Chapter 2. Forecasting

> *"*Analysts predicted in 1980 that one million mobile phones would be used worldwide by the year 2000. They were wrong by 599 million.*" Kurt Hellstrom, president of Ericsson, in an address to Comdex 2000 (PC Magazine, 16-Jan-2000, p. 72).*

2.1. Introduction

Forecasting is an essential and basic activity in any planning process

Effective logistics planning requires accurate estimates of the future activities to be performed by the logistics system or supply chain. These estimates are typically in the form of predictions and forecasts. Many forecasts, especially medium and long-term forecasts, are prepared by other departments in the organization, such as marketing or product development, rather than by the logistics department. However, certain short-term forecasting is performed by logisticians, such as forecasting related to inventory control and individual user demand. In those cases, forecasting future conditions is achieved by extrapolating the historical observations. The underlying pattern is often seasonal corresponding to the natural seasonality of many human and agricultural activities. The focus of this chapter will be on techniques for such short-term forecasting, that are most often performed by logisticians. An often-used alternative term for these activities is demand planning, specifically tactical and operational demand planning.

Only in a few cases of operational logistics planning are the activities that have to be performed already known. Examples are the routing and scheduling of order pickers through a warehouse to collect customer orders after all orders have been received; the routing and scheduling of delivery trucks after the customer stops and quantities are known; and in manufacturing the scheduling of the assembly operations of computers after the customer orders for certain configurations have been received. In all other cases, the future activities have to estimated or forecasted.

Forecasting is an essential activity in any business and governmental organization since it provides the basic information for the planning and execution of all operational activities of the organization. In logistics planning most forecasting is related to forecasting the demands of final and intermediate products. The forecast for the final products of an organization determines then the forecasts for intermediate products, parts, raw materials, and personnel.

Demand Planning Structure

The goal of demand planning is to predict the future consumption of "products", be it physical products, medical services, restaurant meals, or consumption of utilities such as electricity. Demand planning has to support at least the following three dimensions of aggregation and disaggregation: product family tree, geography, and planning horizon. The product family tree is a hierarchical structure relating products to product groups, groups to families, and families to product lines. For example, an individual product or SKU may be a particular model of personal computer with specific processor, video, storage, and memory hardware. A product group may be based on all products with the same processor with different configurations for the other hardware. A product family may be all the product groups intended to be sold to a particular market segment such as the home user or the corporate user. Finally, the product line may be personal computers, which are separated from servers or personal digital assistants. The geographical dimension is also hierarchically structured. It may be divided into areas, regions, and countries. The time dimension is typically structured in years, seasons or quarters, months. If necessary it can be specified to the week or day level. Forecast quantities can be assigned to any intersection of product, geography, and time. Such three-dimensional databases grow very quickly in size even for small and mid-sized companies. Powerful database software and computer hardware has reduced the burden of collecting, maintaining, and processing such large amount of data. A much more difficult problem has been to maintain the internal consistency between various levels and dimensions of the forecast data.

Pattern Classification

When forecasting a particular activity, the most important requirement for generating a high quality forecasts is a fundamental understanding of the underlying pattern of the activity. Depending on its behavior over time, the underlying pattern can be classified as regular or irregular. If the pattern is regular, the future values can be predicted based on the past or historical values. Common regular patterns are a constant pattern, a trend pattern, seasonal pattern, or a combination of trend and seasonal pattern. Classical decomposition decomposes the pattern in a trend, seasonal, and random component. As long as the random variations are small compared to the underlying pattern, accurate forecasts can be obtained by popular mathematical forecasting techniques, such as regression and time series analysis. A pattern can also be irregular, when it is a singular occurrence, intermittent, or highly variable. The pattern is said to be “lumpy”. Such variables are typically very difficult to forecast with any accuracy using popular mathematical models.

Finally, dependent or derived demand patterns constitute a special case when an independent variable causes predictable behavior for a number of other dependent variables. In this case the value of the demand for the final product is stochastic and forecasted with the techniques described above. But once the demand for the final product is known, the demand for all the subcomponents, raw materials, and resources is known with certainty. The derived demand is usually obtained by recursively exploding the *bill of materials* (BOM) of the final product to all of its individual product components. For example, the forecasted independent demand for a car model causes a dependent demand shifted back in time for the associated car engine. Most popular forecasting techniques are based on the assumptions of independent behavior with a random component. The methods are not applicable for forecasting dependent behavior.

The future logistics activities are subject to the actions of many different organizations in and outside the supply chain and to inherent uncertainty caused by truly random events. Therefore, the forecasting should be a collaborative effort among many groups in the supply chain. Marketing can provide advance estimates of the impact of future promotions and product introductions. The expertise of long-term employees may judge the impact that similar conditions have had in the past. The end result of the forecasting process should a consensus-based forecast that is used for all the planning by all organizations in the supply chain.

Classification of Forecasting Methods

Forecasting methods can be classified as quantitative or objective versus qualitative or subjective depending on the fact if an explicit model forms the basis of the forecasting method. Quantitative forecasting models can be further divided into casual and time series models.

Subjective or Qualitative Forecasting Methods

Subjective or qualitative forecasting methods use experts, subjective judgment, intuition, or surveys to produce quantitative estimates about the future. The information on which the forecasts are based is typically non-quantitative and subjective. Historical data may not be available or of little relevance to the forecast. The highly subjective nature of the forecast makes it very difficult to validate the accuracy and to standardize the methods.

These methods are primarily used to predict the demand for a new product or a product in new areas, the impact of policy changes, or the impact of new technology. They are typically used for medium to long-range forecasts. The forecast for cell phone use, which was described in the quote at the start of the chapter, illustrates the difficulty of making these types of predictions.

Quantitative of Objective Forecasting Methods

Objective or quantitative forecasting methods rely on a formalized underlying model to make predictions. They are divided into time series and casual methods.

Time Series Forecasting Models

Time series forecasting methods are based on the fundamental assumption that future estimates are based on prior, historical values of the same variable. This implies that the historical pattern exhibited by the variable to be forecasted will extend into the future. In addition, it is implicitly assumed that historical data are available.

Time series forecasting methods are mostly used to forecast variables for the short term. As such, time series methods are some the forecasting techniques most often used by logisticians.

Casual Forecasting Models

A second major category of forecasting models consists of causal models. The basic assumption for a casual model is that the future value of the forecasted variable can be expressed as a mathematical function of the known values of a set of different variables. For example, the historical sales of a product combined with the historical breakdown rates for this product allow the forecast of the number of breakdowns and required service parts during the coming year. Causal forecasting methods are often used in predicting future economic activity and future social and life science trends. The known variables in the causal model are also called *leading indicators*, since they occur before or lead the forecasted variable and indicate the behavior of the forecasted variable. One leading indicator used often in predicting logistics activity is the purchase of corrugated cardboard boxes. When companies expect an increase in sales, they purchase more shipping supplies such as cardboard boxes. Causal models can be quite good in predicting major changes and trends and are often used in the medium and long-range period.

The major difficulty in applying causal models is of course the development of accurate cause-and-effect relationships, since the true causal relationship is often much more complex than the causal model. Causal models based on regression and economic techniques typically have substantial forecasting errors.

The causal relationship in general is expressed as

$$Y = f\left(X_1, X_2, \ldots X_N\right)$$

Econometric models are a subclass of causal models that use only linear causal relationships, as illustrated in the following expression.

$$Y = a_1 X_1 + a_2 X_2 + \ldots a_N X_N$$

2.2. Forecasting Techniques for Time Series

Time Series Assumptions

A time series is a set of observations x_t each being recorded at a specific time *t*. In a continuous-time time series, the observations are made continuously during a specified time interval. In a discrete-time time series, the observations are made at a discrete set of times. Often the observations are made at a fixed time interval and the time periods are renamed and scaled to the set of integer numbers 1, 2, …N. It is assumed that the value of the forecasted variable is computed as the sum of the expected value of the forecasted variable for that period based on an underlying deterministic pattern and a random component. Furthermore, it is assumed that the expected value of the random component is equal to zero and that the random component is normally distributed with variance σ^2. In other words, if D_t is the forecasted variable in period *t*, Pat_t is the value of the underlying pattern in period *t*, and ε_t is the random component in period *t*, then

$$D_t = Pat_t + \varepsilon_t \tag{2.1}$$

$$\varepsilon_t = N\left(0, \sigma^2\right) \tag{2.2}$$

$$E\left[D_t\right] = Pat_t \tag{2.3}$$

$$Var\left[D_t\right] = Var\left[\varepsilon_t\right] = \sigma^2 \tag{2.4}$$

When forecasting a time series, the first step, as always, should be to determine the underlying pattern of the variable, since it is this pattern that is assumed to extend into the future. Plotting the time series data allows one to determine if a trend, a seasonal component, a long-term cycle, any sudden changes, or any outliers are present. If sudden changes are present, different models may be constructed for each of the homogeneous segments of the series. If outliers are present, they should be further investigated to determine if there is any justification for discarding them. For example, an observation may have been recorded incorrectly or temporary promotion may have significantly increased customer demand.

Any forecasting method for the underlying pattern can be interpreted as a low-pass filter that filters out the high frequency oscillations of the random component. See Brockwell and Davis (1996) for an in depth treatment of time series and forecasting from this point of view.

Forecast Error and Forecast Performance

The forecast error is defined as the algebraic difference between the forecast and the actual realized value for a particular time period. The forecasted value may have been computed the prior period or in an earlier period. The corresponding formulas to compute the forecast error are given next.

$$\begin{aligned} e_t &= F_{t-\tau,t} - D_t \\ e_t &= F_t - D_t \end{aligned} \tag{2.5}$$

Obviously it is desirable that the forecast error for a number of forecast is as close as possible to zero. A number of quality measures for forecasts have been defined in function of the forecast errors. The simplest measure sums the forecast errors to compute the mean error (ME). This method has the disadvantage that large positive and negative errors offset each other and give the false impression of a high quality forecast. To overcome this deficiency, the mean absolute deviation (MAD) computes average absolute value of the forecast error. The mean squared error (MSE) method computes the average value of the squared forecast error. This method also avoids offsetting positive and negative errors. The square root of the MSE is called the root mean squared error (RMSE) and is equal to the standard deviation of the forecast error.

$$\begin{aligned} ME &= \frac{1}{n}\sum_{t=1}^{n} e_t \\ MAD &= \frac{1}{n}\sum_{t=1}^{n} |e_t| \\ MSE &= \frac{1}{n}\sum_{t=1}^{n} e_t^2 \\ RMSE &= \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \end{aligned} \tag{2.6}$$

All of the above quality measures are in absolute units, in other words, in the same units of the variable that is forecasted or those units squared. Many times it is more intuitive to express the forecast error as a percentage of the forecasted value. This is a called a relative forecast error. The quality measure most often used in function of the relative forecast error is the mean absolute percentage error (MAPE).

$$MAPE = \left[\frac{1}{n} \sum_{t=1}^{n} \left| \frac{e_t}{D_t} \right| \right] \cdot 100 \tag{2.7}$$

The MAPE computations can cause numerical difficulties if the realized value of the variable that is forecasted becomes very close to zero or zero.

Confidence Interval

The probability that the forecasted value F_t for the variable under consideration will be exactly equal to the expected value of the forecasted variable Pat_t is very small. Hence, it is more useful to replace the point forecast with an interval forecast. An interval for which we can assert with a specified degree of certainty that it contains the expected value of the forecasted variable is called a *confidence interval*.

Typically, a claim is made that with a probability of $1-\alpha$ the interval

$$\left[F_t - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, F_t + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right] \tag{2.8}$$

contains the expected value Pat_t, where $z_{\alpha/2}$ is derived from the normal distribution. Since the expected value of the forecasted variable is either inside the above interval or not, speaking of a probability in this context seems inappropriate. What is really indicated is that in repeated sampling α percent of confidence intervals computed with the above formula will contain the expected value of the forecasted variable. Although the expected value of the forecasted variable will never be known nor will it be known if this expected value fell inside the computed interval, we can be assured that the method used to compute the interval is α percent reliable, that is, it is expected to work in α percent of the time.

The computation of the confidence interval requires the determination of the standard deviation of the random component of the forecasted variable. Again, since the standard deviation of the random component of the forecasted variable is not known, the forecast error is used as an estimate of the random component and the standard deviation of the forecast error is used, which is computed as

$$s_F = \sqrt{\frac{\sum_{t=1}^{N} (F_t - D_t)^2}{N}} \tag{2.9}$$

Note that this formula implies that the mean of the forecast error is equal to zero. The built-in formulas in computer spreadsheets for computing the standard deviation of a sample will compute the sample mean and then will use this sample mean in the computation of the standard deviation. The formulas will then yield a smaller standard deviation if the computed sample mean is not zero. This implies that you cannot use either the built-in STDEV or the STDEVP functions of Excel. One way to compute the standard deviation in Excel spreadsheets is to use the following built-in functions

$s_F = SQRT\left(SUMSQ\left(e_1 : e_N\right) / COUNT\left(e_1 : e_N\right)\right)$.

The $1-\alpha$ percent confidence interval is then computed as

$$F_t \pm z_{\alpha/2} \cdot s_F \tag{2.10}$$

The values of z in function of the probability $1-\alpha$ can be computed with the Excel function $NORMSINV\left(1-\alpha/2\right)$. The value of z for a number of commonly used probabilities is given in Table 2.1.

Table 2.1. Confidence Intervals for the Standard Normal Distribution

(1−α)	α/2	z	(1−α)	α/2	z
0.5000	0.2500	0.6745	0.9250	0.0375	1.7805
0.6000	0.2000	0.8416	0.9500	0.0250	1.9600
0.6827	0.1587	1.0000	0.9545	0.0228	2.0000
0.7000	0.1500	1.0364	0.9750	0.0125	2.2414
0.7500	0.1250	1.1503	0.9800	0.0100	2.3263
0.8000	0.1000	1.2816	0.9900	0.0050	2.5758
0.8500	0.0750	1.4395	0.9950	0.0025	2.8071
0.9000	0.0500	1.6449	0.9973	0.0014	3.0000

Further information can be found in reference texts on probability and statistics for engineers, such as Hayter (1996). Following are simple but effective forecasting techniques for the most common data patterns, which are the constant, trend, and seasonal data patterns.

