Topic: Time Series Analysis **Name:** Solns

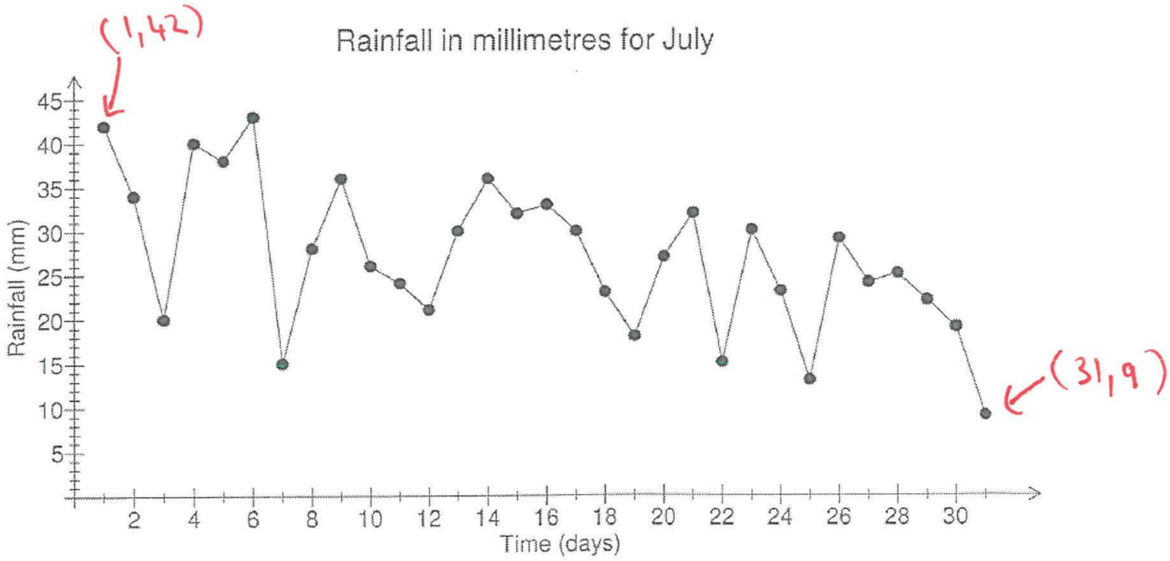
Time series is concerned with the collection of data over a period of time. To enable us to analyse time series data we must take into account (a) the time interval, and (b) any patterns exhibited by the data.

Ex 1.

Construct a time series plot of the daily rainfall in millimetres as recorded for the month of July.

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Rainfall	42	34	20	40	38	43	15	28	36	26	24	21	30	36	32	33
Day	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
Rainfall	30	23	18	27	32	15	30	23	13	29	24	25	22	19	9	

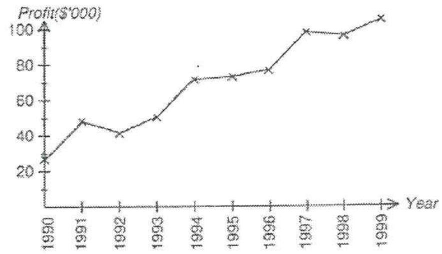
Soln.



Components of time series.

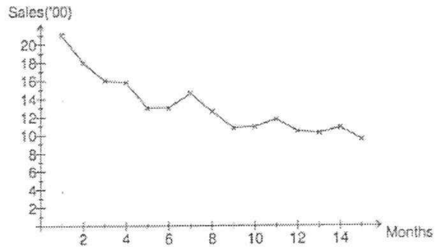
- trend
- seasonal variation or pattern
- cyclical variation of pattern
- irregular or random variation or pattern

➢ TREND



Trend or Secular trend is the long term or overall trend of the data. For example the profits realised by a particular venture may be tending to increase, decrease or even remain constant over an extended time period.

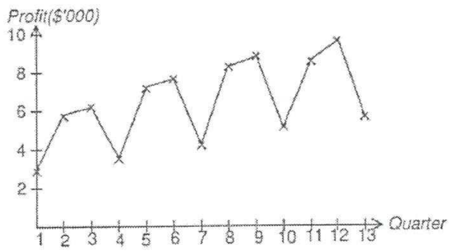
The profit graph on the right indicates that there is an upward trend in profits over the time period under consideration. In such situations we say that the time series plot exhibits a **positive secular trend**.



The sales graph on the right indicates that there is a downward trend in sales over the period of time under consideration. In such situations we say that the time series plot exhibits a **negative secular trend**.

To make a decision on the secular trend of a time series plot make sure that there are a sufficient number of observations in the plot for an overall trend to be identified.

➢ SEASONAL VARIATION OR PATTERN



Seasonal pattern or variation are periodic movements which are related to short term time intervals such as quarters of a year, months of a year, days of the week or hours of a day.

A time series plot with a seasonal pattern is easily identified as it exhibits a pattern of peaks (highs) and troughs (lows) which occur at **regular intervals.**

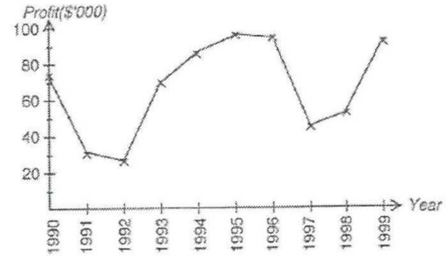
The graph on the right shows a time series plot which has a distinctive seasonal pattern with a regular occurrence of peaks and troughs.

The **period** of a seasonal pattern is the time it takes for the cyclic pattern to repeat itself. The graph shows that there are three 'seasons' which make up one cycle, therefore the period is 3 quarters or 9 months.

Summing up we can say that the time series plot exhibits an overall increasing trend and strong seasonality, which informs us that the profits are predictable as they are calendar-related movements.

➢ CYCLIC VARIATION OF PATTERN

Cyclic patterns are long term fluctuations which are not regular, some may last for a short period and others for a very long period.
A cyclic pattern is evident when peaks and troughs occur at irregular intervals.
The main difference between seasonal and cyclic patterns is that the occurrence of peaks and troughs can be predicted for a seasonal pattern and are unpredictable for a cyclic pattern

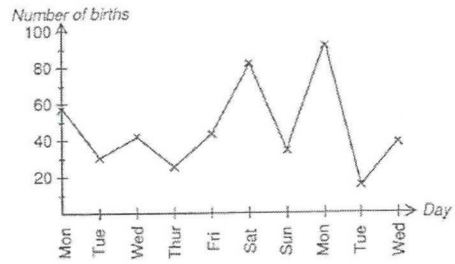


Summing up we can say that the time series plot exhibits an increasing trend and is cyclical.

➢ IRREGULAR OR RANDOM VARIATION OR PATTERN

Irregular or **random** pattern time series data occurs when the data movements are unpredictable short term variations.

Records of the number of natural daily births registered in a city would show this type of pattern.



