

統計學

Spring 2026

政治大學統計系余清祥
2026年04月28日
第十七章：時間數列與預測
<http://csyue.nccu.edu.tw>



Chapter Contents

- 17.1 Time Series Patterns
- 17.2 Forecast Accuracy
- 17.3 Moving Averages and Exponential Smoothing
- 17.4 Trend Projection
- 17.5 Seasonality and Trend
- 17.6 Time Series Decomposition
- Summary

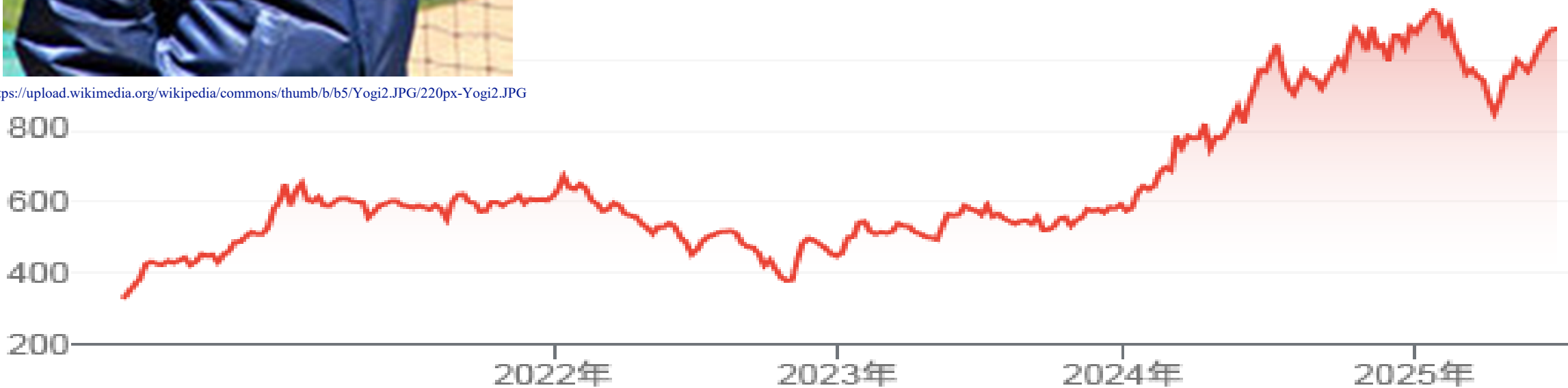
**“It’s tough to make predictions, especially about the future.”
— Yogi Berra**



<https://upload.wikimedia.org/wikipedia/commons/thumb/b/b5/Yogi2.JPG/220px-Yogi2.JPG>



<https://www.relexsolutions.com/resources/measuring-forecast-accuracy>



Introduction

Forecasting methods can be classified as qualitative or quantitative.

- Qualitative methods involve the use of expert judgment to develop forecasts when historical data on the variable being forecast is either unavailable or not applicable.
- Quantitative methods are used when past information about the forecast variable is available, the information can be quantified, and the past may be assumed as a prologue.

In this chapter, we will focus exclusively on quantitative forecasting methods.

When using historical data 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.

We begin the chapter by discussing the various time series patterns. Then, we will introduce different measurements of forecast accuracy, and related advantages and disadvantages.

Later in the chapter, we will learn to forecast time series patterns that are stationary or have a linear trend, and how to incorporate both trend and seasonality into a forecasting model.

Time Series vs. Causal Forecasting Methods

- *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.
- Looking at regression analysis as a forecasting tool, we can view the time series value that we want to forecast as the dependent variable.

If we can identify a good set of related independent, or explanatory, variables, we may be able to develop an estimated regression equation for forecasting the time series.

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.

We leave the discussion on causal forecasting methods to more advanced texts on forecasting.

17-1 Time Series Patterns

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.

The data pattern is important in understanding how the time series behaved in the past.

If such behavior can be expected to continue in the future, we can use it to guide us in selecting an appropriate forecasting method.

To identify the underlying pattern in the data, the 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 graphed on the horizontal axis and the time series variable on the vertical axis.

The objective of time series analysis is to uncover a pattern in the historical data or time series. Later in the chapter, we will learn to extrapolate the pattern into the future (forecast.)

Trend Component

- Tendency of Data to show long term upward or downward movement of data
- Represent Mean Rate of change with respect to time
- Example : Fluctuation in different marked commodity i.e., gold, stock price, exchange Rates etc.

Seasonality Component

- Repetitive Upward and Downward movement from the trend that occurs within calendar year at fixed interval where periodicity is constant
- Reasons : Festivals, Custom within the societies, Business practices etc.

Cyclical Component

- Fluctuation around the trend line at the random interval where periodicity is not constant (more than years interval)
- Examples: Macro-economic changes cause recession, unemployment etc.

Irregular Component

- White noise or random uncorrelated changes that follows normal distribution (mean=0 and constant variance)

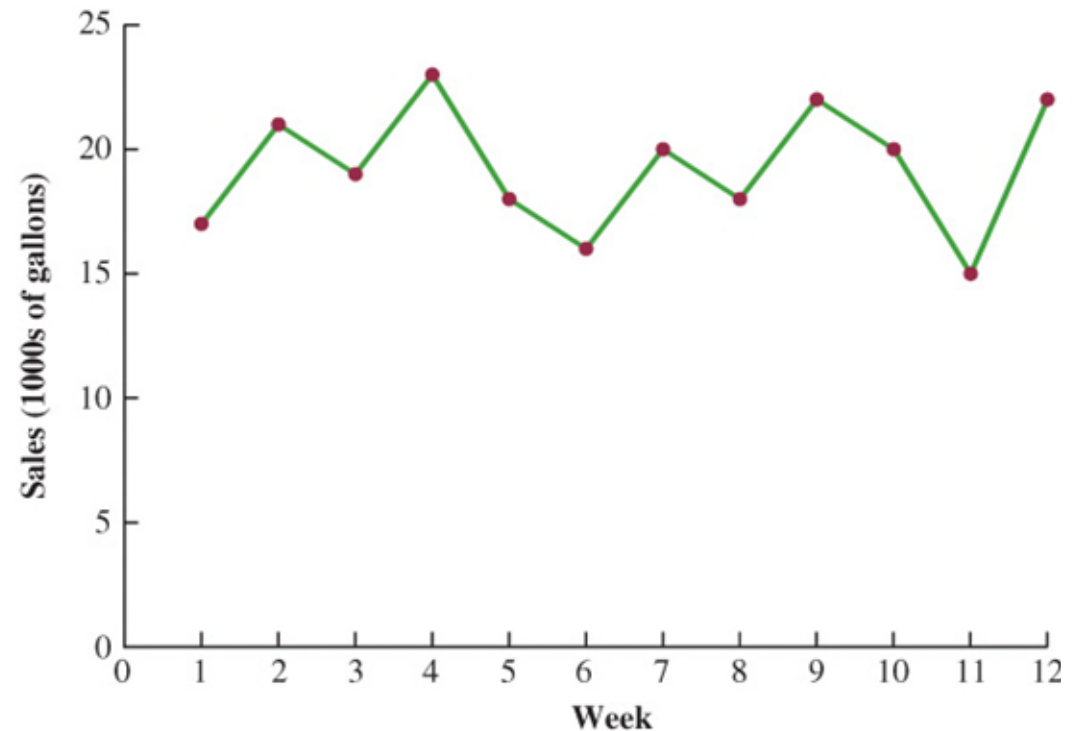
17-1 Horizontal Pattern

A horizontal pattern exists when the data fluctuate randomly around a constant mean over time. To illustrate a time series with a horizontal pattern, consider the 12 weeks of data in the DATAfile *Gasoline*.

These data show the number of gallons of gasoline (in 1000s) sold by a gasoline distributor in Bennington, Vermont, over the past 12 weeks.

The time series plot for these data shows 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.



17-1 Stationary Time Series

The term **stationary time series** 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 with random fluctuations. However, simply observing a horizontal pattern is not sufficient evidence to conclude that the time series is stationary.

In fact, changes in business conditions often result in a time series with a horizontal pattern that shifts to a new level at some point in time.

More advanced texts discuss procedures for determining if a time series is stationary and provide methods for transforming a time series that is nonstationary into a stationary series.

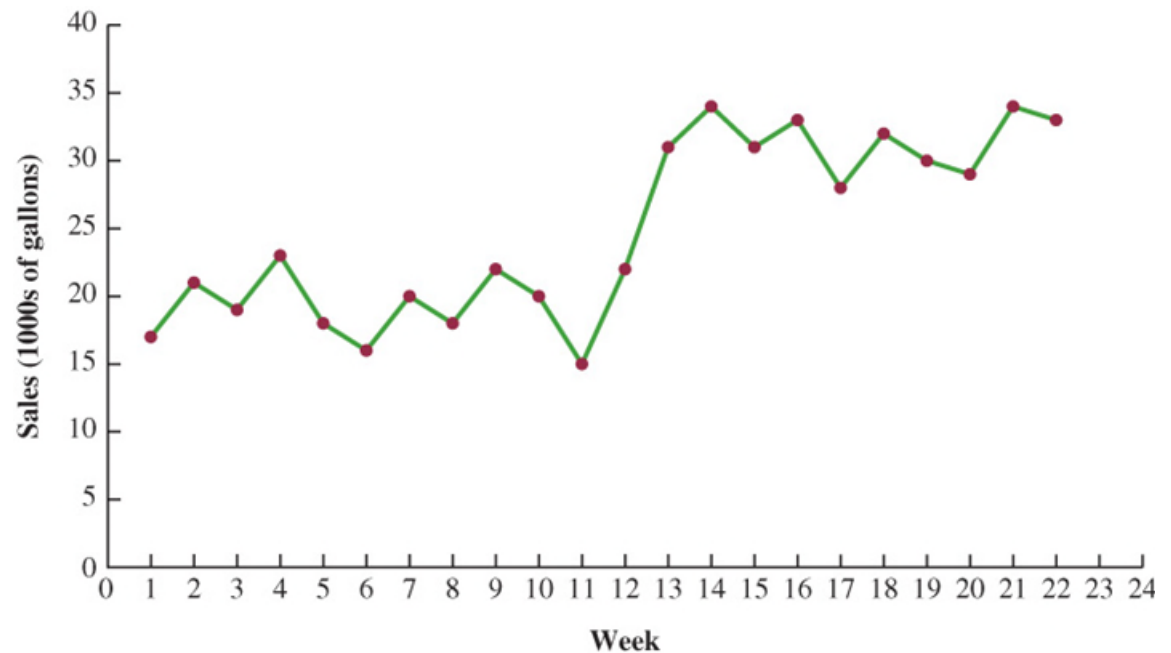
17-1 Shift in Horizontal Pattern

To illustrate a shift in horizontal pattern, consider the 22 weeks included in the DATAfile *Gasolinerevised*.

