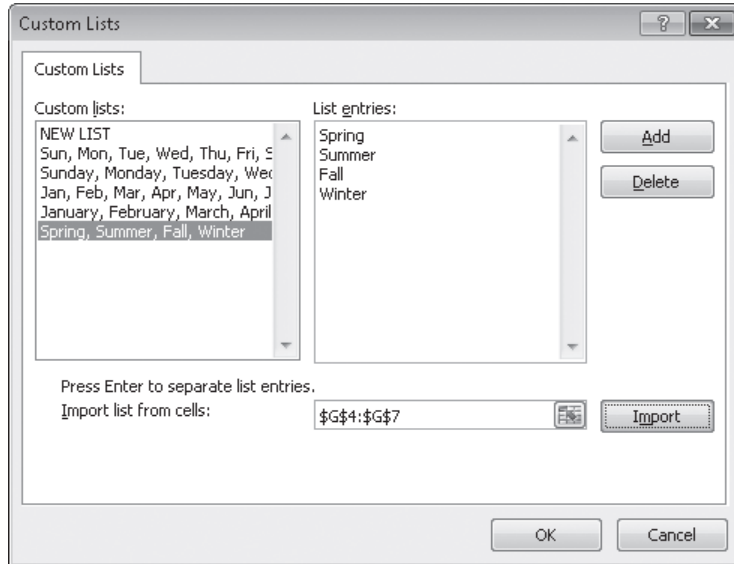
17. Click **OK** twice to close the **Custom Lists** dialog box and the **Excel Options** dialog box.

18. Click cell **C3**.

19. On the **Home** tab, in the **Editing** group, click **Sort & Filter**, and then click **Custom Sort**.

    The Sort dialog box opens, displaying the sorting operation you defined earlier.

20. Click the rule in the **Sort by** row, and then click **Delete Level**.

    The sorting rule disappears.

21. If necessary, in the new **Sort by** row, in the **Column** list, click **Season**.

22. In the same row, in the **Order** list, click **Custom List**.

    The Custom Lists dialog box opens.

23. In the **Custom lists** pane, click the sequence **Spring, Summer, Fall, Winter**.

24. Click **OK** twice to close the **Custom Lists** dialog box and the **Sort** dialog box.

    Excel sorts the data list.

25. Click cell **C3**. Then on the **Home** tab, in the **Editing** group, click **Sort & Filter**, and click **Custom Sort.**

    The Sort dialog box opens.

26. In the **Sort by** row, in the **Column** list, click **Revenue**.

27. In the **Sort on** list, click **Cell Color**.

28. In the new list control that appears in the **Sort by** row, click **On Bottom** to have Excel put the Revenue cells that have no cell color on the bottom.

29. Click **OK**.

    Excel sorts the data list.

✖ **CLEAN UP** Save the ShippingSummary workbook, and then close it.

Organizing Data into Levels

After you have sorted the rows in an Excel worksheet or entered the data so that it doesn't need to be sorted, you can have Excel calculate subtotals or totals for a portion of the data. In a worksheet with sales data for three different product categories, for example, you can sort the products by category, select all the cells that contain data, and then open the Subtotal dialog box. To open the Subtotal dialog box, display the Data tab and then, in the Outline group, click Subtotal.
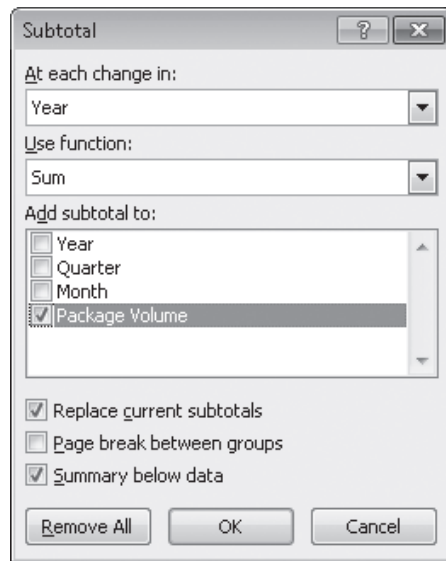
In the Subtotal dialog box, you can choose the column on which to base your subtotals (such as every change of value in the Week column), the summary calculation you want to perform, and the column or columns with values to be summarized. After you define your subtotals, they appear in your worksheet.

