LEARNING OBJECTIVES

Upon completing this session, you should be able to do the following:

- Develop and explain basic forecasting models
- Identify the components present in a time series
- Apply trend-based forecasting models, including linear trend, nonlinear trend, and seasonally adjusted trend
- Use smoothing-based forecasting models

THE IMPORTANCE OF FORECASTING

- Governments forecast unemployment, interest rates, and expected tax revenues for policy purposes
- Marketing executives forecast demand, sales, and consumer preferences for strategic planning
- College administrators forecast enrolments to plan for facilities and for faculty recruitment
- Retail stores forecast demand to control inventory levels, hire employees and provide training

GENERAL FORECASTING ISSUES

- **Forecasting Horizon (Lead Time):**

  The number of future periods covered by a forecast

  Intermediate term – less than one month

  Short term – one to three months

  Medium term – three months to two years

  Long term – two years or more

- **Forecasting period:**

  The unit of time for which forecasts are to be made

- **Forecasting interval:**

  The frequency with which new forecasts are prepared

TIME-SERIES DATA

- Numerical data obtained at regular time intervals
- The time intervals can be annually, quarterly, daily, hourly, etc.

- Example:

Year:	2003	2004	2005	2006	2007
Sales:	75.3	74.2	78.5	79.7	80.2

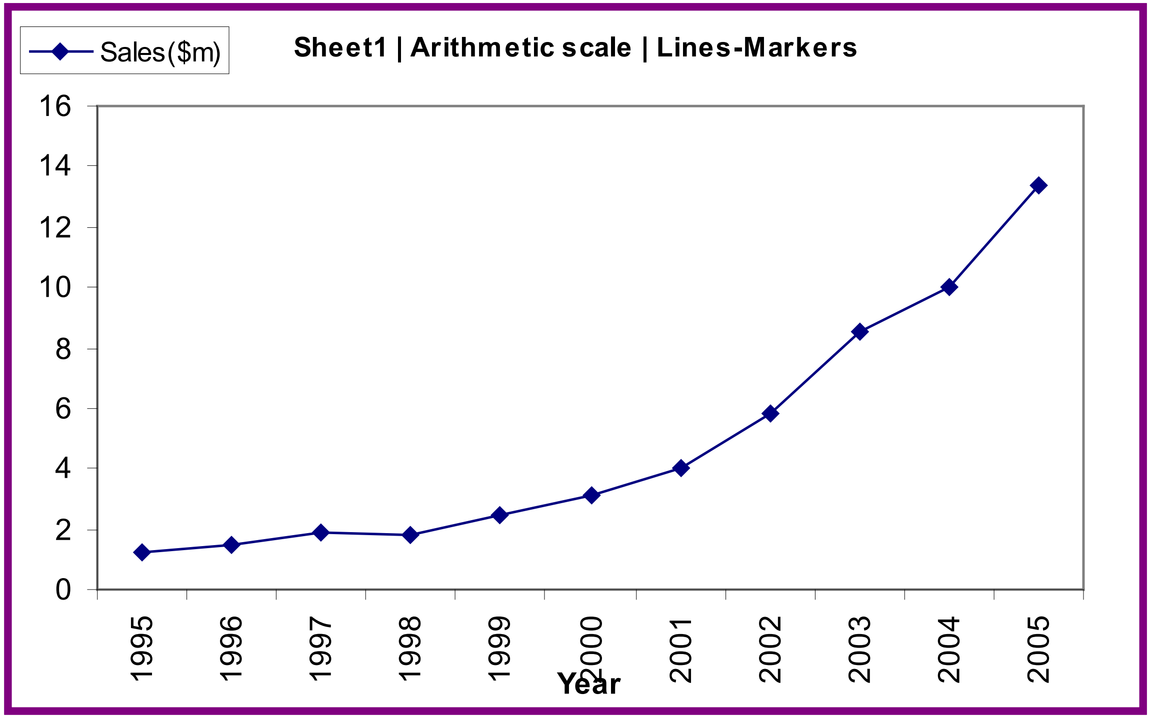
PRESENTATION OF TIME-SERIES DATA

- Table with values of the time series variable listed in chronological order
- The time series table may give an indication of any trends or patterns in the time series, and may highlight any potential abnormalities.

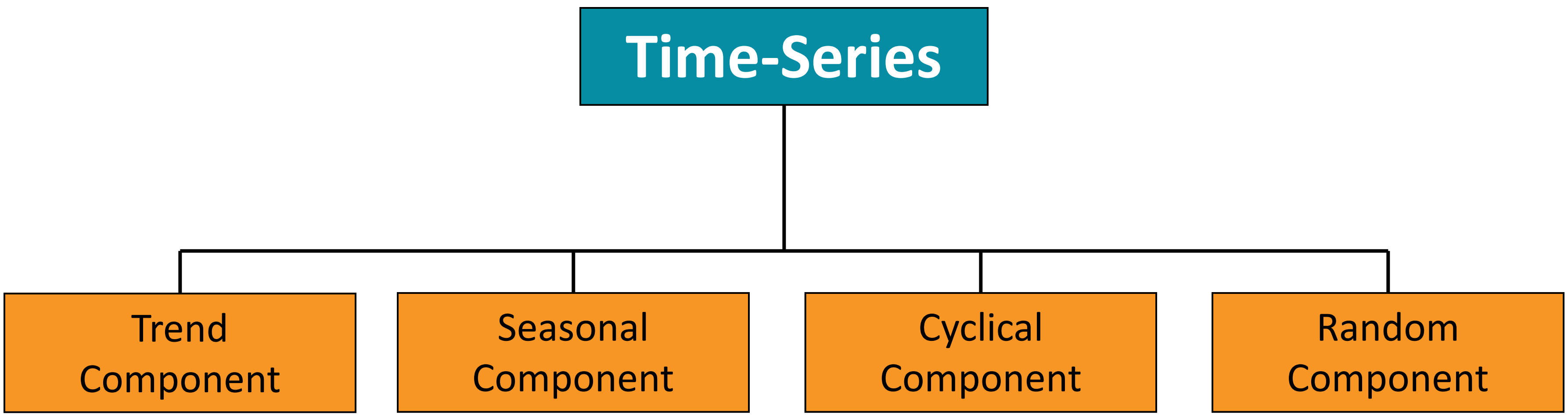
Year	Time X	Y Sales($m)
1995	1	1.2
1996	2	1.5
1997	3	1.9
1998	4	1.8
1999	5	2.5
2000	6	3.1
2001	7	4.0
2002	8	5.8
2003	9	8.5
2004	10	10.0
2005	11	13.4

PRESENTATION OF TIME-SERIES DATA

- Time series data can be drawn as a line graph with time on the horizontal axis.

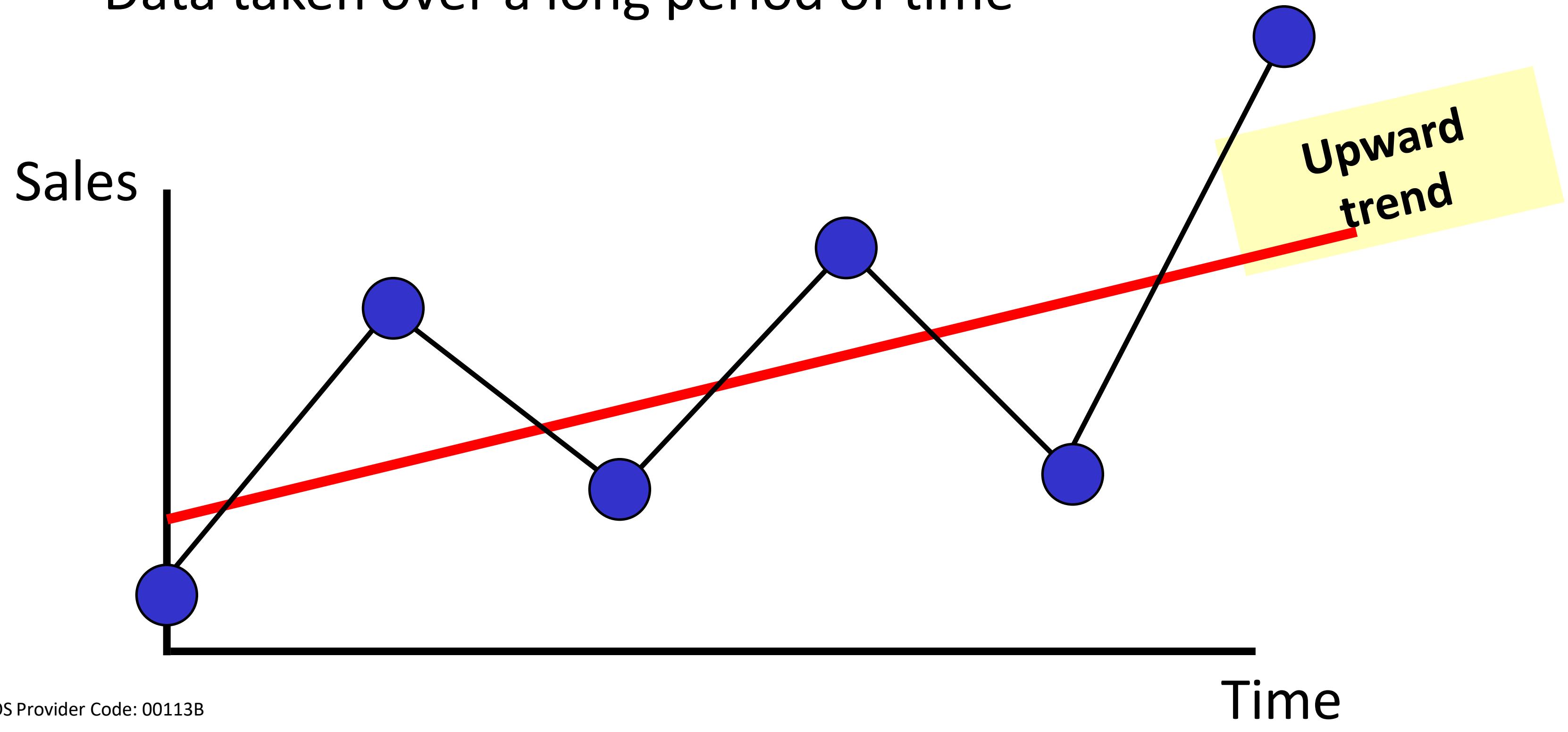


TIME-SERIES COMPONENTS



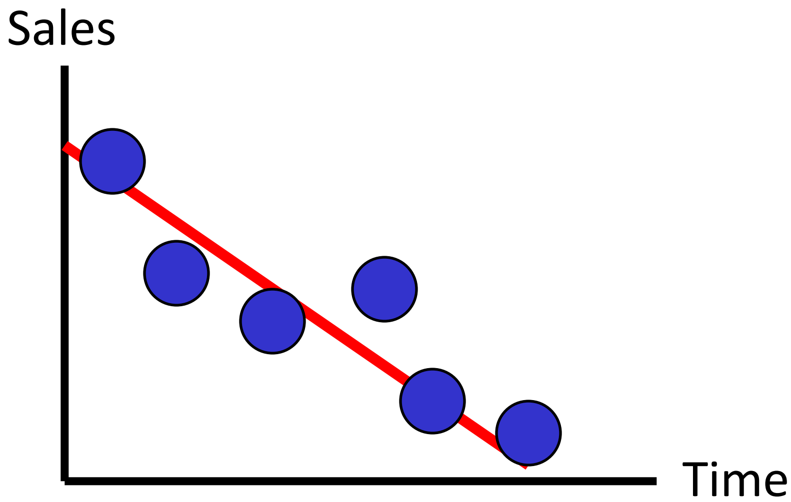
TREND COMPONENT

- Long-run increase or decrease over time (overall upward or downward movement)
- Data taken over a long period of time

