After studying this chapter, you should be able to:

1. Explain how the "percent of sales" method is used to develop pro forma financial statements and how to construct such statements in Excel.
2. Use circular references to perform iterative calculations.
3. Use the Trend function for forecasting sales or any other trending variables.
4. Perform a regression analysis with Excel's built-in regression tools.
5. Determine if a variable is statistically significant in a regression analysis.

Forecasting is an important activity for a wide variety of business people. Nearly all of the decisions made by financial managers are made on the basis of forecasts of one kind or another. For example, in Chapter 3 we’ve seen how the cash budget can be used to forecast short-term borrowing and investing needs. Every item in the cash budget is itself a forecast. In this chapter, we will examine several methods of forecasting. The first, the percent of sales method, is the simplest. We will also look at more advanced techniques, such as regression analysis.
The Percent of Sales Method

Forecasting financial statements is important for a number of reasons. Among these are planning for the future and providing information to the company’s investors. The simplest method of forecasting income statements and balance sheets is the percent of sales method. This method has the added advantage of requiring relatively little data to make a forecast.

The fundamental premise of the percent of sales method is that some, but not all, income statement and balance sheet items maintain a constant relationship with the level of sales. For example, if the cost of goods sold has averaged 65% of sales over the last several years, we would assume that this relationship would hold for the next year. If sales are expected to be $10 million next year, our cost of goods forecast would be $6.5 million (10 million × 0.65 = 6.5 million).

Of course, this method assumes that the forecasted level of sales is already known. There are two primary methods of forecasting sales. The top-down method relies on forecasts of macroeconomic variables (e.g., GDP, inflation rates, etc.) and of the condition of the industry as a whole. These expectations are then converted into a sales forecast for the entire firm, and sales targets for each division or product. The bottom-up method involves discussions with customers to determine the expected demand for each product and expectations regarding prices, which are then summed to calculate a firm-wide sales forecast. Of course, firms can use a combination of the two methods. We will take the sales forecast as a given.

Forecasting the Income Statement

As an example of income statement forecasting, consider the Elvis Products International (EPI) statements that you created in Chapter 2. The income statement is recreated here in Exhibit 5-1. Recall that we have used a custom number format to display this data in thousands of dollars, but that the full-precision numbers are there. Open the workbook that you created for Chapter 2, and make a copy of the Income Statement worksheet. Rename the new worksheet to Pro Forma Income Statement.¹

The level of detail that you have in an income statement will affect the number of items that will fluctuate directly with sales. In general, we will proceed through the income statement line by line asking the question, “Is it likely that this item will change proportionally with sales?” If the answer is yes, then we calculate the percentage of sales and multiply the result by the sales forecast for the next period. Otherwise, we will take one of two actions: Leave the item unchanged, or use other information to change

¹. Pro forma is a Latin word that, for our purposes, can be interpreted to mean “as if.” That is, these forecasted financial statements are presented as if the forecast time period has already happened.
If you don’t know the answer, then you can create a chart that compares the item to sales over the last several quarters or years. It should be obvious if there is a relationship, though you may need to use some of the statistical tools, discussed on page 154, to determine the form of the relationship.

EXHIBIT 5-1
EPI’S INCOME STATEMENTS FOR 2010 AND 2011

For EPI, only one income statement item will clearly change with sales: the cost of goods sold. Another item, SG&A (selling, general, and administrative) expense, is an aggregation of many things, some of which will probably change with sales and some that won’t. For our purposes we choose to believe that, on balance, SG&A will change along with sales.

Changes in the other items are not directly related to a change in sales in the short term. Depreciation expense, for example, depends on the amount and age of the firm’s fixed assets. Interest expense is a function of the amount and maturity structure of debt in the firm’s capital structure. These items may, and probably will, change but we will need additional information. Taxes depend directly on the firm’s taxable income, though this indirectly depends on the level of sales. All of the other items on the income statement are calculated directly.

2. For example, if you know that the lease for the company’s headquarters building has a scheduled increase, then you should be sure to include this information in your forecast for fixed costs.
Before getting started with the forecast, insert a column to the left of column B. Select a cell in column B and click the Insert button on the Home tab and then choose Insert Sheets Columns. Note that a Smart Tag will appear that will give you three choices: (1) Format Same As Left; (2) Format Same As Right; or (3) Clear Formatting. Choose the second option so that the custom number formats and column width will automatically be applied. So that we can experiment later if we choose, enter 40% for the tax rate in B18.

To generate our income statement forecast, we first determine the percentage of sales for each of the prior years for each item that changes. In this case, for 2011 we have:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$3,250,000</td>
<td>84.42%</td>
<td>$3,609,108</td>
</tr>
<tr>
<td>$330,300</td>
<td>8.58%</td>
<td>$334,803</td>
</tr>
</tbody>
</table>

The 2010 percentages (83.45% and 6.99%, respectively) can be found in exactly the same manner. We now calculate the average of these percentages and use this average as our estimate of the 2012 percentage of sales. The forecast is then found by multiplying these percentages by next year’s sales forecast. Assuming that sales are forecasted to be $4,300,000 in 2012 we have:

\[
\text{Cost of Goods Sold 2012 Forecast} = 4,300,000 \times 0.8393 = 3,609,108
\]

\[
\text{SG&A Expense 2012 Forecast} = 4,300,000 \times 0.0779 = 334,803
\]

Exhibit 5-2 shows a forecast of the complete 2012 income statement. To create this forecast in your worksheet, in B4 enter: 2012. Because the 2012 income statement will be calculated in exactly the same way as 2011, the easiest way to proceed is to copy C5:C15 into B5:B15. This will save you from having to enter formulas to calculate subtotals (e.g., EBIT) and will apply the cell borders. Insert a row above row 17, and in A16 type: *Forecast.