Troubleshooting The appearance of buttons and groups on the ribbon changes depending on the width of the program window. For information about changing the appearance of the ribbon to match our screen images, see "Modifying the Display of the Ribbon" at the beginning of this book.

When you add subtotals to a worksheet, Excel also defines groups based on the rows used to calculate a subtotal. The groupings form an outline of your worksheet based on the criteria you used to create the subtotals. For example, all the rows representing months in the year 2009 could be in one group, rows representing months in 2010 in another, and so on. The outline area at the left of your worksheet holds controls you can use to hide or display groups of rows in your worksheet.

Three types of controls can appear in the outline area: Hide Detail buttons, Show Detail buttons, and level buttons. The Hide Detail button beside a group can be clicked to hide the rows in that group. In a worksheet that has a subtotal group consisting of rows 2 through 13, clicking the Hide Detail button next to row 14 would hide rows 2 through 13 but leave the row holding the subtotal for that group, row 14, visible.

When you hide a group of rows, the button displayed next to the group changes to a Show Detail button. Clicking a group's Show Detail button restores the rows in the group to the worksheet.

The level buttons are the other buttons in the outline area of a worksheet with subtotals. Each button represents a level of organization in a worksheet; clicking a level button hides all levels of detail below that of the button you clicked. The following table describes the data contained at each level of a worksheet with three levels of organization.

Level	Description
1	Grand total
2	Subtotals for each group
3	Individual rows in the worksheet

Clicking the Level 2 button in the worksheet would hide the rows with data on, for example, each month's revenue but would leave the row that contains the grand total (Level 1) and all rows that contain the subtotal for each year (Level 2) visible in the worksheet.

If you like, you can add levels of detail to the outline that Excel creates. For example, you might want to be able to hide revenues from January and February, which you know are traditionally strong months. To create a new outline group within an existing group, select the rows you want to group; on the Data tab, in the Outline group, point to Group, and then click Group.

You can remove a group by selecting the rows in the group and then, in the Outline group, clicking Ungroup.

Tip If you want to remove all subtotals from a worksheet, open the Subtotal dialog box, and click the Remove All button.

In this exercise, you'll add subtotals to a worksheet and then use the outline that appears to show and hide different groups of data in your worksheet.

SET UP You need the GroupByQuarter_start workbook located in your Chapter06 practice file folder to complete this exercise. Open the GroupByQuarter_start workbook, and save it as *GroupByQuarter*. Then follow the steps.

1. Click any cell in the data list.

2. On the **Data** tab, in the **Outline** group, click **Subtotal**.

   The Subtotal dialog box opens with the default options to add a subtotal at every change in the Year column, to return the sum of the values in the subtotaled rows, and to add a row with the subtotal of values in the Package Volume column below the final selected row.

3. Click **OK**.

The Subtotal dialog box closes. New rows appear with subtotals for package volume during each year represented in the worksheet. The new rows are numbered 14 and 27. A row with the grand total of all rows also appears; that row is row 28. A new area with outline bars and group-level indicators appears to the left of column A.

4. Click the row heading of row **5**, and drag to the row heading of row **7**.

Rows 5 through 7 are selected.

5. On the **Data** tab, in the **Outline** group, click **Group**.

Rows 5 through 7 are made into a new group. An outline bar appears on a new level in the outline area, and a corresponding Level 4 button appears at the top of the outline area.

6. In the outline area, click the **Hide Detail** button next to row **8**.

Rows 5 through 7 are hidden, and the Hide Detail button you clicked changes to a Show Detail button.

7. In the outline area, click the **Show Detail** button next to row **8**.

Rows 5 through 7 reappear.