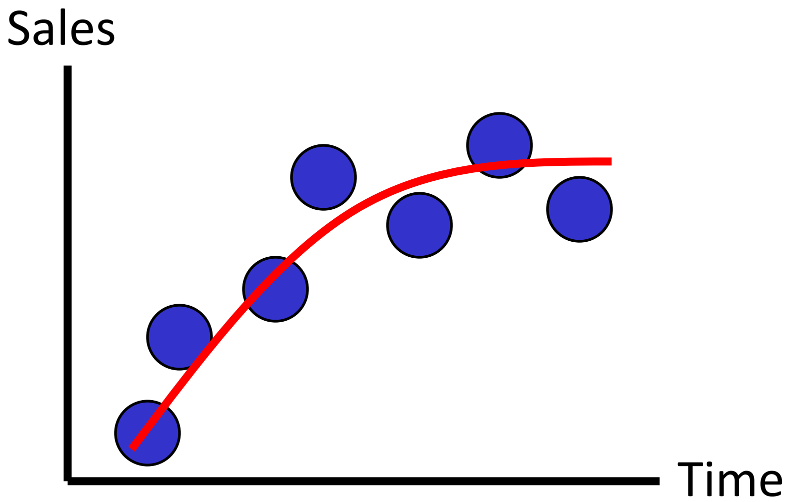


TREND COMPONENT (CONT'D)

- Trend can be upward or downward
- Trend can be linear or non-linear



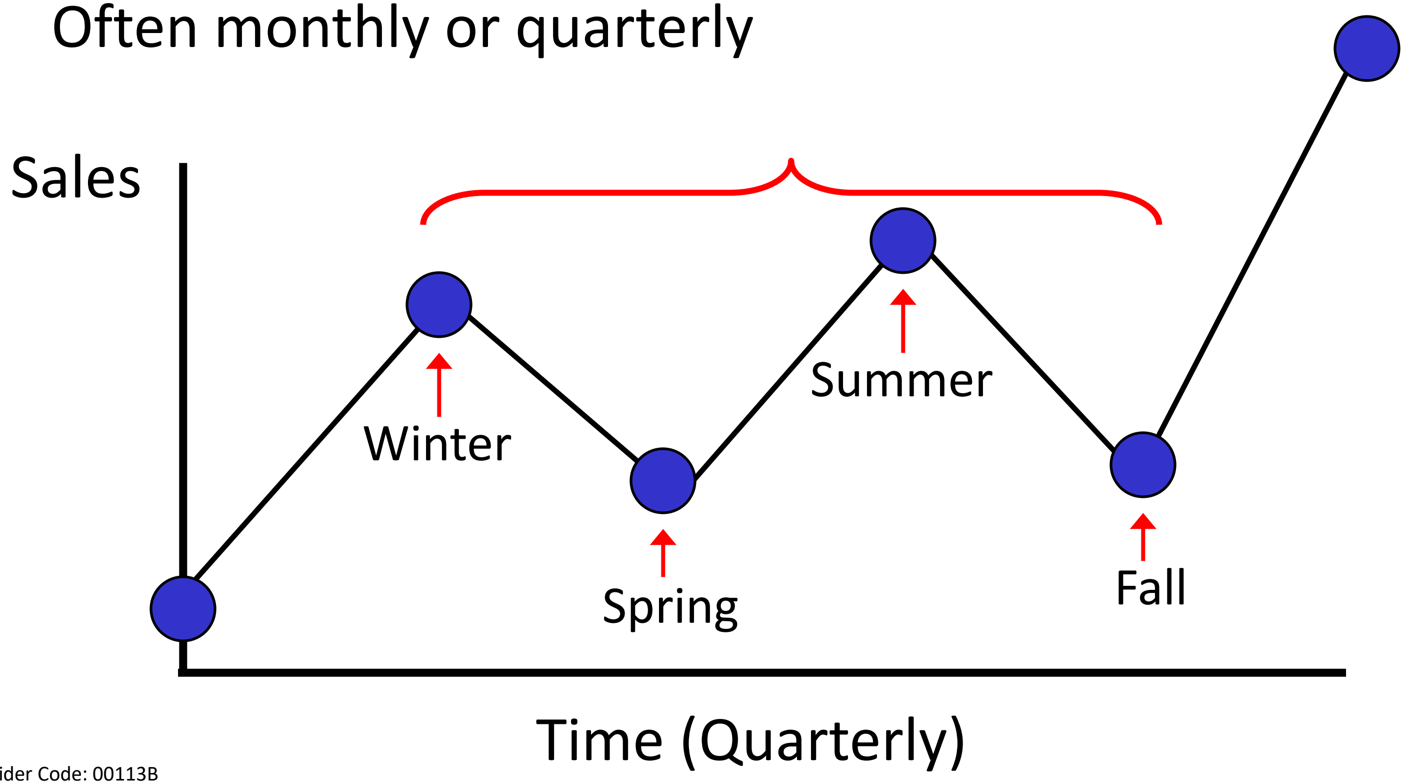
Downward linear trend



Upward nonlinear trend

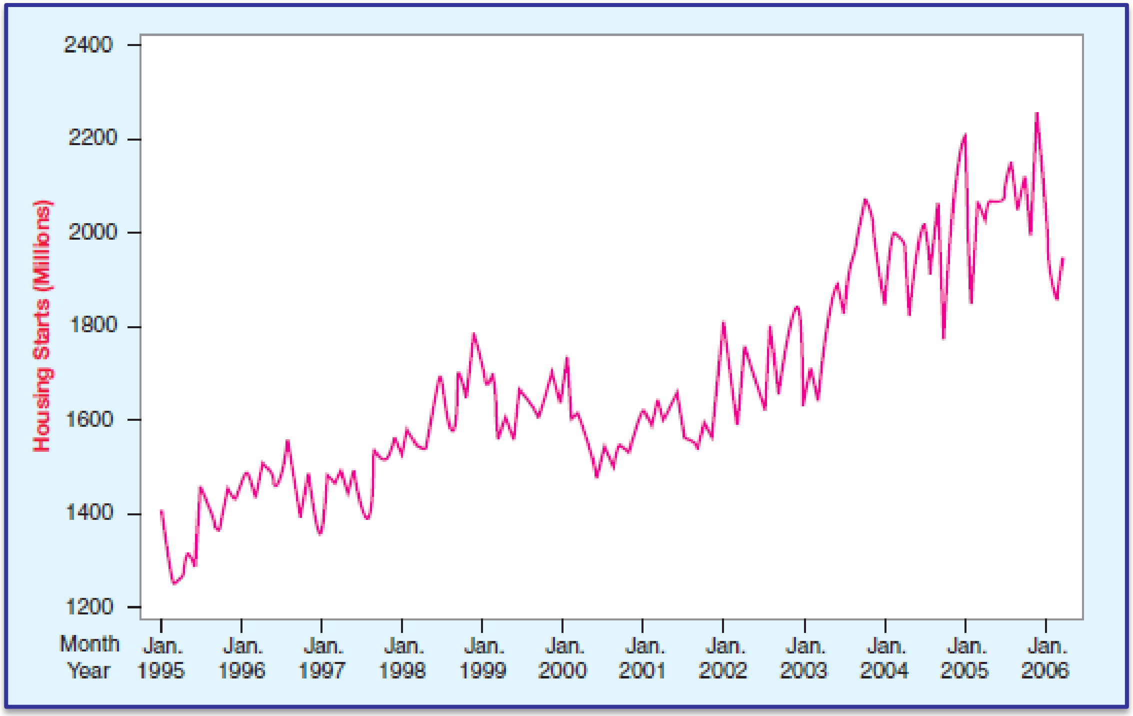
SEASONAL COMPONENT

- Short-term regular wave-like patterns
- Observed within 1 year
- Often monthly or quarterly



CYCLICAL COMPONENT

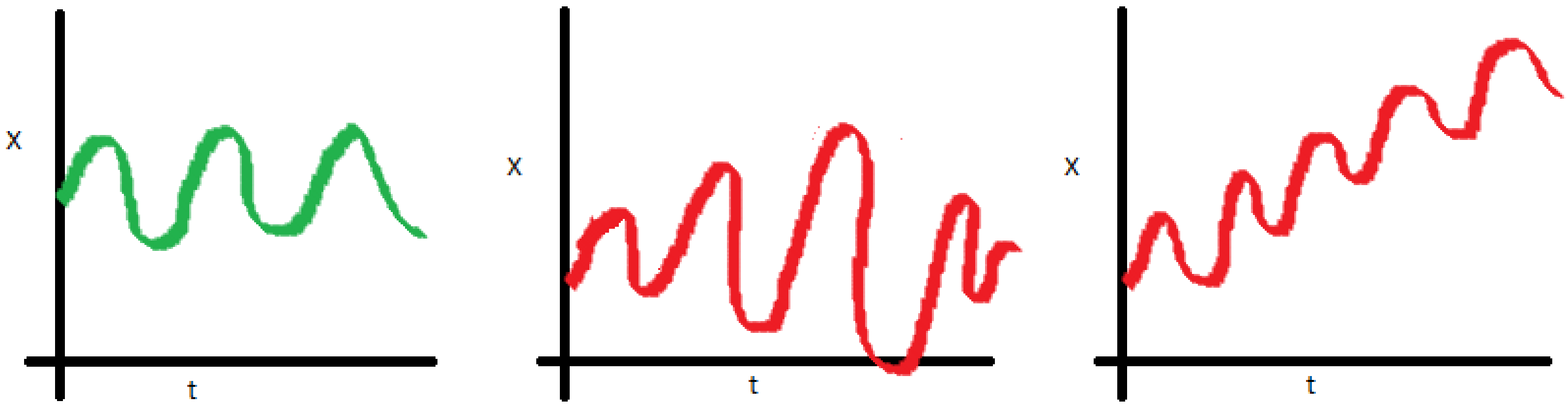
- A wavelike pattern
- It has a recurrence period of more than one year