First, in B5 enter the sales forecast: $4,300,000. Now, we can calculate the 2012 cost of goods forecast in B6 with the formula: =AVERAGE(C6/C$5,D6/D$5)*B$5. This formula calculates the average of the cost of goods as a percentage of sales for the last two

---

3. We have chosen to apply a custom format so that the number has an asterisk to indicate a footnote that informs the reader that these are forecasts. The custom format is #"**".
years and then multiplies it by the sales forecast. The result should be as shown above. Now copy this formula to B8 to get the forecast for SG&A expense.

**EXHIBIT 5-2**

**PERCENT OF SALES FORECAST FOR 2012**

Instead of performing the entire calculation in cells B6 and B8, we could have used a *helper column*. A helper column is used to do intermediate calculations and is sometimes useful. In this case, we could have calculated the average percentage of sales for each item in, say, column K. We would then use these values to perform the final calculation in column B. For example, K6 might contain the formula:  

\[ =\text{AVERAGE}(C6/C$5, D6/D$5) \]

Then the formula in B6 would be:  

\[ =K6*B$5 \]

This technique would allow you to easily see the average percentages (as in a common-size income statement) that are being used to generate the forecast. Although this might be useful, it can be an inefficient use of the spreadsheet unless it is necessary.

Assume that we do not have any information regarding changes in fixed expenses, so copy the value from C9. However, we have been informed that the firm intends to invest $50,000 in fixed assets in 2012. This will cause depreciation expense to rise by $5,000. We need to document this assumption, so in A20 type: *Additional Depreciation*, and in B20 enter: 5,000. We will come back to add a formula in B20 in the next section. Don’t forget to apply the same custom number format to this cell that we used in the others.
The formula to calculate depreciation expense in B10 is: \( =C10+B20 \). Because we don’t yet know how the firm will finance these investments, leave the interest expense at the same level as 2011. To calculate the taxes, in B14, use the formula: \( =B19\times B13 \). Your worksheet should now look like the one in Exhibit 5-2.

**Forecasting Assets on the Balance Sheet**

We can forecast the balance sheet in exactly the same way as the income statement, with some major exceptions. For those items that can be expected to vary directly with sales, our formulas will be similar to those we have already seen. We will explain how to handle the other items below.

Create the percent of sales balance sheet for 2012 by selecting column B and inserting a new column. In B4 type the label: 2012. As before, apply a custom number format to display an asterisk after the number. Like we did with the income statement, we will move, line by line, through the balance sheet to determine which items will vary with sales.

The firm’s cash balance is the first, and perhaps the most difficult, item with which we need to work. Does the cash balance vary, in constant proportion, with sales? Your first response might be, “Of course it does. As the firm sells more goods, it accumulates cash.” This line of reasoning neglects two important facts. The firm has other things to do with its cash besides accumulating it, and because cash is a low-return asset, firms should seek to minimize the amount of their cash balance. For these reasons, even though the cash balance will probably change, it probably will not change by the same percentage as sales. Therefore, we will simply use the cash balance from 2011 as our forecast, so enter: \( =C5 \) into cell B5.

The next two items, accounts receivable and inventory, are much easier. Both of these are likely to fluctuate roughly in proportion to sales. Using the same methodology that we used for the pro forma income statement, we will find the average percentage of sales for the past two years and multiply that amount by our sales forecast for 2012. For accounts receivable, the formula in B6 is: \( =\text{AVERAGE}(C6/'\text{Income Statement}'!B$5:D6/'\text{Income Statement}'!C$5)*'\text{Pro Forma Income Statement}'!$B$5 \). Instead of typing the references to the income statement, it is easier to insert them by displaying both the income statement and balance sheet and selecting the appropriate cells with the mouse. Click the View tab and select New Window. This will create an additional view of the workbook. Next, click on the Arrange All button and choose how you would like the worksheets arranged. In the second view, change to the Income Statement worksheet. Now that both worksheets are visible, it is easier to select cells. Because we will use the same formula for

---

4. Within reason, of course. Firms need some amount of cash to operate, but the amount needed does not necessarily vary directly with the level of sales.
inventory, we can simply copy this formula down to B7. Total current assets in B8 is a calculated value, so we can copy the formula directly from cell C8.

In B9, we have the 2012 gross plant and equipment. This is the historical purchase price of the buildings and equipment that the firm owns. As noted earlier, the firm plans to make net new investments of $50,000 in 2012. We will document this assumption by entering Net Addition to Plant and Equipment in A28 and 50,000 in B28. The formula in B9 is: =C9+B28. Note that this increase is not necessarily due to the expected increase in sales. Although gross fixed assets may rise or fall in any given year, most companies always operate with spare capacity so the changes are not, in the short run, directly related to sales.

We now need to calculate the additional depreciation. We will assume that the expected life of the new equipment is 10 years and that it will be depreciated using the straight line method to a salvage value of zero. In A29 enter the label: Life of New Equipment in Years, and in B29 enter 10. In A30 enter: New Depreciation (Straight Line), and in B30 enter the formula: =B28/B29. The additional depreciation expense will be $5,000. Now, return to the pro forma income statement where we will enter a formula in B20: =Pro Forma Balance Sheet'!B30. This last step allows us to change the amount of the new investment and have the additional depreciation expense reflected on the pro forma income statement.

Now, return to the pro forma balance sheet. Accumulated depreciation will definitely increase in 2012 but not because of the forecasted change in sales. Instead, accumulated depreciation will increase by the amount of the depreciation expense for 2012. To determine the accumulated depreciation for 2012, we will add 2012’s depreciation expense to 2011’s accumulated depreciation. The formula is: =C10+’Pro Forma Income Statement'!B10. To complete the asset side of the balance sheet, we note that both net fixed assets and total assets are calculated values. We can simply copy the formulas from C11:C12 and paste them into B11:B12.

