CHAPTER 18

Time Series Analysis and Forecasting

CONTENTS

STATISTICS *in* PRACTICE

NEVADA OCCUPATIONAL HEALTH CLINIC*
SPARKS, NEVADA

Nevada Occupational Health Clinic is a privately owned medical clinic in Sparks, Nevada. The clinic specializes in industrial medicine. Operating at the same site for more than 20 years, the clinic had been in a rapid growth phase. Monthly billings increased from $57,000 to more than $300,000 in 26 months, when the main clinic building burned to the ground.

The clinic's insurance policy covered physical property and equipment as well as loss of income due to the interruption of regular business operations. Settling the property insurance claim was a relatively straightforward matter of determining the value of the physical property and equipment lost during the fire. However, determining the value of the income lost during the seven months that it took to rebuild the clinic was a complicated matter involving negotiations between the business owners and the insurance company. No preestablished rules could help calculate "what would have happened" to the clinic's billings if the fire had not occurred. To estimate the lost income, the clinic used a forecasting method to project the growth in business that would have been realized during the seven-month lost-business period. The actual history of billings prior to the fire provided the basis for a forecasting model with linear trend

A physician checks a patient's blood pressure at the Nevada Occupational Health Clinic. © Bob Pardue–Medical Lifestyle/Alamy.

and seasonal components as discussed in this chapter. This forecasting model enabled the clinic to establish an accurate estimate of the loss, which was eventually accepted by the insurance company.

*The authors are indebted to Bard Betz, Director of Operations, and Curtis Brauer, Executive Administrative Assistant, Nevada Occupational Health Clinic, for providing this Statistics in Practice.

A forecast is simply a prediction of what will happen in the future. Managers must learn to accept that regardless of the technique used, they will not be able to develop perfect forecasts.

The purpose of this chapter is to provide an introduction to time series analysis and forecasting. Suppose we are asked to provide quarterly forecasts of sales for one of our company's products over the coming one-year period. Production schedules, raw material purchasing, inventory policies, and sales quotas will all be affected by the quarterly forecasts we provide. Consequently, poor forecasts may result in poor planning and increased costs for the company. How should we go about providing the quarterly sales forecasts? Good judgment, intuition, and an awareness of the state of the economy may give us a rough idea or "feeling" of what is likely to happen in the future, but converting that feeling into a number that can be used as next year's sales forecast is difficult.

Forecasting methods can be classified as qualitative or quantitative. Qualitative methods generally involve the use of expert judgment to develop forecasts. Such methods are appropriate when historical data on the variable being forecast are either not applicable or unavailable. Quantitative forecasting methods can be used when (1) past information about the variable being forecast is available, (2) the information can be quantified, and (3) it is

reasonable to assume that the pattern of the past will continue into the future. In such cases, a forecast can be developed using a time series method or a causal method. We will focus exclusively on quantitative forecasting methods in this chapter.

If the historical data are restricted to past values of the variable to be forecast, the forecasting procedure is called a *time series method* and the historical data are referred to as a time series. The objective of time series analysis is to discover a pattern in the historical data or time series and then extrapolate the pattern into the future; the forecast is based solely on past values of the variable and/or on past forecast errors.

Causal forecasting methods are based on the assumption that the variable we are forecasting has a cause-effect relationship with one or more other variables. In the discussion of regression analysis in Chapters 14, 15, and 16, we showed how one or more independent variables could be used to predict the value of a single dependent variable. Looking at regression analysis as a forecasting tool, we can view the time series value that we want to forecast as the dependent variable. Hence, if we can identify a good set of related independent, or explanatory, variables, we may be able to develop an estimated regression equation for predicting or forecasting the time series. For instance, the sales for many products are influenced by advertising expenditures, so regression analysis may be used to develop an equation showing how sales and advertising are related. Once the advertising budget for the next period is determined, we could substitute this value into the equation to develop a prediction or forecast of the sales volume for that period. Note that if a time series method were used to develop the forecast, advertising expenditures would not be considered; that is, a time series method would base the forecast solely on past sales.

By treating time as the independent variable and the time series as a dependent variable, regression analysis can also be used as a time series method. To help differentiate the application of regression analysis in these two cases, we use the terms *cross-sectional regression* and *time series regression*. Thus, time series regression refers to the use of regression analysis when the independent variable is time. Because our focus in this chapter is on time series methods, we leave the discussion of the application of regression analysis as a causal forecasting method to more advanced texts on forecasting.

18.1   Time Series Patterns

TABLE 18.1

GASOLINE SALES TIME SERIES

Week	Sales (1000s of gallons)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22

A **time series** is a sequence of observations on a variable measured at successive points in time or over successive periods of time. The measurements may be taken every hour, day, week, month, or year, or at any other regular interval.[1] The pattern of the data is an important factor in understanding how the time series has behaved in the past. If such behavior can be expected to continue in the future, we can use the past pattern to guide us in selecting an appropriate forecasting method.

To identify the underlying pattern in the data, a useful first step is to construct a **time series plot**. A time series plot is a graphical presentation of the relationship between time and the time series variable; time is on the horizontal axis and the time series values are shown on the vertical axis. Let us review some of the common types of data patterns that can be identified when examining a time series plot.

Horizontal Pattern

A **horizontal pattern** exists when the data fluctuate around a constant mean. To illustrate a time series with a horizontal pattern, consider the 12 weeks of data in Table 18.1. These data

---

[1]We limit our discussion to time series in which the values of the series are recorded at equal intervals. Cases in which the observations are made at unequal intervals are beyond the scope of this text.

FIGURE 18.1   GASOLINE SALES TIME SERIES PLOT
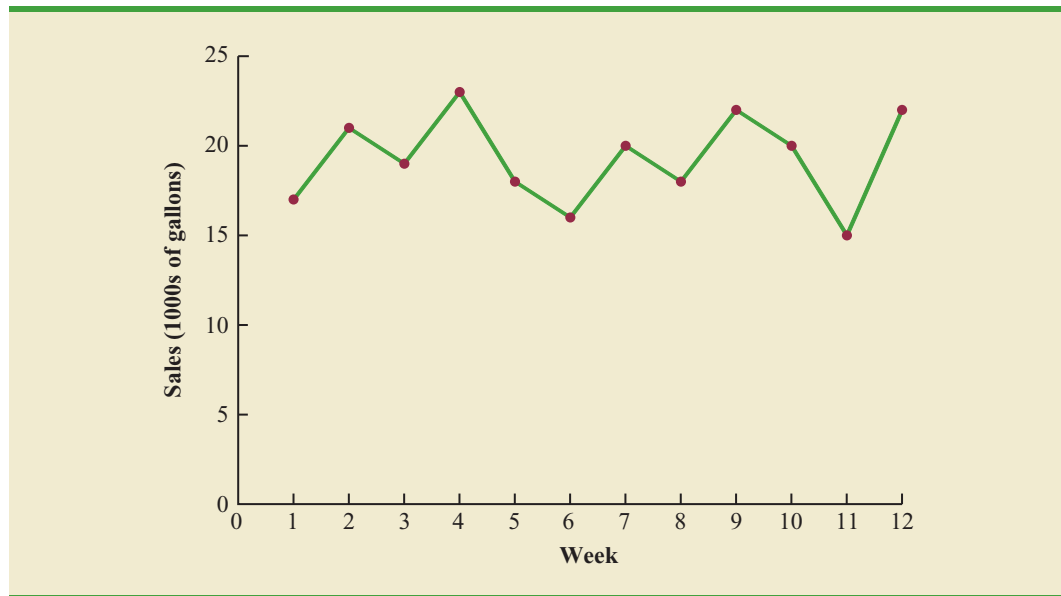
WEB file

GasolineRevised

TABLE 18.2

GASOLINE SALES
TIME SERIES
AFTER OBTAINING
THE CONTRACT
WITH THE
VERMONT
STATE POLICE

Week	Sales (1000s of gallons)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22
13	31
14	34
15	31
16	33
17	28
18	32
19	30
20	29
21	34
22	33

show the number of gallons of gasoline sold by a gasoline distributor in Bennington, Vermont, over the past 12 weeks. The average value or mean for this time series is 19.25 or 19,250 gallons per week. Figure 18.1 shows a time series plot for these data. Note how the data fluctuate around the sample mean of 19,250 gallons. Although random variability is present, we would say that these data follow a horizontal pattern.

The term **stationary time series**[2] is used to denote a time series whose statistical properties are independent of time. In particular this means that

1. The process generating the data has a constant mean.
2. The variability of the time series is constant over time.

A time series plot for a stationary time series will always exhibit a horizontal pattern. But simply observing a horizontal pattern is not sufficient evidence to conclude that the time series is stationary. More advanced texts on forecasting discuss procedures for determining if a time series is stationary and provide methods for transforming a time series that is not stationary into a stationary series.

Changes in business conditions can often result in a time series that has a horizontal pattern shifting to a new level. For instance, suppose the gasoline distributor signs a contract with the Vermont State Police to provide gasoline for state police cars located in southern Vermont. With this new contract, the distributor expects to see a major increase in weekly sales starting in week 13. Table 18.2 shows the number of gallons of gasoline sold for the original time series and for the 10 weeks after signing the new contract. Figure 18.2 shows the corresponding time series plot. Note the increased level of the time series beginning in week 13. This change in the level of the time series makes it more difficult to choose an appropriate forecasting method. Selecting a forecasting method that adapts well to changes in the level of a time series is an important consideration in many practical applications.

_____

[2]For a formal definition of stationary see G. E. P., Box, G. M. Jenkins, and G. C. Reinsell, *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1994, p. 23.

FIGURE 18.2   GASOLINE SALES TIME SERIES PLOT AFTER OBTAINING THE
                   CONTRACT WITH THE VERMONT STATE POLICE
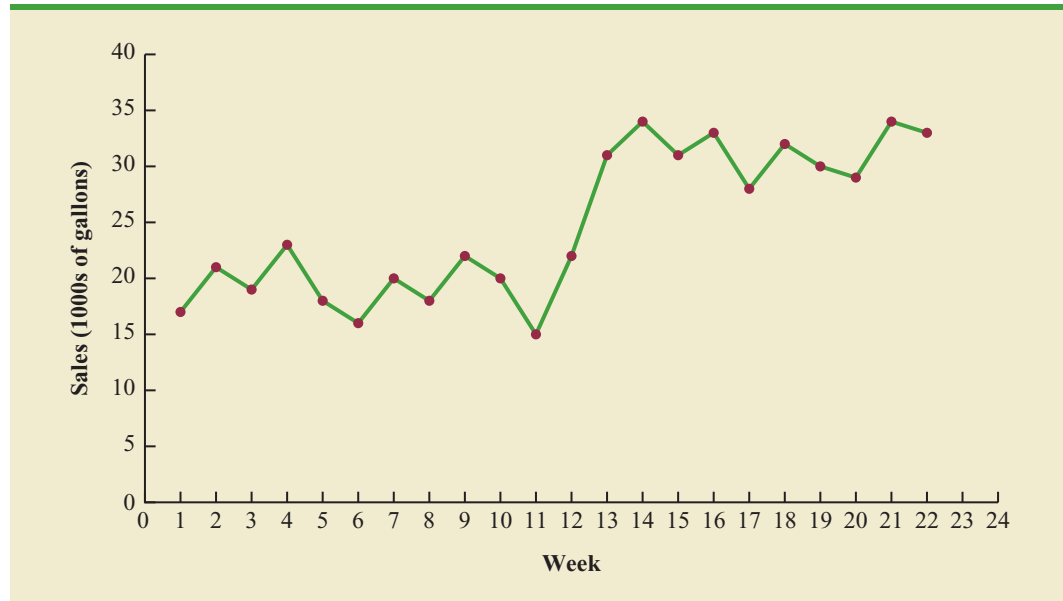
Trend Pattern

WEB file

Bicycle

Although time series data generally exhibit random fluctuations, a time series may also show
gradual shifts or movements to relatively higher or lower values over a longer period of time.
If a time series plot exhibits this type of behavior, we say that a **trend pattern** exists. A trend
is usually the result of long-term factors such as population increases or decreases, changing
demographic characteristics of the population, technology, and/or consumer preferences.

To illustrate a time series with a trend pattern, consider the time series of bicycle sales
for a particular manufacturer over the past 10 years, as shown in Table 18.3 and Figure 18.3.
Note that 21,600 bicycles were sold in year one, 22,900 were sold in year two, and so on.
In year 10, the most recent year, 31,400 bicycles were sold. Visual inspection of the time
series plot shows some up and down movement over the past 10 years, but the time series
also seems to have a systematically increasing or upward trend.

The trend for the bicycle sales time series appears to be linear and increasing over time,
but sometimes a trend can be described better by other types of patterns. For instance, the
data in Table 18.4 and the corresponding time series plot in Figure 18.4 show the sales for
a cholesterol drug since the company won FDA approval for it 10 years ago. The time se-
ries increases in a nonlinear fashion; that is, the rate of change of revenue does not increase
by a constant amount from one year to the next. In fact, the revenue appears to be growing
in an exponential fashion. Exponential relationships such as this are appropriate when the
percentage change from one period to the next is relatively constant.

TABLE 18.3

BICYCLE SALES
TIME SERIES

Year	Sales (1000s)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

Seasonal Pattern

The trend of a time series can be identified by analyzing multiyear movements in histori-
cal data. Seasonal patterns are recognized by seeing the same repeating patterns over suc-
cessive periods of time. For example, a manufacturer of swimming pools expects low sales
activity in the fall and winter months, with peak sales in the spring and summer months.
Manufacturers of snow removal equipment and heavy clothing, however, expect just the

FIGURE 18.3   BICYCLE SALES TIME SERIES PLOT
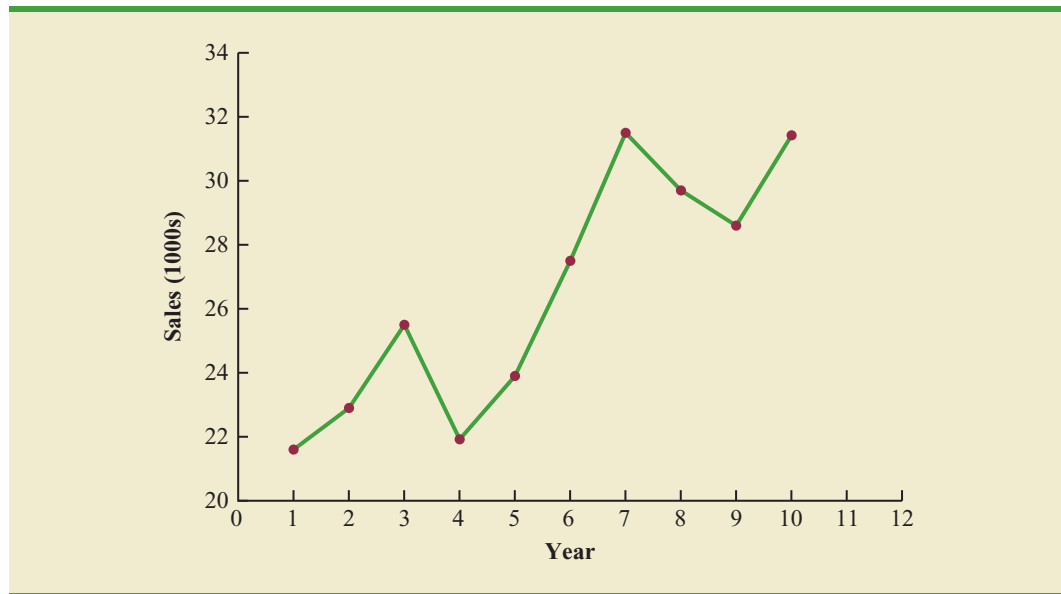
WEB file

Cholesterol

TABLE 18.4

CHOLESTEROL
REVENUE
TIME SERIES
($MILLIONS)

Year	Revenue
1	23.1
2	21.3
3	27.4
4	34.6
5	33.8
6	43.2
7	59.5
8	64.4
9	74.2
10	99.3

opposite yearly pattern. Not surprisingly, the pattern for a time series plot that exhibits a repeating pattern over a one-year period due to seasonal influences is called a **seasonal pattern**. While we generally think of seasonal movement in a time series as occurring within one year, time series data can also exhibit seasonal patterns of less than one year in duration. For example, daily traffic volume shows within-the-day "seasonal" behavior, with peak levels occurring during rush hours, moderate flow during the rest of the day and early evening, and light flow from midnight to early morning.

As an example of a seasonal pattern, consider the number of umbrellas sold at a clothing store over the past five years. Table 18.5 shows the time series and Figure 18.5 shows the corresponding time series plot. The time series plot does not indicate any long-term trend in sales. In fact, unless you look carefully at the data, you might conclude that the data follow a horizontal pattern. But closer inspection of the time series plot reveals a regular pattern in the data. That is, the first and third quarters have moderate sales, the second quarter has the highest sales, and the fourth quarter tends to have the lowest sales volume. Thus, we would conclude that a quarterly seasonal pattern is present.

Trend and Seasonal Pattern

Some time series include a combination of a trend and seasonal pattern. For instance, the data in Table 18.6 and the corresponding time series plot in Figure 18.6 show television set sales for a particular manufacturer over the past four years. Clearly, an increasing trend is present. But, Figure 18.6 also indicates that sales are lowest in the second quarter of each year and increase in quarters 3 and 4. Thus, we conclude that a seasonal pattern also exists for television set sales. In such cases we need to use a forecasting method that has the capability to deal with both trend and seasonality.

Cyclical Pattern

A **cyclical pattern** exists if the time series plot shows an alternating sequence of points below and above the trend line lasting more than one year. Many economic time series exhibit

FIGURE 18.4 CHOLESTEROL REVENUE TIMES SERIES PLOT ($MILLIONS)
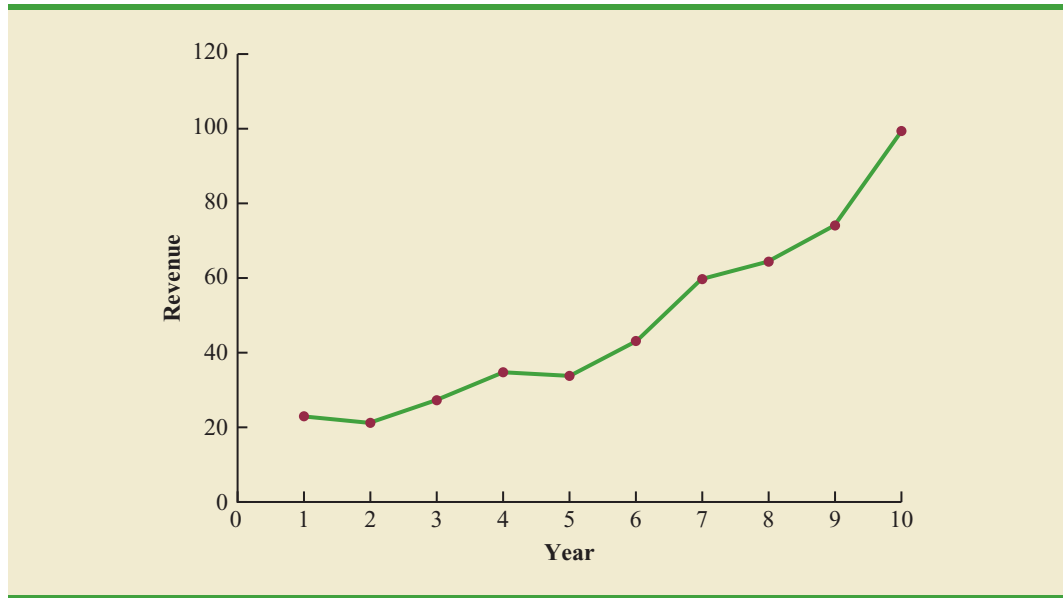
TABLE 18.5 UMBRELLA SALES TIME SERIES

Year	Quarter	Sales
1	1	125
	2	153
	3	106
	4	88
2	1	118
	2	161
	3	133
	4	102
3	1	138
	2	144
	3	113
	4	80
4	1	109
	2	137
	3	125
	4	109
5	1	130
	2	165
	3	128
	4	96

WEB file

Umbrella

cyclical behavior with regular runs of observations below and above the trend line. Often, the cyclical component of a time series is due to multiyear business cycles. For example, periods of moderate inflation followed by periods of rapid inflation can lead to time series that alternate below and above a generally increasing trend line (e.g., a time series for housing costs). Business cycles are extremely difficult, if not impossible, to forecast. As a

FIGURE 18.5   UMBRELLA SALES TIME SERIES PLOT

TABLE 18.6   QUARTERLY TELEVISION SET SALES TIME SERIES

WEB file

TVSales

Year	Quarter	Sales (1000s)
1	1	4.8
	2	4.1
	3	6.0
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6.0
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8.0
	4	8.4

result, cyclical effects are often combined with long-term trend effects and referred to as trend-cycle effects. In this chapter we do not deal with cyclical effects that may be present in the time series.

Selecting a Forecasting Method

The underlying pattern in the time series is an important factor in selecting a forecasting method. Thus, a time series plot should be one of the first things developed when trying to determine what forecasting method to use. If we see a horizontal pattern, then we need to

FIGURE 18.6 QUARTERLY TELEVISION SET SALES TIME SERIES PLOT
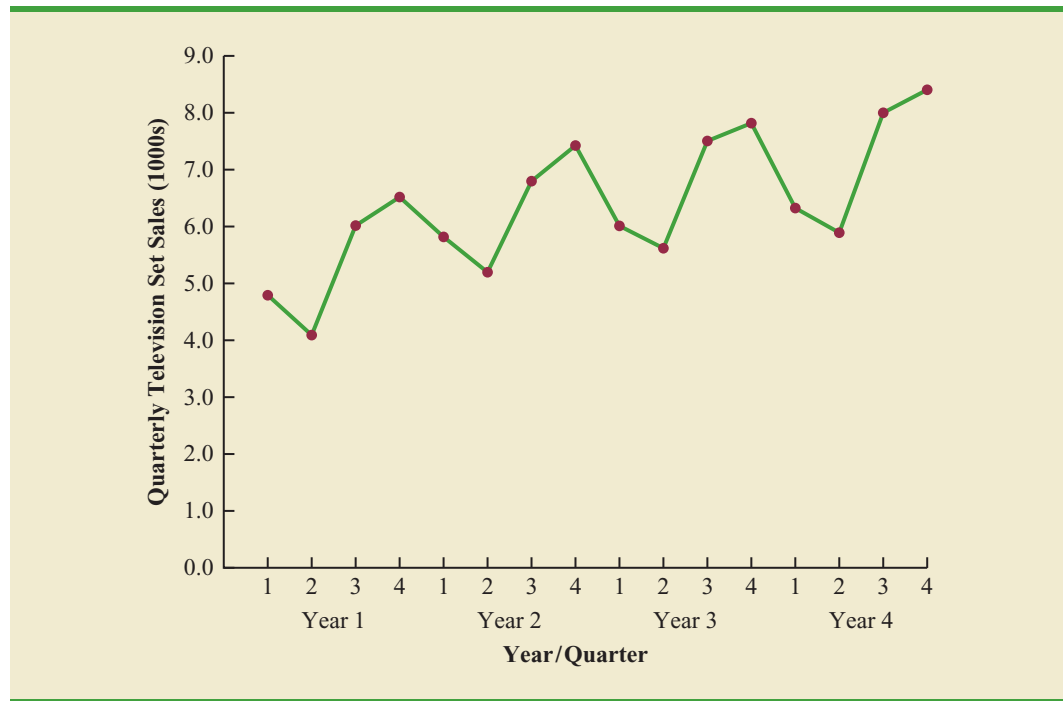
select a method appropriate for this type of pattern. Similarly, if we observe a trend in the data, then we need to use a forecasting method that has the capability to handle trend effectively. The next two sections illustrate methods that can be used in situations where the underlying pattern is horizontal; in other words, no trend or seasonal effects are present. We then consider methods appropriate when trend and/or seasonality are present in the data.