Ex 2.

The sales figures, in thousands of dollars, for a small business over the last three years is tabled below:

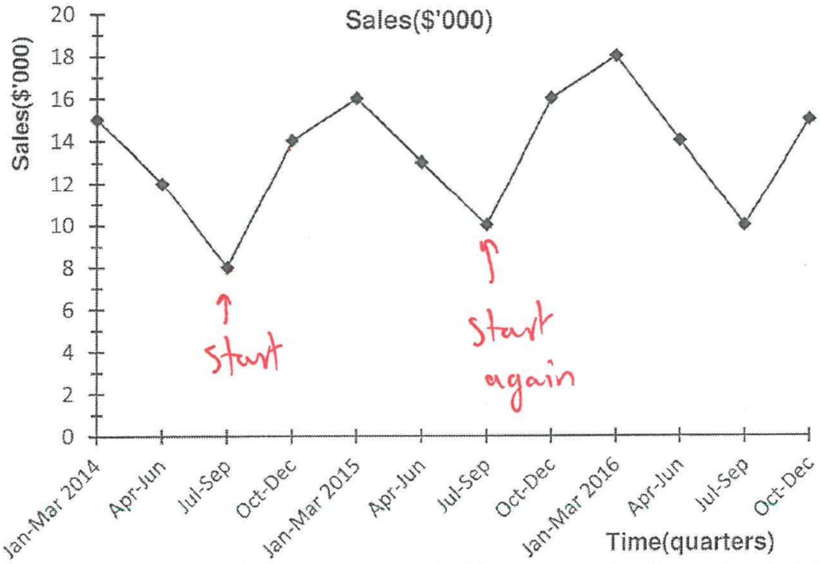
Time Period	Jan-Mar 2014	Apr-Jun 2014	Jul-Sep 2014	Oct-Dec 2014	Jan-Mar 2015	Apr-Jun 2015	Jul-Sep 2015	Oct-Dec 2015
Sales($'000)	15	12	8	14	16	13	10	16

Time Period	Jan-Mar 2016	Apr-Jun 2016	Jul-Sep 2016	Oct-Dec 2016
Sales($'000)	18	14	10	15

(a) Construct a time series plot for the given sales data.
(b) Describe the time series plot in terms of its trend, seasonal, cyclic or random patterns.

Soln.

(a) The required time series plot is shown below.



(b) The time series plot shows a slight positive trend with a seasonal pattern of period 4, peaking during the Jan-Mar quarters

Complete Ex1A

Times Series Analysis

The trend and seasonal components of time series are not random in nature and thus may be studied and measured. We may wish to know what the underlying trend of a time series is so that we can apply this knowledge in the prediction of future values or we may wish to eliminate the trend to enable the study of the seasonal component.

Alternatively we could put to use our understanding of linear regression to determine the line of best fit and reveal the underlying secular trend. Using linear regression in this way is acceptable if the data does not show any seasonal pattern. However, if the time series data under consideration exhibits a seasonal pattern then we must first smooth the given data. Time series data may be smoothed using a moving average.

SMOOTHING TIME SERIES – Moving Averages

The trend of a time series may be determined by the use of a moving average. A moving average involves taking the average value over several data points and by doing so we smooth out the random component and any of the seasonal components. By appropriate selection of the number of terms in the moving average seasonal variation can be removed.

Moving averages tend to reduce the amount of variation present in a data set. This is used to eliminate unwanted fluctuations and this process is called "smoothing of time series."
The result of smoothing of time series is a reasonable indication of the long term trend in the time series.
If the long term trend appears to follow a straight line then the line of best fit may be used to assist in the making of forecasts.

If the time series is seasonal or cyclical we use a ***moving average of size such that it coincides with the number or periods or seasons in the cycle.*** By choosing this moving average the seasonal component will be removed leaving the data trend thus providing suitable data for linear regression.

Summing up we can say that moving averages smooth time series by trying to remove seasonal variation and in doing so reveal any underlying trend.

Ex 3.

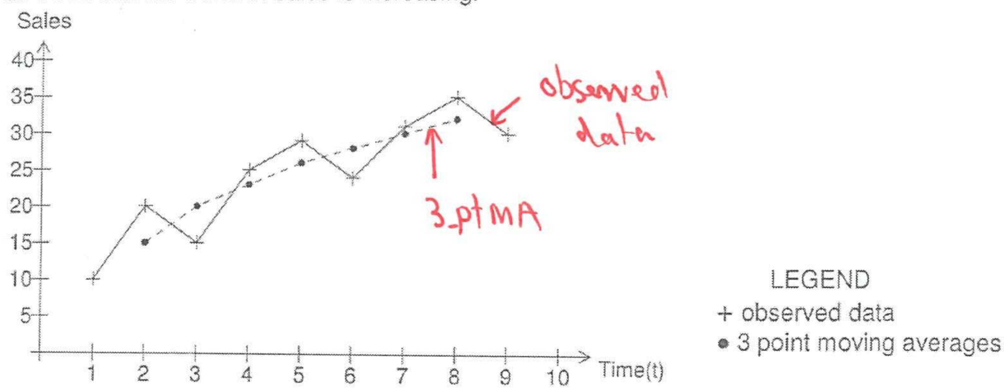
Find the 3 point moving averages of the following sales.

Time period (months)	1	2	3	4	5	6	7	8	9
Sales($)	10	20	15	25	29	24	31	35	30

Soln.

Time Period	Sales	3 point moving averages(ma 3)
1	10	
2	20	$\frac{10+20+15}{3} = 15$
3	15	$\frac{20+15+25}{3} = 20$
4	25	$\frac{15+25+29}{3} = 23$
5	29	$\frac{25+29+24}{3} = 26$
6	24	$\frac{29+24+31}{3} = 28$
7	31	$\frac{24+31+35}{3} = 30$
8	35	$\frac{31+35+30}{3} = 32$
9	30	

The time series and the 3 point moving averages has been plotted below and it can be seen that the 3 point moving average plot is much smoother than the observed data plot and it enables the trend to be seen more clearly. It can be seen that the trend of sales is increasing.



Ex 4.

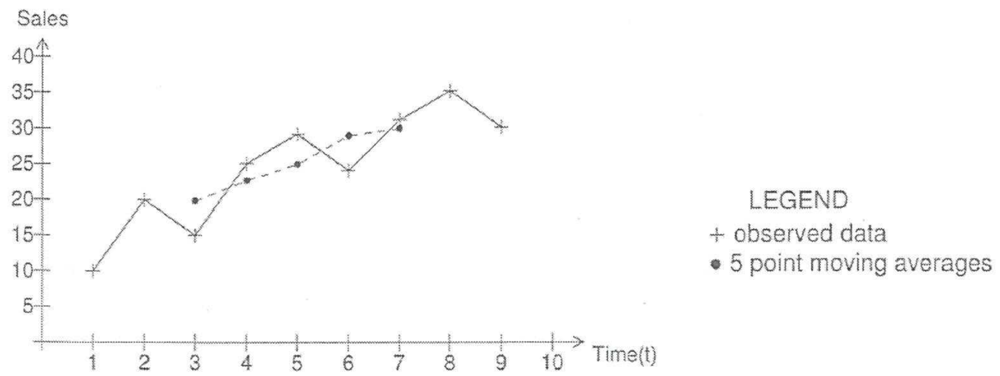
Find the 5 point moving averages of the following sales.

