Chapter 6

August 14, 2016

The preceding three chapters of this book dealt with regression methods applied to cross sectional data. When we have cross sectional data our approach to forecasting is to use the explanatory variables to forecast the value of the dependent variable. This approach to forecasting is often termed *interpolation*.

The other broad method of quantitative business forecasting is to use the history of the data we wish to forecast. We examine the historical pattern of the data we wish to forecast and try to find a discernible pattern. Using this pattern, we forecast forward the future pattern of the data. This method of forecasting is often termed *extrapolation*. The use of the historical pattern of the data to be forecast is known as *time series methods*.

Stage 1 Collection and Analysis of Times Series Data

"The future lies ahead."

— Mort Sahl

A time series of data is data collected sequentially over equal periods of time. There are many forms of time series data. Time series can be collected as daily data (such as the daily closing price of a stock), weekly (such as weekly receipts data), monthly (such as monthly sales data), quarterly (such as quarterly revenue data), or yearly (such as annual profits data).

Remember though, the time period of data collection must be consistent. If our time series data are both in monthly and quarterly form then forecasting can be exceedingly difficult. When one collects time series data, be certain that the time period of collection will remain consistent.

To create a time series of data from our example of Sales Volume data we now collect a series of Sales Volume data from *just one location over consecutive monthly periods.* We collected a data set of four years, or 48 consecutive months (48 periods), of Sales Volume.

Table 1: Four Years of Time Series Data

Period	Year 1	Period	Year 2	Period	Year 3	Period	Year 4
1	95.5	13	255.0	25	325.0	37	440.0
2	127.0	14	263.0	26	333.0	38	472.0
3	118.0	15	253.0	27	348.0	39	494.5
4	153.0	16	263.0	28	355.0	40	512.0
5	175.5	17	274.0	29	248.0	41	519.0
6	180.0	18	278.0	30	363.0	42	526.0
7	202.0	19	284.5	31	370.0	43	530.5
8	190.0	20	300.0	32	388.0	44	558.0
9	220.5	21	305.5	33	412.0	45	557.5
10	245.0	22	310.5	34	435.0	46	590.0
11	260.0	23	372.0	35	490.5	47	636.4
12	320.0	24	412.0	36	532.5	48	677.4

Graphically displaying Time Series Data: Time Series Plot

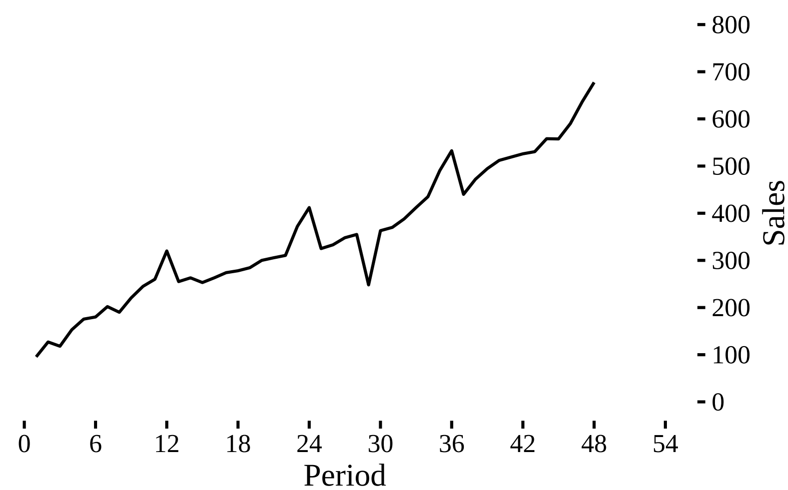

Figure 1: The corresponding time series plot or time series graph of Monthly Sales Volume.

Tips on Graphing Times Series Data: The Practice

Good graphical presentations of time series data and forecasts are both useful and powerful in the presentation of ideas. A good layout for a time series graph is always helpful in data analysis, model fitting, and forecasting.

Notice in our graph, that we allowed for additional space to the right of the last observation of our data series. We have 48 observations but allowed for 54 observations. Notice too that because the time series graph reveals an upward direction of the data series we also allowed for additional space above the last observation. The value of the last observation is 677.4 but allowed for values up to 800. In other words we have provided for plenty of "white space" for our graph.

Along the horizontal or time period axis we created horizontal reference lines at every six periods. It would be too dense to create horizontal reference lines at every of 54 periods. And it would be too sparse to form lines at every 12th period. These are monthly data so every 6 periods denotes a half year. Period 6 is June, Period 12 is December, Period 18 is June again, and so on.

Along the vertical or Sales Volume axis we set vertical reference lines at every 100 units. Units of 100 allow for sufficient spacing so the graph does not become crowded, and yet detailed enough so that we can estimate values with reasonable accuracy.

In summary, when producing time series graph, try to make the graphs uncluttered without sacrificing accuracy or information.

Time Series Data: Time Series Notation

When discussing observations in our data set we think of the data in the following form:

$$Y_1, Y_2, Y_3, \ldots, Y_T$$

Y_1 is the first observation, Y_2 is the second observation, and so on. Y_1 is the oldest observation down through Y_T which is the most recent observation. In our example data series:

$$Y_1 = 95.5$$
$$Y_2 = 127.0$$
$$Y_3 = 118.0$$
$$\vdots$$
$$Y_{48} = 677.4$$

The uppercase letter T of Y_T will always mean the most recent observation in our time series. Y_{T+1} means then an observation one period into the future. Thus, $Y_{T+1}, Y_{T+2}, Y_{T+3}, \ldots$ denote future observations $one, two, three, \ldots$ periods beyond the present time Y_T. Similarly, Y_{T-1} denotes an observation one period in the past, so that $Y_{T-1}, Y_{T-2}, Y_{T-3}, \ldots$, denote observations $one, two, three, \ldots$ periods in the past.

We can generalize this notation to:

$Y_{T+\ell}$: A future observation of ℓ periods ahead.
$Y_{T-\ell}$: A historical observation of ℓ periods back.

As we distinguish between Y_i and $\hat{Y_i}$ in regression models, we shall distinguish between $Y_{T+\ell}$ and $\hat{Y_T}\ell$ in time series models.

$Y_{T+\ell}$ is the future observation of Y, ℓ period ahead, and unknown at time T.

$\hat{Y_T}\ell$ is the forecast of the future observation, ℓ periods ahead, made at time T.

Stage 1 Continued: The Theory of Collecting & Analyzing Data

We devote most of this chapter to Stage 1 of the Forecasting Process. With time series analysis, we recommend considerable investigation to really understand the time series data and what it is “telling us”, before attempting to develop forecasting models. Consequently, we shall begin this time series analysis with some easy, straightforward methods of analysis.

Smoothing Out Time Series Data

Almost all time series data are rough, having a stochastic element to the series. We shall discuss methods of "smoothing out" the data so that we may observe an underlying structure or pattern. The first, and simplest, method of smoothing time series data is through averaging.

Smoothing Out the Data by Averaging

If we believe there exists some underlying trend or pattern in the time series data, then by "smoothing the data" we smooth out (or average out) the random variations to reveal the underlying pattern in the series.

The Practice: Simple Average Smoothing

We begin with a time series of 30 observations, $T = 30$, which has a constant mean, or a slowly changing mean, over time, we call this "horizontal data". As in Table 2 and Figure 2.