Here, the gasoline distributor signs a contract with the Vermont State Police beginning on week 13, which is reflected in the time series plot shown to the side.

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.



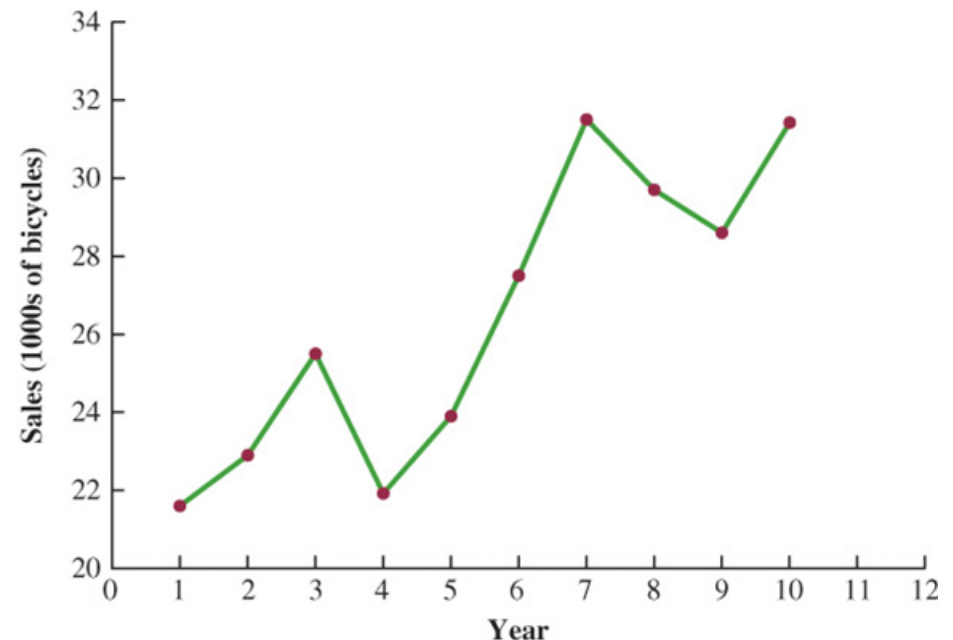
17-1 Linear Trend Pattern

A time series that exhibit show gradual shifts or movements to relatively higher or lower values over a longer period of time it is said to have a **trend pattern**.

A trend is often the result of long-term factors such as population increases or decreases or changes in consumer preferences.

To illustrate a time series with a *linear trend pattern*, consider the 10 years of sales for a particular manufacturer over the past 10 years, and included in the DATAfile *Bicycle*.

The time series plot shows some up and down movement over the past 10 years, but the time series seems also to have a systematically increasing or upward trend.



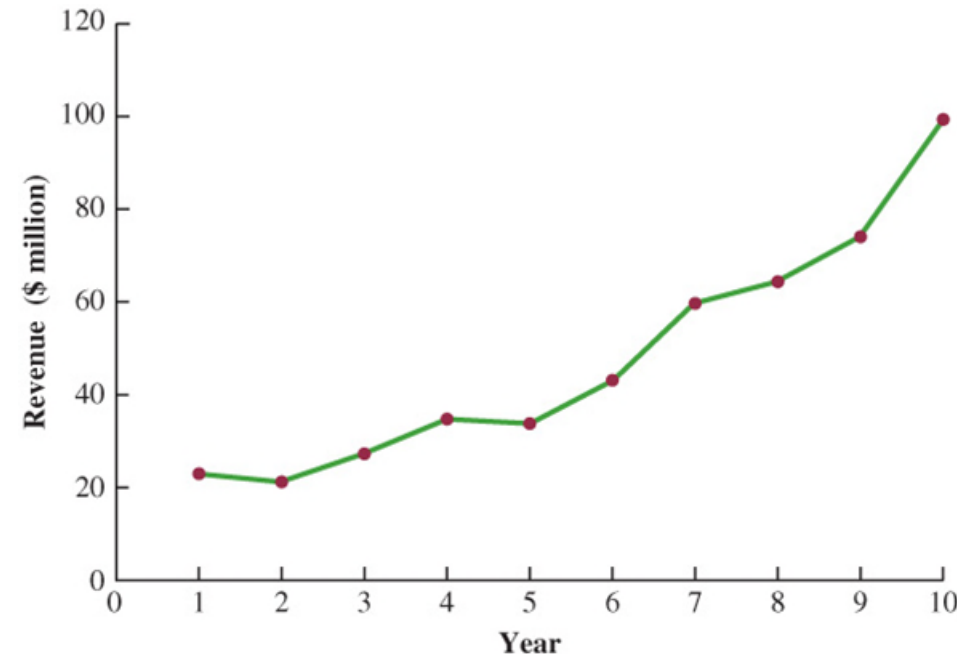
17-1 Non-Linear Trend Pattern

Sometimes, a trend can be described better by a *non-linear trend pattern*.

For instance, consider the DATAfile *Cholesterol* containing the sales revenue for a cholesterol drug since the company won FDA approval for the drug 10 years ago.

The time series plot included here shows a nonlinear increase in which the rate of change of revenue does not increase by a constant amount from one year to the next.

Non-linear trend relationships such as this one, in which the revenue appears to be growing exponentially are appropriate when the percentage change from one period to the next is relatively constant.



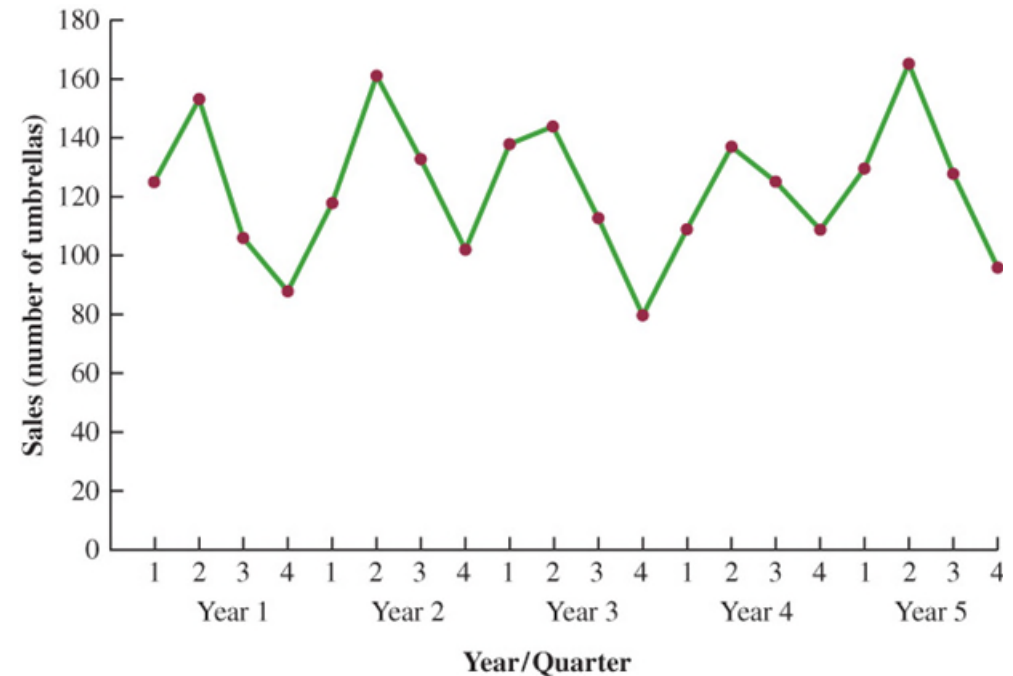
17-1 Seasonal Pattern

A **seasonal pattern** is recognized by observing a recurring pattern over successive periods of time.

Time series data can also exhibit seasonal patterns of less than one year in duration, such as daily traffic flow or weekly restaurant sales.

As an example of a seasonal pattern, consider the quarterly sales of umbrellas at a clothing store over the past five years and included in the DATAfile *Umbrella*.

The time series plot follows a horizontal pattern with seasonal fluctuations within each year, with the highest sales occurring on the second quarter and bottoming on the fourth quarter of each year.



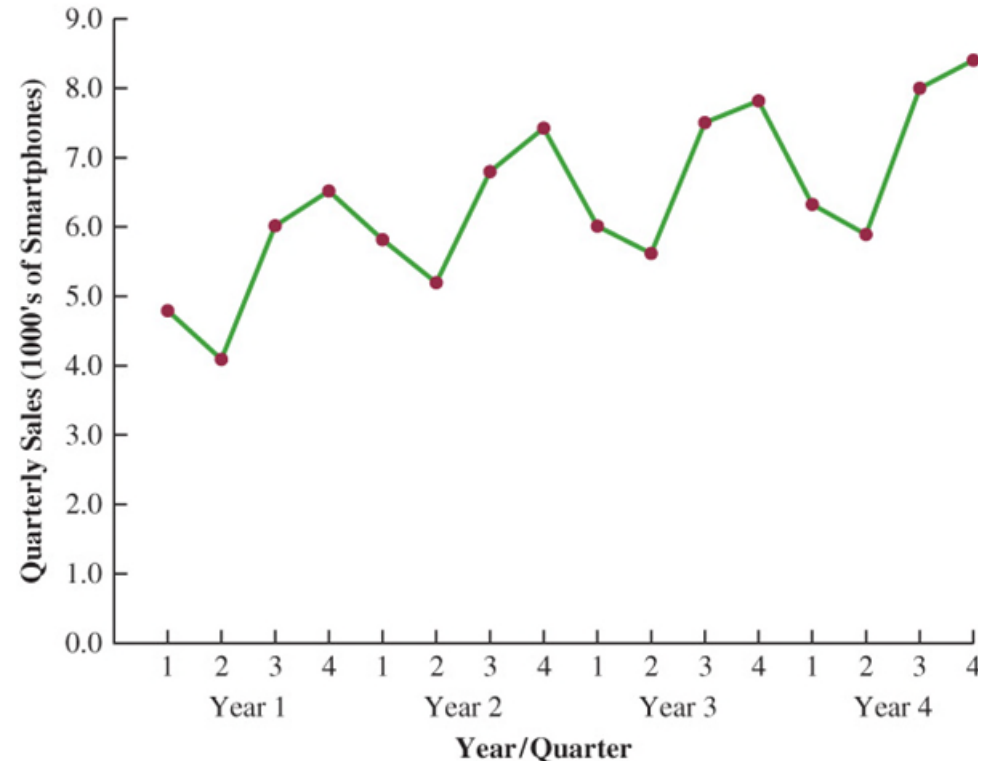
17-1 Trend and Seasonal Pattern

The smartphone sales data for a particular manufacturer over the past four years included in the DATAfile *SmartPhoneSales*.