8. In the outline area, click the **Level 1** button.

All rows except row 1 with the column headings and row 28 with the grand total are hidden.

9. In the outline area, click the **Level 2** button.

The rows with the subtotal for each year appear.

10. In the outline area, click the **Level 3** button.

All rows except rows 5 through 7 appear.

11. In the outline area, click the **Level 4** button.

Rows 5 through 7 reappear.

✖ **CLEAN UP** Save the GroupByQuarter workbook, and then close it.

Looking Up Information in a Worksheet

Whenever you create a worksheet that holds information about a list of distinct items, such as products offered for sale by a company, you should ensure that at least one column in the list contains a unique value that distinguishes that row (and the item the row represents) from every other row in the list. Assigning each row a column that contains a unique value means that you can associate data in one list with data in another list. For example, if you assign every customer a unique identification number, you can store a customer's contact information in one worksheet and all orders for that customer in another worksheet. You can then associate the customer's orders and contact information without writing the contact information in a worksheet every time the customer places an order.

In the case of shipments handled by Consolidated Messenger, the column that contains those unique values, also known as the primary key column, is the ShipmentID column.

If you know a shipment's ShipmentID, it's no trouble to look through a list of 20 or 30 items to find a particular shipment. If, however, you have a list of many thousands of shipments, looking through the list to find one would take quite a bit of time. Instead, you can use the *VLOOKUP* function to let your colleagues type a ShipmentID in a cell and have the corresponding details appear in another cell.

The *VLOOKUP* function finds a value in the leftmost column of a named range, such as a table, and then returns the value from the specified cell to the right of the cell with the found value. A properly formed *VLOOKUP* function has four arguments (data that is passed to the function), as shown in the following definition: *=VLOOKUP(lookup_value, table_array, col_index_num, range_lookup)*.

The following table summarizes the values Excel expects for each of these arguments.

Argument	Expected value
lookup_value	The value to be found in the first column of the named range specified by the *table_array* argument. The *lookup_value* argument can be a value, a cell reference, or a text string.
table_array	The multicolumn range or name of the range or data table to be searched.
col_index_num	The number of the column in the named range with the value to be returned.
range_lookup	A *TRUE* or *FALSE* value, indicating whether the function should find an approximate match (*TRUE*) or an exact match (*FALSE*) for the *lookup_value*. If left blank, the default value for this argument is *TRUE*.

Important When *range_lookup* is left blank or set to *TRUE*, for *VLOOKUP* to work properly the rows in the named range specified in the *table_array* argument must be sorted in ascending order based on the values in the leftmost column of the named range.

The *VLOOKUP* function works a bit differently depending on whether the *range_lookup* argument is set to *TRUE* or *FALSE*. The following list summarizes how the function works based on the value of *range_lookup*:

- If the *range_lookup* argument is left blank or set to *TRUE*, and *VLOOKUP* doesn't find an exact match for *lookup_value*, the function returns the largest value that is less than *lookup_value*.

- If the *range_lookup* argument is left blank or set to *TRUE*, and *lookup_value* is smaller than the smallest value in the named range, an *#N/A* error is returned.

● If the *range_lookup* argument is left blank or set to *TRUE*, and *lookup_value* is larger than all values in the named range, the largest value in the named range is returned.

● If the *range_lookup* argument is set to *FALSE*, and *VLOOKUP* doesn't find an exact match for *lookup_value*, the function returns an *#N/A* error.

As an example of a *VLOOKUP* function, consider the following data, which shows an Excel table with its headers in row 2 and the first column in column B of the worksheet.