Table 2: Horizontal Data

Period	Actual	Period	Actual
1	354	16	332
2	368	17	342
3	329	18	357
4	389	19	357
5	375	20	357
6	375	21	344
7	367	22	361
8	364	23	358
9	379	24	345
10	386	25	367
11	329	26	380
12	334	27	387
13	372	28	346
14	329	29	321
15	320	30	372

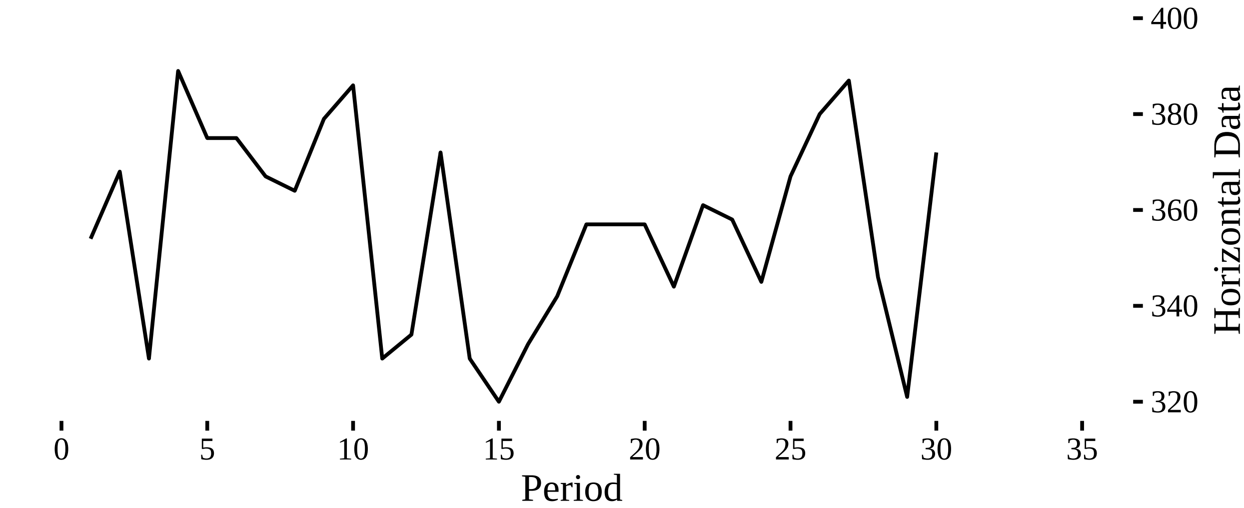

Figure 2: Horizontal Data.

Because the Y_t are varying around a constant mean, a first smoothing of Y_t is just its mean, $\bar{Y}$. In this example, $\bar{Y} = 356.5$.[1]

[1]
$$\bar{Y} = \sum_{t=1}^{T} Y_t$$

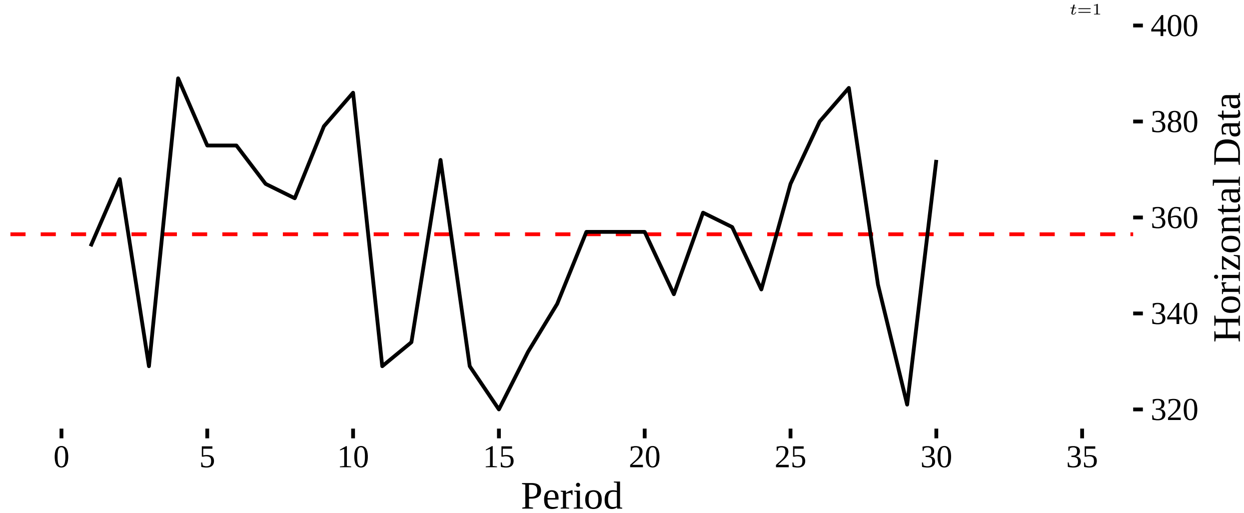

Figure 3: Horizontal Data with Plotted Mean.

Moving Average Smoothings

The simple average discussed above is a smoothing method over all observations. A *moving average* is the technique of creating successive new averages by dropping the "oldest" observation and adding the most recent observation to calculate the new average.

3-period Centered Moving Average Smoothing for period t

To begin with an example, examine the notation in the margin[2]. Next, we shall use the data from Table 2 to illustrate this method. For example, a 3-period Centered Moving Average (at period 7) is:

$$CMA_7(3) = \frac{Y_8+Y_7+Y_6}{3}, \text{ and } CMA_7(3) = \frac{364+367+375}{3} = 368.67$$

Table 3: Horizontal Data subset & $CMA_7(3)$

Period	Actual	Smoothing
6	375	...
7	367	368.67
8	364	...

[2] 3-period Centered Moving Average:

$$CMA_t(3) = \frac{Y_{T+1} + Y_T + Y_{T-1}}{3}$$

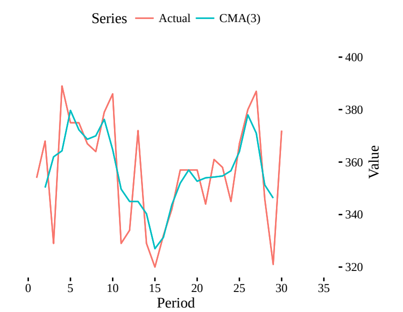

Figure 4: Centered Moving Average(3)

This smoothing method is called a 3-period centered moving average because the computed value is placed at the center, or middle, of the 3 periods being used in the calculations.

Table 4: Horizontal Data with Centered Smoothing Average (3)

Period	Actual	CMA(3)	Period	Actual	CMA(3)
1	354	NA	16	332	331.3
2	368	350.3	17	342	343.7
3	329	362.0	18	357	352.0
4	389	364.3	19	357	357.0
5	375	379.7	20	357	352.7
6	375	372.3	21	344	354.0
7	367	368.7	22	361	354.3
8	364	370.0	23	358	354.7
9	379	376.3	24	345	356.7
10	386	364.7	25	367	364.0
11	329	349.7	26	380	378.0
12	334	345.0	27	387	371.0
13	372	345.0	28	346	351.3
14	329	340.3	29	321	346.3
15	320	327.0	30	372	NA

5-period Centered Moving Average Smoothing for period t

A 5 period moving average smoothing is "smoother" than a 3 period because it uses a larger set of observations.[3]. As an example we smooth period 7 of the data in Table 4.

[3] 5-period Centered Moving Average:

$$CMA_t(5) = \frac{Y_{T+2} + Y_{T+1} + Y_T + Y_{T-1} + Y_{T-2}}{5}$$

$$CMA_7(5) = \frac{Y_9 + Y_8 + Y_7 + Y_6 + Y_5}{5}$$

$$CMA_7(5) = \frac{379 + 364 + 367 + 375 + 375}{5} = 372$$

Table 5: Horizontal data subset with $CMA_t(5)$

Period	Actual	Smoothing
4	389	...
5	375	...
6	375	...
7	367	372
8	364	...
9	379	...
10	386	...

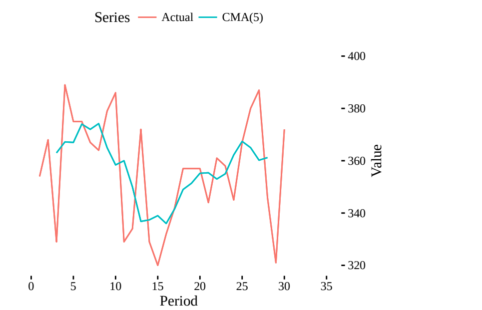

Figure 5: Centered Moving Average(5)

We list below both the 3 and 5 period moving average smoothing.