18.2 Forecast Accuracy

In this section we begin by developing forecasts for the gasoline time series shown in Table 18.1 using the simplest of all the forecasting methods: an approach that uses the most recent week's sales volume as the forecast for the next week. For instance, the distributor sold 17 thousand gallons of gasoline in week 1; this value is used as the forecast for week 2. Next, we use 21, the actual value of sales in week 2, as the forecast for week 3, and so on. The forecasts obtained for the historical data using this method are shown in Table 18.7 in the column labeled Forecast. Because of its simplicity, this method is often referred to as a *naive* forecasting method.

How accurate are the forecasts obtained using this *naive* forecasting method? To answer this question we will introduce several measures of forecast accuracy. These measures are used to determine how well a particular forecasting method is able to reproduce the time series data that are already available. By selecting the method that has the best accuracy for the data already known, we hope to increase the likelihood that we will obtain better forecasts for future time periods.

The key concept associated with measuring forecast accuracy is **forecast error**, defined as

$$\text{Forecast Error} = \text{ActualValue} - \text{Forecast}$$

TABLE 18.7 COMPUTING FORECASTS AND MEASURES OF FORECAST ACCURACY USING THE MOST RECENT VALUE AS THE FORECAST FOR THE NEXT PERIOD

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21	17	4	4	16	19.05	19.05
3	19	21	−2	2	4	−10.53	10.53
4	23	19	4	4	16	17.39	17.39
5	18	23	−5	5	25	−27.78	27.78
6	16	18	−2	2	4	−12.50	12.50
7	20	16	4	4	16	20.00	20.00
8	18	20	−2	2	4	−11.11	11.11
9	22	18	4	4	16	18.18	18.18
10	20	22	−2	2	4	−10.00	10.00
11	15	20	−5	5	25	−33.33	33.33
12	22	15	7	7	49	31.82	31.82
		Totals	5	41	179	1.19	211.69

For instance, because the distributor actually sold 21 thousand gallons of gasoline in week 2 and the forecast, using the sales volume in week 1, was 17 thousand gallons, the forecast error in week 2 is

$$\text{Forecast Error in week 2} = 21 - 17 = 4$$

The fact that the forecast error is positive indicates that in week 2 the forecasting method underestimated the actual value of sales. Next, we use 21, the actual value of sales in week 2, as the forecast for week 3. Since the actual value of sales in week 3 is 19, the forecast error for week 3 is $19 - 21 = -2$. In this case, the negative forecast error indicates that in week 3 the forecast overestimated the actual value. Thus, the forecast error may be positive or negative, depending on whether the forecast is too low or too high. A complete summary of the forecast errors for this naive forecasting method is shown in Table 18.7 in the column labeled Forecast Error.

In regression analysis, a residual is defined as the difference between the observed value of the dependent variable and the estimated value. The forecast errors are analogous to the residuals in regression analysis.

A simple measure of forecast accuracy is the mean or average of the forecast errors. Table 18.7 shows that the sum of the forecast errors for the gasoline sales time series is 5; thus, the mean or average forecast error is $5/11 = .45$. Note that although the gasoline time series consists of 12 values, to compute the mean error we divided the sum of the forecast errors by 11 because there are only 11 forecast errors. Because the mean forecast error is positive, the method is underforecasting; in other words, the observed values tend to be greater than the forecasted values. Because positive and negative forecast errors tend to off-set one another, the mean error is likely to be small; thus, the mean error is not a very useful measure of forecast accuracy.

The **mean absolute error**, denoted MAE, is a measure of forecast accuracy that avoids the problem of positive and negative forecast errors offsetting one another. As you might expect given its name, MAE is the average of the absolute values of the forecast errors. Table 18.7 shows that the sum of the absolute values of the forecast errors is 41; thus,

$$\text{MAE} = \text{average of the absolute value of forecast errors} = \frac{41}{11} = 3.73$$

In regression analysis the mean square error (MSE) is the residual sum of squares divided by its degrees of freedom. In forecasting, MSE is the average of the sum of squared forecast errors.

Another measure that avoids the problem of positive and negative forecast errors offsetting each other is obtained by computing the average of the squared forecast errors. This measure of forecast accuracy, referred to as the **mean squared error**, is denoted MSE. From Table 18.7, the sum of the squared errors is 179; hence,

$$\text{MSE} = \text{average of the sum of squared forecast errors} = \frac{179}{11} = 16.27$$

The size of MAE and MSE depends upon the scale of the data. As a result, it is difficult to make comparisons for different time intervals, such as comparing a method of forecasting monthly gasoline sales to a method of forecasting weekly sales, or to make comparisons across different time series. To make comparisons like these we need to work with relative or percentage error measures. The **mean absolute percentage error**, denoted MAPE, is such a measure. To compute MAPE we must first compute the percentage error for each forecast. For example, the percentage error corresponding to the forecast of 17 in week 2 is computed by dividing the forecast error in week 2 by the actual value in week 2 and multiplying the result by 100. For week 2 the percentage error is computed as follows:

$$\text{Percentage error for week 2} = \frac{4}{21}(100) = 19.05\%$$

Thus, the forecast error for week 2 is 19.05% of the observed value in week 2. A complete summary of the percentage errors is shown in Table 18.7 in the column labeled Percentage Error. In the next column, we show the absolute value of the percentage error.

Table 18.7 shows that the sum of the absolute values of the percentage errors is 211.69; thus,

$$\text{MAPE} = \text{average of the absolute value of percentage forecast errors} = \frac{211.69}{11} = 19.24\%$$

Summarizing, using the naive (most recent observation) forecasting method, we obtained the following measures of forecast accuracy:

$$\text{MAE} = 3.73$$
$$\text{MSE} = 16.27$$
$$\text{MAPE} = 19.24\%$$

These measures of forecast accuracy simply measure how well the forecasting method is able to forecast historical values of the time series. Now, suppose we want to forecast sales for a future time period, such as week 13. In this case the forecast for week 13 is 22, the actual value of the time series in week 12. Is this an accurate estimate of sales for week 13? Unfortunately, there is no way to address the issue of accuracy associated with forecasts for future time periods. But, if we select a forecasting method that works well for the historical data, and we think that the historical pattern will continue into the future, we should obtain results that will ultimately be shown to be good.

Before closing this section, let's consider another method for forecasting the gasoline sales time series in Table 18.1. Suppose we use the average of all the historical data available as the forecast for the next period. We begin by developing a forecast for week 2. Since there is only one historical value available prior to week 2, the forecast for week 2 is just the time series value in week 1; thus, the forecast for week 2 is 17 thousand gallons of gasoline. To compute the forecast for week 3, we take the average of the sales values in weeks 1 and 2. Thus,

TABLE 18.8   COMPUTING FORECASTS AND MEASURES OF FORECAST ACCURACY USING THE AVERAGE OF ALL THE HISTORICAL DATA AS THE FORECAST FOR THE NEXT PERIOD

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21	17.00	4.00	4.00	16.00	19.05	19.05
3	19	19.00	0.00	0.00	0.00	0.00	0.00
4	23	19.00	4.00	4.00	16.00	17.39	17.39
5	18	20.00	−2.00	2.00	4.00	−11.11	11.11
6	16	19.60	−3.60	3.60	12.96	−22.50	22.50
7	20	19.00	1.00	1.00	1.00	5.00	5.00
8	18	19.14	−1.14	1.14	1.31	−6.35	6.35
9	22	19.00	3.00	3.00	9.00	13.64	13.64
10	20	19.33	0.67	0.67	0.44	3.33	3.33
11	15	19.40	−4.40	4.40	19.36	−29.33	29.33
12	22	19.00	3.00	3.00	9.00	13.64	13.64
		Totals	4.53	26.81	89.07	2.76	141.34

$$\text{Forecast for week 3} = \frac{17 + 21}{2} = 19$$

Similarly, the forecast for week 4 is

$$\text{Forecast for week 4} = \frac{17 + 21 + 19}{3} = 19$$

The forecasts obtained using this method for the gasoline time series are shown in Table 18.8 in the column labeled Forecast. Using the results shown in Table 18.8, we obtained the following values of MAE, MSE, and MAPE:

$$\text{MAE} = \frac{26.81}{11} = 2.44$$

$$\text{MSE} = \frac{89.07}{11} = 8.10$$

$$\text{MAPE} = \frac{141.34}{11} = 12.85\%$$

We can now compare the accuracy of the two forecasting methods we have considered in this section by comparing the values of MAE, MSE, and MAPE for each method.

	Naive Method	Average of Past Values
MAE	3.73	2.44
MSE	16.27	8.10
MAPE	19.24%	12.85%

For every measure, the average of past values provides more accurate forecasts than using the most recent observation as the forecast for the next period. In general, if the underlying time series is stationary, the average of all the historical data will always provide the best results.

But suppose that the underlying time series is not stationary. In Section 18.1 we mentioned that changes in business conditions can often result in a time series that has a horizontal pattern shifting to a new level. We discussed a situation in which the gasoline distributor signed a contract with the Vermont State Police to provide gasoline for state police cars located in southern Vermont. Table 18.2 shows the number of gallons of gasoline sold for the original time series and the 10 weeks after signing the new contract, and Figure 18.2 shows the corresponding time series plot. Note the change in level in week 13 for the resulting time series. When a shift to a new level like this occurs, it takes a long time for the forecasting method that uses the average of all the historical data to adjust to the new level of the time series. But, in this case, the simple naive method adjusts very rapidly to the change in level because it uses the most recent observation available as the forecast.

Measures of forecast accuracy are important factors in comparing different forecasting methods, but we have to be careful not to rely upon them too heavily. Good judgment and knowledge about business conditions that might affect the forecast also have to be carefully considered when selecting a method. And historical forecast accuracy is not the only consideration, especially if the time series is likely to change in the future.

In the next section we will introduce more sophisticated methods for developing forecasts for a time series that exhibits a horizontal pattern. Using the measures of forecast accuracy developed here, we will be able to determine if such methods provide more accurate forecasts than we obtained using the simple approaches illustrated in this section. The methods that we will introduce also have the advantage of adapting well in situations where the time series changes to a new level. The ability of a forecasting method to adapt quickly to changes in level is an important consideration, especially in short-term forecasting situations.

Exercises

Methods

1. Consider the following time series data.

Week	1	2	3	4	5	6
Value	18	13	16	11	17	14

Using the naive method (most recent value) as the forecast for the next week, compute the following measures of forecast accuracy.
   a. Mean absolute error.
   b. Mean squared error.
   c. Mean absolute percentage error.
   d. What is the forecast for week 7?

2. Refer to the time series data in exercise 1. Using the average of all the historical data as a forecast for the next period, compute the following measures of forecast accuracy.
   a. Mean absolute error.
   b. Mean squared error.
   c. Mean absolute percentage error.
   d. What is the forecast for week 7?

SELF test

3. Exercises 1 and 2 used different forecasting methods. Which method appears to provide the more accurate forecasts for the historical data? Explain.

4. Consider the following time series data.

Month	1	2	3	4	5	6	7
Value	24	13	20	12	19	23	15

a. Compute MSE using the most recent value as the forecast for the next period. What is the forecast for month 8?
b. Compute MSE using the average of all the data available as the forecast for the next period. What is the forecast for month 8?
c. Which method appears to provide the better forecast?

18.3 # Moving Averages and Exponential Smoothing

In this section we discuss three forecasting methods that are appropriate for a time series with a horizontal pattern: moving averages, weighted moving averages, and exponential smoothing. These methods also adapt well to changes in the level of a horizontal pattern such as we saw with the extended gasoline sales time series (Table 18.2 and Figure 18.2). However, without modification they are not appropriate when significant trend, cyclical, or seasonal effects are present. Because the objective of each of these methods is to "smooth out" the random fluctuations in the time series, they are referred to as smoothing methods. These methods are easy to use and generally provide a high level of accuracy for short-range forecasts, such as a forecast for the next time period.

Moving Averages

The **moving averages** method uses the average of the most recent k data values in the time series as the forecast for the next period. Mathematically, a moving average forecast of order k is as follows:

MOVING AVERAGE FORECAST OF ORDER k

$$F_{t+1} = \frac{\sum(\text{most recent } k \text{ data values})}{k} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-k+1}}{k} \quad \textbf{(18.1)}$$

where

$$F_{t+1} = \text{forecast of the times series for period } t + 1$$
$$Y_t = \text{actual value of the time series in period } t$$

The term *moving* is used because every time a new observation becomes available for the time series, it replaces the oldest observation in the equation and a new average is computed. As a result, the average will change, or move, as new observations become available.

To illustrate the moving averages method, let us return to the gasoline sales data in Table 18.1 and Figure 18.1. The time series plot in Figure 18.1 indicates that the gasoline sales time series has a horizontal pattern. Thus, the smoothing methods of this section are applicable.

To use moving averages to forecast a time series, we must first select the order, or number of time series values, to be included in the moving average. If only the most recent values of the time series are considered relevant, a small value of k is preferred. If more past values are considered relevant, then a larger value of k is better. As mentioned earlier, a time series with a horizontal pattern can shift to a new level over time. A moving average will adapt to the new level of the series and resume providing good forecasts in k periods. Thus, a smaller value of k will track shifts in a time series more quickly. But larger values of k will be more effective in smoothing out the random fluctuations over time. So managerial judgment based on an understanding of the behavior of a time series is helpful in choosing a good value for k.

To illustrate how moving averages can be used to forecast gasoline sales, we will use a three-week moving average ($k = 3$). We begin by computing the forecast of sales in week 4 using the average of the time series values in weeks 1–3.

$$F_4 = \text{average of weeks } 1\text{–}3 = \frac{17 + 21 + 19}{3} = 19$$

Thus, the moving average forecast of sales in week 4 is 19 or 19,000 gallons of gasoline. Because the actual value observed in week 4 is 23, the forecast error in week 4 is $23 - 19 = 4$.

Next, we compute the forecast of sales in week 5 by averaging the time series values in weeks 2–4.

$$F_5 = \text{average of weeks } 2\text{–}4 = \frac{21 + 19 + 23}{3} = 21$$

Hence, the forecast of sales in week 5 is 21 and the error associated with this forecast is $18 - 21 = -3$. A complete summary of the three-week moving average forecasts for the gasoline sales time series is provided in Table 18.9. Figure 18.7 shows the original time series plot and the three-week moving average forecasts. Note how the graph of the moving average forecasts has tended to smooth out the random fluctuations in the time series.

TABLE 18.9   SUMMARY OF THREE-WEEK MOVING AVERAGE CALCULATIONS

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21						
3	19						
4	23	19	4	4	16	17.39	17.39
5	18	21	−3	3	9	−16.67	16.67
6	16	20	−4	4	16	−25.00	25.00
7	20	19	1	1	1	5.00	5.00
8	18	18	0	0	0	0.00	0.00
9	22	18	4	4	16	18.18	18.18
10	20	20	0	0	0	0.00	0.00
11	15	20	−5	5	25	−33.33	33.33
12	22	19	3	3	9	13.64	13.64
		Totals	0	24	92	−20.79	129.21

FIGURE 18.7   GASOLINE SALES TIME SERIES PLOT AND THREE-WEEK MOVING
AVERAGE FORECASTS
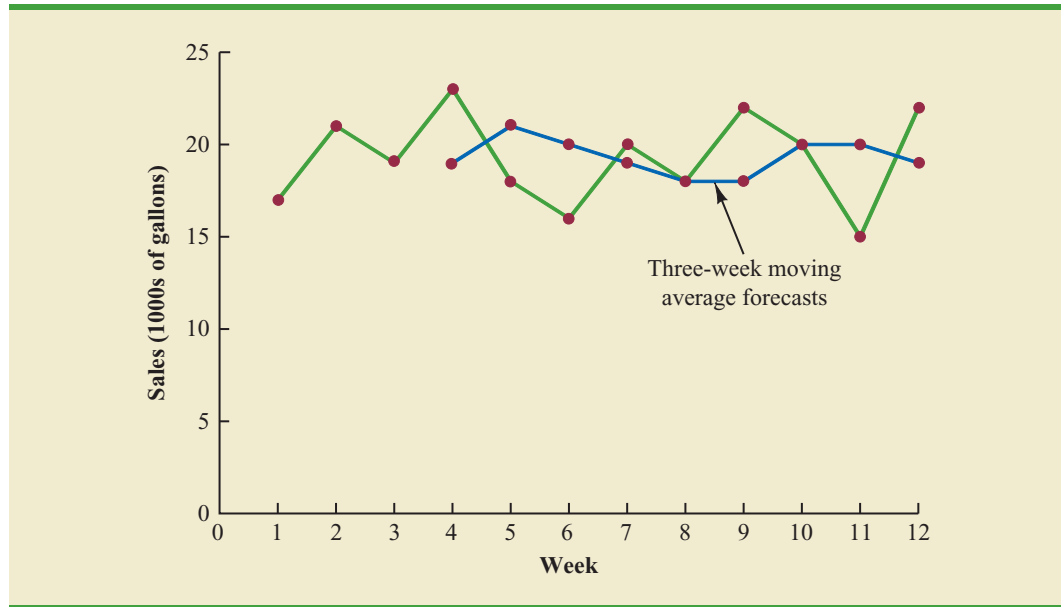
To forecast sales in week 13, the next time period in the future, we simply compute the average of the time series values in weeks 10, 11, and 12.

$$F_{13} = \text{average of weeks 10–12} = \frac{20 + 15 + 22}{3} = 19$$

Thus, the forecast for week 13 is 19 or 19,000 gallons of gasoline.

Forecast accuracy   In Section 18.2 we discussed three measures of forecast accuracy: MAE, MSE, and MAPE. Using the three-week moving average calculations in Table 18.9, the values for these three measures of forecast accuracy are

$$\text{MAE} = \frac{24}{9} = 2.67$$

$$\text{MSE} = \frac{92}{9} = 10.22$$

$$\text{MAPE} = \frac{129.21}{9} = 14.36\%$$

In situations where you need to compare forecasting methods for different time periods, such as comparing a forecast of weekly sales to a forecast of monthly sales, relative measures such as MAPE are preferred.

In Section 18.2 we also showed that using the most recent observation as the forecast for the next week (a moving average of order $k = 1$) resulted in values of MAE = 3.73, MSE = 16.27, and MAPE = 19.24%. Thus, in each case the three-week moving average approach provided more accurate forecasts than simply using the most recent observation as the forecast.

To determine if a moving average with a different order k can provide more accurate forecasts, we recommend using trial and error to determine the value of k that minimizes MSE. For the gasoline sales time series, it can be shown that the minimum value of MSE corresponds to a moving average of order $k = 6$ with MSE = 6.79. If we are willing to

assume that the order of the moving average that is best for the historical data will also be best for future values of the time series, the most accurate moving average forecasts of gasoline sales can be obtained using a moving average of order $k = 6$.

Weighted Moving Averages

A moving average forecast of order $k = 3$ is just a special case of the weighted moving averages method in which each weight is equal to 1/3.

In the moving averages method, each observation in the moving average calculation receives the same weight. One variation, known as **weighted moving averages**, involves selecting a different weight for each data value and then computing a weighted average of the most recent k values as the forecast. In most cases, the most recent observation receives the most weight, and the weight decreases for older data values. Let us use the gasoline sales time series to illustrate the computation of a weighted three-week moving average. We assign a weight of 3/6 to the most recent observation, a weight of 2/6 to the second most recent observation, and a weight of 1/6 to the third most recent observation. Using this weighted average, our forecast for week 4 is computed as follows:

$$\text{Forecast for week } 4 = \tfrac{1}{6}(17) + \tfrac{2}{6}(21) + \tfrac{3}{6}(19) = 19.33$$

Note that for the weighted moving average method the sum of the weights is equal to 1.

Forecast accuracy To use the weighted moving averages method, we must first select the number of data values to be included in the weighted moving average and then choose weights for each of the data values. In general, if we believe that the recent past is a better predictor of the future than the distant past, larger weights should be given to the more recent observations. However, when the time series is highly variable, selecting approximately equal weights for the data values may be best. The only requirement in selecting the weights is that their sum must equal 1. To determine whether one particular combination of number of data values and weights provides a more accurate forecast than another combination, we recommend using MSE as the measure of forecast accuracy. That is, if we assume that the combination that is best for the past will also be best for the future, we would use the combination of number of data values and weights that minimizes MSE for the historical time series to forecast the next value in the time series.

Exponential Smoothing

There are a number of exponential smoothing procedures. The method presented here is often referred to as single exponential smoothing. In the next section we show how an exponential smoothing method that uses two smoothing constants can be used to forecast a time series with a linear trend.

Exponential smoothing also uses a weighted average of past time series values as a forecast; it is a special case of the weighted moving averages method in which we select only one weight—the weight for the most recent observation. The weights for the other data values are computed automatically and become smaller as the observations move farther into the past. The exponential smoothing equation follows.

EXPONENTIAL SMOOTHING FORECAST

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t \tag{18.2}$$

where

$$\begin{aligned}
F_{t+1} &= \text{forecast of the time series for period } t + 1 \\
Y_t &= \text{actual value of the time series in period } t \\
F_t &= \text{forecast of the time series for period } t \\
\alpha &= \text{smoothing constant } (0 \leq \alpha \leq 1)
\end{aligned}$$

Equation (18.2) shows that the forecast for period $t + 1$ is a weighted average of the actual value in period t and the forecast for period t. The weight given to the actual value in period t is the **smoothing constant** α and the weight given to the forecast in period t is $1 - \alpha$. It turns out that the exponential smoothing forecast for any period is actually a weighted average of *all the previous actual values* of the time series. Let us illustrate by working with a time series involving only three periods of data: Y_1, Y_2, and Y_3.

To initiate the calculations, we let F_1 equal the actual value of the time series in period 1; that is, $F_1 = Y_1$. Hence, the forecast for period 2 is

$$
\begin{aligned}
F_2 &= \alpha Y_1 + (1 - \alpha)F_1 \\
&= \alpha Y_1 + (1 - \alpha)Y_1 \\
&= Y_1
\end{aligned}
$$

We see that the exponential smoothing forecast for period 2 is equal to the actual value of the time series in period 1.

The forecast for period 3 is

$$
F_3 = \alpha Y_2 + (1 - \alpha)F_2 = \alpha Y_2 + (1 - \alpha)Y_1
$$

Finally, substituting this expression for F_3 in the expression for F_4, we obtain

$$
\begin{aligned}
F_4 &= \alpha Y_3 + (1 - \alpha)F_3 \\
&= \alpha Y_3 + (1 - \alpha)[\alpha Y_2 + (1 - \alpha)Y_1] \\
&= \alpha Y_3 + \alpha(1 - \alpha)Y_2 + (1 - \alpha)^2 Y_1
\end{aligned}
$$

The term exponential smoothing comes from the exponential nature of the weighting scheme for the historical values.

