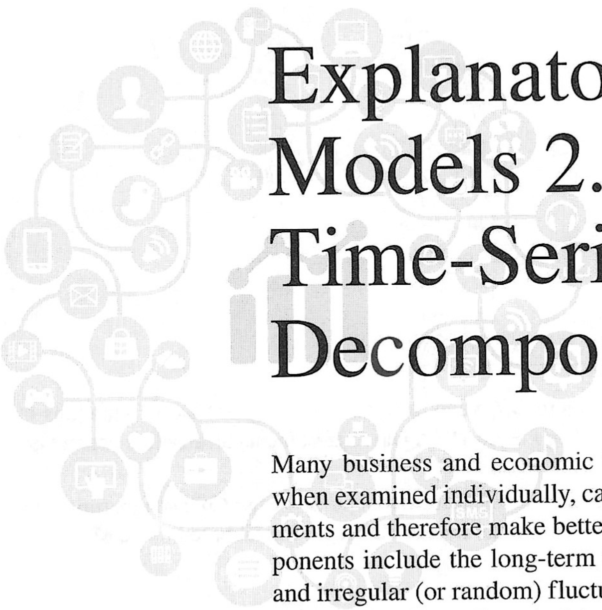
Chapter **Six**

Explanatory Models 2. Time-Series Decomposition



©VLADGRIN/Getty Images

Many business and economic time series contain underlying components that, when examined individually, can help the forecaster better understand data movements and therefore make better forecasts. As discussed in Chapter 2, these components include the long-term trend, seasonal fluctuations, cyclical movements, and irregular (or random) fluctuations. Time-series decomposition models can be used to identify such underlying components by breaking the series into its component parts and then reassembling the parts to construct a forecast.

These models are among the oldest forecasting techniques available and yet remain popular today. Their popularity is due primarily to three factors. First, in many situations, time-series decomposition models provide excellent forecasts. Second, these models are relatively easy to understand and to explain to forecast users. This enhances the likelihood that the forecasts will be correctly interpreted and properly used. Third, the information provided by time-series decomposition is consistent with the way managers tend to look at data and often helps them to get a better handle on data movements by providing concrete measurements for factors that are otherwise not quantified.

The information provided by time-series decomposition is consistent with the way managers tend to look at data and often helps them to get a better handle on data movements by providing concrete measurements for factors that are otherwise not quantified.

There are a number of different methods for decomposing a time series. The one we will use is usually referred to as *classical time-series decomposition* and involves the ratio-to-moving-average technique. The classical time-series decomposition model uses the concepts of moving averages presented in Chapter 3 and trend projections discussed in Chapter 4. It also accounts for seasonality in a multiplicative way that is similar to what you have seen in Winters' exponential smoothing and the way we used seasonal indices in earlier chapters.[1]

[1] Remember that you have also accounted for seasonality using dummy variables in regression models. That method uses additive factors rather than multiplicative ones to account for seasonal patterns.

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

1. Explain the similarity between how time series decomposition and Winters' exponential smoothing deal with seasonality.
2. Explain the four components of a time series. Discuss the trend, the seasonal, and the cyclical components.
3. Explain the difference between seasonal factors (SF) and seasonal indices (SI).
4. Explain how one determines the long-term trend for time-series decomposition.
5. Describe how "cycles" in a business environment differ from true cycles.

THE BASIC TIME-SERIES DECOMPOSITION MODEL

Look at the data on single-family private housing starts (PHS) that are shown in Table 6.1 and Figure 6.1 While the series appears quite volatile, there is also some pattern to the movement in the data. The sharp increases and decreases in housing starts appear to follow one another in a reasonably regular manner, which may reflect a seasonal component. There also appears to be some long-term wavelike movement to the data as well as a slight negative trend. Patterns such as these are relatively common and can best be understood if they can each be isolated and examined individually. The classical time-series decomposition forecasting technique is a well-established procedure for accomplishing this end.

TABLE 6.1 **Single-Family Private Housing Starts (PHS) in Thousands of Units** (c6t1&f1)

Source: Data from Economagic.com

Date	PHS (000)	Date	PHS (000)	Date	PHS (000)
Feb-67	147.1	Feb-84	236.5	Feb-01	273.8
May-67	254.7	May-84	332.6	May-01	373.8
Aug-67	244.2	Aug-84	280.3	Aug-01	341.1
Nov-67	197.8	Nov-84	234.7	Nov-01	284.5
Feb-68	179.9	Feb-85	215.3	Feb-02	293.3
May-68	266.2	May-85	317.9	May-02	386
Aug-68	249	Aug-85	295	Aug-02	360.6
Nov-68	204.2	Nov-85	244.1	Nov-02	318.6
Feb-69	171.1	Feb-86	234.1	Feb-03	304.1
May-69	259	May-86	369.4	May-03	406.3
Aug-69	214.5	Aug-86	325.4	Aug-03	412
Nov-69	165.9	Nov-86	250.6	Nov-03	376.7
Feb-70	136.7	Feb-87	241.4	Feb-04	345.1
May-70	231.6	May-87	346.5	May-04	455.7

(continued on next page)

TABLE 6.1
(continued)

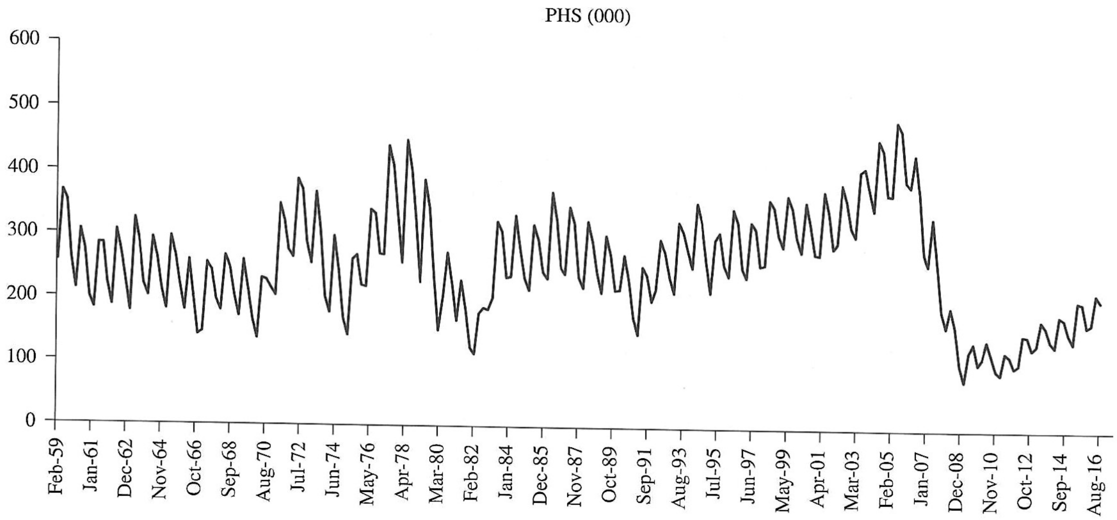
Date	PHS (000)	Date	PHS (000)	Date	PHS (000)
Aug-70	228.8	Aug-87	321.3	Aug-04	439.9
Nov-70	215.8	Nov-87	237.1	Nov-04	369.8
Feb-71	204.8	Feb-88	219.7	Feb-05	368.8
May-71	348.5	May-88	323.7	May-05	484.7
Aug-71	321.5	Aug-88	293.4	Aug-05	470.6
Nov-71	276.2	Nov-88	244.6	Nov-05	391.7
Feb-72	263.9	Feb-89	212.7	Feb-06	382.3
May-72	386.9	May-89	302.1	May-06	432.7
Aug-72	370.9	Aug-89	272.1	Aug-06	372.3
Nov-72	287.6	Nov-89	216.5	Nov-06	278
Feb-73	255.8	Feb-90	217	Feb-07	259.6
May-73	366.9	May-90	271.3	May-07	332.9
Aug-73	306	Aug-90	233	Aug-07	265.3
Nov-73	203.3	Nov-90	173.6	Nov-07	188.3
Feb-74	177.8	Feb-91	146.7	Feb-08	161.9
May-74	297.8	May-91	254.1	May-08	193.9
Aug-74	243.9	Aug-91	239.8	Aug-08	163
Nov-74	168.4	Nov-91	199.8	Nov-08	103.2
Feb-75	142.3	Feb-92	218.5	Feb-09	78.3
May-75	260.9	May-92	296.4	May-09	123.7
Aug-75	268	Aug-92	276.4	Aug-09	138.3
Nov-75	221	Nov-92	238.8	Nov-09	104.7
Feb-76	219	Feb-93	213.2	Feb-10	114.3
May-76	339.6	May-93	323.7	May-10	142.2
Aug-76	333.6	Aug-93	309.3	Aug-10	119
Nov-76	270.1	Nov-93	279.4	Nov-10	95.6
Feb-77	268.7	Feb-94	252.6	Feb-11	89.5
May-77	440.1	May-94	354.2	May-11	123.4
Aug-77	410.3	Aug-94	325.7	Aug-11	117.7
Nov-77	331.8	Nov-94	265.9	Nov-11	99.9
Feb-78	257.5	Feb-95	214.2	Feb-12	105.5
May-78	449.1	May-95	296.7	May-12	151.1
Aug-78	403.9	Aug-95	308.2	Aug-12	150.1
Nov-78	322.9	Nov-95	257.2	Nov-12	128.6
Feb-79	226.6	Feb-96	240	Feb-13	136.1
May-79	386.9	May-96	344.5	May-13	174.1

TABLE 6.1
(continued)

Date	PHS (000)	Date	PHS (000)	Date	PHS (000)
Aug-79	342.9	Aug-96	324	Aug-13	164.9
Nov-79	237.7	Nov-96	252.5	Nov-13	142.6
Feb-80	150.9	Feb-97	237.8	Feb-14	133.8
May-80	203.3	May-97	324.5	May-14	182.6
Aug-80	272.6	Aug-97	314.5	Aug-14	177.6
Nov-80	225.3	Nov-97	256.8	Nov-14	153.8
Feb-81	166.5	Feb-98	258.4	Feb-15	139.9
May-81	229.9	May-98	360.4	May-15	205.4
Aug-81	184.8	Aug-98	348	Aug-15	203.2
Nov-81	124.1	Nov-98	304.6	Nov-15	166.1
Feb-82	113.6	Feb-99	287.2	Feb-16	170.4
May-82	178.2	May-99	366.8	May-16	217.7
Aug-82	186.7	Aug-99	347.2	Aug-16	206.5
Nov-82	184.1	Nov-99	301.3	Nov-16	186.9
Feb-83	202.9	Feb-00	278.2		
May-83	322.3	May-00	357		
Aug-83	307.5	Aug-00	320.5		
Nov-83	234.8	Nov-00	275.2		

FIGURE 6.1 **Single-Family Private Housing Starts in Thousands of Units by Quarter.** (c6t1&f1)
This plot of private housing starts shows the volatility in the data. There are repeated sharp upward and downward movements that appear regular and may be of a seasonal nature. There also appears to be some wavelike cyclical pattern and perhaps a very slight negative trend.

Source: economagic.com.



The model can be represented by a simple algebraic statement, as follows:

$$Y = T \times S \times C \times I$$

where Y is the variable that is to be forecast, T is the long-term (or secular) trend in the data, S is a seasonal adjustment factor, C is the cyclical adjustment factor, and I represents irregular or random variations in the series. Our objective will be to find a way to decompose this series into the individual components.