Table 6: Horizontal data with $CMA_t(3) and CMA_t(5)$

Period	Actual	CMA(3)	CMA(5)	Period	Actual	CMA(3)	CMA(5)
1	354	NA	NA	16	332	331.3	336.0
2	368	350.3	NA	17	342	343.7	341.6
3	329	362.0	363.0	18	357	352.0	349.0
4	389	364.3	367.2	19	357	357.0	351.4
5	375	379.7	367.0	20	357	352.7	355.2
6	375	372.3	374.0	21	344	354.0	355.4
7	367	368.7	372.0	22	361	354.3	353.0
8	364	370.0	374.2	23	358	354.7	355.0
9	379	376.3	365.0	24	345	356.7	362.2
10	386	364.7	358.4	25	367	364.0	367.4
11	329	349.7	360.0	26	380	378.0	365.0
12	334	345.0	350.0	27	387	371.0	360.2
13	372	345.0	336.8	28	346	351.3	361.2
14	329	340.3	337.4	29	321	346.3	NA
15	320	327.0	339.0	30	372	NA	NA

Below is a graph of actual values with a 5-period smoothing superimposed. The following graph adds an additiona 3-period smoothing for comparison.

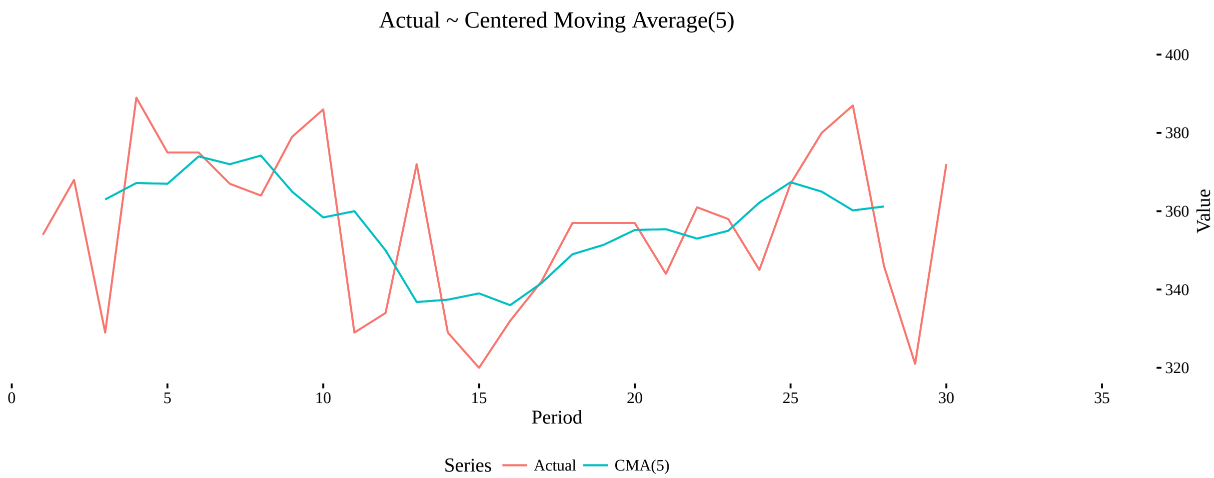

Figure 6: Centered Moving Average(5)

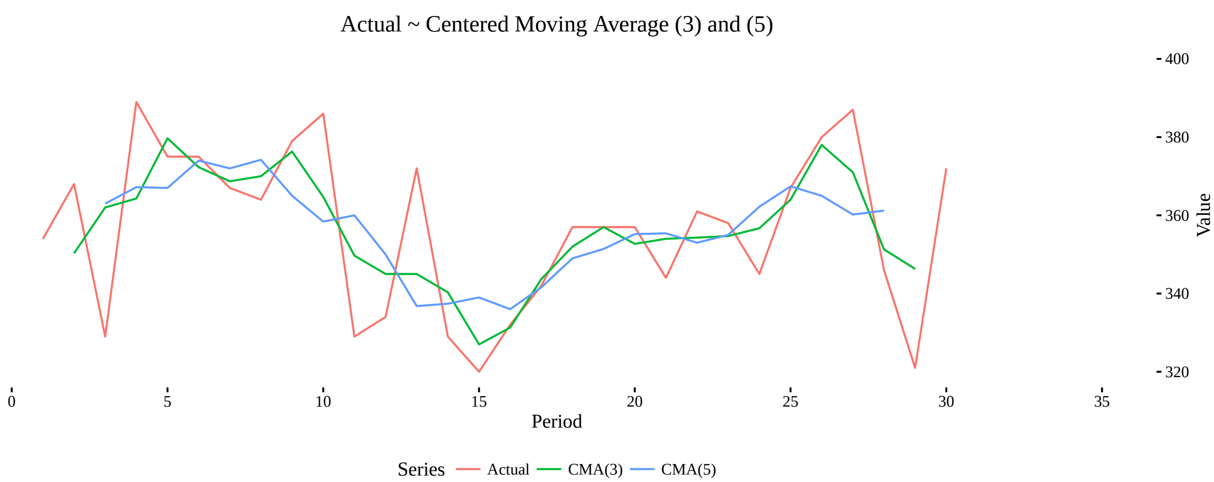

Figure 7: Centered Moving Average(3,5)

4-Period Centered Moving Average Smoothing

A 4-period Centered Moving Average is possible, but the issue of placing the results remains.[4]

[4] 4-period Centered Moving Average:

$$CMA_t(4) = \frac{Y_{T+2} + 2Y_{T+1} + 2Y_T + 2Y_{T-1} + Y_{T-2}}{8}$$

Technically, if the first 4 periods are used, then the placement of $CMA_t(4)$ is between periods 2 and 3 at "period 2.5."

$$CMA_t(4) = \frac{Y_4 + Y_3 + Y_2 + Y_1}{4}$$

$$CMA_{2.5}(4) = \frac{389 + 329 + 368 + 354}{4} = 360$$

The following 4-period centered smoothing is placed at "period 3.5".

$$CMA_t(4) = \frac{Y_5 + Y_4 + Y_3 + Y_2}{4}$$

$$CMA_{3.5}(4) = \frac{375 + 389 + 329 + 368}{4} = 365.25$$

The "adjusted centered smoothing" for period 3 is the average of period 2.5 and 3.5 smoothings. Algebraically, it can be shown that the adjusted moving average of 4 periods, starting with period $t = 3$ is:

$$CMA_3(4) = \frac{360 + 365.25}{2} = 362.63$$

The 4-period Centered Moving Average Smoothing is especially suited for quarterly time series. However, the Centered Moving Average Smoothing formula may be generalized to any even number of periods.

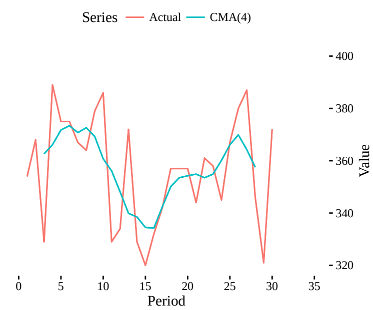

Figure 8: Centered Moving Average(4)

Table 7: Horizontal data with $CMA_t(4)$

Period	Actual	CMA(4)	Period	Actual	CMA(4)
1	354	NA	16	332	334.25
2	368	NA	17	342	342.38
3	329	362.62	18	357	350.12
4	389	366.12	19	357	353.50
5	375	371.75	20	357	354.25
6	375	373.38	21	344	354.88
7	367	370.75	22	361	353.50
8	364	372.62	23	358	354.88
9	379	369.25	24	345	360.12
10	386	360.75	25	367	366.12
11	329	356.12	26	380	369.88
12	334	348.12	27	387	364.25
13	372	339.88	28	346	357.50
14	329	338.50	29	321	NA
15	320	334.50	30	372	NA

6-Period Centered Moving Average Smoothing

$$CMA_t(6) = \frac{Y_{T+2} + 2Y_{T+2} + 2Y_{T+1} + 2Y_T + 2Y_{T-1} + 2Y_{T-2} + Y_{T-3}}{12}$$

12-Period Centered Moving Average Smoothing

$$CMA_t(12) = \frac{Y_{T+6}+2Y_{T+5}+2Y_{T+4}+2Y_{T+3}+2Y_{T+2}+2Y_{T+1}+2Y_T+2Y_{T-1}+2Y_{T-2}+Y_{T-3}+2Y_{T-4}+2Y_{T-5}+Y_{T-6}}{24}$$

Not all smoothings must be placed in the "center" of the data. We have done so only because an even numbered moving average falls in the "middle" of the data. However, for example, one may choose to place a 4 period weekly smoothing at the last period of the four periods smoothed, representing the end of a month.

