MONASH University

Information Technology

FIT1006 Business Information Analysis

Lecture 20
Time Series Analysis and Forecasting (Part 2)

Topics covered:

- Seasonal Indices
- Calculating multiplicative seasonal indices
- Regression based forecasting
- The accuracy of forecast

Lectures 19/20 Motivating problem

- Given the value of building work (quarterly) from Sep 1974 – Dec 2018 .
- Model time series.
- Use historical data to forecast demand for 2019 and 2020.
- Source: ABS.

http://www.abs.gov.au

(File: FIT1006 Lecture 19 and 20.xlsx)

Quarter/Year	Value of Building Work (all sectors) $'Bil
Sep-1974	11.53
Dec-1974	11.06
Mar-1975	9.64
Jun-1975	10.41
Sep-1975	11.15
Dec-1975	10.65
Mar-1976	10.18
Jun-1976	11.37
Sep-1976	11.63
Dec-1976	11.37
Mar-1977	10.14
Jun-1977	11.12
Sep-1977	11.07
Dec-1977	10.57
:	
Mar-2017	27.75
Jun-2017	30.59
Sep-2017	31.52
Dec-2017	31.86
Mar-2018	29.26
Jun-2018	32.84
Sep-2018	32.99
Dec-2018	32.69

Cont.

- If the actual value of building work in 2019 & 2020 is now known (as shown in the table), calculate the accuracy of the forecast.

Quarter/Year	Value of Building Work (All sectors) $'Bil
Mar-2019	29.74
Jun-2019	31.08
Sep-2019	32.17
Dec-2019	30.83
Mar-2020	28.35
Jun-2020	30.14
Sep-2020	30.24
Dec-2020	30.14

Recap from last lecture…

Forecast Accuracy

- One approach to measuring the accuracy of a forecast is to use Mean Absolute Percent Error (MAPE). This is the average error of a series of forecasts.

$$MAPE = \frac{\sum_{i=1}^{n} \frac{\left|\hat{Y}_i - Y_i\right|}{y_i}}{n}$$

$\hat{y}_i = forecast\ at\ period\ i$

$y_i = actual\ value\ period\ i$

$n = number\ of\ terms\ evaluated$

https://flux.qa (Feed code: 3U6KGV)

Question 1

For a forecast value of 15 and an observed value of 20 APE is:

A. – 0.25

✓ B. + 0.25

C. – 0.33

D. + 0.33

E. + 0.75

$$APE = \frac{|\hat{Y}_i - y_i|}{y_i}$$

$$APE = \frac{|15 - 20|}{20}$$

Group Activity

To forecast: $\hat{y}_{t+} \ldots (y_t - \hat{y}_t)$

For forecast accuracy $\ldots = \dfrac{|\hat{Y}_i - y_i|}{y_i}$

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

a = 0.6

Observed (y)	Forecast ($\hat{y}$)	Error	APE
55	55		
59			
53			
48			
44			
50			
52			

The first 'forecast' value is always = 'observed'

Solution

$\boxed{(y_t - \hat{y}_t)}$

$\alpha = 0.6$

Observed (y)	[illegible])	Error	APE
55	55.0	0.0	----------
59	55.0	4.0	0.07
53	57.4	-4.4	0.08
48	54.8	-6.8	0.14
44	50.7	-6.7	0.15
50	46.7	3.3	0.07
52	48.7	3.3	0.06
	50.7	MAPE	0.10

Average

Forecast = 48.7 + 0.6(3.3)