Constant Data Pattern

Constant Data Pattern Characteristics

A variable following a constant data pattern is assumed to have an unvarying expected value. This corresponds to a horizontal line in the time series graph. The actual realizations, which include the random component, are scattered around this horizontal line. An illustration of a constant data pattern is shown in Figure 2.1.

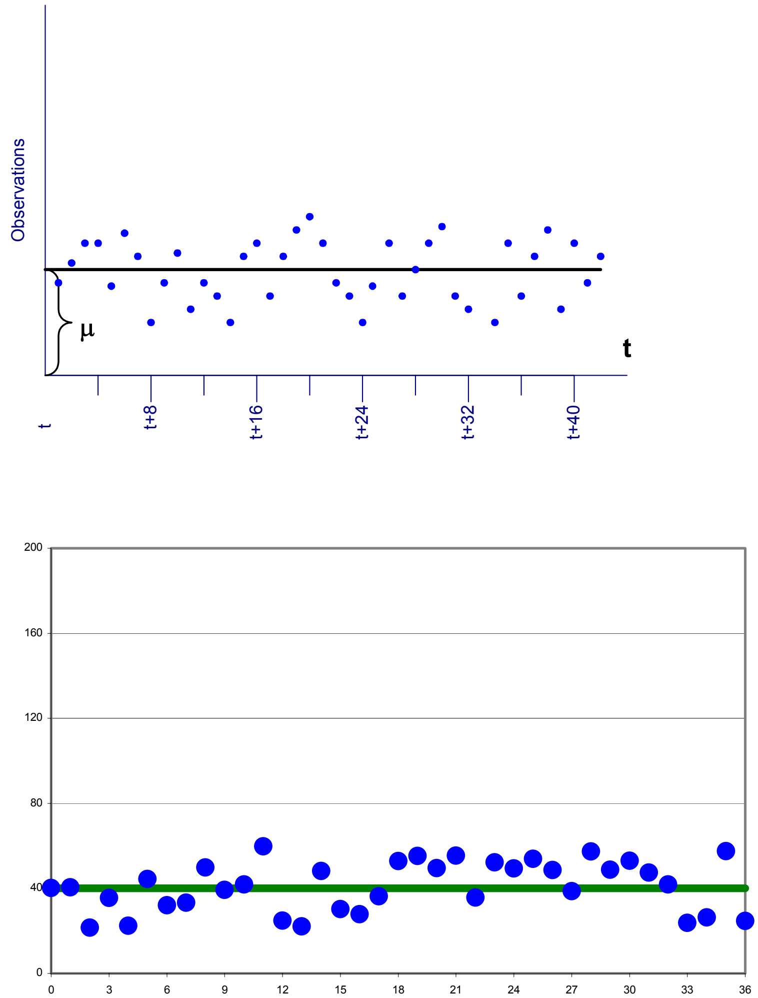


Figure 2.1. Illustration of a Time Series Following a Constant Pattern

A constant pattern has a single constant parameter, which is its mean μ. The forecasted values for all future periods are identical. Since the mean of the forecast error is assumed to be zero, the computation of the standard deviation of the forecast error uses N in the denominator. This is equivalent to the computation of the root mean squared error (RMSE).

$$D_t = \mu_t + \varepsilon_t \tag{2.11}$$

$$F_{t+1} = F_{t+\tau} \qquad \forall \tau \tag{2.12}$$

$$s_e = \sqrt{\frac{\sum_{t=1}^{N} (F_t - D_t)^2}{N}} = \sqrt{\frac{\sum_{t=1}^{N} e_t^2}{N}} \tag{2.13}$$

The forecasted values are illustrated in Figure 2.2. Since the underlying pattern is assumed to be constant, the forecasted values for all the future periods are identical.

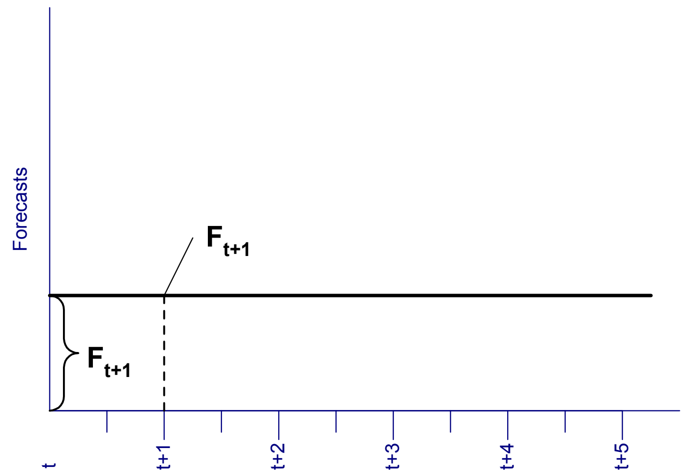


Figure 2.2. Constant Pattern Forecasts

Forecasting Methods for a Constant Data Pattern

Average and Moving Averages

Some of the simplest methods to forecast a variable with a constant data pattern are based on eliminating the random component by averaging the past observations, since the expected value of the random component is equal to zero. Several variants exist depending on the importance or weight attributed to the historical observations.

The average assigns an equal weight to all prior observations.

$$F_t = \frac{1}{N} \sum_{i=1}^{N} D_i \tag{2.14}$$

where N is the total number of available observations for this variable.

One of the disadvantages of the average is that the oldest observations have an equal weight in determining the forecast to more recent observations.

The moving average assigns an equal weight to a limited set of the most recent observations. The number of observations in this set is called the interval and is denoted by N. The weight of each observation is then equal to $1/N$ and the sum of all the weights is equals one.

$$F_t = \frac{1}{N}\sum_{i=1}^{N} D_{t-i} \tag{2.15}$$

The weighted moving averages method assigns an unequal weight to a limited set of most recent observations. The sum of the weights is equal to one. Any observations outside the set can be thought of as having a weight equal to zero.

$$\begin{aligned} F_t &= \sum_{i=1}^{N} \alpha_i D_{t-i} \\ \sum_{i=1}^{N} \alpha_i &= 1 \end{aligned} \tag{2.16}$$

The weighted moving averages method requires that, in addition to the interval N, the N weights must also be determined. The weighted moving averages method may be used to forecast a variable with a growing or declining trend or with a seasonal behavior. Exponential smoothing methods for trend and seasonal patterns are also applicable in those cases and require fewer parameters or are easier to apply. As a consequence, the exponential smoothing methods are used more often than weighted moving averages methods.

Simple Exponential Smoothing

One of the most used methods for short-term forecasting of a variable with an underlying constant pattern is exponential smoothing. To differentiate this variant of exponential smoothing from later variants for trend and seasonal patterns, this variant is also called simple exponential smoothing. Exponential smoothing has several significant advantages. It is mathematically simple, it requires a minimum amount of data, and it is self-adaptive to changes in the underlying data pattern.

In exponential smoothing, the past observations are not given equal weights. The more recent observations are given more weight than earlier observations and it will be shown that the weight decreases by a constant factor when going back farther in the past.

There exist several equivalent expressions for the computation of the forecast in period $t+1$ based on the observed demand during period *t* and the forecast for time period *t*. The constant α is commonly called the exponential *smoothing constant* and $(1-\alpha)$ is referred to as the *damping factor*. The impact of the most recent observation is dampened by the most recent forecast, which incorporates all the previous observations. Equivalently, the new forecast is equal to the previous forecast plus a smoothing term based on the forecast error. The smoothing constant has values strictly between 0 and 1. So the new forecast can also be interpreted as a convex combination of the most recent observation and the most recent forecast.

$$F_{t+1} = \alpha D_t + (1-\alpha)F_t = F_t + \alpha\left(D_t - F_t\right) \tag{2.17}$$

The following initial values are typically used

$$\begin{aligned} F_2 &= D_1 \\ F_1 &= D_1 \end{aligned} \tag{2.18}$$

By repeated substitution of the previous forecasts F_t with expression (2.17), we obtain the following expression for the next forecast. It shows that the weight for previous demand observations decreases geometrically with the constant factor $(1-\alpha)$.

$$F_{t+1} = \sum_{i=0}^{t-1} \alpha(1-\alpha)^i D_{t-i} + (1-\alpha)^t D_1 \qquad t \ge 1 \tag{2.19}$$

The limit for a large number of prior observations of this expression is defined for α strictly between zero and one and is equal to

$$F_{t+1} = \sum_{i=0}^{\infty} \alpha(1-\alpha)^i D_{t-i} \tag{2.20}$$

Since α is strictly smaller than one, the sum of all the weights in the geometric series is defined and equal to one. The next forecast is thus a convex combination of all the previous observations.

$$\sum_{i=0}^{\infty} \alpha(1-\alpha)^i = \frac{\alpha}{1-\left(1-\alpha\right)} = 1 \qquad \alpha < 1 \tag{2.21}$$

using

$$\sum_{i=0}^{\infty} a^i = \frac{1}{1-a} \qquad a < 1 \tag{2.22}$$

However, we do not need to store all previous observations, since the impact of all previous observations is encapsulated in the previous forecast.

Choosing the appropriate value of the exponential smoothing constant requires a degree of judgment based on the understanding of the behavior of the underlying data pattern to balance the responsiveness versus the stability of the forecasting model. A higher value of the smoothing constant assigns more weight to the most recent observations. This allows the forecasting model to respond more quickly to changes in the underlying data pattern. However, if the value of the smoothing constant is increased too much then the model will start tracking the changes in the random component of the time series rather than in the underlying pattern. In this case, the forecast is said to be "nervous". Many times, a more stable forecast is desirable, which implies a smaller value of the smoothing constant. Low values of the smoothing constant provide very "stable" forecasts that are not likely to be influenced by the randomness of the time series. However, the more the smoothing constant is decreased the longer it takes for the forecasting model to adjust to fundamental changes in the underlying data pattern. Higher values of the smoothing constant may be appropriate if the underlying data pattern is likely to change more quickly such as for introduction of a new product, discontinuing of a product, start of a recession or economic boom, or a promotional campaign. However, if one expects that the underlying pattern is fundamentally a trend or seasonal pattern, the exponential smoothing methods appropriate for those cases should be used rather than a larger smoothing constant for the constant data pattern. Those exponential smoothing methods are discussed later in this chapter. The typical range of the exponential smoothing constant is [0.2, 0.4]. The Excel spreadsheet uses 0.3 as the default value of the smoothing constant. If sufficient historical data is available, the smoothing constant can be chosen so that forecast error for the historical period is minimized. In this case, the smoothing constant is set to the value that minimizes the sum of squared errors over the historical period.

Assume that we want start forecasting the sales of a product. The underlying sales pattern for this product is assumed to be constant. The sales for period one were 62 units. Since this is the startup period, a smoothing constant of 0.4 was selected. The forecast for period two is 62 units, but the observed sales equaled 59 units. The forecast for the sales during the third period using exponential smoothing is then

$$F_3 = 0.4 \cdot 59 + (1-0.4) \cdot 62 = 60.8$$

The observed sales for period three are 68 and the forecast for sales during the fourth period is then

$$F_4 = 0.4 \cdot 68 + (1-0.4) \cdot 60.8 = 63.68$$

An example comparing the various forecasting methods for a constant underlying data pattern is given at the end of the section on the constant pattern.

Comparison of the Moving Averages and Exponential Smoothing Methods

Both the moving averages and the exponential smoothing methods use a single parameter to control the responsiveness of the forecasting method to the changes in the observed values. One way to establish a relationship between the interval of the moving averages method and the smoothing constant of exponential smoothing method is compute the average weighted "age" of the observations on which the forecast is based.

For the moving average method, the forecast is based on N observations with equal weight and with ages of 1, 2… N periods. The average weighted age is then

$$\overline{a_{MA}} = \sum_{i=1}^{N} \frac{i}{N} = \frac{N+1}{2} \tag{2.23}$$

using

$$\sum_{i=1}^{N} i = \frac{N\left(N+1\right)}{2} \tag{2.24}$$

For the exponential smoothing method, the forecast is based on all the prior observations with geometrically declining weight. In the limit, the average weighted age is then

$$\overline{a_{ES}} = \sum_{i=1}^{\infty} i \cdot \alpha \left(1-\alpha\right)^{i-1} = \frac{1}{\alpha} \tag{2.25}$$

using

$$\sum_{i=1}^{\infty} i a^{i-1} = \frac{1}{(1-a)^2} \qquad a < 1 \tag{2.26}$$

Making the average weighted age of both methods equal yields the following relationship between the smoothing constant and the interval length.

$$\frac{N+1}{2} = \frac{1}{\alpha} \tag{2.27}$$

Either constant can then be computed from the other with one of the following two formulas

$$\begin{aligned} \alpha &= \frac{2}{N+1} \\ N &= \frac{2-\alpha}{\alpha} \end{aligned} \tag{2.28}$$

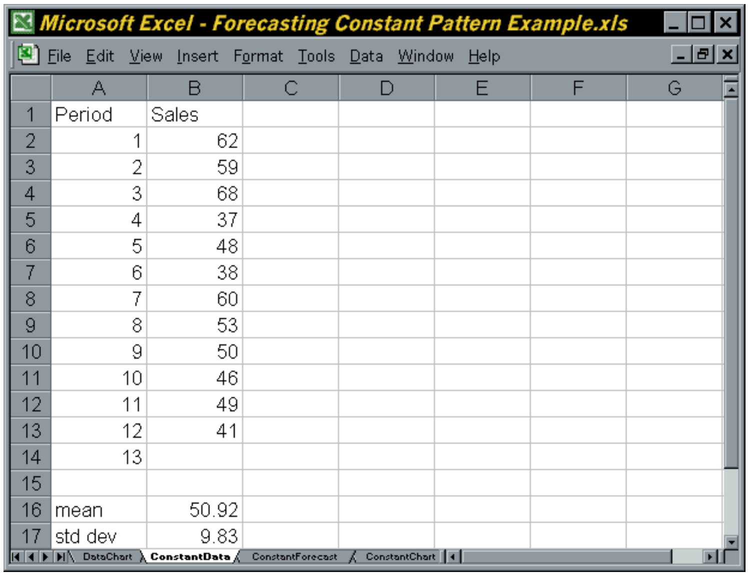
The equivalent exponential smoothing constant for a number of interval lengths of the moving averages method are shown in Table 2.2. Hence, a highly reactive exponential smoothing constant of 0.4 is equivalent to a moving averages method with an interval of 4 if we want the two methods to have the same average age of data. Similarly, a more stable exponential smoothing constant of 0.2 is equivalent to a moving averages method with an interval of 9.