The length of periods of smoothing is usually dependent on the type of time series data being analyzed. For example, some analyses of stock market and bond market time series data use 13-period and 39-period smoothing of weekly data instead of an even numbered smoothing.

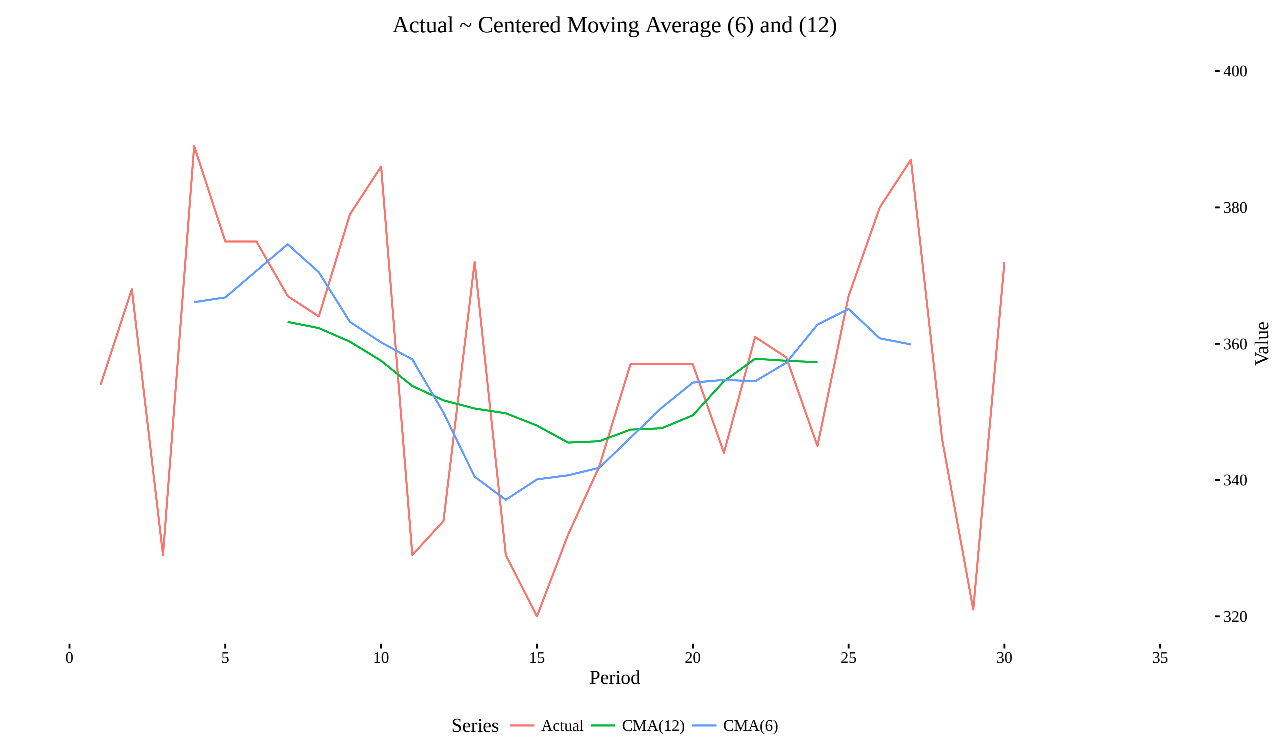

Figure 9: Centered Moving Average(6,12)

Weighted Moving Average Smoothings

Simple moving averages assume an equal weight given to each observation in the calculations. However, There may be situations in which it is better to assign *greater* weight to the most recent observation and *less* weight to observations in the past.

3-period Centered Weighted Moving Average

A 3-period Centered Weighted Moving Average is denoted as $CWMA_t(3)$ and can be found in the margin.[5]

As an example, consider the following $CWMA_t(3)$ notation:

$$CWMA_t(3) = \frac{0.6Y_{T+1} + 0.3Y_T + 0.1Y_{T-1}}{3}$$

All that is required is that the sum of the weights equal 1, as in:

$$0.6 + 0.3 + 0.1 = 1$$

A weighted moving average may be generalized to any number of historical periods and set of weights. The choice of the number of historical periods and the weights used are determined, in part, by the Forecast Analyst who has collected the data and has some appreciation of the data structure.

Table 8: 3 Period Centered Weighted Moving Average

Period	Actual	CWMA(3)	Period	Actual	CWMA(3)
1	354	NA	16	332	336.8
2	368	343.2	17	342	350.0
3	329	368.9	18	357	355.5
4	389	374.6	19	357	357.0
5	375	376.4	20	357	349.2
6	375	370.2	21	344	355.5
7	367	366.0	22	361	357.5
8	364	373.3	23	358	350.5
9	379	381.7	24	345	359.5
10	386	351.1	25	367	372.6
11	329	337.7	26	380	382.9
12	334	356.3	27	387	361.7
13	372	342.4	28	346	335.1
14	329	327.9	29	321	354.1
15	320	328.1	30	372	NA

[5] 3-period Centered Weighted Moving Average. In general terms, we denote the weights by ω_1, ω_2, and ω_3 (omega sub 1, omega sub 2, and omega sub 3).

$$CWMA_t(3) = \frac{\omega_1 Y_{T+1} + \omega_2 Y_T + \omega_3 Y_{T-1}}{3}$$

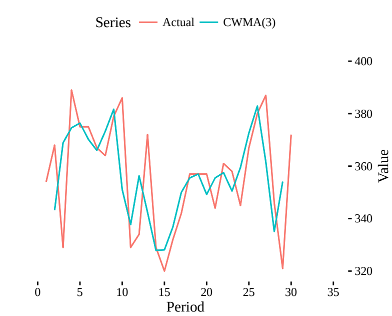

Figure 10: Centered Weighted Moving Average(3)

The Underlying Structure of Time Series Data[6]

[6] Stage 1 continued

Classical Time Series Decomposition

Given an economic time series as shown in Figure 1, one of the most common approaches to understanding such data is through *classical time series decomposition*. The composition of the time series is decomposed into four components: trend, cyclical, seasonal, and irregular components. Considering the four components collectively is the

method of the *classical time series decomposition*. The components can be either added together, or multiplied by each other, to define a time series. If a time series decomposition is additive, all components are expressed in the physical units of Y_t and simply added together. If a times series decomposition is expressed in multiplicative form, component values are expressed in a combination of physical units and percentage of trend T_t, and multiplied together. The actual value Y_t is the *product* of the trend component T_t expressed in physical units, the cyclical component C_t expressed as a percentage of T_t, the seasonal component S_t as a percentage of $T_t \times C_t$, and the irregular component expressed as a percentage of $T_t \times C_t \times S_t$.

Additive Time Series Decomposition

$$Y_t = T_t + C_t + S_t + I_t$$

.

Multiplicative Time Series Decomposition

$$Y_t = T_t \times C_t \times S_t \times I_t$$

.

Trend, seasonal fluctuations, and irregular variations are usually easily identifiable components of a time series. Calculations of trend and seasonality are quite straightforward, and adjustments for irregular variations can be simply handled. Determining the long term cyclical component of a time series is not so direct, and is a known limitation of this method. Long term business cycles are beyond the scope of this topic, and will be discussed in other chapters.

The Trend Component

The trend component of a time series at time t, denoted T_t, is the upward or downward progression of the data over time. Figure 11 below has a trend line superimposed over the actual data, indicating a downward trend. The OLS regression line of T_t is defined in the margin[7]. For our example, the Sales Volume series is a simple OLS regression against time period.

