STAT 2011
Workshop on Data Exploration and Technical Writing

Section 2: Time Series Methods and Forecasting

Reference : David S. M., George P. M,, Layth C. A., Bruce A., The Practice of Statistics for Business and Economics, five edition, W. H. Freeman
Anderson, D. R., Sweeney, D. J., Williams, T. A. Quantitative Methods for Business, latest edition, Cengage Learning.
Tan, S. T. Applied Mathematics for the Managerial, Life, and Social Sciences, latest edition, Brooks /Cole, Cengage Learning

Dr. OUYANG Ming

2023/2024 Term 2

Quantitative Approaches to Forecasting

✓ Quantitative approach to forecasting are based on an analysis of **historical data** concerning one or more **time series**.

A **time series** is a set of observations measured at successive points in time or over successive periods of time.

- If the historical data used are restricted to past values of the series that we are trying to forecast, the procedure is called a **time series method**.

- If the historical data used involve other time series that are believed to be related to the time series that we are trying to forecast, the procedure is called a **causal method**

--------------------Historical Data--------------------------- →

----sample--→

$Y_{BEG}(Y_1)$---------------------------→Y_n---------------------------→$Y_{END}(Y_N)$

$\hat{Y}_1$ ---within-sample forecast--------→ $\hat{Y}_n$ $\hat{Y}_{n+1}$ ----Ex post forecast-----→ $\hat{Y}_N$ $\hat{Y}_{N+1}$ -→Ex ante for.

To build a forecast model for a time series Y_t, a sequence of observations is collected from the available data.

Quantitative Approaches to Forecasting

The ex post period is defined as the time period from the first observation after the end of the sample period to the most recent observation . The most important characteristic of this time period is that the availability of the current values of the time series variable Y_t.

The ex ante forecast period is defined as that time period in which no observations on the time series variable exist.

--------------------Historical Data--------------------------- →

----sample--→

$Y_{BEG}(Y_1)$--------------------------→Y_n--------------------------→$Y_{END}(Y_N)$

$\hat{Y}_1$ ---within-sample forecast--------→ $\hat{Y}_n$ $\hat{Y}_{n+1}$ ----Ex post forecast-----→ $\hat{Y}_N$ $\hat{Y}_{N+1}$ -→ Ex ante for.

To build a forecast model for a time series Y_t, a sequence of observations is collected from the available data.

Measures of Forecast Accuracy

Define the difference between the actual and the forecast values as the **forecast error**

$$e_i = Y_t - \widehat{Y}_t$$

Mean Squared Error

The average of the squared forecast errors for the historical data is calculated. The forecasting method or parameter(s) which minimize this mean squared error is then selected.

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

Mean Absolute Error

The mean of the absolute values of all forecast errors is calculated, and the forecasting method or parameter(s) which minimize this measure is selected.

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

The mean absolute error is less sensitive to individual large forecast errors than the mean squared error measure.

Example: Annual sales

An owner of a small business has been using two models to forecast annual sales during the past four years. Determine which of the two forecast models is better. The data for annual revenues and the forecasts generated by each model labeled model 1 and model 2 are given.

Model 1	Actual	Predicted	Error (e_t)	$\lvert e_t \rvert$	$(e_t)^2$
	15	15.5	-0.5	0.5	0.25
	20	20.0	0.0	0.0	0.00
	19	18.5	0.5	0.5	0.25
	23	27.0	-4	4.0	16
		MAE =	1.25	MSE =	4.125

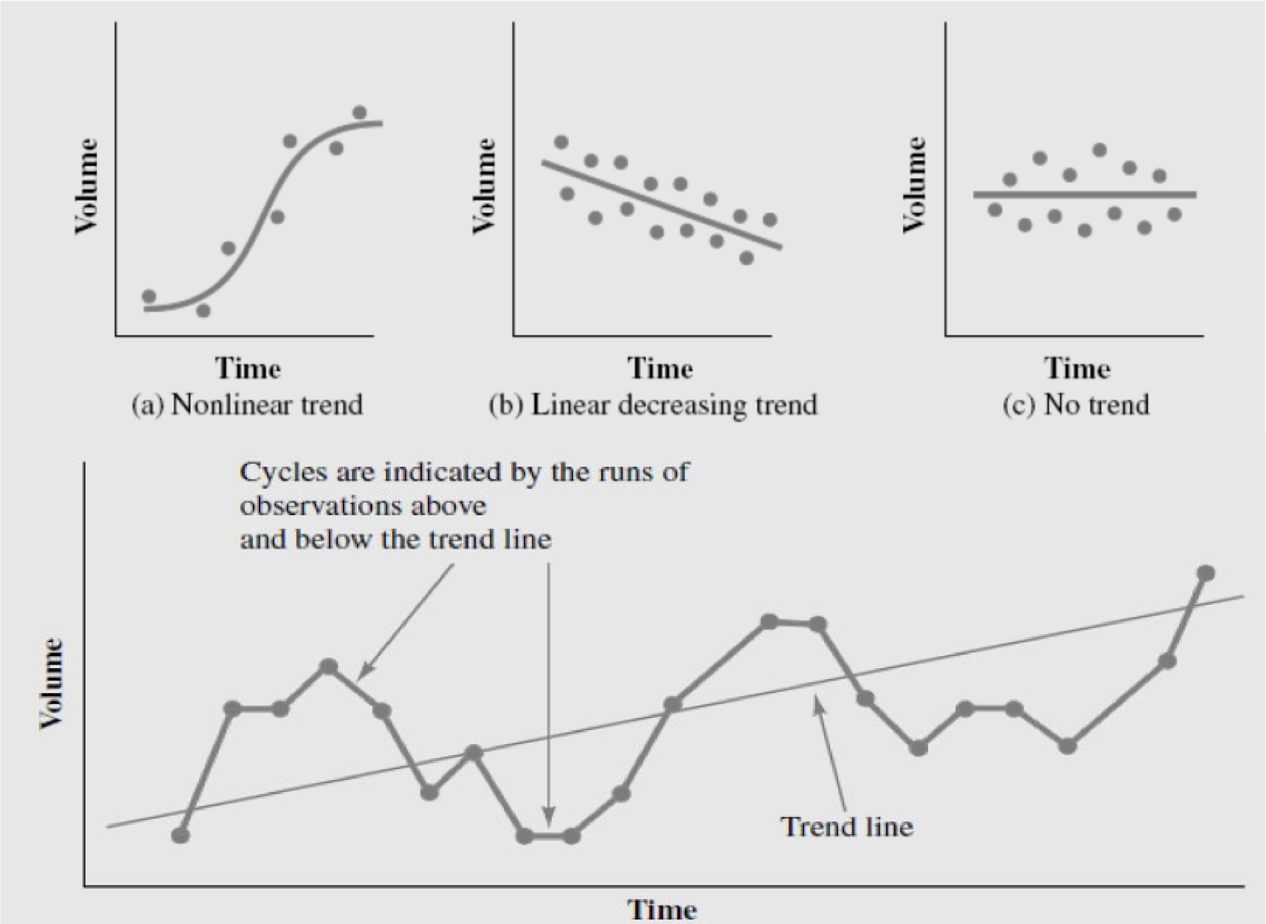
Model 1 is the better model because the forecast values are close to the actual values in all but one time period.

Model 2	Actual	Predicted	Error (e_t)	$\lvert e_t \rvert$	$(e_t)^2$
	15	14	1	1	1
	20	18	2	2	4
	19	21	-2	2	4
	23	24	-1	1	1
		MAE =	1.5	MSE =	2.5

However, the one large error produces a MSE larger than that of model 2.

Components of a Time Series

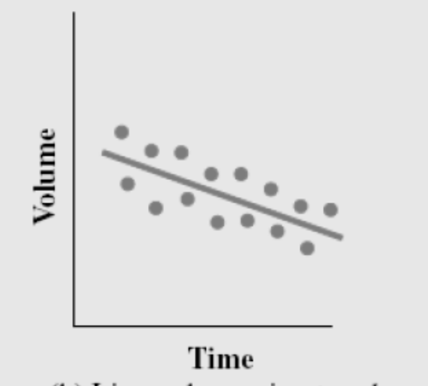
- The trend component accounts for the gradual shifting of the time series over a long period of time.
- Any regular pattern of sequences of values above and below the trend line is attributable to the cyclical component of the series.
- The seasonal component of the series accounts for regular patterns of variability within certain time periods, such as over a year.
- The irregular component of the series is caused by short-term, unanticipated and non-recurring factors that affect the values of the time series. One cannot attempt to predict its impact on the time series in advance.



Trend Projection

If a time series exhibits **a linear trend**, the method of least squares may be used to determine a trend line (projection) for future forecasts

Least squares determines the unique trend line forecast which minimizes the **mean square error** between the trend line forecasts and the actual observed values for the time series



(b) Linear decreasing trend

Compare with linear regression

Using the method of least squares, the formula for the trend projection is

$$\boxed{T_t = b_0 + b_1 t}$$

where T_t = trend forecast for time period t
b_1 = slope of the trend line
b_0 = trend line projection for time 0

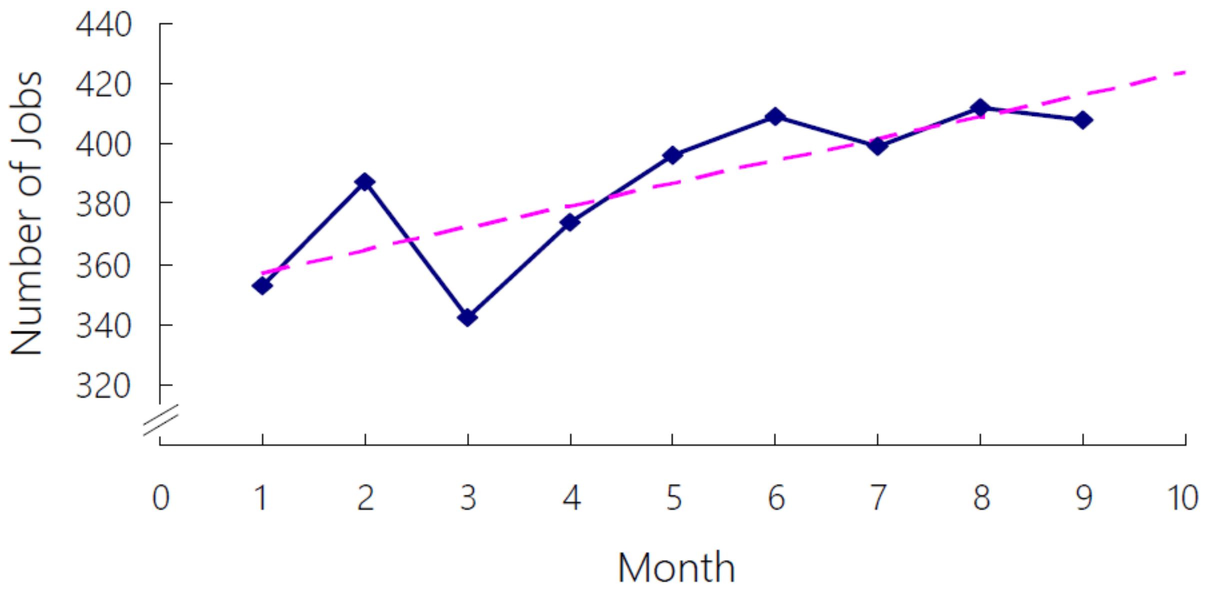
$$b_1 = \frac{n\sum tY_t - \sum t \sum Y_t}{n\sum t^2 - (\sum t)^2} \qquad b_0 = \bar{Y} - b_1 \bar{t}$$

where Y_t = observed value of the time series at time period t
$\bar{Y}$ = average of the observed values for Y_t
$\bar{t}$ = average time period for the n observations