Today…

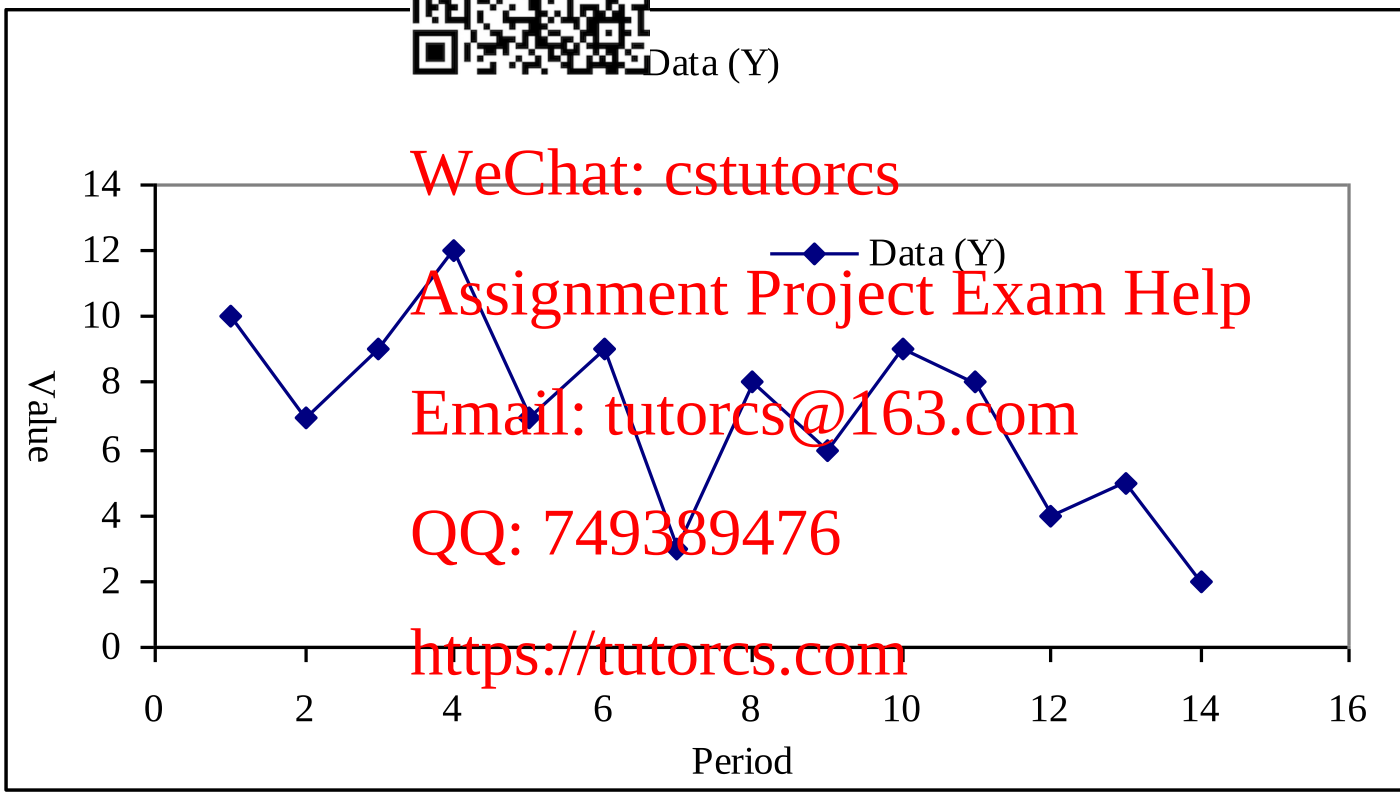
- Lecture 20 Regres[illegible]sed Forecasting

Regression Based Forecasting

- When data has an underlying linear trend, a linear model (equation) using least squares regression can be fitted.
- This approach enables a longer term forecast to be made (in contrast to the one or two step forecasts using exponential smoothing).
- Simple linear models can be extended to include additive and multiplicative seasonality.

Regression Based Forecasting

- Model the follow ear relationship…



Linear Regression of Time Series

- The first step in linear regression is to code the successive observations with an index.
- Typically use numbers, 1, 2, 3 … or 0, 1, 2, ... for this task (assuming equal time intervals).

- Eg, for an example time series, we code:
- Observation: 10, 7, 9, 12, ...
- Period: 1, 2, 3, 4, ...

…

- Then make the [illegible] regression calculations:

	Period (X)	[illegible]	XX	YY	XY
	1	[illegible]	1	100	10
	2	7	4	49	14
	3	9	9	81	27
	4	12	16	144	48
	5	7	25	49	35
	6	9	36	81	54
	7	3	49	9	21
	8	8	64	64	64
	9	6	81	36	54
	10	9	100	81	90
	11	8	121	64	88
	12	4	144	16	48
	13	5	169	25	65
	14	2	196	4	28
Σ	**105**	**99**	**1015**	**803**	**646**

least square equation or regression equation

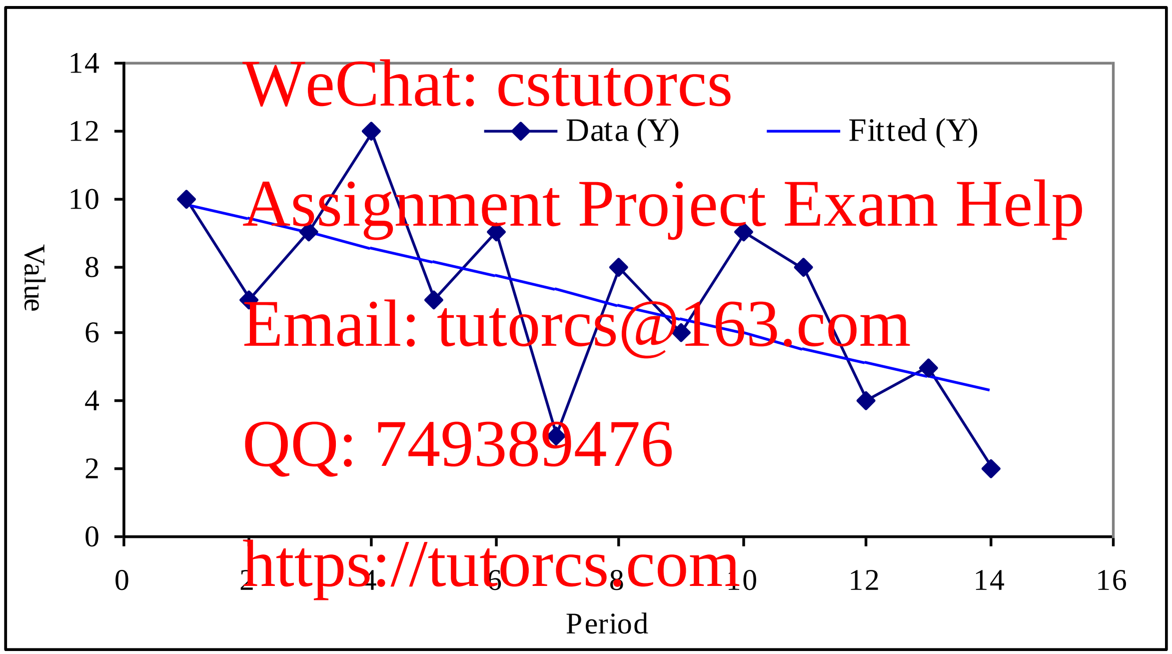
...

- The equation of the line is $Y = 10.25 - 0.42X$.
- From which a table of fitted values can be made.

Period (X)	Data (Y)	Fitted (Y)	
1	10	9.83	=10.25 - 0.42*1
2	7	9.40	=10.25 - 0.42*2
3	9	8.98	=10.25 - 0.42*3
4	12	8.56	=10.25 - 0.42*4
5	7	8.13	...
6	9	7.71	
7	3	7.28	
8	8	6.86	
9	6	6.44	
10	9	6.01	
11	8	5.59	
12	4	5.16	
13	5	4.74	
14	2	4.31	

…

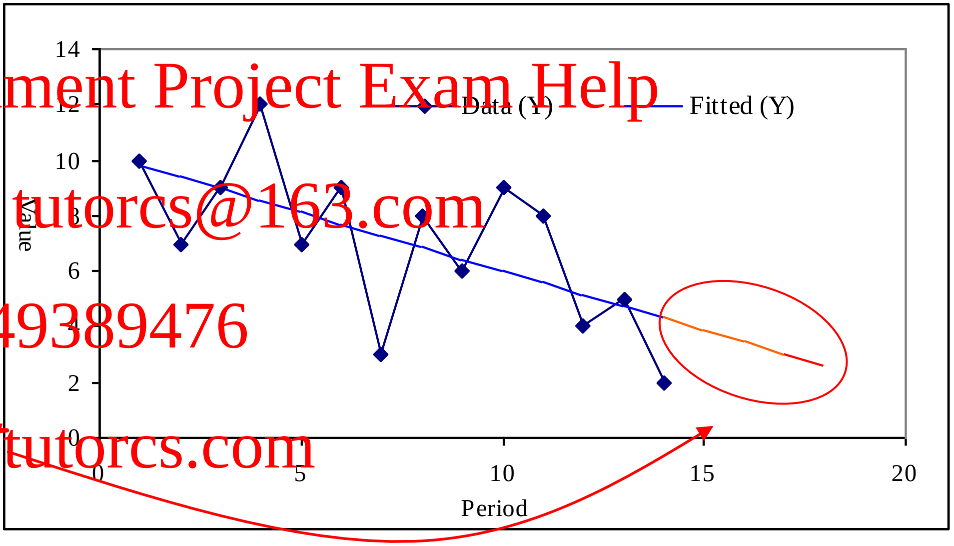
- A plot of observed values *vs* Least Squares fitted values.