We now see that F_4 is a weighted average of the first three time series values. The sum of the coefficients, or weights, for Y_1, Y_2, and Y_3 equals 1. A similar argument can be made to show that, in general, any forecast F_{t+1} is a weighted average of all the previous time series values.

Despite the fact that exponential smoothing provides a forecast that is a weighted average of all past observations, all past data do not need to be saved to compute the forecast for the next period. In fact, equation (18.2) shows that once the value for the smoothing constant α is selected, only two pieces of information are needed to compute the forecast: Y_t, the actual value of the time series in period t, and F_t, the forecast for period t.

To illustrate the exponential smoothing approach, let us again consider the gasoline sales time series in Table 18.1 and Figure 18.1. As indicated previously, to start the calculations we set the exponential smoothing forecast for period 2 equal to the actual value of the time series in period 1. Thus, with $Y_1 = 17$, we set $F_2 = 17$ to initiate the computations. Referring to the time series data in Table 18.1, we find an actual time series value in period 2 of $Y_2 = 21$. Thus, period 2 has a forecast error of $21 - 17 = 4$.

Continuing with the exponential smoothing computations using a smoothing constant of $\alpha = .2$, we obtain the following forecast for period 3:

$$
F_3 = .2Y_2 + .8F_2 = .2(21) + .8(17) = 17.8
$$

Once the actual time series value in period 3, $Y_3 = 19$, is known, we can generate a forecast for period 4 as follows:

$$
F_4 = .2Y_3 + .8F_3 = .2(19) + .8(17.8) = 18.04
$$

Continuing the exponential smoothing calculations, we obtain the weekly forecast values shown in Table 18.10. Note that we have not shown an exponential smoothing forecast

TABLE 18.10   SUMMARY OF THE EXPONENTIAL SMOOTHING FORECASTS
                   AND FORECAST ERRORS FOR THE GASOLINE SALES TIME SERIES
                   WITH SMOOTHING CONSTANT $\alpha = .2$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	−1.03	1.06
6	16	18.83	−2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	−0.61	0.37
9	22	18.49	3.51	12.32
10	20	19.19	0.81	0.66
11	15	19.35	−4.35	18.92
12	22	18.48	3.52	12.39
		Totals	10.92	98.80

or a forecast error for week 1 because no forecast was made. For week 12, we have $Y_{12} = 22$ and $F_{12} = 18.48$. We can we use this information to generate a forecast for week 13.

$$F_{13} = .2Y_{12} + .8F_{12} = .2(22) + .8(18.48) = 19.18$$

Thus, the exponential smoothing forecast of the amount sold in week 13 is 19.18, or 19,180 gallons of gasoline. With this forecast, the firm can make plans and decisions accordingly.

Figure 18.8 shows the time series plot of the actual and forecast time series values. Note in particular how the forecasts "smooth out" the irregular or random fluctuations in the time series.

Forecast accuracy In the preceding exponential smoothing calculations, we used a smoothing constant of $\alpha = .2$. Although any value of α between 0 and 1 is acceptable, some values will yield better forecasts than others. Insight into choosing a good value for α can be obtained by rewriting the basic exponential smoothing model as follows:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$
$$F_{t+1} = \alpha Y_t + F_t - \alpha F_t$$
$$F_{t+1} = F_t + \alpha(Y_t - F_t) \tag{18.3}$$

Thus, the new forecast F_{t+1} is equal to the previous forecast F_t plus an adjustment, which is the smoothing constant α times the most recent forecast error, $Y_t - F_t$. That is, the forecast in period $t + 1$ is obtained by adjusting the forecast in period t by a fraction of the forecast error. If the time series contains substantial random variability, a small value of the smoothing constant is preferred. The reason for this choice is that if much of the forecast error is due to random variability, we do not want to overreact and adjust the forecasts too quickly. For a time series with relatively little random variability, forecast errors are more likely to represent a change in the level of the series. Thus, larger values of the smoothing

FIGURE 18.8   ACTUAL AND FORECAST GASOLINE SALES TIME SERIES
WITH SMOOTHING CONSTANT $\alpha = .2$
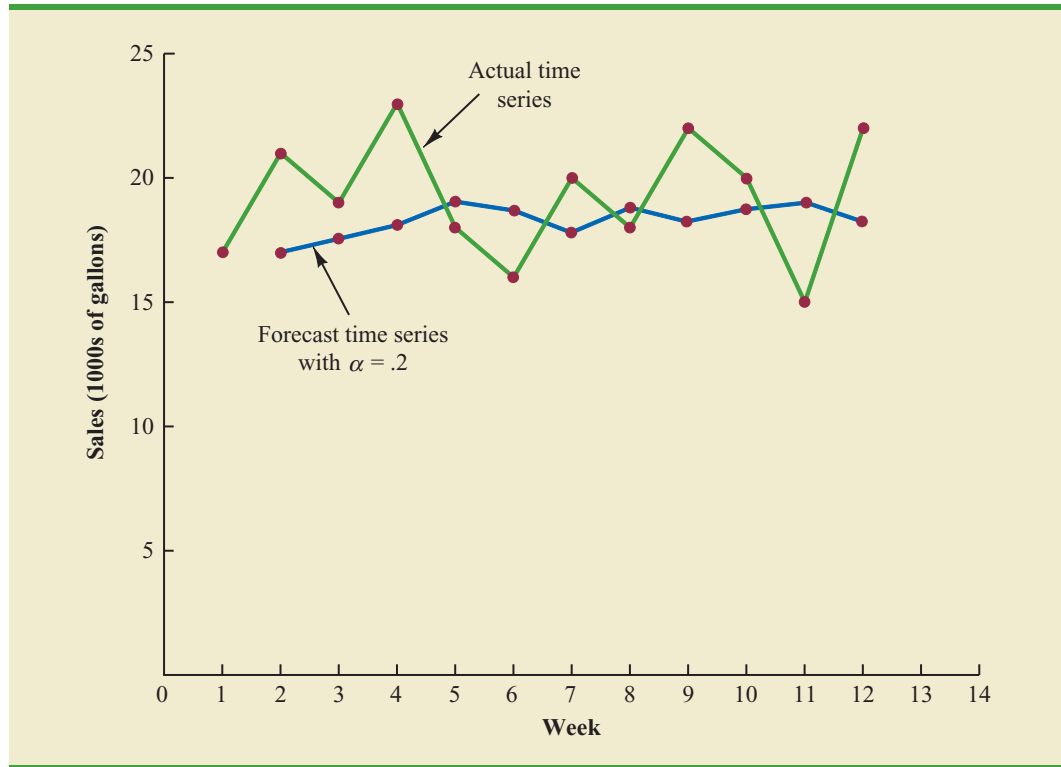
constant provide the advantage of quickly adjusting the forecasts; this allows the forecasts to react more quickly to changing conditions.

The criterion we will use to determine a desirable value for the smoothing constant α is the same as the criterion we proposed for determining the order or number of periods of data to include in the moving averages calculation. That is, we choose the value of α that minimizes the MSE. A summary of the MSE calculations for the exponential smoothing forecast of gasoline sales with $\alpha = .2$ is shown in Table 18.10. Note that there is one less squared error term than the number of time periods because we had no past values with which to make a forecast for period 1. The value of the sum of squared forecast errors is 98.80; hence MSE $= 98.80/11 = 8.98$. Would a different value of α provide better results in terms of a lower MSE value? Perhaps the most straightforward way to answer this question is simply to try another value for α. We will then compare its mean squared error with the MSE value of 8.98 obtained by using a smoothing constant of $\alpha = .2$.

The exponential smoothing results with $\alpha = .3$ are shown in Table 18.11. The value of the sum of squared forecast errors is 102.83; hence MSE $= 102.83/11 = 9.35$. With MSE $= 9.35$, we see that, for the current data set, a smoothing constant of $\alpha = .3$ results in less forecast accuracy than a smoothing constant of $\alpha = .2$. Thus, we would be inclined to prefer the original smoothing constant of $\alpha = .2$. Using a trial-and-error calculation with other values of α, we can find a "good" value for the smoothing constant. This value can be used in the exponential smoothing model to provide forecasts for the future. At a later date, after new time series observations are obtained, we analyze the newly collected time series data to determine whether the smoothing constant should be revised to provide better forecasting results.

TABLE 18.11   SUMMARY OF THE EXPONENTIAL SMOOTHING FORECASTS AND
FORECAST ERRORS FOR THE GASOLINE SALES TIME SERIES WITH
SMOOTHING CONSTANT $\alpha = .3$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	18.20	0.80	0.64
4	23	18.44	4.56	20.79
5	18	19.81	−1.81	3.28
6	16	19.27	−3.27	10.69
7	20	18.29	1.71	2.92
8	18	18.80	−0.80	0.64
9	22	18.56	3.44	11.83
10	20	19.59	0.41	0.17
11	15	19.71	−4.71	22.18
12	22	18.30	3.70	13.69
		Totals	8.03	102.83

NOTES AND COMMENTS

1. Spreadsheet packages are an effective aid in choosing a good value of α for exponential smoothing. With the time series data and the forecasting formulas in a spreadsheet, you can experiment with different values of α and choose the value that provides the smallest forecast error using one or more of the measures of forecast accuracy (MAE, MSE, or MAPE).

2. We presented the moving average and exponential smoothing methods in the context of a stationary time series. These methods can also be used to forecast a nonstationary time series which shifts in level but exhibits no trend or seasonality. Moving averages with small values of k adapt more quickly than moving averages with larger values of k. Exponential smoothing models with smoothing constants closer to one adapt more quickly than models with smaller values of the smoothing constant.

Exercises

Methods

5. Consider the following time series data.

Week	1	2	3	4	5	6
Value	18	13	16	11	17	14

   a. Construct a time series plot. What type of pattern exists in the data?
   b. Develop the three-week moving average forecasts for this time series. Compute MSE and a forecast for week 7.
   c. Use $\alpha = .2$ to compute the exponential smoothing forecasts for the time series. Compute MSE and a forecast for week 7.

   d.   Compare the three-week moving average approach with the exponential smoothing
        approach using $\alpha = .2$. Which appears to provide more accurate forecasts based on
        MSE? Explain.
   e.   Use a smoothing constant of $\alpha = .4$ to compute the exponential smoothing forecasts.
        Does a smoothing constant of .2 or .4 appear to provide more accurate forecasts based
        on MSE? Explain.

6. Consider the following time series data.

Month	1	2	3	4	5	6	7
Value	24	13	20	12	19	23	15

   Construct a time series plot. What type of pattern exists in the data?
   a.   Develop the three-week moving average forecasts for this time series. Compute MSE
        and a forecast for week 8.
   b.   Use $\alpha = .2$ to compute the exponential smoothing forecasts for the time series. Com-
        pute MSE and a forecast for week 8.
   c.   Compare the three-week moving average approach with the exponential smoothing ap-
        proach using $\alpha = .2$. Which appears to provide more accurate forecasts based on MSE?
   d.   Use a smoothing constant of $\alpha = .4$ to compute the exponential smoothing forecasts.
        Does a smoothing constant of .2 or .4 appear to provide more accurate forecasts based
        on MSE? Explain.

WEB file

Gasoline

7. Refer to the gasoline sales time series data in Table 18.1.
   a.   Compute four-week and five-week moving averages for the time series.
   b.   Compute the MSE for the four-week and five-week moving average forecasts.
   c.   What appears to be the best number of weeks of past data (three, four, or five) to use
        in the moving average computation? Recall that MSE for the three-week moving
        average is 10.22.

WEB file

Gasoline

8. Refer again to the gasoline sales time series data in Table 18.1.
   a.   Using a weight of 1/2 for the most recent observation, 1/3 for the second most recent
        observation, and 1/6 for third most recent observation, compute a three-week weighted
        moving average for the time series.
   b.   Compute the MSE for the weighted moving average in part (a). Do you prefer this
        weighted moving average to the unweighted moving average? Remember that the
        MSE for the unweighted moving average is 10.22.
   c.   Suppose you are allowed to choose any weights as long as they sum to 1. Could you
        always find a set of weights that would make the MSE at least as small for a weighted
        moving average than for an unweighted moving average? Why or why not?

WEB file

Gasoline

9. With the gasoline time series data from Table 18.1, show the exponential smoothing fore-
    casts using $\alpha = .1$.
   a.   Applying the MSE measure of forecast accuracy, would you prefer a smoothing con-
        stant of $\alpha = .1$ or $\alpha = .2$ for the gasoline sales time series?
   b.   Are the results the same if you apply MAE as the measure of accuracy?
   c.   What are the results if MAPE is used?

10. With a smoothing constant of $\alpha = .2$, equation (18.2) shows that the forecast for week 13
     of the gasoline sales data from Table 18.1 is given by $F_{13} = .2Y_{12} + .8F_{12}$. However, the
     forecast for week 12 is given by $F_{12} = .2Y_{11} + .8F_{11}$. Thus, we could combine these two
     results to show that the forecast for week 13 can be written

$$F_{13} = .2Y_{12} + .8(.2Y_{11} + .8F_{11}) = .2Y_{12} + .16Y_{11} + .64Y_{11} + .64F_{11}$$

   a.   Making use of the fact that $F_{11} = .2Y_{10} + .8F_{10}$ (and similarly for F_{10} and F_9), continue
        to expand the expression for F_{13} until it is written in terms of the past data values Y_{12},
        Y_{11}, Y_{10}, Y_9, Y_8, and the forecast for period 8.

b. Refer to the coefficients or weights for the past values Y_{12}, Y_{11}, Y_{10}, Y_9, Y_8. What observation can you make about how exponential smoothing weights past data values in arriving at new forecasts? Compare this weighting pattern with the weighting pattern of the moving averages method.

Applications

11. For the Hawkins Company, the monthly percentages of all shipments received on time over the past 12 months are 80, 82, 84, 83, 83, 84, 85, 84, 82, 83, 84, and 83.
    a. Construct a time series plot. What type of pattern exists in the data?
    b. Compare the three-month moving average approach with the exponential smoothing approach for $\alpha = .2$. Which provides more accurate forecasts using MSE as the measure of forecast accuracy?
    c. What is the forecast for next month?

12. Corporate triple-A bond interest rates for 12 consecutive months follow.

    9.5   9.3   9.4   9.6   9.8   9.7   9.8   10.5   9.9   9.7   9.6   9.6

    a. Construct a time series plot. What type of pattern exists in the data?
    b. Develop three-month and four-month moving averages for this time series. Does the three-month or four-month moving average provide more accurate forecasts based on MSE? Explain.
    c. What is the moving average forecast for the next month?

SELF test

13. The values of Alabama building contracts (in $ millions) for a 12-month period follow.

    240   350   230   260   280   320   220   310   240   310   240   230

    a. Construct a time series plot. What type of pattern exists in the data?
    b. Compare the three-month moving average approach with the exponential smoothing forecast using $\alpha = .2$. Which provides more accurate forecasts based on MSE?
    c. What is the forecast for the next month?

14. The following time series shows the sales of a particular product over the past 12 months.

Month	Sales	Month	Sales
1	105	7	145
2	135	8	140
3	120	9	100
4	105	10	80
5	90	11	100
6	120	12	110

    a. Construct a time series plot. What type of pattern exists in the data?
    b. Use $\alpha = .3$ to compute the exponential smoothing forecasts for the time series.
    c. Use a smoothing constant of $\alpha = .5$ to compute the exponential smoothing forecasts. Does a smoothing constant of .3 or .5 appear to provide more accurate forecasts based on MSE?

15. Ten weeks of data on the Commodity Futures Index are 7.35, 7.40, 7.55, 7.56, 7.60, 7.52, 7.52, 7.70, 7.62, and 7.55.
    a. Construct a time series plot. What type of pattern exists in the data?
    b. Compute the exponential smoothing forecasts for $\alpha = .2$.
    c. Compute the exponential smoothing forecasts for $\alpha = .3$.
    d. Which exponential smoothing constant provides more accurate forecasts based on MSE? Forecast week 11.

16. The Nielsen ratings (percentage of U.S. households that tuned in) for the Masters Golf Tournament from 1997 through 2008 follow (*Golf Magazine*, January 2009).

Year	Rating
1997	11.2
1998	8.6
1999	7.9
2000	7.6
2001	10.7
2002	8.1
2003	6.9
2004	6.7
2005	8.0
2006	6.9
2007	7.6
2008	7.3

WEB file

Masters

The rating of 11.2 in 1997 indicates that 11.2% of U.S. households tuned in to watch Tiger Woods win his first major golf tournament and become the first African American to win the Masters. Tiger Woods also won the Masters in 2001 and 2005.

a.   Construct a time series plot. What type of pattern exists in the data? Discuss some of the factors that may have resulted in the pattern exhibited in the time series plot for this time series.

b.   Given the pattern of the time series plot developed in part (a), do you think the forecasting methods discussed in this section are appropriate to develop forecasts for this time series? Explain.

c.   Would you recommend using the Nielsen ratings for only 2002–2008 to forecast the rating for 2009, or should the entire time series from 1997–2008 be used? Explain.

18.4 Trend Projection

We present three forecasting methods in this section that are appropriate for time series exhibiting a trend pattern. First, we show how simple linear regression can be used to forecast a time series with a linear trend. We then illustrate how to develop forecasts using Holt's linear exponential smoothing, an extension of single exponential smoothing that uses two smoothing constants: one to account for the level of the time series and a second to account for the linear trend in the data. Finally, we show how the curve-fitting capability of regression analysis can also be used to forecast time series with a curvilinear or nonlinear trend.

WEB file

Bicycle

Linear Trend Regression

TABLE 18.12

BICYCLE SALES TIME SERIES

Year	Sales (1000s)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

In Section 18.1 we used the bicycle sales time series in Table 18.3 and Figure 18.3 to illustrate a time series with a trend pattern. Let us now use this time series to illustrate how simple linear regression can be used to forecast a time series with a linear trend. The data for the bicycle time series are repeated in Table 18.12 and Figure 18.9.

Although the time series plot in Figure 18.9 shows some up and down movement over the past 10 years, we might agree that the linear trend line shown in Figure 18.10 provides a reasonable approximation of the long-run movement in the series. We can use the methods of simple linear regression (see Chapter 14) to develop such a linear trend line for the bicycle sales time series.

FIGURE 18.9   BICYCLE SALES TIME SERIES PLOT
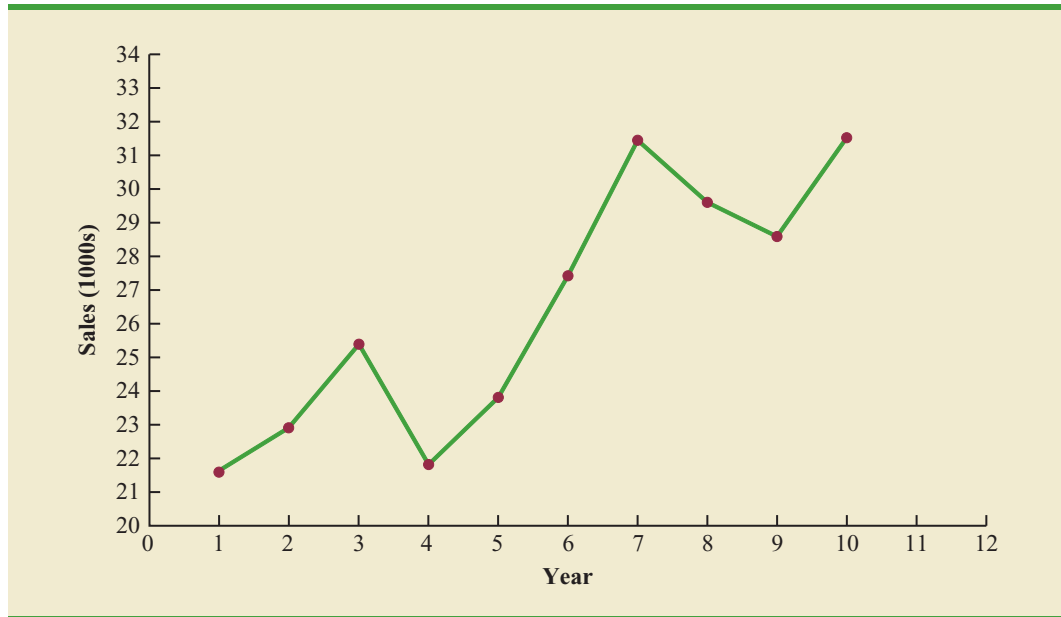
FIGURE 18.10   TREND REPRESENTED BY A LINEAR FUNCTION FOR THE BICYCLE SALES TIME SERIES
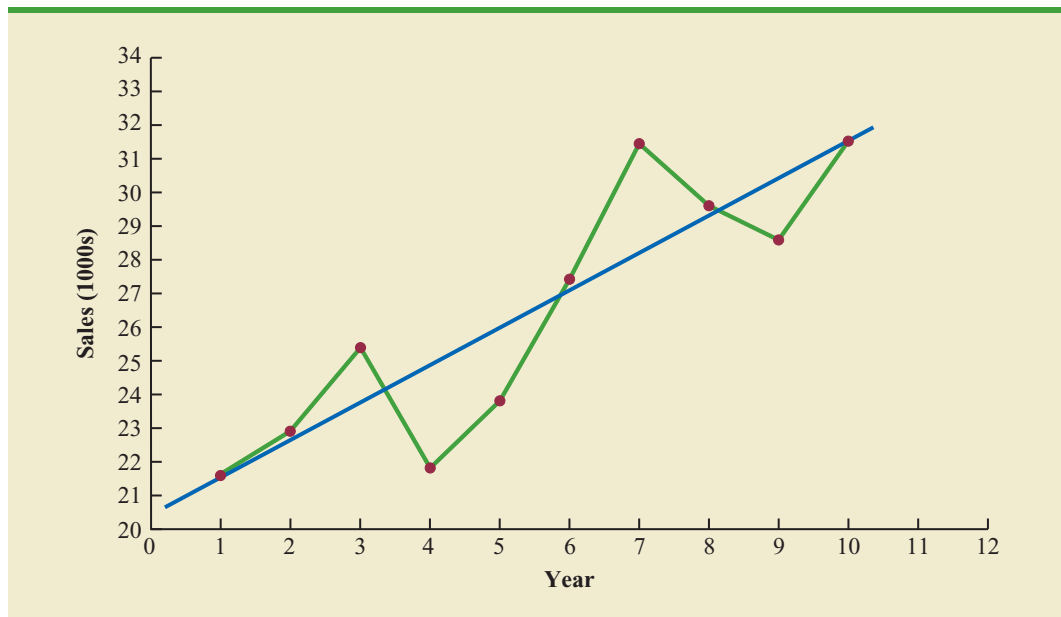
In Chapter 14, the estimated regression equation describing a straight-line relationship between an independent variable x and a dependent variable y is written as

$$\hat{y} = b_0 + b_1 x$$

where $\hat{y}$ is the estimated or predicted value of y. To emphasize the fact that in forecasting the independent variable is time, we will replace x with t and $\hat{y}$ with T_t to emphasize that we are estimating the trend for a time series. Thus, for estimating the linear trend in a time series we will use the following estimated regression equation.