Example: Auger’s Plumbing Service

The number of plumbing repair jobs performed by Auger's Plumbing Service in each of the last nine months is listed below. Forecast the number of repair jobs Auger's will perform in December using the least squares method

Month	Jobs	Month	Jobs
March	353	August	409
April	387	September	399
May	342	October	412
June	374	November	408
July	396		



Example: Auger's Plumbing Service

The number of plumbing repair jobs performed by Auger's Plumbing Service in each of the last nine months is listed below. Forecast the number of repair jobs Auger's will perform in December using the least squares method

Trend Projection

(month)	t	Y_t	tY_t	t^2
(Mar)	1	353	353	1
(Apr)	2	387	774	4
(May)	3	342	1026	9
(June)	4	374	1496	16
(July)	5	396	1980	25
(Aug)	6	409	2454	36
(Sep)	7	399	2793	49
(Oct)	8	412	3296	64
(Nov)	9	408	3672	81
Sum	45	3480	17844	285

$$\bar{t} = \frac{45}{9} = 5$$

$$\bar{Y} = \frac{3480}{9} = 386.667$$

$$b_1 = \frac{n\sum tY_t - \sum t \sum Y_t}{n\sum t^2 - (\sum t)^2} = \frac{9(17844) - 45(3480)}{9(285) - (45)^2} = 7.4$$

$$b_0 = \bar{Y} - b_1 \bar{t} = 386.667 - 7.4(5) = 349.667$$

$$T_{10} = b_0 + b_1 t = 349.667 + (7.4)(10) = 423.667$$

The trend component yields a forecast of 423.667 for month 10

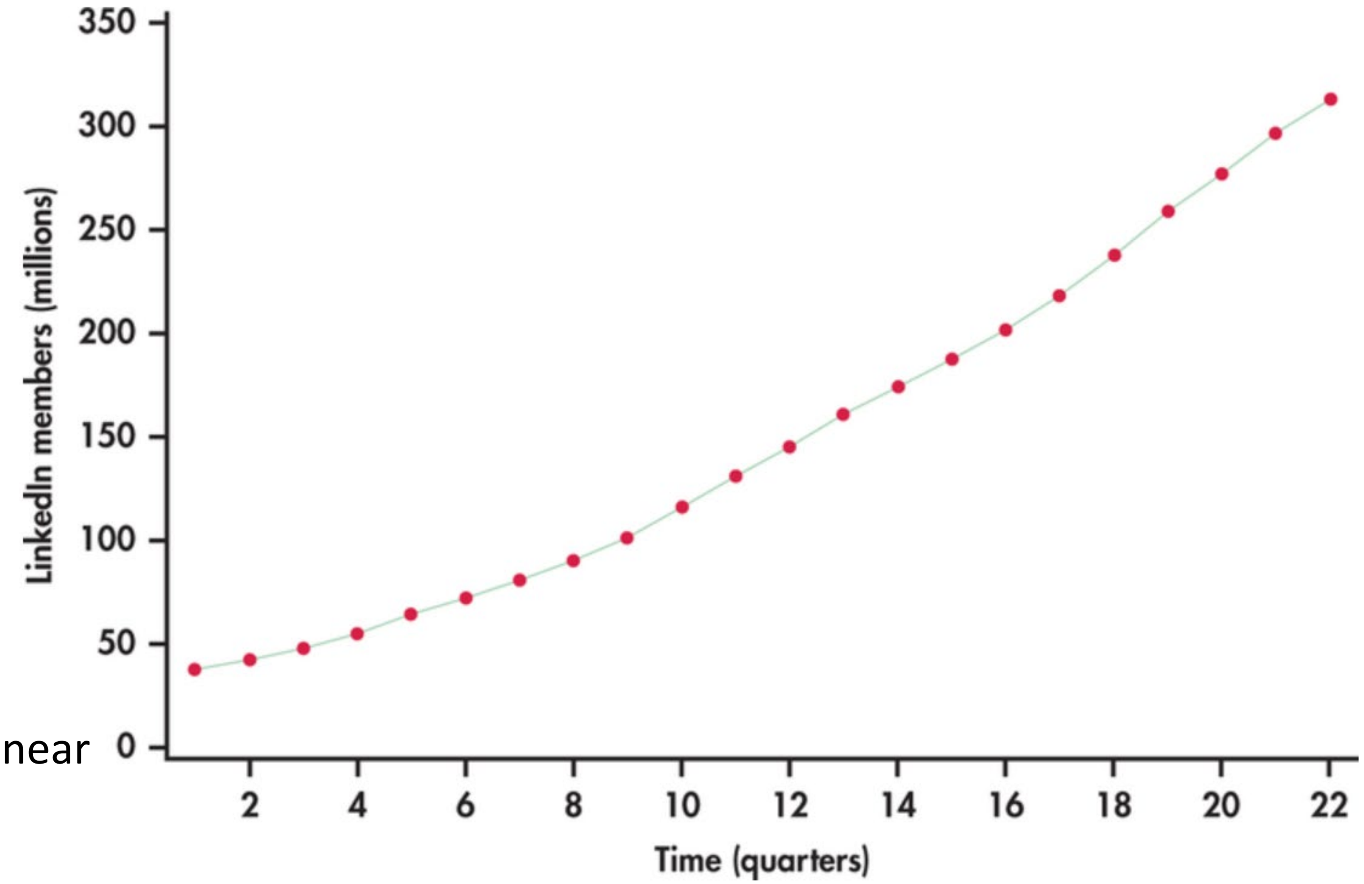
Example: LinkedIn Members

Figure shows the number of LinkedIn members (in millions) by quarter starting from the first quarter of 2009 and ending with the second quarter of 2014. The number of members are clearly increasing with time. **However, the trend appears not to be linear.**

The linear model is not able to capture the curvature in the series. One possible approach to fitting curved trends is to introduce the square of the time index to the model:

$$\boxed{T_t = b_0 + b_1 t + b_2 t^2}$$

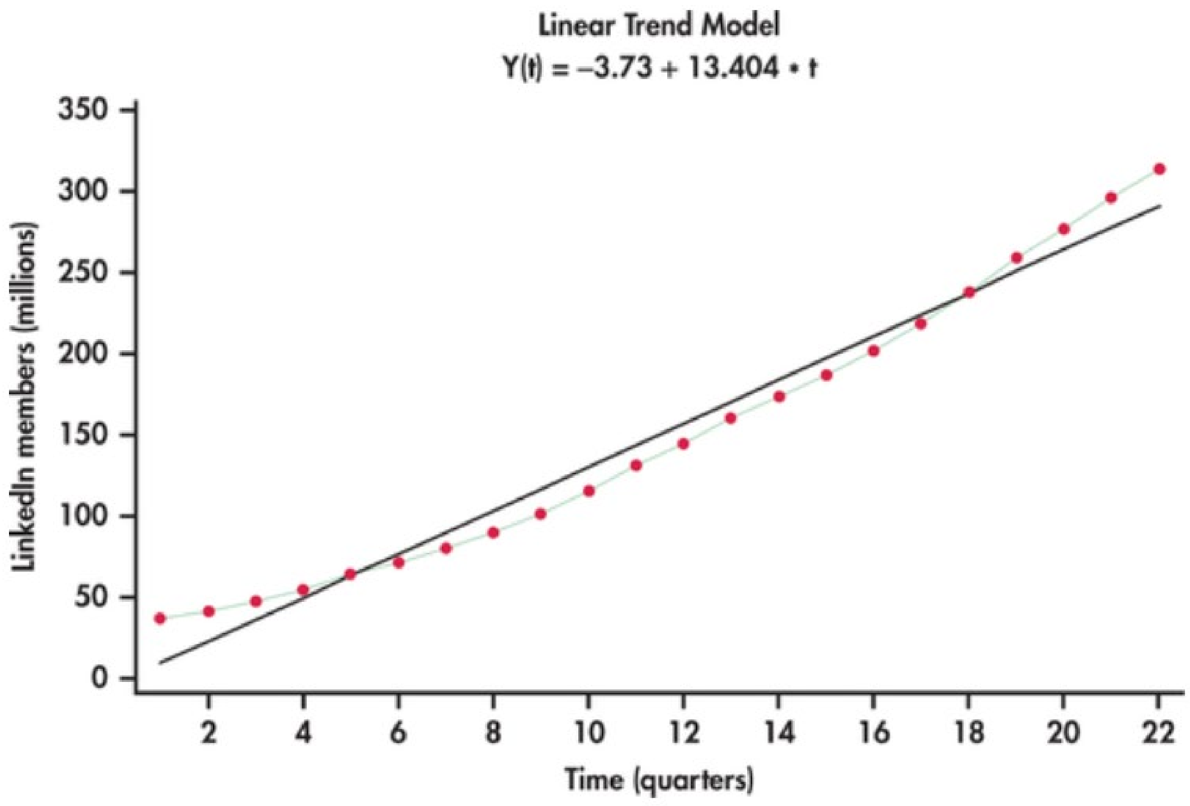
Quadratic trend model

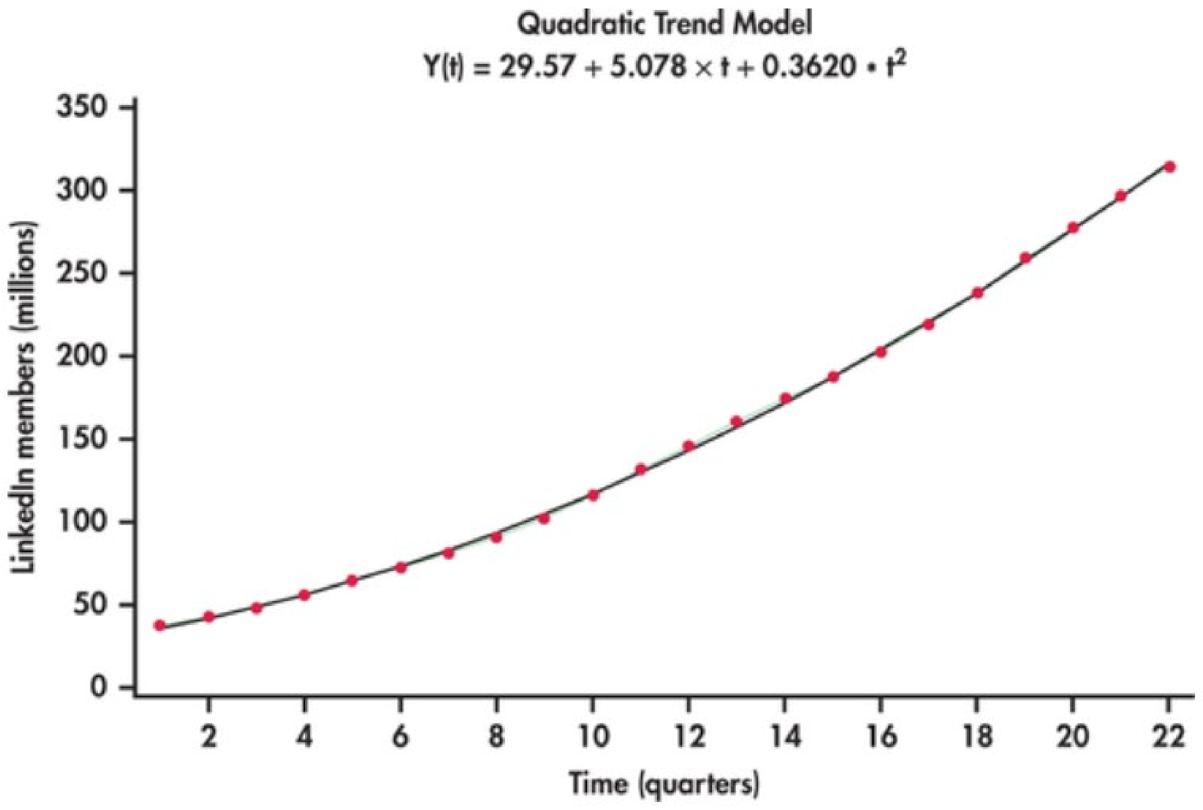


Quadratic trends can be quite flexible is adapting to a variety of curvature in practice but there can be nonlinear patterns that challenge the quadratic model