Forecasting

- To forecast, ext[illegible]e model beyond the observed data. [illegible]ecast for periods 15, 16, 17 and 18 is:

Period (X)	Data (Y)	Fitted (Y)
3	9	8.98
4	12	8.56
5	7	8.13
…	…	…
13	3	4.74
14	8	4.31
15		3.89
16		3.47
17		3.04
18		2.62



using least square equation, substitute x values for period 15, 16, 17 & 18

Accuracy of the Forecast

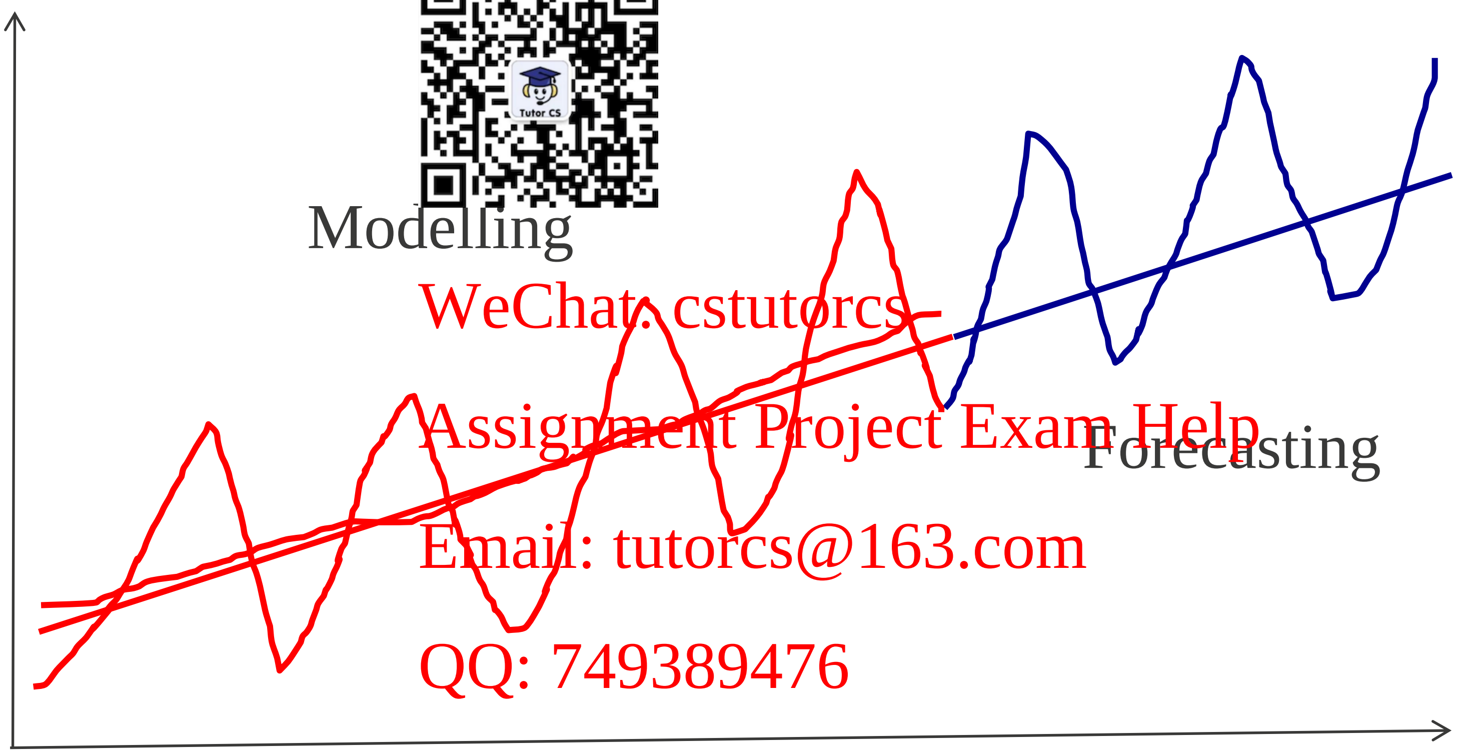
- If the observed … at periods 15 to 18 are subsequently f… be 4, 6, 5, 2, then MAPE is:

Period (X)	Data (Y)	Fitted (Y)	APE
3	9	8.98	
4	12	8.56	
5	7	8.13	
…	…	…	
13	3	4.74	
14	8	4.31	
15	4	3.89	0.03
16	6	3.47	0.42
17	5	3.04	0.39
18	2	2.62	0.31
		MAPE	

APE = Abs Error/Actual

$= \frac{|4 - 3.89|}{4} = 0.03$

Forecasting: general process



Forecasting Seasonal Data

- When forecasting seasonal data we need to observe whether the model follows an additive or multiplicative model.

- If the underlying model is additive then multiple regression is the usual approach to modelling the time series. *(not covered in this course)*

- If the underlying model is multiplicative, then seasonal indices can be determined and the deseasonalised series can be forecasted.

Question 2

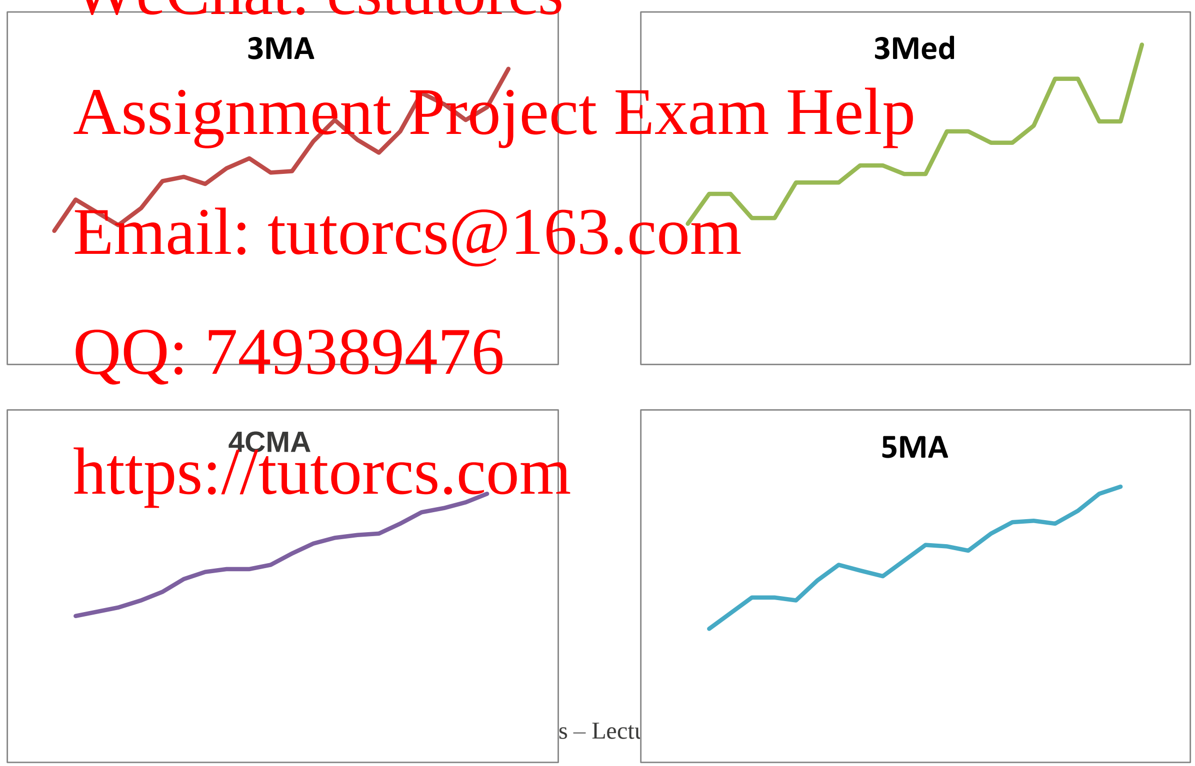
For motivating [illegible] data the best smoother to remove seasonal effect is:

A. 3 MA

B. 3 Med

✓ C. 4 CMA

D. 5 MA



Multiplicative Model: Calculating Seasonal Indices

- The following steps are used to calculate the seasonal index
- Ratios to moving average method.
  - The time series is smoothed. (Use 4 CMA for quarterly data).
  - Divide each observation by its corresponding moving average.
  - Calculate the average ratio for each season.
  - Normalise ratios (to have an average of 1)
  - Method can be adapted for periods of any length.

Example

=Average(average(362, 385, 432,341), average(385, 432,341,382))
= 382.5

Calculate the seasonal indices for the following data:

Quarter	Sales	Centred 4 Period MA	Obs/MA
1	362		
2	385		
3	432	382.50	1.13
4	341	388.00	0.88
1	382	399.25	0.96
2	409	413.25	0.99
3	498	430.38	1.16
4	387	454.75	0.85
1	473	478.25	0.99
2	513	499.63	1.03
3	582	519.38	1.12
4	474	536.88	0.88
1	544	557.88	0.98
2	582	580.63	1.00
3	681	601.50	
4	557	627.63	

The observed value is 113% of what the trend predicts it to be. ie 432/382.5 = 1.13.

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		

Each average is multiplied by 4/3.99 to get calculate the index.

De-seasonalising Data

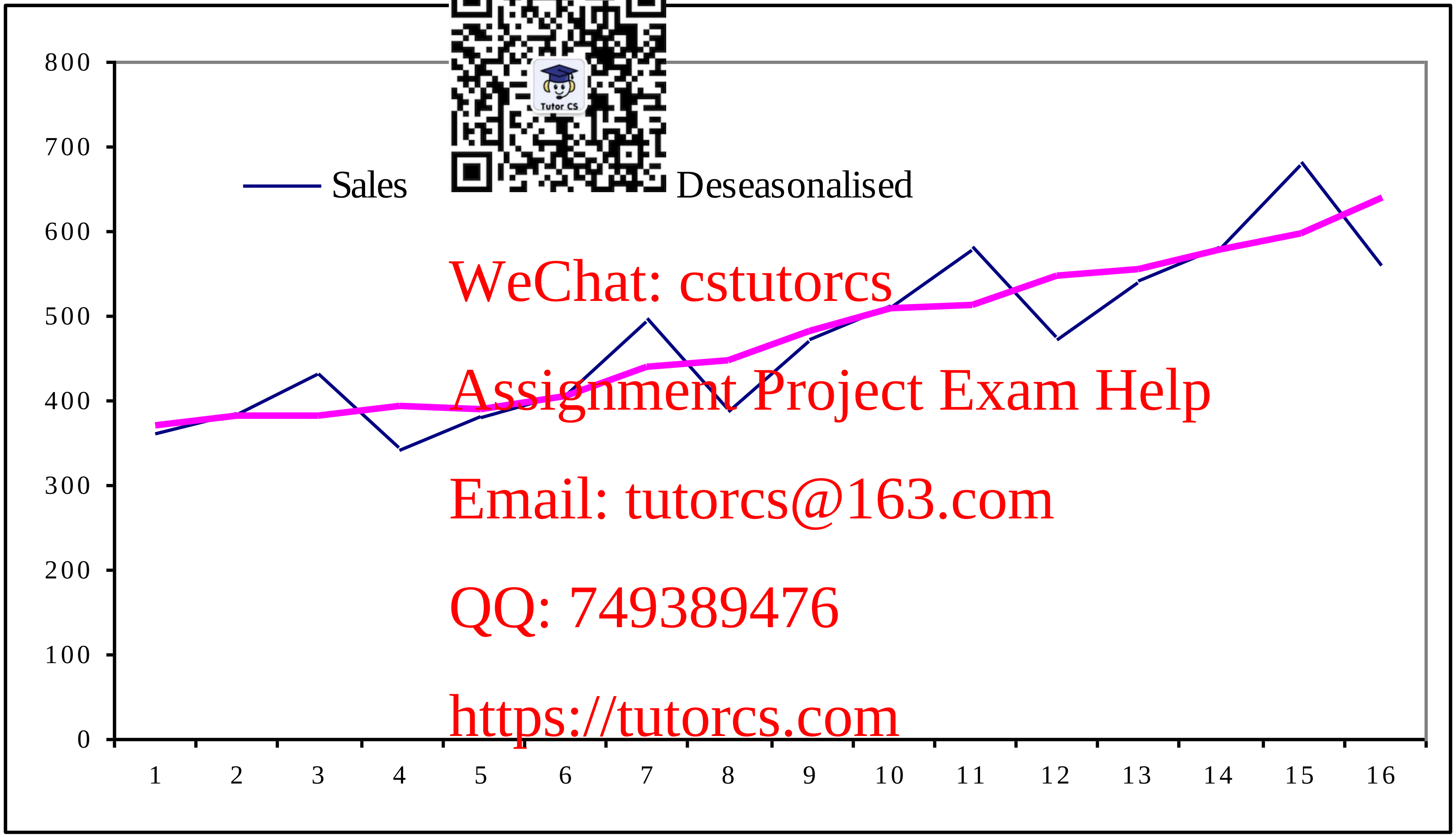
- De-seasonalis[illegible]ime series by dividing each observation b[illegible]asonal factor.

