Dr. Saeed A. Dobbah Alghamdi

Department of Statistics
Faculty of Sciences
Building 90, Office 26F41
King Abdulaziz University
http://saalghamdy.kau.edu.sa

Dr. Saeed's Website

Statistics
for
Business & Economics

David R. Anderson
Dennis J. Sweeney
Thomas A. Williams

Time Series Analysis and Forecasting

Chapter 14

Learning Objectives

LO1 Define the components of a *time series*

LO2 Define forecast and forecast accuracy

LO3 Compute *moving average*

LO4 Determine *trend equations*

LO5 Determine and interpret a set of *seasonal indexes*

LO6 Deseasonalize data using a seasonal index

Introduction

- The purpose of this chapter is to provide an introduction to time series analysis and forecasting.
- Quantitative forecasting methods can be used when
  - past information about the variable being forecast is available
  - the information can be quantified
  - it is reasonable to assume that the pattern of the past will continue into the future.
- The objective of time series analysis is to
  - discover a pattern in the historical data or time series
  - extrapolate the pattern into the future (forecasting)
- The forecast is based solely on past values of the variable and/or on past forecast errors.
- 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.

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 pattern of the data is an important factor in understanding how the time series has behaved in the past. If such behavior can be expected to continue in the future, we can use the past pattern to guide us in selecting an appropriate forecasting method.
- To identify the underlying pattern in the data, a useful first step is to construct a **time series plot**.
- A **time series plot** is a graphical presentation of the relationship between time and the time series variable; time is on the horizontal axis and the time series values are shown on the vertical axis.

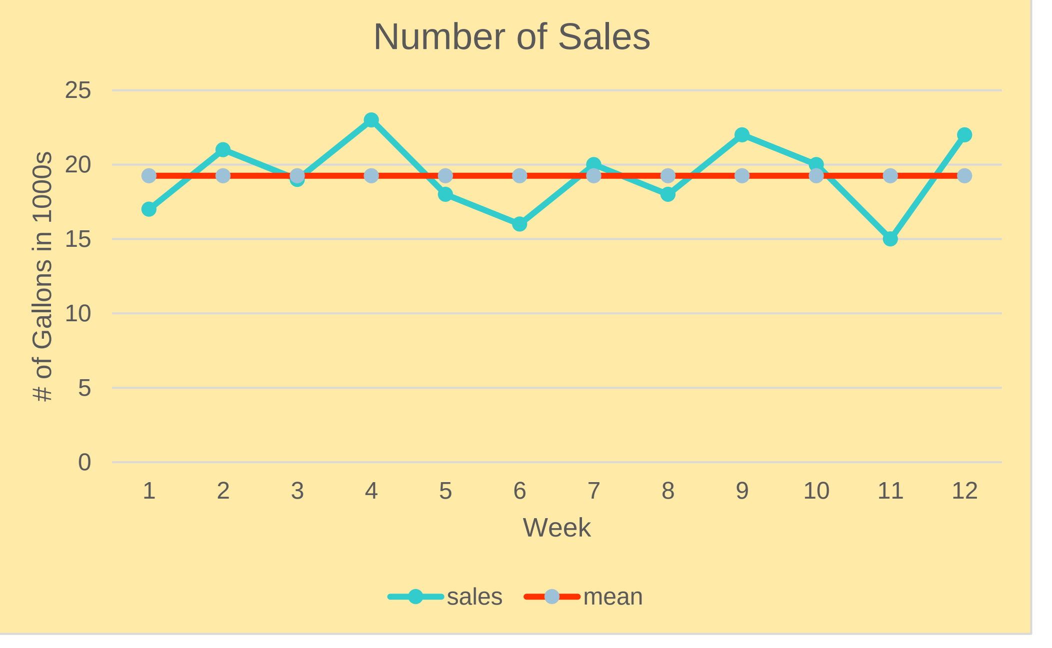
Horizontal Pattern

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

Example

The number of gasoline sold in 1000s of gallons weekly over the pas 12 weeks in a gas station. The average value or mean for this time series is 19.25 gallons per week. Note how the data fluctuate around the sample mean. Although random variability is present, we would say that theses data follow a horizontal pattern.

Week	Sales
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22



Horizontal Pattern

- 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**. But simply observing a horizontal pattern is not sufficient evidence to conclude that the time series is stationary. More advanced texts on forecasting discuss procedures for determining if a time series is stationary and provide methods for transforming a time series that is not stationary into a stationary series.
- Changes in business conditions can often result in a time series that has a horizontal pattern shifting to a new level.

Trend Pattern

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

Example

The time series of bicycle sales in 1000s for a particular manufacturer over the past 10 years, as shown in the table and the figure.

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

Visual inspection of the time series plot shows some up and down movement which seems to be systematically increasing or upward trend. The trend appears to be linear and increasing over time, but sometimes a trend can be described better by other types of patterns.

Seasonal Pattern

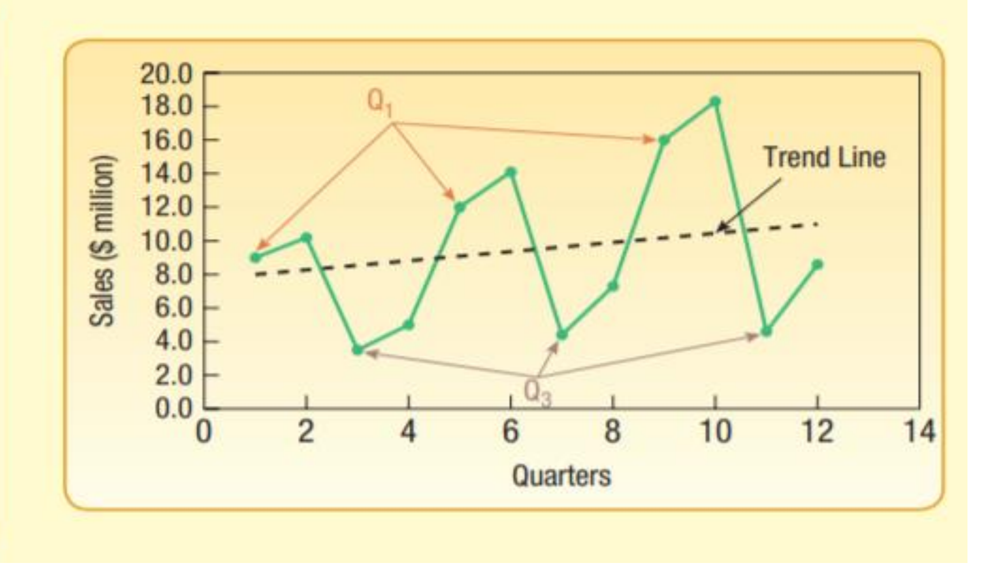
- The trend of a time series can be identified by analyzing multiyear movements in historical data. **Seasonal patterns** are recognized by seeing the same repeating patterns within one year over successive periods of time. For instance, a manufacturer of swimming pools expects low sales activity in the fall and winter months, with peak sales in the spring and summer months.