[7] $T_t = \beta_0 + \beta_1 TIME_t$

$$Sales_t = \beta_0 + \beta_1 Period_t$$

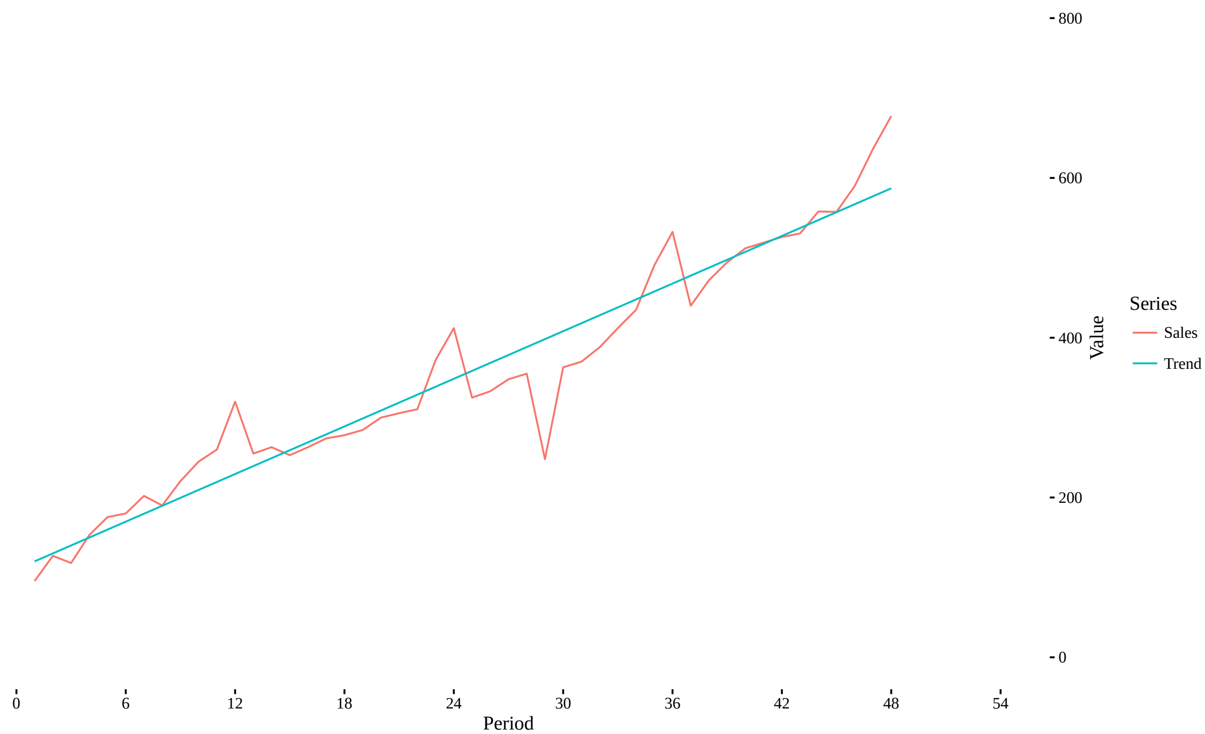

Figure 11: Sales Volume Data with Plotted OLS Trend line.

The Cyclical Fluctuations Component

The cyclical component of a time series at each time step t, denoted C_t, and illustrated below in Figure 12 and 13, are the broad up-and-down swings of the series around it's trend line. These cycles of high and low can last more than a year, having different lengths and amplitudes. The cyclical component of a time series is usually attributable to some larger aspect of the economy or business cycle by which this particular data have been affected.

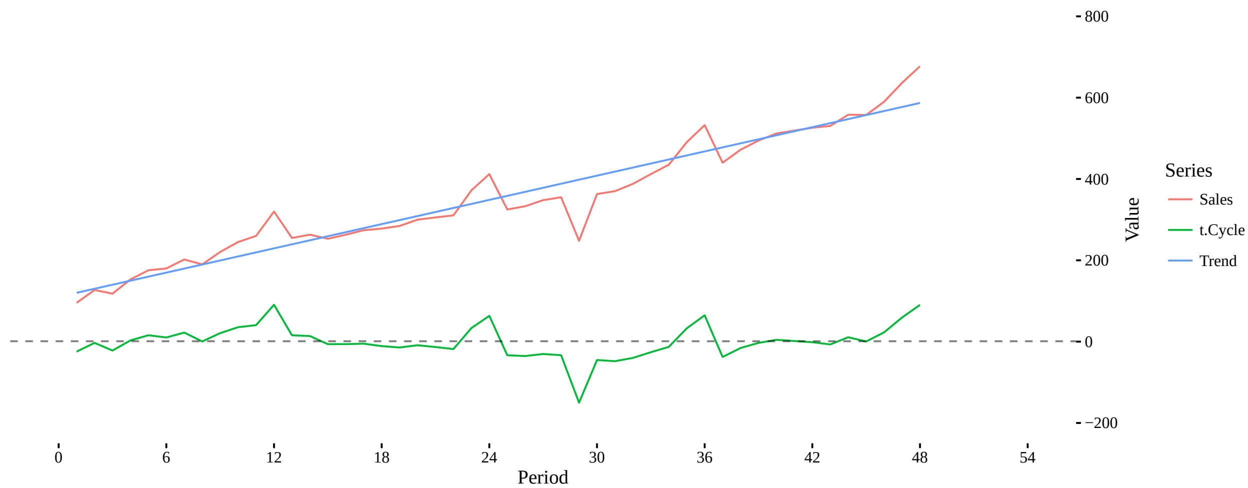

Figure 12: Sales Volume Data with Trend line and Cycle.

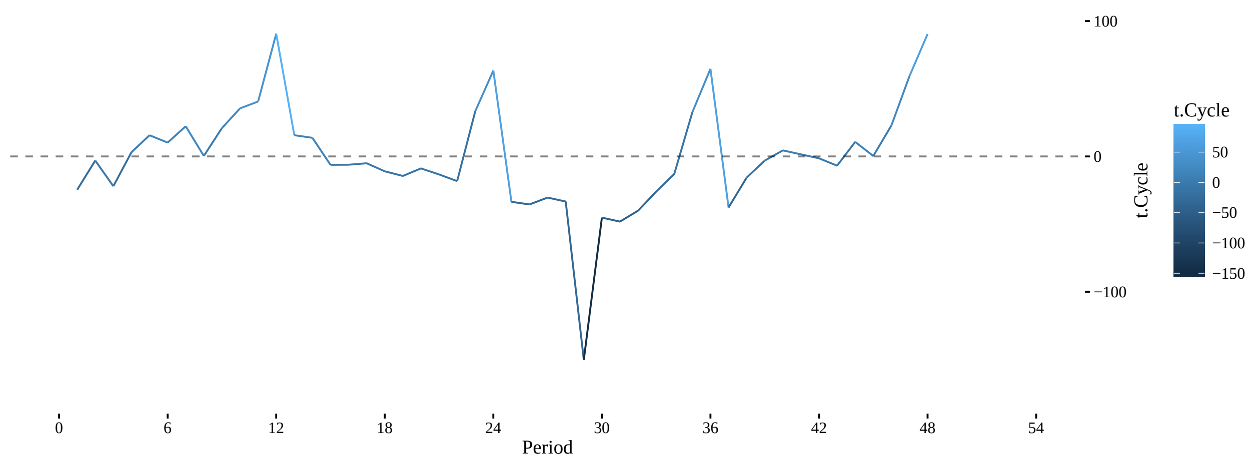

Figure 13: Sales Volume Data, Cycle component.

The Seasonal Fluctuations Component

The seasonal component of a time series at time t, denoted S_t, illustrated in the following example, are the reoccurring fluctuations within a year around the trend/cyclical components. Most business data reveal seasonal fluctuations, or *seasonality*, so it is important for businesses to know and plan for changes in sales or economic activity that will be more or less normal due to the seasonal nature of their product. For example, if sales are low or high due to seasonality, as oppose to trend, business may better adjust their expectations accordingly.

Seasonal Indices: The Ratio-to-Moving-Average Method

There are many occasions in the analysis of time series when we wish to isolate the seasonality in the data. This is known as determining the *seasonal indices* and then *deseasonalizing the data*. One method is the ***Ratio-to-Moving-Average*** method.

This was first developed in 1922 by Frederick Macaulay at the National Bureau of Economic Research and then adopted and promoted by Julius Shiskin of the U.S. Bureau of the Census. Often referred to as the Census II decomposition, this method has evolved and now in its most recent version, it is the Census **X-13 ARIMA-SEATS decomposition**. As the name suggests, the topic has become considerably more computationally sophisticated. The following basic example using the ratio-to-moving-average method offers a solution, and it is the initial step for some of the more advanced methods in use today.