**Forecasting Liabilities on the Balance Sheet**

Once the assets are completed, the rest of the balance sheet is comparatively simple because we can mostly copy formulas already entered. Before continuing, however, we need to distinguish among the types of financing sources. We have already seen that the types of financing that a firm uses can be divided into three categories:

- Current liabilities
- Long-term liabilities
- Owner’s equity

These categories are not sufficiently distinguished for our purposes here. Instead, we will divide the liabilities and equity of a firm into two categories:
• **Spontaneous sources of financing**—These are the sources of financing that arise during the ordinary course of doing business. One example is accounts payable. After a credit account is established with a supplier, no additional work is required to obtain credit; it just happens spontaneously when the firm makes a purchase. Note that not all current liabilities are spontaneous sources of financing (e.g., short-term notes payable, long-term debt due in one year).

• **Discretionary sources of financing**—These are the financing sources that require a large effort on the part of the firm to obtain. In other words, the firm must make a conscious decision to obtain these funds. Furthermore, the firm’s upper-level management will use its discretion to determine the appropriate type of financing to use. Examples of this type of financing include any type of bank loan, bonds, preferred stock, and common stock (but not retained earnings).

Generally speaking, spontaneous sources of financing can be expected to vary directly with sales. Changes in discretionary sources, on the other hand, will not have a direct relationship with changes in sales. We always leave discretionary sources of financing unchanged for reasons that will soon become clear.

Returning now to our forecasting problem, the first item to consider is accounts payable. As noted above, accounts payable is a spontaneous source of financing and will, therefore, change directly with sales. To enter the formula, all that is necessary is to copy the formula from one of the other items that we have already completed. Copy the contents of B6 (or B7, it doesn’t matter which) and paste it into B14. The result should indicate a forecasted accounts payable of $189.05.

The next item to consider is the short-term notes payable. Because this is a discretionary source of financing, we will leave it unchanged from 2011. In reality, we might handle this item differently if we had more information. For example, if we knew that the notes would be retired before the end of 2012, we would change our forecast to zero. Alternatively, if the payments on the notes include both principal and interest, our forecast would be the 2011 amount less principal payments that we expect to make in 2012. Because we are leaving it unchanged, the formula in B15 is: =C15.

If we assume that the “other current liabilities” item represents primarily accrued expenses, then it is a spontaneous source of financing. We can, therefore, simply copy the formula from B14 and paste it into B16. The forecasted amount is $163.38.

Long-term debt, in B18, and common stock, in B20, are both discretionary sources of financing. We will leave these balances unchanged from 2011. In B18 the formula is: =C18 and in B20 the formula is: =C20.
The final item that we must consider is retained earnings. Recall that retained earnings accumulates over time. That is, the balance in any year is the accumulated amount that has been added in previous years plus any new additions. The amount that will be added to retained earnings is given by:

\[
\text{Change in Retained Earnings} = \text{Net Income} - \text{Dividends}
\]

where the dividends are those paid to both the common and preferred stockholders. The formula for retained earnings will require that we reference forecasted 2012 net income from the income statement and the dividends from the statement of cash flows (see Exhibit 2-7, page 59). Note that we are assuming that 2012 dividends will be the same as the 2011 dividends. We can reference these cells in exactly the same way as before, so the formula is:

\[
\text{Change in Retained Earnings} = \text{C21} + \text{'Pro Forma Income Statement'!B15} + \text{'Statement of Cash Flows'!B19}
\]

The results should show that we are forecasting retained earnings to be $297.04 in 2012.

At this point, you should go back and calculate the subtotals in B17, B19, and B22. Finally, we calculate the total liabilities and owner’s equity in B23 with \( =B19+B22 \).

**Discretionary Financing Needed**

Sharp-eyed readers will notice that our pro forma balance sheet does not balance. Although this appears to be a serious problem, it actually represents one of the purposes of the pro forma balance sheet. The difference between total assets and total liabilities and owner’s equity is referred to as discretionary financing needed (DFN, also called additional funds needed or required new funds). In other words, this is the amount of discretionary financing that the firm thinks it will need to raise in the next year. Because of the amount of time and effort required to raise these funds, it is important that the firm be aware of its needs well in advance. The pro forma balance sheet fills this need. Frequently, the firm will find that it is forecasting a higher level of assets than liabilities and equity. In this case, the managers would need to arrange for more liabilities and/or equity to finance the level of assets needed to support the volume of sales expected. This is referred to as a deficit of discretionary funds. If the forecast shows that there will be a higher level of liabilities and equity than assets, the firm is said to have a surplus of discretionary funds. Remember that, in the end, the balance sheet must balance. The “plug figure” necessary to make this happen is the DFN.

We should add an extra line at the bottom of the pro forma balance sheet to calculate the DFN. Type Discretionary Financing Needed in A25, and in B25 add the formula \( =B12-B23 \). This calculation tells us that EPI expects to need $38,119.50 (displayed as 38.12 with the custom number format) more in discretionary funds to support its forecasted level of assets. In this case, EPI is forecasting a deficit of discretionary funds. Apply the custom number format to this number and to the rest of the balance sheet.
CHAPTER 5: Financial Forecasting

EXHIBIT 5-3
EPI’S PRO FORMA BALANCE SHEET FOR 2012

To make clear that this amount is a deficit (note that the sign is the opposite of what might be expected when using that word), we can have Excel inform us whether we will have a surplus or deficit of discretionary funds. Use an IF statement and realize that if the DFN is a positive number, then we have a deficit; otherwise we have a surplus or DFN is zero. So the formula in C25 is: =IF(B25>0,"Deficit", IF(B25<0,"Surplus", "Balanced")). Your balance sheet should now resemble that in Exhibit 5-3.

5. You could also design a custom number format. One possible format is: #,###.00,"Deficit";#,###.00,"Surplus". The benefits of this approach are that you don’t need to use a separate cell and you don’t need to enter a formula.
**Using Iteration to Eliminate DFN**

Circular errors result when a formula refers back to itself, either directly or indirectly through another formula. A simple example would be if the formula in B18 was =B18. Excel cannot calculated this because the result depends on itself (it is self-referential). In most cases, this is undesirable even if the formula eventually converges to a solution. However, there are circumstances that are necessarily self-referential and cannot be solved in any other way.