The included time series plot exhibits not only an increasing trend, but it also indicates that sales are lowest in the second quarter of each year and highest in quarters 3 and 4.

Thus, we conclude that a seasonal pattern also exists for smartphones sales.

In such cases, we need to use a forecasting method able to deal with both trend and seasonality.



17-1 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.

Economic time series commonly exhibit cyclical behavior with regular runs of observations below and above the trend line.

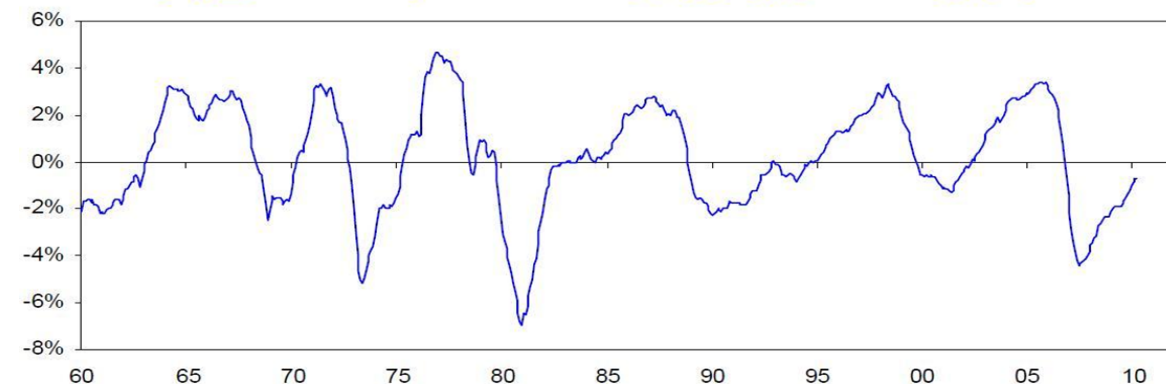
Often, the cyclical component of a time series is due to multi-year business cycles.

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).

Because business cycles are extremely difficult, if not impossible, to forecast, 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.

美國1960年~2010年期間的GDP缺口



17-1 Selecting a Forecasting Method

Because the underlying pattern in the time series is an important factor in selecting a forecasting method, a time series plot should be one of the first analytic tools employed when trying to determine which forecasting method to use.

When we see a horizontal pattern, we need to select an appropriate method.

Similarly, when we observe a trend in the data, we need to use a forecasting method that has the capability to handle the trend effectively.

In the next two sections, we illustrate methods for assessing forecast accuracy and consider forecasting models that can be used in situations for which the underlying pattern is horizontal. In other words, no trend or seasonality is present.

We then consider methods appropriate when trend and/or seasonality are present in the data.

17-2 Forecast Error

We begin developing forecasts for the gasoline time series included in the DATA file *Gasoline* using the simplest of all the forecasting methods:

We use the most recent week's sales volume as the forecast for the next week.

Because of its simplicity, this method is often referred to as a naïve forecasting method (column labeled Forecast in the table.)

The key concept associated with measuring forecast accuracy is **forecast error**. If we denote Y_t and \hat{Y}_t as the actual and forecasted values of the time series for period t , respectively, the forecasting error for period t is

$$e_t = Y_t - \hat{Y}_t$$

That is, the forecast error for time period t is the difference between the actual and the forecasted values for period t .

Week	Time Series Value	Forecast	Forecast Error
1	17		
2	21	17	4
3	19	21	-2
4	23	19	4
5	18	23	-5
6	16	18	-2
7	20	16	4
8	18	20	-2
9	22	18	4
10	20	22	-2
11	15	20	-5
12	22	15	<u>7</u>
		Total	5

17-2 Mean Absolute Error

The **mean absolute error**, denoted MAE (or MAD for deviation), is a measure of forecast accuracy that avoids the problem of positive and negative forecast errors offsetting one another. MAE is the average of the absolute values of the forecast errors:

$$MAE = \frac{\sum_{t=k+1}^n |e_t|}{n - k}$$

The table to the right shows that the sum of the absolute values of the forecast errors is 41. Thus

$$MAE = 41/11 = 3.73$$

As we do not have any data to produce a naïve forecast for the first week of this time series, we must adjust our calculations in the numerator and denominator, accordingly.

For the naïve forecast of the gasoline data, we have:

$$k = 1 \quad \text{and} \quad n - k = 11$$

Week	Time Series Value	Forecast	Forecast Error	Abs. Value Forecast Error
1	17			
2	21	17	4	4
3	19	21	-2	2
4	23	19	4	4
5	18	23	-5	5
6	16	18	-2	2
7	20	16	4	4
8	18	20	-2	2
9	22	18	4	4
10	20	22	-2	2
11	15	20	-5	5
12	22	15	<u>7</u>	<u>7</u>
		Total	5	41

17-2 Mean Squared Error

Another measure that avoids the problem of positive and negative 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.

$$MSE = \frac{\sum_{t=k+1}^n e_t^2}{n - k}$$

The table to the right shows that the sum of the squared forecast errors is 179. Thus

$$MSE = 179/11 = 16.27$$

Because the size of MAE and MSE depends upon the scale of the data, it is difficult to make comparisons for different time intervals or across different time series.

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17	4	16
3	19	21	-2	4
4	23	19	4	16
5	18	23	-5	25
6	16	18	-2	4
7	20	16	4	16
8	18	20	-2	4
9	22	18	4	16
10	20	22	-2	4
11	15	20	-5	25
12	22	15	-7	49
		Total	5	179

17-2 Mean Absolute Percentage Error

To make comparisons that do not depend upon the scale of the data, we need to work with relative or percentage error measures. One such measure is the **mean absolute percentage error**, denoted as MAPE, and calculated as

$$MAPE = \frac{\sum_{t=k+1}^n \left| \left(\frac{e_t}{Y_t} \right) 100 \right|}{n - k}$$

Where $(e_t/Y_t)100$ is the percentage error for the forecast at time t .

The table to the right shows that the sum of the absolute values of the percentage errors is 211.69. Thus

$$MAPE = \frac{211.69}{11} = 19.24\%$$

Week	Time Series Value	Forecast	Abs. Value		Percentage Error	Abs. Value Percentage Error
			Forecast Error	Forecast Error		
1	17					
2	21	17	4	4	19.05	19.05
3	19	21	-2	2	-10.53	10.53
4	23	19	4	4	17.39	17.39
5	18	23	-5	5	-27.78	27.78
6	16	18	-2	2	-12.50	12.50
7	20	16	4	4	20.00	20.00
8	18	20	-2	2	-11.11	11.11
9	22	18	4	4	18.18	18.18
10	20	22	-2	2	-10.00	10.00
11	15	20	-5	5	-33.33	33.33
12	22	15	<u>7</u>	<u>7</u>	<u>31.82</u>	<u>31.82</u>
		Total	5	41	1.19	211.69

17-2 Compare Forecasting Models

Forecast accuracy helps measure how well the forecasting method forecasts historical values of the time series but cannot be used to measure the accuracy of future forecasts.

When comparing different forecasts, and we think that the historical pattern will continue into the future, we should select the forecasting method that fits best the historical data.

As an example, suppose we forecast the gasoline time series data using the average of past values. Thus, if we use the notation F_t for the forecast at time period t , we have:

$$F_2 = Y_1 = 17 \quad F_3 = (Y_1 + Y_2)/2 = 19 \quad F_4 = (Y_1 + Y_2 + Y_3)/3 = 19 \quad \text{and so on}$$

We can now proceed with the calculations of MAE, MSE, and MAPE as we did for the naïve method. The results are shown to the right.

When comparing the accuracy of the two forecasting methods, we see that the average of past values method is more accurate for each measure of forecasting accuracy.

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

17-3 Smoothing Methods

In this section we discuss three forecasting methods that are appropriate for a time series with a horizontal pattern:

- Moving averages
- Weighted moving averages
- Exponential smoothing.

These methods are also capable of adapting well to changes in the level of a horizontal pattern such as what we saw with the extended gasoline sales time series.

However, without modification, they are not appropriate when considerable trend, cyclical, or seasonal effects are present.

Because the objective of each of these methods is to “smooth out” 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.

17-3 Moving Average

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. A **moving average** forecast of order k is:

$$F_{t+1} = \frac{\sum(\text{most recent } k \text{ data values})}{k} = \frac{\sum_{i=t-k+1}^t Y_i}{k} = \frac{Y_{t-k+1} + \dots + Y_{t-1} + Y_t}{k}$$

Where, F_{t+1} is the forecast of the time series for period $t + 1$, Y_t is actual value of the time series in period t , and k is the number of periods used to generate the forecast.

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.

Thus, the periods over which the average is calculated *move* with each ensuing period.

To use moving averages to forecast a time series, we must first select the order k , or number of time series values to be included in the moving average.

The value of k is selected based on the number of past values that are considered relevant. The greater the number of relevant past values, the larger the value selected for k .

17-3 Moving Average of the Gasoline Time Series

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.

- A smaller value of k will track shifts in a time series more quickly (the naïve approach is actually a moving average with $k = 1$).
- Larger values of k will be more effective in smoothing out random fluctuations.
- Thus, managerial judgment based on an understanding of the behavior of a time series is helpful in choosing an appropriate value of k .

A summary of a three-week ($k = 3$) moving average calculations for the gasoline time series data follows (*see notes.)