Time period (months)	1	2	3	4	5	6	7	8	9
Sales($)	10	20	15	25	29	24	31	35	30

Soln.

Time Period	Sales	5 point moving averages (ma 5)
1	10	
2	20	
3	15	$\frac{10+20+15+25+29}{5} = 19.8$
4	25	$\frac{20+15+25+29+24}{5} = 22.6$
5	29	$\frac{15+25+29+24+31}{5} = 24.8$
6	24	$\frac{25+29+24+31+35}{5} = 28.8$
7	31	$\frac{29+24+31+35+30}{5} = 29.8$
8	35	
9	30	

The time series and the 5 point moving averages have been plotted below and it can be seen that the 5 point moving average plot is much smoother than the observed data plot and it enables the trend to be seen more clearly. It can be seen that the trend of sales is increasing over the time periods under consideration.



Ex 5.

The following table shows the share price of a mining company over a period of 4 years.

Period	Time period (t)	Price (cents)	3 ma
2012 April	1	35	
August	2	44	39
December	3	38	44
2013 April	4	50	40
August	5	32	45
December	6	p	41
2014 April	7	q	46
August	8	47	42
December	9	41	43
2015 April	10	41	r
August	11	50	43
December	12	38	

Use "Solver"

(a) Find the values p, q and r.

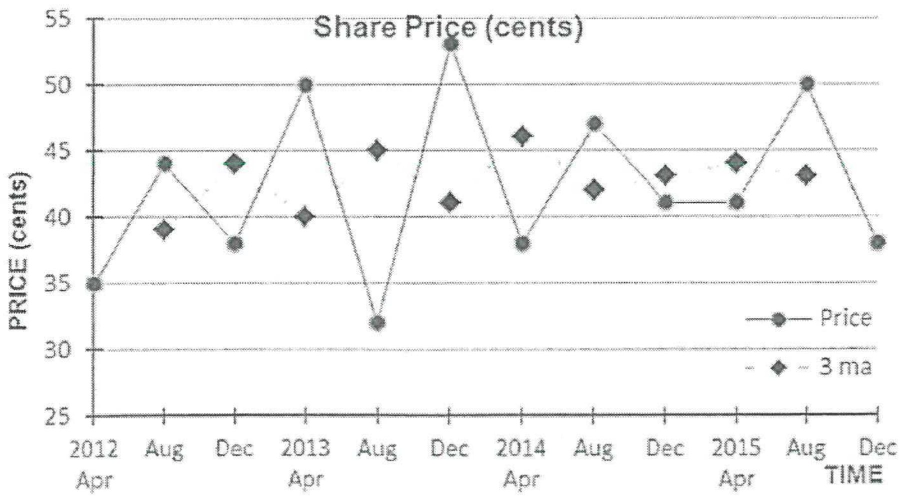
$\frac{50+32+p}{3} = 45 \Rightarrow p = 53$ ¢

$\frac{q+47+41}{3} = 42 \Rightarrow q = 38$ ¢

$\frac{41+41+50}{3} = r \Rightarrow r = 44$ ¢

p= 53 cents, q = 38 cents, r = 44 cents

(b)



(c) The 3 point moving average has not smoothed the time series data to any large extent, There are still troughs and peaks in the 3 point moving average. This informs us that the time series data does not have a period of 3. Examining the pattern of the data it appears that the given data has a random pattern and not a seasonal pattern. This will account for the data not being smoothed using the 3 point moving average.

Complete Ex1B

"Even" Number of data

SMOOTHING TIME SERIES - Centred Moving averages

In the previous section we calculated the moving averages of an odd number of data points and aligned the calculated moving average with the middle data value which corresponded to a given time period.

When calculating the moving average of an even number of data values the centre of these data values does not correspond to one of the given time periods of the given data set. To overcome this problem we take a 2 point moving average of the already smoothed values which will line up with the given time periods. This process is called **centring.**

Ex 6.

Find the **4-point moving averages** (4-pt MA) and the **4-point centred moving averages** (4-pt CMA) of the following sales data.

Time period (months)	1	2	3	4	5	6	7	8	9
Sales($)	10	20	15	25	29	24	31	35	30

Soln.

Method 1

Time Period	Sales	4 point moving averages (ma 4)	4 point centred moving averages (cma 4)
1	10		
2	20		
		$\frac{10+20+15+25}{4} = 17.5$	
3	15		$\frac{17.5+22.25}{2} = 19.875$
		$\frac{20+15+25+29}{4} = 22.25$	
4	25		$\frac{22.25+23.25}{2} = 22.75$
		$\frac{15+25+29+24}{4} = 23.25$	
5	29		$\frac{23.25+27.25}{2} = 25.25$
		$\frac{25+29+24+31}{4} = 27.25$	
6	24		28.5
		$\frac{29+24+31+35}{4} = 29.75$	
7	31		29.875
		$\frac{24+31+35+30}{4} = 30$	
8	35		
9	30		

Centred moving averages

Centred moving averages may be determined using the technique discussed above, however it is probably done more efficiently by the following method.

In order to calculate the 4 point centred moving average we need to use ***five consecutive*** data points. The centred moving averages are determined by taking the sum of, half the first value, the next three values and half the sum of the fifth value and then dividing this sum by four.

For the given time series data, the first 4 point centred moving average corresponding to time period 3 is found by considering the five consecutive sales entries that is 10, 20, 15, 25 and 29.

This 4 point centred moving average is given by $\frac{(\frac{1}{2}\text{ of }10) + 20 + 15 + 25 + (\frac{1}{2}\text{ of }29)}{4} = 19.875$

Using this technique we can find the 4 point centred moving averages for the time series data as shown below.

Method 2

4-pt CMA
⇒ Use 5points

Time Period	Sales	**4 point centred moving averages** (cma 4)
1	$\boxed{10}$	
2	20	
3	15	$\frac{(\frac{1}{2}\text{ of }10) + 20 + 15 + 25 + (\frac{1}{2}\text{ of }29)}{4} = 19.875$
4	25	$\frac{(\frac{1}{2}\text{ of }20) + 15 + 25 + 29 + (\frac{1}{2}\text{ of }24)}{4} = 22.75$
5	$\boxed{29}$	$\frac{(\frac{1}{2}\text{ of }15) + 25 + 29 + 24 + (\frac{1}{2}\text{ of }31)}{4} = 22.25$
6	24	$\frac{(\frac{1}{2}\text{ of }25) + 29 + 24 + 31 + (\frac{1}{2}\text{ of }35)}{4} = 28.5$
7	31	$\frac{(\frac{1}{2}\text{ of }29) + 24 + 31 + 35 + (\frac{1}{2}\text{ of }30)}{4} = 29.875$
8	35	
9	30	

Ex 7.