LINEAR TREND EQUATION

$$T_t = b_0 + b_1 t \tag{18.4}$$

where

T_t = linear trend forecast in period t
b_0 = intercept of the linear trend line
b_1 = slope of the linear trend line
t = time period

In equation (18.4), the time variable begins at $t = 1$ corresponding to the first time series observation (year 1 for the bicycle sales time series) and continues until $t = n$ corresponding to the most recent time series observation (year 10 for the bicycle sales time series). Thus, for the bicycle sales time series $t = 1$ corresponds to the oldest time series value and $t = 10$ corresponds to the most recent year.

Formulas for computing the estimated regression coefficients (b_1 and b_0) in equation (18.4) follow.

COMPUTING THE SLOPE AND INTERCEPT FOR A LINEAR TREND*

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} \tag{18.5}$$

$$b_0 = \bar{Y} - b_1\bar{t} \tag{18.6}$$

where

Y_t = value of the time series in period t
n = number of time periods (number of observations)
$\bar{Y}$ = average value of the time series
$\bar{t}$ = average value of t

_____
*An alternate formula for b_1 is

$$b_1 = \frac{\sum_{t=1}^{n}tY_t - \left(\sum_{t=1}^{n}t \sum_{t=1}^{n}Y_t\right)/n}{\sum_{t=1}^{n}t^2 - \left(\sum_{t=1}^{n}t\right)^2/n}$$

This form of equation (18.5) is often recommended when using a calculator to compute b_1.

To compute the linear trend equation for the bicycle sales time series, we begin the calculations by computing $\bar{t}$ and $\bar{Y}$ using the information in Table 18.12.

$$\bar{t} = \frac{\sum_{t=1}^{n}t}{n} = \frac{55}{10} = 5.5$$

$$\bar{Y} = \frac{\sum_{t=1}^{n}Y_t}{n} = \frac{264.5}{10} = 26.45$$

TABLE 18.13   SUMMARY OF LINEAR TREND CALCULATIONS FOR THE BICYCLE SALES TIME SERIES

t	Y_t	$t - \bar{t}$	$Y_t - \bar{Y}$	$(t - \bar{t})(Y_t - \bar{Y})$	$(t - \bar{t})^2$
1	21.6	−4.5	−4.85	21.825	20.25
2	22.9	−3.5	−3.55	12.425	12.25
3	25.5	−2.5	−0.95	2.375	6.25
4	21.9	−1.5	−4.55	6.825	2.25
5	23.9	−0.5	−2.55	1.275	0.25
6	27.5	0.5	1.05	0.525	0.25
7	31.5	1.5	5.05	7.575	2.25
8	29.7	2.5	3.25	8.125	6.25
9	28.6	3.5	2.15	7.525	12.25
10	31.4	4.5	4.95	22.275	20.25
Totals 55	264.5			90.750	82.50

Using these values, and the information in Table 18.13, we can compute the slope and intercept of the trend line for the bicycle sales time series.

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{90.75}{82.5} = 1.1$$

$$b_0 = \bar{Y} - b_1\bar{t} = 26.45 - 1.1(5.5) = 20.4$$

Therefore, the linear trend equation is

$$T_t = 20.4 + 1.1t$$

The slope of 1.1 indicates that over the past 10 years the firm experienced an average growth in sales of about 1100 units per year. If we assume that the past 10-year trend in sales is a good indicator of the future, this trend equation can be used to develop forecasts for future time periods. For example, substituting $t = 11$ into the equation yields next year's trend projection or forecast, T_{11}.

$$T_{11} = 20.4 + 1.1(11) = 32.5$$

Thus, using trend projection, we would forecast sales of 32,500 bicycles next year.

To compute the accuracy associated with the trend projection forecasting method, we will use the MSE. Table 18.14 shows the computation of the sum of squared errors for the bicycle sales time series. Thus, for the bicycle sales time series,

$$\text{MSE} = \frac{\sum_{t=1}^{n}(Y_t - F_t)^2}{n} = \frac{30.7}{10} = 3.07$$

Because linear trend regression in forecasting uses the same regression analysis procedure introduced in Chapter 14, we can use the standard regression analysis procedures in Minitab or Excel to perform the calculations. Figure 18.11 shows the computer output for the bicycle sales time series obtained using Minitab's regression analysis module.

TABLE 18.14   SUMMARY OF THE LINEAR TREND FORECASTS AND FORECAST
ERRORS FOR THE BICYCLE SALES TIME SERIES

Year	Sales (1000s) Y_t	Forecast T_t	Forecast Error	Squared Forecast Error
1	21.6	21.5	0.1	0.01
2	22.9	22.6	0.3	0.09
3	25.5	23.7	1.8	3.24
4	21.9	24.8	−2.9	8.41
5	23.9	25.9	−2.0	4.00
6	27.5	27.0	0.5	0.25
7	31.5	28.1	3.4	11.56
8	29.7	29.2	0.5	0.25
9	28.6	30.3	−1.7	2.89
10	31.4	31.4	0.0	0.00
			Total	30.70

In Figure 18.11 the value of MSE in the ANOVA table is

$$\text{MSE} = \frac{\text{Sum of Squares Due to Error}}{\text{Degrees of Freedom}} = \frac{30.7}{8} = 3.837$$

MSD in Minitab's Trend Analysis output is the mean squared deviation, the average of the squared forecast errors.

This value of MSE differs from the value of MSE that we computed previously because the sum of squared errors is divided by 8 instead of 10; thus, MSE in the regression output is not the average of the squared forecast errors. Most forecasting packages, however, compute MSE by taking the average of the squared errors. Thus, when using time series packages to develop a trend equation, the value of MSE that is reported may differ slightly from the value you would obtain using a general regression approach. For instance, in Figure 18.12, we show the graphical portion of the computer output obtained using Minitab's Trend Analysis time series procedure. Note that MSD = 3.07 is the average of the squared forecast errors.

FIGURE 18.11   MINITAB REGRESSION OUTPUT FOR THE BICYCLE SALES
TIME SERIES

```
The regression equation is
Y = 20.4 + 1.10 t

Predictor    Coef    SE Coef        T        p
Constant    20.400     1.338    15.24    0.000
t            1.1000    0.2157     5.10    0.001

S = 1.95895   R-sq = 76.5%   R-sq(adj) = 73.5%

Analysis of Variance

SOURCE            DF      SS       MS       F       p
Regression         1   99.825   99.825   26.01   0.001
Residual Error     8   30.700    3.837
Total              9  130.525
```

FIGURE 18.12   MINITAB TIME SERIES LINEAR TREND ANALYSIS OUTPUT
FOR THE BICYCLE SALES TIME SERIES

Holt's Linear Exponential Smoothing
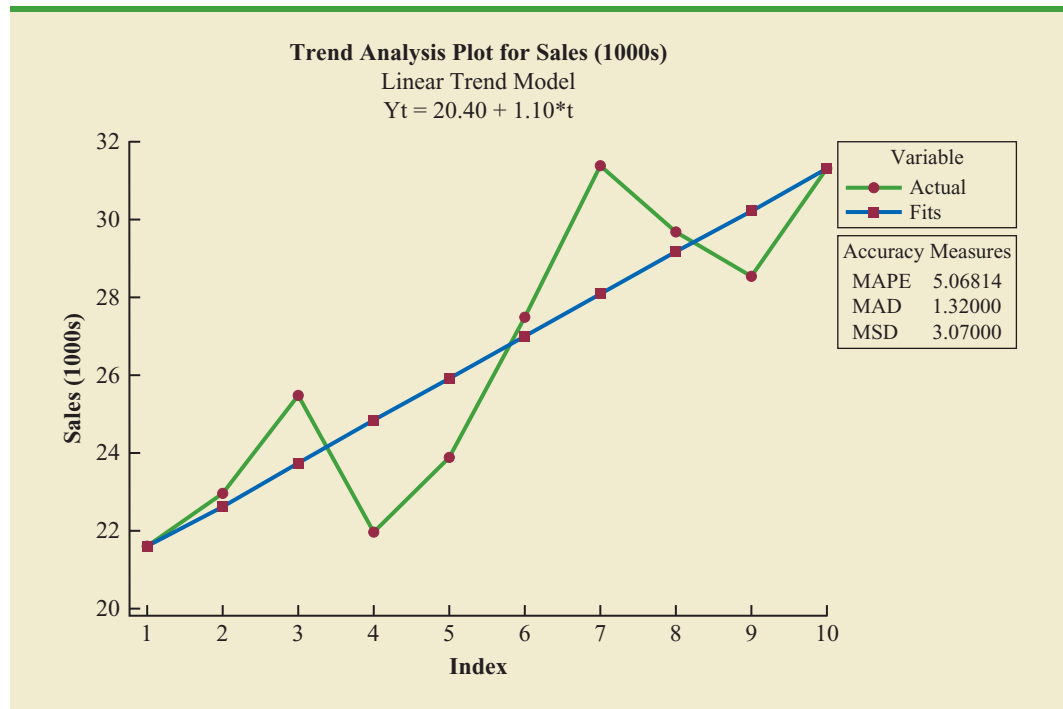
Holt's linear exponential smoothing is often called double exponential smoothing.

Charles Holt developed a version of exponential smoothing that can be used to forecast a time series with a linear trend. Recall that the exponential smoothing procedure discussed in Section 18.3 uses the smoothing constant α to "smooth out" the randomness or irregular fluctuations in a time series; and, forecasts for time period $t + 1$ are obtained using the equation

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

Forecasts for Holt's **linear exponential smoothing** method are obtained using two smoothing constants, α and β, and three equations.

EQUATIONS FOR HOLT'S LINEAR EXPONENTIAL SMOOTHING

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \qquad \textbf{(18.7)}$$
$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta) b_{t-1} \qquad \textbf{(18.8)}$$
$$F_{t+k} = L_t + b_t k \qquad \textbf{(18.9)}$$

where

L_t = estimate of the level of the time series in period t

b_t = estimage of the slope of the time series in period t

α = smoothing constant for the level of the time series

β = smoothing constant for the slope of the time series

F_{t+k} = forecast for k periods ahead

k = the number of periods ahead to be forecast

Let us apply Holt's method to the bicycle sales time series in Table 18.12 using α = .1 and β = .2. To get the method started, we need values for L_1, the estimate of the level of the time series in year 1, and b_1, the estimate of the slope of the time series in year 1. A commonly used approach is to set $L_1 = Y_1$ and $b_1 = Y_2 - Y_1$. Using this startup procedure, we obtain

$$L_1 = Y_1 = 21.6$$
$$b_1 = Y_2 - Y_1 = 22.9 - 21.6 = 1.3$$

Using equation (18.9) with $k = 1$, the forecast of sales in year 2 is $F_2 = L_1 + b_1 = 21.6 + 1.3(1) = 22.9$. Then we move on using equations (18.7) to (18.9) to compute estimates of the level and trend for year 2 as well as a forecast for year 3.

First we use equation (18.7) and the smoothing constant α = .1 to compute an estimate of the level of the time series in year 2.

$$L_2 = .1(22.9) + .9(21.6 + 1.3) = 22.9$$

Note that $21.6 + 1.3$ is the forecast of sales for year 2. Thus, the estimate of the level of the time series in year 2 obtained using equation (18.7) is simply a weighted average of the observed value in year 2 (using a weight of α = .1) and the forecast for year 2 (using a weight of $1 - \alpha = 1 - .1 = .9$). In general, large values of α place more weight on the observed value (Y_t) whereas smaller values place more weight on the forecasted value ($L_{t-1} + b_{t-1}$).

Next we use equation (18.8) and the smoothing constant β = .2 to compute an estimate of the slope of the time series in year 2.

$$b_2 = .2(22.9 - 21.6) + (1 - .2)(1.3) = 1.3$$

The estimate the slope of the time series in year 2 is a weighted average of the difference in the estimated level of the time series between year 2 and year 1 (using a weight of β = .2) and the estimate of the slope in year 1(using a weight of $1 - \beta = 1 - .2 = .8$). In general, higher values of β place more weight on the difference between the estimated levels, whereas smaller values place more weight on the estimate of the slope from the last period.

Using the estimates of L_2 and b_2 just obtained, the forecast of sales for year 3 is computed using equation (18.9):

$$F_3 = L_2 + b_2 = 22.9 + 1.3(1) = 24.2$$

The other calculations are made in a similar manner and are shown in Table 18.15. The sum of the squared forecast errors is 39.678; hence MSE = 39.678/9 = 4.41.

Will different values for the smoothing constants α and β provide more accurate forecasts? To answer this question we would have to try different combinations of α and β to determine if a combination can be found that will provide a value of MSE lower than 4.41, the value we obtained using smoothing constants α = .1 and β = .2. Searching for good values of α and β can be done by trial and error or using more advanced statistical software packages that have an option for selecting the optimal set of smoothing constants.

TABLE 18.15   SUMMARY CALCULATIONS FOR HOLT'S LINEAR EXPONENTIAL SMOOTHING FOR THE BICYCLE SALES TIME SERIES USING $\alpha = .1$ AND $\beta = .2$

Year	Sales (1000s) Y_t	Estimated Level L_t	Estimated Trend b_t	Forecast F_t	Forecast Error	Squared Forecast Error
1	21.6	21.600	1.300			
2	22.9	22.900	1.300	22.900	0.000	0.000
3	25.5	24.330	1.326	24.200	1.300	1.690
4	21.9	25.280	1.251	25.656	−3.756	14.108
5	23.9	26.268	1.198	26.531	−2.631	6.924
6	27.5	27.470	1.199	27.466	0.034	0.001
7	31.5	28.952	1.256	28.669	2.831	8.016
8	29.7	30.157	1.245	30.207	−0.507	0.257
9	28.6	31.122	1.189	31.402	−2.802	7.851
10	31.4	32.220	1.171	32.311	−0.911	0.830
					Total	39.678

Note that the estimate of the level of the time series in year 10 is $L_1 = 32.220$ and the estimate of the slope in year 10 is $b_1 = 1.171$. If we assume that the past 10-year trend in sales is a good indicator of the future, equation (18.9) can be used to develop forecasts for future time periods. For example, substituting $t = 11$ into equation (18.9) yields next year's trend projection or forecast, F_{11}.

$$F_{11} = L_{10} + b_{10}(1) = 32.220 + 1.171 = 33.391$$

Thus, using Holt's linear exponential smoothing we would forecast sales of 33,391 bicycles next year.

Nonlinear Trend Regression

WEB file

Cholesterol

The use of a linear function to model trend is common. However, as we discussed previously, sometimes time series have a curvilinear or nonlinear trend. As an example, consider the annual revenue in millions of dollars for a cholesterol drug for the first 10 years of sales. Table 18.16 shows the time series and Figure 18.13 shows the corresponding time series plot. For instance, revenue in year 1 was $23.1 million; revenue in year 2 was $21.3 million; and so on. The time series plot indicates an overall increasing or upward trend. But, unlike the bicycle sales time series, a linear trend does not appear to be appropriate. Instead, a curvilinear function appears to be needed to model the long-term trend.

TABLE 18.16

CHOLESTEROL REVENUE TIME SERIES ($MILLIONS)

Year (t)	Revenue ($millions)
1	23.1
2	21.3
3	27.4
4	34.6
5	33.8
6	43.2
7	59.5
8	64.4
9	74.2
10	99.3

Quadratic trend equation A variety of nonlinear functions can be used to develop an estimate of the trend for the cholesterol time series. For instance, consider the following quadratic trend equation:

$$T_t = b_0 + b_1 t + b_2 t^2 \tag{18.10}$$

For the cholesterol time series, $t = 1$ corresponds to year 1, $t = 2$ corresponds to year 2, and so on.

The general linear model discussed in Section 16.1 can be used to compute the values of b_0, b_1, and b_2. There are two independent variables, year and year squared, and the dependent variable is the sales revenue in millions of dollars. Thus, the first observation is 1,

FIGURE 18.13 CHOLESTEROL REVENUE TIMES SERIES PLOT ($MILLIONS)
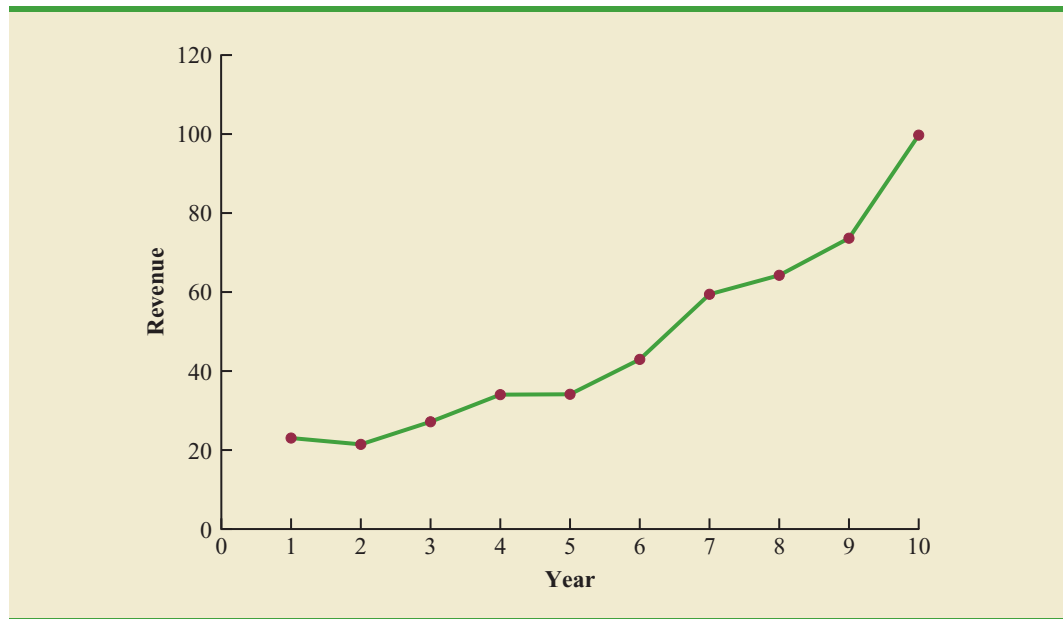
1, 23.1; the second observation is 2, 4, 21.3; the third observation is 3, 9, 27.4; and so on. Figure 18.14 shows the Minitab multiple regression output for the quadratic trend model; the estimated regression equation is

$$\text{Revenue (\$millions)} = 24.2 - 2.11 \text{ Year} + 0.922 \text{ YearSq}$$

where

$$\text{Year} = 1, 2, 3, \ldots, 10$$
$$\text{YearSq} = 1, 4, 9, \ldots, 100$$

FIGURE 18.14 MINITAB QUADRATIC TREND REGRESSION OUTPUT FOR THE BICYCLE SALES TIME SERIES

```
The regression equation is
Revenue = 24.2 - 2.11 Year + 0.922 YearSq

Predictor      Coef    SE Coef          T        p
Constant     24.182      4.676       5.17    0.001
Year         -2.106      1.953      -1.08    0.317
YearSq       0.9216      0.1730       5.33    0.001

S = 3.97578   R-Sq = 98.1%   R-Sq(adj) = 97.6%

Analysis of Variance

SOURCE            DF       SS      MS        F       p
Regression         2   5770.1  2885.1   182.52   0.000
Residual Error     7    110.6    15.8
Total              9   5880.8
```

Using the standard multiple regression procedure requires us to compute the values for year squared as a second independent variable. Alternatively, we can use Minitab's Time Series—Trend Analysis procedure to provide the same results. It does not require developing values for year squared and is easier to use. We recommend using this approach when solving exercises involving using quadratic trends.

Exponential trend equation Another alternative that can be used to model the non-linear pattern exhibited by the cholesterol time series is to fit an exponential model to the data. For instance, consider the following exponential trend equation:

$$T_t = b_0(b_1)^t \qquad\qquad (18.11)$$

To better understand this exponential trend equation, suppose $b_0 = 20$ and $b_1 = 1.2$. Then, for $t = 1$, $T_1 = 20(1.2)^1 = 24$; for $t = 2$, $T_2 = 20(1.2)^2 = 28.8$; and for $t = 3$, $T_3 = 20(1.2)^3 = 34.56$. Note that T_t is not increasing by a constant amount as in the case of the linear trend model, but by a constant percentage; the percentage increase is 20%.

Minitab has the capability in its time series module to compute an exponential trend equation and it can then be used for forecasting. Unfortunately, Excel does not have this capability. But, in Section 16.1, we do describe how, by taking logarithms of the terms in equation (18.11), the general linear model methodology can be used to compute an exponential trend equation.

Minitab's time series module is quite easy to use to develop an exponential trend equation. There is no need to deal with logarithms and use regression analysis to compute the exponential trend equation. In Figure 18.15, we show the graphical portion of the computer output obtained using Minitab's Trend Analysis time series procedure to fit an exponential trend equation.

FIGURE 18.15   MINITAB TIME SERIES EXPONENTIAL GROWTH TREND ANALYSIS OUTPUT FOR THE CHOLESTEROL SALES TIME SERIES
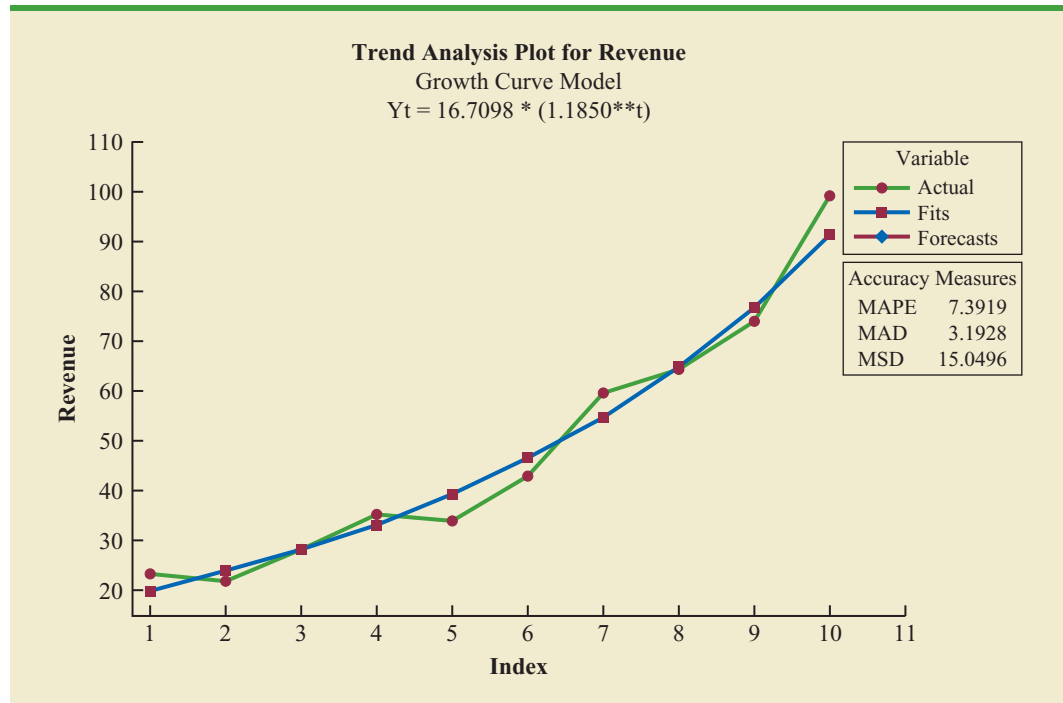
NOTES AND COMMENTS

Linear trend regression is based upon finding the estimated regression equation that minimizes the sum of squared forecast errors and therefore MSE. So, we would expect linear trend regression to outperform Holt's linear exponential smoothing in terms of MSE. For example, for the bicycle sales time series, the value of MSE using linear trend regression is 3.07 as compared to a value of 3.97 using Holt's linear exponential smoothing. Linear trend regression also provides a more accurate forecast using the MAE measure of forecast accuracy; for the bicycle sales time series, linear trend regression results in a value of MAE of 1.32 versus a value of 1.67 using Holt's linear method. However, based on MAPE, Holt's linear exponential smoothing (MAPE = 5.07%) outperforms linear trend regression (6.42%). Hence, for the bicycle sales time series, deciding which method provides the more accurate forecasts depends upon which measure of forecast accuracy is used.