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
		Total	0	24	92	-20.79	129.21

17-3 Moving Average Forecast Accuracy

Using the three-week moving average calculations from the last slide, the values for the three measures of forecast accuracy are:

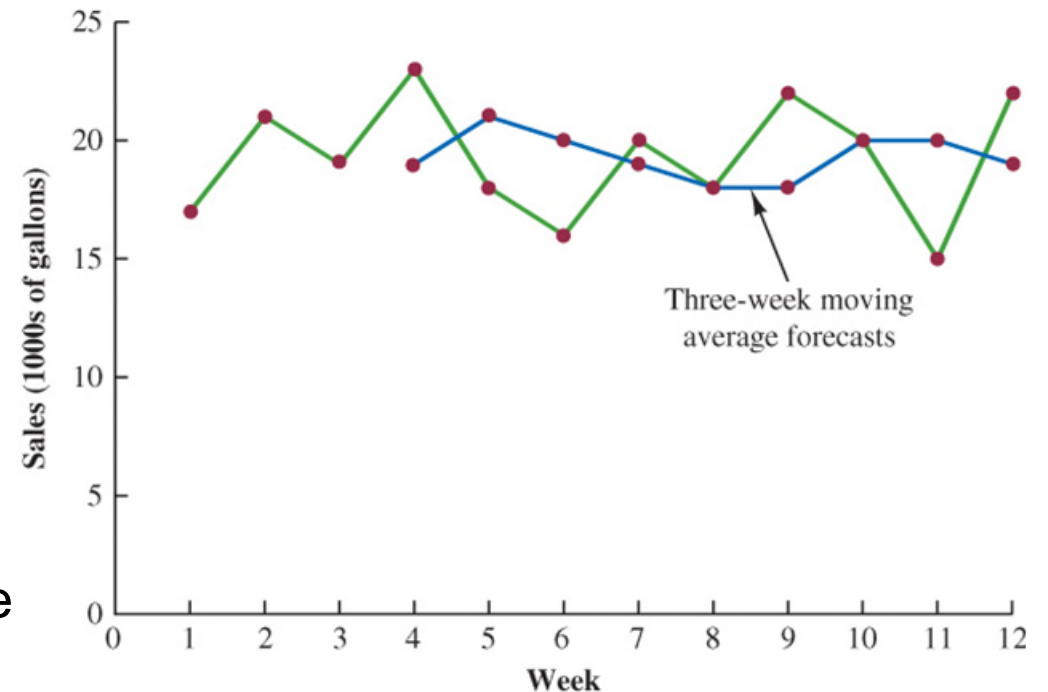
$$MAE = \frac{\sum_{t=4}^{12} |e_t|}{12 - 3} = \frac{24}{9} = 2.67$$

$$MSE = \frac{\sum_{t=4}^{12} e_t^2}{12 - 3} = \frac{92}{9} = 10.22$$

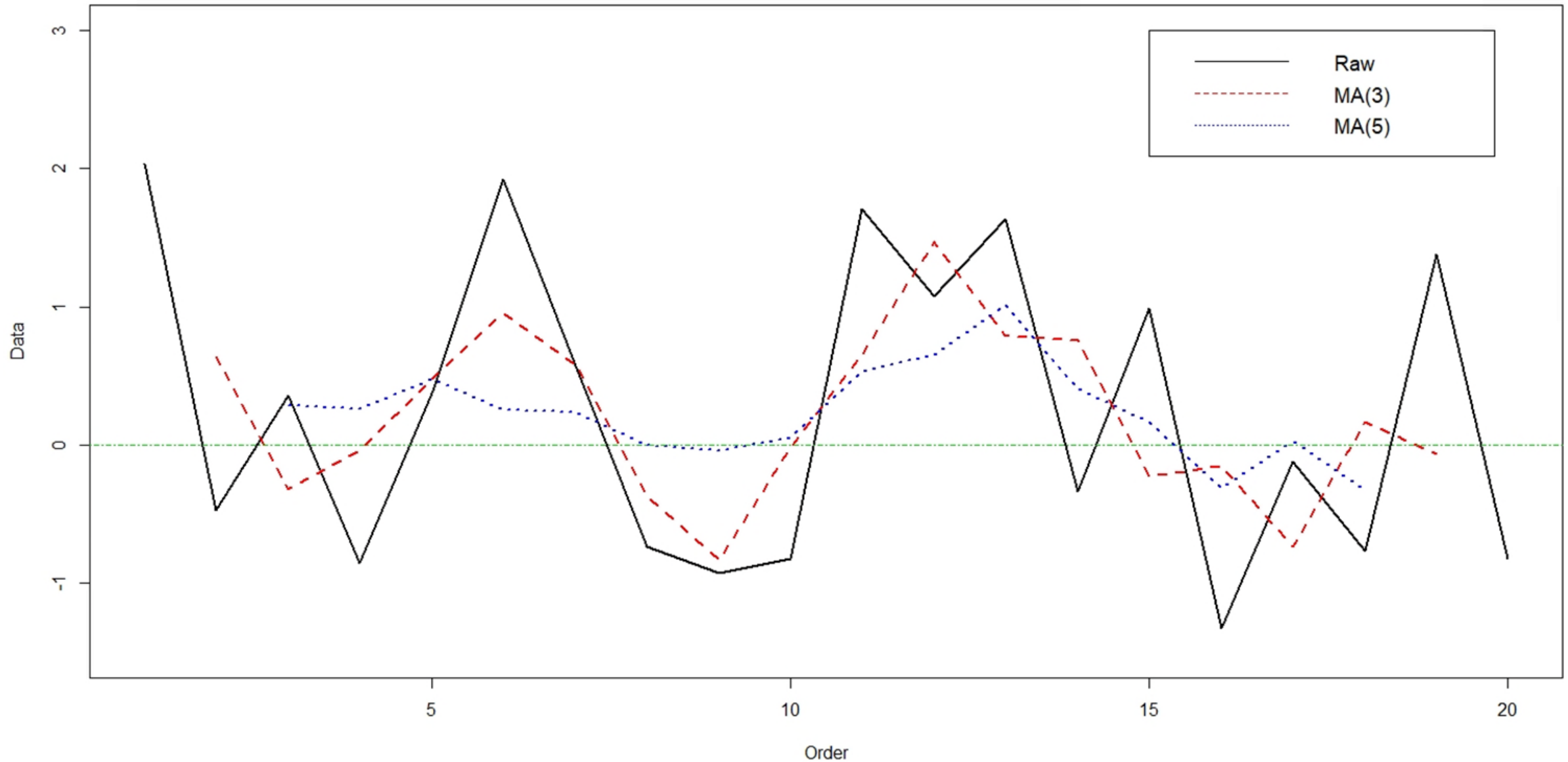
$$MAPE = \frac{\sum_{t=4}^{12} \left| \left(\frac{e_t}{\bar{Y}_t} \right) 100 \right|}{12 - 3} = \frac{129.21}{9} = 14.36\%$$

Thus, in each case the three-week moving average approach provides more accurate forecasts than simply using the most recent observation as the forecast.

Trial and error reveals that forecast accuracy may be further improved to a minimum value of $MSE = 6.79$ with a moving average of order $k = 6$.



Example of Moving Average, $N(0,1)$ Random Numbers



17-3 Weighted Moving Average

One variation of the moving averages method, known as **weighted moving averages**, involves selecting a different weight for each data value in the moving average and then computing the forecast as a weighted average of the most recent k values.

$$F_{t+1} = w_t Y_t + w_{t-1} Y_{t-1} + \cdots + w_{t-k+1} Y_{t-k+1}$$

Where F_{t+1} is the forecast of the time series for period $t + 1$, Y_t is the actual value of the time series in period t , w_t is the weight applied to Y_t , and k is the number of periods used to generate the forecast.

Generally, the most recent observation receives the largest weight, and the weight decreases with the relative age of the data values.

For a three-week weighted moving average, we could assign weights of $w_t = 3/6$, $w_{t-1} = 2/6$, and $w_{t-2} = 1/6$, with the sum of the weights equal to 1.

Using these parameters, the forecast for week 4 of the gasoline time series data would be:

$$F_4 = 3/6 (19) + 2/6 (21) + 1/6 (17) = 19.33$$

17-3 Weighted Moving Average Forecast Accuracy

In the weighted moving averages method, we 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 preferable.

The selected weights must be non-negative and their sum must equal 1.

To determine whether one combination of data values and weights provides a more accurate forecast than another one, we recommend using MSE as the measure of forecast accuracy.

That is, if we assumed that the combination that is best for the past will also be best for the future, we would use the combination of data values and weights that minimized MSE for the historical time series to forecast the next value in the time series.

17-3 Exponential Smoothing

Exponential smoothing 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.

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

Where F_{t+1} is the forecast of the time series for period $t + 1$, Y_t is the actual value of the time series in period t , F_t is the forecast of the time series for period t , and α is the smoothing constant ($0 \leq \alpha \leq 1$).

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 $1 - \alpha$.

17-3 Exponential Smoothing as Weighted Average

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 the concept with a time series involving three periods of data: Y_1 , Y_2 , and Y_3 .

We begin by letting 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

$$F_2 = \alpha Y_1 + (1 - \alpha)F_1 = \alpha Y_1 + (1 - \alpha)Y_1 = Y_1$$

Thus, the 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

$$F_4 = \alpha Y_3 + (1 - \alpha)F_3 = \alpha Y_3 + \alpha(1 - \alpha)[\alpha Y_2 + (1 - \alpha)Y_1] = \alpha Y_3 + \alpha(1 - \alpha)Y_2 + (1 - \alpha)^2 Y_1$$

We see that F_4 is a weighted average of Y_1 , Y_2 , and Y_3 , and the sum of the weights equals 1.

17-3 Exponential Smoothing of the Gasoline Time Series

Let us again consider an application of the exponential smoothing approach to forecasting to the gasoline sales time series.

Because there is no forecast available for Week 1, it follows that: $F_2 = Y_1 = 17.00$

We can add each actual value Y_t and each previously calculated forecast F_t to the calculation of F_{t+1} with a $\alpha = 0.2$.

$$F_3 = 0.2(21) + (1 - 0.2)(17.00) = 17.80$$

$$F_4 = 0.2(19) + (1 - 0.2)(17.80) = 18.04$$

⋮

$$F_{12} = 0.2(15) + (1 - 0.2)(19.35) = 18.48$$

A summary of the exponential smoothing calculations for the gasoline time series data with smoothing constant $\alpha = 0.2$ is shown.

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
		Total	10.92	98.80

17-3 Exponential Smoothing Forecast Accuracy

Although any value of α between 0 and 1 is acceptable, insight into choosing a more suitable value for α can be obtained by rewriting the basic exponential smoothing model as follows:

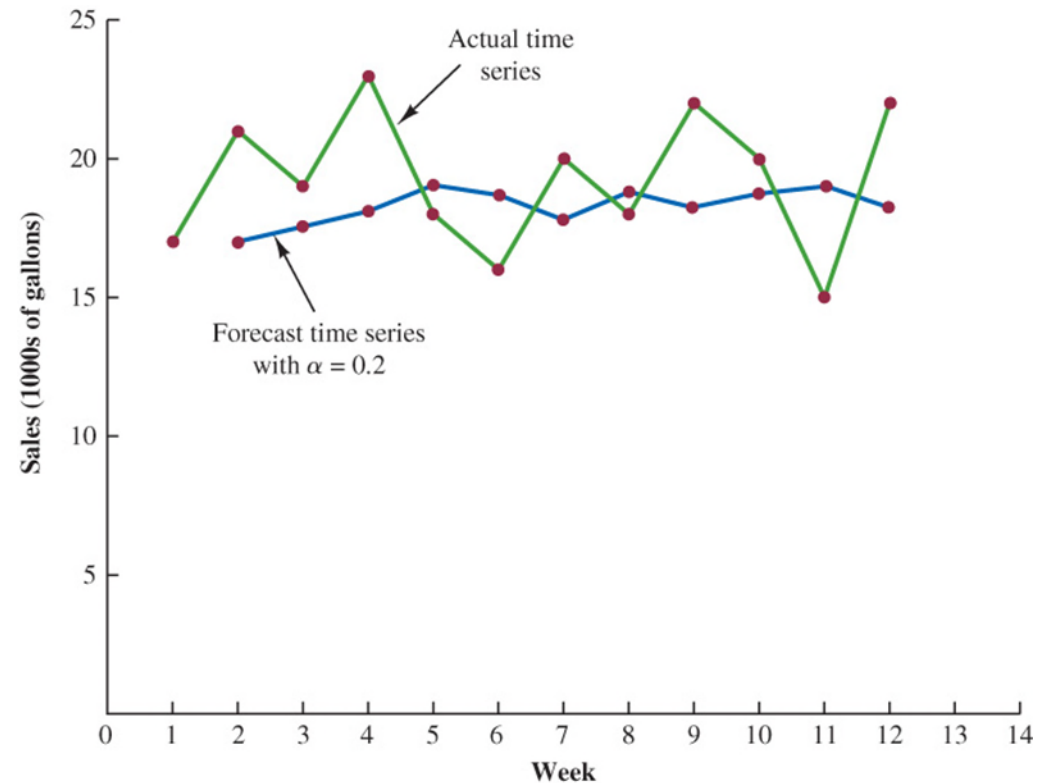
$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

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

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

Thus, the new forecast is equal to the previous forecast plus the smoothing constant times the most recent forecast error.