Example

The sales in millions of a clothing store over the past three years are reported quarterly as shown in the table and the figure.

Note: the time series include a combination of increasing trend and seasonal pattern.

Year	Quarter	Code	Sales
2014	Q1	1	9
	Q2	2	10.2
	Q3	3	3.5
	Q4	4	5
2015	Q1	5	12
	Q2	6	14.1
	Q3	7	4.4
	Q4	8	7.3
2016	Q1	9	16
	Q2	10	18.3
	Q3	11	4.6
	Q4	12	8.6

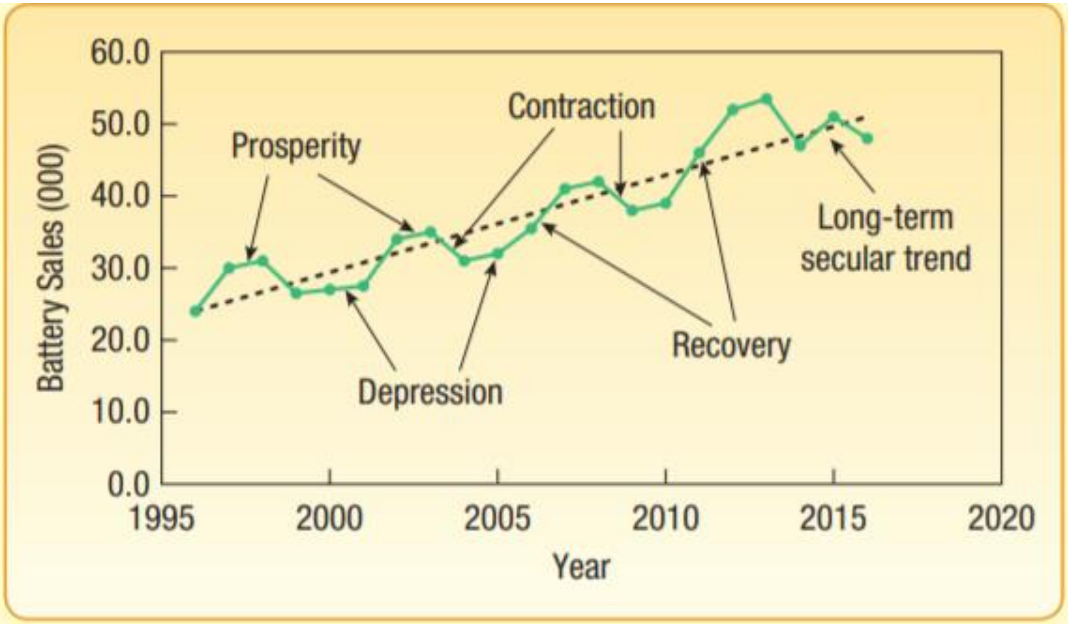


Cyclical Pattern

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

Example

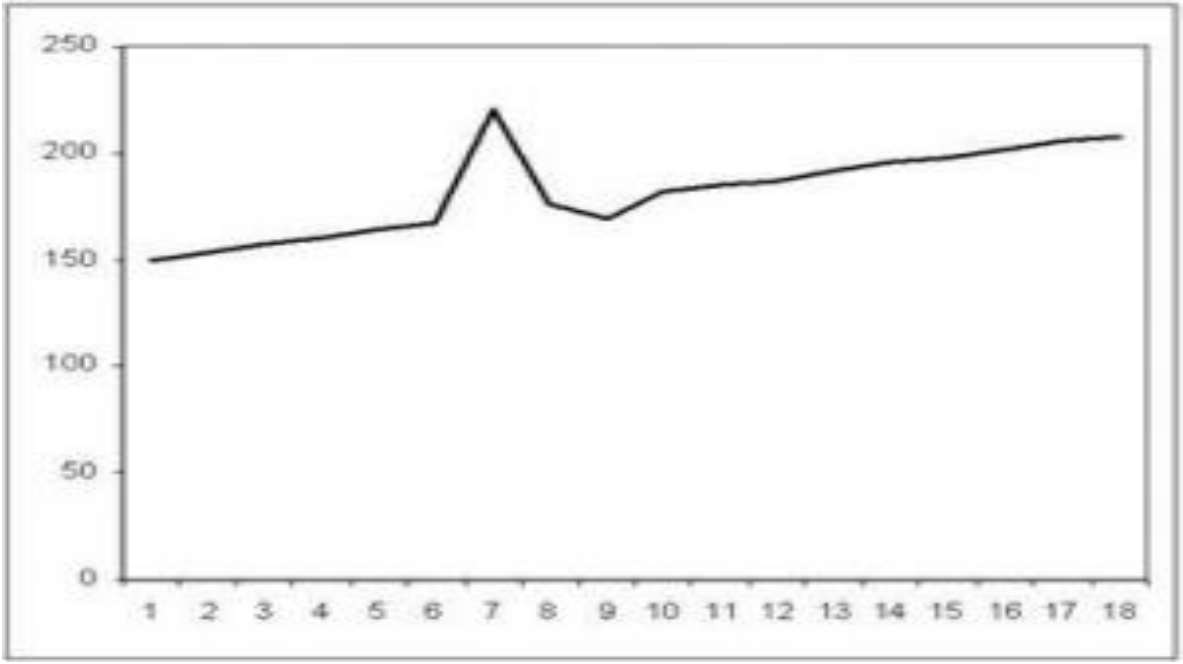
Year	Battery Sales (000)	Year	Battery Sales (000)
1996	24.0	2007	41.0
1997	30.0	2008	42.0
1998	31.0	2009	38.0
1999	26.5	2010	39.0
2000	27.0	2011	46.0
2001	27.5	2012	52.0
2002	34.0	2013	53.5
2003	35.0	2014	47.0
2004	31.0	2015	51.0
2005	32.0	2016	48.0
2006	35.5		



Irregular Pattern

- An **irregular pattern** exists when unexplained changes of the time series over time and don't relate to the effect of the other patterns.

Example



Forecast

- **Forecast** is the process of making predictions based on past and present data and most commonly by analysis of patterns or trends.
- The underlying pattern in the time series is an important factor in selecting a forecasting method. Thus, a time series plot should be one of the first things developed when trying to determine what forecasting method to use. If we see a horizontal pattern, then we need to select a method appropriate for this type of pattern. Similarly, if we observe a trend in the data, then we need to use a forecasting method that has the capability to handle trend effectively.
- We will illustrate methods that can be used in situations where the underlying pattern is horizontal; in other words, no trend or seasonal effects are present. We then consider methods appropriate when trend and/or seasonality are present in the data.