For example, if we wish to eliminate the DFN deficit, then the firm must raise that amount of money. Suppose that any discretionary funds will be raised with long-term debt. Simply adding $38.12 to the long-term debt in B18 will not quite solve the problem because that will lead to other changes. Specifically, additional long-term debt will increase interest expense and result in lower net income. In turn, this will reduce retained earnings and still leave us with a (smaller) deficit of funding. This new DFN can then be added to long-term debt again, setting off the same chain of calculations. We repeat this cycle as many times as necessary until DFN is equal to zero (or within some allowable tolerance).

By default, Excel will not allow such calculations because the result may not converge. This would lead to an infinite loop of calculations that would tie up your computer in an endless series of calculations. However, if we know that the result will converge (as it will in this case) we can enable these kinds of self-referential, or iterative, calculations. To do so, click Options in the File tab and then go to Formulas. Check the Enable iterative calculation option. Note that we can set the maximum number of iterations as well as the convergence criteria. The default settings will cause the calculation to stop after 100 iterations or if the change in the result is 0.001 or less. Because we should need only a few iterations, leave these at their default settings.

Before we can eliminate the DFN, we need to make a few changes to the pro forma income statement and balance sheet. On the pro forma income statement, we need to add an interest rate. In A21 add the label: Interest Rate and then type 11.70% into B21. This will allow us to calculate the total interest expense as the amount of debt changes. In B12 we will calculate the interest expense for 2012 with the formula: =B21*('Pro Forma Balance Sheet'!B15+'Pro Forma Balance Sheet'!B18). Note that the interest expense is 11.70% of the sum of short-term notes payable and long-term debt. At this point, the value in B12 should be the same as before (76.00).

On the pro forma balance sheet, we need to add our self-referential formula. Our goal is to have the long-term debt (in B18) increase by the amount of the DFN (in B25). However, we can’t just set the formula in B18 to =B25. If we did, then the long-term debt would be 38.12, which would lead to a bigger DFN. This would then cause the debt to grow and the DFN to shrink, which would then cause the debt to shrink and the DFN to grow. It will never converge and will bounce back and forth forever.
To solve this problem by hand, we would start with the current amount of long-term debt (424.61) and then add the DFN to that. This will increase long-term debt, increase interest expense, lower net income, and reduce retained earnings leading to a lower DFN. We now start over again by adding the new DFN amount to long-term debt and the cycle will repeat. If we do this three or four times, DFN will get very close to zero. It may take 20 or 30 cycles for DFN to converge to exactly zero.6

Fortunately, we don’t have to do this by hand. With the right formula for long-term debt, we can make the amount accumulate over many cycles. In B18 enter the formula: \(=B18+B25\). This formula will take the current amount of long-term debt and add the DFN. This will lead to a chain of calculations that will lead to lower DFN. This amount will then be added to the long-term debt, and so on. Eventually, it will converge so that DFN equals zero and long-term debt is 465.61. Note also that interest expense is 80.80, net income is 90.17, and retained earnings is 294.16. The pro forma balance sheet should now look similar to the one in Exhibit 5-4 on page 153, except that we have a couple of important modifications to make.

This whole process will occur very rapidly, and you may not even see the changes taking place. It will be instructive to step through the process one iteration at a time. To do this, go to the Formulas tab in Options and set the Maximum Iterations to 1 (the default is 100). Now, re-enter the formula in B18 (you must do this to reset the calculation). You should see that long-term debt is now 0.00, and DFN is 465.61. To step through the calculation, simply press the F9 key. This will cause the workbook to recalculate one cycle of the iterative formula. Long-term debt will now be 465.61 and DFN will be −32.69. Press F9 again to repeat the calculation and you will see how the numbers change. Keep pressing F9 until DFN goes to zero. Make sure to go back and reset the maximum number of iterations to 100 or more before continuing.

Let’s now improve our iterative calculations a bit. It is very helpful to have the capability to enable or disable the iterative calculations. This can be done as discussed above, but that is tedious. Instead, we can use a cell value (0 or 1) combined with IF statements to do the job. In A31, enter: Iteration and in B31 enter: 0. This will disable iteration, while a 1 will enable iteration. Now, in B18 change the formula for long-term debt so that it is:  
\(=IF(B31=1,B18+B25,C18)\). If iteration is turned on (B31 = 1) then the formula will be the same as before. If iteration is off then long-term debt will be the same as it was in 2011. It can also be helpful to have a note appear when iteration is on. So, in C31 enter the formula:  
\(=IF(B31=1,"Iteration is ON","")\).

6. You are strongly urged to try doing this by hand. This exercise will greatly improve your understanding of the process.
One final change is necessary. We would like to know exactly how much new financing is required. It should be clear that the original $38.12 is not the correct answer because each time that we iterate we add more long-term debt. So, we need a cell to calculate the accumulated DFN. Select row 26 and insert a row. Now, in A26 enter the label: Total Accumulated DFN, and in B26 enter the formula: =IF(B32=1,B26+B25,B25). If iteration is on, this formula will keep track of the additions to DFN. Otherwise, it will be equal to the DFN without iteration. Experiment by changing B32 to 0 and back to 1 to see the effect of these changes.

**EXHIBIT 5-4**
**THE PRO FORMA BALANCE SHEET AFTER ITERATION**

This worksheet could be further refined in several ways. As one example, instead of raising all of the DFN using long-term debt, we could allocate some of it to new equity. In this case, we might use the long-term debt ratio to determine how much should be long-term debt. The
balance would be allocated to equity. Note that additional equity would result in more dividends, which would complicate the situation a bit.

Using circular references should be the last resort. They should be used only when absolutely necessary, as in this case. If your calculation does not converge to a single value, then Excel will eventually stop trying to calculate it and you will have wrong answers. Furthermore, this technique is quite calculation intensive and will cause recalculation of a large spreadsheet to slow to a crawl. If at all possible, you should try to find another method of solving the problem that doesn’t involve circular references.  