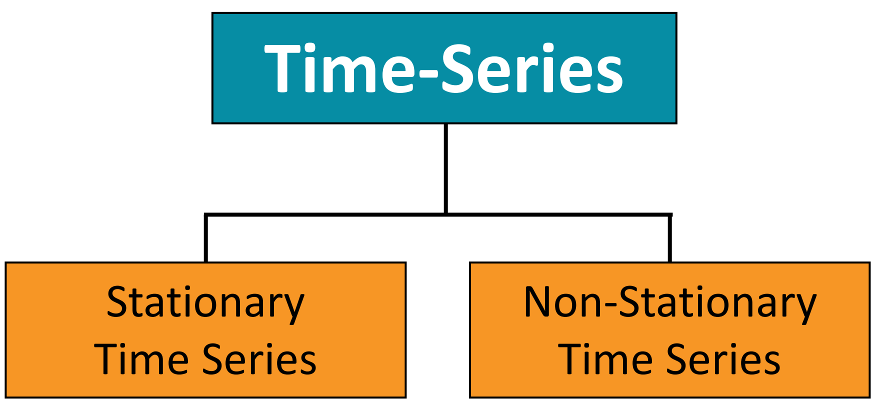


RANDOM COMPONENT

- Unpredictable, random, “residual” fluctuations
- Due to random variations of
  - Nature
  - Accidents or unusual events
- “Noise” in the time series

TYPES OF TIME SERIES





STATIONARY TIME SERIES

- A time series that shows no tendency to grow or decline over time.
- It tends to plot more or less as a horizontal line, with random fluctuations above and below the long term average.
- A stationary time series does NOT contain a long term upward or downward trend.
- However, it may contain:

  Seasonal, Cyclical and Irregular movements.

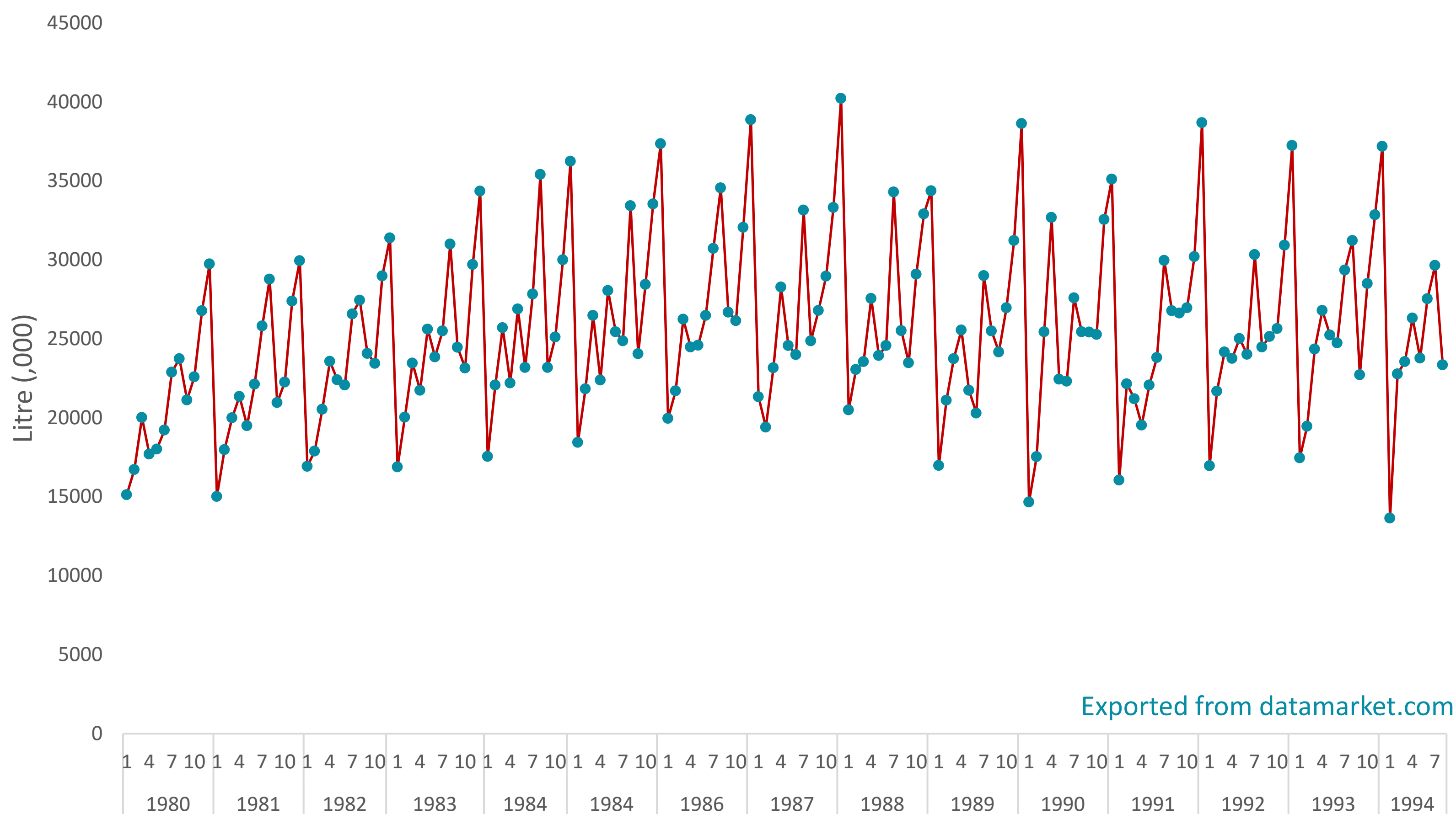
NON-STATIONARY TIME SERIES

- Generally one that tends to increase over time (that is, it has an upward trend) or decrease over time (that is, it has a downward trend).
- A non-stationary time series can also contain the other time series components:

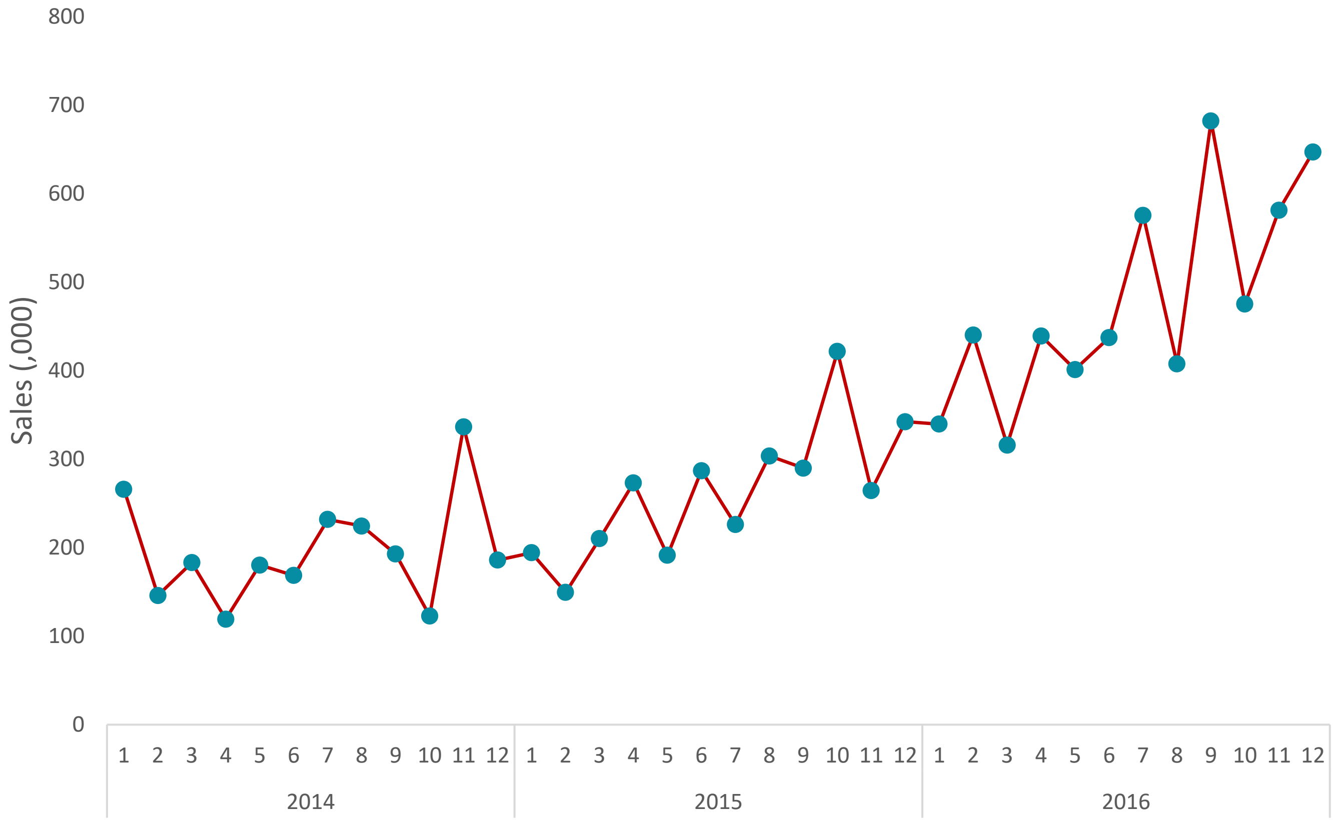
  Seasonal, Cyclical and Irregular movements.