The table below shows the number of prescriptions filled by a pharmacy over a 4 year period. Calculate and plot the 4 point centred moving averages.

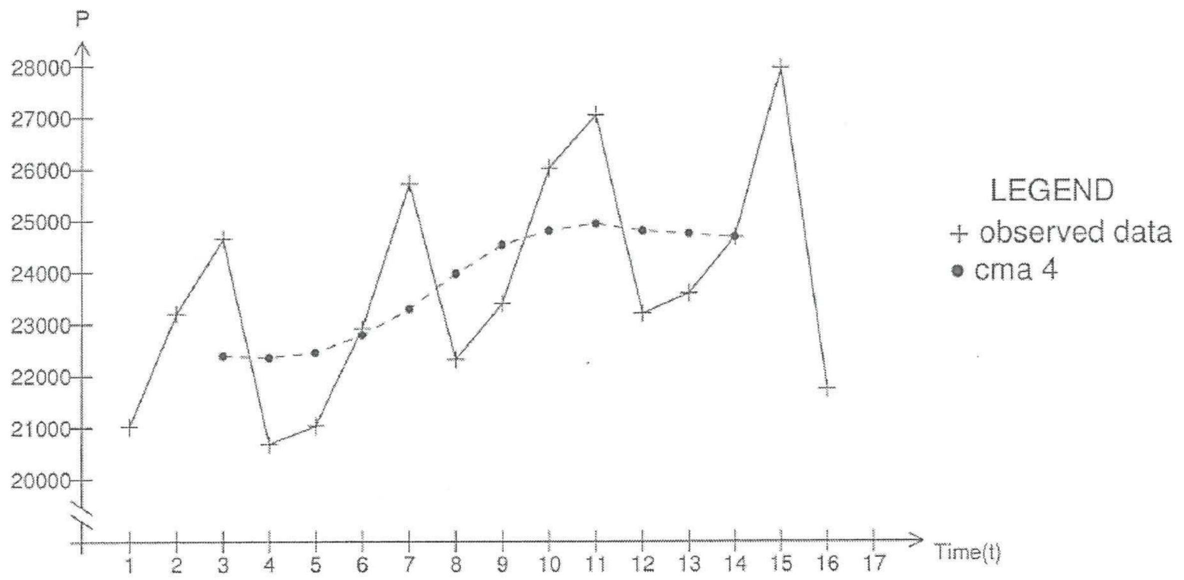
	MARCH	JUNE	SEPTEMBER	DECEMBER
1994	21 031	23 194	24 645	20 674
1995	21 029	22 905	25 711	22 308
1996	23 395	26 014	27 047	23 200
1997	23 589	24 683	27 959	21 730

Soln.

"Using spreadsheet"

Quarter	Time period (t)	No. of Prescriptions(P)	4 point centred moving averages(cma4)
March 1994	1	21 031	
June	2	23 194	
September	3	24 645	22 385.75
December	4	20 674	22 349.375
March 1995	5	21 029	22 446.5
June	6	22 905	22 784
September	7	25 711	23 284
December	8	22 308	23 968.375
March 1996	9	23 395	24 524
June	10	26 014	24 802.5
September	11	27 047	24 938.25
December	12	23 200	24 796.125
March 1997	13	23 589	24 743.75
June	14	24 683	24 629
September	15	27 959	
December	16	21 730	

$$\frac{0.5\times 21031+23194+24645+20674+0.5\times 21029}{4}$$



Complete Ex1C

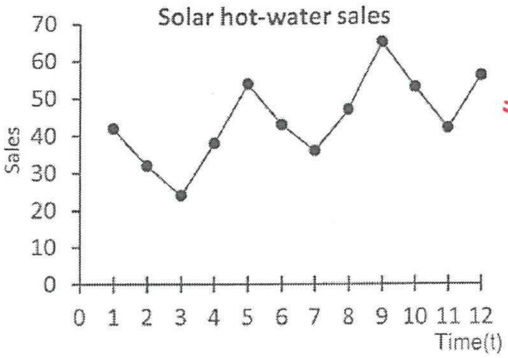
PREDICTIONS AND THE MOVING AVERAGE, SEASONAL INDICES

Having smoothed the observed time series data to reveal the underlying trend we can use the trend line for prediction purposes. Let us consider the following situation.
The following table shows the number of solar hot-water systems sold over a three year period.

	JANUARY	APRIL	JULY	OCTOBER
2014	42	32	24	38
2015	54	43	36	47
2016	65	53	42	56

Graphing the time series will reveal the nature of the time series data.
The time series graph indicates that a seasonal pattern exists and that the period of the seasonal cycle is 4, that is the pattern of points repeats every 4 points as can be seen in the graph of the time series. The pattern of four points or seasons is repeated annually.

In order to see the underlying trend we need to smooth the data and this is best achieved by using the 4 point centred moving average because the pattern of repetition is 4 points.



repeats every 4 points
$\Rightarrow$ 4-pt CMA

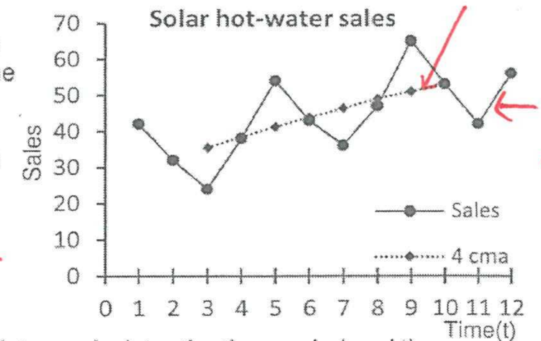
Quarter	Time (t)	Sales (S)	cma 4 (M)
Jan 2014	1	42	
April	2	32	
July	3	24	35.5
October	4	38	38.375
Jan 2015	5	54	41.25
April	6	43	43.875
July	7	36	46.375
October	8	47	49
Jan 2016	9	65	51
April	10	53	52.875
July	11	42	
October	12	56	

$\frac{0.5\times 42+32+24+38+0.5\times 54}{4}$

Jan 2017 13

Line of best fit

Graphing the 4 point centred moving average against the time will give us an indication of the underlying trend of the observed data.
The graph of the 4 point centred moving average (dotted line) indicates that the sales of solar hot-water systems is increasing at a steady rate.
Now as the long term trend indicated by the 4 point centred moving average approximates a line, a line of best fit would be appropriate for prediction purposes



The line of regression (best fit) is determined by entering into a calculator the time period and the corresponding smoothed value, that is the ordered pairs (t, cma 4). For this example the following ordered pairs need to be entered; (3, 35.5), (4, 38.375), (5, 41.25), . . .