Other Forecasting Methods

The primary advantage of the percent of sales forecasting method is its simplicity. There are many other more sophisticated forecasting techniques that can be implemented in a spreadsheet program. In the rest of this chapter we will look at techniques based on linear regression analysis.

Linear Trend Extrapolation

Suppose that you were asked to perform the percent of sales forecast for EPI. The first step in that analysis requires a sales forecast. Because EPI is a small company, nobody regularly makes such forecasts and you will have to generate your own. Where do you start?

<table>
<thead>
<tr>
<th>Year</th>
<th>Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>1,890,532</td>
</tr>
<tr>
<td>2008</td>
<td>2,098,490</td>
</tr>
<tr>
<td>2009</td>
<td>2,350,308</td>
</tr>
<tr>
<td>2010</td>
<td>3,432,000</td>
</tr>
<tr>
<td>2011</td>
<td>3,850,000</td>
</tr>
</tbody>
</table>

Your first idea might be to see if there has been a clear trend in sales over the past several years and to extrapolate that trend, if it exists, to 2012. To see if there has been a trend, you

---

first gather data on sales for EPI for the past five years. Table 5-1 presents the data that you have gathered. Add a new worksheet to your EPI workbook and rename it “Trend Forecast” so that it can be easily identified. Enter the data from Table 5-1 into your worksheet beginning in A1.

The easiest way to see if there has been a trend in sales is to create a chart that plots the sales data versus the years. Select A1:B6, and then insert a scatter chart by clicking the Insert tab and choosing “Scatter with Straight Lines and Markers” as the chart type. Once the chart is created, go to the Layout tab and insert the chart title: EPI Sales for 2007 to 2011. Your worksheet should resemble that in Exhibit 5-5.

Examining the chart leads to the conclusion that sales have definitely been increasing over the past five years, but not at a constant rate. There are several ways to generate a forecast from this data, even though the sales are not increasing at a constant rate.

**EXHIBIT 5-5**

**EPI TREND FORECAST WORKSHEET**

One method is to let Excel draw a linear trend line. That is, let Excel fit a straight line to the data and extrapolate that line to 2012 (or beyond). The line generated is in the form of:

\[ Y = mX + b \]

which you should recognize as the same equation used in algebra courses to describe a straight line. In this equation, \( m \) is the slope and \( b \) is the intercept.

To determine the parameters for this line (\( m \) and \( b \)) Excel uses regression analysis, which we will examine later. To generate a forecast based on the trend, we need to use the **TREND** function which is defined as:

---

8. A line chart would also work. However, because our x-axis labels (years) are numeric, a scatter chart is the better choice.
TREND(KNOWN_Y'S, KNOWN_X'S, NEW_X'S, CONST)

In the TREND function definition, KNOWN_Y'S is the range of the data that we wish to forecast (the dependent variable) and KNOWN_X'S is the optional range of data (the independent variables) that we want to use to determine the trend in the dependent variable. Because the TREND function is generally used to forecast a time-based trend, KNOWN_X'S will usually be a range of years, though it can be any set of consecutive numbers (e.g., 1, 2, 3, ...). NEW_X'S is a continuation of the KNOWN_X'S for which we don't yet know the value of the dependent variable. CONST is a boolean (True/False) variable that tells Excel whether or not to include an intercept in its calculations (generally this should be set to true or omitted).

To generate a forecast for 2012, first enter 2012 into A7. This will provide us with the NEW_X'S value that we will use to forecast 2012 sales. Next, enter the TREND function as: =TREND(B$2:B$6,A$2:A$6,A7,TRUE) into B7. The result is a sales forecast of $4,300,000, which is the same sales forecast that we used in the percent of sales forecasting method for the financial statements.

We can extend our forecast to 2013 and 2014 quite easily. To do this, first enter 2013 into A8 and 2014 into A9. Now copy the formula from B7 to B8:B9. You should see that the forecasted sales for 2013 and 2014 are $4,825,244 and $5,350,489, respectively.

Adding Trend Lines to Charts

An interesting feature of charts in Excel is that we can tell Excel to add a trend line to the chart. Adding this line requires no more work than making a menu choice; we do not have to calculate the data ourselves. To add a trend line to our chart, select the data series in the chart and click on it with the right mouse button. Click Add Trendline and then click on the Close button on the resulting dialog box to see the default linear trend line. You can also show trend lines that aren’t linear. For example, if sales had been increasing at an increasing rate, you might want to fit an exponential trend instead of a linear one. Excel also offers five other trend lines that it can calculate, including a moving average of user-determined length.

Excel can even do a forecast automatically in the chart! (Note that you will not get the actual numerical forecast using this method.) First, delete the trend line that we added by selecting it and then pressing the Delete key on your keyboard, or right-click the trend line and choose Delete from the shortcut menu. Now, select the original line again and insert a linear trend line. Before clicking the Close button, look at the Forecast section and set Forward to 1 unit.

---

9. While KNOWN_X'S and NEW_X'S are technically optional arguments, they should not be omitted in most cases. If both are omitted, then TREND returns an array of values on the trend line instead of the next forecasted value. In this case, if you neglect to array enter (Ctrl+Shift+Enter) and select multiple cells then the result will be the first value on the trend line.
After clicking on the Close button, you will see a trend line that extends to 2012. We could also extend the forecast to 2013 or 2014 by setting Forward to 2 or 3.

Note that you don’t have to first delete the trend line before showing the forecast. Instead, you could right-click the existing trend line, choose **Format Trendline**, and enter the forecast period as before.

Recall that we said that Excel generates the equation for the trend line and uses this equation to make the forecast. We can have Excel show this equation on the chart by selecting the appropriate options. Right-click on the trend line and choose **Format Trendline** from the shortcut menu. Near the bottom of the dialog box click on **Display** **Equation** **on** **Chart**. Click on the Close button and you should see the equation appear on the chart.