Exercises

Methods

SELF test

17. Consider the following time series data.

t	1	2	3	4	5
Y_t	6	11	9	14	15

   a. Construct a time series plot. What type of pattern exists in the data?
   b. Develop the linear trend equation for this time series.
   c. What is the forecast for $t = 6$?

18. Refer to the time series in exercise 17. Use Holt's linear exponential smoothing method with $\alpha = .3$ and $\beta = .5$ to develop a forecast for $t = 6$.

19. Consider the following time series.

t	1	2	3	4	5	6	7
Y_t	120	110	100	96	94	92	88

   a. Construct a time series plot. What type of pattern exists in the data?
   b. Develop the linear trend equation for this time series.
   c. What is the forecast for $t = 8$?

20. Consider the following time series.

t	1	2	3	4	5	6	7
Y_t	82	60	44	35	30	29	35

   a. Construct a time series plot. What type of pattern exists in the data?
   b. Using Minitab or Excel, develop the quadratic trend equation for the time series.
   c. What is the forecast for $t = 8$?

Applications

SELF test

21. Because of high tuition costs at state and private universities, enrollments at community colleges have increased dramatically in recent years. The following data show the enrollment (in thousands) for Jefferson Community College from 2001–2009.

Year	Period (t)	Enrollment (1000s)
2001	1	6.5
2002	2	8.1
2003	3	8.4
2004	4	10.2
2005	5	12.5
2006	6	13.3
2007	7	13.7
2008	8	17.2
2009	9	18.1

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop the linear trend equation for this time series.
c. What is the forecast for 2010?

22. The Seneca Children's Fund (SCF) is a local charity that runs a summer camp for disadvantaged children. The fund's board of directors has been working very hard in recent years to decrease the amount of overhead expenses, a major factor in how charities are rated by independent agencies. The following data show the percentage of the money SCF has raised that was spent on administrative and fund-raising expenses for 2003–2009.

Year	Period (t)	Expense (%)
2003	1	13.9
2004	2	12.2
2005	3	10.5
2006	4	10.4
2007	5	11.5
2008	6	10.0
2009	7	8.5

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop the linear trend equation for this time series.
c. Forecast the percentage of administrative expenses for 2010.
d. If SCF can maintain their current trend in reducing administrative expenses, how long will it take them to achieve a level of 5% or less?

23. The president of a small manufacturing firm is concerned about the continual increase in manufacturing costs over the past several years. The following figures provide a time series of the cost per unit for the firm's leading product over the past eight years.

Year	Cost/Unit ($)	Year	Cost/Unit ($)
1	20.00	5	26.60
2	24.50	6	30.00
3	28.20	7	31.00
4	27.50	8	36.00

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop the linear trend equation for this time series.
c. What is the average cost increase that the firm has been realizing per year?
d. Compute an estimate of the cost/unit for next year.

24. FRED® (Federal Reserve Economic Data), a database of more than 3000 U.S. economic time series, contains historical data on foreign exchange rates. The following data show

the foreign exchange rate for the United States and China (Federal Reserve Bank of St.Louis website). The units for Rate are the number of Chinese yuan to one U.S. dollar.

WEB file

ExchangeRate

Year	Month	Rate
2007	October	7.5019
2007	November	7.4210
2007	December	7.3682
2008	January	7.2405
2008	February	7.1644
2008	March	7.0722
2008	April	6.9997
2008	May	6.9725
2008	June	6.8993
2008	July	6.8355

a.   Construct a time series plot. Does a linear trend appear to be present?
b.   Using Minitab or Excel, develop the linear trend equation for this time series.
c.   Use the trend equation to forecast the exchange rate for August 2008.
d.   Would you feel comfortable using the trend equation to forecast the exchange rate for December 2008?

25.   Automobile unit sales at B. J. Scott Motors, Inc., provided the following 10-year time series.

Year	Sales	Year	Sales
1	400	6	260
2	390	7	300
3	320	8	320
4	340	9	340
5	270	10	370

a.   Construct a time series plot. Comment on the appropriateness of a linear trend.
b.   Using Minitab or Excel, develop a quadratic trend equation that can be used to forecast sales.
c.   Using the trend equation developed in part (b), forecast sales in year 11.
d.   Suggest an alternative to using a quadratic trend equation to forecast sales. Explain.

26.   Giovanni Food Products produces and sells frozen pizzas to public schools throughout the eastern United States. Using a very aggressive marketing strategy they have been able to increase their annual revenue by approximately $10 million over the past 10 years. But increased competition has slowed their growth rate in the past few years. The annual revenue, in millions of dollars, for the previous 10 years is shown.

WEB file

Pasta

Year	Revenue
1	8.53
2	10.84
3	12.98
4	14.11
5	16.31
6	17.21
7	18.37
8	18.45
9	18.40
10	18.43

a. Construct a time series plot. Comment on the appropriateness of a linear trend.
b. Using Minitab or Excel, develop a quadratic trend equation that can be used to forecast revenue.
c. Using the trend equation developed in part (b), forecast revenue in year 11.

27. *Forbes* magazine ranks NFL teams by value each year. The following data are the value of the Indianapolis Colts from 1998 to 2008 (Forbes website).

WEB file

NFLValue

Year	Period	Value ($millions)
1998	1	227
1999	2	305
2000	3	332
2001	4	367
2002	5	419
2003	6	547
2004	7	609
2005	8	715
2006	9	837
2007	10	911
2008	11	1076

a. Construct a time series plot. What type of pattern exists in the data?
b. Using Minitab or Excel, develop the quadratic trend equation that can be used to forecast the team's value.
c. Using Minitab or Excel, develop the exponential trend equation that can be used to forecast the team's value.
d. Using Minitab or Excel, develop the linear trend equation that can be used to forecast the team's value.
e. Which equation would you recommend using to estimate the team's value in 2009?
f. Use the model you recommended in part (e) to forecast the value of the Colts in 2009.

18.5 Seasonality and Trend

In this section we show how to develop forecasts for a time series that has a seasonal pattern. To the extent that seasonality exists, we need to incorporate it into our forecasting models to ensure accurate forecasts. We begin by considering a seasonal time series with no trend and then discuss how to model seasonality with trend.

Seasonality Without Trend

As an example, consider the number of umbrellas sold at a clothing store over the past five years. Table 18.17 shows the time series and Figure 18.16 shows the corresponding time series plot. The time series plot does not indicate any long-term trend in sales. In fact, unless you look carefully at the data, you might conclude that the data follow a horizontal pattern and that single exponential smoothing could be used to forecast sales. But closer inspection of the time series plot reveals a pattern in the data. That is, the first and third quarters have moderate sales, the second quarter has the highest sales, and the fourth quarter tends to be the lowest quarter in terms of sales volume. Thus, we would conclude that a quarterly seasonal pattern is present.

WEB file

Umbrella

In Chapter 15 we showed how dummy variables can be used to deal with categorical independent variables in a multiple regression model. We can use the same approach to model a time series with a seasonal pattern by treating the season as a categorical variable.

Recall that when a categorical variable has k levels, $k - 1$ dummy variables are required. So, if there are four seasons, we need three dummy variables. For instance, in the umbrella sales time series season is a categorical variable with four levels: quarter 1, quarter 2, quarter 3, and quarter 4. Thus, to model the seasonal effects in the umbrella time series we need $4 - 1 = 3$ dummy variables. The three dummy variables can be coded as follows:

TABLE 18.17

UMBRELLA SALES
TIME SERIES

$$Qtr1 = \begin{cases} 1 \text{ if Quarter 1} \\ 0 \text{ otherwise} \end{cases} \quad Qtr2 = \begin{cases} 1 \text{ if Quarter 2} \\ 0 \text{ otherwise} \end{cases} \quad Qtr3 = \begin{cases} 1 \text{ if Quarter 3} \\ 0 \text{ otherwise} \end{cases}$$

Year	Quarter	Sales
1	1	125
	2	153
	3	106
	4	88
2	1	118
	2	161
	3	133
	4	102
3	1	138
	2	144
	3	113
	4	80
4	1	109
	2	137
	3	125
	4	109
5	1	130
	2	165
	3	128
	4	96

Using $\hat{Y}$ to denote the estimated or forecasted value of sales, the general form of the estimated regression equation relating the number of umbrellas sold to the quarter the sales take place follows:

$$\hat{Y} = b_0 + b_1 \, Qtr1 + b_2 \, Qtr2 + b_3 \, Qtr3$$

Table 18.18 is the umbrella sales time series with the coded values of the dummy variables shown. Using the data in Table 18.18 and Minitab's regression procedure, we obtained the computer output shown in Figure 18.17. The estimated multiple regression equation obtained is

$$\text{Sales} = 95.0 + 29.0 \, Qtr1 + 57.0 \, Qtr2 + 26.0 \, Qtr3$$

We can use this equation to forecast quarterly sales for next year.

Quarter 1: Sales $= 95.0 + 29.0(1) + 57.0(0) + 26.0(0) = 124$
Quarter 2: Sales $= 95.0 + 29.0(0) + 57.0(1) + 26.0(0) = 152$
Quarter 3: Sales $= 95.0 + 29.0(0) + 57.0(0) + 26.0(1) = 121$
Quarter 4: Sales $= 95.0 + 29.0(0) + 57.0(1) + 26.0(0) = 95$

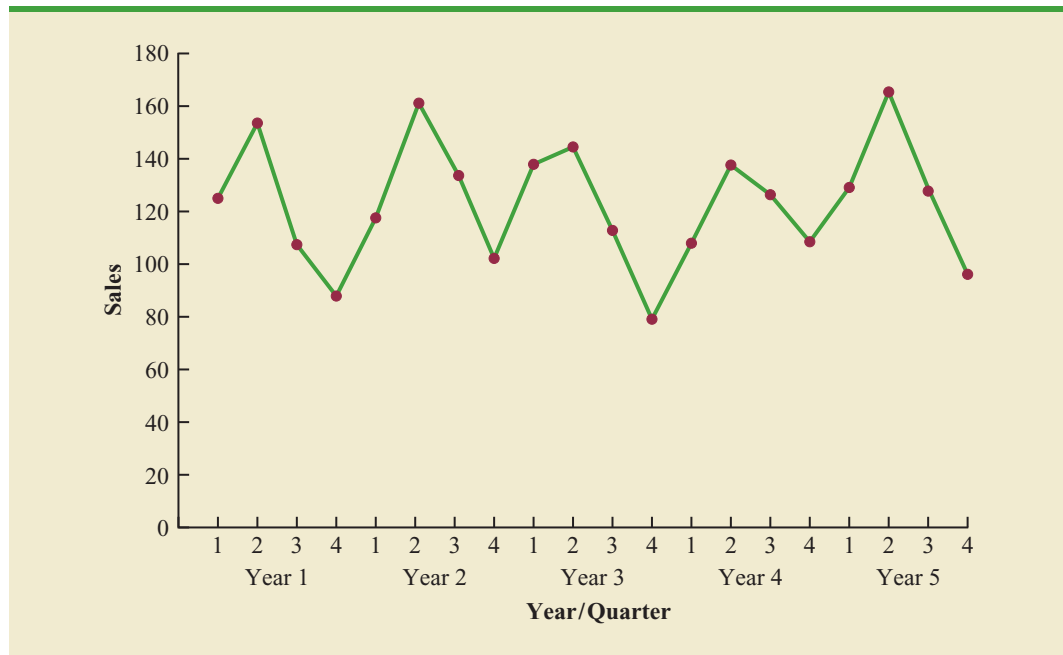**FIGURE 18.16**   UMBRELLA SALES TIME SERIES PLOT

TABLE 18.18   UMBRELLA SALES TIME SERIES WITH DUMMY VARIABLES

Year	Quarter	Qtr1	Qtr2	Qtr3	Sales
1	1	1	0	0	125
	2	0	1	0	153
	3	0	0	1	106
	4	0	0	0	88
2	1	1	0	0	118
	2	0	1	0	161
	3	0	0	1	133
	4	0	0	0	102
3	1	1	0	0	138
	2	0	1	0	144
	3	0	0	1	113
	4	0	0	0	80
4	1	1	0	0	109
	2	0	1	0	137
	3	0	0	1	125
	4	0	0	0	109
5	1	1	0	0	130
	2	0	1	0	165
	3	0	0	1	128
	4	0	0	0	96

It is interesting to note that we could have obtained the quarterly forecasts for next year simply by computing the average number of umbrellas sold in each quarter, as shown in the following table.

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	125	153	106	88
2	118	161	133	102
3	138	144	113	80
4	109	137	125	109
5	130	165	128	96
Average	124	152	121	95

Nonetheless, the regression output shown in Figure 18.17 provides additional information that can be used to assess the accuracy of the forecast and determine the significance of the

FIGURE 18.17   MINITAB REGRESSION OUTPUT FOR THE UMBRELLA
SALES TIME SERIES

```
The regression equation is
Sales = 95.0 + 29.0 Qtr1 + 57.0 Qtr2 + 26.0 Qtr3


Predictor    Coef   SE Coef      T       P
Constant    95.000    5.065   18.76   0.000
Qtr1        29.000    7.162    4.05   0.001
Qtr2        57.000    7.162    7.96   0.000
Qtr3        26.000    7.162    3.63   0.002
```

results. And, for more complex types of problem situations, such as dealing with a time series that has both trend and seasonal effects, this simple averaging approach will not work.

Seasonality and Trend

WEB file

TVSales

Let us now extend the regression approach to include situations where the time series contains both a seasonal effect and a linear trend by showing how to forecast the quarterly television set sales time series introduced in Section 18.1. The data for the television set time series are shown in Table 18.19. The time series plot in Figure 18.18 indicates that sales are lowest in the second quarter of each year and increase in quarters 3 and 4. Thus, we conclude that a seasonal pattern exists for television set sales. But the time series also has an upward linear trend that will need to be accounted for in order to develop accurate forecasts of quarterly sales. This is easily handled by combining the dummy variable approach for seasonality with the time series regression approach we discussed in Section 18.3 for handling linear trend.

TABLE 18.19

TELEVISION SET
SALES TIME
SERIES

Year	Quarter	Sales (1000s)
1	1	4.8
	2	4.1
	3	6.0
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6.0
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8.0
	4	8.4

The general form of the estimated multiple regression equation for modeling both the quarterly seasonal effects and the linear trend in the television set time series is as follows:

$$\hat{Y}_t = b_0 + b_1\,\text{Qtr1} + b_2\,\text{Qtr2} + b_3\,\text{Qtr3} + b_4 t$$

where

$\hat{Y}_t$ = estimate or forecast of sales in period t

Qtr1 = 1 if time period t corresponds to the first quarter of the year; 0 otherwise

Qtr2 = 1 if time period t corresponds to the second quarter of the year; 0 otherwise

Qtr3 = 1 if time period t corresponds to the third quarter of the year; 0 otherwise

t = time period
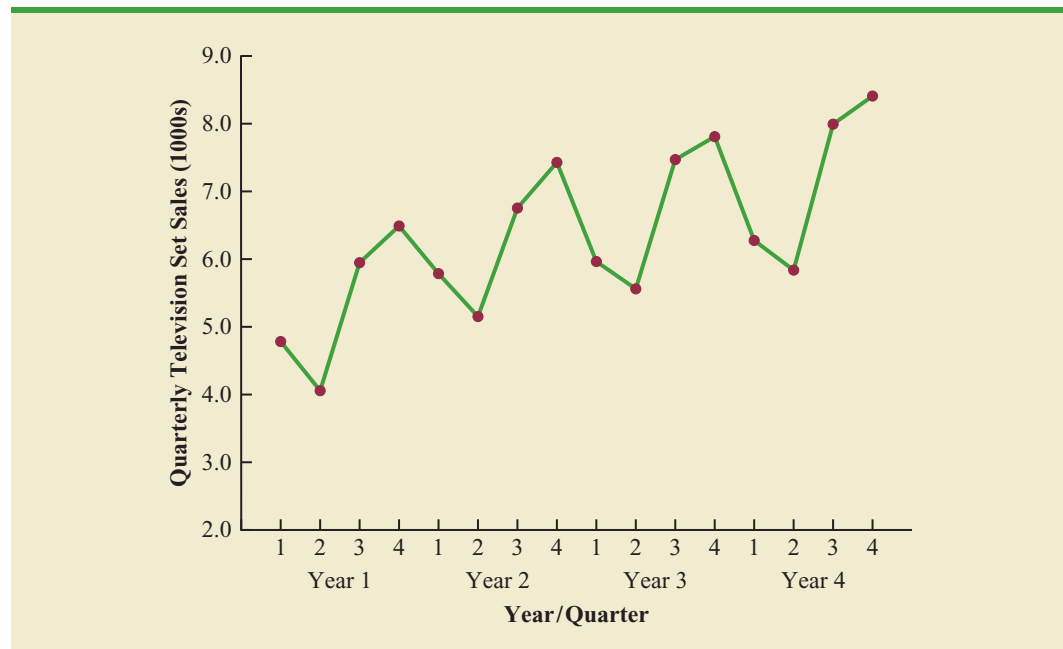
FIGURE 18.18   TELEVISION SET SALES TIME SERIES PLOT

TABLE 18.20 TELEVISION SET SALES TIME SERIES WITH DUMMY VARIABLES AND TIME PERIOD

Year	Quarter	Qtr1	Qtr2	Qtr3	Period	Sales (1000s)
1	1	1	0	0	1	4.8
	2	0	1	0	2	4.1
	3	0	0	1	3	6.0
	4	0	0	0	4	6.5
2	1	1	0	0	5	5.8
	2	0	1	0	6	5.2
	3	0	0	1	7	6.8
	4	0	0	0	8	7.4
3	1	1	0	0	9	6.0
	2	0	1	0	10	5.6
	3	0	0	1	11	7.5
	4	0	0	0	12	7.8
4	1	1	0	0	13	6.3
	2	0	1	0	14	5.9
	3	0	0	1	15	8.0
	4	0	0	0	16	8.4

Table 18.20 is the revised television set sales time series that includes the coded values of the dummy variables and the time period t. Using the data in Table 18.20, and Minitab's regression procedure, we obtained the computer output shown in Figure 18.19. The estimated multiple regression equation is

$$\text{Sales} = 6.07 - 1.36 \text{ Qtr1} - 2.03 \text{ Qtr2} - .304 \text{ Qtr3} + .146t \qquad \textbf{(18.12)}$$

We can now use equation (18.12) to forecast quarterly sales for next year. Next year is year 5 for the television set sales time series; that is, time periods 17, 18, 19, and 20.

Forecast for Time Period 17 (Quarter 1 in Year 5)

$$\text{Sales} = 6.07 - 1.36(1) - 2.03(0) - .304(0) + .146(17) = 7.19$$

Forecast for Time Period 18 (Quarter 2 in Year 5)

$$\text{Sales} = 6.07 - 1.36(0) - 2.03(1) - .304(0) + .146(18) = 6.67$$

FIGURE 18.19 MINITAB REGRESSION OUTPUT FOR THE UMBRELLA SALES TIME SERIES

```
The regression equation is
Sales (1000s) = 6.07 - 1.36 Qtr1 - 2.03 Qtr2 - 0.304
                    Qtr3 + 0.146 Period

Predictor      Coef   SE Coef       T      P
Constant     6.0688    0.1625   37.35  0.000
Qtr1        -1.3631    0.1575   -8.66  0.000
Qtr2        -2.0337    0.1551  -13.11  0.000
Qtr3        -0.3044    0.1537   -1.98  0.073
Period      0.14562   0.01211   12.02  0.000
```

Forecast for Time Period 19 (Quarter 3 in Year 5)

$$\text{Sales} = 6.07 - 1.36(0) - 2.03(0) - .304(1) + .146(19) = 8.54$$

Forecast for Time Period 20 (Quarter 4 in Year 5)

$$\text{Sales} = 6.07 - 1.36(0) - 2.03(0) - .304(0) + .146(20) = 8.99$$

Thus, accounting for the seasonal effects and the linear trend in television set sales, the estimates of quarterly sales in year 5 are 7190, 6670, 8540, and 8990.

The dummy variables in the estimated multiple regression equation actually provide four estimated multiple regression equations, one for each quarter. For instance, if time period t corresponds to quarter 1, the estimate of quarterly sales is

Quarter 1: Sales $= 6.07 - 1.36(1) - 2.03(0) - .304(0) + .146t = 4.71 + .146t$

Similarly, if time period t corresponds to quarters 2, 3, and 4, the estimates of quarterly sales are

Quarter 2: Sales $= 6.07 - 1.36(0) - 2.03(1) - .304(0) + .146t = 4.04 + .146t$
Quarter 3: Sales $= 6.07 - 1.36(0) - 2.03(0) - .304(1) + .146t = 5.77 + .146t$
Quarter 4: Sales $= 6.07 - 1.36(0) - 2.03(0) - .304(0) + .146t = 6.07 + .146t$

The slope of the trend line for each quarterly forecast equation is .146, indicating a growth in sales of about 146 sets per quarter. The only difference in the four equations is that they have different intercepts. For instance, the intercept for the quarter 1 equation is 4.71 and the intercept for the quarter 4 equation is 6.07. Thus, sales in quarter 1 are $4.71 - 6.07 = -1.36$ or 1360 sets less than in quarter 4. In other words, the estimated regression coefficient for Qtr1 in equation (18.12) provides an estimate of the difference in sales between quarter 1 and quarter 4. Similar interpretations can be provided for -2.03, the estimated regression coefficient for dummy variable Qtr2, and $-.304$, the estimated regression coefficient for dummy variable Qtr3.

Models Based on Monthly Data

Whenever a categorical variable such as season has k levels, k − 1 dummy variables are required.

In the preceding television set sales example, we showed how dummy variables can be used to account for the quarterly seasonal effects in the time series. Because there were 4 levels for the categorical variable season, 3 dummy variables were required. However, many businesses use monthly rather than quarterly forecasts. For monthly data, season is a categorical variable with 12 levels and thus $12 - 1 = 11$ dummy variables are required. For example, the 11 dummy variables could be coded as follows:

$$\text{Month1} = \begin{cases} 1 \text{ if January} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Month2} = \begin{cases} 1 \text{ if February} \\ 0 \text{ otherwise} \end{cases}$$

.