Seasonal Index, Basic Example

We shall use the Sales Volume time series from Table 2 and Figure 1 for this example. Because they are monthly data and there are 12 months in a year, we first determine the 12-period Centered Moving Average Smoothing of the data. As there are 48 observations, or 4 years of data, we start with period 7 and end at period 42. Although the Centered Moving Average method does not allow one to calculate the cyclical component at the beginning and end of a time series, this is not an issue when creating Seasonal indices provided at least 2 years of data exist.

Step 1: Determine the 12-period Center Moving Average Smoothing of the Data. We use the formula on page 12.

Step 2: Calculate the Ratio of the Actual to the Moving Average (hence, the name of the method).

$$Ratio_t = \frac{Sales_t}{CMA_t(12)}$$

Table 9: Ratio to Moving Average Method

Period	Month	Sales	CMA12	Ratio	Period	Month	Sales	CMA12	Ratio
1	Jan	95.5	NA	NA	25	Jan	325.0	333.27	0.98
2	Feb	127.0	NA	NA	26	Feb	333.0	340.50	0.98
3	Mar	118.0	NA	NA	27	Mar	348.0	348.60	1.00
4	Apr	153.0	NA	NA	28	Apr	355.0	358.23	0.99
5	May	175.5	NA	NA	29	May	248.0	368.35	0.67
6	Jun	180.0	NA	NA	30	Jun	363.0	378.31	0.96
7	Jul	202.0	197.19	1.02	31	Jul	370.0	388.12	0.95
8	Aug	190.0	209.50	0.91	32	Aug	388.0	398.71	0.97
9	Sep	220.5	220.79	1.00	33	Sep	412.0	410.60	1.00
10	Oct	245.0	231.00	1.06	34	Oct	435.0	423.25	1.03
11	Nov	260.0	239.69	1.08	35	Nov	490.5	441.08	1.11
12	Dec	320.0	247.87	1.29	36	Dec	532.5	459.17	1.16
13	Jan	255.0	255.40	1.00	37	Jan	440.0	472.65	0.93
14	Feb	263.0	263.42	1.00	38	Feb	472.0	486.42	0.97
15	Mar	253.0	271.54	0.93	39	Mar	494.5	499.56	0.99
16	Apr	263.0	277.81	0.95	40	Apr	512.0	512.08	1.00
17	May	274.0	285.21	0.96	41	May	519.0	524.62	0.99
18	Jun	278.0	293.71	0.95	42	Jun	526.0	536.74	0.98
19	Jul	284.5	300.46	0.95	43	Jul	530.5	NA	NA
20	Aug	300.0	306.29	0.98	44	Aug	558.0	NA	NA
21	Sep	305.5	313.17	0.98	45	Sep	557.5	NA	NA
22	Oct	310.5	320.96	0.97	46	Oct	590.0	NA	NA
23	Nov	372.0	323.71	1.15	47	Nov	636.4	NA	NA
24	Dec	412.0	326.17	1.26	48	Dec	677.4	NA	NA

Step 3: We collect the Ratios in a table by Month and determine the Median value of the ratios. These Median values are often termed the *Unadjusted Seasonal Indices*.

Table 10: Ratios of Actual Sales to CWMA(12) Smoothing, with Median Monthly values (Unadjusted Seasonal Index)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Year 1	NA	NA	NA	NA	NA	NA	1.02	0.91	1.00	1.06	1.08	1.29
Year 2	1.00	1.00	0.93	0.95	0.96	0.95	0.95	0.98	0.98	0.97	1.15	1.26
Year 3	0.98	0.98	1.00	0.99	0.67	0.96	0.95	0.97	1.00	1.03	1.11	1.16
Year 4	0.93	0.97	0.99	1.00	0.99	0.98	NA	NA	NA	NA	NA	NA
Median	0.98	0.98	0.99	0.99	0.96	0.96	0.95	0.97	1.00	1.03	1.11	1.26

Step 4: We expect the average of the seasonal indices to be 1, so with 12 periods in the season, the sum of the seasonal indices must equal 12. Because their sum is 12.18, we multiply each Unadjusted Seasonal Index by $\frac{12}{12.18}$. For example, for Month 1 (January):

$$\frac{12}{12.18} \times 0.98 = 0.96$$

The sum of the Adjusted Seasonal Indices equals 12, so their average equals 1. As one might have expected, Sales are seasonally strongest in December, the 12th month, and weakest in July, the 7th month. The Adjusted Seasonal index values reflect this.

Table 11: Adjusted Seasonal Index

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.96	0.96	0.98	0.98	0.95	0.95	0.94	0.96	0.98	1.01	1.1	1.24

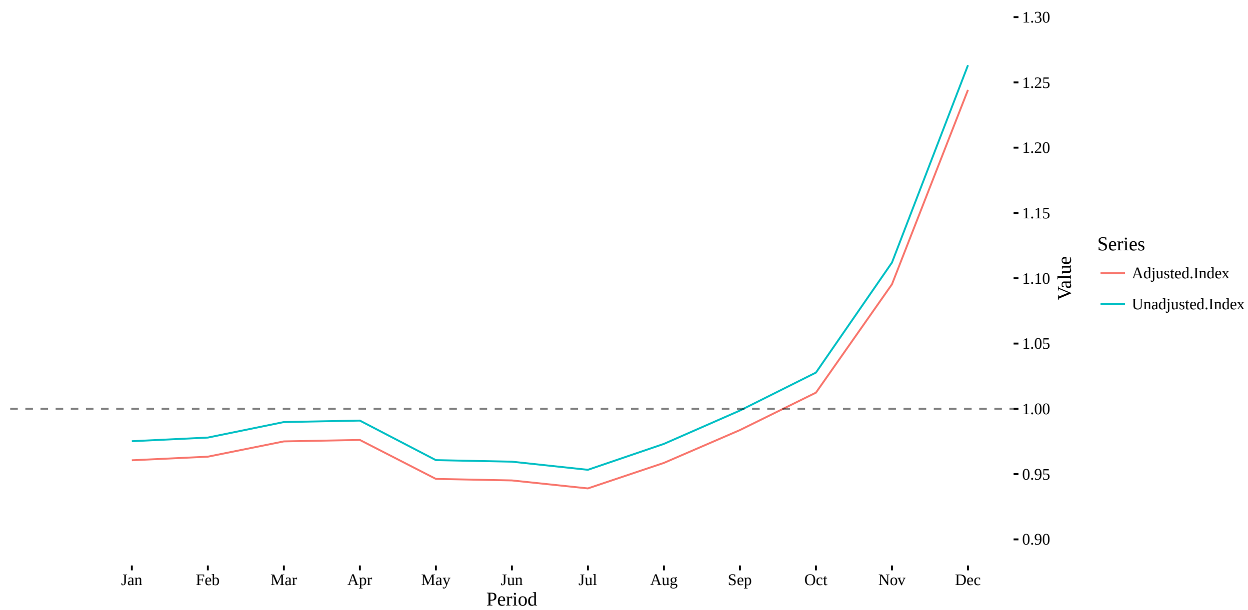

Figure 14: Sales Volum Data, Seasonal Index.

Step 5: The original, Actual, values are divided by the Seasonal Indices, creating a De-Seasonalized Series.

$$DeSeasonalized_t = \frac{Sales_t}{SeasonalIndex_t}$$

Notice in the Figure below that the periods of (11,12), (23,24), (35,36), and (47,48) are reoccurring local peaks in Sales Volume. This is the seasonality of Sales.