EXAMPLE:

MONTHLY AUSTRALIAN WINE SALES: BY WINE MAKERS IN BOTTLES ≤ 1 LITRE

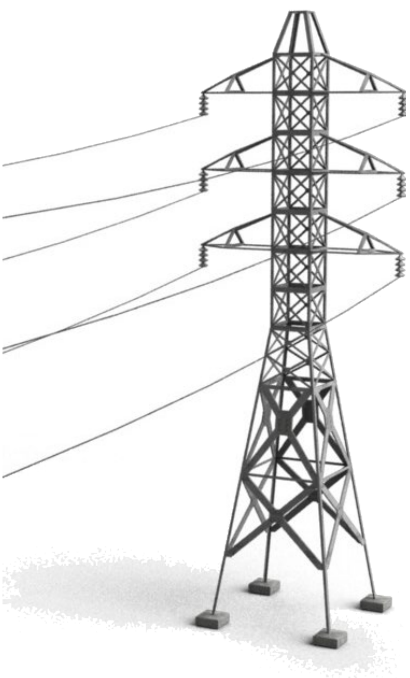


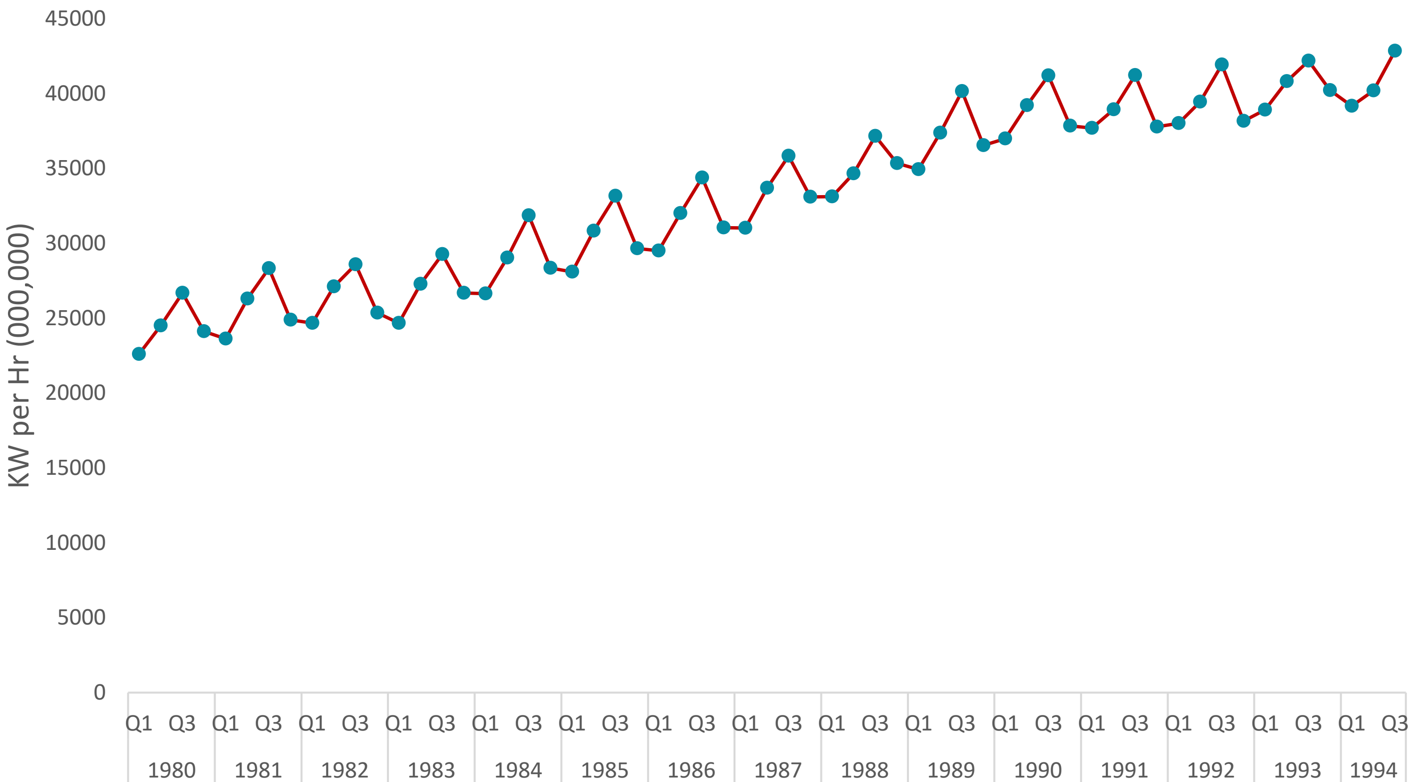
EXAMPLE:
SALES OF SHAMPOO OVER A THREE YEAR PERIOD



Exported from datamarket.com

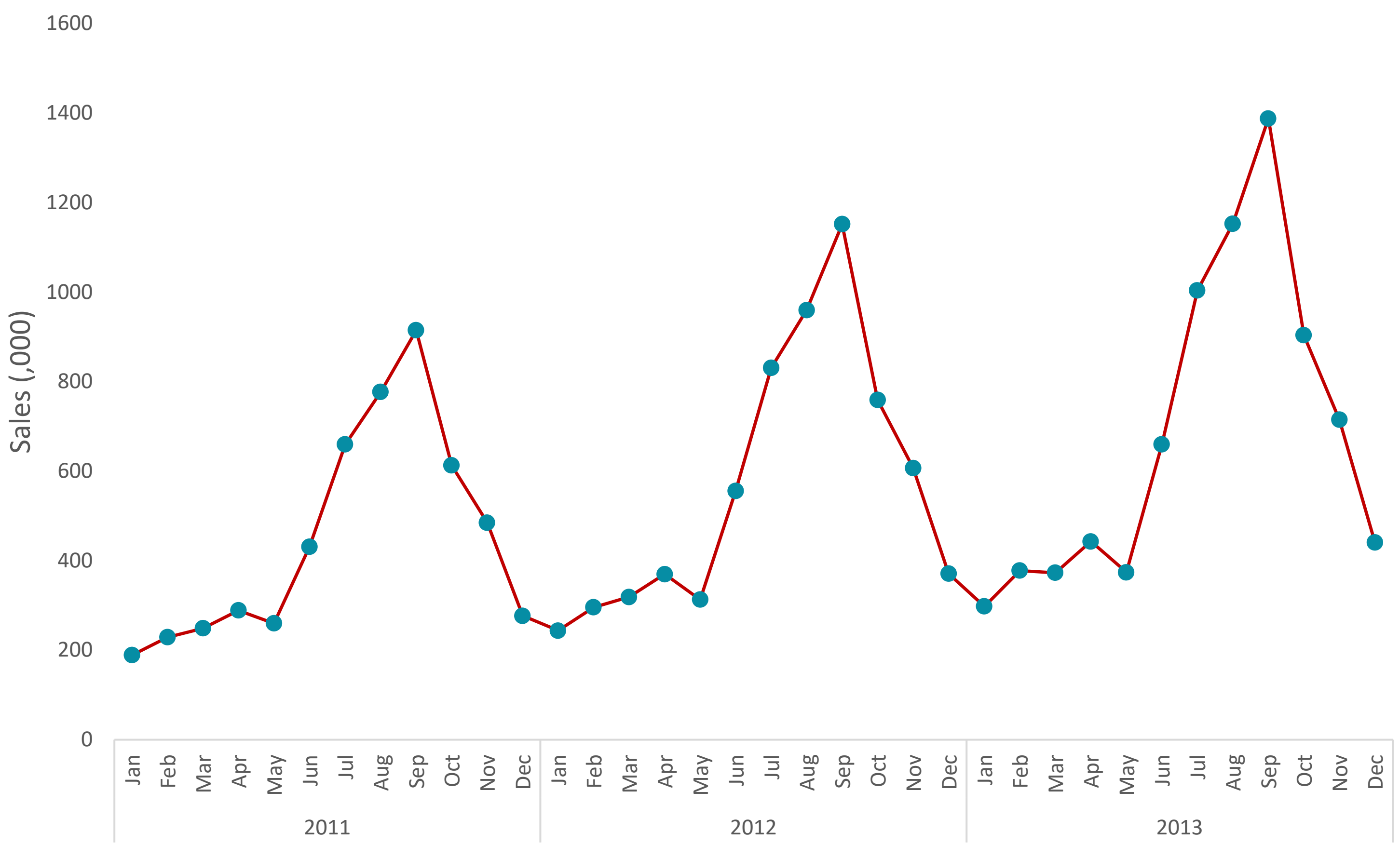
EXAMPLE:
QUARTERLY ELECTRICITY PRODUCTION IN AUSTRALIA





Exported from datamarket.com

EXAMPLE:
MONTHLY SALES OF TASTY COLA (JAN 2011 – DEC 2013)



Exported from datamarket.com

DEVELOPING A FORECASTING MODEL

- **Steps in forecast modelling:**
  - model specification
  - model fitting
  - model diagnosis

- **Goal:**

  Use the simplest available model that meets forecasting needs

WHAT IS SMOOTHING?