DESEASONALIZING THE DATA AND FINDING SEASONAL INDICES

The first step in working with this model is to remove the short-term fluctuations from the data so that the longer-term trend and cycle components can be more clearly identified.

The first step in working with this model is to remove the short-term fluctuations from the data so that the longer-term trend and cycle components can be more clearly identified. These short-term fluctuations include both seasonal patterns and irregular variations. They can be removed by calculating an appropriate moving average (MA) for the series. The moving average should contain the same number of periods as there are in the seasonality that you want to identify. Thus, if you have quarterly data and suspect seasonality on a quarterly basis, a four-period moving average is appropriate. If you have monthly data and want to identify the monthly pattern in the data, a 12-period moving average should be used. The moving average for time period t (MA_t) is calculated as follows:

For quarterly data:

$$MA_t = (Y_{t-2} + Y_{t-1} + Y_t + Y_{t+1})/4$$

For monthly data:

$$MA_t = (Y_{t-6} + Y_{t-5} + \cdots + Y_t + Y_{t+1} + \cdots + Y_{t+5})/12$$

The moving average represents a "typical" level of Y for the year that is centered on that moving average.

The moving average for each time period contains one element from each of the seasons. For example, in the case of quarterly data, each moving average would contain a first-quarter observation, a second-quarter observation, a third-quarter observation, and a fourth-quarter observation (not necessarily in that order). The average of these four quarters should therefore not have any seasonality. Thus, the moving average represents a "typical" level of Y for the year that is centered on that moving average. When an even number of periods is used in calculating a moving average, however, it is really not centered in the year. The following simple example will make that clear and also help you verify your understanding of how the moving averages are calculated.

Let Y be the sales of a line of swimwear for which we have quarterly data (we will look at only six quarters of the data stream). MA_3 is the average of quarters 1 through 4. To be centered in the first year, it should be halfway between the second and third quarters, but the convention is to place it at the third quarter ($t = 3$). Note that each of the moving averages shown in the following example contains a first-, second-, third-, and fourth-quarter observation. Thus, seasonality in the data is removed. Irregular fluctuations are also largely removed, since such variations are random events that are likely to offset one another over time.

	Time Index	Y	Moving Average	Centered Moving Average
Year 1:				
First quarter	1	10	MISSING	MISSING
Second quarter	2	18	MISSING	MISSING
Third quarter	3	20	15.0(MA_3)	15.25(CMA_3)
Fourth quarter	4	12	15.5(MA_4)	15.75(CMA_4)
Year 2:				
First quarter	5	12	16.0(MA_5)	MISSING
Second quarter	6	20	MISSING	MISSING

$$MA_3 = (10 + 18 + 20 + 12)/4 = 15.0$$
$$MA_4 = (18 + 20 + 12 + 12)/4 = 15.5$$
$$MA_5 = (20 + 12 + 12 + 20)/4 = 16.0$$

As noted, when an even number of periods is used, the moving averages are not really centered in the middle of the year. To center the moving averages, a two-period moving average of the moving averages is calculated.[2] This is called a *centered moving average.* The centered moving average for time period t (CMA_t) is found as follows:

$$CMA_t = (MA_t + MA_{t+1})/2$$

For the swimwear data used in our example, we have:

$$CMA_3 = (15.0 + 15.5)/2 = 15.25$$
$$CMA_4 = (15.5 + 16.0)/2 = 15.75$$

This second moving average further helps to smooth out irregular or random fluctuations in the data.

Note the "MISSING" that appears under the moving average and centered moving average columns in the data table. With just six data points, we could not calculate four-period moving averages for the first, second, or sixth time period. We then lose one more time period at the end of the data stream when calculating the centered moving average. Thus, the smoothing process has a cost in terms of the loss of some data points. If an n-period moving average is used, $n/2$ points will be lost at each end of the data series by the time the centered moving averages have been calculated. This cost is not without benefit, however, since the process will eventually provide clarification of the patterns in the data.

By comparing the actual value of the series in any time period (Y_t) with the deseasonalized value (CMA_t), you can get a measure of the degree of seasonality.

The centered moving averages represent the deseasonalized data (i.e., seasonal variations have been removed through an averaging process). By comparing the actual value of the series in any time period (Y_t) with the deseasonalized value (CMA_t), you can get a measure of the degree of seasonality. In classical

[2] If the number of periods used is odd, the moving averages will automatically be centered, and no further adjustment is usually made.

time-series decomposition, this is done by finding the ratio of the actual value to the deseasonalized value. The result is called a *seasonal factor* (SF_t). That is:

$$SF_t = Y_t/CMA_t$$

A seasonal factor greater than 1 indicates a period in which Y is greater than the quarterly average for the year, while the reverse is true if SF is less than 1. For our brief swimwear sales example, we can calculate seasonal factors for the third and fourth time periods as follows:

$$SF_3 = Y_3/CMA_3 = 20/15.25 = 1.31$$
$$SF_4 = Y_4/CMA_4 = 12/15.75 = 0.76$$

We see that the third period (third quarter of year 1) is a high-sales quarter, while the fourth period is a low-sales quarter. This makes sense, since swimwear would be expected to sell well in July, August, and September but not in October, November, and December.

When we look at all of the seasonal factors for an extended time period, we generally see reasonable consistency in the values for each season. We would not expect all first-quarter seasonal factors to be exactly the same, but they are likely to be similar. To establish a seasonal index (SI), we average the seasonal factors for each season. This will now be illustrated for the private housing starts data shown initially in Table 6.1 and Figure 6.1.

The data for private housing starts are reproduced in part in Table 6.2. Only the beginning and near the end of the series are shown, but that is sufficient to illustrate all of the necessary calculations. The four-period moving average for private housing starts is denoted as PHSMA4 (private housing starts four-period moving average) and is shown in the fourth column of Table 6.2. The elements included in two values of PHSMA are shown by the brackets in the table and are calculated by adding the corresponding four quarters and then dividing by four.

TABLE 6.2 Time-Series Decomposition of Private Housing Starts (c6t2&f2)

Date	Time Index	PHS (000)	PHSMA4	PHSCMA	PHSCMAT	CF	SF	SI
Feb-59	1	256.7			266.20			**0.82**
May-59	2	367.3			266.11			1.19
Aug-59	3	350.8	308.50	303.09	266.01	1.14	1.16	1.11
Nov-59	4	259.2	297.68	289.96	265.92	1.09	0.89	0.89
Feb-60	5	213.4	282.25	272.70	265.83	1.03	0.78	**0.82**
May-60	6	305.6	263.15	255.90	265.73	0.96	1.19	1.19
Aug-60	7	274.4	248.65	244.86	265.64	0.92	1.12	1.11
Nov-60	8	201.2	241.08	238.45	265.55	0.90	0.84	0.89
Feb-61	9	183.1	235.83	237.05	265.45	0.89	0.77	**0.82**
May-61	10	284.6	238.28	240.93	265.36	0.91	1.18	1.19

TABLE 6.2 (continued)

Date	Time Index	PHS (000)	PHSMA4	PHSCMA	PHSCMAT	CF	SF	SI
Aug-61	11	284.2	243.58	244.13	265.27	0.92	1.16	1.11
Nov-61	12	222.4	244.68	247.24	265.17	0.93	0.90	0.89
Feb-62	13	187.5	249.80	248.31	265.08	0.94	0.76	**0.82**
May-62	14	305.1	246.83	247.33	264.99	0.93	1.23	1.19
Aug-62	15	272.3	247.83	246.66	264.90	0.93	1.10	1.11
Nov-62	16	226.4	245.50	247.90	264.80	0.94	0.91	0.89
•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•
Feb-13	217	136.1	147.23	149.08	246.07	0.61	0.91	**0.82**
May-13	218	174.1	150.93	152.68	245.97	0.62	1.14	1.19
Aug-13	219	164.9	154.43	154.14	245.88	0.63	1.07	1.11
Nov-13	220	142.6	153.85	154.91	245.79	0.63	0.92	0.89
Feb-14	221	133.8	155.98	157.56	245.69	0.64	0.85	**0.82**
May-14	222	182.6	159.15	160.55	245.60	0.65	1.14	1.19
Aug-14	223	177.6	161.95	162.71	245.51	0.66	1.09	1.11
Nov-14	224	153.8	163.48	166.33	245.41	0.68	0.92	0.89
Feb-15	225	139.9	169.18	172.38	245.32	0.70	0.81	**0.82**
May-15	226	205.4	175.58	177.11	245.23	0.72	1.16	1.19
Aug-15	227	203.2	178.65	182.46	245.13	0.74	1.11	1.11
Nov-15	228	166.1	186.28	187.81	245.04	0.77	0.88	0.89
Feb-16	229	170.4	189.35	189.76	244.95	0.77	0.90	**0.82**
May-16	230	217.7	190.18	192.78	244.85	0.79	1.13	1.19
Aug-16	231	206.5	195.38		244.76			1.11
Nov-16	232	186.9			244.67			0.89

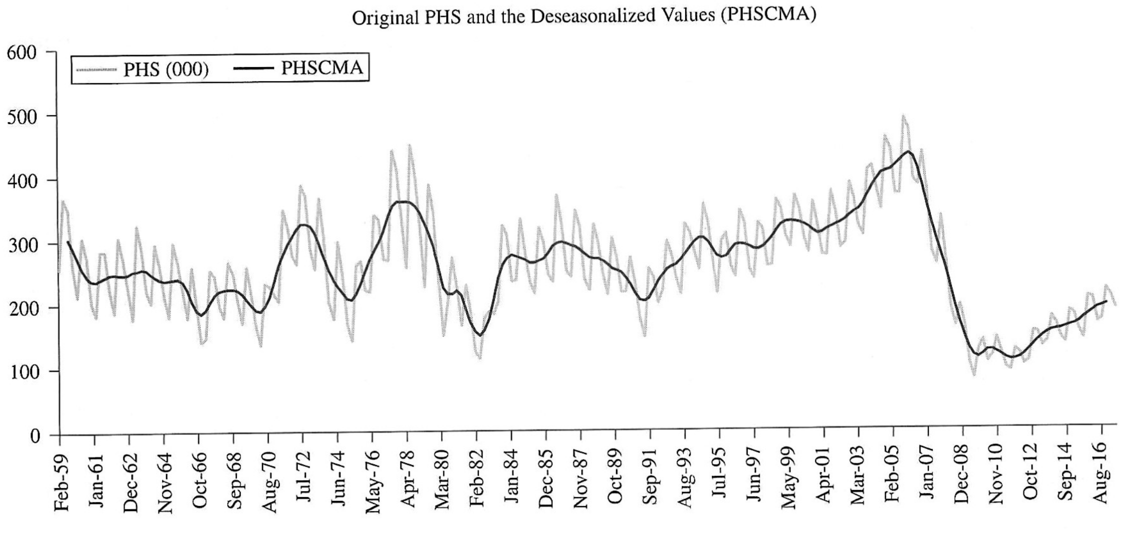
PHS = Private housing starts (in thousands)
PHSMA4 = Private housing starts four-period moving average
PHSCMA = Private housing starts centered moving average
PHSCMAT = Private housing starts centered moving-average trend (trend component)
CF = Cycle factor (PHSCMA4/PHSCMAT)
SF = Seasonal factor (PHS/PHSCMA4)
SI = Seasonal indices (normalized mean of seasonal factors)

The centered moving average (PHSCMA) is shown in the fifth column. The calculation of PHSCMA for Aug-59 is:

$$\text{PHSCMA} = (308.50 + 297.689)/2 = 303.09$$

Notice that for PHSCMA, there is no value for each of the first two and last two quarters. This loss of four quarters of data over 232 observations is not too

FIGURE 6.2 Private Housing Starts (PHS) with the Centered Moving Average of Private Housing Starts (PHSCMA) in Thousands of Units. (c6t2&f2)
The centered moving-average series, shown by the darker line, is much smoother than the original series of private housing starts data (lighter line) because the seasonal pattern and the irregular or random fluctuations in the data are removed by the process of calculating the centered moving averages.



severe. The two lost quarters that are most critical are the last two, since they are the closest to the period to be forecast.