Table 2.2. Equivalent Parameters for Moving Averages and Exponential Smoothing

N	α	N	α
2	0.67	8	0.22
3	0.50	9	0.20
4	0.40	10	0.18
5	0.33	11	0.17
6	0.29	12	0.15
7	0.25		

Constant Pattern Forecasting Example

In the example we will determine the sales forecast for the next month of a product, which is assumed to have a constant sales pattern. The available data are the monthly sales values for the past year. The numerical data are shown in Figure 2.3 and the graph of the historical data is shown in Figure 2.4.



Microsoft Excel - Forecasting Constant Pattern Example.xls

File Edit View Insert Format Tools Data Window Help

	A	B	C	D	E	F	G
1	Period	Sales					
2	1	62					
3	2	59					
4	3	68					
5	4	37					
6	5	48					
7	6	38					
8	7	60					
9	8	53					
10	9	50					
11	10	46					
12	11	49					
13	12	41					
14	13						
15							
16	mean	50.92					
17	std dev	9.83					

DataChart ConstantData ConstantForecast ConstantChart

Figure 2.3. Historical Data for the Constant Pattern Example

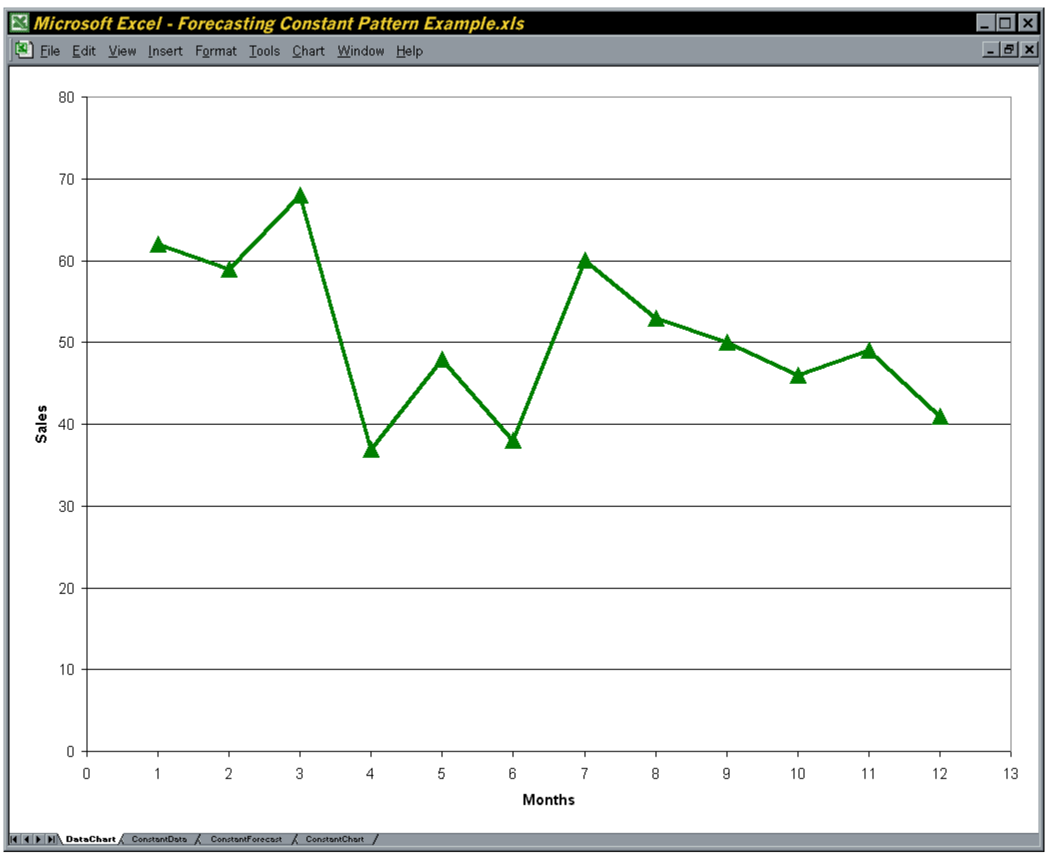


Figure 2.4. Graph of the Historical Data for the Constant Pattern Example

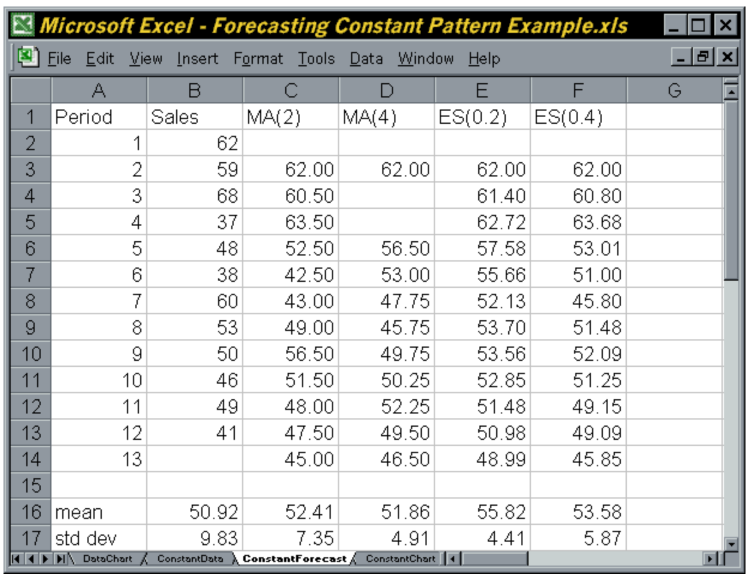
We will use four methods to forecast the next value based on the historical data: moving averages with an interval of 2, denoted by MA(2), moving averages with an interval of 4, denoted by MA(4), (simple) exponential smoothing with a smoothing constant of 0.2, denoted by ES(0.2), and exponential smoothing with a smoothing constant of 0.4, denoted by ES(0.4). The moving averages methods use partial data during the initial periods of the forecasting process to start up their process. The forecasts were computed with the formulas introduces above and collected in a spreadsheet, which is shown in Figure 2.5. The graph of the forecasts is shown in Figure 2.6. The forecasted values for period 13 range from 45 to 49 depending on the forecasting method used. The computations of the forecast for period 13 by the four different methods are given next.

$$F_{13}^{MA(2)} = \frac{49+41}{2} = 45.00$$

$$F_{13}^{MA(4)} = \frac{50+46+49+41}{4} = 46.50$$

$$F_{13}^{ES(0.2)} = 0.2 \cdot 41 + 0.8 \cdot 50.98 = 48.99$$

$$F_{13}^{ES(0.4)} = 0.4 \cdot 41 + 0.6 \cdot 49.09 = 45.85$$



	A	B	C	D	E	F	G
1	Period	Sales	MA(2)	MA(4)	ES(0.2)	ES(0.4)	
2	1	62					
3	2	59	62.00	62.00	62.00	62.00	
4	3	68	60.50		61.40	60.80	
5	4	37	63.50		62.72	63.68	
6	5	48	52.50	56.50	57.58	53.01	
7	6	38	42.50	53.00	55.66	51.00	
8	7	60	43.00	47.75	52.13	45.80	
9	8	53	49.00	45.75	53.70	51.48	
10	9	50	56.50	49.75	53.56	52.09	
11	10	46	51.50	50.25	52.85	51.25	
12	11	49	48.00	52.25	51.48	49.15	
13	12	41	47.50	49.50	50.98	49.09	
14	13		45.00	46.50	48.99	45.85	
15							
16	mean	50.92	52.41	51.86	55.82	53.58	
17	std dev	9.83	7.35	4.91	4.41	5.87	

Figure 2.5. Numerical Forecasts for the Constant Pattern Example

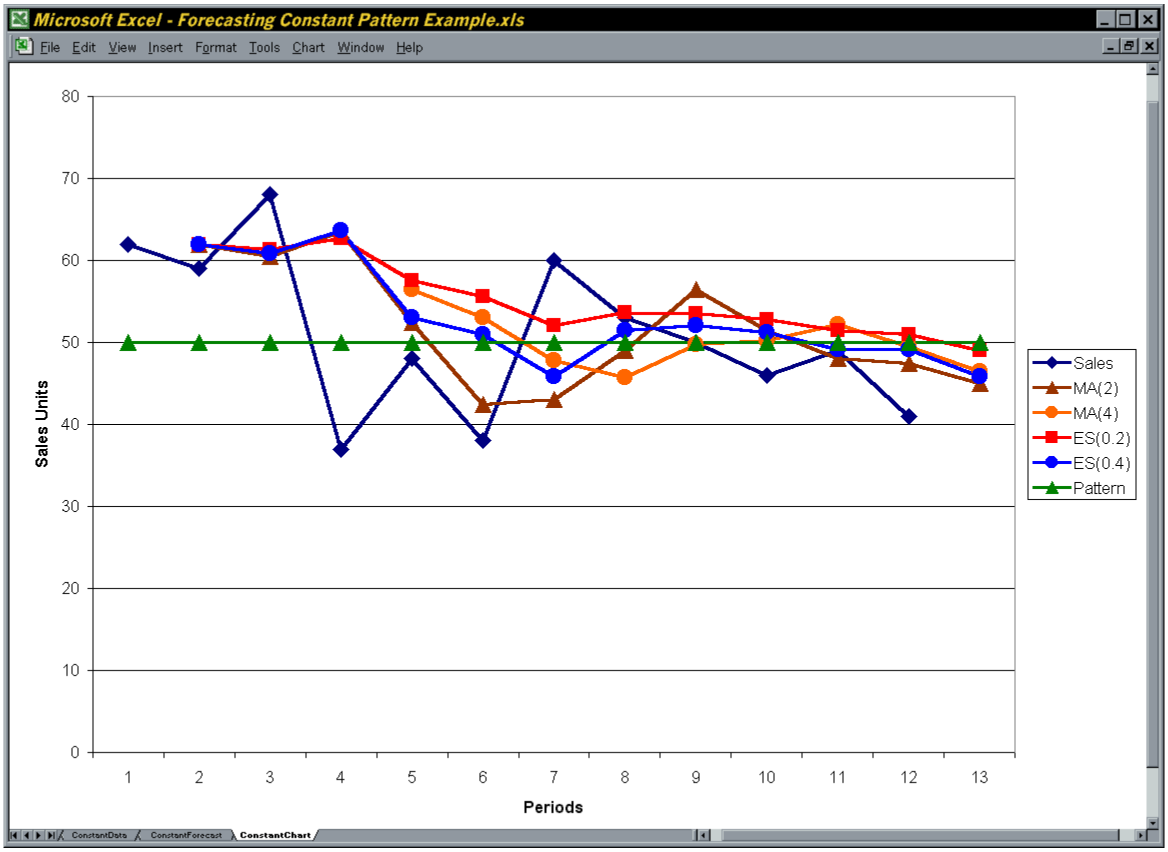


Figure 2.6. Graph of Forecasts for the Constant Pattern Example

The obvious question is which forecast is "correct" or the best? In the next spreadsheet the standard measures of forecast performance are computed based on the forecast errors for each method and period based on the formulas introduced above. The computations for the forecast errors for period 12 by the four different forecasting methods are shown next.

$$e_{12}^{MA(2)} = 47.50 - 41 = 6.50$$
$$e_{12}^{MA(4)} = 49.50 - 41 = 8.50$$
$$e_{12}^{ES(0.2)} = 50.98 - 41 = 9.98$$
$$e_{12}^{ES(0.4)} = 49.09 - 41 = 8.09$$

The forecasting method with the smallest mean square error (MSE) and mean absolute percentage deviation (MAPE) is the moving averages method based on two data points MA(2). This method also has the smallest average error. So all performance measures for this example indicate the selection of MA(2) as the preferred forecasting method.

Microsoft Excel - Forecasting Constant Pattern Example.xls

	I	J	K	L	M	N	O	P	Q	R
1	Period	Sales	MA(2)	eMA(2)	MA(4)	eMA(4)	ES(0.2)	eES(0.2)	ES(0.4)	eES(0.4)
2	1	62								
3	2	59	62.00	3.00	62.00	3.00	62.00	3.00	62.00	3.00
4	3	68	60.50	-7.50	60.50	-7.50	61.40	-6.60	60.80	-7.20
5	4	37	63.50	26.50	63.00	26.00	62.72	25.72	63.68	26.68
6	5	48	52.50	4.50	56.50	8.50	57.58	9.58	53.01	5.01
7	6	38	42.50	4.50	53.00	15.00	55.66	17.66	51.00	13.00
8	7	60	43.00	-17.00	47.75	-12.25	52.13	-7.87	45.80	-14.20
9	8	53	49.00	-4.00	45.75	-7.25	53.70	0.70	51.48	-1.52
10	9	50	56.50	6.50	49.75	-0.25	53.56	3.56	52.09	2.09
11	10	46	51.50	5.50	50.25	4.25	52.85	6.85	51.25	5.25
12	11	49	48.00	-1.00	52.25	3.25	51.48	2.48	49.15	0.15
13	12	41	47.50	6.50	49.50	8.50	50.98	9.98	49.09	8.09
14	13		45.00		46.50		48.99		45.85	
15										
16	Avg	50.92		2.50		3.75		5.91		3.67
17	St Dev	9.83		10.57		11.05		11.06		10.74
18	MSE			111.70		122.01		122.32		115.29
19	MAPE			17.06		19.52		19.69		17.69

ConstantData / ConstantForecast / ConstantChart

Figure 2.7. Forecast Errors and Performance Measures for the Constant Pattern Example

This does not imply that we should report that 45 as the forecasted value of the sales in period 13. We can however state that the expected value and median of forecasted sales is 45 and that there is a 68 % change that the sales in period 13 will fall in the one σ range 45 ± 10.6 or [34, 56], a 95 % change that sales will fall in the two σ range 45 ± 21.2 or [23, 67], and almost a 100 % change that the sales will fall in the three σ range 45 ± 31.8 or [13, 77].