The resulting linear regression model is given by $\hat{M} \approx 2.5060t + 28.4926$ where $\hat{M}$ is the predicted 4 point centred moving average sales figure and t is the time period.

G.C. Classpad

List1	List2
3	35.5
4	38.375
⋮	⋮
10	52.875

$\Rightarrow$ Calc $\downarrow$ Regression $\downarrow$ Linear Reg $\downarrow$ ok

$\hat{M} = 2.5060t + 28.4926$

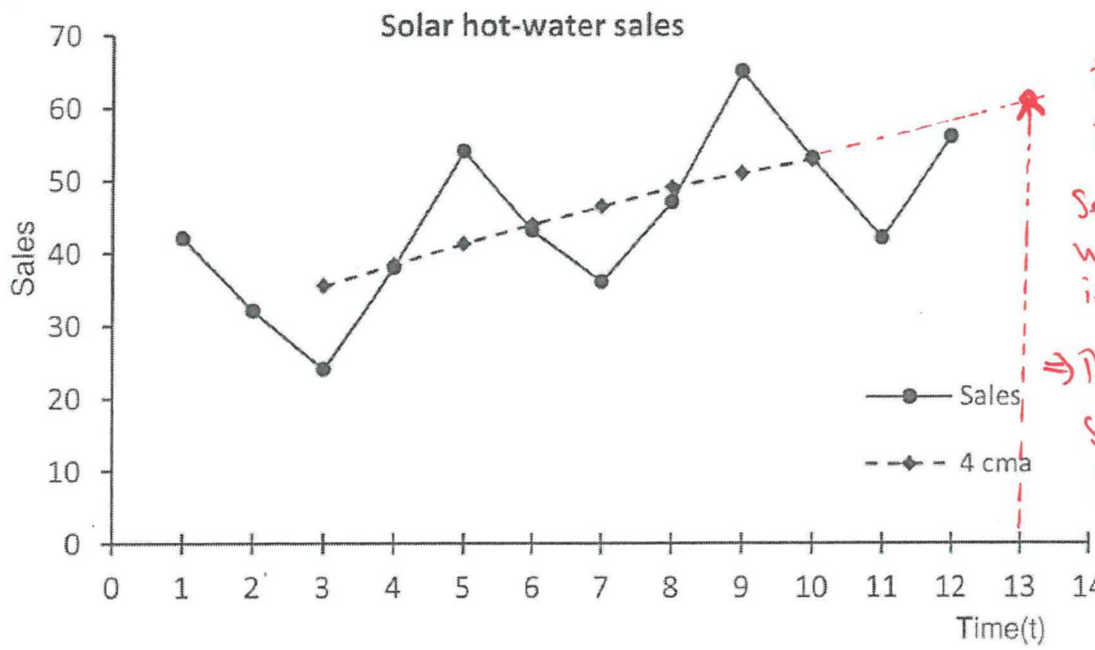
We can now use this linear regression model to predict the expected 4 point centred moving average sales for Jan 20017

Now Jan 2017 is time period 13, hence the expected 4 point centred moving average sales figure is given by

$$\hat{M} \approx 2.5060(13) + 28.4926 \approx 61.07$$

Expected 4 point centred moving average sales figure for the Jan 2017 quarter is 61 solar hot-water systems.

In the example above we used the smoothed values (cma 4) to predict sales for Jan 2017, hence our prediction is a moving average figure as we used a least squares regression line based on the 4 point centred moving averages. This prediction is not a good indicator of the sales figure for Jan 2017 as no allowance has been made for the seasonal nature of the observed data.



$t = 13$
$\hat{M} \approx 61$
Seasonal effect was not taken into account!
⇒ The predicted Sales should be higher for January

Examination of the graph indicates that the predicted sales figure of 61 hot water systems for time period 13 (January 2017 quarter) given by the cma 4 regression line will be too low, as time period 13 is in the same part of the periodic cycle as time periods 1, 5 and 9.

Thus to obtain a more realistic prediction we need to factor back in the seasonal effect for the January quarter.

This seasonal effect that needs to be factored back in for each season is found using the **average percentage method** and it gives us a measure of the effect each season has on the sales.

This measure, found by using the average percentage method, is called the **seasonal effect**, **seasonal component** or more commonly the **seasonal index.**

CALCULATING SEASONAL INDICES

Let us consider the solar hot water systems situation discussed on the previous page.

	JANUARY	APRIL	JULY	OCTOBER
2014	42	32	24	38
2015	54	43	36	47
2016	65	53	42	56

Step 1 Rearrange the given time series data in the following table form to make the calculation of the seasonal indices easier.

Year	Quarter	Time (t)	Sales (S)	Quarterly mean for the year	Percentage of yearly quarterly mean
2014	Jan	1	42		
	April	2	32		
	July	3	24		
	October	4	38		
2015	Jan	5	54		
	April	6	43		
	July	7	36		
	October	8	47		
2016	Jan	9	65		
	April	10	53		
	July	11	42		
	October	12	56		

Step 2 Calculate the quarterly means for 2014, 2015 and 2016 and enter in the table.

To calculate the quarterly mean for 2014 we sum the quarterly sales for 2014 and then divide this sum by 4, the number of quarters in 2014.

Hence the quarterly mean for 2014 = $\frac{42+32+24+38}{4} = \frac{136}{4} = 34$

This calculation of the quarterly mean for 2014 has been entered in the table together with the quarterly means for 2015 and 2016

Year	Quarter	Time (t)	Sales (S)	Quarterly mean for the year	Percentage of yearly quarterly mean
2014	Jan	1	42	$\frac{42+32+24+38}{4} = 34$	
	April	2	32		
	July	3	24		
	October	4	38		
2015	Jan	5	54	$\frac{54+43+36+47}{4} = 45$	
	April	6	43		
	July	7	36		
	October	8	47		
2016	Jan	9	65	$\frac{65+53+42+56}{4} = 54$	
	April	10	53		
	July	11	42		
	October	12	56		

In simple terms we have found the yearly average of the quarterly sales figures for 2014, 2015 and 2016 which are 34, 45 and 54 respectively.

Step 3 Calculate the quarterly proportions for 2014 and express them as a percentage.

Jan 2014 sales figure of 42, as a proportion of the quarterly mean for 2014 = $\frac{42}{34}$ = 1.2353 (4 d.p.)

Expressing this as a percentage = 1.2353 x 100% = 123.53%.

April 2014 sales figure of 32, as a proportion of the quarterly mean for 2014 = $\frac{32}{34}$ = 0.9412 (4 d.p.)

Expressing this as a percentage = 0.9412 x 100% = 94.12%.

July 2014 sales figure of 24, as a proportion of the quarterly mean for 2014 = $\frac{24}{34}$ = 0.7059 (4 d.p.)