Forecast Accuracy

- There are several measures of **forecast accuracy** which measures how well a particular forecasting method is able to reproduce the time series data that already available.
- By selecting the method that has the best accuracy for the data already known, we hope to increase the likelihood that we will obtain better forecasts for future time periods.
- The key concept associated with measuring forecast accuracy is **forecast error**, defined as

Forecast Error = Actual Value – Forecast

Absolute Forecast Error = |Actual Value – Forecast|

Squared Forecast Error = (Actual Value – Forecast)^2

Percentage Error = (Forecast Error/Actual Value)*100

Absolute Percentage Error = |(Forecast Error/Actual Value)*100|

Forecast Accuracy

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

MAE = mean of the absolute value of forecast errors

MSE = mean of the squared forecast errors

MAPE = mean of the absolute value of percentage forecast errors

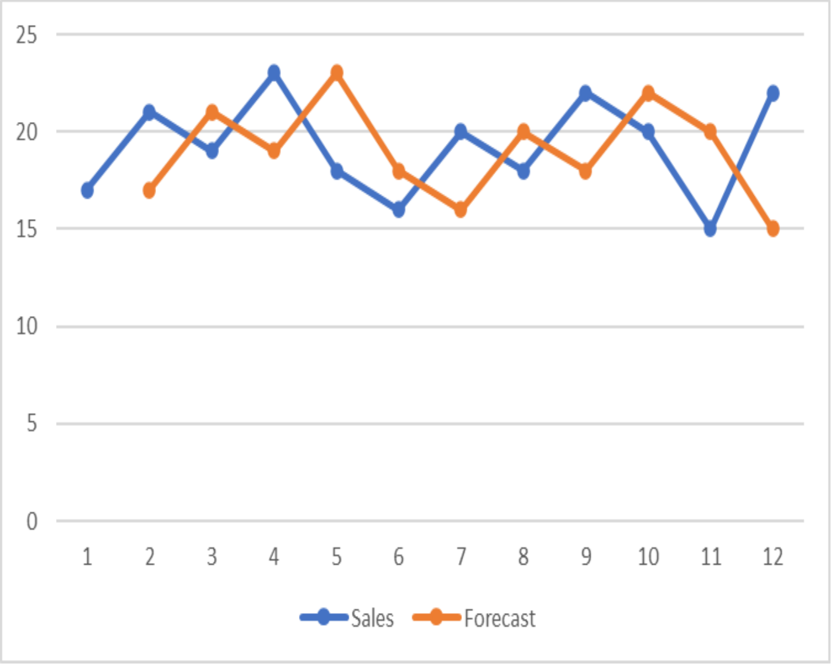
Naive Forecasting Method

- A **naïve forecasting method** uses the most recent value as the forecast for the next period.

Example

The number of gasoline sold weekly over the past 12 weeks in a gas station.

Week	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	17				
2	21	17	4	16	19.05
3	19	21	2	4	10.53
4	23	19	4	16	17.39
5	18	23	5	25	27.78
6	16	18	2	4	12.50
7	20	16	4	16	20.00
8	18	20	2	4	11.11
9	22	18	4	16	18.18
10	20	22	2	4	10.00
11	15	20	5	25	33.33
12	22	15	7	49	31.82
		mean	3.73	16.27	19.24



MAE = mean of the absolute value of forecast errors = 3.73

MSE = mean of the squared forecast errors = 16.27

MAPE = mean of the absolute value of percentage forecast errors = 19.24%

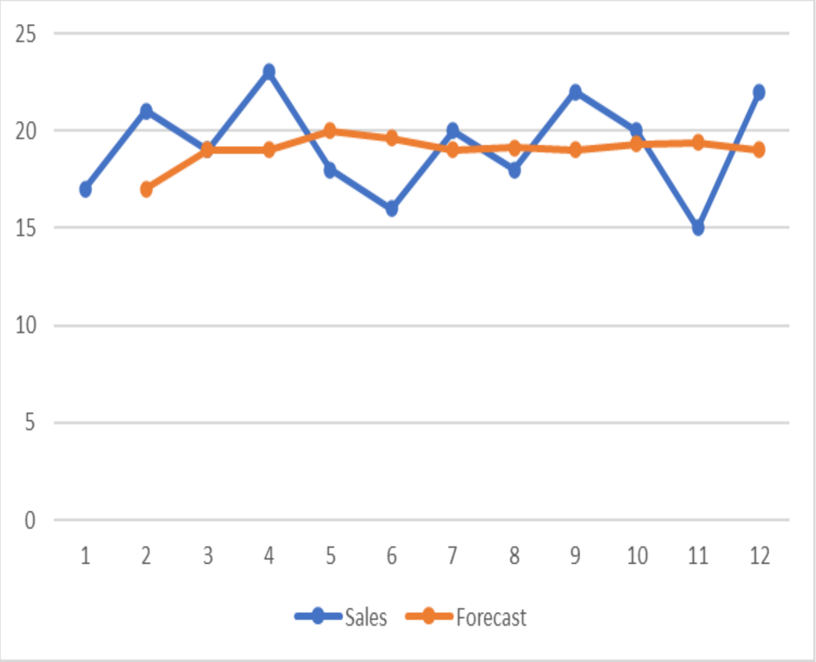
Average of Past Values Forecasting Method

- An **average of past values forecasting method** uses the average of all the historical data as the forecast for the next period.

Example

The number of gasoline sold weekly over the past 12 weeks in a gas station.

Week	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	17				
2	21	17.00	4.00	16.00	19.05
3	19	19.00	0.00	0.00	0.00
4	23	19.00	4.00	16.00	17.39
5	18	20.00	2.00	4.00	11.11
6	16	19.60	3.60	12.96	22.50
7	20	19.00	1.00	1.00	5.00
8	18	19.14	1.14	1.31	6.35
9	22	19.00	3.00	9.00	13.64
10	20	19.33	0.67	0.44	3.33
11	15	19.40	4.40	19.36	29.33
12	22	19.00	3.00	9.00	13.64
		mean	2.44	8.10	12.85