Example: LinkedIn Members

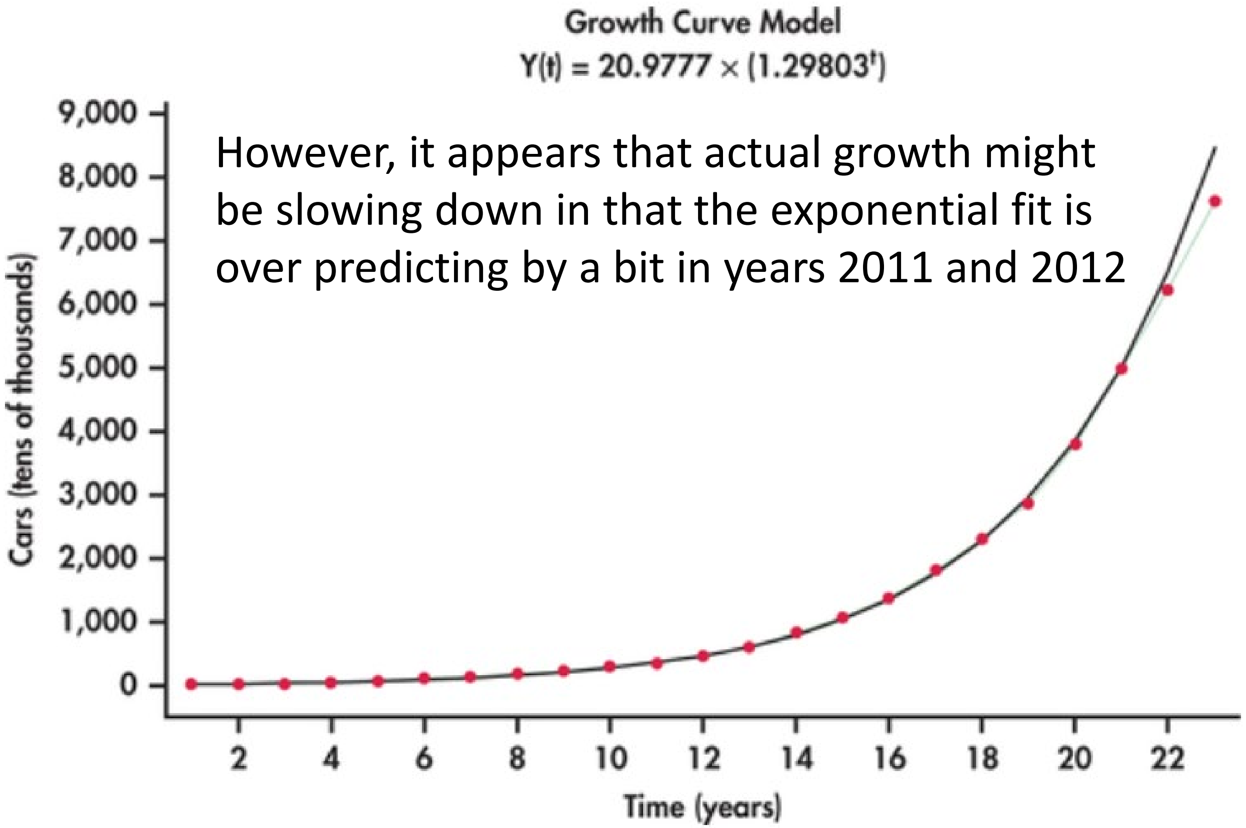
Figure shows the number of LinkedIn members (in millions) by quarter starting from the first quarter of 2009 and ending with the second quarter of 2014. The number of members are clearly increasing with time. **However, the trend appears not to be linear.**





Example: Chinese Car Ownership

China's rapid economic growth can be measured on numerous dimensions. Consider Figure, which shows the time plot of the number of passenger cars owned (in tens of thousands) in China from 1990 to 2012.



Exponential trend model

$$\boxed{T_t = b_0 * b_1{}^\wedge t}$$

$$\boxed{\log(T_t) = b_0{}' + b_1{}'t}$$

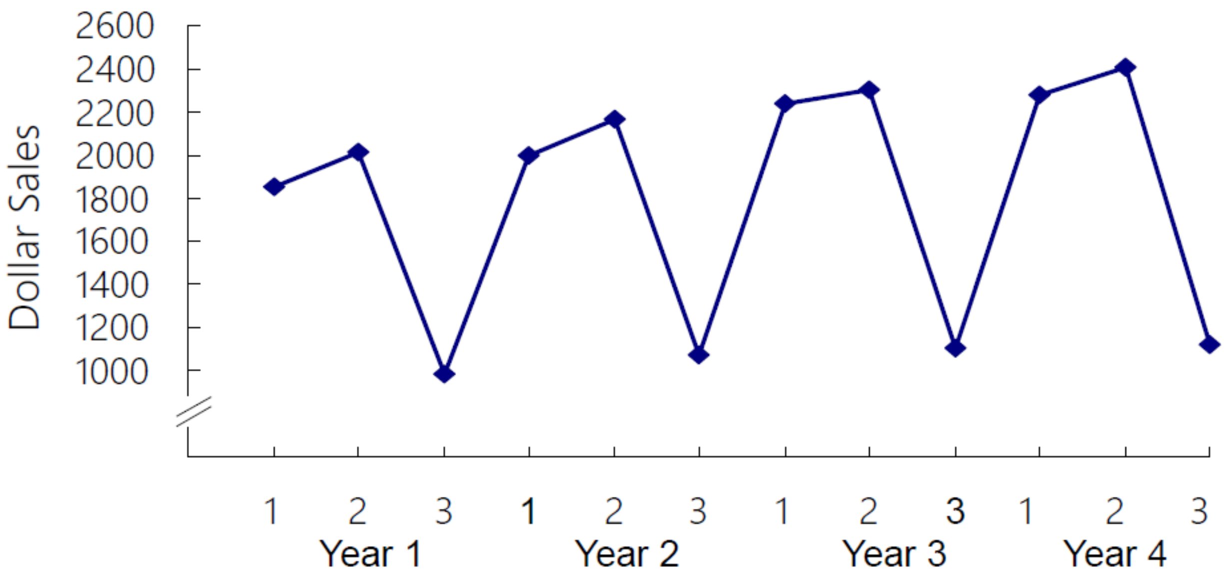
The estimated model has direct interpretation on the growth rate. As the time index t increases by one unit, we multiply the estimate for the number of cars from the previous year by 1.29803. This implies that the yearly growth rate in the number of passenger cars is estimated to be 29.8%.

Seasonal patterns/component

- Variables of economic interest are often tied to other events that repeat with regular frequency over time. Agriculture-related variables will vary with the growing and harvesting seasons. Sales data may be linked to events like regular changes in the weather, the start of the school year, and the celebration of certain holidays.

- As a result, we find a repeating pattern in the data series that relates to a particular "season," such as **month of the year, day of the week, or hour of the day**. In the applications to follow, we see that to improve the accuracy of our forecasts, we need to account for seasonal variation in the time series

Graph of seasonal sales time series

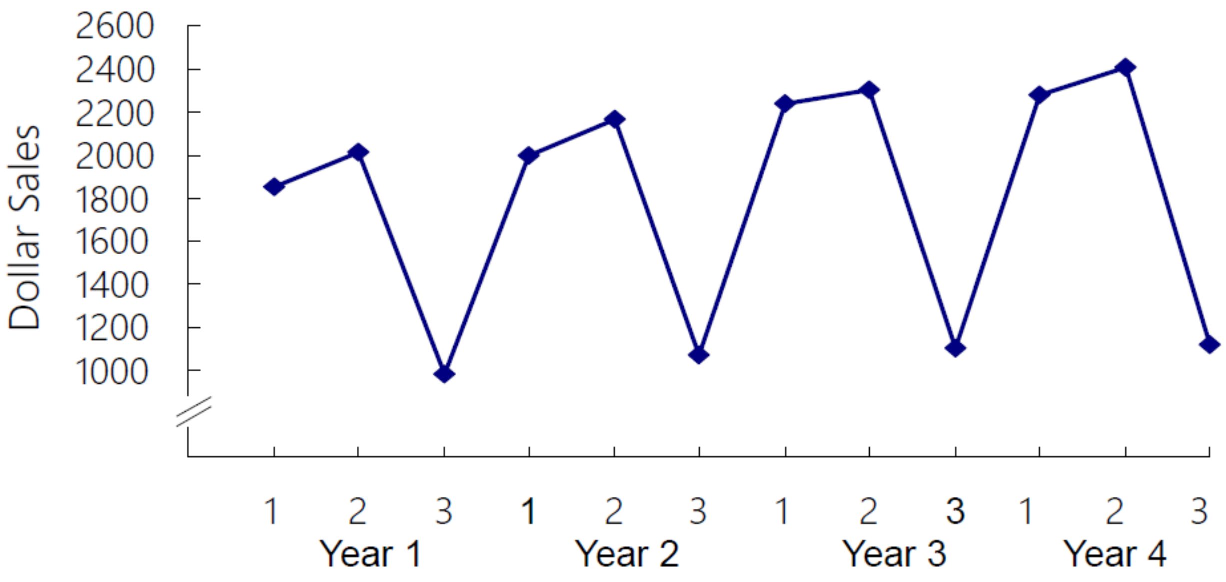


Example: Terry’s Tie Shop

Business at Terry's Tie Shop can be viewed as falling into three distinct seasons: (1) Christmas (November-December); (2) Father's Day (late May -mid-June); and (3) all other times. Average weekly sales ($) during each of the three seasons during the past four years are shown below.

	Year			
Season	1	2	3	4
1	1856	1995	2241	2280
2	2012	2168	2306	2408
3	985	1072	1105	1120

Graph of seasonal sales time series



Using indicator/dummy variables

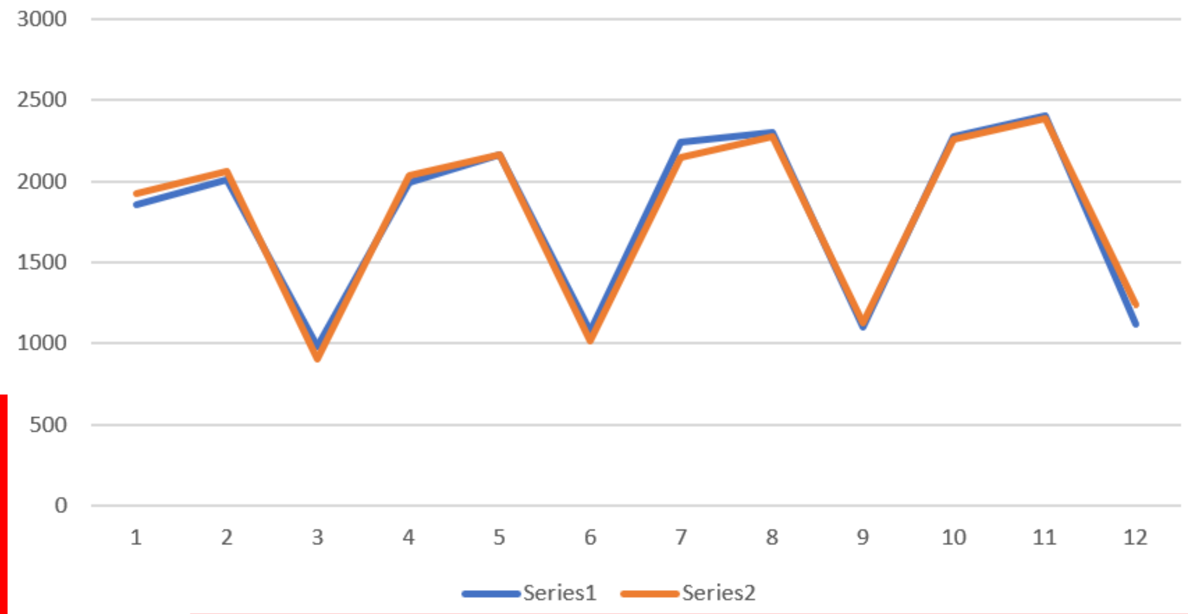
$$\boxed{T_t = b_0 + b_1 t + b_2 S_2 + b_3 S_3}$$

$$S_2 = \begin{cases} 1\ if\ the\ season\ is\ 2 \\ 0\ otherwise \end{cases} \qquad S_3 = \begin{cases} 1\ if\ the\ season\ is\ 3 \\ 0\ otherwise \end{cases}$$