Period	Quarter	Sales	Index	Deseasonalised
1	1	362	0.98	369.39
2	2	385	1.01	381.19
3	3	432	1.14	378.95
4	4	341	0.87	391.95
5	1	382	0.98	389.80
6	2	409	1.01	404.95
7	3	498	1.14	436.84
8	4	387	0.87	444.83
9	1	473	0.98	482.65
10	2	513	1.01	507.92
11	3	582	1.14	510.53
12	4	474	0.87	544.83
13	1	544	0.98	555.10
14	2	582	1.01	576.24
15	3	681	1.14	597.37
16	4	557	0.87	640.23

De-seasonalised value $= \frac{362}{0.98} = 369.39$

Trend and Error

De-Seasonalised Time Series



Seasonal Forecasting

- Having de-seasonalised the data, we can fit a least squares line of best fit, this will create a non-seasonal forecast or trend equation.

- We can use this equation to create a trend for future periods.

- We then re-seasonalise the trend by multiplying by the seasonal indices.

- The next slide shows all steps.

Period	Quarter	Sales	Index	Deseasonalised	Trend	Forecast
1	1	362	0.98	369.39	339.30	
2	2	385	1.01	381.19	357.50	
3	3	432	1.14	378.95	375.70	
4	4	341	[illegible]	391.95	393.90	
5	1	382	[illegible]	389.80	412.10	
6	2	409	[illegible]	404.95	430.30	
7	3	498	[illegible]	436.84	448.50	
8	4	387	0.87	444.83	466.70	
9	1	473	0.98	482.65	484.90	
10	2	513	1.01	507.92	503.10	
11	3	582	1.14	510.53	521.30	
12	4	474	0.87	544.83	539.50	
13	1	544	0.98	555.10	557.70	
14	2	582	1.01	576.24	575.89	
15	3	681	1.14	597.37	594.09	
16	4	557	0.87	640.23	612.29	
17	1		0.98		630.49	617.88
18	2		1.01		648.69	655.18
19	3		1.14		666.89	760.26
20	4		0.87		685.09	596.03
21	1		0.98		703.29	689.23
22	2		1.01		721.49	728.71
23	3		1.14		739.69	843.25
24	4		0.87		757.89	659.36

Slope	18.20
Intercept	321.10

Using excel function to find the gradient and intercept of regression equation or using the least squares formula (in Lecture 8)

Equation of line:
$Y = 18.20x + 321.1$
For Period 17:
$Y = 18.2 \times 17 + 321.1$
$= 630.5$

Re-seasonalise trend:
- Multiply with index
For Period 17 → Qtr 1
Index for Qtr 1 = 0.98
$= 630.49 \times 0.98$
$= 617.88$

Plot of Data, Trend and Forecast

Summary

- You should be a[illegible]
- Calculate the le[illegible]uares regression model for a linear time series.
- De-seasonalise data using the ratio to moving average method.
- Make a de-seasonalised and seasonal forecast using regression.
- Calculate the accuracy of your forecast using MAPE.