MAE = mean of the absolute value of forecast errors = 2.44

MSE = mean of the squared forecast errors = 8.10

MAPE = mean of the absolute value of percentage forecast errors = 12.85%

Moving Average Method

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

MOVING AVERAGE FORECAST OF ORDER *k*

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

where

F_{t+1} = forecast of the times series for period $t + 1$

Y_t = actual value of the time series in period t

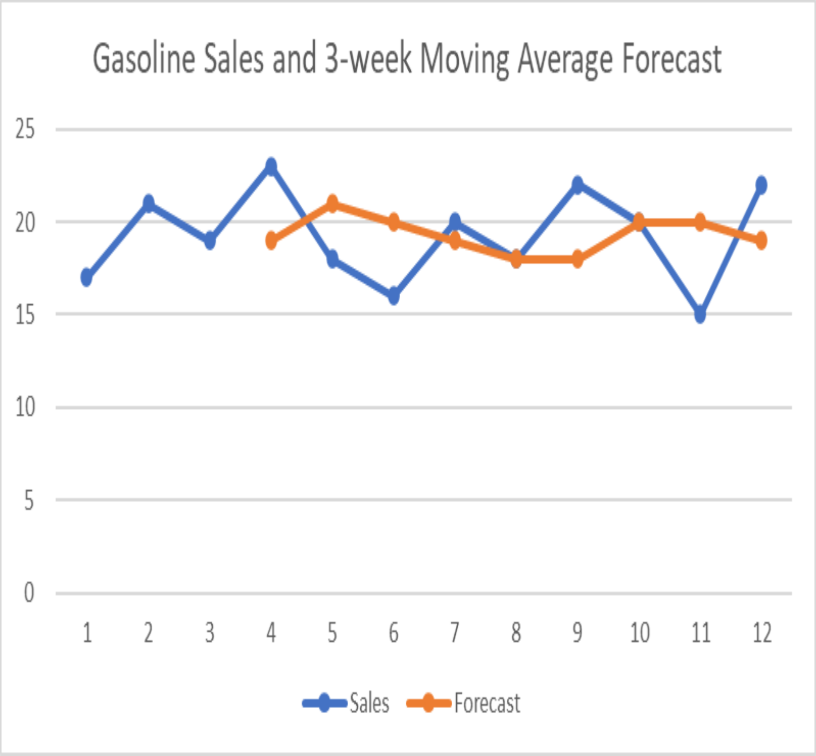
- The term *moving* is used because every time a new observation becomes available for the time series, it replaces the oldest observation in the equation and a new average is computed. As a result, the average will change, or move, as new observations become available.
- To use moving averages to forecast a time series, we must first select the order, or number of time series values, to be included in the moving average. If only the most recent values of the time series are considered relevant, a small value of ***k*** is preferred. If more past values are considered relevant, then a larger value of ***k*** is better.

Moving Average Method

Example

The number of gasoline sold weekly over the past 12 weeks in a gas station. A three weeks moving average calculations are as the following:

Week	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	17				
2	21				
3	19				
4	23	19.00	4.00	16.00	17.39
5	18	21.00	3.00	9.00	16.67
6	16	20.00	4.00	16.00	25.00
7	20	19.00	1.00	1.00	5.00
8	18	18.00	0.00	0.00	0.00
9	22	18.00	4.00	16.00	18.18
10	20	20.00	0.00	0.00	0.00
11	15	20.00	5.00	25.00	33.33
12	22	19.00	3.00	9.00	13.64
		mean	2.67	10.22	14.36



MAE = mean of the absolute value of forecast errors = 2.67

MSE = mean of the squared forecast errors =10.22

MAPE = mean of the absolute value of percentage forecast errors = 14.36%

Weighted Moving Average Method

- In the **moving averages** method, each observation in the moving average calculation receives the same weight. The **weighted moving averages** method involves selecting a different weight for each data value and then computing a weighted average of the most recent k values as the forecast. In most cases, the most recent observation receives the most weight, and the weight decreases for older data values.
- To use the weighted moving averages method, we must first select the number of data values to be included in the weighted moving average and then choose weights for each of the data values. In general, if we believe that the recent past is a better predictor of the future than the distant past, larger weights should be given to the more recent observations. However, when the time series is highly variable, selecting approximately equal weights for the data values may be best. The only requirement in selecting the weights is that their sum must equal 1.

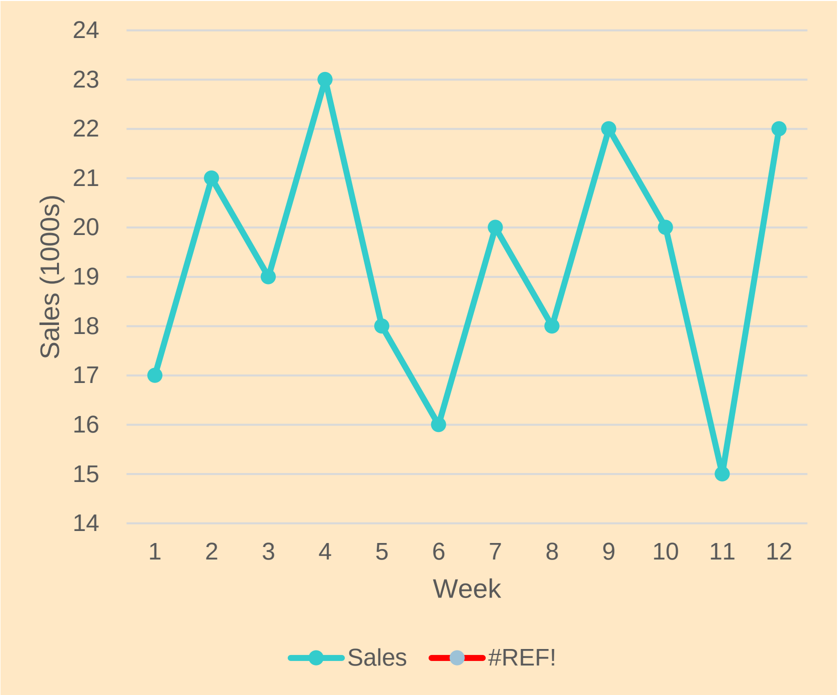
Weighted Moving Average Method

Example

The number of gasoline sold weekly over the past 12 weeks in a gas station.

A three weeks weighted moving average with weights 0.2,0.3 and 0.5 respectively are as the following:

Week	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	17				
2	21				
3	19				
4	23	19.20	3.80	14.44	16.52
5	18	21.40	3.40	11.56	18.89
6	16	19.70	3.70	13.69	23.13
7	20	18.00	2.00	4.00	10.00
8	18	18.40	0.40	0.16	2.22
9	22	18.20	3.80	14.44	17.27
10	20	20.40	0.40	0.16	2.00
11	15	20.20	5.20	27.04	34.67
12	22	17.90	4.10	16.81	18.64
		mean	2.98	11.37	15.93