Figure 6.2 shows a plot of the original private housing starts (PHS) data (lighter line) along with the deseasonalized data (darker line) represented by the centered moving averages (PHSCMAs). Notice how much smoother the data appear once seasonal variations and random fluctuations have been removed.

The process of deseasonalizing the data has two useful results:

The deseasonalized data allow us to see better the underlying pattern in the data.

1. The deseasonalized data allow us to see better the underlying pattern in the data, as illustrated in Figure 6.2.
2. It provides us with measures of the extent of seasonality in the form of seasonal indices.

The seasonal factors (SF) for each quarter are shown in the eighth column of Table 6.2. Recall that the seasonal factors measure the extent to which the observed value for each quarter is above or below the deseasonalized value (SF > 1 and SF < 1, respectively). For this example:

$$SF_t = PHS_t / PHSCMA_t$$

For the first two and the last two quarters, seasonal factors cannot be calculated, since there are no centered moving averages for those quarters. The calculations of the seasonal factor for Feb-13 is:

$$SF = 136.1/149.08 = 0.91$$

It makes sense that winter would have a low SF (less than 1), since this is often not a good period in which to start building. The reverse is true in the spring and summer.

Since the seasonal factors for each period are bound to have some variability, we calculate a seasonal index (SI) for each period, which is a standardized average of all of that period's seasonal factors. As shown below, the seasonal indices for the second and third quarter are above 1 and indicate that these quarters are generally high for private housing starts.

SI
0.82
1.19
1.11
0.89

These add to 4.00, as expected. The warmer spring and summer months are the strongest seasons for housing starts, whereas the fall and winter months are low.

As shown above, the private housing starts' seasonal index for the first quarter is 0.82. This means that the typical first-quarter PHS is only 82 percent of the average quarterly value for the year. Thus, if the housing starts for a year totaled 400, we would expect 82 to occur in the first quarter. The 82 is found by dividing the yearly total (400) by 4 and then multiplying the result by the seasonal index $[(400/4) \times 0.82 = 82]$.

FINDING THE LONG-TERM TREND

The long-term trend is estimated from the deseasonalized data for the variable to be forecast. Remember that the centered moving average (CMA) is the series that remains after the seasonality and irregular components have been smoothed out by using moving averages. Thus, to find the long-term trend, we estimate a simple linear equation as:[3]

$$\begin{aligned} \text{CMA} &= f(\text{TIME}) \\ &= a + b(\text{TIME}) \end{aligned}$$

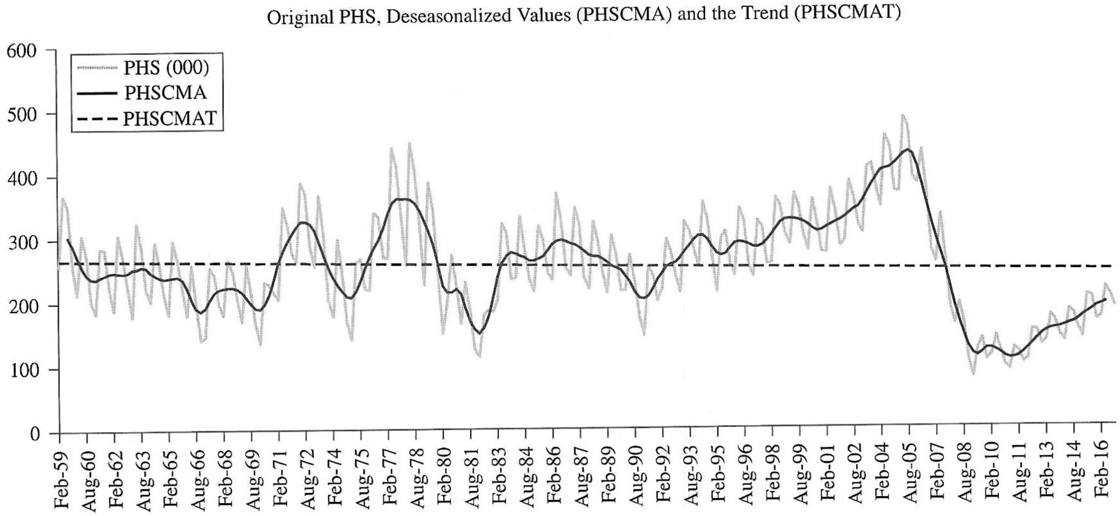
where TIME = 1 for the first period in the data set and increases by 1 each quarter thereafter. The values of a and b are normally estimated by using a computer regression program, but they can also be found quickly on most hand-held business calculators.

Once the trend equation has been determined, it is used to generate an estimate of the trend value of the centered moving average for the historical and forecast periods. This new series is the centered moving-average trend (CMAT).

For our example involving private housing starts, the linear trend of the deseasonalized data (PHSCMA) has been found to be slightly negative. The centered

[3] A linear trend is most often used, but a nonlinear trend may also be used. Looking at a graph such as the one shown in Figure 6.2 is helpful in determining which form would be most appropriate for the trend line.

FIGURE 6.3 Private Housing Starts (PHS) with Centered Moving Average (PHSCMA) and Centered Moving-Average Trend (PHSCMAT) in Thousands of Units. (c6t2&f2)
The long-term trend in private housing starts is shown by the straight dotted line (PHSCMAT). The lighter line is the raw data (PHS), while the wavelike dark line is the deseasonalized data (PHSCMA). The long-term trend is seen to be slightly negative. The equation for the trend line is: PHSCMAT = 266.293 -0.093(TIME).



moving-average trend for this example is denoted PHSCMAT, for "private housing starts centered moving-average trend." The equation is:

$$PHSCMAT = 266.293 - 0.093(TIME)$$

where TIME = 1 for for the first quarter of 1959. Because there are no data for the first two quarters or the last two quarters for PHSCMA, the linear trend is estimated using quarters 3 through 230. This line is shown in Figure 6.3, along with the graph of private housing starts (PHS) and the deseasonalized data (PHSCMA).

MEASURING THE CYCLICAL COMPONENT

The cyclical component of a time series is the extended wavelike movement about the long-term trend. It is measured by a cycle factor (CF), which is the ratio of the centered moving average (CMA) to the centered moving-average trend (CMAT). That is:

$$CF = CMA/CMAT$$

Looking at the length and amplitude of previous cycles may enable us to anticipate the next turning point in the current cycle.

A cycle factor greater than 1 indicates that the deseasonalized value for that period is above the long-term trend of the data. If CF is less than 1, the reverse is true.

The cycle factor is the most difficult component of a time series to analyze and to project into the forecast period. If analyzed carefully, however, it may also be the component that has the most to offer in terms of understanding where the industry may be headed. Looking at the length and amplitude of previous cycles may enable us to anticipate the next turning point in the current cycle. This is a major advantage of the time-series decomposition technique. An individual familiar with an industry can often explain cyclic movements around the trend line in terms of variables or events that, in retrospect, can be seen to have had some importance. By looking at those variables or events in the present, we can sometimes get some hint of the likely future direction of the cycle component.

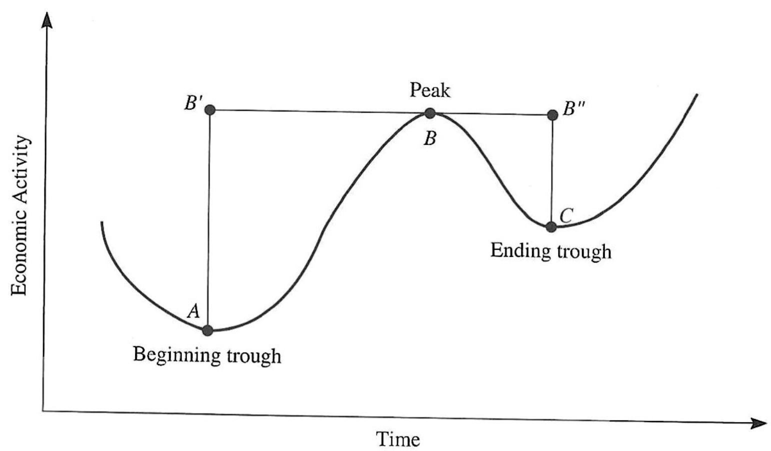
Overview of Business Cycles

Business cycles are long-term wavelike fluctuations in the general level of economic activity. They are often described by a diagram such as the one shown in Figure 6.4. The period of time between the beginning trough (A) and the peak (B) is called the *expansion phase,* while the period from peak (B) to the ending trough (C) is termed the *recession,* or *contraction, phase.*

The vertical distance between A and B' provides a measure of the degree of the expansion. The start of the expansion beginning at point A is determined by three consecutive months of increase in economic activity. Thus, the preceding recession is only officially over three months after the economy has turned around. Similarly, the severity of a recession is measured by the vertical distance between B'' and C, and the official beginning of the recession is dated as the first of three consecutive months of decline.

If business cycles were true cycles, they would have a constant amplitude. That is, the vertical distance from trough to peak and peak to trough would always be the same. In addition, a true cycle would also have a constant periodicity. That would mean that the length of time between successive peaks (or troughs) would always be the same. However, with economic and business activity, this degree of regularity is unlikely. As you will see when we look at the cyclical component for private housing starts, the vertical distances from trough to peak (or peak to trough) have considerable variability, as does the distance between successive peaks and successive troughs.

FIGURE 6.4
The General Business Cycle.
A business cycle goes through successive periods of expansion, contraction, expansion, contraction, and so on.



Business Cycle Indicators

There are a number of possible business cycle indicators, but three are particularly noteworthy:

1. The index of leading economic indicators
2. The index of coincident economic indicators
3. The index of lagging economic indicators

The individual series that make up each index are shown in Table 6.4.

TABLE 6.4 U.S. Business Cycle Indicators

Components of the Composite Indices
Leading Index
1 Average weekly hours, manufacturing
2 Average weekly initial claims for unemployment insurance
3 Manufacturers' new orders, consumer goods, and materials
4 ISM® new orders index
5 Manufacturers' new orders, nondefense capital goods excl. aircraft
6 Building permits, new private housing units
7 Stock prices, 500 common stocks
8 *Leading Credit Index*™
9 Interest rate spread, 10-year Treasury bonds less federal funds
10 Avg. consumer expectations for business conditions
Coincident Index
1 Employees on nonagricultural payrolls
2 Personal income less transfer payments
3 Industrial production
4 Manufacturing and trade sales
Lagging Index
1 Inventories to sales ratio, manufacturing and trade
2 Average duration of unemployment
3 Consumer installment credit outstanding to personal income ratio
4 Commercial and industrial loans
5 Average prime rate
6 Labor cost per unit of output, manufacturing
7 Consumer price index for services

The Conference Board, https://www.conference-board.org/pdf_free/press/US%20LEI%20-%20Tech%20Notes%20Jun%202017.pdf.

It is possible that one of these indices, or one of the series that makes up an index, may be useful in predicting the cycle factor in a time-series decomposition. This could be done in a regression analysis with the cycle factor (CF) as the dependent variable. These indices, or their components, may also be quite useful as independent variables in other regression models, such as those discussed in Chapters 4 and 5.