Lectures 19/20 Motivating problem

- Given the val[illegible]ilding work (quarterly) fro[illegible]1974 – Dec 2018 .
- Model time series.
- Use historical data to forecast demand for 2019 and 2020.
- Source: ABS.

http://www.abs.gov.au

(File: FIT1006 Lecture 19 and 20.xlsx)

Quarter/Year	Value of Building Work (all sectors) $'Bil
Sep-1974	11.53
Dec-1974	11.06
Mar-1975	9.64
Jun-1975	10.41
Sep-1975	11.15
Dec-1975	10.65
Mar-1976	10.18
Jun-1976	11.37
Sep-1976	11.63
Dec-1976	11.37
Mar-1977	10.14
Jun-1977	11.12
Sep-1977	11.07
Dec-1977	10.57
:	
Mar-2017	27.75
Jun-2017	30.59
Sep-2017	31.52
Dec-2017	31.86
Mar-2018	29.26
Jun-2018	32.84
Sep-2018	32.99
Dec-2018	32.69

Cont.

- If the actual v[illegible] building work [illegible]9 & 2020 is now known (as shown in the table), calculate the accuracy of the forecast.

Quarter/Year	Value of Building Work (All sectors) $'Bil
Mar-2019	29.74
Jun-2019	31.08
Sep-2019	32.17
Dec-2019	30.83
Mar-2020	28.35
Jun-2020	30.14
Sep-2020	30.24
Dec-2020	30.14

Solution

Forecast

Solution (Ref: FIT1006 Lecture 19 and 20.xlsx)

Season/Year	Season	Time Index	Value of Building Work (all sectors) $B	[illegible]	Ratio of [illegible]rved [illegible]oving [illegible]rage	Seasonal Indices	Deseasonalised Series	Fitted Linear Model	Fitted Seasonal Model	Forecast	APE
Sep-1974	Sept	1	11.53	[illegible]		1.036	11.1	7.9	8.2		
Dec-1974	Dec	2	11.06	[illegible]		1.035	10.7	8.0	8.3		
Mar-1975	Mar	3	9.64	[illegible]	0.908	0.919	10.5	8.1	7.5		
Jun-1975	Jun	4	10.41	10.5	0.990	1.010	10.3	8.3	8.3		
Sep-1975	Sept	5	11.15	10.5	1.059	1.036	10.8	8.4	8.7		
Dec-1975	Dec	6	10.65	10.7	0.994	1.035	10.3	8.5	8.8		
:											
Mar-2017	Mar	171	27.75	30.8	0.902	0.919	30.2	27.1	24.9		
Jun-2017	Jun	172	30.59	30.8	0.994	1.010	30.3	27.2	27.4		
Sep-2017	Sept	173	31.52	30.9	1.021	1.036	30.4	27.3	28.3		
Dec-2017	Dec	174	31.86	31.0	1.029	1.035	30.8	27.4	28.3		
Mar-2018	Mar	175	29.26	31.0	0.945	0.919	31.8	27.5	25.3		
Jun-2018	Jun	176	32.84	30.9	1.062	1.010	32.5	27.6	27.9		
Sep-2018	Sept	177	32.99	30.9	1.066	1.036	31.8	27.7	28.7		
Dec-2018	Dec	178	32.69	30.9	1.057	1.035	31.6	27.8	28.8		
Mar-2019	Mar	179	29.74			0.919				27.3	0.08
Jun-2019	Jun	180	31.08			1.010	Slope	0.11		28.1	0.09
Sep-2019	Sept	181	32.17			1.036	Intercept	7.80		28.5	0.12
Dec-2019	Dec	182	30.83			1.035				28.6	0.07
Mar-2020	Mar	183	28.35			0.919				27.8	0.02
Jun-2020	Jun	184	30.14			1.010				28.6	0.05
Sep-2020	Sept	185	30.24			1.036				28.9	0.04
Dec-2020	Dec	186	30.14			1.035				29.0	0.04
										MAPE	0.06

Reading/Questions (Selvanathan)

- Reading: Time Se
  - 7th Ed. Se 3, 17.5, 17.6, 17.8
- Questions: Time Series
  - 7th Ed. Questions 17.12, 17.14, 17.26, 17.34 (linear models only).
  - Tutorial 11 Questions.