The Excel spreadsheet includes in its analysis tools the moving averages and (simple) exponential smoothing methods. It also uses the term *damping factor* as the complement of the exponential smoothing constant. In other words, the sum of the smoothing constant and the damping factor is by definition equal to one. You must specify the damping factor as the parameter of the exponential smoothing method. Excel correctly suggest a range of 0.2 to 0.4 for the exponential smoothing constant, but it erroneously states that the default damping factor is 0.3, when in fact the default damping factor is 0.7 and the default smoothing constant is 0.3. The Excel spreadsheet computes the standard deviation of the forecast error only on the last three observations and forecasts for the exponential smoothing method and on the observations in the interval for the moving averages method.

As a final observation, the historical sales values were sampled from a normal distribution with a mean equal to 50 and a standard deviation of 10, which yields a coefficient of variation equal to 0.2. The exponential smoothing method ES(0.2), with smoothing constant equal to 0.2, generated a sales forecast closest to the true expected value of the sales in period 13. However, the underlying data pattern is almost never known in real-life forecasting projects.

Trend Data Pattern

Trend Data Pattern Characteristics

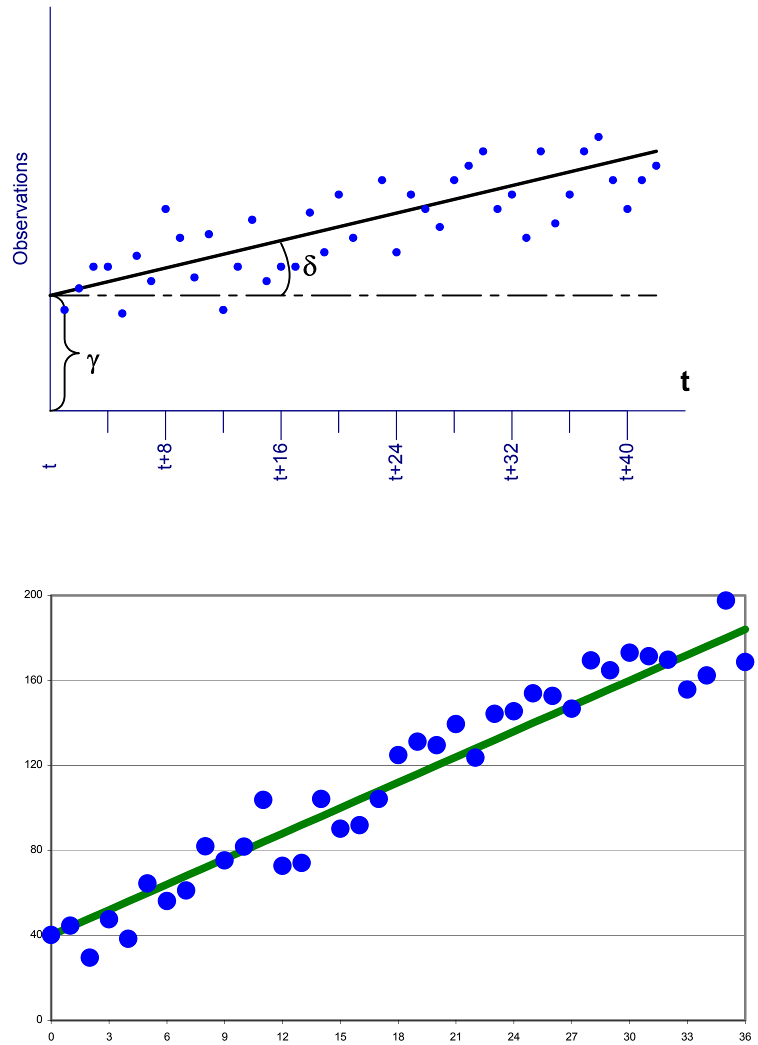


Figure 2.8. Illustration of a Time Series Following a Trend Pattern

A trend pattern has two parameters, the offset of the trend at time zero, which is denoted by γ, and the slope, which is denoted by δ. The forecasted values for future periods are not identical. Since the mean of the forecast error is assumed to be zero, the computation of the standard deviation of the forecast error uses N in the denominator. The forecasted values are illustrated in Figure 2.9.

$$\begin{aligned} D_t &= \gamma + \delta t + \varepsilon_t \\ \varepsilon_t &= N\left(0, \sigma^2\right) \end{aligned} \tag{2.29}$$

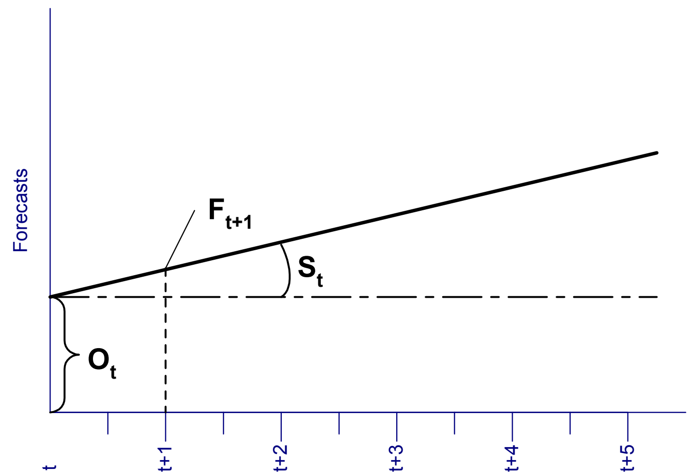


Figure 2.9. Trend Pattern Forecasts

Forecasting Methods for a Trend Data Pattern

Polynomial Fitting

Polynomial fitting methods determine the coefficients of a polynomial that fits the observations most closely by minimizing the sum of squares of the deviations between the polynomial and observed values.

In case of an underlying trend pattern, a first-degree polynomial is used. This implies that two parameters have to be determined, which are called the offset (O) and the slope (S).

$$F_t = O + t \cdot S \tag{2.30}$$

The underlying pattern does not have to be a linear trend, since higher-level polynomials can be used. However, high-level polynomials increase the volatility or nervousness of the forecast and should be used with great care. Further information on polynomial fitting can be found in Montgomery et al.

(2001). The Excel software can be used to perform linear curve fitting with the functions INTERCEPT and SLOPE that compute the intercept and slope of the fitted line, respectively.

Double Exponential Smoothing (Holt's Method)

The double exponential smoothing method develop by Holt (1957) for data with a constant linear trend data uses two smoothing constants to adjust the intercept and slope. At the end of period *t*, when the real demand for this period is known, the parameters are updated based on the following formulas:

$$\begin{aligned} O_t &= \alpha D_t + (1-\alpha)(O_{t-1} + S_{t-1}) \\ S_t &= \beta(O_t - O_{t-1}) + (1-\beta)S_{t-1} \end{aligned} \tag{2.31}$$

The following initial values are typically used

$$\begin{aligned} O_1 &= D_1 \\ O_2 &= D_2 \\ S_2 &= D_2 - D_1 \end{aligned} \tag{2.32}$$

This requires that at minimum two observations or data points are available. If a larger number of observations is available, linear regression is typically used to estimate the slope and intercept of the underlying pattern. The forecasts are then computed with the following expressions

$$\begin{aligned} F_{t+1} &= O_t + S_t \\ F_{t,t+\tau} &= O_t + \tau \cdot S_t \end{aligned} \tag{2.33}$$

The structure of the update process for the offset of forecast variable O_t and for the slope of the variable S_t is identical. The new value is computed as the sum of the last observed value multiplied by the corresponding smoothing constant and the previous value multiplied by one minus the smoothing constant or damping factor. The new forecast for the next period is equal to the offset plus one times the slope. The new forecast for periods further in the future is now different from the forecast for the next period and obtained by multiplying the slope with τ, the number of periods between the future period and the current time period.

Just as in the simple exponential smoothing method for the constant data pattern, the values of the two smoothing constants in the double exponential smoothing method should be based on careful judgment of the behavior of the underlying data pattern. Larger smoothing constants allow faster reaction of the forecasting model, which results in a smaller lag but a more "nervous" forecasting method. Smaller smoothing constants delay the reaction of the forecast model and provide for a more "stable" forecasting method. Typically, values for α are in the [0.1, 0.4] range with 0.2 as a typical starting value. It is usually desirable to have more stable behavior of the slope, so 0.1 is often chosen as the starting value for β. The values of the smoothing parameters can also set to minimize the sum of the squared errors over an historical interval.

The double smoothing process requires a value for the initial offset and the initial slope. If enough historical data are available, then these values can be computed with a linear regression model fitted through the historical observations. The double exponential smoothing process is then used to update the offset and slope and generate the next forecasts.

Consider the following small example. It is assumed that the underlying pattern follows a linear trend. The first four observations are 53, 47, 59, and 70. The linear regression through these four data points yields an intercept of 41.50 and a slope of 6.30. The offset of the line at period four is then equal to 66.70. The initial values of both smoothing constants were set equal to 0.2. The forecast for period five based on the offset and slope of period four is then

$$F_5 = O_4 + S_4 = 66.70 + 6.30 = 73$$

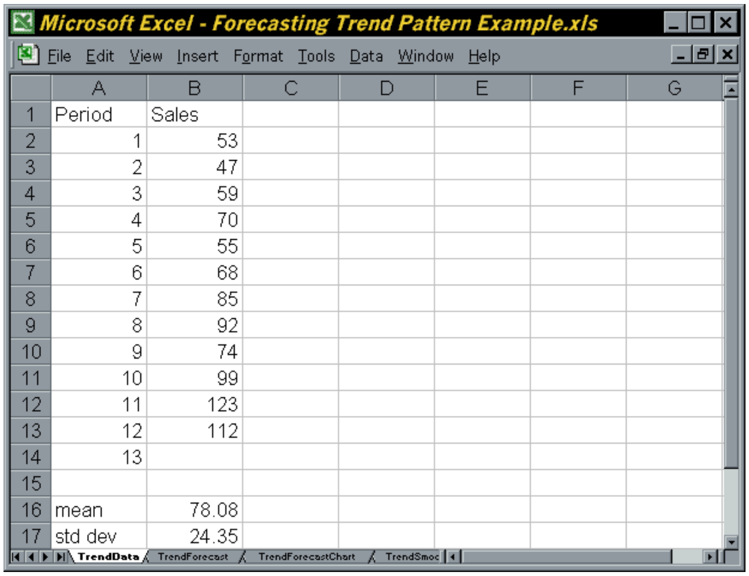
The new observed value is equal to 55, which implies that the forecast error for period five is equal to 18. The offset and slope for period five are now updated. Finally, the new forecast for period six is computed.

$$\begin{aligned}
O_5 &= \alpha D_5 + (1-\alpha)(O_4 + S_4) = 0.2 \cdot 55 + 0.8 \cdot (66.70 + 6.30) = 69.40 \\
S_5 &= \beta(O_5 - O_4) + (1-\beta)S_4 = 0.2 \cdot (69.40 - 66.70) + 0.8 \cdot 6.30 = 5.58 \\
F_6 &= O_5 + S_5 = 69.40 + 5.58 = 74.98
\end{aligned}$$

This process is repeated when future observations become available.

The complete example comparing double exponential smoothing methods with different smoothing constants for an underlying data trend pattern is given at the end of the section on the trend pattern.

Trend Pattern Forecasting Example


Microsoft Excel - Forecasting Trend Pattern Example.xls

	A	B
1	Period	Sales
2	1	53
3	2	47
4	3	59
5	4	70
6	5	55
7	6	68
8	7	85
9	8	92
10	9	74
11	10	99
12	11	123
13	12	112
14	13	
15		
16	mean	78.08
17	std dev	24.35

Figure 2.10. Historical Data for the Trend Pattern Example

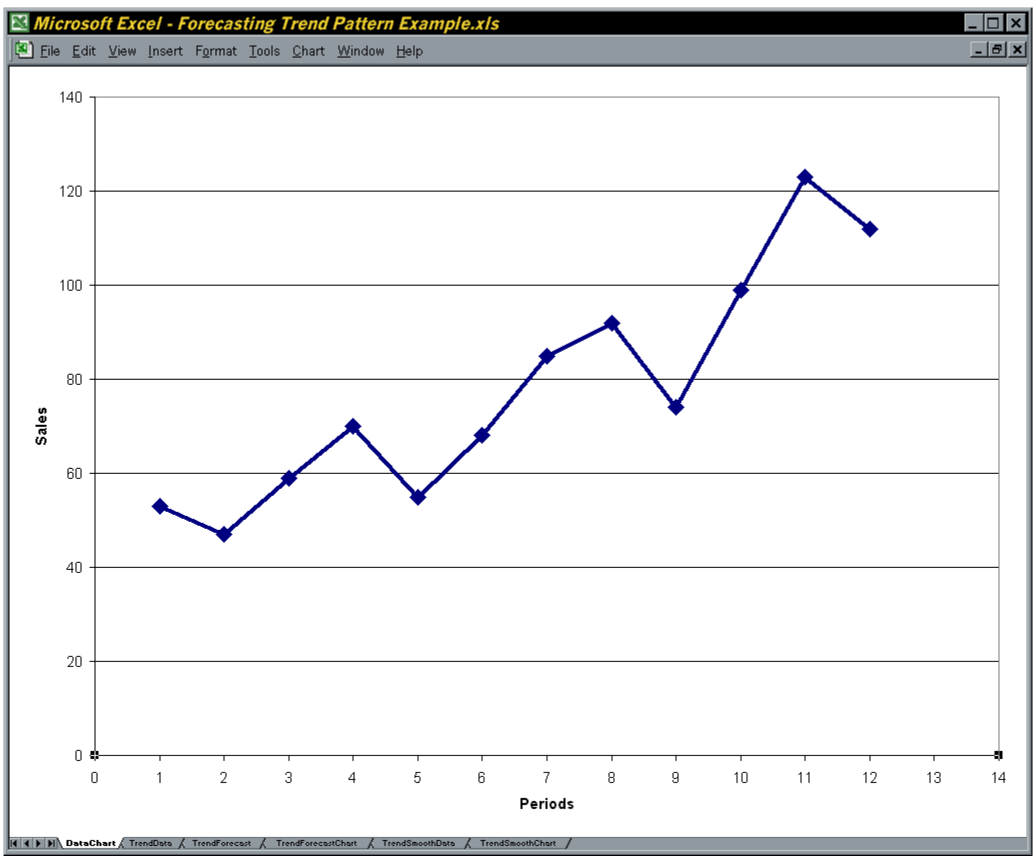
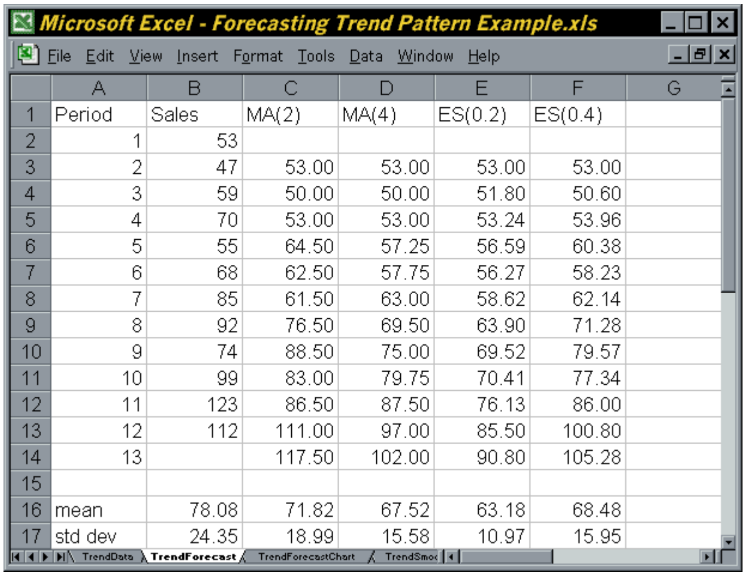