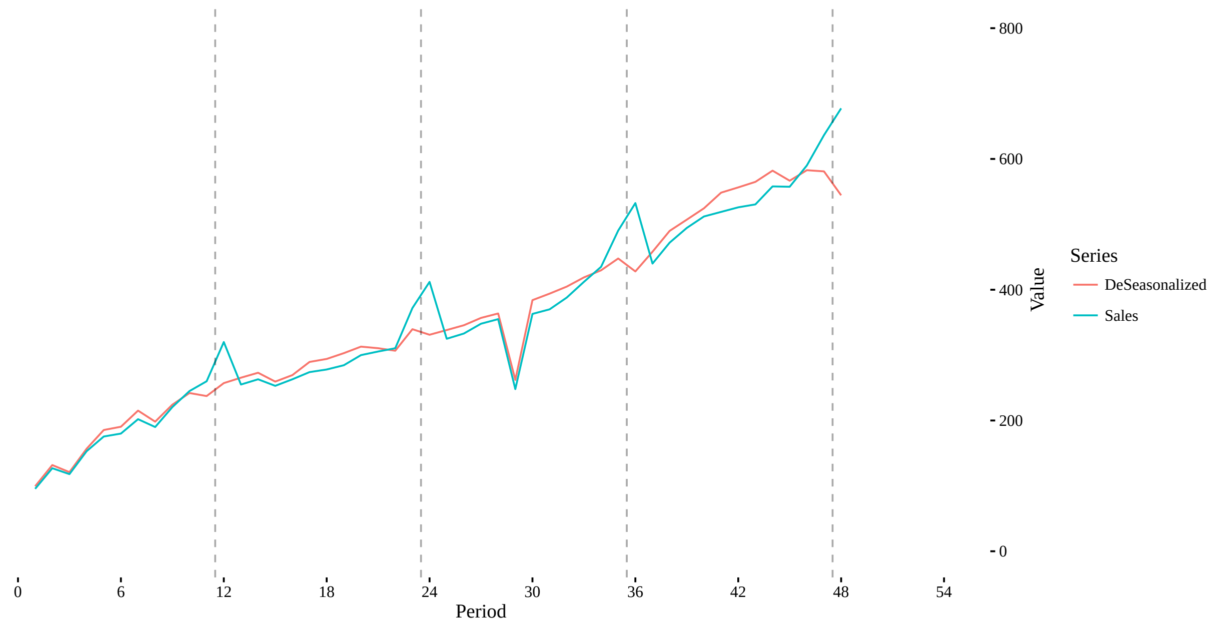

Figure 15: Sales Volume with DeSeasonalized Data.

Table 12: Actual Sales, CWMA(12) Smoothing, Ratio, Seasonal Index, and DeSeasonaled Sales

Period	Month	Sales	CMA12	Ratio	Seasonal	DeSeasonalized
1	Jan	95.5	NA	NA	0.96	99.42
2	Feb	127.0	NA	NA	0.96	131.83
3	Mar	118.0	NA	NA	0.98	121.02
4	Apr	153.0	NA	NA	0.98	156.74
5	May	175.5	NA	NA	0.95	185.45
6	Jun	180.0	NA	NA	0.95	190.44
7	Jul	202.0	197.19	1.02	0.94	215.11
8	Aug	190.0	209.50	0.91	0.96	198.21
9	Sep	220.5	220.79	1.00	0.98	224.15
10	Oct	245.0	231.00	1.06	1.01	242.00
11	Nov	260.0	239.69	1.08	1.10	237.36

Period	Month	Sales	CMA12	Ratio	Seasonal	DeSeasonalized
12	Dec	320.0	247.87	1.29	1.24	257.18
13	Jan	255.0	255.40	1.00	0.96	265.46
14	Feb	263.0	263.42	1.00	0.96	273.01
15	Mar	253.0	271.54	0.93	0.98	259.47
16	Apr	263.0	277.81	0.95	0.98	269.42
17	May	274.0	285.21	0.96	0.95	289.54
18	Jun	278.0	293.71	0.95	0.95	294.13
19	Jul	284.5	300.46	0.95	0.94	302.97
20	Aug	300.0	306.29	0.98	0.96	312.96
21	Sep	305.5	313.17	0.98	0.98	310.55
22	Oct	310.5	320.96	0.97	1.01	306.70
23	Nov	372.0	323.71	1.15	1.10	339.60
24	Dec	412.0	326.17	1.26	1.24	331.12
25	Jan	325.0	333.27	0.98	0.96	338.33
26	Feb	333.0	340.50	0.98	0.96	345.67
27	Mar	348.0	348.60	1.00	0.98	356.90
28	Apr	355.0	358.23	0.99	0.98	363.67
29	May	248.0	368.35	0.67	0.95	262.07
30	Jun	363.0	378.31	0.96	0.95	384.06
31	Jul	370.0	388.12	0.95	0.94	394.02
32	Aug	388.0	398.71	0.97	0.96	404.77
33	Sep	412.0	410.60	1.00	0.98	418.81
34	Oct	435.0	423.25	1.03	1.01	429.68
35	Nov	490.5	441.08	1.11	1.10	447.78
36	Dec	532.5	459.17	1.16	1.24	427.97
37	Jan	440.0	472.65	0.93	0.96	458.05
38	Feb	472.0	486.42	0.97	0.96	489.96
39	Mar	494.5	499.56	0.99	0.98	507.15
40	Apr	512.0	512.08	1.00	0.98	524.51
41	May	519.0	524.62	0.99	0.95	548.44
42	Jun	526.0	536.74	0.98	0.95	556.52
43	Jul	530.5	NA	NA	0.94	564.94
44	Aug	558.0	NA	NA	0.96	582.11
45	Sep	557.5	NA	NA	0.98	566.72
46	Oct	590.0	NA	NA	1.01	582.78
47	Nov	636.4	NA	NA	1.10	580.98
48	Dec	677.4	NA	NA	1.24	544.42

The Irregular Variations Component

The irregular component of a time series at time t, denoted I_t, is illustrated in Figure 16 below. They are the random, unexpected deviations from the trend/cyclical/seasonal components. Irregular changes or "shocks" to a time series are considered nonrecurring chance events such as strikes. They can be the result of unusually good (The movie E.T. and Reese's Pieces Candy) or unusually bad (Tylenol poisoning) publicity about a product.

In Figure 16 below, an irregular variation occurs around period 30 as the series drops abruptly and then returns to its upward trend.

$$Irregular_t = \frac{Sales_t}{Trend_t \times SeasonalIndex_t}$$

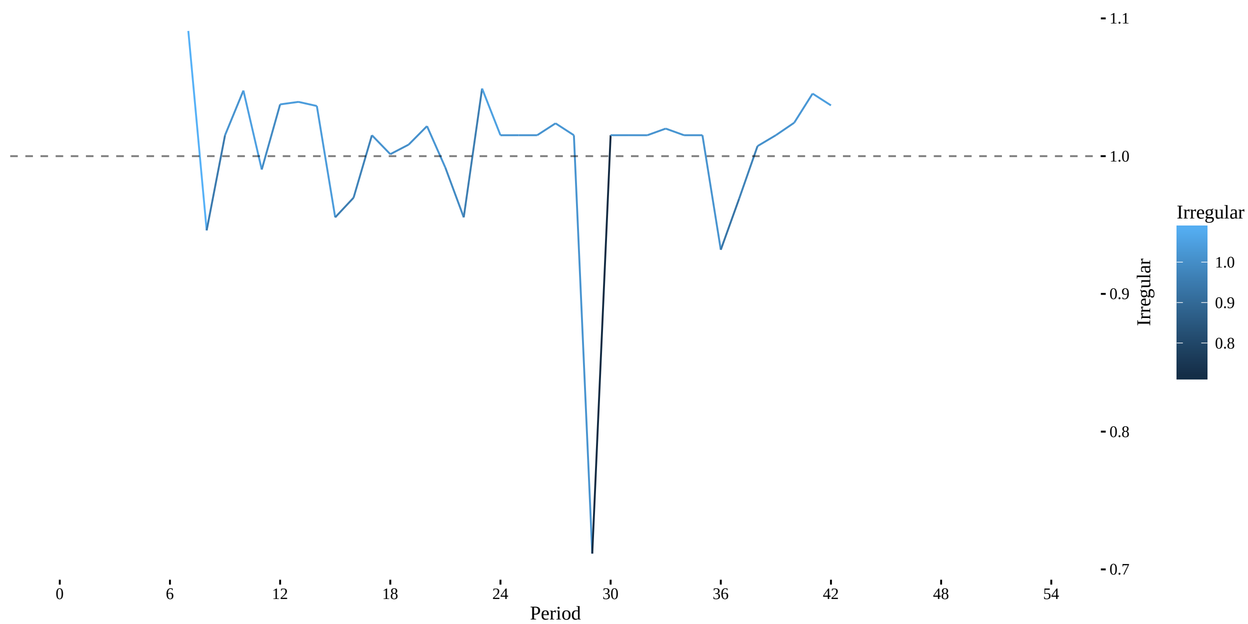

Figure 16: Sales Volume with DeSeasonalized Data.