Figure 6.5 shows what are considered the official business cycles for the U.S. economy in recent years. The shaded vertical bars identify the officially designated periods of recession.

The Cycle Factor for Private Housing Starts

Let us return to our example involving private housing starts to examine how to calculate the cycle factor and how it might be projected into the forecast period. In Table 6.2, the cycle factors (CF) are shown in column seven. As indicated previously, each cycle factor is the ratio of the deseasonalized data (CMA) to the trend value (CMAT). For the private housing starts data, we have:

$$CF = PHSCMA/PHSCMAT$$

The actual calculation for Feb-05 is:

$$CF = 419.64/249.05 = 1.68$$

You can see in Figure 6.3 that in Jun-06, the centered moving average was above the trend line.

The cycle factor is plotted in Figure 6.6. You can see that the cycle factor (CF) moves above and below the line at 1.00 in Figure 6.6 exactly as the centered moving average moves above and below the trend line in Figure 6.3. By isolating the cycle factor in Figure 6.6, we can better analyze its movements over time.

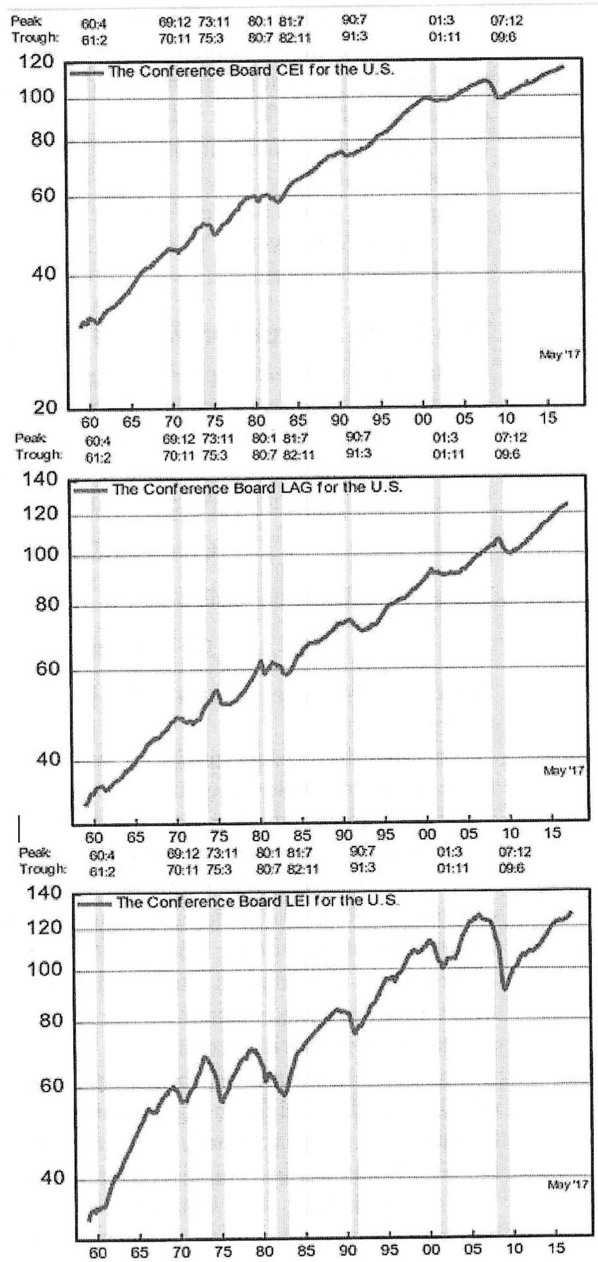
You see that the cyclical component for private housing starts does not have a constant amplitude or periodicity. The dates for peaks and troughs are shown in Figure 6.6b, along with the values of the cycle factor at those points. Identification of these dates and values is often helpful in considering when the cycle factor may next turn around (i.e., when the cycle factor turns from having a positive slope to a negative slope or vice versa).

For example, for the CF, the average height of the four peaks shown in Figure 6.6b is 1.38. That might be a good value to use for what you might expect for the height of the next peak. It would be reasonable to argue that the August 2005 peak (CF=1.73) was unusually high, so you might use the average of the other three peaks. That average is **1.27**. This would give you an expectation that the next peak would have a CF of 1.27. There is no perfect way to make this judgment. Others might suggest dropping both the highest and lowest values. But you need to have a reasonable logic for whatever decision you make. Once you have a decision about the expected height of the next peak, you need to decide when that next peak is likely to occur.

One way to approach this is to look at the historical peaks and see how many quarters there were between peaks. From the August 1972 to February 1978, the distance was 22 quarters. The next peak-to-peak distance was 34 quarters, and the

FIGURE 6.5
Official Business Cycles in the United States.
For each graph, the periods identified as official recessions are indicated by the grey vertical bands.

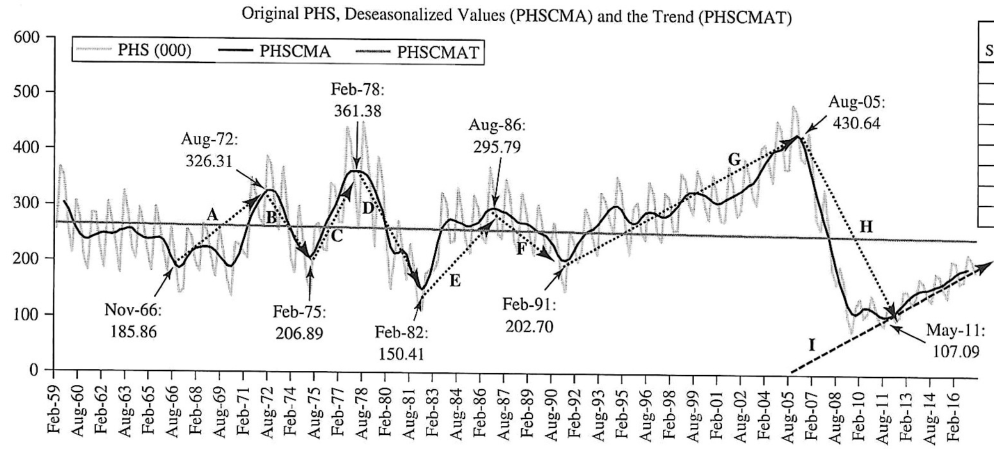
The Conference Board, https://www.conference-board.org/pdf_free/press/US%20LEI%20-%20Tech%20Notes%20Jun%202017.pdf. Content reproduced with permission.



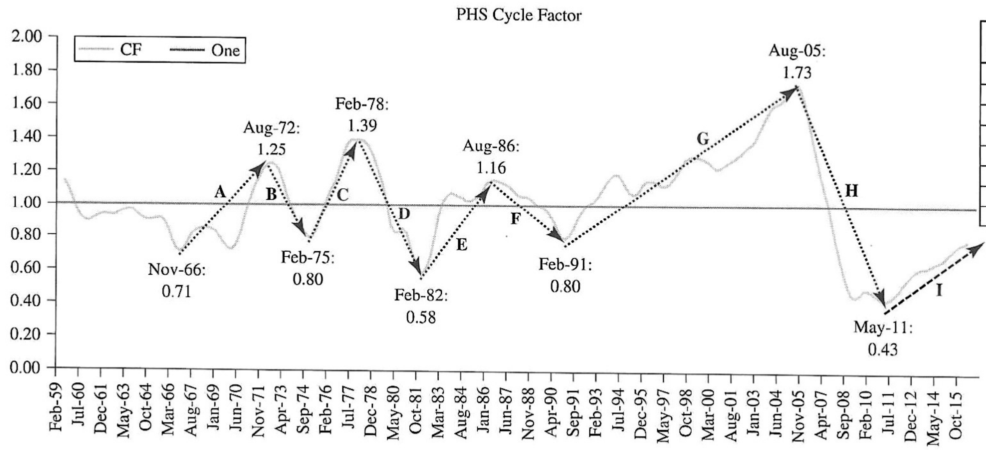
last peak-to-peak was 76 quarters. The latter covered an unusually long period from Feb-91 through Oct-05.[4] The average of these three is 44 quarters. Again one might suggest that because the last expansion was unusually long, only the

[4] As of the date of this edition of the text, a new peak had not yet been identified.

FIGURES 6.6A AND 6.6B. Cycle Factor (Cf) for Private Housing Starts. (c6f6a and c6f6b)
The cycle factor is the ratio of the centered moving average to the long-term trend in the data. As the upper graph (6.6a) shows, the deseasonalized values move slowly around the trend line with little regularity. In the lower graph (6.6b), you see that the actual cycle factor moves above and below the line at one in the same manner as PHSCMA moves around the trend. Dates and values of cycle factors at peaks and troughs are shown in the boxes in 6.6b.



Line Segment	Rise (+) or Fall (–)
A =	140.45
B =	–119.42
C =	154.49
D =	–210.97
E =	145.38
F =	–93.09
G =	227.94
H =	–323.55



Line Segment	Rise (+) or Fall (–)
A =	0.54
B =	–0.45
C =	0.59
D =	–0.81
E =	0.58
F =	–0.36
G =	0.93
H =	–1.3

average of the first two should be used. This would be 28 quarters. Again you see that there is considerable room for judgment involved.

Another thing to consider is the slope of the dotted lines representing the rate of expansion or contraction in the cycles. Look at the line segments labeled G and compare the slopes with the line segments I. They have slopes that are almost exactly the same based on a visual inspection of the two lines. Since we do not know the next peak we can only base this on what has already been observed. Based on Figure 6.6b, for the 76 quarters between Feb-91 and Aug-05, the change in the cycle factor was 0.93. Thus, over that period, the slope was 0.012 (0.93/76 = 0.012).

Now consider the slope of segment I in Figure 6.6b. We might expect the slope of I to remain at something close to 0.012. We see **1.27** is the height of the next peak, and we see the most recent trough is 0.43. The difference is 0.840 (1.27 - 0.43). At an increase of 0.012 per quarter, it will take an estimated 70 quarters to the next peak. Thus, our prediction would be that the next peak will be at a CF of 1.27 in the third quarter of 2028.

The seasonal index for the first quarter (designated by Feb here) is 0.82. Multiplying the predicted deseasonalised value (the centered moving average) by 0.82, we have: 327.83*0.82=268.82 as the prediction for private housing starts in the first quarter of 2025.

Here we made the judgment that the rate of recovery will be similar to the rate of the last recovery based on what we have observed thus far for the slope of line segment I. You can see that there is room for much debate about this prediction. Other analysts may focus more on the distances between peaks and troughs. We only know what is the best prediction in retrospect. The important thing is for you to have a logical and well-documented process to support your prediction.

It is important to recognize that there is no way to know exactly where the cycle factor will be in the forecast horizon, and there is no a priori way to determine the best technique for projecting the cycle factor. A thorough review of the past behavior of the cycle factor, along with alternative forecasts, should be evaluated for consistency and congruity before selecting values of the cycle factor for the forecast horizon.

Perhaps most frequently the cycle factor forecast is made on a largely judgmental basis by looking carefully at the historical values, especially historical turning points and the rates of descent or rise in the historical series. You might look at the peak-to-peak, trough-to-trough, peak-to-trough, and trough-to-peak distances by dating each turning point, such as we show in Figure 6.6. Then, you could calculate the average distance between troughs (or peaks) to get a feeling for when another such point is likely. You can also analyze the rates of increase and/or decrease in the cycle factor as a basis on which to judge the expected slope of the forecast of the cycle factor.