CustomerID	Customer
CI01	Fabrikam
CI02	Northwind Traders
CI03	Tailspin Toys
CI04	Contoso

If the *=VLOOKUP (E3, Table1, 2, TRUE)* formula is used, when you type *CI02* in cell E3 and press Enter, the *VLOOKUP* function searches the first column of the table, finds an exact match, and returns the value *Northwind Traders* to cell F3.
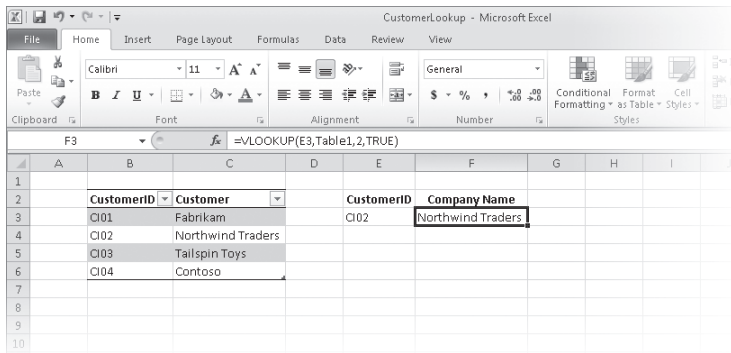
Tip The related*HLOOKUP* function matches a value in a column of the first row of a table and returns the value in the specified row number of the same column. The letter "H" in the *HLOOKUP* function name refers to the horizontal layout of the data, just as the "V" in the *VLOOKUP* function name refers to the data's vertical layout. For more information on using the *HLOOKUP* function, click the Excel Help button, type *HLOOKUP* in the search terms box, and then click Search.

Important Be sure to give the cell in which you type the *VLOOKUP* formula the same format as the data you want the formula to display. For example, if you create a *VLOOKUP* formula in cell G14 that finds a date, you must apply a date cell format to cell G14 for the result of the formula to display properly.

In this exercise, you'll create a *VLOOKUP* function to return the destination postal code of deliveries with ShipmentIDs typed in a specific cell.

SET UP You need the ShipmentLog_start workbook located in your Chapter06 practice file folder to complete this exercise. Open the ShipmentLog_start workbook, and save it as *ShipmentLog*. Then follow the steps.

1. In cell **C3**, type the formula **=VLOOKUP(B3, Shipments, 5, FALSE)**.

   Cell B3, which the formula uses to look up values in the Shipments table, is blank, so the *#N/A* error code appears in cell C3.

2. In cell **B3**, type **SI3049224**, and press Enter.

   The value *51102* appears in cell C3.

3. In cell **C3**, edit the formula so that it reads **=VLOOKUP(B3, Shipments, 2, FALSE)**.

   The formula now finds its target value in table column 2 (the CustomerID column), so the value *CI512191* appears in cell C3.

4. In cell **C3**, edit the formula so that it reads **=VLOOKUP(B3, Shipments, 4, TRUE)**.

   Changing the last argument to *TRUE* enables the *VLOOKUP* formula to find an approximate match for the ShipmentID in cell B3, whereas changing the column to *4* means the formula gets its result from the OriginationPostalCode column. The value *14020* appears in cell C3.

5. In cell **B3**, type **SI3049209**.

   The value in cell B3 is smaller than the smallest value in the Shipments table's first column, so the *VLOOKUP* formula displays the #*N/A* error code in cell C3.

6. In cell **B3**, type **SI3049245**.

   The ShipmentID typed into cell B3 is greater than the last value in the table's first column, so the *VLOOKUP* formula displays the last value in the target column (in this case, the fourth column). Therefore, the value *44493* appears in cell C3.

   ✖ **CLEAN UP Save the ShipmentLog workbook, and then close it. If you are not continuing directly to the next chapter, exit Excel.**

Key Points

● You can rearrange the data in a worksheet quickly by clicking either the Sort Ascending or Sort Descending button in the Sort & Filter group on the Data tab.

● Don't forget that you can sort the rows in a worksheet by using orders other than alphabetical or numerical. For example, you can sort a series of days based on their order in the week or by cell color.

● If none of the existing sort orders (days, weekdays, and so on) meets your needs, you can create your own custom sort order.

● You can divide the data in your worksheet into levels and find a subtotal for each level.

● Creating subtotals enables you to show or hide groups of data in your worksheets.