Closing remarks

The astute reader may have noticed that we used the 12-Period Centered Moving Average to assist in calculating the seasonal component, not the OLS linear trend used in our initial example. The 12-Period Centered Moving Average discussed on pg. 12 offers a more dynamic fit for measuring the monthly Sales trend than the rigid OLS line. This is a topic we will cover in future chapters. For now, the reader is left to consider the benefits of using a smoothing function to measure the trend of a time series.

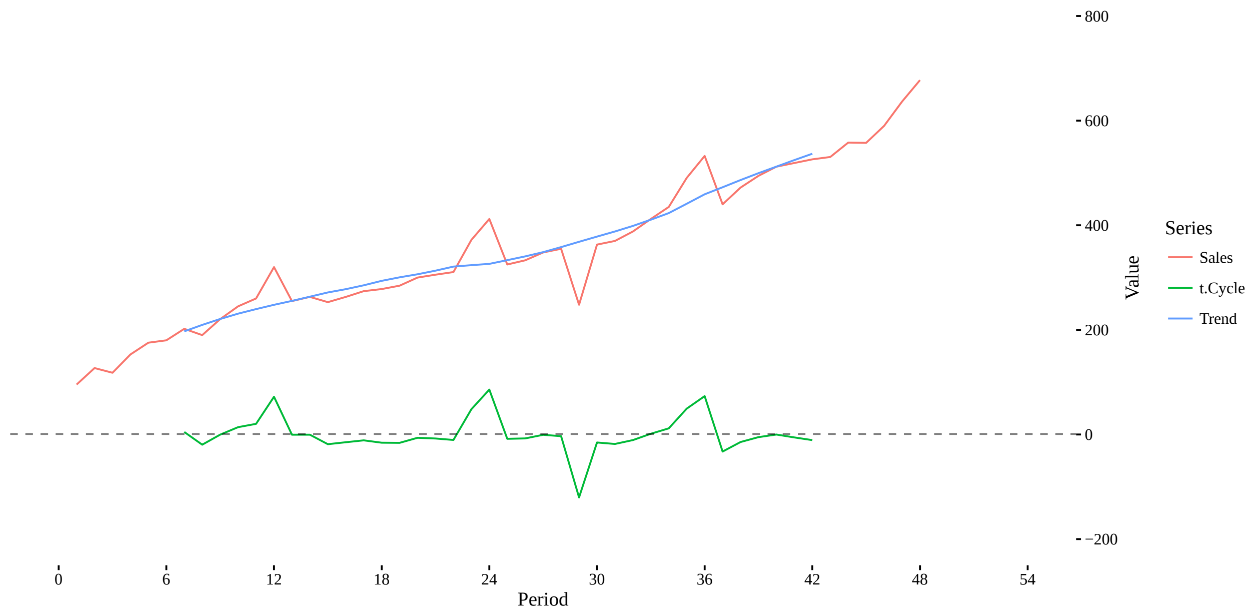

Figure 17: Sales Data with CWMA(12) Trend and Cycle.

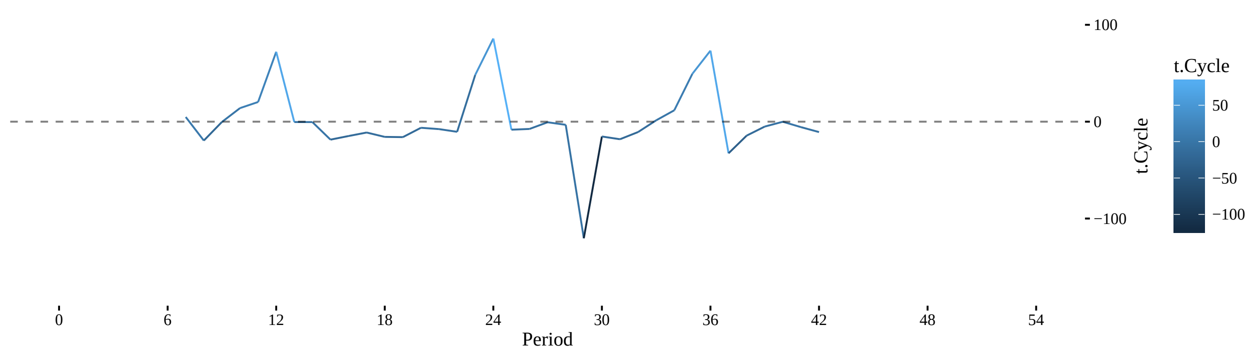

Figure 18: Sales Data, CWMA(12) Cycle component.

PROBLEMS AND QUESTIONS

Constant Mean Data

1. Listed below is a time series of 24 observations. Plot the data and determine the mean and variance of the the data set.

Period	Observation	Period	Observation
2	100	14	86
3	87	15	86
4	123	16	112
5	90	17	85
6	96	18	101
7	75	19	135
8	78	20	120
9	106	21	76
10	104	22	115
11	89	23	90
12	83	24	92

Smoothing of Data

2. Overlay on the time series graph of the data

a. The simple average of the data, as a smoothing.
b. A three period centered moving average.
c. A five period centered moving average.

3. Overlay on the time series graph of the data.

a. A four period centered moving average.
b. A three period weighted centered moving average with weights of .6 for the most recent, .3 for one period back, and .1 for two periods back.

De-seasonalizing Data

4. The following 24 observations of quarterly data contain trend and seasonality.

Period(t)	Quarter(Q_t)	Actual(Y_t)	Period(t)	Quarter(Q_t)	Actua(Y_t)
1	1	409	13	1	349
2	2	415	14	2	350
3	3	468	15	3	443
4	4	261	16	4	315
5	1	308	17	1	362
6	2	326	18	2	411
7	3	431	19	3	447
8	4	254	20	4	295
9	1	346	21	1	310
10	2	370	22	2	368
11	3	430	23	3	442
12	4	290	24	4	302

a. Plot the data.
b. Determine a four period centered moving average of the data and overlay it on the original data.
c. Determine the ratio-to moving average seasonal indices for the four quarters.
d. Re-plot the original data and create an overlay plot of the de-seasonalized data.

One-step ahead forecasts

5. Determine a set of one-step ahead forecasts of the data, using...

a. Three period moving average forecasts.
b. Five period moving average forecasts.
c. Four period moving average forecasts.
d. Three period weighted moving average forecasts; weights .6, .3 and .1

6. Determine a set of one-step ahead forecasts of the data, using...

a. Single exponential smoothing forecasts with $\alpha = .1$.
b. Single exponential smoothing forecasts with $\alpha = .2$.
c. Single exponential smoothing forecasts with $\alpha = .7$.

7. Using the one-step ahead forecasts of Problem 6, determine the forecasts for period 25 in... a., b., and c. above.

8. Determine the mean squared error (MSE) of each of the one-step ahead forecasts in Problem 6. Which is the better forecast? Explain.

Data with Trend

9. Listed below is a time series of 24 observations. Plot the data and determine the mean and variance of the the data set.

Period	Observation	Period	Observation
2	70	14	120
3	85	15	134
4	60	16	121
5	88	17	93
6	68	18	113
7	106	19	125
8	75	20	136
9	86	21	142
10	124	22	117
11	122	23	132
12	87	24	141

10. Determine the one-step ahead forecasts using...

a. Three period moving average
b. Weighted moving average with weights .5, .3, .2. What are the forecasts for period 25 in each case above? What is the MSE in each case above?

11. Determine the one-step ahead forecasts using..

a. Single exponential smoothing with $\alpha = 0.3$
b. What is the forecast for period 25?
c. What is the MSE of this forecasting model?

12. Determine the one-step ahead forecasts using...

a. Double exponential smoothing with $\alpha = 0.3$
b. What is the forecast for period 25?
c. What is the forecast for periods 26 to 30?
d. Determine the 95% confidence intervals of forecast for periods 25 to 30. e. What is the MSE of the forecasting model?

Linear Regression of a Time Series

13. Determine a linear regression model...

a. For forecasting periods 25 to 30.
b. Determine the 95% confidence intervals of forecast for periods 25 to 30.
c. What is the MSE of this forecasting model?

14. How does your "line of sight" forecast compare with the double exponential smoothing model?
With the linear regression model?

15. Which is the better model, double exponential smoothing or linear regression? Explain.