It is important to recognize that there is no way to know exactly where the cycle factor will be in the forecast horizon, and there is no a priori way to determine the best technique for projecting the cycle factor. A thorough review of the past behavior of the cycle factor, along with alternative forecasts, should be evaluated for consistency and congruity before selecting values of the cycle factor for the forecast horizon.

THE TIME-SERIES DECOMPOSITION FORECAST

You have seen that a time series of data can be decomposed into the product of four components:

$$Y = T \cdot S \cdot C \cdot I$$

where Y is the series to be forecast. The four components are:

T = The long-term trend based on the deseasonalized data. It is often called the *centered moving-average trend* (CMAT), since the deseasonalized data are centered moving averages (CMA) of the original Y values.

S = Seasonal indices (SI). These are normalized averages of seasonal factors that are determined as the ratio of each period's actual value (Y) to the deseasonalized value (CMA) for that period.

C = The cycle component. The cycle factor (CF) is the ratio of CMA to CMAT and represents the gradual wavelike movements in the series around the trend line.

I = The irregular component. This is assumed equal to 1 unless the forecaster has reason to believe a shock may take place, in which case I could be different from 1 for all or part of the forecast period.

Previous sections of this chapter have illustrated how these components can be isolated and measured.

To prepare a forecast based on the time-series decomposition model, we simply reassemble the components. In general terms, the forecast for Y (FY) is:

$$\text{FY} = (\text{CMAT})(\text{SI})(\text{CF})(I)$$

For our private housing starts example, we will denote the forecast value based on the model as PHSFTSD. Thus,

$$\text{PHSFTSD} = (\text{PHSCMAT})(\text{SI})(\text{CF})(I)$$

where PHSCMAT is the private housing starts centered moving-average trend, SI is the seasonal indices, and Cf is the cycle factor. The irregular factor (I) is assumed equal to 1, since we have no reason to expect it to be greater or less than 1 because of its random nature. The actual and forecast values for private housing starts are shown Figure 6.7. The actual values (PHS) are shown by the lighter solid line; fitted values are shown by the dashed line; and forecast values are shown by the dotted line. The forecast calculations are shown in Table 6.5 for

FIGURE 6.7 **Private Housing Starts (PHS) and A Time-Series Decomposition Forecast.** (c6t4&f7) The actual values for private housing starts are shown by the lighter blue line, the time-series decomposition fitted values are shown by the green dashed line, and a four-year forecast is shown by the dotted red line.

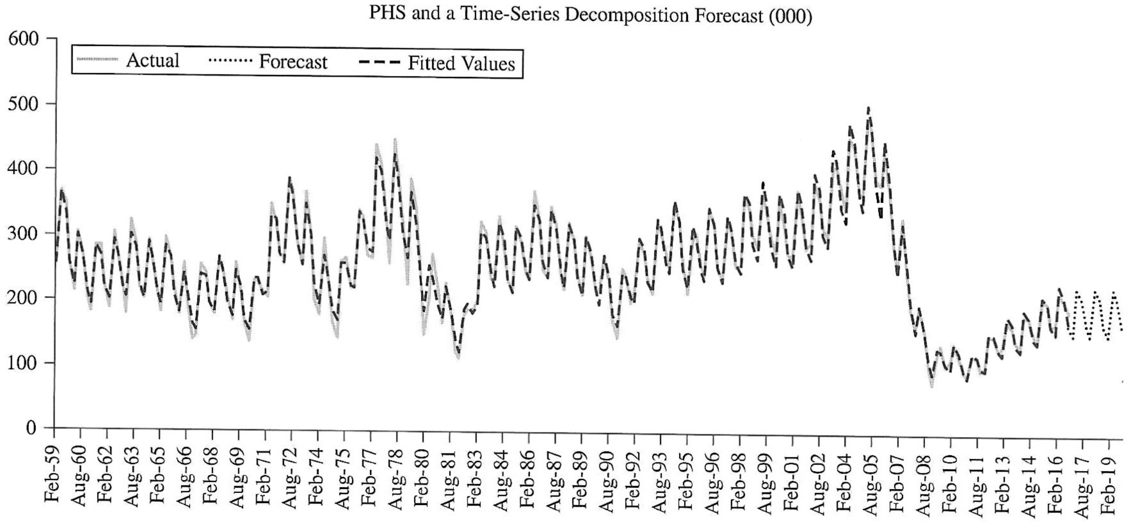


TABLE 6.5 **PHS Time-Series Decomposition Forecast** (c6t4&f7)

Date	Original Data	Forecasted Data	Centered Moving Average	CMA Trend	Seasonal Indices	Cycle Factors
Feb-1959	256.70				0.82	
May-1959	367.30				1.19	
Aug-1959	350.80	335.13	303.09	266.01	1.11	1.14
Nov-1959	259.20	258.04	289.96	265.92	0.89	1.09
Feb-1960	213.40	222.55	272.70	265.83	0.82	1.03
May-1960	305.60	304.08	255.90	265.73	1.19	0.96
Aug-1960	274.40	270.75	244.86	265.64	1.11	0.92
Nov-1960	201.20	212.20	238.45	265.55	0.89	0.90
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
Feb-2015	139.90	140.68	172.38	245.32	0.82	0.70
May-2015	205.40	210.46	177.11	245.23	1.19	0.72
Aug-2015	203.20	201.75	182.46	245.13	1.11	0.74
Nov-2015	166.10	167.14	187.81	245.04	0.89	0.77
Feb-2016	170.40	154.87	189.76	244.95	0.82	0.77
May-2016	217.70	229.07	192.78	244.85	1.19	0.79
Aug-2016	206.50	207.90		244.76	1.11	0.77
Nov-2016	186.90	168.56		244.67	0.89	0.77
Feb-2017		154.91		244.57	0.82	0.78
May-2017		225.57		244.48	1.19	0.78
Aug-2017		209.08		244.39	1.11	0.77
Nov-2017		168.51		244.29	0.89	0.78
Feb-2018		154.52		244.20	0.82	0.78
May-2018		224.85		244.11	1.19	0.78
Aug-2018		209.06		244.02	1.11	0.77
Nov-2018		168.25		243.92	0.89	0.78
Feb-2019		154.24		243.83	0.82	0.78
May-2019		224.47		243.74	1.19	0.78
Aug-2019		208.79		243.64	1.11	0.78
Nov-2019		167.99		243.55	0.89	0.78
Feb-2020		153.99		243.46	0.82	0.78
May-2020		224.13		243.36	1.19	0.78
Aug-2020		208.48		243.27	1.11	0.78
Nov-2020		167.73		243.18	0.89	0.78

Forecasting Winter Daily Natural Gas Demand at Vermont Gas Systems

1

Mike Flock, Distribution Engineer, Vermont Gas Systems, Inc.

Vermont Gas Systems is a natural gas utility with approximately 26,000 residential, business, and industrial customers in 13 towns and cities in northwestern Vermont. Vermont Gas Systems' Gas Control Department forecasts the gas demand and arranges the gas supply and transportation from suppliers in western Canada and storage facilities along the Trans-Canada Pipeline that deliver the gas to our pipeline. The quantities of gas must be specified to the suppliers at least 24 hours in advance. The Gas Control Department must request enough natural gas to meet the needs of the customers but must not over-request gas that will needlessly and expensively tax Trans-Canada Pipelines' facilities. Because Vermont Gas Systems has the storage capacity for only one hour's use of gas as a buffer between supply and demand, an accurate forecast of daily natural gas demand is critical.

Source: Flock, Mike, "Forecasting Winter Daily Natural Gas Demand at Vermont Gas Systems," *Journal of Business Forecasting* 13, no. 1 (Spring 1994), p. 23.

the first and last parts of the series. You will note that this method takes the trend (PHSCMAT) and makes two adjustments to it: the first adjusts it for seasonality (with SI), and the second adjusts it for cycle variations (with CF).

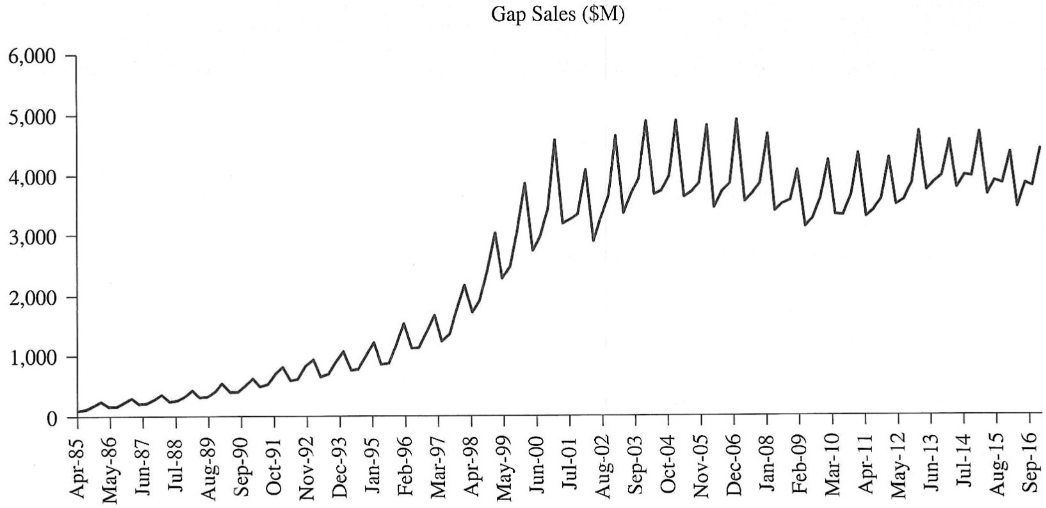
Because time-series decomposition models do not involve a lot of mathematics or statistics, they are relatively easy to explain to the end user. This is a major advantage, because if the end user has an appreciation of how the forecast was developed, he or she may have more confidence in its use for decision making.

Integrative Case

The Gap

FORECASTING GAP SALES DATA WITH TIME-SERIES DECOMPOSITION

For time-series decomposition, we need a long data stream to be able to identify cycles. Therefore, in this Gap case, we have extended the data back to 1985. Previously, we have started with 2006. The sales of Gap for the 128 quarters covering fiscal 1985 through the end of Gap's 2016 fiscal year are shown below. (c6Gap)



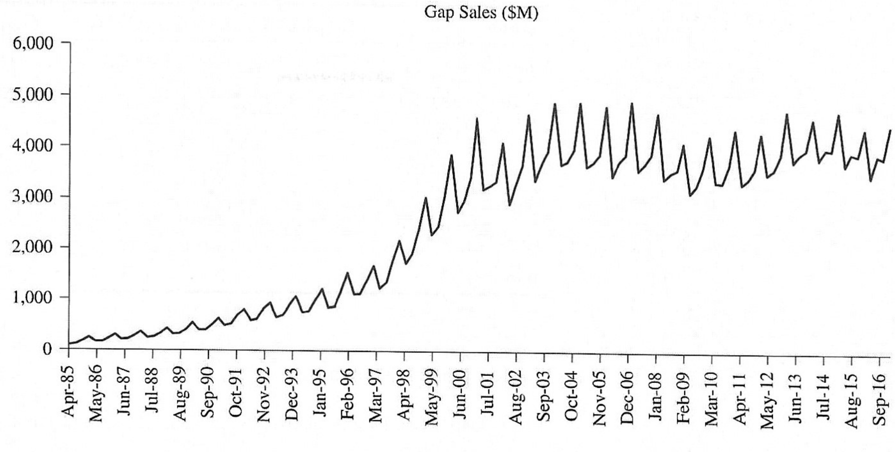
Case Questions

1. Based on the graph of Gap sales, describe what you see in the 1985–2016 Gap sales using the terms used in time-series decomposition: *trend, seasonality,* and *cycle.*
2. For this question, use only Excel. Using Gap sales data for 1985Q1 through 2016Q4, calculate the four-period moving average (GapMA4), the centered moving average (GapCMA), and the CMAT using a time index that goes from 1 for 1985Q1 through 128 for 2016Q4. Extend the trend through the 2017 fiscal year.
3. Plot Gap sales, GapCMA, and GapCMAT on the same graph for the period from 1985Q1 through 2017Q4. There will be some quarters for which not all variables will have values.
4. Calculate the seasonal indices (SF). Are they consistent with your expectations? Explain.
5. Calculate the cycle factors (CF). Plot CF along with a horizontal line at one. Describe where you see peaks and/or troughs. Do you think you have enough data to clearly identify the cycle for Gap sales? Explain.
6. Now use ForecastX™ to prepare a forecast of Gap sales for the 12 quarters of the 2017 through 2019 fiscal years. Show the audit report plot of the actual, fitted, and forecast sales. Does the forecast look reasonable? Why or why not?
7. What are the four seasonal indices calculated by ForecastX™? Do these appear consistent with the seasonal factors that you calculated in Excel? Why or why not?