.

.

$$\text{Month11} = \begin{cases} 1 \text{ if November} \\ 0 \text{ otherwise} \end{cases}$$

Other than this change, the multiple regression approach for handling seasonality remains the same.

Exercises

Methods

SELF test

28. Consider the following time series.

Quarter	Year 1	Year 2	Year 3
1	71	68	62
2	49	41	51
3	58	60	53
4	78	81	72

  a. Construct a time series plot. What type of pattern exists in the data?
  b. Use the following dummy variables to develop an estimated regression equation to account for seasonal effects in the data: Qtr1 = 1 if Quarter 1, 0 otherwise; Qtr2 = 1 if Quarter 2, 0 otherwise; Qtr3 = 1 if Quarter 3, 0 otherwise.
  c. Compute the quarterly forecasts for next year.

29. Consider the following time series data.

Quarter	Year 1	Year 2	Year 3
1	4	6	7
2	2	3	6
3	3	5	6
4	5	7	8

  a. Construct a time series plot. What type of pattern exists in the data?
  b. Use the following dummy variables to develop an estimated regression equation to account for any seasonal and linear trend effects in the data: Qtr1 = 1 if Quarter 1, 0 otherwise; Qtr2 = 1 if Quarter 2, 0 otherwise; Qtr3 = 1 if Quarter 3, 0 otherwise.
  c. Compute the quarterly forecasts for next year.

Applications

30. The quarterly sales data (number of copies sold) for a college textbook over the past three years follow.

Quarter	Year 1	Year 2	Year 3
1	1690	1800	1850
2	940	900	1100
3	2625	2900	2930
4	2500	2360	2615

  a. Construct a time series plot. What type of pattern exists in the data?
  b. Use the following dummy variables to develop an estimated regression equation to account for any seasonal effects in the data: Qtr1 = 1 if Quarter 1, 0 otherwise; Qtr2 = 1 if Quarter 2, 0 otherwise; Qtr3 = 1 if Quarter 3, 0 otherwise.

c. Compute the quarterly forecasts for next year.

d. Let $t = 1$ to refer to the observation in quarter 1 of year 1; $t = 2$ to refer to the observation in quarter 2 of year 1; ... and $t = 12$ to refer to the observation in quarter 4 of year 3. Using the dummy variables defined in part (b) and t, develop an estimated regression equation to account for seasonal effects and any linear trend in the time series. Based upon the seasonal effects in the data and linear trend, compute the quarterly forecasts for next year.

31. Air pollution control specialists in southern California monitor the amount of ozone, carbon dioxide, and nitrogen dioxide in the air on an hourly basis. The hourly time series data exhibit seasonality, with the levels of pollutants showing patterns that vary over the hours in the day. On July 15, 16, and 17, the following levels of nitrogen dioxide were observed for the 12 hours from 6:00 A. M. to 6:00 P. M.

WEB file

Pollution

July 15:	25	28	35	50	60	60	40	35	30	25	25	20
July 16:	28	30	35	48	60	65	50	40	35	25	20	20
July 17:	35	42	45	70	72	75	60	45	40	25	25	25

a. Construct a time series plot. What type of pattern exists in the data?

b. Use the following dummy variables to develop an estimated regression equation to account for the seasonal effects in the data.

Hour1 $= 1$ if the reading was made between 6:00 A.M. and 7:00 A.M.; 0 otherwise

Hour2 $= 1$ if if the reading was made between 7:00 A.M. and 8:00 A.M.; 0 otherwise

.

.

.

Hour11 $= 1$ if the reading was made between 4:00 P.M. and 5:00 P.M., 0 otherwise.

Note that when the values of the 11 dummy variables are equal to 0, the observation corresponds to the 5:00 P.M. to 6:00 P.M. hour.

c. Using the estimated regression equation developed in part (a), compute estimates of the levels of nitrogen dioxide for July 18.

d. Let $t = 1$ to refer to the observation in hour 1 on July 15; $t = 2$ to refer to the observation in hour 2 of July 15; ... and $t = 36$ to refer to the observation in hour 12 of July 17. Using the dummy variables defined in part (b) and t, develop an estimated regression equation to account for seasonal effects and any linear trend in the time series. Based upon the seasonal effects in the data and linear trend, compute estimates of the levels of nitrogen dioxide for July 18.

32. South Shore Construction builds permanent docks and seawalls along the southern shore of Long Island, New York. Although the firm has been in business only five years, revenue has increased from $308,000 in the first year of operation to $1,084,000 in the most recent year. The following data show the quarterly sales revenue in thousands of dollars.

WEB file

SouthShore

Quarter	Year 1	Year 2	Year 3	Year 4	Year 5
1	20	37	75	92	176
2	100	136	155	202	282
3	175	245	326	384	445
4	13	26	48	82	181

a. Construct a time series plot. What type of pattern exists in the data?

b. Use the following dummy variables to develop an estimated regression equation to account for seasonal effects in the data. Qtr1 $= 1$ if Quarter 1, 0 otherwise; Qtr2 $= 1$

if Quarter 2, 0 otherwise; Qtr3 = 1 if Quarter 3, 0 otherwise. Based only on the seasonal effects in the data, compute estimates of quarterly sales for year 6.

c. Let Period = 1 to refer to the observation in quarter 1 of year 1; Period = 2 to refer to the observation in quarter 2 of year 1; . . . and Period = 20 to refer to the observation in quarter 4 of year 5. Using the dummy variables defined in part (b) and Period, develop an estimated regression equation to account for seasonal effects and any linear trend in the time series. Based upon the seasonal effects in the data and linear trend, compute estimates of quarterly sales for year 6.

33. Electric power consumption is measured in kilowatt-hours (kWh). The local utility company offers an interrupt program whereby commercial customers that participate receive favorable rates but must agree to cut back consumption if the utility requests them to do so. Timko Products has agreed to cut back consumption from noon to 8:00 P.M. on Thursday. To determine Timko's savings, the utility must estimate Timko's normal power usage for this period of time. Data on Timko's electric power consumption for the previous 72 hours are shown below.

WEB file

Power

Time Period	Monday	Tuesday	Wednesday	Thursday
12–4 A.M.	—	19,281	31,209	27,330
4–8 A.M	—	33,195	37,014	32,715
8–12 noon	—	99,516	119,968	152,465
12–4 P.M.	124,299	123,666	156,033	
4–8 P.M.	113,545	111,717	128,889	
8–12 midnight	41,300	48,112	73,923	

a. Construct a time series plot. What type of pattern exists in the data?
b. Use the following dummy variables to develop an estimated regression equation to account for any seasonal effects in the data.

Time1 = 1 for time period 12–4 A.M.; 0 otherwise
Time2 = 1 for time period 4–8 A.M; 0 otherwise
Time3 = 1 for time period 8–12 noon; 0 otherwise
Time4 = 1 for time period 12–4 P.M; 0 otherwise
Time5 = 1 for time period 4–8 P.M; 0 otherwise

c. Use the estimated regression equation developed in part (b) to estimate Timko's normal usage over the period of interrupted service.
d. Let Period = 1 to refer to the observation for Monday in the time period 12–4 P.M.; Period = 2 to refer to the observation for Monday in the time period 4–8 P.M; . . . and Period = 18 to refer to the observation for Thursday in the time period 8–12 noon. Using the dummy variables defined in part (b) and Period, develop an estimated regression equation to account for seasonal effects and any linear trend in the time series.
e. Using the estimated regression equation developed in part (d), estimate Timko's normal usage over the period of interrupted service.

34. Three years of monthly lawn-maintenance expenses ($) for a six-unit apartment house in southern Florida follow.

WEB file

AptExp

Month	Year 1	Year 2	Year 3
January	170	180	195
February	180	205	210
March	205	215	230
April	230	245	280
May	240	265	290

Month	Year 1	Year 2	Year 3
June	315	330	390
July	360	400	420
August	290	335	330
September	240	260	290
October	240	270	295
November	230	255	280
December	195	220	250

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop an estimated regression equation that can be used to account for any seasonal and linear trend effects in the data. Use the following dummy variables to account for the seasonal effects in the data: Jan = 1 if January, 0 otherwise; Feb = 1 if February, 0 otherwise; Mar = 1 if March, 0 otherwise; . . . Nov = 1 if November, 0 otherwise. Note that using this coding method, when all the 11 dummy variables are 0, the observation corresponds to an expense in December.
c. Compute the monthly forecasts for next year based upon both trend and seasonal effects.

18.6 Time Series Decomposition

In this section we turn our attention to what is called **time series decomposition**. Time series decomposition can be used to separate or decompose a time series into seasonal, trend, and irregular components. While this method can be used for forecasting, its primary applicability is to get a better understanding of the time series. Many business and economic time series are maintained and published by government agencies such as the Census Bureau and the Bureau of Labor Statistics. These agencies use time series decomposition to create deseasonalized time series.

Understanding what is really going on with a time series often depends upon the use of deseasonalized data. For instance, we might be interested in learning whether electrical power consumption is increasing in our area. Suppose we learn that electric power consumption in September is down 3% from the previous month. Care must be exercised in using such information, because whenever a seasonal influence is present, such comparisons may be misleading if the data have not been deseasonalized. The fact that electric power consumption is down 3% from August to September might be only the seasonal effect associated with a decrease in the use of air conditioning and not because of a long-term decline in the use of electric power. Indeed, after adjusting for the seasonal effect, we might even find that the use of electric power increased. Many other time series, such as unemployment statistics, home sales, and retail sales, are subject to strong seasonal influences. It is important to deseasonalize such data before making a judgment about any long-term trend.

Time series decomposition methods assume that Y_t, the actual time series value at period t, is a function of three components: a trend component; a seasonal component; and an irregular or error component. How these three components are combined to generate the observed values of the time series depends upon whether we assume the relationship is best described by an additive or a multiplicative model.

An **additive decomposition model** takes the following form:

$$Y_t = \text{Trend}_t + \text{Seasonal}_t + \text{Irregular}_t \qquad \textbf{(18.13)}$$

where

$$\text{Trend}_t = \text{trend value at time period } t$$
$$\text{Seasonal}_t = \text{seasonal value at time period } t$$
$$\text{Irregular}_t = \text{irregular value at time period } t$$

The irregular component corresponds to the error term ε in the simple linear regression model we discussed in Chapter 14.

In an additive model the values for the three components are simply added together to obtain the actual time series value Y_t. The irregular or error component accounts for the variability in the time series that cannot be explained by the trend and seasonal components.

An additive model is appropriate in situations where the seasonal fluctuations do not depend upon the level of the time series. The regression model for incorporating seasonal and trend effects in Section 18.5 is an additive model. If the sizes of the seasonal fluctuations in earlier time periods are about the same as the sizes of the seasonal fluctuations in later time periods, an additive model is appropriate. However, if the seasonal fluctuations change over time, growing larger as the sales volume increases because of a long-term linear trend, then a multiplicative model should be used. Many business and economic time series follow this pattern.

A **multiplicative decomposition model** takes the following form:

$$Y_t = \text{Trend}_t \times \text{Seasonal}_t \times \text{Irregular}_t \qquad \textbf{(18.14)}$$

where

$$\text{Trend}_t = \text{trend value at time period } t$$
$$\text{Seasonal}_t = \text{seasonal index at time period } t$$
$$\text{Irregular}_t = \text{irregular index at time period } t$$

The Census Bureau uses a multiplicative model in conjunction with its methodology for deseasonalizing time series.

In this model, the trend and seasonal and irregular components are multiplied to give the value of the time series. Trend is measured in units of the item being forecast. However, the seasonal and irregular components are measured in relative terms, with values above 1.00 indicating effects above the trend and values below 1.00 indicating effects below the trend.

Because this is the method most often used in practice, we will restrict our discussion of time series decomposition to showing how to develop estimates of the trend and seasonal components for a multiplicative model. As an illustration we will work with the quarterly television set sales time series introduced in Section 18.5; the quarterly sales data are shown in Table 18.19 and the corresponding time series plot is presented in Figure 18.18. After demonstrating how to decompose a time series using the multiplicative model, we will show how the seasonal indices and trend component can be recombined to develop a forecast.

Calculating the Seasonal Indexes

Figure 18.18 indicates that sales are lowest in the second quarter of each year and increase in quarters 3 and 4. Thus, we conclude that a seasonal pattern exists for the television set sales time series. The computational procedure used to identify each quarter's seasonal influence begins by computing a moving average to remove the combined seasonal and irregular effects from the data, leaving us with a time series that contains only trend and any remaining random variation not removed by the moving average calculations.

Because we are working with a quarterly series, we will use four data values in each moving average. The moving average calculation for the first four quarters of the television set sales data is

$$\text{First moving average} = \frac{4.8 + 4.1 + 6.0 + 6.5}{4} = \frac{21.4}{4} = 5.35$$

Note that the moving average calculation for the first four quarters yields the average quarterly sales over year 1 of the time series. Continuing the moving average calculations, we next add the 5.8 value for the first quarter of year 2 and drop the 4.8 for the first quarter of year 1. Thus, the second moving average is

$$\text{Second moving average} = \frac{4.1 + 6.0 + 6.5 + 5.8}{4} = \frac{22.4}{4} = 5.60$$

Similarly, the third moving average calculation is $(6.0 + 6.5 + 5.8 + 5.2)/4 = 5.875$.

Before we proceed with the moving average calculations for the entire time series, let us return to the first moving average calculation, which resulted in a value of 5.35. The 5.35 value is the average quarterly sales volume for year 1. As we look back at the calculation of the 5.35 value, associating 5.35 with the "middle" of the moving average group makes sense. Note, however, that with four quarters in the moving average, there is no middle period. The 5.35 value really corresponds to period 2.5, the last half of quarter 2 and the first half of quarter 3. Similarly, if we go to the next moving average value of 5.60, the middle period corresponds to period 3.5, the last half of quarter 3 and the first half of quarter 4.

The two moving average values we computed do not correspond directly to the original quarters of the time series. We can resolve this difficulty by computing the average of the two moving averages. Since the center of the first moving average is period 2.5 (half a period or quarter early) and the center of the second moving average is period 3.5 (half a period or quarter late), the average of the two moving averages is centered at quarter 3, exactly where it should be. This moving average is referred to as a *centered moving average*. Thus, the centered moving average for period 3 is $(5.35 + 5.60)/2 = 5.475$. Similarly, the centered moving average value for period 4 is $(5.60 + 5.875)/2 = 5.738$. Table 18.21 shows a complete summary of the moving average and centered moving average calculations for the television set sales data.

What do the centered moving averages in Table 18.21 tell us about this time series? Figure 18.20 shows a time series plot of the actual time series values and the centered moving average values. Note particularly how the centered moving average values tend to "smooth out" both the seasonal and irregular fluctuations in the time series. The centered moving averages represent the trend in the data and any random variation that was not removed by using moving averages to smooth the data.

Previously we showed that the multiplicative decomposition model is

$$Y_t = \text{Trend}_t \times \text{Seasonal}_t \times \text{Irregular}_t$$

By dividing each side of this equation by the trend component T_t, we can identify the combined seasonal-irregular effect in the time series.

The seasonal-irregular values are often referred to as the de-trended values of the time series.

$$\frac{Y_t}{\text{Trend}_t} = \frac{\text{Trend}_t \times \text{Seasonal}_t \times \text{Irregular}_t}{\text{Trend}_t} = \text{Seasonal}_t \times \text{Irregular}_t$$

For example, the third quarter of year 1 shows a trend value of 5.475 (the centered moving average). So $6.0/5.475 = 1.096$ is the combined seasonal-irregular value. Table 18.22 summarizes the seasonal-irregular values for the entire time series.

Consider the seasonal-irregular values for the third quarter: 1.096, 1.075, and 1.109. Seasonal-irregular values greater than 1.00 indicate effects above the trend estimate and values below 1.00 indicate effects below the trend estimate. Thus, the three seasonal-irregular values for quarter 3 show an above-average effect in the third quarter. Since the year-to-year fluctuations in the seasonal-irregular values are primarily due to random error, we can

TABLE 18.21 CENTERED MOVING AVERAGE CALCULATIONS FOR THE TELEVISION
SET SALES TIME SERIES

Year	Quarter	Sales (1000s)	Four-Quarter Moving Average	Centered Moving Average
1	1	4.8		
1	2	4.1		
			5.350	
1	3	6.0		5.475
			5.600	
1	4	6.5		5.738
			5.875	
2	1	5.8		5.975
			6.075	
2	2	5.2		6.188
			6.300	
2	3	6.8		6.325
			6.350	
2	4	7.4		6.400
			6.450	
3	1	6.0		6.538
			6.625	
3	2	5.6		6.675
			6.725	
3	3	7.5		6.763
			6.800	
3	4	7.8		6.838
			6.875	
4	1	6.3		6.938
			7.000	
4	2	5.9		7.075
			7.150	
4	3	8.0		
4	4	8.4		

average the computed values to eliminate the irregular influence and obtain an estimate of the third-quarter seasonal influence.

$$\text{Seasonal effect of quarter 3} = \frac{1.096 + 1.075 + 1.109}{3} = 1.09$$

We refer to 1.09 as the *seasonal index* for the third quarter. Table 18.23 summarizes the calculations involved in computing the seasonal indexes for the television set sales time series. The seasonal indexes for the four quarters are .93, .84, 1.09, and 1.14.

Interpretation of the seasonal indexes in Table 18.23 provides some insight about the seasonal component in television set sales. The best sales quarter is the fourth quarter, with sales averaging 14% above the trend estimate. The worst, or slowest, sales quarter is the second quarter; its seasonal index of .84 shows that the sales average is 16% below the trend estimate. The seasonal component corresponds clearly to the intuitive expectation

FIGURE 18.20 QUARTERLY TELEVISION SET SALES TIME SERIES AND CENTERED MOVING AVERAGE
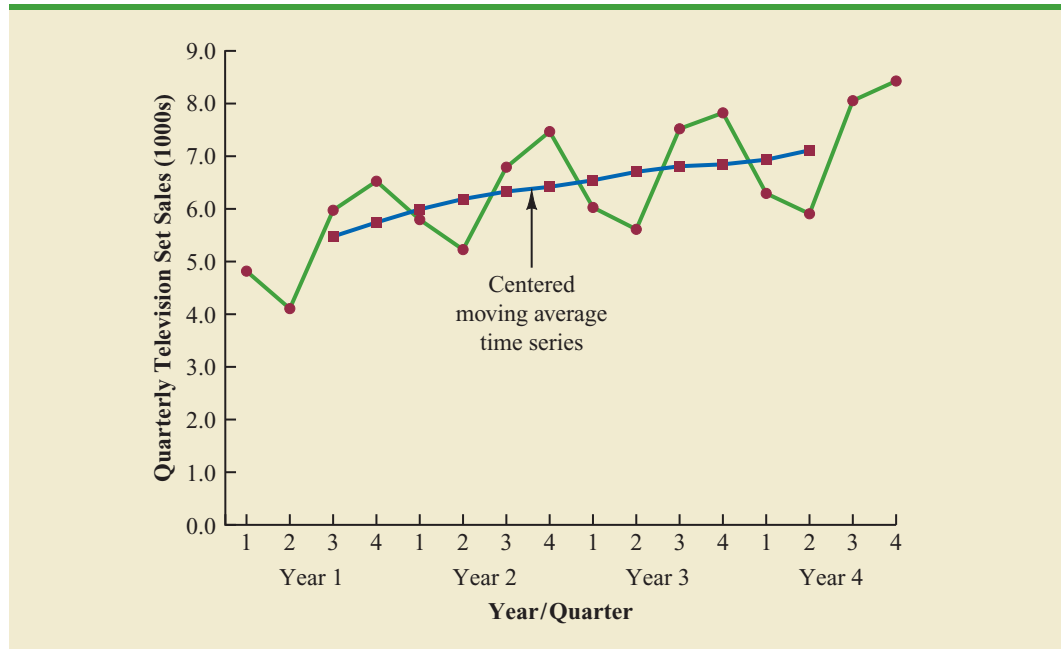
that television viewing interest and thus television purchase patterns tend to peak in the fourth quarter because of the coming winter season and reduction in outdoor activities. The low second-quarter sales reflect the reduced interest in television viewing due to the spring and presummer activities of potential customers.

TABLE 18.22 SEASONAL IRREGULAR VALUES FOR THE TELEVISION SET SALES TIME SERIES

Year	Quarter	Sales (1000s)	Centered Moving Average	Seasonal-Irregular Value
1	1	4.8		
1	2	4.1		
1	3	6.0	5.475	1.096
1	4	6.5	5.738	1.133
2	1	5.8	5.975	0.971
2	2	5.2	6.188	0.840
2	3	6.8	6.325	1.075
2	4	7.4	6.400	1.156
3	1	6.0	6.538	0.918
3	2	5.6	6.675	0.839
3	3	7.5	6.763	1.109
3	4	7.8	6.838	1.141
4	1	6.3	6.938	0.908
4	2	5.9	7.075	0.834
4	3	8.0		
4	4	8.4		

TABLE 18.23   SEASONAL INDEX CALCULATIONS FOR THE TELEVISION SET
SALES TIME SERIES

Quarter	Seasonal-Irregular Values			Seasonal Index
1	0.971	0.918	0.908	0.93
2	0.840	0.839	0.834	0.84
3	1.096	1.075	1.109	1.09
4	1.133	1.156	1.141	1.14

One final adjustment is sometimes necessary in obtaining the seasonal indexes. Because the multiplicative model requires that the average seasonal index equal 1.00, the sum of the four seasonal indexes in Table 18.23 must equal 4.00. In other words, the seasonal effects must even out over the year. The average of the seasonal indexes in our example is equal to 1.00, and hence this type of adjustment is not necessary. In other cases, a slight adjustment may be necessary. To make the adjustment, multiply each seasonal index by the number of seasons divided by the sum of the unadjusted seasonal indexes. For instance, for quarterly data, multiply each seasonal index by 4/(sum of the unadjusted seasonal indexes). Some of the exercises will require this adjustment to obtain the appropriate seasonal indexes.

Deseasonalizing the Time Series

Economic time series adjusted for seasonal variations are often reported in publications such as the Survey of Current Business, The Wall Street Journal, *and* BusinessWeek.

A time series that has had the seasonal effects removed is referred to as a **deseasonalized time series**, and the process of using the seasonal indexes to remove the seasonal effects from a time series is referred to as deseasonalizing the time series. Using a multiplicative decomposition model, we deseasonalize a time series by dividing each observation by its corresponding seasonal index. The multiplicative decomposition model is

$$Y_t = \text{Trend}_t \times \text{Seasonal}_t \times \text{Irregular}_t$$

So, when we divide each time series observation (Y_t) by its corresponding seasonal index, the resulting data show only trend and random variability (the irregular component). The deseasonalized time series for television set sales is summarized in Table 18.24. A graph of the deseasonalized time series is shown in Figure 18.21.

Using the Deseasonalized Time Series to Identify Trend

The graph of the deseasonalized television set sales time series shown in Figure 18.21 appears to have an upward linear trend. To identify this trend, we will fit a linear trend equation to the deseasonalized time series using the same method shown in Section 18.4. The only difference is that we will be fitting a trend line to the deseasonalized data instead of the original data.