The equation that Excel displays, using scientific notation, is:

\[ y = 525245x - 1E+09 \]

which is Excel’s way of saying:

\[ y = 525,245x - 1,000,000,000 \]

However, you should be suspicious of rounding problems any time you see scientific notation. In some cases the rounding isn’t important, but in this case it is. We can fix the problem by right-clicking on the equation and choosing **Format Trendline** **Label** from the shortcut menu. Apply another format and you should now see that the equation is:

\[ y = 525,244.60x - 1,051,441,646.20 \]

---

**EXHIBIT 5-6**

**EPI TREND FORECAST WORKSHEET**

We can see that this equation does indeed generate the forecast for 2012 by substituting 2012 for \( x \) in the equation. At this point, your worksheet should look like the one in Exhibit 5-6.
Regression Analysis

The term regression analysis (also known as ordinary least squares or OLS regression) is a sophisticated-sounding term for a rather simple concept: fitting the best line to a data set. As simple as it sounds, the mathematics behind regression analysis is beyond the scope of this chapter. Excel can easily handle quite complex regression models, and we will make use of Excel’s regression tools without delving too deeply into the underlying mathematics.

As we’ve noted, regression analysis is a technique for fitting the best line to a data set: a very powerful tool for determining the relationship between variables and for forecasting. You could simply plot the data and draw a line that appears to best fit the data, but there is no guarantee that the line you draw is actually the best line. In regression analysis, the best line is defined as the one that minimizes the sum of the squared errors (SSE). The errors are the difference between the actual data points and those predicted by the model.

In our previous example, we used regression analysis (disguised within the TREND function) to forecast EPI’s level of sales for 2012. Aside from forecasting, the second major use of regression analysis is to understand the relationship between variables. In this section, we will use Excel’s regression tool to perform a regression analysis.10

Consider the following example in which we will make use of regression analysis to try to get an alternative forecast of next year’s cost of goods sold for EPI. Table 5-2 provides the historical data for sales and cost of goods sold. Note that the data for the variables must be in columns when using the regression tool. Excel will misinterpret the data if it is in rows.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sales</th>
<th>Cost of Goods</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>$1,890,532</td>
<td>$1,570,200</td>
</tr>
<tr>
<td>2008</td>
<td>2,098,490</td>
<td>1,695,694</td>
</tr>
<tr>
<td>2009</td>
<td>2,350,308</td>
<td>1,992,400</td>
</tr>
<tr>
<td>2010</td>
<td>3,432,000</td>
<td>2,864,000</td>
</tr>
<tr>
<td>2011</td>
<td>3,850,000</td>
<td>3,250,000</td>
</tr>
</tbody>
</table>

Recall that we previously calculated the average percentage of sales for 2010 and 2011 and used that average to generate our forecast for 2012. Suppose, however, that you are 10. The regression tool is not a built-in function in the same sense as TREND. Instead, it is a part of the data analysis tools included with Excel. There is a regression function, LINEST. However, this function is more complex to use because it returns an array of values instead of a single value. Furthermore, the return values are not labeled. See the online help for more information.
Other Forecasting Methods

concerned that there may possibly be a more systematic relationship between sales and cost of goods sold. For example, it is entirely possible that as sales rise, cost of goods sold will rise at a slower rate. This may be due to efficiencies in the production process, quantity discounts on materials, and so on. Alternatively, there may be another relationship, or none at all. Regression analysis can help us to gain a better understanding of the historical relationship and, hopefully, generate better forecasts of the future cost of goods sold.

Before running the regression, let’s create a chart of the data to help get a visual picture of the historical relationship. Enter the data from Table 5-2 into a new worksheet beginning in cell A1. Now select B1:C6 and create a scatter chart of the data. To facilitate our visualization, change the scale on each axis as follows: Right-click the y-axis and choose Format Axis. On the Axis Options tab, change the Minimum to 1,000,000, the Maximum to 4,000,000, and the Major unit to 1,000,000. Repeat those settings for the x-axis. This will ensure that the scale of each axis is the same, which makes it much easier to see the relationship between our two variables.

**FIGURE 5-1**
**CHART OF COST OF GOODS SOLD VS. SALES**

The chart in Figure 5-1 shows what appears to be a pretty consistent relationship. Furthermore, the slope of the line is something less than 45 degrees so we know that a change in sales of $1 will lead to a change of less than $1 in cost of goods sold (as we would expect). We can’t know the exact relationship from reading the chart, but we can run a regression analysis on the data to find the exact slope and intercept of best-fitting line for this data.

Excel provides several functions to calculate the parameters of a regression equation. For example, the **INTERCEPT**, **SLOPE**, and **LINEST** functions all return the parameters of a regression line, while the **TREND** and **FORECAST** functions use linear regression to generate forecasts. There are also functions for nonlinear regression (e.g., **GROWTH** and **LOGEST**). However, Excel also includes another method that we will cover here: the regression tool in the Analysis ToolPak add-in. This tool works very much like any statistical program that you may have used. It will ask for the data and then output a table of the regression results,
including diagnostic data that is used to determine whether the relationship between the variables is statistically significant.

Make sure that the Analysis ToolPak add-in is installed and enabled on your PC. Click the File tab and go to Options, and then click Add-Ins. Look for Analysis ToolPak under “Active Application Add-ins.” If it is listed, then the add-in is ready to use. If it isn’t, then check to see if it is listed under “Inactive Application Add-ins.” If so, then you will need to enable the add-in by clicking the Go button and then placing a check mark next to the add-in name. If you don’t see the add-in listed in either location, then you will need to do a custom install from the Office 2010 installation media.

To run the regression tool, click the Data Analysis button on the Data tab. Next, select Regression from the list of analysis tools that are available. Figure 5-2 shows the dialog box with the data ranges and other options already entered.