MAE = mean of the absolute value of forecast errors = 2.98

MSE = mean of the squared forecast errors = 11.37

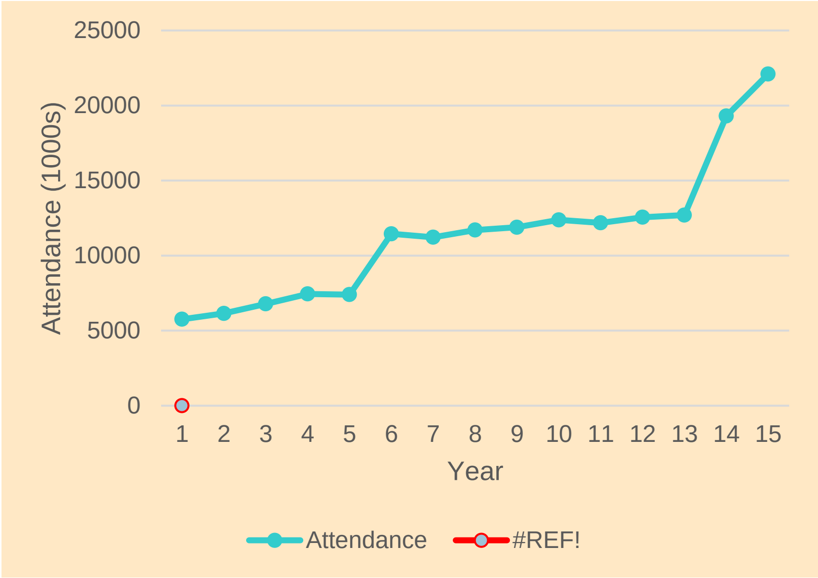
MAPE = mean of the absolute value of percentage forecast errors = 15.93%

Weighted Moving Average Method

Example

Cedar Fair operates several amusement parks. Its combined attendance (in 000s) for the last 12 years as given in the table. Compute a three-year weighted moving average with weights of 0.2, 0.3, and 0.5 for successive years.

Year	Attendance	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1993	5761				
1994	6148				
1995	6783				
1996	7445	6388.10	1056.90	1117037.61	14.20
1997	7405	6987.00	418.00	174724.00	5.64
1998	11450	7292.60	4157.40	17283974.76	36.31
1999	11224	9435.50	1788.50	3198732.25	15.93
2000	11703	10528.00	1175.00	1380625.00	10.04
2001	11890	11508.70	381.30	145389.69	3.21
2002	12380	11700.70	679.30	461448.49	5.49
2003	12181	12097.60	83.40	6955.56	0.68
2004	12557	12182.50	374.50	140250.25	2.98
2005	12700	12408.80	291.20	84797.44	2.29
2006	19300	12553.30	6746.70	45517960.89	34.96
2007	22100	15971.40	6128.60	37559737.96	27.73
		mean	1940.07	8922636.16	13.29



MAE = 1940.07 , MSE = 8922636.16 and MAPE = 13.29%

Exponential Smoothing

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

EXPONENTIAL SMOOTHING FORECAST

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

where

F_{t+1} = forecast of the time series for period $t + 1$

Y_t = actual value of the time series in period t

F_t = forecast of the time series for period t

α = smoothing constant $(0 \leq \alpha \leq 1)$

Exponential Smoothing

Example

The number of gasoline sold weekly over the past 12 weeks in a gas station.

The exponential smoothing with smoothing constant 0.2 and initial value equal to the actual value of period one is as the following:

Week	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	17				
2	21	17.0	4.00	16.00	19.05
3	19	17.8	1.20	1.44	6.32
4	23	18.0	4.96	24.60	21.57
5	18	19.0	1.03	1.07	5.73
6	16	18.8	2.83	7.98	17.66
7	20	18.3	1.74	3.03	8.70
8	18	18.6	0.61	0.37	3.38
9	22	18.5	3.51	12.34	15.97
10	20	19.2	0.81	0.66	4.05
11	15	19.4	4.35	18.94	29.01
12	22	18.5	3.52	12.38	15.99
		mean	2.60	8.98	13.40

MAE = 2.6 , MSE = 8.98 and MAPE = 13.4%

Linear Trend Regression

- Simple linear regression can be used to forecast a time series with a linear trend.
- The estimated regression equation describing a straight-line relationship between an independent variable *x* and a dependent variable *y* is written as $\hat{y} = b_0 + b_1 x$
- To emphasize the fact that in forecasting the independent variable is time, we will replace *x* with *t* and $\hat{y}$ with T_t to clarify that we are estimating the trend for a time series. Thus, for estimating the linear trend in a time series we will use the following estimated regression equation.

LINEAR TREND EQUATION

$$T_t = b_0 + b_1 t$$

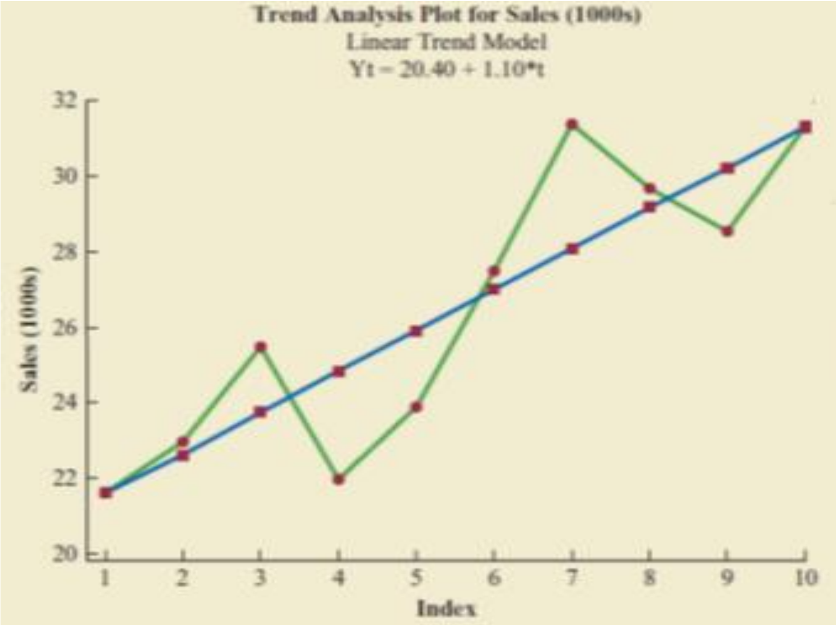
where

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

Linear Trend Regression

Example

The time series of bicycle sales for a particular manufacturer over the past 10 years illustrates a time series with a trend pattern. Use MegaStat (Time Series / Forecasting, Trendline Curvefit, Linear) to calculate Predicted (Forecast).