Example: Terry's Tie Shop

Business at Terry's Tie Shop can be viewed as falling into three distinct seasons: (1) Christmas (November-December); (2) Father's Day (late May -mid-June); and (3) all other times. Average weekly sales ($) during each of the three seasons during the past four years are shown below.

$$T_t = 1892.43 + 36.47 * t \quad \textit{plus} \quad + 94.03 * S_2 + 1095.43 * S_3$$



Additive seasonal effects

Using indicator/dummy variables

$$T_t = b_0 + b_1 t + b_2 S_2 + b_3 S_3$$

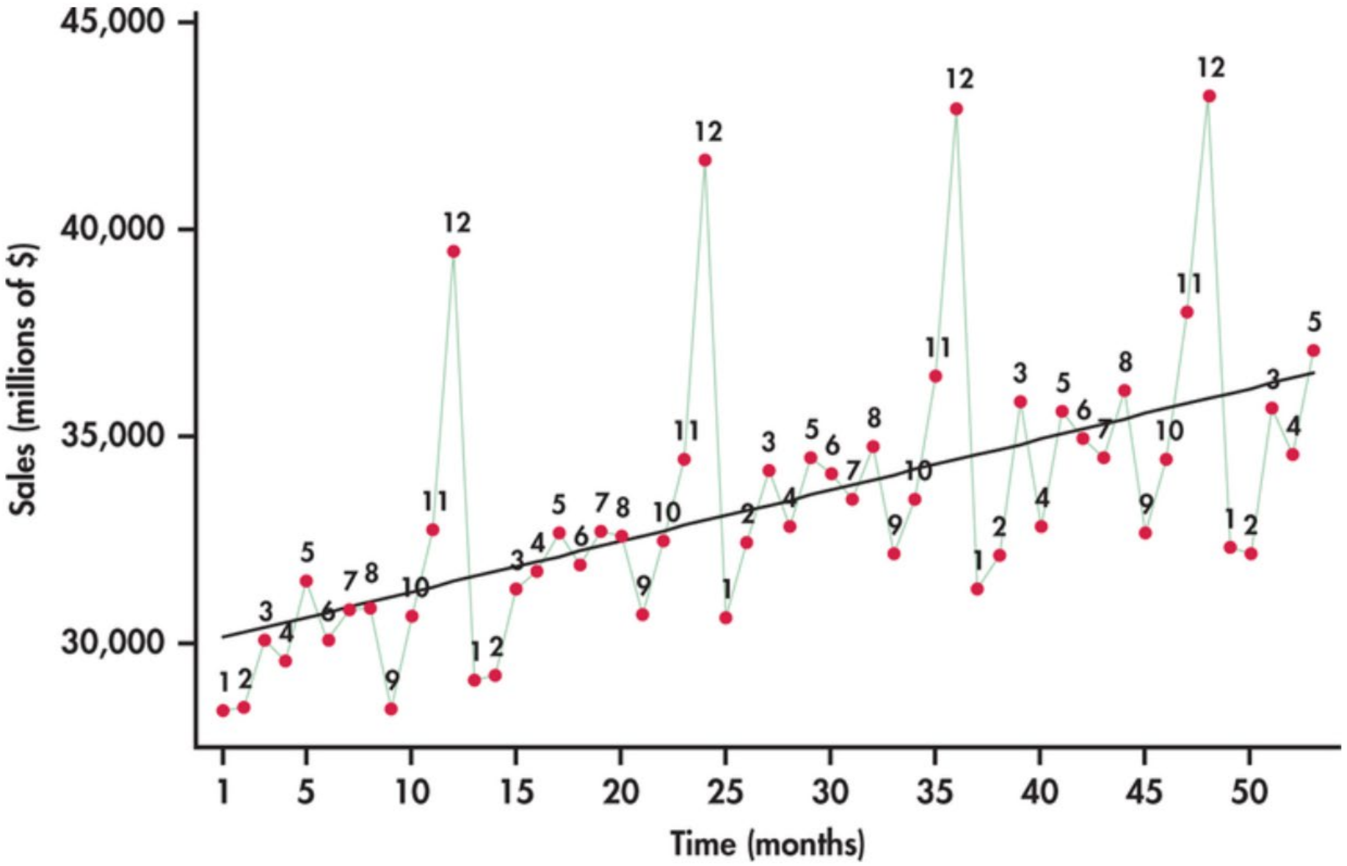
$$S_2 = \begin{cases} 1 \; if \; the \; season \; is \; 2 \\ 0 \; otherwise \end{cases} \qquad S_3 = \begin{cases} 1 \; if \; the \; season \; is \; 3 \\ 0 \; otherwise \end{cases}$$

The implication is that the average amount of increase or decrease in sales for a given month around the trend line is the *same* from year to year.

Example: Monthly warehouse Club and Superstore Sales

The Census Bureau tracks a variety of retail and service sales using the Monthly Retail Trade Survey. Consider, in particular, monthly sales (in millions of dollars) from January 2010 through May 2014 for warehouse clubs (Costco, Sam's Club) and superstores (Target and Walmart).

- Sales are **increasing** over time. The increase is reflected in the superimposed trend line fit.
- **A distinct pattern repeats itself every 12 months**: January, February, and September sales are consistently below the trend line; sales pick up in the spring months and seem to level off; and, finally, there is an initial increase in November sales followed by a more dramatic increase to a peak in December.



Example: Monthly warehouse Club and Superstore Sales

The Census Bureau tracks a variety of retail and service sales using the Monthly Retail Trade Survey. Consider, in particular, monthly sales (in millions of dollars) from January 2010 through May 2014 for warehouse clubs (Costco, Sam's Club) and superstores (Target and Walmart).

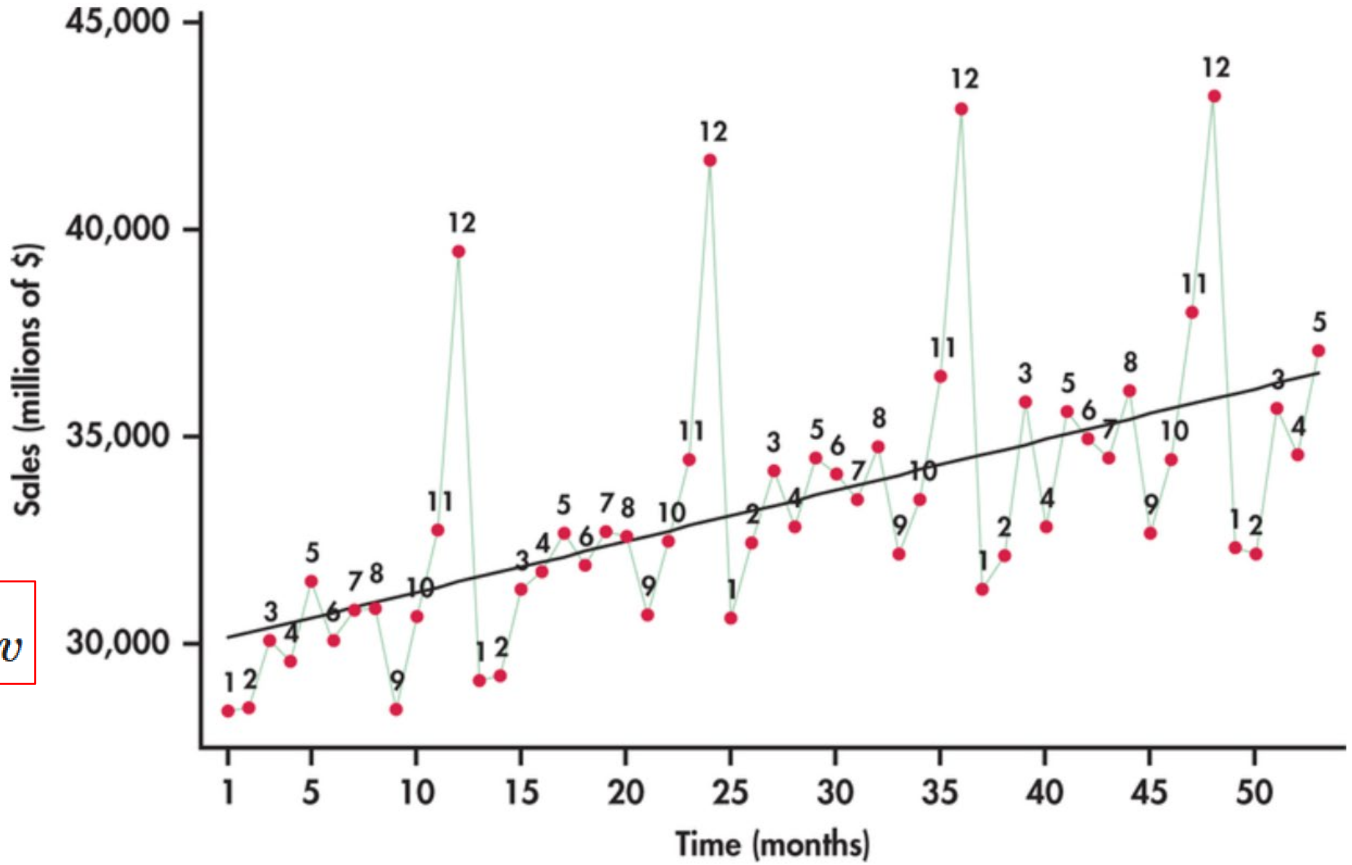
$$Jan = \begin{cases} 1 & \text{if the month is January} \\ 0 & \text{otherwise} \end{cases}$$

$$Feb = \begin{cases} 1 & \text{if the month is February} \\ 0 & \text{otherwise} \end{cases}$$

$$\vdots$$

$$Nov = \begin{cases} 1 & \text{if the month is November} \\ 0 & \text{otherwise} \end{cases}$$