Figure 2.11. Graph of the Historical Data for the Trend Pattern Example



Microsoft Excel - Forecasting Trend Pattern Example.xls

	A	B	C	D	E	F	G
1	Period	Sales	MA(2)	MA(4)	ES(0.2)	ES(0.4)	
2	1	53					
3	2	47	53.00	53.00	53.00	53.00	
4	3	59	50.00	50.00	51.80	50.60	
5	4	70	53.00	53.00	53.24	53.96	
6	5	55	64.50	57.25	56.59	60.38	
7	6	68	62.50	57.75	56.27	58.23	
8	7	85	61.50	63.00	58.62	62.14	
9	8	92	76.50	69.50	63.90	71.28	
10	9	74	88.50	75.00	69.52	79.57	
11	10	99	83.00	79.75	70.41	77.34	
12	11	123	86.50	87.50	76.13	86.00	
13	12	112	111.00	97.00	85.50	100.80	
14	13		117.50	102.00	90.80	105.28	
15							
16	mean	78.08	71.82	67.52	63.18	68.48	
17	std dev	24.35	18.99	15.58	10.97	15.95	

Figure 2.12. Numerical Forecasts for the Trend Pattern Example

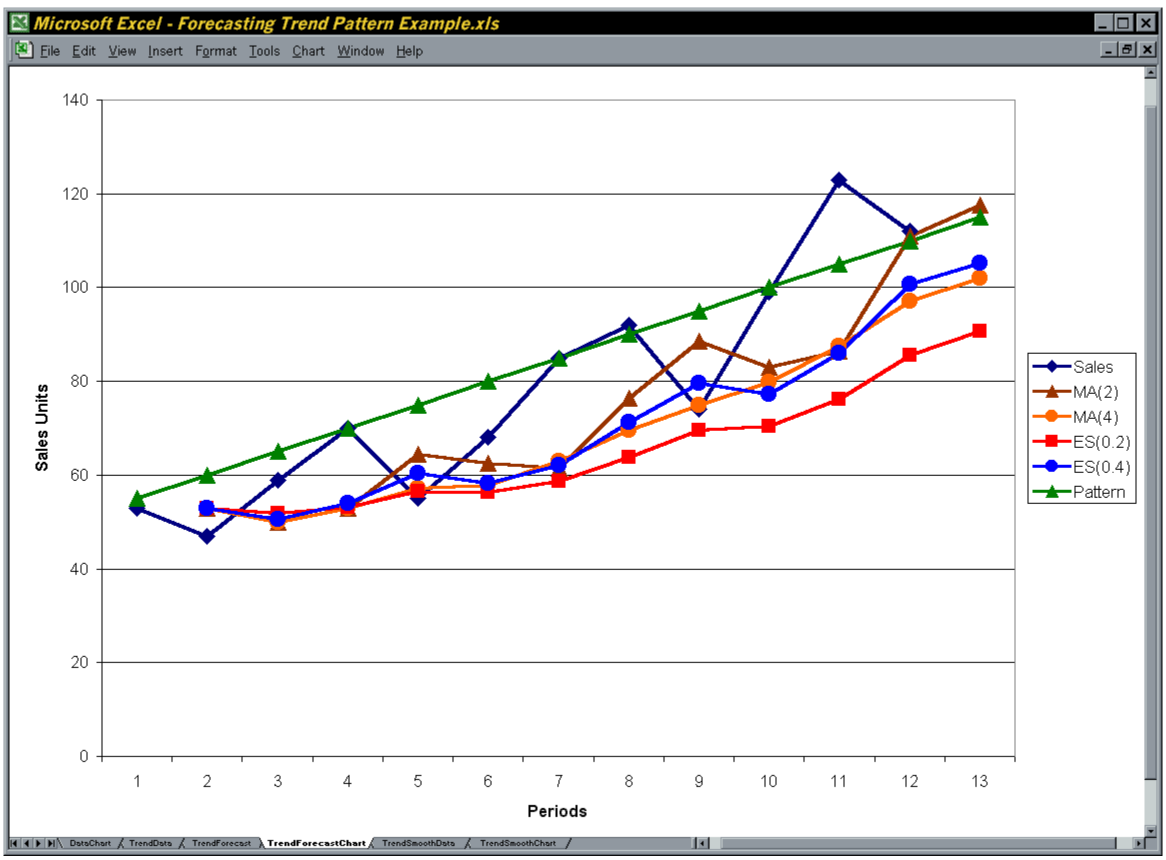


Figure 2.13. Graph of Forecasts for the Trend Pattern Example

Microsoft Excel - Forecasting Trend Pattern Example.xls

	I	J	K	L	M	N	O	P	Q	R
1	Period	Sales	MA(2)	eMA(2)	MA(4)	eMA(4)	ES(0.2)	eES(0.2)	ES(0.4)	eES(0.4)
2	1	53								
3	2	47	53.00	6.00	53.00	6.00	53.00	6.00	53.00	6.00
4	3	59	50.00	-9.00	50.00	-9.00	51.80	-7.20	50.60	-8.40
5	4	70	53.00	-17.00	53.00	-17.00	53.24	-16.76	53.96	-16.04
6	5	55	64.50	9.50	57.25	2.25	56.59	1.59	60.38	5.38
7	6	68	62.50	-5.50	57.75	-10.25	56.27	-11.73	58.23	-9.77
8	7	85	61.50	-23.50	63.00	-22.00	58.62	-26.38	62.14	-22.86
9	8	92	76.50	-15.50	69.50	-22.50	63.90	-28.10	71.28	-20.72
10	9	74	88.50	14.50	75.00	1.00	69.52	-4.48	79.57	5.57
11	10	99	83.00	-16.00	79.75	-19.25	70.41	-28.59	77.34	-21.66
12	11	123	86.50	-36.50	87.50	-35.50	76.13	-46.87	86.00	-37.00
13	12	112	111.00	-1.00	97.00	-15.00	85.50	-26.50	100.80	-11.20
14	13		117.50		102.00		90.80		105.28	
15										
16	Avg	78.08		-8.55		-12.84		-17.18		-11.88
17	St Dev	24.35		16.84		17.49		22.82		17.67
18	MSE			283.50		305.74		520.98		312.23
19	MAPE			17.14		16.81		20.67		17.54

DataChart / TrendData / TrendForecast / TrendForecastChart / TrendSmoothData / TrendS

Figure 2.14. Constant Forecast Errors and Performance Measures for the Trend Pattern Example

Microsoft Excel - Forecasting Trend Pattern Example.xls

	A	B	C	D	E	F	G	H	I	J
5	Period	Sales	DES(0.2)	eDES(0.2	offset(0.2)	slope(0.2)	DES(0.4)	eDES(0.4	offset(0.4)	slope(0.4)
6	1	53								
7	2	47								
8	3	59								
9	4	70			66.70	6.30			66.70	6.30
10	5	55	73.00	18.00	69.40	5.58	73.00	18.00	65.80	4.86
11	6	68	74.98	6.98	73.58	5.30	70.66	2.66	69.60	4.65
12	7	85	78.88	-6.12	80.11	5.55	74.24	-10.76	78.55	5.51
13	8	92	85.65	-6.35	86.92	5.80	84.05	-7.95	87.23	6.14
14	9	74	92.72	18.72	88.98	5.05	93.38	19.38	85.63	4.59
15	10	99	94.03	-4.97	95.02	5.25	90.22	-8.78	93.73	5.30
16	11	123	100.27	-22.73	104.82	6.16	99.03	-23.97	108.62	7.21
17	12	112	110.98	-1.02	111.18	6.20	115.83	3.83	114.30	6.91
18	13		117.38				121.21			
19										
20	mean	78.08		0.31				-0.95		
21	std dev	24.35		12.96				13.93		
22	MSE			167.91				193.98		
23	MAPE			0.10				0.11		

TrendData / TrendForecast / TrendForecastChart / TrendSmoothData / TrendSmoothChart

Figure 2.15. Numerical Trend Forecasts for the Trend Pattern Example

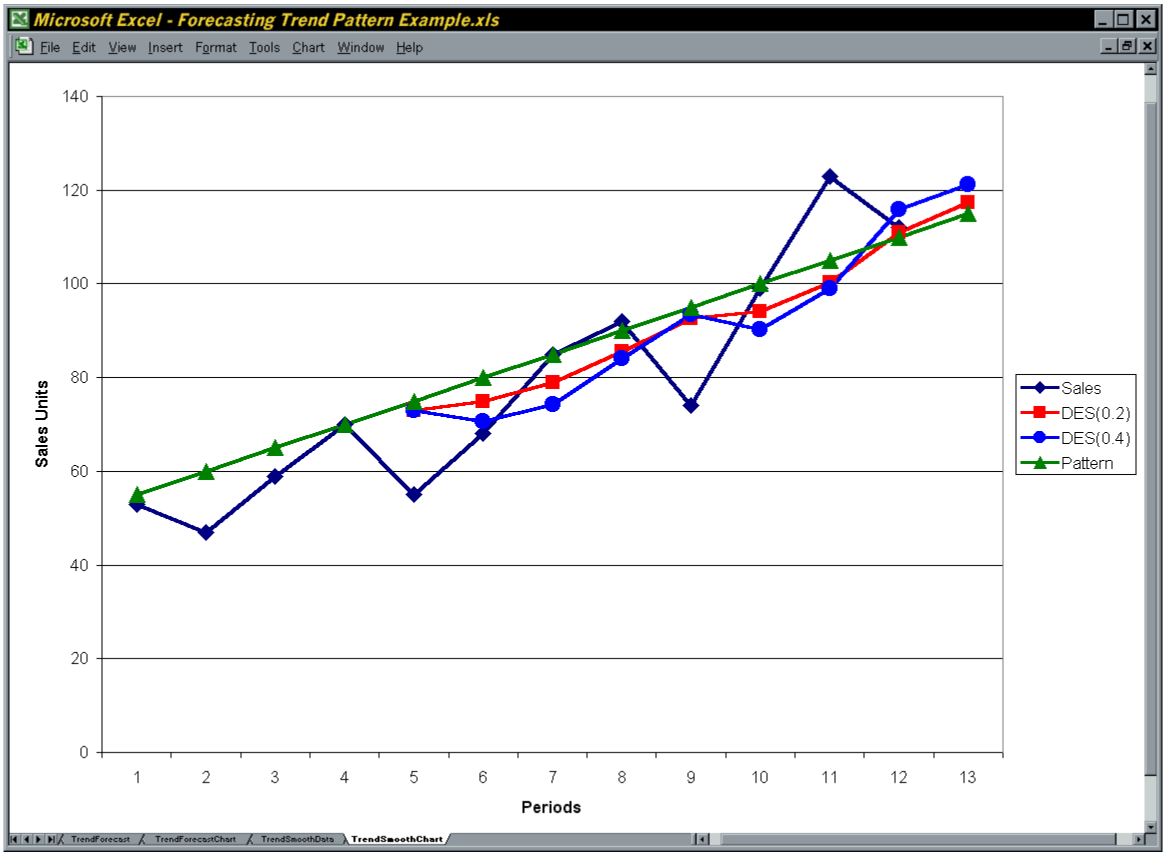


Figure 2.16. Graph of Trend Forecasts for the Trend Pattern Example

It can be observed in the above figure that the exponential smoothing with 0.4 as smoothing constant reacts quicker to the changes in the observed values than the exponential smoothing with 0.2 as smoothing constant. For example, at period eleven both forecast are very close together. However, the observed value for period eleven is much larger than forecasted. The forecast by DES(0.4) for period twelve is then larger than the forecast by DES(0.2).

Seasonal Data Pattern

Seasonal Data Pattern Characteristics

Forecasting Methods for a Seasonal Data Pattern

Classical Decomposition

The classical additive decomposition assumes that the data are generated by the following process

$$
\begin{aligned}
& D_t = Pat_t + C_t + \varepsilon_t \\
& \sum_{t=1}^{P} C_t = 0 \qquad (2.34) \\
& \varepsilon_t = N\left(0, \sigma^2\right)
\end{aligned}
$$

where Pat_t is the trend component, C_t is the seasonality component with a known period, and ε is the random component with a zero expected value. The data is "deseasonalized" by subtracting the seasonal component, applying a forecasting method for a trend based pattern, and then "reseasonalized" by adding the forecast of the seasonal component.