Expressing this as a percentage = 0.7059 x 100% = 70.59%.

Oct 2014 sales figure of 38, as a proportion of the quarterly mean for 2014 = $\frac{38}{34}$ = 1.1176 (4 d.p.)

Expressing this as a percentage = 1.1176 x 100% = 111.76%.

Year	Quarter	Time (t)	Sales (S)	Quarterly mean for the year	Percentage of yearly quarterly mean
2014	Jan	1	42		$\frac{42}{34}$ x 100% = 123.53%
	April	2	32	$\frac{42+32+24+38}{4} = 34$	$\frac{32}{34}$ x 100% = 94.12%%
	July	3	24		$\frac{24}{34}$ x 100% = 70.59%
	October	4	38		111.76%

2015	Jan	5	54	$\frac{54+43+36+47}{4} = 45$	
	April	6	43		
	July	7	36		
	October	8	47		

2016	Jan	9	65	$\frac{65+53+42+56}{4} = 54$	
	April	10	53		
	July	11	42		
	October	12	56		

For the 2014:

- the first quarter; (Jan, Feb and March) has a seasonal effect that lifts the sales of these hot water systems 23.53% above the quarterly average.
- the second quarter; (April, May June) has a seasonal effect that drops the sales of these hot water systems 5.88% below the quarterly average (94.12% is 5.88% below 100%).
- the third quarter; (July Aug and Sept) has a seasonal effect that drops the sales of these hot water systems 29.41% below the quarterly average (70.59% is 29.41% below 100%).
- the fourth quarter; (Oct , Nov and Dec) has a seasonal effect that lifts the sales of these hot water systems 11.76% above the quarterly average.

Hence we have found the seasonal indices for 2014.

The seasonal index for the Jan quarter of 2014 is 1.2353 or 123.53%,
The seasonal index for the April quarter of 2014 is 0.9412 or 94.12%
The seasonal index for the July quarter of 2014 is 0.7059 or 70.59%
The seasonal index for the October quarter of 2014 is 1.1176 or 111.76%

Note: The average of the seasonal indices for 2014 should be 100% which means that with four seasons the total of the seasonal indices should be 400% (or 4).
Hence 123.53% + 94.12% + 70.59% + 111.76% = 400%

Our time series data covers 3 years of hot water system sales, that is 2014, 2015 and 2016 and hence we will need to factor in the years 2015 and 2016 to arrive at an overall seasonal index for each of the quarters.
To find the overall seasonal indices for the given time series we need to calculate the seasonal indices for each quarter of 2015 and 2016 and then find the average or mean of each quarter for the three years 2014, 2015 and 2016. This averaging over the three years is termed the average percentage method.

Step 4 Calculate the seasonal indices for the given time series.

The table below shows the completed percentage of quarterly means for 2015 and 2-16.

Year	Quarter	Time (t)	Sales (S)	Quarterly mean for the year	Percentage of yearly quarterly mean
2014	Jan	1	42	$\frac{42+32+24+38}{4} = 34$	$\frac{42}{34} \times 100\% = 123.53\%$
	April	2	32		$\frac{32}{34} \times 100\% = 94.12\%\%$
	July	3	24		$\frac{24}{34} \times 100\% = 70.59\%$
	October	4	38		111.76%
2015	Jan	5	54	$\frac{54+43+36+47}{4} = 45$	$\frac{54}{45} \times 100\% = 120\%$
	April	6	43		$\frac{43}{45} \times 100\% = 95.56\%$
	July	7	36		80%
	October	8	47		104.44%
2016	Jan	9	65	$\frac{65+53+42+56}{4} = 54$	$\frac{65}{54} \times 100\% = 120.37\%$
	April	10	53		98.15%
	July	11	42		77.78%
	October	12	56		103.70%

Having found the percentage of quarterly mean for each of the three years we will need to average these results to determine a measure of the effect each season has on the sales, that is the **seasonal indices** for this time series.

The seasonal indices may be found by arranging the percentage quarterly means or the quarterly proportions in table form.

	January quarter	April quarter	July quarter	October quarter
2014	123.53%	94.12%	70.59%	111.76%
2015	120%	95.56%	80%	104.44%
2016	120.37%	98.15%	77.78%	103.70%
Seasonal Index	$\frac{123.53\%+120\%+120.37\%}{3}$ = 121.3%	$\frac{94.12\%\%+95.56\%+98.15\%}{3}$ = 95.94% (2 d.p.)	76.12% (2 d.p.)	106.63% (2 d.p.)

NOTE:

(i) The seasonal indices may be written as a proportion or as a percentage. For example the seasonal index for the January quarter may be written as 1.213 or 121.3%

(ii) The seasonal indices for every time series will always add up to the number of time periods of the seasonal cycle.
In this case we have quarters of a year, hence the time series has 4 time periods in the year and the seasonal indices will add up to 4 if we report them as a proportion or 400% if we report using percentages.

(iii) Adding the seasonal indices we have 121.3% + 95.94% + 76.12% + 106.63% = 399.99%. This sum is just short of the expected 400% due to rounding.

(iv) If the time periods in a time series are months and as there are 12 of them in a year then the seasonal indices will add up to 12 or 1200%. If the time periods are days of the week in a time series and there are 7 time periods then the seasonal indices will add up to 7 or 700%

(v) To calculate the seasonal indices for a time series the **average percentage method** may use the **mean** or the **median** of the percentages. When required to determine the seasonal indices of a time series we should always assume that the mean is to be used. If the median is to be used then specific instruction will indicate this.

(vi) For the given example if we were instructed to determine the seasonal indices using the **median** then the required seasonal indices would be determined as follows:

	January quarter	April quarter	July quarter	October quarter
2014	123.53%	94.12%	70.59%	111.76%
2015	120%	95.56%	80%	104.44%
2016	120.37%	98.15%	77.78%	103.70%
Seasonal Index	= 120.37% median	**95.56%**	**77.78%**	**104.44%**

We are now in a position to make a better estimate of the sales for the January 2017 quarter.
The expected 4 cma sales figure for the January 2017 quarter was found to be 61.07.
To obtain a more realistic figure we need to factor back in the seasonal effect for the January quarter.
The calculated seasonal index for the January quarter was found to be 121.3% which tells us that generally the quarterly sales of hot water systems are 21.3% more than in an average quarter.
Therefore we need to lift the sales figure of 61.07 by 21.3%.

* Predicted sales for the 2017 January quarter = 61.07 x Seasonal index for Jan quarter
= 61.07 x 1.213
= 74.08 (2 d.p.)

Hence the expected sales for January 2017 quarter is 74 solar hot-water systems.

Predicted value = trend (Regression line) × Seasonal index

Ex 8.

The table shows the monthly sales figures and the seasonal indices for Unitek.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sales($000s)	156	256	243	207	165	106	166	215	203	209	178	165
Seasonal index	0.83	1.35	1.29	1.09	0.87	0.56	0.88		1.07	1.11	0.94	0.87

Soln.