Solutions to Case Questions

1. The Gap sales exhibit an increasing positive trend over the time frame being evaluated and a very clear seasonal pattern that repeats itself year to year. It appears that the seasonality may be more pronounced in the more recent years than it was in the early years. From this graph, it is not clear that there are the long-term swings that are normally associated with a cyclical pattern.

(c6Gap)



2. Below is an abbreviated table of the requested series:

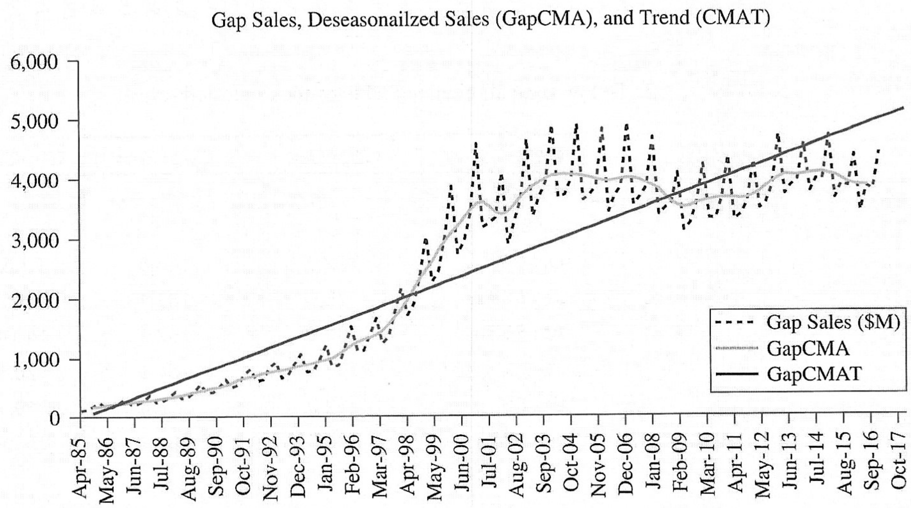
Date	GapMA4	GapCMA	GapCMAT
Apr-85			
Jul-85			
Oct-85	162.0	168.8	61.818
Jan-86	175.5	181.1	101.15
Apr-86	186.8	192.1	140.482
Jul-86	197.5	204.8	179.814
Oct-86	212.0	218.4	219.146
Jan-87	224.8	231.4	258.478
.	.	.	.
.	.	.	.
.	.	.	.
Apr-13	4,046.5	4,060.5	4388.338
Jul-13	4,074.5	4,055.8	4427.67
Oct-13	4,037.0	4,042.6	4467.002
Jan-14	4,048.3	4,062.4	4506.334
Apr-14	4,076.5	4,076.0	4545.666
Jul-14	4,075.5	4,092.1	4584.998
Oct-14	4,108.8	4,094.1	4624.33
Jan-15	4,079.5	4,069.1	4663.662
Apr-15	4,058.8	4,044.4	4702.994
Jul-15	4,030.0	3,989.6	4742.326
Oct-15	3,949.3	3,921.9	4781.658
Jan-16	3,894.5	3,888.6	4820.99
Apr-16	3,882.8	3,875.4	4860.322

(continued on next page)

(continued)

Date	GapMA4	GapCMA	GapCMAT
Jul-16	3,868.0	3,873.5	4899.654
Oct-16	3,879.0		4938.986
Jan-17			4978.318
Apr-17			5017.65
Jul-17			5056.982
Oct-17			5096.314
Jan-18			5135.646

3. The requested graph is shown below.



4. A sample of the calculation of seasonal factors is shown below:

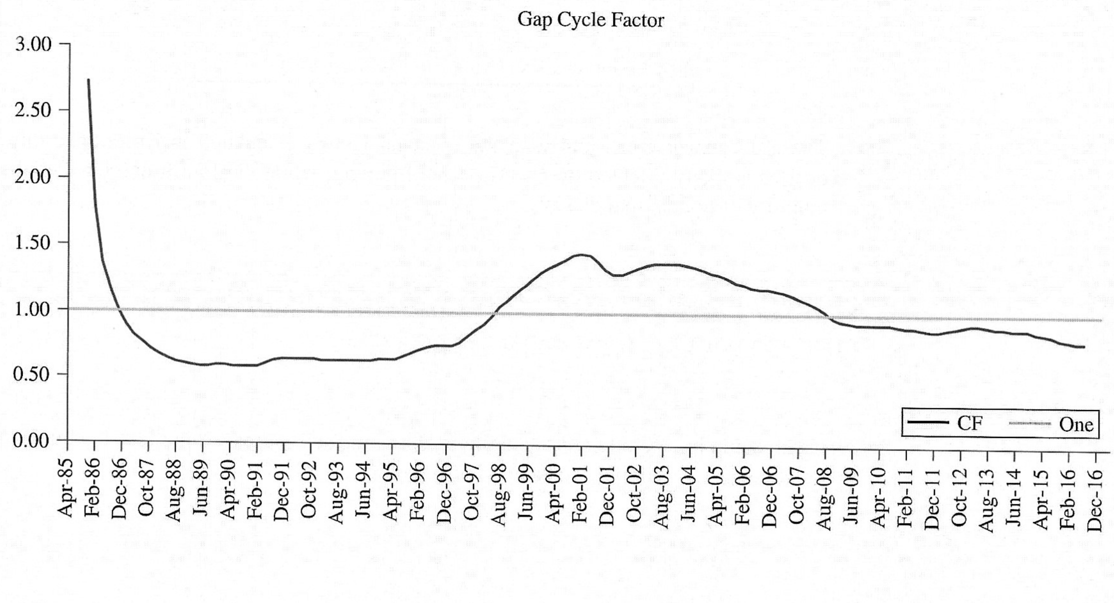
Date	GapSales ($M)	GapCMA	SF = GapSales/ GapCMA
Apr-85	106		
Jul-85	120		
Oct-85	182	168.75	1.08
Jan-86	240	181.13	1.33
Apr-86	160	192.13	0.83
Jul-86	165	204.75	0.81
Oct-86	225	218.38	1.03
Jan-87	298	231.38	1.29

(continued on next page)

(continued)

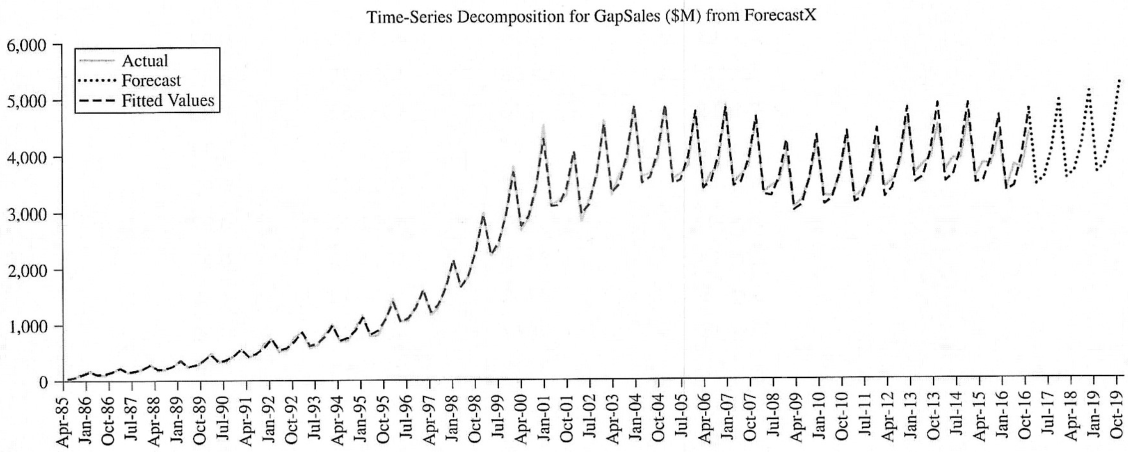
Date	GapSales ($M)	GapCMA	SF = GapSales/ GapCMA
.	.	.	.
.	.	.	.
.	.	.	.
Apr-13	3729	4060.50	0.92
Jul-13	3868	4055.75	0.95
Oct-13	3976	4042.63	0.98
Jan-14	4575	4062.38	1.13
Apr-14	3774	4076.00	0.93
Jul-14	3981	4092.13	0.97
Oct-14	3972	4094.13	0.97
Jan-15	4708	4069.13	1.16
Apr-15	3657	4044.38	0.90
Jul-15	3898	3989.63	0.98
Oct-15	3857	3921.88	0.98
Jan-16	4385	3888.63	1.13
Apr-16	3438	3875.38	0.89
Jul-16	3851	3873.50	0.99
Oct-16	3798		
Jan-17	4429		

5. The cycle factors are calculated as: CF = GapCMA/GapCMAT. The lighter line in the graph represents a horizontal line at one that is helpful when evaluating a cycle. The dark line is the cycle factor (CF) for each quarter.



Here we see only one trough at April 1989, where CF=0.58, and one peak at July 2001, where CF=1.39. In this case, even with the extended Gap sales data, there is not enough information to identify a cycle.

6. The ForecastX™ Audit Report graph of the time-series decomposition follows:



Everyone will not agree about whether the forecast looks reasonable. Many will think that the forecast looks high due to the declining cycle factors.