The multiplicative decomposition assumes that the data are generated by the process in which the pattern is multiplied by the seasonal effect, where c_t is the seasonality factor with a known period. The data is "deseasonalized" by dividing by the seasonal factor, applying a forecasting method for a trend based pattern, and then "reseasonalized" by multiplying by the forecast of the seasonal factor.

$$
\begin{aligned}
& D_t = Pat_t \cdot c_t + \varepsilon_t \\
& \sum_{t=1}^{P} c_t = P \qquad (2.35) \\
& \varepsilon_t = N\left(0, \sigma^2\right)
\end{aligned}
$$

In the additive variant, the amplitude of the seasonal variations stays constant over time, while in the multiplicative variant the amplitude of the seasonal variations is proportional to the amplitude of the underlying linear trend. The additive and multiplicative seasonal patterns are illustrated in the next two figures.

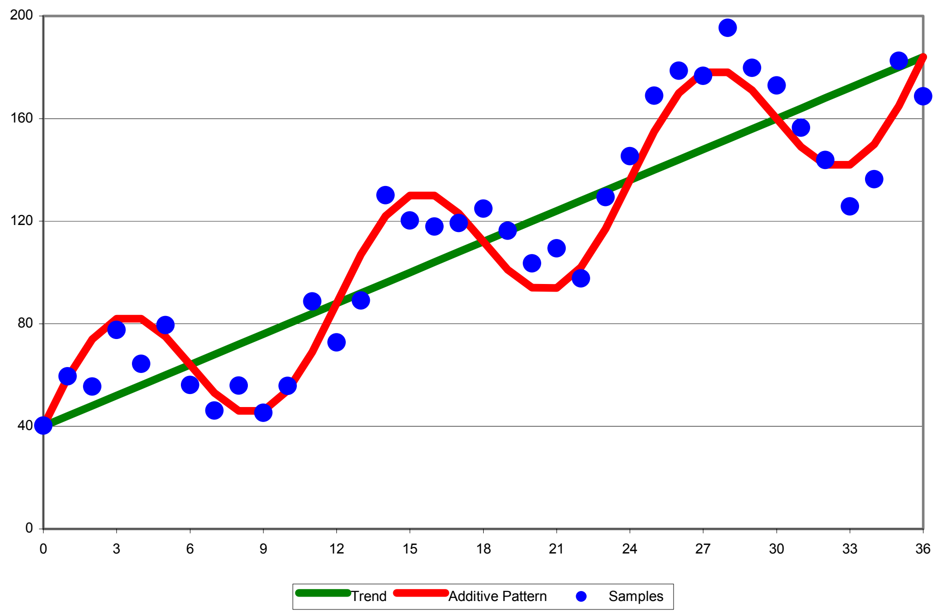


Figure 2.17. Additive Seasonal Pattern Illustration

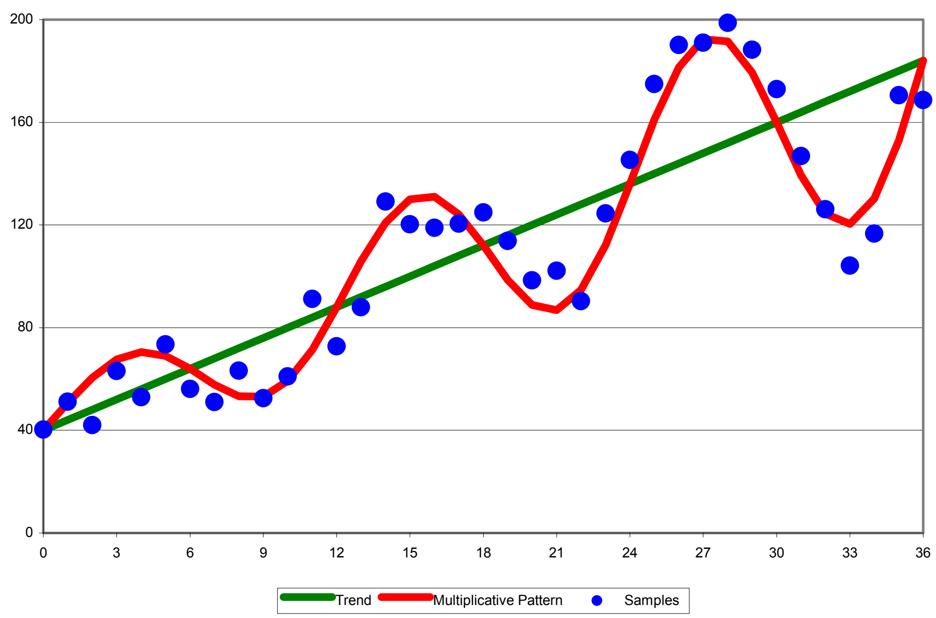


Figure 2.18. Multiplicative Seasonal Pattern

Triple Exponential Smoothing (Winter's Method)

There exist several variants of the triple exponential smoothing method developed by Winter (1970) depending on the assumption about the behavior of the underlying process. In each variant, the triple exponential smoothing method uses three smoothing constants to adjust the intercept, slope, and seasonality component C_t or seasonality factor c_t corresponding to the current period.

For the additive variant the trend and the seasonal components are added together and the underlying process is assumed to be

$$\begin{aligned} D_t &= Pat_t + C_t + \varepsilon = O + t \cdot S + C_t + \varepsilon_t \\ \varepsilon_t &= N\left(0, \sigma^2\right) \end{aligned} \tag{2.36}$$

At the end of period *t*, when the real demand for this period is known, the parameters are updated based on the following formulas:

$$\begin{aligned} O_t &= \alpha\left(D_t - C_{t-P}\right) + (1-\alpha)(O_{t-1} + S_{t-1}) \\ S_t &= \beta(O_t - O_{t-1}) + (1-\beta)S_{t-1} \\ C_t &= \gamma\left(D_t - O_t\right) + (1-\gamma)C_{t-P} \end{aligned} \tag{2.37}$$

The starting values are computed with

$$\begin{aligned} O_t &= D_t \\ S_t &= \left(D_t - D_{t-P}\right)/P \\ C_i &= D_i - \left(D_1 + \left(i-1\right) \cdot S_t\right) \qquad\qquad i = 1..P \end{aligned} \tag{2.38}$$

Observe that this implies the initial estimate of $C_1 = 0$.

The forecasts for the next periods are then computed with

$$\begin{aligned} F_{t+1} &= O_t + S_t + C_{t+1-P} \\ F_{t,t+\tau} &= O_t + \tau S_t + C_{t+\tau-P} \end{aligned} \tag{2.39}$$

To compute the initial values of the intercept, slope, and seasonality components, at least one full cycle plus one additional observation of historical data are required.

Just as in the simple exponential smoothing method for the constant data pattern, the values of the three smoothing constants in the triple exponential smoothing method should be based on careful judgment of the behavior of the underlying data pattern. Typically, values for α are in the [0.1, 0.4] range with 0.2 as a typical starting value. It is usually desirable to have a more stable behavior of the slope and seasonality factors or components, so 0.1 is often chosen as the starting value for β and γ. The values of the smoothing parameters can also set to minimize the sum of the squared errors over an historical interval.

The double smoothing process requires a value for the initial offset and the initial slope. If enough historical data are available, then these values can be computed with a linear regression model fitted through the historical observations. The triple exponential smoothing process is then used to update the offset and slope and generate the next forecasts.

The multiplicative variant of the triple exponential smoothing method for seasonal data assumes that the data are generated by the process in which the pattern is multiplied by the seasonal effect.

$$D_t = Pat_t \cdot c_t + \varepsilon_t \tag{2.40}$$

At the end of period t, when the real demand for this period is known, the parameters are updated based on the following formulas, where || indicates the modulo operator:

$$\begin{aligned} O_t &= \alpha\left(D_t / c_t\right) + (1-\alpha)(O_{t-1} + S_{t-1}) \\ S_t &= \beta(O_t - O_{t-1}) + (1-\beta)S_{t-1} \\ c_{t\|P} &= \gamma\left(D_t / O_t\right) + (1-\gamma)c_{t-P\|P} \end{aligned} \tag{2.41}$$

All the seasonality factors are then normalized so that their sum equals the number of periods in the cycle and the forecast for the next and future periods is then computed based on the normalized seasonality factors.

$$c_t = c_t \cdot \frac{P}{\sum_{j=1}^{P} c_j} \tag{2.42}$$

$$\begin{aligned} F_{t+1} &= \left(O_t + S_t\right) c_{t+1\|P} \\ F_{t,t+\tau} &= \left(O_t + \tau S_t\right) c_{t+\tau\|P} \end{aligned} \tag{2.43}$$

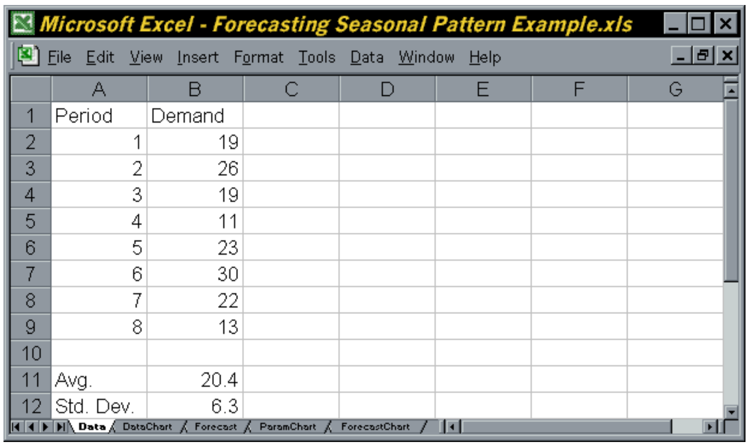
To compute the initial values of the intercept, slope, and seasonality factors, at least two full cycles of historical data are required.

There exists also a multiplicative version of the triple exponential smoothing method, which assumes that the underlying process is given by

$$D_t = Pat_t \cdot c_t \cdot \varepsilon$$
$$E(\varepsilon) = 1 \tag{2.44}$$

Further information on forecasting based on classic time series decomposition for this underlying seasonal pattern can be found in Ballou (1998).

Seasonal Pattern Forecasting Example



	A	B	C	D	E	F	G
1	Period	Demand					
2	1	19					
3	2	26					
4	3	19					
5	4	11					
6	5	23					
7	6	30					
8	7	22					
9	8	13					
10							
11	Avg.	20.4					
12	Std. Dev.	6.3					

Figure 2.19. Historical Data for the Seasonal Pattern Example

As the first step in understanding the underlying pattern, the numerical data is displayed graphically. The pattern appears to be seasonal with an upward trend. It is not clear if the amplitude of the seasonal oscillations increases with the linear trend. In real-life cases this indicates the simultaneous use of both the additive and multiplicative variants. The example will continue using the multiplicative variant.

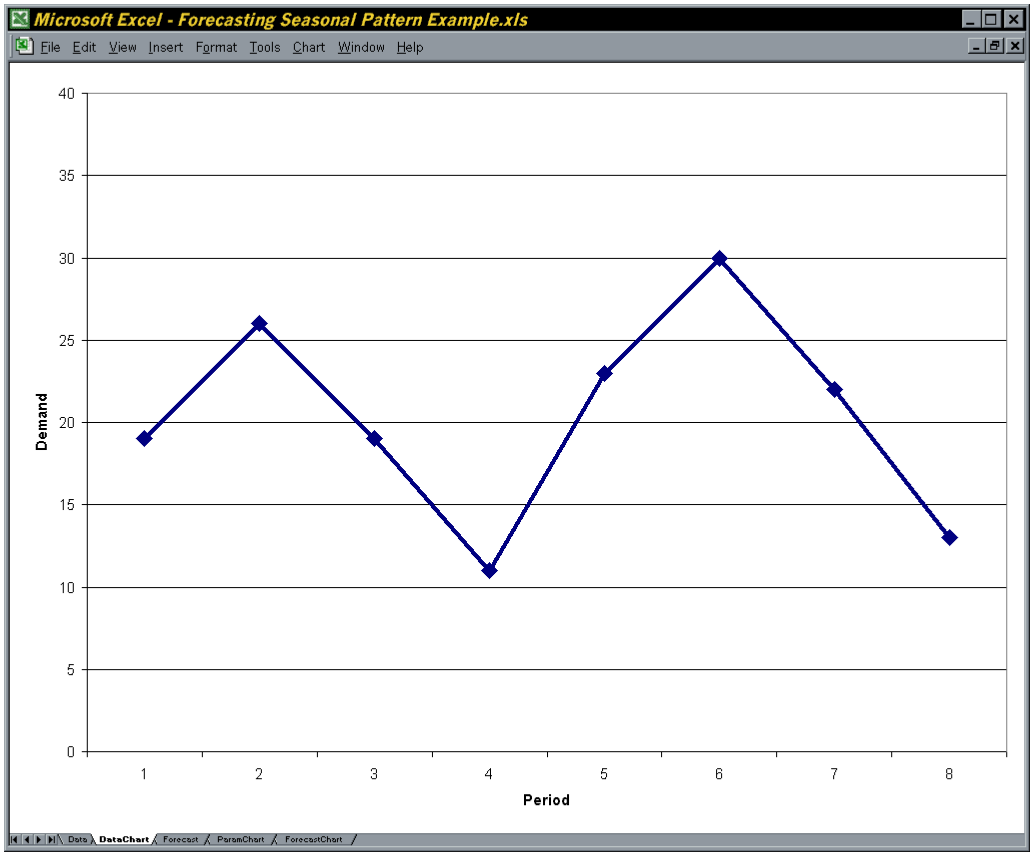


Figure 2.20. Graph of Historical Data for the Seasonal Pattern Example

The first task is to determine the cycle length of the underlying seasonal pattern. In this particular example this task is not difficult. The time interval between two successive high values or peaks is four periods and the time interval between two successive low values or valleys is also four periods. So we assume that the underlying seasonal pattern has a cycle length of four periods.