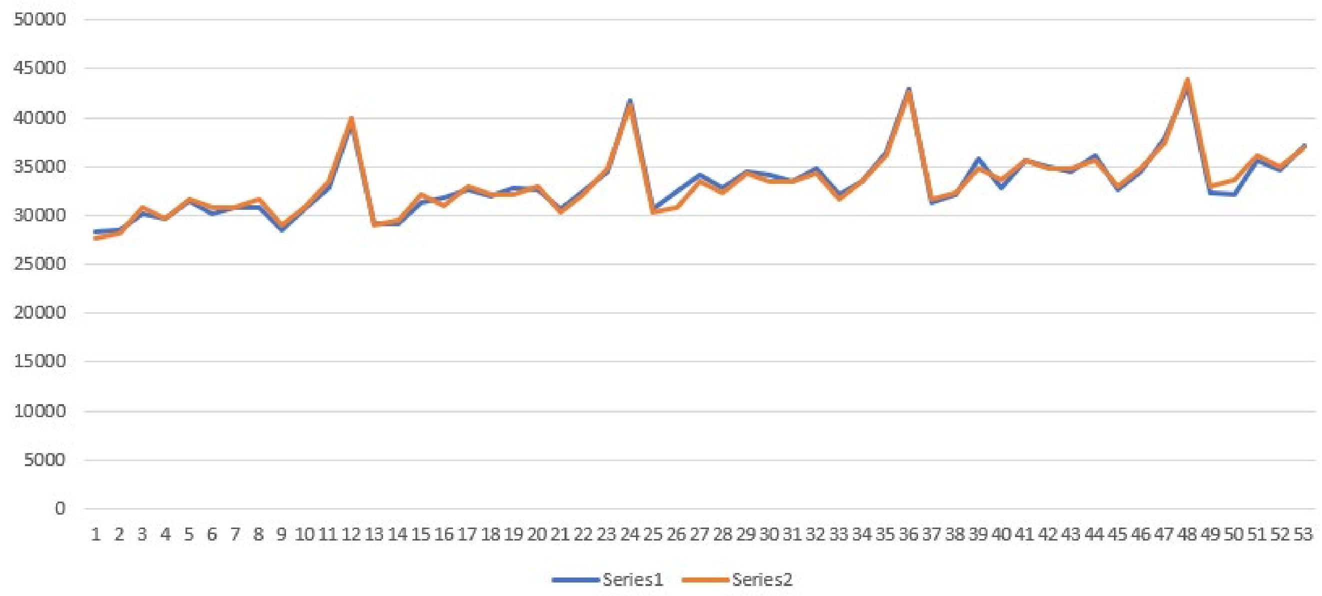
$$\boxed{\text{SALES}_t = \beta_0 + \beta_1 t + \beta_2 Jan + \cdots + \beta_{12} Nov}$$



Example: Monthly warehouse Club and Superstore Sales

The Census Bureau tracks a variety of retail and service sales using the Monthly Retail Trade Survey. Consider, in particular, monthly sales (in millions of dollars) from January 2010 through May 2014 for warehouse clubs (Costco, Sam's Club) and superstores (Target and Walmart).

Notice from the regression output that all the monthly coefficients are negative. When all of these 0's are substituted into the model, we obtain a baseline trend model fit for the Decembers. Each of the other months have an estimated trend line a certain amount below December depending on the magnitude of the month's coefficient.

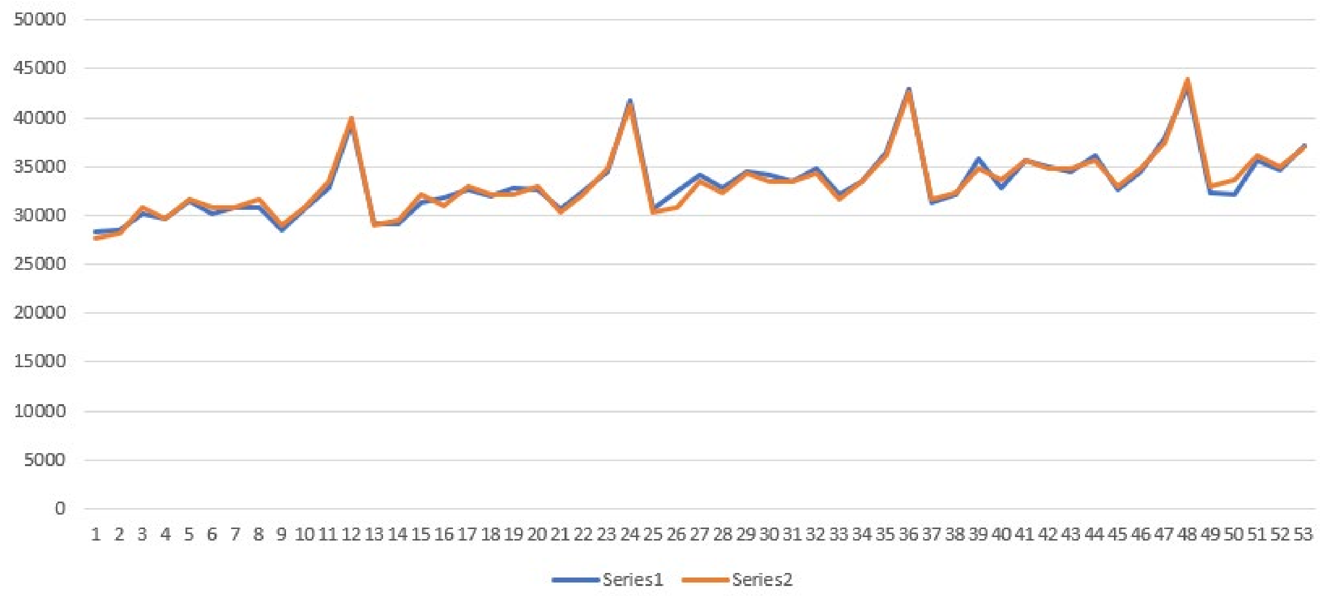


-6306.9	-8864.6	-10530	-7816.7	-8410.4	-8410.4	-7448.2	-9309.6	-8081.8	-10503	-10940	110.814	38512.6

Example: Monthly warehouse Club and Superstore Sales

The Census Bureau tracks a variety of retail and service sales using the Monthly Retail Trade Survey. Consider, in particular, monthly sales (in millions of dollars) from January 2010 through May 2014 for warehouse clubs (Costco, Sam's Club) and superstores (Target and Walmart).

Notice from the regression output that all the monthly coefficients are negative. When all of these 0's are substituted into the model, we obtain a baseline trend model fit for the Decembers. Each of the other months have an estimated trend line a certain amount below December depending on the magnitude of the month's coefficient.



-6306.9	-8864.6	-10530	-7816.7	-8410.4	-8410.4	-7448.2	-9309.6	-8081.8	-10503	-10940	110.814	38512.6

Example: Amazon Sales

We noted that sales is increasing at a greater rate over time. This implies the need for a nonlinear trend model. We also noted that the fourth-quarter seasonal surge increases over time.

It is evident that an additive seasonal model is not appropriate for modeling the Amazon series. Instead, we need to consider the situation in which each particular season is some proportion of the trend.

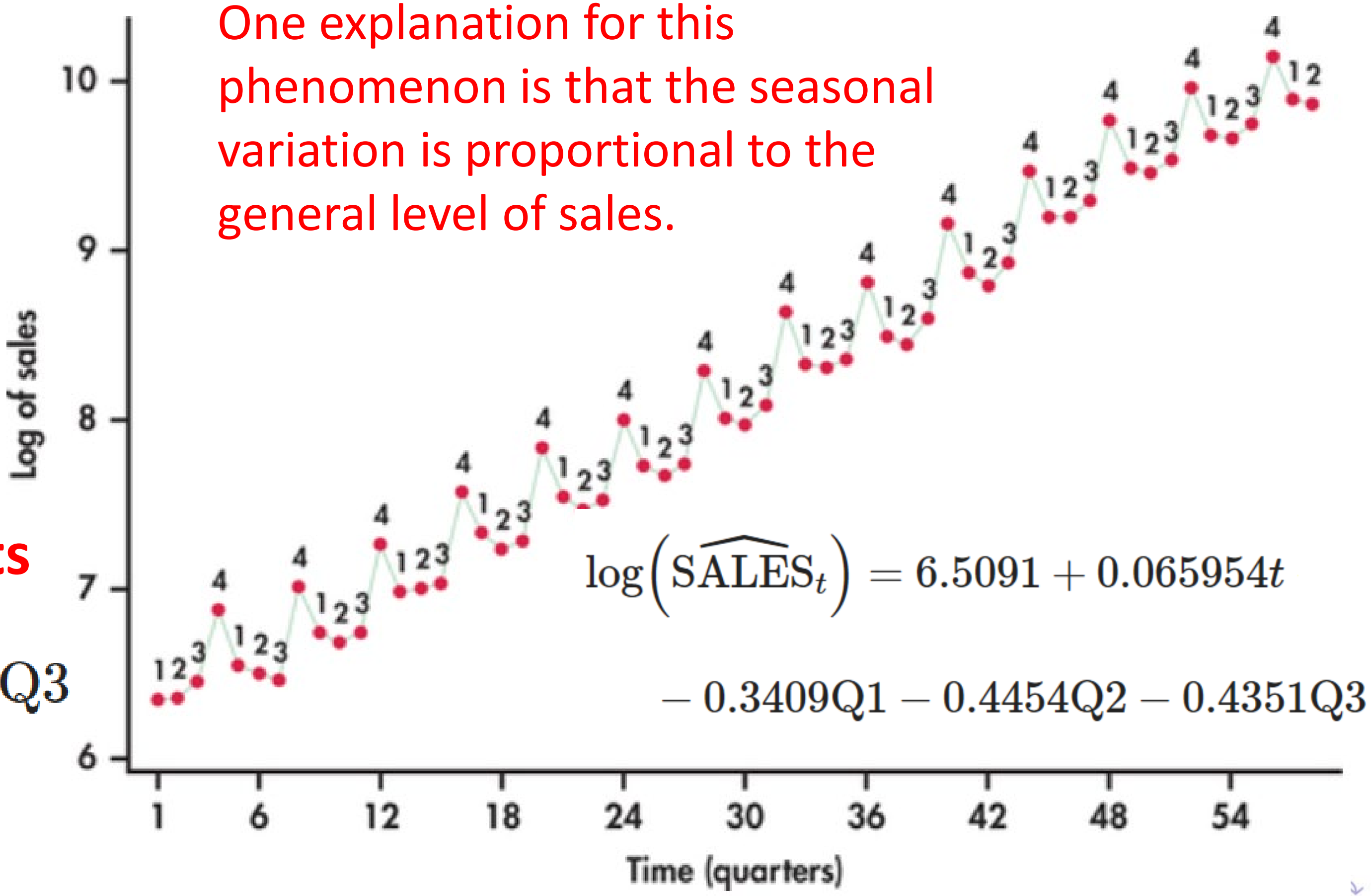
$$\boxed{\hat{y} = \mathrm{TREND} \times \mathrm{SEASON}}$$

Multiplicative seasonal effects

$$\log(\mathrm{SALES}_t) = \beta_0 + \beta_1 t + \beta_2 \mathrm{Q1} + \beta_3 \mathrm{Q2} + \beta_4 \mathrm{Q3}$$

Q1, Q2, and Q3 are indicator variables for quarters 1, 2, and 3

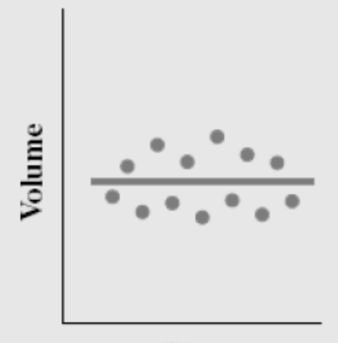
One explanation for this phenomenon is that the seasonal variation is proportional to the general level of sales.



$$\log\left(\widehat{\mathrm{SALES}_t}\right) = 6.5091 + 0.065954t - 0.3409\mathrm{Q1} - 0.4454\mathrm{Q2} - 0.4351\mathrm{Q3}$$

Smoothing Methods

In cases in which the time series is **fairly stable** and has no significant trend, seasonal, or cyclical effects, one can use smoothing methods for forecasting. These methods can average out the irregular components of the time series.



(c) No trend

Common smoothing methods are:

- Moving averages
- Centered moving averages
- Weighted moving averages
- Exponential smoothing
- Exponential smoothing DOUBLE/TRIPLE

Moving Average Method

The moving average model uses the average of the last/most recent k values of the time series as the forecast for time period t.

$$\hat{y_t} = \frac{1}{k}(y_{t-1} + y_{t-2} + \cdots + y_{t-k})$$

The number of preceding values included in the moving average is called the **span** of the moving average.

As a general rule, larger spans smooth the time series more than smaller spans by averaging many ups and downs in each calculation. Smaller spans tend to follow the ups and downs of the time series.