7. The seasonal indices from the ForecastX™ results are:

Date	Seasonal Indices
First Quarter	0.88
Second Quarter	0.90
Third Quarter	1.01
Fourth Quarter	1.22

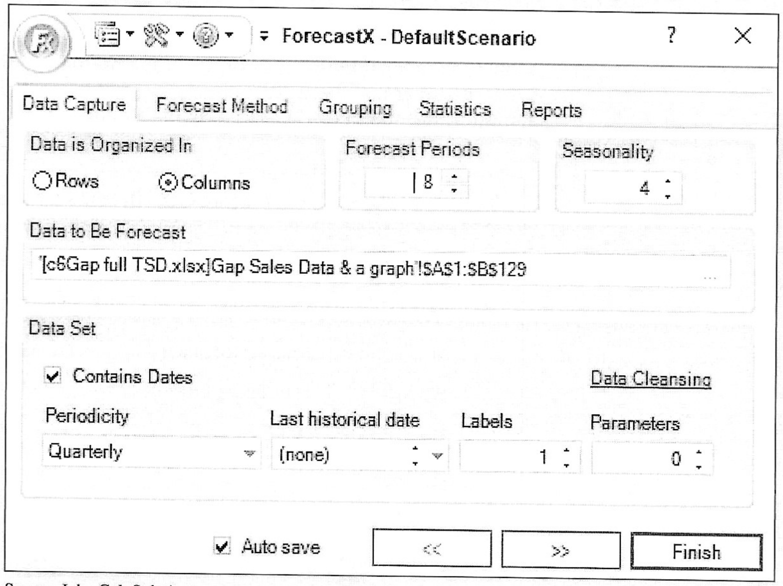
These do appear consistent with the seasonal factors calculated in Excel. The third quarter is likely affected by back-to-school buying, while the fourth quarter sees the majority of holiday purchases.

USING FORECASTX™ TO MAKE TIME-SERIES DECOMPOSITION FORECASTS

As usual, begin by opening your data file in Excel and select any cell in the data you want to forecast. In this example, we have selected cell B7.

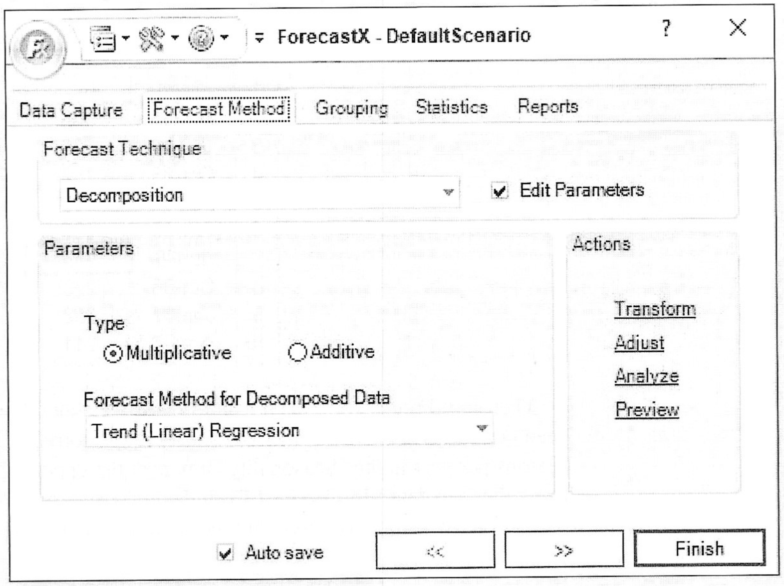
	A	B
1	**Date**	**Gap Sales ($M)**
2	Apr-85	106
3	Jul-85	120
4	Oct-85	182
5	Jan-86	240
6	Apr-86	160
7	Jul-86	165
8	Oct-86	225
9	Jan-87	298
10	Apr-87	211

Then start ForecastX™. In the **Data Capture** dialog box, check to be sure the data has been correctly identified. If the seasonality is not correctly identified, enter the number of seasons per year in the "Seasonality" box near the upper-right corner (four in this example using the Gap data). In **Forecast Periods,** make sure that you enter the number of periods that you want to forecast (eight in this example: a two-year forecast by quarters).



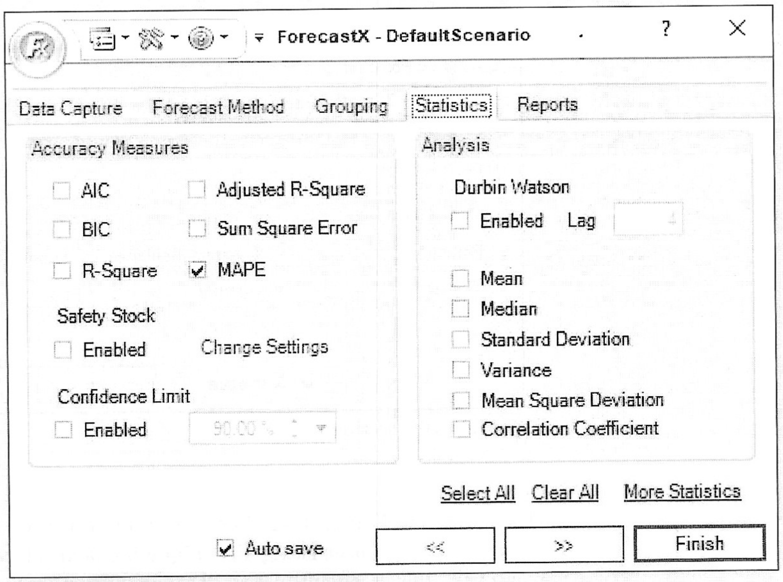
Source: John Galt Solutions

Then click the **Forecast Method** tab. In the **Forecast Method** dialog box, click the down arrow in the **Forecasting Technique** box and select **Decomposition** (note that in ForecastX™, this is simply called "Decomposition" rather than "Time-Series Decomposition"). Click **Multiplicative** and select **Trend (Linear) Regression** as the **Forecast Method for Decomposed Data.**



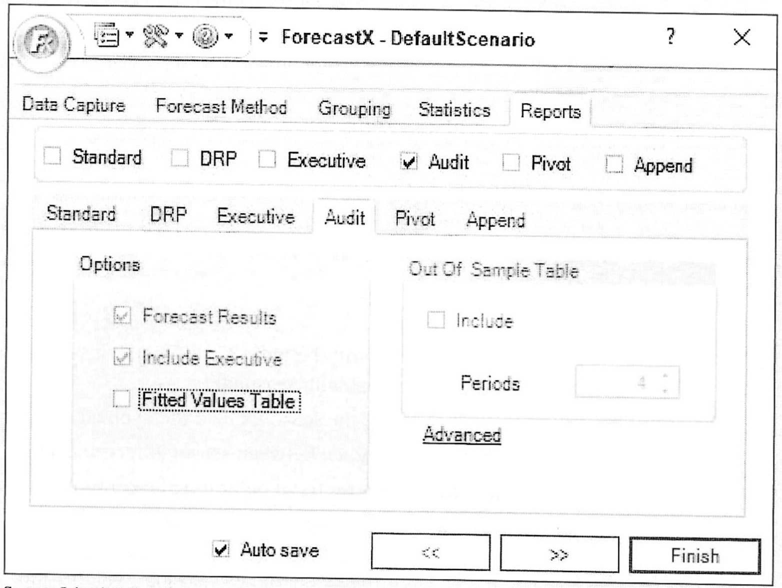
Source: John Galt Solutions

Then click the **Statistics tab.** In the Statistics dialog box, select the statistics that you desire. Remember that there are more choices if you click the **More Statistics** button near the bottom right. We have selected only the MAPE.



Source: John Galt Solutions

After selecting the statistics you want to see, click the **Reports** tab. In the **Reports** box, select those you want. The typical selection for decomposition would be the **Audit Report,** as shown here. The **Fitted Values Table** can be very long when you have an extended series for decomposition, so you may not want that long table when doing a time-series decomposition.



Source: John Galt Solutions

Then click the **Finish** button.

Suggested Readings

Aczel, Amir D.; and Jayavel Sounderpandian. "Time Series, Forecasting and Index Numbers." In *Complete Business Statistics.* 6th ed. Boston: McGraw-Hill/Irwin, 2006, pp. 582–602.

Austin, John S. "How to Use and Interpret Seasonal Factors." *Business Economics* 16, no. 4 (September 1981), pp. 40–42.

Campbell, Jeffrey R. "Entry, Exit, Embodied Technology, and Business Cycles." *Review of Economic Dynamics* 1 (1998), pp. 371–408.

Chatterjee, Satyajit. "From Cycles to Shocks: Progress in Business-Cycle Theory." *Business Review,* Federal Reserve Bank of Philadelphia (March/April 2000), pp. 27–37.

Chatterjee, Satyajit. "Productivity Growth and the American Business Cycle." *Business Review,* Federal Reserve Bank of Philadelphia, September/October 1995.

Espasa, Antoni; and Daniel Pena. "The Decomposition of Forecast in Seasonal ARIMA Models." *Journal of Forecasting* 14, no. 7 (December 1995), pp. 565–83.

Keating, Barry P. "Recession Forecasting: Using Leading Indicators and Decomposition." *The Journal of Business Forecasting*. Fall 2009, Volume 28, Number 3, pp. 23–30.

Layton, Allan P. "Dating and Predicting Phase Changes in the U.S. Business Cycle." *International Journal of Forecasting* 12, no. 3 (September 1996), pp. 417–28.

Majani, Bernard E. "Decomposition Methods for Medium-Term Planning and Budgeting." In *The Handbook of Forecasting*. Eds. Spyros Makridakis and Steven C. Wheelwright. New York: John Wiley & Sons, 1982, pp. 153–72.

Makridakis, Spyros; Steven C. Wheelwright; and Victor E. McGee. *Forecasting Methods and Applications.* 2nd ed. New York: John Wiley & Sons, 1983, pp. 130–78.

Sommers, Albert T. *The U.S. Economy Demystified.* Rev. ed. Lexington, MA: Lexington Books, 1988.

Temin, Peter. "The Causes of American Business Cycles: An Essay in Economic Historiography." In *Beyond Shocks: What Causes Business Cycles?* Eds. Jeffrey Fuhrer and Scott Schuh. Federal Reserve Bank of Boston, 1998.

Veloce, William. "An Evaluation of the Leading Indicators for the Canadian Economy Using Time Series Analysis." *International Journal of Forecasting* 12, no. 3 (September 1996), pp. 403–16.

Zellner, Arnold, ed. *Seasonal Analysis of Economic Time Series.* U.S. Department of Commerce, Bureau of the Census, 1978.

Exercises

1. Explain the similarity between how time-series decomposition and Winters' exponential smoothing deal with seasonality.
2. Discuss the trend, the seasonal, and the cyclical components.
3. What is the difference between seasonal factors and seasonal indices?
4. How is the long-term trend determined for a time-series decomposition model?
5. How do true cycles and the cycles typically found in business data differ?
6. Using your own words, write a description of each of the four components of the classic time-series decomposition technique. Avoid using mathematical relationships and technical jargon as much as possible so that your explanations can be understood by almost anyone.
7. Suppose that sales of a household appliance are reported to be 13,000 units during the first quarter of the year. The seasonal index for the first quarter is 1.24. Use this information to make a forecast of sales for the entire year. Actual sales for the year were 42,000 units. Calculate your percentage error for the year. What percentage error would result if you forecast sales for the year by simply multiplying the 13,000 units for the first quarter by 4?
8. In a time-series decomposition of sales (in millions of units), the following trend has been estimated:

$$\text{CMAT} = 12.315 + 0.196(T)$$

The seasonal indices have been found to be:

Quarter	Seasonal Index
1	1.27
2	1.02
3	0.73
4	0.98

For the coming year, the time index and cycle factors are:

Quarter	*T*	*CF*
1	21	1.01
2	22	1.04
3	23	1.06
4	24	1.04

a. From this information, prepare a forecast for each quarter of the coming year.

b. Actual sales for the year you forecast in part (*a*) were 17.2, 13.2, 10.8, and 14.2 for quarters 1, 2, 3, and 4, respectively. Use these actual sales figures along with your forecasts to calculate the mean absolute percentage error for the forecast period.