The next task is to compute initial estimates of the underlying trend. We will compute an average value for each of the cycles and fit a straight line through those points. Cycle one covers periods one through four and its average time coordinate is thus 2.5, cycle two covers periods five through eight and its average time coordinate is thus 6.5. The computations for the average demand for each cycle and the slope and offset in period eight are shown next and are illustrated in the following figure.

$$\overline{x_{C1}} = \frac{1+4}{2} = 2.5$$

$$\overline{y_{C1}} = \frac{19+26+19+11}{4} = 18.75$$

$$\overline{x_{C2}} = \frac{5+8}{2} = 6.5$$

$$\overline{y_{C1}} = \frac{23+30+22+13}{4} = 22$$

$$S_8 = \frac{22-18.75}{6.5-2.5} = 0.81$$

$$O_8 = 22 + (8-6.5)\cdot 0.81 = 23.22$$

If more than two cycles of data are available the slope and intercepts can be compute with linear regression techniques.

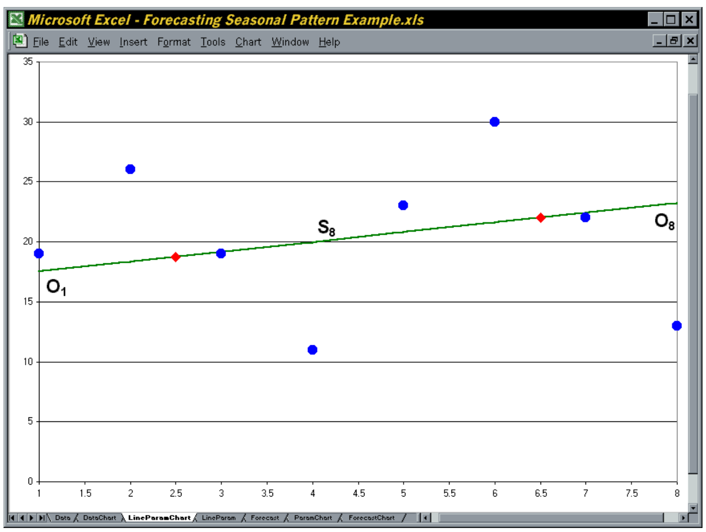
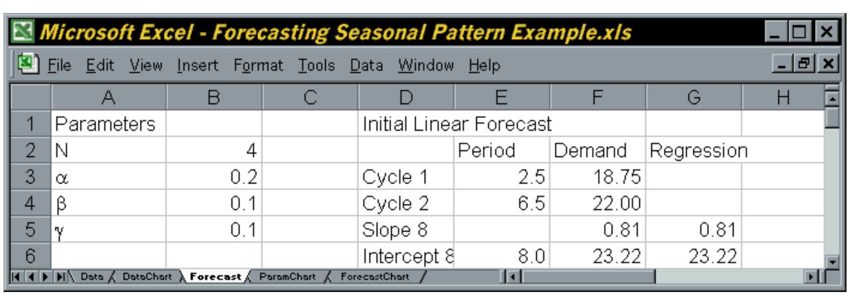


Figure 2.21. Initial Linear Forecast Parameters Graph for the Seasonal Pattern Example



	A	B	C	D	E	F	G	H
1	Parameters			Initial Linear Forecast				
2	N	4			Period	Demand	Regression	
3	α	0.2		Cycle 1	2.5	18.75		
4	β	0.1		Cycle 2	6.5	22.00		
5	γ	0.1		Slope 8		0.81	0.81	
6				Intercept 8	8.0	23.22	23.22	

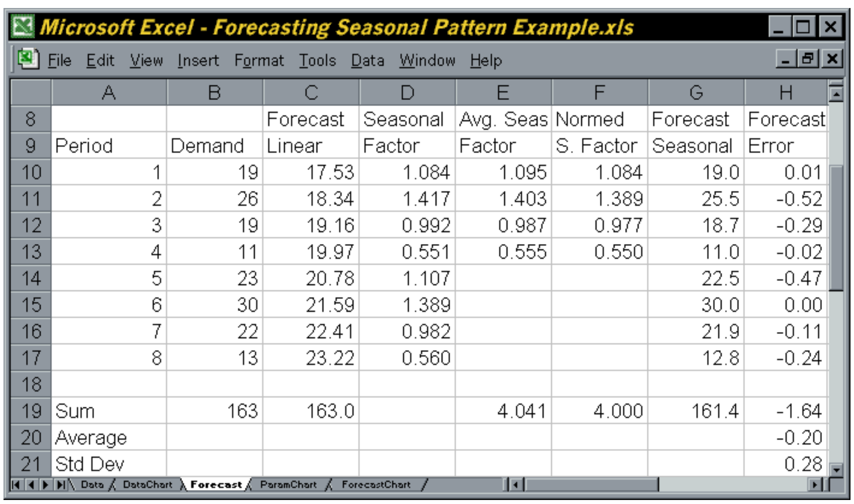
Figure 2.22. Initial Linear Forecast Parameters for the Seasonal Pattern Example

The next task is to compute the initial seasonality factors. There exists one factor for each of the time periods in the cycle. The factors are computed as the ratio of the observed value divided by the linear forecast. The computations for the calculation of seasonality factor of the fourth period in the cycle are shown next, including the normalization of the seasonality factors so that all factors in a cycle sum up to the number of periods in the cycle.

$$c_4 = \frac{11}{18.75 + 1.5 \cdot 0.81} = 0.551$$
$$c_8 = \frac{13}{23.22} = 0.560$$
$$\overline{c_4} = \frac{0.551 + 0.560}{2} = 0.555$$
$$\sum_{t=1}^{4} c_t = 4.041$$
$$c_4 = 0.555 \frac{4}{4.041} = 0.550$$

The computations for the seasonal forecast for period eight are shown next and the seasonal forecast for all periods and the corresponding errors are shown in the next figure.

$$F_8 = O_8 \cdot c_{8\|4} = 23.22 \cdot 0.550 = 12.8$$
$$e_8 = F_8 - D_8 = 12.8 - 13 = -0.2$$



	A	B	C	D	E	F	G	H
8			Forecast	Seasonal	Avg. Seas	Normed	Forecast	Forecast
9	Period	Demand	Linear	Factor	Factor	S. Factor	Seasonal	Error
10	1	19	17.53	1.084	1.095	1.084	19.0	0.01
11	2	26	18.34	1.417	1.403	1.389	25.5	-0.52
12	3	19	19.16	0.992	0.987	0.977	18.7	-0.29
13	4	11	19.97	0.551	0.555	0.550	11.0	-0.02
14	5	23	20.78	1.107			22.5	-0.47
15	6	30	21.59	1.389			30.0	0.00
16	7	22	22.41	0.982			21.9	-0.11
17	8	13	23.22	0.560			12.8	-0.24
18								
19	Sum	163	163.0		4.041	4.000	161.4	-1.64
20	Average							-0.20
21	Std Dev							0.28

Figure 2.23. Initial Forecast for the Seasonal Pattern Example

The forecast of the historical values using the linear trend and seasonality factors computed above is shown in the next figure. The forecasted values match the observed values very closely, which is numerically demonstrated by an average error of –0.20 or a relative error of less than one percent. We can use the average error in this case because all forecast errors basically have the same sign.

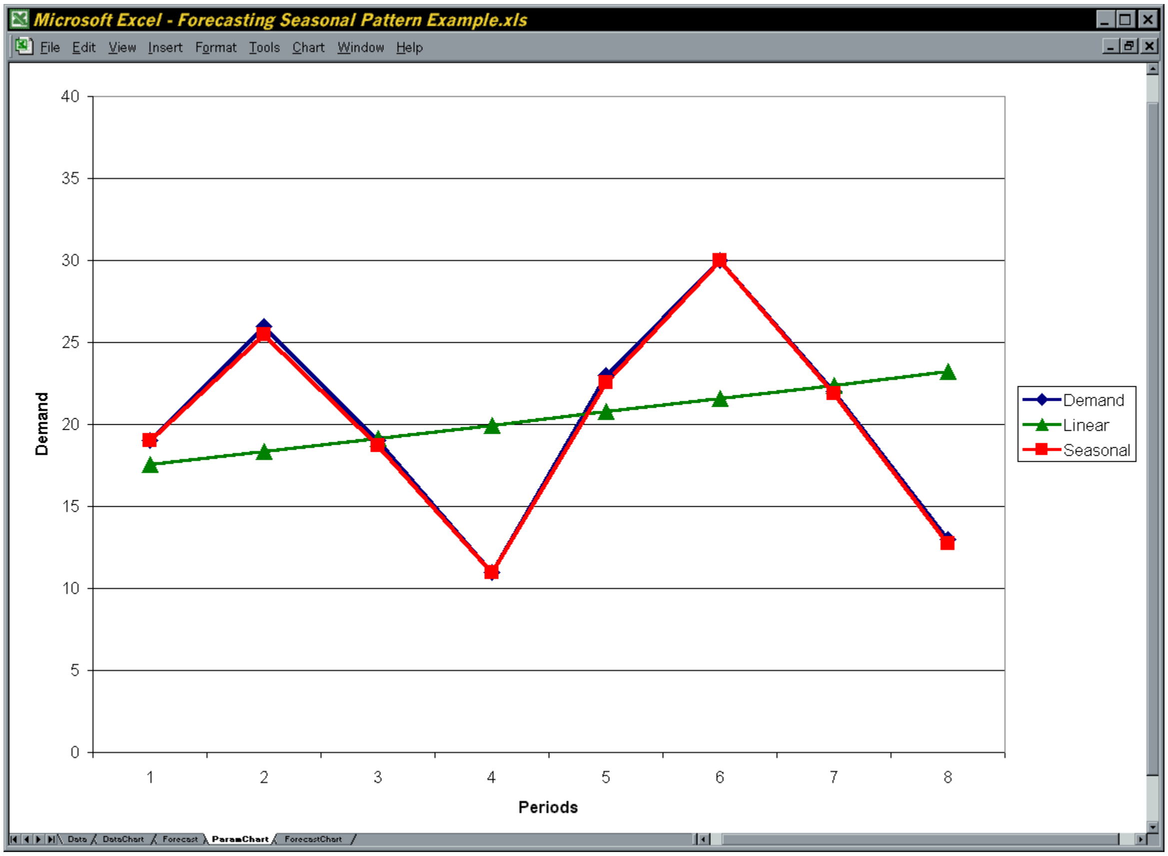


Figure 2.24. Graph of Initial Forecast for the Seasonal Pattern Example

The computations for the forecast and for the updates of the trend parameters and seasonality factor for period nine are shown next. Observe that only the seasonality factor associated with this period is updated directly based on the new observed value, but that the normalization of the seasonality factors may also changes the other factors.

$$F_9 = (O_8 + S_8) \cdot c_{9\|4} = (23.22 + 0.81) \cdot 1.084 = 26.1$$

$$O_9 = \alpha (D_9 / c_1) + (1 - \alpha)(O_8 + S_8)$$

$$= 0.2(27/1.084) + 0.8(23.22 + 0.81) = 24.21$$

$$S_9 = \beta (O_9 - O_8) + (1 - \beta) S_8$$

$$= 0.1(24.21 - 23.22) + 0.9 \cdot 0.81 = 0.83$$

$$c_1 = \gamma (D_9 / O_9) + (1 - \gamma) c_1$$

$$= 0.1(27/24.21) + 0.9 \cdot 1.084 = 1.087$$

$$\sum_{t=1}^{P} c_t = 4.003$$

$$c_1 = c_1 \frac{P}{\sum_{t=1}^{P} c_t} = 1.087 \frac{4}{4.003} = 1.086$$

Microsoft Excel - Forecasting Seasonal Pattern Example.xls

	A	B	C	D	E	F	G	H	I	J
23	Period	Demand	Demand				Average	Std.Dev	Abs. Rel	MAPE
24		Actual	Forecast	intercept	slope	error	error	error	error	
25	1	19	19.0							
26	2	26	25.5							
27	3	19	18.7							
28	4	11	11.0							
29	5	23	22.5							
30	6	30	30.0							
31	7	22	21.9							
32	8	13	12.8	23.22	0.81					
33	9	27	26.1	24.21	0.83	-0.9	-0.95	0.945	3.501	3.501
34	10	35	34.8	25.07	0.83	-0.2	-0.60	0.691	0.711	2.106
35	11	26	25.3	26.05	0.85	-0.7	-0.64	0.700	2.764	2.325
36	12	16	14.8	27.35	0.89	-1.2	-0.79	0.865	7.711	3.672
37	13		30.6							
38	14		40.4							
39	15		29.3							
40	16		17.1							

Figure 2.25. Forecast for the Seasonal Pattern Example

	K	L	M	N	O	P	Q	R	S	T	U
23	Period	Seasonality Factors					Normalized Seasonality Factors				
24		c1	c2	c3	c4	sum	c1	c2	c3	c4	sum
25	1										
26	2										
27	3										
28	4										
29	5										
30	6										
31	7										
32	8	1.095	1.403	0.987	0.555	4.041	1.084	1.389	0.977	0.550	4.000
33	9	1.087	1.389	0.977	0.550	4.003	1.086	1.388	0.976	0.549	4.000
34	10	1.086	1.389	0.976	0.549	4.001	1.086	1.389	0.976	0.549	4.000
35	11	1.086	1.389	0.978	0.549	4.002	1.086	1.388	0.978	0.549	4.000
36	12	1.086	1.388	0.978	0.553	4.004	1.085	1.387	0.977	0.552	4.000
37	13										
38	14										
39	15										
40	16										

Figure 2.26. Seasonality Factors for the Seasonal Pattern Example

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
23	Period	Demand	Demand				Average	Std.Dev	Abs. Rel	MAPE	Seasonality Factors					Normalized Seasonality Factors				
24		Actual	Forecast	intercept	slope	error	error	error	error		c1	c2	c3	c4	sum	c1	c2	c3	c4	sum
25	1	19	19.0																	
26	2	26	25.5																	
27	3	19	18.7																	
28	4	11	11.0																	
29	5	23	22.5																	
30	6	30	30.0																	
31	7	22	21.9																	
32	8	13	12.8	23.22	0.81						1.095	1.403	0.987	0.555	4.041	1.084	1.389	0.977	0.550	4.000
33	9	27	26.1	24.21	0.83	-0.9	-0.95	0.945	3.501	3.501	1.087	1.389	0.977	0.550	4.003	1.086	1.388	0.976	0.549	4.000
34	10	35	34.8	25.07	0.83	-0.2	-0.60	0.691	0.711	2.106	1.086	1.389	0.976	0.549	4.001	1.086	1.389	0.976	0.549	4.000
35	11	26	25.3	26.05	0.85	-0.7	-0.64	0.700	2.764	2.325	1.086	1.389	0.978	0.549	4.002	1.086	1.388	0.978	0.549	4.000
36	12	16	14.8	27.35	0.89	-1.2	-0.79	0.865	7.711	3.672	1.086	1.388	0.978	0.553	4.004	1.085	1.387	0.977	0.552	4.000
37	13		30.6																	
38	14		40.4																	
39	15		29.3																	
40	16		17.1																	

Figure 2.27. Forecast and Seasonality Factors for the Seasonal Pattern Example

The computations for the forecasts made at the end of period 12 for all the periods in the next cycle are shown next and forecasts for the next cycle are shown in the next figure.

$$F_{12,13} = \left(O_{12} + S_{12}\right) c_{13\|4} = \left(27.35 + 0.89\right) \cdot 1.085 = 30.6$$
$$F_{12,14} = \left(O_{12} + \tau S_{12}\right) c_{14\|4} = \left(27.35 + 2 \cdot 0.89\right) \cdot 1.387 = 40.4$$
$$F_{12,15} = \left(O_{12} + \tau S_{12}\right) c_{15\|4} = \left(27.35 + 3 \cdot 0.89\right) \cdot 0.977 = 29.3$$
$$F_{12,16} = \left(O_{12} + \tau S_{12}\right) c_{16\|4} = \left(27.35 + 4 \cdot 0.89\right) \cdot 0.552 = 17.1$$

Finally, the computations for the confidence intervals for the forecasts for the periods in the next cycle with a 95 % confidence level are shown next using the 0.865 as the standard deviation.