Example: Rosco Drugs

Sales of Comfort brand headache medicine for the past ten weeks at Rosco Drugs are shown below. If Rosco Drugs uses a 3-period moving average to forecast sales, determine the MSE. What is the forecast for Week 11?

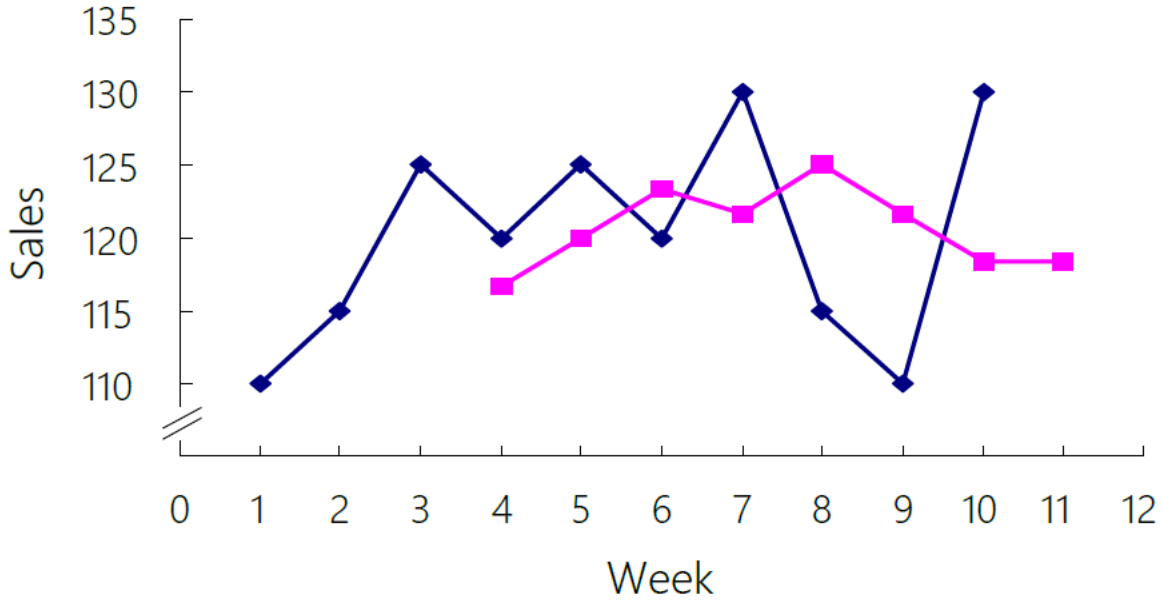
Week	Sales	Week	Sales
1	110	6	120
2	115	7	130
3	125	8	115
4	120	9	110
5	125	10	130

Example: Rosco Drugs

Sales of Comfort brand headache medicine for the past ten weeks at Rosco Drugs are shown below. If Rosco Drugs uses a 3-period moving average to forecast sales, determine the MSE. What is the forecast for Week 11?

Week	Time Series Values	Moving Average Forecast	Forecast Error	Squared Forecast Error
1	110			
2	115			
3	125			
4	120	116.67	3.33	11.11
5	125	120.00	5.00	25.00
6	120	123.33	- 3.33	11.11
7	130	121.67	8.33	69.44
8	115	125.00	-10.00	100.00
9	110	121.67	-11.67	136.11
10	130	118.33	11.67	136.11
		Mean	0.48	69.84

MSE = 69.84 and the forecast for Week 11 is 118.33



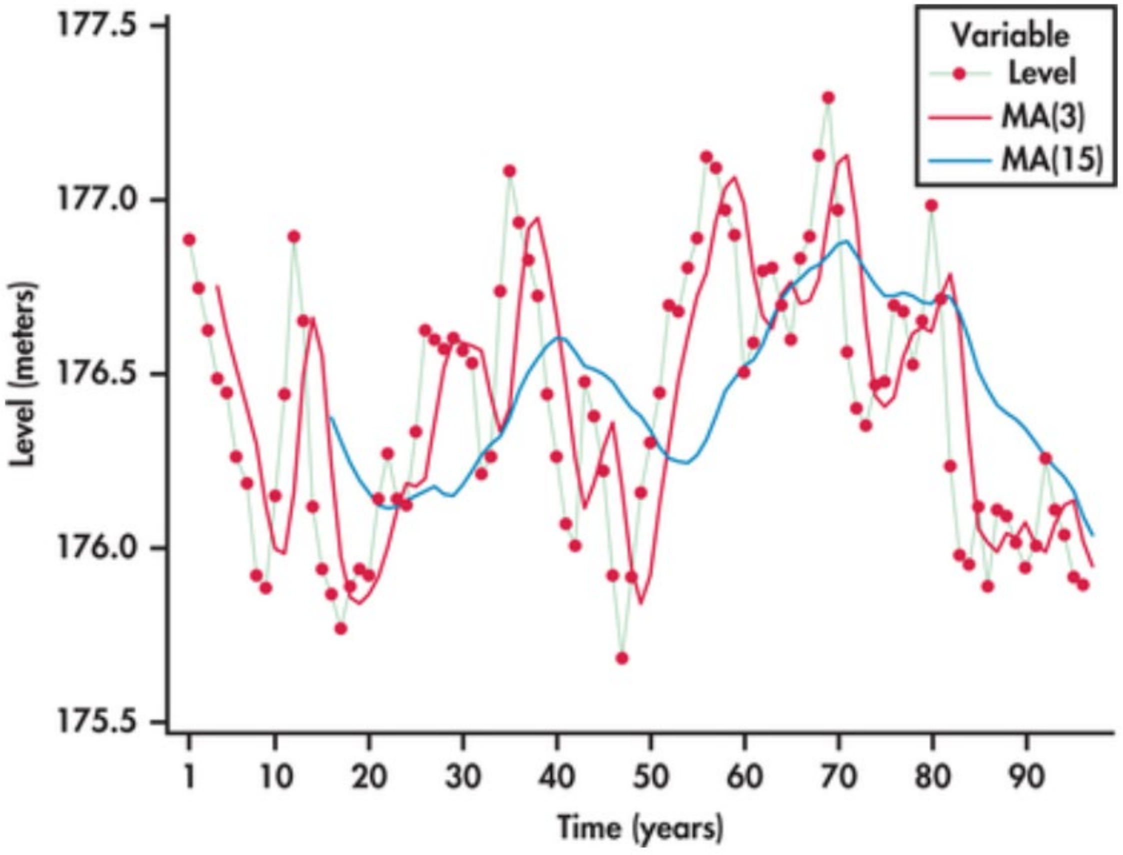
Example: Great Lakes Water Levels

Consider again the annual average water levels of Lakes Michigan and Huron

$$\hat{y}_t = \frac{1}{3}(y_{t-1} + y_{t-2} + y_{t-3})$$

$$\hat{y}_t = \frac{1}{15}(y_{t-1} + y_{t-2} + \cdots + y_{t-15})$$

The 15-year moving averages provide a long-term perspective of the cyclic movements of the lake levels. The 3-year moving averages are better able to follow the larger ups and downs while smoothing the smaller changes in the time series.

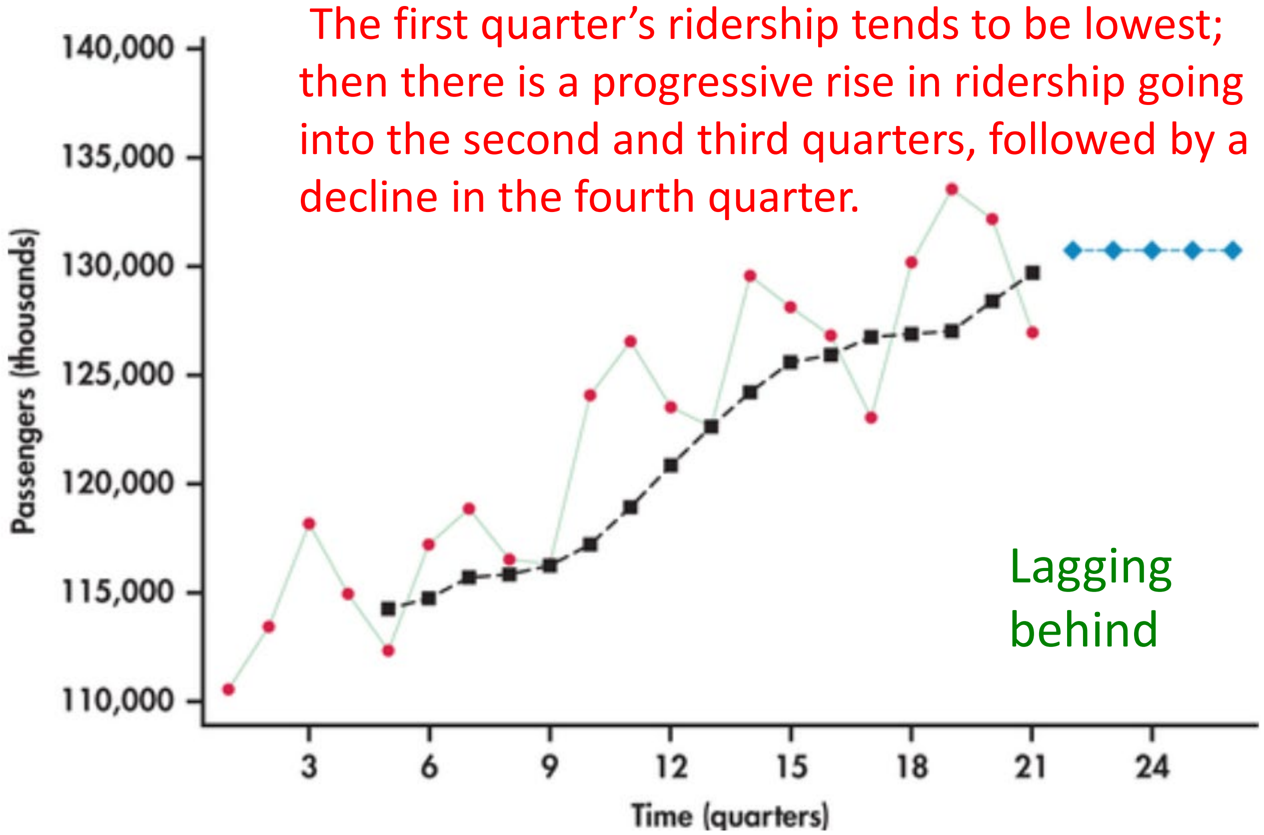


Example: Light Rail Usage

Quarterly number of U.S. passengers (in thousands) using light rail as a mode of transportation. The series begins with the first quarter of 2009 and ends with the first quarter of 2014

$$\hat{y}_t = \frac{1}{4}(y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4})$$

The moving averages are a smoothed-out version of the original time series, reflecting only the general trending in the series, which is upward



Centered Moving Average Method

The centered moving average method consists of computing an average of n periods' data and **associating it with the midpoint of the periods**. For example, the average for periods 5, 6, and 7 is associated with period 6. This methodology is useful **in the process of computing season indexes**