It follows that if the time series contains substantial random variability, a small value of the smoothing constant is preferred (*see notes.)



17-4 Linear Trend Regression

Regression analysis may be used to forecast a time series with a linear trend as is the case for the previously introduced bicycle sales time series (DATAfile: *Bicycle*).

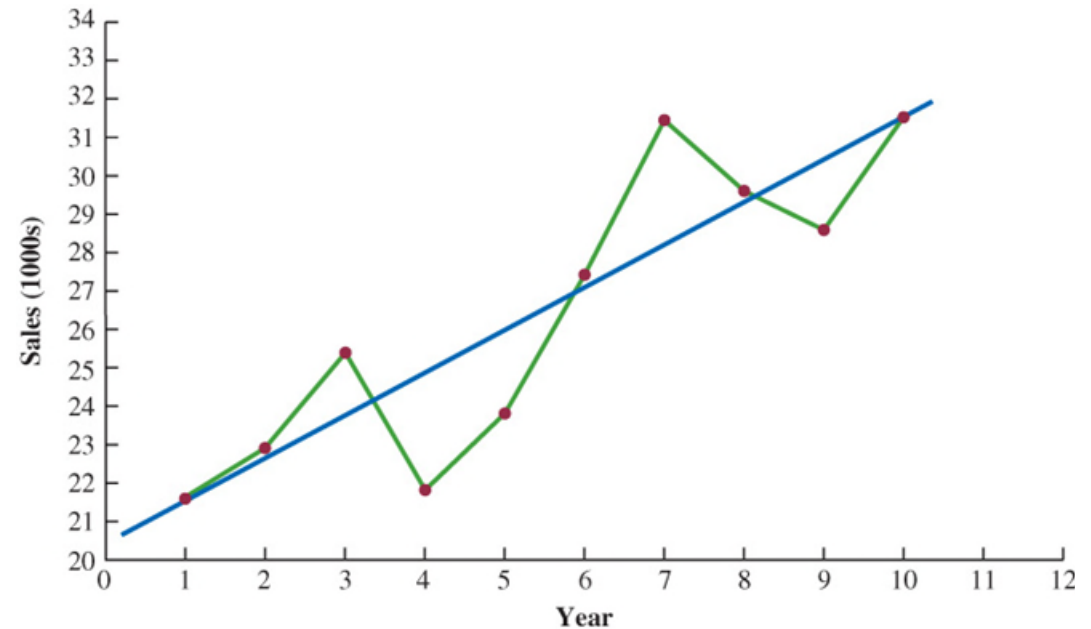
Regression analysis uses known values of variables to estimate the relationship between the **dependent variable** and one or more **independent variables**.

Then, we can use values of the independent variables to estimate the associated value of the dependent variable.

Estimating the linear relationship between x and y requires finding the values of the parameters b_0 and b_1 for the straight line

$$\hat{y} = b_0 + b_1x$$

so that the MSE is minimized, as shown for the bicycle time series data.



17-4 Linear Trend Equation: Line of Best Fit

The notation for the linear trend equation is:

$$T_t = b_0 + b_1 t$$

Where t is the time period, T_t is the linear trend forecast in period t , b_0 is the intercept of the linear trend line, and b_1 is the slope of the linear trend line.

It can be shown that the equations for computing the values of b_0 and b_1 that minimize the MSE are:

$$b_1 = \frac{(t - \bar{t})(Y_t - \bar{Y})}{\sum(t - \bar{t})^2} \quad \text{and} \quad b_0 = \bar{Y} - b_1 \bar{t}$$

Where Y_t is the value of the time series in period t , n is the number of time periods, \bar{Y} is the average value of the time series, and \bar{t} the average value of t .

The calculations necessary for computing the values of b_0 and b_1 for the Bicycle data set are shown in the next slide.

17-4 Bicycle Data Linear Trend Forecast

Let us calculate b_0 and b_1 for the Bicycle data set:

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

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

$$b_1 = \frac{\sum (t - \bar{t})(Y_t - \bar{Y})}{\sum (t - \bar{t})^2} = \frac{90.75}{82.50} = 1.10$$

$$b_0 = \bar{Y} - b_1 \bar{t} = 26.45 - 1.10(5.5) = 20.40$$

The linear trend regression equation is:

$$T_t = 20.4 + 1.10t$$

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	<u>31.4</u>	4.5	4.95	<u>22.275</u>	<u>20.25</u>
55	264.5			90.750	82.50

17-4 Prediction with the Linear Trend Forecast

The slope, $b_1 = 1.10$ (thousands of units/year), can be interpreted as:

Over the past 10 years, the firm has experienced an average growth of 1,100 units per year.

We can use the trend equation to project the trend component of the time series for year 11:

$$T_{11} = b_0 + b_1t = 20.4 + 1.10(11) = 32.5$$

The linear model yields a sales forecast of 32,500 bicycles for next year.

We can also use the trendline to forecast sales farther into the future. For instance, using the trend equation, we develop annual forecasts for two and three years into the future as

$$T_{12} = 20.4 + 1.10(12) = 33.6$$

$$T_{13} = 20.4 + 1.10(13) = 34.7$$

Note that the forecasted value increases by 1,100 bicycles in each year.

17-4 Regression Forecast Error

The table further below shows the computation of the minimized sum of squared errors for the bicycle sales time series.

Minimizing the sum of squared errors also minimizes measure of accuracy, MSE.

For the bicycle sales time series, we have:

$$\text{MSE} = \sum_{t=1}^n \frac{(Y_t - T_t)^2}{n} = \frac{30.70}{10} = 3.07$$

Note that, in this example, we are not using past values of the time series to produce forecasts, and so $k = 0$.

Thus, we can produce a forecast for each period of the time series (*see notes.)

	Sales (1000s)		Forecast	Squared Forecast
Week	Y_t	Forecast \hat{Y}_t	Error	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	<u>0.00</u>
			Total	30.70

17-4 Nonlinear Trend Regression

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.

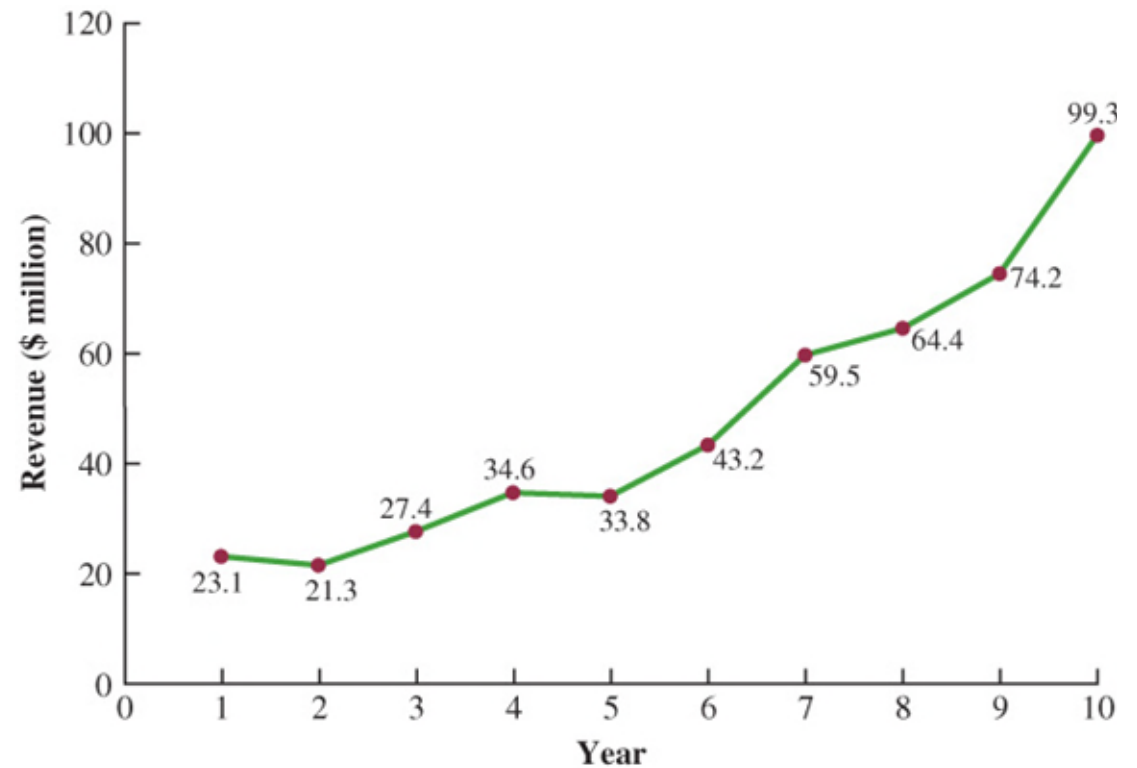
DATAfile: *Cholesterol*

The time series plot indicates an overall increasing or upward trend.

However, unlike the bicycle sales time series, a curvilinear function appears to be suited to model the long-term trend.

A variety of nonlinear functions can be used to develop an estimate of the trend for curvilinear time series.

We will consider two nonlinear trend equations: quadratic and exponential.



17-4 Quadratic Trend Regression

The quadratic trend equation is written as follows.

$$T_t = b_0 + b_1t + b_2t^2$$

For the cholesterol time series, we set:

$$t = \text{Year} = 1, 2, 3, \dots, 10$$

$$t^2 = \text{YearSq} = 1^2 = 1, 2^2 = 4, 3^2 = 9 \dots 10^2 = 100$$

The computer output of a standard multiple regression output for the quadratic trend model produces the estimated regression equation.

$$\text{Revenue} = 24.18 - 2.11\text{Year} + 0.922\text{YearSq}$$

The goodness of fit is excellent ($R^2 = 98.12\%$) and tests of significance show the relationship between Revenue and the quadratic term to be significant.