Year	Sales	Forecast	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	21.6	21.50	0.10	0.01	0.46
2	22.9	22.60	0.30	0.09	1.31
3	25.5	23.70	1.80	3.24	7.06
4	21.9	24.80	2.90	8.41	13.24
5	23.9	25.90	2.00	4.00	8.37
6	27.5	27.00	0.50	0.25	1.82
7	31.5	28.10	3.40	11.56	10.79
8	29.7	29.20	0.50	0.25	1.68
9	28.6	30.30	1.70	2.89	5.94
10	31.4	31.40	0.00	0.00	0.00
		mean	1.32	3.07	5.07

MAE = mean of the absolute value of forecast errors = 1.32

MSE = mean of the squared forecast errors = 3.07

MAPE = mean of the absolute value of percentage forecast errors = 5.07%

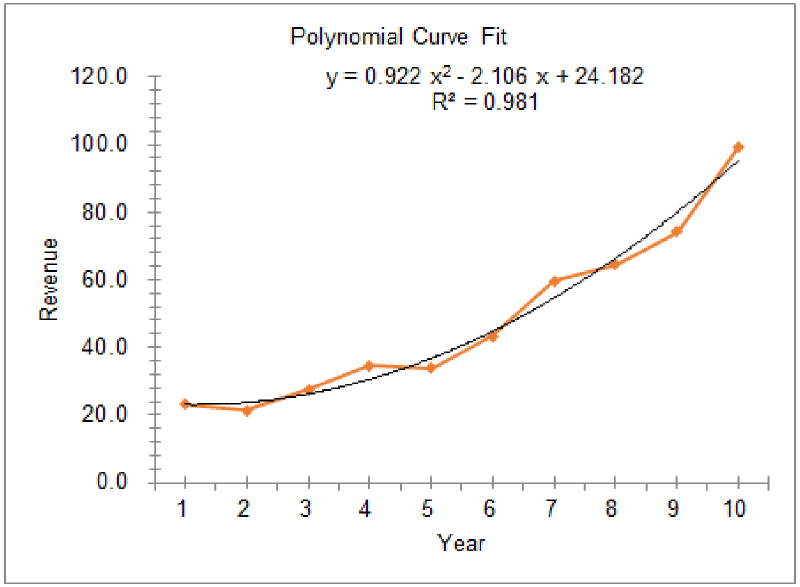
Nonlinear Trend Regression

- The use of a linear function to model trend is common. However, sometimes time series have a curvilinear or nonlinear trend.
- A variety of nonlinear functions can be used to develop an estimate of the trend, For instance, consider the following quadratic trend equation:

$$T_t = b_0 + b_1 t + b_2 t^2$$

Example: Consider the annual revenue in millions of dollars for a cholesterol drug for 10 years of sales. Use MegaStat (Time Series / Forecasting, Trendline Curvefit, 2nd degree polynomial) to calculate Predicted (Forecast).

Year	Revenue	Predicted	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	23.1	23.00	0.10	0.01	0.44
2	21.3	23.66	2.36	5.55	11.06
3	27.4	26.16	1.24	1.54	4.53
4	34.6	30.50	4.10	16.78	11.84
5	33.8	36.69	2.89	8.36	8.55
6	43.2	44.72	1.52	2.32	3.53
7	59.5	54.60	4.90	24.03	8.24
8	64.4	66.32	1.92	3.67	2.97
9	74.2	79.88	5.68	32.22	7.65
10	99.3	95.28	4.02	16.15	4.05
		Mean	2.87	11.06	6.29



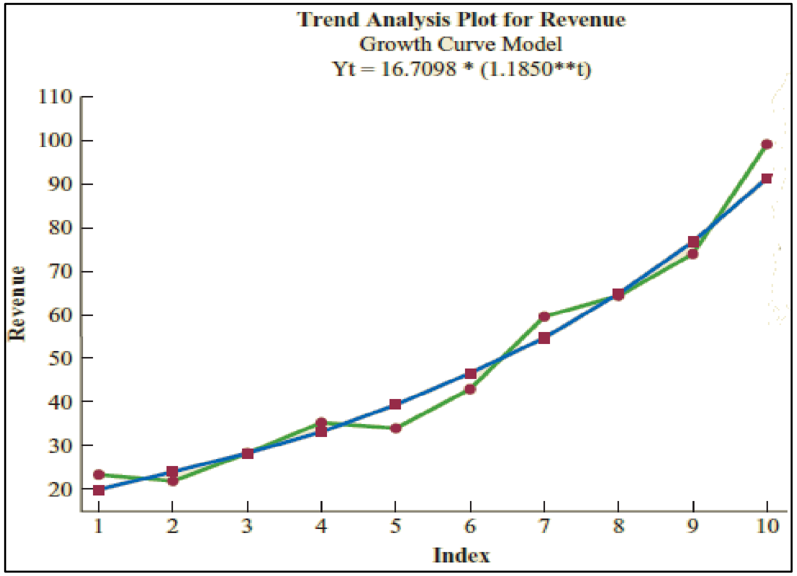
MAE = 2.87, MSE = 11.07 and MAPE = 6.29%

Nonlinear Trend Regression

- Another alternative that can be used to model a nonlinear pattern is to fit an exponential model to the data. For instance, consider the following exponential trend equation: $T_t = b_0(b_1)^t$

Example, consider the annual revenue in millions of dollars for a cholesterol drug for 10 years of sales. Use MegaStat (Time Series / Forecasting, Trendline Curvefit, Exponential (Ln)) to calculate Predicted (Forecast).

Year	Revenue	Predicted	Absolute Forecast Error (AFE)	Squared Forecast Error (SFE)	Absolute Percentage Error (APE)
1	23.1	19.80	3.30	10.88	14.28
2	21.3	23.46	2.16	4.68	10.16
3	27.4	27.80	0.40	0.16	1.48
4	34.6	32.95	1.65	2.73	4.78
5	33.8	39.04	5.24	27.49	15.51
6	43.2	46.27	3.07	9.40	7.10
7	59.5	54.82	4.68	21.87	7.86
8	64.4	64.97	0.57	0.32	0.88
9	74.2	76.98	2.78	7.75	3.75
10	99.3	91.22	8.08	65.22	8.13
		Mean	3.19	15.05	7.39



MAE = 3.19, MSE = 15.05 and MAPE = 7.39%

Seasonality Without Trend

- To the extent that seasonality exists, we need to incorporate it into our forecasting models to ensure accurate forecasts.
- Dummy variables can be used to deal with categorical independent variables in a multiple regression model. We can use the same approach to model a time series with a seasonal pattern by treating the season as a categorical variable. Recall that when a categorical variable has *k* levels, *k* - 1 dummy variables are required.
- As an example, consider the number of umbrellas sold at a clothing store over the past five years. to model the seasonal effects in the umbrella time series we need 4 – 1 = 3 dummy variables. The three dummy variables can be coded as follows:

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

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

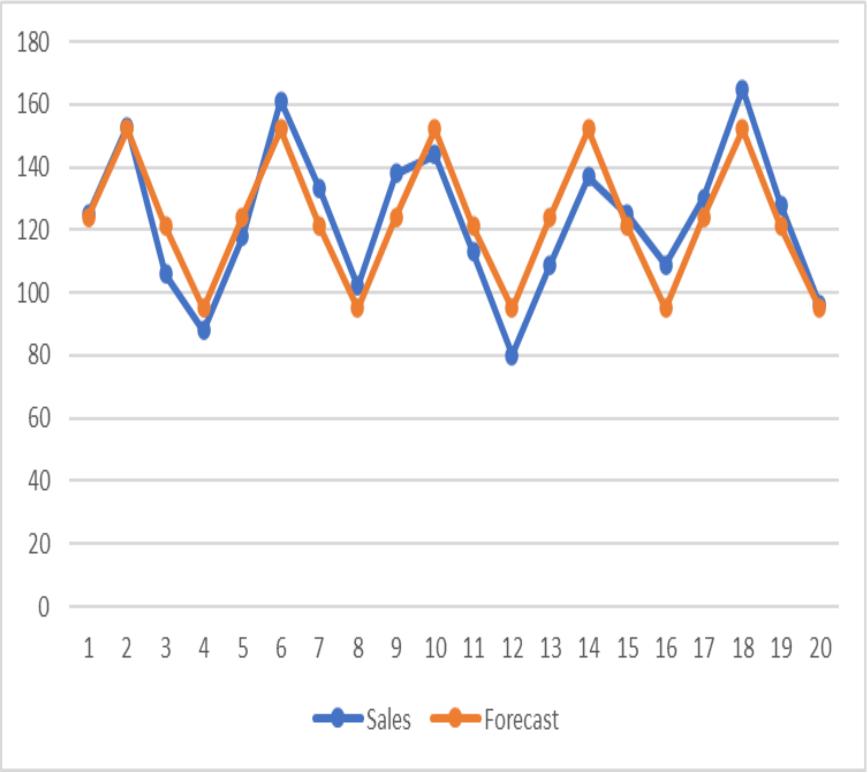
$$\hat{Y} = b_0 + b_1 \text{ Qtr1} + b_2 \text{ Qtr2} + b_3 \text{ Qtr3}$$

Seasonality Without Trend

Umbrella example

Use MegaStat (Correlation / Regression, Regression, X (q1, q2, q3), Y (Sales)) to get the coefficient of the equation. $\text{Sales} = 95.0 + 29.0\ \text{Qtr1} + 57.0\ \text{Qtr2} + 26.0\ \text{Qtr3}$

Year	Quarter	q1	q2	q3	Sales	Forecast	AFE	SFE	APE
1	1	1	0	0	125	124	1	1	0.80
	2	0	1	0	153	152	1	1	0.65
	3	0	0	1	106	121	15	225	14.15
	4	0	0	0	88	95	7	49	7.95
2	1	1	0	0	118	124	6	36	5.08
	2	0	1	0	161	152	9	81	5.59
	3	0	0	1	133	121	12	144	9.02
	4	0	0	0	102	95	7	49	6.86
3	1	1	0	0	138	124	14	196	10.14
	2	0	1	0	144	152	8	64	5.56
	3	0	0	1	113	121	8	64	7.08
	4	0	0	0	80	95	15	225	18.75
4	1	1	0	0	109	124	15	225	13.76
	2	0	1	0	137	152	15	225	10.95
	3	0	0	1	125	121	4	16	3.20
	4	0	0	0	109	95	14	196	12.84
5	1	1	0	0	130	124	6	36	4.62
	2	0	1	0	165	152	13	169	7.88
	3	0	0	1	128	121	7	49	5.47
	4	0	0	0	96	95	1	1	1.04
						Mean	8.9	102.6	7.57



MAE = 8.9, MSE = 102.6 and MAPE = 7.57%

Seasonality With Trend

- The regression approach can be extended to include situations where the time series contains both a seasonal effect and a linear trend.
- The general form of the estimated multiple regression equation for modeling both the quarterly seasonal effects and the linear trend in time series is as follows:

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

where

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

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

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

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

t = time period

- For example, The estimated multiple regression equation for the quarterly television set sales time series which has a seasonal pattern with an upward linear trend that will need to be accounted for in order to develop accurate forecasts of quarterly sales.

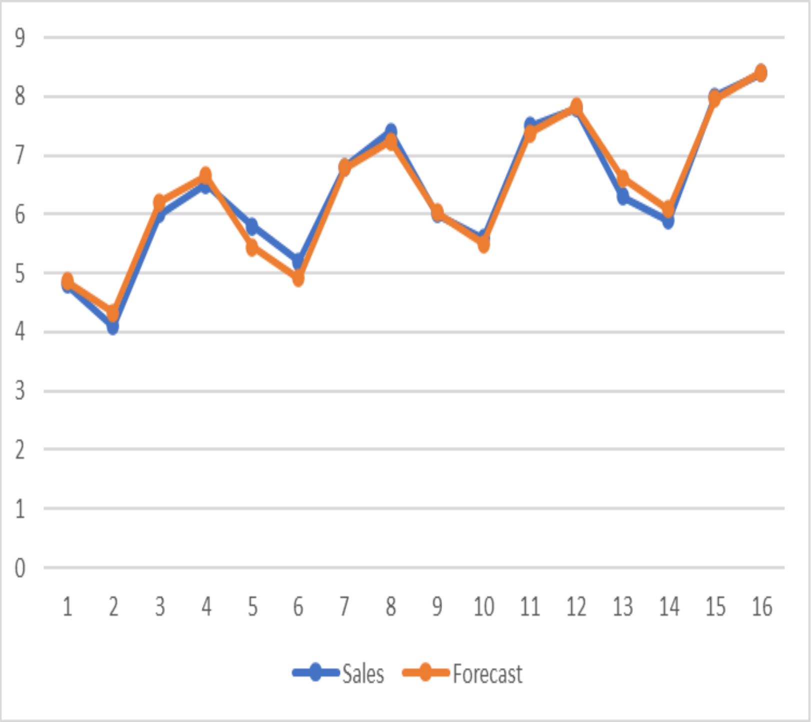
Seasonality With Trend

Television example

Use MegaStat (Correlation / Regression, Regression, X (q1, q2, q3, t), Y (Sales)) to get the coefficient of the equation.

$$\text{Sales} = 6.07 - 1.36\ \text{Qtr1} - 2.03\ \text{Qtr2} - .304\ \text{Qtr3} + .146t$$

Year	Quarter	q1	q2	q3	t	Sales	Forecast	AFE	SFE	APE
1	1	1	0	0	1	4.8	4.85	0.05	0.00	1.07
	2	0	1	0	2	4.1	4.33	0.23	0.05	5.52
	3	0	0	1	3	6	6.20	0.20	0.04	3.35
	4	0	0	0	4	6.5	6.65	0.15	0.02	2.33
2	1	1	0	0	5	5.8	5.43	0.37	0.13	6.31
	2	0	1	0	6	5.2	4.91	0.29	0.08	5.60
	3	0	0	1	7	6.8	6.78	0.02	0.00	0.24
	4	0	0	0	8	7.4	7.23	0.17	0.03	2.25
3	1	1	0	0	9	6	6.02	0.02	0.00	0.27
	2	0	1	0	10	5.6	5.49	0.11	0.01	1.94
	3	0	0	1	11	7.5	7.37	0.13	0.02	1.78
	4	0	0	0	12	7.8	7.82	0.02	0.00	0.21
4	1	1	0	0	13	6.3	6.60	0.30	0.09	4.74
	2	0	1	0	14	5.9	6.07	0.17	0.03	2.94
	3	0	0	1	15	8	7.95	0.05	0.00	0.64
	4	0	0	0	16	8.4	8.40	0.00	0.00	0.01
							Mean	0.14	0.03	2.45