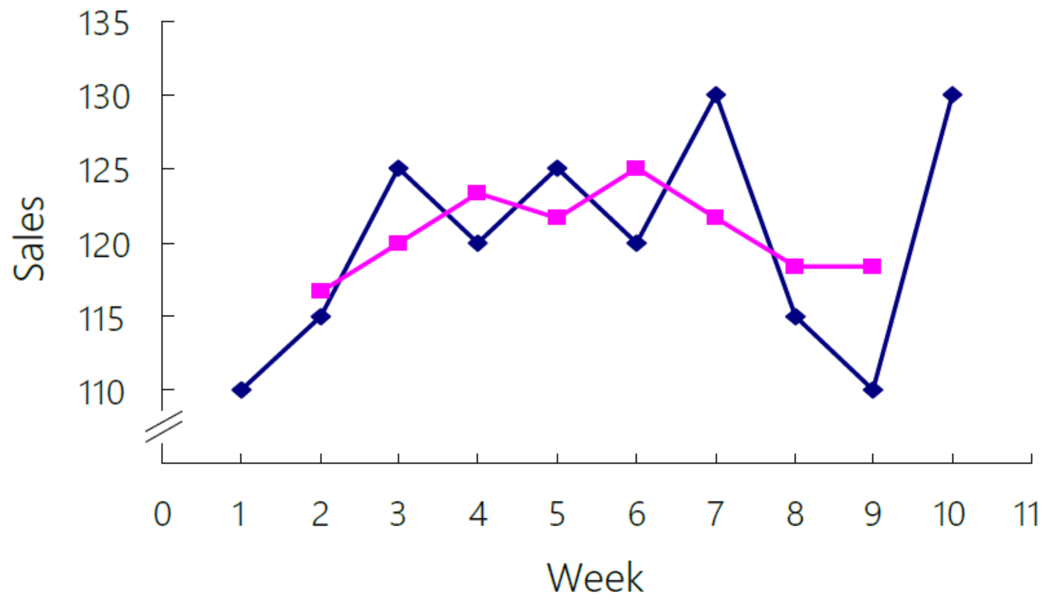
Week	Time Series Values	Moving Average Forecast	Forecast Error	Squared Forecast Error
1	110			
2	115			
3	125			
4	120	116.67	3.33	11.11
5	125	120.00	5.00	25.00
6	120	123.33	- 3.33	11.11
7	130	121.67	8.33	69.44
8	115	125.00	-10.00	100.00
9	110	121.67	-11.67	136.11
10	130	118.33	11.67	136.11
		Mean	0.48	69.84

MSE = 69.84 and the forecast for Week 11 is 118.33

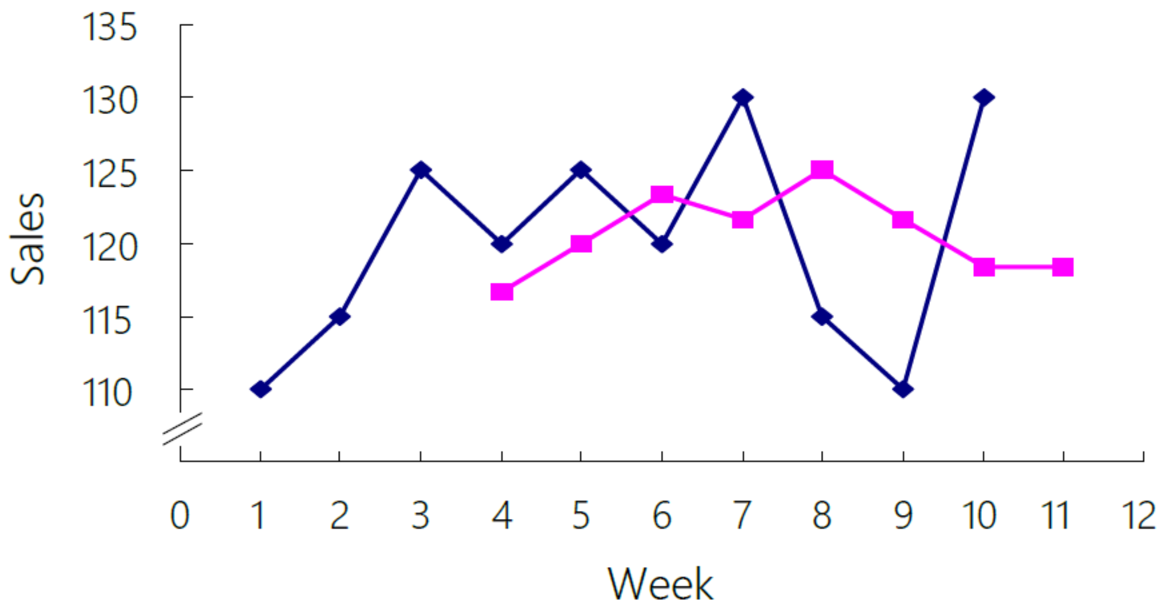
Week	Time Series Values	Centered Moving Average Forecast	Forecast Error	Squared Forecast Error
1	110			
2	115	116.67	-1.67	2.78
3	125	120.00	5.00	25.00
4	120	123.33	-3.33	11.11
5	125	121.67	3.33	11.11
6	120	125.00	-5.00	25.00
7	130	121.67	8.33	69.44
8	115	118.33	-3.33	11.11
9	110	118.33	-8.33	69.44
10	130			
		Mean	-0.63	28.13

Example: Rosco Drugs

Graph of sales time series and centered moving average forecast



The approach utilizes the moving averages to provide a baseline for the general level of the series, not to provide forecasts. Instead of projecting the moving averages into the future, we use them as a summary of the past



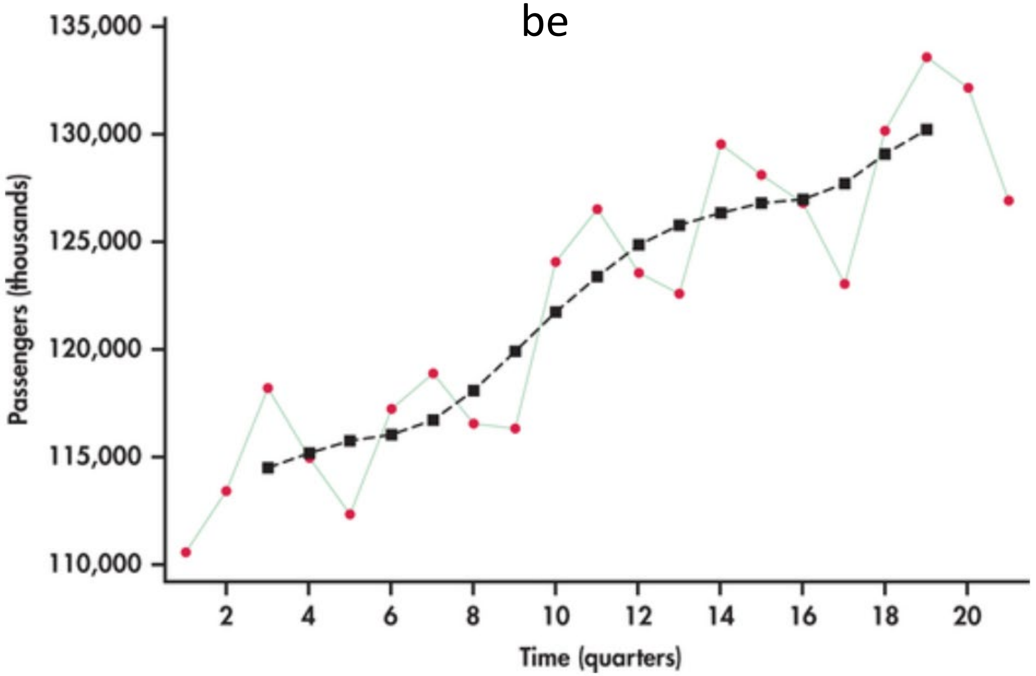
Example: Light Rail Usage and Centered Moving Averages

Year	Quarter	t	Passengers		CMA0.5	CMA
2009	1	1	110569			
2009	2	2	113433	#N/A	114279.3	
2009	3	3	118183	#N/A	114721.3	114500.25
2009	4	4	114932	#N/A	115669	115195.125
2010	1	5	112337	114279.25	115839	115754
2010	2	6	117224	114721.25	116244.5	116041.75
2010	3	7	118863	115669	117232	116738.25
2010	4	8	116554	115839	118945.3	118088.625

$$\frac{y_1 + y_2 + y_3 + y_4}{4}$$

$$\frac{y_2 + y_3 + y_4 + y_5}{4}$$

These averages are shifted back in time to provide an estimate of where the level of the process was as opposed to a forecast of where the process might be



If we were to use this average as a forecast, then it would forecast period 5. Instead, we look at the average as representing the past level of the time series.

There is another approach for estimating the seasonal effect by using CMA and without the use of regression.

Multiplicative seasonal model $\hat{y} = \text{TREND} \times \text{SEASON}$

Example: Light Rail Usage and Centered Moving Averages

seasonal ratio

Year	Quarter	t	Passengers		CMA0.5	CMA	Season	Average	Center	
2009	1	1	110569						0.969378	114061.7
2009	2	2	113433	#N/A	114279.3				1.015514	111700.1
2009	3	3	118183	#N/A	114721.3	114500.25	1.032164	1.022349	1.022086	115629.3
2009	4	4	114932	#N/A	115669	115195.125	0.997716	0.993278	0.993022	115739.7
2010	1	5	112337	114279.25	115839	115754	0.970481	0.969628	0.969378	115885.6
2010	2	6	117224	114721.25	116244.5	116041.75	1.010188	1.015776	1.015514	115433.2
2010	3	7	118863	115669	117232	116738.25				
2010	4	8	116554	115839	118945.3	118088.625				

Scale the seasonal factors: divide each seasonal factor by the average of the seasonal factors

$$\frac{y}{\mathrm{TREND}} = \mathrm{SEASON}$$

seasonal ratio

With seasonal ratios in hand, the next step is to seasonally adjust the series so that we can estimate the overall trend for forecasting purposes.

There is another approach for estimating the seasonal effect by using CMA and without the use of regression.

Multiplicative seasonal model $\hat{y} = \mathrm{TREND} \times \mathrm{SEASON}$

Weighted Moving Average Method

In the weighted moving average method for computing the average of the most recent k periods, the more recent observations are typically given more weight than older observations. For convenience, the weights usually sum to 1.

$$\hat{y}_t = \frac{a_1}{k} y_{t-1} + \frac{a_2}{k} y_{t-2} + \cdots + \frac{a_k}{k} y_{t-k} \qquad \sum_{i=1}^{k} a_i = k$$

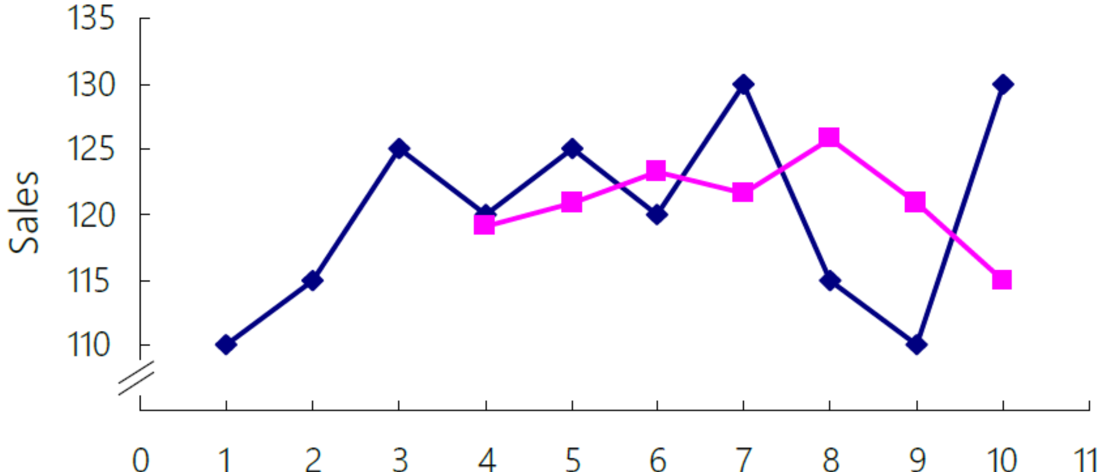
For instance, consider the most recent observation receiving a weight three times as great as that given the oldest observation, and the next observation receiving a weight twice as great as the oldest.

$$\hat{y}_t = \frac{3}{6} y_{t-1} + \frac{2}{6} y_{t-2} + \frac{1}{6} y_{t-3}$$

Weighted Moving Average Method

Week	Time Series Values	Weighted Moving Average Forecast	Forecast Error	Squared Forecast Error
1	110			
2	115			
3	125			
4	120	119.17	0.83	0.69
5	125	120.83	4.17	17.36
6	120	123.33	- 3.33	11.11
7	130	121.67	8.33	69.44
8	115	125.83	-10.83	117.36
9	110	120.83	-10.83	117.36
10	130	115.00	15.00	225.00
		Mean	0.48	79.76

Graph of sales time series and weighted moving average forecast



In many settings, the current value of a time series depends more on the most recent value and less on past values. In many settings, the current value of a time series depends more on the most recent value and less on past values.