Recall that for a linear trend the estimated regression equation can be written as

$$T_t = b_0 + b_1 t$$

where

$$T_t = \text{linear trend forecast in period } t$$
$$b_0 = \text{intercept of the linear trend line}$$
$$b_1 = \text{slope of the trend line}$$
$$t = \text{time period}$$

TABLE 18.24   DESEASONALIZED VALUES FOR THE TELEVISION SET SALES TIME SERIES

Year	Quarter	Time Period	Sales (1000s)	Seasonal Index	Deseasonalized Sales
1	1	1	4.8	0.93	5.16
	2	2	4.1	0.84	4.88
	3	3	6.0	1.09	5.50
	4	4	6.5	1.14	5.70
2	1	5	5.8	0.93	6.24
	2	6	5.2	0.84	6.19
	3	7	6.8	1.09	6.24
	4	8	7.4	1.14	6.49
3	1	9	6.0	0.93	6.45
	2	10	5.6	0.84	6.67
	3	11	7.5	1.09	6.88
	4	12	7.8	1.14	6.84
4	1	13	6.3	0.93	6.77
	2	14	5.9	0.84	7.02
	3	15	8.0	1.09	7.34
	4	16	8.4	1.14	7.37

FIGURE 18.21   DESEASONALIZED TELEVISION SET SALES TIME SERIES
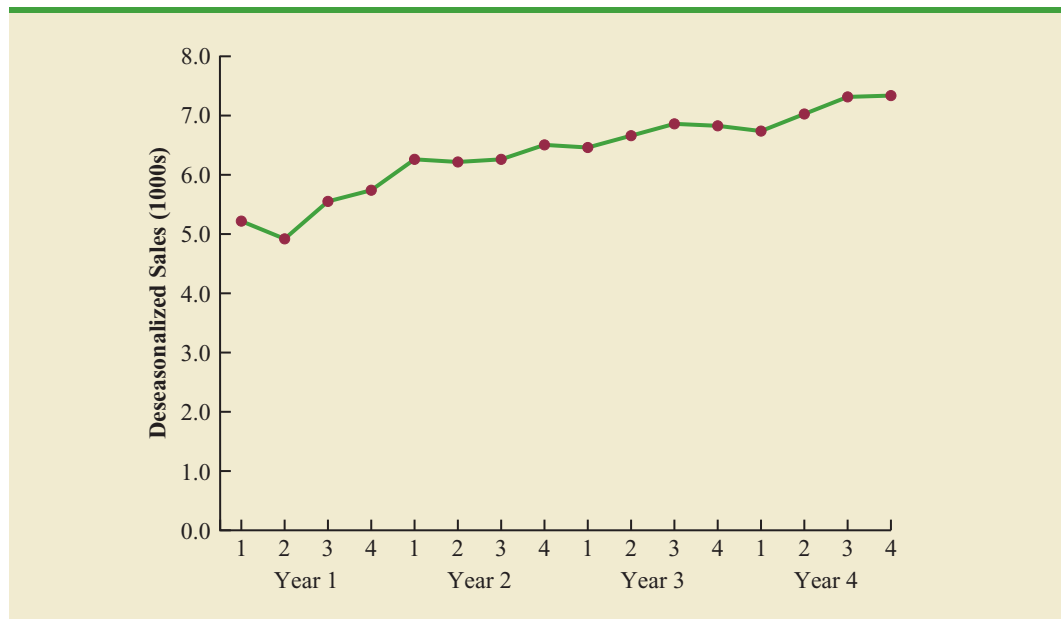
In Section 18.4 we provided formulas for computing the values of b_0 and b_1. To fit a linear trend line to the deseasonalized data in Table 18.24, the only change is that the deseasonalized time series values are used instead of the observed values Y_t in computing b_0 and b_1.

Figure 18.22 shows the computer output obtained using Minitab's regression analysis procedure to estimate the trend line for the deseasonalized television set time series. The estimated linear trend equation is

$$\text{Deseasonalized Sales} = 5.10 + 0.148\ t$$

FIGURE 18.22 MINITAB REGRESSION OUTPUT FOR THE DESEASONALIZED TELEVISION SET SALES TIME SERIES

```
The regression equation is
Deseasonalized Sales = 5.10 + 0.148 Period

Predictor      Coef  SE Coef       T      P
Constant     5.1050   0.1133   45.07  0.000
Period      0.14760  0.01171   12.60  0.000

S = 0.215985   R-Sq = 91.9%   R-Sq(adj) = 91.3%

Analysis of Variance

Source            DF      SS      MS       F      P
Regression         1  7.4068  7.4068  158.78  0.000
Residual Error    14  0.6531  0.0466
Total             15  8.0599
```

The slope of 0.148 indicates that over the past 16 quarters, the firm averaged a deseasonalized growth in sales of about 148 sets per quarter. If we assume that the past 16-quarter trend in sales data is a reasonably good indicator of the future, this equation can be used to develop a trend projection for future quarters. For example, substituting $t = 17$ into the equation yields next quarter's deseasonalized trend projection, T_{17}.

$$T_{17} = 5.10 + 0.148(17) = 7.616$$

Thus, using the deseasonalized data, the linear trend forecast for next quarter (period 17) is 7616 television sets. Similarly, the deseasonalized trend forecasts for the next three quarters (periods 18, 19, and 20) are 7764, 7912, and 8060 television sets, respectively.

Seasonal Adjustments

The final step in developing the forecast when both trend and seasonal components are present is to use the seasonal indexes to adjust the deseasonalized trend projections. Returning to the television set sales example, we have a deseasonalized trend projection for the next four quarters. Now we must adjust the forecast for the seasonal effect. The seasonal index for the first quarter of year 5 ($t = 17$) is 0.93, so we obtain the quarterly forecast by multiplying the deseasonalized forecast based on trend ($T_{17} = 7616$) by the seasonal index (0.93). Thus, the forecast for the next quarter is 7616(0.93) = 7083. Table 18.25 shows the quarterly forecast for quarters 17 through 20. The high-volume fourth quarter has a 9188-unit forecast, and the low-volume second quarter has a 6522-unit forecast.

TABLE 18.25 QUARTERLY FORECASTS FOR THE TELEVISION SET SALES TIME SERIES

Year	Quarter	Deseasonalized Trend Forecast	Seasonal Index	Quarterly Forecast
5	1	7616	0.93	(7616)(0.93) = 7083
	2	7764	0.84	(7764)(0.84) = 6522
	3	7912	1.09	(7912)(1.09) = 8624
	4	8060	1.14	(8060)(1.14) = 9188

Models Based on Monthly Data

In the preceding television set sales example, we used quarterly data to illustrate the computation of seasonal indexes. However, many businesses use monthly rather than quarterly forecasts. In such cases, the procedures introduced in this section can be applied with minor modifications. First, a 12-month moving average replaces the four-quarter moving average; second, 12 monthly seasonal indexes, rather than four quarterly seasonal indexes, must be computed. Other than these changes, the computational and forecasting procedures are identical.

Cyclical Component

Mathematically, the multiplicative model of equation (18.14) can be expanded to include a cyclical component.

$$Y_t = \text{Trend}_t \times \text{Cyclical}_t \times \text{Seasonal}_t \times \text{Irregular}_t \qquad \textbf{(18.15)}$$

The cyclical component, like the seasonal component, is expressed as a percentage of trend. As mentioned in Section 18.1, this component is attributable to multiyear cycles in the time series. It is analogous to the seasonal component, but over a longer period of time. However, because of the length of time involved, obtaining enough relevant data to estimate the cyclical component is often difficult. Another difficulty is that cycles usually vary in length. Because it is so difficult to identify and/or separate cyclical effects from long-term trend effects, in practice these effects are often combined and referred to as a combined trend-cycle component. We leave further discussion of the cyclical component to specialized texts on forecasting methods.

NOTES AND COMMENTS

1. There are a number of different approaches to computing the seasonal indexes. In this section each seasonal index was computed by averaging the corresponding seasonal-irregular values. Another approach, and the one used by Minitab, is to use the median of the seasonal-irregular values as the seasonal index.

2. Calendar adjustments are often made before deseasonalizing a time series. For example, if a time series consists of monthly sales values, the value for February sales may be less than for another month simply because there are fewer days in February. To account for this factor, we would first divide each month's sales value by the number of days in the month to obtain a daily average. Since the average number of days in a month is approximately $365/12 = 30.4167$, we then multiply the daily averages by 30.4167 to obtain adjusted monthly values. For the examples and exercises in this chapter, you can assume that any required calendar adjustments have already been made.

Exercises

Methods

SELFtest

35. Consider the following time series data.

Quarter	Year 1	Year 2	Year 3
1	4	6	7
2	2	3	6
3	3	5	6
4	5	7	8

    a.   Construct a time series plot. What type of pattern exists in the data?

    b.   Show the four-quarter and centered moving average values for this time series.

    c.   Compute seasonal indexes and adjusted seasonal indexes for the four quarters.

36. Refer to exercise 35.

    a.   Deseasonalize the time series using the adjusted seasonal indexes computed in part (c) of exercise 35.

    b.   Using Minitab or Excel, compute the linear trend regression equation for the deseasonalized data.

    c.   Compute the deseasonalized quarterly trend forecast for Year 4.

    d.   Use the seasonal indexes to adjust the deseasonalized trend forecasts computed in part (c).

Applications

37. The quarterly sales data (number of copies sold) for a college textbook over the past three years follow.

WEB file

TextSales

Quarter	Year 1	Year 2	Year 3
1	1690	1800	1850
2	940	900	1100
3	2625	2900	2930
4	2500	2360	2615

    a.   Construct a time series plot. What type of pattern exists in the data?

    b.   Show the four-quarter and centered moving average values for this time series.

    c.   Compute the seasonal and adjusted seasonal indexes for the four quarters.

    d.   When does the publisher have the largest seasonal index? Does this result appear reasonable? Explain.

    e.   Deseasonalize the time series.

    f.   Compute the linear trend equation for the deseasonalized data and forecast sales using the linear trend equation.

    g.   Adjust the linear trend forecasts using the adjusted seasonal indexes computed in part (c).

38. Three years of monthly lawn-maintenance expenses ($) for a six-unit apartment house in southern Florida follow.

WEB file

AptExp

Month	Year 1	Year 2	Year 3
January	170	180	195
February	180	205	210
March	205	215	230
April	230	245	280
May	240	265	290
June	315	330	390
July	360	400	420
August	290	335	330
September	240	260	290
October	240	270	295
November	230	255	280
December	195	220	250

a. Construct a time series plot. What type of pattern exists in the data?
b. Identify the monthly seasonal indexes for the three years of lawn-maintenance expenses for the apartment house in southern Florida as given here. Use a 12-month moving average calculation.
c. Deseasonalize the time series.
d. Compute the linear trend equation for the deseasonalized data.
e. Compute the deseasonalized trend forecasts and then adjust the trend forecasts using the seasonal indexes to provide a forecast for monthly expenses in year 4.

39. Air pollution control specialists in southern California monitor the amount of ozone, carbon dioxide, and nitrogen dioxide in the air on an hourly basis. The hourly time series data exhibit seasonality, with the levels of pollutants showing patterns over the hours in the day. On July 15, 16, and 17, the following levels of nitrogen dioxide were observed in the downtown area for the 12 hours from 6:00 A.M. to 6:00 P.M.

WEB file

Pollution

July 15:	25	28	35	50	60	60	40	35	30	25	25	20
July 16:	28	30	35	48	60	65	50	40	35	25	20	20
July 17:	35	42	45	70	72	75	60	45	40	25	25	25

a. Construct a time series plot. What type of pattern exists in the data?
b. Identify the hourly seasonal indexes for the 12 readings each day.
c. Deseasonalize the time series.
d. Using Minitab or Excel, compute the linear trend equation for the deseasonalized data.
e. Compute the deseasonalized trend forecasts for the 12 hours for July 18 and then adjust the trend forecasts using the seasonal indexes computed in part (b).

40. Electric power consumption is measured in kilowatt-hours (kWh). The local utility company offers an interrupt program whereby commercial customers that participate receive favorable rates but must agree to cut back consumption if the utility requests them to do so. Timko Products cut back consumption at 12:00 noon Thursday. To assess the savings, the utility must estimate Timko's usage without the interrupt. The period of interrupted service was from noon to 8:00 P.M. Data on electric power consumption for the previous 72 hours are available.

WEB file

Power

Time Period	Monday	Tuesday	Wednesday	Thursday
12–4 A.M.	—	19,281	31,209	27,330
4–8 A.M	—	33,195	37,014	32,715
8–12 noon	—	99,516	119,968	152,465
12–4 P.M.	124,299	123,666	156,033	
4–8 P.M.	113,545	111,717	128,889	
8–12 midnight	41,300	48,112	73,923	

a. Is there a seasonal effect over the 24-hour period?
b. Compute seasonal indexes for the six 4-hour periods.
c. Use trend adjusted for seasonal indexes to estimate Timko's normal usage over the period of interrupted service.

Summary

This chapter provided an introduction to the basic methods of time series analysis and forecasting. First, we showed that the underlying pattern in the time series can often be identified by constructing a time series plot. Several types of data patterns can be distinguished,

including a horizontal pattern, a trend pattern, and a seasonal pattern. The forecasting methods we have discussed are based on which of these patterns are present in the time series.

For a time series with a horizontal pattern, we showed how moving averages and exponential smoothing can be used to develop a forecast. The moving averages method consists of computing an average of past data values and then using that average as the forecast for the next period. In the exponential smoothing method, a weighted average of past time series values is used to compute a forecast. These methods also adapt well when a horizontal pattern shifts to a different level and resumes a horizontal pattern.

An important factor in determining what forecasting method to use involves the accuracy of the method. We discussed three measures of forecast accuracy: mean absolute error (MAE), mean squared error (MSE), and mean absolute percentage error (MAPE). Each of these measures is designed to determine how well a particular forecasting method is able to reproduce the time series data that are already available. By selecting a method that has the best accuracy for the data already known, we hope to increase the likelihood that we will obtain better forecasts for future time periods.

For time series that have only a long-term linear trend, we showed how simple time series regression can be used to make trend projections. We also discussed how an extension of single exponential smoothing, referred to as Holt's linear exponential smoothing, can be used to forecast a time series with a linear trend. For a time series with a curvilinear or non-linear trend, we showed how multiple regression can be used to fit a quadratic trend equation or an exponential trend equation to the data.

For a time series with a seasonal pattern, we showed how the use of dummy variables in a multiple regression model can be used to develop an estimated regression equation with seasonal effects. We then extended the regression approach to include situations where the time series contains both a seasonal and a linear trend effect by showing how to combine the dummy variable approach for handling seasonality with the time series regression approach for handling linear trend.

In the last section of the chapter we showed how time series decomposition can be used to separate or decompose a time series into seasonal and trend components and then to deseasonalize the time series. We showed how to compute seasonal indexes for a multiplicative model, how to use the seasonal indexes to deseasonalize the time series, and how to use regression analysis on the deseasonalized data to estimate the trend component. The final step in developing a forecast when both trend and seasonal components are present is to use the seasonal indexes to adjust the trend projections.

Glossary

Time series A sequence of observations on a variable measured at successive points in time or over successive periods of time.

Time series plot A graphical presentation of the relationship between time and the time series variable. Time is shown on the horizontal axis and the time series values are shown on the verical axis.

Horizontal pattern A horizontal pattern exists when the data fluctuate around a constant mean.

Stationary time series A time series whose statistical properties are indepepndent of time. For a stationary time series the process generating the data has a constant mean and the variability of the time series is constant over time.

Trend pattern A trend pattern exists if the time series plot shows gradual shifts or movements to relatively higher or lower values over a longer period of time.

Seasonal pattern A seasonal pattern exists if the time series plot exhibits a repeating pattern over successive periods. The successive periods are often one-year intervals, which is where the name seasonal pattern comes from.

Cyclical pattern A cyclical pattern exists if the time series plot shows an alternating sequence of points below and above the trend line lasting more than one year.

Forecast error The difference between the actual time series value and the forecast.

Mean absolute error (MAE) The average of the absolute values of the forecast errors.

Mean squared error (MSE) The average of the sum of squared forecast errors.

Mean absolute percentage error (MAPE) The average of the absolute values of the percentage forecast errors.

Moving averages A forecasting method that uses the average of the most recent k data values in the time series as the forecast for the next period.

Weighted moving averages A forecasting method that involves selecting a different weight for the most recent k data values values in the time series and then computing a weighted average of the values. The sum of the weights must equal one.

Exponential smoothing A forecasting method that uses a weighted average of past time series values as the forecast; it is a special case of the weighted moving averages method in which we select only one weight—the weight for the most recent observation.

Smoothing constant A parameter of the exponential smoothing model that provides the weight given to the most recent time series value in the calculation of the forecast value.

Linear exponential smoothing An extension of single exponential smoothing that uses two smoothing constants to enable forecasts to be developed for a time series with a linear trend.

Time series decompostition A time series method that is used to separate or decompose a time series into seasonal and trend components.

Additive model In an additive model the actual time series value at time period t is obtained by adding the values of a trend component, a seasonal component, and an irregular component.

Multiplicative model In a multiplicative model the actual time series value at time period t is obtained by multiplying the values of a trend component, a seasonal component, and an irregular component.

Deseasonalized time series A time series from which the effect of season has been removed by dividing each original time series observation by the corresponding seasonal index.

Key Formulas

Moving Average Forecast of Order k

$$F_{t+1} = \frac{\sum (\text{most recent } k \text{ data values})}{k} \tag{18.1}$$

Exponential Smoothing Forecast

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t \tag{18.2}$$

Linear Trend Equation

$$T_t = b_0 + b_1 t \tag{18.4}$$

where

$$b_1 = \frac{\sum_{t=1}^{n} (t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n} (t - \bar{t})^2} \tag{18.5}$$

$$b_0 = \bar{Y} - b_1 \bar{t} \tag{18.6}$$

Holt's Linear Exponential Smoothing

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \tag{18.7}$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)\, b_{t-1} \tag{18.8}$$

$$F_{t+k} = L_t + b_t\, k \tag{18.9}$$

Quadratic Trend Equation

$$T_t = b_0 + b_1 t + b_2 t^2 \tag{18.10}$$

Exponential Trend Equation

$$T_t = b_0 (b_1)^t \tag{18.11}$$

Additive Decomposition Model

$$Y_t = \text{Trend}_t + \text{Seasonal}_t + \text{Irregular}_t \tag{18.13}$$

Multiplicative Decomposition Model

$$Y_t = \text{Trend}_t \times \text{Seasonal}_t \times \text{Irregular}_t \tag{18.14}$$

Supplementary Exercises

41. The weekly demand (in cases) for a particular brand of automatic dishwasher detergent for a chain of grocery stores located in Columbus, Ohio, follows.

Week	Demand	Week	Demand
1	22	6	24
2	18	7	20
3	23	8	19
4	21	9	18
5	17	10	21

   a. Construct a time series plot. What type of pattern exists in the data?
   b. Use a three-week moving average to develop a forecast for week 11.
   c. Use exponential smoothing with a smoothing constant of $\alpha = .2$ to develop a forecast for week 11.
   d. Which of the two methods do you prefer? Why?

42. The following table reports the percentage of stocks in a portfolio for nine quarters from 2007 to 2009.

Quarter	Stock %
1st—2007	29.8
2nd—2007	31.0
3rd—2007	29.9
4th—2007	30.1
1st—2008	32.2
2nd—2008	31.5
3rd—2008	32.0
4th—2008	31.9
1st—2009	30.0

a. Construct a time series plot. What type of pattern exists in the data?
b. Use exponential smoothing to forecast this time series. Consider smoothing constants of $\alpha = .2, .3$, and .4. What value of the smoothing constant provides the most accurate forecasts?
c. What is the forecast of the percentage of stocks in a typical portfolio for the second quarter of 2009?

43. United Dairies, Inc., supplies milk to several independent grocers throughout Dade County, Florida. Managers at United Dairies want to develop a forecast of the number of half-gallons of milk sold per week. Sales data for the past 12 weeks follow.

Week	Sales	Week	Sales
1	2750	7	3300
2	3100	8	3100
3	3250	9	2950
4	2800	10	3000
5	2900	11	3200
6	3050	12	3150

a. Construct a time series plot. What type of pattern exists in the data?
b. Use exponential smoothing withf $\alpha = .4$ to develop a forecast of demand for week 13.

44. To avoid a monthly service fee in an interest-bearing checking account, customers must maintain a minimum average daily balance. Bankrate's 2008 survey of 249 banks and thrifts in the top 25 metropolitan areas showed that you need to maintain an average balance of $3,462 to avoid a monthly service fee. With an average fee of $11.97 and an average interest rate of only 0.24 percent, customers with interest-bearing checking accounts are not getting much value for basically providing the bank with a line of credit equal to the average monthly balance required to avoid the monthly service fee (Bankrate website, October 27, 2008). The following table shows the minimum average balance required to avoid paying a monthly service fee from 2001–2008.

Year	Balance ($)
2001	2435
2002	2593
2003	2258
2004	2087
2005	2294
2006	2660
2007	3317
2008	3462

a. Construct a time series plot. What type of pattern exists in the data?
b. Using Minitab or Excel, develop a linear trend equation for the time series. Compute an estimate of the average balance required to avoid a monthly service fee for 2009.
c. Using Minitab or Excel, develop a quadratic trend equation for the time series. Compute an estimate of the average balance required to avoid a monthly service fee for 2009.
d. Using MSE, which approach provides the most accurate forecasts for the historical data?
e. For these data would you recommend that the forecast for 2009 be developed using the linear trend equation or the quadratic trend equation? Explain.

45. The Garden Avenue Seven sells CDs of its musical performances. The following table reports sales (in units) for the past 18 months. The group's manager wants an accurate method for forecasting future sales.

Month	Sales	Month	Sales	Month	Sales
1	293	7	381	13	549
2	283	8	431	14	544
3	322	9	424	15	601
4	355	10	433	16	587
5	346	11	470	17	644
6	379	12	481	18	660

WEB file

CDSales

a. Construct a time series plot. What type of pattern exists in the data?
b. Use exponential smoothing with $\alpha = .3, .4,$ and $.5$. Which value of α provides the most accurate forecasts?
c. Use trend projection to provide a forecast. What is the value of MSE?
d. Which method of forecasting would you recommend to the manager? Why?

46. The Mayfair Department Store in Davenport, Iowa, is trying to determine the amount of sales lost while it was shut down during July and August because of damage caused by the Mississippi River flood. Sales data for January through June follow.

Month	Sales ($1000s)	Month	Sales ($1000s)
January	185.72	April	210.36
February	167.84	May	255.57
March	205.11	June	261.19

a. Use exponential smoothing, with $\alpha = .4$, to develop a forecast for July and August. (*Hint*: Use the forecast for July as the actual sales in July in developing the August forecast.) Comment on the use of exponential smoothing for forecasts more than one period into the future.
b. Use trend projection to forecast sales for July and August.
c. Mayfair's insurance company proposed a settlement based on lost sales of $240,000 in July and August. Is this amount fair? If not, what amount would you recommend as a counteroffer?