● Use the *VLOOKUP* function to look up a value in one column of data and return a value from another column in the same row.
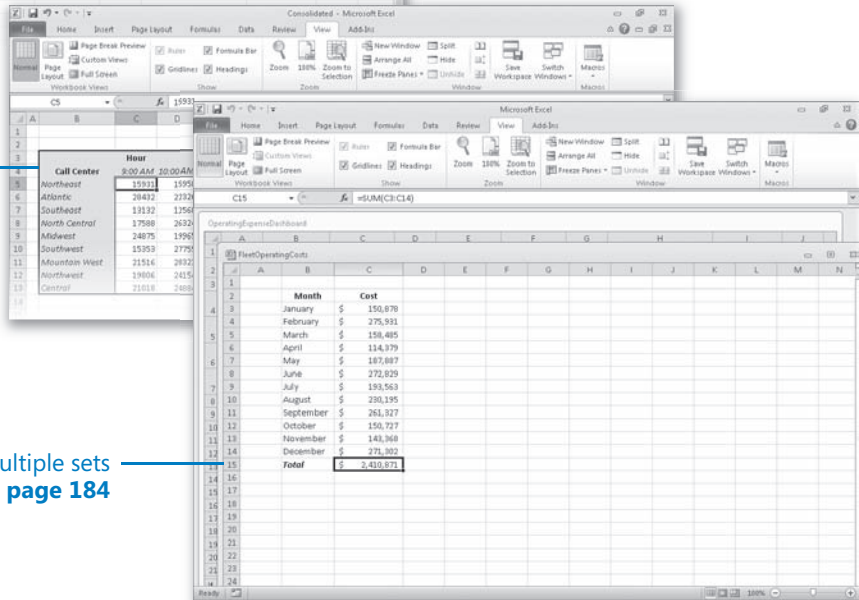
Chapter at a Glance

Use workbooks as templates for other workbooks, **page 168**
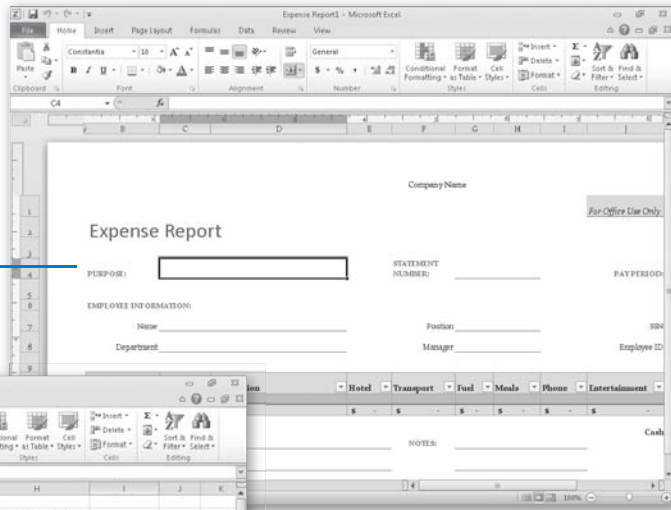
Link to data in other worksheets and workbooks, **page 175**
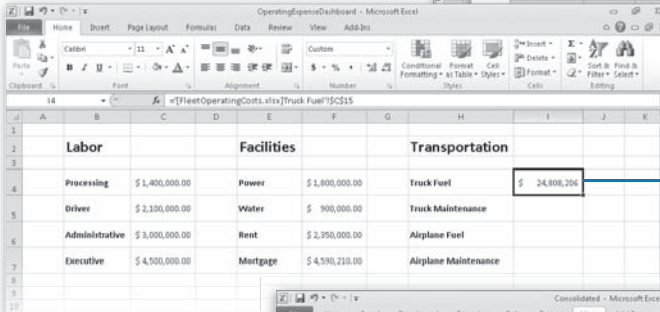
Consolidate multiple sets of data into a single workbook, **page 180**

Group multiple sets of data, **page 184**