9. A tanning parlor located in a major shopping center near a large New England city has the following history of customers over the last four years (data are in hundreds of customers):

(c6p9)

	Mid-Month of Quarter				
Year	**Feb**	**May**	**Aug**	**Nov**	**Yearly Totals**
2012	3.5	2.9	2.0	3.2	11.6
2013	4.1	3.4	2.9	3.6	14.0
2014	5.2	4.5	3.1	4.5	17.3
2015	6.1	5.0	4.4	6.0	21.5

a. Construct a table in which you show the actual data (given in the table), the centered moving average, the centered moving-average trend, the seasonal factors, and the cycle factors for every quarter for which they can be calculated in years 2012 through 2015.

b. Look at the seasonal index for each quarter as calculated in ForecastX™. Do they make sense to you? Explain why or why not.

c. Make a forecast of the number of customers for each quarter of 2016.

d. The actual numbers of customers served per quarter in 2016 were 6.8, 5.1, 4.7, and 6.5 for quarters 1 through 4, respectively (numbers are in hundreds). Calculate the MAPE for 2016.

e. Using the results provided in the tables produced by ForecastX™, prepare a time-series plot of the actual data, the centered moving averages, the long-term trend, and the values predicted by your model for 2012 through 2016 (where data are available).

10. Carl Lipke is the marketing VP for a propane gas distributor. He would like to have a forecast of sales on a quarterly basis, and he has asked you to prepare a time-series decomposition model. The data for 2005 through 2016 follow:(c6p10)

Propane Gas Sales in Millions of Pounds (Total at End-Month of Each Quarter)

Year	March	June	September	December
2005	6.44	4.85	4.67	5.77
2006	6.22	4.25	4.14	5.34
2007	6.07	4.36	4.07	5.84

(continued on next page)

(continued)

Year	March	June	September	December
2008	6.06	4.24	4.20	5.43
2009	6.56	4.25	3.92	5.26
2010	6.65	4.42	4.09	5.51
2011	6.61	4.25	3.98	5.55
2012	6.24	4.34	4.00	5.36
2013	6.40	3.84	3.53	4.74
2014	5.37	3.57	3.32	5.09
2015	6.03	3.98	3.57	4.92
2016	6.16	3.79	3.39	4.51

a. To help Carl Lipke see how propane gas sales have varied over the 12-year period, prepare a time-series plot of the raw data and the deseasonalized data (i.e., the centered moving averages).

b. Use ForecastX™ to find seasonal indices for quarters 1 through 4. Write a short paragraph in which you explain to Carl Lipke exactly what these indices mean.

c. Plot the values of actual sales, the centered moving averages, and the trend. All of these can be found in your ForecastX™ results.

d. From ForecastX™, get a forecast for 2017Q1 through 2017Q4 based on the time-series decomposition model. Enter your forecast values into an Excel sheet that you set up like in the table shown below. Given the actual values shown in the table, calculate the mean absolute percentage error (MAPE) for 2017.

Date	Actual Sales	Forecast Sales	Error	Absolute Error	Absolute % Error
Mar-17	5.39				
Jun-17	3.56				
Sep-17	3.03				
Dec-17	4.03				
				MAPE =	

11. The Bechtal Tire Company (BTC) is a supplier of automotive tires for U.S. car companies. BTC has hired you to analyze its sales. For this problem, do all the work in ForecastX™ and be sure to request the MAPE in the statistics tab. Data from 1995Q1 through 2016Q4 are given in the following table (in thousands of units):

(c6p11) **BTC Sales of Tires**

Year	Q1	Q2	Q3	Q4
1995	2,029	2,347	1,926	2,162
1996	1,783	2,190	1,656	1,491
1997	1,974	2,276	1,987	2,425
1998	2,064	2,517	2,147	2,524

(continued on next page)

(continued)

Year	Q1	Q2	Q3	Q4
1999	2,451	2,718	2,229	2,190
2000	1,752	2,138	1,927	1,546
2001	1,506	1,709	1,734	2,002
2002	2,025	2,376	1,970	2,122
2003	2,128	2,538	2,081	2,223
2004	2,027	2,727	2,140	2,270
2005	2,155	2,231	1,971	1,875
2006	1,850	1,551	1,515	1,666
2007	1,733	1,576	1,618	1,282
2008	1,401	1,535	1,327	1,494
2009	1,456	1,876	1,646	1,813
2010	1,994	2,251	1,855	1,852
2011	2,042	2,273	2,218	1,672
2012	1,898	2,242	2,247	1,827
2013	1,669	1,973	1,878	1,561
2014	1,914	2,076	1,787	1,763
2015	1,707	2,019	1,898	1,454
2016	1,706	1,878	1,752	1,560

a. Write a report to Bechtal Tire Company in which you explain what a time-series decomposition analysis shows about its tire sales. Include in your discussion seasonal, cyclical, and trend components.

b. Show a time-series graph with the actual data and the values that the time-series decomposition model would predict for each quarter from 1995Q1 through 2017Q4 (some data will be missing for certain historical quarters, and, of course for 2017, you will have only the forecast values).

c. If actual sales for 2017 were Q1 = 1,445.1, Q2 = 1,683.8, Q3 = 1,586.6, and Q4 = 1,421.3, what MAPE would result from your 2017 forecast? How does this MAPE compare to the MAPE ForecastsX™ calculated for the historic period?

12. A regional supplier of jet fuel is interested in forecasting its sales. These sales data are shown for the period from 2002Q1 to 2017Q4 (data in billions of gallons):

(c6p12)

Jet Fuel Sales (Billions of Gallons)

Year	Q1	Q2	Q3	Q4
2002	23.86	23.97	29.23	24.32
2003	23.89	26.84	29.36	26.30
2004	27.09	29.42	32.43	29.17
2005	28.86	32.10	34.82	30.48
2006	30.87	33.75	35.11	30.00

(continued on next page)

(continued)

Year	Q1	Q2	Q3	Q4
2007	29.95	32.63	36.78	32.34
2008	33.63	36.97	39.71	34.96
2009	35.78	38.59	42.96	39.27
2010	40.77	45.31	51.45	45.13
2011	48.13	50.35	56.73	48.83
2012	49.02	50.73	53.74	46.38
2013	46.32	51.65	52.73	47.45
2014	49.01	53.99	55.63	50.04
2015	54.77	56.89	57.82	53.30
2016	54.69	60.88	63.59	59.46
2017	61.59	68.75	71.33	64.88

a. Prepare a time series graph of these data. What, if any, seasonal pattern do you see in the plot? Explain.

b. Use ForecastX^{TM} to make a time series decomposition forecast for 2018. Write a brief report explaining your forecast. Include a graph of the fitted values, the forecast values, and the actual sales.

c. Develop two other forecasts of jet fuel sales using the following methods:

1. A Winters' exponential smoothing model; and
2. A regression model using just time and quarterly dummy variables.

Compare the MAPEs for the three models you have developed, and comment on what you like or dislike about each of the three models for this application.

13. The following table contains quarterly data on Upper Midwest car sales (CS) in thousands for 1996Q1 through 2016Q4:

(c6p13) **Upper Midwest Car Sales (CS)**

Year	Q1	Q2	Q3	Q4
1996	407.6	431.5	441.6	306.2
1997	328.7	381.3	422.6	369.4
1998	456.3	624.3	557.5	436.7
1999	485.0	564.3	538.3	412.5
2000	555.0	682.7	581.3	509.7
2001	662.7	591.1	616.9	529.7
2002	641.2	632.7	576.6	475.0
2003	542.8	558.9	581.7	537.8
2004	588.1	626.5	590.9	580.1
2005	589.2	643.2	593.9	612.2

(continued on next page)

(continued)

Year	Q1	Q2	Q3	Q4
2006	586.1	699.4	734.4	753.8
2007	691.6	793.4	864.9	840.8
2008	653.9	754.8	883.6	797.7
2009	722.2	788.6	769.9	725.5
2010	629.3	738.6	732.0	598.8
2011	603.9	653.6	606.1	539.7
2012	461.3	548.0	548.4	480.4
2013	476.6	528.2	480.4	452.6
2014	407.2	498.5	474.3	403.7
2015	418.6	470.2	470.7	375.7
2016	371.1	425.5	397.3	313.5

a. Prepare a time-series plot of Upper Midwest car sales from 1996Q1 through 2016Q4.

b. Use ForecastX™ to do a time-series decomposition forecast for 2017 (be sure to request the MAPE). In the results, you see the seasonal indices. Do they make sense? Why or why not?

c. ForecastX™ calculated the historic MAPE as a measure of fit. Write a short explanation of what this MAPE means to a manager.

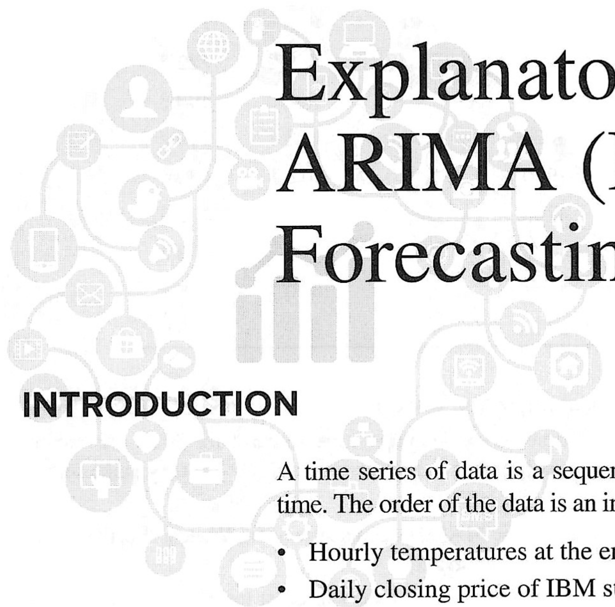
d. Now calculate the MAPE for the 2017Q1–2017Q4 forecast horizon as a measure of accuracy, given that the actual values of CS for 2017 were:

2017Q1	301.1
2017Q2	336.7
2017Q3	341.8
2017Q4	293.5

e. Prepare a Winters' exponential smoothing forecast of CS using data from 1996Q1 through 2016Q4 as the basis for a forecast of 2017Q1–2017Q4. Compare these results in terms of fit and accuracy with the results from the time-series decomposition forecast.

Chapter **Seven**

Explanatory Models 3. ARIMA (Box-Jenkins) Forecasting Models



©VLADGRIN/Getty Images

INTRODUCTION

A time series of data is a sequence of numerical observations naturally ordered in time. The order of the data is an important part of the data. Some examples would be:

- Hourly temperatures at the entrance to Grand Central Station
- Daily closing price of IBM stock
- Weekly automobile production by the Chevrolet Division of General Motors
- Data from an individual firm: sales, profits, inventory, back orders
- An electrocardiogram

When a forecaster examines time-series data, two questions are of paramount importance:

1. Do the data exhibit a discernible pattern?
2. Can this pattern be exploited to make meaningful forecasts?

We have already examined some time-series data by using regression analysis to relate sequences of data to explanatory variables. Sales (as the dependent variable), for instance, might be forecast by using the explanatory (or independent) variables of product price, personal income of potential purchasers, and advertising expenditures by the firm. Such a model is a structural or causal forecasting model that requires the forecaster to know in advance at least some of the determinants or predictors of sales. But in many real-world situations, we do not know the determinants of the variable to be forecast, or data on these predictor variables are not readily available. It is in just these situations that the ARIMA technique has a decided advantage over standard regression models. ARIMA may also be used as a benchmark for other forecasting models; we could use an ARIMA model, for example, as a benchmark for comparison with our best structural regression model. The acronym ARIMA stands for autoregressive integrated moving