Exponential Smoothing

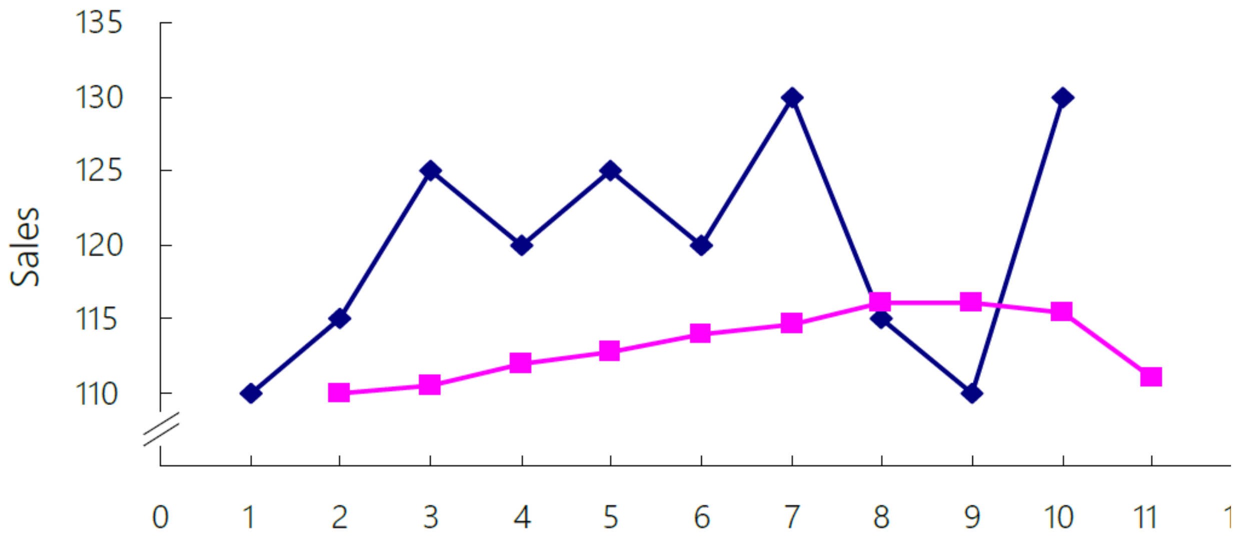
Using exponential smoothing, the forecast for the next period is equal to the forecast for the current period plus a proportion (α) of the forecast error in the current period.

$$\widehat{y_t} = \alpha y_{t-1} + (1-\alpha)\widehat{y_{t-1}} \qquad 0 \le \alpha \le 1$$
$$= \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + \cdots$$
$$= \alpha[y_{t-1} + (1-\alpha) y_{t-2} + (1-\alpha)^2 y_{t-3} + \cdots$$

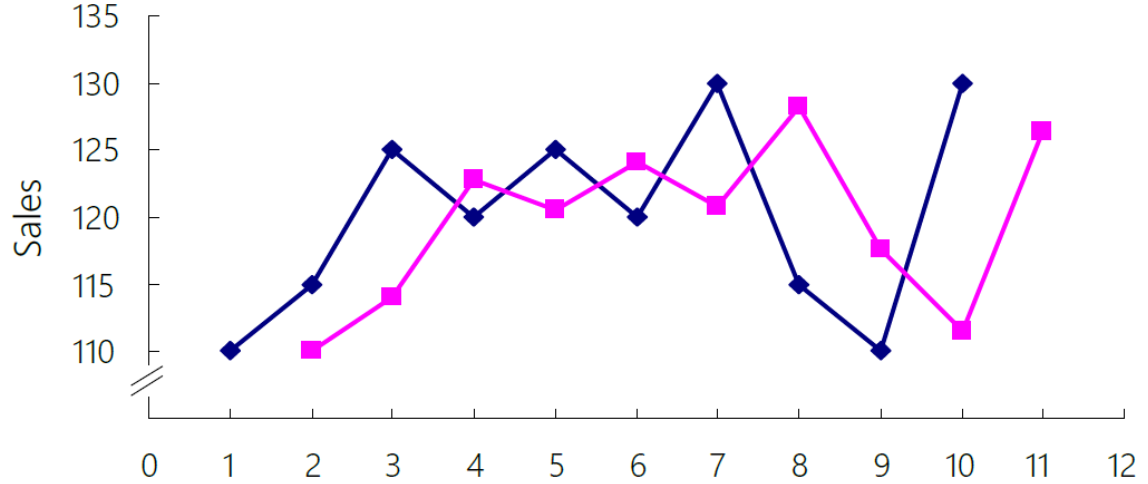
Smaller values of α correspond to greater smoothing of movements in the time series. Larger values of α put most of the weight on the most recent observed value, so the forecasts tend to be close to the most recent movement in the series.

Example: Rosco Drugs

Graph of actual and forecast sales time series with smoothing constant $\alpha = 0.1$



Graph of actual and forecast sales time series with smoothing constant $\alpha = 0.8$



Example: Rosco Drugs

		α = 0.1		α = 0.8	
Week	Y_t	F_t	$(Y_t - F_t)^2$	F_t	$(Y_t - F_t)^2$
1	110				
2	115	110.00	25.00	110.00	25.00
3	125	110.50	210.25	114.00	121.00
4	120	111.95	64.80	122.80	7.84
5	125	112.76	149.94	120.56	19.71
6	120	113.98	36.25	124.11	16.91
7	130	114.58	237.73	120.82	84.23
8	115	116.12	1.26	128.16	173.30
9	110	116.01	36.12	117.63	58.26
10	130	115.41	212.87	111.53	341.27
		Sum	974.22	Sum	847.52
		MSE	108.25	MSE	94.17

- **Exponential Triple Smoothing (ETS)**

FORECAST.ETS

FORECAST.ETS.CONFINT

FORECAST.ETS.SEASONALITY

FORECAST.ETS.STAT

With Trend and Season

Exponential Smoothing DOUBLE

$$s_t = \alpha y_t + (1-\alpha)(s_{t-1} + b_{t-1})$$

$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$

trend smoothing factor

$$\widehat{y_{t+1}} = s_t + b_t$$

Exponential Smoothing TRIPLE

$$s_t = \alpha(y_t - c_{t-L}) + (1-\alpha)(s_{t-1} + b_{t-1})$$

a cycle of seasonal change of length

$$c_t = \gamma(y_t - s_{t-1} - b_{t-1}) + (1-\gamma)c_{t-L}$$

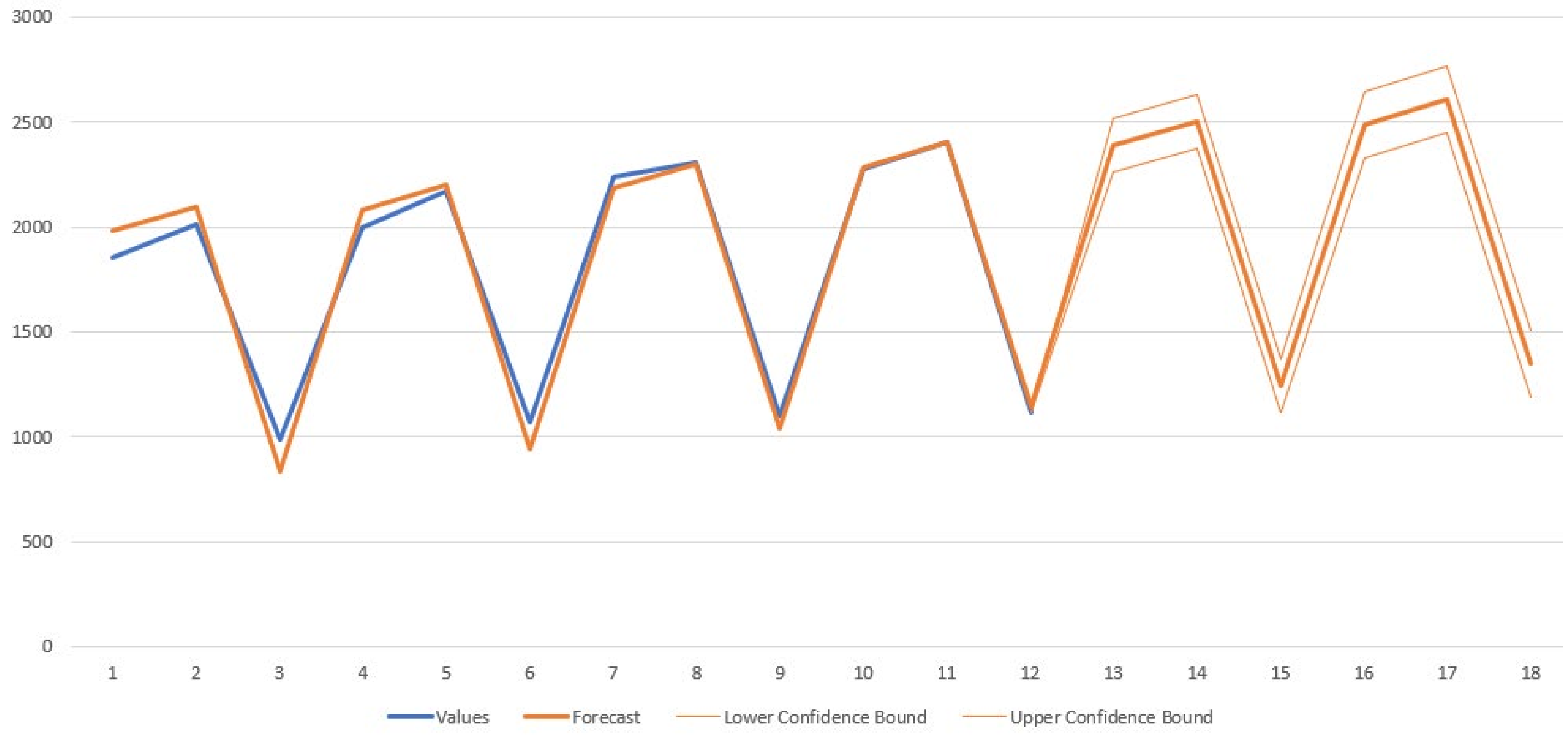
$$\widehat{y_{t+1}} = s_t + b_t + c_{t-L+1}$$

seasonal change smoothing factor

Example: Terry’s Tie Shop

Business at Terry's Tie Shop can be viewed as falling into three distinct seasons: (1) Christmas (November-December); (2) Father's Day (late May -mid-June); and (3) all other times. Average weekly sales ($) during each of the three seasons during the past four years are shown below.

Statistic	Value
Alpha	0.00
Beta	0.00
Gamma	0.75
MASE	0.08
SMAPE	0.04
MAE	52.81
RMSE	64.18



Extension

- Autoregressive-based models/Lag Model

First-Order Autoregressive Model

A **first-order autoregressive model** specifies a linear relationship between successive values of the time series. The shorthand for this model is AR(1), and the equation is

$$y_t = \beta_0 + \beta_1 y_{t-1} + \epsilon_t$$

- Autoregressive moving-average model
- Autoregressive integrated moving average