<i>Regression Statistics</i>					
Multiple R	99.05%				
R Square	98.12%				
Adj. R Square	97.58%				
Standard Error	3.97578				
Observations	10				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>
Regression	2	5770.13	2885.06	182.52	0.000
Residual	7	110.65	15.81		
Total	9	5880.78			
	<i>Coeffs</i>	<i>Std. Err.</i>	<i>t Stat</i>	<i>P-value</i>	
Intercept	24.18	4.68	5.17	0.001	
Year	-2.11	1.95	-1.08	0.317	
YearSq	0.922	0.17	5.33	0.001	

17-4 Exponential Trend Regression

The exponential trend equation is written as follows.

$$T_t = b_0 e^{b_1 t}$$

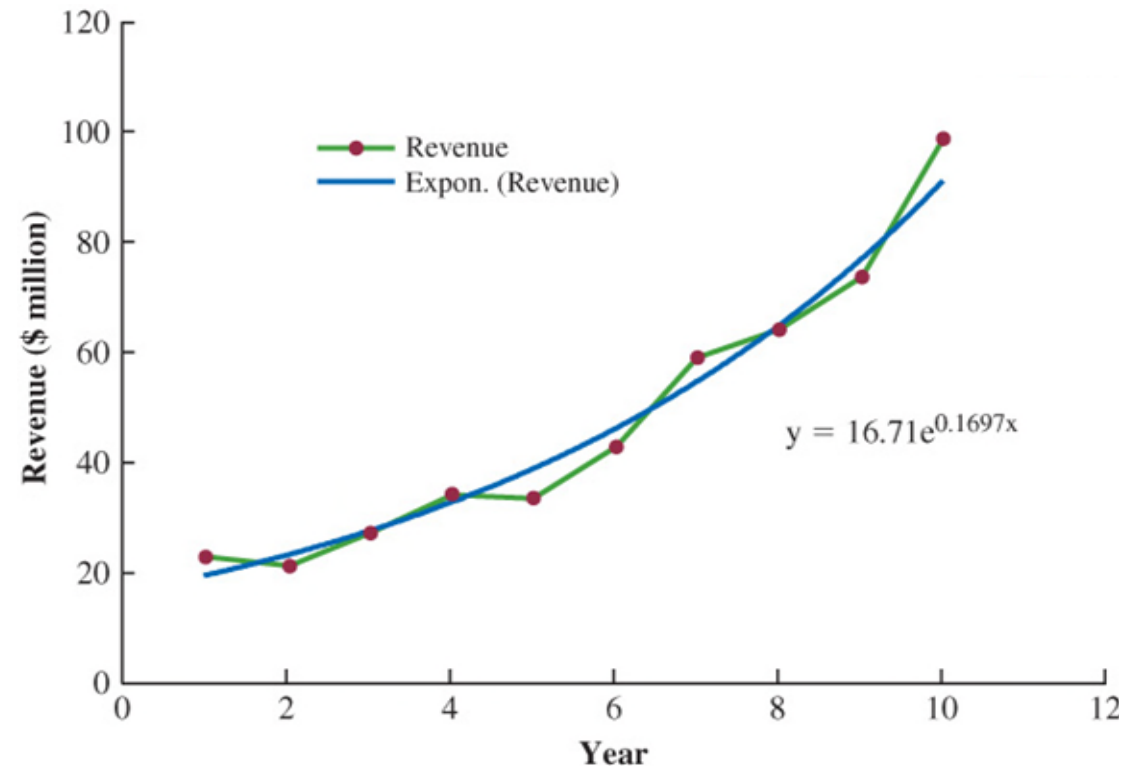
Most statistical software packages compute an exponential trend equation directly.

However, applying a natural log transformation to both sides of the equality in the equation, produces the equivalent linear form.

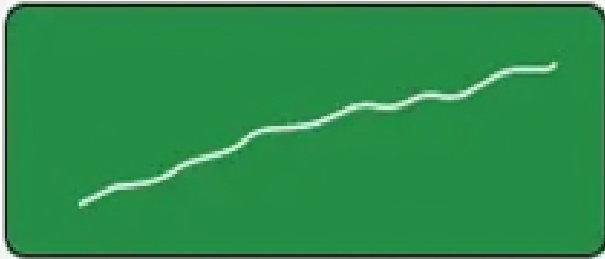
$$\ln T_t = \ln b_0 + b_1 t$$

Computations for the cholesterol time series produce the following estimated regression equation (*see notes.)

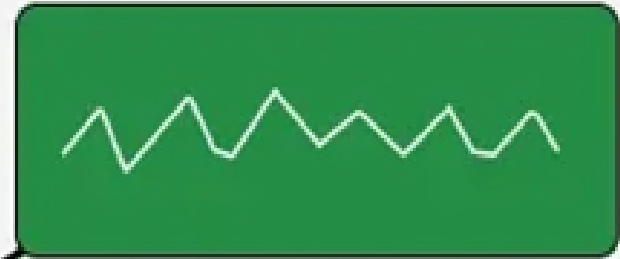
$$T_t = 16.71e^{0.1697x}$$



Trend

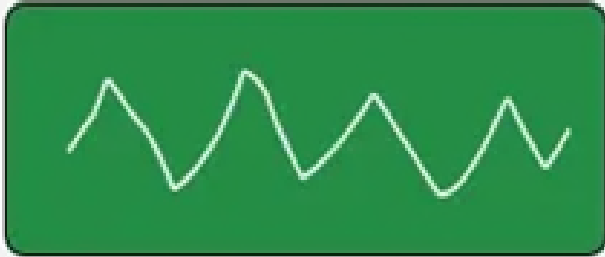


Sesonality



**Components of
Time Series**

Cyclical



Irregular



17-5 Seasonality Without Trend

The quarterly sales of umbrellas at a store over the past five years (DATA file *UmbrellaSales*) are an example of a quarterly seasonal pattern with no long-term trend in sales.

We can model a time series with a seasonal pattern by treating the season as a categorical variable with four levels, one for each quarter.

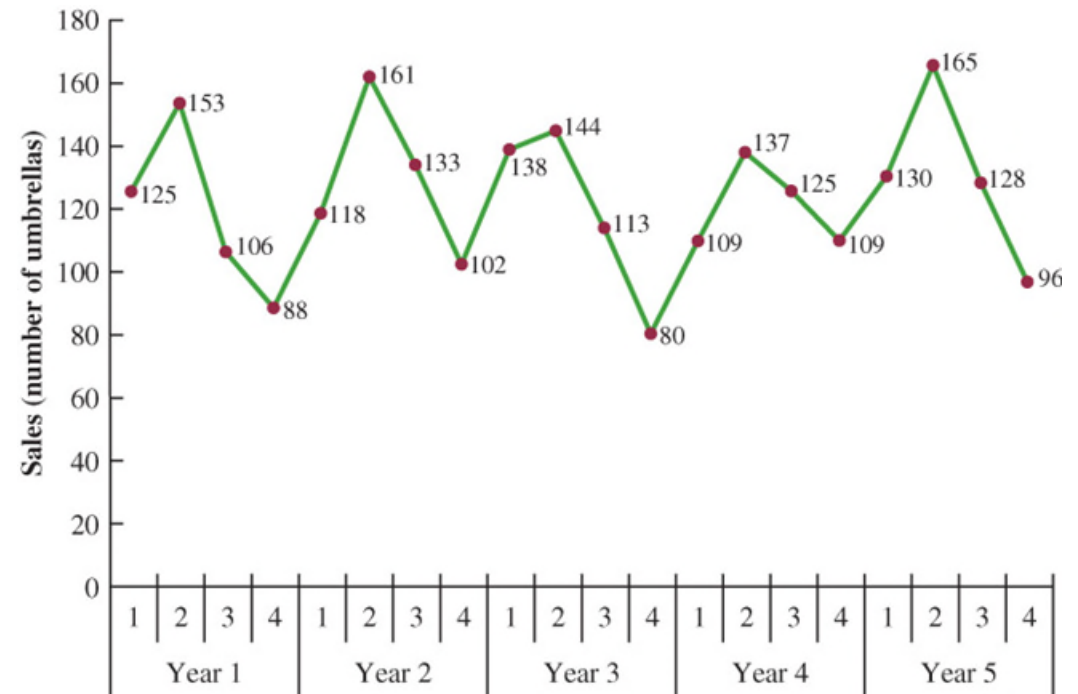
When a categorical variable has k levels, $k - 1$ dummy variables are required.

In this case, we need three dummy variables:

$$Q_{tr1} = \begin{cases} 1 & \text{if period } t \text{ is a Quarter 1} \\ 0 & \text{otherwise} \end{cases}$$

$$Q_{tr2} = \begin{cases} 1 & \text{if period } t \text{ is a Quarter 2} \\ 0 & \text{otherwise} \end{cases}$$

$$Q_{tr3} = \begin{cases} 1 & \text{if period } t \text{ is a Quarter 3} \\ 0 & \text{otherwise} \end{cases}$$



17-5 Seasonality Without Trend: Regression Equation

The general form of the equation relating the quarterly number of umbrellas sold is:

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

We can use the table with the coded values of the dummy variables to the right and the computer output for multiple regression to find the values of b_0 , b_1 , b_2 , and b_3 that minimize the sum of squared errors:

$$\hat{Y}_t = 95.0 + 29.0 Qtr1 + 57.0 Qtr2 + 26.0 Qtr3$$

And use it to forecast quarterly sales for next year:

$$\text{Qtr1 sales} = 95 + 29(1) + 57(0) + 26(0) = 124$$

$$\text{Qtr2 sales} = 95 + 29(0) + 57(1) + 26(0) = 152$$

$$\text{Qtr3 sales} = 95 + 29(0) + 57(0) + 26(1) = 121$$

$$\text{Qtr4 sales} = 95 + 29(0) + 57(0) + 26(0) = 95$$

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

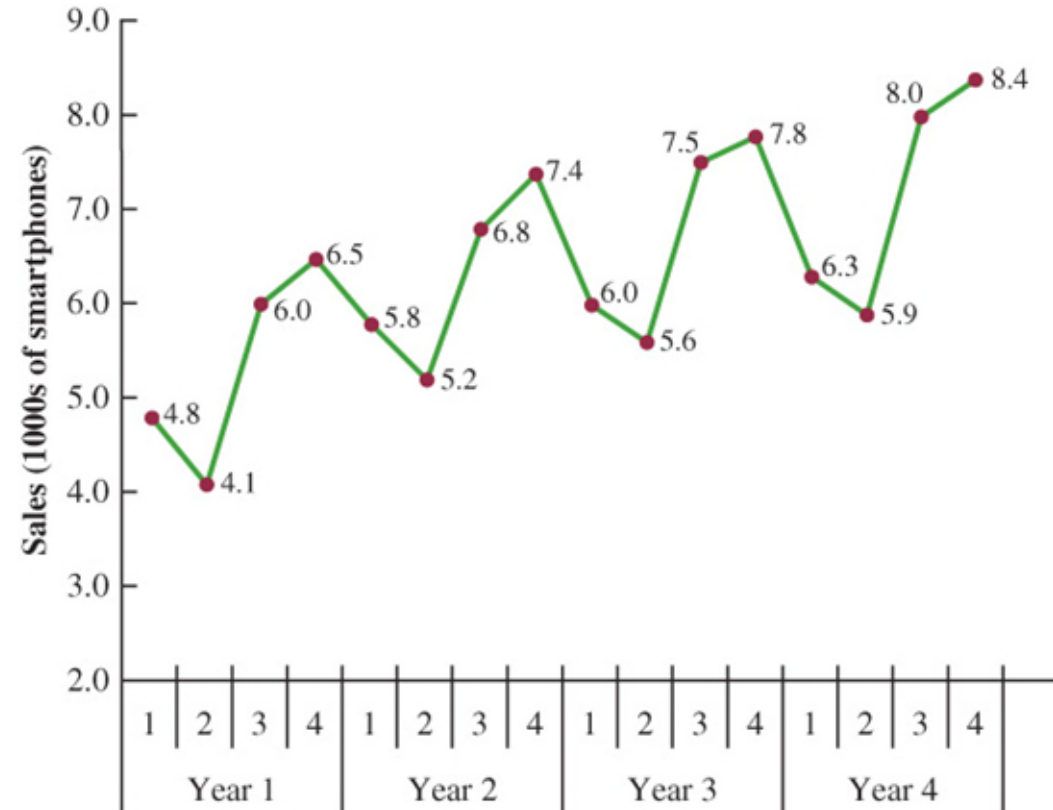
17-5 Seasonality With Trend

The smartphone sales data for a particular manufacturer over the past four years (DATA file: *SmartPhoneSales*) are an example of a quarterly seasonal pattern with a long-term trend.

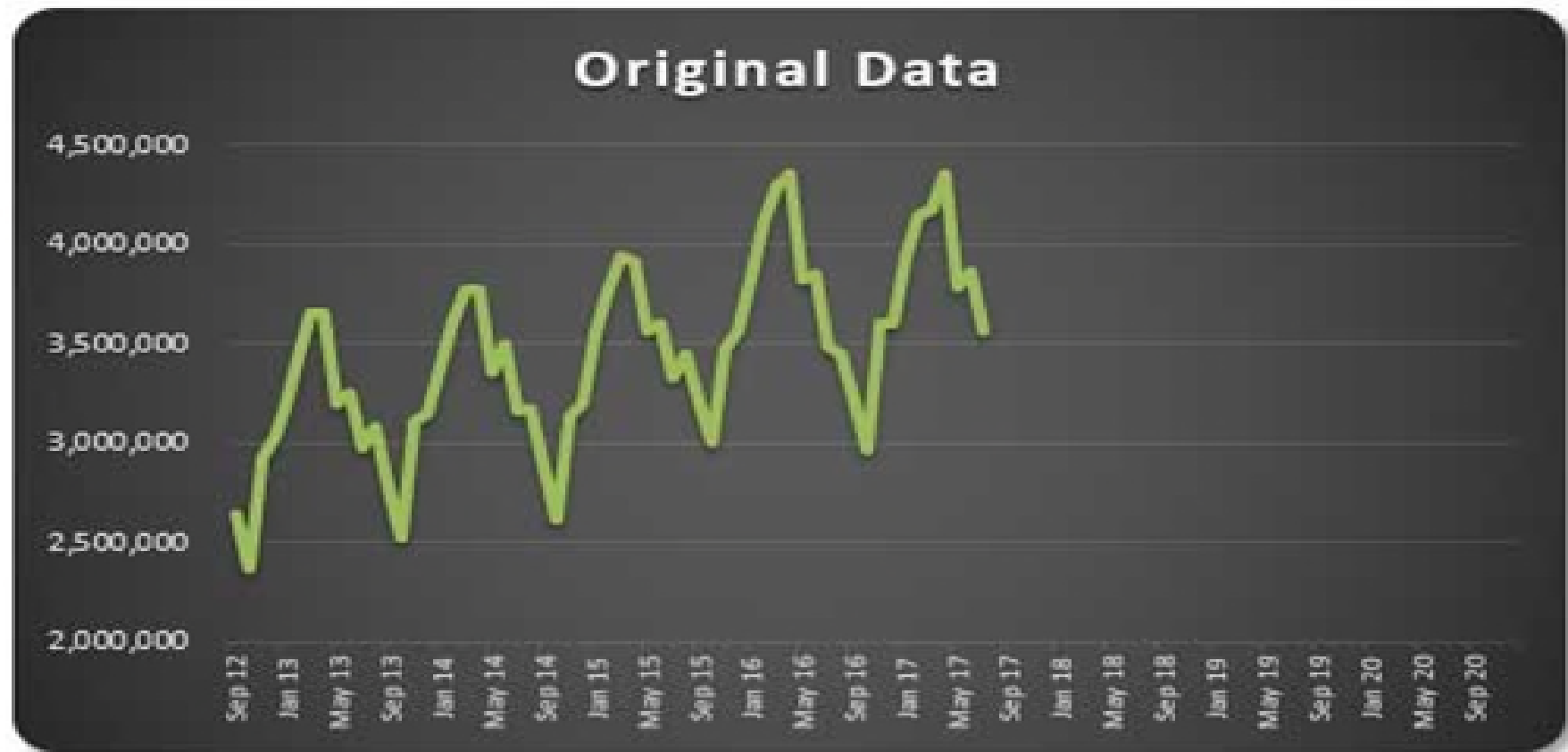
We can model the smart phone sales using three dummy variables as we did for the umbrella sales, with the addition of a time period t to account for the linear trend, as shown in this general equation:

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

Where t is the time period, \hat{Y}_t is the forecast of sales in period t , and $Qtr1$, $Qtr2$, $Qtr3$, are the seasonal dummy variables, defined as in the seasonality without trend demonstration.



Date	Sales
30 Sep 12	2,644,539
31 Oct 12	2,359,800
30 Nov 12	2,925,918
31 Dec 12	3,024,973
31 Jan 13	3,177,100
28 Feb 13	3,419,595
31 Mar 13	3,649,702
30 Apr 13	3,650,668
31 May 13	3,191,526
30 Jun 13	3,249,428
31 Jul 13	2,971,484
31 Aug 13	3,074,209
30 Sep 13	2,785,466
31 Oct 13	2,515,361
30 Nov 13	3,105,958
31 Dec 13	3,139,059
31 Jan 14	3,380,355
28 Feb 14	3,612,886
31 Mar 14	3,765,824
30 Apr 14	3,771,842
31 May 14	3,356,365
30 Jun 14	3,490,100
31 Jul 14	3,163,659
31 Aug 14	3,167,124
30 Sep 14	2,883,810
31 Oct 14	2,610,667
30 Nov 14	3,129,205
31 Dec 14	3,200,527
31 Jan 15	3,547,804



Seasonality and Trend

Source: https://www.fm-magazine.com/content/fmm-home/news/2017/dec/how-to-use-excel-in-seasonal-forecasting-201718012/_jcr_content/article/articleparsys/image.img.jpg/1513353870032.jpg

17-5 Seasonality With Trend: Regression Equation

We can use the coded values of the dummy variables and the computer output for multiple regression to find the values of b_0 , b_1 , b_2 , b_3 , and b_4 that minimize the sum of squared errors:

$$\text{Sales} = 6.069 - 1.363 Qtr1 - 2.034 Qtr2 - 0.304 Qtr3 + 0.1456t$$

And use it to forecast quarterly sales for year 5:

$$\text{Qtr1 sales} = 6.069 - 1.363 (1) - 2.034 (0) - 0.304 (0) + 0.1456(17) = 7.18$$

$$\text{Qtr2 sales} = 6.069 - 1.363 (0) - 2.034 (1) - 0.304 (0) + 0.1456(18) = 6.66$$

$$\text{Qtr3 sales} = 6.069 - 1.363 (0) - 2.034 (0) - 0.304 (1) + 0.1456(19) = 8.53$$

$$\text{Qtr4 sales} = 6.069 - 1.363 (0) - 2.034 (0) - 0.304 (0) + 0.1456(20) = 8.98$$

Thus, accounting for the seasonal effects and the linear trend in smartphone sales, the estimates of quarterly sales in Year 5 are 7180, 6660, 8530, and 8980 sets.

The slope of the trendline indicates a consistent growth in sales of about 146 sets per quarter.

17-5 Models Based on Monthly Data

In the preceding umbrella sales and smartphone sales examples, we showed how dummy variables can be used to account for the quarterly seasonal effects in the time series.

With four levels for the categorical variable season, three dummy variables were required.

However, many businesses use monthly rather than quarterly forecasts.

For monthly data, season is a categorical variable with 12 levels. Thus $12 - 1 = 11$ dummy variables are required.

For example, the 11 dummy variables could be coded as follows:

$$\text{Month1} = \begin{cases} 1 & \text{if January} \\ 0 & \text{otherwise} \end{cases} \quad \text{Month2} = \begin{cases} 1 & \text{if February} \\ 0 & \text{otherwise} \end{cases} \quad \dots \quad \text{Month11} = \begin{cases} 1 & \text{if November} \\ 0 & \text{otherwise} \end{cases}$$

Month 12, December, is obtained by setting all 11 dummy variables equal to 0.

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

17-6 Time Series Decomposition

Time series decomposition assumes that Y_t , the time series value at time t , is a function of three components: trend component; seasonal component; and irregular (or error) component.

Government agencies such as the Census Bureau and the Bureau of Labor Statistics, use time series decomposition to create deseasonalized time series.

Time series components can be combined into additive or multiplicative models (*see notes.)

We will restrict our discussion of time series decomposition to showing how to develop trend and seasonal components for a **multiplicative decomposition model**, taking the form:

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

Where

Trend_t is the trend at time period t , measured in the units of the time series variable.

Seasonal_t and Irregular_t , at time period t , are indexes measured in relative terms, with values indicating effects either above the trend (> 1.00) or below the trend (< 1.00 .)

17.6 Centered Moving Average

To decompose a time series with a multiplicative model, we will use the quarterly smartphone sales time series.

Because the time series has a quarterly frequency, we will use 4-quarter moving averages to remove the combined seasonal and irregular effects from the data.