(a) Calculate the seasonal index for August.
Sum of seasonal indices equals the number of seasons, in this example there are 12 seasons.
Therefore 0.83 + 1.35 + 1.29 + 1.09 + 0.87 + 0.56 + 0.88 + x + 1.07 + 1.11 + 0.94 + 0.87 = 12
Now x = 1.14, hence the seasonal index for August is 1.14.

(b) The seasonal index for May is 0.87. Explain what this means.
In May Unitek records 13% fewer sales than in an average month.

Ex 9.

The table shows sales of i-pads from a electrical retail outlet. Calculate the seasonal indices rounded to two decimal places.

Year	March quarter	June quarter	September quarter	December quarter
2011	310	260	290	340
2012	350	300	340	390
2013	400	350	380	430

Soln.

From working earlier examples the seasonal index can be defined as follows:

Seasonal index = $\frac{\text{value for season}}{\text{seasonal average}}$ where the season is the month, the quarter, the day etc and the seasonal average is the yearly monthly average, the yearly quarterly average, the weekly daily average etc.

Step 1 State the formula for the seasonal index in context.

Seasonal index = $\frac{\text{value of the quarter}}{\text{yearly quarterly average}}$

Step 2 Find the yearly quarterly average for 2011.

Yearly quarterly average = $\frac{310+260+290+340}{4} = 300$

Step 3 Find the seasonal index of each season (quarter) for 2011

Seasonal index for March quarter $= \frac{310}{300} = 1.0333$ (4 d.p.)

Seasonal index for June quarter $= \frac{260}{300} = 0.8667$ (4 d.p.)

Seasonal index for September quarter $= \frac{290}{300} = 0.9667$ (4 d.p.)

Seasonal index for December quarter $= \frac{340}{300} = 1.1333$ (4 d.p.)

Step 4 Find the yearly quarterly average for 2012.

Yearly quarterly average = $\frac{350+300+340+390}{4} = 345$

Step 5 Find the seasonal index of each season (quarter) for 2012

Seasonal index for March quarter $= \frac{350}{345} = 1.0145$ (4 d.p.)

Seasonal index for June quarter $= \frac{300}{345} = 0.8696$ (4 d.p.)

Seasonal index for September quarter $= \frac{340}{345} = 0.9855$ (4 d.p.)

Seasonal index for December quarter $= \frac{390}{345} = 1.1304$ (4 d.p.)

Step 6 Find the yearly quarterly average for 2013.

Yearly quarterly average = $\frac{400+350+380+430}{4} = 390$

Step 7 Find the seasonal index of each season (quarter) for 2013

Seasonal index for March quarter $= \frac{400}{390} = 1.0256$ (4 d.p.)

Seasonal index for June quarter $= \frac{350}{390} = 0.8974$ (4 d.p.)

Seasonal index for September quarter $= \frac{380}{390} = 0.9744$ (4 d.p.)

Seasonal index for December quarter $= \frac{430}{390} = 1.1026$ (4 d.p.)

Step 8 Find the seasonal indices of the three years by finding the average seasonal index of each quarter.

Seasonal index for March quarter $= \frac{1.0333+1.0145+1.0256}{3} = 1.02$ (2 d.p.)

Seasonal index for June quarter $= \frac{0.8667+0.8696+0.8974}{3} = 0.88$ (2 d.p.)

Seasonal index for September quarter $= \frac{0.9667+0.9855+0.9744}{3} = 0.98$ (2 d.p.)

Seasonal index for December quarter $= \frac{1.1333+1.1304+1.1026}{3} = 1.12$ (2 d.p.)

Note (i) The sum of the seasonal indices in each case is equal to the number of seasons.
(ii) We found the seasonal indices for each quarter using the **average percentage method** in step 8.

Complete Ex1D

SEASONALLY ADJUSTED DATA

Having found the seasonal indices we now can remove the seasonal component from each observed value in the time series and obtain data giving the best estimate of the long term trend. Removal of the seasonal component results in seasonally adjusted data or deseasonalised data.

** $$\textbf{Seasonally adjusted data (or deseasonalised data)} = \frac{\text{actual value}}{\text{seasonal index}}$$

When using the deseasonalised linear regression model for prediction we must factor back in the seasonality as the prediction is a deseasonalised value. The model to be used is as follows:

** **Prediction = trend(given by deseasonalised linear regression model) x appropriate seasonal index**

PREDICTING OR FORECASTING

To predict in time series using a linear regression model we have a number of choices:

1. **Using the observed data.**
   Fitting a line of best fit by eye or using a least squares regression line will generally result in a very unreliable prediction because of the make up of time series data.
2. **Using the moving average.**
   If this regression model is used the immediate result is the predicted moving average figure. To obtain the predicted figure we must factor back in the seasonal effect. This is achieved by multiplying the predicted moving average by the appropriate seasonal index.
3. **Using the seasonally adjusted data.**
   If this regression model is used the immediate result is the predicted seasonally adjusted figure. To obtain the predicted figure we must factor back in the seasonal effect. This is achieved by multiplying the predicted moving average by the appropriate seasonal index.

* **Ex 10.**

Attendances at a city health club are recorded quarterly as shown in the table below.

	Feb	May	Aug	Nov
2014	3648	1826	1358	5404
2015	2980	1518	1098	3784
2016	2176	1032	846	3086

(a) Calculate the seasonal indices for each quarter.
(b) Deseasonalise the observed data.
(c) Confirm that linear regression is appropriate for the deseasonalised figures.
(d) Find the equation of the least squares regression line for the seasonally adjusted attendances.
(e) Determine the seasonally adjusted attendances for 2017. How reliable are these figures.
(f) Predict the expected attendance for the Feb quarter of the year 2018.
(g) The owners of the health club are aware of the falling attendance and can keep the club open provided that the predicted quarterly attendance is above 500. When will the club be forced to shut down?