47. Canton Supplies, Inc., is a service firm that employs approximately 100 individuals. Managers of Canton Supplies are concerned about meeting monthly cash obligations and want to develop a forecast of monthly cash requirements. Because of a recent change in operating policy, only the past seven months of data that follow are considered to be relevant.

Month	1	2	3	4	5	6	7
Cash Required ($1000s)	205	212	218	224	230	240	246

a. Construct a time series plot. What type of pattern exists in the data?
b. Use Holt's linear exponential smoothing with $\alpha = .6$ and $\beta = .4$ to forecast cash requirements for each of the next two months.
c. Using Minitab or Excel, develop a linear trend equation to forecast cash requirements for each of the next two months.
d. Would you recommend using Holt's linear exponential smoothing with $\alpha = .6$ and $\beta = .4$ to forecast cash requirements for each of the next two months or the linear trend equation? Explain.

48. The Costello Music Company has been in business for five years. During that time, sales of pianos increased from 12 units in the first year to 76 units in the most recent year. Fred Costello, the firm's owner, wants to develop a forecast of piano sales for the coming year. The historical data follow.

Year	1	2	3	4	5
Sales	12	28	34	50	76

    a. Construct a time series plot. What type of pattern exists in the data?

    b. Develop the linear trend equation for the time series. What is the average increase in sales that the firm has been realizing per year?

    c. Forecast sales for years 6 and 7.

49. Consider the Costello Music Company problem in exercise 48. The quarterly sales data follow.

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total Yearly Sales
1	4	2	1	5	12
2	6	4	4	14	28
3	10	3	5	16	34
4	12	9	7	22	50
5	18	10	13	35	76

    a. Use the following dummy variables to develop an estimated regression equation to account for any seasonal and linear trend effects in the data: Qtr1 = 1 if Quarter 1, 0 otherwise; Qtr2 = 1 if Quarter 2, 0 otherwise; and Qtr3 = 1 if Quarter 3, 0 otherwise.

    b. Compute the quarterly forecasts for next year.

50. Refer to the Costello Music Company problem in exercise 49.

    a. Using time series decomposition, compute the seasonal indexes for the four quarters.

    b. When does Costello Music experience the largest seasonal effect? Does this result appear reasonable? Explain.

51. Refer to the Costello Music Company time series in exercise 49.

    a. Deseasonalize the data and use the deseasonalized time series to identify the trend.

    b. Use the results of part (a) to develop a quarterly forecast for next year based on trend.

    c. Use the seasonal indexes developed in exercise 50 to adjust the forecasts developed in part (b) to account for the effect of season.

52. Hudson Marine has been an authorized dealer for C&D marine radios for the past seven years. The following table reports the number of radios sold each year.

Year	1	2	3	4	5	6	7
Number Sold	35	50	75	90	105	110	130

    a. Construct a time series plot. Does a linear trend appear to be present?

    b. Using Minitab or Excel, develop a linear trend equation for this time series.

    c. Use the linear trend equation developed in part (b) to develop a forecast for annual sales in year 8.

53. Refer to the Hudson Marine problem in exercise 52. Suppose the quarterly sales values for the seven years of historical data are as follows.

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total Yearly Sales
1	6	15	10	4	35
2	10	18	15	7	50
3	14	26	23	12	75
4	19	28	25	18	90
5	22	34	28	21	105
6	24	36	30	20	110
7	28	40	35	27	130

a. Use the following dummy variables to develop an estimated regression equation to account for any season and linear trend effects in the data: Qtr1 $=$ 1 if Quarter 1, 0 otherwise; Qtr2 $=$ 1 if Quarter 2, 0 otherwise; and Qtr3 $=$ 1 if Quarter 3, 0 otherwise.
b. Compute the quarterly forecasts for next year.

54. Refer to the Hudson Marine problem in exercise 53.
a. Compute the centered moving average values for this time series.
b. Construct a time series plot that also shows the centered moving average and original time series on the same graph. Discuss the differences between the original time series plot and the centered moving average time series.
c. Compute the seasonal indexes for the four quarters.
d. When does Hudson Marine experience the largest seasonal effect? Does this result seem reasonable? Explain.

55. Refer to the Hudson Marine data in exercise 53.
a. Deseasonalize the data and use the deseasonalized time series to identify the trend.
b. Use the results of part (a) to develop a quarterly forecast for next year based on trend.
c. Use the seasonal indexes developed in exercise 54 to adjust the forecasts developed in part (b) to account for the effect of season.

Case Problem 1   Forecasting Food and Beverage Sales

The Vintage Restaurant, on Captiva Island near Fort Myers, Florida, is owned and operated by Karen Payne. The restaurant just completed its third year of operation. Since opening her restaurant, Karen has sought to establish a reputation for the Vintage as a high-quality dining establishment that specializes in fresh seafood. Through the efforts of Karen and her staff, her restaurant has become one of the best and fastest growing restaurants on the island.

To better plan for future growth of the restaurant, Karen needs to develop a system that will enable her to forecast food and beverage sales by month for up to one year in advance. Table 18.26 shows the value of food and beverage sales ($1000s) for the first three years of operation.

Managerial Report

Perform an analysis of the sales data for the Vintage Restaurant. Prepare a report for Karen that summarizes your findings, forecasts, and recommendations. Include the following:

1. A time series plot. Comment on the underlying pattern in the time series.
2. An analysis of the seasonality of the data. Indicate the seasonal indexes for each month, and comment on the high and low seasonal sales months. Do the seasonal indexes make intuitive sense? Discuss.
3. Deseasonalize the time series. Does there appear to be any trend in the deseasonalized time series?
4. Using the time series decomposition method, forecast sales for January through December of the fourth year.
5. Using the dummy variable regression approach, forecast sales for January through December of the fourth year.
6. Provide summary tables of your calculations and any graphs in the appendix of your report.

Assume that January sales for the fourth year turn out to be $295,000. What was your forecast error? If this error is large, Karen may be puzzled about the difference between your forecast and the actual sales value. What can you do to resolve her uncertainty in the forecasting procedure?

TABLE 18.26   FOOD AND BEVERAGE SALES FOR THE VINTAGE
                  RESTAURANT ($1000s)

Month	First Year	Second Year	Third Year
January	242	263	282
February	235	238	255
March	232	247	265
April	178	193	205
May	184	193	210
June	140	149	160
July	145	157	166
August	152	161	174
September	110	122	126
October	130	130	148
November	152	167	173
December	206	230	235

WEB file

Vintage

Case Problem 2   Forecasting Lost Sales

The Carlson Department Store suffered heavy damage when a hurricane struck on August 31.
The store was closed for four months (September through December), and Carlson is now
involved in a dispute with its insurance company about the amount of lost sales during the
time the store was closed. Two key issues must be resolved: (1) the amount of sales Carlson
would have made if the hurricane had not struck and (2) whether Carlson is entitled to any
compensation for excess sales due to increased business activity after the storm. More than
$8 billion in federal disaster relief and insurance money came into the county, resulting in
increased sales at department stores and numerous other businesses.

   Table 18.27 gives Carlson's sales data for the 48 months preceding the storm. Table 18.28
reports total sales for the 48 months preceding the storm for all department stores in the county,
as well as the total sales in the county for the four months the Carlson Department Store was
closed. Carlson's managers asked you to analyze these data and develop estimates of the lost
sales at the Carlson Department Store for the months of September through December. They
also asked you to determine whether a case can be made for excess storm-related sales during

TABLE 18.27   SALES FOR CARLSON DEPARTMENT STORE ($MILLIONS)

Month	Year 1	Year 2	Year 3	Year 4	Year 5
January		1.45	2.31	2.31	2.56
February		1.80	1.89	1.99	2.28
March		2.03	2.02	2.42	2.69
April		1.99	2.23	2.45	2.48
May		2.32	2.39	2.57	2.73
June		2.20	2.14	2.42	2.37
July		2.13	2.27	2.40	2.31
August		2.43	2.21	2.50	2.23
September	1.71	1.90	1.89	2.09	
October	1.90	2.13	2.29	2.54	
November	2.74	2.56	2.83	2.97	
December	4.20	4.16	4.04	4.35	

WEB file

CarlsonSales

TABLE 18.28   DEPARTMENT STORE SALES FOR THE COUNTY ($MILLIONS)

Month	Year 1	Year 2	Year 3	Year 4	Year 5
January		46.80	46.80	43.80	48.00
February		48.00	48.60	45.60	51.60
March		60.00	59.40	57.60	57.60
April		57.60	58.20	53.40	58.20
May		61.80	60.60	56.40	60.00
June		58.20	55.20	52.80	57.00
July		56.40	51.00	54.00	57.60
August		63.00	58.80	60.60	61.80
September	55.80	57.60	49.80	47.40	69.00
October	56.40	53.40	54.60	54.60	75.00
November	71.40	71.40	65.40	67.80	85.20
December	117.60	114.00	102.00	100.20	121.80

WEB file

CountySales

the same period. If such a case can be made, Carlson is entitled to compensation for excess sales it would have earned in addition to ordinary sales.

Managerial Report

Prepare a report for the managers of the Carlson Department Store that summarizes your findings, forecasts, and recommendations. Include the following:

1. An estimate of sales for Carlson Department Store had there been no hurricane.
2. An estimate of countywide department store sales had there been no hurricane.
3. An estimate of lost sales for the Carlson Department Store for September through December.

In addition, use the countywide actual department stores sales for September through December and the estimate in part (2) to make a case for or against excess storm-related sales.

Appendix 18.1   Forecasting with Minitab

In this appendix we show how Minitab can be used to develop forecasts using the following forecasting methods: moving averages, exponential smoothing, trend projection, Holt's linear exponential smoothing, and time series decomposition.

Moving Averages

WEB file

Gasoline

To show how Minitab can be used to develop forecasts using the moving averages method, we will develop a forecast for the gasoline sales time series in Table 18.1 and Figure 18.1. The sales data for the 12 weeks are entered into column 2 of the worksheet. The following steps can be used to produce a three-week moving average forecast for week 13.

  Step 1. Select the **Stat** menu
  Step 2. Choose **Time Series**
  Step 3. Choose **Moving Average**
  Step 4. When the Moving Average dialog box appears:
        Enter C2 in the **Variable** box
        Enter 3 in the **MA length** box

> Select **Generate forecasts**
> Enter 1 in the **Number of forecasts** box
> Enter 12 in the **Starting from origin** box
> Click **OK**

Measures of forecast accuracy and the forecast for week 13 are shown in the session window. The mean absolute error is labeled MAD and the mean squared error is labeled MSD in the Minitab output.

Exponential Smoothing

WEB file

Gasoline

To show how Minitab can be used to develop an exponential smoothing forecast, we will again develop a forecast of sales in week 13 for the gasoline sales time series in Table 18.1 and Figure 18.1. The sales data for the 12 weeks are entered into column 2 of the worksheet. The following steps can be used to produce a forecast for week 13 using a smoothing constant of $\alpha = .2$.

> **Step 1.** Select the **Stat** menu
> **Step 2.** Choose **Time Series**
> **Step 3.** Choose **Single Exp Smoothing**
> **Step 4.** When the Single Exponential Smoothing dialog box appears:
>> Enter C2 in the **Variable** box
>> Select the **Use** option for the Weight to Use in Smoothing
>> Enter 0.2 in the Use box
>> Select **Generate forecasts**
>> Enter 1 in the **Number of forecasts** box
>> Enter 12 in the **Starting from origin** box
>> Select **Options**
> **Step 5.** When the Single Exponential Smoothing-Options dialog box appears:
>> Enter 1 in the **Use average of first K observations** box
>> Click **OK**
> **Step 6.** When the Single Exponential Smoothing dialog box appears:
>> Click **OK**

Measures of forecast accuracy and the exponential smoothing forecast for week 13 are shown in the session window. The mean absolute error is labeled MAD and the mean squared error is labeled MSD in the Minitab output.*

Trend Projection

WEB file

Bicycle

To show how Minitab can be used for trend projection, we develop a forecast for the bicycle sales time series in Table 18.3 and Figure 18.3. The year numbers are entered into column 1 and the sales data are entered into column 2 of the worksheet. The following steps can be used to produce a forecast for year 11 using trend projection.

> **Step 1.** Select the **Stat** menu
> **Step 2.** Choose **Time Series**
> **Step 3.** Choose **Trend Analysis**

---

*The value of MSD computed by Minitab is not the same as the value of MSE that we computed in Section 18.3. Minitab uses a forecast of 17 for week 1 and computes MSD using all 12 weeks of data. In Section 18.3 we compute MSE using only the data for weeks 2 through 12, because we had no past values with which to make a forecast for week 1.

Step 4. When the Trend Analysis dialog box appears:

>   Enter C2 in the **Variable** box
>   Choose **Linear** for the Model Type
>   Select **Generate forecasts**
>   Enter 1 in the **Number of forecasts** box
>   Enter 10 in the **Starting from origin** box
>   Click **OK**

The equation for linear trend, measures of forecast accuracy, and the forecast for the next year are shown in the session window. The mean absolute error is labeled MAD and the mean square error is labeled MSD in the Minitab output. To generate forecasts for a quadratic or exponential trend select **Quadratic** of **Exponential growth** instead of **Linear** in Step 4.

Holt's Linear Exponential Smoothing

WEB file

Bicycle

To show how Minitab can be used to develop forecasts using Holt's linear exponential smoothing method, we develop a forecast for the bicycle sales time series in Table 18.3 and Figure 18.3. In Minitab, Holt's linear exponential smoothing method is referred to as Double Exponential Smoothing. The year numbers are entered into column 1 and the sales data are entered into column 2 of the worksheet. The following steps can be used to forecast sales in year 11 using Holt's linear exponential smoothing with $\alpha = .1$ and $\beta = .2$.

Step 1. Select the **Stat** menu
Step 2. Choose **Time Series**
Step 3. Choose **Double Exp Smoothing**
Step 4. When the Double Exponential Smoothing dialog box appears:

>   Enter C2 in the **Variable** box
>   Select the **Use** option for the Weights to Use in Smoothing
>   Enter .1 in the **level** box
>   Enter .2 in the **trend** box
>   Select **Generate forecasts**
>   Enter 1 in the **Number of forecasts** box
>   Enter 10 in the **Starting from origin** box
>   Click **OK**

Measures of forecast accuracy and Holt's linear exponential smoothing forecast for year 11 are shown in the session window. The mean absolute error is labeled MAD and the mean square error is labeled MSD in the Minitab output.

Time Series Decomposition

WEB file

TVSales

To show how Minitab can be used to forecast a time series with trend and seasonality using time series decomposition, we develop a forecast for the television set sales time series in Table 18.6 and Figure 18.6. In Minitab, the user has the option of either a multiplicative or additive decomposition model. We will illustrate how to use the multiplicative approach as described in section 18.6. The year numbers are entered into column 1, the quarter values are entered into column 2, and the sales data are entered into column 3 of the worksheet. The following steps can be used to produce a forecast for the next quarter.

Step 1. Select the **Stat** menu
Step 2. Choose **Time Series**
Step 3. Choose **Decomposition**
Step 4. When the Decomposition dialog box appears:

>   Enter C3 in the **Variable** box
>   Enter 4 in the **Season Length** box

Select **Multiplicative** for Method Type
Select **Trend plus Seasonal** for Model Components
Select **Generate forecasts**
Enter 1 in the **Number of forecasts** box
Enter 16 in the **Starting from origin** box
Click **OK**

The seasonal indexes,[†] measures of forecast accuracy, and the forecast for the next quarter are shown in the session window. The mean absolute error is labeled MAD and the mean square error is labeled MSD in the Minitab output.

Appendix 18.2   Forecasting with Excel

In this appendix we show how Excel can be used to develop forecasts using three forecasting methods: moving averages, exponential smoothing, and trend projection.

Moving Averages

WEB file
Gasoline

To show how Excel can be used to develop forecasts using the moving averages method, we will develop a forecast for the gasoline sales time series in Table 18.1 and Figure 18.1. The sales data for the 12 weeks are entered into worksheet rows 2 through 13 of column B. The following steps can be used to produce a three-week moving average.

Step 1. Click the **Data** tab on the Ribbon
Step 2. In the **Analysis** group, click **Data Analysis**
Step 3. Choose **Moving Average** from the list of Analysis Tools
       Click **OK**
Step 4. When the Moving Average dialog box appears:
       Enter B2:B13 in the **Input Range** box
       Enter 3 in the **Interval** box
       Enter C2 in the **Output Range** box
       Click **OK**

The three-week moving averages will appear in column C of the worksheet. The forecast for week 4 appears next to the sales value for week 3, and so on. Forecasts for periods of other length can be computed easily by entering a different value in the **Interval** box.

Exponential Smoothing

WEB file
Gasoline

To show how Excel can be used for exponential smoothing, we again develop a forecast for the gasoline sales time series in Table 18.1 and Figure 18.1. The sales data for the 12 weeks are entered into worksheet rows 2 through 13 of column B. The following steps can be used to produce a forecast using a smoothing constant of $\alpha = .2$.

Step 1. Click the **Data** tab on the Ribbon
Step 2. In the **Analysis** group, click **Data Analysis**
Step 3. Choose **Exponential Smoothing** from the list of Analysis Tools
       Click **OK**
Step 4. When the Exponential Smoothing dialog box appears:
       Enter B2:B13 in the **Input Range** box
       Enter .8 in the **Damping factor** box

[†]The results differ slightly from the results shown in Table 18.12 because Minitab computes the seasonal indexes using the median of the seasonal-irregular values.

> Enter C2 in the **Output Range** box
> Click **OK**

The exponential smoothing forecasts will appear in column C of the worksheet. Note that the value we entered in the Damping factor box is $1 - \alpha$; forecasts for other smoothing constants can be computed easily by entering a different value for $1 - \alpha$ in the Damping factor box.

Trend Projection

WEB file

Bicycle

To show how Excel can be used for trend projection, we develop a forecast for the bicycle sales time series in Table 18.3 and Figure 18.3. The data, with appropriate labels in row 1, are entered into worksheet rows 1 through 11 of columns A and B. The following steps can be used to produce a forecast for year 11 by trend projection.

> **Step 1.** Select an empty cell in the worksheet
> **Step 2.** Select the **Formulas** tab on the Ribbon
> **Step 3.** In the **Function Library** group, click **Insert** Function
> **Step 4.** When the Insert Function dialog box appears:
> > Choose **Statistical** in the Or select a category box
> > Choose **Forecast** in the Select a function box
> > Click **OK**
> **Step 5.** When the Forecast Arguments dialog box appears:
> > Enter 11 in the **x** box
> > Enter B2:B11 in the **Known y's** box
> > Enter A2:A11 in the **Known x's** box
> > Click **OK**

The forecast for year 11, in this case 32.5, will appear in the cell selected in step 1.

Appendix 18.3 # Forecasting Using StatTools

In this appendix we show how StatTools can be used to develop forecasts using three forecasting methods: moving averages, exponential smoothing, and Holt's linear exponential smoothing.

Moving Averages

WEB file

Gasoline

To show how StatTools can be used to develop forecasts using the moving averages method we will develop a forecast for the gasoline sales time series in Table 18.1 and Figure 18.1. Begin by using the Data Set Manager to create a StatTools data set for these data using the procedure described in the appendix in Chapter 1. The following steps will generate a three-week moving average forecast for week 13.

> **Step 1.** Click the **StatTools** tab on the Ribbon
> **Step 2.** In the **Analyses Group**, click **Time Series and Forecasting**
> **Step 3.** Choose the **Forecast** option
> **Step 4.** When the StatTools-Forecast dialog box appears
> > In the **Variables** section, select **Sales**
> > Select the **Forecast Settings** tab
> > > In the **Method** section, select **Moving Average**
> > > In the Parameters section, enter 3 in the **Span** box
> > Select the **Time Scale** tab

      In the **Seasonal Period** section, select **None**
      In the **Label Style** section, select **Integer**
  Click **OK**

The following output will appear in a new worksheet: three measures of forecast accuracy; time series plots showing the original data, the forecasts, and the forecast errors; and a table showing the forecasts and forecast errors. Note that StatTools uses the term "Mean Abs Err" to identify the value of MAE; "Root Mean Sq Err" to identify the square root of the value of MSE; and "Mean Abs Per% Err" to identify the value of MAPE.

Exponential Smoothing

WEB file

Gasoline

To show how StatTools can be used to develop an exponential smoothing forecast, we will again develop a forecast of sales in week 13 for the gasoline time series shown in Table 18.1 and Figure 18.1. Begin by using the Data Set Manager to create a StatTools data set for these data using the procedure described in the appendix in Chapter 1. The following steps will produce a forecast using a smoothing constant of $\alpha = .2$.

    Step 1. Click the **StatTools** tab on the Ribbon
    Step 2. In the **Analyses Group**, click **Time Series and Forecasting**
    Step 3. Choose the **Forecast** option
    Step 4. When the StatTools-Forecast dialog box appears
           In the **Variables** section, select **Sales**
           Select the **Forecast Settings** tab
             In the **Method** section, select **Exponential Smoothing (Simple)**
             Remove the check mark in the **Optimize Parameters** box
             In the Parameters section, enter .2 in the **Level (a)** box
           Select the Time Scale tab
             In the **Seasonal Period** section, select **None**
             In the **Label Style** section, select **Integer**
           Click **OK**

The following output will appear in a new worksheet: three measures of forecast accuracy; time series plots showing the original data, the forecasts, and the forecast errors; and a table showing the forecasts and forecast errors. Note that StatTools uses the term "Mean Abs Err" to identify the value of MAE; "Root Mean Sq Err" to identify the square root of the value of MSE; and "Mean Abs Per% Err" to identify the value of MAPE.

Holt's Linear Exponential Smoothing

WEB file

Bicycle

To show how StatTools can be used for trend projection, we develop a forecast for the bicycle sales time series in Table 18.3 and Figure 18.3 using Holt's linear exponential smoothing. Begin by using the Data Set Manager to create a StatTools data set for these data using the procedure described in the appendix in Chapter 1. The following steps will produce a forecast using smoothing constants of $\alpha = 1$ and $\beta = .2$.

    Step 1. Click the **StatTools** tab on the Ribbon
    Step 2. In the **Analyses Group**, click **Time Series and Forecasting**
    Step 3. Choose the **Forecast** option
    Step 4. When the StatTools-Forecast dialog box appears
           In the **Variables** section, select **Sales**
           Select the **Forecast Settings** tab
             In the **Method** section, select **Exponential Smoothing (Holt's)**
             Remove the check mark in the **Optimize Parameters** box

> In the Parameters section, enter .1 in the **Level (a)** box
> In the Parameters section, enter .2 in the **Trend (b)** box
> Select the Time Scale tab
>> In the **Seasonal Period** section, select **None**
>> In the **Label Style** section, select **Integer**
> Click **OK**

The following output will appear in a new worksheet: three measures of forecast accuracy; time series plots showing the original data, the forecasts, and the forecast errors; and a table showing the forecasts and forecast errors. Note that StatTools uses the term "Mean Abs Err" to identify the value of MAE; "Root Mean Sq Err" to identify the square root of the of MSE; and "Mean Abs Per% Err" to identify the value of MAPE. The StatTools output differs slightly from the results shown in Section 18.4 because StatTools uses a different approach to compute the estimate of the slope in period 1. With larger data sets the choice of startup values is not critical.