Before running the analysis, we need to determine the theoretical relationship between the variables of interest. In this case, we are hypothesizing that the level of sales can be used to predict the cost of goods sold. Therefore, we say that the cost of goods sold is dependent on sales. So the cost of goods sold is referred to as the dependent (Y) variable, and sales is the independent (X) variable. Our mathematical model is:

11. Many regression models have more than one X variable. These models are known as multiple regressions and Excel can handle them just as easily as our bivariate regression. The only restriction is that your X variables must be in a single contiguous range.
Other Forecasting Methods

\[
\text{Cost of Goods Sold}_i = \alpha + \beta(\text{Sales}_i) + \tilde{e}_i
\]  \hspace{1cm} (5-1)

where \(\alpha\) is the intercept, \(\beta\) is the slope, and \(\tilde{e}\) is the random error term in period \(t\).

There are many options on this dialog box, but for our simple problem we are only concerned with four of them. First, we need to tell Excel where the dependent (Y) variable data are located. In the “Input Y Range” edit box enter \$C$1:$C$6, or select this range with the mouse. In the “Input X Range” edit box enter \$B$1:$B$6. Because we have included the labels in our input ranges, we must make sure to check the Labels box. Finally, we want to tell Excel to create a new worksheet within the current workbook for the output. Click on the box to the left of “New Worksheet Ply:” in the Output section, and type Regression Results in the edit box to give a name to the new worksheet.

After clicking the OK button, Excel will calculate the regression statistics and create a new worksheet named “Regression Results.” We could also have Excel enter the output in the same worksheet by specifying the Output Range. Note that you only need to specify the upper left corner of the area where you want the output. (Beware that Excel has a minor bug. When you click on the radio button for the Output Range, the cursor will return to the edit box for the Y range. Before selecting your output range, you must click in the proper edit box, otherwise you will overwrite your Y range. This bug has existed in the past several versions of Excel.)

**EXHIBIT 5-7**

**REGRESSION RESULTS**

![Regression Results](image-url)
Exhibit 5-7 shows the output of the regression tool (it has been reformatted to make it a bit easier to read). The output may appear to be complex if you are not familiar with regression analysis. However, we are primarily concerned with the output, which gives the parameters of the regression line. In cells B17:B18 are the parameters of the regression equation. If we substitute these numbers into equation (5-1) we find:

\[
\text{Cost of Goods Sold}_t = -63,680.82 + 0.8583(Sales_t) + \tilde{e}_t
\]

The equation tells us that, all other things being equal, each $1 increase in sales will lead to an $0.8583 increase in cost of goods sold.

**Statistical Significance**

Before we use this equation to make our forecast, we should make sure that there is a statistically significant relationship between the variables. If the relationship is not significant, then any forecast would be of dubious quality. Furthermore, in a multiple regression it is possible that some \(X\) variables are significant while others are not.

We will begin by looking at the \(R^2\) in cell B5. The \(R^2\) is the coefficient of determination and tells us the proportion of the total variation in the dependent variable that is explained by the independent variable(s). In this case, changes in sales are able to explain nearly 100% of the variability in the cost of goods sold. That is a stronger relationship than you will normally find, but it indicates that this equation is likely to work very well, as long as we have a good forecast of next years’ sales.

It is important to understand that \(R^2\) does not indicate statistical significance. Indeed, it can be increased by simply adding an additional independent variable; even a random variable. This problem can be avoided by using the adjusted \(R^2\), which modifies the original \(R^2\) to account for the number of independent variables. The adjusted \(R^2\) will only increase if the additional variables actually improve the predictive abilities of the model.

To judge the statistical significance of the individual \(X\) variables, we look at the \(t\)-statistics for our regression coefficients (D18; normally we aren’t too concerned with the significance of the intercept). Usually we want to know whether a coefficient is statistically distinguishable from zero (i.e., “statistically significant”). Note that the magnitude of the coefficient is not the issue. If the coefficient for sales is significantly different from zero, then we know that sales is useful in predicting cost of goods sold. The \(t\)-statistic tells us how many standard errors the coefficient is

---

12. We are not trying to minimize the importance of this other output. On the contrary, it would be foolish to attempt to use regression methods for any important purpose without understanding the model completely. We are merely trying to illustrate how Excel can be used for this type of analysis as simply as possible.
away from zero. The higher this number, the more confidence we have that the coefficient is different from zero. In this case, the $t$-statistic is 41.81. A general rule of thumb is that, for large samples, a $t$-statistic greater than about 2.00 is significant at the 95% confidence level or more. Even though we don’t have a large sample, we can be quite sure that the coefficient for sales is significant. Note that we can also use the $p$-value (E18) to determine the exact confidence level. Simply subtract the $p$-value from 1 to find the confidence level. Here, the $p$-value is 0.00003, so we are essentially 100% (actually, 99.997%) confident that our coefficient is significant.

In a multiple regression analysis we can judge the significance of the entire model, as opposed to individual variables, by looking at the $F$ statistic. A high $F$ statistic indicates that the model is significant. To judge the $F$ statistic without consulting statistical tables, Excel provides the Significance $F$ in F12. As with the $p$-value, discussed above, the closer this value is to 0 the better the model. Generally, we look for Significance $F$ to be less than 0.05. In the case of a single $X$ variable, the $F$ statistic provides the same information as the $t$-statistic. Like the adjusted $R^2$, the $F$ statistic will only increase if additional variables add value to the model.

We are very confident that the coefficient for sales is not zero, but we don’t know for sure if the correct value is 0.8583. That number is simply the best point estimate given our set of sample data. Note that in F18:G18 we have numbers labeled “Lower 95%” and “Upper 95%.” This gives us a range of values between which we can be 95% sure the true value of this coefficient lies. In other words, we can be 95% confident that the true change in cost of goods sold per dollar change in sales is between $0.7929 and $0.9236. Of course, there is a small chance (5%) that the true value lies outside of this range.\footnote{Again, we are using quite a small sample with only five observations. This reduces our confidence somewhat and widens the 95% confidence interval. It would be preferable to use higher frequency data such as quarterly sales and cost of goods sold.}

As an aside, note that the 95% confidence range for the intercept contains 0. This indicates that we cannot statistically distinguish the intercept coefficient from zero. This is also confirmed by the rather high $p$-value, and low $t$-statistic, for the intercept. However, because we are merely using this model for forecasting, the significance of the intercept is not important.