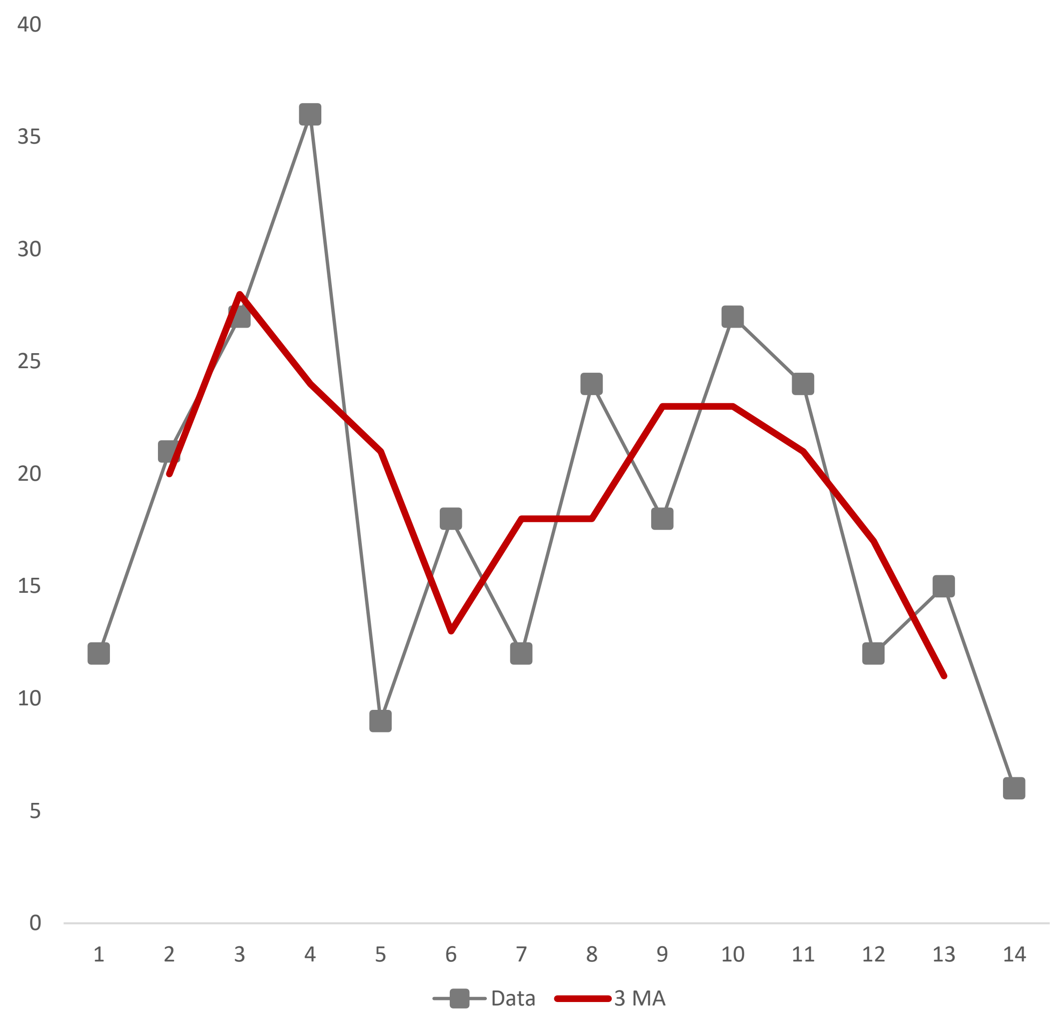
- Smoothing is usually done to help us better see patterns, trends in time series.
- Generally smooth out the irregular roughness to see a clearer signal.
- For seasonal data, we might smooth out the seasonality so that we can identify the trend.
  - Moving Average
  - Single / Double Exponential Smoothing

MOVING AVERAGE

- A moving average works by successively taking observations over a number of periods and averaging to smooth the data.
- M-period moving average – the average of the M data points.
- Moving average creates a bit of a difficulty when we have an even number of time periods in the seasonal span (as we usually do).

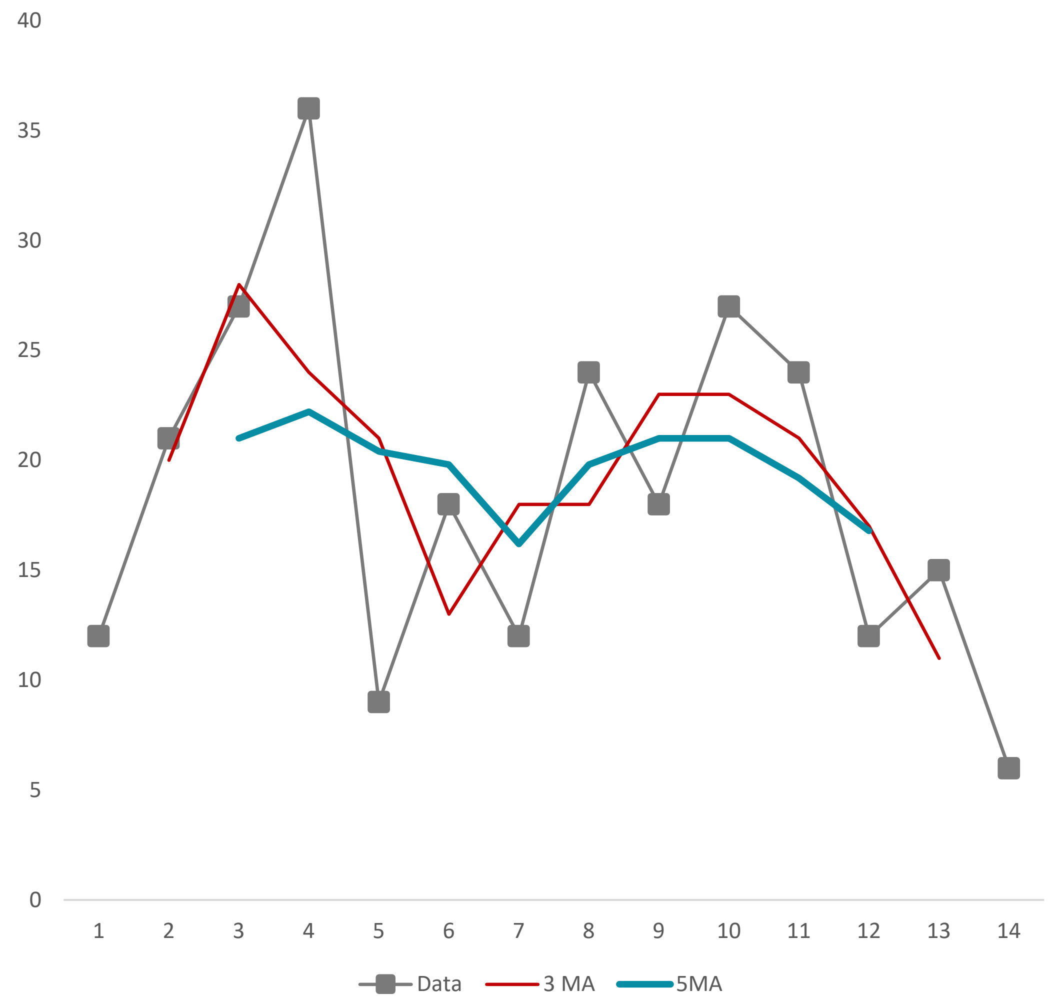
3-PERIOD MOVING AVERAGE

Data	3MA	3 MA
12		
21	=(12+21+27) / 3	20
27	=(21+27+36) / 3	28
36	=(27+36+9) / 3	24
9	=(36+9+18) / 3	21
18	…	13
12	…	18
24		18
18		23
27		23
24		21
12		17
15		11
6		



5-PERIOD MOVING AVERAGE

Data	5 MA	5 MA
12		
21		
27	=(12+21+27+36+9) / 5	21.00
36	=(21+27+36+9+18) / 5	22.20
9	=(27+36+9+18+12) / 5	20.40
18	…	19.80
12	…	16.20
24		19.80
18		21.00
27		21.00
24		19.20
12		16.80
15		
6		



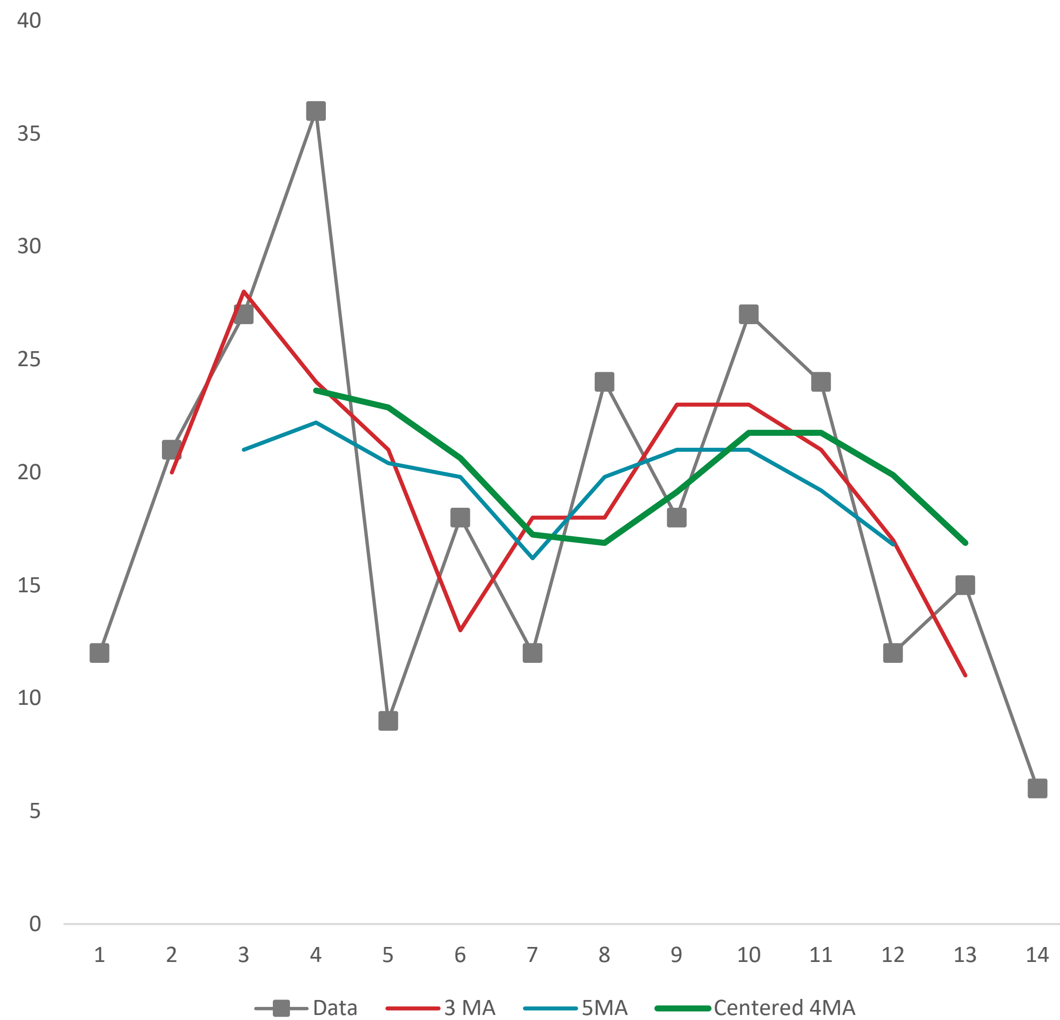
CENTRED 4-PERIOD AVERAGE

- A centred 4-Period MA is used for data with cycles of 4 periods (Ex: Summer, Autumn, Winter, Spring)
- The centre (average) of a 4-period timespan is at period 2.5 (i.e. between 2nd and 3rd observation) which does not match the original periods.
- To overcome this issue, take the average of pairs of off-centred average values to re-centre them.