MAE = 0.14, MSE = 0.03 and MAPE=2.45%

Time Series Decomposition

- An attention to **time series decomposition** should be done. Time series decomposition can be used to separate or decompose a time series into seasonal, trend, and irregular components. While this method can be used for forecasting, its primary applicability is to get a better understanding of the time series. Many business and economic time series are maintained and published by government agencies such as the Census Bureau and the Bureau of Labor Statistics. These agencies use time series decomposition to create de-seasonalized time series.
- Time series decomposition methods assume that *Yt*, the actual time series value at period *t*, is a function of three components: a trend component; a seasonal component; and an irregular or error component. How these three components are combined to generate the observed values of the time series depends upon whether we assume the relationship is best described by an additive or a multiplicative model.
- If the sizes of the seasonal fluctuations in earlier time periods are about the same as the sizes of the seasonal fluctuations in later time periods, an **additive model** is appropriate. However, if the seasonal fluctuations change over time, growing larger as the sales volume increases because of a long-term linear trend, then a **multiplicative model** should be used.

Time Series Decomposition

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

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

where

$$\text{Trend}_t = \text{trend value at time period } t$$

$$\text{Seasonal}_t = \text{seasonal value at time period } t$$

$$\text{Irregular}_t = \text{irregular value at time period } t$$

- The irregular or error component accounts for the variability in the time series that cannot be explained by the trend and seasonal components.
- A **multiplicative decomposition model** takes the following form:

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

- Note that, trend is measured in units of the item being forecast. However, the seasonal and irregular components are measured in relative terms, with values above 1.00 indicating effects above the trend and values below 1.00 indicating effects below the trend.
- The multiplicative decomposition model is used most often in practice especially in business and economic time series.

Time Series Decomposition

Example: We will work with the quarterly television set sales time series. Use MegaStat (Time Series / Forecasting, Deseasonalization)

- Centered moving average calculations for the television set sales time series are shown in the following table with seasonal indexes and sales de-seasonalized.

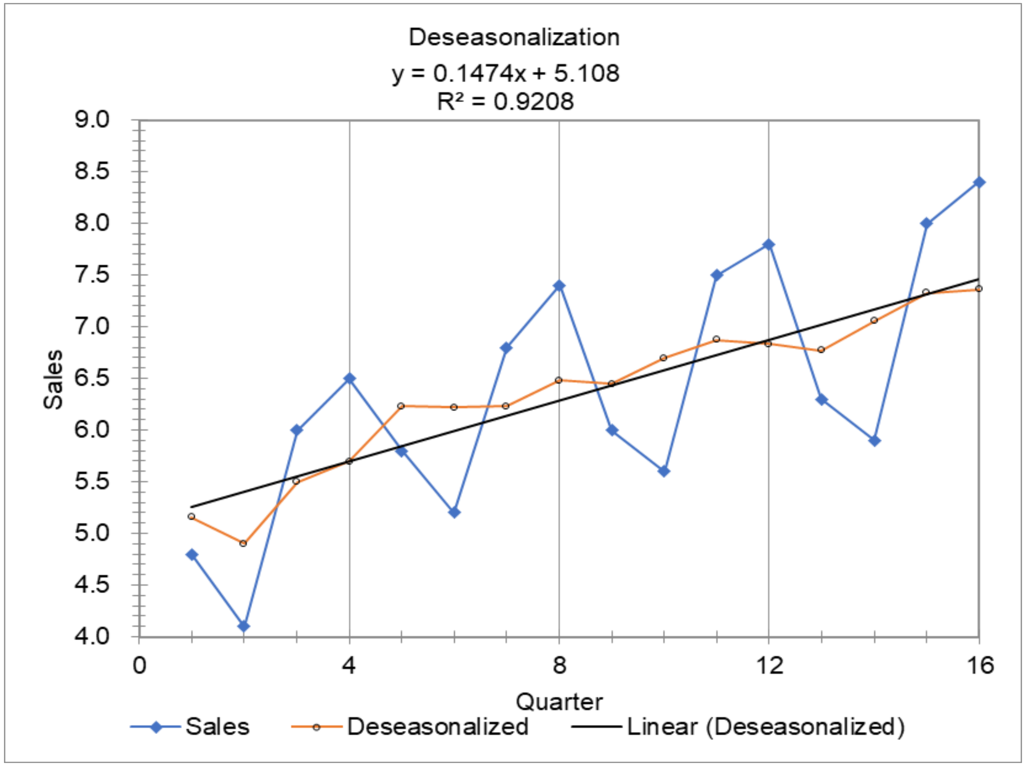
Note that the seasonal indexes for third and fourth quarters are above 1 indicating above sales trend effect.

t	Year	Quarter	Sales	Centered Moving Average	Ratio to CMA	Seasonal Indexes	Sales Deseasonalized
1	1	1	4.8			0.931	5.16
2	1	2	4.1			0.836	4.90
3	1	3	6.0	5.475	1.096	1.092	5.50
4	1	4	6.5	5.738	1.133	1.141	5.69
5	2	1	5.8	5.975	0.971	0.931	6.23
6	2	2	5.2	6.188	0.840	0.836	6.22
7	2	3	6.8	6.325	1.075	1.092	6.23
8	2	4	7.4	6.400	1.156	1.141	6.48
9	3	1	6.0	6.538	0.918	0.931	6.45
10	3	2	5.6	6.675	0.839	0.836	6.70
11	3	3	7.5	6.763	1.109	1.092	6.87
12	3	4	7.8	6.838	1.141	1.141	6.83
13	4	1	6.3	6.938	0.908	0.931	6.77
14	4	2	5.9	7.075	0.834	0.836	7.05
15	4	3	8.0			1.092	7.33
16	4	4	8.4			1.141	7.36

Time Series Decomposition

- The graph for the de-seasonalized sales time series show the original sales with and without seasonal effects along with a sales trend $T_t = b_0 + b_1 t$
- The seasonal indexes provide some insight about the seasonal component in television set sales. The best sales quarter is the fourth quarter, with sales averaging 14% above the trend estimate. The worst, or slowest, sales quarter is the second quarter; its seasonal index of .84 shows that the sales average is 16% below the trend estimate.

Calculation of Seasonal Indexes					
	1	2	3	4	
1			1.096	1.133	
2	0.971	0.840	1.075	1.156	
3	0.918	0.839	1.109	1.141	
4	0.908	0.834			
mean:	0.932	0.838	1.093	1.143	4.007
adjusted:	0.931	0.836	1.092	1.141	4.000



Summary

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

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

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

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

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

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

Summary

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

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

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

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

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

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

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

Summary

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

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

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

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

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

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

Summary

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