We are now quite confident that our model is useful for forecasting cost of goods sold. To make a forecast for the 2012 cost of goods sold, we merely plug our 2012 sales forecast into the equation:

$$\text{Cost of Goods Sold}_{2012} = -63,680.82 + 0.8583(4,300,000) = 3,626,854.68$$
Recall that using the percent of sales method our forecast for 2012 cost of goods sold was $3,609,107.56. Our regression result agrees fairly closely with this number, so either number is probably usable for a forecast. However, note that both of these methods depend critically on our sales forecast. Without a good forecast of sales, all of our other forecasts are questionable.

To generate this forecast yourself, return to your worksheet with the data from Table 5-2. In A7 enter: 2012 for the year and in B7 enter the sales forecast of 4,300,000. Now, calculate the forecast by using the regression output. The equation in C7 is: =$'Regression Results'!B17+$'Regression Results'!B18*B7.

As we did with the TREND function, we can replicate this regression directly in the XY chart that was completed earlier. Simply right-click on one of the data points and choose Add Trendline. Now, place the equation on the chart and have the trend line extended to forecast one period ahead. Your worksheet should now look like the one in Exhibit 5-8.
Summary

In this chapter, we have examined three methods of forecasting financial statements and variables. We used the percent of sales technique to forecast the firm’s income statement and balance sheet based upon an estimated level of sales. We used a time-trend technique to forecast sales as an input to the percent of sales method. Finally, we looked at regression analysis to help generate a better forecast of the cost of goods sold by using the relationship between that and sales over the past five years.

We have barely scratched the surface of forecasting methodologies. However, we hope that this chapter has stimulated an interest in this important subject. If so, be assured that Excel, either alone or through an add-in program, can be made to handle nearly all of your forecasting problems. Please remember that any forecast is almost assuredly wrong. We can only hope to get reasonably close to the actual future outcome. How close you get depends upon the quality of your model and the inputs to that model.

Table 5-3
Functions Introduced in This Chapter

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Function</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast future outcomes based on a time trend</td>
<td>TREND(KNOWN_Y'S, KNOWN_X'S, NEW_X'S, CONST)</td>
<td>156</td>
</tr>
</tbody>
</table>

Problems

1. Using the data in the student spreadsheet file P&G.xlsx (to find the student spreadsheets for Financial Analysis with Microsoft Excel, sixth edition, go to www.cengage.com/finance/mayes) forecast the June 30, 2011, income statement and balance sheet for Procter & Gamble. Use the percent of sales method and the following assumptions: (1) Sales in FY 2011 will be $81,000; (2) The tax rate will be 27.26%; (3) Each item that changes with sales will be the five-year average percentage of sales; (4) The preferred dividend will be 219; and (5) The common dividend payout ratio will be 42% of income available to common stockholders.

   a. What is the discretionary financing needed in 2011? Is this a surplus or deficit?

   b. Assume that the DFN will be absorbed by long-term debt and that the total interest rate is 4.50% of LTD. Set up an iterative worksheet to eliminate it.
c. Create a chart of cash vs. sales and add a linear trend line. Is the cash balance a consistent percentage of sales? Does the relationship fit your expectations?

d. Use the regression tool to verify your results from part c. Is the trend statistically significant? Use at least three methods to show why or why not.

e. Turn off iteration, and use the Scenario Manager to set up three scenarios:
   1) Best Case — Sales are 5% higher than expected.
   2) Base Case — Sales are exactly as expected.
   3) Worst Case — Sales are 5% less than expected.
   What is the DFN under each scenario?

2. Use the same data as in Problem 1.

   a. Recalculate the percentage of sales income statement, but this time use the \text{TREND} function to forecast other income and interest expense.

   b. Recalculate the percentage of sales balance sheet, but this time use the \text{TREND} function to forecast cash, gross property plant and equipment, gross intangibles, and other long-term assets.

   c. Do these new values appear to be more realistic than the original values? Does this technique make sense for each of these items? Might other income statement or balance sheet items be forecasted in this way?


   a. Create a scatter plot to show the relationship between the returns on Cymer and the S&P 500. Describe, in words, the relationship between the returns of Cymer and the S&P 500. Estimate the slope of a regression equation of this data. Repeat for Contrafund.
b. Add a linear trend line to the chart, and place the equation and $R^2$ on the chart. Does this equation confirm your guess from part a? How much of the variability in Cymer returns can be explained by variability in the broad market? Repeat for Contrafund.

c. Using the Analysis ToolPak add-in, run a regression analysis on this data. Your dependent variable is the Cymer returns, and the independent variable is the S&P 500 returns. Does this confirm the earlier results? The slope coefficient is Cymer’s beta. Is the beta of this stock statistically significant? Explain.

d. Repeat part c using the returns on Contrafund and the S&P 500. Compare the $R^2$ from both regressions. What conclusions can you draw from the difference?

**Internet Exercises**

1. Because you are reading this after the end of Procter & Gamble’s fiscal year 2011, how do your forecasts from the previous problems compare to the actual FY 2011 results? Does it appear that more information would have helped to generate better forecasts? Insert Procter & Gamble’s actual sales for 2011 into your forecast. Does this improve your forecast of earnings?

2. Choose your own company and repeat Problem 3. The data can be easily obtained from Yahoo! Finance (http://finance.yahoo.com). Enter a ticker symbol and get a stock price quote. On the left side of the page click the link for “Historical Prices.” Set the dates for a five-year period and the frequency to monthly. Click the link at the bottom of the page to load the data into Excel. Now, repeat the steps using the ticker symbol SPY (an exchange traded fund that mimics the S&P 500). Now, combine the monthly closing prices onto one worksheet and calculate the monthly percentage changes. You should now have the data necessary to answer the questions from Problem 3. Note that to improve your results, you can also get the dividends and calculate the monthly total returns.
This page intentionally left blank