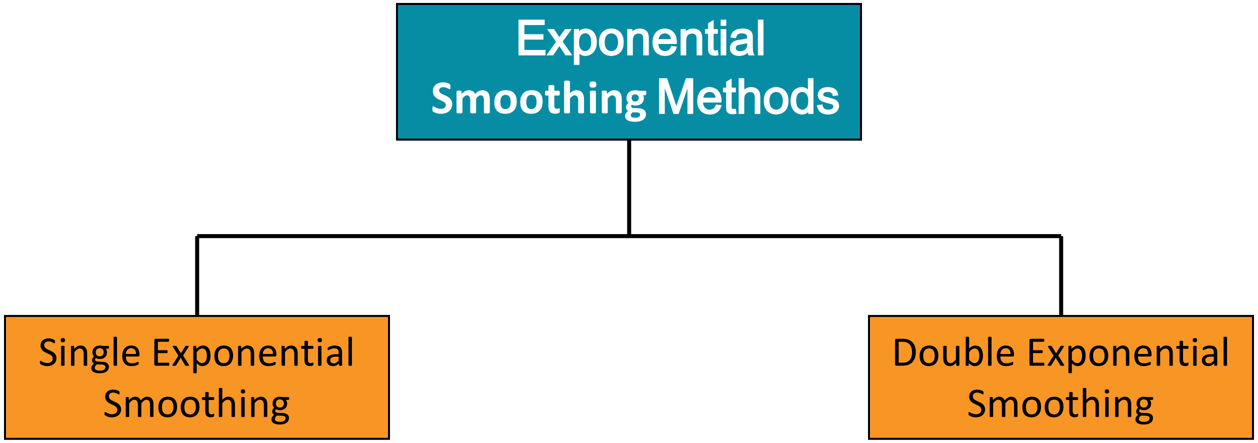
- Same methodology can be applied for any even numbered seasons.

CENTRED 4-PERIOD MA

Time	Data	4 MA	Centred 4 MA
1	12		
2	21		
3	27	24.00	23.63
4	36	23.25	22.88
5	9	22.50	20.63
6	18	18.75	17.25
7	12	15.75	16.88
8	24	18.00	19.13
9	18	20.25	21.75
10	27	23.25	21.75
11	24	20.25	19.88
12	12	19.50	16.88
13	15	14.25	
14	6		



EXPONENTIAL SMOOTHING METHODS



SINGLE EXPONENTIAL SMOOTHING

- A weighted moving average
  - Weights decline exponentially
  - Most recent observation weighted most
- Used for smoothing and short term forecasting

SINGLE EXPONENTIAL SMOOTHING

- The weighting factor is α
  - Subjectively chosen
  - Range from 0 to 1
  - Smaller α gives more smoothing, larger α gives less smoothing
- The weight is:
  - Close to 0 for smoothing out unwanted cyclical and irregular components
  - Close to 1 for forecasting

SINGLE EXPONENTIAL SMOOTHING MODEL

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

New Forecast = Previous Forecast + α (Previous Actual – Previous Forecast)

OR

New Forecast = previous Forecast – α (Error)

SINGLE EXPONENTIAL SMOOTHING MODEL

- Suppose we use weight $\alpha = 0.2$

Quarter (t)	Sales (y_t)	Forecast from prior period	Forecast for next period (F_{t+1})
1	23	NA	23
2	40	23	23+0.2(40-23) = 26.4
3	25	26.4	26.4 + 0.2 (25 – 26.4) =26.12
4	27	26.12	26.12 + (0.2)(27 - 26.12)=26.29
5	32	26.296	26.29 + (0.2)(32 - 26.29)=27.437
6	48	27.437	= 31.549
7	33	31.549	= 31.840
8	37	31.840	= 32.872
9	37	32.872	= 33.697
10	50	33.697	= 36.958
etc...	etc...	etc...	etc...

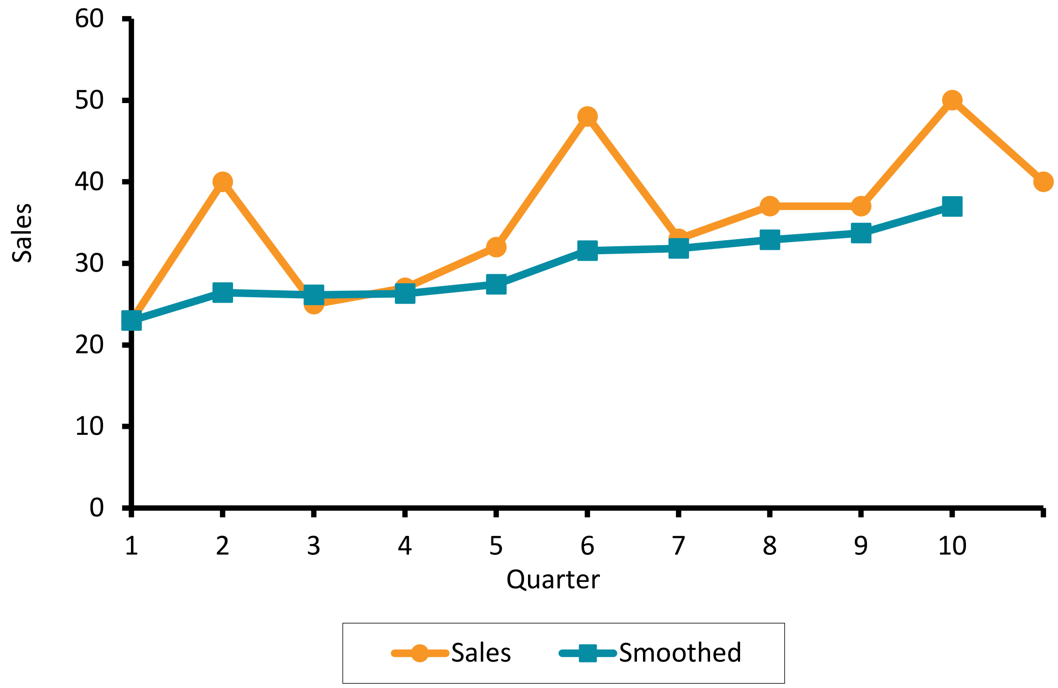
$F_1 = y_1$ since no prior information exists

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

SINGLE EXPONENTIAL SMOOTHING MODEL

- Seasonal fluctuations have been smoothed.
- NOTE:

  The smoothed value in this case is generally a little low, since the trend is upward sloping and the weighting factor is only 0.2.



DOUBLE EXPONENTIAL SMOOTHING

- Double exponential smoothing is sometimes called exponential smoothing with trend
- If trend exists, single exponential smoothing may need adjustment
- Add a second smoothing constant to account for trend
- [[Double exponential smoothing might be used when there's trend (either long run or short run), but no seasonality]]

DOUBLE EXPONENTIAL SMOOTHING

$$\boxed{C_t = \alpha y_t + (1-\alpha)(C_{t-1} + T_{t-1})}$$

$$\boxed{T_t = \beta(C_t - C_{t-1}) + (1-\beta)T_{t-1}}$$

$$\boxed{F_{t+1} = C_t + T_t}$$

where:

y_t = actual value in time t

α = constant-process smoothing constant

β = trend-smoothing constant

C_t = smoothed constant-process value for period t

T_t = smoothed trend value for period t

F_{t+1}= forecast value for period t + 1

t = current time period

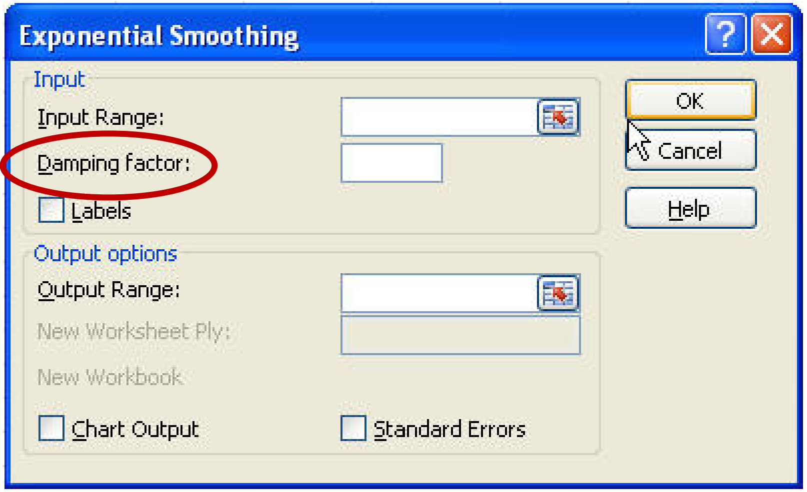
DOUBLE EXPONENTIAL SMOOTHING

- Double exponential smoothing is generally done by computer
- Use larger smoothing constants α and β when less smoothing is desired
- Use smaller smoothing constants α and β when more smoothing is desired

EXPONENTIAL SMOOTHING IN EXCEL

- Use: Data → data analysis → exponential smoothing

The "damping factor" is (1 - α)



TREND-BASED FORECASTING

- When data has an underlying linear trend, a linear model (equation) using least squares regression can be fitted.
- Simple linear models can be extended to include seasonality (additive and multiplicative).

Year	Time Period (t)	Sales (y)
2003	1	20
2004	2	40
2005	3	30
2006	4	50
2007	5	70
2008	6	65

- Use time (t) as the independent variable:

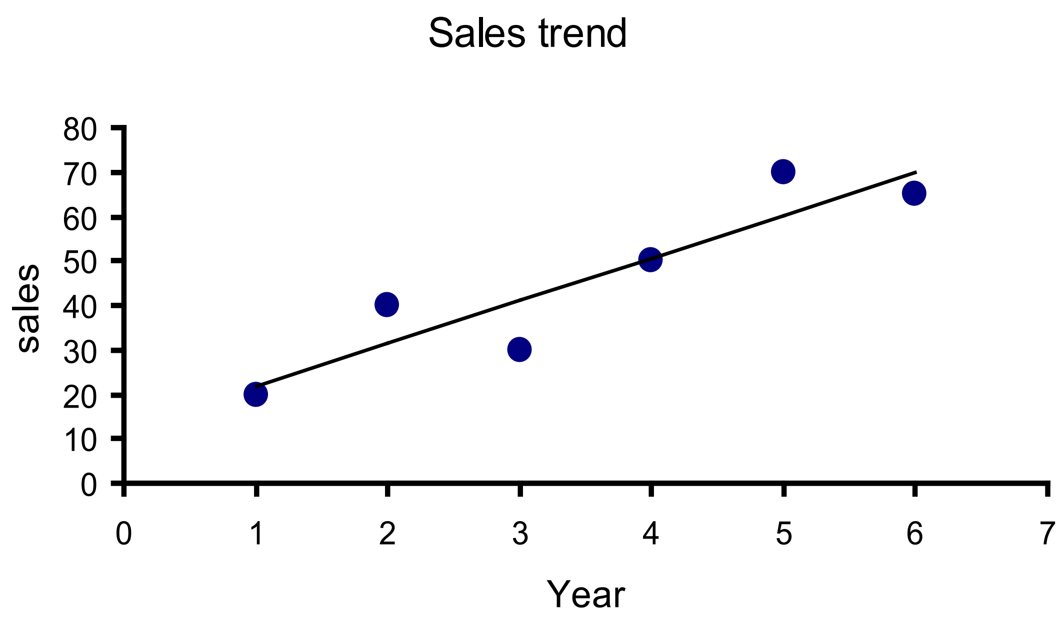
$$\hat{y} = b_0 + b_1 t$$

TREND-BASED FORECASTING (CONT'D)