However, the first moving average for quarters 1-4 year 1, is actually centered at quarter 2.5. The second moving average for quarters 2-4 year 1, and quarter 1 year 2, is centered at quarter 3.5.

To remedy this issue, we introduce the *centered moving average* (CMA_t), defined as (*see notes.)

$$CMA_t = \frac{Y_{t-2} + 2(Y_{t-1} + Y_t + Y_{t+1}) + Y_{t+2}}{8}$$

Year	Quarter	Sales (1000s)	Centered Moving Average
1	1	4.8	
1	2	4.1	
1	3	6.0	5.475
1	4	6.5	5.738
2	1	5.8	5.975
2	2	5.2	6.188
2	3	6.8	6.325
2	4	7.4	6.400
3	1	6.0	6.538
3	2	5.6	6.675
3	3	7.5	6.763
3	4	7.8	6.838
4	1	6.3	6.938
4	2	5.9	7.075
4	3	8.0	
4	4	8.4	

17-6 Seasonal-Irregular Values

The time series plot shows how the centered moving average values “smooth out” both the seasonal and irregular components of the time series.

By dividing each side of the multiplicative decomposition model by the component $Trend_t$, we can identify the combined seasonal-irregular effect in the time series.

$$\frac{Y_t}{Trend_t} = Seasonal_t \times Irregular_t$$

For example, in quarter 3, we have:

$$Seasonal_3 \times Irregular_3 = \frac{Y_t}{Trend_t} = \frac{6.0}{5.475} = 1.096$$

The seasonal-irregular values for the entire time series are summarized in the next slide.

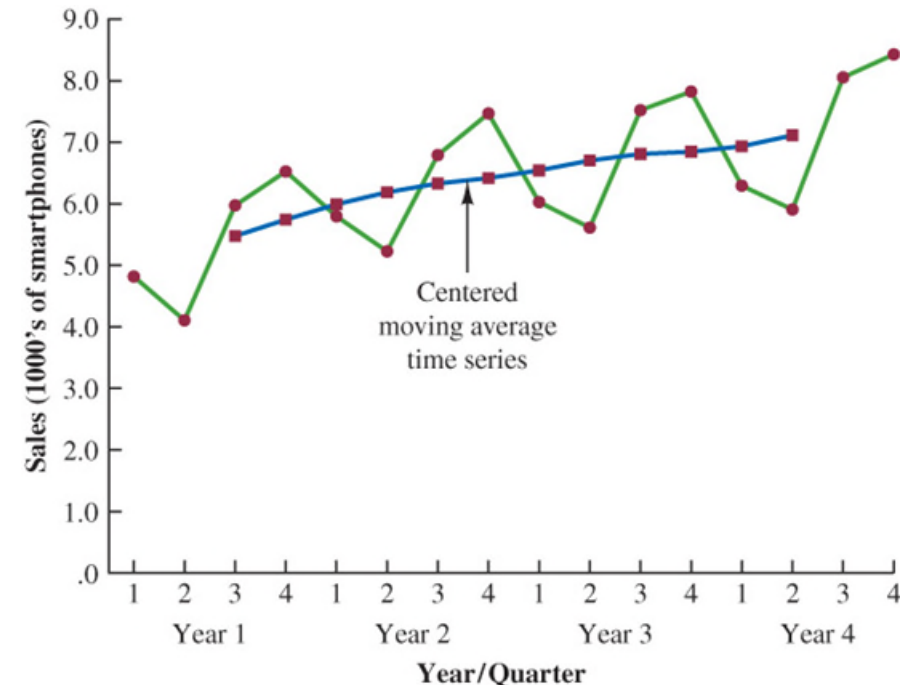


Table 17.16 Centered Moving Average Calculations for the Smartphone Sales Time Series

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

註：過程說明

→ 先計算四期MA平均

➤ $Y_1 + Y_2 + Y_3 + Y_4 \Rightarrow Y_{2.5}$

➤ $Y_2 + Y_3 + Y_4 + Y_5 \Rightarrow Y_{3.5}$

➤ $Y_3 + Y_4 + Y_5 + Y_6 \Rightarrow Y_{4.5}$

➤ $Y_4 + Y_5 + Y_6 + Y_7 \Rightarrow Y_{5.5}$

➤ ○ ○ ○ ○

接著計算中央MA平均

$Y_{2.5} + Y_{3.5} \Rightarrow Y^*_3$

$Y_{3.5} + Y_{4.5} \Rightarrow Y^*_4$

$Y_{4.5} + Y_{5.5} \Rightarrow Y^*_5$

○ ○ ○ ○

17.6 Seasonal Indexes

To calculate the seasonal index for a given quarter, we average the seasonal-irregular values shown to the right for the same quarter over the years available in the time series.

For example, to calculate the *seasonal index* for quarter 3, we average the three quarter 3 seasonal-irregular values

$$\text{Seasonal Index Qtr3} = \frac{1.096 + 1.075 + 1.109}{3} = 1.09$$

The quarterly seasonal indexes are (*see notes.)

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

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		

17.6 Deseasonalizing the Time Series

A time series that has had the seasonal effects removed is referred to as a **deseasonalized time series**.

Deseasonalizing the time series refers to the process of using the seasonal indexes to remove the seasonal effects.

In a multiplicative decomposition model, we deseasonalize a time series by dividing each observation by its corresponding seasonal index (*see notes.)

The table shows the deseasonalized quarterly sales obtained dividing each Y_t by the corresponding seasonal index.

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

17-6 Identify Trend with a Deseasonalized Time Series

The deseasonalized smartphone sales time series appears to have an upward linear trend. Fitting a linear trend line to the deseasonalized time series with simple linear regression produces the estimated linear trend equation.

$$\text{Deseasonalized Sales} = 5.104 + 0.1476t$$

$b_1 = 0.1476$ 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 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

$$T_{17} = 5.104 + 0.1476(17) = 7.613$$

Thus, the deseasonalized trend forecast for next quarter (period 17) is 7613 smartphones.

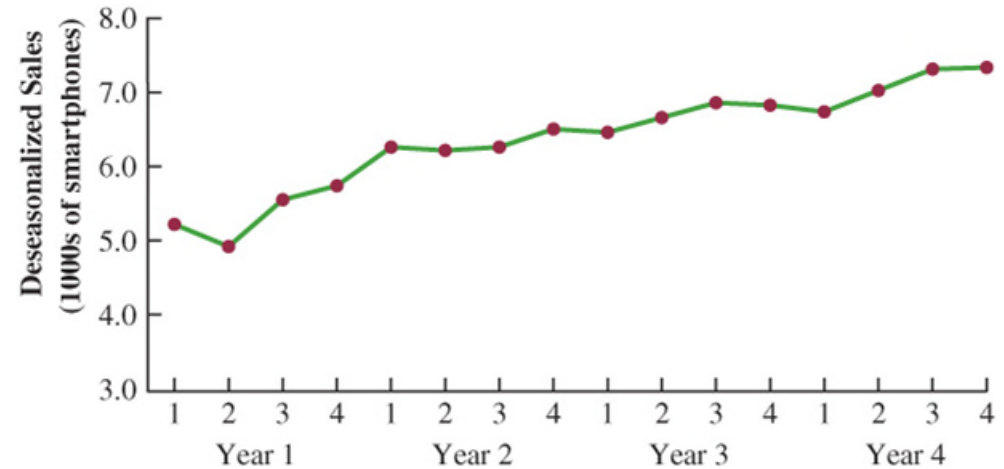


Figure 17.22 Regression Output for the Deseasonalized Smartphone Sales Time Series

Analysis of Variance

Source	DF	Adj SS	Adj MS	F Value	p-Value
Regression	1	7.4089	7.40893	157.05	0.000
Error	14	0.6604	0.04717		
Total	15	8.0694			

Model Summary

S	R-sq	R-sq(adj)
.217197	91.82%	91.23%

Coefficients

Term	Coef	SE Coef	t Value	p-Value
Constant	5.104	0.114	44.81	0.000
t	0.1476	0.0118	12.53	0.000

Regression Equation

$$\text{Deseasonalized Sales} = 5.104 + 0.1476 t$$

註：去掉季節性因素後，再以迴歸分析找出未來趨勢，並求出不加季節性的預測值。

→最後再乘上季節調整因子，得出完整的預測值。

17.6 Seasonal Adjustments

Similar calculations of trend forecast for quarters 2-4 of year 5 (periods 18-20), yield

$$T_{18} = 7.761 = 7761 \text{ units} \quad T_{19} = 7.908 = 7908 \text{ units} \quad T_{20} = 8.056 = 8056 \text{ units}$$

We obtain the forecast for the next four quarters by multiplying each deseasonalized quarterly forecast based on trend by its respective seasonal index.

For example, for quarter 3 (period 19), we have

$$\text{Forecast quarter 3} = T_{19} \times (\text{Seasonal Index Quarter 3}) = (7908)(1.09) = 8620 \text{ units}$$

The complete quarterly forecast is shown below.

Year	Quarter	Deseasonalized Trend Forecast	Seasonal Index	Quarterly Forecast
5	1	7613	0.93	$(7613)(0.93) = 7080$
	2	7761	0.84	$(7761)(0.84) = 6519$
	3	7908	1.09	$(7908)(1.09) = 8620$
	4	8056	1.14	$(8056)(1.14) = 9184$

17.6 Cyclical Component

The multiplicative decomposition model can also be expanded to include a cyclical component

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

The cyclical component, like the seasonal component, is expressed as a percentage of trend.

However, incorporating the cyclical component in a decomposition model presents challenges:

- A cyclical component refers to a multi-year time period. Thus, obtaining enough relevant data to estimate the cyclical component is often difficult.
- Cycles usually vary in length.

Because of the challenges involved with identifying and/or separating cyclical effects from long-term trend effects, these effects are often combined and referred to as a combined trend-cycle component.

We leave discussion of the cyclical component to specialized texts on forecasting methods.

Summary

- This chapter introduced basic methods of time series analysis and forecasting.
- We first showed that the underlying pattern in the time series, such as a horizontal pattern, a trend pattern, or a seasonal pattern, can often be identified by building a time series plot.
- We also discussed three measures of forecast accuracy: mean absolute error (MAE), mean squared error (MSE), and mean absolute percentage error (MAPE).
- We showed three different smoothing methods for developing a stationary forecast of a horizontal pattern: moving average, weighted moving average, and exponential smoothing.
- For time series that have only a long-term linear trend, we showed how regression analysis can be used to make trend projections.
- We showed how dummy variables can be added to a multiple regression model to develop an equation that forecasts seasonal effects without or with a long-term trend.
- Finally, we showed how time series can be decomposed into trend and seasonal components, deseasonalize time series, and use seasonal indexes for forecasting.