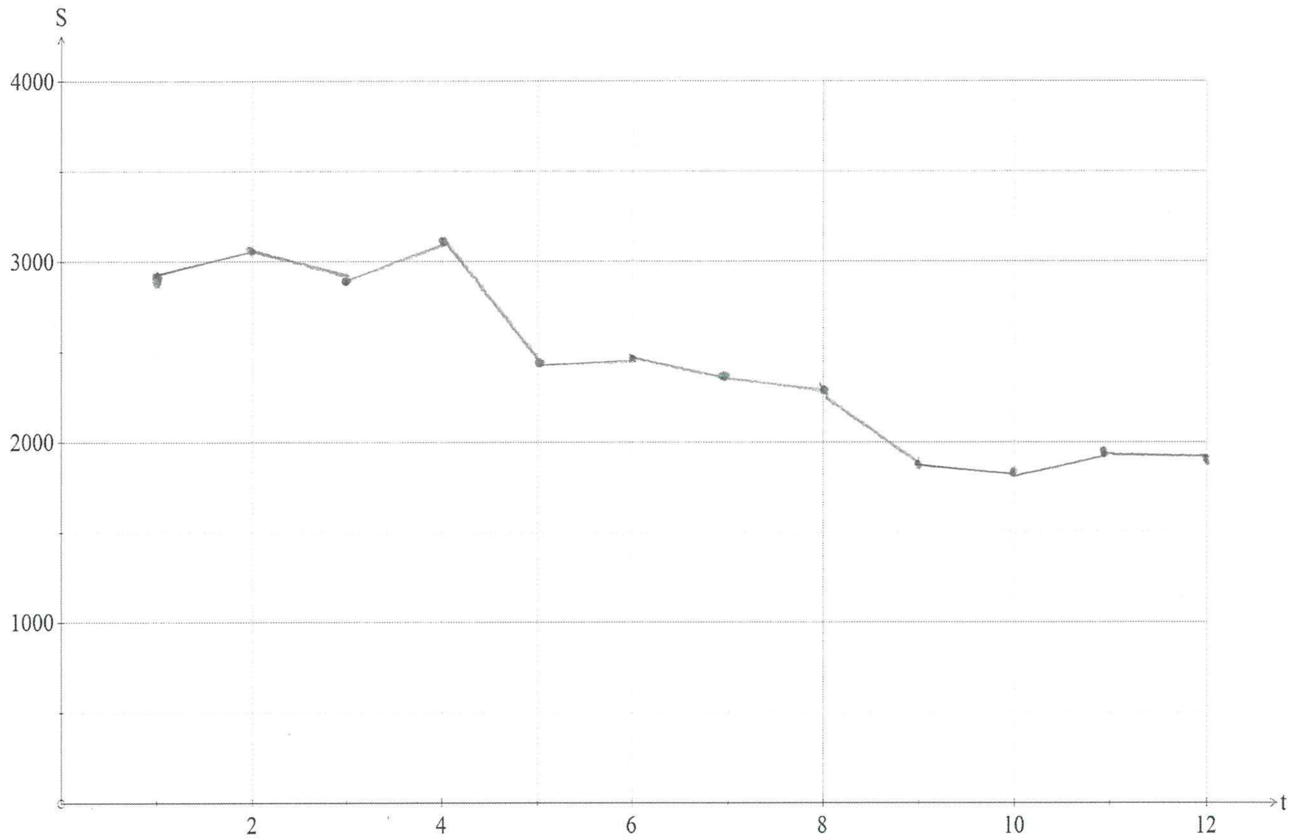
a) "Using spreadsheet"

OR Deseasonalised data $= \dfrac{\text{actual value}}{\text{Seasonal index}}$

(C1) Quarter	(C2) Time Period (t)	(C3) Attendance (A)	(C4) Quarterly mean for the year	(C5) Percentage of year's quarterly mean	(C6) Seasonally Adjusted Attendance (S)
Feb 2014	1	3648	(3648 + 1826 + 1358 + 5404) ÷ 4 = 3059	$\frac{3648}{3059} \times 100 = 119.25\%$	$\frac{3648}{1.2275} = 2971.89$
May 2014	2	1826		59.69%	$\frac{1826}{0.6075} = 3005.76$
Aug 2014	3	1358		44.39%	$\frac{1358}{0.4620} = 2939.39$
Nov 2014	4	5404		176.66%	$\frac{5404}{1.7031} = 3173.04$
Feb 2015	5	2980	2345	$\frac{2980}{2345} \times 100 = 127.08\%$	2427.70
May 2015	6	1518		64.73%	2498.77
Aug 2015	7	1098		46.82%	2376.62
Nov 2015	8	3784		161.37%	2221.83
Feb 2016	9	2176	1785	$\frac{2176}{1785} \times 100 = 121.91\%$	1772.71
May 2016	10	1032		57.82%	1698.77
Aug 2016	11	846		47.40%	1831.17
Nov 2016	12	3086		172.89%	1811.99

	Feb. quarter	May quarter	Aug. quarter	Nov. quarter
2014	119.25%	59.69%	44.39%	176.66%
2015	127.08%	64.73%	46.82%	161.37%
2016	121.91%	57.82%	47.40%	172.89%
Seasonal Index	$\frac{119.25 + 127.08 + 121.91}{3} = 122.75\%$	60.75%	46.20%	170.31%
	(1.2275)	(0.6075)	(0.4620)	(1.7031)

b) Deseasonalised data $= \dfrac{\text{actual value}}{\text{Seasonal index}}$ $\Rightarrow$ Complete column 6 c)



The graph of deseasonalised data, S, plotted against t, the time periods, shows a reasonable linear trend. Hence, linear regression is appropriate in this case.

d) Using graphic calculator,

Time period t → List 1 | List 2 ← Seasonally Adjusted Attendance (S)

List 1	List 2
1	2971.89
2	3005.76
3	2939.39
⋮	⋮
12	1811.99

G.C. → Cal → Regression → Linear Reg → ok

** Note: to find the equation of the least squares regression line

⇒ Use time period (t) and the seasonally adjusted figures (S)

$$\hat{S} = -139.0159t + 3297.7709$$

e) Feb 2017 $\Rightarrow t = 13$ then $\hat{S} = -139.0159(13) + 3297.7709 \approx 1491$

May 2017 $\Rightarrow t = 14$ then $\hat{S} = -139.0159\,(14) + 3297.7709 \approx 1352$

Aug 2017 $\Rightarrow t = 15$ then $\hat{S} = -139.0159\,(15) + 3297.7709 \approx 1213$

Nov 2017 $\Rightarrow t = 16$ then $\hat{S} = -139.0159(16) + 3297.7709 \approx 1074$

Hence, the seasonally adjusted attendances for 2017 are

Feb quarter = 1491 ; May quarter = 1352 ;

Aug quarter = 1213 ; Nov quarter = 1074

These predictions are <u>reliable</u> as they are not far removed from the given data.

f) Feb 2018 $\Rightarrow t = 17$ then $\hat{S} = -139.0159(17) + 3297.7709 \approx 935$

(934.5006)

"SAF"

Seasonally Adjusted Figure.

Note: $\boxed{\text{Predicted Value} = \underset{\text{(Regression line)}}{\text{trend}} \times \text{appropriate seasonal index}}$

Hence, attendance for Feb 2018 = $934.5006 \times 1.2275 \approx 1147$

"SAF" Seasonal index for Feb.

$\Rightarrow$ Feb 2018 attendance is expected to be 1147.

g) May 2018 $\Rightarrow t = 18$ then $\hat{S} = -139.0159(18) + 3297.7709 \approx 795$

(795.4847)

Attendance for May 2018 = $795.4847 \times 0.6075 \approx 483$ (< 500)

Hence, the club will be forced to shut down after the end of Feb 2018 quarter.