Table 2.3. Forecasts based on a 95 % Confidence Interval for the Seasonal Example

Period	Forecast Mean	Lower Bound	Upper Bound
13	30.6	28.9	32.3
14	40.4	38.7	42.1
15	29.3	27.6	31.0
16	17.1	15.4	18.8

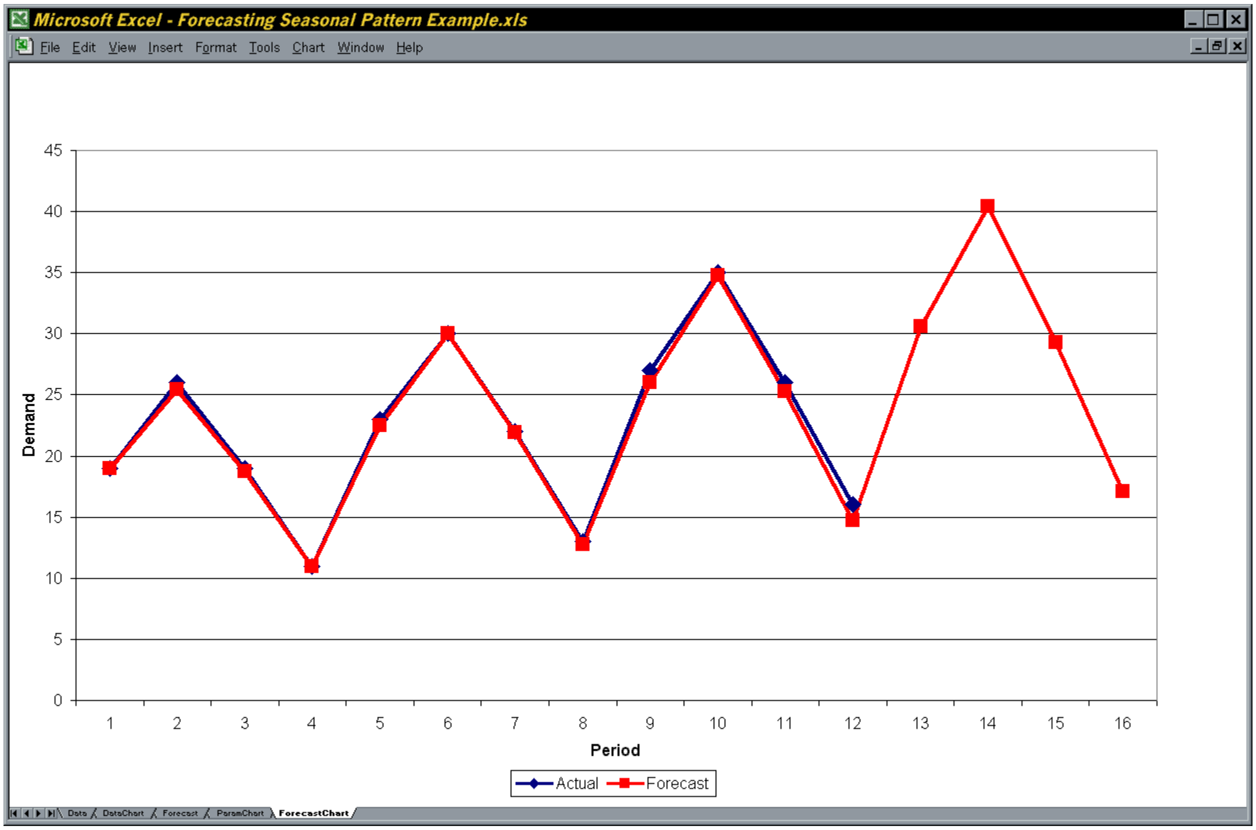


Figure 2.28. Graph of Forecast for the Seasonal Pattern Example

2.3. Summary

Forecasting of future conditions is an essential planning activity for almost any organization. Strategic, long-range forecasts are based on qualitative data or quantitative data with very large uncertainties. Long-range forecasts may be obtained with qualitative or very simple quantitative forecasting methods. Shorter-term tactical and operational forecasts are usually based on historical data. Many powerful and highly mathematical forecasting techniques exist, such as regression, time series analysis, and forecasting methods for autocorrelated data. However, the most important requirement for the generation of a high quality forecast is an understanding of the underlying data pattern of the forecasted variable. The correct forecasting method can then be selected for this underlying pattern. Applying correctly the wrong forecasting method for a data pattern will only generated wrong forecasts. One simple method to understand the underlying pattern is to graph the historical values over time. Since the underlying data pattern is almost never fully understood or remains unchanged for extended periods of time, most forecasts will have significant errors. Hence, it does not make much sense to report only the expected value of the forecasted variable. Based on the root mean square forecast error, confidence intervals can be established, which provide much more meaningful forecast information. A simple method to judge the quality of the forecast is to graph the forecast errors over time.

The generation of a high-quality forecast requires the understanding of the underlying pattern, the application of the correct method for that pattern, and the reporting of the results with their stochastic characteristics. The forecasting results generated by the blind application of very sophisticated forecasting techniques should be highly suspect. The further in the future the forecasts are, the easier it is to make very large forecasting errors.

In most logistics planning, forecasting cannot be avoided. So logistics professionals should forecast as well as possible and realize that they probably did not do it well enough.

2.4. Exercises

True-False Questions

The confidence interval of a demand forecast is a critical datum when judging the quality of the short-term forecast of the future value of the demand, (T/F)_____(1).

A cyclic demand pattern has cycles of varying length and magnitude, (T/F)_____(2).

Time series forecasting attempts to predict the future value of a variable based on the currently observed values of different variables, (T/F)_____(3).

A forecast created by the moving averages method lags any trend the forecasted variable may have, (T/F)_____(4).

In exponential smoothing the weights of the different historical values form a geometrically decreasing pattern when the values go farther back in time, (T/F)_____(5).

In exponential smoothing a larger smoothing constant will create a forecast that more rapidly tracks a current trend in the forecasted variable, (T/F)_____(6).

ARIMA models are used for forecasting autocorrelated data, (T/F)_____(7).

Using forecasting techniques based on time series analysis, different forecasters will arrive at the same forecast if they use the same data, (T/F) _____(8).

Using the moving averages forecasting method, only the last observation or realized demand needs to be kept, (T/F) _____(9).

A higher value of the alpha parameter in exponential smoothing forecasting increases the stability of the forecast, (T/F) _____(10).

The sum of the normalized seasonality factors in the multiplicative variant of Winter's exponential smoothing method must add up to the number of periods in the cycle, (T/F) _____(11).

Holt's exponential smoothing method requires three smoothing constants, (T/F) _____(12).

At least two full cycles of data are required to use the multiplicative variant of Winter's exponential smoothing method, (T/F) _____(13).

When the data are autocorrelated, they are assumed to be independent samples of a probability distribution, (T/F) _____(14).

A professionally prepared forecast will include the expected value and a confidence interval for the forecast, (T/F) _____(15).

A forecast of the weekly demand for products in a grocery store tends to be more accurate, as measured by the coefficient of variation, than the forecast for the daily demand for the same products, (T/F)______(16).

A forecast for the sales in a home improvement store of batteries of a single manufacturer tends to be more accurate, as measured by the coefficient of variation, than the forecast for the sales of batteries of all manufacturers combined, (T/F)______(17).

Given that all other parameters remain the same, a larger mean squared forecast error corresponds to a larger confidence interval, (T/F)______(18).

The moving averages forecasting technique is the appropriate forecasting method when the underlying data exhibits a linearly growing trend, (T/F)______(19).

PPC Triple Exponential Smoothing

This question has the additional purpose to familiarize you with the structured answers you would need to provide to a similar question during an examination.

The Polychromatic Paper Company (PPC) produces high quality, glossy, wrapping paper for gifts. They sell primarily to discount merchandisers and drug store chains. They have collected quarterly data on the sales of their top of the line paper grade, expressed in millions of rolls. The data for the last two years are given in Table 2.4. You are asked to provide a sales forecast for the same product for the next year, periods 9 through 12, with an 80 % confidence interval. Use two significant digits after the decimal point in all your calculations.

Table 2.4. Historical Sales Data

Period	Demand
1	8
2	12
3	15
4	32
5	6
6	8
7	9
8	22

Identify the underlying data pattern and write down the algebraic expression for it. Provide a concise and complete definition of the variables that are used in the algebraic expression.

Use the appropriate exponential smoothing method for this underlying data pattern to forecast the demand for the next year. Use the standard values for the smoothing constant(s). Compute the forecast offset (or intercept), the slope, the linear forecast, the seasonality factors, and the seasonal forecast based on the data. Summarize your calculations in the following table. Not all the rows, columns, or cells of the table have to be used. Show the computations to derive the intercept and slope at end of period eight.

Period	Demand	Linear Forecast			Seasonality Factor	Seasonal Forecast	Forecast Error
1	8						
2	12						
3	15						
4	32						
5	6						
6	8						
7	9						
8	22						

Compute the forecasts for the next year. For each period you must show the linear forecast, the seasonality factor used, and the expected value, lower and upper bound of the confidence interval of the seasonal forecast. Summarize your calculations in the next table. Compute and show the standard deviation of the forecast error, the number of standard deviations corresponding to the confidence limits, and the size of the confidence interval.

Period	Linear Forecast	Seasonality Factor	Seasonal Forecast	Lower Conf. Int.	Upper Conf. Int.
9					
10					
11					
12					

At then end of period 9 the company contacts you again with the new sales data for this period and asks you to update your forecast for the next four periods (periods 10, 11, 12, and 13). The actual sales for period 9 were 4 million rolls. Compute the updated values for the slope, intercept, seasonality factors, and standard deviation of the forecast error. Give the formulas for the computation of the update intercept, slope, and seasonality factors. Compute the updated values for the intercept, slope, and seasonality factors, standard deviation of the forecast error, and the size of the confidence interval. Finally, show your computations for the update forecasts in the next table.

Period	Linear Forecast	Seasonality Factor	Seasonal Forecast	Lower Conf. Int.	Upper Conf. Int.
10					
11					
12					
13					

Assume that the company reported to you that the sales for period 9 were 7 million rolls. In other words, the actual sales of 4 million rolls never happened and the actual sales were 7 million rolls. What would be your forecasting report to PPC state in this case?

High Voltage Electrical Consumption

The Industrial Metals Foundry is in the process of budgeting for their energy costs during the next year. Their energy costs are primarily related to their consumption of high-voltage electricity for their electric arc smelters. They have recorded their energy costs for the past three years on a quarterly basis. They have asked you to provide them with a forecast of their energy costs for the next year with a 90 % confidence level. Clearly explain your general approach and state the assumptions that you have used to arrive at this forecast. Use the answer structure provided in the previous exercise to provide the details of your response to their request. The data is shown in the next table.

Period	Demand
1	27000
2	70000
3	41000
4	13000
5	30000
6	73000
7	48000
8	15000
9	34000
10	82000
11	51000
12	16000

2.5. References

1. Ballou, R. H., (1998), Chapter 9 in **Business Logistics Management**, 4th Edition, Prentice Hall, Englewood Cliffs, New Jersey.
2. Box, G. E. and G. M. Jenkins, (1970), **Time Series Analysis, Forecasting and Control**, Holden Day, San Francisco
3. Brockwell, P. and R. Davis, (1996), **Introduction to Time Series and Forecasting**, Springer-Verlag.
4. Gross, C. W. and R. T. Peterson, (1983), **Business Forecasting**, 2nd Edition, John Wiley and Sons, New York

5. Hayter, A., (1996). **Probability and Statistics for Engineers and Scientists**. PWS Publishing Company, Boston, Massachusetts.

6. Holt, C. C., (1957), "Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages", Office of Naval Research Memo, No. 52.

7. Kutner, M., C. Nachtschiem, W. Wasserman, and J. Neter, (1996), **Applied Linear Statistical Models**, McGraw-Hill, Boston, Massachusetts.

8. Montgomery, D. C. and L. A. Johnson, (1976), **Forecasting and Time Series Analysis**, McGraw-Hill, New York

9. Montgomery, D., E. Peck, and G. Vining, (2001), **Introduction to Linear Regression Analysis**, John Wiley & Sons, New York, New York.

10. Nahmias, S., (2000), Chapter 2 in **Production and Operations Analysis**, 4th Edition, Irwin/McGraw-Hill, Boston, Massachusetts.

11. Wilson, J. H. and B. Keating, (1990), **Business Forecasting**, Richard D. Irwin, Homewood, IL.

12. Winters, P. R., (1960), "Forecasting by Exponentially Moving Averages", ***Management Science***, Vol. 6, pp. 324-342.