Year	Time Period (t)	Sales (y)
2003	1	20
2004	2	40
2005	3	30
2006	4	50
2007	5	70
2008	6	65

- The linear trend model is:

$$\hat{y} = 12.333 + 9.5714\, t$$

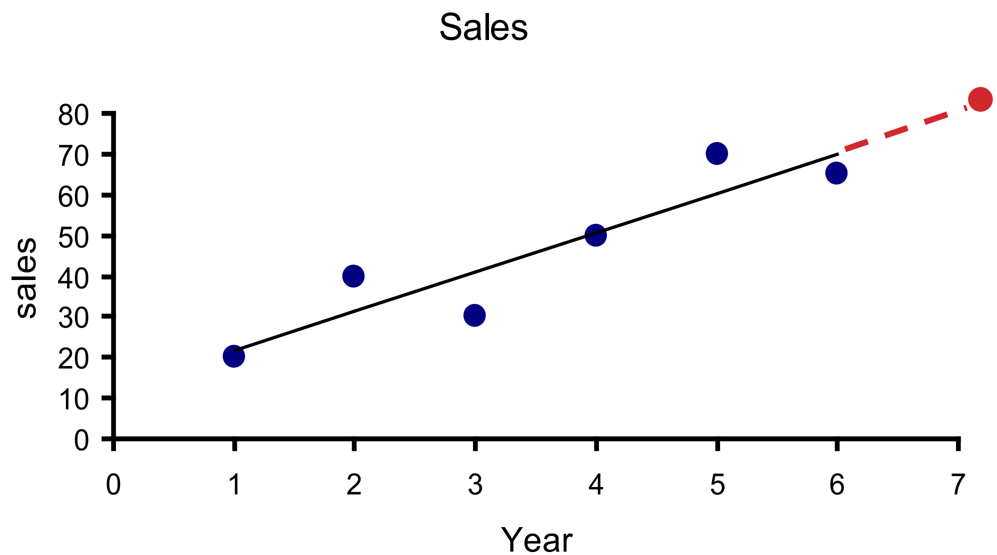


TREND-BASED FORECASTING
(CONT'D)

Year	Time Period (t)	Sales (y)
2003	1	20
2004	2	40
2005	3	30
2006	4	50
2007	5	70
2008	6	65
2009	7	???

- Forecast for time period 7:

$$\hat{y} = 12.333 + 9.5714\,(7)$$
$$= \boxed{79.33}$$



ACCURACY OF THE FORECAST
COMPARING FORECAST VALUES TO ACTUAL DATA

- The forecast error or residual is the difference between the actual value in time *t* and the forecast value in time *t*

- Error in time t:

$$\boxed{e_t = y_t - F_t}$$

TWO COMMON MEASURES OF FIT

- Measures of fit are used to gauge how well the forecasts match the actual values.

MSE (mean squared error)

$$MSE = \frac{\sum (y_t - F_t)^2}{n}$$

- Average squared difference between y_t and F_t

MAD (mean absolute deviation)

$$MAD = \frac{\sum | y_t - F_t |}{n}$$

- Average absolute value of difference between y_t and F_t
  Less sensitive to extreme values, error is in scale of the variable.

MEAN ABSOLUTE PERCENT ERROR (MAPE)

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}}{n}$$

Where:

y_i = Actual value of y

$\hat{y}_i$ = Model forecasted values of y

n = number of observations

NONLINEAR TREND-BASED FORECASTING

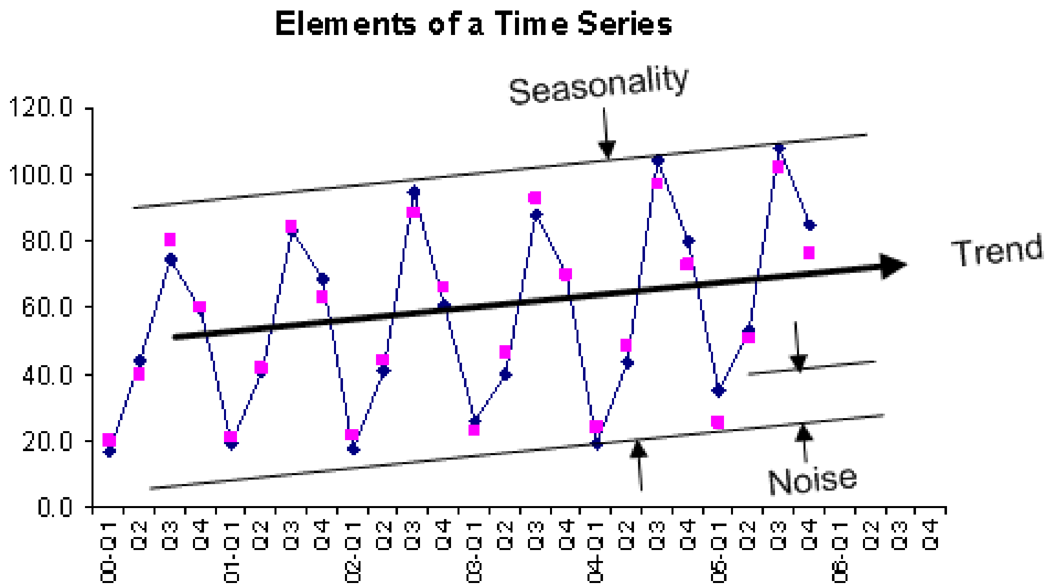
- A nonlinear regression model can be used when the time series exhibits a nonlinear trend
- One form of a nonlinear model:

$$\boxed{y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon}$$

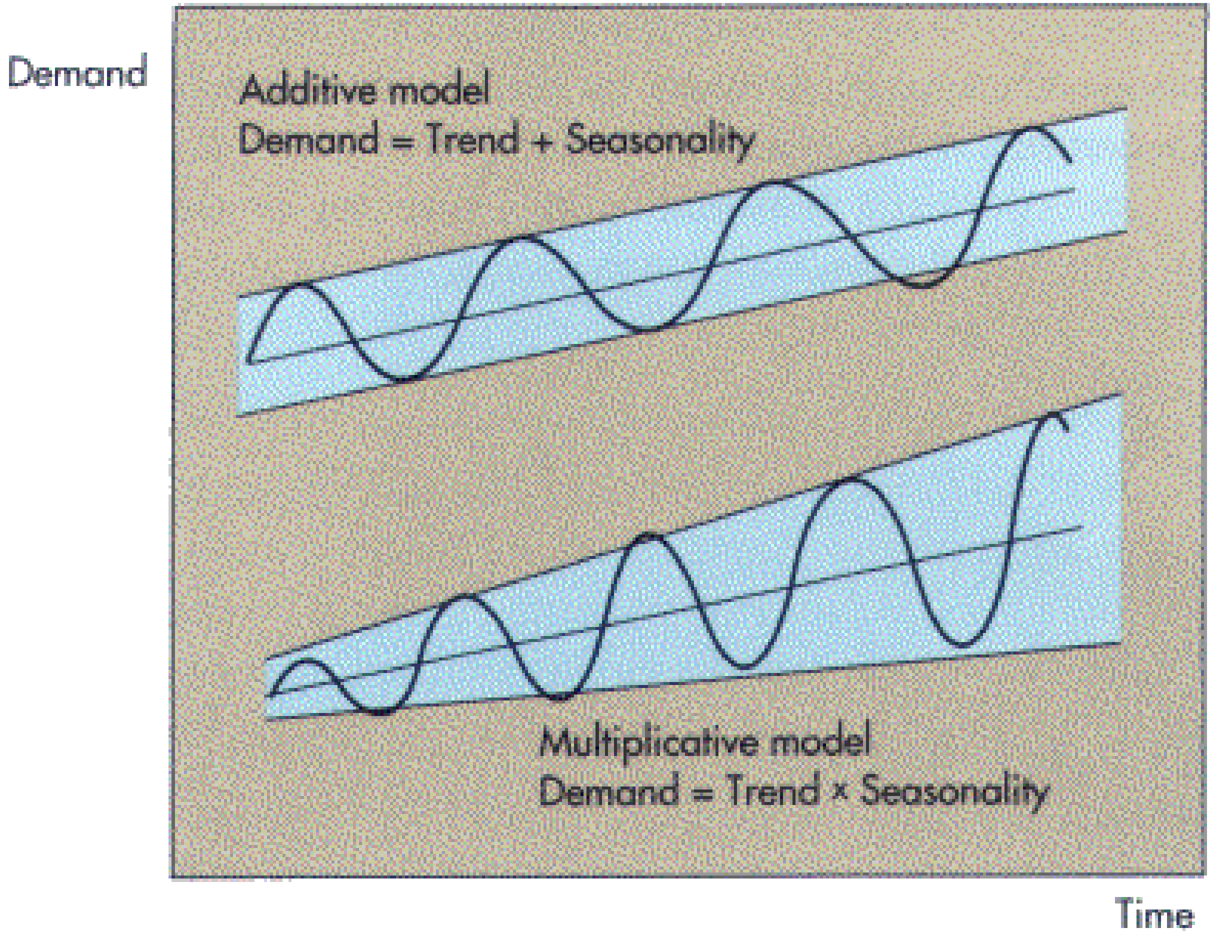
- Compare R^2 and s_ε to that of linear model to see if this is an improvement
- Can try other functional forms to get best fit.

TREND-BASED FORECASTING FOR DATA WITH SEASONAL COMPONENT

- Components of a time series:
  - Trend
  - Seasonal
  - Cyclic element
  - A random component



LINEAR TREND MODELS WITH SEASONAL COMPONENT



FORECASTING SEASONAL DATA

- Need to observe whether the model follows an additive or multiplicative model.
- If the underlying model is multiplicative, then seasonal indices can be determined. Then the de-seasonalised series can be forecast. Seasonality can be added to the final prediction.
- If the underlying model is additive then multiple regression is the usual approach to modelling the time series.

  (not covered in this course)

MULTIPLICATIVE TIME-SERIES MODEL

- Used primarily for forecasting
- Allows consideration of seasonal variation
- Observed value in time series is the product of components

$$\boxed{y_t = T_t \times S_t \times C_t \times I_t}$$

where T_t = Trend value at time t

S_t = Seasonal value at time t

C_t = Cyclical value at time t

I_t = Irregular (random) value at time t

MULTIPLICATIVE TIME-SERIES MODEL

- Four main steps in building multiplicative time-series models:

  1. Calculating Seasonal Indices
  2. De-seasonalising Data
  3. Non-seasonal forecast
  4. Re-seasonalise the forecast

MULTIPLICATIVE TIME SERIES MODEL

- Ratios to moving average method
  1. Smooth the time series (Use 4 centred MA for quarterly data)
  2. Divide each observation by its corresponding MA
  3. Calculate average ratio for each season
  4. Normalise ratios (to have an average of 1)

FINDING SEASONAL INDICES

- Calculate seasonal indices for the following data

Quarter	Sales	Centred 4 MA	Ratio (Obs /MA)
1	724		
2	770		
3	864	765	1.13
4	682	776	0.88
1	764	798.5	0.96
2	818	826.5	0.99
3	996	860.75	1.16
4	774	909.5	0.85
1	946	956.5	0.99
2	1026	999.25	1.03
3	1164	1038.75	1.12
4	948	1073.75	0.88
1	1088	1115.75	0.98
2	1164	1161.25	1.00
3	1362		
4	1114		

= 864 / 765 = 1.13
Quarter 3 sales are 13% higher than the annual average sales

= 764 / 798.5 = 0.96
Quarter 1 sales are 4% lower than the annual average sales

Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

NORMALISING SEASONAL INDICES

- Normalising seasonal indices

Quarter	Sales	Centred 4 MA	Ratio (Obs /MA)
1	724		
2	770		
3	864	765	1.13
4	682	776	0.88
1	764	798.5	0.96
2	818	826.5	0.99
3	996	860.75	1.16
4	774	909.5	0.85
1	946	956.5	0.99
2	1026	999.25	1.03
3	1164	1038.75	1.12
4	948	1073.75	0.88
1	1088	1115.75	0.98
2	1164	1161.25	1.00
3	1362		
4	1114		

Quarter	1	2	3	4	
			1.13	0.88	
	0.96	0.99	1.16	0.85	
	0.99	1.03	1.12	0.88	
	0.98	1.00			
Average	**0.97**	**1.01**	**1.14**	**0.87**	**3.99**
Normed Indices	**0.98**	**1.01**	**1.14**	**0.87**	**4.00**

= 0.97 * 4.0 / 3.99 = 0.98 Each average is multiplied by 4/3.99 to get the index

Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

DE-SEASONALISING

- The data is de-seasonalised by dividing the observed value by its seasonal index

$$\boxed{T_t \times C_t \times I_t = \frac{y_t}{S_t}}$$

- This smooths the data by removing seasonal variation

Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

DE-SEASONALISING

Period	Quarter	Sales	Index	De-seasonalised
1	1	724	0.98	738.78
2	2	770	1.01	762.38
3	3	864	1.14	757.89
4	4	682	0.87	783.91
5	1	764	0.98	779.59
6	2	818	1.01	809.90
7	3	996	1.14	873.68
8	4	774	0.87	889.66
9	1	946	0.98	965.31
10	2	1026	1.01	1015.84
11	3	1164	1.14	1021.05
12	4	948	0.87	1089.66
…		…	…	

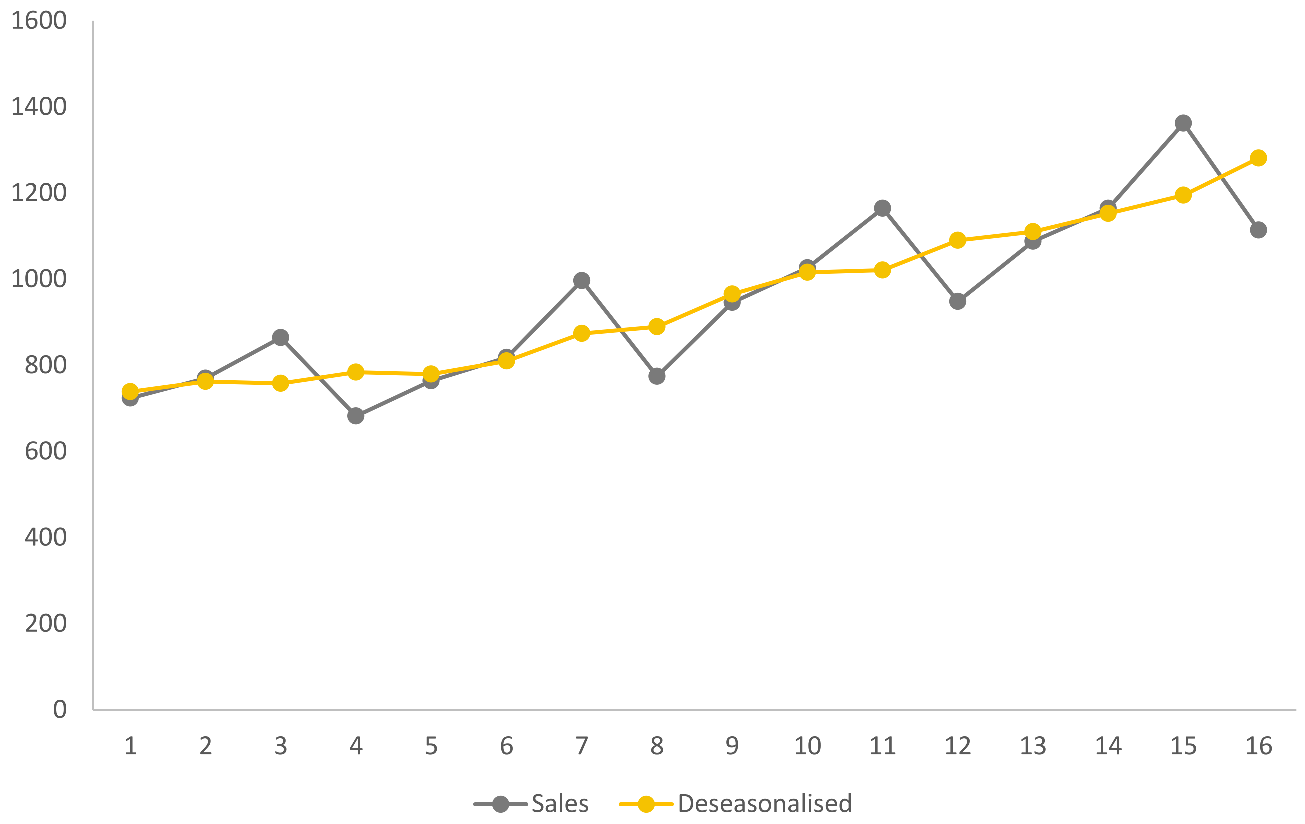
Example:

$$\boxed{738.78 = \frac{724}{0.98}}$$

etc…

Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

UNSEASONALISED V. SEASONALISED

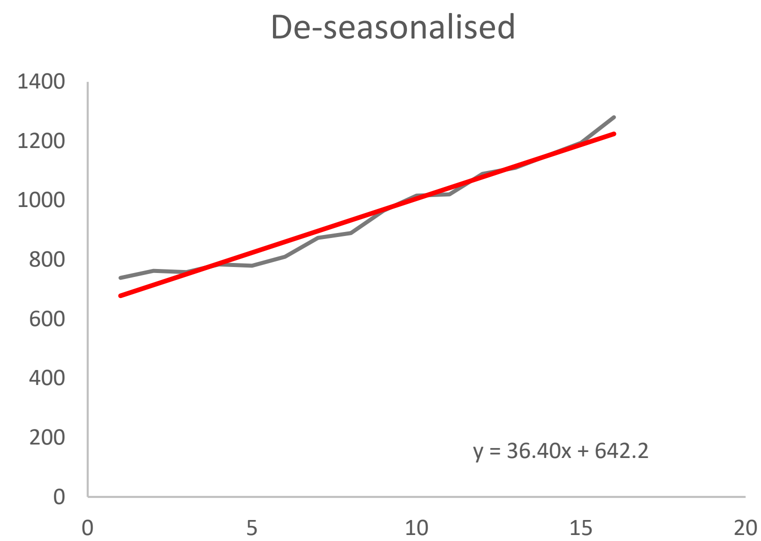


- Calculating seasonal indices
- De-seasonalising Data
- Non-seasonal forecast
- Re-seasonalise the forecast

FORECASTING

- Can find the best fitted line on de-seasonalised data.
- Can be used to generate the non-seasonal forecast

Period	Quarter	Sales	Index	De-seasonalised	Trend Y = 36.40 x + 642.2
1	1	724	0.98	738.78	678.6
2	2	770	1.01	762.38	715
3	3	864	1.14	757.89	751.4
4	4	682	0.87	783.91	787.8
5	1	764	0.98	779.59	824.2
6	2	818	1.01	809.90	860.6
7	3	996	1.14	873.68	897
8	4	774	0.87	889.66	933.4
9	1	946	0.98	965.31	969.8
10	2	1026	1.01	1015.84	1006.2
11	3	1164	1.14	1021.05	1042.6
12	4	948	0.87	1089.66	1079
		…	…	…	…



Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

FORECASTING

- Forecasting for the next eight quarters

Period	Quarter	Sales	Index	Deseasonalised	Trend Y = 36.40 x + 642.2	Forecast
12	4	…	…	…	…	
13	1	1088	0.98	1110.2	1115.4	
14	2	1164	1.01	1152.48	1151.8	
15	3	1362	1.14	1194.74	1188.2	
16	4	1114	0.87	1280.46	1224.6	
17	1		0.98		1261	**1235.78**
18	2		1.01		1297.4	**1310.374**
19	3		1.14		1333.8	**1520.532**
20	4		0.87		1370.2	**1192.074**
21	1		0.98		1406.6	**1378.468**
22	2		1.01		1443	**1457.43**
23	3		1.14		1479.4	**1686.516**
24	4		0.87		1515.8	**1318.746**

1261 * 0.98 = 1235.78

Calculating seasonal indices → De-seasonalising Data → Non-seasonal forecast → Re-seasonalise the forecast

PLOT OF DATA, TREND AND FORECAST